Exploitation dynamique des données de production pour améliorer les méthodes DFM dans l'industrie Microélectronique

Muhammad Kashif Shahzad

UNIVERSITÉ DE GRENOBLE

Thèse

Pour obtenir le grade de

Docteur de l'Université de Grenoble

Spécialité : **Génie Industriel : Conception et Production**

Arrêté ministériel : 07 Août 2006

Présentée par

Muhammad Kashif Shahzad

Thèse dirigée par **Michel Tollenaere**

et encadrée par **Ali Siadat et Stéphane Hubac**

préparée au sein du **Laboratoire G-SCOP (Laboratoire des Sciences pour la Conception, l'Optimisation et la Production de Grenoble - UMR5272)** et de l'**École doctorale I-MEP2 (Ingénierie - Matériaux Mécanique Énergétique Environnement Procédés Production)**

Exploitation dynamique des données de production pour améliorer les méthodes DFM dans l'industrie Microélectronique

Thèse soutenue publiquement le **05 Octobre 2012**,
devant le jury composé de :

Prof. Stéphane Dauzère-Pérès École des mines de Saint-Étienne, France	Président
Prof. Bernard Grabot Ecole Nationale d'Ingénieur de Tarbes, France	Rapporteur
Prof. Hervé Panetto Université de Lorraine -Nancy, France,	Rapporteur
Prof. Michel Tollenaere Université de Grenoble, France	Directeur de Thèse
Dr. Ali Siadat MCF, ENSAM Arts et Métiers ParisTech, Metz, France	Co-encadrant
Mr. Stéphane Hubac Process Control Expert, STMicroelectronics, Crolles, France	Directeur Industriel

The Dissertation Committee for Muhammad Kashif Shahzad Certifies that this is the approved version of the following dissertation:

<< Exploitation dynamique des données de production pour améliorer les méthodes DFM dans l'industrie Microélectronique >>

Thesis Committee:

Prof. Dauzère-Pérès, Stéphane (Président)

Prof. Grabot, Bernard (Rapporteur)

Prof. Panetto, Hervé (Rapporteur)

Prof. Michel Tollanere, Academic Supervisor

Prof. Ali Siadat, Academic Co-Supervisor

Mr. Stéphane Hubac, Industrial Supervisor

Dedication

To my family, who offered me unconditional love and support throughout the course of this thesis.

*T*he beginning of knowledge is the discovery of something we don't understand.

- Frank Herbert

Preface

Everything started in Pakistan...

I completed my two bachelors' degrees in mechanical engineering and computer sciences in 1999 and started my career as a computer instructor at the Petromen Training Institute (ministry of science & technology), Pakistan. I served here till 2001 and found that there is a big gap between academia and industry for the industrialization of latest research to uplift the performance. I joined Pakistan Information Technology Company (PITC), Pakistan in 2001 as Assistant Director (programmer/system analyst) to bridge this gap with latest IT revolutions. I served PITC till 2007 and contributed to my best. I completed MS degree in 2006 with specialization in total quality management (TQM) to update myself with the latest research and developments.

During these academic and industrial experiences I learned that research is the backbone for industrial performance, so I started looking for a PhD position and received MS+PhD scholarship from the HEC (higher education commission), Pakistan. I selected France among many European countries for my PhD based on the quality of education, research and its rich cultural heritage. I profited from the beautiful French language and after 5 years of my stay in France, it seems like my second home.

I completed my M2R degree (Industrial Engineering) at the ENSGI (Ecole Nationale Supérieure de Génie Industriel) in 2008 and did my research internship at the Laboratoire G-SCOP (Sciences pour la Conception l'Optimisation et la Production) with Professor Michel Tollenaere. I was very fortunate that Professor Tollenaere offered me a PhD position at STMicroelectronics that exactly matched my objectives to do research and deploy it for the industrial performance. I started my PhD with the registration at I-MEP2 doctoral school in December, 2008. Four years later, the end of my PhD is a good opportunity to take a look backwards and meditate about these four years. This PhD has changed my life and gave me an opportunity to dive into an exciting and challenging domain of semiconductor manufacturing. It is a wonderful experience where I participated and presented my research articles that gave me a chance to visit different cities in France and different countries including Portugal, Holland, Italy, Canada and USA.

During M2R I attended the courses in global supply chain management, operational research and optimization, production planning, industrial information systems, research methodology, global industrialization, management information systems and artificial intelligence and knowledge management. My M2R research internship "supply chain configuration modeling under the influence of product family architecture" at the Laboratoire G-SCOP was an exciting experience that lasted for 5 months. I owe my special gratitude to my co-supervisor Khaled Hadj-Hamu. It is all his motivation, guidance and supervision that I published my first conference (INCOM-2009) and journal (JIM-2011) articles. This research training and the courses at ENSGI updated me with the latest research in the domain of supply chain configuration and optimization. During the PhD (2008-2012) I spent 70% of my time at STMicrlelectronics and 30% at G-SCOP. This division helped me a lot in getting the professional and academic experience with advanced R&D teams at STMicroelectronics and Laboratoire G-SCOP. I have undertaken many courses at the training center of STMicroelectronics on advanced semiconductor manufacturing, process control and data analysis tools. Besides this I also got the opportunity to receive one week six-sigma green belt training in Holland followed by two advanced design of experiment and engineering data modeling courses at MIT, USA.

I appreciate the French education system that offers CIFR (industrial PhD scholarships) in close collaboration between research labs and the industries. I was fortunate to achieve this distinct scholarship at STMicroelectronics in collaboration with two labs *(i)* G-SCOP, Grenoble and *(ii)* LCFC, Metz France. I have really enjoyed my whole stay (3 years at STMicroelectronics and 1 year at Laboratories G-SCOP and LCFC)

where I was able to industrialize the research carried out through the course of this PhD experienced an exceptionally advanced R&D environment.

Finally I would like to express my words of gratitude for my academic supervisor Dr. Michel Tollenaere, co-supervisor Dr. Ali Siadat and my industrial research director Mr. Stéphane Hubac for their guidance in developing the present work and also for being so supportive, encouraging, and understanding in good and bad times. It is because of their support and guidance that I was able to publish 7 international conferences along with 4 journal submissions currently under review. My special thanks for my colleagues at STMicroelectronics and G-SCOP with home I participated in different case studies and important discussions that helped me in understanding the complex semiconductor manufacturing processes.

This is not the end but the beginning and there is a long way to go...

Acknowledgement

I feel necessary to personally thank all those who have helped and guided me throughout my PhD.

This research and thesis could not have been completed without the help and guidance from many people. First and foremost I have to thank my academic supervisors Dr. Michel Tollenaere and Dr. Ali Siadat. They were always there to guide me in my research as well as in solving problems in general. Under their guidance I was able to grow from an under graduate student to a Doctor of Philosophy. Although I spent 30% of my time at the research lab, but still my supervisors were always there for me when needed. It is really an honor for me to work at the advanced R&D center Crolles, of the 7th largest semiconductor manufacturer (2012) STMicroelectronics. I owe my special gratitude for Mr. Stéphane Hubac, my industrial director, who helped me in integration and getting familiar with the complex semiconductor manufacturing and R&D processes in the most simple and beautiful way. I found him as an intelligent and highly competent person with exceptional managerial skills and his training and guidance shall help all the way throughout my career.

I worked with different R&D teams and I would like to acknowledge their support, guidance and participation. First of all I am grateful to Renaud Nicolas, Lidwine Chaize and Jean-Claude Marin from T2D (technology to design) Aples team for their participation and we together proposed MAM (mapping and alignment) model for die/site level correlations to capture spatial correlations. It is a significant contribution towards more effective DFM methods. I cannot forget my experience with the yield management group particularly Celine Julien and Roberto Gonella where we together proposed SPM (spatial positioning) model to accurately capture the spatial variations. This proposed solution is a significant contribution towards accurately capturing the spatial variations. My special thanks go to IT teams at STMicroelectronics because without their support it would have not been possible to analyze the production database issues and propose the ROM2I2 (referential ontology Meta model for information integration). I am grateful to Hugues Duverneuil, Guillaume Chezaud, Brigitte Boulloud, David Rozier, Stephane Coquio and Bruno Sauvage for giving me access to the manufacturing database and data storage architecture.

I appreciate the contribution and support of Jerome Altieri, Thomas Chaillou and Hugues Kejikian from the equipment engineering team at STMicroelectronics. We together developed a 3-step yield aware sampling strategy (YASS) to enhance the inspection capacities for more R&D measurements. I nust appreciate the team lead by Mr. Herve Jaouen in the development of SMA (spice model alignment) tool for the process integration team. I personally thank all the team members in the process integration team Raul-andres Bianchi, Gregory Bidal, Claire Gallon, Dominique Golanski, Guillaume Ribes, Jerome Bonnouvrier and Nicolas Planes for their support and help.

Finally I want to thank my colleagues at STMicroelectronics: Francois Pasqualini, Philippe Vialletelle, Alain Delporte and Aymen Mili with whom I did not work in a team but their support has helped me a lot. I also thank all of my colleagues at G-SCOP (Mhamed Sahnoun, Mohammed Farouk Bouaziz) and LCFC labs for their encouragement and support.

Exploitation dynamique des données de production pour améliore les méthodes DFM dans l'industrie Microélectronique

Muhammad Kashif SHAHZAD, PhD (Industrial Engineering)
University of Grenoble, 2012

Academic Supervisor: **Prof. Michel Tollenaere**
Academic Co-supervisor: **Dr. Ali Siadat**
Industrial Supervisor: **Mr. Stephane Hubac**

+

Abstract: The DFM (design for manufacturing) methods are used during technology alignment and adoption processes in the semiconductor industry (SI) for manufacturability and yield assessments. These methods have worked well till 250nm technology for the transformation of systematic variations into rules and/or models based on the single-source data analyses, but beyond this technology they have turned into ineffective R&D efforts. The reason for this is our inability to capture newly emerging spatial variations. It has led an exponential increase in technology lead times and costs that must be addressed; hence, objectively in this thesis we are focused on identifying and removing causes associated with the DFM ineffectiveness. The fabless, foundry and traditional integrated device manufacturer (IDM) business models are first analyzed to see coherence against a recent shift in business objectives from time-to-market (T2M) and time-to-volume towards (T2V) towards ramp-up rate. The increasing technology lead times and costs are identified as a big challenge in achieving quick ramp-up rates; hence, an extended IDM (*e*-IDM) business model is proposed to support quick ramp-up rates which is based on improving the DFM ineffectiveness followed by its smooth integration. We have found *(i)* single-source analyses and *(ii)* inability to exploit huge manufacturing data volumes as core limiting factors (failure modes) towards DFM ineffectiveness during technology alignment and adoption efforts within an IDM. The causes for single-source root cause analysis are identified as the *(i)* varying metrology reference frames and *(ii)* test structures orientations that require wafer rotation prior to the measurements, resulting in varying metrology coordinates (die/site level mismatches). A generic coordinates mapping and alignment model (MAM) is proposed to remove these die/site level mismatches, however to accurately capture the emerging spatial variations, we have proposed a spatial positioning model (SPM) to perform multi-source parametric correlation based on the shortest distance between respective test structures used to measure the parameters. The *(i)* unstructured model evolution, *(ii)* ontology issues and *(iii)* missing links among production databases are found as causes towards our inability to exploit huge manufacturing data volumes. The ROMMII (referential ontology Meta model for information integration) framework is then proposed to remove these issues and enable the dynamic and efficient multi-source root cause analyses. An interdisciplinary failure mode effect analysis (*i*-FMEA) methodology is also proposed to find cyclic failure modes and causes across the business functions which require generic solutions rather than operational fixes for improvement. The proposed e-IDM, MAM, SPM, and ROMMII framework results in accurate analysis and modeling of emerging spatial variations based on dynamic exploitation of the huge manufacturing data volumes.

Keywords: design for manufacturing, effective root cause analysis, information integration, metrology coordinates mapping and alignment

Table of Contents

List of Tables

List of Figures

List of Acronyms

AAP	Average Application Probability
a.k.a.	also known as
BEOL	Back-End-Of-Line
BPR	Business Process Reengineering
CAP	Current Application Probability
CMOS	Complementary Metal Oxide Semiconductor
C&E	Cost and Effect Matrix
CTP	Common Technology Platform
DBMS	Database Management Systems
DFA	Design for Assembly
DFD	Design for Design
DFF	Design for Fabrication Facilities
DFM	Design for Manufacturing
DFMA	Design for Manufacturability and Assembly
DFMEA	Design Failure Mode Effect Analysis
DFP	Design for Product
DFR	Design for Reliability
DFSM	Design for Semiconductor Manufacturing
DFT	Design for Test
DFX	Design for All
DFY	Design for Yield
DL	Description Logic
DM	Data Mart
DMS	Data-Method-Stat Triangle
DOE	Design of Experiment
DRC	Design Rule Check
DRM	Design Rule Manual
DSP	Digital Signal Processing
DTD	Document Type Definition
DWH	Data Warehouse
EBR	Edge Bevel Removal
ECAD	Electrical Computer Aided Design
EDA	Engineering Data Analysis
EEPROM	Electronically Erasable Programmable Read Only Memory
EHF	Equipment Health Factor
ER	Entity Relationship
EWS	Electrical Wafer Sort

FDC	Fault Detection and Classification
FE	Focus Exposure
FEOL	Front-End-Of-Line
FIS	Federated Information System
FMEA	Failure Mode Effect Analysis
FMECA	Failure Mode Effect and Criticality Analysis
GDS	Graphic Design System
HOLAP	Hybrid Online Analytical Processing
IA	Impact Analysis
IDM	Integrated Device Manufacturer
IO	Input Output
IP	Intellectual Property
IPD	Integrated Product Development
LIFO	Last In First Out
LSI	Large Scale Integration
LSL	Lower Specification Limit
LVS	Layout versus Schematic
MCU	Micro Controller Unit
MDD	Manufacturing Driven Design
MOLAP	Multidimensional Online Analytical Processing
MPU	Micro Processing Unit
MSI	Medium Scale Integration
ODS	Operational Data Stores
OEM	Original Equipment Manufacturer
OLAP	Online Analytical Processing
OLTP	Online Transactional Processing
OOC	Out of Control
OPC	Optical Proximity Correction
OWL	Ontology Web Language
PCM	Process Control Monitor
PCS	Process Control Structures
PDA	Personal Digital Assistant
PLC	Product Life Cycle
PMB	Process Monitoring Box
POP	Package on Package
PFMEA	Process Failure Model Effect Analysis
PLM	Product Life Cycle Management
PSM	Phase Shift Masking
QA	Quality Assurance

QFD	Quality Function Deployment
RAM	Random Access Memory
RDF	Resource Description Framework
ROLAP	Relational Online Analytical Processing
ROM	Read Only Memory
RPN	Risk Priority Number
RTL	Register Transfer Level
SAE	Society of Automotive Engineers
SCAN	Strategic Creative Analysis
SI	Semiconductor Industry
SIP	System in Package
SIPOC	Supplier, Input, Process, Output and Customer
SOC	System on Chip
SPICE	Simulation Program with Integrated Circuit Emphasis
SSI	Small Scale Integration
SWFMEA	Software Failure Mode Effect Analysis
SWOT	Strengths, Weakness, Opportunity and Threat Analysis
TCAD	Technology Computer Aided Design
TRO	Top Ranked Objective
TSMC	Taiwan Semiconductor Manufacturing Company
T2D	Technology to Design
T2M	Time to Market
T2V	Time to Volume
T2Q	Time to Quality
ULSI	Ultra Large Scale Integration
UPT	Unified Process Technology
USL	Upper Specification Limit
VLSI	Very Large Scale Integration
WIP	Work in Process

Chapter 1: Introduction

The objective of this chapter is to introduce readers with the problem background, research questions and methodology along with a brief description of the industrial and scientific contributions. We start with an introduction to the semiconductor industry (SI), current trends and challenges. The biggest challenge faced today is the increasing technology lead times and costs due to ineffective R&D efforts against newly emerging manufacturability and yield loss mechanisms. The design for manufacturing (DFM) approach being used against these issues is presented and discussed for its economic benefits and limitations. The single-source wafer/lot level analysis is found as the core limiting factor towards more effective DFM methods, resulting in extended lead times and costs. Based on this discussion, four research questions are formulated followed by a graphical representation of research methodology and timeline. The industrial and scientific contributions are also briefly presented to complete the improvement cycle and finally we conclude the chapter with typographic convention used throughout this thesis and the thesis organization.

Contents

1.1 INTRODUCTION TO SEMICONDUCTOR INDUSTRY (SI)

The history of the semiconductor can be traced back to 1947 with the invention of first transistor by John Bardeen, Walter Brattain and William Shockley. The transistor acts like an on/off switch and is the basic building block for an electronic circuit where thousands of transistors are manufactured and interconnected on the silicon wafer to form an integrated circuit (IC) chip. The SI is responsible for an efficient and effective manufacturing of the IC chips and was borne with the invention of first integrated circuit in 1959. The first IC was developed by Jack Kilby at the Texas Instruments (TI) in 1959; however the Fairchild lab is credited for the invention of first commercial IC logic gate in 1961 by Robert Noyce [Brinkman et al., 1997]. The first highly integrated circuit had 2300 transistors and it was manufactured by the Intel (4bit microprocessor 4004) in 1971 [Kumar, 2008]. Since then, the transistor count has exponentially increased till today. This potential growth was first predicted by Gordon E. Moore, cofounder of Intel, in 1965 who postulated that the transistor count shall double every 18-24 months at reduced cost, area and power [Moore, 1998]. This prediction along with the industrial slogan "smaller, faster and cheaper" has been accepted by the SI as a standard. The "smaller" means more transistors in the same area resulting in faster flow of current at reduced power and "cheaper" refers to continuously decreasing costs due to increasing yield.

The semiconductor industry (SI) has revolutionized our daily lives with IC chips and on the average we are using more than 250 chips and 1 billion transistors per day per person. These chips are installed in almost all the equipments around us ranging from dish washers, microwave ovens and flat screens to office equipments. The use of semiconductors in cars, trains, aircraft and ships is constantly expanding and PCs, servers and pocket calculators also owe their existence to these chips. The global electronics business is 1.04T$ industry [Dummer, 1997] including 300+ B$ share of the semiconductor manufacturing (2012) during last 50 years. The global semiconductor manufacturing (1998 till 2011) sales revenues characterize SI with the cyclic demand patterns and a +ve cumulative annual growth rate (CAGR) of 8.72% (Figure 1.1). This positive CAGR ensures that the demand driven downfalls follow a cumulative demand growth; hence, it motivates the SI to continuously invest in the R&D for a new technology to capture the maximum market share from cumulative growths. The R&D investment in SI is expected to reach 18% of the total revenues by the year 2012. It is evident that the success of SI lies in the effectiveness of these R&D efforts which has a strong impact on the lead times and costs associated with the new technology development, its derivative or simply the improvement initiatives.

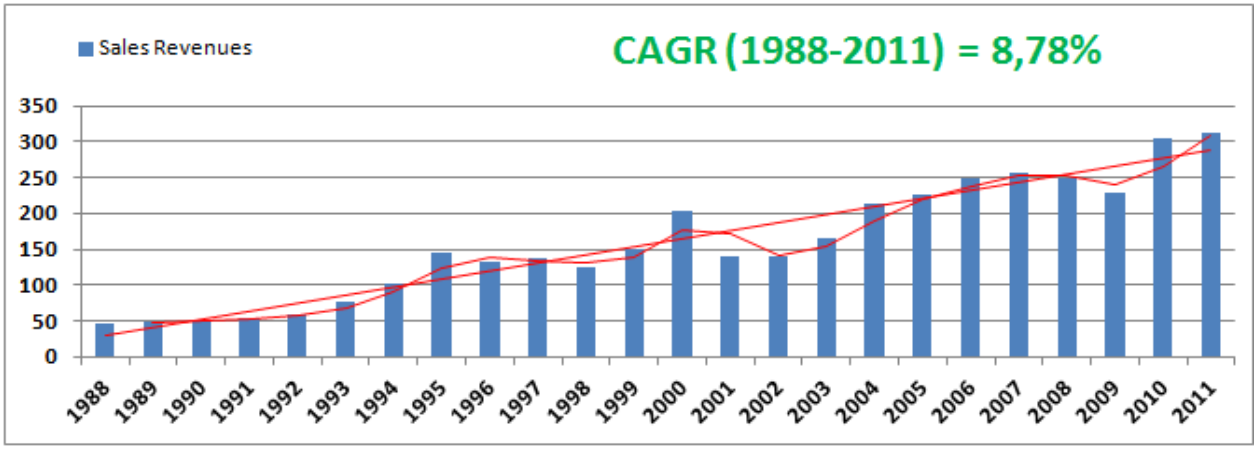


Figure 1.1 - Global sales revenues of SI[1]

The SI has kept its pace as per Moore's law with the continuous introduction of new technologies, every 2 to 3 years. The increase in transistor count for new technologies has led exponential and linear increases in the technology and product development costs (Figure 1.2) respectively. As a matter of fact, we need new technology every 2 years to keep up the pace with the Moore's law but doubling transistors count add manufacturing complexities that result in new manufacturability and yield loss mechanisms. The existing DFM methods do not

[1] The data is collected from the well known technology research centers *(i)* Gartner {www.gartner.com} and *(ii)* isuppli {www.isuppli.com}

provide solution to these new yield loss mechanisms; hence, extensive R&D efforts along with innovation in the material, process and equipment are must for new technologies. It often results in increasing *(i)* technology costs and *(ii)* technology lead times.

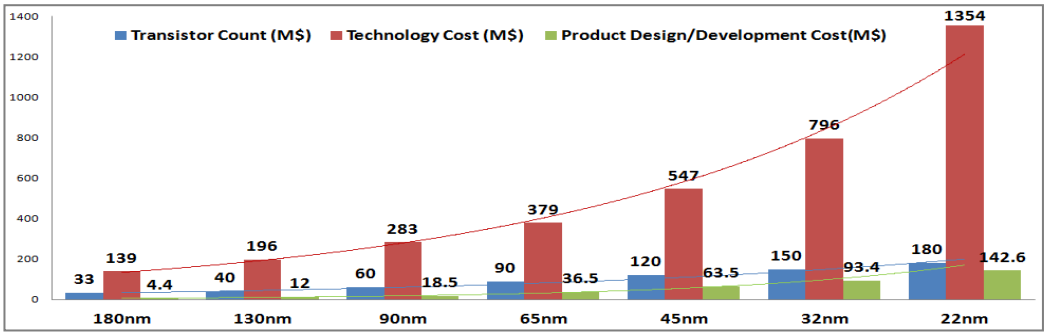


Figure 1.2 - Global product and technology costs for SI[2]

The DFM (design for manufacturing) philosophy is a well known methodology for the yield and manufacturability assessments. It is subjectively focused on transforming systematic manufacturing variations into rules and/or models for their subsequent use in CAD simulations. The DFM methods play a significant role in the technology development; hence, it has got focus of the SI industry to help in coping up with the Moore's law. These methods were adopted by the SI industry in 1980 as a yield enhancement strategy that worked very well till 250nm technology, but beyond it has turned into a high cost ineffective R&D effort. The DFM inefficiencies due to these ineffective R&D efforts are compensated by two means: *(i)* developing a technology in an alliance to share R&D costs for its timely introduction to the market and *(ii)* material, equipment and process based innovations. The success of the SI lies in their ability to quickly transfer this new technology in alliance partners' business models and its subsequent continuous derivative and improvement initiatives followed by alignment (validation) and adoption (customization). The cyclic demand patterns and CAGR guarantees an equal opportunity to all the stake holders, resulting in R&D investments even in downtimes. The Samsung electronics is the best example where it raised revenues in just one year by 10.34B$ where 80% of these revenues came from DRAM and NAND flash memories (Figure 1.3) and it resulted from 32nm to 30nm technology derivative (alignment and adoption) efforts.

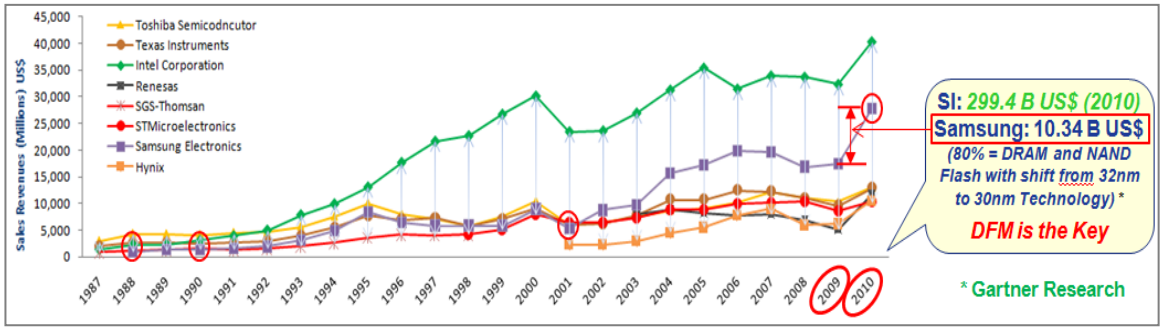


Figure 1.3 - Global sales revenues of SI[2]

The SI has followed the **I**ntegrated **D**evice **M**anufacturer (IDM) business model since 1980 where products are designed and manufactured in the same facility. To address the increasing technology costs and lead times that often result in an opportunity loss, the SI has transformed into fablite and fabless business models. The fablite business model is similar to an IDM model with an exception that new technology is developed in an alliance and is transferred to alliance partners' business models prior to its alignment and adoption. The fabless model advocates the design and production as specialized business function in the separate facilities with an objective to address the emerging manufacturability and yield loss mechanisms and to reduce the exponential increase in technology alignment and adoption costs. The IDM-fablite is known as the best SI business model in competitive environment because it is coherent to reduce the

[2] The data is collected from the well known technology research centers *(i)* Gartner {www.gartner.com}, *(ii)* isuppli {www.isuppli.com} and (iii) International Business Strategies, Inc., report: Key trends in technology and supply for advanced features within IC industry (October 1, 2009)

technology alignment and adoption lead times (e.g. 30nm, Figure 1.3) and costs [Shahzad et al., 2011b]. The ineffective DFM methods are a big challenge for an IDM-fablite model and in this thesis we are focused on putting DFM back on track by removing the causes for ineffectiveness.

In this thesis we are focused on *(i)* fast technology transfer from an alliance to the IDM-fablite business model and *(ii)* continuous technology alignment and adoption for its subsequent derivatives and improvements. It is based on the fact that an IDM-fablite business model provides coherent platform for knowledge capitalization based on production data exploitation and analysis (effective root cause analysis) for technology alignment and adoption lead times and costs improvements. The technology is generally defined by the minimum feature size that can be manufactured with it e.g. 250nm, 180nm, 130nm, 90nm, 65nm, 45nm, 32nm and 22nm. It comprises of the design flow, **S**imulation **P**rograms with **I**ntegrated **C**ircuit **E**mphasis (SPICE) models, **D**esign **R**ules **M**anual (DRM), **C**omputer **A**ided **D**esign (CAD) tools, process flow, equipments, recipes, **S**tatistical **P**rocess **C**ontrol (SPC) and **R**un **t**o **R**un (R2R) feed forward or feed backward strategies. The DRM includes the rules and/or models to address the manufacturability and yield loss mechanisms for a given technology. These rules and/or models are programmed in the form of process and DFM kits which are used during the CAD simulations to assess the final printed features and associated drifts. These drifts in the geometric shapes of the features are then analyzed with **C**ritical **A**rea (CAA) and hot spot **A**nalyses to find manufacturability and yield limitations. It helps us in solving the yield issues early in design phase to avoid design respins and waste of resources during prototyping, which result in lead times and costs improvements.

The technology alignment is a process to validate design, process and DFM kits and is equally applicable for new technology, its derivative or simple improvement initiatives. It has a strong impact on the reusability concept as all the design libraries (pre-designed circuits) need to be requalified which is not trivial. The technology adoption refers to the customization of base technology for each product to achieve target yield levels and it uses **A**dvance **P**rocess (APC) and **E**quipment (AEC) **C**ontrol techniques. The DFM methods are objectively focused on the manufacturability and yield but subjectively they are focused on finding root causes against deviations/drifts encountered during the technology alignment and adoption efforts. It can be concluded that the success of SI lies in the effective R&D efforts to identify root causes; hence, effective R&D is a key to improve existing ineffective DFM methods resulting in improved lead times and costs.

1.2 ROLE OF DFM AND ECONOMIC BENEFITS

The economic benefits associated with the DFM methods are presented in Figure 1.4, which are divided in technology alignment and/or adoption efforts and economic benefits. The green and red curves represent cumulative cash flows with and without DFM efforts. The DFM methods help in improving design, development and ramp-up periods (lead times) by quickly finding the root causes against newly emerging manufacturability and yield loss mechanisms. It results in cost reduction and provides an early penetration into the market with higher profit margins; however the biggest gain is the long selling period in the SI against the continuously shortening product life cycles.

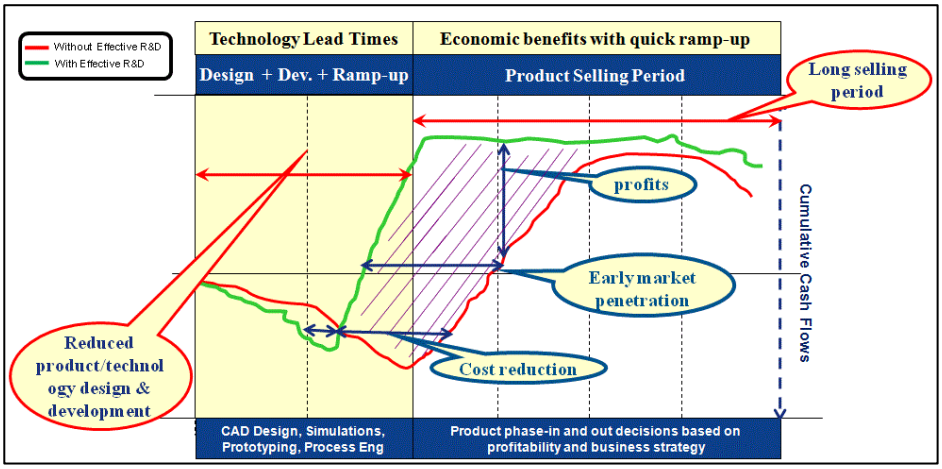


Figure 1.4 - The role of DFM and economic benefits

It is evident that in this competitive environment with continuously shortening product life cycles, we must focus and improve the technology lead times. This improvement results in the maximum market share in an equal opportunity 300+ B$ industry along with long selling periods. It is important to note that we are introducing new technologies at every 2-3 years at an exponentially increasing technology costs (Figure 1.2) and lead times where DFM ineffectiveness is attributed towards ineffective R&D efforts. The ineffective DFM methods are compensated with the innovation in design, process, material and equipment which result in extended technology lead times and costs. These innovations are of course necessary for the changing requirements and to establish leader position, however local R&D efforts within IDM-fablite models for technology derivative and improvement alignment and adoption must be improved to achieve true DFM economic benefits.

1.3 DFM CHALLENGES AND LIMITATIONS

It is very important to start with the basic understanding of semiconductor design and manufacturing processes prior to discuss the DFM role and challenges in SI. Let us start with the design side where an IC is characterized by electrical parameters (functions) where it undergoes a complex manufacturing process with approximately 200+ operations, 1100+ steps and 8 weeks of processing, however the number of steps and operations vary with the selected technology. A chip is designed using CAD tools and design libraries (reusable blocks of circuits). It follows the design simulation steps where design rules are validated and electrical and parasitic (unwanted) parameters are extracted using SPICE models and technology files. The drifts and variations are adjusted through layout optimization. The design is further simulated using DFM rules and/or models (CAA and hotspot analyses) to find out potential drifts in drawn features and printed images that could result in manufacturability and yield losses. These potential failures are addressed by either changing the design or layout optimization. Upon validation, the design moves to the mask preparation step. The masks are glass plates with an opaque layer of chrome carrying the target chip layout. For one product the mask set consists of 15 to 35 individual masks depending on the technology. They are used to fabricate thousands of transistors and a network of interconnected wires to form an electronic chip (Figure 1.5) on silicon wafer.

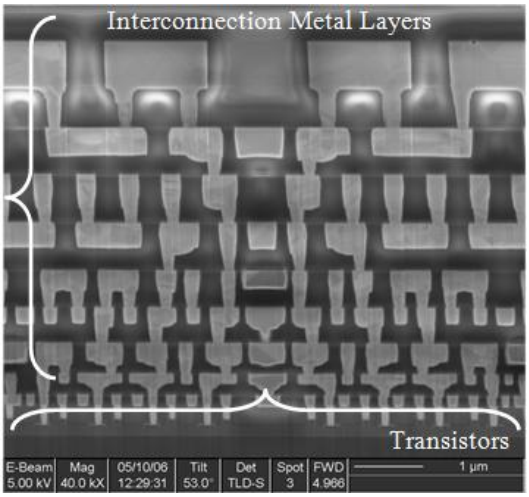


Figure 1.5 - Cross section of an electronic chip (transistors and interconnects)

The transistors and interconnects are the geometric shapes which are fabricated through repetitive sequence of deposition, lithography, etching, polishing, measurement, SPC etc. process operations. It is known that the process imperfection results in drift of geometric shapes of transistors a.k.a. devices and interconnects which can ultimately lead to the product failure. The DFM methods are focused on finding such potential drifts and its subsequent resolution by the designer before it enters in the production line. It is quite evident that the role of DFM methods is highly critical as it allows only manufacturable and yieldable designs to move on. In the design flow, we can simply say that the DFM rules and/or models are used to distort the drawn features followed by electrical characterization to ensure the product functionality.

The DFM is inserted in the design flow upon circuit layout completion by the designer, where geometric

shapes are drifted based on the DFM rules and/or models for a given technology. The electronic design automation tools are then used to characterize and extract the electrical and parasitic parameters. The design and wire layout is optimized until these parameters comply with the technology and product specification. To ensure volume production, the DFM methods are used to optimize product masks where geometric shapes are compensated against process drifts to keep printed layout as close as possible to the optimized layout that passed the CAD simulations.

A chip is manufactured on a wafer (Figure 1.6) made from silicon (Si). The wafer is divided in horizontal/vertical lines crossing each other known as the scribe lines. These scribe lines serve dual purpose: *(i)* they contain test structures used for metrology and/or inspection and *(ii)* they are used to cut the wafer and separate individual dies (chips). The notch is a cut in the wafer and it is used to describe the crystal orientation and wafer position during process and metrology operations. A site (field) is composed of individual dies whereas the number of dies in a site is characterized by the product mask. In order to monitor intra-die variations the test structures could be placed within the field at different positions. The electronic product undergoes metrology and inspection steps to ensure product quality where decisions are made based on the parametric yield to either scrap or move the wafer to next steps. These measurements are done on the test structures in scribe lines or fields; however, the product itself goes for a functional test at the end of the manufacturing process to sort the bad and good chips.

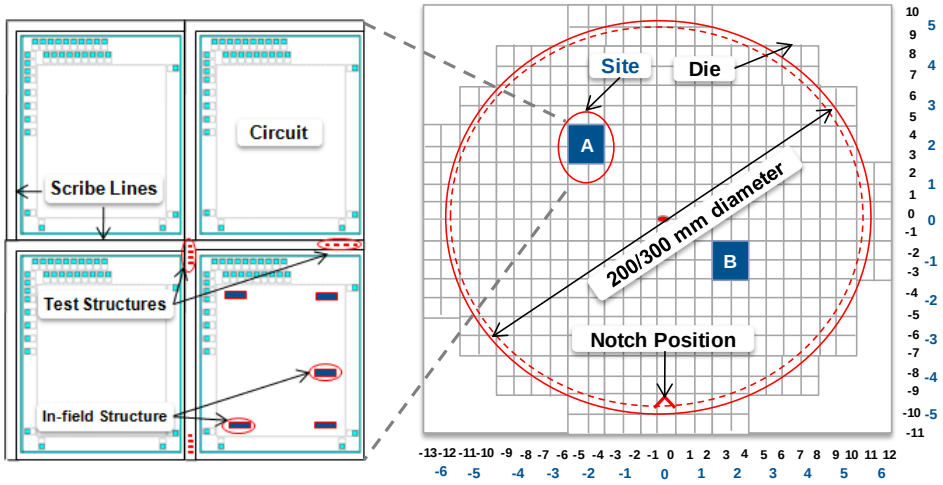


Figure 1.6 - Structure of the product wafer

Let us discuss the basic DFM concept and its evolution to the current challenges (Figure 1.7 and Figure 1.8). The designers design a new IC chip and send for manufacturing to assess its manufacturability. The design is then improved based on the recommendations by the manufacturing plant because process imperfection results in printed layout which is deviated from the drawn layout. It results in design respins, delays and costs. To avoid and/or minimize these respins, we simulate the designed chips using design rules and models on the manufacturability, yield and cost criteria. The design rules, DFM rules and models are extracted from the data collected across the production line against significant drifts and variations and this process is equally applicable for the technology alignment and adoption efforts. The success of SI lies in our ability to make the feedback loop more efficient and effective so that newly emerging spatial drifts and variations are quickly analyzed based on huge data volumes collected across the production lines. It shall help in the technology lead times and costs reduction.

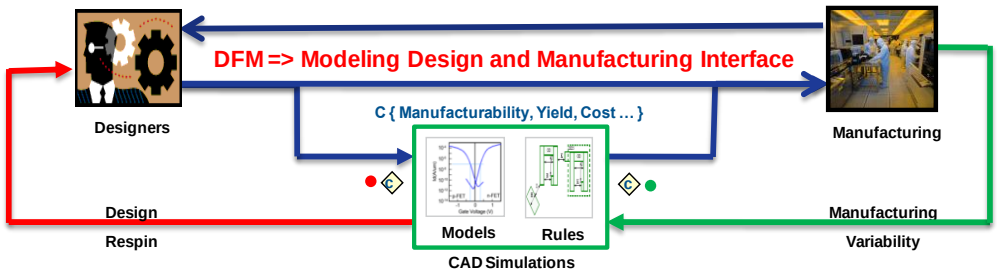


Figure 1.7 - The role of DFM in SI

The concept equivalent to "DFM" was coined by Mr. LeBlanc (French) and Mr. Eli Whitney (American) in 1778 and 1788 respectively [Dummer, 1997] by proposing a system for the production of musket. It received an industry wide recognition as "producibility through interchangeable parts". Mr. Roger W. Bolz is credited for organized DFM methodology [Bolz, 1958] as an alternative term for "producibility", introduced in his book "The producibility handbook"; however **D**esign for **M**anufacturability "DFM" received industry wide acceptance around 1960 [Boothroyd, 1968]. Let us formally extend the DFM concept. In the SI it is defined as the ability to reliably predict downstream life cycle needs and issues during early phases of design [Herrmann et al., 2004]. It is focused on economic benefits from the volume production by trading off cost-quality-time triangle [Raina, 2006] and is classified as [Mehrabi et al., 2002] product DFM (producing manufacturable designs within defined processes) and process DFM (developing processes with less rework and high manufacturability). The most appropriate classification of DFM methods in the SI is physical and electrical DFM [Appello et al., 2004]. The physical DFM refers to the process variations that result in geometric shape drifts during manufacturing whereas electrical DFM is focused on the characterization of parametric and functional product yields. The parametric characterization refers to the extraction of key electrical and parasitic parameters whereas functional characterization mainly refers to the signal timing and delays that result in the faulty products to ensure product functionality.

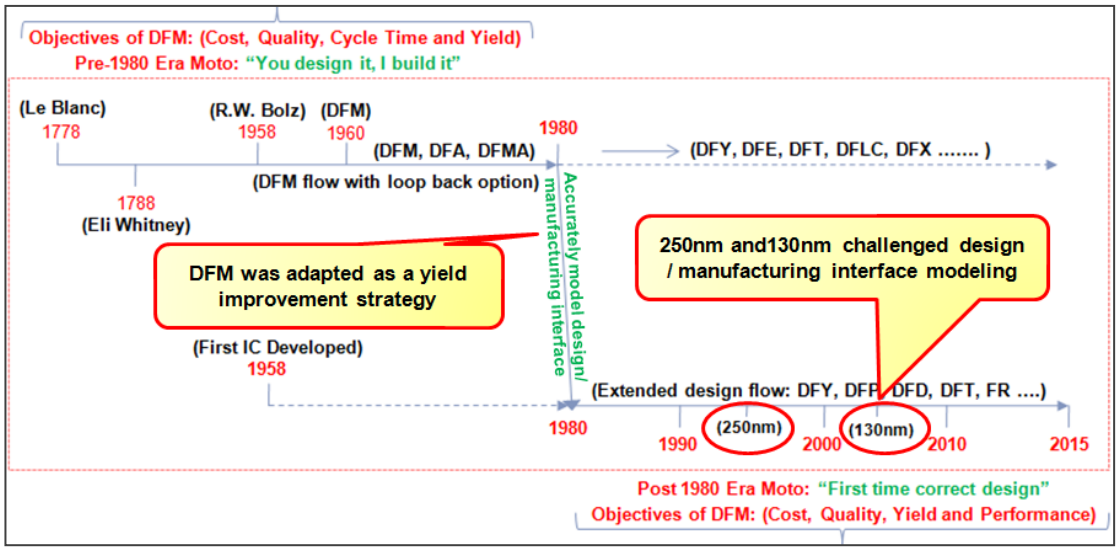
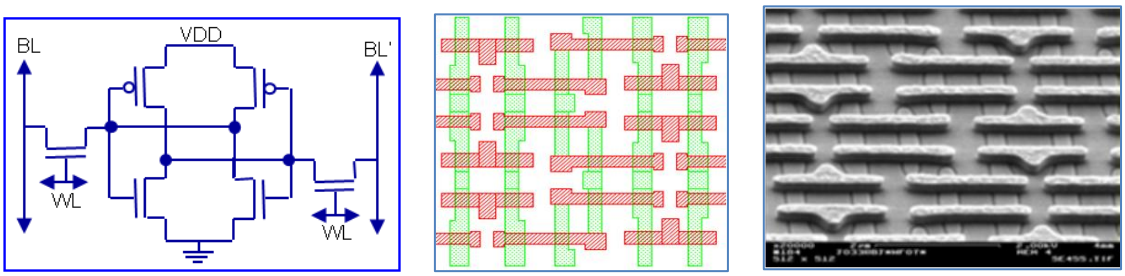


Figure 1.8 - Historical evolution of the DFM

It is evident that most of the manufacturability and yield issues are related with the process imperfection and/or equipment drifts due to which we are not able to accurately print the design layout on silicon wafer (Figure 1.9). Such drifts either result in electrical failures (electrical DFM) or physical failures (physical DFM) ultimately resulting in product failure. The process imperfection leads to the manufacturability and yield loss mechanisms which are classified as systematic or random. The random fault mechanisms can neither be modeled nor controlled but they can be minimized by following robust/recommended rules, however the systematic fault patterns can be transformed into rules and/or models for their subsequent use during CAD simulations. Please refer to section-B.6 of Annexure-B for detailed understanding of most common manufacturability and yield loss mechanisms.



a) IC chip design b) Design layout c) printed layout

Figure 1.9 – Drifts in drawn features and printed images

In 1980 the DFM concept was adapted as a yield enhancement strategy in the SI (Figure 1.8). It went very well till 250nm technology [Cliff, 2003 and Radojcic et al., 2009] but after that the increasing complexity of the circuit layout and shrinking sub wavelength lithography resulted in multiple respins and yield losses with 193nm stepper and 130nm node (Figure 1.10). The introduction of the compensation techniques like **O**ptical **P**roximity **C**orrection (OPC) and **R**esolution **E**nhancement **T**echnique (RET) emerged as an extended design flow (DFM flow) to mitigate the yield loss mechanisms. These methods are used during the mask data preparation to ensure manufacturability and yield. Beyond this point the manufacturability and yield losses can only be controlled through process control and recipe adjustments, however new systematic drifts patterns in the printed design layouts resulting in parametric and functional yield losses can be modeled for the technology improvements.

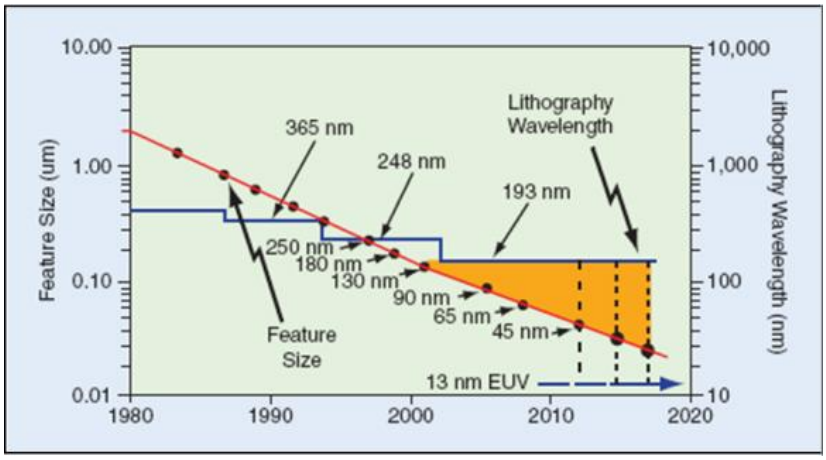


Figure 1.10 - Lithography and feature size [Radojcic et al., 2009]

From above facts it is evident that stretching the CMOS technology below 45nm size is difficult without extending the traditional DFM methodology. The DFM has evolved from design rules to DFM rules like (layout/routing rules) and DFM models like **C**ritical **A**rea **A**nalysis (CAA), **C**hemical **M**echanical **P**olishing (CMP), Shape, Yield, Leakage and [McGregor, 2007] **S**tatistical **S**tatic **T**iming **A**nalysis (SSTA) to mitigate the potential yield losses. Our efforts are ideally focused to produce first time correct design that can be ramped without manufacturability issues and yield loss. The reasons behind the unsuccessful DFM implementation [Seino et al., 2009] are lack of awareness on the importance of DFM by the product designers [Ahmed and Abdalla, 2000], less understanding of designs influence on the manufacturing, wrong variations analysis, inability to perform multi-source root cause analyses and different perceptions of the engineers/managers. There is one thing common in these DFM methods and it is "Data".

In this thesis we are primarily focused on improving effectiveness of the R&D effort so that DFM can be put back on track in the technology alignment and adoption processes to exploit the economic benefits. Let us formally define the potential DFM challenges within technology alignment and adoption processes (Figure 1.11). In technology alignment we are focused on process alignment, based on test (representative) products, where data captured across the production line is analyzed using DFM methods (R&D efforts) to find root cause against the yield limitations and model to hardware gaps. These root causes are further classified as systematic or random and transformed into rules and/or models. New rules and/or models are fed back to the technology models for their subsequent use in CAD simulations; however processes are also improved to achieve the target yields.

The design libraries are qualified as per new rules and models (design rules, DFM rules, models) for their subsequent use by the designers as reusable components to reduce the technology lead times. In technology adoption process, we use product prototypes and follow the process alignment link like the one we have seen in the technology alignment in the above paragraph. The data captured from the process is analyzed with **M**anufacturing **F**or **D**esign (MFD) methods a.k.a. APC/AEC methods to find causes against the drifts and yield losses. The MFD methods differ from the DFM in the sense that they are focused on the process, equipment or recipe adjustments to ensure the manufacturability of a given design whereas DFM provides us with new rules and/or models to ensure a first time correct design. The feedback to the process results in quick and rapid technology adoption for given yield

targets, however the feedback link to improve the technology do not exist because design kits are frozen at this stage and any proposed change shall need requalification of all design libraries which is costly and time consuming.

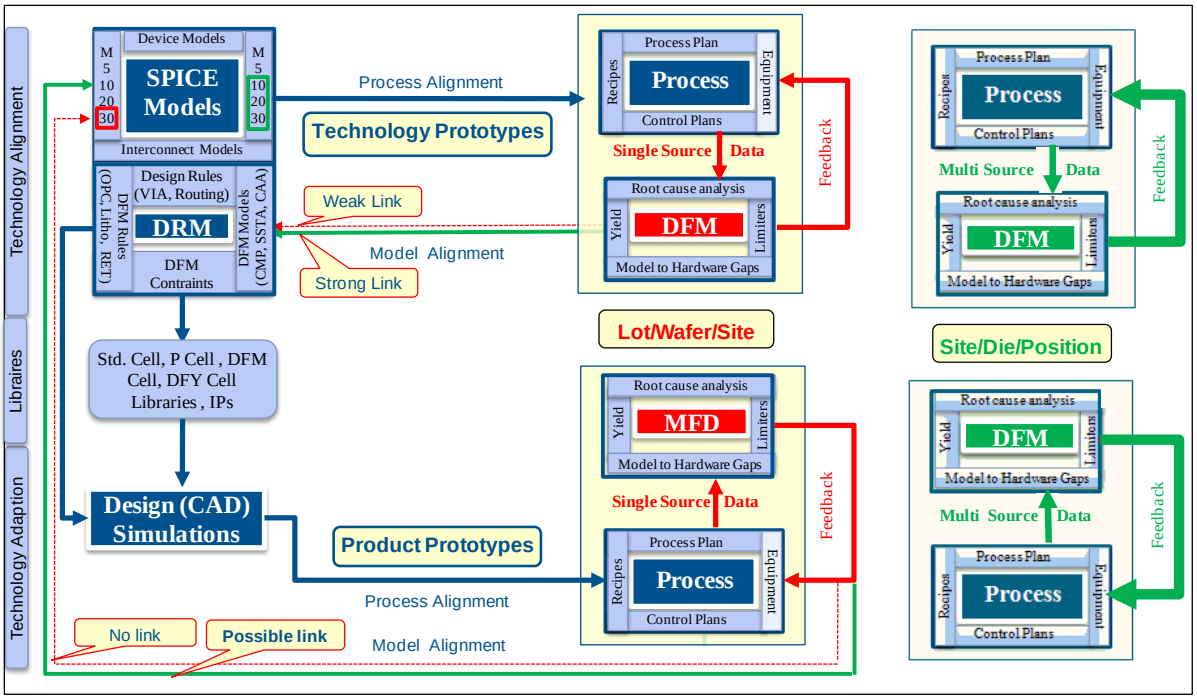


Figure 1.11 - The role of DFM in technology alignment and adoption processes

It is evident that the effectiveness of R&D efforts depends on the input data which at present is the single-source (lot/wafer/site); hence, our engineers are not able to exploit the huge data volumes collected across the production line. The goal of this thesis is to find and remove the limiting factors associated with the single-source root cause analysis and provide generic scientific contributions to enable multi-source (lot/wafer/site/die) and test structure position based dynamic data exploitation. It shall result in efficient root cause analysis and effective DFM methods. If the analyses results are quickly transformed into rules and/or models followed by its inclusion in the technology models, the designs shall result in higher yield and manufacturability. It shall also result improvements in technology alignment and adoption lead times and costs. So in order to put DFM back on track we suggest *(i)* a shift from MFD to DFM efforts and *(ii)* single-source to multi-source site/die/position based root cause analysis within the technology alignment and adoption processes.

1.4 RESEARCH QUESTIONS

We have seen that the DFM ineffectiveness has resulted in an exponential increase in the technology lead times and costs. The 8.78% CAGR is the main reason behind increasing R&D investments in compliance with the Moore's law and the industrial slogan smaller, faster and cheaper. It has resulted in the shift of SI business objectives from time-to-market (T2M) and time-to-volume (T2V) towards ramp-up rate a.k.a. time-to-quality (T2Q) to ensure economic benefits. The traditional IDM has transformed into IDM-fablite and fabless business models to achieve SI business objective by reducing technology lead times and R&D costs. This new structural transformation has ensured the introduction of new technologies every 2-3 years but it has highlighted ineffectiveness to continuously improve these technologies with local DFM efforts. It is because of the fact that besides the availability of huge data volumes and dimensions collected across production line, we are unable to find drift/variation patterns, classify them as systematic or random and transform them into rules and/or models against emerging manufacturability and yield loss mechanisms. So our first basic research question is to analyze the existing business models and find out if they are able to support the shift in the business objectives knowing that, at present we are not able to continuously improve new technologies which are developed in technology alliances. The formal first research question is presented below:

Q1: *What are the top ranked business objectives in the SI, today? What is the best strategy to achieve these top ranked objectives? Do we have a business model coherent to achieve these objectives, if no then what it should be?*

We have seen that new technology, which is developed in an alliance, is transferred into the business model where it follows technology alignment and adoption processes. The DFM methods can play a significant role in the alignment and adoption that has become ineffective resulting in increased lead times and costs. It depicts our inabilities to model the emerging drifts and variations, besides the availability of huge data volumes and dimensions (multi-source). We believe that if we can find and fix the limitations in multi-source data exploitation, it shall result in more effective DFM methods. So, the second research question is to identify the key limitations that result in our inability to dynamically exploit the production data sources for R&D purposes. The second research question is formally presented as under:

Q2: *What are the true DFM challenges (limiting factors and failure modes) and respective root causes within technology alignment and adoption processes?*

The identified limitations are further investigated to find relevance with an organization or a domain; hence, the third research question is about generic solutions (scientific contributions) to solve the problems associated with dynamic exploitation of production data sources. The third research question is formally presented below:

Q3: *What are the generic solutions to remove these root causes and put DFM back on track for the technology lead time and costs improvements?*

Finally, it is important to assess the potential industrialization and post industrialization challenges of proposed generic solutions. The proposed solutions provide an opportunity to exploit multi-source data; hence, it raises need for more metrology and inspection data for R&D purposes. It is likely to reduce the metrology/inspection capacities reserved for normal production. We cannot buy new tools to increase the capacities (fixed cost); hence, we need an intelligent way to spare metrology/inspection capacities for R&D purposes. The fourth research question is formally presented below:

Q4: *What are the consequence of proposed solutions industrialization and how we can resolve it to allow smooth integration of the proposed solutions?*

1.5 RESEARCH METHODOLOGY AND SCHEMATIC

In this section we graphically present the methodology and timeline (Figure 1.12) as a block diagram. It summarizes that how the research was carried out along with scientific and industrial contributions.

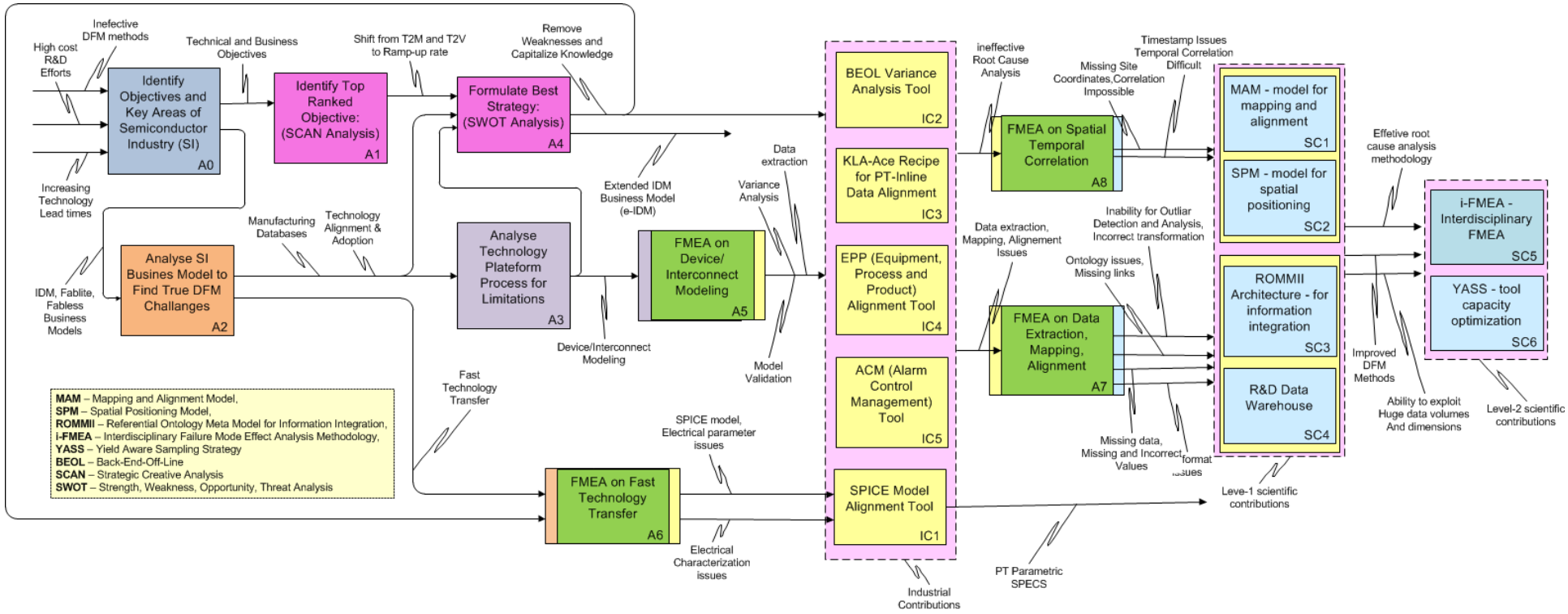


Figure 1.12 - The Research schematic and contributions at a glance

We started with brainstorming sessions (A0) with engineers and managers at STMicroelectronics and highlighted potential business (T2M, T2V, ramp-up-rate) and technical (area, power, timing, leakage) objectives. The SCAN analysis (A1) resulted in leadership position and ramp-up-rate as the top ranked objectives and traditional IDM business model analysis (A2) resulted in the *(i)* fast technology transfer, *(ii)* manufacturing databases and *(iii)* effective root cause analysis (R&D) as key improvement areas. These findings were used as input in SWOT analysis (A3) and we agreed on the strategy to remove weaknesses and capitalize the knowledge. As we know that the technology lead times and costs are the key factors for success in SI; hence, we analyzed technology derivative and improvement processes (A4) to find limitations within the scope of key improvement areas as identified in A2. The key failure modes were found as *(i)* data extraction, *(ii)* alignment and *(iii)* pre-processing due to ontology issues and missing database links. This analysis resulted in the proposition of an extended IDM-fablite (*e*-IDM) business model with integrated DFM methodology for quick ramp-up-rate and leadership position. The analyses blocks A0→A4 shows that existing IDM, IDM-fablite, fabless and foundry business models do not support recent shift in the business objectives (ramp-up rate); hence, we propose an extended IDM-fablite (*e*-IDM) business model to achieve quick ram-up rate. It also improves the internal manufacturing window while keeping intact its backward compatibility with technology, initially developed in an alliance. It provides an answer to the first research question.

We have used a well known **F**ailure **M**ode **E**ffect **A**nalysis (FMEA) methodology to find potential root causes against initially identified failure modes (A4). The initial root causes identified are ontology issues, missing links between databases, missing values and varying measurement reference coordinates that restrict our engineers to single-source root cause analysis. It is because of the fact that wafer is often rotated prior to site/die level measurements due to test structure orientation in the scribe lines and fields. This rotation changes the x, y coordinates resulting in varying measurement coordinates which requires an accurate alignment prior to perform multi-source site/die level analyses. We developed 5 tools (industrial contributions: ICs) *(i)* **S**pice **M**odel **A**lignment (SMA) for fast technology transfer (IC1), *(ii)* **BEOL**-variance analysis to analyze parametric drifts in back-end-of-line interconnect modeling process (IC2), *(iii)* KLA-Ace recipe for the mapping and alignment of coordinates to enable multi-source site/die level analysis (IC3), *(iv)* **E**quipment, **P**roduct, **P**rocess (EPP) life cycle extraction tool to extract multi-source contextual data to support drift/excursion analysis (IC4) and *(v)* **A**larm **C**ontrol and **M**anagement (ACM) tool to extract, analyze and manage alarm data for tool capacity optimization (IC5). These tools were developed and provided at the disposition of **T**echnology to **D**esign (T2D), **E**quipment **E**ngineering (EE) and European project IMPROVE teams, being crucial for the successful and quick technology derivative improvements with local DFM efforts. The objective was to find new failure modes which are not evident at this stage.

While using these tools, the teams identified new failure modes that were taken into account and the tools were rectified and updated accordingly. This hybrid approach helped us to find the core failure modes: *(i)* unstructured data model evolutions, *(ii)* missing data dimensions and *(iii)* wrong correlation due to test structure positions. These failure modes are grouped as *(i)* ineffective root cause analysis (infield and scribe line test structure positions) and *(ii)* data extraction, mapping and alignment, were identified to propose generic solutions (A6 and A7). We also performed FMEA on the technology transfer step and found issues between SPICE model and measured electrical parameter relationship. The SPICE models are the mathematical equations used to extract the electrical parameters based on the transistors and interconnect geometries and process technology selected for the manufacturing. The name of the parameters in these SPICE models are generic whereas the names used by the test engineers while writing test programs for the test structure on the wafer are different; hence, the mismatch creates a significant problem and delays the model validation resulting in extended lead times. The initial root causes that were found are incorrect and error prone SPICE-PT parameter alignment due to manual data alignment and pre-processing; hence, we initially developed SMA (spice model alignment) tool (IC1) that resulted in automation and removed the associated issues. The further use of this software tool identified unstructured naming conventions for the metrology parameters and varying formats of the CAD simulation results as the key root causes. These industrial contributions IC1 to IC6 are added in appendices for reference. The analysis blocks A6-A7 and industrial contributions IC1→ IC5 partially answers the second question, established in section 1.4.

A detailed FMEA analysis is performed on the *(i)* ineffective root cause analysis and *(ii)* data extraction,

mapping and alignment failure modes. The resulting root causes are used to propose generic solutions a.k.a. scientific contributions (SCs). The key root causes linked with "ineffective root cause analysis" failure modes are *(i)* missing and varying coordinate due to varying metrology reference frames, *(ii)* rotation of wafer prior to the measurement steps and *(iii)* the infield/scribe line test structure positions. A generic MAM (mapping and alignment) model (SC1) is proposed to remove die/site level mismatches and provide means for an accurate correlation. The SPM (spatial positioning) model (SC2) is also proposed to enable correlation between parameters, based on the shortest distance between test structures, used for metrology. The ROMMII (relational/referential ontology Metal model for information integration) framework (SC3) and R&D **D**ata **W**arehouse (DWH) model (SC4) are proposed to address the causes identified against data extraction, mapping and alignment failure modes. The word mapping in this failure mode refers to the mapping between multiple databases (links). It provides the ability to perform pre-failure analysis on the potential impact of any structural change in data model over existing users and applications. These scientific contributions (SC1 to SC4) refer to the question-3 whereas the identified failure modes and root causes complement the answer to question-2. It is very important to note that root causes identified fall in different business functions and in our case it is **I**nformation **T**echnologies (IT) and **E**ngineering **D**ata **A**nalysis (EDA), so based on the experience during PhD, we have also proposed a 4-step *i*-FMEA (interdisciplinary failure mode effect analysis) approach (SC5) to capture true challenges that might fall in other business functions. The *i*-FMEA methodology completes our answer to question-2.

At the end, a brain storming session on the industrialization of proposed scientific contributions highlighted potential impacts. It was found that it shall result in extended analysis and metrology demands by R&D engineers for an effective root cause analyses. It is likely to reduce inspection/metrology tools capacities for normal production lots; hence, to address this issue a yield aware sampling (YASS) strategy (SC6) is proposed which is based on information fusion (alarms, states and meters data) to intelligently predict the production lots with likely yield loss. So the objective is to inspect bad or suspected lots and allow good lots to move to the next production steps. The additional metrology and inspection capacities can be then used for the R&D purposes. This YASS strategy provides answer to the question-5.

1.6 MAJOR CONTRIBUTIONS

Let us summarize the major contributions made during this thesis. The contributions are divided in two categories as *(i)* industrial contributions and *(ii)* scientific contributions.

1.6.1 Analysis of Overall System and Industrial Contributions

We made 5 industrial contributions in the form of software tools for R&D engineers as *(i)* BEOL-variance analysis, *(ii)* KLA-Ace Recipe, *(iii)* EPP (equipment, product, process) life cycle tool, *(iv)* SMA (spice model alignment) tool and *(v)* ACM (alarm control and management) tool. A brief description of these contributions is presented as under:

a) **SMA (spice model alignment) Tool (IC1):** This tool is developed as an operational fix for the process integration (PI) team, working for technology transfer, alignment and adoption. It is developed to remove the causes associated with the failure modes identified and presented in Chapter-4. It facilitates R&D engineers in aligning the SPICE model parameters against measured electrical parameters and generates SPECS based on manufacturing capabilities.

b) **BEOL Variance Analysis Tool (IC2):** This tool is developed for **T**echnology **t**o **D**esign (T2D) team to support the device/interconnect modeling process by quickly pre-processing data followed by parasitic variance analysis that result in the parametric and functional yield losses. The significant variations are further analyzed using KLA-ACE recipe for the PT-Inline site level correlations to find root causes against these drifts.

c) **KLA-Ace Recipe *for* PT-Inline Correlation (IC3):** In this recipe we enable multi-source correlation using PT-Inline data sources by mapping the data at site levels. Both of the data sources, **P**arametric **T**ests (PT) and Inline measurements are captured at Site level. The PT is a type of electrical

measurements whereas Inline corresponds to the physical measurements. The PT database has site numbers and x,y coordinates against each measurement; however, the Inline data has only site number. These site numbers cannot be used for site level mapping because during measurements the Wafer is rotated based on the position of the test structure; hence, every measurement has different coordinates. It means that until and unless we have wafer position defined by notch (a cut on the wafer edge used for alignment) and its x,y coordinates, site level mapping cannot be performed; hence, multi-source root cause analysis is not possible. In this recipe we use the mask level information to normalize the site level identifiers so that quick mapping can be performed to provide engineers an ability to perform multi-source root cause analysis.

d) **EPP (Equipment, Product, Process) Life Cycle Tool (IC4):** The results obtained from the BEOL and KLA-Ace recipe justify the parasitic drifts against geometric specification variations; however, it does not provide an answer if the drift was caused due to process or equipment variation. In order to perform an in depth analysis we developed this tool that exploits the manufacturing data sources and generates product and equipment life cycles. The EPP tool directly connects with the maintenance (TGV) and out of control (OOC) databases for the equipment related data extraction whereas product and process data is extracted from the process database using KLA-Ace recipes. All these data are input to the EPP tool that perform consistency checks and populates them into a multidimensional database. It provides a user friendly interface to extract customized equipment and product life cycles.

e) **ACM (Alarm Control and Management) Tool (IC5):** This tool is developed for the lithography equipment engineering team. At present engineers do not have information about alarms categorization based on machine states; hence, we have linked INGENIO (equipment alarms) database with maintenance and process databases to classify alarms based on the machine state. The extracted data is further used to develop predictive models for the likely yield loss across the production processes with an objective to skip good lots and inspect the lost with suspicion of bad yield. It is used in the YASS (yield aware sampling strategy) as presented in section 1.6.2e.

These tools have helped us in finding true DFM challenges (failure modes) and associated root causes as a part of our proposed *i*-FMEA methodology. The brief description for these tools is presented in the appendices D to H.

1.6.2 Scientific Contributions

In this thesis, 6 scientific contributions are proposed as *(i)* MAM model, *(ii)* SPM Model, *(iii)* ROMMII platform and R&D DWH Model, *(iv) i*-FMEA methodology and (v) YASS strategy. A brief description is presented as under:

a) **MAM (mapping and alignment) Model (SC1):** It is a generic model for site/site and die/die mapping of metrology data along with die/site qualification. The objective is to enable our R&D engineers to perform multi-source root cause analysis to find root cause against newly emerging spatial drifts and variations. It is presented and discussed in detail in Chapter-5.

b) **SPM (spatial positioning) Model (SC2):** It is a fact that all metrology measurements are performed on test structure and these test structures are located in scribe lines and/or in the fields. To capture and better understand newly emerging spatial variations, it is likely to perform analysis using measurements coming from the test structures with shortest distance. This model performs mapping on different data sources based on the test structure position.

c) **ROMMII (referential ontology Meta model for information integration) Architecture and R&D Data Model (SC3 and SC4):** The ROMMII architecture provide a strong control over unstructured data model evolution and at the same time flexibility to continuously evolve the data

models to include new data dimensions. The R&D data model is proposed to support the need of R&D engineers to ensure that 1 year R&D data is always available for analysis purposes.

d) ***I*-FMEA (Interdisciplinary failure mode effect analysis) Methodology (SC5):** It is 4-step methodology which is primarily based on the traditional FMEA approach but subjectively focused on finding cyclic failure modes and root causes. We conclude that the cyclic failure modes and root causes must be addressed with the generic R&D solutions (SC1 to SC5) than operational fixes (IC1 to IC5). In our case, highlighted cyclic failure modes and root causes are classified as the true DFM challenges requiring generic R&D solutions. It enables the R&D engineers to extract accurate and timely value by dynamically exploiting the available huge data volumes and dimensions. The result is a shift from data driven ineffective DFM efforts towards information and knowledge driven DFM initiatives.

e) **YASS (yield aware sampling strategy) Strategy (SC6):** This intelligent sampling strategy predicts the production lots as good, bad or suspected lots based on the predictive state and alarms model. These models are learned from data extracted using ACM tool. The objective is to increase the metrology capacities so that it can be spared for the R&D purposes. It is important to support the industrialization of our proposed generic scientific contributions because of the fact that ability to model abnormal drifts shall result in huge demand for more metrology for R&D analysis.

1.7 THESIS ORGANIZATION

This thesis is divided into 7 chapters and 8 appendixes. The contents of this thesis are organized so that the readers do not have to read all the chapters to get the information they need. Therefore, some concepts are repeated across several chapters. The following list briefly describes each chapter and appendix:

- **Chapter-2** [***Literature Review***]: We present a comprehensive literature review across 3 dimensions *(i)* semiconductor industry and challenges *(ii)* role of DFM methods in SI and *(iii)* information integration challenges. We suggest that readers with little or no background of semiconductor design must go through appendices B and C where we have discussed in detail the semiconductor design and manufacturing processes with an example of simple CMOS inverter design. The objective is to introduce the readers with potential design and manufacturing interface complexities and emerging manufacturability and yield loss mechanisms.
- **Chapter-3** [**An Extended IDM (*e*-IDM) Business Model**]: In this chapter, SI business models and objectives are analyzed followed by the identification of key improvement areas. The SCAN and SWOT analysis techniques are used to find a strategy to exploit opportunities. The technology development process is further analyzed for the identification of limiting factors that result in DFM ineffectiveness. The challenges found here along with the key improvement areas forms the basis for an extended IDM (*e*-IDM) business model with the improved DFM methods.
- **Chapter-4** [***i-FMEA Methodology for True DFM Challenges***]: The objective of this methodology is to identify true DFM challenges against those found in the literature review. A 4-step *i*-FMEA methodology is proposed to find root causes against failure modes in other business functions. The identified root causes are grouped as *(i)* ineffective root cause analysis and *(ii)* data extraction, mapping and alignment which serve as basis to remove DFM ineffectiveness and ensure the success of our proposed *e*-IDM business model.
- **Chapter-5** [***MAM- SPM Models for Mapping, Alignment and Positioning***]: In this chapter, the two groups of root causes "ineffective root cause analysis" and "mapping and alignment" are addressed that directly impacts DFM ineffectiveness and restrict R&D engineers to single-source wafer/die/site level analysis. The MAM and SPM models are proposed to remove die/site level mismatches and enable a multi-source die/site and position based root cause analysis.
- **Chapter-6** [***ROMMII Architecture and R&D Data Model for Information Integration***]: The second group of root causes "data extraction" is addressed here and ROMMII (referential/relational ontology Meta

model for information integration) framework is proposed that allows the data model evolution and enables us to exploit huge data volumes as well as inclusion of new data dimensions. The proposed R&D data model removes different data retention issues with multiple production databases.

- **Chapter-7** [***YASS (yield aware sampling strategy) for tool capacity optimization***]: The industrialization of proposed contributions (chapter-3, chapter-5 and chapter-6) does result in new demands for the additional metrology to capture and model variations. It is quite difficult in an IDM business model to spare metrology/inspection capacities for R&D purposes; hence, the YASS strategy spares tool capacities for R&D. The proposed strategy advocates and facilitates industrializations of e-IDM fablite, MAM, SPM and ROMMII contributions.

We conclude our thesis with discussion and critical analysis on the proposed solutions against alternative solutions. We also present some industrial recommendations for SI on the potential industrialization of the proposed scientific contributions and end up with future research directions to improve the DFM methods.

- **Appendix A** (Publications): provides a list of publications in international conferences and journals along with abstracts and keywords.
- **Appendix B** (Semiconductor Design, Mask and Manufacturing Processes): provides a brief review on design, mask preparation and manufacturing processes along with key challenges and limitations and most common manufacturing and yield loss mechanisms.
- **Appendix C** (CMOS Inverter Design and Manufacturing: practical example): simulates the CMOS inverter design, masks and manufacturing steps.
- **Appendix D** (SMA-Tool): presents a brief description of SMA tool used by the PI team to support fast technology transfer efforts.
- **Appendix E** (BEOL-Tool): presents a brief description of the BEOL-Tool for T2D team to support quick variance analysis during interconnects modeling process.
- **Appendix F** (KLA-Ace Recipe): presents a KLA-Ace Recipe that uses the mask data to complement the missing x, y coordinates for inline data and enables PT/Inline site level correlation analysis.
- **Appendix G** (EPP-Tool): presents brief description of the tool that provide product and equipment life cycle data extraction, used by IMPROVE project team for the computation of **E**quipment **H**ealth **F**actor (EHF).
- **Appendix H** (ACM-Tool): presents brief description of the tool that provides extraction of alarms and states data to be used for PAM and PSM prediction models. It is used by the equipment engineers to analyze and manage the alarm controls to identify the alarms to be controlled on priority.

This thesis ends with a list of references used during research and an index of important terms. To help the readers, the first paragraph of each chapter provides a synopsis of that chapter's contents.

1.8 TYPOGRAPHIC CONVENTIONS

The typographic conventions are used in the thesis to improve users' readability. A comprehensive list of acronyms that appear throughout this thesis is provided and if an acronym appear in the thesis, it is not emphasized in bold type e.g. DFM, IDM, etc. For the standard concepts, if an acronym appears very first time then its initial letters are put in capital with acronym in brackets, e.g. **D**esign **F**or **M**anufacturing (DFM); however for the proposed concepts the acronym precedes the description e.g. MAM (**M**apping and **A**lignment **M**odel) etc. The citations are shown in italic type and between the quotation marks ("") e.g. "*A data warehouse is a subject oriented, integrated, time-variant, nonvolatile collection of data in support of management's decisions*".

We have tried to write each chapter as a whole so that they can be read independently; however there are situations in which we have referred or repeated contents to some extent that has previously appeared or that shall appear in the following chapter(s).

Chapter 2: Literature Review

The readers not familiar with the SI design and manufacturing flows are recommended to first read the appendices B and C. The appendix B provides a brief review on the design, mask preparation and manufacturing flows along with most common manufacturability and yield loss mechanisms. The appendix C presents a simulation of a simple CMOS inverter (2 transistors) design and manufacturing steps for an in depth understanding of the complexity in SI design and manufacturing processes. The literature review in this chapter is divided across 3 dimensions *(i)* the semiconductor industry and challenges *(ii)* the role of DFM methods in semiconductor industry and *(iii)* the information integration issues to support our methodology and scientific contributions.

Contents

2.1 SEMICONDUCTOR INDUSTRY (SI): BACKGROUND AND CHALLENGES

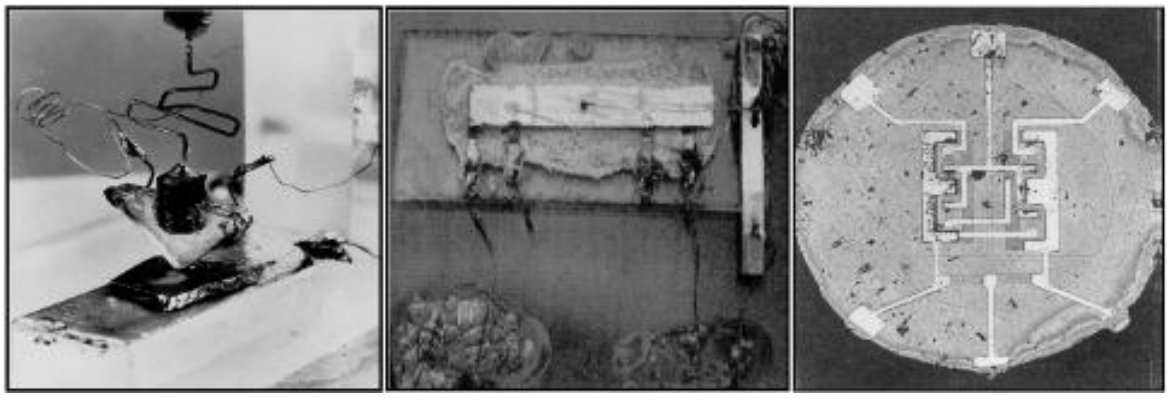
The Semiconductor industry (SI) is characterized by the fastest change in smallest period of time and has become a 300+ B$ industry in less than 60 years [Stamford, 2012 and Dale, 2012]. It has followed the industrial slogan smaller, faster and cheaper in compliance with Moore's law that predicted the doubling of transistor every 18 to 24 months [Moore, 1998]. This miniaturization has led serious design and manufacturing challenges which are resolved with the continuous introduction of new technologies, processes and production equipments at the cost of huge R&D investments. The increasing R&D costs have been compensated by increasing wafer sizes and reducing technology alignment and adoption lead times. The SI has seen many circuit integration eras and series of technology nodes to manufacture more complex integrated circuits. The DFM methods were initially introduced around 1980 as a yield enhancement strategy that worked very well up until the 250nm technology node but beyond this it became a high cost R&D activity. The newly emerging variations have resulted in increasing the technology alignment and adoption lead times and associated costs. We need to reduce these lead times so that associated R&D costs are reduced and high profit gains with early penetration into the market.

The objective of this section is to briefly review the history of the semiconductor industry, evolution, business models, current trends and major challenges. It shall provide us with an overview of the market and its influence on the semiconductor industry, further we shall also see the evolution of technical and business challenges that led the structural transformation of the IDM structure into fablite and fabless business models. This section shall highlight the need to analyze the existing evolutions and find or propose a business model (see chapter-1, section 1.4) to achieve the new shift in objectives from T2M and T2V towards ramp-up rate a.k.a. **T**ime-to-**Q**uality (T2Q), essential for success [Carrillo and Franza, 2006].

2.1.1 Historical Background

The electronics business is a trillion dollars industry that has developed over a 100 years in a number of steps. It started with the invention of 2 and 3 elements vacuum tubes by John Ambrose Flemming (1904) and Dr. Forest (1906) that led the development of radios and televisions, very first electronic products to be sold commercially [Dummer, 1997]. The invention of transistor (1947) and integrated circuits (1950s) revolutionized and led the industrial specialization as *(i)* consumer electronics and *(ii)* semiconductor industries [Braun and MacDonald, 1982]. A transistor is a semiconductor device that is functionally equivalent to a vacuum tube but is smaller, robust, operates at lower voltages, consumes less power and produces less heat. It is the building block for integrated circuits (ICs) in SI that allows integration of several circuits and has pushed electronics industry towards miniaturization. The electronics industry started with a 200 million $ market share in 1927 [Dummer, 1997] and it has seen an exponential growth to 1.04 T$ including 300+ B$ from the semiconductor industry in 2012 [Lefkow, 2012]. We are focused on the SI that lives and dies with the industrial slogan: smaller, cheaper and faster. Being smaller is a key towards success as it increases the number of transistor per unit area and speed up the current flow due to transistors proximity, resulting in more circuits being manufactured at the higher performance but reduced cost.

The emergence of the SI can be traced back to [Brinkman et al., 1997] John Bardeen, William Shockley and Walter Brattain that invented a solid-state transistor at Bell Telephone Laboratories (1947). In 1954, William Shockley set up Shockley Semiconductor Laboratory to industrialize the silicon transistors. Gordon E. Moore (a chemist), joined Shockley Semiconductors in 1957 but a group of 8 engineers including Gorden E. Moore and Robert Noyce left Shockley Semiconductors in 1958. They offered the idea of silicon based semiconductors to Sherman Fairchild, the founder of the Fairchild Camera and Instrument. They set up a new division called Fairchild Semiconductor [Tuomi, 2004] and started its operations with $3,500 capital (1957). The first integrated circuit was developed by Jack Kilby at the Texas Instruments (TI) in 1959; however the Fairchild lab is credited for the invention of the first commercial IC logic gate in 1961 by Robert Noyce [Brinkman et al., 1997]. These benchmarked revolutions are presented in Figure 2.1.



a) First Transistor (1947) Bell Labs | ***b)*** First IC Device (1959) Texas Instruments | ***c)*** Silicon IC Chip (1961) Fairchild Lab

Figure 2.1 - Benchmarks in semiconductor technology evolution [Quirk and Serda, 2000]

2.1.2 Evolution of Semiconductor Industry

The demand for semiconductors is mainly driven by end-user markets: data processing, consumer electronics, communications, automotive industry and industrial sector [Ballhaus et al., 2009]. The SI forms a part of this complex interaction among multiple industrial sectors [Yoon et al., 2010 and Kumar, 2008]. In general, the semiconductors demands do not generate directly from end users, but it is determined by the related end-customer market. We can analyze the market supply and demand by dividing it on product type, application segments and geographical regions.

Understanding market structure based on products is better understood as two categories as *(i)* non memory products and *(ii)* memory products (Figure 2.2). It is evident that memory is the biggest market in semiconductor till today. The **M**icroprocessor (MPU) and **M**icrocontroller (MCU) **U**nits have lowest relative market demand but these products are classified as high value products; hence return on investment is also high on these products.

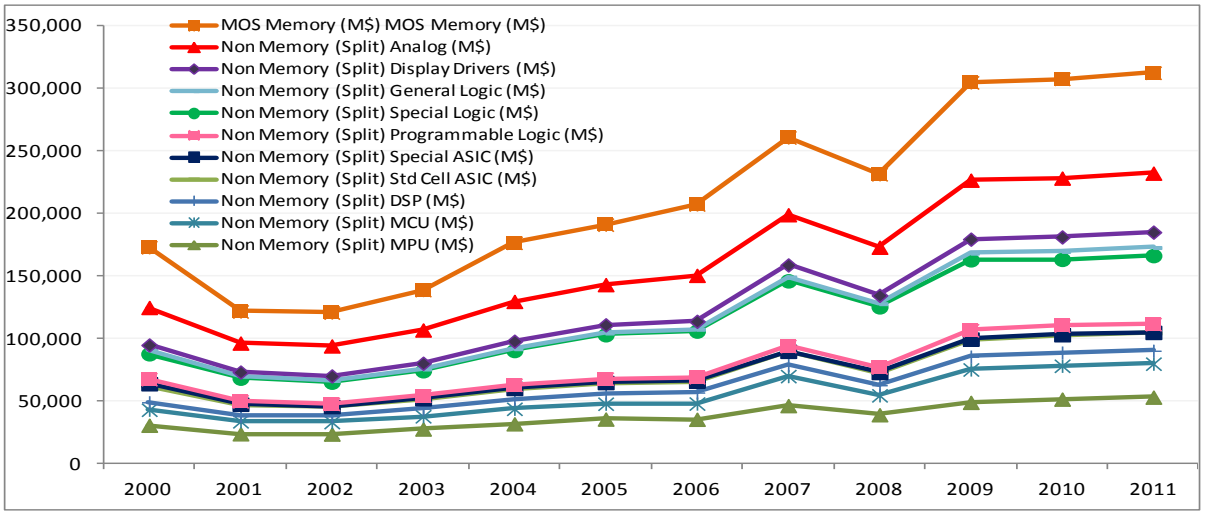


Figure 2.2 – Product based market structure[1]

The market structure with reference to the application segments is presented in Figure 2.3 that shows consumer electronics and wireless communication as one of the leading market segments, however automotive applications is the emerging market segment which is included in others category. At present, the automotive market is only 8% of the total SI market but is expected to dominate in the future.

[1] The data is used from International Business Strategies, Inc., report: Key trends in technology and supply for advanced features within IC industry (October 1, 2009), Gartner {www.gartner.com} and isuppli {www.isuppli.com}

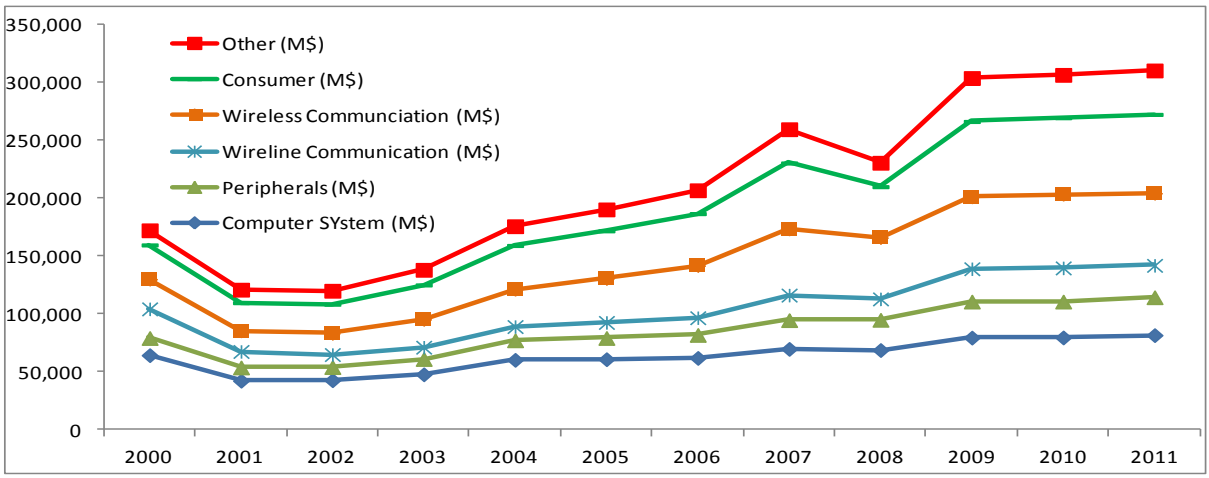


Figure 2.3 – Application segment based market structure[4]

The geographic position based segmentation (Figure 2.4) clearly shows a rising demand in China followed by North America. This geographic segmentation had influenced the semiconductor business model evolutions, to be discussed in section 2.2.

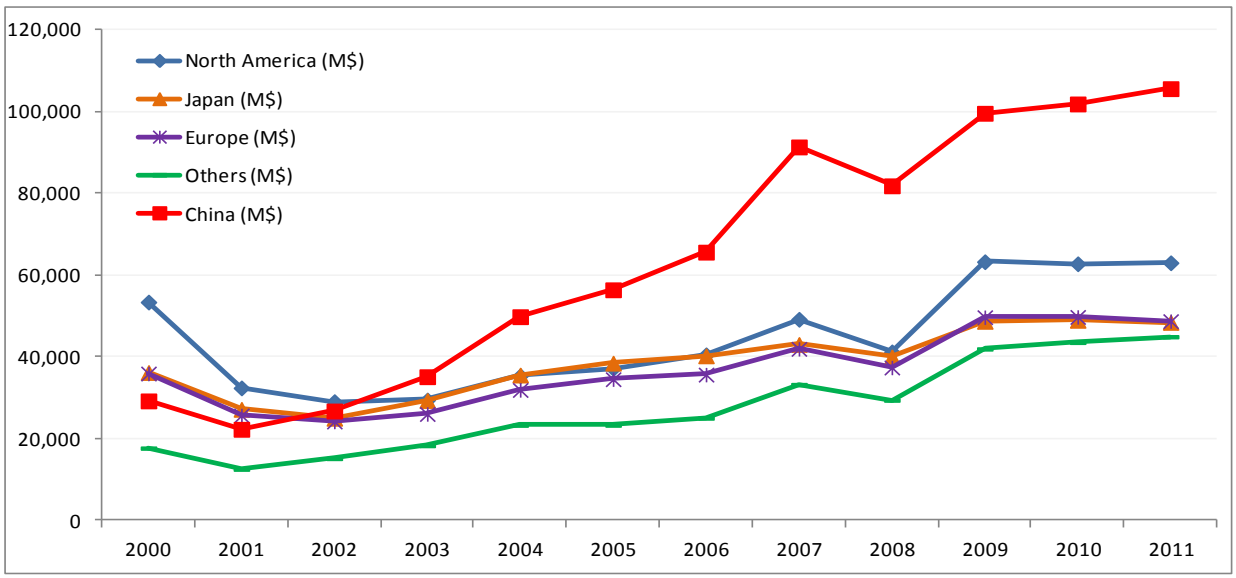


Figure 2.4 – Geographical position based market structure[5]

2.1.3 Role of Moore's Law in Semiconductor Industry

The technology scaling has been the primary driver behind improving the performance characteristics of IC's. Gordon E. Moore made an observation in 1965 that the number of transistors shall double in logic circuits every 24 months [Moore, 1998]. It was named as Moore's law of miniaturization by Caltech Professor Carver A. Mead; however this projection was altered by Gordon E. Moore himself in 1975 by changing the duration from 24 to 18 months. This projection proved accurate and was accepted as an industrial standard. Robert Noyce (1977) proposed that miniaturization is less likely to be limited by the laws of physics then by the laws of economics and he named it as the Moore's 2nd law, however Mr. Ross in 1995 stated that Moore's 2nd law shall end around 2005 [Ross, 2003].

[4,2] The data is used from International Business Strategies, Inc., report: Key trends in technology and supply for advanced features within IC industry (October 1, 2009) , Gartner {www.gartner.com} and isuppli {www.isuppli.com}

We saw that this projection proved wrong and Moore's 2nd law still prevails; however the cost associated with new technologies is exponentially increasing.

During last 50 years, the SI has followed the industrial slogan smaller, faster and cheaper that resulted in shrinking the device geometries, increasing chip sizes (Figure 2.5) and rapid improvements at reduced costs. The period from 1960 to 1970 is credited for the emergence of new technologies that lead the **S**mall and **M**edium **S**cale **I**ntegrations (SSI, MSI) but actual competition started around 1970 with **L**arge and **V**ery **L**arge **S**cale **I**ntegrations (LSI, VLSI) to capture maximum market share with the support of automation technologies that were introduced around 1980 [Quirk and Serda, 2000]. The circuit integration era beyond 1999 is marked as **U**ltra **L**arge **S**cale **I**ntegration (ULSI) and is attributed to the start of volume production efforts [Kumar, 2008]. These circuit integration eras are attributed to the Moore's law that has pushed the SI towards smaller, cheaper and faster devices.

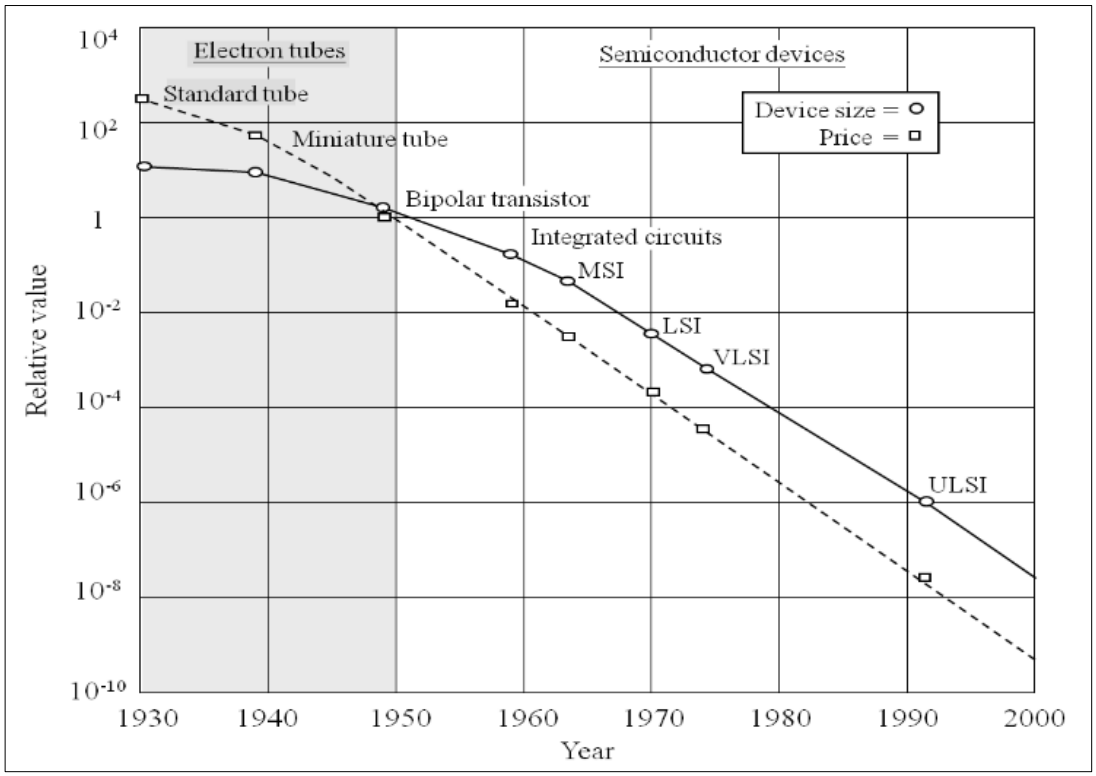


Figure 2.5 – Circuit integration eras and reducing costs (Chang and Sze, 1996)[6]

The well-known Moore's law has diversified into *(i)* equivalent scaling and *(ii)* functional scaling as presented in Figure 2.6 [Kahng, 2010]. The equivalent scaling is based on the principle of miniaturization that has led the emergence of **S**ystem **O**n **C**hip (SoC) where functionalities performed by different chips e.g. memory, **IO** (input output), processing unit etc. are put on the same chip to get higher performance at reduced area and cost. The major SoC application areas are information and digital contents processing e.g. memories (ROM, RAM, EEPROM, Flash) microcontrollers, microprocessors, **D**igital **S**ignal **P**rocessors (DSP) etc. It has resulted in the design complexity and requires new technologies to be regularly introduced to cope up with the pace defined by Moore's law. The functional scaling is an important area for the advancement of semiconductor technology. It is focused on grouping multiple functionalities (circuits systems) in a single chip (hetero-integration of digital and non-digital contents). These systems are also known as **S**ystem in **P**ackage (SiP). The current trends include functional and equivalent scaling at the same time which has resulted in the emergence of new type of systems called **P**ackage on **P**ackage (PoP). Multiple chips are stacked on one another in the **PoP** which give rise to more complex design and manufacturing interface complexities (variations). The most common examples for such systems are the mobile phones, digital cameras and **P**ersonal **D**igital **A**ssistants (PDA). These systems require innovation in terms of new devices as well as materials; hence results in the most complex manufacturing environment.

[6] Redrawn from C. Chang and S. Size, McGraw-Hill, ULSI Technology, (New York: McGraw-Hill, 1996).

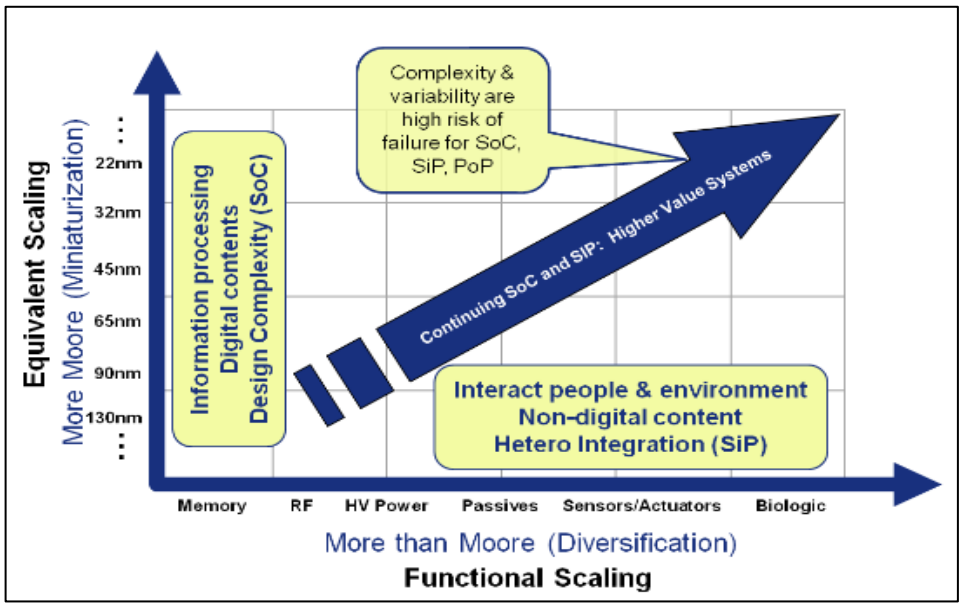


Figure 2.6 – Diversification in Moore's law [more Moore and more than Moore]

2.1.4 Nanometer vs. Micrometer Semiconductor Technologies

The integrated circuit components (transistor, capacitors, resistors etc.) have different sizes, but it is of interest to define the technology based on the smallest feature (dimension) size that can be manufactured on the silicon wafer. It is called node and it has reduced since 1965 every 2 to 3 years. The node sequences successfully achieved till today (Table 2.1) are: 1.5 µm, 1 µm, 0.8 µm, 0.6 µm, 0.5 µm, 0.35 µm, 0.25 µm, 0.18 µm, 0.13 µm, 90 nm, 65 nm, 45 nm, 32nn and 22nm. The two latest nodes are known as the leading edge, three immediately before as the mainstream and others as the oldest [Kumar, 2008].

Year	1988	1992	1995	1997	1999	2002	2005	2008	2011	2013
Feature size (nm)	1000	500	350	250	180	130	90	65	45	32
Mtrans/cm²					7	14-26	47	115	284	701
Chip size (mm²)					170	170-214	235	269	308	354
Signal pins/chip					768	1024	1024	1280	1408	1472
Clock rate (MHz)					600	800	1100	1400	1800	2200
Wiring levels					6-7	7-8	8-9	9	9-10	10
Power supply (V)					1.8	1.5	1.2	0.9	0.6	0.6
High-perf power (W)					90	130	160	170	174	183
Battery power (W)					1.4	2.0	2.4	2.0	2.2	2.4

Table 2.1 – CMOS technology scaling and characteristics

The terms used to represent the technology sizes (nm or µm) are always confused and there are no set rules to follow or express it. The term nanotechnology was coined by the national science foundation (NSF) in USA which states that the technologies used to manufacture devices with at least one dimension in the range below 100 nm shall be expressed and measured in nanometers. The micrometer scaled has been used to express the technologies till 0.13 µm but the important question is how small is the nanometer or micrometer? A comparative scale to imagine the size of the current technologies is presented in Figure 2.7. It can be seen that the smallest feature manufactured by the recent 32nm technology is even smaller than the length of a bacteria which is not visible with the naked eye [Jones, 2012].

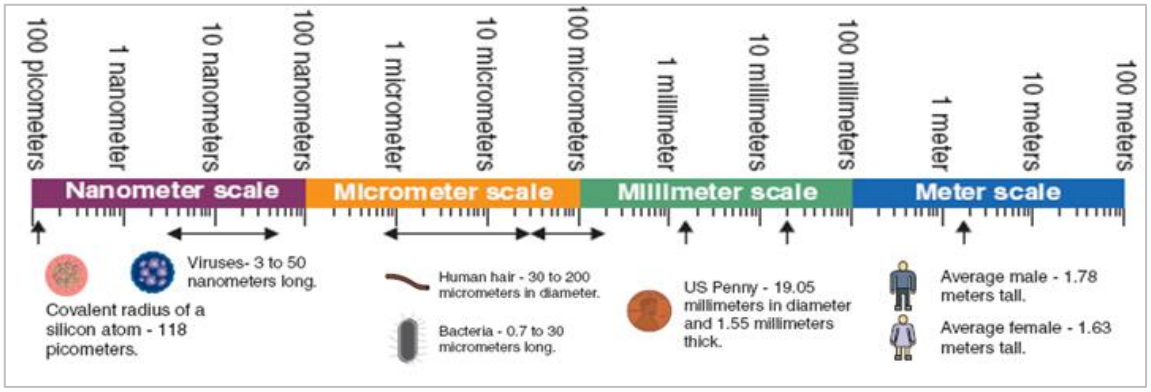


Figure 2.7 – Technology size scale, how small is small? [Jones, 2012]

2.1.5 Semiconductor Business Model and its Evolution

The SI has manufactured its own production equipments and products till 1980, however in 1980 these functions were separated into *(i)* **O**riginal **E**quipment **M**anufacturers (OEM) and *(ii)* **I**ntegrated **D**evice **M**anufacturer (IDM) models to handle the increasing complexities and costs [Wu, 2003 and Hurtarte, 2007]. The DFM approach helped the SI to enhance the yield till 250nm but it has turned into an ineffective R&D activity with high costs. To address these increasing DFM costs and resulting increased technology adoption and alignment lead times, the IDM has transformed into fablite and fabless business models with an objective to integrated DFM within the design flow [Kumar, 2008 and Moore, 1998]. Let us analyze the value chain of the integrated circuits comprising of three major steps *(i)* IC design, *(ii)* manufacturing and *(iii)* assembly/test. The semiconductor business models in this value chain are classified as the *(i)* IDM, *(ii)* fablite, and *(iii)* fabless models (Table 2.2). The transformation of these models has resulted in the emergence of the **E**ngineering **D**ata **A**analysis (EDA)[7] and **I**ntellectual **P**roperty (IP) companies to help automate the design and manufacturing flows and integration of DFM in the design flow. These are the strategic decisions; however the objective is to reduce the increasing new technology lead times and associated costs.

	IC Design	Manufacturing	Assembly / Test
IDM	Yes	Yes	Yes
Fablite	Yes	Yes	
Fabless	Yes		
Foundry		Yes	
EDA	Yes	Yes	
IP	Yes		

Table 2.2 – Semiconductor business models

- **Integrated Device Manufacturer (IDM) Model:** The IDM's are companies that operate along the entire value chain in the semiconductor manufacturing. An exponential increase in the technology R&D costs and production equipments has resulted in the transformation towards fablite business model where new technology is developed in a technology alliance to share high R&D costs and reduce the new technology lead times e.g. Intel etc.
- **Fablite Model:** These are exactly similar in their operations with the traditional IDM models, however the new technology is developed in an alliance to share the high R&D costs and reduce the technology lead times.

[7] The acronym EDA stands for engineering data analysis whereas electronic design automation is always written in full.

The production capacities are also outsourced to the technology alliance if required. It is a modern IDM model which is still holding the maximum market share e.g. Samsung, Toshiba, STMicroelectronics.

- **Fabless and Foundry Model:** This model is being promoted by ITRS based on the fact that the new emerging technology challenges require strong collaboration in this competitive environment to reduce the lead times and costs. The fabless companies do not carry out the product development; they simply design the products based on the technology which is developed in an alliance. The designs produced are equally manufacturable in the technology alliance partners manufacturing facilities. These companies have been successful as there is no capital investment required but the success of these companies relies on IP circuits. In comparison to these fabless, the foundries mostly operate large and modern production facilities with high levels of capacity utilization e.g. TSMC, UMC etc.
- **EDA and IP Companies:** The EDA companies are involved in the design phase of the value chain and they operate in strong collaboration with IDM, fabless and foundries. These companies devote themselves exclusively to licensing (intellectual property or IP companies) and they specialize in the design of certain modules and license the resulting intellectual property to their customers. Unlike fabless companies, IP companies do not have sales operations and license their design and development services exclusively to third parties. There are also companies that focus on electronic design automation. Compared with the other business models, the volume of sales generated by IP and EDA companies is a small part of the overall market but without any initial investment.

The reasons behind structural transformations have been reported as potentially increasing technical and business challenges that have resulted in increasing technology lead times and costs. The objective is to capture the maximum market share; hence the shift in business objective from T2M and T2V towards ramp-up rate a.k.a. T2Q is attributed to be the key driver in this transformation. The fabless and foundry business models are being promoted since 1999 [Kumar, 2008] where design, manufacturing and electronic design automation are separated as the core business functions and emphasis is put on the collaboration and alliances to address the exponentially increasing costs. IDM business model has been demonstrated as the best business model in this competitive environment because it is coherent to reduce the technology alignment and adoption lead times [Shahzad et al., 2011a] and benefit from the cyclic demand patterns to capture maximum market share. We believe that success lies in our ability to quickly ramp-up the production that requires a mature technology and continuous improvements. The IDM platform provides R&D engineers with all the data to be used for the multi-source analysis. The problem today is that engineers - even with the availability of large data sets - are not able to exploit this data, resulting in ineffective DFM methods. Removing these challenges to utilize the data while performing the technology alignment and adoption activities will support the lead times and cost reduction efforts.

2.1.6 Trends in Semiconductor Industry

It is very important to analyze the important trends followed by the semiconductor industry in compliance with the Moore's laws before moving towards the challenges faced by the industry (Figure 2.8). We have seen the transistor count doubling every 18 months in compliance with Moore's law (Figure 2.8a). The first 1 billion transistor processor was developed by many manufacturers in early 2010, but Intel holds the distinction of manufacturing it with the latest technology and on the smallest area (Core i7 [6 core] Gulftown with 32nm and 240mm^2). Intel holds the credit for 2 billion transistor processor in 2011 (Core i7 [6 core] Sandy Bridge-E) with 32nm and 434mm^2. The SI seems to lag behind from projected technology scaling and till today (2012) 32nm is not yet a mature technology for its compliance with quick ramp-up-rates (Figure 2.8b).

To benefit from the technology scaling industry has shifted across different wafer sizes for the economies of volume production (Figure 19-c). The increase in wafer size can be seen almost every 10 to 12 years that has played a significant role in the reduction of transistor costs (Figure 19-d). It gave rise to the notion of volume production to compensate the technology and higher entry costs of the chips. The curves for 200mm and 300 mm wafers present the clear gain in costs by just extended the wafer size. The production equipment costs (Figure 19-e) are continuously increasing because they are required to support

the technology scaling and the increasing wafer sizes. It is usually carried out in the equipment alliances to share R&D costs and provide the partners an early penetration in the market with new technologies and wafer sizes. It is evident that such high R&D costs even if shared, result in higher entry costs (Figure 19-f) for smaller chips but it is compensated with the volume production resulting from the increasing wafer sizes [Kumar, 2008 and Quirk and Serda, 2000].

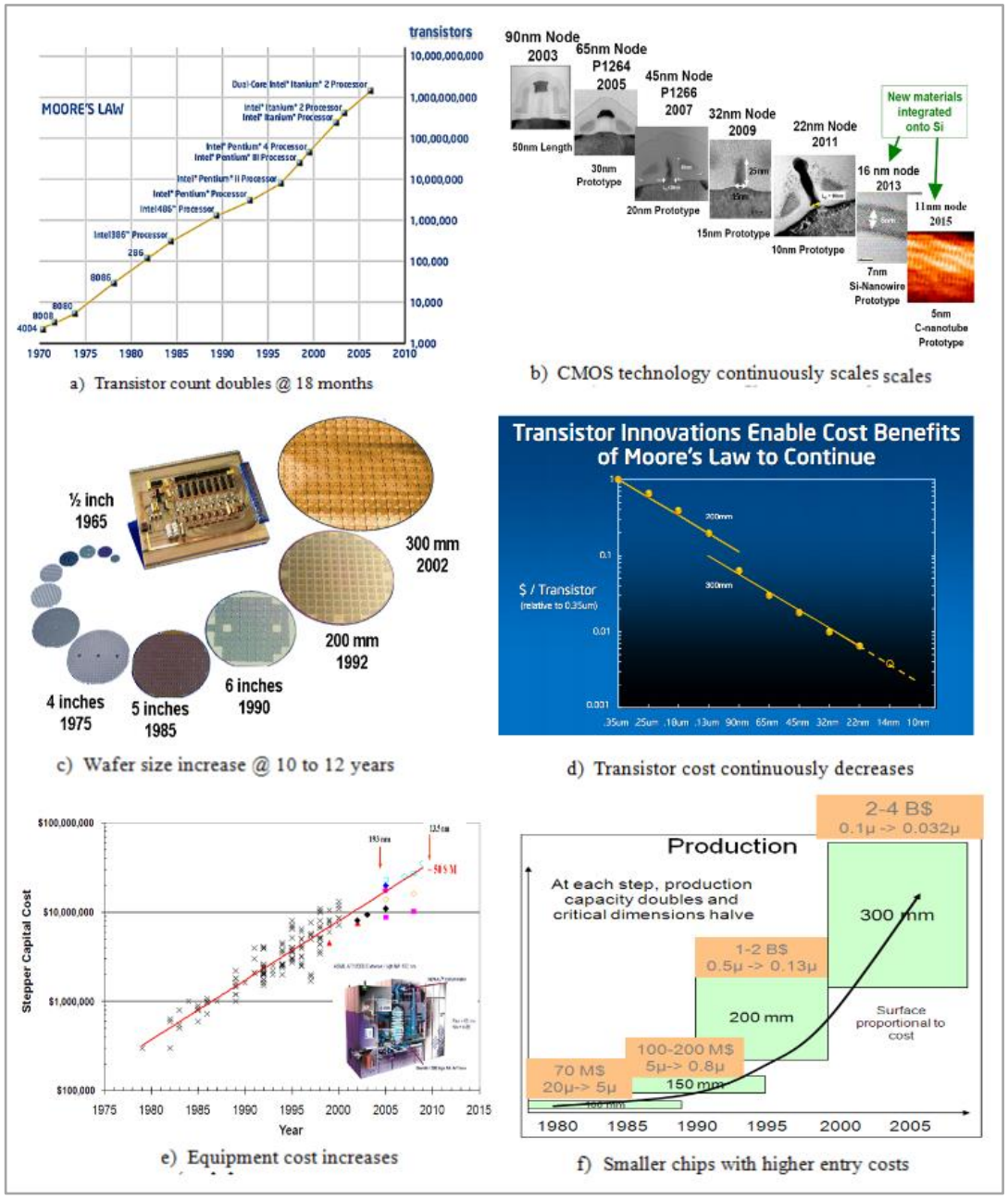
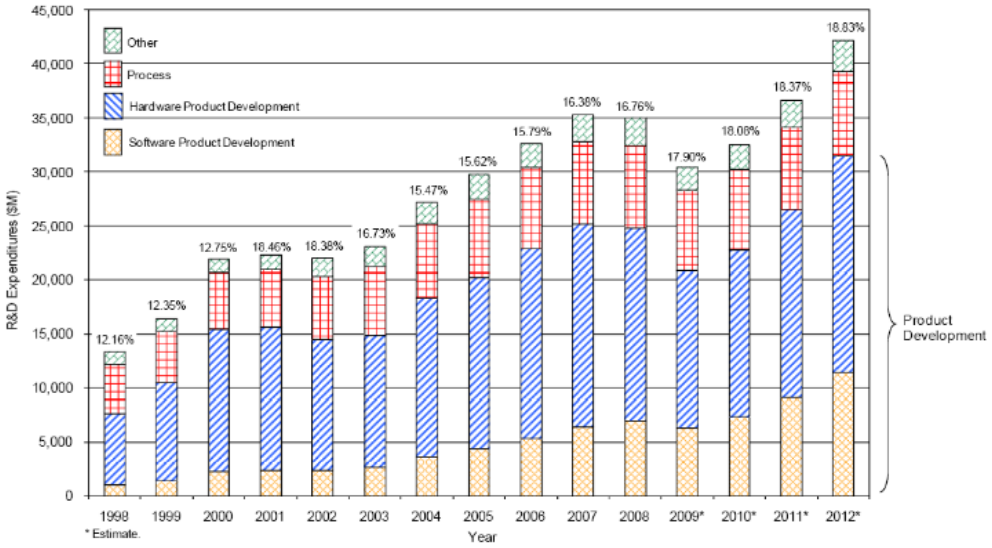


Figure 2.8 – Trends in semiconductor industry within last 50 years

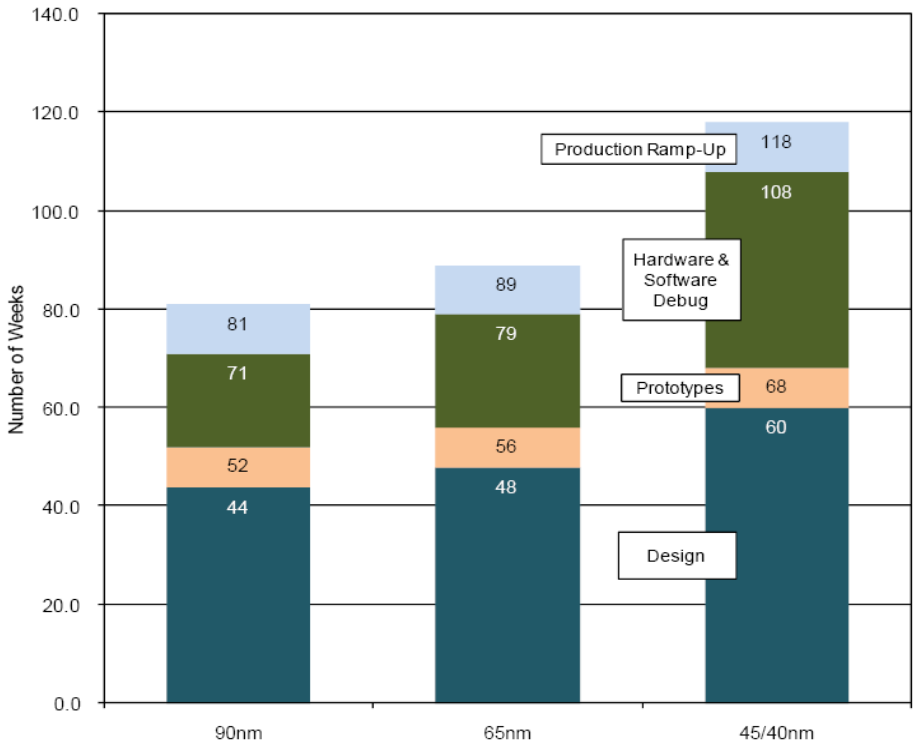
2.1.7 Challenges faced by Semiconductor Industry

The SI is characterized by the cyclic demand patterns (Figures 1.1, 2.2, 2.3, and 2.4) and there have been multiple downward dips, but it always came up with cumulative demand growth which is guaranteed by a positive CAGR rate. It motivates the SI for R&D investments to comply with Moore's law where new technologies capture maximum market share and provide an equal opportunity for all stake holders. What now follows is a summary of the key challenges faced by the industry in an effort to its compliance with the Moore's law.

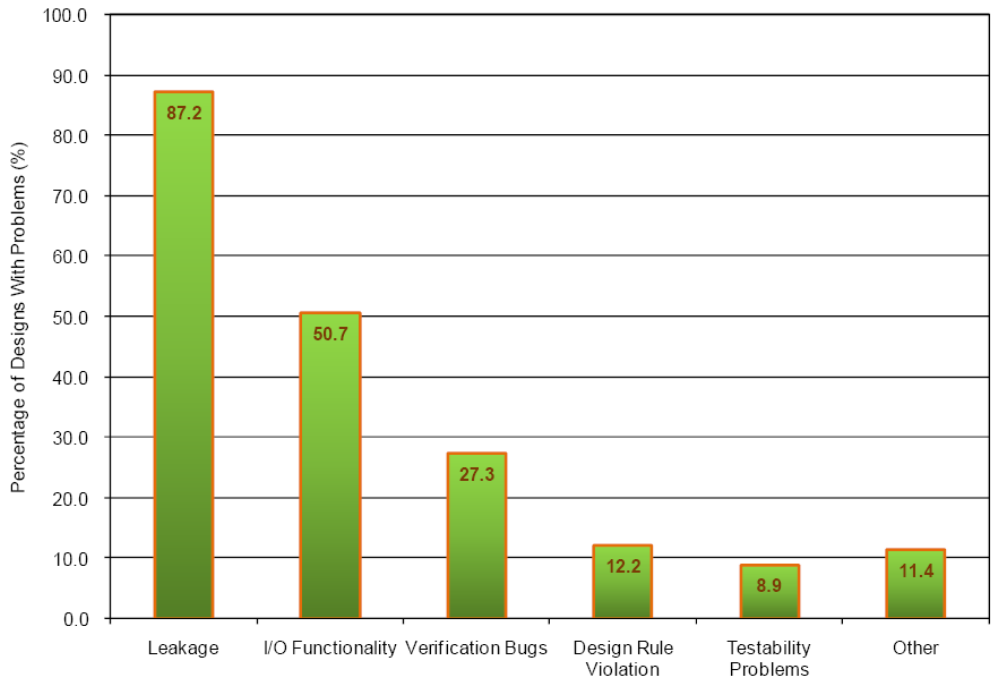
Besides the cyclic downfalls, [International Business Strategies, 2009] investment in R&D within the SI has continuously increased and is expected to reach 18% of total revenues by the year 2012 (Figure 2.9a). Major portion of these R&D costs is focused on the development of new technology platforms (2-7 billion USD from 65nm to 22nm) and IP/library qualifications for predictable design/manufacturing interfaces. Major designs respin reasons (Figure 2.9b) faced by the SI are leakage (87%), I/O functionality (50%), verification bugs (27%), design rule violations (12%) and testability issues (8%). One of the important challenges faced by SI is exponentially increasing costs during technology derivative improvements (alignment) and adoption (product design/development) efforts (Figure 1.2). The design/debug periods for the 90nm, 65nm and 45/40nm are 44/9, 48/23 and 60/40 weeks respectively and prototype/volume production takes 8/10 weeks respectively for all nodes (Figure 2.9c).



a) Continuously increasing R&D investments



b) Design respin costs



c) Design, debug and prototype periods

Figure 2.9 – Major design challenges faced by semiconductor industry [International Business Strategies, 2009]

The IC industry is in a highly challenging phase and IC business continues to have many opportunities with IP (intellectual proprietary) and cell libraries (reusable circuit components) qualifications. These IPs and reusable components reduces the product design and development time subject to the ability to quickly adopt the given technology against new products. The key for success lies in the ability to quickly design, develop and ramp-up the production which at present is challenged by the manufacturing variations. The DFM (design for manufacturing) methods are the best tools to quickly model these manufacturing variations to reduce the technology alignment and adoption lead times for economic benefits (Figure 1.4).

Let us conclude this section and formally propose the research question-1 (chapter-1, section 1.4). The market is characterized by the demand that depends on the population and economy. The world population and demands are continuously increasing whereas the demand patterns depend on economic cycles. The CAGR (+8.72%) guarantees a cumulative demand; hence the SI is obliged to respond to the growing market with new, faster and high value but low cost products. It has led a shift in business objective from T2M and T2V towards T2Q and the emergence of design and manufacturing interface complexities. The SI can only respond the market growth by introducing new technologies every 2 to 3 years, but the R&D cost for the development of new technologies has exponentially increased. In order to address the increased technology lead times and R&D costs, SI model (IDM) has transformed into a fablite and fabless business models. But still, IDM fablite model is reported to be the best models in terms of revenue generation. The IDM fablite model provides a coherent platform for the knowledge capitalization from production line and our ability to improve the challenges faced by the R&D engineers shall result in the continuous improvement efforts for the technology alignment or adoption within the business model. It shall reduce the lead times, associated costs along with an opportunity for early penetration in the market with higher profits.

The SI is a 300+ B$ industry with an equal opportunity to capture maximum market for its stake holders. In order to maximize maximum share, the SI objectives need to be revisited and ranked to formulate a strategy that is truly in line with the top ranked objective. This can be accomplished by further analyzing the existing business models (IDM, Fablite and Fabless) to either find or propose an extension in the existing

models to get maximum market share in line with the top ranked business objectives. These arguments formulate the first research question, which is addressed in chapter-3.

Readers who are not familiar with the SI design and manufacturing flows should read the appendices B and C before moving to section 2.2. These appendices provide a brief review on the design, mask preparation and manufacturing flows in the SI along with a CMOS inverter example that simulates these steps. The objective is to better understand manufacturability and yield limitations prior to investigate the DFM ineffectiveness beyond the 250nm node. It is to remember that the DFM ineffectiveness has resulted in an exponential increase in the technology development costs followed by their extended alignment and adoption lead times. It is a challenge towards accomplishing quick ramp-up rates with the existing SI business models. The next section reviews literature about DFM and its evolution in the SI to find reasons for the DFM ineffectiveness.

2.2 ROLE OF DFM METHODS IN SEMICONDUCTOR INDUSTRY AND EVOLUTION

The SI is a most complex, competitive and technologically fastest growing manufacturing domain. The success of the SI is attributed to the DFM methods that helped us to continuously move towards miniaturization at the reduced cost but beyond 250nm technology these methods have turned into a high cost ineffective R&D activity. It has resulted in the transformation of the business models and today a new technology is developed in an alliance to share the costs and reduce technology lead times. This strategy ensures timely introduction of the new technologies to cope up with the pace of Moore's law in compliance with the industrial slogan smaller, faster and cheaper. In this section we shall analyze the scope and evolution of the DFM methods and shall try to establish key challenges and reasons that led this ineffectiveness.

First, we provide a general understanding about few key words: design, manufacturing, DFM and MFD. A design comprises of conceptual/behavioural, embodiment/structural and detailed/layout design flow stages whereas manufacturing represent a manufacturing system [McGregor, 2007] as *(i)* **D**edicated **M**anufacturing **L**ines (DML), *(ii)* **F**lexible **M**anufacturing **S**ystems (FMS) and *(iii)* **R**econfigurable **M**anufacturing **S**ystem (RMS). The DFM refers to the intermittent operations focused on concurrent design/process selection and prototype development for the technical/economical design evaluation [Herrmann et al., 2004]; whereas MFD is an effort focused on controlling the repetitive operations dedicated to normal production using **A**dvanced **P**rocess (APC) and **E**quipment (AEC) **C**ontrol operations. The DFM objectively follows the design to manufacturing information flow but MFD focuses on the manufacturing to design information flow (process control).

The DFM is defined as the ability to reliably predict downstream life cycle needs and issues during early phases of design (Committee on New Directions in Manufacturing 2004). It is focused on economic benefits from the volume production by trading off cost-quality-time triangle [Raina, 2006] and is classified as [Mehrabi et al., 2002] product DFM (producing manufacturable designs within defined processes) and process DFM (developing processes with less rework and high manufacturability). It is further classified as soft and hard DFM [Liebmann, 2008], where soft DFM is related to the existing technologies focused on design rule clean layouts and hard DFM focused on the design-technology co-optimization, ultra-regular layouts, extreme resolution enhancement techniques (RET) and process/device modeling. DFM is further classified as physical and electrical DFM [Appello et al. 2004] based on rules and models respectively. DFM in the manufacturing industry is also categorized on its intended purpose [Honma et al., 2009] like DFP (performance), DFD (design), DFF (fab), DFY (yield), DFR (reliability) and DFT (testability). One frequent DFM alternative (DFY) is focused on via/wire optimization, metal density uniformity and cell swapping in SI [Raina, 2006]. Manufacturing issues that cannot be predicted accurately without the product geometry and specifications gets more complicated when 60-80% components are outsourced [Rezayat, 2000], leaving designers and R&D engineers with limited data.

2.2.1 SI Challenges and Rise of Interest in DFM

Initially DFM efforts were based on rough estimates of the downstream effects and rest was expected to be controlled through advanced process control (APC) and advanced equipment control (AEC). This worked well until 250nm but after that increasing complexity of circuit layout and shrinking sub wavelength lithography eventually resulted multiple respins and yield losses. The 130nm node is considered the cut-off point where the need for DFM was felt to tackle increasing feature and design limited yield losses [Cliff, 2003]. From the designer's perspective, things are getting more difficult because the process windows they are getting back from manufacturing are so tight that they are having a hard time getting the design methodologies to work [Peters 2005 and Dauzere-Peres et al., 2011].

The concept of DFM has emerged in a diversification of terms like DFY, DFV and DFT etc. but all terms come under the umbrella of DFM along the product life cycle (PLC) as DFX having similar objectives of cost, quality, yield and time-to-market [Anderson, 2006] where X refers to the various stages in **P**roduct **L**ife **C**ycle (PLC). As the new **D**esign **F**or **A**ll (DFX) methods are being explored, the definition of DFM has become synonymous with DFX and [Herrmann et al., 2004] concurrent engineering (simultaneous development of a design and the supporting life cycle processes). These methods could be written rules or simulation tools for cost/performance estimation. The DFX tools are focused to provide designer, predictability information on multiple issues within down side of the product life cycle. These tools are applied directly on the CAD designs and the provide advice on the product performance (qualitative, quantitative and binary).

The design and manufacturing interface modeling has been straightforward till 250nm node [Cliff, 2003 and Radojcic et al., 2009]; however hardware to model gaps (variability) started increasing with the industrial shift towards 193nm stepper to manufacture features less than the wavelength of light source. The compensation techniques like OPC and RET emerged as an extended flow (DFM flow) to mitigate these manufacturability and yield loss mechanisms. The efforts to move towards 157nm wavelength light source have been abandoned; hence we have to live with 193nm stepper for the smaller process nodes e.g. 45nm, 32nm, 22nm etc. It is one of the major causes for manufacturability and yield related issues and we have seen the rise of interest in DFM methods followed by aggressive use of compensation techniques like OPC, RET, **P**hase **S**hift **M**asking (PSM) and immersion lithography [Venkataraman, 2007] during design flow.

From the above facts, it is evident that stretching CMOS technology for the smaller process nodes without extending the traditional DFM methods is not possible. It is important that we put DFM back on track because of the investments made in the equipment, material, process and design (IP/libraries) for economical benefits. The DFM methods in last two decades has evolved from design rules to DFM rules like (layout/routing) and DFM models like CAA, CMP, Shape, Yield, Leakage and SSTA to mitigate yield losses [McGregor, 2007]. The new business objective of the SI “ramp-up rate” links our success with the first time correct design or a design that could be ramped-up without yield loss. Increasing complexities and newly emerging spatial variations along with the associated costs has been the core reason for the structural transformation of the SI business models. We need to put DFM back on track for the future smaller nodes, otherwise the Moore’s 2nd which was predicted to end in 2005 [Ross, 2003] shall become true now (see section 2.1.3).

2.2.2 A Comparison of DFM Efforts in SI and Manufacturing Industries

Let us start by comparing DFM efforts between manufacturing industries and the SI (Figure 2.10). The design stages in manufacturing industries (conceptual, embodiment and detail design) can be mapped partially to the extended design flow in SI (behavioral, structural/schematic and layout design) respectively. Extended design flow in addition to the design stages includes DRC/LVS checks based on timing/delay data from intended technology. Prototype and pre-production stages are present in both industries as a design verification stage prior to the normal production. Design flows are directly impacted by the CAD tools used to simulate design against cost, cycle time and manufacturability. Prototype and pre-production being

sensitive are evaluated against yield loss limiters and are strongly impacted by the **C**omputer **A**ided / **I**ntegrated **M**anufacturing (CAM/CIM) tools and the manufacturing system (DML, FMS, RMS).

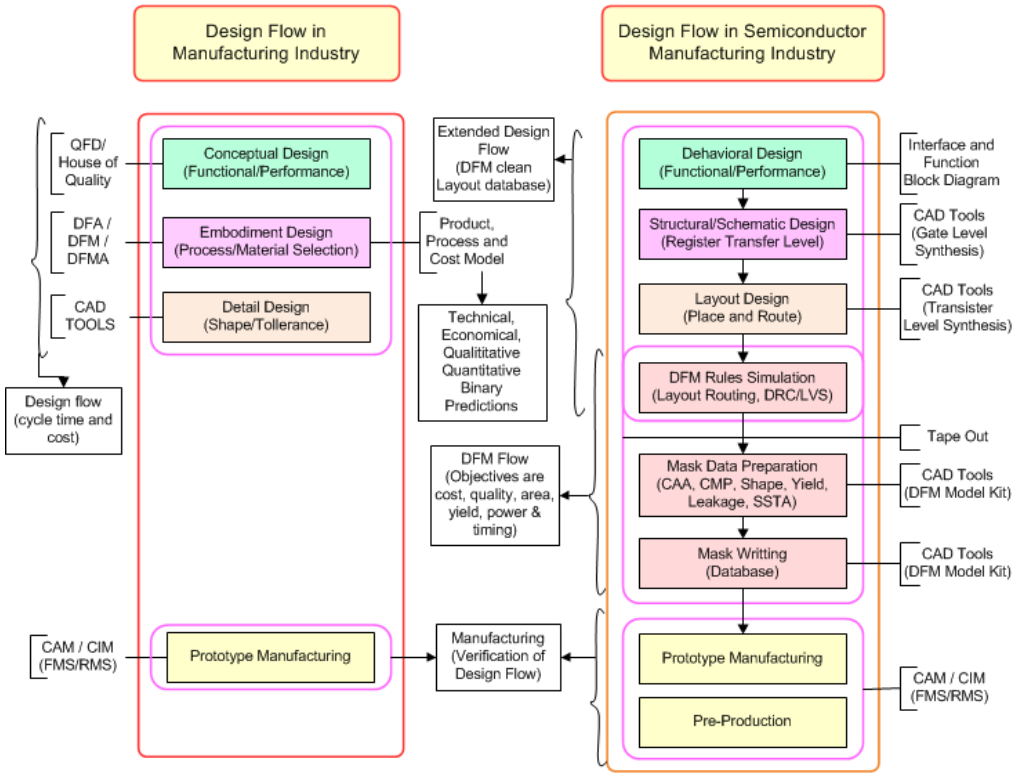


Figure 2.10 - Comparison of the design flows in manufacturing industries and SI

Within SI we have shifted from rule based to model based DFM (Figure 1.7, section 1.3) flow where DFM models are applied to perform hotspot analysis, CMP simulations and geometrical shape predictions. We can observe that both DFM and MFD efforts are present in the design and development stages. The DFM knowledge includes DRC/LVS to estimate the device behavior and manufacturability and the DFM flow (compensation techniques like RET, OPC, MEEF) to accurately predict and correct potential defect driven yield limiters. It is capitalized as predictive models to simulate designs whereas partial DFM efforts are applied to adapt the selected process for each product that might help in improving existing DFM knowledge. The design driven defects are spotted and corrected through design respins which is less costly, however additional steps in the DFM flow result gigabit designs which are computationally expensive in simulation time. The prototype and pre-production are the phases where partial DFM efforts are applied by the process engineering and integration teams to adapt the base process from the technology platform. Later stages in the PLC (growth, maturity and decline) include only MFD efforts (APC/AEC) to ensure the process target. During design and development phases technology platform (TP) serves as a base technology and its customization is ensured through process adoption at prototype and pre-production phases. Identification of the design/defect driven yield loss mechanisms provide an opportunity to improve the technology platform, shrink to new process node or to move towards a new technology platform and we argue that every abnormal drift or variation within manufacturing system is an opportunity to learn variability and to improve Design rules/DFM models.

The DFM methodology seems quite different in the SI and manufacturing industries; let us make a brief comparison to understand the high level similarities and low level differentiations about DFM in these two domains. Both industries are subjectively focused on the integration of DFM early in the design flow [Vliet and Luttervelt, 2004] so that manufacturable products move to the production line. The correct geometries are one of the significant factors for the success of DFM methodology in both industries. Incorrect geometries and systematic physical defects in the ICs often result in the timing, signal and delay, leading to a functional failure. The SI manufacturing process requires 1100+ steps and 8 weeks of

processing; hence it is critical to find out the potential issues as early as possible in design flow to save the production and inspection capacities. The process based DFM initiatives in SI can be referred as a technology; however its maturity has a strong impact on the potential yield and manufacturability. A common approach in DFM philosophy as stated by [Das and Kanchanapiboon, 2009] is the use of parametric models and knowledgebase library which is similar to device and interconnects **S**imulation **P**rogram with **I**ntegrated **C**ircuit **E**mphasis (SPICE) models and design rules.

In manufacturing industry we are focused on integrating DFM efforts within the design flow to improve effectiveness of DFM philosophy [Vliet and Luttervelt, 2004]. It is exactly the same objective that we are following in SI to put DFM back on track against its emerging ineffectiveness, hence manufacturability and yield related issues can be attributed to incapable or ineffective DFM methods. The SI is focused on building design libraries (pre-simulated circuit reusable components) that are used by the designers during the design process. It is assumed that all the manufacturing data is available [Myint and Tabucanon, 1998] for its use in the automated DFM evaluations but at present besides the huge data volume our engineers are not able to exploit it for the root cause analysis against newly emerging drifts or variation phenomenon. The biggest differences we can find between the semiconductor and manufacturing industries in the evolution of DFM methods are *(i)* the change of pace for the new technologies and in SI it is characterized by the fastest change in the smallest period of time and *(ii)* the inability to exploit the huge data volume and dimensions in SI that result in increased technology lead times.

We can see that objectively both industries use the DFM philosophy to find manufacturing and yield related issues as early as possible. The SI is different than other manufacturing industries in the sense that it is characterized by the fastest change in the smallest period of time, hence our success lies in our ability to quickly analyze the emerging spatial variations and transform them into rules and/or models. It can be seen that DFM evaluations in both industries are supported with the automated CAD tools but in the SI, R&D engineers are not able to exploit huge data volume and dimensions to find root causes against new variations and drifts. This ineffectiveness in the DFM methods has resulted from our inability to exploit these sources and is resulting in high costs and extended technology lead times.

2.2.3 DFM Techniques (pre-1980 era)

The manufacturing industry has followed the motto "I designed it; you build it!" [Anderson, 2006] for a long time as a sequential design flow, before they shifted to the DFM/DFA approaches for the economic benefits. The DFM analyzes a design for potential cost and manufacturability prediction and evaluate its alternatives on yield, quality and performance; whereas **D**esign **F**or **A**ssembly (DFA) is focused on reducing number of parts in an assembly to ultimately reduce cycle time. The DFM, being a bit older concept than DFA was focused at the component level by collecting material information and available processes with estimated costs but proved unfruitful in terms of assembly cycle time [Bolz, 1958]; hence DFA was established as a technique to simplify the product structure [Boothroyd and Dewhurst, 1990] before DFM can be applied for the cost and cycle time reductions. The manufacturing decision within DFM/DFA resides on production/assembly operations and factors that contribute to it are material, geometry and tolerance [Boothroyd, 1994 and Boothroyd and Dewhurst 1990]. The DFA provides an opportunity to trade-off potential improvements based on the type of assembly operation (manual, semi-automated and fully automated), evaluated on merging or separating the components or assembly. During the decade (1989-1999), notion of DFM existed as a technology, a methodology and a philosophy. The DFM is applied across the design and development phases in order to achieve economical benefits, if it is applied accurately, it can directly reduce the design respin, time to market, development costs. Traditionally, DFA methods evaluate the ease of assembly and DFM methods evaluate the feasibility and cost of manufacturing the product [Anderson, 2006 and Bralla, 1998 and Corbett et al., 1991].

The design costs consume about 10% of the total product design and development budget, whereas 70-80% of manufacturing costs are determined at the time of design [Boothroyd and Dewhurst 1990],

[Anderson, 2006]. The Sequential design flows are not appropriate for quality improvements and cost reduction, **D**esign **F**or **M**anufacturability and **A**ssembly (DFMA) methodology with loopback option is a solution guaranteed for the reduction in cost, assembly time, parts in assembly and time to market**.** The term DFMA was coined by Boothroyd as a methodology [Boothroyd, 1994] and implemented as a simulation tool that proved beneficial for manufacturing industries in terms of cost, design spin and cycle time reduction. The generic DFMA flow (Figure 2.11) defines the scope of initial DFM efforts limited to simplify the product structure followed by material and process selection from the existing knowledge. It is important to note that accurate technical and economical prediction is not possible until the geometry and specification of the products are complete [Herrmann et al., 2004]; however during DFMA cycle, prototype validate such predictions and ensures smooth transition toward volume production.

These facts are equally applicable to the SI as well, for example if we look at the IC design the CAA, Hotspot and SSTA analyses are applied only after the transistor level synthesis where the geometric shapes of the devices and interconnects are well defined. The cycle time reduction by simplifying the product structure is less applicable in SI, however layout optimization is performed in order to avoid manufacturbility and yield issues. We have seen that technology plays a critical role in lead times and costs and success depends on its matruity which can be attributed to DFM effectiveness.

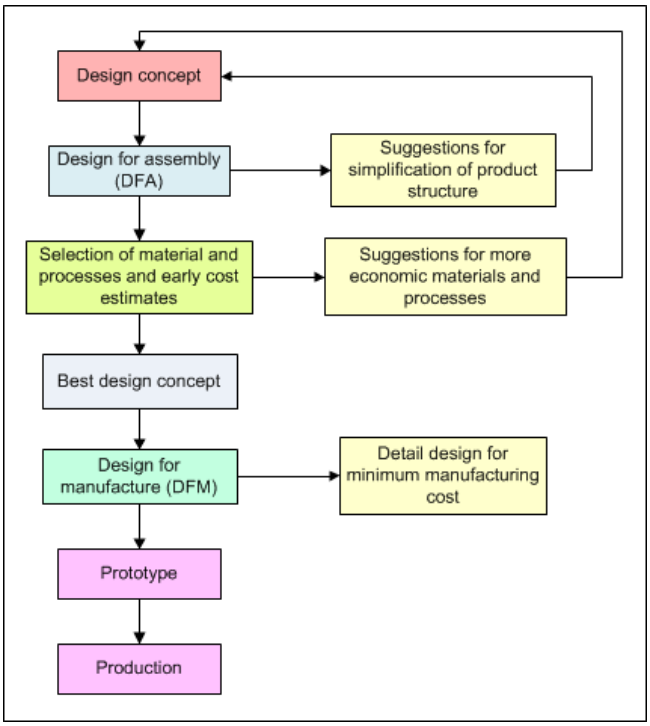


Figure 2.11 - DFMA (design for manufacturability and assembly) [Herrmann et al., 2004]

The product, process and cost models are proposed for a detailed manufacturability analysis [Ramana and Rao, 2002]**.** The product and process models describe product as design parameters (geometric and non geometric information) and process as the capabilities associated with geometric and non geometric parameters respectively. The cost models are classified as the scaling, activity-based and statistical models which provide precise cost estimation based on the detailed design. The measures for the manufacturability are classified as binary (0/1: if one or more DFM guidelines or rules are violated [Zuozhi and Jami, 2005], design is not manufacturable), qualitative (difficult, easy or moderate with an ordinal ranking among candidate designs) and quantitative (cycle time and cost against allocated budget and time). The manufacturability assessment is performed at two stages, high level (parametric matching) and detailed level (interpolation/extrapolation). The parametric matching system uses a large database of processes characterized by capabilities whereas matching is done while designers provide limits to the design attributes for appropriate candidate processes.

The manufacturing system palys the most important role in achieveing DFM objectives [Koren et al., 1999 and Mehrabi et al., 2002]. In production, the DML is a fixed automated line used to produce the core technology parts in volume with low cost per unit and is successful until demand exceeds supply. The FMS is focused on product variety with changeable volume/mix whereas RMS suport rapid change in system structure adjust for the production capacity in response to product mix. The DML and FMS are systems whereas RMS is a dynamic system that combines DML's high throughput and FMS's flexibility. In SI, all the equipments used fall in RMS category; however they are limited by the size of the wafer being used for production e.g. production lines for 200mm wafer and 300 mm wafers. In these production lines we can use multiple technologies by process and equipment adjustments. The production database are the key source for first hand knowledge to provide an accurate feeback to design, process and product engineers. It provides means to find the root causes against abnormal behaviors. It can be concluded from the above discussion that DFM effectiveness depends on the ability to access and exploit production databases. Integrated DFM tools have also been proposed to benefit from the available manufacturing data because the relevant information extraction and availability to apply DFM is a challenging part **[**Eskelinen, 2001**].** A universal data model with an open access technology is needed as design flows have turned up data-centric with the inclusion of extreeme RET [Cottrell and Grebinski, 2003] techniques.

The above proposed solutions are industrialized and open access technologies are developed for the extraction of relevant information. These systems worked well and have led the emergence of EDA (engineering data analysis) companies to develop these technologies. The problem is that these tools become obselete quickly and the customers are required to purchase the new upgraded versions. New equipments and metrology techniques provide additional information on production processes, but propriatry nature of production data sources, extraction and analysis tools restricts our ability to include additional information in existing data analysis framework which results in ineffective DFM methods.

2.2.4 Adaption and Diversification of DFM to SI (post-1980 era)

The DFM concept existed in its mature form around 1980's when SI started realizing associated benefits. In the early days of DFM adoption in SI, design flows included loop back option (Figure 2.12) for optimization focused on feeding back the manufacturing data as soon as it becomes available for new designs [Reid and Sanders, 2005].

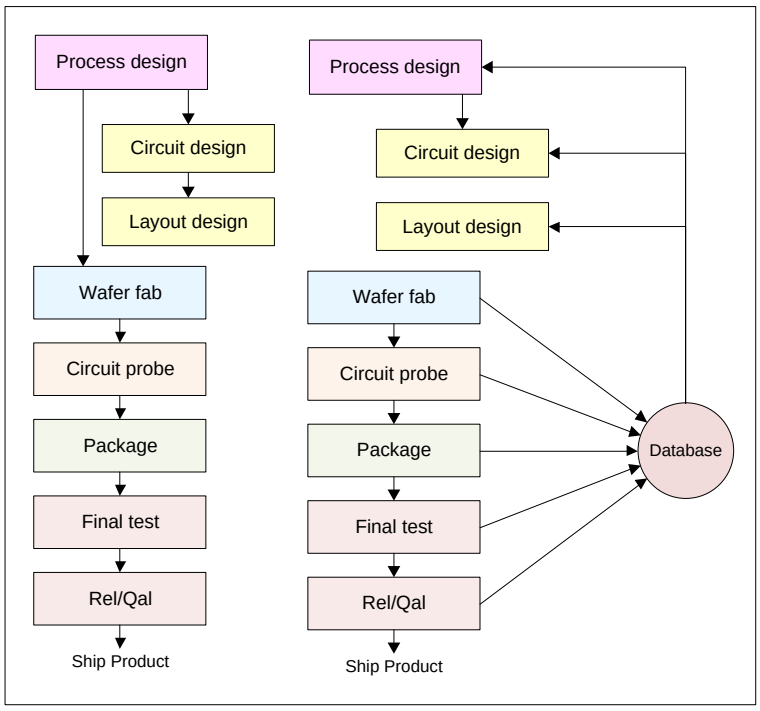


Figure 2.12 - Early semiconductor design flows with loop back [Reid, and Sanders, 2005]

Let us discuss the general IC design [Reddy, 2000] flow (Figure 2.13). It starts by transforming customers' basic requirements into design specifications followed by behavioral description (interface and functional block diagrams) which is simulated to ensure target functionality. The most commonly used languages at this stage are **H**ardware **D**escription **L**anguage (HDL) and Verilog that offers the advantage of concurrency over other high level programming languages. The design is further detailed out at functional and schematic levels by gate level synthesis using **R**egister **T**ransfer **L**evel (RTL) language at structural/schematic levels. The netlist is generated in step-5 using automated EDA tools. It follows the timing and delay simulations on netlist where process, timing and delay models are taken from intended processes. The objective of the timing and delay simulations is to assess functionalities as well as manufacturability with **L**ayout **V**ersus **S**chematic (LVS) checks. The **D**esign **R**ule **C**heck (DRC) is performed before LVS to ensure the designs compliance with **D**esign **R**ules (DR) for a given technology. It verifies spacing, shapes etc. of devices and interconnections to ensure manufacturability of a design during production. These steps are repeated until all DR violations are removed and the parametric performance is achieved. The final design is extracted in **G**raphic **D**atabase **S**ystem (GDS) format, also known as tape-out. It is similar to netlist with additional information and it is generic enough to be used within any manufacturing facility for prototyping or production.

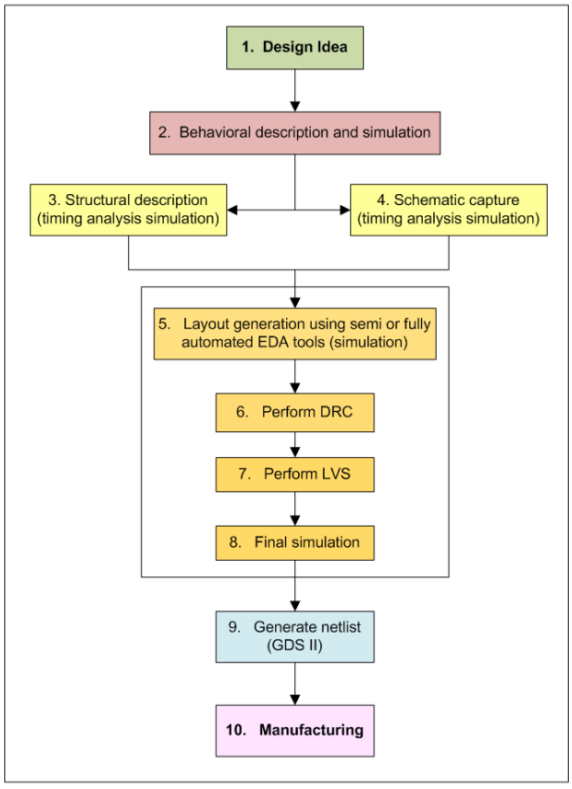


Figure 2.13 - Typical design flow within SI [Reddy, 2000]

The SI follows an extended design flow (traditional design flow and DFM flow) supported by design enablement (design rules), DFM rule and DFM models. Design rules ensure the potential manufacturability whereas DFM rules are an extension of design rules, classified as recommended or robust rules, to ensure yield. The DFM model kit refers to the process variation linked with the performance or capabilities of production facilities. The sub-wavelength lithography has resulted in increasing gap in model-to-hardware matching (gap between device layout as viewed by designer and final shape in silicon) due to the influence of the neighboring shapes and the precision with which geometries can be manufactured. 130nm node is the cut-off point where need for DFM was felt to tackle potentially increasing feature limited and design limited yield loss. OPC aware and OPC friendly design turned to be a specific focus, however **O**ptical **R**ule **C**heck (ORC) is the key for chip level analysis to ensure reliability and yield [Cliff, 2003 and Turakhia et al., 2009]. The knowledge based DFM systems (Figure 2.14), with product information base, manufacturing knowledge

base, expert system and concurrent system design elements; are developed with an objective to trade-off cost, quality and yield for the integrated product design and process selection [Yuyin et al., 1996].

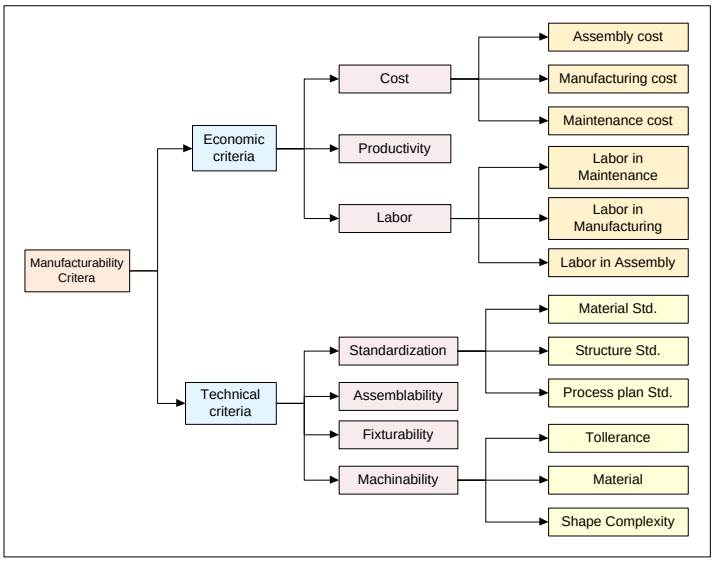


Figure 2.14 - Manufacturability criteria for IC designs [Yuyin et al., 1996]

The DFM is a key business strategy under timely driven market with cost reduction and continuous introduction of the new products [White et al., 1995 and Kasem, 1997 and Corbett et al., 1991). The Idea of **M**anufacturing **D**riven **D**esign (MDD) as an information integration model is proposed to share design and manufacturing information across all design engineers [Cho and Hsu, 1997]. It is focused on capturing design and resources knowledge at all stages of the product design, modeling the relationship and putting them in a Meta database which provides information to all design engineers and is a similar approach to DRC used in SI industry. **I**ntegrated **P**roduct **D**evelopment (IPD) was proposed to capture manufacturing data as the key competitiveness and strategic element (Figure 2.15):

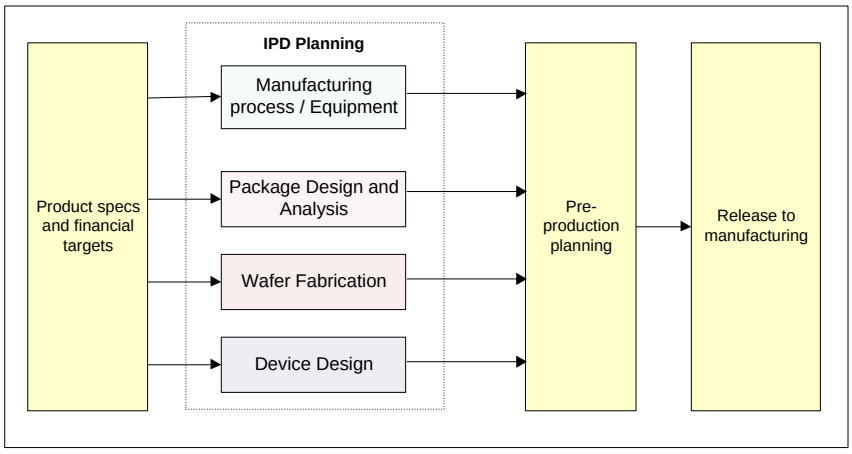


Figure 2.15 - Integrated product development framework [Cho and Hsu, 1997]

Ultimate diversification of the DFM within IC industry has emerged as a yield improvement strategy focused on an ideal objective ‘first time correct design’ achieved by proactively mitigating design driven (sub-micron lithographic printability issues) and defect driven (process or technology issues) yield losses [Redford 2009 and Raina, 2006]. We need to proactively solve systematic defect driven yield related issues using DFM methodology in an efficient and effective manner from cell to chip level. These methodologies are based on data mining and are applied [Ouyang et al., 2000 and Schuermyer et al., 2005 and Appello et al., 2004] to learn yield limiters but the extent to which such methods can be applied on the production lines is still a question; however they highlighted features with high correlation and those which were predicted to have no correlation, turned up with strong correlation requiring correction in the models. The DFY elements

so far taken into consideration are grouped under physical DFM (via and wire optimization, metal density uniformity and cell swapping) whereas areas which require to be integrated are the electrical DFM (model based guidelines) to continue CMOS scaling [Raina, 2006].

Before we proceed further, let us summarize the discussion by defining the scope of DFM rules and or models [DFM Dictionary, 2004] within design and manufacturing flows (Figure 2.16). The presented flow clearly demonstrates that once the design is ready, it moves to the mask data preparation step which is the second line of defense against the manufacturability and yield loss mechanisms. Both scribe and device level Boolean operations are performed along with aggressive RET compensation. The design moves to the production where it is controlled using the advanced process and equipment control methods (MFD). The design rules, DFM rules and model improvements are based on the data collected from the production lines. The analysis against the significant drift updates the rules and models to ensure manufacturability and yield gains. It can be concluded that this analyses is dependent on our ability to access and exploit the production databases which is useless if the data models are not updated continuously with new data dimensions.

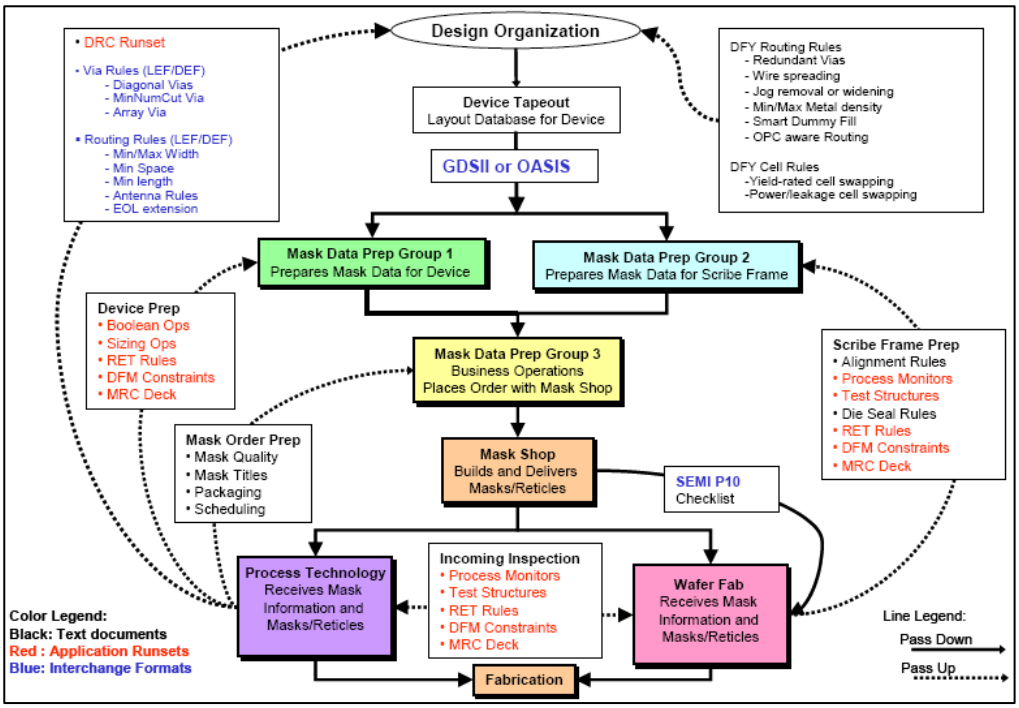


Figure 2.16 Scope of design rules, DFM rules and DFM models [DFM Dictionary, 2004]

2.2.5 DFM Challenges and ECAD/TCAD Tools

The OPC and RET techniques being used extensively at 90nm/65nm nodes with 193nm steppers turn 100MB design data in several GB which is time consuming and complex when simulated [Hurat and Cote, 2005]. The ECAD and TCAD (electrical and technology CAD) tools help designers through extended design flow [Radojcic et al., 2009 and Trybula, 1995], however the objective is to apply DFM methods in an organized way to compensate deficiencies from design layout (GDSII). Authors propose an idea of clear shape predictive model dynamically synchronized with the manufacturing data [Hurat and Cote, 2005]. [Honma et al., 2009] presented a DFM workflow focused on addressing downstream design and manufacturing issues. A similar methodology is proposed by [Hurat et al., 2006] as a Genuine DFM approach named as DFM0 (design rules), DFM1 (design guidelines like redundant vias etc.), DFM2 (new technology platform), DFM3 (RET/OPC) and DFM4 as an effort to solve issues in DFM3. Statistical metrology based on variability dissection in two directions was proposed to learn variability [Boning et al., 2008]: *(i)* downstream across process flow (APC/AEC) with model based feedback and feed forward system to control manufacturing variations and *(ii)* upstream across design flow integrated with ECAD/TCAD tools as design rules, DFM rules or DFM models. Authors have proposed a web based 3-layer interface for

collaborative engineering and data flow between design and manufacturing [Morinaga et al., 2006]. Proposed architecture suggests distributed relational databases where interface layer is focused on providing the alternative naming conventions and data sharing across design and manufacturing flow.

Most semiconductor manufacturers expect 193nm immersion lithography to remain the dominant patterning technology at 32nm and 22nm technology nodes [Venkataraman, 2007]. The primary goal of DFM is to enlarge process yield window [Raina, 2006] while the primary goal of MFD (APC/AEC) is to keep the manufacturing process in that yield window. The RET to overcome diffraction effects including OPC, PSM and immersion lithography are being stretched, resulting in restrictions on layout and an explosive growth in the number of design rules being employed to enable manufacturable designs. The DFM techniques that analyze design content, flag areas of the design that limit yield, and make changes to improve yield are being developed and employed [Venkataraman, 2007]; however it is extremely difficult, if not impossible, to predict all the problems that are likely to occur and what impact they will have on a product.

From the above discussion, we extend our argument that we are objectively focused on fining challenges and limitations that resulted in DFM ineffectiveness followed by generic solutions to resolve these issues to put DFM back on track. In this thesis we are not working on improving the simulation time with extended design and DFM rules; however we are working on the extraction of these rules and improving the DFM models. Please refer to Figure 2.40, the feedback loop presented here is meant for the new technology which is frozen upon its transfer from an alliance and we are not authorized to change it. To be more competitive, we have to extend these feedback loops for the technology derivative or improvement initiatives (alignment) and product design and development (adoption). It requires efficient data access, dynamic exploitation and effective root cause analysis.

2.2.6 Increasing Design Size and DFM Realization Challenges

The Chip design at 32nm is capable of producing a billion-gate design; however most of the CAD tools we use do break at 100 million gates [Lapedus, 2008]. Shifts in various technology platforms led a test of fabless business model vs. IDM in terms of robust control over DFM. For example, **T**aiwan **S**emiconductor **M**anufacturing **C**ompany (TSMC) that has turned itself into a foundry in the Fabless business model, has made a very smart strategy by launching a new DFM methodology called unified DFM where clients are provided with the model based DFM design kit (lithography checker, CMP, CAA, OPC, hot spot …) and they can use the simulation resources of TSMC to quickly come up with the first design. It is a test for traditional IDMs as TSMC with such services to its customers has turned itself in a virtual IDM with no obligations but extended support and collaboration with the customer.

The DFM information has traditionally existed as design rules (robust/recommended/ restrictive) [Mason, 2007]. The **D**esign **F**or **S**emiconductor **M**anufacturing (DFSM) system suggest circuit level simulations [White and Trybula, 1996] where DFM kits are integrated with the CAD simulation tools to simulate device physics, process physics, layout and interconnects. The role of electronic design automation has taken an utmost importance for the effective DFM integration in extended design flow to ensure design enablement [Radojcic et al., 2009] a.k.a. manufacturability and yield. The CAD tools being developed in view of design flow and DFM flow fall under the category ECAD and TCAD [Trybula, 1995]. The ECAD tools are used to design digital circuit and move through different phases in the design flow whereas TCAD tools are used during the DFM flow to verify the circuit against DFM methodologies like parasitic extraction, LVS, DRC etc. to assess its manufacturability and yield. One other classification of DFM tools [Wong et al., 2008] do exists as the design and manufacturing side DFM tools and they can be mapped to the ECAD and TCAD tools respectively. The DFM tools can also be categorized based on the critical layers: cells, IP blocks, I/O, memory (metal1/poly/via/diffusion) and above metal1 taking full chip simulations mitigated by the DFM aware place and route tools. The Objective is to design using DFM compliant cells and IPs to avoid later corrections [Trybula, 1995] and faster first time correct design and this concept is equivalent to reusable components.

Factors that lead the realization of DFM concept are *(i)* the CAD/CAE tools to ensure proper linkage between the design and manufacturing, *(ii)* relational databases and computer networking capabilities for efficient and effective data exploitation, *(iii)* process health feedback (SPC) with possible preventive maintenance predictions and *(iv)* the yield/performance trade-off modeling and simulation tools.

2.2.7 Role of SI Business Models in DFM Evolution and Adaption

Besides above factors SI business model and its evolution has a strong impact on the knowledge capitalization through R&D activities. We have seen in the section 2.1 that the SI started as an IDM that designed, fabricated, marketed and manufactured equipment; however in mid 1970s they separated their business functions as IDM and equipment manufacturers [McGregor, 2007]. The downfall in SI around mid 1980s forced IDMs to outsource low value production operations to the more efficient manufacturers. In late 1980s the idea of fabless organization (design companies) emerged as a less resource and capital intensive business model focused on innovating new designs rather than solving manufacturing issues. It has shifted industry's supply chain structure from centrally coordinated to vertically disintegrated configuration. Asian innovative semiconductor manufacturing companies turned into foundry business model (1980) in which foundry could possibly manufacture its own design or external design and share its manufacturing capabilities with innovative smaller design companies referred as fabless.

To remain competitive, [Morinaga et al., 2006] SI need to send products quickly to the market with highest production yield whereas product cycles are getting shorter and shorter with each new process node and rapidly declining selling prices. The SI evolution is dependent on the competitiveness supported by technology platforms, equipment and materials. The DFM plays a significant role in technology development and continuous manufacturing process improvements. Technology platform (characterized by device and interconnect models) has turned up as a single factor capability index for competitiveness and requires continuous improvement; however cost for technology platform development is increasing (expected 32nm technology platform cost is 7 billion USD). Every new technology node should have 2x transistor density, ability to quickly ramp up for high volumes with multiple designs and yield to be as good as or better than previous node. We have seen that no changes can be made after the technology is released for production due to requalification of existing design libraries but Intel permits the change in design rules. They have demonstrated with a case study that almost 100% design rules got changed before chip tape out to ensure highest yield [Webb, 2008] resulting in the highest yield. The best explanation for this is that Intel primarily focus on the single product line i.e. Microprocessor which might permit the design rules changes. It could also be due to the fact that they are able to exploit production data source for quick knowledge capitalization. It can be concluded that effective knowledge capitalization is the key which can help us improving the existing technologies for better yield.

The **C**ommon **T**echnology **P**latform (CTP) is a new strategy adapted by the existing IDMs to share R&D efforts [McGregor, 2007] at lower process nodes (<90nm). IDMs also combine the manufacturing facilities to distribute production orders with established yield to support their business strategy as an intelligent move to share IPs, libraries and tools defining core competitiveness. Cost reduction per unit function with increased R&D costs is balanced by increasing yield through enhanced wafer area (4"-1975, 5"-1985, 6"-1990, 200mm-1992, and 300mm-2002). Wafer manufacturers have planned to move to 450mm wafers but neither the industry nor the equipment are ready for this transition. Shifting to 450mm is going to be a big jump and precisely saying won't be possible without moving towards predictive and prescriptive modeling.

One axis that is strongly influenced by DFM competitiveness is technology axis characterized by process, material and equipment innovation. [Quirke et. al, 1992] demonstrated that change in the material for specific components could drastically impact the performance of critical parameters like cost and quality. Process or material change with the latest product generations has presented huge challenges to the industry and forced a close integration of design, layout, process, and manufacturing [James, 2009]**.** The SI under the

significant shifts of Moore, more Moore and more than Moore has benefited 5 technology shrinks from 180nm to 45nm within last decade resulting from improvements in *(i)* Lithography (248 nm, 193 nm, 193 nm immersion) *(ii)* Materials (Al, Cu, low-k, high-k dielectrics, metal gate, resist chemistry) *(iii)* strain techniques (stress liners, embedded SiGe, shallow trench isolation) and *(iv)* Sensitivity to random effects/defects (dopant fluctuation, line edge roughness). In addition to this [Trybula 1995] electronic design automation has empowered industry with capabilities to bridge not only the increasing gap between design and manufacturing but also the cost reduction with computer aided design tools (ECAD/TCAD). The DFM efforts are beneficial if they exceed implementation costs like software license cost, process characterization and model calibration resources, loss of design efficiency and time to market opportunity and loss of layout density [Liebmann, 2008]. Variability has always been present in IC's [Aitken et al., 2008], but the most advanced data analysis tool used to manage variability has been (and still is) "the spreadsheet". There is no single DFM methodology integrated within EDA flow that can distribute residual variance; however authors do agree that DFM has evolved from rule based to model based (physical to electrical DFM) and is defined as broad set of practices that helps produce compelling products at high yield levels on a competitive schedule.

It is evident from the above discussion that DFM is critical for the success of the SI but its effectiveness depends on our ability to exploit the production data source by continuously adding new data dimensions. In the SI business models engineers still use excel for a variety of analysis. It is an important question that even in the presence of advances and sophisticated GUI tools for data extraction and analysis, why R&D engineers are using excel as an intermediate tool. The most logical answer that we received from the engineers in the production line is that newly emerging variations need multi-source analysis, hence they use excel to align multi-source data which is captured at different levels (wafer/site/die). This methodology is not working, because they are not able to find root causes against abnormal drifts or variations.

2.2.8 Industry wide Understanding of the DFM Concept

Let us review how the DFM concept is viewed and understood by the engineers and managers in the SI. The survey conducted in 1989 on DFM success showed that companies have not been successful in implementing the DFM concept because engineers and managers had different perceptions. The reasons are listed as lack of awareness on the importance of DFM by product designers, less understanding of design influence on manufacturing and improper variations analysis [Seino et al., 2009]. This study has proposed technology management methodology with suggestions to enlighten designers with the necessity of DFM providing appropriate manufacturing knowledge, establishing and providing design rules reflecting manufacturability and clarification/correction proposals from manufacturing to design. It is also important to note that the Industrial engineering methodologies cannot prove fruitful in case of outsourcing to the developing countries lacking production skills [Shingo and Dillion, 1989 and Young and Murray, 2007].

We have analyzed DFM scope/evolution in manufacturing and semiconductor industries and concluded that defect and design related yield limiters are the key focus to achieve T2M and T2V. IT revolution has lead significant improvements in DFM methodology to address **C**omplementary **M**etal **O**xide **S**emiconductor (CMOS) scaling challenges through CAD/CAE tools, databases and modeling; however industry needs to stretch CMOS till 22nm to justify investment made in equipment and process R&D. Industrial competitiveness has significantly shifted towards DFM compliant cells and IPs to avoid later corrections and faster first time correct design [Trybula, 1995] that limit design innovations. It is the key reason for freezing the technology once it is transferred to the business model because any change shall require requalification of the reusable design resources.

Today competitiveness is defined as the ability to stretch CMOS through innovation i.e. process, equipment, material, design and EDA. The DFM information is traditionally stored as design rules and efforts these days within DFM are focused on simulation time resulting from gigabit designs [Mason 2007 and Cliff, 2003]. The CAD simulation tools are helpful in this regard but yield related issues keep on

emerging with shrinking technologies. The overwhelming amount of manufacturing data is turning into a potential challenge where knowledge extraction is highly temporal and success lies in timely extraction. A potential solution requires focus not only on the data and its representation but also on the analysis methodologies to achieve benefits. Industry has to shift from data towards information and/or knowledge capitalization with descriptive, predictive and prescriptive models as proposed in this article. We believe that the IDM has an ideal structure to shift from data driven DFM towards information/knowledge driven DFM efforts in comparison with fabless to provide designers sufficient margins.

From the above discussions we can summarize that DFM is focused on manufacturability and yield along with 4 improvement axis (Figure 2.17) as *(i)* technology (process, equipment and material innovations), *(ii)* product design (design driven yield loss) and *(iii)* manufacturing (defect driven yield loss) axes. The defect driven yield losses are controlled through APC/AEC (MFD) efforts characterized as descriptive models whereas design driven yield losses require data mining (DFM) efforts as being predictive/prescriptive models. The MFD efforts map to the horizontal axis (technology scale) whereas DFM efforts are focused on design and manufacturing (vertical axes) and partially technology axis (process and equipment engineering). The process, equipment and material axis constitute the technology scale whereas design and manufacturability axes correspond to the product and variability respectively. The MFD efforts when integrated with DFM, in a fully automated system, provide significant benefits not only in terms of yield improvement and robust process control but also in pattern learning while moving towards process improvements, techno shrink and future nodes. The technology scale directly impacts the design driven defects and design in turn challenge our computational capabilities for potential manufacturability and yield. To have an industry wide uniform understanding of the DFM concept, we can categorize it as *(i)* data driven, *(ii)* information driven and *(iii)* knowledge driven efforts that correspond to descriptive, predictive and prescriptive models.

We can simplify the discussion by defining our objective to move from the data driven DFM efforts (as today) towards information and knowledge driven DFM initiatives. It requires dynamic exploitation of the production data sources to find systematic patterns that can be transformed into rules and/or models. The DFM effectiveness depends on the quality of the input data followed by information integration to generate knowledge. This argument forms the basis for our 2nd and 3rd research question that *(i)* what are the true DFM challenges, *(ii)* what are the limiting factors in the dynamic exploitation of the production data that has led DFM ineffectiveness and *(ii)* what are the generic solutions to address those limitations? The answers to these questions can be found in the chapters 4 and 5.

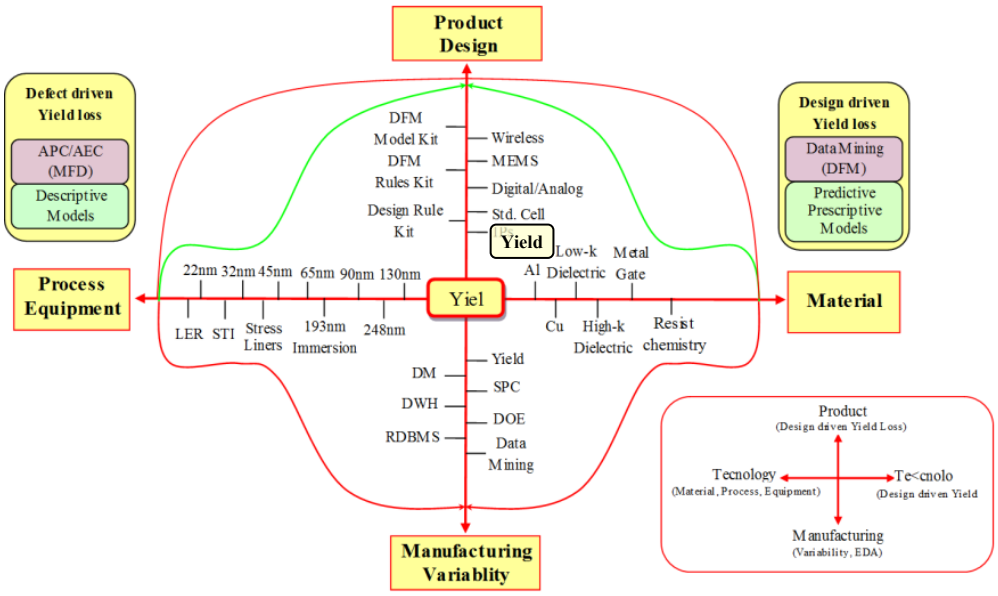


Figure 2.17 - 4-dimensional innovation framework for unified agile DFM system

2.3 INFORMATION INTEGRATION CHALLENGES TOWARDS MORE EFFECTIVE DFM METHODS

From the above discussion we have established a unified understanding of the DFM concept based on the shift from data driven DFM efforts towards information and knowledge driven DFM initiatives. The unified DFM concept is based on the [Zeigler et al., 2000] knowledge models as *(i)* descriptive, *(ii)* predictive and *(iii)* prescriptive where transition among these models require a shift from data towards information and knowledge. The data analysis framework where engineers transform data into information and knowledge is best explained by the **D**ata-**M**ethod-**S**tat (DMS) triangle (Figure 2.18). Data is generated from the methods (manufacturing processes) and is stored at three levels, *(i)* **O**perational **D**ata **S**ources (ODS), *(ii)* **D**ata **W**arehouse (DWH) and *(iii)* **D**ata **M**arts (DM). The statistics includes data analysis methodologies that generate information from data and further transform it into knowledge with machine learning algorithms. The results from this analysis are used to control, align or fix the drifting process so that manufacturability and yield can be ensured. The DMS triangle follows a cyclic improvement pattern to transform data into information and knowledge.

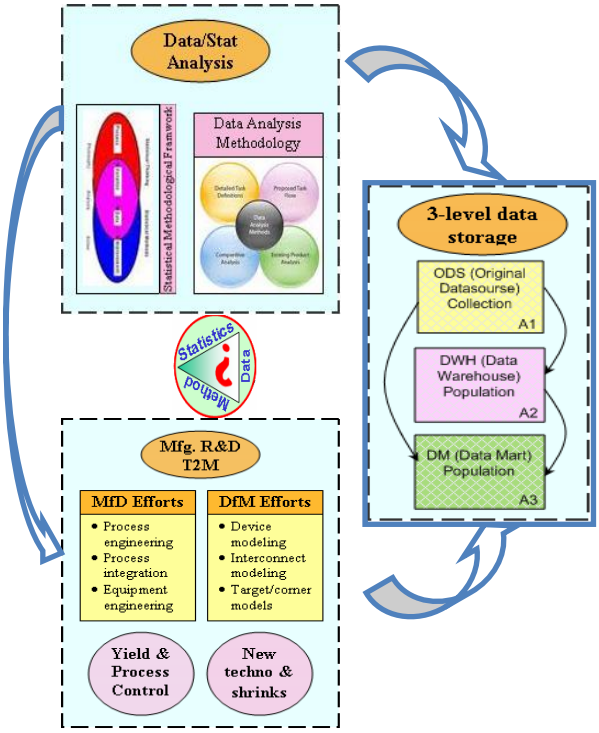


Figure 2.18 - Data-method-stat triangle

The **O**nline **T**ransaction **P**rocessing (OLTP) and **O**nline **A**nalytical **P**rocessing (OLAP) are the key concepts around database management domain where enterprise wide information is stored at three levels ODS, DWH and DM. OLAP is further categorized as ROLAP (relational OLAP), MOLAP (multidimensional OLAP) and HOLAP (hybrid OLAP). The ODS is characterized by the relational database normalized for the insertion, deletion and updation anomalies whereas DWH comprises of Star and Snowflake schemas and DM's are specialized single unrelated tables (views) for routine reports. The selected **D**ata**b**ase **M**anagement **S**ystem (DBMS) plays a very important role in managing the architecture of the data storage and improving query responses by either indexing or creating views for frequently accessed data.

In section 2.2, we have established that the potential cause for the DFM ineffectiveness is the input data, hence we need to find a way to accurately and efficiently integrate data/information so that it can be quickly transformed into information and knowledge. We have also seen that the SI is capable of huge data volume and dimensions, so integration among heterogeneous data sources shall result in effective multi-source root cause analysis for DFM effectiveness. In this section, our objective is to analyse the existing data, information and database integration approaches so that if not best fit is found then an appropriate proposition in line with our scenario can be proposed.

2.3.1 Data/Information Integration Issues

The database technology to store data/information was coined around 1960 and since then engineers have always been complaining about data volume and dimensions to perform accurate statistical analysis. Recent revolutions in IT technologies have enabled huge data volumes and dimensions along with decreasing storage and computational costs. It has raised the need for data integration and first integration approach was introduced in 1980 as Multibase [Landers and Rosenberg, 1982 and Hurson and Bright, 1991]. The data dimension here refers to a specific type of data and in technical terms a new table in the database.

The data collected across the production line is of heterogeneous nature; hence it requires integration prior to multi-source analysis. The data integration is defined as unifying data that share common semantics but originates from different sources. It is expected to address 4 types of heterogeneities: *(i)* structural heterogeneity, involving different data models; *(ii)* syntactical heterogeneity, involving different languages and data representations; *(iii)* systemic heterogeneity, involving hardware and operating systems and *(iv)* semantics heterogeneity, involving different concepts and their interpretations [Gruber, 1993 and Busse et al., 1999]. The new database technologies have solved the syntactic and systematic heterogeneity but semantic and structural heterogeneities are still to be addressed. The semantic heterogeneity is mainly focused in research with reference to data/information integration and it deals with three types of concepts: *(i)* the semantically equivalent concepts, *(ii)* the semantically unrelated concepts, and *(iii)* the semantically related concepts [Cui and O'Brien, 2000]. The Ontology and Metadata are two approaches used till date to remove such type of heterogeneities for an efficient data and information integration.

2.3.2 Ontology from Philosophy to Computer Science

The term ontology has existed within the philosophy and computer domains (artificial intelligence, knowledge representation, databases etc.); however its first existence is traced to Bailey's dictionary of 1721, that defines it as 'an Account of being in the Abstract' [Welty and Smith, 2001]. It was officially introduced by [Gruber, 1993] as an "explicit specification of conceptualization". The conceptualization refers to an abstract model of how people commonly think about a real thing in the world; and explicit specification means that concepts and relationships of an abstract model get explicit names and definitions that can be used in an integration task to describe the semantics of the information sources. It also provides us vocabulary and the relationships between entities within the domain to address the semantic heterogeneity [Goh, 1997 and Cui and O'Brien, 2000].

The ontology in computer science is an agreed and shared understanding (i.e. semantics) of a certain domain with an objective to exchange information and ensure interoperability across autonomous and distributed applications [Tzitzikas et al., 2001 and Spyns et al., 2002]. This concept exist and has been applied in almost all domains e.g. E-commerce [Guarino and Persidis, 2003], Semantic Web [Glaser et al., 2004], Database design [Guarino, 2002], Database integration [Wand et al., 1999], Information access and retrieval [Abdelali et al., 2003], Software engineering [Deridder and Wouters, 2000]. These applications can be grouped under the domains *(i)* artificial intelligence (for knowledge sharing) and *(ii)* database (semantic data models. The ontology development is referred to a group activity, especially in the environment where heterogeneous domains has emerged under a similar domain [Klein and Noy, 2003].

2.3.3 Data/Information Integration

Information systems evolution (autonomous/heterogeneous/distributed/flexible) has driven the need for data/information integration shortly after the introduction of databases, over last three decades and it is an active research area [Wache et al., 2001]. Research in 80's was focused on centralized databases in order to have integrated data/information model however the same converged in 90's to combine data/information within multiple databases [Heimbigner and McLeod, 1985].Today focus has shifted towards loose or tight collaboration of databases under the title of federated database [Sheth and Larson, 1990]. Further it has given rise to the classification of information systems as centralized, distributed, heterogeneous and federated

information systems [Corcho and Gomez-Perez, 2000]. The federated information systems supported with federated database provides a tightly coupled federation layer between information system and federated databases, tightly coupled through schema matching (metadata) and single federated unified schema. Ontology has to deal with the all type of information like structured (databases), semi-structured (XML) and non-structured (web pages). Increasing web based non-structured and semi-structured information have pushed all the concentration in ontology to focus on removing semantic heterogeneity for web resulting in significant developments like **W**eb **O**ntology **L**anguage (OWL), **D**escription **L**ogic (DL), SPARKLE (ontology based language), **R**esource **D**escription **F**ramework (RDF), **D**ocument **T**ype **D**efinition (DTD) etc…

2.3.3.1 Metadata based data/information Integration approaches

The ontology (explicit specification of a conceptualization [Gruber, 1993]) and Metadata (explicitly managed data describing other data or system elements to support reusability and interoperation) based approaches to address semantic heterogeneity are very well circled research areas. The Metadata approach is further classified as [OMG, 1997 and Busse et al., 1999]:

a) **Technical metadata:** provides technical access methods like protocol, speed of connection, cost of queries, query capabilities to address technical heterogeneity.

b) **Logical metadata:** are schemas and logical relationships like data dictionaries in **R**elational **D**atabase (RDB), class diagrams in **O**bject **O**riented **D**atabase (OODB) and a global schema that captures relationships and dependencies between several schemas e.g. schema metadata in **F**ederated **I**nformation **S**ystems (FIS).

c) **Metamodels:** metadata supports the interoperability of schemas in different data models and address data model heterogeneity.

d) **Semantic metadata:** describe semantics of concepts (domain-specific descriptions) through ontologies.

e) **Infrastructure metadata:** helps users to find relevant data through navigational aids like annotated bookmarks as well as a thesaurus structure.

f) **User-related metadat**a: describes responsibilities and preferences of information system users (user profiles).

The ontologies and database schemas (semantic data model) are closely related [Cui and Brien, 2001], with few tangible differences, but exceptions do hold. The key difference is that ontology represents common shared knowledge whereas schema is focused on the contents or data instances; hence, both schema and ontology are important for heterogeneous data/information integration. One common solution to address the ontology vs. schema approaches for integration is to use three layer storage architecture i.e. original data sources, data warehouses and data marts where information/analysis needs of the users are fulfilled from DWH and DM [Henderson, 1998]. They provide a limited solution, but it depends on the schema architecture designed to incorporate a solution for heterogeneity or be an aggregated information store.

The semiconductor industry is high-tech with potentially growing volumes of temporal data and urgent need for data alignment, extraction and integration require a solution that includes remedy for semantic and structural heterogeneity with model evolution management. The important aspect to be noted here is that schema (metadata) based integration approaches have led to the development of technology standards like DTD and XML Schema which provides an opportunity for heterogeneous information integration.

2.3.3.2 Ontology based data/information Integration approaches

The ontology development, mapping and management are very well defined and matured domains demonstrated for the data/information integration. There are three major ontology architectures *(i)* single

ontology (global ontology for shared vocabulary for semantics e.g. SIMS [Arens et al., 1996]), *(ii)* multiple ontology (multiple ontologies no common shared vocabulary e.g. OBSERVER [Mena et al., 1996]) and *(iii)* hybrid ontology (multiple ontologies with shared vocabulary e.g. COIN [Goh, 1997]). It is also used as a global query schema, e.g. SIMS [Arens et al., 1996] user query is expressed in terms of ontology which is reformulated in respective sub-queries followed by results combination. The ontology acts as a global schema for query; hence it requires an automated tool otherwise it would be difficult for the user to formulate a query from this schema without knowing each and every detail of individual schema [Wache et al., 2001]. There are many languages for the representation of the ontology as: description logic (DL), terminological (T-Box: vocabulary of the application domain), non-terminological (A-Box: assertions about vocabulary) and Classical Frame based languages.

[Buccella and Cechich, 2003] proposed a three step approach for data integration where first of all shared vocabulary is developed, followed by local ontology development and then associated mapping. This approach is only focused on reducing semantic heterogeneity. The case study from an automobile industry [Maier et al., 2003] demonstrates the benefits of using ontology for efficient and reliable data integration in **P**roduct **L**ife **C**ycle **M**anagement (PLM). It is a 6-step methodology: *(i)* schema import (table is taken as concept), *(ii)* creating relations (logical relations between tables), *(iii)* create mappings (concept to concept, attribute to attribute and attribute to concept), *(iv)* business logic (deductive logic), *(v)* rule based modeling, *(vi)* inferencing and schema export. It is recommended as a beneficial methodology for heterogeneous data integration.

[Buccella et al., 2005] compared 7 systems (SIMS, OBSERVER, DOME, KRAFT, Carnot, InfoSleuth and COIN) for ontology based distributed and heterogeneous data integration including structured and unstructured data sources (web). Framework used for evaluation purpose is architecture, semantic heterogeneity (semantically equivalent concepts, semantically unrelated concepts and semantically related concepts) and query resolution [Cui and Brien, 2000]. Many ontology based data-integration surveys can be found, that focus on ontology development and mapping whereas other focus on languages used to represent ontologies [Wache et al., 2001 and Corcho and Gomez-Perez, 2000].

The query resolution is an important but neglected factor within ontology based data integration tools/systems, however optimization approaches could be paramount in overall ontology based system efficiency for integration. The SI have three-level storage architecture hence ODS is expected to store the measurement/observation time as transaction time whereas aggregated information is recorded through ETL routines in DWH and DM's. Risk of database servers being overloaded due to heavy queries from users might result in temporal data mismatch resulting in an opportunity loss. It is necessary to find a solution for an efficient data extraction and integration considering temporal nature of data in such a way that scheduled ETL (extraction, transformation and loading) routines are least impacted. This 3-layer storage architecture provides an opportunity to handle temporal issue but still require semantic and structural heterogeneity to be addressed.

2.3.4 Ontology based Database-Integration Approaches

Let us move one step ahead and analyze ontology based approaches used for database integration. The database integration provides system interpretability and requires first of all mapping between schemas and then integration system to answer the user queries [Doan and Halevy, 2005 and Lenzerini, 2002]. The schema mapping provides a model that can also be used for inter-schema correspondence [Rahm and Bernstein, 2001]. [Dou and Lependu, 2006] used two databases (Nwind and Store) from similar domain to first learn ontology from respective schema and then perform mapping "ontology merging" using TOOL called OntoGrate with a focus on optimizing query performance results. In this approach user queries are translated from WEB-PDDL to SQL and executed for faster results and methodology is named "inferential data integration".

[Aparício et al., 2005] has demonstrated that ontologies help us in addressing semantic heterogeneity better than building global schema, hence database integration if supported with building ontology (single or multiple) followed by query answering using global schema proves its benefits. In this approach the global schema is presented with database as SQL view to enhance the performance of system. [Fiedler et al., 2005] present a very interesting debate and argues that the solution heterogeneous databases integration is difficult but it can be achieved with assumptions; however they suggest that collaboration instead of integration is better approach in such scenarios. The idea is very simple, that global schema provides possible opportunities to merge the databases resulting in a global view for the global users. The coordination concept presented here is similar to the correspondence table between databases and provide means to communicate, coordinate and cooperate through SQL views. Such ideas can be tested with few small databases but serious problems shall arise in case of huge temporal databases with semantic heterogeneity.

So far all above approaches talk about developing ontology and then using it to find similarities between schema for query model but no approach talk about possible real time live databases where schema or concepts could be added on as and when required basis; hence it results in the need for dynamic resynchronization. In our case, the scenario is quite complex, overall domain is semiconductor manufacturing, but diversity and heterogeneity in comparison to the evolution during last two decades, presents a challenge for information integration. In order to keep-up pace with the technological evolutions it has become important to push the concept of data/information integration and knowledge acquisition for competitive advantage.

2.3.5 RDB Integration based on Schema Matching

Schema matching is a critical step towards integration (RDB to data-warehouse and schema matching between heterogeneous data warehouses) but it differs from ontology in a sense that it is focused on explicit information contents whereas ontology provides unified concepts and associated attributes. Schema matching shall result in global schema to be used as query model and it is a difficult task. Till today there exist very few efforts on developing autonomous schema matching efforts and mostly it is done in a semi-automatic or manual fashion with the help of domain experts. Semantic common points identified within schema matching shall result as integration points not only for global schema but also for data integration [Palopoli et al., 2000]. The proposed approaches are automatic and semi-automatic based on large schema having 19-21 attributes but real time online applications where ODS and DWH schemas have 50-100 tables and 10-50 attributes each, results in a huge problem. It becomes more critical when added with temporal nature of data being stored through ETL routines using advanced automation systems.

The ontology usage is increasing within the web applications (B2B, B2C & C2C) running (365*12*7) to search, navigate, browse and share knowledge, with scalability, availability, reliability and performance [Ramakrishnan and Gehrke, 2003 and Das et al., 2002]. The ontology management with increasing need to incorporate all emerging data sources has emerged as one critical research area that require special focus and attention not only on the security, best ontology and knowledge representation techniques but also on the storage and efficient querying mechanisms. The RDBMS are being used to achieve this performance level over OODBMS [Abernethy and Altman, 1998 and Stoffel et al., 1997]. Many ontology development environments have been proposed with varying strengths to address specific needs but most significant developments so far have been on the web for knowledge sharing across heterogeneous data sources.

The ontology concepts have been applied in clinical research and medical domain with an effort for knowledge sharing and common vocabulary development. Manufacturing R&D industries like SI are competitors in terms of market share and associated sales revenues; hence the streamlined ontology concepts for the knowledge sharing on the web as presented in the literature are difficult to be used in the SI because of their conflicting business objectives.

2.3.6 RDB Schema to Ontology Mapping Approaches

Due to increasing volume of data across web, relational database to ontology mapping has emerged as a significant domain for knowledge sharing and exchange of data [Konstantinou et al., 2006 and Liddle et al., 2003 and Calvanese et al., 2001]. [Zuling and Hongbing, 2005] proposed a methodology to map relational databases (using MySQL) to ontology by mapping its schema and instances to T-BOX and A-BOX (using "Protégé" software) and then user request is parsed into RDQL (RDF query language) which is transformed to formulate SQL query for execution. There are certain questions which are still unanswered like in a situation if the instances are in the size of terabytes: are we obliged to bring database instances within the developed ontology domain where SPARKLE or any other web semantic query language can be used to answer user requests? Perhaps it is not a good solution as far as query performance and storage space is concerned.

[Calvanese et al., 2001] proposed the notion of ontology of integration where we intend to develop a global ontology from local ontologies for information extraction. [Zuling and Hongbing, 2005] proposed an interesting idea to implement ontology database with user defined rule based querying and reasoning. [Astrova et al., 2007] propose to store the web ontologies into relational databases in order to benefit from computational efficiencies. [Curé, 2005] proposed DBOM (database to ontology mapping) framework for semantic web data integration. [Nyulas et al., 2007] developed a plug-in to be used with "protégé" to import database schema with first hand mapping to OWL and provides a strong feature where you can and cannot import data along with the schema, in case if data is not imported it can be accessed through OWL-database bridge.

2.3.7 Ontology Driven Data Extraction Tools

The ontology development and management environment (freeware and commercial) exploded as the interest in web information integration increased. It has resulted in the emergence of technologies like RDF, XML schema, DTD, OWL etc. Our focus is on data alignment, integration and extraction of temporal manufacturing data; hence we shall focus on the integrated environments developed for data/information extraction and integration based on 3-layer storage (ODS, DWH and DM). Two commercial and three freeware ontology tools are selected out of 100+ available based on storage schema (database) as under for possible integration and utilization of schema matching and ontology management [Youn and McLeod, 2006].

a) **Link-Factory Workbench (commercial tool):** It is a 3-tier architecture for large medical ontologies, user interface is client-side application that is used to connect to the business tier (implemented as ontology server) to manage or query data-tier i.e. relational database. Input output formats supported are XML, RDF, OWL and multi user capabilities with information extraction whereas ontology stored in the database is just used for linking purpose. It does not support graphical output view.

b) **K-Infinity (commercial tool):** Knowledge builder utility with two major components, graph and concept editors, facilitates the knowledge engineers to build objects, relations, network of knowledge, concepts, individual attributes and relations.

c) **Protégé 2000 (freeware):** It is developed by Stanford University, USA and it is a freeware having a lot of functionalities. It is a graphical tool for domain ontology development and management supported with plug-ins to graphically view the tables and diagrams. Further it can be used to learn ontology from relational database schema with and without importing instances. Query requests require an understanding of RDF and SPARKLE. The I/O format is RDF, XML Schema and Java [Arens et al., 1996, Mena et al., 1996 and Ramakrishnan and Gehrke, 2003].

d) **WebODE (freeware):** It is one of the powerful freeware developed to serve three purposes *(i)* ontology development, *(ii)* ontology integration *(iii)* ontology based application development, all three purposes are met through graphical user interface provided with its 3-tier ontology editor

WebODE. It is developed in JAVA; hence support CORBA, XML and RMI. It provides GUI, web and multi user support including prominent web formats.

e) **ICOM (freeware):** It is a powerful tool in terms of capabilities of conceptual **E**ntity **R**elation (ER) modeling with an objective to ensure the visualization of all constraints as one single conceptual model. It could also be used for the translation of ER conceptual models developed in ontologies for various purposes. ICOM tool reasons with (multiple) diagrams by encoding them in a single description logic knowledge base, and shows the result of any deductions such as inferred links, new constraints, and inconsistent entities or relationships. Theoretical results guarantee the correctness and the completeness of the reasoning process.

One thing is apparent that ontology has evolved with a focus on sharing and exchanging web knowledge. There is no approach proposed so far that takes into account the model evolution and its dynamic synchronizations. In SI, the databases are growing at an immense pace in data volume and dimension; hence we are obliged to continuously restructure the data models so that new dimensions can be timely updated. Our success lies in our ability to dynamically exploit all data volumes and dimensions to find root causes against the abnormal drifts emerging from spatial variations. So we need an ontology based system that can pre-assess the failures of potential changes in the data models. These assessments must be made at the application and user levels. Any potential changes must follow the compliance loop (updation) where end user application is affected by this change and end users must be informed for all structural changes in data dimensions. This argument complement 2nd research question that how we can enable dynamic restructuring of data models to continuously support the addition and deletion of data dimensions? The answer to this question can be found in chapter-6.

2.4 SUMMARY AND CONCLUSIONS

In this chapter the literature has been analyzed across three dimensions *(i)* semiconductor industry background and its evolution, *(ii)* role of DFM methods and current challenges and *(iii)* information integration approaches for dynamic data exploitation. It is observed that the SI is characterized by the cyclic demand patterns with positive CAGR (+8.72%) that guarantees a cumulative demand; hence we are obliged to respond to the growing market with new, fast and high value but low cost products. It has led a shift in business objective (T2M, T2V to ramp-up rate) and emergence of new technical challenges (design and manufacturing interface complexities). The SI can only respond market growth by introducing new technologies every 2 to 3 years or quickly maturing the technology derivative and improvement efforts with local DFM efforts. The DFM methods had been adopted by the SI around 1980 as a yield enhancement strategy which worked well till 250nm technology node, but beyond this node DFM has resulted in high cost and ineffective R&D efforts, creating a challenge in accomplishing quick ramp-up rates with the existing SI business models.

In order to address the increased technology lead times and R&D costs, the SI model (IDM) has structurally transformed into fablite and fabless business models. But still, IDM fablite model is reported to be the best models in terms of revenue generation. We argue that the IDM fablite model provides a coherent platform for the knowledge capitalization from production line and our ability to improve the challenges faced by the R&D engineers shall result in the continuous improvement efforts for technology alignment or adoption. It shall reduce technology lead times, associated costs along with an opportunity for early penetration in the market with higher profits.

The SI is a 300+ B$ industry with an equal opportunity to capture maximum market for its stake holders. In order to get maximum share, it is necessary to revisit and rank the SI objectives so that a strategy can be formulated truly in line with the top ranked objective. The existing business models (IDM, Fablite and Fabless) must be analyzed to evaluate their potential to support the newly formulated business strategy. In case if none of the existing model is capable to support it, a new model or an extension to the existing

business models must be proposed to ensure its compliance with the new strategy in line with the top ranked business objectives. This argument formulates our first research question, which is addressed in chapter-3.

The SI emerged as one of the most complex, competitive and technologically fastest growing manufacturing domain. From the literature review on the evolution of the DFM methods, it is evident that success lies in the ability to put DFM back to track. Few significant factors identified being major causes for DFM ineffectiveness *(i)* inability to exploit the production data for multi-source root cause analysis, *(ii)* improper analysis due to difficulties in data alignment because of the varying metrology coordinates and *(iii)* heterogeneous understanding of the DFM concept. In this chapter, a unified understanding of the DFM concept as the data driven R&D effort is proposed with an objective to shift these data centered efforts towards information and knowledge centered R&D initiatives. It requires dynamic exploitation of the production data sources to find systematic patterns that can be transformed into rules and/or models. The DFM effectiveness depends on the quality of input data followed by information integration and analysis to generate knowledge. This argument forms the basis for 2nd and 3rd research questions as presented in section 1.4. They can be further divided into smaller questions as *(i)* what are the true DFM challenges, *(ii)* what are the limiting factors in exploitation of the production data sources that has led the DFM ineffectiveness and *(ii)* what are the generic solutions to address those limitations? The answers to these questions can be found in the chapters 4 and 5.

The data analysis framework where engineers transform data collected across the production line into information and knowledge is best explained by the data-method-stat (DMS) triangle. The data is generated from methods (processes) and is stored at three levels, *(i)* operational data sources (ODS), *(ii)* data warehouse (DWH) and *(iii)* data marts (DM). The Statistics includes analysis methodologies that generate information from the data and further transform it into knowledge with advanced statistical and/or machine learning algorithms. The results from these analyses are used to control, align or fix the drifting processes so that manufacturability and yield can be ensured. The DMS triangle follows an improvement cycle to transform the data into information and knowledge. The online transaction (OLTP) and analytical processing (OLAP) are key concepts which are built around the methodologies to store and process the data to generate information and knowledge. The enterprise wide information is stored at three levels ODS, DWH and DM. The OLTP system is based on the relational databases where end-user queries are optimized for transactional processing (insert, delete, and update). In Comparison to the OLTP, OLAP systems are based on the DWH and DM concepts and are focused on the query performance and quick transformation of data into information and knowledge for decision making. The data is stored in a multidimensional array in the OLAP systems which results in fast slicing, dicing and drill up/down operations. The primary objective of the OLAP system is on fast data aggregation. In the SI, R&D engineers are primarily focused on multi-source analysis not for the purpose of aggregation but for mapping and alignment so that effective root cause analyses can be performed to correct the drifts processes.

The above presented OLAP and OLTP systems do not match with the objectives of R&D engineers; however both systems offer interesting benefits which can be used to improve the R&D productivity. The OLAP system and multidimensional modeling can be used to quickly access multi-source data from multiple databases accumulate in huge volumes for the purpose of mapping and alignment instead of aggregation. The slicing, dicing and drill up/down are of very high interest which can help engineers to find root causes across multiple dimensions of the data sources. The OLTP systems offer transactional efficiency; hence we can use this quality to manage data model evolutions. The difficult part of the SI is that ODS and DWH data sources are of proprietary nature; hence they cannot be changed easily. It has resulted in serious issues about the inclusion of new data dimensions and huge data volumes. The consequence of this fact is that R&D engineers are not able to exploit the production data sources which indirectly contribute to the DFM ineffectiveness.

The data collected across the production line is of heterogeneous nature; hence it requires integration prior to multi-source analysis. The data integration is defined as unifying data that share common semantics but originates from the different sources. The SI being high-tech with potentially growing volumes of temporal data and urgent need for data alignment, extraction and integration require a solution that includes remedy for the semantic and structural heterogeneities with model evolution management. The ontology and Metadata based data, information and database integration approaches are used most commonly for the sharing and exchange of web data. These approaches neither support the huge data volumes nor the dynamic restructuring of data models.

In SI, the databases are growing at an immense pace in data volume and dimensions; hence success lies in the ability to dynamically exploit huge data volumes and dimensions for an effective root cause analysis against the abnormal drifts emerging from the spatial variations. So we need an ontology based system that can pre-assess the failures of potential changes in the data models. These assessments must be made at the application and user levels. Any potential changes must follow the compliance loop (updation) where end user application is affected by this change and end users must be informed for all structural changes across all data dimensions. This argument helps us in complementing 2nd research question that how we can enable dynamic restructuring of data models to continuously support the addition and deletion of new data dimensions? The answer to this question can be found in chapter-6.

Chapter 3: An Extended IDM (*e*-IDM) Business Model [8]

In this chapter, SI business models are analyzed for their support to a recent shift in the business objectives towards quick ramp-up-rate. The **S**trategic **C**reative **AN**alysis (SCAN) and **S**trength, **W**eakness, **O**pportunity and **T**hreat (SWOT) analyses techniques are used in addition to the brain storming sessions to *(i)* identify key improvement areas in today's most successful IDM-fablite business model and *(ii)* formulate business strategy to achieve maximum economic benefits. The technology development/improvement process is further analyzed to find the limitations that have led to the DFM ineffectiveness and extended ramp-up rates. The challenges found here along with the key improvement areas form the basis for the proposition of an extended IDM business (*e*-IDM) model with integrated DFM methodology.

Contents

[8] SHAHZAD M.K., HUBAC S., SIADAT A., TOLLENAERE M., An Extended Business Model to Ensure Time-to-Quality in Semiconductor Manufacturing Industry, International Conference on Enterprise Information Systems, 2011, Portugal

3.1 INTRODUCTION

The traditional SI business model (IDM) has structurally transformed into IDM-fablite, fabless and foundry business models in order to share the exponentially increasing technology R&D costs and lead times. It is very important to understand that to cope with the cumulative demand growth (+8.78%), new IC technologies must be introduced every 2-3 years. The design and manufacturing interface complexities (new manufacturability and yield loss mechanisms) due to miniaturization are rising and being compensated with the design, process, material and equipment innovations. It requires more human resource and experimentations that give rise to technology development costs to ensure its timely introduction in the market. So it is evident from the above facts that structural transformation was just to share the increasing R&D costs. In the current situation an IDM-fablite is classified as the best model for associated economic opportunities.

The recent shift in the SI business objectives from T2M and T2V towards the ramp-up rate needs to be reassessed in the context of potential opportunities associated with cumulative demand growth (+8.78 CAGR). So in simple words, one needs to reconsider that the best IDM-fablite model is capable to sustain this growth in demand for maximum economic benefits. This assessment has been divided into three parts as *(i)* identify business strategy in compliance with increasing demands and economic opportunities, *(ii)* assess existing business model for possible compliance with the quick ramp-up objective and *(iii)* find key improvement areas and/or propose a business model.

3.1.1 Strategic Planning and Analysis

Strategic planning and analysis is a process in which the business environment, its operations, and its interactions are identified to formulate a strategy that can improve the organizational efficiency and effectiveness by increasing the organization's capacity to deploy its resources intelligently [Worrall, 1998 and Business-Directory, 2007]. There is a wide range of analytical models/tools that can be used for the formulation of a strategy including *(i)* SWOT analysis, *(ii)* PEST analysis, *(iii)* Porter's five forces analysis, *(iv)* Four corner's analysis, *(v)* Value chain analysis, *(vi)* Early warning scans, *(vi)* SCAN analysis and *(vii)* war gaming. It is highly important to select the most appropriate tool for the strategy formulation; hence each tool's strengths and constraints are assessed as a first step.

a) **SWOT analysis:** It is most widely used approach in which the **S**trengths, **W**eaknesses, **O**pportunities and **T**hreats associated with the business activity are identified. The first step is to define business objective of the activity and find internal and external factors important to achieve the objectives. The strengths and weaknesses are usually internal whereas opportunities and threats are always external. It is a generic tool and has a wide application for the formulation of a business strategy to achieve the objectives [Mindtools, 2007].

b) **PEST analysis:** It is a tool used to understand the **P**olitical, **E**conomic, **S**ocio-cultural and **T**echnological environment of the organization and is commonly used to analyze the market growth, decline or such factors to align the future business directions. The four PEST factors can be treated like opportunities and a threat as in SWOT analysis but its emphasis is more on the socio-economic factors. Its application is highly dependent on the type of business and or activity being analyzed [Businessballs, 2006].

c) **Porter's five forces:** This concept is based on the five forces and is used to find out the power center in the current business situation. These five forces are *(i)* suppliers power, *(ii)* buyer power, *(iii)* competitive rivalry, *(iv)* threat of substitution and (*v)* threat of new entry. By understanding power center, this concept can also be used to find the areas of strength, to improve weaknesses and to avoid mistakes [Porter, 1996 and Porter, 1998].

d) **Four corners analysis:** It is a predictive tool designed by Michael E. Porter that helps in determining a competitor's course of action and is based on the basic principle of knowing what motivates the competitor. This dimension helps in finding a more accurate and realistic assessment of the competitors possible moves [Gilad, 2011]. The four corners are *(i)* motivation (drivers), *(ii)* motivation (management), *(iii)* actions (strategy) and *(iv)* actions (capabilities).

e) **Value chain analysis:** It is based on the concept that all the activities in the organization create and add a value. This methodology is used to identify those activities which do not add value; hence they can either be removed or replaced with more efficient activities. Each value adding activity is considered a source of competitive advantage [Porter, 1996].

f) **Early warning system:** The purpose of strategic early warning system is to detect or predict strategically important events as early as possible. They are often used to identify the first strategic move from competitor or to assess the likelihood of a given scenario becoming reality. The seven components of this system are market definition, open systems, filtering, predictive intelligence, communicating intelligence, contingency planning and the cyclic process [Comai and Tena, 2007].

g) **War gaming:** This technique is used to find competitive vulnerabilities and wrong internal assumptions about competitors' strategies. They are used for critical strategic decisions but depend on the correct assessment of the competitors moves [Treat et al., 1996].

h) **SCAN:** It is used for the formulation of business strategies by first identifying and listing the business objectives. These business objectives are ranked and then the top ranked business objective is selected for further investigation using SWOT analysis that result in potential set of strategies. The most appropriate strategy is selected and is always based on the strengths, weaknesses, opportunities and threats [Winer, 1993].

If above listed commonly used tools are analyzed for the formulation of a business strategy, it can be determined that early warning system, war gaming, four corner analysis and peters' five forces system are specifically designed for marketing environment in which one is obliged to assess the moves of his competitors to take an appropriate action. Although they help in the formulation of strategic moves, they do not directly comply with our generic needs. The value chain analysis is an excellent tool used to identify inefficient activities so that they can be improved afterwards. But in this case, it is quite difficult to quantify the value against each activity; hence, value chain does not seem appropriate either. The PEST analysis emphasize on socio-economic elements for the business strategy which are not relevant in our case. The SCAN and SWOT analysis being generic for the business strategy formulation are applicable in our case. The SCAN analysis has an advantage over SWOT as it starts by ranking the objectives followed by SWOT analysis to formulate a business strategy in line with the top ranked objective.

3.1.2 SCAN Analysis

The SCAN analysis is the approach selected in this case based on its inherent ability to first rank the business objectives followed by the business strategy formulation based on the SWOT which is truly in line with the top ranked objective [Winer, 1993]. Below is a short description of the two steps in the SCAN analysis:

3.1.2.1 Ranking Business Objectives (Step-1)

Listing the objectives, being followed by organization, is the first step of the analysis. The SCAN analysis is primarily dependent on the selected objective(s); hence potential **T**op **R**anked **O**bjective (TRO) must be carefully selected. An objective is chosen from the list that seems to be an ultimate goal of the organization and then a question is posed: "Why is the organization pursuing this objective?". The answer is to be found from the list of objectives and selected objective is placed at a higher ranking (level). This process continues until an answer cannot be found and this objective is called the top ranked objective. Once the TRO is found, the same process is repeated by asking the question, "How we can achieve this objective?" with TRO from

top to bottom. This step results in the readjustment of certain objectives. The conclusion is that by repeatedly asking the "Why and How?" questions, we find TRO which is the objective of SCAN Analysis.

3.1.2.2 SWOT Analysis (Step-2)

The business strategies are the means through which goals and missions are accomplished. A successful strategy focuses the four elements to generate competitive advantage as *(i)* strengths, *(ii)* weaknesses, *(iii)* opportunities and *(iv)* threats, a.k.a. SWOT analysis. It is believed to have started with the term SOFT (**S**atisfactory, **O**pportunity, **F**ault and **T**hreat) that resulted from the research work on corporate planning conducted at Stanford Research Institute (SFI) from 1960-1970. It is known that SOFT analysis was presented in a seminar at Zurich in 1964 and Urick and Orr changed F to W and called it SWOT [Humphrey, 2005] analysis. Its history can be accurately traced back to Ken Andrews in the early 1970s [Andrews, 1980] which was subsequently modified in the format of a 2*2 matrix, matching the internal factors (strengths and weaknesses) of an organization with its external factors (opportunities and threats) to systematically generate strategies [Weihrich, 1982].

The SWOT analysis does not provide any guidelines that how the strengths and weaknesses can be identified; hence its accuracy is dependent on the manager's perception. However it is an excellent tool to bring in line the business strategy with top ranked business objectives to quickly generate the competitive advantage.

3.2 SCAN ANALYSIS: PART-1 (TOP RANKED BUSINESS OBJECTIVES IN SI)

Keeping in view the objective to establish the business strategy that can comply with the recent shift in the business objective towards ramp-up rate, brainstorming sessions with engineers and managers were held to highlight the potential objectives. The prepared list of objectives includes technical as well as business objectives, independent of any specific technology. The list of candidate potential objectives and the ranking of business objectives are presented as under:

Sr.	Objective Statement	Sr.	Objective Statement
1.	Get leadership position for ST in SI	12.	Introduce new technologies every 2 to 3 years
2.	Develop new technology alliances/partnerships	13.	Reduce R&D costs for new technology development
3.	Improve sales revenues	14.	Get maximum market share
4.	Quick ramp-up rate	15.	Reduce technology alignment lead times
5.	Improve silicon based device model validation	16.	Improve silicon based interconnect model validation
6.	Improve manufacturing flows	17.	Improve design platform
7.	Improve DFM methods ineffectiveness	18.	Improve R&D efforts efficiencies
8.	Achieve customer satisfaction	19.	Reduce product cost
9.	Improve yield	20.	Reduce time-to-market and time-to-volume
10.	Improve quality and reliability	21.	Reduce technology adoption lead times and costs
11.	Improve MFD (APC/AEC) effectiveness	22.	Improve root cause analysis

Table 3.1 – List of SI Objectives

The Initial SCAN analysis results are presented in the Figure 3.1 below. It is evident from the analysis results that "to get leadership position" is identified as a top ranked objective; however, other appropriate objective e.g. "quick ramp-up-rate" is also selected for further analysis to support the top ranked objective, using SWOT technique.

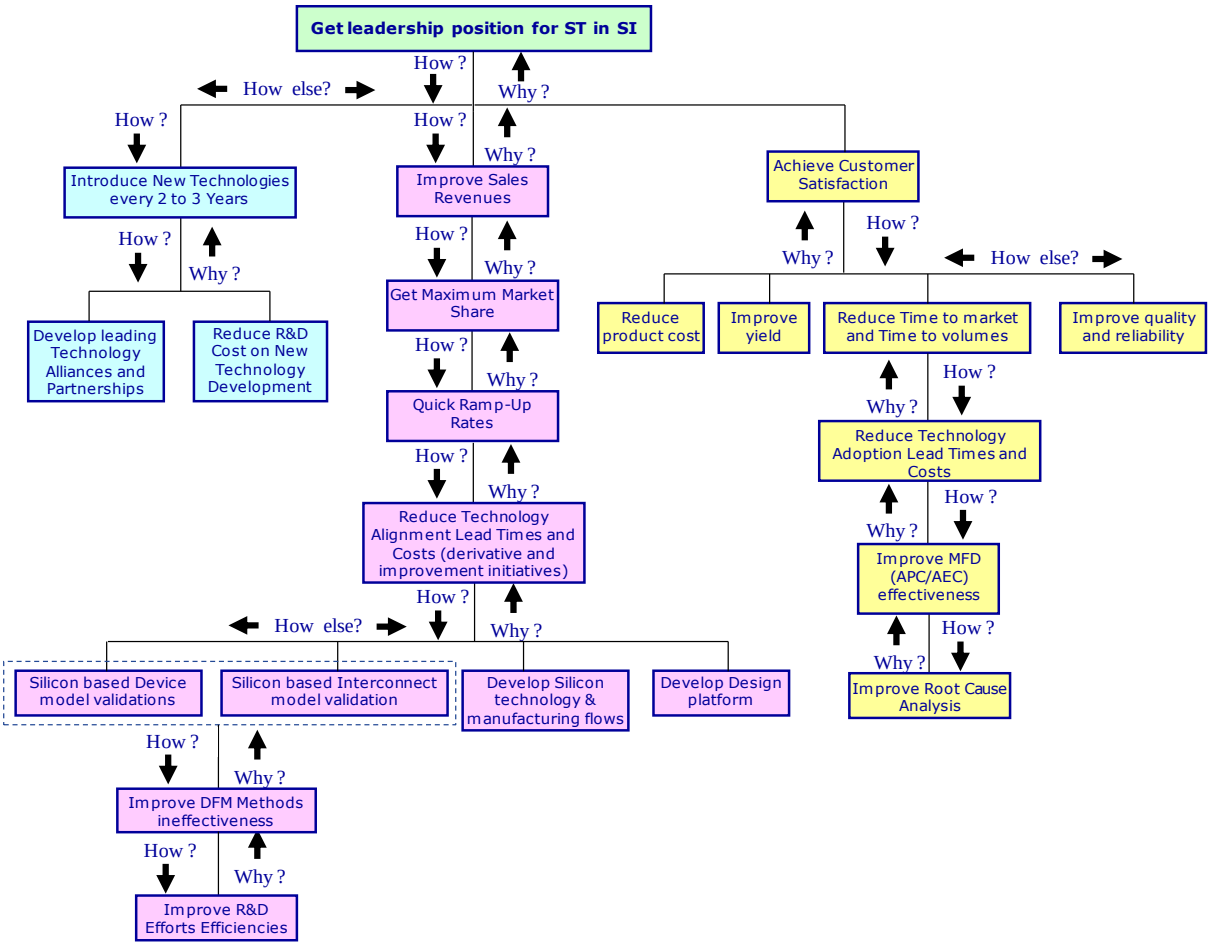


Figure 3.1 – Ranking of SI business objectives

In this analysis, the listed objectives are presented in rectangular boxes and they can be read either from bottom to top by asking question "why is this objective being pursued?" or from top to bottom by asking question "how can this objective be achieved?". Answer to these questions can be found in terms of another objective on the end of the link. The starting question that was posed was why R&D efficiencies should be improved, and the answer is that they should be improved so that the DFM methods can be improved. The next question is: why DFM methods should be improved? The answer is that, it should be improved because the quick silicon based device and interconnect models validation are desirable. This process continues until the top ranked objective << to get a leadership position and quick ramp-up rate >> are reached. Now, looking at the example from top to bottom flow, a question is posed that how a leadership position for STMicroelectronics can be achieved, and the answer is that by introducing new technologies every 2-3 years, maximizing sales revenues and total customer satisfaction. The "how else" part supplements the answer by providing evidence for the additional objectives that can contribute towards the target objective.

3.3 KEY IMPROVEMENT AREAS IN IDM-FABLITE BUSINESS MODEL

Before moving towards the SWOT analysis (2nd step in SCAN analysis) to formulate a business strategy, the IDM-fablite business model needs to be assessed for the strengths and weaknesses. A presentation of the IDM-fablite business model to analyze its key elements and operations is an appropriate first step (Figure 3.2).

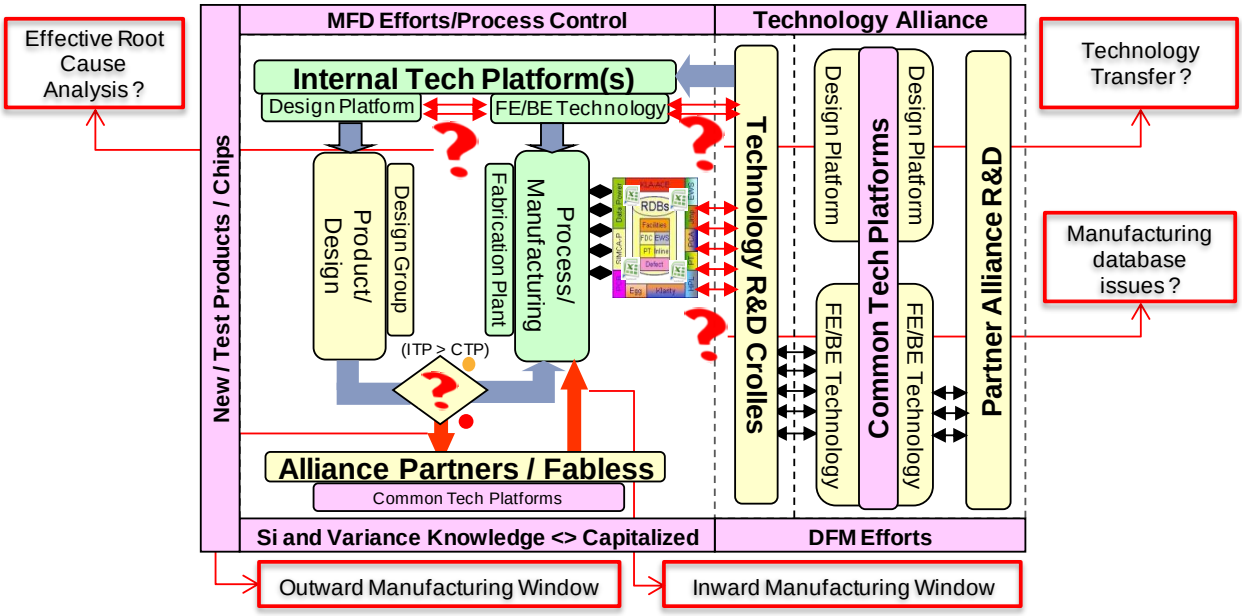


Figure 3.2 – IDM-fablite business model operations

In the above IDM-fablite business model, it is evident that new technology a.k.a. technology platform is developed in an alliance where front-end and back-end (FE/BE) technologies are developed using DFM methods. The FE refers to the technology used for fabrication of transistors and interconnects whereas BE refers to the technology used for packaging and chip integration. It is to remember that every new technology explores the possible potential design, process, equipment and material innovations that give rise to extensive experimentation and R&D analysis and success of DFM methods depends on effective R&D efforts. The new technology is transferred into IDM business model where it is referred as **I**nternal **T**echnology **P**latform (ITP). This newly transferred technology is first aligned on the internal manufacturing resources; hence, it is identified as the first improvement area where our objective is the fast technology transfer. Its alignment requires huge R&D efforts to perform silicon based validation of device and interconnect models, design rules, DFM rules and DFM models.

Once aligned, this technology is further used by the design groups to design new products. The next step is to make a decision that where this newly designed product shall be manufactured. It is based on criteria ITP > **C**ommon **T**echnology **P**latform (CTP) which means that if ITP is more robust in terms of yield and reliability (indicators computed from the R&D efforts effectiveness) then the product is manufactured in our facilities (internal manufacturing window) otherwise it shall be manufactured in alliance partners facilities (external manufacturing window). To remain competitive, SI needs to send products quickly to the market with highest production yield and this is not possible without a robust/mature technology [Morinaga et al., 2006]. The success of this IDM-fablite models requires that every new technology should have 2x transistor densities, ability to ramp-up quickly (Design rules and DFM rules) and yield to be as good as or better [Webb, 2008] than previous nodes (trading off DFM constraints).

The DFM plays a significant role in the new technology development; however it can also be equally used for the technology derivative and improvement initiatives with the support of the engineering information systems (EIS). The principle design of these EISs is coherent with the operational efficiencies and supports only the data driven DFM efforts (single-source root cause analysis). The objectives of an IDM-fablite model are to *(i)* ensure ITPs' backward compatibility with alliance partners' CTP (keep intact outward manufacturing window) and *(ii)* continuously inject competitiveness by improving ITP (enlarge inward manufacturing window). In a technology alliance, partners have access to CTP, similar equipment and material but still one partner get more market share and enjoys high profit margins than others, why? Answer to this question is not trivial, so two questions are formalized and used within brain storming sessions during SWOT and use-case analyses as under:

a) Do methods exist (Fig.3.2) to improve device and interconnect models (technology alignment) and product development process (technology adoption)?

b) Can manufacturing databases and EIS support continuous improvement in the ITP (Fig.3.2)?

3.4 TECHNOLOGY DERIVATIVE/IMPROVEMENT PROCESS ANALYSIS

In the context of above defined questions in section 3.3, technology derivative or improvement process is analyzed. The answers are further used during SWOT analysis for the formulation of a business strategy. The **I**ntegrated **DEF**inition (IDEF0) model for the said analysis is presented in Figure 3.3. In this process the method, business functions and sub-functions have been represented with different actors involved. The interactions between actors and functions are presented using arrows whereas interaction between functions and sub-functions are shown by <<uses>> and <<extends>> arrows. The <<uses>> type arrow defines the relationship that the function always uses the sub-function prior to its interaction with the actors defined in the IDEF0 model whereas the <<extends>> type refers to potentially extending the functionality of the source function/sub-function to the target function/sub-function.

The **P**rocess **I**ntegration (PI), technology R&D, device engineering, **C**entral **C**AD **D**esign **S**olution (CCDS), DFM, Design, EDA and interconnect modeling teams are directly involved in the process. It can be seen that the need for FE technology development/improvement could arise as per company policy to move towards technology shrinks (55nm, 50nm, 40nm) or customer feedback for technology improvements. The technology shrink a.k.a. technology derivative is a process where new technology is derived from the existing technology and is capable of manufacturing smaller features (40nm is derived from 45nm, 55nm and 50nm are derived from 65nm). These derived technologies are based on design, process and/or material innovations and result in the gain of area and costs. On the other hand, the technology improvement refers to the efforts fully dependent on process innovations that enable to manufacture same design in smaller area. These improvements enable to put features more closely but feature sizes are not reduced.

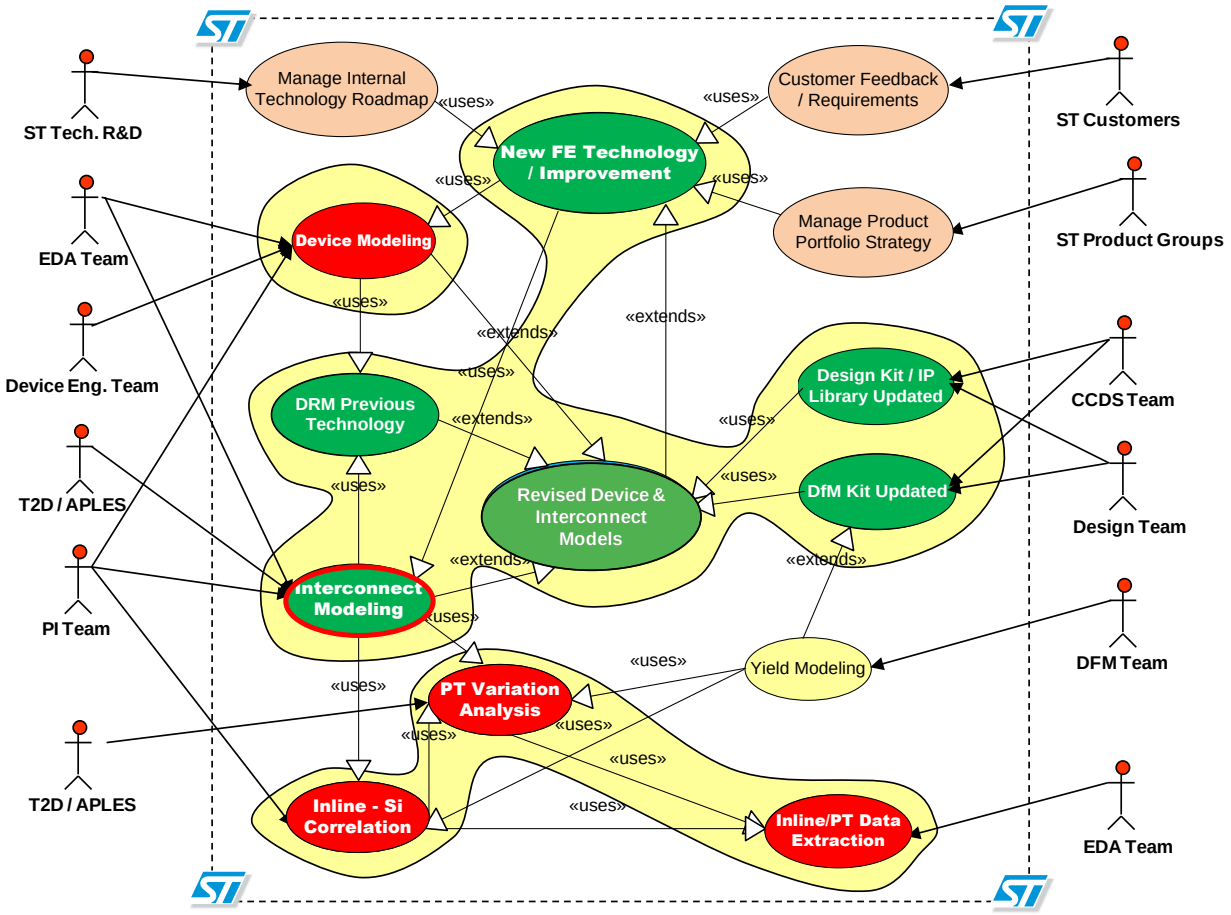


Figure 3.3 – Technology derivative/improvement process

The device and interconnect modeling are the key processes in the technology derivative/ improvement initiatives where SPICE models (device and interconnect models) are developed through CAD simulations and validated based on the measured electrical (PT) results. This process starts with the definition of target values on critical and key parameters. The physical parametric stack is constructed based on the source technology design rule manual (DRM) and is simulated until the electrical and parasitic extraction satisfies the SPICE model corners. This stack is revised to validate until the electrical and parasitic parameters comply with SPICE models. The geometric stack is identified through CAD simulations which are validated through test products.

The focus here is on interconnect modeling where key sub-processes are listed as *(i)* data extraction, *(ii)* drift analysis and *(iii)* silicon based correlation. The test products undergo the manufacturing process flow and the data is collected across the production line. The first step in silicon based validation of the interconnect models is the extraction of the data from multiple data sources for analysis purposes. These measured parameters are benchmarked against the simulated results and significant deviations are highlighted (single-source analysis). Any significant deviations are further investigated for the root cause analysis (multi-source analysis) to find the source cause followed by model corrections. In this process, success lies in the ability to quickly resolve the conflicts that arise from the stack validated vs. simulated results that shall reduce the respins and technology lead times. These results are used to revise the device and interconnect models in the DRM.

The DFM teams are focused on identifying the manufacturability and yield loss mechanisms where the drifts/variations leading to parametric or functional yield loss are further investigated for their classification as systematic or random mechanisms. These drifts and variations are transformed into rules and/or models for inclusion in the DFM kits and subsequent use during CAD simulations. The design and DFM kits are used by the designers in addition to the design libraries to quickly design, simulate and assess its manufacturability and yield as a first time correct design. The similar process is required in the device modeling; hence, they result in a chain reaction of improvements ultimately leading to a new technology derivative/improvement effort (alignment or adoption) with reduced lead times and costs.

3.5 KEY CHALLENGES IN TECHNOLOGY DERIVATIVE/IMPROVEMENT PROCESS

The IDEF0 model is followed in case of *(i)* a new technology transfer from an alliance (alignment), *(ii)* local technology derivative and improvement initiatives (alignment), and *(iii)* during the product design and development (adoption). The success lies in the ability to capitalize the manufacturing knowledge by dynamically exploiting the data collected across the production line to quickly identify the significant drifts/variations followed by its root causes. Let us answer the two questions proposed in section 3.3. The answer to our first question is quite evident that in an IDM-fablite business model we have methods and processes for quick technology alignment and adoption but answer to the second question is not trivial because at present lead times associated with technology alignment and adoption are increasing which is the biggest challenge towards ramp-up rate. So to find an answer to this question, a common project was conducted with the T2D group to better understand the limitations faced by the R&D engineers and to assess that why they are not able to quickly validate the SPICE models and find root causes against the drifts and variations. In this project a back-end-of-line variance analysis tool (BEOL-tool)[9] was developed and deployed to analyze the limitations. The results are presented on sub-process basis in the following sub-sections:

3.5.1 Data Extraction Issues

It was found that there are 6+ data sources and 11+ engineering data analysis tools being used to support this technology derivative/improvement process (Figure 3.3). The following very interesting observations were also made:

[9] The complete UML model, GUI, Algorithms and experimental results can be found in Appendix-E

a) Multiple manufacturing data sources (relational databases) are dedicated to operational excellence and do not support DFM/MFD efforts; hence, engineers spend a lot of time in extraction, cleaning and alignment before analysis and in most cases, it results in zero value addition.

b) Manufacturing data resources have serious ontology issues (same parameter with different semantics in different databases), as a consequence it becomes difficult to align and correlate multi-source data resulting in a missed opportunity.

c) Unstructured evolution of local databases has resulted missing links which are the key to perform a multivariate or predictive modeling.

It is evident from these observations that multi-source data extraction has serious issues which restrict engineers to find root causes using single source analysis resulting in extended lead times and in most cases no improvements.

3.5.2 Variance Analysis Challenges

The variance analysis is the first step towards root cause analysis and provides initial signals about potential drifts or variations. It is based on the electrical parametric data collected across the production line. The key observations made during this project are listed as under:

a) R&D engineers may apply ± σ filters during single-source data extraction to remove outliers that could possibly highlight a significant drift. The resulting data when checked against its compliance with the normality law, is often found in compliance.

b) Excel is the widely used tool in SI besides multiple advanced statistical tools but engineers prefer excel because they could easily handle and manipulate data in excel that might result in misleading conclusions

c) The data collected across the production line is huge and the tools being used for the data pre-processing are too slow to handle it because of their relational database limitations.

It is evident from these observations that when different analysis tools are used, including excel, they are like black boxes in terms of the algorithm used for the computation of statistics. It was found that the same data when used for certain statistics computation with different tools, often result in slight variation that might be misleading. If the data is filtered then it is not possible to improve or go for a technology derivative.

3.5.3 Silicon Based Correlation Limitations

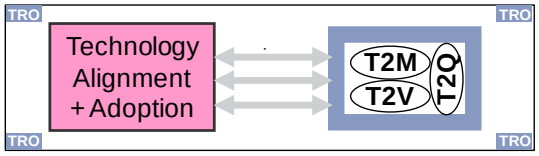
This sub-process is not yet implemented because R&D engineers are not able to align the extracted data at site/die and test structure position levels. It means that it is not possible to perform multi-source analysis to find root causes against drifts/variations. It can simply be concluded that the knowledge capitalization is at its minimum in this sub-process. Now there exists sufficient information about the strengths and weaknesses of an IDM-fablite business model to move towards SWOT analysis so that an appropriate business strategy can be agreed prior to propose modification in the business model to ensure its compliance with a recent shift in the business objective i.e. ramp-up rate.

3.6 SWOT Analysis on IDM-fablite Business Model

The SWOT analysis is focused on the top ranked objective <<to achieve leadership position in SI and quick ramp-up rate>> and the 2*2 box diagram is presented in Figure 3.4. In addition to this objective, ramp-up rate, technology alignment and adoption have also been taken up as key objectives to achieve leadership position.

Questionnaire and brainstorming were used with technology R&D, device engineering and process integration teams in this regard. It has resulted in the proposition of four strategies as under:

***a)* Strength/Opportunity Option:** This option suggests joint ventures with top ranked IDMs, foundries and fabless companies to best exploit the strengths e.g. intellectual capital, state of the art equipment, data and methods against potential opportunities (high revenues and market share).



SWOT Analysis *for* TRO
"Achieve leadership position in SI"

		INTERNAL	
		Strength	**Weaknesses**
		• *Intellectual Capital in R&D* • *State of the art equipment* • *Quality standards & procedures* • *Huge manufacturing Data* • *Product/test chip characterization, FDC, APC, Device and Interconnect modeling methodologies* • *11+ EDA tools & 6+ databases*	• *Ontology Issues* • *Missing links between Databases* • *Excel (widely used) + Data Filtering* • *Interconnect & Device modelling based on previous DRM* • *Min knowledge capitalization*
EXTERNAL	**Opportunities** **Environmental Factors** • *Huge demand increase is expected* • *Higher revenues with high value products* • *TCAD Alliances to reduce R&D Costs* **Competitive Factors** • *Capitalize methodological knowledge* • *Stretch CMOS to the limits* • *Strength of Euro in comparison with Dollar & other currencies*	**SO Options:** **"Establish joint venture with TOP IDM's to reduce R&D costs and capitalize knowledge for higher revenues and profits."**	**WO Options:** **"Capitalize local production knowledge & improve coordination b/w R&D functions to understand information needs"**
	Threats **Environmental Factors** • *Dynamic customer requirements* • *Physics laws are limiting to answer variability* • *Fabless business model* **Competitive Factors** • *Technology Platform development & competitiveness* • *Backward compatibility with common technology platform* • *ITRS Pressure*	**ST Options:** **"Innovation in material, design & process along with efficiency and effectiveness at methodologies"**	**WT Options:** **"Remove ontology issues, missing DB links & base device & interconnect modelling on Si Data"** **<< DFM is a solution >>**

Figure 3.4 – SWOT analysis results

***b)* Strength/Threat Option:** This option suggest focus on the design, process, equipment and material innovation to mitigate threats like limiting physics laws, technology platform development and backward compatibility and dynamic customer requirements.

***c)* Weakness/Opportunity Option:** The ontology issues, missing database links, usage of excel for data analysis, minimum knowledge capitalization (correlation between geometric and electrical measurements) are characterized as the core weaknesses; hence, in order to exploit the opportunities there must be a focus on the knowledge capitalization and improved coordination between R&D functions.

***d)* Weakness/Threat Option:** This option suggests mitigating the threats by eliminating weaknesses: ontology issues, establishing missing links between database and tuning EIS by transforming relational data sources to multidimensional data structures coherent for advanced statistical analysis.

The strategies a & b are already deployed at the STMicroelectronics where joint ventures with the top ranked IDMs' to share R&D costs and intellectual capital are established. The significant number of PhD students in strong collaboration with LABS and industrial partners are inducted to carry out latest research. The main objective is to best utilize the strengths and equal opportunity SI behavior to exploit potential opportunities emerging from cumulative growth in demand that cannot be achieved without removing the weaknesses. It has been discussed above that success lies in the ability to quickly exploit huge manufacturing data for the root cause analysis; hence added value come from the weakness mitigation. It shall strengthen opportunity window by minimizing threats; hence strategies c & d are adopted. The ontology issues and missing links need to be removed and silicon (results from the wafer measurement) knowledge capitalization needs to be increased.

3.7 PROPOSED EXTENDED IDM (E-IDM) BUSINESS MODEL

It can be concluded that multi-source data extraction, alignment and mapping is critical to improve the effectiveness of DFM and MFD efforts during new technology transfer (alignment), technology derivative and improvement efforts (alignment) and product design and development (adoption). Based on these facts and business strategies (c & d) it can also be concluded that existing IDM-fablite mode is not coherent with the recent shift in the business objective (ramp-up rate). So an extended IDM-fablite business model has been proposed which provides true coherence with recent shift in business objectives (ramp-up rate) and top ranked SI objective (Figure 3.5).

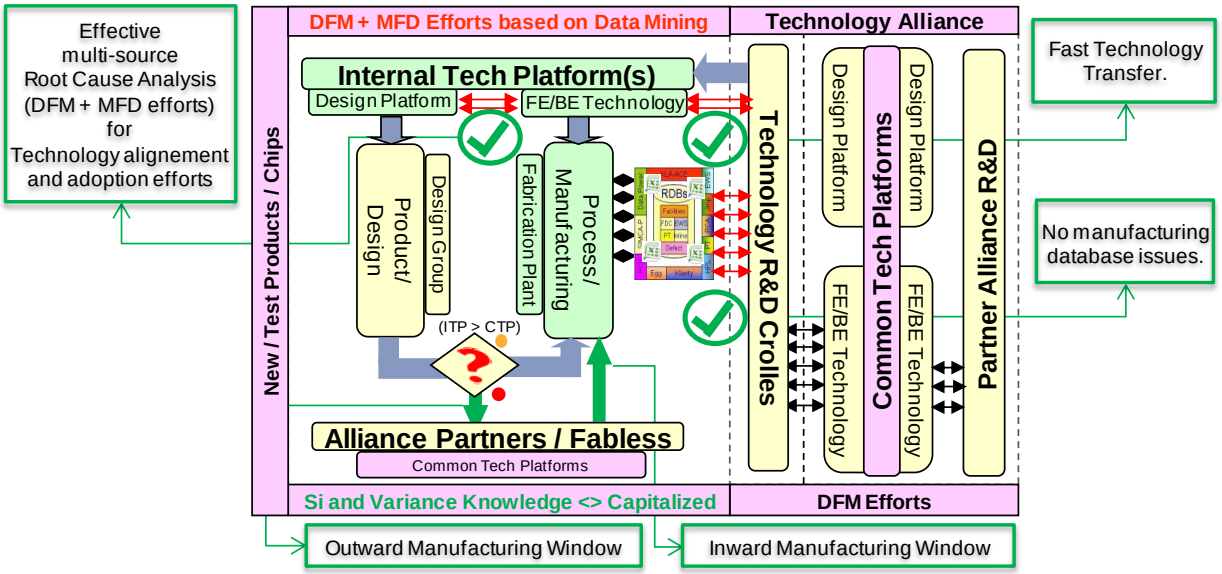


Figure 3.5 – Proposed extended IDM-fablite business (e-IDM) model

The DFM efforts are normally applied to find manufacturability and yield loss mechanisms. It is computationally and financially expensive in terms of measurement data and experimentation; hence, they are restricted to be used during the new technology development efforts in technology alliance. In the e-IDM model the integration of DFM and MFD efforts have been proposed because both are based on the R&D efforts (Figure 1.11) and need multi-source data for an effective root causes analysis. This integration gives rise to the need for efficient data extraction, mapping and alignment which means that engineering information systems must be tuned for effective data exploitation.

On the basis of business strategies c & d and the key challenges identified in technology derivative/improvement process, scientific contributions *(i)* MAM (mapping and alignment model), *(ii)* SPM (spatial positioning model) and the *(iii)* ROMMII (referential ontology Metal model for information integration) have been proposed in chapters 5 and 6. The proposed contributions enables R&D engineers to perform multi-source root cause analysis at site/die and the test structure position based analysis. It shall result in the fast technology transfer from an alliance followed by continuous technology derivative/improvement initiatives for alignment purposes. The product development lead times (technology adoption) shall also significantly reduce along with an opportunity to improve the technology, based on any indentified yield loss mechanism.

The most remarkable thing about this proposed model is that it is not structurally very different than traditional IDM-fablite model but it offers a shift from single-source data driven R&D efforts to multi-source information and knowledge driven initiative. It enables swift silicon based knowledge capitalization. In the previous IDM-fablite model the inward manufacturing window has never existed but the proposed model ensures the inward manufacturing window where one can get orders from alliance partners design centers based on the fact that internal technology is robust in yield and reliability than the common technology platform. This inward manufacturing window shall serve the organization in downtimes as well. A very good

example of inward manufacturing window is recently witnessed when the revenues of STMicroelectronics dropped by 6% whereas the worldwide sales revenues have increased. The reason was that Nokia fired 5000 employs and it had a significant impact on the sales orders for STMicroelectronics.

3.8 RESEARCH SCHEMATIC AND ADVANCEMENTS (E-IDM MODEL)

The research schematic as presented in Figure 3.6, is used to demonstrate the logical flow and sequence of activities performed that resulted in the proposed scientific contributions. This chapter has addressed the blocks A0 to A4. The blocks A0 and A1 have resulted in the identification of TRO as <<to achieve the leadership position in SI and quick ramp-up rate>> which is not possible without introducing new technologies every 2-3 years, quick ramp-up rate and technology alignment and adoption efforts. Next, A2 was followed where three key improvement areas were identified as *(i)* fast technology transfer, *(ii)* database issues and *(iii)* technology alignment and adoption. Further block A3 was considered because it is necessary to identify the strengths and weaknesses to be used within SWOT analysis. The weaknesses identified in technology derivative/improvement are *(i)* data extraction, *(ii)* mapping and *(iii)* alignment. It restricts engineers to single-source root cause analysis which is not sufficient in the context of newly emerging spatial drifts/variations. The SWOT analysis is performed based on the input gathered from the blocks A0, A1, A2 and A3. Three strategies have been identified, however strategies adopted focus the removal of weaknesses to mitigate threats and exploit the opportunities associated with the cumulative demand growth. An extended IDM-fablite business model has been proposed which removes these weaknesses from the technology alignment and adoption efforts.

This proposed e-IDM model shall serve as the basis for whole thesis. The next chapters shall identify the key root causes associated with the challenges identified in block A3 in the context of three identified key improvement areas. Based on the root causes identified, further generic scientific contributions MAM, SPM and ROMMII frameworks have been proposed to enable this e-IDM business model.

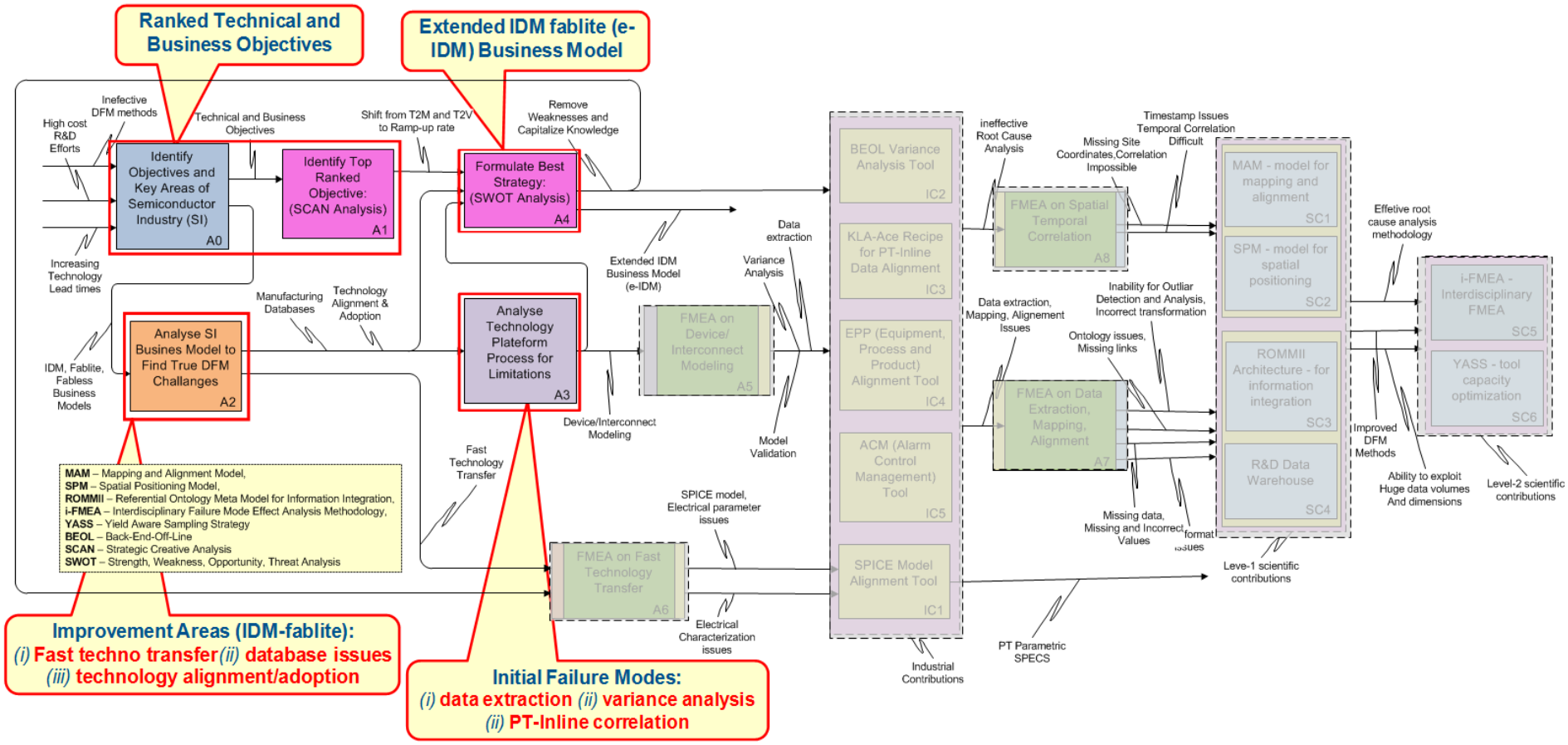


Figure 3.6 - The research schematic and advancement with e-IDM business model

3.9 SUMMARY AND CONCLUSIONS

This chapter has analyzed the most successful IDM-fablite business model from the SI to assess its support towards a recent shift in the business objectives "to achieve leadership position and quick ramp-up rate". The first step was the identification of business (T2M, T2V, ramp-up-rate) and technical (area, power, timing, leakage) objectives using brainstorming sessions held with engineers and managers, followed by their subsequent ranking using SCAN analysis. The top ranked business objective "to achieve leadership position and quick ramp-up rate" and key improvement areas as the *(i)* fast technology transfer, *(ii)* manufacturing databases and *(iii)* effective root cause analysis (R&D) from IDM-fablite model, are further investigated using SWOT analysis. The objective is to align business strategy to assess its potential compliance with IDM-fablite business model against identified TROs. The technology development/improvement process was further investigated to assess potential challenges faced during its compliance with the TROs and *(i)* data extraction, *(ii)* alignment and *(iii)* pre-processing due to ontology issues and *(iv)* missing database links are found as key failure modes.

The conclusions highlighted that the existing IDM-fablite model do not fully support the TRO objective. It is because of the fact that our R&D engineers are not able to exploit multi-source data collected across production line for the root causes analysis which has led ineffectiveness in the DFM methods. The DFM methods play a critical role in technology development, alignment and/or adoption. It is important to improve technology alignment and adoption lead times and associated costs; hence an extended IDM-fablite (*e*-IDM) business model is proposed with the integrated DFM efforts. The proposed model provides compliance with TROs and enhances inward manufacturing window while maintaining its backward compatibility with the technology developed in alliance.

In next chapter we shall find out true DFM challenges within the proposed *e*-IDM model so that associated failure modes and root causes against DFM integration are removed. The objective is to ensure the success of proposed *e*-IDM model so that recent shift in business objectives quick-ramp-up rate is achieved for potential economic benefits and leadership position.

Chapter 4: *I*-FMEA Methodology for True DFM Challenges[10]

In the previous chapter, an extended IDM-fablite (*e*-IDM) model is presented with DFM integration to continuously improve the technology derivative/improvement alignment and adoption. The propose *i*-FMEA methodology helps in identifying and removing true DFM challenges, which are critical for the success of *e*-IDM model. The *i*-FMEA is different than traditional FMEA as it searches failure modes and root causes across business functions. It is applied on two groups of failure modes as *(i)* ineffective root cause analysis (infield and scribe line test structure positions) and *(ii)* data extraction, mapping and alignment. It has been seen that most of the cyclic root causes (repeating causes) are traced back to IT and EDA business functions. The identified root causes form the basis for generic scientific contributions as *(i)* MAM, *(ii)* SPM, *(iii)* ROMMII and *(iv)* YASS. The experience learned during the identification of true DFM challenges has led the proposition of 4-step *i*-FMEA methodology which is capable of finding cyclic failure modes and root causes, which require generic solutions rather than operational fixes.

Contents

[10] Shahzad M.K., Hubac S., Siadat A. and Tollenaere M., An interdisciplinary FMEA methodology to find true DFM challenges, 12th European APCM Conference, Grenoble France 2012

4.1 INTRODUCTION

The previous chapter has proposed an extended IDM-fablite (e-IDM) business model to address the recent shift in business objectives "ramp-up rate" in compliance with top ranked business objective "to achieve the leadership position". The success of the proposed model depends on the ability to remove DFM ineffectiveness across three key improvement areas: *(i)* fast technology transfer, *(ii)* ineffective root cause analysis and *(iii)* extraction, mapping and alignment to improve technology alignment and adoption efforts. In this chapter, the objective is to identify true failure modes and root causes for the DFM ineffectiveness for generic solutions rather than operational fixes. The proposed *i*-FMEA methodology is based on well known approach used for the concept, design and process known as **F**ailure **M**ode **E**ffect **A**nalysis (FMEA). This approach is limited by the expert's knowledge and scope because failure modes and root causes are searched at system/subsystem/ component levels. The root causes associated with the failure modes are removed with operational fixes, which are not permanent solution. The proposed 4-step interdisciplinary FMEA (*i*-FMEA) methodology is superior to the traditional FMEA, because it searches for cyclic failure modes and root causes across all the business functions. It is objectively focused on the continuous scanning of business environment for any potential change followed by the identification of key challenges and failure modes with cyclic causes. The results are very promising as the key failure modes and root causes identified were never thought to be the source of the problem.

4.2 HISTORICAL EVOLUTION OF FMEA METHODOLOGY

The FMEA approach was initially proposed by US military in the standard MIL-P-1629, about 60 years ago, with an objective to identify different failure modes of system components, evaluate their effect on the system and propose proper counter-measures. Besides the military based applications, it was first adopted by the aerospace industry and was named as the failure effect analysis technique (MIL-18372). The earliest formal description as the failure mode effect analysis (FMEA) was given by New York Academy of Sciences [Coutinho, 1964]. The FMEA approach was further extended as the **F**ailure **M**ode **E**ffect and **C**riticality **A**nalysis (FMECA) by NASA to ensure the desired reliability of the space systems [Jordan, 1972]. A lot of diversifications in the traditional FMEA methodology have been seen [Reifer, 1979] e.g. **S**oftware **F**ailure **M**ode and **E**ffects **A**nalysis (SWFMEA), industry-wide FMEA standard J-1739 issued by the **S**ociety of **A**utomotive **E**ngineers (SAE), etc.

4.2.1 FMEA Process and Evolution

The traditional 5-step FMEA process is presented in Figure 4.1. It starts with the clear description of the scope e.g. design, product or process, followed by the identification of important functions for further analysis (step-1). The most commonly used methods for this step are as *(i)* SIPOC (**S**upplier, **I**nput, **P**rocess, **O**utput and **C**ustomers) and process maps for the scope definition, and *(ii)* QFD (**Q**uality **F**unction **D**eployment) and C&E (**C**ost-**E**ffect) matrix for functional analysis. A cross functional team is established based on the scope definition and functional analysis results. The potential failure modes, effects and root causes are identified and listed along with occurrences, severity and detection (step-2). The brainstorming method is the most widely used technique at this step along with a likert scale from 1-10 to quantify potential risk against occurrences, severity and detection. The **R**isk **P**riority **N**umber (RPN) is computed by multiplying the severity, occurrence and detection (step-3). A threshold RPN value is used as criteria to initiate operational fixes against a particular failure. The failure modes with RPN value above than the threshold value are focused and operational fixes are applied to remove the respective root causes (step-4). Finally the operational fixes effectiveness is reviewed and cycle is repeated until the RPN number falls below the threshold value.

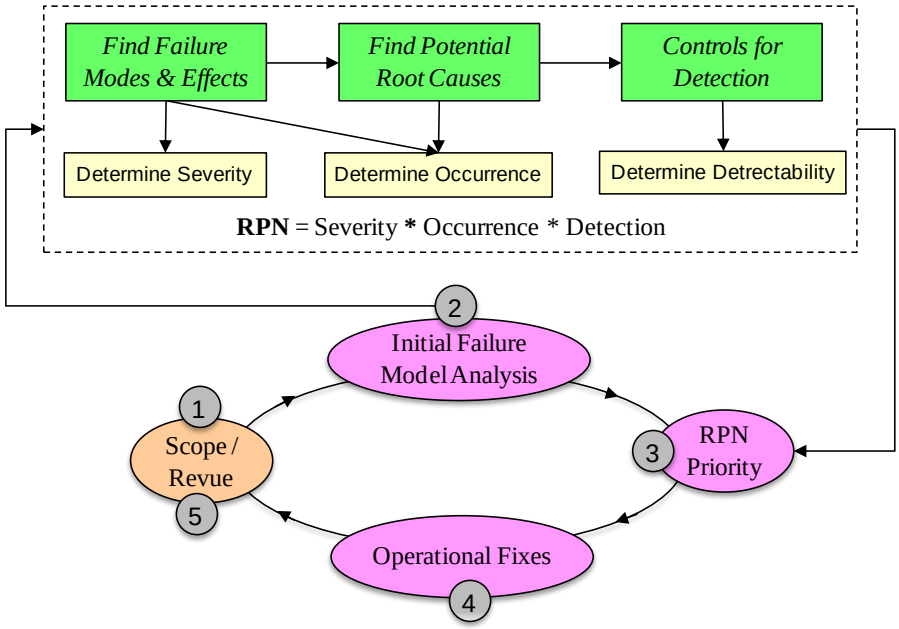


Figure 4.1 – Traditional 5-step FMEA process

The FMEA approach has received an industry wide acceptance and applicability; hence, it is very common to find diversified standards in each domain. The most common forms of FMEA are: *(i)* DFMEA: FMEA applied in the product design a.k.a. design FMEA, *(ii)* PFMEA: FMEA applied in the process a.k.a. process FMEA, *(iii)* S/C FMEA: FMEA applied in a system or a concept a.k.a. system or concept FMEA, *(iv)* SWFMEA: FMEA applied in software development or information system a.k.a. software FMEA, etc. [Ireson et al., 1996]. It is not in the scope of this thesis to discuss and compare FMEA evolutions across different domains; hence, this discussion restricts itself to the traditional FMEA approach to benchmark it against the proposed *i*-FMEA methodology.

4.2.2 Basic FMEA Vocabulary

Let us discuss some basic vocabulary [Dhillon, 1999] to better understand the traditional FMEA approach and its comparison with our proposed *i*-FMEA methodology as presented in section 4.3.

Failure mode: It is defined as the ways or modes in which something might fail and could be any potential or actual errors or defects.

Criteria: It is the objective, one wants to achieve and for which a non compliance results in a failure.

Failure Effect: It is simply referred as the consequences of the failures.

Failure Cause: These are the reasons for the failures that start the drift and end up in a failure.

Severity: It is the failure mode's consequence based on the worst case scenario.

Occurrence: It is the frequency of the occurrence of a failure mode or a cause.

Detection: It refers to the existing controls ability to detect potential failures.

4.2.3 Benefits and Limitation of Traditional FMEA Approach

The FMEA approach is focused on identifying potential risk that might lead to the product, process and/or a system failure. It is a continuous improvement process; however, success lies in the ability to accurately identify the scope and functional analysis. It do not provide a thorough and systematic analysis and it is limited by the scope and selected functions based on the expert's knowledge. The estimation of frequencies and severity from the users input and identification of operational fixes are not trivial. The biggest advantage of FMEA approach is that it can be applied in any context to assess potential risks and apply operational fixes to avoid failures. The most important FMEA applications are: *(i)* when product, process or service is

being designed or redesigned, *(ii)* when an existing product, process or service is being applied in different way, *(iii)* when developing control plans for new or modified process, *(iv)* when improving the existing products, processes or services, *(v)* when analyzing failures of an existing process and *(vi)* periodically throughout life of process, product or service.

4.3 PROPOSED INTERDISCIPLINARY FMEA (I-FMEA) METHODOLOGY

The *i*-FMEA approach is proposed to overcome the limitations of traditional FMEA approach as discussed in section 4.2.3. The proposed methodology is presented in Figure 4.2 and is based on the existing FMEA approach.

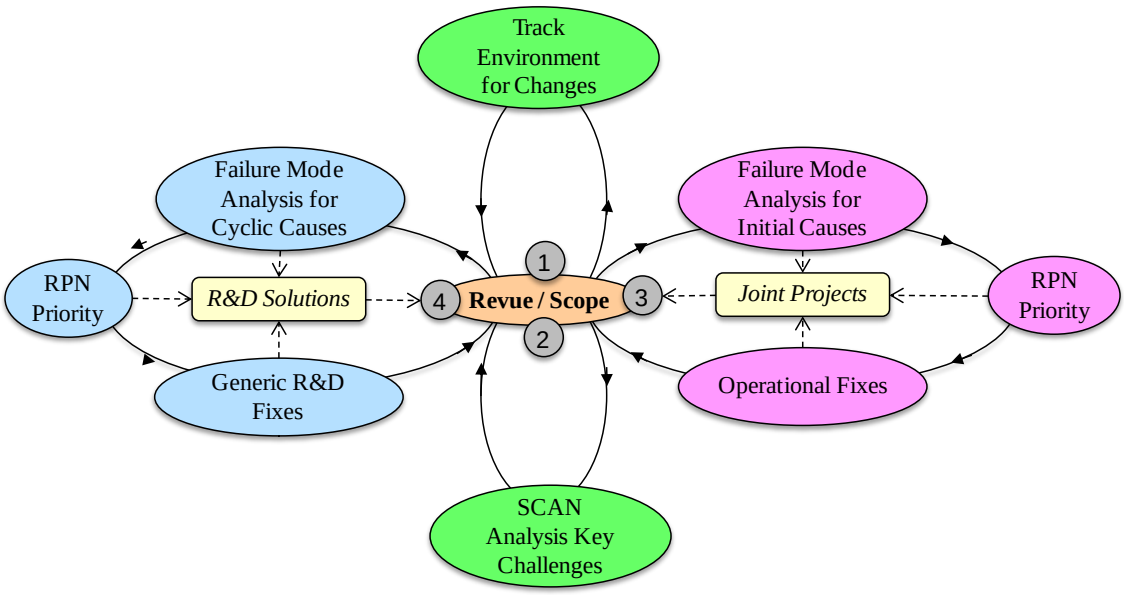


Figure 4.2 – 4-step *i*-FMEA methodology

It is a 4-step methodology and is based on the continuous scanning of business environment for any potential change (step-1). Upon the detection of a potential change, the SCAN analysis (step-2) is performed to find potential areas with key challenges. These areas along with challenges are further reviewed prior to move towards step-3. This step in the *i*-FMEA methodology is exactly same as the traditional FMEA approach, where we find potential failure modes, root causes and their detectability along with the estimation of severity, occurrence and detection on 1-10 likert scale. The potential risk is computed as RPN and operational fixes are applied. In *i*-FMEA methodology, operational fixes are developed in joint projects with respective end users based on expert's knowledge. These fixes are continuously monitored and reviewed because they often lead to the failure modes with cyclic causes which are input for the step-4. If no cyclic causes exist then one can loop back to steps 1 and 2. The cyclic failure causes are the converged failure mechanisms that emerge from the initial failure modes through operational fixes. These cyclic causes require in depth investigation similar to step-2 followed by its RPN value based prioritization. The generic R&D solutions are the only permanent option against cyclic failure modes and causes.

4.3.1 Comparison of i-FMEA with Traditional FMEA Approach

The comparison between traditional FMEA and proposed *i*-FMEA approaches is presented below:

a) The traditional FMEA is a 5-step approach used to identify potential failure modes based on the criticality assessment whereas *i*-FMEA is a 4-step methodology which is based on the traditional FMEA approach but it is focused on finding the cyclic failure modes and causes that have the potential to become causes for other failure modes.

b) The success of traditional FMEA depends on our ability to precisely define the scope and functional analysis before applying the basic steps. It requires expert's knowledge and could easily lead to the waste of time and resources; however, *i*-FMEA methodology continuously analyze the business environment for any potential change followed by SCAN analysis to find out critical processes that might get effected. This methodology is a better approach as it emphasizes to align all business functions to quickly respond to any potential change in the business environment.

c) The traditional FMEA approach is subjectively focused on applying operational fixes to improve the design, process or system imperfection but *i*-FMEA methodology is focused on identifying failure modes with cyclic causes followed by generic R&D fixes so that any chances for associated causes to become causes for other failure modes can be eliminated.

The above discussed potential differences clearly defines the advantages of our proposed *i*-FMEA methodology over traditional FMEA approach, which is primarily/subjectively focused on aligning the business processes to quickly respond to business environmental changes for maximum market share and higher profits.

4.3.2 *i*-FMEA Methodology and Thesis Schematic

The proposed *i*-FMEA methodology can be easily identified from the thesis schematic as presented in Figure 4.3 below. The same schematic is used as presented in chapter-1 to highlight the activities grouped into 4-steps of *i*-FMEA methodology. The business environment scanning in this thesis highlighted a shift in the SI business objective from T2M and T2V towards ramp-up rate, increasing R&D costs, ineffective DFM methods and exponentially increasing technology lead times. In the step-2 the business model has been analyzed using the SCAN analysis to find out the key challenges and this step resulted in the technology development process as the key challenging area (A0, A1, A2 and A4). In the step-3 the initially identified failure modes (A3) are found as *(i)* data extraction, *(ii)* data mapping and *(iii)* data alignment. The operational fixes (IC1, IC2, IC3, IC4 and IC5) are provided for the root causes identified against these failure modes in the joint projects.

The operational fixes in terms of industrial tools further resulted in three failure modes as: *(i)* fast technology transfer issues, *(ii)* ineffective spatial correlation and *(iii)* mapping and alignment issues (step-4). The cyclic root causes identified against these failure modes emerged across all the teams using developed operational fixes. The generic solutions are then proposed as *(i)* MAM model, *(ii)* SPM Model, *(iii)* ROMMII platform and *(iv)* R&D data model. These generic scientific contributions (SC1, SC2, SC3 and SC4) provide us means to remove weaknesses from our existing processes so that they are aligned to quickly respond to the changes in business environment.

4.4 *I*-FMEA METHODOLOGY RESULTS

The step-1 results have been presented in chapter-3. The industrial contributions (operational fixes) are presented in appendices D to H whereas generic scientific contributions *(i)* MAM and *(ii)* SPM models are presented in chapter-5 and *(iii)* ROMMII platform and *(iv)* R&D data model are presented in chapter-6. In this section the results of failure mode analysis with initial and cyclic causes are presented.

4.4.1 Step-2: Initial Failure Modes and Root Causes

The step-1 of the proposed *i*-FMEA methodology has resulted technology derivative/improvement process as the key area which is impacted by a recent shifts in the business environment. In this process, the technology derivative and improvement initiatives and fast technology transfer have been identified as the core business functions. These sub-processes shall be investigated with the core functions using traditional FMEA approach to find initial failure modes and subsequent causes for the operational fixes.

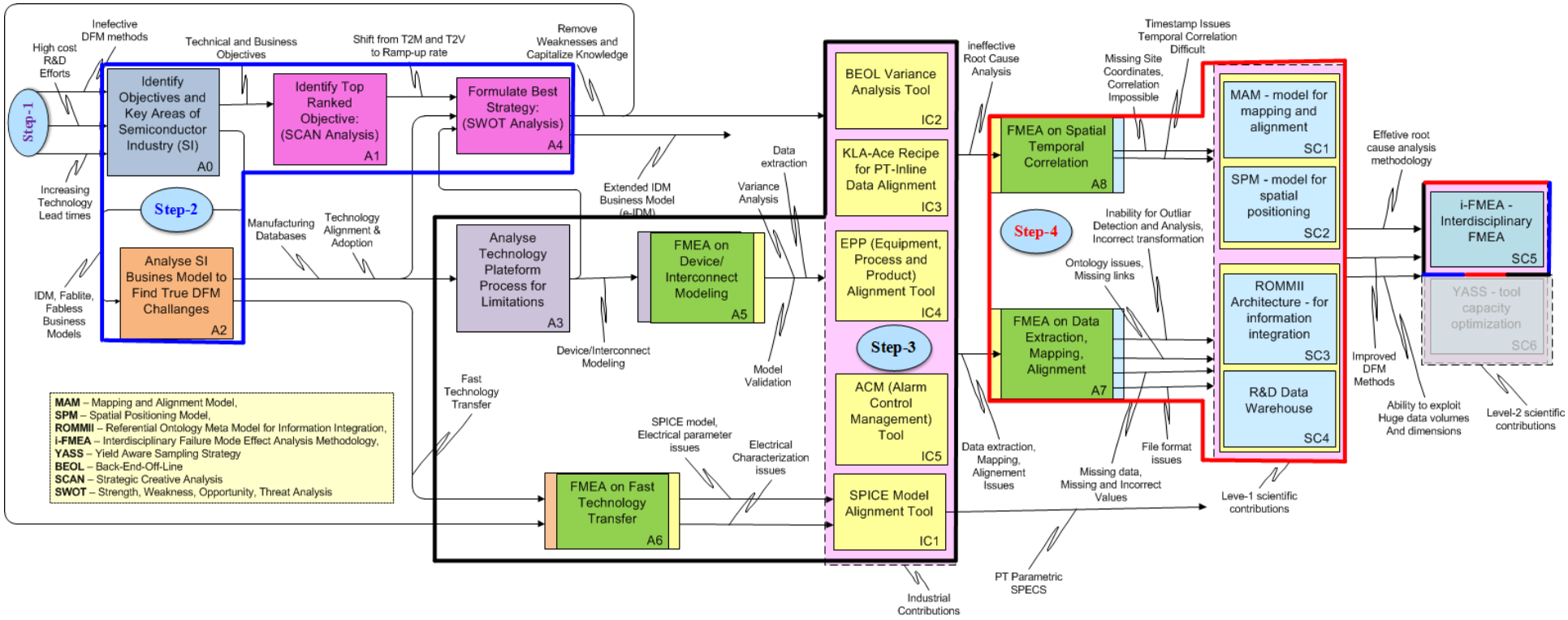


Figure 4.3 – 4-Step *i*-FMEA methodology and research thesis schematic

4.4.1.1 Technology Derivative/Improvement Initiative

This process comprises of four sub-processes: *(i)* device/interconnect modeling, *(ii)* data extraction and pre-processing, *(iii)* variance analysis and *(iv)* PT-Inline correlation. The traditional FMEA results are presented in tables 4.1 and 4.2. Let us discuss the FMEA results of device/interconnect modeling (Table 4.1) where sub-processes are benchmarked as *(i)* to generate typical stack, *(ii)* to generate stack for corner models, *(iii)* QA and impact analysis, *(iv)* design impact analysis, *(v)* variance analysis and *(vi)* model validation. The cutoff value for RPN based prioritization is taken as 200. It is known that *(i)* variance analysis and *(ii)* model validation are the critical sub-processes; hence, the discussion will be restricted to these two critical sub-processes. The simulated target and corner models against typical and corner stack definitions are validated where significant deviations are further investigated to find root causes so that the process, equipment or SPICE models can be adjusted. The variance analysis and model validation steps are then the most important tasks that have direct impact on the technology lead times and associate costs.

The most important and critical root causes identified are incorrect tool, incorrect statistics, incorrect measurements, wrong data pre-processing and site level data alignment for the electrical and physical geometric measurements. The fact that the extracted data is not understandable by the R&D engineers makes it almost impossible to analyze the variances and/or model validations. If the preventions and proposed actions are closely analyzed, it can be easily concluded that the biggest challenge is data extraction in right format.

The critical causes highlighted in device/interconnect modeling are as *(i)* data extraction, *(ii)* alignment and *(iii)* pre-processing, which are further investigated as presented in the table 4.2 to ascertain the proposed corrections. It can be concluded that these sub-processes are very critical for device/interconnect model variance analysis in technology derivative/improvement initiatives based on the fact that computed RPN values are above the cutoff value (200). These critical elements are further discussed as under:

a) Data Extraction Sub-Process

The SI production line data is stored in 4-layer data storage architecture: *(i)* flat files, *(ii)* operational data sources or stores (ODS), *(iii)* data warehouses (DWH) and *(iv)* data marts (DM). The higher storage layers (DWH and DM) represent aggregation and integration of the data from lower layers (flat files and ODS). The industry normally uses multiple data extraction tools which are classified as single-source or multi-source data extraction and analysis utilities. More often it is found that either data is not available or users are not able to properly extract it using these utilities. In both cases neither variance analysis nor model validation could be performed, which result in high severity. The most apparent root causes identified against these limitations are: *(i)* the databases have different data retention periods and *(ii)* users are not able to use multiple tools due to training issues. The users use multiple tools for data extraction from different data sources that result in data incoherencies and pose severe data alignment issues because ODS have longer data retentions than DWH and DM. It is often seen that, multi-source data extracted by end users has varying formats and is not understandable due to different vocabulary used in different databases. The databases are often not supported with up-to-date data dictionaries to present semantics of data and dimensions. All of these limitations result in data alignment; hence, the single-source variance analysis is likely possible but multi-source analysis due to data alignment issues is almost impossible. The proposed solutions could be to either change DWH/DM databases or develop automated data extraction tools.

Sr.	Block	Sub-Block	Function	Criteria	Failure Mode	Effect	SEV	Cause	Occurrence	OCC	Prevention	Detection	DET	RPN	Recommended Actions	Rensposible Person and Target Date
1	Technology Derivative / Improvement Initiatives	Device / Interconnect Modelling	**Generate Inline Stack for Typical Model**	First time correct inline stack for all metal layers with geometries and materials that give us a typical (target) SPICE model specifications	incorrect stack defination	deviated model target	8	Selection of the previous DRM stack which does not correspond to - the new product performance requirements, new materials, new pitch and densities	It alwas occurs whenever device or interconnect R&D engineer starts model target and corner validaation steps during technology alignment as well as adoption efforts	8	Supporting R&D engineers with mutli-source data extraction so that we can learn to predict the new stack in addition to previous stack definitions from DRM	It is always the case in the present process.	2	128	Capitalize the manufacturing knowledge so that engineers can extrapolate or use DOE (design of exxperiments) technique to define initial stack	IT, EDA and Training Center, as soon as possible
								Incorrect Simulation Tool	It is a rare activity, as prior to its use, we test a tool in depth for its accuracy and correctness	1	Benchmark simulation tools and flow	Detection is done by R&D engineers only in case of significant or abnormal parametric drifts	4	32	Benchmark all charecterization tools before usage in the flow	Device and interconnect modeling R&D teams, as soon as possible
			Construct Model Corners	Stack definition for model corner definition (slow and fast) that defines the spread and ultimately the area of the chip		deviated corners		Incorrect SPICE Model	It is a rare activity	1	Automate the SPICE model inclusions		4	32	Automate the SPICE model inclusion	IT, EDA Sections, as soon as possible
						respins		wrong test simulation conditions	It could happen quite often, the wrong simulation condition might result in different target and corners	4	Generate test cases or automate SPICE simulations		4	128	Automate the process	IT, EDA Sections, as soon as possible
2			**QA and Impact Analysis**	Assess the manufacturability by simulating eletcrical parameters of a test design against the model target and corners specifications	misleading manufacturing assessments	It shall result in significant manufacturability issues and yield losses and wrong characterisation that shall lead to difficult model validation and correction	8	selection of an inappropriate test structure	It can happen any time as no test strcture is a true representative of the product or a new technology	5	Select apropriate design test structures for a given technology	Detection is done by R&D engineers only in case of significant or abnormal parametric drifts	4	160	Design new test structures which are complex enough to test the new technology challenges	Device and interconnect modeling R&D teams, continous efforts
						lead time delays		Incorrect SPICE Models and Tools	It is a rare activity	1	Automate the SPICE model inclusions		4	32	Automate the SPICE model inclusion	IT, EDA Sections, as soon as possible
						increasing cost and waste of resources		wrong test simulation conditions	It could happen quite often, the wrong simulation condition might result in different target and corners	4	Generate test cases or automate SPICE simulations		4	128	Automate the process	IT, EDA Sections, as soon as possible
3			**Design Impact Analysis**	Assess the yield of an IP design taken from internal design library by simulating the eletcrical parametrs against target and corners specifications	incorrect yield assessments	misleading characterization and validation of model corners which shall result in yield losses when used with actual products during production	8	selection of an IP which is not true representative	It can happen any time so its detection is not apparent until prototyping	5	Select appropriate IP from design library which is closely representative of a product (technology adoption) and/or complex enough for a technology (technology alignment)	Detection is often at the Process Integration team level	6	240	Design new IPs which are complex enough to test new technology challenges	Device/interconnect modeling and process integration R&D teams, continous efforts
						lead time delays		Incorrect SPICE Models and Tools	It is rare but is likey possible	1	Automate the SPICE model inclusions		4	32	Automate the SPICE model inclusion	IT, EDA Sections, as soon as possible
						increasing cost and waste of resources		wrong test simulation conditions	It is not often but if engineers don't put attention, often result in late detection	4	Generate test cases or automate SPICE simulations		4	128	Automate the process	IT, EDA Sections, as soon as possible
4			**Variance Analysis (site to site)**	site to site PT variance analysis	unrealistic PT dispersion leading to enlarged model corners	model cannot be corrected (accepted by the designers) until and unless supported by the silicon results	8	data is not understandable (data extraction issues) and inappropriate data extraction tool with different formats	frequent occurrence	8	appropriate data extraction tool and training	Detection is at the R&D engineer level	6	384	Automate the process	IT, EDA Sections, as soon as possible
				scribe line vs. infield PT variance analaysis	wrong site to site mapping	impossible to differentiate between systematic and random patterns that leads to no model improvements even in the case of changing device behaviour	8	inappropriate data pre-processing (data pre-processing issues)	frequent occurrence	8	appropriate data extraction tool and training		5	320	Automate the process	IT, EDA Sections, as soon as possible
				site to site PT-Inline correlation	no/unrealistic/weak PT-Slicin correlation	reliable root cause analysis not possible during exercusions and crises within production line	8	incorrect site to site mapping (data alignment issues)	frequent occurrence	8	appropriate data extraction tool and training		5	320	Automate the process	IT, EDA Sections, as soon as possible
5			**Model Validation**	Compare the Models targets and dispersion with the targets and corners received from the prodution line Silicon (Wafer)	large significant variation between simulation and silicon results	respins	8	incorrect tool	frequent occurrence	6	appropriate data extraction tool and training	Detection is at the R&D engineer level	6	288	Automate the process	IT, EDA Sections, as soon as possible
						enlarged corner models leading to large chip area, power consumption and	8	incorrect measurements	frequent occurrence	6	Remove missing links between databases		6	288	Automate the process	IT, EDA Sections, as soon as possible
						improved yield if corners are large enough	8	incorrect statistics	frequent occurrence	6	appropriate training		6	288	Automate the process	IT, EDA Sections, as soon as possible
								wrong data preprocessing	frequent occurrence	6	Automate preprocessing		6	288	Automate the process	IT, EDA Sections, as soon as possible

Table 4.1 – Device / Interconnect modeling FMEA result

a) Data Alignment Sub-Process

The multi-source data alignment has emerged as a core limiting factors in finding root causes where missing common shared identifiers, identifiers with different names and insufficient identifiers are the key failure modes resulting in the data misalignment. The missing data dictionaries, inconsistent ETL routines, ontology issues (same concept with different semantics in different data sources) and database up gradation are the key root causes. Unfortunately there is no prevention possible until an overall database up-gradation is performed, which is not a practical solution; hence, the proposed solutions include: *(i)* database audits to find potential links between multiple data source, *(ii)* ETL routines audit to ensure accurate and complete data population in data sources, *(iii)* upgrade data sources to remove ontology issues and *(iv)* add more data dimensions along with an automated data extraction and alignment process.

b) Data Pre-Processing Sub-Process

Besides the fact that data is perfectly extracted, the incorrect data pre-processing might destroy the value within data. The root cause analysis is performed with an objective to quickly get the value from data as it degrades with time. The above proposed data extraction and alignment solutions do contribute in timely extraction of the value. The key failure modes in pre-processing are outliers skipping, inaccurate filtering and inappropriate data transformations, which refer to the inability to select and apply correct outlier detection and removal process. It requires an automated process or good end user training so that accurate value can be quickly extracted from the data.

4.4.1.2 Fast Technology Transfer

The fast technology transfer into SI business model provides an opportunity for early penetration into the market; however, success depends on the ability to quickly align new technology SPICE models against manufacturing resources. It is important to understand the difference between front-end (FE), back-end (BE), front-end-of-line (FEOL) and back-end-of-line (BEOL) technologies. The FE and BE technologies correspond to the manufacturing processes related with the fabrication of ICs and their packaging and assembly operations respectively whereas FEOL and BEOL operations are FE processes for the manufacturing of transistors and their interconnections respectively. In this fast technology transfer, the focus is put on FE technology; however BE technology is not different then FE operations. In FE technology, further emphasis is put on the SPICE model alignment which is one of the key steps in technology alignment. It is focused on the identification of PT parametric specifications (LSL, Target and USL). The FMEA results are presented in table 4.3.

It can be noted that the computed RPN values are above the cutoff value, so this function is critical for the success of fast technology transfer; hence, an appropriate method must be used to remove the associated causes. The principle objective is to find L/W for the geometries, which when simulated using newly received SPICE models gives us target and corners within the model specifications. It is further validated using the test products and comparing measured PT parametric results against simulated specifications. The data extraction, alignment and pre-processing again becomes critical for the success of this particular step.

The key failure modes identified are: *(i)* missing SPICE simulation results, *(ii)* varying simulation output formats and *(iii)* wrong simulation results. The first cause is identified as the missing geometries in the simulated files which do not allow us to go for either variance analysis or correlation purposes. The wrong test simulation conditions is an often encountered cause which is very hard to detect in normal procedure and has the worst effects on the SPICE model alignment efforts. The simulated parametric scale is often different than the measurement scale on production line; hence appropriate normalization formula is to be applied to analyze appropriate drifts.

Block	Sub-Block	Functions	Criteria	Failure Mode	Effect	SEV	Causes	Occurrence	OCC	Prevention	Detection	DET	RPN	Recommended Actions	Responsible and Target Date
Technology Derivative / Improvement Initiatives	Data Extraction	Data extraction (using operational data sources (ODS), data warehouse (DWH) and data marts (DM))	To extract data from multiple data sources *(i)* PT, *(ii)* Inline, *(iii)* WIP, *(iv)* Maintenance, *(v)* Defectivity, *(vi)* EWS, *(vii)* Facilities Data using *(a)* PT-Analysis, *(b)* Egg, *(c)* EWS, *(d)* BO, *(e)* PL/SQL, *(f)* KLA-ACE, *(g)* Klarity-Defect tools.	data not available in the data sources (ODS, DWH, DM)	device/interconnect model validation through prototyping is not possible	8	data is archived e.g. EDA data retenion period is 3 months whereas for FDC retension perios is 60 days	It is a factual position which always holds good	8	data retention period must be uniform for all the data sources	It is detected by R&D engineers	5	320	Increase or change database servers and storage space	IT and EDA Sections, as soon as possible
					lead time delays and waste of resources		operational data sources has long retention period, hence engineers prefer ODS then DWH and DM that result in incoherences as they don't have access to all the data extraction utilities	It is normal procedures for engineers as there is no SOP available in this regard	8	Engineers are trained on DWH and DM based data extraction utilities	End user detect it most often while doing extraction of the data using multi source data extraction tools like KLA-ACE, PL-SQL, BO	5	320	Automate data extraction process	IT and EDA Sections, as soon as possible
				end users are unable to extract data	no data preprocessing, no model validation based on the data collected across the production line, no variance analysis	9	End users do not have access or knowledge to effectively use data extraction tool(s) as data is extracted using diversified set of tools e.g. Egg, PT, EDA	For engineers it is quite frequent because during crises or process improvement efforts they are obliged to dig into the root cause, hence they extract multiple source data and align it for correlation analysis	8	end users must be given advance user training for the data extraction TOOLS at training center and refresher courses for continous support. As today engineers require multi source data to find out root cause, so is it practical to train engineers to use a diversified set of data extraction tools and tuilities?	It is often detected by the end user, as they are provided basic user training and they can do limited fucntions with the databases	5	360	Automate data extraction process	IT and EDA Sections, as soon as possible
			Usable data format (csv, dat, kdf, klf, stdf), if same data is extracted from different data sources e.g. (PT data is present in IDEA-PT database as well as EDA database, we have PT-Analysis and KLA-ACE tool for the data extraction). Even within one specified format we need to ensure that all extraction tools uses the same file format otherwise it shall require additional manual processing prior to data aligned.	File format not compatible to align the data, even csv format can have multiple versions, default version works fine but changing the platform raises this issues even harder.	difficult data alignment acroess multiple data sources, root cause assessment not possible, waste of time, resources and efforts	6	Incompatibility issue even within the same format files due to operating system and the tool used (e.g. even to generate csv from data extraction utility we have multiple compatibility formats). Default format shall work fine untill and unless we have same operating system like windows XP, with the change of the operating system compatibility issues can become critical.	It is a rare event for end users, as most of the time we have same platform and use default settings but multi source data extraction requires similar format for further analysis	6	no prevention	End user detect it during data alignement process across data extracted from multiple data sources	6	216	Unify data extraction formats through process automation	IT and EDA Sections, as soon as possible
			data extracted is understandable by all end users (PT, Inline, Equipment etc....) to perform root cause, variance, correlation and predictive modeling	data extraction does not carry description of the parameters and their nomenclature	difficult data alignment acroess multiple data sources, root cause assessment not possible, waste of time, resources and efforts	5	ETL routines do not extract the explaination of the these parameter names while data is extracted and data extraction utilities are not adapted to incorporate this information as wel	it is an always event	8	no prevention	End users while multi source data extraction and alignment prior to data analysis	5	200	Automate the data extraction process that by passes the need for data dictionaries	IT and EDA Sections, as soon as possible
	Data Alignment	align multi sources data, based on common shared identifiers for advanced analysis	multi source data aligned without any potential bias in the statistical analysis	missing common shared identifiers (missing database links), it means that within two data dimensions (data sources) we dont have common data columns that can be used to join these two dimensional data sources.	difficult multi source data alignment	8	missing data dictionaries against data model evolution	It is always the case	8	no prevention	end users detection, however could be detected by IT people	5	320	Audit the databases for potential links	IT and EDA Sections, as soon as possible
				empty identifier, it means that data columns as common identifier does exist but in the actual databases, thses columns are empty; hence cannot be used for alignement	unreliable root cause analysis and in some cases no root cause and correlation analysis	8	inconsistant ETL (extraction, transformation and loading) routines	It is frequent when ever we need to align data from multiple data sources	8	no prevention	detectable by IT people	5	320	Audit the ETL routines	IT and EDA Sections, as soon as possible
				identifiers with different name and nomenclature (ontology issue), it means that data columns identified as shared identifiers across multiple dimensions does have different names or nomenclature	waste of resources and efforts	8	ontology issues	data columns which are empty are always empty whereas other columns are always present	6	no prevention	end user detection	8	384	restructre the databases so that ontology issues are removed	IT and EDA Sections, as soon as possible
				insufficient identifier to align multi-dimensional data, it means that we dont have complete set of common identifiers (data columns) in all dimensions that can be used for full alignement, hence require some hypothesis adding data inconsistancy	Ineffective root cause analysis	8	database upgradations	Frequent	8	no prevention	detectable by IT people	8	512	add more data dimensions	IT and EDA Sections, as soon as possible
	Data pre-processing	clean data from outliers and apply distribution fitting for each parameter to get reliable and accurate analysis results	n-dimensional clean and noise free data	outliers not detected	noisy data, shall mislead interpretations and more probably no significant correlation or root cause analysis	8	inappropriate outlier detection process	Frequent	8	no prevention	end user detection	5	320	Automate data extraction process	IT and EDA Sections, as soon as possible
				significant deviation filtered as an outlier	inconsistent data, opportunity loss to learn from a significant deviation	8	misalignement of the data collected from multi-sources	Frequent	8	no prevention	end user detection	5	320	end users trainings	IT and EDA Sections, as soon as possible
				outliers removed without assignable cause		8	lack of expertise in outlier selection in a given data set	Frequent	8	no prevention	end user detection	5	320	end users trainings or new statistica analysis tools	IT and EDA Sections, as soon as possible
				lack of understanding data from different data sources	inability to classify and clean the outliers in other data sources	8	misalignement of the data collected from multi source	Frequent	8	no prevention	end user detection	5	320	end users trainings or new statistica analysis tools	IT and EDA Sections, as soon as possible
				incorrect transformation applied on the data to ensure distribution copmliance	faulty and misleading conclusions	8	lack of knowledge in statistics and method	moderate	8	no prevention	end user detection	5	320	end users trainings or new statistica analysis tools	IT and EDA Sections, as soon as possible

Table 4.2 – Data extraction, alignment and pre-processing FMEA Results

Block	Sub-Block	Functions	Criteria	Failure Mode	Effect	SEV	Causes	Occurrence	OCC	Prevention	Detection	DET	RPN	Recommended Actions	Responsible and Target Date
Fast Technology Transfer	Front End Technology	SPICE model alignment	Compute electrical parameters SPECS (LSL, Target, USL)	Missing SPICE simulation reuslts	missing PT specs	8	Missing geometries in the simulated output files	very frequent	8	no prevention	easily detected by R&D engineers when they dont find simulated PT parametrs in the output file	4	256	Automate the process because most of the steps occur in regular intervals and can be easily inducted into a customized network software which is shared by the PI and R&D engineers	IT and EDA Sections, as soon as possible
				Varying simulation output file formats	waste of time and resources		time consuming computations because of manual data alignment PT-Spice parameters and normalization	It is always the case	9	no prevention	It is always experienced by R&D engineers	3	216		IT and EDA Sections, as soon as possible
					increased lead times		difficult to compute and apply normalization formulas because simulated PT parameters and measurements have different measurement scales	It is always the case	9	no prevention	It requires utmost care otherwise it is undeteced	6	432		IT and EDA Sections, as soon as possible
				wrong simulation results	mis-interpretation of the PT-specs		incorrect PT-SPICE parameter relationship	frequent	6	no prevention	It also requires utmost care otherwise it is undeteced	6	288		IT and EDA Sections, as soon as possible
					wrong PT specs		wrong normalization	It is not very frequent, but upon chaning the SPICE model, it is highly likely to happen	5	no prevention	It is easily detecedt only if the resulting value has significant deviation	5	200		IT and EDA Sections, as soon as possible
					yield losses		error in LW scaling factor	It depends on our expertise in design of experiment (DOE) and might turn up frequent	5	no prevention	It is difficult to detec as we are already in the phase of optimum selection of geometries for the simulation	5	200		IT and EDA Sections, as soon as possible
					svere manufacturability issues		Wrong simulation and test conditions	It is likely to be frequent	6	no prevention	Detection only in case of significant deviation other wise mostlty it is undeteced	6	288		IT and EDA Sections, as soon as possible

Table 4.3 – Fast technology transfer FMEA results

The most critical root cause is the misalignment of PT and SPICE model parameters that result in inaccurate PT-Specs. The PT and SPICE model parameters are the same electrical parameters but they have different names because SPICE models are developed in new technology alliance and then are transferred to partners' alliance manufacturing facilities for alignment. It is the responsibility of the receiving plant to maintain the PT-SPICE parametric relationship against each revision or maturity level of the model. The LW scaling refers to different geometries and potential shrink due to process imperfection. The L/W (length/width) of geometries requires expertise in the **D**esign of **E**xperiment (DOE) which might be an issue without proper training and competence. The proposed solution is an automated software tool that removes the human intervention and generates PT aligned specs as accurately as possible. It results in quick alignment and validation of the SPICE (device and interconnect) models and technology lead times and costs are significantly improved.

From the above discussion, it can be concluded that the core limitations are associated with the ability to extract, align and pre-process the production line huge data volumes and dimensions. It is an important step prior to variance analysis and model validation steps during technology alignment adoption efforts. The proposed solutions as the operational fixes are the automated data extraction, alignment and pre-processing utilities that provide clean aligned data for the root cause analysis.

4.4.2 Operational Fixes through Joint Projects

Five operational fixes have been made through joint projects with the teams working in the technology derivative/improvement initiatives. These operational fixes are the automated software tools that provide end users with unified and agreed format based multi-source data extraction, alignment and pre-processing facilities for accurate root cause analysis, which is critical for DFM effectiveness. The developed software tools are: *(i)* BEOL-variance analysis, *(ii)* KLA-Ace Recipe, *(iii)* EPP (equipment, product, process) life cycle, *(iii)* SMA (spice model alignment) and *(v)* ACM (alarm control and management) tools. A brief description of these tools, is presented as under; however, detail description can be found in the appendices D to H a.k.a. industrial contributions (IC1, IC2, IC3, IC4 and IC5).

a) SMA (spice model alignment) Tool (IC1)

This tool is developed for the process integration team that actively participates in the technology transfer and its subsequent alignment and validation steps. The primary responsibility of the engineers in this team is to validate the SPICE models for the new technology or its derivative. They perform characterization using CAD simulations for different geometric specifications (test structures) using SPICE models provided by the technology alliance. These simulated parametric results are validated against the metrology and inspection results from the production line. This step is likely error prone and time consuming as the SPICE model and metrology parameters does not follow the same naming nomenclature and units. It makes it quite difficult for the PI team to compare them with the simulated SPICE parameters. The simulation results have format issues that result in the data which is either not usable or lead to incorrect results. On average, each technology has 2000+ parameters to be controlled and monitored and an engineer is given the responsibility of 200 parameters. Till now this process is manual and the R&D engineers maintain relationship files between SPICE and test program files and excel is the most likely tool being used for analysis.

Our proposed tool facilitates engineers in computing and aligning the SPICE parameters with the simulated and measured electrical parameters for different technologies. It is a network driven utility being shared by multiple engineers and it supports them to change normalization formulas on geometric specification variations to analyze the potential impact in the simulated results. This tool has highlighted two significant failure modes *(i)* unstructured naming conventions and *(ii)* varying file format output.

b) BEOL Variance Analysis Tool (IC2)

This tool is developed for T2D team to analyze and identify significant parasitic components; resistance (R) capacitance (C) and RC, during the BEOL interconnect modeling process. It takes input, the measured parasitic

elements and computes, inter die, intra die and scribe line inter metal layer variances. The R&D engineers use excel to analyze the data and it is found that they apply ± σ filter while data extraction. This step screens out the significant deviations which are potential suspects; hence the resulting data closely follow the normal distribution which is not the case. A multi-dimensional data model has been proposed to perform requisite computations along with a comparison of results obtained with and without filters. This tool has resulted in the identification of missing values, wrong filters and missing coordinates to compute the geometric specifications of interconnect metal lines as the causes to our inability to model the newly emerging spatial variations or finding an answer to the yield drifts or excursions. The proposed tool has significantly reduced the processing time and resulted in error free computations.

c) KLA-Ace Recipe for PT-Inline Correlation (IC3)

The results obtained from the BEOL-Tool highlight significant parasitic variations which are further investigated for the root causes. The most probable justification for these variances could be the variations in the geometric specifications and/or defectivity. So it is needed to perform multi-source root cause (correlation) analysis at the site levels to answer the newly emerged behavior, an uncontrolled process or equipment drifts. It requires measurement coordinates so that transistors or interconnects geometries can be computed for the correlation purposes. It is found that electrical data is measured at site level with site number and (x, y) coordinates but physical measurements in database are supported with only site numbers. It is not possible to compute geometries because the wafers are rotated during the metrology steps based on the test structure orientation; hence same site number allotted to two different measurements is not likely to be the same site on the surface of silicon wafer. The information from MASK data is used to develop KLA-ACE recipe for the coordinates generation followed by its mapping with the PT data. The analysis further highlighted that there are a lot of missing values because there are certain metrology steps where only 9 sites are tested due to time constraints. In such case there are only 9 sites data available to perform statistical inference which cannot be trusted; hence it is recommend that 17 sites metrology must be ensured, otherwise effective root cause analysis might not give correct results. This recipe enables R&D engineers to perform an effective root cause analysis to find answers against variations at the site levels. The said tool has further highlighted the need to extend the data alignment from site to die levels and its scope from PT and Inline data to defectivity and EWS (electrical wafer sort) data. The multi-source root cause analysis requires capturing and modeling newly emerging spatial variations.

d) EPP (Equipment, Product, Process) Life Cycle Tool (IC4)

The results obtained from the BEOL and KLA-Ace recipe justify the parasitic drifts against geometric specification variations, however it does not provide an answer if the drift was caused due to process or equipment variation. In order to perform an in depth analysis, the proposed tool exploits the manufacturing data sources and generate product and equipment life cycles. This tool equally serves the WP3 (work package-3) in the European project IMPROVE where **E**quipment **H**ealth **F**actor (EHF) can be computed to predict maintenance events prior to its failure. This EHF indicates if the equipment was in good condition when a particular process step was performed. The EPP tool directly connects with the maintenance (TGV) and **O**ut **o**f **C**ontrol (OOC) databases for the equipment related data extraction whereas product and process data is extracted from the process database using KLA-Ace recipes. All these data are input to the EPP tool that perform consistency checks and populates it into a multidimensional database. It provides a user friendly interface to extract customized equipment and product life cycles. This tool provided us a real insight to the limitations while processing data across multiple databases and serious ontology issues, missing links, unstructured data model evolutions and the huge data volumes are found as significant limiting failure modes.

e) ACM (Alarm Control and Management) Tool (IC5)

This tool is developed for the Lithography equipment engineering team. At present engineers do not have information about alarms categorization based on machine states; hence INGENIO (equipment alarms) database is linked with maintenance and process databases to classify alarms based on the machine state. It reduces number of alarms as well as unnecessary corrective actions that might cause serious issues in the normal production. This tool provides engineers an opportunity to classify and extract only relevant tools and respective lot, wafer, process and SPC information. It also enables our engineers to generate alarms and states data on the equipment at module levels

to be used in the predictive alarms (PAM) and states (PSM) models. These models are the heart of the YASS strategy (SC6, Chapter-7) because they predict likely yield loss for each production lot. The lots predicted bad or suspected are inspected to control the bad lots passing to the next steps. It results in the availability of additional metrology and inspection tools capacities to be used for R&D purposes.

The above presented tools were deployed as operational fixes in technology derivative improvement/ initiatives in an anticipation of problem resolution, but they further highlighted cyclic failure modes and root causes that require generic R&D solutions. These cyclic failure modes and root causes are presented in section 4.4.3. The complete UML models for these tools (operational fixes) along with data models are presented in the appendices D to H for reference.

4.4.3 Step-3: Cyclic Failure Modes and Root Causes

The fast technology transfer is not further included for the investigation of cyclic failure modes and causes because of the fact that new technology is transferred to alliance partners' manufacturing facilities where success lies in our ability for its quick alignment and subsequent adoptions. It can be said that it is about understanding the manufacturability and yield loss mechanisms and its quick modeling so that first time correct design can be achieved. In step-2 of the *i*-FMEA methodology, 5 operational fixes have been applied based on initial failure modes and associated root causes. These tools further highlighted and converged into cyclic failure modes and root cause that require generic R&D solutions. The step-3 is performed on *(i)* effective root cause analysis and *(ii)* data extraction, mapping and alignment and results are presented in Table 4.4.

The failure modes and root cause identified here are called cyclic because of the fact that besides the operational fixes they are repeated on regular basis; hence it requires a generic R&D solution so that we can empower our engineers with the ability to dynamically exploit the huge production data volumes and dimension. The analysis on the effective root cause analysis function resulted that besides site/site level mapping for PT and Inline data sources we are not able to find the root causes against significant drifts. In order to address this situation we need to enlarge the scope of multi-source root cause analysis from PT and Inline towards defectivity and EWS data sources. It further highlighted the issues of die to die mapping and die to site qualification because PT and Inline data is captured at site level but defectivity and EWS data is available at die level with no site information that we can use to map it with site level data sources. Besides these facts it is a fact that all metrology measurements during the process are made on the test structures which might be present in the scribe lines or fields; hence to rightly capture the spatial variations it is important to shift our analysis based on the test structure position and mapping is done based on the test structures with shortest distances. It is highly critical to capture the spatial variations emerging in new technologies whereas excursions, drifts and variations require site/die level correlation analysis. The generic model SPM (spatial position modeling) is proposed for this purpose which is presented and discussed in chapter-5.

The multisource data extraction, mapping and alignment are the other functions where we have encountered cyclic failure modes and root causes. The root causes identified are inconsistent data models, unstructured data models, missing data dimensions (insufficient identifiers) and data retention periods. The tools that we deployed for different R&D teams were reported no more working after some period and we found that IT people have changed the data model without any analysis of its potential impact on existing data extraction and analysis tools. We were obliged to develop new versions of the tools so that engineers can continuously use the tools for multi-source data extraction in unified format. We faced huge difficulties in finding the changes made by the IT people and we encountered that unstructured data model evolution is the biggest issue today.

Block	Sub-Block	Functions	Criteria	Failure Modes	Effect	SEV	Causes	Occurrence	OCC	Prevention	Detection	DET	RPN	Recommended Actions	Responsible and Target Date
Technology Alignment and Adoption Efforts	Effective Root Cause Analysis	Manufacturability and Yield Loss Mechanisms	Efficient and effective (die/site) level root cause analysis	no root cause identified	increasing technology lead times and costs	9	newly emerging spatial variations due to miniaturization resulting from Moore's law	Frequent	9	The failure mode and its root causes are cyclic in nature and operational fixes as provided earlier proved that such fixes shall not work in future and require generic R&D solutions	It is very difficult to detect without efficient and effective root cause analysis	5	405	Generic model for test structure position based data mapping and correlation analysis (SPM: Spatial Positioning Model, Chapter-5)	IT and EDA Sections, as soon as possible
					waste of resources and time	9	multi source data extraction, alignment and mapping (site or die level) issues	Frequent	9				405		
				no significant correlation identified	no rules as constraints to enhance manufacturability or yield	9	missing scribe line and infield test structure positions	Frequent	9				405		
					no variance model or improvements	9	reticle/die level mapping does not take into account spatial variations	Frequent	9				405		
	Data Extraction, Mapping and Alignment	Multi-Source Data Extraction	Quality data	noise	mis-interpretation	8	wrong measurements	Frequent	9	Automate the process and avoid ETL inconsistancies	Its detection is at R&D engineers level, but its too late when it is deteceid	6	432	Generic model to pre-assess the query failures and execute the query against inconsistancies (ROMMII: Referential Ontology Meta Model for Information Integration, Chapter-6)	IT and EDA Sections, as soon as possible
					wrong results	8	incorret ETL routines	Frequent	9				432		
							incorrect outlier filters and no statistical knowledge	Frequent	9				432		
			No missing values	missing values	few data for statistics	7	skipped measurements	Frequent	8	data model inconsistancies must be checked against the data extraction and analysis tools	It is normally not possible to make data models consistant withwhat we have as data collection across the production line	8	448	Continous restructuring of the data models to take into account new valuable data dimensions taht might help engineers in finding true root causes (ROMMII: Referential Ontology Meta Model for Information Integration, Chapter-6)	IT and EDA Sections, as soon as possible
					pre-processing not possible	8	inconsistant data models, sever ontology issues	Frequent	9				576		
							computed parameters	Frequent	7				448		
			multi-source data extraction using e xisting tools	user unable to extract mult-source data	no analysis	9	access to multiple data sources and training	Frequent	9	unstructured data model evolution must be controlled with uniform data retention	It is always the case and it results in highest inconsistancy	9	729	Generic model to manage the data model evolutions and R&D data model for uniform data retention (ROMMII: Referential Ontology Meta Model for Information Integration, Chapter-6)	IT and EDA Sections, as soon as possible
				missing data dimensions	no alignment	9	data retention period	Frequent	9				729		
					waste of time and resources	9	unstructured data model evolution	Frequent	9				729		
		Multi-Source Data Alignment and Mapping	Multi-Source data alignment at Site/Die levels	no multi-source data alignment	no site/site or die/die level mapping (PT and Inline data)	9	missing measurement coordinates	Frequent	9	There is no operational fix that shall help us, so at present no prevention	It is obvious and is encountered on daily basis by the engineers but its impact is huge as it do not support the mult-source effective root cause analysis	9	729	Generic R&D solution for the multi-source data mapping and alignment (MAM: mapping and alignment model, Chapter-5)	IT and EDA Sections, as soon as possible
				no multi-source root cause analysis	no die/site alignment (defectivity and EWS data vs. PT and Inline Data)	9	non-standard coordinates	Frequent	9				729		
				only single-source root cause analysis	waste of resources and time resutling in increased lead times	9	insufficient identifiers	Frequent	9				729		

Table 4.4 – FMEA results on cyclic failure modes and root causes

We also noted that with the new revolution in IT technology, new metrology equipment are capable of providing new additional data dimensions which are not updated or included in the existing data models based on the fact that those databases are of proprietary nature and managed by the external companies. It has resulted in inconsistent data models with a lot of missing dimensions that might help engineers in effective root cause analysis as well as means to effectively bridge the missing links between databases. The size of the database is the biggest issue not because of the cost or storage capacities but subsequent data exploitation. It is the reason that data retention period in ODS is longer than the DWH and DM. The ODS hold single source data; hence our R&D engineers spend huge amount of time in data extraction using single-source data sources followed by manual mapping and alignment. One more convincing reason is that today the production and R&D engineers' use same data sources where they have different data needs. The production and process engineers do not need data more than the product life cycle which is 8-12 weeks whereas R&D engineers' need data for at least 1 year. We can conclude that today's engineering and production databases, data extraction tools and analysis utilities are tuned to support production and process engineers whereas R&D engineers also try to benefit from the same source but without success. It requires a separate R&D data source which is designed and controlled by our internal IT section and we can structure its evolution on as and when need basis without impacting the performance of existing R&D data extraction tools. This situation shall persist until and unless a generic R&D solution is searched and applied.

The cyclic root causes as discussed above are true challenges that result in ineffectiveness of the DFM methods. Generic solutions have been proposed as *(i)* ROMMII (referential ontology Meta Model for information integration) and *(ii)* R&D data model to address inconsistent data model, unstructured evolution and data retention period issues. These solutions are presented and discussed in detail in the chapter-6.

4.4.4 Generic R&D Solutions

The brief description of proposed generic R&D solutions a.k.a. scientific contributions (SC) is presented as under:

a) MAM (mapping and alignment) Model (SC1)

There are four types of metrology data which is most commonly used by the R&D engineers to perform root cause analysis. It includes electrical (PT), physical (inline), EWS and defectivity data. The PT and inline data is captured at the site level from the test structures in scribe lines whereas defectivity and EWS data is measured at the die levels. The varying metrology coordinates for these measurements and missing die to site references make it almost impossible to perform accurate multi-source root cause analysis. A generic MAM model is proposed to resolve these die/site level mismatches. The polar coordinate system has been used as a principle to develop the said model. It enables our engineers to accurately compute the geometries of transistors and interconnect metal lines followed by correlation against drifts and excursions.

The proposed model have resulted in a dream come true for the engineers because till now they have been trying to find causes against newly emerging spatial variations using single-source wafer, site or die level analyses that contributed towards the inefficient DFM methods. These inefficiencies are being compensated in the technology alliances that result in exponential increase in the technology costs and extended technology adoption and alignment lead times. The proposed MAM model enables the local DFM efforts for continuous technology derivative alignment and adoption improvements. It removes the die/site level mismatches as well as generates die to site qualification for multi-source die/site level effective root cause analyses.

b) SPM (spatial positioning) Model (SC2)

The results obtained from the MAM model enables us to base our root cause analysis at die or site levels. It is important to note that the parameters are measured using scribe line or infield test structures. These test structures are not true products but are taken as the representatives of the actual products and are spread across the surface and scribe lines. The newly emerging drifts are due to spatial variations; hence root cause analyses performed based on the die/site level might not accurately capture these spatial variations. The SPM model is proposed to perform

correlation based on the shortest distance between test structures used for the measurements, to ensure that spatial variations are well taken into consideration. It computes the shortest distance and identifies the parameters along with coordinates for the correlation purposes. It is important to note that with the increasing number of parameters and test structures the computational cost increases exponentially. An optimized algorithm is presented that improves the efficiency by 50%. This model is important for the technology alignment and adoption efforts specially when improving the yield excursions.

c) ROMMII (referential ontology Meta model for information integration) Architecture and R&D Data Model (SC3 and SC4)

The proposed architecture enables us to remove the limitation associated with the unstructured evolutions of data models. It allows us to add new data dimensions when needed to ensure that engineers are provided with up-to-date dimension for an accurate root cause analysis. This platform learns the Meta model for the R&D data model and performs an accurate versioning against all potential evolutions. Any data query spanning over multiple periods is divided into sub-queries to avoid errors and resulting data is merged prior to its distribution to end user.

The R&D data model (SC4) is proposed to avoid the inconsistencies in the retention periods for different databases. The growth in the data sources often result in archiving the old data, however the archived data can be uploaded to the respective databases if needed. To make it more convenient for the engineers, the database must hold 12 months data. Today, the production databases are used for R&D purposes; hence 12 months period data retention is not likely possible due to volume and performance issues. It must also be noted that most of the existing production databases are of proprietary nature, hence any structural change is to be requested that further delays our efforts as well increases the costs. This R&D data model is to be maintained in house by IT or R&D people to allow its evolution and restructuring on as per need basis under quality control loop.

The algorithms with this ROMMII platform perform pre-failure analysis upon any potential structural change in the data model. It computes statistics at application and user levels and upon any potential change performs a risk analysis on the likely impacted users and applications. The potential users are also intimated about the newly added data dimension to ensure that all users are aware of any new changes.

4.5 RESEARCH SCHEMATIC AND ADVANCEMENTS (I-FMEA METHODOLOGY)

The research schematic is presented in Figure 4.4 to show the advancement. The proposed *i*-FMEA methodology holds most of the advancements except the YASS (yield aware sampling strategy). The block A0 correspond to step-1 of the proposed methodology and blocks A1 to A4 correspond to the step-2. These blocks have already been discussed in the chapter-3 where based on the analysis results an extended IDM (e-IDM) fablite model is proposed. The blocks A5 and A6 present the traditional FMEA results (step-3) along with 5 industrial contributions a.k.a. operational fixes from IC1 to IC5. These operational fixes are highlighted as screen shots of the developed and deployed software tools. The blocks A7 and A8 correspond to the step-4 which is focused on identifying the cyclic failure modes and root causes followed by generic solutions a.k.a. R&D fixes as SC1 to SC4. The SC5 (*i*-FMEA) is the scientific contribution summarized and presented in this chapter to find cyclic failure modes and root causes.

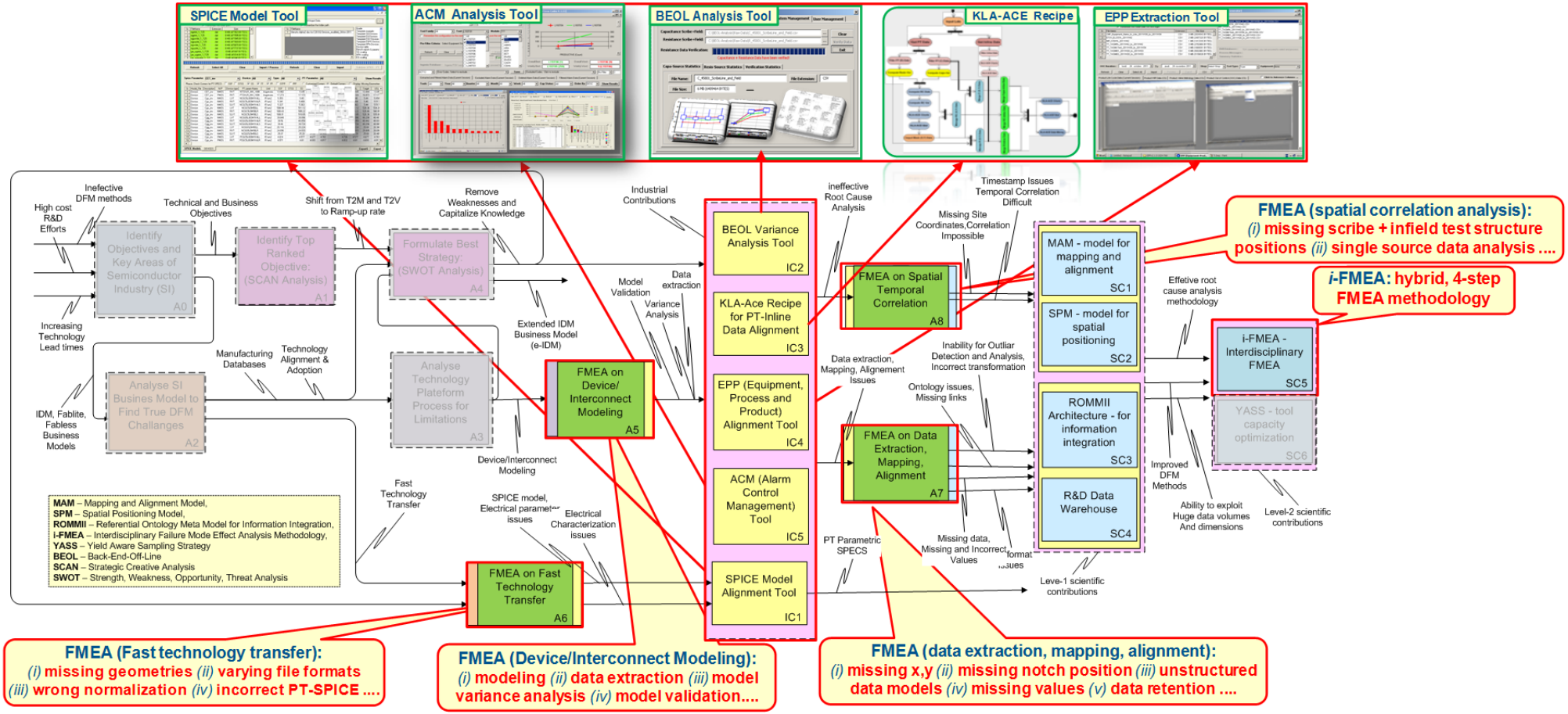


Figure 4.4 - The research schematic and advancement with *i*-FMEA methodology

4.6 SUMMARY AND CONCLUSIONS

In this chapter we used our proposed *i*-FMEA approach to find true DFM challenges. In this approach, the first and second steps are attributed to chapter-2 where the top ranked business objectives are identified as the "leadership position and quick ramp-up rate" along with data extraction, alignment and pre-processing as initial failure modes. The initial root causes identified are ontology issues, missing links between databases, missing values and varying measurement reference frames that restrict engineers to the single-source root cause analysis. We have developed 5 different tools *(i)* BEOL-variance analysis to analyze parametric drifts in back-end-of-line interconnect modeling process, *(ii)* KLA-Ace recipe for the mapping and alignment of coordinates to enable multi-source site level analysis, *(iii)* EPP (equipment, product, process) life cycle extraction tool to extract multi-source contextual data to support drift/excursion analysis and *(iv)* ACM (alarm control and management) tool to extract, analyze and manage alarm data for the tool capacity optimization. These tools address the above initially identified root causes. The objective is to either confirm the existing or identify new failure modes and it was surprising to find new failure modes as: *(i)* unstructured data model evolutions, *(ii)* missing data dimensions and *(iii)* wrong correlation due to test structure positions.

These failure modes were grouped as *(i)* ineffective root-cause analysis (infield and scribe line test structure positions) and *(ii)* data extraction, mapping and alignment, which need generic solutions. A detailed FMEA analysis on these two grouped failure modes was performed and resulting root causes were used to propose generic solutions. The new root causes linked with the ineffective root cause analysis failure mode are found as missing and varying coordinates due to varying metrology reference frames, rotation of the wafer prior to measurement steps and infield/scribe line test structure positions. It is very important to note that the root causes identified could fall in different business function and in our case these are IT and EDA.

The FMEA is also performed on the technology transfer that resulted in issues between SPICE model and PT parameters relationship. The names of the parameters in these SPICE models are generic whereas the names used by the test engineers are customized; hence the mismatch creates a significant problem. The initial root causes that were found are incorrect and error prone results because of manual data alignment and pre-processing; hence, we developed SMA (spice model alignment) tool that resulted in automation and removed the associated inefficiencies. The further use of this software tool identified unstructured naming convention for the metrology parameters and the varying format for the CAD simulation results as key root causes.

The generic scientific contributions have been proposed based on the identified root causes. The generic solutions are presented in chapters 5 and 6 as *(i)* MAM and SPM models and *(ii)* ROMMII and R&D DWH model. The proposed *i*-FMEA approach has evolved based on the experience through these phases where it is noted that true failure modes and root causes can be traced to other business functions. These steps are formulated into 4-step *i*-FMEA methodology as *(i)* identify initial failure modes using FMEA, *(ii)* identify cyclic failure modes through joint projects, *(iii)* identify root causes against failure modes and *(iv)* propose generic R&D solutions. In the next chapter 5 and 6 we present *(i)* MAM and SPM models and *(ii)* ROMMII platform and R&D DWH model respectively as generic solutions.

Chapter 5: Measurement Coordinates Mapping, Alignment and Positioning

[11,12,13,14]

All the measurements during production processes are performed on the test structures, which are placed in the scribe lines and/or in the fields. The wafer position is marked by its notch position and it is rotated prior to measurements depending on test structures orientation; hence, measurement coordinates potentially vary with each metrology step. It is critically important to map and align the multi-source measurement data prior to the root cause analysis. At present, engineers spend huge amount of time in data alignment because of the missing x,y coordinates for measurements in the database, issues with metrology reference frames and missing identifier between site and die levels data. A generic MAM (mapping and alignment) model for site/site and die/die levels coordinates mapping and alignment followed by die/site qualification is proposed. In order to capture newly emerging spatial variations, it is important to perform correlation analysis based on the position of the test structure used to measure the source and target parameters. The SPM (spatial positioning) model is also presented that perform mapping based on the shortest distance between test structures on the wafer surface.

Contents

[11] Shahzad M.K., Tollenaere M., Hubac S., Siadat A., Extension des Méthodes DFM pour l'industrialisation de produits microélectroniques, 9e Congrès Internationale de Génie Industriel Montréal, 2011, Canada

[12] Shahzad M.K., Hubac S., Siadat A. and Tollenaere M., *MAM (mapping and alignment model) for inspection data in semiconductor industry*, 12th European APCM Conference, Grenoble France

[13] 2012Shahzad M.K., Hubac S., Siadat A. and Tollenaere M., *SPM (spatial positioning model) model to improve DFM methods*, 12th European APCM Conference, Grenoble France 2012

[14] Shahzad M.K., Siadat A. Tollenaere M. and Hubac S. (2012), Towards more effective DFM methods in Semiconductor Industry (SI), International Journal of Production Research, (submitted) TPRS-2012-IJPR-0638

5.1 INTRODUCTION TO DEVICE/INTERCONNECT MODELING

It is seen that the success of SI lies in our ability to quickly align and adopt new technology transfer followed by continuous technology derivative/improvement efforts. The device and interconnect modeling have (section 3.4, Figure 3.3) been demonstrated as the key processes. In chapter-3 and 4, a detailed discussion is done on the data extraction, alignment and pre-processing limitations but here interconnect modeling process will be analyzed to highlight the data mapping and alignment problem. The basic interconnect modeling process is presented in the Figure 5.1.

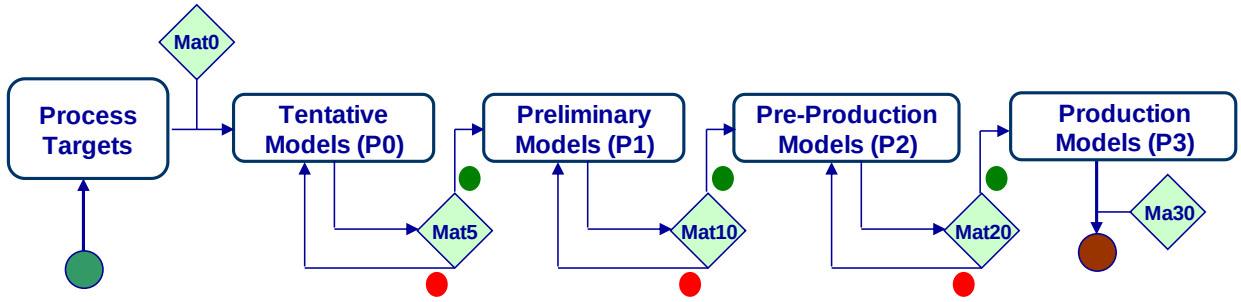


Figure 5.1 – BEOL - Interconnect modeling process

This process consists of 5 maturity levels defined for the SPICE models to ensure its potential manufacturability and yield robustness. It starts with tentative model step where interconnect stack is defined against the typical and corner models (Figure 5.2). The typical model refers to the target electrical and parasitic value whereas the corner models refer to slow and fast devices and their respective permissible parasitic or electrical parametric values to ensure the product functionality. The stack defines the geometric specifications of the metal lines and vias used to interconnect the devices on silicon surface. These stacks are simulated using SPICE (device /interconnect) models. The electrical parameters and parasitic are extracted for the validation purposes. This process is repeated until the geometric interconnect stack is found that fulfills the SPICE models target and corner specifications. The **Q**uality **A**ssurance (QA) is performed using defined stacks against the test product and results are benchmarked against SPICE target and corner models. The **T**echnology **t**o **D**esign (T2D) team is responsible for the design **I**mpact **A**nalysis (IA) using an IP from the existing design libraries to ensure the manufacturability and yield related issues prior to prototyping. The approval committee approves the results and the model for its inclusion in the **U**nified **P**rocess **T**echnology (UPT) database where the SPICE models and technology definitions are stored along with maturity levels.

A well known **B**usiness **P**rocess **R**eengineering (BPR) methodology is used to identify non value added activities. The objective is to propose efficient and effective process modifications. It can be noted that most of thee activities are classified as non value added activities as there is no transformation of the data into information. The model corner step is the most important step to be focused as it has 1...n frequency with transformation of data but no value is added. These repetitions often result in the technology lead times and waste of resources. The search of previous **D**esign **R**ule **M**anual (DRM) values to start the basic stack definition step, DRM freezing meeting, typical and corner model correction, acceptance meeting and the model production approval steps being non value added steps besides data transformation must be improved or removed for the overall process efficiency.

The preliminary and pre-production models (Figure 5.3) follow the same process and steps as the tentative mode. The only difference is that in the preliminary models 1 lot of 25 wafers is used and in pre-production model 3 lots are used to validate the SPICE models. The design respins often exist and the interconnect stacks are redefined until it complies with target and corner model specifications. In these sub-processes, model correction is the most critical step as R&D engineers needs data to answer drifts and variations but they have limited access due to which at this process we are unable to find root causes and correct models. It results in lead time delays, costs and waste of resources.

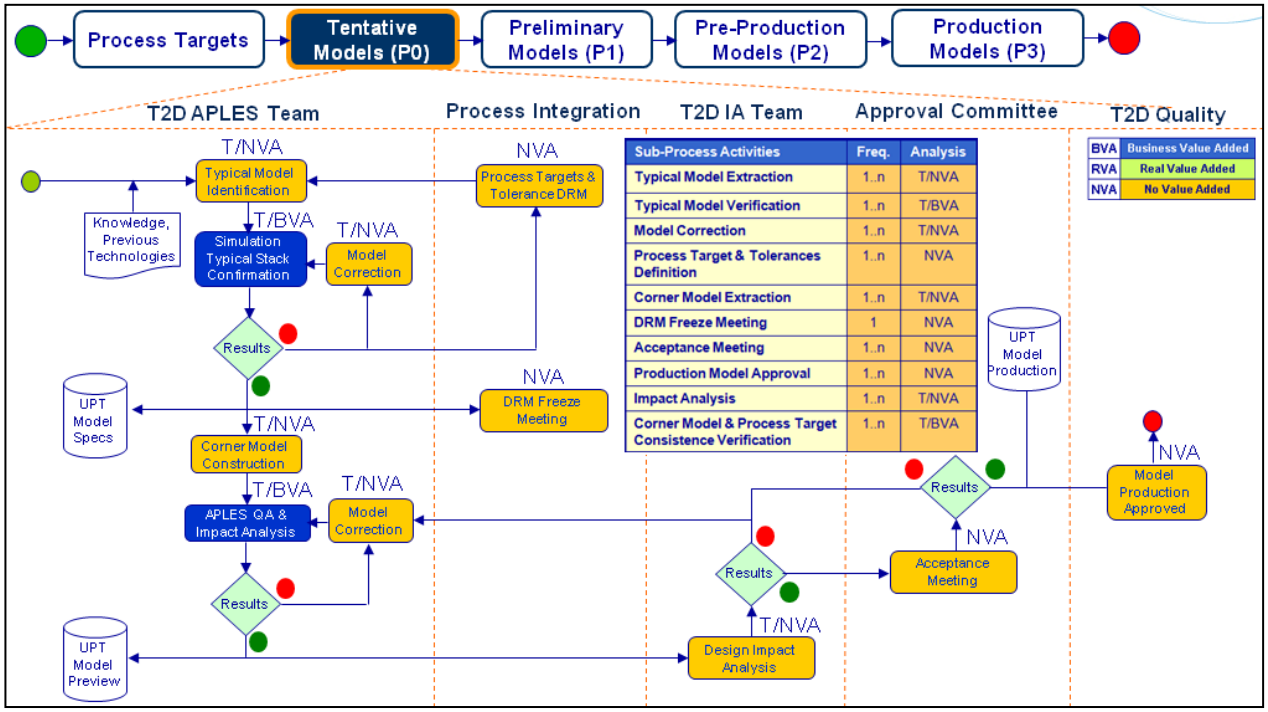


Figure 5.2 – The BEOL - Interconnect modeling process (tentative model)

The production model (Figure 5.4) is the last phase in interconnect modeling process where test structures (PCM: process control monitors, PCS: process control structure) are finalized that shall be used to capture the process and metrology information during normal production time line. A comparison between simulated and measured electrical and parasitic parameters is carried out with a formal report. This report is based on the results and analysis received from the PI team on 3 lots.

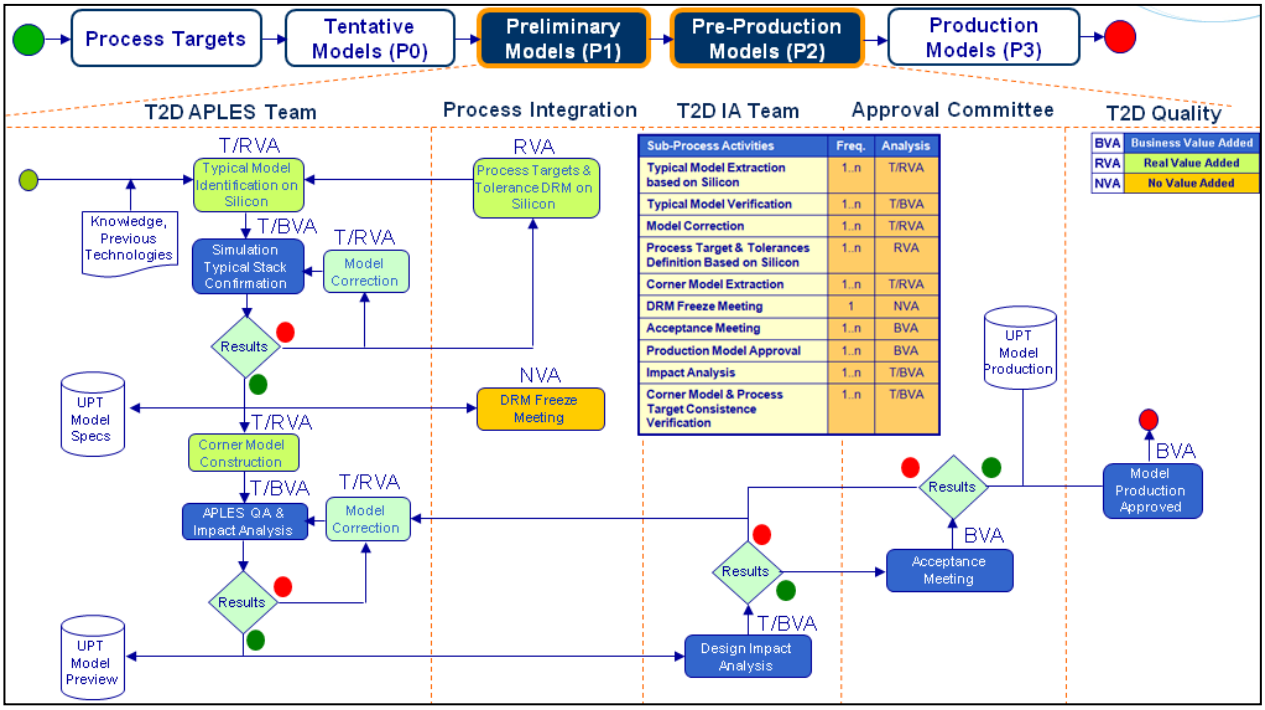


Figure 5.3 – The BEOL-Interconnect modeling process (preliminary/pre-production models)

On the basis of above interconnect modeling process description it is evident that most of the business value and real value are added during preliminary and pre-production models where the data extracted from the production line is analyzed to find root cause against drifts. The model correction loops

for both typical and corner models are time consuming tasks that directly increase the technology lead times. The improved process is presented in Figure 5.5.

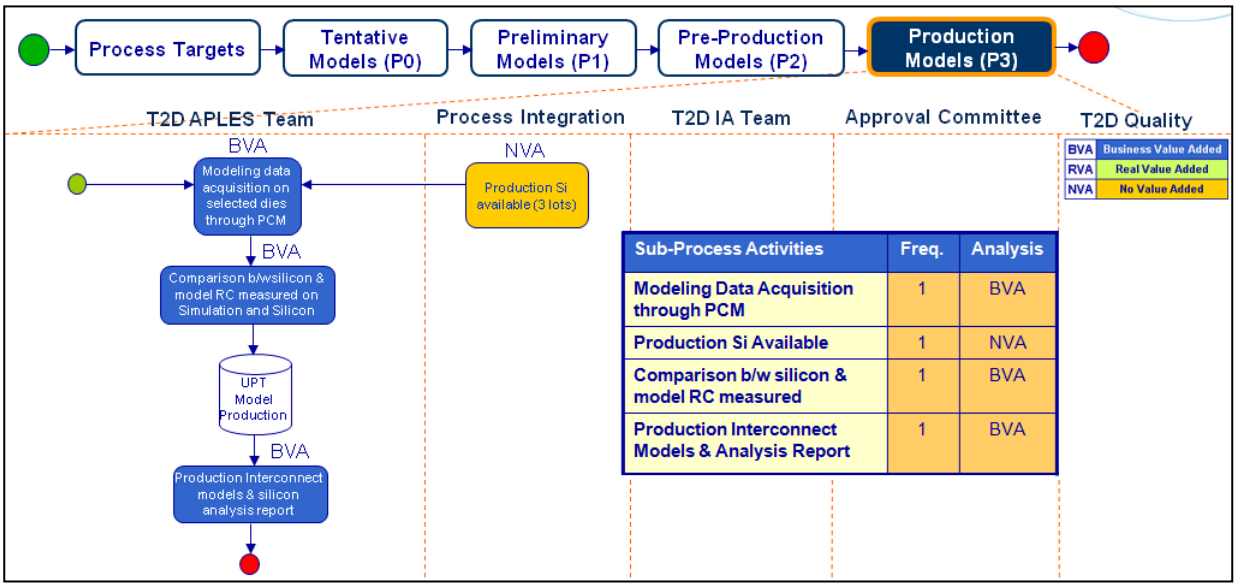


Figure 5.4 – The BEOL - Interconnect modeling process (production models)

The new process includes R&D data model with data retention period of +1 years. It acts as a knowledge capitalization repository where root causes and correlations against variance and drifts analysis are capitalized. It is further included during phase-0 (P0) where initial stacks are defined and simulated against the SPICE models. The second proposed improvement is to be made at the effective root cause analysis during model correction and validation steps. The proposed MAM model in this chapter provides us site and die levels coordinate mapping and alignment to support effective R&D efforts in the interconnect modeling. This ability shall reduce technology lead times and associated costs.

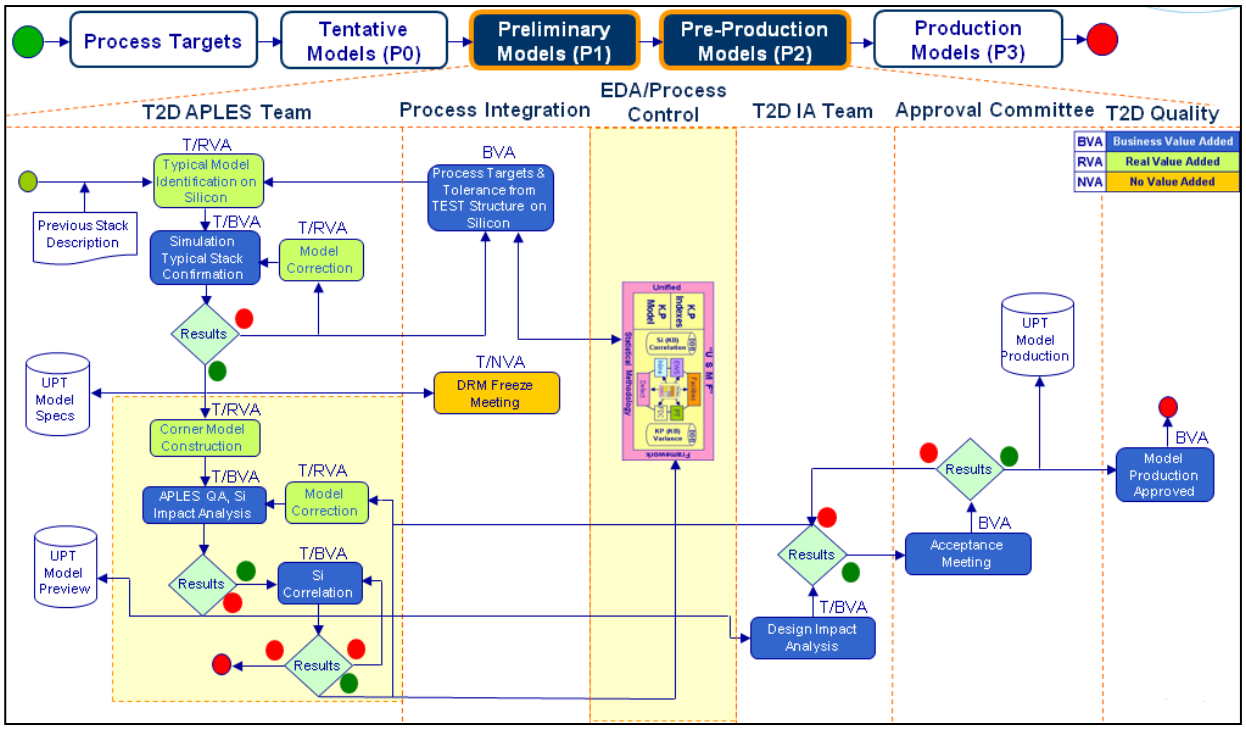


Figure 5.5 – Proposed BEOL - Interconnect modeling process

5.2 CHALLENGES IN MULTI-SOURCE DATA ANALYSIS

The SI manufacturing operations are grouped as Front-End-Of-Line (FEOL) and Back-End-Of-Line (BEOL) operations to manufacture the transistors and interconnections between them respectively. The production line data is classified as the contextual and inspection data. The contextual data includes maintenance, process recipe, Fault Detection and Classification (FDC), Work in Progress (WIP) and Out of Control (OOC) data. The inspection data comprises of PT, Inline, Electrical Wafer Sort (EWS) and defectivity data. The FEOL and BEOL operations can be divided into 1 to n operations where defectivity and inline data is collected across all operations whereas the PT data is collected at the end of metal-1 and metal-7 when contacts are manufactured for subsequent use of electrical tests. The EWS data is collected at the end of all processes to sort dies (chips) into good and bad dies. The most commonly used type of data during root cause analysis is the inspection data. The overall process is presented in the Figure 5.6.

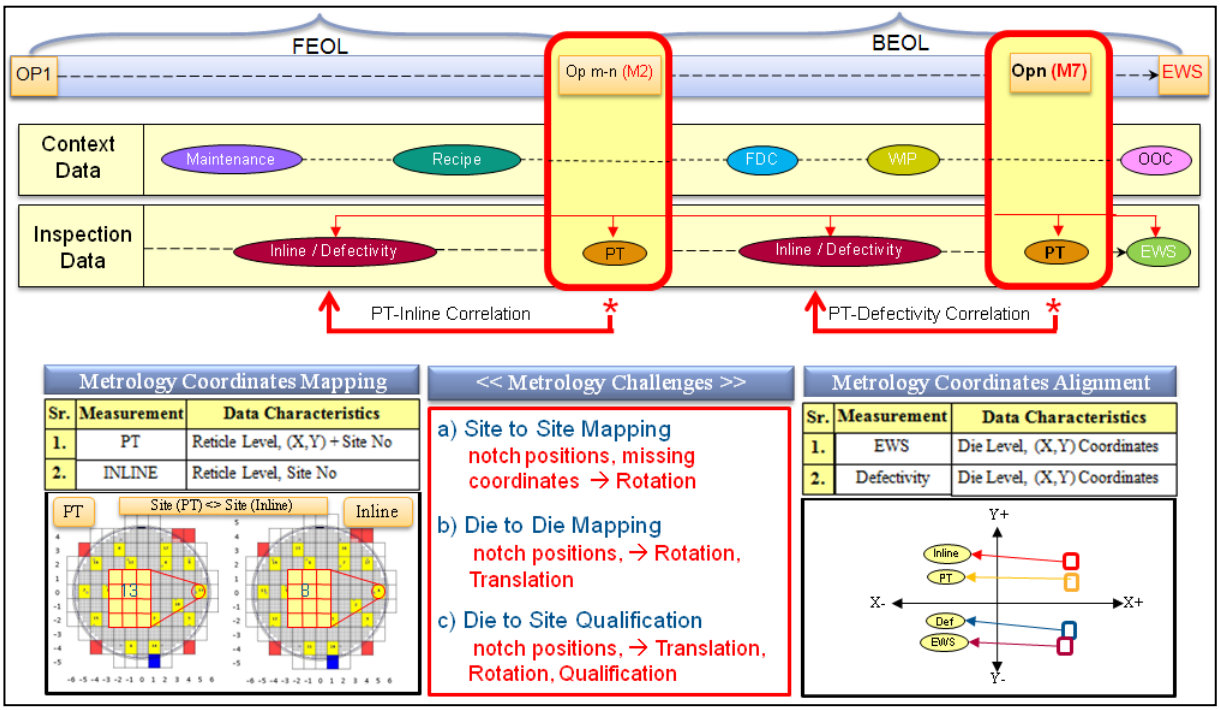


Figure 5.6 – Multi-Source data analysis challenges

The BEOL process is presented in Figure 5.7 for the manufacturing of a simple CMOS (complementary metal oxide semiconductor) inverter circuit with two transistors. The CMOS is a technology that enables the manufacturing of P-type and N-type transistors together on the same silicon wafer. The BEOL interconnections can be divided in three distinct parts as metal-1, metal-2 to metal-5 and metal-6 to metal-7. All the metal lines undergo the repetitive steps of geometric stack deposition, patterning (vias and lines), etching and metal filling. The number of metal lines depends on process technology being used and number of transistors in the circuit.

To ensure the product quality, metrology, defectivity, parametric and the inspection steps are added in production processes. The collected data enables us to compute geometric stack variations and physical interconnect geometries (width and thickness) followed by its correlation with PT (electrical) and defectivity (inspection) data as shown in the Figure 5.8 The parametric (PT) data is measured at metal2 and metal7 and is used to analyze the product functional compliance. It is important to note that increase in the transistor density requires more interconnect wires resulting in an exponential increase in the global interconnect length for the same chip area. These multi-step metal lines are separated by a dielectric, so they offer resistance to the current flow and act as a capacitor to cause signal delays. The parasitic (resistance/capacitance) drifts must be quickly analyzed to find root causes so that the drifts can be modeled to avoid the product failures and reduce the respins.

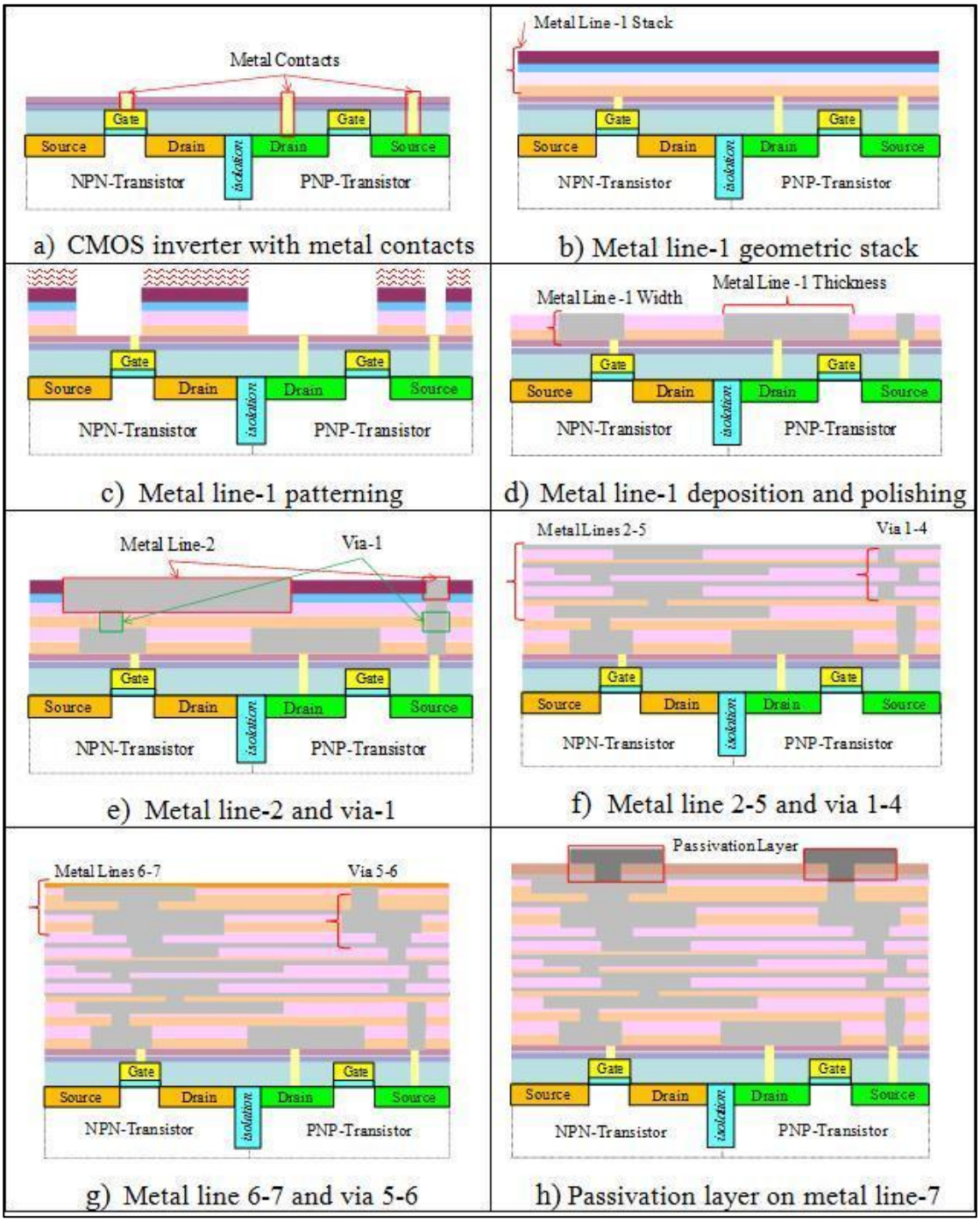


Figure 5.7 – The BEOL – process for interconnections

The possible type of multi-source root cause analysis could be easily classified based on the levels on which they are collected across the production lines. The PT and Inline data is collected at site level whereas defectivity and EWS data are collected at die levels (Figure 5.6). The potential root cause analysis could be performed at the site/site, die/die and die/site levels. The further investigation on existing production data sources highlight that all PT measurements are supported with the x,y coordinates and indices whereas inline data is supported only with the site numbers. It is also seen in the chapters 1 and 4 that wafers are rotated prior to the measurements based on the position of the test structure on the wafer. It is because of the fact that metrology equipment cannot move right or left to align itself with the test structure input pins or pads. It is evident that PT and Inline data collected across the wafer cannot be mapped because the site numbers present in both measurements correspond to different coordinates. In the absence of x,y coordinates for inline data it can be concluded that site level mapping is not highly difficult, time consuming and error prone. The die level x,y coordinates do exist in the source data files, however neither these values

nor the position of the wafer are uploaded in the database; hence it is highly difficult for the R&D engineers to perform even die/die level correlation analysis. The die/site level analysis is also not possible because there do not exist common identifiers that can be used for die/site level mapping.

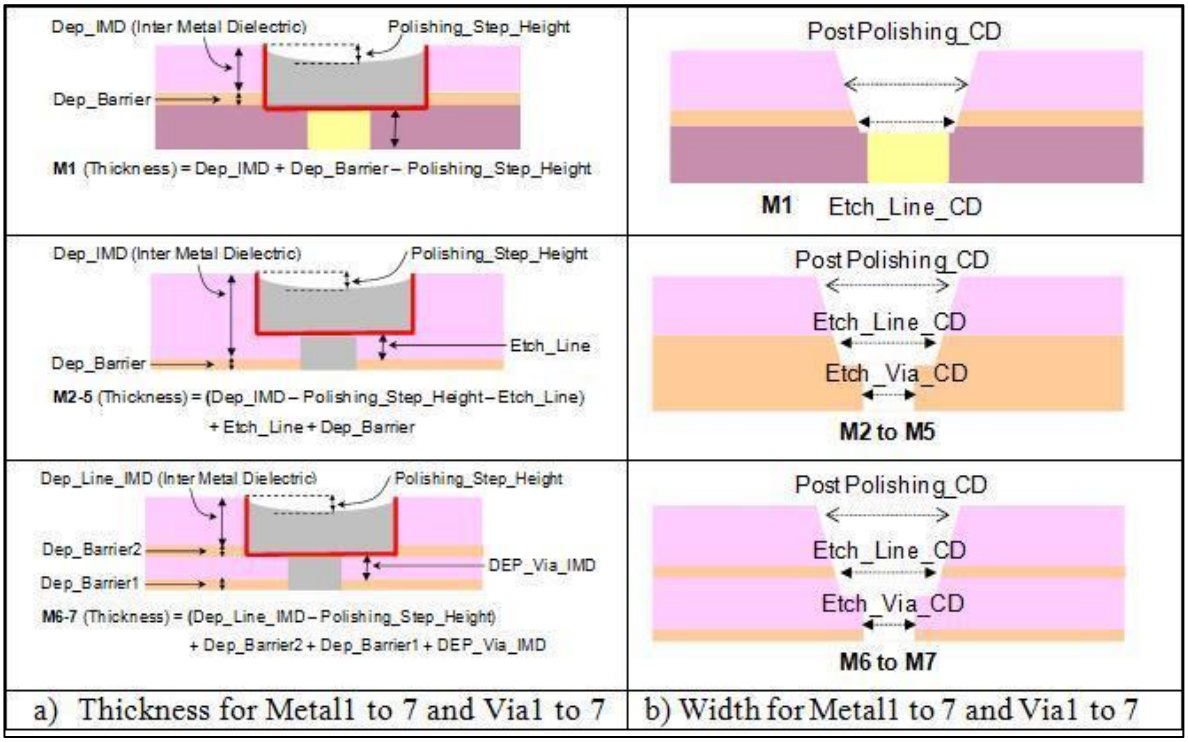


Figure 5.8 – The BEOL – geometric computations in BEOL process

It is seen that critical measurement information as *(i)* wafer position, *(ii)* x,y coordinates are either not available or they are not uploaded in the databases for the end users. The reason is that existing data analysis approaches are based on single-source data analysis but even if this information is made available to the engineers they are not able to deal with the level of complexity involved while performing coordinates normalization. It is because of the fact that the reference frames used during measurement operations are specific to the equipment and its manufacturer. In order to improve the productivity of the R&D engineers, these missing data dimensions must be uploaded in the data sources by restructuring the data models. The generic R&D solutions are required to perform die and site level mapping and alignment so that effective multi-source root cause analysis can be enabled to improve DFM ineffectiveness.

5.3 SITE/DIE LEVEL MISMATCH PROBLEM

The data is measured at the predefined site/die positions on the wafer using one of the four potential reference frames as presented in Figure 5.9. These reference frames are defined and fixed by original equipment manufacturer (OEE). The wafer rotation in addition to these reference frames makes the situation very complex to resolve site/die level mismatches for the accurate computation of geometries and its correlation with the drifts and inspection data.

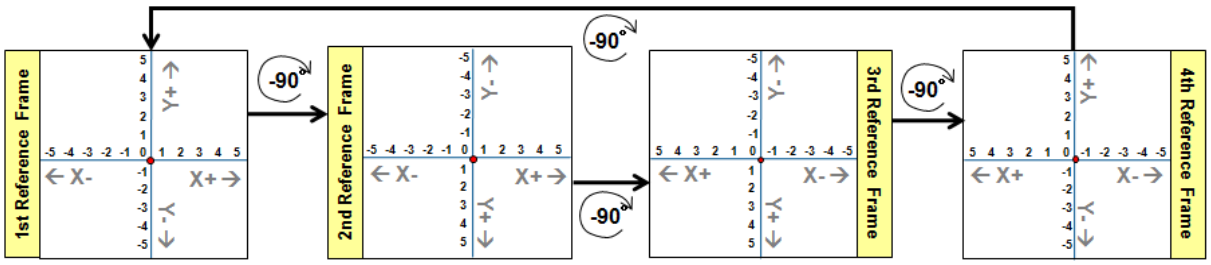
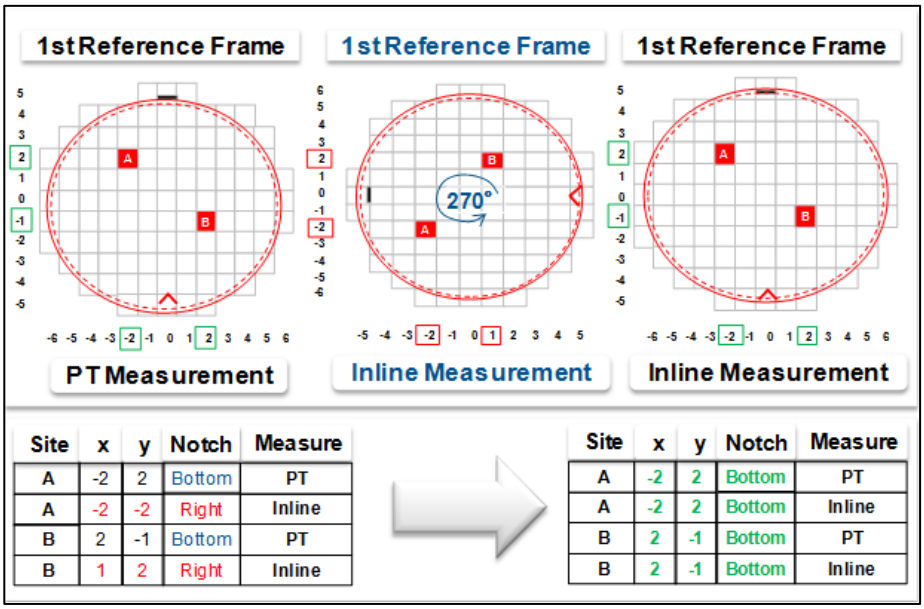


Figure 5.9 – Metrology reference frames

The reference frames used for the metrology can be shifted from one to another by a clockwise rotation to accurately map source and target site/die level measurements. The site/site level mismatches are presented in the Figure 5.10 where PT and inline measurements with notches at bottom and right respectively follow the 1st reference frame. Hence, notch position of the inline measurement is rotated by 270° anticlockwise to shift its notch position downwards for site/site mapping.

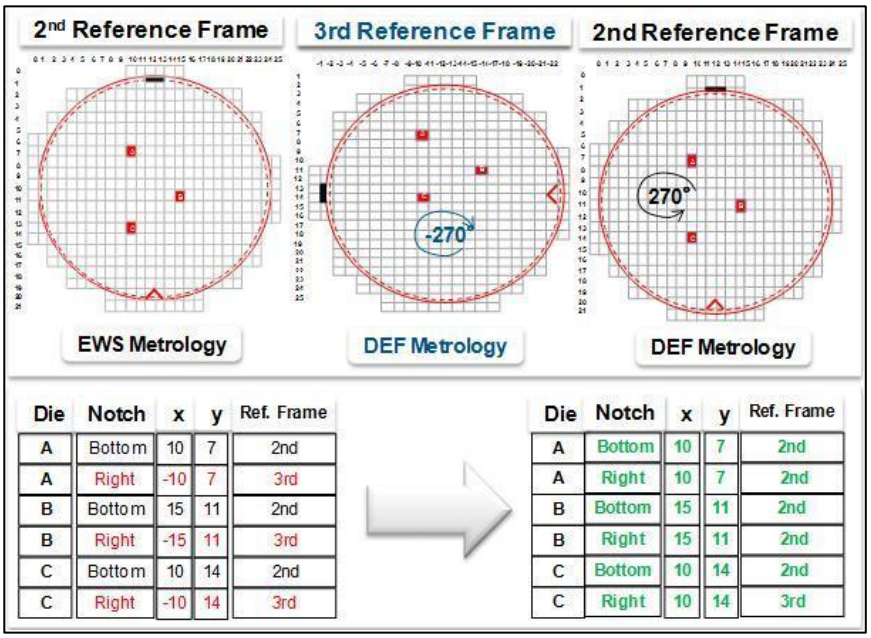


Site	x	y	Notch	Measure
A	-2	2	Bottom	PT
A	-2	-2	Right	Inline
B	2	-1	Bottom	PT
B	1	2	Right	Inline

Site	x	y	Notch	Measure
A	-2	2	Bottom	PT
A	-2	2	Bottom	Inline
B	2	-1	Bottom	PT
B	2	-1	Bottom	Inline

Figure 5.10 – Site/Site level mismatches due to notch position

Similar mismatches can also be found at die/die level measurements as presented in Figure 5.11 where the measurement coordinates of electrical wafer sort (EWS) and defectivity (inspection) data vary in their reference frames and notch positions. It is required to rotate the measurement coordinates of defectivity data by 270° clockwise to change reference frame from 3rd to 2nd frame and then 270° anticlockwise to move its notch position down for the die/die level mapping.



Die	Notch	x	y	Ref. Frame
A	Bottom	10	7	2nd
A	Right	-10	7	3rd
B	Bottom	15	11	2nd
B	Right	-15	11	3rd
C	Bottom	10	14	2nd
C	Right	-10	14	3rd

Die	Notch	x	y	Ref. Frame
A	Bottom	10	7	2nd
A	Right	10	7	2nd
B	Bottom	15	11	2nd
B	Right	15	11	2nd
C	Bottom	10	14	2nd
C	Right	10	14	3rd

Figure 5.11 – Die/Die levels mismatches due to reference frames

The die/site qualification is another problem that restricts our ability to perform the correlation between PT or Inline (site level) and EWS or defectivity (die level) data. It is presented with an example in the Figure 5.12 where inline (site) and EWS (die) data follow the 1st and 2nd reference frames with the

different notch positions and missing die/site level qualification data. First of all, it is required to rotate the die level wafer data by 270° clockwise and 270° anticlockwise to bring it to the 1st reference frame with notch position at the bottom. The x and y components of sites (inline data) associated with each die in EWS data must be found (Figure 5.12) for the die/site qualification.

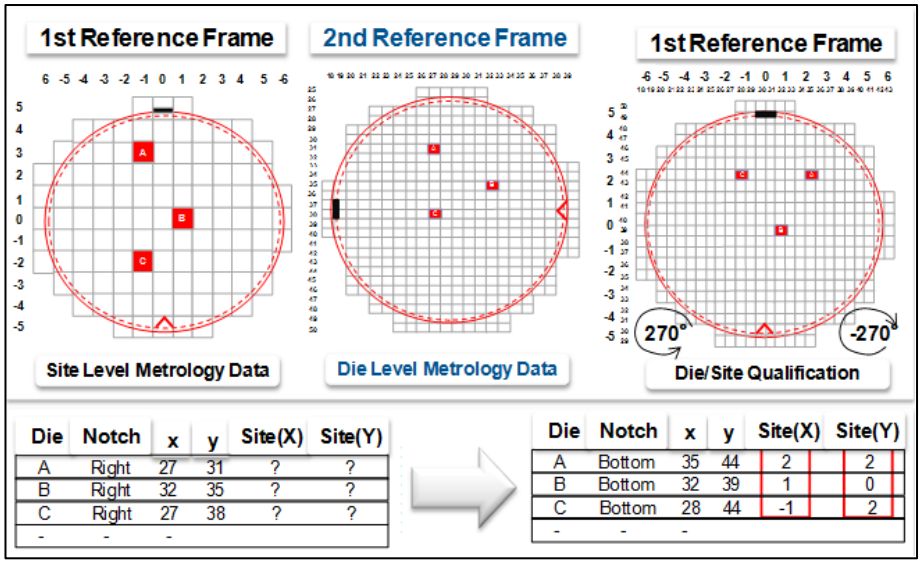


Figure 5.12 – Die/Site level qualifications and varying reference frames

5.4 Proposed Die/Site Level Mapping, Alignment and Qualification (MAM) Model

From the problem context discussed above in section-3, the problem is separated in two parts as *(i)* site/site, die/die mapping and alignment and *(ii)* die/site qualification. Let us first start with the site/site, die/die mapping and alignment issue.

5.4.1 Site/Site or Die/Die Level Mapping and Alignment

Our proposed mapping and alignment model is based on the transformation of the polar into Cartesian coordinates where each point is represented by its x and y coordinates and the rotation angle as presented below (Figure 5.13).

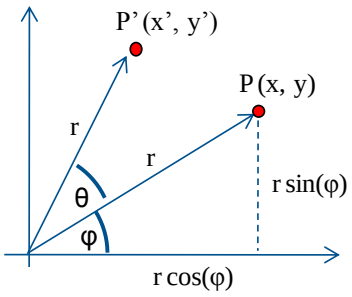


Figure 5.13 – Polar coordinates formulation and rotation

Any point P(x,y,φ) in the polar coordinates is represented in Cartesian coordinates as under:

(1)

(2)

If this point P(x,y,φ) is further rotated with reference to the same origin by θ° to P′(x',y′, φ+θ) then it can be represented as under:

(3)

(4)

The new coordinates of the point P′(x’,y′ ,φ+θ) with the substitution can be rewritten in the matrix notation as under:

(5)

(6)

(7)

In site/site and die/die mapping and alignment problem each site/die within the source and target wafers is taken as a point located in different reference frames with different notch positions. So the above presented concept is used to always rotate the coordinates of our sites and dies to the 1st reference frame with notch at bottom. In the proposed model is taken as the clockwise angle between two reference frames and as an anticlockwise angle that rotates the sit/die level coordinates to move the notch position at bottom.

To better understand the steps involved in this proposed formulation let us present an example as shown in the Figure 5.14 where a wafer is located in the 2nd reference frame and the 4th quadrant so all die or site level coordinates are known with reference to its origin. It is important to find these coordinates with reference to the origin located in the 1st reference frame. The original coordinates of the wafer are rotated by = - 270º to get new coordinates in the 1st reference frame. The resulting wafer is found in 4th quadrant and 1st reference frame but to move the notch position at bottom the wafer is further rotated around its center by θ = 90º followed by its translation to center.

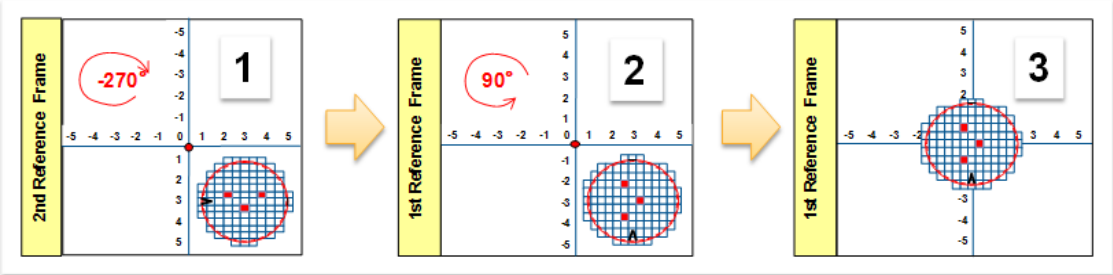


Figure 5.14 – Reference frame and notch position rotation and translation

The problem is generalized such that the point is located in the 2nd reference frame and the coordinates of this point are known with reference to its origin .We have to find the coordinates of the point with reference to the new origin located in the target reference frame where this point shall be further rotated by the angle to get the notch position at bottom. Finally it shall be translated to the center of the reference frame. This formulation is presented in Figure 5.15.

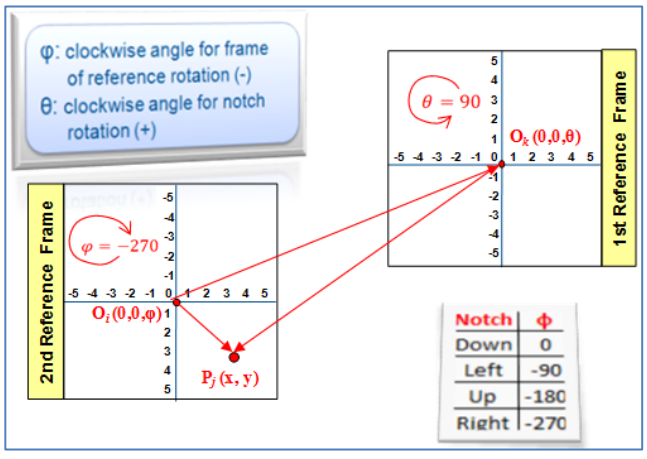


Figure 5.15 – Generic formulation for reference frame and notch position rotation

The points , and can be represented as vectors. The said point is known with reference to the point and our objective is to compute the same point with reference to . The both origins and have x and y coordinates as (0,0) and are superimposed on one another hence the angle between them is taken as 0. The new coordinates of the point with reference to can be found by taking a vector sum as under:

(8)

In this formulation the is a vector between points and , is a vector between points and and is a vector between points and . The matrix representation of the vector sum based on the Cartesian coordinates (transformed polar coordinates) is as under:

(9)

(10)

In this formulation and shall rotate the reference frame and the notch position respectively followed by its translation to centre given by and a.k.a. alignment factor. Now the algorithm developed to normalize the given wafer coordinates is presented below. Let us start with the description of the variables (Table 5.1).

Notations	Description	Notations	Description
	S → source, t → target		Reference frame for metrology
	Computed wafer center		Notch position
	No of sites/dies on x-axis		No of sites/dies on y-axis
	Rotation angle for Notch position		Rotation angle for Reference frame

Table 5.1 – Description of the MAM model variables

The proposed generic 4-step algorithm for site/site or die/die mapping and alignment is presented as under:

Step-1: Initialize

Step-2: Compute ( ), , ,

Step-3: For all measurements points in source and target,

Step-4: Add and to x,y coordinates of target wafer

The step-4 is very important to be focused as it provides an adjustment if the centers of the source/target wafer are different due to different number of dies/sites on their x and y axes. After this adjustment, they shall exactly map on one another. Now we present an example to validate the proposed model (Figure 5.16). In this example the source wafer is in the 3rd reference frame and is required to be

transformed to the 1st reference frame followed by its rotation of the notch position at bottom and translation towards the center. We shall transform P(-4,2) ' P'(x',y') [= -180, = 90].

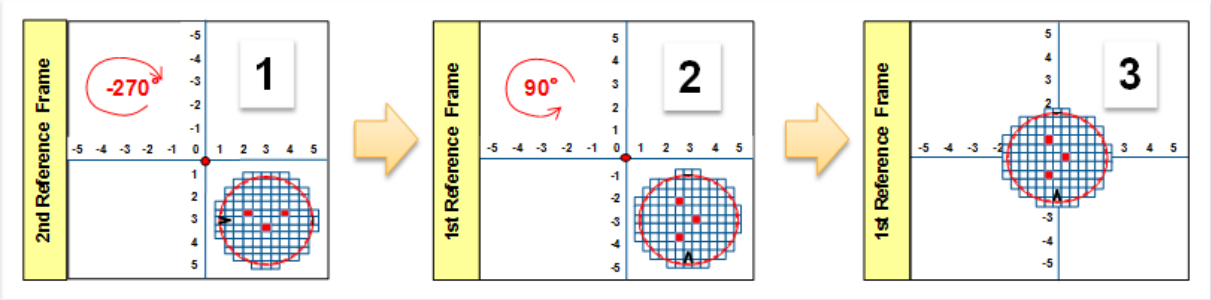


Figure 5.16 – Reference frame and notch position rotation with translation example

After transformation of all the coordinates, the results are presented in the Figure 5.17.

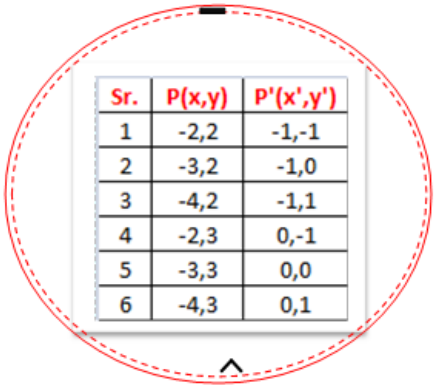


Figure 5.17 – Full map transformations

5.4.2 Die/Site level Qualification

The metrology and inspection data are captured at site and die levels, hence more robust and effective root cause analysis requires the need for die to site qualification so that an accurate correlation between die and site level data is possible. There are two distinct steps in the proposed algorithm for die to site qualifications *(i)* compute the maximum site index on both X and Y axes and *(ii)* perform die to site qualification. The description of the variables used in the proposed algorithm is presented below (Table 5.2):

Notations	Description	Notations	Description
	x,y components to compute wafer center		Site x,y dimensions
	min, max [x,y] distance of central site		Site count along x,y axes
EBR	Edge bevel removal		Diameter of the wafer

Table 5.2 – Description of the Die/Site qualification variables

The product mask provides us a set of basic information like x and y dimensions of the site (Step-X, Step-Y), number of dies on the Step-X and Step-Y and Shift-X and Shift-Y (Figure 5.18). It is important to note that the center of the central die/site is not always the center of the wafer; hence shift (x,y) is used to

find the exact wafer center followed by X,Y (min/max) distances from this computed wafer center in the central site/die. It is the key information to accurately count the number of sites or dies along X+, X-, Y+ and Y- axes. The step-1 to step-3 in the algorithm corresponds to these computations. The counted sites and dies are further used to find the X and Y site components for each die as per step-4 and step-5 in the algorithm.

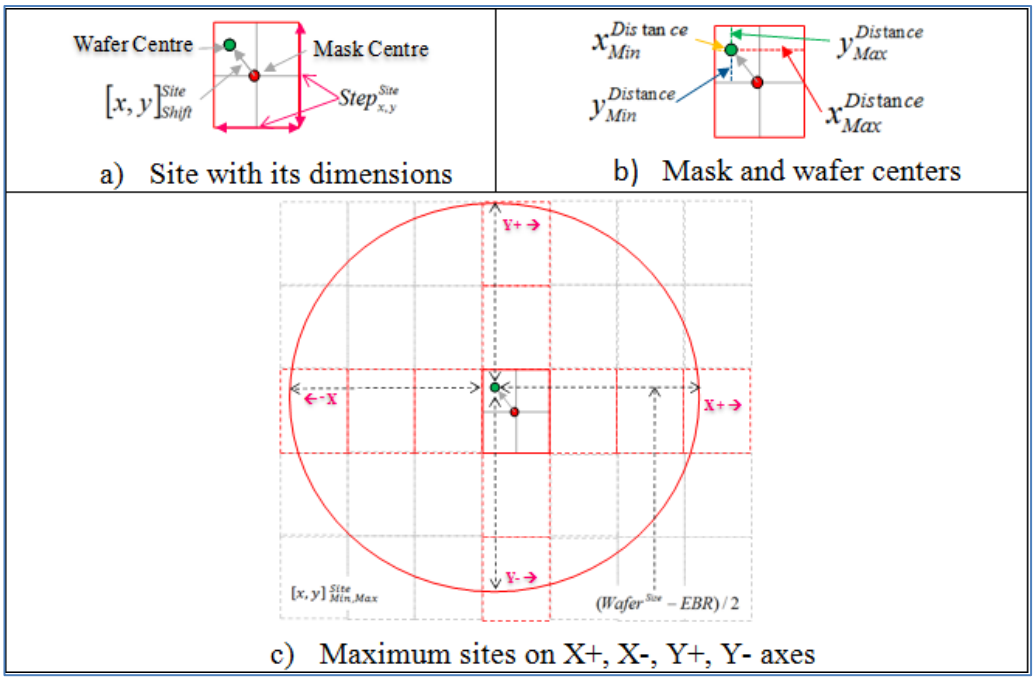


Figure 5.18 – Die/Site qualification, mask and wafer center and site counts

Step-1: Compute the center of the wafer in site source

If then

else

Step-2: Initialize → 0

Step-3: Compute

→ While ,

,

→ While ,

,

→ While ,

,

→ While ,

,

Step-4: Die to Site Qualification for X component

Update

Step-5: Die to Site Qualification for Y component

Update

The **E**dge **B**evel **R**emoval (EBR) plays a very critical role during production and is defined as the outer most area on the wafer surface which is not used during production, so this area must be excluded prior to any computations. It can be seen in the Figure 5.19 below as the distance between the outer solid and dotted lines. The above algorithm takes an assumption that source and target wafers are already mapped and aligned to the 1st reference frame with notch position at bottom based on our model as presented in section 4.1. To validate the proposed die/site qualification algorithm, we start with an example (Figure 5. 18) where site (x, y) and shift (x, y) have the coordinates (10, 15) and (3.5, 6.5) respectively. The wafer under consideration has a 300 mm diameter and 3mm (EBR) edge base removal. The shift (x,y) is used to compute the center of the central die/site from the actual wafer center; hence the wafer center is simply defined by shift(-x,-y) as (-3.5,6.5). This wafer center is used to compute the minimum and maximum distances along x and y axes as Min(x)=1.5, Max(x)=8.5, Min(y)=14 and Max(y)=1. The number of sites along x and y axes are computed with reference to the wafer center as per step-3 of the algorithm.

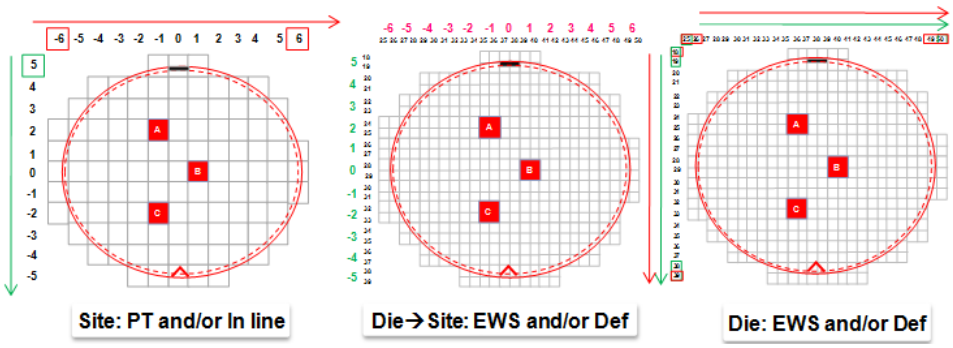
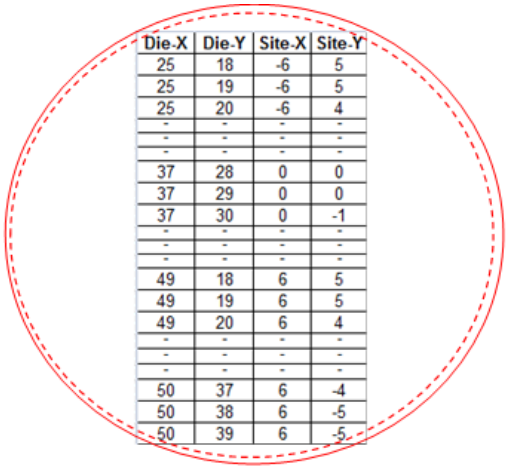


Figure 5.19 – Die/Site qualification example

We use double loop to assign site x and y components to each die in the target wafer data. The full map die/site qualification from the proposed algorithm is presented in Figure 5.20.



Die-X	Die-Y	Site-X	Site-Y
25	18	-6	5
25	19	-6	5
25	20	-6	4
-	-	-	-
-	-	-	-
-	-	-	-
37	28	0	0
37	29	0	0
37	30	0	-1
-	-	-	-
-	-	-	-
-	-	-	-
49	18	6	5
49	19	6	5
49	20	6	4
-	-	-	-
-	-	-	-
-	-	-	-
50	37	6	-4
50	38	6	-5
50	39	6	-5

Figure 5.20 – Full map Die/Site qualification example

5.5 TEST STRUCTURE POSITION BASED MAPPING AND ALIGNMENT (SPM) MODEL

Let us review the structure of the wafer, sites and dies as presented in Figure 5.21. A site is referred by its respective x and y coordinates and can contain multiple dies defined by product mask. The dies itself are also referred by x and y coordinates. The reference coordinates x and y for the site and die as presented in the figure appears normalized as 0^{th} row is always at the center but in actual practice it is not the case. Site A is expanded in the Figure 5.20 and it can be noted that **P**rocess **C**ontrol **M**onitors (PCM), **P**rocess **C**ontrol **S**tructure (PCS) and **P**rocess **M**onitoring **B**ox (PMB) test structures are placed on horizontal/vertical scribe lines and inside the dies as well. The flat and notch position are commonly used to align the wafer during production processes and as we said earlier the orientation of the test structure defines the rotation of the wafer prior to metrology steps.

There are thousands of test structures on a single wafer and are used for different types of inspections e.g. PT, Inline, Defectivity, EWS etc. It is also seen that multiple inline measurements must be mapped to compute geometric specifications which can be further correlated with PT, EWS or defectivity data to find root causes and drifts. The site and die level mapping, alignment and qualification Model (MAM) is provided which exactly fulfill the need for R&D engineers. We have discussed above that test structure might be placed in horizontal or vertical scribe lines and inside dies. It is a fact that all the tests before final test are performed on these test structures, hence it is highly likely possible that die/site level mapping might not provide accurate results because of the fact that test structures used to measure the parameters are placed at distance. The newly emerging spatial variations can be more accurately modeled if mapping and alignment is based on the shortest distance between test structures.

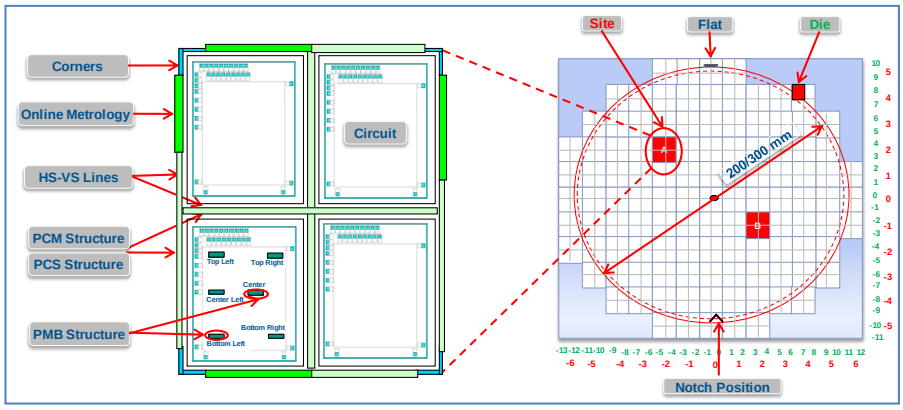


Figure 5.21 – Structure f wafer, sites and dies

5.5.1 SPM (spatial positioning) Problem (Source/Target → 1*1) Context

At present test structure position based multi-source correlation is not possible because of the fact that test structure positions are not available in the databases. Upon further investigation it was found that source files either directly or through an alternative data model can be effectively used for the position based mapping and alignment. The problem description for one source and one target is presented in the Figure 5.22.

The Figure 5.22(a) shows a simple site in the wafer with both infield and scribe lines test structures. The 5.22(b) shows the source and target parameters that must be mapped based on the site level mapping and alignment. The position of the test structures with reference to the center point is known; hence the shortest distance between source and target parameters can be easily computed as presented in 5.22(c). In an anticipation of alternative mapping, a step-circle is drawn and distance between source parameter of source site 'E' is computed with all the sites in step circle as shown in 5.22(d) and 5.22(e). Based on the shortest distance formulae the source parameter of source site 'E' is mapped with the target parameter of site 'F'. It is more realistic in an anticipation of tracking the spatial variations emerging especially in the development of new technologies.

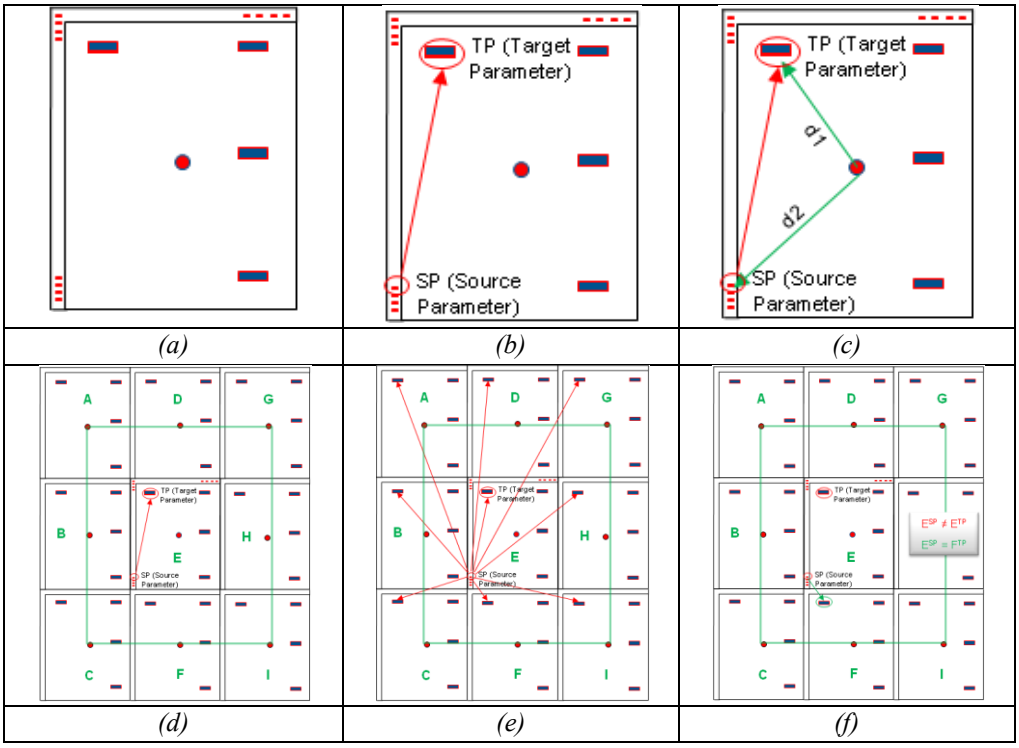


Figure 5.22 – Structure of wafer, sites and dies

5.5.2 Step-Circle based Basic Algorithm [Source/Target (1*1)] for Mapping

Let us start with the variable description used in the proposed algorithm as presented in the Table 5.3.

Notations	Description	Notations	Description
	Wafer Min/Max on X-Y axes		X,Y dimensions of Site
	X,Y coordinates of target parameter(s) w.r.t site center	Results[]	Results matrix
	Source to target distance		S → source, t → target
	X,Y coordinates w.r.t. site center of source parameter		Source or target site x and y components

Table 5.3 – Description of the Step-Circle (B) Variables

The step-circle based basic algorithm for site to site mapping is presented as under:

Step-1:

Step-2:

,

Step-3:

5.5.3 Example for Basic Step Circle Algorithm [Source/Target (1*1)]

The example starts with the presentation of basic information that center is at the center of the site and coordinates for the source and target parameters are SP (-4.5, -7.0) and TP (-3.5, 6.5) respectively (Figure 5.23). The step x, y are taken as 10 and 15 respectively and we focused on the source site 'A'.

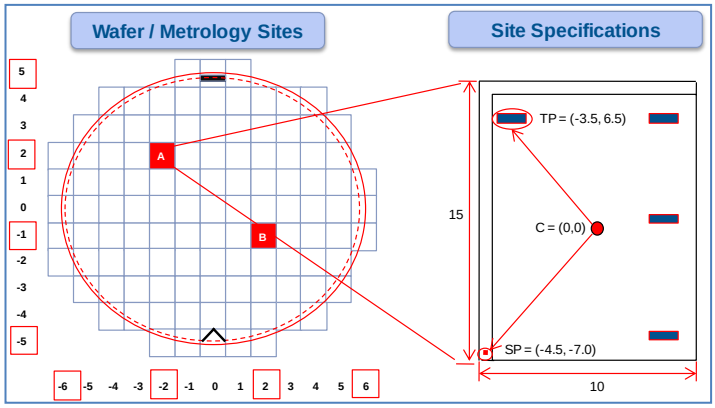


Figure 5.23 – The example of initialization variables (step-1)

The step-2 navigates through all the sites on wafer surface as source site. The two first 'for' loops generate the step-circle as presented in Figure 5.24.

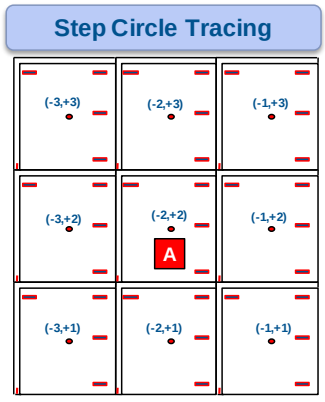


Figure 5.24 – The example of basic step-circle algorithm (step-2)

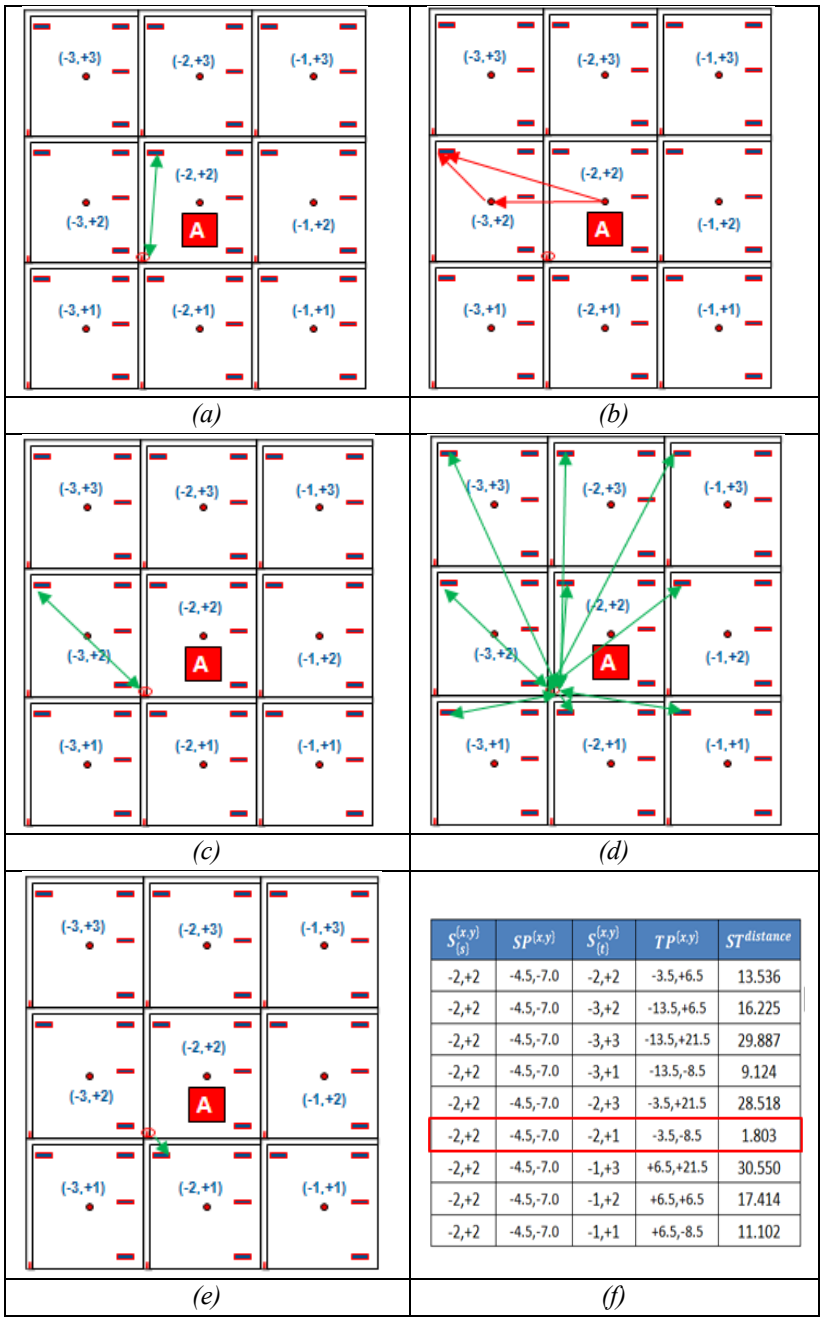
The next step is to compute the distance between source and target parameters in the source site followed by the distance computation with target parameters across all sites in the step circle. The sample computation of distance between source and target parameters for the source site A (-2, 2) are presented as under:

(11)

(12)

$$\text{——————————} = 13.536 \quad (13)$$

The graphical description of this computation is presented in the Figure 5.25(a). The computation of source and target distance between source site and target step-circle sites is presented in Figure 5.25(b), 5.25(c), 5.25(d) and 5.25(e). All computed distances are presented in the Figure 5.25(f). The Step-3 of the proposed algorithm finds the shortest distance resulting in mapping of the source site (-2, +2) and target site (-2, +1).



$S_{(s)}^{(x,y)}$	$SP^{(x,y)}$	$S_{(t)}^{(x,y)}$	$TP^{(x,y)}$	$ST^{distance}$
-2,+2	-4.5,-7.0	-2,+2	-3.5,+6.5	13.536
-2,+2	-4.5,-7.0	-3,+2	-13.5,+6.5	16.225
-2,+2	-4.5,-7.0	-3,+3	-13.5,+21.5	29.887
-2,+2	-4.5,-7.0	-3,+1	-13.5,-8.5	9.124
-2,+2	-4.5,-7.0	-2,+3	-3.5,+21.5	28.518
-2,+2	-4.5,-7.0	-2,+1	-3.5,-8.5	1.803
-2,+2	-4.5,-7.0	-1,+3	+6.5,+21.5	30.550
-2,+2	-4.5,-7.0	-1,+2	+6.5,+6.5	17.414
-2,+2	-4.5,-7.0	-1,+1	+6.5,-8.5	11.102

Figure 5.25 – The example of basic step-circle algorithm (step-3)

5.5.4 SPM (spatial positioning) Problem (Source/Target → 1*n) Context

In this section, the effectiveness of step-circle based basic algorithm will be analyzed in the context of one source and multiple target parameters. The example to demonstrate the level of complexity by increasing multiple targets is presented in Figure 5.26. It is evident from the figure that the coordinates of the source and target parameters are known with reference to the center of the site. The number of computations significantly increases even for a single site.

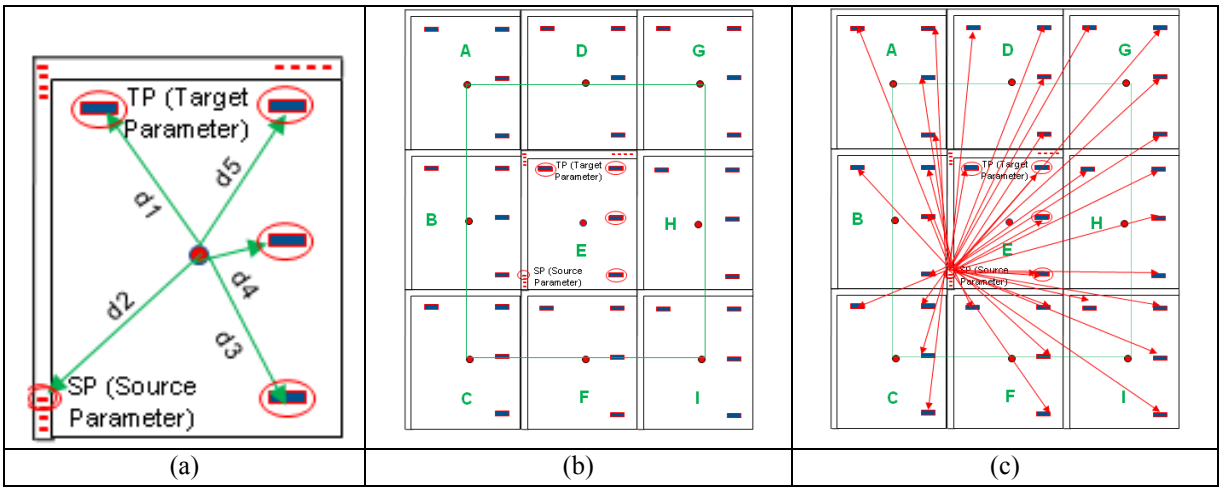
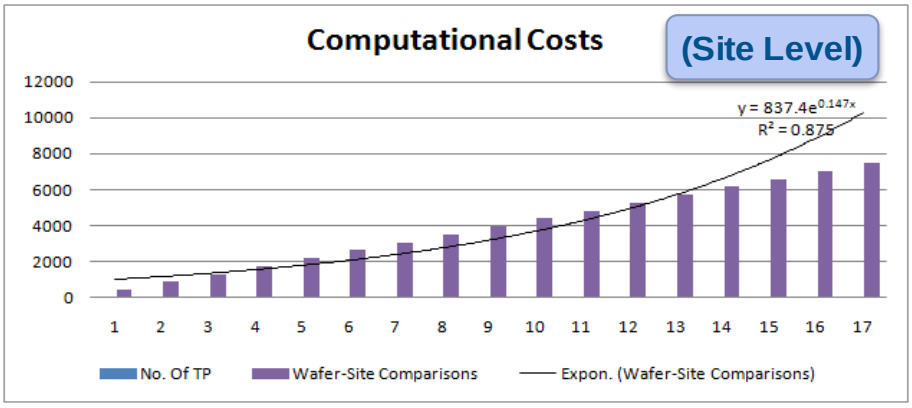
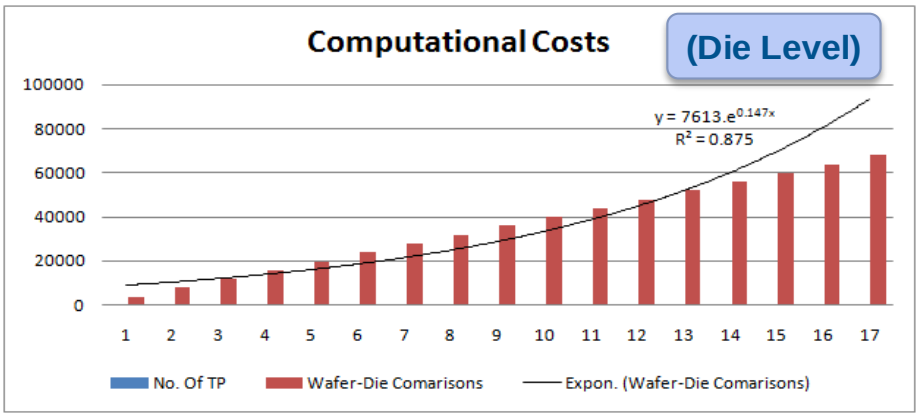


Figure 5.26 – The SPM problem source/target (1*n) context

The computational costs associated with increasing number of target parameters are presented in Figure 5.27(a) and 5.27(b) for both site/die levels which strongly highlight the need for optimization.



(a)



(b)

Figure 5.27 – Computational costs with increasing target parameters

5.5.5 Optimized Step-Circle Based Algorithms (i*n) Problem

The variable declaration in the table 5.3 for the basic algorithms hold valid in this section, however additional variables are presented as under in Table 5.4.

Notations	Description	Notations	Description
	Distance from the source parameter to the top left (TL), bottom left (BL), top right (TR) and bottom right (BR) corners		Direction of traversing

Table 5.4 – Description of step-circle (O) algorithms

The proposed algorithm is presented as under:

Step-1:

Step-2:

Step-3:

Step-4:

Step-5: Compute Distance

,

5.5.6 Example for Optimized Step Circle Algorithm [Source/Target (1*n)]

We start with the same example as presented in Figure 5.22. The proposed algorithm is functionally similar to basic step-circle algorithm; however it is more intelligent and finds the potential direction of the likelihood of target parameters. The step-2 in proposed algorithm computes the coordinates of the four corners of the source site a.k.a. TL (top left), TR (top right), BL (bottom left) and BR (bottom right). Once these coordinates are available, the distance from our source parameter to these corners is computed. These steps can be seen in the presented Figure 5.28(a), 5.28(b) and 5.28(c). The shortest distance is computed among them which define the directionality for the step-circle. The step-3 in this algorithm computes directional

semi step-circle based on the identified directionality as shown in the Figure 5.28(d) and 5.28(e). The distance computations from the step-4 and step5 clearly result in the 50% reduction in computational efforts.

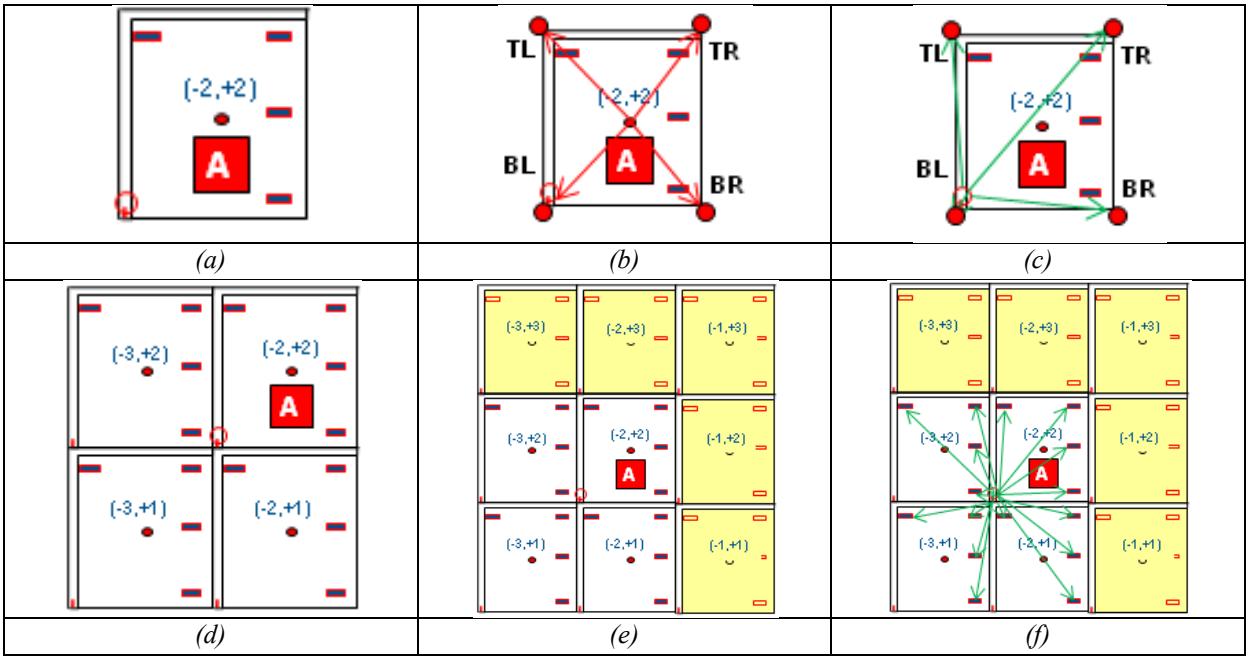


Figure 5.28 – Computational costs with increasing target parameters

5.6 Data Model for Position Based Site/Site Mapping

To support an automated process for position based site/site mapping, a data model is presented which is populated using the existing source files generated by inspection equipments, test program specifications and test structure layout files. The data model is presented in two parts (Figures 5.29 a & b) for the purpose of description and understanding.

Let us start with the presentation of data model proposed to support PT-Inline (site level) and EWS-Defectivity (die level) data based spatial variation modeling. We can see that a lot is composed of wafers and every wafer is further composed of site level wafer map (site_full_map class) that includes the sites used for the PT and Inline measurements. An association class die_full_map provides die level coordinates for each wafer against a wafer and a site level wafer map to ensure die to site qualification for PT and Inline measurements against EWS and Defectivity data. Further, it can be seen that each lot is composed of 1 or many products (multi product wafer), however each product is associated with a mask set where we find all the key information to compute the actual center of the wafer and its die and site level coordinates with TL (top left), BL (bottom left), TR (top right), BR (bottom right) information. The x,y coordinates and associated site nos. as found in the measurements class might be different based on the test type, hence mes_coordinates class and test_type classes support the identification of actual x, y coordinate and generic site numbers for PT-Inline measurements. Measurement class holds the PT-Inline measurements performed against the metrology within the wafer_steps class. The defects and yield classes hold the EWS and Defectivity measurements (die level) associated with the wafer steps and the product tests where product tests are defined at the product level. To compute the positions of the test structures on the wafer for spatial variation modeling the test_structure_map class provides the accurate position. One of the biggest issues in the spatial variation modeling is to find the parameters being measured with the specific test structures so, a test_program_specs class is added the holds information of all the EWS parameters along with their specification limits and bin definition in the bin_defination_program class for each test structure defined in test_structure_map class. We have also presented class_lookup and defects classes. The defectivity measurements class holds the list of defects to be measured against defectivity steps in wafer_steps class.

The proposed data model fully supports the spatial variation modeling on the wafer surface for PT, Inline, EWS and defectivity measurements data.

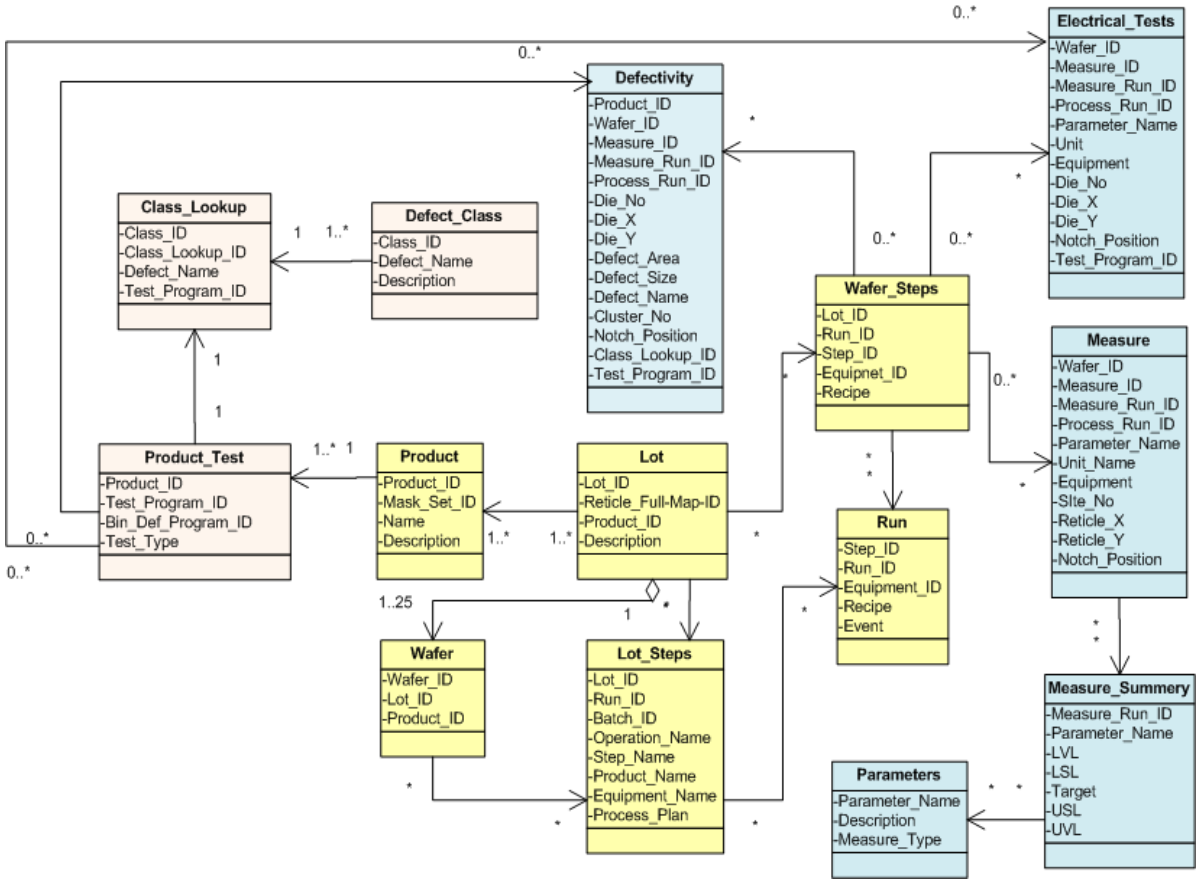


Figure 5.29(a) – The data model for SPM model

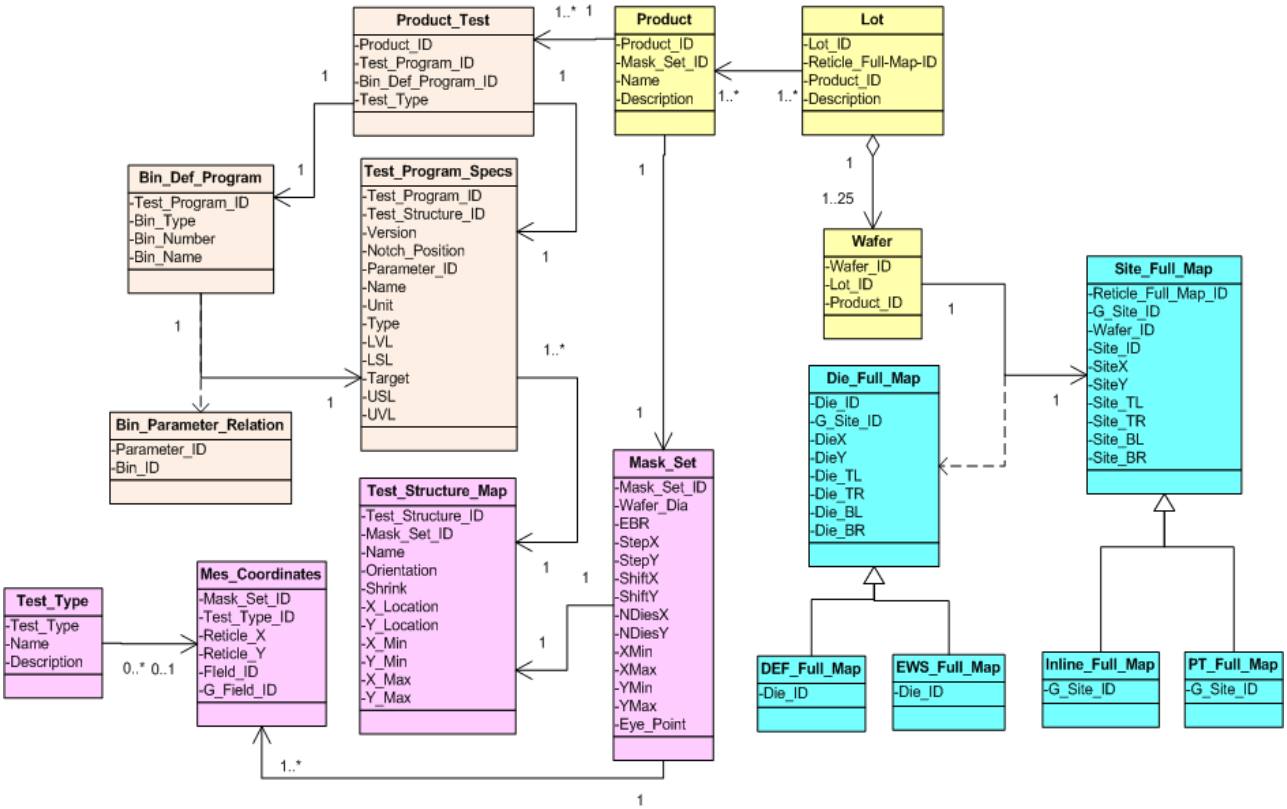


Figure 5.29(b) – The data model for SPM model

5.7 RESEARCH SCHEMATIC AND ADVANCEMENTS (*MAM AND SPM MODELS*)

The research schematic and advancement is presented in Figure 5.30. The MAM (SC1) and SPM (SC2) models are key generic contributions made in this thesis which lead us towards a shift from data driven ineffective DFM efforts towards information and knowledge driven efficient DFM efforts. The proposed generic scientific contributions address the cyclic failure modes and root causes identified in the step-3 of *i*-FMEA methodology presented in chapter-4. Our proposed *i*-FMEA methodology has significant advantages as it helps in identifying the

Cyclic failure modes and root cause which are repeated until and unless they are fixed with generic solutions. The proposed MAM and SPM models ate the partial contributions in solving the issues associated with data extraction, mapping and alignment during technology alignment and adoption efforts.

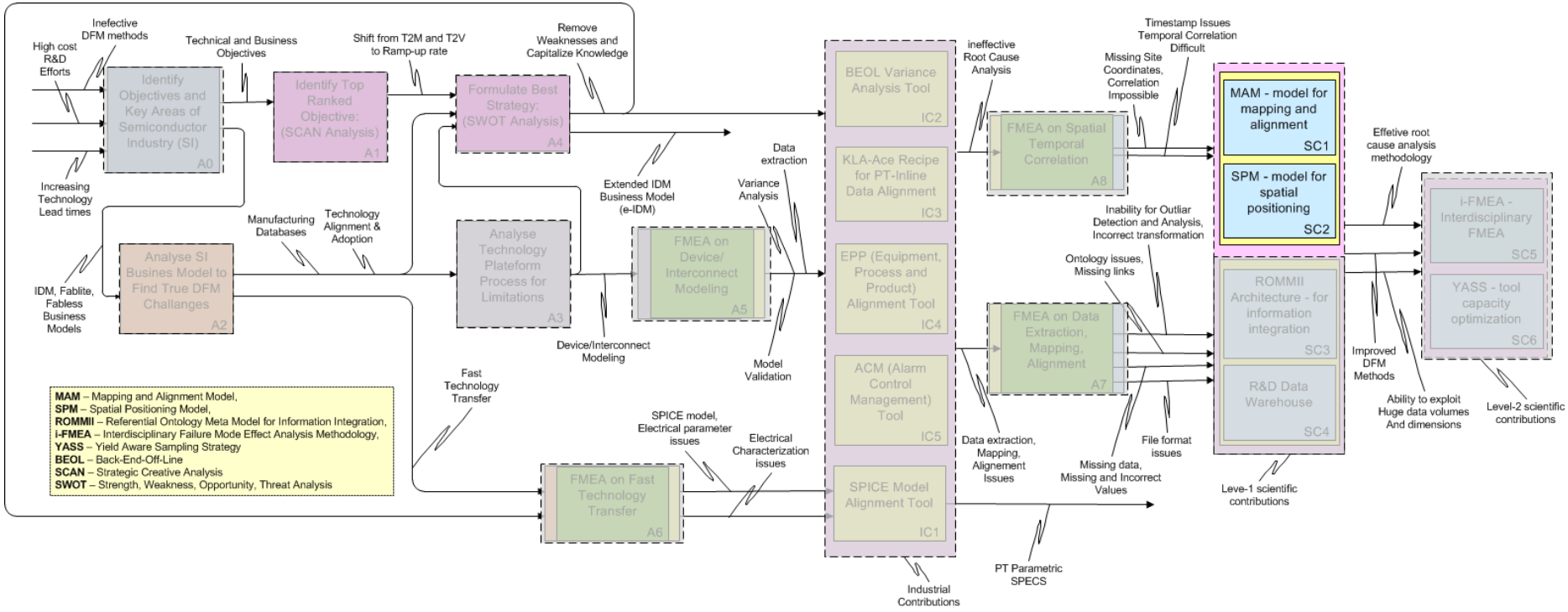


Figure 5.30 - The research schematic and advancement with MAM and SPM models

5.8 SUMMARY AND CONCLUSIONS

In this chapter, we have analyzed the traditional interconnect modeling process being one of the critical processes in technology alignment and adoption initiatives. The key objective is to analyze the effect of cyclic failure modes and causes as: *(i)* data extraction, *(ii)* pre-processing and *(iii)* multi-source root cause analysis, so that generic R&D solutions can be proposed. The BPR analysis on the interconnect modeling process has resulted that the cyclic failure modes and root causes have a significant influence on the model correction and interconnect stack definitions steps that lead to an exponential increase in technology lead times and associated costs. We have proposed a revised interconnect modeling process by adding a knowledge base, which provides an initial feedback for the first time correct stack definition followed by quick model correction based on effective root cause analysis.

The effective root cause analysis is the key to improve existing ineffectiveness in the DFM methods and is defined as the ability to analyze data to find answers against drifts and/or variations. These drifts and variations are further classified as systematic or random for their transformation into rules and/or models. At present, engineers are facing huge difficulty in the data mapping and alignment to perform multi-source analyses based on fact that the measurements result in varying coordinate system due to test structure orientations. The wafer is rotated prior to the metrology steps and its position is monitored using notch or flat positions. The varying coordinates issue is further complicated with different metrology reference frames. These issues have a significant impact on stack definition, model correction and validation steps.

We have proposed mapping and alignment (MAM) model and spatial positioning (SPM) model. The MAM model is capable of performing site/site, die/die and die/site levels data mapping and alignment. It is highly important in correlation analysis because inspection data are captured at different levels (PT and Inline are captured at Site and defectivity and EWS data are captured at die levels). The proposed MAM model enables multi-source root cause analysis resulting in quick knowledge capitalization. We know that during production the wafers undergo multiple inspection steps and all the tests are performed on test structures placed inside the field and/or scribe lines. To accurately capture the spatial variations we have proposed SPM model, which enables the mapping of source and target parameters based on distance between tests structures used for measurements. The analyses results are more accurate and help us in capturing the spatial variations emerging due to miniaturization. The MAM model is suggested for mapping and alignment during technology adoption efforts whereas SPM model is a best fit during technology alignment initiatives.

We concluded the chapter with a data model, which is filled using source metrology, test program specifications and layout files, provided by the **R**eticule **A**ssembly **T**eams (RAT). The data model provides an efficient way of internal processing prior to data mapping and alignment. The solution for missing data dimensions and continuous data models restructuring for the successful evolutions is provided in next chapter, which results in a generic solutions, focused on removing weaknesses and improving DFM effectiveness. It enables an effective multi-source root cause analyses for the R&D and/or product/process/equipment engineers.

Chapter 6: ROMMII and R&D Data Model for Information Integration[15]

In this chapter, the cyclic failure modes and root causes associated with multi-source data extraction function are addressed. The key causes are identified as *(i)* unstructured data model evolution, *(ii)* ontology issues and *(ii)* data retention periods. The ROMMII platform is presented to address unstructured data model and ontology issues where the R&D model provides an efficient solution to data retention issues easing production data sources. The ROMMII platform enables us to exploit huge data volumes and dimensions by removing *(i)* model inconsistencies, *(ii)* pre-failure assessments to avoid extraction and analysis utilities failures, and *(iii)* information diffusion to perspective end users about the inclusion of new data dimensions. The R&D data model resolves the issue of different data retention periods in existing production sources and proprietary constraints associated with multiple production data sources.

Contents

[15] Shahzad M.K., Hubac S., Siadat A., Tollenaere M., ROMMII (referential ontology meta model for information integration) Architecture for Dynamic Restructuring of DWH data models, 12th European APCM Conference, Grenoble France, 2012

6.1 INTRODUCTION

Looking at the historical evolution of data needs by R&D engineers as presented in Figure 6.1, it can be seen that in past engineers have been complaining about few data volumes and dimensions for their inability to capitalize knowledge. The memory and computational costs were high in comparison to the performance and efficiency. However recent revolutions in the IT technologies have resulted in continuous reduction in the memory and computational costs, and increase in the computation power. It is because of this fact that today we are able to store large data volumes and can exploit them at high efficiency BUT R&D engineers are still complaining and this time it is not about the volume and dimensions but their inability to exploit these huge volumes and dimensions. As a consequence, the productivity of our R&D engineers is low which impacts the competitiveness of SI resulting in extended technology alignment and adoption lead times, so it is necessary to work on solutions to facilitate engineers in their efforts to quickly transform the data into information and then knowledge for technology lead times and costs reductions.

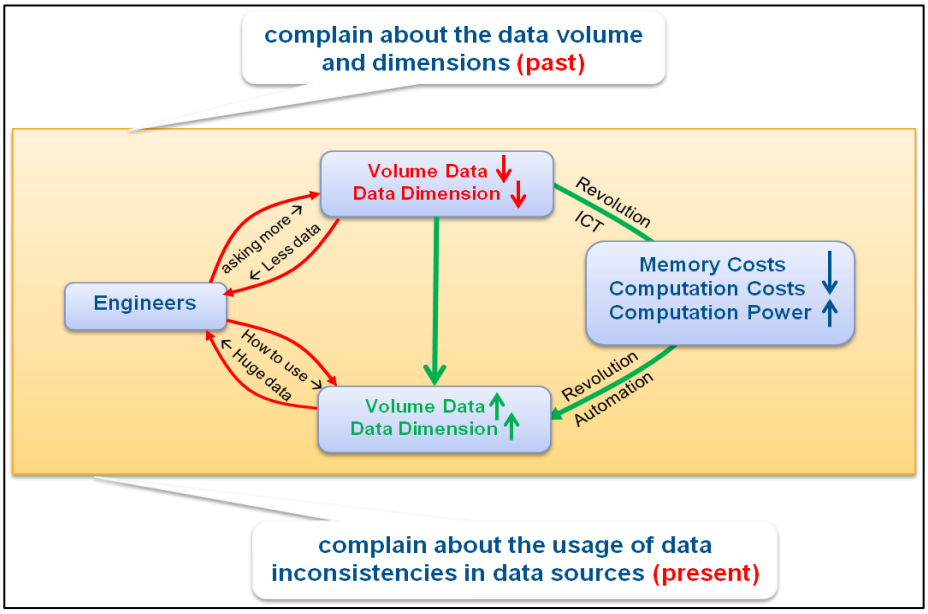


Figure 6.1 – Existing data extraction and analysis challenges

6.2 HISTORICAL EVOLUTION FROM UNSTRUCTURED TOWARDS STRUCTURED DATA STORAGE

The data is defined as the collection of raw facts and figures and it is processed in a specific context by engineers, middle manager and managers to generate information. The data is collected across the production process and is classified as the *(i)* process and *(ii)* contextual data (Figure 5.6) where the process data is used to control the product quality and the contextual data serves diverse purposes ranging from planning to equipment control and engineering. The database is defined as a structured way of storing the data and information for an efficient response through a query language; however the evolution of data storage architectures can be broadly classified as *(i)* flat files *(ii)* relational and *(iii)* dimensional eras as discussed in the following sub-sections.

6.2.1 Flat Files Database Era (1890 till 1968)

The flat file databases are defined as the files containing delimited records of varying lengths in the form of rows and columns with no relationship or link between records or fields. The concept of flat files can be traced back to Herman Hollerith who conceived the idea of storing census data in holes punch-card for the tabulation purposes; hence, the census of 1890 was completed by the US Census Bureau with punch-card systems. Herman Hollerith left US census bureau in 1896 and started his company that is known today by the name of IBM [Blodgett and Schultz1969]. The most common data storage medium was the magnetic tape

with sequential, indexed and random access methods. The third generation programming languages (e.g. COBOL, BASIC) were used to write programs to read, write, delete and update the data in these file. These file systems are still used today for the system configuration files but they are not being used for commercial purpose because of the limitations as *(i)* weak security, *(ii)* no sharing, *(iii)* likely duplication and *(iv)* high maintenance costs for data consistency and access controls [Impagliazzo, 2012].

6.2.2 Non-Relational Database Era (1968-1980)

In this era the emphasis was to overcome the limitations associated with the initial file systems and IBM again took lead and proposed the first non-relational databases management system called IMS (information management system). The IMS is a hierarchical database with a tree like storage architecture having parent-child as one-to-many relationship. It was developed in a joint project with Rockwell and Caterpillar (1966-1968) to manage large bill-of-material (BOM) for the Apollo space shuttle. The network database model has an advantage of efficient searching over flat file databases. It also results in less data redundancy, security and integrity; however the biggest disadvantage is the difficulty in implementing the many-to-many relationships [Blackman, 1998].

The second major contributions in this era are the network database models the *(i)* CODASYL DBTG and *(ii)* IDS model. The CODASYL (conference on data systems languages) also credited for the first general purpose business programming language (COBOL) in 1968. The database task group (DBTG) proposed the specifications for the network database model in 1969 along with interfaces in the COBOL language for data definition (DDL) and data manipulation languages (DML). The network database model (IDS at Honeywell) uses directed acyclic graph with nodes as records and edges as relationships [Bachman, 1973]. It has the similar advantages like IMS with many-to-many relationship; however, it lacks structural independence [Blackman, 1998].

6.2.3 Relational Database Era (1970 till present)

The relational database is the first major breakthrough of the century that has changed the way in which data could be stored and used for the decision support systems. There is no doubt that the rich database efforts in terms of hierarchical and network based databases resulted the relational databases. The idea of relational database model was initially proposed by Ted Codd at IBM in 1970 [Codd, 1970]. This model provides advantages over all above discussed database models as *(i)* no redundancy, *(ii)* security and integrity, *(iii)* all types of possible relations and *(iv)* the economies of scale whereas, the only associated advantage is the size and cost of the DBMS. It has also led the emergence of object oriented database management systems (OODBMS) around 1985; however, it has resulted in little success because the conversion cost is too high for the billions of bytes of data [Blackman, 1998]. The examples of the RDBMS in this era are *(i)* Ingress which ended into Informix and *(ii)* System R by IBM which resulted in DB2.

6.2.4 Dimensional Database Era (1990 till present)

The revolution in the database technology for efficient data storage and subsequent retrievals was supported with the decreasing memory and computational costs. It has resulted in the availability of huge volume of data. The drawback of size associated with the relational databases has resulted in the inefficient data exploitation with rise of interest in the methods for quicker response queries. The dimensional database era started (1990) with the emergence of enterprise resource planning (ERP) and management resource planning (MRP) concepts. The biggest advantage is that data is stored in multi-dimensional cubes which results in quicker aggregation and faster response to the end user queries [Blackman, 1998].

The above discussion on the evolution from flat-files towards dimensional databases indicates that the driving force behind all these efforts is the ability to efficiently store and exploit data resources with the security and integrity. In today's competitive environment where knowledge is the key for success, it is needed to transform data into information and knowledge so that varying processes and business strategies can be reformulated for competitive gains. We hypothesize that the principle objective is data/information

integration for the knowledge capitalization. Now we shall focus in detail on the dimensional database systems in order to assess its appropriateness for our proposed e-IDM business model (section 3.7, chapter-3).

6.3 EXISTING DATA/INFORMATION INTEGRATION SYSTEMS

The existing information systems used for the data/information extraction/integration efforts are classified as online transaction processing (OLTP) and online analytical process (OLAP) systems. The OLTP systems are supported with the relational databases and are focused on efficient insert, update and delete queries whereas OLAP systems require DWH/DM architecture with emphasis on data retrieval. It is also important to note that all business intelligence (BI) tools require the ability to quickly aggregate, dice and slice data for decision making purposes using OLAP cubes which are built from the DWH/DM databases. A brief comparison is presented below for reference:

OLAP Systems (DWH Database)	OLTP Systems (Relational Databases)
It uses dimensional database structures	It uses normalized databases
Designed for analysis of business measures by category and attributes	These systems are design for real time (transactional) business operations
It supports few concurrent connections	It supports multiple concurrent connections and database operations
The consistent and validated data is bulk loaded for large, complex and unpredictable queries	The database is optimized for the validation of data and transactional processes
The query response is optimized with multiple indices with few joins	It uses few indices as the emphasis is on the accurate transactions with many joins
It follows periodic updates through extraction, transformation and loading (ETL) routines with few frequent modifications	It requires frequent modifications
It operates with huge data volumes and large data aggregations	It operates on small data volumes and few aggregations

Table 6.1 – OLAP vs. OLTP Systems

The OLAP systems are further classified as MOLAP, ROLAP and HOLAP. Let us briefly review these systems:

- **MOLAP:** The MOLAP based BI tools generates cubes which are aggregated from multi-dimensional database (DWH). These tools generate quick reports because data is already pre-aggregated. The MOLAP cubes are disconnected with the underlying DWH, hence the drill down is not possible.
- **ROLAP:** The ROLAP based BI tools also generate cubes from DWH, but they are supported with underlying relational database structures. It allows end users to perform drill down operations and accurately search the relevant information.
- **HOLAP:** It overcomes the disadvantages associated with both MOLAP and ROLAP systems as discussed above and provides an aggregated OLAP cube and access to the transactional data from the DWH databases.

It is evident from the above discussion that DWH and DM databases are the core of the BI tools. These tools are subjectively focused on converting data into information and knowledge for quick decision making. The OLAP cube along with drill down features facilitates the engineers to find most relevant and accurate information. We shall analyse in depth the DWH and DM architecture in the next section; however, it can be easily concluded that the relational databases can be equally used for OLAP like features but at the cost of

performance and efficiency. The DWH and DM can be implemented using relational database management systems (e.g. Oracle, SQL Server) or purely dimensional database management systems (e.g. TeraData). The relational database systems provide additional utilities to build OLAP cubes whereas dimensional database systems provide an edge on relational databases in terms of query performance and ability to treat huge data volumes.

6.4 DWH-DM: INFORMATION INTEGRATION AND BUSINESS INTELLIGENCE PLATFORM

The growth in the availability of huge data volumes and technologies to store and process them has resulted significant challenges in the integration. The DWH and DM databases as discussed above provide us an opportunity to efficiently store and exploit the huge data volumes. The latest BI tools use the DWH and/or DM architecture to quickly respond to the varying information needs of R&D engineers. The DWH and DM are defined as under:

- **DWH:** The DWH is a subject oriented, integrated, non-volatile and time-variant collection of data to support management decisions [Inmon, 2005]. It is a relational database that uses ETL routines to populate the DWH with historical data which is collected and validated from multiple data sources. It takes off the processing load from the transactional databases (OLTP) for in-depth but computationally expensive analysis.
- **DM:** The DM can be simply defined as a smaller version of a DWH with the data from one source e.g. sales. The data stored in the data marts is highly or partially summarized.

Let us start with the basic definitions and concepts prior to discuss the new database architecture for DWH and DM.

6.4.1 Basic Definitions and Concepts

Let us start with basic concepts and definitions to better understand the problem and further discuss the topic:

- **Facts:** It is a type of table in the dimensional database which includes two types of columns *(i)* facts (quantitative measures) and *(ii)* foreign keys to dimension tables. The facts are numerical measurements of a certain business process and they are aggregated in the OLAP cubes e.g. aggregated sales volume against country, region or city.
- **Dimensions:** The dimensions are the tables that store records related to the particular dimension. These are the attributes by which the facts in the fact table are grouped. The dimensions that change over time are referred as slowly changing dimensions and can be treated by *(i)* overwriting the old value, *(ii)* adding a new column and *(iv)* adding new row and version. The dimensions are classified as confirm dimensions if it is linked to multiple fact tables whereas the single dimensions, with few attributes having yes/no values, are referred as junk dimensions. They are grouped into single dimension to reduce number of referred dimensions in respective fact tables.
- **Attribute:** It refers to the fields within fact and dimensions tables.
- **Hierarchy:** The dimensions are further decomposed into sub-dimensions and it is referred to as dimensional hierarchy e.g. time dimension (year, quarter, and month) and geographic dimension (country, sales region, state, city and store).
- **Drill up/down:** The drill up/down are specific operations applied on the OLAP cube where the data is presented at higher or lower hierarchies based on the dimensions. It requires either ROLAP or HOLAP cubes (see section 6.3).
- **OLAP Cube:** It is defined as a multidimensional representation of data from the fact and dimension tables. The internal storage of OLAP cube is different but efficient as show below.

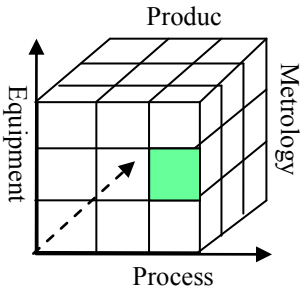


Figure 6.2 – OLAP cube architecture

- **Slicing and Dicing:** The slicing operation is defined as the process of retrieving data from OLAP cube by filtering it on a given dimension whereas dicing operation filters the original cube data across all dimensions as show below.

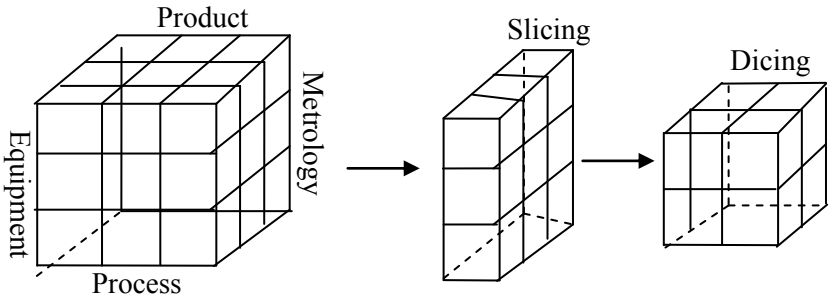
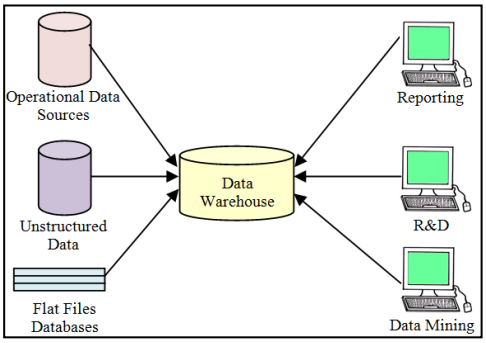


Figure 6.3 – Slicing and dicing operations on OLAP cube

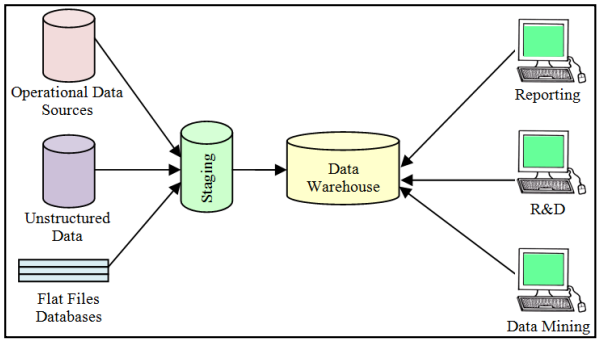
6.4.2 The DWH Architectures, Models and Schemas

In order to better understand the advantages associated with the DWH and DM databases, it is important to discuss its potentially possible architectures, models and schemas as under:

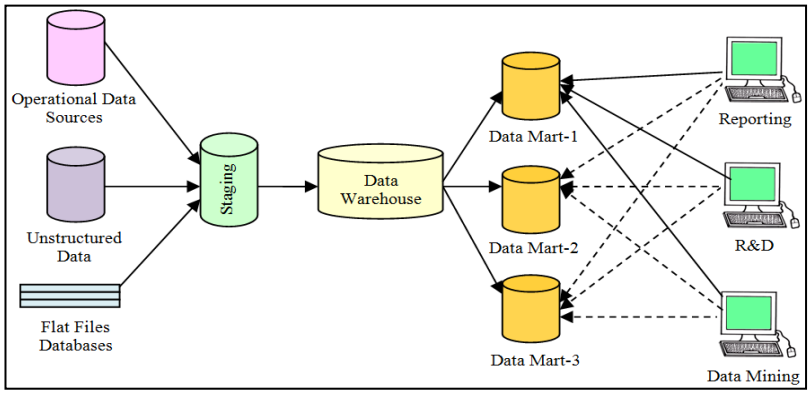
- **DWH Architectures/Frameworks:** The most commonly used DWH architectures are as *(i)* basic DWH architecture (Figure 6.4a), *(ii)* DWH with staging (Figure 6.4b) and *(iii)* DWH and DM with staging (Figure 6.4c). In the basic DWH structure, the data from different data sources is directly moved to the data warehouse through ETL routines and end users are provided with direct access over the DWH through OLAP based data extraction and analysis utilities. The disadvantage associated with this architecture is the inconsistencies in the ETL routines; hence, data validation could raise serious issues during subsequent analysis by end users. The staging is a temporary storage which is used for the validation of data prior to its permanent storage in the DWH; hence, it overcomes the drawbacks associated with the basic DWH architecture (Figure 6.4 b & c). The typical DWH architecture being used in most of the BI tools, where data from the DWH is further aggregated to the DM and end users are provided with access to these data marts for efficient query responses (Figure 6.4c).



a) Basic DWH Architecture



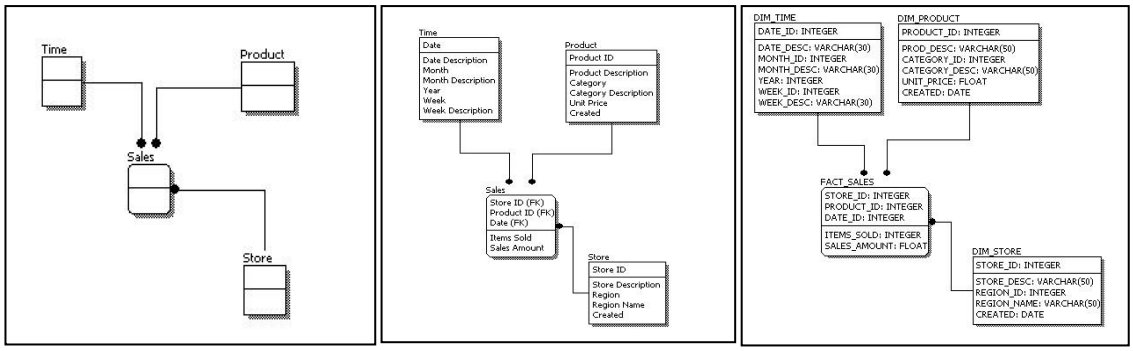
b) DWH with Staging Architecture



c) DWH and DM with Staging Architecture

Figure 6.4 – Principle DWH architectures and frameworks

- **DWH Models:** The DWH follows three modelling levels as the *(i)* conceptual, *(ii)* logical and *(iii)* physical models (Figure 6.5). The conceptual model provides the highest level of abstraction where only names of entities and their relationships are considered. The logical level in comparison to conceptual level describes the system in much more detail; however, the physical implementation is still an abstraction. It includes identification of attributes and primary and foreign keys followed by normalization process. The physical model takes into account the target database management system (Oracle, SQL Server, MySQL, Access etc.) constraints and data types and domains are detailed for each attribute. The physical model can be demoralised at this level based on the user requirements.



a) Conceptual b) Logical c) Physical

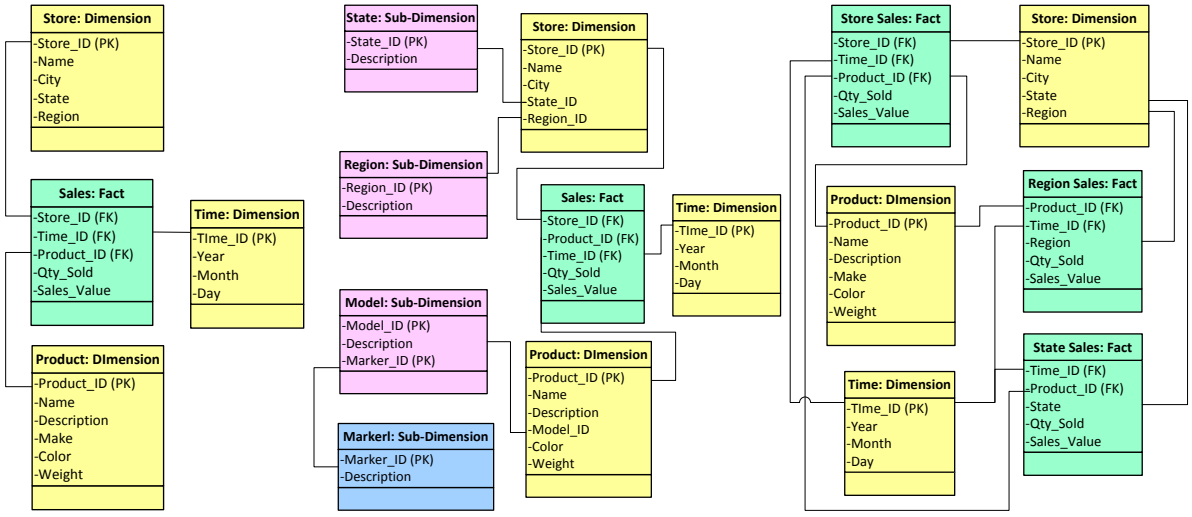
Figure 6.5 – DWH data models

A tabular comparison of three levels of DWH models is presented as under in Table 6.2 that clearly represents that moving from conceptual towards physical models minimizes and finally removes abstraction.

Function	Conceptual	Logical	Physical
Entities	Yes	Yes	-
Relations	Yes	Yes	-
Attributes	-	Yes	-
Primary Keys	-	Yes	Yes
Foreign Keys	-	Yes	Yes
Data Types	-	-	Yes
Data Domains	-	-	Yes
Data Lookup	-	-	Yes
Constraints	-	-	Yes

Table 6.2 – Comparison of DWH data models

- **DWH Schemas:** The data warehouse design and development use three commonly used schemas as *(i)* Star, *(ii)* Snowflake and *(iii)* Constellation Schemas. The Star schema (Figure 6.6a) consists of fact and dimension tables; however, there are no hierarchies for dimensions or the fact tables. The fact tables do hold quantitative or additive facts because primarily the DWH is focused on aggregation on different dimensions and periods. These are huge tables with millions of records where dimensions are used to build OLAP cubes which are further sliced, dices or drilled up/down to search relevant information. The queries used within the OLAP cube are SQL queries with multiple joins. It is very important to note down that as the number of joins increases, queries performance degrades but DWH architecture has a multi-dimensional storage with multiple indices for fast query results. The major difference between Star and Snowflake schemas is that the snowflake schema allows hierarchical dimensions; however, all dimensions are linked to the central fact table (Figure 6.6b). The advantage is that multi-dimensional hierarchy results in data aggregation at multiple levels where information is sliced and/or diced with the precise granularity. The multiple fact/dimension tables' joins affect the query performance but such limitations are compensated with additional indexing. The third most commonly used DWH schema is the Constellation schema (Figure 6.6c) which is also called the hybrid of Star and Snowflake schemas. In this schema, the facts tables are divided in multiple fact tables and hierarchical dimensions are shared by these fact tables.



a) Star Schema b) Snowflake Schema c) Constellation Schema

Figure 6.6 – DWH schemas

In the above schemas, it can be seen that it is a data warehouse about sales data. In star schema, the sales fact table is connected with three dimensions as *(i)* time, *(ii)* product and *(iii)* store. The primary keys of dimensions are accumulated in the fact table where they act either as composite key or simply foreign keys. In case, if they are selected as foreign keys then the fact table ID can be added as a primary key; however, it is not necessary. In the snowflake schema, the store dimensions is further hierarchically divided into state and region sub-dimensions and similarly the product dimension follows two levels sub-dimensions as *(i)* model and *(ii)* marker. The presented snowflake schema allows aggregating sales at regional and state levels through multiple join SQL queries. The constellation schema is an improved form of Snowflake schema, that allows the dimensions to be linked to multiple fact tables and as a result sales can be aggregated at the region and state levels with more efficiency.

The above DWH schemas can be implemented in the traditional relational database management systems e.g. Oracle and SQL server. These DBMSs are designed for the OLTP systems to optimize transactional processing; however, the vendors have added additional OLAP modules that can be used to implement the DWH as a relational database. These modules provide users with interfaces that can be used

to build OLAP cubes and subsequent query processing. The newly emerged DBMS 'Teradata' is a special database system which is designed only for the DWH and DM databases; hence, it is obvious that it has significant advantages over traditional RDBMS in terms of storage and query response. The major portion of work within DWH management is to monitor the ETL routines and ensure its consistencies with the DWH data model.

A generic Meta model for the data warehouse (Figure 6.7) and its schemas is presented [Darmont, et al., 2007] as under. It can be seen that a data warehouse is composed of multiple fact tables where each fact table can have multiple dimensions. The hierarchy levels are associated with the dimensions which are shared between the fact tables. Each table has attributes (fields) and extend to tuples (records) where tuples and attributes are linked with the values. All of the above defined schemas can be easily derived from the proposed Meta model.

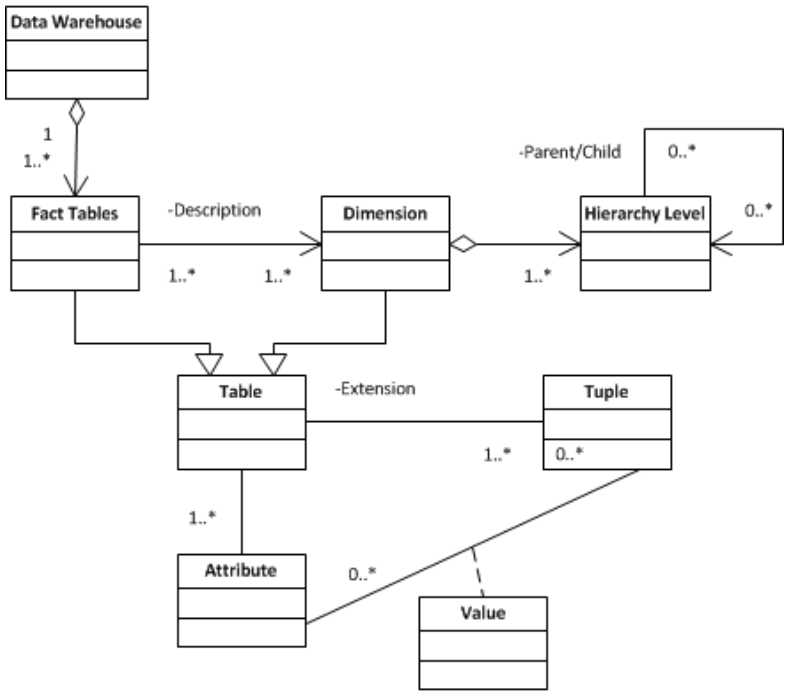


Figure 6.7 – Data warehouse Meta model for business intelligence [Darmont et al., 2007]

6.4.3 Inmon and Kimbell DWH Philosophies

William Inmon and Ralph Kimball are well known in the domain of information management for decision support and have played a pivotal role in changing the information management concept. Mr. William Inmon is known as the father of data warehousing and is credited for proposing the concept and term 'Data Warehouse' in 1991. Mr. Ralph Kimball is known as the father of business intelligence and is credited for the concept 'Data Marts', star schema and snowflake schema. These legends have contributed most to the domain of data warehousing and business intelligence tools. In order to better understand the contributions of these great people let us define the term business intelligence (BI) as BI = Inmon's Corporate DWH + Kimball's Data Marts + Data Mining. It can be said that both subjectively focus on developing means (DWH+DM) to efficiently exploit the huge data volumes to ultimately support the decision making process.

Inmon's advocates that it is necessary to start building large enterprise wide data warehouse followed by multiple data marts to meet with the customized or specific needs of different business functions within the organization. In comparison to this philosophy, Kimball proposes that we must start by building data marts to serve data analysis requirements of respective organizational business functions followed by its virtual integration for enterprise wide data warehouse. These approaches by Inmon and Kimball can be stated as the 'top down' and 'bottom up' philosophies respectively; however, both are based on the dimensional modeling.

The Kimball approach is more realistic and follows the natural sequence of enterprise information evolution where departments started modeling and managing their information needs as per business requirements and later the need to interlink these information sources resulted in the enterprise wide information systems. The Inmon's approach is difficult to follow because of the fact that it is difficult to potentially forecast the information needs in advance; hence, even if historical data/information is stored we are not sure, if it shall ever be used? In comparison, Kimball's method is more practical as it give rise to the evolution of the real information needs and once modeled as DWH and DM serve the basis for business intelligence. The only threat to this approach is that constantly changing business requirements and needs requires frequent modification in the data models which results in high level of complexity because of cross version queries. The Inmon's model in longer run requires fewer changes and can come up as a robust and reliable source for analyses purposes.

6.4.4 The DWH Challenges

The interest in DWH and BI has risen in the last decade especially in the domains with continuously growing data volumes. The organizational success lies on our ability to efficiently analyse the data volumes and formulate business strategies. The DWH/DM architectures along with three schemas enable efficient data mining capabilities to generate information and knowledge from data sources. There is no doubt that in present information era, BI tools and DWH/DM architectures are playing a critical role in the knowledge capitalization but let us look at the challenges faced by BI/DWH projects:

- The DWH project is thought to be a magic trick. It is expected to resolve all the varying information needs of the organization, which is not true. The DWH project has an average of 3 years of duration where business requirements changes significantly; hence, the success of the BI/DWH project depends on the clear identification of business objectives.
- The changing business requirements must be addressed quickly. It requires frequent changes in the DWH data model and associated ETL routines. These changes can be addressed during the DWH project development resulting in higher costs and lead times.
- These changes might emerge after the deployment of DWH project, if so then any change in the DWH model might result in severe consequences on existing historical data. The most common changes are the modifications of dimension tables which are addressed by *(i)* overwriting the old value, *(ii)* adding a new column or *(iv)* adding new row and version. The complexity arises afterwards when there are different versions of DWH data model and a cross version query arrives. The schema versioning and cross schema version queries have been very well treated in the literature [Johann et al., 2002 and Matteo et al., 2006].
- It is also highly important to analyse and assess the need for a DWH and BI. We have seen that industries are adopting these solutions without proper pre-assessments but the results are not justified against the project costs.
- Inappropriate user interfaces to exploit (slice, dice and drill up/down) the OLAP cube data by the end users might lead to the failure of a DWH project.

It is evident that the common failure arises when we fail to assess the need for BI/DWH project and the changing business requirements which lead to DWH model evolutions. These evolutions are well addressed with schema versioning and cross version OLAP queries for accurate results. The BI tools are supported with the outstanding ETL routines for the validation and transformation of input data prior to its storage into DWH. It is important to note that moderate level BI needs can be easily addressed using summary tables within the relational databases, however increasing volumes and unstructured data model evolution in the relational databases are the limiting factors which give rise to the interest in DWH projects. It is known that the relational databases are designed and developed for transactional processing (OLTP) optimization and they are subjectively focused on efficient insertion, deletion and modification of the records. These relational

databases are not built for the fast retrieval of summarized data from huge data collection; hence, even upon the implementation of summary tables, retrieval efficiency cannot be compared with DWH projects. The DWH projects are focused on fast data aggregation and analyses across multiple dimensions to fulfil the varying information needs.

6.5 PROPOSED R&D DWH MODEL

In this section, the SI scenario and its varying information needs is presented along with its growing data volumes. We shall further assess the need for DWH project and the possibility to add summary tables within existing relational databases against the growing information needs for the R&D engineers. The objective is to justify the choice with reasons to go for one of the above proposed solutions.

The SI has grown to 300+ B$ industry in last 60 years with a cumulative +ve growth rate of 8.73%. This positive growth rate suggests that we must continuously invest in the R&D efforts even in the downtimes to prepare ourselves for the good times. The objective is early penetration in the market resulting in maximum market share which cannot be achieved without robust technology. In this thesis we are focused on technology derivative/improvement efforts because new technology is developed in the technology alliance to share exponentially increasing R&D costs and reduce the lead times. Upon the transfer of this new technology into alliance partners manufacturing facilities require quick alignment and adoption followed by subsequent technology derivative/improvement alignment and adoption with local R&D efforts. Today the local R&D efforts have turned into high cost activities; hence, the root causes for this ineffectiveness are identified as *(i)* unstructured data model evolution and *(ii)* missing links between databases which restrict R&D engineers from multi-source root cause analysis. These ineffective R&D efforts are key for DFM ineffectiveness, so we need to remove these limitations.

It is important to note that both contextual and process data collected across the production line is stored in more than 6 operational data stores (ODS) which are proprietary databases. The end users and R&D engineers are provided with single-source data extraction and analyses utilities to control the respective production process to ensure product quality. The root cause analysis requires multi-source data analysis which is not possible because of the missing links between databases. The data retention in these RDBS is more than one year and they store the production as well as engineering and R&D data. R&D engineers are not able to extract data from 6+ databases because *(i)* they do not have access to those databases and *(ii)* they are not familiar with the other domains data and the extraction tool. Beside the RDBS, there are also data warehouses but data analysis utilities provided to the end users are difficult to use and often crashes against complex queries. The granularities of the stored data are not accurately captured at site/die levels; hence, die/site or test structure position based data analysis is almost impossible with the existing data resources.

The biggest issues to be addressed are the *(i)* retention period and *(ii)* multi-source data extraction issues. The RDBS and DWH are implemented using Oracle, Ingress and SQL Server; however, existing DWH do not carry data from all dimensions. There are two possible solutions *(i)* design and deploy a separate R&D relational database for analysis purpose supported with summary tables for efficient query responses and *(ii)* design and develop R&D DWH for fast data retrieval. The first option is likely possible solution but in this case summary table will not help us because during correlation analyses, accurately mapped and aligned data is required instead of aggregated information. The relational databases are designed for efficient insert, delete and update operations; hence R&D relational data model shall not solve the issues. The DWH is primarily focused on fast data aggregations across dimensions and efficient query performances. The R&D DWH is suggested as the best solution where inherent architecture with facts and dimensions shall be used to exploit huge data volumes to extract, map and align multiple-sources data instead of aggregation. In this case, the DWH structure can be evolved and modified in order to take into account the new business needs. The selection of DWH architecture is not based on efficient aggregation but on the efficient exploitation of huge data volumes for multi-source data mapping and alignment.

The proposed R&D data warehouse is presented as under in the Figure 6.8 a & b. The basic concepts of facts and dimensions tables and DWH architecture are used in the context of multi-source data extraction, mapping and alignment from huge data volumes. It is very important to note that at present, there are a number of excellent data analysis tools based on advanced data mining and artificial intelligence techniques but they cannot give results until they have multi-source input data in the correct format. The concept of multiple fact tables linked together like hierarchical dimensions has been proposed with an objective to accurately map and align different facts for further analyses.

The presented schema is the logical representation of DWH model without cardinalities and associations because as per our best knowledge till now there do not exist a standard modelling language for the DWH/DM schemas. The above presented DWH model is partial representation of actual schema because of confidentiality issues with contextual facts and dimensions, however, the above model is sufficient to present and demonstrate the concept. In this representation, [S.D], [D], [C.F] and [M.F] notations are used for sub-dimension, dimension, contextual facts and measure facts respectively. The product dimension has product_type, mask_set, process and mask sub-dimensions. The process dimension summarizes the process information for the given product and can be used for further exploitation. The mask_set and mask sub-dimensions provide critical information about wafers' structural description with quantitative measures. The process_plan dimension is composed of brick, operation, EDC_plan (engineering data collection) and step sub-dimensions. It provides complete description of the total steps including process and measurement during production operations for each product. The measurement steps are linked with the EDC plan which further links it with the test_program dimension to list out parameters to be measured in this respect. The lot_wafer dimension provides the list of effected lots and names of the wafers included in the lot. The chamber-tracking and run are the additional contextual facts linked to the lot_wafer_steps central contextual fact. The lot_wafer_steps is the central contextual fact which provides the actual execution of each production process step in detail for slicing/dicing operations during root cause analysis and knowledge capitalization efforts.

In this model we have included 4 measure fact tables as *(i)* PT_Measure, *(ii)* EWS_Measure, *(iii)* Inline_measure and *(iv)* Defectivity_Measure facts. Each measure fact is linked with the fact in the lot_wafer_steps fact table. The defectivity_measure fact is linked with the defectivity_program dimension and class_lookup and class programs sub-dimensions. The facts collected in this case are at the die levels and it can be seen that there are no traces of site level information at this level. The PT, EWS and Inline measure facts have the test_program dimension and test_bin_def, test_bin_relation, test_program_specs, test_structure_map and parameter sub-dimensions. The test_bin and test_program_specs provide list of parameters that can be tested against the given test_program along with specification limits. The test_structure-map sub-dimension is critical because it provides us access to the test structures and their location which are used during the measurement of the parameters. The additional contextual facts and dimensions which are helpful in establishing effective R&D efforts e.g. fault detection and classification (FDC) information is not presented in the above model being highly confidential. This and similar information can be added any time in the similar structure using our proposed Meta model for knowledge capitalization (Figure 6.8).

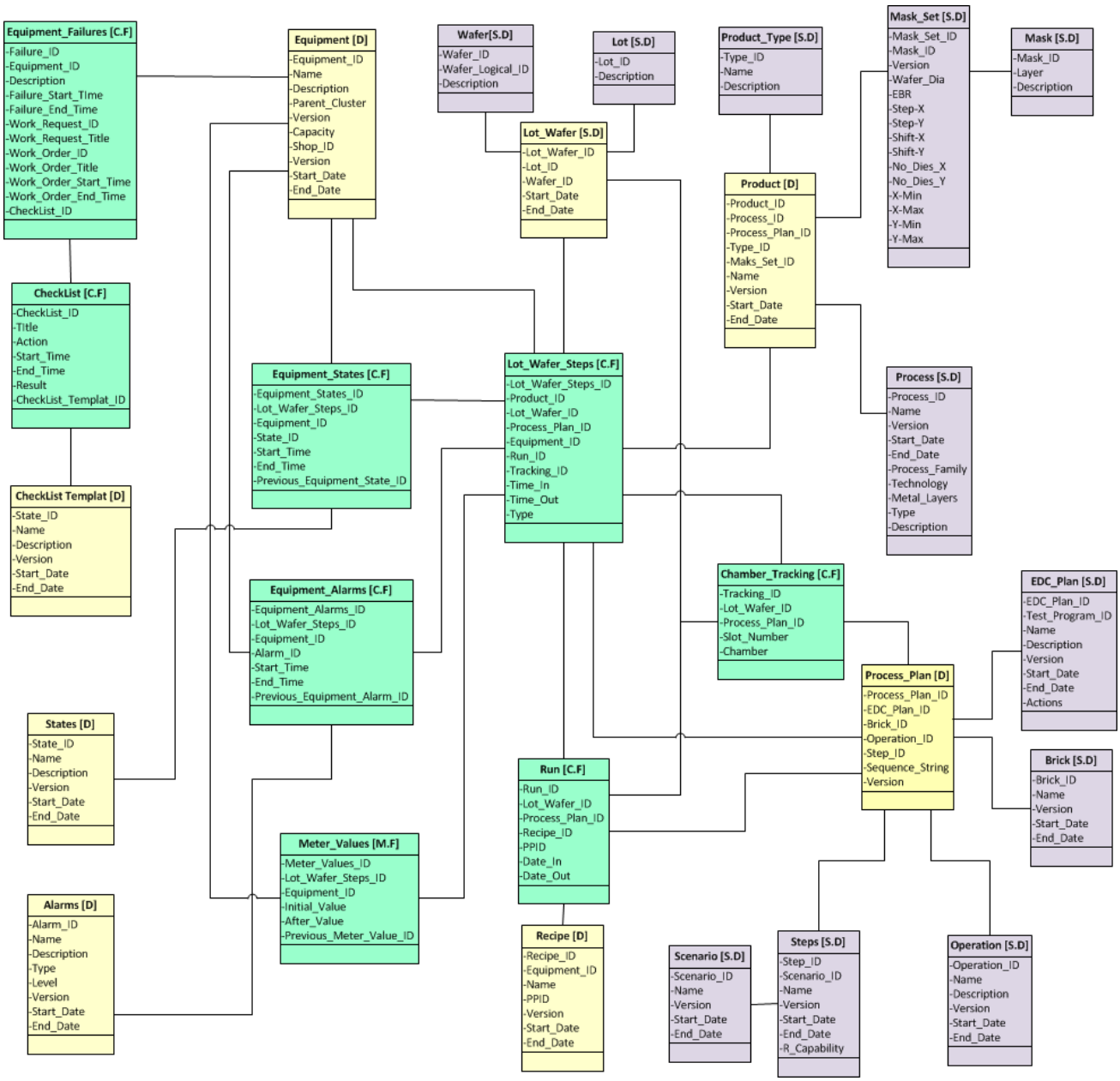


Figure 6.8(a) – Proposed R&D data warehouse for knowledge capitalization

While designing the R&D DWH model, there were two options as *(i)* the measured fact data could be treated prior to its storage for coordinates mapping and alignment using our proposed MAM and SPM models or *(ii)* store data in original shape with supplement or missing information like notch position, measure reference frame and wafer structural information and treat the data upon its extraction. The second option was chosen because it is believed that the data should be kept in its original format and end users must be given an option to transform it on as required basis. The advantage is that data is in its original shape and it can be used for different transformations as per choice of the end users. The proposed R&D DWH model is implemented using relational databases (MS Access); however, the rise of data volume for R&D and engineering tasks dictates us to select a commercial RDMS like Oracle, SQL Server or TeraData. The end users can use the above model to start the analysis from individual measured facts (PT, Inline, EWS, and Defectivity) and subsequently move towards its correlation with other measured facts and contextual facts. It is not possible to accurately define the potential size of the database to justify the use of commercial database management systems, but for the R&D and engineering initiatives, the estimated size for one year R&D data is expected to rise above 1.5 TB (tera bytes).

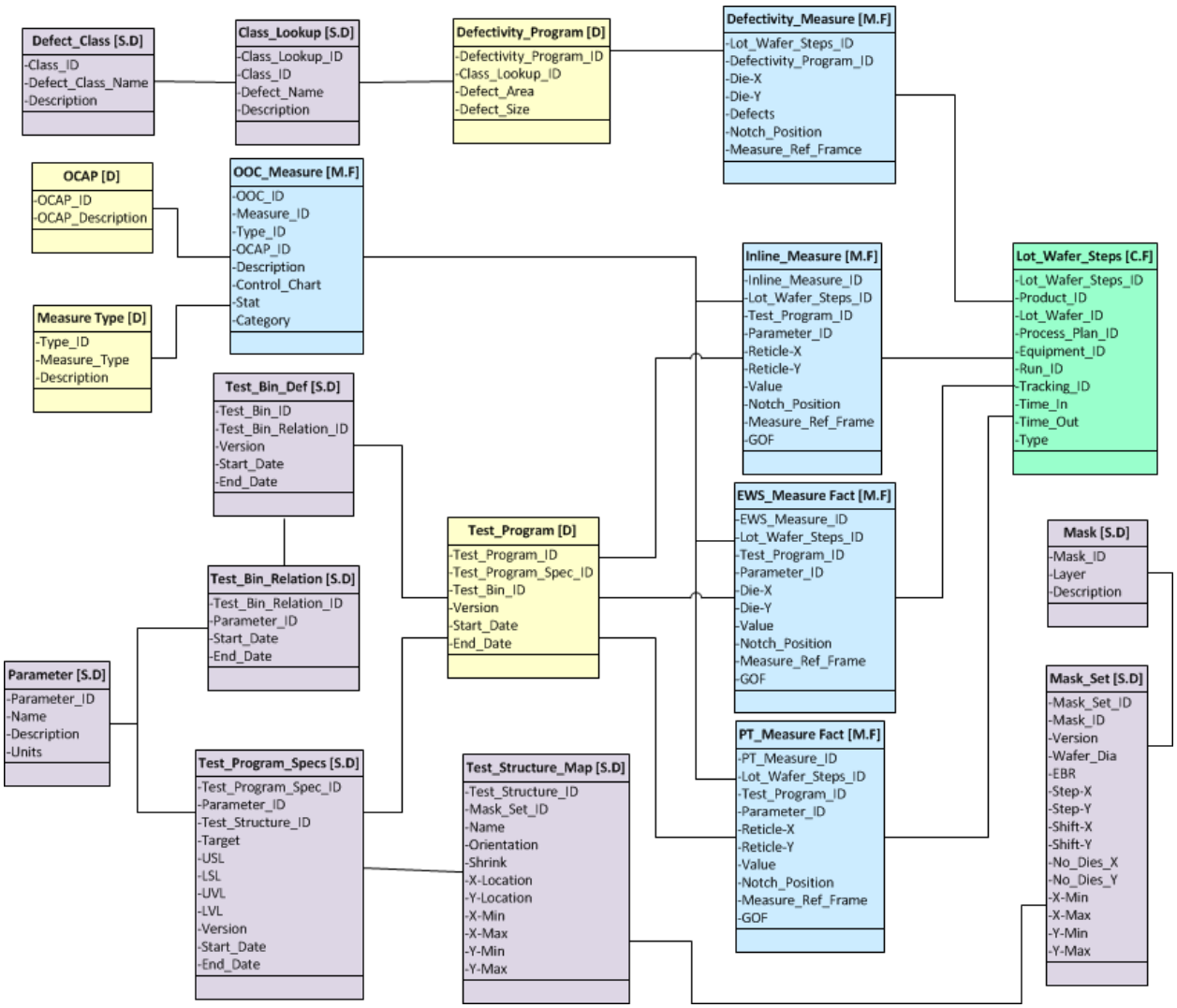


Figure 6.8(b) – Proposed R&D data warehouse for knowledge capitalization

The above presented data model for R&D DWH is a simple logical presentation with notations as [S.D], [D], [C.F] and [M.F] and without the declaration of primary and foreign keys. It is done on purpose to highlight that to the best of our knowledge there do not exist a standard DWH modelling language like UML. The DWH data model is not normalized and to certain extent duplication of the data is allowed to improve data retrieval efficiency. If a DWH is implemented in RDBMS then indexing and views can be used to improve the data retrieval efficiency; however, the best way is to use OLAP tools provided by the respective vendors to slice/dice and drill up/down the OLAP cubes for knowledge capitalization. It is also important to note that no primary or foreign keys are added or highlighted because duplication is allowed. The fact tables include keys from dimensional tables (foreign keys). They are used as composite key in the fact tables; however, if required, a simple identifier can be added within the fact tables. The same rule is also applicable for the dimension tables.

A Meta model could be extracted (Figure 6.9) for the knowledge capitalization from the above presented R&D DWH model to improve R&D effectiveness. In this schema, there are central fact tables that hold the contextual data whereas child fact tables associated with the respective central fact table are the quantitative measured facts. The central and child fact tables are connected in a star formation whereas the dimensions are shared across these fact tables and they are placed in hierarchal association resulting in the formation of flower schema. This schema has commonalities with star and constellation schemas as *(i)* star schema and *(ii)* shared hierarchal dimensions. The major differences in existing and proposed schemas are that in our proposed schema *(i)* contextual data is placed in central fact tables and *(ii)* quantitative measures or facts are stored in child tables in the star schema. The objective of the proposed flower schema is to enable

huge data volume exploitation for the purpose of mapping and alignment across multiple dimensions for effective root cause analyses (knowledge capitalization) but existing schemas are focused on efficient and fast aggregation of quantitative measures.

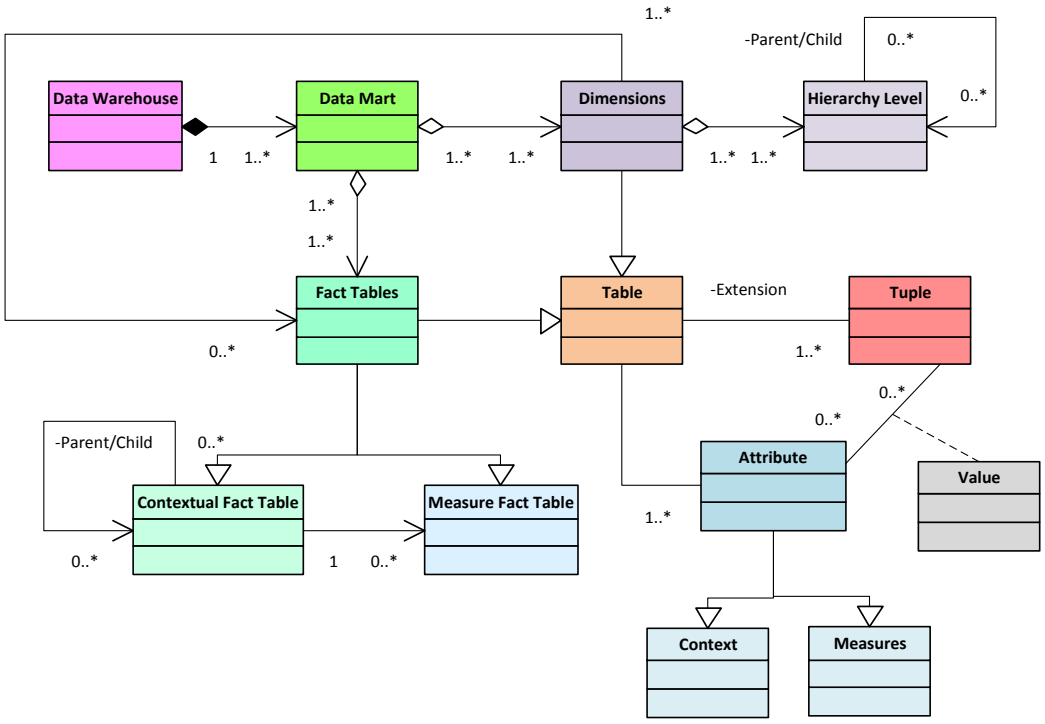


Figure 6.9 – Meta model for knowledge capitalization

In this proposed Meta model for knowledge capitalization a data warehouse is composed of multiple data marts which are further composed of multiple facts and dimensions tables. The fact tables are classified as contextual and measure fact tables. The contextual fact tables exist in the parent/child relationship and are supported with many measure fact tables. These tables are further supported with multiple dimensions. A table can have multiple attributes of the measure or context types. If we formulate a DWH from the above Meta model with generous extensions, we shall get the schema as presented below (Figure 6.10). We have used C.F, M.F, D and S.D notations for the contextual fact, measure fact, dimension and sub-dimension tables. The resulting logical structure of facts and dimensions is quite similar to flower; hence, we have proposed its name as flower schema that supports knowledge capitalization from huge data volumes across multiple dimensions.

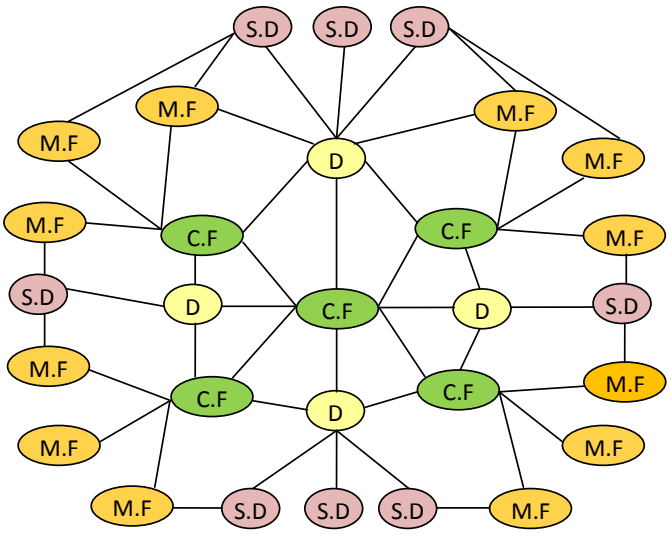


Figure 6.10 – Flower schema for knowledge capitalization

The above proposed R&D DWH architecture and respective Meta model demonstrates that multi-dimensional modelling concepts can also be used for multi-source data mapping and alignment in addition to the aggregation over huge data volumes. It is biggest advantage is that it lies next to relational data models and provide us with an ability to extract multi-source data for an effective root cause analysis. We have full control over its structural evolution and it can be modified on as per need basis in terms of new dimensions and facts. The scope of proposed R&D DWH is not limited to R&D teams as it can be equally used by all engineering teams. Its final implementation using relational or dimensional databases is left on the discretion of end users; however, based on the fact that focus has been put on mapping and aligning data, its implementation in relational database shall be equally good at reduced cost and effort.

Let us present two end user requests for the extraction of PT and EWS data based on *(i)* site and die levels and *(ii)* based on test structure positions. The end users are concerned with the lot QS93005 and wafer QS93005.5 for PT and EWS measurements carried out against metrology steps *(i)* CUMEAS_PARAMETRIC_T1-01 and *(ii)* TEST_PRO45V1_T1-01 respectively. The step *(i)* is performed at metal-2 and step *(ii)* is performed at the end of metal-7. The electrical tests performed at metal-2 are known as PT tests whereas these tests when performed at the end metal-7 are called EWS tests. Now the above presented DWH model (Figure 6.7) with the SQL queries will be used to extract the relevant data for further processing using MAM and SPM models.

a) Site and Die Level PT and EWS Data extraction:

Select Lot_Wafer_Steps.Lot_Wafer_Steps_ID, Lot_Wafer_Steps.Process_Plan_ID, Process_Plan.Step_ID, PT_Measure. From Lot_Wafer_Steps (Inner Join Lot_Wafer ON (Lot_Wafer_Steps.Lot_Wafer_ID = Lot_Wafer.Lot_Wafer_ID) Where Lot_Wafer.Lot_ID = 'QS93005' and Lot_Wafer.Wafer_ID = 'QS93005.5') Inner Join Process_Plan On (Lot_Wafer_Steps.Process_Plan_ID = Process_Plan.Process_Plan_ID) Inner Join Steps On (Steps.Step_ID = Process_Plan.Step_ID) Where Process_Plan.Step_ID IN ('CUMEAS_PARAMETRIC_T1-01', 'TEST_PRO45V1_T1-01') Inner Join PT_Measure ON (Lot_Wafer_Steps.Lot_Wafer_Steps_ID = PT_Measure.Lot_Wafer_Steps_ID)*

The above SQL multiple join query shall result all PT and EWS parameters IDs and values.

b) Test Structure Position Based PT and EWS Data Extraction:

Select Lot_Wafer_Steps.Lot_Wafer_Steps_ID, Lot_Wafer_Steps.Process_Plan_ID, Process_Plan.Step_ID, PT_Measure., Test_Program.Test_Program_Spec_ID, Test_Program_Specs.*, Test_Structure_Map.* From Lot_Wafer_Steps (Inner Join Lot_Wafer ON (Lot_Wafer_Steps.Lot_Wafer_ID = Lot_Wafer.Lot_Wafer_ID) Where Lot_Wafer.Lot_ID = 'QS93005' and Lot_Wafer.Wafer_ID = 'QS93005.5') Inner Join Process_Plan On (Lot_Wafer_Steps.Process_Plan_ID = Process_Plan.Process_Plan_ID) Inner Join Steps On (Steps.Step_ID = Process_Plan.Step_ID) Where Process_Plan.Step_ID IN ('CUMEAS_PARAMETRIC_T1-01', 'TEST_PRO45V1_T1-01') Inner Join PT_Measure ON (Lot_Wafer_Steps.Lot_Wafer_Steps_ID = PT_Measure.Lot_Wafer_Steps_ID) Inner Join Test_Program On (PT_Measure.Test_Program_ID = Test_Program.Test_Program_ID) Inner Join On (Test_Program.Test_Program_Spec_ID = Test_Program_Specs.Test_Program_Spec_ID) Inner Join Test_Structure_Map On (Test_Structure_Map.Test_Structure_ID = Test_Program_Specs.Test_Structure_ID)*

The above example returns the same PT and EWS data at die and site levels but along with the test structure positions. These test structure positions are further used in the SPM models to perform accurate mapping based on site and die levels generated through MAM model. The key is that this data model provides and generates all type of measured and contextual data needs. The data extracted is further processed using MAM and SPM models for the coordinate normalizations. The above presented SQL queries are not the only set of queries, stored procedures can also be built as per user needs which can be further supported with user define SQL functions for slicing, dicing and drill up/down operations. The size of one year's data storage is estimated to be 1.5 TB which means that multiple join queries are likely to crash; hence, OLAP supported RDBS (Oracle, SQL Server) or purely dimensional DBMS as TeraData are needed for efficient and fast query results. To best serve the end users, it is suggested that based on this R&D DWH

model, customized data extraction utilities must be developed and provided to R&D engineers for the productivity and improvement.

6.6 Problem Context and Current Challenges

Let us discuss current challenges in the data extraction and analyses framework (Figure 6.11) where there are multiple engineering and R&D teams working for the technology derivative/improvement alignment/adoption and process control efforts and initiatives. The R&D efforts by these teams require data extraction and analysis for which there exist a large number of customized applications CA1, CA2 … CAn. The end users use these applications to access the data which is stored in 3-level storage architecture *(i)* ODS, *(ii)* RDBS and *(iii)* DWH and DM. The data extraction requests are classified as single-source or multi-source data analysis initiatives. At present, the single-source data extraction and subsequent analysis is likely possible and is being carried out but multi-source data extraction is not possible because of the missing database links, missing values, incompatible data formats and users' inability to use multi-source data extraction utilities. The IT revolutions have resulted in huge data volumes and availability of new facts which must be constantly updated in existing data resources to improve the R&D effectiveness. This is not possible because of the fact that the operational data sources are of proprietary nature and its structure cannot be changed. It results in a situation where multi-source analysis is never possible and often results in the waste of time and resources when engineers try to align and map the data using excel sheets.

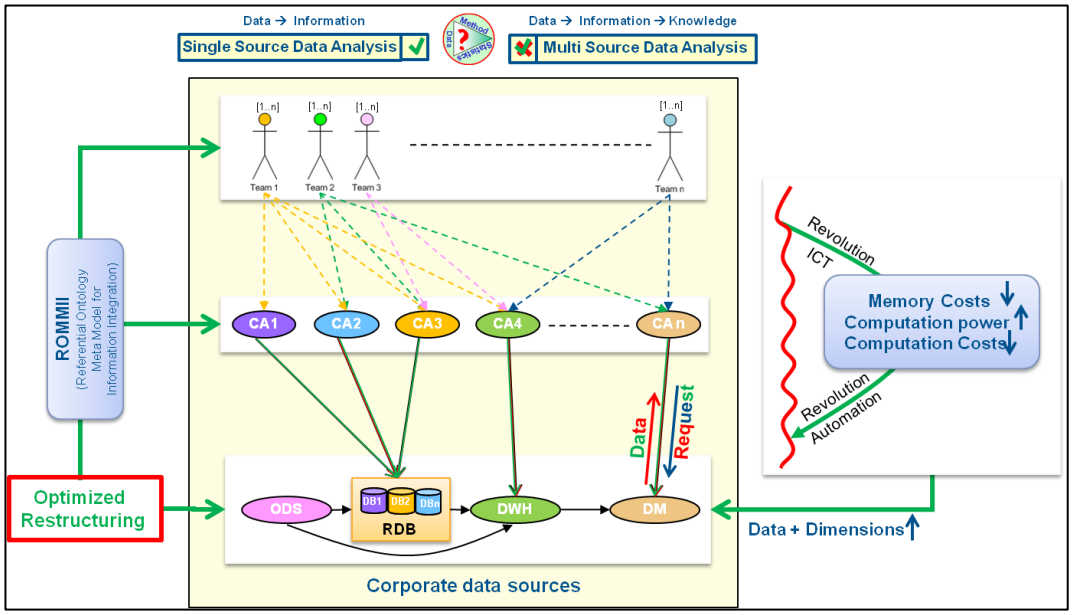


Figure 6.11 – Current data extraction and analysis challenges

The proposed R&D DWH solution to exploit huge data volumes for mapping and alignment is sufficient to provide means for effective root cause analyses; however, it does not provide solution for unstructured evolution of data models that have been established as critical factor for effective R&D efforts.

6.7 Proposed ROMMII Framework

The proposed R&D DWH does not provide solution for the relational data model evolution because they are proprietary databases and changing the database structure shall result in crashing end-user applications. It is also not possible to redesign and developing the existing relational databases and applications; hence, we have proposed ROMMII (**R**elational/**R**eferential **O**ntology **M**eta **M**odel for **I**nformation **I**ntegration) framework to bridge the gap. This framework provides a high level Meta model to implement the ontology for information integration without changing the structure of existing data sources. It is implemented as a relational database instead of using ontology based languages. This is because of the fact that SQL query performance for the validation and optimization of the end-user requests is much higher if manipulated as a

relational model. The relational database is also easy to manage and control against the rapid and frequent modifications.

The proposed framework is called Meta Model because it precisely defines specifications of the relational model which is used to track the modifications and missing links between databases. It is referred as ontology because of the fact that it offers a shared collection of missing knowledge across the databases for subsequent use during query pre-processing and its reformulation for the query optimization. The word "information integration" represents the objective of the proposed framework besides different semantics and syntaxes through alias, feed forward and feed backward transformation functions. The objective is to optimally use the existing relational data resources supplemented with additional relational or DWH databases to provide end users with updated and correct information for knowledge capitalization.

6.7.1 Use Case Diagram for ROMMII Platform

The use case diagram for ROMMII platform is presented in the Figure 6.12. The proposed system is composed of 4 main use cases as *(i)* learn Meta model, *(ii)* modify Meta model, *(iii)* execute query and *(iv)* compute user statistics. The "learning Meta-model" is the first step in the proposed framework and is to be carried out by the IT administrator because this step requires the definition of feed forward and feed backward functions to establish potential and missing links between databases. The "modify data model" allows the IT administrators to evolve the data models on need basis which imperatively uses 'learn Meta model' and pre-failure assessment functions. The pre-failure assessment function can be initiated by the data administrators with an objective to identify the potentially effected applications, users and concerned application and IT administrators. In this use case diagram the roles of IT and data administrators have been differentiated. The data administrator is responsible for all the data, its structure, format and storage within organization whereas IT administrators are responsible for the efficient management of data sources.

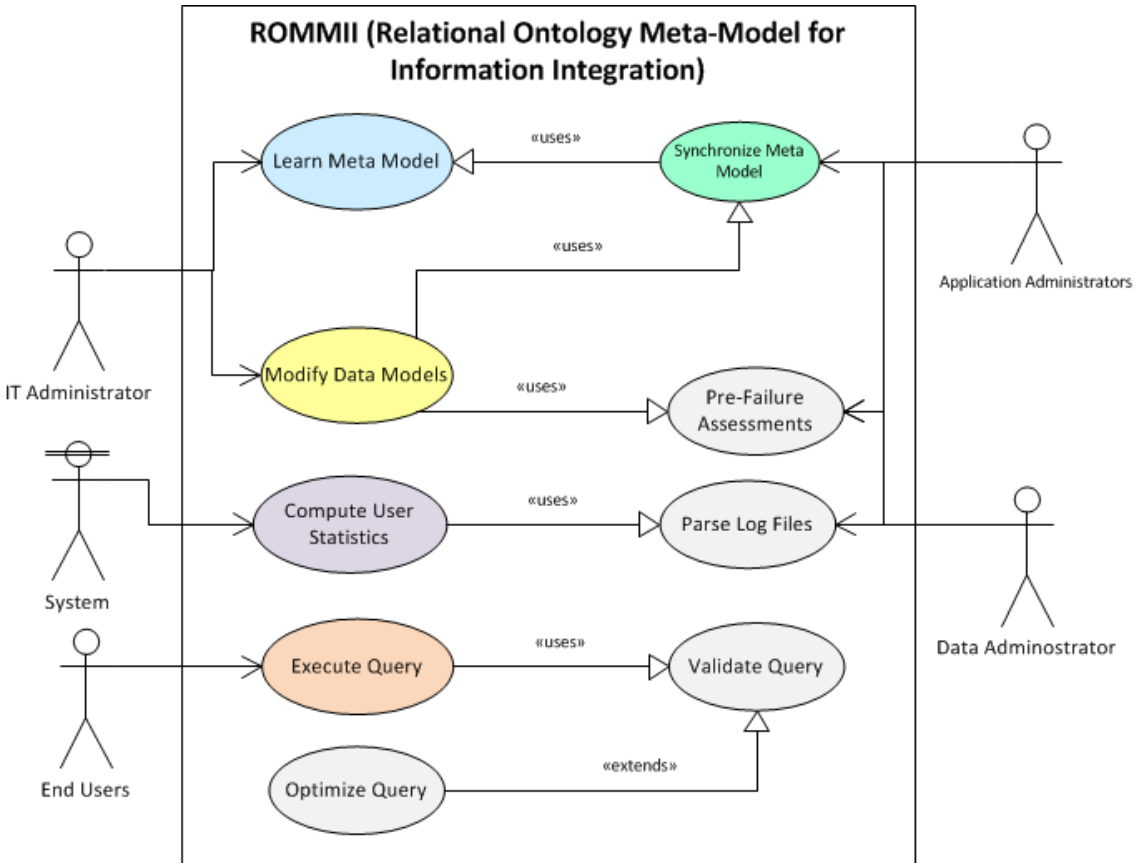


Figure 6.12 – Use case diagram for ROMMII platform

The end users execute the SQL queries in search of data/information but these queries must be validated against the evolved model for consistencies. The 'synchronize data model' function ensures that upon any modification the end user applications are properly updated by application administrators. The query optimization function extend its services to 'validate query' function where SQL query is reformulated for *(i)* retrieval efficiency, *(ii)* exclusion of deleted fields/tables and *(iii)* inclusion of new fields if they are not synchronized with end user applications. The 'compute user statistics' function computes and updates the user statistics based upon the data queries executed by end users through respective applications. It is important to note that based on ROMMII platform, customized end user data extraction applications are suggested.

6.7.2 Meta Model for ROMMII Platform

The Meta model for ROMMII platform is presented in Figure 6.13. In this Meta model 'database' is composed of tables and can have 1 or many database versions. Each new 'database version' (a) results from the sequence of structural 'modifications' (b) like, addition, deletion or modification of tables/fields. The 'tables' are further composed of 'fields' which can be either 'key' or 'non-key' attributes. The fields have multiple 'sample' values which are used in identifying fields that can be potentially used to establish a 'relation' between tables. These relations are classified as potential missing links between databases. The 'fields' are associated with the domain and can have multiple 'alias' that could likely be a part of 'potential links' (c) between 'tables'. The 'relation' between tables is associated with the 'transformation rules' (d) which are classified as feed forward or feed backward. The 'user statistics' (e) is based on historical patterns of the query execution by end users through applications. The potential risk is computed and stored in 'users_risk_summery' based on type of queries executed by the end users and associated risk levels defined in 'query_risk_levels'. The data model constructed from the Meta model below is not presented here due to confidentiality reasons. The data model corresponding to the Meta model for ROMMII platform is referred as the Metabase in this chapter.

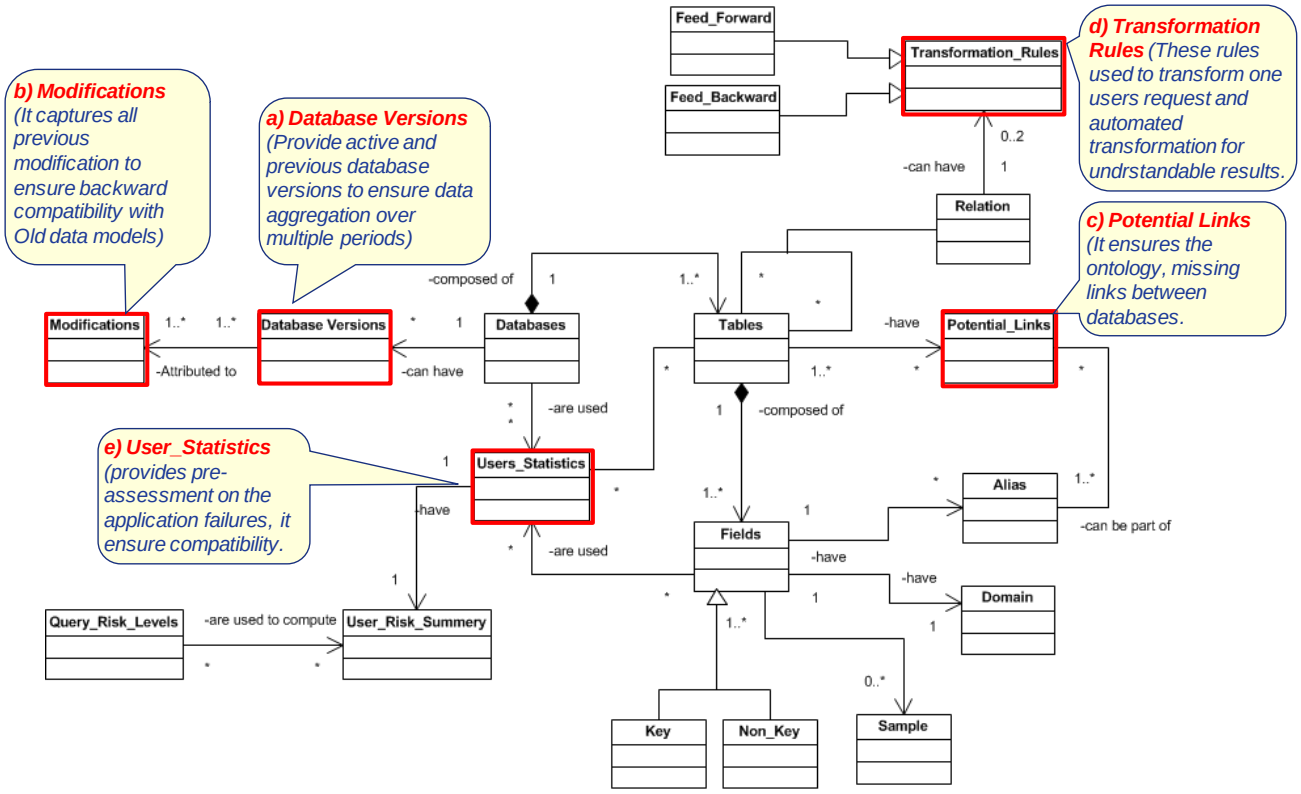


Figure 6.13 – Meta model for ROMMII platform

6.7.3 Activity and Sequence Diagrams against Use Cases

In this section, the activity diagrams against uses cases are presented as *(i)* learn Meta model, *(ii)* synchronize Meta model, *(iii)* query validation and optimization, *(iv)* log file parsing and users statistics computations and *(v)* pre-failure assessments. The modify data model structure (model evolution) use cases are based on the pre-failure assessment and learn Meta model use cases; hence, they are included in the sequence diagrams but not in the activity diagrams.

a) Learn Meta Model

The objective of this use case is multifold; it learns the Meta model for selected database and allows subsequent inclusion of other databases. The new database addition or modification requires Meta model learning and synchronization sub-processes. The activity and sequence diagrams for the "learning Meta model" use, case are presented in Figures 6.14 and 6.15 respectively.

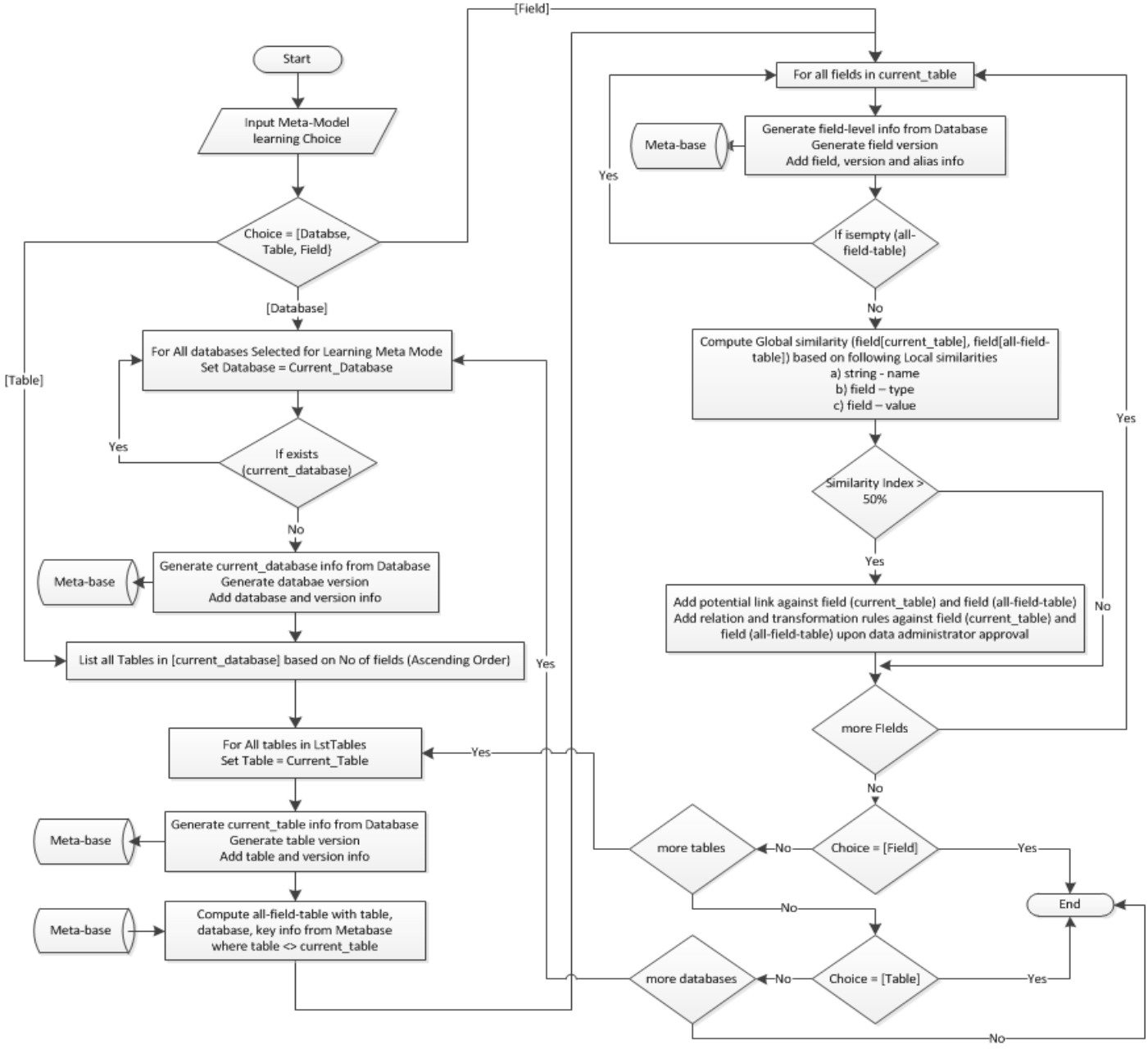


Figure 6.14 – Learning Meta model for new database, table and/or fields

The learn Meta model activity starts with a choice for the database/data administrators and redirects itself based on the choices made as the database, table or field. A list if attributes and paths for each selected database is prepared. We loop through each database and verify its duplication from existing Metabase. The databases with duplication are skipped and we proceed to the next database in list. If the database is not in the Metabase, then the database and its version info is generated and added to the Metabase. A list of tables in t heselected database is prepared and looped followed by the addition of its version info into Metabase. At this stage, a list of all fields from the Metabase is generated excluding the table being processed for the identification of potential missing links. We, then start by first adding information of each field in the table being processed, into Metabase along with alias, local and global similarity indices. If the similarity index computed on two fields is > 50%, then they are suspected to have potential link. The data/database administrators make decisions about the inclusion of this as a potential link, if accepted, the potential link info along with feed forward and/or feed backward transformation functions are developed and added by the database administrators.We continue until all fields, tables or databases are looped through and added into Metabase for query validation and optimizations efforts. The role of each participant in the ROMMII platform is explained in their relevant activity diagram.

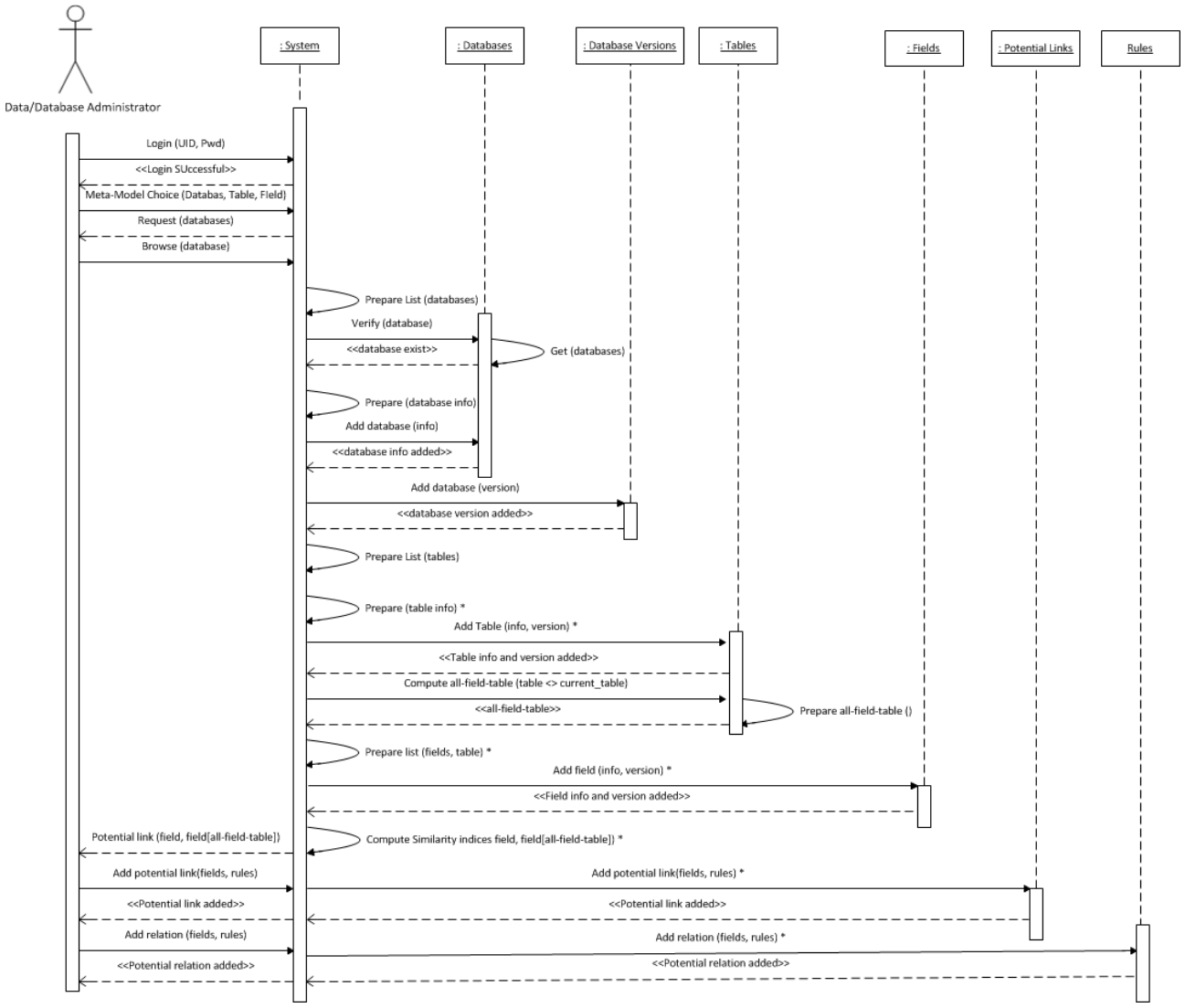


Figure 6.15 – Sequence diagram to learn Meta model use case

The interactions between data/database administrators and the system objects are presented above in the Figure 6.15. The data/database administrators identify themselves with a UID (user identifier) and PWD (password) to get appropriate rights for the execution of "learn Meta model" use

case. The presented sequence diagram is compatible with the activity diagram and is self explanatory. It is important to note that potentially identified links between databases and/or tables are pointed by the system whereas their inclusion in the Metabase is always at the discretion of the administrators. All potential links are further supported with the identification of feed forward and/or feed backward transformation rules, which are programmed as internal database functions to improve query performance.

i) Synchronize Meta Model

The "synchronize Meta model" is an important use case that extends its services to modify the data models. It uses "pre-failure assessment" and "computes user statistics" use cases. The basic objective of this use case is to manage and resolve all the inconsistencies that exist between the end user applications used for data extraction and the Metabase. It happens when application administrators fail to comply with the notifications from ROMMII for the application modifications. This use case can be initiated by the application administrators to synchronize model modifications and ensure that no failures are associated with respective user applications. The activity and sequence diagrams are presented in the Figure 6.16 and 6.17 respectively.

In the Figure 6.16(a), database/data administrators are provided with an option to modify table or field. The table option further requires the selection of modification type as add, delete or modify the table. The modify table option is similar to the modify field option because in both cases users can add, delete or modify the fields. The modify field or table options are presented in Figure 6.16(b). The table deletion is the most simple process where the table status is set to deleted in the Metabase along with start and end valid times. The table is not deleted physically from the Metabase which help us in SQL query optimization. This step requires similar deletion of corresponding alias, domains, samples, potential links, relations and transformation rules. The add table option is simple as it requires just new entries. It starts with the generation of table properties and its version info which are then added to the Metbase. The "all-field-table" is generated from the Metabase which is further used to find out potential links with the table fields being added into the Metabase. The table fields are added to the Metabase with alias, potential links, relations and transformation rules. These rules help us to remove data type and/or semantic inconsistencies.

The modify table option requires to check if the table exists in Metabase otherwise the next table is selected. The modification in the table is classified as addition, deletion or modification of field type. The addition and modification of field type requires deletion of their current versions along with addition of new information and version. The new field is then checked against the existing fields for potential links and relations. All these changes are recorded in the Metabase which generates new database version. The actions performed are stored against each new database version which can be easily explored and reversed to generate the previous database versions.

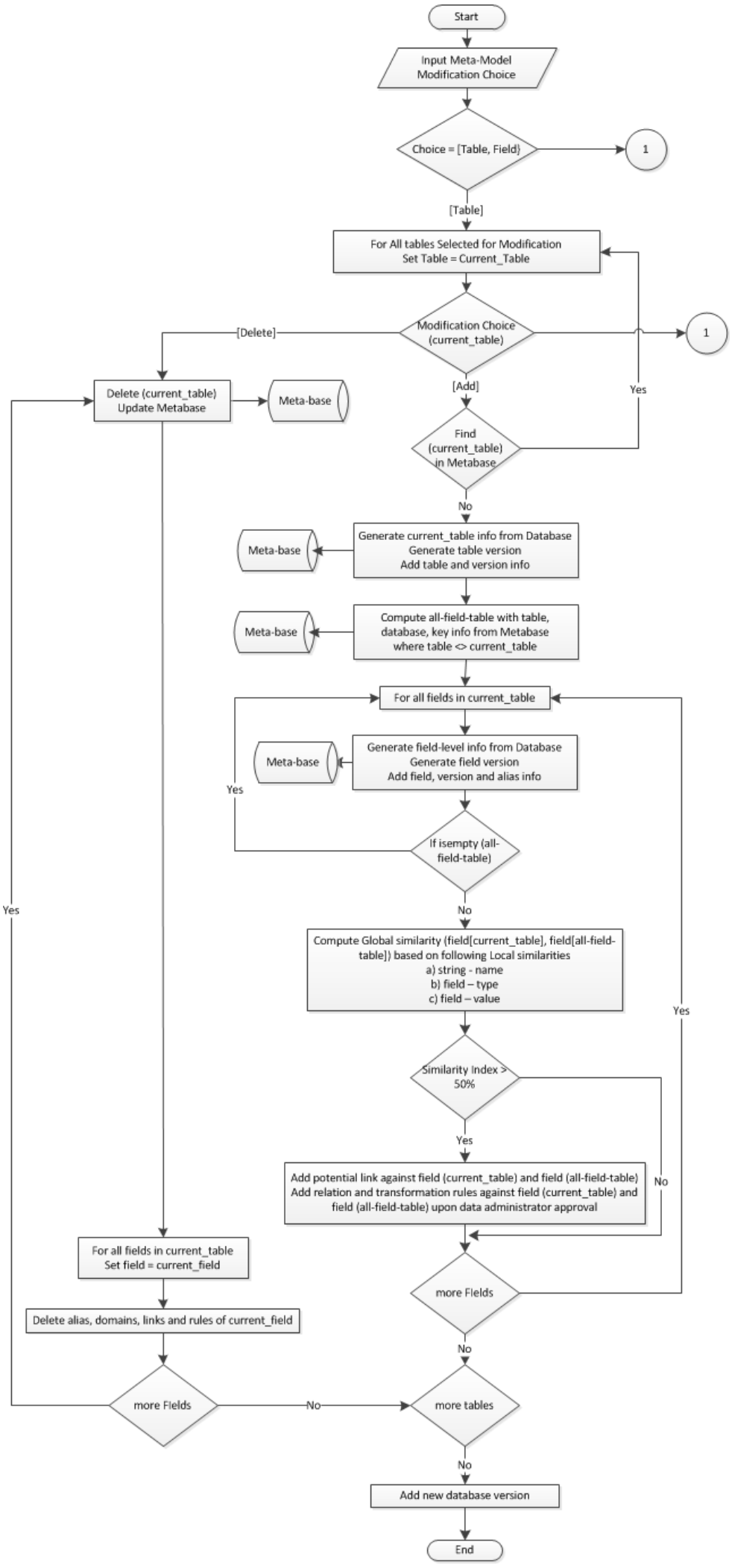


Figure 6.16(a) – Modification and synchronization activity diagram

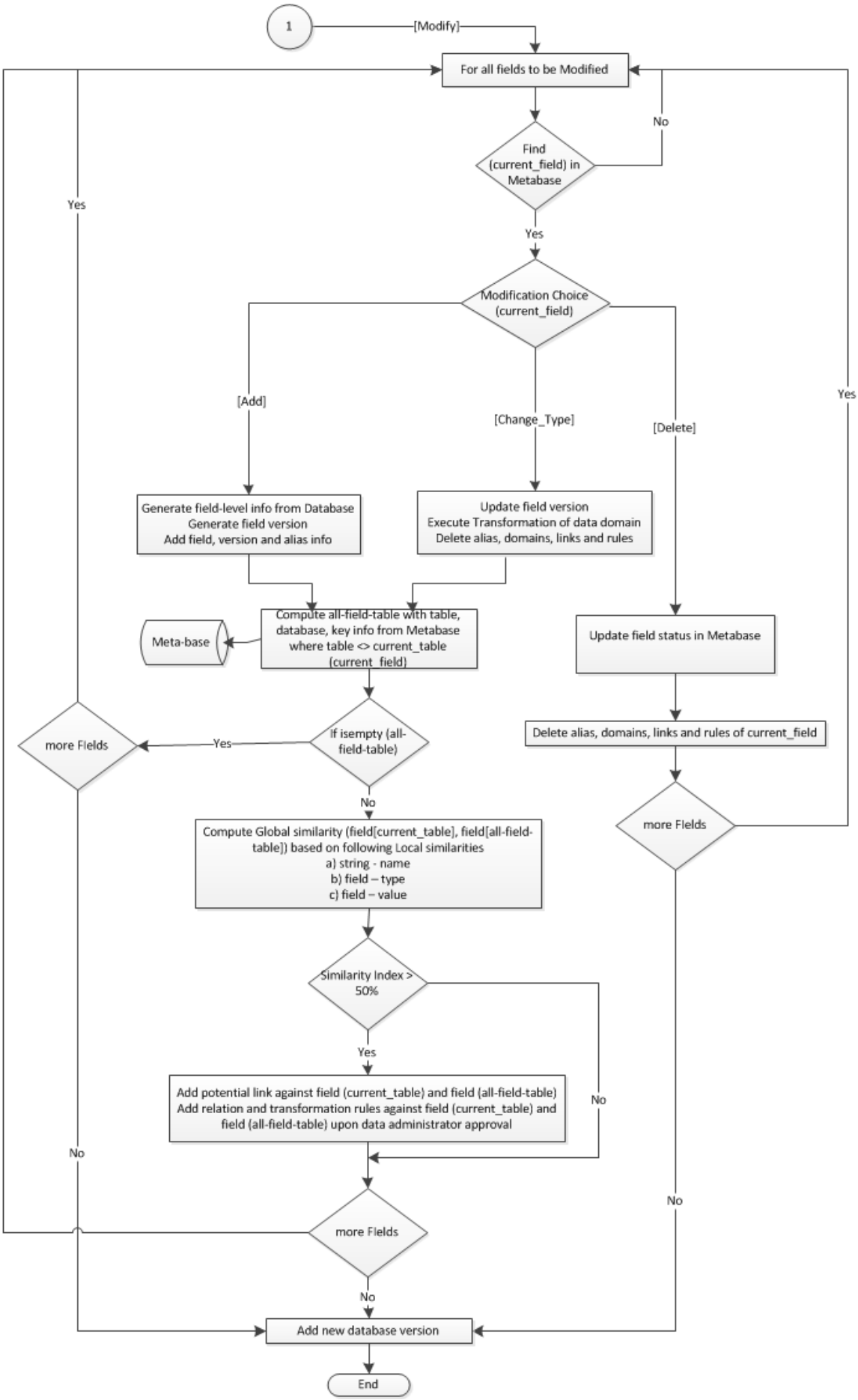


Figure 6.16(b) – Modification and synchronization activity diagram

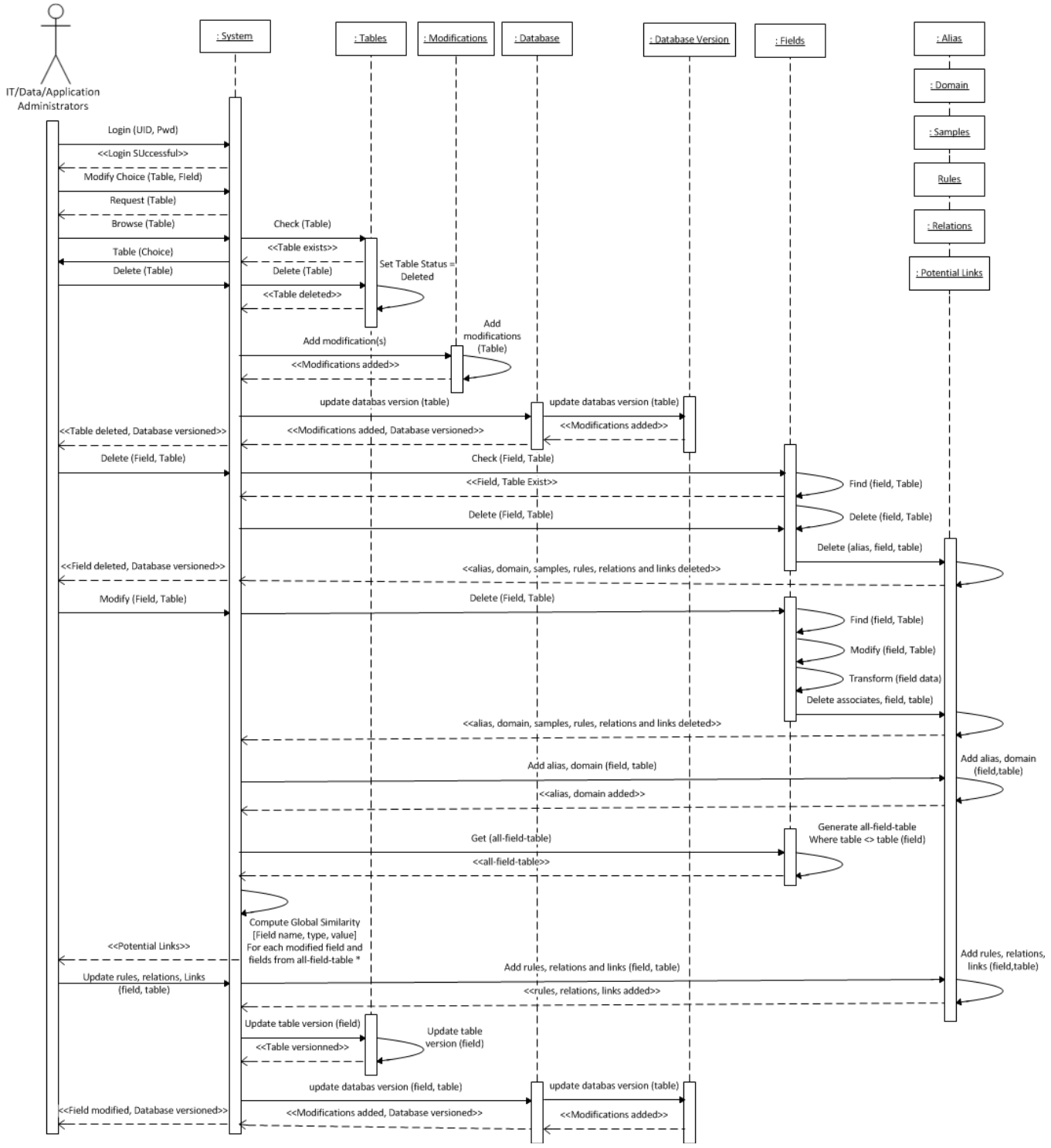


Figure 6.17 – Sequence diagram to modify and synchronize Meta model

ii) Query Validation and Optimization

This is one of the most important and key use case in the ROMMII platform where all the end users queries for data extraction are first validated followed by its optimization. It is important to note that user express their data needs through complex GUI (graphical user interfaces) where upon the user selections, a complex SQL query is generated at the back end. This query is required to be validated to ensure that it shall return valid results. If inconsistencies at this level are found then optimization algorithm reformulates SQL query to avoid the application crashes. It provides an opportunity to move towards application agility; however all the queries might not be optimized e.g. if the inner join or where clause criteria involves a table or field which have been deleted or modified. The activity and sequence diagrams for this use case

are presented in the Figures 6.18 and 6.19 respectively. In the proposed ROMMII platform, the focus is on data manipulation language (DML) which includes select, insert, update and delete SQL queries with 'inner joins' and 'where' clauses; however, it can be equally upgraded to include data definition (DDL) and control (DCL) languages.

The primary objective of this use case is to validate and optimize query prior to execution. It can be seen in Figure 6.17 that we start by parsing the SQL query in three lists as *(i)* where_fields, *(ii)* inner_join_fields and *(iii)* select_fields lists. These lists hold the respective fields in the format of 'database.table.field'. The tables from where_fields and inner_join_fields are checked in the latest database and respective table versions. It is important to note that if 'where' and 'inner join' clauses have wrong or deleted fields the execution of the SQL query shall result in a failure. If any of these tables are missing then we do not proceed and inform the user about non executable SQL query; however, if all the tables are traced back in the Metabase then list of all the fields from the select_fields list is prepared. These fields are checked against the latest field versions in respective database; however if a field is missing it is traced back in the older versions. The fields not found in the older databases versions are removed from the SQL queries whereas traced fields are kept for "select" type SQL queries. The traced fields are removed for insert/update and delete queries. The SQL query is updated with the fields, which were newly added in the current version besides the fact that they were not selected in user defined queries. Finally the numbers of fields in the validated and optimized SQL query are computed. If the field count is < 1 then user is informed and query execution is aborted but for field count >=1, end users are provided with the requested data. The optimization in this process is defined as steps *(i)* remove missing fields, *(ii)* keep fields not found in current version but available in the older versions and *(iii)* add new fields in current version from the previous version. It is referred as optimization because it ensures that the SQL query is executed even if it includes previous version fields. It provides end users with additional fields which are recently added in new versions but not selected by the end user for reference.

The sequence diagram for the query validation and optimization is presented in Figure 6.18. It is initiated by the end-user where he selects the data as per his need and requirement through a GUI. Based on this selection, the SQL query is automatically generated with "where" and "inner join" clauses, which requires the validation and optimization before execution. This sequence diagram includes only two objects *(i)* tables and *(ii)* fields, and is quite efficient in its execution; however, database and database version objects can also be included if multiple database query is to be executed. In our case, it is assumed that the users are working with single flexible DWH database. The execution flow is almost similar as explained above.

The list of excluded fields from the SQL query are also reported back to the end user along with the inclusion of new fields in the current version to ensure that end users are familiar with new type of data for its potential use during analysis efforts. This activity diagram complements the "pre-failure assessment" use case as it provides means to compute potential failure in terms of RPN (risk priority number) using 'query risk levels' object from the Metabase. The RPN is not used to prioritize the potential failures but for the relative comparison of gains achieved through ROMMII platform.

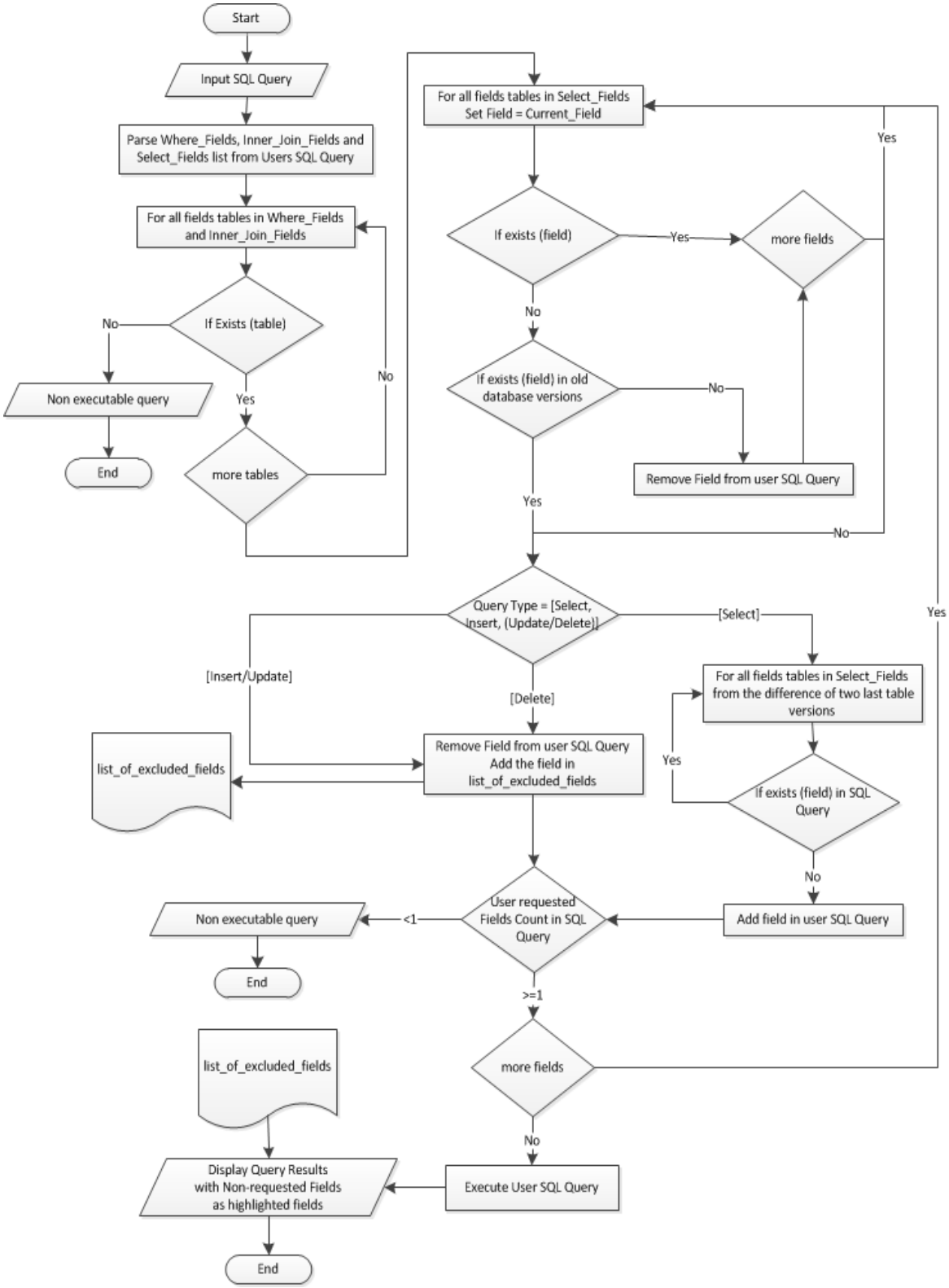


Figure 6.18 – Query validation and optimization activity diagram

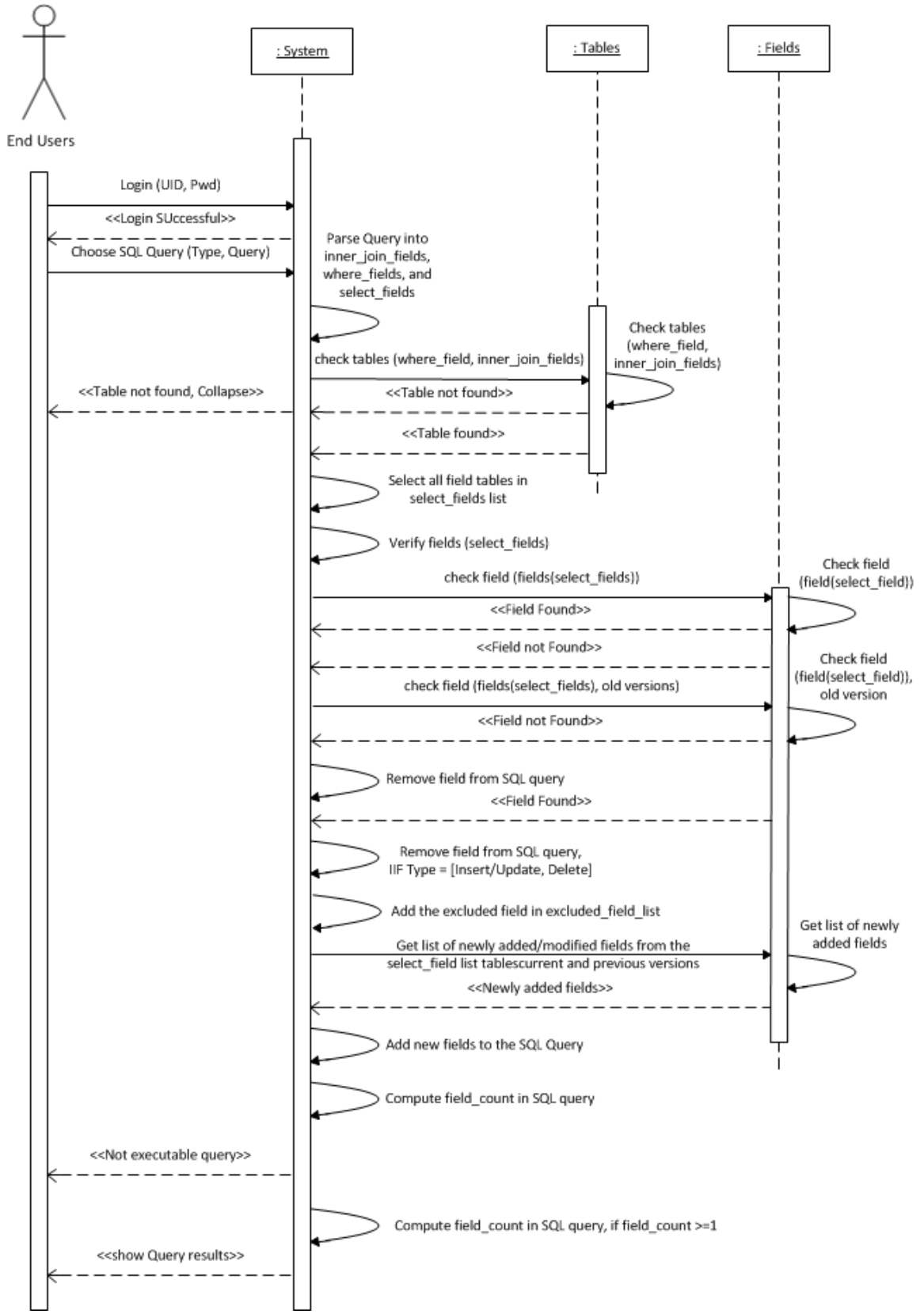


Figure 6.19 – Query validation and optimization sequence diagram

iii) Log file Parsing and Users Statistics Computations

The "log file" parsing and "user statistics computation" use cases generate knowledge that shall be used during pre-failure assessments. The log files are generated by database server on 24 hours basis using automated routine and accordingly user statistics are updated. These log files can be adjusted by database/data administrators on weekly basis e.g. if it is changed to weekly basis the log files are still generated on 24 hours being sequenced as 1 to 7 for the given week. The activity diagram for these use cases is presented in the Figures 6.20; however, sequence diagram does not seem to contribute or elaborate the working of proposed ROMMII platform so we have restricted ourselves to the activity diagram only.

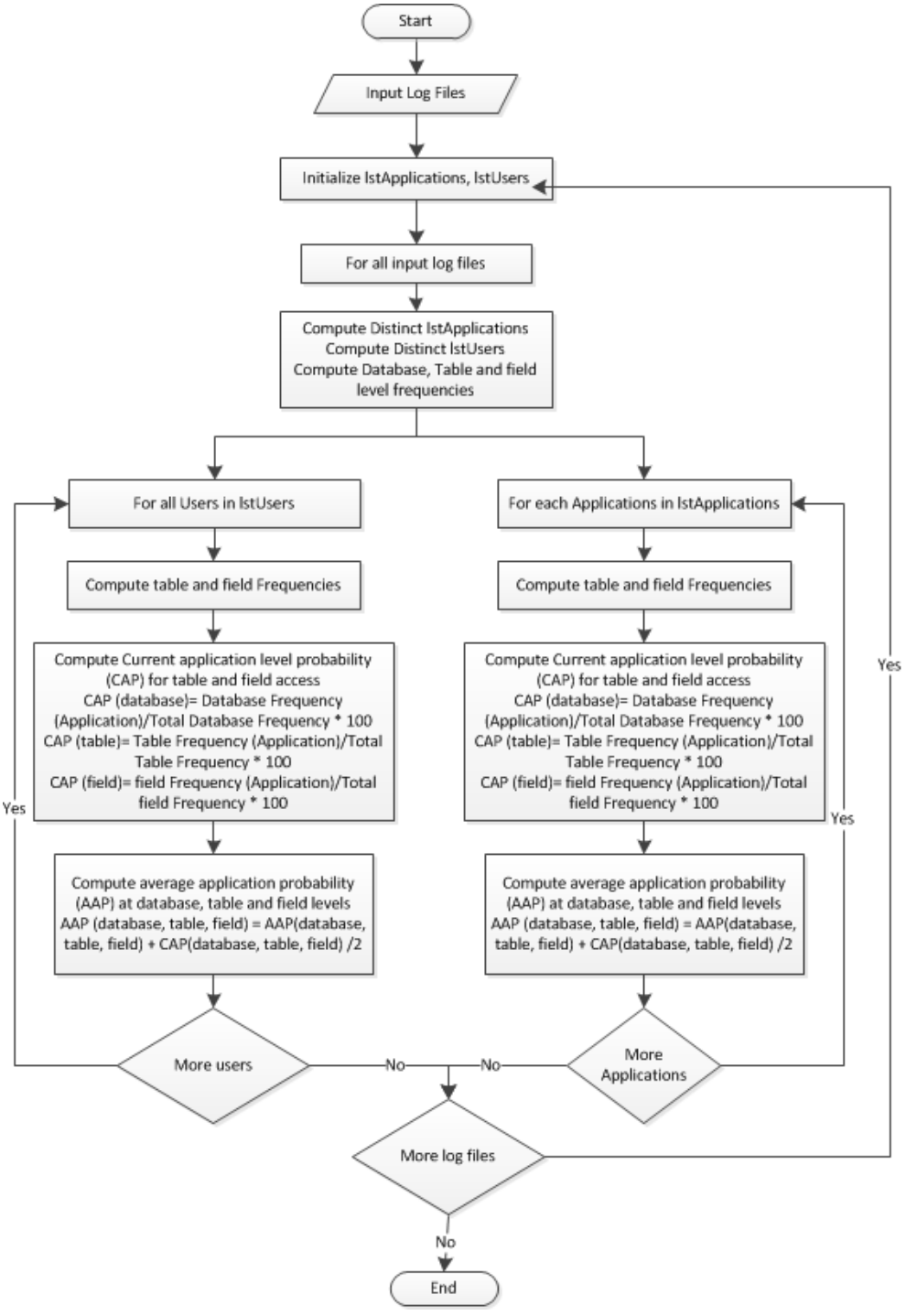


Figure 6.20 – Log File parsing and user statistics computation activity diagram

In the above presented activity diagram (Figure 6.21), we start by importing log files. The distinct users and applications lists are computed by parsing log files one by one along with database, table and field level frequencies. From this step, we follow two parallel flows as *(i)* application level statistics and *(ii)* user level statistics. In application level statistics, we compute database, table and field level frequencies followed by the computation of current application probability (CAP). This step follows an average application probability (AAP) computation where simple average of the two computed statistics is taken. These steps are repeated for each application identified from the current log file being processed. The CAP and AAP are computed for user level statistics against all distinct users identified from the log file. These computations are repeated until all the log files are processed.

iv) Pre-failure assessments

The "pre-failure assessment" is another important use case in the ROMMII platform, where situations can be identified when end user requests for the data cannot be fulfilled. This use case supports the "modify data model" and "synchronize" use cases. The activity diagram is presented in Figures 6.21. The pre-failure assessment is quite similar to the query validation and optimization activity diagrams with the only difference that upon a non-executable SQL query the RPN is computed and reported to the users; however, in case if optimized query is executable then excluded field list and initial count of select_fields list is used to compute the probabilistic occurrence instead of fetching it from the Metabase. The risk is reported to the user and can be used in many possible ways. It is important to note that SQL query is not executed in this use case. The sequence diagram has been purposefully excluded as it is an extension of the sequence diagram in the Figure 6.19. The only difference is the addition of 'query risk levels' object for the computation of RPN (risk-priority-number). The RPN number provides an idea about the potential pre-failure; however, the real value added is the optimized query and excluded list of fields.

The implementation of proposed methodology is not trivial because the existing data sources might have been implemented in heterogeneous database management systems. The learning Meta model, log file parsing and query optimization and validation are the use cases which require in depth knowledge and understanding of the internal database architecture.

Let us summarize the contributions made in this section. The ROMMII platform has been presented to address challenges of agility. The increasing heterogeneous data needs of the engineers require continuous evolution of data models which is not possible at present because any change in data model results in the failure of front-end user applications. One of the biggest reasons is the proprietary nature of the data sources and front-end applications, which restricts us from data model evolutions in accordance with the change in organizational data needs. We have seen the decreasing prices for storage and computation of data, which has made possible the exploitation of huge data volumes with data dimensions. To take full advantage of the present opportunities, the ROMMII platform is proposed that provides ontology based Meta model to manage structural evolutions. The proposed platform ensures backward compatibility of data extraction and analysis tools to meet with the multi-source data extraction needs of engineers. The proposed platform is tested using MS Access 2003, but implementation is not made public due to confidentially reasons.

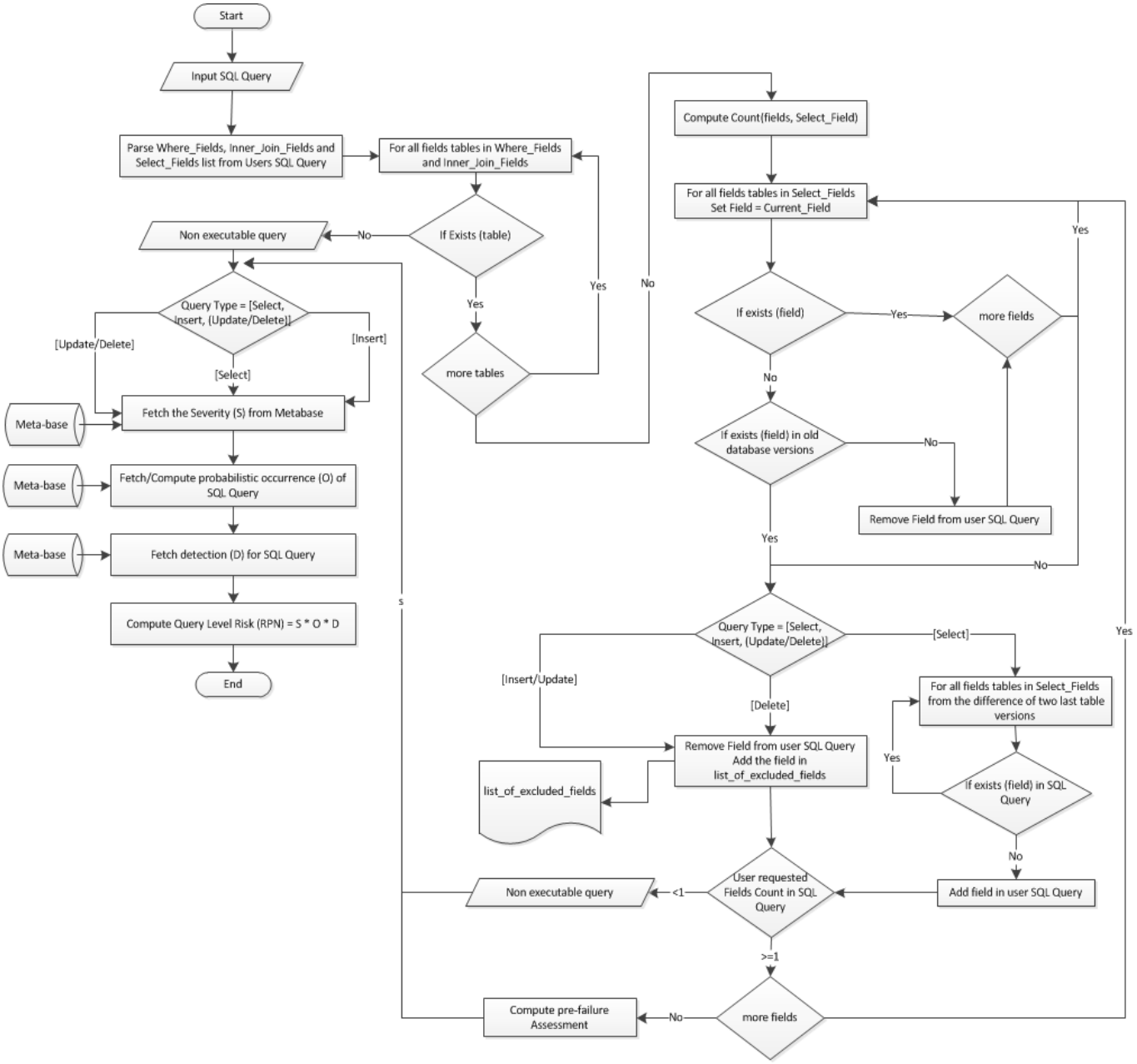


Figure 6.21 – Pre-failure assessment activity diagram

6.8 THE BIG PICTURE OF ROMMII PLATFORM

The big picture for the ROMMII platform is presented below in Figure 6.22. It can be seen that ROMMII platform is responsible to manage the proposed R&D data model which is supported with the MAM and SPM models to ensure multi-source data extraction, mapping and alignment for analyses purposes. The proposed R&D data model is populated using existing data sources through data wrappers. The data wrappers are ETL routines used to populate DWH from existing data sources. The R&D data model addresses the issue of data retention period and provides one year storage across multiple data sources. The proposed ROMMII platform allows the restructuring of R&D DWH, because it is not proprietary and we simply need to adjust data wrappers against the structural changes in data models. The ROMMII can also include the additional data structures including existing operational, DWH and DM data sources. This architecture advocates that we must move towards customized user applications to address the specific needs of R&D engineers to improve productivity.

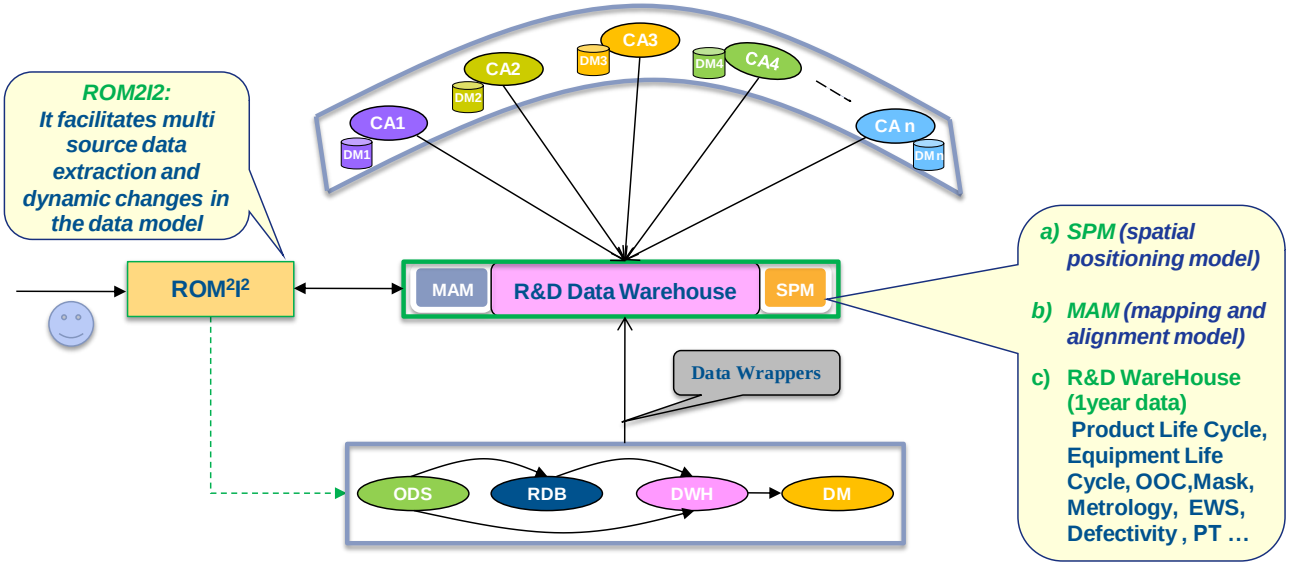


Figure 6.22 – The Big picture with ROMMII platform

6.9 RESEARCH SCHEMATIC AND ADVANCEMENTS (*ROMMII FRAMEWORK AND R&D DWH DATA MODEL*)

The research schematic and advancement are presented in the Figure 6.23. The ROMMII (SC3) platform and R&D data model (SC4) models are the key generic contributions made in this thesis which enable us to exploit huge data volumes and dimensions, and continuously evolve the existing data models for the inclusion of new dimensions. The significance of these contributions is that it performs pre-failure assessments on the existing single or multi-source data extraction and analyses utilities against potential structural changes in the data models. It ensures that all the appropriate measures are taken prior affecting the model evolution, resulting in the agility of existing utilities. In addition to this, it also ensures that all structural changes are well communicated to potentially interested users based on the user and application historical statistics. The ability to exploit huge data volumes and dimensions along with our proposed MAM and SPM models shall ensure timely value extraction and knowledge capitalization, which is a key to improve technology alignment and adoption lead times and costs.

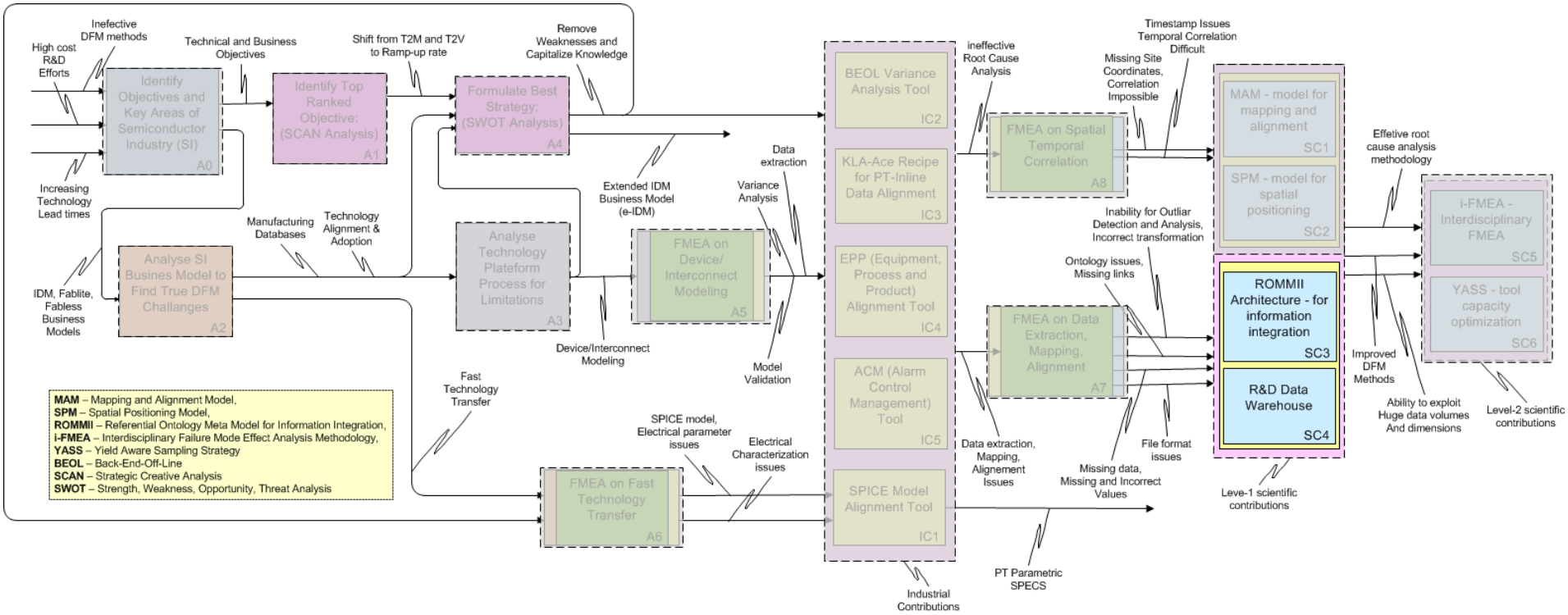


Figure 6.23 - The research schematic and advancement with ROMMII platform and R&D DWH model

6.10 SUMMARY AND CONCLUSIONS

The R&D engineers have always been complaining about the availability of insufficient data and dimensions to perform statistical analysis due to high storage and computational cost. The recent IT revolutions have resulted in decreasing storage and computational costs at higher performance. It has ensured the availability of huge data volumes and dimensions for R&D engineers, but they are still complaining that besides the availability, now they are not able to exploit these huge data volumes due to inconsistent data models and their unstructured evolutions. It is because of the fact that root cause analysis has shifted from the single-source to multi-source analyses and if multiple data sources have ontology issues and missing common identifiers then it is not possible to extract, map and align the multi-source data for effective root cause analysis. The production databases have converged to data archiving due to increasing data volumes; hence, they have varying data retention periods. It results in a big issue for the R&D engineer, because they need at least one year's data for the R&D analyses. The reason is that existing production data resources are equally being used for the R&D as well as production management and engineering purposes. Moreover they are of proprietary nature; hence, we do not have the right to alter or modify the data models. Any such effort directly results in the failure of associated data extraction and analysis tools.

The proposed ROMMII platform shifts unstructured data model evolutions towards structured evolutions and provides a pre-failure assessment upon any intended potential change in the data models. The impact of potential change is computed at the end users and application administrator levels and they are respectively intimated for potential changes in the applications to avoid failures and for information purposes. The application administrators comply with the potential changes and evolutions in the model and change the application to support extraction and analysis efforts to avoid potential failures. It also ensures that new data dimensions are always available for the engineers and moreover they are intimated upon any inclusion or exclusions of data dimensions. To avoid the data retention and proprietary issues, we have proposed R&D DWH data model, which motivates customized application development for engineering teams.

The proposed R&D DWH data model is a multi-dimensional data warehouse whereas the ROMMII model is implemented as a relational database. The ROMMII platform is modeled and presented using the UML and process flow diagrams. Its industrialization is highly specific to the type of database management system being used for the R&D data model and existing databases. It leads to a new research area as well that, how the agility provided by our proposed ROMMII and R&D data model can be extended, if R&D data model comprises of different databases for the efficiency and performance.

The potential industrialization of proposed generic solutions as *(i)* MAM, *(ii)* SPM and *(iii)* ROMMII model, is likely to result in increasing demand for metrology/inspection of engineering and R&D lots. It has direct impact on the existing metrology/inspection tools capacities, which are dedicated for the production lots. It is unlikely that we shall purchase new metrology equipments; hence, we need to spare the tool capacities if we want smooth integration of the proposed generic solutions. In next chapter, we present a yield aware sampling strategy, which spares the inspection capacities to be used for engineering and R&D purposes.

Chapter 7: Yield Aware Sampling Strategy (YASS) for Tool Capacity Optimization[16]

In the previous chapters, generic R&D solutions are proposed as the *(i)* MAM and SPM Models and *(ii)* ROMII Platform and R&D DWH model. The potential industrialization of these proposed models in an IDM-fablite business model shall result in additional requests for R&D lots metrology and inspection. It is highly likely that the measurement capacities run out because production lots have the priority to use these capacities to ensure product quality. In order to ensure that proposed solutions are industrialized, IDM-fablite business model must be provided with the methodology to generate additional measurement capacitates. The existing metrology/inspection strategies are being optimized using static, dynamic and smart sampling strategies based on the risk and associated delays. In this chapter we have proposed a yield aware sampling strategy, which is objectively focused on finding and inspecting the bad lots while moving the good lots to the next production steps. The equipment used in the SI are highly sophisticated and hold three different key levels of information as *(i)* alarms, *(ii)* states and *(iii)* meters data. This data provides key information about potential health of the equipment while it was processing production lot(s); hence, this data could easily be used to learn the predictive models for subsequent predictions. These reliable predictions in comparison with the static, dynamic and smart sampling strategies shall result in additional capacities to be used for R&D purposes.

Contents

[16] Shahzad M.K., Chaillou Thomas, Hubac S., Siadat A., Tollenaere M., A yield aware sampling strategy for inspection tools capacity optimization, International Conference on Artificial Intelligence, Las Vegas, USA 2012

7.1 INTRODUCTION

The proposed solutions as *(i)* MAM model, *(ii)* SPM Model, *(iii)* ROMMII Platform and *(iv)* R&D data model exactly match with the needs of the R&D engineers. It provides an opportunity for our engineers to perform accurate root cause analysis taking into account the newly emerging spatial variations. To best capitalize this opportunity, additional dies/sites must be inspected in addition to the existing 9-17 sites metrology/inspection strategy. This additional metrology requires more inspection tools or the available capacities that primarily serve to control production lots in an *e*-IDM business model. As a result we have two options: *(i)* either purchase new metrology/inspection tools or *(ii)* optimize the inspection capacities. It must be noted that the inspection/metrology tools in an IDM primarily serve the production to control the product quality with 100% inspection at each step hence, inspection tools quickly run out of capacities resulting in the production cycle delays. These equipments are costly and add up fixed cost, which is not welcomed. To best utilize the available inspection capacities, we use sampling strategies (static, dynamic and smart), which are based on the risk, delays and capacities. Industrialization of such blind strategies is still a big question that might result in skipping bad lots to move to the next production steps, resulting in waste of resources and customer dissatisfaction.

The industrialization and success of proposed generic R&D solutions depend on the additional inspection capacities; hence, in order to generate these inspection capacities we move one step ahead and propose a yield aware sampling strategy that predicts all production lots as good, bad or suspected lots. These predictions are made based on the likely yield loss with predictive state (PSM) and alarm (PAM) models. The steps involved in this strategy are: *(i)* classify potentially suspected lots, *(ii)* cluster and/or populate suspected lots in the priority queues and *(iii)* apply **L**ast **i**n **F**irst **O**ut (LIFO) to optimize capacities. It provides sufficient spare metrology/inspection capacities that can be used for extended R&D purposes.

7.2 METROLOGY/INSPECTION AND PRODUCTION TOOLS CAPACITIES ISSUES

The SI has revolutionized our daily lives with electronic chips that can be found in almost all the equipments around us and follows the slogan smaller, faster and cheaper driven by Moore's law [Moore, 1998]. It postulates that the number of transistors shall double in every 18 to 24 months at reduced cost and power. Since then the SI has kept its pace as per Moore's law by continuously investing in R&D for the new technologies. New equipments are being manufactured to support and keep up with the emerging demands and the pace defined by the Moore's law. The equipments are highly expensive; hence, decisions to purchase new production equipments are based on business strategy and estimated ROI (return on investment). The metrology/inspection tools carry fixed costs and often phase out or need changes to cope up with new technologies; hence, capacity optimization strategies are used to balance inspection load instead of purchasing new metrology/inspection tools.

7.2.1 Why do we need 100% inspection?

The IC chip manufacturing has become a complex but expensive production process resulting in process control challenges to find lots with yield issues before they consume the expensive production resources. An electronic chip undergoes approximately 200 operations, 1100+ steps and 8 weeks of processing prior to packaging and assembly. To ensure the product quality, metrology /inspection steps are added within the manufacturing flow, almost after every manufacturing step. The design of an economical control for a production process has always been an interesting research area that was initiated by [Duncan, 1956] but it lacks robust results and requires a balance between controls and their costs. Hsu proposed a control plan with the concept of skipping by decreasing control frequency as compared to a 100% inspection plan that indirectly added risk of skipping bad lots to the next steps [Hsu, 1977]. Reynolds et al, (1988) presented, that sampling size and frequency as two levels of controls are the better solution to quickly identify the issues while minimizing the cost of errors [Reynolds et al., 1988]. Sanos introduced the concept of SPC in the semiconductor industry that served as the basis for an updated control plan [Spanos, 1991]. Many approaches have been proposed for an adaptive control: *(i)* sampling strategy based on the number of wafers passed on metrology tools [Raaij and Verhallen, 2005], *(ii)* updating control plan based on process excursions [Mouli

and Scott, 2007] and *(iii)* updating control plan based on risk encountered during productions. Coledani and Tolio have designed a buffer for control machines taking into account the quality and cycle time expectations and it is the latest contribution in research regarding control plans [Colledani, 2008] and [Colledani and Tolio, 2009].

7.2.2 Why additional capacities?

It is known that inspection tools have limited capacities; hence, optimal capacity utilization in a high product mix is a key for success. There are some critical steps as well where the delay due to inspection capacity limitation might have a strong impact on the next production step, so these priority products must be inspected before other products in queue. Limited metrology/inspection tools capacities has a strong impact on the production cycle times; hence, an efficient sampling and control strategy is required to optimize the capacities and exploit economic benefits. In addition to this, there is also a strong need for the additional metrology/inspection capacities based on the shift from MFD to local DFM efforts in our proposed extended IDM (e-IDM) model. The objective is to capture newly emerging spatial variations and model the systematic variations and drifts into rules and/or models for subsequent use during CAD simulations to assess yield and manufacturability.

7.2.3 What is wrong with the sampling strategies?

The existing strategies are classified as static and dynamic where static sampling [Lee, 2002] selects the same numbers of lots but dynamic sampling [Raaij and Verhallen, 2005] selects the number of lots for inspection, based on the overall production. Smart sampling is a new approach that samples lots by taking into account the risk associated with production tools, inspection tools capacities and delays to dynamically minimize the wafers at risk [Dauzere-Peres et al., 2011and Sahnoun et al., 2011]. It is a better approach than the static and dynamic strategies. In this strategy, if a lot in the waiting queue is controlled and it passes the inspection step then all lots in the waiting queue, processed before this lot are removed with a confidence that they are good lots. However none of them provide an evidence for a likely yield loss against sampled lots resulting in skipping the suspected lots to move to the next process steps. We need a yield aware strategy to classify good, bad and suspected lots to reduce the inspection load significantly followed by an optimization strategy that exploits the production resources against limiting inspection tools capacities.

A generic production process is presented in Figure 7.1 where lots are processed, controlled and/or skipped at the production and inspection tools to avoid bad lots moving to the next steps. The sampling and scheduling strategies in this production process are focused on priority queues and lots are sampled based on the risk levels associated with the product or process. These strategies are an effort to balance the difference between production and inspection tools capacities. These strategies can be viewed as a blind strategy with a high risk of skipping bad lots to the next steps. Our proposed YASS strategy empowers the control with PSM and PAM models to filter good and bad lots followed by capacity optimization. We argue that the focus should be shifted to find and control bad products rather than inspecting 100% product using blind inspection strategies that do not differentiate between bad/good products.

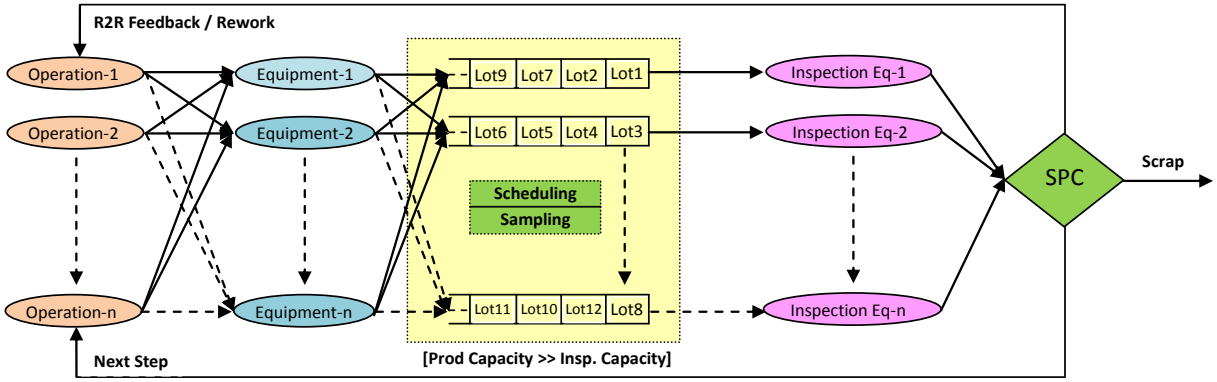


Figure 7.1 - Generic production process with existing sampling strategies

It is evident that the inspection capacity allocation problem is linked with the process control plans in the semiconductor industry. If mistakes are committed like skipping the bad lots to the next production steps then the consequences are evident in terms of customer dissatisfaction and costs. It is due to the fact that existing approaches do not provide any evidence of the likely yield loss. The additional capacities can be spared from the production control if a reliable inspection strategy is adopted that avoids the inspection of good lots. The existing sampling strategies do not guarantee a likely yield loss; hence, to ensure smooth integration of the scientific contributions (SC1, SC2 and SC3) in the e-IDM model a 3-step yield aware sampling strategy is presented that provides us with the additional inspection capacities. In this strategy, the focus is on the identification of bad or suspected lots while moving good lots to the next production steps. It shall not only increase the optimal utilization of the inspection capacities, but also provide us additional capacities for the R&D purposes. Our proposed yield aware 3-step strategy uses the predictions made by the PSM and PAM models. Based on the predictive output combinations, lots are either added to the priority inspection queues or moved to the next production steps followed by the capacity optimization.

7.3 PROPOSED 3-STEP YIELD AWARE SAMPLING STRATEGY (YASS)

The proposed yield aware sampling strategy is based on the principle that good lots are skipped for the next production steps and bad lots are strictly controlled to avoid resource wastage and customer dissatisfaction. We start with the classification of production lots as good, bad or suspected and only bad or suspected lots are potentially inspected while permitting the good lots to move to the next production steps. The waiting queues are established and LIFO based optimization is applied on the lots, which are clustered based on the product type and process recipe within the queues. If a potential lot results in a bad lot, then all lots in the same cluster manufactured before the inspected lots are scraped.

The proposed strategy is presented in Figure 7.2. In our scenario, the equipment is composed of 1...n modules in the parent child relationship. These modules are further classified as critical/non-critical and shared/non-shared elements. All the modules are characterized by meters, alarms and states, which are recorded in the databases during the production operations. The alarms are generated at the module level; however, based on these alarms the automation system changes the states of the child and parent modules. The changes in the states of the child modules serve as the basis for change in the state of parent modules. It represents the equipment health and such information can accurately predict a likely yield loss in production due to unpredictable equipment behavior. The meter data triggers the preventive maintenance operations; however, the module level meter data shall be included in the future to compute the weighted probabilities for the state and alarm level predictions. The alarms and states data represents the status of the production process and equipment health respectively; hence, they are used to learn predictive alarm [PAM] and state [PSM] models.

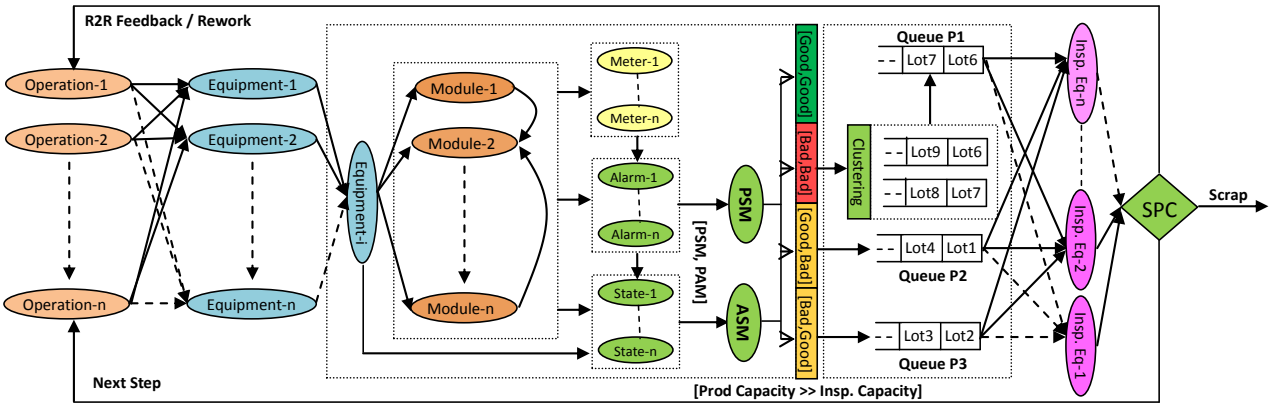


Figure 7.2 - Methodology with predictive state (PSM) and alarm (PAM) models

We have used only the states and alarms data; however, the module level meter data shall be included in the future to compute the weighted probabilities to refine the state and alarm level predictions. The first step in the inspection strategy is to classify the good, bad and/or suspected lots. In this step, we start

with the exploitation of the historical equipment states, alarms and SPC (statistical process control) data from the process, maintenance and alarms databases to build predictive state [PSM] and alarm [PAM] models. These models [PSM, PAM] are then used to classify the new production lots as good and/or bad lots and generate four possible outputs: *(i)* [good, good], *(ii)* [good, bad], *(iii)* [bad, good] and *(iv)* [bad, bad].

The good production lots [good, good] are moved to the next production steps without metrology and bad lots [bad, bad] undergo the 100% inspection and their results are used to update the prediction [PAM] and [PSM] models. The suspected lots [good, bad] or [bad, good] are clustered (2nd step) based on the equipment, product and recipe. It is followed by a priority queue allocation algorithm (3rd step) that enters the suspected lot clusters into priority queues for further optimization based on LIFO (last in first out) principle. It states that if a suspected lot defies the prediction upon inspection then all the lots in the same cluster shall be subjected to 100% inspection otherwise, the cluster members are skipped. The predictive state and alarm models [PSM, PAM] are updated with the feedback against all coherent and incoherent predictions as good and/or bad examples. It provides us an intelligent way to reliably sample only bad or potentially suspected lots followed by priority queuing and optimization for the economic benefits. The step-1, where predictive [PSM, PAM] models are learned, is implemented with a heuristic algorithm as presented in section 7.4. These models classify new production lots as good, bad or suspected lots. These PAM and PSM models correspond to the step-1 of the proposed methodology. The success of this approach depends on the accuracy of learning and classification algorithm for PSM and PAM models. We propose two heuristics, first for the PSM and PAM models and second to cluster the suspected lots and applying the queue optimization. A tuning parameter is also provided at the discretion of the user to control the PSM and PAM prediction confidence levels.

7.3.1 Heuristic Algorithm for [PAM, PSM] Models [Step-1]

The alarms and states data collected from the equipment is not trivial to use with the existing classification and pattern sequence learning algorithms; hence, a heuristic algorithm is proposed to predict the given wafer W*j* as a good or bad. The time required for a production operation varies from 30 minutes to 4 hours and during operation a series of a set of equipment states and alarms are generated. It is evident that data collected is a matrix of sets of alarms and states. Our proposed algorithm is simple and is based on probabilistic likelihood with the previous inspections. We start with the presentation of the variables used in the algorithm as under (Table 7.1):

Variable	Description	Variable	Description
	G → good, B → Bad, C → Confusion		E^i → Equipment (i), W^j → Wafer (j)
	Matrixes of [G, B, C] alarms or states data for all equipment $E^{1..n}$ and Wafers $W^{1..n}$		Global Support of alarms and states for equipment E^i and wafer W^j
	Local Support for each set of alarms and states in the Matrix for equipment E^i and wafer W^j		Target defined by the user for each equipment E^i to be used for model predictions
	Matrix for alarms and states, historical data for equipment E^i		Matrix of alarms or states sequence for equipment E^i and wafer W^j

Table 7.1 - Description of [PAM, PSM] models variables

The T[Ei] is a tuning parameter and is defined by the user to tighten the prediction and it is used to optimize the inspection capacities utilization, if required. For example, if user sets its value to 50%, it means that computed likelihood shall be compared with this target prior to [good, bad] predictions. It is also important to note that the historical data shall be used to populate alarms and states matrixes for all equipments; however, a confusion matrix [C] shall consist of duplicated states and alarms sets found in both good and bad matrixes.

Proposed algorithm for the prediction of [PSM, PAM] model is as under. We start with the predictive model learning by computing good, bad and confusion matrixes of alarms and states sequence sets for all equipments as shown in step-1. It shall be used during the computation of the local support. In Step-2, the local support for each set of alarms and states sequence for a given wafer Wj is computed by counting the similar set of alarms or states sequences in [G,B,C] matrixes. This count is divided by the respective count of sequences in [G,B,C] matrixes to get the local support values for each sequence of alarms and states in wafer Wj . Further global support for the wafer Wj is computed by multiplying the local supports computed for each set of alarms and states sequence against [G,B,C] matrixes with the sum of computed local supports. The results are summed at [G,B,C] level for the prediction by comparing it against the T[Ei] in step-4. If the global support is smaller than the T[Ei] then the benefit of the doubt is given to the prediction by adding the global support for [C]. Said algorithm for [PSM, PAM] models is presented as under:

Step-1: Compute for state and alarms data for each equipment E^i

Step-2: Compute local support for each set of alarms/states sequences for a given wafer W^j and equipment E^i

Step-3: Compute global support for a given wafer W^j and equipment E^i

Step-4: Predict [PSM, PAM] output for the given wafer W^j

7.3.2 Example for [PAM, PSM] predictions

To demonstrate the said algorithm, a simple example is presented with [G,B,C] matrixes for alarms. The tables 7.2 and 7.3 present an example of the alarms matrix for equipment Ei for good and bad yield respectively. The table 7.4 presents a confusion matrix that accumulates alarms found in both good and bad yield; hence, they cannot be directly used during the prediction of good or bad lots.

Alarm Matrix for Equipment Ei [Good Yield]				
T_ID	Module1	Module2	Module3	Module4
1	107(0x6B)			
2	384(0xF04)			
3		384(0xA04)		988(0xA04)
4		384(0xA04)		
5		384(0xC04)		
6				
7	384(0xF04)	384(0xD04)		
8				84(0xC09)
9	384(0xF04)	384(0xD04)		38(0xF14)
10	384(0xF04)	384(0xD04)		38(0xF14)

Table 7.2 - Alarm matrix for equipment Ei [Good Yield]

Alarm Matrix for Equipment Ei [Bad Yield]				
T_ID	Module1	Module2	Module3	Module4
1	107(0x6B)	384(0xA04)		
2	384(0xF04)			84(0xC09)
3		388(0xA04)		988(0xA04)
4		374(0xA04)		
5	14(0xC09)	384(0xC04)		24(0xC09)
6				
7	384(0xF04)	312(0xD04)		
8				84(0xC09)
9	384(0xF04)	172(0xD04)		38(0xF14)
10				34(0xF22)

Table 7.3 - Alarm matrix for equipment Ei [Bad Yield]

Alarm Matrix for Equipment Ei [Confusion]				
T_ID	Module1	Module2	Module3	Module4
1	107(0x6B)			988(0xA04)
2		384(0xF04)	988(0xA04)	
3				988(0xA04)
4	38(0xF04)			
5		384(0xC04)		
6				
7	384(0xF04)			312(0xD04)

Table 7.4 - Alarm matrix for equipment Ei [confusion]

Alarm Matrix for Equipment Ei, Wafer Wj				
T_ID	Module1	Module2	Module3	Module4
1	384(0xF04)	172(0xD04)		38(0xF14)
2		384(0xA04)		988(0xA04)
3		384(0xC04)		
4	38(0xF04)			84(0xC09)
5				34(0xF22)
6		384(0xC04)	988(0xA04)	

Table 7.5 - Alarm matrix for equipment Ei, wafer Wj

These tables are constructed to be further used by the PAM and PSM models during predictions, local and global support computations. The computed results against the potential lot Wj for prediction are detailed in Table 7.5 with 54% likelihood of the lot being a good lot (Table 7.6). The computation of global support for the PSM prediction is same as the PAM model.

Local Support			Global Support		
[G]	[B]	[C]	[G]	[B]	[C]
0.2	0	0	0.27		
0.1	0	0	0.13		
0.1	0	0	0.13		
0	0.1	0		0.13	
0	0.1	0		0.13	
0	0	0			
0	0	0.143			0.19
Total = 0.743			0.54	0.27	0.19

Table 7.6 - Local and global support for wafer Wj [54%, Good]

7.3.3 Clustering and priority queue allocation [Step-2 and Step-3]

Based on the predictions from [PSM, PAM] models (section 3.1), we follow the 2nd and 3rd step in the proposed yield aware inspection strategy where suspected lots are clustered and added to the priority queues followed by LIFO optimization. It is presented with a simple flow chart (Fig.5). All production lots with

prediction combination [PSM'Good and PAM'Good] are simply skipped whereas other lots are populated in the priority queues P1, P2 and P3 where P1>P2>P3. Lots with the combination [PSM'Bad] and PAM'Bad] are first clustered based on the similarity of product, technology and recipe followed by the population of last lot from each cluster in P1. If an inspected lot from the P1 validates the model prediction then its respective cluster members are simply scrapped; otherwise, each member of the cluster is inspected and the predictive [PSM, PAM] models are updated. In case of differences in the model predictions, lots are declared as suspected and are populated in the priority queues P2 [PSM'Good and PAM'Bad] and P3 [PSM'Bad and PAM'Good]. The lots from P2 are sequentially inspected; if it defies the models then all lots processed before the inspected lot in P2 are given the benefit of the doubt and are skipped. If a lot inspected from the P3 defies the model then all the respective cluster members are inspected and predictive [PSM, PAM] models are updated for coherences and incoherencies.

It is evident from the above discussion that the [PAM] predictions have higher priority than the [PSM] predictions based on the two facts, (i) child modules influence the statesof their parent modules, hence prediction model developed at the module level might have a dual impact and (ii) alarms count and duration result in the change of state of the modules. The states data is an aggregation of the alarms data at module level; hence, alarms data provide more low level detailed information with no influence on the alarms of parent modules. Based on these facts in this proposed methodology we have given higher priority to PAM prediction while performing information fusion of modules alarms and states data. The [PSM, PAM] prediction weights shall be defined in the future by including the meter data. The meter data is very critical because the values of the meters initiate the preventive maintenance actions on modules. We believe that it shall play a pivotal role in defining the prediction accuracies of the PAM model. The alarms and resulting prediction shall get more weight when meter data falls within the distribution of previous preventive maintenance actions.

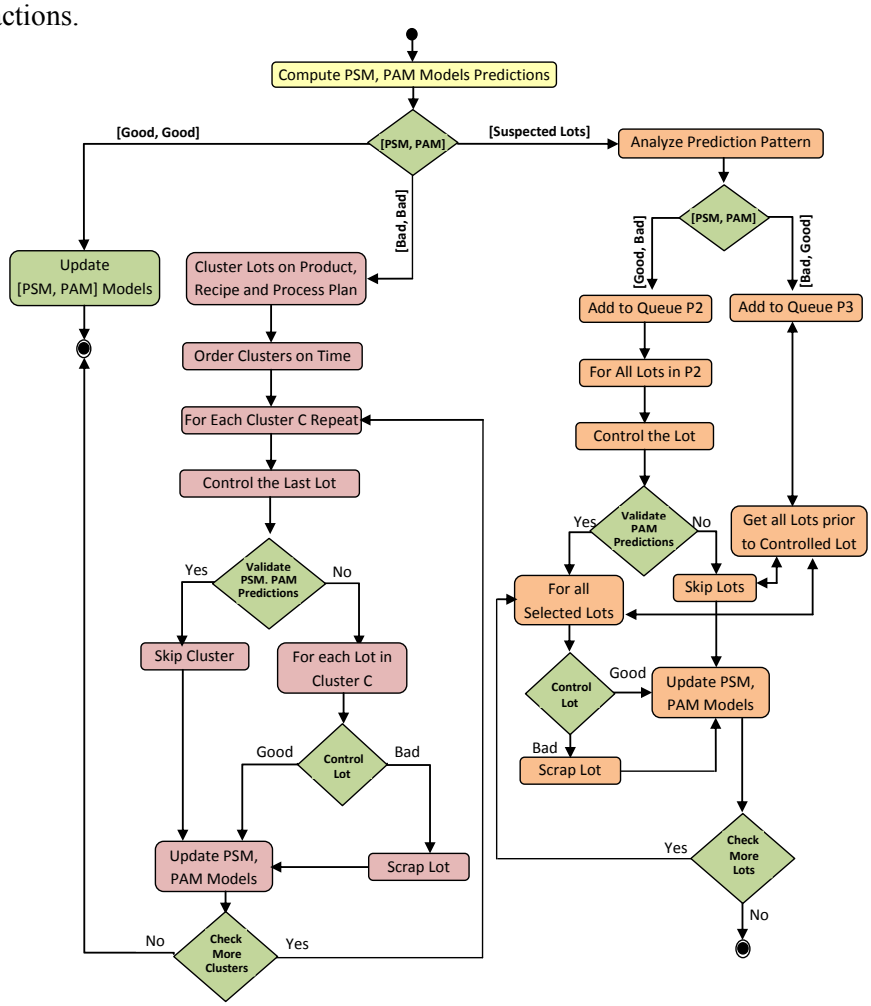


Figure 7.3 - Flow chart for clustering and queue allocations

7.4 DATA MODEL TO SUPPORT [PAM, PSM] MODELS

The biggest challenge in building and deploying these PAM and PSM predictive models is the multi-source data extraction, alignment and preprocessing from SPC, maintenance, process and alarms data sources. To facilitate this, a data model (Fig.6) is proposed with the ASCM (Alarm and State Control Management) tool that allows engineers a quick extraction and alignment of the data. In this data model, the equipment (equipment class) is composed of modules (module class) and every module has a state (state_history class) and alarms (module_alarms class) history. The usage meter data is also available (usagemeter class) but in this paper this data has not been used. The parent-child associations between modules are captured by the parent_child_relation class. A product (product class) is manufactured using multiple lots (lot class) but follows a single process plan (process_plan class). The process plan has multiple operations (process_operation) and each operation can have multiple steps (process_steps). The process step undergoes different step runs (step_run class) as the production or metrology runs. The production step runs are associated with the equipment that has the capability (equipment_capability class) to perform the process steps. This data model is translated into a relational database, which is implemented using SQL server to support the ASCM tool (IC5, Appendix F).

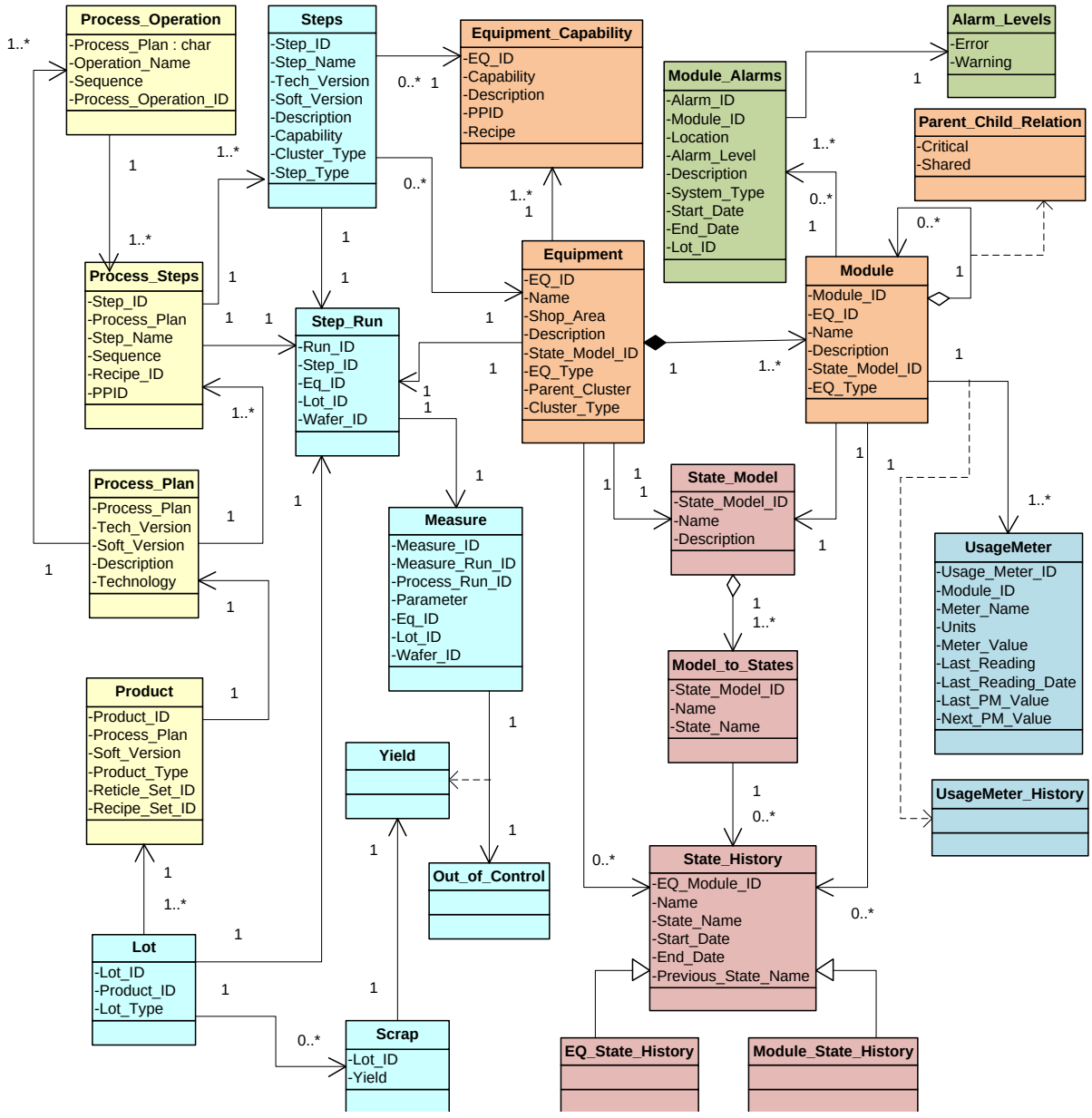


Figure 7.4 - Data model to support [PAM, PSM] models

7.5 RESEARCH SCHEMATIC AND ADVANCEMENTS (*YASS STRATEGY*)

The research schematic and advancement is presented in Figure 7.5. The YASS strategy is the key contribution made in this chapter with an objective to generate additional measurement capacities. The YASS is a 4-step methodology and it is supported by PSM and PAM predictive models. These models predict each and every production lot as good, bad or suspected and based on proposed clustering and queue optimization algorithms, we skip, inspect, scrap or partially control production lots. The skipped lots are the good ones and this strategy is based on the simple principle that we should find and inspect the bad lots whereas good lots must be pushed up to the next processing steps to reduce cycle times. In addition to cycle time gains we get additional inspection capacities that shall be used to support the industrialization of our proposed generic R&D models. The success of this strategy depends on the accuracy of PSM and PAM models. These models are learned in the due course of production. We have full inspection load in the start of the production because PSM and PAM models have minimum accuracy but as more and more production lots are processed, we get accurate PSM and PAM models, that result in additional inspection capacities. The granularity of the PSM and PAM models are still to be determined. At present we have hypothesized that PSM and PAM models shall be learned at production step level for each equipment; however, in future the similarity between production plans can also be used to cluster these PSM and PAM models.

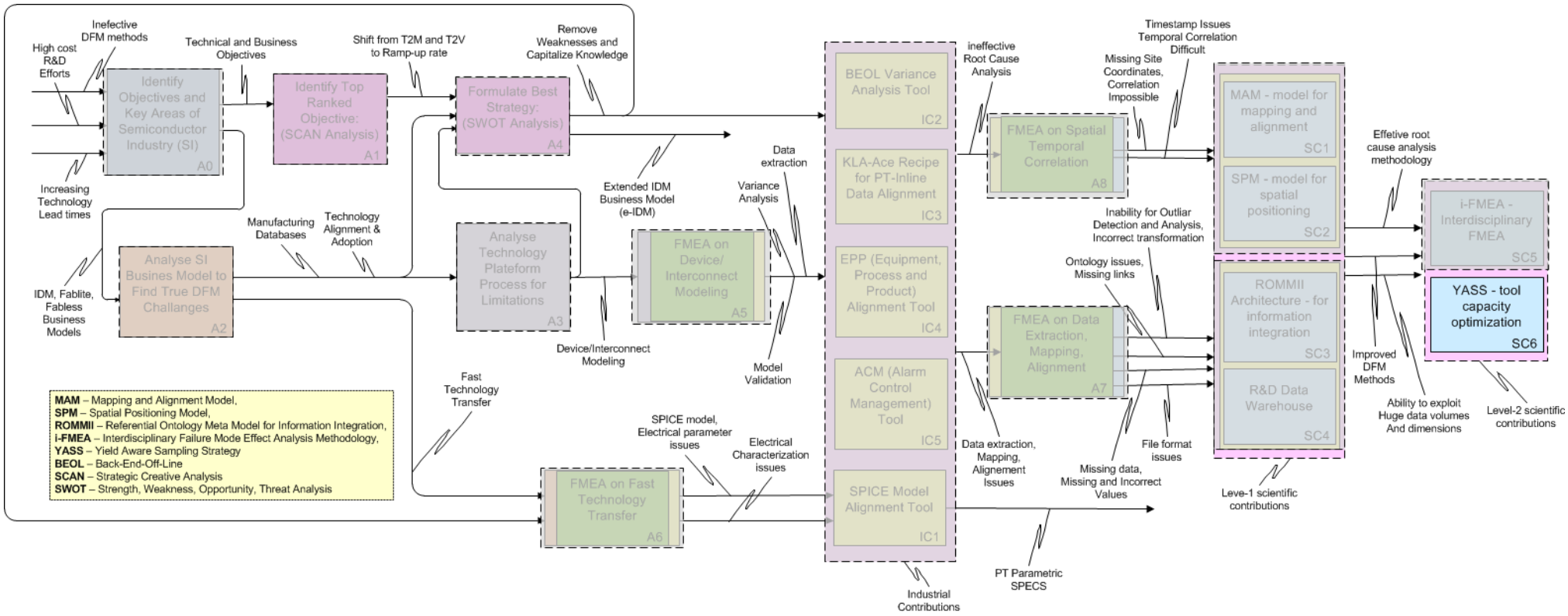


Figure 7.5: The research schematic and advancement with YASS sampling strategy

7.6 SUMMARY AND CONCLUSIONS

The YASS strategy has significant advantages over existing static, dynamic and smart sampling strategies as it identifies the production lots with likely yield loss; hence, good lots are skipped to next production steps resulting in additional tools capacities. This strategy also provides a confidence and is directly linked with the alarms and states data coming from the equipment. The alarms and states data describes a precise equipment condition and evolution; hence, it could be easily linked with its impact on the yield. The advantage is not limited to additional capacities but it also provides an optimized utilization of the inspection capacities. The $T^{[Ei]}$ (tuning) parameter provides a control to the end users if they want to tighten the sampling or decrease the sampling rates in good periods.

It has one disadvantage of granularity of the prediction models, which we believe has to be at the product and equipment levels. In the start these models are blank; however, they can be used for rough predictions based on the equipment and process similarities. The product plays a significant role; hence, these models need to be adopted for each new product. So in all cases to adopt these models we need to start with 100% inspection till the maturity of the models. The long term benefit for the capacities optimization can be utilized by the R&D engineers. To extend the proposed approach in future, we believe that equipment must be made capable with built-in modules to automatically learn and manage these models at the product level and as soon as a product finishes a process step, it is tagged for inspection or a skip.

Discussions and Conclusions

The semiconductor industry is the most fragile, fastest growing and most competitive manufacturing domain. It is characterized by the cyclic demand patterns with a positive CAGR (+8.72) that guarantees cumulative demand growth. It motivates us to continuously invest in R&D efforts so that we are ready to respond to the growing market with new, fast and high value but low cost products. The SI, since its birth, has followed the industrial slogan << smaller, faster and cheaper >> driven by Moore's law that postulates doubling transistor density every 18-24 months. This miniaturization trend has led the emergence of technical (design and manufacturing interface) challenges that often result in serious manufacturability and yield level issues. It has also resulted in the shift of business objectives from time-to-market and time-to-volume towards ramp-up-rate, which can be achieved by introducing *(i)* new technologies every 2 to 3 years, *(ii)* technology derivative(s) and/or *(iii)* current technology improvements.

The new technology is developed in an alliance to share R&D costs and model emerging design and manufacturing complexities, supported with innovation in design, material, process and equipment. The cost and lead times involved in a new technology are 5-7 B$ and 2-3 years respectively, as modeling design and manufacturing issues are highly complex and has turned into a high cost R&D activity. The competitiveness, however requires a quick technology alignment and/or adoption whether it is the new technology, technology derivatives or improvements. The manufacturing data collected across the production line is the key towards technology derivatives and improvement efforts success, besides new technologies. It requires dynamic and effective production data exploitation so that the new technology derivatives and improvements can be quickly aligned and adapted at reduced costs.

The design for manufacturing (DFM) is a well-known approach in the manufacturing domain where manufacturing variations are transformed into rules and/or models for subsequent use during CAD simulations to assess potential manufacturability and yield prior to prototyping or production. It helps us in finding the first time correct design, which effectively reduces the lead times and avoids design respin costs. It was adopted by SI in 1980 as a yield enhancement strategy that worked very well till 250nm technology, but beyond this it has become ineffective. The ineffectiveness means that it is difficult to classify manufacturing drifts and variations as random or systematic patterns followed by their transformation into rules and/or models. There is nothing wrong with the DFM methods, rather, the issues lies with the quality of input data, which is fed to these methods. So, if we can improve the quality of the data, we can put DFM back on track to help us in reducing the lead times and costs associated with the technology derivatives and improvement initiatives for the competitive advantage.

The above discussion highlights the need for literature review across three domains as *(i)* the SI business model evolutions and challenges, *(ii)* role of DFM methods in the success of SI and *(iii)* information integration needs to enable dynamic exploitation of the manufacturing data collected across the production line. We found transformation of traditional IDM business model into fabless and fablite models to capture maximum market share in the SI market demand with a positive CAGR. This transformation is truly in line with shifts in business objectives to address the growing market needs but recent shift towards the ramp-up-rate needs its realignment. The miniaturization driven by Moore's law has resulted in the emergence of new design/manufacturing interface complexities. The spatial variation on the wafer surface is a recent phenomenon, which is critical to the ineffectiveness of the DFM methods. The data collected across the production line is categorized as measurement data (PT, Inline, EWS and Defectivity, etc.) and contextual data (WIP, product/equipment life cycle, SPC, etc.). The existing DFM methods are based on the single source analyses, which are insufficient to capture the spatial variations; hence, the need to shift from single to multi-source data analyses to improve DFM effectiveness. It is to be further strengthened by moving from the site level analyses towards die and test structure position based analyses as well. The wafer orientation due to test structure positions and measurement reference frames often results in varying site/die level coordinates, which requires a generic solution because engineers spend huge amount of time in mapping and aligning site/die level data. It often results in bad alignment, which adds to the DFM ineffectiveness. The

DFM methods cannot be improved in the present emerging spatial variations until the shift from single to multi-source data analysis at site, die and test structure position based analyses. The methodology for data mapping and alignment cannot work alone until it is supported with the data. The recent revolutions in IT have enabled low cost storage and exploitation of huge data volumes. It provided the opportunity to continuously restructure the data models so that new data dimensions can be added to be further used during analyses.

This thesis is objectively focused in putting DFM back on track. There has been a transformation of SI business models and shift in business objectives based on market dynamics; hence, the need to first analyze the existing business models against a recent shift towards ramp-up-rate. The objective is to propose an extension in the existing business models to support the new objective as well as a methodology to synchronize the business models with the future shifts in business objectives. The SCAN analysis is selected as the methodology to identify the top rank objectives followed by its alignment with SI business model based on extension or modification. The SCAN analysis highlighted the leadership position and quick ramp-up-rate as the top ranked objectives. The SWOT analysis resulted in the adoption of strategies to mitigate weaknesses and exploit the opportunities. The best known IDM-fablite business model is bench marked to assess its potential success against the shift in business objectives and we identified *(i)* fast technology transfer, *(ii)* manufacturing databases and *(iii)* effective root cause analysis (R&D) as the key improvement areas for its compliance with recent shift in business objectives along with strategy to mitigate weaknesses. The analysis was further extended to technology development process, which highlighted the *(i)* data extraction, *(ii)* alignment and *(iii)* pre-processing due to ontology issues and *(iv)* missing database links as the key weaknesses for mitigation. These weaknesses are classified as the main reasons for DFM ineffectiveness based on the fact that due to these challenges R&D engineers are not able to exploit manufacturing data sources besides huge data volumes and dimensions. The proposed e-IDM fablite business model is an extension of existing IDM-fablite model where these weaknesses are removed and engineers can potentially exploit the production resources to perform multi-source site/die/position based root cause analyses.

Compliance with the proposed e-IDM business model necessitated finding root causes against respective identified weaknesses (failure modes) grouped as *(i)* ineffective root-cause analysis (infield and scribe line test structure positions) and *(ii)* data extraction, mapping and alignment. We have proposed and used *i*-FMEA approach to identify the cyclic root causes against these failure modes. The proposed 4-step methodology helps identify the cyclic root causes that result in failure and need generic R&D solutions instead of operational fixes as proposed by traditional FMEA approach. The root causes clearly demonstrate the inability of R&D engineers to exploit huge volumes of data and allow the inclusion of new data dimensions in existing data models. The proposed *i*-FMEA approach helps in tracing back the potential root causes in other business functions. The key cyclic root causes identified are the *(i)* data format issues, (ii) unstructured evolution of data models and *(iii)* missing database links and dimensions that need generic R&D fixes. The SCAN and *i*-FMEA methodologies are suggested to be used upon the detection of any change in the business environment to quickly align the business model and find weaknesses along with respective root causes for, which generic R&D solutions are critical.

The effective root cause analysis is the key to improve the existing ineffectiveness in the DFM methods and can be defined as the ability to analyze the data to find answers against drifts and/or variations. These drifts and variations are further classified as systematic or random for their transformation into rules and/or models. At present, engineers are facing huge difficulties in data mapping and alignment to perform multi-source analyses. It is based on the fact that measurements result in varying coordinate systems due to test structure orientation on the wafer. The wafer is rotated prior to the metrology steps and its position is monitored using notch or flat positions. The varying coordinates issue is further complicated with the different metrology reference frames. These issues have a significant impact on stack definition, model correction and validation steps during the interconnect modeling process.

The proposed generic R&D solutions include MAM and SPM models, which take into account the information from the wafer, its mask and test structure positions, and normalize the site (PT and Inline) and die (EWS and Defectivity) level measurements. The MAM model provides fixes to sites and dies level measurements so that they can be used for the correlation purposes. The SPM model provides ability to correlate parameters based on the position of the respective test structures whereas proposed die/site qualification model enables R&D engineers to perform site/die level root cause analyses. The above generic models remove all inconsistencies with the measurement coordinates so that engineers could perform multi-source root causes analysis. Recent IT technologies have enabled huge data volumes storage and exploitation. The engineers need multi-source data coming from multiple data sources for effective root cause analysis; hence, the biggest challenge faced at this level is to allow restructuring of data models.

The SPM model is highly efficient; however, there is still room for improvement to obviate computing the positions w.r.t. each wafer for all the test structures. The structural position map can be computed and saved as a search pattern, which can be used for mapping and alignment of measurement data prior to data analyses. We can apply a similar search pattern based idea for the MAM model based on the fact that wafer position and measurement reference frame have discrete values, which can be used as a reference for mapping and alignment in an effort to reduce the computational costs and efforts.

The R&D engineers have always been complaining about the availability of insufficient data volumes and dimensions to perform statistical analysis due to storage and computational cost issues. The recent IT revolutions have resulted in decreasing storage and computational costs at higher performance. It has enabled the availability of huge data volumes and dimensions to the R&D engineers, but they are still complaining that besides the availability, they are not able to exploit huge data volumes and dimensions due to unstructured data model evolutions and inconsistent data models. It is because of the fact that root cause analysis has shifted from single-source analysis to multi-source analysis. If multiple data sources have ontology issues and missing common identifiers then it is not possible to extract multi-source data for the root cause analysis. The production databases have converged to data archiving due to increasing data volumes; hence, they have different retention periods. It causes a big issue for the R&D engineer because they need at least one year's data available all the time for the R&D purposes. The reason is that existing production data resources are equally being used for the R&D and moreover they are of proprietary nature; hence, the inability or right to alter or modify the model. Any such effort directly results in the failure of associated data extraction and analysis tools, which is not less than a disaster.

The proposed ROMMII platform shifts unstructured model evolutions towards structured evolutions and provides a pre-failure assessment upon any intended potential change in the data model. The impact of potential change is computed at end users and application administrator levels and they are respectively intimated for potential changes in the applications to avoid failures and for information purposes. The application administrators comply with potential change, and evolution in the model is ensured to support the extraction and analysis utilities to avoid failures. It also ensures that all new data dimensions are available at all times for the R&D engineers and moreover, they are intimated of any inclusion or removal of data dimension from the R&D data sources. To avoid the data retention period and proprietary issues, we have proposed an R&D data model, which shall result in the customized application development for each R&D group. It shall be managed and controlled by ROMMII framework where we are able to manage the data dimensions.

The R&D data model is a multi-dimensional data warehouse whereas the ROMMII model is implemented as a relational database. The ROMMII platform is modeled and presented using the UML and process flow diagram. A simple database with three tables is used to present the potential results; however, its industrialization is highly specific to the type of database management system being used for the R&D data model; hence, its industrialization requires the brainstorming sessions on the choice of DMS first. It leads to a new research area as well, that how the agility provided by proposed ROMMII platform can be extended if R&D data model comprises of different databases.

In this thesis we have demonstrated that the existing IDM-fablite business model does not support the recent shift in the business objective i.e. ramp-up-rate; hence, we have proposed the e-IDM fablite business model to fully comply with the new business objectives. The root causes against weaknesses found during the business model analysis are identified using our proposed *i*-FMEA methodology. The generic R&D solutions as the MAM, SPM, ROMMII and R&D DWH model are proposed instead of operational fixes. The industrialization of these solutions shall result in extended requests from the R&D engineers for additional measurements at site and/or die levels. It shall challenge the production and inspection capacities of an IDM. The proposed solutions can only be industrialized if they do not add additional fixed cost; hence, we have further analyzed the inspection capacities and existing strategies to assess if we could spare inspection capacities for R&D purposes. The existing static, dynamic and smart sampling strategies for the inspection are blind strategies and are based on the risk, delay and capacities. The proposed 3-step YASS strategy uses PAM and PSM prediction algorithms based on alarms and states data. The alarms and states patterns likely predict the potential bad or suspected lots and reduce inspection load. The additional capacities generated can be used for R&D purposes to support the smooth industrialization of proposed scientific contributions.

It is highly important to note that in the present organizational structure the role of IT is limited to database management where control and administration of these resources are outsourced. The proprietary nature of database resources further complicates the issue with regards to data models restructuring to add new dimensions. The proposed ROMMII framework suggests the need for the role of data administrator to be created in parallel to IT. The administrator must be responsible for all data sources in the organization along with data dictionaries and any potential change must pass through this role to ensure consistency of concepts and new data dimension along with naming standards and description. The proposed solutions suggest the development of customized data extraction and analysis utilities for different teams in order to improve the productivity of R&D engineers. These customized applications must be updated upon the request of the end user for additional data analysis capabilities. These utilities must be able to use the data analysis algorithms developed and tested by the local R&D teams. The proposed solutions in this thesis are critical for the IDM-fablite business models to maintain their leadership position by capturing maximum market share, which can be achieved by putting DFM back on track. It requires enhancing the productivity of R&D engineers by enabling them to use and exploit the huge production data volumes for effective root cause analyses.

Appendix A: List of Publications

We have published scientific contributions in the well known international conferences and journals. The count for these publications includes Journal *(1)*, Conference *(7)* and others *(2)*. In this section we present the list of the articles published/accepted/submitted along with the abstract for the readers' interest.

A.1 Journal Publications:

1. **Shahzad M.K.**, Siadat A., Tollenaere M. and Hubac S. (2012), ***Towards more effective DFM methods in Semiconductor Industry (SI)***, International Journal of Production Research, (submitted) TPRS-2012-IJPR-0638

Abstract: The DFM methods are used during the technology alignment and adoption processes in the semiconductor industry (SI) for manufacturability and yield assessments. These methods have worked well till 250nm technology for the transformation of systematic variations into rules and/or models based on single-source analysis but afterwards they have become ineffective R&D efforts. The site/die level mismatches due to metrology reference frames and test structure orientations are the main causes for this ineffectiveness. It restricts us from modeling newly emerging spatial variation resulting from miniaturization that often results in increasing the yield losses and lead times. The purpose of this paper is to improve the DFM effectiveness; hence we present a generic coordinates mapping, alignment and qualification model to remove the site/die level mismatches. It provides us an accurate computation of the physical dimensions followed by correlation against the electrical and inspection data at site/die levels. This ability shall effectively transform the newly emerging spatial variations into DFM rules and/or models, based on the multi-source root cause analysis. The presented model is further integrated in BEOL (back-end-of-line) interconnect modeling and an extended BEOL (*e*-BEOL) process is proposed for the yield and lead time improvements based on effective DFM methods.

Keywords: design for manufacturing (DFM); coordinates mapping and alignment; BEOL interconnect modeling and root cause analysis

A.2 International Conference Publications:

1. **Shahzad M.K.**, Hubac S., Siadat A., Tollenaere M., ***An Extended Business Model to Ensure Time-to-Quality in Semiconductor Manufacturing Industry***, International Conference on Enterprise Information Systems, 2011, Portugal

Abstract: Semiconductor manufacturing industry (SMI) has shifted from an IDM (integrated device manufacturer) to a fabless structure where technology is developed in an alliance to share high R&D costs and address time to market and time to volume challenges. In this fabless structure, electronic design automation has emerged as a key stake holder to model increasing design and manufacturing interface complexities and its integration within design flow, but collaboration within alliances have resulted information sharing and technology transfer as the key challenges. We argue that IDM model is superior to a fabless structure due to its inherent ability for faster/superior knowledge capitalization. We benchmarked and analyzed a world reputed IDM with use-case and SWOT (strength, weakness, opportunity, threat) analyses to identify the limiting factors that led this transformation and found data and statistics as the core issues. We have proposed an extended IDM business model where engineering information systems (EIS) are tuned for design for manufacturability (DFM) compliance to achieve time to quality (time to volume, time to market) and yield ramp up rate at low cost but effective R&D efforts.

Keywords: SMI business model; time-to-market (T2M); time-to-volume (T2V); design for manufacturing (DFM); yield ramp-up rate

2. **Shahzad M.K.**, Tollenaere M., Hubac S., Siadat A., ***Extension des Méthodes DFM pour l'industrialisation de produits microélectroniques***, 9e Congrès Internationale de Génie Industriel Montréal, 2011, Canada

Abstract : Semiconductor manufacturing industry (SMI) is characterized by the fastest change in smallest period of time; hence to address time-to-market and time-to-volume challenges, DFM was included in design flow (1980) as a yield enhancement strategy. It has become an industrial standard to assess yield/manufacturability of the design. Test chip is used to validate the geometric stack against resulted specs and models are frozen and distributed to the CAD department for inclusion in design and DFM kits. This paper proposes a DFM methodology to include geometric measurements, which could impact significantly electrical test results making it difficult to adapt the target models. It requires site to site mapping on the wafer, which is not trivial because wafer center is different than the mask center and the test structures PCS/M (process control/monitoring structure) could be present in the horizontal or vertical scribe lines. A case study on the interconnect modeling is performed in a top ranked SMI and an extended methodology to rapidly align local interconnect models on target the source specs is proposed. BPR (business process reengineering) and IDEF0 are used for analysis and newly proposed methodology along with a data model, which is implemented in a tool for R&D engineers.

Keywords: design for manufacturing (DFM), manufacturing for design (MFD), time to market, (T2M), time to volume (T2V) and time to quality (T2Q), engineering data analysis (EDA)

3. **Shahzad M.K.**, Hubac S., Siadat A., Tollenaere M., ***An Interdisciplinary FMEA methodology to find true DFM challenges***, 12th European APCM Conference, Grenoble France, 2012

Abstract: The FMEA (failure mode effect analysis) is a widely used and well known approach for the concept, design and process improvements; however it is limited by the expert's knowledge and its scope. In this article we have used the FMEA approach to identify and remove root causes from the DFM (design for manufacturing) inefficiencies for faster technology improvements and derivatives, resulting in a quick ramp-up-rate. We propose a 4-step interdisciplinary FMEA (*i*-FMEA) methodology to find root causes across the business functions based on the fact that performance of a function depends on inputs, which are outputs from other functions. A case study is conducted in a reputed IDM (integrated device manufacturer) and results are the causes that we have never considered as the limiting factors.

Keywords: deisgn for manufacturing (DFM) challenges, effective root cause analysis, FMEA approach

4. **Shahzad M.K.**, Hubac S., Siadat A., Tollenaere M., ***MAM (mapping and alignment model) for inspection data in semiconductor industry***, 12th European APCM Conference, Grenoble France, 2012

Abstract: Increasing model to hardware gaps has turned technology development process into a high cost R&D activity, hence a new technology is developed in an alliance which is transferred, aligned and adapted for every product. Our engineers are focused on reducing the technology adaption and alignment lead times based on an efficient and effective root cause analysis but they spend significant amount of time in data extraction, mapping and alignment because available inspection data vary in format and coordinate system depending on tool and vendor. Today we have a huge volume of data in multiple dimensions but database issues *(Shahzad et. al, 2011)* limit our capabilities leading to an opportunity loss. In this article we present MAM (mapping and alignment model) for inspection data to ensure site to site mapping between PT and Inline data, die to die

mapping and alignment between EWS and defectivity data and die to site qualification between PT/Inline and EWS/Defectivity data. It empowers our engineers to quickly find the root causes, classify them as systematic or random and transform them into rules and models for the faster ramp-up-rate.

Keywords: coordinates mapping and alignment, effective root cause analysis, knowledge capitalization

5. **Shahzad M.K.**, Hubac S., Siadat A., Tollenaere M., ***ROMMII (referential ontology meta model for information integration) Architecture for Dynamic Restructuring of DWH models***, 12th European APCM Conference, Grenoble France, 2012

Abstract: Semiconductor industry (SI) is facing difficulties in data/information integration while deploying DFM (design for manufacturing) methods for an efficient root cause analysis. Recent IT developments have resulted in the availability of huge volume of data across multiple dimensions, however we are still unable to fully exploit data for a knowledge discovery. The database technologies have addressed platform heterogeneity and efficiency BUT unstructured data model evolution, ontology issue, missing links and the risk of failure of a single source and/or multi source data analysis tool are the key limitations towards an efficient root cause analysis. In this article we present a software framework supported with ROMMII architecture to address these key challenges and ensure dynamic restructuring of the data model against varying needs to ensure the availability and access of every single data element without any risk of the tool failures.

Keywords: data model evolution, software agility, multi-source root cause analysis

6. **Shahzad M.K.**, Hubac S., Siadat A., Tollenaere M., ***SPM (spatial positioning model) model to improve DFM methods***, 12th European APCM Conference, Grenoble France, 2012

Abstract: Scribe line and infield test structures are used to model the design and manufacturing interface complexities resulting from technology scaling, manufacturing and fundamental limitations *(Duane et. al, 1997 and Duane and James, 1996)*. The biggest challenge in accurately capturing these model to hardware gaps is our inability to accurately map, align and position the inspection data (PT, Inline, EWS, Defectivity) collected at die and field levels. The test structures are assumed to be the true representative of the products and are physically located at different positions, hence mapping and alignment do not capture the spatial variations which could answer even the random drifts. In this article we focus the challenge faced by the engineers in implementing DFM with the spatial variation based on test structures positions and the shortest distance between test structures measuring inspection parameters. We propose the SPM model that computes the test structure positions across the wafer and our intelligent algorithm find the target test structure in closest vicinity for spatial correlations. SPM model shall directly impact the productivity of engineers and empower engineers for an effective multi source correlation.

Keywords: spatial variation analysis, yield management, design for manufacturing (DFM)

7. **Shahzad M.K.**, Thomas C., Hubac S., Siadat A., Tollenaere M., ***A yield aware sampling strategy for inspection tools capacity optimization***, International Conference on Artificial Intelligence, USA 2012

Abstract: The product quality in semiconductor manufacturing is ensured with 100% inspection at each process step; hence inspection tools quickly run out of capacities resulting in the production cycle delays. To best utilize the production and inspection capacities, existing sampling (static, dynamic and smart) strategies are based on the risk and delays. These strategies, however do not guarantee a reliable lot sample that represents a likely yield loss and there is a high risk of moving a

bad production lot to next production steps. We present a 3-step yield aware sampling strategy to optimize inspection capacities based on the likely yield loss with the predictive state (PSM) and alarm (PAM) models as: (i) classify potentially suspected lots, (ii) cluster and/or populate suspected lots in the priority queues and (iii) apply last in first out (LIFO) to optimize capacities. This strategy is implemented with two heuristics. We also present a data model with ASCM (Alarm and State Control Management) tool for the multisource data extraction, alignment and pre-processing to support the validation of [PSM, PAM] predictive models.

Keywords: sampling strategy, tool capacity optimization, yield prediction

A.3 Other Publications:

1. **Shahzad M.K.** and Hadj-Hamou K., ***Supply chain configuration modeling under the influence of product family architecture***, INCOM, 2009

Abstract: Supply chain partners strive hard for operational business excellence, enhanced integrated value chain and sustainable competitive advantage under mass-customization/ globalization challenges. In this paper new notions, GBOP (generic bill-of-product: set of product family variants), GBOP/GSCS (generic supply chain structure) interface and GBOP architectural constraints have been introduced that shall empower supply chain the flexibility to rapidly reconfigure under business environmental dynamism and quickly respond to the varying customer needs with economies of scope. Further a mathematical model is proposed to investigate the influence of GBOP on supply chain configuration, relationship between GBOP and GSCS architectures, optimal redefinition of GBOP/GSCS and decisions related to opening or closing of market segments under cost minimization and profit maximization objectives.

Keywords: supply chain management (SCM); integer linear programming; quantitative modeling of facility design

2. **Shahzad M.K.** and Hadj-Hamou K., **Integrated supply chain and product family architecture under highly customized demand**, Journal of Intelligent Manufacturing DOI: 10.1007/s10845-012-0630-0.

Abstract : Mass customization efforts are challenged by an unpredictable growth or shrink in the market segments and shortened product life cycles which result in an opportunity loss and reduced profitability; hence we propose a concept of sustainable mass customization to address these challenges where an economically infeasible product for a market segment is replaced by an alternative superior product variant nearly at the cost of mass production. This concept provides sufficient time to restructure the product family architecture for the inclusion of a new innovative product variant while fulfilling the market segments with the customer delight and an extended profitability. To implement the concept of sustainable mass customization we have proposed the notions of generic-bill-Of-products (GBOP: list of product variants agreed for the market segments), its interface with generic-supply-chain-structure and strategic decisions about opening or closing of a market segment as an optimization MILP (mixed integer linear program) model including logistics and GBOP constraints. Model is tested with the varying market segments demands, sales prices and production costs against 1 to 40 market segments. Simulation results provide us an optimum GBOP, its respective segments and decisions on the opening or closing of the market segments to sustain mass customization efforts.

Keywords : Supply chain configuration; Mass customization; Product family architecture; Generic-bill-of-products

Appendix B: Semiconductor Design, Mask and Manufacturing Processes

We briefly introduce the design, mask and manufacturing process flows followed by key challenges and associated manufacturability and yield limiting phenomenon. It is important because the DFM methods are focused on modeling and transforming these challenges into rules and/or models. The resulting solutions are also discussed to provide a better understanding about the design rules (DR), DFM rules and models which are used during the CAD simulations to pre-assess the manufacturability and yield. In order to provide more clarity and in depth understanding to the reader, we have taken an example of a simple CMOS inverter and simulated the design, mask and manufacturing steps. This example is presented in Appendix-C; however the readers with good understanding on these processes can continue to chapter-2, section 2.2.

B.1 Electronic Circuits and its Elements

An electronic circuit is all about regulating and controlling the flow of current (electrons). These electrons are the atomic particles around the nucleolus of an atom and carries negative charge; hence the flow of electrons refers to the flow of negative electrical charge (current). It is driven by difference in the positive and negative charge (a.k.a. potential measured in voltages [v]) between two locations and follows the positive to negative direction. The materials that allow this flow are referred as the conductors whereas the materials blocking the flow are called insulators. The resistance offered to the flow of electrons is called resistance and it defines the conductivity or resistivity of the materials. In comparison to this simplest definition, an integrated circuit (IC) is a collection of thousands of transistors fabricated and interconnected with one another in a semiconductor material whereas the purpose remains the same i.e. to regulate and control the flow of electrons. The semiconductor materials fall within the category of conductors and insulators and their properties can be altered between conductors and insulators. These materials are transformed into n-type and p-type semiconductors by doping (adding or removing) electrons from the semiconductor material respectively. Such materials when brought in contact with one another forms a difference in potential and starts conducting the electrical current.

These electronic circuits can be classified as either analog or digital circuits. The analog circuits regulate and control the flow of electrons that varies continuously over a range of voltage, current and power values. The digital circuits in comparison to the analog circuits operate at two distinct voltage levels (high and low). The components used in the electronic circuits fall in the category of active and passive structures. The passive components conducts electrons regardless how they are connected e.g. resistors and capacitors whereas the active components control the direction of current flow and also act as an amplifying device e.g. diodes and transistors. The diodes are different than transistor in an aspect that they allow the current flow only in one direction where transistors can direct and regulate the current in both directions. A diode is formed when n-type and p-type semiconductor material are brought in contact, an initial flow of current due to potential difference results in depletion region that results in a barrier voltage (threshold voltage) which is required prior to the current flow.

At present there are three types of transistor technologies being used in the integrated circuit manufacturing [Quirk and Serda, 2000] as *(i)* bipolar junction transistor (BJT), *(ii)* field-effect transistors (FET) and *(iii)* a mix of BJT and FET a.k.a. BiCMOS (Figure B.1). The transistors in the BJT technology have two PN junctions that can be configured as NPN or PNP transistors a.k.a. current amplifying device. Here P refers to the semiconductor material region, which is in deficit of electrons and N defines the abundance of electrons. These regions are created using a doping process that adds or removes the electrons from a region. The FETs are the voltage amplifying devices and are more compact and power efficient than BJT devices. They are further classified as n-MOS and p-MOS where MOS (metal oxide semiconductor) refers to the thin oxide gate located above the depletion regions to control the flow of current. These MOSFET devices have converged into complementary metal oxide semiconductor (CMOS) technology that enables the use of both n-MOS and p-MOS devices in the same circuit and since 1980, it has been the most prominent technology to manufacture integrated circuits. The BiCMOS technology uses the best features of both BJT and CMOS.

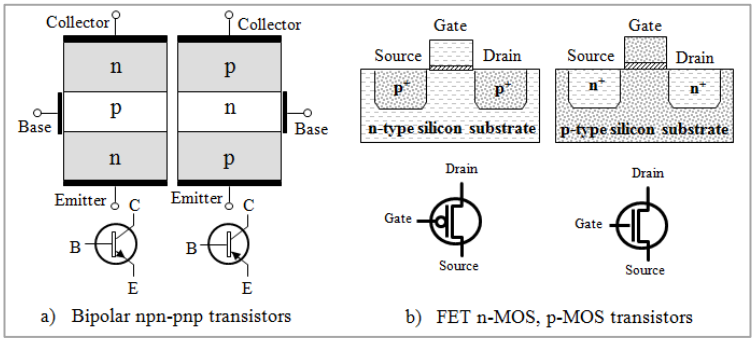


Figure B.1 – Bipolar and FET transistors structure [Quirk and Serda, 2000]

B.2 Reusability and modularity in electronics Design

Like other manufacturing domains, the concept of modularity and reusability do exist in the electronics circuit designing with an objective to reduce the design time. It is for sure that all the electronic circuits are composed of thousands and thousands of active components (n-MOS and p-MOS transistors) and passive components (resistors, capacitors). The devices (n-MOS and p-MOS) are integrated to construct reusable circuits that further integrate to form gates, modules and systems (Figure B.2). When a designers starts a new designs based on the design specifications, he selects these predesigned, simulated and qualified circuits to design a new chip. It is important to note that designers use automated CAD tools for the circuit designing that enable a quick design.

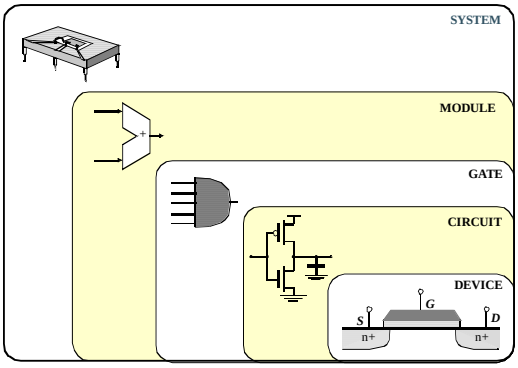


Figure B.2 – Reusability and modularity concept in electronics circuits

B.3 Integrated Circuit Design Process

The design flow is the sequence of steps in designing an IC (Integrated Circuit) chip where almost all steps are automated with the CAD tools [Sagar2001]. These flows are specific for each technology (0.18µm, 0.13µm …) and vary with the design companies. The integrated circuit design process starts with design specifications provided from the customers and follows *(i)* conceptual design, *(ii)* physical design and *(iii)* validation phases. The design flow also vary depending on the type the integrated circuit (analog or digital), however we present a generic digital design flow that takes into account the major steps in the conceptual and physical design phases (Figure B.3).

Based on the design specifications the design team starts with a general block diagram that contains the major functional blocks, this step do not require the usage of CAD tools [Sagar, 2001and Tiri, 2006 and Tomas, 1997]. During high level synthesis (HLS) step, the design team uses a hardware description language (HDL) to model and simulate the block diagram. This model is a high level abstraction of the functional description for each block to ensure the integrity of design idea. The commonly used programming languages at this step are C, vHDL etc. The structural description step goes in the detail of each functional block and circuit is drawn at the logic gates and standard cells level. This step is still the conceptual design phase; hence the details of these components are still conceptual (function based) and independent of the technology. These are modeled and simulated using “verilog” language to ensure intended design outputs. The transistor level synthesis (TLS) is the first step towards physical design where the electronics

components modeled using verilog in the previous step are replaced by the reusable components with physical description from a given technology. This information is used to take into account the effect of the technology against timing, delay and power estimations. The floor and power planning, and place and route (P&R) are the steps where the physical positioning of the electronic components are readjusted and optimized against the added constraints. The extraction of electrical parameters is based on the SPICE Models provided by the chosen technology. The SPICE models are the mathematical equations that use the physical geometries of the electrical components to compute the target indicators (timing, delay and power). The validation phase uses the variation models to compute the unwanted parameters (resistance and capacitance). The design rules check (DRC) and layout versus schematic (LVS) steps ensure the design manufacturability against the given technology. The design is finally tapped out in the GDSII format and is referred as the netlist , which is used during the mask preparation phase prior to manufacturing.

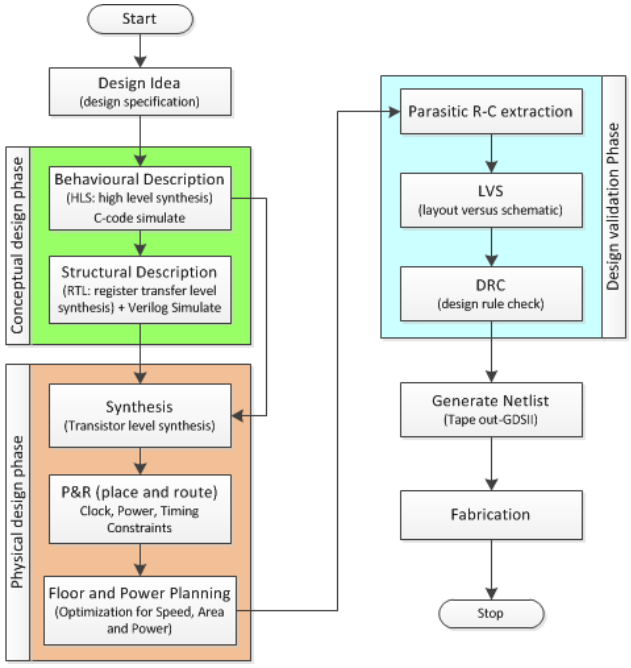


Figure B.3 – Generic design flow for an integrated circuit

It is important to note that these steps are supported by CAD tools from different companies e.g. Cadence, Mentor, Synopsys etc. These tools are plugged with SPICE Models, design, process and DFM kits. The netlist is simulated using SPICE models to characterize the integrated circuits against defined timing, power and delay indicators. The design goes through the DRC simulation and all violations are either exempted or removed by the designer by optimizing the circuit layout. The critical area (CAA) and Hotspot analyses are performed on the circuits' netlist using DFM rules and models to find potential manufacturability and yield issues. These kits are developed and maintained for each technology but as we know that a new technology is developed in a technology alliance, hence these kits are defined and managed at the technology level.

The new technologies are designed and developed in alliances using test products because it is not possible to test all potentially possible designs; hence these kits are developed based on the data analysis from these test products. We also know that the same customer specifications can be met with multiple designs so every new product designed and developed using a given technology might result in drifts, variations and model to hardware gaps. If we properly analyze these variations and transform them into rules and/or models then it shall result in our ability to continuously update the design and DFM kits. In the normal practice, the new technology from the alliance is transformed to the alliance partners' business models where it is declared as frozen and changes are not authorized (section 1.3, Figure 1.10). The reason is that the general purpose and standard cell libraries are developed and qualified for each technology, if this improvement link is authorized then all these existing design libraries are again qualified and validated,

which is expensive and time consuming. It might result in significant delays. The designers are willing to accept any potential changes recommended through local DFM efforts if they are supported with the gains in parametric and functional yields.

More often the process integration teams come up with the changes in the design kits based on the analysis of the parametric data collected from production line. The proposed changes are not supported because they are not supported with the gains in parametric or functional yields. We need to provide our engineers the ability to support their proposed changes in the design and DFM kits along with respective yield improvements or degradation results. It shall extend our knowledge capitalization ability to support the technology alignment and adoption improvement efforts.

B.4 Mask Preparation Prior to Manufacturing

The masks a.k.a. reticles are high precision plates containing microscopic images of electronic circuits or optical devices ("geometries"). They are made from very flat pieces of glass with a layer of chrome on one side. They are used in wafer fabrication, mostly to build IC's (integrated circuits) but also to make flat panel displays, microsystems and optical devices. These masks are used in photolithography process (Figure B.4) to project photo mask image on wafer and if image is projected several times side to side on wafer (stepping), then mask is called a reticle [Hwaiyu, 2005 and Quirk and Serda, 2000].

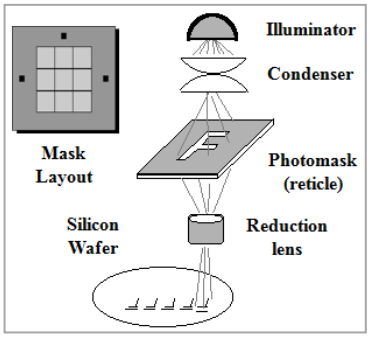


Figure B.4 - The mask layout and its use in photolithography

The reticle manufacturing process is divided in the front-end and back-end flows as presented in Figure B.5. The front end flow is responsible for both scribe lines as well as device data preparation. In this step the GDSII is converted to MEBES (accepted by most litho equipment) format a.k.a. fracturing process. In this process the customer's data is converted into "write tool language" using basic shapes rectangles and trapezoids and it is also known as OPC. The back-end flow follows the physical manufacturing of masks as well as their inspection and repairs as an iterative activity.

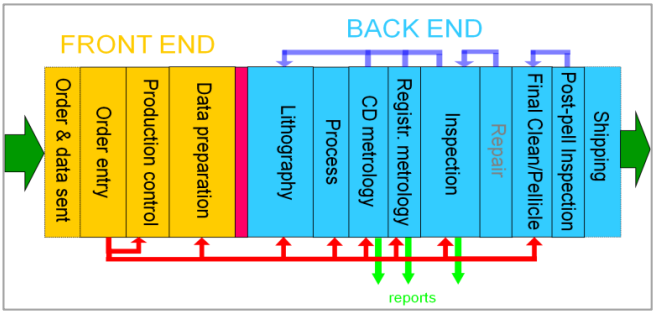


Figure B.5 - mask fabrication flow

When a tapeout is complete, GDSII is sent to two groups in Mask Data Prep. Group 1 prepares the mask data for the device itself while Group 2 prepares the Scribe Frame that surrounds the device (Figure B.6). During fabrication, the scribe frame holds the process manufacturing information and after fabrication the wafer is cut along scribe lines to cut these die from wafer. Mask layers are derived from the GDSII drawing layers via resizing and Boolean operations. Dummy metal fill is added to open areas not already

filled by chip assembly and RET is performed to improve litho printability and device is verified with Litho rule (LRC) and mask rule (MRC) checks [Luo, 2010].

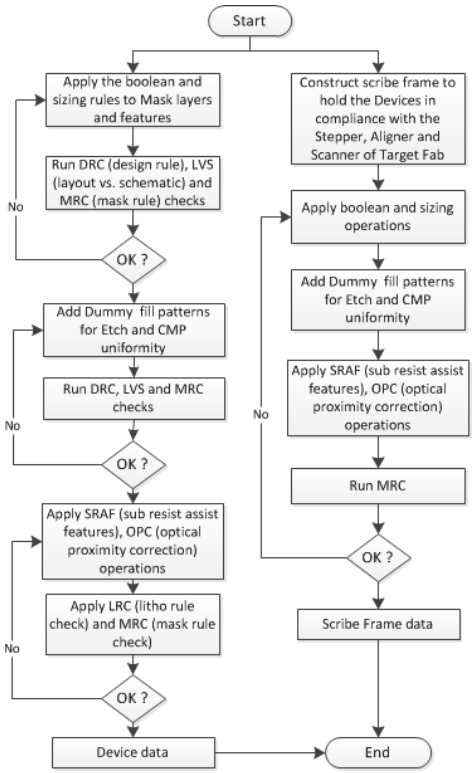


Figure B.6 – Mask data preparation flow

The manufacturing process steps in the mask preparation back-end flow are further detailed as under (Figure B.7). The steps 0 to 5 represent resist flow, used to etch the geometric shapes on the surface of photo mask followed by metrology, clean and inspection steps from 6 to 19.

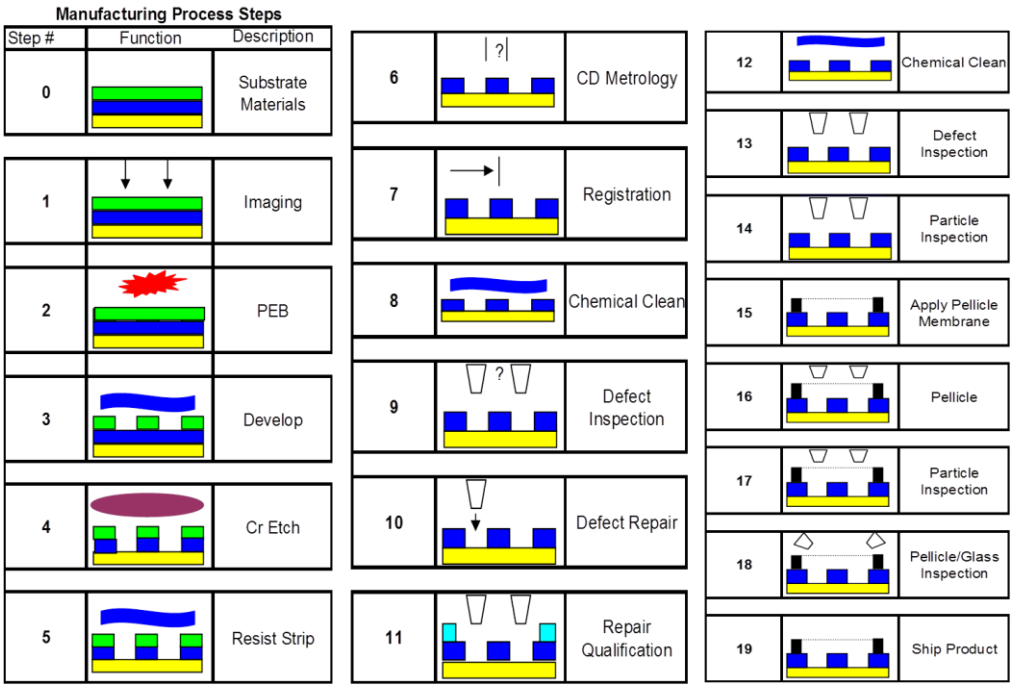


Figure B.7 - back-end mask manufacturing flow

The most common defects for the masks are chrome (Cr) spots, extensions, bridging, pinholes, clear extensions and breaks etc. (Figure B.8)

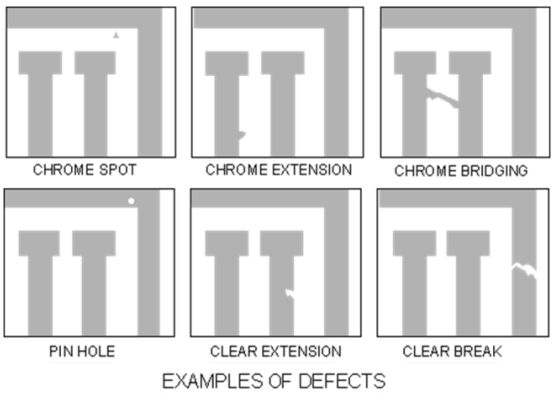
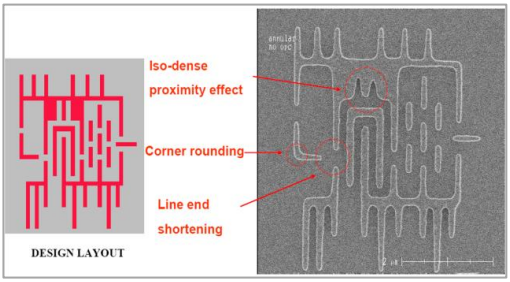
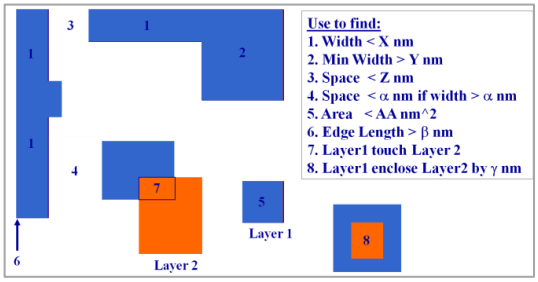


Figure B.8 - mask manufacturing defects

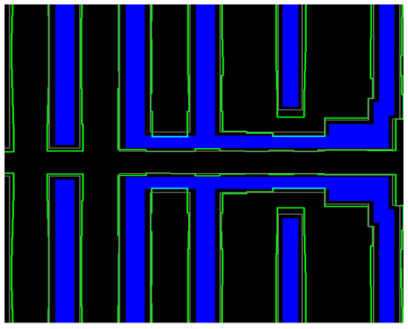
The mask data preparation is the last step where we can put our efforts to improve the design for manufacturability and yield issues. This step if not properly addressed shall result in worst yield issues. The no of masks to be used depends and are defined by the technology being used; hence the rules are applied on both integrated circuit designs (devices and interconnects) and scribe frames. As we know that the biggest challenge faced today is the sub wavelength manufacturing so a mask provides us a safeguard against those yields limiting factors by ensuring the devices against litho and mask rule checks. The SRAF features are special features, which are added to the hot spots to avoid the effects during lithography, etching or CMP operations. The reasons to apply OPC due to lithography issues are presented in the Figure B.9a. It is evident that due to sub wavelength lithography we are not able to exactly fabricate the shapes as described in design layout; hence OPC rules must be defined to ensure a clean design layout (Figure B.9b). The OPC based layout is presented in Figure B.9c where bleu shapes are after etch targets, grey shapes are after litho target and green correspond to actual mask layout [Chiang, 2010 and Orshansky, 2008 and Wong, 2008].



a) Lithography effects and OPC



b) OPC Rules for effective layouts



c) Lithography effects and OPC

Figure B.9 - mask manufacturing defects

It is also evident that the emergence of new technologies has added to the complexities and resulted in increasing number of mask count. The increasing use of OPC methods highlights the complexity on rise with every new technology (Figure B.10).

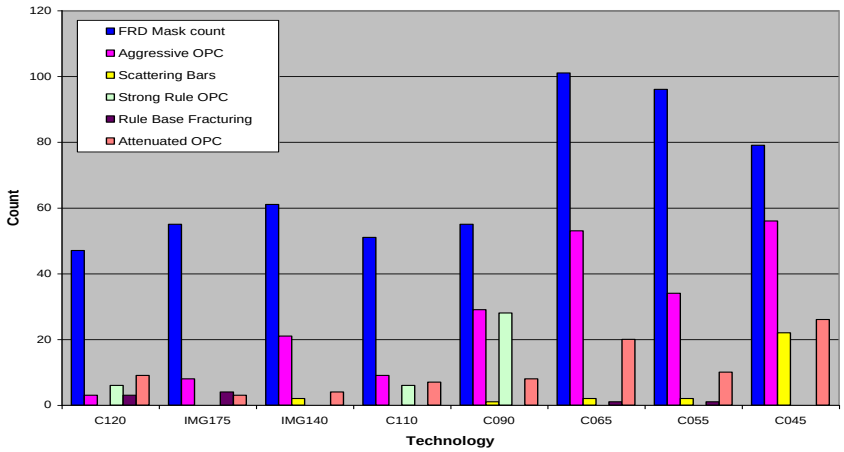


Figure B.10 - Mask manufacturing defects

B.5 Semiconductor manufacturing process

The semiconductor material used in the manufacturing of ICs is the key; hence let us start by summarizing the semiconductor materials, which are being used and those that can be potentially used during the manufacturing.

B.5.1 What is a Semiconductor?

The semiconductor is a material that falls between the conductors and insulators and it can change its properties if electrons are added or removed using doping operation. It can be used for the manufacturing of active as well as passive components in the electronic circuits. The highly refined silicon is used for wafer fabrication and is termed as semiconductor or electronic grade silicon. The ultra-high purity of semiconductor-grade silicon is obtained from a multi-step process referred to as the Siemens process. The silicon makes up about 25% of the earth's crust by weight and it is the second most abundant element. The crystalline silicon was first prepared by Deville in 1854 and it has metallic luster and a grayish color. It is positioned in Group IV-A of the periodic table as shown in Figure B.11. The other elements in group IV-A are carbon (C), germanium (Ge), tin (Sn), and lead (Pb). The carbon is strictly nonmetallic, silicon is essentially metallic, germanium is a metalloid, and tin and lead are true metals [Quirk and Serda, 2000].

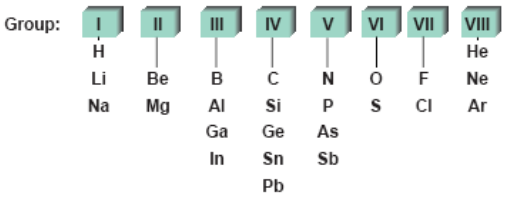


Figure B.11 – The silicon in the periodic table

The group IV elements are elemental semiconductors whereas group V and III elements are n-type and p-type dopants used to change the properties of the silicon. The silicon is preferred over other metals due to its superiority in terms of melting point (1415ºC) and accurate oxidation properties. This oxidation layer is used as a dielectric for the transistor gate to control and regulate the flow of current within transistors.

B.5.2 The MOSFET Transistors Structure and Operations?

We have seen n and p type MOSFET transistor (Figure B.1), let us discuss their structure and how does they Operate? The structural and operational description of the npn and pnp transistors is presented in the figures B.12 and B.13 respectively. It consists of a substrate, source, drain and a gate. The current flows from the source towards drain based on the capacitor environment formed by the gate, substrate and their internal dielectric (gate oxide). When a charge is applied on the gate it either pulls or pushes the electrons in the substrate across the channel length. These

electrons if accumulated towards the gate cannot jump into gate because of the oxide shield, however depletion region is extended that ultimately allows the flow of electrons.

The npn MOSFET is presented in Figure B.12 with its two operation modes. The npn transistor shall work as an electronic switch if we are able to establish a conductive path between the source and drain under the gate. A battery (3v) is connected to a small lamp with external wires. If there is no voltage applied on the gate then there is no flow of the current between source and drain resulting in the open gate situation and the bulb remains off, however a small voltage at the gate is all what is need to turn it on. The gate to source voltage (V_{gs}) is applied on the gate that creates an electrical field, which pulls electrons from the source, drain and substrate. These electrons get collected underneath the gate oxide a.k.a. the depletion region. This depletion region gets extended as V_{gs} is increased and when it reaches the threshold voltage (V_{th}) the depletion regions extends to the source and drain resulting in the flow of drain to source current (I_{ds}). The flow of current from drain to source results in the completion if the circuit and the lamp is turned on. The depletion region in npn transistor is referred as n-channel resulting in drain to source current flow.

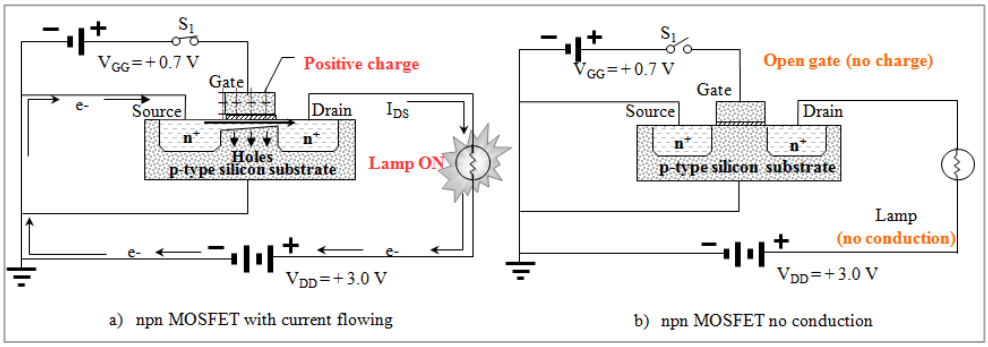


Figure B.12 – npn MOSFET structure and current flow [Quirk and Serda, 2000]

The structural and operational mode for the pnp MOSFET is presented in Figure B.13. The open gate situation prevails until no voltage is applied on the gate and there is no current flow between drain and source. When we apply a negative voltage on the gate then electrons are pushed back from the gate oxide resulting in the accumulation of holes underneath the oxide. As soon as the applied voltage reaches the threshold level, the depletion region gets extended towards drain and source resulting in flow of current from source to drain. This depletion region is also known as the p-channel and current flow is attributed to the flow of positive charges.

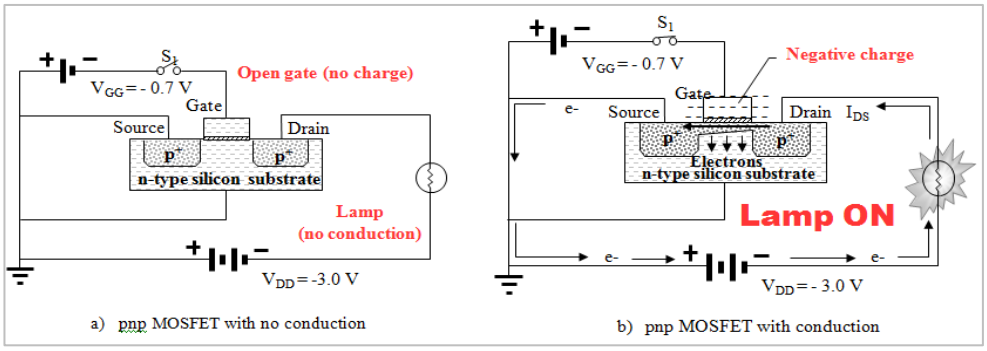


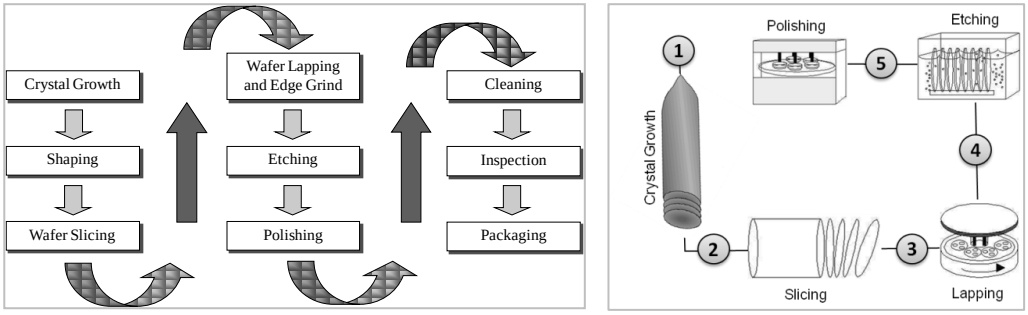
Figure B.13 – pnp MOSFET structure and current flow [Quirk and Serda, 2000]

B.5.3 An Overview of Semiconductor Manufacturing

The high level semiconductor manufacturing process can be divided in 5 steps *(i)* wafer or substrate manufacturing and processing, *(ii)* fabrication of integrated circuits on the wafer, *(iii)* probing or testing, *(iv)* packaging and *(v)* final test.

a) **Wafer or Substrate manufacturing (Figure B.14):** The wafers or substrates manufacturing starts from the crystal growth using Czochralski (CZ) process. It is grown in the cylindrical shape and then this cylinder is sliced into wafers. To smoothen the sliced surface and the edges we perform

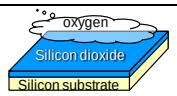
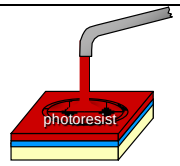
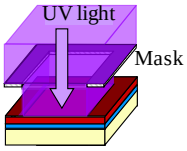
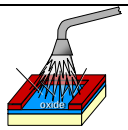
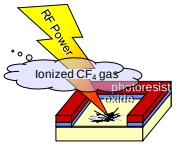
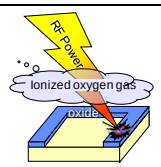
wafer lapping, etching and polishing operations. It follows cleaning of the surface, inspection and packaging of the wafers for shipment to the SI [Hwaiyu, 2005and Quirk and Serda, 2000 and Rakesh, 2008].



a) Generic wafer manufacturing process b) 5-step wafer preparation

Figure B.14 – Wafer manufacturing wafer processing

b) **Fabrication of the integrated circuit:** The major tasks used in the manufacturing of a transistor are presented in the table B.1 [Michael, 2004].

Thik Films	 a) Oxidation (Field oxide)	This is the first step in the IC manufacturing where an oxide layer is grown on the silicon wafer surface in horizontal or vertical furnaces at about 1100ºC using dry/wet operations. It is grown using chemical reaction between Si and Oxygen as SiO2.
Photo Lithography	 b) Photoresist Coating	The photo resist coating follows surface treatment to drive off moisture, coating of small amount of photo resist (mm) and pre-bake that stabilizes the film prior to pattern transfer step in lithography process. The +ive and –ive photo resists become soluble and insoluble upon exposure to light. We use spin, spray, dip or roller coating processes for this key step in patterning on the silicon wafer.
Photo Lithography	 c) Mask-Wafer Alignment and Exposure	The exposure step transfers the pattern from the mask to photo resists coating on wafer surface. The mask patterns are 5x or 4x larger than features being manufactured. These features are optically shrunk before reaching the surface of wafer. The mask and misalignment errors are the key challenges faced by the industry that often result in the physical defects (functional failures).
	 d) Exposed Photoresist	The exposure results the photo resists material soluble or insoluble resulting in the features being transferred to the silicon wafer surface.
	 e) Photoresist Develop	This step dissolves the area exposed to light more quickly in +ive photo resist and dissolves the area not exposed to the light for the –ive photo resist material. The objective is to prepare the pattern to follow etch and Ion implantation steps.
Etch	 f) Oxide Etch	Initial etching operation used wet etch (isotropic) process which is based on the liquid chemicals. It etches material in all directions at the same rate, hence while etching down a film it etches underneath the edge of photo resist as well. This process worked very well when the line widths were large but shrinking of the features has resulted in serious defects. Now the industry has moved towards dry etching technique that uses ionized gases to perform etching which is faster and directional. It is also referred as an anisotropic etching. The ionized gases used during dry etch operations are tetrafluoromethane (CF4), pertluoropropane (C3F3) or fluoroform (CHF3). In this process we use high frequency energy to split up the gas molecules in a low pressure chamber that creates reactive products.
	 g) Photoresist Strip	

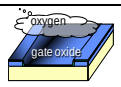
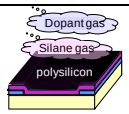
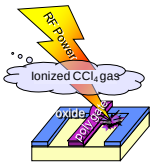
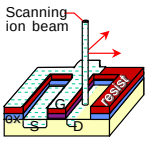
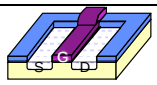
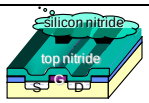
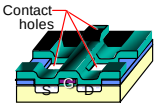
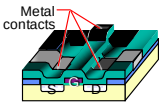
Thin Films	 h) Oxidation (Gate oxide)	These gate oxidation and poly silicon depositions (thin films) are very critical operations a.k.a. thin films; hence it results in the formation of gate oxide that acts as a dielectric between gate and wafer substrate. It plays an important role in the formation of channel between source and drain regions. The poly silicon results in gate formation and is used to switch the transistor on and off. The most important challenge faced today is the uniformity while reducing the thickness of these thin films.
	 i) Poly silicon Deposition	
Photo and Etch	 j) Polysilicon Mask and Etch	This step is used to transfer the mask pattern for the gate with photolithography operation followed by etching. The poly silicon gate is a key element in whole operation of the MOSFET transistors. We have resolved the issues of isotropic etching, but under etch and over etch issues are still a big challenge for the R&D engineers.
Diffusion	 k) Ion Implantation	The ion implantation step is used to change the substrate so that they can act as n-type or p-type source and drain regions. It is alternative to diffusion. In diffusion, dopant atoms are moved from surface into silicon substrate by thermal means and in ion implantation they are forcefully moved into the surface with ionized beams.
	 l) Active Regions	The active regions are those regions on the surface of silicon substrate where poly gates are formed along with sources and drains. The operations used for the active regions are called front-end-of-line (FEOL) operations.
Thin Film	 m) Nitride Deposition	By this time the transistor is ready and we move towards metallization process (inter connections) a.k.a. back-end-of-line (BEOL) operations. This film is deposited to save the interaction between two layers of operations (FEOL and BEOL).
Photo and Etch	 n) Contact Etch	This step is the start of the BEOL operations and first step involves creating contacts to the source, drain and the gate so that potential can be applied at the transistor level to turn it act like a switch. Here again we follow the photo lithography and etch operations. It is known as Metal1.
	 o) Metal Deposition and Etch	These steps involve repeating the photo and etch operations to construct vias and metal lines starting from the contact (BEOL). The number of metal lines may vary depending on technology, however reduction in the geometric specifications are serious challenges against potential defects.

Table B.1 – Steps in transistor manufacturing process

Let us analyze the production flow in the semiconductor manufacturing plant where the production line is divided in different bays and these bays might contain production equipments from different manufacturers. It takes about 1100+ operations and 8 weeks of processing. The flow and movement of the wafers within these production bays are presented in the Figure B.15.

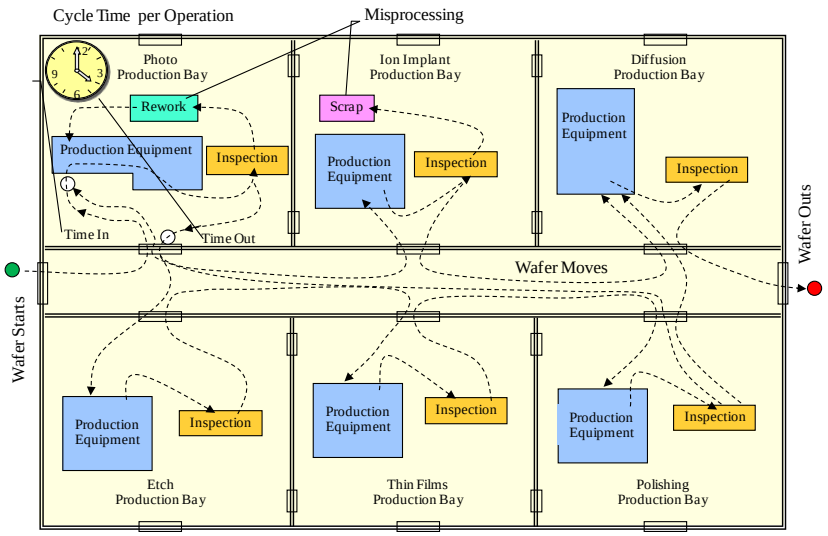


Figure B.15 – Steps in transistor manufacturing process [Michael, 2004]

It acts as a second line of defense against manufacturability and yield issues and we use manufacturing for design (MFD) efforts to put yield drifts and excursions back on track. It can be noted that every production bay is supported with the inspection/metrology equipments, so the products leaving the production are inspected for defects or drifts. There are two major type of tests performed at this level *(i)* defectivity (inspections) and *(ii)* metrology (geometric measurements). These measurements are performed on the process control monitors (PCM) and the data is available at the site level, however the defectivity is performed on sampled locations throughout the wafer and is available at die levels. These measurements are tagged with x,y coordinates for the die or site with reference to the wafer notch position. The analysis against drifts and excursions at wafer level is possible, but due to emerging spatial variation components engineers need die and site level analysis. Such type of multi-source root cause analysis is not possible at the moment due to varying coordinate systems. It has significantly contributed to DFM ineffectiveness and is based on the fact that root cause analysis provides reasons for drifts and systematic occurrences are further transformed into rules and/or models.

a) **Probing and Testing:** In this phase, wafer is tested and electrical measurements are made at wafer level for the specifications compliance. There are 5 types of electrical tests performed *(i)* IC design verification, *(ii)* in-line parametric test, *(iii)* electrical wafer sort, *(iv)* reliability tests and *(v)* final tests [Hwaiyu2005, Quirk and Serda, 2000].

The parametric tests are performed early in the manufacturing phase at metal1 right after the composition of the transistors. They are performed on the test structures (a.k.a. PCM: process control monitors) located within circuit and scribe lines. They measure different parameters (e.g. leakage current, critical dimensions, threshold voltage, resistance, etc.) to identify process problems, establish wafer pass/fail criteria, collect data, assess special tests and obtain wafer level reliability data. These tests are performed using probe card that interfaces with test structures (Figure B.16) and the test equipments a.k.a. automated test equipments (ATE). The coordinates for each measurement are stored and depends on the orientation of the wafer prior to its measurements. These measurements on individual basis are used to perform wafer level variations analysis but computation of new parameters from the measurements and multivariate analysis is not possible. The R&D engineers first align the data having different measurement coordinates so that analysis can be performed at the site and die levels instead of wafer level.

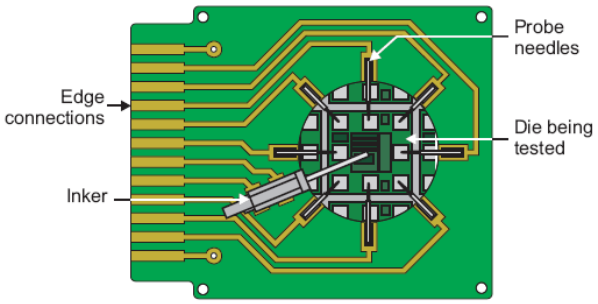


Figure B.16 – Probe interface card and electrical tests

The objectives of EWS test are the chip functionality and sorting good and bad dies, so evidently this test is performed at die level. The test results are categorized by assigning different bin numbers to the wafers. It includes continuity test, output checks and functional tests. The only issue with this type of test is the time required, hence fault models are used to sort the dies and optimize the time. Few reasons for the bad products has been identified as the larger wafer diameters, increased die size, increase in number of process steps, shrinking features, process maturity and crystal defects. It is really interesting to note that R&D engineers are specially struggling in performing reverse engineering to trace back the root cause for bad product yield and variation in the drifts.

b) **Assembly and Packaging:** In this phase, the good dies are separated from the wafer, assembled and are packaged. The assembly involves backgrind, die separation, die attach and wire bonding. The backgrind reduces the wafer thickness and die separation cuts each die from the wafer and it is followed by physical attachment on the substrate. Wirebonding finally attaches wires from die bonding pads to the lead frame of the substrate for the external connections [Quirk and Serda, 2000 and Scotten, 2012].

The traditional packaging consists of plastic or ceramic packaging. The ceramic packaging is used for the state of the art ICs that require a maximum reliability or high power. The two main types of ceramic packaging are either a refractory (high temperature) ceramic or ceramic DIP (CERDIP) technology. The main objective is to protect the dies from environment, interconnections, physical support and heat dissipation. The new packaging designs are being introduced to provide more reliable, faster and higher-density circuits at lower cost. The most common technology is the ball grid array (BGA). It uses ceramic/plastic substrate with an array of solder balls to connect substrate to the circuit (Figure B.17).

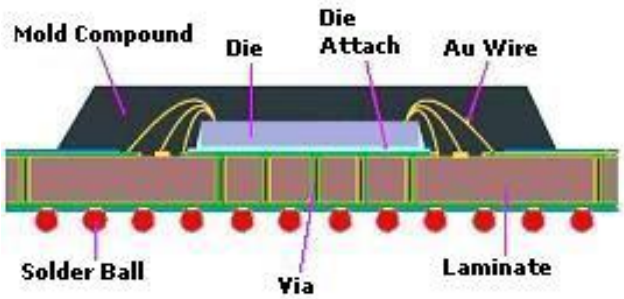


Figure B.17 – Probe interface card and electrical tests (Scotten, 2012)

c) **Final Test:** It is likely that during packaging the die may have been damaged and its failure due to packaging issues is more than 1%. The final test is 100% on all assembled and packaged chips to insure that any IC improperly packaged is not shipped.

B.6 Key Manufacturing Processes and Challenges

Before we move to the role of DFM in the semiconductor industry, let us rapidly summarize the key manufacturing processes *(i)* oxidation, *(ii)* photolithography, *(iii)* etch, *(iv)* ion implantation, *(v)* chemical vapor deposition (CVD) for thins layer and *(vi)* polishing a.k.a. CMP.

i) **Oxidation:** It is a batch process where multiple wafer (200+) are processed together and S_iO_2 is grown on the silicon wafer surface between 900ºC and 1200ºC in the oxidation furnace. The wafers are heated in the furnaces containing oxidant (process gas), usually O_2, steam or N_2O that result in the formation of thin deposition of oxide. The schematic of the oxidation furnace is presented in Figure B.18 [Hwaiyu, 2005 and Michael, 2000].

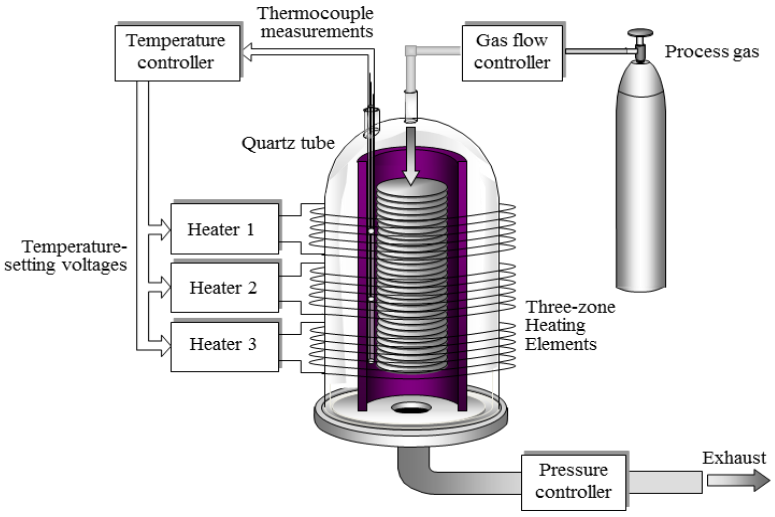


Figure B.18 – Structural description of the oxidation furnace

ii) **Photo Lithography:** It is one of the mostly used steps in the manufacturing process. The structural schematic description of litho stepper equipment is presented in Figure B.19. The wafers in this process undergo the resist coat followed by the soft or hard bake steps depending on resist type being used. The transfer sections are responsible for the transfer of wafers to/from alignment exposure sections and resist coat and development sections. These wafers are loaded back to the wafer cassette (Lot) which is a box that carries 25 wafers at a time [Hwaiyu, 2005].

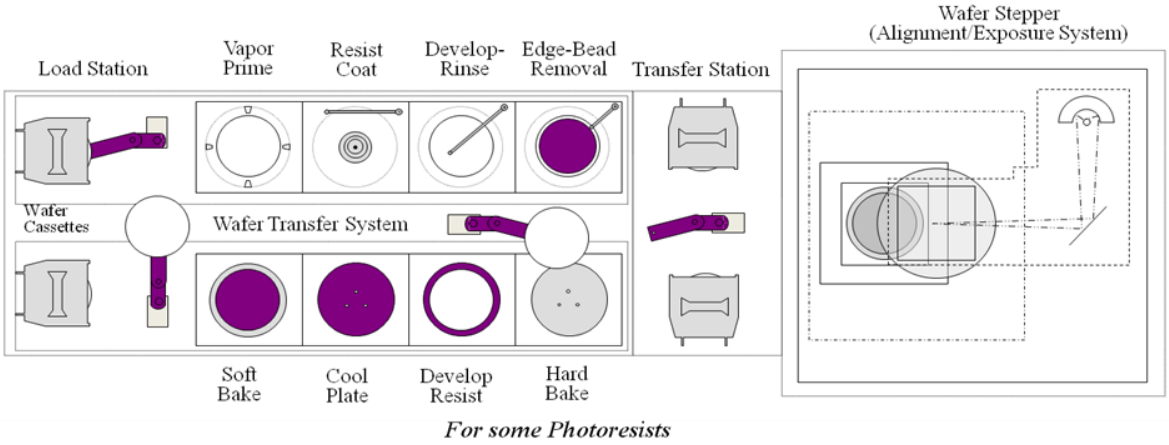


Figure B.19 – Structural description of the lithography equipment

This operation is very critical to insure the manufacturability and yield issues due to mask misalignment errors. The most common errors associated with this process are grouped into focus exposure misalignment (Table B.2) and illumination (Table B.3) errors.

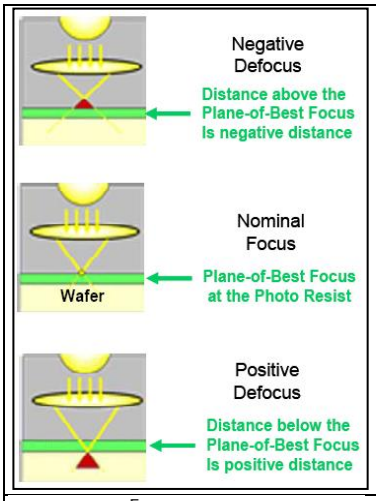
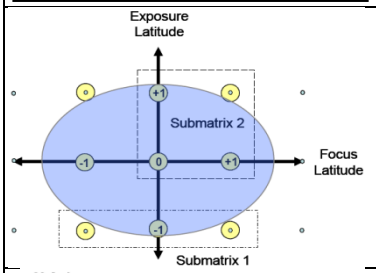
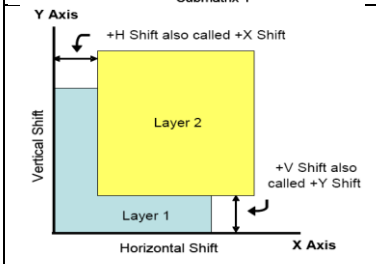
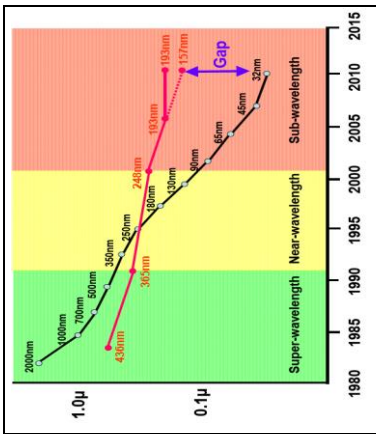
	**Defocus:** The defocus conditions denote simulation conditions where the light from the stepper's projection system are not at the nominal focal point. There are three types of defocus conditions *(i)* negative focus where the focal point is above the plane of best focus for the resist development, *(ii)* nominal focus where the focal point is directly on the plane of best focus and have best contrast between illuminated and non-illuminated resists and *(iii)* positive defocus where the focal point is below the plane of best focus. The circuit features with significant height, such as poly gates of transistors or interconnect structures cannot have all the features at Nominal Focus. The positive or negative defocus errors often result in the resist material left undeveloped or resulting in damaging the surface below the resist material.
	**Focus Exposure Matrix (FEM):** The nominal condition FoEo for FEM is at location (0,0) and different FE conditions are denoted by their row and column position. The number of FE conditions in the FE matrix is variable with 5, 9 and 15 conditions specified in the FE matrix as the most common qualification set. The qualified process windows for each technology is defined and used as a reference to trace back the deviations.
	**Mask Misalignment Error (Overlay Error):** During fabrication there is an error distribution for Mask Misalignment a.k.a. Overlay Error between the two consecutive masks that must be overlaid to produce a circuit feature. The overlay errors are called H (X) Shift and V (Y) Shift which denote the relative mask misalignments to each other. This misalignment is critical that often result in coverage issues in contact and vias formation.

Table B.2 – important focus exposure misalignment factors [DFM Dictionary, 2004]

	**Sub wavelength:** Shortly after the turn of the century, a stepper using 248nm illumination wavelength (λ) was able to produce 130nm features on the silicon wafer and193nm λ illumination produced 90nm features. Aggressive RET enabled feature sizes less than half of the wavelength of the illumination, but as silicon features continued to shrink, the sub-wavelength challenge resulted in added complexity. The 193nm node has become the end of the road for optical wavelength reduction. The efforts to build 157nm λ stepper system were abandoned due to the technical problems and the steppers shifted towards immersion technology and water instead of air, at the interface between the lens and wafer.

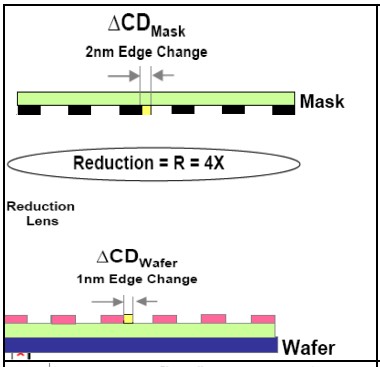
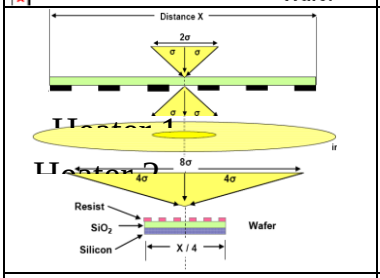
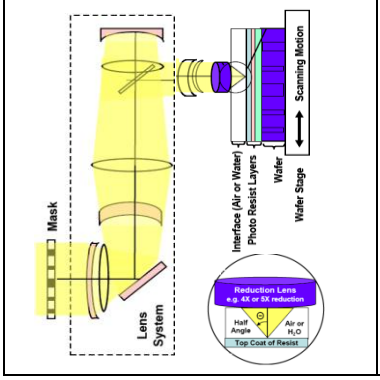
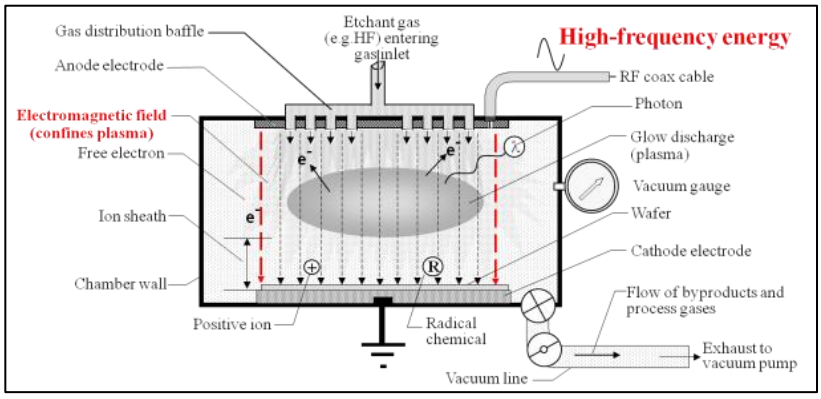
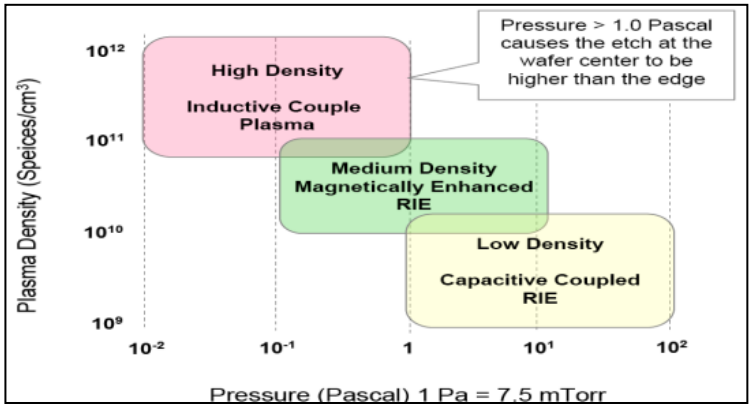
	**Mask Error Enhancement Factor (MEEF):** It is because of the fact that wavelength of light used to expose resist is larger than the feature size; hence edge placement errors (EPE) are amplified. So, during OPC modeling, an edge movement on a mask feature can result in a larger edge movement on the silicon image. The change in the edge placement on the mask is reduced by the reduction ratio (R) of the projection system and amplified by MEEF so that a 2 nm edge shift on the mask, even reduced by 4X can result in a 1 nm shift on the wafer.
	**Reduction Factor:** Projection stepper systems typically reduce the image on the mask (reticle) by 4X or 5X before exposing the resist on the wafer. The image reduction by the lens in the projection system allows the patterns on a mask to be 4X to 5X larger than the images realized on the wafer. The Numerical Aperture of the Project System is NAO = NALENS (1 + 1/R) where R > 1.
	**Numerical Aperture (NA)**: It is a measure of total light projection power of the stepper lens system. It is equal to the maximum half-angle of light that can make it through the projection lens system multiplied by the real index of refraction (n) of the medium at the lens and resist interface. The half-angle Θ is half of the angle from the focal point to on the wafer to the edge of the stepper lens. It is computed as N.A = n.sinΘ. A moderate half-angle of 450 produces a NA = 0.70 (= 1.0 · sine 45) with an air interface. An aggressive Θ of 72O produces a NA = 0.95 (= 1.0 · sine 72) for an air interface and NA = 1.36 (= 1.4 · sine 72) for a water interface.

Table B.3 – Important illumination factors [DFM Dictionary, 2004]

iii) **Etch:** It is the process step which is repeated with photolithography and is critical in controlling geometries of the features. The plasma etching is the most commonly used technique as it provides a combination of chemical and physical etching. The physical component provides good anisotropy with little undercutting and chemical component provides good selectivity. The low plasma density operates at a relatively high chamber pressure between 1 and 100 Pascal (CCP). In contrast, the high density plasma provided by ICP has an upper pressure limit of 1 Pascal to keep the etch rates across the wafer consistent (Figure B.20).



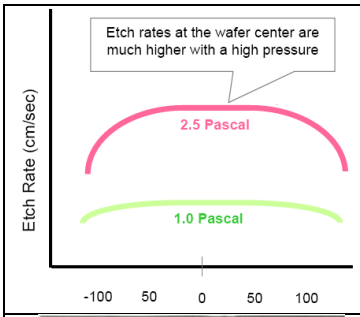
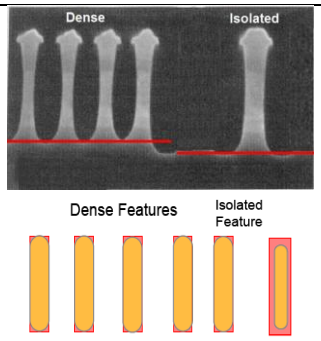
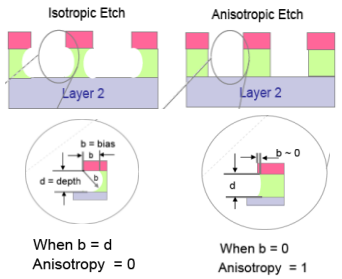
a) Etch equipment and process



b) Plasma as Etch gas

Figure B.20 – Structural description of the Etch process

This step in combination with lithography is responsible for most of the errors and defects during manufacturing, however the most common defect related phenomenon [Hwaiyu, 2005 and Orshansky, 2008 and Wong, 2008) are presented as under (Table B.4):

	**Etch Macro-Loading:** Macroscopic loading occurs when the etch rate varies across large distances (1 – 1000 mm) such as the wafer position in the reactor or with the die position on the wafer. When more wafers are in the reactor or more total area is exposed by the mask pattern the etchant species may become depleted lowering the overall etch rate. For example, with the high plasma density, when the gas pressure is too high the etch rate can vary 50% from the center of the wafer to the edge. Macro-loading is difficult to control.
	**Etch Micro-loading:** It is defined as the phenomenon in which the etch rate varies across small distances (0.1um – 10um) due to variations in the resist pattern such as feature widths, aspect ratios and area densities. The narrow wires have higher aspect ratios than wider wires and slower etch rate a.k.a. Aspect Ratio Dependent Etching (ARDE). The probability of the etchant species making it all the way down to surface is lower. A dense patterned area may locally deplete etchant species which lowers the etch rate but isolated patterns which have plenty of etchant species located around them may become over etched.
	**Etch Anisotropy:** Etch Anisotropy is determined by the ratio of the lateral and vertical etch rates of a single layer whereas isotropic etch is non-directional so $etch_{lat}$ = $etch_{vert}$ which means $etch_{lat}$ / $etch_{vet}$ = 1 and the Anisotropy of the etch is equal to 0. A physical etch process, using ion bombardment, is highly directional (anisotropic) but it has poor selectivity since the ions etch all materials at similar rates. In contrast, a chemical etch processes is highly selective but it is also very non-directional (isotropic).

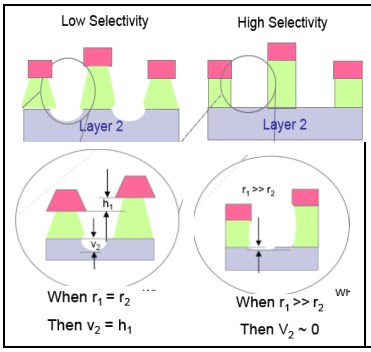
	**Etch Selectivity:** Etch selectivity is determined by the ratio of the etch rates of two different materials. For example, the selectivity of a chemical etch for silicon dioxide compared to silicon is 20x times faster. In the example on left, we need to etch through the silicon dioxide down to silicon substrate. However, the center area of silicon dioxide is 100nm (= h1) higher than the left or right edges. Removing all of the oxide in the center means the other areas will be over etched. If the etch rate of silicon is 20X less than silicon dioxide, the silicon valley V2 (shown on the right) will only be over etched by 5 nm (= 100/20).

Table B.4 – Important etch phenomenon [DFM Dictionary, 2004]

iv) **Ion Implantation:** The Ions a.k.a. charged molecules or atoms are created by a strong electrical field that strips away the electron resulting in an Ion [Hwaiyu, 2005 and Wong, 2008]. These Ions are filtered (lighter Ions are separated) and then accelerated towards a target wafer (Figure B.21). The depth of implantation depends on the acceleration energy (voltage).

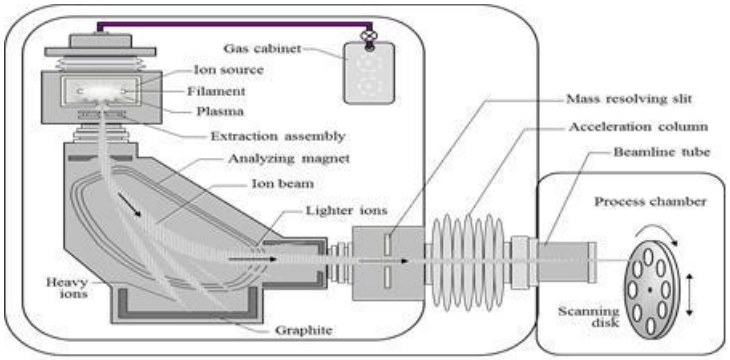


Figure B.21 – Structural description of the Etch process

This process is used in parallel with the diffusion step, but it offers a precise control of the dose and is quite fast (a wafer can take 6 seconds) and multi energy implants. It has certain disadvantages as well e.g. deep and shallow profiles are difficult, often uses toxic gas sources (arsine, phosphine) and it is quite expensive.

v) **Chemical Vapor Deposition (CVD):** The CVD is a critical process which is used to deposit thin film, high purity and high performance layers. The precursor gases (often diluted in carrier gases) are delivered into the reaction chamber at ambient temperatures. As they pass over or come into contact with a heated substrate, they react or decompose into solid phase which are deposited on substrate. The substrates temperature is critical and can influence what reactions will take place. The structural process is presented in Figure B.22 [Michael, 2004].

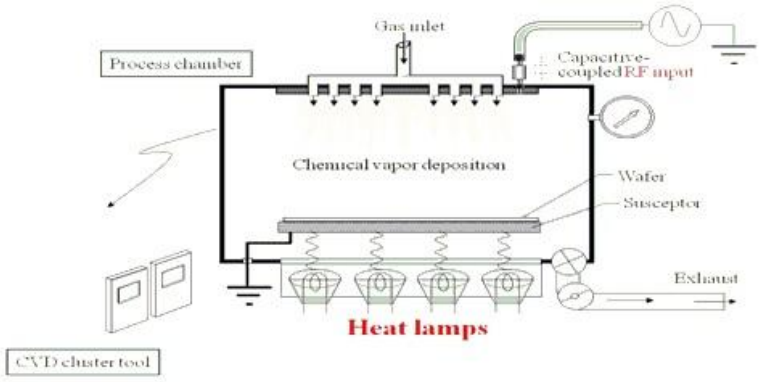


Figure B.22 – Structural description of the CVD process.

vi) **Chemical Mechanical Polishing (CMP):** The CMP process is very important in the metallization and BEOL interconnects process for the planarization of wafers using slurry (alumina/silica plus metal oxidizer chemicals). We need to smooth the surface to allow more layers to be deposited on top. In this process the water interacts with the glass forming a hydroxyl bond (Si-OH) which decreases oxide hardness. The mechanical motion also decreases the oxide strength and abrasive particulates grind off the surface glass. In this process the metal is first oxidized using NH_4OH (for Cu) forming Cu(OH)2 and then slurry abrades the new oxides from surface. The CMP mechanism is presented in Figure B.23 as under [Orshansky, 2008 and Wong, 2008].

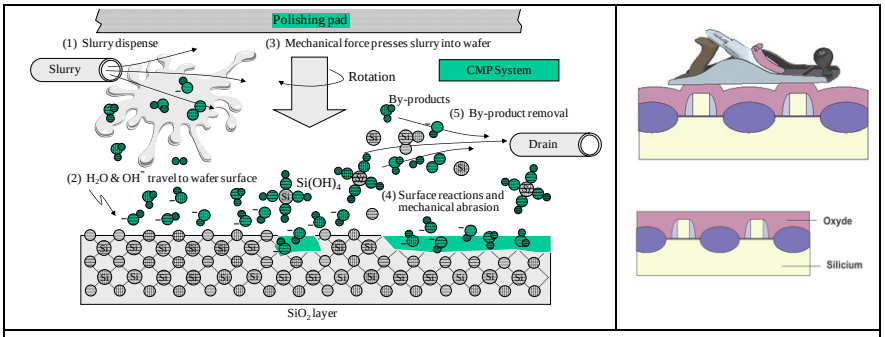


Figure B.23 – CMP mechanism

The most commonly used metal for metallization is copper a.k.a. soft metal. The wide wires can experience copper dishing in the center of metal lines which is removed during polishing. A similar problem called erosion occurs on the dielectric oxide that separates wires, hence the dishing is related with the wire widths and erosion is related to difference in wire density (Figure B.24).

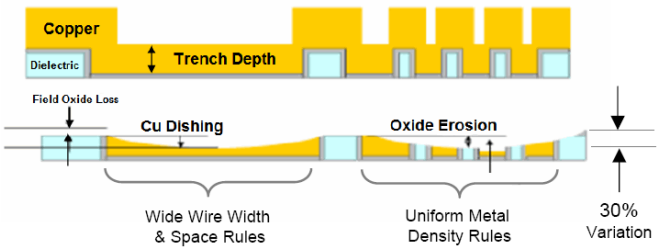


Figure B.24 – Dishing and erosion mechanisms.

The metal thickness and oxide thickness can vary by -30% to +30% without additional routing rules. DFM aware routers provide a greater uniformity of metal widths and keep the routing density within a {min max} density range. If we follow the DFY, the routing guidelines can reduce reduces the thickness variation by 3X to produce a ~ 10% lower variation in the copper and/or oxide thickness.

B.7 Most Common DFM and Yield Limiting Mechanisms

We have seen that the key behaviors within the IC manufacturing processes (section B.6) might result in significant geometric variations with strong impact on the parametric and functional yields. We also know that, SPICE models (Appendix B.1) need geometric dimensions in addition to process technology parameters to extract the electrical and parasitic components. So, it is the right time to introduce the design rule manual (DRM) and DFM Rules and Models. The DRM is the bible for the designers and contains design rules for given technology. These design rules (geometric specifications and allowed variations) are programmed into design kits which are used during CAD simulations to verify their compliance (DRC). It

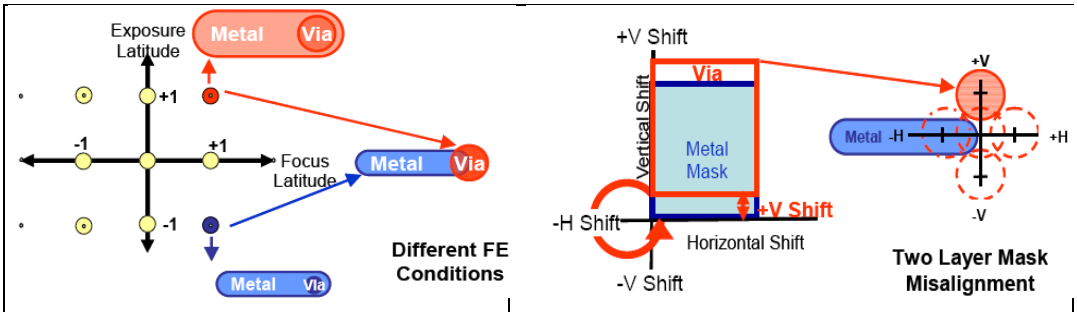
ensures the product functionality against the given specification; hence any deviation from these defined design rules shall result in product functional yield loss. In comparison to the DRM, the DFM rules are basically an extension to the design rules and are focused on insuring the printability of the features on silicon wafer. The DFM models are used to perform CAA and hotspot analyses to identify the printability issues in the design layout and the compensations are added to result the design layout which is more manufacturable with less yield issues. These compensations are added to the mask which acts as an interface between design and manufacturing processes.

An electronic product (chip) is characterized by electrical parameters; hence an unexpected drift/variation (beyond model corners) is to be investigated for its root causes. The major cause for these drifts can be traced to uncontrolled variations in the geometric shapes but it is not always the case and such variations might present the newly emerging (spatial) behaviors, not included in the SPICE models, design rules, DFM rules or DFM models. Our success lies in the ability to quickly analyze and transform these new variations into models and/or rules during technology transfer and technology derivative initiatives (alignment) and product design and development (adoption) efforts to reduce the technology lead times and associated costs.

Let us summarize different yield and manufacturability issues faced by R&D engineers on daily basis arising from the key manufacturing variation mechanisms (section B.6). If we do not come up with the correct response then we are obliged to increase the area of the design layout to insure the product functionality which is against the industrial slogan smaller, faster and cheaper. The most common DFM issues are classified as systematic or random defects. The random defects cannot be controlled; however the systematic defects can be easily controlled and modeled. Before proceeding into the role of DFM challenges in SI, let us quickly summarize the most well-known DFM defects and respective rules and/or models used in the CAD simulations.

B.7.1 Manufacturability Issues

The most common manufacturability issues irrespective of the technology are discussed in order to demonstrate the impact of the key process behaviors on the resulting IC chips. The SPICE models take into account the effect of geometric feature variations and predict its influence on the electrical parameters of the product. In the example below, at the F1 E1 condition, the images are printed larger than normal because of the higher light intensity with greater diffusion exposes more of the resist and the opposite occurs at the F1 E-1 condition (Figure B.25). It is also likely possible that contact over the via is misaligned and totally misses the metal resulting in a functional failure [Hwaiyu, 2005).



a) Worst case for exposure focus | b) Overlay error resulting misalignment

Figure B.25 – Extreme process conditions

These worst process conditions when supported with varying etch conditions, unreliable depth of focus and CMP effect; often result in systematic manufacturability issues (Table B.5):

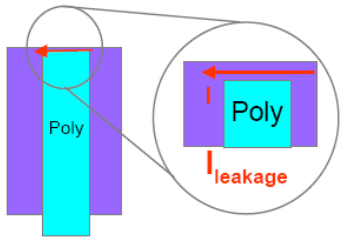
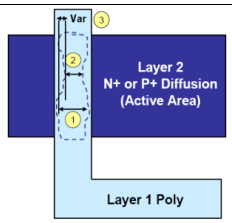
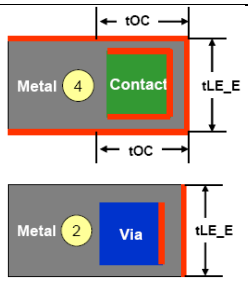
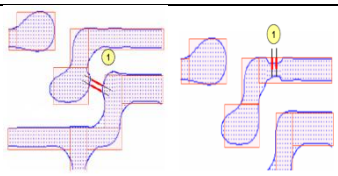
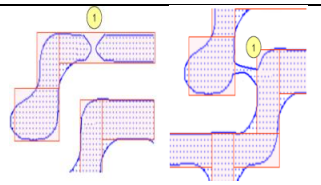
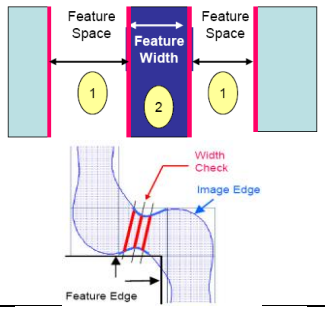
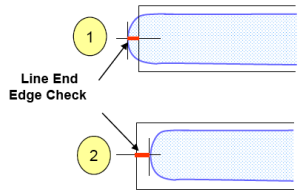
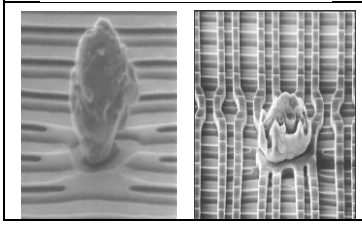
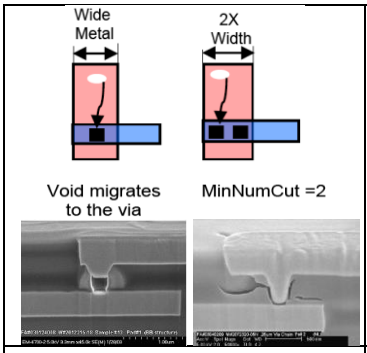
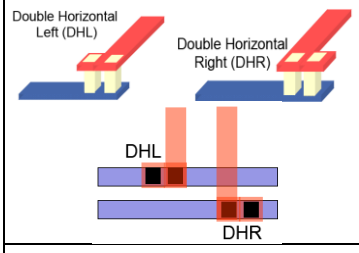
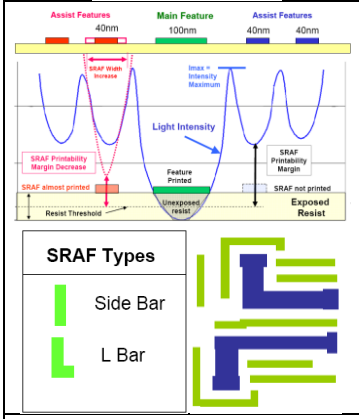
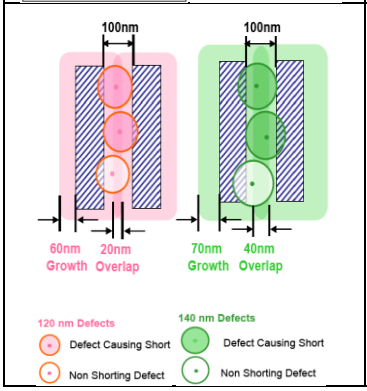
	**Transistor end cap extension:** The gate poly of the transistor must extend beyond the active area to make sure that there are no current leakage issues. This extension is known as end cap extension and the mask, overlay and exposure focus error often result in the reduction of end cap extensions.
	**Gate Length:** The Transistor gate length is a critical factor for both timing and power. The gateLengthMin and gateLengthMax checks measure the transistor's gate lengths at several places along the transistor's channel and trigger the detector if the measurements fall within the specified {min max) measurement range.
	**Via and contact coverage issues:** The coverage is defined as the area of overlap, which can range from 0% to 100%, depending on the intersection of the two images on each layer. We can see in example that these errors often or might result in missing contact and vias causing serious issues in interconnection and contact formation steps. The tOC and tLE_E refers to outer corner and line end extension and are known as feature fragments.
	**Bridging and necking:** These are defined as if the space between two images or width of an image falls within a specified {min, max} range. It must be a positive value to avoid opens or shorts that result in functional product failure. The red lines represent the drawn features after OPC whereas the bleu lines represent the silicon image printed on the silicon.
	**Open and short:** The open defect is said to be formed if the distance across an image falls within the range of {0, max} and a short is formed when the distance between two or more images is greater than 0. It has to be a positive distance at the nominal focus exposure (FE) conditions.
	**Space and width defect:** The Feature Space is the distance from an edge fragment outward to nearest feature whereas the feature width is the shortest distance across an image from an Image Edge to the opposite Image Edge. The feature edges are composed of several different Fragment Types e.g. tE (edge), tOC (outer corner), tIC (inner corner) and tLE_E (line end edge) where tLE_E > tOC > tIC>tE.
	**Line end edge issues:** If image pushes out beyond the feature it is called Line-End Push Out (LEPO) whereas shrinking is referred as Line-End Pull Back (LEPB). These issues can become swear if not controlled specially for the contact and vias and current leakage.
	**Random particles:** The incidental particles are the major causes for opens and shorts on the layer being fabricated. They can randomly fall anywhere on the integrated circuit resulting in its functional failure are called random defects whereas systematic defects (hot spots) can be easily modeled and controlled.

Table B.5 – Systematic and random manufacturing defects [DFM Dictionary, 2004]

The deviations between intended and realized physical pattern(s) on silicon are called physical faults resulting in the functional failure of the product and deviations in the electrical characteristics are referred as the electrical faults. The physical deviations often result from the imperfection in manufacturing processes and the mechanisms behind these are classified as incidental and base line. The incidental mechanisms (gap fills, linear deposition etc.) can be resolved and controlled whereas the baseline mechanisms (flakes, particles in/on resist, incomplete develop etc.) are very hard to model.

We have discussed in chapter-1, section 1.2 that technology alignment and adoption lead times are critical for the success of SI; hence it is our ability to analyze and model the faults and associated mechanisms (Table B.6) that result in the success or failures. Let us summarize generic approaches (rules/models) adopted to avoid these manufacturability issues and yield limiters prior to discuss the role of DFM in the evolution of SI.

	**Minimum Number of Via Cuts:** In order to avoid the coverage issues that mismatches in the metal contacts, the min number of via cut rules are defined for each technology. Multiple metal contacts ensure current carrying capacity and reliability; however besides this fact multiple vias help us in avoiding the voids that migrate towards vias due to thermal stresses trapped in fabrication processes. In case of multiple vias, a single via can trap the void preventing its further migration to other vias that provide the interlayer connection.
	**Redundant Via Type:** According to the position of the redundant via, we categorize redundant vias by their order of redundancy and their orientation. In the example, we list the names for double redundant vias. The naming scheme is easily extensible to triple redundant vias or via arrays. Its orientation can be *(i)* horizontal right, *(ii)* horizontal left, *(iii)* vertical right and *(iv)* vertical left.
	**SRAF (sub-resist assist) Features:** These features provide a good control over the fabrication of the critical features and are used against the limitations of the lithography. They are very small in size and do not get printed on the silicon wafer; however they are added to the mask. These features diffract the light such that the illumination for the main feature turns out to be dense resulting in higher printability. The width, type and space of these SRAF features define the light intensity required for accurate feature printing. These feature are small in size and space is adjust so that they are not printed at all but the depth of focus of the illumination system becomes dense resulting in improved feature printing.
	**Critical Area:** The wires in example have a 100nm width and a 200nm pitch. The Critical Area for a 120 nm defect can be determined by growing the area around a wire by 60nm (the radius of the defect) in all directions. The result is a 20nm overlap between the two grown areas. If the center of a 120nm defects falls anywhere in this area, a short between the wires occurs. If the defect center is outside of the critical area, the defect doesn't cause a short. The Critical Area for a 140nm defect is larger than a 120nm defect because the overlap area is wider.

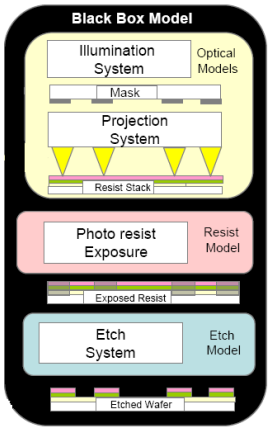
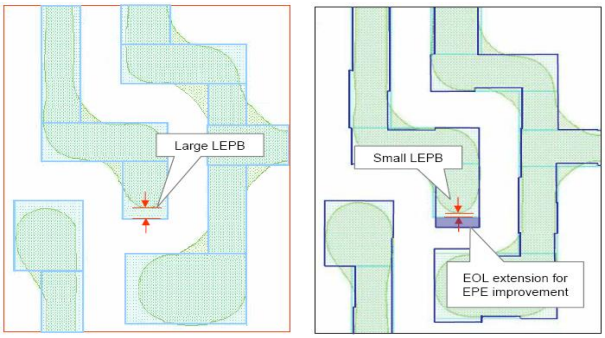
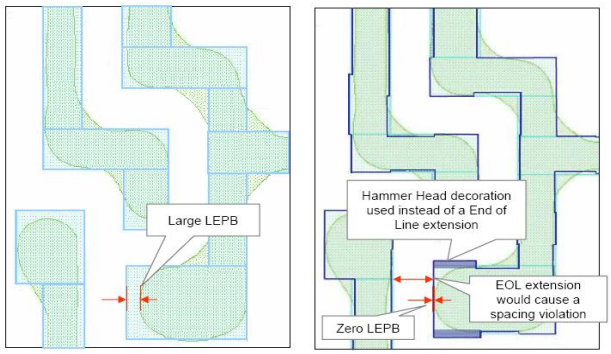
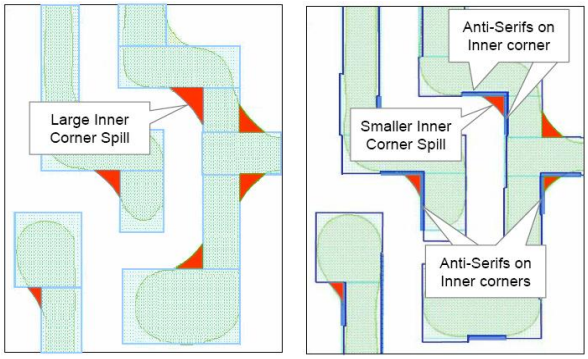
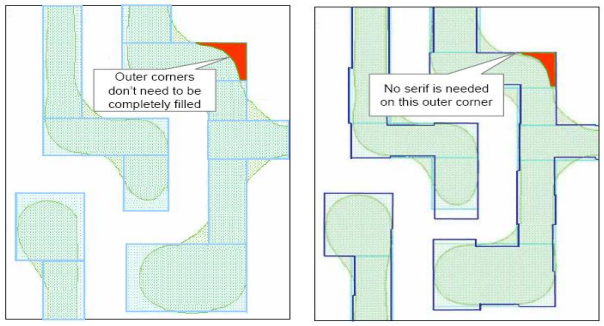
	**Process and Proximity Compensation:** The model based Optical Proximity Correction (OPC) and Process and Proximity Compensation (PPC), both rely upon physical models to identify the faults that occur during pattern transfer from a mask (reticle) to the wafer. A Black Box model lumps all these non-idealities in a single, multi-dimensional model. Unfortunately, the effects of one parameter change (a mask error, defocus condition, over exposure condition, under etch condition) can either compensate or multiply the effects of another parameter change. Physics based models; calculate these effects separately in the sequential flow. The various process models are calibrated with both measured and/or simulated data within a predefined range.

Table B.6 – Generic approaches to address failure mechanisms [DFM Dictionary, 2004]

The above list is just a summary of key failures and remedies (Table B.5 and Table B.6) from an exhaustive list. It is evident that engineers are always focused on addressing these manufacturing and yield limiting factors as discussed above. We have also seen that the efforts to address these issues either result in some rules (design/DFM rules) and/or models (SI2007). Let us take an example of a circuit and follow the most common techniques as discussed above to improve silicon features printability on the silicon image using OPC (Table B.7).

a) The end of Line (EOL) extension improves Line End Push Out (LEPO)	b) Hammer Head improves Line End Pull Back (LEPB)
	
c) Inner Corners (IC) have a negative offset to improve corner spill	d) Outer Corners (OC) do not have excessive corner fill
	
e) Via Coverage (Two Layer Overlap Area) is symmetric	f) Long straight lines have been biased correctly so OPC is required

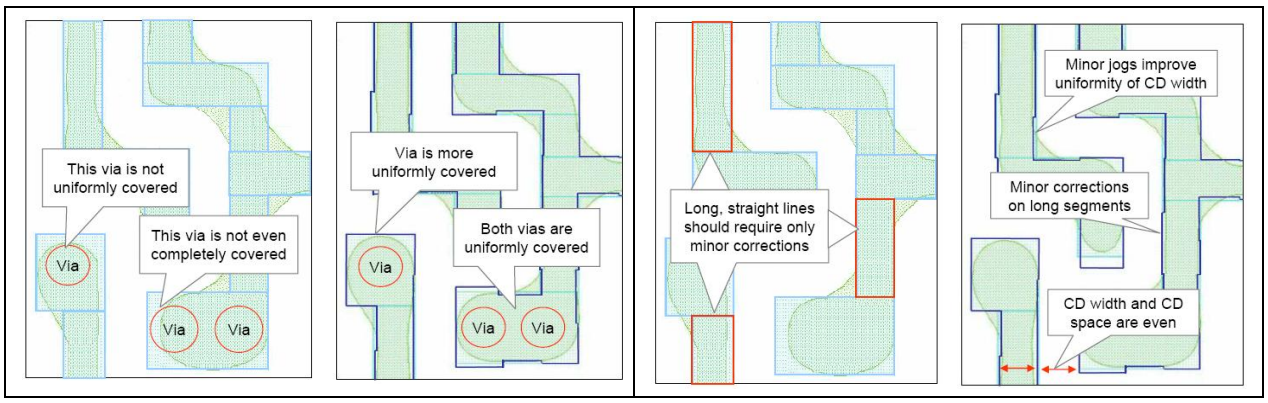


Table B.7 – Layout improvement efforts using OPC technique [DFM Dictionary, 2004]

The DFM optimization helps to make layouts more robust against both systematic and random yield loss mechanisms. Robustness against *systematic* yield loss mechanisms is particularly relevant to help avoid unforeseen design-process marginalities, especially during technology alignment and adoption efforts. The ramp-up experience has shown that design-process marginalities occur almost exclusively at layout configurations that are exactly at minimum design rule; hence they can be strongly reduced by designing just slightly (5-10%) above minimum design rule, where possible. The minor corrections can make a major difference for systematic yield.

The robustness against *random* yield loss mechanisms is relevant in volume production, where a small yield difference on a large number of lots represents a significant financial value. The defect density distribution is commonly used in conjunction with critical area models for yield predictions. The practical DFM implication is that minor corrections also give minor effects where random yield is concerned; hence the DFM corrections should be fairly large: via doubling, metal width/space > +30%, etc. It must be noted that the DFM actions for random yield should not lead to an increase in chip area, because an improvement in random yield will almost never outweigh the gross die decrease.

B.8 Summary and Conclusion

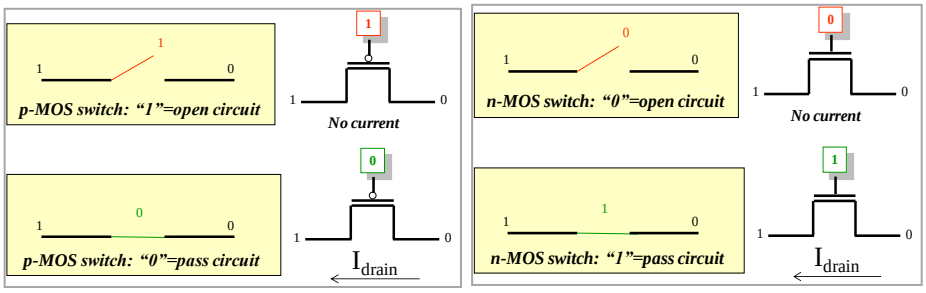
In this section we have reviewed that the semiconductor design, mask and manufacturing processes in detail. It provides us clear evidence that design CAD simulations, mask preparation and process control are the three lines of defense against manufacturability and yield losses. If the design and DFM kits are mature enough we can easily detect the manufacturability and yield issues as early as possible in the design phase. The mask provides us an interface between design and manufacturing where process imperfections is compensated to address potential yield losses; however mask errors could cause serious issues. We have also seen that design and DFM rules are never perfect because the DRM is developed during the new technology development in an alliance using test products. It is frozen upon its transfer to alliance partners' business models and changes are not authorized within the existing IDM fablite model. It is a potential opportunity where we can simply feedback rules and/or models from the newly emerging spatial variations to technology derivatives and improvement initiatives (alignment) and product design/development (adoption) efforts. Besides the fact that, any potential change in the DRM requires requalification of design libraries, the design engineers are ready to accept these changes if they are supported with the gains and/or losses in the potential yield.

Appendix C: CMOS Inverter Design and Manufacturing (An Example)

We have briefly reviewed the SI design and manufacturing process in Chapter-2 to provide users a glimpse of complexity involved in the integrated circuits. In order to strengthen the IC design and manufacturing knowledge and for the readers who wish to see those processes in more detail, we took a simple CMOS inverter (a circuit with 2 transistors) example and completed the design, mask and manufacturing steps. The important phenomenon that might raise the DFM challenges are also discussed and commented to improve the understanding of the DFM methods and their evolution.

C.1 The CMOS Inverter Design

The objective of this section is to go through the design and manufacturing steps of a simple CMOS inverter circuit and identify the major challenges being faced by the engineers at all steps. The inverter is truly the nucleus of all digital designs and is the basis of logic gates for complex electronics functions for the digital and analog designs. It is composed of n-MOS and p-MOS transistors together and its basic function is to inverse the signal. Before proceeding with the example, let us first simplify the operations of n-MOS and p-MOS transistors (Figure C.1). It is easy to interpret that when a voltage (potential) is applied to the gate, p-MOS results in an open circuit i.e. no current flows from the drain to source whereas in case of no voltage the current do flow.



a) p-MOS operation *b)* n-MOS operation

Figure C.1 – n-MOS and p-MOS current flows

The structure of the CMOS inverter is presented below (Figure C.2) along with its corresponding cross section of npn and pnp transistors for the manufacturing propose. In this inverter the output received is the inverse of the input. So we shall use CMOS technology to manufacture this CMOS inverter later. We shall not put emphasis on its design life cycle as it is a simple inverter. Our objective is not to pass through the whole design process but to highlight the key challenges faced by the designers, manufacturers and R&D engineers. It is also important to note that the design flow changes with the process technology and comprises of CAD tools and the sequence of these tools to simulate the design prior to manufacturing. The SPICE models are the key for the CAD simulations of the design. These models are defined at the device level and are composed of mathematical equations to compute the electrical parameter to characterize the functionality of the design. If we minutely look at the function of an electronic circuit, we can simply conclude that it is all about flow of current. The physical defects, parasitic (unwanted resistance and capacitance) and uncontrolled variations in the geometries of the devices and interconnects are the key reasons for the manufacturability and yield issues.

A simple example of the SPICE model for a device is the mathematical equation that computes the Ion and Ioff currents. Ion is the current that flows from source to drain when threshold voltage (Vth) is applied on the gate and Ioff is the amount of the current flowing through source and drain even when there is not voltage at the gate. It is also called leakage current that consumes unwanted power. The amount of these currents depends on the gate oxide thickness (w), poly gate length (L) and the Ion implantation. The W and L are controllable parameters whereas implantation is not completely systematic. The general equation for the Ion and Ioff computation is presented below in Figure C.3. These and like equations used in the SPICE models are totally technology dependent and are delivered with each new technology. We can see that Ion and Ioff are proportional to W/L ratio, typically L is kept to the

minimum dimension allowed in the technology and is to be drawn as indicated in the schematic but W is allowed vary. The μ_o, q, K are the process technology constants.

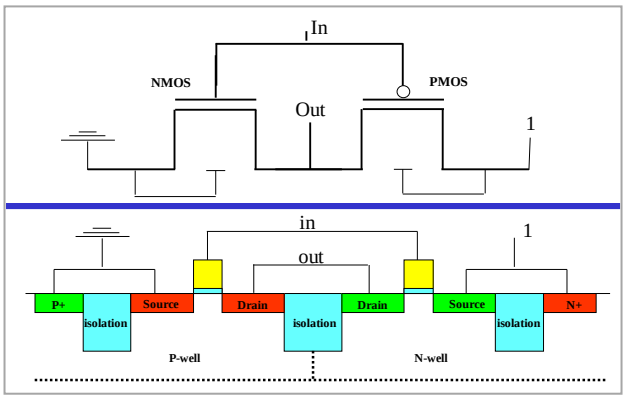


Figure C.2 – CMOS inverter structural presentation

The design is simulated using these SPICE models and process variations that might affect the geometric shapes, resulting in likely significant drifts. The computed parameters are analyzed against the target specifications and corrections are made until these parameters fall within the model corners, which is essential for the correct operation of the design.

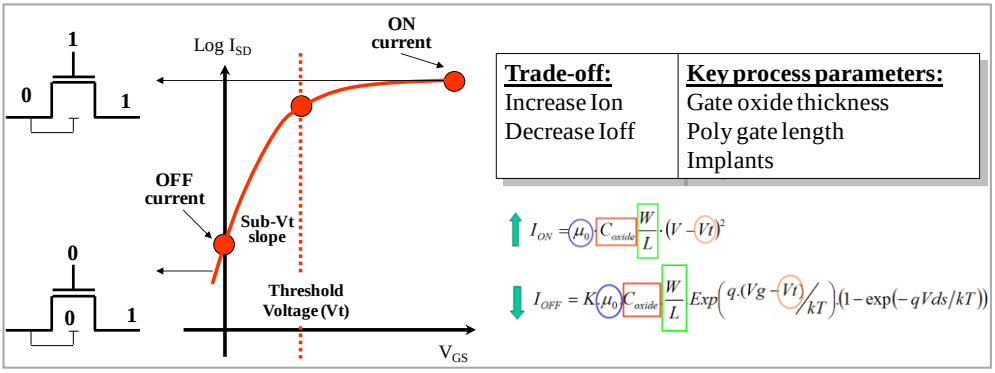
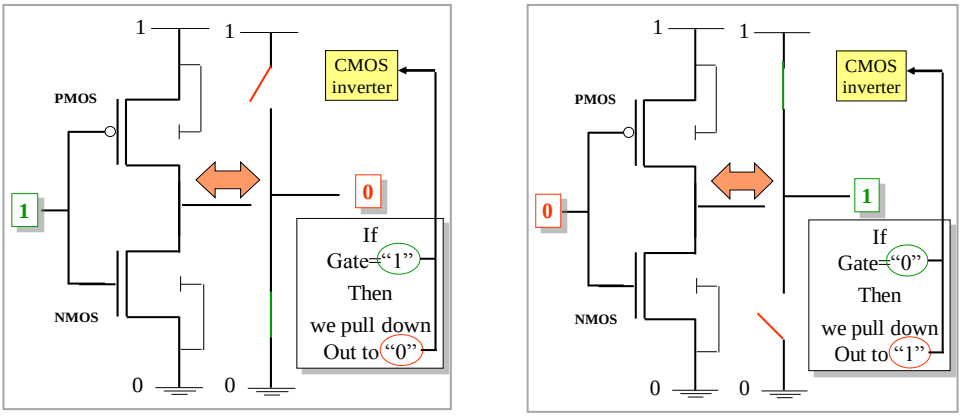


Figure C.3 – Ion and Ioff currents and trade off

Let us analyze the design of the CMOS inverter according to the design phases defined in Chapter-2, section 2.2.3. The schematic (logical) design and intended simulated functionality of the CMOS inverter is presented in Figure C.4. It is self explanatory that if 1 is applied at the input, the output is inversed (0) and visa versa.



a) CMOS inverter with input 1 b) CMOS inverter with input 0

Figure C.4 - Structure of the CMOS inverter and operational outputs

Let us now define symbols to be used in the CAD design for the above [Lee, 2005] CMOS inverter (Figure C.5).

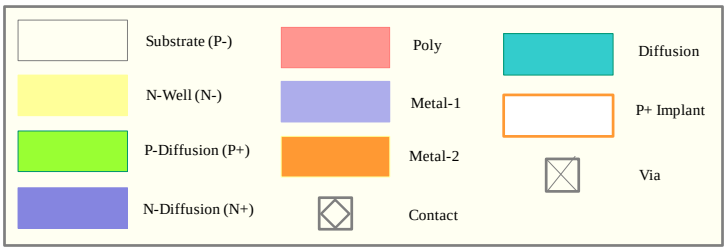
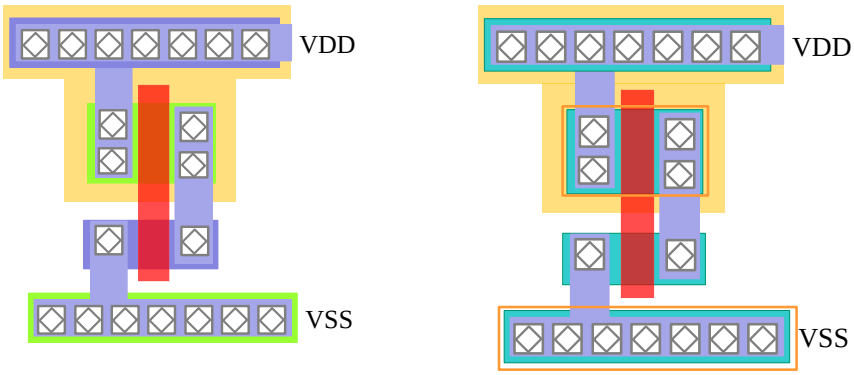


Figure C.5 – Colors and notations for CMOS inverter CAD design

We start with two potential physical designs of the CMOS inverter. The design as presented in Figure C.6a uses the p and n diffusion layers for the formation of n and p MOS transistors and p-type substrate whereas the design presented in C.6b uses the diffusion and p-implant in the structure of the inverter.



a) CMOS inverter with p-diffusion b) CMOS inverter with p-implant

Figure C.6 – Physical design of CMOS inverter [Lee, 2005]

Both physical designs (Figure C.6) are different in construction but they refer to the same CMOS inverter schematic (logical design). The next step as per design flow is the optimization of this layout and extraction of the electrical parameters. It is the good time to introduce the concept of DRM (design rule manual) which is referred as the bible for designers. The DRM includes all the physical design layout rules for active, passive components and metal interconnections at each layer (BEOL). These design rules are included in the design kits specific for each technology and new product designs are simulated to find design rules violations. If a design is constructed using design libraries where reusable circuits are already qualified against these rules, then design passes the DRC test easily. The power, timing and area constraints force modification in the design manually that requires the extraction of parasitic/electrical parameters to ensure its functionality, manufacturability and yield. It is also evident from the design (Figure C.6) that source, target and gate are covered with maximum contacts; it is because of the fact that the mask misalignment could lead to a missing contact resulting in failure of the circuit; hence maximum coverage as contacts is a normal practice. It helps us in reducing the resistance between the contacts and diffusion as well as increasing the flow of current.

The DRM for the technology is a confidential document, so we cannot present here the DRM itself; however some generic design rules are presented here to demonstrate how a design is improved throughout the physical design phase. We start with a generic transistor example (Figure C.7) with 0.2µm channel length (L) and 20µm width (W). The layout optimization starts by adding maximum number of contacts for the source and drain as shown in C.7a. The second improvement made in the transistor layout is about its aspect ratio (ratio of length to width of the feature), so we divide this transistor into four transistors (connected parallel and 5um width) to ensure good aspect ratio that is manufacturable with fewer defects. The transistor

2 and 4 from left to right are flipped to have reversed their sources and drains to transform the transistor layout into a compact transistor as shown in the figures C.7b and C.7c respectively.

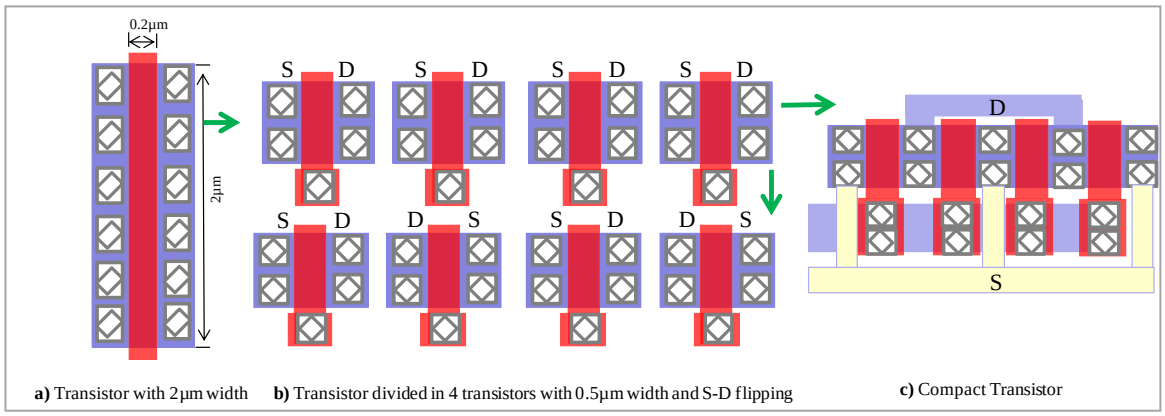


Figure C.7 – Layout optimization of CMOS inverter [Lee, 2005]

We know that the layout cannot be fabricated exactly as drawn in the layout due to mask misalignments and manufacturing processes imperfection. The objective of the DFM methods is to capture these limitations and put them in the DRM so that during CAD simulation low yield and impossible to manufacture layouts can be identified and corrected. The main parameters of these transistor model (e.g. see Ion and Ioff current equations at Figure C.3) variations are DW and DL that shows the delta difference of drawn width and length from the effective width and length.

Now let us discuss another example where a design layout is qualified/validated to ensure its functionality, manufacturability and yield. This is an example where imperfect process could result in effective channel length variations, ultimately affecting the speed of the current that cause serious timing delays and/or power losses (Figure C.8). Such variations often result in fast or slow devices that might have very serious impact on the circuit timing.

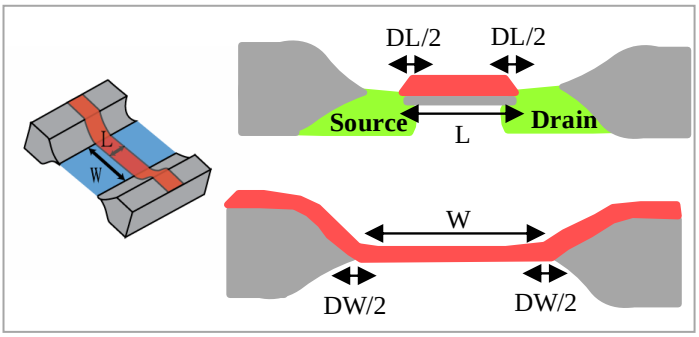


Figure C.8 – Delta in drawn length (DL) and width (DW) of transistor [Lee, 2005]

Let us consider an example where DL is 0.015um and DW is 0.045um. The fast transistor is simulated with narrower L and wider W (negative DL, -0.015um) whereas the slow transistor is modeled using a wider L and smaller W (negative DW, -0.045um). The following table (Table C.1) shows L and W of the transistors from the first and second optimization efforts, and the geometric variations at the slow, typical and fast corners that emulate the manufacturing process tolerances.

		Fast	Typical	Slow
Maximum Contacts	*L(um)*	0.185	0.200	0.215
	W(um)	20.045	20.000	19.955
Compact Transistor	*L(um)*	0.185	0.200	0.215
	W(um)	20.180	20.000	19.820

Table C.1 – Model corners for fast, typical and slow devices

The impact of layout optimization techniques do not seem to have a strong variation in the two optimization steps on width; however the length undergoes significant variation due to under or over etch and a compact transistor formation. The impact of the length variation (0.18um) between slow or fast corner against typical corner is 4x in comparison to the one without compact transistor. This could pose a circuit performance deviation for the circuit designers if left unaccounted. The designers perform Monte Carlo simulation by adding small statistical variation to W and L of every transistor in the circuit with an objective to center design for the yield improvement. The layout can be challenging for the circuit designer to optimize it for better yield.

The subsequent layout optimizations methods applied on the circuit layout are *(i)* Speed up the transistor with higher frequency response by reducing the parasitic components (resistance and capacitance), *(ii)* clean up the substrate disturbances from minority carrier and coupling noise, (iii) balancing area, speed and noise, *(iii)* stress relief (insert dummy transistors), *(iv)* protect the gate (control charge accumulation at the poly gate a.k.a. antenna effect) and *(v)* improve yield (avoiding single via or contact, spacing, end-of-line variations, metal coverage of contacts and vias). All these optimization efforts result in design rules programmed into design kits.

In this section we have drawn the logical circuit (schematic) for CMOS inverter followed by its transformation into physical design. The CAD simulation using technology based SPICE Models for the extraction of electrical parameters ensure manufacturability and functionality. The physical (transistor level synthesis) requires further simulations based on manufacturability variations. It is based on the fact that we are not able to manufacture the geometric shapes like they are drawn in the schematic; hence layout optimization is required followed by the CAD simulation to ensure timing, power and leakage. We have not physically simulated the logical and physical designs because it shall result in more than 2000+ electrical parameters; however we have presented and introduced the whole process. We have optimized the layouts manually whereas other optimization techniques are listed for reference. The CAD simulations follow an automated process using software tools which are supported with design kits, process kits and DFM kits. It is also important to note that we have not included discussion on the BEOL interconnections but the CAD simulation flow is exactly similar with the only difference of SPICE models. The validated designs result in the generation of netlist which is used as an input for the mask preparation.

C.2 Mask Set for CMOS Inverter

Transistors are interconnects are manufactured using masks, hence these masks act as an interface between design and manufacturing domains. They are the second line of defense against yield limiting defects after CAD simulations. The mask numbers are technology dependent and are manufactured with very high precision. The structural frame of the mask is presented in Figure C.9 where most crucial, from the data analysis point of view, are the metrology and test patterns. They are added on the mask in horizontal and vertical scribe lines which means that wafer rotation is a must prior to the measurements, which result in varying coordinates. The coordinates provided by the machine offer coherent platform for a single-source analysis but multi-source analysis is impossible by the R&D engineers.

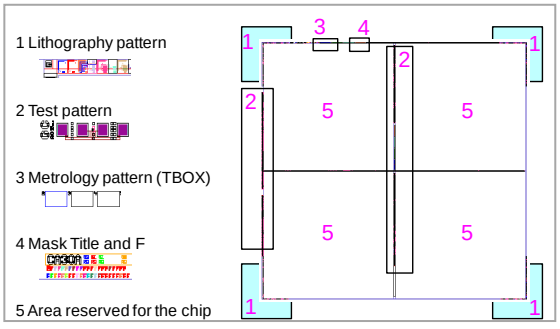


Figure C.9 – Frame structure of a mask

The most important issues towards manufacturing and yield are printability (lithography), polishing and etching (under/over etch) issues that result in geometric shape variations. It directly impacts the electrical parameters (2000+) resulting in functional or parametric yield losses. The cross section view of the CMOS inverter fabrication is presented in Figure C.10. In this circuit, Vdd refers to the drain voltage and GND is the ground connection. The y is the field oxide that separates two transistors and A refers to the poly gate contacts.

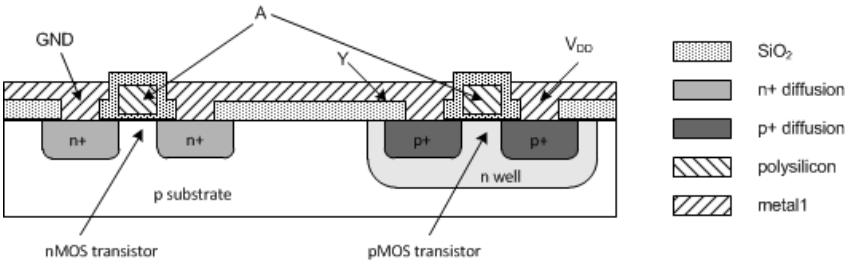
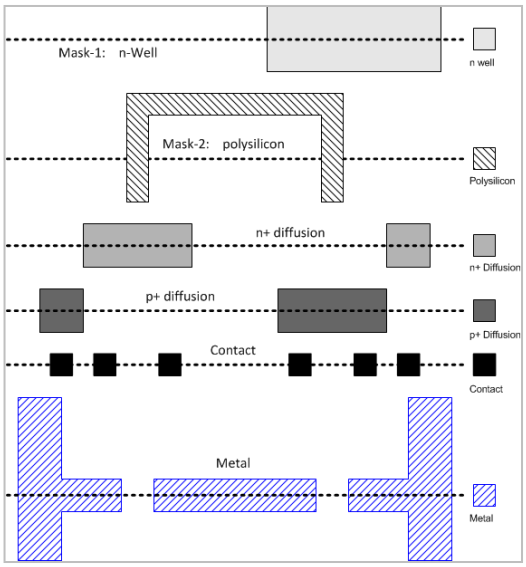
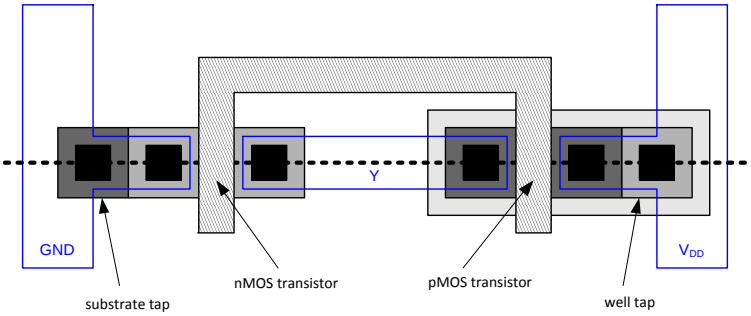


Figure C.10 – CMOS inverter's structural cross section view

A specific technology flow being highly confidential cannot be presented, however a generic flow is presented as under that shall help us understanding the corresponding mask requirements. This process includes *(i)* creation of the well formation, *(ii)* field and gate oxide growths, *(iii)* deposit and pattern polysilicon layer, *(iv)* implant source, drain regions, substrate and *(v)* create metal contact and *(vi)* create metal1 to metal 7 interconnects [James, 2000]. We shall need at least 6 masks for this CMOS inverter (Figure C.11).



a) Individual masks for CMOS inverter manufacturing



b) Combined masks and resulting CMOS inverter structure

Figure C.11 – Mask set for CMOS inverter

C.3 CMOS Inverter Manufacturing Steps

We start with a raw wafer and first manufacturing process step is the oxidation step (Figure C.12) where S_iO_2 (dielectric) is grown on the wafer surface at 1200ºC in an oxidation furnace with H_2O and O_2. The principle objective here is to protect the surface of the wafer from the subsequent manufacturing steps. The wafers are heated in the furnaces containing oxidant, usually O_2, steam or N_2O resulting in the deposition layer of an oxide. It is important to note that no mask is used during this step.

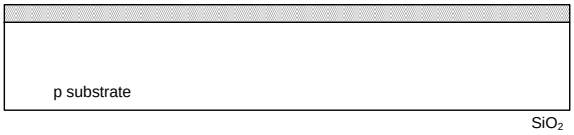
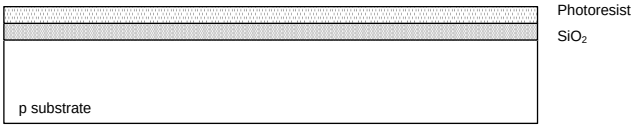
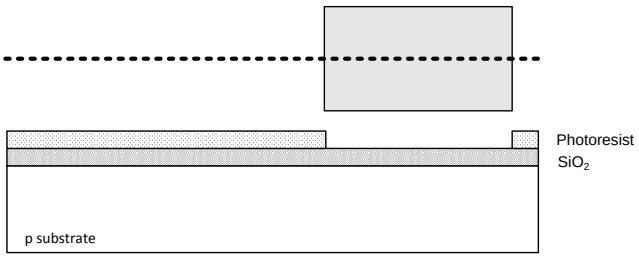


Figure C.12 – Oxidation layer on the raw wafer

The next step in the manufacturing flow is the deposition of the photoresist material of about 1mm thickness. The positive photoresist material is normally used for this step while transferring the patterns from mask to wafer (Figure C.13). The n-well mask is used for the light exposure that transfers the pattern of n-well on the photoresist.



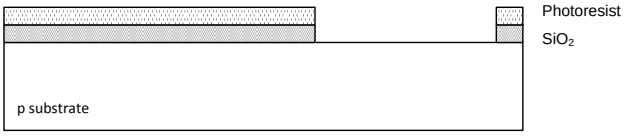
a) Photoresist layer deposited prior to light exposure



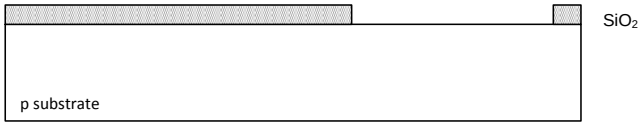
b) n-Well mask is used for the pattern transfer

Figure C.13 – Photolithography for n-Well

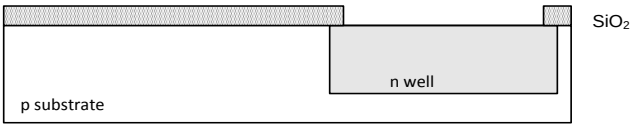
We etch the S_iO_2 so that the patterned area can be used for diffusion or implantation to form the n-well in our CMOS inverter. We use etch oxide with hydrofluoric acid (HF) that seeps through S_iO_2 (oxide) but it attacks only where resist has been exposed (Figure C.14a). We strip off remaining photoresist using mixture of acids called piranha etch (Figure C.14b). The n-well is formed by placing the wafer in furnace with arsenic gas and heating until atoms diffuse into exposed Si. The Ion implantation is performed with beam of As (arsine) ions that enter only the Si exposed are and are blocked by S_iO_2 (Figure C.14c). The last step is stripping off the oxide layer for the next operations (Figure C.14d).



a) Etching of SiO2 for n-Well diffusion



b) Stripping off the photoresist material

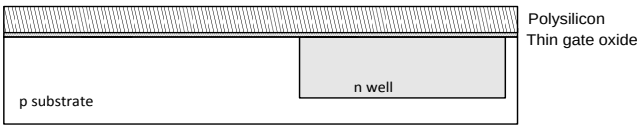


c) N-well diffusion and/or ion implantation

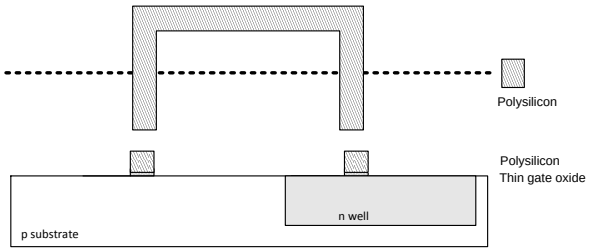
d) Stripping Oxide layer

Figure C.14 – n-Well formation

The next step is the polysilicon deposition (Figure C.15) using Mask 2. We start by depositing a very thin layer of gate oxide < 20 Å (6-7 atomic layers) in the furnace. This step follows the Chemical Vapor Deposition (CVD) for the deposition of polysilicon. The wafer is placed in furnace with Silane gas (SiH4) that forms many small crystals called polysilicon.



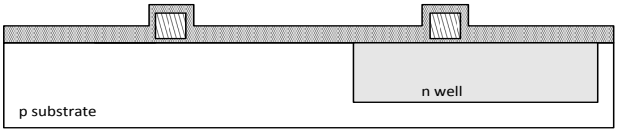
a) Gate oxide and polysilicon



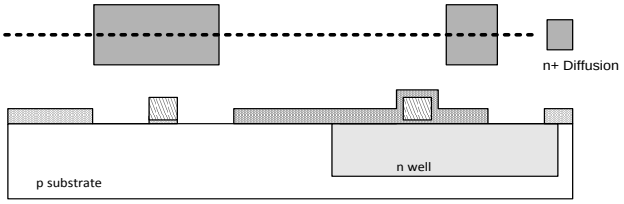
b) Lithography process for polysilicon

Figure C.15 – n-Well formation

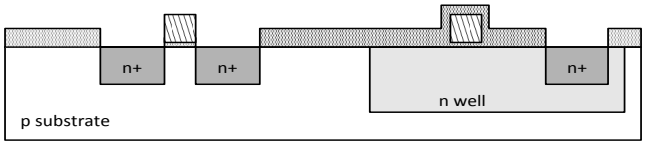
The n+ diffusion step is a self aligned process where we repeat the oxide (Figure C.16a) and masking steps for n+ dopants diffusion to form nMOS source, drain and n-well contact (Figure C.16b). The polysilicon step is better than the metallization step for self aligned gates because it does not melt during processing. The n+ diffusion is made (Figure C.16c) using diffusion process where impurities are absorbed on and beneath the wafer surface using heat. It is replaced with the ion implantation due to better control but for the source and drain regions it is still called diffusion. The oxide is stripped off to complete the patterning step (Figure C.16d).



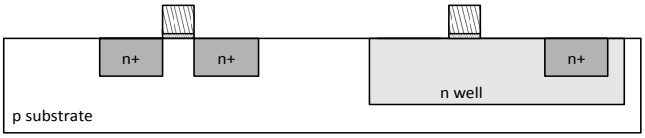
a) Oxidation layer



b) p-well mask and pattern transfer



c) n+ diffusion and source, drain and n-well contact



d) stripping off oxide layer

Figure C.16 – n+ diffusion layer

The next step is a p+ diffusion that forms the regions for pMOS source, drain and substrate contact (Figure C.17).

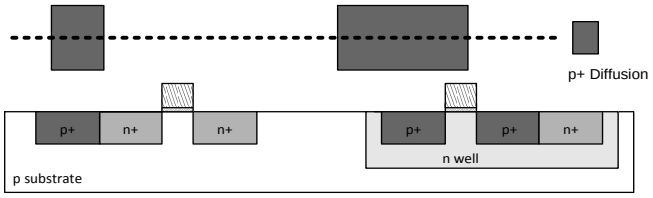
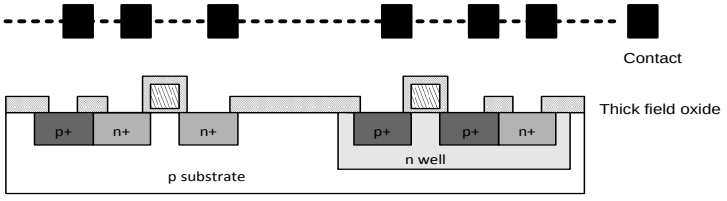
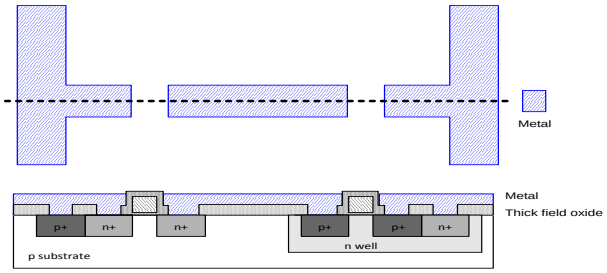


Figure C.17 – p+ diffusion layer

The final step includes the deposition of contacts and metallization to wire the devices n-MOS and p-MOS transistors. The chip is covered with thick field oxide and the contacts are etched (Figure C.18a). The aluminum is sputtered over the whole wafer and patterned to remove excess metal leaving behind the wires (Figure C.18b).



a) metal contacts



b) line metallization

Figure C.18 – Contacts and metallization steps

C.5 Summary and Conclusion

In this section we have simulated a simple CMOS inverter (2 transistor circuit) example against the design, mask and manufacturing steps with an objective to provide an overview of the design and manufacturing lows. It is a very simple example, however in reality designers and engineers use sophisticated tools to accomplish their tasks.

Appendix D: SMA (Spice Model Alignment) Tool

The proposed **S**pice **M**odel **A**lignment (SMA) tool is focused on helping R&D engineers for new front-end (FE) technology transfer, developed in the international semiconductor development alliance (ISDA). The use case diagram for technology transfer is presented in Figure D.1 where ISDA, technology R&D, process integration (PI) and design groups are the key stake holders. The ISDA alliance develops new technology and defines the perimeters. These perimeters are the key technology parameters along with their target and specification limits. The technology R&D group in the receiving plant is responsible for alignment of these parameters which is further classified as the technology maturity levels 10, 20, 30 etc. The PI team plays a key role during the alignment of these parameters which includes process flow, design rule check, and design rule manual alignment against device and interconnect SPICE model parameters. The design group develops design kits taking into account the manufacturing variations encountered during the alignment process to ensure manufacturability and yield during prototyping and normal production.

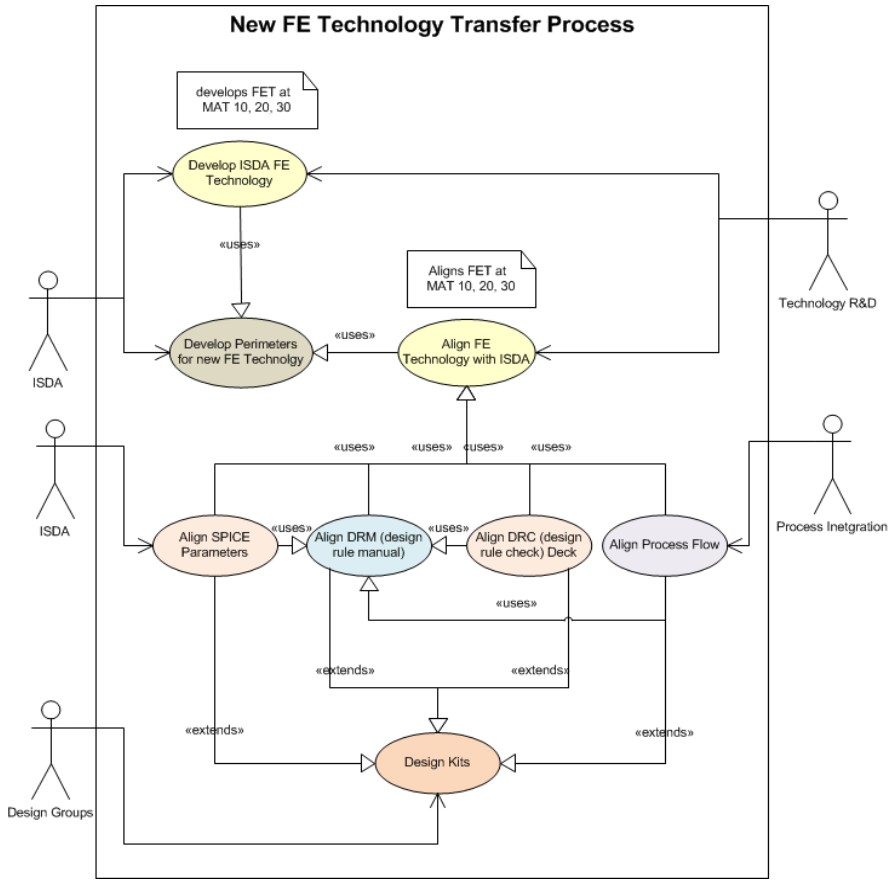


Figure D.1 - New front-end (FE) Technology Transfer

The scope of SMA tool is around the process flow alignment against the SPICE model parametric alignment by the ISDA alliance. The device and interconnect SPICE models maturity requires respective technology parameters alignment against the changes made by the ISDA alliance. It includes the management of standard and special conditions to be used during design CAD simulations of technology parameters using test product. The test structures a.k.a. vehicles are used during these simulations and are specially designed to represent the complexity of the target products to be developed using this technology. The process starts with the definition of device geometries [W, L] scaling to be tested through simulations for its compliance against target technology parameters and subsequent validation by manufacturing on silicon wafers. The simulated SPICE parameters computation often requires normalization in order to format the simulated results into unified measurement units. These SPICE model parameters are then related with

the PT parameters to generate PT specs. These generated PT specs are used for the process alignment. It is very important to normalize the SPICE models' simulated parameters because the PT specs measured on the production line have specific formats depending on the type of metrology equipment used. The lower part of the use case diagram is the part of proposed system where during process alignment any drifts or variations from the simulated parameters are further investigated using correlation between geometric shape variations. The challenges faced within this process are as under:

a) The names of SPICE model parameters from the simulations are not standard; hence PT-SPICE parameter correlation is to be manually controlled for each technology and each maturity level of SPICE models. Every technology has almost 2000+ parameters and normally engineers are allotted with 200 parameters each to manage and control the maturity and respective process alignment.

b) The computation of SPICE parameters using SPICE Models result in the tab limited data files (dat, csv etc.) file formats. These simulation files are hard to manipulate and engineers spend huge amount of time because multiple spice parameters are used and normalized to compute technology parameters. The varying file format of simulation files is a big issue along with the computation and normalization.

In order to address these issues and help PI and Technology R&D engineers, the SMA tool is developed and deployed to achieve following objectives which result in quick technology transfer and alignment.

- Manage SPICE-PT parameter relationship
- Generate PT-SPECS applying normalization and choice of corners
- Simulate parameters and edit normalization formulas

The sequence diagrams, data model and deployment diagrams are not presented here, however the tool is explained through different functionalities using GUIs. The reason for this is the confidentiality as well as space restrictions in the thesis. The SMA tool is versioned (X.Y.Z) upon each modification requested by the engineers where X → Major Version, Y → Minor Version and Z → Revisions. It is developed using Visual Basic 6.0 professional edition and MS-Access 2003 for prototyping. It is 2-tier, multi-user database application where all users are connected to the central shared database for authentication and profile management. It is a relational database and also serves as the share point for pre-simulated SPICE parameters among the application users. If the engineers requested SPICE parametric simulation results are in the parameters which are shared then those parameters are directly extracted and are not simulated. The functionality of the tool is presented through GUI explanations. The proposed tool is unique in the sense that based on the functionality selected; the local database is dynamically generated where analysis results are stored until and unless it is removed by the end users. The end users can share the simulation results in the central database for the other tool users.

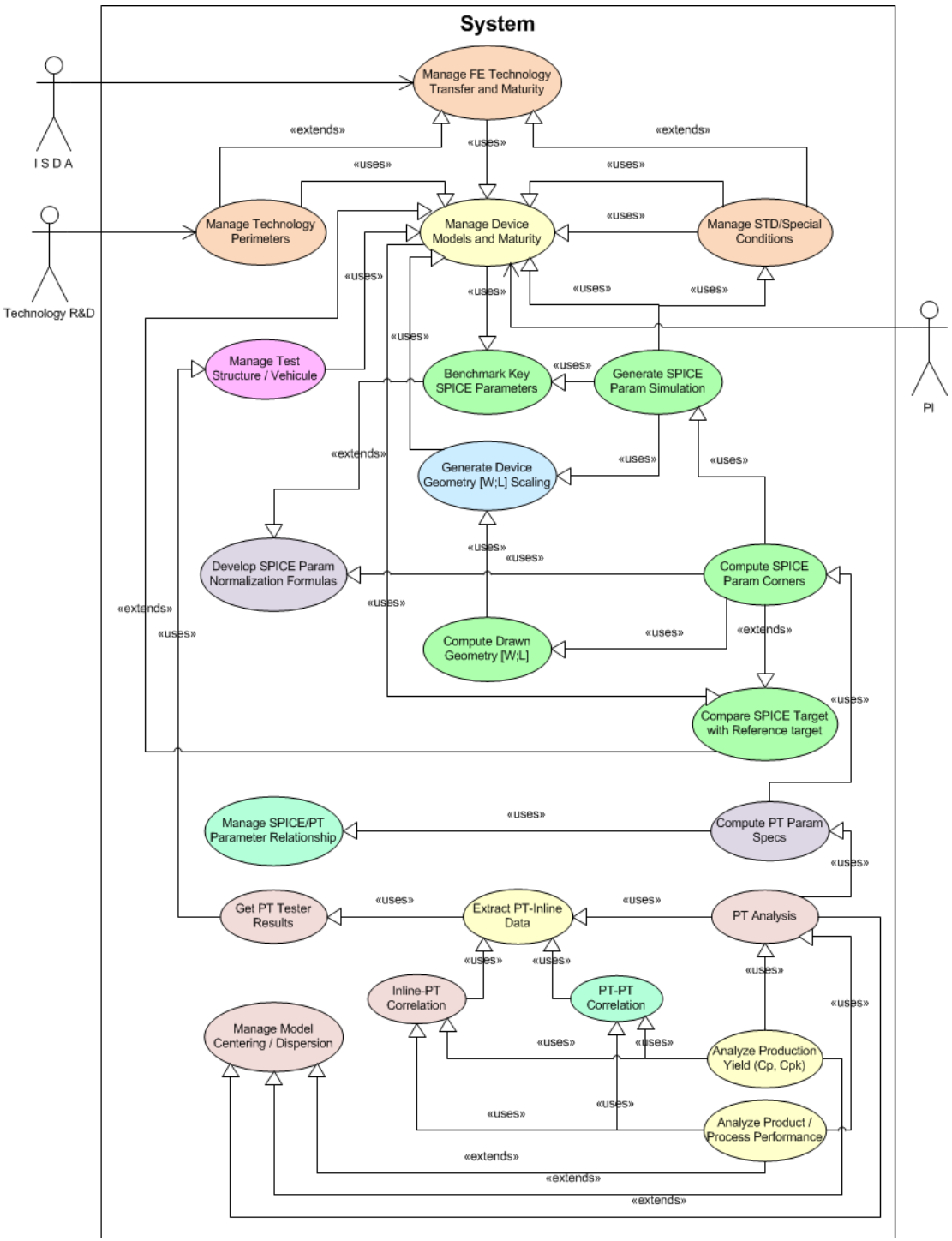


Figure D.2 – SPICE Model Maturity and Process Alignment

The first GUI is presented in Figure D.3 (a & b) where end users can browse simulation result files (SPICE parameters) by selecting the folder. The files are listed with the description of each file in folder/subfolders with its extensions and size. These files can be selected by double clicking on the files which are then added to the list box on the right; however user can select all files by clicking the button << Select All >>. The PT specs file is required to be browsed by each engineer for his own set of SPICE parameters which provide relationship between SPICE and PT parameters along with the normalization and computation formulas. Once selected, users import these relationships by clicking the button << import >>. The end users are also provided with option to filter Spice parameters, devices and PT parameters by selecting from the respective list boxes.

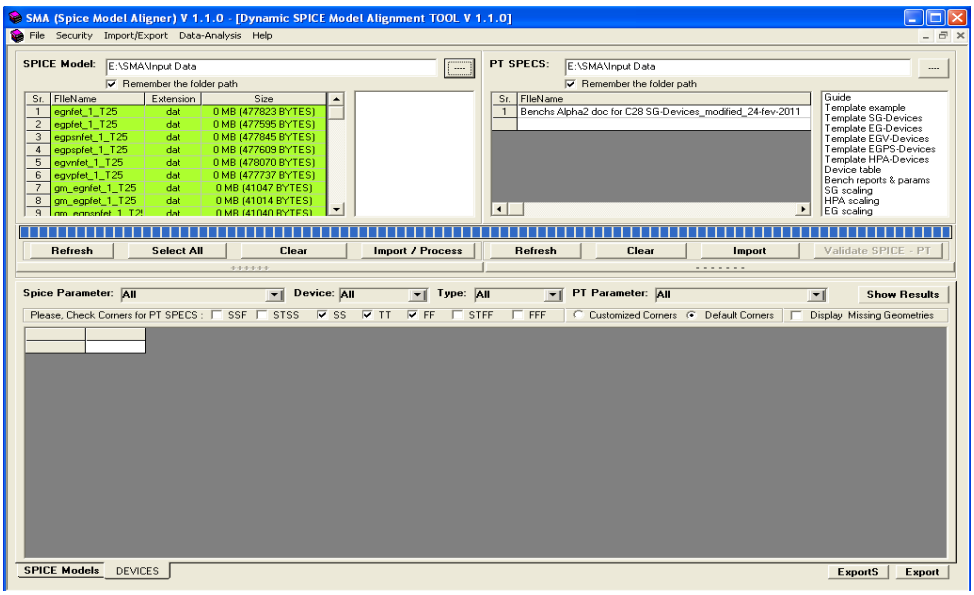


Figure D.3 (a) – SMA Tool Main GUI

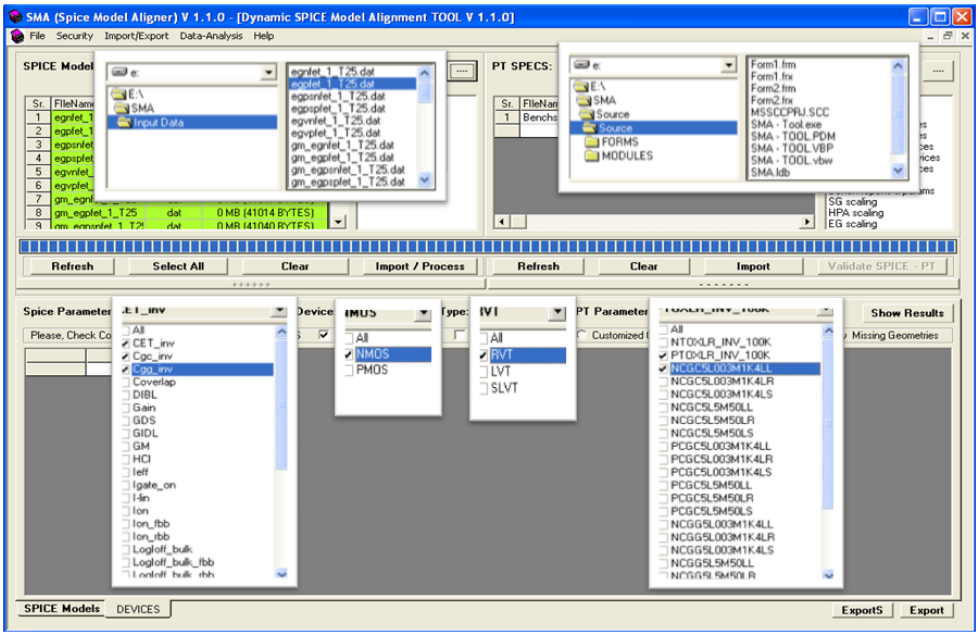


Figure D.3 (b) – SMA Tool Main GUI with Options

The PT-SPICE relationship file when loaded is presented in the Figure D.4 (a) where model file and description refers to the SPICE parameters along with respective PT parameter name and units. The Figure D.4 (b) presents formulae for PT parameters along with normalization. The users can change these formulas at any time where respective results are computed instantly based on the new formulas. The changes made can be made permanent for future. The normalization formulas do include the units and can be easily modified to ensure the final resulting units for compliance with the measure PT parameters. The user can double click on any of the listed formulas where he is provided option as a text box along with previous formulae for change. Once the changes are made, they can be made permanent by clicking the enter key. The users are required to select option "yes' from the message box to save it permanently, however these changes are made in the dynamic data base created on the local computer of the user. If user want to keep the relationship file as an excel sheet, he can extract the whole file by clicking the button on the bottom << Export >>.

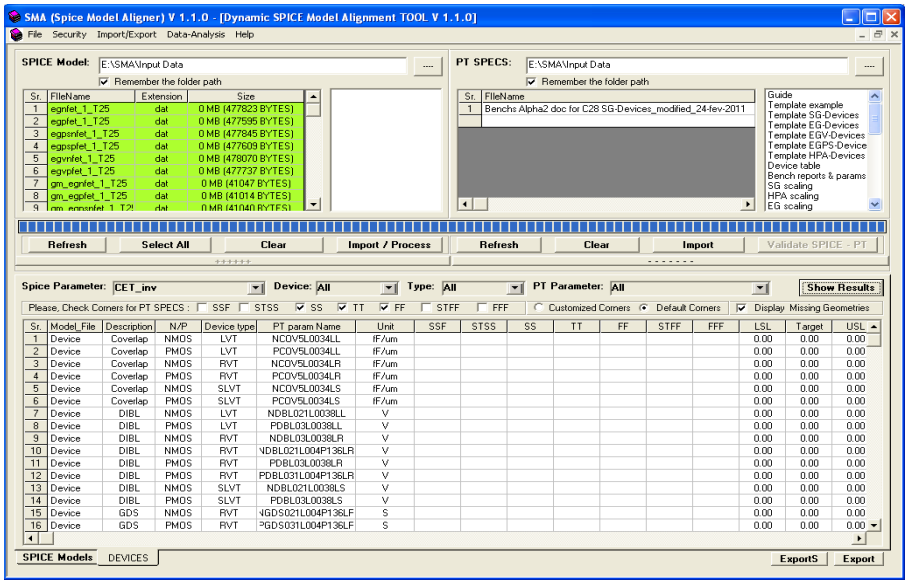


Figure D.4 (a) – Import of PT-SPICE Relationship File

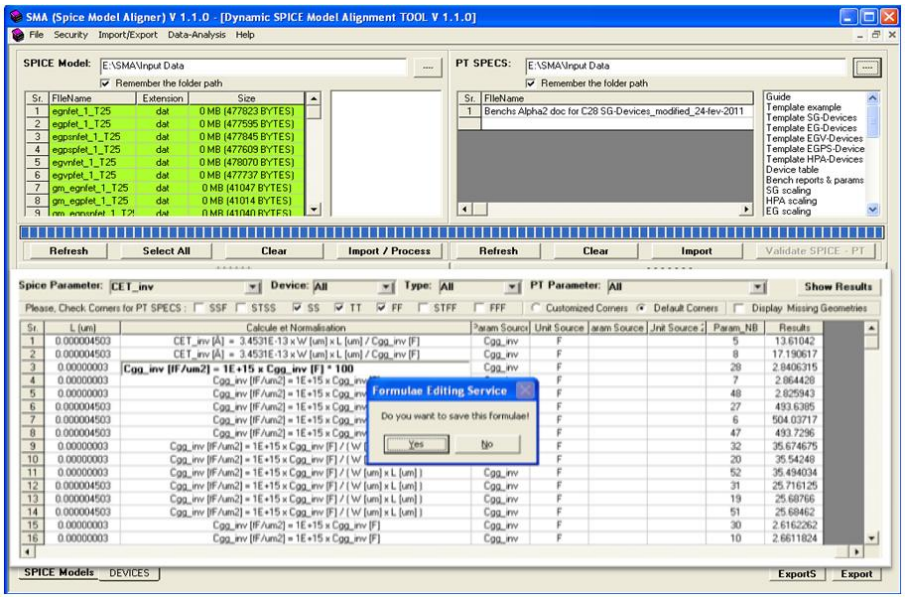


Figure D.4 (b) – Normalization Formulas and Instant Changes

One of the important algorithms implemented in SMA tool, is computation of the PT specs based on one or multiple simulated SPICE parameters. The implemented algorithm is based on the concept of "STACK" data structure which is traversed on last in first out (ILFO) principle. The algorithm cannot be presented, however simulation of the algorithm with an example formulae, is presented below in Figure D.5 (a & b). The formulae is presented in infix format as (1 + (2 * ((3 + (4 * 5)) * 6))) which is first converted into postfix notation as 1 2 3 4 5 * + 6 * * +. This post fix notation is then loaded into a stack and traversed for the computation of formulae. This algorithm has two steps as *(i)* infix to postfix notation and *(ii)* computation of postfix notation expression. The infix notation computation is presented in Figure D.5 (a) where we search for the inner most parentheses and transform it into post fix notation. The algorithm continues until the parentheses are finished. The parentheses are removed from the final post fix notation which is used in the expression evaluation presented in Figure D.5 (b).

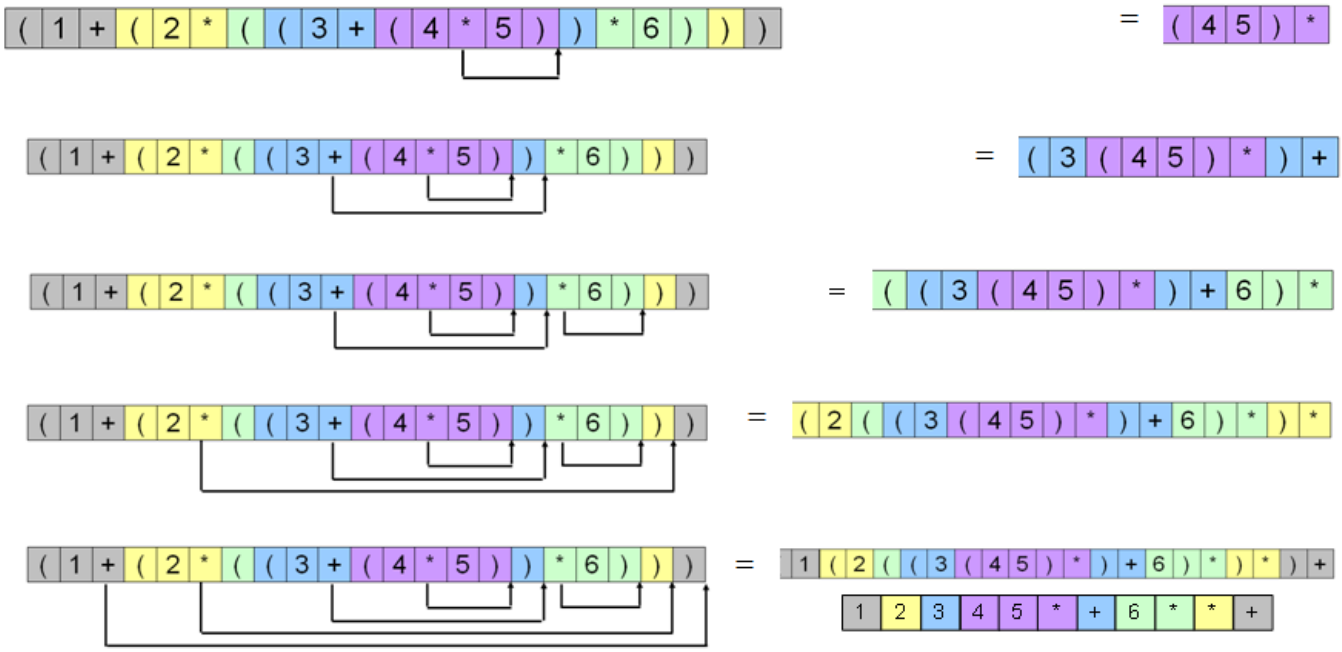
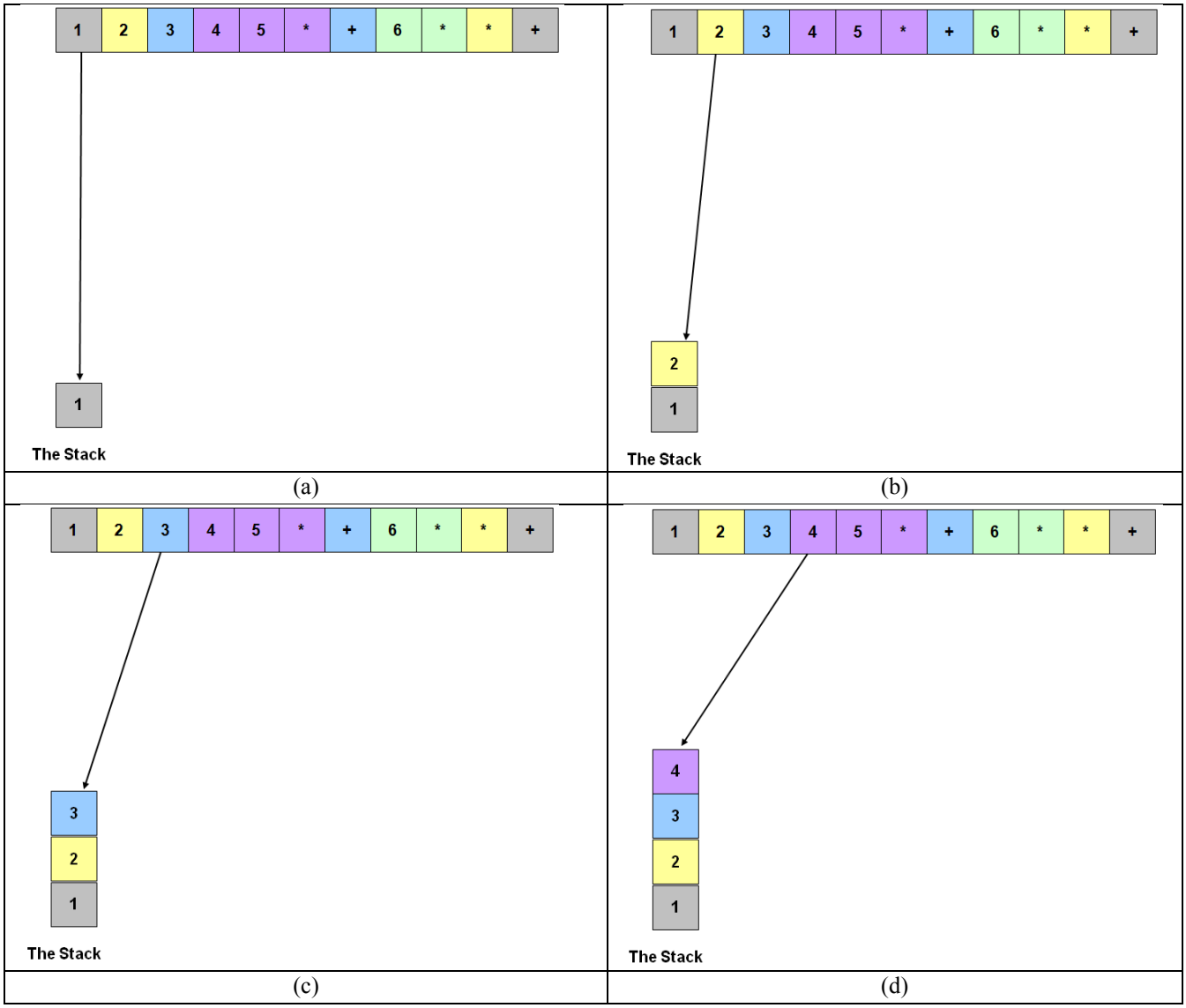
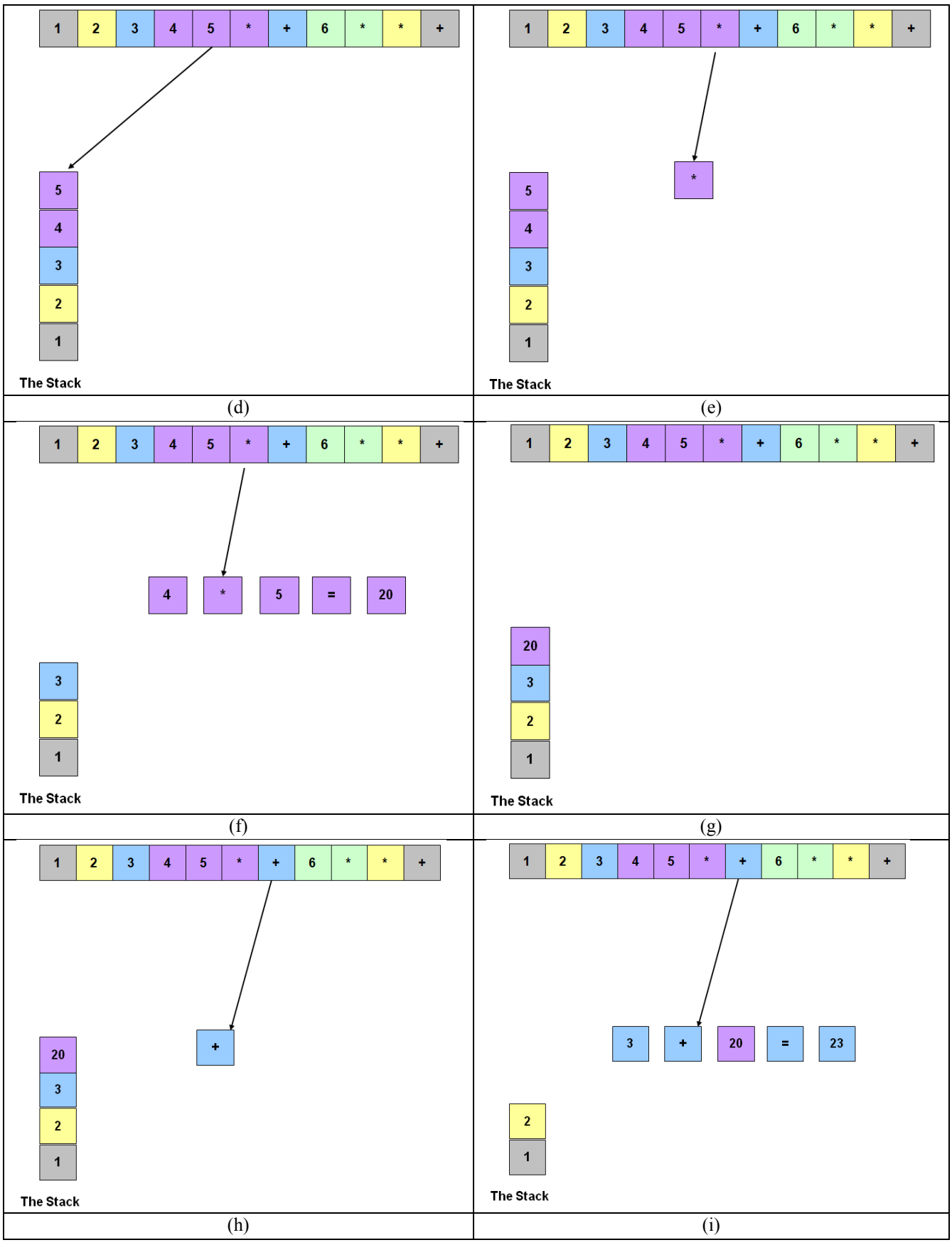
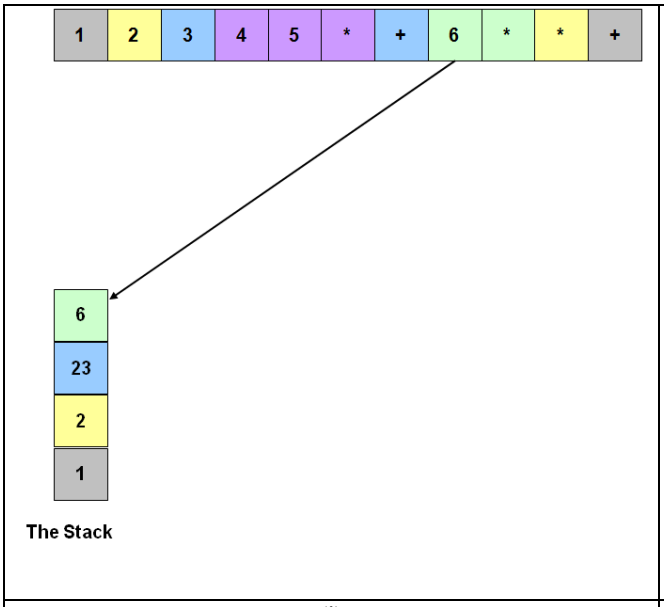


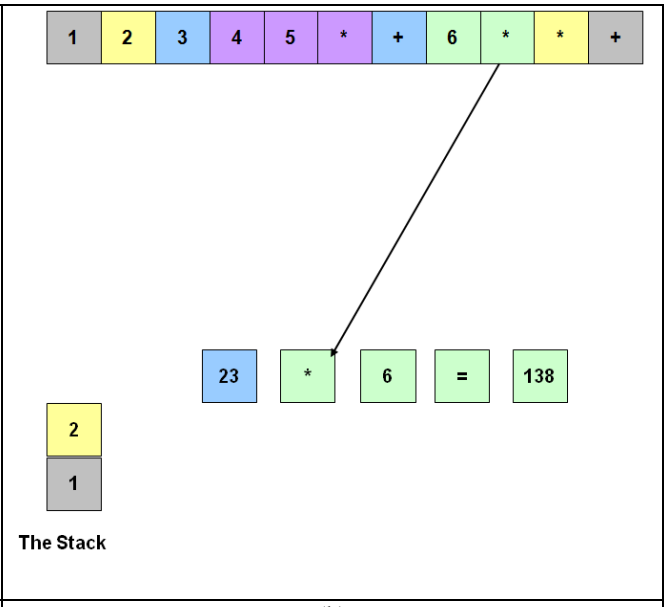
Figure D.5 (a) – Normalization Formulas and Instant Changes



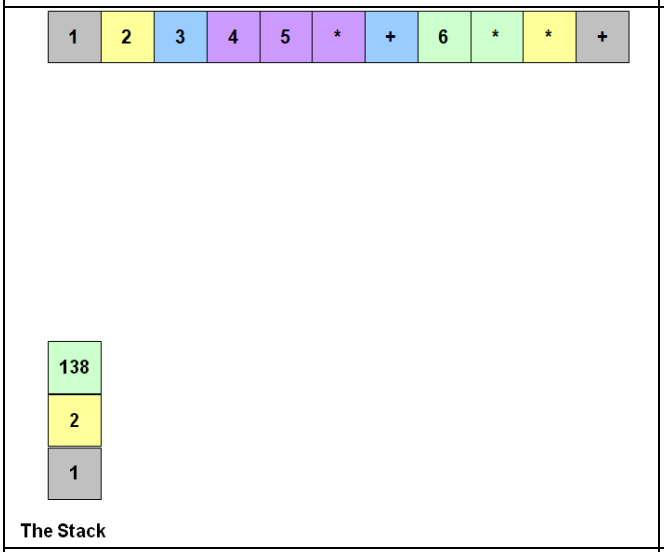

1 2 3 4 5 * + 6 * * +
5
4
3
2
1
The Stack
(d)
1 2 3 4 5 * + 6 * * +
*
5
4
3
2
1
The Stack
(e)
1 2 3 4 5 * + 6 * * +
4 * 5 = 20
3
2
1
The Stack
(f)
1 2 3 4 5 * + 6 * * +
20
3
2
1
The Stack
(g)
1 2 3 4 5 * + 6 * * +
+
20
3
2
1
The Stack
(h)
1 2 3 4 5 * + 6 * * +
3 + 20 = 23
2
1
The Stack
(i)



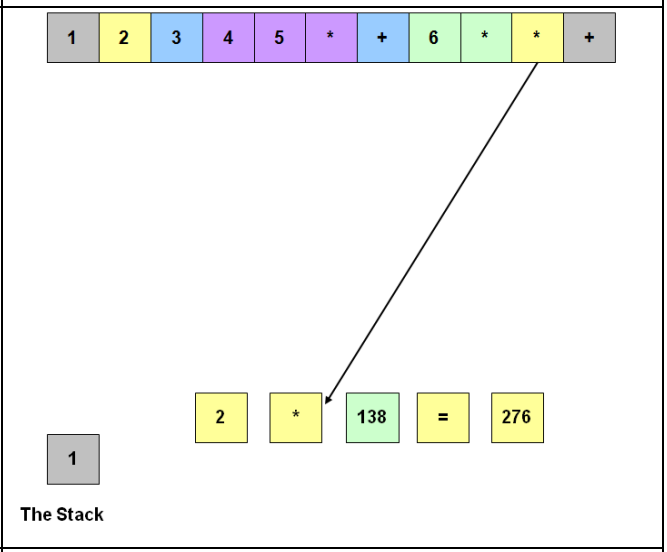
(j)



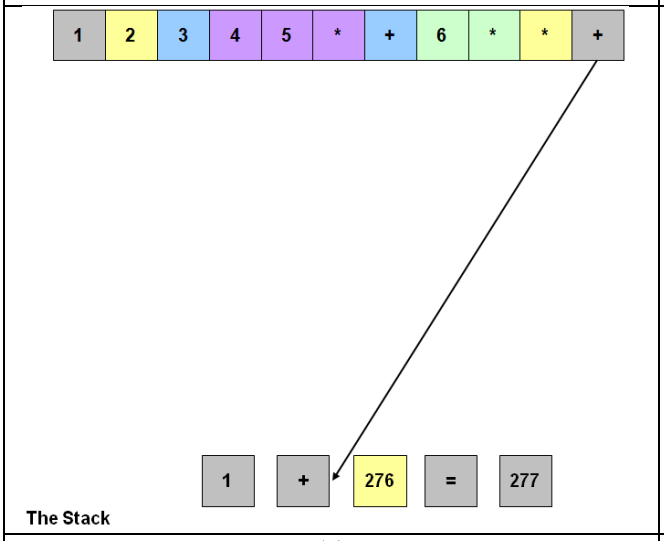
(k)



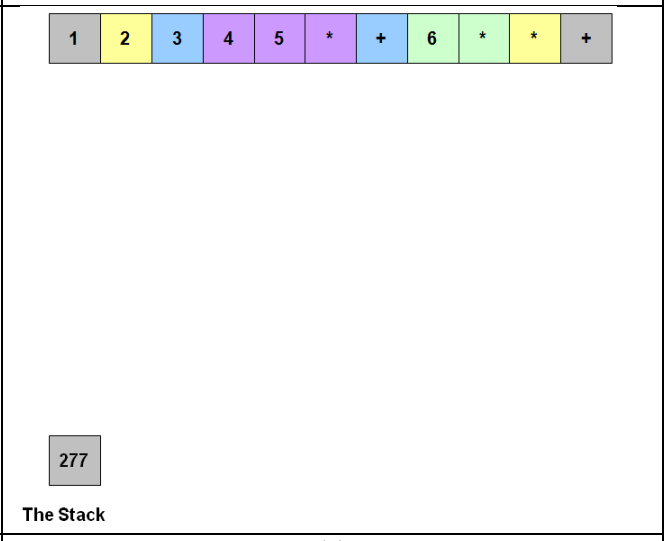
(l)



(m)



(n)



(o)

Figure D.5 (b) – Stack Structure based Expression Evaluation

The Figure D.5 (b) describes the sequence in which postfix notation based mathematical formula is read and evaluated. In the Figure D.4 (a) the options << customized corners >> and << default corners >> provide users to define the scope based on what PT specs are generated. The SPICE parameters are simulated based on the model corners as SS (slow), FF (fast) and TT (target). The ST prefix stands for STMicroelectronics where respective model used for the simulation is the company's internal model. The notations SSF and FFF stands for slow-fast and fast-fast corners. The user clicks on the button << Show Results >> to compute PT specs as LSL, Target and USL. The target PT spec is the target computed from the simulated SPICE parameters whereas LSL and USL specs corresponds to the minimum and maximum model corners as shown in Figure D.6 (a).

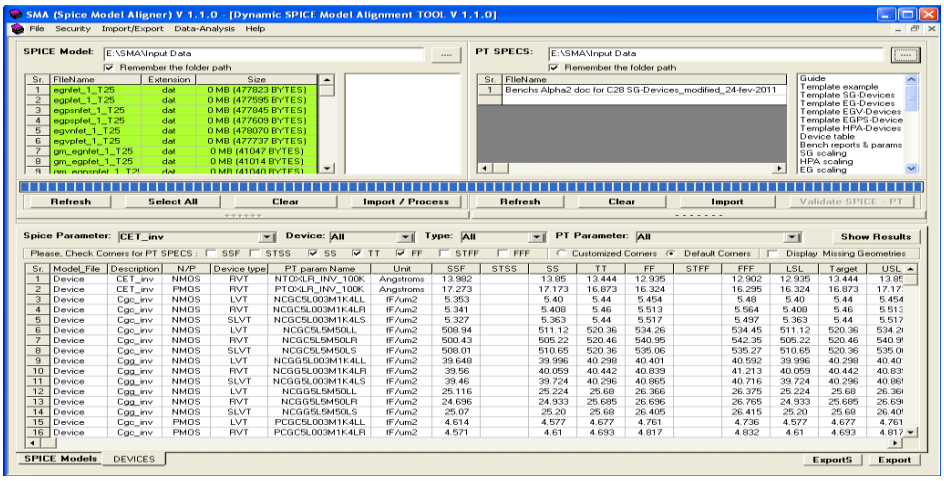


Figure D.6(a) – Computed PT Specs LSL, Target and USL

The users can extract the computed specs as whole or by selecting few lines by clicking in the buttons <<ExportS>> and <<Export>> where "ExportS" stands for the selected export as shown in Figure D.7. In addition to these functions the SMA tool has a lot of other interesting functionalities which cannot be presented here for space limitation; however the key algorithm and functions are briefly explained. The computed PT specs are used to analyze the PT measurement made during the prototyping on the test structures. If significant deviations are found, they are investigated prior to the adjustments in the process or equipment.

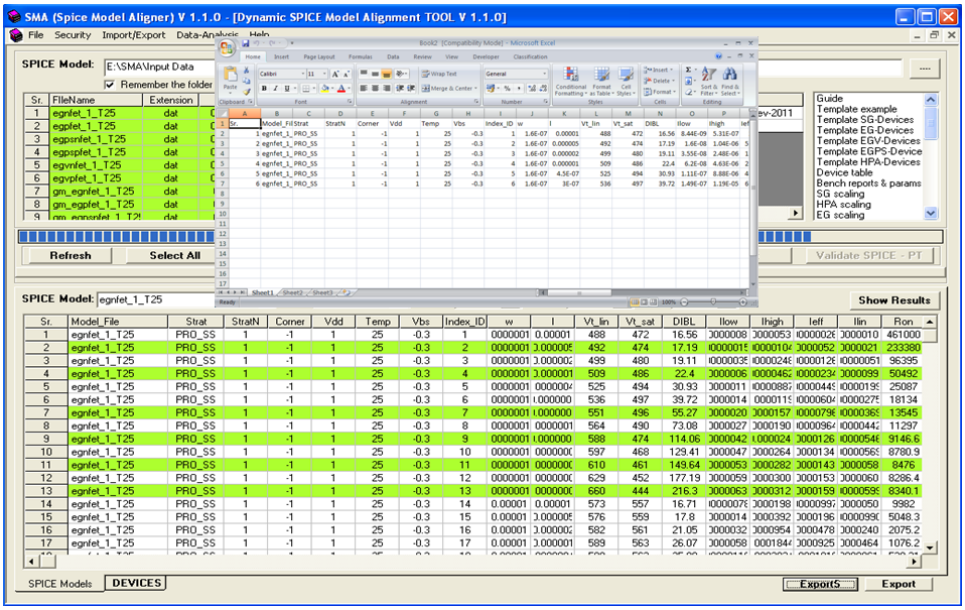


Figure D.7 – Computed PT Specs LSL, Target and USL

Appendix E: BEOL (back-end-of-line) Variance Analysis Tool

The BEOL variance analysis tool is developed for T2D aples team who is primarily responsible for the PT-variance analysis followed by silicon correlation to answer drifts and excursions. The IDEF model at two levels is presented in Figure E.1 below.

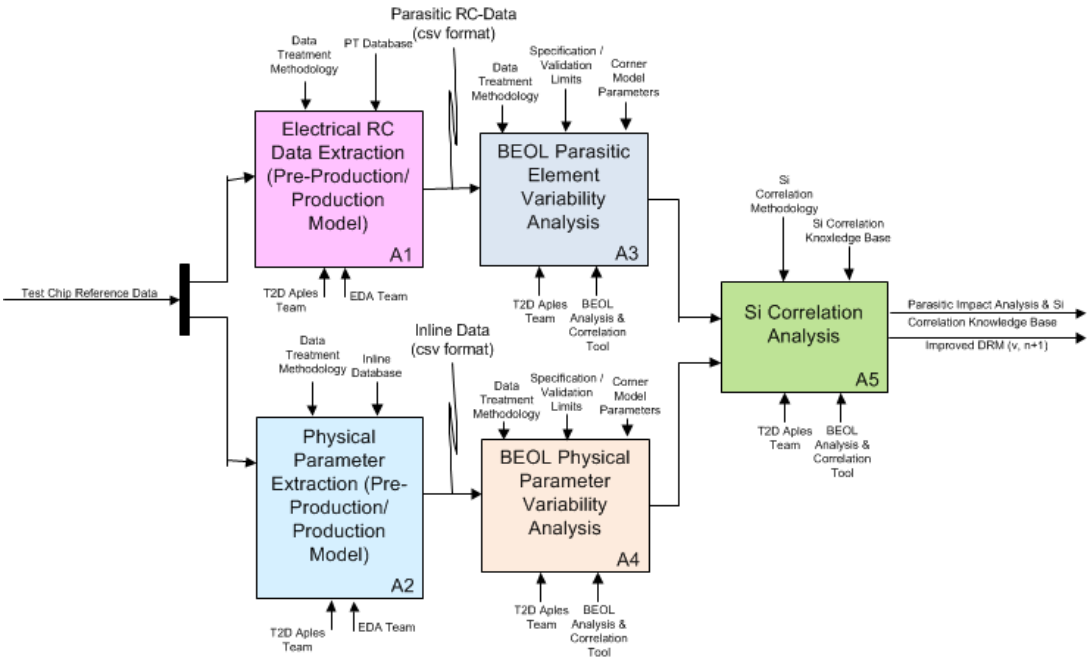


Figure E.1 – IDEF0 Model for BEOL Variance Analysis Tool

The most import tasks in this process are (i) PT data extraction (A1) followed by (ii) PT variance analysis a.k.a. BEOL parasitic variability analysis (A3). The physical parametric data extraction and variability analysis (A3 and A4) refers to the geometric measurements made on the silicon wafer which leads to the possibility of Si correlation analyses (A5). The BEOL variance analysis tool addresses the challenges faced by the T2D aples team during the steps A1 and A3. The data extracted using data extraction utilities; hence we do not find any format issues but data volume is quite high. The end users also filter data upon its extraction; hence the resulting data is normally distributed and do not exhibit significant variation. The engineers apply manual filter i.e. delete by hand certain values as outliers which might remove the potential drifts. The computation of different PT correlations based on scribe line and metal layers, is tedious job and engineers spend huge amount of time in data extraction and analysis to find the drifts which can be further investigated to find root causes.

The proposed tool offer following advantages to the R&D engineers:

a) It verifies the data format and completeness issues
b) perform filtered and non filtered data analysis
c) generate PT variance analysis graphs to be used for further investigation

The BEOL-Variance analysis tool is presented with its GUIs due to confidentiality issues, however it uses multidimensional data model at the back end to improve the computational efficiency. It has reduced the hectic work of weeks into 5 minutes processing and as a result, engineers focus on the analysis and interpretation for decision making rather than pre-processing and generating graphs. The tool starts with the user authentication as presented in Figure E.2.

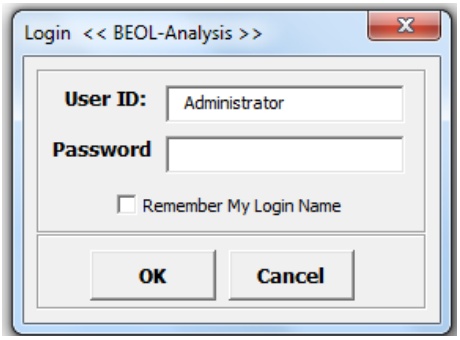


Figure E.2 – Users Login and Authentication Step

The first step is to browse the capacitance and resistance scribe and field files. The users can validate the selected files against the format, syntax and completeness by clicking on the button << Verify Data >>. The basic input file characteristics are computed and displayed to user for reference (Figure E.3).

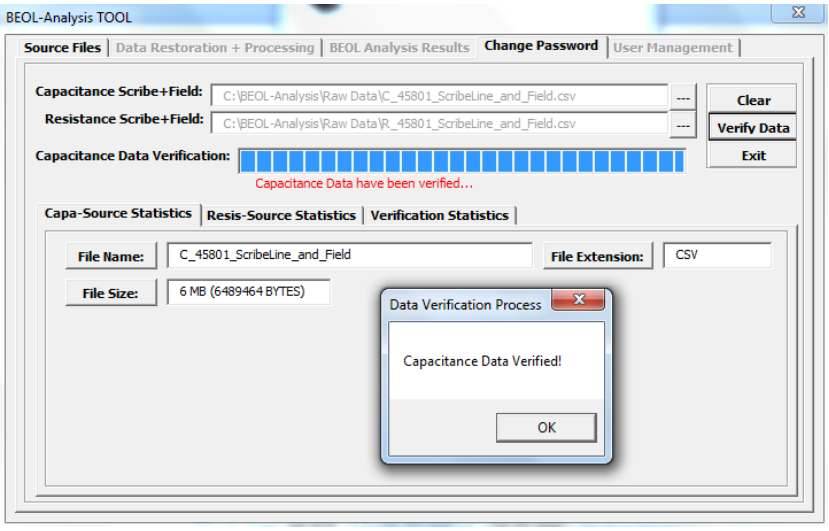


Figure E.3 – Input File Validation and Characteristics

Upon files validation, the users move to the next step where they first restore the data by clicking on the button << Restore Data >>. It reads data from validated tab delimited input file formats and then restore it to the multidimensional data model implemented using MS Access 2003. It is performed prior to data processing to avoid lengthy computation times. The restoration progress bar highlights the completed task along with messages upon the completion of the task.

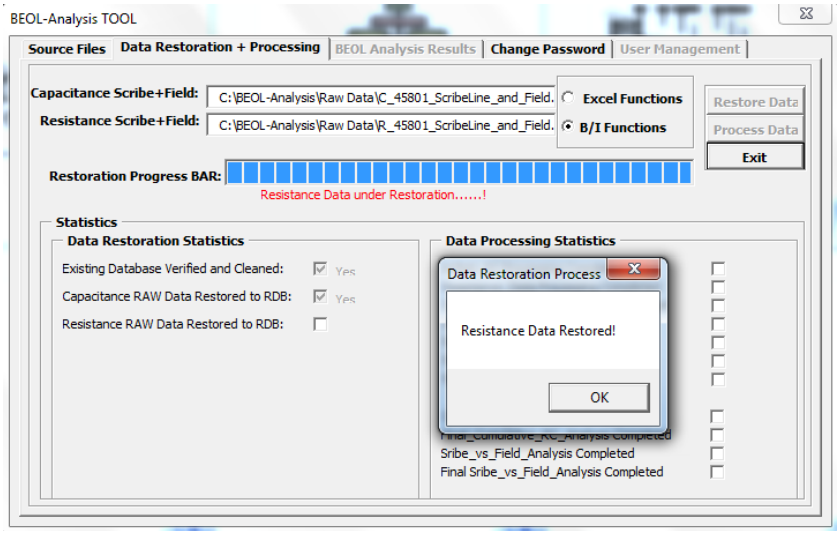


Figure E.4 – Input data Restoration to Multidimensional Data Model

At this level, users have two options to compute different statistics as *(i)* use excel or *(ii)* built-in functions. This option is provided to demonstrate that algorithms used for computation of statistics do result in slightly different statistics which might be misleading. The B/I functions are programmed in this tool based on the standard algorithms. The successfully computed results are highlighted as done under the data processing statistics block and user moves to BEOL analysis results tab (Figure E.5).

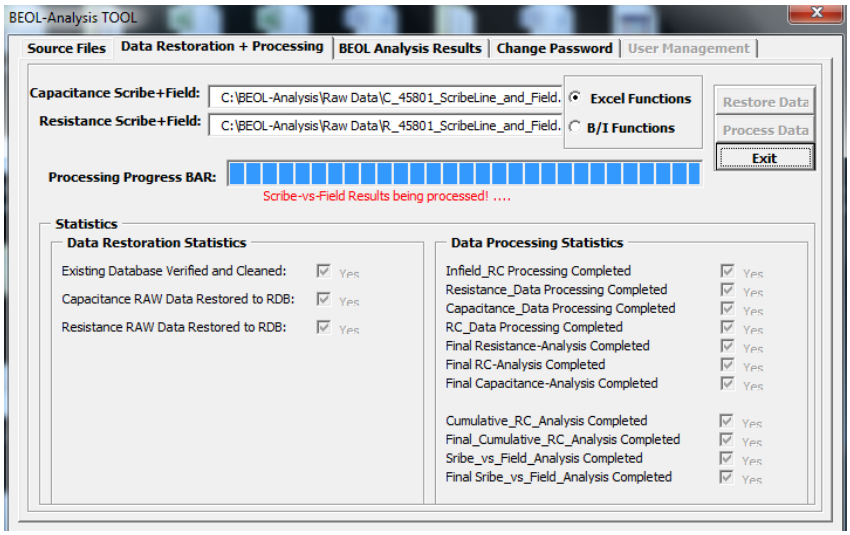


Figure E.5 – Computation of Statistics with Excel and Built-in Functions

The results are saved at users selected paths in an excel sheet as presented in Figure E.6 (a). The parasitic element based analysis using excel and built-in functions are presented in the Figure E.6 (b). It can be seen that the role of algorithms in statistics computation is highly critical. The results provide a comparison of no filter, LSL-USL filter and LVL-UVL filters and it can be seen that when filters are applied and number of counts are less the difference between the computed statistics using excel and built-in functions is quite significant.

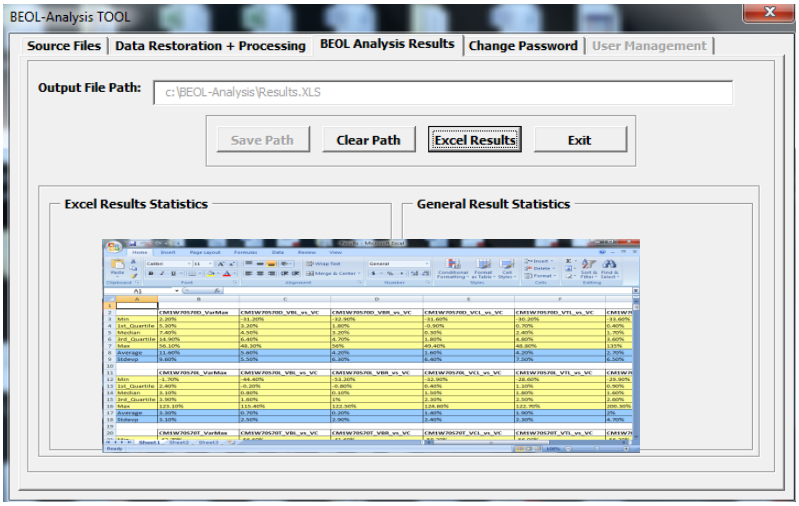


Figure E.6 (a) – Computation of Statistics with Excel and Built-in Functions

Layer	Layer_ID		No Limits Filter										LSL-USL Filter										LVL-UVL Filter									
			Programmed Functions						Built in Functions				Programed Functions						Built in Functions				Programed Functions						Built in Functions			
			Count	μ	Median	σ	Min	Max	μ	Median	σ	Min	Max	Count	μ	Median	σ	Min	Max	μ	Median	σ	Min	Max	Count	μ	Median	σ	Min	Max	μ	Median	σ	Min	Max
Capacitance	CM1W70S70D		4421	134.73	133.06	7.67	113.36	266.02	134.725	133.06	7.672	113.36	266.02	599	147.846	145.065	7.201	141.57	188.304	147.85	145.07	7.201	142	188	4421	134.725	133.06	7.672	113.4	266	135	133.06	7.672	113.4	266
	CM1W70S70L	M1	4422	167.81	167.54	6.21	77.27	185.24	167.814	167.535	6.21	77.274	185.24	4421	167.834	167.535	6.059	137.61	185.242	167.83	167.54	6.059	138	185	4422	167.814	167.54	6.21	77.27	185.2	168	167.535	6.21	77.27	185.2
	CM1W70S70U	Scribe	4417	297.66	293.35	22.24	223.24	411.05	297.663	293.347	22.235	223.24	411.05	2147	277.896	279.069	8.687	223.24	292.111	277.9	279.07	8.687	223	292	4417	297.663	293.35	22.24	223.2	411.1	298	293.347	22.24	223.2	411.1
	CM1W70S70T		4423	256.39	254.89	7.52	115.52	282.75	256.394	254.892	7.519	115.52	282.75	4202	255.601	254.453	6.107	211.8	269.481	255.6	254.45	6.107	212	269	4423	256.394	254.89	7.519	115.5	282.7	256	254.892	7.519	115.5	282.7
	CM2W70S70L	Mx Lateral	30614	152.97	152.73	7.26	51.76	243.72	152.969	152.728	7.263	51.762	243.72	29394	152.588	152.461	5.39	103.98	164.367	152.59	152.46	5.39	104	164	30614	152.969	152.73	7.263	51.76	243.7	153	152.728	7.263	51.76	243.7
	CM3W70S70L	Scribe + Field	30904	147.13	147.19	5.34	51.81	215.82	147.127	147.192	5.335	51.809	215.82	30831	147.233	147.195	4.14	103.95	164.346	147.23	147.2	4.14	104	164	30904	147.127	147.19	5.335	51.81	215.8	147	147.192	5.335	51.81	215.8
	CM4W70S70L		30946	154.32	153.98	5.27	138.82	238.86	154.323	153.982	5.271	138.82	238.86	30577	154.148	153.897	4.991	138.82	166.047	154.15	153.9	4.991	139	166	30946	154.323	153.98	5.271	138.8	238.9	154	153.982	5.271	138.8	238.9
	CM5W70S70L		30950	167.54	168.07	5.99	122.04	258.16	167.54	168.072	5.993	122.04	258.16	29615	166.996	167.743	5.41	122.04	176.789	167	167.74	5.41	122	177	30950	167.54	168.07	5.993	122	258.2	168	168.072	5.993	122	258.2
	CM1W70S70U		4417	297.66	293.35	22.24	223.24	411.05	297.663	293.347	22.235	223.24	411.05	2147	277.896	279.069	8.687	223.24	292.111	277.9	279.07	8.687	223	292	4417	297.663	293.35	22.24	223.2	411.1	298	293.347	22.24	223.2	411.1
	CM2W70S70U	Mx Vertical	4412	300.15	296.87	27.32	147.36	369.34	300.148	296.87	27.324	147.36	369.34	1688	272.104	273.014	11.21	147.36	288.52	272.1	273.01	11.21	147	289	4412	300.148	296.87	27.32	147.4	369.3	300	296.87	27.32	147.4	369.3
	CM3W70S70U	Scribe	4423	304.09	302.21	18.57	186.06	377.45	304.091	302.213	18.571	186.06	377.45	1016	280.983	282.295	7.449	186.06	288.934	280.98	282.3	7.449	186	289	4423	304.091	302.21	18.57	186.1	377.5	304	302.213	18.57	186.1	377.5
	CM4W70S70U		4422	282.82	281.83	19.98	242.55	354.04	282.819	281.826	19.983	242.55	354.04	1941	264.154	264.574	8.216	242.55	277.6	264.15	264.57	8.216	243	278	4422	282.819	281.83	19.98	242.5	354	283	281.826	19.98	242.5	354
	CM2W70S70T		26388	233.46	233.45	14.11	77.88	363.23	233.456	233.452	14.105	77.881	363.23	21719	231.613	231.731	6.209	154.84	242.637	231.61	231.73	6.209	155	243	26388	233.456	233.45	14.11	77.88	363.2	233	233.452	14.11	77.88	363.2
	CM3W70S70T	Mx Total	26529	220.13	220.75	9.00	80.50	297.10	220.127	220.75	8.998	80.503	297.1	26426	220.402	220.768	6.741	155.15	242.176	220.4	220.77	6.741	155	242	26529	220.127	220.75	8.998	80.5	297.1	220	220.75	8.998	80.5	297.1
	CM4W70S70T	Field	26520	244.75	244.24	8.05	91.57	353.41	244.748	244.238	8.053	91.571	353.41	9264	236.628	237.042	2.925	222.68	240.968	236.63	237.04	2.925	223	241	26520	244.748	244.24	8.053	91.57	353.4	245	244.238	8.053	91.57	353.4
	CM5W70S70T		26530	212.75	213.01	7.12	87.29	316.53	212.752	213.01	7.121	87.288	316.53	25867	212.297	212.82	6.267	159.68	225.931	212.3	212.82	6.267	160	226	26530	212.752	213.01	7.121	87.29	316.5	213	213.01	7.121	87.29	316.5
	CM5W70S70U		4409	68.80	68.80	0.99	63.76	80.22	68.8	68.797	0.994	63.758	80.215	1	68.8	68.797	0.994	63.758	80.215	68.8	68.797	0.994	63.8	80.2	4409	68.8	68.797	0.994	63.76	80.22	68.8	68.797	0.994	63.76	80.22
	CM6W400S400L	M6	4423	231.42	232.89	7.89	192.86	258.91	231.415	232.885	7.893	192.86	258.91	4423	231.415	232.885	7.893	192.86	258.907	231.42	232.89	7.893	193	259	4423	231.415	232.89	7.893	192.9	258.9	231	232.885	7.893	192.9	258.9
	CM6W400S400U	Scribe	4412	69.32	69.38	1.06	65.85	88.70	69.323	69.376	1.063	65.85	88.701	1	69.323	69.376	1.063	65.85	88.701	69.323	69.376	1.063	65.9	88.7	4412	69.323	69.376	1.063	65.85	88.7	69.3	69.376	1.063	65.85	88.7
	CM6W400S400T		4423	358.04	359.25	9.00	313.08	376.98	358.037	359.251	9.003	313.08	376.98	1821	349.584	351.58	6.961	313.08	357.466	349.58	351.58	6.961	313	357	4423	358.037	359.25	9.003	313.1	377	358	359.251	9.003	313.1	377
	CM6W400S400U	M7	4412	69.32	69.38	1.06	65.85	88.70	69.323	69.376	1.063	65.85	88.701	1	69.323	69.376	1.063	65.85	88.701	69.323	69.376	1.063	65.9	88.7	4412	69.323	69.376	1.063	65.85	88.7	69.3	69.376	1.063	65.85	88.7
	CM7W400S400L	Scribe	4423	265.99	268.12	8.36	225.72	282.15	265.985	268.121	8.364	225.72	282.15	4423	265.985	268.121	8.364	225.72	282.154	265.99	268.12	8.364	226	282	4423	265.985	268.12	8.364	225.7	282.2	266	268.121	8.364	225.7	282.2
	CM7W400S400T		4423	322.91	325.09	8.64	282.10	340.67	322.906	325.09	8.64	282.1	340.67	4252	322.406	324.707	8.433	282.1	333.857	322.41	324.71	8.433	282	334	4423	322.906	325.09	8.64	282.1	340.7	323	325.09	8.64	282.1	340.7
Resistance	RLM1W70S70	RLM	4422	3695.32	3684.76	176.81	3042.82	4200.97	3695.32	3684.76	176.81	3042.8	4201	4422	3695.32	3684.76	176.8	3042.8	4200.97	3695.3	3684.8	176.8	3043	4201	4422	3695.32	3684.8	176.8	3043	4201	3695	3684.76	176.8	3043	4201
	RLM2W70S70		4423	3288.51	3297.04	174.75	2762.50	3769.43	3288.51	3297.04	174.75	2762.5	3769.4	4412	3289.63	3297.37	173.5	2866.9	3769.43	3289.6	3297.4	173.5	2867	3769	4423	3288.51	3297	174.7	2763	3769	3289	3297.04	174.7	2763	3769
	RLM3W70S70		4422	3311.67	3302.25	127.64	2830.50	3758.42	3311.67	3302.25	127.64	2830.5	3758.4	4421	3311.78	3302.25	127.5	2885.2	3758.42	3311.8	3302.2	127.5	2885	3758	4422	3311.67	3302.2	127.6	2830	3758	3312	3302.25	127.6	2830	3758
	RLM4W70S70		4423	3102.30	3073.28	204.57	2599.42	3633.55	3102.3	3073.28	204.57	2599.4	3633.5	3896	3144.49	3110.99	179.4	2866.9	3633.55	3144.5	3111	179.4	2867	3634	4423	3102.3	3073.3	204.6	2599	3634	3102	3073.28	204.6	2599	3634
	RLM5W70S70		4423	3337.23	3313.04	183.67	2871.54	3940.76	3337.23	3313.04	183.67	2871.5	3940.8	4423	3337.23	3313.04	183.7	2871.5	3940.76	3337.2	3313	183.7	2872	3941	4423	3337.23	3313	183.7	2872	3941	3337	3313.04	183.7	2872	3941
	RLM6W400S400		4423	52.30	51.89	1.79	49.22	64.01	52.297	51.885	1.788	49.22	64.011	4423	52.297	51.885	1.788	49.22	64.011	52.297	51.885	1.788	49.2	64	4423	52.297	51.885	1.788	49.22	64.01	52.3	51.885	1.788	49.22	64.01
	RLM7W400S400		4423	52.33	51.85	1.95	49.07	63.65	52.333	51.845	1.945	49.065	63.653	4423	52.333	51.845	1.945	49.065	63.653	52.333	51.845	1.945	49.1	63.7	4423	52.333	51.845	1.945	49.07	63.65	52.3	51.845	1.945	49.07	63.65

Figure E.6 (b) – Parasitic Elements (Resistance/Capacitance) Analysis

The Figures E.6 (c & d) presents the graphs generated by tool as the box whisker plots for scribe line or field metal resistances and capacitances. As it is said earlier that significant variation or drifts are further investigated for their root causes. The first and prime suspect of these drifts is the variation in geometric shapes of transistors; hence we need to further correlate the PT drifts against these geometric variations. This PT-Inline correlation needs site level information that shall act as a pivot for multi-source data analysis for root cause analysis. It is presented in appendix F, however proposed solution (KLA-ACE recipe) takes into account the site level information for both PT and Inline data and also provides the end users to create new variable to be used during analyses.

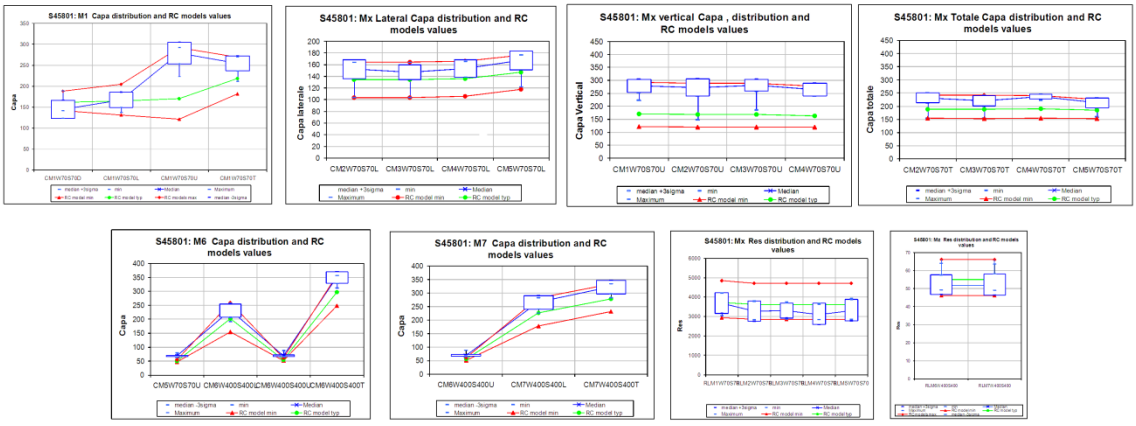


Figure E.6 (c) – Parasitic Elements (Resistance/Capacitance) Analysis

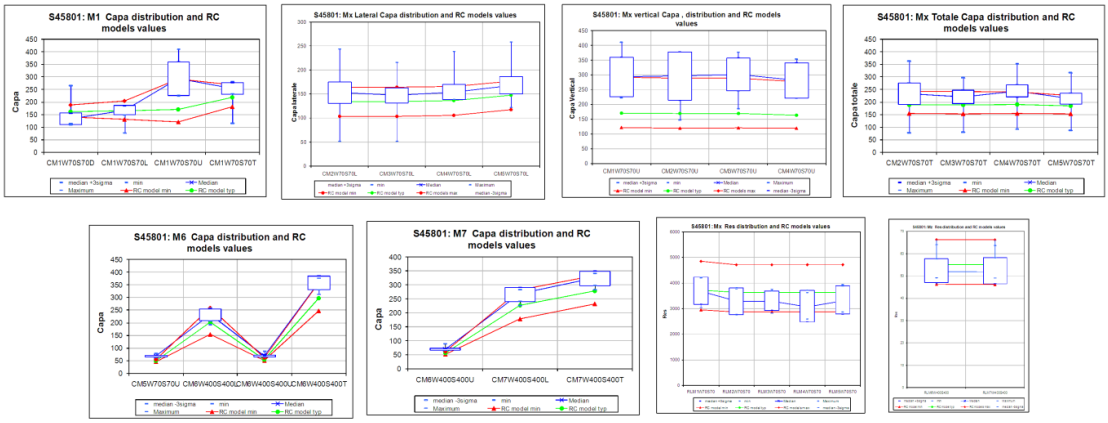


Figure E.6 (d) – Parasitic Elements (Resistance/Capacitance) Analysis

Appendix F: KLA-Ace Recipe for PT-Inline Correlation

The objective of PT-Inline correlation is to find root causes against PT drifts and variations. This section is the continuation of the appendix-E where we are focused on finding PT variance analysis using our proposed BEOL-Variance analysis tool followed by root cause analysis against drifts. We have developed KLA-ACE Recipe that follows the following activity diagram for the mapping and alignment of multi-source data at site levels. The KLA-ACE is a well-known multi-source data analysis tool used in the semiconductor industry. It is a commercial product by KLA Tencor which is based on workflow engine concept. The end users create recipes using different data extraction, pre-processing and alignment nodes. These recipes are created offline but when executed in one go or steps results in designed analysis.

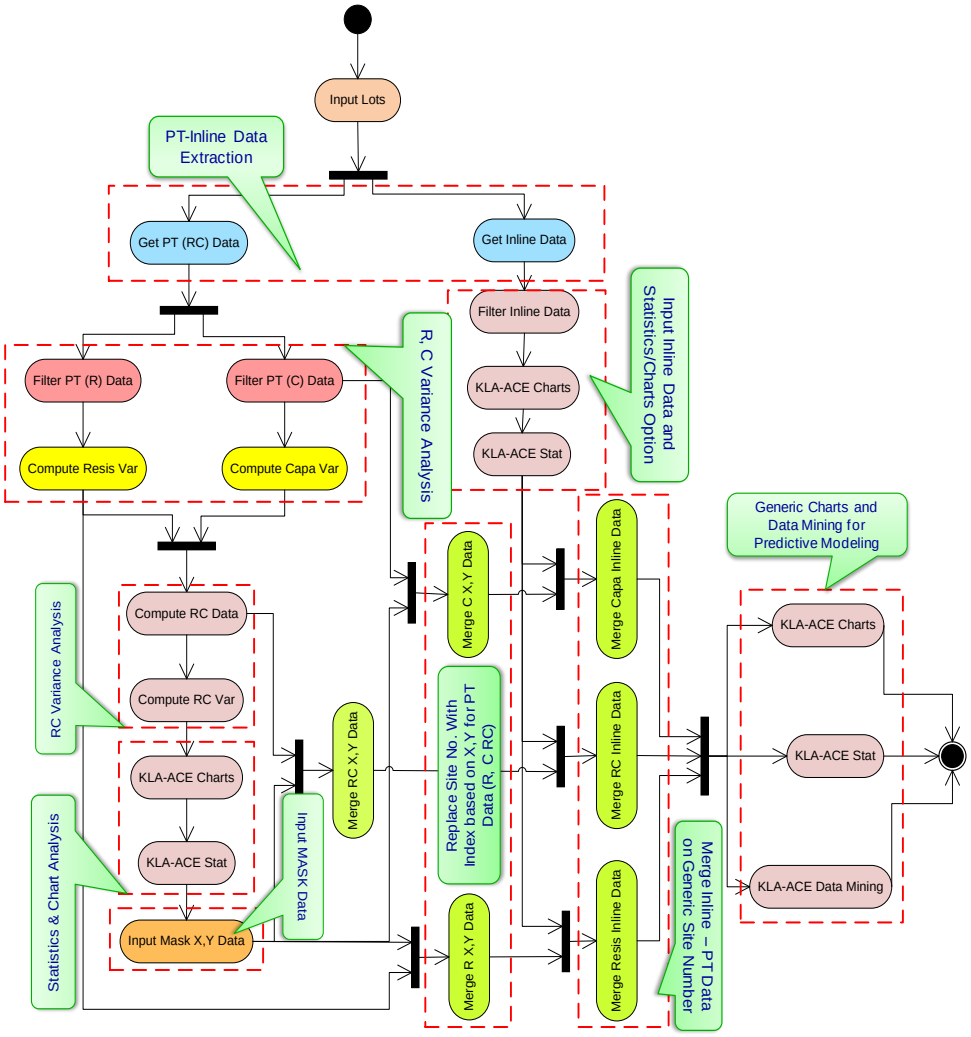


Figure F.1 – PT-Inline Mapping Flow Chart

The workflow for the recipe as presented above is self explanatory; however for the convenience to read and understand, the logical functions are group together. The flow runs parallel for PT-Inline data extraction (for pt-inline data extraction from databases) until we use Mask data to merge these two. Within PT flow we perform R, C, RC variance analysis followed by computation of different statistics and graph/chart analysis options. In parallel action user is provided variance analysis and graph/chart options for the inline data. The next step is to apply site normalization and use this to merge pt-inline site-to-site mapped data. Users are provided with an option for the basic predictive modelling using KLA-Ace pre-defined and built-in nodes.

The comparison of the proposed KLA-ACE recipe and BEOL-tool is very important at this stage. The KLA-ACE database is limited by the amount of data retained. It offers only three months of data

whereas we can use the BEOL tool on any period of data. The BEOL tool helps us quickly finding out significant drifts as we do not always need KLA-ACE recipe to perform correlations. The KLA-ACE recipe is very important when we need to further investigate excursions or drifts as it requires site level mapping of multi-source data. At present the PT data extracted from databases have x, y coordinates along with site indices whereas the inline data accompanies the site indices. The x, y coordinates along with notch position are missing at this level; hence site level mapping for correlation is not possible as shown in the Figure F.2.

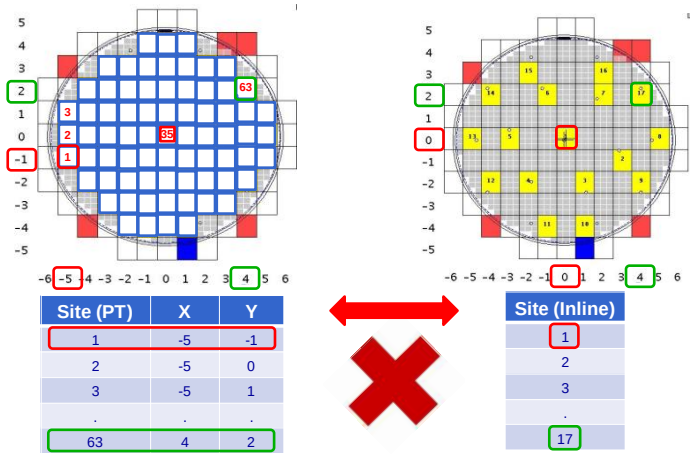


Figure F.2 – Site Level PT-Inline Mapping

We use mask level information to find site level coordinates of the sites tested for geometric measurements (inline). These x,y coordinates are then helpful in performing site level mapping for correlation purposes as show in the Figure F.3. The proposed method is almost similar to the hard coded methods for the mapping and alignment. The need for generic model for the site and die level mapping and die to site qualification was felt very badly when we observed varying measurement reference frames and notch positions for different metrology tests which often result in varying site and die level x,y coordinates. We then propose MAM and SPM models as presented in chapter-5 for generic site/die level mapping and alignment, and die to site qualification for accurate analyses. The said recipe is still valid and being used for PT-Inline site level correlation analyses.

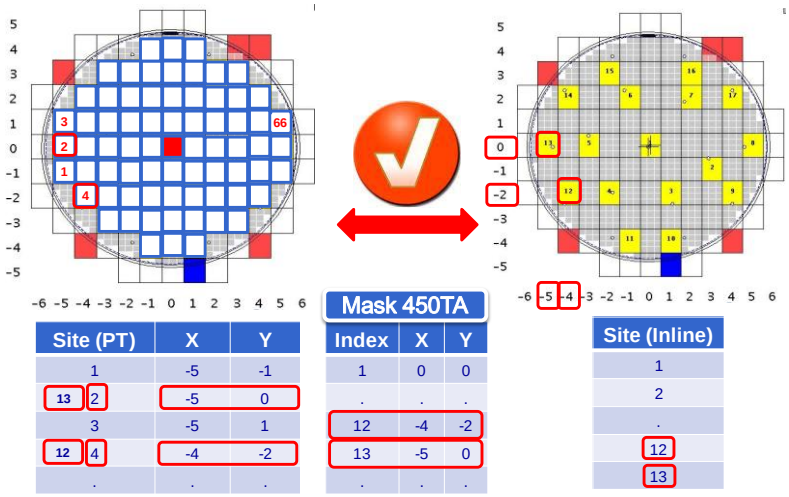


Figure F.3 – Site Level PT-Inline Mapping

The said core flow as presented in Figure F.1 also provides the end users with the ability to create new variables from existing variables. It is very important function because the width and height against a geometric shape are computed from multiple inline parameters. It facilitates users with variety of correlations.

Appendix G: EPP (Equipment, Product, Process) Life Cycle Tool

The EPP tool is primarily developed for the IMPROVE (Implementing Manufacturing Solutions to Increase Equipment Productivity and FAB Performance) project. The main objective is to provide contextual data to support effective root cause analyses efforts. It is a 2-tier and multi-user database application developed using VB6.0 professional and SQL Server 2008. This EPP is developed with a multi-dimensional database, however in this appendix the data model and detailed UML model are not presented due to confidentiality and space limitations. The proposed application is presented using the GUIs.

The EPP starts with users' authentication with id and password. This application during its startup identifies the best resolution and adjusts all objects on the GUIs accordingly (Figure G.1). The users start by selecting the duration of the data extraction, however default period is of recent one week. It is important to note that output of this tool is classified as the product life cycle or equipment life cycle.

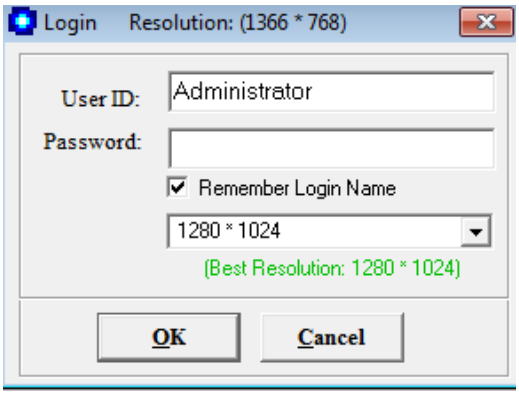


Figure G.1 – User Authentication in EPP Tool

The users do have an option to select type of workshop, equipment or module(s). The modules are considered as the sub-equipments which are associated with the parent equipment in parent-child relationship. This tool is directly connected with the maintenance data warehouse sp we start by clicking the << Import >> button. It fetches data from data warehouse into local multidimensional database (Figure G.2). The data comprises of "Equipment States", "Equipment Failure to Work Request/Order" and "Equipment Checklist Steps". The << Process >> button aligns data and generate equipment life cycle for the given period which is presented in the Tab << Equipment Life Cycle Data >>. The users have an option to scroll and go through other data sets generated in three tabs. The option "Arrange Columns" automatically adjusts the length and width of the columns for user readability.

The internal algorithm which aligns and generates the equipment life cycle is not presented here because of the confidentiality reasons as it involves the internal data warehouse schema. The equipment life cycle data is objectively focused on tracing the equipment states, respective failure modes which lead to the work request/order and information about the check list tasks performed as corrective or preventive maintenance against these work orders. This data provides us a possible connection with OOC (out of control) production lots to assess whether equipment health is related to the yield loss or not. It also helps us in building models for predictive maintenance to avoid the expensive corrective maintenances.

The << WIP (Lots) >> button generates the list of the production lots which were being processed when the equipment was undergoing different states, failures and actions. This list of production lots helps us to extract product life cycle from the process data warehouse. Three KLA-ACE recipes are developed to extract WIP, Metrology and Out of control data from the process, metrology and OOC databases. The data extracted from these databases is linked to generate the product life cycle which is highly useful for modeling different aspects. We start by running the KLA-ACE recipe to extract WIP and Metrology information of the production lots; however OOC information is selected based on the duration directly from

the OOC database. The extracted WIP and Metrology data files (csv extracts) are browsed by the users as presented in the GUI at Figure G.3. The users select all the files by clicking on button << Select All >> or by clicking on individual files in the file list. The << Import >> button extracts the information into multidimensional database to improve processing times. The << Process >> button extracts OOC data and generates the product life cycle information which is presented to the user in the "Product Life Cycle Data" tab, however the source data can be found in other three tabs. The "Autosize" option adjust the column widths and heights for readability and this data can be exported in the csv format.

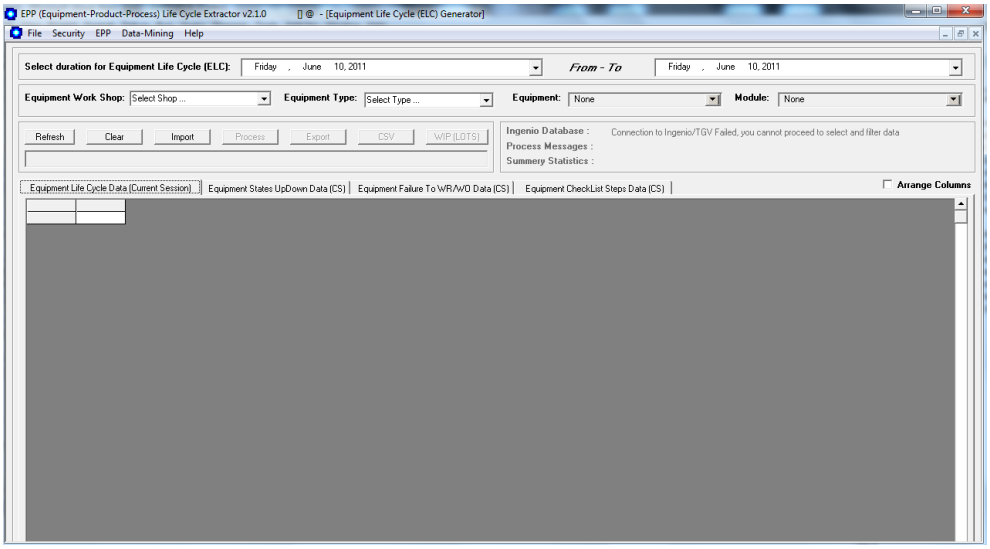


Figure G.2 – Equipment Life Cycle Data

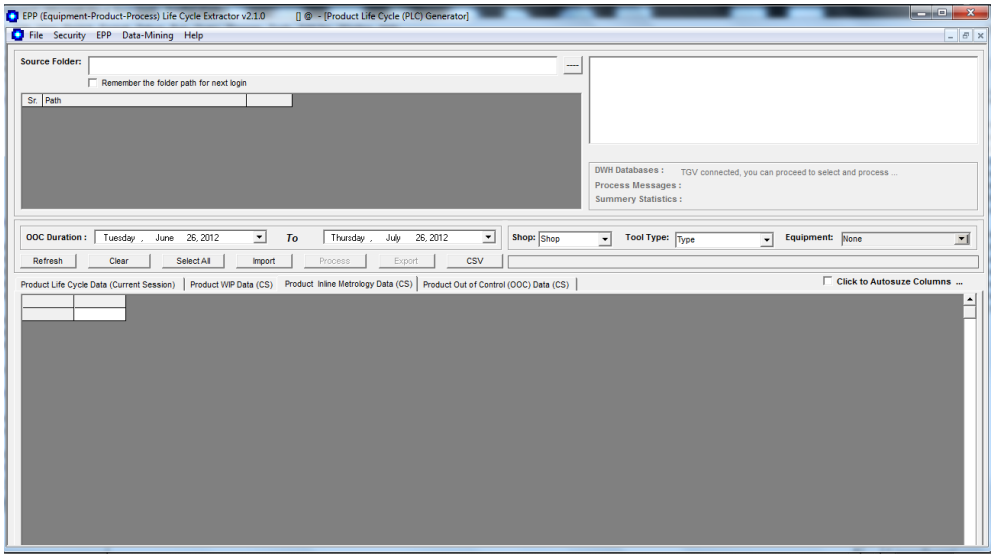


Figure G.3 – Equipment Life Cycle Data

We have presented the basic functionality of the tool regarding product and equipment life cycle generation. It is contextual data which helps a lot in establishing the root cause against drifts and variations. The other functionalities like data mining, security and help are not presented here because of the space limitations. The said tool serves not only the IMPROVE project but also all the engineering and R&D departments.

Appendix H: ACM (Alarm Control Management) Tool

The proposed ACM tool is subjected to the equipment engineering teams with an objective to help them in equipment alarms management to reduce unnecessary equipment stops. The goals of the proposed tool are to *(i)* improve equipment engineers' efficiency and *(ii)* empower equipment engineers to learn changing equipment behavior. The objectives targeted with this tool are as under:

a) Analyze, Manage and Control Alarms and Warning coming from TOOL
b) Learn Alarm Patterns linked with Product, Process or Equipment
c) Predict Cost and Yield related risk against significant Alarm Patterns

The ACM tool links the proprietary databases of the individual equipments with the other data sources as presented in Figure H.1 below. The automation system collects the alarms generated by equipment and stores them into alarms database. The ACM tools provides slicing, dicing and drill down/up operations on the data extracted from these proprietary databases. The data collected from the production line, stored in multiple data sources, is linked with the alarms generated by the equipments to build the PAM and PSM models as presented in chapter-7. The objective is to learn the alarm patterns linked with bad and good yields for subsequent use in the prediction to spare the capacities.

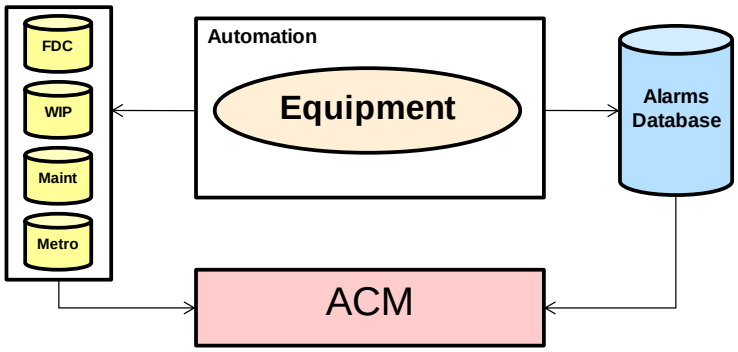


Figure H.1 – Scope of ACM Tool

The ACM tool is developed using VB 6.0 professional and MS SQL server and it is a 2-tier multi-user database application. The database used at backend is a dimensional data warehouse and OLAP queries are implemented as simple SQL queries. The query responses are optimized through views and indexing @ the cost of space. The key features of the tools are presented as under:

a) Input Alarm files, pre-process and filter Alarms in the database
b) Provide user Global Pareto and Best/Worst equipment statistics based on weekly and overall basis
c) Alarm pre-processing prior to analysis on historical data
d) Generate Pareto on Count/Duration
e) Display Missing geometries within each simulation
f) Export the SPICE parameters, PT specs and model output to excel
g) Export selected SPICE/Model output parameter to excel

The tool is not presented as UML model due to space limitation; however a brief description is presented through key and important GUIs. The users have an option to connect directly to the equipment alarms proprietary database or input the alarms data from the alarms data extracted through data extraction utilities. The duplicate records if found are separated during data input step. The main GUI is presented at Figure H.2(a). The "Litho-Tool" name presents the type of equipment family under processing. The users browse the input files which are then selected for import. The files with background color as green are

already pre-processed files. The button << Import >> results in importing raw data along with pre-processing into multidimensional data warehouse. The alarms data in its raw format needs processing to find out its start and end times for the computation of duration. The right side statistics presents the historical evolution of alarms on the basis of error count and duration. The best and worst tools are highlighted for special consideration of equipment engineer.

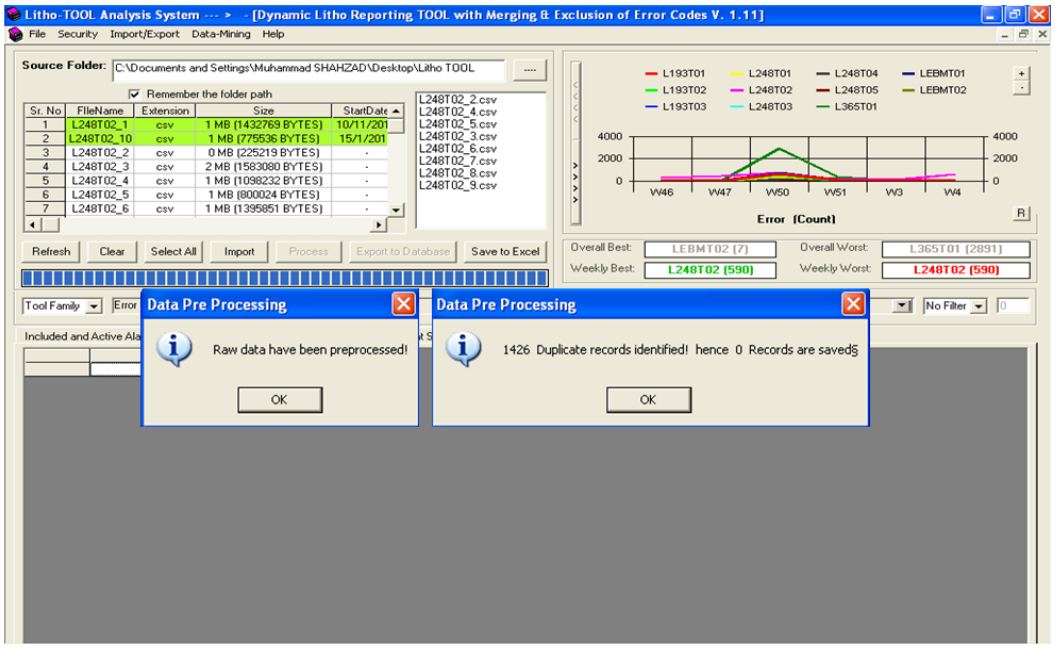


Figure H.2 (a) – Input Alarms Data for ACM Tool

The similar data input which is directly connected with the equipment alarms database is presented at Figure H.2(b). The user selects the duration, tool family, tool and module(s). The users in this interface have the option to memorize the data extraction configuration. The pre-filter option is an excellent criterion to filter out the alarms when the equipment was under the maintenance or qualification. The chart configuration can be used to analyze the historical error count or duration at multiple granularities e.g. days, hours, minutes. The chart can also be sliced to the type of alarms e.g. productive which means the alarms when the equipment was in the state of production. The algorithms implemented at the back end of the tool provide us an opportunity to slice/dice the alarms data into different charts along with the assessment of best and worst tool. The tool offers two types of key statistics about the best and worst tool as *(i)* overall statistics computed among all tools across the historical data and *(ii)* last week statistics. If the data extraction period selected by the user overlaps the existing data in the dimensional database then the duplicate records are filtered out in respective tabs. The available tabs are *(i)* extracted and filtered alarm data, *(ii)* excluded alarm data, *(iii)* deleted alarms data and *(iv)* filtered alarms data. The filters are managed separately where user groups in different alarms so that during the data extraction they are filtered out being non significant. Te users do have an option to delete different alarms or filter them as per need, however no alarm data is physically deleted but its status is changed for subsequent analyses.

It is very important to filter out and manage the alarms which we do not want to include because they are insignificant and have very high count but very small durations. Te users can select the tool family to group the excluded and included alarms. The users can use the buttons >>> and <<< to exclude and include the alarms respectively. The filtered alarms are internally stored and used during new data extractions.

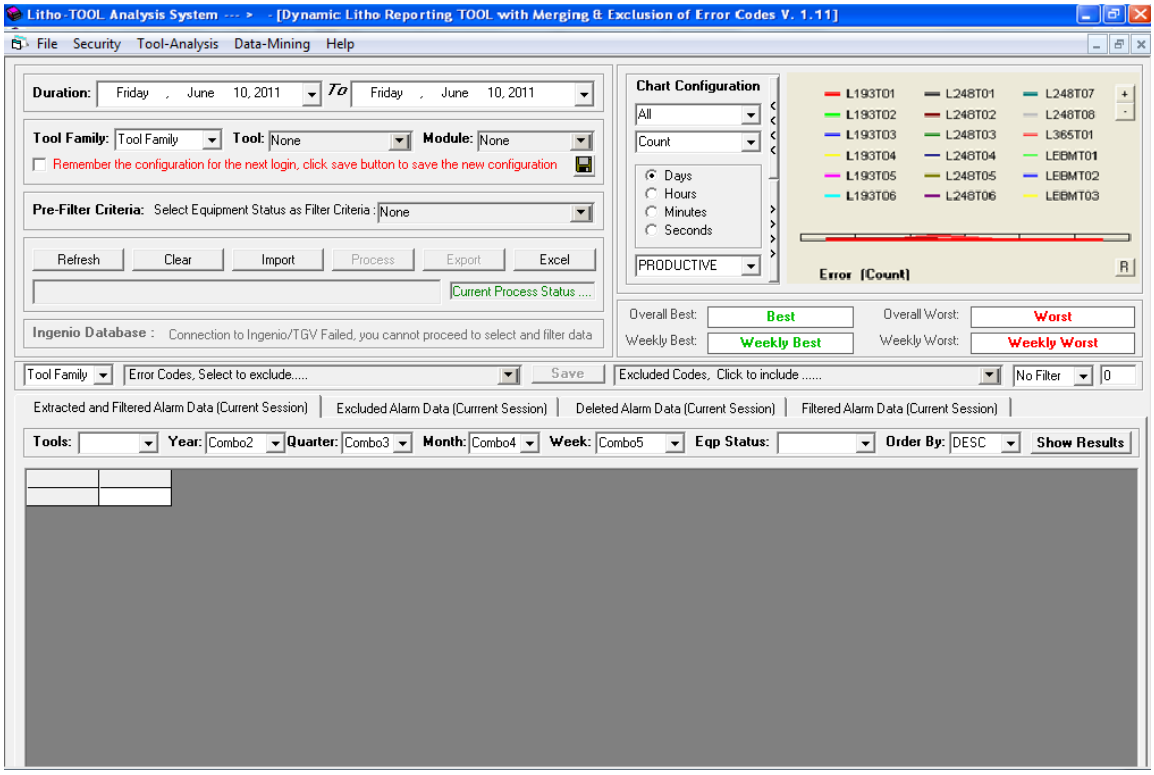


Figure H.2 (b) – Input Alarms Data for ACM Tool

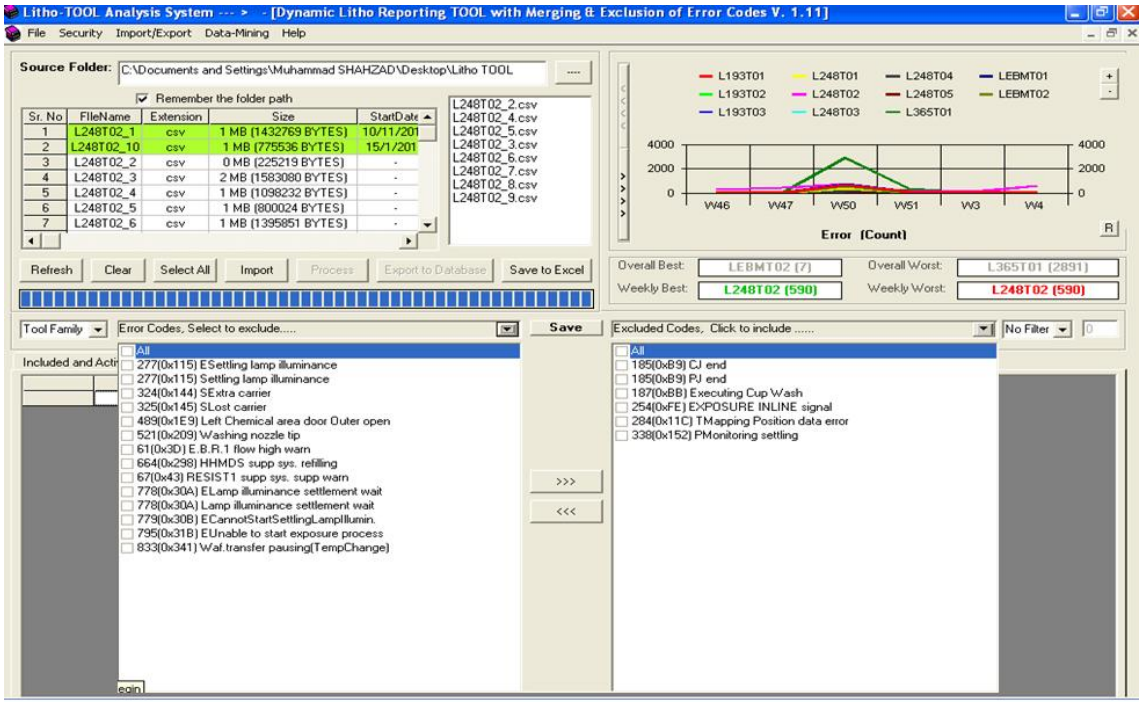


Figure H.3 – Input Alarms Data for ACM Tool

The users can right click on the graphs and can select full screen view for better visibility. The alarm information is highlighted as tool tip when user scrolls his mouse over the equipment alarms line as shown in the Figure H.4.

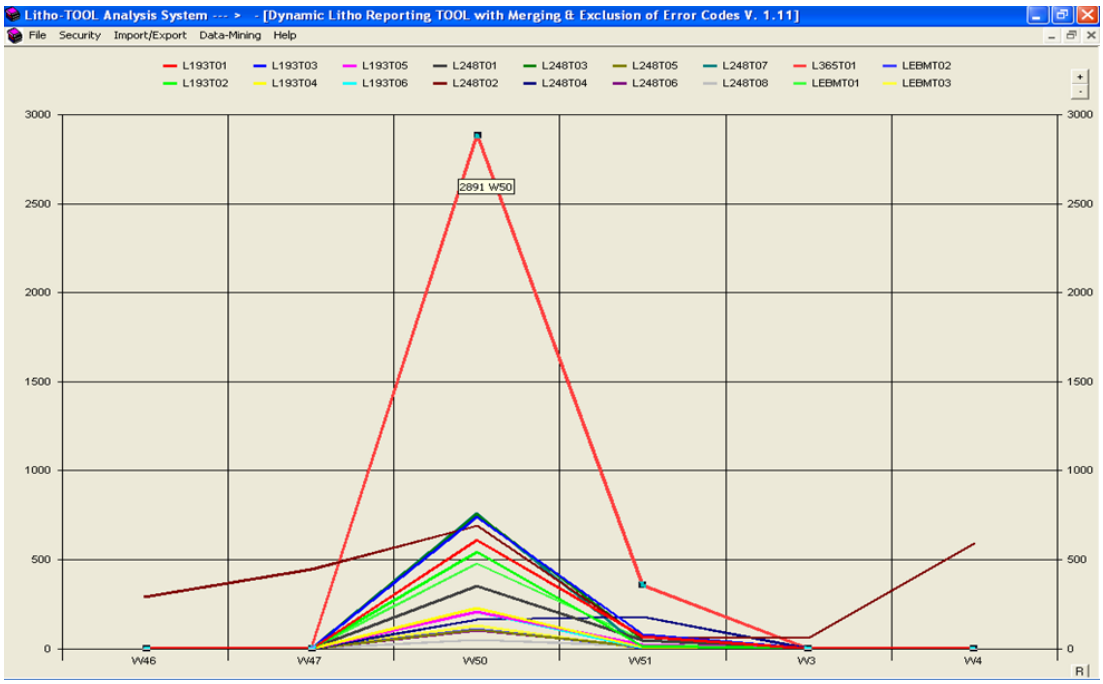


Figure H.4 – Chart Options with Tool Tip Information

An example of data extraction is presented at Figure H.5 where user is informed upon the data extraction about alarms exclusions, duplications and deletion. The users also have the option to filter and further exclude, delete and/or include the alarms using alarm levels as error, warning or others. Te option export to database is very critical as it allows the engineers to export the data as presented in the flexgrid into the database whereas the same data can be exported to an excel sheet by clicking on the button << Save to Excel >>. The data filtered across 4 tabs can be moved around by right clicking and then selecting the options as delete, exclude, include or filter.

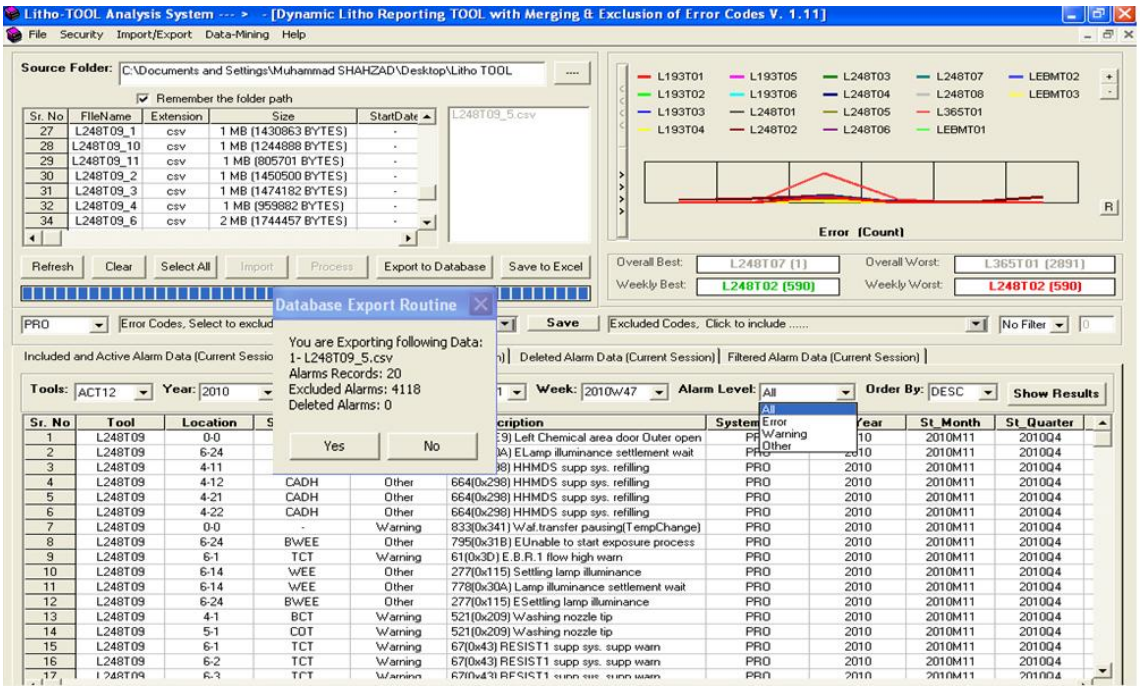
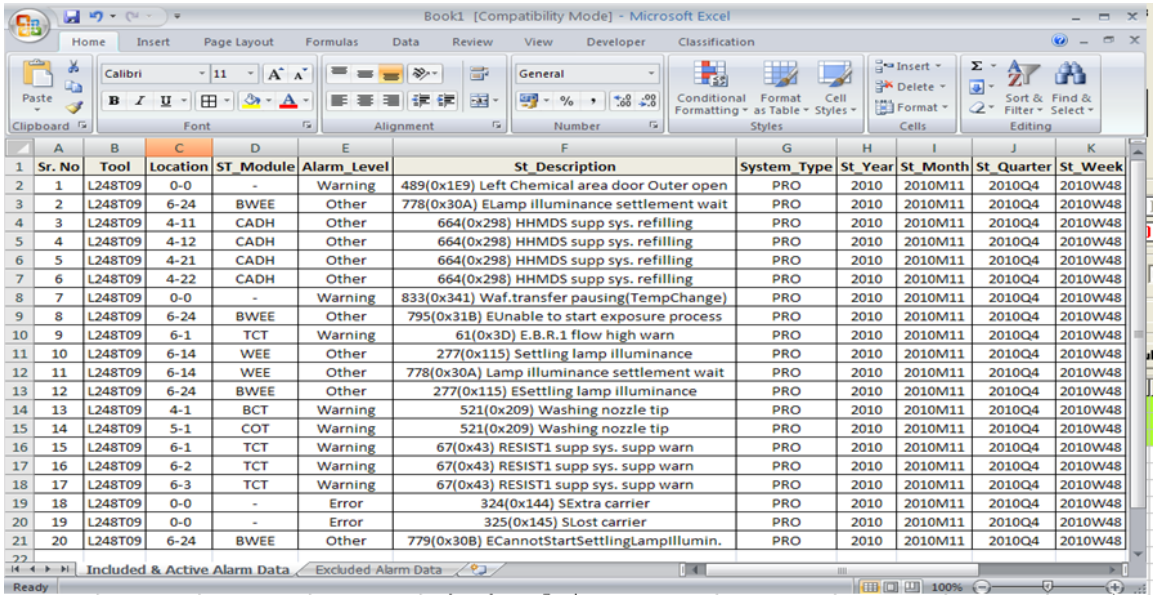


Figure H.5 – An Example of Data Extraction for ACM Tool

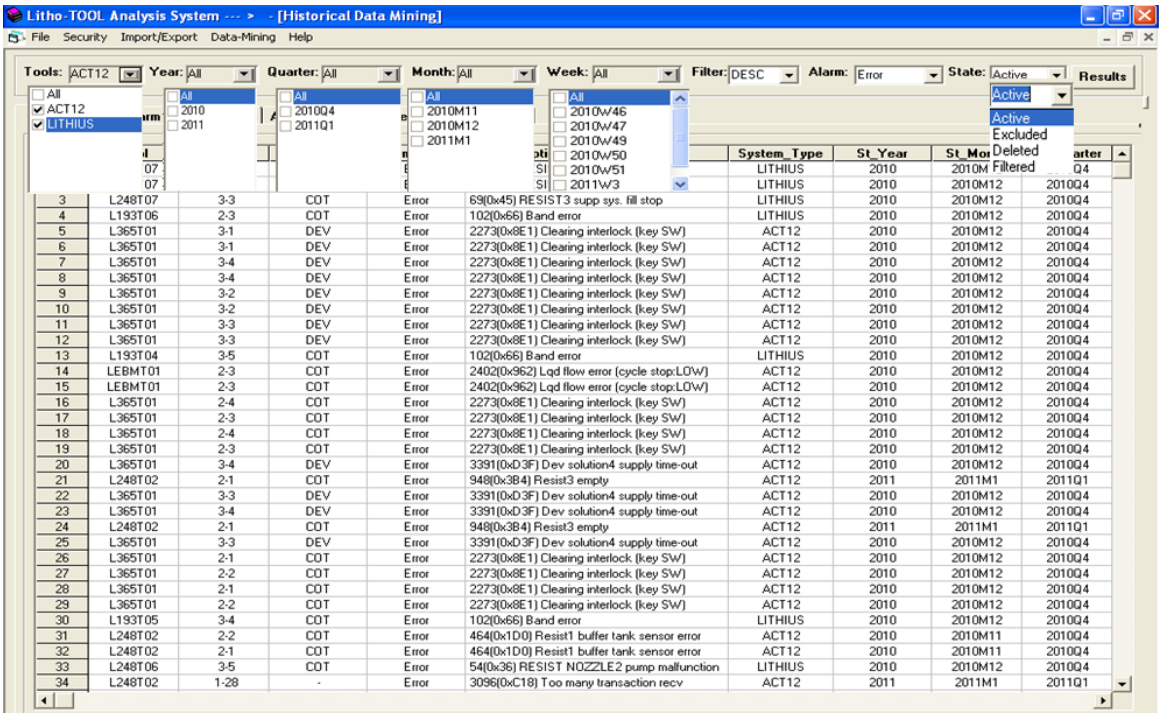
The data extracted into excel sheet from the GUI is presented at Figure H.6 below. It can be seen that the extracted data is properly formatted when extracted.



Sr. No	Tool	Location	ST_Module	Alarm_Level	St_Description	System_Type	St_Year	St_Month	St_Quarter	St_Week
1	L248T09	0-0	-	Warning	489(0x1E9) Left Chemical area door Outer open	PRO	2010	2010M11	2010Q4	2010W48
2	L248T09	6-24	BWEE	Other	778(0x30A) ELamp illuminance settlement wait	PRO	2010	2010M11	2010Q4	2010W48
3	L248T09	4-11	CADH	Other	664(0x298) HHMDS supp sys. refilling	PRO	2010	2010M11	2010Q4	2010W48
4	L248T09	4-12	CADH	Other	664(0x298) HHMDS supp sys. refilling	PRO	2010	2010M11	2010Q4	2010W48
5	L248T09	4-21	CADH	Other	664(0x298) HHMDS supp sys. refilling	PRO	2010	2010M11	2010Q4	2010W48
6	L248T09	4-22	CADH	Other	664(0x298) HHMDS supp sys. refilling	PRO	2010	2010M11	2010Q4	2010W48
7	L248T09	0-0	-	Warning	833(0x341) Waf.transfer pausing(TempChange)	PRO	2010	2010M11	2010Q4	2010W48
8	L248T09	6-24	BWEE	Other	795(0x31B) EUnable to start exposure process	PRO	2010	2010M11	2010Q4	2010W48
9	L248T09	6-1	TCT	Warning	61(0x3D) E.B.R.1 flow high warn	PRO	2010	2010M11	2010Q4	2010W48
10	L248T09	6-14	WEE	Other	277(0x115) Settling lamp illuminance	PRO	2010	2010M11	2010Q4	2010W48
11	L248T09	6-14	WEE	Other	778(0x30A) Lamp illuminance settlement wait	PRO	2010	2010M11	2010Q4	2010W48
12	L248T09	6-24	BWEE	Other	277(0x115) ESettling lamp illuminance	PRO	2010	2010M11	2010Q4	2010W48
13	L248T09	4-1	BCT	Warning	521(0x209) Washing nozzle tip	PRO	2010	2010M11	2010Q4	2010W48
14	L248T09	5-1	COT	Warning	521(0x209) Washing nozzle tip	PRO	2010	2010M11	2010Q4	2010W48
15	L248T09	6-1	TCT	Warning	67(0x43) RESIST1 supp sys. supp warn	PRO	2010	2010M11	2010Q4	2010W48
16	L248T09	6-2	TCT	Warning	67(0x43) RESIST1 supp sys. supp warn	PRO	2010	2010M11	2010Q4	2010W48
17	L248T09	6-3	TCT	Warning	67(0x43) RESIST1 supp sys. supp warn	PRO	2010	2010M11	2010Q4	2010W48
18	L248T09	0-0	-	Error	324(0x144) SExtra carrier	PRO	2010	2010M11	2010Q4	2010W48
19	L248T09	0-0	-	Error	325(0x145) SLost carrier	PRO	2010	2010M11	2010Q4	2010W48
20	L248T09	6-24	BWEE	Other	779(0x30B) ECannotStartSettlingLampIllumin.	PRO	2010	2010M11	2010Q4	2010W48

Figure H.6 – An Example of Data Extraction in Excel Sheet

The data extracted can be further used during data mining and similar operations using data mining options as presented below in Figure H.7. The data can be filtered on multiple time dimensions and status of the alarms.



3	L248T07	3-3	COT	Error	69(0x45) RESIST3 supp sys. fill stop	LITHIUS	2010	2010M12	2010Q4
4	L193T06	2-3	COT	Error	102(0x66) Band error	LITHIUS	2010	2010M12	2010Q4
5	L365T01	3-1	DEV	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
6	L365T01	3-1	DEV	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
7	L365T01	3-4	DEV	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
8	L365T01	3-4	DEV	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
9	L365T01	3-2	DEV	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
10	L365T01	3-2	DEV	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
11	L365T01	3-3	DEV	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
12	L365T01	3-3	DEV	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
13	L193T04	3-5	COT	Error	102(0x66) Band error	LITHIUS	2010	2010M12	2010Q4
14	LEBMT01	2-3	COT	Error	2402(0x962) Lqd flow error (cycle stop:LOW)	ACT12	2010	2010M12	2010Q4
15	LEBMT01	2-3	COT	Error	2402(0x962) Lqd flow error (cycle stop:LOW)	ACT12	2010	2010M12	2010Q4
16	L365T01	2-4	COT	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
17	L365T01	2-3	COT	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
18	L365T01	2-4	COT	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
19	L365T01	2-3	COT	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
20	L365T01	3-4	DEV	Error	3391(0xD3F) Dev solution4 supply time-out	ACT12	2010	2010M12	2010Q4
21	L248T02	2-1	COT	Error	948(0x3B4) Resist3 empty	ACT12	2011	2011M1	2011Q1
22	L365T01	3-3	DEV	Error	3391(0xD3F) Dev solution4 supply time-out	ACT12	2010	2010M12	2010Q4
23	L365T01	3-4	DEV	Error	3391(0xD3F) Dev solution4 supply time-out	ACT12	2010	2010M12	2010Q4
24	L248T02	2-1	COT	Error	948(0x3B4) Resist3 empty	ACT12	2011	2011M1	2011Q1
25	L365T01	3-3	DEV	Error	3391(0xD3F) Dev solution4 supply time-out	ACT12	2010	2010M12	2010Q4
26	L365T01	2-1	COT	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
27	L365T01	2-2	COT	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
28	L365T01	2-1	COT	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
29	L365T01	2-2	COT	Error	2273(0x8E1) Clearing interlock (key SW)	ACT12	2010	2010M12	2010Q4
30	L193T05	3-4	COT	Error	102(0x66) Band error	LITHIUS	2010	2010M12	2010Q4
31	L248T02	2-2	COT	Error	464(0x1D0) Resist1 buffer tank sensor error	ACT12	2010	2010M11	2010Q4
32	L248T02	2-1	COT	Error	464(0x1D0) Resist1 buffer tank sensor error	ACT12	2010	2010M11	2010Q4
33	L248T06	3-5	COT	Error	54(0x36) RESIST NOZZLE2 pump malfunction	LITHIUS	2010	2010M12	2010Q4
34	L248T02	1-28	-	Error	3096(0xC18) Too many transaction recv	ACT12	2011	2011M1	2011Q1

Figure H.7 – Data Mining Alarms Data with ACM Tool

The historical data analysis is very important for equipment engineers to monitor the equipment performance. The data extracted using GUI at Figure H.7 is further used for graphical analysis as presented at Figure H.8. The users have multiple options as the graphical analysis can be sliced at the tool(s), module(s) or alarm(s) levels. It can also be switched to error count or duration levels along with the historical trends. The individual data used along with the chart can be exported to excel as presented at Figure H.9. The slicing and dicing operations can be dropped to the level of location. It can be switched from alarms count to the alarms duration with similar options.

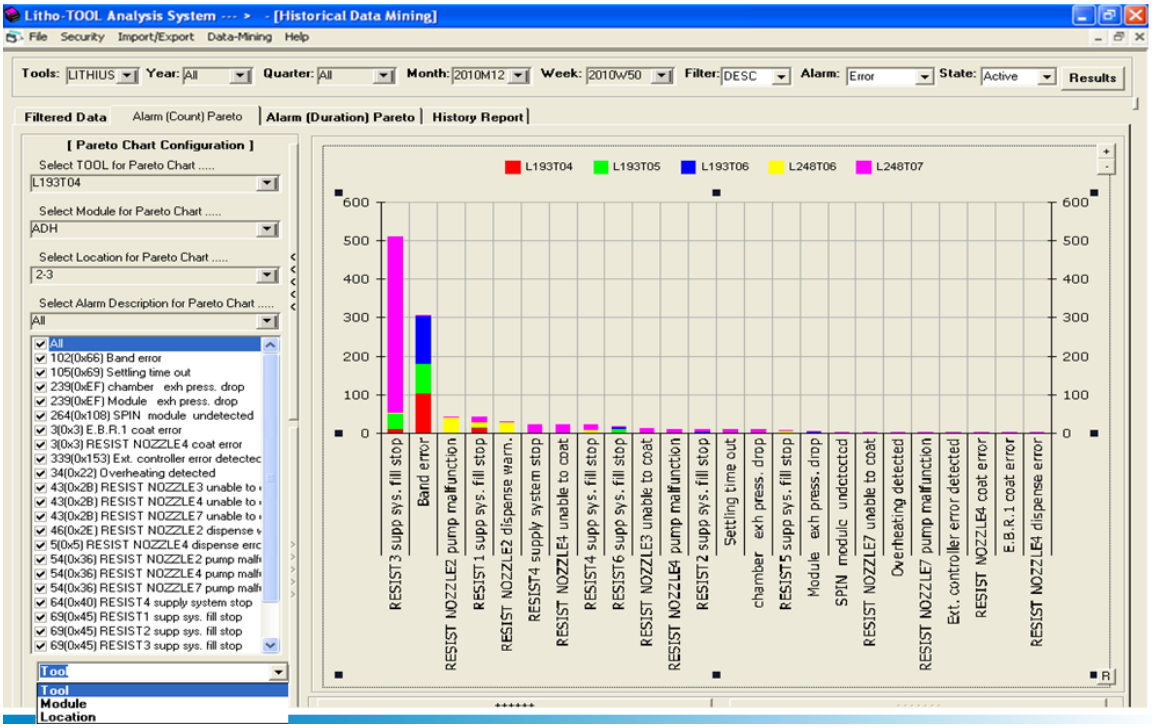


Figure H.8 – Graphical Analysis with Slicing and Dicing Options

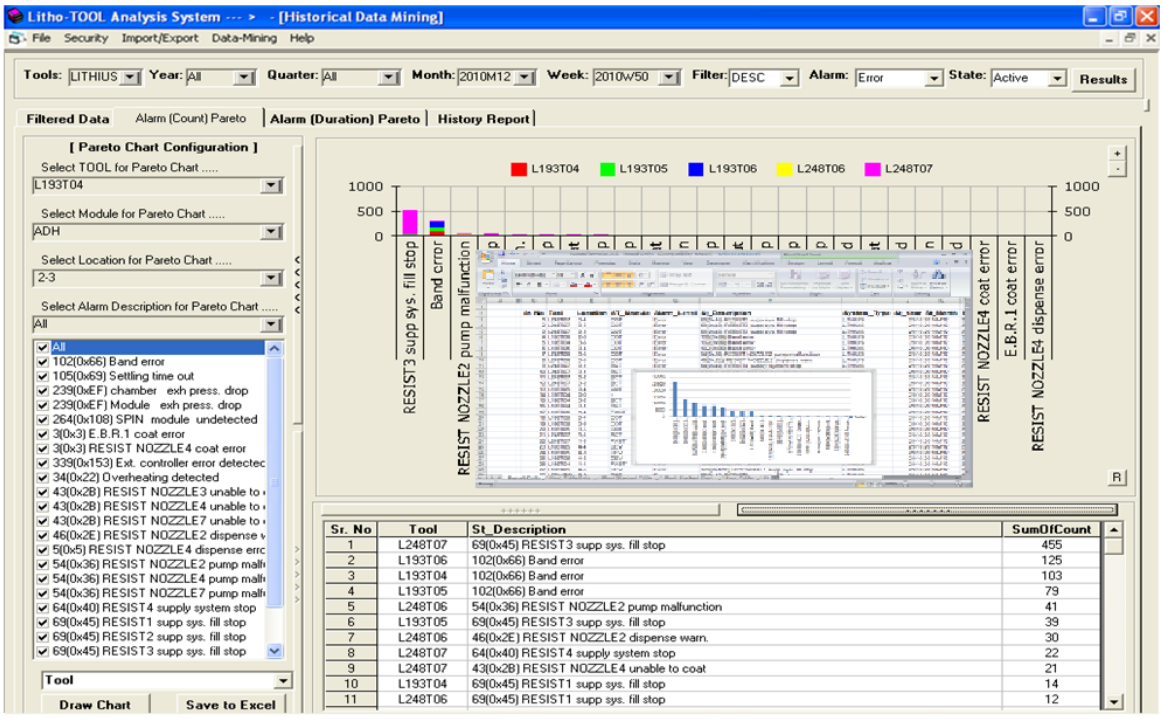


Figure H.9 – Exporting Individual Data with Charts

The historical report option provides users with the alarms and counts over last 13 weeks. If user clicks on any of the tool then sub chart is generated along with data in the bottom as presented at Figure H.10. The data with graph can be exported into excel sheet (H.11). The user, however has the slicing, dicing and drill down/up options and can analyze data in a variety of different ways for equipment management.

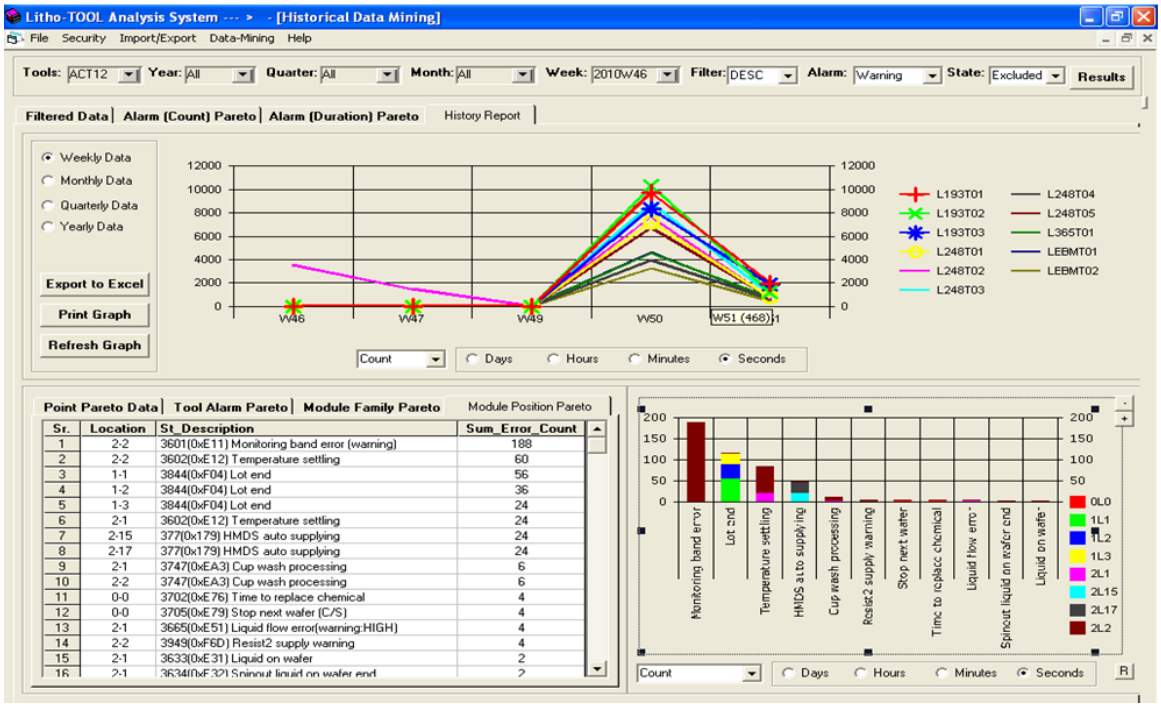


Figure H.10 – Historical Reporting for Alarms Count and Duration

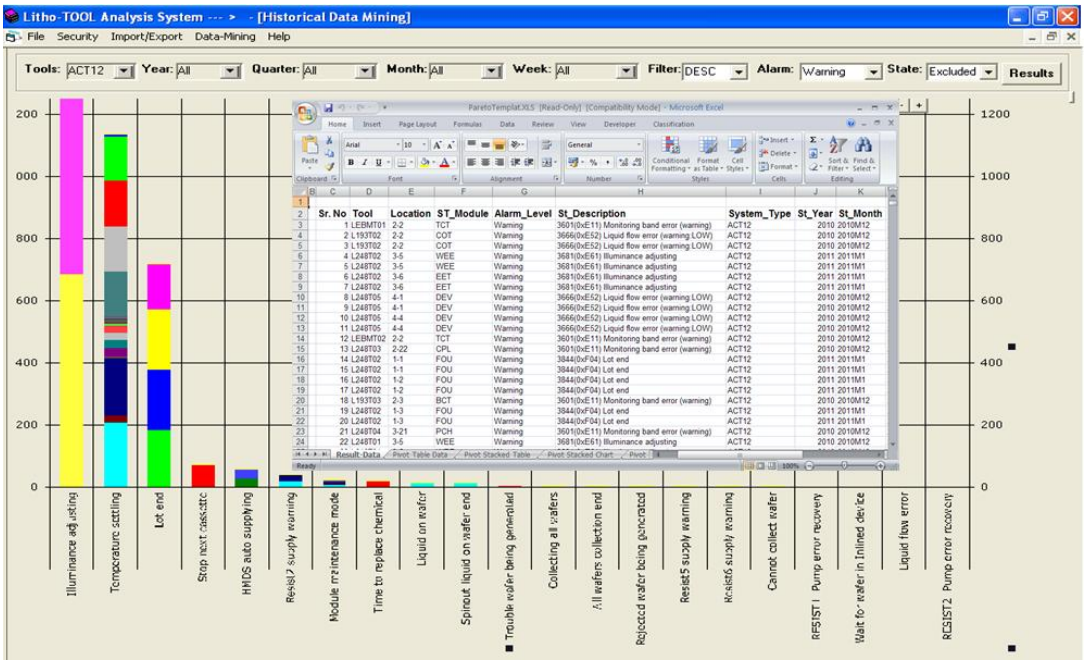


Figure H.11 – Exporting Historical Alarms Data into Excel

References

[A]

[Abdelali et al., 2003] A. Abdelali, J. Cowie, D. Farwell, B. Ogden and S. Helmreich, Cross-Language Information Retrieval using Ontology. *Proceedings of TALN Batz-sur-Mer*. France, 2003

[Abernethy and Altman, 1998] N. F. Abernethy and R. B. Altman, "SOPHIA: Providing basic knowledge services with a common DBMS*", Proceedings of the 5th KRDB Workshop*, Seattle, WA, 1998

[Ahmed and Abdalla, 2000] A.M. Ahmed and H.S. Abdalla, Beyond competition: A framework for the 21st century. *International Journal of Production Research*, 38(15):3677–3709, 2000

[Aitken et al., 2008] R. Aitken, L. Capodieci, F. Klass, C. Hou and V. Singh, *DFM in Practice: Hit or Hype? 45th Annual ACM IEEE Design Automation Conference*, 2008

[Andrews, 1980] K. R. Andrews. *The Concept of Corporate Strategy*. R. D. Irwin, 1980

[Anderson, 2006] D.M. Anderson, Design for Manufacturability & Concurrent Engineering; How to Design for Low Cost, Design in High Quality, Design for Lean Manufacture, and Design Quickly for Fast Production, 448 pages; CIM Press, 2006

[Appello et al., 2004] D. Appello, A. Fudoli and K. Giarda, Understanding Yield Losses in Logic Circuits, *IEEE Design & Test of Computers*, pp.208-215, 2004

[Aparício et al., 2005] A.S. Aparicio, O.L.M Farias, and D.N. Santos, Applying Ontologies in the Integration of Heterogeneous Relational Databases', *In Proceeding of the Australasian Ontology Workshop (AOW)*, Sydney, Australia, pp. 11-16, 2005

[Arens et al., 1996] Y. Arens, C.N. Hsu, and C.A. Knoblock, Query processing in the SIMS information mediator. *In Advanced Planning Technology*. AAAI Press, California, USA, 1996

[Astrova et al., 2007] I. Astrova, N. Korda, and A. Kalja, Storing OWL Ontologies in SQL Relational Databases, *World Academy of Science, Engineering and Technology*, 2007

[B]

[Bachman, 1973] C.W. Bachman, The programmer as navigator. *Communications of the ACM*, 16(11):653–658, 1973

[Ballhaus *et al.*, 2009] W. Ballhaus, A. Pagella, and C. Vogel, A change of pace for the semiconductor industry? *Technical report, German Technology, Media and Telecommunications*, November 2009

[Blackman, 1998] K. R. Blackman. Technical note: IMS celebrates thirty years as an IBM product. *IBM Systems Journal*, 37(4):596–603, 1998

[Blodgett and Schultz1969] J.H. Blodgett and C.K. Schultz, Data processing pioneer. *Journal of the Amerrican Society for Information Science and Technology*, 20(3):221–226, 1969

[Bolz, 1958] Bolz R.W. (Ed.); *Metals Engineering Processes*, McGraw-Hill, New York, 1958

[Boning *et al.*, 2008] D.S. Boning, K. Balakrishnan, C. Hong, N. Drego, A. Farahanchi, K.M. Gettings, L. Daihyun, A. Somani, H. Taylor, D. Truque, and X. Xiaolin, Variation. *Semiconductor Manufacturing, IEEE Transactions on*, 21(1):63–71, 2008

[Boothroyd and Redford, 1968] G. Boothroyd, A.H. Redford, *Mechanized Assembly*, McGraw-Hill, London, 1968

[Boothroyd and Dewhurst, 1990] G. Boothroyd, and P. Dewhurst, *Product Design for Assembly,* Wakefield, RI, USA, 1990

[Boothroyd, 1994] G. Boothroyd, Product Design for Manufacture and Assembly, *Computer Aided Design*, Vol. 26, Issue 7, 1994

[Bralla, 1998] J.G. Bralla, *Design for Manufacturability Handbook*, Second Edition - McGraw-Hill Professional, 1998

[Braun and MacDonald, 1982] E. Braun and S. MacDonald, *Revolution in miniature: The history and impact of semiconductor electronics*. Cambridge University Press, 1982

[Brinkman *et al.*, 1997] W.F. Brinkman, D.E. Haggan and W.W. Troutman, A history of the invention of the transistor and where it will lead us. *IEEE Journal of Solid-State Circuits*, 32(12):1858 –1865, 1997

[Buccella and Cechich, 2003] A. Buccella and A. Cechich, An ontology approach to data integration. *JCS&T*, Vol. 3 (No.2), 2003

[Buccella *et al.,* 2005] A. Buccella, A. Cechich, N. Brisaboa, Ontology-Based Data Integration Methods: a Framework for Comparison*, Revista Colombiana de Computaci´on*, Vol. 6, No. 1, 2005

[Business-Directory, 2007] BNET Business-Directory. *Strategic analysis*, 2007

[Businessballs, 2006] Businessballs. *Pest market analysis tool*, 2006

[Busse *et al.,* 1999] S. Busse, R.D. Kutsche, U. Leser, and H. Weber, Federated Information Systems: Concepts, Terminology and Architectures. *Technical Report*. Nr. 99-9, TU Berlin, 1999

[C]

[Calvanese *et al.,* 2001] D. Calvanese, G. De. Giacomo, and M. Lenzerini, In Proc. of the 14th Int. Workshop on Description Logic (DL 2001), volume 49 of CEUR Electronic Workshop Proceedings, http://ceur-ws.org/, pages 10-19, 2001

[Carrillo and Franza, 2006] J.E. Carrillo and R.M. Franza, Investing in product development and production capabilities: The crucial linkage between time-to-market and ramp-up time. *European Journal of Operational Research*, 171(2):536 – 556, 2006

[Chang and Sze, 1996] C.Y. Chang and S.M. Sze, ULSI technology, McGraw-Hill, Materials science and process technology series, 1996

[Cho and Hsu, 1997] J. Cho and C. Hsu, Manufacturing driven design, *IEEE International Conference on Systems, Man and Cybernetics*, 1997

[Cliff, 2003] M. Cliff, DFM - An Industry Paradigm Shift, *International Test Conference (ITC'03)*, 2003

[Codd, 1970] E.F. Codd, A relational model of data for large shared data banks. *Communications of the ACM*, 13(6):377–387, 1970

[Colledani and Tolio, 2009] M. Colledani and T. Tolio, Performance evaluation of production systems monitored by statistica process control and off-line inspections. *International Journal of Production Research*, 120(2):348–367, 2009

[Colledani, 2008] M. Colledani, Integrated analysis of quality and production logistics performance in asynchronous manufacturing lines. In *IFAC World Congress*, 2008

[Comai and Tena, 2007] A. Comai, and J. Tena, Early warning systems for your competitive landscape. *Society of Competitive Intelligence Professionals*, 10(3), June 2007

[Corbett, 1991] J. Corbett, M. Dooner, J. Meleka, and C. Pym, *Design for Manufacture Strategies, Principles and Techniques*, Addison-Wesley, 1991

[Corcho and Gomez-Perez, 2000] O. Corcho, and A.G. Perez, Evaluating knowledge representation and reasoning capabilities of ontology specification languages. In *Proceedings of the ECAI 2000 Workshop on Applications of Ontologies and Problem-Solving Methods*, Berlin, 2000

[Cottrell and Grebinski, 2003] D.R. Cottrell, T.J. Grebinski, Interoperability Beyond Design: Sharing Knowledge between Design and Manufacturing, isqed, pp.214, *Fourth International Symposium on Quality Electronic Design*, 2003

[Curé, 2005] O. Curé, Mapping Databases to Ontologies to Design and Maintain Data in a Semantic Web Environment, *Systemic, Cybernetics and Informatics* Volume-4, number-4, 2005

[Coutinho, 1964] J.S. Coutinho, Failure effect analysis. *Trans. NY Acad. Sci.*, 26(II):564–584, 1964

[Cui and Brien, 2000] Z. Cui, and P. O'Brien, Domain Ontology Management Environment. In *Proceedings of the 33rd Hawaii International Conference on System Sciences*, 2000

[Cui and Brien, 2001] Z. Cui, D. Jones, and P. O'Brien, Issues in Ontology-based Information Integration. *IJCAI* – Seattle, USA, 2001

[D]

[Dale, 2012] F. Dale, Intel's semiconductor market share surges to more than 10-year high in 2011, March 2012

[Darmont *et al.*, 2007] J. Darmont, F. Bentayeb, O. Boussaid, Benchmarking Data Warehouses, *International Journal of Business Intelligence and Data Mining*, 2007

[Das and Kanchanapiboon, 2009] S. Das, and A. Kanchanapiboon, A multi-criteria model for evaluating design for manufacurability. *International Journal of Production Research*, 49(4):1197–1217, 2009

[Dauzere-Peres *et al.*, 2011] S. D. Peres, J.L. Rouveyrol, C. Yugma, and P. Vialletelle, A smart sampling algorithm to minimize risk dynamically. In *Advanced Semiconductor Manufacturing Conference*, 2011

[Henderson, 1998] D. Henderson, Overcoming data integration problems, Info Management Direct, 1998

[Deridder and Wouters, 2000] D. Deridder, B. Wouters, The Use of an Ontology to Support a Coupling between Software Models and Implementation, *European Bibliography Conference on Object-Oriented Programming (ECOOP'00)*, International Workshop on Model Engineering, 2000

[DFM Dictionary 2004] Si2, DFM Dictionary, http://www.si2.org/openeda.si2.org/dfmcdictionary/index.php/Main_Page, retrieved on 12th july 2011

[Dhillon, 1999] B.S. Dhillon, *Design Reliability, Fundamentals and Applications*. CRC Press LLC, Florida USA, 1999

[Doan and Halevy, 2005] A. Doan and A.Y. Halevy, Semantic-integration research in the database community: a brief survey. *AI Magazine*, 26(1):83–94, 2005

[Dou and Lependu, 2006] D. Dou and P. LePendu, Ontology-based Integration for Relational Databases. In *Proceedings of ACM Symposium on Applied Computing (SAC)*, pages 461-466, 2006

[Dummer, 1997] G.W.A. Dummer, *Electronics Inventions and Discoveries*. Taylor & Francis, 1997.

[Duncan, 1956] A.J. Duncan, The economic design of X-bar charts used to maintain current control of a process, *Journal of the American Statistical Association*, vol. 51(274), pp. 228-242, 1956

[E]

[Eskelinen, 2001] H. Eskelinen, Improving the productivity of complex electronic systems design by utilizing applied design methodologies, *IEEE AESS Systems Magazine*, Vol. 16, Issue 10, 2001

[F]

[Fiedler et al., 2005] G. Fiedler, T. Raak, and B. Thalheim, Database collaboration instead of integration. In *APCCM'05*, 2005

[G]

[Gilad, 2011] B. Gilad, (President Academy of Competitive Intelligence Boca Raton Florida USA). Strategy without intelligence, intelligence without strategy. *Business Strategy Series*, 12(1):4–11, 2011

[Guarino, 2002] N. Guarino, Ontology-Driven Conceptual Modeling, Tutorial at *21st International Conference on Conceptual Modeling* (ER'02). Tampere, Finland, 2002

[Glaser *et al.,* 2004] H. Glaser, H. Alani, L. Carr, S. Chapman, F. Ciravegna, A. Dingli, N. Gibbins, S. Harris, M.C. Schraefel, N. Shadbolt, CS AKTive Space: Building a Semantic Web Application. In First European Web Symposium (ESWS'04). Springer Verlag, pp. 417–432, 2004

[Goh, 1997] C.H. Goh, Representing and Reasoning about Semantic Conflicts in Heterogeneous Information Sources. Phd, MIT, 1997

[Gruber, 1993] T. Gruber, A translation approach to portable ontology specifications. *Knowledge Acquisition* 1993 –Vol. 5(2), 199–220, 1993

[Guarino and Persidis, 2003] N. Guarino, A. Persidis, Evaluation Framework for Content Standards. Deliverable 3.5, OntoWeb EU project (IST-2000-29243), 2003

[H]

[Heimbigner and McLeod, 1985] D. Heimbigner and D. McLeod, A Federated Architecture for Information Management, *ACM Transactions on Office Information Systems*, Vol. 3, No. 3, pp. 253-278, July 1985

[Herrmann *et al.,* 2004] J.W. Herrmann, J. Cooper, S.K. Gupta, C.C. Hayes, K. Ishii, D. Kazmer, P.A. Sandborn and W.H. Wood, New Directions in Design for Manufacturing, *Proceedings of DETC'04 ASME 2004 design engineering technical conferences and computers and information in engineering conference*, 2004

[Hsu, 1977] J.I.S. Hsu, A cost model for skip-lot destructive sampling. *IEEE Transactions on Reliability*, 26(1):70–72, 1977

[Humphrey, 2005] A.S. Humphrey, SWOT analysis for management consulting, 2005 (newsletter) from http://www.sri.com/about/alumni

[Hurat and Cote, 2005] P. Hurat and M. Cote, DFM for Manufacturers and Designers, 25th Annual BACUS Symposium on Photomask Technology, *Proceedings of SPIE,* Volume 5992 pp. 126-132, 2005

[Hurat *et al.,* 2006] P. Hurat, M. Cote, C.M. Tsai, J. Brandenburg, A Genuine Design Manufacturability Check for Designers, *Proc. of SPIE* Vol. 6156, 2006

[Hurson and Bright, 1991] A.R. Hurson, and M.W. Bright, Multidatabase Systems: An Advanced Concept in Handling Distributed Data. *Advances in Computers*, 32:149–200, 1991

[Hurtarte, 2007] J. S. Hurtarte, *Understanding Fabless IC Technology*. San Diego, CA, Elsevier, 2007

[Honma et al., 2009] M. Honma, H. Itoh, N. Iriki, S. Kobayashi, S. Kuwabara, N. Miyazaki, H. Suzuki, N. Yoshioka, S. Arima and K. Kadota, Organized DFM for total optimization in semiconductor manufacturing, *AEC/ APC Symposium* Asia 2009

[I]

[Impagliazzo, 2012] J. Impagliazzo. A brief history of database systems. Retrieved July 20th, 2012 from http://www.comphist.org/computing_history/new_page_9.htm

[Inmon, 2005] W.H. Inmon, *Building the Data Warehouse*. Wiley; 4 edition (October 7, 2005), 2005

[International Business Strategies, 2009] International Business Strategies, *Key trends in technology and supply for advanced features within IC industry*. Technical report, by International Business Strategies, Inc., ibs_inc@ix.netcom.com, October 2009

[Ireson *et al.*, 1996] G.W. Ireson, C.F. Coombs, and R.Y. Moss, *The Handbook of Reliability Engineering and Management*. McGraw-Hill New York, 1996

[J]

[James, 2009] D. James, Design-for-manufacturing features in nanometer logic processes – a reverse engineering perspective, *IEEE Custom Integrated Circuits Conference (CICC)*, 2009

[Jones, 2012] S.W. Jones, Introduction to integrated circuit technology, retrieved on 20th July 2012, http://www.icknowledge.com/freecontent/Introduction%20to%20IC%20technology%20rev%205.pdf

[Johann et al., 2002] E. Johann, K. Christian and M. Tadeusz, The COMET Metamodel for Temporal Data Warehouses, In Proc. of the *14th Int. Conference on Advanced Information Systems Engineering (CAISE'02)*, 2002

[Jordan, 1972] W.E. Jordan, Failure modes, effects and criticality analyses. In *Proc. Annu. Reliability Maintainability Symp.*, pages 30–37, 1972

[K]

[Kahng, 2010] A.B. Kahng, Scaling: More than Moore's law, *IEEE Design & Test Computers*, Vol. 27, pp. 86-87, 2010

[Kasem, 1997] M. Kasem, Integrated development and manufacturing methodology for advanced power MOSFETs, *IEE/CPMT Int'l Electronics Manufacturing Technology Symposium*, 1997

[Kumar, 2008] R. Kumar, *Fabless Semiconductor Implementation*. USA, McGraw-Hill Professional, 2008

[Konstantinou et al., 2006] N. Konstantinou, D. Spanos, M. Chalas, E. Solidakis, and N Mitrou, VisAVis: An Approach to an Intermediate Layer between Ontologies and Relational Database Contents. *International Workshop on Web Information Systems Modeling* (WISM2006) Luxembourg, 2006

[Koren et al., 1999] Y. Koren, U. Heisel, F. Jovane, T. Moriwaki, G. Pritschow, G. Ulsoy and H. Van Brussel, Reconfigurable Manufacturing Systems, Annals of the CIRP Vol. 48/2/1999, Retrieved from: http://www-personal.umich.edu/~ulsoy/pdf/RMS_CIRP99_Keynote.pdf

[L]

[Lapedus, 2008] M. LaPedus, TSMC's 32-nm DFM model aims to unify design flow, retrieved from: http://www.eetimes.com/showArticle .jhtml?articleID=208402578, 2008

[Landers and Rosenberg, 1982] T. Landers, and R.L. Rosenberg, An Overview of MULTIBASE. In *Second International Symposium on Distributed Data Bases*, pages 153–184, Berlin, Germany, Holland, 1982

[Lee, 2002] J. Lee, Artificial intelligence-based sampling planning system for dynamic manufacturing process. *Expert Systems with Applications*, 22(2):117–133, 2002

[Lefkow, 2012] C. Lefkow, Global electronics sales 'to surpass 1 *trillion in* 2012', *January* 2012

[Lenzerini, 2002] M. Lenzerini, Data integration: A theoretical perspective. In *PODS*, pages 233–246, 2002

[Liddle et al., 2003] S.W. Liddle, K.A. Hewett and D.W. Embley, An Integrated Ontology Development Environment for Data Extraction, In *Proceedings of the 2nd International Conference on Information System Technology and its Applications (ISTA2003)*, Lecture Notes in Informatics, vol. P-30, pp. 21-33, 2003

[Liebmann, 2008] L. Liebmann, DFM, the teenage years. *Proc. of SPIE*, page 6925, 2008

[M]

[Maier et al., 2003] A. Maier, H. P. Schnurr and Y. Sure, "Ontology-based Information Integration in the Automotive Industry", In *Proc. of ISWC* 2003, pp. 897-912

[Mason, 2007] M.E. Mason, DFM EDA Technology A Lithographic Perspective, IEEE Symposium on VLSI Technology, 2007

[Matteo et al., 2006] M.G. Stefano and L. Jens, R. Stefano, and V. Gottfried, Schema versioning in data warehouses: Enabling cross-version querying via schema augmentation. *Data & Knowledge Engineering*, 59(2):435–459, 2006

[McGregor, 2007] J. McGregor, The common platform technology: A new model for semiconductor manufacturing. *Technical Report*, 2007

[Mehrabi et al., 2002] M.G. Mehrabi, A.G. Ulsoy, Y. Koren, P. Heytler, Trends and perspectives in flexible and reconfigurable manufacturing systems, *Journal of Intelligent Manufacturing*, 13, 135-146, 2002

[Mena et al., 1996] E. Mena, V. Kashyap, A. Sheth, and A. Illarramendi, Observer: An approach for query processing in global information systems based on interoperability between pre-existing ontologies. In *Proceedings 1st IFCIS International Conference on Cooperative Information Systems (CoopIS '96)*. Brussels, 1996

[Mindtools, 2007] SWOT analysis: discover new opportunities. Manage and eliminate threats, Retrieved on 20th July 2012, from http://www.mindtools.com/pages/article/newTMC_05.htm.

[Moore, 1998] G.E. Moore, Cramming more components onto integrated circuits. *Proceedings of the IEEE*, 86(1):82–85, January 1998

[Morinaga *et al.,* 2006] H. Morinaga, H. Kakinuma, T. Ito, and T. Higashiki, Development of a platform for collaborative engineering data flow between design and manufacturing. In *IEEE International Symposium* on *Semiconductor Manufacturing*, pages 45 –48, 2006

[Mouli and Scott, 2007] C. Mouli, and M. Scott, Adaptive metrology sampling techniques enabling higher precision in variability detection and control. *In Advanced Semiconductor Manufacturing Conference*, 2007

[Myint and Tabucanon, 1998] S. Myint, and M.T. Tabucanon, The framework for an expert system to generate alternative products in concurrent engineering design. *The framework for an expert system to generate alternative products in concurrent engineering design*, 37(2):125–134, 1998

[N]

[Nyulas et al., 2007] C. Nyulas, M. O. Connor, and S. Tu. DataMaster – a plug-in for importing schemas and data from relational databases into Protégé. In Proc. of Protégé 2007, Stanford Medical Informatics, 2007

[O]

[OMG, 1997] Object Management Group, Meta Object Facility (MOF) Specification, Joint Revised Submission, OMG Document ad/97-08-14, 1997 retrieved from http://www.omg.org

[Ouyang et al., 2000] C. Ouyang, H. T. Heineken, J. Khare, S. Shaikh and M. d'Abreu, Maximizing wafer productivity through layout optimizations, *13th International Conference on VLSI design*, 2000

[P]

[Palopoli et al., 2000] L. Palopoli, G. Terracina and D. Ursino: The System DIKE: Towards the Semi-Automatic Synthesis of Cooperative Information Systems and Data Warehouses. *ADBIS-DASFAA* 2000, Matfyzpress, 108-117

[Peters, 2005] L. Peters, DFM Worlds Collide, Then Cooperate, *EETimes Magazine*, 2005 retrieved from: http//www. eetimes.com

[Porter, 1996] M. E. Porter, What is strategy? Harvard Business Review, pages 61–78, November-December 1996

[Porter, 1998] M. E. Porter, Competitive advantage: creating and sustaining superior performance, PB: Free Press, 1998

[Q]

[Quirke et al., 1992] M.T. Quirke, J.J. Barrett and M. Hayes, Planar Magnetic Component Technology A review, *IEEE Transactions on Components, Hybrids and Manufacturing Technology*, Vol. 15, No. 5 (884-892), 1992

[Quirk and Serda, 2000] M. Quirk, and J. Serda, *Semiconductor Manufacturing Technology*. Prentice Hall; 1 edition (November 19, 2000), 2000

[R]

[Raaij and Verhallen, 2005] V.W.F. Raaij, and T.M.M. Verhallen, *Adaptive sampling methodology for in-line defect inspection*, 2005

[Radojcic et al., 2009] R. Radojcic, D. Perry and M. Nakamoto, Design for manufacturability for fabless manufactuers. *Solid-State Circuits Magazine*, IEEE, 1(3):24–33, 2009

[Rahm and Bernstein, 2001] E. Rahm and P.A. Bernstein, A survey of approaches to automatic schema matching. *VLDB* J., 10(4):334–350, 2001

[Ramakrishnan and Gehrke, 2003] R. Ramakrishnan and J. Gehrke: *Database Management Systems* 3rd Edition, McGraw Hill, Chapter 3: The Relational Model, pages 57-99, 2003

[Ramana and Rao, 2002] K.V. Ramana and P.V.M. Rao, Automated manufacturability evaluation of sheet metal components: a review, *ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (IDETC/ CIE2002), 2002

[Raina, 2006] R. Raina, What is DFM, DFY and why should i care? *IEEE International Test Conference, ITC-2006*, Santa Clara, CA, USA, 2006

[Redford, 2009] M. Redford, J. Sawicki, P. Subramaniam, C. Hou, Y. Zorian and K. Michaels, DFM—Don't Care or Competitive Weapon?, *46th ACM/IEEE Design Automation Conference (DAC'09)*, pp. 296-297, 2009

[Reid and Sanders, 2005] R.D. Reid and R.N. Sanders, *Operations Management: An integrated approach*. 2nd Edition © Wiley 2005

[Reddy, 2000] S.V. Reddy, Digital design flow options, *Master Thesis submitted at Ohio State University*, retrieved from: http:// www.ohio.edu, 2000

[Reifer, 1979] D.J. Reifer, Software failure modes and effects analysis. *IEEE Transactions on Reliability*, R-28(3):247–249, 1979

[Reynolds *et al.*, 1988] M.R. Reynolds, R.W. Amin, J.C. Arnold, and J.A. Nachlas, X charts with variable sampling intervals. *Technometrics*, 30(2):181–192, 1988

[Rezayat, 2000] M. Rezayat, Knowledge-based product development using xml and kcs. *Computer Aided Design*, 32(5):299–309, 2000

[Ross, 2003] P.E. Ross, Commandments [technology laws and rules of thumb], *Spectrum, IEEE*, 40(12):30–35, 2003

[S]

[Sahnoun *et al.*, 2011] M. Sahnoun, B. Bettayeb, P. Vialletelle, A. Mili, and M. Tollenaere, Impact of sampling on w@r and metrology time delay, In *Intel European Research & Innovation Conference*, 2011

[Schuermyer *et al.,* 2005] C. Schuermyer, K. Cota, R. Madge, B. Benware, Identification of systematic yield limiters in complex ASIC through volume structural test fail data visualization and analysis, *IEEE International Test Conference Proceedings* (ITC), 2005

[Shahzad *et al.*, 2011a] M. K. Shahzad, S. Hubac, A. Siadat, and M. Tollenaere, An Extended IDM Business Model to Ensure Time-to-Quality in Semiconductor Manufacturing Industry, volume 219 of *Communications in Computer and Information Science*, pages 118–128. Springer Berlin Heidelberg, 2011. 10.1007/978-3-642-24358-5_12

[Shahzad *et al.*, 2011b] M.K. Shahzad, S. Hubac, A. Siadat, and M. Tollenaere, An extended business model to ensure time-to-quality in semiconductor manufacturing industry, *In International Conference on Enterprise Information Systems*, 2011

[Sheth and Larson, 1990] A.P. Sheth and J.A. Larson, Federated Database Systems for Managing Distributed, Heterogeneous, and Autonomous Databases, *ACM Computing Surveys*, Vol. 22, No. 3, pp. 183-236, 1990

[Shingo and Dillion, 1989] S. Shingo and A. P. Dillion, *A Study of the Toyota Production System*, Productivity Press, 1989

[Seino et al., 2009] T. Seino, S. Honda and T. Tanaka, Manufacturing technology management to accelerate "design for manufacturability", *Portland International Conference on Management of Engineering & Technology (PICMET)* pp. 1137-1142, 2009

[Spanos, 1991] C.J. Spanos, Statistical process control in semiconductor manufacturing. *Microelectronics Engineering*, 10(3-4):271–276, 1991

[Spyns *et al.,* 2002] P. Spyns, D. Oberle, R. Volz, J. Zheng, M. Jarrar, Y. Sure, R. Studer, R. Meersman, OntoWeb - a Semantic Web Community Portal. In: Karagiannis, D., Reimer, U., (eds.): *Proceedings of the Fourth International Conference on Practical Aspects of Knowledge Management* (PAKM'02), LNAI 2569, Springer Verlag. pp. 189–200, 2002

[Stamford, 2012] C. Stamford, Market *share analysis: Total semiconductor revenue, worldwide*, April 2012.

[Stoffel et al., 1997] K. Stoffel, M. Taylor and J. Hendler, "Efficient Management of Very Large Ontologies", *Proceedings of American Association for Artificial Intelligence Conference, (AAAI-97)*, AAAI/MIT Press, 1997

[T]

[Tzitzikas *et al.,* 2001] Y. Tzitzikas, N. Spyratos, P. Constantopoulos, Mediators over Ontology-based Information Sources. *Proceedings of the Second International Conference on Web Information Systems Engineering* (WISE'01), 2001

[Treat *et al.*, 1996] J. Treat, G. Thibault, and A. Asin, Dynamic competitive simulation: war gaming as a strategic tool. *Strategy and Business*, pages 46–54, 1996

[Trybula, 1995] W. J. Trybula, The evolution of design for manufacturing, *IEEE Technical Applications Conference and Workshops (Northcon95)*, pp. 332-334, 1995

[Turakhia *et al.*, 2009] R. Turakhia, M. Ward, S. K. Goel and B. Benware, Bridging DFM Analysis and Volume Diagnostics for Yield Learning – A Case Study, *27th IEEE VLSI Test Symposium (VTS'09)* pp. 167-172, 2009

[Tuomi, 2004] I. Tuomi, Industrial structure and policy choice: Notes on the evolution of semiconductors and open source. *Technical report*, Joint Research Centre, Institute for Prospective Technological Studies Spain, 2004

[V]

[Venkataraman, 2007] S. Venkataraman, Does Test Have a Greater Role to Play in the DFM process, *IEEE International Test Conference (ITC)*, pp. 1-1, 2007

[Vliet and Luttervelt, 2004] H. W. V. Vliet, and K. V. Luttervelt, Development and application of a mixed product/process-based DFM methodology. *International Journal of Integrated Manufacturing*, 17(3):224–234, 2004

[W]

[Wache et al., 2001] H. Wache, T. Vögele, U. Visser, H. Stuckenschmidt, G. Schuster, H. Neumann, and S. Hübner, Ontology - based Integration of Information - A Survey of Existing Approaches. In: Proceedings of IJCAI-01 *Workshop: Ontologies and Information Sharing, Seattle*, WA, 108-117, 2001

[Wand et al., 1999] Y. Wand, V. Storey, R. Weber, An Ontological Analysis of the relationship Construct in Conceptual Modeling, *ACM Transactions on Database Systems*, Vol. 24, No. 4, pp. 494– 528, 1999

[Webb, 2008] C. Webb, Intel design for manufacturing and evolution of design rules. *Intel Technology Journal*, 12(2), 2008

[Welty and Smith, 2001] C. Welty and B. Smith (eds.), Formal Ontology and Information Systems, New York, *ACM Press*, 2001

[Weihrich, 1982] H. Weihrich, The TOWS matrix – A tool for situational analysis. *Long Range Planning*, 15:54–66, 1982 retrieved from http://www.usfca.edu/fac_staff/weihrichh/docs/tows.pdf

[White et al., 1995] K.P. White, R.N. Athay and W.J. Trybula, Applying DFM in the Semiconductor Industry, *17th IEEECPMT International Electronics Manufacturing Technology Symposium*, pp. 438-441, 1995

[White and Trybula, 1996] K.P. White, W.J. Trybula, DFM for the next generation, *19th IEEE/CPMT Electronics Manufacturing Technology Symposium*, pp. 109-116, 1996

[Winer, 1993] L. Winer, The Strategic Creative Analysis (SCAN) Process. *Case studies Center for Applied Research*, Lubin School of Business, Pace University, 1993

[Wong et al., 2008] B.P. Wong, A. Mittal, G.W. Starr, F. Zach, V. Moroz, and A. Kahng, Nano-CMOS Design for Manufacturability: *Robust Circuit and Physical Design for Sub-65nm Technology Nodes*. John Wiley & Sons, 2008

[Worrall, 1998] L. Worrall, Strategic analysis: a scientific art, Wolverhampton Business School, Management Research Centre, 1998

[Wu, 2003] S. L. Wu, Industry Dynamics within Semiconductor Value Chain IDM, *Foundry and Fabless, Technical report*, USA, 2003

[Y]

[Yoon et al., 2010] M. Yoon, and F. Malerba, Technological interrelatedness, knowledge generality and economies of scale in the evolution of firm boundaries: A history-friendly model of the fabless ecosystem. *In International Conference on Opening Up Innovation: Strategy, Organization and Technology* at Imperial College London Business School, 2010

[Youn and McLeod, 2006] S. Youn, and D. McLeod, (Eds.). Ontology Development Tools for Ontology-Based Knowledge Management, In M. Khosrow-Pour (Ed.), *Encyclopedia of e-commerce, e-government and mobile commerce* (pp. 858-864) Hershey, PA: Idea Group 2006

[Young and Murray, 2007] D. Young, S. Murray, Industrial Engineering FE Exam Preparation, Kaplan Education, Dearborn Publishing, Inc. 2007

[Yuyin et al., 1996] S. Yuyin, Z. Bopeng, C. Fuzhi and M. Qingguo, A Knowledge-based Design For Manufacture System, *IEEE International Conference on Systems, Man and Cybernetics*, pp. 1220-1224, 1996

[Z]

[Zeigler *et al.,* 2000] B.P. Zeigler, H. Praehofer, and T.G. Kim, Theory of Modeling and Simulation, Academic Press, San Diego, CA, 2000

[Zuling and Hongbing, 2005] K. Zuling and W. Hongbing, Implementation and Application of Ontology Databases with User-Defined Rules (UDR) Supported, *Proceedings of the First International Conference on Semantics, Knowledge, and Grid*, 2005

[Zuozhi and Jami, 2005] Z. Zuozhi, and S.J. Jami, Domain independent shell for DFM and its applications to sheet metal forming and injection molding, *Computer Aided Design*,Vol.37, No. 9, pp 881-898, 2005

Vita

Mr. Muhammad Kashif Shahzad was born in small town of Kasur in Pakistan. He completed his bachelor degrees in mechanical engineering and computer science from university of engineering and technology, Lahore and Allama Iqbal Open University, Islamabad (1995-2000). He started his career as Lecturer of Computer Science at Petroman Training Institute, Lahore (1999-2001) and then he joined WAPDA (water & power development authority) as Assistant Director (Programmer/System Analyst) where he spent 6 years and held diverse responsibilities ranging from network design and deployment to the programming with 4-GL languages. Besides his professional responsibilities, he also completed Master degree in total quality management (TQM) from University of Punjab (2006). He decided to come to France in 2007 for PhD studies in industrial engineering to effectively merge his knowledge and experience of mechanical engineering, quality management and computer science. He completed his Master degree in industrial engineering from ENSGI, INP Grenoble in 2008 and started PhD with STMicroelectronics titled <<dynamic exploitation of production databases to improve DFM methods in semiconductor industry>>. He has attended advanced training on engineering data analysis (EDA) and modeling and design of experiments training (DOE) from MIT, USA and is also certified green belt of design for six sigma (DFSS). His research interest includes DFM methods, data warehouse modeling, system modeling, business process reengineering, capacity optimization semiconductor manufacturing processes.

Permanent address (or email): 20, Yasrib Street, Akbar Shaheed Road, Kot Lakhpat, Lahore
Pakistan, shahzad347@gmail.com

This dissertation was typed by Muhammad Kashif Shahzad.