Will the Telco survive to an ever changing world ? Technical considerations leading to disruptive scenarios

Roberto Minerva

THESE DE DOCTORAT CONJOINT TELECOM SUDPARIS et L'UNIVERSITE PIERRE ET MARIE CURIE

Spécialité: Informatique et Télécommunications

Ecole doctorale: Informatique, Télécommunications et Electronique de Paris

Présentée par :
Roberto Minerva

Pour obtenir le grade de
DOCTEUR DE TELECOM SUDPARIS

LES OPERATEURS SAURONT-ILS SURVIVRE DANS UN MONDE EN CONSTANTE EVOLUTION? CONSIDERATIONS TECHNIQUES CONDUISANT A DES SCENARIOS DE RUPTURE

Soutenue le 12 Juin 2013

devant le jury composé de :

Noel CRESPI	Professeur, HDR, IMT, TSP	Directeur de thèse
Stéphane FRENOT	Professeur, HDR, INSA Lyon	Rapporteur
Yvon KERMARREC	Professeur, HDR, Telecom Bretagne	Rapporteur
David HAUSHEER	Professeur, TU Darmstadt, Allemagne	Examinateur
Guy PUJOLLE	Professeur, HDR, UPMC	Examinateur
Joerg WIDMER	Chief Researcher, IMDEA	Examinateur
Max MICHEL	Docteur, Networks and IT Coordination Director, Group IT, Orange	Examinateur

Thèse n° 2013TELE0011

Dedication

To my kids: Mauro, Fabrizio and Sveva, evoking their appalling past and wishing them a bright future; and to Stefania; their courageous mother and my beautiful wife.

Acknowledgments

During 2008 I was researching in the "usual way" within the Innovation department of my company. That was the period of a deep transformation from Research to Innovation. At that time, researchers were put within very confined and limited terrains to explore. It was a sort of intellectual "prison" that many Operators were constructing in order to tackle the challenges of the Web world. Instead of opening up minds and research, the approach was to leverage the assets and competitive advantages of the Operators. This approach was aggravating even more the "attitude" of Operators to feel better and more capable of other Competitors. I felt a compelling need to break these glass walls and breath fresh air. On the other side, the success of Web Companies was so stimulating and challenging that there was (is) the need to understand how to build the future of communication service. One thing was puzzling me: how much the needs and rights of users were disregarded. So I decided to put on my "*User Hat*" and figure out if a more user centric web was possible. This was the main personal motivation to enter in a Doctorate.

This thesis is the result of many lucky circumstances, the first is the encounter with Professor **Noël Crespi** that greatly helped me in finding a way towards the enrollment in a Doctorate and constantly supported me with patience and scientific rigor towards the completion of this thesis; to Dr. Roberto Saracco that allowed me to endeavor this journey within my company, Telecom Italia, demonstrating once more his open mind; to my colleagues and friends, Antonio Manzalini and Corrado Moiso that supported me with a lot of technical discussions and cooperative work in writing papers and technical documentation. Other colleagues have variously helped me during these years, from Dr. Giovanni Colombo to Ing. Carlo Eynard that encouraged me and helped in finding good research topics. Thanks also to Prof. **Djamal Zeghlache** for showing me viable paths and objectives. I want also to thank Prof. **Tiziana Margaria** and Prof. **Yvon Kermarrec** for their valuable and enlighten suggestions during the mid-term review.

Other important people were: Prof. **Stephane Frenot** and Prof. **Yvon Kermarrec** that, as Rapporteurs, devoted valuable time in reviewing and suggesting a lot of improvements to this work. I'm pretty sure that their insights helped a lot in increasing the value of this document. My gratitude goes also to all the Examiners that spent time in reading and commenting the document.

I want to thank also all the people in the Institute given a hand in sorting out all kinds of problems and issues and acknowledge the **people of TMSP and UPMC** for easing the administrative matters.

Finally thanks to all my family that patiently has supported me in undertaking this enterprise and has waited for me to finish.

.

Roberto Minerva

12 June 2013

Abstract

The telecom sector is going through a very challenging phase due to radical technological developments (even transformations) -- driven mainly by the constant evolution of the Internet. -- These transformations will have a huge impact on the telecom industry as a whole and, as a consequence, on the future deployment of new networks, platforms, and services. These new technologies and new solutions can be differently shaped according to the specific views, paradigms and models proposed by the various actors in the communication industry. Much of this innovation is promoted and introduced by prominent actors in their efforts to conquer new shares of the Internet market. This approach is particularly fierce in the mobile sector, where the rapid innovation in mobile platforms (e.g., new operating systems such as iOS or Android) has led, on one hand, to the explosion of the mobile internet nurtured by the boom of so called “smartphones”, while simultaneously contributing to the creation of closed environments that are fully controlled (at the client and at the server level) by a single company. As a consequence, the (mobile) Internet is highly fragmented and users are usually trapped into walled gardens.

The situation of mobile internet users is just one case, as the whole Internet is going through several major transformations. For instance, some of the disruptions are determined by the intertwining of technological issues that different disciplines are analyzing and solving. This yields to multiple solutions produced by different industries; solutions that can be adopted in different contexts. One clear-cut example is the case of NBIC (Nano, bio, information and cognitive) technologies.

It obviously will be quite useful to identify when these transformations will accelerate in accordance with the maturation of technological and industry factors. The industry is heavily impacted and hence transformed during these rapid changes, which occur during what this thesis defines as Inflection Points. A first goal is to identify a set of potential and future inflection points made possible by the evolution of technologies. The evolution of the Internet has a particularly strong impact on Telecomm Operators (TelCos). In fact, the telecommunication industry is on the verge of substantial changes due to many factors: e.g., the progressive commoditization of connectivity, the increasing importance of software-based solutions and the flexibility that introduces (compared to the static TelCo systems), and the dominance in the service realm of Web Companies (WebCos).

The thesis is not centered on a socio-political discussion, but elaborates, proposes and compares possible inflection points in evolutionary contexts in order to determine and consolidate possible scenarios based on solutions and approaches that are technologically viable (i.e., that follow a plausible technological trajectory).

The identified scenarios cover a broad range of possibilities:

1) Traditional TelCos, an evolutionary path: the TelCos will essentially maintain their role in connectivity and related services (even if it must be in a consolidated market -- i.e., a market that is not growing, but essentially maintaining the same perimeter);
2) TelCos as Bit Carriers: this scenario is made possible by the progressive decrease in the value of connectivity and related services. Connectivity is a commodity and TelCos will build networks capable of conveying large amount of data with minimal infrastructural costs. The TelCo will merely forward data from edge to edge with the simplest possible network;
3) TelCos as Platform Enablers: Operators will be able to accommodate for the needs of customers by putting together different classes of resources such as connectivity, processing, storage and sensing/actuation. The TelCo will be able to create a networked platform offering resources and control/management functions that customers will be using to create overlay solutions on top of this infrastructure. This scenario can be seen as a logical evolution of the introduction of the XaaS in the TelCo Networks, where X stands for {N= network, P = platform, I = infrastructure}. Services will mainly be provided by third parties;
4) TelCos as Service Providers: this is an optimistic scenario in which Operators are able to position themselves as successful Service Providers. In this scenario, the networked platform is used by the TelCo to directly provide desirable services to their customers. These services are communication services, as well as new classes of services made possible by technological evolution (e.g., the Internet of Things); and
5) Disappearing TelCos: this scenario is the most disruptive with respect to the current situation. It depicts how the evolution of some technologies could potentially lead to a situation where connectivity will be completely commoditized and become a "right" of the citizen. Services will be freely created at the edge of

networks and “networks of networks” will be created between end points representing communities, closed groups or even service providers.

For each scenario a viable platform (from the TelCo’s perspective) is described, together with its potential benefits and the portfolio of services that could be provided.

List of Publications

I. **R. Minerva** “On Some Myths about Network Intelligence” ICIN 2008 Conference, October 2008, Bordeaux France

II. **R. Minerva** “Un duplice approccio all’innovazione dei servizi per superare il divario fra Web 2.0 e le Telecomunicazioni” Notiziario Tecnico Telecom Italia Anno 17, Numero 3 Dicembre 2008 (in Italian) available at http://www.telecomitalia.com/content/dam/telecomitalia/it/archivio/documenti/Innovazione/NotiziarioTecnico/2008/fd_numero03/p15_30.pdf

III. **R. Minerva**, I Paradossi della Rete, come sopravvivere in un Oceano Rosso e Sognare un Oceano Blu, ”, Notiziario Tecnico Telecom Italia (ISSN 2038-193X), Anno 18, Numero 3 – 2009 (in Italian, available at http://www.telecomitalia.com/content/dam/telecomitalia/it/archivio/documenti/Innovazione/NotiziarioTecnico/2009/fd_numero03/04_Paradossi_rete.pdf)

IV. A. Manzalini, **R. Minerva**, C. Moiso, Autonomics in the home, in Proceedings ICIN2009, 26-29 ottobre 2009, Bordeaux (France).

V. A. Manzalini, **R. Minerva**, C. Moiso, If the Web is the platform, then what is the SDP?, in Proceedings ICIN2009, 26-29 ottobre 2009, Bordeaux (France). Best Paper Award 4.

VI. Manzalini, A., **Minerva, R.**, & Moiso, C. (2010). Exploiting P2P Solutions in Telecommunication Service Delivery Platforms. In N. Antonopoulos, G. Exarchakos, M. Li, & A. Liotta (Eds.), Handbook of Research on P2P and Grid Systems for Service-Oriented Computing: Models, Methodologies and Applications (pp. 937-955). Hershey, PA: Information Science Reference. doi:10.4018/978-1-61520-686-5.ch040.

VII. C. Moiso, I. Carreras, B. Fuentes, J.A. Lozano, A. Manzalini, **R. Minerva**, D. Miorandi, Towards a Service Ecology for Pervasive Networked Environments, in Proceedings Future Network and Mobile Summit 2010, 16-18 June 2010, Firenze (Italy), ISBN: 978-1-905824-16-8.

VIII. A. Manzalini, P. H. Deussen, S. Nechifor, M. Mamei, **R. Minerva**, C. Moiso, A. Salden, T. Wauters, F. Zambonelli, Self-optimized Cognitive Network of Networks, in Proceedings Future Network and Mobile Summit 2010, 16-18 June 2010, Firenze (Italy), ISBN: 978-1-905824-16-8.

IX. A. Manzalini, **R. Minerva**, C. Moiso, Towards Resource-Aware Network of Networks, in Proceedings IEEE International Symposium on Wireless Pervasive Computing (ISWPC 2010), 5-7 May 2010, Modena (Italy), (IEEE, 2010), 221-225.

X. A. Manzalini, C. Moiso, **R. Minerva**, X. Hesselbach, J.S. Pareta, J. F. Botero, Generalized Cognitive Networks of Virtual Resources, in Proc. 9th Workshop in G/MPLS networks (WGN9), 5-6 july 2010, Girona (Spain).

XI. Manzalini, Antonio, Minerva, Roberto and Moiso, Corrado. "Autonomic clouds of components for self-managed service ecosystems." Journal of Telecommunications Management 3, no. 2 (2010): 164-180.

XII. A. Manzalini, **R. Minerva**, C. Moiso, The Inner Circle: How to Exploit Autonomic Overlays of Virtual Resources for Creating Service Ecosystems, in Proceedings ICIN2010, 11-14 october 2010, Berlin (Germany), IEEE Xplore.

XIII. A. Manzalini, **R. Minerva**, C. Moiso, P. Baglietto, M. Maresca, M. Stecca, Analysis of Design Patterns for Composite TelCo Services, in Proceedings ICIN2010, 11-13 october 2010, Berlin (Germany) , IEEE Xplore.

XIV. **R. Minerva**, A. Manzalini, C. Moiso, Which way to QoS in future networks: distributed or centralized decision logic?, in 50th FITCE Conference "ICT bridging an ever shifting digital divide", 31 august - 3 September 2011, Palermo (Italy).

XV. A. Manzalini, **R. Minerva**, C. Moiso, Towards 0-Touch Networks, in Proceedings Technical Symposium at ITU Telecom World 2011 (ITU WT11 2011), 25-27 October 2011, Geneva (Swiss).

XVI. **R. Minerva**, N. Crespi, Unleashing the Disruptive Potential of User Controlled Identity Management, in Proceedings Technical Symposium at ITU Telecom World 2011 (ITU WT11 2011), 25-27 October 2011, Geneva (Swiss)

XVII. N. Brgulja, A. Manzalini, **R. Minerva**, C. Moiso, "Autonomic nature-inspired eco-systems", in Springer Transactions on Computational Science XV, Special Issue on "Advances in Autonomic Computing: Formal Engineering Methods for Nature-Inspired Computing Systems", Marina L. Gavrilova, C. J.Kenneth Tan and Cong-Vinh Phan Eds, LNCS 7050 (Springer, January/February 2012), 158-191, 978-3-642-28524-0. 58.

XVIII. R. Minerva, A. Manzalini, C. Moiso, "Towards an expressive, adaptive and resource aware Network Platform", in "Advances in Next Generation Services and Service Architectures", editors A. Prasad, J. Buford, V. Gurbani, Vol. 1 (River Publisher, 2011), 43-63, ISBN 978-87-92329-55-4.

XIX. A. Manzalini, **R. Minerva** "Towards Halos Networks, Ubiquitous Networking and Computing at the Edge" - in Proceedings 2012 16th Int. Conf. Intelligence in Next Generation Networks: Realising the power of the network; Enabling the Internet of everything (ICIN 2012), 8-10 ottobre 2012, Berlino (Germania), ISBN 978-1-4673-1525-8, IEEE Xplore

XX. C. Moiso, **R. Minerva**, Towards a User-Centric Personal Data Ecosystem – The Role of the Bank of Individuals' Data, in Proceedings 2012 16th Int. Conf. Intelligence in Next Generation Networks: Realising the power of the network; Enabling the Internet of everything (ICIN 2012), 8-10 ottobre 2012, Berlino (Germania), 202-209, ISBN 978-1-4673-1525-8, IEEE Xplore. Best Paper Award.

XXI. A. Manzalini, **R. Minerva**, C.Moiso "Bio-Inspired Ecosystems for Telecommunications Pervasive Networks" in Pervasive Computing, Editor: Mary L. Howard, Nova Publisher: https://www.novapublishers.com/catalog/product_info.php?products_id=18241

XXII. A. Manzalini, **R. Minerva**, C. Moiso, Verso le reti 0-touch, in Notiziario Tecnico di Telecom Italia, n. 1/2011 (April, 2011), 38-55 (in Italian)

XXIII. **R. Minerva**, C. Moiso , A. Manzalini, N. Crespi, "Cloud Computing as an Enabler for new Ecosystems of Services" accepted by Springer as a book chapter,

XXIV. **R. Minerva**, A. Manzalini, C. Moiso, N. Crespi, "Virtualization, an Enabler for a new TelCo Platform" accepted by Springer as a book chapter

XXV. G. Balboni, **R. Minerva** " Nuove Reti e nuovi servizi", (in italian Notiziario Tecnico Telecom Italia (ISSN 2038-193X), Numero 3 – 2012 (available at http://www.telecomitalia.com/content/dam/telecomitalia/it/archivio/documenti/Innovazione/NotiziarioTecnico/2012/n3-2012/N2.pdf)

XXVI. A. Manzalini, N. Crespi, **R. Minerva**, U. Emad, Chapter "Software Networks At the Edge" IGI Book title: Evolution of Cognitive Networks and Self-Adaptive Communication Systems; Approved, to be published

XXVII. A. Manzalini, **R. Minerva**, F. Callegati, W. Cerroni, A. Campi "Clouds of Virtual Machines at the Edge", IEEE Com Mag Future Carriers Networks; Under approval.

Other contributions appeared in not referred journals are:

XXVIII. R. Minerva "Crossing the Web 2.0 Chasm: A winning Strategy for Operators", White Paper in SDP 2.0 Service Delivery Platforms in the Web 2.0 Era, Moriana Group, September 2008

XXIX. R. Minerva, How Web 2.0 Will Generate New Revenue Streams for Operators, Appeared in Telecoms 2.0, Oct/Nov 2009. available in http://www.telecom2.com/309x

XXX. R. Minerva, "I Paradossi della Rete, ovvero come navigare in Oceani Rossi ed Oceani Blu", appeared in Rivista AEIT, October 2010 (in Italian)

XXXI. R.Minerva and S. Bell, "Boundary blurring between Telecom and the Internet", appeared in Connect World available at http://www.connect-world.com/articles/recent_article.php?oid=EMEA_2010_02 5.

XXXII. R. Minerva, "An Open Innovation Approach as an Enabler for new Business Models in Telecomms" in European Commission Service Innovation Yearbook 2009-2010 (Spring 2010), available at http://elivinglab.org/ServiceInnovationYearbook_2009-2010.pdf

XXXIII. A. Manzalini, R. Minerva, V. Gonzales "Halos Networks: a competitive way to Internet of-with Things" http://www.comstrat.org/en/Digiworld/Communications-Strategies_41_.html

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

This chapter outlines the context, motivation, contributions and organization of this thesis.

1. 1. Context

In the digital environments services are becoming more and more relevant. Providers differentiate themselves for the capability to provide to customers a relevant set of easy to use functionalities. Services should be considered from several perspectives: in terms of functions and interfaces offered to the final users, in terms of supporting architectures and systems, and in terms of proposed business models and related ecosystems. Currently, customers are seen as final users of the services, but a number of new trends are changing this belief. In fact users are now active actors of the service ecosystem. They not only "consume" services and determine the success of commercial offerings, they also contribute to the development, and testing of them; and, in some cases, they also provide the execution platform for deployment of them.

This is deeply changing the rules and the foundations on top of which services have been developed and offered so far. These changes have technological and business impacts, but also social and economic consequences.

This transformation is occurring in parallel with another makeover: the traditional world of telecommunications has been drastically changed by the advent of the Internet in terms of new technologies, new approaches to services and new ecosystems and business opportunities. The intertwining of changes at the service and the network level have still to be fully understood. For instance the equation that a fall in revenue from communication services can be compensated by major revenues generated by new offering is highly questionable and will be one of the point dismantled by this work. One of the major changes in this traditional world is due to the "interpretation" that the Telecom industry has given to the move from circuits to packets and the definition and adoption of "all-IP" networks. This passage was mainly viewed as an opportunistic move towards cheaper, more programmable and more aligned to the Internet infrastructures. The end to end principle has been considered in a very religious way, with some advocate of the BellHeads (i.e., the traditionalists and old fashioned Operators using the IP technologies to refurbish their network) and the NetHeads (Isenberg 1998). In reality this conflict is more intricate and difficult to grasp; (Minerva and Moiso 2004) discussed the issues emerging from this argument.

The telecom industry is exposing many adjustments: consolidated business models and ecosystems together with architectural solutions and technologies are falling apart. For instance historical companies do not exist anymore (e.g., Nortel), once fierce competitors are now merged and integrated (Alcatel and Lucent, Nokia and Siemens), new comers emerge and are capable of obtain large marker share (Huawei, ZTE). Some paradoxes emerge pinpointing to the contradictions of this industry. These could be seen as symptoms of consolidation, or signals of even bigger transformation to come. This "consolidation vs. structural change" is one of the issues to understand in order to proper position the architectural challenges in the service architectures and in the telecom industry in general. Paraphrasing (Kuhn 1996), this could be seen as a period rich in anomalies and paradoxes that are leading to new technical results (that today are not better of current ones, but that promise for the future better solutions): in other words, this could be a paradigm shift. As stated in (Kuhn 1996):

The success of the paradigm... is at the start largely a promise of success ... Normal science consists in the actualization of that promise... Mopping up operations are what engage most scientists throughout their careers. They constitute what I am here calling normal science... That enterprise seems an attempt to force nature into the preformed and relatively inflexible box that the paradigm supplies. No part of the aim of normal science is to call forth new sorts of phenomena; indeed those that will not fit the box are often not seen at all. Nor do scientists normally aim to invent new theories, and they are often intolerant of those invented by others. And:

The transition from a paradigm in crisis to a new one from which a new tradition of normal science can emerge is far from a cumulative process, one achieved by an articulation or extension of the old paradigm. Rather it is a reconstruction of the field from new *fundamentals, a reconstruction that changes some of the field's most elementary theoretical generalizations as well as many of its paradigm methods and applications. During the transition period there will be a large but never complete overlap between the problems that can be solved by the old and by the new paradigm. But there will also be a decisive difference in the modes of solution. When the transition is complete, the profession will have changed its view of the field, its methods, and its goals*.

Thomas Kuhn

This seems to be the case of Operators. Their goal is to reuse assessed and well consolidate technologies in order to compete with new comers. They are competing in two different technological fields: the network and the service architectures. These architectures follow a release logic: any time a new architecture is standardized and released, it is built on top of previously achieved results; any time new technologies are considered and introduced in the framework, they are modified in such a way that they fit in this continuity line (and they lose effectiveness). This is clear in the evolution of Next Generation Networks and their standards: they are a continuous reworking and improving of the concept of the Intelligent Network.

1. 2. Motivation

In order to determine whether the transformation phase of the Telecomms industry is leading to a paradigm shift, it is important to consider the current technologies used by the Operators and those used by their major competitors and to see what are the respective advantages and capabilities. In addition, the evolution path of strongly related technologies in the Information and Communication Technologies, ICT, can give a detailed insight of capabilities and possibilities that will be available in the next ten years.

In this timeframe, pervasiveness of "smart objects", i.e., objects capable of interaction with the external world in order to create a (physical) environment appropriate for providing personalized services to humans and systems, will be well consolidated.

In this document, this high level definition of service is used: a service is a time-perishable, intangible experience performed for a customer acting in the role of a co-producer (Fitzsimmons and Fitzsimmons 2003). In other terms, a generic service refers to processes, functions, performances, or experiences that one person or organization does for the benefit of another. When the service is provided by means of a software system, the adopted definition is: "(Web) services provide a standard means of interoperating between different software applications, running on a variety of platforms and/or frameworks" (Booth, et al. 2004).

The following definition can be considered: "A Service System is a complex system, i.e., a system in which the parts interact in a non-linear way. As such, a service system is not just the sum of its parts, but through complex interactions, the parts create a system whose behavior is difficult to predict and model. In many cases, a main source of complexity in a service is its people, whether those at the client, those at the provider, or those at other organizations (Spohrer, et al. 2007).

Nowadays the most used control paradigm is the well-known client - server model. Other two relevant models are: the network intelligence, NI, and the peer to peer, P2P, paradigms. They are used to provide a large number of successful services. Other models are currently emerging especially for supporting new classes of services: e.g., the PubSub model. The type of distributed infrastructure and the control paradigms to be used in order to provide services are a major concern and a differentiator between competitor. It is important **to understand what are the technologies and the control paradigms to use in order to implement services in an effective way, i.e., providing a rich set of functionalities, supporting extensibility of them and allowing different level of programmability**. This should be considered also from the point of view of the **expressive power of the solutions** [XVIII]. That means to identify to what extend a control paradigm can fulfill richness, extensibility and simplicity requirements to implement services and functionalities in highly distributed and pervasive systems.

This bring to another major issue: **what is the right architecture for (new) classes of services**? A part of the Telecoms industry has been striving in order to promote horizontal solutions, i.e., general purpose platforms over imposed to the network infrastructure that are to be used in order to provide any kind of service. Many attempts to provide a generic and full horizontal architecture have failed, e.g., the TINA architecture (Berndt and Minerva 1995) had this ambitious goal and it failed to reach it. Today, the "web as a platform" trend is showing that the web technologies and the related infrastructure are a viable way to propose a large number of vertical services. In more details, **this work aim at identify the fittest control paradigms to be used to support the infrastructure of the future**. In addition, **it**

will try to give a glimpse of the possible classes of services that are technical and economically viable for the future.

It is important to stress out that in the generic definition of a service the user has a proactive role and is directly involved in the provision of the service. This aspect is not obvious in the definition of service for the ICT industry, in fact the user is not properly identified as a proactive actor of the provision of services. From this standpoint, another characterizing point of this work is the “User centric approach”. Generally the user centric approach is intended as a method used to identify the requirements of users and improve the way to satisfy them in building up a product or a software solution (Eriksson, Niitamo and Kulkki 2005). In this document, the “user centric approach” goes further and embraces the definition of service and service system: it emphasizes the possibilities and the capabilities that a user could make available directly in order to cooperate in the creation, development, and offering of services and platforms. In other terms, the ICT capabilities are such that users can be able to create their own infrastructure independently form other actors for using services and applications. Why is this important? Many times, the rights and the capabilities of the user are disregarded and neglected in favor of the (economic) interests of specific actors of an ecosystems. As explained in (Clark, et al. 2005), many internet solutions are implemented not in order to achieve the best technical result in the interest of the users, but for exploiting and consolidate techno-economic advantages of a particular actor. Many services could be offered in very different and more effective ways, if the user is really posed at the center of the service design. **One objective of this work is to approach the service architecture definition with a User Centric approach and hence to identify the tussles that impacts the users**.

In order to determine how services and service architectures will evolve, it is important to identify roadmaps for future technologies that can have an impact. One of the major question is: will the foreseen evolution of technologies bring to different way to provide services ? How will these new technologies reshape the network and service frameworks ? In order to answer to these questions, **the document will extensively discuss the impact of new technologies on service and network architectures**. A byproduct of this investigation is to try to **determine (at a very general level) what the values of the network and services will be in the future**. In fact, many disruptive technologies are emerging (e.g., cognitive radio, software defined networks) that can drastically redefine the value of the network asset. **One**

hypothesis that will be investigated throughout the technological analysis is how the combination of general purpose computing and software will impact on the telecoms infrastructures. This possibility is well represented by the concept defined in (Andreessen 2011) that "**software will eat everything**".

Another qualifying point of the thesis is related to identify some impacts of the technological evolution and transformation on the ecosystem of the telecoms industry. This analysis will be conducted **by figuring out some scenarios centered on how the evolution of technologies and the choice of different service control paradigms can impact the Ecosystems and the business models of the industry**. This scenario construction will be considered starting from a major question: What Operators can do in this changing world ? Operators seem to be the dinosaurs that are condemned to disappear because of a paradigm shift in the telecoms and ICT industries.

1. 3. Contribution of the Thesis

The future communications environment will be different from the current one. The integration of different capabilities will determine and shape the evolution towards a new smart environment supporting and enriching the human needs to "communicate". This smart environment will put together communications, processing, storage and sensing/actuating capabilities. This integration will likely shake the traditional foundations of the telecoms industry. **Figure 1** depicts the convergence of these four areas into a powerful and general purpose enabling environment. The extension of this enabling environment with advanced software technologies will lead to the creation of a new sector, the Smart Environment, that exceeds in terms of functionalities each one of the constituents.

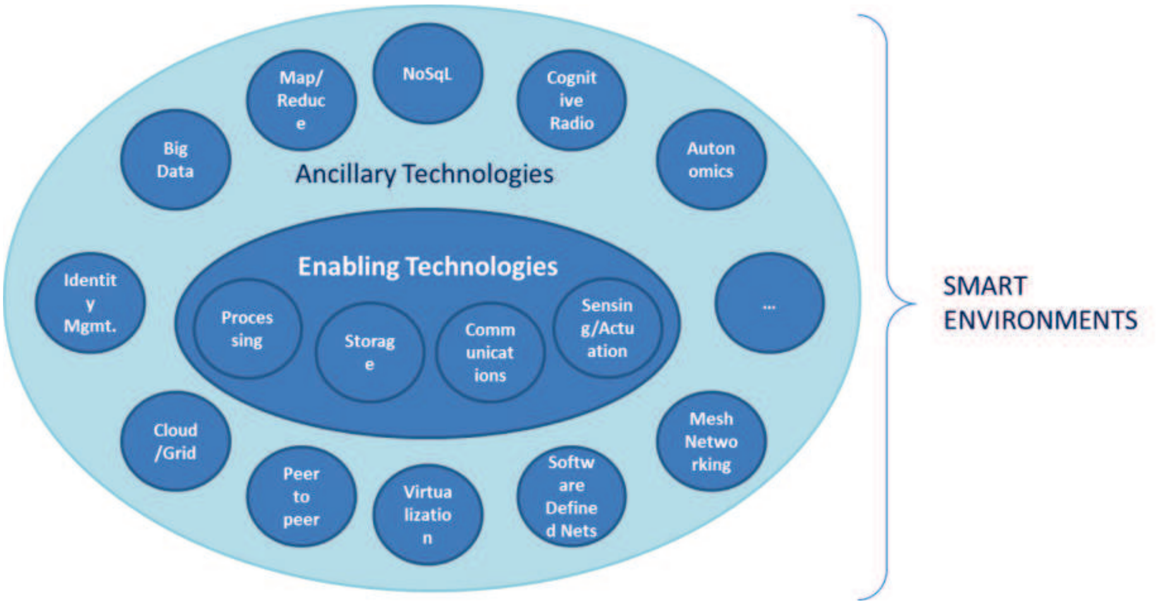


Figure 1: Convergence of Technologies will Lead to Smart Environments

On top of the basic infrastructure (made out of enabling technologies) the ancillary technologies can bring further value, for instance cognitive and artificial intelligence, distributed processing solutions, cloud and grid computing, virtualization, new means to deal with data, and the like can shape a different intelligent environment for the users to be immersed in.

The main contribution of the thesis is related to determining **whether the telecoms industry is undergoing under a structural transformation** or a consolidation towards smart environments. The main hypothesis is that **there is a paradigm shift underway**. This change is manifest and evident in terms of:

- Evolution of essential technologies, i.e., the thesis identifies and shows the general technological evolution and specific disruptive technologies that could alter the status quo;
- modifications of the ecosystems, i.e., the thesis identifies the assets of different actors with major roles in the industry and discuss how they are reacting to this swing. Special attention is paid at how TelCo can respond to this turmoil;
- advent of new classes of services and their appeal, i.e., the work identifies service opportunities. They are analyzed in terms of needed control paradigms and related service architectures for supporting these services.

Another major attempt is in identify **the possibilities of enforcing a factual User Centric approach** that really empowers the users and is aligned with their rights and capabilities. This means to demonstrate when and how the users could have a proactive role in the service provision and in the construction of related platform. The whole contributions of this thesis could be broadly categorized as show in **Figure 2**.

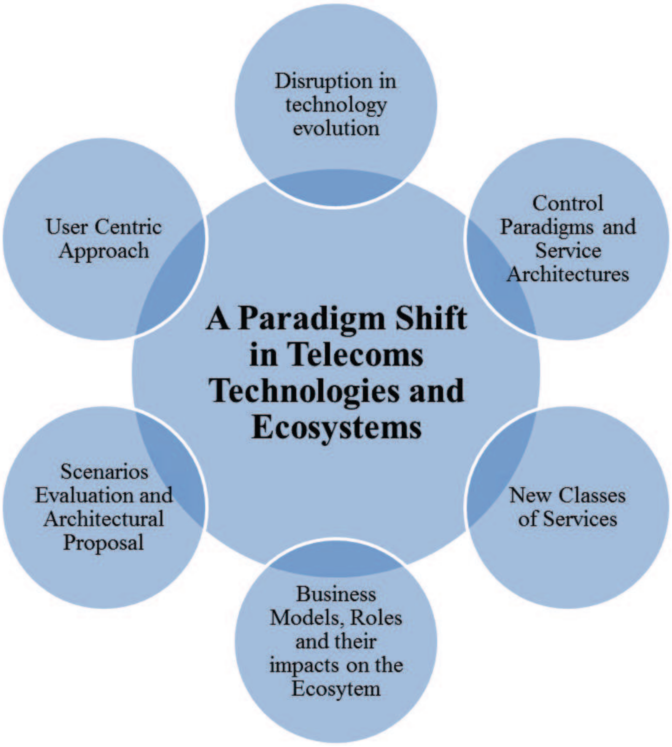


Figure 2: Contributions

In particular these factors are considered and thoroughly analyzed:

- *Technology evolution and roadmaps*: i.e., the current evolution of basic technologies (as processing, storage and communications) as well as the progression of network and service architectures, together with technologies stemming in adjacent fields can lead to playground that is very different from today. For instance, the possibility to use general purpose hardware for mission critical processing could be a disruptive factor for the entire telecoms industry.
- *Control paradigms for the provision of services:* in a highly pervasive and distributed environment, the way objects communicate and interact is of paramount importance. Adopting the wrong interaction paradigm can undermine the entire service offering and hamper a smooth adoption of meaningful services.

- *Networks and service architecture technologies for building viable and effective solutions*: service and network architectures should be capable of supporting the integration of processing, storage, communications and sensing capabilities. Many of the current architectures are not designed to allow this needed integration. There is a need to reshape, refurbish and even to throw away current architecture in order to support the new smart environments. The actors that have better architectures will have an edge. This analysis will especially focus on the possibilities of TelCos to offer appropriated and viable platforms to address the future communications environment.
- *New classes of services*: the combination of new technologies with the proper control paradigms will lead to the creation of new classes of services that will be substantially different from the current ones. The most important aspect of this convergence is that services will arise also from other sciences (e.g., biology, neuroscience, physics) and their combination with the smart environments will drastically challenge the way services are currently thought of. It will not be the Internet of Things, but every single real world object will be deeply embedded in the Internet.
- *Mutation of the ecosystem, new business models and roles*: the smart environments and the technological evolution will have a tremendous impact on the current way services and applications are provided. New relationships between actors will emerge and the role of users could change as well. New actors could play relevant roles (e.g. local communities). One important proposition here is to identify and describe roles and opportunities that enable the User to play a proactive role in the new environment. In general this mutation of the ecosystem will have a huge impact on the traditional business model of the TelCos. However, new opportunities will emerge such as the servitization capabilities (i.e., the possibility to transform products/goods into services).

In order to systematize the approach, the thesis (essentially in the third part) will focus on this aspects:

- *Construction and evaluation of future scenarios*: this scenarios will take into account the possibilities of TelCos to play a relevant role in the "smart environment" ecosystem. Different scenarios will be constructed and analyzed ranging from the most conservative one (nothing changes) up to the most disruptive one (e.g., the disappearing of TelCos).

- *Architecture proposition*: for each viable role a supporting architecture will be sketched out in order to emphasize the assets and the opportunity that a TelCo could put into the field.

Thesis contributions are summarized in **Table 1**.

Table 1: Summary of contributions

	General Contexts: Technologies and Impacts on Ecosystem, and Role of the Operator
Control Paradigms	Papers I, II, III, XI, XII, XVII have discussed the control paradigms at the technological, ecosystems and role of the Operator levels
Service Architectures	Papers I, II, III, IV, V, VI VII, XXII, have considered the Service Architectures at the technological, ecosystem and Role of the Operator levels.
Network Architectures	Papers VIII, IX, X, XIII, XIV, XVI, XVIII, XX, XXIII, and XXV have considered the different perspective of new networks.
New Classes of Services	Papers II, III, IV, XIII, XXII, XXIII, XXIV discuss the new service classes considering different perspective of the technologies, the ecosystems and the role of the Operator.
User Centric	XV and XIX have considered the technological implications, the impact on the ecosystem and the Role of the Operator within a User Centric approach.

One underlying consideration that is underpinning all the reasoning of this work is that software is acquiring a principal role in almost any sector of the different industries. As clearly put forward by Mark Andreessen in (Andreessen 2011), the question is “Why Software will eat everything”. If one sentence should be used to describe this thesis, then “software will eat everything” is the one to use.

1. 4. Organization of the Thesis

Based on research contributions, this dissertation is organized into three parts.

The first part covers the technology evolution in terms of essential technologies, new network architectures, new classes of services, Contextually this part describes also the different control paradigms and how they are leveraged in order to support the solutions. In this part, a broad overview of literature (0) is presented and then the evolution of enabling technologies (Chapter 3.), new network architectures (Chapter 4.) and emerging Services and their control paradigms (Chapter 5.) are discussed.

The second part of dissertation primarily focuses on the tussles of the Telecoms worlds and business models and relationships within the ecosystem. This part consists of two chapters as given below.

Chapter 6. presents an overview of how the network vale is perceived, some paradoxes that characterize the current ICT evolution and the general approaches that TelCos are undertaking in terms of Lean and/or Smart Operator.

Chapter 7. analyzes how the Operators pursue their traditional business model by means of mainstream architectures and applying the traditional paradigm of the network intelligence.

In the third part of dissertation, a summary of architectural considerations and a viable proposition for the role of platform provider is presented. This part includes these chapters as given below.

Chapter 8. presents how the scenarios have been defined and put together. It also describes the five scenarios from the TelCo perspective.

Chapter 9. analyses and consolidates the research contributions and compares them with the initial challenges presented in Section 1. 2. It provides a high level blueprint of a platform that enables new services. It is designed in such a way to encompass the contribution and the resources of the customers. In addition this chapter describes potential future works and outlines some conclusions and perspectives on the use of this work.

PART 1. TECHNOLOGY EVOLUTION

Introduction

The evolution of ICT technologies combined with solutions stemming from other disciplines is reaching such a power that it is possible to foresee great changes for the telecoms industry. These changes could be represented by Inflections Points, i.e., specific points in time in which the evolution of the technology makes viable solutions that up to some time ago were deemed impracticable. The notions of Service and Network themselves can be drastically revisited due to new possibilities offered by ICT technologies. Networks are closer in meaning to smart environments capable of adapting to the requests of user than a set of systems to be used to carry signal between two end points. Services are offering unlimited capabilities and ICT technologies are providing a framework for many new classes of services and applications.

This part consists of four chapters:

Chapter 2, in which the State of the Art related to Control Paradigms, Service and Networks architectures and the classes of services is depicted.

Chapter 3. Chapter 4. and Chapter 5. present the possible evolution of enabling technologies, Service architectures, Network architectures and Services integrating original contributions of the thesis respect to the most innovative trends of the academy and industry in a time span of about ten years.

Chapter 2. Background: Services, Control Paradigms, and Network Architectures

Highlights

- How the notion "Services" is treated in ICT (and especially in telecoms)
- Interaction paradigms and their expressive power
- Service and Network architectures from the software perspective

2. 1. Introduction

In this chapter, the background for analysing the service realm from an ICT and specifically telecoms perspective is given. The terms of discourse and the definitions are broad and usually they have been coined in a specific environment (e.g., IT, Web, or Telecoms environments). These definitions can be adopted in larger contexts, however, when ambiguity rises from the extension of the meaning to a broader context, the document will clarify the issues..

Intuitively, services in this context are a set of discrete and valuable functions provided by a system. Such a system, in the context of communications, will be distributed, i.e. parts and functions of it are not placed in the same location. In addition a multitude of users will be requesting for the service. The distributed system parts as well as the system and the users need to interact and hence some interaction models and mechanisms should be defined in order to well support the provision of the service to users. Since services are provided by a software infrastructure, a general architecture of these systems (and their subsystems) needs to be identified in order to facilitate the design and the life cycles of services. ICT and Communications services presume also the ability to exchange information and hence networks are to be used.

Services can be interpreted in different ways depending on the perspective, the Telecommunication Operators, TelCos, see services as an extension of basic communication means, Web Companies, WebCos, see services as functionalities that they can provide/sell by means of web capabilities, other IT actors consider services as software packages that can be provided and sold to users that will use them in order to accomplish a set of tasks or functionalities (e.g., a spreadsheet). It is important to have a common ground for expressing and discussing the meaning of services.

Services are accessed by the users in very different way and the internal working of services is characterized also by different control means (i.e., ways in which functionalities are coordinated within the "service system" and they made available to the users). Also from this perspective the views on services are profoundly different: TelCos see services as sets of functionalities that deeply rely on the network. Actually the network is driving the service provision by requesting the provision of functionalities that the control layer of the network is not capable to deliver. WebCos considers services as a set of functions requested in a client – server functions (the WebCos acting as Server) that can be accomplished with a complex and highly distributed back end system. Typically a WebCo is seen as a "server". Other views on services, for instance in peer to peer or overlay models, consider different interaction and control paradigms that try to integrate valuable functions provided by other peers or entities and the functions needed to support a logical view of the system that is delivery the functions themselves. From this standpoint, also a system devoted to deliver services is conceptually and practically very different depending on the stakeholder that is offering the service functionalities and is building and operating a system supporting the delivery of functions. These differences results in distinct service architectures that do emphasize the approach pursued by the specific Actor and do leverage the specific assets of the Stakeholder (e.g., the network, the data center, the logical overlaying of functions).

The relationship between the network and the services is in fact an important aspect to consider and define in order to properly understand the intricacies of ICT and Communications services. For TelCos it is the major asset to leverage in providing services (actually services can be seen as "extensions" or enrichments of the major capability offered by the network: the network is providing the basic service, i.e., connectivity, and services rely on it and functionally extend its properties. For WebCos, the network is just a pipe that is providing a single functionality: the best effort connectivity. All the relevant functionalities are to be provided to the edge and in particular by powerful servers. In peer to peer systems, the network is logically abstracted and its relevant properties are offered as services to be integrated with other capabilities and functions offered by the peers.

In order to shed a light on these different approaches and their consequences a set of examples of some services already implemented adhering to different approaches is presented. [XVIII] and (Minerva 2008) address many of the topics of this chapter and could be considered as a sort of guide to these issues. It also considers the interaction and control paradigms for their expressive power, i.e., the capability to represent and implement the services.

2. 2. Services

As a starting point for a Service Definition, the W3C definition is (Booth, et al. 2004) is considered because it generally applicable to represent the concept of a service: "*a service is an abstract resource that represents a capability of performing tasks that represents a coherent functionality from the point of view of provider entities and requester entities. To be used, a service must be realized by a concrete provider agent*[1]". A service is then a complex concept that involves many actors/agents that offer resources supporting valuable functions. These functions can cooperate in order to achieve the goals of a service . According the W3C, a Service can be represented as in **Table 2**:

Table 2: W3C Service Definition

Property of a Web Service	Description
• **a service is a resource**	• A resource is an entity that has a name, may have reasonable representations and which can be owned. The ownership of a resource is critically connected with the right to set policy on it
• **a service performs one or more tasks**	• A task is an action or combination of actions that is associated with a desired goal state. Performing the task involves executing the actions, and is intended to achieve a particular goal state.
• **a service has a service description**	• a description is a set of documents that describe the interface to and semantics of a service

[1] An agent is a program acting on behalf of person or organization.

• **a service has a service interface**	• an interface is the abstract boundary that a service exposes. It defines the types of messages and the message exchange patterns that are involved in interacting with the service, together with any state or condition implied by those messages.
• **a service has a service semantics**	• it is the behavior expected when interacting with the service. The semantics expresses a contract (not necessarily a legal contract) between the provider entity and the requester entity. It expresses the intended effect on the real-world caused by invoking the service. A service semantics may be formally described in a machine readable form, identified but not formally defined, or informally defined via an "out of band" agreement between the provider entity and the requester entity.
• **a service has an identifier**	• an identifier is a unique unambiguous name for a resource.
• **a service has one or more service roles in relation to the service's owner**	• a service role is an abstract set of tasks which is identified to be relevant by a person or organization offering a service. Service roles are also associated with particular aspects of messages exchanged with a service.
• **a service may have one or more policies applied to it**	• a policy is a constraint on the behavior of agents or people or organizations
• **a service is owned by a person or organization**	• an agent has the right and authority to control, utilize and dispose of the service
• **a service is provided by a person or organization**	
• **a service is realized by a provider agent**	• A provider agent is an agent that is capable of and empowered to perform the actions associated with a service on behalf of its owner — the provider entity.

Figure 3 depicts the complexity of the service definition.

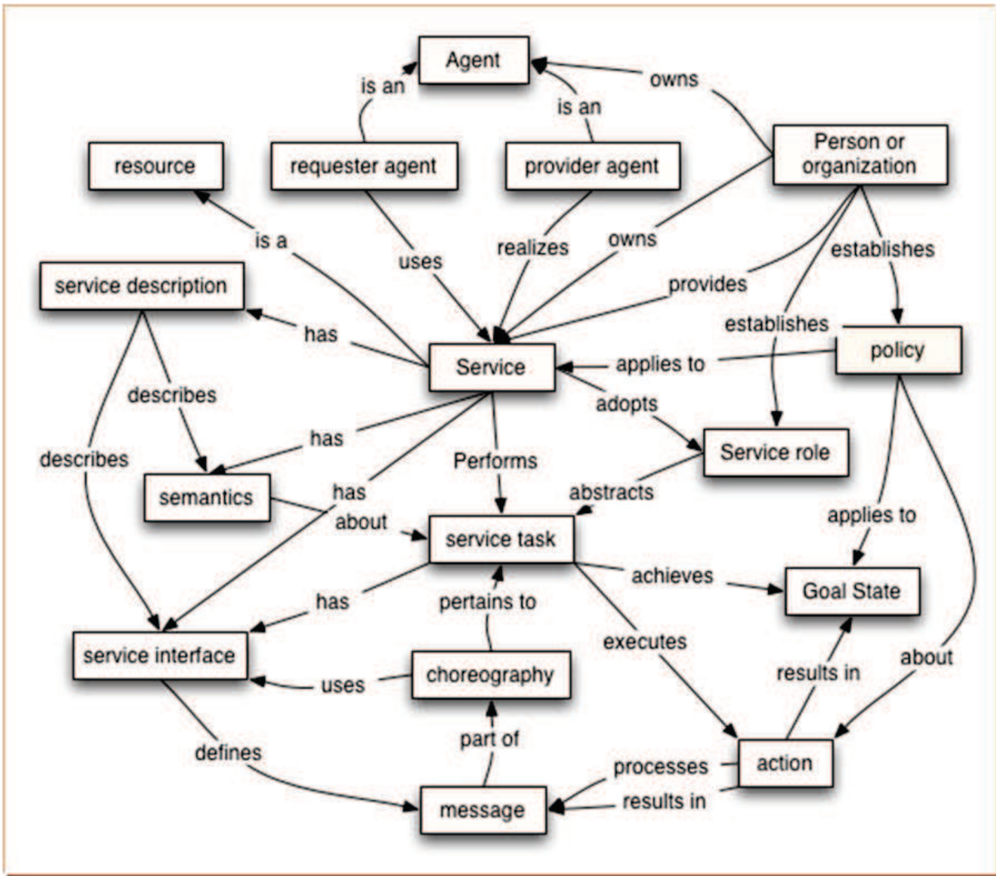


Figure 3: The W3C Service Definition

The W3C definition of a service is a generic one, but it contains some elements that pertain to the specific context and interaction model that are underpinning the web services. For instance, even if it is not cited, the underlying interaction model is client - server in which a requester agent (a client) will send messages (requests) to a provider agent (i.e., a server). The service is a resource that is strictly owned by a person or an organization, this approach is straightforward for the web context (based on a client - server interaction paradigm), but not necessarily fits for other interaction paradigms (e.g., the peer to peer, P2P). The actions performed by a service aims at achieving a goal state (a predetermined situation favorable to the requester and/or the provider), this is not always the case for other systems based on other interaction paradigms.

Also the Telecom related definitions of service are in a way the brain child of a specific vision of the service world, for instance: "*telecommunications service: 1. Any service provided by a telecommunication provider. 2. A specified set of user-information transfer capabilities provided to a group of users by a telecommunications system. Note: The telecommunications service user is responsible for the information content of the message.*

The telecommunications service provider has the responsibility for the acceptance, transmission, and delivery of the message" (NTIA 1996). In addition the term "service" in a telecoms world very frequently is interpreted as a "network" or at least as the transport service that a network providers, for instance: "*3G: Third-generation mobile network or service. Generic name for mobile network/service based on the IMT-2000 family of global standards*" (ITU 2011). Often also the literature follows this approach such as (Znaty and Hubaux 1997) that stated:

- "*Telecommunication services is a common name for all services offered by, or over, a telecommunication network*";
- "*Value added services are services that require storage/transformation of data in the network and are often marketed as stand-alone products. Examples of such services are freephone, premium rate, virtual private network and telebanking. Many value added services can be offered by special service providers connected to the network*";
- "*Teleservices include capabilities for communication between applications. Teleservices are supplied by communication software in the terminals. Telephony, audio, fax, videotex, videotelephony are examples of teleservices*".

All these definitions bring in the concept of a network supporting a major service on top of which "added value" functions and services can be instantiated as represented in **Figure 4**.

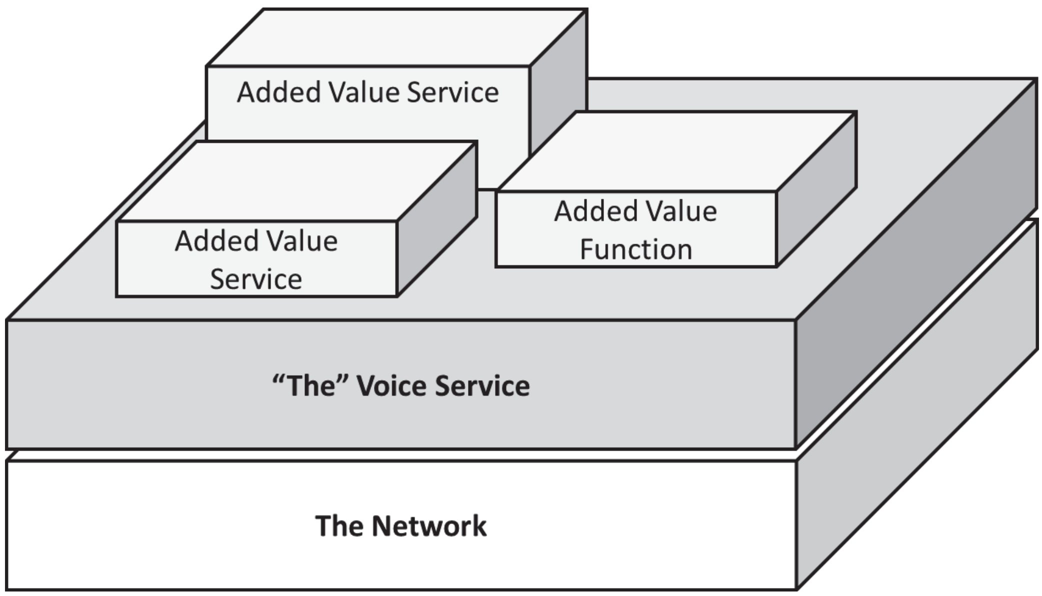


Figure 4: The Network, The Service and Added Value Functions/Services

These definitions reflect the vision of services: the network is the main asset and it provides a main service, THE voice service, on top of which several functions or services can be provided. They add value to the unique service at hand.

A more detailed service definition that is also considering the integration of different paradigms is given in (Bertin and Crespi 2013).

Also in the Peer to Peer sector, the service definition strongly depends on the "view" of the system organization for supporting functionalities. For instance (Gradecki 2002) defines a service as "*predefined functionality that can be utilized by peers*" where a peer is "*an application, executing on a computer, that has the ability to communicate with similar peer applications*".

These definitions introduce another issue: the differences between services and applications. Actually services and application are sometimes used as synonymous, so an application can be defined as a set of functions that have meaning and usefulness within a specific domain or a software program. In order to differentiate the two concepts a service is an application that can serve other applications, i.e., a service is an application that offers reusable functionalities that other applications can use over time.

Simply put, the definition of a service is a complex matter that deserves more than a mere technology approach, as stated in (S. Jones 2005): "*A service's intention is to undertake certain functions to provide value to the business; its specification isn't just the direct service it provides but also the environment in which it undertakes those functions*". This statement points to the problem that services have also nonfunctional goals and they could be so important to determine the meaning and the context of the service itself. Also the previous definition, for instance, could be seen as biased: it mention the value for business, but a service could have a value for a community independently from any commercial value and business. A user centric view on services could also change the perspective. Paraphrasing the previous sentence, it could be stated that "*a service's intention is to undertake certain functions to provide value to a person, a group of persons, a community or a business organization and its representatives; its specification isn't just the direct service it provides but also the environment in which it undertakes those functions*".

Services are becoming a crucial topic for many industries, as a consequence a new area is attracting interest from researchers: the Service Science (Service Science 2013). It is an interdisciplinary approach to the study, design, and implementation of services systems (Service Science 2013). In this context a service is even more difficult to define, because the definition encompasses business, social, technology and ecosystems issues. A service system, i.e., a system that can support services is a complex system in which specific arrangements of people and technologies take actions that provide value for others.

As stated in (Cardoso, Voigt and Winkler 2009), "*the set of activities involved in the development of service-based solutions in a systematic and disciplined way that span, and take into account, business and technical perspectives can be referred to as service engineering: Service Engineering is an approach that provides a discipline for using models and techniques to guide the understanding, structure, design, implementation, deployment, documentation, operation, maintenance and modification of electronic services (e-services)*". They propose a new approach to services: services are tradable and can be supported by Internet systems. This combination leads to an Internet of Services that needs to be supported by appropriated Service Architectures and by the Service Engineering.

Summarizing, services strongly depends on many factors like its goals, foreseen users, the ecosystem in which it is used, the system used to be provided and the supporting technologies and mechanisms for using it.

2. 3. Interaction and Control Paradigms

Services considered in this document refer to functions and applications that can be realized by means of ICT technologies. In this context a service system can comprise different subsystems each one providing specific functions and executing particular tasks. In addition the system as a whole needs to interact with its users. Actually the way in which a service system interacts with users is a characterizing factor of the system itself. In fact one first consideration is that users can be seen as external (**Figure 5**) to the service system or as part of it (**Figure 6**).

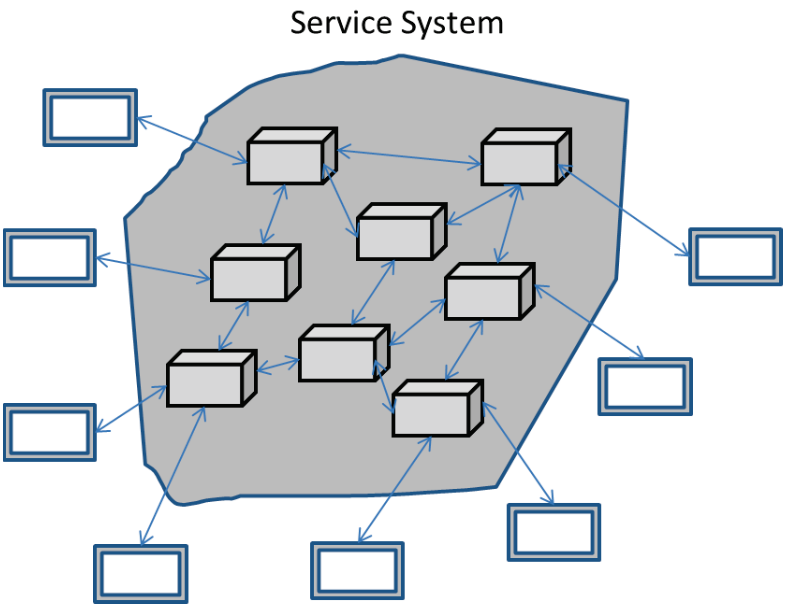


Figure 5: External Users interaction with a Service System

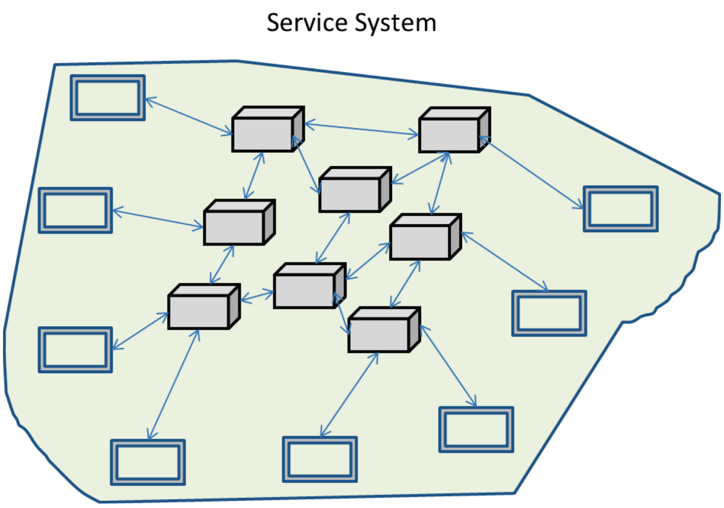


Figure 6: Users as Part of a Service System

In this document the way in which parts and components of the whole service system interact and cooperate is seen as the "control paradigm", while the way in which external entities interact with the service system is termed interaction paradigm. In the case of control paradigm the emphasis is on the architecture of the service system, in the interaction paradigm the focus is on the way users are accessing the system functions. This separation is not a hard one, in fact interaction and control paradigms often are synonymous, e.g., in the case of client - server systems, C-S. In addition, a complex service system could be implementing several control paradigms between its components: many web companies in fact are offering services implementing a highly distributed control paradigm between internal components and parts, and a simple client - server one for user interactions.

With reference to **Figure 5** and **Figure 6**, the difference is relevant: in the first case the relationship is clearly of a dependence, users can access to the system functionalities, that are owned, provided and managed from a specific actor (the provider agent). In the second case the ownership of the service system is blurring, since the users are providing functions and resources to the service system. It is difficult to see the users as external, without them the system would not even able to exist. This is a great difference that points to the client - server , C-S, and the peer to peer, P2P, interaction paradigms as represented in **Figure 7**.

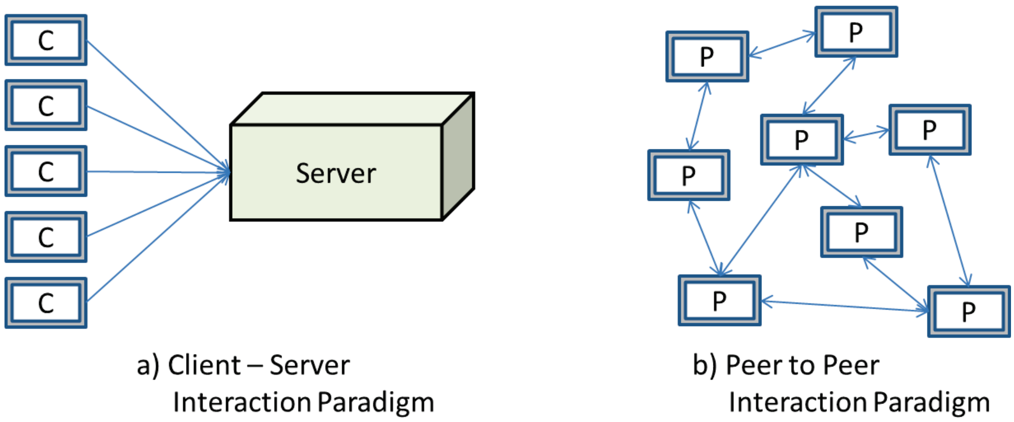


Figure 7: Client – Server and Peer-to-Peer Interaction Paradigms

Graphically the difference is evident, in the client - server the central role is played by the server (that usually is seen as the service system) that is the centralized entity providing valuable functions to the clients that in this relationships have a lesser role. They can request functions and they expect to receive a response. In the P2P paradigm, each single entity or node is equal to the others; there are not privileged relationships and each peer can request services from any other peer. Peers are designed and programmed for cooperation and the service system is totally decentralized. Also the communication mechanisms of P2P and C-S systems are different. In the P2P case, the communication between two communicating peer is direct in the sense that each peer knows the address of the other (even if the communication between them could go through other nodes, multi-hop), in the C-S case the communication is brokered, i.e., the server is in between the clients that do not have a direct reference of each other and are forced to pass through the server (Taylor and Harrison 2009). There is an extensive literature related to P2P and C-S, for example (Pavlopoulos and Cooper 2007), (Orfali, Harkey and Edwards 2007), (Taylor 2005), (Maly, et al. 2003), and (Leopold 2001).

Also the increasing wireless capabilities have raise the question of interaction and control paradigms (Datla, et al. 2012).

The C-S paradigm is based on a very simple interaction between the clients and the server; a client send a request (essentially a structured message) to a server and expects (non blocking) a response from the server. **Figure 8** depicts the simple interaction.

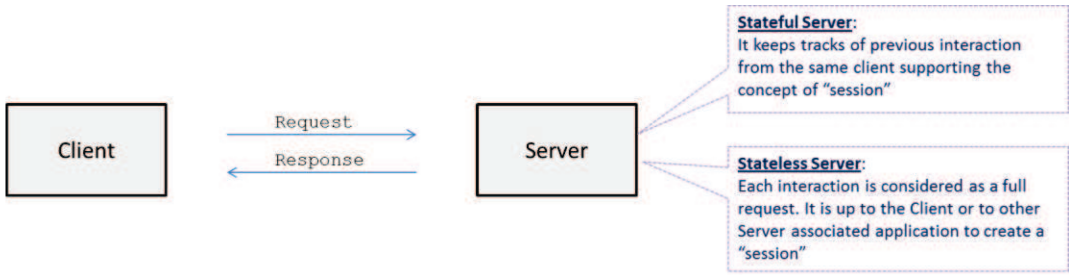


Figure 8: The Client-Server Interaction Paradigm

The Server is assumed to send a `response` to the client, however the client cannot assume that it will be necessarily received. However the `request` and `response` primitives are related and the client and server systems have to take care of this relationship. This means that a `response` cannot be received without a prior `request`. At the same time a client not receiving the `response` could start polling the server in order to get an answer. These examples indicate that the interaction model is not balanced and there are several limitation. The Server could be Stateful or Stateless, the difference is whether the system keep track of the previous interactions with clients and has a finite state machine associated to the interactions going on. A stateful server is more complicated to manage especially if many clients are requesting in parallel the functions of the server. There is the REST architectural proposition (Fielding and Taylor 2002) related to web services and architecture that is promoting with a lot of success the idea of having stateless server as a principle of the architectural design.

Some interaction paradigms in proper sense are [in italics excerpts from (Foster 1995), (Cachin, Guerraoui and Rodrigues 2011)]:

- **Tasks and channels**. It is a computational model designed by (Foster 1995) based on distributed Tasks with logical channels connecting them. A task consists of a Program, local memory, and a collection of I/O ports. Tasks interact by sending messages through channels. A task can send local data values to other tasks via output ports. A task can receive data values from other tasks via input ports, ports are connected by means of channels. The local memory contains the program's instructions and its private data.

- **Message passing**. *Message-passing programs, create multiple tasks, with each task encapsulating local data. Each task is identified by a unique name, and tasks interact by sending and receiving messages to and from named tasks. The message-passing model does not preclude the dynamic creation of tasks, the execution of multiple tasks per processor, or the execution of different programs by different tasks. However, in practice most message-passing systems create a fixed number of identical tasks at program startup and do not allow tasks to be created or destroyed during program execution. These systems are said to implement a single program multiple data (SPMD) programming model because each task executes the same program but operates on different data.* The difference with Tasks and Channels model is that a task needs only a queue to store messages and a queue can be shared among all the tasks that need to communicate with a specific task, while in the Tasks and Channels model, a channel (a queue) is created for supporting the communication between two tasks.
- **Data Parallelism**. *Data parallelism exploits concurrency by applying the same operation to multiple elements of a data structure. A data-parallel program consists of a sequence of such operations. As each operation on each data element can be thought of as an independent task, the natural granularity of a data-parallel computation is small, and the concept of "locality" does not arise naturally. Hence, data-parallel compilers often require the programmer to provide information about how data are to be distributed over processors, in other words, how data are to be partitioned into tasks. The compiler can then translate the data-parallel program into an SPMD formulation, thereby generating communication code automatically.*
- **Shared Memory**. *In the shared-memory programming model, tasks share a common address space, which they read and write asynchronously. Various mechanisms such as locks and semaphores may be used to control access to the shared memory. An advantage of this model from the programmer's point of view is that the notion of data "ownership" is lacking, and hence there is no need to specify explicitly the communication of data from producers to consumers. This model can simplify program development. However, understanding and managing locality becomes more difficult, an important consideration*

(as noted earlier) on most shared-memory architectures. It can also be more difficult to write deterministic programs.

The message passing model is represented in **Figure 9**. It is based on a very simple organization in which a sender forwards messages by means of a queue to a recipients.

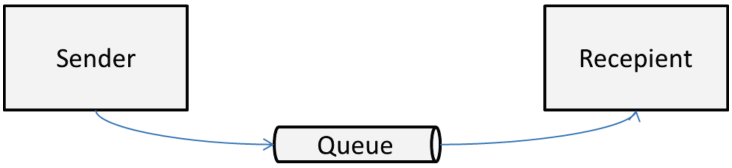


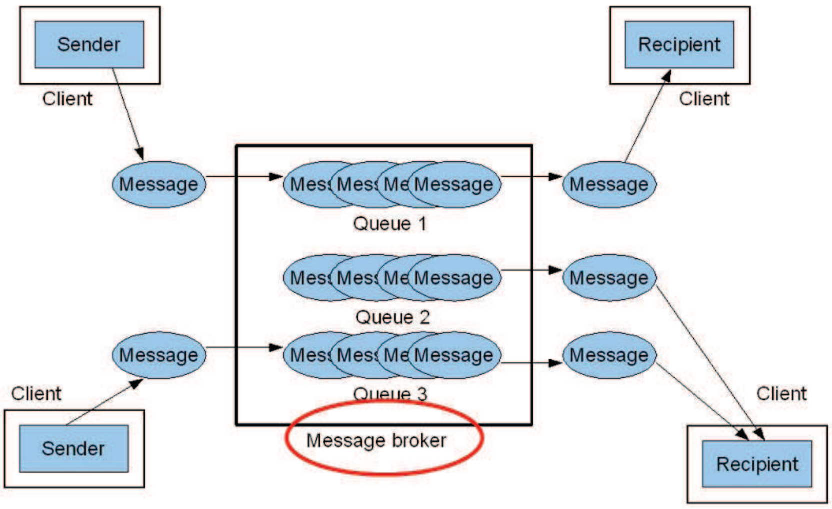
Figure 9: Message Passing Model

Obviously the different mechanisms have advantages and drawback. For a deeper analysis (Kubiatowicz 1998) and (Foster 1995) provide more details and discussions. A short analysis of pros and cons is provided in the following:

- Message Passing: MP paradigm can be interpreted as an "interrupt with data", i.e., event are queued for processing together with associated data (they could also be large bulk of data). Receiving tasks can then operate on data received when the message is taken from the queue. There is the need of an explicit treatment of communication (e.g., "send message to task n") and management of data (data are copied from data structures in the source of the message, transmitted to a queue and then copied to a data structure at the sink destination). Assembling and disassembling of messages can be cumbersome.
- Shared Memory: one of the major advantages is that the communication details between programs and the "shared memory" are hidden to the programmers, freeing them from the problem of dealing with communication and (in the shared memory) with the needs to deal with data management (replication, caching, data coding and encoding, optimization of distribution of data). The programmers can actually focus on the parallelism and the issues to solve. However, the programs have to adopt a sort of polling in order to synchronize with new data and new events (and this can be time and resource consuming), in addition data management is optimized having in mind a system optimization, specific applications needing other optimization policy will find difficult to implement specific application oriented policies. This could be an example of abstraction that eases programming, but sometimes it is detrimental for applications that need to do "different" things.

- Tasks and Channels: this paradigm can be seen as a variation of the message passing. According to (Foster 1995), the "task/channel" paradigm enforces the programmer to better conceptualize the communication of a parallel and distributed program. This could be provide advantages in design a large distributed system, but it also introduces more overhead in the need to control individual channels and in creating a relationship between channels and related tasks.
- Data Parallelism: while this paradigm can lead to significant advantages when the same instructions can be applied to different data, its general application is not easy, in fact not all the application domain offer problems that can be effectively solved with this paradigm. Actually the major drawback of this paradigm is its applicability to more restricted set of applications compared to the previous ones.

In large MP systems, senders and recipients share a common infrastructure made out of queues, these can be organized as a sort of Message Broker in charge for receiving and dispatching (in an asynchronous manner) messages between sources and sinks (**Figure 10**).



http://middleware.sovibox.fi/message-brokers/

Figure 10: A Message Passing MP System

There are at least three important features:

- the model can be synchronous or asynchronous, i.e., in the first case, the sender waits for the receiver to deal with the message, in the second case once a message has been sent, the sender can proceed with its computation. The choice of synchronicity and asynchronicity depends on the applications requirements. The introduction of queues in order to store messages for later processing is a case in which asynchronous mechanisms are implemented.

- routing, the system (especially by means of the Broker Functions) can route the messages even if the sender has not provided a full path to the destination.
- conversion, the Broker functions can also make compatible different messages format in such a way to support the interoperability between different messaging systems.

Communication between the Sender and Receivers is not “brokered”, because the event message manager enables an asynchronous but full communication between the two parties.

The message passing, MP, and the C-S paradigms have communalities: i.e., a sender and a receiver can be easily mapped on a client and a server. However in a MP system the role of sender and receiver are less constraining than in a C – S system, in fact the MP model is more granular (in the sense that the sender could be just sending messages – events – to a receiver without expecting any “response” or any service from the server). In the MP paradigm, messages are not necessarily correlated between a client and a server, and actually each entity can be a sender (and not necessarily a receiver). Under this perspective, the MP paradigm can be used in order to support a C-S model: `request` and `response` primitives can be implemented as two separate messages exchanged between a sender and a recipient. On the other side, using a C – S system for implementing a MP system introduces the logical concept that the receiving party will offer a service to the sender or at least will respond to the received message. This means that the expressiveness of the MP is greater than the C-S paradigm even if at the cost of further complications[2] (**Figure 11**).

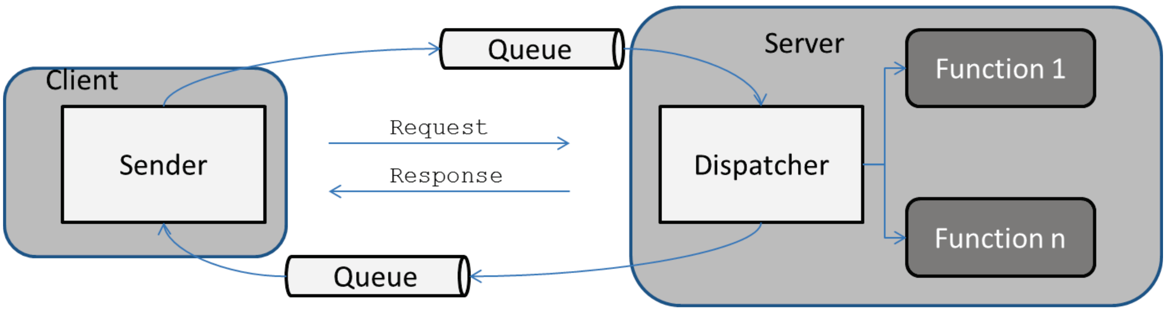


Figure 11: A C-S Relationship Implemented by means of a MP system

[2] Actually a double C-S system could provide the functionalities of a MP paradigm, however the difference between the server side and the client one is such that many client are not capable to host both the client and the server functions.

The P2P paradigm is very close to the Message Passing and the Task and Channel models, in fact each peer can be seen as an entity capable of being a sender and a receiver and in order to communicate is uses channels for sending and receiving data (messages, events, and flows). Actually some P2P protocols, e.g., gnutella (Ripeanu, Foster and Iamnitchi 2002), seem to be designed having the MP paradigm in mind while others are taking the Task/Channel model as the founding concept: e.g., the i2p tunneling mechanisms (Aked 2011).

Other interaction mechanisms are emerging especially for data intensive processing, for instance (Grelck, Scholz and Shafarenko 2010), (Margara and Cugola 2011) analyses the relations between stream processing and the Complex Event Processing.

2. 4. Network Architectures and Distributed Platforms

Even if the interaction paradigms are independent from the network capabilities, in reality they are strongly influenced by the underlying communication infrastructure in several ways. As stated by (Deutsch 1995) and further elaborated by (Rotem-Gal-Oz 2006), (Thampi 2009), there are network fallacies that should not be ignored in designing networked applications. Robustness and resilience of networks cannot be taken for granted and hence service design should take care of utilizing interaction, and control paradigms capable of coping with these fallacies as well as the right partition of functions within different administrative domains (including users domain).

2. 4. 1. Distributed Platforms as a Means to Cope with Network Fallacies

As seen for the message passing MP paradigm, there is the need to design how messages (or events) are passed between the involved entities. In addition, some extra entities besides the communication ones are usually involved in the interactions. For instance, queues and message brokers are networked elements supporting the interaction paradigm. The client - server paradigm strongly relays on the capabilities of supporting communication protocols. There are two major flavors of the Client-Server paradigm: the IT one and Web one.

The Client – Server paradigm as supported by IT technologies

The first one is based on Remote Procedure Call mechanisms as depicted in **Figure 12**.

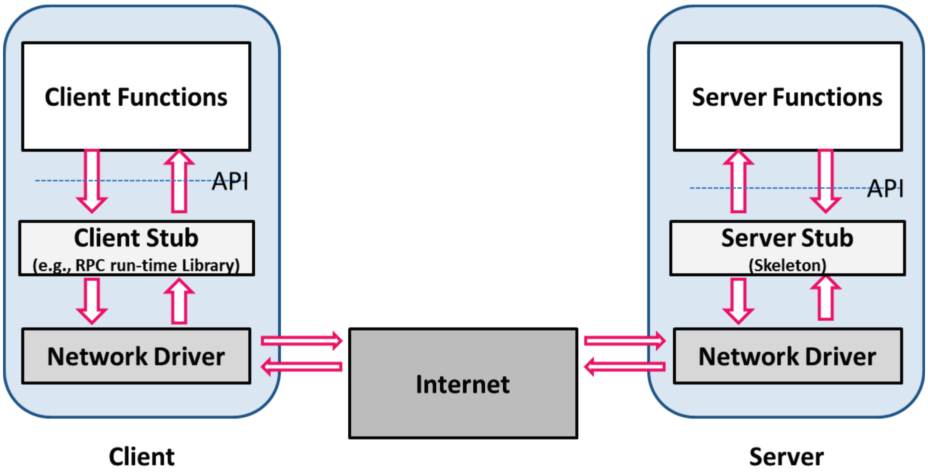


Figure 12: The Client-Server Paradigm and Remote Procedure Call Mechanisms

A few observations are useful: if the client and the server functions are synchronous, the underlying mechanisms are based on events/messages, in fact the network drivers will fire events and messages as soon as they arrive; the underlying network is based on IP communications with the two options of connection oriented TPC or connectionless UDP protocols, this means that requests can pop up asynchronously and the network driver is essentially a sort of internal queue manager; the Upper Layer functions are usually provided by means of Application Programming Interfaces, API. An API (magic 2012), (API 2013), is a source code-based specification intended to be used as an interface by software components to communicate with each other. An API may include specifications for routines, data structures, object classes, and variables. A difference between an API and a protocol is that the protocol defines a standard way to exchange requests and responses between functional entities (to be seen as black boxes) while APIs expose a part of the software infrastructure (and organization within the functional entity). An API can be:

- language-dependent, meaning it is only available by using the syntax and elements of a particular language, which makes the API more convenient to use.
- language-independent, written so that it can be called from several programming languages. This is a desirable feature for a service-oriented API that is not bound to a specific process or system and may be provided as remote procedure calls or web services.

This approach paved the way towards middleware platforms, the initial RPC systems were in fact proprietary and confined to well defined operating systems with machines essentially interconnected by enterprise networks. From this stage, the introduction of middleware platforms, i.e., software functions in between applications and operating system capable of creating a sort of (distributed) unified platform supporting these functions as stated in (Krakowiak 2003):

- *Hiding distribution, i.e. the fact that an application is usually made up of many interconnected parts running in distributed locations;*
- *Hiding the heterogeneity of the various hardware components, operating systems and communication protocols;*
- *Providing uniform, standard, high-level interfaces to the application developers and integrators, so that applications can be easily composed, reused, ported, and made to interoperate;*
- *Supplying a set of common services to perform various general purpose functions, in order to avoid duplicating efforts and to facilitate collaboration between applications.*

This approach was the basis for DCOM (Krieger and Adler 1998) and CORBA (Object Management Group 2006), (Chung, et al. 1998), (He and Xu 2012). More recently this approach has led to the definition of the OSGi platforms (Kwon, Tilevich and Cook 2011), (Kwon, Tilevich and Cook 2010) as represented in **Figure 13**.

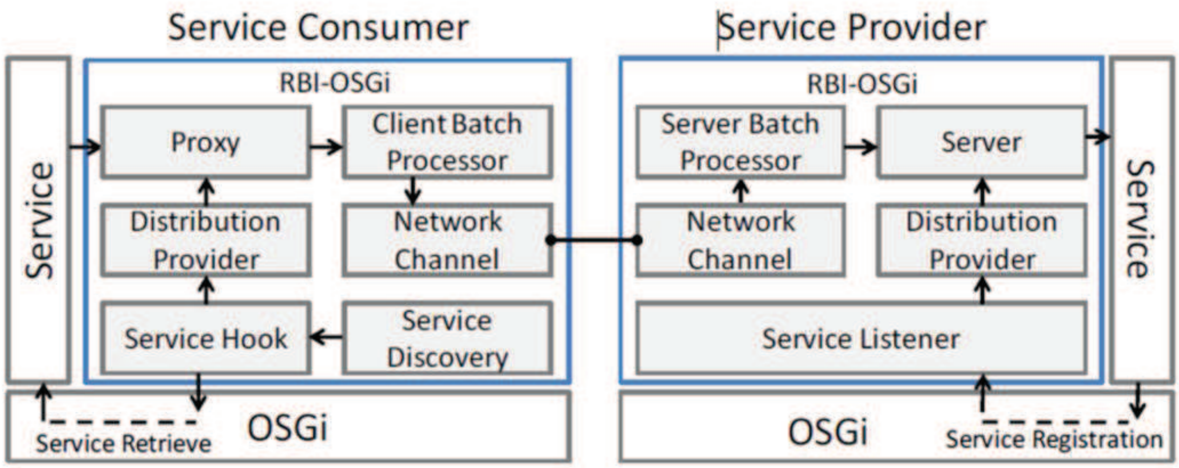


Figure 13: RBI/OSGi Architecture

The Client – Server paradigm as supported by Web technologies

The other trend within Client - Server paradigm is based on Web Technologies. In this case the basis is the http protocol used as a means to send and receive requests and responses between a remote client and a server (**Figure 14**). XML_RPC (Laurent, et al. 2001) is a very simple and effective protocol for C-S interactions based on Http for transport and XML for information representation. Remote procedure invocation and responses are coded with XML and the payload is forwarded by means of HTTP.

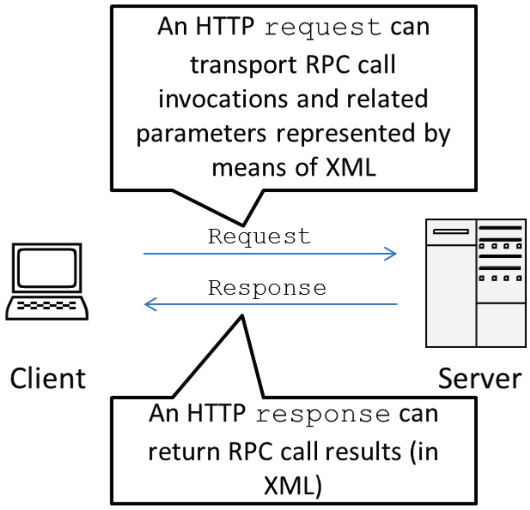


Figure 14: RPC Interactions based on Http

The simple idea that an RPC mechanism can be built using XML and HTTP (Richards 2006) has started a new view of creating web services, i.e., the possibility to invoke dynamically services in the web directly from a client remote application. This approach has led to the definition of new protocols such as SOAP as well as to new approaches in the software architectures (e.g., REST). For a while these two approaches diverged: a lot of effort was put on specification of Web Services Architecture Service Oriented Architectures. However the increasing importance of the REST approach has suggested the need to integrate the two approaches under the web services definition (see 2. 5.).

2. 4. 2. Control Paradigms and Network Architectures

The combination of the communication issues, i.e., how communications capabilities are supported, offered, and used by systems that implement services, and the interaction paradigms themselves define a set of control paradigms:

- the client – server one, C-S, that exploits protocols, architectures and capabilities of the web and IT industry in order to support services by means of networked resources;
- the network intelligence one, NI, that combines event based processing with the network capabilities and resources;

- the peer to peer approach, P2P, that combines the message passing mechanisms (or the event processing one) with distributed capabilities offered control by the different peers over an overlay network.

As stated in (Santoro 2006), the C-S paradigm is not properly a distributed processing paradigm: "*Incredibly, the terms "distributed systems" and "distributed computing" have been for years hijacked and (ab)used to describe very limited systems and low-level solutions (e.g., client - server) that have little to do with distributed computing*". The C – S paradigm can be seen as a by product of a set of protocols, functionalities and mechanisms that allow the distributed communication, interaction and coordinated processing. The C –S paradigm does not really contributes to the advancement of distributed processing, it just relies on those techniques. However, it is considered in here for its profound impact and importance in the evolution of service platforms.

Services are built around interaction and control mechanisms. Users can access and take advantage of service functionalities by interacting with a system. For instance the C-S one is very simple but it requires the "client" to activate the service (pull) and waiting for a response. Generally, it is not envisaged that a "server" will autonomously initiate an interaction towards the user (a push). Thus is obviously a limitation, services are designed in such a way to adhere to this interaction paradigm, and not all the services can be appropriately provided in a client -server fashion. For instance a simple service like "inform the client when a condition on certain data is met" (event notification) can be tricky to implement in a client – server fashion, but it could be simple to realize with another paradigm (e.g., a P2P implementation based on a message passing mechanism). The issue with the client – server paradigm is a well know one (see for instance (Adler 1996): "Important examples include task allocation and event notification in collaborative workgroup systems, and task sequencing and routing in workflow applications") and it has an impact especially when users/tasks need to coordinate several threads of processing and information distribution. Sometimes, in order to circumvent the issue a continuous polling of the client on the server can be implemented, but this is not an optimized solution. Solutions like Websockets (Hickson 2011) can be seen as a way to solve this problem in the context of HTTP client server communication. Here a socket is created and the server can forward on the socket asynchronous information. In this way the "difference" between the client and the server is reduced and this architecture can be seen as a move towards P2P solutions also for the Web. The other paradigm, the Network Intelligence one, is even more limited. It is based on the network capability to determine that the user is requesting service functionalities. The network "control" then triggers (i.e., it requests for assistance to a set of service layer functionalities) to the specific service. In this case the service is confined to the network context in which the user is operating, it is also confined by the protocols and mechanisms of the network. The P2P paradigm seems to be the most adaptive, it supports a granular message based (or event processing) approach that allows to design the service functions very effectively at a cost of representing a logical network of relationships between peers.

The paradigms have different "expressive power" ((Minerva 2008) and [XVIII]), i.e., different capabilities to represent and support the control needs of a service. Actions and interactions within the parts of a service system can be represented and supported in different fashions and with different mechanisms, expressive power means here how easily the control paradigm can represent and support the needed communications and control of the components. Introducing functions to circumvent problems that could be easily solved with other paradigms could be a sign of low expressive power, examples are the polling mechanisms for C - S systems compared to message passing systems or event driven programming, or the oversimplification of ParlayX APIs that do not allow the support of anonymity in services. An exemplification of the lack of expressive power in C – S paradigm is represented by WebSocket definition (Fette and Melnikov 2011): a new mechanism that extends the paradigm capabilities had to be introduced in order to allow the server side to autonomously notify the occurrence of events of partial data to the client side. Lack of expressive power can also yield to a phenomena called Abstraction Inversion: users of a function need functionalities implemented within that are not exposed at its interface, i.e., the need to recreate functions that are available at lower levels. Abstraction Inversion, too much abstraction, and wrong control paradigm can lead to anti-patterns in intelligent systems (Laplante, Hoffman and Klein 2007), with consequences in the ability of platforms to support services, increase in cost of software development, longer processes to delivery phase. Generally put, if a control paradigm requires too much adaptation and development of specific functions for representing and support the control flows, that means that that paradigm is not well suited for that service.

The issue of expressive power is a fundamental one: the possibility to implement services with the best model for supporting the interactions between its parts should have a large importance when platforms and systems are designed, implemented or simply bought. Instead it is often neglected and service offerings are created on the available platform totally disregarding how the final service will be provided. This has led (especially in the telecommunication world) to aberration in which highly interactive and cooperative services have been implemented with a very constraint paradigm (Minerva 2008) (Minerva 2008, 1) such as Network Intelligence (see for instance the attempt to implement Video On Demand, VoD, over IMS (Riede, Al-Hezmi and Magedanz 2008)). The Interaction and Control paradigms are at the very heart of the service design and implementation, and this issue is one of the crucial considerations for determining how to create effective service platforms.

The considered control paradigms are depicted in **Figure 15** and are briefly discussed in this section.

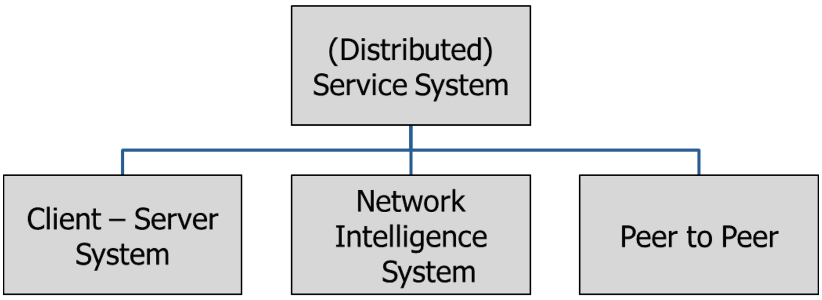


Figure 15: Control Paradigms in the Context of Networked Interactions

The C-S control paradigm

The C-S control paradigm is the simplest one, it is based on the concept that the network is essentially transparent to the functional interactions between the clients and the server. It clearly and cleverly exploits the end-to-end principle (Saltzer, Reed and Clark 1984) of the Internet in order to opportunistically use the network capabilities. This principle states that "*mechanisms should not be enforced in the network if they can be deployed at end nodes, and that the core of the network should provide general services, not those tailored to specific applications*". So the C-S paradigm is a means to move the intelligence at the edges of the network and confining the network to provide generic and reusable services related to transport. This has given rise to the concept of Stupid Network as proposed by (Isenberg 1998) (**Figure 16**). Represents the concept of the stupid network.

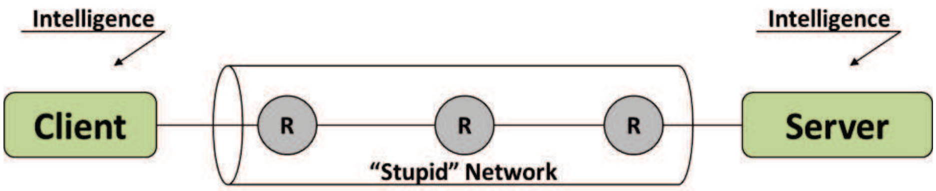


Figure 16: The Stupid Network and the Aggregation of Intelligence at the Edges

As stated in [XVIII], the C-S control paradigm has several merits:

- Simple control pattern
- Implemented in stateless or stateful manner
- Decoupling of application functions from the network intricacies (service deperimeterization)

Its major drawback is that not always the network has an ideal behavior and there is often the need to orchestrate resources and capabilities in order to provide compelling services.

The Network Intelligence Control Paradigm

It is the typical control paradigm used by the telecoms industry in order to provide services. It leverages the capabilities of the network in order to provide services. **Figure 17** illustrates the basic mechanisms behind it. In this case clients can interact with the control functions of the "network" and only when it is not possible to serve the request the service is "triggered". There are two interaction models coupled in this control paradigm: a sort of client – server between the users of the "network service" and an event driven model (Dabek, et al. 2002) for the interaction between the control functions and the (added value) service functions.

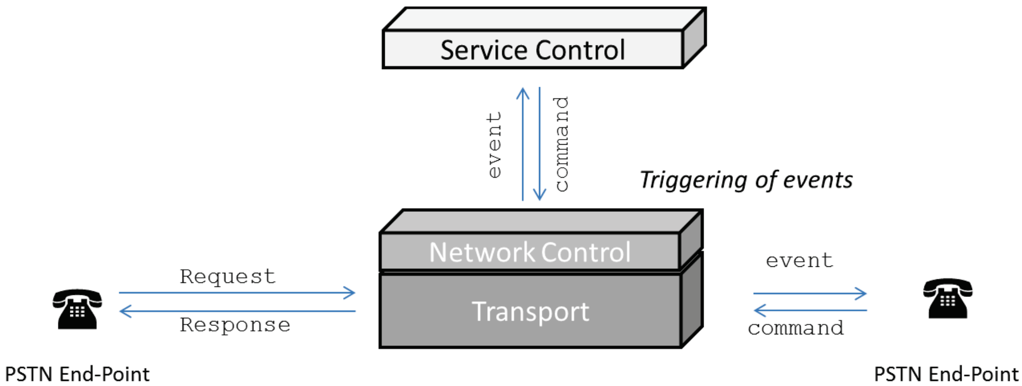


Figure 17: Triggering of Events as an Interaction Paradigm

The composition of interaction model is a bit complicated in this control paradigm. The caller (or the initial requestor) could be considered in a stateful C-S relation with the network. The network control triggers events towards a service control entity when it has no instructions on how to proceed with the service request. This relation is not client - server , because more (correlated) events can be sent by one party to the other. It is more an event driven approach. Also the callee or the destination of the communication is in an event driven relationship with the "service system". One evidence is that the network control is in charge for the orchestration of services. There is no way for the end-points to request directly services and functions to the service control. This layer is mediated by the "network". This approach was meaningful when the end point were stupid, by nowadays, terminals have plenty of intelligent capabilities and they could be capable of benefitting from a direct interaction with the service control and even from the direct communication with network resources. This dependence of the services from the network is a major drawback. In addition the complexity of this model is even complicated by the fact that all the interactions are stateful and organized around very complex finite state machines (Minerva and Moiso 2004). The FSMs are associated to the concept of call and more recently session control model. Typically a service progress along the states and the possibilities offered by a very complex and controlled model that does not offer too much flexibility to the programmers (Minerva 2008, 1), in addition the call and sessions are strongly coupled with network resource allocation. For instance in **Figure 18**, yellow resources are allocated to a session.

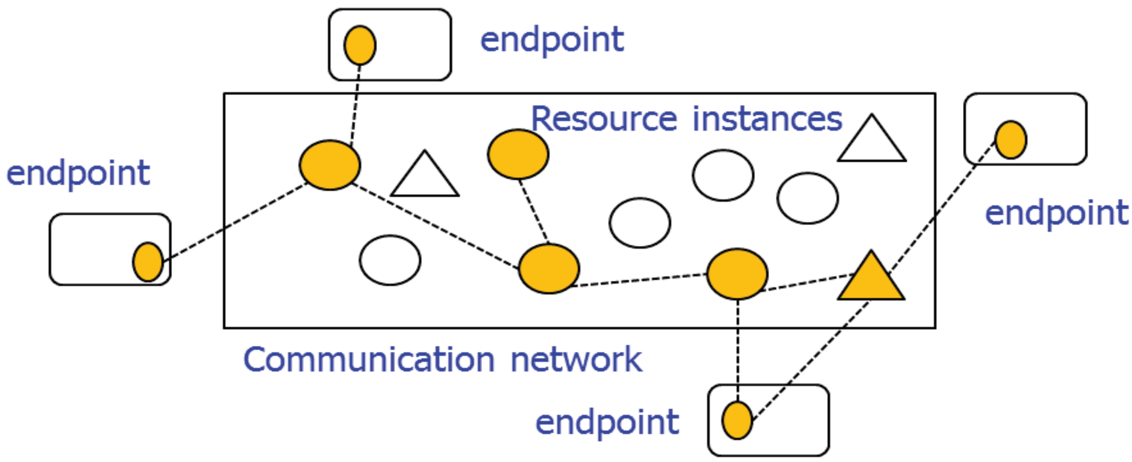


Figure 18: The Session Concept and the Allocation of Resources

Resources are made available only within a session in order to access to another resource, the session control has to negotiate for it, while the endpoints have only the possibility to interact with the session control in order to request a mediated control of new resources. The possibility for endpoints (with increasing level of intelligence) of directly interact with a specific resource is totally neglected by the Network Intelligence, NI, control paradigm. This leave outside of the service control those elements that have the major requirements to do so.

This approach has been successful for the Intelligent Network (Faynberg, et al. 1996), and it has been reiterated for many of the new control architectures for telecoms such as IP Multimedia SubSystem, IMS, (Camarillo and Garcia-Martin 2007), and Next Generation Networks in general (Copeland 2009). The IMS in particular has been considered as the new control platform to be used in order to offer the integration between telecoms and internet services. This assumption is based on the following features offered by the IMS:

- Multi access Networks, i.e., the capability to support several heterogeneous access networks
- Layering and decoupling of functions at the transport, control and service level
- Reuse of the successful GPRS functional architecture

IMS is then seen as a reference architecture for Next Generation Networks. This approach has some pros and cons, like:

- Regulated and Standardized
- Support of interworking
- Constrained and Limited
  - New add-ons are difficult to introduce in the software
  - Stateful model
  - Session Control is just an evolution of the call control
- Complexity increases with number of new functional entities and states (for instance the presence of different types of Call State Control Functions, CSCF, give rise to the need of several protocol interaction in order to sort out the Functions orchestrating the service).

This is a very traditional approach that disregards the issues in interaction and control paradigms, the evolution of terminals and edge functionalities, the current development of the major projects in the Future Internet such as GENI (Peterson and Wroclawski 2007), AKARI (Aoyama 2009) and others (Pan, Paul and Jain 2011). These new architectures consider different interaction and control paradigms (highly distribution of functions, P2P, mesh networking) that greatly differ from traditional ones. For a good overview of Network

Intelligence the interested reader could refer to the ICIN conference that collects many contributions in this field. For a traditional approach and its mainstream developments in the exploitation of Network Intelligence good source are (NGNI Group 2005), (NGNI Group 2006), (NGNI group 2009), and (NGNI Group 2012). **Figure 19** represents a (desired) mainstream evolution of technologies and system from a telecoms perspective.

Figure 19: The NGNI Playground Evolution

Summarizing the pros and cons of the NI approach are:

- Mix of different mechanisms (C/S and event – commands)
- Generally implemented in a stateful manner
- Strong dependence of application functions from network ones (e.g., the triggering mechanism is the way in which services can be activated creating a strong dependence of services on the status of the network).

Some other major concerns in the NI approach are: the perimeterization of services (i.e., services are tightly coupled with the network functionalities and its geographical extension) and the assumption that QoS is a major requirement for service. They have been discussed in [XXXI] and [XIV] showing how Operators cannot compete at a global level because their services cannot be provided outside of the deployed and owned network and arguing that QoS and more bandwidth are not always practical solutions for solving service issues (in fact, Google solved the problem of long map downloading time by introducing AJAX solutions based on asynchronous download of data[3]).

More about the traditional approach of TelCos can be found in 8. 6.

The Peer To Peer Control Paradigm

[3] Actually this is another evidence that the expressive power of the C – S paradigm is not sufficient to provide services that do require responsive control on the flow of data. Introducing asynchronous (and invisible to the users) calls between the client and the server is pointing to the need for more capable control mechanisms. Actually those mechanisms are also provided in WebSocket that definitely change the C – S paradigm (in fact the server can use now the socket to notify events and messages to the client)

Apart from the different topologies and characterization of structured, unstructured and hybrid peer to peer networks (Lua, et al. 2005) (Merwe, Dawoud and McDonald 2007), two concepts are central to P2P networks: equality of peers, i.e., each node is potentially a client and a server of the network; overlaying and virtualization, i.e., the creation of a logical virtual network on top of physical resources. Another important aspect is the dynamicity of the network: nodes leave and join dynamically and the average uptime of individual nodes is relatively low. The topology of an overlay network may change all the time. Once a route is established, there is no guarantee of the length of time that it will be valid. Routing in these networks is therefore very problematic. In other terms the P2P control paradigm is complicated by the need to manage and keep track of a highly dynamic infrastructure that changes over time, on the other side the interaction between peer could be implemented adopting the most appropriated (for the specific service) interaction mechanism. So the a P2P infrastructure could be able to encompass several interaction paradigms offering to services the most suitable mechanisms.

P2P network Systems are characterized by:

- Complexity of the entire system and simplicity of the single node
    - Nodes are both clients and servers
    - P2P systems are real distributed systems, they need to govern an overlay network (resource naming, addressing, location)
- Scalability and Robustness
    - P2P systems can scale to millions of nodes
    - Problems in detecting the source of issues in case of malfunctions
- Different way of programming
    - P2P systems can be implemented as event-driven system but at the end the communication between two terminal nodes can be client-server
- Social flavor of the solutions (Koskela, et al. 2013). The willingness to share and the thresholds at which an altruistic behavior is triggered have been studied (Ohtsuki, et al. 2006). They are particularly meaningful to understand when a P2P network will became valuable to the single user.

Summarizing the P2P approach seems to be very flexible and promising; some pros and cons are:

- Able to support many control patterns
- Generally implemented in a stateful manner (the overlay control)
- Intertwining of application functions with network ones.

2. 5. Service Architectures

The interaction and control mechanisms have to be framed and supported by software architectures, i.e., a software architecture consists of a set of concepts, principles, rules and guidelines for constructing, deploying, operating and withdrawing services. It also describes the environment in which such services operate. The service architecture identifies components to build services, describe the way they are combined, and the way they interact. The combination of organization principles of the architecture, the underlying technology platform and the enabled interaction and control paradigms determine the expressive power of the service architectures and its capabilities in supporting several classes of services. The expressive power is the capability of services and application to represent the real world interactions and to organize the available functions in such a way to accommodate the applications and users needs.

There are major attempts to define software architectures (Proper 2010), (Lankhorst 2013) and their expressive power with respect to the problem domain representation and the involved stakeholders.

The Web and IT views on Service Architectures

The evolution trajectory of Web and IT service architecture is defined by different factors such as:

- Evolution of web servers towards application server architectures (Minch 2009);
- Evolution of Web Service (Booth, et al. 2004) and Service Oriented Architectures (Partridge and Bailey 2010) and the competition between SOAP and REST based solutions;
- The development of advanced solutions within data centers that have led to Cloud Computing (Gong, Ying and Lin 2012) opposed to Grid Computing (Foster, Zhao, et al. 2008), (Sadashiv and Kumar 2011);
- Virtualization of resources (Chowdhury and Boutaba 2010), (Sahoo, Mohapatra and Lath 2010) and the Software as a Service models (Prodan and Ostermann 2009).

An attempt to standardize the mechanisms behind the simple XML_RPC approach has been done with SOAP (Simple Object Access Protocol, (Box, et al. 2000)). SOAP provides a simple and lightweight mechanism for exchanging structured and typed information between peers in a decentralized, distributed environment using XML. SOAP consists of three parts:

- The SOAP envelope construct defines an overall framework for expressing what is in a message; who should deal with it, and whether it is optional or mandatory.
- The SOAP encoding rules defines a serialization mechanism that can be used to exchange instances of application-defined data types.
- The SOAP RPC representation defines a convention that can be used to represent remote procedure calls and responses.

The definition of SOAP was instrumental to the growth of the Web Service Architecture, Web services provide a systematic and extensible framework for application-to-application interaction, built on top of existing Web protocols and based on open XML standards (Curbera, Duftler, et al. 2002). The Web services framework is divided in three areas: communication protocols, service descriptions, and service discovery. They correspond to three protocol specifications that are the cornerstones of the Web Services Architecture (Curbera, Leymann, et al. 2005):

- the mentioned simple object access protocol which supports communication between Web services;
- the Web Services Description Language (WSDL, (Chinnici, et al. 2007)), which specifies a formal, computer-readable description of Web services;
- the Universal Description, Discovery, and Integration (UDDI, (Clement, et al. 2004)) directory, which is a registry of Web services descriptions.

These protocols form the skeleton of the well-known Web Service Architecture represented in **Figure 20**. The Service Provider uses the UDDI functions in order to register and advertise its services on the Universal Registry. A Service Requestor can access to the UDDI functionality and search for the most suitable service for its needs. It gets a WDSL description that can be used in order to correctly access to the service. At this point the service requestor and the service provider can be linked and can use the SOAP Protocol in order to access and provide the service functions.

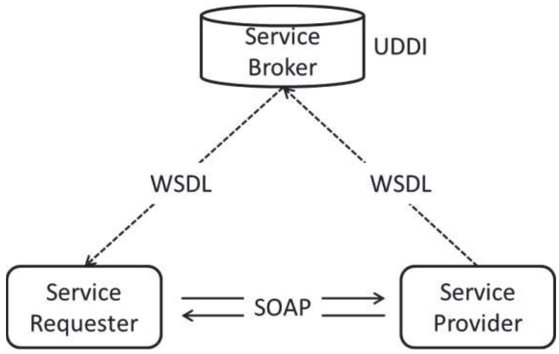


Figure 20: The Web Services Architecture

The specification has been extensive and the original simplicity of the solution has led to a complex set of documents and to overall complex implementations. Essentially the UDDI service, originally intended for use in a dynamic way, is used off-line in order to keep track of the software components made available. The SOAP protocol itself has been disregarded in favor of simplest solutions based on the REST approach (Fielding and Taylor 2002). REST is a set of architectural principles for designing and implementing distributed systems. It proposes the use of web-based interfaces that are described in XML and handled through HTTP. The REST principles are:

- A stateless client-server protocol
- A set of well-defined operations (e.g., GET, POST, PUT, and DELETE) to perform functions on resources
- A Universal Syntax for resource identification (resources are identified by URIs)
- Use of XML or HTML for describing pieces of information and their links.

REST is at the core of Web Application Development as shown by **Figure 21**. This figure shows that whenever a simple solution is offered and it is based on well-known mechanisms then the programmers will adopt it. This is the case for REST: it is based on established principles of the web programming (each function is an addressable resource and the mechanisms to interact are those supported by the http protocol) and it is simpler than other protocol and related solutions (like SOAP). Web programmers use it much more than other solutions.

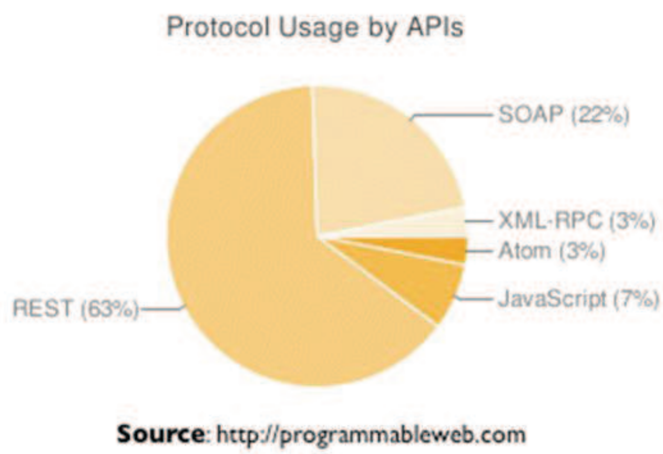


Figure 21: Protocol Usage by APIs

The Web Service Architecture has mingled with the definition of Service Oriented Architectures, SOA (Sward and Boleng 2012), (SOA-Wikipedia 2013). Service-Oriented Architecture is a software architectural concept that defines how to use services to support the requirements of software users. In a SOA environment, nodes on a network make resources available to other participants in the network as independent services that the participants access in a standardized way.

Most definitions of SOA identify the use of Web services (i.e. using SOAP or REST) in their implementation. However, SOA can be implemented using any service-based technology.

Unlike traditional object-oriented architectures, SOAs comprise loosely coupled (joined), highly interoperable application services. Because these services interoperate over different development technologies (such as Java and .NET), the software components become very reusable, due to the virtue of the interface definition being defined in a standards compliant manner (WSDL) which encapsulates / hides the vendor / language specific implementation from the calling client / service. The relations between SOA and Web Services were first of competition and then of integration, nowadays SOAs can be implemented as Web Service Architectures.

The availability within data center of storage and processing capabilities has led to the development of Cloud Computing, i.e., "*a large-scale distributed computing paradigm that is driven by economies of scale, in which a pool of abstracted, virtualized, dynamically-scalable, managed computing power, storage, platforms, and services are delivered on demand to external customers over the Internet*" (Foster, Zhao, et al. 2008). This concept seems to be very close to the grid computing. Very similar in scope, Cloud Computing and Grid Computing differ in at least to features that are important for this document:

- The standardization and openness of the approach
- The control and optimization of the underlying resources with special emphasis on the network ones.

The Grid in fact has pursued a standardization of its solution from its incipit and it has considered resources with holistic and broad view. **Figure 22** represents the two different approaches: the grid aims at controlling resources and groups of them also deployed in different administrative domains, while the Cloud (at least originally) was aiming to homogeneous and single administrative domains. In addition the Grid is aiming at granular control on the resources and it also fosters the control on connectivity as one major goal.

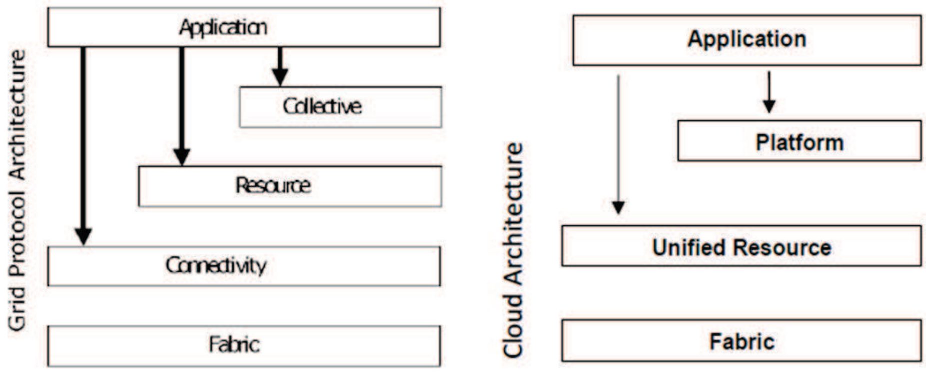


Figure 22: Grid and Cloud Computing Architectures

Only recently, cloud computing (and the Web Companies behind it) have started to work with the goal to achieve a better control of connectivity. For instance Google has unveiled their effort to build a large scale controllable network (Vahdat 2012).

Contextually to the development of Cloud Computing, virtualization technologies have made such progresses to become a major technology trends in several application domains. Resource virtualization is the possibility to create a virtual machine that acts and behave like a real resources on top of another hardware and software resource. Software executed on these virtual machines is separated from the underlying hardware resources. The benefits resides on the possibility to decouple even further the software infrastructure from the underlying hardware and to segment on a cluster of machines different applications and resources limiting the occurrence that the fault of a (virtual) machine could impact on other machines and resources (Metzler 2011). Virtualization was instrumental to the possibility to introduce

Cloud Computing (Lenk, et al. 2009) and the Software as a Service (Turner, Budgen and Brereton 2003) model to a large audience. Its extension has led to the new trend of virtualizing everything as a service (XaaS, (Schaffer 2009)) and to offer these features by means of a cloud infrastructure.

Summarizing, the Web and IT world is aiming at very large structures that can virtualize processing, storage resources and can offer compelling services developed and organized in reusable building block. Resources are addressable with web mechanisms and they offer APIs that can be mashed up and combined in order to provide new services and applications. In spite of standardization these solutions aim at proprietary environments in order to capitalize a technology investment and advantage.

The Network Intelligence view on Service Architectures

The evolution of service architecture in a Network Intelligence context is characterized by a two competing approaches and visions:

- the traditional TelCo one that can be seen as an evolutionary path from the ISDN, IN and IMS architectures;
- an innovative one that stems from open platforms and in particular TINA and goes through Parlay and the Service Delivery Platform in order to leverage Network APIs.

The two approaches have often been competing for driving the evolution of the Next Generation Networks, NGN, but nowadays they stick together because they have become rather traditional and generally obsolete. The windows of opportunity for opening up network interfaces, building a new programmable architecture and offering new services was around 2000-2002 (along the initial deliveries of Web 2.0 platforms). After that period, the window of opportunity closed as discussed in several papers that have identified the decline of the network intelligence (Minerva, Moiso and Viviani 1998), (Minerva and Moiso 2000), (Licciardi, Minerva and Cuda 2000), (Minerva and Moiso 2000), (Minerva and Moiso 2004) and [V].

The current attempts to repositioning the Operators in the Web 2.0 wave (Maes 2010) are probably useless and are consuming resources that should be spent for other scenarios (see for instance 8. 8. and 8. 10.)

The main contributions in this sector are:

- Architectural definitions for TINA (Berndt and Minerva 1995), (H. a. Berndt 1999), (Yagi, et al. 1995), Parlay (Walkden, et al. 2002), Open Service Access (OSA, (Moerdijk and Klostermann 2003)) and Service Delivery Platform, SDP (Moriana Group 2004), (Baglietto, et al. 2012);
- Definition of Open Interfaces like Parlay (Di Caprio and Moiso 2003), ParlayX (Parlay Group 2006) and oneAPI (GSM Association 2010);
- Definition of the IMS architecture (Knightson, Morita and Towle 2005) and usage of the SIP protocol (Johnston 2009) as building blocks for a capable NGN (TISPAN 2009);
- Integration of IMS and SDP (Pavlovski 2007).

A service delivery platform (or SDP) is a centralized hub for the creation and integration of all applications and services offered within a network. It is implemented by or on behalf of a network operator, and resides in the "IT" part of the operator's infrastructure.

The SDP allows applications to be abstracted from bearers, channels, enablers and operational support systems. It goes hand-in-hand with an architectural framework which describes how it interfaces with the other systems in the operator's infrastructure, including external system. An SDP is used in order to "expose" APIs. They can be open also to third parties. The reason for adopting the SDP is on the TelCo's infrastructure and the possibility to monetize the service exposure, i.e., the offering of APIs, also to third parties developers. The SDP is filling a gap in the definition of the IMS at the service layer. IMS service layer is essentially a collection of different application server (SIP, OSA, CAMEL) not operating in a synergistic way. The goal of SDP is to integrate the different functions of these servers and to integrate them with the management infrastructure of the TelCo. To simplify the reasoning behind the equation NGN = IMS + SDP should represent the evolutionary strategy of Operators. This assumption will be questioned in 8. 6.

The Peer to Peer View on Service Architectures

Peer to Peer systems have the goal to optimize the usage of peer resources in order to meet the requirements of a community of users. For instance the Invisible Internet Project, I2P, aims at making the communication between peers secure and anonymized. In order to achieve this goal, the overlay logical network is organized around a few building blocks that can be combined to create an infrastructure on top of which services and applications can be executed guaranteeing the users to remain anonymous (Aked 2011). Any peer will try to create encrypted tunnels towards the peers it wants to communicates to. Tunnels will be outbound and inbounds. The needed peers for creating tunnels are selected by means of a distributed Network database. When the communication has to be created a Lease will be issued so that the inbound and inbound tunnels can be connected and the end parties can communicate. **Figure 23** illustrates the mechanisms.

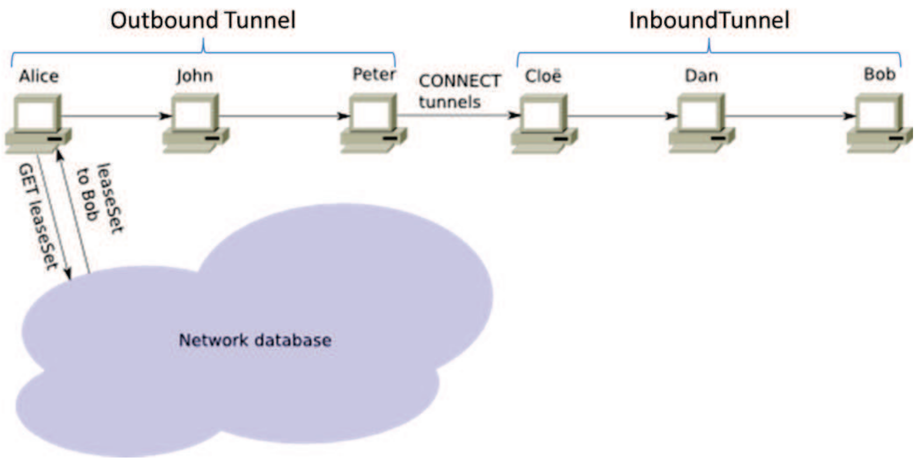


Figure 23: Invisible Internet Project, I2P

Once the network between peers has been constructed, then common applications can be launched and used taking advantage of the new infrastructure.

Another relevant P2P architecture is JXTA (Gradecki 2002), whose goal is to support the creation and the usage of large distributed application using the functionalities offered by specialized peers. JXTA decouples basic functions from services and the different applications. At the lower levels, mechanisms for creating pipes between peers, group them, identify them are provided. On a higher level services for discovering of peers, determining the group associated to a specific peer and others are provided. At the upper level, applications are created and used exploiting the JXTA functionalities. **Figure 24** depicts the architecture.

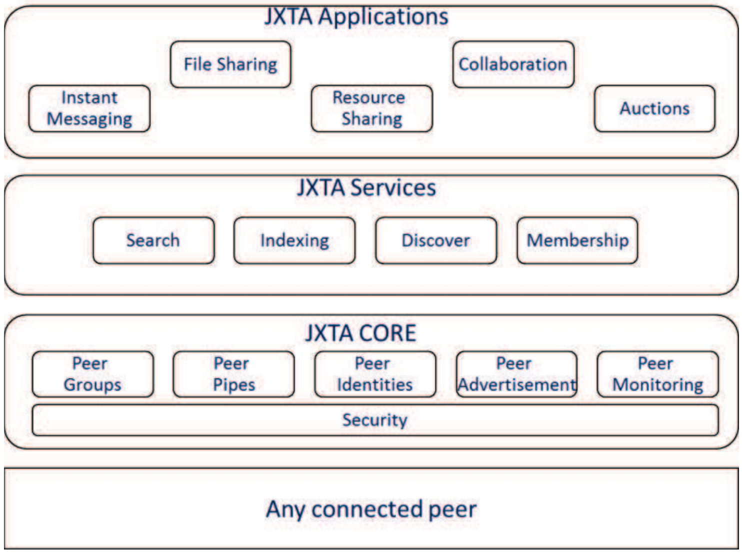


Figure 24: JXTA Architecture

Some consideration about P2P architectures are:

- Complexity of the entire system and simplicity of the single node
    - Nodes are both clients and servers
- P2P systems are real distributed systems,
    - they need to govern an overlay network (resource naming, addressing, location)
    - the overlay network is totally decoupled by the underlying physical infrastructure. This can introduce inefficiency in how network resources are used.
- Scalability and Robustness
    - P2P systems can scale to millions of nodes
    - Problems in detecting the source of issues in case of malfunctions
- Different way of programming
    - P2P systems can be implemented as event-driven system but at the end the communication between two terminal nodes can be client-server

An important property of P2P systems is their power (Delaney, and Catarci and Little 2001), [VI], and in particular scalability (Hu and Liao 2004), (Krishnamurthy, Wang and Xie 2001) but also availability (Bhagwan, Savage and Voelker 2003) and other properties (Pouwelse, et al. 2005). Peer to Peer networks offer the possibility to aggregate computing, communication and storage capabilities made available by participating peers and to scale up to very large systems. In comparison a Data Center requires a lot of investments and planning in order to achieve the desired/needed capacity and scalability.

The total power of a data center is fixed and is given by the number of its servers, and related processing and storage capabilities as well as the total bandwidth associated to the data center, in other terms the "power" can be indicated as Power(Client-Server System) = $\{b_S, s_S, f_S, p_S\}$

Where:

b_S = bandwidth of the Server System

s_S = storage of the Server System

p_S = processing in the Server System

In a P2P system, the total power is given by the aggregation of individual contributions of peers, so it grows with the number of participants, in a way the "power" could be indicated asPower (P2P System) = $\sum (b_i, s_i, f_i, p_i)$

Where:

b_i = bandwidth of node i

s_i = storage of node i

p_i = processing of node i

The issue of optimization of usage of underlying physical resources has been tackled (V. Gurbani, V. Hilt and I. Rimac) defining simple protocols for the allocation of resources. Virtualization and self-organization could be other meaningful ways to lessen the problem [VI, VII, VIII, IX].

A Taxonomy of services and application of P2P (Minerva 2008) is depicted in **Figure 25**.

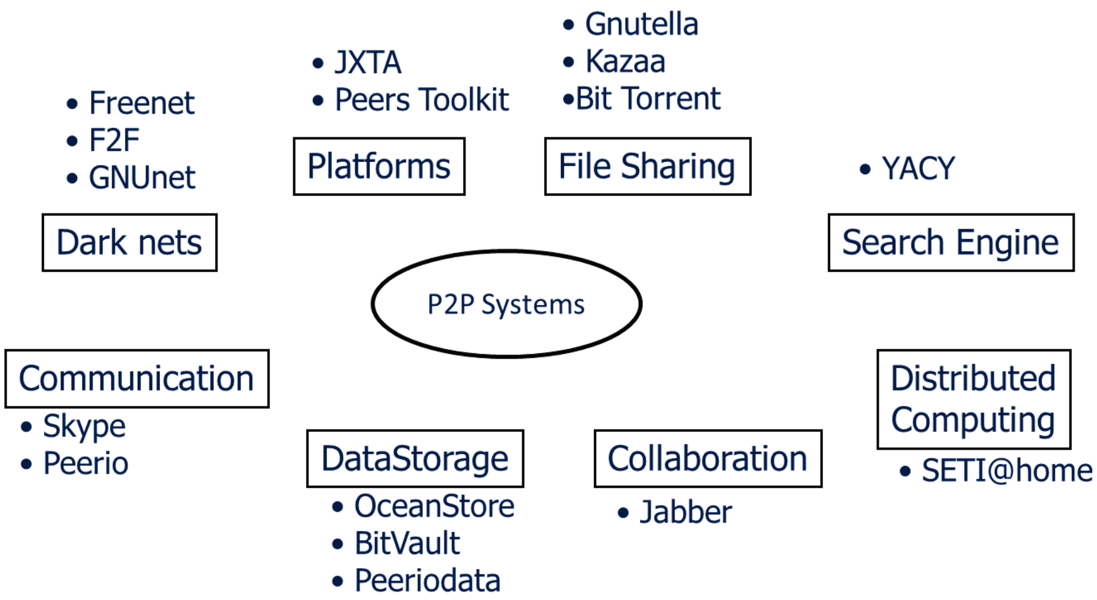
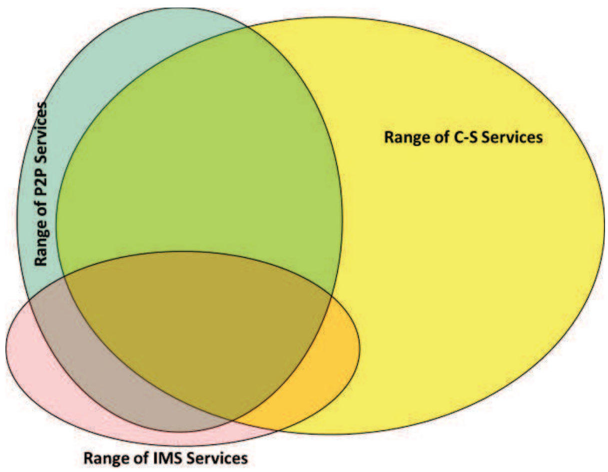


Figure 25: A Taxonomy of Current P2P Services

The current coverage of services with the three analyzed control paradigm is quite different as depicted in **Figure 26**. However, P2P seems to be usable for providing services in a widely distributed manner and its potential capabilities are similar to the C-S paradigm, especially if a set of standardized mechanisms could be generally implemented (e.g., distributed hash tables). The value of P2P paradigm is also related to its capability to deperimetrize services[4] and to disrupt the current status quo. Operators cold try to benefit to revert a critical situation in the realm of service offering.



[4] Services in a P2P system can be implemented wherever there is processing and storage power (provided by users), individual peers provide also connectivity that is usable independently from the Operator that is physically supporting it. Services can be deployed and provided by making use of resources that are not necessarily attached to the specific TelCo Network, actually they can be geographical distributed.

Figure 26: The Service Span of the Considered Control Paradigm and Architectures

2. 6. Findings

This chapter has introduced some of the approaches that are undertaken by major Actors in order to provide services. The distinctive standpoints of C – S, NI, and P2P approaches will be used in Chapter 8. In order to identify and discuss some possible scenarios that could occur in the future. It is very significant from a business and scientific perspective, the increasing interest that is growing around services. The definition of a "service science" comprising several application areas is important because it is paving the way to more precise mechanisms, methodologies and systems that can fulfill the service requirements. Under this perspective, it should be noted that the user interests and rights are still not fully considered: in fact many the "user-centered" approaches try to understand how to better offer service to user, and not in guaranteeing to users fair services that deliver value for the right price. This is a deficiency that will be emphasized in the entire document. Chapter 6. will deal with this important aspect of services: i.e., Business Models and the way they are used by Actors of the ICT industry to exploit and take advantage of customers and users. In addition in this chapter an essential concept related to services appeared: the more and more proactive role of users. This is a changing factor in the industry because users can play a more relevant role in delivery services and this can have deep social and economic impacts.

Chapter 8. will consider some of the service issues sketched in this chapter and will present scenarios that show how the Stakeholders (especially the TelCos) can leverage Services and Business Models in order to increase or consolidate their positions in the business or to change the "status quo" yielding to new opportunities. In addition the user based model could lead to disruptive scenarios in which traditional Stakeholders can be bypassed in favor of service and network infrastructures totally based on user provided resources. Interaction and control paradigms have been introduced; they are important because they affect the way services are implemented and provided. Choosing a control paradigm "a priori" without considering the service, the users and their needs can push towards cumbersome solutions that instead of focusing clearly on solving problems related to the service delivery are tackling the issue of how a platform supporting a specific paradigm can be used in order to deliver a specific service. This lack of flexibility in service design and implementation is a major drawback of many platforms that do impose very stringent control means on services and applications. As seen, Network Intelligence is especially prone to this problem because of the tight coupling of network and service functionalities. The C – S and the P2P paradigms are more flexible and can support a variety of possible interaction and control mechanisms reducing this dependency. The expressiveness of a paradigm is of paramount importance in order to create efficiently and effectively new services. Section 2. 5. has shown that the expressive power of a paradigm and its supporting infrastructure has a direct consequence on the possibility to create a large number of interesting services. This leads to the issue of service architectures. In this chapter, the basic service architecture approaches have been considered showing that the C – S one has currently and advantage, while the P2P is more complicated, but it could support in a very flexible and open way many services. The NI paradigm and related architectures seem the one that is less flexible and reusable for providing new services.In the following chapters, the issues related to these architectures will be further elaborated and in Chapter 8. some scenarios based on different architectural propositions will be discussed. Eventually in Chapter 9. a new flexible architecture based on findings of this study will be presented.

Chapter 3. Technological Evolution of the ICT Sector

Highlights

- Evolution of the Technologies of the telecoms ecosystem
- Virtualization of resources
- The importance of data and related technologies
- Distributed Processing Technologies and their applications
- Terminals

3. 1. Introduction

The goal of this chapter is to represent the current trends and developments characterizing the evolution of the ICT industry. The rationale behind this is to look at how the industry is progressing from the technological point of view. In fact the evolution of some basic technologies like processing storage, sensing and others will be briefly analysed in order to project it in a time frame of ten years. This should provide a vision of what will be possible at the technological level. In addition the evolution of some relevant software mainstream technologies will be considered aiming at evaluating how they will shape the future. This chapter should provide a sort of common ground on top of which to reason about how and why to use available technologies and solutions. On top of this, some initial considerations will be put forward in order to identify some inflection points (i.e., some points in time in which the technical evolution can enable new services and solutions that where out of reach up to a few years ago: e.g., pervasiveness of networks and devices. They are not singularity points, but these "inlfections" could identify changes in the linear evolution and can anticipate the singularities) that can enable new services and new technologies and some possible disruptions in this "mainstream" approach. Actually this chapter tries to analyse the major technologies forming the mainstream and to identify whether they could be superseded by better ones on the basis of the possible evolution of enabling technologies.

Under this respect the evolution of data centre, virtualization and cloud computing seem to represent the most likely evolution towards the creation and consolidation of large systems capable of serving millions of users by means of the C – S paradigm. The prevailing vision is the one that foresees the concentration of services in "hyper" clouds made out of millions of servers. The technologies supporting this concentration are consolidating today and services will be moved to the cloud.

This move of services and applications in the cloud comes with an manifest consequence: large quantities of data will be gathered, collected, processed within the cloud. In addition new data stemming from the integration of different data sets will be created in these massive data centers. Almost any action of users could be observed, screened and recorded. This Big Data approach could be pursued and supported in different ways in various parts of the world. Europe seems to adopt an approach respectful of the privacy of the users, but this attitude may have very heavy consequences on an industrial level (i.e., giving an advantage to the US industry). New approaches to identity management together to a user centred approach in data could be one of the inflections points.

Finally this chapter will briefly analyse the evolution of (mobile) terminals. The reason for this is that new terminals are very capable in terms of processing, storage, communications and sensing. Their power will grow also in the future making them an essential factor in the evolution of the services. Actually personal terminals could be seen as service platforms capable of delivery many important functions to the users.

3. 2. The ICT Ecosystem

The technological evolution of the ICT sector is confirming its pace of developments: the Moore Law (i.e., the number of transistors on integrated circuits doubles approximately every 18 months (Schaller 1997)), the Kryder Law (i.e., magnetic disk areal storage density increases at a pace much faster than the doubling transistor count occurring every 18 months in Moore's Law (Walter 2005)), the Nielsen's Law (stating that network connection speeds for high-end home users would increase 50% per year, or double every 21 months (Nielsen 1998)) and the Koomey Law (Koomey 2010) (i.e., the number of computations per joule of energy dissipated has been doubling approximately every 1.57 years) are holding valid. There is not an acceleration but a constant growth, nevertheless processing, storage and communications are reaching capabilities that can change the way in which ICT is used and perceived by users. Even if the ICT industry is not reaching singularity points, it is approaching situations in which the software technical solutions can substitute hardware based ones. These can be considered inflection points, i.e., a context in which a technical solution is viable thanks to the increased processing, storage and communications capabilities. These inflection points anticipate singularity points and will determine and influence the technological evolution of the ICT industry. Some main considerations are worth at this point:

1. General purpose hardware will be more and more capable (in the sense that it will provide more processing power combined with the possibility to store more information) and will progressively substitute specialized one.
2. The power consumption of hardware will be optimized and will bring benefits both to big data centers and to portable user equipment.
3. Software will be the distinctive factor of any device, service and system.

So far there are not laws governing the technological evolution of sensors and/or actuators, but it is likely that sensing and actuation capabilities will benefit from the general evolution of processing, storage and communication capabilities. Determining how much they will progress needs still to be determined. However also this technological branch will influence the evolution of ICT. Some forecasts discussed in (Zaslavsky, Perera and Georgakopoulos 2013) put in front the hypothesis of having up to 100 billion of sensors. This means that thousands of sensors could surround each human. It is not clear when this situation will be common

Keeping an outlook to 2020 and beyond, the forecasts depicted in **Table 3** can be considered (following the current pace of evolution) for some interesting technologies in the ICT sector:

Table 3: Technology Roadmaps

Technology Sector	Advancements and perspective to 2020 and beyond
Processing	• Capacity doubles every 18 months • Reduction in power consumption • Considerable cost reduction • Possibilities:

	o Multicore systems to the general public o Use of ARM based systems also in Data Centers • Development of non-VonNeumann architectures
Storage	• Capacity doubles every 18 months • Cost reduction (over 375.000 times over initial developments) • Possibilities: o Flash memories with 1 to 5 Terabytes o Hard Drives with tens of TBs o Integration of flash memory with wireless technologies o New technologies such as memristor (could help in decreasing the space needed for electronic components increasing their processing speed)
Fiber	• The progress in optoelectronics and the possible decrease in price due to production volumes will make fiber based solution convenient. This will be emphasized by the possibilities to: o Multiplication of the available Lamba channels o Increase the capacity of the single fiber (up to 100 Terabits/s) o New technological advances (e.g., attoseconds optics)
Radio	• Radio technologies are close to reach the theoretical limits of the spectrum, this implies • The need to reduce the radius of Cells (higher distribution of devices) • Reuse and allocation of new portion of spectrum (e.g., 700Mhz) • The introduction of new techniques for interference resolution (Cognitive Radio) • The usage of new technologies such as: lowPower and nanoimpulse systems or OAM (orbital angular momentum) to cram theoretically infinite transmission channels per given frequency (in lab 2.5 Tbs)
Local Networks	• Local connectivity on the increase especially for WiFi based solutions with speed that by the end of decade will exceed 1Gbs • Embedded communication in consumer electronics and household appliances and possible also in car and vehicles • Mesh networking and direct communication capable of supporting applications and services

Visualization	• The evolution of screen technologies (e.g., 8K screens) will need communication channels over 180 and up to 500 Mps • Screens of ebooks will have a definition comparable to paper and will be as thin and flexible • Micro projection will be available in terminals • Screen with embedded camera (e.g., Apple patent) • Wide deployment of embedded screens in the environments (e.g., tag price will be substituted by screens) • Semitransparent Screens for windshields and glasses • Gesture and natural interaction with screens
Sensors	• Billions of sensors deployed and active by the end of the decade • Sensors as native parts of products • Virtual Sensing (i.e., extraction of information from sensor data)
Terminals	• Terminals will benefit from the pace of technical evolution in terms of processing, storage and communication capabilities. • Terminals will be a service platform in the hands of users. • Always best connected strategies will be implemented by the terminal and not by the networks. • FON like networking will be used especially in denser populated areas • Open environments will compete with walled garden solutions • New interactive terminals will appear (goggles, watches and the like)

So the environment in which users will be embedded will be visual, highly interactive and responsive. ICT capabilities will be taken for granted and embedded communication will promote the idea of free connectivity to customers. On the other hand the progress of local area communications and direct communication in the terminals will enforce this capability.

Terminals will be a differentiated factor in the provision of services and they will drive the developments in terms of services and applications. In this sense, a competition between closed and open environments will going on. With heavy consequences on the share of the application market. However any winning technology will be related to terminals in a way or another (because it will run on terminals or because it will enable new services for terminals).

In the rest of this chapter more details on specific aspects of the technological evolutions will be briefly discussed.

3. 3. Data Center Evolution

Data Centers are one of the battlefield for the future of communication market. Actually they are for many global providers one of the two horns for the market strategy. The application market is dominated by the Client - Server paradigm, so a strong presence on the server side or having a prominent market share on it is fundamental. For this reason, many web companies don't consider Data Centers as cost center. They are instead considered Innovation points. In fact companies like Google put a substantial effort and investments on the development of such infrastructure (Platform), (Barroso, Dean and Holzle 2003), (Dean and Barroso 2013), (Minerva and Demaria 2006). The technological evolution of data centers is dominated by some issues that owners try to solve:

- Power consumption and heat production;
- Replication, distribution and impacts of communication costs;
- Increase and use of processing and storage;
- Programmability and leverage of highly distributed systems (Ghemawat, Gobioff e Leung 2003), (Pike, et al. 2005), (Burrows 2006), (Chang, et al. 2008).

These issues are intertwined and determine many interesting approaches carried out by the data center owners: for instance the need to reduce power consumptions and reduce space allocation is pushing towards the usage of ARM based architectures (see for instance Dell and HP related announcements[5,6]). In addition, the need of controlling how connectivity between different data center is used has pushed towards the usage of novel communications solutions that move to communications infrastructure some concepts already applied with success in the processing and storage: i.e., virtualization and use of low cost general purpose hardware (Open Networking Foundation 2012). It is not a case that Google is a pioneer of the Software Defined Networking with his own solution (Vahdat 2012), Amazon is following this route as well and its research and experiments in this area are consolidating[7] (Miller, Brandwine and

---

[5] http://slashdot.org/topic/datacenter/dell-calxeda-develop-second-arm-server/ last accessed on April 9th, 2013.

[6] http://www.crn.com/news/data-center/231902061/hp-intros-low-energy-data-center-platform-based-on-arm-servers.htm?itc=xbodyjk last accessed on April 9th, 2013.

[7] http://billstarnaud.blogspot.it/2012/02/amazons-simple-workflow-and-software.html last accessed April 9th 2013.

Doane 2011). Facebook is active in this area (see its participation in Open Networking Foundation[8]) and it is experimenting on a small scale system[9] the benefits of SDN.

The strive for reaching new levels of distribution in processing and storage capabilities have led to interesting techniques in data management (from Google's Big Table (Chang, et al. 2008) to NoSQL databases (Leavitt 2010), (Stonebraker 2010) up to Amazon's Dynamo (DeCandia, et al. 2007), and the like), in parallel computing (e.g., Google's Map-Reduce (Dean and Ghemawat 2010), Hadoop (White 2010), and others).

Facebook approach is more conservative and it aims at the definition and possibly standardization of a viable general purpose based data center infrastructure, the Open Compute Project (Hsu and Mulay 2011), to be defined with the support of the community and other web companies.

In trying to solve their problems, the Web Companies have taken the approach to ensure an answer to the "clients" in the shortest possible time (e.g., search queries). In terms of the Consistency, Availability and Partition Tolerance, CAP, theorem (Brewer 2000), (Brewer 2012) these companies (e.g., Google and Amazon) have chosen to enforce distribution of functions and availability instead that guaranteeing consistency of data. This is a quite new approach that can be acceptable by final customers of search engines, and social networks. This approach has shaped the architectures of the infrastructure of Google, Amazon and Facebook. Even Twitter's newer architecture is adopting similar approaches[10] (Krishnamurthy, Gill and Arlitt 2008), (Mathioudakis and Koudas 2010). Actually Twitter is a very interesting architecture because it is pushing for a quasi-real time approach in retrieving data and derive information by them to be exposed to end customers.

For a general discussion about the Google architecture see (Barroso, Dean and Holzle 2003), (Ghemawat, Gobioff and Leung 2003), and (Minerva and Demaria 2006).

Summarizing the data center approach can be characterized by:

- The adoption of disruptive solutions in their design to achieve efficiency, reliability and scalability, by means of modularity and "self-similarity";

[8] See Open Networking Foundation at www.opennetworking.org last accessed 30th May 2013

[9] http://gigaom.com/2013/01/31/facebook-experiments-with-small-scale-software-defined-networking/ last accessed April 9th, 2013.

[10] See for instance http://engineering.twitter.com/2010/07/cassandra-at-twitter-today.html for the usage of a nosql solution in Twitter. Last accessed April 9th, 2013.

- The adoption of a Do It Yourself, DIY, attitude, almost the majority of big Data Center owners have built their own infrastructures (including buildings) using cutting edge technologies (from cooling to power distribution, from servers' architecture to networking protocols) and positioning their infrastructure as a means to cope with and differentiate form competitors;
- Leveraging any progress in processing and storage technology (e.g., ARM-based clusters);
- The mastering of software. Actually the web companies have mastered their own platforms, starting from general purpose hardware and introducing innovation in networking, software, and communications by also adopting open source software. They are capable of developing their own software solutions and bring them to the market for large scale use;
- The development and adoption of new programming paradigms (e.g., MapReduce, Pregel (Malewicz, et al. 2010), etc.).

Data centers can be seen as sets of "centralized" servers, but actually behind a client-server based front end, they are the epitome of distributed computing. In fact many technologies have been developed, tested and tuned up into data centers (e.g., big table, chubby, dynamo and other distributed solutions). So they have a sort of double face: a traditional one in terms of servers receiving and treating requests form a multitude of clients, and a distributed one in which parallel computation, distribution of data and the newest distributed computing technologies are invented, tried and nurtured (see **Figure 27**). The Data Centers are factories of innovation (when a Do It Yourself, DIY, approach is used). For this reason adopting the classical Telecom Operator approach to Data Center (the one of intelligent buyer) is not a winning one, in fact the most of innovation is created within DIY Data center and is usable after some time in commercial solutions for data center. For this reason the technological gap between Web Companies and Telecom Providers in this sector will increase over time. TelCos will be lower cost provider of regional solutions because their solutions will be the one offered by the market and the only competitive asset for differentiating will be the price.

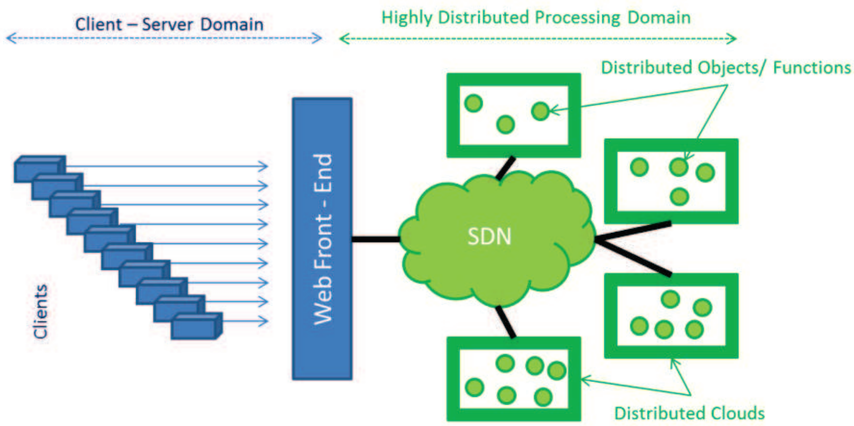


Figure 27: The Double Face of Data Centers

On the other side the Data center winning players are moving toward terminals. Each one of them is addressing this market and this technological area. The attempt is clear: to dominate the two aggregation points of intelligence in the Internet: the client and the server side. Doing this the Web companies are besieging Telecom Operators reducing them to a mere commodity and accruing the problems of investment on the communication platform.

3. 4. Resource Virtualization

A lot of effort has been spent and devoted to the possibility to virtualize systems and applications on a general purpose infrastructure. The increasing processing power and storage capabilities help in in adopting more and more virtualization technologies within data centers and in distributed systems.

In a quest for flexibility in resource allocation, virtualization techniques are playing an important role. Virtualization "means to create a virtual version of a device or resource, such as a server, storage device, network or even an operating system where the framework divides the resource into one or more execution environments..[11]". Virtualization can be applied at different levels, e.g., hardware, operating system and application/service. **Figure 28** represents different perspectives on virtualization.

[11] For a definition of Virtualization see http://www.webopedia.com/TERM/V/virtualization.html

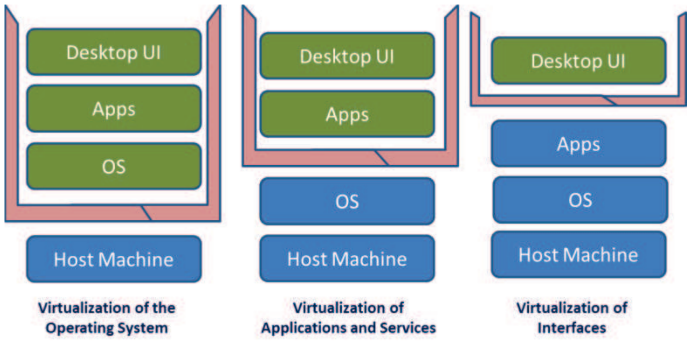


Figure 28: Examples of Virtualization at Different Layers

Generally speaking, Virtualization allows to:

- Expand hardware capabilities, allowing each single machine to do more simultaneous work;
- control costs and to simplify management through consolidation of servers;
- control large multiprocessor and cluster installations, for example in server farms;
- improve security, reliability, and device independence by using hypervisor architectures;
- to run complex, OS-dependent applications in different hardware or OS environments.

These possibilities are quite important and are enabling a new kind of offering: the XaaS (everything as a service) paradigm is heavily based on the virtualization techniques. In the Data Center area, virtualization is used to support a whole new approach for the organization of software: every software application can be virtualized in the network and being offered to customers as an on-demand service. New businesses have emerged and also some problems. Some of these issues are strongly related to connectivity: for instance what happens if the Data Center supporting business critical applications is perfectly running, but there is not connectivity to it? Virtualization, connectivity and Data center structure are seen by the customer as a whole, but these features can be sold and operated independently from each other. Another issue is related to the software instances: with traditional software, the customer can use the software even if the producer of it has gone bankrupt, or has not updated the application for a while. With XaaS, if the Provider of the application is not properly running the environment, the customer may not have access to the critical functionalities. This problem could be even exacerbated if the Providers goes out of business abruptly without giving the time to the customer to migrate to other platforms. One solution could be to run virtualized instances of the application within the domain of the customer.

Another issue to be considered in Virtualization is that even if the hardware progress is constant and new capable hardware platforms will succeed during time, the virtualization is another additional layer that will consume processing and storage capabilities. Virtualization has to consider the May's Law[12] and to compensate for the inefficiencies introduced by software.

In spite of this, in many ICT areas virtualization is already being successfully used and it comes with a lot of disruption as well as new opportunities for creating services and application. In the Network Control area, for instance, the Network Virtualization Function initiative, sponsored by ETSI (NFV 2012) is using virtualization for promoting higher levels of functional aggregation and organization of the network functions.

3. 5. Distributed Applications and Technologies

Currently distributed processing is seen as strongly related to cloud computing (Foster, Zhao, et al. 2008), (Di Costanzo, De Assuncao and Buyya 2009), this is due for two reasons: for the possibility to create within a single cloud a large distributed system (e.g., Google, Amazon, Twitter) and for the possibility to create large distributed applications on the intercloud (Buyya, Ranjan and Calheiros 2010). However Cloud computing is deployed and used quite differently by several actors. For instance some prominent web companies use these technologies in order to leverage the infrastructure needed to carry out their principal business (e.g., market place for Amazon and search/advertisement for Google). Their solution of cloud have stemmed as a means to leverage the idle time of their powerful data centers. Many Operators have seen a potential revenue generator in this kind of offering and they have put in place cloud computing solutions. The major difference between the Operators approach and the web companies one is related to a recurring concept in the deployment of ICT solutions: deperimeterization of services[13]. In a few words, the web companies (especially the big ones) have a global footprint, i.e., their services are well-known worldwide and customers that use the Internet can typically make use of their offering. Simply put, the Web companies can leverage a long tail effect on the global market, i.e., services provided can be globally

[12] Software efficiency halves every 18 months, compensating Moore's Law [http://en.wikipedia.org/wiki/David_May_%28computer_scientist%29]

[13] i.e., the capability to provide services that are globally accessible and independent from a specific network infrastructure. Typically the services offered by a web company like www.WebCompany.com are accessible to all the users of the web, while services provided by a TelCo are available only to the subscribers of the specific service on the specific network.

accessed by users independently from the geographical location of the users. This allows the WebCos to make (small profits in many locations and to reduce the idle time of their infrastructure. Deperimeterization in this c)ase means that world-wide users can access the web company cloud offering irrespectively of the location of the user. The Telecoms proposition is somehow different: even if their offering could be somehow deperimeterized, these companies have not a global footprint (even the biggest ones are Regional – China Mobile – or international). As a matter of fact their business is bounded to be regional and is confined to locations where the specific TelCo has a network infrastructure deployed. As discussed in [XXIII] in this area the approach of TelCo has to be totally different and even more technical challenging compared to Web Companies. TelCos have to emphasize their regionalism trying to work out an offering (coupled with a technology platform) that is customizable for the specific regional needs of the customer, in addition, the solution should be supporting some levels of interoperability with other regional solution possibly not owned by the same TelCo. In this context the concept of Federation of Clouds is of the paramount importance. This means to provide a regional solution that is adhering to laws and practices of that specific location (together to the management of the platform) but open to integration with other clouds of resources. Federation is a difficult technical topics that comprises not only IT and software issues, but also communications. In fact regional federated clouds need to have a good level of interconnection in order to allow applications and data to flow. In other terms TelCos have to try to overcome their regionalism by federating different clouds possibly pertaining to different provides. Web companies cannot count on the local presence, but they can provide "normalized" and cheap solutions (costs are keep down because of the big volumes) that are accessible globally. More considerations and a proposal for a taxonomy of cloud solution can be find in the already mentioned paper [XXIII].

However, distributed computing is not and will not be confined to cloud computing. In fact, there are a number of different technologies under this technological area. One first observation is that generally speaking, Cloud Computing is disregarding the issues related to communication. It is suggested to have good level of Service Level Agreements in order to cope with issues and problems in connectivity. This is not correct and actually even the first specifications of Grid were trying to cope with network organization for adequately support the processing and storage capabilities (Baroncelli, et al. 2005). This approach has led to specifications related to the possibility to control by means of managers the connection establishment, update and tear down. His trend has then been adopted by the OpenFlow (McKeown 2008) initiative and has led to a new revamp of the concept of programmable networks: the software defined networks (Lantz, Heller and McKeown 2010), (Gude, et al. 2008). Similar considerations have driven Web Companies to use these technologies for the optimization of their connectivity capabilities, for instance the G-Scale solution developed by Google (Vahdat 2012).

However the evolution of highly distributed systems is approaching a new cornerstone: i.e., the integration of social driven connectivity with small range connectivity can bring to the massive adoption of new mechanisms. In fact, the proliferation of capable terminals[14] will permit to exploit the real capabilities of distributed processing. Terminals will be nodes of a distributed environments and they will be able to determine how to connect each other, what functions to execute locally and how and where to store data. P2P technologies are a clear demonstration of the capability to create huge infrastructures capable of supporting different classes of services. As discussed in [XVIII], the peer to peer technology (compared to network intelligence and client-server) is the one that can support many computation paradigms (event based, tuple space, or other) so that different applications classes can be supported. This flexibility requires the introduction of some functions (with respect to the easiest paradigm: the client – server) in order to create and support the overlay network functions needed to nodes for correctly operating and benefit of the distributed infrastructure. This paradigm has been proposed for the creation of global service oriented platforms. For example the Nanodatacenter[15] project (Laoutaris, Rodriguez and Massoulie 2008), (Valancius, et al. 2009) was defined in order to use P2P and BitTorrent (Izal, et al. 2004) for the distribution of data in a distributed network formed by access gateways and access points. In this case an Operator could try to exploit processing, storage and communications capabilities of those types of devices already deployed in the customers' homes.

These edge distributed systems will be characterized by the large number of devices connected and by the high dynamicity of the availability of the single node. Actually nodes can come and go in the infrastructure bringing in and out their processing, storage, communications and sensing capabilities. These systems will not be managed with traditional

[14] i.e., terminals that have a considerable processing power and storage as well as the ability to use different of connectivity means (e.g., 3 or 4 G, WiFi, Bluetooth).

[15] See http://www.nanodatacenters.eu last accessed May 30th 2013

mechanisms. There is the stringent need to have self-organizing capabilities that have to entirely substitute and make it useless the human intervention. The dynamic of entering in these network will be so short that no human configuration of the terminals could be fast enough to allow the integration of the nodes. Autonomic capabilities are a need, they are essential to ensure the existence of these type of infrastructures. Due to the spontaneous nature of these aggregations of resources, there is the need to understand how to promote the resources sharing and avoid as much as possible opportunistic behaviors. Studies in social networking, e.g. (Ohtsuki, et al. 2006), show that when the benefits over the costs of sharing exceed a parameter k (representing the number of neighbors, or links to them), i.e., B/C > k then an altruistic behavior emerges in communities. In order to reach this situation and maintain a value for the major part of participating nodes some strategies could be implemented based on game theory in order to increase the global value for all and to optimize the benefits for the single nodes. It is clear that these aggregation of nodes will act and behave as complex systems and programming their behavior will require a lot of innovation and research.

When the technology will reach this point a sort of very powerful and highly pervasive platform will be available. In this situation, applications will be organized as dynamic "coalitions" of cooperating service components with these features:

- deployed on dynamic and pervasive "cloud" of computing resources (clusters of servers, users' devices, sensors, etc.)
- provided by multiple actors (users, service providers, enterprises, equipment providers, sensor network providers, etc.)
- "ad hoc" assembled and adapted (according to situated needs, component and resource availability, and their changes)

These coalitions will be the framework for building smart environments that enable new business opportunities beyond traditional value chains. A possible effect is that this new organization can contrast "oligopolies" in the service delivery market. These innovation came from a bottom up approach and leverages the end user capabilities. For this reason, a similar trend is difficult to spot and possibly it could emerge as a " change of phase" , i.e., when the number of terminals and communication capable devices will reach a critical level this new way of interacting will suddenly emerge bringing disruption to the communication status quo.

In addition cognitive and software defined radio technologies could give a boost to this approach making it a sort of pervasive collaborative platform.

3. 6. Harvesting Data: Big Data and Personal Data

Technology advancements in several sectors (devices, sensors, cloud computing, pervasive connectivity, online services, process automation, etc.) extend the possibility to collect, process and analyze data. So far this information is scattered and mirrored in different silos pertaining to different companies. These data however provide hints and clues about the behavior of users and organization (and things associated to humans) with respect to:

- Activities carried out in the physical world and the virtual worlds
- the activities performed inside the organizations
- the "digital footprint" of people, i.e., the collection of data and logs that described the interaction of users with digital services.

By their nature, "Big Data" are characterized by high Volumes, Velocity and Variety, requiring scalability for managing and analyzing them. Aiming at supporting the 3V properties, many technologies have been developed such as:

- NoSQL Data Bases (Leavitt 2010), (Cattell 2011);
- Map-Reduce processing framework (Dean and Ghemawat 2008), (Dean and Ghemawat 2010);
- Real-time data stream processing (Stonebraker, Cetintemel and Zdonik 2005) (e.g., Twitter (Bifet and Frank 2010));
- Semantic representations (Auer, et al. 2007);
- Improved Data Mining for Analytics (Piatetsky-Shapiro 2007);
- Elastic resource platforms (Marshall, Keahey and Freeman 2010).

The current trend for many companies is being a hub for big data collection in order to exploit the information gathered while conducting business with customers. For instance TelCos are eager to collect and possibly exploit the information collected about the users (surfing habits, connectivity related data and the like). Collecting personal information and profiling people is a promising business. This aggregation role enables new business opportunities because of the possibility to analyze data patterns in order to derive new information, the possibility to mash-up different sources of data and information and the possibility to distribute this information to interested customers. Actually Data availability and mash-up capabilities enable an ecosystem of applications bringing benefits to people, public and private organizations. Many future applications will leverage these features in order to provide meaningful services to people and organizations. On a more scientific side, the data collection enables the data-intensive Scientific Discovery (as proposed in the Fourth Paradigm (Hey, et al. 2009), (A. J. Hey 2011)). Many new scientific discoveries will emerge by the capability to access to a large database of differentiated data. For instance collateral effects of combined therapies and mix of medicines are hard to determine in laboratories. Accessing health records of people being cured with these combinations could result in determining if there are counter-indication for certain pathologies of groups of patients. Data analysis and cross-checking in fact gives new info and new models. For instance in (González, Hidalgo and Barabasi 2008), some scientists were able to cross-check the mobility patterns of people and determining their habits.

Operators have plenty of useful data: they range from the data call record that logs all the calls from and to a user, to location information about a user, to surfing habits and web sites visited of users. These data sets could be used in order to determine and improve the understanding of social interactions or the goods flows within a city. However these data have to be dealt with in respect to the ownership of them and the privacy of users. A user centric based ownership model should be enforced by law in order to let the user decide how to leverage their own data.

The World Economic Forum (World Economic Forum 2011) sees the area of personal data as the new oil for the digital market. However, user ownership should be enforced and guaranteed. In this context an approach as the one promoted in [XX] seems to be promising in terms of potential business exploitation, but still very respectful of the user ownership. Simply put, data should be collected and stored into banks of data. Users will then decide if they just want to store them for personal advantage (retrieving and controlling the interaction in the web, retrieve a specific transaction or a document and the like) or to exploit their value in a controlled market. So users can establish policies for using personal data and open up accordingly the access to third parties.

Another trend will emerge: a sort of reverse engineering, i.e., even if data are anonymized, tracking capabilities can be put in place in order to determine who the producer of that data was or whose the data was referring to. Privacy and security will be an integral part of the offering of personal data stores.

3. 7. Disruptions in Identity Management

Digital Identity is another issue that has the potential to disrupt the current technical and economical evolution of the Internet. As stated by (Clark, et al. 2005), the identity management is a tussle more of a technological problems. The companies that can manage the identities of users are in the position to infer and derive the actions and the behaviors of users in their digital interaction within the cyberspace (Landau and Moore 2011). In fact identity management is strongly related to profiling of users, the link between them is an enabler for relating actions and behaviors to the individual. The issue of identity is also related to the rights of the citizen and/or customer: the Internet is global by nature and data "flow" where it is easier to deal with them. There is a great difference in the law enforcement between Europe and USA in privacy issues. In Europe privacy is considered an important right of the Citizen, in USA privacy is seen as a right of the Customer. The consequences of these two divergent approaches are very profound: personal data and identity management in Europe are so regulated that many services (e.g., the study of fluxes in a smart city) are not possible unless all the involved parties give the consensus to the treatment of data (e.g., the call data record can be used for a service if and only if all the involved parties, i.e., the caller and the callee) agree to its usage. In USA it is enough to give a general disclaim in the service offering and personal data can be used. As a consequence, the American Companies are in a more favorable position to offer these kind of services over the European ones. In addition the collection and aggregation of data is performed out of Europe in order to avoid this more stringent regulation. The consequence on competition is evident: US companies can do more that European ones. There is the need to enforce the regulation on all the data of European citizens, but that is very tricky and difficult.

There is a basic observation that should drive the discussion about identity. Who is the owner of the personal identity? In [XVI], this issue has been tackled with a disruptive approach in mind. If the owner of the identity (and the related identifiers) is the user, than a great change could take place in the communications and internet sectors. Just to name two examples: if the phone number is property of the user, then roaming, number portability should be radically change under the concept that the user identity can be dynamically associated for a small period of time to another network or to a communication environment. Same thing for instance in the Internet, if the email address is not owned by the Provider but by the user, then e-mail addresses should be portable from a system to another. This could change the long established market of email. Identity is also intertwined with security. Users should be able to decide how their identity should be dealt with, in www.identitywoman.net a spectrum of identity modes is given (see **Figure 29**)[16].

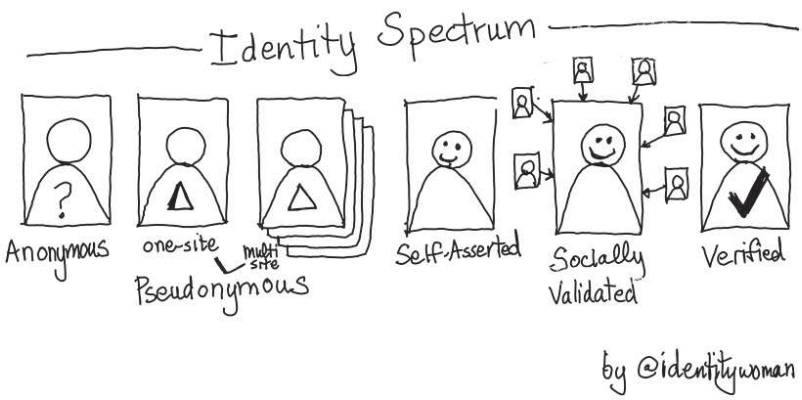


Figure 29: The Different Modes that Constitute the Identity Spectrum

User should have the right to be anonymous without any Provider trying to impose registration to site or identification of the person; other times the user should be let free to use pseudonymous because there is more than a digital identity associated to an individual; the user should also be capable of self-asserting its identity by means of statement without any authority trying to validate those or requesting proof of statements; a user could also identified by its friends and then being socially validated; the usual case of validation by a third party should by the exception rather than the norm (as it is today) and this form of identification should be used in limited circumstances (e.g., when using home banking applications). This identity spectrum shows how users are spoiled of their rights and how different identity management in the web is from the real world.

This means that the service provider should ensure to deal with personal identity in one of the chosen mode. However for security reasons, these modes (e.g., anonymous) should be supersede if an authority (e.g., the police under a mandate from the Justice office) requires so. There is a need for a strong regulation of the Digital Identity in order to guarantee the fundamental rights of the citizens, but also to allow tracking of users under specific and important (from the perspective of law enforcement) conditions. The need for an Identity Layer has been advocated several times in the Internet community (e.g., (Cameron 2005)), but there is not an agreement in the industry because of the commercial relevance of the issue. In

[16] See http://www.identitywoman.net/the-identity-spectrum

principle the identity should be user centric, i.e., the user decides session by session the identity mode and associates his identity to a specific Identity Provider (in certain case it could be the user himself). This layer could be natively integrated in the network in such a way that access to a network will be granted to the individual based on the preferred mode, but with the possibility to strongly identify the user under the request of authorities. The interesting point is that, as happens in most of the social relations in the real world, the user is often his own Identity Provider by means of assertion about himself. These assertion could be socially validated by friends or people related to the individual. In this way part of the value of the ownership of Identity is given back to individuals.

Similar approaches are undertaken by Dark Networks or very specialized P2P networks like i2p (Herrmann and Grothoff 2011) that try to protect and hide the users interactions from control of other parties. Identity management could be dealt with in a similar way as information are handled in the i2p network, i.e., through onion routing, and nested encryption by means of a set of Private/Public Keys. **Figure 30** represents a possible chain of encrypted Identifiers/token that can be decrypted by a specific Identity Provider and passed through to a more secure and low lever Identity Provider and Enforcer. The last one could be a national authority.

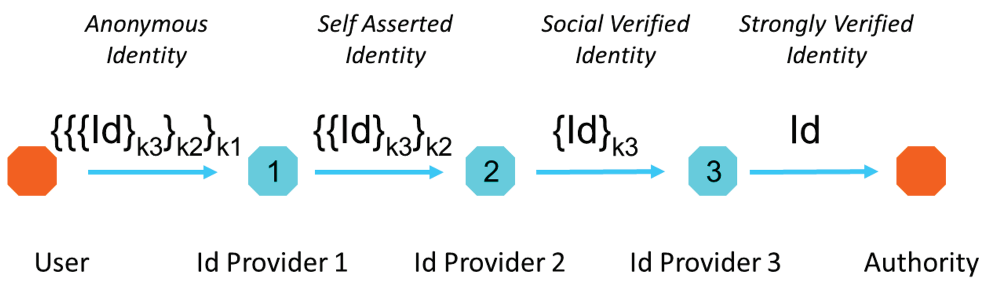


Figure 30: A Chain of Identity Keys

At each step, the Identity provider can encrypt only the relevant information at his level, there is noneed to disclose more information (that in fact is kept encrypted and is passed to the next Provider, and so forth).

3. 8. Cloud Computing and Beyond

The term Cloud Computing is nowadays a common jargon for users and providers. It refers to a computing environment that provides computational services typically in a client - server fashion. Services can range from on demand infrastructural capabilities like storage or computing capabilities to applications and services like Customer Relationship Management applications.

"The National Institute of Standards and Technology (NIST) defines cloud computing as a "*pay-per-use model for enabling available, convenient and on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction*" (Mell and Grance 2011).

Figure 31 depicts a typical configuration of a Cloud Computing infrastructure.

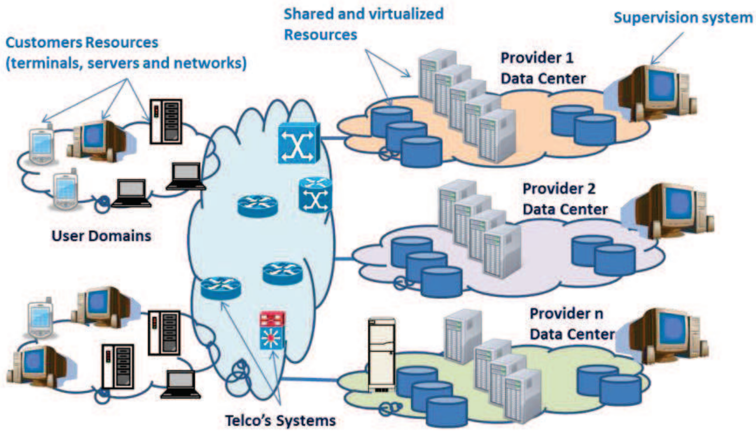


Figure 31: A Typical Cloud Computing Configuration.

In this representation, user systems can access to "clouds of resources/services/applications" by means of the Internet. Each cloud can provide capabilities on demand (users buy resources, platforms and services just for the time they need them). Relationships between the different clouds can vary according to business relationships among the providers of the infrastructures.

According to NIST (Mell and Grance 2011), a Cloud Computing system is characterized by a set of essential characteristics, such as:

- On-demand self-service, i.e., the capability offered to a user to directly manage all the needed infrastructure.
- Broad network access, i.e., the ability to access to Cloud Services by means of common (Internet based) mechanisms independently from the underlying networks (fixed, mobile) and compatibly with the most common devices (PC, Mobile phones, tablet and the like).

- Resource pooling, i.e., the Providers can dynamically integrate needed resources in order to satisfy customers' needs. Examples of resources are storage, processing, memory, and network bandwidth.
- Rapid elasticity, i.e., the capability to flexibly allocate the needed resources according to availability and customer's demand.
- Measured service, i.e., the Providers should make available to customers a precise accounting of resources allocated and used.

The features and capabilities of a Cloud system can be summarized into a well renowned model (for instance in (Vaquero, et al. 2009), (Lenk, et al. 2009)) that foresees three majors Service Models:

- Software as a Service (SaaS), i.e., services and applications are delivered to users by means of a web browsers and /or specific client applications.
- Platform as a Service (PaaS), i.e., all the typical functionalities of a software platform (e.g., libraries, tools, services) are provided to the users by means of a browser or a client application.
- Infrastructure as a Service (IaaS), i.e., basic capabilities, like processing, storage, and connectivity, are provided to the user that can configure them (e.g., through a web browser of client applications) in order to deploy and execute his/her own services and applications.

From an ecosystem point of view, the NIST definition implies a very simple business model: the pay per use one. It could be implemented by obvious Web Companies (like Google and Amazon), by relevant IT Companies and by Telecom Operators (TelCos).

From a deployment perspective, the NIST definition includes four options:

- Private cloud. A full infrastructure (comprising management capabilities) is offered to a single organization.
- Community cloud. The infrastructure is offered and provisioned for exclusive use by a specific community of consumers.
- Public cloud. The cloud infrastructure is offered and provisioned for open use by the general public. This refers mainly to SMEs and residential (but not only) customers.

- Hybrid cloud. The cloud infrastructure is an integration of different cloud infrastructures that remain separated, but are capable of interoperating by means appropriated technology and business goals.

Figure 32 is derived from the taxonomy as defined by NIST.

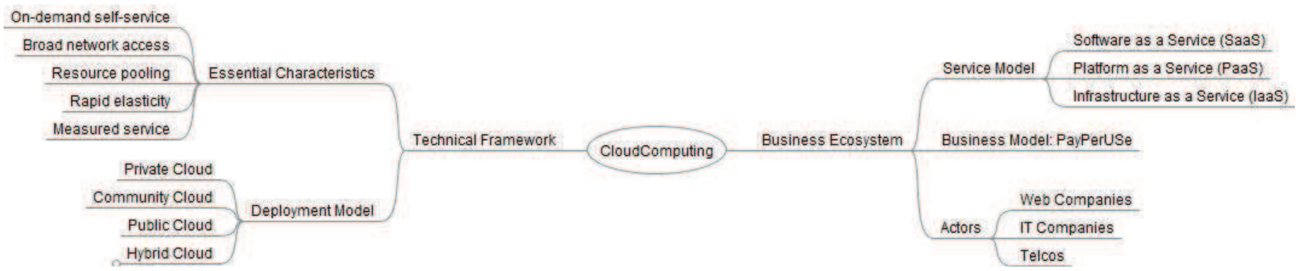


Figure 32: A Cloud Computing Taxonomy that Enriches the NIST Definition

This chapter provides a wider view at the technological and business level of this cloud computing definition aiming at correctly positioning the TelCos proposition in the market and in the technological scenario.

3. 8. 1. A Market Driven Evolution of Cloud Computing

The recent evolution of cloud computing (actually cloud computing is a derivative of old and well-known ideas related to utility computing (Vouk 2008)) has been heavily influenced and led from the innovation of web service platforms as provided by big companies like Google, Amazon and others. These companies have been instrumental in the technological transformation of Application Servers in very complex (and highly distributed) data centers. They had to put in place high capable and available data centers able to provide services (e.g., search or selling of goods) to a large audience and with a high variability in demand. Their infrastructure was dimensioned in such ways to be able to provide an answer to each worldwide customer in less than a second. Their infrastructures count for hundreds of thousands of general purpose computing machines (Greenberg, et al. 2008).

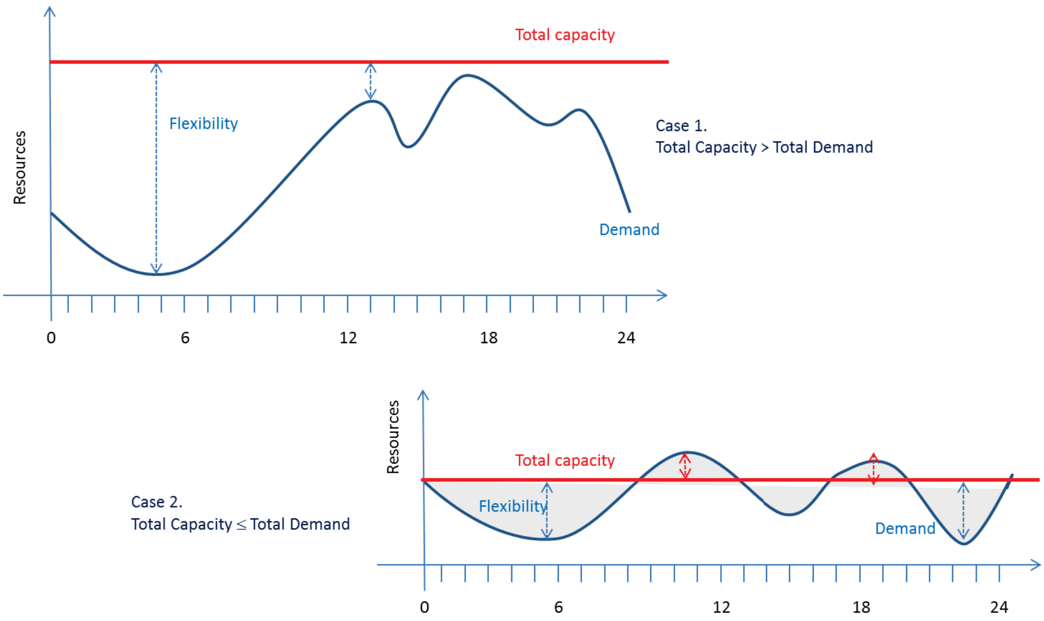


Figure 33: Flexibility in Capacity Allocation in Data Centers

The opportunistic approach of these giants is based on the fact that they deployed an enormous capacity that is seldom totally used for providing in-house services. There is a lot of spare capacity that they can reuse or can offer to clients. An example is Gmail, the e-mail service offered by Google. In order to index and organize information, the Californian companies had developed over the years a gigantic infrastructure that is capable of storing a large part of the known web. They have spare capacity and they can use it flexibly in order to provide to user large repositories for collecting mails. The variance between the deployed capabilities and the real usage of them is the key for providing cloud computing services. **Figure 33** depicts two typical situations, the first one in which the total capacity is always greater than the demand for resources (in this case the business goal is to sell the spare capacity, i.e., shown as flexibility in Case 1); the second one depicts a situation in which sometimes all the resources are over-allocated, a sort of negative flexibility that can hamper the functioning of the system (in this case a better allocation strategy is to be implemented, e.g., able to optimize SLAs and to reduce penalties).

3. 8. 2. Some Technology Trends in Cloud Computing

The technological panorama of Cloud Computing is vast, in this section a few aspects of its evolution will be taken into consideration. They are relevant for understanding how the current market propositions are put forwards by major actors and their implications from a service ecosystem point of view.

Emphasis on Data Management

The current offering of services and applications in cloud computing is derived largely by the capability of some companies in dealing with huge data sets. Two of the major actors in this field, namely Google and Amazon, were pioneers of new ways for dealing with large datasets and the use of advanced techniques. Google has been using the MapReduce approach (Dean and Ghemawat 2008) in order to implement an indexing mechanism able to perform on a large infrastructure of general purpose machines. The MapReduce[17] method consists of a Map() procedure that filters and sorts a sequence of data and a Reduce() procedure that executes a combination of available results of Map operations. There are two steps in this method:

"Map" step: The master node takes the input, splits it into smaller sub-problems, and allocates them to worker nodes. The worker node executes the smaller problem, and returns the answer to the master node.

"Reduce" step: The master node then gathers the answers to all the sub-problems and combines them to form the result.

In this way, large set of data are reduced into smaller chunks and each chunk is dealt with in parallel by a worker. Intermediate results are sorted out and combined by reduce processes in an ordered sequence (as depicted in **Figure 34**).

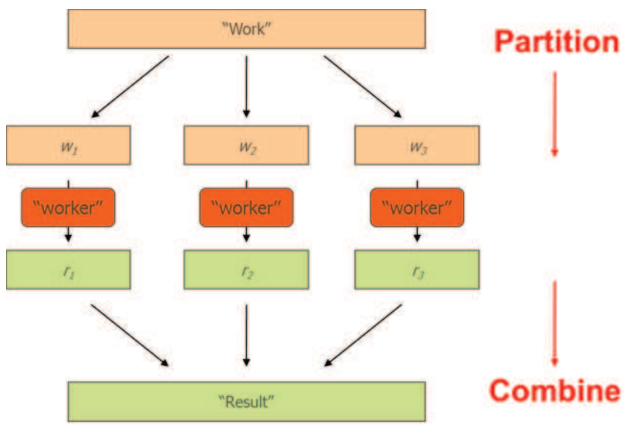


Figure 34: A MapReduce Example

[17] See for instance http://en.wikipedia.org/wiki/MapReduce last accessed May 30th 2013

The Google approach was inspirational and this has led to the well know approach of the Hadoop Open Source platform (White 2010). MapReduce and Hadoop mechanisms are often offered as a service in cloud computing platforms (e.g., Amazon Elastic MapReduce Service (Gunarathne, et al. 2010)).

Another example of data processing innovation is the Amazon platform. It is intended to support the large variety of services offered by Amazon (see **Figure 35**).

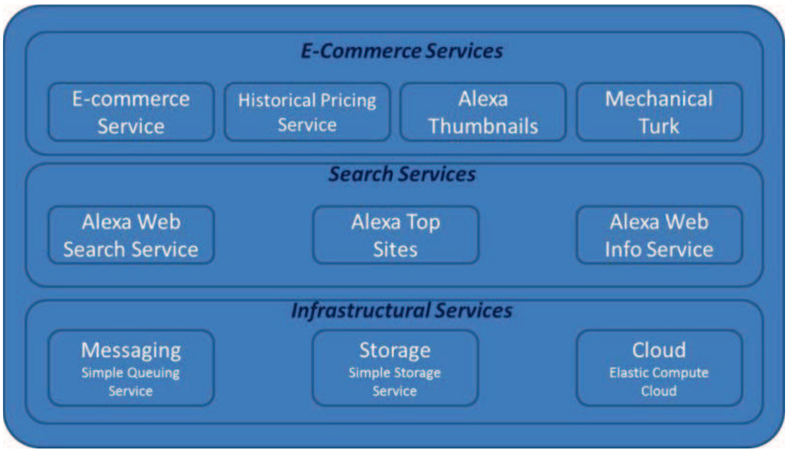


Figure 35: The Amazon Cloud Infrastructure

Particularly interesting in the realm of data manipulation is the Amazon solution for the Simple Storage Service (S3). Dynamo (DeCandia, et al. 2007) is a highly available, proprietary distributed storage system. Its goal is to provide database services in a highly distributed environment. In order to reach the goal it is based on a key-value approach and it uses distributed hash tables (DHTs) for pointing to data. Functions similar to Dynamo are offered through SimpleDB web service (Sciore 2007) by Amazon (Murty 2009) which also offers elastic basic storage services through S3 (Brantner, et al. 2008). Dynamo is one example of a consistent trend in database evolution named noSQL, it is not following a traditional relational database management system approach, instead it supports the high distribution and partition of huge datasets by means of a distributed, fault-tolerant architecture. Other example of this trend are the Facebook's internal platform Cassandra and the already cited MapReduce and Hadoop systems. Actually the design choices for Dynamo are: scalability in order to add new systems to the network minimizing their impact; symmetry, i.e., each node has no special roles, in fact all features are in all nodes; decentralization, the Dynamo design do not foresee any Master node(s); Highly Availability, data are replicated in the network nodes and they must be always available; speed, the system should provide access to distributed data in a very quickly and consistently with user requirements. These design guidelines have privileged partition and availability of data over the consistency of data (or better: dynamo adopted a “weak consistency” model according to Brewer’s CAP theorem (Brewer 2012)[18]. In other terms, data will be distributed and always available even if replicas of data could be in different (and inconsistent) states. Another interesting feature is that data are always writable because conflicts are dealt with during “reads”.

Another relevant example of availability and timeliness in providing information is given by the functions implemented by the Twitter architecture, that is able to provide in real-time information about the activities and the information that users performs or share within the system. In this case the problem is not only indexing information in order to allow a fast retrieval to users, but also to tag the information and make it available in the shorter time possible to user that are checking specific hashtags. In fact the Twitter engine is based on a PubSub model (Fontoura, Maxim and Marcus) and each single event to be published is indexed and delivered to interested users. The Twitter architecture integrates a message queue engine, Kestrel, a Hadoop Base Content Store with a noSQL metadata store based on the Cassandra solution (Lakshman and Malik 2010). Also in this service, consistency of data is a minor requirement compared to quasi real/time availability and partition.

Generally speaking, an approach that favors availability instead of consistency can deal with huge data and can provide very fast response time by discounting the needs of consistency of all the data replicas. However the consistency requirement could be a major need (e.g., in financial and transactional related applications). In these cases, a network able to provide high availability features and to keep delays within specific intervals could be needed. Depending on services, consistency could be pursued in high distributed systems by relaxing requirements either on high availability or on partition of data (centralization). In this case, appropriated control paradigms should be chosen: dynamo and big table have chosen

[18] From Wikipedia: “Brewer's theorem, states that it is impossible for a distributed computer system to simultaneously provide all three of the following guarantees: a) Consistency (all nodes see the same data at the same time); b) Availability (a guarantee that every request receives a response about whether it was successful or failed); c) Partition tolerance (the system continues to operate despite arbitrary message loss or failure of part of the system). According to the theorem, a distributed system cannot satisfy all three of these guarantees at the same time.

P2P, Twitter has chosen PubSub, consistent networking could be supported by the Network Intelligence paradigm.

Virtualization

Virtualization is widely used in cloud computing solutions. Actually the progress of the virtualization techniques is not penalizing too much from the performance point of view. The concept of Hypervisor and related technologies have matured so much that now different options are possible and each of them is not penalizing too much the overall performance of the hosting machines. Typical configurations are depicted in **Figure 36**.

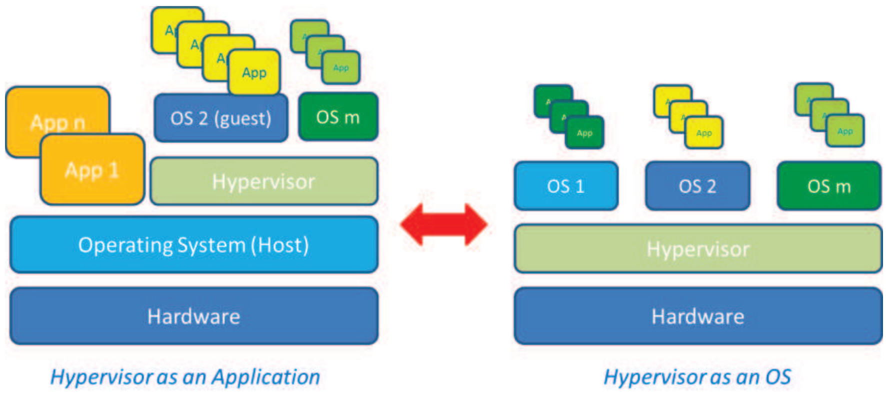


Figure 36: Examples of Hypervisor Applicability.

Virtualization has been applied mainly to processing and storage changing the face of utility computing and taking progressively advantage of multicore systems (Kumar, et al. 2007).

OpenNebula (Sempolinski and Thain 2010) is an interesting open source solution because it provides a uniform and homogeneous view of virtual resources, abstracting away from different virtualization technologies by means of drivers (new drivers can be created to add support to new virtualization technologies). It uses a scheduler, which can be easily tailored or changed, to take VM placement decisions (e.g. to balance the workload or to consolidate servers). In addition it is able to support the Virtual Machines migration and this makes Openebula a good solutions for experimenting in highly dynamic environment like those inspired to P2P or those posed at the edge of the network and supported by edge nodes.

Focalization on Perimeterized Solutions

Cloud computing comes also with a number of drawbacks, for example clouds are designed to interact within an homogeneous environment. Usually providers prefer to impose a close and proprietary environment instead of looking for interoperability with other systems. Providers of Cloud solutions have generally adopted a walled garden approach, i.e., a cloud is internally homogeneous in terms of approach, machine virtualization, interfaces and it is not interoperable with other clouds. This implies that the user that needs services from different clouds has to cope with interoperability issues within the customer domain (e.g., interoperability between resources in Cloud Domain 1 and in Cloud Domain N has to be sorted out within the Customer Domain). **Figure 37** depicts one of such situations in which different clouds are homogeneous (right) but they cannot interwork directly (left) and it is a task of the user to integrate the functions provided by the two different cloud services.

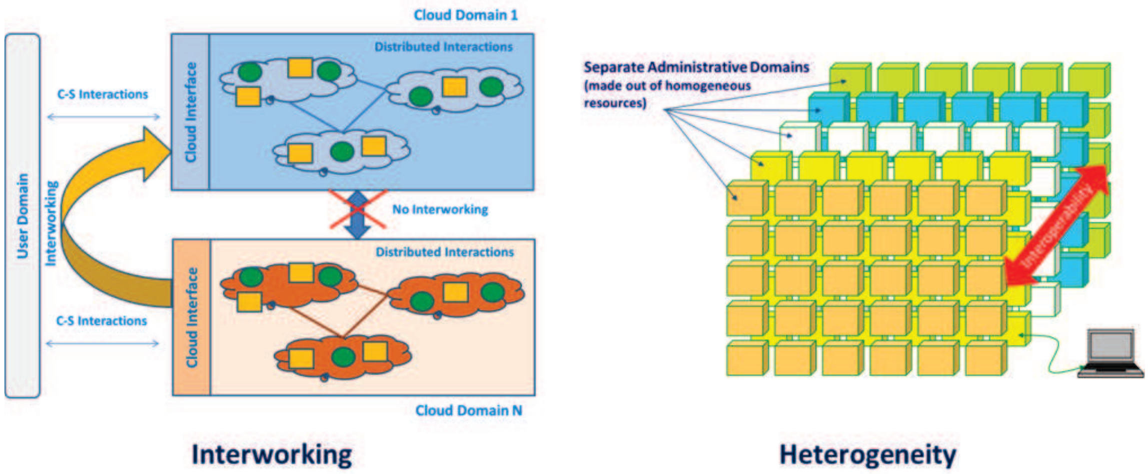


Figure 37: Cloud Computing: Current Interoperability (*left*) and Heterogeneity (*right*)

Actually the interoperability is a major difference between cloud and grid computing (Dillon, Wu and Chang 2010), the latter is more complicated in terms of interfaces and mechanisms but it can support interworking of heterogeneous systems. The concept behind grid computing is the possibility to put together heterogeneous resources provided by multiple providers in order to integrate them into a virtual organization. Resources are negotiated for and are chosen according to specific needs of the applications. Resources can be used as single elements or they can aggregated and used as collections/groups. In addition, grid computing is standardizing a set of programming interfaces in order to allow the development of customized applications and services fulfilling particular needs and requiring specific arrangements of resources. Obviously this architectural characteristic comes with a price: programming at collective or single resource level implies more complexity and a clear knowledge of how to compose and organize resources.

On the other side, in cloud computing, interoperability at customer level can be alleviated by means of virtualization, in fact if all the involved domains are providing virtualized resources of the same type, the customer applications can have an homogenous view on available resources independently from the specific cloud. In any case, the creation of walled gardens is a market strategy of Cloud providers in order to segment and perimeterize the Cloud offering.

Lack of solutions considering network issues

Another major issue of cloud computing is the lack of any references to the underlying networks. Networks are abstracted and networking resources are not made visible to users and applications by means of interfaces. This diminishes the flexibility of the cloud because the integration of processing, storage and communications allows to build platform that are more capable to adapt to the dynamic needs of the applications imposed by the execution context. The assumption is that connectivity is granted and it will be provided according to a best effort arrangement. It is also assumed that the capacity of the supporting networks is sufficient for services and application to deliver the expected behavior. Actually this could be a big issue as pointed out in (Deutsch 1995) and (Rotem-Gal-Oz 2006). The assumption that the network per sè will always be available providing the expected services is wrong and dangerous and it could lead to disruptive effects on services. Many cloud computing solutions are designed in such a way to cope with dynamic behavior of networks and they try to trade off the unreliable behavior of the network by increasing processing and storage capabilities at the edges (in the user domain and in the cloud). As an example of this trend, the Amazon Silk browser dynamically splits computation between the servers and the end device (in this case a tablet) in order to optimize the resources and the processing load between components. This is done to mediate the adverse cases of a malfunctioning network status. Actually some programming languages like Ambient Talk (Dedecker, et al. 2006) have been designed in order to cope with network fallacies. It is based on the possibility of communication processes to keep working while hiding to the programmers the need to check the network connectivity status.

A major difference between cloud and grid computing is the view on resources: the grid has defined and uses interfaces that allow to manage connectivity according to the needs of the distributed applications. If an application needs to transfer a large amount of data between different computing nodes (e.g., specialized in the analysis of medical images), this requirement can be better fulfilled if the application can control the connectivity between nodes and decide when and how to transfer the files. While current commercial solutions for cloud computing are mainly focusing on processing and storage (giving simple representation and access to virtualized images of these types of entities), grid computing is representing resources at different layers and by means of well-defined interfaces. **Figure 38** represents a high level architecture for grid computing.

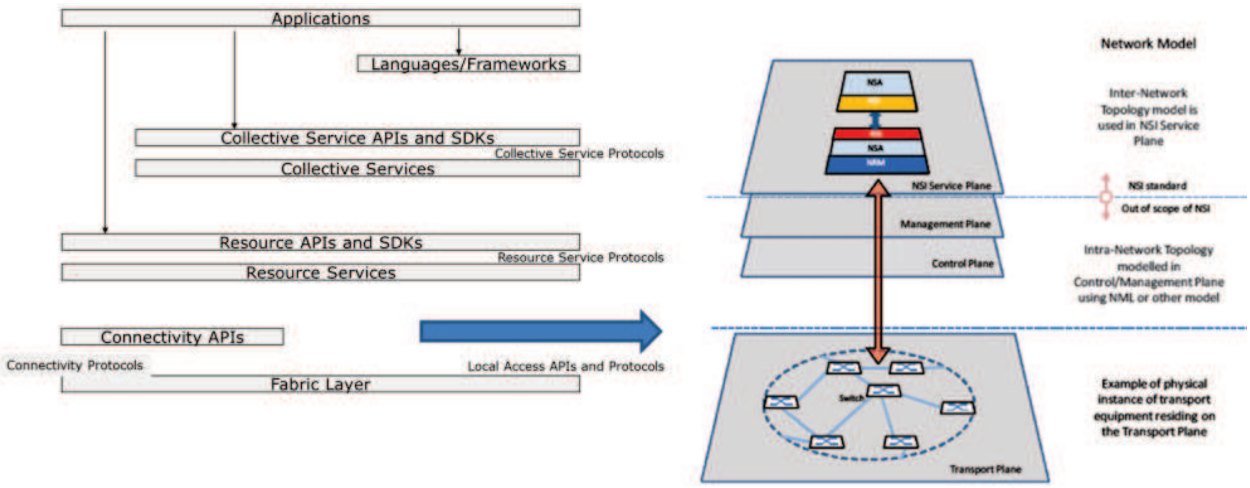


Figure 38: Grid Computing Layered Model and the Network Representation

Some research projects are investigating the combined view of IT and network virtual resources. For instance, IRMOS[19] (Cucinotta, et al. 2010) project has introduced the concept of Virtual Service Network (VSN), which consists in an aggregation of VMs, virtual links and virtual storage nodes. The VSN description, an ontology-based graph model (OWL), specifies hardware and QoS requirements for each of the virtual resources integrating the network. In the IRMOS platform, a VSN is created for each application as part of an automatic SLA negotiation process. During such process, an ISONI provider (i.e. provider of virtual resources) able to fulfil the hardware and QoS requirements of the application is dynamically discovered and selected (all virtual resources will be provided by the same provider). Resources are then reserved and integrated into the VSN where the application will run. At run-time, the initially signed SLA can be renegotiated, but only the amount of resources assigned to the VSN can be modified (not the type of resources, and not their provider). SLA renegotiation, and the subsequent VSN modification, can be triggered because the application

[19] See http://www.irmosproject.eu/

end-user changed the QoS requirements of the application, because the initially amount of reserved resources was not enough to satisfy the signed SLA, or as a response to some scalability rules specified for the application.

Applications at the upper layer have different interfaces that can be used to control and manage resources (or group of resources termed as collective) at different layers and with different granularity. In this way an application can negotiate and allocate a collective resource (an aggregate of functionalities provided by a group of resources that can be controlled and managed as a single entity), or it can access to resources and to connectivity for linking them together. In the specification (Roberts G. 2010) of the Open Grid Forum, a model for a network of network is given in order to present to applications and services an interface for requesting and controlling the composition of connectivity among different networks. A simpler specification (Metsch T. 2011) is provided within the framework of OGF by the Open Cloud Computing Interface, OCCI, initiative. It provides an object model for describing how resources can be connected by means of links and network interfaces. The goal is to allow applications and Cloud Infrastructures to cope with the complexity of supporting networks and to orchestrate the needed resources on a dynamic basis.

Another relevant initiative for the integration of network virtualization in cloud infrastructure is related to OpenFlow (McKeown 2008) and to the definition of a Network Operating System (Gude, et al. 2008) and its applicability to data centers (Tavakoli, et al. 2009). OpenFlow is the parent of new initiatives related to the so called software defined networking. The goal is to allow the opening up interfaces within network resources in order to allow virtualization and programmability in a similar way as cloud and grid computing are offering with processing and storage entities.

Software Defined Networking (SDN) is about virtualizing network equipment and decoupling them from network management and control; not only this, a key facet of SDN is introducing API for programming network services. In principle, this could mean morphing routers into commodity (low cost) programmable boxes controlled and programmed (through API) by an outside source. This research track may have a deep impact on Cloud Computing. For example, in (Armbrust, et al. 2010), it is mentioned how two-thirds of the cost of WAN bandwidth is the cost of the high-end routers, whereas only one-third is the fiber cost. So, simpler network nodes (e.g. routers) built from commodity components (as SDN is planning to have) deployed in WAN, may provide costs dropping more quickly than they have had historically, enabling new paradigms of interactions Cloud – Network.

In this direction, OpenStack (Pepple 2011) is an open source cloud project and community with broad commercial and developer support. OpenStack is currently developing two interrelated technologies: OpenStack Compute and OpenStack Object Storage. OpenStack Compute is the internal fabric of the cloud creating and managing large groups of virtual private servers and OpenStack Object Storage is software for creating redundant, scalable object storage using clusters of commodity servers to store terabytes or even petabytes of data. Interestingly, OpenStack has a network connectivity project named Quantum[20]. Quantum looks to provide "network connectivity as a service" between interface devices managed by other OpenStack services. Quantum itself does not talk to nodes directly: it is an application-level abstraction of networking. It requires additional software (in the form of a plug-in) and it can talk to SDN via an API.

3. 9. Terminals and Devices Evolution

The basic key technologies (see **Table 3**) having an impact on the evolution of terminals are rapidly progressing and in the future years they will have meaningful effects on terminals. For instance:

- New Radio Interfaces (beyond LTE advanced) and multicore technologies will bring the capability to have more bandwidth, in fact each core could process flows at a specific frequency, and the aggregation of different point to point connections at specific frequencies will exceed the current limitations;
- Each terminal will be capable of handling different technologies for exploiting local communications.
- Always best connected services will be dictated by the terminals that will use available spectrum (and cognitive radio capabilities) to get the best connectivity possible in the operation context
- New Interfaces like Touch, Voice, Gestures, Bio-neural will allow the creation of natural and easy to use interfaces. Some interfaces could be portable from terminals to terminals becoming a distinctive market offering;

[20] Information available at https://wiki.openstack.org/wiki/Quantum last accessed May 29th 2013

- M2M and IoT capabilities will be such that each terminal will be part of crowdsensing communities, i.e., aggregation of terminals that gather information on the context and the environment and return it to a community that will use it to provide services or simply monitor a part of a city;
- Displays (e.g., HD, 3D, Flexible, Pico-projectors) will support a high quality almost indistinguishable from the reality and this will enable new services. In addition the terminal itself will be able to project in the environment video and images.
- Disappearing of SIMs; Identity will be based on different mechanisms (bio-Id) that will allow for higher levels of Security, Profiling, Services because the terminal will be capable to recognize the actual user and to setup the personal environment in cooperation with cloudified information and applications;
- Storage capabilities in terminals will continue to progress and it will be abundant also in low level terminals. The combination of large storage capabilities, availability of connectivity and function provided in the cloud, and increased processing power as well as multimedia features will make possible to record almost any moment of life of users.
- Smart Materials, and Nanotechnologies will allow the creation of flexible, adaptable and wearable terminals that will act also as sensors for users.
- Batteries limitations will remain and likely will be the major barrier to further developments.

Terminals will be more and more "personal", they will be worn all the time, they will be disassembled and made extremely portable (some components could be deployed in or use the human body or clothes). Almost any aspect of personal life will be measured and stored. The positioning of Internet companies in the control of terminals/services will become even more stringent and there will be the need to have Regulation that safeguard privacy and establish fair rules for the usage of data captured by terminals. The current battle for supremacy in the mobile operating system sector will have taken new forms, but it will be fundamental also in the future. Likely, application portability and the breaking of closed walled garden will become a necessity for users, but also for the industry. There will be a movement for a standardization of terminal capability that could lead to a situation similar to the current web, i.e., a few stable protocols and mechanisms that enable users to take advantage of the web. In this context there will be an increasing synergy between the terminals and the cloud in order to make readily available to the users all their data and environments. Already now some browsers (like Amazon’s Silk and Apple iCloud) are exploiting synergies with cloud applications, this trend will consolidate even more in the future (Bahl, et al. 2012).

Users will strive to access in the same way the same service from several different terminals and will look for the same user experience. Virtualization of terminals and applications within the cloud will be a possible trend that will enable the users to break the walled garden approach taken so far by mobile terminal vendors.

Terminal will also be the means to provide augmented and virtual reality applications that will make possible the creation of new compelling and immersive services. These services should be greatly customized in order to present to users the desired information.

Terminals will also be a major means to support new forms of payment, from prepaid card systems up to newer ones. Virtual and digital money will be a major issue. So far all the money transaction on the Internet are tracked, there is not the possibility to spend money in an anonymous way for buying goods like it happens in the real world. Possibly this will change if digital currencies like Bitcoin (Barber, et al. 2012) will have success. In any case, mobile terminals will be the new wallet of the future.

The combination of all these features and challenges will make the terminals even more central , if possible, in future communications ecosystems. Having a grip on terminals, will mean essentially to have a chance to own the users. On the other side, the progress of technologies and the awareness of Users and Regulators should also help in limiting unfair usage of portable devices. In any case, users should acquire much more awareness of their possibility to control this fundamental element of the ICT ecosystem.

3. 10. Findings

As seen in Section 3. 2. , the evolution of basic technologies will continue at the same pace of today. In a ten years timeframe, the processing power will double any 18 months leading to an increase of more than 50 times while the power consumption will be cut down. Solutions like RasperryPi will not be an exception and be widely deployed. Storage will follow a similar path. Screen technologies will support a “natural” interaction with humans. Pervasiveness of computing will be supported by the wide availability of low cost connectivity granted by improved local capabilities and the availability to manage frequencies at the local level in freer modes. These developments give a clear idea of the possibilities offered by tomorrow technologies well beyond the current status. Data Centers will become commodities and distribution of functionalities over different administrative domains will become a stringent requirement. Virtualization will allow the creation of adaptive and virtualized computing environments. Their possibilities will be further considered and described in Section 5. 5.

Data will play a fundamental role: owning data and being able to infer information from them will be a distinctive advantage of some Actors. This advantage should be mitigated by appropriated and fair policies for guaranteeing to users (to be interpreted as citizens and customers) some levels of privacy and ownership on their data.

With respect to data center and cloud computing technologies, this chapter has shown that WebCos have a technology advantage over many other Actors. This advantage is built on the C – S approach as an interaction paradigm (simplicity and efficiency in getting the service functionalities). It will be difficult to close the gap with them for any company willing to have a predominant role in the “cloud”. These companies are advancing the technologies behind the server front-end towards highly distributed systems preparing for “switching” to even further level of distribution of functions. Following them on this terrain is extremely risky, time consuming and investment heavy: Actors like Google and Amazon have made their technical infrastructure an asset that is daily increased by adding new functions and new solutions. In addition they have a total control on the systems having built it piece by piece. TelCos have not this capability and this fact has consequences as further elaborated in Chapter 8. scenarios.

The evolution of terminals is driven by some major factors: the progress of basic technologies, the disposition of users to invest into terminals, the need of terminals in order to access to services. Terminals are now seen by WebCos as a means to control the Client side of the C – S model. Entering in this area allows them to have a full control on the value chain of the services leaving to TelCos only the commoditized part of the connectivity.

In this evolution, the User is constrained to a passive role: consumer of services, consumer of terminals and consumer of data. Users can contribute to services in terms of profiles and user generated content. However the increasing capabilities of terminals could be an element that will give some freedom to users: in fact mobile terminals can be used to directly communicate (Direct to Direct communication can be seen as the new frontier of some communications services that further disrupt the TelCos market); they can be used to aggregate into local clouds capable of elaborating local data and providing local services to the participants; terminals can collect and process large quantities of personal data preserving privacy of users. Terminals capabilities will be fundamental in order to access and support functionalities of highly adaptive and dynamic smart environments.

In a linear projection of technology evolution, the power of current solutions will allow WebCos to keep an edge over the other competitors: services could be considered as a natural basin of these companies. This has the consequence that the role of TelCos will remain essentially the same also for the future: traditional TelCos or connectivity providers. The technologies and the solutions as well as the business motivations behind these scenarios will be further elaborated in the following chapters and will be collectively described in Section 8. 6. and Section 8. 7.

However the growing capabilities of terminals, the increasing awareness of some individuals for a different approach to privacy and management of personal data could be interpreted as two possible inflection points capable of determining new paths in the evolution of the ICT industry. These two capabilities are important because they may enable new scenarios for the service provision (e.g., Section 8. 10.).

Chapter 4. The New Networks

Highlights

- Ossification of the Internet and Requirements for its evolution
- New Approaches for Networks: Virtualization, Software Defined Networks, Information Centric Networks, Cognitive Networks
- Disruptive Approaches

4. 1. Introduction

The main purpose of this chapter is presenting some possible paths for the evolution/transformation of the network. It is clear that the new network will have to cope with a gigantic increase in data communications demand. Cisco has come out in 2012 with “dramatic” forecasts on the increase of bandwidth request (CISCO 2012), that in the course of 2013 have been reduced. In order to keep pace with this surging demand requires flexibility and reduction of costs in network equipment. It is not economical sustainable to invest large amount of money in technologies that do not guarantee to be valid for the years to come. There is the need to introduce in the networks low cost technologies that have a Return of Investments, ROI, in a few years. This is clearly in sharp contrast with the investments in access networks and in the infrastructure that the TelCos have to face with the next years. However the access networks constitute a sort of monopoly and, in spite of a huge capital investment need, they can leverage this advantage. The control and service infrastructure, instead, are more characterized by software and as such they are prone to be improved, modified, extended and replaced in shorter cycles. There is an increasing discrepancy between the evolution of the network infrastructure and the service layer.

The network evolution has always been regulated by standards and by a progressive and smooth adoption of them. Operators have had time to align to a controlled and predictable way of evolving the network. In these days, however, something is changing under the pressure of the increasing demand for data (especially mobile data). The proliferation of “half generation" solutions like 2.5G, or 3.5G are examples of how the market is requesting operators to deploy more and more capable solutions. Keeping aligned the evolution of demand and standardization is not easy and it is becoming more difficult. Standards take a relevant time to be specified and implemented. Surprisingly the definition and adoption of new technologies in the access is faster in the access than in the core (and the service realm). For instance the LTE specification is more advances in the access than in the solutions proposed for the service control (i.e., the IP Multimedia Subsystem). There is an increasing risk of deploying in the network solutions and architectures that have been designed several years ago and that are not capable or viable to cope with current needs and problems.

In addition there is the increasingly important issue of the Internet ossification, i.e., Internet mechanisms that have been solid for years are now becoming obsolete and need to be updated also from the technological and architectural points of view. In this section, Future Internet requirements and needs will be discussed stemming from existing research initiatives (e.g., GENI (Peterson and Wroclawski 2007), Fire (Gavras, et al. 2007), Akari (Aoyama 2009). These projects are proposing an extremely flat and programmable network infrastructure that exceeds the concept of communication-only with the integration of «distributed processing, storage, and sensing/actuation» within a common infrastructure.

Driven by the Future Internet, the evolution of the Networks towards a more software driven approach based on general purpose systems is capturing the interest of many TelCos. The reason is that these seems technologies capable of reducing the investment costs and allowing the TelCos to stay aligned with the technological evolution cycles. Three major trends will be considered, namely Virtualization, Software Defined Networking and Information Centric Networking. They present features that can be implemented in a mainstream fashion (i.e., aligned to the foreseen and standardized evolution of the network) or they can be used in much more disruptive ways. These approaches will be further discussed in this Chapter.

In the last part of the section a few divergent trends are considered with special attention to their disruptive aspects. They aim at showing that the future networks could be extremely different from how the linear progression proposed by the standardization describe the infrastructure. They introduce and use technologies in a different way creating new possibilities.

4. 2. Ossification of the Internet and Future Internet

The current Internet is based on autonomous systems that decide internally how to allocate communication resources (in terms of routing capabilities); these resources offer a best effort communication capabilities. **Figure 39** represents independent networked functionalities that work separately without any possible strict coordination, but just using a best-effort approach. All the issues and problems in delay or lose of information, as well as security and mobility are to be solved by the end nodes and the end applications.

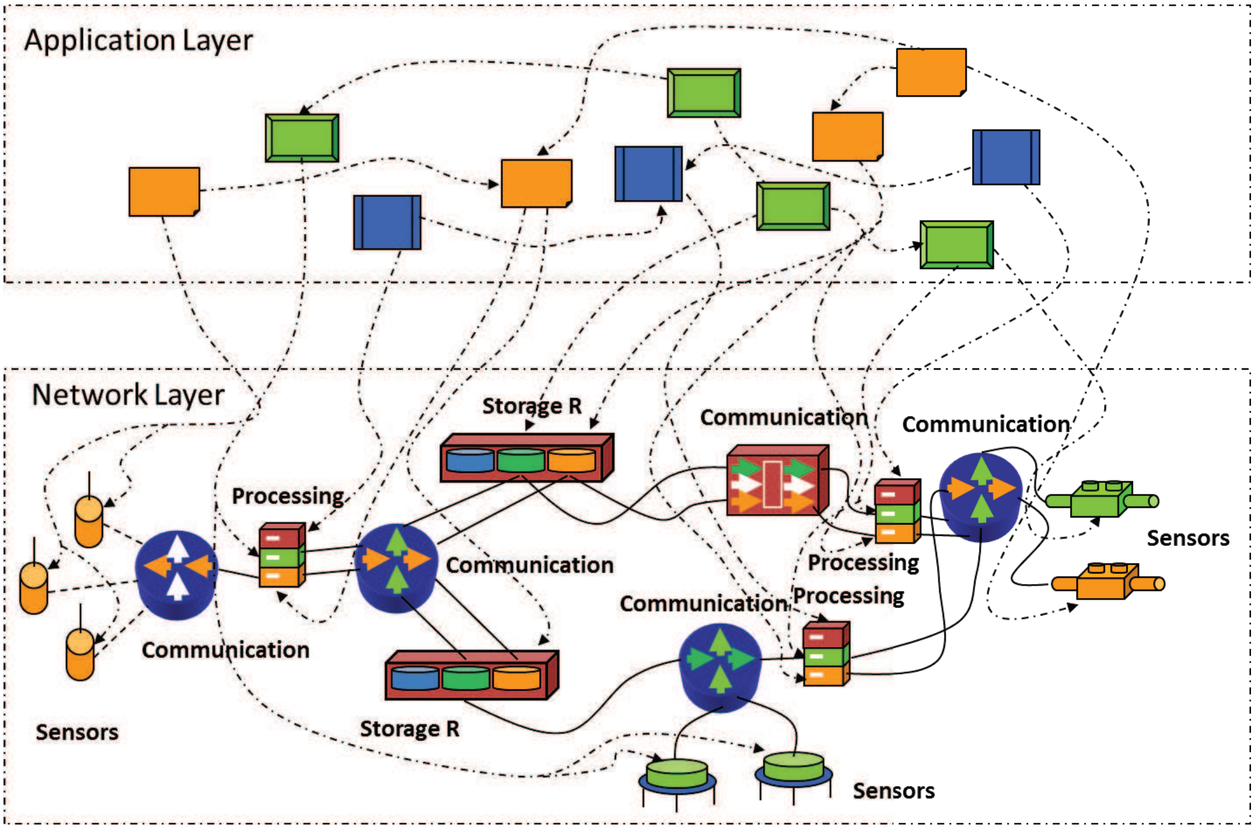


Figure 39: Independent Autonomous Systems Supporting Best-effort Capabilities

A few issues are related to the IP ossification. As stated by (Turner and McKeown 2008): "*Today, many aspects appear to be set in stone, such as difficulty in getting IP multicast deployed, major obstacles to deployment of IPv6, not optimized routes (problems with BGP), extensions to routing— such as multicast, anycast, . . .*". This poses the need to change the Internet to allow for natively encompassing security, mobility, better resource management and pervasiveness features. It is quite difficult to make any changes in the infrastructure of the internet and following the end-to-end principle (Clark, et al. 2005), major changes are possible only at end nodes, not in routers. In addition, this monolithic approach has been exploited by a small number of companies to create a sort of monopoly on the IP layer based on closed routers and systems. As a consequence, more flexible solutions have been implemented by means of overlaying mechanisms (i.e. building virtual networks on top of the Internet, e.g., P2P). These overlay networks clearly show programmability and flexibility points as well as the possibility to improve the underlying protocols. Examples are the JXTA P2P architecture (Wilson 2002) or the many solutions created on the BitTorrent protocol (Pouwelse, et al. 2005).

The major Internet rigidities are:

- Security, the current internet has not been designed to cope with security issues caused by the availability of many and distributed computers. Classical examples of Internet flaws in this are the Distributed Denial of Service attacks DDOS that exploit the possibility of different sources to concentrate traffic and requests on a single or a few server nodes causing these to go out of service. The network as a system has not capabilities to detect and act against these and other threats.
- Availability/Reliability, having adopted an autonomous system approach with limited cooperation between nodes and adjacent subsystem, the reliability of the solution resides in the capability to route information over other paths. There is no assurance that a system will be up and running and it will provide a certain level of service. The entire services developed over the Internet are unreliable by definition.
- Predictability, systems do not offer any means to determine or to control whether a path or a set of subsystems will be available, nor that a service can be provided in a predictable way. There is not the possibility to predict beforehand how a service or a subsystem will behave.
- Manageability, typically the nodes are closed systems with limited management capabilities. Management functions are exerted essentially by means of Client Line Management (CLI). This requires a lot of work from human operators and often, important management functions are to be executed manually. A big source of issues in the routing is the human based configuration of nodes.
- Mobility, in an increased mobile world with plenty of mobile terminals and devices, the Internet and the mobile networks are not fully supporting the possibility to access objects and functions on a very dynamic base. Roaming and its issues like the need to have an anchor network (the home network) for accessing to services is an example, another one is

the availability of plenty communication capabilities and devices and not being able to use and share them. In addition, the problem of locating applications and objects and to create a local supporting environment for optimizing the experience are all concepts out of scope for current networks.

- Sensing, also sensors are not well integrated in the Internet (some solutions have even proposed to move to different protocol than IP (Wan, Campbell and Krishnamurthy 2002), (Dunkels, Alonso and Voigt 2003), because it is cumbersome and over specified for small devices). Sensor networks will also need different types of transport capabilities and probably they will benefit from reliability and predictability of network services in order to support mission critical applications.
- Scalability, large scale system are difficult to maintain and operate over an Internet. There is the need to support mechanisms and solutions for easing the goal of integrating and let the different components working together. This is even more true considering that the future network will not provide simple connectivity but it will evolve towards the integration of communication, processing, storage and sensing/actuation. And these features are not integrated nor supported natively by the internet.

The internet is a major building block of an increasing critical infrastructures for people and companies. The previous features need to be present in an new infrastructure in order to improve the effectiveness of the internet and its usage. Those features point also to persistent issues that cannot and have not been solved with incremental improvements to the current architecture (Roberts 2009), (Rexford 2010). There is a lot of work going on in order to find out new paradigms that can prove more powerful, on top of which to build the so-called Future Internet.

Another important issue that is not dealt with in the current Internet is the clear separation between the “network” stratum and the application one. In principle it is a good separation, but it creates problems because the network cannot adapt or fulfill the dynamic changes of the application. A level of interfacing between the network and the application can guarantee that the resources of the network are not wasted, but they are cost-effectively used in order to reach two optimization goals: a) the usage of resources on the application side, i.e., the resources are allocated in order to fully satisfy the needs of the application; b) the resources are optimized at the global level, i.e., only the needed resources are allocated to competing applications and a fair strategy is applied in order to minimize the usage and allocation of resources to application. This twofold problem needs to be supported by the Future Internet in order to solve some of the previous issues (e.g., manageability and predictability).

With respect to **Figure 39**, there is the need to introduce a higher level of cooperation between systems and nodes in order to offer a better quality service (it is not the case to introduce guaranteed levels of services, but instead to guarantee that resources are utilized in such a way that the solution adopted can be considered "quasi-optimal" and that during time (if conditions stay stable) the solution can converge versus an optimal one. Simple algorithms and protocols can be used in order to allow changes in status of resources so that they can progressively find optimal solutions. (Jelasity, Montresor and Babaoglu 2005) demonstrated that under this condition a complex system can converge towards the optimal solution. **Figure 40** illustrates the difference between a best-effort approach and a cooperative one.

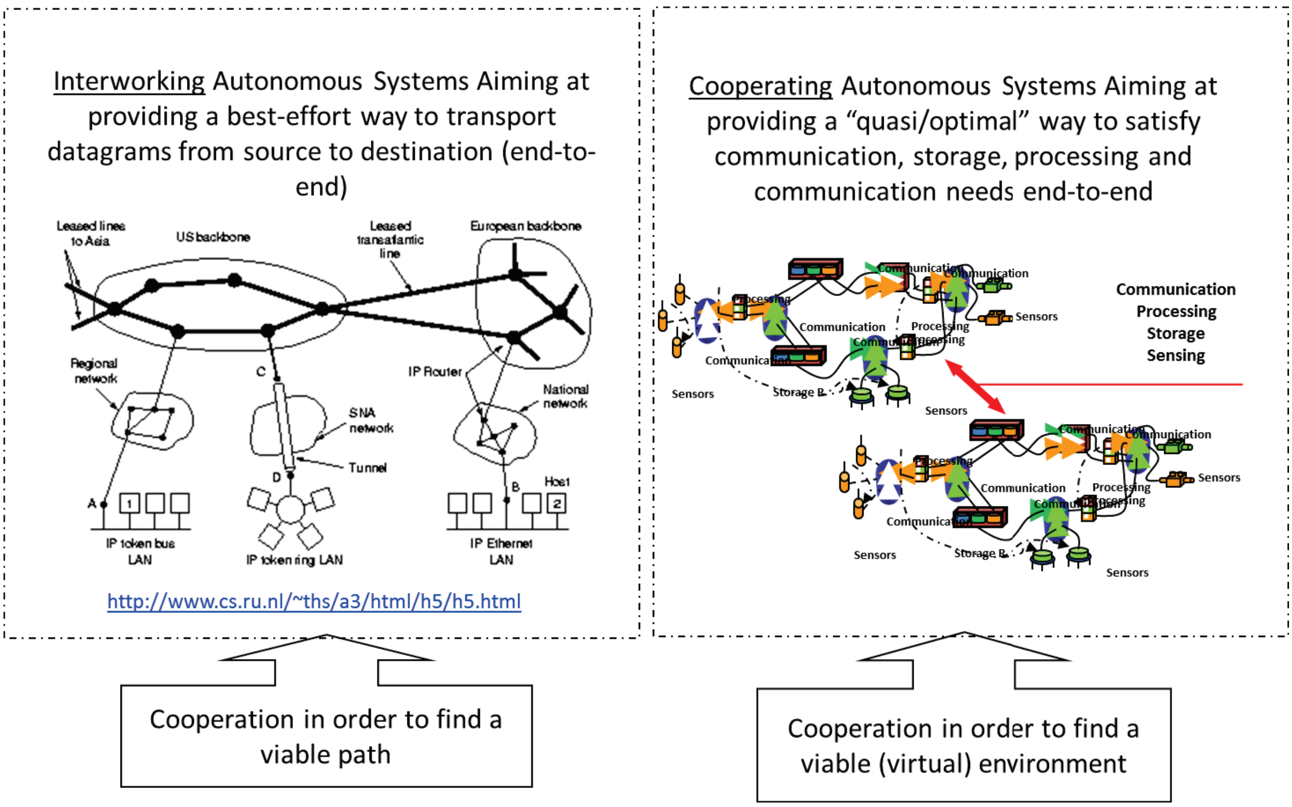


Figure 40: Cooperative Systems in which Each Node has a Simple Behavior

Two major properties can are needed in order to further innovate the Internet Architecture:

- Programmability in the nodes; if programmable functions are instantiated in Internet nodes then a high level of network flexibility and adaptability can be reached.

Programmable nodes are the building blocks that allow to make the Internet less monolithic and are the key to introduce richer functions and features in it.

- In order to keep pace with the solutions and the mechanisms needed to program and control a complex infrastructure, virtualization could be widely used. It brings the ability to co-locate multiple instances of network services on the same hardware – each running in a different virtual machine, but also the capability to confine specific applications in a virtual machine. Crashes and security issues of an application can be confined to a specific virtual machine, in addition the set of virtualized applications and functionalities can be increased depending on the demand of users by replicating virtual machines in the system. So virtualizations seems to be a useful mechanism to move towards the Future Internet.

This path to change could be interpreted in different ways, some initiatives like GENI (Peterson and Wroclawski 2007), Akari (Aoyama 2009), or FIRE (Gavras, et al. 2007) are adopting a clean slate approach, i.e., they advocate the need of a totally new approach for defining the Future Internet. Many Vendors, instead, are more conservative in this approach and they propose the progressive opening up of routing interfaces to the application layer. These two trends will be further discussed in the next sections.

4. 3. Virtualized Networks

Virtualization has widely being exploited by the IT industry, a working definition of Virtualization is “the ability to run multiple operating systems on a single physical system and share the underlying hardware resources” (VMWare 2006). Virtualization carries in a number of advantages such as:

- Expanding hardware capabilities, allowing each single machine to do more simultaneous work
- Control of costs and simplification of management through consolidation of servers
- Higher level of control of large multiprocessor and cluster installations, for example in server farms
- Improvement of security, reliability, and device independence possible thanks to hypervisor architectures
- Ability to run complex, OS-dependent applications in different hardware or OS environments

But why and how to approach virtualization in the Network ? Today there are already some low levels of virtualization of the network. For examples, solutions for managing and virtualize connectivity have introduced functionalities related to the creation and support of virtual networks based on MPLS or GMPLS and VPNs. This virtualization, however, offers a coarse-grained network link virtualization that is far from allowing the management of fully fledged virtual networks. Examples of virtualized networks are overlays networks (as P2P networks over the Internet). They can be seen as virtual networks at the application level. One issue is the relations between the virtual network and the underlying physical infrastructure. This missing nexus is highly detrimental either at the application level either to the physical level, in fact applications cannot control how networked resources can be allocated and used, while the physical network resources are not optimized because the overlay network can refer and use resources that are far from the actual location. The issue (tackled by P4P (Xie, et al. 2008) and ALTO (Gurbani, et al. 2009)) is that the separation between applications and the network is not providing optimized solutions, resources are misused and the service returned is not always adequate for the applications.

A network wide virtualization (using the same paradigm used for IT resources) would allow:

- To optimize the use of physical resources (as previously discussed)
- To integrated deeply IT and Net resources in virtual networks tailored to applications requirements, This is a giant step with respect to actual situation just focusing on the connectivity.
- To operate independent virtual networks "dedicated" to different Users and migrate them when necessary. So applications and services can be segmented and can use specialized networks to reach their goals. One example is the ability to create independent network slices "dedicated" to different Players (e.g., virtual network operators, application service providers for video streaming, CDN, etc)

Recently, a group of Operators has started a new initiative in ETSI called the Network Function Virtualization, whose goal is to foster the wide adoption of virtualization within the network. The grand plan is to create a general purpose computing infrastructure on top of which to instantiate several images of specific nodes. This has the advantage of decoupling the hardware procurement from the software one. In addition, instances of specific nodes could be acquired by different vendors, and deployed remotely by using standard IT solutions. **Figure 41** represents these possibilities.

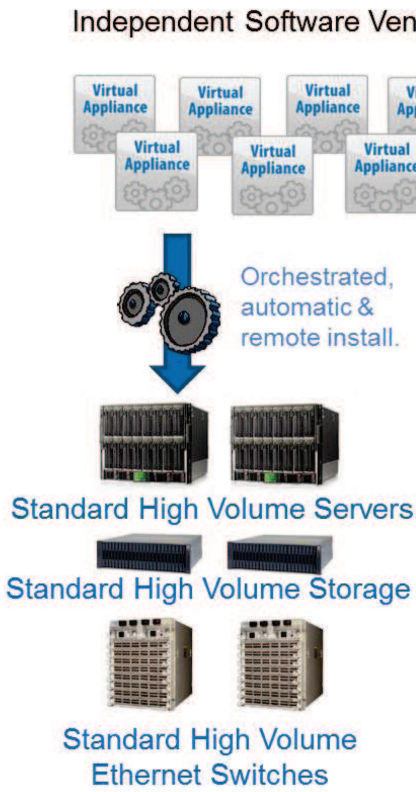


Figure 41: Network Virtualization Approach

4. 4. Towards Software Defined Networks

Together with the Virtualization another trends is acquiring momentum in the industry: the software defined networking was spawn in 2008 by studies carried out in Stanford for a clean slate approach towards the evolution of the Internet (Yap, et al. 2010), (Koponen, et al. 2011). This approach draws its root into older approaches such as programmable networks (Campbell, et al. 1999) or even TINA (Berndt and Minerva 1995).

In traditional IP networks, a node has its own control plane and the management/ policy plane actuate configurations with CLI (Command Line Interface), there is automatic support for dealing with the complexity of a whole system and avoiding configuration issues. This is also a major source of errors and issues in the network. In a Software Defined Network, SDN, control and data planes are decoupled, so network control and states are logically centralized, and the underlying network infrastructure is abstracted from the applications. This means that the routing nodes can be controlled by means of specific protocols such as OpenFlow

(McKeown 2008) and all the intelligence can be moved somewhere else. In addition the switching capabilities can be exerted by very low cost devices that can be bought for low prices. In this way, the Control Level becomes programmable because SDN offers programmable interfaces (i.e., APIs) to the network. It is then possible to implement new "routing" protocols (e.g., customize paths for network traffic engineering), to intelligently allocate resources to the needed applications. The rich and intelligent functions move to a programmable infrastructure that changes the way in which networks (and autonomous systems) are governed and orchestrated. Network wide policies specific for customers or services can be implemented and controlled. The network can be controlled and managed with a holistic approach. This means that cognitive capabilities can be applied and the network will became highly adaptive an responsive to changes in its usage.

The next step is to create a Network Operating System (**Figure 42**), i.e., a set of functions and interfaces that programs can use in order to access services of the infrastructure. The consequence is that routing mechanisms, but also multicast, processing and computing became services that can be mashed up and composed. The definition of a Network Operating System will allow applications to orchestrate how the network responds to the requests of the applications. The rigid separation between the network level and the application one is reconcile by means of APIs and programmability.

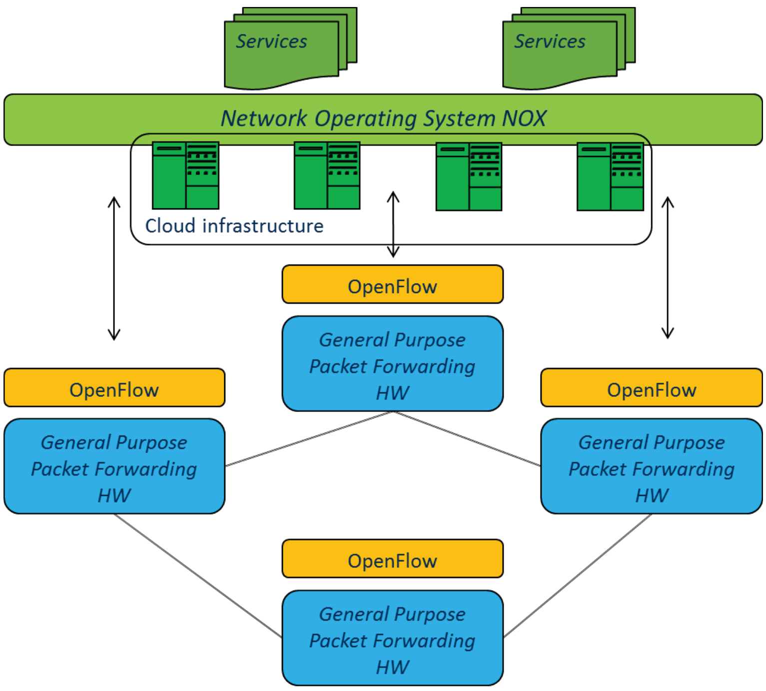


Figure 42: A Network Operating System, NOX, based on OpenFlow Interfaces

This trend will have a deep impact on the industry and IP protocols. Operators willing to provide services and platforms have to consider the availability of these platforms. Operators interested in being Bit Carrier, will prefer to remain at level L1-L2 and to exploit the optical technologies in order to save on energy, processing and intelligence in their network. This is one of the most important inflection point (and decision point) in the evolution of technology because it shows that large networks with a considerable bandwidth can be managed and operated by functions of L1 –L2. Moving at upper layers will cost money because systems are more intelligent (and more expensive) and they consume more power. A Bit Carrier will try to use as much as possible lower level equipment and it will avoid to move to upper levels. In order to fully appreciate the innovation carried by the software defined networking, an example is explicative. Using the concept of NOX an Operator A can offer to Operator B the usage of its resources by means of open Interfaces. Operator B could even decide that the most important pieces of logic will remain remotely located in its data center, but the effect of control capabilities will be exerted to the local level (to the Openflow enabled resource). Roaming could be totally reshaped with this approach as well as the relationships between the Operators and other infrastructure providers. In fact a company can offer its programmable resources to Operators that can implement on top of it services for their clients. **Figure 43** tries to represent a few of these capabilities.

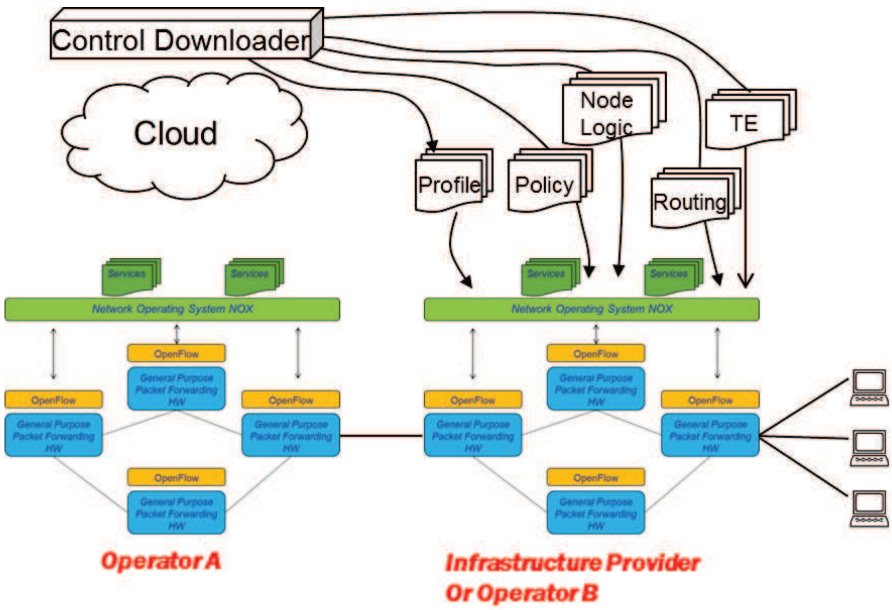


Figure 43: Deperimeterization of Networks by Means of SDN

This example shows how programmability and opening up of interfaces of resources leads to deperimeterization of services and networks. Actually a provider could negotiate dynamically the allocation of needed resources independently from the location and to build a solution to support the connectivity and the services of its customers. Services are not anymore strongly tied to the physical network infrastructure. Perimeterization is a major drawback that operators have to solve in order to compete worldwide. The consequences and the business models behind this new approach have still to be fully understood and explored.

4. 5. Information Centric Networks

In the context of Future Internet studies there is a trend focusing on the importance of accessing to data and information. In a seminal presentation, Van Jacobson (Jacobson, et al. 2009) presented the concept and justified it by saying that over 90 per cent of communication was human to data (i.e., a human accessing to data) instead than a human to human (or a point to point) communication. Adopting a data oriented view leads to the idea to reshape the networking capabilities focusing on how to access to indexed data. Data should be identified and stored in a sort of large P2P network and information accessed according to several policies established by the owner of the information and the preferences and rights of the user. From an architectural view (Rothenberg 2008) this approach could lead to a three layer functional architecture. At the lower level a data plane for the control and transport of chunks of data and information. At an intermediate level, an information plane is in charge for dealing with the information, i.e., naming and addressing, indexing of content, management of replication, mediation of formats, data mining and correlation, reasoning and inferring of information from available data and information. At the upper level the functions are related to how data are accessed and made available to users, to security of data. Another important set of functions could be related to how the data can be certified, updated and even cancelled. This aims at lessen the problem of not authorized duplication of data. At this level, solutions like Vanish (Geambasu, et al. 2009), that allows the users to have control on data that they “post” on the Internet, could find a natural context for usage. In addition this layer could be in charge for dealing with how providers can control and establish policies for usage. On top of this infrastructure Application Programming Interfaces could be provided in such a way to allow the creation of applications or simply the management of data. **Figure 44** depicts a possible architectural view of the ICN considering this different plans.

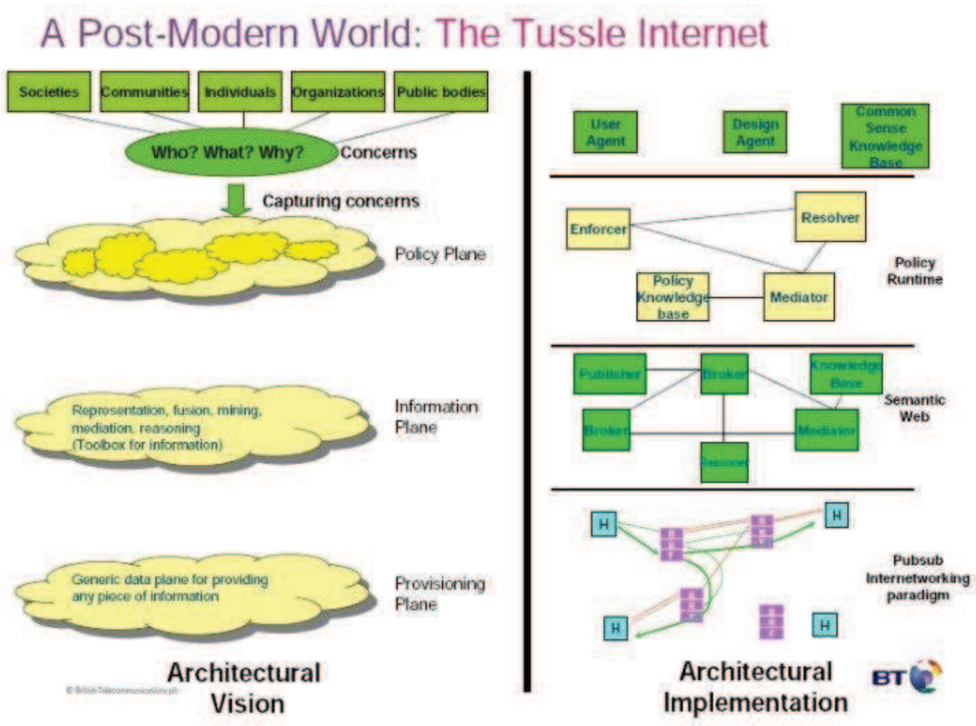


Figure 44: Architectural Vision and Architectural Implementation (Source: BT)

From the point of view of implementation the architecture can be based on a number of available technologies: at the lower level a P2P infrastructure for data exchange (e.g., based on bit torrent) could take care of the optimization of the bit exchange, while a PubSub (Fontoura, et al. 2013) infrastructure could help in notifying the availability of new data or duplication or update of existing data. In order to give a glimpse of possible solutions, an architecture could be organized as follows:

- At the information Plane level, a brokering infrastructure could find a viable application, for instance an infrastructure based on Distributed Hash Table (and the federation of them) could help in tracking the available data and information, also NoSqL solutions could be used to store indexing and pointers to data objects.
- Data mining solutions could also be integrated at this level.
- At the upper layer there is the need to implement a sort of Policy Based Systems (Follows and Straeten 1999). At this level the Policy Decision Points will decide and determine how to grant access and the permitted usage of Data while at the underlying levels, Policy Decision Points will retrieve, format and present the data accordingly.

The topic of “Information Centric Networks” had a moment of hype in the past years and currently it is not so popular, however it points to meaningful and important issues of New Networks that deserve research and thinking. At least these aspects are relevant:

- The need of highly distributed architecture for dealing with information storage and replication and fast delivery to users (Content Delivery Networks, CDN)
- Optimize architecture for information retrieval especially in relation to the personal data issue
- The issues related to policing of access to data and information (e.g., the mentioned Vanish approach)

Current CDN architectures are a sort of overlay network not integrated within the networking infrastructure. In the future the network will be a programmable combination of processing, storage, communications and sensing capabilities. The ideal case is to have solutions that can integrate the SDN capabilities with intelligent caching and processing. Actually new mechanisms for sharing algorithms and programs that produce the wanted data as result of calculation (Katti, et al. 2008) can be considered as an alternative to transport an overwhelming quantity of data. Transferring a program and initial data could be more efficient that transport the entire mass of data. The user processing capabilities can be used to execute the algorithm for deriving the desired data. Networks could be totally reshaped by similar approaches.

From an Operator point of view, a possible strategy for dealing profitably with data could be the one to optimize the resources for transferring the bits (at the data layer) and to develop differentiated mechanisms for dealing with personal data (consumer market), company data (business market) and Content Provider data (e.g., the big web companies). In each case, there is the possibility to optimize the level of experience of the customer. The underlying need of different customer segments are the same: data are generated as results of processes. They can be personal processes (e.g., shopping, travelling, community experience) that can be made evident to the user in order to help him to understand his behavior; or they can be company processes (how a company deals with data produced by its customers, data related to internal processes and the like); or they can be related to how data flows are generated and consumed by customers of content providers. Being able to optimize how data are managed and handled in an enterprise means to understand the internal business processes. Many companies have the need to improve processes and to save money by making those process more effective.

Acquiring this knowledge is a change of perspective that gives new opportunities to the Operators. This is another inflection point: winning companies (and Operators) will be able to intercept data and help the customers to improve their internal and external processes.

4. 6. Cognitive Networks

Many technologies fall under the umbrella of "Cognitive Networks". They range from Software Defined Radio (Jondral 2005) to Autonomic Computing (Marinescu and Kroger 2007). The communality is the attempt to introduce higher levels of intelligence in order to optimize the usage of complex resources. [XIX] presents an extensive treatment of the technologies, and their evolution. The interest and the trend towards cognitive networks is consolidating at the standardization level (e.g., the AFI initiative (Chaparadza, et al. 2009)) and within the industry (for instance the new NoOps approach (Fu, et al. 2010) aiming at leveraging automation, from development through deployment and beyond, to increase the speed at which applications can be released in a cloud environment). The cognitive capabilities are considered under two perspectives: 0-touch networks [XV] (or systems), i.e., how self-configuring capabilities can help in coping with highly dynamic changes in configurations of networks and systems; Networks of Networks, i.e., how intelligent functions can be used in order to allow the self- organization of complex systems [IX]. Future Networks will integrate at least processing, storage, communications and sensing capabilities provided by highly distributed systems. These systems will be highly heterogeneous and they will pertain to several administrative domains. Management will result more and more complex because the systems will need to integrate functions that so far have been separated: in fact, processing and storage from communications and sensing and actuation are three different functional groups with their own specific requirements. The single resource will be forged by the integration of hardware (proprietary and general purpose) and basic and specific software (possibly by different providers). In addition, systems will depend from different management authorities, each one with its policies and management requirements. Another important point is that systems will aggregate an work together in a very dynamic way: the concept of ephemeral networking make this concept very explicit [XXVII, XIX]. For these reasons it will be extremely important to approach the problem of highly distributed and heterogeneous systems in a different way: cognitive technologies could be extremely helpful in changing the way if how systems are managed. Each single resource should be able to detect its own behavior and optimize it with respect to its intended working and to the environment in which it is integrated. This results in two control loops used to optimize the behavior of the resource: a local one intended to optimize the internal behavior and a global one used to optimize the behavior of the resource with respect to the external environment. **Figure 45** depicts this concept.

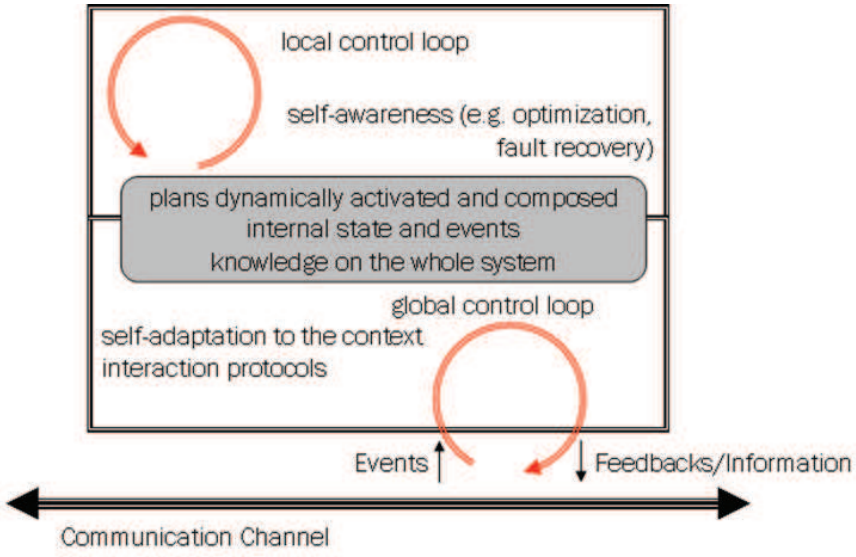


Figure 45: Local and Global Control Loop

In order to devote the large part of processing capabilities to the functional goal of a resource, the approach should favor the lightness of the management capabilities. Under this respect, gossiping based solutions seem to be attractive because they reduce the “behavioral load” of the resource, in addition it has been demonstrated that gossiping algorithms converge to the optimal solutions, if enough time and iterations between objects is given (Jelasity, Montresor and Babaoglu 2005). This is important because simplicity can help to converge towards optimal solutions. However a system made out of many cooperating resources (and not all of them are reliable) is changing very frequently and computing the optimal solutions cannot be practical or even useful. So a few iterations of the system can lead to a refinement of the solution (a quasi-optimal one) that is good enough for the system and save processing time. A number of optimization algorithms and solutions have been studied in [VI, VII, VIII, XI].

Another important feature of autonomic systems is that they can contribute to make related systems more reliable. In [XII] , a situation in which a set of inner autonomics resources use their feature to "enforce" some levels of autonomics to unreliable resources has been proposed. A similar situation can be depicted as in **Figure 46**.

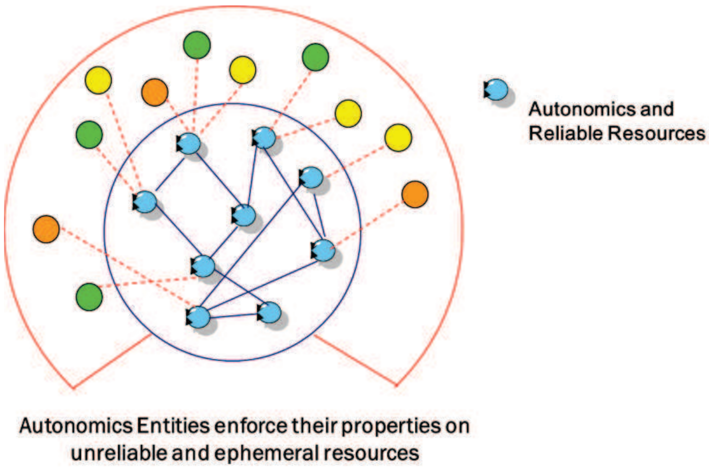


Figure 46: An Inner Circle of Reliable Resources.

The autonomic reliable resources can be used as a sort of reliable set of resources that unreliable resources (at the crown of the circle) can use to store, process and integrate data in an intelligent and robust way.

Figure 47 depicts three cases in which physical resources cooperate by means of gossiping protocols in order to ease the supervising and orchestration of resources for upper layers. A virtual resource can have autonomic properties (case a) and b)) or can be unreliable (case c)).

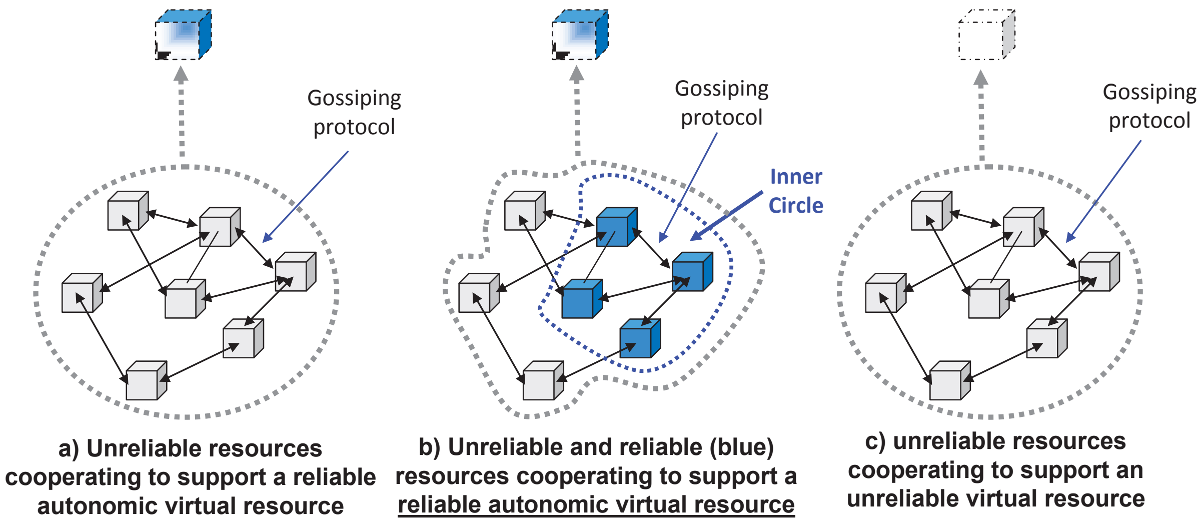


Figure 47: Creating Autonomic Virtual Resources

An Inner Circle could be composed by peers managed by different administrative domains, provided that they (demonstrate to or are certified to) offer suitable levels of reliability and trustiness. In a certain sense, the Inner Circle properties can be such that they can be inherited (or exported towards) external entities and environments (see **Figure 47** case b)). For example, the supervision features implemented on the peers of the Inner Circle can be used to export these properties towards applications which use also resources provided by other domains that do not show the same stability and reliability (e.g., nodes are executed in a highly unstable environment, peers are not completely reliable and trusted and their functions should be carefully used, and the like): they can be used to enforce reliability policies and to identify contingency plan in case of need. **Figure 48** represents the aggregation and orchestration of (virtual) resources in order to fulfill non-functional requirement at the level of virtualized autonomic overlay of resources. The Inner Circle concept is "recursive"; it can be adopted at several layers of the architecture. This property guarantees the openness of the architecture and the possibility to integrate different types of resources from different domains (de-perimeterization).

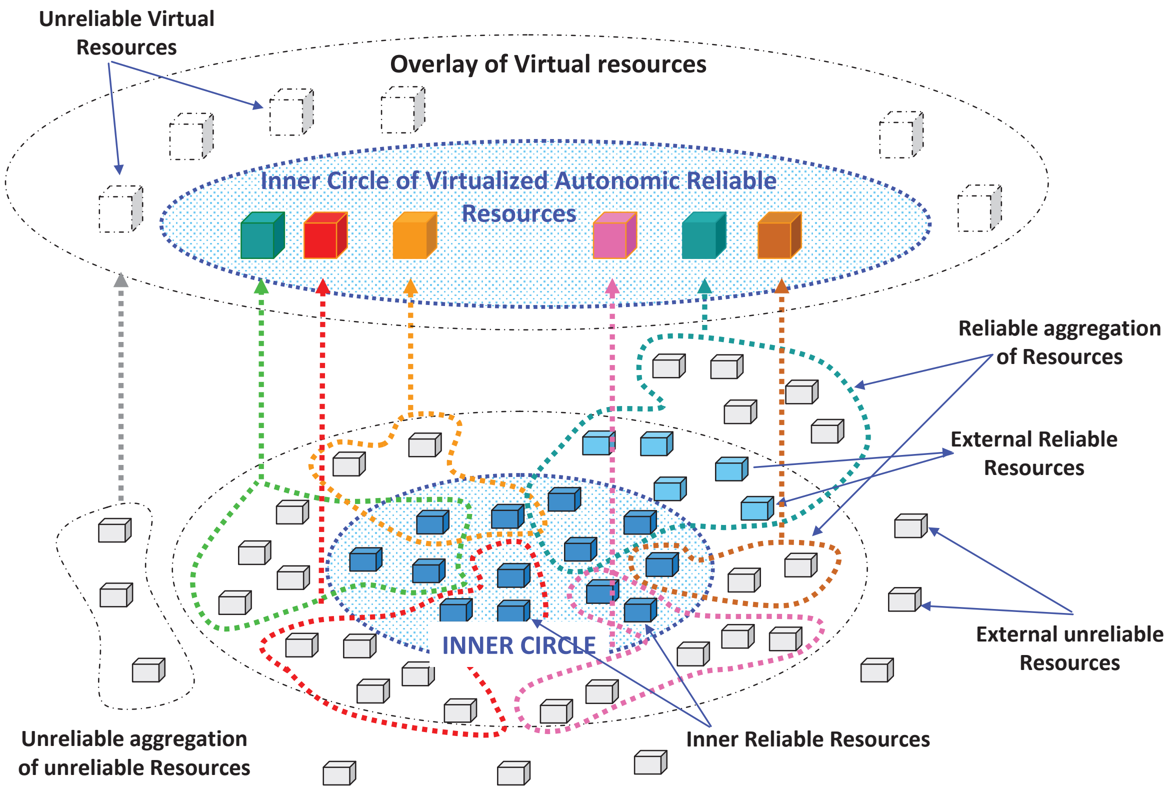


Figure 48: Inner Circle at the Physical and Resource Aware Levels

Operators and Players aiming at providing services can exploit the Inner Circle in order to allocate (virtual) resources for the creation of a reliable and trusted distributed execution and networking environment. They are specialized through the deployment of software modules for service logic, logic for controlling objects, and DB/repository schema, and interconnecting them according to the needed communication patterns. In particular, the resources can be aggregated according to a distributed architecture able to deliver applications according to the control paradigm most suited for each of them. This will also guarantee the availability of a minimal set of always available resources that will support a minimal set of vital functions for services and applications.

4. 6. 1. General Principles

The proposed architecture (**Figure 49**) relies on a small set of principles shaping the envisaged structure.

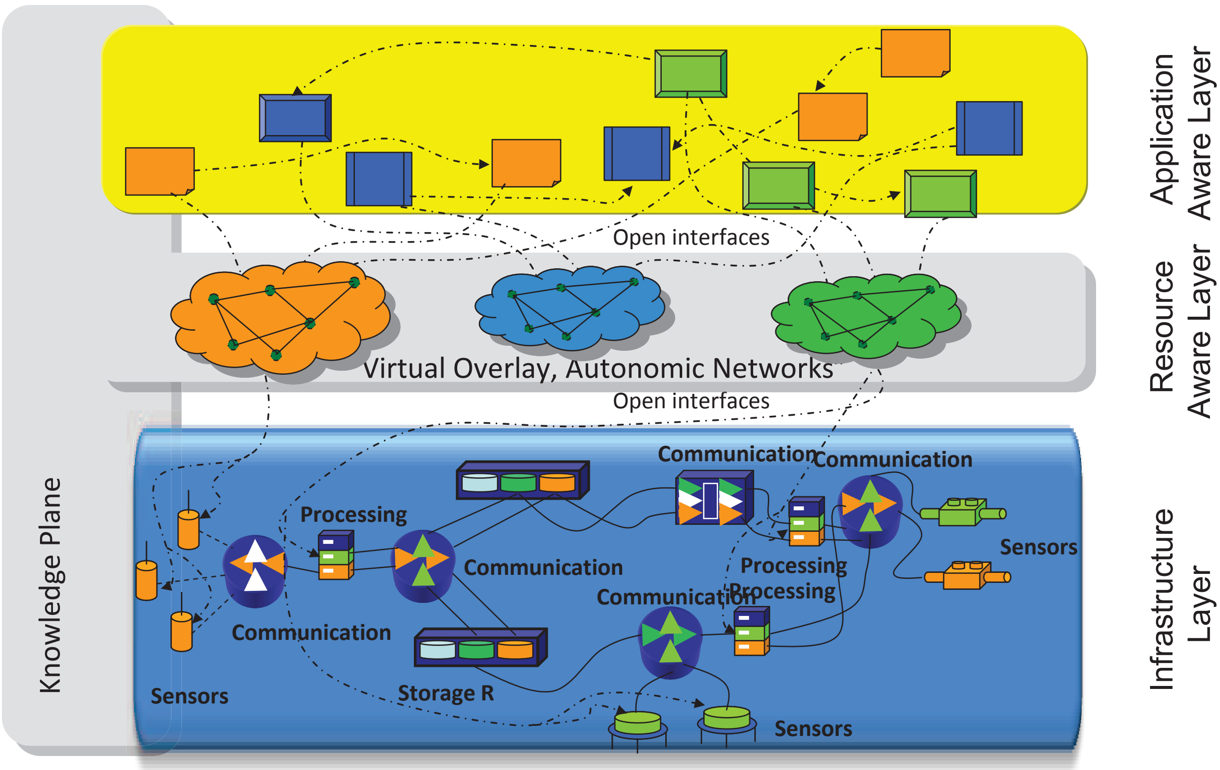


Figure 49:A System View of Cognitve Service Platform

Entities and related functions are organized in three layers: **Infrastructure Layer** comprises entities that provide a virtualized view for communication, storage, processing and sensing/actuation (things) resources; the **Resource Aware Layer** comprises entities and functions needed to supervise and optimize the usage of the virtual resources; **the Application Layer** is made out of applications that use the resources of the Infrastructure. All the entities in Cognitive Service Platform are represented as *resources*, i.e., entities which provide capabilities and can be allocated to/used by applications. Different types of resources can be envisaged. For instance, *infrastructural resources* provide features (implemented by physical entities, such as servers, communication network nodes, machines) to create distributed execution/networking environments for deploying and delivering services; *application resources* provide logical functions/data to be composed and aggregated for creating applications (or other composed application resources); they could rely on capabilities of infrastructural resources allocated to them. Capabilities of physical and logical resources are virtualized in order to ensure secure, personalized and isolated usage. Virtualization includes features for abstracting (e.g., simplifying, copying with heterogeneity) the interfaces to access/use resource capabilities, for sharing them (e.g., by partitioning their capabilities in isolated portions singularly allocable), and for providing a formalism for their description (e.g., to be used by allocation/negotiation functions). Each resource type can adopt specific *virtualization paradigm*. The behavior of the (virtualized) resources is enriched with *autonomic features*: self-awareness and self-organization capabilities (e.g., by adopting control-loops mechanisms) make the entire infrastructure more robust, controllable and resilient. Autonomic resources can self-adapt their behavior to achieve self-management, to react to internal/external events, and coordinate with other (nearby) resources. As demonstrated in (Jelasity, Montresor and Babaoglu 2005), a global behavior emerges from local decisions and interactions (a typical mechanism adopted for the governance of complex adaptive systems).

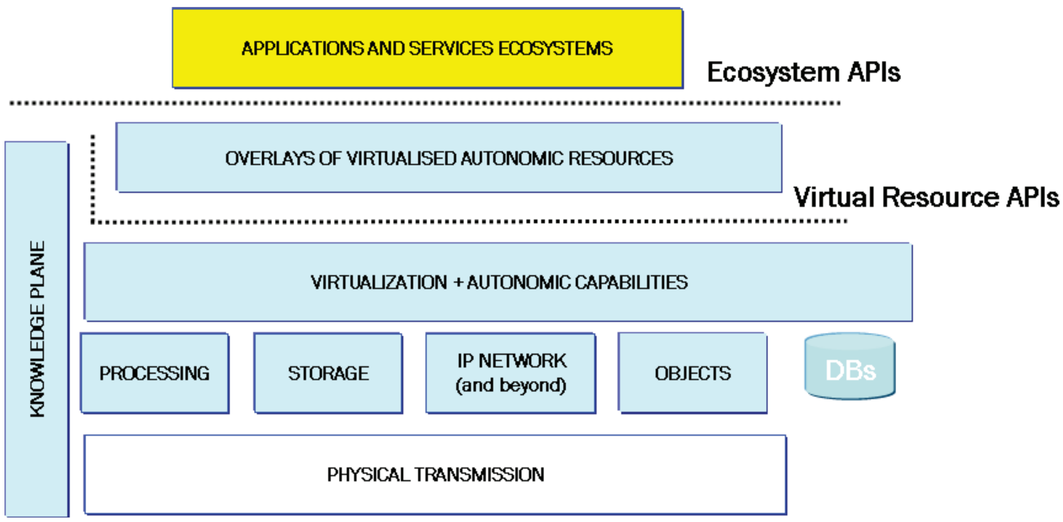


Figure 50: A Layered View of Cognitive Service Platform

The (virtualized) resources are clustered into *overlay networks*, used for the exchange of messages among the resources (which have to implement logic for joining, maintaining and possibly optimizing the overlays to which have to participate). Overlays implement several control and supervision algorithms, such as, resource discovery, resource allocation, load balancing, fault detection and recovery, dissemination of system information. The protocols adopted for their implementation are based on *gossiping*. They are based on iterative information exchange: during each protocol step, a node exchanges with (a small subset of) its neighbors in the overlay a small amount of data, and combine them to update its local state. In order to fully orchestrate a complex infrastructure that collects several networks and a multitude of end points, cognitive approach to infrastructure governance is undertaken. A Knowledge Plane (KP) is introduced aiming at handling the available cross-layer knowledge gathered from all the protocols being used and from all the resources (and their virtualized representation). The KP is fundamental for effectively create virtualized systems that fully satisfy Users and applications requirements. These principles are embodied in the layered architecture depicted in **Figure 50**. On top of this infrastructure, applications and services can be built and combined by aggregating the virtual resources and extending them by introducing new resources. This enable the creation of an ecosystem where several players (e.g., Individuals, Enterprises, Service and Network Providers) can interact in an open and collaborative way to produce/consume/ sell/buy/trade data, contents functions, and applications.

4. 7. Disruptive Approaches

There are plenty of activities aiming at re-defining and implementing new solutions for the network architecture. Here some of them are sketched out to give a glimpse of the underlying complexity and richness of research around the theme of new networks.

4. 7. 1. Local Area Networks

They are important because there is a continuous increase usage of this technology. For instance as stated in GIGAOM[21] the usage of WiFi networks has exceed the usage of cellular nets. Many Operators are trying to integrate WiFi Networks with cellular ones in order to off-load traffic from the mobile infrastructure to the more locally capable ones. Some fixed Operators (e.g., BT) have tried for long to use WiFi technologies in order to deploy wireless networks in alternative to mobile infrastructure. Another interesting initiative is FON[22] that

[21] http://gigaom.com/2012/07/10/we-already-use-wi-fi-more-than-cellular-why-not-continue-the-trend/ last accessed April 10th, 2013

[22] http://corp.fon.com/ last accessed April 10th, 2013

tries to share part of the wireless bandwidth of a single user with members of a community (in this case the FON network). Behind these initiatives there are the studies and the experimentations aiming at creating and supporting mesh networking. These initiatives are based on the possibility of dealing with access point as resources that can be hacked and programmed. In fact there are some distribution of software (e.g., openWRT (Fainelli 2008) and DD-WRT (Weiss 2006)) that can be deployed on existing access points making them programmable and configurable to the needs of the users. For instance FON is based on openWRT. Along this mesh networking studies, the need to optimize the routing between mesh nodes has brought to develop new routing mechanisms capable of coping with the dynamicity of mesh systems. One interesting protocol (that anticipate somehow the needs of the mentioned autonomic systems) is the B.A.T.M.A.N. protocol (Johnson, Ntlatlapa and Aichele 2008) that uses a sort of gossiping mechanism in order to create routing paths within mesh networks. Another interesting experiment is Netsukuku (Lo Pumo 2007) that aims at create a network without any central authority and not controlled by an ISP. These two experiments are particularly interesting because they have leveraged the availability of open source for access points in order to create new communities networks that have a great disruptive potential. Studies in WiFi and local area networks are then very important for radical alternative to public and strictly controlled networks.

4. 7. 2. Community Networks

"Home with Tails" (Slater and Wu 2009) and the seminal work of (St Arnaud, Wu and Kalali 2003) and more recently on Free Fiber to the Homes[23] propose a very disruptive model for the access network: the user is the owner of the fiber connecting the home to the cabinet (and maybe even to the central switch). This simple proposition is strongly user centric (and hence the interest in it) and it has a very high disruptive potential: people can manage their connectivity and can create communities to share the communication capabilities. Within large condominium, the infrastructure is owned by the community and internal communication can be totally free of charge. Communities can even create small data centers in order to support the processing and storage needs of the users.

An interesting and related concept is the one of Open Access Network (Battiti, et al. 2003), i.e., a network organization that clearly separate the ownership of the infrastructure

[23] Available at http://free-fiber-to-the-home.blogspot.it/ last accessed April 9th 2013

from the service provision. An example of this type of network is Guifi.net. As stated in the web site[24]: "*guifi.net is a telecommunications network, is open, free and neutral because is built through a peer to peer agreement where everyone can join the network by providing his connection, and therefore, extending the network and gaining connectivity to all*".

Another interesting experiment and community is Village TelCo (Adeyeye and Gardner-Stephen 2011). It is based on a mesh network made up of Wi-Fi mini-routers combined with an analogue telephone adaptor (aka 'Mesh Potato' (Rowe 2009)). This infrastructure connects the village to a SIP based server that can be seen as a gateways towards other networks. There is the possibility to use FreeSwitch[25], i.e., an open source implementation of a Softswitch, for controlling the connectivity of the local set of SIP phones and the interworking with other networks. In principle these type of implementation could even make use of openBTS, i.e., an open source implementation of a GSM base station (Burgess, Samra and others 2008) for creating an open GSM network.

4. 7. 3. Military Communication (MilCom)

It comprises several projects aiming at the improvement of communication technologies for military usage. For instance Direcnet[26] is an attempt to build a military communication architecture based on open standards and existing technologies. It is promoting the integration and exploitation of self-configuration with a mesh, sensor, peer to peer, cognitive radio networking technologies. This infrastructure creates performing networks able to provide considerable bandwidth to the single soldier (DirecNet) especially if the number of nodes per area increases. Other projects aim at ensure communication and data consistency in several MANET (e.g., (Bastos and Wietgrefe 2010)).

These solutions are for military usage now, but as stated for Direcnet they have the objectives of being open and using COTS systems, they could reach also civil usage in the near future because the envisaged technologies are already available and they just need to be standardized. Using this technologies with a user centric approach, i.e., giving the power to

[24] Available http://guifi.net/en/what_is_guifinet last accessed April 9th 2013

[25] Available at http://www.freeswitch.org/ last accessed April 9th 2013

[26] Available at https://www.opengroup.us/direcnet/documents/16322/2012_MILCOM_DirecNet_Technical_Panel.pdf last accessed April 9th 2013

the final users, could allow the creation of very powerful networks capable of competing in terms of features and capability with the traditional TelCo infrastructures..

Direct Communication has the potential to substitute, for some applications, traditional networks exploiting the increasing processing and communication capabilities of new terminals.

4. 8. Findings

This Chapter has investigated the possible evolutions of the network. A first remark is that the Internet will go through a substantial change in order to correct and limit the ossification issue. This will bring into the network a lot of interesting functionalities and capabilities like more cooperative approach between different subsystems and a progressive adoption of virtualization capabilities. In addition the gap between the network and the application will be closed by means of Application Programming Interfaces. This is a long pursued goal but in the future it will be finally reached guaranteeing to application the possibility to adapt the network to the real communications needs of services. Actually the network itself will not anymore comprise only communications resources but it will encompass processing, storage and progressively also sensing and actuation elements.

Information Centric Networking (even if the hype of this approach is over) point to a big problem to consider by the TelCos: the usage and transport of data is prevalent with respect to the human to human communication. This requires specific attention in order to tailor the network to this new prevailing goal. The point of supporting a better “information” retrieval within the network is still a priority in the ICT industry and the companies that will be capable of accommodating in an effective way this task will have a competitive advantage. Virtualization of specialized networks (for instance virtualized content delivery networks) for supporting a specific task could result in new business opportunities.

Software Defined Networking is a means to control and optimize the usage of available resources. Virtualization is a means to run logical networks on top of a physical infrastructure made out of general purpose machines. Both technologies have a high disruption potential, but when integrated their capability to change the status quo further increases. Some TelCos are reluctant to fully exploit the disruptive potential. In fact the combination of the two technologies will be used by TelCos in two different fashions (this is drawn on discussions with different TelCos and the analysis of a few proposals of European Projects[27]):

- The most innovative TelCos will try to use SDN (and the related Virtualization techniques) in order to create different networks and to adopt new control paradigms as well as cooperation model. The first implementations will be deployed in confined areas (e.g., small country in which the Operator is active and it is deploying a new infrastructure). This approach will potentially have an impact on existing business model and established relationships between Operators (see for instance **Figure 43**).
- Traditional Operators that see the SDN and the virtualization capabilities as a means to reduce the investments on new infrastructure. In this case the disruption that SDN is bringing is limited. Also fear the SDN and Virtualization can lowering the entrance barriers to such a level that many new Actors can afford the deployment of a network infrastructure. Consequently they will act in such a way to limit the applicability of SDN and Virtualization.

A byproduct of SDN and Virtualization is a possible change in the Telecommunications equipment market. In fact the value moves to software and programmability. New Vendors (e.g., coming from the IT market) can offer appealing propositions displacing the current incumbent Vendors. Established Vendors in the realm of networking as CISCO and Juniper can attempt to mitigate the impact of SDN and Virtualization by exposing and opining up a few APIs in their systems. Traditional telecoms Vendors (e.g., Alcatel – Lucent, Nokia Siemens Networks, Ericsson and others) could try to preserve the value of their solutions by adopting a similar approach. IT Vendors (like HP and IBM) can offer a better proposition because they do not need to preserve existing product lines. In addition a market of new Vendors can emerge (e.g., VMware with the acquisition of Necira could play an important role in this). In other terms, a disruptive implementation for SDN and Virtualization can leas in a few years to a restructuring of the TelCos' ecosystems introducing new Vendors and transforming existing ones.

From the point of view of a TelCo, the combination of SDN and Virtualization poses the problem of Mastering of software. It refers to the possibility to directly programming the infrastructure in order to differentiate from the other competitors. The capability to deal with successful software project is questionable within TelCos, so they should make a relevant

[27] This information cannot be openly disclosed

effort in converting their IT department into efficient software factories capable of creating new infrastructural solutions as well as new services and application on top of them. This transformation requires skills and attitude that not all the TelCos have or are ready to use. This issue will strongly influence the possibilities of TelCos to play different roles in the service provision. Section 8. 8. and Section 8. 9. describe respectively the TelCos as a Platform Provider scenario and the TelCo as a Service Provider scenario. They exploit the technologies discussed in this Chapter in order to support higher (compared to today situation) level of resources and network programmability. Programmability is a requirement for playing the Service and Platform Provider roles.

Another result of this Chapter is the emphasis put on the need to create a relation between the complexity of the networked environment and the need to introduce cognitive behaviors in the networks. As said, networks will encompass different functionalities (processing, storage and sensing besides communications). Networks will became richer in functionalities, but they will be also more complex in terms of allocation and usage of resources and functionalities. In order to cope with this complexity, autonomic and cognitive solutions will be progressively introduced in the network environments so that resources will contribute to the self-organization of the infrastructure. Even if core networks will became flatter and functionally simpler, there will be an increased need to integrate other type of resources and to integrate them even if they pertain to different administrative domains. The infrastructure should be capable of integrating into a virtualized logical network the resources of the TelCo as well as the resources of different Enterprises that will dynamically join in order to meet a business goal. Self-organization, federation and integration of different types of resources will be characterizing features of many scenarios of Chapter 8.

Chapter 5. Emerging Service and Applications Scenarios

Highlights

- From Cloud Computing to XaaS
- Bank of User Data
- Smart Environments
- Internet with Things

5. 1. Introduction

The Technological evolution described in previous chapters will lead to new opportunities in terms of services. These opportunities are related to new classes of services that could not be proposed in the past due to technological limitations (e.g., augmented and virtual reality), to current services like the Cloud, And to existing classes of services (like communications services that could be totally reshaped by the new technology). This chapter does not aim at covering in an exhaustive manner the possibilities, it tries to evaluate a few interesting classes of services that could have a value from a TelCo perspective (see for instance Section 6. 5. in which some lines of action for the Operators are presented) the impact that new technologies could have on existing services, the possible evolution of existing services and the introduction of new classes of services. One feature of these classes of services is that they have a potential large market that could help the TelCo in finding alternative markets to its traditional communication service market. Still it should be clear that there are also other possibilities like multimedia services and evolution of IPTV, or Virtual and Augmented reality applications, or even e-government related applications. All of them are large potential market and they deserve a careful attention. The classes of services discussed in this chapter are those under close evaluation of a number of Operators (e.g., Telecom Italia and Telefonica together with MIT for personal data), the evolution of machine to machine towards Internet of

Things undergoing within numerous Operators. The cloud represent a specific example: many Operators need to correctly position their offering in this area in order to understand whether the network and ICT assets of a TelCo can be exploited. Smart Environments is instead an initiative undergoing within Telecom Italia whose goal is to understand how edge environments (and direct to direct communication) could be integrated within the TelCo Operator with an open and possibly user center approach.

The objectives of this chapter could be summarized as follows:

- To present some new classes of services that can provide new opportunities to several Actors (the analysis is focused on the TelCos)
- To highlight the way these services can be implemented and the merit that interaction and control paradigms for supporting them could have (and vice versa the difficulties that can be met in trying to implement services choosing a paradigm "a priori").
- To identify if possible new user centric approaches and their viability in the provision of new services. These aspects are particular relevant in the personal data management that is considered by many Actors as one new important source of revenue and a sort of test-bed to establish new forms of relationships with Users.

The chapter focuses on the opportunities and the challenges that Cloud Computing is posing to TelCos. This is a main trend, in fact many TelCos want to pursue this business in order to increase their value proposition to customers. The Chapter will indeed analyze what conditions and what approaches are indicated in order to take advantage from these technologies. In addition some considerations on the Cloud Ecosystems are presented.

In this section a few Service and Application Scenarios are presented in order to frame the possible technological evolution in viable deployments or use cases. This section is not necessarily advocating the merits of a service or class of them from the specific point of view of a stakeholder (e.g., a Web Company, and Operator, …), it offers mainly a view on what could be done at the technological level. Whenever it is possible, the chapter emphasizes a user centered proposition, i.e., a proposal that puts the users and their rights at the center more that the specific economic return of a stakeholder. However, some examples could have a TelCo oriented approach as a means to identify some viable services and applications that can bring value to the customers and can help the Operators to find different roles and business models in the ICT markets. Service scenarios do not aim at being exhaustive nor complete, they just cover a set of areas that promise interesting developments and give an idea of how solutions could be forced by specific business model or approaches.

TelCos are and will be interested in communications services, so they will pursue the goal of creating new communications capabilities. There are several attempts going on to exploit existing means in order to increase the appeal and the number of functionalities offered by communication services. The term Rich Communication Service refers to the possibility to aggregate around the "data transport" capability a set of appealing functions like VoIP, Presence, Instant Messaging and the like. This is another example of the general category of Unified Communications. The TelCos are striving to revive this market and they are trying to exploit existing architectures (like IMS with the Rich Communications Suite) or new ones (like the rising WebRTC solutions).

Another recurrent intention is to enter into the Cloud Computing market offering several XaaS possibilities. One of the assets to capitalize, in this case, seems to be the possibility to locally serve and support the customers. In addition TelCos are interested in offering federated cloud functionalities in order to integrate customer solutions in several different countries. There is a need to understand whether these proposals are interesting for the user and how they can be differentiated and improved with respect to offers made by WebCos. In addition there is the issue related to the technological gap between WebCos and the other possible players of the Cloud. In this case completion has to be played at the technological and business level in a very aggressive fashion.

On the new services side, there are some opportunities that the TelCo may undertake. The Bank of User Data is one of them: personal data will be collected according to policies and mechanisms established by the user and the ownership of data remains of the user. The TelCo can play the role of Storage Provider and possible the one of Broker of data if the user enables this possibility. Plenty of services could be created and offered on personal data infrastructures. Another example are the Smart Environments, i.e., intelligent systems that can adapt their resources in order to support the communication, processing and storage needs of several users in a dynamic way. Finally another opportunity is presented: the Internet with Things, i.e., the capability to create a networked environment in which any physical resource can be virtualized in the clouds creating a sort of relationship between the atoms (of the physical resource) and the bits (of the virtualized one). Actions on the physical or virtual world have an impact also in the “other world”. This approach is very different from the Machine to Machine, M2M, one. Today TelCos are using M2M platform to earn from connectivity from distributed devices and from management services (dynamic activation and deactivation of SIM cards). The level of functionalities provided is very limited and with the increase in pervasiveness of other solutions this opportunity can vanish if it is not supported by a richer set of functionalities. Internet with Things seems to be a vast domain into which to offer new compelling services.

Determining the opportunities of providing new services is a means to understand if there are significant business returns and a means to understand how much to invest into service architectures. One of the issues limiting investments in horizontal platforms for instance is the need to justify the platform costs with respect to vertical solutions on a service by service basis. Horizontal platforms make sense if and when the number of services that can be potentially offered is huge and these services are easily supported on a delivery platform. These classes of services are promising also from this point of view.

5. 2. New Communication Services

Since the inception of Voice over IP solutions (e.g., Voicetec in the ninety) the world of telecommunications has seen these technologies with suspicion and fear. It was clear that they were capable of changing the way the major services of Telecommunications networks were provided to customers and were paid for.

Even new technologies are often modified and conducted to a specific model: the Network Intelligence. This is a constant attitude of many Operators that try to smooth the disruption of new technologies with respect to communication services. There is a long tradition in this approach.

The major attempts to deal with Voice over IP, VoIP, technologies were to frame them in the usual architectural vision represented in **Figure 51** (Vemuri 2000).

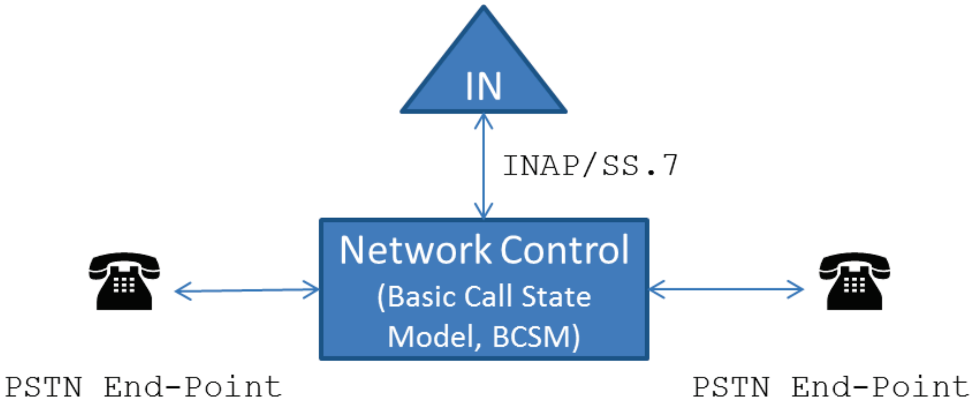


Figure 51: The Network Intelligence Paradigm

Services are value added functions provided by specialized Nodes (in this case Intelligent Network Controllers) and service functionalities are specified and provided according to a standardized Finite State Machine for representing call progress.

For instance, the H.323 protocol stack was designed to be compatible and to be used as a mere evolution of the voice architecture. Actually H.323 protocol stack is an adaptation of the Telephony signaling protocols over an IP signaling infrastructure. The paradigm behind it is essentially the same traditional one represented in **Figure 52**.

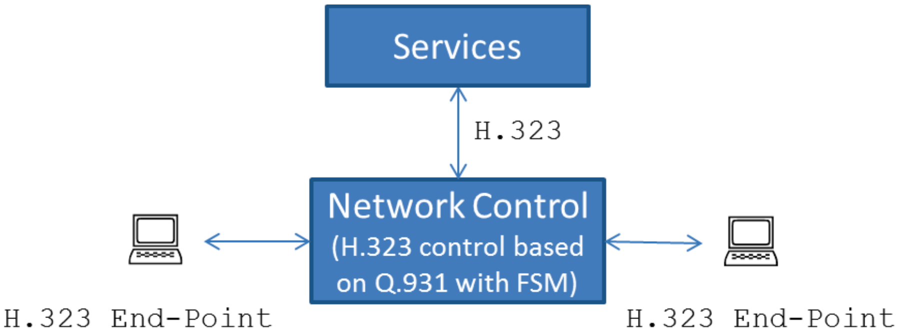


Figure 52: The H.323 Control Architecture.

Even the Session Initiation Protocol, SIP (Rosenberg, et al. 2002), has been modified and changed within 3GPP in order to better fit into the paradigm of the Network Intelligence. Actually H.323 (Thom 1996) and SIP (Schulzrinne and Rosenberg 1998) have been considered as the new signaling system, i.e. a means for end points to ask to the network how to orchestrate resources in order to put in place a "call" or in more modern terms a "session". Under this respect, call and session are synonymous. They refer to the fact that endpoints and network functions/resources are correlated and associated with the goal to create a suitable communication environment for the time requested by the participants". H.323 was already a stateful protocol with its own call model. SIP can be stateful and stateless, i.e., its servers could keep track of the interactions with end points and behave according to a complex session state machine. This approach has been pursued in the 3GPP specifications leading to some changes in the SIP definition for adapting this protocol to the "network intelligence" paradigm, i.e., services are to be provided within the context of a session, they are related to the allocation of some network resources, service logic is triggered by the session control function when it has the need to do so. The result of this redefinition of the SIP protocol is showed in **Figure 53**.

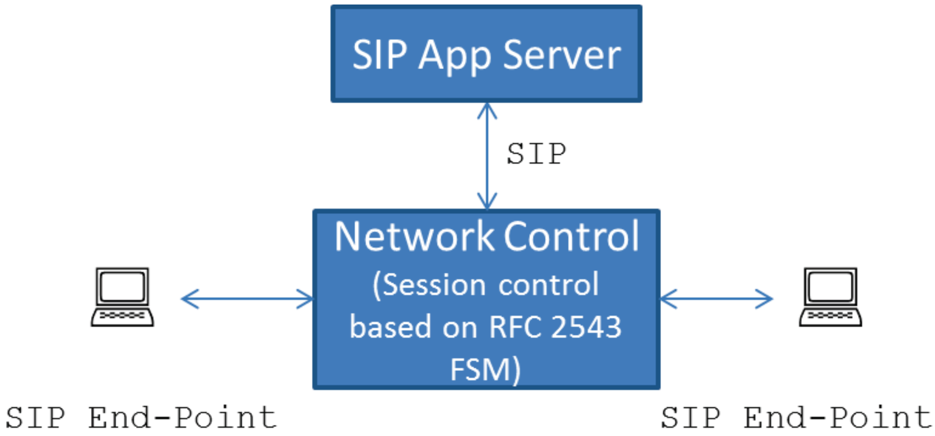


Figure 53: The Network Intelligence Approach to SIP.

This attitude and philosophy has left room for others to innovate and adopt more distributed approaches. Skype is paradigmatic from this point of view and it shows how P2P mechanisms can support large scale communication systems. It is based on a hybrid and structured peer to peer approach in which the end to end communication takes place between the peer nodes, while a part of the signaling and the organization for the overlay network relays on the functions offered by "super-nodes" and the identification, authorization and billing functionalities are centralized into a few servers directly under the control of Skype (**Figure 54**). The architectural choices are quite smart, the valuable and important functions from a business perspective (Identity, authorization and billing) are kept in a centralized and well controlled infrastructure. Other control functions (like addressing, routing, grouping of peers, their organization and "keepalive") are demanded to a set of supernodes (i.e., trusted node collocated in "friendly" locations). The peers have the control of local linkage to supernodes and the end to end communication with other peers.

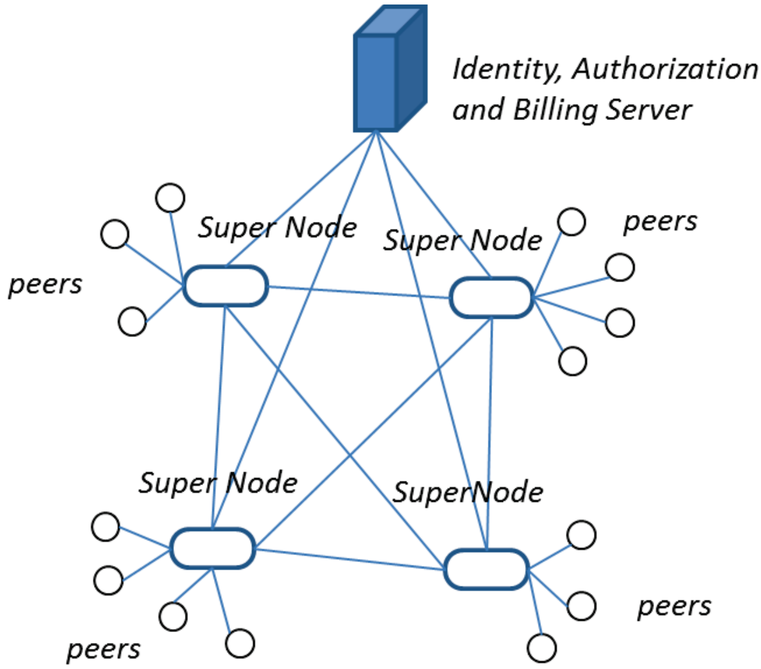


Figure 54: The Skype Architecture

The interesting fact is that one of the fathers of the SIP protocol had to go through a reverse engineering study in order to determine the structure of Skype (Baset and Schulzrinne 2006). The work was aiming at determining the reasons of the success of the P2P solution. Contextually in the same period the SIP community was trying to propose the usage of SIP as a viable protocol to create P2P systems (Singh 2005). The Skype implementation is based on: the consolidated Joltid[28] platform used in Kazaa[29] for supporting a large scale P2P system, the economic implementation, the quality of codec used. All these features were missing at the time to the SIP protocol (that once again was used in a very traditional manner). Currently Skype is dominating the VoIP market and has reached a rewarding position as a collector of international traffic (around 34% of international traffic with a record of over 50 million concurrent users according to (Skype, Wikipedia 2013)) and it has added over time many interesting features like presence, video call, video conferencing and desktop sharing.

Operators are still striving in determining how to regain a grip on users. There is an effort in positioning the Rich Communication Suite, RCS (Henry, Liu and Pasquereau 2009), (Lin and Arenzana Arias 2011), as designed and put together by the GSM Association (GSM Association 2009). The attempt is to provide an interoperable solution between several

[28] http://joltid.com/ last accessed May 30th 2013
[29] http://en.wikipedia.org/wiki/Kazaa last accessed May 30th 2013

Operators able to support presence, instant messaging, synchronization with address book and other value added functions by means of APIs and using the IMS infrastructure. The solution targets smart phone with the intent to provide services to advanced users. The issue behind RCS are several: the architecture is not prove itself so flexible and the issues in interoperability between different IMS systems are causing interoperability problems (that major Web Companies do not have); services are not differentiating from those offered already by WebCos and, in addition, they suffer from a bit of invasiveness (e.g., the presence feature was originally offered in such a way to violate the users privacy); the service is not based on a clear definition of a community (like Facebook) and hence they are very similar to the "old way" of making phone calls. Simply put: from a user centric view, the RCS does not offer any novelty neither from the service functionalities point of view neither from the business perspective (it is using the existing billing models) nor from the platform one (the IMS and service exposure combination). In addition, also the Phone makers are "tepid" in the initiative.

In more recent time, there is interest in the TelCos community for the applicability of the WebRTC (Loreto and Romano 2012) solutions for supporting the real time communication between browsers. WebRTC is under definition by the World Wide Web Consortium, W3C. The specification is based on an API and supporting communication protocols enabling "browser to browser" applications for voice call, video chat and P2P file sharing without plugins. In a way, WebRTC expands the capability of the WebSockets, i.e. mechanisms that are supported in HTML5 to allow the bidirectional communication between a browser and a WebServer avoiding the continuous polling. The envisaged use of WebRTC is by means of a Server that sits in between the peers for helping in establishing the communication, once this goal has been achieved, the peers (the browsers) can directly communicate for exchanging video, audio and data. **Figure 55** shows the described mechanisms. It should be noted that the definition of WebRTC supports also the possibility to have a direct signaling between the web browsers (a real P2P solution), however many actors like Web Companies and Operators will find more convenient to have a brokering function between different peers. Obviously this is a replication of existing models reproduced in order to have a controlling position over users.

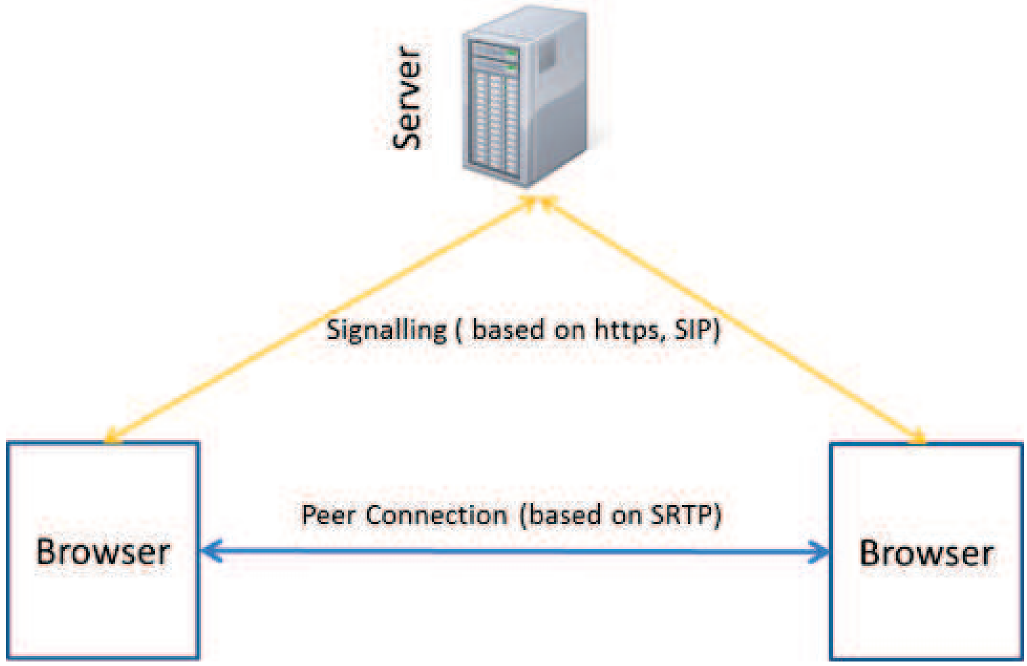


Figure 55: The WebRTC Architecture

From the programmability side, the WebRTC solution is interesting, in fact if offers an API integrated within the browser so that applications can use it to control the setting up of the communication between the peers (with the support of a server) and the transport of data, video and audio between peers. **Figure 56** shows the software architecture within a browser.

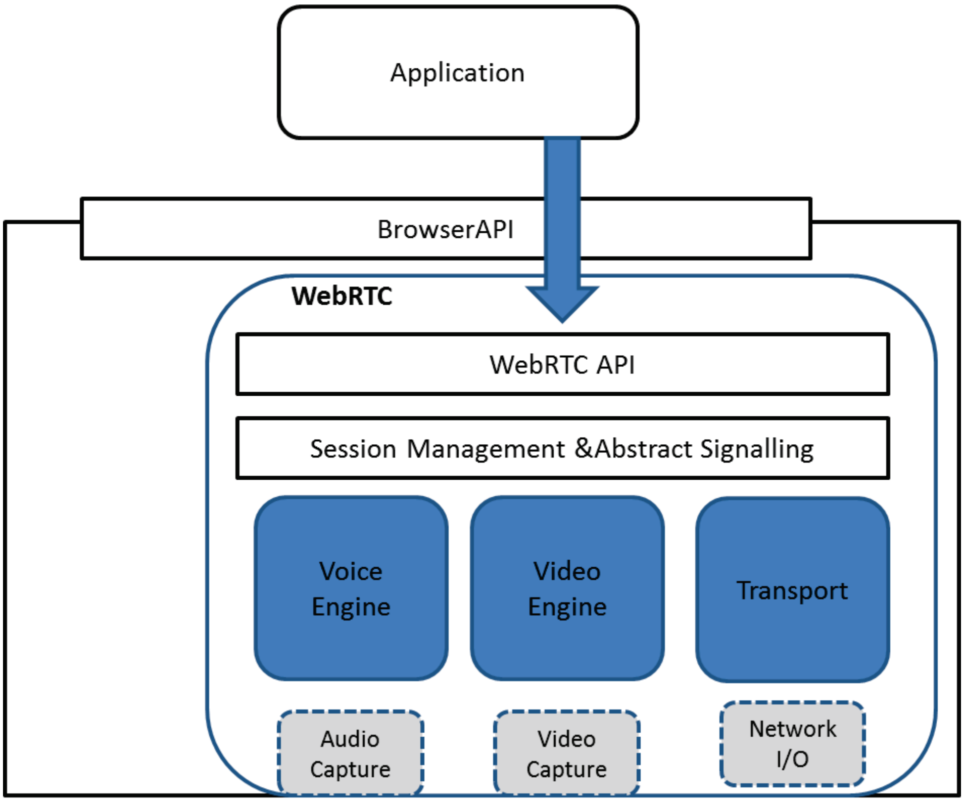


Figure 56: The WebRTC Software Architecture within a Browser

It is pretty obvious to think to extension of this architecture in order to make it highly distributed and consistent with a structured P2P network as Chord. In this case a Distributed Hash Table can be used to find the address of a specific user and then to connect directly without the need to have intermediation by a server (Werner 2013).

Figure 57 shows the possibility to use a DHT (like openDHT and its interfaces (Rhea, et al. 2005)) for storing and retrieving information needed to connect to a specific peer.

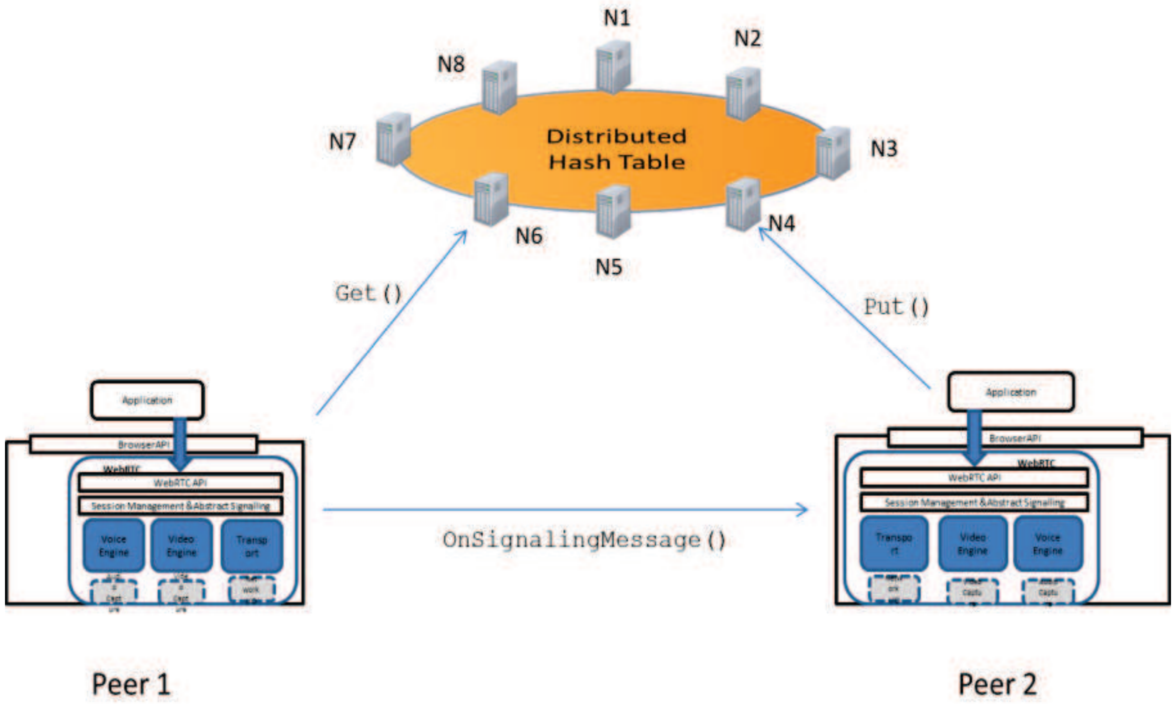


Figure 57: Using WebRTC in a DHT Context

In this example, `Peer 2` stores its information in the DHT (Put() operation). `Peer 1` retrieves the information from the DHT (Get() operation) and then sends an invitation to the other peer for establishing the connection. This example, very far from being complete, just shows a possible trend that many web programmers could endeavor aiming at a more user centric approach. A few considerations emerge from the analysis of the approach of TelCos to new communications services:

- P2P paradigm is capable to compete with the more traditional Network intelligence one as shown by Skype
- WebRTC can be easily used in a P2P fashion, leaving out the Operators from brower to browser voice communications

- The application of the Network Intelligence approach do not give advantages to user in terms of service functionalities, it just help the TelCo to keep a grip on customers.

5. 3. From Cloud to Everything as a Services (XaaS)

This section analyses the Cloud technologies and its service opportunities mainly from a TelCo perspective in order to understand whether the Operators have a chance to introduce innovation or disruption in a market that (at least for the residential customers) is in the hands of WebCos. Many TelCos see the Cloud and its evolution towards the approach Everything as a Service (XaaS) as an opportunity to enter into adjacent markets especially when the Cloud is associated with the network asset (an advertisement from Telecom Italia says: "the Cloud with the network inside"). This approach is quiet controversial because (as seen in Section 3. 3.) leading WebCos have a considerable technological advantage over competition and there are doubts that leveraging the network asset will give to TelCos an advantage. In these section an analysis of what TelCos could really do and achieve in this technological and service realm is presented. One real point, as discussed in Section 6. 4. clearly show the benefits that the Operator can offer to customers by means of the "cloud + network" proposal.

5. 3. 1. Cloud Computing Services from a TelCo Perspective

As seen the cloud computing technological and market scenarios are largely dominated by Web and Information Technology companies. The technological pace is determined by needs and solutions stemming from the web companies that were able to create walled gardens with proprietary technologies. Each major web player is also able to directly and autonomously develop and master its own specific solution. From this perspective the technical gap between the Web and the Telecom industries is striking and probably insurmountable. In addition, major web companies have a significant footprint in the provision of services to residential users and their services are deperimeterized (i.e., they can be accessed independently from an owned network, see Section 6. 5.).

Competition under these circumstances is hard especially if TelCos are continuously playing the role of "intelligent buyers" and do deliver services on platforms developed by IT or Telecommunications companies.

In order to improve this situation, TelCos should change the rules of the game at the technological and at the business level.

In the following sections a divergent perspective on Cloud Computing for TelCos is presented and discussed. It is based on the assumptions that market and customer differentiation is a premium, that there is not only a "pay per use" business model behind cloud computing, that residential and business market segments have different needs and expectations. From a technical perspective the integration of connectivity within private and federated clouds could be a key element for bringing cloud solutions to enterprises, that network programmability is a means to support enterprise requirements and a pass thru for delivering better distributed services that can support consistency of data when customers require such a feature.

Approaching the Cloud Computing in this different way means also to have another view on the taxonomy of NIST (Liu, et al. 2011). In fact new dimensions and aspects of the Cloud Computing proposition could be introduced. **Figure 58** illustrate a few new viewpoints that an Operator must consider in entering in the cloud competition.

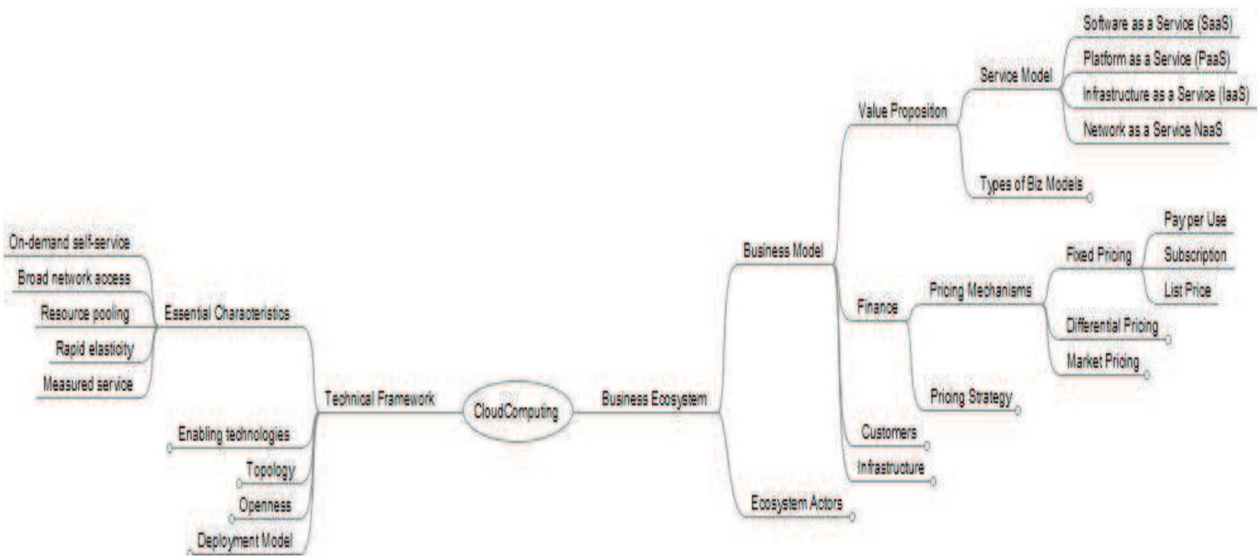


Figure 58: New Viewpoints on the Cloud Taxonomy

From a business perspective some new considerations for the cloud market are related to the value proposition to associate to the Service Models, some considerations related to pricing mechanisms and even more important to pricing strategies. A better definition of target customers could help in tailoring solutions that fit the expectations of the users. The business infrastructure identifies the mechanisms and the processes that a TelCo can leverage to pursue a cloud related offering. And finally the Actors of the ecosystems are those stakeholders that have relationship with the TelCo and can help or contribute or have to be involved in order to make a viable business.

On the technical side, new dimensions are related to enabling technologies to be used for a viable implementation of a cloud infrastructure able to support the business objectives. Topology deals with the organization and the kind of infrastructure to be controlled. Openness point to a major feature of the platform: the ability of the proposed platform to be extensible and flexible in such a way to extend and improve the functionalities and services provided over time. The deployment model extend a bit the one proposed by NIST and tries to figure out some viable and meaningful deployment from the TelCo perspective.

5. 3. 2. On the Cloud Business Ecosystem

Two aspects of the Business Ecosystem will be briefly sketched in this section: the aggregated actors and the business models and opportunities reasonably pursuable by a TelCo. They strongly characterize the possibilities and the approaches that a TelCo can attempt.

The Ecosystems Actors

The ecosystem of a cloud offering (from a TelCo perspective) is quite complex because TelCos do need to have a direct link with customers and because the construction of the TelCo Cloud requires a lot of links with other stakeholders. In addition TelCos are not relying on a "make", but on a "buy" approach and then the construction phase of a cloud solution is made also of relationships and integration with other entities. TelCos have to seek cooperation of a large number of other stakeholders in order to put in place a cloud offering. **Figure 59** (bottom links) depicts a possible set of stakeholders for a cloud platform.

Central to the approach is a clear and valuable relation with the users. It is mandatory to be able to have a different approach to customers compared to the one established by web companies: in this case there is the need to have a direct communication with customers in order to support them, to integrate their systems and to fulfill their requirements by means of a day by day cooperation. Clients in this case are also Enterprises that seek a greater integration of their private platforms into a cloud. Communities (in a large sense) are also important in order to grasp requirements, to promote the solution and the functionalities to a large audience, and for extending and tuning the offered capabilities (a sort of beta test). From a development point of view, the internal IT and Network organizations of a TelCo have to cooperate in order to design and agree the best specification for a cloud platform, they should also cooperate in order to define a target list of services and the conditions for integrating the platform and its systems into the TelCos processes and workflow. Other actors involved in the definition, design and implementation of the platform are technologies vendors, developers and integrators. Here the goal is to avoid as much as possible a lock in situation in which the TelCo is forced to follow the design and product evolution decisions of a specific vendors. In such a competitive market (in which the TelCo is not the primary choice for many customers) flexibility and readiness to modify and extend the capabilities is of paramount importance. Advisors and consultancies agents should cooperate in this phases in order to advice on trends and best practices of the industry. From a more commercial point of view, Resellers and even other TelCos can be useful to enlarge the potential market of the cloud platform in order to exceed the rigid boundaries determined by the need to deploy networks in specific geographic areas. Government and Regulation Authorities have a role in governing the possibilities and the limits of TelCos in this market. Governments can also be seen as potential customers for usage of cloud solutions in many situations.

The current business model of many WebCo is the one of a Walled Garden. TelCos have to nurture new ecosystems in the field of cloud computing allowing a more open model for application development and for the integration within customers systems. Interoperability between different systems and environments and new interfaces are important in order to catalyze new developments that are portable over different cloud computing platforms. The expected contribution from enterprises and developers is the creation of a federated environment that is open and can exploit and leverage the contribution of each stakeholder.

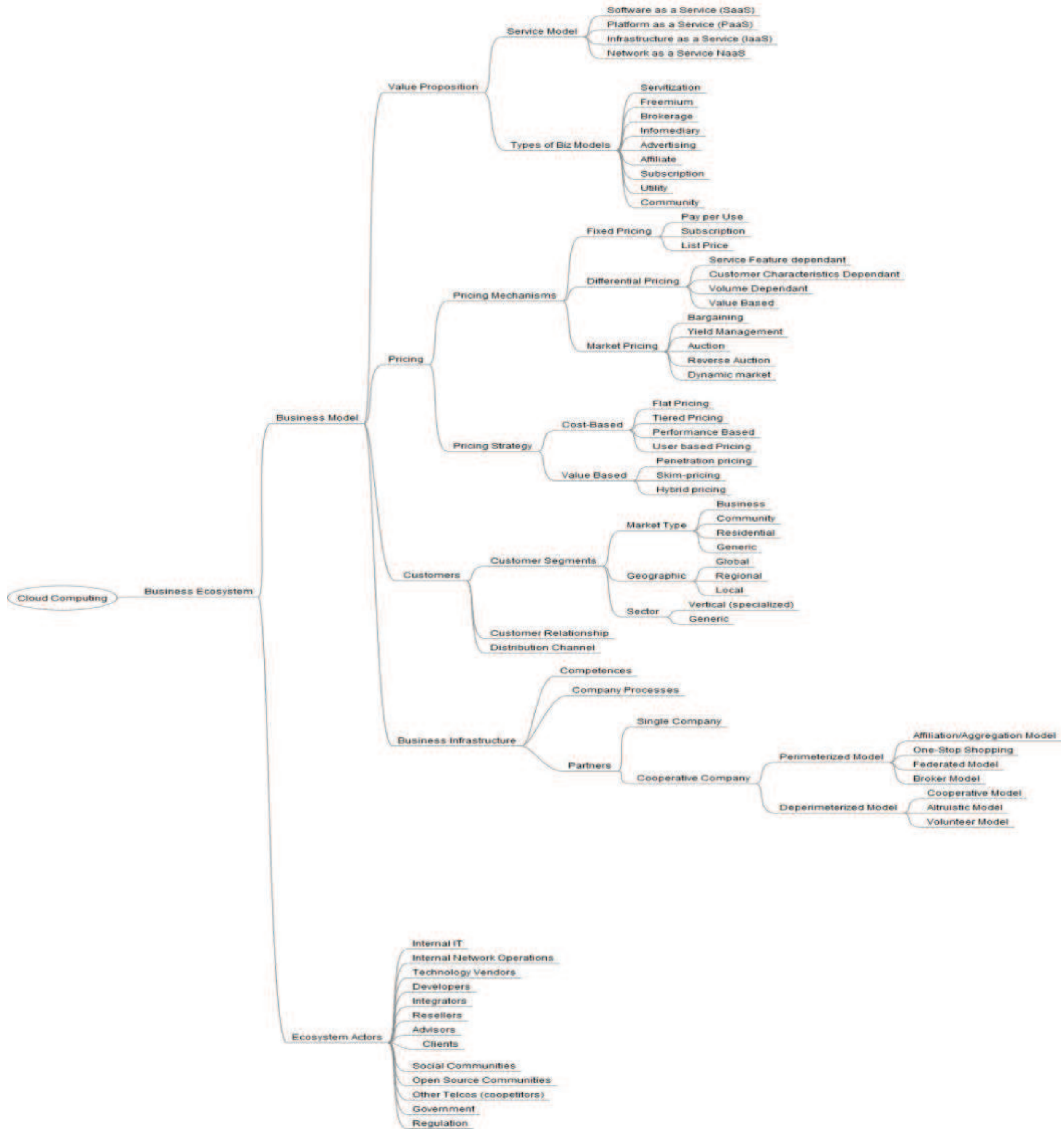


Figure 59: Taxonomy of Cloud Computing from a Business and TelCo Perspective.

The Business Model

The Business Model definition should encompass at least a convincing value proposition, i.e., a clear definition of the value chain and the perceived benefits for customers and the other stakeholders. This is a combination of the service model (XaaS) and the types of Business Model applied to the service model. As described in (Armbrust, et al. 2010), there are several business models for cloud computing and the Internet. For a complete list refer to (Rappa

2004). The TelCo proposition should be aligned to the real possibilities that an Operator has in the market. Some types of business models are out of scope (such as Manufacturer, or others) while the Broker one seems to fit well in the tradition and skills of TelCos. Another viable option is the possibility to help customers to better enter in the realm of servitization by means of a cloud solution, i.e., the TelCo, the customer and the users can establish a Business to Business to Customer, B2B2C, relationship in which the cloud platform enables the customer to move from the selling of products into the selling of product-related services. The capillary presence of TelCos in the territory and the support of customer care department can make this even more appealing and possible.

Pricing is another important aspect of the construction of a sustainable business model. Different pricing schemas can be applied to Cloud services. They can have a fixed price structure in which the user pays for resources usage, or a subscription fee. However, cloud computing can be charged also according to the perceived value of the customer of the services provided by the TelCo. **Figure 60** illustrates the value that customers give to the different service models.

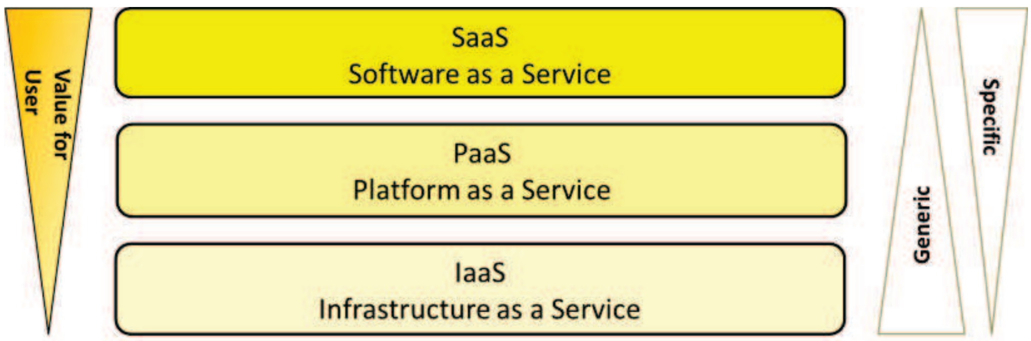


Figure 60: The Value of Cloud Layers and Functions as Perceived by Users

A hybrid pricing model could be offered to customers: basic and generic functionalities of the Infrastructure and platform could have a fixed price or a differential one (depending on quantities and volumes), while specific and tailor services could be feature dependent. Other interesting options are related to the possibility to dynamically make the price by means of auctions or bargaining with customers for resources and features made dynamically available. For instance, Google is using a complex auction mechanism in bidding for personalized advertisement, a similar approach could be adopted for allocation of valuable resources of the cloud. Other pricing strategies should be carefully analyzed in order to align the pricing schemas to the cost structure of the cloud solution that the TelCo is building.

Another relevant issue is determining the target customers. Web companies have a supremacy in the customer market and a strong grip on Small and Medium Enterprise, SME, market, but sometimes they lack the same hold on larger businesses. One possible step for the TelCo is to address mainly the business market by leveraging its local customer management and channel distribution capabilities. In addition, TelCos could differentiate services and features of the cloud platform in terms of vertical markets, i.e., the cloud offering could be instrumental for many businesses for better cover and exploit specific markets (such as Public Administration, e-health, smart cities and the like). Another major point is the possibility to focus on a national market or to have an international footprint. In this latter case, the TelCo should create a network of relationships and allies in those markets in which a direct presence is not possible. In addition, the TelCo could take the opportunity to leverage different deployment models in order to create a global cloud solution. As a rule of thumb, a TelCo should adopt and promote the federation of cloud solutions in order to create a global coverage of the market, at the same time, it should act locally by promoting hybrid solutions in local markets. The combination of a federated and a hybrid approach allows to provide to customers a large portfolio of standardized and generic services. These services could be locally tailored to the needs of the specific market or even customers. This flexibility and elasticity should be supported at the technical infrastructure level by a high degree of programmability and composition of services.

Eventually, a TelCo should take care of its Business Infrastructure, i.e., the combination of skill/competences, processes and company attitude in doing business. Processes and IT skills are quite important for a successful deployment of cloud infrastructures, however they should also be supported by the right capability of doing business in a cooperative way (e.g., by involving partners or with a "doing yourself" approach). Another dimension of the Business Infrastructure is the willingness to pursue an open or a closed model for the cloud platform. A closed way of operating in the market naturally excludes deperimeterized coverage because the company is reluctant to operate in within a dynamic scenario of short term or opportunistic relationships and alliances. In this case a perimeterized model (and market) is more appropriated. The previous **Figure 59** summarizes some of the aspects related to the Business model dimension.

5. 3. 3. On the Cloud Technical Framework

The NIST technical framework (Liu, et al. 2011) under which cloud solutions can be designed and implemented is extended in this section. The reasons for this broadening lay in the need to better leverage from a TelCo perspective the network assets and to promote them to the general attention. This leverage is not pursued with a traditional perspective (i.e., the network has value and it provides Quality of Service related features), instead the networking capabilities are framed within a highly distributed environment compatible with the end to end principle of the Internet.

One of the beliefs is that networking aspects will be more considered in the future of cloud computing and will not be treated as minor issues in the provision of cloud solutions. If Nielsen's Law holds true (Nielsen 1998) (i.e., a high-end user's connection speed grows by 50% per year[30]) then the users will be limited by the network capabilities. This means that the growth in bandwidth will lag behind the grow in processing power. From a cloud computing perspective, this law could have interesting consequences: the commoditization effect on processing will be faster than on bandwidth (i.e., 18 months for doubling the processor power vs. 21 months to double the available bandwidth); bandwidth could maintain a premium value over computing. If this law holds valid than a sort of **paradox** could emerge: Cloud Computing Providers started by providing a value added service, but they will end up providing a commodity service, while TelCos are starting by providing a commodity service and will end up offering a value added connectivity service needed to the whole cloud ecosystem. Scarcity of bandwidth could be an important factor for the optimization of network resources.

The enabling technologies of cloud computing cover a broad spectrum. They range from virtualization up to data management. Virtualization has been applied so far mainly to processing and storage. New emerging technologies are bringing virtualization benefits also to other kind of resources: networking resources have been covered discussing the advancements of projects like OpenFlow in Section 4. 4. Smart objects and sensors will be virtualized as well. This will lead to a decoupling of local proprietary sensor solutions from the virtual representation in the cloud of virtual and smart objects. The combination of virtualized communication, processing and storage capabilities coupled with smart objects

[30] See http://en.wikipedia.org/wiki/Nielsen%27s_law#Contributions last accessed May 30th 2013

within the cloud will make possible new kinds of applications in the fields of Internet of Things and Ambient Intelligence. Any real object could be in the near future virtually represented by means of a clone in the cloud. The relationship between a real and a virtualized object creates a sort of continuum that allows to interact, manipulate and govern real objects that can be augmented in terms of features and intelligence. The number of distributed smart objects will increase over time and it will soon become difficult to control and manage them by means of human intervention. These objects will progressively expose intelligent behavior and the ability to self-organize in complex situations becoming autonomics. Autonomics capabilities and ubiquitous communication will make these object pervasive. Pervasiveness will determine the possibility to be involved and actually support and control large parts of production life cycles, or processes in the home or in the enterprises. Their strong relation with cloud computing will bring an increase in the perceived value of cloud based application and services. Objects will need to communicate each other in order to adapt to the execution context and to the desiderata of end users. The topology of cloud infrastructures will change because pervasiveness, heterogeneity of intelligent entities and objects governed by cloud services, dynamicity and elasticity of service environments will require the ability to integrate different resources into autonomic and intelligent systems. They will be arranged in a distributed fashion, but for specific applications there will be the need to centralize resources and architectures (e.g., government related applications). In addition, (virtualized) resources are to be programmed and controlled, extensibility and programmability of resources will be a common requirement. In environments operating closely with the final customers there will be an increasing need for trust also in the software development. Open source implementations could find a further boost because they are controlled and extended by large communities of programmers that continuously check for bugs and malicious developments. Deployment models will greatly vary from closed environment (e.g., private clouds) to federated solutions or even “clouds of clouds” (i.e., interclouds) in which capabilities and resources of heterogeneous and different infrastructures will be negotiated and then integrated in a seamless platform. This evolution will emphasize the need for interoperability between different clouds, the value of connectivity for creating dynamic links between resources and consequently will require trusted and regulated stakeholders. TelCos will have a chance to play a relevant role in this context by leveraging assets such as connectivity, identity management and a regulated behavior that is respectful of users privacy.

Cloud computing will also extend its essential capabilities by offering choices between consistency and availability of data. Programmers and developers of services will be able to choose between different solutions for guaranteeing transactions and consistency to financial services or high availability and real-time speed data stream management. Terminals will be more and more integrated in the cloud, initially as simple clients but progressively they will become active nodes in the provisioning of services.

Figure 61 is providing a view on the technical framework of the upcoming cloud computing.

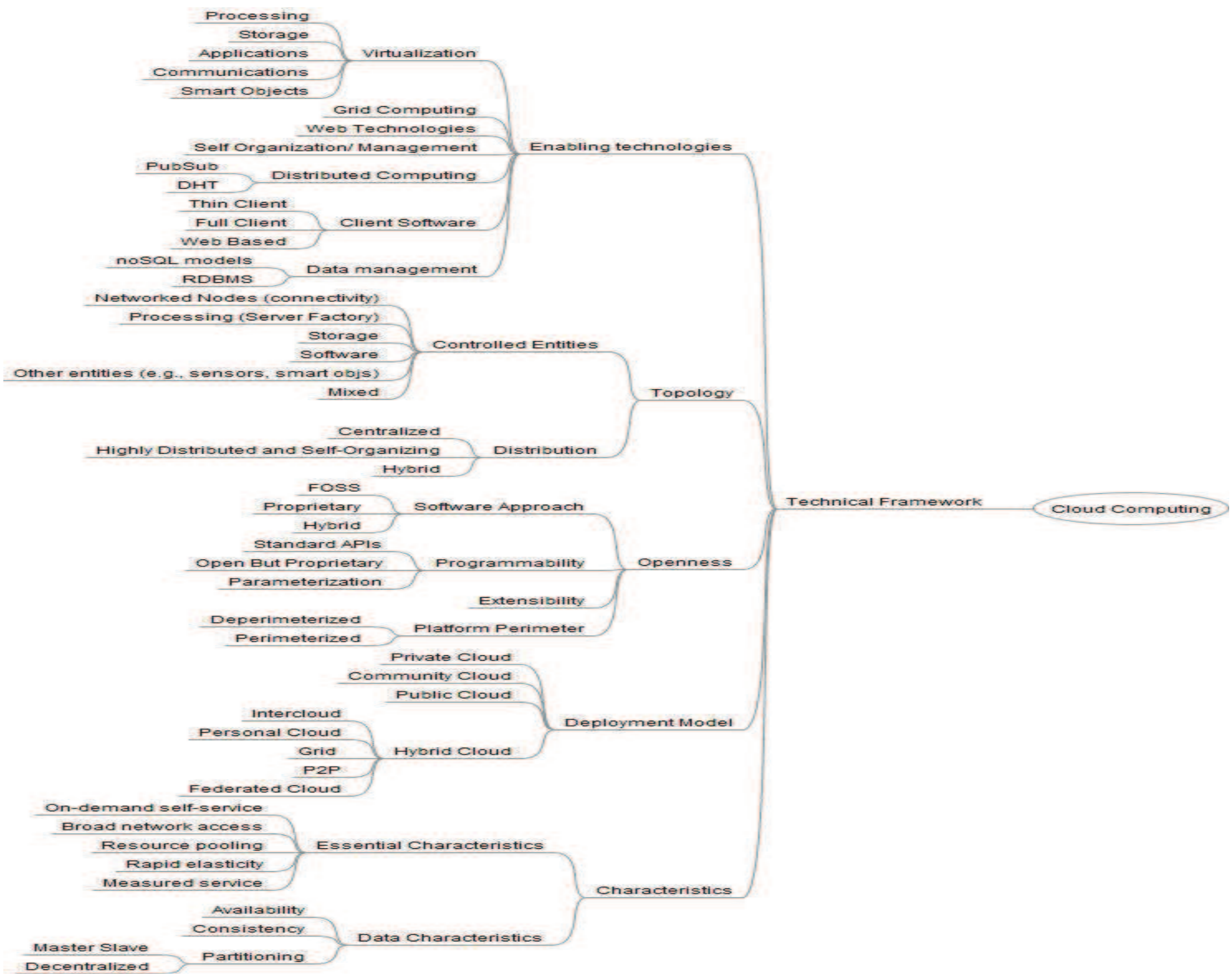


Figure 61: A Taxonomy of Technical Aspects of Cloud Computing

5. 3. 4. Examples of TelCo Services in a Cloud Context

TelCos have to use the Cloud approach for providing services mainly to enterprises and business customers. This choice is dictated by the following reasons: business customers are more interested in creating a long lasting relationship with the provider, this link can also leverage customer relationship management systems; the existing billing relationship for connectivity services can be exploited in order to promote a sort of one stop shop approach for many services; the TelCo can leverage the local footprint and to provide integration capabilities at a global level, for certain enterprises this can capabilities can make a difference.

The TelCo's approach to the cloud has to promote and leverage the interoperability and integration of different complex IT systems into a federated infrastructure with a rich service portfolio. However this large infrastructure should be flexible enough to accommodate for private systems. Enterprises should be able to decide which processing, storage, communication and sensing capabilities to keep in-house and which ones to externalize.

In the following section, examples of Cloud services (in general according to the Platform as a Service, PaaS or Software as a Service, SaaS models) based on these foundations are given. A Business to Business to Customer approach is usually pursued in order to leverage the assumed capabilities of a TelCo's cloud platform.

Intermediary in the Internet with Things

The Internet of Things is an environment in which some intermediary functionality can have value. (Bohli, Sorge and Westhoff 2009) propose that TelCos play the role of intermediary between small (wireless) sensor network providers and final customers. The Operators could even promote the wide adoption of (wireless) sensor networks by subsidizing the small sensor providers. The goal is to collect a set of meaningful data that can be exploited by determining macro trends within a community or a specific location.

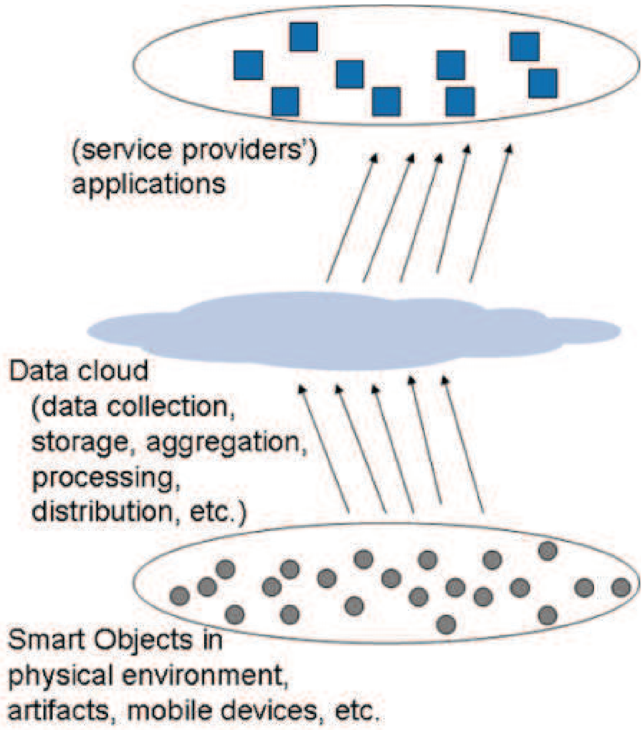


Figure 62: The Intermediary Service in an Internet of Things Context

In particular the Cloud platform could be extended for offering functions aiming at a «Data Cloud»:

- to collect data from sensors networks geographically dispersed and to integrate them with other data sets (e.g., from the Public Administration in an open data fashion, from users' devices, and automotive embedded devices);
- to store large data sets to be dynamically updated and to analyze (data mining) and aggregate data into information;
- to provide message based engines and to derive information from complex events management;
- to distribute to Service providers' applications relevant and updated data with the requested aggregation level and according to the agreed dispatching mechanism (e.g., pub-sub, continuous query, big tables,…).

This service could be characterized by:

- frequent interaction with aggregation systems of distributed sensor networks;
- elasticity in processing and storage allocation depending on data and processing to be dealt with;
- interaction with centralized and traditional applications;

- Data cloud capabilities offered as SaaS and SaaS.

Figure 62 describes the service.

Massively Multiplayer Online Games (MMOG)

MMOG Providers at a global level could request cloud computing services aiming at:

- Optimizing the processing and storage load balancing;
- Optimizing the network features in order to provide a better service to their clients, e.g., optimization of latencies for the users, dynamic increase in bandwidth allocation (e.g., to download large files), etc.

The Cloud Computing Provider, in addition, could provide specific functions for the MMOG Provider such as:

- Load prediction service, in order to evaluate the servers load based on a prediction of distribution of entities and players;
- Resource allocation service: to provide local servers and rebalance the processing load due to an increasing number of players.

The service should aim (maybe in a Silk fashion) at rebalancing the processing and storage load in order to fulfill real-time constraints and real time capabilities as negotiated with the users.

Features of this service are:

- support of QoS parameters in the interactions with user terminals;
- load balancing and optimization capabilities.

Similar features could also be used to better support multimedia service provision within a specialized cloud.

They could be packaged in a specific PaaS offering for MMOG or generic Multimedia Providers.

Virtual Terminals

Physical devices could be virtualized in the cloud and augmented with additional features or capabilities (e.g., more processing power or more storage). The physical device could use its virtual image for the execution of background processing; for the migration (teleporting) of tasks that do require too many physical device resources (Chun, Ihm, et al. 2011); migration of tasks that requires more resources than those available in the physical device or that have nonfunctional requirements (e.g., performance, security, reliability, parallel execution) that the physical terminal cannot satisfy; delegation of tasks to be executed when the physical device is not connected to a network; extension of storage for keeping all the events forwarded or generated by the terminal. Feasible scenarios are related to the usage of this service in the context of network PC provided to an Enterprise for supporting Teleworking capabilities.

The major features of the service are:

- elasticity in allocation of processing and storage capabilities as a consequence of the dynamic needs of running applications
- migration of the virtual terminal based on the actual location of the corresponding physical mobile device.

Also in this case the Virtual Terminal capabilities can be packaged as PaaS and/or SaaS offering.

5. 3. 5. An Agenda for TelCo Oriented Cloud Platforms

As seen, web companies have an edge from the technological and the market perspective over the TelCos. In order to recover the gap there is the need of a coordinated and standardize set of actions aiming at the definition of an open, programmable and federated cloud platform.

Following the offering of the web companies using proprietary platforms (maybe acquired by IT companies) does not solve the problem, Also IT companies are lacking behind and to many proprietary solutions do even fragment the cloud computing market. Operators should take actions in order to move towards a grid based approach. This means to embrace the heterogeneity of the platform components and the ability to mix and match resources pertaining to different administrative and technological domain. The task is more complex that aiming at the construction of a single and proprietary solution, but the risk is to be kept at the margins of an ever increasing market.

In addition, virtualization capabilities are emerging also in the realm of sensors and smart objects as well as in the networking itself. This will configure a situation in which the dynamic combination and programmability of processing, storage, sensing and communication resources will move intelligence, services, applications and infrastructures towards the edge of the network and within the customers' domain. The TelCos can offer to this edge environments programmable and on demand connectivity as well as the ability to support mechanism for self-management of complex edge networks as well as complement the locally missing resources with virtualized ones. This is a daunting objective and it should be approached with a step to step strategy. The starting point is interesting from a technical and marketing point of view: the integration of Enterprise IT systems within a federated cloud approach. From a marketing perspective this means to create a sort of hybrid cloud in which the enterprise IT systems maintain their importance but can cooperate with remote systems and access to specialized applications in order to improve and enrich their functions. Such an hybrid cloud platform could become a sort of cooperative environment in which different enterprises can implement or integrate companies' processes, functions, applications, services and market places for conducting business.

The cooperative cloud is then instrumental to create an ecosystem in which different enterprises contribute in terms of resources and final customers can access a large set of specialized services.

In order to make such a platform a reality, TelCos should be instrumental to the definition of open architectures within important standard bodies. The more time elapses without this standardization effort the more the cloud computing business is in the hands of web companies proposing walled gardens.

Some core specifications already exist and they are in the field of Open Grid Forum and Web Services. TelCos should somehow endorse them and to promote a further level of standardization and visibility. The concept to "Virtual Organization" behind the grid computing fits well with the possible business proposition that TelCos should pursue and strive for. In addition support to network virtualization (e.g., OpenFlow) should be guaranteed. Software defined networks are the next step towards a change in the connectivity and communication proposition of Operators. The possibility to create virtualized views on network capabilities and to offer them in an integrated manner with Virtual Private Networks or even better with Virtual Organizations will provide a viable business framework for the Operators. A large ecosystem of application developers, process system integrators and IT companies could exploit the capabilities offered by such a different cloud platform.

5. 4. Bank of User Data

Personal data are gathered and exploited by many WebCos in order to support their Business Model (usually related to advertisement). Data are collected in several ways, e.g., by monitoring the activities of users (e.g., cookies), by storing the information they put in their profiles (e.g., Facebook) or by storing and associating the search history to the account of users. There is a continuous trend to get more and more data from users and then to derive information (and possibly sell it) by applying Big Data and Business Analysis related technologies. The Users are constantly spoiled in their rights and property. However there is an increasing awareness of the importance of the issue, for instance (Benkler 2006) has inspired a group of students of the New York University to develop and run a competitor of Facebook called Diaspora (**Figure 63**). This awareness is leading to a new approach in dealing and using the personal data (World Economic Forum 2011), (World Economic Forum 2013).

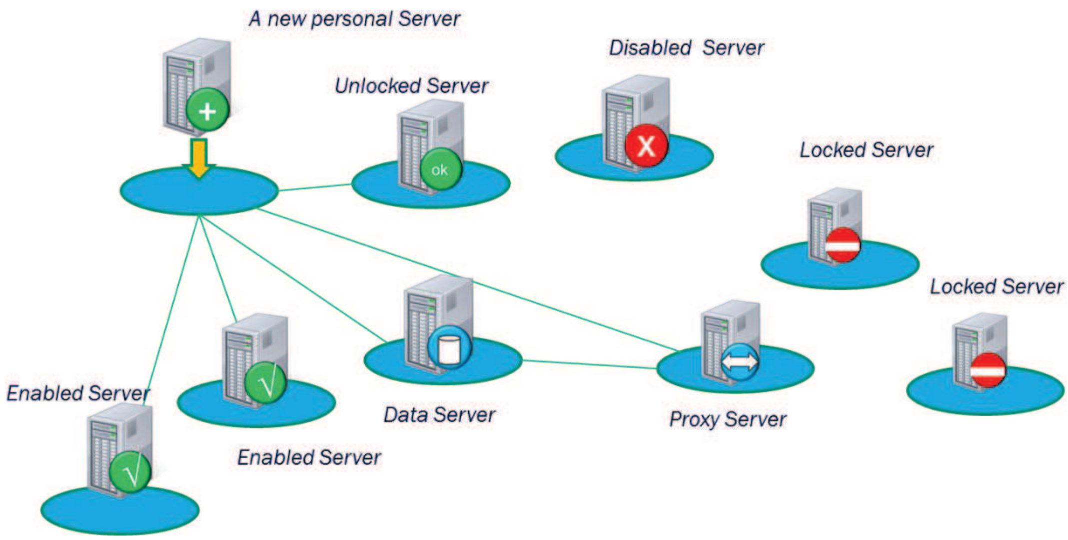


Figure 63: Diaspora, a Social Network Respectful of Personal Data Ownership

For instance, the Diaspora system is characterized by the fact that data are and remain solely propriety of the users. Users can also decide if personal data are to be stored and shared in general servers, or if they have to remain and be shared directly from the user devices.

5. 4. 1. The Value of (User) Data

This is an example of how to deal with data under the push of privacy requirements of aware users. The World Economic Forum (World Economic Forum 2011) considers the personal data as a big opportunity for the digital economy. Its proposition is to engage the user in such a way to get the usage permission in a more aware fashion.

However this is not enough, the personal data should be dealt with as the “money” of users is treated. The metaphor should be the one of the bank, a bank of the user data. The foundation of this is that data are owned by the user and they have to be protected and made available to the user whenever it asks for that. If the user wants to invest some of its data, i.e., s/he wants to allow others to use personal data, it can ask the Bank to find out a possible market for selling the data and to state the conditions under which the user will grant the access to its data. In other terms, the bank of user data can act as a sort of Broker that guarantees the investment to its client. Data will be processed, stored and manipulated according to rules and requirements decided by the user and not by the providers. Under this condition, a return of investment could also be negotiated between the user and the Broker or the final exploiter of the data. In addition, personal data could also be made available in an anonimized fashion and the Broker could play the role of aggregator of meaningful data sets, guaranteeing on one side the privacy of the users (by anonimization) and the return of investment, and, on the buyer side, it guarantees the statistical relevance and aggregation of data. **Figure 64** depicts a possible ecosystem around the concept of Bank of User Data.

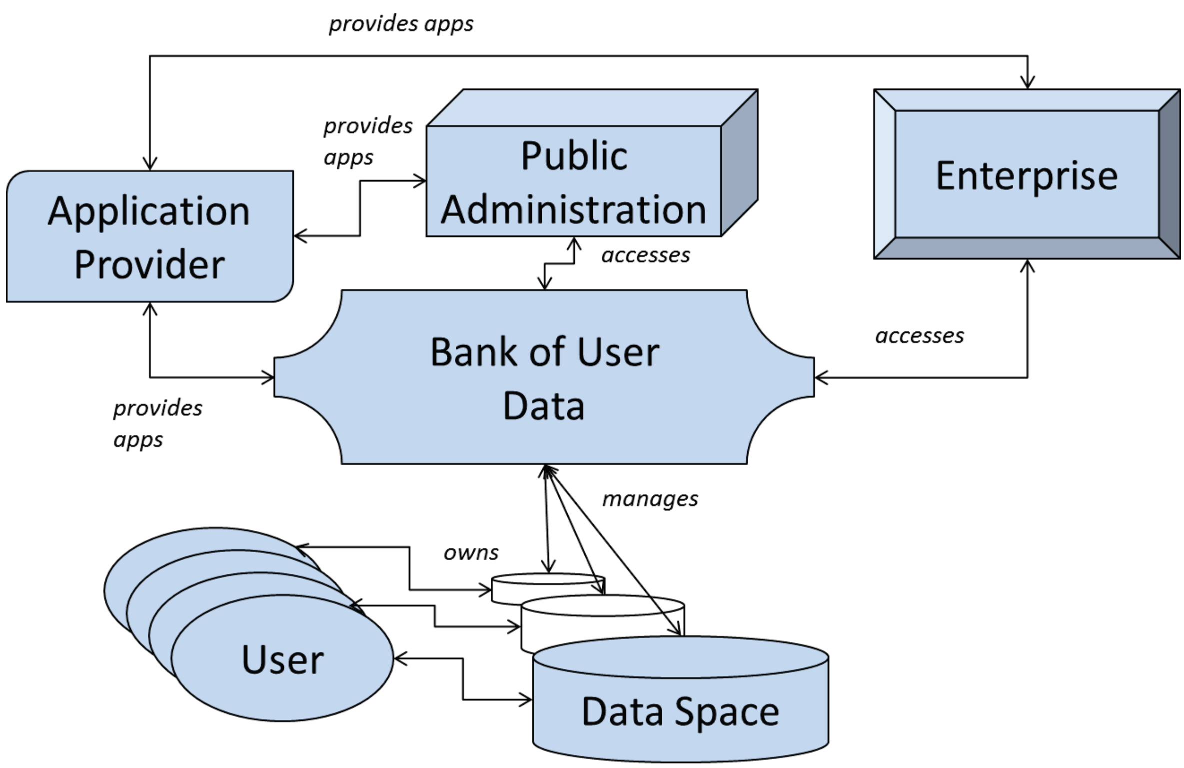


Figure 64: The Bank of User Data

Some simple applications could be considered as examples of the valuable features returned to the user:

- The user can have access to the full list of its digital expenditures and in addition it can get information (type, model, warranty) of the products that has been bought. In case of a fault of a product, the user can have an easy way to find out all the documents for a reimbursement or the repair assistance under the warranty. In addition the user can collect expenditures from several sources and not only from its bank or credit card provider. So its view will be an holistic one.
- The user can provide its anonimized or clear data in return for money or for free services. The contract is agreed considering also the requirements and the policy set forth by the user.

Pushing this approach to the extreme consequences, the user interactions with digital services (even the most common ones) should be transactional, i.e., for any request of the user to a server or a service, a transaction together with the results of the query (in a client – server model) should be returned in order to monitor and collect all the relevant information from the user side. **Figure 65** depicts this occurrence in the case of a search query. It also shows that the profile of the user should be built by the user itself.

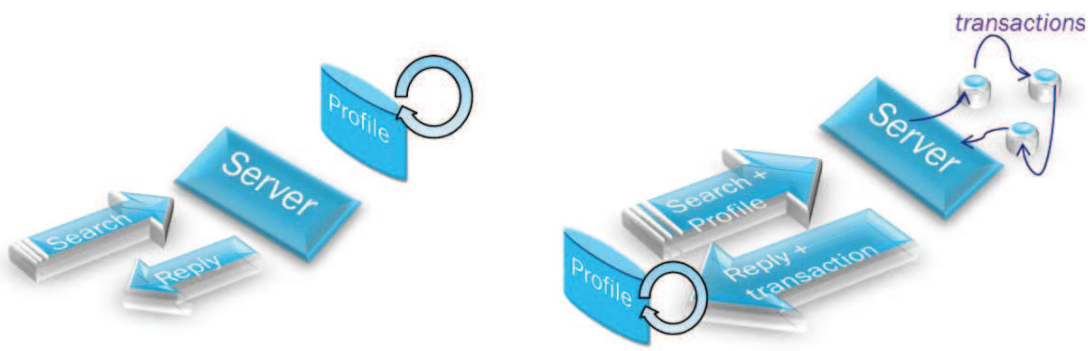


Figure 65: Adding Transactions to the Web Interactions

The value of this approach is that the user could collect an holistic view of its actions, behavior and habits. Extending the example in **Figure 65,** the user could allow the search provider to use a relevant part of the user profile in order to better satisfy the user request. The user profile could contain a list of interests of the user, a part of the aggregate search history (spanning over several search engines), and the like. The service could be more scoped down to the real interest of the user and the collection of responses and transactions could further enrich the user profile to be used for future interactions. Only the user will have access to the whole set of personal data. Entering in this transactional Internet will be more and more important for empowering the user over merchants and service providers. More information on the Bank of User Data can be found in [XVI, XX].

5. 4. 2. Towards an Architecture for Dealing with Aggregated Data

In order to capitalize the value of data, there is the need to create an ecosystem supporting meaningful value chains. The approach advocated is the one termed “user centric”; i.e., data generated by users are owned by users and they could be managed by other parties if and only if the user decides to share them according to precise policies negotiated with the other party. These relationships should be supported and made possible by means of an enabling platform. Since data pertain to users, in principle data can be scattered over many different user systems. For this reason, the usage environment has to be intended as a highly distributed platform able to track and integrate distributed data. **Figure 66** represents a logical data management platform.

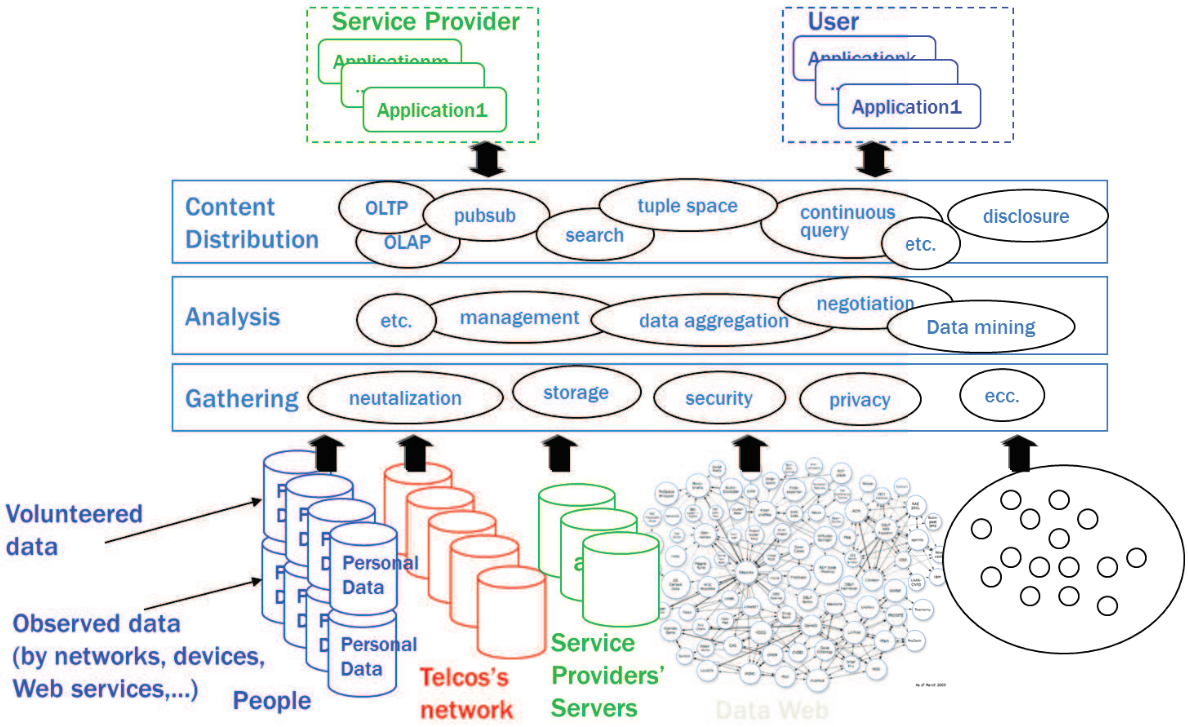


Figure 66: An Architectural Framework for a Data-centric Ecosystem

Data will be collected in different fashions, some data can be directly provided by the data generators, other data can be collected by observation of activities or behaviors of devices and users, other data can be retrieved by accessing data offered by service providers (e.g., TelCos or web companies), finally other data can be semantically derived by the web and specific data sets made available by environments or communities. One first important level is represented by the set of functions that allows the collection, gathering and organization of raw data. At this level, it is important to provide good mechanisms in order to support capabilities, but even more important are those mechanisms that allow the neutralization, the privacy and the security of data. Once data have been gathered and organized, the second level is related to the analytics of the available data, here several functions and mechanism can take place. Data Mining is the primary function, but also the possibility to correlate at real time the data is important (e.g., NoSql techniques). The upper layer is supporting different application oriented mechanisms for accessing to the available information. Access to these data and information will follow different paradigms such as: transaction oriented, i.e., data can be inserted, or read, or modified or deleted in a single transaction; stream oriented, i.e., data and information are produced and are continuously forwarded to the consumers as in a stream (in this case mechanisms such a PubSub (Tarkoma 2012) can be offered). Other mechanisms can be offered such as tuple spaces and blackboards or continuous query. It should be noted that data availability and usage is changing, for instance there is a trend in considering new way of dealing with data. Streaming computing (as proposed by IBM (Turaga, et al. 2010)) is an interesting case of how to process streams of data using a chain (or a graph) of cooperating nodes. The data streaming processing (e.g., S4 from Yahoo (Neumeyer, et al. 2010), Deduce from IBM (Kumar, et al. 2010) and in general complex event processing solutions (Buchmann 2009)) are paving the way for new paradigms for dealing with large sets of data.

The platform depicted in **Figure 67** should be capable of supporting several mechanisms to deal with data. One interesting concept that will be more and more important is the need to store the data and to make them available by publishing or forwarding them according to the needs of the applications. Many technologies can be used, but the concept of "store and share" is fundamental in order to enabling compelling applications.

In this field it is worth to mention the Twitter architecture (Krishnamurthy, Gill and Arlitt, A few chirps about twitter 2008) that could be considered the best implemented example on how real-time data are stored, processed and made available to the users. **Figure 67** depicts two important components of the Twitter architecture.

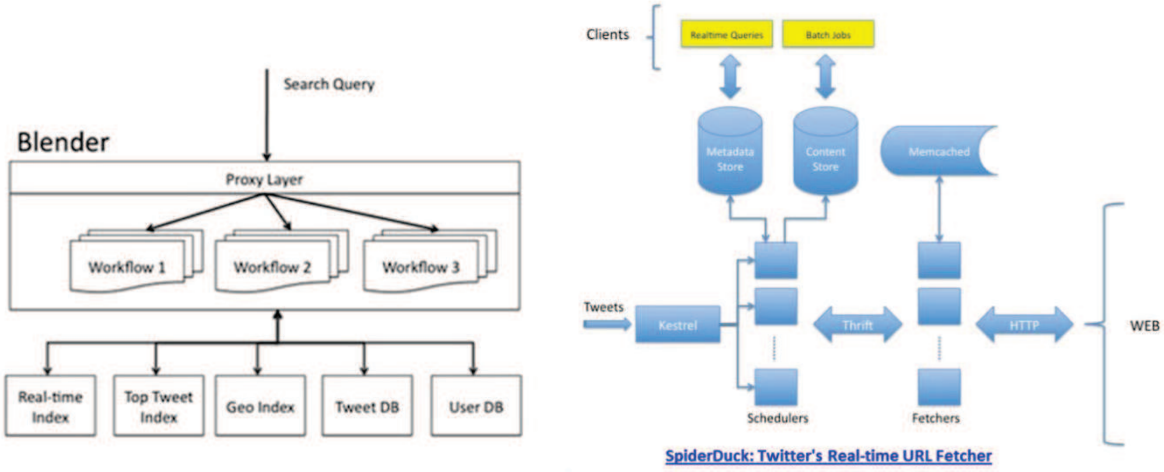


Figure 67: Blender and SpiderDuck, Two Building-blocks of Twitter Technology

The Twitter architecture is mainly based on Open Source technologies that the web company integrates in an innovative way in order to deal with real time data. To the best knowledge of the author, there is not a clear and complete description of the Twitter architecture, but mainly a description of its components. For example:

- Blender: Queries from the website, API, or internal clients at Twitter are issued to Blender via a hardware load balancer. Blender parses the query and then issues it to back-end services, using workflows to handle dependencies between the services. Finally, results from the services are merged and rendered in the appropriate language for the client.
- SpiderDuck: it is a service at Twitter that fetches all URLs shared in Tweets in real-time, parses the downloaded content to extract metadata of interest and makes that metadata available for other Twitter services to consume within seconds.

Twitter is particularly important because its capability to deal with real-time stream of data can be used to develop very compelling applications in several domains. In the future, the Twitter "engine" and its architecture could be more important as application enabler than the Google search platform and it is using a different paradigm than C – S.

5. 5. Smart Environments

A smart environment is "*a physical world that is richly and invisibly interwoven with sensors, actuators, displays, and computational elements, embedded seamlessly in the everyday objects of our lives, and connected through a continuous network*" (Weiser, Gold and Brown 1999). Smart environments will be fostered by societal efforts (e.g., Smart Cities), commercial offerings (transportation, home automation, …) and end users endeavors (e.g., www.bwired.nl, or even FON[31]).

Smart Environments integrate the pervasiveness of processing/storage/sensing resources with intelligent behaviors and the ability to understand the "context". Smart environments will be characterized by:

- the intelligence of the ambient, i.e., the capability of the system to analyze the execution context; to adapt the ambient resources and the user devices in order to accommodate the user expectations; to learn from user and resources behavior;
- the awareness, i.e., the capability to identify, locate, and orchestrate resources according to the identified users of the ambient (the humans) and their understood needs and intentions.
- The ability to melt data stemming from different sources in order to virtualize or augment the user perception of the surrounding environment.

The concept of context is central to the smart environments. The User Context is made out of different components/resources of various networks that are integrated in order to create a single environment that fits the dynamic user requirements. There are a few typical characteristics of the user context such as:

- Physical context: i.e., lighting, noise, traffic condition, temperature and the like;
- Time Context: such as time of a day, week, month, season of a year;
- Computing context: it is defined in terms of processing elements, storage capabilities, network connectivity, communication cost, communication bandwidth, nearby communication resources;
- User context: it is defined in terms of user profile (the collection of preferences and data related to the user), location, social situation;
- Emotional context: it should encompass feelings, psychology, ...
- Social Context: in terms of People, Conventions, Etiquette and so on.

A few of these characteristics are referring to time and space or to physical objects (e.g., the computing context), while the most difficult to grasp and understand context aspects are related to users and their personality and social links and behaviors. Context awareness (Baldauf, Dustdar and Rosenberg 2007) is one of the examples of the coordinated work that

[31] http://corp.fon.com/

different sciences have to carry out in order to achieve meaningful results. Cognitive aspects will have an increasing relevance in here, in order to understand the human behavior and the intentions, for creating new compelling interactions modes, for "injecting" intelligence within computing environments. Actually, three main trends in cognitive studies can be identified in ICT

- Cognitive architectures (mainly related to AI studies) aiming at providing reasoning and adaptation to changing situations;
- Cognitive networks (e.g., (Clark, et al. 2003)) aiming at providing a cognitive view on networked resources. They work towards a Knowledge plan for the Internet;
- Autonomic and Self-organization architectures, aiming at the ability for each node to self-organize and work in complex environments by means of a control loop.

A cognitive approach has been extensively applied to specific problem domains (reasoning about a situation) but just recently it has been proposed for specific problem domains AND the supporting infrastructure (especially in FP7 projects and Internet of Things (IOT-A 2011), iCore[32]). This means that cognition will be applied to humans as well as to environments. In fact, in the long term, Cognition and Intelligence for Digital Ambients will emerge as a viable technology.

Many European projects are putting considerable effort in this area aiming at easing the Self – CHOP management of systems of increasing complexity:

- Self-CONFIGURING, i.e., a system can dynamically adapt to changing environments
- Self-HEALING, i.e., a system can discover, diagnose and react to disruptions
- Self-OPTIMIZATION, i.e., a system can monitor and tune resources automatically
- Self-PROTECTION, i.e., can anticipate, detect, identify and protect against threats from anywhere.

Availability of these solutions at industrial level will radically transform the way ICT systems are managed and operated, with great impact on OPEX and redefinition of internal processes as well as a great advantage for users that will be able to relay on smart environments able to adapt to real needs.

Smart Environments will be built around the combination of four major features supported by environment's nodes: processing, storage, communication and sensing/actuation

[32] Documentation available at http://www.iot-icore.eu/public-deliverables last accessed April 9th 2013

capabilities. In principle, the combination of these features plus their spatial extension (e.g., small scale communications, small nodes like RasperryPi (Upton and Halfacree 2012) vs. the use of large scale communications offered by operators plus the access to large datacenters) will determine the effective range of usage of the smart environment. They could range from very small and specific environments in the home, in a lab, up to large environments covering an entire city (smart city) or an entire nation. However even small smart environments could be interconnected in such a way to create a sort of "network of networks" whose capabilities and whose extension can go beyond the local and small extension. Using a combination of store and forward capabilities, data generated in a local and smart environment could be transmitted hop by hop exploiting opportunistic connectivity. Data could be downloaded and stored in a car and then their destination tagged. The car travelling in the country side could act as a sort of DHL carrier. Simple calculations have been done for determining the best way to forward large bulks of data, and the use of carriers like Fedex is better than the use of networks. In a way, cars, trains and ships in smart environments could be cheap and available means to transfer in a "social" way large amount of data. When a well-connected spot is reached, these data (or a part of them) could be forwarded by the car to the new actor by means of the available network. In areas that are densely populated the small node connectivity could in principle be a sort of pervasive fabric of connectivity capabilities offered by a great variety of devices. Data can "float" in these environment or can travel through it in order to reach destinations far away from the source. The entanglement of data [XVI] is based on similar concepts. Data could be stored and made permanent even if the nodes of the P2P network offing this service are not. Each node could store and host a copy of the data and replicate it over other nodes. If not all the nodes are turned off simultaneously or in a short period of time, there is the possibility to make the shared data persistent.

Another interesting capability is that all the transport links could be based on small range connectivity between close and adjacent nodes. This means that a "network of networks" could represent an alternative service to large scale public network for certain applications (not real time and with simple QoS expectation).

In order to achieve this interconnection, a sort of social " factor" should be present and exploited. As previously seen one of the problem in similar environments (e.g., P2P networks) is the opportunistic behavior, i.e., a set of nodes take advantage of the features and capabilities offered by the network without returning any value to the community. There are many studies tackling the problem to enforce or at least to promote an altruistic behavior. In (Ohtsuki, et al. 2006) such a behavior occurs when the number of links between a node and its neighbors exceed the ratio between the benefits / cost associated to the sharing of resources. In a jam-packed environment the occurrence of altruistic behaviors should be facilitated and promoted by the large number of connected nodes. Nodes of such a network of networks could be characterized as social nodes, in the sense that they are aware of their importance for ensuring the consistency of the network of network.

These smart environments could be characterized by these features:

- Smart Environments could be seen as new communication entities (a kind of new terminal/network);
- Smart environments can be composed by: a) simple objects that are controlled intelligently by the user device; or b) smart objects that can cooperate and offer intelligent functions.

In any case, the aggregation and integration of these environments will present hard to dominate issues. They are similar to those exposed by complex systems. In this sense, strategies for optimizing the usage of resources cannot follow traditional approaches and have to integrate dynamic networking features (as those supported by P2P networks) and autonomic and cognitive ones in order to enforce an intelligent behavior to the whole system. Terminals will play a major role in providing this intelligence, however, the Operators could offer optimization functionalities at the edge of their networks in order to help in coping with the complexity. Actually the public network edge nodes could behave as anchor points in order to introduce some linearity and stability to very dynamic systems.

From an application point of view, end users will be fundamental actors of smart environments: their increased awareness for social and environment issues (e.g., reduction of CO carbon footprint, energy savings, and the like) and the willingness to take advantage of technological evolution for easing life will create opportunities for the development of these smart environments and especially for applications and services. From the implementation and deployment points of view, social communities will be fundamental for giving a viral push towards the use of smart environments and solutions; developers communities as well could help in order to reduce barriers to deploy and use smart applications that integrate, make use and support smart objects. Smart objects will be more and more available thanks to low cost sensors and actuators and by the increase capability of end users to build smart objects themselves. Programmability issues will also lessen because of new specific solutions or programming languages for these kind of environments (e.g., the scratch language (Maloney, et al. 2010) for kids is able to control sensors or AmbientTalk for dealing with not well connected environments (Dedecker, et al. 2006)). As demonstrated by a number of success stories, crowdsourcing for developing hardware and software solutions tailored to customers' needs could find its way as an established way of promoting new products or elicit the user requirements.

From the communication stand point, it should be noted that mobile terminals are already offering (and they will support even further this option) the possibility to integrate sensors (from NFC up to more complex ones). This feature together with the continuous development of low cost or cheap pervasive systems, gives rise to the possibility to use various communications means (direct communication, WiFi, Zigbee, mobile wireless) for interconnection of different environments. In addition there are already some communication engines (e.g., Pachube/Cosm, but also Twitter) that support new paradigms of communication well beyond the client - server. The elements for creating compelling services and to personalize these for the specific customer on top of smart environments are technologically emerging. The major issues have a non-technical nature: how to accommodate for open ecosystems supporting this complexity, what are the advantages for companies to adopt this approach and the like. Once more the service offering will be determined not by the fulfillment of the customer requirements in an open market, but by the resolution of tussles within an ecosystem dominated by proponents of closed environments (e.g., the web companies, the terminal manufactures, the consumer electronics industry, the operators).

Smart environments will benefit a great deal from data manipulation techniques (e.g., big data) or from graphical and semantic technologies (to be used to present to the user the right information content)

5. 6. Internet with Things

Internet of Things is a "catch-all" name referring to several technical and business possibilities. Usually Internet of Things is seen as a set of vertical application domains that share a limited number of common functionalities. In this document, the undertaken approach is to identify a number of common functionalities that can be used to define a common platform. In order to set the problem domain, this definition is adopted: “The Internet of Things, IoT, is a self-configuring and adaptive complex system made out of networks of sensors and smart objects whose purpose is to interconnect “all” things, including every day and industrial objects in such a way to make them intelligent, programmable and more capable of interacting with humans”. Plenty of opportunities and technical challenges can be envisaged in this sector.

5. 6. 1. The value of the Internet of Things

Internet of Things is rewarded by many as a fruitful application domain in order to generate revenues. Many Operators are trying to capitalize these possibility by means of initiatives under the broad umbrella of machine to machine, M2M, applications. They are essentially characterized by the fact that a set of sensors can use an aggregator device equipped with a communication board and a SIM. Access to a mobile network is used in order to dispatch relevant data to an application center. Often there is not a need for full mobility (e.g., a device controlling the working of a lift does not really need to by mobile), however the low cost of mobility boards and their convenience for sending data are considered as facilitators for building these M2M applications. There are a number of issues that are hampering a wide development of M2M applications even from the perspective of the Operators:

- The revenue generated by sending a few bytes of data is not considered an interesting business proposition.
- SIMs should be activated/deactivated at will by the service provider without the intervention of the TelCo, this means to give up the control of an asset that is considered extremely valuable by Operators.
- M2M applications have showed that traffic patterns can have important consequences on the signaling networks, e.g., when container ships approach an harbor, there could be spikes of traffic due to the simultaneous attempt of thousands of SIMs (one per container or more) to connect to a mobile network (Shafiq, et al. 2012). This sudden increase in signaling traffic can cause problems to networks without providing considerable revenues.
- Management functions of these type of SIMS (e.g., billing, recharge of credit and the like) are sometimes more costly than the revenue generated.

One major point is that the data traffic for these services is not usually sufficient to justify a huge involvement of Operators in this domain. For instance **Figure 68** represents the amount of traffic generated by aggregators (i.e., those devices/systems that receive data from many sensors, and manipulated them and deliver the aggregated data to a service center). The assumption is to keep the size of the single message reasonably small (**Figure 68** shows data for messages 10 kB long) and to consider a large (10 Millions) base of aggregators. If an aggregator forwards few messages (from 3 up to 100 messages per day) then the generated traffic is less that 4% of the traditional traffic generated by phone calls (over a customer base of 20 Millions of phone users). Obviously if the rate of message forwarding increases than also the traffic generated will grow. With more than 1000 messages per day (1440 messages per day means one message per minute) the two types of traffic are comparable.

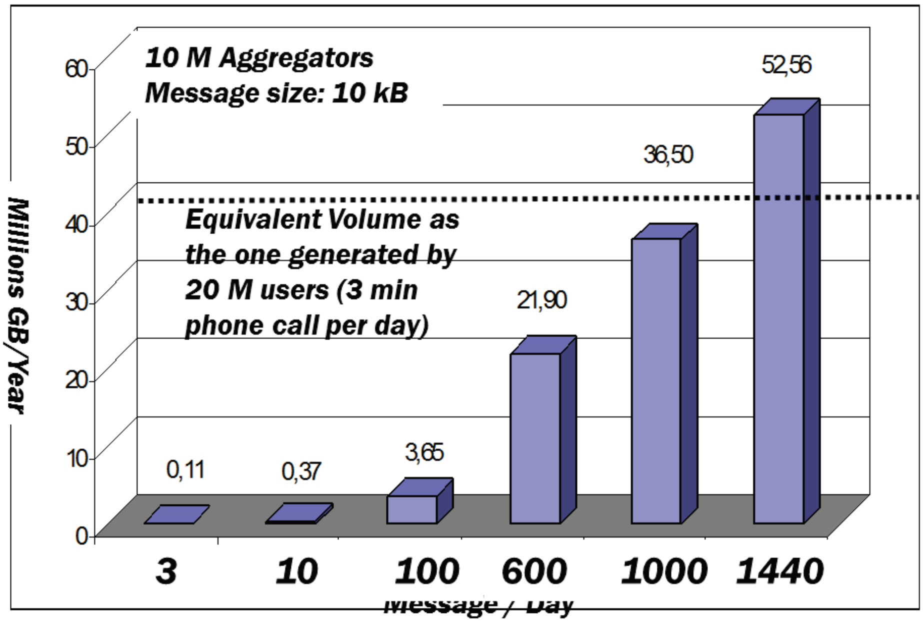


Figure 68: Examples of Traffic Generated by Sensor Aggregators.

In the future it is envisaged that both the message rate and the number of aggregators will grow, but in order to have a substantial traffic increase , sensors and aggregators have to deal and manipulate multimedia information. This means that the size of the single message has to increase (e.g., a photograph) or the single message has to change into a stream of data (i.e., camera recording). So also for sensors, the traffic pattern will drastically change when devices will deal with multimedia data. On the other side, connectivity is becoming a commodity and the Operators cannot base their plans in the area of Internet of Things on future traffic increases. There must be something more substantial to this.

The Internet of Things is based on the possibility to measure phenomena and to deal with large data sets. The major value is in the ability to deal with data. The data management has several aspects in this case: the ability to collect data, store them and forward them to the right recipients, the capability to interpret streams of data and to derive relevant information, to aggregate several sources of information and integrate them in meaningful ways. An exemplification of this new way of dealing with massive data is represented by the Forth Paradigm (A. J. Hey 2011).

How to gain this value is a major topic for leveraging Internet of Things. **Figure 69** represents an hypothesis on how the value of the market of IoT will be segmented.

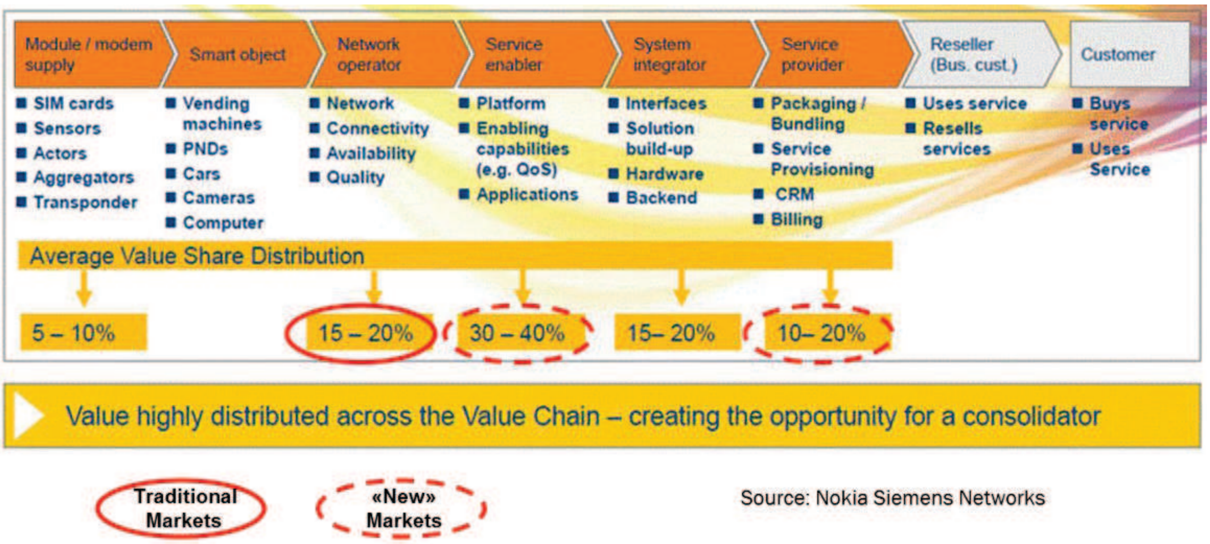


Figure 69: Segmenting the IoT Market Value

A traditional business model focusing on communication can aim at getting a share of about 15-20%, while the major value resides in the platform, its integration and the creation and delivery of services. The major single source of value is in fact the platform, this is probably the starting point for any successful initiative in IoT. Operators could be effective actors of an IoT ecosystem because they have skills and competences in several areas: the communication, the platform and the possibility to manage and leverage large systems (Service Provider). The value of the platform could be further augmented by the capacity to collect and derive information form data been dealt with by the system itself. An example of leveraging real time data is Twitter, it is a valuable platform capable of providing real time information about the hot topics discussed in the Internet and in the twitter community. Twitter in fact is a potential infrastructure capable of seizing and leveraging the Internet of

Things value. In the next section some high level principles for building a valuable IoT platform are discussed.

5. 6. 2. Towards a Platform for the Internet with Things

A viable proposition for the Operators in the realm of Internet of Things are large scale systems, i.e., those systems that comprises thousands and even millions of sensors and actuators devoted to provide useful services to customers and citizens. “Smart cities” will certainly follow in this area. These systems will be characterized by the fact that sensors and actuators will be deployed in a very irregular pattern, i.e., certain areas and environments will be covered with plenty of sensors, while others will have a scarce deployment of them. In addition sensors will pertain to several administrative domains. This means that complexity will lay in the sheer number of sensors, in the uneven deployment, and in the variety of administration domains. These features can be coped with by means of virtualization and the willingness of users and providers to share resources (Bohli, Sorge and Westhoff 2009). In such an environment, the willingness to make available a virtual representation of the sensor will have a paramount importance. Service and infrastructure providers should support open platforms in order to allow the brokering and the dynamic integration of heterogeneous resources. TelCos can be trusted providers that integrate the different systems into a meaningful larger environment.

Virtualization is a means to allow high degrees of adaptation and integration of several resources. Each administrative domain in fact could make use of sensors and actuators based on different technologies, communication paradigms, protocols and software platforms. In addition sensors and actuators could be used in order to “substitute” or provide similar functions of specialized sensors (not deployed or available in a specific area). For instance, using cameras with visual detection can be a substitute for location or tracking capabilities as depicted in **Figure 70**; noise sensors could substitute counting sensors (the level of rumor generated by a crowd can be used to approximately determine the number of users in a location), and the like.

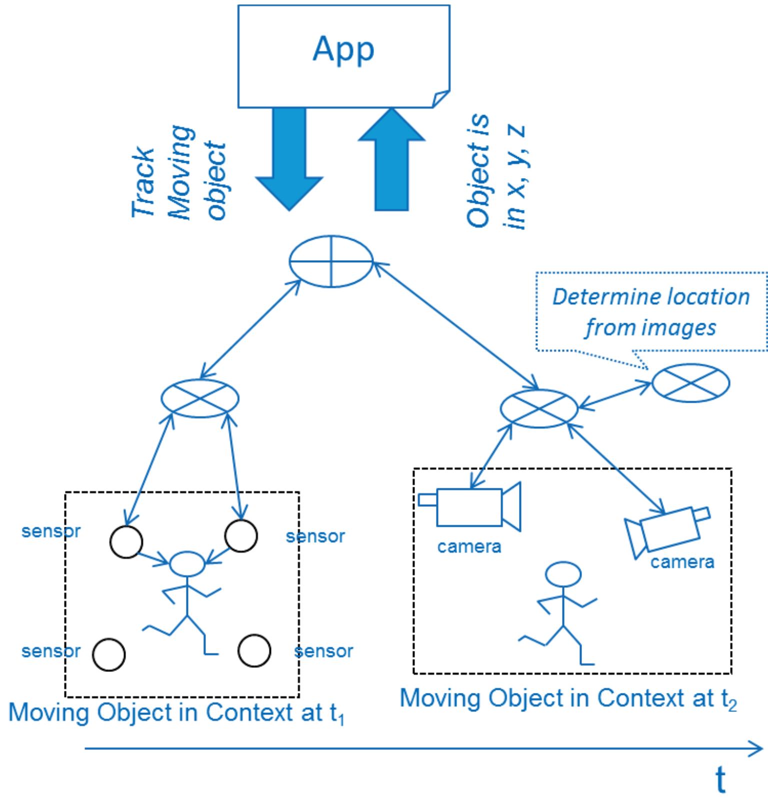


Figure 70: Using Real World Objects to "Deputize for" Missing Sensors

These examples point to the need to integrate sensing capabilities and cognitive behaviors in order to support and open up to a whole wealth of applications. Virtualized objects and their functionalities should be transparently mapped and adapted to provide to applications the expected logical functions. In doing this the IoT platform should be capable of determine whether available resources and their capabilities can be allocated and whether the available functions can be integrated in order to approximate the logical requested functions. These mechanisms are all based on the comprehension of the execution context and on the adaptation, and integration of functions in order to approximate the needed capabilities. Cognition, reasoning and programming of virtual objects are fundamental ingredients for a viable IoT architecture (iCore 2013). **Figure 71** represents how to adapt the available basic functions by means of reasoning and situation awareness to the requests of the application.

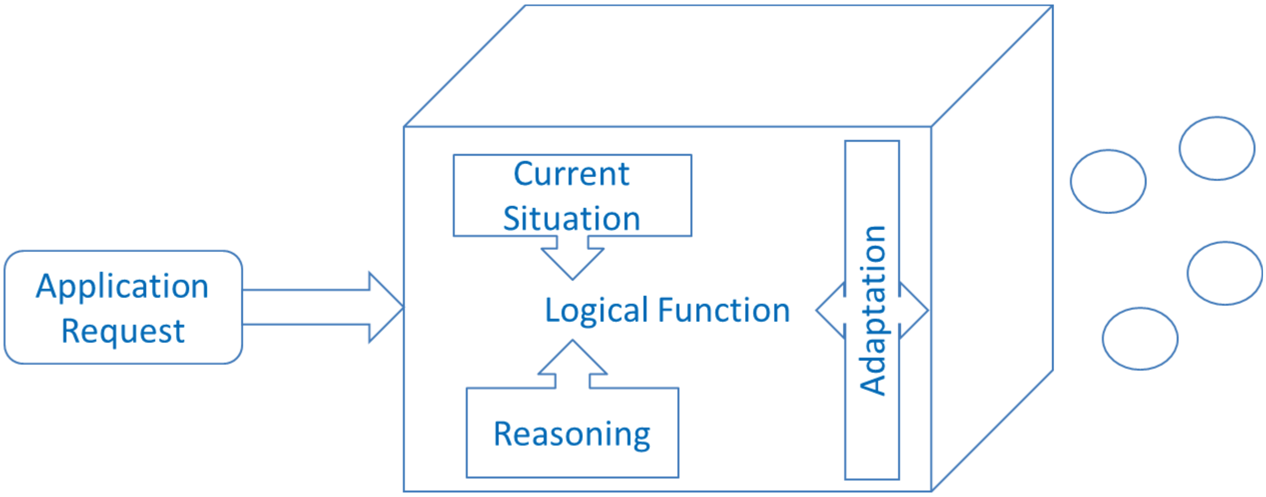


Figure 71: Flexible Adaptation of Functions by Means of Cognition

Actually these mechanisms point to a very crucial aspect of cognitive IoT platforms: the need to maximize the satisfaction of requests of many competing applications, and the need to globally minimize the usage of available systems resources. This pushes for two different perspectives (and goals) of the platform: an application view, i.e., the set of functionalities and mechanisms needed to facilitate the creation and the execution of applications; and a system view, i.e., the set of mechanisms and functions that allow the platform to optimize the allocation and the consumption of resources according to availability and requests. It is interesting to note that both these functional views can be supported by cognitive mechanisms, a set for reasoning (and learning) about the functions needed by applications and another set for reasoning (and learning) about the working of the system itself.

There is an interesting capability to leverage in order to reach a high penetration of sensing features: the crowdsensing (Ganti 2011), i.e., the possibility to use sensors in mobile devices in order to collect and provide data to applications. In spite of many technical issues, crowdsensing is an attempt to overcome one of the major issues in Internet of Things: the wide deployment of sensors. For instance in order to make a city a Smart City, plenty of (wireless) sensor networks should be deployed. Even if the costs of sensors and machinery is rapidly decreasing, the cost of deployment of the infrastructure and its maintenance are still high. Extensive deployments require a lot of investments without the assurance of any economic return. Crowdsensing may be seen as a way to involve users in sharing the data they can collect about a specific environment or a city. At the same time, the individual user has to be rewarded for participating to the data gathering. The collection should be respectful of preferences and policies stated by the users, but in any case, users will be a primal source of data about the environment.

As seen, data gathering will be very heterogeneous and data themselves will be collected in different formats. There is a stringent need to "normalize" these interfaces, layers and data formats in order to create a unique virtual platform on top of which to build applications and services. **Figure 72** (from (iCore 2013)) depicts how adaptation and virtualization can help to build such a homogeneous view on data.

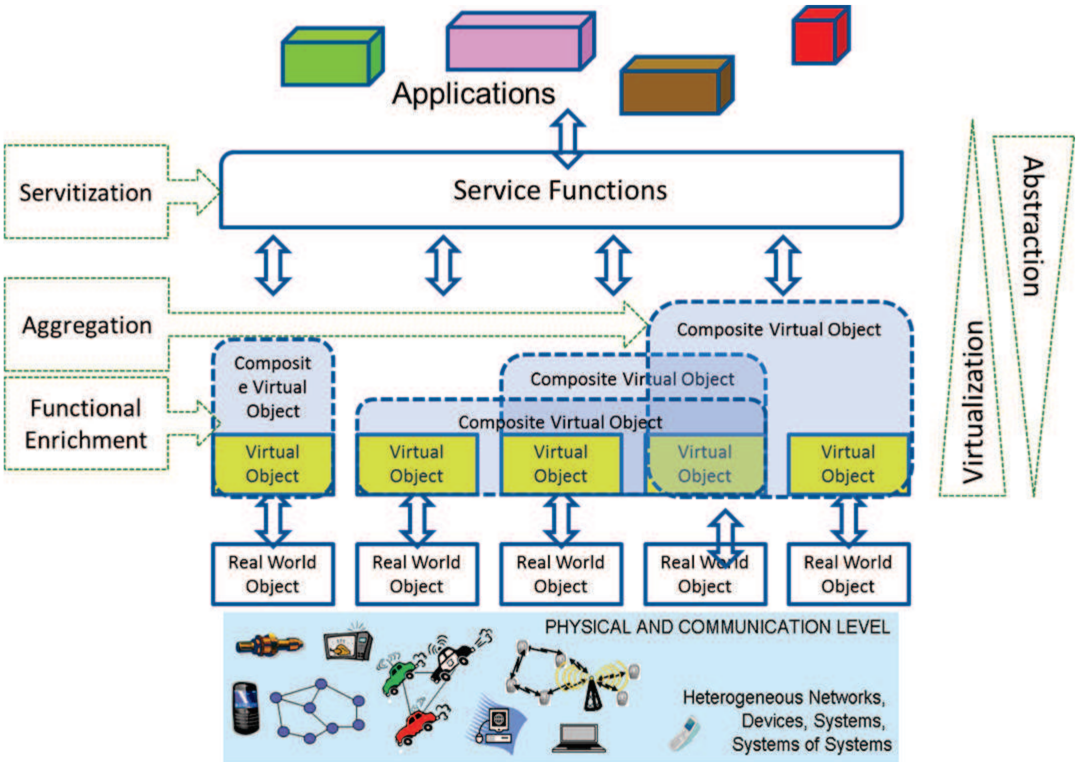


Figure 72: A Generic View of an IoT Platform (iCore Architecture)

It should be noted that aggregation of functionalities leads to abstraction (that generally speaking is a good property for a software system), however this property has to be carefully used because abstraction means also to have a coarse grain control on resources and sometimes applications need to exert a very granular and specific control over resources' functionalities. The issue of abstraction versus granularity has been discussed in [I]. One viable solution is to offer different levels of interfaces and APIs that support different degrees of abstraction. This approach is depicted in **Figure 73**.

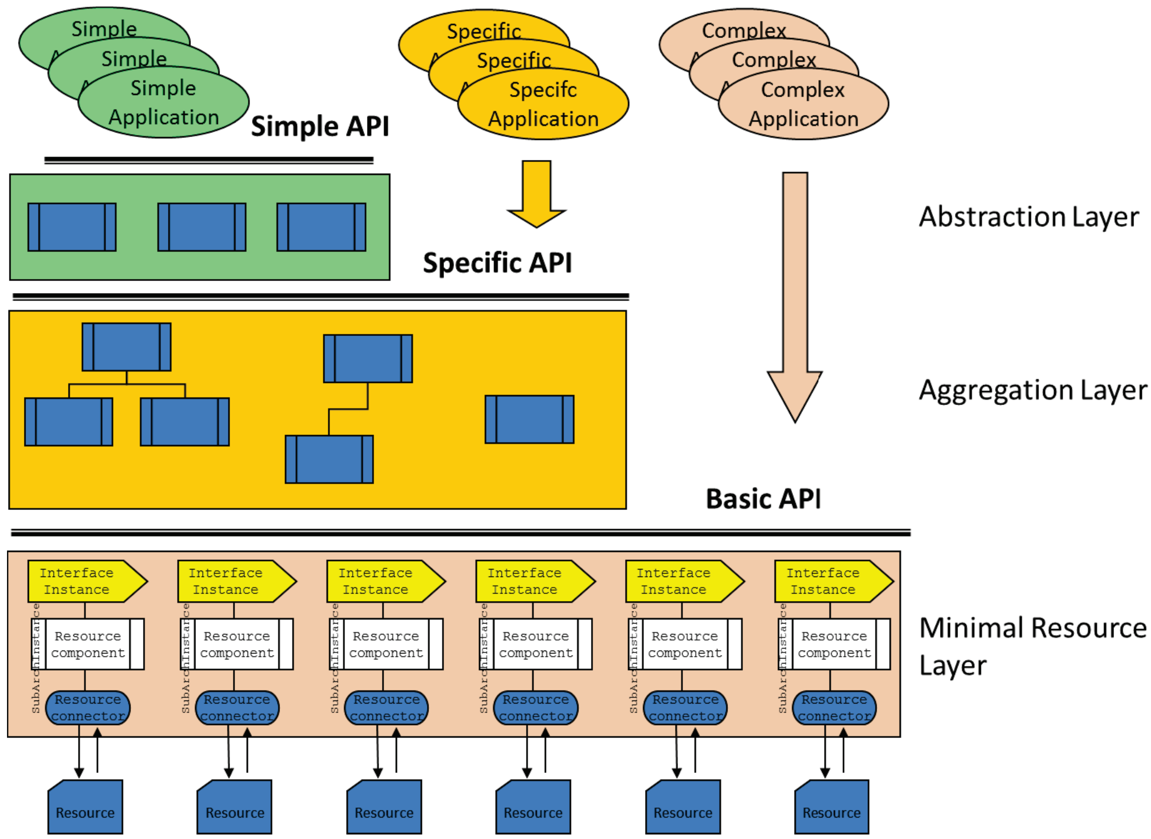


Figure 73: Composition of Interfaces and Programmability.

As a simple rule of thumb, it is important to expose for each resource type all the interfaces it is offering as a stand-alone interfaces (not mixing up different interfaces and/or protocols). These basic interfaces can be integrated and mixed if the services require a simplified and aggregate access to underlying functionalities. At the upmost level, APIs and interfaces can abstract functionalities in favor of simplicity. Simple applications will use simple interfaces, those requiring a granular control over the specific features of the resource will have the possibility to access to them at the cost of dealing with more complexity.

Virtualization, complex systems technologies (e.g. autonomics) and cognition are capabilities that probably will change the meaning of Internet of Things, in fact they allow for a huge transformation: any single object (even physical or immaterial) can be represented in a cloud and being used, mashed up and transformed. In this case it is possible to move from Internet of Things to an Internet **with** Things. To clarify this concept, Internet of Things can be seen as an environment built by the combination of low-cost standard sensors, short range communication, capillary and macro networks, vertical services, data aggregation in cloud, and the possibility through specialized interfaces of third party development. The Internet with Things can be seen as a large horizontal environment comprising “Virtual Objects” that mirror “Things” in cloud, that extend objects with semantics, allow extended data integration and aggregation, support federation and portability of different data formats, and uses the Cloud as a development platform. These differences enables a lot of new applications and domains. Some of those are sketched in the next subsections.

Another interesting aspect of the Internet with things is represented by the capability to deal with streams of data in an efficient way. This can yield to the spread of different paradigms beyond client – server. Actually one communication paradigm that is interesting and is applicable to IoT as well as to social relationships is the Publish/Subscribe (PubSub) one (Tarkoma 2012). The working of PubSub is strongly based on the message passing paradigm. Recipients subscribe to topics/data they are interested in and senders send messages to the queue associated with the topic/datastream. A broker then forwards messages to the subscribers that can use the receive data. **Figure 74** represents this communication paradigm.

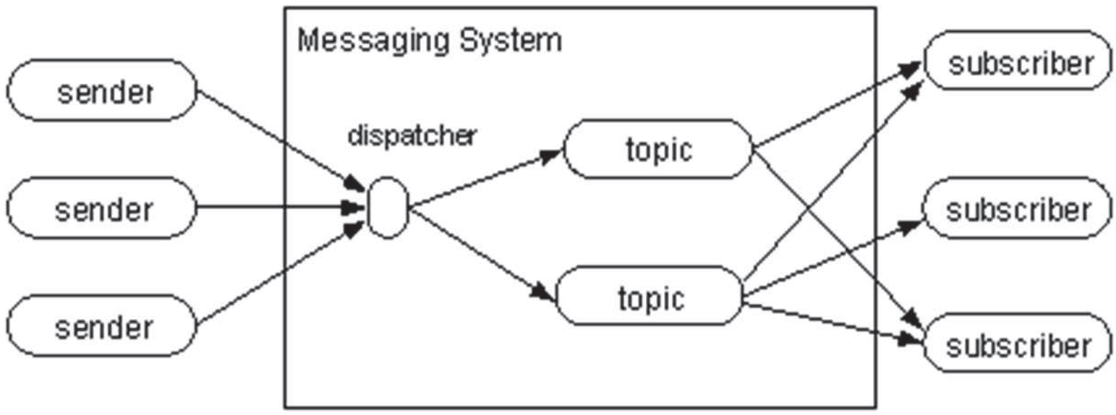


Figure 74: A Typical PubSub Mechanism

Similar mechanisms can be used in order to store data (the queued data can be send to a storage function for instance) in order to have an historical trace of the stream of data that applications can use for some analytics or for log purposes. In addition, the messaging system can try to do some real-time analysis on data passed in order to derive some information. For instance in the case of monitoring sensors, if the messaging system receives many alarms it can try to related them in order not overflow the queues or to correlate them so that different causes can be skimmed out.

Actually this communication paradigm is at the base of Twitter. It is interesting to note that this platform is aiming at sharing information and news within communities, but some developments (e.g., (Kranz 2010)) are using this platform for providing IoT related applications. Two considerations are important:

- The real time web (and in particular the PubSub model) could support new classes of services and enabling new platforms and providers;
- The PubSub mechanisms applied to Internet with Things enable the Brokering role for information exchanged by smart objects and the extraction of relevant information.

Figure 75 represents a possible extension of the PubSub functionalities that recalls in a way Twitter.

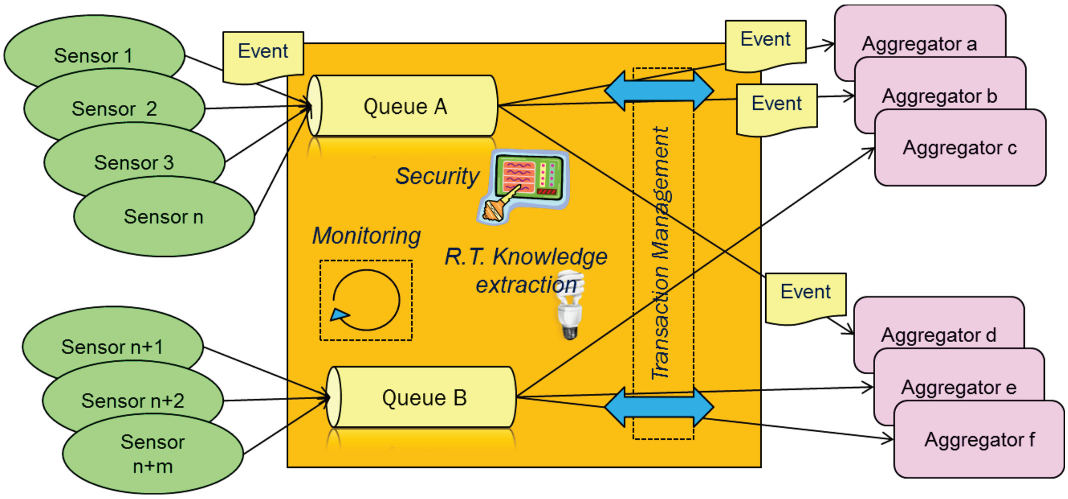


Figure 75: A Transactional Complex Event Processing Derived from Twitter

This platform can be considered a sort of Twitter of Things, because it is focusing on the real-time exchange of information and data between sensors and smart objects and a set of applications. In addition it is a Complex Event Processing engine that can extract knowledge by analyzing the type of data being dealt with. As said, in the Internet of/with Things a fundamental role is the one of Platform Provider, this Twitter of Things platform can be quite important for playing this role and the broker one.

5. 6. 3. Virtualization Enabled Scenarios

Virtualization is a powerful concept that can be used to develop a number of disruptive or at least innovative services. An interesting point is that virtualization can enable servitization, i.e., the capability to transform objects and products into services. On the other hand, virtualization of objects is strongly tighten to Identity Management, i.e., how objects can be related to a user of them and under what conditions and privacy features. Virtual Objects can be created, extended and mashed up by several users that will share the final outcome, each one owning a part of the composed virtual object and the related environment. Creating a synergy between the virtual continuum and the identity management systems offers the possibility to highly personalize an environment to the needs of a specific customer or its community. In the following, two examples of virtualization capabilities are provided.

Virtual Continuum

Virtualization can be exerted to such a level to allow the representation of almost any physical object in the "cloud" . Each clone object could:

- enrich the characteristics of physical devices, in terms of processing, storage, reliability, connectivity, battery usage, etc.;
- break the fences of the walled gardens of device vendors, by creating an alternative/independent service ecosystem;
- enable the virtualization of "physical device ownership", by enabling people to use shared devices/objects, as their private objects (e.g., personalization).

This level of Virtualization is quite interesting from the point of view of the creation of new applications and services, in fact it creates a virtual continuum [XXIV] between real world objects and their representations in a digital environment that can be seen as a bridge between the physical world and cyberspace. Objects can be functionally extended, but the major point is that each physical object, e.g., a product, can become a service. This approach, called Servitization (Baines, et al. 2009), tries to transform any good, product into a service. A product is presented and made available as a set of composable functionalities that the user can activate/deactivate in order to customize the product and to have it to fully satisfy dynamic requirements. In a way, the single product is wrapped in a functional shell that allows to deliver the free and premium functionalities on demand.

The Virtual Continuum is the constant entanglement between real objects and their representation in the network. Events, actions, data on a physical objects will be represented in the virtual world and vice versa. The Virtual Continuum makes possible the close relation between atoms and bits.

Introducing virtualization in the Internet of Things context could have many merits. For instance it could help in overcoming the heterogeneity of many proprietary architectures and systems enabling the possibility to run several proprietary IoT applications into a single platform. Virtualization of sensors and actuators is also important because it allows to represent real world objects into the network. This enables the possibility of representing and programming a real world object in the iCore platform and to control/govern/integrate the virtual instance in a programmable environment. Virtualization is the first step towards the virtual continuum, i.e., the possibility to create a link between any physical object and its representation in the cloud. The virtualized object can extend the functionalities, the features and the capabilities offered by the real one. For instance a real world object, a car, can be virtualized in the cloud and its functions can be controlled and managed in the "internet". The car functions can be extended by applications and services running in the cloud. The virtualized car can be enriched with control functionalities (e.g., for a trip in the mountains, the driver can enable the 4wheel drive capability for a couple of days by interacting with the car and paying for that functionality for the required time. Additional power can be bought by means of the virtualized representation of the car for a specific travel, and so on). The car is not anymore a product sold to a client, but a service that can enable and disable premium functionalities depending on the needs of users. This is an example of how virtualization can support Servitization. Different objects could represent a single real object allowing for sharing of its functionalities in different virtual environments. The same physical sensor can be virtualized and framed into different contexts of usage. Valuable resources can be virtualized and shared in such a way to increase their usage and to help in limiting the wide deployment of similar sensors by different providers. On the other side valuable functions can be derived by virtualizing the real sensor capabilities: e.g., a camera can be used as a location device, a noise sensor can be used to determining the number of people present in a street, and so forth. Virtualization allows to derive usable functions by many heterogeneous devices deployed in the real world.

This virtual continuum (**Figure 76**) can be seen as a sort of platform that enables the migration of control from the device to the cloud. In this way the "ownership" moves from the device producers to cloud providers and it allows users to benefit from a larger application ecosystems.

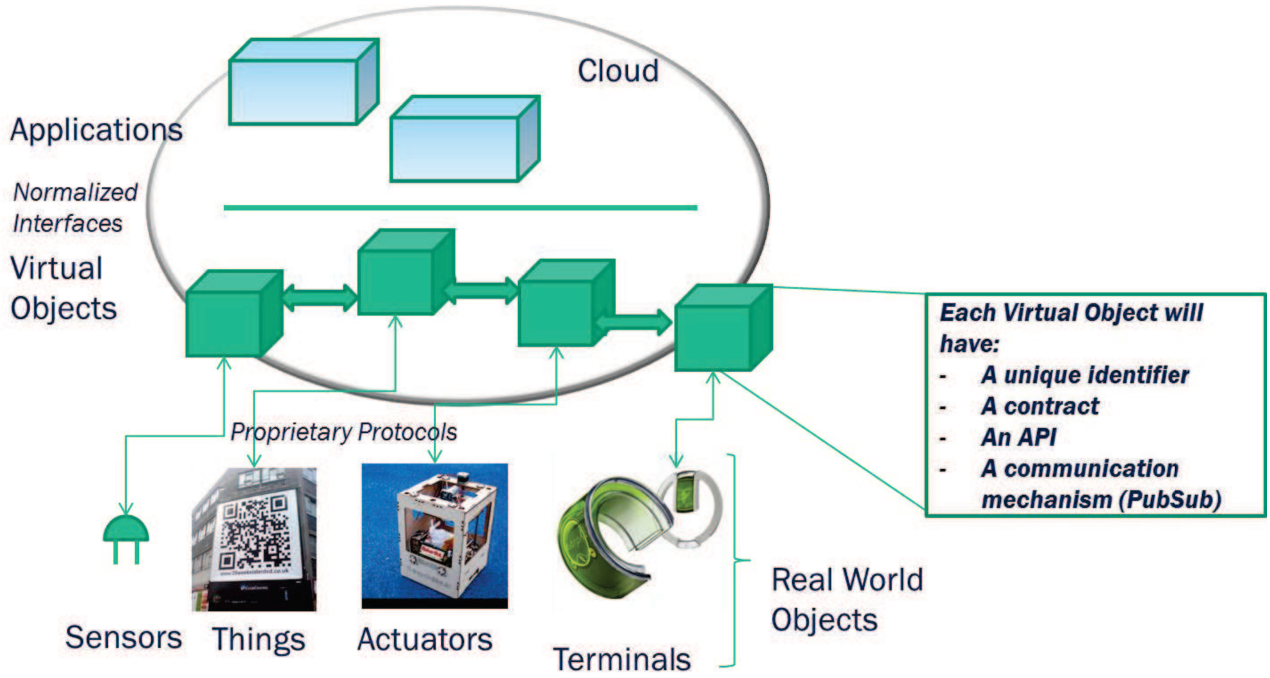


Figure 76: A Representation of the Virtual Continuum

Some features of the Virtual Continuum are:

- Decoupling of physical objects and virtual ones
- Normalized applications can be built on Virtual Objects (having an impact on Physical Objects)
- Plenty of Applications can be built on top of virtualized and extended objects
- Leverage of Cloud infrastructure

Virtual Terminal

Another example of the possibilities introduced and offered by virtualization capabilities is the Virtual Terminal paradigm [XXIV]: a terminal can be cloned in the cloud and it can act on behalf of the physical terminal. It can also be "composed" by the integration of several physical terminals. Enabling such a possibility could free the user from the need to run applications within a proprietary and confined walled garden. An example of this trend is given by the Silk solution developed by Amazon: the terminal and the server can dynamically determine (considering the network conditions and the terminal capabilities and status) where is the right place to execute relevant functions. Sometimes the cloud can take over the processing burden, some other the processing is entirely executed in the terminal.

The concept of virtualization of mobile phones is getting traction worldwide. There are already some initiatives related to the definition of Virtual Terminals and Virtual Environments. For instance the Clone Cloud project within Intel Research Center (Chun and Maniatis 2009), (Chun, Ihm, et al. 2011) aims at "clone the entire set of data and applications from the smart-phone onto the cloud and selectively execute some operations on the clones, reintegrating the results back into the smart-phone". Also NTT is working on a Virtual Terminal (Itoh, Chen and Kusumoto 2010) with the idea of offloading many time consuming processing tasks from the terminal into the network .

A Virtual Terminal is a software feature that provides a perfect mirror, constantly updated, of all the actions, data and occurrences of a specific terminal or a group of them. A Virtual Environment is a software feature that allows customers to use an entire (virtualized) computational and communication environment tailored to their specific needs.

The concept behind these projects is very simple: to integrate data and execution environments of the terminal into the cloud, i.e., to provide a functional extension of the terminal capabilities into the cloud (see **Figure 77**).

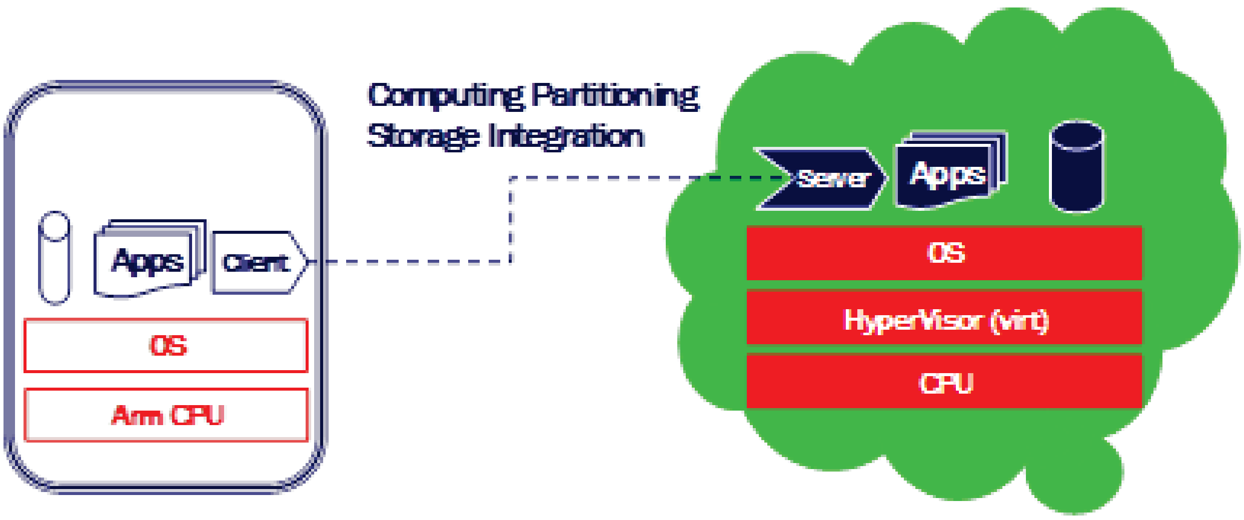


Figure 77: Virtual Terminal as a Functional Enrichment of a Terminal

The objective of this initiative is to drive the definition of a platform capable to a) support the virtualization of terminals and other smart objects, b) to allow the TelCos to play the role of aggregator and broker of virtualized (communication, processing, storage) resources and applications; c) to allow the creation and management of virtual environments tailored to the specific needs of a customer. **Figure 78** sketches the problem domain.

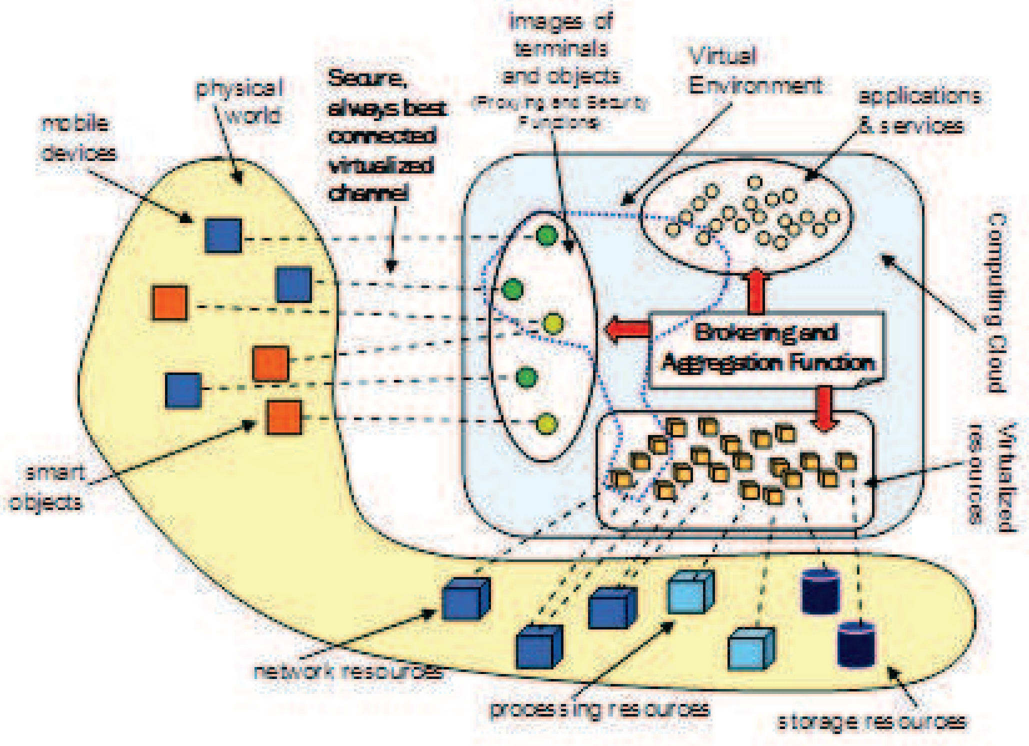


Figure 78: Towards a Platform for Virtual Environments

Some Examples

Moving the SIMs into the virtual terminal. A user could map one or more physical terminals onto a (subscribed) virtual terminal in order to synchronize them or to extend their capabilities. The user decides how to deal with communications and processing services, however the different terminals can create a sort of mesh network and can partition information and computational tasks among them. See **Figure 79**.

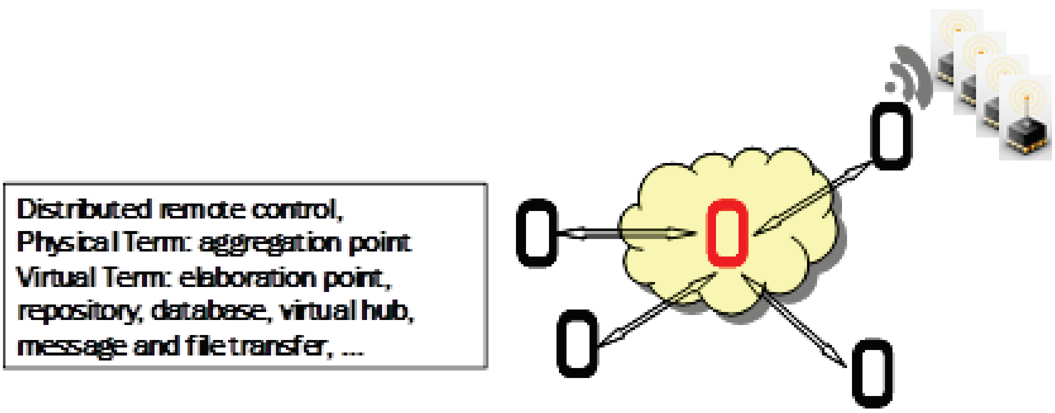


Figure 79: 1-to-n Mapping

Another situation is depicted in **Figure 80**: a user has many virtual terminal subscriptions and it associates the unique physical terminal to all its virtual images in order to integrate the services into a single end-point.

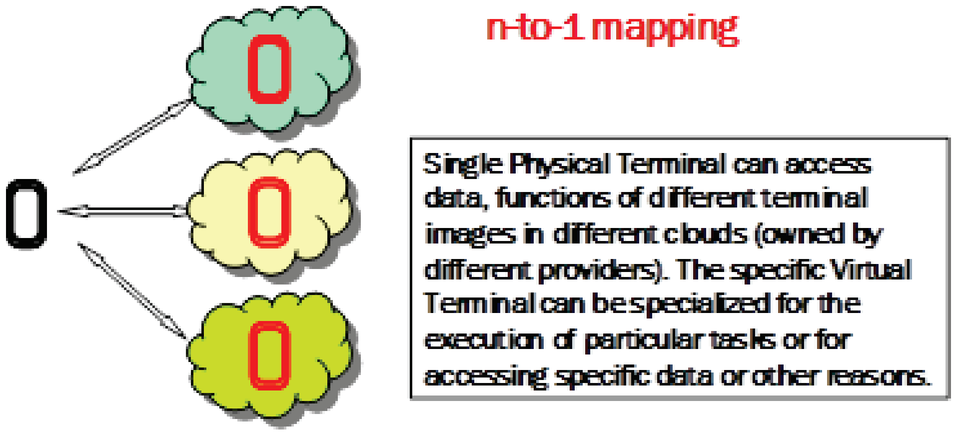


Figure 80: n-to-1 Mapping

My trip scenario. A person can create a virtual object called “my trip” that represents all the needed information related to a travel. This object is created in the virtual environment. This object comprises the set of information, data, and alerts needed to make the trip easier or more rewarding for the user. The user first visit virtually by means of a street view, the destination. S/he bookmarks some points of interest with alerts. All the booking and tickets are aggregated to the virtual object “my trip” (aggregation of information). At the check-in the user has just to share with the clerk a pointer to the ticket and gets in return another object (the check-in object) to associate with the “my trip” object. At the destination the user can start wandering and can ask the navigation support to “my trip” and can get also some useful information about monuments, restaurant and the like. This instance of “my trip” (plus some addition location information) can be stored in the system. Old information can be used to recall the user of things done or seen in that place.

5. 7. Findings

This chapter has presented some opportunities for providing new services. Other services are considered as well by the ICT industry. These services have been considered in order to give feedbacks on traditional services (as the rich communication ones), on current proposals of many TelCos (like the cloud computing) and on possible future opportunities like personal data, smart environments (many Operators are starting to work on smart cities) and Internet of Things (all the major TelCos are trying to understand how to move from machine to machine services to Internet of Things). Some of them will be further considered in Section 6. 5. as examples of new opportunities that “smart operators” can exploit in order to open up and take the leadership in new markets. Other services are newer and deserve more study and experimentation in order to find out whether they are actually viable.

Communication services will not provide considerable improvements with respect to the current situation, the RCS and even the WebRTC based services do not offer anything that is new for the users. They are a collection of already available functions that the TelCos are making more interoperable between their domains. In reality users have already interoperable communications services, in fact each major WebCo is offering these services. They allow the user to communicate within very large communities (think to Facebook) and interoperability is reached by switching to a different applications. Actually the value in rich communication services is not too much in the communications functions, it is in the community, ,i.e., how many people can be reached by that application/service. This is an important lesson to keep in mind. Many Operators still consider interoperability a major value, but in reality it is not if the community associated to the service is big enough. The revenues generated by these services will not generally be high and possibly they will not even compensate for the investment put forward to create a service delivery solution. Probably a better approach is to compete in terms of tariffing on existing capabilities.

Many TelCos consider Cloud Computing and the provision of XaaS functionalities as a major business opportunity. As discussed in this Chapter, however, the Cloud market is already controlled by other Actors (like the WebCos) and the role for Operators is restricted to certain types of services and functionalities. A first insight is that there is the need to clearly differentiate the markets: residential is a very competitive one. The Business market is large and is very valuable but it is also demanding in terms of attention the customers. It is also difficult because customers are more used to determine the best and more competitive offering and they can switch from one provider to the other very easily. In addition, issues posed by this type of customers could be challenging. For example the local cloud provider could be different country by country. Providing an interoperable federated cloud could be very difficult and the revenues generated by it have to be shared between different partners. From a technical perspective the adopted solutions will lag behind those deployed and offered by the WebCos, And so customers could continuously ask for improvements.

The newer services offer some possibilities: the Bank of user data can be appealing if a vast majority of user will recognize the need to be protected and not exploited by others that do manage the user personal data. Another viable way of proposing the service is trying to return to the user a part of the revenue generated by selling the user data. An important point that deserves a lot of attention is the identification of parameters and criteria to evaluate the effective value of the data of a single user and the aggregated value of data of several users in the network. However this service seems to be important because is based on a real user centered principle: the data pertain to the users and they have to be managed according to rules and policy determined by the user itself. Even if there is no business value behind it, this is an important indication that consumer associations and authorities worldwide should carefully consider. As previously seen, personal data and identity are two faces of the same coin and so similar issues related to identity ownership can be put forwards. They will be important issues to be solved at the technical level as well as at economic and regulation levels. A few of these issues are considered in Chapter 6. and in the most disruptive scenario of Chapter 8.

Smart environments seems to be the natural evolution of the integration of Telecommunications and IT. It is important to start very soon in experimenting, providing and supporting this class of services in order to grasp the user requirements and to provide captivating services. In this area, new interaction paradigms are more important than ever in order to engage the users. In addition, the large distribution of functions and the fragmentation of devices will oblige for distributed solutions adopting new paradigms. For instance the PubSub one can be a good one in this context in order to forward events to a set of subscriber processes that need to react to the occurrence of an event. P2P solutions are a natural choice for this type of services, but they do not offer standardized solutions and in addition they need a lot of programming. Once more there is a compelling push to Operators to enter into different levels of mastering of software. The same issue is of capital importance for some scenarios like in Section 8. 8. , Section 8. 9. , and Section 8. 10.

Internet with Things is the most appealing service scenario for the TelCos. It is a broad opportunity falling over several markets and industries. It has impacts on many processes and for this reason can have an economic relevance for many customers (improving the efficiency of some business processes is a viaticum to success). This application domain allows also to introduce disruption in existing providers and it also can be approached with new interaction and control paradigms. In other terms it is a market in an initial state and the Provider that will identify and provide the best technologies has a good chance to be an important player in a rich market of the future. TelCos can use it also instrumentally for experimenting new paradigms that displace the WebCos (e.g., P2P propositions) and play as technological forerunners for a first time since a long period. In addition Internet with Things has a large Business market characterization that can be exploited by the TelCos in order to start to position themselves with respect to WebCos. The virtual continuum create also the possibility to virtualize terminals and could offer the opportunity to reduce the power of terminal Vendors with respect to Operators. Users can be offered virtualized and interoperable environments that free them from strong dependences from specific terminals and operating systems. Once more a user centric approach could be a viable way to revert compromised business positions in favor of new ones more open and favorable to the users.

Different paradigms have been considered in the services described in this Chapter. Some of them are well known and fall in the basic ones (i.e., C – S, NI, and P2P). However different models are emerging especially in the realm of Smart Environments and IoT. In these cases there is the need to deal with events in order to allow a granular control of functionalities to applications and the different stakeholders of the service. In addition, these services will comprise many different resources interacting in several ways each other. There is the need to organize resources and their interaction in such a way that autonomic behaviors can be enforced for limiting the human intervention. These services need to be supported by cognitive solutions (as those described in Section 4. 6.) that will make simpler to program and use newer service architectures (see for instance the scenarios of Section 8. 8. , Section 8. 9. , Section 8. 10. , and to a larger extend Chapter 9.). One of major finding is that services have to be provided with appropriated solutions that do not require a lot of adaptation of the architecture in order to represent the service interactions, but exactly the opposite: the service architecture has to be chosen according to its capabilities (its expressive power) to represent and support in the easiest possible way the interaction and control needs of new services. Flexibility of platforms and architectures, the ability to encompass new control paradigms are major advantages of future solutions.

Internet with Things and smart environments can be considered as inflection points in which new technologies open the path to new markets and approaches. These classes of services can be considered by TelCos as means to enter into processes and markets with a winning proposal. Chapter 6. will further elaborate on how to enter into adjacent markets with innovative approaches. Scenarios in Chapter 8. will use some of these findings in order to show what roles can be played by Smart Operators.

Part 2: Ecosystems and the Network Value

Introduction

This part of the thesis pays attention to the changes that the technological evolution can bring to the telecoms ecosystem and to the major asset of the TelCos: the network. This part clarifies that the term Over the Top, OTT, so frequently used in the literature and by Operators is misleading and it represents and old conception of the Business relationships. In addition some Network paradoxes are identified and discussed (for an agile reading the interested person can refer to[XXXI] and [XXXII]). The difficulties of TelCos in effectively entering into adjacent markets are scrutinized and discussed together with the contraposition of the Lean and Smart Operator. The aims is to identify if the Network Asset is still a value to leverage in the competition with other actors of the ICT contexts.

In this part of thesis, the focus is on the Tussles and the conflicts going on in the telecoms industry. In Chapter 6. , the current struggle of the TelCos with new comers is represented and discussed as well as the possible opportunities that the TelCo can seize in a changing market. In Chapter 7. , the current reaction of the TelCo to the competition with Web Companies is represented, i.e., how the TelCos are leveraging their assets.

Chapter 6. A Tussle of Business Models and the Positioning of ICT Players

Highlights

- The “Value” of the Network
- Network Paradoxes
- Opportunities and Applications
- Comparison between TelCos and Other Competitors

6. 1. Introduction

The growth of the Internet has led to a new ecosystem of players very different from the traditional one in which the TelCos were used to play. The success of very innovative Web Companies like Google and Facebook and their dominance in specific service classes has displaced the TelCos in such a way that often these companies are seen as the most dangerous competitors.

Innovative WebCos have brought to the market a large number of new (and old) services offering to users the capability to use them for very low prices or even for free. Free services are made possible because their provision is guaranteed by other business models. In many cases, services are not the goal of the business (like for many TelCos’ services) but a means to engage the users and to get from them information that can be used in order to profile the users. Users are becoming the “goods”. In this new context the competition is essentially based on the leverage of assets of the different competitors: simply put, the Web Companies push for leveraging the Server infrastructure, the TelCos want to leverage the network and the terminal providers are enforcing the value of terminals. This struggle is represented in **Figure 81**. The client - server paradigm is used in order to describe the segmentation of the market of ICT services and to exemplify the forces at play, in fact in the picture some major Competitors are represented together to their moves towards different segments.

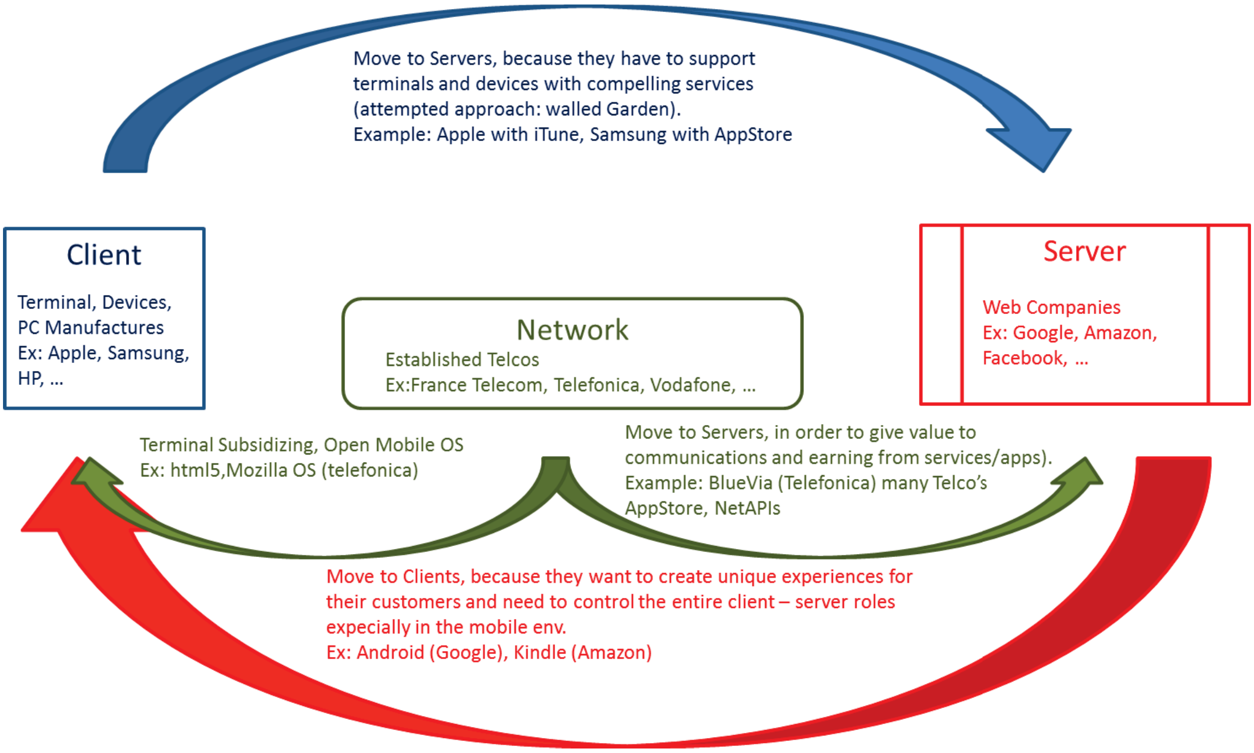


Figure 81: Internet Market Positioning of some Actors

It is evident (from **Figure 81**) that the role of TelCos is less relevant that the others: the network is seen as a pipe and all the value is moving to the edges. Controlling the Clients and the Servers gives a lot of power to the companies that are able to play this global role. In addition this approach enables (it is based upon it) the possibility to create "walled gardens" that are not interoperable with other environments (especially those of the competitors). This is why many web companies are trying to greatly differentiating themselves even when they use the same infrastructure (e.g., the case of Amazon usage of Android).

In this competition the value of the network is greatly revisited and if it is analyzed under the "end to end principle" (Clark, et al. 2005) it is going to shrink and be reduced to the value of a commodity. The End to end principle states that: "mechanisms should not be enforced in the network if they can be deployed at end nodes, and that the core of the network should provide general services, not those tailored to specific applications". This has led to the concept of the Stupid Network (Isenberg 1998) and to the even stronger "Best

Network Paradox[33]" which states that "*The best network just moves bits and the best network is the hardest one to make money running*". Actually these discussions bring to an important point: what is the value of the network in this new context ?

Before entering in some detail in this topic, it is worth to mention that the Stupid Network assumption are not all correct in all their typical arguments. For instance it is not true that the E2E principle means necessarily the Network Neutrality (see next section), in fact the differentiation of traffic flows at the edge would be meaningless while differentiated Quality of Service could be seen as a network service offered to edge applications. The ideas that the Network is Stupid (it is just a pipe) and that all the Intelligence is only at the edges do not hold true. The network can be extremely intelligent (e.g., adaptive and cognitive networks) and focusing on specialized tasks, e.g., optimized transport functions. What is missing (and this is a big issue) in the Internet today is the possibility for the applications to require to the network a set of functions useful for help the task of the application (e.g., Content Delivery Network functions), the clear separation and decoupling between the applications and the network has to be overcame in favor of a cooperation between the two levels. An interface between the two has to be defined.

This Chapter will tackle the issues related to the value of the network, i.e., how much it is worth and how the value can sustain investments in new infrastructures. In this section some paradoxes are identified and they seem to point to the fact that a structural change is surfacing. Paradoxes and inflection points are two forces that can change the way the market is addressed by different Actors, a part of this chapter is then devoted to show how the competition is forcing the TelCos to change their role. In this context it is extremely important to re-consider how WebCo are approaching the market and how they are positioned (over the top or at the edge of the network). Another related factor is the capability to properly execute strategies and the deployment of new offering by some of the major Actors in the ICT field.

Many Operators are considering the possibility to leverage the connectivity in order to enter into adjacent markets. In order to be successful in this move, there is the need to bring value to users. Section 6. 4. will discuss the rationale and the possibilities to enter into new markets with winning propositions. In this context an Operator could be implementing a lean strategy (i.e., reducing costs and OPEX) or could be implementing a smart strategy

[33] See http://netparadox.com/ last accessed 31st March 2013

(i.e., entering into new “big” markets as those described in previous section, e.g., the Internet of Things). A lean strategy tries to preserve the status quo will a smart strategy is more aggressive and it is trying to change the situation.

6. 2. The Value of the Network

A tough quarrel has been started in order to re-balance the market power and to take advantage of the assets of each of the competitors. This implies a different perception of the “value of the network”. The most evident effect of this dispute is the network neutrality one. According to (Net Neutrality, Wikipedia 2013): “*Net neutrality (also network neutrality or Internet neutrality) is the principle that Internet service providers and governments should treat all data on the Internet equally, not discriminating or charging differentially by user, content, site, platform, application, type of attached equipment, and modes of communication*”. This issue is not to be dealt with in here[34], it is mentioned in order to highlight the struggle between different position in the ICT industry with respect to the vision of the network and, consequently, its value. Another issue that recursively occurs is related to the Internet economy with specific respect to the Internet Transit and Peering Models. As stated in (DrPeering 2010), the Transit fees are approaching to 0$/Mbit p second, and peering costs have to take into account this new reality. The prediction is that the Internet interconnectivity models will drastically change in favor of private peering or transit. According to (DrPeering 2010) also the CDN industry will drastically change in order to adapt to the lower cost of transit. This will bring even more value to the “access networks” that are the means of users to access to the internet. The access networks seem to be the last resort for many Operators, i.e., the possibility to charge for accessing to the Internet services. However the Capex in the access networks have to find a reasonable return in investment and this is a difficult point due to the decreasing revenues generated today by the networks. It seems to be a sort of closed circuit in which the value of the network is progressively decreasing. In addition, a variety of technologies (especially for the mobile access) seem to be available and they do not fall necessarily only in the domain of the Operator. As clearly explained in (Paolini 2012), small size cells seem to have a more

[34] It is opinion of the Author, however, that the network neutrality should consider aspects related to the “tragedy of Commons”: a non-regulated usage of the network and its resources could be detrimental for the entire ecosystem especially in the case in which a few bandwidth hogs (consumers of high bandwidth) allocate much of the resources devoted to large communities. Some regulation beyond best effort should be put in place in order to guarantee a fair usage of available resources. This should be especially considered in cases in which communication resources are scarce.

advantageous "Total Cost of Ownership" and this is especially true for WiFi. The investments on the network access in this case should carefully balance the minor structural investment (Capex) due to the lower costs of small cells with the Opex needed to manage and operate this more distributed infrastructure. **Figure 82** (courtesy of SenzaFili Consulting) shows clearly the difference in terms of Total Cost of Ownership.

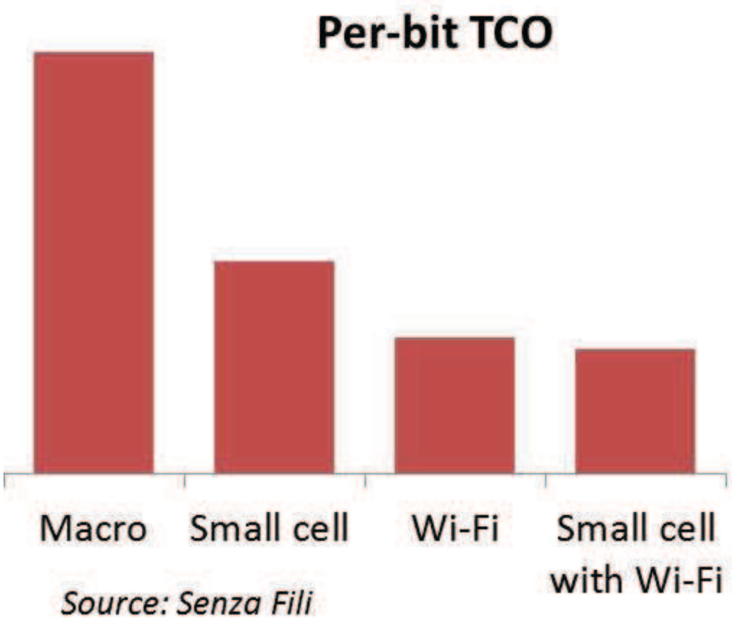


Figure 82: Total Cost of Ownership (per Bit) in Access Networks

The possibility to use WiFi, however, introduces the possibility for many alternative Operators to enter somehow in the wireless business[35]. One example is FON, they allow users to share bandwidth of individual access point as a sort of offload for other networks. They are cooperating with large Operators (e.g., BT) in order to create a vast infrastructure.

On the fixed side, the fiber access network of Google[36] in Kansas City is an example of how new competitors can leverage their assets in order to displace the traditional business of the Operators. Google Fiber exemplifies how the fiber access can be used in order to leverage the Web Company assets, in fact the offering is a full bundle of connectivity, storage capabilities (up to 1 TByte in the cloud and 2 TByte locally) plus video services (and related Boxes).

[35] In this document, the issues related to the spectrum management and technologies like software define and Cognitive Radio are left aside, but undeniably they have the potential to greatly impact the telecommunication structure if the Authorities will open to the possibilities offered by the technologies.

[36] See https://fiber.google.com/about/]

6. 2. 1. How Competition is Evolving and Roles of TelCos and Web Companies

In many countries, there has been a continuous effort to introduce higher levels of competition between Actors in the communications sector. One example is FCC[37], but essentially all the National Authorities in developed Countries have adopted similar approaches. This has brought to a fierce competition between Operators (especially in the Mobile segment) and users have benefitted of reduced tariffs and improved services. There is a sort of spiraling effect that is pushing (mobile) Operators to reduce tariffs in order to attract new customers. This has also increased the churn of users that are moving from one Operator to the other in order to benefit from the latest offers. The net effect is that the entire Telecommunication industry has competed on communications prices. On the other side, the service segment has been left practically untouched and totally allotted to the Web Companies, that have entered in this area taking advantage of the IP technologies and the related “end to end principle”. Essentially the Web Companies have taken advantage of the fight between Operators in order to take a grip on the less valuable market of services. This market is less valuable because applications are bought for a few dollars and then used for a long period. However this market is interesting for them because it is synergistic with their main business model.

The current Business Model of Operators is depicted in **Figure 83**.

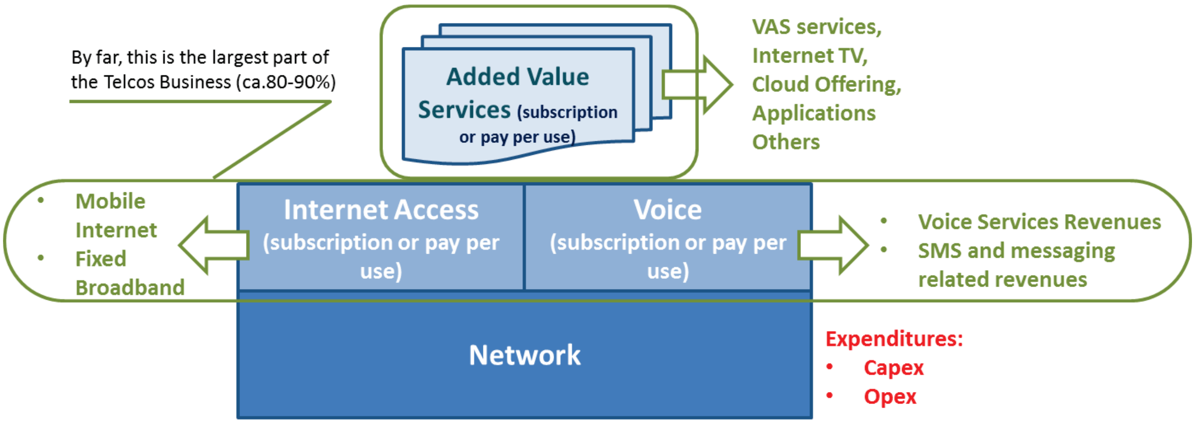


Figure 83: Current Business Models of TelCos

Two evidences emerge:

[37] See http://transition.fcc.gov/connectglobe/sec5.html last accessed 31st March 2013

1. the gross of the revenues come from Voice (with revenues declining) and broadband Internet access (on the increase, but not enough to compensate the Voice loss)
2. the service revenue are very low compared to the traditional ones, i.e., Revenues (Network) >> Revenues (Services)

This business model will keep valid until this equation is kept valid: Revenues (Network) + Revenue (Services) >= Capex + Opex. When this will not hold true, there is a serious risk for the Business. Actually the margin of the Telecommunications Industry are still high (simply put: the difference between Revenues (from network services and added services) and the sum of Opex and Capex is positive and higher than 10%). A large part of the Opex is due to site lease and backhaul (in mobile networks) while the major individual source of Capital Expenditure is due to RAN infrastructure. In comparison, the Core network infrastructure cost is almost negligible[38].

Some Operators are entering into the logic of trying to compensate the losses by means of services, however the percentage of the annual decline of voice revenues is very high (from 2% to 5% and more per year). For large companies, compensating this loss with the revenues generated by services and applications can be a daunting task especially for Operators that do not have an international footprint. The basic business models applied by Operators are the subscription and the "pay per use" models. They are very aligned to the idea that communications capabilities are a sort of commodity to pay for in order to access to other services. Under this respect, Amazon and AT&T have joined in order to support the embedded communication approach adopted by Kindle: communications means are already paid for the users that stay in the Amazon Walled Garden.

The business models of Web Companies are different and more diversified. Here two different business model will be briefly introduced: the Google one and the Amazon one.

Google bases its revenues essentially on the advertising. Its two most revenue generating applications are AdSense and AdWords. These two applications are organized in such a way to automatically conduct auction between advertisers and to optimize their exposure to users according to the search being queried and the user profile, as well as the monthly budget of the advertiser and its maximum cap per single bid. This is done by means of very secret algorithms and solutions. All the business of Google is built around these two revenue engines.

---

[38] See for instance http://techneconomyblog.com/tag/opex/

Figure 84 depicts the Google Business Model (Minerva and Demaria 2006). Many services are offered for free in order to gather personal information of the users and to further detail their personal profile. Some premium services are also offered in order to satisfy the requests of specific users that need more service functionalities or simply more processing or storage capabilities. This combination of free and premium offering is now well-known under the name of Freemium Business Model (Teece 2010).

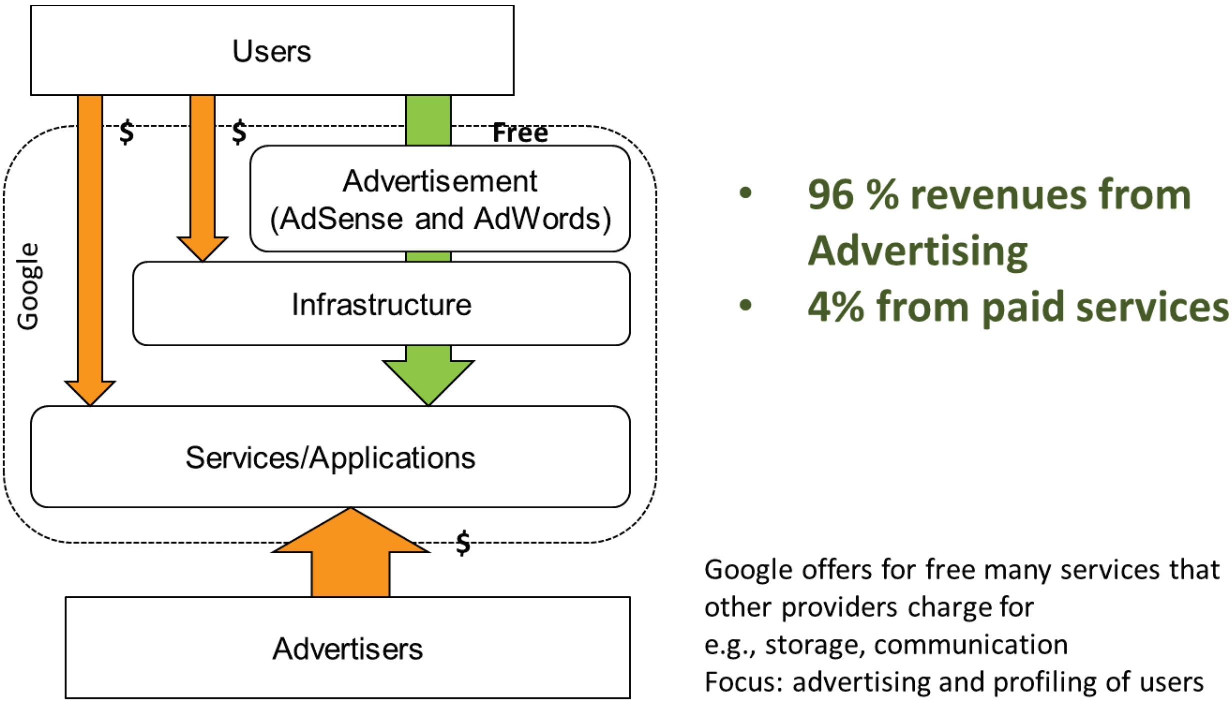


Figure 84: Google Business Model

The largest part of the Google Revenues comes from Advertisers, while the selling of functions or apps covers just a small portion of the multi-billion dollars business of Google.

Amazon aims at being the global market place for the Internet, i.e., the preferred shop for buying goods or products. In addition Amazon has leveraged its infrastructure introducing the cloud computing capabilities. It is a developer and advocate of the Infrastructure as a Service (IaaS) paradigm. In fact a relevant part of the revenues comes from the selling of processing and storage capabilities. **Figure 85** depicts the approach to Business of Amazon.

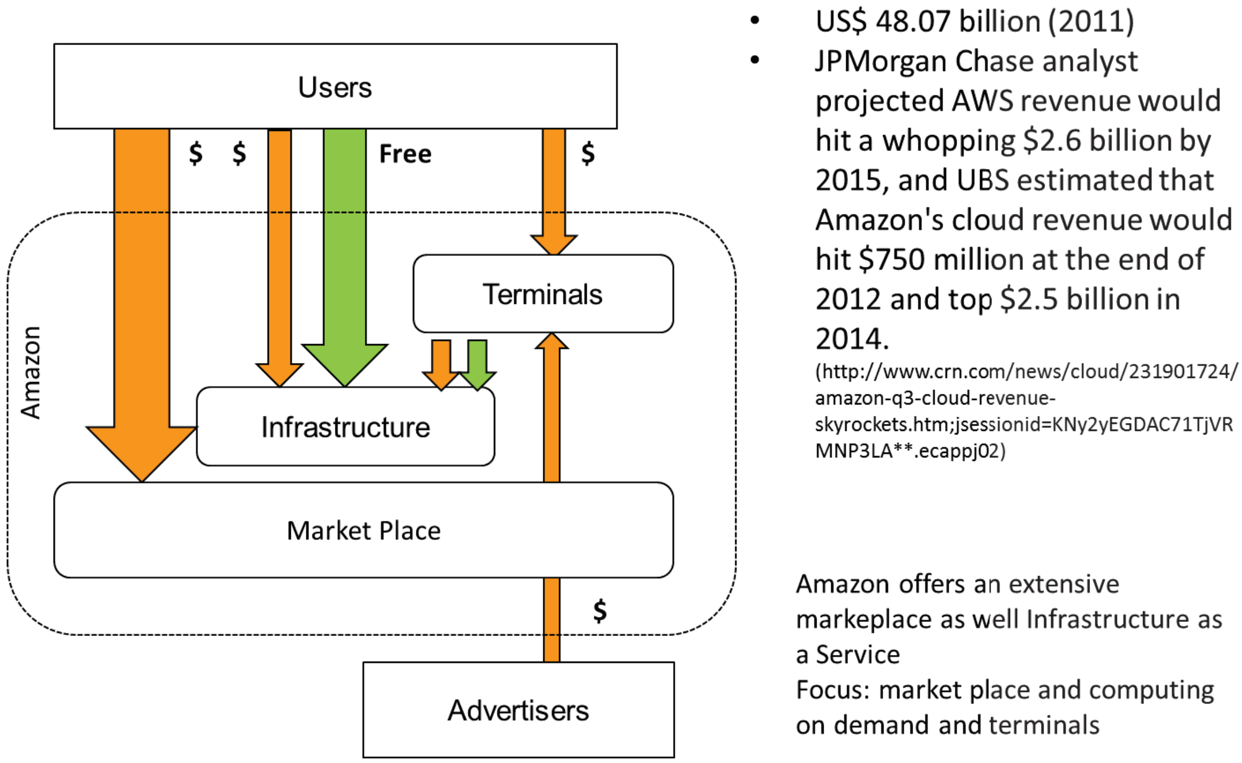


Figure 85: Amazon Business Model

Amazon is also leveraging its device offering in order to push customers to use its offering. Terminals are offered at a reduced price, if users accept advertising of products sold in the Amazon Marketplace. The terminals themselves are using the cloud capabilities of Amazon in order to improve the quality of experience offered to users (e.g., the Silk browser and the related technology). The leitmotiv behind this integrated approach is to keep the customers as much as possible inside the Amazon Walled Garden.

It is evident so far that the approach of Web Companies is strongly based on the leverage of a main business model and the need to confine the customers into a specific controlled environment. Unfortunately this approach is much appreciated by users and neglected by Regulators, in fact there is no initiative to free the App Markets places (a sort of Market Place Neutrality) and also the big issue of the gathering and ownership of personal data is greatly left to a few visionaries people.

6. 2. 2. Why Over The Top (OTT) Are Not OTT

In the ICT industry, the term Over the Top is often used in order to define companies like Google, Amazon and others that offer services in s strong competition with TelCos. The idea behind this name is that these companies are using the network to their own advantage and they "sit" on top of it as represented in **Figure 86**.

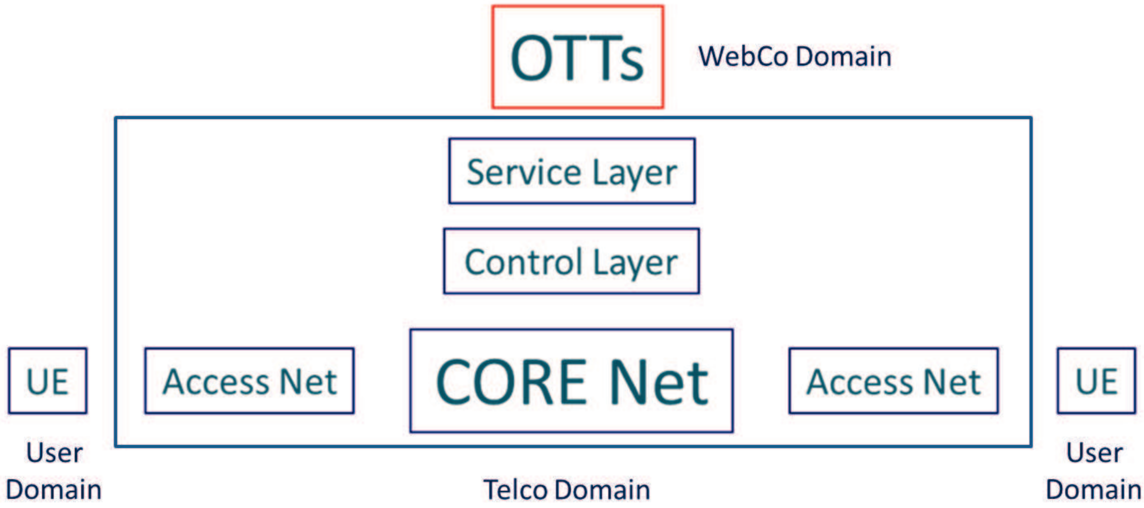


Figure 86: The Concept of Over The Top (OTT)

This concept pinpoints also to the concept of parasites, the Web Companies are seen as exploiter of the huge investments of TelCos. WebCos are making a lot of money exploiting for very low prices of connectivity without returning to TelCos enough value. Actually each WebCo pays a huge amount of money to the large connectivity providers in order to get an Internet access and some of them play relevant money to Operators in order to use their data center infrastructure (for instance Twitter with NTT Data). The issue of the usage (or exploitation) of the network by the WebCos is not well posed in this way, in fact one issue to tackle is the Internet Interconnection and Peering economy and the other is the wrong interpretation that the TelCos have given to the Internet market service space. Connectivity prices are determined by the dynamics of the Internet Interconnection market and the Web Companies are paying according to the market value for connectivity. Sometimes they even found more convenient to invest in building their own network (e.g., WebCos are major sponsors of submarine connections). The second point refers to the end to end principle, Internet services fundamentally rely on the edge capacity and intelligence. The Operators have not approached in the right way this opportunities following for many years the concept of "network intelligence" leaving room to WebCos to address this market and technological segment. WebCos have fully leveraged the capabilities at the edge favored by the introduction of IP networks. They have done this exploiting the client – server paradigm (initially on the server side by building monstrous data centers and now addressing also the client space especially in the mobile area). In this sense the WebCos are not Over the Top, but, as depicted in **Figure 87**, “at the edge”.

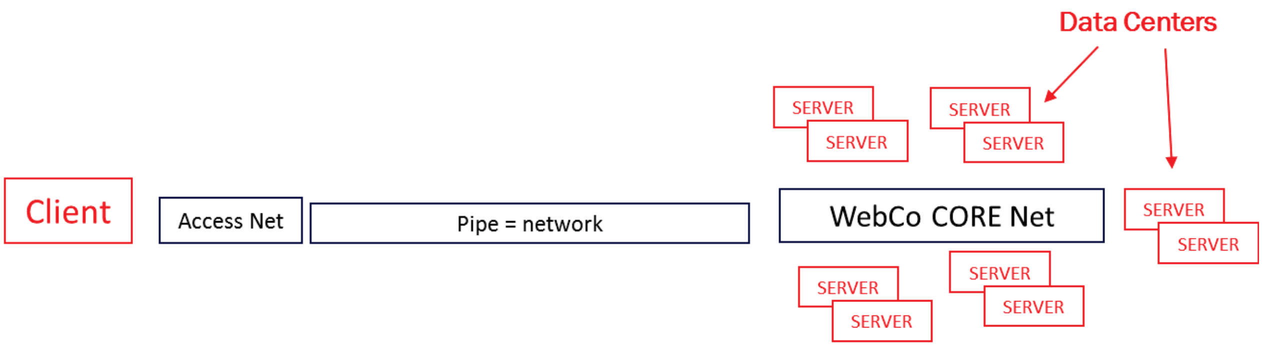


Figure 87: Web Companies Seen “At The Edge”

This can be seen as a rather “philosophical issue”, but it is not. From a strategic point of view it is important to clearly identify the field in which a competitor is operating. An example is the attempt to position the Network Interfaces (NetAPIs) as an appealing offering to the WebCos. This attempt is represented in **Figure 88**.

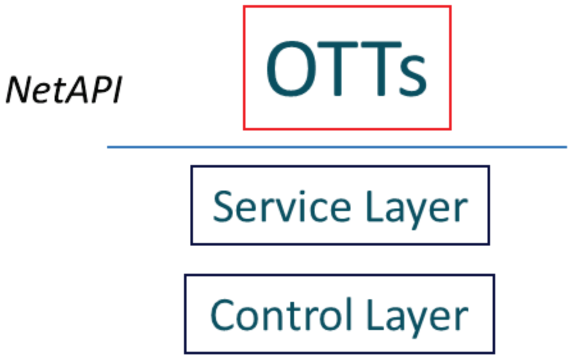


Figure 88: The NetAPI Proposition of TelCos for OTTs

Actually this proposition is not taking over, because the WebCos do not relay on the service and control capabilities of the network, They are at the edge and they use the Internet (and its mechanisms) as a platform. There is no reason nor value for them to give up with their very successful Data Center proposition for moving “on top of the network” [V, XIII]. Even the interesting offering related to Content Delivery Networks have to be declined in a different way (not adopting a Network Intelligence paradigm) but offering them as a native Internet capability to be accessed, used and paid for according to IP and Internet approaches.

6. 3. The Network Paradox

As seen, the Telecommunication industry market is going through a challenging period: on one side, the usage of networks is increasing (more fixed and mobile calls, more messaging and more (mobile) broadband data traffic); on the other side the revenue generated by the increasing traffic is steadily declining each year. Some have identified the issues (Holland 2013), and others, like (Chetan Sharma 2012) have carried out a detailed analysis per major countries and for type of services. This analysis are aligned to the mainstream that shows how revenues for voice (ARPU per user), for messaging and in the next future also for (mobile) broadband decline steadily.

This paradigmatic situation[39] has been described very early in [II, III, XXVIII, XXIX, XXXI, XXXII] and it can be summarized as in **Figure 89**:

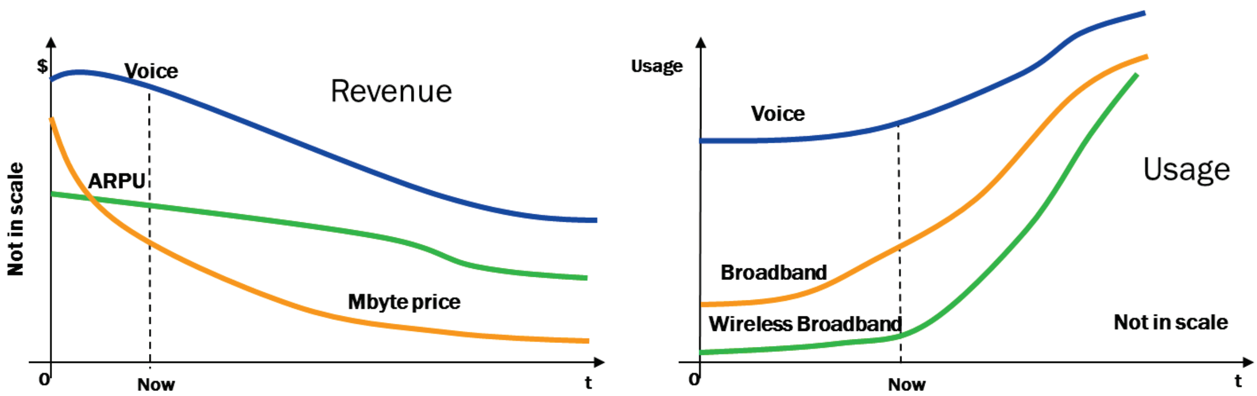


Figure 89: Decreasing Revenues and Increasing Usage of Bandwidth.

Investments by operators in LTE licenses and infrastructure as well as in fiber deployment are massive and it is difficult to find out a way for returning of the investments in a secure and appropriate time frame. This situation leads to the Revenue Paradox: "*the more capacity is offered by the network the less it produces revenues*". This means that investments by Network Operators should be carefully planned and executed in order to determine a break event point and to have a ROI in a predictable time.

---

[39] As depicted in **Figure 89**, the volume of calls is not growing, and actually in many country the volume of fixed calls is diminishing (except in France) and some countries (UK and Japan) have seen a decrease also in the mobile call volume during 2011, http://stakeholders.ofcom.org.uk/binaries/research/cmr/cmr12/icmr/ICMR-2012.pdf. Contextually, the Mobile Data Volume has increased steadily in many countries (it was 13% in Japan that has the lowest growth rate). These could be signs that voice service usage are declining in favor of communication services offered by the Internet (from VoIP to video calls and the like), in fact Ovum (http://ovum.com/press_releases/ovum-reveals-death-of-telephone-is-exaggerated-as-ott-voip-predicted-to-cost-TelCos-479bn-by-2020/) is predicting that voice traffic is shifting towards VoIP, but the total erosion of the voice market by web companies will reach less that 10% in 2020.

Consequently to this situation, Operators are leveraging the technological evolution in order to reduce the investments in infrastructure and systems. Operators try to cut costs and to become a Lean company. This move has a negative effect in the whole telecommunications ecosystem: many consolidated Vendors have seen a reduction of their revenue due to the lack of Operators investments as well as the raise of new competitors coming from developing Countries (especially from China). Some of them have tried to offer outsourcing capabilities to the Operators in such a way that the TelCo has demanded the investments and the management of the network to a Vendor. Someone can see this occurrence as a shift into a service market strongly dominated by software (i.e., Operators are buying the outsourcing of their network as a software service), but actually it is a detrimental occurrence for the Operator (that is reduced to a mere marketing branch), for the Vendor itself (that has to deal with shrinking margins due to the competition on systems and solutions by new comers) and for users (the network of an Operator is bounded more than ever before to technological evolution roadmap of a single provider). Many operators are trying to grasp another possibility: the creation of new business opportunities in areas close to the communication one. This trend is represented in **Figure 90**:

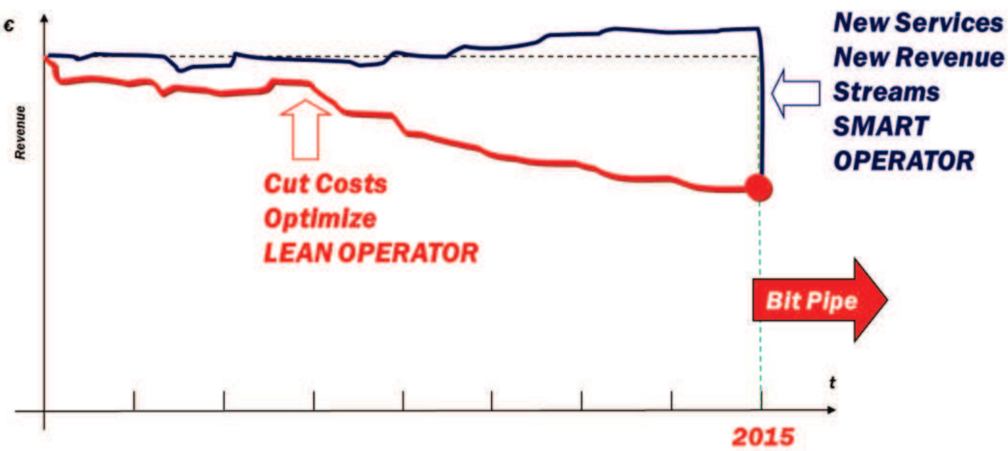


Figure 90: Lean Operator Versus Smart Operator

The **Lean Operator** will try to cope against the situation shrinking the costs and saving more resources. At a certain point in time this attitude will bring to be a Bit Carrier and possible (especially for national Carrier) the Governments will need to intervene in order to defend a national assets as the Telecommunication network. The **Smart Operator** will try to identify new Services and new Revenue Streams in order to compensate the loss of revenue due to the fall of voice, messaging and data revenues.

6. 4. On the Need to Enter into New Markets

The need to enter into new adjacent markets has been proposed by many experts (e.g., (Chetan Sharma 2012), (Delta Partner Group A 2012)) or undertaken by some Operators (e.g., Telefonica with its new branch Telefonica Digital that aims at new markets and new businesses).

Entering in existing markets or creating new ones is difficult. As discussed in (Chan 2005) established markets have such a level of completion that it is extremely difficult and expensive to enter into them without being ready to pay a relevant effort and price. On the other side, opening new markets is far more profitable, but it is also difficult to find a "blue ocean" without competitors. Competing on existing markets can be just a matter of prices, and how a price can be lowered in order to acquire customers. Many Operators entering in the Digital TV market are doing exactly this: the user proposition is to provide the same bouquet of channels over the internet for a discounted or similar price. In the health sector, many Operators are acting as resellers of medical equipment in order to bundle this machinery with communication offering. In simple terms, this approach is lessening the value of the existing market introducing a greater fragmentation and bringing to the customer a marginal advantage.

6. 4. 1. How to Enter into a (New) Market

Entering in a market with a disruptive proposition can be more rewarding, but it is also more difficult to identify the right level of innovation or disruption to introduce. With respect to the Information Communications Technology (ICT) market at least four disruptive factors should be considered and possibly at least one should be really introduced as a major feature for a market offering:

- Technological novelty: For instance the success of the "WII gamepad" is due to the introduction of a technological differentiator, the sensors in the joystick. This has changed the way the games were carried out and the users have greed this evolution with succeed;
- New Business model: entering in a market by adopting the prevalent business model is usually not convenient. The other Competitors are familiar with it and they know its limits and advantages. This means that they can predict to a large extend the moves of

the new entrant and to successfully cope with it. A new Business proposition appealing for the user is a first step towards a blue ocean or at least to an advantage. For instance the Freemium business model has been a successful way for Google and other Web Companies for entering into existing markets bringing disruption and enforcing another (more relevant) business model (e.g., advertising);

- A New Ecosystem: entering into a market without having an ecosystem to back up the service offering is bound to loose. For instance the Apple ecosystem, has been carefully designed in order to have Apple to concentrate on the lucrative part of the business (the terminals) while the rest of value chain (from the applications down to the terminal covers) have been left open. This has created a huge ecosystem that is also used to help in customer churn;
- User Experience: this is the most disregarded aspect by TelCos in the provision of services[40]. Having an appealing user experience is of paramount importance. If Nintendo with the WII was just focusing on the pure technological aspects, probably it would not have been able to create a market success. Users have demonstrated that a nice user experience counts for a large part of the success, more than the functional richness of the solution (applying Pareto: 80% is user experience, 20% is functional richness). Actually a device or a product functionally rich could even be seen as "difficult" and not appealing for customers. A trade-off is always needed.

6. 4. 2. Bits vs. Data

Telecom Operators have also another issue to cope with: many competitors (the Web Companies) are capable of dealing with data and information, while the TelCos focus essentially on bits. This means that services offered by the Web Companies are directly visible and appreciated by the users. A search engine returns information and links to data or further information. Operators provide a communication capability that is seen as a commodity (e.g., in the Kindle terminals, communication comes embedded with the terminal and the user does not pay for it if he stays in the Amazon ecosystem). Actually the attitude to think in terms of bits and calls is detrimental for the TelCos, they are not capable

[40] TelCos have a tradition in providing services that do not require complex interaction with the users, their interface are more for technicians that have to configure the Operator services. If an Operator wants to address the service market, this attitude have to change.

to spot the right functionalities of services and the difficult to replicate data that are the core of the Web 2.0 proposition.

There is a huge difference between the Business Models implemented by the Operators and the Web Companies. Roughly speaking and with reference to the Internet Business models[41] as defined in (Rappa 2004), the TelCos are capable of applying to different Business Models: the Subscription model, in which customers are charged a periodic fee to subscribe to a service (e.g., unlimited calls for a month) or the utility model, in which a customer is charged according to the effective usage of a service (or per units of service). Web Companies cover a much larger spectrum of business propositions: Brokerage, Advertising, Infomediary, Merchant, Manufacturer (Direct), Affiliate, Community, up to Subscription and Utility Models. This diversity makes the Web Company quicker and smarter in applying the right business proposition of a combination of them. Surprisingly, it is a long time that the TelCos are aiming at becoming Brokers on behalf of their users, however this role is now strongly in the hands of a small group of Companies (e.g., Google (Minerva and Demaria 2006)). This denotes a difficulty in changing attitude and internal processes.

Many times the Web Companies are offering Free Services for benefitting from other revenues sources, while the business of Operators is providing services and functions for a price. This is the case of profiling of users: many services are provided for free in order to better and better profile the user. The advantage relies on the fact the advertisers will receive more precise information on the specific user (and his/her relations with others) and hence the advertising will cost more (because the user becomes a more precise advertisement target). As a matter of fact this leads to a Business Model Paradox: "what is a valuable (and billable) service for an Operator, is a service given for free by WebCos".

Figure 91 depicts how the profiling capabilities of Facebook have changed over time[42].

---

[41] See also http://digitalenterprise.org/models/models.html
[42] See http://mattmckeon.com/facebook-privacy/

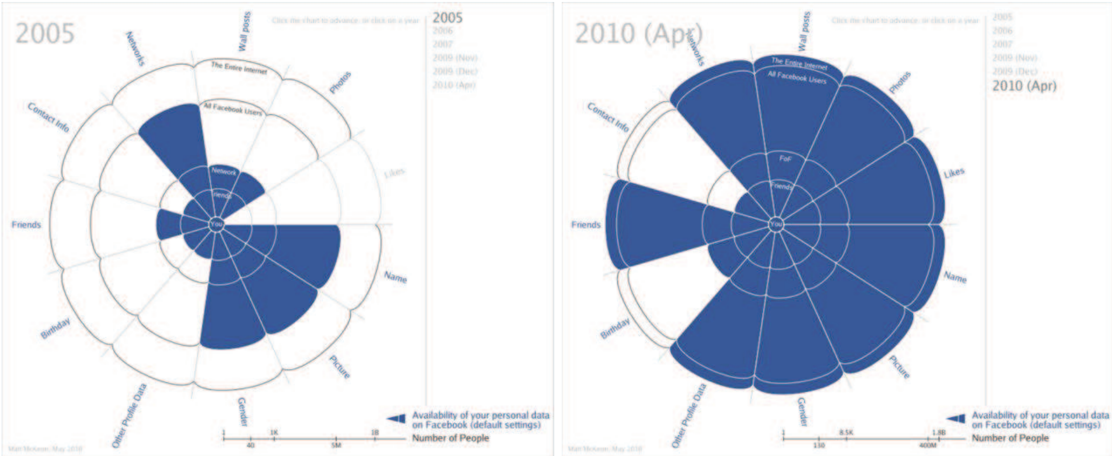


Figure 91: Facebook Profiling from 2005 up to 2010 (source: Matt Mckeown)

6. 4. 3. The Size of the Market: i.e., How Long is your Long Tail ?

Many Operators are striving to enter into the so called long tail markets aiming at addressing the market of small communities interested in particular topics or services. There is the idea that making a number of successful services will provide substantial revenue. This reasoning is wrong and dangerous. The Web Companies (even the smaller one) have really a long reach, in fact even an individual (national) company with nice services can find a relevant market comprised of customers from all over. Instead even the greater Network Operators[43] are bounded to geographical limitation. The long tail of the Operators is limited by two factors: the service offering is limited to the regions in which the Operator has a "network", single regional markets could have different interests in services and became too small for developing a viable strategy. This yield to the Business Span Paradox: "*the services provided by an Operator are local, Services provided by WebCos are global*". Pursuing a strategy of selling services and applications in Regional markets could prove unsuccessful and even a waste a money due to development and management costs. Actually develop and deploying services in platforms (like the Web Companies data center) capable of supporting global services implies that potentially these costs will be spread over the big long tail. For Operators, the ratio between addressable customers, the development and the platform costs could have a meaningful difference and have a great impact on amortization. Big size in long tail and platform can make an important difference between Regional Operators and the global Web Companies. In addition current Web Companies solutions have a technological plus over those of the Operators.

[43] See http://en.wikipedia.org/wiki/List_of_mobile_network_operators last accessed March 31st 2013

A possible approach is to develop interesting applications disregarding specific problem domains in the hope to create a set of appealing application capable of attracting the interest of paying users. This approach ("a set of million dollar applications") is a bit naïve and tries to emulate the capabilities of Web Companies in creating interesting applications in several application domains. It ignores some of the lesson of the Web 2.0 approach: the need to make available difficult to replicate data, the need to leverage on meaningful interfaces and available resources, associate the application to a winning business model, and to use the web as a platform [V]. If these points are not well implemented, then the creation of compelling and successful applications is like searching a needle in a haystack. From the Operators point of view it is much better to first understand whether the service dimension can be a viable option in terms of capabilities to build global services, and to create a strong ecosystem around them. In other terms, the Operator should determine whether it wants to be a Lean Operator focusing on a very streamlined and efficient connectivity provider or if it wants to enter into the "forth dimension" (Chetan Sharma 2012), i.e., the service competition.

If the service path is a viable option, then it is more future proof and profitable in the long run to undergo through a platform approach, i.e., to put in place a viable de-perimeterized platform to be used to create applications in a specific problem domain. This platform should be based on the availability of a well consolidated and difficult to replicate set of data. Related business models should emphasize either an existing asset (e.g., the connectivity) or even better to disrupt or bring innovation into new markets.

Figure 92 depicts the two possible options:

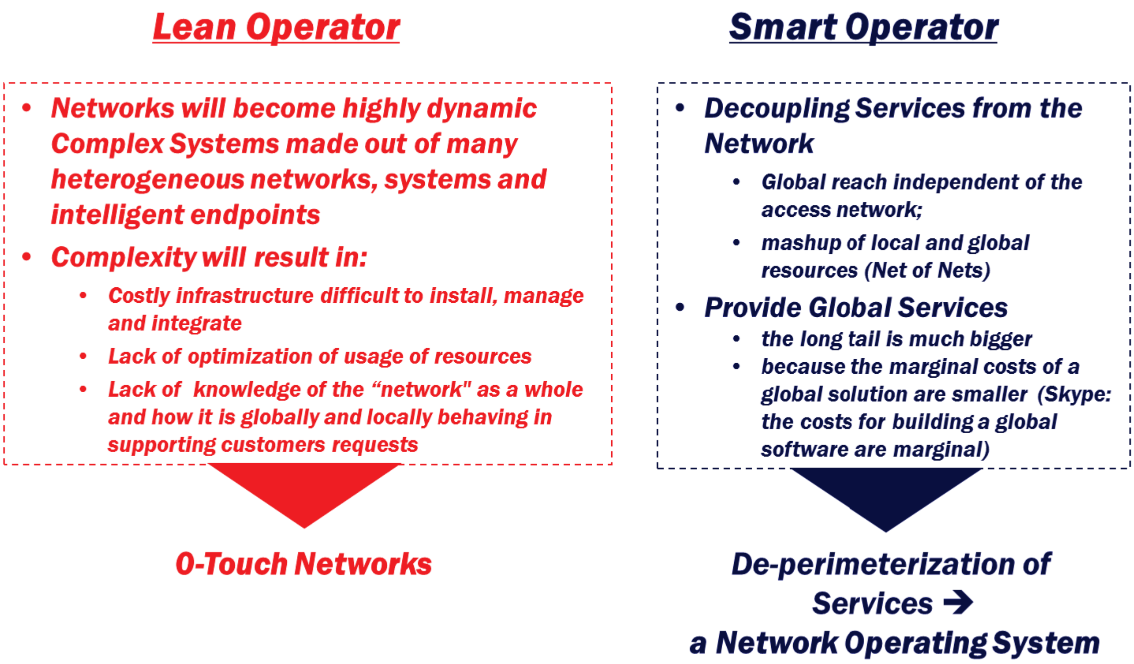


Figure 92: Lean vs. the Smart Operator

- The Lean Operator will focus on the network as a platform and will try to make the network very flat, efficient and cheap in order to reduce its complexity. It also will try to make it 0-touch in order to limit the OPEX for managing it. With respect to the increasing complexity at the edge, it will behave simply as a pipe supporting the best possible connectivity at the best prices possible.
- The Smart Operator, instead, will try to decouple its services from the specific network (de-perimeterization) looking for a global footprint, its network will be open in such a way to encompass and easily integrate other resources and entire networks (the Networks of Networks approach). In addition it will try to capitalize to the largest possible market the services in such a way to make marginal the costs for the software development and the needed infrastructure.

6. 5. A Quadrant of Opportunities

While the evolution of technologies is somehow predictable, the acceptance of technologies to the market, and even more the success of services in a market are much more unpredictable. The success of specific services is extremely difficult to pre-determine because it depends on the satisfaction of real requirements, its technological consistency (the proper working), the user experience behind it and a well formulated and appealing business models. Sometimes also partially missing one of these points make a service less appealing to the market. There is a major difference between the services that can be of interest to a Lean and a Smart Operator: communication services or loosely related services. The more the service is diverging from the beaten path, the more it is difficult for an Operator to identify at the first attempt the right solutions and platforms. In addition there is an attitude towards the creation of them that can characterize the service offering: a traditional focus on services vs. a more advanced approach to them. **Figure 93** illustrates these concepts and pinpoints to a few business opportunities considering also the possibility to introduce new Business models different from Subscription and Utility ones.

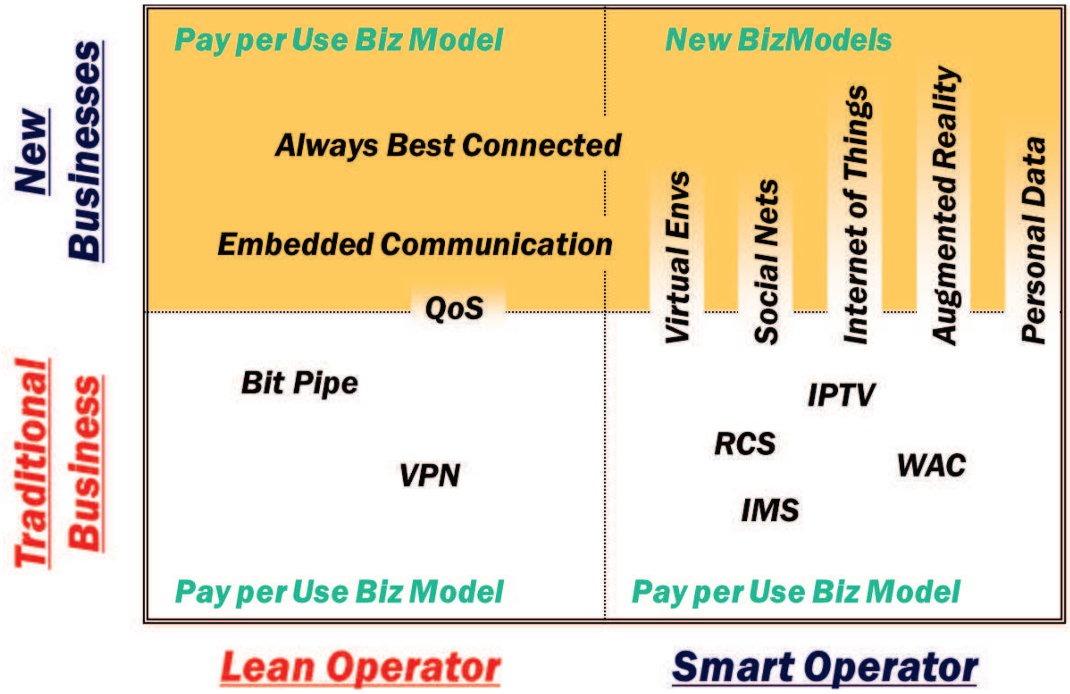


Figure 93: An Example of a Possible Quadrant of Opportunities.

It should be noted that a Lean Operator can be traditional (focusing on the current situation and needs of the customers) or being proactive and innovative trying to enter into new courses in order to promote and leverage connectivity. An example can be the well-known service “Always Best Connected”, the ABC service. In this case, the service dynamically determines the best way to connect the user depending on the context of usage of the terminal. In a traditional approach the TelCo would directly determine from its own perspective (and not the one of the customer) the best available network in a very similar way as in roaming. In order to be innovative, and pragmatic with respect to the increasing capabilities of the terminals, a TelCo could approach the problem in a different way, i.e., by providing support in negotiation with different networks the best tariffs or conditions for the user. This Brokering is quite different than imposing to the user a walled garden of federated networks. Another example in this realm is the ability to allow the user to make use (within the same Operator) of the best network even if the user does not usually have the right to use it: user A has a fixed broadband connection subscription with Telecom A. If there is a problem in this connection, the Operator could allow the user to make use for a certain period of time (the time needed to solve the fixed network problem) of an alternative network without any additional expenditure.

On the embedded side, the approach should leverage connectivity with the best possible quality: for instance if a service is experiencing problems in the connection, the Operator should try to offer new conditions to the user and if the problem persists, to return credit to the user.

In this connectivity based approach, the business models are the usual ones: subscription and pay per use. However it could be the case (for instance in the case of the ABC service, that some forms of Brokerage can be appreciated by the users and paid for.

On the Smart Operator side, there are more meaningful differences. A traditional Operator will try to leverage the current approach as well as the existing platforms in order to enter into new markets (Blum, et al. 2011). This is the case of IMS, a platform designed and developed as the evolution of the Intelligent Network and now being proposed for several services. As discussed in [XVIII], network intelligence based platforms are not the best solutions for services in the realm of Machine to Machine, M2M, and the Internet of Things (in fact ETSI is proposing a different solution) or Virtual Smart environments or Personal data. These classes of services deserve different approaches based on PubSub or different interaction paradigms. Pursuing such an approach has the value to leverage the existing platform, but it has the demerit of approaching new classes of services with mechanisms that do not fit with the requirements of the services. For example: when a container ship enters in a harbor, then suddenly many of the M2M devices connect to the network and try to establish communications links overwhelming the signaling networks with spikes of requests. The signaling network (and the SIP protocol) are not ideal for this kind of usage. Probably using the data layer for transmitting these requests fits better the type of traffic. Unfortunately the IMS architecture has been designed in order to establish and create session between communications entities, the traffic pattern of these classes of services does not fit in this architectural paradigm.

Similar issues are raised for the usage of IMS for providing on demand video services. Ad hoc solutions have proved to be more performing and more adequate to this specific class of services. The approach that an architecture fits all can work at the experimental level and in field trials, but in the real deployment, there is the need to have highly specialized platform for the specific classes of services. Operators should invest first in some application domains and then to invest on the platforms. A more general approach (than IMS) could be tried, TelCos already have data center and cloud based solutions, this infrastructure (the IaaS, infrastructure as a Service) can be the basis for developing specific platforms for coping with specific problem domains (PaaS, Platform as a Service). In this way the TelCo could integrate at the infrastructural level different solutions, but keeping the due specialization and on an upper level to allow the development of services and applications that can integrate specific functions offered by different platforms. **Figure 94** exemplifies this approach.

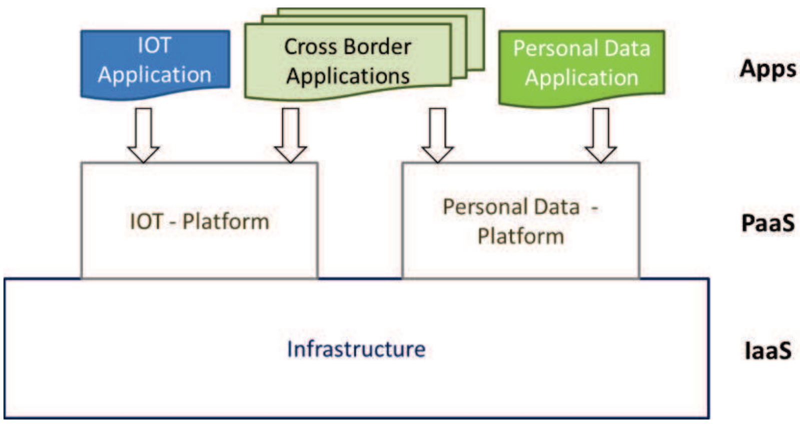


Figure 94: Infrastructure as a Mediation Level

This solution has the merit to provide common functionalities at the infrastructure level, but to allow personalization and specialization at the platform level (specific interaction, data management, functional mechanisms and knowledge can be implemented in several (virtualized) platforms). Services and applications can use and integrate the functionalities in order to support requirements and needs of users. Each independent platform is developed in order to support at the best the typical domain applications. virtualization and programmability of specialized resources can be achieved with greater consistency than adapting an old and cumbersome architecture (as IMS is! (Pesch, Pous and Foster 2005), (Agrawal, et al. 2008)) for providing specific problems.

Summarizing, a traditional Operator will try to leverage the existing platforms and architecture (mainly the IMS) for developing new services. Even if IMS is a good platform for communications services with multimedia capabilities, the delays in definition and implementation have made the platform obsolete also for traditional services. This is the case of Rich Communication Services, RCS (GSM Association 2009), (Henry, Liu and Pasquereau 2009), (Lin and Arenzana Arias 2011). These types of services should have been the new frontier of “one to one” communication by enriching the user experience with new and advanced functionalities. Unfortunately, the specification of these functionalities, their implementation on a cumbersome architecture, and the goal to achieve interoperability between solutions of different Operators (and supported by different vendors) has proved a failure. Users that want to have this rich communication experience have to use other apps such as Skype, Whatsup, and the like. Actually this points to another big difference between the service and application development in a Web vs. a TelCo environment. Web Companies are not seeking for interoperability, in fact they try to impose in the market their solutions as standards de facto (Google has not the objective to make its messaging system interoperable with Facebook’s one, they have the goal to position their solution as the market leader that the others have to follow). Operators, due to their regional footprint and for historical reasons, always try to reach interoperability with others. This delays the deployment and reduces the market grip. This is exactly what happened to RCS.

Coming back to the Smart Operator, the “new business” option of **Figure 93** comprises a number of classes of services to be deployed over specific platform in a very quick manner. If TelCos want to succeed in this environment, they should not relay anymore on platforms and solutions that established Vendors will make available (if Google wants to enter into a new market, they either buy and existing company, or they start building a platform, or both). This is a big change of attitude because many Operators have lost the grip and the control on technological evolution and they have transformed themselves from technological companies into “intelligent buyers” totally depending on Vendors solutions. These solutions are not unique, in fact they can be bought by any competitor, in this way the platforms are not differentiating factors for TelCos, while platform are still key enablers for Web Companies.

6. 6. Are Applications a Way to Compensate the Loss of Revenue for Operators?

Many Operators have tried to enter in the Mobile Application market as a sort of panacea for compensating the loss in traffic generated revenues. There have been attempts to create app store of the single Operator, but this approach was segmenting too much the market, so under the support of GSM Association, the Wholesale Application Community, WAC, (Goncalves, Walravens and Ballon 2010) was created in 2010 with the objective to attract developers of applications and to help them to market their results towards the different carriers. The WAC should act as a sort of warehouse in which developers were storing their applications and in which the Operators were cherry-picking those applications that were considered appealing for their market and their business. Applications should have been developed according to common APIs in order to not fragment even further the application market. **Figure 95** depicts the flow of activities and the flow of money (in red) of a WAC solution.

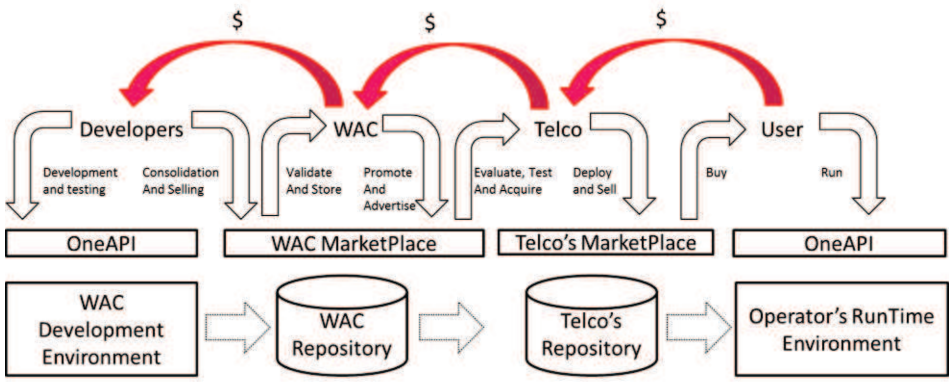


Figure 95: Control and the Money Flows in a WAC Solution

The solution was based on the need to develop applications over a unique interface in such a way to create a uniform market. These applications were then validate, advertised and sold by the WAC to the interesting TelCos. TelCos should have their own market place for selling directly to Users. This mechanism had the form of a B2B2C long chain.

This attempt has failed (the WAC has been sold to an external company[44]). The issues related to WAC are the complexity of the flow of control and money with respect to proprietary application stores (like Apple). One other missing point is the lack of appeal in

[44] See http://techcrunch.com/2012/07/17/wac-whacked-telecom-backed-alliance-merges-into-gsma-assets-acquired-by-api-management-service-apigee/

developing solutions for a "white label" instead than for well-known and established brands. In fact in several classification of Brands the Web Companies are in better position that the TelCos[45].

The market of Apps is quite fragmented and the competition is also on the development environments (in fact Apple decided to push developers to use a specific programming language, ObjectiveC, in order to fully differentiate the APIs, the programming language and the Software Development Kit with respect to other similar initiatives). In this way it is particular cumbersome to develop the same application for different Application Stores. This fragmentation brings developers to be loyal to a specific brand and the related development environment. In addition, as already said, the TelCos have a regional footprint and their Brand recognition is lower than the one of Web Companies. This results in the willingness of developers to stay with consolidated App Stores.

But there is another important factor to be considered: App stores do produce some considerable revenues (Kimbler 2010), but their goal for big Web Companies is to support their real business model: advertising for Google and Terminals for Apple. The major Web Company that has a real interested in making profitable APP Stores is Amazon, in fact the selling of applications is a natural extension of their market place, but even in this case, the web company is offering free applications for nurturing its business of terminals (e.g., the kindle) that is a means to access to their market place.

There are other two aspects of the App Stores to evaluate: there is a proliferation of AppStores that reduces the value of the market (competition between different very similar applications) and the declining price of Apps. The value of applications is an important factor as well as the sustainability of the business behind the application development: up to 59% of the mobile applications "do not produce enough revenue to break even their development costs" and "80% of application do not generate enough revenue to support a stand-alone business"[46].

So App development is more a means to create a lively ecosystem that attracts users and gives them reasons for remaining in a specific proprietary environment (e.g., Apple) than a

[45] See for instance http://www.forbes.com/powerful-brands/ or http://www.brandindex.com/ranking/us/2012-annual/top-buzz-rankings

[46] See http://visual.ly/wake-call-truth-behind-mobile-app-success

business in itself (with exceptions). The class of revenue generating apps is small and probably it falls in a few domains like gaming, and productive tools (editors, and the like).

6. 7. Capability of Execution of the Different Actors

The Actors competing in the service sector are very heterogeneous and they have different skills and assets to leverage in this tussle. In this short analysis we focus on these classes of companies:

- the Internet Companies, i.e., the big players of the Web (Google, Facebook, Amazon and the like) as well as the application specific companies like SAP and Salesforce and the small internet companies that introduce many new services and applications in the Internet;
- the TelCos, i.e., the communication service providers that offer communication capabilities to the market;
- the Software Companies, these are the well established companies that do provide basic software (especially for the desktop and server systems);
- the IP companies, i.e., those companies that provide equipment and solutions for the evolution of networking infrastructure (Cisco, Juniper, and to a certain extend also the telecommunication vendors);
- the Hardware and Infrastructure Companies, these are the companies whose major part of the business resides in selling hardware products to users. They comprise traditional computer makers as HP but to certain extend also Apple and the terminal manufacturers.

This classification is not a hard one, in fact some of the considered actors are able to play different roles in this classification. Actually this is a first lesson in this analysis: some actors are capable of differentiating their market in order to better support their main business model. This is the case of Google, Amazon and Apple. They are capable to offer high level and appealing solutions in different areas. This means that these companies are very flexible and their know-how covers many technological areas, or they are able to acquire very rapidly the needed know-how. Other companies are more static and they are quite good in support their established business model, but they have difficulties in entering in adjacent sectors. These companies will be considered under a small number of business perspective:

- How the market perceives the Company

- the value of their "main business model" and how they interpret it
- the flexibility of providing cross-sector solutions and offering
- the capability to introduce innovation and the level of mastering of software (i.e., the capability to use software as a differentiator for their business
- the span of their business
- the capability to leverage local business

These are not to be intended as the parameters for a due diligence of these companies, they are meant to provide an indication of how these companies can move in the ICT market.

Some of the main considerations are represented in **Table 4** and they are partially drawn from (Acker, et al. 2012).

Table 4: Differentiators between some of the ICT Players

	Internet Companies	TelCos	Software Companies	IP Companies	HW and Infrastructure Companies
Market perception	- Many high regarded brands - Small companies are appreciated for innovation value	- Greedy and exploiting - Solid and trustable	- Monopolists (e.g., Microsoft) - Needed but they not always deliver what expected	- Enabler of the communication world - Some interesting capabilities (e.g., CISCO for Business market)	- Appealing (Apple, Samsung, and in general terminal vendors) - Old and not innovative

Robustness of the business model	- Diversity in application of Business Models - Some WebCos are very specialized in a single Biz Model (e.g., Google) - Others (e.g., Amazon) are capable to integrate several Biz Models - Capacity to apply a full range of Biz Models	- Two Major Biz Models: subscription and Pay per Use - Processes of the TelCos are all oriented towards these two BizModels - Difficulty in moving towards other BizModels (e.g., Brokering Model)	- They have been essentially based on Licenses Biz Model - Now they are moving to the concept of Software as a Service and they are interpreting it in a broad sense offering a large capability of outsourcing	- Entering into the traditional TelCo Market (Video Conferencing) with advanced offering - Their traditional monopoly of infrastructure selling is declining in favor of new comers (e.g., Chinese companies) or new paradigms (SDN and virtualization) - Need to find alternatives	- Major business model is manufacturer business model (selling equipment) - Capability and shift towards Software services and apps - Possibility to move to SaaS models
Cross-sector solution	- Some WebCos are already entering in adjacent markets (e.g., eHealth, transportation, etc.) with disruptive solutions - Capability to leverage their “general purpose” infrastructure to enter in new markets	- Communication is assumed as the major business, other business are considered as a means to sell connectivity - Entering into adjacent markets with cross sector offering is usually deemed as not rewarding because the current communication market is huge compared to the new one	- They adopt an approach similar to WebCos, some Software companies (IBM) are aggressively addressing the IoT market with global solutions Flexibility and knowledge of different markets that use ICT	- Leveraging IP networking in order to address the edge market - Cisco move to IoT leveraging IP networking	- Possible exploitation of edge capabilities - Move to new markets leveraging specific technologies (HP with sensors) -

Innovation capability	- Very High levels of Innovations and high flexibility - Free minded and aligned with the Internet speed	- Very conservative and centered on preserving the status quo - Difficult to introduce new approaches	- Flexible and with good possibility to innovate - Good know-how of fields near ICT (e.g., smart cities)	- Very centered on IP networking - Possible expansion towards Internet applications	- Very Flexible and capability to innovate - Possibility to innovate with a bottom up approach
Span of Business	- Global	- Regional - Some have an large international footprint	- Global - Some companies have regional relevance only	- Global	- Global
Leverage of local business	- Low, but they leverage the "market of one"	- Very High - Possibility to leverage local presence	- High penetration (commercial) in many Countries	- High penetration (commercial) in many Countries - Specific in the Business market - Low penetration at the residential level	- High penetration (commercial) in many Countries

From Table 4, it emerges that WebCos are well positioned to play a major role in future service development. They have advanced and innovative service platforms, they are perceived as innovative and they are flexible and quick in grasping new opportunities. Second are the Software Companies, they have similar advantages of WebCos (maybe non their appeal on the customers) and can leverage on the mastering of software that seems to be an essential ingredient for the success in the ICT competition.

The other actors (especially the Terminal Manufactures) have some possibilities but they need to push their innovation capabilities and a stronger grip on service development. The global presence seems to be a major advantage, however the local presence (especially from the Operators point of view) offer the possibility of a direct contact with customers and the opportunity to better present and leverage the company offering by means of personalized contacts.

These few considerations show how difficult is the role of Operators in the current context, the shrinkage of the traditional business models is not so fast for the time being and this leaves some time to TelCos to reinvent themselves and find new opportunities. However the traditional important position given by the ownership of the network is diminishing and there is the need to make choices related to future business and the role that the single TelCo wants to play.

6. 8. A New Trend: Servitization

One important trend that Operators could leverage for entering in an innovative way into new businesses is Servitization. Servitization is a means to create a continuum between a product and a set of remote functionalities that augment the capabilities and the properties of the product and relate it to other products and functions either in the physical or virtualized environment. The new service offering should be structured in such a way to identify very early a set of generic services (i.e., a set of widely reusable and accessible functionalities and capability stemming from the functions offered by Virtualized objects) and to provide them in a generically accessible and secure infrastructure (a cloud). A generic definition can be as follows: Servitization is the capability of creating a link between a (physical) product and a set of services and enriched functionalities that extend, complement, and add value to the product itself (see **Figure 96**).

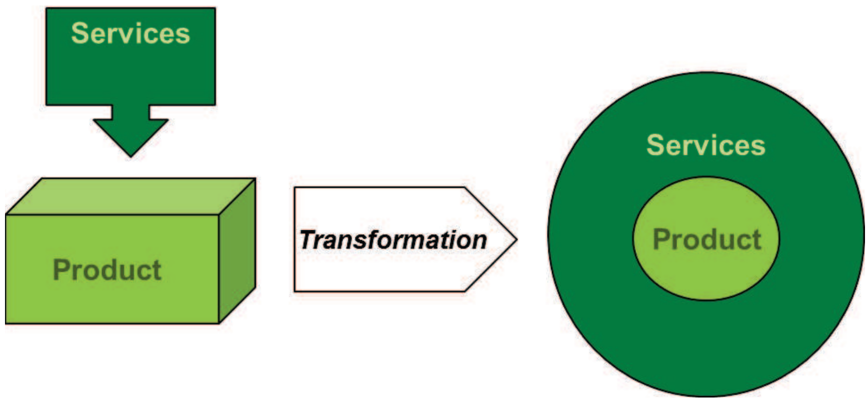


Figure 96: Servitization

Sometimes servitization is directly related to the so called Network Effect, i.e., the increase of value of a product or service depending on the number of people that use it. The IoT market could be an example of servitization acquiring value by means of the network effect given by the availability of many objects.

This capability is strongly dependent on the capability to establish a constant link between the real world object and its representation on the cloud. This property can be ensure by services like "always best connected", i.e., the capability to ensure connectivity between the real world object and its virtual counterpart in a seamless way and independently from the location of the specific object. Examples of servitization have been given in previous sections.

Servitization could allow the TelCos to cooperate with totally new actors that are not willing to cope with the intricacies of the communications and that need to leverage and move their products into the Internet and the Virtual Continuum. This property must be suitably leveraged and exploited by means of new relationships with producers and consumers communities in order to create a viable ecosystem of Actors and Stakeholders that will benefit from it.

6. 9. Findings

This Chapter has shown that TelCos are at a cross-road, the occurrence of many paradoxes is a signal that a structural change is taking place. However this transformation is not so abrupt and some time will pass before things will drastically change. In the meantime the TelCos have to be schizophrenic, on one side they should try to leverage and exploit the current infrastructure and the traditional business models on the other side they have some time to prepare for a significant switch to new infrastructures, business models and even markets.

This period is characterized by the growth of softwarization of functionalities and the capability of mastering the software is of paramount importance (as discussed in Chapter 4. and Chapter 5.). Servitization is an example of how products can be transformed by means of software into services. One possible role of the Operator is to be a facilitator of this transformation and to occupy some relevant role in the new value chains.

Another important fact that emerged from the analysis carried out in this Chapter is that in order to enter into new markets there is the need to bring some disruption. This could be at the technological, business model, or user interaction levels. Disrupting the status qui requires capability of execution, a clear perspective and the will to succeed. The typical Operators have to improve under several aspects in order to implementing this strategy.

Another important aspect emerged from the analysis is that it is extremely important to understand what the competitors are doing. Under this respect the attitude of considering the WebCos as a sort of parasites that are freely exploiting “the network” is very dangerous more for the Operators that the WebCos themselves. In fact WebCos are not operating on the top of the network, they are exploiting the “end to end principle” and they operate at the edge of the network. In addition they have greatly contributed to the success of broadband and mobile networks by attracting users. In fact customers are not interested in connectivity “per se”, they are interested in services and functionalities that can be used in the network. A major disruption that WebCos are bringing in is the service area: many services that were sold by the Operators are now provided for free because they are remunerated by different business models.

Other important aspects emerged by this analysis are related to:

- The need to move towards the leverage of data instead of the bit transport. This means that dealing with data become more important than dealing with transport services.
- The Operators do not have a comparable footprint with respect to WebCos, this means that in terms of addressable markets the WebCos have the capabilities to exploit bigger markets. A possibility is that TelCos have to give up to a traditional characteristic of their business: the deperimeterization of services. This means that Operators that want to compete in the service area should decouple services from the underlying network. So far, this relationship has been a very tight one that has to be relaxed if services are the business target. This is also strongly related to some of the finding of Chapter 5. that showed how services should be provided with well-designed and flexible architectures and not over imposed on the available platform. Once again the choice of control paradigms can have an impact on the successful development and deployment of services.

Chapter 7. How the TelCos Sustain their Connectivity based Business Model in the Wave of Disruption Generated by the Technological Evolution ?

Highlights

- Network Intelligence as a means to Compete
- Net APIs
- Networks and QoS/QoE

7. 1. Introduction

Many Operators are trying to enforce the usual business models and the traditional approach to services and infrastructure control. This is reflected, for example, in the evolution of the specification of 3GPP for the Long Term Evolution, LTE, and the following releases. From the service perspective the major novelty of LTE is in the full switch to packet switching and the lay down of circuit switching, while simpler technologies for radio access are enabler for supporting the broadband IP connectivity. At the service and control level, LTE is not introducing relevant differences with respect to IMS. In the fixed network side, a similar approach is followed in ETSI TISPAN (TISPAN 2009). This is can be seen as the willingness to reuse and capitalized the current investment in the network control and service infrastructure. However this attitude, especially in a period of rapid technological transformation, is inappropriate. The technological basis of these architectures are very old (more than ten years old) and new technologies and solutions have emerged and seized large sectors of the ICT market. Hadoop and the noSQL technologies can be an example, as well as different programming languages and paradigms (agile programming). Pursuing with this old perspectives give a technological advantage to the Web Companies and other competitors. Instead of running after the WebCos technologies, TelCos should jump ahead and experiment, develop and use alternative technologies.

In this chapter, an analysis of the most consolidated and used technologies "in the network" is provided in order to understand whether they are fit for supporting the TelCos in the daunting task of competing with innovative and technology oriented WebCos. Firstly, an analysis of current network platforms based on the Network Intelligence is carried out. These platforms should allow for easy development of services capable of competing in richness of functions and easiness of usage with those offered by WebCos or by P2P systems. The interaction and control paradigms play once more an important role. In addition the strong tie of the network and the services has to be evaluated and it should be understood if this is a real asset or a weakness in the development of services. The association of IMS and Service Delivery Platform, SDP, and the progressive move towards Service Oriented Architecture, SOA, within a TelCo environment must be evaluated from the perspective of new classes of services.

A second scrutiny is devoted to the Exposure, i.e., the possibility to open up the TelCo's systems in terms of Application Programming Interfaces that can be freely used by third parties in order to develop networked applications. Are they a real business enabler? Is there a practical usage of network related APIs ? What kind of interfaces should be exposed? Actually, there is a property of the APIs that needs to be considered: the more the API abstract from the underlying resources the simpler is programming but there is also a deficiency in expressive power (i.e., the API provide too simple and sometimes useless functionalities).

A third issue to consider is the assumption that users want Quality of Service, QoS, and Quality of Experience, QoE, from the network. This is a recurring argument and it is profoundly determining the strategies and the attempts of many Operators to be successful in the services realm. Is there any perceived value from the user stand point in network services offering some guarantees in the Quality of Service? Is it appealing form the residential users? Are these features appealing to the business market ?

The underlying issue analyzed in this Chapter is to determine whether the traditional properties of the "network" are still assets or are simply an attitude that the TelCos should get rid of.

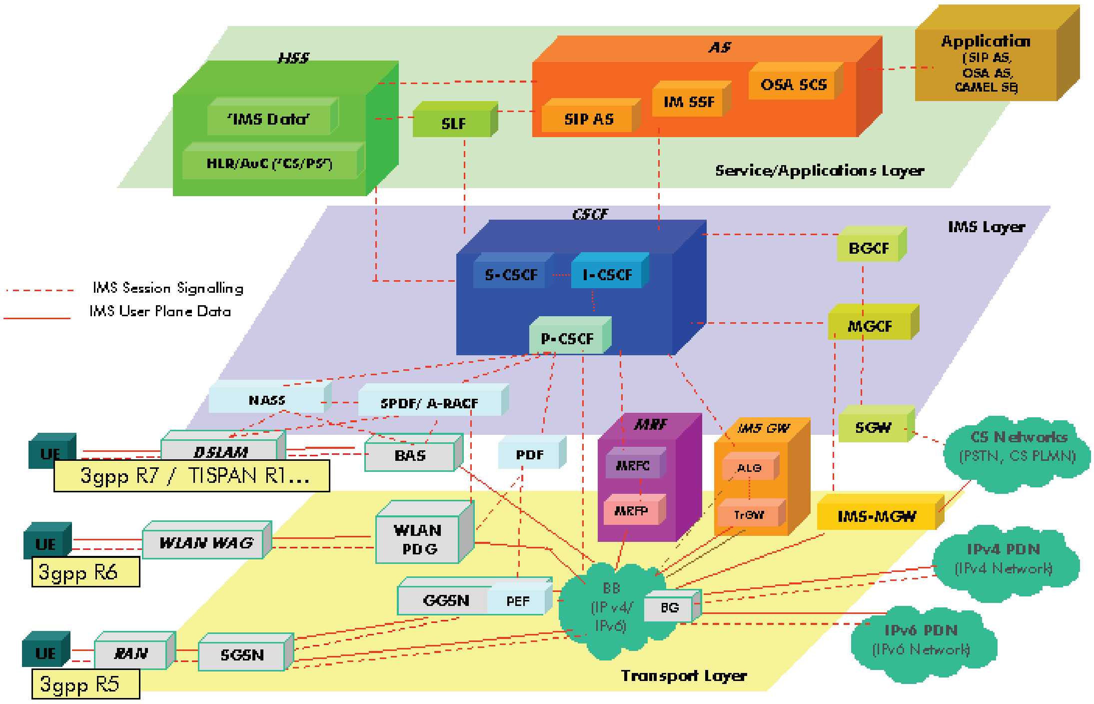
7. 2. The Network Intelligence Legacy

The IP Multimedia Subsystem (Copeland 2009) (first defined around 1999) is a specification for an IP mobile network originally devised by the 3GPP as an evolution of the GPRS architecture (Bettstetter, Vogel and Eberspacher 1999). The general goal of IMS was to provide a reference architecture for a new system able to encompass voice and multimedia services in an all-IP infrastructure. In particular its goals were:

- To decouple the telecoms infrastructure into different layers (namely Transport, Control and Service Layers) in order to guarantee the evolution and independence of networked functionalities;
- To provide voice and multimedia services over a controlled (mobile) IP network;
- To support different access networks (e.g., UMTS but also WLAN, WiMax, etc.);
- To use as many IETF defined protocols as possible to get a full compatibility with the IP and Internet standards.

Surprisingly the definition of IMS first caught the attention of Fixed Line Operators. At that time, fixed network issues were focusing on the definition of a viable architecture able to integrate the circuit and the packet switched infrastructures. The fixed line community was stuck in discussions about what a softswitch should be. The emergence of the IMS architecture was a help in resolving these questions, as a result that the fixed line community embraced the IMS concepts. Soon ETSI TISPAN was working on the definition of an IMS architecture and its extension supporting Fixed Mobile Convergence.

Figure 97 depicts the IMS architecture now accepted also by ETSI TISPAN as a framework for Fixed-Mobile Convergence.



from wikipedia.com

Figure 97: IP Multimedia Subsystem

The IMS architecture is based upon the separation of functionalities at transport, control and service levels. The main protocols are IETF protocols (with some changes): SIP (Rosenberg, et al. 2002), (Johnston 2009) for the control of sessions of services and Diameter (Calhoun, et al. 2003) (for management and QoS related operations).

The transport layer is derived from the GPRS architecture. The Mobile Access Network collects IP streams and forwards them towards a couple of nodes, i.e., the Serving and Gateway GPRS Support Nodes (SGSN and GGSN) that guarantee the transport of the IP packets from the access network to the Internet. The IMS transport layer has been extended in order to allow for the integration of fixed access (through DSLAMs and Border Gateways) and WLAN networks. These forward IP packets directly into the core IP network. In order to provide some functions related to QoS, the Diameter protocol and its control architecture have been adopted. There are Policy Decision Points (logically pertaining to the control infrastructure), that decide which policies to apply for a certain stream of IP packets) and Policy Enforcement Points that execute those policies. In order to interact with existing PSTN networks a media gateway function is defined, it allows for the "translation" of VoIP streams into analog voice. In addition Media Resource Functions are defined as a means to provide multimedia related functions.

The Control Layer is organized around a number of servers. They are related essentially to the control and management of sessions (the Call Session Control Function, CSCF), QoS and the interworking with the PSTN signaling system. The CSCF Servers are duplicated in order to allow for roaming and mobility. This is essentially the equivalent of the call control mechanism implemented in the PSTN.

The Service Layer is organized around a number of disjoined servers for the management of user data and identity, application servers for providing SIP, OSA or IN services and a sort of exposure gateway enabling third party application providers to use functions and services of the IMS infrastructure.

This architecture has been promoted as a step forward in the field of multimedia communication by the Telecommunication industry. However it has received a number of criticisms (e.g., (Waclawsky 2005), [V], (Minerva 2008, 2)). In fact, the ability of the IMS architecture to provide new services has been questioned. The IMS is defined on the assumption that services will be "centralized" and controlled by the Operator (as in the old Intelligent Network architecture). This approach tends to reapply the "known" business models and solutions in the new context of Next Generation Networks, where established Operators and Vendors push for a "controlled" transition to a more Internet based control of services but still defending their legacy networks and product lines. IMS proposes the usage of a few protocols (mainly SIP) in order to provide session related capabilities. SIP seems to be a new signaling system, but the traditional framework for providing services remains the same. In IMS the session concept is synonymous of "call", in fact services are built on top of the concept of establishing a session. The session is "stateful", i.e., the states are stored in the network in order to keep track of the progress of processing in terms of Finite state machine. Call Session Control Function, CSCF, are essentially an evolution of the Basic Call Model by means of the SIP functionalities. In addition, the IMS architecture (based on the successful GPRS one) is introducing the traditional mechanisms for the (artificial) perimeterization of services within the domain of an Operator. In this sense, the CSCF functionalities are replicated in order to move the control to the Operators that has a contract with the customer requesting services to a network (possibly in roaming). This organization has a major drawback in the complexity, in fact many SIP interactions take place in order to execute a control functionality and this number is very high if roaming functionalities are needed.

From a service perspective, the greatest drawback comes from the ambiguous definition in IMS of the Service Layer. It is generically referenced as an environment in which different applications server coexist (SIP, OSA and Camel), they use also different protocols, but only SIP is considered. This means that a lot of “normalization” has to take place. There is not the possibility to directly organize the network resources from the service layer if a session is not taking place (no COPS based interactions). The interaction paradigm between the control and the service layer is still the traditional one of Intelligent Networks, i.e., services are triggered by network requests, the service logic is started and it is executed until the value added functions are requested, then the control can be returned to the Control Layer for the normal execution of services. This is quite different and much more constrained than the freedom and the expressiveness capabilities offered by the Interne Application Servers by means of much easier interfaces (e.g., based on the REST architecture).

In order to overcome this major shortcoming of the IMS architecture, many Operators have seen a solutions in a better definition of the Services Layer, in particular the concept of Service Delivery Platform has been advocated as a solution to a flawed definition of the service environment within IMS.

Service Delivery Platforms, SDP (Ohnishi, et al. 2007), (Moriana Group 2004), are a way to create and govern services over an IP Platform. SDPs (see **Figure 98**) are telecom oriented platforms that use Service Oriented Architecture, SOA (Sward and Boleng 2012), and Web Services (Snell, Tidwell and Kulchenko 2009) technologies in order to enable a framework for building new services. The concept spawns from the application server definition: they embed within the IT framework “connectors” for interfacing with telecom networks and telecom devices. The idea is to be able to bridge two classes of servers: the telecom control servers (such as Parlay Gateway) and the IT application servers. The goal is to create a sophisticated environment to develop, manage and deliver services, as well as reducing the development and deployment costs associated with telecom services.

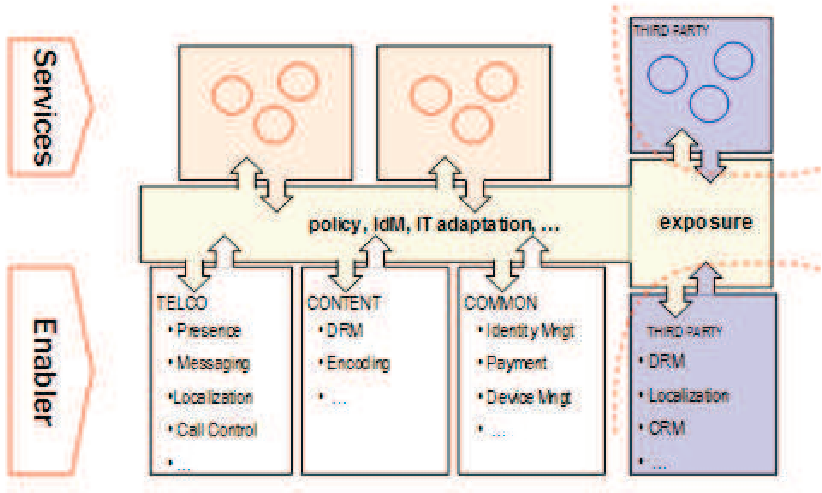


Figure 98: General Structure of a Service Delivery Platform

An SDP is built around a mechanism for passing and controlling events between a number of processes, each one representing a reusable functionality. Such a "service bus" allows communication and synchronization (by means of exchange of messages, i.e., events and commands) between cooperating processes and functionalities. The "bus" allows the flow of some messages also towards external domains, so enabling the communication with third parties applications. Functionalities (and related processes) are divided into: Enablers, that provide common and general purpose functionalities; and Services, that use the general functionalities to build business logics and services. Some Services and even some Enablers (through the exposure interfaces) can be provided by third parties. In this way the service portfolio and the functionalities offered by an SDP can be greatly extended.

There is interest in combining SDP with IMS. The reason seems to be that the definition of the Service Layer of IMS has not adequately defined a common service architecture (i.e. a software architecture is needed and not a functional one). In other words, the IMS service layer is more a collection of systems used to provide services on the top layer of IMS, than a real service architecture. SDP seems to cover (in a traditional operators' approach) the lack of specification for a well-structured and well defined Service Layer.

The combination of IMS and SDP is viewed by many operators as a winning combination to provide a rich set of session and content based services over the new all-IP network. In addition, the SDP approach is a means to leverage the possibility for Operators to offer Network APIs that could appeal developers and foster the creation of a service ecosystem centered around the assets and the traditional approach of the Operator. As deeply discussed in [V], this approach is not working because, simply put, the real development platform is the Web. It is in the web context and with the web protocols and mechanisms that services and related will grow. In addition as discussed in paper on paradigms], the programming paradigm chosen for the IMS and SDP combination are very far from the predominant ones of the internet (essentially based on a client – server model and not on triggering mechanisms stemming from a complex finite state machine as proposed by IMS). To make things even worst for the TelCos, the technological gap between the Web Companies platforms and the IMS based one is increasing and not reducing because the developer's communities are bringing a lot of new ideas and initiatives. The success of TelCos ecosystems based on NetAPIs and SDP is very limited even in the consolidated realm of Voice Services (e.g., the integration of Rebbit in BT).

The SDP promise was also aiming at the integration of cloud and XaaS approaches into the realm of the TelCos. As discussed in [XXIII], TelCos have limited possibilities also in the Cloud market. If they want to success they have to aim at the business market and specifically to those companies that want to integrate their existing processing and storage infrastructure at an international level and have a certain level of QoS with respect to interconnection between systems distributed in several Countries. The Federated Cloud approach is the most promising one for the Operators because it allows to leverage several assets as: local presence, continuous support to the customer at a national level, some level of guaranteed communication services, the need and the possibility to interwork and integrate different clouds. This is a small part of the entire cloud business that gives and advantage to the Web Companies if customers are looking for cheaper price, simplicity of usage and greater availability of resources.

In order for the Operators to be competitive it is necessary to disrupt the market at the service proposition and at the technological level by introducing discontinuity in the status quo. The SDP + IMS approach is a mainstream approach that points to give a programmable control on traditional service but does not support innovative service deployment. IMS in particular is a traditional platform that suits the value added service approach. Being a horizontal platform it may serve for providing also other classes of service, but it lacks behind specialized and more IT and web oriented platforms. An example is the Video

Services provision that can be implemented with SDP + IMS, but the rigidity of this platform makes it less adequate than other specific platforms for the provision of TV like service.

[XXV] discusses how a Telecom Operator can introduce new services and what are the new technological paradigms that could be followed. One direction is aiming at user centric services like Personal Data, Identity Management and Security adopting a disruptive approach: helping the user to better cope with an unfriendly Internet interested in spoiling the user's rights. The second direction is to move from Internet of Thing (IoT) to the Internet with Things, i.e., the environment capable of representing each real world object and to allow the creation of many applications and the servitization of products. The third direction is towards the smart environment and the edge intelligence, i.e., the possibility for the Operator to act as a sort of anchor point (a hub) around which edge node can aggregate in order to create dynamic and very smart environments. In order to take these directions, there is the need to rethink the traditional Operator approach and to develop new platforms that disrupt the current technological mainstream. In this sense, the new platform can be based on overlaying (especially using P2P technologies), on cognitive behavior of nodes and the entire infrastructure, on virtualization of resources and functions. This is a new combination of existing technologies that can help in competing with the WebCos. In [XXIV and XIV] a detailed description of the envisage architecture can be found. Here the most important properties of it are briefly sketched out. A new Service Enabler Platform (a network of networks) will strongly depend on:

- the virtualization of resources (communication, processing, storage, sensing);
- the cooperative orchestration of single or subsystems' resources;
- the introduction of self-organization capabilities in order to achieve an autonomic, cognitive [XVII] behavior of applications and resources (the highly dynamic and unpredictable behavior of a network of networks requires the real-time adaptation to different contexts);he introduction of different (from client - server) paradigms for the cooperation of distributed (virtual) objects. Overlay networking is a real distributed processing paradigm and it fits properly in this dynamic environment. The combination of these technologies will lead to a programmable networked environment such as the one represented in **Figure 99**.

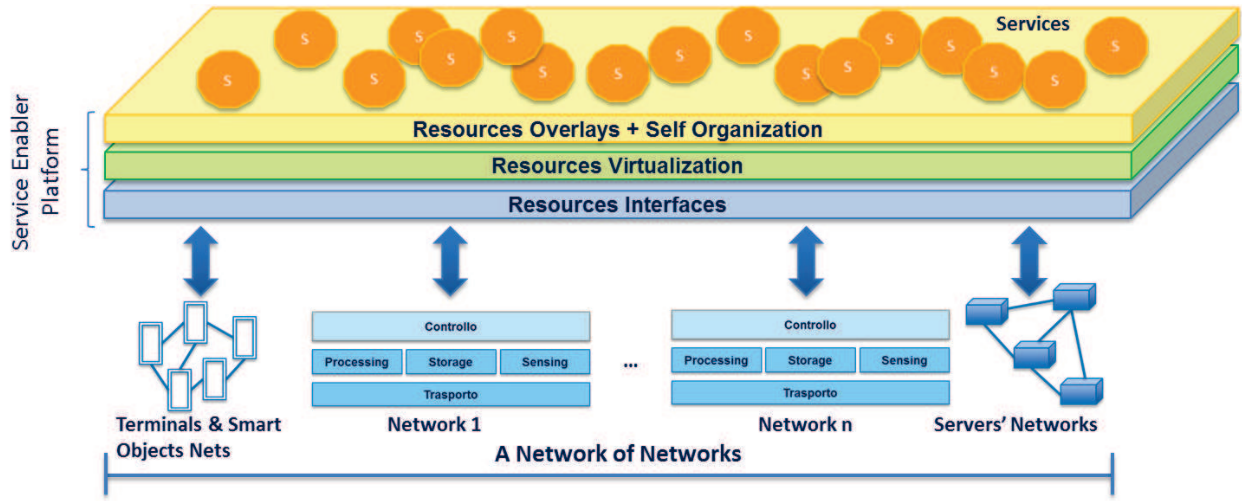


Figure 99: A Service Enabling Environment and its Basic Principles and Layering.

The Service Enabler platform is a sort of Network Operating System, that, through the representation and virtualization of networked resources spanning across many subsystems and different administrative domains, will allow applications to negotiate for "virtualized" and autonomic resources, to allocate them, to control and program their functionalities according to the specific needs and requirements. The upper layer is made out of overlay network that comprises basic resources. These basic resources can be extended or can be integrated with new specialized ones in order to allow for the provision and offering of many services. It is important to stress out the role of end-users terminals and networks. They provide to the entire network a set of capabilities and the possibility to the entire network to rapidly grow (similarly to P2P networks in which end users contribute to the resources of the system that can scale up).

Actually virtualization and opening up of interfaces (Application Programming Interfaces, APIs) can be combined in order to build a very flexible and programmable platform for services. **Figure 100** depicts this possibility from a clean slate approach. It introduces the concept of a knowledge based layer that represents the status of the network that can be used in order to optimize the resources for the single application, as well as for exerting a global level of optimization. Predictive mechanisms could be also considered in order to prepare the whole system to tackle upcoming issues and problems. For a more detailed discussion refer to [XIV].

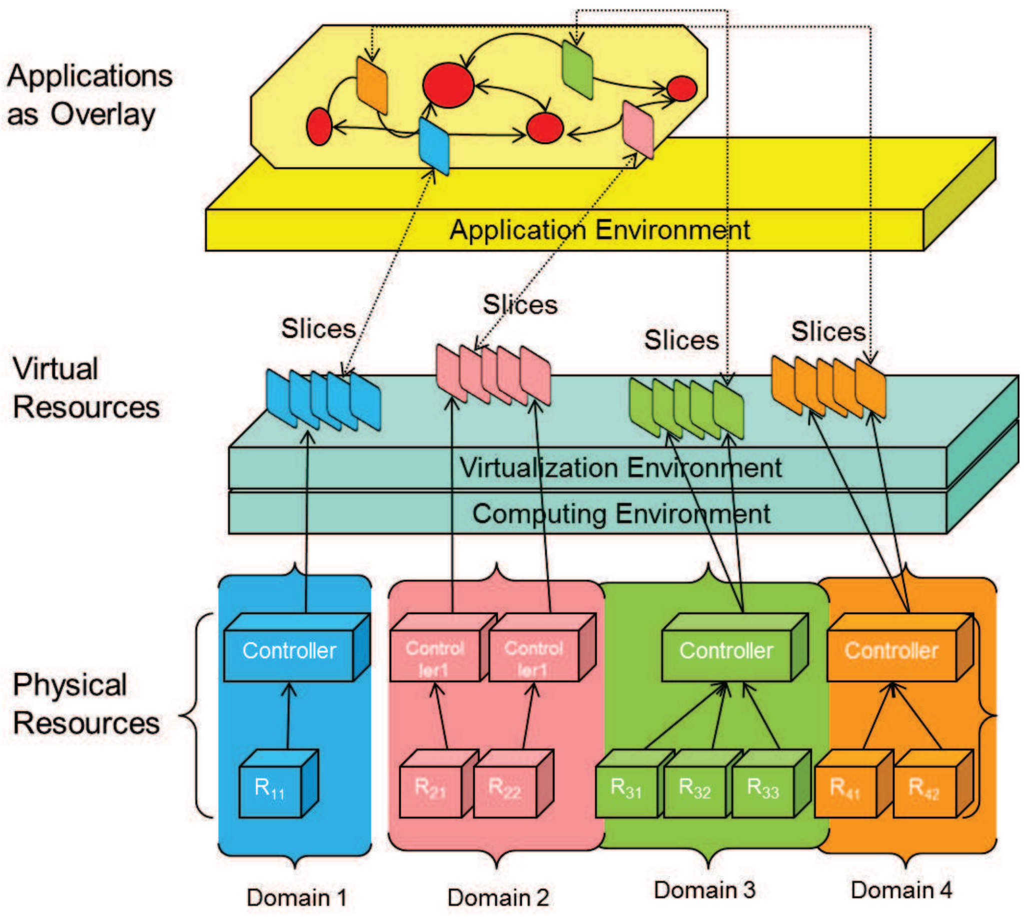


Figure 100: Layered Cognitive Architecture for the Future Internet.

One important feature of this platform is its capability to span over different administrative domains exceeding the limitations of the current strong coupling of service layer and network layer.

7. 3. About Network APIs

Many TelCos see the possibility to monetize the network by means of APIs and related functionalities. Virtualization and the progressive softwarization of network functions lead to new issues and new opportunities for positioning the APIs in the market. One major point related to softwarization is its huge impact on the value of the network. Up to now, network deployment is a capital intensive business, i.e., creating a new network means to put upfront a relevant amount of money and to have a return of investment in the long term (actually many business plans of Operators have shown that the payback period for new deployment in LTE and Fiber is increasing from 4-5 year to over 10 years and usually it scale up to 15 years). The softwarization of the network can break this barrier and many more Actors can be ready to invest in the deployment of infrastructure. Many Operators see this as a direct menace to the current status quo and an excessive opening competition. Opening up interface, in a sense, is enabling this scale down of barrier. TelCos can open up interfaces at different levels. **Figure 101** shows the most conservative one: the client just pay for application that are executed in the framework of the Operator.

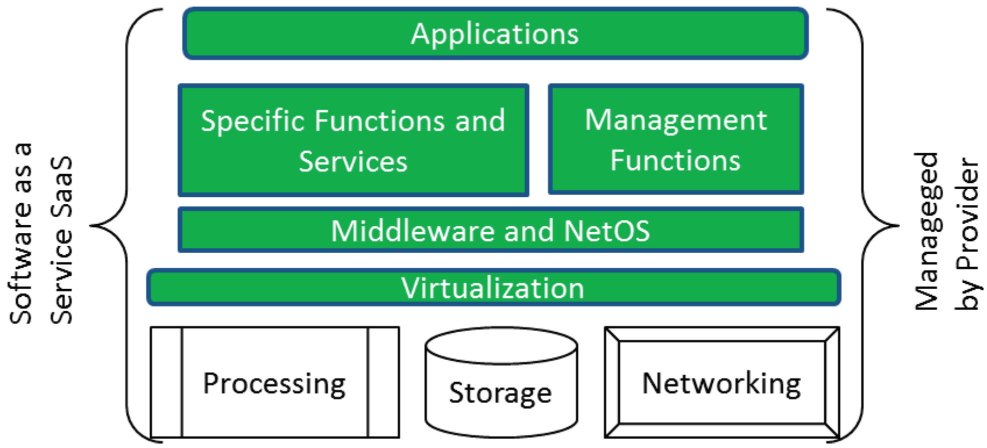


Figure 101: Network and Service functions Seen as Software as a Service

In this case the clients use the services offered by the network as such, the capability to modify the functions are to be negotiated directly with the Provider, while the direct access to resources and low level functionalities is not offered and the "ownership" of the infrastructure is kept in the hands of Operators. This represents the situation in which a TelCo sees its infrastructure and service offering as a sort of Walled Garden totally closed and controlled by the Operator. Another approach is to position the infrastructure as a platform that offers value functionalities and open up APIs for programming them. **Figure 102** depicts this possibility.

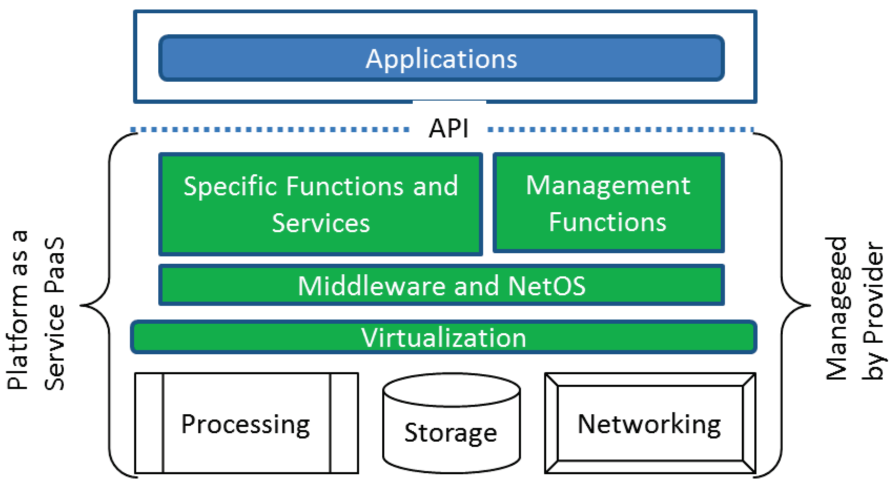


Figure 102: Network as a Platform (PaaS)

In this case, the programmable part of the network comprises value added functions they are specific networking, storage or processing functions that the applications can use and orchestrate in order to reach their computational and communications goals. One example of this possibility is in the context of content delivery networks (or information centric networks): a web company can access and use these functionalities in order to build an abstract CDN according to specific needs. The Operator does not actually disclose the network topology, but it offers the possibility to create a virtual CDN according to a number of local and geographical requirements. The advantage is a higher level of programmability and abstraction for the Client and, for the Operator, a major grip on the actual network infrastructure (that is not disclosed at a physical level).

Figure 103 represents the need to have different APIs in order to satisfy different requirements from potential customers.

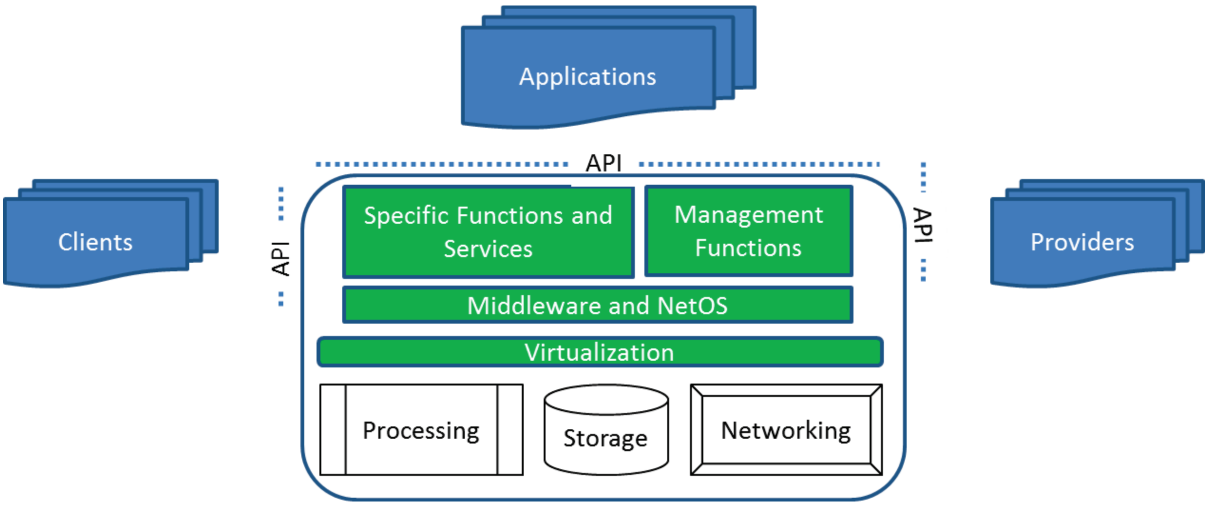


Figure 103: North, East and West bound interfaces for a Network Platform

Figure 103 shows that three different interfaces can be offered: the North interface is following the traditional approach of the Operators, the network infrastructure triggers events and the application running in the Cloud can use this interaction paradigm in order to execute services and applications. The East interface is offered to Providers according to Web technologies and interfaces. Providers can use this interface in order to organize an abstract network according to their needs. The well-known web interactions (client-server relations) will be used in order to regulate the usage of value added functions. In this case, the network functions are seen as services and functions offered by servers, i.e., the providers are “the clients” in a client – server relationship. The West Interface is devoted to final users that can use it in order to adapt the network to its specific needs. Examples of this usage fall in the realm of changing dynamically the characteristics of the access to the network (varying some parameters, e.g., a gamer can dynamically request to the network to change the value of maximum latency). The real commercial value of this interface is questionable, however some P2P applications or communities can use this interface in order to better serve the specific community requirement. Also in this case the API should be exposed making usage of web and/or P2P interfaces and technologies. This arrangement of interfaces preserve to the Operators the entire value of the infrastructure (made out of processing, storage and communications capabilities).

Figure 104 shows a more granular and programmable approach. Here the resources are virtualized and they expose direct control interfaces.

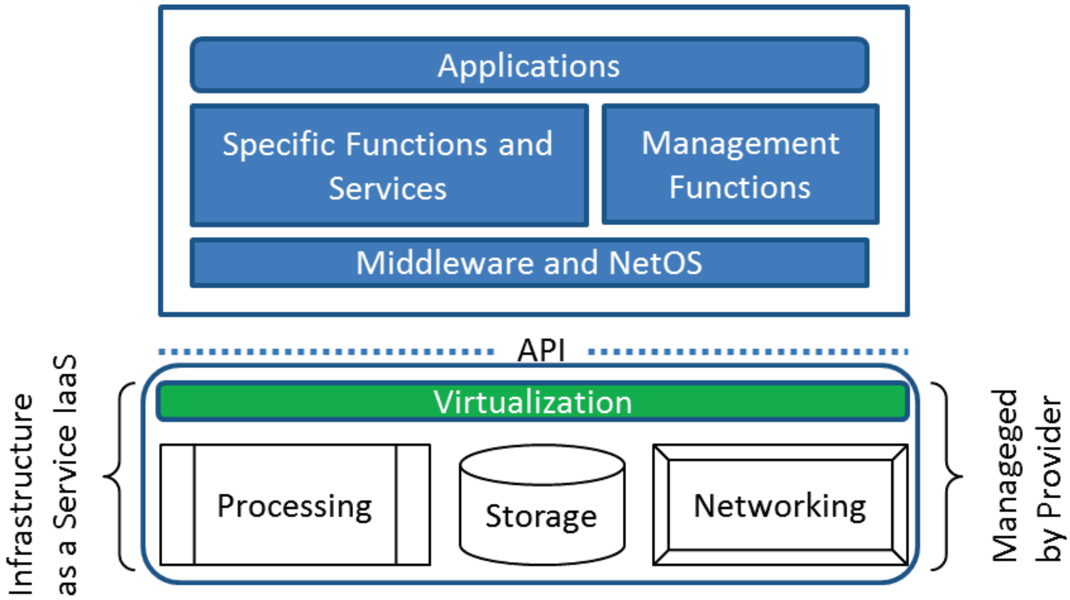


Figure 104: "Network as a Service"

In this case, the programmers have direct access to virtualized instances of network resources. They can program and orchestrate the resources in order to create their own virtual infrastructure. The control on each allocated virtual resource granular. This means that the Network functions are totally spoiled and the control is moved to the programmability infrastructure. In terms of APIs, the possibilities are similar to those discussed for the Network as a Platform case. **Figure 105** shows the three identified APIs.

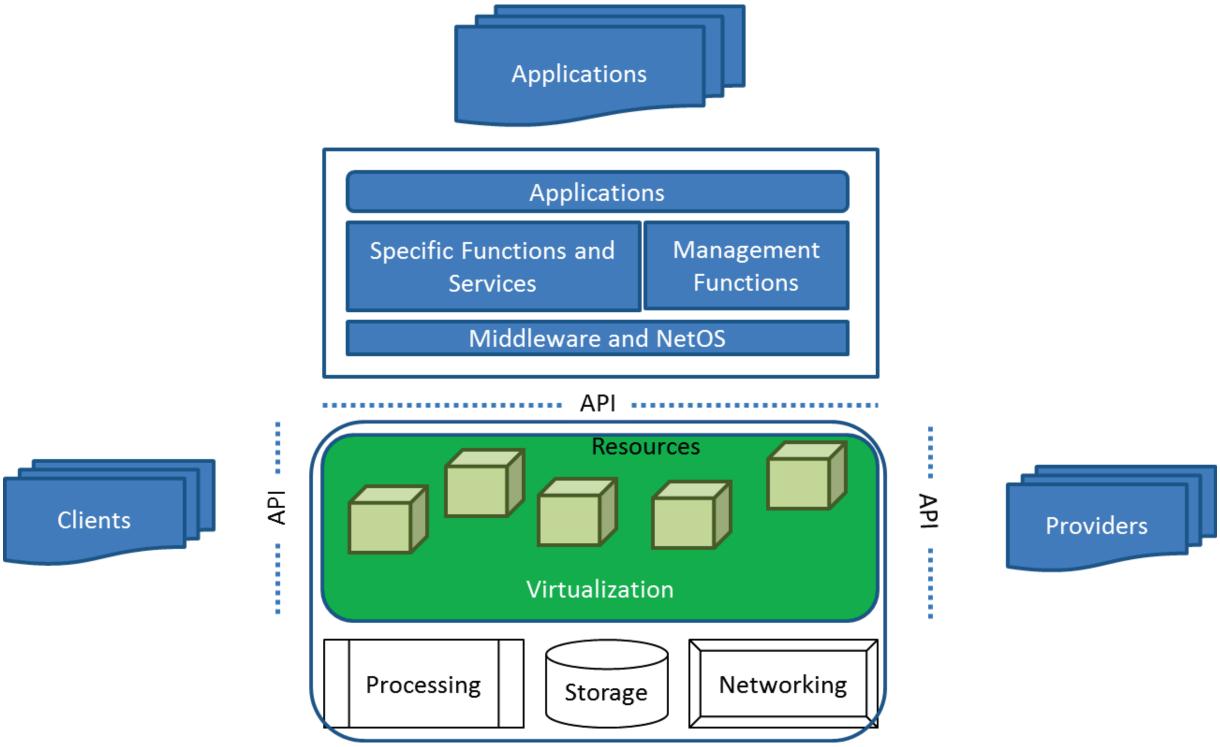


Figure 105: Network Resources as a Service

In this case the North bound interface is defined and used in a "network intelligence" based paradigm, i.e., events and commands are passed though APIs that are capable to deal with this interaction model. Applications and middle level functionalities can be placed in a cloud (eventually hosted by the Operator itself). In this case the control is very granular and the programmers have a full control of the allocated resources. The East interface is offered to Providers that can use this functions in order to allocate and control specific resources and to orchestrate them for the benefit of the global goals of the Provider. This interface is directly determined by the type of resource and eventually it could be offered according to the preferred paradigm (even based, client-server, or peer to peer). The West interface (and also in this case its commercial use is questionable) should offered similar mechanisms of the East bound one. Its envisioned used is for prosumers or clients that want to favor the development of P2P or community applications. APIs at this level provide a granular control of each virtualized resource and they give a great deal of "ownership" to the users. The effect of this is a lowering of the importance of owning the physical resources and the reproducibility of resources. As said, this can break the barrier of intensive capital investment on the network. The combination of softwarization and the generalized usage of general purpose computing can lead to many new entrant to create alternative network solutions or part of networks.

7. 4. The Recurring Myth of Quality of Service and Quality of Experience

The provision of negotiable and dynamic Quality of Service for All-IP Networks is an objective for many architectural definitions. IMS has been defined around the possibility to govern network elements in order to provide differentiated classes of services. Policy Decision Function nodes are able to instruct Policy Enforcement Function nodes on bandwidth management, class of service allocation and other QoS related characteristics. This approach is typically “network intelligence” based, i.e., a small number of nodes (at the control and service level) provide intelligent functions in order to govern several underlying simple network elements. Nodes at the upper layers decide the policy to be executed, while the transport elements enforce the chosen policies. This is a sort of Event – Command paradigm in which the network elements are capable to detect triggering conditions and to send events to the upper layers. Nodes at this level are decision points that determine the logic to be executed and instruct the enforcement points on how to proceed by means of commands (commands can also be pre-defined scripts created using a QoS definition language). This control model is not aligned with the basic concepts of IP communication, e.g., the introduction of a supervising layer of governing nodes is not aligned with the independence of autonomous systems. This approach is difficult to introduce for design, technical, and economical reasons and it could lead to a misalignment in the industry: on one side Operators will force a modification in IP networking, and the networking community will enforce the best effort principle. The best effort mode has allowed the creation of large networks made out of cooperating autonomous systems and the creation of many applications that (using asynchronous functionalities) are able to provide an acceptable level of end-to-end quality of experience. The cooperation of Autonomous Systems (AS) is the key advantage of IP networking, different networks (often operated and controlled by different entities) cooperate in order to find a viable path for routing packets. The shared goal of AS is to maximize the number of packets that can be effectively routed towards a destination (either an end point or another network). This leads to two considerations:

- Users perceive quality of service as the result of cooperation between independent systems (**Figure 106**);

- Enforcing quality of service characteristics into a single network does not guarantee the enforcement of the quality in all the chain of cooperating subsystems.

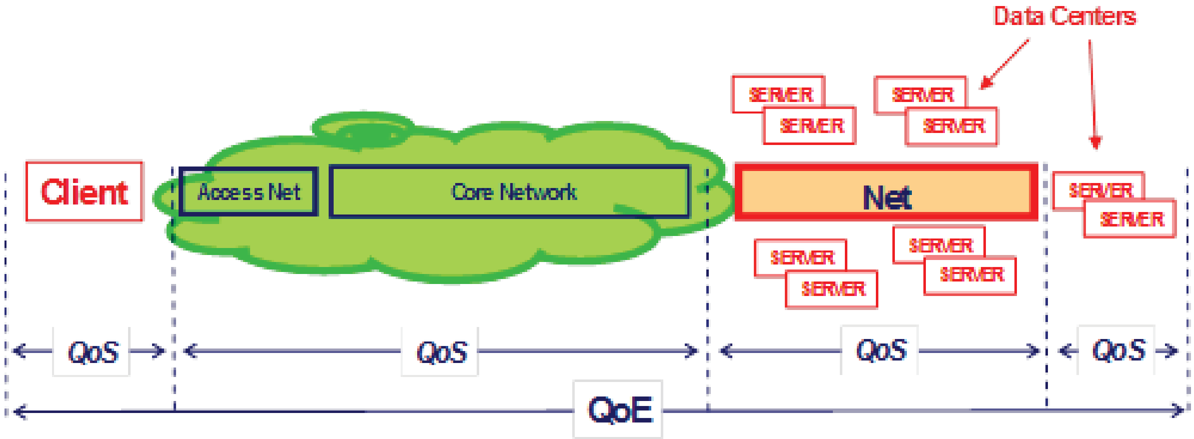


Figure 106: Quality of Experience as Perceived by the User

In the future, these subsystems will not only provide communication, but also processing, storage and "sensing" capabilities. This transformation will exacerbate the problem of governing control and management states of resources allocated to AS. There are many management dependencies among resources. These relationships are not fully update in an automatic manner. Increasing the number of resources will create an even greater number of dependencies and the need to automatically update them will clearly emerge. Moreover, high-level network goals (e.g., connectivity matrix, load-balancing, traffic engineering goals, survivability requirements, etc.) are translated into low-level configuration commands individually executed on the network elements (e.g., forwarding table, packet filters, link-scheduling weights, and queue-management parameters, as well as tunnels and NAT mappings) which are mostly hand-made. For example, IP and Ethernet originally embedded decision logic for path-computation in distributed protocols, whose complexity incrementally grew with the growth of Internet. Growing complexity and dynamicity in future networks pose serious doubts about the efficiency in extending further the distributed control protocols. On one side, the network decision logic should become scalable and more effective. On the other side, low level configurations should be made adopting self-configuration capabilities (triggered by global network decision logic).

Quality of Service and different Networks

The problem of guarantee a system-wide Quality of Experience depends on the ability of the entire communication and computation environment to support the requested features and constraints. As seen in **Figure 106**, this capability has impact on each subsystems (from terminals to public and private networks): each component should be able to support the expected level of quality, otherwise the entire environment will be providing a quality of experience depending from the lowest quality provided by the less capable subsystem. This degradation effect is not the only one to cope with: even if all the involved subsystems were individually capable of providing the expected quality of service, this does not guarantee the final result because each subsystem has to coordinate and cooperate with other "neighbors" in order to achieve the expected result. Except in environments totally under the control of a single provider, the QoE is an issue that has to be tackled with a different perspective aiming at the cooperation of independent systems. The QoE problem should be tackled through the cooperation of Autonomous Systems aiming at providing not only communication paths, but also at cooperating for supporting processing, storage and "sensing" needs of connected subsystems. This is a daunting task, but it is one in line with the Internet approach and it is also encompassing needs and requirements of networked computing (grid and cloud computing) as well as those of the Internet of things. A federated approach in which a subsystem is able to provide a set of functionalities to others and its cooperation for achieving a common goal should be the basis for approaching the QoE problem. There is another issue related to QoE in this large distributed systems: complexity. A network of networks tends to collect a number of heterogeneous systems and nodes that create the need for very complex integration and management. With the increase in number of nodes, the traditional management of large system will fail short, so there is the need to determine how the components of a network of network should be organized, configured and managed in order to provide better QoE. Actually virtualization and segmentation of functionalities can allow the Provider to request to the Network Provider to allocate virtual resources that support certain level of expectations. These resources can span over different administrative domains and can interoperate in order to provide the required level of service (even following a best effort approach, in fact the single environment will adhere to the current way of providing services, but the virtual infrastructure could be over equipped in terms of (virtual) resources so that the negotiated QoS parameters can be meet (at least statistically).

7. 5. Findings

This section has examined some of the core beliefs of many TelCos. Network Intelligence, Exposure and QoS are persistent icons of the past. Some TelCos believe that they are still key differentiators with respect to new services. In this perspective the important thing is to show that their proper usage can return a lot of value to Operators allowing them to recover importance (and revenues) in the market. This Chapter, however, has shown that the current platforms are obsolete (their definition is more than ten years old) and there are new technologies (such virtualization and software defined networking) that can provide better programmability and can support the deperimeterization of services (from the network). The IMS has been conceived as a novel architecture for providing session based services. For session based services, the old traditional telephony services are mainly considered. IMS can fully support them, however newer architectures like P2P have clearly demonstrated that other technologies can now be more efficient and less costly in order to support these services. IMS has been also proposed for multimedia services (Jain and Prokopi 2008). Surely the platform is capable of sealing with them, but the market is much more oriented towards specialized solutions devoted to IPTV and multimedia services in general. IMS has also been considered as a viable platform for supporting other services (like the M2M ones). Also in this case, newer and more advanced solutions can be used. For instance Twitter could be transformed into an efficient enough engine for many IoT applications. The control paradigm based on a network capable of triggering events towards the services layer is adaptable to many situations and services, but the fact that services are depending on connectivity and related basic call model is too much of a burden. NetAPIs or Exposure is another TelCo activity that started too soon and never found a convincing path to the market. Many TelCos were ready in early 2000 (just when the new Web 2.0 trend was taking shape) to release Parlay based APIs. What stopped a large development of them? Essentially three considerations emerged: why to offer network control functions if nobody was pressing for having them ? What is the right charge for these functionalities ? What are the security risks ? These three consideration stopped almost any initiative in this area. It is singular that after more than ten years, many Operators are re-proposing very similar approaches even if it is evident that two of the major attempts in this area (rebbit and bluvia) have failed. Proposing this approach now is not a winning proposition: even less people are interested in call or session based control (and if they are, they find appealing interfaces somewhere else), There is a lack of standardization and APIs are fragmented. Initiatives like oneAPI of GSMA are lagging behind and they are not delivering interesting enough APIs for developers to use. Lastly, the Quality of Service issue is a driving concept within many TelCos. There is still the opinion that QoS is a property requested by the market that differentiate the TelCos from other providers. It is difficult to change this attitude. And many TelCos do not want to learn lessons from the past. For instance, when the Google maps were first released, users were experiencing very long waiting times before the maps were downloaded. And when a user was scrolling to fast a part of the map, there was a long delay before the new part of the map was shown. At the time this was taken as an evidence that the Internet needed to have more bandwidth and it should be also supported by QoS properties. IN other terms Operators were saying that in order to solve the problem the network should provide more capacity and this capacity should be governed by means of signaling and policing. The same problem has been tackled and solved by the web in a totally different way: the browser should asynchronously request the most probable requested pieces of the map before an explicit request of the user. Simply put the browser is downloading more data that actually shown to the user and they are used as buffers while other data are downloaded in order to fulfill a user request. This asynchronous request was the basis for a winning web technology: AJAX. This example shows that the attitudes of WebCos and TelCos are deeply different and one is promoting solutions at the edge, while the other is enforcing policing and capacity in the network. With the increasing capacity offered by the new networks and with the growing capabilities of terminals and servers there is the possibility (as demonstrated by the Amazon’s Silk solution) that the cooperation between the terminal and the cloud can solve many communication issues. Actually this seems to be a very promising possibility, depending on the available network resources, processing and flow of data can move towards the terminals (if the conditions allow this) or towards the cloud that will take care of optimizing the delivery of information to the terminal. The real progress in distributed applications and for new classes of services relies on the fact that the network resources have to be programmable and they should be virtualized. By means of these two properties, virtual networks tailored for specific needs of applications can be created and programmed. For instance a Content Delivery Network can be virtualized and can be deployed over an SDN based infrastructure. Nodes become programmable and they can be migrated also towards processing infrastructures not related to the Operator. In this way deperimeterization of the services from the network can be pursued. As a conclusion it could be stated that neither the NI Platforms, nor the API Exposure nor the QoS platforms are mechanism capable of solving the issues that Operators are facing nowadays.

Part 3. Scenarios, Viable Architectures and Future Work

Introduction

The technology and business evolution of the telecoms industry can lead to different outcomes depending on the ability of the Operators to Leverage their assets. In this part some scenarios are constructed based on the likelihood of the success of different technologies and business model. They are the driving forces of change, i.e., those capable to revert the mainstream (or to create one) giving new opportunities to the Actors of the telecoms environment. The capability of Operators to play innovative roles and deploy innovative platforms is addressed (and questioned). The role of Operators is incarnated by specific architectures that fully support the needed functions to play a specific role in the ecosystems.

This part contains two chapters:

Chapter 8. presents the driving forces shaping the evolution of the telecoms industry and from them five possible scenarios (from the perspective of an Operator) are examined in terms of enabling platform.

Chapter 9. aims at providing further details of the most promising architecture and maps the thesis achievements with some fundamental questions posed in the initial chapter. Finally Future works and directions for this thesis are introduced and commented.

Chapter 8. What Are the Possible Scenarios that Could Occur as a Consequence of the Technological Evolution?

Highlights

- Driving Forces of Change
- Scenario Identification and Construction
- Scenarios: TelCo Roles and related Architectures

8. 1. Introduction

Predicting the evolution of technologies is difficult, but predicting the combination of technology, market, and society is a daunting task. In fact, many good technologies that could have an appeal on the market have been disregarded by the users in favor of poorer ones (e.g., Betamax versus VHS). In the success of a technology on the market unpredictable behaviors and attitudes (as well as cultural differences) of the users have a great impact. However having discussed in the previous sections the technological evolution and the current business status of the Telecommunication Industry and, at a lesser extent, the one of the ICT market, an initial understanding of the possible evolutions, on the long run, of this sector in turmoil is possible. This could be useful to plan strategies on how to cope with competitors and act accordingly.

This Chapter first tries to determine simple mechanisms and methodology for defining a few interesting scenarios. They are based on an analysis of the forces that are driving the transformation of the ICT industry. Some of them have been discussed together with their potentiality in the previous Chapters. These drivers of change and trends are collected and organized in order to evaluate their likelihood and the disruptive impact they could have on the industry. The most likely will define the “linear evolution” of the Industry. So it is more interesting to identify those driving forces and trends that are less probable but that carry more disruptive potential. For instance some of the technologies discussed from 0up to Chapter 5. are very disruptive: the combined effect of the processing power made available in the future and the possibility to virtualize and to move to software functionalities that so far have been implemented in hardware are surely defining a set of inflection points that can determine an innovative path in the ICT evolution. These and other trends will be taken into consideration in order to determining a set of transformation scenarios.

The time frame chosen for this analysis is ten years. This guarantees that the scenario finding are not too short in time and new emerging technologies have enough time to deploy and make manifest their disruptive potential.

The foreseen scenarios take into consideration a conservative approach in order to understand whether the current approach of many TelCos can hold valid in this time frame, a more pessimistic scenario in which all the decreasing economic indicators will keep decreasing, two scenarios that consider whether and how a TelCo could play the specific and rewarding roles of Platform or Service Provider, and, eventually, a very disruptive scenario in which the TelCo will disappear. This range of scenarios cover a large spectrum of possibilities and also consider a large quantity of identified trends and driving forces. Surely the scenarios are not complete but they can be used in order to reason about strategic options in future roles and possibilities viable for TelCos.

These scenarios should be considered as tools to be specialized and further extended depending on the level of analysis that needs to be achieved. In addition the scenarios should be somehow interpreted on the basis of the stakeholders and the type of TelCo. Small and Regional TelCos will probably consider these scenarios in different way compared to large and internationalized TelCos. A few hints related to the possible usage of the scenarios tailored on the type (and the footprint) of Operators are given in the various Sections. Also the geographical provenance can play a fundamental role in the scenario analysis. It is likely that some of this scenarios are not (yet) applicable to many TelCos operating in developing Countries. In those contexts the markets are still growing (especially in the mobile sector) and the issue of revenue decrease is not becoming evident. Insights of this analysis are devoted to a medium-large TelCo (i.e., a TelCo that is the incumbent in some countries, it is operating in some saturated markets (developed countries), it has an international footprint). Some considerations can also apply to Large Operators present in several markets and to some extend to small regional TelCos. For larger Operators there could be the case to potentially play different roles in different markets. However this is a dangerous proposition because fragmenting the platforms and their usage decreases their value.

8. 2. Scenario Planning

Scenario planning is often used in order to reason about different perspectives and possibilities that an industry or a (political) situation can lead to. In literature there are many examples of it (Varum and Melo 2010). In this document a light approach to scenario definition is chosen, it is based on the Mobena approach (Battistella, Colucci and Nonino 2012).

Throughout this section a simplified methodology is used in order to determine the possible evolution of the Telecommunications industry. The perspective used is the one of the (established) Telecommunication Operator, TelCo. The simplification is essentially related to business and technological considerations and analysis. Actually, technological issues and related decision points are the focus of the scenarios definition.

In this analysis, a Scenario is a description of forces and trends that can shape the evolution of an industry. They have been determined and analyzed in Part I and Part II as technologies and approaches to services and markets. herein this section, they are organized in a homogeneous way in order to draw some insights about future directions of the Telecommunication industry.

The basic points of the scenario analysis are:

- A challenging question to answer (the reason for scenario analysis)
- The selection of a meaningful time frame for analysis (giving the right time to phenomena to occur and show their impact). A typical time frame is ten years.
- The involved stakeholders
- A map of driving forces and trends. This is focused on technology issues, but it tries to include business, regulatory/legal, and societal trends.

8. 3. Timeframe for Scenarios

The chosen time frame for analysis is 2023 and beyond (for those scenarios in which the technological evolution is more blurred). The reason for this is that a period of ten years seems to be long enough to see the effects and impacts of current technologies and business trends without including trends and developments that are not seeded today.

In addition, this time frame is consistent with the possible “endurance” of current approaches of established TelCos. If the erosion of the current major business model (and revenue) is keeping its pace, e.g., between 2% per year, in the period between 2020 and 2022 an erosion between 20% and over 50% of the total revenue will have heavy impacts on the possibility of many TelCos to stay in the market. **Figure 107** depicts this occurrence with several percentage of decrease.

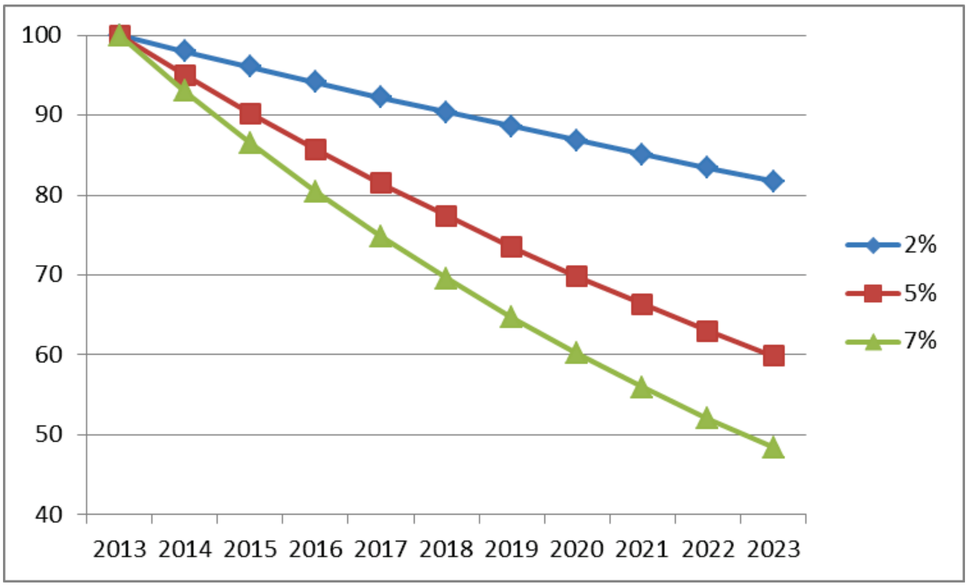


Figure 107: Continuous Decrease of TelCos Revenues

In addition, in the same time period, TelCos are required to largely invest on infrastructure because of the evolution of networks (from 4th generation on for mobile, and for fiber networks in the fixed). TelCos are investing heavily especially on access technologies (either mobile and fixed) aiming at bringing a lot of bandwidth to customers. As already discussed, the pure offering of connectivity has negative effects on the traditional business model of Operators. In fact many established TelCos are experiencing a new state of play: on one side, the price for Mb/s is constantly decreasing[47] (and the “*all you can eat*” approach is applied also in the mobile market that traditionally has given a high return to

[47] For example http://www.itu.int/ITU-D/ict/ipb/index.html

Operators); on the other side, the use of networks is constantly increasing (CISCO 2012). This structural change in the telecoms market and industry has already been described as the “network paradox”: the more capacity is provided by the network, the less it produces revenues.

In **Figure 108**, the hypothesis that a TelCo[48] with 57.4% of Gross Profit Margin (i.e., Revenues – (Capex + Opex)= 57,4) is increasing its Capex + Opex (i.e., it is investing in new infrastructure or is spending more on operating the network), it is keeping it constant (i.e., investments are compensated by cost saving on operations) and it is reducing it (i.e., the Operator saves more that it is investing in new infrastructure) are presented and associated with possible decreases in revenues. A margin of around 50% is lower that the value obtained by the best in class, but it seems a reasonable average between several Operators[49].

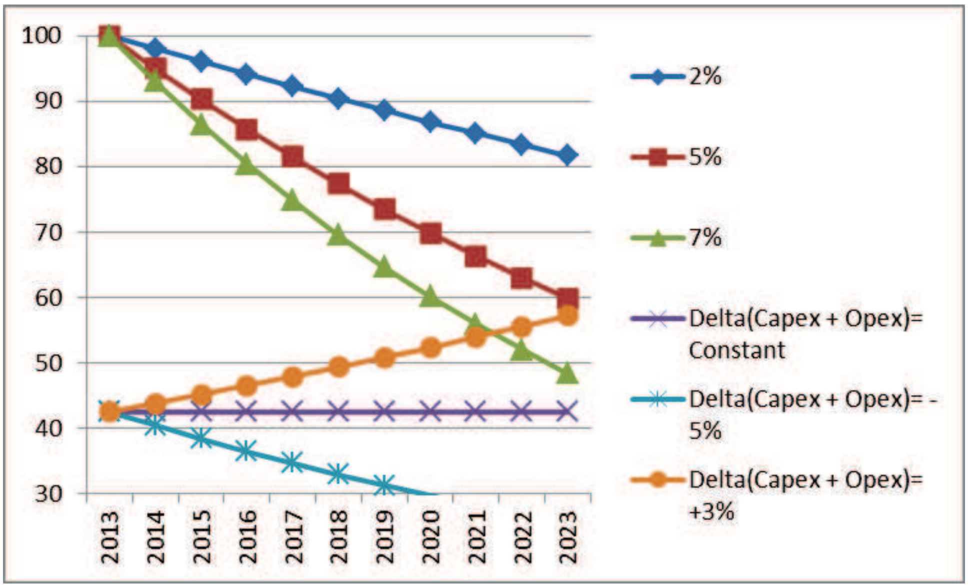


Figure 108: Possible Relation between Revenues and Capex + Opex

It is evident that the decrease of revenues together with the increase in Capex + Opex make the TelCo Business unsustainable in the time period considered in this document. The approach of the Lean Operator (see previous sections) is compatible with sustainability of the business even in presence of a decrease of revenue. However, the progressive reduction of margins make the Operator business less appealing to stakeholders and to the market. In the longer terms, beyond 2023, if the revenues curves do not change direction, the business

[48] According to [http://ycharts.com/rankings/industries/Telecom%20Services/gross_profit_margin], France Telecom has a margin of %7,4, Telecom Italia is on 56,3% while Telefonica is at 70.815 and BT is at 14.10%.

[49] See http://ycharts.com/rankings/industries/Telecom%20Services/gross_profit_margin

will move towards commoditization and after 2013 a heavy intervention of the Regulator or National Governments will be needed in order to support the TelCos role. As such, this is a structural transformation of an industry. This situation will not occur if the telecommunication industry finds new ways for generating revenues keeping the same structure and ecosystem.

This state of things is very similar to what happened in the computer industry due to the Moore Law's: a continuous increase in processing power and lowering of associated production costs have eroded the margins. This evolutions brought out of market many historical computing companies like Burroughs, Sperry, Digital, and the like. The combination of Moore's, Kryder's (for memory) and Koomey's laws promise an ICT world full of capabilities and able to consume less and less power. This is particularly appealing and important for the device and mobile industry that will be able to provide to customers highly performing systems at lower prices. For instance the projected needs of bandwidth in the household is questioned by many Actors of the Telecommunications Market (Saracco 2008). In fact some of them foreseen an incremental growth of fixed bandwidth request that can be satisfied by means of new technologies that leverage the copper wiring instead than the fiber. If this proves true, some Operators can reduce the Capex and delay a massive fiber investments after 2020 (Loveland 2012).

In the future there will be an increasing difficulty in sustaining the current (i.e., traditional) business model in Telecoms. This means that either the TelCos will adapt to be commoditized or they re-invent themselves. One possible (and conservative) way is to find a means to invert the trend and to revitalize the connectivity business. The other way is to adapt to the new scenario and move competition and efforts in new sectors by leveraging new business models and leveraging the available connectivity. These two scenarios are the fulcrum of this section and they are important to determine the scenarios.

8. 4. Driving Forces

According to the Mobena methodology (Battistella, Colucci and Nonino 2012), the driving forces of evolution need to be considered. The most likely of them are part of the mainstream and will determine a conservative scenario, a scenario foreseeable by applying the current understanding of the trends. More interesting scenarios can be defined by pulling together less likely evolution trends and driving forces. Their combination could let emerge disruption and big transformations. These scenario are more difficult to determine and describe, but are those richer in terms of insights and foresights.

Some major technological and market trends and assumptions emerged from the discussion of the previous chapters, the major ones are summarized below:

1. Commoditization will lead to software solutions for Networks (SDN, Autonomics), and the barrier for network deployment will be lowered;
2. TelCos will fight against commoditization by pushing for QoS related connectivity;
3. TelCos will try to leverage current platforms and approaches (SDP and IMS) in order to provide network based services. These platforms will be progressively "cloudified";
4. Borders between Data Centers and the Network will blur;
5. TelCos will try to enter into new markets leveraging connectivity;
6. TelCos will try to become Platform or Service Providers;
7. New Services Classes will emerge as market champions;
8. IOT will be one major business challenge for TelCos;
9. Terminals will be nodes of capable Networks;
10. New Devices will enable new smart environment and edge networks;
11. IoT and Pervasive communication will create a new gap between different areas (well-connected cities vs. rural areas)
12. Network will be regulated by Utility ("pay as you go") or subscription business models (unlimited connectivity for a fixed fee), Regulation could intervene in order to protect Network Investments;
13. Software and virtualization will drive the future of communications;

A larger number of issues and topics emerge from the analysis of the ICT market and evolution. They are not exhaustive of the industry analysis but they represent a some of the possibilities given by the foreseen future of this industry.

They are briefly described in **Table 5**.

Table 5: Driving Forces in the ICT Sector

Driving Force	Explanation
Terminals as Personal Platforms	Terminal capabilities will be so sophisticated that they will become a preferred platform for delivery of services

Powerful Terminals	Increased capabilities of terminals (increasing storage, processing, communications and sensing capabilities)
Data Increase	Data traffic will steadily increase in fixed and mobile network. High capability in network a stringent requirement
Free Voice	Voice and instant communications will be essentially free and provided by Web Companies
Decreasing price for Mb/s	Price for Mb/s will tend to decrease giving a smaller margin (considering the return of investment for new access networks)
BitCarriers	Telcos will essentially be Bit Carriers
Direct Comm	Direct communications will have a grip on user becoming a major feature
Regulation protects NetsInves	Regulation could impose rules to safeguard the value of investments in networks
Edge Services	Services will be provided essentially at the edge and sometimes they will be provided exclusively by users
SrvcsDeperimetrization	Telcos will be able to deperimetrize their services from the network becoming able to provide global services
DCs as Nets	Data Centers will encompass a large part of networked information and resources. Blurring the borders between Clouds and networks
P2P Services	P2P services will keep growing and will consolidate as a means to create and provide highly distributed services
P2P Platforms	P2P platforms will emerge as alternatives to Client Server DataCenters
Network Services	Telcos will be able to promote networked functionalities as successful services
IoT Services	IoT services will be a major goals for many Providers (IOT Market growth)
QoS	Telcos will sell QoS to customers (mainly businesses)
New Service Classes	New Service Classes will emerge as market champions (e.g., e-gov, augmented/virtual reality, social media, IoT, …)
IMS +SDP	SDP and IMS will be able to control the network provisioning of new classes of services
SDP Services	SDP and Telco Platforms will provide compelling services (mainly communications)
NetAPIs	Network APIs will be (widely) used in order to build networked services
IOT Platforms	Telcos will provide IoT platforms that provide appealing services beyond connectivity
New Markets	Telcos will enter in other markets leveraging Connectivity and Platform capabilities
SWEE	“Software will eat everything” (interview to Marc Andreessen), SWEE
Telcos as SP	Telcos will play a significant role as Service Providers, i.e., they will be capable of offering compelling networked applications that the market will pay for
PlatfEnablers	Telcos will play a significant role as Platform Enablers, i.e., Operators will be able to offer compelling networked functions that the market will be willing to pay for
BW increase	Increasing bandwidth for fixed and wireless networks, i.e., the Telcos will be capable of deploy and successfully operate new networks

Privacy	Privacy of data will be an issue having an impact on how Enterprises conduct their business
Legal Issues in Privacy	Privacy will be a major concern (see other point in the data section). Two opposing trends will face: privacy as a customer requirement (customers will pay for it) vs. privacy as a citizen right. Machine tracking will offer even more mechanisms to profile the customers
Edge Complexity	Networking complexity will move to the edge where terminals and heterogeneous access networks will compete in order to offer connectivity and other value added functions to customers
Pervasiveness	Pervasiveness of communications, processing, storage and sensing capabilities
Devices for Artifacts	New devices will help people to "create" things and artifacts (e.g., micromachinery)
BM: Value to Ideas	New business models will emerge in order to leverage the possibility to create objects and artifacts (value to ideas)
EU Companies penalized	The different legal approaches for the Internet around the world will allow to many Providers to bypass the most stringent regulation (European Web Companies will be penalized compared to US and Asian one)
Artifacts created by Apps	Services and Apps will help people in creating or personalizing their of artifacts
Net Resources Virt	Extensive virtualization of Network Resources
New Devices	New classes of devices (based on sensors and actuators) will emerge
ABC (terminals)	Capabilities to detect and use the best networks by terminals will lead to a new way of providing the Always Best Connected, ABC, service
PSCAS	Integration in the network of virtualized processing, storage, communication and sensing/actuating resources PSCAS
Several Access Nets	Competing heterogeneous access networks (especially for Wireless)
BM: Servitization	Servitization will be a major trends and a means to keep an edge on competition giving value to easily reproducible artifacts
SmartCities ==> digital gap	There will be an increasing digital gap between cities and rural areas (smart city initiatives will contribute to increase the digital divide)
Autonomic Devices	Autonomics and cognitive capabilities in the terminals and edge networks (i.e., self organization)
Virtualization in the cloud	Trend towards virtualization of resources in the "cloud"
Ubiqutous&Free Local Connectivity	Local connectivity will be essentially free and ubiquitous in highly developed and inhabited areas
SDN	Software Defined Networking will be widely deployed in order to save OPEX and CAPEX
Pay as you go	"All you can eat" tariffing will be superseded by pay as you go
Autonomic Nets	Autonomics networking will be a predominant technological trend
Machine traffic > human	Traffic generated by machines will not exceed the traffic generated by human until machines will not exchange multimedia information (forecast: machine generated traffic will exceed the traffic generated by humans in the considered timeframe)

Virtual Terminals	Virtualization of devices in the "cloud" (Virtual Terminal)
Crowdsourcing for artifacts	Crowdsourcing will emerge as an important factor for creating and producing artifacts
Flat Net	Public networks will be streamlined and simplified (streamlined network architecture)
Social Nets key for users	Social Networks as a key interest of users for accessing to local, global, contextualized information and services
Terminals as media sources	(Mobile) Terminals as major means to produce and consume media data
More Machines than Humans	There will be more machines connected to the internet than personal devices
Walled Gardens	Actors will strive to impose to the market their Walled Gardens (e.g., Internet of Things will be dominated by proprietary solutions)
Crowdsourcing for services	Crowdsourcing will emerge as an important factor for providing services
Social M as next dialtone	Social Media as the next dial tone (people will want to get always in touch with their correspondents) and will want to share continuously their info
All Service in cloud	Services will all be cloudified and made available through web mechanisms
Quantified Self	Services and Apps will help users in controlling their personal habits, functions and data (Quantified Self)
DC issues	Data Centers increasing their size and power, issues in needs for communication between sites and to customers, power consumption and heating. This will lead to the need to find more sustainable solutions
Global AppStores	Application Stores will have a global dimension and market or will lose importance. A few big App Stores will prevail
Social Flavor of Apps/services	Services and Apps will maintain and increase their social flavor
Availability of Personal Data	Increasing importance in personal data and their ubiquitous and continuous availability
Freemium	Freemium will remain a prevalent business model for Internet services
Specialized AppStores	Only very specialized application stores will survive (and possibly they will compete with the most important brands ones)
Decrease in Apps price	Prices of Apps will generally decrease
Premiums Apps keep an edge	Premium services and Apps (as games) will keep an edge
Smart Environments	Services will be provided in the context of smart environment capable to contextualize the services to the needs of the specific user.
Piracy in Dark Nets	Piracy will be constrained into dark networks (uncertainty)

These trends are represented in this **Figure 109** in an organized view: they are ranked for their importance from the impact point of view (X-axis) and for their probability (Y-axis).

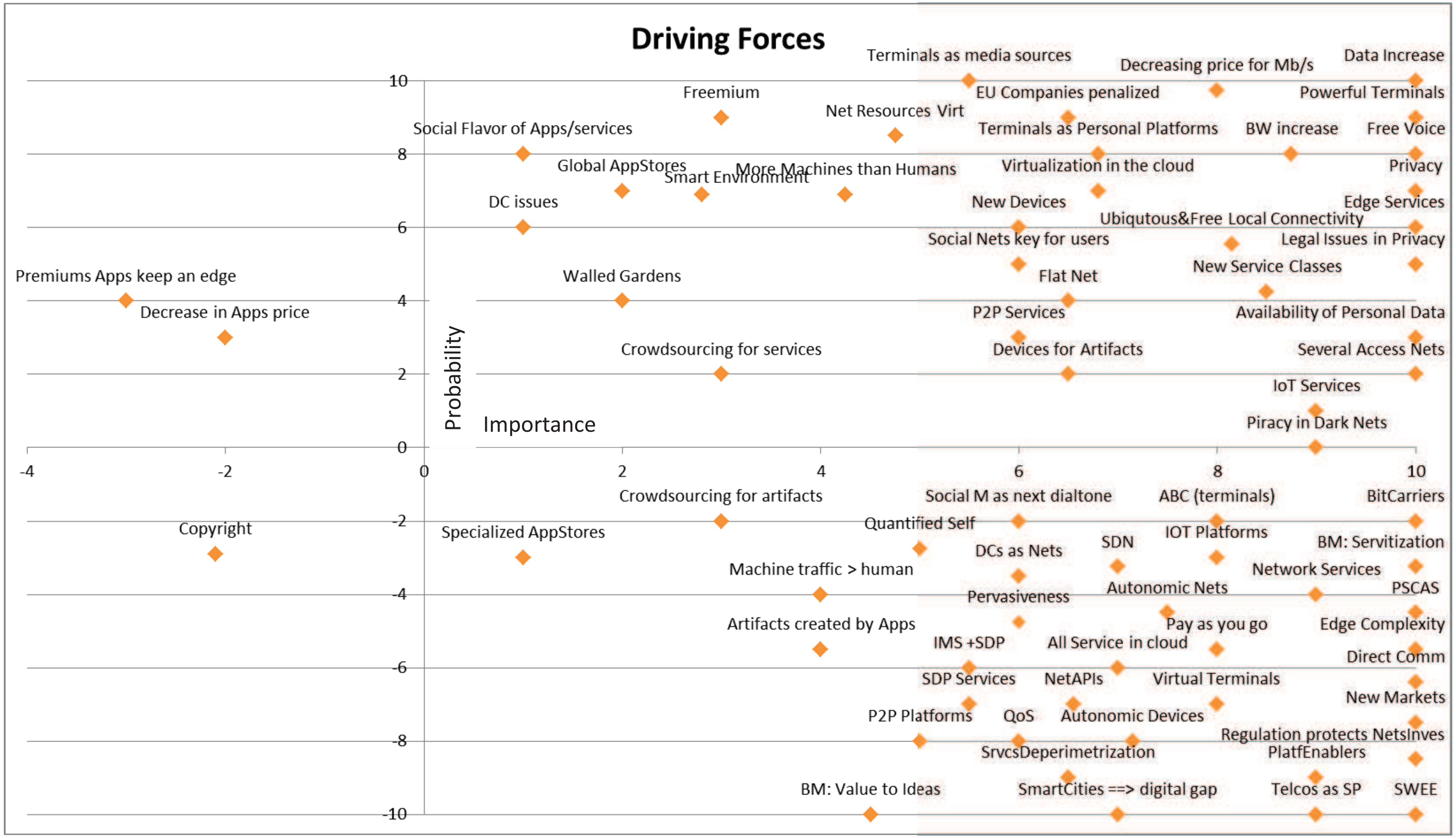


Figure 109: Some Possible Techno-economic Driving Forces

8. 5. Scenarios Characterization

In order to characterize the possible scenarios, some trends have been considered with respect to the possibility for TelCos to play relevant roles in the ICT industry.

Depending on the role that the TelCo intends to play and the relevance of its assets and the likely technological and market evolutions (the trends), five scenarios have been identified and considered in this section:

- a conservative one in which the TelCos will retain a certain importance in the value chain of ICT services and infrastructure pushing for QoS based communication;
- a pessimistic one in which the TelCos will become commodity providers; an optimistic one in which TelCos can act as successful Service Providers;
- a positive one in which TelCos can position their assets as a platform for other to build services;
- a disruptive one in which advances in alternative networking technologies will change the communication industry.

Figure 110 depicts the identified scenarios.

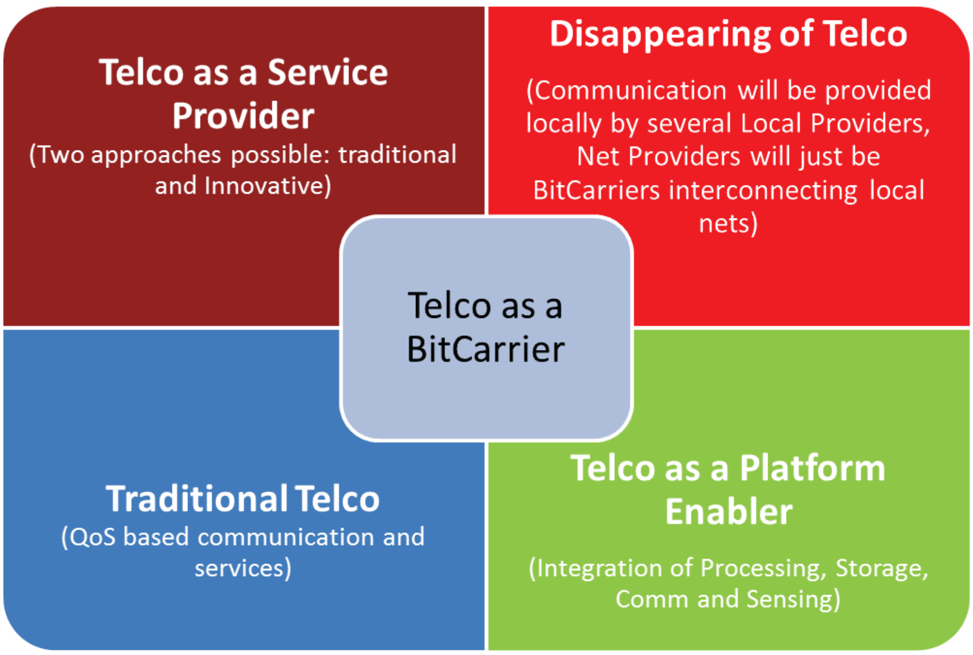


Figure 110: Identified Scenarios

8. 6. Traditional TelCo: A Quality of Service Based Scenario

Many TelCos are trying to maintain the current role in the value chain of ICT industry in spite of the changes it is undergoing. This type of TelCos are trying to leverage the traditional network assets and, in addition, they are pushing for a Quality of Service approach. In other terms, they are trying to "sell" managed connectivity to customers (essentially to companies and providers) together with some network based services. The trends and technological evolution impacting on this scenario are many (see **Figure 111**).

8. 6. 1. Favorable Trends

Such an Operator can count on some assets such as increase demand for bandwidth, the possible interest for negotiated QoS, the always best connected features, the possible success of NetAPI and the combination of SDP + IMS. First of all, the increase in bandwidth consumption and the possible scarcity of connectivity in certain densely populated areas can request bandwidth as a managed service, i.e., some users and especially service providers could request the reservation and availability of a part of the available bandwidth for their customers, leaving the rest to the best effort internet. In this context in which the bandwidth is a scarce resource, the network infrastructure could be orchestrated in ways to provide resources to premium customers. IMS is capable of providing these feature and hence it could be a viable solution and architecture. On top of it some added value services could be created in order to differentiate the service offering. In this case, the integration of the SDP technologies into the cloud could have the benefit to provide the SDP services and functionalities in a XaaS manner. New services could include multimedia capabilities, but also new classes of services like IoT. In fact the traditional TelCo could benefit from the evolution of M2M technologies towards the IoT and propose its infrastructure as the basis for this kind of services. Also Network APIs could be leveraged in this very traditional scenario. A good example of these possibilities is given by the Fraunhofer playgrounds related to IMS (NGNI Group 2005), SDP (NGNI Group 2006) and now also smart communication (NGNI Group 2012).

The business models supporting this scenario are the well-known and stable ones: subscription and "pay per use". Also in this case the advantage is the TelCo are leveraging the existing infrastructure and know-how. This approach could be pursued maintaining the current ecosystems of vendors.

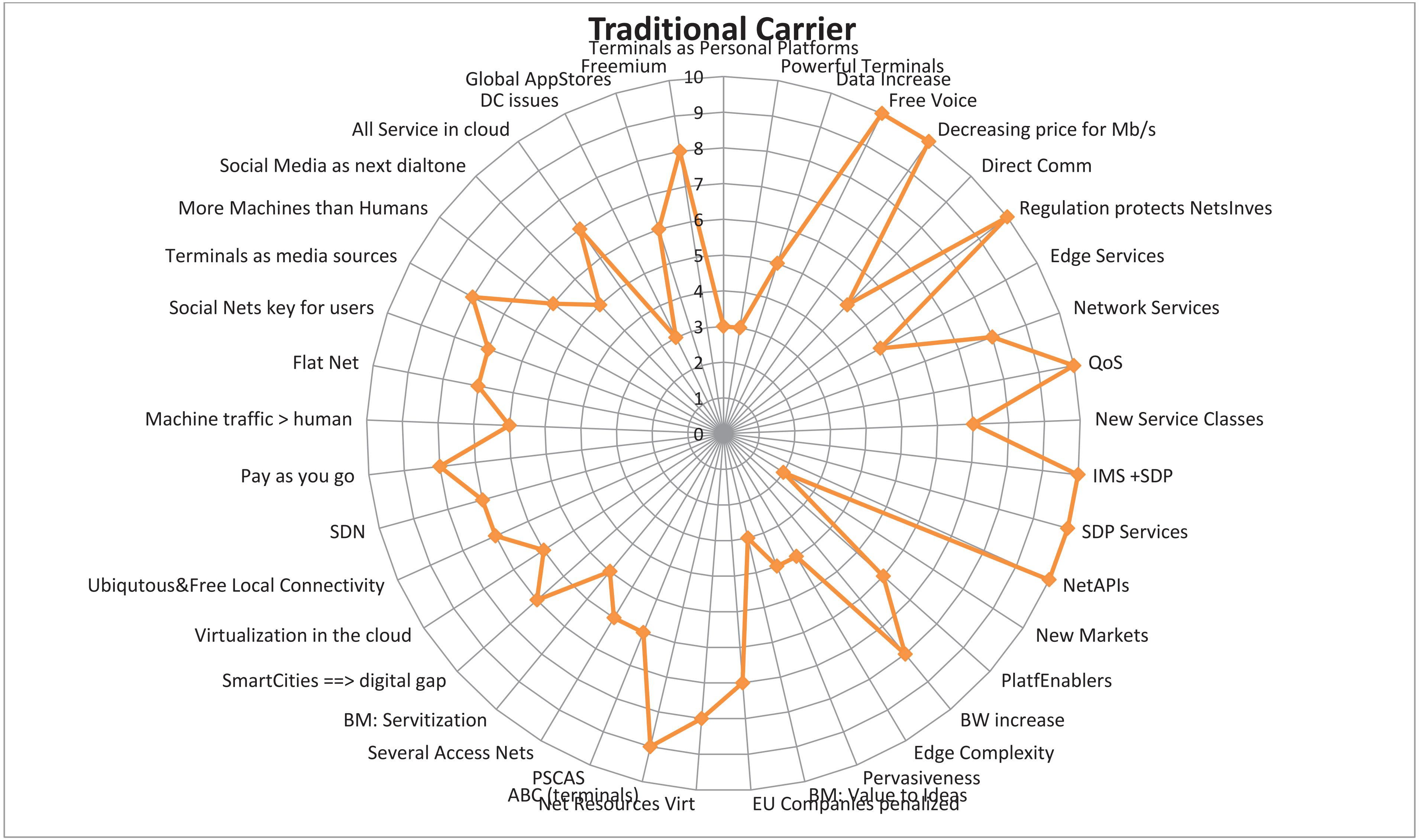


Figure 111: Some Trends Impacting the Traditional TelCo Scenario

8. 6. 2. Critical Trends

This scenario suffers from the current problems of the TelCo approach. Social networking is not properly grasped and nurtured, communities are and will remain outside of the perimeter of the TelCo. Web Companies are not going to use this infrastructure (IMS) because it is technically too far from their technologies and needs, and it is not leveraging the Internet, but the specific Operator architecture. Free and pervasive communications will flourish at the edge of this network intelligence based infrastructure and clients servers will be capable of bypass all the intelligent functions and relay on the end – to – end principle: intelligent functions will move to the edge. This will be especially true for the APPs market, and the trend will be exacerbated because the IMS infrastructure of several Operators are fragmented. In addition, TelCo's services will be regionalized and confined into the domain of the Operator networks. Deperimeterization will remain of property of the Web Companies. Services will be implemented with a network flavor and as already evident, only professional developers will try to take advantage of the Net APIs. The SDP + IMS framework can be used to develop new classes of services, but this usage is not straightforward and clean. In fact a lot of adaptation and complexity will be to be managed. The cost of developing services will be kept at the normal standard of telecommunications costs (higher that IT) and usually the development costs of services will not justify the investments in terms of expected returns. The network intelligence paradigm is definitely very expensive and difficult to adapt.

8. 6. 3. The Envisaged Architecture for This Scenario

In order to achieve the goal of a telecom oriented control platform capable of governing and optimizing connectivity resources and service functions, the network intelligence paradigm and platforms can be pursued and used. This infrastructure can be differently implemented, it is based on a few macro-components like IMS, Service Delivery Platforms, and a set of well-formed APIs (i.e., NetAPIs) possibly integrated in a Cloud Computing infrastructure. There is quite some reasoning behind this integration of macro-components:

- IMS is seen as a means to control the underlying network infrastructure and governing the connectivity resources in such a way to optimize their usage and to provide the agreed connectivity to customers. The type of control exerted on these resources is of the type master - slave (in which physical resources are slaves and all the logic and intelligence is moved towards networked control systems). IMS is key for leveraging the network resources.

- Service Delivery Platform is a distributed platform able to provide programmability and service composeability. In coordination with IMS it could provide a means to leverage network resources and provide network centered services. It is also seen as a means to integrate and leverage BSS and OSS capabilities (e.g., accounting and billing).
- Net APIs is a set of Network supported APIs that can be used by the TelCo or by developers for composing services taking advantage of the infrastructure and network functionalities. This is seen as a key aspect for creating a vast ecosystem of developers associated to the TelCo. Actually some Operators are differentiating between mass and business market sectors: services for mass market will be strongly dominated by Web Companies, while the business market is still largely supported by TelCos. So the idea is to develop facilities and functions that “professional” developers can use in order to provide compelling services to SME and large companies.
- Cloud computing is appealing for TelCos in this scenario, because while the computing is clearly a commodity whose price will rapidly go down, connectivity seems to have some value in this context (after all if there is not guaranteed connectivity even the most resilient data centers cannot provide services). SO cloud computing is seen as a means to sell or offering to Data Centers Providers guaranteed connectivity at a valuable price.

The possible Network infrastructure of this type of TelCo is represented in **Figure 112**.

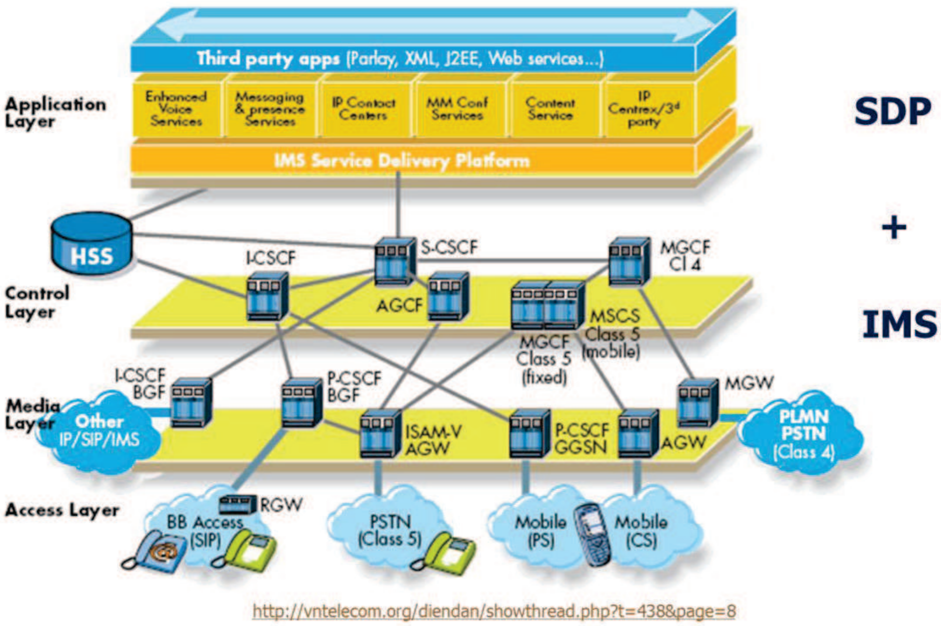


Figure 112: Target Architecture for a Traditional TelCo Scenario

Such a TelCo will try to use the connectivity capabilities and its service infrastructure for competing with Web Companies offering communication related services. In addition this TelCo will offer to its customer connectivity under very stringent Service Level Agreement (SLAs). This will possible give the "controlled" connectivity a higher value respect to the best effort connectivity.

The assumptions behind this approach are that the QoS based connectivity will provide to customers a much better experience in terms of services and capabilities compared to best effort connectivity. In addition, Service providers and Users should find the ratio between "service experience" and "connectivity price" viable and convenient with respect to the price of best effort connectivity.

It is doubtful that the final customer will be willing to pay for QoS based connectivity. This approach is mainly devoted to cope with the issue related to Web Companies. Regional or national Web Companies (i.e., Web Companies that are providing services within the same regional market of the TelCo) are already paying for "guaranteed" connectivity. Offering them a means to dynamically control the QoS of connectivity could even have negative effects on TelCos. Providing QoS control requires a more complex infrastructure and resources, those resources have to be programmed by the Web Company as well as by the TelCo. These effect will bring to an increase on CAPEX and especially OPEX. Even with the assumption that the infrastructure investments on QoS are negligible (IMS and SDP are or will be in place and they are aiming at providing control and services QoS related), the costs associated to dealing with this feature will increase in terms of new processes and programming. These costs will be sustained by the TelCo and by the Service Provider. In the regional case, the introduction of QoS means should be well balanced with respect to the increase of complexity in the TelCo and the Provider infrastructure.

On a global scale and contrarily to what expected, the QoS issue has not a deep impact on the peering system. An example can help in identify a few issues. Currently YouTube is paying to its "regional" provider up to 1 million $ per day in connectivity. The problem is that due to peering other global providers transport the traffic generated by YouTube essentially for free in their networks. The problem is also related to a new regulation of peering rather than to a mere issue in connectivity payment by the single customer.

Assuming that peering will be totally reshaped, some hypotheses can be formulated:

- A revenue sharing between connectivity providers, i.e., a regional carrier will share the revenue generated by a global Web Companies with other carriers that distribute that traffic locally;
- Other carriers will block the traffic generated by the Global Web Company. The Web Company will have an interest to move its content towards the most important markets by stipulating agreements with regional TelCos.

These hypotheses are not strongly related to any QoS issue, actually even considering a best effort network, the same hypotheses could be formulated. With respect to the second bullet, this does not seem to be much different from what is happening already: many Web Companies are asking for caching space to providers like Akamai or directly to TelCos.

The QoS issue could have a global impact if the Web Company is asking to its regional Provider to guarantee certain QoS parameters on a worldwide scale. In this case, the TelCo should sign agreements with many Providers in order to support the expected levels of QoS. This could have a bigger impact on the peering system. However also in this case there could be issues in the provision of end to end QoS.

In addition the revenues generated by services will probably not by enough to justify investments on such an obsolete infrastructure strongly based on a unique paradigm. Successful services are few (as the Intelligent Network experience has shown) and they will provide the major revenue of the platform.

8. 6. 4. Tailoring of the insights

A medium Operator under these conditions will strive to keep the revenue generated by traditional services. It could also attempt to leverage its infrastructure (IMS and SDP) in order to enter into new markets, but the competition (especially in developed Countries) is so intense and probably based on more modern technologies that the success of this move is very questionable. Due to its international market and the presence in less developed countries, the Operator can try to leverage the traditional platform in these markets. However the rapid growth of smartphones also in these areas requires a lot of infrastructural investments as well as the need to offering new services that the platform cannot fully support. The traditional platform, in spite of its increasing obsolescence can hold valid for a while giving the possibility to play the Service Provider role especially if in the original area this Operator has already developed services.

A large and international Operator can leverage the international infrastructure to provide updated services in many different countries. However these services will lack regionalization and specialization to the specific needs of certain geographical areas. Services in less developed country will be provided basing them on the consolidated experience gained in more advanced markets. However in large parts of the market the most innovative customers and enterprises will move towards other offering provided with more appropriated infrastructure and by more focused solutions. The typical customer of this Operator could become the average skilled users, while younger clients and technology savvy ones will move to other more advanced Providers.

Small regional Operators will have hard times in order to operate a cumbersome platform that will not generate enough revenue to justify itself. Federation with larger providers of specialization on a few services will result from this situation. Preferred Federation partners will be medium-large Operators because they will not be perceived as too much aggressive towards the smaller partners. In addition the technology platforms can be complementary.

8. 7. TelCo as a Bit Carrier: A Commodity Scenario

Some TelCos could see “pure” connectivity as a viable target business. They can focus on means and infrastructure to bring the best possible connectivity to their customers disregarding value added services. The goal is to have a very streamlined network that is optimized for transporting bits at the highest speed and the lowest time from source to destination. This type of TelCo will try to optimize its network and forward large bulks of data in the most efficient way (quicker and shorter, i.e., these TelCos will try to deliver the data minimizing the consumption of network resources). Only a few services (very instrumental to the delivery of data) will be considered in the network. Likely the transport infrastructure will be optical. This is because remaining at lower layers (optical layers L1 and L2) will be less expensive than dealing with L3 and up. **Figure 113** represents the major factors that have an impact on this scenario.

8. 7. 1. Favorable Trends

Some technology and market trends make this scenario somehow appealing. The network infrastructure will be streamlined and many 0-touch functionalities could be introduced in the network making the TelCo infrastructure more agile. In addition the evolution of optical technologies makes faster and cheaper to transport large bulks of data. This TelCo will try to exploit this feature remaining as much as possible to the lower layers of the OSI stack (physical and L2) and just when needed for interconnection it will move the layer 3 functions and equipment. Resources at this level (L3 and up) are more expensive either in terms of power consumption than in terms of management and hence a Bit Carrier will try to optimize their usage adopting the principle of transporting the data as much as possible at the L1-L2 layers. Because of the increasing demand for fixed and mobile bandwidth, such a TelCo will try to provide a flexible infrastructure in the highly densely populated area and they will try to combine the deployment of the fiber with the deployment of micro-cells in order to increase the bandwidth offered to customers. The increased capabilities of terminals and the new Bandwidth hungry services will be exploited by this TelCo type to make available the requested bandwidth. Technologies (from optical ones up to SDN) will help to address the Lean Operator approach. Investments in services will be minimal (just those services that are essential for providing connectivity). The possible quest for Quality of Service / Quality of Experience can be met by providing managed services by means of a slim control infrastructure and by the replications of transport capabilities (that will cost less and less). Any increase in traffic will be beneficial for this kind of Operators.

8. 7. 1. Critical Trends

The business perspective of this kind of TelCos is reduced to the Bit Pipe, so increase in traffic is important by the trends towards a less and less favorable price for Mbit/s transported over the network will bring this TelCo to be essentially commoditized. For this reason, the network architecture has to be very flat and very cheap. Sometimes the Direct Communication between terminals could even be a major drawback of this TelCo. The evolution of the computing and communications capabilities at the edge is a major risk and could spoil a part of the business in the long run. The increase of traffic due to new services, e.g., M2M, is not enough until these new services are not going multimedia. For these reasons, this kind of scenario will drive the TelCo to a commodity role, big Operators will have to squeeze down in terms of revenues, people and infrastructure. But is not all, the Vendor ecosystem will be impacted as well. Actually the outsourcing of the "network" could be interpreted a s a signal in this direction: the network ownership does not make anymore the difference, and the infrastructure could be managed directly by a provider. Networks will not differentiate on services, but just on prices.

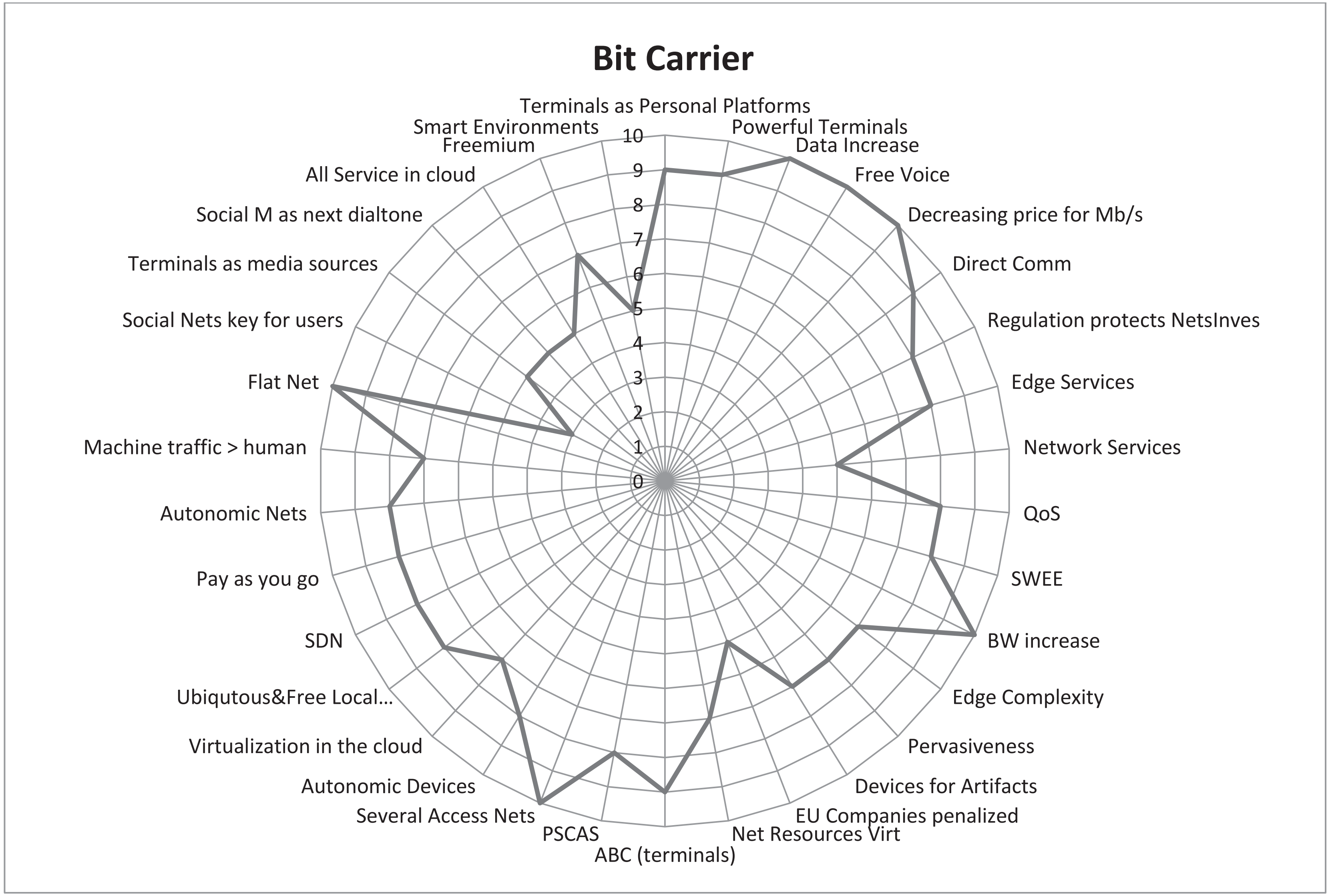


Figure 113: Some Trends Impacting the BitCarrier Scenario

8. 7. 2. The Envisaged Architecture for a BitCarrier

The architecture for such a TelCo has to be simple and efficient, it should be optimized to deliver large bulks of data out of the infrastructure as quickly as possible. Lean and 0-touch solutions will be privileged in order to reduce the Opex. Each investment in new technologies will be carefully scrutinized and new technologies will be delivered only in those areas that allow a fast return in investments. Long term plans will be sustained only to keep the footprint on the market.This TelCo will have a network similar to the one showed in **Figure 114**.

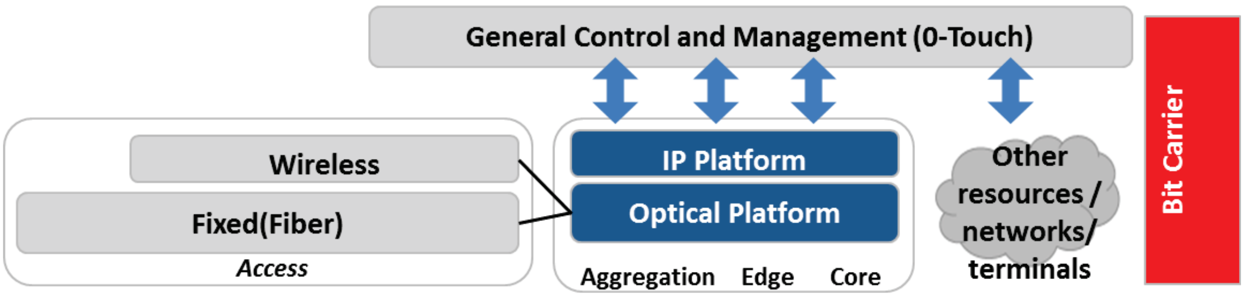


Figure 114: BitCarrier Architecture

The most important part of the infrastructure will be the access that will remain the asset of this TelCo as far as local communication technologies will take over. Investments will concentrate on this part of the network. The Optical platform is the second architectural block of value: keeping the traffic at this level will maintain the architecture cost competitive, the general control and management will be highly automated and functions at this level will be focused on providing managed connectivity services. Example of functions at this level are:

- Subscriber management, accounting and billing in order to determine how a user access to the network, what resources has used and what the billing is compared to one of the consolidated business models (subscription or pay per use).
- Identity, AAA. These function are needed in order to properly manage how the user accesses the infrastructure ad in order to provide basic services like private networks.
- Quality of service functions could be needed in order to provide added value functionalities to the enterprises that need to have an assured level of service.
- Addressing could also be used in order to understand how to better route data flows in order to minimize the routing and transport resources.

- Mobility is a service that has to be provided in order to allow people to freely move in the covered areas of the network.

These functions are depicted in **Figure 115**.

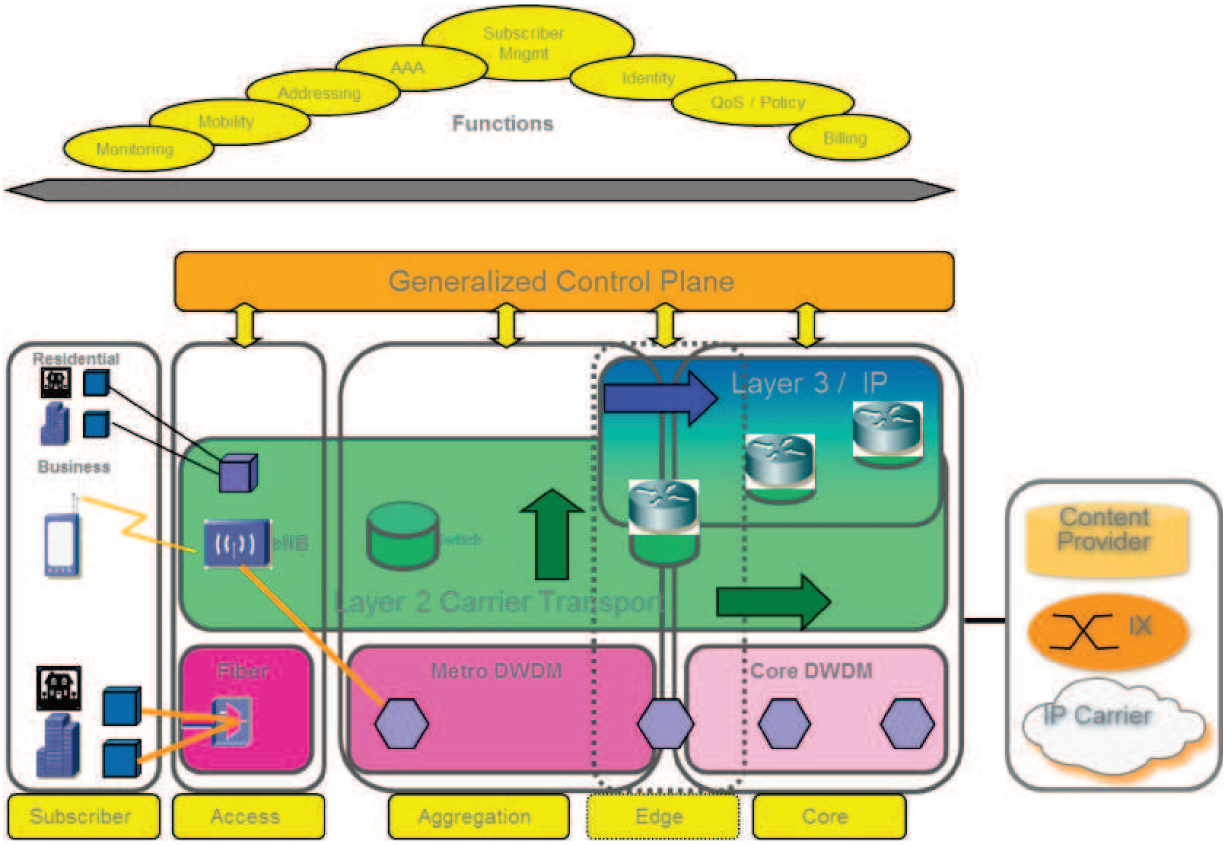


Figure 115: High Level Functions Provided by a BitCarrier.

The sustainability of this business is strongly related to:

- The value that the market will assign to connectivity;
- The CAPEX of the (optical) infrastructure;
- The OPEX of the network.

Approximately, the relation between the revenues generated by this infrastructure and its costs and investments can be represented as f(Revenue)/f(OPEX)*f(CAPEX). It is out of scope for this document to determine the exact formulation of a sustainability function, it is only important to note that if it is lower than 1 than that business is not sustainable. This relation brings in some very interesting consequences:

- Such a TelCo will try to cut down investment to the minimum to sustain the maximum envisaged traffic and will also try to reduce the OPEX (that means personnel and processing for managing the infrastructure) to a minimum;

- Those areas that will not provide a sure return of investment will be kept outside of the infrastructure creating an even more profound digital divide. The focus will be towards dense inhabited areas.

Such a TelCo will have a Lean approach, its goal will always be to cut down investments and costs while bringing connectivity to very demanding areas and markets. This TelCo will strive to keep at a minimum the CAPEX and the OPEX of its infrastructure since it is likely that the value of connectivity will decrease over time. It is foreseeable that Regulator will intervene in order to stabilize the price of the bandwidth in such a way that it will sustain the survivability of the Operator.

8. 7. 3. Tailoring of the insights

A medium-large Operator that decides to be a bit carrier should base its activity on a very large customer base that only large Countries can guarantee. A medium European Country could not support more than one aggressive BitCarrier. This is due to the need of the BitCarrier to have a large customer base that can only be acquired and maintained by aggressive price offering (more that by means of a Quality based service). At an international level the medium –large Operator cannot offer "continental" plans and so it has to operate on a single country base. The foreseen growth in these countries allows this Operator to "gain time" before that also in developing country the saturation of the market is evident. A large Operator with a wide presence on several markets can be well positioned in the medium term for exploiting the growth in developing countries and largely compensate for decreases in established markets. In addition its global footprints can attract many customers because of the easiness in using this Operators services in several markets.

A small regional Operator will strive in order to obtain cash flow (connectivity price is decreasing almost everywhere). It should be very efficient and cost savvy in order to compete in the market. The participation in alliances will not solve the problem.

8. 8. TelCo as a Platform Enabler

Many TelCos have tied over the year to play a role as Service Providers. Many attempts have failed over time, and some TelCos have decided to have a different role in the value chain of networked services, i.e., the role of enabler. The rationale behind this move is to try to give value by enabling other companies to develop services and functions in several and different context. This means to aim at a general purpose infrastructure that could enable the creation of many ICT services by capitalizing the TelCo offering of communications and processing, storage and sensing capabilities. This is a radical move towards the integration of several technologies into an enabling infrastructure, it also imply to make it easy for others to develop services and functionalities on top of this infrastructure.

Playing this role means essentially to open the infrastructure to collaboration with other platforms and systems. There is the need of capability to help the federation/integration of processing, storage and communication capabilities at the edge. This role is strongly based on the possibility to extend and leverage the edge capabilities of large Provider infrastructures (at the east bound border) and those of user communities at the edge (on the west bound border). **Figure 116** depicts some of the impacting trends on the TelCo acting as a Platform Provider.

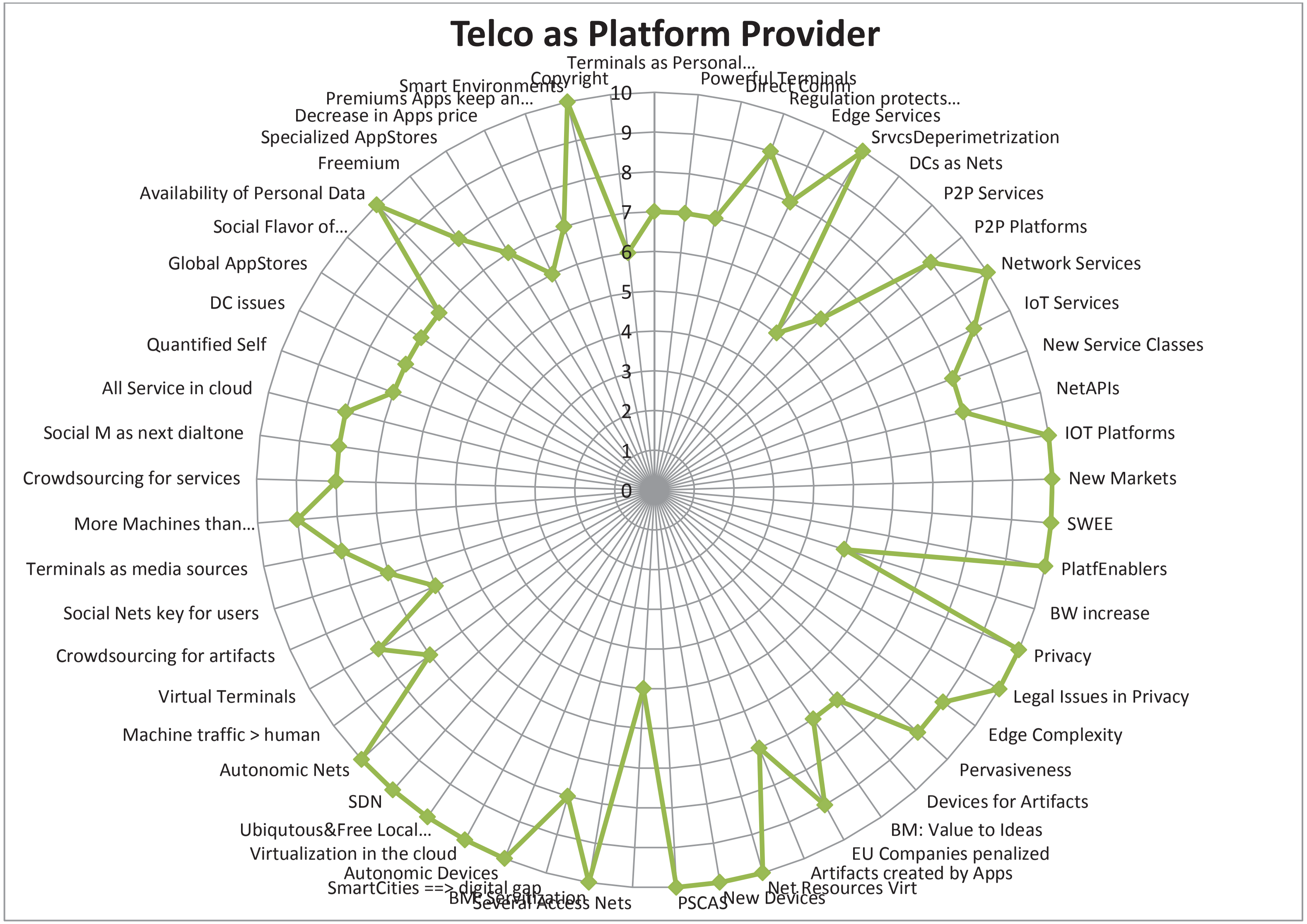


Figure 116: Trends Impacting the Role of TelCo as a Platform Enabler

8. 8. 1. Favorable Trends

The technologies of cloud computing and in general those of distributed processing are generally available (even if some WebCos keep an edge on those). Peer to peer and overlay technologies are not fully exploited by WebCos and they could be used by a Platform enabler in order to create a cooperative platform that extend its reach to terminals and enterprises infrastructures. Virtualization and Software defined Networks are essential for reaching this goal. New (general purpose platforms supporting several interaction paradigms, and not just the network intelligence one) are of paramount importance in order to enable fully distributed platforms encompassing several functionalities at the edges. All the trends emphasizing the edge intelligence can be exploited by the enabling role. Autonomic and cognitive functionalities are extremely important to cope with the complexity of the task and automate the configuring of the platform. The TelCo could have also the role of a "trusted anchor" that enables users to move from one smart environment to the other and also enable to move from a local environment to a more general one. In this sense a TelCo playing as an Enabler can leverage the trust of customers and can create more reliable communities. In addition the Platform provider could try to leverage the personal data by proposing a Brokering model that will enable other actors to access user data and users to defend, protect and "invest" their data according to policies defined by the users themselves. The Platform Provider role fits perfectly with this approach because the applications developed on top of personal data will essentially fall in the domain of a multitude of Service Provider.

Applying a platform wide approach allows to minimize the need to find a killer application. This is a task for the Service Providers, while the goal of the Platform provider is to enable a rich class of new services. These classes of services can address a vertical market segment or can be "horizontal". In any case the Platform Providers aims at the many services that can be designed and deployed. From a developer point of view, the availability of a rich platform could help in minimize the costs for developing a deployable applications. Having many small applications running and generating some revenue could lead to reach a breakeven point in reasonable time even with niche applications.

The major favorable feature of the technological evolution is the possibility of new technologies to allow the deperimeterization of services and networks. In fact the softwarization of the infrastructure can allow the TelCo to provide service out of its traditional borders. This makes the TelCo a potential global player capable of creating an infrastructure on demand. Another important feature that could help the platform proposition is the need for companies producing goods to enter into the servitization business model. Such a platform can be of great help of them by providing the easy integration of processing, storage, communication and sensing.

8. 8. 2. Critical Trends

The role of platform provider is very complex and is based on the possibility of integrate different platforms. This role has to be gained in the field by demonstrating that the data are dealt with in a user centric manner and also the infrastructure is created and supported in a fair (to users) manner. Special attention has to be paid to the programmability of the platform this means to maintain the interfaces and the Network APIs in such a way to guarantee the backward compatibility and to allow a graceful evolution. It also requires attention to programmers and to their ecosystems. Technical issues have to be explained and solved in a very short time. Programmers are the real client of the platform and a lively environment for them has to be created in order to share solutions to specific problem. This ecosystem needs to be nurtured and supported in order to enrich the functionalities of the platform and to augment the number of applications and services that are available.

Complexity is a fierce problem, especially at the edge. So the TelCo should be able to master the software and the issues that raise at the edge of complex systems and this is a role that Operators have never played so far. In addition new services will be initially provided into specialized solutions (siloed) and they porting onto a platform could require some time before proving economically viable.

Two major points of attention are: a) many consolidated players (like Google, Twitter) are already very competitive and appreciated platform providers for specific services; b) the softwarization of networks will lower the capital barrier and many new comers could aggressively address the communication market with new solutions.

8. 8. 3. The Envisaged Architecture for a TelCo Playing the Service Enabler Role

The network infrastructure needed to play the role of Service Enabler has to be radically different from the one of the Carrier. Three major differences must be emphasized: a) the communication focus moves from the lower layers (L1/L2) to higher layer L3 and up (i.e., the network has to deal with IP and application player protocols), b) the network is not only based on communication resources, it has to integrate and offer processing, storage and sensing/actuating capabilities; 3) the focus moves from bits to data and information, the network has to provide means to gather, process and delivery to customers real time information. The networked infrastructure for playing the Service enabler role is represented in **Figure 117**.

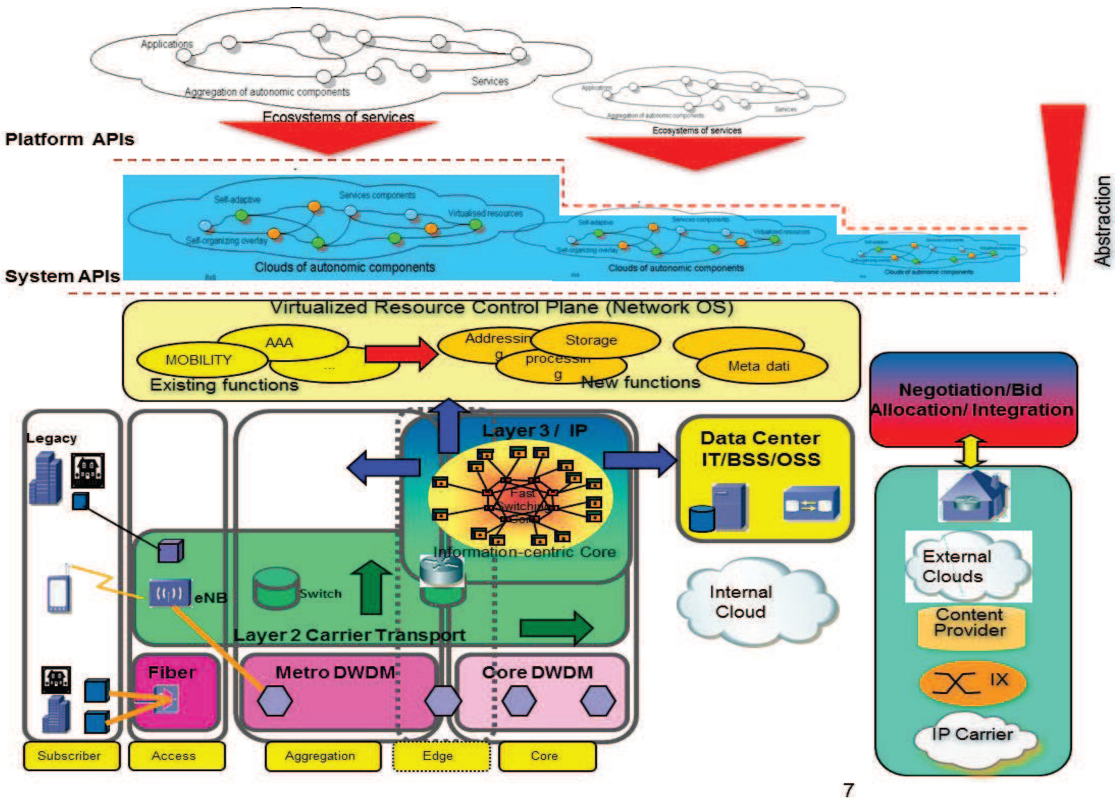


Figure 117: A Possible Network Architecture for the Service Enabler

This platform is characterized by the fact that many of its functionalities and services can be abstracted and virtualized, they can be componentized and offered by means of API to the external world. Internally the resources will span a much larger intent that the simple communication capability: processing, storage and sensing/actuating will be integrated. The network will be open to extension and integration, i.e., external resources could be asked for and included in the entire infrastructure in order to allow the integration of customers network and systems into a common framework. For instance, banks and other financial entities that have a large system infrastructure will not move to the cloud, instead they will request for extensibility and integration (under their direct control) of networked capabilities. Virtualization and abstraction come to help. The Service Enabler will provide a set of virtualized capabilities that will extend the infrastructure of the financial institution.

Also end user capabilities could be integrated and used in order to create complex computing and storage capabilities. For example a Nanodatacenter (Laoutaris, Rodriguez and Massoulie 2008) could be dynamically built by integrating end user terminals into a virtualized data center. User could be remunerated for allowing the usage of their resources for a certain period of time. The integration relation is based on contract that can have a limited in time validity.

The platform should be able to accommodate also for smart objects, this pushes for introducing different communication paradigms that differ from the prevailing one (i.e., the client – server). The Service Enabler network should be able to provide a Complex Event Processing (CEP) engine capable of dispatching in real time events and commands to a multitude of smart objects as well as to derive usable information by analyzing event patterns. This real time event engine will be a sort of twitter of things where each smart thing will be able to send (and receive) events about its perception about the status of an environment, resources or system. Transactions, i.e., a set of events that describe the occurrence of a functionality, will became more and more important. Users will want to be informed of the final status of actions performed on their behalf from financial systems, or security related applications. Users will want to trace all the actions that they, or their terminals, or smart objects temporarily associated to the person are performing on behalf of the individual. These transaction will be able to represent the digital experience of users, and they need to be collected and passed to the users in real-time. A PubSub engine (Tarkoma 2012) seems to be a technical possibility to allow this transaction based digital world.

All these functions will be provided by means of componentized functions that will be instantiated as programmable entities of an overlay network. They will be resilient and self-organizing in order to allow the developers to focus on the functionalities of the service instead than coping with the complexity of the system organization. Each entity will provide specialized APIs in order to offer functions and programmability to the applications. Different levels of APIs will be requested in order to exploit different levels of abstraction that the Service Enabler network can support. The differentiation of API (represented in **Figure 117** as a ladder) are useful for allowing the service designer to choose the level of granularity and control on networked resources that the service needs.

8. 8. 4. Tailoring of the insights

This role for a medium-large Operator requires the strategic decision to move to a different role. The platform is a means in order to become an effective partner of many national and international companies. This role requires the ability to master the software of the platform and for size and skill such TelCo could make it. There is, however, the need to transform this Operator into a technological actor capable of developing compelling software solutions. There is not room in mediun Countries for more than a TelCo acting as a Platform Provider because this role still requires a large market share in order to be sustainable. This role can also lead to the internationalization of the offering and hence to compete with large global companies. The differentiator factors will be innovative solutions and appealing prices.

Large Operators could have a great interest in this role. Being capable of operating a platform can allow to address a large market comprising different types of country. Each market can benefit and could use the platform according to the technological maturity level of that market. However infrastructural costs can be shared among different subsidiaries of the global companies leading to some saving (large platforms cost less in proportion).

In order to succeed, small Regional Operators reside in a developed country and be very aggressive in terms of offering and innovation. Competing with large Operators and very specialized ICT companies will not be easy and the small market not necessarily will back up the Operator in terms of users and revenues.

8. 9. TelCo as a Service Provider

This is a role highly looked for by several TelCos. In this document this role is played by a TelCo that has totally refurbished its control and service layer by getting rid of the obsolete and awkward IMS platform in favor of one infrastructure more aligned with the recent trends in ICT. Roughly speaking a Service Provider could use the Enabler Provider platform and adding on top of that some mechanisms, tools and capabilities for directly develop services. Being a service provider is a difficult role for a TelCo. The first step to take is increasing the capability to master the software in order to quickly create compelling services. A second step is to address services with a disruptive approach either at the technological or the market level. These approaches are very difficult for an established Operator. In addition, as showed by **Figure 118**, there are many trends impacting the capability to play this role.

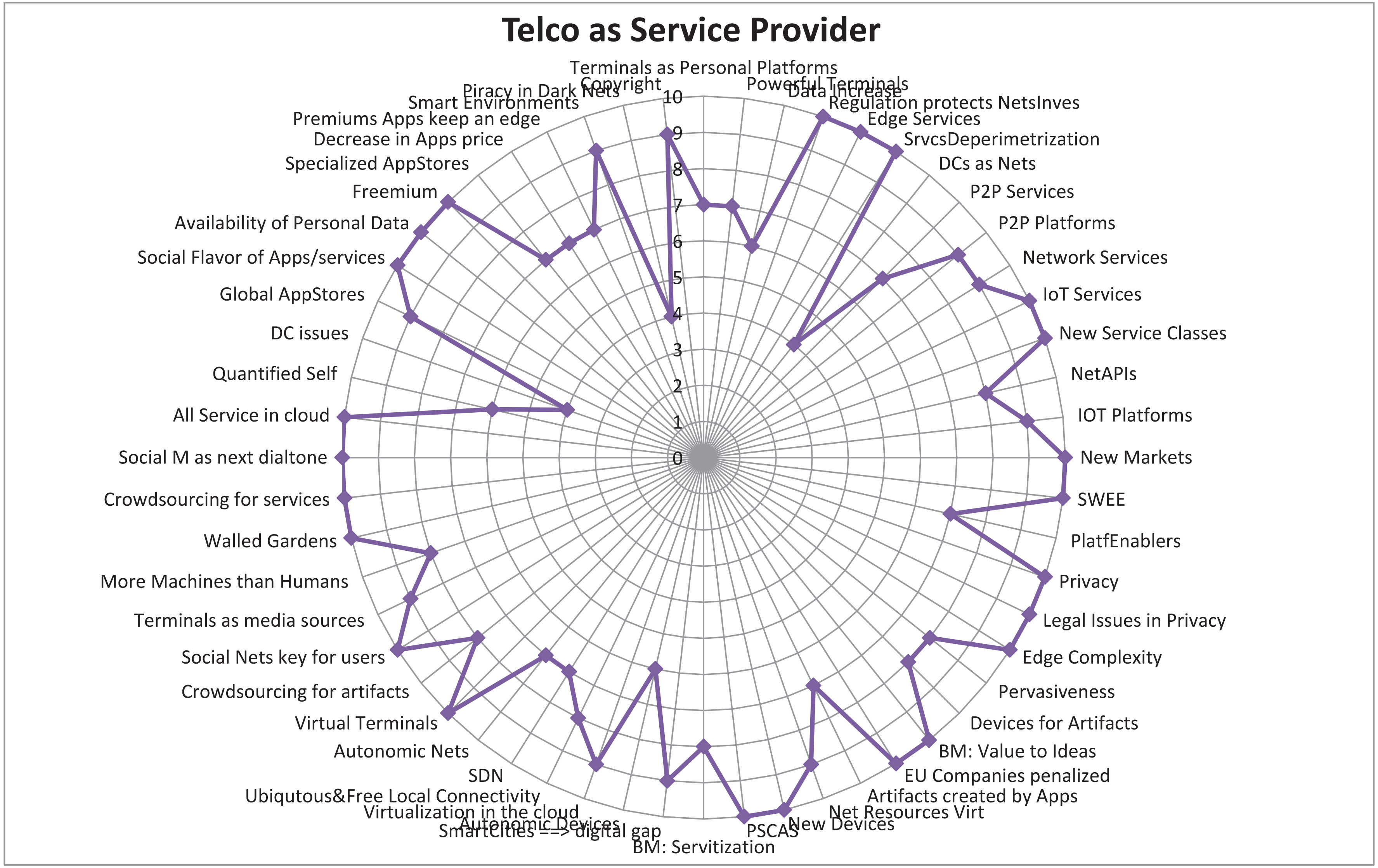


Figure 118: Trends Impacting the Ability to Play the Service Provider Role.

8. 9. 1. Favorable Trends

The softwarization of networks is a strong enabler for quickly moving towards the integration at the service level of processing, storage, communication and sensing functionalities. This will create the basis on top of which very specific services can be build. However the platform level has to adopt catting edge technologies in order to give an tangible advantage to the service provider. This is not necessarily true from a Platform provider that could point to leveraging the stability, availability and the know-how of a well-established platform. Mastering of software has to be established at two different levels, the platform one and the service one. They seem similar, but they diverge in skills. Acting as a platform provider means focusing on functionalities and components that are to be reused, and selling them a services. Acting as a service provider means to focus on the specific requirements of the final customers and not those of developers and prosumers. In the second case the user experience is extremely important and the service provider has to take care also of the interaction and solution on the client side (in the terminals). Many new classes of services are possible and possibly the service provider should correctly position its offering in several vertical markets.

8. 9. 2. Critical Trends

There are many drawback to a Service Provider role for a TelCo. The first one could be the need to justify the investment on the platform and the service development in terms of revenue generated by services. In other terms, the Service Providers need to find from the very beginning a set of killer applications that will help to start the business. It is clear that there are not many killer applications that have not been provided, so this could be a daunting task. Reaching the breakeven point with applications generated by one single Service Provider is difficult and aiming at niche applications requires the knowledge of specific problem domains that a single provider does not have. The competition on services is played more at the user experience level than at the technological level, so existing platforms and service providers using the C - S approach still can be major competitors. Another important factor is that many Service Provider have created walled garden solutions and gaining users from those ecosystems could be very expensive and difficult.

8. 9. 3. The Service Provider Architecture

The Operator playing the role of Service Provider will opt for a different infrastructure with respect to the BitCarrier and the Platform Enabler. Its value is in the creation of applications, so all the platform will be customized at several levels (from the resource level to the upper layers) in such a way to take advantage by special feature and capabilities. The platform will lose generality and will be more and more targeted to the specific markets addressed by the Operator. The service and application creation will be highly standardized and a well-defined registry of application will be introduced. This process of specialization will be further exacerbated by the need to support Application Stores.

Figure 119 represent a possible architecture for a Service Provider

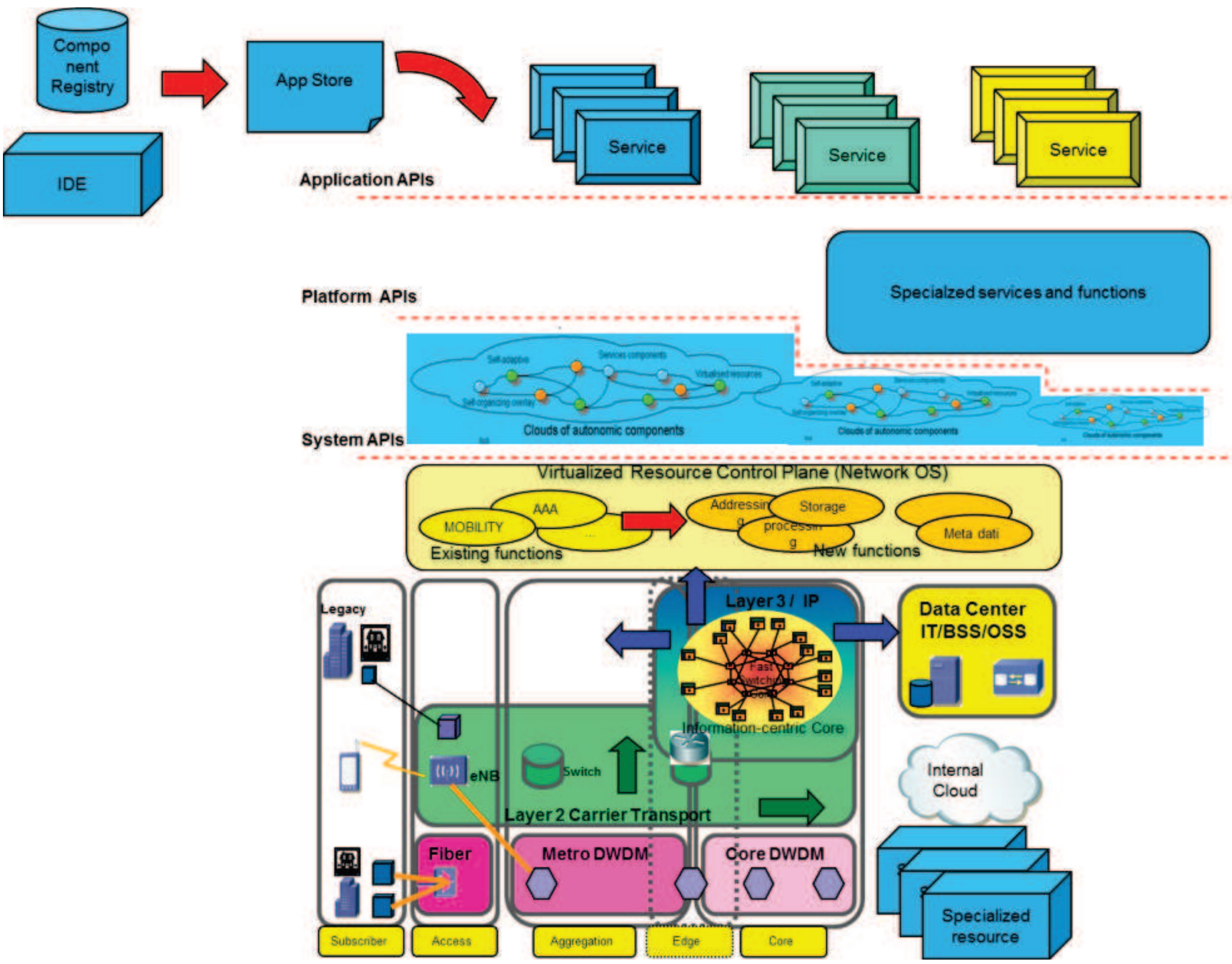


Figure 119: A Network Architecture Supporting the Role of Service Provider.

8. 9. 4. Tailoring of the Insights

A large-medium Operator should put a lot of effort in successfully play this role.There is the need to set up the platform and to populate it with services that produce considerable revenue. As discussed previously, Operators do not have “a long tail” so this effort is daunting. Achieving survivability by service on a few markets is really difficult and prone to a lot of competition from small companies as well as from big players. This role is also technology intensive, because the software skills required are even more stringent than those of platform provider. There is the need to take care of the platform and to build new services. These are two activities that deserve particular attention. This Operator will try to fully exploit the business market and will try to find alliances and deals with other similar Operators or Companies in order to expand the market.

Large Operator will try to play this role in order to exploit its service portfolio in several markets capitalizing the development cost. A large number of clients is needed and the coverage of many countries could be a success factor. The dimension of this TelCo could also be such that it could be possible to have good deals with terminal vendors. In this case the TelCo could try to create walled gardens in order to lower the user churn. However this approach imposes to the Provider to be capable of creating very successful services that many users want to utilize.

8. 10. Disappearing TelCo

This scenario is by far the most difficult to predict, but not necessarily the less plausible. This scenario envisage a situation in which software capabilities and general purpose hardware can be used to create dynamically a communication infrastructure capable of allowing a general communications between end points. If any single object is capable of routing information in the nearby domain, then local networks can grow so big to became regional interconnected networks that support the needs of large communities. For instance if a FON like approach is used in the northern flatland in Italy, then there is the possibility to create a sort of very large mesh network connecting large cities and densely populated areas. In addition cars and vehicles can help in transporting large data masses. Data can travel from car to car and from access point to access point until they reach their final destination. In addition if local providers are willing to cover certain areas with resources (general purpose communication and processing resources), these islands of well interconnected objects could act as a sort of island that mobile terminals can be using in order to extend the capabilities of a specific island and in order to interconnect different islands. Access to an island of interconnection can be granted in exchange for some resources that the new node can make available or for As said, the infrastructural costs for deploying (wireless) network could fall so much that local operators could become a reality, in addition technologies like cognitive radio can be unleashed also from the Authorities and allow the creation of very large and powerful networks. An example of this kind of network could be the mesh networks, but at a much larger scale, DirecNet is an example of how to combine different elements into a powerful, scalable and mobile network[50]. **Figure 120** depicts some major trends impacting on this scenario.

8. 10. 1. Critical Trends

The business perspective of this kind of TelCos is reduced to the Bit Pipe, so increase in traffic is important by the trends towards a less and less favorable price for Mbit/s transported over the network will bring this TelCo to be essentially commoditized. For this reason, the network architecture has to be very flat and very cheap. Sometimes the Direct Communication between terminals could even be a major drawback of this TelCo. The evolution of the computing and communications capabilities at the edge is a major risk and could spoil a part of the business in the long run. The increase of traffic due to new services, e.g., M2M, is not enough until these new services are not going multimedia.

For these reasons, this kind of scenario will drive the TelCo to a commodity role, big Operators will have to squeeze down in terms of revenues, people and infrastructure. But is not all, the Vendor ecosystem will be impacted as well. Actually the outsourcing of the "network" could be interpreted a s a signal in this direction: the network ownership does not make anymore the difference, and the infrastructure could be managed directly by a provider. Networks will not differentiate on services, but just on prices.

This scenario is drawn upon cutting edge technologies that have still to prove their applicability on the large scale, nevertheless they are so promising and potentially capable of transforming the "concept" of network that they deserve a particular attention.

[50] Actually DirecNet is a mobile network in two senses: it supports wireless connectivity between its nodes and it "moves" with its nodes (airplanes, tanks, military vehicles, warships and soldiers).

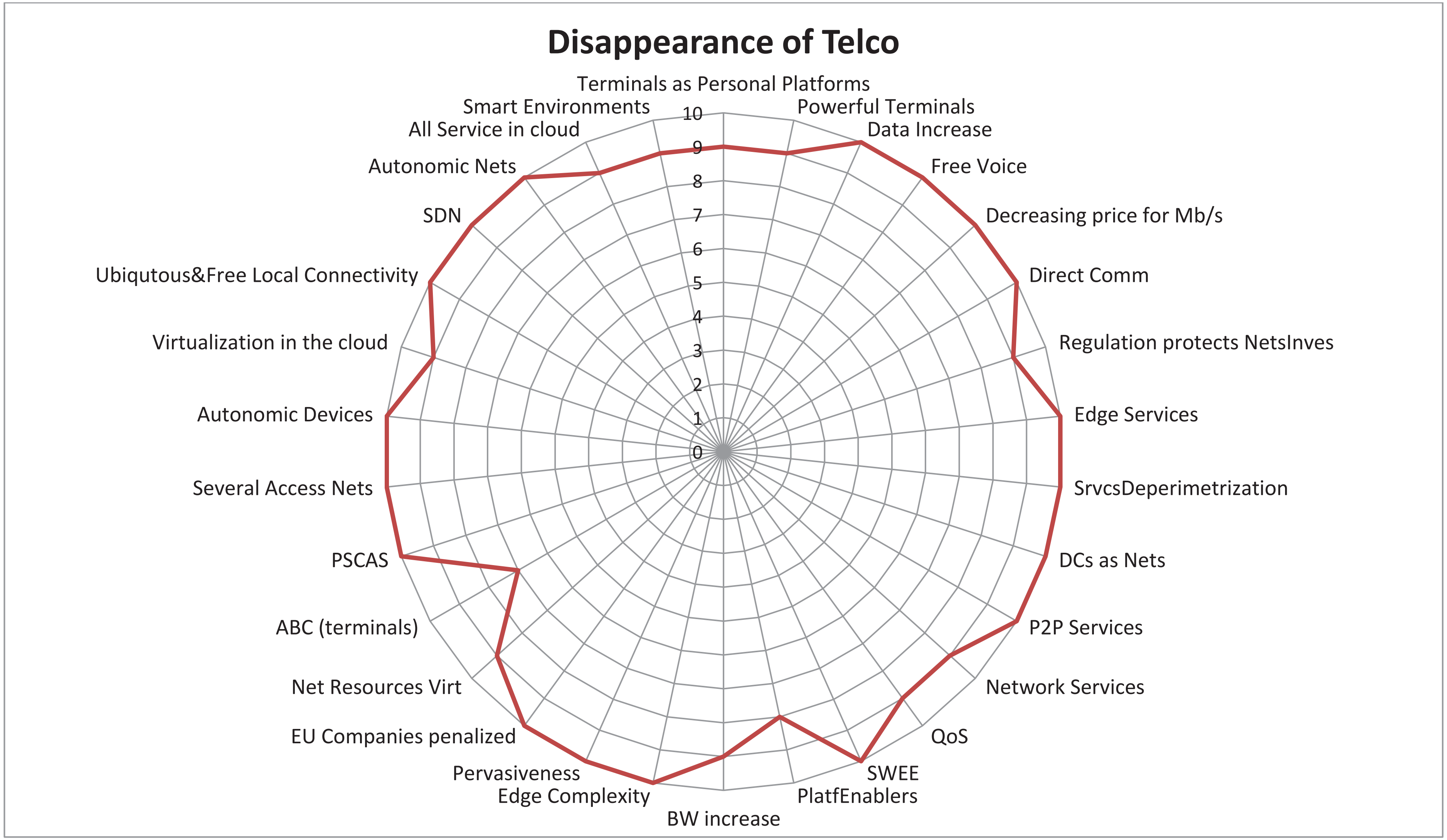


Figure 120: Major Trends for the "Disappearing of the TelCo" Scenario

8. 10. 2. Favorable Trends

Autonomics networking, virtualization and software defined networking together with overlay and mesh technologies are some of the ingredient of this new kind of network. If cognitive radio is added then the technological disruption is such to be possible to reshape the concept of "network". A network is a set of cooperating processing, storage, communications and sensing resource capable to adapt to the specific needs of their customers. The software will be the enabler of this approach, so any actor willing to play this game should be capable of mastering the software intricacies of software development. In addition these infrastructure made out of many unreliable nodes are "complex system" and they follows the dynamic of these systems in particular respect to the possibility of "a prompt change of phase" that can bring the system in a totally different state that let emerge a very intelligent behavior. This typically occurs when the number of objects is very high. Companies able to dominate this complexity will lead the ICT developments for many years becoming change player by introducing a lot of disruption in the status quo. The technological trends are yielding toward this direction. This approach is so disruptive that also the business models will change in favor of more participative ones, and the disappearing of large infrastructure providers that "own" the users. In fact the users themselves will become resource and infrastructure providers. For instance **Figure 121** represents the role of the user in a similar ecosystem.

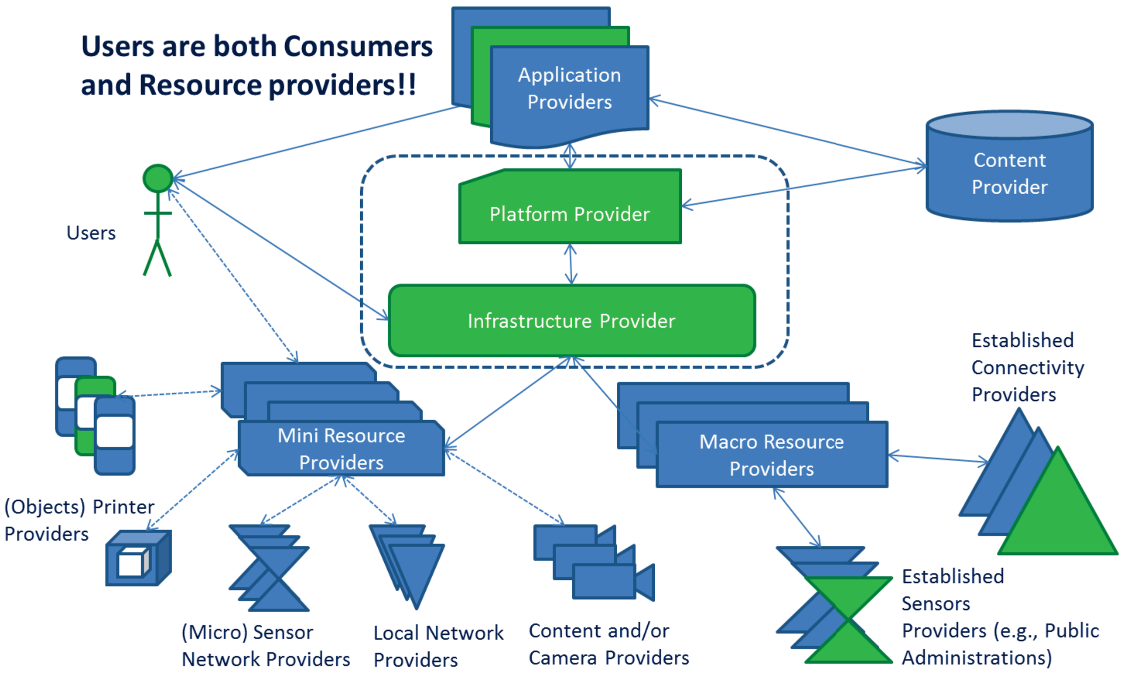


Figure 121: The Ecosystem Envisaged for the Service Enabler Platform.

In this case, customers will be resource providers and some of them can even provide a considerable large wealth of resources.

8. 10. 3. Critical Trends

The major issue concerning this scenario derive from the disruptive changes introduced into the current ICT ecosystem. The vanishing of traditional player like the Operators in favor of local communities putting together a viable infrastructure is so full of innovation that the traditional Actors will violently react to this occurrence. Regulation can play a significant role in one sense or the other (favoring or blocking the wave of change). Even in this case, the technologies will find a large usage at the local level and the local communication will greatly improve. Many services will become local because it were they pertain (e.g., file exchange, interaction within smart environments, and the like) impacting anyhow with the current approach.

8. 10. 4. The Disappearing TelCo Architecture

The disappearing scenario is based on a few basic assumptions:

- the move of intelligence (i.e., added value functions) from the network to the edge of it;
- the paramount importance of the terminals that will be the leading factor for providing intelligent functions and services to users (that will determine and in a sense dictate how the infrastructure should accommodate terminal requests);
- the software advantage (over hardware – i.e., due to technology progress many solutions that required specialized hardware and a lot of performance tuning can now be executed on general purpose computers by means of very sophisticated software).

The last point introduces a new concept: "software will eat everything" (Andreessen 2011), i.e., the possibility that many objects of common daily usage (from computers to vehicles, from the fridge to a manufacturing machinelike a robot) will be controlled and functionally extended by software capabilities. But this is not all, many of these software capabilities will be available by means of open source solutions and/or crowdsourced by active and dedicated communities. As an example, the DIY movement is growing under the idea that "printers of objects" can be programmed and they could even "print" a large part of their components (a sort of auto-reproduction model). This trend can be quite disruptive for consolidated ways and processes of manufacturing and be a technology enabling a turning point for a whole industry. These assumptions have led to a scenario in which terminals and applications running on them can negotiate and allocate the available resources in order to fulfill their communication, processing and storage needs. This is not all, in fact terminals could also drive the optimal (from their point of view) allocation of resources, and using the interchangeability and mutual support of storage, processing and communication in order to achieve their goals. An example of today is Amazon's Silk, i.e., the capability to determine the context in which a terminal is operating and consequently allocate the best configuration of resources in the terminal, in the cloud and in the network in order to accomplish the terminal task. In this scenario terminals, communication, processing, storage resources as well as actuators and sensors can dynamically cooperate at the edge of public networks in order to create the best environment for terminals and application execution. Software adaptation capabilities (ranging from autonomic processing and communication to cognitive radio techniques and software defined radio) can be used in order to support highly dynamic edge " platforms" . Two current trends could lead to this scenario: virtualization of resources and the so-called software defined network. Virtualization means that meaningful functionalities can be virtualized almost everywhere in the network and migrate accordingly to the roaming of terminals/users. Software defined networking is freeing the networking platform from technological monopolies and walled gardens. Now routers and other networking elements could be mainly implemented on software and be executed on general purpose hardware. These two characteristics are the basis for a highly distributed and always available communication and execution platform very different from current ones. New actors will emerge that will be using software programmability and flexibility as added value over communication and processing prices (that will be provided as commodity). In such a situation any processing or storage capable machine could become a pawn in the puzzle of creating "personal networks" adapting to the goals of the user and his terminals. Software capabilities will became a competitive advantage as well as the capability to aggregate on demand a broad set of resources needed to sustain the platform. New business models (involving also the final users) will emerge. An example could be FON in which a few users make available part of their resources to other to use. This scenario can be seen as a two steps process: in the first step, small and community based networks at the edge will emerge in order to satisfy the local communications and processing needs of a dynamic community (**Figure 122.a**). In the second step the interconnected islands will be so powerful and capable to create a very powerful network capable of integrating different control mechanisms acting on a cloud of mobile entities (**Figure 122.b**).

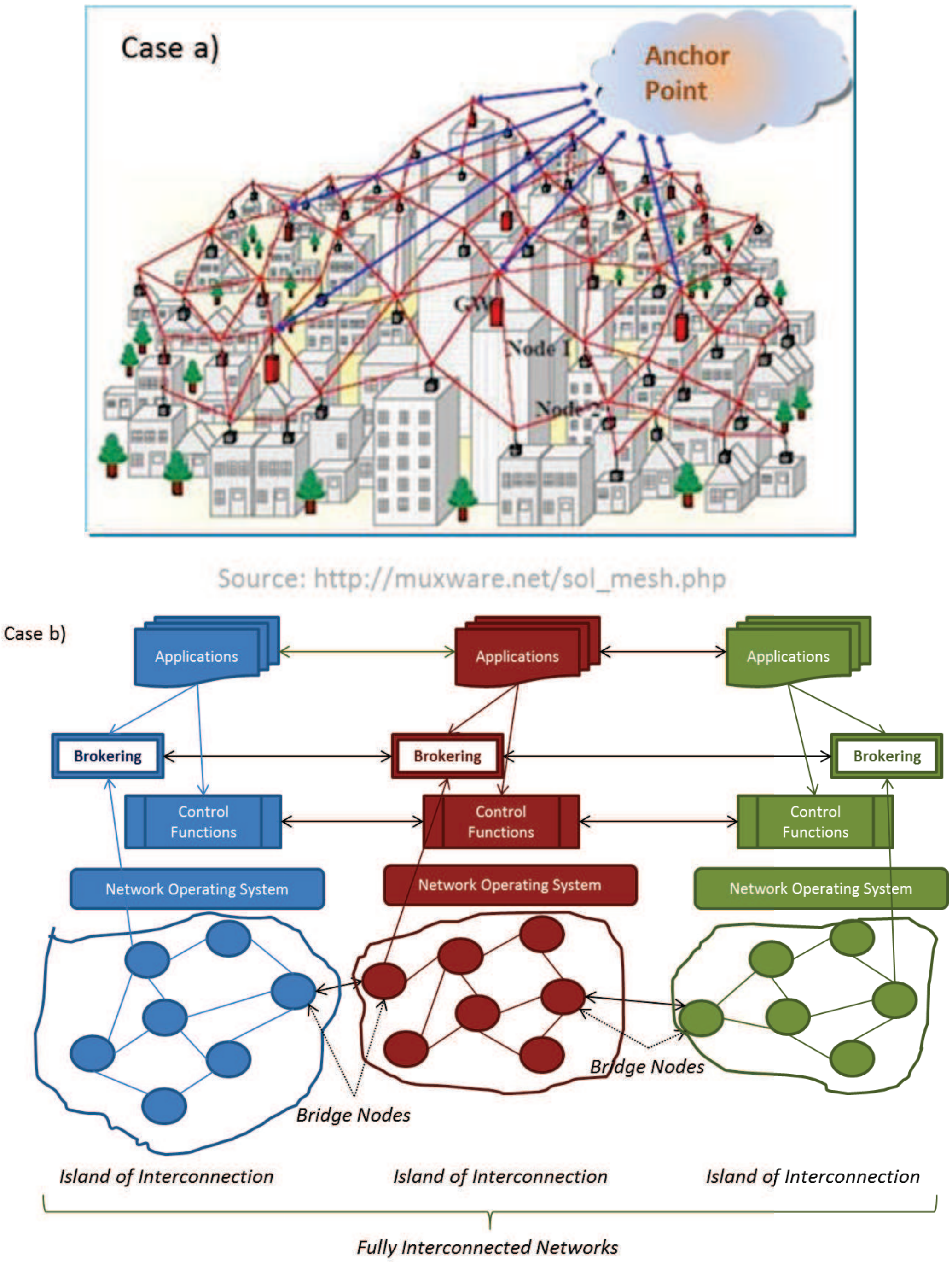


Figure 122: Towards an Architecture Without a TelCo

This kind of architecture is based on the concept of sharing resources, actually a Brokering role is essential for gathering and putting together all the resources made available. In addition the broker function will be extremely useful for negotiating the resources and the prices related to them. The broker function is needed at the early stage, when resources are just local and made available by a relatively large group of users, and in the second step, when many resources are available and there is the need to negotiate resources between different islands. Another important feature of these architectures is the virtualization: resources can be virtualized so that abstract specific networks can be created on top of the available and dynamic infrastructure. These specific networks can be specialized for services (e.g., providing a virtual CDN for a Content Provider). Actually the large number of participating resources will make (as in the P2P networks) the entire infrastructure more stable.

8. 10. 5. Tailoring of the Insight

In this scenario, the Operators are very small in terms of people and these people are mainly developers. This is exactly the scenario that all the Operators are trying to avoid. More than an Operator this type of solution is operated by a technology company that is capable of programming large distributed systems spanning from a (reduced) set of network resources to a large number of terminals. The deperimeterization of services and the possibility to arrange networks by aggregating genral purpose resources should offer the possibility to leverage the deperimeterization of services. In other terms this Operator should aim at becoming a global provider of services exploiting the fact that a large general purpose infrastructure is already deployed and usable by many companies.

8. 11. Findings from Scenario Analysis

In the previous sections five scenarios have been described and some of the enabling and critical points have been emphasized. The last step is evaluatng the likelihood of the scenarios and to prepare for the future. The **Traditional TelCo** scenario will be instrumental to slow down the decline of the Operators business trying to leverage the concept of network services and the related one of Quality of Service. It will not help in solving the major issues of the Operators in the service field: the regionalization and perimeterization of services. This approach is strongly coupled with a traditional network infrastructure and does help in introducing the requested software agility. The development of new classes of services on the IMS and SDP architecture (even if possible) will be so cumbersome and heavy compared to other solutions that will introduce further delay in the service provision (see for instance the case of the WAC or the next failure of the Rich Communication initiative). The decision to stay on this infrastructure will increase the technological gap between the Web Companies and TelCo in favor of the former. The traditional scenario has a high probability to be implemented by many conservative Operators. In the short/medium term it will be instrumental to preserve the current business and the related “value added communication services”. It will not contribute with new revenues and it will not be successful for allowing the Operator in entering into new business. The goal of using SDP and IMS for instance for multimedia services or for M2M applications will lead to failure, because the other competitors will be able to use more advanced platforms and related interaction paradigms. For instance the twitter platform is already more appealing than IMS for developing IoT related applications because of the capability to use a PubSub model. Operators trying to adopt this approach to services and platform are simply trying to postpone the time when they will totally out of the services and applications markets. In other terms, the Traditional TelCo approach will lead to the **BitCarrier scenario**. This one could be appealing for regional Operators or for new entrants, they have new network infrastructures that do require less Opex and are operated by less people. For big TelCo, this approach will lead to a very considerable shift down in terms of business and organization that probably some companies will fail. However the most likely perspective for a large number of Operators is essentially the one to become Bit Carrier. Their power will be preserved essentially by the Regulation because of the huge acquisition in the past of portion of spectrum. Authorities will try to safeguard these investments in order to not completely disrupt the current market. It is also likely that in the context of a unique European Telecommunications Market, some companies will merge in order to play the transnational role of Bit Carrier. The technology evolution will help to keep infrastructural cost down, but on the other side, the revenue generated by connectivity will squeeze even further reducing the current margins and leading to a commoditization of the communications. The **TelCo as a Service Provider Scenario** and **TelCo as a Platform Enabler scenario** seem to be complementary in the sense that a Service Provider needs to have a Platform and hence a TelCo could play both roles. This approach is not true, being a platform provider is very different from being a service provider. Skills and processes have to be very different and also the market approach has to be different: as said a service provider will compete on services and it is important for it to figure out as soon as possible a winning and differentiating service portfolio. The platform provider instead has to create the conditions to have as many as possible Service Providers using the platform . In a sense the Service Provider needs to be on the Top of the value chain, while the Platform Providers will earn if the (many) others will earn also from niche services. The platform itself is very different. On the Service Provider side, the architecture is such that tools and mechanisms to cut down the design and development time are prominent, while for the Platform provider the possibility to help its customers in the entire life cycle of services is most important. The goal of the Platform provider (and its related mechanisms) is to populate the platform of service and to attract many developers. The focus is on the ecosystem and not on a few winning services. Openness of the platform, accessibility, well-formed and stable interfaces are a priority for the Platform Provider. Actually its goal is to promote API at the north, east and west bounds. The goal of the Service provider is to rationalize and make effective the usage of the north bound interfaces (those towards the systems of the Service Provider). IN a way the Service Provider scenario could be degrade (if the technological choices are too much traditional) towards the traditional TelCo Scenario and then to the BitCarrier one. **So the most interesting scenario for an innovative TelCo is the one of becoming a Platform Provider** (an enabler) and using the coming technologies to deperimeterize the infrastructure and the services. A broader view of the architecture for a Platform Provider is given in **Figure 123**.

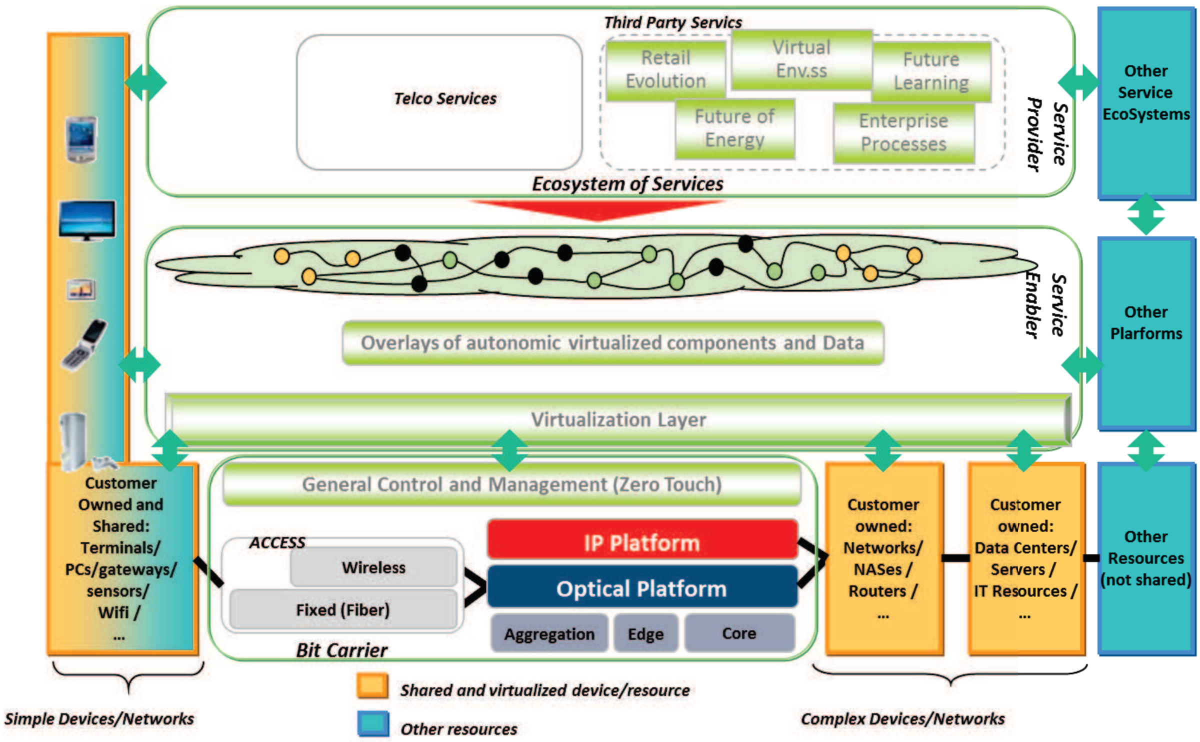


Figure 123: A Service Enabler Platform

This platform could be considered as a sort of blueprint that an innovative TelCo can progressively implement, deploy and improve according to the technical and market evolution. The need to integrate the TelCo platform with external ones is of paramount importance, because it is the key to exceed the offering of the web companies. The integration and openness is a key differentiator as well as the ability to integrate several networks in such a way to "de-perimeterize" the communications and processing infrastructure. Virtualization is another key enabler, different classes of resources can be virtualized and offered as a infrastructural service to developers and Service Providers. Another possibility (more challenging) is to try to sell to Service providers functionalities instead than virtualized resources. This step is a challenge, but the envisaged architecture can dependably support this market goal. The virtualization level sits on top of the transport capabilities and try to extend the basic functionalities of the underlying infrastructure and to make them programmable. On top of this virtualized level, the virtualized resources will be gathered, aggregated and organized into overlay networks in order to create the logical infrastructure needed to operate the service. Components at this level need to be autonomics and supporting self-organizing capabilities. This will limit the programmers burden in dealing with the intricacies of the networked environments. The overlay networks and their resources will offer open interfaces to the upper layer. In this way services could be create on top of the overlay infrastructure supporting the service. It is important to offer the capability to organize overlay according to different interaction paradigms like client – server, peer to peer, PubSub (or blackboards, tuple space and the like) in order to accommodate the different styles of programming of the developers. This platform can be used in a more appropriate way than the IMS one to develop new and current classes of services. The communication paradigm (i.e., the network intelligence one) is not forced on the users (the developers) of this platform. Actually the platform can accommodate the different interaction paradigms giving a high level of freedom to the programmers. Service can be developed using the platform functionalities. If the Platform provider perceives an opportunity, it is able to create its own portfolio of services, but this should be carefully considered because service providers do not want to have a platform provider that become a direct competitor and in addition the owner of the platform that they are using. So the Platform provider could use new applications in an instrumental way: i.e., for opening new market initiatives and to create an initial ecosystem around a new class of services or a vertical market. The last scenario, the **Disappearing of the TelCo scenario**, could be of interest for the TelCos that want to really entering into new markets. This approach could be undertaken by spin-off of the TelCos addressing new market initiatives. They could act disruptively on purpose, in order to gain a different market position in nascent markets and business opportunities. These initiatives need to be created with a long term perspective in order to create a new market and to help it to grow. At the end of this path the company that undertook it will be totally transformed and it will difficult to consider it as a traditional TelCo. This path is a big challenge, but the enabling technologies are all available even if at an initial stage.

Chapter 9. Research Contribution, Future Work and Conclusions

This thesis has been carried out during a long period of years considering the evolution of technologies and the changes in strategies and approaches to the competition of major (European) Operators. This analysis has brought to identify a number of scenarios (discussed in 8. 5. and following sections) that represent possible strategies for Operators. The scenarios emphasize the assets and strategic role that the TelCo can leverage in order to achieve the business goal. For each scenario an enabling platform that exploits the assets at hands has been presented in terms of principles and guidelines (an high level architecture) that the TelCos should follow in order to build a platform enabling them to play the chosen different role. Contrarily to what has been proposed in (El-Darwiche, et al. 2010), the roles chosen by the Operator and to a certain extend the business models should not be mixed up in order to focus on the platform and the supported functions. An Operator should characterize itself as the best in class in one of the role in order to obtain a recognition from the market, while an ambiguous proposition in the market could lessen the impact of the strategy.

The scenarios of Chapter 8. could be further personalized and adapted to the current situation in which the specific TelCo is operating. In certain areas of the world, the traditional TelCo Scenario could still be valid and rewarding because the market, especially the mobile one, is on the rise and because Web services are not yet available to a vast part of the population (Kimbler and Taylor 2012). The TelCo as a Service Provider scenario can be appealing in particular countries like Korea and Japan in which the role of the Operators in services has been consolidated over the years. In other markets, the Bit Carrier scenario could be the most likely one, due to the difficulty in positioning the Operator in the service real. A paraphrased quote of K. Marx, “A specter is haunting Europe — the specter of disappearing TelCos” could describe a the possibility to a radical transformation, a revolution, of the consolidated TelCo business. Each scenario, though, should be considered as a starting point and it should be specialized and detailed according to the specific objectives of the analysis and the uniqueness of the specific TelCo. Some general considerations with respect to the scenarios and the issues presented in this document are presented and discussed in 9. 3.

The most appealing scenario from a technical and business perspective is the one of the TelCo as a Platform Provider. This could be a proposition that will not be particularly evident to final customers, but it could be particularly rewarding to a TelCo because of the central role. The platform can leverage both local and global presence of the Operator and the whole infrastructure could be used to deperimetrize the services. Drawn on considerations of section 8. 8. , the envisaged architecture is further discussed in sections 9. 1. and 9. 2.

Finally , the future work is presented and discussed in 9. 4.

9. 1. A Viable Architecture

One first requirement for modern service architectures is the need to have a high expressive power and being capable of supporting powerful control paradigms. Under this respect one of the major outcome of the analysis is that highly distributed platform such as P2P networks are capable of offering a good combination of expressive power, the possibility to optimize the networked resources usage and are simple enough to program. These networks (and in general the smart environments) will be constituted by a large number of nodes. There is the constraint of making each individual node capable to adapt to the current context and to cooperate with other for achieving a global goal (as well as reaching the individual goal of optimization of usage). In addition, networks have to exceed the limitation imposed by a strict perimeterization of resources in a single administrative domain. Actually resources are to be used whenever they are available and they economically and technically fit with the goals of users within a smart environment. In this sense it is important to see the smart environments as a sort of Network of Networks, i.e. a dynamic association of different types of physical resources made available by cooperative networks.

In an open implementation of resource aware network of networks, the real winner is the final user that will have at least two benefits: on one side an optimized and fair access and usage of networked resources without the obligation to get the service from a specific provider; and on the other side, the users themselves could contribute to the empowerment of the network of networks by providing resources (i.e., users could become temporary resource providers). This is a major shift in role with respect to current rigid business models.

In order to fulfill the requirement of extreme flexibility, programmability and self-organization put forwards in the previous sections, the need for a novel networked platform for NGNs and Future Internet is advocated. It aims at the optimization of resource usage, the selection of the best combination of control pattern for each service, and the integration, in a seamless way, of Operators' resources with those of Users and external Providers. It is made out of layered overlays comprising virtualized resources, enhanced with autonomic [XXI] and cognitive [VIII] capabilities to achieve self-awareness and self-organization. Generally speaking there is a strong requirement to decouple the network functions (natively supporting mobility, resource awareness and the other requirements) from the service control. However this sharp separation should not lead (as it is today) to two separate and incommunicable worlds: services will better perform if they are able to adapt to the available network capabilities. In other terms, a clear interface between the network layer and the service layer has to be defined even in the Future Internet.

The proposed architecture relies on three layers (**Figure 124**):

• the lower layer, the physical layer, comprises pervasive heterogeneous resources spanning from network nodes and servers, to Users' devices and "smart things" (e.g., sensors, actuators); they are interconnected through several types of (wired and wireless) networks;

• the intermediate layer comprises clouds of interacting software components (Manzalini 2006) which provide virtualization of functions and features of underneath resources and networks; they are grouped in overlays, and have autonomic capabilities to provide self-adaptation, self-management, and self-organization;

• the upper layer is an ecosystem of composable applications, offered and used by different actors (e.g., Network Providers, LEs, SMEs, Prosumers), by leveraging the features of the underlying clouds of autonomic components [XI].

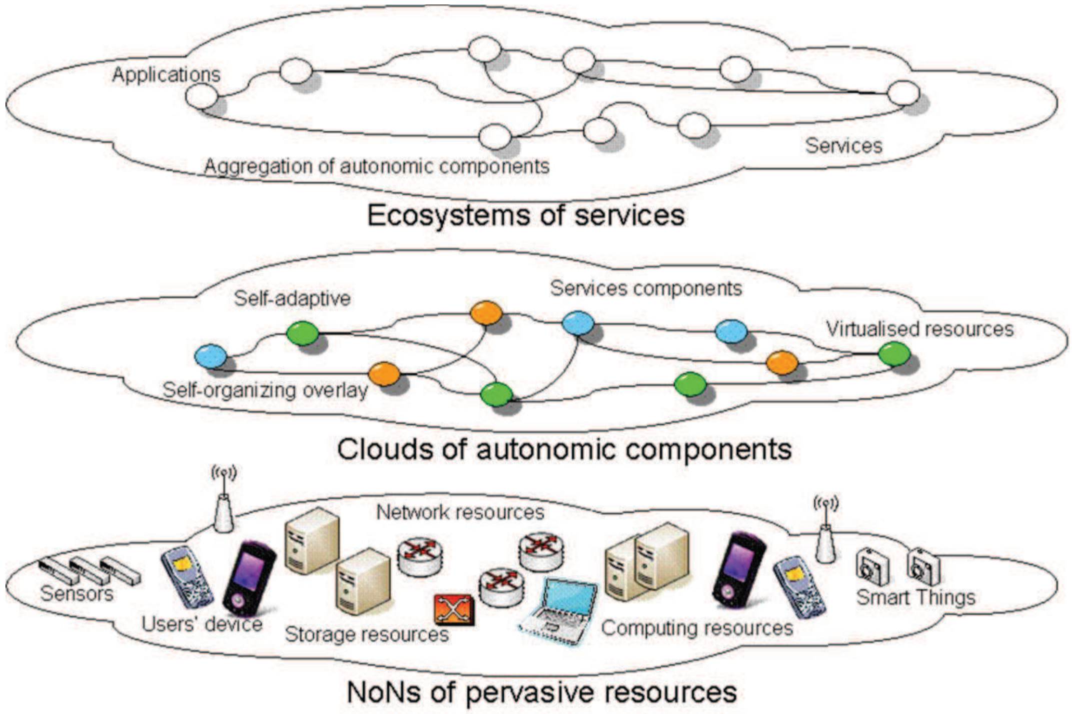


Figure 124: Layered Service Platform (for Network of Networks)

The upper layer is more a development and execution platform for services. This section provides a high level description of the functional characteristics and enabling technologies of the lower two layers, which are in charge for providing the capabilities to enable the creation and the maintenance of future application ecosystems.

9. 1. 1. Pervasive Resources in a Network of Networks

The considered network and services architecture (a Network of Networks) encompasses pervasive heterogeneous resources spanning from network element (transport nodes, switches, routers and even home gateways) to servers, to Users' devices (e.g. laptop, smartphones, etc.) and "smart things" (e.g., sensors, actuators); said resources are interconnected through several types of (wired and wireless) networks.

At the connectivity level (specifically Layer 3, 2 and 1), it is advisable to consider technologies and architectures (such as Generalized Multi Protocol Switching) capable of extending the control of packet switching to time division (e.g., legacy circuits), to wavelength (e.g. DWDM) and to spatial switching (e.g., for incoming port or fiber to outgoing port or fiber). The introduction of autonomic and cognitive capabilities in this context will allow the creation and maintenance of optimized overlay networks of storage and processing virtual resources. In particular, the idea is that autonomic and cognitive capabilities will complement the Network Control Plane (performing the connection management functions for the data plane) and the Service Control.

The basic motivation is that if a network node is enhanced with certain level of knowledge of the overall status of the network, autonomic operations decisions should be at least as good, if not better (in terms of global effectiveness and optimization) than those made in absence of data and knowledge.

The introduction of autonomic and cognitive capabilities, while owing credits to previous pioneering ideas and findings in the area of knowledge-based network management and of cognitive network paradigms, attempts at leveraging a novel architectural approach, more specifically suited for the outlined future network scenarios. In order to gather and represent the knowledge of the Network of Networks, nodes have to be described and represented. Properties of nodes like intended usage, linkage and aggregations with other nodes should be represented as well. Finally relationships between the different nodes, aggregates of nodes and other resources and users should be described. This approach should be for "networking" what the semantic web is for the Internet: the description and aggregation of information will have a paramount importance in the working of the Future Internet. From this perspective there is the need to develop a language for resources and services representation and request. This language should be a major mechanism that enables the creation of a Cognitive Plane (in this particular case a dynamic collection of information about resources operating in a Network of Networks and their respective relationships). The Cognitive Plane is a major enabler for the resource awareness and for the exchange of information between the service and the network layers.

9. 1. 2. Overlays of Autonomic Virtualized Resources

This layer consists of ensembles of software components that create a uniform shared substrate of virtualized resources, dynamically allocable to applications. The virtual resources rely on the capabilities provided by the underlying "physical" resources, according to a virtualization model, such as those based on hypervisors adopted to create and manage the virtual machines (Marinescu and Kroger 2007). The components have to implement mechanisms for partitioning the resource capabilities in "isolated" slices; the slices are the allocation unit and provide interfaces (at the required level of detail) to access the resource capabilities. Moreover, they have to implement hypervisor functions to monitor the state of the resource and of its slices, to detect critical situations, and trigger a decision process to select and activate corrective actions. Finally, they include functions for negotiating, allocating, de-allocating and, possibly, migrating virtual resources, i.e., slices, according to requests performed by applications/users.

These virtual resources are enhanced with autonomic capabilities. They must be able to react to (unplanned) internal or external events and self-adapt their behavior (e.g., through the dynamic change of their internal execution plans) and the one of the controlled resources (e.g., configuring them by using their controlling protocols). The autonomic logic is exerted by decision modules that determine the actions needed for ensuring self-CHOP (i.e., Self - configuration, healing, optimization, protection) behavior and for tuning slices allocation and configuration. The single autonomic logic is expected to strongly cooperate at the local level with those executed by other nearby components in order to implement decentralized algorithms able to exert supervision and resource allocation at global level (Ferrari, et al. 2009) and able to propagate the new behaviors at a global one.

In order to deal with scalability, each component must communicate only with a few nearby neighbors, interconnected through a (self-managed) overlay network. In particular, overlays are used to guarantee scalability, reliability, and abstraction from underlying networks. They aggregate resources sharing similar functions (e.g., all the resources which provide storage in a given geographical area) and support the spreading of cooperation messages among the local autonomic components. The distribution of local messages allows achieving distributed and global management policies. Moreover, overlay techniques by interacting with the Cognitive Plane (e.g., by means of information hashing in a DHT like fashion) allow interconnecting virtual resources allocated to a given application/user. These application-specific overlays create a uniform computational and networking environment able to decouple applications from actual pervasive physical resources and networks. Self-organization mechanisms can be used to optimize and maintain these overlays: by self-adaptation of their configurations according to dynamically evolving conditions on available resources and network conditions allows an optimization of the usage.

The features of this layer of autonomic virtual resources enable the applications to achieve a suitable level of resource awareness; through these capabilities, the applications are able to self-adapt and self-organize (with limited human intervention) by taking into account the availability of resources, the applications' requirements, and their continuous changes.

A network of networks tends to collect a number of heterogeneous systems and nodes that create the need for very complex integration and management. With the increase in number of nodes, the traditional management of large system will fail short, so there is the need to determine how the components of a network of network should be organized, configured and managed. The basic point is to start from the development of autonomic systems and to apply the technology in this realm in order to achieve self-CHOP features. Under the perspective of QoE, the self-CHOP features have a double importance: on one side there is the need to ensure that the entire network of networks is properly used, managed and optimized; on the other side, each single network or networked environment will try to optimize the usage of resources by allocating them in such a way to maximize its convenience. In this way a double problem arise: how to maximize the users returns and at the same time minimize the usage of available (and scarce) resources.

9. 1. 3. More About a Network Operating System

In the previous two subsection some general principles on how to organize the layers of a Network of Networks have been discussed. This subsection aims at providing a deeper description of the Network Operating system (essentially the level 2 of the proposed architecture). Future communications environment (the Network of Networks) will strongly rely on:

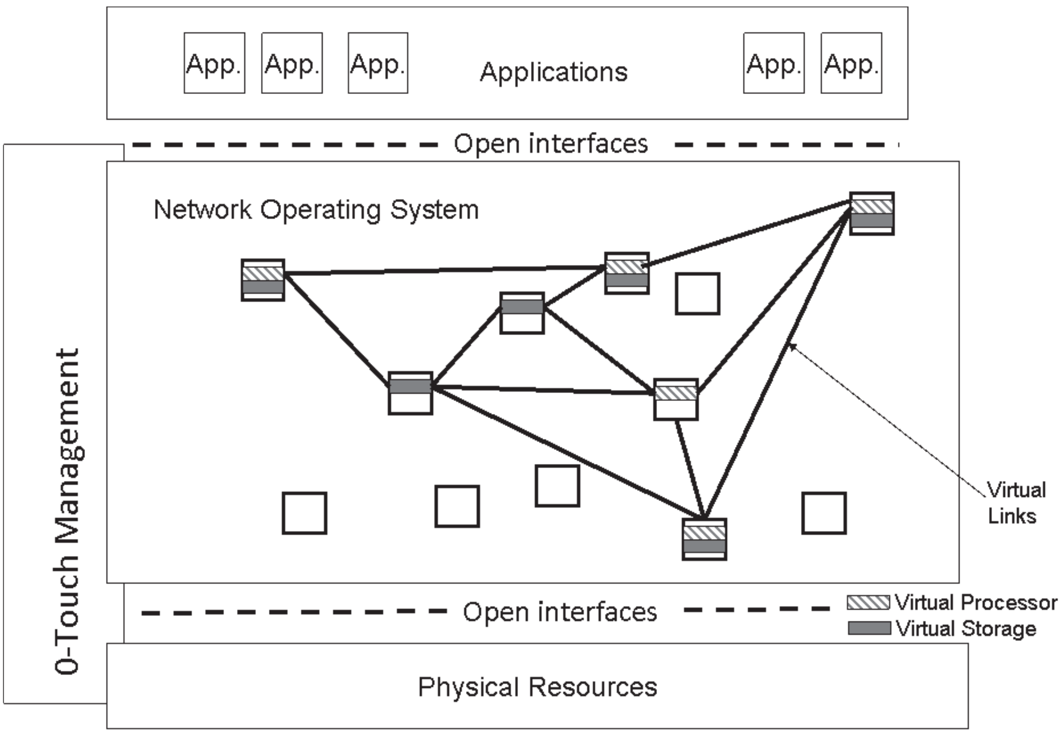


Figure 125: A 0-Touch Network Layering for the TelCo Domain

- the extension of the goal of cooperation between autonomous systems from the cooperative forwarding of packets to the cooperative orchestration of subsystems' resources (communication, processing, storage, sensing);
- the virtualization of resources;
- the introduction of self-CHOP properties in order to achieve an autonomic behavior of applications and resources.

This yields to a programmable networked environment, i.e., a Network Operating System, which, through representation and virtualization of networked resources spanning across many subsystems and different administrative domains, allows applications to negotiate for "virtualized" resources, to allocate them, to control and program their functionalities according to the specific needs and requirements. The Network Operating Systems is based on a decoupling between the physical resources and their virtualization. Logical resources will offer to applications open interfaces for negotiating, allocating, controlling, monitoring, programming and using resources (**Figure 125**). The implications of this approach are many:

- resource virtualization, based on abstraction for coping with heterogeneity and on the definition of dynamic slices for multiple allocations. An example is the router virtualization as described in OpenFlow (McKeown 2008);
- autonomic features and bio-inspired algorithms (e.g., self-organization ones), to deal with complexity in organization and structuring of virtualized environments [XVII, XXI];
- an open language for describing resources without over imposing a specific model and open interfaces for requesting functionalities to (virtual) resources.

It is important to present a network view that applications can manipulate (through open interfaces) in order to achieve from the networked environment the appropriated level of system availability, QoS, granular programmability.

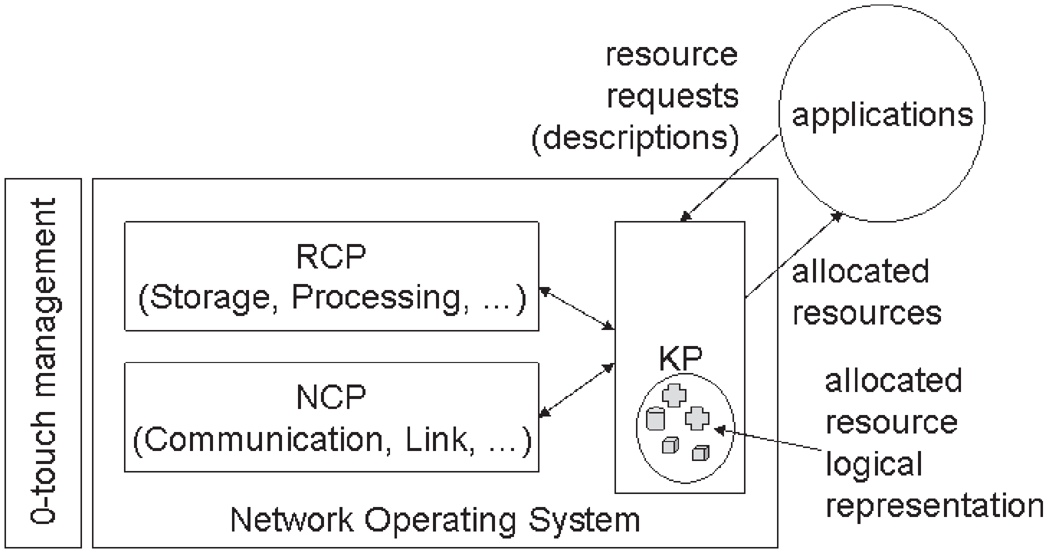


Figure 126: Architecture for Generalized Cognitive Networks

The proposed functional architecture is capable of hooking together and managing different types of virtual resources (even pertaining to the User domain), such as computing, storage, network resource (e.g., servers, routers, nodes, laptops, mobile phones), "smart things" organized in overlays networks. The architecture is based on three planes, which cooperate in the allocation and organization of virtualized resources and their management and exploitation (**Figure 126**):

- Network Control Plane (NCP) performs connection management for the data plane (where the actual forwarding logic stands) for all network interfaces (both packet-switched and non-packet-switched);
- Resource Control Plane (RCP) performs the allocation of IT-based virtual entities according to the requirements put forward by applications; allocations are decided by considering the distribution of resources and applying algorithms that enforce autonomic features; the virtual entities represent either infrastructural resources (e.g., computing and storage) or service elements (e.g., to access application functions or data);
- Knowledge Plane (KP) receives from the applications the requests of resources and interacts with the RCP (e.g., for allocating storage and computing resources) and the NCP (for communication resources); moreover, it maintains a logical representation of all the resources allocated to applications, e.g., used for supervision purposes.

In general terms, upon receiving an application request for resources, the KP interacts with the RCP to get a set of (virtual) resources candidate for the allocation. RCP returns a set of resources fulfilling the application requirements, each of them labeled with some additional information useful for optimizing the selection according to the adopted criteria (e.g., performance, cost reduction, or power savings). KP, by using the updated information on the availability of network resources (e.g., obtained through NCP) and the information previously allocated to the application, selects the "best" subset of virtual resources returned by RCP, and requests NCP the provisioning of virtual links interconnecting the selected virtual resources. NCP, by means of a distributed network signaling (e.g., with RSVP-TE) allocates the network resources, fulfilling the application requirements (e.g., QoS parameters) and returns to the KP the data on the allocated virtual links. In case of failure, KP can either select a "sub-optimal" subset of virtual resources, or negotiate with the application a refinement of the requirements. The KP is also in charge to represent to the specific application and to its users a virtualized representation of the allocated resources, This representation could be very detailed in order to allow a very granular control over specific virtualized resources or being quite general in order to allow the application to exert a "light" control over the networked virtual system (e.g., the entire network represented as a pipe with some parameters that can be toggled by the application).

Form the programmability point of view, the network operating system (**Figure 127**) should be structured in such a way to allow for:

- The management of allocated virtual resources; actually the interfaces offered should support functions for requesting (at run time) new resources types, in order to satisfy the changing needs of the applications;
- The possibility to negotiate and allocate the resources so that an economics of the network can be provided (providers of resources should have a kind of return); the possibility to allocate and use the available resources.

Different levels of programmability (**Figure 127**) are needed in order to satisfy the granularity of control required to implement services. The lower level interface should ensure the possibility to deal directly with specific infrastructure resources and functionalities. At upper level the resources are more and more represented as aggregation of functionalities and abstraction of physical systems. Eventually a view on the Knowledge Plane could be offered to application in order to control the availability of resources according to specific needs and semantic. This could be a means to confine the possible resource allocation for applications insisting on a specific environment or for regulating how a specific service/platform provider wants to use the "network of networks".

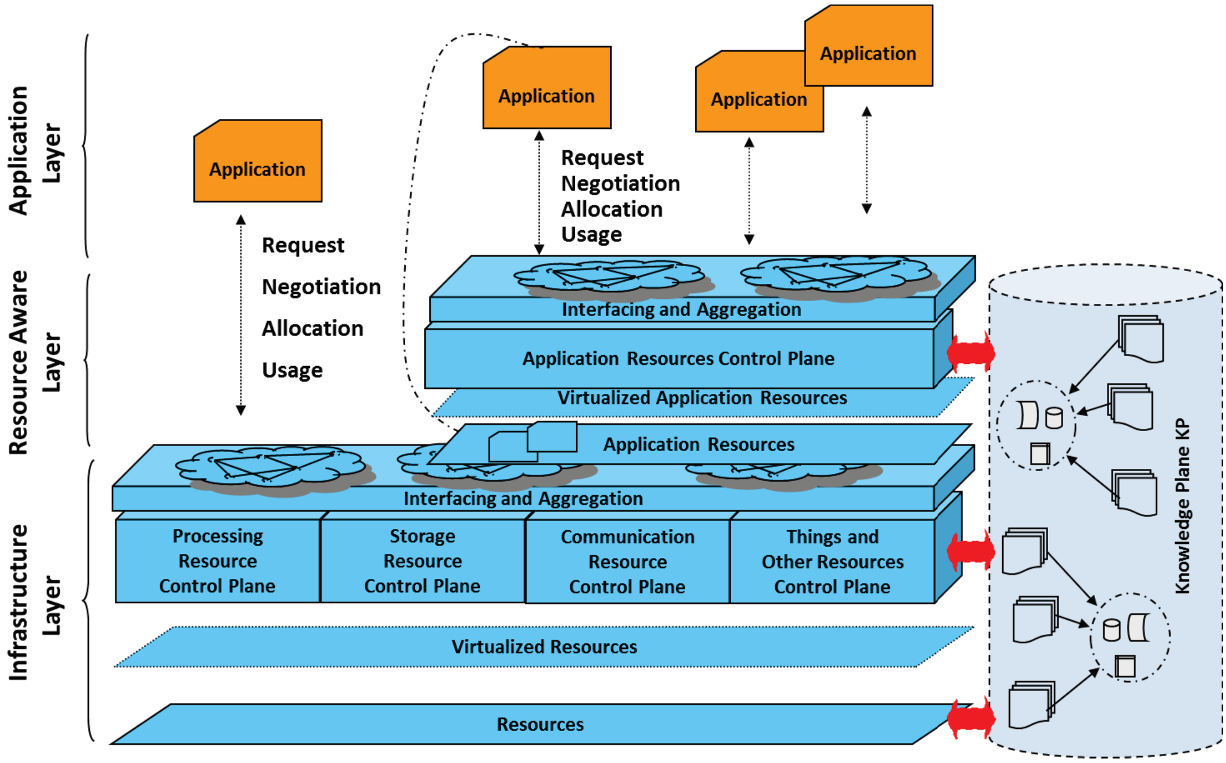


Figure 127: Programmability View of the Network Operating System

The Network Operating System itself could be implemented in two different ways: as a centralized system monitoring and exerting control over distributed physical resources, or in a widely distributed manner, in which each virtualized resource is an object executing on a distributed environments. Each of the three planes (KP, NCP, RCP) could even be implemented in a different fashion, for instance the KP in a strongly centralized way and the NCP and the RCP in a distributed manner, e.g., as a set of instances running on different nodes, interacting through distributed processing mechanisms. All the instances of a plan are interconnected through an overlay; its links are in charge of transporting signaling-like messages, for enabling the interworking and the cooperation of the instances. Overlays are suitable means to abstract the heterogeneity of the underlying computing and communication infrastructure, cope with the dynamic evolution of the nodes in the system, and to achieve scalable solutions. Even autonomic/cognitive capabilities could be implemented in different fashions: on one side, each plan can be running on servers that do guarantee autonomic features, on the "distributed side" autonomic and self-* features should be locally implemented in each distributed node hosting virtual resources and on the virtual resources themselves. The plans work on virtual resources mapped onto physical ones according to a virtualization model. Network virtualization is a powerful technique as it provides flexibility, promotes diversity, and promises security and increased manageability:

by allowing multiple heterogeneous network architectures to cohabit on a shared physical substrate, network virtualization is a diversifying attribute of the future inter-networking paradigms. The considered model enables to create, starting from a shared set of physical resources, a uniform substrate of virtual entities, dynamically allocable to applications and providing the expected view and control functions. The model is general as it can be applied to different types of resources, such as computing, storage, network components (e.g., bridges, NAT, host-only, virtual switches, virtual DHCP server, and virtual network adapters), "things", service components, etc. A virtualization model includes:

- an abstract view of the features/functions provided by the resources, in order to (1) simplify the access to capabilities, and (2) mediate the heterogeneity on interfaces/protocols actually provided by physical resources to control/access their features;
- a mechanism to group the capabilities of the resource (e.g., in terms of processing, storage, bandwidth) in isolated partitions; the partitions are to be used as allocation units of virtual resources; configuration mechanisms allow to associate, in a flexible way, the capabilities of physical resources (e.g., storage dimension, execution cycles, queue lengths, optional functions) to partitions, and to configure QoS parameters, to fit the negotiated SLA; each partition is an isolated context to protect the physical resource and the other partitions to an incorrect/malicious behaviour of the application and create an environment for the enforcement of negotiated SLA, by avoiding (functional/non-functional) interferences among different partitions;
- a formalism for requests and negotiation of virtual resources, e.g., in terms of dimension/quantity of resources to be allocated and description of non-functional features; an associated protocol is adopted to request/negotiate virtual resources.

Prior-art (Chowdhury and Boutaba 2010) is already providing interesting proposals of network virtualization at different layers: examples are L1-UCL, L2-VNET, L3-AGAVE, VIOLIN, etc.. However all these examples suffer the lack of node performance and programmability, sufficient isolation, cross-layer and cross-domain interoperability, optimization of physical resources and other features. The adopted model is based on a Virtual Machine Monitor (VMM) able to host several Virtual Images (e.g., virtual routers) able to interconnect to the data plan. A "virtual-machine" model, similar to the one based on hypervisor proposed by XEN (Barham, et al. 2003) is adopted for virtualizing IT-based resources. More discussion on these technical choices and their implication are available in

[X]. Virtual resources allocated to a specific application are organized in monitoring "views" (e.g, overlays in the case of distributed control plans, group of agents and components in the case of centralized plans). KP, which is in charge to keep track of all the resources allocated to an application, is in charge to create, maintain and check the consistency of these logical views. For instance, monitoring overlays can be used by an application in order to have a global control over all the virtual resources allocated to it: by means of such a monitoring overlay, the application can, for instance, perform some supervision/management activities, e.g., in order to optimize its performance or the use of the allocated resources (e.g., in order to reduce the costs), or to detect/recover some failures. Monitoring overlays can also be used by KP in order to supervise allocated virtual resources. Monitoring overlays interconnect software entities in charge to monitor/supervise the virtual (IT-based or network-based) entities. For instance, they could interconnect, the "slices" of a virtual machine, or the software representation of (GMPLS) virtual connections. The links interconnecting distributed elements of monitoring overlays are used to transport signaling messages. These signaling links are set-up as virtual links, by using the NCP capabilities. Architecture entities (e.g., local instances of the plans, virtualized resources) must be enriched with autonomic features in order to:

- promptly react to critical/unplanned situations by identifying contingency plans;
- optimize resource configurations and allocations;
- tame complexity of huge amount of entities, by imposing aggregations and cooperation.

Autonomic features are introduced by means of autonomic control loops within each virtualized resource:

- internal control loops: they are in charge of achieving self-awareness, to control the internal state and events, to detect critical and/or unplanned situations, and to trigger a decision process to select and actuate the corrective actions;
- external control loops: they are in charge of achieving self-organization and self-adaptation of the autonomic entity with respect to its "environment", by processing events and messages sent by other entities (e.g., resources, local instances of the plans, and applications) interconnected though overlays.

Autonomic entities in the architecture can be implemented as ensembles of interacting autonomic components, like those provided by CASCADAS ACE Toolkit (Deussen, et al. 2010). External control loops can be used to implement a cooperative behaviour among peer entities to perform distributed decision algorithms or to create a "non-local" vision of the state of the system. The organization of the entities can be distributed or centralized in a single execution environment depending on the nature of the problem to solve, design choices and the possibility to actually centralize virtual resources (some domains could prefer to have under their control the virtualized resources and not demanding to other domains the control; an example could be the virtualization of resources pertaining to different Operators).

Some Use Cases

The examples in this section focus on a business scenario in which a company or an aggregation of companies (i.e., a consortium or companies in a customer – provider relationship) want to create a networked environment able to provide connectivity and computing services according to actual needs of users and able to integrate in a meaningful way the companies infrastructures and IT resources. The features of this network of networks are challenging: the system comprises heterogeneous private networks, user terminals that are interconnected through one or more public networks; the request bandwidth and computing/storage capabilities are highly dynamic and not always predictable depending on the actual needs of users and endpoints that act in a distributed fashion on such a large system. The use cases try to show how the distributed and centralized paradigm could be implemented. In the conclusion some hints on their value is further discussed. The two proposed scenarios are related to: a) routing path allocation, and b) computing/storage allocation. They could be combined as an integrated service proposition for businesses.

Flexible Routing Path Allocation: for this scenario, the possibility to control routers à la OpenFlow is assumed. A number of open routers in the different subsystems (private, public networks and even on some terminals) can expose control interfaces towards external controllers that instruct the specific router on what kind of policies to apply or how to deal with IP packet flows. Virtualized representations of a router with a pre-assigned range of capabilities (maybe representing the minimum and the maximum performance that the virtual router can cope with for that specific usage). **Figure 128** represents two possibilities: to fully distribute the logic preserving the domain independence (each single companies manages its domain) and enforcing the cooperation by means of specific protocols/interfaces among domains vs. a more centralized solutions in which virtual representations of routers (and even controllers) can be localized into a single domain (or in a single machine).

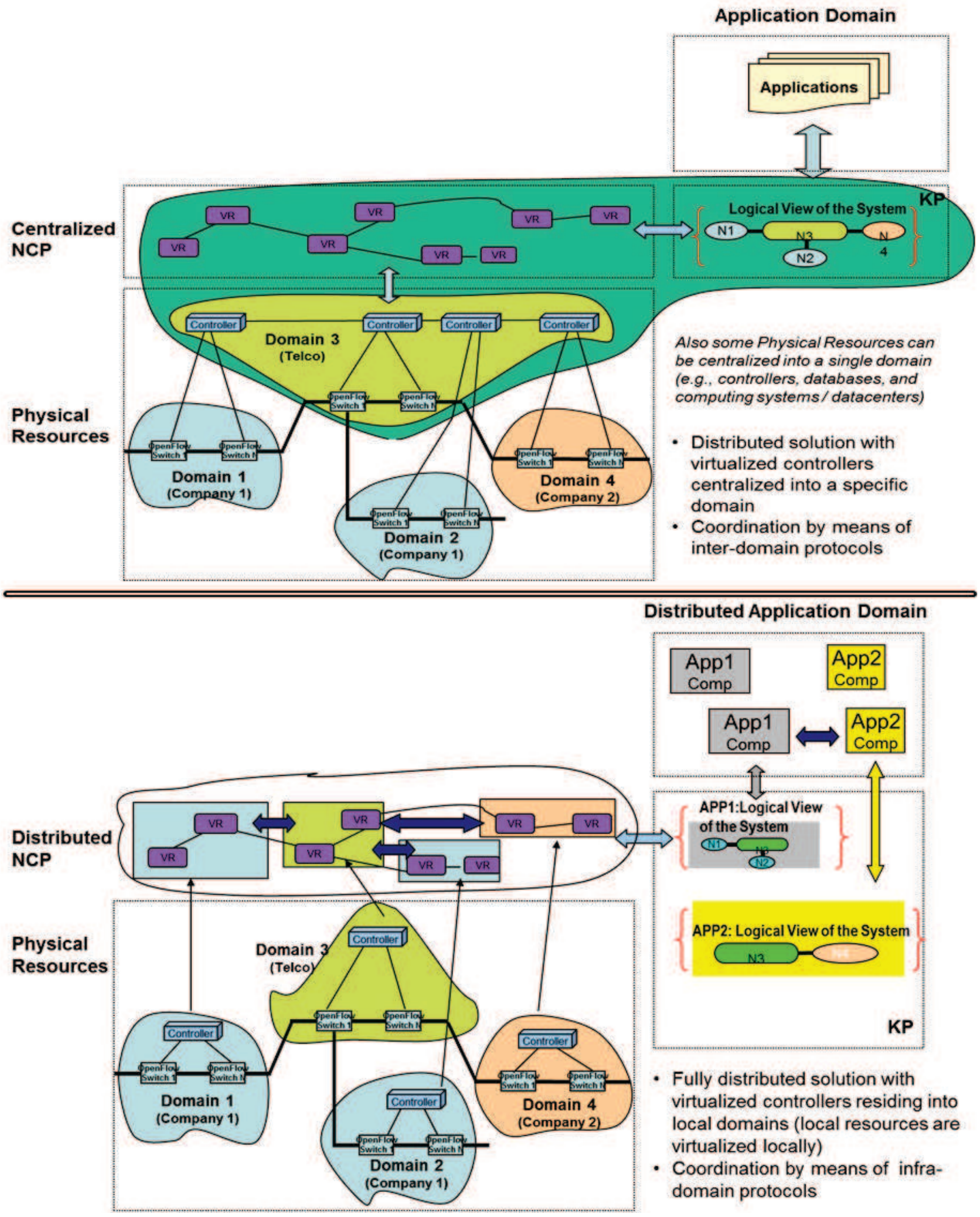


Figure 128: Centralized vs. Distributed Approach for Path Control

Especially in the case in which nodes are relatively small in number, the centralization of controllers and virtual slices could be quite effective (for instance many new deployment in the realm of Next Generation Networks are aiming at drastically reducing the number of nodes in the network. This flat architecture is a good candidate for a centralized implementation. In case of more fragmented environments (comprising several networks with a very dynamic behavior) coordination between different controllers/virtual router should be exerted by means of inter-domains programming means. Each domain can be organized according to local design choices and implementation paradigms, but it will provide to other domains means and protocols for managing the allocated resources. The applications can operate on abstract objects requesting the change of communication related parameters (e.g., more bandwidth from/to a certain domain, a certain jitter or others. The level of granularity should allow applications to fully exploit the environment. A “one to one” granularity between physical resource, virtualized resource and logical representation can be offered to applications. Applications themselves can be implemented in centralized or distributed fashions. In the centralized case, application will essentially run on a single (and homogeneous) computational environment possibly offered by a single provider that will guarantee security, protection and full availability of the environment. In the decentralized case, components of the application will run locally and will globally interact with others in order to coordinate and governing the allocated virtual resources in a pretty similar way as overlay networks do. The centralized implementation could make use of agents that cooperate in order to implement the application logic by operating on the underlying resources, while an overlay application could directly operate on resources. The logical view of resources is also able to support the dynamic creation or allocation, the monitoring, and update of resources (whatever a new resource can be defined at the application level: a new network, a new router and the like). In this way control applications can dynamically change the topology of their communication space by attaching new resources and to create virtual communication links between them. At the physical level this means to create the right connectivity between different end points according to parameters and feature requested by the upper layer.

Dynamic Storage Computing Allocation: this scenario is based on the possibility of creating a distributed data center by aggregating different computing systems ranging from established cloud of computing resources to single resources made available by the single user. The presence of small resources possibly available for a short period of time (e.g., this is typical for P2P networks) request to have a large number of them in order to guarantee the availability of “some resources”. On the other side, this approach can enable the creation of very large data centers by reusing capabilities offered by largely distributed (user) systems. The “service” aims at dynamically providing processing and storage power on demand. User equipment is considered as a resource and some computing tasks can be allocated to them. The approach taken is similar to the one used in the Google’ s data center: to replicate data and processing over different cluster of computer and statistically expect that at least one will complete the computation or will return the data stored. Slices of storage or processing can be negotiated (considering the replication factor) and allocated to users. In this case the difference between a centralized approach and distributed one is more evident: in the centralized one the allocation of resources is directly controlled by the KP. Each available resource is able to communicate with the KP directly and provide all the info related to the life cycle. This is done in a sort of master-slave fashion.

The KP then is able to monitor all the instances of resources that are available and on application demand it is able to allocate the resources that better match the request. The distributed case is more complicated, in fact the KP is distributed: this means that the collection of information is different from the master-slave one. Each resource (a peer) should inform its peers of the status and its availability. This information has to be collected and made hashable or addressable by means of widely distributed mechanisms. Distributed Hash Tables have been largely used for implementing distributed file systems and the same mechanisms could be implemented for supporting the distributed KP.

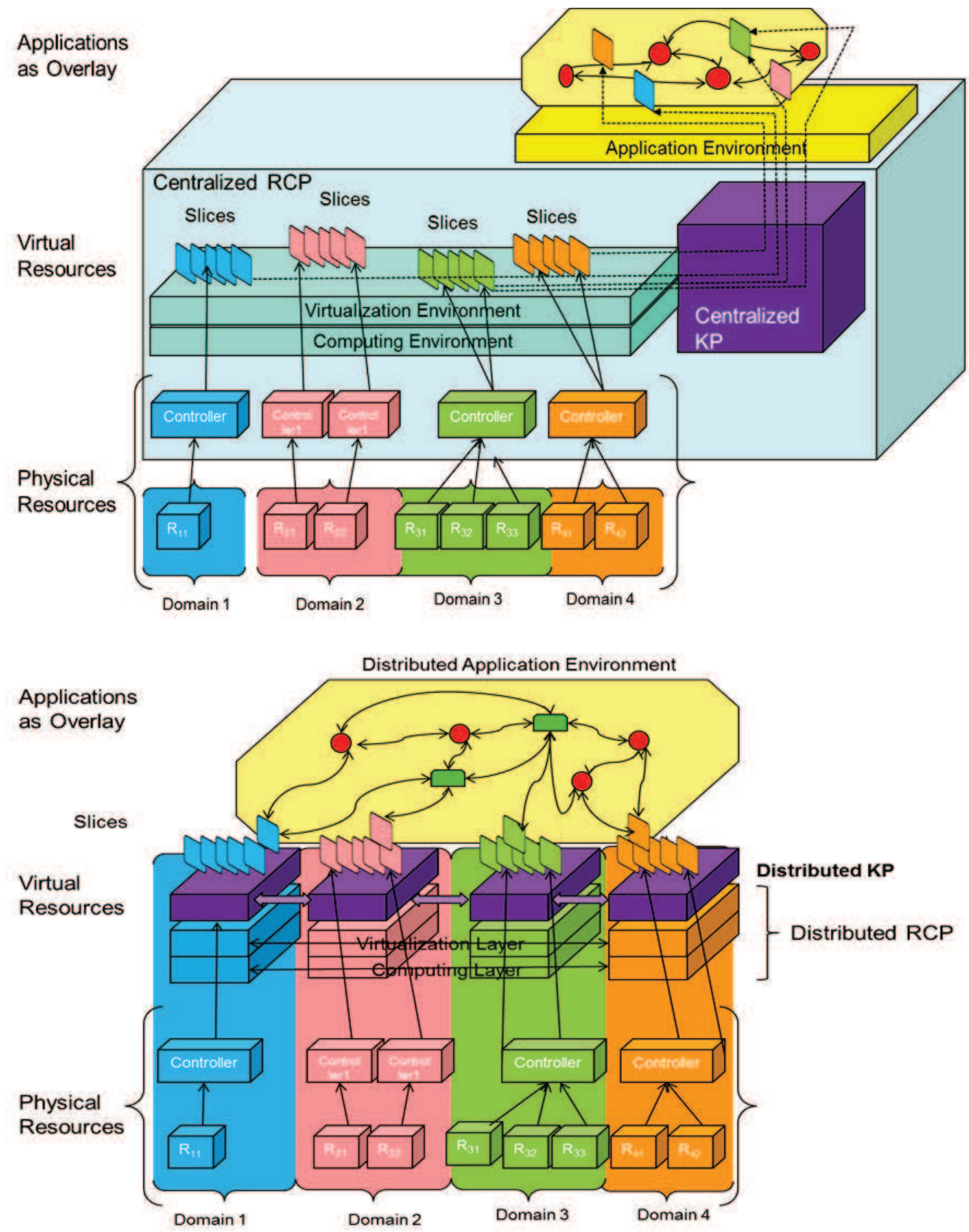


Figure 129: Centralized and Distributed models for RCP and KP

The centralized or distributed implementations (**Figure 129**) of the RCP show some interesting issues: in the centralized one, the computing and the virtualization environment can be normalized and using the same model. The entire organization of the RCP and KP are easy and straightforward, they just request a simple event handler in the resource environment. In addition an application programming environment can be provided with similar mechanisms adopted for the Network Operating System. In the case of distribution of functions, the local computing and virtualization can be different from domain to domain. In addition the distributed KP could be using the same local infrastructure for operating. The interworking has to be reach by well-defined communication mechanisms and protocols that have an inter-domain value. Slice of resource made available to applications can reside on the local domain. At the application level, the example considers distributed applications that make use of slices. In the centralized example, the application environment is homogeneous and application can be distributed over a uniform platform. In the case of distribution, the application components are distributed directly on the local environment.

Centralized VS. Distributed Network Operating System.

Some initial considerations can be made in order to guide further development in the area of these studies. The issues related to QoE require the coordination of different subsystem. Under this respect the centralization has the merit of providing a simple enough and intuitive way of programming and coordinating the different subsystems, however it has the drawback of imposing to some subsystem to handover the control of local functions to a centralized domain. Some subsystems)even if eager to cooperate) could still value the "independence" of their behavior, Currently many public networks are aiming at the decrease of nodes and a centralized approach could be extremely meaningful because the number of resources to manage is low (even if the number of slices could be very high). On the other side, the number of intelligent nodes on the edge of networks is increasing creating new problems in managing the complexity of dynamic and unstable subsystem. In this case, the distribution of logics together with the introduction of some levels of self-organization could be beneficial for dealing with larger and larger edge systems exploiting their increasing power. At the application level, the issue related to centralization or distribution has be solved on a case by case approach. Some applications are distributed by nature (e.g., P2P systems) and so they fits naturally in the distributed approach. Others can benefit from a level of centralization especially in the case of simple applications that do not need to enforce a granular control over networked resources. In this case the application designer should be able to choose the pattern that better fits its application without having to many constraints imposed by the architecture. This flexibility is another important factor, the network operating system should not force the application towards a specific design pattern, but it should instead accommodate and support several paradigms. This is probably the best lesson learned so far.

9. 1. 4. Small Cooperating Objects Can Do Great Things

Autonomic behavior should be enforced in smart objects and nodes. The functionalities and the processing complexity required to achieve an autonomic behavior should be commensurate to the global capabilities of the object itself. Implementing heavy algorithms on objects having limited processing and storage capabilities will not give the desired results. So the processing complexity of the (distributed) autonomic capabilities should be kept at a minimum and algorithms should be simple in order to allow objects to execute simple functionalities. The simultaneous execution of these functionalities could result in autonomic behaviors.

Figure 130 depicts the case of a set of objects (some of them could have an ephemeral existence) that cooperate in order to store a value and to make it available permanently. The resilience of this system is based on the concept of the Active/Active configurations for redundant servers (Availability Digest 2006). The more objects are available and the more likely is the ability to store the value. In case of update of the value, the change has to be propagated to all the linked objects that represent a sort of distributed memory with different replicas.

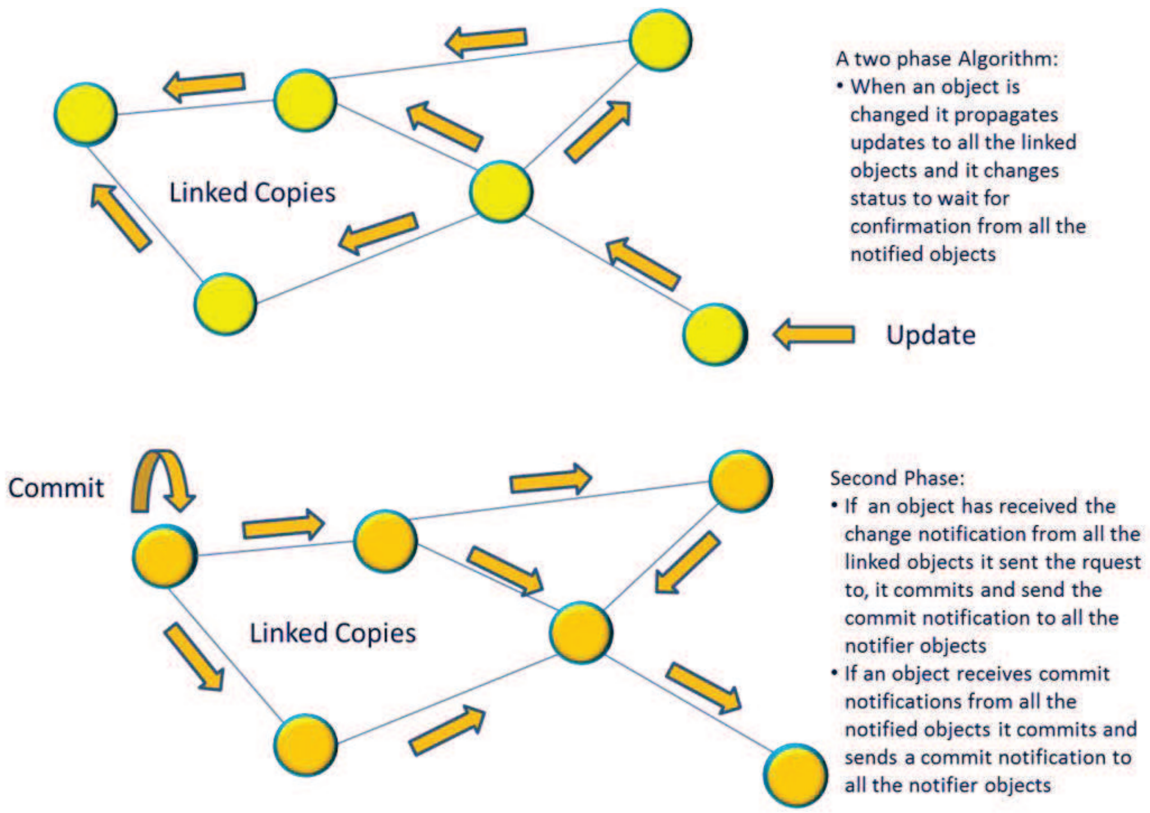


Figure 130: Gossiping Algorithm for Data Entanglement

The idea behind the example [XVI] is to create a sort of data entanglement effect, i.e., any modification to a value is rapidly spread among all the object storing the replicas. A simple gossiping algorithm is sketched out in **Figure 130**. Similar algorithms have been proposed in [VII] and [XVII] for optimizing the resource allocation of processing nodes in big clusters of data. These algorithms show that simple mechanisms can allow great system to converge towards good solutions. This is another pillar on which to build new architectures done with thousands of nodes. In addition these solutions (and the one proposed by Web Companies like Amazon and Google) point to a different approach with respect to Consistency, Availability and Partition Tolerance, (it is the Brewer's Theorem or CAP Theorem). The network intelligence approach points to guarantee consistency of all the information and its availability, the cost is a more centralized approach (i.e., a much lower Partition Tolerance). The peer to peer approach instead promotes Partition Tolerance and Availability, while Consistency is sometimes disregarded. This means that information can be not aligned or consistent, but it is provided anyway (it is available) and distributed in order to guarantee to users the possibility to access to information. This is a great difference in the approach and represents how Web Companies have a winning proposition: they provide data/information anyhow to users, while the Operators try to enforce consistency (ensured by means of a less distributed infrastructure).

9. 2. How to Migrate From Existing Platforms

As a further example of applicability of some of these concepts that emphasize how these conceptual framework can encompass current architectures, the case of the Internet with Things and the Virtual Continuum are considered. The value proposition in this approach relies on extended communication capabilities:

- The connection between a real object and a virtualized one is based on Always Best Connected communications, i.e., the communication infrastructure (the object itself and the network of networks) will strive to provide a reliable link offering the maximum capability possible depending on the context of use of the resource.
- The platform provides a set of functions and means for exerting the expected behavior of the virtual object, e.g., processing, storage and communication capabilities (e.g., PubSub engine, data store). In addition it provides means and functions for controlling, managing and monitoring the virtual objects (e.g., the object repository for brokering of objects and resources).

- A set of Application Programming Interfaces, APIs, for allowing high level of programmability of virtual objects and platform functions/objects. APIs are fundamental for service construction and allow to programmers the capability to create new services and applications. Programmers can create general solutions as well as vertical ones that are tailored for specific markets and businesses.
- Generic services and applications can be created on top of at least two different levels of APIs: virtual Object APIs for directly control the virtualized instance of a real object, and higher level APIs for benefiting of all the rich functionalities offered by the platform. These API are used by developers (Internally and externally to the TelCo) for creating services. In this respect, the servitization capabilities can leverage a number of innovative applications.
- A set of already developed and tested applications by a large ecosystem (here the network effect can have its greater outcome).

Figure 131 represents the possible value propositions with respect to the architecture.

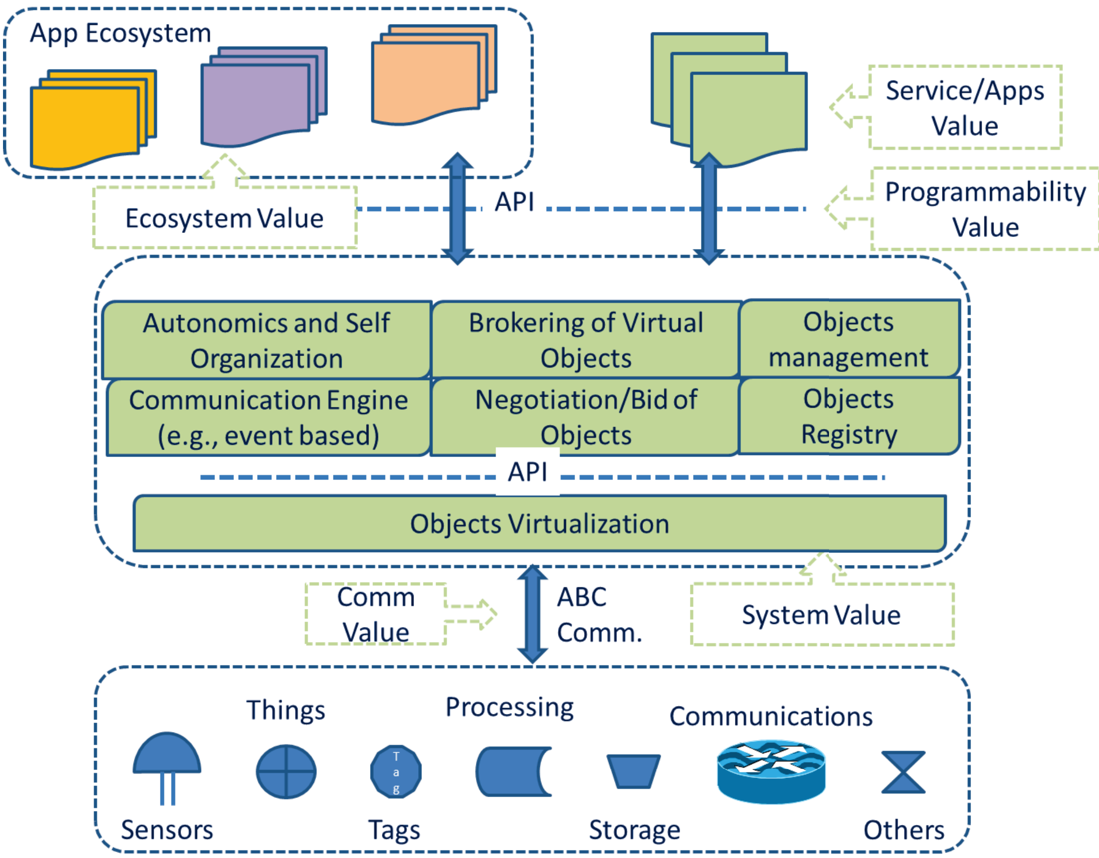


Figure 131: Features of a Service Enabler Platform.

This architecture is based on a control model based on events. It is the most appropriate to deal with potentially millions of sensors emitting a considerable number of signals to be received, manipulated, dispatched and even "understood". A PubSub mechanism is the chosen one because it is capable of dealing in various ways with the events (from a simple forwarding to the right receiver to the extraction of information from the aggregation of "events"). It could be a viable mechanisms for supporting stream processing capabilities as those proposed by IBM [stream processing]. The bulk of the platform can be built on already existing functionalities of current platforms, for instance the object management is a major value in current implementation of M2M solutions as well as the Registry of objects (in order to determine the active SIMs). The introduction of autonomics functions for allows to make simple sensors more reliable (by means of aggregation and the characteristics of the inner circle [XII]. Non reliable objects can be made more robust by means of virtualization and the representation of their functionalities by means of aggregated objects. Also security mechanisms could be introduced in this fashion creating a sort of shield around simple sensors. These functions are represented in **Figure 132**.

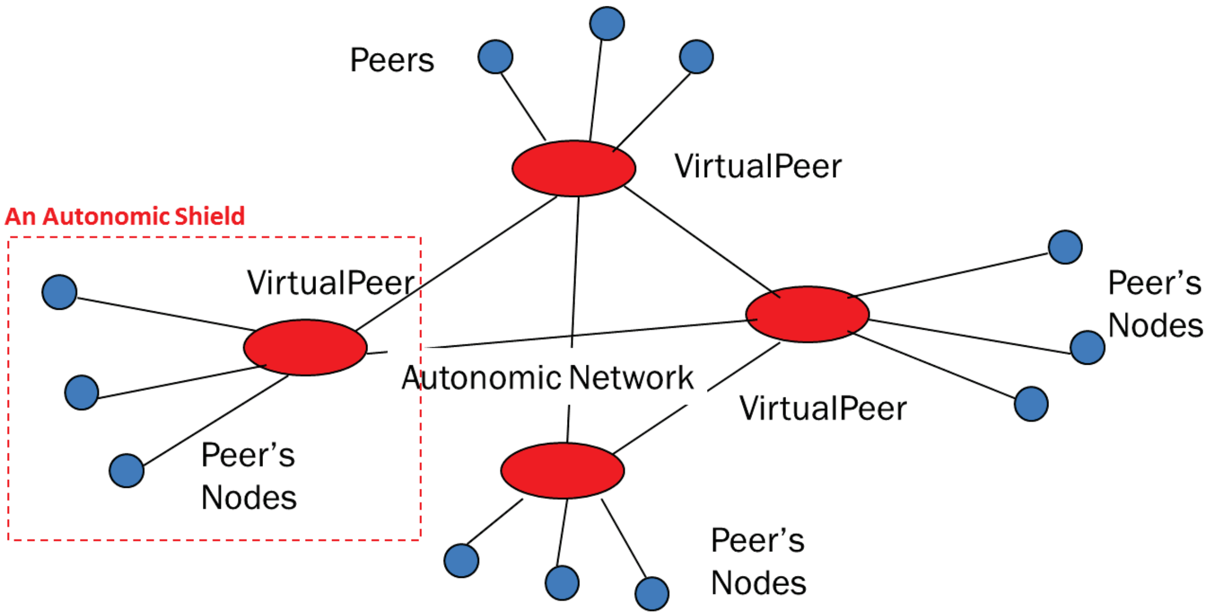


Figure 132: Virtualization for Introducing Autonomic Capabilities

Brokering of resources and functions will instead extend the existing capabilities enlarging the reach of the platform by deperimetrize it and allowing other providers to contribute with resources and functions. The Brokering function to a certain extend will help in promoting new and more open business models.

It is worth to consider the difference of this approach respect to the traditional attitude of TelCos. In a traditional business proposition, the TelCo was trying to create a full offering by acting a number of different roles in the value chain of a service (from infrastructure, to platform up to service provider). This was done by leveraging the network and its capabilities. In this approach, the TelCo will focus mainly on the role of Platform Providers and will act in order to collect and integrate different infrastructures (processing, storage and communications) as well as to allow different service development capabilities. For certain services, the approach still allows the TelCo to play the full set of roles (i.e., the service provider) or to focus on the communication infrastructure (connectivity provider). This flexibility will guarantee the possibility to choose the best role depending on the service proposition.

A representation of the new ecosystem is given in **Figure 121**.

One of the major differences is the role of the user: it is not only a customer, it can be a provider of resources and be a partner of the TelCo. The platform in fact is built also by putting together and integrating resources pertaining to different administrative domains (and end users will have a key role in this). One requirement of the platform is to be able to easily and dynamically integrate resources, systems and networks of other actors in such a way to cooperate in the creation of a synergic network of networks.

9. 3. Summary of Conclusions

This thesis was an attempt to determine a number of questions that will shape the future of the telecoms industry with particular emphasis on the network and service architectures. In this fast changing situation, even formulating the right and long-lasting questions can be considered as an important contribution to research. Some part of this thesis has been the object of some discussion with experts in the telecoms industry and the underlying consideration related to the future vision of this industry are deemed worthy of consideration.

Reconsidering the question posed in section 1. 2. and section 1. 3. , some conclusion can be drawn:

Best Control Paradigms. The Client-Server paradigm seems to be inadequate to the increasing distribution of functionalities and objects. Actually even the champions of this paradigm have two sides in their systems: the front end is organized (for simplicity) in a Client – Server fashion, while the "service engine" (e.g., in Google search or in Twitter) is organized in a fairly distributed manner and a combination of distributed processing techniques are used (from peer to peer, e.g., Amazon Dynamo, to logic programming means like in Map/Reduce). Even the web is adopting mechanisms to solve the drawbacks of the

Client – Server paradigm. For instance, the introduction of the Websocket mechanisms (Fette and Melnikov 2011) is an attempt to solve the one way communication between the client and the server, with the server that cannot send results in more than one transaction. This lead to the consideration that in the future a large number of architectures and services will expose simple (and limited functionalities) by means of client – server mechanisms, while the richer processing will be carried out in highly distributed environment using more adequate techniques. The Network Intelligence approach seems to be a remembering of the past and its applicability for providing deperimeterized services has been questioned several times [V], [VI], [XVIII]. The main drawback is that network intelligence architectures are strongly coupling network functions with service behavior and capabilities. In addition, services are to be executed in the administrative domain that owns the user subscription. The peer to peer paradigm epitomizes those architectures that support the possibility to adopt different control paradigm (e.g., event based) within a networked environment and doing so they provide a good compromise that results in a high expressive power combined to the capability to leverage the edge functions. Obviously the best control paradigm strongly depends on the service at hand, but in general, the peer to peer paradigm support the power of distributed processing, with the capability to abstract the functions and nodes at the right level, in addition it support several mechanisms for object to interact. This comes with the price to deal with a bit of intricacies due to manage an overlay network

Expressive power of the solutions. Under this perspective [XVIII] has discussed the issue in terms of capability of a control paradigm to represent functions. The event passed interaction is very simple and powerful because it can be used to represent all the other mechanisms, e.g., the client – server paradigm can be represented in terms of two related events, and the trigger mechanisms of the Network Intelligence can be represented as a set of events that represent either trigger notifications towards the service layer or commands to the control layer. The peer to peer paradigm is usually based on event exchange between peers (for instance the gnutella protocol (Ripeanu, Foster and Iamnitchi 2002)). Simplicity at this level means that in order to represent complex situation, there is the need to relate events each other, and to dispatch them to the right nodes. However an event based system can be very expressive (because each event is treated singularly as an element of information/control) or can support abstraction (putting together a number of events exchanged between peers). In order to deal with the complexity and the requirements of smart environments the even interaction model and the P2P control paradigm are the two elements that guarantee the greatest capability to represent and program the complexity of future systems. Also APIs are prone to the problems of expressive power [I], [XVIII]. In fact, when API abstract too much the functionalities of a protocol or merge functions of two or more protocols there is the possibility to lose control on the basic functionalities offered by the specific protocols. Abstraction is in alternative to granular control. Future programmable architecture should be designed in such a way to have different abstraction capabilities (and related API in order to cope with the protocol dependence of the middleware (Messerschmitt 1996), [I].

New classes of services. They rely on virtualization of functionalities, like in the case of Virtual Continuum and in general the services of the Internet of Things, or in the ability to properly manage the fluxes of data from different sources (e.g., the personal data case). Smart environments subsume for a lot of the services discussed in this document. Here it is important to stress out that virtual reality or even augmented reality based services could open new service scenarios strongly determined by the capability of presenting a nice and compelling user interface to user. All these classes of services are characterized by social factors (i.e., the users are involved in the provision of the services as content providers or as resource providers) and by the high distribution of data, resources and functions. Dealing with them in traditional fashions (e.g., with client - server approaches like in the Web of Things, or with network intelligence based platforms (Blum, et al. 2011)) is possible but very cumbersome. Mapping the events occurrences generated by a set of sensors to the flow of the SIP protocol or similar ones is difficult and, because of the middleware dependence of the protocol, it can drastically reduce the information content offered by each event and impairing the whole service provision. A platform like Twitter is much more adequate (and a rational choice) for these services than any implementation of the IMS.

Right architectures for new classes of services. As discussed in this document, forcing the services to the architecture at hand can drastically reduce the service functionalities and the capability of the system to represent the complexity and the richness of interactions and relations between entities involved in the service. The combination of virtualization, overlaying and autonomic networking seems to guarantee the right combination for implementing the new classes of services (as well as the traditional ones). In addition, virtualization can offer the possibility to segment networks and service environments according to the needs of specific applications or customers. Flexibility of the architecture and the ability to readily encompass resources from different networks (in a network of networks fashion) are of paramount importance. The new architectures must support the deperimeterization of services (i.e., the capability to offer and support services independently from the presence of the specific network resources strongly associated to the user and the provider). This feature can easily be supported by the client – server and the overlay paradigms while it is a major drawback of the network intelligence platforms. The client – server ,as already pointed out, is becoming a sort of front-end for simplifying the interaction of the users with the bulk of the platform.

A viable User Centric approach. User centricity will be a main trends for the year to come. As clearly put forward by (Benkler 2006): "*the move to a communications environment built on cheap processors with high computation capabilities, interconnected in a pervasive network—the phenomenon we associate with the Internet. It is this … shift that allows for an increasing role for nonmarket production in the information and cultural production sector, organized in a radically more decentralized pattern than was true of this sector in the twentieth century*." In addition (Benkler 2006) states that "*The networked information economy improves the practical capacities of individuals along three dimensions: (1) it improves their capacity to do more for and by themselves; (2) it enhances their capacity to do more in loose commonality with others, without being constrained to organize their relationship through a price system or in traditional hierarchical models of social and economic organization; and (3) it improves the capacity of individuals to do more in formal organizations that operate outside the market sphere*". This autonomy of the user is rapidly moving, thanks to the evolution of the ICT technologies, to the networks and the communities, enabling smart communities environments and will soon reach the capability to produce goods (e.g., printers of objects). Finally Benkler states that "*If the transformation I describe as possible occurs, it will lead to substantial redistribution of power and money from the twentieth-century industrial producers of information, culture, and communications—like Hollywood, the recording industry, and perhaps the broadcasters and some of the telecommunications services giants—to a combination of widely diffuse populations around the globe, and the market actors that will build the tools that make this population better able to produce its own information environment rather than buying it ready-made. None of the industrial giants of yore are taking this reallocation lying down. The technology will not overcome their resistance through an insurmountable progressive impulse. The reorganization of production and the advances it can bring in freedom and justice will emerge, therefore, only as a result of social and political action aimed at* *protecting the new social patterns from the incumbents' assaults. It is precisely to develop an understanding of what is at stake and why it is worth fighting for that I write this book*". In other terms, the User Centric approach is a sort of natural right of billions of users, but the incumbent giants controlling the straits of information and technology will try to reduce to a minimum the autonomy of the users. They will be capable of impose their approach only by means of awareness and sharing of information, resources and ideas.

The Value of the Network. Networks will radically change in the future, virtualization and software defined networking will be two main drives for this revolution, in addition if cognitive radio solutions will be enabled by technology and regulation then the barrier costs of the infrastructure will considerably drop. Under this perspective, the programmability of these solutions comes with a price that some Operators are not willing to pay for. However other actors can see this as an opportunity for building competing infrastructures. The costs of the networks will aggregate in the edges, i.e., in the systems needed to offer the access to the transport capability to users. Access networks are and will be the most expensive part of networks and they will remain the major value of the Operators. As discussed in the Disappearing TelCo scenario (see chapter 8), users can collectively join and create alternative network for satisfying local communication needs. The possibility of the public network to be a sort of anchor point allowing the exchange of information between different smart environment far away will remain and important functions that users will deem valuable and will pay for. The revenue generated by access networks, however, will decline over time and will reach a threshold close to a commodity. The value of the core network instead will rapidly decrease together with the value of its intelligence. Transport of bulks of data will be economical and affordable. The TelCos will try to move the network intelligence toward the edge (in the access part of the network), for instance providing functionalities to self-organize the access resources as in the case of Self Organizing Networks of 3GPP (Hu, et al. 2010), (Feng and Seidel 2008).

Transformations of the telecoms ecosystem. The turmoil of the telecom ecosystem is manifest and it has been described in chapter 6 and chapter 7: there is a huge impact on the numbers of the industry (less revenues, less employees, less companies). The traditional business models (subscription and pay per use) are challenged by new offers, in fact services that before were only offered by Operators are now provided (often for free) by several companies like Skype, Facebook and Google. TelCos are usually irrelevant in many of the new services (e.g., social environments) and are struggling to position themselves in the cloud computing markets (finding a viable niche essentially in the Business Market). These are evidences of a difficulty in finding appropriate strategies for revamping the traditional business. In addition, the technological evolution brings in new possibilities for the competing companies. The capability to dominate and master the software development will be determinant for the success of future ventures. TelCos are not strong enough in this area, because they have reduced the expenditure in research and innovation, and have approached the technological evolution with the "intelligent buyer" approach. Technologies have been nurtured and cultivated only in order to propose favorable solutions in standards. In order to cope with the evolution it is mandatory to be capable to develop advanced solutions. This capability is clearly in the hands of competitors (Web and IT companies) while traditional vendors are trying to preserve their investments in traditional approaches and they induce the TelCos along old paths and systems. Under these conditions, it is hard to enter into new markets domains. So TelCos are doomed to a long decline towards the commoditization. They will pass from a Traditional TelCo scenario (based on the network Intelligence) to a BitCarrier Scenario. The possibility for a TelCo to play a major role as described in the Service Provider scenario are very low and essentially confined into the business market realm. The TelCo as a Platform Provider Scenario is feasible if the Operators increase their capability to deal with software. It is impossible to compete with innovative companies that address new market opportunity using software as an enabler. TelCos are still requesting established Vendors to provide general solutions. This increases the value for the Vendor, but it diminishes the possibility to compete for each single customer (the common infrastructure becomes irrelevant from the differentiation point of view and competition is played at the price level). Web Companies are more prone to build and master their software platforms, the winning ones have an edge on the competition and become "standard de facto" for many developers that flock to develop new services on the enabling infrastructure.

Is this a revolution or a consolidation of the market ? Many signals (the raising of numerous paradoxes, the impossibility to keep the business at the current level, the technological advantage of Web companies) indicate that the telecoms industry is at a cross road. While the TelCos maintain their approach of "buyers", the web companies pushes on the pedal of innovation (they are "doers") in order to displace the competitors. Software is a major differentiator factor. Only the companies that are capable to dominate the software development of huge platform can try to have a winning service strategy. Voice services are more and more offered by means of software solutions. Movies, films, music are all sharable over the network and users are not capable of providing user generated content with high definition devices. In the future the networks will be put in place by means of general purpose systems plus open source software. Objects are not designed by software systems and then printed by means of printers of objects. Physibles (i.e., objects that can be downloaded) have been created by the Pirate Bay as an example of how the networks can now be used to produce and share physical objects. Some printers, such as the RepRap (Jones, et al. 2011), are designed in such a way that they can be used to print a large part of their own components. This recall the pioneering work of Von Neumann about Self-Replication Machines and their implementations (Pesavento 1995). This works are strongly based on the work of A. Turing about the Universal Machine (Turing 1936). Since any aspect of the daily life is governed by software and this will occur even more in the future, then the prediction that “software will eat everything” can found some background also on classical theoretical works.

9. 4. Future Work Directions

The findings of this thesis should be corroborated by extensive developments and deployments. Only proof of concepts, demonstrators, prototypes and simulators of the proposed architectures have been developed, e.g., in European projects like CASCADAS (Manzalini 2006), or Univerself (Tsagkaris, et al. 2011) or internally to Telecom Italia. There is the need to enter into a virtual cycle of developments that can lead to a verification on the field of the advocated advantages of the approaches. Nevertheless some of the concepts are emerging either at the academic and industrial level (for instance the developments in the field of Software defined networks) or at the standardization level (e.g., the Network Virtualization Function initiative of ETSI, to whom the author has actively participated).

One important topic to be further analyzed and experimented is related to User Centricity. It would be interesting to try in the field the attitude of people to share their resources and data in favor of communities. There are plenty of studies in this direction (e.g., (Ohtsuki, et al. 2006)) that try to determine when an altruistic behavior is triggered in human being. In order to create a powerful architecture at the edge of the TelCos infrastructures, there is the need to determine the validity of these behavior and how much people are willing to share for creating a free and “democratic” platforms.

Another interesting topic to investigate is the economic value proposition of a user Centric approach and the viability of the TelCo as a Platform Provider. This work is needed in order to determine to what extent the approach can create a return of investment consistent with the expectations and capable of repaying the OPEX and CAPEX expenditures needed to develop it.

One central problem is the need to verify internally to the TelCos the acceptance of these scenarios and the possibility for them to try to implement solutions similar to those proposed in this section. Actually some concepts of this thesis have be used to provoke discussion internally to telecoms companies about possible strategies and approaches for the future. The typical reactions are two: a fierce opposition to these proposition (usually this position comes from middle management) or an attentive evaluation of the scenarios and their consequences (typically from high level management, whose general comments refer to the time frame of the commoditization of the Operators[51]). Respect to the architectural propositions (represented by the TelCos as a Platform Provider, and the Disappearing of the TelCos) there are different positioning within TelCos: many people consider the role of service and platform provider as interchangeable (one of the conclusion of this thesis is that they are not) and hence they promote the service proposition[52], i.e. the creation of a vast service portfolio comprising many services that generate “some revenue”; other people consider the Disappearing TelCo Scenario as totally unrealistic; a minority of people see the Disappearing of TelCo and the BitCarrier Scenarios as strongly related and likely; just a few TelCo people believe that the Traditional TelCo Scenario (and the implicated IMS leverage) is a winning proposition.

With respect to new classes of services, it would be extremely interesting to start developments of the Virtual Continuum. The author is currently involved in the definition of a viable architecture for the leverage of the Italian Cultural Heritage based on the concept of Virtual Continuum and the exploitation of the servitization business model. This is a promising development area that could look at several dimensions:

- The digitalization of the Cultural Heritage and its artifacts

[51] Some Managers believe that a ten years span period is too long and the evidence of some of the predictions will occur before.

[52] A current view of some TelCos (e.g., Telefonica Digital could be an example) is that they should try to provide as many as possible “small services” in a OTT fashion in order to increase the revenue generated by services. This approach has been criticized by this thesis because often the focalization on single services increases the costs of platform developments. The TelCo should focus on classes of services and the enabling platforms.

- The semantic representation of relationships between the artifacts and their history. In addition further relationship could be created between objects apparently loosely related.
- The definition of new interfaces and way to offer the cultural artifacts to non-expert people by means of new interaction technologies (from augmented reality to 3D printing). This should aim at a better fruition of the cultural artifacts.
- The creation of services and applications that exploit the cultural heritage within the relationships of an ecosystem of actors. In order to be succesfull, the ecosystem should conjugate tourisms operators, city and regional governments, museums and cultural institutions, creating a virtual organization aiming at the leverage of culture.
- The possibility to stimulate creativeness and the development of new works of art by means of ICT technologies. In this case, virtual reality, new graphical interfaces but also the possibility to quickly prototype objects could be offered to artists, designers and prosumers.

The latter point is addressing a new challenge: the possibility to create new physical objects by means of ICT technologies and by mashing up existing features and components. This possibility is opening a whole wealth of research and industrial initiatives that can find an important ally in the Do it yourself (DIY) movement. This could be a means to implement User Centric approaches in the area of production of goods. It could be a means to challenge the current trend of moving factories and production facilities in the developing countries. This trend could be inverted in favor of bringing production and ideas in the garage of users.

Bibliography

- Acker, Olaf, Hans and Geerdes, Germar and Gröne, and Florian Schröder. *The Global ICT 50: The Supply Side of Digitization.* Booz & Co Briefing, Booz&co, 2012.
- Adeyeye, Michael, and Paul Gardner-Stephen. "The Village Telco project: a reliable and practical wireless mesh telephony infrastructure." *EURASIP Journal on Wireless Communications and Networking* (Springer) 2001, no. 1 (2011): 1-11.
- Adler, Richard M. "Distributed coordination models for client/server computing." *Computer* (IEEE) 28, no. 4 (1996): 14-22.
- Agrawal, Prathima, Jui-Hung Yeh, Jyh-Cheng Chen, and Tao Zhang. "IP multimedia subsystems in 3GPP and 3GPP2: overview and scalability issues." *IEEE Communications Magazine* (IEEE) 46, no. 1 (2008).
- Aked, Symon. "An investigation into darknets and the content available via anonymous peer-to-peer file sharing." *9th Australian Information Security Management Conference.* Perth, Australia: Edith Cowan University, 2011.
- Andreessen, Marc. "Why Software Is Eating The World." *Wall Street Journal*, 2011: available at http://online.wsj.com/article/SB10001424053111903480904576512250915629460.html Last accessed May 2013.
- Aoyama, Tomonori. "A new generation network: beyond the internet and NGN." *IEEE Communications Magazine* (IEEE) 47, no. 5 (2009).
- API, wikipedia. *Application Programming Interface.* available at https://en.wikipedia.org/wiki/Application_programming_interface Last accessed may 2013, wikipedia, 2013.
- Armbrust, Michael, et al. "A view of cloud computing." *Communications of the ACM* (ACM) 53, no. 4 (2010): 50-58.
- Auer, Soren, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. "Dbpedia: A nucleus for a web of open data." *The semantic web.* Springer, 2007. 722-735.
- Availability Digest. *What is Active/Active?* White Paper, The Availability Digest, 2006.

- Baglietto, Pierpaolo, Massimo Maresca, Michele Stecca, and Corrado Moiso. "Towards a CAPEX-free service delivery platform." *16th International Conference onIntelligence in Next Generation Networks (ICIN).* Berlin: IEEE, 2012.
- Bahl, Paramvir, Richard Y Han, Li Erran Li, and Mahadev Satyanarayanan. "Advancing the state of mobile cloud computing." *Proceedings of the third ACM workshop on Mobile cloud computing and services.* Low Wood Bay, Lake District, United Kingdom: ACM, 2012. 21 - 28.
- Baines, Tim S, Howard W Lightfoot, Ornella Benedettini, and John M Kay. "The servitization of manufacturing: a review of literature and reflection on future challenges." *Journal of Manufacturing Technology Management* (Emerald Group Publishing Limited) 20, no. 5 (2009): 547 - 567.
- Baldauf, Matthias, Schahram Dustdar, and Florian Rosenberg. "A survey on context-aware systems." *International Journal of Ad Hoc and Ubiquitous Computing* (Inderscience) 2, no. 4 (2007): 263-277.
- Barber, Simon, Xavier Boyen, Elaine Shi, and Ersin Uzun. "Bitter to Better—How to Make Bitcoin a Better Currency." Edited by Angelos D. Keromytis. *16th Conference on Financial Cryptography and Data Security.* Kralendijk, Bonaire: Springer, 2012. 399-414.
- Barham, Paul, et al. "Xen and the art of virtualization." *ACM SIGOPS Operating Systems Review* (ACM) 37, no. 5 (2003): 164-177 .
- Baroncelli, Fabio, Barbara Martini, L Valcarenghi, and P Castoldi. "A service oriented network architecture suitable for global grid computing." *Conference on Optical Network Design and Modeling.* Milan, Italy: IEEE - IP TC6, 2005. 283 - 293.
- Barroso, Luiz Andrè, Jeffrey Dean, and Urs Holzle. "Web search for a planet: The Google cluster architecture." *IEEE Micro* (IEEE Computer Society Press) 23, no. 2 (2003): 22-28.
- Baset, Salman A, and Henning Schulzrinne. "An analysis of the skype peer-to-peer internet telephony protocol." *INFOCOM 2006. 25th IEEE International Conference on Computer Communications.* Barcelona, Spain: IEEE, 2006. 1 - 11.
- Bastos, Luis, and Hermann Wietgrefe. "Challenges and proposals for WiMAX-based military hybrid communications networks." *MILITARY COMMUNICATIONS CONFERENCE, 2010-MILCOM 2010.* San Jose, USA: IEEE, 2010. 2020-2025.
- Battistella, Cinzia, Katia Colucci, and Fabio Nonino. "Methodology of Business Ecosystems Network Analysis: A Field Study in Telecom Italia Future Centre." In *Information Systems: Crossroads for Organization, Management, Accounting and Engineering*, by Marco De Marco, Dov Te'eni, Valentina Albano, & Stefano Za, 239-249. Heidelberg, Berlin: Springer, 2012.

➢ Battiti, Roberto, Renato Lo Cigno, Fredrik Orava, and Bjorn Pehrson. "Global growth of open access networks: from warchalking and connection sharing to sustainable business." *Proceedings of the 1st ACM international workshop on Wireless mobile applications and services on WLAN hotspots.* New York, NY, USA: ACM, 2003. 19-28.

➢ Benkler, Yochai. "The wealth of networks: How social production transforms markets and freedom." Yale University Press, 2006.

➢ Berndt, Hendrik and Darmois, E and Dupuy, F and Inoue, Y and Lapierre, M and Minerva, R and Minetti, R and Mossotto, C and Mulder, H and Natarajan, N and others. *The TINA Book: A Co-operative Solution for a Competitive World.* Edited by Yugi Inoue, Martine Lapierre, & Cesare Mossotto. Prentice Hall, 1999.

➢ Berndt, Hendrik, and Roberto Minerva. *Definition of Service Architecture.* Deliverable, Red Bank, N.J.: TINA Consortium, 1995.

➢ Bertin, Emmanuel, and Noel Crespi. *Architecture and Governance for Communication Services.* London, UK: Wiley-ISTE, 2013.

➢ Bettstetter, Christian, H-J Vogel, and Jorg Eberspacher. "GSM phase 2+ general packet radio service GPRS: Architecture, protocols, and air interface." *IEEE Communications Surveys & Tutorials* (IEEE) 2, no. 3 (1999): 2 - 14 .

➢ Bhagwan, Ranjita, Stefan Savage, and Geoffrey M Voelker. "Understanding availability." *Peer-to-Peer Systems II.* Springer, 2003. 256-267.

➢ Bifet, Albert, and Eibe Frank. "Sentiment knowledge discovery in twitter streaming data." Edited by Bernhard Pfahringer, Geoff Holmes, & Achim Hoffmann. *13th International Conference on Discovery Science.* Canberra, Australia: Springer, 2010. 1 - 15.

➢ Blum, Niklas, Jens Fiedler, Lajos Lange, and Thomas Magedanz. "Application-driven Quality of Service for M2M communications." *Intelligence in Next Generation Networks (ICIN), 2011 15th International Conference on.* Berlin: ICIN, 2011. 41 - 45.

➢ Bohli, Jens-Matthias, Christoph Sorge, and Dirk Westhoff. "Initial observations on economics, pricing, and penetration of the internet of things market." *ACM SIGCOMM Computer Communication Review* (ACM) 39, no. 2 (2009): 50-55.

➢ Booth, David, et al. *Web Services Architecture.* W3C Working Group Note, W3C, W3C, 2004.

➢ Box, Don, et al. *Simple object access protocol (SOAP) 1.1.* Standard, W3C, 2000.

- Brantner, Matthias, Daniela Florescu, David Graf, Donald Kossmann, and Tim Kraska. "Building a database on S3." *Proceedings of the 2008 ACM SIGMOD international conference on Management of data.* ACM, 2008. 251-264.
- Brewer, Eric. "Pushing the CAP: strategies for consistency and availability." *Computer* (IEEE) 45, no. 2 (2012): 23-29.
- —. "Towards robust distributed systems." *Proceedings of the Annual ACM Symposium on Principles of Distributed Computing.* Portland: ACM, 2000. 7.
- Buchmann, Alejandro and Koldehofe, Boris. "Complex event processing." *it-Information Technology* (Oldenbourg Wissenschaftsverlag GmbH) 51, no. 5 (2009): 241 - 242.
- Burgess, David A, Harvind S Samra, and others. *The OpenBTS Project.* Report available at http://openbts.sourceforge.net, OpenBTS.org, 2008 .
- Burrows, Mike. "The Chubby lock service for loosely-coupled distributed systems." *Proceedings of the 7th symposium on Operating systems design and implementation.* Berkeley: USENIX Association, 2006. 335-350.
- Buyya, Rajkumar, Rajiv Ranjan, and Rodrigo N Calheiros. "Intercloud: Utility-oriented federation of cloud computing environments for scaling of application services." *ICA3PP'10 Proceedings of the 10th international conference on Algorithms and Architectures for Parallel Processing* . Springer, 2010. 13-31.
- Cachin, Christian, Rachid Guerraoui, and Luâis Rodrigues. *Introduction to reliable and secure distributed programming.* Berlin Heidelberg: Springer verlag, 2011.
- Calhoun, Pat, John Loughney, Erik Guttman, Glen Zorn, and Jari Arkko. *Diameter base protocol.* RFC 3588, Fremont California, USA: IETF, 2003.
- Camarillo, Gonzalo, and Miguel-Angel Garcia-Martin. *The 3G IP multimedia subsystem (IMS): merging the Internet and the cellular worlds.* New York: John Wiley and SOns, 2007.
- Cameron, Kim. *The laws of identity.* Redmond, USA, Microsoft, 2005.
- Campbell, Andrew T., Herman G. De Meer, Michael E. Kounavis, Kazuho Miki, John B. Vicente, and Daniel Villela. "A survey of programmable networks." *SIGCOMM Comput. Commun. Rev.* (ACM) 29, no. 2 (1999): 7-23.
- Cardoso, Jorge, Konrad Voigt, and Matthias Winkler. *Service Engineering for the Internet of Services.* Vol. 19, in *Enterprise Information Systems - Lecture Notes in Business Information Processing*, by

Joaquim Filipe, & José Cordeiro, 15-27. Berlin Heidelberg: Springer Lecture Notes in Business Information Processing, 2009.

- Cattell, Rick. "Scalable SQL and NoSQL data stores." *ACM SIGMOD Record* (ACM) 39, no. 4 (2011): 12-27.
- Chan, W and Mauborgne, Renee. *Blue ocean strategy: how to create uncontested market space and make competition irrelevant.* Harvard, USA: Harvard Business Press, 2005.
- Chang, Fay, et al. "Bigtable: A distributed storage system for structured data." *ACM Transactions on Computer Systems (TOCS)* 26, no. 2 (2008): article n 4.
- Chaparadza, Ranganai, et al. "ETSI Industry Specification Group on Autonomic Network Engineering for the Self-managing Future Internet (ETSI ISG AFI)." *Web Information Systems Engineering-WISE 2009.* Berlin Heidelberg: Springer, 2009. 61 - 62.
- Chetan Sharma. *Operator's Dilemma (And Opportunity): The 4th Wave.* Issaquah, WA, USA : Chetan Sharma Consulting, 2012.
- Chinnici, Roberto, Jean-Jacques Moreau, Arthur Ryman, and Sanjiva Weerawarana. *Web services description language (wsdl) version 2.0 part 1: Core language.* Recommendation, Boston - Geneva: W3C, 2007.
- Chowdhury, NM, and Raouf Boutaba. "A survey of network virtualization." *Computer Networks* (Elsevier) 54, no. 5 (2010): 862-876.
- Chun, Byung-Gon, and Petros Maniatis. "Augmented smartphone applications through clone cloud execution." *Proceedings of the 12th conference on Hot topics in operating systems.* Berkeley, USA: USENIX Association, 2009. 8.
- Chun, Byung-Gon, Sunghwan Ihm, Petros Maniatis, Mayur Naik, and Ashwin Patti. "Clonecloud: elastic execution between mobile device and cloud." *Proceedings of the sixth conference on Computer systems.* ACM, 2011.
- Chung, P Emerald, et al. "DCOM and CORBA side by side, step by step, and layer by layer." *C++ Report* 10, no. 1 (1998): 18-29.
- CISCO. *Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2011-2016.* White Paper, San Jose, Ca, USA: CISCO, 2012.

- Clark, David D, Craig Partridge, J. Christopher Ramming, and John T. Wroclawski. "A knowledge plane for the internet." *SIGCOMM '03 Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications.* New York: ACM, 2003. 3-10.
- Clark, David D., John Wroclawski, Karen R. Sollins, and Robert Braden. "Tussle in cyberspace: defining tomorrow's internet." *IEEE/ACM Transactions on Networking (TON)* (IEEE/ACM) 13, no. 3 (2005): 462 - 475.
- Clement, Luc, Andrew Hately, Claus Rogers, Tony von Riegen, and and others. *UDDI version 3.0. 2.* Specification technical committee draft, UDDI Org, 2004.
- Copeland, Rebecca. *Converging NGN wireline and mobile 3G networks with IMS.* London: CRC Press, 2009.
- Cucinotta, T, et al. "Virtualised e-learning with real-time guarantees on the irmos platform." *IEEE International Conference on Service-Oriented Computing and Applications (SOCA).* Perth, WA, Australia: IEEE, 2010. 1 - 8.
- Curbera, Francisco, Frank Leymann, Tony Storey, Donald Ferguson, and Sanjiva Weerawarana. *Web services platform architecture: SOAP, WSDL, WS-policy, WS-addressing, WS-BPEL, WS-reliable messaging and more.* Englewood Cliffs: Prentice Hall PTR, 2005.
- Curbera, Francisco, Matthew Duftler, Rania Khalaf, William Nagy, Nirmal Mukhi, and Sanjiva Weerawarana. "Unraveling the Web services web: an introduction to SOAP, WSDL, and UDDI." *IEEE Internet Computing* (IEEE) 6, no. 2 (2002): 86-93.
- Dabek, Frank, Nickolai Zeldovich, Frans Kaashoek, David Mazières, and Robert Morris. "Event-driven programming for robust software." *Proceedings of the 10th workshop on ACM SIGOPS European workshop.* Saint-Emilion, France: ACM, 2002. 186-189.
- Datla, D., et al. "Wireless distributed computing: a survey of research challenges." *IEEE Communications Magazine* (IEEE ComSoc) 50, no. 1 (2012): 144-152.
- Dean, Jeffrey, and Luiz André Barroso. "The Tail at Scale." Edited by ACM. *Communications of the ACM* 56 (2013): 74-80.
- Dean, Jeffrey, and Sanjay Ghemawat. "MapReduce: a flexible data processing tool." *Communications of the ACM* (ACM) 53, no. 1 (2010): 72-77.
- Dean, Jeffrey, and Sanjay Ghemawat. "MapReduce: simplified data processing on large clusters." *Communications of the ACM* (ACM) 51, no. 1 (2008): 107-113.

- DeCandia, Giuseppe, et al. "Dynamo: amazon's highly available key-value store." *ACM Symposium on Operating Systems Principles: Proceedings of twenty-first ACM SIGOPS symposium on Operating systems principles.* ACM, 2007. 205-220.
- Dedecker, Jessie, Tom Van Cutsem, Stijn Mostinckx, Theo D'Hondt, and Wolfgang De Meuter. "Ambient-oriented programming in ambienttalk." *ECOOP 2006--Object-Oriented Programming.* Nantes, France: Springer, 2006. 230 - 254.
- Delaney, Ben, T and Catarci, and TDC Little. "The power of P2P." *IEEE MultiMedia* (IEEE) 8, no. 2 (2001): 100-103.
- Delta Partner Group A. *TELECOM AT A CROSSROADS: RE-INVENT YOURSELF NOW OR RISK LOSING HALF YOUR CASH FLOW.* white paper, Barcelona, Spain: Delta Partner Group, 2012.
- Deussen, Peter H, Edzard Hofig, Matthias Baumgarten, Maurice Mulvenna, Antonio Manzalini, and Corrado Moiso. "Component-ware for Autonomic Supervision Services." Edited by IARIA. *International Journal On Advances in Intelligent Systems* 3, no. 1-2 (2010): 87 - 105.
- Deutsch, Peter . *Fallacies of Distributed Computing.* WHITE PAPER, wikipedia, 1995.
- Di Caprio, Gaetano, and Corrado Moiso. "Web Services and Parlay: an architectural comparison." *Proc. ICIN.* Bordeaux, France: ICIN, 2003. 1 - 6.
- Di Costanzo, Alexandre, Marcos Dias De Assuncao, and Rajkumar Buyya. "Harnessing cloud technologies for a virtualized distributed computing infrastructure." *IEEE Internet Computing* (IEEE) 13, no. 5 (2009): 24-33.
- Dillon, Tharam, Chen Wu, and Elizabeth Chang. "Cloud computing: issues and challenges." *24th IEEE International Conference on Advanced Information Networking and Applications (AINA).* Perth, Australia: IEEE, 2010. 27 - 33.
- DrPeering. *Internet Transit Prices - Historical and Projected.* white paper, San Meteo, USA: DrPeering, 2010.
- Dunkels, Adam, Juan Alonso, and Thiemo Voigt. *Making TCP/IP viable for wireless sensor networks.* SICS Research Report, Stockholm: Swedish Institute of Computer Science, 2003.
- El-Darwiche, Bahjat, Roman Friedrich, Pierre Péladeau, and Karim Sabbagh. *The Future of Telecom Operators - Capabilities for Rapid Change.* White paper, New York: Booz & C, 2010.

- Eriksson, Mats, Veli-Pekka Niitamo, and Seija Kulkki. *State-of-the-art in utilizing Living Labs approach to user-centric ICT innovation-a European approach.* White Paper, Lulea: Lulea University of Technology Sweden, 2005.
- Fainelli, Florian. "The OpenWrt embedded development framework." *Proceedings of the Free and Open Source Software Developers European Meeting.* Brussles: FOSDEM, 2008. 1.
- Faynberg, Igor, Nitin J Shah, Lawrence R Gabuzda, and Marc P Kaplan. *The Intelligent Network Standards: their application to services.* New York: McGraw-Hill Professional, 1996.
- Feng, Sujuan, and Eiko Seidel. *Self-organizing networks (SON) in 3GPP long term evolution.* White Paper, Munich, Germany: Nomor Research GmbH, 2008.
- Ferrari, Laura, Antonio Manzalini, Corrado Moiso, and Peter H Deussen. "Highly distributed supervision for autonomic networks and services." *Fifth Advanced International Conference on Telecommunications, 2009. AICT'09.* Venice, Italy: IEEE, 2009. 111 - 116.
- Fette, Ian, and Alexey Melnikov. *The websocket protocol.* RFC 6455, Boston - Geneva: W3C, 2011.
- Fielding, Roy T., and Richard N, Taylor. "Principled design of the modern Web architecture." *ACM Transactions on Internet Technology (TOIT)* (ACM) 2, no. 2 (2002): 115-150.
- Fitzsimmons, James A., and Mona J. Fitzsimmons. *Service Management: Operations,Strategy, and Information Technology.* New York: McGraw-Hill, 2003.
- Follows, Jonathan, and Deflet Straeten. *Application driven networking: Concepts and architecture for policy-based systems.* Red Books. White Planes, USA: IBM Corporation, 1999.
- Fontoura, Alexander, Shraer Maxim, Gurevich Marcus, and Vanja Josifovski. "Top-k Publish-Subscribe for Social Annotation of News." *Proceedings of the VLDB Endowment.* VLDB , 2013. to appear.
- Foster, Ian. *Designing and building parallel programs.* Boston: Addison-Wesley Reading, 1995.
- Foster, Ian, Yong Zhao, Ioan Raicu, and Shiyong Lu. "Cloud computing and grid computing 360-degree compared." *Grid Computing Environments Workshop.* Austin, TX, USA: IEEE, 2008. 1 - 10.
- Fu, Jicheng, Wei Hao, Michael Tu, Biao Ma, Joshua Baldwin, and Farokh B Bastani. "Virtual services in cloud computing." *6th World Congress on Services (SERVICES-1).* Miami, FL, USA: IEEE, 2010. 467 - 472.
- Ganti, Raghu K and Ye, Fan and Lei, Hui. "Mobile crowdsensing: Current state and future challenges." *IEEE Communications Magazine* (IEEE) 49, no. 11 (2011): 32 - 39 .

- Gavras, Anastasius, Arto Karila, Serge Fdida, Martin May, and Martin Potts. "Future internet research and experimentation: the FIRE initiative." *ACM SIGCOMM Computer Communication Review* (ACM) 37, no. 3 (2007): 89 - 92.
- Geambasu, Roxana, Tadayoshi Kohno, Amit Levy, and Henry M Levy. "Vanish: Increasing data privacy with self-destructing data." *Proc. of the 18th USENIX Security Symposium.* USENIX, 2009. 299-316.
- Ghemawat, Sanjay, Howard Gobioff, and Shun-Tak Leung. "The Google file system." *ACM SIGOPS Operating Systems Review* (ACM) 37, no. 5 (2003): 29-43.
- Goncalves, Vania, Nils Walravens, and Pieter Ballon. "How about an App Store?" Enablers and Constraints in Platform Strategies for Mobile Network Operators." *Ninth International Conference on Mobile Business and 2010 Ninth Global Mobility Roundtable (ICMB-GMR).* Athens, Greece: IEEE, 2010. 66 - 73.
- Gong, Yadong, Zongquan Ying, and Meihong Lin. "A Survey of Cloud Computing." *Proceedings of the 2nd International Conference on Green Communications and Networks 2012 (GCN 2012).* Gandia, Spain: Springer, 2012. 79-84.
- González, Marta C., Cesar Hidalgo, and Albert-Laszlo Barabasi. "Understanding individual human mobility patterns." *Nature* (Nature) 453, no. 7196 (2008): 779–782.
- Gradecki, Joseph D. *Mastering Jxta: Building Java Peer-to-Peer Applications.* New York: John Wiley \& Sons, Inc., 2002.
- Greenberg, Albert, James Hamilton, David A Maltz, and Parveen Patel. "The cost of a cloud: research problems in data center networks." *ACM SIGCOMM Computer Communication Review* (ACM) 39, no. 1 (2008): 68-73.
- Grelck, Clemens, Sven-Bodo Scholz, and Alex Shafarenko. "Asynchronous stream processing with S-Net." *International Journal of Parallel Programming* (Springer) 38, no. 1 (2010): 38-67.
- GSM Association. *Home-3rd Party Access Project-OneAPI.* White Paper, London: GSM Association, 2010.
- GSM Association. *Working Group:"Rich Communication Suite (RCS) & Rich Communications Ecosystem (RCE).* White paper, London, UK: GSMA, 2009.
- Gude, Natasha, et al. "NOX: towards an operating system for networks." *ACM SIGCOMM Computer Communication Review* (ACM) 38, no. 3 (2008): 105-110.

- Gunarathne, Thilina, Tak-Lon Wu, Judy Qiu, and Geoffrey Fox. "MapReduce in the Clouds for Science." *IEEE Second International Conference on Cloud Computing Technology and Science (CloudCom).* Indianapolis, IN, USA: IEEE, 2010. 565 - 572.
- Gurbani, V, Volker Hilt, Ivica Rimac, Marco Tomsu, and Enrico Marocco. "A survey of research on the application-layer traffic optimization problem and the need for layer cooperation." *IEEE Communications Magazine* (IEEE) 47, no. 8 (2009): 107-112.
- He, W., and L. Xu. "Integration of Distributed Enterprise Applications: A Survey." *IEEE Transactions on Industrial Informatics* (IEEE) PP, no. 99 (2012): 1.
- Henry, Katell, Qunkai Liu, and Severin Pasquereau. "Rich Communication Suite." *13th International Conference on Intelligence in Next Generation Networks, 2009. ICIN 2009.* Bordeaux: ICIN, 2009. 1 - 6.
- Herrmann, Michael, and Christian Grothoff. "Privacy-implications of performance-based peer selection by onion-routers: a real-world case study using I2P." *PETS'11 11th international conference on Privacy enhancing technologies.* Waterloo, ON, Canada: Springer, 2011. 155-174.
- Hey, Anthony JG and Tansley, Stewart and Tolle, Kristin Michele and others. "The fourth paradigm: data-intensive scientific discovery." *Proceedings of the IEEE* (IEEE - Microsoft) 99, no. 8 (August 2011): 1334 - 1337.
- Hey, Anthony JG, Stewart Tansley, Kristin Michel Tolle, and and others. *The fourth paradigm: data-intensive scientific discovery.* Book, Redmont: Microsoft Research, 2009.
- Hickson, Ian. *The websocket api.* Working Draft WD - websockets, Boston - Geneva: W3C, 2011.
- Holland, Gary. "A Changing Market Environment Demands a High Leverage Network." *TechZine*, January 29th, 2013.
- Hsu, Paul P.E., and Veerendra Mulay. *Open Compute Project--Server and Data Center Design.* Panel presentation, Silicon Valley Leadership Group SVLG, 2011.
- Hu, Honglin, Jian Zhang, Xiaoying Zheng, Yang Yang, and Ping Wu. "Self-configuration and self-optimization for LTE networks." *Communications Magazine* (IEEE) 48, no. 2 (2010): 94 - 100.
- Hu, Shun-Yun, and Guan-Ming Liao. "Scalable peer-to-peer networked virtual environment." *Proceedings of 3rd ACM SIGCOMM workshop on Network and system support for games.* Portland, OR, USA: ACM, 2004. 129-133.
- iCore. *Functional Architecture.* Deliverable D2.3, Brussels: iCore Consortium, 2013.

- IOT-A. *Initial Architectural Reference Model for IoT.* Project Deliverable D1.2, Brussles: IOT-A Consortium, 2011.
- Isenberg , David I. "The dawn of the “stupid network”." *netWorker*, March 1998: 24-31.
- Itoh, Mitsutaka, Eric Y Chen, and Tetsuya Kusumoto. "Virtual Smartphone over IP." *NTT Technical Review* (NTT) 8, no. 7 (2010): 1 - 5.
- ITU. "ITU and its Activities Related to Internet Protocol (IP) Networks." *http://www.itu.int.* Edited by ITU. April 4th, 2011. http://www.itu.int/osg/spu/ip/glossary.html (accessed May Last accessed 26th, 2013).
- Izal, Mikel, Guillaume Urvoy-Keller, Ernst W Biersack, Pascal A Felber, Anwar Al Hamra, and Luis Garces-Erice. *Dissecting bittorrent: Five months in a torrent’s lifetime.* Vol. 3015, in *Passive and Active Network Measurement - Lecture Notes in Computer Science*, by Chadi and Pratt, Ian Barakat, 1 - 11. Berlin Heidelberg: Springer, 2004.
- Jacobson, Van, Diana K Smetters, Nicholas H: Plass, Michael F Briggs, Paul Stewart, James D Thornton, and Rebecca L Braynard. "VoCCN: voice-over content-centric networks." *Proceedings of the 2009 workshop on Re-architecting the internet.* Rome, Italy: ACM, 2009. 1 - 6.
- Jain, Menuka, and Maria Prokopi. "The IMS 2.0 service architecture." *The Second International Conference on Next Generation Mobile Applications, Services and Technologies, 2008. NGMAST'08.* Cardiff, Wales: IEEE Computer, 2008. 3-9.
- Jelasity, Mark, Alberto Montresor, and Ozalp Babaoglu. "Gossip-based aggregation in large dynamic networks." *ACM Transactions on Computer Systems (TOCS)* (ACM) 23, no. 3 (2005): 219 - 252.
- Johnson, David, Ntsibane Ntlatlapa, and Corinna Aichele. *Simple pragmatic approach to mesh routing using BATMAN.* Report presented at Wireless Africa, Pretoria: Meraka Institute, 2008.
- Johnston, Alan B. *SIP: understanding the session initiation protocol.* London: Artech House Publishers, 2009.
- Jondral, Friedrich K. "Software-defined radio: basics and evolution to cognitive radio." *EURASIP Journal on Wireless Communications and Networking* (Hindawi Publishing Corp.) 2005, no. 1 (2005): 275-283.
- Jones, Rhys, et al. "Reprap - the replicating rapid prototyper." *Robotica* (Cambridge Univ Press) 29, no. 1 (2011): 177-191.

- Jones, Steve. "Toward an acceptable definition of service [service-oriented architecture]." *IEEE Software* 22, no. 3 (May-June 2005): 87-93.
- Katti, Sachin, Hariharan Rahul, Wenjun Hu, Dina Katabi, Muriel Médard, and Jon Crowcroft. "XORs in the air: practical wireless network coding." *IEEE/ACM Transactions on Networking (TON)* (ACM/IEEE) 16, no. 3 (2008): 497 - 510.
- Kimbler, Kristofer. "App store strategies for service providers." *14th International Conference on Intelligence in Next Generation Networks (ICIN).* Berlin: ICIN, 2010. 1 - 6.
- Kimbler, Kristofer, and Mac Taylor. "Value added mobile broadband services Innovation driven transformation of the 'smart pipe'." *16th International Conference on Intelligence in Next Generation Networks (ICIN).* Berlin: ICIN, 2012. 1 - 6.
- Knightson, Keith, Naotaka Morita, and Thomas Towle. "NGN architecture: generic principles, functional architecture, and implementation." *IEEE Communications Magazine* (IEEE) 43, no. 10 (2005): 49-56.
- Koomey, Jonathan G. "Outperforming Moore's law." *IEEE Spectrum* (IEEE) 47, no. 3 (2010): 68.
- Koponen, Teemu, et al. "Architecting for innovation." *ACM SIGCOMM Computer Communication Review* (ACM) 41, no. 3 (2011): 24-36.
- Koskela, Timo, Otso Kassinen, Erkki Harjula, and Mika Ylianttila. "P2P group management systems: A conceptual analysis." *ACM Comput. Surv.* (ACM) 45, no. 2 (February 2013): 20:1-20:25.
- Krakowiak, Sacha . "What is Middleware." *www.objectweb.com.* 2003. http://middleware.objectweb.org/ (accessed May Last accessed 26th , 2013).
- Kranz, Matthias and Roalter, Luis and Michahelles, Florian. "Things that twitter: social networks and the internet of things." *What can the Internet of Things do for the Citizen (CIoT) Workshop at The Eighth International Conference on Pervasive Computing (Pervasive 2010).* Helsinki, Finland: Helsinki Institute for Information Technology HIIT and University of Helsinki, 2010. 1 - 10.
- Krieger, David, and Richard M Adler. "The emergence of distributed component platforms." *IEEE Computer* (IEEE) 31, no. 1 (1998): 43-53.
- Krishnamurthy, Balachander, Jia Wang, and Yinglian Xie. "Early measurements of a cluster-based architecture for P2P systems." *Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement.* Burlingame, CA, USA: ACM, 2001. 105 - 109.
- Krishnamurthy, Balachander, Phillipa Gill, and Martin Arlitt. "A few chirps about twitter." *Proceedings of the first workshop on Online social networks.* Seattle, WA, USA: ACM, 2008. 19-24.

- Kubiatowicz, John David. *Integrated Shared-Memory and Message-Passing Communication in the Alewife Multiprocessor.* PhD Thesis, Boston: MIT, 1998.
- Kuhn, Thomas S. *The Structure of Scientific Revolutions.* 3rd. Chicago: University of Chicago Press, 1996.
- Kumar, Sanjay, Himanshu Raj, Karsten Schwan, and Ivan Ganev. "Re-architecting VMMs for multicore systems: The sidecore approach." *Workshop on Interaction between Opearting Systems and Computer Architecture (WIOSCA).* San Diego, California, USA: University of FLorida, USA, 2007. paper 3.
- Kumar, Vibhore, Henrique Andrade, Buugra Gedik, and Kun-Lung Wu. "DEDUCE: at the intersection of MapReduce and stream processing." *Proceedings of the 13th International Conference on Extending Database Technology.* Lausanne, Switzerland: ACM, 2010. 657-662.
- Kwon, Young-Wo, Eli Tilevich, and William R Cook. "Which middleware platform should you choose for your next remote service?" *Service Oriented Computing and Applications* (Springer) 5, no. 2 (2011): 61-70.
- Kwon, Young-Woo, Eli Tilevich, and William R Cook. "An assessment of middleware platforms for accessing remote services." *IEEE International Conference on Services Computing (SCC), 2010.* Miami, FL, USA: IEEE, 2010. 482-489.
- Lakshman, Avinash, and Prashant Malik. "Cassandra: a decentralized structured storage system." *ACM SIGOPS Operating Systems Review* (ACM) 44, no. 2 (2010): 35-40.
- Landau, Susan, and Tyler Moore. "Economic tussles in federated identity management." *Proceedings of the Tenth Workshop on the Economics of Information Security (WEIS).* George Mason University, USA, 2011.
- Lankhorst, Marc. *Enterprise architecture at work: Modelling, communication and analysis.* Berlin Heidelberg: Springer, 2013.
- Lantz, Bob, Brandon Heller, and Nick McKeown. "A network in a laptop: rapid prototyping for software-defined networks." *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks.* Monterey, CA, USA: ACM, 2010. article n 19.
- Laoutaris, Nikolaos, Pablo Rodriguez, and Laurent Massoulie. "ECHOS: edge capacity hosting overlays of nano data centers." *ACM SIGCOMM Computer Communication Review* (ACM) 38, no. 1 (2008): 51-54.

- Laplante, Phil, Robert R Hoffman, and Gary Klein. "Antipatterns in the creation of intelligent systems." *IEEE Intelligent Systems* (IEEE) 22, no. 1 (2007): 91-95.
- Laurent, Simon St, Joe Johnston, Edd Dumbill, and Dave Winer. *Programming web services with XML-RPC.* Sebastopol, CA, USA: O'Reilly Media, Incorporated, 2001.
- Leavitt, Neal. "Will NoSQL databases live up to their promise?" *IEEE Computer* (IEEE) 43, no. 2 (2010): 12 - 14.
- Lenk, Alexander, Markus Klems, Jens Nimis, Stefan Tai, and Thomas Sandholm. "What's inside the Cloud? An architectural map of the Cloud landscape." *Proceedings of the 2009 ICSE Workshop on Software Engineering Challenges of Cloud Computing.* Vancouver, Canada: IEEE, 2009. 23-31.
- Leopold, Claudia. *Parallel and Distributed Computing: A survey of models, paradigms, and approaches.* New York: Wiley, 2001.
- Licciardi, CarloA., Roberto Minerva, and Alberto Cuda. "TINA is dead, long live TINA: toward programmable solutions for Next Generation Services." *TINA Conference.* Paris: TINA_C, 2000. 16–20.
- Lin, Min, and J Arenzana Arias. "Rich Communication Suite: The challenge and opportunity for MNOs." *15th International Conference on Intelligence in Next Generation Networks (ICIN).* Berlin: ICIN, 2011. 187 - 190.
- Liu, Fang, Jin Tong, Jian Mao, Robert Bohn, John Messina, and Lee: Leaf, Dawn Badger. *NIST cloud computing reference architecture.* NIST Special Publication - 500-292, Washington, USA: NIST Special Publication, 2011, 292.
- Lo Pumo, Andrea. *Overview of the Netsukuku network.* preprint arXiv:0705.0815, arXiv , 2007, 1- 5.
- Loreto, Salvatore, and Simon Pietro Romano. "Real-Time Communications in the Web: Issues, Achievements, and Ongoing Standardization Efforts." *IEEE Internet Computing* (IEEE) 16, no. 5 (2012): 68-73.
- Loveland, Richard. "It’s Not a Dream -- It’s REAL! VDSL2 Vectoring in the Real World." *OSP Magazine*, 2012.
- Lua, Eng Keong, Jon Crowcroft, Marcelo Pias, Ravi Sharma, and Steven Lim. "A survey and comparison of peer-to-peer overlay network schemes." *IEEE Communications Surveys and Tutorials* (IEEE) 7, no. 2 (2005): 72-93.
- Maes, Stèphane H. *Next Generation Telco Service Providers: Telco 2.0 and Beyond.* Huawei White Paper, Huawei, 2010.

- magic, instructional media +. "Application programming interface." *http://www.immagic.com/.* 2012. http://www.immagic.com/eLibrary/ARCHIVES/GENERAL/WIKIPEDI/W120623A.pdf (accessed May Last accessed 26th, 2013).
- Malewicz, Grzegorz, et al. "Pregel: a system for large-scale graph processing." *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data.* Indianapolis, Indiana, USA: ACM, 2010. 135-146.
- Maloney, John, Mitchel Resnick, Natalie Rusk, Brian Silverman, and Evelyn Eastmond. "The scratch programming language and environment." *ACM Transactions on Computing Education (TOCE)* (ACM) 10, no. 4 (2010): article n. 16.
- Maly, Robin Jan, Jan Mischke, Pascal Kurtansky, and Burkhard Stiller. *Comparison of Centralized (Client-Server) and Decentralized (Peer-to-Peer) Networking.* Semester thesis, Zurich, Switzerland: ETH Zurich,, 2003, 1-12.
- Manzalini, Antonio and Zambonelli, Franco. "Towards autonomic and situation-aware communication services: the cascadas vision." *IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications.* Prague, Cekia: IEEE, 2006. 383 - 388.
- Margara, Alessandro, and Gianpaolo Cugola. "Processing flows of information: from data stream to complex event processing." *Proceedings of the 5th ACM international conference on Distributed event-based system.* New York: ACM, 2011. 359-360.
- Marinescu, Dan, and Reinhold Kroger. *State of the art in autonomic computing and virtualization.* Report, Wiesbaden: Distributed Systems Lab, Wiesbaden University of Applied Sciences, 2007.
- Marshall, Paul, Kate Keahey, and Tim Freeman. "Elastic site: Using clouds to elastically extend site resources." *CCGRID '10 Proceedings of the 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing.* Melbourne, Australia: IEEE Computer Society, 2010. 43-52.
- Mathioudakis, Michael, and Nick Koudas. "Twittermonitor: trend detection over the twitter stream." *Proceedings of the 2010 international conference on Management of data.* Indianapolis, IN, USA: ACM, 2010. 1155-1158.
- McKeown, Nick, Tom Anderson, Hari Balakrishnan, Guru Parulkar, Larry Peterson, Jennifer Rexford, Scott Shenker, and Jonathan Turner. "OpenFlow: enabling innovation in campus networks." *ACM SIGCOMM Computer Communication Review* (ACM) 38, no. 2 (2008): 69-74.
- Mell, Peter, and Timothy Grance. *The NIST definition of cloud computing (draft).* NIST Special Publication 800-145, Washington, USA: NIST, 2011.

- Merwe, Johann Van Der, Dawoud Dawoud, and Stephen McDonald. "A survey on peer-to-peer key management for mobile ad hoc networks." *ACM computing surveys (CSUR)* (ACM) 39, no. 1 (2007): article n. 1.
- Messerschmitt, David G. "The convergence of telecommunications and computing: What are the implications today ?" *Proceedings of the IEEE* (IEEE) 84, no. 8 (1996): 1167 - 1186.
- Metsch T., Edmonds A.,. *Open Cloud Computing Interface – Infrastructure.* Specification GFD.184, Lemont, Illinois, USA : Open Grid Forum, 2011.
- Metzler, Jim. *Virtualization: Benefits, Challenges, and Solutions.* White Paper, San Francisco: Riverbed Technology, 2011.
- Miller, Kevin C., Jason E. Brandwine, and Andrww j. Doane. *Using Virtual Networking Devices to manage routing communications between connected computer networks.* US Patent 7953865 B1, United States Patent Office, 2011.
- Minch, Robert. *Oracle's E-Business Suite: An N-Tier, Networked Application.* ITM305-003, Boise City, ID, USA: Boise State University, 2009.
- Minerva, Roberto. "On Some Myths about Network Intelligence." *Proceedings of International Conference on Intelligence in Networks-ICIN2008 (October2008).* Bordeaux, France: ICIN, 2008. 1 - 6.
- Minerva, Roberto. "On the art of creating services: Do different paradigms lead to different services?" *Journal of Telecommunications Management* (Henry Stewart Publications) 1 (2008): 33-45.
- Minerva, Roberto. *On the Importance of Numbers and Names for a 4G Service Architecture.* Vol. 3, in *Annual Review of Wireless Communication*. IEC, 2008.
- Minerva, Roberto, and Corrado Moiso. "Next Generation Networks: should The New Service Architecture Approximate The IN?" *International Switching Symposium, XVII World Telecommunications Congress.* Birmingham, UK: ISS, 2000.
- —. "The Death of Network Intelligence?" *Proc. International Symposium on Services and Local Access 2004 (ISSLS),.* Edinburgh, 2004.
- Minerva, Roberto, and Corrado Moiso. "Will the" circuits to packets" revolution pave the way to the" protocols to APIs" revolution?" *CSELT TECHNICAL REPORTS* (CSELT) 28, no. 2 (2000): 213-226.
- Minerva, Roberto, and Tiziana Demaria. "There is a Broker in the Net... Its name is Google." *Preceeding of ICIN Conference.* Bordeaux: ADERA - ICIN, 2006. 1 - 6.

- Minerva, Roberto, Corrado Moiso, and Gabriele Viviani. "The Middleware at the verge between Internet and Telecom services." *CSELT Technical Reports* (CSELT) 26 (1998): 657--672.
- Moerdijk, A-J, and Lucas Klostermann. "Opening the networks with Parlay/OSA: standards and aspects behind the APIs." *IEEE Network* (IEEE) 17, no. 3 (2003): 58-64.
- Moriana Group. *Service Delivery Platforms and Telecom Web Services - an Industry Wide Perspective.* THOUGHT LEADER Report, Egham, Surrey, UK: Moriana Group, 2004.
- Murty, James. *Programming Amazon Web Services: S3, EC2, SQS, FPS, and SimpleDB.* Sebastopol, CA, UK: O'Reilly Media, 2009.
- Net Neutrality, Wikipedia. *Network Neutrality.* 2013. http://en.wikipedia.org/wiki/Net_neutrality (accessed May Last Accessed 26th, 2013).
- Neumeyer, Leonardo, Bruce Robbins, Anish Nair, and Anand Kesari. "S4: Distributed stream computing platform." *IEEE International Conference on Data Mining Workshops (ICDMW).* Sydney, Australia: IEEE, 2010. 170 - 177.
- NFV. *Network Functions Virtualisation.* White Paper, Sophia Antipolis, France: ETSI, 2012.
- NGNI group. *FUSECO Playground.* Report, Berlin: Fraunhofer Fokus, 2009.
- NGNI Group. *Open IMS Playground.* Report, Berlin: Fraunhofer Fokus, 2005.
- NGNI Group. *Open SOA Telco Playground.* Report, Berlin: Fraunhofer Fokus, 2006.
- NGNI Group. *Smart Communications Playground.* Report, Berlin: Fraunhofer, 2012.
- Nielsen, Jakob. *Nielsen's law of internet bandwidth.* Report on line, Useit, 1998.
- NTIA. *Telecommunication Service.* web Page, National Telecommunications & Information Administration, Washington, USA: NTIA, 1996.
- Object Management Group. *CORBA Component Model 4.0.* Specification formal/06-04-01, Boston: OMG, 2006.
- Ohnishi, Hiroyuki, Yoji Yamato, Masashi Kaneko, Takaaki Moriya, Miki Hirano, and Hiroshi Sunaga. "Service delivery platform for telecom-enterprise-Internet combined services." *Global Telecommunications Conference, 2007. GLOBECOM'07.* Washington, DC, USA: IEEE, 2007. 108 - 112.
- Ohtsuki, H, C. Hauert, E Lieberman, and A. M. Nowak. "A simple rule for the evolution of cooperation on graphs and social networks." *Nature* (Nature) 441, no. 7092 (2006): 502–505.

- Open Networking Foundation. *OpenFlow-Enabled Cloud Backbone Networks Create Global Provider Data Centers.* White Paper, Palo Alto: Open Networking Forum, 2012.
- Orfali, Robert, Dan Harkey, and Jeri Edwards. *Client/server survival guide.* New York: Wiley, 2007.
- Pan, Jianli, Subharthi Paul, and Raj Jain. "A survey of the research on future internet architectures." *IEEE Communications Magazine* (IEEE) 49, no. 7 (2011): 26-36.
- Paolini, Monica. *The economics of small cells and Wi-Fi offload.* Report, Sammamish, WA, USA: SenzaFili Consulting, 2012.
- Parlay Group. *ParlayX 2.1 Specification.* Specification, London, UK: The Parlay Group, 2006.
- Partridge, Chris, and Ian Bailey. *An Analysis of Services.* Report Unclassified, Model Future, London, UK: UK Ministry of Defence, 2010.
- Pavlopoulos, A., and R. Cooper. "Towards a Survey of the Deployment Space for Internet Applications." *BNCOD '07 24th British National Conference on Databases.* Glasgow, UK: Springer, 2007. 110-119.
- Pavlovski, Christopher J. "Service delivery platforms in practice [IP Multimedia Systems (IMS) Infrastructure and Services]." *IEEE Communications Magazine* (IEEE) 45, no. 3 (2007): 114-121.
- Pepple, Ken. *Deploying OpenStack.* Sebastopol, CA, USA: O'Reilly Media, 2011.
- Pesavento, Umberto. "An implementation of von Neumann's self-reproducing machine." *Artificial Life* (MIT Press) 2, no. 4 (1995): 337-354.
- Pesch, Dirk, Maria Isabel Pous, and Gerry Foster. "Performance evaluation of SIP-based multimedia services in UMTS." *Computer Networks* (Elsevier) 49, no. 3 (2005): 385-403.
- Peterson, Larry, and John Wroclawski. *Overview of the GENI architecture.* GENI Design Document GDD-06-11, GENI: Global Environment for Network Innovations, Washington, DC, USA: NSF, 2007.
- Piatetsky-Shapiro, Gregory. "Data mining and knowledge discovery 1996 to 2005: overcoming the hype and moving from “university” to “business” and “analytics”." *Data Mining and Knowledge Discovery* (Springer) 15, no. 1 (2007): 99-105.
- Pike, Rob, Sean Dorward, Robert Griesemer, and Sean Quinlan. "Interpreting the data: Parallel analysis with Sawzall." *Scientific Programming* (IOS Press) 13, no. 4 (2005): 277-298.
- Platform, Google. *Google platform.* wikipedia topic, Wikipedia, Last accessed May 2013.
- Pouwelse, Johan, Pawel Garbacki, Dick Epema, and Henk Sips. "The bittorrent p2p file-sharing system: Measurements and analysis." *Peer-to-Peer Systems.* Heidelberg Berlin: Springer, 2005. 205-216.

- Prodan, Radu, and Simon Ostermann. "A survey and taxonomy of infrastructure as a service and web hosting cloud providers." *10th IEEE/ACM International Conference on Grid Computing.* Banff, Alberta, Canada: IEEE, 2009. 17-25.
- Proper, Erik. "Trends in Enterprise Architecture Research." Edited by S. Aier, M. Ekstedt, F. Matthes, E. Proper, & J.L. Sanz. *5th TEAR Workshop.* Delft: Springer, 2010.
- Rappa, Michael A. "The utility business model and the future of computing services." *IBM Systems Journal* (IBM) 43, no. 1 (2004): 32 - 42.
- Rexford, Jennifer and Dovrolis, Constantine. "Future Internet architecture: clean-slate versus evolutionary research." *Communications of the ACM* (ACM) 53, no. 9 (2010): 36-40.
- Rhea, Sean, et al. "OpenDHT: a public DHT service and its uses." *ACM SIGCOMM Computer Communication Review.* New York: ACM, 2005. 73-84.
- Richards, Robert. "XML-RPC." *Pro PHP XML and Web Services* (Apress), 2006: 595-631.
- Riede, Christian, Adel Al-Hezmi, and Thomas Magedanz. "Session and media signaling for IPTV via IMS." *Proceedings of the 1st international conference on MOBILe Wireless MiddleWARE, Operating Systems, and Applications.* Brussels: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008. Article n. 20.
- Ripeanu, Matei, Ian Foster, and Adriana Iamnitchi. *Mapping the gnutella network: Properties of large-scale peer-to-peer systems and implications for system design.* arXiv preprint cs/0209028, arXiv, 2002.
- Roberts G., Kudoh T., Monga I., Sobieski J., Vollbrecht J.,. *Network Services Framework v1.0.* Specification of the Open Grid Forum GFD.173, Lemont, Illinois, USA : Open Grid Forum, 2010.
- Roberts, James. "The clean-slate approach to future internet design: a survey of research initiatives." *Annals of telecommunications* (Springer) 64, no. 5 (2009): 271-276.
- Rosenberg, Jonathan, et al. *SIP: session initiation protocol.* RFC 3261, Internet Engineering Task Force, 2002.
- Rotem-Gal-Oz, Arnon. "Fallacies of distributed computing explained." http://www. rgoarchitects. com/Files/fallacies.pdf, 2006, 20.
- Rothenberg, Christian Esteve . "The increasing role of the Semantic Web & Ontologies in new generation network architectures." *Report online: http://chesteve.wordpress.com/2008/07/01/the-role-of-semantic-web-ontologies-in-newnext-generation-internetworking-architectures/*, 2008.
- Rowe, David. "The mesh potato." *Linux Journal* (Belltown Media) 2009, no. 188 (2009): Article n. 5.

- Sadashiv, Naidila, and SM Dilip Kumar. "Cluster, grid and cloud computing: A detailed comparison." *6th International Conference on Computer Science & Education (ICCSE).* Singapore: IEEE, 2011. 477-482.
- Sahoo, J., S. Mohapatra, and R., Lath. "Virtualization: A Survey on Concepts, Taxonomy and Associated Security Issues." *Second International Conference on Computer and Network Technology (ICCNT),.* Bangkok, Thailand: IEEE, 2010. 222 - 226.
- Saltzer, Jerome H., David P. Reed, and David D. Clark. "End-to-end arguments in system design." *ACM Transactions on Computer Systems (TOCS)* (ACM) 2, no. 4 (1984): 277-288.
- Santoro, Nicola. *Design and analysis of distributed algorithms.* Vol. 56. New York: Wiley-Interscience, 2006.
- Saracco, Roberto. "From value chains to ecosystem: New opportunities for telecommunications and new challenges for managing networks and services." *NOMS.* Salvador, Bahia, Brazil: IEEE, 2008. xxiii.
- Schaffer, H.E. "X as a Service, Cloud Computing, and the Need for Good Judgment." *IT Professional* (IEEE) 11, no. 5 (2009): 4-5.
- Schaller, Robert R. "Moore's law: past, present and future." *IEEE Spectrum* 34, no. 6 (1997): 52-59.
- Schulzrinne, Henning, and Jonathan Rosenberg. "A Comparison of SIP and H. 323 for Internet Telephony." Report, 1998, 162 - 167.
- Sciore, Edward. "SimpleDB: a simple java-based multiuser syst for teaching database internals." *ACM SIGCSE Bulletin* (ACM) 39, no. 1 (2007): 561-565.
- Sempolinski, Peter, and Douglas Thain. "A comparison and critique of eucalyptus, opennebula and nimbus." *IEEE Second International Conference on Cloud Computing Technology and Science (CloudCom).* Indianapolis, IN, USA: IEEE, 2010. 417-426.
- Service Science, Wikipedia. "Service science, management and engineering." *Wikipedia.* 2013. http://en.wikipedia.org/wiki/Service_science (accessed May Last accessed 27th , 2013).
- Shafiq, Muhammad Zubair, Lusheng Ji, Alex X Liu, Jeffrey Pang, and Jia Wang. "A first look at cellular machine-to-machine traffic: large scale measurement and characterization." *Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE joint international conference on Measurement and Modeling of Computer Systems.* London, UK: ACM, 2012. 65-76 .

- Singh, Kundan, Schulzrinne, Henning. "Peer-to-peer internet telephony using SIP." *Proceedings of the international workshop on Network and operating systems support for digital audio and video.* Skamania, Washington, USA: ACM, 2005. 63-68.
- Skype, Wikipedia. "Skype." *Wikipedia.* 2013. http://en.wikipedia.org/wiki/Skype (accessed May Last accessed 27th , 2013).
- Slater, Derek, and Tim Wu. *Homes with Tails-What if You Could Own Your Internet Connection.* White Paper, CommLaw Conspectus, 2009.
- Snell, James, Doug Tidwell, and Pavel Kulchenko. *Programming Web services with SOAP.* Sebastopol, CA, USA: O'Reilly Media, 2009.
- SOA-Wikipedia. *Definition of Service Oriented Architecture.* 2013. http://en.wikipedia.org/wiki/Service-oriented_architecture (accessed May Last accessed 26th, 2013).
- Spohrer, Jim, Paul P. Maglio, John Bailey, and Daniel Gruhl. "Steps Toward a Science of Service Systems." *Computer* (IEEE Computer Society) 40, no. 1 (2007): 71-77.
- St Arnaud, Bill, Jing Wu, and Bahman Kalali. "Customer-controlled and-managed optical networks." *Journal of Lightwave Technology* (IEEE) 21, no. 11 (2003): 2804 - 2810.
- Stonebraker, Michael. "SQL databases v. NoSQL databases." *Communications of the ACM* (ACM) 53, no. 4 (2010): 10-11.
- Stonebraker, Michael, Uğur Cetintemel, and Stan Zdonik. "The 8 requirements of real-time stream processing." *ACM SIGMOD Record* (ACM) 34, no. 4 (2005): 42-47.
- Sward, Ricky E, and Jeff Boleng. "Service-oriented architecture (SOA) concepts and implementations." *ACM SIGAda Ada Letters.* New York: ACM, 2012. 3-4.
- Tarkoma, Sasu. *Clean-Slate Datacentric Pub/Sub Networking.* Sebastopol, Ca, USA: Wiley Online Library, 2012.
- Tavakoli, Arsalan, Martin Casado, Teemu Koponen, and Scott Shenker. "Applying NOX to the Datacenter." *HotNets.* New York City, NY : ACM, 2009.
- Taylor, Ian J. *From P2P to Web services and grids: peers in a client/server world.* Berlin Heidelberg: Springer, 2005.
- Taylor, Ian J, and Andrew B Harrison. *From P2P and grids to services on the web: evolving distributed communities.* Heidelberg Berlin: Springer, 2009.

- Teece, David J. "Business models, business strategy and innovation." *Long range planning* (Elsevier) 43, no. 2 (2010): 172–194.
- Thampi, Sabu M. *Introduction to Distributed Systems.* preprint arXiv:0911.4395, arXiv preprint arXiv:0911.4395, 2009.
- Thom, Gary A. "H. 323: the multimedia communications standard for local area networks." *IEEE Communications Magazine* (IEEE) 34, no. 12 (1996): 52 - 56.
- TISPAN. *NGN Functional Architecture.* ETSI ES 282 001, Sophia Antipolis: ETSI, 2009.
- Tsagkaris, Kostas, et al. "Operator-driven framework for establishing and unifying autonomic network and service management solutions." *GLOBECOM Workshops (GC Wkshps), 2011 IEEE.* Houston, TX, USA: IEEE, 2011. 684 - 689 .
- Turaga, Deepak, et al. "Design principles for developing stream processing applications." *Software: Practice and Experience* (Wiley Online Library) 40, no. 12 (2010): 1073-1104.
- Turing, Alan M. "On computable numbers, with an application to the Entscheidungsproblem." *Proceedings of the London mathematical society* (London mathematical society) 42, no. 2 (1936): 230--265.
- Turner, Jon, and Nick McKeown. "Can overlay hosting services make ip ossification irrelevant." *PRESTO: Workshop on Programmable Routers for the Extensible Services of TOmorrow.* 2008.
- Turner, Mark, David Budgen, and Pearl Brereton. "Turning software into a service." *Computer* (IEEE) 36, no. 10 (2003): 38 - 44.
- Upton, Eben, and Gareth Halfacree. *Meet the Raspberry Pi.* New York: Wiley, 2012.
- Vahdat, Amin. "Symbiosis in scale out networking and data management." *Proceedings of the 2012 international conference on Management of Data.* Scottsdale, Arizona, USA: ACM, 2012. 579-580.
- Valancius, Vytautas, Nikolaos Laoutaris, Laurent Massoulie, Christophe Diot, and Pablo Rodriguez. "Greening the internet with nano data centers." *Proceedings of the 5th international conference on Emerging networking experiments and technologies.* Rome, Italy: ACM, 2009. 37-48.
- Vaquero, Luis M, Luis Rodero-Merino, Juan Caceres, and Maik Lindner. "A break in the clouds: towards a cloud definition." *ACM SIGCOMM Computer Communication Review* (ACM) 39, no. 1 (2009): 50-55.
- Varum, Celeste Amorim, and Carla Melo. "Directions in scenario planning literature--A review of the past decades." *Futures* (Elsevier) 42, no. 4 (2010): 355–369.

- Vemuri, Kumar. "Sphinx: A study in convergent telephony." *Proceedings of the IP Telecom Services Workshop 2000 (IPTS2000).* Atlanta, Georgia, USA: ATT Lab Research & Pulver.com, 2000. 9 - 18.
- VMWare. *Virtualization overview.* White Paper, Palo Alto, Ca, USA: VMWare, 2006.
- Vouk, Mladen A. "Cloud computing—Issues, research and implementations." *30th International Conference on Information Technology Interfaces, 2008. ITI 2008.* Dubrovnik, Croatia: IEEE, 2008. 31 - 40.
- Waclawsky, John G. "IMS: a critique of the grand plan." *Business Communications Review* (BCR ENTERPRISES INC) 35, no. 10 (2005): 54 - 58.
- Walkden, M, et al. *Open service access: advantages and opportunities in service provisioning on 3G mobile networks definition and solution of proposed Parlay/OSA specification issues.* OSA Specification issues, Project P1110 Technical Information EDIN, Heidelberg, Germany: Eurescom, 2002.
- Walter, Chip. "Kryder's law." Edited by Scientfic American. *Scientific American* (Nature Publishing Group) 293, no. 2 (2005): 32-33.
- Wan, Chieh-Yih, Andrew T Campbell, and Lakshman Krishnamurthy. "PSFQ: a reliable transport protocol for wireless sensor networks." *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications.* Atlanta, GA, USA: ACM, 2002. 1-11 .
- Weiser, Mark, Rich Gold, and John Seely Brown. "The origins of ubiquitous computing research at PARC in the late 1980s." *IBM systems journal* (IBM) 38, no. 4 (1999): 693 - 696.
- Weiss, Aaron. "DD-WRT Tutorial 3: Building a Wireless Bridge." *WiFi PLanet.* October 20th, 2006. www.wi-fiplanet.com/tutorials/article.php/3639271 (accessed May Last accessed 27th, 2013).
- Werner, Max Jonas. *Peer-to-Peer Networking using Open Web Technologies.* Research Report, Hamburg, Germany: Faculty of Engineering and Computer Science of Hamburg University of Applied Sciences, 2013.
- White, Tom. *Hadoop: The definitive guide.* Sebastopol, CA, USA: O'Reilly Media, Inc., 2010.
- Wilson, Brendon J. *Jxta.* San Francisco, CA, USA: New Riders - Pearson Publishing, 2002.
- World Economic Forum. *Personal Data: the emergence of a new asset class.* White Paper, Cologny - Geneva: World Economic Forum, 2011.
- World Economic Forum. *Unlocking the Value of Personal Data: From Collection to Usage.* White Paper, Colgny - Geneva: WEF, 2013.

- Xie, Haiyong, Y Richard Yang, Arvind Krishnamurthy, Yanbin Grace Liu, and Abraham Silberschatz. "P4p: provider portal for applications." *ACM SIGCOMM Computer Communication Review* (ACM) 38, no. 4 (2008): 351-362 .
- Yagi, H, K Ohtsu, M Wakano, H Kobayashi, and R Minerva. "TINA-C service components." *Proc. TINA95 Integrating Telecommunications and Distributed Computing-from Concept to Reality.* Melbourne, Australia: TINA_C, 1995.
- Yap, Kok-Kiong, et al. "Blueprint for introducing innovation into wireless mobile networks." *Proceedings of the second ACM SIGCOMM workshop on Virtualized infrastructure systems and architectures.* New Delhi, India: ACM, 2010. 25-32.
- Zaslavsky, Arkady, Charith Perera, and Dimitrios Georgakopoulos. *Sensing as a service and big data.* arXiv:1301.0159v1, arXiv preprint, 2013.
- Znaty, Simon , and Jean-Pierre Hubaux. "Telecommunications Services Engineering: Principles, Architectures." *ECOOP Workshops.* Jyväskylä, Finland: Springer, 1997. 3-11.