



HAL
open science

Simulation et optimisation du transport automatisé dans la fabrication de semi-conducteurs.

Téwendé Jean-Etienne Arthur Kiba

► To cite this version:

Téwendé Jean-Etienne Arthur Kiba. Simulation et optimisation du transport automatisé dans la fabrication de semi-conducteurs.. Autre. Ecole Nationale Supérieure des Mines de Saint-Etienne, 2010. Français. NNT : 2010EMSE0582 . tel-00611812

HAL Id: tel-00611812

<https://theses.hal.science/tel-00611812>

Submitted on 27 Jul 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NNT : 2010 EMSE 0582

THÈSE

présentée par

Téwendé Jean-Etienne KIBA

pour obtenir le grade de
Docteur de l'École Nationale Supérieure des Mines de Saint-Étienne

Spécialité : **Génie Industriel**

SIMULATION ET OPTIMISATION DU TRANSPORT AUTOMATISE DANS LA FABRICATION DE SEMI-CONDUCTEURS

soutenue à Gardanne, le 10 Novembre 2010

Membres du jury

Président	Dominique FEILLET	Professeur, Ecole des Mines de Saint-Etienne
Rapporteurs	Pierre CASTAGNA	Professeur, Université de Nantes
	Christian PRINS	Professeur, Université Technologique de Troyes
	Oliver ROSE	Professeur, Technology University of Dresden, Allemagne
Examineurs	Christian ARTIGUES	Chargé de Recherche au LAAS-CNRS, Toulouse
	Gilles LAMIABLE	Ingénieur, STMicroelectronics
Invité	Leon VERMARIEN	Ingénieur, STMicroelectronics
Directeur de thèse	Stéphane DAUZERE-PERES	Professeur, Ecole des Mines de Saint-Etienne
Co-encadrant	Claude YUGMA	Maître-Assistant, Ecole des Mines de Saint-Etienne

Spécialités doctorales :

SCIENCES ET GENIE DES MATERIAUX
 MECANIQUE ET INGENIERIE
 GENIE DES PROCEDES
 SCIENCES DE LA TERRE
 SCIENCES ET GENIE DE L'ENVIRONNEMENT
 MATHEMATIQUES APPLIQUEES
 INFORMATIQUE
 IMAGE, VISION, SIGNAL
 GENIE INDUSTRIEL
 MICROELECTRONIQUE

Responsables :

J. DRIVER Directeur de recherche – Centre SMS
 A. VAUTRIN Professeur – Centre SMS
 G. THOMAS Professeur – Centre SPIN
 B. GUY Maître de recherche – Centre SPIN
 J. BOURGOIS Professeur – Centre SITE
 E. TOUBOUL Ingénieur – Centre G2I
 O. BOISSIER Professeur – Centre G2I
 JC. PINOLI Professeur – Centre CIS
 P. BURLAT Professeur – Centre G2I
 Ph. COLLOT Professeur – Centre CMP

Enseignants-chercheurs et chercheurs autorisés à diriger des thèses de doctorat (titulaires d'un doctorat d'État ou d'une HDR)

AVRIL	Stéphane	MA	Mécanique & Ingénierie	CIS
BATTON-HUBERT	Mireille	MA	Sciences & Génie de l'Environnement	SITE
BENABEN	Patrick	PR 1	Sciences & Génie des Matériaux	CMP
BERNACHE-ASSOLLANT	Didier	PR 0	Génie des Procédés	CIS
BIGOT	Jean-Pierre	MR	Génie des Procédés	SPIN
BILAL	Essaïd	DR	Sciences de la Terre	SPIN
BOISSIER	Olivier	PR 1	Informatique	G2I
BORBELY	Andras	MR	Sciences et Génie des Matériaux	SMS
BOUCHER	Xavier	MA	Génie Industriel	G2I
BOUDAREL	Marie-Reine	PR 2	Génie Industriel	DF
BOURGOIS	Jacques	PR 0	Sciences & Génie de l'Environnement	SITE
BRODHAG	Christian	DR	Sciences & Génie de l'Environnement	SITE
BURLAT	Patrick	PR 2	Génie industriel	G2I
COLLOT	Philippe	PR 1	Microélectronique	CMP
COURNIL	Michel	PR 0	Génie des Procédés	SPIN
DAUZERE-PERES	Stéphane	PR 1	Génie industriel	CMP
DARRIEULAT	Michel	IGM	Sciences & Génie des Matériaux	SMS
DECHOMETS	Roland	PR 1	Sciences & Génie de l'Environnement	SITE
DESRAYAUD	Christophe	MA	Mécanique & Ingénierie	SMS
DELAFOSSÉ	David	PR 1	Sciences & Génie des Matériaux	SMS
DOLGUI	Alexandre	PR 1	Génie Industriel	G2I
DRAPIER	Sylvain	PR 2	Mécanique & Ingénierie	SMS
DRIVER	Julian	DR 0	Sciences & Génie des Matériaux	SMS
FEILLET	Dominique	PR 2	Génie Industriel	CMP
FOREST	Bernard	PR 1	Sciences & Génie des Matériaux	CIS
FORMISYN	Pascal	PR 1	Sciences & Génie de l'Environnement	SITE
FORTUNIER	Roland	PR 1	Sciences & Génie des Matériaux	SMS
FRACZKIEWICZ	Anna	DR	Sciences & Génie des Matériaux	SMS
GARCIA	Daniel	MR	Génie des Procédés	SPIN
GIRARDOT	Jean-Jacques	MR	Informatique	G2I
GOEURIOT	Dominique	MR	Sciences & Génie des Matériaux	SMS
GRAILLOT	Didier	DR	Sciences & Génie de l'Environnement	SITE
GROSSEAU	Philippe	MR	Génie des Procédés	SPIN
GRUY	Frédéric	MR	Génie des Procédés	SPIN
GUY	Bernard	MR	Sciences de la Terre	SPIN
GUYONNET	René	DR	Génie des Procédés	SPIN
HERRI	Jean-Michel	PR 2	Génie des Procédés	SPIN
INAL	Karim	PR 2	Microélectronique	CMP
KLÖCKER	Helmut	DR	Sciences & Génie des Matériaux	SMS
LAFOREST	Valérie	CR	Sciences & Génie de l'Environnement	SITE
LERICHE	Rodolphe	CR CNRS	Mécanique et Ingénierie	SMS
LI	Jean-Michel	EC (CCI MP)	Microélectronique	CMP
LONDICHE	Henry	MR	Sciences & Génie de l'Environnement	SITE
MALLIARAS	George Grégory	PR 1	Microélectronique	CMP
MOLIMARD	Jérôme	MA	Mécanique et Ingénierie	SMS
MONTHEILLET	Frank	DR 1 CNRS	Sciences & Génie des Matériaux	SMS
PERIER-CAMBY	Laurent	PR 2	Génie des Procédés	SPIN
PIJOLAT	Christophe	PR 1	Génie des Procédés	SPIN
PIJOLAT	Michèle	PR 1	Génie des Procédés	SPIN
PINOLI	Jean-Charles	PR 0	Image, Vision, Signal	CIS
STOLARZ	Jacques	CR	Sciences & Génie des Matériaux	SMS
SZAFNICKI	Konrad	MR	Sciences & Génie de l'Environnement	SITE
THOMAS	Gérard	PR 0	Génie des Procédés	SPIN
TRIA	Assia		Microélectronique	CMP
VALDIVIESO	François	MA	Sciences & Génie des Matériaux	SMS
VAUTRIN	Alain	PR 0	Mécanique & Ingénierie	SMS
VIRICELLE	Jean-Paul	MR	Génie des procédés	SPIN
WOLSKI	Krzysztof	DR	Sciences & Génie des Matériaux	SMS
XIE	Xiaolan	PR 1	Génie industriel	CIS

Glossaire :

PR 0	Professeur classe exceptionnelle
PR 1	Professeur 1 ^{ère} classe
PR 2	Professeur 2 ^{ème} classe
MA(MDC)	Maître assistant
DR	Directeur de recherche
Ing.	Ingénieur
MR(DR2)	Maître de recherche
CR	Chargé de recherche
EC	Enseignant-chercheur
IGM	Ingénieur général des mines

Centres :

SMS	Sciences des Matériaux et des Structures
SPIN	Sciences des Processus Industriels et Naturels
SITE	Sciences Information et Technologies pour l'Environnement
G2I	Génie Industriel et Informatique
CMP	Centre de Microélectronique de Provence
CIS	Centre Ingénierie et Santé

A ma mère ...

Remerciements

Les travaux présentés dans cette thèse ont été réalisés dans le cadre d'une Convention Industrielle de Formation par la Recherche (CIFRE n^o 1139/2007) accordée par l'Association Nationale de la Recherche Technique (ANRT), en collaboration avec STMicroelectronics. Ce travail fait aussi partie du projet « Rousset 2003-2008 » et « Nano 2012 ».

Je voudrais remercier Stéphane Dauzère-Pérès pour son accueil très convivial, ses conseils et nos échanges constructifs. À Claude Yugma pour ses encouragements et sa disponibilité. Un gros merci à Gilles Lamiable pour son ouverture, son professionnalisme et sa bonne humeur. Je garderai en mémoire nos *weekly* (très redouté), formidables moments d'échanges.

Je me suis enrichi de plusieurs cultures, par la rencontre de différentes personnes, durant ces 10 années passées en France (ou devrais-je dire en Europe?). *Chaque rencontre à sa façon fut unique.* Il est difficile de mentionner tous les noms sans courir le risque d'en oublier une grande partie ... Je m'attacherai donc seulement à indiquer que j'ai apprécié les moments passés et l'apprentissage culturel avec mes camarades, amis et amies durant mon séjour à Paris, Gardanne et Grenoble. Tout particulièrement, merci à Manuel C. pour sa bonne humeur et son ouverture d'esprit. À Carole S. pour m'avoir supporté dans notre box (n^o215AS2) à ST. Merci à tous les membres de l'équipe AMHS pour leurs disponibilités. Merci à tous ceux qui ont cru en moi et m'ont soutenu dans l'équipe SFL. Merci à Gracien C. pour sa disponibilité et à Bernard D. pour sa simplicité et pour la relecture de ce manuscrit.

Merci à Aységul S., complice et amie de tous les jours avec qui j'ai beaucoup appris. Merci du fond du coeur à Joëlla F. et Adèle L. pour avoir rendu moins difficile mes premiers jours en France. À Marie C., Farès Y., Ronan F. et Fabien M. pour leur amitié sans faille depuis que je suis en France. Merci également à Claudine et Nadège, Camille & Patrick pour leurs encouragements. Enfin, un GRAND merci à **Céline Ewencyk** compagne de tous les jours pour sa patience, son écoute et son soutien.

Merci à mon père, mes frères et soeurs pour leur soutien tout au long de mes études. Je ne saurais commencer ce mémoire de thèse sans remercier du fond de mon coeur, celle par qui tout ça est possible, ma **mère** qui m'a toujours assuré d'un soutien *indéfectible*.

« **Kiib rît medga, tì Wënd pund bïim** ».

Jean-Etienne KIBA
Gardanne le 10 Novembre 2010

*La **théorie**, c'est quand on sait tout et que rien ne fonctionne. La **pratique**, c'est quand tout fonctionne et que personne ne sait pourquoi. Ici, nous avons réuni **théorie et pratique** : rien ne fonctionne...et personne ne sait pourquoi!*

Albert Einstein

Table des matières

Introduction générale	1
I CONTEXTE, ANALYSE DE LA PROBLÉMATIQUE	5
1 Fabrication de semi-conducteurs	7
1.1 Résumé du chapitre	8
1.2 Fabrication des puces	9
1.2.1 Introduction	9
1.2.2 De l'idée au circuit	10
1.2.3 Environnement de production	10
1.2.4 Étapes de fabrication	12
1.3 Configuration du système de transport et de stockage	15
1.3.1 Introduction	15
1.3.2 Système automatique : les motivations	16
1.3.3 Layout	17
1.3.4 Véhicules	19
1.3.5 Stockage	23
1.4 Transport dans les unités de fabrication	25
1.4.1 Transport : un service pour la production	27
1.4.2 Politiques de gestion des véhicules	30

1.4.3	Gestion du stockage	33
1.5	Quelques indicateurs de performance	35
1.5.1	Introduction	35
1.5.2	Indicateurs de production	35
1.5.3	Indicateurs de transport	37
1.5.4	Indicateurs de stockage	41
1.6	Conclusion	42
2	Présentation et analyse du cas industriel	43
2.1	Résumé du chapitre	44
2.2	Fonctionnement de l'usine 300mm de Crolles	45
2.2.1	Système de transport	47
2.2.2	Processus de fabrication	53
2.3	Particularités industrielles	55
2.3.1	Système de transport unifié	55
2.3.2	Gestion des requêtes de transport par « service minimum » garanti	57
2.3.3	Repositionnement des véhicules	61
2.3.4	Regroupement des emplacements de stockage	66
2.4	Conclusion	69
3	Positionnement dans la littérature	71
3.1	Résumé du chapitre	72
3.2	Corrélation : transport et configuration	73
3.3	Corrélation : transport et production	75
3.4	Corrélation : transport et stockage	82
3.5	Notre positionnement	87
3.6	Conclusion	91

II	MODELISATION ET SIMULATION	93
4	Extension des travaux de Montoya-Torres	95
4.1	Résumé du chapitre	96
4.2	Rappels et limitations des travaux	97
4.2.1	Introduction	97
4.2.2	Rappels	99
4.2.3	Limitations	103
4.3	Extensions des travaux	106
4.3.1	Problématiques	106
4.3.2	Challenges	107
4.3.3	Extensions des travaux	108
4.3.4	Résultats	109
4.4	Conclusion	115
5	Modélisation et Simulation de Crolles 300	117
5.1	Résumé du chapitre	118
5.2	Modélisation de Crolles 300	119
5.2.1	Introduction	119
5.2.2	Définitions	121
5.2.3	Modélisation de la production	122
5.2.4	Modélisation du transport	129
5.2.5	Modélisation du stockage	134
5.2.6	Externalisation des données	138
5.3	Simulation de Crolles 300	143
5.3.1	Procédure « what next »	144
5.3.2	Procédure « where next »	145
5.3.3	Procédure « vehicle management »	147
5.3.4	Procédure « Equipment Management »	150

5.4	Validation de la simulation	152
5.4.1	Définitions et critères de validation	152
5.4.2	Validation du modèle	156
5.4.3	Données aléatoires	158
5.5	Plan d'expérience	160
5.5.1	Introduction	160
5.5.2	Complexité du nombre de tests	161
5.5.3	Etude du temps d'assignation	163
5.5.4	Étude du Carrier Exchange Time CET	165
5.5.5	Comparaison entre politique de « service minimum » et « service classique »	168
5.5.6	Difficultés de maintenir l'adéquation entre ressources de transport et besoins de production	172
5.5.7	Conclusion sur le plan d'expérience	174
5.6	Conclusion	177

III AMELIORATION DE LA POLITIQUE DE « SERVICE MINIMUM » GARANTI 179

6	Analyse du système global de fabrication 181
6.1	Prototype 183
6.2	Introduction 183
6.2.1	Motivation 184
6.2.2	Principe et algorithme 184
6.3	Etude de la politique du « service minimum » garanti 188
6.3.1	Impacts du nombre de véhicules 189
6.3.2	Impacts du nombre de lots 197
6.4	Véhicules dédiés 205
6.4.1	Introduction 205

6.4.2	Les véhicules dédiés	206
6.4.3	Résultats sur les véhicules dédiés	207
6.5	Conclusion	213
6.5.1	Conclusion	215
6.5.2	Perspectives	216
A		219
A.1	Écart type des équipements internes en fonction des véhicules.	219
A.2	Écart type des équipements externes en fonction des véhicules.	220
A.3	Écart type des équipements <i>internes</i> en fonction de la charge des lots.	221
A.4	Écart type des équipements <i>externes</i> en fonction de la charge des lots.	224
Bibliographie		234

Liste des tableaux

2.1	Paramètres techniques des véhicules SRC330.	51
2.2	Éléments d'information dont dispose un controleur local de véhicules.	63
2.3	Éléments d'information dont dispose le contrôleur de véhicules.	64
2.4	Table <i>Zone send, Zone receive</i>	64
3.1	Revue de la littérature : modélisation du transport avec fortes hypothèses simplificatrices sur la production.	76
3.2	Revue de la littérature : modélisation de la production avec fortes hypothèses simplificatrices sur le transport.	80
3.3	Revue de la littérature : sélection d'articles étudiant les stockeurs.	83
3.4	Pourcentage de publications par quadrant.	88
4.1	Résumé du niveau tactique : approche de Montoya-Torres [70]	100
4.2	Résumé du niveau opérationnel : approche de Montoya-Torres [70].	103
4.3	Temps moyen de cycle des lots.	110
4.4	Temps moyen de cycle des lots.	112
4.5	Temps moyen de chargement.	113
4.6	Pourcentage total de lots chargés en plus de 300 secondes.	114
5.1	Modélisation des équipements.	123
5.2	Répartition des lots en lithographie (lots/jour).	125
5.3	Modélisation des lots.	126

5.4	Liste des attributs des lots	127
5.5	Liste des routes modélisées.	128
5.6	Modélisation du suport de transport.	129
5.7	Liste des attributs du support de transport modélisé.	130
5.8	Listes des attributs des codes-barres.	130
5.9	Résultats de la comparaison des deux tronçons (<i>en secondes</i>).	132
5.10	Modélisation de la flotte de véhicules.	134
5.11	Modélisation d'un stockeur.	136
5.12	Liste des attributs d'un emplacement unitaire de stockage.	138
5.13	Exemple d'externalisation des informations sur le stockeur par défaut.	141
5.14	Les principales sous-procédures du « what next ».	145
5.15	Les principales sous-procédures du « where next ».	147
5.16	Les principales sous-procédures de gestion des véhicules.	149
5.17	Les principales sous-procédures de gestion des équipements.	150
5.18	Indicateurs locaux utilisés pour la validation du modèle.	158
5.19	Indicateurs globaux utilisés pour la validation du modèle.	158
5.20	Tableau des données aléatoires.	159
5.21	Valeurs du LWM et du HWM des 9 zones de Crolles 300.	163
5.22	Caractéristiques de la simulation pour la comparaison entre la politique de service « minimum » et « classique ».	170
5.23	Politique de service « minimum » vs. « classique » : temps moyen pour changer de lots [équipement interne].	171
5.24	Politique de service « minimum » vs. « classique » : temps moyen pour changer de lots sur un équipement externes de la zone étudiée.	172
5.25	Répartition de la fréquence du nombre de véhicules dans la zone étudiée en fonction du LWM.	174
6.1	Caractéristiques de la simulation pour le « service minimum ».	189
6.2	temps moyen de cycle en fonction des véhicules.	192

6.3	Écart type du temps moyen de cycle (en %) en fonction des véhicules.	192
6.4	Productivité en fonction des véhicules.	194
6.5	<i>Carrier Exchange Time</i> moyen en fonction des véhicules [équipements internes].	195
6.6	<i>Carrier Exchange Time</i> moyen en fonction des véhicules [équipements externes].	197
6.7	temps moyen de cycle en fonction de la charge des wafers en entrée.	200
6.8	Écart type du temps de cycle en fonction de la charge des wafers en entrée.	200
6.9	Productivité en fonction de la charge des wafers en entrée.	201
6.10	<i>Carrier Exchange Time</i> en fonction de la charge des wafers en entrée [équipements internes].	202
6.11	<i>Carrier Exchange Time</i> en fonction de la charge des wafers en entrée [équipements externes].	203
6.12	Plan d'expérience pour les véhicules dédiés.	207
6.13	Temps moyen de cycle et productivité avec et sans véhicules dédiés.	209
A.1	Écart type <i>Carrier Exchange Time</i> moyen en fonction des véhicules [équipements internes].	220
A.2	Écart type du <i>Carrier Exchange Time</i> moyen en fonction des véhicules [équipements externes].	221
A.3	Écart type du <i>Carrier Exchange Time</i> en fonction de la charge des wafers en entrée [équipements internes].	223
A.4	Écart type du <i>Carrier Exchange Time</i> moyen en fonction de la charge des wafers en entrée [équipements externes].	224

Table des figures

1.1	Wafers et barres de silicium.	9
1.2	Opérateurs en salle blanche.	11
1.3	Configuration <i>inter-Bay/intra-Bay</i>	18
1.4	Configuration partiellement liée.	19
1.5	Configuration unifiée.	20
1.6	Véhicule sur rails.	21
1.7	Véhicules guidés par l’humain.	22
1.8	Véhicule guidé automatiquement.	22
1.9	Véhicules de transport aérien.	23
1.10	Stockeur en 3D.	24
1.11	Lots sur les ports de chargement d’un équipement.	27
1.12	Décomposition du temps total d’opération.	28
1.13	Décomposition temps de chargement et de déchargement d’un lot.	39
1.14	Décomposition du <i>Carrier Exchange Time</i>	40
2.1	Agencement des rails de l’unité de fabrication de Crolles 300.	46
2.2	Interactions entre les éléments d’une unité de fabrication.	47
2.3	Modes d’échange d’informations entre les acteurs de l’unité de fabrication.	49
2.4	Découpage en zones de l’unité de fabrication.	50
2.5	Dépôt latéral d’un lot par un véhicule au plafond.	51

2.6	Principaux acteurs d'une zone.	52
2.7	Coexistence homme-machine.	54
2.8	Cas où la somme des LWMs dépasse le nombre de véhicules.	60
2.9	Recommandation d'avoir deux véhicules par zone.	61
2.10	Différents types de stockage unitaire.	67
2.11	Liens entre équipements, stockeurs par défaut et <i>alternates</i>	69
3.1	Différentes configurations de l'unité de fabrication.	84
3.2	Configuration 3 - configuration liée.	84
3.3	Comparaison de deux règles pour le destockage de lots dans un stockeur : FIFO <i>vs.</i> priorité des lots.	85
3.4	Classement des niveaux de détails en modélisation d'un système de fabrication et de production de semi-conducteurs, Jimenez <i>et al.</i> [37].	88
4.1	Agencement des rails de l'unité de fabrication utilisé dans [70].	98
4.2	Découpage en zones logiques de l'unité de fabrication de [70].	101
4.3	Découpage de l'unité de fabrication utilisée dans [70].	102
4.4	Illustration du parking unique utilisé dans [70].	104
4.5	Illustration de centres de zone « mal » placés.	105
4.6	Comparaison du temps moyen de cycle entre les politiques FCFS et CL-FEFS.	111
4.7	Comparaison du temps moyen de chargement d'un lot entre la politique FCFS et CL-FEFS.	113
4.8	Comparaison entre FCFS et CL-FEFS : proportion de lots chargés sur un équipement en plus de 5 minutes.	114
5.1	Illustration des équipements de production sous AutoMod.	124
5.2	Illustration d'un lot de production sous AutoMod.	126
5.3	Comparaison de deux tronçons pour un tour complet.	131
5.4	Illustration d'un véhicule aérien sous AutoMod.	133
5.5	Importance des temps d'accès au stockage.	135

5.6	Stockeur vu de profil.	135
5.7	Stockeur vu de face.	136
5.8	Organigramme de la procédure « what next ».	144
5.9	Organigramme de la procédure « where next ».	146
5.10	Organigramme de la procédure « vehicle management ».	148
5.11	Organigramme de la procédure « Equipment Management ».	152
5.12	Illustration du danger d'utiliser un seul indicateur (1/2)	154
5.13	Illustration du danger d'utiliser un seul indicateur (1/2).	155
5.14	Illustration du danger d'utiliser un seul indicateur (2/2).	156
5.15	Comparaison des temps moyen de cycle.	156
5.16	Validation : productivité des équipements.	157
5.17	Découpage de l'unité de fabrication de Crolles 300 en zones.	161
5.18	Répartition du temps d'assignation à Crolles 300.	164
5.19	Décomposition [57] du temps mis pour changer de lot sur un équipement (CET).	166
5.20	Temps moyen de changement d'un lot sur un équipement : variation en cloche.	167
5.21	Politique de service « minimum » vs. « classique » : comparaison du temps moyen de cycle et de la productivité.	171
5.22	Répartition de la fréquence du nombre de véhicules dans la zone étudiée (1/2).	173
5.23	Répartition de la fréquence du nombre de véhicules dans la zone étudiée (2/2).	173
6.1	Prototype de pré-calcul du <i>Low Water Mark</i>	186
6.2	temps moyen de cycle.	190
6.3	temps moyen de cycle.	191
6.4	Productivité en fonction du nombre de véhicules.	193
6.5	temps moyen de cycle.	199
6.6	Écart type du temps moyen de cycle.	199

6.7	Productivité en fonction de la charge des wafers en entrée.	200
6.8	Illustration de véhicules dédiés.	206
6.9	Enchainements possibles de tâches dans la zone étudiée.	208
6.10	Productivité : comparaison entre avec et sans véhicules dédiés.	209
6.11	Temps moyen de cycle avec et sans véhicules dédiés.	210
6.12	Illustration des futurs rails.	211

INTRODUCTION GÉNÉRALE

Présentation du problème Pour fabriquer des composants micro-électroniques deux grandes étapes sont nécessaires. Le *Front-End* qui rassemble les étapes de fabrication des éléments de base de la puce et, le *Back-End* celui des interconnexions entre les différents éléments de base, puis l'encapsulation de la puce. Dans cette thèse, seule l'étape *Front-End* sera considérée.

Les processus de fabrication des composants micro-électroniques sont très complexes. Le processus réentrant, c'est-à-dire la possibilité pour un lot de visiter plusieurs fois le même équipement, entraîne un grand nombre d'étapes (jusqu'à 700) de fabrication. De plus cette fabrication exige un environnement de fabrication ultra propre (cent fois plus propre qu'une salle chirurgicale) et coûtent très chers, jusqu'à 70 millions d'euros pour un scanner en lithographie.

L'objectif global d'une unité de fabrication de semi-conducteur est de faire un produit le plus rapidement possible, à moindre coût et en grande quantité.

La compétition entre fabricants de semi-conducteurs est d'autant plus rude, qu'elle accentue le dilemme entre des décisions contradictoires qui existent au sein des entreprises. Par exemple, pour un commercial, un produit de qualité avec un délai de fabrication court se vend très bien. Tandis que pour la production, garantir la qualité exige des délais de fabrication plus longs et donc coûtera plus cher.

À ces contraintes décisionnelles, il faut ajouter des contraintes techniques qui, de notre point de vue, constituent une source de motivations pour l'étude des problèmes liés au transport dans les unités de fabrication des semi-conducteurs :

- *Retour sur investissements.* Pour amortir le coût très élevé des équipements, il faut faire passer le plus de lots possible sur l'équipement, donc réduire les temps d'enchaînement des tâches sur les équipements. Cette politique s'appelle le

continuous processing. Elle est assurée par les opérateurs en salle blanche dans les usines non automatisées, et de plus en plus par des *véhicules automatisés* dans les usines avancées.

- *Attentes des lots*. Si cela était possible, on mettrait les équipements de production les uns à côté des autres, cela favoriserait le *continuous processing*. Mais, en pratique, ce n'est pas possible pour diverses raisons (maintenance du park équipements, etc.). Ainsi, l'agencement des équipements de production induit que l'on accepte l'attente des lots. Le *transport automatisé* est un moyen de réduire cette attente.
- *Allocation des ressources*. Très souvent, il y a plus d'équipements de production que de ressources de transport (opérateurs ou véhicules). Donc si tous les équipements de production émettent en même temps une requête de transport, elles ne pourront pas toutes être satisfaites. L'activité de production est déséquilibrée, il faut donc la rééquilibrer et le système de *transport automatisé* le permet.

Objectifs de la thèse Dans cette thèse, nous nous intéressons au système de transport automatisé de lots dans les unités de fabrication de semi-conducteurs.

Les travaux précédents dans la littérature des semi-conducteurs, plus particulièrement sur le transport automatisé, se sont focalisés sur les règles de priorités des véhicules (quels véhicules choisir et pour quelle tâche?) et sur des méthodes de recherche de trajets réalisables et libres de tout conflit (routage). Cependant, ces règles de priorités et méthodes de recherche de trajets sont dépendantes du type de produits, de la configuration du réseau de transport, du nombre de véhicules, etc. De plus il a été montré qu'il n'y avait pas une règle surpassant toutes les autres [20].

De plus, les approches utilisées sont le plus souvent analytiques ou par simulation à événements discrets. Pour ces dernières, la complexité des modèles fait que le temps de développement est très long. Ceci a pour conséquence le développement de modèles « agrégés », permettant une compréhension globale du système. Mais dès que l'on veut faire une étude plus fine (estimation du temps d'assignation des véhicules par exemple), les modèles agrégés ne suffisent plus [74].

L'objectif de cette thèse est de déterminer par rapport à l'existant, de nouvelles

conditions nécessaires à la mise en adéquation entre ressources de transport et besoins de la production. Nous utiliserons une approche par simulation à événements discrets car, la complexité du problème et le niveau de détails recherchés ne pourraient être capturés par un modèle analytique « simple ». Nous simulerons le plus fidèlement possible les trois aspects entrant en compte dans les unités de fabrication ; le transport, la production et le stockage.

Ce travail de thèse a été réalisé dans le cadre d'une convention CIFRE¹. Ces travaux permettent d'affiner les connaissances théoriques sur la problématique du transport dans les unités de fabrication des semi-conducteurs. Ils contribuent aussi à la compréhension du système industriel (peu documenté par le fournisseur) et au développement de nouvelles solutions. En outre, il y a augmentation du niveau de confiance entre industriel et académique par le partage commun d'un prototype fidèle, compris, ajusté et validé par les deux parties.

Plan de lecture Cette thèse est divisée en trois parties.

- La première partie est composée de 3 chapitres. Le chapitre 1 présente et analyse les problèmes sur les systèmes de transport automatisés. Le chapitre 2 décrit plus précisément notre problématique en analysant le cas industriel. Il nous permettra de modéliser les contraintes industrielles. Le chapitre 3 présente notre positionnement par rapport à la littérature, ce qui nous permettra d'affiner notre sujet de thèse.
- La seconde partie se compose du chapitre 4 sur les extensions des travaux de Montoya-Torrès [70]². Le chapitre 5 présente la modélisation et la simulation de l'unité de fabrication de Crolles 300mm.
- La troisième partie qui correspond également au dernier chapitre, traite de l'analyse et du calcul des valeurs des paramètres critiques de l'unité de fabrication. Les conclusions de ce travail de thèse ainsi que les perspectives de recherche sont présentées à la fin du mémoire.

¹Conventions Industrielles de Formation par la REcherche

²Étudiées en première année de thèse et avant de démarer la thèse en convention CIFRE

Première partie

CONTEXTE, ANALYSE DE LA PROBLÉMATIQUE

Chapitre 1

Fabrication de semi-conducteurs

Ce chapitre traitera :

- de la fabrication des puces (section 1.2),
- de la configuration du système de transport (section 1.3),
- du transport dans les unités de fabrication (section 1.4) et,
- des indicateurs de performances de l'unité de fabrication (section 1.5).

Nous répondrons aux deux questions suivantes : **que fait le système global de production, de transport et de stockage et, comment mesure-t-on ses performances ?**

1.1 Résumé du chapitre

Ce premier chapitre présente l'environnement des semi-conducteurs. Les différentes étapes et les conditions de fabrication sont abordées de manière à fixer les idées de bases.

La caractéristique particulière de la fabrication des semi-conducteurs est le phénomène réentrant, c'est-à-dire le fait de visiter plusieurs fois le même équipement. Cela est dû au fait qu'il y a plus d'étapes de fabrication que de machines. L'enchaînement de ces différentes étapes est réalisé par des opérateurs ou des systèmes de transport automatique. La performance de ces systèmes dépend majoritairement de la configuration de l'unité de fabrication (*layout*) et des politiques de transports utilisées.

L'introduction des systèmes de transport automatique est le résultat de l'augmentation du diamètre des wafers (loi de Moore [73]) et donc de leurs poids. La taille des containers dans lesquels sont transportés les wafers augmente aussi et cette augmentation entraîne l'apparition de problèmes ergonomiques et de risques musculosqueletiques pour les opérateurs devant les transporter.

Les principales qualités de tous systèmes automatiques sont la *répétabilité* ou la capacité du système à reproduire la même chose et, la *fiabilité* ou la capacité du système à posséder une probabilité de bon fonctionnement pendant un temps donné. Maîtriser ces deux qualités est le challenge de tous les managers des unités de production.

Enfin, il est très important de pouvoir évaluer *sans ambiguïté* les performances du système dans l'objectif de l'améliorer. C'est l'objet de la dernière section. Mais remarquons que dans [94] les auteurs suggère que « changer de point de vue » (point de vue équipements, opérateurs ou véhicules) permettrait d'améliorer les performances des unités de fabrication de semi-conducteurs.

1.2 Fabrication des puces

Un microprocesseur est obsolète dès le début de sa fabrication en série.

Loi de Murphy, 1947

1.2.1 Introduction



FIGURE 1.1 – Wafers et barres de silicium.

La figure 1.1 est une photo de semi-conducteurs utilisés comme élément de base en microélectronique : le silicium. Un semi-conducteur est un matériau qui a les caractéristiques électriques d'un isolant, mais pour lequel la probabilité qu'un électron puisse contribuer à un courant électrique, quoique faible, est suffisamment importante. Cette propriété est à la base du fonctionnement des composants de l'électronique moderne : diodes, transistors, etc. Issue de l'anglais, *transconductance varistor* (résistance variable de transconductance), le transistor est un composant électronique actif fondamental en électronique, utilisé principalement comme interrupteur commandé. C'est en 1947, que Bell Labs produisirent le premier transistor et sont à l'origine de la naissance de l'industrie des semi-conducteurs [31]. Cofondateur de la société Intel, Gordon Moore avait affirmé dès 1965 que le nombre de transistors

par circuit de même taille allait doubler, à *prix constant*, tous les ans. Il rectifia par la suite en portant à dix-huit mois le rythme de doublement. Il a cependant déclaré en 1997 que cette croissance des performances des puces se heurterait aux environs de 2017 à la limite physique : celle de la taille des atomes. La loi de Moore n'a pas encore été remise en cause, mais les contraintes de miniaturisation la rendent de plus en plus difficile à respecter. Elle a également un impact économique très important : le prix d'un mégabit de mémoire électronique coûtait 75000 euros en 1973 contre 1 centime d'euro en 2008 [83].

Constitution d'un semi-conducteur : un dispositif semi-conducteur est constitué de deux parties. La première est un coeur de silicium, un carré de quelques millimètres de côté, fragile et très mince : **la puce** ou *chips*. La seconde est un **boîtier** ou (*package*) qui le protège de son environnement externe et qui en facilite le montage dans les systèmes finis.

1.2.2 De l'idée au circuit

La puce, comme tout autre produit doit d'abord être conçue. Il est important que le concepteur ait une excellente connaissance du système dans lequel le circuit devra être inséré. Cela permet de mieux caractériser les systèmes (TV, automobiles, ordinateurs...). Les puces sont conçues à l'aide des instruments CAD/CAE¹, qui permettent de définir leur structure logique et fonctionnelle, de simuler leur fonctionnement et de définir leur structure physique. A la fin de cette phase de conception, le système CAD/CAE génère les masques utilisés pour la première partie du cycle de production : le *Front-End*. A ce stade, le résultat fini est une puce qui, testée, est prête à être introduite dans le boîtier.

1.2.3 Environnement de production

L'élément de base pour le démarrage de la production c'est la plaquette de silicium ou *wafers* : une tranche de forme presque circulaire, de quelques millimètres

1. CAD signifie Computer-Aided-Design : Ensemble d'outils et de programmes informatiques permettant d'assister l'ingénieur dans la conception d'un produit.

d'épaisseur. On obtient les puces à partir des plaquettes. Les circuits élémentaires qui réalisent la fonction (analogique ou numérique) à laquelle est destinée la puce sont gravés sur les plaquettes selon des techniques et des installations sophistiquées pour graver des traits de l'ordre du micron (millionième de mètre). Pour graver des traits aussi fins, les appareils de production doivent être installés dans un environnement ultra stable, sans vibrations et ultra propre. Dans les *process*² avec des géométries submicroniques, c'est-à-dire avec des traits inférieurs au micron, une simple particule de poussière sur une plaquette peut endommager complètement le circuit. C'est pourquoi la production est réalisée dans un environnement où la poussière est rigoureusement évitée : la *salle blanche* (figure 1.2). Les opérateurs doivent revêtir blouses, chaussures et coiffes spéciales qui limitent au maximum la contamination involontaire.



FIGURE 1.2 – Opérateurs en salle blanche.

2. Ensemble des étapes ou transformations nécessaires à la fabrication d'un produit. Ils peuvent être manuels, mécanisés voire complètement automatisés. Généralement spécifique à chaque entreprise et couvert par le secret de fabrication.

1.2.4 Étapes de fabrication

Les étapes de fabrication peuvent se diviser en deux grandes parties : le *Front-End* et le *Back-End*. Le *Front-End* regroupe les étapes de fabrication des éléments de base de la puce tandis que le *Back-End* regroupe celui des interconnexions des éléments de bases.

Cette thèse s'intéresse aux *Front-End*, le *Back-End* n'étant pratiquement plus fait en Europe.

***Front-End* (de la plaquette à la puce) :** le processus de production *Front-End* peut varier en fonction du type de technologie et de circuit. Il existe quelques étapes élémentaires de *process* communs à la plupart des technologies :

- ***Oxydation*** : la plaquette est oxydée sur toute sa surface. Pour cette opération, on utilise des fours spécifiques.
- ***Dépôt de résine*** : sur la couche d'oxyde on dépose de façon uniforme une couche de résine photo sensible qui se transforme sous l'action de lumière.
- ***Photolithographie*** : à ce stade entrent en jeu les masques, c'est-à-dire le résultat des phases de conception. Un masque est aligné sur la plaquette et le tout est exposé à une source de lumière. La résine « s'impressionne », comme une pellicule photographique normale, dans les zones laissées libres par le masque.
- ***Développement*** : après avoir « photographié » le masque sur la résine on procède au développement de façon analogue à ce qui se produit dans un processus photographique habituel : la résine qui a pris la lumière est enlevée.
- ***Gravure*** : en appliquant des acides ou des gaz déterminés, on enlève l'oxyde laissé libre par la résine, sans attaquer le silicium de départ.
- ***Dopage*** : il s'agit maintenant d'introduire, dans les zones exposées, des éléments chimiques pour modifier les caractéristiques du silicium, et le rendre capable de conduire les signaux électriques. Ces éléments (dopants) sont introduits dans la structure cristalline du silicium par diffusion ou implantation ionique.

Le cycle décrit ci-dessus se répète plusieurs fois : autant de fois qu'il y a des masques. Les technologies les plus complexes demandent plus de vingt masques. Le résultat final est une plaquette sur laquelle sont dessinées les puces, normalement identiques.

Les puces sont prêtes mais ... fonctionneront-elles correctement ? Pour le vérifier, la plaquette est soumise à un test électrique, dit EWS (*Electrical Wafer Sort*), durant lequel on contrôle les paramètres fonctionnels de chaque puce et on identifie les rebuts. Le processus de production *Front-End* terminé, la plaquette est prête pour être envoyée à la phase suivante : le *Back-End*.

Back-End (de la puce au produit fini) : l'objectif du processus de *Back-End* est d'encapsuler les puces qui ont passé le test EWS³ dans un boîtier approprié, d'effectuer le contrôle final et de produire un circuit prêt à la vente. Le boîtier a beaucoup d'importance. Il doit protéger le silicium, fournir un support mécanique adapté, fournir la connexion électrique avec le monde extérieur et favoriser la dispersion de chaleur générée pendant le fonctionnement. Il existe un nombre exponentiel de types de boîtiers. Les différentes étapes du *Back-End* sont les suivantes :

- ***Découpe*** : la plaquette contrôlée est placée sur un support adhésif flexible. Les puces sont séparées les unes des autres au cours d'une opération mécanique de sciage, effectuée avec un appareil de haute précision.
- ***Collage*** : les bonnes puces sont prélevées et soudées sur une structure métallique : la grille. Cette grille permettra de conduire à l'extérieur les signaux électriques, en donnant la possibilité au dispositif de réagir avec le reste du système dans lequel il sera introduit.
- ***Câblage*** : la connexion entre les points d'entrée et de sortie des signaux électriques de la puce et la grille (*frame*) se fait par l'intermédiaire de fils très fins, normalement en or, ancrés sur des zones précises de la puce et reliés aux extrémités de la grille.
- ***Encapsulage, découpe et mise en forme*** : l'ensemble est prêt pour être introduit dans une enveloppe protectrice de résine, céramique ou autre matière, d'où sortent les extrémités de la grille qui deviennent les *pieds* du circuit.
- ***Marquage et test*** : nous sommes à la fin du processus de *Back-End*. Le circuit est testé et sur son corps on estampille un sigle qui, outre le logo du producteur, identifie aussi ses fonctions et ses prestations ainsi que d'autres indications éventuelles.

3. *Electrical Wafer Sort* ou test électrique des wafers afin d'écarter les wafers les moins performants.

À présent nous savons comment sont fabriqués les puces, nous allons maintenant voir comment sont agencés les équipements de production et comment les supports de fabrication des puces (wafers) sont transportés d'une étape à l'autre.

1.3 Configuration du système de transport et de stockage

« *Tout espace est défini par les réseaux qui le desservent et l'organisent.* »

LEPETIT B. 1998

1.3.1 Introduction

La fabrication d'une « puce » requiert environ 600 étapes⁴ qui ne se déroulent pas physiquement toutes au même endroit. Par exemple, après une insolation dans la zone de photolithographie, les lots subissent généralement une inspection dans une autre zone, celle de métrologie. Ce déplacement physique du lot d'une zone à une autre est un transport. De manière plus générale, un transport est le passage d'un lot d'une étape à une autre et il doit garantir l'intégrité physique du lot.

Par le passé, le transport des lots était habituellement assuré par des opérateurs. Mais de nos jours, avec l'augmentation du diamètre des wafers, les lots sont plus lourds à transporter et les distances plus grandes à parcourir. Il y a alors augmentation des risques de troubles musculosquelettiques⁵ des opérateurs.

Les systèmes de transports *automatiques* se présentent comme une excellente alternative à l'humain, non pas seulement à cause de la suppression des risques musculosquelettiques pour les opérateurs, mais parce que tout système automatique doit posséder en théorie au moins deux importantes propriétés : celle de « fiabilité », relative à la confiance dans le système et celle de « répétabilité », relative à la capacité du système à refaire la même chose.

Dans le cas du transport, un système fiable doit pouvoir *garantir un nombre de transports par jour*, tandis qu'un système répétable doit pouvoir *garantir un temps de transport* par exemple, un temps de livraison pour 95% des lots.

Outre ces propriétés, il est important de souligner que la configuration du support de transport, c'est-à-dire la disposition des rails en lignes droites ou en courbes, influence également ses performances. Par exemple, un système ayant peu de lignes

4. Une étape de fabrication correspond au passage d'un lot d'une machine à une autre

5. Trouble affectant les muscles et les os, souvent provoqué par des mouvements répétitifs.

droites sera moins performant en terme de temps de transport. Nous confirmerons cette affirmation dans le chapitre consacré au plan d'expériences.

La suite de cette section sera découpée de la manière suivante. Nous rappellerons les principales motivations ayant conduit à l'introduction des systèmes de transport automatique dans les unités de fabrication de semi-conducteurs dans la section 1.3.2. On répondra ainsi à la question : qu'est que l'on attend du système? Nous présenterons ensuite dans la section 1.3.3 les différents types de configurations de systèmes de transport, en insistant sur leurs avantages et inconvénients. La section 1.3.4 décrira les différents types de véhicules, on identifiera en particulier les véhicules considérés dans cette thèse. Différents types de stockages seront présentés dans la sous-section 1.3.5 car, même s'ils sont peu traités, dans la littérature, ils font partie intégrante des systèmes de transport.

1.3.2 Système automatique : les motivations

Il y a deux motivations majeures qui justifient l'introduction des systèmes automatiques dans les usines de semi-conducteurs : des motivations humaines et des motivations financières.

Motivations humaines

Les wafers sont transportés par lots dans des cassettes appelées FOUP⁶. Un FOUP comprend au maximum 25 wafers. L'augmentation du diamètre des wafers rend plus lourd le transport de FOUPs par les opérateurs. En effet, un FOUP vide de wafers 300mm pèse environ 4.5 kg contre 7.7 kg quand il est plein. Transporter environ 8kg⁷ de manière répétée augmente les risques d'accident musculosquelettique des opérateurs. Les systèmes automatiques répondent à ce problème car les lots sont transportés par des véhicules qui assurent l'intégrité physique des lots.

6. Front Opening Unified Pod

7. contre 4.5 kg pour les FOUPs 200mm.

Motivations financières

Le coût par plaquette augmente avec l'augmentation du diamètre des wafers, car il y a en moyenne 1,5 fois plus de puces à chaque augmentation, et jusqu'à 2,5 fois plus quand on passe de 200 à 300mm. Un wafer 300mm en fin de parcours peut valoir entre \$ 3500 et \$ 4500 selon le prix du marché. Ainsi, plus on produit de lots, plus on rentabilise son investissement. Il faut donc un système fiable, c'est-à-dire présentant la variabilité la plus petite possible et donc, une continuité dans l'enchaînement des tâches sur les équipements de production. Les systèmes automatiques de transport répondent à ce problème.

1.3.3 Layout

Le layout constitue le « squelette » de la fab. Il représente la manière dont sont agencés les équipements de production et de stockage.

La disposition des équipements a une grande importance sur les performances de la fab [7]. Un problème de base complexe est connu dans la littérature sous le nom de *Quadratic Assignment Problem* (QAP) [56]. Il existe différentes approches pour déterminer une solution optimale de ce problème : par exemple [29] est basé sur la technique du *branch and bound* et [1] sur des méthodes de linéarisation. Mais il a été prouvé que le QAP est un problème NP-difficile⁸, donc toutes ces méthodes peuvent difficilement résoudre des problèmes de très grande taille [82] c'est-à-dire qui ont un très grand nombre d'instances ($n > 30$).

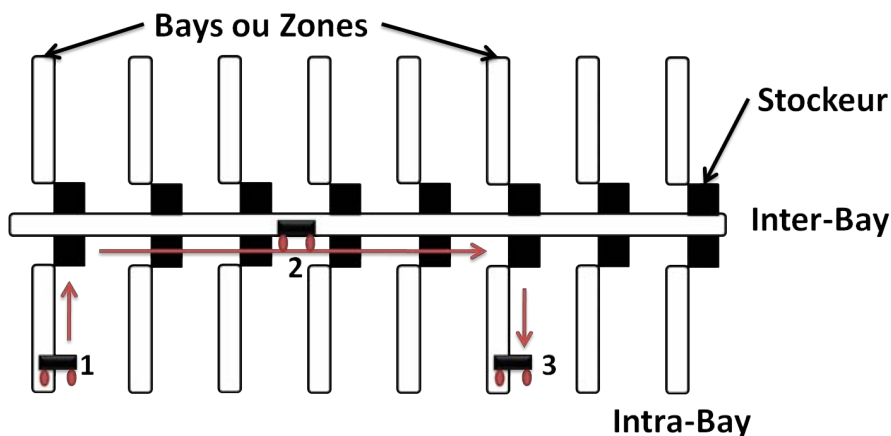
Dans la pratique il existe deux approches pour agencer des équipements dans les unités de fabrication de semi-conducteurs [80] : la première consiste à positionner les équipements selon la séquence des étapes de process. Cette approche minimise à l'évidence les temps entre étapes successives. La seconde approche, la plus utilisée dans l'industrie, consiste à agencer les équipements en fonction de leur type afin de faciliter les opérations de maintenance.

8. C'est-à-dire qu'il n'y a pas d'algorithme connu pouvant résoudre ce problème en un temps polynomial, et même des petites instances requiert un temps de résolution long

Les différents types de configuration

L'unité de fabrication est découpée en ateliers répartis dans différentes zones, c'est-à-dire que des équipements d'un même atelier peuvent être physiquement dans des zones différentes. Ce regroupement en zone permet de distinguer deux types de transports : le transport *inter-bay*, c'est-à-dire d'une zone à l'autre et le transport *intra-bay*, c'est-à-dire à l'intérieur de la même zone. En se basant sur ces deux types de transports, les auteurs dans [7] ont relevé dans la littérature trois configurations possibles que nous reprenons ici :

- **Configuration inter-bay/intra-bay** [*non unified AMHS*] (figure 1.3). Dans cette configuration toutes les zones sont séparées les unes des autres et le système comporte généralement un couloir central. Cette configuration ne permet pas à un véhicule de passer d'une zone à l'autre. Ce sont les stockeurs qui servent d'interface d'échange entre les zones. Un lot doit utiliser au moins trois systèmes de transport distincts pour passer d'une zone à l'autre : le premier à l'intérieur de la zone pour amener le lot vers le stockeur qui est le point d'échange avec le couloir central. Le second dans le couloir central en direction du stockeur servant de point d'échange avec la zone de destination et, le troisième système de transport permettant d'amener le lot à sa destination.



Les véhicules se déplacent uniquement à l'intérieur de chaque zones (bays).
Il faut 3 systèmes de transport différents pour aller d'une zone à une autre.

FIGURE 1.3 – Configuration *inter-Bay/intra-Bay*.

1.3. CONFIGURATION DU SYSTÈME DE TRANSPORT ET DE STOCKAGE¹⁹

D'après [7] et sous les mêmes conditions, la configuration *inter-bay/intra-bay* est celle qui utilisera le plus de stockeurs par rapport à toutes les autres,

- **Configuration partiellement liée** [*linked bays*] (figure 1.4). Dans cette configuration certaines zones peuvent communiquer directement entre elles sans passer par un intermédiaire (couloir central ou stockeurs). Cela permet de gagner une étape de transport par rapport à la configuration précédente,
- **Configuration unifiée** [*unified AMHS*] (figure 1.5). Les véhicules peuvent se déplacer partout dans la salle blanche, c'est-à-dire d'une zone à une autre. Elle ne requiert donc qu'une seule étape de transport pour aller d'un équipement à un autre, même si les équipements sont situés dans différentes zones. Toujours d'après [7], c'est cette configuration qui demande le moins de stockeurs et est meilleure, en terme de performance de transport, que la configuration *inter-bay/intra-bay*.

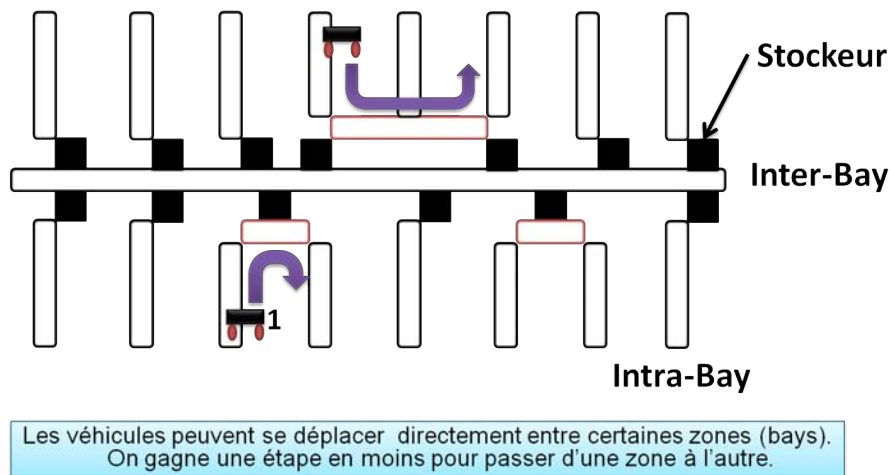


FIGURE 1.4 – Configuration partiellement liée.

1.3.4 Véhicules

Les véhicules ont tous la même fonction, celle de transporter des entités d'un point à un autre. Mais, en fonction de la situation, certains types de véhicules sont mieux adaptés que d'autres. Par exemple l'avion est mieux adapté pour le transport

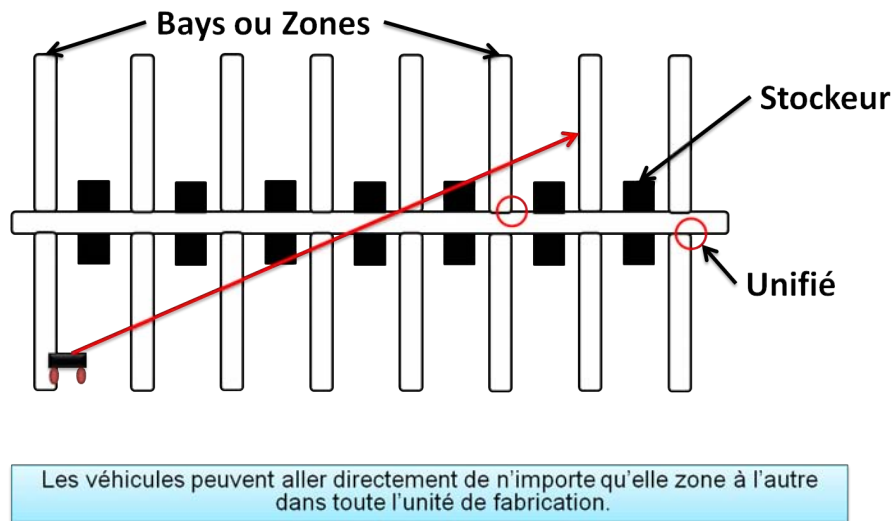


FIGURE 1.5 – Configuration unifiée.

intercontinental que le train. De même, en fonction du type de zone (*inter-bay* ou *intra-bay*), il existe différents types de véhicules.

Les systèmes de transport dans les unités de fabrication de semi-conducteurs sont divisés de manière générale en système de transport *intra-bay* et en système de transport *inter-bay* [48]. Bien que cette distinction perde de son sens avec les systèmes unifiés, elle continue d'être utilisée car elle facilite l'analyse des systèmes de transport. Elle permet par exemple aux auteurs dans [60] de classer les véhicules en quatre types dans les systèmes de type *intra-bay* et en un seul type dans les systèmes *inter-bay*.

Véhicules dans les systèmes *intra-bay*

- **Véhicules sur rails** [*Rail Guided Vehicles*] (figure 1.6, source MURATEC) : ce sont des véhicules utilisés en général pour le transport des produits dans les industries. Ils se déplacent souvent sur les rails et sont indépendant de l'humain. La particularité de ces véhicules est qu'ils sont utilisés pour de court trajet (entre 2 équipements en général) mais ils ne couvrent jamais toute l'unité de fabrication.
- **Véhicules guidés par l'humain** [*Personnal Guided Vehicles*] (figure 1.7, source

1.3. CONFIGURATION DU SYSTÈME DE TRANSPORT ET DE STOCKAGE²¹

DAIFUKU) : ce sont des véhicules, généralement à roulette et complètement dirigés par l'humain. Ils sont utilisés par les opérateurs en salle blanche pour acheminer les lots particulièrement aux endroits où il n'y a pas de système automatique de transport.

- **Véhicules guidés automatiquement** [*Automated Guided Vehicles*] (figure 1.8, source SPACECAMP) : ces véhicules se déplacent également au sol mais sont complètement indépendant de l'humain.
- **Véhicules de transport aérien** [*OverHead Transport Vehicles*] (figure 1.9, source MURATEC) : dans cette configuration les véhicules se déplacent au plafond. On gagne ainsi de la place au sol, et cela permet d'avoir plus d'équipements de production à l'intérieur de l'unité de fabrication.



FIGURE 1.6 – Véhicule sur rails.

Véhicules dans les systèmes *inter-bay*

Dans les systèmes *inter-bay*, on dénote un seul type de véhicules : les *Overhead Shuttle Vehicles*. Ce sont des véhicules aériens permettant le transfert de lots d'une zone à une zone dans les systèmes non-unifiés. Avec l'arrivée des systèmes unifiés,



(a) Véhicule guidé par l'humain.



(b) Opérateur et véhicule guidé par l'humain.

FIGURE 1.7 – Véhicules guidés par l'humain.



FIGURE 1.8 – Véhicule guidé automatiquement.

plus efficaces que les systèmes partiellement liés [7], les *Overhead Shuttle Vehicles* initialement utilisés pour le transport *intra-bay* ont été remplacés par des *OverHead transport Vehicles* (figure 1.9). Ces véhicules sont capables d'effectuer à la fois du

1.3. CONFIGURATION DU SYSTÈME DE TRANSPORT ET DE STOCKAGE²³

transport *inter-bay* et *intra-bay*.



FIGURE 1.9 – Véhicules de transport aérien.

1.3.5 Stockage

Si on pouvait, on mettrait les équipements de production les un à côté des autres en fonction de la séquence des étapes de process. Mais pour divers raisons (facilitation de la maintenance, réduction des trajets, etc ...), les équipements sont placés souvent « loin » les uns des autres : cela signifie que l'on accepte l'attente, il faut donc la gérer. Ainsi, lorsqu'il y a inadéquation temporelle entre la fin d'une tâche et le début d'une autre tâche sur un même équipement, il est nécessaire de stocker le lot en attendant que l'équipement soit disponible.

Dans le domaine des semi-conducteurs, outre les raisons d'implantation des équipements de production, d'autres facteurs favorisent cette inadéquation temporelle à savoir : la variabilité des temps de traitements des lots par les équipements, le temps de disponibilité de l'équipement qui est fonction du temps de maintenance, etc.

De plus, avec le flux réentrant, c'est-à-dire le fait pour un produit de visiter plusieurs fois le même équipement, les lots de production à différents stades de leur fabrication sont très souvent en concurrence pour le même équipement. Il se forme alors une file d'attente et il devient très difficile de déterminer l'ordre de passage des lots sur l'équipement. Les lots en attentes sont stockés en attendant leur tour.

Les systèmes de stockage peuvent être classés selon deux critères : **capacité et temps d'accès**. De manière générale, la capacité des emplacements de stockage est

inversement proportionnelle au temps d'accès, c'est-à-dire que, plus le temps d'accès est rapide, plus l'équipement de stockage est de petite capacité et vice-versa.

Les stockeurs

Un stockeur (figure 1.10, source [17]) est formé d'un robot, sorte de bras mécanique, et d'étagères de stockage, le tout enfermé dans un grand container [17]. Il possède une grande capacité de stockage, c'est-à-dire une centaines d'emplacement unitaire, mais un très long temps d'accès. Cela est dû au fait que le robot à l'intérieur du stockeur possède son propre temps de cycle et ne peut déplacer qu'un seul produit à la fois.

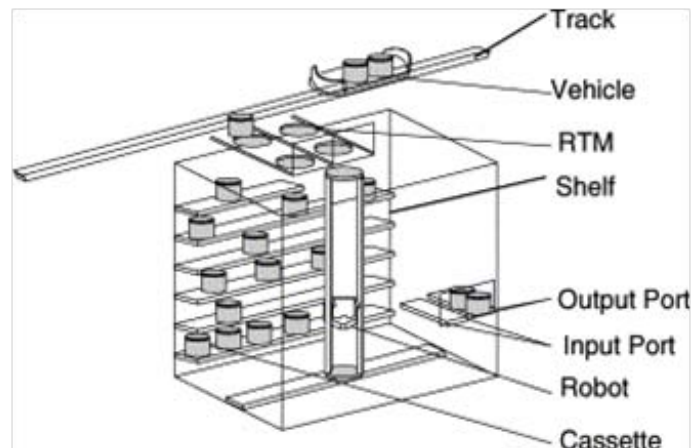


FIGURE 1.10 – Stockeur en 3D.

À cause de son long temps d'accès, le stockeur peut donc devenir un « goulot d'étranglement » ou *bottleneck*. Il faut donc réduire son utilisation au minimum car il peut être limitant pour un système, particulièrement en cas de panne. Nous verrons par la suite comment y remédier dans le chapitre sur le cas industriel.

1.4 Transport dans les unités de fabrication de semi-conducteurs

L'imagination est la meilleure compagnie de transport au monde.

Roger Fournier, 1965

Motivations.

La « vie » d'un lot est une répétition de trois états : en attente, en traitement, et en transport. Les états d'attente et de transport sont considérés comme étant des *temps de non-production* car le wafer ne subit aucun traitement particulier. On dit alors que le transport n'apporte pas de valeur ajoutée.

Cependant, même si le transport des lots n'apporte pas de valeur ajoutée, elles y contribuent car, elle permet la continuité de l'enchaînement des tâches de production sur les équipements⁹. Il y a alors augmentation de la productivité des équipements et donc de la productivité globale. Ainsi, intégrer les contraintes de production dans la gestion du transport c'est non seulement avoir une représentation la plus fidèle possible du système global [55], mais c'est surtout réunir toutes les conditions nécessaires à l'amélioration de la production. *Etudier finement le transport nécessite d'étudier la production.* Pour ce faire, il est nécessaire de comprendre le fonctionnement des scénarios de transport.

Les scénarios de transport des lots.

L'objectif globale d'une unité de fabrication est de faire un produit le plus rapidement possible, avec le moins d'équipements et en quantité énorme. Les systèmes de transport automatique contribuent à la réalisation de cet objectif. Cependant, la performance des systèmes de transport automatique, même de très hautes technologies, dépend fortement de l'adéquation entre les scénarios de transport implémentés dans leurs logiciels et les besoins réels de la production. Par exemple, un système de transport qui traite tous les lots avec la même priorité ne prend pas en compte les

9. permettant ainsi un retour sur investissement plus efficient

besoins de la production, qui sont eux tournés vers les exigences clients (respect des délais). Puisque le transport est un service pour la production, il est nécessaire que cette adéquation se face en fonction des besoins de la production.

Point de vue opérateur. L'opérateur en salle blanche se soucie très peu du transport pure. Ce qu'il recherche, c'est que le lot soit présent au moment où il en a besoin ou, que le lot libère l'équipement au moment où il veut en placer un autre sur le port de chargement. Ce point de vue montre deux points importants au niveau transport : le chargement et le déchargement des lots.

Point de vue équipement. L'objectif des équipements de production est de traiter le plus rapidement possible le maximum de lots. Cela signifie pour un point de vue équipement que l'objectif principal est d'avoir un temps à vide le plus petit possible, voir nul. Pour ce faire, elle attend du système de transport que le temps de remplacement d'un lot par un autre soit le plus court possible. Cet point de vue montre également deux points importants : le chargement et le déchargement des lots.

Point de vue transport. Le transport « ne sait » pas qu'il existe des lots prioritaires, il répond à une requête de transport. De même, le transport « ne sait pas » qu'il existe des équipements prioritaires, c'est-à-dire des équipements dont le temps de cycle est très court (de l'ordre de 4 mn), il sert plus particulièrement ces équipements si on le lui demande explicitement.

Pour être en adéquation avec les exigences de la production (priorité des lots, des équipements, etc.), le transport a besoin de scénarios de chargement et de déchargement adaptés. On remarque que quelque soit le point de vue, deux éléments importants reviennent dans le transport des unités de fabrication des semi-conducteurs : le chargement et le déchargement des lots. De manière générale, tous transport dans une unité de fabrication de semi-conducteurs peut se décomposer en deux parties :

1. le déchargement des lots, dont l'objectif principal est de rapprocher le lot au plus près des équipements de production de la prochaine étape. Il permet également de libérer plus rapidement le port de chargement,

2. le chargement des lots, dont l'objectif principal est de charger un lot le plus rapidement possible sur le port de chargement (figure 1.11) de la machine afin de réduire le temps de non-production.

Le scénario de déchargement des lots répond à la question : *quel est l'emplacement de stockage le plus proche ?* Tandis que celui du chargement des lots répond à la question : *quel lot choisir ?*



FIGURE 1.11 – Lots sur les ports de chargement d'un équipement.

Très peu d'articles traitent du couplage fin du transport et de la production dans la littérature sur la simulation des unités de fabrication de semi-conducteurs, car jugé trop coûteuse en temps de conception et de calcul, en particulier quand l'horizon de décisions est à moyen ou court terme.

1.4.1 Transport : un service pour la production

Motivations.

L'objectif de la production est de faire passer le plus de lots possibles (mais aussi les bons) sur les équipements de production. Pour ce faire, il est nécessaire de réduire le temps de non-production des équipements. Il faut pour cela comprendre de quoi est composé le temps total opératoire d'un équipement. On peut se rendre

compte (voir figure 1.12) que ce temps total opératoire comporte un temps effectif de production de l'équipement appelé *OEE*¹⁰ pour *Overall Equipment Effectiveness*.

On s'aperçoit (figure 1.12), que le challenge pour la production est d'améliorer le temps de l'OEE afin de rendre les équipements plus efficaces. Le transport, considéré comme sans valeur ajoutée, **ne doit en aucun** cas être un élément limitant. Cela signifie que, si la production arrive à augmenter le temps de fonctionnement effectif de l'équipement, le transport ne doit pas contribuer à diminuer ce temps de fonctionnement effectif. En effet, ne pas amener le bon produit **au bon moment** réduit considérablement le temps de fonctionnement effectif de l'équipement.

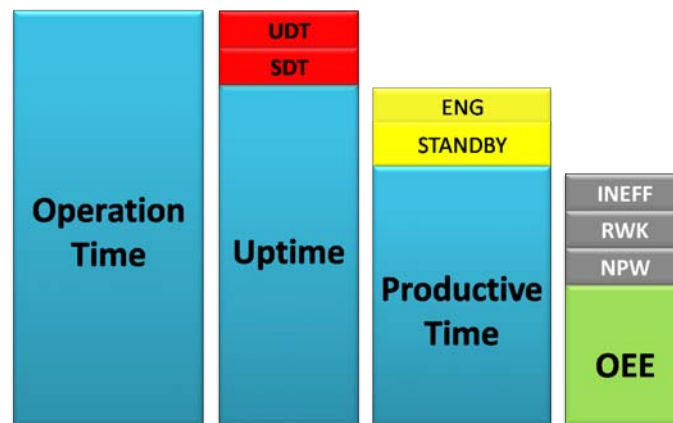


FIGURE 1.12 – Décomposition du temps total d'opération.

Le temps de fonctionnement effectif d'un équipement ou OEE

La décomposition du temps total d'une opération de production ou *Operation Time* comporte les éléments suivants :

1. Le temps pendant lequel l'équipement est en marche ou *uptime*,
2. Le temps pendant lequel l'équipement est en arrêt ou *downtime*,

Décomposition du *downtime*. (figure 1.12)

10. Norme SEMI-E10-E79

1. Arrêts non planifiés ou *unscheduled downtime* (UDT).
2. Arrêts planifiés ou *scheduled downtime* (SDT).

Dans la suite, nous ne nous intéresserons qu'aux temps où l'équipement n'est pas en test ou en arrêt. C'est-à-dire à l'*uptime*.

Décomposition du *uptime*. (figure 1.12)

1. Temps de production ou *productive time* qui représente le temps effectif pendant lequel l'équipement remplit la fonction pour laquelle il est choisi.
2. Temps de tests de lots d'*engineering* ou *engineering time*. C'est le temps pendant lequel l'équipement est dans les conditions de process mais, est utilisé pour des expérimentations.
3. Temps représentant tous les arrêts autres que les arrêts non planifiés ou *standby*. Durant les temps de *standby* l'équipement est dans les conditions de process mais n'est pas utilisé (par exemple : temps où l'opérateur n'est pas devant la machine).

Décomposition du *productive time* : (figure 1.12)

- *Inefficiency* noté INEFF. Il est calculé uniquement quand l'équipement fonctionne à vitesse constante et qu'interviennent des « petits temps d'arrêts ». Exemples : changement d'une recette, temps d'échange entre deux lots.
- *Rework* ou RWK. C'est le temps perdu à cause de la destruction de lots (*scraps*) ou le temps de retraitement de certains lots.
- *Non Product Wafer*, noté NPW, qui représente le temps perdu pour la qualification des équipements de production. Un équipement est qualifié quand il est capable de traiter correctement la recette liée à un lot. Les NPWs sont en fait des lots qui testent si l'équipement est prêt à être utilisé.
- ***Overall Equipment Efficiency*** noté OEE, permet de connaître la durée du temps de fonctionnement effectif de l'équipement. Il relie trois éléments en un seul pourcentage : le temps pendant lequel la machine est en marche, le nombre total de lots traités par l'équipement et le nombre de « bon » lots qui ont été traités.

Le challenge de la production est d'avoir un OEE par équipement le plus élevé possible¹¹. Pour le lecteur désireux d'en savoir plus, je recommande la lecture de [81] qui traite des problèmes de perte de capacité d'équipements.

Cette décomposition montre que la gestion du *productive time* est un élément important dans la maximisation du temps de fonctionnement effectif de l'équipement. Dans la littérature, les problèmes de *productive time* sont des problématiques de gestion du choix et d'ordonnancement des lots, dont celui du regroupement des lots (ou *batching*). Ces problématiques sont expliquées dans ce qui suit :

- **Gestion du choix des lots** (*lot release control*) : ce problème est celui de la stabilisation et de la non-détérioration de la productivité de l'unité de fabrication. Il consiste à déterminer à quel moment et quels lots choisir, parmi une liste de lots.
- **L'ordonnancement des lots** (*lot scheduling problem*) : il consiste à déterminer la séquence de passage de lots en attente devant des d'équipements en parallèle¹². Je recommande la lecture de [50] et [9] pour les problèmes d'ordonnements de multiples produits.
- **L'ordonnancement des batches** (*batch scheduling*) : l'objectif du *batching* est triple : (1) *regrouper* des lots (*batch*) en attente devant une station d'équipements, (2) *déterminer à quel moment* et (3) **dans quelle séquence** doivent être processés les *batches*. Pour le lecteur désireux d'en savoir plus, je recommande la lecture de [95] qui traite de la rentabilité du process par batch et [68] pour les problèmes d'ordonnement de batches.

1.4.2 Politiques de gestion des véhicules

Motivation.

Le transport étant un service pour la production, les politiques de gestion des véhicules doivent donc être pensées en fonction des exigences de la production. Le transport ne doit en aucun cas être un élément qui limite la production. Pour ce

11. Dans tous les cas, de ne pas le dégrader.

12. les stations d'équipements en parallèle sont un ensemble d'équipements pouvant traiter des lots ayant la même recette. Exemple : les équipements de photolithographie

faire, il est nécessaire de choisir la politique de transport qui fournisse la meilleure adéquation entre ressource de transport et besoin en production car, il en existe plusieurs [69].

Cependant, choisir la politique de transport adaptée n'est pas une chose facile car, elle est fonction de plusieurs paramètres (tels que le nombre de véhicules, la priorité des lots, la maintenance, la taille de l'unité de fabrication, etc.). Bien que l'on utilise l'article « la » pour parler de la politiques de transport des véhicules, il est important de souligner qu'une politique de transport est constituée de deux sous-politiques : une sous-politique permettant de garantir un niveau de service ou *vehicles dispatching policy*, et une sous politique permettant de trouver le plus court chemin, libre de tous conflits de transports ou *routage policy*.

Politiques de *dispatching* des véhicules

Le *dispatching* des véhicules consiste à trouver un ordonnancement permettant de traiter toutes les demandes de transport. L'objectif est de garantir un niveau de service en sélectionnant une requête de transport parmi toutes celles qui attendent un transfert, et l'affecter à un véhicule disponible. Le *dispatching* se déclenche dans les trois cas suivant :

- 1 - juste après le dépôt par un véhicule d'un lot (le véhicule devient libre),
- 2 - lorsqu'un véhicule arrive sur une place de parking,
- 3 - et enfin lors de la requête de transport d'un lot.

Le *dispatching* permet donc de minimiser le temps d'attente des lots, donc la longueur la file d'attente, de maximiser la productivité du système et de garantir un niveau de service à un équipement donnée [5]. Nous reprenons ici les principales politiques de *dispatching* de [70].

Cas d'une requête de transport et plusieurs véhicules

- *Nearest Vehicle* (NV). Dès l'émission d'une requête de transport, c'est le critère de distance qui est évalué entre les véhicules libres et la station. Le véhicule choisi est celui dont la distance est la plus proche. Des variantes à cette politique prennent en compte la vitesse des véhicules (proximité temporelle). Dans ce cas, le véhicule choisi sera celui qui prendra le moins de temps.

Cette règle permet donc de minimiser les distances parcourues à vide, car elle sélectionne le couple <véhicule, produit> dont la phase de récupération est la plus courte.

- **Longest Idle Vehicle** (LIV). Le véhicule qui est resté inactif le plus longtemps sera le véhicule affecté. *Cette règle assure un équilibre de la charge de travail pour tous les véhicules* dans le système.
- **Least Utilized Vehicle** (LUV). Le véhicule sélectionné sera celui qui aura le moins servi à l’instant de la prise de décision, c’est-à-dire celui dont le taux d’utilisation est le plus petit. *Cette règle assure un équilibre de la dégradation des véhicules* dans le système en terme de maintenance techniques. On évite ainsi qu’un véhicule donné soit plus souvent en réparation que les autres.

Cas d’un véhicule et plusieurs requêtes de transport

- **Shortest Travel Time/Distance** (STT/D) : la station desservie sera celle qui est la plus proche du véhicule en termes de temps ou de distance de trajet.
- **First Come First Serve** (FCFS) : le véhicule est affecté à la requête qui attend depuis le plus longtemps.
- **Unit Load Shop Arrival Time** (ULSAT) : le véhicule dessert la station dont la pièce devant être transportée est entrée en fabrication depuis le plus longtemps.

Ces règles ne couvrent pas tous les cas, par exemple le cas de plusieurs requêtes avec plusieurs véhicules n’est pas traité. Dans ce cas, le système ne peut pas évaluer toutes les possibilités, il va réduire le nombre de possibilités (par des critères de distance, de niveau de trafic, etc.) jusqu’à se retrouver dans le cas de l’une des deux situations précédentes et ainsi pouvoir appliquer l’une d’entre elle. Pour le lecteur désireux d’avoir plus d’informations, je recommande la lecture de [4].

Les politiques de routage

L’objectif du routage est de trouver un trajet le plus court en termes de distance ou de temps. Une décision de routage comprend trois aspects :

- la recherche de l’*existence de chemins* qui mènera le véhicule de son origine vers sa destination,

- le *chemin doit être réalisable* c'est-à-dire libre de conflits, de blocages et sans congestion,
- le *chemin doit être optimal* ou *pseudo-optimal* selon une fonction objective.

Pour le lecteur désireux d'en savoir plus, je recommande la lecture de [35] pour le problème du plus court chemin, [23] concernant la recherche dynamique du plus court chemin ou enfin [45] pour les problèmes de routage avec fenêtre de temps.

1.4.3 Gestion du stockage

Rappel : A cause de l'agencement des équipements, l'attente des lots est une donnée incontournable. Dans la section 1.4.3, nous avons dit que les dispositifs de stockage pouvaient se classer selon des critères de capacité et de temps d'accès.

Dans cette thèse nous distinguerons principalement deux types de stockage classés en fonction de ces deux critères : les emplacements unitaires de stockage ou *bins*, possédant un rapide temps d'accès et les containers pouvant contenir plusieurs emplacements unitaires de stockage (stockeurs), ayant un long temps d'accès.

Les motivations

On appelle *temps de cycle du stockeur* la somme du temps de cycle du robot interne et du temps d'entrée/sortie du lot sur le port du stockeur. Ainsi, un lot à l'intérieur d'un stockeur provoquera un temps d'attente supplémentaire dû, au temps de cycle du robot. Ceci aura pour conséquence d'augmenter le risque d'attente à vide des équipements de production, donc du temps de non-production. En particulier, appelons T le temps à partir duquel l'équipement n'aura plus aucun lot sur ses ports de chargement. Si le temps d'attente du lot (temps de cycle du robot + temps pour sortir du stockeur) est supérieur à T alors, nécessairement le stockeur devient un élément limitant de l'unité de fabrication. Néanmoins, le principal avantage du stockeur est qu'il possède une grande capacité (en termes d'emplacement) de stockage.

Mais *est-il possible d'avoir les avantages du stockeur (grande capacité de stockage) sans les inconvénients (long temps de cycle) ?* Nous tenterons de donner une réponse à cette question dans le chapitre suivant.

Fonctionnement du système de stockage

Les fonctionnements des systèmes de stockage sont très peu traités dans la littérature, sans doute parce qu'ils sont considérés comme des temps de non-production. Pourtant, l'agencement même des équipements de production impose aux managers d'accepter l'attente. Mais cette attente, même si elle fait partie du temps de non production influe pour beaucoup sur le temps de cycle des lots.

Lorsqu'un lot à l'intérieur d'un stockeur émet une requête de transport au système, le robot à l'intérieur du stockeur connaît l'identité du lot et sa position. Ce dernier va se déplacer pour aller chercher le lot afin de l'amener sur le port de déchargement interne du stockeur. Le lot est ensuite acheminé vers le port de déchargement extérieur à l'aide d'une sorte de tapis roulant. Le véhicule peut alors prendre le lot pour l'amener à sa destination.

Nous verrons dans la suite qu'une mauvaise synchronisation entre politique du choix des lots et gestion du robot, conduit à de mauvaises performances en termes de transport des lots.

1.5 Quelques indicateurs de performance

Baromètre. Instrument ingénieux qui indique la sorte de temps que nous sommes en train de subir.

Ambrose Bierce

1.5.1 Introduction

Les indicateurs présentés dans cette section sont ceux généralement utilisés dans le domaine des semi-conducteurs et en particulier dans l'entreprise STMicroelectronics.

Un indicateur permet de mesurer de façon objective un phénomène étudié : c'est un outil décisionnel synthétique, c'est-à-dire facile à établir et pratique à utiliser. Il doit être :

- *Pertinent* : c'est-à-dire exactement approprié au phénomène à observer. Il doit apporter une valeur ajoutée pour la détermination de l'efficacité ou non du système de management,
- *Précis* : c'est-à-dire fournir des informations univoques,
- *Fiable* : dans sa définition, sa collecte et son exploitation,
- *Rapide à collecter et à traiter* : c'est-à-dire qu'il ne doit pas gêner l'activité en cours,
- *Communicable* : pourquoi cet indicateur, dans quels buts ? Pour en faire quoi ?

Un indicateur mesure donc l'efficacité d'un dispositif mis en place. C'est un outil de dialogue entre les utilisateurs ayant des cultures et des préoccupations différentes.

1.5.2 Indicateurs de production

Temps de cycle

Le temps de cycle de production, encore appelé temps de séjour ou délai de fabrication, est la durée totale nécessaire pour fabriquer un lot de wafers. Il mesure le temps écoulé entre le moment où le lot entre dans l'unité de fabrication et le moment où il sort. Les entreprises avec des temps de cycle courts sont capables de lancer

souvent de nouveaux produits, de pénétrer plus rapidement de nouveaux marchés, de réagir plus efficacement aux changements de la demande et de délivrer à temps leurs produits. Elles peuvent aussi offrir à leurs clients des prix plus compétitifs car elles ont rationalisé leurs process et elles maîtrisent le nombre de wafers se trouvant dans l'unité de fabrication à un instant t (niveau d'encours, voir section 1.5.2). Selon [70], une réduction du temps de cycle équivaut à une augmentation du niveau de satisfaction des clients.

Pour des raisons de confidentialité, on préfère utiliser un autre indicateur : le *X-factor*. Il représente le rapport du temps de cycle total sur le temps total de process. Il est donné par la relation suivante : $\frac{\text{temps de cycle}}{\text{temps total de process}}$ [66]. Avec cet indicateur, l'objectif recherché est d'avoir un X-factor le plus proche possible de l'unité. La façon la plus économique (sans investissements financiers) d'y parvenir est de réduire le temps de cycle des produits. En effet, investir dans les équipements coûte extrêmement cher (ex : près de 70 millions d'euros pour un scanner). Pour le lecteur désireux d'en savoir plus je recommande l'excellent article de [71].

Remarque : le calcul du temps de cycle prend en compte des étapes additionnelles, non communes à toutes les fabs. Par exemple, le nombre d'étapes de métrologie (mesures de plaques) est variable d'une unité de fabrication à l'autre. Cet indicateur n'est donc pas le plus pertinent pour comparer les performances des fabs. Les industriels utilisent à la place le *day-per-mask*. Il représente le temps nécessaire pour le changement d'un « masque » à l'étape de photolithographie. Cet indicateur permet de comparer plus efficacement les performances entre les fabs. En effet, un même produit est caractérisé par un nombre de masques bien précis suivant la « *technologie* » utilisée. Ainsi, l'indicateur temps de changement d'un masque est toujours le même quelque soit l'unité de fabrication.

Niveau d'encours

Le niveau d'encours (*(work-in-progress)*, WIP) est défini comme le nombre de wafers de production se trouvant dans l'unité de fabrication à un instant donné. Le nombre de wafer pris en compte est aussi bien sur une activité de production que

sur une activité de non-production (transport et attente des lots).

Les encours représentent du capital immobilisé dans les ateliers, donc trop d'encours diminue la trésorerie disponible. En régime stationnaire, c'est-à-dire si on attend « suffisamment longtemps », la moyenne du WIP devient proportionnelle à la moyenne du temps de cycle (CT). Cette loi est connue dans la littérature sous le nom de loi de Little [61] et donnée par la formule : $WIP = \lambda CT$, où λ représente le taux moyen d'arrivée des lots par unité de temps. Ainsi, d'après cette loi, pour λ fixé et en régime stationnaire, il n'est pas possible de réduire la moyenne du WIP sans réduire la moyenne du temps de cycle et *vice versa*.

Débit du système ou *throughput rate*

Le débit du système, appelé « *throughput rate* », est défini comme le nombre de wafers finis qui sortent de l'unité de fabrication par unité de temps. Cet indicateur permet de savoir si le système est stable, c'est-à-dire s'il n'y a pas d'accumulation de produits quelque part dans le système. En particulier, lorsque le taux de lancement de lots est très proche du taux de production.

Considérons une ligne de fabrication idéale, c'est-à-dire sans aucune variabilité de temps d'arrivée ou de process. Avec un niveau d'encours bas, le temps de cycle est égal au temps de process total théorique auquel on rajoute le temps de transport. Tant que la ressource est inférieure à la capacité de la ressource goulot, le temps de cycle reste à cette valeur, sans aucun temps d'attente. A partir du moment où le *throughput* arrive à sa valeur maximale, les temps d'attente commencent à apparaître. A partir de ce point, toute augmentation du niveau des encours se traduit par une augmentation du temps de cycle [16].

1.5.3 Indicateurs de transport

Temps de livraison

Le temps de livraison est le temps écoulé entre la prise d'un lot à un point A et le dépôt du lot à un point B. Notre définition inclut le temps de prise (*pick-up time*) et de dépôt (*drop-down time*) du lot. On distingue mieux ainsi sur un même

trajet, les véhicules transportant des lots de ceux qui n'en transportent pas. Il suffit de comparer leur temps de transport sur le trajet.

Le temps de livraison est donné dans la littérature par sa moyenne et son écart type. Ces indicateurs permettent seulement de savoir la plage de temps de livraison des lots à savoir *moyenne* \pm *ecart type*. On ne peut répondre de manière précise à la question : quel est le pourcentage de lots livrés en un temps t ? Pourtant, le transport est un service pour production et celle-ci est tournée vers le besoin client. Ce dernier a besoin de garantis. Calculer le temps de livraison sous forme de fréquences cumulées croissantes permet de fournir cette garantie. On peut ainsi connaître le temps de livraison pour 50%, 85% et 95% des lots et, cela est plus parlant aux clients.

Le temps de livraison est un indicateur de la santé du système AMHS, mais, il ne permet pas de savoir comment, ni quand un problème survient.

Temps d'assignation

Il représente le temps entre l'émission de la requête de transport par le lot et l'attribution de cette requête à un véhicule de transport. Ce temps va fortement dépendre :

1. du temps de traitement informatique (protocoles de communication),
2. de la qualité des algorithmes d'attributions,
3. de la politique de gestion de transport (dispatching des véhicules) et,
4. du nombre de véhicules, car plus il y a de véhicules, plus 1 et 2 seront petits.

1 et 2 possèdent une très grande variabilité et 3 est en général fixé. Donc, le temps d'assignation possède une très grande variabilité. Réduire cette variabilité est un problème complexe. Dans la pratique le temps d'assignation est très difficile à calculer. Une des raisons principales est que les constructeurs n'aiment pas communiquer sur les performances réelles de leurs systèmes AMHS. On peut consulter [74] pour une méthode analytique de calcul du temps d'assignation dans une zone.

Cet indicateur intervient lors de tout déplacement d'un véhicule dans l'unité de fabrication. Il contribue donc fortement à la réduction du CET. La figure 1.13 donne une décomposition du temps pour échanger un ancien lot par un nouveau lot sur un même port de chargement.

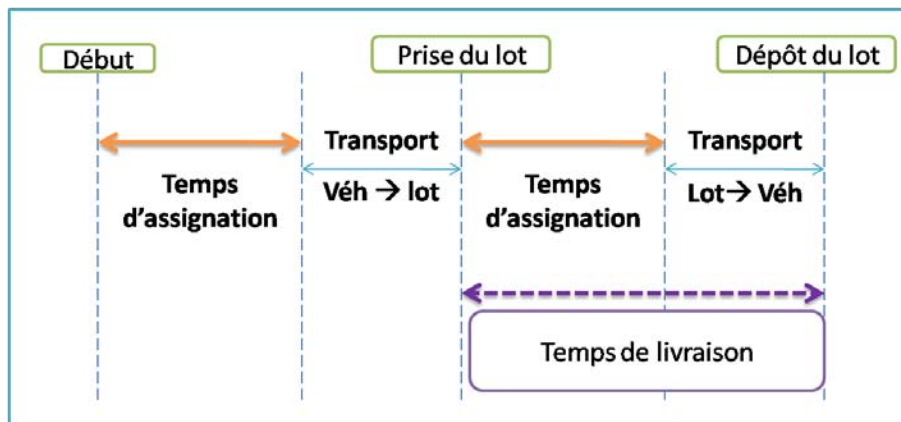


FIGURE 1.13 – Décomposition temps de chargement et de déchargement d'un lot.

Taux de charge des véhicules

Le *taux de charge* est le rapport du temps en charge d'un véhicule sur le temps total des déplacements. Il est donné par la formule suivante : $\frac{\text{temps total en charge}}{\text{temps total des déplacements}}$. Le temps en charge d'un véhicule est la somme des temps pendant lesquels un véhicule non vide se déplace. Cet indicateur a une double fonction : connaître l'efficacité des politiques de transport utilisées et la qualité de la répartition des véhicules dans l'unité de fabrication. Un déséquilibre du taux de charge des véhicules peut traduire le fait qu'il n'y a pas les bons véhicules aux bons endroits. Si on est certain de la qualité des politiques de transport, le taux de charge des véhicules permet de savoir s'il y a trop ou pas assez de véhicules. Cela se traduit par une sous utilisation des véhicules. Attention cependant car cet indicateur peut donner des résultats difficilement exploitables. Par exemple, si on force les véhicules à rester au repos, il faut en tenir compte dans le calcul du taux de charge des véhicules.

Temps d'échange des lots sur un port de chargement

Le *Carrier Exchange Time* (CET) est le temps mis par le système pour remplacer un lot sur un port de chargement. Prenons comme instant initial un lot A ayant terminé son opération de production sur un équipement. Ce lot signale alors au système qu'il est prêt à être enlevé du port de chargement (*ready to unload*). Le

Carrier Exchange Time est la somme du temps de déchargement du lot A et, de chargement d'un lot B sur le même port de chargement (Il est aussi appelé *white space* figure 1.14).¹³

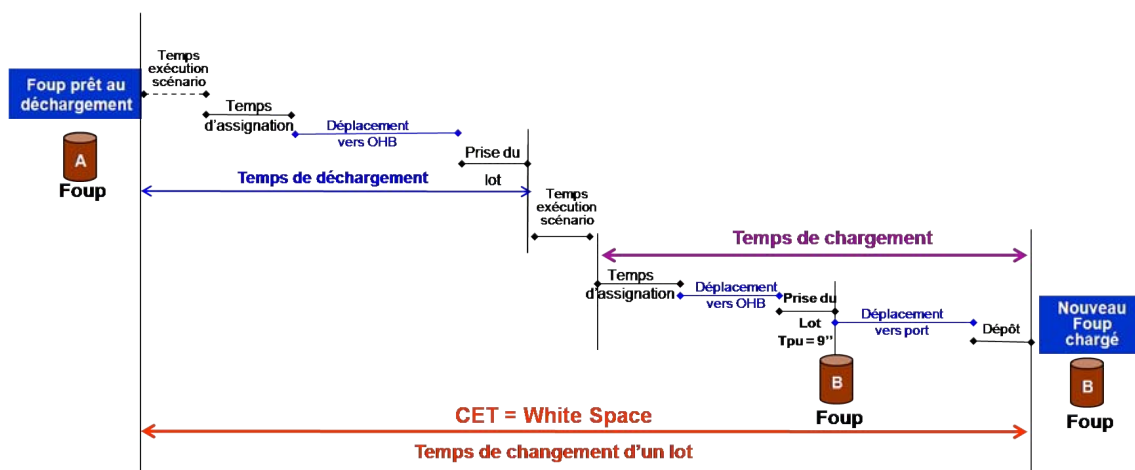


FIGURE 1.14 – Décomposition du *Carrier Exchange Time*.

C'est un indicateur théorique très important. En effet, il faut faire passer un nombre maximum de lots sur les équipements afin d'amortir les investissements sur l'achat de ces derniers qui, on le rappelle, coûtent très cher. Cette approche porte le nom de *continuous processing*. Elle permet idéalement de faire tendre vers 0 le temps à vide du port de chargement (*loadport idle time*). Le *Carrier Exchange Time* permet de mesurer la cadence du *continuous processing*. Bien que théoriquement très important [52], le CET reste très difficile à mesurer en pratique. Sa définition montre qu'il est fonction des besoins (temps de *process*, de *setup*, etc ...) propres à chaque équipement. Le challenge consiste donc à ne jamais laisser un équipement vide à l'intérieur d'une zone. On respecte ainsi le CET objectif de chaque équipement. En fonction de leur temps de process, donc de leur CET, on peut ainsi classer les équipements en deux groupes :

- les équipements prioritaires : nécessitant un CET « court »,
- et les équipements non prioritaires.

13. Source : G. Lamiable

C'est la coexistence de ces équipements au sein d'une même zone qui caractérise la difficulté d'assurer le *continuous processing*.

1.5.4 Indicateurs de stockage

Temps d'accès à un emplacement de stockage

Il représente le temps que met un lot pour physiquement se retrouver sur un emplacement de stockage, le point de départ étant le début du dépôt ou de la prise d'un lot sur ce espace de stockage. Ce temps est fonction du type de stockage. Par exemple, pour les stockeurs unitaires, il correspond au temps de prise (*pick-up*) ou de dépôt du lot (*drop-down time*) qui est de l'ordre de 8 à 9 secondes environ. Pour les stockeurs, il faut ajouter à ces temps le temps de cycle du robot interne (20 secondes) et celui d'un « tapis roulant » interne (10 secondes). Ce qui fait que le temps d'accès d'un stockeur est 5 à 6 fois plus élevé que celui d'un stockeur unitaire encore appelé *OverHead Buffer*. C'est un critère très important qu'il faut prendre en compte lors de l'optimisation du système AMHS.

Taux d'utilisation des emplacements unitaires de stockage

Cet indicateur représente le temps pendant lequel le stockeur unitaire n'est pas vide. Il correspond au rapport du temps total d'utilisation du stockeur sur le temps total observé. Il est donné par la formule $\frac{\text{temps total utilisation stockeur}}{\text{temps total observe}}$.

C'est un indicateur très important, en particulier, pour le cas des stockeurs unitaires. Car il permet de détecter les lots *holder*, c'est-à-dire des lots qui sont sortis de leur route. Ils ne sont alors plus « visibles » par le système et changent provisoirement de statut. Ils vont ainsi rester sur la zone de stockage jusqu'à leur « réintégration » informatique dans le système. Ils occupent ainsi inutilement et pendant longtemps des zones de stockages au détriment des lots de production.

1.6 Conclusion

Dans ce chapitre nous avons présenté les bases nécessaires à notre étude. La présentation du contexte dans lequel se fabriquent les puces, ne permet pas seulement de mieux appréhender les contraintes de fabrication, mais surtout, celles du transports liées à la fabrication. Les problèmes de transports [26] sont assez bien identifiés dans la littérature, mais très souvent sans forte corrélation avec la production et le stockage [53]. Il suffit de regarder l'étude bibliographique de [37]. Le stockage, beaucoup plus *négligé* dans la littérature, est pourtant très important [39], surtout pour détecter les problèmes de *sous-capacité*¹⁴ de l'unité de fabrication.

Dans ce premier chapitre, nous avons répondu aux deux questions :

1 - *Que fait le système ?* Le système transporte, traite et stocke des produits dont la fabrication est très complexe.

2 - *Comment mesurer ses performances ?* À l'aide d'au moins deux indicateurs (par exemple, le couple CET et temps de cycle) car un seul indicateur n'est pas suffisant [96].

Dans la suite nous tenterons de répondre aux questions suivantes :

- *Quel est la nature des relations entre transport, production et stockage ?*
- *Quel est la force de ces relations ?*

Comment allons-nous procéder ? Nous allons nous servir de la littérature existante, analyser les relations entre les trois éléments et utiliser un plan d'expérience pour étayer nos affirmations. Avant cela, nous allons présenter les particularités du cas industriel de cette thèse. C'est l'objet du chapitre suivant.

14. La sous-capacité de l'unité de fabrication représente la situation où il y a plus de lots que le nombre d'équipements ne peut traiter.

Chapitre 2

Présentation et analyse du cas industriel

Ce chapitre traitera :

- du processus de fabrication et du système de transport automatisé de l'usine 300 mm de STMicroelectronics à Crolles (section 2.2),
- des particularités du fonctionnement du système de transport automatisé de l'usine 300mm de STMicroelectronics à Crolles (section 2.3).

Nous répondrons aux questions : *comment fonctionne l'usine de fabrication 300mm de STMicroelectronics à Crolles ? Quelles sont les particularités de cette usine ?*

2.1 Résumé du chapitre

Dans une ligne de fabrication mono-produit, les contraintes d'enchaînements de tâches sont différentes de celles d'une ligne multi-produits (*high mix*). Plus généralement, les performances d'une unité de fabrication sont fortement liées à sa configuration [53]. Il est donc important, une fois que l'on sait ce que fait le système, de savoir comment il le fait, quelles sont ses limites et contraintes particulières. Ce second chapitre présente le cas industriel de notre étude : l'unité de fabrication 300mm de Crolles que nous nommerons Crolles 300.

La première partie décrit le mode de fonctionnement du système de transport des lots. Elle montre qu'à Crolles 300, l'on est en mode hybride, c'est-à-dire coexistence homme-machine. La difficulté, dans ce mode de fonctionnement, vient de la « perturbation » que peut engendrer l'humain sur le système automatique.

La seconde partie décrit les particularités du système et l'on peut en retenir quatre importantes.

- **Le système de transport est unifié.** Tous les rails sont liés et les véhicules peuvent aller directement d'une machine à l'autre.
- **Le système gère ses requêtes de transport par une approche utilisant le « service minimum » garanti.** La politique de « service minimum » consiste à toujours garder un nombre minimal prédéfini de véhicules dans une zone donnée, afin d'être réactif aux requêtes de transport. De cette politique, vont beaucoup dépendre les temps d'assignation d'une requête de transport à un véhicule.
- **Le système utilise une approche particulière de repositionnement des véhicules,** qui consiste à « envoyer » ou à « appeler » des véhicules afin de respectivement désengorger ou compenser la zone concernée en véhicules.
- **Le système utilise une approche consistant à regrouper en « stockeur par défaut »,** les emplacements unitaires de stockage au plafond pour la gestion du stockage des lots. Cette approche permet de distinguer les espaces de stockage à accès rapide, de ceux à accès moins rapide. Cette distinction a un rôle important sur les performances du système de transport.

Toutes ces particularités nous permettront de mieux comprendre le contexte industriel et d'affiner notre problématique.

2.2 Fonctionnement de l'usine 300mm de Crolles

Mieux vaut comprendre peu que comprendre mal.

Anatole France

Objectifs du début de thèse : L'unité de fabrication réelle sur laquelle j'ai travaillé était (et est) en phase de construction (à certains endroits) ou d'agrandissement (à d'autres endroits) en même temps que ce déroulait la thèse. De ce fait, la première problématique industrielle consistait à modéliser au plus vite le système existant afin de pouvoir l'accompagner de la manière la plus efficace¹ possible dans son extension. En outre, le fonctionnement du système lui-même n'étant pas bien connu par l'équipe² AMHS, il n'était pas très facile de comprendre comment fonctionnait réellement le système. La modélisation du système va donc permettre de fixer, infirmer ou valider les hypothèses sur le fonctionnement et l'amélioration du système.

Introduction

Les travaux présentés dans cette thèse ont été réalisés en partenariat avec *ST-Microelectronics*, une compagnie Franco-Italienne de semi-conducteurs. L'unité de production considérée est la seule unité de fabrication 300mm de l'entreprise et est localisée à Crolles (France). Cette unité, en plein travaux d'extension, possède déjà 1800 m de rails utilisés en 2007 (figure 2.1). Notre travail s'inscrit donc dans une volonté de l'entreprise d'avoir des réponses aux questions du type *que faire si ... ?* Et d'identifier les conditions optimales de gestion du transport de wafers. Par exemple, comment réagit le système de transport si on augmente le nombre de lots ?

Toutes les unités de fabrication de semi-conducteurs ne possèdent pas nécessairement les mêmes processus, les mêmes systèmes de transport ou de pilotage. Par exemple, l'unité de fabrication de Crolles 300 se distingue de celles rencontrées dans la littérature par son fonctionnement en mode hybride, c'est-à-dire la coexistence entre l'homme et la machine. Le système de transport utilise des véhicules aériens

1. notamment en s'aidant de la possibilité de représentation 3D dynamique.

2. le fournisseur étant propriétaire des logiciels.

et de capacités unitaires. Lorsqu'un véhicule rencontre sur son passage un véhicule au repos, le système de transport permet au véhicule de « pousser » le véhicule au repos afin de libérer le passage. Enfin, la politique de gestion du système de transport est particulière : elle se fait par « service minimum », qui consiste à conserver des véhicules toujours prêts et proches des besoins.

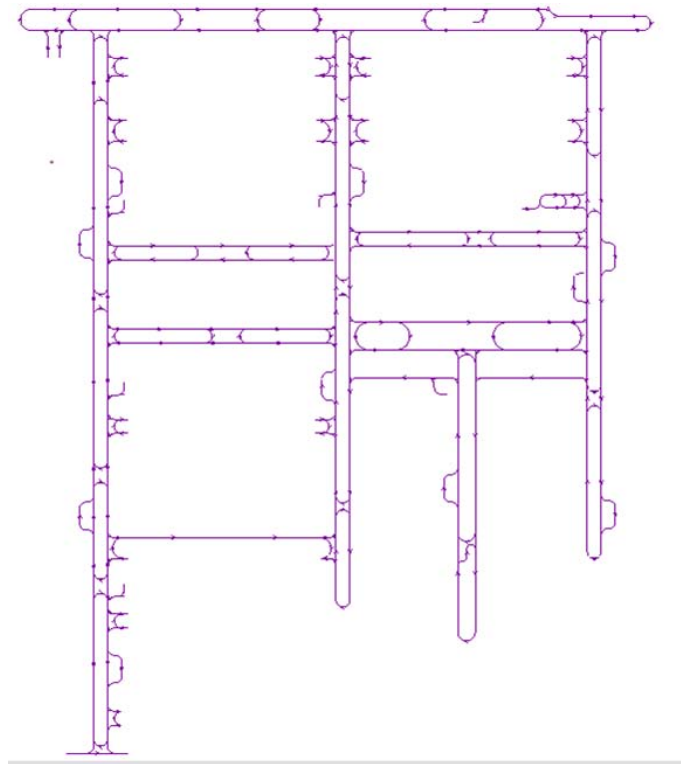


FIGURE 2.1 – Agencement des rails de l'unité de fabrication de Crolles 300.

La suite de cette section est découpée en deux parties. Nous présenterons le système de transport de Crolles 300mm et son fonctionnement global dans la section 2.2.1. Nous décrirons ensuite le processus de fabrication vue d'un point de vue transport à la section 2.2.2.

2.2.1 Système de transport

Avant de donner les détails du système, nous allons commencer par donner une vue globale du système de pilotage du transport des plaquettes de Crolles 300 mm.

Pilotage global du système de transport

On considère généralement trois acteurs essentiels au pilotage d'un système de transport (qui est le cas pour celui de Crolles 300 mm).

1. *le Manufacturing Execution System (MES) — le donneur d'ordre*
2. *l'Execution Control System (ECS) — le manager de l'information*
3. *l'Automated Material Handling System (AMHS) — l'exécuteur d'ordre*

L'interaction entre ces trois facteurs permet le pilotage du système, comme montré sur la figure 2.2.

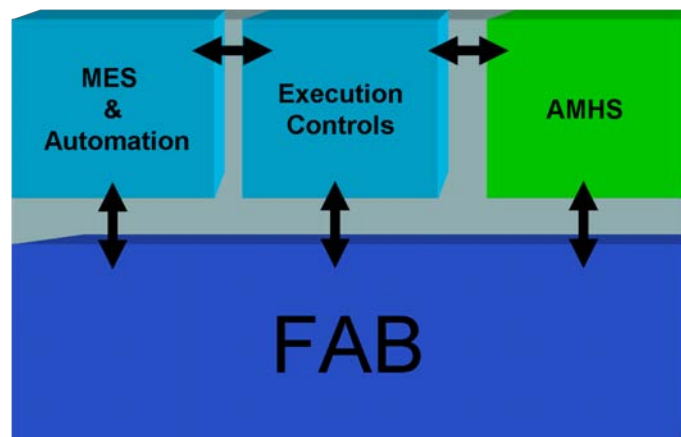


FIGURE 2.2 – Interactions entre les éléments d'une unité de fabrication.

Nous allons à présent présenter le fonctionnement de chaque acteurs.

Le Manufacturing Execution System (MES) possède les informations sur les gammes opératoires des lots, sur les équipements qualifiés pour exécuter une tâche, sur leurs emplacements et sur les noms des espaces de stockage. Plus précisément,

il possède des informations sur l'identité, la source, la destination et la priorité des lots. Le MES est le donneur d'ordre qui permet de faire passer un lot d'une étape n à une étape $n + 1$. Le MES échange et met à jour les informations sur le système à l'aide de l'**automation** à travers l'équipements interface. L'automation est le centre de gestion de l'information. Il est l'informateur du MES.

L'Execution Control System (ECS) est le centre de traitement de l'information. Il est essentiellement composé de *l'activity manager*, qui est le cœur décisionnel du système car c'est lui qui gère les différents scénarios. Il est responsable des décisions de transport, d'ordonnancement des lots grâce aux mécanismes de calcul de priorité des lots implémentés dans l'outil de dispatching, etc. La performance de l'unité de fabrication est liée à la qualité des règles implémentées dans *activity manager*. Ces règles sont basées sur la logique « premier entré, premier sorti », couplée à la priorité des lots.

L'Automated Material Handling System (AMHS) est en charge du transport grâce au contrôleur de véhicules et du stockage des lots. Le contrôleur de véhicules gère le transport à l'aide d'unités de contrôle local propres à chaque zone. Les contrôleurs locaux échangent des informations toutes les 5 secondes avec le contrôleur de véhicules. Ce dernier possède les informations sur le nombre de véhicules présents dans chaque zone, donc de leurs besoins en termes de transport. Il est en charge de mettre en adéquation les véhicules et avec les requêtes de transport des lots de la zone. En effet, il connaît les zones où il y a des véhicules libres disponibles et les zones où il y a des besoins non satisfaits en termes de transport.

Mode de communication. Le moyen de communication utilisé entre tous les acteurs est « l'écoute ». Cela signifie que les différents acteurs ne s'envoient pas d'information directement, mais la diffusent dans un réseau auquel sont reliés tous les autres acteurs. Chaque acteur reçoit l'information et ne traite que celle qui le concerne. La figure 2.3 illustre ce mode de communication.

Nous allons à présent nous intéresser plus précisément au pilotage du système de transport à l'échelle d'une zone.

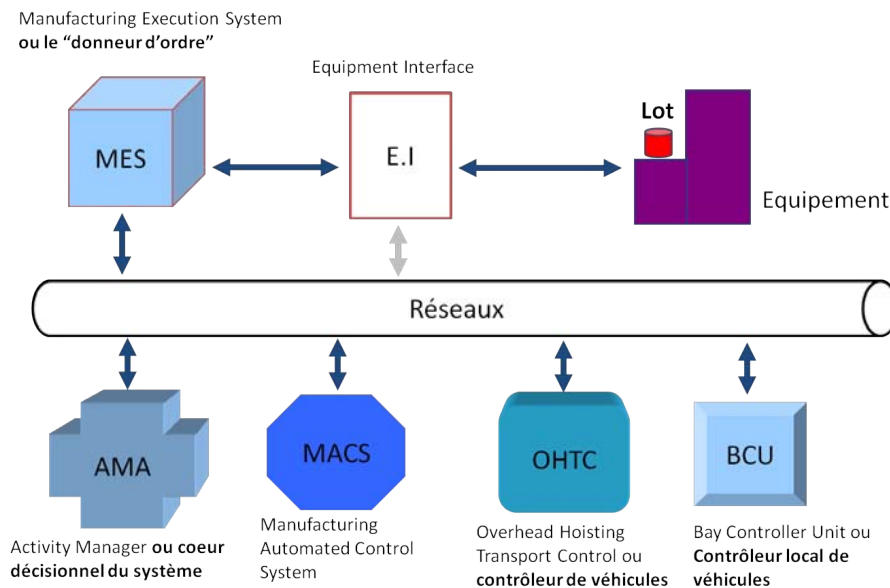


FIGURE 2.3 – Modes d'échange d'informations entre les acteurs de l'unité de fabrication.

Pilotage local du système de transport

L'unité de fabrication de Crolles 300 mm est découpée en zones comme représenté sur la figure 2.4. A quelques équipements près, les zones correspondent à des ateliers.

Le pilotage du système de transport à Crolles 300mm se fait par zone. [70] a montré que le découpage en zones permet une meilleure gestion de sous-systèmes plus petits et moins complexes que l'unité de fabrication. La zone sera un élément essentiel dans la compréhension du fonctionnement du système industriel. Il est donc nécessaire de bien connaître, la « population » physique d'une zone et les interactions entre les acteurs de la zone et d'autres qui sont extérieurs à la zone.

Constitution de la « population » physique d'une zone. Elle se compose de :

- **Rails** sur lesquels se déplacent les véhicules. Les rails sont suspendus au plafond et alimentés électriquement par un câble dont le rayon d'action est de 60 m. Le nombre maximum de véhicules dans une zone est limité à 12 pour des

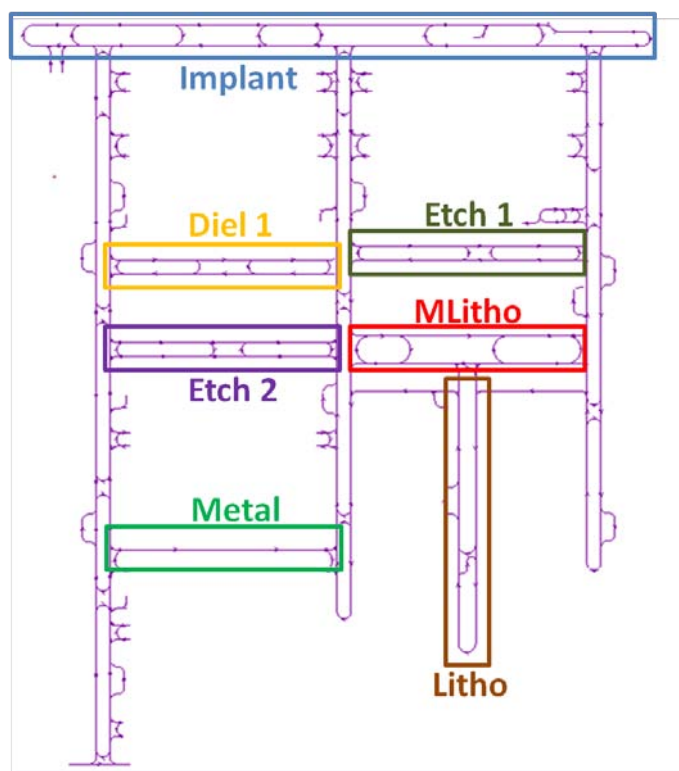


FIGURE 2.4 – Découpage en zones de l'unité de fabrication.

raisons de puissance électrique d'alimentation.

- **Véhicules** pour transporter les lots. Les véhicules sont aériens, ce qui permet une meilleure utilisation de l'espace au sol [89] et ils se déplacent sans contact sur les rails. Ils possèdent leur propre système de propulsion, alimenté par induction électromagnétique, ce qui permet de s'affranchir des problèmes de recharge ou de changement de batterie rencontrés avec d'autres types de véhicules. Ils sont tous du même type (modèle SRC 330 de chez MURATEC), de capacité unitaire c'est-à-dire qu'ils ne peuvent transporter qu'un seul lot. Ils possèdent deux bras mécanique : un bras vertical, permettant la prise/dépôt du lot par sa partie supérieure et, un bras latéral, permettant la prise/dépôt du lot sur des emplacements latéraux de stockage (figure 2.5). Deux capteurs infrarouges, un à l'avant, un à l'arrière, permettent aux véhicules d'adapter leurs paramètres (tableau 2.1) en fonction du trafic.



(a) Véhicule au repos.

(b) Dépôt latéral

FIGURE 2.5 – Dépôt latéral d'un lot par un véhicule au plafond.

Paramètres	Valeurs	Unités
Vitesse maximale	3,3 m/s	m/s
Accélération	1	m/s ²
Décélération	3	m/s ²

TABLE 2.1 – Paramètres techniques des véhicules SRC330.

- **Codes-barres** qui permet aux véhicules de se déplacer, se repérer et servent également de point de chargement/déchargement. Ils sont disposés à intervalles presque constants les uns des autres sur les rails. Pour se déplacer, les véhicules émettent des requêtes successives aux différents codes-barres. Les codes-barres servent également de point de parking pour les véhicules.
- **Equipements de production** pour le traitement des lots. Ils possèdent entre deux et quatre ports de chargement et une ou plusieurs chambres de process.
- **Emplacements de stockage** pour stocker les lots de production. On distingue les emplacements de stockage unitaire au plafond, auxquels les véhicules peuvent accéder latéralement et les gros (en termes de places) stockeurs au sol.

Nous connaissons à présent l'ensemble des éléments d'une zone. Nous allons à présent décrire comment interagissent ces éléments.

« **Interactions à l'intérieur d'une zone.** La figure 2.6 illustre l'agrandissement d'une zone avec ses différents acteurs. Remarquons que ce sont les véhicules qui servent à faire le lien physique entre les différents acteurs de la zone. Le résultat de l'interaction avec les autres éléments de la zone est le passage des véhicules par différents états. Nous donnons dans la suite la définition de ces différents états.

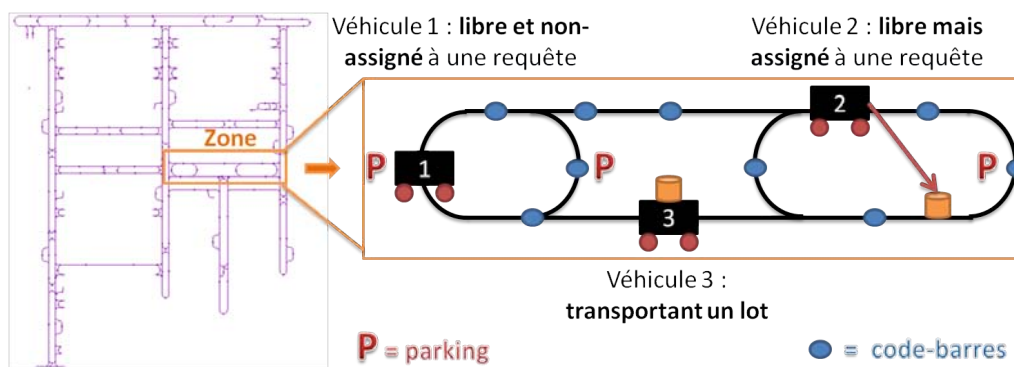


FIGURE 2.6 – Principaux acteurs d'une zone.

Dans toutes les définitions qui vont suivre, il est important de noter que *les véhicules doivent être de capacité unitaires.*

Définition d'un véhicule non-assigné : un véhicule *non-assigné* à une requête de transport est un véhicule qui ne possède aucune tâche de transport à effectuer parmi sa liste des tâches à faire, ni dans sa tâche courante. Cela ne signifie pas que le véhicule est immobile. Un véhicule peut être non-assigné mais en mouvement (exemple, véhicule ayant terminé toutes ses tâches et se dirigeant vers un parking).

Définition d'un véhicule libre : un véhicule libre est un véhicule qui ne transporte pas de lot. Avec cette définition, un véhicule non assigné n'est pas nécessairement libre. Par exemple, un véhicule en train de terminer sa dernière livraison et n'ayant aucune requête de transport en attente. Inversement, un véhicule libre n'est pas nécessairement non-assigné. Par exemple, un véhicule qui répond à une requête de transport est libre (il ne transporte pas de lot) mais n'est pas assigné (car il détient dans sa liste de tâches courantes une requête de transport).

Ainsi, dans une zone, il y a quatre états différentes pour un véhicule : (1) les véhicules libres et non-assignés, (2) les véhicules libres mais assignés, (3) les véhicules non-libres et assignés et (4) les véhicules non-libres et non-assignés. Nous verrons plus loin (chapitre 2.3 « particularités industrielles ») que ces derniers véhicules sont ceux que le système va comptabiliser comme candidats possibles d'une requête de transport de la zone de destination du véhicule.

Chaque fois qu'un véhicule se trouve dans l'un de ses états, il a interagi avec un des acteurs de l'unité de fabrication. Nous terminerons cette section en précisant qu'un véhicule est capable de « pousser » un autre véhicule gênant afin de continuer sa route. Cette caractéristique très importante permet de réduire les temps de transport des véhicules [47] et [46]. En effet, dans les systèmes où les véhicules ne possèdent pas cette propriété, le véhicule recherchera un chemin de contournement, s'il n'en trouve pas, il restera bloqué jusqu'à ce que le véhicule devant lui reçoive une requête de transport.

Le mode de recherche de trajets des véhicules est celui du plus court chemin en distance. Lors de leurs trajets, les véhicules sont capable d'éviter les zones à fort trafic ou d'embouteillage car ils sont à « l'écoute » du système.

2.2.2 Processus de fabrication

Le système de transport considéré dans cette thèse a une **configuration unifiée** (figure 2.1), c'est-à-dire que les véhicules peuvent aller directement d'un équipement à l'autre et *vice-versa*. Cette configuration couvre, au moment de l'écriture de ce manuscrit, **dix zones** regroupant différents ateliers de production. Mais à terme il y aura environ **une vingtaine de zones**.

Bien que l'on soit en configuration unifiée, nous utiliserons parfois les termes *interbay* et *inrabay* par souci de précision lors de nos descriptions.

Unité de fabrication 300mm de Crolles

Dans notre unité de fabrication, il y a environ 350 équipements regroupés en une centaine de familles de stations d'équipements. Chaque famille d'équipements comporte un ou plusieurs équipements identiques, ou capables de réaliser le même

traitement sur un lot. Chaque équipement possède un temps de préparation avant le traitement du lot ou *setup time*.

L'unité de fabrication est en mode **hybride** (figure 2.7), c'est-à-dire qu'il y a coexistence entre l'homme et la machine. Les opérateurs interviennent dans les zones où il n'y a pas encore de système de transport automatisé.

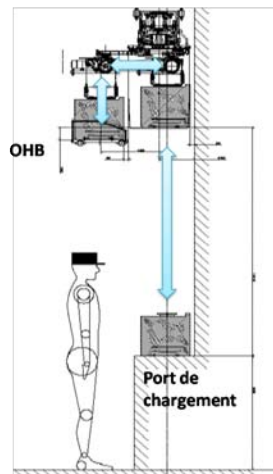


FIGURE 2.7 – Coexistence homme-machine.

La taille des lots varie entre 20 et 25³ wafers en moyenne. Les gammes opératoires (nombre d'étapes) considérées peuvent avoir jusqu'à **700 étapes** pour un seul produit. En particulier dans l'usine considérée, un lot visite environ **39** fois l'atelier de photolithographie et un peu plus de **152** fois l'atelier de métrologie. Les lots NPW⁴ sont également pris en compte. On a calculé qu'ils représentaient environ 15% du nombre de lots démarrés dans l'unité de fabrication.

3. Un lot contient rarement 25 wafers [97].

4. *Non Product Wafers*, lots servant à la qualification des équipements et/ou aux tests lors des pannes équipements

2.3 Particularités industrielles

Observer attentivement, c'est se rappeler distinctement.

Edgar Allan Poe

Introduction

Nous présentons dans cette section toutes les particularités qui distinguent notre unité de fabrication de celles rencontrées dans la littérature.

Ces particularités permettent de mieux appréhender la complexité de notre problématique d'un point de vue pratique et de fournir des solutions industrielles adaptées. D'un point de vue théorique, il permet d'affiner la recherche d'articles spécifiques (quand ils existent). Dans le cas d'une thèse CIFRE, ces particularités constituent une excellente passerelle entre recherche théorique et pratique.

Dans la suite nous donnerons une liste non exhaustive des particularités les plus importantes, de notre point de vue, de l'unité de fabrication 300mm de Crolles.

2.3.1 Système de transport unifié

On a vu dans la section 1.3.3 que les systèmes unifiés étaient plus performants que les systèmes non-unifiés [7], [53]. Pour rappel, l'unité de fabrication 300mm de Crolles étudiée dans cette thèse est entièrement unifiée, donc les véhicules peuvent se déplacer partout. Avant de montrer les conséquences d'une telle particularité, quelques définitions s'imposent.

Définition d'un parking. Un parking est un emplacement de l'unité de fabrication (représenté physiquement par un code-barres), où les véhicules viennent se positionner dans l'attente d'une nouvelle requête de transport. Si possible, il doit être le plus proche possible d'un maximum de point de chargement/déchargement. L'emplacement des parkings est stratégique.

Décrivons à présent, différentes particularités liées à la nature unifiée du système de transport en général et, de Crolles 300 mm en particulier :

- ***Gestion de collisions aux intersections*** : dans les systèmes unifiés, l'embranchement des rails finit par se rejoindre à un moment ou à un autre, afin de permettre aux véhicules de pouvoir se déplacer en n'importe quels points de la salle. Les lieux où les embranchements se rejoignent sont appelés *intersections convergentes*. Différents véhicules venant de divers embranchements peuvent se croiser aux intersections convergentes. Il est donc nécessaire de gérer l'ordre de passage des véhicules, un peu comme il y a des règles de priorités pour une entrée d'autoroute. Cette gestion de la priorité réduit les risques de collisions aux intersections convergentes car l'AMHS ne doit jamais être un élément bloquant.
- ***Gestion de l'emplacement des parkings*** : dès l'émission d'une requête de transport, le temps d'attente du lot doit être idéalement le plus petit possible. Dans les systèmes unifiés, les véhicules peuvent aller partout dans l'unité de fabrication, ce qui veut dire qu'ils peuvent être proches, comme très éloignés de l'emplacement de la requête. Il est donc important que le véhicule se retrouve au plus près possible des requêtes de transport. Mais, avec environ 350 équipements répartis en 10 zones, le problème est le suivant : *d'où viendra la prochaine requête de transport ?* De la zone 2, 7, ou 9 ? etc. Les parkings fournissent une solution acceptable (voir définition ci-dessus). Cependant, il ne suffit pas que le parking soit le plus près possible du maximum de points, il doit gêner le moins possible le passage des autres véhicules. Le choix de l'emplacement des parkings est un problème dont l'optimal n'est pas facile à trouver.
- ***Gestion des congestions*** : une congestion est une forte intensité de transport en un nœud du réseau, comme par exemple, un embouteillage routier. Dans le cas des semi-conducteurs, par exemple, les équipements en traitement thermique, à cause de leur temps de traitement très long (jusqu'à 4 heures et plus), attendent le plus souvent d'avoir un maximum de lots appelés *batch* pour les traiter. Une fois que les lots ont terminé leurs traitements, ils émettent en *même temps*, le même nombre de requêtes de transport que de lots. Mais, puisque le système est unifié, les véhicules s'ils sont disponibles, peuvent tous venir en même temps. Cela engendre la formation d'une file et augmente le risque de congestion du trafic. Il est donc nécessaire de bien gérer ce risque.

2.3.2 Gestion des requêtes de transport par « service minimum » garanti

Il y a environ 350 équipements de production pour 26 véhicules. Si tous les équipements émettent une requête de transport en même temps, elles ne pourraient pas toutes être servies. Donc, il y aura nécessairement inadéquation, à un moment ou à un autre, entre ressources de transport et besoins de la production.

Le rôle d'un système de transport est de mettre en adéquation les ressources de transport et les besoins de la production en transport. Les approches traditionnelles (par exemple [41] et [70]) imposent d'associer la bonne politique d'ordonnancement des requêtes de transport des véhicules à celle de routage. La difficulté avec ces approches réside dans le nombre de combinaisons possibles entre politiques de routage et d'ordonnancement des requêtes de transport. Il est donc nécessaire de réduire le nombre de combinaisons.

Quelque soit la politique trouvée, un problème majeur demeure, c'est celui de la disponibilité des véhicules. Une approche consiste à conserver constamment un certain nombre de véhicules dans une zone donnée, afin d'être réactif en cas de requête de transport. Une solution industrielle a été développée par l'entreprise MURATEC, par l'introduction de paramètres de contrôles du nombre de véhicules présents dans une zone tels que le *Low Water Mark* (LWM) et le *High Water Mark* (HWM). Ces paramètres sont des éléments essentiels dans la compréhension du système de Crolles 300mm.

Low Water Mark (LWM) : c'est le nombre minimum de véhicules *vides et non assignés* que la zone vise à conserver en permanence. Par exemple, LWM=2 signifie que la zone vise à conserver en permanence deux véhicules vides et non assignés dans la zone. La définition impose que **LWM \geq 0**.

Si le nombre de véhicules vides et non assignés est strictement inférieur à LWM, le contrôleur de véhicules fait appel à des véhicules extérieurs pour compenser ce manque. Il va tester toutes les zones (suivant un ordre prédéfini) afin de trouver s'il y a des véhicules disponibles. Lors du test d'une zone, le système va évaluer par anticipation, l'état de la zone si l'on retire un véhicule et prendre une décision. Ce

dernier retranche virtuellement, un véhicule du total des véhicules de la zone cible et prend une décision en fonction du résultat. Si le nombre de véhicules restants est supérieur ou égal à LWM alors, la zone peut céder le véhicule. Sinon, le nombre de véhicule restant est strictement inférieure au LWM et la zone cible ne peut pas céder de véhicule. Dans ce dernier cas, le contrôleur de véhicules s'adressera à une autre zone et recommencera l'algorithme décrit ci-dessus. Lorsque toutes les zones ont été testées sans succès, on recommence avec la première zone et ainsi de suite jusqu'à trouver un véhicule libre. Cet algorithme nous montre que l'on ne peut pas « appeler » de véhicules à partir d'une zone où le nombre de véhicules est inférieur au LWM.

Trouver la valeur du LWM, c'est définir le nombre de véhicules correspondant aux tâches moyennes de l'activité de la zone. *Le LWM permet de repositionner rapidement les véhicules afin qu'ils soient au plus près du besoin.* C'est ce que nous avons appelé **service minimum**.

High Water Mark (HWM) : il représente le nombre maximal de véhicules *vides et non assignés* autorisés à rester dans la zone. Par exemple, HWM=3 signifie que pas plus de trois véhicules vides et non assignés ne sont autorisés à rester la zone. La définition impose que $\mathbf{HWM} \geq \mathbf{0}$.

Si le nombre de véhicules vides et non assignés est strictement supérieur à HWM, le contrôleur de véhicules « pousse » l'excédant de véhicules vers des zones extérieures. Le choix des zones où envoyer l'excédant de véhicules se fait selon un ordre prédéfini. L'algorithme de rejet des véhicules s'effectue par anticipation comme suit : un véhicule est virtuellement ajouter à la première zone cible. Si le nombre total de véhicules de cette zone cible est inférieur ou égale à son HWM alors, le véhicule en « trop » de la zone source est autorisé à se rendre dans la zone cible. Sinon, le nombre total de véhicules de la zone cible est strictement supérieur à HWM. Dans ce dernier cas le système recherche parmi sa liste de zones prédéfinies, la prochaine zone à tester. Lorsque toutes les zones ont été testées sans succès, on recommence avec la première zone et ainsi de suite jusqu'à trouver un véhicule libre. Cet algorithme nous montre que l'on ne peut pas « pousser » de véhicules vers une zone où le nombre de véhicules est inférieur au LWM. *Le HWM permet d'éviter les problèmes*

de congestion.

Propriétés du *Low Water Mark*

Nous allons à présent présenter quelques propriétés du paramètre *Low Water Mark*. L'intérêt de cette étude est de permettre une meilleure compréhension sur le fonctionnement du paramètre.

P1 : $\mathbf{LWM} \leq \mathbf{HWM}$, car par définition le LMW représente un nombre minimal et le HWM un nombre maximal de véhicules dans une zone donnée.

P2 : $\mathbf{LWM} \geq 1$. Le cas $\mathbf{LWM} = 0$ signifie que la zone ne requiert aucun véhicule, donc pas de transport au départ de cette zone. Ce cas présentant très peu d'intérêt dans la pratique, nous l'écartons.

P3 : $\sum_{i=1}^{\text{Nombre de zones}} \mathbf{LWM}_i \geq \text{Nombre de zones}$, où \mathbf{LWM}_i représente le *Low Water Mark* de la zone i . P3 est obtenu à partir de P2 et en faisant une somme membre à membre sur le nombre de zones.

P4 : $\sum_{i=1}^{\text{Nombre de zones}} \mathbf{LWM}_i \leq (2 * \text{Nombre de zones})$ qui correspond à **une condition pour garantir une faible dispersion** dans le temps pour remplacer un lot par un autre sur un équipement.

Nous allons prouver la propriété P4 en deux temps. Nous allons d'abord montrer (par l'absurde) qu'il est nécessaire d'avoir la condition $\sum_{i=1}^{\text{Nombre de zones}} \mathbf{LWM}_i \leq (\text{Nombre de vehicules})$ pour que le système puisse se rééquilibrer tout seul. Supposons que

$\sum_{i=1}^{\text{Nombre de zones}} \mathbf{LWM}_i \geq (\text{Nombre de vehicules})$ (illustré en exemple par la figure 2.8). Mais sous cette condition le système de rééquilibrage des véhicules ne fonctionne pas, car il n'est pas possible d'« appeler » des véhicules. Rappelons qu'il est interdit d'appeler un véhicule d'une zone si cet appel fait passer le nombre total de véhicules en dessous de la valeur du LWM. Donc, on a nécessairement

$\sum_{i=1}^{\text{Nombre de zones}} \mathbf{LWM}_i \leq (\text{Nombre de vehicules})$.

Il reste à déterminer par zone, le nombre de véhicules, nécessaire pour remplacer un lot par un autre sur un équipement de production. Prenons comme référence la zone où se trouve l'équipement. Le remplacement d'un lot par un autre comprend

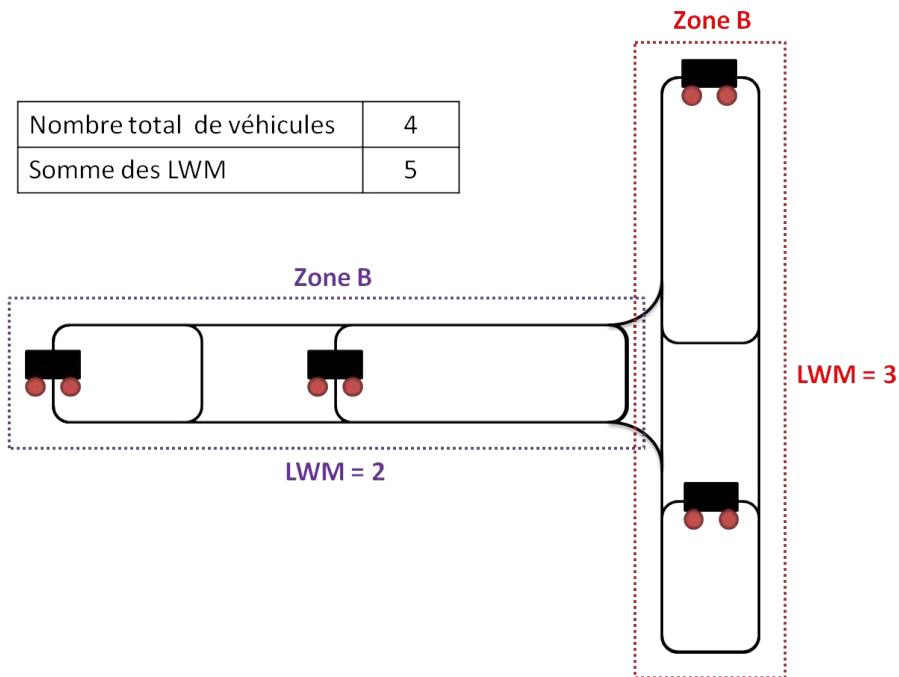


FIGURE 2.8 – Cas où la somme des LWMs dépasse le nombre de véhicules.

deux parties : le déchargement du lot et le chargement du nouveau lot. Nous allons à présent montrer qu'il faut au moins deux véhicules si on veut un temps de remplacement le plus court et si on veut réduire la dispersion de ce temps.

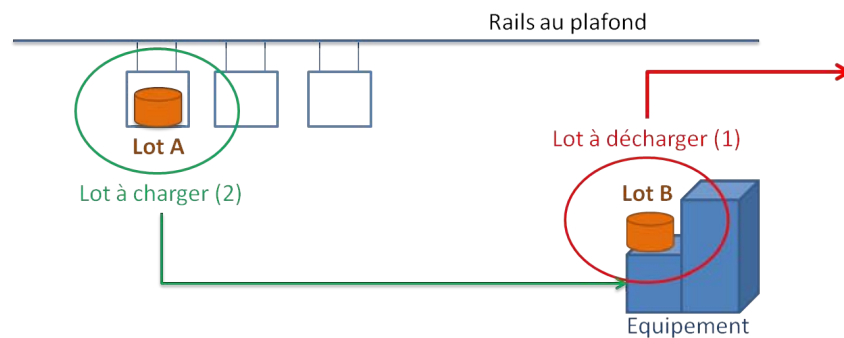
- **1 véhicule pour le déchargement du lot** : l'assignation d'une requête de transport ne peut se faire qu'à l'intérieur de la zone. Il faut donc au moins un véhicule présent dans la zone, sinon on augmente le temps d'assignation de la tâche (augmentation correspondant au temps de recherche d'un véhicule).
- **1 véhicule pour le chargement du nouveau lot** : nous distinguerons deux cas pour le chargement du lot. Dans le premier cas, le lot à charger se trouve dans la même bay que l'équipement. Un autre véhicule pour assurer cette tâche (la prochaine étape du déchargement étant susceptible d'être dans une autre zone, le véhicule correspondant au déchargement peut sortir de la zone). Dans le second cas, le lot à charger ne se trouve pas dans la même zone que l'équipement. Dans ce cas, la zone où se trouve le lot doit posséder un véhicule prêt à quitter la zone. Cette dernière remarque appliquée à notre

zone de référence, recommande la présence d'un second véhicule pour assurer le chargement d'un lot présent dans la zone ou celui d'un autre à destination d'une autre zone.

En conclusion on a : *Nombre de vehicules* = 2 * *Nombre de zones* **si on veut assurer un temps changement de lot le plus court possible**. Appliqué à notre inégalité on obtient donc :

$$\sum_{i=1}^{\text{Nombre de zones}} LWM_i \leq (2 * \text{Nombre de zones}).$$

Cette relation **n'est pas une relation absolue**, elle donne une condition nécessaire pour réduire la dispersion du temps de changement d'un lot sur un équipement, dans le cas où l'assignation de tâche ne peut se faire que dans la zone.



- **Tout chargement d'un lot A est précédé par un déchargement d'un lot B.**
- Le lot A a besoin d'un véhicule et le lot B à besoin d'un autre véhicule.

FIGURE 2.9 – Recommandation d'avoir deux véhicules par zone.

2.3.3 Repositionnement des véhicules

Comme dit dans la section précédemment, avec 350 véhicules et 26 véhicules, il y aura nécessairement un déséquilibre de l'activité de production. Ce déséquilibre se traduit par un temps d'assignation élevé. Soulignons dès à présent que **l'assignation de ressources de transport à Crolles 300 mm n'est possible qu'à l'intérieur de la zone d'où part la requête**. Autrement dit, *un lot ne peut pas « appeler » un*

véhicule qui n'est pas dans sa zone. Nous y reviendrons plus en détails par la suite. Dans le cas de Crolles 300mm, le temps d'assignation se décompose en deux parties :

- Ta1 - le temps pour trouver la ressource,
- Ta2 - le temps pour amener la ressource dans la zone considérée.

Si le niveau d'encours est important, c'est-à-dire le nombre de lot dans l'unité de fabrication, c'est le temps pour trouver la ressource (Ta1) qui croît. Si, par contre, la zone de la requête est éloignée des ressources disponibles, c'est le temps pour amener la ressource dans la zone de destination (Ta2) qui croît. La problématique est que les ressources sont soit trop loin, soit elles ne sont pas disponibles. Le repositionnement des véhicules est une solution à ce problème. Il consiste à choisir des emplacements stratégiques (parkings) où les véhicules doivent venir se replacer afin de pouvoir répondre le plus rapidement possible aux requêtes de transport. Une fois les parkings identifiés, il faut rendre le plus efficient possible, le temps de recherche et de déplacement des véhicules vers la zone de la requête.

Deux grands acteurs interviennent dans le repositionnement des véhicules à ST-Microelectronics Crolles :

- le **contrôleur de véhicules** qui assure une gestion globale des véhicules et,
- les **contrôleurs locaux de véhicules** qui assurent une gestion locale des véhicules dans chaque zone.

Le premier organise le transport et gère les besoins des différentes zones en termes de véhicules. C'est lui qui prend les décisions. Il est en charge de l'équilibre entre ressources de transport et requêtes de transport entre les différentes zones. Le second informe fréquemment (toutes les 5 secondes) le premier sur la situation de sa zone et est responsable de l'assignation des ressources aux requêtes des lots. Hors mis l'assignation d'une requête à un véhicule, il ne peut prendre aucune décision.

Fonctionnement du contrôleur local de véhicules. Il a principalement un rôle d'informateur. Un tableau est associé à chaque contrôleur local de véhicules. Un exemple de ces tableaux est présenté par le tableau 2.2.

- **Numéro de zone** : un numéro de zone est associé à chaque zone. Cela permet une meilleure gestion informatique.
- **LWM et HWM** : le tableau contient les informations sur valeur du *Low*

Numéro de zone	LWM	HWM	Zones Send	Zones Receive
5	1	4	2, 4, 7, 9	11, 7, 9

TABLE 2.2 – Éléments d'information dont dispose un contrôleur local de véhicules.

Water Mark et du *High Water Mark*.

- **Zones Send** : ce sont les numéros (dans l'ordre hiérarchique) des zones où l'on peut envoyer le surplus de véhicules. Pour rappel, le surplus de véhicules correspond au nombre de véhicules vides et non assignés au-dessus de la valeur du HWM. Dans l'exemple du tableau 2.2, si le HWM de la zone 5 est atteint, le surplus de véhicules est « poussé » vers la zone 2. Si le HWM de la zone 2 est atteint, on passe à la zone 4, puis à la zone 7, puis à la zone 9 si nécessaire. Si on ne trouve aucune zone, on revient au début et on recommence jusqu'à ce que l'on trouve une zone pouvant accepter le véhicule.
- **Zones receive** : Ces numéros représentent, dans l'ordre, les zones d'où l'on peut faire un appel de véhicules si un besoin se fait sentir. Ce besoin se traduit, soit par le fait que le nombre de véhicules vides et non assignés est en dessous du LWM, soit par le fait que la somme du nombre de requêtes de transports de la zone dépasse le nombre total de véhicules de la zone (en charge ou vides). Dans l'exemple du tableau 2.2, si le LWM de la zone 5 n'est pas atteint, les véhicules manquants seront appelés à partir de la zone 11. S'il n'y en a pas assez ou que le LWM de cette zone ne le permet pas, alors les véhicules manquants sont appelés à partir de la zone 7, puis de la zone 9. S'il n'y a toujours pas assez de véhicules, on recommence jusqu'à ce que l'on trouve une zone pouvant céder un véhicule. Ce temps de recherche augmente le temps d'assignation des lots (Ta1).

Fonctionnement du contrôleur de véhicules. C'est lui qui prend les décisions. Toutes les 5 secondes chaque contrôleur local de véhicules lui envoie des informations sur de l'état de sa zone. Seul le contrôleur de véhicules peut faire appel aux véhicules extérieurs à une zone. Les contrôleurs locaux ne peuvent pas le faire. Cependant, si un véhicule vide et non assigné traverse une zone ayant une requête de transport

en attente, alors il peut être réquisitionné par le contrôleur local de véhicules (assignation). La décision du contrôleur de véhicules d'envoyer ou d'appeler un véhicule dans une zone ou vers une zone se font en fonction des tableaux associés à chaque contrôleur local de véhicules (tableau 2.2). Les tableaux 2.3 et 2.4 expliquent sur quoi sont basées les décisions prise par le contrôleur de véhicules.

	HWM	LWM	Nombres de lots	Nombre de véhicules	
				Chargés	Au repos
Zone 1	10	3	1	1	0
Zone 2	10	2	6	2	3
Zone 3	9	2	1	1	10

TABLE 2.3 – Éléments d'information dont dispose le contrôleur de véhicules.

	Target(1)	Target(2)	...	Target(N-1)	Target(N)
Zone 1	Zone 2	...	Zone 3	Zone N-1	Zone N
Zone 2	Zone 4	...	Zone N	Zone 2	Zone N-1
Zone 3	Zone N	...	Zone N-1	Zone 5	Zone 8

TABLE 2.4 – Table *Zone send*, *Zone receive*.

Observation 1 : la première ligne du tableau 2.3 (« Zone 1 ») représente une situation où, LWM=3 avec une requête de transport (1 lot), un véhicule transportant un lot (1 véhicule chargé) dans la zone et aucun véhicule libre et non assignés (0 véhicule au repos). Au total pour LWM=3, il y a 1 véhicule dans la zone (Nombre de véhicules Chargés + Nombre de véhicules au repos), ce qui correspond à un déficit de 2 véhicules.

Décision de l'observation 1 : appel de deux véhicules pour équilibrer le LWM. L'appel se fait en fonction du tableau associé à la zone dont un exemple est donné par le tableau 2.2.

Observation 2 : la seconde ligne du tableau 2.3 (« Zone 2 ») représente une situation où, LWM=2 avec six requêtes de transport (6 lots), deux véhicules transportant

un lot (2 véhicules chargés) dans la zone et trois véhicules libres et non assignés (3 véhicules au repos). Au total, il y a 5 véhicules dans la zone (Nombre de véhicules Chargés + Nombre de véhicules au repos) pour $LWM=2$. Bien que le nombre de véhicules soit supérieur à LWM , il y a un déficit de véhicules car il y a 5 véhicules mais 6 requêtes de transport, or les lots sont prioritaires.

Décision liée à l'observation 2 : appel d'un véhicules pour renforcer la capacité de transport.

Observation 3 : la troisième ligne du tableau 2.3 (« Zone 3 ») représente une situation où, $LWM=2$ avec 1 requête de transport (1 lot), 1 véhicule transportant un lot (1 véhicule chargé) dans la zone et dix véhicules libres et non assignés (10 véhicules au repos). Au total, il y a 11 véhicules dans la zone (Nombre de véhicules Chargés + Nombre de véhicules au repos) pour $LWM=2$. Le nombre de véhicules est supérieur à $LWM=2$ mais aussi à $HWM=9$. Il y a 2 véhicules en trop dans la zone.

Décision liée à l'observation 3 : expulsion de deux véhicules pour désengorger la zone.

Bilan sur l'utilisation des LWMs et des HWMs. L'étude du repositionnement des véhicules permet d'obtenir plusieurs informations sur le *Low Water Mark* et le *High Water Mark*.

Les avantages.

- mettre en adéquation des ressources de transport avec les besoins de productions,
- rééquilibrer un système qui se déséquilibre lui même,
- réduire le temps d'assignation des véhicules.

Quelques inconvénients.

- **le système immobilise parfois inutilement les véhicules** dans la zone pour qu'ils soient prêts à répondre aux requêtes de transport.
- **une zone peut se « vider » de ses véhicules**, ce qui augmente le temps d'assignation des requêtes de transport à l'origine de cette zone. Ainsi, le

LWM fournit une condition nécessaire mais pas suffisante pour réduire le temps d'assignation.

- **il peut biaiser certains indicateurs.** Par exemple, le taux d'occupation des véhicules fournira des résultats inexploitable, car on force les véhicules à attendre (à vide) dans la zone.

2.3.4 Regroupement des emplacements de stockage

Nous avons annoncé dans le chapitre 1 que l'on pouvait classer les emplacements de stockage en deux groupes. Un groupe en fonction du nombre d'emplacements de stockage (capacité de stockage) et un autre en fonction du temps d'accès. En général, *le temps d'accès au stockage est inversement proportionnel à la capacité de stockage*. Cette remarque nous permettra de n'utiliser qu'un seul critère pour parler des différents types de stockage puisque, les deux critères sont reliés. On rappelle qu'un stockeur est un container géant renfermant plusieurs emplacements unitaires de stockage et un robot (bras mécanique servant à déplacer les lots).

Ainsi, sur le critère de la capacité de stockage, il existe à Crolles 300 mm deux types d'emplacement de stockage unitaire comme représenté sur la figure 2.10

- **Side Track Buffers** ou *STB*. Stockeurs unitaires aérien placés sur le côté des rails. Ils présentent deux avantages majeurs : (1) un *temps d'accès cinq fois plus rapide* qu'un stockeur et (2) très peu d'électronique embarquée, ce qui fait qu'*ils tombent rarement en panne*. Cependant, leur nombre est limité par la longueur des rails de la zone.
- **Under Track Buffer** ou *UTB*. Stockeurs unitaires aériens placés juste en-dessous des rails. Ils ont un *temps d'accès plus rapide que celui des STB* mais, sont plus proches du sol que les STB. Il est souvent difficile de placer des équipement en dessous des UTB, ce qui génère une perte de place au sol. Ils ne conviennent pas forcément partout.

Si on pouvait, on mettrait les équipements de production côte à côte, mais pour des raisons pratiques (maintenances, réduction des temps entre opérations, etc.) ce n'est pas le cas. En acceptant de ne pas mettre les équipements les uns à côté des autres, on accepte, d'une certaine façon, l'attente (temps de trajet plus long). Mais

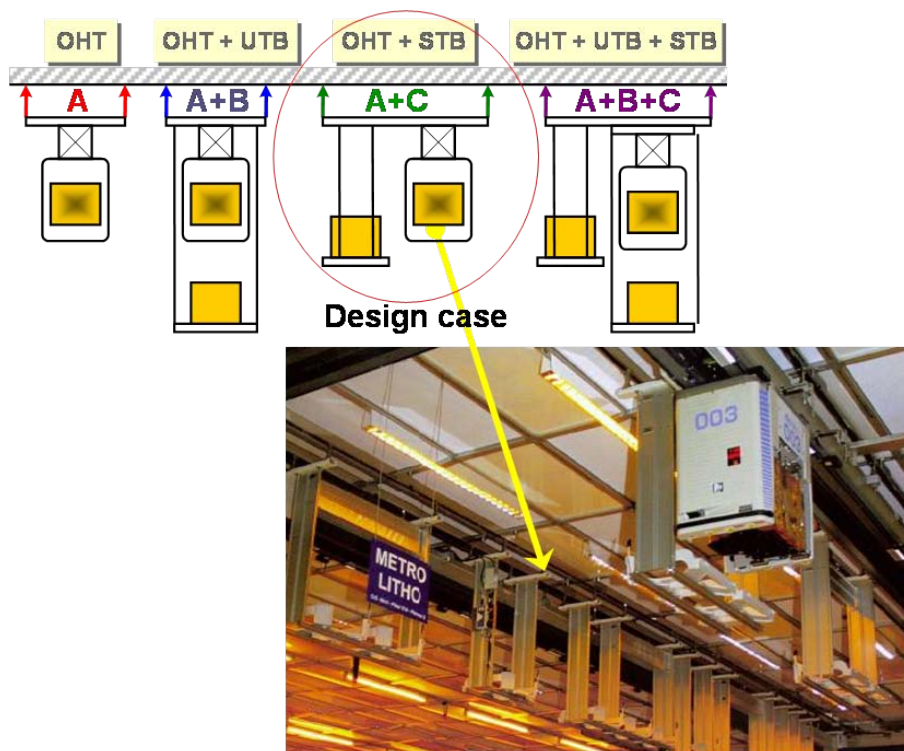


FIGURE 2.10 – Différents types de stockage unitaire.

toutes les attentes ne sont pas liées à la position géographique de l'équipement. Par exemple, attente pour accéder au port de chargement, attente pour que la bonne « recette » soit chargée, attente due à la maintenance, etc.

En général, un lot attend pour passer sur un équipement de production. Pour mieux gérer cette attente, chaque équipement de production à Crolles 300mm est associé à une et une seule *zone* de stockage mais, une zone de stockage peut être associée à un ou plusieurs équipements de production.

La zone de stockage doit posséder les propriétés suivantes : (p1) **être au plus près des équipements qui l'utilisent**, sachant qu'ils ne sont pas tous physiquement aux mêmes endroits. Cette propriété est particulièrement importante pour les équipements qui nécessitent un temps de changement de lots très court (équipements de mesures par exemple), (p2) **être de capacité suffisante**, dans le sens qu'elle ne doit pas se remplir trop vite et (p3) **avoir un temps d'accès le plus court possible**. Par exemple, il ne doit pas contribuer à augmenter le temps de

changement du lot sur un port de chargement.

Les emplacements de stockage unitaires respectent la propriété (p3) mais pas (p2). Pour y remédier, à Crolles 300mm, les ingénieurs ont regroupé les emplacements de stockage unitaires pour former des « stockeurs par défaut ». Chaque équipement de production possède son *stockeur par défaut*. Le problème de dimensionnement des stockeurs par défaut reste un problème ouvert. À cause de la propriété (p3), un stockeur est rarement le *stockeur par défaut* d'un équipement de production.

Que faire si le *stockeur par défaut* est plein? Toujours à cause de la propriété (p3), on préférera un autre *stockeur par défaut* comme espace de stockage, plutôt qu'un stockeur qui est plus long d'accès. Cet autre *stockeur par défaut* qui est une alternative de stockage en cas de remplissage du *stockeur par défaut* initial est appelé *alternate*.

Ainsi, tout équipement de l'unité de fabrication possède son stockeur par défaut et chaque stockeur par défaut une liste d'espace alternatif de stockage. Notons que parce qu'ils possèdent un temps d'accès plus long, les stockeurs dans une liste d'*alternate* se trouvent toujours en fin de liste et ne sont jamais *défaut stockeur* d'un équipement. La figure 2.11 illustre la correspondance entre équipements de production, stockeurs par défaut et espaces alternatifs de stockage (*alternates*).

Nous terminerons cette section par la description d'une difficulté liée aux regroupements d'emplacements de stockage unitaire à Crolles 300mm : le **dimensionnement des stockeurs par défaut**. Ce problème concerne la capacité (en nombre d'emplacement) de la zone de stockage. La difficulté est de déterminer un nombre « optimal » d'emplacement de stockage tel que : s'il n'y en a pas assez, le stockeur par défaut se remplira très vite ainsi que les alternates. S'il y en a trop, il faut ajouter au temps d'accès à l'emplacement de stockage, le trajet des véhicules. On peut alors atteindre voire, dépasser le temps d'accès à un stockeur.

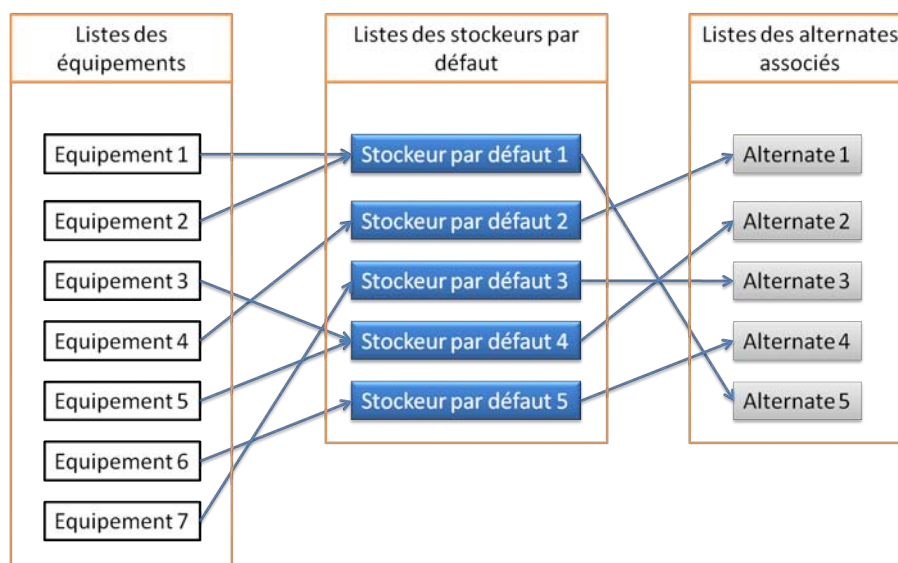


FIGURE 2.11 – Liens entre équipements, stockeurs par défaut et *alternates*.

2.4 Conclusion

Dans ce chapitre nous avons présenté les « rouages » du transport au sein de l'unité de fabrication de Crolles. La présentation du fonctionnement de cette unité de fabrication permet d'affiner la problématique industrielle de cette thèse. Tandis que les particularités industrielles montrent, d'une part la complexité de notre problème et d'autre part les besoins industriels.

Nous avons répondu aux deux questions suivantes :

- 1 - *Comment fonctionne l'usine de fabrication de STMicroelectronics à Crolles ?* Le système fonctionne en mode semi-automatique.
- 2 - *Quelles sont les particularités de cette usine ?* Il y en a plusieurs mais les quatre plus importantes de notre point de vue sont : le **système unifié** qui utilise la politique de **service minimum** pour le dispatching des véhicules. Une **politique particulière de repositionnement** des véhicules qui permet la réduction du temps d'assignation et le **regroupement d'espaces de stockage unitaires** en stockeur par défaut.

On peut constater que ces particularités industrielles touchent les problématiques de transport (système unifié, repositionnement des véhicules), de production (*service*

minimum) et de stockage (regroupement des *emplacements de stockage unitaire*). Nous allons, dans la suite, montrer que le transport, la production et le stockage sont liés et que toute étude n'incluant pas ces trois éléments pourrait conduire à des résultats donc l'impact décisionnel peut être catastrophique.

Comment allons-nous procéder ? Nous allons nous servir d'une revue de la littérature et d'un plan d'expérience pour étayer nos affirmations. Avant de faire le plan d'expérience, nous allons d'abord analyser les relations entre les trois éléments (transport, production et stockage). Cela va ensuite nous permettre de nous positionner dans la littérature. C'est l'objet du chapitre suivant.

Chapitre 3

Positionnement dans la littérature

Ce chapitre traitera :

- De la dépendance entre transport automatisé et configuration du système (section 3.2),
- De la dépendance entre transport automatisé et production (section 3.3),
- Du stockage, intermédiaire entre le transport et la production (section 3.4) et,
- De notre positionnement dans la littérature scientifique (section 3.5).

Nous répondrons aux questions suivantes : *comment sommes-nous positionnés par rapport à la littérature ? Quelles sont les originalités de cette thèse ?*

3.1 Résumé du chapitre

La puissance des ordinateurs a longtemps été la principale limite à l'utilisation de modèles complexes de simulation. En conséquence, les modèles simplifiés se sont imposés comme une solution. Cependant l'évolution de la complexité des produits, donc de leur fabrication, rend de plus en plus difficile une description analytique complète des modèles qui les décrivent. En même temps on note que le développement de la puissance des ordinateurs a considérablement réduit les temps de calcul. Malgré cette évolution, les temps de construction du modèle et de calcul sont souvent encore trop longs en particulier pour des décisions au niveau opérationnel [90].

La « vie » d'un lot est une succession de trois états : en transport, en production ou en stockage . Il en va de même pour la simulation des unités de fabrication de semi-conducteurs. Donc si l'on veut avoir le rendu le plus fidèle possible en simulation, il est essentiel de modéliser ces trois éléments fondamentaux que sont le transport, la production et le stockage. Pourtant, dans la littérature la majorité des simulations ne considère que soit un ou deux des éléments fondamentaux, soit les couples : « transport, production », « transport, stockage » ou « production, stockage ». Soulignons que nous avons rencontré très peu de d'articles qui traitent de la dernière combinaison. *Et, quand le couple « production, stockage » est considéré, les capacités de stockage sont généralement considérées comme étant infinies.* De plus, *les hypothèses simplificatrices font qu'un certain niveau de détails n'est pas pris en compte dans les modèles.* Nous montrerons donc les limitations de telles hypothèses sur les résultats et décisions pouvant en découler.

Nous proposons dans cette thèse de travailler avec les trois acteurs en même temps, à savoir *transport + production + stockage* afin de modéliser *le plus fidèlement possible* l'unité de fabrication. A notre connaissance, il n'existe pas dans la littérature une telle approche qui intègre en plus les complexités réelles des systèmes de stockage. Nous positionnerons notre travail par rapport à un recueil de la littérature entre 1999 et 2008 [37].

3.2 Corrélation entre transport automatisé et configuration du système

Le monde dans lequel chacun vit dépend de la façon de le concevoir.

Arthur Schopenhauer

Corrélation économique. Le coût unitaire d'une unité de fabrication 300mm est environ 45% plus élevé qu'une unité de 200mm [3]. Une simple amélioration de 1 à 2% du rendement d'une unité de fabrication 300mm génère des profits de plusieurs millions de dollars [19]. Etant donné que les performances du système AMHS¹ sont liées à sa configuration [26], améliorer la configuration de l'unité de fabrication permet de faire des gains de l'ordre de 10 à 30% [67]. L'investissement est donc considérablement amorti. Mais, améliorer la configuration du système seul ne suffit pas. En effet, le système AMHS n'offre pas les mêmes performances selon que l'on soit en configuration unifiée ou non-unifiée [7], les systèmes unifiés ayant de meilleures performances [7]. Mais, dans [53], les auteurs ont montré que le basculement vers les systèmes unifiés ne pouvait se faire **sans dégradation du niveau d'encours**. Ils ont en outre montré que si $(TU + TL) \leq (LP - 1) \times CT$, avec TU = durée de déchargement du lot, TL = durée de chargement + temps du choix de lot à livrer, CT = temps de cycle de l'équipement et LP = nombre de port de chargement, alors l'AMHS n'affectait pas le temps de cycle des lots. Faire de la livraison de machine à machine n'est pas l'objectif en soi. L'objectif est la **réduction du temps de cycle tout en conservant le niveau d'encours de la fab.**

Corrélation technique. L'AMHS est devenu un élément incontournable de l'industrie des semi-conducteurs, en particulier depuis le passage du diamètre des wafers de 200 à 300mm [14]. La réalité est que la compétition entre fabricants exige à la fois : des équipements de production de très haute performance utilisant les dernières technologies, et des temps de livraison courts utilisant des technologies de transport automatisé de pointe. Pour le lecteur désireux d'en savoir plus, [19] et [92] traitent de l'importance des technologies dans la compétition industrielle.

1. Automated Material Handling System

Cependant, on peut observer que certaines technologies semblent atteindre leurs limites. Par exemple, dans [54], les auteurs ont montré que l'amélioration du temps de transport semble atteindre ses limites physiques. Plus généralement, dans [94], c'est l'existence d'*idées reçues*² sur les performances des systèmes de transport automatisé qui est dénoncée. Il semble donc judicieux d'incorporer, lors de la conception de l'unité de fabrication (niveau stratégique), la configuration de l'unité de fabrication et du système de transport automatisé comme préconisé dans [25], [13].

Conclusion. D'une part, l'amélioration de la configuration de l'unité de fabrication permet de générer des bénéfices très importants et, d'autre part, la dénonciation des idées reçues sur le système de transport automatisé montre qu'il ne peut seul améliorer les performances du système. Donc, *les performances du système de transport automatisé sont fortement dépendantes de la configuration de l'unité de fabrication.*

L'union fait la force. Nous pensons que, pour améliorer les performances de l'unité de fabrication, la prise en compte du système de transport automatisé devrait fournir de meilleures performances que chaque système pris séparément. C'est cette idée qui sera à la base de notre approche.

Nous allons montrer dans la suite que la prise en compte du triplet (transport, production et stockage) permet d'avoir de meilleurs résultats en termes de modélisation et de simulation des systèmes de fabrication de semi-conducteurs.

2. par exemple, avoir des véhicules plus rapides réduit le temps de livraison ou la livraison de machine à machine éliminera les files d'attente, etc.

3.3 Corrélation entre transport automatisé et production

« Chaque système est vrai par ce qu'il propose et faux par ce qu'il exclut. »

Roger Caillois

Introduction

Dans la littérature, le transport est considéré comme n'apportant aucune valeur ajoutée, mais il est souvent incontournable [32] (enchaînement des opérations de production).

Le transport automatisé des lots permet de respecter les limites de tolérance de chocs et de vibration [3], [87] et réduit la contamination des wafers due à la manipulation humaine [77]. *Il vise à amener le bon lot, au bon endroit et au bon moment.* En outre, l'augmentation du poids des lots fait qu'ils ne sont plus ergonomiquement adaptés au transport manuel. Il est donc très important d'améliorer les performances des systèmes de transport automatisé.

Un des leviers d'action pour améliorer les performances du transport est l'utilisation de politiques « optimisées » de gestion des véhicules. Cependant, il n'y a pas forcément de politique générique de transport surpassant les autres [20]. L'excellente revue de littérature de [69] décrit différentes politiques de gestion du transport, mais *très peu d'articles se sont intéressés à l'intégration des contraintes de production* [42].

En se référant aux problématiques traitées dans la littérature, on peut citer : la priorité et la destination des lots, l'ordonnancement et le regroupement de lots, la variabilité des temps de process, le nombre de ports de chargement et le nombre de plaquettes par lot.

Modélisation du système de transport, avec fortes hypothèses simplificatrices sur la production

Le tableau 3.1 résume cette section en présentant différentes études bibliographiques avec de fortes hypothèses simplificatrices sur la production dans la modéli-

sation d'une unité de fabrication.

Auteurs	Année	Hypothèses sur la production
Dong-Yeh <i>et al.</i> [22]	2002	Pas de prise en compte de la priorité et de la destination des lots.
Luo <i>et al.</i> [62]	2005	Variabilité des temps de process non considérée
Da-Yin and Hsien-Sheng [21]	2004	Pas de prise en compte du temps de réallocation
Montoya [70]	2005	Nombre limité de ports de chargement non considéré
Kiba <i>et al.</i> [42]	2008	
Paprotny <i>et al.</i> [76]	2000	Non prise en compte des caractéristiques des équipements
Taush and Hennessey [88]	2002	
Shah <i>et al.</i> [86]	2006	
Campbell <i>et al.</i> [15]	1999	Non prise en compte de la taille des lots
Hideto [33]	2006	
Greenberg [30]	2007	
Adusumilli and Wright [2]	2007	
Nazzal and McGinnis [74]	2007	Non prise en compte du temps de non production

TABLE 3.1 – Revue de la littérature : modélisation du transport avec fortes hypothèses simplificatrices sur la production.

Pas de prise en compte de la priorité et de la destination des lots. Dans [22], les auteurs ont comparé deux logiques de gestion des véhicules : PULL et PUSH. Dans le concept PULL, l'AMHS est conçu de manière à connaître l'identité du lot et sa destination. Tandis que, dans le concept PUSH, le transport du lot est prédéfini dans l'ordinateur central, et est déclenché par le système AMHS. L'inconvénient majeur du concept PUSH est que le lot est transporté au plus proche du prochain équipement, *sans prise en compte de la priorité des lots*. Les auteurs ont démontré la supériorité de la logique PULL sur la logique PUSH. Ils ont relevé une amélioration variant entre 7 et 9% de la capacité de transport et une réduction de 30% du temps d'attente des lots.

Mais, connaître l'identité et la destination du lot ne suffit pas, *il faut être le plus*

proche possible du lot le plus prioritaire à transporter (respect des délais clients, du temps de contamination etc.).

Négligence de la variabilité des temps de fabrication. Dans [62], les auteurs ont traité des problématiques de routage. Ils ont proposé un algorithme évolutionnaire (colonies de fourmis) afin de déterminer en dynamique, le plus court chemin entre deux points. *Les durées opératoires sur les équipements sont considérées fixes.* Cette hypothèse permet de prévoir la fin d'un processus de production. Connaître la fin de la durée opératoire d'un équipement permet de distinguer les nœuds du réseau concentrant le plus de trafic. On peut ainsi contourner ces points de congestions dans la recherche du plus court chemin.

Cependant, dans la pratique, les temps des traitements opératoires ne sont souvent pas fixes. Il y a des variabilités pouvant être dues à la disponibilité de l'équipement (par exemple, l'équipement tombe momentanément en panne ou n'a pas la bonne recette, etc.). Il est donc difficile de prévoir la fin d'un processus. *Les performances de l'algorithme pourraient être biaisées, car elles dépendent entièrement de l'emplacement des nœuds concentrant le plus de trafic.*

Pas de prise en compte du temps de réallocation. Le temps de réallocation des véhicules a été négligé. Le temps de réallocation est le temps nécessaire pour trouver un véhicule libre et non-assigné à une tâche, afin de le ramener à une position où il pourrait répondre à un maximum de requêtes de transport, en un minimum de temps. Dans [21], les auteurs ont traité des problèmes d'allocation et de dispatching des véhicules dans une unité de fabrication de grande taille. L'approche utilisée est à deux étapes : (1) sélection de la meilleure politique de dispatching et (2) détermination du nombre de véhicules.

Cette définition montre bien que le temps d'assignation est fortement dépendant du temps de réallocation des véhicules. Or, un temps d'assignation élevé augmente l'inadéquation entre transport et planning de production. Par exemple, mettre trop de temps pour trouver un véhicule peut fortement perturber l'enchaînement des tâches sur les équipements de production.

Nombre de ports de chargement non considéré. Dans sa thèse, [70] a traité des problèmes d'allocation et de dispatching des véhicules avec les hypothèses simplificatrices suivantes : deux ports de chargement et une durée opératoire fixe pour tous les équipements. Ces travaux ont été améliorés par [42], par la mise en place d'une politique de gestion des véhicules non assignés à une requête de transport. Pourtant, *le nombre de ports de chargement est un paramètre très important, intervenant dans le calcul du temps de changement d'un lot sur un port de chargement* [52]. Le temps pour changer un lot par un autre sur un port de chargement est un indicateur essentiel pour mesurer la qualité du service fournit par un système de transport.

Non prise en compte des caractéristiques des équipements. Deux types de transports ont été comparés dans [88] et [76]. Les convoyeurs aériens (*Continuous Flow Transport ou CFT*) et les transports utilisant des véhicules aériens unitaires. Les critères de comparaison sont le temps de livraison et le niveau d'encours. Certaines caractéristiques d'équipements, comme par exemple les buffers internes, n'ont pas été prises en compte. Pourtant, dans [86], on relève une meilleure performance du CFT quand on ajoute des buffers internes sur les équipements de production. *Ces buffers permettent de réduire le nombre de transports et le temps de non-production.*

Non prise en compte de la taille des lots. Différentes études ont montré que *diminuer le nombre de wafers de 25 à 13 dans un lot, conduit à une réduction du temps de cycle de 10% [15] ou 16% [30] et du niveau d'encours de 20% [2]*. De plus, si le système de transport peut le supporter, [33] prévoit une réduction de 70% du temps de livraison. On a donc un impact sur des indicateurs très importants tels que le temps de cycle et le niveau d'encours. La taille des lots est une contrainte de production à ne pas négliger.

Non prise en compte du temps de non production. Dans [74], les auteurs ont développé une méthode analytique, basée sur les chaînes de Markov, pour estimer le temps de réponse des véhicules ou temps d'allocation dans une boucle fermée. Les temps de transport sont supposés déterministes, car l'objectif est d'avoir un modèle analytique rapide. *Les interactions entre les véhicules ont été ignorées pour faciliter*

les calculs.

Si on distingue en deux parties le temps de réponse d'un véhicule (temps pour trouver le véhicule et temps du trajet pour aller au lot), alors le temps du trajet pour aller au lot va fortement dépendre des conditions de trafic, c'est-à-dire du niveau d'encombrement. Dans le cas d'une boucle fermée, cas d'étude des auteurs, ces encombrements se limitent (sauf cas exceptionnel) à la prise ou au dépôt des lots. Mais, dans un système unifié où le véhicule doit parcourir un « grand » trajet, il y a plus de variabilités (congestion en un point du réseau à fort trafic, contraintes unidirectionnelles de déplacement, etc.). Il peut y avoir erreur d'estimation du modèle, ce qui peut conduire à une surestimation du temps de réponse des véhicules.

Conclusion. Ne pas considérer à minima les contraintes de production peut conduire à des résultats partiels et donc à des décisions erronées. Mais, est-il toujours bénéfique, en termes de temps de calcul et de pertinence des résultats de simulation, de descendre à un important niveau de détails de la production ? C'est dans ce choix que réside toute la difficulté.

Modélisation de la production, avec fortes hypothèses simplificatrices sur le transport

Le tableau 3.2 est un résumé de cette section. Il reprend l'étude bibliographique montrant les impacts de fortes hypothèses simplificatrices sur le transport dans la modélisation d'une unité de fabrication de semi-conducteurs.

La modélisation d'une fab est très complexe. Cette complexité des processus de production pousse les auteurs à faire des hypothèses simplificatrices sur le système de transport lors de la modélisation d'une unité de fabrication. Les auteurs peuvent ainsi mieux se focaliser sur la qualité de l'innovation apportée. Le plus souvent, les simplifications sont considérées par les auteurs comme minimales. Mais, qu'en est-il vraiment ? Surtout que, d'une part, le comportement humain est très difficilement prévisible et, d'autre part, les opérateurs sont de plus en plus remplacés par des véhicules automatiques. Ainsi, les hypothèses sur le transport automatisé sous-estiment les problèmes de gestion de répétitivité et de fiabilité des systèmes automatiques.

Auteurs	Année	Hypothèses sur la production
Panwalkar and Iskander [75]	1977	Non prise en compte de l'ordonnancement et du regroupement des lots
Blackstone <i>et al.</i> [12]	1982	
Kim <i>et al.</i> [50]	2001	
Zhao <i>et al.</i> [95]	2006	
Barketau <i>et al.</i> [9]	2008	
Schmidt and Rose [84]	2008	Non prise en compte de la variabilité du transport
Brain and Lin [14]	2007	Non prise en compte du nombre maximum de véhicules

TABLE 3.2 – Revue de la littérature : modélisation de la production avec fortes hypothèses simplificatrices sur le transport.

Non prise en compte de l'ordonnancement et du batching. Dans [50], les auteurs ont amélioré des modèles existants³ d'ordonnancement, de regroupement par *batch* (nouvelles heuristiques) et de choix des lots (règles de dispatching). En ordonnancement, trois règles pour estimer le « manque d'activité », ou *slack time*, des lots ont été proposées et comparées aux règles existantes [49], [12], [75]. Notons que le choix des *batches* s'effectue en se basant également sur le *slack time* des lots.

Il n'y a aucune hypothèse sur les contraintes de transport. Pourtant, dans [63], les auteurs ont montré par simulation l'existence d'une relation entre la capacité de transport des véhicules et la taille des *batches*. *Ils ont conclu que la capacité des véhicules devrait être inférieure ou égale à un multiple entier de la taille du batch de l'équipement.* Le problème de *batching* a une très forte corrélation avec les problèmes de transport.

De ce fait, la meilleure politique de regroupement des lots n'atteindra pas les performances attendues, si le système automatique de transport ne suit pas. C'est le cas notamment dans [95] et [9]. Il est donc important d'incorporer les contraintes de transport dans la modélisation de la production.

3. Le logiciel utilisé est un logiciel de simulation FACTOR/AIM développé par Pritsker.

Non prise en compte de la variabilité du transport. Dans [84], les auteurs ont mis en exergue les bénéfices d'utiliser des lots de petite taille, des *mini-batches* ou un wafer unique. L'objectif de l'utilisation de lots de petite taille est de diminuer les encours.

Les temps de transport ont été considérés comme identiques pour tous les lots afin de simplifier. Pourtant, les lots de petite taille engendrent plus de transport et donc de trafic. Il y a également augmentation du risque de congestion. Les hypothèses ainsi faites sur le transport sont donc lourdes de conséquences. Il semble nécessaire d'inclure un minimum de contraintes de transport.

Non prise en compte du nombre maximum de véhicules. Dans [14], les auteurs ont étudié l'importance d'utiliser des véhicules « libres ». Les véhicules libres sont des véhicules qui ne sont affectés à aucune zone. Ils ont montré qu'avec une vitesse se situant entre 2m/s et 3m/s, on réduit, de manière non négligeable, le temps de livraison dans une unité de fabrication avec gros volume de production et un grand nombre de *hot lots*⁴.

Les véhicules libres permettent de réduire les temps d'allocation et de réallocation des véhicules lorsqu'ils sont disponibles. Cependant, l'utilisation de véhicules libres accroît le nombre total de véhicules dans le système. D'après [70], on sait qu'il existe une limite supérieure sur le nombre total de véhicules au-delà de laquelle, il y a dégradation du temps de cycle.

Conclusion. Le transport est un service pour la production, donc **d'un point de vue modélisation et simulation, transport et production doivent être considérés avec suffisamment de détails**. Les exemples cités ci-dessus et tirés de la littérature étayent nos conclusions. On retrouve beaucoup de travaux dans la littérature traitant du transport et de la production mais privilégiant le plus souvent l'un ou l'autre. Pourquoi? La réponse se trouve-t-elle dans le temps de calcul trop long des ordinateurs, ou plutôt dans la nécessité de détails qui contribue à complexifier les temps de conception de modèle et de simulation ?

4. les *hot lots* sont des lots de très grande priorité, devant idéalement avoir un temps d'attente nul.

3.4 Corrélation entre transport automatisé et stockage

Le génie consiste à voir ce que tout le monde a vu et à penser ce que personne n'a
pensé.
Jean-Sairien

Introduction

Le stockage fait partie intégrante d'un système de transport automatisé. Dans la littérature, il est traité comme « simple » intermédiaire entre transport et production. Ceci est particulièrement vrai dans les systèmes de transports ayant des configurations *interbay/intrabay* [26] car, le stockage constitue le seul moyen d'échange entre les zones. Notons qu'il existe plusieurs types de stockage [85]. Nous traiterons dans cette section uniquement des stockeurs de grande capacité. Le cas des stockeurs de capacité unitaire n'est pas traité dans la littérature à notre connaissance.

De manière générale, il y a très peu d'articles traitant des performances du système de stockage [36]. Il s'en suit que très peu d'études ont été effectuées sur l'impact du manque de fiabilité du système de stockage sur l'ensemble des performances de la fab.

L'objectif de cette section est d'évaluer l'impact et l'importance des systèmes de stockage par rapport à la production et au transport. Cela permet d'évaluer « l'intensité » de la corrélation pouvant exister.

Étude de la littérature : impact des stockeurs sur l'unité de fabrication.

Le tableau 3.3 est un résumé de cette section. Il reprend les principaux travaux ayant traité des problèmes de stockeurs à notre connaissance.

Gestion distribuée versus gestion centralisée. Dans [64], les auteurs ont étudié l'impact du nombre de déplacements par heure des véhicules, de l'utilisation des

Auteurs	Année	Thèmes étudiés
Cardarelli and Pelagage [17]	1995	Interaction entre transport et production sur le stockeur
Pillai <i>et al.</i> [79]	1999	Relations stockeurs, rails et nombre de véhicules
Mackulak and Savory [64]	2001	Gestion distribuée versus gestion centralisée
Jang <i>et al.</i> [36]	2005	Modélisation analytique des stockeurs
Wiethoff and Swearingen [93]	2006	Intégration des règles de production dans la gestion du stockeur

TABLE 3.3 – Revue de la littérature : sélection d’articles étudiant les stockeurs.

stockeurs et du temps moyen de livraison sur différentes configurations du système de stockage (distribuée ou centralisée). Les résultats de simulation montrent que le temps moyen de livraison dans le système distribué est meilleur que celui du système centralisé. De plus, le taux d’utilisation du robot intérieur au stockeur varie de **24%** à **38,4%** pour les systèmes distribués, et de **17%** à **52%** pour les systèmes centralisés. Ils ont conclu que les systèmes distribués sont préférables aux systèmes centralisés. Le stockage n’est donc pas qu’un simple intermédiaire entre transport et production.

Relations entre stockeur, agencement des rails et nombre de véhicules.

Dans [79], les auteurs ont discuté de l’emplacement optimal des stockeurs en fonction de plusieurs types de configuration des rails : configurations 1, 2 et 3 (voir figures 3.1 et 3.2). D’après les auteurs, puisque le nombre de stockeurs est un élément important dans le coût total de l’AMHS, il est préférable de choisir la configuration utilisant le moins de stockeurs. Ainsi, la configuration 3 est préférée aux configurations 1 et 2. La configuration 3 utilise simplement **60%** des stockeurs de la configuration 1 et **70%** des stockeurs de la configuration 2 pour les mêmes performances. Cependant, pour tirer un profit maximum de la configuration 3, il faut **105%** de véhicules en plus par rapport aux configurations 1 et 2 : c’est le prix à payer.

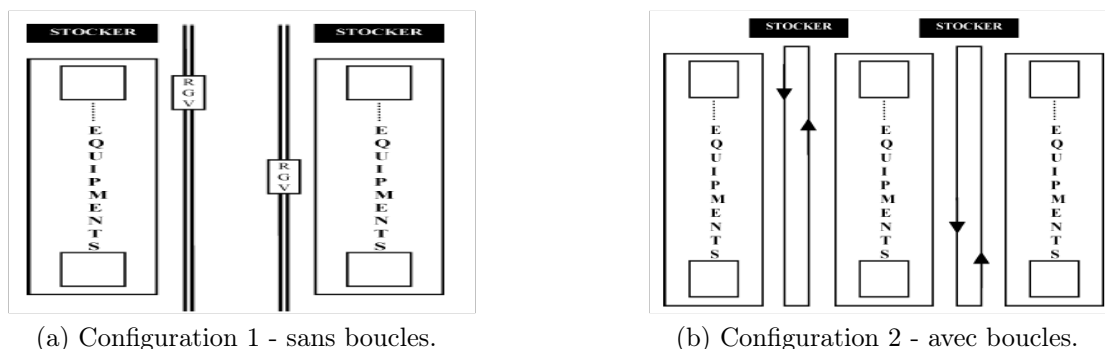


FIGURE 3.1 – Différentes configurations de l'unité de fabrication.

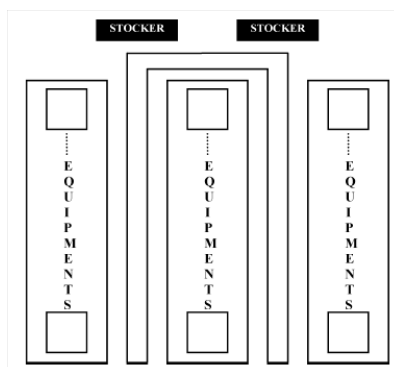


FIGURE 3.2 – Configuration 3 - configuration liée.

Interaction entre transport et production sur le stockeur. Dans [17], les auteurs ont travaillé sur la conception et l'optimisation des systèmes *interbay* et de stockage (méthodes probabilistes). Si le lot est à l'intérieur du stockeur, il dépend du temps de cycle du robot et de la disponibilité du port extérieur de chargement. Le robot du stockeur et le port de chargement extérieur sont, eux-mêmes, dépendants de la disponibilité des opérateurs. Dans les autres cas, le lot dépend du trafic.

Les auteurs ont conclu qu'il était difficile d'aboutir avec ces méthodes à des résultats probants sur la capture des aspects dynamiques (dépendance des opérateurs, disponibilité des ports, densité du trafic, etc.). Ce qui est intéressant à retenir dans cette étude, *c'est que les auteurs démontrent l'existence d'une interaction entre*

transport, production et stockage et, de plus, que cette interaction n'est pas évidente à maîtriser.

Intégrer les règles de production dans la gestion du stockeur. Dans [93], les auteurs ont démontré l'importance de tirer profit de la priorité des lots à l'intérieur des stockeurs. Pour ce faire, ils ont comparé la règle de FIFO, avec une règle de sélection des lots par priorité. Ils ont trouvé que la priorité des lots réduit considérablement le temps d'attente des lots. Par exemple, sur la figure 3.3, on voit que des lots A et B passeront les premiers si l'on prend des critères de priorité (P1) alors qu'ils sont de rang 4 avec la règle FIFO. Donc, prendre en compte les règles de priorité des lots de production permet une meilleure efficacité de la gestion des lots du stockeur. Il est important d'en tenir compte dans toute optimisation du système.

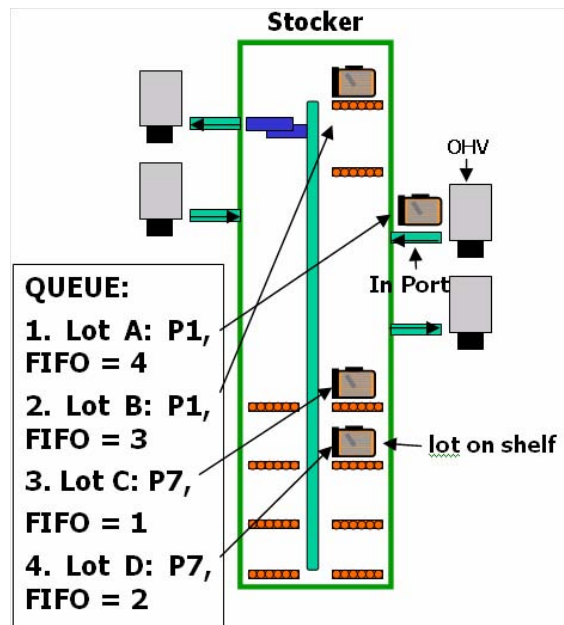


FIGURE 3.3 – Comparaison de deux règles pour le destockage de lots dans un stockeur : FIFO *vs.* priorité des lots.

Modélisation analytique des stockeurs. Dans [36], les auteurs ont développé une méthode de modélisation et d'analyse du système de stockeurs. C'est une mé-

thode analytique basée sur une modification de la « *Decomposition method* » proposée préalablement par [27]. *Ils ont conclu que le taux de panne du robot dans le stockeur est une variable particulièrement critique de la performance de tout le système.* Faisons le rappel suivant afin de mieux comprendre la seconde conclusion des auteurs. Il existe un mini-convoyeur permettant de faire passer le lot du port extérieur au port intérieur du stockeur et *vice-versa*. Ce mini-convoyeur est appelé transporteur. On appelle *Mean Time To Failure*, ou MTTF, le temps moyen d'occurrence d'une panne. Les auteurs ont montré que, quand le MTTF du transporteur est petit, *la taille des ports de chargement joue un rôle crucial dans le taux de livraison global.*

Conclusion. Le stockeur est un élément bien plus important que l'attention qui lui est donnée dans la littérature. C'est un élément incontournable de la fab. Nous avons montré que le stockeur est en relation directe avec le transport, la configuration du système et la production. Nous avons également mis en exergue que ces relations ne sont pas « faibles », c'est-à-dire qu'elles ont un impact conséquent. Mais, comme on l'a décrit précédemment, en termes de modélisation, cela impose d'inclure plus de détails. Cependant, il y a un compromis à faire entre détails et qualité de la modélisation. En effet, *un modèle est une représentation d'un système actuel, devant contenir suffisamment de détails pour répondre aux questions posées, sans prendre en compte plus de détails que nécessaires.*

Nous avons ainsi montré les limitations que l'on rencontre dans la littérature sur les modèles de fab. De manière générale, les impacts des liens entre transport, production et stockage ne sont pas considérés dans les modèles. Nous allons nous positionner par rapport à cela.

3.5 Notre positionnement par rapport à la littérature

« *On ne connaît pas complètement une science tant qu'on n'en sait pas l'histoire* ».
Auguste Comte

Dans [37], les auteurs ont effectué une classification des publications, en fonction du niveau de détails considérés dans la modélisation des unités de fabrication de semi-conducteurs, entre 1999 et 2008. Ils ont regroupé en quatre quadrants les différents niveaux de détails autant sur la capacités que sur le transport 3.4. Le quadrant 1 représente le niveau le plus agrégé, par exemple la simulation agrégée sous Excel, tandis que le quadrant 4 représente le niveau le plus détaillé (simulation détaillée du système AMHS et de la production). Entre les deux, le quadrant 3 représente un grand niveau de détails pour la production mais un faible niveau pour l'AMHS. Le quadrant 2 représente l'inverse. Nous avons complété cette bibliographie en l'étendant de 2008 à 2010.

Par rapport au classement effectué par [37], le travail de cette thèse se positionne dans le quadrant 4, c'est-à-dire que nous allons privilégier un important niveau de détails autant sur l'AMHS que sur la production. **L'innovation majeure de cette thèse est la prise en compte simultanée du transport, de la production et du stockage dans un modèle de simulation le plus *fidèle*.** De plus, nous considérons les stockeurs avec leurs capacités réelles. Dans le cas de cette thèse, l'unité de fabrication considérée est l'unité 300mm de l'entreprise STMicroelectronics située à Crolles, en France.

Il existe dans la littérature très peu d'articles se positionnant dans le quadrant 4. En effet, on peut remarquer que les quadrants 2 et 3 correspondent à environ **80%** de la littérature jusqu'en 2008 (voir tableau 3.4).

Intérêt de simuler une unité de fabrication le plus fidèlement possible. Les différents éléments fournis dans les sections précédentes (sections 3.2, 3.3 et 3.4) permettent de répondre sans ambiguïté. En effet, notre étude bibliographique sur le transport, la production et le stockage a montré qu'étudier séparément ces trois

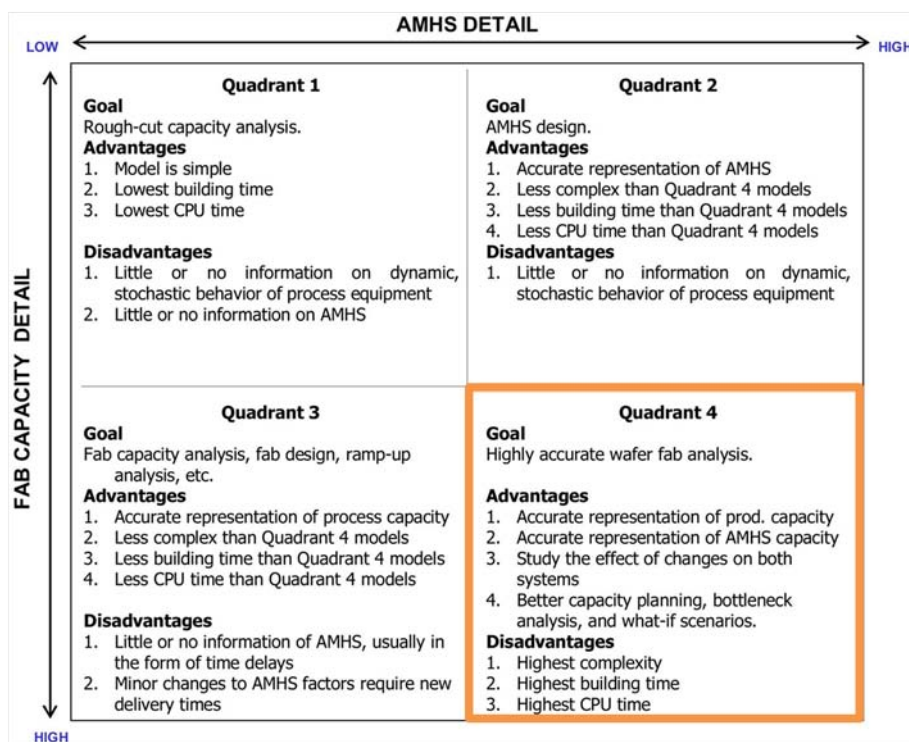


FIGURE 3.4 – Classement des niveaux de détails en modélisation d’un système de fabrication et de production de semi-conducteurs, Jimenez *et al.* [37].

Quadrant 1	Quadrant 2
12%	34%
Quadrant 3	Quadrant 4
46%	8%

TABLE 3.4 – Pourcentage de publications par quadrant.

éléments peut conduire à des résultats et décisions biaisés. Cette étude recommande de travailler en combinant *au moins* deux de ces éléments si l’on veut réduire les erreurs, soit le couple « transport, production », « transport, stockage » et « production, stockage ».

Dans cette thèse nous allons plus loin. Nous proposons de travailler en combinant les trois éléments : **transport** + **production** + **stockage**. Nous espérons ainsi

trouver une valeur de Pareto du niveau de détails qui, de notre point de vue, permet d'avoir une étude théorique de qualité reflétant les réalités industrielles. **Nous ajoutons une contrainte forte : la capacité réelle des stockeurs.** Cette contrainte va nous permettre d'augmenter le niveau de fidélité de notre modèle.

Nous fournissons ainsi une avancée scientifique au niveau des techniques de modélisation et de simulation, mais également, une avancée industrielle par une meilleure appréhension de ces réalités et la proposition de solutions adaptées. Cette dernière avancée permet d'effectuer de meilleures estimations financières des solutions déployées.

Pourquoi une approche par simulation ? Par exemple, couvrir en transport toutes les zones *interbay* et *inrabay* demande des logiciels complexes de simulation à événements discrets [18], [28] et [78].

De manière générale, *la simulation est mieux adaptée aux systèmes tellement complexes, qu'une formule n'est pas suffisante pour prédire des événements dans le système.* C'est le cas des unités de fabrication de semi-conducteurs où, à chaque événement, plusieurs phénomènes peuvent se produire en même temps.

D'après [8], [6] un modèle à événements discrets est une représentation des composants d'un système et de leurs interactions. Tandis qu'un modèle mathématique est une représentation sous forme de formules des composants d'un système, dans un objectif de « résolution ». Les modèles à événements discrets font « tourner » un modèle mais ne le « résolvent » pas. Leurs résultats sont basés sur les interactions entre les composants du système. De plus, les modèles à événements discret sont **dynamiques**, dans le sens que, c'est le changement de temps qui joue un rôle essentiel. A contrario, la **plupart** des modèles mathématiques sont statiques, dans le sens, qu'ils représentent le système en un point précis et à un instant donné. Cela est principalement dû au fait que les modèles de simulation à événements discrets, intègrent beaucoup plus naturellement les contraintes dynamiques par rapport aux modèles mathématiques. *C'est dans ce sens que la simulation est un bon moyen d'expérimenter de nouvelles approches.*

Simulation fidèle : temps de conception et de simulation très long. Outre la complexité des unités de fabrications (phénomène réentrant, nombre important d'étapes, etc.), les difficultés majeures de la simulation **fidèle** d'une unité de fabrication sont *le temps très long de conception et de simulation*.

Par exemple, dans [10], les auteurs utilisent la théorie du *MaxFlow* pour déterminer la disponibilité des véhicules. Le calcul de cette disponibilité est très gourmand en temps de calcul mais, l'intérêt du *MaxFlow* est qu'elle est indépendante du temps de calcul.

Dans [24], pour simuler fidèlement une unité de fabrication, les auteurs ont combiné deux logiciels commerciaux : AutoMod pour le système de transport automatisé et AutoSched AP pour l'ordonnancement de la production. *Ils ont conclu que la fréquence des échanges entre les deux modèles était le principal obstacle pour réduire le temps d'exécution*. [55] propose une simulation à deux étapes : (1) simulation de la production et (2), simulation du transport. Dans la première étape, le niveau d'en-cours de la ligne de production est prédit puis, dans la seconde étape, les capacités du système de transport automatisé sont prédites.

Notre originalité. Bien que nous utilisions une approche par simulation, nous nous démarquons des différents auteurs de la littérature par trois points :

1. **Utilisation d'un logiciel unique** simulant à la fois transport, production et stockage. Nous contournons ainsi le problème rencontré par [24], à savoir l'utilisation de deux logiciels ralentissant les temps d'exécution. Le logiciel que nous allons utiliser est AutoMod, un logiciel de simulation à événements discrets développé par *Applied Materials*.
2. **Prise en compte de la capacité réelle du stockage.** Cela nous permet de nous démarquer de [55] par exemple.
3. **Externalisation des variables critiques du modèle.** La notion de variable critique ici est en lien avec la simulation. Ce sont les variables qui seront le plus fréquemment modifiées (mises à jour) pour garantir la pérennité du modèle. A notre connaissance, en terme de simulation de transport, nous n'avons pas trouvé dans la littérature la notion de variable critique à externaliser.

3.6 Conclusion

Dans ce chapitre nous avons répondu aux questions :

– **Comment sommes-nous positionnés par rapport à la littérature ?**

Nous avons une approche de résolution par simulation *fidèle*, en prenant en compte à la fois *le transport, la production et le stockage*. L'approche par simulation à événements discrets se justifie par la volonté de capturer l'aspect dynamique du problème, trop complexe pour un modèle analytique.

– **Quelle est l'originalité de cette thèse ?** La prise en compte simultanée du transport, de la production et du stockage dans la modélisation **fidèle** de la fab. Elle se justifie par la double volonté d'avoir, d'une part, le plus petit écart possible entre théorie et pratique et, d'autre part, de détecter des interactions cachées entre les différents éléments. C'est l'innovation majeure de cette thèse.

Avant de passer à la modélisation et à la simulation de notre modèle, nous allons d'abord présenter les travaux entrepris durant la première année de thèse qui nous ont servi pour la suite de notre travaux. D'une part, ce travail a permis une meilleure prise en main du logiciel AutoMod et d'autre part, une introduction aux rudiments de fonctionnement théorique d'une unité de fabrication.

Deuxième partie

MODELISATION ET SIMULATION

Chapitre 4

Extension des travaux de Montoya-Torres

Ce chapitre traitera :

- des rappels et des limitations des travaux de Montoya-Torres [70] (section 4.2),
- de l'extension de ces travaux (section 4.3).

4.1 Résumé du chapitre

Dans les grandes unités de fabrication de semi-conducteurs, les distances et donc les temps de transport des lots de production sont plus élevés. Du point de vue lot, cela se traduit par des temps d'attente plus grand. Il est donc nécessaire de réduire les temps de transport de manière générale et en particulier les temps d'attente des lots. Réduire le temps de transport, c'est trouver une stratégie de gestion qui augmente l'adéquation entre transport et production.

Dans [70], les auteurs ont proposé une stratégie (*diviser pour régner*) pour la gestion des véhicules dans les *grandes unités* de fabrication de semi-conducteurs. C'est une approche à deux niveaux, un niveau tactique et un niveau opérationnel. L'objectif du niveau tactique est de définir la distribution des véhicules qui permettra de diminuer au maximum le temps de service d'une demande de transport. Celui du niveau opérationnel est de prendre en compte l'évolution dynamique du système de fabrication, par la définition et la sélection de politiques de gestion opérationnelles des véhicules. Leurs résultats ont montré que la division et la gestion de l'unité de fabrication en zones logiques permettaient une meilleure gestion de l'unité de fabrication.

Cependant, l'unité de fabrication utilisée comporte un seul parking. Lorsqu'ils n'ont plus de requêtes de transport à effectuer, les véhicules se rendent à ce parking pour attendre y une requête de transport. Il y a donc augmentation du transport à vide des véhicules, et donc perte de capacité. Le temps de trajet pour venir prendre un lot est également croissant.

Nous proposons de réduire le temps de transport à vide des véhicules. Pour ce faire, nous augmenterons le nombre de parkings (deux parkings par zone), puis nous ferons continuellement tourner les véhicules dans la zone jusqu'à la prochaine requête de transport. L'objectif est de rendre les véhicules plus réactifs aux requêtes de transport et de réduire le temps de transport à vide.

4.2 Rappels et limitations des travaux

Le progrès en art ne consiste pas à étendre ses limites, mais à les mieux connaître.
Georges Braque

Les travaux présentés dans cette section ont été réalisés durant la première année de thèse, et avant que je ne bénéficie du dispositif CIFRE avec STMicroelectronics. Ils sont le prolongement des travaux réalisés dans [70] dans le cadre d'un autre dispositif CIFRE (2002 à 2005 avec STMicroelectronics). Ces travaux s'inscrivent dans le cadre d'une étude en amont, concernant la gestion des véhicules dans les grandes unités de fabrication de semi-conducteurs.

4.2.1 Introduction

Les auteurs sont partis du constat que, dans les grandes ¹ unités de fabrication de semi-conducteurs, les véhicules parcourent une plus grande distance pour exécuter les requêtes de transport. La figure 4.1 illustre l'unité de fabrication sur laquelle ont travaillé les auteurs.

En effet, avec l'augmentation de la taille des wafers (300mm aujourd'hui et 450mm demain), la taille des équipements de production augmente aussi. Ils occupent plus de surface au sol et cela augmente la taille de l'unité de fabrication. La conséquence est l'**augmentation** :

- *du temps de livraison* des lots car les trajets sont plus longs à parcourir,
- *de la variabilité des temps de transport* car le véhicule disponible peut aussi bien être à proximité du lot que très éloigné,
- *du temps de transport à vide des véhicules*, donc une perte de capacité de transport en nombre de véhicules. En effet, un véhicule avec un temps de transport à vide élevé constitue une ressource en moins dans l'unité de fabrication.

Le constat des auteurs de [70] est très important car l'augmentation du temps de transport influe directement sur le temps de cycle des lots et donc sur la productivité de l'unité de fabrication. Afin de réduire le temps de transport, surtout le

1. Longueur=100m, largeur=50m

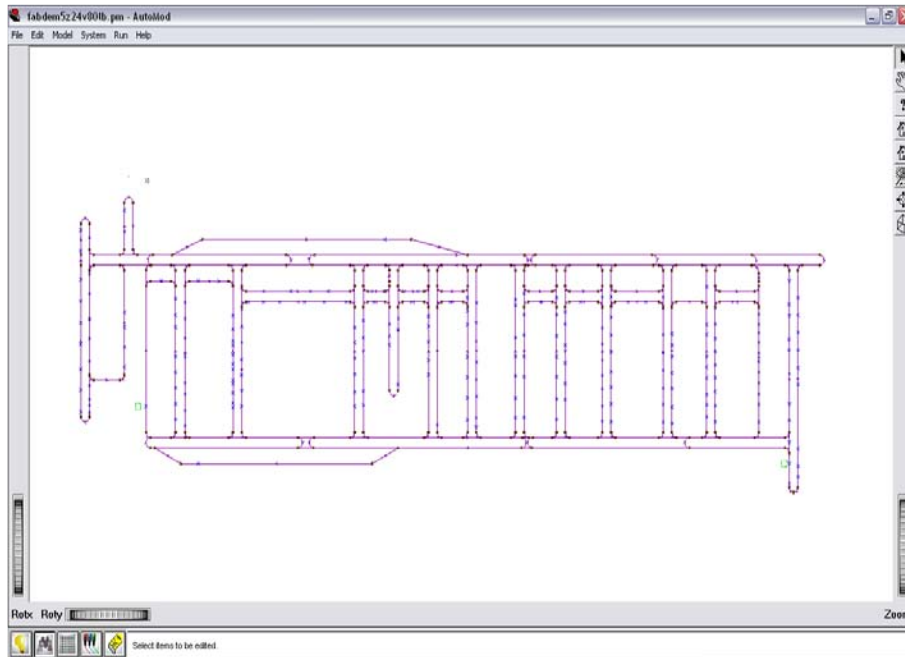


FIGURE 4.1 – Agencement des rails de l’unité de fabrication utilisé dans [70].

temps de transport à vide des véhicules, les auteurs ont utilisé une approche consistant à minimiser le temps maximal d’attente des lots pour tout point de l’unité de fabrication.

Ainsi, à un temps maximal d’attente correspond un « rayon d’action » des véhicules, d’où la notion de zone. Les auteurs ont considéré que l’unité de fabrication entière correspondait à une zone (1-zone). Ils ont ensuite comparé, pour les grandes unités de fabrication, les performances du système de transport lorsque l’unité de fabrication est divisée en x -zones ($x > 1$) par rapport à 1-zone. L’approche utilisée est une approche à deux niveaux :

- **Tactique** - Il s’agit d’étudier le problème d’un point de vue global, afin de décider de la « meilleure » façon de gérer les véhicules dans le système. L’objectif est de définir la distribution des véhicules qui permettra de diminuer au maximum le temps de service d’une demande de transport. La solution du niveau tactique permet d’obtenir un découpage logique du réseau de transport en zones de gestion et d’affecter un nombre de véhicules à chacune des zones

pour satisfaire les requêtes de transport.

- **Opérationnel** - Il s'agit de coupler la solution optimale du niveau tactique avec un modèle permettant de prendre en compte l'évolution dynamique du système de fabrication. Les auteurs ont défini des politiques pour la gestion opérationnelle des véhicules. Ils ont également défini et comparé plusieurs stratégies, chacune caractérisée par des critères dynamiques pour la ré-affectation des véhicules.

Les résultats obtenus ont montré que la gestion par zones est bien plus intéressante lorsqu'une *stratégie appropriée* de dispatching de véhicules est implémentée. Ainsi, les auteurs ont pu voir que la règle de dispatching FCFS (*First Come First Serve*) ou FIFO (*First In First Out*) qui est la plus utilisée dans la littérature, donne de très mauvaises performances si le réseau n'est pas découpé en zones. Avec le découpage en zones et une politique intelligente de gestion des véhicules, la règle FCFS peut être très intéressante.

L'approche utilisée par les auteurs permet donc d'améliorer l'adéquation entre ressources et besoins de transport par une réduction de la zone à couvrir. Soulignons que les problèmes d'allocations sont bien connus en recherche opérationnelle (problème de p-centre [34], etc.) ou en statistiques (problème du k-means [65] [38], etc.).

Dans la suite de ce chapitre, nous rappellerons brièvement les approches utilisées par les auteurs et leurs principaux résultats (section 4.2.2). Nous détaillerons ensuite les limitations de ces travaux (section 4.2.3). Puis nous présenterons les extensions de ces travaux dans la section 4.3. La section 4.4 conclura ce chapitre et on y proposera des perspectives de recherche.

4.2.2 Rappels

L'objectif global d'une unité de fabrication est de faire un produit le plus rapidement possible, à moindre coût et en grande quantité. Pour atteindre cet objectif, les managers disposent de plusieurs outils : les systèmes de production, de communication, etc, parmi lesquels le système de transport automatique.

Le transport qui est un service pour la production, a pour objectif d'amener le

bon lot au bon endroit et au bon moment. La problématique du transport est donc de mettre en place des politiques qui assurent une adéquation temporelle et spatiale entre les ressources et les besoins en transport des lots. Dans le cas de l'unité de fabrication (théorique) étudiée par Montoya-Torres [70], la recherche de cette adéquation se traduit par : la division de l'unité de fabrication en zones logiques, l'affectation des véhicules à ces zones, puis la détermination de la « meilleure » politique de transport des véhicules. Les auteurs ont utilisé une approche à deux niveaux, tactique et opérationnel, dont la cohérence a été démontrée dans [72]. En effet, ils ont montré que des détériorations au niveau tactique impactaient négativement le niveau opérationnel. Pour mieux comprendre l'approche des auteurs dans [70], nous détaillons chaque niveau de leur étude.

4.2.2.1 Niveau tactique

Le tableau 4.1 résume le niveau tactique développé dans [72].

Situation	Grande unité de fabrication
Travail à faire	Affectation des ressources de transport
Action effectuée	Découpage de l'unité de fabrication en zones logiques
Résultats	5 zones et un besoin de 24 véhicules
Méthode	Programmation mathématique
Modèle	Variante du problème du p-centre [34]

TABLE 4.1 – Résumé du niveau tactique : approche de Montoya-Torres [70]

L'idée développée au niveau tactique est de « diviser pour régner ». L'unité de fabrication est divisée en *zones logiques*. La taille des zones est déterminée de manière à équilibrer la charge de travail entre les zones, et à minimiser la distance maximale parcourue par les véhicules à l'intérieur d'une zone.

Plusieurs nombres de zones ont été testés afin d'étudier le comportement du système. Un programme linéaire en nombres entiers (variante du p-centre) est alors implémenté. Il utilise comme données d'entrée le même nombre de zones que ceux utilisés dans [70], la charge de travail de chaque point de l'unité de fabrication, exprimée en temps de transport par heure. Rappelons que l'objectif du programme

linéaire en nombres entiers est de diviser l'unité de fabrication en zones logiques. Les résultats du programme linéaire en nombre entiers sont les centres de zones et les points appartenant à chaque zone. L'appartenance d'un point à une zone est déterminée en fonction de la charge de transport du point considéré et de sa distance au centre considéré.

Soulignons un cas très intéressant pour la suite de notre étude. Supposons qu'un point soit proche d'un centre $C1$ d'une zone, mais que ce point soit très chargé, c'est-à-dire que le nombre de transports vers ce point est élevé. Dans le souci d'équilibrer la charge en nombre de transports des zones, le modèle peut alors affecter ce point à un autre centre $C2$ plus éloigné afin de diminuer la charge de centre $C1$. Cette remarque permet alors de se rendre compte que les zones peuvent avoir n'importe quelle forme géométrique comme le montre la figure 4.2.

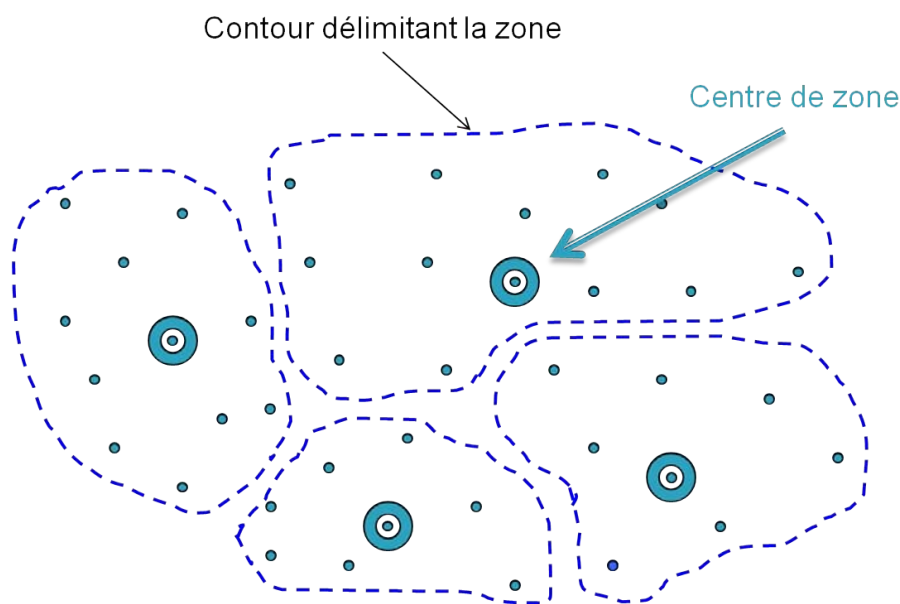


FIGURE 4.2 – Découpage en zones logiques de l'unité de fabrication de [70].

Les résultats exploités par les auteurs utilisent un découpage logique en 5 zones et 24 véhicules suffisent pour obtenir des performances acceptables. La figure 4.3, illustre un exemple de découpage logique de l'unité de fabrication en 5 zones et une répartition des véhicules.

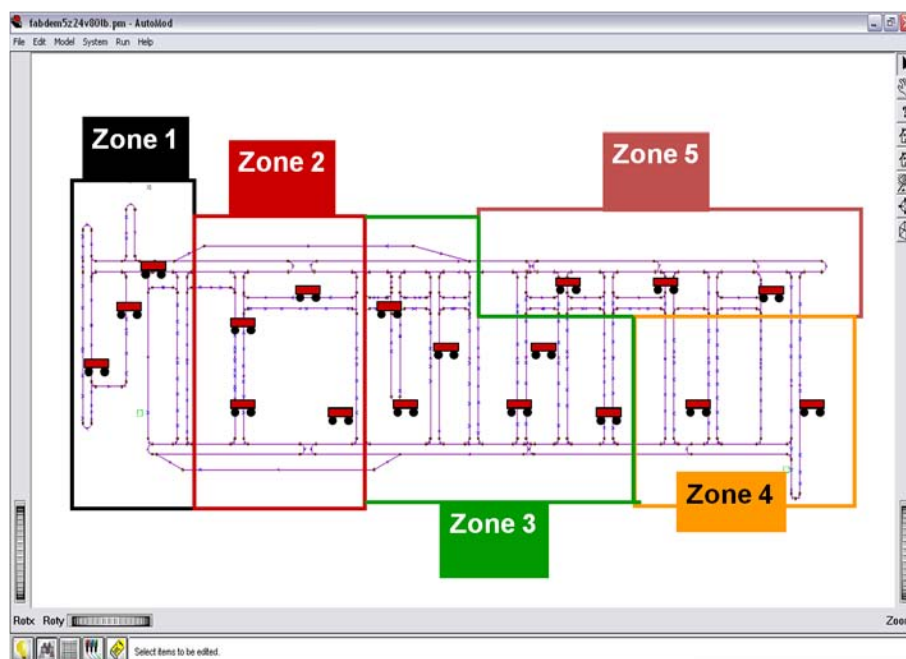


FIGURE 4.3 – Découpage de l’unité de fabrication utilisée dans [70].

Avec ces résultats, les auteurs vont maintenant déterminer la « meilleure » politique de gestion des véhicules au niveau opérationnel.

4.2.2.2 Niveau opérationnel

Le tableau 4.2 résume le niveau opérationnel développé dans [72]. Les idées développées au niveau opérationnel sont, d’une part, de trouver la politique la mieux adaptée pour répondre aux requêtes de transport des lots, et d’autre part, en se basant sur cette politiques de trouver les conditions permettant d’améliorer l’adéquation entre ressources et besoins de transport.

Afin de trouver la politique la mieux adaptée pour répondre aux requêtes de transport, les auteurs ont testé plusieurs politiques de transport et ont finalement retenu la politique *ZVBDemFab* : « (1) vérifier s’il n’y a pas de requêtes de transport au point où se trouve le véhicule, sinon (2) utiliser la règle premier arrivé, premier servi (CL-FCFS, *Current Location - First Come First Served*). Pour plus de détails,

Situation	Nombres de zones et de véhicules prédéfinis
Travail à faire	Trouver des politiques de transport adaptées
Action effectuée	Tests de plusieurs politiques de transport
Résultats	Politique : CL-FCFS
Méthode	Simulation à évènements discrets

TABLE 4.2 – Résumé du niveau opérationnel : approche de Montoya-Torres [70].

voir [70].

Pour augmenter l'adéquation entre ressources et besoins de transport, les auteurs ont mis en place une règle de demande d'accès des zones par les véhicules. L'objectif est d'éviter de surcharger une zone en véhicules afin d'éviter les embouteillages. Pour ce faire, un nombre maximal de véhicules est prédéfini dans chaque zone. Avant d'accepter une requête de transport, le véhicule interroge la zone de destination pour savoir si le nombre maximal de véhicules de la zone est atteint. Si c'est le cas, le véhicule refuse la requête de transport et reste dans sa zone. Si aucune requête de transport ne parvient au véhicule, ce dernier se dirige vers le *seul* parking de l'unité de fabrication.

4.2.3 Limitations

4.2.3.1 Un seul parking

Dans les unités de fabrication de semi-conducteurs, lorsqu'un véhicule termine toutes les tâches de transport qui lui ont été attribuées, il se dirige vers un parking, afin d'attendre d'autres requêtes de transport.

L'unité de fabrication considérée par [70] ne comprend qu'un seul parking (figure 4.4), qui n'appartient à aucune zone du découpage. L'idée est de ne pas regrouper tous les véhicules dans une seule zone afin de ne pas engendrer d'encombres.

Lorsque les véhicules n'ont plus de requêtes de transport à satisfaire, ils sortent de la zone afin de rejoindre le parking. Ce dernier n'est pas à égale distance de tous

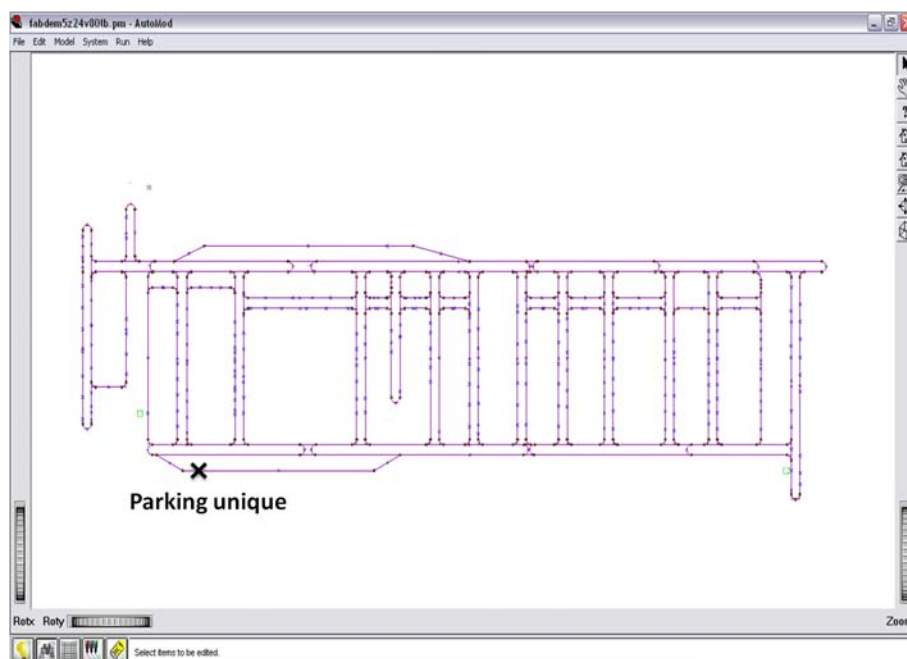


FIGURE 4.4 – Illustration du parking unique utilisé dans [70].

les points de l'unité de fabrication, en particulier du point où le véhicule a effectué sa dernière livraison. Une fois qu'un véhicule démarre en direction d'un point de l'unité de fabrication, le véhicule est *obligé* de terminer ce transport. Il n'y a pas de possibilité de redirection des véhicules dans le modèle considéré. Ainsi, si une requête de transport arrive juste après le démarrage d'un véhicule à destination du parking, le véhicule est obligé d'aller d'abord au parking. *Il y a donc augmentation du temps de transport à vide des véhicules, ce qui conduit à une perte de capacité de transport.*

4.2.3.2 Un centre de zone « mal » placé

Le programme linéaire en nombres entiers utilisé par les auteurs garantit que chaque point appartenant à une zone se trouve à un temps maximal T_{max} du centre de la zone. L'objectif du programme linéaire est de minimiser T_{max} . Cependant, avec uniquement une contrainte de temps, les zones formées peuvent être déséquilibrées

en termes de charge de travail. Le programme va y remédier en excluant de la zone un ou plusieurs points de « faible » charge, pour garder un point de forte charge. Une conséquence à cela est la modification de la forme géométrique des zones.

Puisque l'objectif est de minimiser T_{max} , le maximum est au moins atteint par une zone. Cette zone impose aux autres zones un temps de transport d'au plus T_{max} . Il est donc possible d'effectuer un temps de transport plus court dans d'autres zones, sauf si toutes les zones sont identiques. Ce constat nous amène à remettre en cause les centres des zones concernées, c'est-à-dire des zones où l'on peut faire de la livraison en un temps inférieur à T_{max} . Le centre de certaines zones déterminé par le programme mathématique n'est pas le centre « optimal ». La figure 4.5 illustre le cas de centres de zone mal placés par rapport à la distance maximale T_{max} .

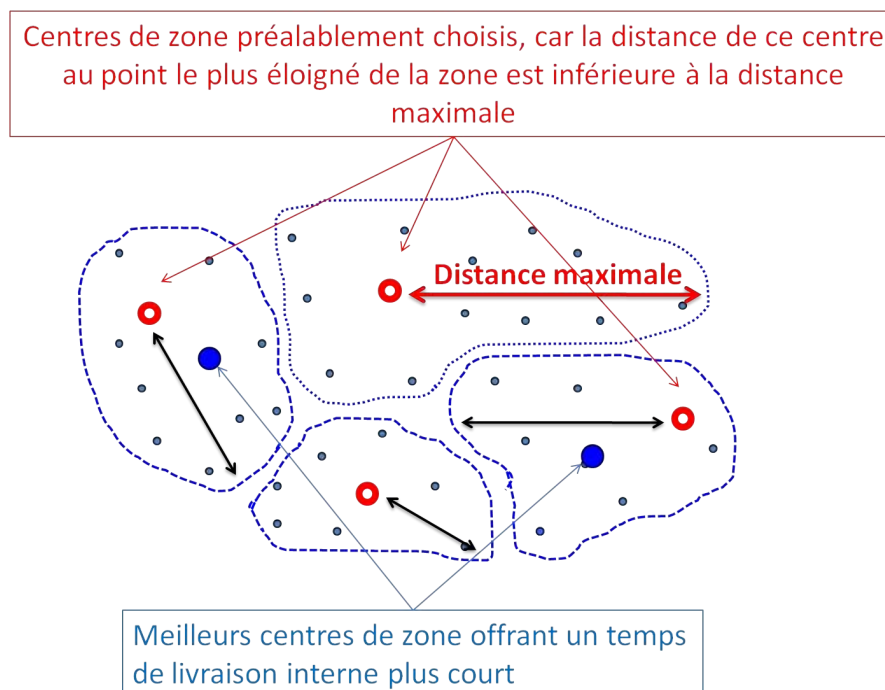


FIGURE 4.5 – Illustration de centres de zone « mal » placés.

4.3 Extensions des travaux

*« Dans une vie, on peut au mieux inventer une chose, mais on peut en perfectionner mille. »
Proverbe japonais.*

4.3.1 Problématiques

Nous allons nous intéresser à la limitation due au parking unique. La configuration de l'unité de fabrication utilisée par les auteurs comporte un seul parking (figure 4.4) n'appartenant à aucune des zones logiques du découpage. De plus, ce parking n'est pas à égale distance de toutes les zones.

Chaque fois qu'un véhicule devient libre et n'est assigné à aucune requête de transport, il se dirige vers le parking dans l'attente d'une nouvelle requête de transport. Les véhicules circulent donc à vide pour rejoindre le parking. Si une nouvelle requête de transport est émise, ils repartent à vide du parking vers le point qui en a fait la demande. Si le parking est proche de la zone, ce temps de déplacement à vide des véhicules peut être « négligeable ». Sinon, le temps de déplacement à vide est élevé et cela génère une perte de capacité en nombre de véhicules.

Notons que, dans le modèle considéré, il n'est pas possible d'affecter une requête de transport à un véhicule déjà en circulation. Il faut que le véhicule soit en arrêt ou en stationnement sur un point de l'unité de fabrication. Donc un véhicule qui a terminé un transport et ne trouve pas de nouvelle requête ira au parking, et la destination de ce dernier ne peut être changée durant le trajet. *Avec un parking unique, dès qu'il y a une « rupture » dans l'enchaînement des tâches, le système devient moins performant.* C'est donc lorsqu'il y a des enchaînements de requêtes de transport que les grands systèmes à parking unique sont les plus performants.

S'il y avait au moins un parking dans chaque zone, le temps de trajet à vide des véhicules serait réduit. Cependant, les véhicules stationnant dans la zone gêneraient le passage des autres véhicules, en particulier ceux traversant la zone puisqu'il n'y a pas de voie de contournement.

Remarque. Lorsqu'un lot émet une requête de transport, le temps écoulé entre l'émission de la requête et la prise du lot par le véhicule peut se décomposer en deux parties : le temps pour trouver un véhicule (temps d'assignation) et le temps pour amener le véhicule au lot. Lorsque le nombre de requêtes de transport devient inférieur au nombre de véhicules, il y a création d'une file d'attente au niveau du parking. Dans ce cas, *le temps pour trouver une ressource est toujours nul (temps d'assignation) s'il y a au moins un véhicule dans le parking*. Le temps pour amener le véhicule au lot est peu variable dans un système où le trafic est fluide.

Cette remarque est très importante, car elle pourrait expliquer en partie pourquoi les auteurs dans la littérature se sont peu intéressés au calcul ou à l'estimation du temps d'assignation des véhicules. Ajoutons également que, dans le cas des « petites » unités de fabrication, ces temps ont « peu » d'importance et donc souvent négligés.

4.3.2 Challenges

Nous avons identifié dans la section précédente une problématique liée à l'approche de Montoya-Torres [70]. On a montré qu'avoir un seul parking dans une grande unité de fabrication augmente le temps pour amener un véhicule au point qui émet une requête de transport. *Le parking unique dans les grandes unités de fabrication augmente donc la perte en capacité de transport des véhicules*.

Pour réduire le temps de transport à vide des véhicules libres et non assignés, il est nécessaire de « rapprocher » le parking de l'endroit où il y a des requêtes de transport, c'est-à-dire des zones. Mais puisqu'il y a plusieurs zones, il faut donc *au moins* autant de parkings que de zones.

Nous avons aussi montré que les centres des zones trouvés par le programme linéaire en nombres entiers de [70], assurent un temps maximal de déplacement mais pas un temps moyen de déplacement dans la zone.

Le challenge est donc double :

1. *déterminer la « meilleure » position des parkings* afin de minimiser le temps moyen de déplacement,

2. *réduire la gêne mutuelle des véhicules* à l'intérieur d'une zone afin de fluidifier le transport.

4.3.3 Extensions des travaux

Cette section a été publiée dans un chapitre de livre [42].

Nous allons à présent présenter les extensions des travaux précédents. La problématique abordée concerne la gestion des véhicules libres et non assignés au repos. Dans la littérature, on rencontre très peu d'auteurs qui traitent de la gestion des véhicules aux repos.

4.3.3.1 Détermination de la « meilleure » position des parkings

Pour déterminer les centres de zones, les auteurs ont utilisé un programme linéaire en nombres entiers qui est une variante du problème du p-centre. Nous avons amélioré le programme linéaire en ajoutant dans la fonction objectif une nouvelle variable. Cette dernière permet de corriger le choix du centre de la zone pour avoir un temps moyen de couverture minimal. Les détails de la modélisation sont présentés dans [42]. La variable introduite représente le temps de couverture de la zone à partir de chacun de ces points. Elle permet ainsi de choisir comme centre le point ayant le plus petit temps de couverture de zone.

4.3.3.2 Réduction de la gêne mutuelle des véhicules d'une zone

Il existe des travaux dans la littérature traitant de différents problèmes de blocage des véhicules. Pour le lecteur désireux d'approfondir ce sujet, je recommande la lecture de [91] et [59]. Dans [47], les auteurs ont montré, par simulation, que changer l'ordre d'attribution des requêtes de déchargement pour des véhicules déjà en route, permet d'améliorer la productivité sous différentes politiques de dispatching ([43] et [44]). Insistons sur le fait que, dans le cas qui nous concerne, il n'est pas possible de ré-affecter les tâches.

Nous ne traiterons que du *cas des véhicules qui sont au repos* à un emplacement de parking. Lorsqu'un véhicule est en déplacement, la « gêne » occasionnée par ce

dernier sur les autres fait partie intégrante du déplacement (distance de sécurité, ralentissement, arrêt, etc.). Donc dans une zone, un véhicule en déplacement est beaucoup moins « gênant » pour les autres véhicules qu'un véhicule en stationnement. Nous allons utiliser cette idée pour réduire le niveau de gêne des véhicules. Ainsi, plutôt que de faire stationner un véhicule dans une zone, nous allons le faire « tourner continuellement » à l'intérieur de la zone jusqu'à la prochaine requête de transport.

La mise en œuvre technique nécessite de déterminer un *deuxième* parking dans la zone, afin que les véhicules puissent faire le plus petit va-et-vient possible entre les deux parkings. Le choix de ce deuxième parking doit cependant respecter les deux propositions suivantes :

1. réduire le temps de transport à vide des véhicules puisqu'ils vont continuellement tourner dans la zone,
2. réduire le temps mis par le système pour choisir un véhicule afin de lui assigner une requête de transport, car il est impossible d'affecter une requête de transport à un véhicule déjà en déplacement. Donc, si les deux parkings sont très éloignés l'un de l'autre, il faut attendre qu'un des deux véhicules qui est en train de tourner atteigne un des deux parkings.

Le second parking est donc choisi en fonction de sa distance par rapport au premier parking.

4.3.4 Résultats

L'approche utilisée est la simulation à événements discrets en utilisant le logiciel AutoMod de Applied Materials. Nous avons repris le modèle de simulation de [70], ainsi toutes les conditions de simulation sont les mêmes. Le temps total de notre étude est de 180 jours répartis en : 120 jours pour stabiliser le système pendant lesquels on ne recueille pas de statistiques, et 60 autres jours pendant lesquels on recueille les données pour les étudier. Les résultats présentés sont normalisés.

Nous donnons les significations des différentes politiques de dispatching :

- *CL-FEFS*, *Current Location - First Encounter First Serve* : regarder d'abord sur le point où l'on se trouve s'il n'y a pas de requête de transport, ensuite

- premier rencontré, premier servi.
- *FCFS*, *First Come First Serve* : premier arrivé premier servi.
 - *TAZ*, *Turn Around in Zone* : tourner continuellement dans la zone.
 - *RP*, *Return to Parking* : retour au parking.

Supériorité du *Turn Around in Zone*. Le tableau 4.3 présente les résultats obtenus en comparant la politique TAZ avec la politique originale utilisée par les auteurs RP. La politique de dispatching utilisée est FCFS. Les résultats montrent une domination de la politique TAZ dans trois cas, autant pour le temps moyen de cycle que pour l'écart type. Les gains obtenus pour le temps moyen de cycle, 3,5% pour 20 véhicules, 2,5% pour 24 véhicules et 0,4% pour 30 véhicules. Cette réduction s'explique par la proximité des véhicules dans la zone.

Nb. de véhicules	20		24		30	
	RP	TAZ	RP	TAZ	RP	TAZ
Temps de cycle	1,085 X	1,049 X	X	0,975 X	0,966 X	0,962 X
Ecart type	0,088 X	0,074 X	0,055 X	0,047 X	0,041 X	0,040 X

TABLE 4.3 – Temps moyen de cycle des lots.

Les résultats des auteurs ont montré que le système est sensible à la variation du nombre de véhicules lorsque celui-ci varie entre 18 et 24. En dessous, le système est très mauvais, au delà le système est peu sensible et il y a risque de congestion.

Conclusion : *la politique TAZ fournit de meilleurs résultats sur le temps moyen de cycle. Il est donc préférable de tourner dans la zone plutôt que d'avoir un parking unique.*

Comparaison des politiques de dispatching. Pour la comparaison des politiques de dispatching nous avons fait varier le nombre de véhicules entre 18 et 24, parce que c'est dans cet intervalle que les variations sont les plus significatives d'après [70]. L'idée est de comparer l'efficacité de l'approche TAZ lorsqu'on utilise deux politiques de dispatching FCFS et CL-FEFS.

La figure 4.6 illustre les résultats obtenus en comparant le temps moyen de cycle des lots. Ils montrent la *supériorité de la politique FCFS sur CL-FEFS lorsque l'on applique la politique TAZ*. Il y a une différence de 2% entre les deux politiques de dispatching, avec un avantage pour FCFS. On peut donc dire sans se tromper que les lots passent moins de temps dans l'unité de fabrication avec la politique FCFS. Toutes les valeurs numériques sont présentées dans le tableau 4.4. Pour les deux politiques, le temps de cycle décroît lorsque le nombre de véhicules augmente. L'écart maximal du temps moyen de cycle entre 18 et 24 véhicules est de 16%. Soulignons que 2% de réduction sur un temps de cycle global n'est pas négligeable car le temps de fabrication d'une plaquette se compte en mois.

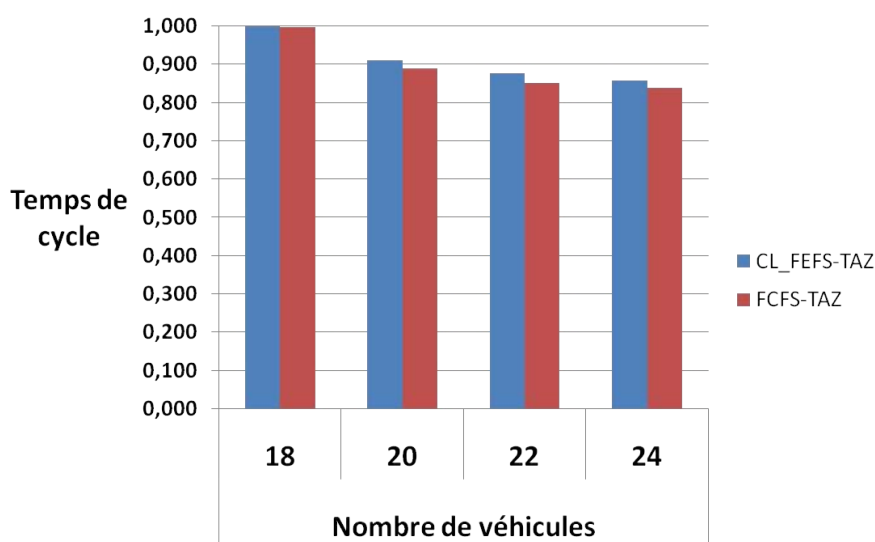


FIGURE 4.6 – Comparaison du temps moyen de cycle entre les politiques FCFS et CL-FEFS.

Si les lots passent moins de temps dans l'unité de fabrication avec la politique FCFS et que le temps de traitement des lots par les équipements n'a pas changé, alors c'est donc le temps de transport des lots qui a été réduit. Mais les véhicules ont conservé les mêmes caractéristiques que dans [70], c'est-à-dire qu'ils possèdent les mêmes performances. *C'est donc le temps d'attente des lots qui a été amélioré.* Pour étayer cette affirmation, nous avons étudié un indicateur local qui est le temps

	Nombre de véhicules			
	18	20	22	24
CL-FEFS et TAZ	0,998	0,911	0,876	0,857
Ecart type	11%	8%	6%	5%
FCFS et TAZ	0,998	0,889	0,852	0,839
Ecart type	19%	7%	5%	5%

TABLE 4.4 – Temps moyen de cycle des lots.

moyen de chargement d'un lot sur un équipement. Les résultats sont illustrés par la figure 4.7. Hormis le cas à 18 véhicules, les véhicules mettent moins de temps pour charger les lots avec la politique FCFS qu'avec CL-FEFS. Cela est numériquement confirmé dans le tableau 4.5 qui présente le temps moyen de chargement des lots sur les équipements de production. Ils montrent une réduction de 4% sur le temps moyen de chargement des lots lorsque le nombre de véhicules varie de 18 à 24.

Avec la politiques CL-FEFS, si un lot très éloigné des zones de fort trafic, donc de véhicules, il mettra plus de temps pour « trouver » un véhicule. Il y a augmentation du temps d'attente de ce lot. Une telle politique ne garantit pas qu'un véhicule libre et non assigné sera attribué à un lot en attente. Cette politique favorise visiblement les zones à fort trafic ou les zones fermées, comme les boucles. Tandis qu'avec la politique FCFS, un tel lot « trouvera » plus rapidement un véhicule car il « appelle » le premier véhicule qui devient libre et qui est non assigné à une requête de transport. On peut résumer cette remarque comme ceci : *un véhicule qui est « réveillé » par un lot n'est pas certain d'être pris par ce lot avec la politique CL-FEFS. C'est le cas avec la politique FCFS.*

Enfin, on observe un écart maximal de 17% pour la politique CL-FEFS et 23% pour FCFS si on fait varier le nombre de véhicules entre 18 et 24. Donc, plus il y a de véhicules, plus le niveau d'adéquation entre capacité de transport et besoins en transport de la production augmente.

Nous présentons enfin un dernier indicateur qui est le pourcentage de lots qui a été chargé sur un équipement en plus de 5 minutes. Les résultats sont illustrés par la figure 4.8 et les résultats numériques sont dans le tableau 4.6. L'information majeure apportée par ces résultats est, qu'avec la politique CL-FEFS, les lots chargés en plus

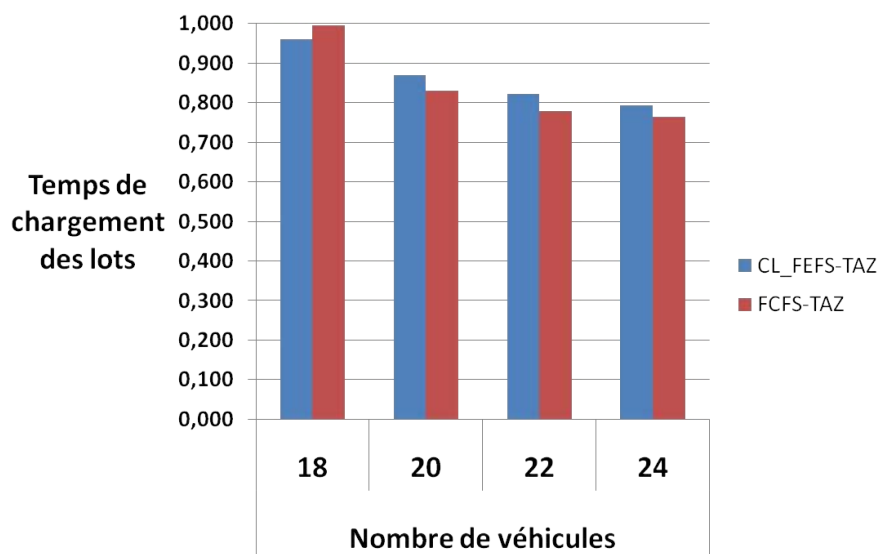


FIGURE 4.7 – Comparaison du temps moyen de chargement d’un lot entre la politique FCFS et CL-FEFS.

	Nombre de véhicules			
	18	20	22	24
CL-FEFS et TAZ	0,961	0,869	0,822	0,794
FCFS et TAZ	0,995	0,830	0,779	0,764

TABLE 4.5 – Temps moyen de chargement.

de 5 minutes sont plus nombreux qu’avec la politique FCFS. On a 6% de lots en moins chargés en plus de 5 minutes avec FCFS et 18 véhicules, 20% en moins avec 20 véhicules, 36% avec 22 véhicules et 47% avec 24 véhicules. La proportion de lots qui attendent longtemps est ainsi réduite de près de moitié quand on passe de 18 à 22 véhicules.

Conclusion : L’approche TAZ est mieux adaptée à la politique FCFS qu’à la politique CL-FEFS.

On retrouve dans [46] des études comparables à ce que nous avons développé dans [42]. Nous nous démarquons de [46] dans la définition du statut « *idle* » des

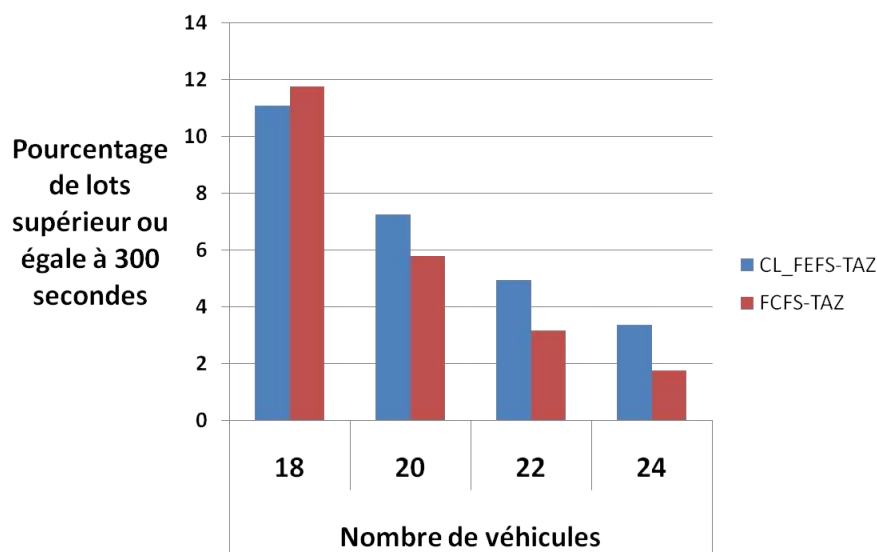


FIGURE 4.8 – Comparaison entre FCFS et CL-FEFS : proportion de lots chargés sur un équipement en plus de 5 minutes.

	Nombre de véhicules			
	18	20	22	24
CL-FEFS et TAZ	11,11	7,26	4,95	3,35
FCFS et TAZ	11,76	5,78	3,15	1,76

TABLE 4.6 – Pourcentage total de lots chargés en plus de 300 secondes.

véhicules. Pour [46], un véhicule est dans le statut « *idle* » lorsqu'il est en stationnement au parking, est en déplacement sans requête de transport en attente, ou a terminé sa dernière livraison. Alors que, dans [42], un véhicule qui a le statut « *idle* » est un véhicule au repos sur un point de parking. Les études de [46] complètent les nôtres dans le sens où les auteurs ont testé la politique de « tourner continuellement dans la zone » en faisant varier la taille des zones. Leurs résultats suggèrent que cette politique est efficace si la zone où elle est appliquée est de taille « moyenne ».

4.4 Conclusion

Sur la base des travaux effectués dans [70], nous avons identifié quelques limitations.

Le parking unique dans les grandes unités de fabrication augmente la perte en capacité des véhicules dans l'unité de fabrication. Afin de bénéficier des avantages de la division en zones logiques de l'unité de fabrication, nous avons défini deux parkings à l'intérieur de chaque zone où les véhicules libres et non assignés tournent continuellement. Les résultats ont conforté la supériorité de l'approche TAZ qui consiste à tourner continuellement dans la zone. Nous avons également comparé deux politiques de dispatching. Les résultats suggèrent que la politique FCFS est la mieux adaptée.

Les résultats obtenus sont encourageants mais méritent un approfondissement. Il faut par exemple déterminer les conditions optimales de l'utilisation de l'approche TAZ. On pourrait par exemple utiliser les travaux de [46] pour poursuivre l'étude.

Chapitre 5

Modélisation et Simulation de Crolles 300

Ce chapitre traitera :

- de la modélisation de l'unité de fabrication de Crolles 300 de STMicroelectronics (section 5.2)
- de la simulation de l'unité de fabrication de Crolles 300 de STMicroelectronics (section 5.3)
- de la validation de la simulation (section 5.4)
- du plan d'expérience (section 5.5)

Nous répondrons aux questions suivantes : **à quoi sert un modèle ? Comment effectuer des changements rapides de scénarios dans les systèmes complexes ? Quels sont les paramètres clés de notre système ?**

5.1 Résumé du chapitre

Un modèle est l'ensemble des relations décrivant un processus, et une simulation est une succession de modèles s'échangeant des informations de manière dynamique. Comprendre une simulation c'est comprendre comment fonctionnent les différents modèles qui la composent. Nous présentons dans cette section les modèles de transport, de production et de stockage qui constituent l'unité de fabrication de Crolles 300 de STMicroelectronics.

Puisque les modèles s'échangent des informations entre eux, nous mettrons en exergue à travers la construction des modèles, les différentes interactions existant entre les modèles de transport, de production et de stockage. Ces interactions mettent également en évidence l'importance de la précision des données dans la conception des modèles et sera démontrée aux travers d'exemples concrets.

Un modèle est toujours entaché d'erreurs. Il est donc important de pouvoir « mesurer » ces erreurs afin de pouvoir évaluer le niveau de confiance du modèle. Pour ce faire, on utilise des indicateurs, d'une part pour mesurer les performances du modèle et, d'autre part pour évaluer les erreurs avec le système réel. Le choix des indicateurs doit permettre de mesurer le plus rapidement et le plus précisément possible les performances du modèle.

La validation est le procédé permettant de déterminer si un modèle de simulation est une représentation précise du système, *pour l'objectif particulier de l'étude considérée* [58]. Nous détaillerons donc le mode et l'approche de validation de notre modèle de simulation et nous fixerons un niveau de confiance du modèle à 15% de variabilité.

Nous présentons ensuite une technique *d'externalisation des données* qui, pour le cas des semi-conducteurs fait gagner énormément en flexibilité et en maintenance des modèles. Dans ce chapitre nous répondons à la question « quelles sont les données les plus critiques à externaliser ? ». Soulignons que l'externalisation des données est particulièrement très bien adaptée aux unités de fabrication en pleine extension. Afin de pouvoir aborder le chapitre suivant un plan d'expérience est mis en oeuvre afin d'obtenir un maximum de renseignements sur les paramètres clés avec le minimum d'expérimentations.

5.2 Modélisation de l'unité de fabrication de Crolles 300

« Tous les modèles sont faux, mais certains sont utiles ».

Georges Box.

Rappelons que l'objectif de la modélisation est double :

- **accompagner le manager dans la compréhension du système existant** en modélisant le plus fidèlement possible ce que l'on a compris du système. Cela se traduit concrètement par des mises à jour permanentes du modèle par rapport au système réel tout au long de la thèse.
- **fournir un outil fiable afin d'améliorer les paramètres critiques** du système. La fiabilité se traduisant par la confiance du manager et de l'équipe AMHS de Crolles 300 au modèle.

5.2.1 Introduction

La simulation est de plus en plus utilisée dans tous les secteurs de la vie courante. Dans le secteur automobile par exemple, elle est utilisée pour les crashs test car c'est moins cher que le test réel. En météorologie pour la prévisions de catastrophes naturelles permettant de sauver des vies humaines. En aviation pour les simulateurs de vol. Cette montée en puissance de la performance des ordinateurs est particulièrement utilisée par les scientifiques. En effet, les modèles de simulation numérique sont maintenant soutenus par les modèles de simulation à événements discrets et la 3D, formidable outil de communication.

Il est important que les modèles de simulation puissent être facilement reproductibles. De nos jours, les différents domaines de spécialité sont à l'interface les unes des autres. Un modèle dans un domaine donné est susceptible d'être modifié et utilisé dans un autre domaine. Les modèles doivent donc être le plus explicite possible.

D'après [11], les relations mathématiques contiennent des informations sur différents types de processus et, connaissant une partie des informations de cette relation,

on peut en déduire l'information manquante. Par exemple dans la relation $F = M \times a$ (masse fois accélération), on est capable connaissant la masse de l'objet et l'accélération qu'il subit, d'en déduire la force qui lui est appliquée **sans même avoir dû la mesurer** (dû à l'égalité). Ainsi pour plusieurs relations successives décrivant un processus plus complexe, l'information donnée par une relation peut à son tour servir comme nouvelle information de la relation suivante. De proche en proche on arrive donc à connaître toutes les informations de toutes les relations qui décrivent le processus. **L'ensemble des relations décrivant ainsi un processus donné s'appelle le *modèle* ou « copie mathématique ».** *La modélisation est donc la conception d'un modèle.*

Le fait de faire fonctionner ce modèle sur un ordinateur s'appelle une « **simulation** ». On « simule » donc quelque chose de réel par quelque chose de mathématique. La simulation d'un modèle sur ordinateur permet donc d'obtenir toutes les informations désirées sur un processus en n'ayant plus besoin de tester le processus réel lui-même. Le modèle ainsi créé permet d'effectuer tout type de tests, de simulations, d'analyses et de mises au point, rapidement et sans limites ou contraintes, sans risques et à coûts très réduits.

La suite de cette section est divisée comme suit. Nous présenterons les définitions que nous avons adoptées pour la modélisation et la simulation à la section 5.2.2. Trois systèmes sont essentiels au fonctionnement de l'unité de fabrication de semi-conducteur : le transport, la production et le stockage. Si on veut avoir un modèle le plus complet possible il faut modéliser ces trois systèmes. La modélisation du système de production est décrite à la section 5.2.3. Deux équipements échangent des lots à l'aide du système de transport dont le modèle sera décrit à la section 5.2.4. Les lots en attente doivent être stockés. La section 5.2.5 décrit le système de stockage. Nous terminerons cette section par la présentation de notre approche d'externalisation des variables décrite à la section 5.2.6 et particulièrement bien adaptée pour les unités de fabrication en construction.

5.2.2 Définitions

Nous donnons dans cette section les définitions que nous avons adoptées et qui proviennent de [6], [8].

Modèle : représentation actuelle d'un système. Il doit contenir suffisamment de détails pour répondre aux questions que l'on se pose, mais pas plus que l'on en a besoin.

Évènement : occurrence qui change l'état d'un système. On distingue deux types d'évènements : les évènements internes appelés *endogènes* et ceux externes appelés *exogènes*.

Modèle à évènements discrets : représentation des composants d'un système et de leurs interactions.

Modèle mathématique : représentation mathématique (sous forme de formule) des composants d'un système, dans un objectif de « résolution ».

Entité : représente un objet nécessitant une définition explicite. Elle peut être dynamique ou statique. Par exemple un lot est une entité dynamique et une ressource une entité statique.

Nous utiliserons un modèle de simulation à évènements discrets. Afin de bien fixer les idées un modèle contient plusieurs entités. Les caractéristiques d'une entité sont décrites à l'aide d'attributs. A l'intérieur d'un modèle les entités interagissent entre elles à l'aide de procédures qui utilisent leurs attributs.

Nous ne prenons pas en compte les pannes d'équipements, de véhicules ou de stokeurs. Néanmoins, il est tout à fait possible d'arrêter un équipement durant la simulation afin d'observer et de récolter des données. Pour les véhicules, c'est plus difficile car un véhicule à l'arrêt peut bloquer le transport et parfois toute la simulation. Nous ne modélisons pas non plus les règles de calculs de priorité¹ des lots. Nous supposons donc que les lots choisis ont déjà la bonne priorité.

1. elles influent surtout sur la variabilité du temps de cycle des lots.

5.2.3 Modélisation de la production

Le parc d'équipements de l'unité de fabrication de Crolles 300 se compose d'environ 350 machines, dont plus d'une centaine sont accessibles par le système de transport automatisé. Les autres machines n'étant accessibles que par les opérateurs.

Les équipements sont regroupés en huit ateliers, et chaque atelier est divisé en une ou plusieurs zones qui occupent différentes positions géographiques. Par exemple les trois zones Etch 1, Etch 2, Etch 3 de l'atelier de gravure ne sont pas aux mêmes endroits. Chaque équipement possède entre deux et quatre ports de chargement/déchargement et une ou plusieurs chambres de traitement de lots. Les équipements sont des entités statiques. Dans la suite nous utiliserons le terme port de chargement pour désigner le port de chargement/déchargement.

Modélisation des équipements

De manière générale, les équipements sont modélisés sous forme de ressources et les ports de chargement sous forme de files d'attente de capacité unitaire. Un équipement de production se compose de ports de chargement, d'un bras mécanique interne servant à prendre les wafers, et d'une ou de plusieurs chambres de traitement.

Les ports de chargement et les chambres de traitement des lots ont été modélisés sous forme de files d'attente à capacité finie. Le statut d'un port de chargement est modélisé sous forme binaire (0 libre, 1 occupé). Le bras mécanique interne de l'équipement sert à prendre les wafers après l'ouverture du container pour les mettre dans la chambre de traitement. Il est modélisé sous la forme d'une ressource de capacité unitaire.

Toutes les durées opératoires sont modélisées sous forme de temps d'attente. Le temps écoulé entre le dépôt du container sur le port de chargement et son ouverture par l'équipement s'appelle le *setup time*. Il est lui aussi modélisé par un temps d'attente. Une fois la durée opératoire terminée, l'équipement referme le container et le lot recule légèrement et attend d'être déchargé. Ce temps est appelé *undock time*.

Le tableau 5.1 résume les différentes entités modélisées sur les équipements de

production.

Noms	Modélisé
Port de chargement	oui
Port de déchargement	oui
Statut du port de chargement	oui
Bras mécanique interne	oui
Chambre de traitement	oui
Durée opératoire	oui
Durée opératoire par chambre	non
<i>Setup time</i>	oui
<i>Undock time</i>	oui

TABLE 5.1 – Modélisation des équipements.

Certains équipements, en général ceux de métrologie², ne traitent pas tous les wafers d'un lot, mais seulement un nombre prédéfini. On dit que l'équipement fait du *sampling*. Il est modélisé en ajustant le temps de traitement des lots sur les machines. Par exemple pour un *sampling* de deux, c'est-à-dire qu'on ne mesure que le premier et le dernier wafers, le temps de traitement sera de : *sampling* multiplié par la durée opératoire d'un wafer exprimée en minutes. Les équipements utilisant le *sampling* ont en général une durée opératoire courte, c'est-à-dire de quelques minutes.

Les données sur le temps de traitement par chambre sont difficiles à obtenir pour certains équipements. Les temps de traitement n'ont donc pas été pris en compte dans le modèle de production.

La figure 5.1 illustre deux équipements de production de Crolles 300 modélisés sous AutoMod. L'éditeur utilisé est appelé **ACE** et permet la création d'images avec l'extension *.cell*³.

Importance de la modélisation des équipements. Un point très important dans la modélisation des équipements de production est que le modèle doit être

2. Équipement de mesure

3. Format ASCII représentant une entité en 3D.

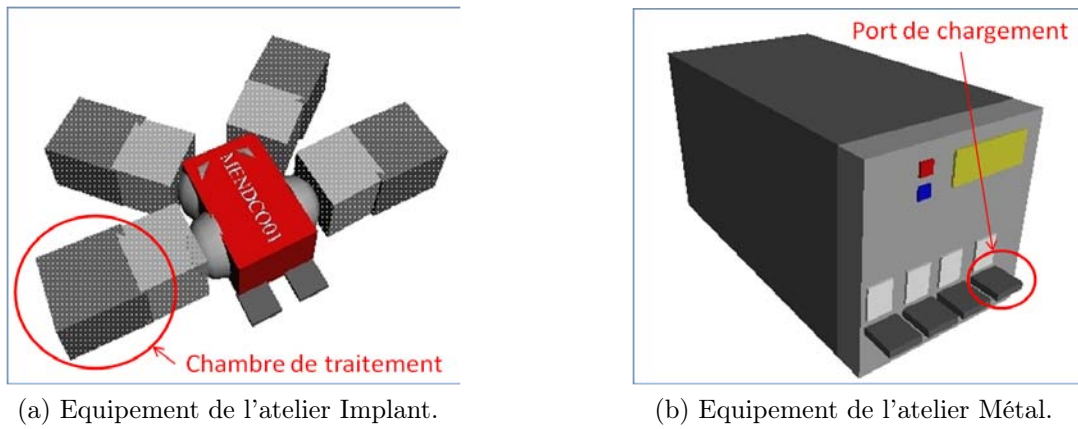


FIGURE 5.1 – Illustration des équipements de production sous AutoMod.

capable de respecter la cadence des lots qu'elle traite. Autrement dit, la répartition de la charge des lots sur les équipements. Si un niveau de détails n'est pas atteint l'équipement risque de surestimer ou de sous estimer le nombre de lots qu'il traite. Nous y reviendrons au moment de la validation du modèle de simulation. Le tableau 5.2 représente une comparaison de la charge de répartition des lots sur des équipements à fort encours. Ce tableau a été obtenu durant la période de validation du modèle. Il montre que le nombre de lots traités par jour et pour les équipements choisis de Crolles 300 (colonne « Unité réelle ») est supérieur à ceux obtenus par simulation (colonnes « Simul[1] » et « Simul[2] »). La différence entre les deux simulations est que dans la seconde simulation, il y a amélioration de la répartition de la charge des lots. On peut s'en rendre compte par les colonnes « Facteur[1] » et « Facteur[2] ». Elles représentent respectivement le facteur multiplicatif entre les résultats de Crolles 300 et ceux des simulations. L'amélioration est illustrée par le passage d'un facteur multiplicatif moyen de 8 à 2.

Le nombre de lots traités par un équipement représente indirectement l'intensité du trafic de la zone pour une période donnée. Moins il y a de lots traités dans la zone sur une période de temps, moins il y a de trafic. Donc la répartition de la charge des lots peut renseigner sur une éventuelle congestion ou embouteillage : équipements moins alimentés ou plus alimentés que d'autres. Mais on ne peut pas dire où passe

	Unité réelle	Simul[1]	Facteur[1]	Simul[2]	Facteur[2]
Equipement 1	75	9	8	37	2
Equipement 2	79	7	11	33	2
Equipement 3	45	24	2	-	-
Equipement 4	24	24	1	-	-
Equipement 5	118	14	8	53	2
Equipement 6	123	17	7	60	2
Equipement 7	114	14	8	54	2
Total	578	109	-	291	-

TABLE 5.2 – Répartition des lots en lithographie (lots/jour).

cet embouteillage. En outre, respecter la répartition de la charge des lots permet de détecter et de corriger d'autres types de problèmes très importants. Par exemple, la somme totale des lots sur les sept équipements est de 578 lots pour la colonne « Unité réelle » contre 109 lots pour « Simul[1] » et 237 lots pour « Simul[2] ». Réduire ces différences, c'est apporter une réponse à une ou plusieurs des questions suivantes :

- (a) l'unité de fabrication est-elle assez chargée en lots ? Si on a confiance dans le modèle alors une mauvaise répartition des lots peut être due au taux de charge des lots en entrée.
- (b) le régime stationnaire de la simulation est-il atteint ? Si le régime stationnaire n'est pas atteint la cadence des lots sur les équipements n'est pas encore stabilisée.
- (c) les lots sont-ils retardés quelque part dans la simulation ? Si un ou plusieurs lots sont retardés, les équipements directement en amont ou en aval seront eux aussi impactés.

Conclusion sur la modélisation des équipements de production. Reproduire l'unité de fabrication à l'identique est l'idéal recherché par les managers des unités de fabrication. Avec 350 équipements à prendre en compte la difficulté est énorme. Mais bien que les capacités des ordinateurs soient en constante progression, la simulation requiert un long temps de calcul afin de produire des résultats statistiques valides [74]. D'autre part, comme on vient de le montrer, négliger un certain nombre de détails a d'importantes conséquences sur la validité du modèle.

Cela traduit ainsi la *complexité de la modélisation fidèle d'une unité de fabrication*.

Modélisation des lots

Le fonctionnement de l'unité de fabrication réelle est basé sur le **contrôle par le produit**, c'est-à-dire que c'est le produit, donc le lot, qui est l'élément déclencheur des processus. Un lot est modélisé (figure 5.2) de manière générale par son type, sa technologie et ses attributs. Le type du lot permet de distinguer les lots de production, des lots de test, ou des lots engineering par exemple. La technologie du lot détermine la route du lot, c'est-à-dire le nombre d'étapes nécessaires pour la réalisation du produit final. Les attributs du lot représente les différentes informations nécessaires à son interaction et à son traitement par d'autres entités. Ces éléments sont résumés dans la tableau 5.3.

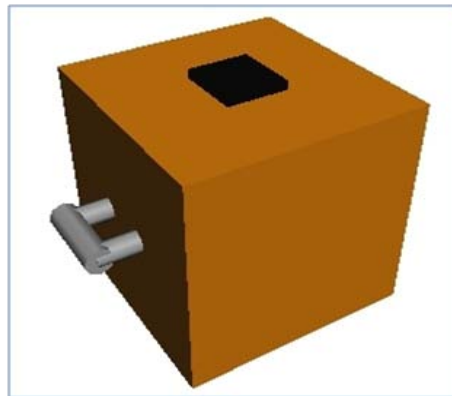


FIGURE 5.2 – Illustration d'un lot de production sous AutoMod.

Noms	Modélisé
Type du lot (production, engineering, etc.)	oui
Technologie du lot	oui
Attributs du lot	oui

TABLE 5.3 – Modélisation des lots.

La modélisation du contrôle par le produit consiste à affecter aux produits les in-

formations nécessaires à leurs parcours. Dans ce mode de fonctionnement, à chaque passage d'un lot d'une procédure à une autre, d'une fonction à une autre ou d'un modèle à un autre, le lot doit être capable de fournir diverses informations afin de pouvoir être traité. Par exemple, un lot doit pouvoir fournir sa position physique, son nom, la position de l'emplacement unitaire de stockage où il se trouve, le nom de l'équipement de destination, son numéro d'étape, le port de chargement de destination, etc. Le lot étant considéré comme une entité, il possède des attributs par défaut et d'autres créés par le programmeur. Tous les attributs utilisés sont regroupés dans le tableau 5.4. Il n'y a pas une liste standard d'attributs de lot. Les attributs sont choisis en fonction du niveau d'interaction avec le modèle et du choix des informations que désire le programmeur.

Attributs	Modélisé
Position en cours du lot	oui
Nom de la dernière position du lot	oui
Nom du stockage unitaire trouvé	oui
Identité de l'emplacement unitaire dédié au lot	oui
Code-barres de l'emplacement unitaire dédié au lot	oui
Numéro de l'emplacement unitaire	oui
Nom du stockeur par défaut	oui
Nom du stockeur (si associé)	oui
Nom de l'alternative de stockage	oui
Equipement devant traiter le lot	oui
Numéro de l'étape	oui
Numéro de l'étiquette du lot	oui
Temps de cycle du lot	oui
Temps de transport du lot	oui
Temps de livraison du lot	oui
Etat de la liste des tâches à faire	oui
Etat du port de chargement	oui
Numéro du port de chargement du lot	oui
Temps de changement de lot sur un port de chargement	oui
Nom de la liste des tâches à faire de l'équipement	oui

TABLE 5.4 – Liste des attributs des lots

Modélisation de la route du lot

La « route » d'un produit est la succession des différentes étapes de fabrication de ce produit et est fonction du type de produit à fabriquer. Il existe donc plusieurs routes de produits, et le plus petit commun multiple de ces routes s'appelle la **route générique**. Tandis que la route du produit le plus fabriqué dans l'unité de fabrication s'appelle la route du **big runner**. Nous avons principalement modélisé ces deux routes (tableau 5.5).

Noms	Modélisé
Route générique	oui
Route du <i>big runner</i>	oui

TABLE 5.5 – Liste des routes modélisées.

Une route se modélise en faisant correspondre à chaque étape de fabrication un numéro d'étape. Ces numéros sont ensuite alignés de manière séquentielle à l'intérieur d'un fichier pour former la route du lot. Dans ce fichier, un phénomène réentrant correspond à une répétition de numéro d'étapes. Nous avons considéré dans cette thèse **trois types de produits** à cause de la complexité du problème. On réduit alors le nombre de lots différents en compétition devant un équipement de production.

La route générique rend plus difficile la mesure de l'écart entre simulation et réalité, car le lot visite plus d'étapes avec cette route. Elle ne correspond à aucune route utilisée dans la réalité. Les résultats fournis par des indicateurs comme la répartition de la charge des lots, le temps de cycle, de livraison, etc... sont utilisés mais ils ne peuvent pas être comparés à la réalité. C'est un peu comme si on compare les indicateurs de fabrication d'une voiture et d'un train. Pour réduire cet écart de mesure, on utilise la route d'un lot spécifique, celle du *big runner* qui elle est mesurable.

Conclusion sur la modélisation de la production. La modélisation de la production est très importante dans la modélisation d'une unité de fabrication car,

c'est le niveau de détails du modèle de production qui va « fixer » celui du modèle de transport. Rappelons que le transport est un service pour la production.

5.2.4 Modélisation du transport

La modélisation du transport peut se diviser en deux parties : la modélisation du support de transport et la modélisation de la flotte de véhicules.

Modélisation du support de transport

Le support de transport est modélisé à l'aide de rails situés au plafond à un peu plus de quatre mètres du sol. Chaque rail possède trois éléments essentiels : son sens, son lien de transfert et son rayon de courbure. Un rail peut autoriser le déplacement unidirectionnel, bidirectionnel ou latéral des véhicules. **Tous les rails modélisés dans notre modèle sont unidirectionnels car c'est le cas de Crolles 300.** Le lien de transfert permet aux véhicules de passer d'un rail à un autre sans interruption. Le rayon de courbure permet de déterminer très précisément l'angle d'un virage. Il vaut 0 pour les lignes droites et permet de déterminer à quelle vitesse un véhicule doit négocier un virage. Outre le transport les rails servent également de support pour les codes-barres. Le tableau 5.6 représente un résumé des éléments modélisés pour le support de transport.

Noms	Modélisé
Rails	oui
Courbure du rail	oui
Codes-barres	oui

TABLE 5.6 – Modélisation du support de transport.

Le rail est une entité donc possède des attributs représentés par le tableau 5.7. Le *navigation factor* est un nombre qui multiplie la longueur d'un rail. Il est utilisé pour la détermination du plus court chemin par les véhicules. Sa valeur par défaut est 1. L'augmentation du *navigation factor* permet de forcer un véhicule à éviter un segment de rail à moins que sa destination ne se trouve sur ce rail. Par exemple un

rail de 10 m avec un *navigation factor* de 2, est perçu par les véhicules comme un rail de 20 m.

Noms	Modélisé
Chemin bidirectionnel	non
Chemin unidirectionnel	oui
<i>Navigation factor</i>	oui
Vitesse limite sur les rails	oui
Rayon de courbure du rail	oui

TABLE 5.7 – Liste des attributs du support de transport modélisé.

Un point de contrôle est une location sur un rail où le véhicule peut s'arrêter et recevoir des informations. Les codes-barres sont des entités modélisés par des points de contrôle et donc ils possèdent des attributs. Pour se déplacer, les véhicules interrogent successivement les codes-barres qu'ils rencontrent. Un trajet représente donc une succession de code-barres. Le *nombre d'appels* d'un code-barres détermine le nombre de véhicules qui interrogent simultanément ce code-barres à un instant donné. Tandis que, la *capacité du code-barres* détermine le nombre maximum d'appels simultanés autorisés sur le code-barres. Par défaut la capacité d'un code-barres est infini, ce qui permet à un nombre non limité de véhicules de se déplacer sur le segment de rail concerné. On a toujours *nombre d'appels du code-barres* \leq *capacité du code-barres*. Les codes-barres servent également de points de chargement de lots. Le tableau 5.8 représente la liste de attributs considérés pour les codes-barres.

Noms	Modélisé
capacité du code-barres	oui
distance de libération du code-barres	oui
nombre d'appel du code-barres	oui

TABLE 5.8 – Listes des attributs des codes-barres.

Importance des détails dans la modélisation du support de transport.

Un véhicule n'aborde pas un virage d'angle $\alpha \neq \beta$ de la même manière (vitesse,

accélération). À vitesse identique, plus l'angle est petit, plus la force centrifuge est forte. Les spécifications du constructeur MURATEC recommandent que les wafers à l'intérieur du lot ne subissent pas une force centrifuge de plus de 3G, sinon le système de transport ne sera plus capable d'assurer l'intégrité physique du lot. Donc minimiser la courbure des rails par exemple dans la modélisation, c'est prendre le risque de surestimer les capacités en temps de déplacement des véhicules donc des performances du système de transport. Cette surestimation n'est pas facile à corriger car, très difficile à détecter après modélisation.

Pour illustrer l'importance des détails dans les rails, nous présentons une étude demandée en début de thèse, par un spécialiste en configuration de rails de Crolles 300. L'objectif est de comparer les performances de deux tronçons de rails, *layout A* et *layout B*, par évaluation du temps mis par les véhicules pour faire un tour. Le tracé du *layout B* nécessite de déplacer un équipement de production. Tandis que le *layout A* permet de contourner cet équipement. L'enjeu est que déplacer un équipement est beaucoup plus fastidieux que de le contourner. Les deux tronçons se différencient par l'existence d'un double virage sur le trajet du *layout A*. Les deux tronçons sont illustrés par la figure 5.3. Les véhicules se déplacent dans le sens des aiguilles d'une montre.

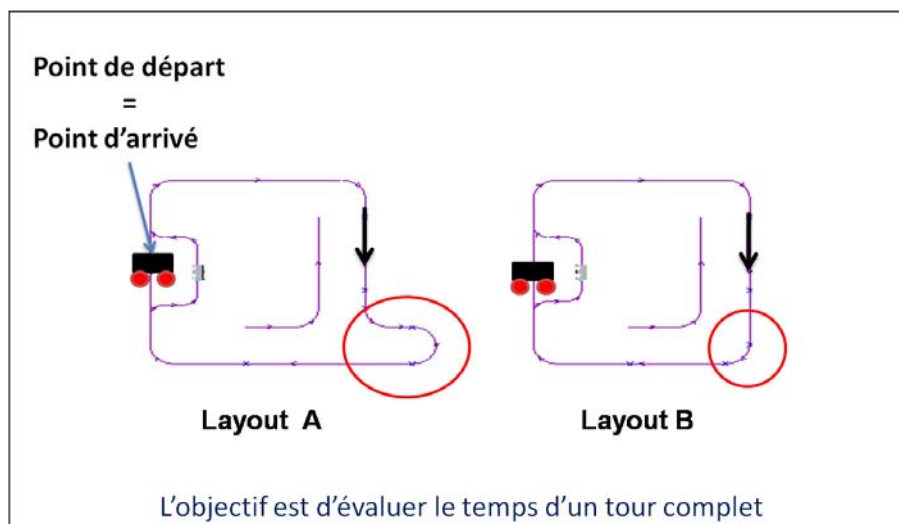


FIGURE 5.3 – Comparaison de deux tronçons pour un tour complet.

Le tableau 5.9 représente les résultats numériques de cette comparaison et les valeurs sont exprimées en *secondes*. La colonne « Temps de transport » représente le temps pour faire le tour complet d'un tronçon. La colonne « Temps du virage », le temps perdu sur la portion de virage étudiée (partie encadrée dans le *layout* A). Enfin, la colonne « % temps total » représente le pourcentage du temps perdu sur la portion étudiée par rapport au temps total nécessaire pour faire un tour. Les résultats montrent que l'on perd 51% du temps total d'un tour complet dans le double virage du *layout* A. Cet exemple illustre les enjeux et l'importance de bien modéliser le support de transport. Au final, l'équipement sera déplacé.

	Temps de transport	Temps du virage	% temps total
Configuration A	5,44	2,76	51%
Configuration B	3,28	0,80	24%
Différence	2,16	1,96	-

TABLE 5.9 – Résultats de la comparaison des deux tronçons (*en secondes*).

Conclusion sur l'importance des détails dans la modélisation du support de transport. Tenir compte de détails sur la modélisation du support de transport, c'est augmenter le niveau de confiance des décideurs sur le modèle. Dans cette étude *l'angle des virages et la longueur des arcs de virage* sont les éléments clés de la validité du modèle. Lorsque le support de transport est bien modélisé, une partie des performances du système repose sur la flotte de véhicules, c'est ce que nous verrons dans la suite.

Modélisation de la flotte de véhicules et du routage

Nous venons de voir qu'un véhicule n'a pas les mêmes performances (vitesse, accélération) sur une ligne droite ou sur un virage. Le support de transport influe donc sur les performances des véhicules. La performance du système de transport est fortement liée aux caractéristiques des véhicules [14]. Un véhicule est une entité et donc possède des attributs. Puisque l'on utilise du contrôle par le produit, le

véhicule utilise les mêmes attributs que ceux des lots. La figure 5.4 représente une illustration d'un véhicule sous AutoMod.

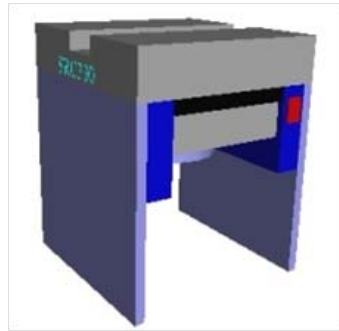


FIGURE 5.4 – Illustration d'un véhicule aérien sous AutoMod.

Par défaut, le véhicule choisit le chemin le plus court. En cas d'encombrement du chemin, le véhicule doit pouvoir changer de trajet. Pour évaluer le niveau d'encombrement d'un chemin, on utilise la notion de *navigation factor*. On rappelle que c'est un facteur multiplicatif de la longueur du rail. Plus le *navigation factor* est grand, plus le risque d'embouteillage est élevé. Le *navigation factor* est incrémenté d'une unité dès qu'il y a un certain nombre de véhicules sur la portion de rail. Ce nombre de véhicules est appelé *véhicule seuil*. Il est fonction de la longueur du rail et de la distance de freinage des véhicules. Il est ensuite décrémenté dès que les véhicules en « trop » quittent le rail. Les éléments utilisés pour la modélisation de la flotte de véhicules sont présentées dans le tableau 5.10.

Conclusion sur la modélisation de la flotte de véhicules. Les véhicules sont des entités mobiles de l'unité de fabrication. Ils servent d'interface entre les lots et les équipements. Leurs performances sont fortement dépendantes des caractéristiques prises en compte dans la modélisation de la flotte de véhicules. Les véhicules utilisent les mêmes attributs que ceux des lots car on est dans un contrôle par le produit.

Noms	Modélisé
Type de véhicules	oui
Temps de chargement	oui
Temps de déchargement	oui
Vitesse en ligne droite	oui
Accélération	oui
Décélération	oui
Distance de freinage	oui
Distance d'arrêt	oui

TABLE 5.10 – Modélisation de la flotte de véhicules.

5.2.5 Modélisation du stockage

Un véhicule, le plus performant soit-il, voit ses performances réduites si le temps gagné en transport est perdu ailleurs. En effet, le temps d'accès aux emplacements de stockage est un facteur pouvant être important pour les temps des scénarios de chargement et de déchargement de lots, comme le montre la figure 5.5. Le stockage est donc un élément agissant sur les performances du système de transport.

Les stockeurs sont modélisés dans la littérature sous forme de files d'attente de capacité infinie [37], car les auteurs s'intéressent le plus souvent à la politique de gestion des véhicules ou de gestion des lots. Mais dans cette thèse nous modélisons la capacité réelle des stockeurs : **tous les stockeurs sont donc de capacité finie**, c'est-à-dire qu'ils ont un nombre limité d'emplacements unitaires de stockage (entre 100 et 500) dans le cas du système de Crolles 300.

Modélisation des stockeurs

Un stockeur inclut un robot, sorte de bras mécanique, qui permet de déplacer les lots depuis les ports internes vers les ports externes et vice-versa. Le port interne est relié au port externe par un mini convoyeur, sorte de tapis roulant dont le temps de cycle est d'environ 10 secondes. Le robot à l'intérieur du stockeur possède aussi son propre temps de cycle qui est d'environ 20 secondes. La figure 5.6 présente une vue du stockeur de profil.

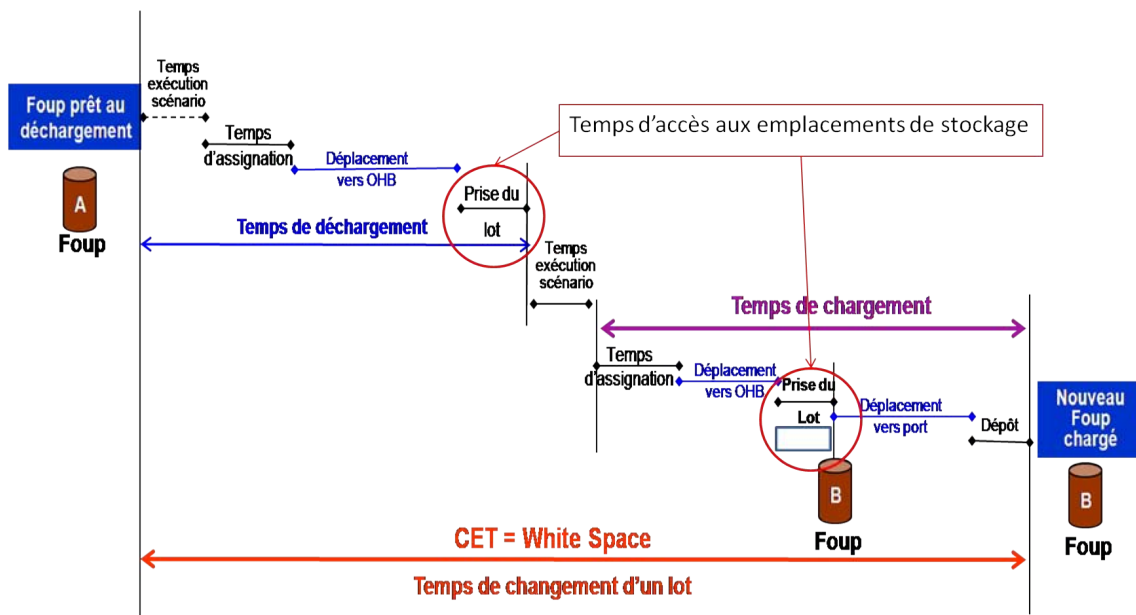


FIGURE 5.5 – Importance des temps d'accès au stockage.

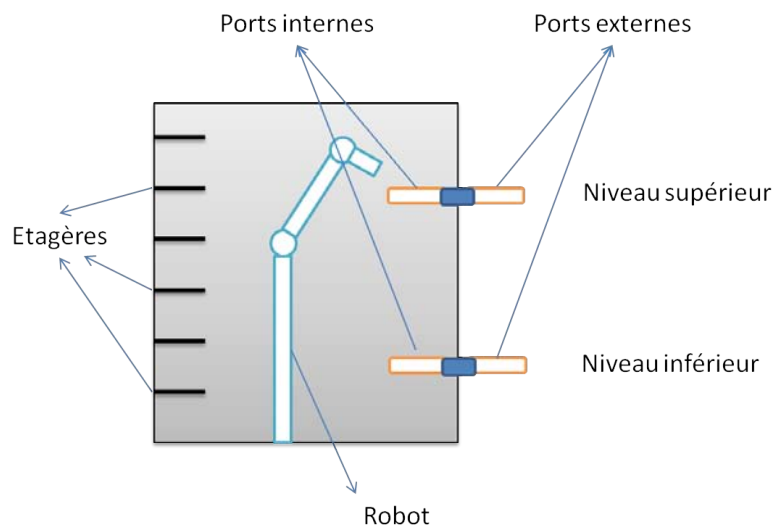


FIGURE 5.6 – Stockeur vu de profil.

Chaque stockeur de Crolles 300 possède quatre ports : deux ports au niveau in-

férier pour le chargement/déchargement manuel et deux ports du niveau supérieur pour le système de transport automatique. Dans notre modèle, lorsqu'un opérateur veut prendre un lot, il passe par le stockeur. Le véhicule dépose le lot sur le port de chargement supérieur, le robot transfère le lot sur le port inférieur et l'opérateur récupère le lot sur le port de déchargement inférieur. Cette coexistence entre humains et machines correspond au mode hybride. La figure 5.7 représente une vue du stockeur de face.

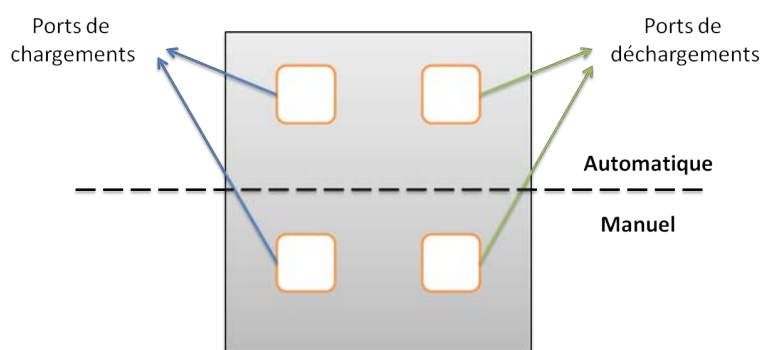


FIGURE 5.7 – Stockeur vu de face.

Les différentes entités et caractéristiques qui sont intervenues dans la modélisation d'un stockeur sont représentées dans la tableau 5.11.

Noms	Modélisé
Robot à l'intérieur du stockeur	oui
Temps de cycle du robot	oui
Ports internes	oui
Ports externes	oui
Temps de cycle du mini convoyeur	oui
Capacité du stockeur	oui
Emplacements de stockage internes du stockeur	non

TABLE 5.11 – Modélisation d'un stockeur.

Le stockeur peut devenir un goulot d'étranglement. Pour le démontrer, étudions un cas très courant du système réel, celui d'un véhicule qui vient prendre un lot se trouvant à l'intérieur d'un stockeur. Il faut compter **un minimum de 50 secondes**. En effet, il faut :

- 20 secondes au robot du stockeur pour aller chercher le lot,
- 20 secondes pour le ramener sur le port interne et,
- 10 secondes pour que le lot rejoigne le port externe.

Pendant, le lot n'est toujours pas sur le véhicule, il faut encore ajouter le temps de prise du lot sur un stockeur qui est de 12 secondes, soit **un total 1 minute environ**. Ce temps est relativement énorme surtout si l'on tient compte de l'ordre de grandeur dans lequel il intervient. Par exemple considérons un équipement dont le temps de changement de lot est de 200 secondes. Une minute, c'est 30% du temps de changement d'un lot sur un équipement. Si, de plus, on ajoute à ce temps :

- le temps de transport,
- le temps d'assignation qui peut être très variable,
- le rang de la requête de transport, c'est-à-dire son numéro dans l'ordre des requêtes. Si c'est la troisième requête, il faut d'abord que les deux premières s'exécutent.

On peut alors dépasser le temps de changement d'un lot sur un port de chargement d'un équipement. On augmente ainsi le temps de non-production de l'équipement. Le stockeur peut ainsi rapidement devenir un goulot d'étranglement. Dans la mesure du possible, il doit donc être utilisé de manière pertinente.

Modélisation du regroupement des emplacements unitaires de stockages

Les emplacements de stockage unitaire sont latéralement accrochés aux rails sur lesquels se déplacent les véhicules. Ils sont physiquement repérables à l'aide du code-barres situé sur les rails et en face de chaque emplacement de stockage. Les emplacements unitaires de stockage ont un temps de chargement/déchargement de 8 secondes, soit **6 fois** plus rapide qu'un stockeur. Cette rapidité d'accès en fait un élément privilégié pour les lots de production.

Un *stockeur par défaut* est un espace de stockage possédant **au minimum** un emplacement unitaire de stockage. Cette définition implique qu'un stockeur peut

être un stockeur par défaut. **Tout équipement de l'unité de fabrication est associé à au moins un stockeur par défaut** et, plusieurs équipements peuvent partager un même stockeur par défaut.

Un emplacement unitaire de stockage est une entité modélisée par une file d'attente de capacité unitaire. A Crolles 300 ces emplacement ne possèdent aucune mécanique, les attributs sont donc ceux par défaut du logiciel. Le regroupement des emplacements unitaires de stockage se fait à l'aide de fichier externe. Ce fichier contient le nom de tous les emplacements unitaires de stockage formant le stockeur par défaut. Les fichiers externes sont présentés dans la section suivantes. Le tableau 5.12 donne la liste des attributs des emplacements unitaires de stockage.

Noms	Modélisé
Nom du stockage unitaire	oui
Position stockage unitaire (code-barres)	oui
Nom du default stockeur associé	oui

TABLE 5.12 – Liste des attributs d'un emplacement unitaire de stockage.

5.2.6 Externalisation des données

Un modèle est valide pour une application particulière si son algorithme (sa logique) est correct **et** s'il utilise les données appropriées. Il est très important que la personne qui analyse les données comprenne le procédé par lequel elles sont produites, plutôt que de traiter les données juste comme des informations abstraites. On peut s'en rendre compte à travers une description des principales erreurs fournies par [58] :

- **Données non représentatives par rapport aux besoins du modèle** : par exemple les données recueillies sur un test de champ de bataille ne sont pas représentatives des combats réels à cause de la différence du comportement des troupes et du manque de fumée par exemple,
- **Données n'ayant pas le format désiré** : par exemple, le temps entre 2 pannes d'un équipement doit inclure le temps pendant lequel la machine est au repos (non production) et, le temps de changement d'équipe.

- **Données pouvant contenir des mesures, des enregistrements ou des erreurs** : par exemple les arrondis des temps de traitement qui peuvent affecter le choix de la distribution de probabilité,
- **Données pouvant être biaisées** : par exemple les équipes de maintenance augmentent parfois la fiabilité de certains équipements afin d’être bien vu par leurs supérieurs
- **Données récoltées dans des conditions inconnues** : par exemple les données de la réparation d’un satellite.

Dans la littérature, il est très difficile de trouver des articles qui traitent de la problématique de la mise à jour des données dans la simulation des semi-conducteurs. Cela peut s’expliquer par la quasi absence de modèles de simulation complets d’une unité de fabrication réelle.

Une fois les données (de grandes tailles) introduites en « dur » dans un logiciel, il est très fastidieux, **quelque soit le logiciel utilisé**, d’effectuer **certaines** modifications. On aimerait pouvoir faire par exemple des mises à jour des modèles de simulation le plus *facilement* et le plus *rapidement* possible. L’intérêt majeur de la mise à jour des modèles est double, il permet :

1. **un gain de temps** : en effet, on n’est pas obligé de complètement recoder le modèle pour prendre en compte les nouvelles données. Le système industriel considéré dans cette thèse a nécessité *pas moins de 6 mois* de développement.
2. **une maîtrise de l’erreur entre simulation et réalité** : une simulation est une « photo » à un instant donnée de la réalité. Si on supprime un des éléments sur la photo, celle-ci est moins représentative de la réalité : on augmente ainsi le niveau d’erreurs entre simulation et réalité. Maîtriser ces erreurs c’est être capable de donner une valeur à tout instant de l’erreur globale. Par exemple, elle est de 15%⁴ (en accord avec le manager) pour le cas de la simulation du système industriel étudié dans cette thèse. Nous y reviendrons plus en détails dans la section validation de la simulation.

4. choisi afin de prendre en compte deux aspects : (1) la fréquence importante des mises à jour qui entraîne une forte variabilité entre le modèle et la réalité et (2) le mode « hybride » (cohabitation homme-machine) de l’unité de fabrication également source de variabilité

Externalisation. Une solution aux problèmes de mise à jour des données est l'externalisation des données. Cela permet au code d'avoir accès à des données situées dans un fichier, sous un format plus simple à manipuler. L'externalisation des données s'oppose au codage en dur, qui consiste à introduire les données directement dans le code. Dans le cas de ce dernier, la maintenance du modèle est très difficile surtout quand on a des données de grandes tailles. On s'affranchit ainsi de la connaissance de la logique du code si on veut effectuer des mises à jour des données ou des changements de scénarios.

Mais toutes les données ne sont pas nécessairement externalisables. Celles qui le sont, sont appelées **données sensibles**. *Une donnée sensible, pour notre cas, est une donnée dont toute modification entraîne un changement lourd en temps, et parfois en logique, du code informatique.* Par exemple, redimensionner le nombre d'emplacement unitaire de stockage d'un stockeur par défaut, c'est supprimer ou ajouter un ou plusieurs de ces emplacements. La difficulté se situe dans le processus de mise à jour des attributs des emplacements unitaires de stockage. En effet, ajouter un nouvel emplacement, c'est fournir au système, son nom (identifiant), son code-barres associé (sa position géographique) et le nom du stockeur par défaut auquel il appartient. Sans une de ces informations le système serait incapable de retrouver l'emplacement unitaire de stockage ajouté.

Données sensibles, le cas de Crolles 300. L'usine de Crolles 300 est en pleine extension, ce qui signifie qu'il y a **des déplacements ou des ajouts d'équipements** afin d'augmenter la capacité de l'unité de fabrication. Le *redimensionnement des stockeurs par défaut* doit permettre de maintenir ou d'augmenter l'encours de production, le **changement de la chaîne de stockage alternatif** doit permettre de réduire les temps d'accès aux lots, notamment en affectant les stockeurs en fin de liste, et enfin l'**ajout de rails** permet de passer en mode *Full-Auto*, c'est-à-dire sans interventions humaines.

A partir de ces remarques, on en déduit alors que les *données sensibles* de notre modèle concernent les données sur :

- les stockeurs par défaut (regroupement d'emplacements unitaires de stockage),
- les alternates (alternative de stockage),

- les routes des lots (gamme opératoire),
- les équipements ou groupes d'équipements.

L'externalisation des données constitue une part non négligeable du travail de modélisation de cette thèse. Nous avons utilisé 28 fichiers externes pour les stockeurs par défaut, 28 autres fichiers pour les alternates, 4 fichiers pour les routes de lots et 101 fichiers pour les équipements regroupés en familles de stations. Au total, nous utilisons donc pour notre modèle **161 fichiers externes**.

Exemple de fichier externe. Le tableau 5.13 représente un fichier texte qui correspond à l'externalisation des informations d'un stockeur par défaut. Les lignes sont entrées en fonction de l'ordre auquel on rencontre les emplacements unitaires de stockage sur les rails. Les deux premières lignes du fichier sont respectivement, le nom du stockeur par défaut et sa capacité en nombre d'emplacement (OHB).

En partant de la gauche, la première colonne représente le type d'emplacement unitaire de stockage, la seconde colonne représente le nom de l'emplacement unitaire de stockage en chaîne de caractère. La dernière colonne représente le nom du code-barres associé à l'emplacement unitaire de stockage.

Nom		
Nombre total d'OHB		
Type de l'emplacement	Nom de l'emplacement	Code-barres
OHB_1	$OHBname_1$	$CBid_1$
OHB_2	$OHBname_2$	$CBid_2$
—	—	—
—	—	—
OHB_{21}	$OHBname_{21}$	$CBid_{21}$
OHB_{22}	$OHBname_{22}$	$CBid_{22}$

TABLE 5.13 – Exemple d'externalisation des informations sur le stockeur par défaut.

Conclusion. L'externalisation des variables est particulièrement bien adaptée aux systèmes en construction mais, aussi pour la maintenance du modèle et le déploiement rapide de scénarios. En effet, avec les données externalisées, un simple copier-

coller suffit à l'ajout ou à la suppression d'un emplacement unitaire de stockage ou d'un équipement. Par exemple, dans le tableau 5.13, ajouter un emplacement unitaire de stockage revient à insérer une ligne contenant le nom de l'emplacement de stockage unitaire et les types correspondants, la position géographique du code-barres et d'ajouter 1 au nombre total d'emplacement de stockage unitaire du stockeur par défaut.

5.3 Simulation de Crolles 300

C'est l'affaire de quelques lignes de code...
Un employeur n'ayant jamais programmé.

Logiciel utilisé

Le logiciel AutoMod de APPLIED MATERIALS, est un outil informatique de modélisation et de simulation de **flux en 3D**. Le choix du logiciel se justifie par le fait que **AutoMod utilise un outil CAD**⁵, et permet de construire une **représentation fidèle** du système au 1/100^e de mm et de le simuler au 1/100^e de seconde nécessaire. Toutes ces fonctionnalités sont nécessaires autant théoriquement qu'industriellement pour modéliser l'unité de fabrication le plus fidèlement possible. Par exemple, la simulation de flux en 3D va nous permettre de pouvoir interagir plus facilement avec le manager à l'aide du support visuel et de faire de la simulation en temps réel.

Procédures et fonctions

La simulation est le fait de faire fonctionner un modèle sur un ordinateur. Dans cette thèse nous allons faire fonctionner les modèles de transport, de production et de stockage décrits précédemment. Nous nous baserons sur **quatre** grandes procédures :

1. *what next* ou chargement des lots sur un port,
2. *where next* ou déchargement des lots d'un port,
3. *vehicle management* ou gestion du fonctionnement des véhicules et,
4. *equipment management* ou gestion des machines et des stockeurs.

Les lots à forte priorité, *hot lots*⁶, ne sont pas traités dans cette thèse. Néanmoins la flexibilité du modèle permet d'intégrer ces lots.

5. CAD signifie Computer-Aided-Design : Ensemble d'outils et de programmes informatiques permettant d'assister l'ingénieur dans la conception d'un produit.

6. Un *hot lot* est un lot qui ne doit pas attendre, il est donc prioritaire et tous les lots doivent le laisser passer.

5.3.1 Procédure « what next »

C'est la procédure de chargement des lots sur le port d'un équipement. Un équipement qui ne traite pas de lot est difficilement rentabilisable à cause de son coût très élevé (près de \$75 millions pour un scanner en lithographie). Il faut donc veiller à ce que l'équipement, ne soit jamais désalimenté en lots. L'objectif de la procédure « what next » est d'apporter le plus rapidement possible, le bon lot sur le port vide d'un équipement. Il vise donc à réduire le temps d'attente à vide de l'équipement. La figure 5.8 illustre le fonctionnement de la procédure « what next ».

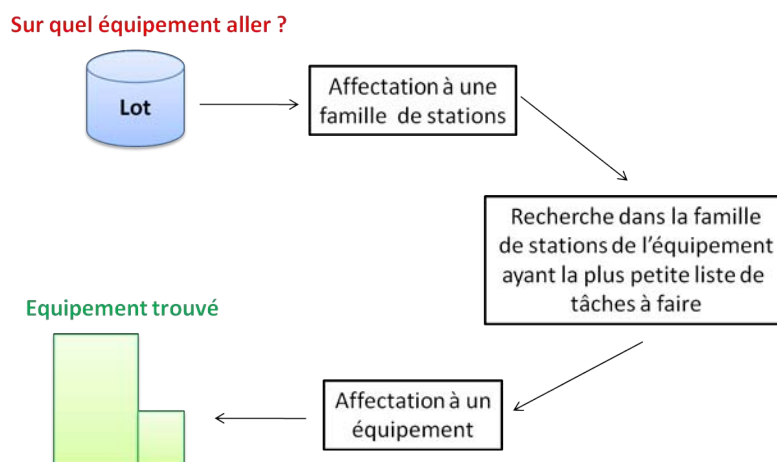


FIGURE 5.8 – Organigramme de la procédure « what next ».

Une famille de stations d'équipements est un regroupement d'équipements pouvant effectuer le même traitement à une étape donnée. Dans notre modèle et pour la route du « big runner », un lot visite 102 familles de stations, réparties en 602 étapes au cours de sa vie. Le nombre d'étapes étant supérieur aux nombres de familles de stations, les lots vont nécessairement visiter plusieurs fois une même famille de stations. Il existe donc une *surjection* entre le nombre d'étapes de fabrication et le nombre de familles de stations. Le tableau 5.14 résume les principales sous-procédures de la procédure « what next ».

La procédure « P_WhatNext » va attribuer un numéro de famille de stations au lot correspondant à son étape de fabrication. Le lot connaît à présent la famille de

Procédure	Objectifs
P_WhatNext	affectation d'un lot à une famille de stations
FewestQueue	recherche de la plus petite liste de tâches à faire
MACS	affectation d'un lot à un équipement

TABLE 5.14 – Les principales sous-procédures du « what next ».

station où il sera traité. Il faut maintenant trouver à l'intérieur de cette famille un équipement disponible pour le traiter. L'objectif est d'équilibrer la charge de travail des équipements dans une famille de station. Dans notre modélisation, chaque équipement possède une liste de tâches à faire. On recherche parmi les équipements pouvant traiter le lot, celui qui a la plus petite liste de tâches à faire (procédure « FewestQueue »). On se retrouve alors dans différents cas de figure :

- *la liste des tâches à faire est vide et l'équipement est disponible.* Dans ce cas l'équipement est directement sélectionné (procédure « MACS ») et le transport se déclenche.
- *la liste des tâches à faire est vide mais l'équipement n'est pas disponible.* Dans ce cas, le lot est mis en attente dans la liste des tâches à faire de l'équipement.
- *la liste des tâches à faire n'est pas vide.* Dans ce dernier cas, le lot est mis en attente dans la liste des tâches à faire.

Cas exceptionnel des équipements prioritaires. Un équipement est dit prioritaire lorsqu'il est très sollicité et possède un temps de cycle très court. C'est le cas par exemple de certains scanners en lithographie, ou d'équipement de mesure en métrologie. Pour ce type d'équipement, on ne choisit pas l'équipement uniquement en fonction de sa liste de tâche à faire, mais aussi en fonction de sa vitesse de traitement des lots. Les équipements les plus rapides seront choisis en premier.

5.3.2 Procédure « where next »

C'est la procédure de déchargement d'un lot d'un équipement. Lorsqu'un lot termine son traitement sur un équipement, il doit être déchargé du port de chargement de l'équipement. Dans ce cas, la destination du lot déchargé est un autre équipement

(s'il est disponible), du moment où ce n'est pas la dernière étape du lot. Bien souvent l'équipement n'est pas disponible. L'objectif de la procédure « where next » est de rapprocher le lot au plus près des équipements susceptibles de pouvoir le traiter. Si l'équipement est disponible, alors le lot est transporté directement sur l'équipement. La figure 5.9 illustre le fonctionnement de la procédure « where next ».

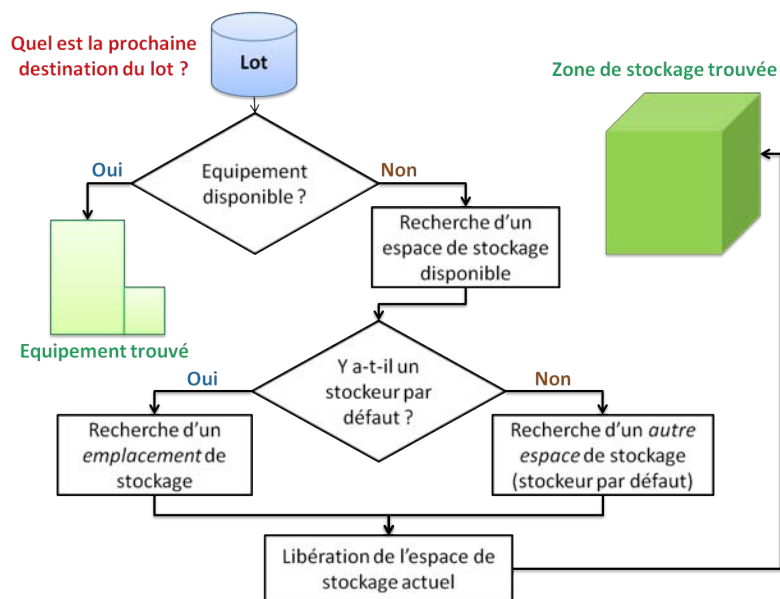


FIGURE 5.9 – Organigramme de la procédure « where next ».

Pour rappel, il y a deux types d'espace de stockage dans notre modèle. Les stockeurs (gros espace de stockage) et les stockeurs par défaut (regroupement d'emplacement de stockage unitaire). Chaque équipement est associé à un espace de stockage. Si cet espace de stockage est plein, il est remplacé par une liste d'espaces alternatifs de stockage. Du fait des temps d'accès plus courts, les stockeurs par défaut sont préférés aux stockeurs, en particulier pour les lots de production. Un résumé des différentes sous-procédures intervenant dans la procédure « where next » est donné dans le tableau 5.15.

Le lot recherche si un emplacement de stockage est disponible dans le stockeur par défaut associé à l'équipement (procédure « FindDefSto »). S'il y a une place disponible, le lot récupère les informations nécessaires à l'aide de la procédure « Sear-

Procédures	Objectifs
DefSto	gestion globale du stockage
SearchOHB	gestion des stockeurs par défaut (et alternatifs)
FindDefSto	recherche d'un emplacement de stockage
FindAlternates	recherche d'espace alternatif de stockage
ReleaseOHB	libération de l'espace de stockage d'un emplacement unitaire

TABLE 5.15 – Les principales sous-procédures du « where next ».

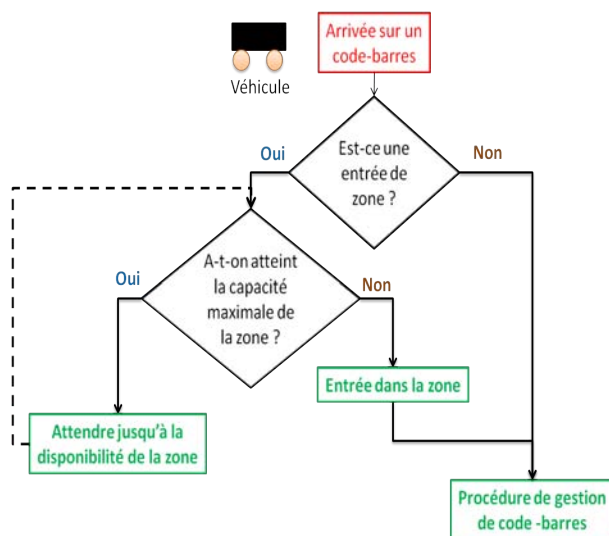
chOHB ». Sinon, il n'y a pas de d'emplacement disponible, et le lot étend sa recherche parmi la liste des espaces alternatifs de stockage à l'aide de la procédure « FindAlternates ».

Lorsqu'un lot libère un emplacement de stockage unitaire, il informe le système que cet emplacement est maintenant disponible à l'aide de la procédure « ReleaseOHB ». Soulignons que la mise à jour de variables entre deux instants très courts. C'est une problématique majeure dans le cas de la simulation à événements discrets. En effet, il y a risque d'affectation d'un lot à un espace de stockage alors que celui-ci est déjà attribué. Ce problème est réglé dans la modélisation en introduisant un temps d'attente d'une seconde avant et après chaque mise à jour de d'espace de stockage.

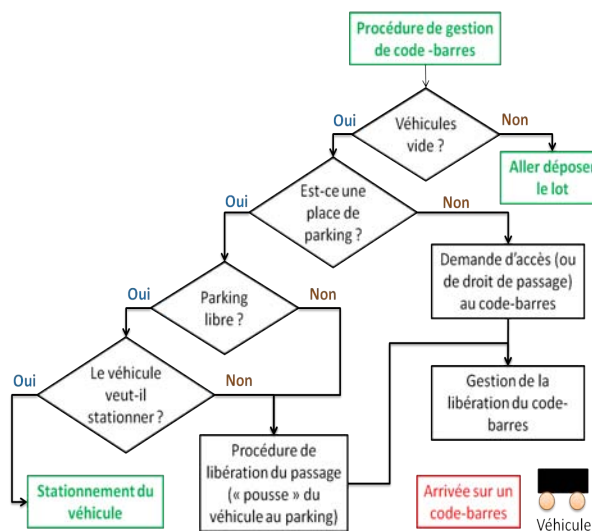
5.3.3 Procédure « vehicle management »

C'est la procédure qui gère le fonctionnement des véhicules. Elle est décrite par la figure 5.10.

Pour se déplacer, les véhicules émettent une succession de requêtes d'autorisation de passage sur les codes-barres situés sur les rails. Un trajet libre est donc un trajet ayant obtenu toutes les autorisations de passage des codes-barres de ce trajet. Si un code-barres change d'état (libre ou occupé) durant le déplacement d'un véhicule, celui-ci doit attendre que le code-barres se libère avant de pouvoir continuer sa route. Le véhicule se déplace donc de code-barres à code-barres. Les sous-procédures entrant dans la gestion du fonctionnement des véhicules sont regroupées dans le tableau 5.16. La gestion des distances de freinage et d'arrêt, de modification des



(a) Etape 1.



(b) Etape 2.

FIGURE 5.10 – Organigramme de la procédure « vehicle management ».

véhicules, etc. est prise en compte et gérée par la sous procédure « StnReleaseFunc ». Il peut arriver qu'un code-barres ne réponde pas à un véhicule qui émet une requête de passage. Ce cas de figure est corrigé par la procédure « StnNotClaimedFun » 5.16.

Sous procédures	Objectifs
ArrivingAtStation	gère l'évènement : arrivée sur un code-barres
StnReleaseFunc	gestion de la libération des codes-barres
StnNotClaimedFun	gestion d'un code-barres qui ne répond pas
ParkOkFunc	gestion des places de parkings
Bump	gestion de la pousse des véhicules en stationnement
OHTC	différence entre LWM ⁷ et Nb.Veh. ⁸ de la zone
TriggerManagement	déclenchement de la « pousse » des véhicules
VehicleInitialization	initialisation des attributs des véhicules

TABLE 5.16 – Les principales sous-procédures de gestion des véhicules.

Les codes-barres sont également un moyen de gestion événementielle (dépôt, prise de lot, animation 3D, recueil d'informations, etc.) par l'intermédiaire de procédure comme par exemple « ArrivingAtStation ». Ils permettent également, l'attribution ou le refus d'une place de parking (procédure « ParkOkFunc »). Soulignons que dans le cas de notre modèle, cette attribution s'effectue en fonction de paramètres bien définis tels que : le Low Water Mark, le High Water Mark et la disponibilité des parkings dans la zone. C'est la procédure « OHTC » qui est en charge de déterminer l'état, en manque de véhicules ou pas, d'une zone. Par exemple la différence entre le *Low Water Mark* et le nombre de véhicules dans la zone.

Les places de parking sont choisies de manière à gêner le moins possible le déplacement des autres véhicules. Mais comme les parkings sont situés sur les rails et qu'il n'y pas pas de voie de contournement, alors, nécessairement à un moment ou à un autre ils se retrouvent sur le trajet d'un véhicule. Il faut donc anticiper cette rencontre entre véhicules.

Lorsqu'un véhicule V_1 rencontre un véhicule V_2 en stationnement sur un parking, le véhicule V_1 a la possibilité de « pousser » le véhicule V_2 en stationnement, afin que ce dernier libère le passage. Cette possibilité de **pousser les véhicules est très importante pour les performances des systèmes de transport**. Elle permet d'éviter par exemple les détours (transport à vide) qui font perdre de la capacité en transport. La gestion de la « pousse » des véhicules est gérée par la procédure « Bump ». Pour le lecteur désireux d'en savoir plus sur la gestion des véhicules aux

repos et du blocage des véhicules, je recommande la lecture de [41], [46] et [47].

5.3.4 Procédure « Equipment Management »

C'est la procédure de gestion des équipements et des stockeurs. Elle intervient avant le transport du lot vers l'équipement ou le stockeur, et après son dépôt par le système de transport.

Pour rappel, la procédure de déchargement *where next* a pour objectif d'amener le lot au plus près des équipements de l'étape suivante. S'il y a un équipement disponible, le véhicule dépose directement le lot sur le port libre de l'équipement (*tool to tool delivery*). Sinon, plusieurs procédures (tableau 5.17) sont mises en oeuvre. Nous allons les décrire dans ce qui suit.

Sous procédures	Objectifs
P_ProcEqtx	traitement des lots
FindLPEqptx	recherche d'un port de chargement libre
P_WaitEqptx	gestion des lots de la liste des tâches à faire
P_ProcessEqptx	traitement des lots sortant de la TTL ⁹
HSTOMx	gestion des équipements hors rails

TABLE 5.17 – Les principales sous-procédures de gestion des équipements.

Rappelons que le système de Crolle 300mm fonctionne en mode hybride, c'est-à-dire que ce ne sont pas tous les équipements qui sont sous les rails. Donc ils ne sont pas tous accessibles par le système de transport automatique. C'est la première chose que vérifiera le système pour tous les équipements. *Nous nous plaçons donc volontairement dans le cas où l'équipement n'est pas disponible dans la suite.*

Situation où l'équipement n'est pas sous les rails. C'est le cas par exemple de l'atelier de traitement thermique qui n'est pas accessible par le système de transport automatisé. Il faut donc une intervention humaine. On se base sur l'observation du cas réelle. Elle consiste à utiliser le stockeur le plus proche de la destination, comme lieu d'échange entre les véhicules et les opérateurs. On réduit ainsi, la distance à parcourir par l'opérateur, qui est moins rapide que les véhicules. C'est la procédure

« HSTOMx » qui est en charge de gérer ce cas particulier. Elle gère la prise, le transport, le dépôt des lots par l'opérateur, ainsi que le traitement des lots par l'équipement.

Situation où l'équipement est sous les rails. On vérifie d'abord que la liste des tâches à faire est vide ou pas (procédure « P_ProcEqtx »).

- *Si elle est vide*, on vérifie que le port de chargement de l'équipement est libre ou pas (procédure « FindLPEqptx »).
- *Si le port de chargement est libre*, alors il y a transport et traitement du lot par l'équipement.
- *Si le port de chargement n'est pas libre*, alors le lot est mis en attente dans la liste de tâches à faire par l'équipement. Lorsqu'un équipement devient disponible alors le lot ayant le plus attendu dans la liste est choisi pour être traité.
- *Si la liste des tâches à faire n'est pas vide*, alors le lot est mis en attente dans la liste de tâches à faire par l'équipement. Lorsqu'un équipement devient disponible alors le lot ayant le plus attendu dans la liste est choisi pour être traité.

Les procédures en charge de la gestion des lots dans la liste des tâches à faire sont : « P_WaitEqptx » pour les entrées et « P_ProcessEqptx » pour les sorties des lots de la liste. Soulignons que ces deux procédures peuvent être améliorées par introduction de règles de gestion plus complexes. La figure 5.11 illustre le déroulement de la procédure « equipment management ».

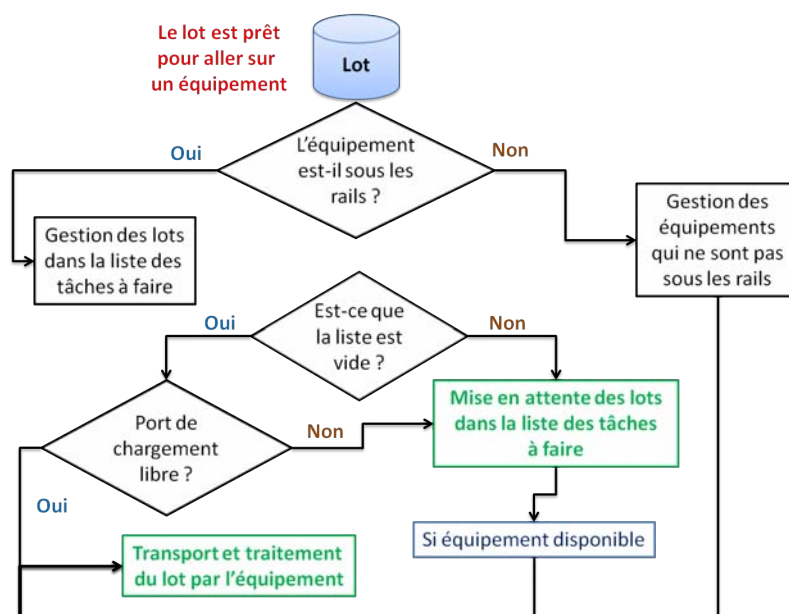


FIGURE 5.11 – Organigramme de la procédure « Equipment Management ».

5.4 Validation de la simulation

« À la source de toute erreur dont on accuse l'ordinateur, on trouvera au moins 2 erreurs humaines, dont celle de rejeter la faute sur l'ordinateur. »

Première loi de Gilb de la non-fiabilité.

5.4.1 Définitions et critères de validation

Définitions

Si un modèle n'approxime pas « suffisamment » le système actuel, toute conclusion sur le modèle est entachée d'erreurs et peut conduire à de coûteuses décisions. La validation devrait et doit être effectuée pour tous les modèles, que le modèle existe sous une forme ou qu'il soit à construire dans le futur. Il faut donc réfléchir et élaborer une procédure de validation lors de la conception du modèle (et même avant).

La validation est le procédé permettant de déterminer si un modèle de simulation est une représentation précise du système, *pour l'objectif particulier de l'étude considérée* [58]. Un modèle de simulation pour un système complexe ne peut être qu'une *approximation* du système réel, quelque soit le temps et l'argent dépensés pour sa construction.

Un modèle de simulation ainsi que ses résultats sont dits **crédibles**, si les décideurs et autres personnes clés du projet les acceptent comme étant *corrects*. Il faut noter cependant qu'un modèle crédible n'est pas nécessairement valide et vice versa. D'après [58], les éléments suivants aident à l'établissement de la crédibilité d'un modèle :

- la compréhension et l'acceptation des décideurs avec les hypothèses du modèle,
- la démonstration que le modèle a été validé et vérifié (exemple, absence de « bugs »),
- l'implication et le degré d'appartenance au projet des décideurs,
- la réputation du développeur,
- une animation convaincante.

Au vu de ces éléments, nous pouvons affirmer que le modèle présenté dans cette thèse est crédible. En effet, le modèle est compris et accepté aussi bien par l'encadrant industriel que par les encadrants scientifiques, tous très impliqués. Plusieurs discussions avec l'équipe AMHS, à l'aide de l'animation 3D, ont permis d'ajuster et d'affiner le comportement du modèle, jusqu'à atteindre un niveau d'animation convaincant. Les principales incohérences telles que la congestion, les stockeurs qui se remplissent trop vite, les équipements qui sont sur-utilisés, etc. ont été corrigées. Rappelons cependant, que le système réel est toujours en pleine extension, donc en modification, au moment de la rédaction de ce manuscrit .

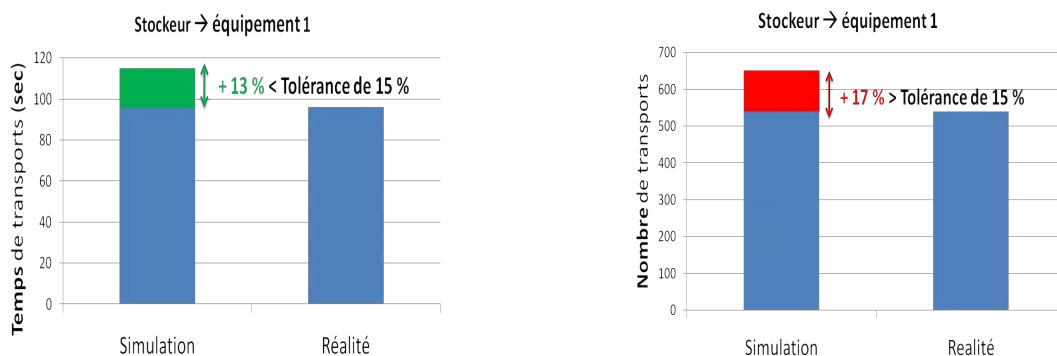
Critères de validation

Nous avons décidé avec le manager industriel, d'*un seuil de validation du modèle fixé à $\pm 15\%$ autant au niveau global que local*. Ce seuil de 15% traduit la prise en compte du mode hybride de Crolles 300, c'est-à-dire la cohabitation homme-machine,

qui génère des variabilités¹⁰ importantes. Par exemple, il peut y avoir inadéquation sur la priorité du lot, entre un opérateur et l'ordinateur central, ce qui aura pour conséquence un retard sur le choix et le transport du lot ou des autres lots.

Pour rappel, les auteurs dans [96] recommandent de ne pas utiliser un seul indicateur pour mesurer les performances d'une unité de fabrication, car il serait nécessairement insuffisant. Nous pouvons confirmer ces recommandations aux travers de différents cas que nous avons rencontré lors de nos différentes étapes de validation. Nous présentons ici le cas de deux indicateurs utilisés pour la validation : le temps de livraison sur les équipements et le nombre total de transports dans une zone. L'objectif est de comparé les résultats obtenus par simulation avec ceux de Crolles 300.

Cas 1 : *le temps de transport respecte le seuil mais pas le nombre de transport.* La figure 5.12 illustre le cas où l'indicateur temps de transports est dans les spécifications $13\% < 15\%$ en valeur absolue, mais l'indicateur nombre de transports dépasse les spécifications, $17\% > 15\%$ en valeur absolue.



(a) Comparaison du temps de transport.

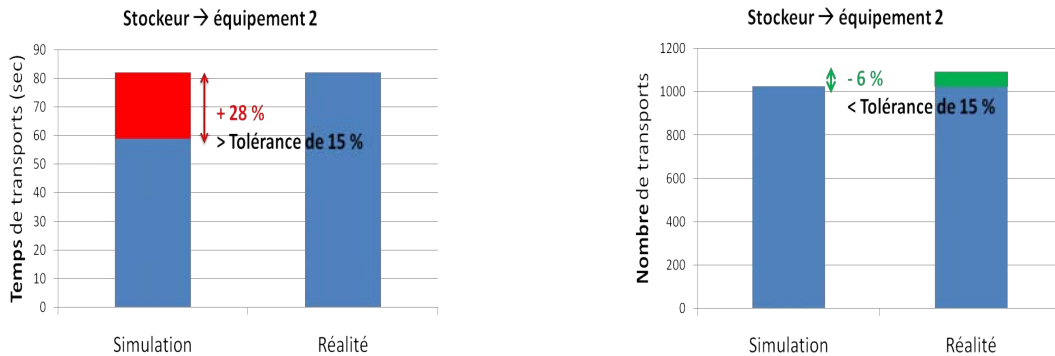
(b) Comparaison du nombre de transport.

FIGURE 5.12 – Illustration du danger d'utiliser un seul indicateur (1/2)

Cas 2 : *le nombre de transport respecte le seuil mais le temps de transport.* La figure 5.13 illustre la situation inverse, c'est-à-dire que l'indicateur nombre de

10. l'humain étant incapable de répétabilité parfaite.

transports respecte les spécifications, $6\% < 15\%$ en valeur absolue, alors que le temps de livraison est hors spécifications, $28\% > 15\%$ en valeur absolue.



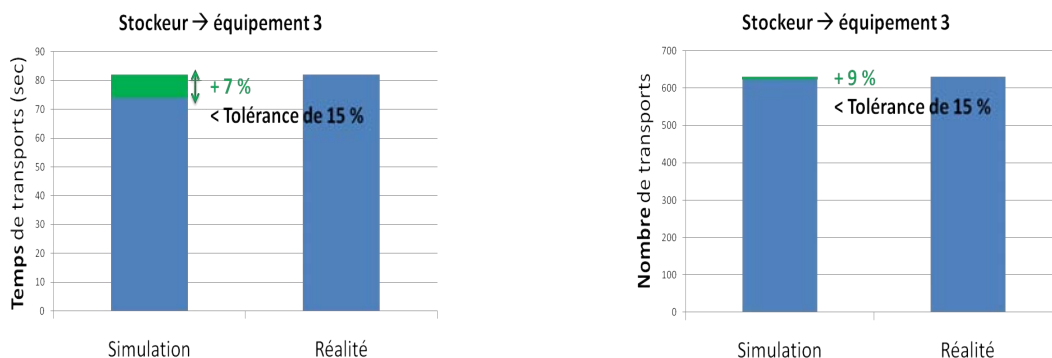
(a) Comparaison du temps de transports.

(b) Comparaison du nombre de transports.

FIGURE 5.13 – Illustration du danger d'utiliser un seul indicateur (1/2).

Cette étude justifie les recommandations émises par [96] et liées au risque d'utiliser un seul indicateur. Nous ajoutons à cette recommandation, qu'il faut pour toute simulation fidèle de l'unité de fabrication, effectuer une validation locale (sur un équipement ou un groupe d'équipements par exemple) et une validation globale du système.

Cas 3 : le nombre de transports et le temps de transport respectent le seuil. La figure 5.14 illustre le cas où tous les indicateurs locaux respectent les spécifications : $7\% < 15\%$ pour le temps de transport et $9\% < 15\%$ pour le nombre de transport. Cependant le temps de cycle des lots, qui est un indicateur global, présente un écart avec le système réel de près de $31\% > 15\%$. Cet exemple montre que ce n'est pas parce que les indicateurs locaux respectent les spécifications que les indicateurs globaux le respectent aussi et inversement. La figure 5.15



(a) Comparaison du temps de transports.

(b) Comparaison du nombre de transports.

FIGURE 5.14 – Illustration du danger d'utiliser un seul indicateur (2/2).

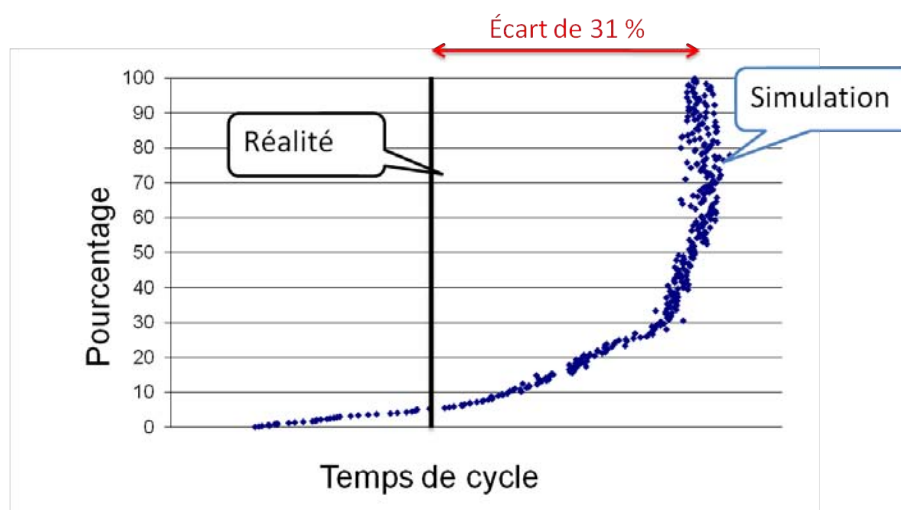


FIGURE 5.15 – Comparaison des temps moyen de cycle.

5.4.2 Validation du modèle

Taille du modèle : le modèle obtenu se compose de **1651** variables, **345** procédures, **151** fonctions, de **528** sous-routines¹¹ et d'environ **45000** lignes de codes. A

11. les sous-routine étant des fonctions ne renvoyant pas de valeur. Elles ne font que des calculs

cela, il faut ajouter à la partie construction du modèle, soit **600** rails tracés, **857** jonctions de rails, **837** codes-barres et **161** fichiers externes.

Deux critères sont utilisés pour la validation du modèle de simulation : la validation par indicateurs globaux et locaux, indispensables pour avoir confiance dans le modèle comme nous l'avons vue précédemment.

Indicateurs locaux. Le temps de livraison représente un temps de livraison pour 80% des lots. Le niveau de productivité représente la moyenne des lots traités par tous les équipements entre la simulation et l'unité de fabrication réelle. La différence respecte la limite des spécifications ce qui est illustré par la figure 5.16.

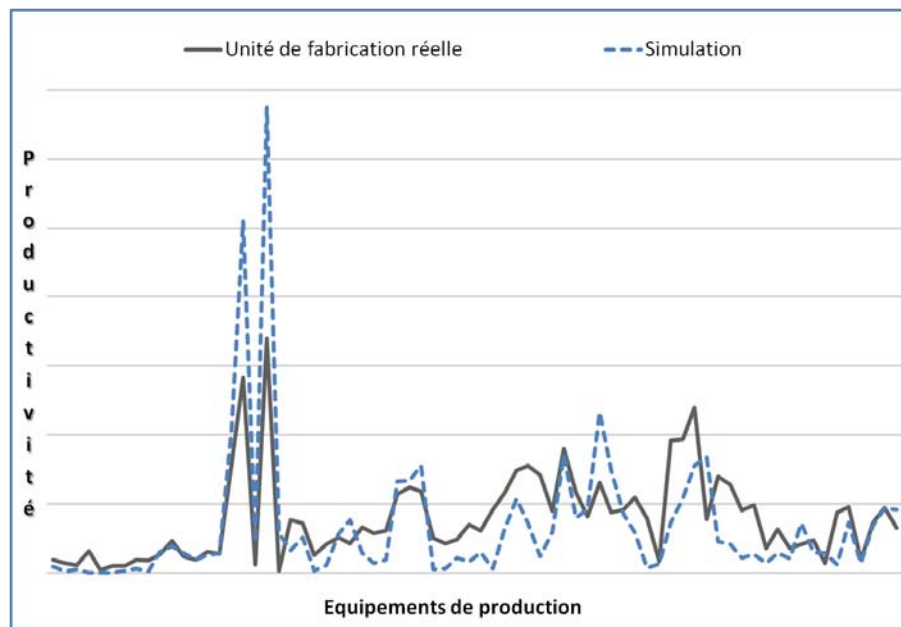


FIGURE 5.16 – Validation : productivité des équipements.

Les valeurs numériques sont résumées dans le tableau 5.18. Pour des raisons de confidentialité, nous ne pouvons pas donner les valeurs réelles.

Indicateurs globaux utilisés pour la validation du modèle. *Le modèle a été validé par l'entreprise.* Le seuil d'erreurs toléré est de **15%** entre simulation et

Indicateurs locaux	Système réel	Simulation	Différences en %
Temps de livraison	K	K*0,90	10 %
Niveau de productivité	X	X*1,15	15 %

TABLE 5.18 – Indicateurs locaux utilisés pour la validation du modèle.

réalité. Ce seuil traduit le mode hybride de l'unité de Crolles 300 de (coexistence homme-machine). Le tableau 5.19 représente les résultats utilisés pour la validation globale. Ces résultats montrent que l'on respecte les spécificités établies. On est à 11% d'écart avec le temps moyen de cycle réel, avec un écart type de 28%.

Indicateurs globaux	Unité de fabrication réelle	Simulation	Différences en %
Temps de cycle	Y	Y*1,11	11 %
Écart type	21%	28%	-

TABLE 5.19 – Indicateurs globaux utilisés pour la validation du modèle.

Avec cette validation, nous avons dans un premier modèle de base, qui a été affiné. Ce modèle a permis au manager et à l'équipe AMHS d'avoir un niveau de confiance suffisant au modèle mais surtout, de pouvoir connaître avec précision le niveau d'erreur maximal du modèle. Bien entendu ce niveau d'erreur est appelé à être réduit dans le futur.

5.4.3 Données aléatoires

Le tableau 5.20 représente les données aléatoires utilisées dans la simulation.

Chaque équipement de production possède son propre temps de traitement. Le temps réel de traitement est connu et est introduit dans la simulation. Mais à cause des différents aléas qui peuvent survenir en situation réelle, nous avons choisi de le faire correspondre à la moyenne d'une loi de probabilité gaussienne, avec une variabilité que nous allons préciser. Notons que d'après [6], la loi gaussienne¹² est bien adaptée pour les temps de traitement des équipements.

12. typiquement utilisée pour modéliser les événements ayant une variabilité limitée.

Un lot dans un FOUP¹³ peut être constitué d'au maximum 25 wafers, mais dans la pratique un lot en contient généralement moins car, durant le processus de fabrication les wafers peuvent être endommagés, on peut être amené à constituer des lots pour des traitements bien particuliers à une étape donnée, etc.

	Loi de probabilité	Variabilité
Temps de traitement des lots	Gaussienne	20%
Temps de déplacement des opérateurs	Gaussienne	20%

TABLE 5.20 – Tableau des données aléatoires.

Mais **avec moins de wafers, le temps de traitement est plus court**. Cette variation correspond à l'écart type de la loi gaussienne utilisée précédemment. Sa valeur est de 20% par rapport à la moyenne et est déterminée en fonction d'un historique du nombre de wafers dans un lot fourni par l'entreprise. Pour notre simulation, nous avons estimé qu'un lot contient 23 wafers.

Nous avons fait le choix d'un temps de marche opérateur d'environ 1 minute avec une variabilité de 20% afin de refléter les aléas du comportement humain.

Notre simulation dure 360 jours, répartie en 180 jours pour suffisamment charger l'unité de fabrication afin que les équipements de production prennent une cadence la plus régulière possible pour le traitement des lots. Les 180 jours restant correspondent au temps de notre étude, c'est-à-dire le temps pendant lequel on recueille des données statistiques.

Ce « long » temps de simulation permet de nous affranchir des répliques, nécessaires à la validation des modèles de simulation avec des paramètres aléatoires. En effet, plus le temps de simulation est long et plus l'aléa généré par le système sera significatif comme dans les répliques. De plus, nous n'avons aucune variabilité au niveau du nombre de lots lancés (données d'entrée) car il est quasiment fixe. Cela est dû au mode de fonctionnement entre STMicroelectronics et ses clients.

13. Front Opening Unified Pod

5.5 Plan d'expérience

« *L'expérience est une observation provoquée dans le but de faire naître une idée.* »
Emile Auguste Chartier.

5.5.1 Introduction

Le plan d'expérience peut se définir comme une suite ordonnée d'essais élémentaires d'une expérimentation qui accompagne une recherche scientifique ou des études industrielles. *Elle permet de faire l'expérience de l'expérience* et réduit considérablement la complexité du problème abordé, ce qui signifie avoir le maximum de renseignements avec le minimum d'expériences.

En outre, le plan d'expérience permet de répondre aux questions suivantes :

- Quels sont les facteurs les plus influents ?
- Existe-t-il des interactions entre les facteurs (corrélations) ?
- Comment minimiser le nombre de points de mesure du processus afin d'obtenir le maximum d'informations ?
- Existe-t-il des biais entre les résultats des mesures ?

Pour tenter de répondre à ces questions, nous présenterons tout d'abord la complexité du problème dans la section 5.5.2 et nous justifierons notre approche pour réduire la combinatoire des tests. A l'aide de mesures réelles de l'unité de fabrication, nous étudierons ensuite deux indicateurs : le temps d'assignation (section 5.5.3) et le temps moyen mis pour changer de lot d'un équipement (section 5.5.4). Cette étude permettra de mieux comprendre les principales sources de variabilité des systèmes de transport automatisé. Nous comparerons ensuite deux politiques de gestion des véhicules. Celle utilisée à Crolles 300 que nous nommerons « service minimum », et celle qui est la plus utilisée dans la littérature que nommerons « service classique » (section 5.5.5). Nous terminerons ce chapitre en répondant aux questions posées précédemment dans la définition du plan d'expérience.

5.5.2 Complexité du nombre de tests

L'unité de fabrication de Crolles 300 est découpée en 10 zones au moment de l'écriture de ce manuscrit comme le montre la figure 5.17. Pour rappel, chaque zone possède deux valeurs : une pour le *Low Water Mark* (LWM) et l'autre pour le *High Water Mark* (HWM).

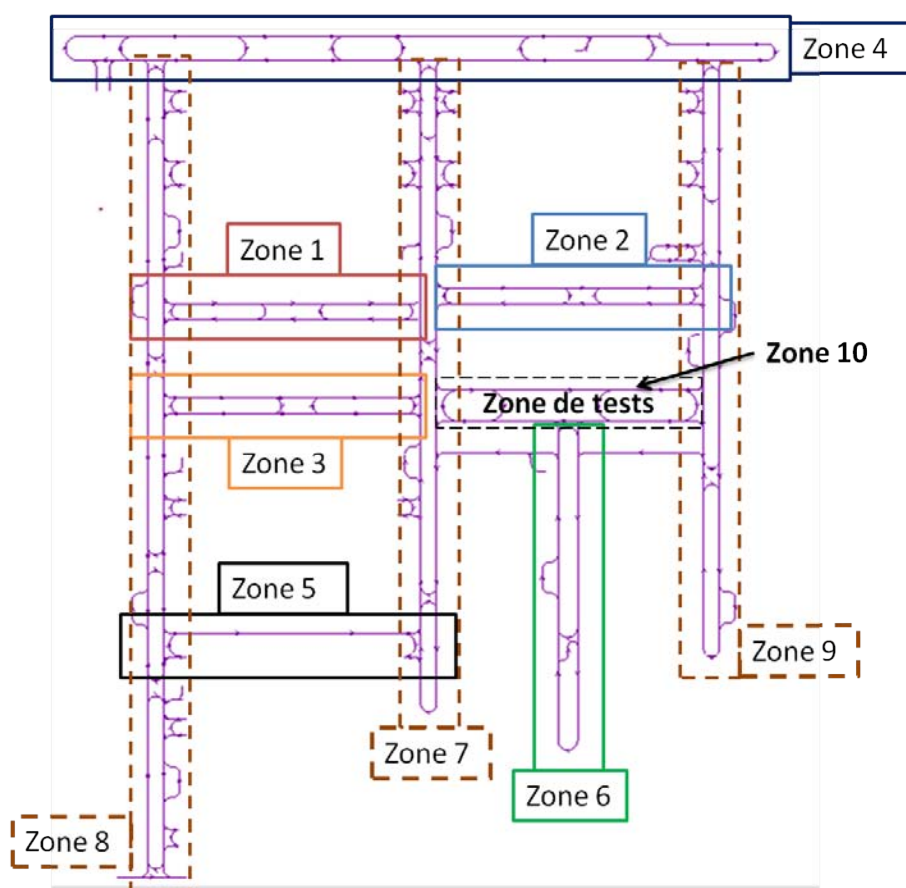


FIGURE 5.17 – Découpage de l'unité de fabrication de Crolles 300 en zones.

D'un point de vue théorique, le *High Water Mark* prend ses valeurs dans l'intervalle $[0, +\infty]$ mais en pratique il est limité. La première limitation est la longueur des rails de la zone. Elle fixe le nombre de véhicules pouvant rester dans la zone en fonction de la distance de sécurité entre les véhicules et du risque de congestion. La seconde limitation est la capacité électrique maximale que peut supporter une zone.

À Crolles 300 elle correspond à au plus 12 véhicules par zone. La valeur maximale de l'intervalle des valeurs du *High Water Mark* se réduit à $[0, 12]$ par zone.

Une autre information permettant de limiter les valeurs du *High Water Mark* est le nombre total de véhicules qui est de 26. Il correspond au nombre de véhicules au moment de la conception du modèle. D'après les historiques des valeurs du *High Water Mark* mis en place par les ingénieurs de Crolles 300, on note que la valeur maximale atteint rarement 5. Une explication est qu'il ne faut pas conserver trop de véhicules dans une zone. Nous avons donc limité notre intervalle de valeurs possibles du LWM à $[0, 4]$.

Le *Low Water Mark* variera aussi entre 1 et 4 par la relation $LWM \leq HWM$. Le cas $LWM = 0$ n'est pas traité ici, et ne semble pas très pertinent de notre point de vue à cause des caractéristiques de la zone étudiée. Elle représente environ 1/4 du trafic total de l'unité de fabrication de Crolles 300, avec 4 à 5 déposes/enlèvements de lots sur équipements par minute. Demander à une telle zone de ne garder ou de n'appeler aucun véhicule ne nous semble pas très pertinent.

Un test représente une instance lancée avec des valeurs de *Low Water Mark* et de *High Water Mark* fixe. La combinatoire du nombre de tests en fonction des valeurs du *Low Water Mark* et du *High Water Mark* est élevée. Par exemple, si on fixe la valeur maximale du *High Water Mark* à 4, il y a 8100 possibilités pour 10 zones. Sachant qu'il faut 6 heures par test, cela fait 48600 heures, soit environ **6 ans de simulation**.

Pour réduire la combinatoire, nous nous sommes focalisés sur une zone clé, celle de la métrologie, juste après l'étape de lithographie. Ce choix s'explique par le fait que cette zone est la plus sollicitée à Crolles 300, 1/4 du trafic total sur seulement 55 m de rails installés. Elle est également la zone qui échange le plus avec les autres zones, puisqu'elle correspond à une étape de mesure. On est ainsi assuré que toute variation survenant dans cette zone aura un impact *significatif* sur l'unité de fabrication de Crolles 300. On travaille ainsi sans perte de généralités. *Dans la suite, tous les tests seront effectués en ne modifiant que les paramètres de cette zone.* On réduit ainsi le nombre de tests de 95%, qui passe de 8100 à 4 et le temps total de simulation de 6 ans à 24 heures. *Tous les autres paramètres resteront fixes* et ont été choisis à l'aide d'un prototype de pré-calcul du *Low Water Mark*.

Le tableau 5.21 résume les valeurs du LWM et du HWM de toutes les zones sauf la zone d'étude puisque les valeurs varieront. La zone 5 correspond à une zone où il n'est pas possible de faire attendre des véhicules sans gêner le passage d'autres véhicules. Elle est en extension au moment de l'écriture de ce manuscrit.

	LWM	HWM
zone 1	1	2
zone 2	2	2
zone 3	2	2
zone 4	2	4
zone 5	0	0
zone 6	3	5
zone 7	4	4
zone 8	3	4
zone 9	2	3

TABLE 5.21 – Valeurs du LWM et du HWM des 9 zones de Crolles 300.

Conclusion : dans toute la suite la valeur maximale du *High Water Mark* est fixée à 4 et l'on fera varier la valeur du *Low Water Mark*. La zone cible choisie est la zone 10 correspondant à la zone de métrologie.

5.5.3 Etude du temps d'assignation

Prenons comme point de départ temporel l'instant où un véhicule est choisi par le système pour répondre à une requête de transport. Appelons t le temps de transport du lot de son origine vers sa destination et Δt la variation du temps de transport pouvant apparaître sur cette portion de trajet. Par exemple, des variations dues au ralentissement à cause d'un autre véhicule, au dépôt ou à la prise de lots par un autre véhicule, etc. On peut affirmer sans se tromper, sauf cas exceptionnels de congestion, de panne nécessitant un arrêt, etc., que la valeur de Δt ne fluctuera pas beaucoup car le temps calculé ici est du transport pur. Par analogie, le temps du trajet Paris-Berlin en train, à quelques minutes près sensiblement, est toujours le même, sauf cas exceptionnel.

Cependant, si on déplace le référentiel de départ, c'est-à-dire que l'on ajoute au temps de transport pur t le temps entre l'émission de la requête de transport et le moment où un véhicule est choisi par le système, alors la variabilité Δt peut être importante. Cela se comprend de manière intuitive car, par analogie, la durée d'attente d'un taxi par un client est variable. Mais si on regarde de plus près ce temps d'attente, on se rend compte qu'il se décompose en deux parties : (1) le temps pour trouver la ressource (le véhicule) et (2) le temps pour amener la ressource dans la zone concernée. Rappelons que dans le cas particulier de Crolles, une requête de transport ne peut être attribuée à un véhicule que s'il est à l'intérieur de la zone. Ceci explique pourquoi Δt augmente si on prend en compte ce nouveau repère dans le calcul du temps de transport.

Donc le temps d'assignation représente le temps entre le début d'une requête de transport jusqu'au moment de l'affectation de la requête à un véhicule. Il dépend de plusieurs paramètres tels que le nombre total de véhicules, l'encours de l'unité de fabrication, la position des parkings et les politiques de gestion des véhicules.

La figure 5.18 représente les mesures du temps d'assignation effectuées en salle blanche à Crolles 300. A ce stade de l'étude, notre motivation principale est de comprendre le pourquoi du comment de la variabilité du temps d'assignation.

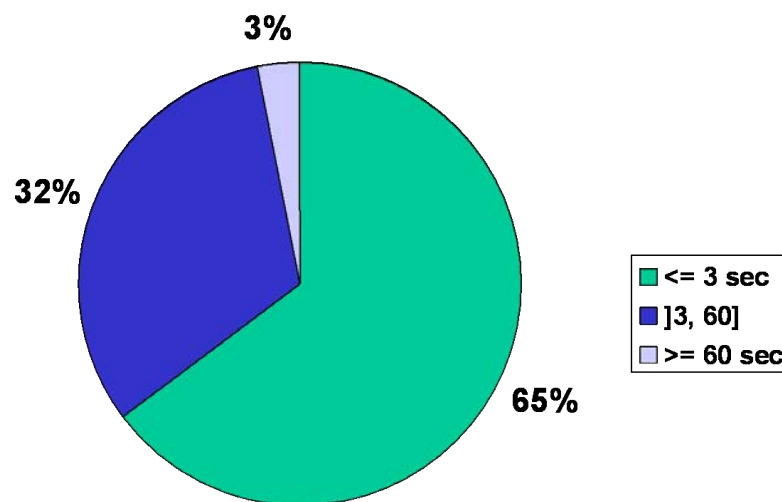


FIGURE 5.18 – Répartition du temps d'assignation à Crolles 300.

Ces mesures montrent que, 65% des lots trouvent un véhicule en moins de 3 secondes, 32% entre 3 et 60 secondes et le reste des lots, soit 3%, trouvent un véhicule en plus de 60 secondes. Donc 97% des lots trouvent un véhicule en moins de 60 secondes. Le problème concerne les 3% des lots restants, qui eux, mettent plus de 60 secondes pour trouver un véhicule. Si on se ramène à l'échelle du temps de changement d'un lot sur le port de chargement d'un équipement, le temps des 3% des lots restant impactent considérablement l'enchaînement des tâches de manière générale et le temps de livraison en particulier.

Pour un temps de changement d'un lot sur un équipement de l'ordre de 300 secondes par exemple, 60 secondes représentent 20% de ce temps et 120 secondes 40%. Ainsi, à lui seul, le temps d'assignation peut représenter plus de la moitié du temps de changement d'un lot sur un équipement, ce qui justifie notre étude.

Remarque : le temps d'assignation est difficile à mesurer en pratique (c'est un temps informatique qui peut être très bref) et à estimer en simulation car non pris en compte dans la plupart des logiciels commerciaux. On peut parfois l'estimer en retranchant du temps de transport le temps du début de la requête.

Conclusion : *le temps d'assignation est un indicateur très variable qui permet de mesurer en local le niveau d'adéquation entre ressources et besoins.*

5.5.4 Étude du Carrier Exchange Time CET

Un des objectifs du système de transport est de contribuer à maximiser le taux d'utilisation des équipements, en apportant toujours le bon lot au bon endroit et *au bon moment*. En effet, l'enchaînement des tâches sur un équipement de production permet, d'une part, d'avoir un temps de cycle des lots le plus court (accroissement du rendement) et, d'autre part, d'amortir le prix élevé des équipements de la salle blanche (retour sur investissement).

Donc, en théorie, le système de transport *doit au moins pouvoir garantir* un temps de changement de lot réduisant au maximum le temps d'attente à vide du port de chargement. Dans la pratique, puisque les équipements ont des temps différents de traitement des lots, ils doivent être définis selon les besoins propres de chaque

équipement.

Ainsi, l'indicateur permettant de mesurer la cadence de passage des lots sur les équipements de production est le temps mis pour changer un lot sur le port de chargement d'un équipement appelé *Carrier Exchange Time* ou CET. Sur la base de l'analyse des temps de traitement des lots par équipement, on peut regrouper les équipements en deux catégories principales : les équipements de haute priorité nécessitant un temps court de changement de lots, et les équipements standards qui sont tous les autres équipements.

La figure 5.19 représente une décomposition du temps mis pour changer de lot sur le port de chargement ou *CET*, dans le cas où tout se passe bien (*happy flow*). Le stockeur considéré est un emplacement unitaire situé au plafond (OHB sur la figure). Cette décomposition montre que le temps mis pour changer un lot sur le port de

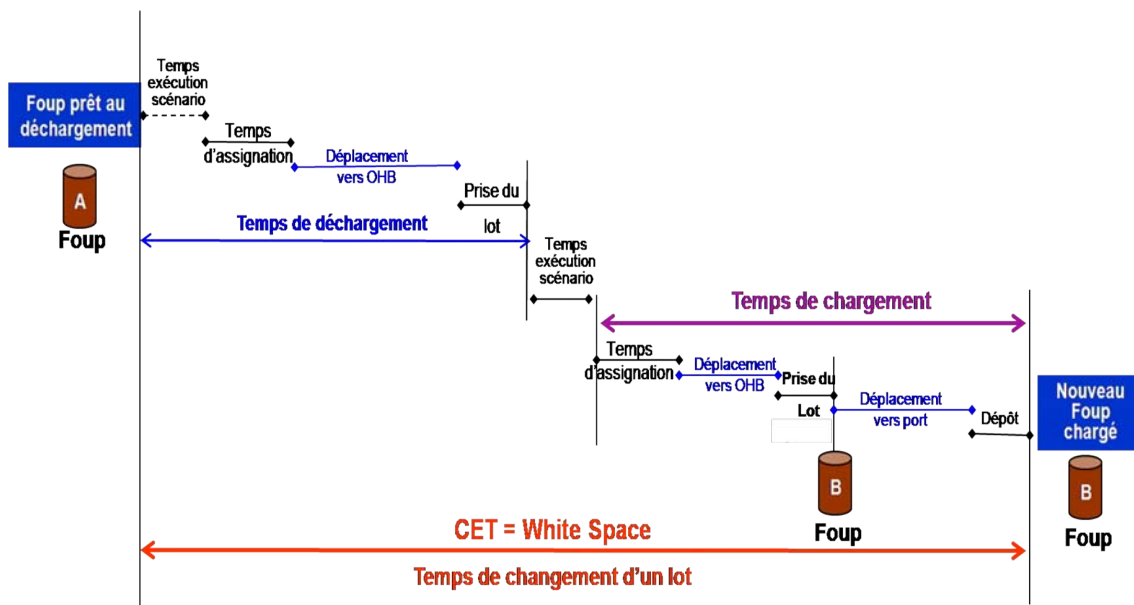


FIGURE 5.19 – Décomposition [57] du temps mis pour changer de lot sur un équipement (CET).

chargement d'un équipement donné se compose de **temps fixes**, tels que le temps d'exécution des scénarios informatiques, les temps de prise et de dépôt des lots et de **temps variables**, tels que les temps de déplacement des véhicules et d'assignation

d'une requête de transport à un véhicule.

La figure 5.20 illustre la variation du temps mis pour changer de lot sur le port de chargement d'un équipement, en fonction du *Low Water Mark* pour deux équipements de la zone étudiée. La valeur du *High Water Mark*¹⁴ est fixée à 4. En dessous de l'axe des abscisses, nous avons le nombre de transports correspondant au nombre de mesures effectuées pour ce test et sur une période de 90 jours. Le temps pour atteindre le régime stationnaire est de 180 jours.

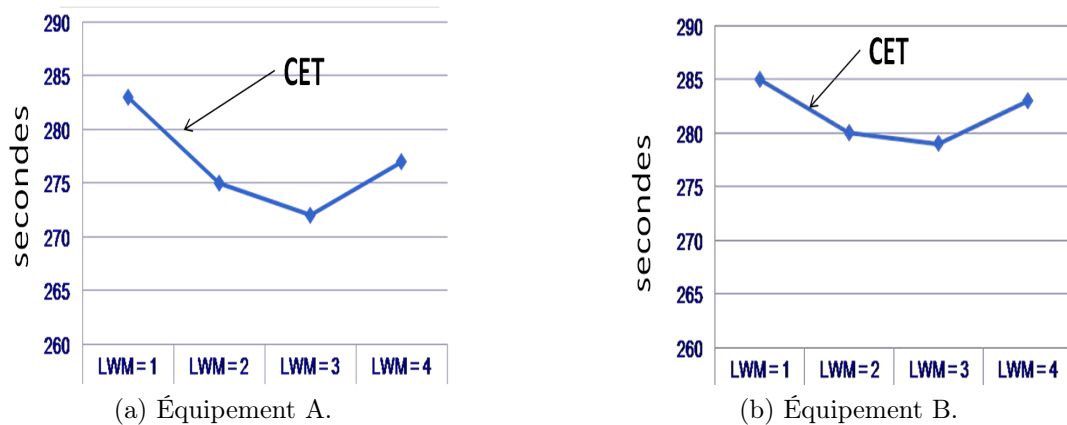


FIGURE 5.20 – Temps moyen de changement d'un lot sur un équipement : variation en cloche.

Les résultats montrent que le temps mis pour changer un lot sur le port de chargement d'un équipement, est maximal pour la valeur $LWM = 1$ et minimale pour les valeurs $LWM = 3$. Ceci à été observé sur la majorité des équipements. Cette courbe traduit le fait qu'avec peu de véhicules ($LWM = 1$) ou trop de véhicules ($LWM = 4$) dans la zone, le temps mis pour changer de lot sur un équipement se dégrade. *Ce phénomène a été confirmé en situation réelle dans l'unité de fabrication de Crolles 300.*

Donc, pour une zone donnée, le problème qui consiste à déterminer la valeur du LWM qui minimise le temps de changement d'un lot pour un équipement donné,

14. Nombre maximal de véhicule non assignés autorisés à rester dans la zone.

admet une solution unique. Cependant, cette solution peut ne pas convenir à un équipement voisin (à cause de la différence des temps de traitement des lots par les équipements) ou cette solution peut ne pas convenir à d'autres équipements externes à la zone, mais en relation avec l'équipement étudié.

Le transport étant un service pour la production, cet indicateur permet de mesurer la qualité de ce service.

Conclusion : *Le temps mis pour changer de lot sur un équipement permet de mesurer la cadence de l'enchaînement de tâches sur les équipements de production.*

5.5.5 Comparaison entre politique de « service minimum » et « service classique »

Cette section a été publiée dans un chapitre de livre : Progress In Material Handling Research [40].

Une politique de transport qui marche pour une unité de fabrication A ne convient pas nécessairement à une unité de fabrication B. Il faut donc trouver, pour chaque unité de fabrication, la politique de transport correspondante. Le plus souvent, il faut tester plusieurs politiques avant de trouver la « bonne ». Les possibilités sont nombreuses car il faut, d'une part, faire la bonne association entre le routage et le dispatching et, d'autre part, prendre en compte différents paramètres tels que le type de produits fabriqués, l'agencement des rails, etc.

Les motivations de cette section sont doubles : (1) montrer qu'il est possible de réduire considérablement l'espace de recherche des « bonnes politiques », et (2) montrer que la politique de « service minimum » domine une des politiques les plus utilisées dans la littérature.

Le « service minimum » : c'est une politique de dispatching qui consiste à assurer, *presque tout le temps*, la présence d'un nombre prédéfini de véhicules non-assignés à l'intérieur de chaque zone. Ce nombre prédéfini de véhicules s'appelle le *Low Water Mark* (LWM). Cette politique permet ainsi de minimiser le temps d'assignation des requêtes de transport aux véhicules.

Les principaux avantages de cette politique sont, d'une part, la fiabilité et la rapidité avec laquelle les véhicules répondent aux requêtes de transport car presque toujours dans la zone, et d'autre part, la propriété de répétabilité des performances du système de transport, due au repositionnement dans la zone d'un nombre prédéfini de véhicules.

Par la suite, nous dirons qu'une zone est *stable* si le nombre de véhicules non-assignés à une requête de transport à l'intérieur de la zone est supérieur ou égal à la valeur du LWM sur une période donnée. Un des objectifs du service minimum est de garantir cette période de temps la plus longue possible.

Cependant, la difficulté de cette politique réside dans la capacité du système à maintenir les zones stables. La stabilité des zones est gérée à l'aide du contrôleur de véhicules dont la mission est de réduire les différences entre la valeur du LWM et le nombre de véhicules de la zone. Pour ce faire, le contrôleur de véhicules fait appel aux véhicules situés dans les autres zones. Mais cet appel est soumis à une condition : si elle rend instable la zone qui cède le(s) véhicule(s) alors elle est interdite par le système et le contrôleur doit s'adresser à une autre zone. Cette restriction est une condition forte comme nous le verrons par la suite.

Si une requête de transport dont l'origine est externe à la zone est affectée à un véhicule de la zone, le « service minimum » peut perdre sa performance de rapidité et de répétabilité. En effet, les véhicules en moins dans la zone représentent une perte de capacité de transport, donc un allongement potentiel du temps d'assignation. Pour être efficace, l'attribution de requêtes de transport aux véhicules utilisant la politique de « service minimum », ne doit se faire qu'uniquement à l'intérieur de la zone. Cela signifie qu'un lot d'une zone A ne peut directement faire appel à un lot d'une zone B.

L'efficacité d'un « service minimum » se mesure par sa capacité à maintenir un nombre prédéfini de véhicules constamment dans chaque zone de l'unité de fabrication.

Le « service classique » : c'est une politique de dispatching consistant à affecter des véhicules aux différentes zones de l'unité de fabrication, puis à attribuer les requêtes de transport au véhicule le plus proche. Cette politique minimise le transport

à vide des véhicules, car le temps de fonctionnement à vide des véhicules représente une perte de capacité du système en nombre de véhicules.

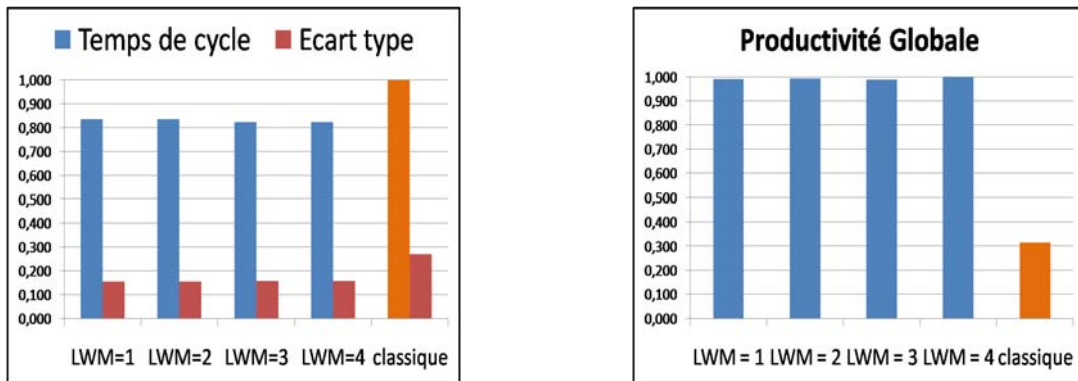
L'avantage majeur de cette politique est que les véhicules sont complètement autonomes pour la recherche de tâches. L'inconvénient majeur est qu'elle ne tient pas compte de la priorité des lots à transporter. Le tableau 5.22 résume les caractéristiques des deux simulation. La seule différence est la politique de dispatching utilisée.

	Classique	Minimum
Temps total simulation	270 jours	270 jours
Période de chargement des lots	180 jours	180 jours
Période de simulation	90 jours	90 jours
Nombre d'équipements	250	250
Nombre de zones	10	10
Nombre de véhicules	26	26
Nombre de parkings	35	35
Nombre de lots en entrée	2898 wafers/semaine	2898 wafers/semaine
Politique de routage	plus court chemin	plus court chemin
Politiques de dispatching	lot le plus proche	LWM

TABLE 5.22 – Caractéristiques de la simulation pour la comparaison entre la politique de service « minimum » et « classique ».

La figure 5.21 représente les temps moyen de cycles pour différentes valeurs du LWM dans une gestion par « service minimum » et par « service classique ». Le temps moyen de cycle est plus long dans la gestion par « service classique » (figure 5.21a). Les wafers mettent donc plus de temps pour sortir de l'unité de fabrication, ce qui diminue la productivité de l'unité de fabrication (figure 5.21b).

Le tableau 5.23 représente, *pour les équipements internes*, le temps mis pour changer de lot sur leur port de chargement, obtenu en comparant « service minimum » et « service classique ». L'indicateur prend sa plus grande valeur (en gras dans le tableau 5.23) dans la gestion par « service classique » et ce, *indépendamment de la position des équipements* (en gras dans le tableau 5.24 *pour les équipements externes* à la zone étudiée).



(a) Comparaison des temps moyens de cycles et des écarts types.

(b) Comparaison de la productivité.

FIGURE 5.21 – Politique de service « minimum » vs. « classique » : comparaison du temps moyen de cycle et de la productivité.

Noms des équipements	Variation du LWM				
	LWM=1	LMW=2	LWM=3	LMW=4	Classique
Équipement 1	0,574	0,550	0,550	0,558	0,891
Équipement 2	0,729	0,705	0,705	0,713	1,000
Équipement 3	0,682	0,620	0,620	0,620	0,915
Équipement 4	0,721	0,659	0,659	0,659	0,783
Équipement 5	0,659	0,597	0,597	0,581	0,837
Équipement 6	0,721	0,698	0,698	0,705	0,822
Équipement 7	0,736	0,721	0,721	0,729	0,860

TABLE 5.23 – Politique de service « minimum » vs. « classique » : temps moyen pour changer de lots [équipement **interne**].

On prend donc plus de temps pour changer un lot sur les ports de chargement des équipements dans la gestion par « service classique ». Ceci s'explique par une présence plus fréquente de véhicules dans une gestion par « service minimum ».

De plus remarquons que, *même pour la plus mauvaise valeur* (LWM=1) de la politique de « service minimum », elle reste meilleure que la politique de « service classique ». Ce résultat apporte une information très importante pour la recherche des politiques de gestion des véhicules. Il fournit une solution de départ de bonne

Noms des équipements	Variation du LWM				Classique
	LWM=1	LMW=2	LWM=3	LMW=4	
Equipement 1	0,678	0,601	0,607	0,552	0,934
Equipement 2	0,738	0,705	0,705	0,842	0,934
Equipement 3	0,743	0,721	0,710	0,727	0,967
Equipement 4	0,732	0,732	0,738	0,727	1,000
Equipement 5	0,738	0,705	0,694	0,716	0,923
Equipement 6	0,492	0,497	0,503	0,497	0,503

TABLE 5.24 – Politique de service « minimum » vs. « classique » : temps moyen pour changer de lots sur un équipement **externes** de la zone étudiée.

qualité pour la recherche de politique de gestion, et permet en industrie de garantir une meilleure productivité, donc un meilleur retour sur investissement.

Conclusion : *l'enchaînement de tâches sur les équipements de production est plus efficace avec une politique de « service minimum ».*

5.5.6 Difficultés de maintenir l'adéquation entre ressources de transport et besoins de production

Pour rappel, le *Low Water Mark* (LWM) est le nombre minimal de véhicules non assignés devant être constamment dans la zone. Si le nombre de véhicules dans la zone est en dessous de la valeur du LWM, le système compense la différence en appelant des véhicules d'une autre zone. Cependant cette compensation n'est possible que *si cet appel respecte le LWM de la zone d'où l'on veut prendre des véhicules*. Les figures 5.22 et 5.23 représentent **les fréquences du nombre de véhicules** dans une zone à fort trafic. En abscisses on a le nombre de véhicules qui varie de 1 à 10. En ordonnées la fréquence du nombre de véhicules. On constate que pour un *Low Water Mark* donné, la fréquence du nombre de véhicules dans une zone est une courbe en cloche. Le maximum de la courbe en cloche est translaté vers la droite lorsque la valeur du LWM augmente. Si la fréquence du nombre de véhicules correspondant à la valeur du LWM est très éloigné du maximum, on peut penser

que l'activité de la zone demande plus de véhicules.

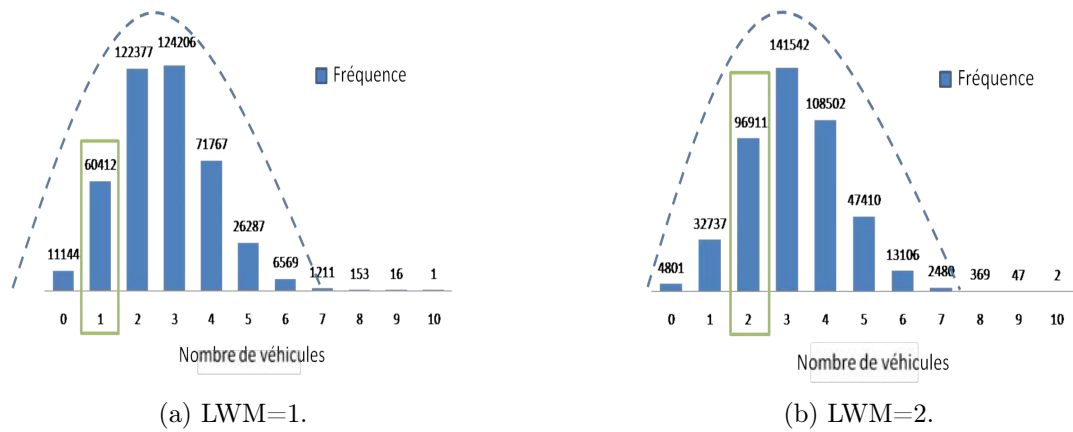


FIGURE 5.22 – Répartition de la fréquence du nombre de véhicules dans la zone étudiée (1/2).

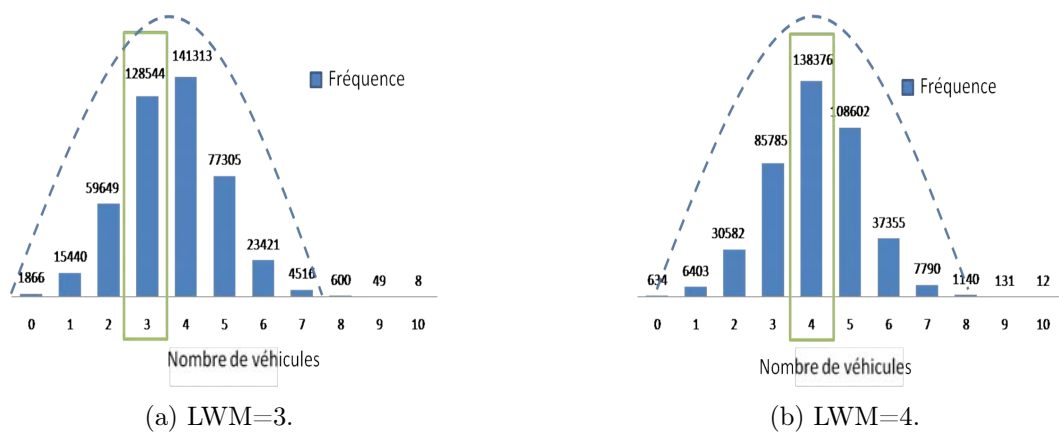


FIGURE 5.23 – Répartition de la fréquence du nombre de véhicules dans la zone étudiée (2/2).

Remarque : le nombre de véhicules devant être un entier, la notation 3^- (tableau 5.25) signifie qu'il y a en moyenne entre 2 et 3 véhicules, mais que l'on est plus proche de 3 que de 2.

	Moyenne	Ecart type	Max	Min	Mode	% Nb. Veh = 0
LWM = 1	3 ⁻	1	10	0	3	2,63%
LWM = 2	3	1	10	0	3	1,07%
LWM = 3	3 ⁺	1	10	0	4	0,41%
LWM = 4	4	1	10	0	4	0,15%

TABLE 5.25 – Répartition de la fréquence du nombre de véhicules dans la zone étudiée en fonction du LWM.

Ces résultats nous permettent d’une part, de valider notre modélisation du fonctionnement de la gestion du *Low Water Mark* et d’autre part, de montrer une limitation de l’approche par « service minimum » : la difficulté du système à maintenir *parfois* au moins un véhicule dans la zone (voir colonne [Nb. Veh = 0] du tableau 5.25).

Cette difficulté a également été constatée en situation réelle dans l’unité de fabrication de Crolles 300.

Conclusion : *la gestion des véhicules par « service minimum » telle qu’elle est implémentée dans l’unité de fabrication de Crolles 300, peut vider une zone complètement de ses véhicules.*

5.5.7 Conclusion sur le plan d’expérience

Nous allons maintenant répondre à toutes les questions que nous avons posé précédemment dans l’introduction de cette question.

Quels sont les facteurs les plus influents ?

Les différents tests montrent que les facteurs les plus influents sont :

- le *Low Water Mark*, parce que c’est l’élément clé du « service minimum ». Il fixe le nombre moyen de ressources que l’on pense avoir besoin afin de répondre aux requêtes de transport des lots de la zone.

- le temps d'assignation des requêtes de transport, parce que c'est un des paramètres les plus aléatoires et qu'il peut influencer de manière très importante le temps d'enchaînement des tâches.

Existe-t-il des interactions entre les facteurs ?

Puisque le temps d'assignation est difficilement mesurable et estimable, il n'est pas facile de prouver les interactions avec les autres facteurs, même si intuitivement on le sent. De notre point de vue, le sujet nécessite un approfondissement.

Comment minimiser le nombre de points de mesure du processus afin d'obtenir le maximum d'informations ?

Parmi les indicateurs que nous venons d'étudier, le temps mis pour remplacer un lot par un autre sur un équipement (*Carrier Exchange Time*), permet de mesurer la cadence de passage des lots sur les équipements de production. La décomposition de cet indicateur (voir figure 5.19) montre qu'il prend en compte plusieurs autres indicateurs tels que temps de livraison, le temps de déplacement à vide, le temps de prise ou de dépôt d'un lot, le temps d'assignation.

Cet indicateur local couplé à d'autres indicateurs globaux comme le temps moyen de cycle et la productivité, permet de minimiser le nombre de points de mesure avec un maximum d'informations car, on surveille à la fois les informations globales et locales de l'unité de fabrication.

Existe-t-il des différences entre les résultats des mesures ?

Les résultats des tests menés à la section 5.5.5 ont montré que le temps de cycle et la productivité étaient inversement croissant, c'est-à-dire que quand l'un augmente l'autre diminue et inversement. Donc, il n'existe clairement pas de relation entre ces deux mesures.

Cependant, ce n'est pas parce que le temps mis pour changer de lot sur un équipement est grand que le temps de cycle sera nécessairement grand, puisque ce

dernier indicateur est local. Mais pour les équipements à forte activité, les informations fournies par cet indicateur ont un impact sur les résultats globaux. Par exemple, si le temps mis pour changer de lot sur un équipement de forte activité est grand, donc moins de lots passent sur cette machine, alors on devrait le ressentir sur le temps de cycle total et la productivité.

Il faut donc faire attention car il peut exister un biais entre les résultats fournis par le temps mis pour changer de lot et le temps de cycle ou la productivité.

5.6 Conclusion

Dans ce chapitre nous avons présenté nos choix de modèles nécessaires à la simulation précise de l'unité de fabrication de semi-conducteurs Crolles 300. Nous avons divisé en trois modèles l'unité de fabrication : (1) un modèle de transport, (2) un modèle de production et (3) un modèle de stockage. L'étude de ces trois modèles nous a permis non seulement d'avoir le niveau de précision désiré par modèle, mais aussi de pouvoir mieux mettre en exergue les relations existantes entre transport, production et stockage. On a, au travers d'exemples, pu évaluer certains impacts entre les différentes entités et montrer le niveau d'erreurs qu'il génère.

Nous avons à l'aide de la validation du modèle établi le niveau de confiance global de notre modèle à 80%, même si localement on atteint les 100%. Cela est important pour la suite de notre étude. D'un point de vue industriel elle permet surtout au manager de pouvoir travailler avec un modèle auquel il a « confiance ». Dans ce chapitre, nous avons répondu aux questions suivantes :

- *À quoi sert un modèle ?* Lorsque l'on a à traiter un problème complexe, il est conceptuellement impossible de l'appréhender d'un seul bloc, dans son ensemble. Notre esprit a besoin de « dégrossir » le problème afin de pouvoir le comprendre petit à petit. Une fois le problème découpé en sous-problèmes de tailles plus petites, l'analyse de chacun de ces sous-problèmes nécessite de commencer à en comprendre les grandes lignes puis d'affiner sa compréhension pour enfin comprendre tous les détails. C'est à cela que sert un modèle.
- *Comment effectuer des changements rapides de scénarios dans les systèmes complexes ?* Dans les systèmes complexes, les modifications qui accompagnent les changements de scénarios sont le plus souvent plus complexes que le scénario lui même et prennent beaucoup de temps. Une solution proposée dans cette thèse est l'externalisation des données sensibles. Les résultats ont fait leurs preuves par l'étude de différents scénarios effectués à Crolles 300.
- *Quels sont les paramètres clés de notre système ?* Ce sont le *Low Water Mark* et le *High Water Mark* parce qu'ils sont des paramètres clés sur lesquels se base le système afin de maintenir une adéquation entre les ressources et les besoins.

Le rôle des contrôleurs de véhicules est de maintenir une certaine adéquation entre les ressources disponibles et les besoins des lots. Nous allons dans la suite

proposer une méthode permettant de renforcer les politiques des contrôleurs de véhicules afin d'améliorer la qualité du service.

Troisième partie

AMELIORATION DE LA POLITIQUE DE « SERVICE MINIMUM » GARANTI

Chapitre 6

Analyse du système global de fabrication

Ce chapitre traitera :

- d'un prototype de pré-calcul du *Low Water Mark* (section 6.1)
- de l'étude de la politique du « service minimum » (section 6.3)
- de l'optimisation de la politique du « service minimum » (section 6.4)

Nous allons répondre aux questions suivantes : quelle est la valeur optimale du LWM ? Quelle solution pour la limitation de la politique de gestion par « service minimum » ?

Résumé du chapitre

Ce chapitre étudie de manière plus précise le « comportement » du *Low Water Mark*(LWM). Puisque c'est un des paramètres clé, on se propose de déterminer une valeur du LWM qui maximise la productivité.

Le nombre de combinaisons possibles pour déterminer la valeur du LWM est très élevé et prend un temps considérable si on veut tous les tester. Pourtant c'est un des paramètres critiques de l'unité de Crolles 300. De plus, dans la pratique il est déterminée de façon empirique suivant l'expérience du manager et de son équipe. Pour y remédier, nous proposons un prototype permettant d'avoir un premier résultat satisfaisant. Ce prototype se base sur le nombre de transports internes et sortants de la zone.

Une fois les premières solutions déterminées, nous allons pouvoir les affiner par une étude de la politique de gestion par « service minimum ». Nous utiliserons principalement trois types d'indicateurs : temps de cycle, productivité et temps pour changer un lot par un autre sur le port de chargement d'un équipement. Nous ferons varier deux paramètres à savoir, le nombre de lots en entrée de l'unité de fabrication et le nombre de véhicules. Cette section nous permettra de déterminer une valeur plus précise du LWM.

Cependant, même avec une valeur « optimale » du LWM, la politique de gestion par « service minimum » possède une limitation : *une zone donnée peut se vider totalement de ses véhicules*. Pourtant, l'attribution d'une requête de transport ne peut se faire qu'à l'intérieur d'une zone. Sans véhicules dans la zone, il y a augmentation des temps de transport. Pour y remédier, nous proposons de dédier des véhicules dans une zone. Les véhicules dédiés ne pouvant jamais sortir de la zone.

6.1 Prototype de pré-calcul du *Low Water Mark*

« On peut faire des prévisions sur tout, sauf sur l'avenir ».

LAO Tseu.

6.2 Introduction

Rappelons que la zone d'étude considérée dans ce chapitre est la zone 10 dans le découpage de Crolles 300. Elle correspond à la zone de métrologie juste après l'étape de lithographie. Elle représente 1/4 du trafic total sur 55m de rail et possède la plus forte densité de transport. Elle contient 5 parkings mais ne peut conserver que 4 véhicules au plus à cause de la valeur du *High Water Mark* qui a été fixé à 4.

Le *High Water Mark* est majoritairement utilisé par le système dans deux cas. Le premier cas permet de réduire le risque de congestion d'une zone en « poussant » hors de la zone le surplus de véhicules. En « poussant » les véhicules hors de la zone, le *High Water Mark* concourt au rééquilibrage des zones. Mais ce dernier cas dépend aussi de l'ordre dans lequel on recherche les zones où envoyer les véhicules. Dans le second cas il fixe la condition pour qu'une zone puisse recevoir des véhicules. Si sa valeur est trop petite, la zone refusera assez fréquemment les véhicules. Si par contre sa valeur est trop grande, elle acceptera beaucoup de véhicules.

Le *Low Water Mark* quant à lui représente le besoin minimum en transports de la zone. Il est majoritairement utilisé par le système dans deux cas. Le premier cas sert à fournir un niveau de service minimum entre les ressources et les besoins de transports de la zone. Si sa valeur est trop petite l'inadéquation sera importante. Si par contre elle est trop grande la zone conservera des véhicules dont elle n'a pas besoin, ce qui conduit à une perte en capacité de transport. Dans le second cas, il est utilisé comme seuil à partir duquel le système est autorisé à « appeler » un véhicule d'une zone. Si par exemple $LWM=3$ et qu'il y a quatre véhicules, le système ne pourra emprunter qu'un seul véhicule dans la zone. De manière générale le système ne peut emprunter que $(\text{Nombre de véhicules} - LWM)$ véhicules dans une zone.

Donc le *Low Water Mark* est l'élément clé de la zone. Il détermine son besoin minimum, et c'est de lui que va dépendre la qualité du rééquilibrage des véhicules.

Le *High Water Mark* intervient de manière à fluidifier le trafic. Enfin, la relation $LWM \leq HWM$ montre que fixer la valeur du *High Water Mark* ne donne aucune information sur celle du *Low Water Mark* donc des besoins de la zone. Cela motive notre étude sur le *Low Water Mark* qui de notre point de vue, est un paramètre plus « important » que le *High Water Mark*.

6.2.1 Motivation

L'origine du développement de ce prototype provient du constat suivant : les valeurs du *Low Water Mark* (LWM) dans l'unité de fabrication de Crolles 300mm sont déterminées sur la base de l'expérience du manager et de son équipe, puis ajuster en fonction des besoins de la zone. On aimerait déterminer ces valeurs avec une approche déterministe, tout en fournissant une solution moyenne qui traduise une adéquation entre ressources de transport et besoins en transport des lots de la zone. Ce prototype a donc été développé pour servir d'outil d'aide à la décision, à la compréhension et fournir une première solution « approchante ».

Nous distinguerons par la suite trois types de transports pour une zone : (1) les *transports entrants* qui ont pour origine une position extérieure à la zone et pour destination une position de la zone, (2) les *transports sortants* dont l'origine est interne et la destination externe à la zone et enfin (3) les *transports internes* dont l'origine et la destination sont à l'intérieur de la zone. L'**objectif** de ce prototype est de déterminer des valeurs du LWM en fonction du nombre de transports entrants, sortants et internes.

6.2.2 Principe et algorithme

Principe de l'algorithme du prototype. On est parti du constat que, dans une zone la destination d'une requête de transport est interne ou externe à la zone. Or, le transport est un service pour la production, il doit donc y avoir « suffisamment » de ressources de transport pour supporter les besoins de la production.

Les transports internes correspondent à un enchaînement de tâches internes à la zone. Ces enchaînements s'effectueront plus rapidement s'il y a des véhicules disponibles dans la zone. La première partie de l'algorithme de prévision est en charge

de déterminer le nombre de véhicules nécessaires pour satisfaire la demande interne moyenne en requêtes de transport. Tandis que la seconde partie de l'algorithme de prévision est en charge de déterminer le complément de véhicules nécessaires pour palier aux transports sortants.

le prototype comporte deux fonctions, une fonction apprentissage et une fonction prévision.

L'**apprentissage** permet d'estimer la proportion des différents types de transport de la zone en fonction du nombre total de transports de la zone. Par exemple, en se basant sur les données d'une semaine d'une zone, on obtient que : 20% du transport total correspond au transport interne, 40% au transport entrant et 40% au transport sortant. L'intérêt d'avoir une approximation de la proportion du nombre de transports est le gain de temps qu'il procure dans le temps de calcul. En effet, le calcul précis du nombre de transports par type et par jour est une tâche fastidieuse, tandis que le calcul du nombre total de transports de la zone est bien plus facile à calculer.

La **prévision** utilise les proportions obtenues dans l'apprentissage, afin de *prévoir* le nombre de transports internes, entrants et sortants des différentes zones de la fab. Une fois connu le nombre de transports, la valeur du LWM est déterminé en se basant sur le nombre de transports par heure que peut effectuer un véhicule, d'après les spécifications du constructeur.

La figure 6.1 est une capture d'écran du prototype développé et interfacé sous Visual Basic pour Applications (VBA). Le choix de VBA se justifie par le fait que c'est un des outils connus et utilisés dans l'entreprise, donc plus facilement maintenable.

Les données d'entrées utilisées sont les suivantes :

- **NbJob** : nombre de requêtes de transport au départ de la zone étudiée, vers un emplacement de stockage ou un équipement.
- **Target** : nombre de transports par jour que peut effectuer un véhicule dans des conditions moyennes (c'est une donnée du constructeur).

L'*algorithme 1* détermine la valeur du LWM_i de la $zone_i$ en fonction du nombre de requêtes de transport interne de la zone, c'est-à-dire les transports dont l'origine et la destination sont dans la zone. *Il assure que l'on disposera, en fonction des données*

FIGURE 6.1 – Prototype de pré-calcul du *Low Water Mark*.

Algorithm 1 détermination du LWM en fonction du nombre de véhicules **internes**.

Require: $LWM \neq 0$ {au moins 1 véhicule par zone}
for $i = 0$ to $NbBay$ **do**
 while $(NbJob_i / LWM_i) > Target$ **do**
 $LWM_i \leftarrow LWM_i + 1$
 end while
end for

du constructeur, d'un nombre suffisant de véhicules pour répondre aux requêtes de transport de la zone. Chaque fois que le rapport entre le nombre de requête de transports et la valeur du LWM_i est supérieure à la donnée du constructeur (i.e manque de véhicules), on incrémente le LWM_i d'une unité jusqu'à la condition d'arrêt.

L'algorithme 1 fixe les valeurs du LWM_i en fonction des besoins propres de la zone. L'algorithme 2 va ajuster les valeurs du LWM_i trouvées par l'algorithme 1

Algorithm 2 détermination LWM en fonction du nombre de véhicules **sortants**.

Require: $LWM \neq 0$ {au moins 1 véhicule par zone}

```

for  $i = 0$  to NbBay do
  while  $(LWM_i - NbVehSortant_i) \leq 0$  do
     $LWM_i \leftarrow LWM_i + 1$ 
  end while
end for

```

en fonction de la demande sortante. Soulignons que la demande entrante ne nous intéresse pas dans ce prototype, car les transports entrants alimentent la zone en véhicules, et c'est ce que l'on recherche.

Les véhicules réquisitionnés par l'algorithme 1 peuvent sortir de la zone. C'est le cas lorsqu'il y a plus de transports sortants et cela diminue le niveau d'adéquation des ressources de transport face aux besoins de la zone.

Le **processus de validation du prototype** a consisté à comparer le nombre de transports prévus par le prototype avec ceux obtenus dans l'unité de fabrication réelle. *Les résultats finaux sont les valeurs du LWM pour les différentes zones.* Ce prototype a été validé par l'équipe AMHS de Crolles 300.

Remarque. L'idée développée dans ce prototype est utilisable pour tout système de transport géré par zone. Il permet d'avoir des premières solutions acceptable à affiner par la suite.

6.3 Etude de la politique du « service minimum » garanti

Le minimum pourrait être défini comme la perfection qu'atteint un objet lorsqu'il n'est plus possible de l'améliorer par soustraction.

John Pawson

Motivation : l'objectif du système de transport est d'augmenter adéquation entre les ressources de transport (les véhicules) et les besoins en transport de la production (les lots). Pour atteindre cet objectif, l'utilisation et l'amélioration de politiques de gestion des véhicules sont très utilisées dans la littérature.

A travers notre plan d'expérience, nous avons montré dans la section 5.5.5 la dominance de la politique de « service minimum » sur la politique de « service classique ». Nous avons également identifié le *Low Water Mark* et le *High Water Mark* comme étant des paramètres critiques de réglages. Nous voulons à présent étudier les impacts du choix du LWM sur les performances du système :

- en fonction du nombre de véhicules, section 6.3.1
- en fonction de la charge du système en nombre de lots, section 6.3.2.

Dans la suite de notre étude, nous distinguerons deux types d'équipements pour chaque zone : les équipements se trouvant à l'intérieur de la zone ou *équipements internes* (numérotés de 1 à 7) et les équipements à l'extérieur de la zone ou *équipements externes* (numérotés de 8 à 13). La motivation de cette distinction s'explique par le fait que, nous voulons évaluer l'impact du choix de la valeur du paramètre critique autant à l'intérieur, qu'à l'extérieur de la zone. Le tableau 6.1 représente les caractéristiques de la simulation utilisées dans la suite de notre étude.

Tous les résultats de cette section sont normalisés afin de respecter la confidentialité des résultats. Dans toute la suite nous utiliserons le terme *écart maximum* pour désigner la différence entre la plus grande et la plus petite valeur parmi un ensemble de données.

	Caractéristiques
Temps total de simulation	360 jours
Temps de stabilisation	180 jours
Temps de simulation	180 jours
Nombres d'équipements	250
Nombres de zones	10
Nombres de véhicules	26
Nombres de parkings	35
Politique de routage	Plus court chemin
Politiques de dispatching	<i>Low Water Mark</i>

TABLE 6.1 – Caractéristiques de la simulation pour le « service minimum ».

6.3.1 Impacts du nombre de véhicules

Dans cette section, nous allons étudier comment se comporte le système lorsque le nombre de ressources de transport varie, c'est-à-dire lorsque le nombre de véhicules augmente ou diminue. Cela nous permettra de répondre à la question suivante : « comment choisir la valeur du *Low Water Mark* pour faire un maximum de service avec un minimum de véhicules » ?

6.3.1.1 Temps de cycle et productivité en fonction du nombre de véhicules

Nous avons choisi de commencer les tests avec 22 véhicules. Ce chiffre est motivé par le fait que deux véhicules sont nécessaires par zone. Un pour le chargement et l'autre pour le déchargement de lots. Avec 10 zones, il faut un minimum de 20 véhicules. Mais avec 20 véhicules, le LWM=HWM et il y a rarement de cas où le nombre de véhicules dépasse le LWM. Donc ne peut plus faire du rééquilibrage car il ne peut plus « emprunter » de véhicules dans une zone. Nous avons donc choisi 22 véhicules pour permettre un minimum de rééquilibrage, mais aussi pour tester les limites du système. Les résultats en *italiques* dans tous les tableaux de cette section représentent la meilleure valeur à LWM constant, c'est-à-dire en ligne, tandis que les résultats en **gras** représente la meilleure valeur à nombre constant de lots à charger,

c'est-à-dire en colonne.

Les figures 6.2 et 6.3 illustrent les variations de l'écart-type du temps moyen de cycle lorsque le nombre de véhicules varie de 22 à 27, et pour différentes valeurs du LWM.

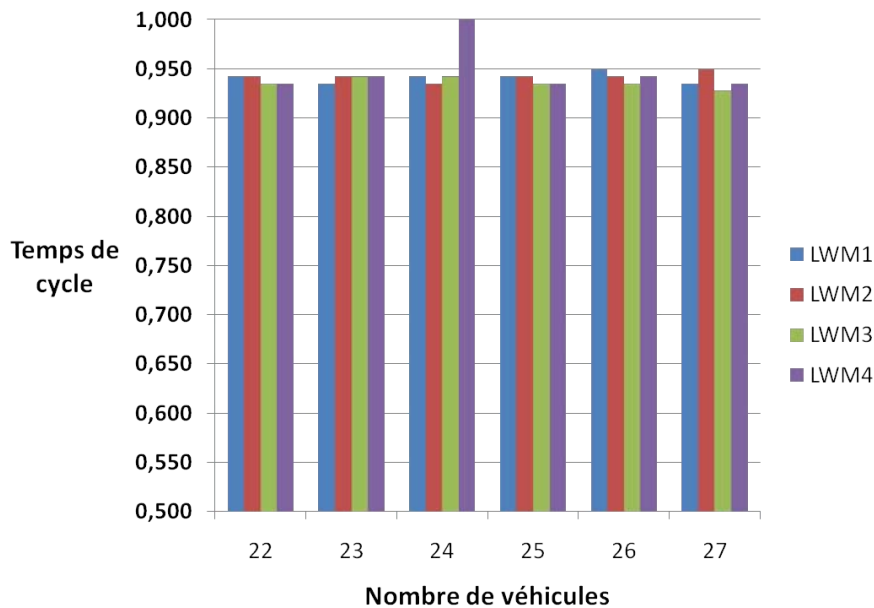


FIGURE 6.2 – temps moyen de cycle.

Afin de mieux comprendre les résultats, fixons quelques idées. Supposons que le temps de cycle d'un produit est de 95 jours avec une variation de ± 4 jours. Une réduction de 6% de ce temps de cycle correspond à 6 jours. Cette amélioration n'est significative dans les délais de livraison que si le nombre de jours gagnés est supérieur à la variabilité du temps de cycle. Tandis qu'avec une réduction de 1,5%, soit 1 jour gagné, l'amélioration n'est pas significative car ce jour gagné peut être inclus dans la variabilité du temps de cycle. Il y a juste réduction de la variabilité du temps de cycle. Si on l'on se réfère aux résultats du tableau 6.2, on observe que :

À **Low Water Mark constant**, le nombre de véhicules a une *influence très faible* le temps moyen de cycle des lots. En effet, pour les valeurs 1, 2 ,3 du LMW, la différence entre la plus grande et la plus petite valeur (écart maximum) du temps

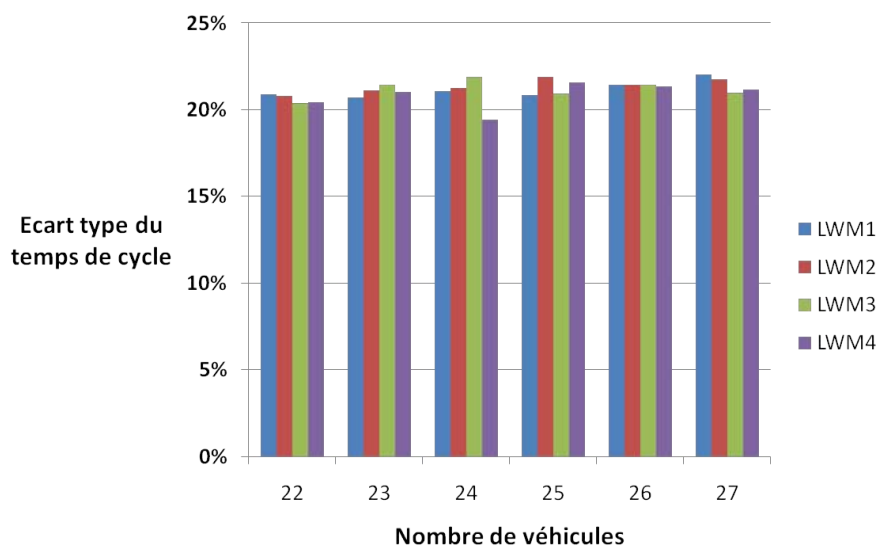


FIGURE 6.3 – temps moyen de cycle.

moyen de cycle est de 1%. Pour LWM=4 l'écart maximum est de 7% à cause du cas 24 véhicules qui est difficilement explicable. Avec un seul véhicule dans la zone (LWM=1) réquisitionné il est difficile d'assurer le chargement et le déchargement des lots. Le système reste donc insensible à l'ajout de véhicule, puisque ce dernier n'est pas ajouté dans la zone d'étude. Pour le cas LWM=2 et 3 le chargement et le déchargement de la zone est assuré mais il faut faire attention à la particularité de la zone d'étude : 1/4 du trafic total, 4 à 5 déposes/enlèvements de lots par minute. Il semble donc que l'on est en-dessous de la demande effective de la zone avec LWM=2, et le meilleur résultat est obtenu est LWM=3 et 27 véhicules. Ces résultats viennent confirmer la tendance observée durant le plan d'expérience (section 5.21). La nouveauté dans ces résultats est qu'ils montrent que : *dans une gestion par zones de l'unité de fabrication, il est nécessaire de satisfaire en premier le niveau de demande de chaque zone*. Le problème de la bonne répartition des véhicules est un élément important pour la gestion des véhicules.

À nombre de véhicules constant, le LWM influe *très faiblement* sur le temps moyen de cycle des lots. L'écart maximum est de 1% pour les un nombre de véhicules de 22 à 27 sauf, pour le cas de 24 où il est de 7%. Ce dernier cas est difficilement

explicable. Faire varier la valeur du LWM de 1 à 4 lorsque le nombre de véhicules est constant, c'est emprunter des véhicules dans des zones pour les mettre dans la zone d'étude. On déséquilibre les autres zones. Ce constat permet d'expliquer la faible variation des résultats obtenus. Ils vont fortement dépendre du niveau de déficit en transports des zones d'où on a pris des véhicules. La nouveauté dans ce résultat est que *l'amélioration seule du niveau de demande d'une zone (augmentation du LWM), sans les autres zones, n'améliore pas le temps de cycle*. L'unité de fabrication est un milieu où tout est lié. Cela peut s'expliquer par le fait que le LWM est un paramètre d'ajustement local, permettant d'améliorer l'adéquation entre ressources de transport et besoins en production.

LWM	Nombre de véhicules					
	22	23	24	25	26	27
1	0,942	0,935	0,942	0,942	0,949	<i>0,935</i>
2	0,942	0,942	0,935	0,942	0,942	0,949
3	0,935	0,942	0,942	0,935	0,935	0,928
4	0,935	0,942	1,000	0,935	0,942	<i>0,935</i>

TABLE 6.2 – **temps moyen de cycle** en fonction des véhicules.

LWM	Nombre de véhicules					
	22	23	24	25	26	27
1	21	21	21	21	21	22
2	21	21	21	22	21	22
3	20	21	22	21	21	21
4	20	21	19	22	21	21

TABLE 6.3 – **Écart type** du temps moyen de cycle (en %) en fonction des véhicules.

Les résultats sur l'écart type du temps moyen de cycle présentés dans le tableau 6.3 confirment notre précédente analyse. La politique de gestion par « service minimum » présente très peu de variabilité que ce soit à nombre de véhicules constant ou à *Low Water Mark* constant.

La productivité ici représente le nombre de produits terminés sur la période de simulation considérée qui est de 6 mois. Elle est illustrée par la figure 6.4. Les

observations de cette figure confirment les résultats obtenus précédemment avec la figure 6.2 sur le temps de cycle. En effet, le temps de cycle et la productivité sont inversement proportionnels. Un temps de cycle long signifie moins de lots en sortie, donc faible productivité. Si le temps de cycle varie très peu, la productivité aussi variera très peu.

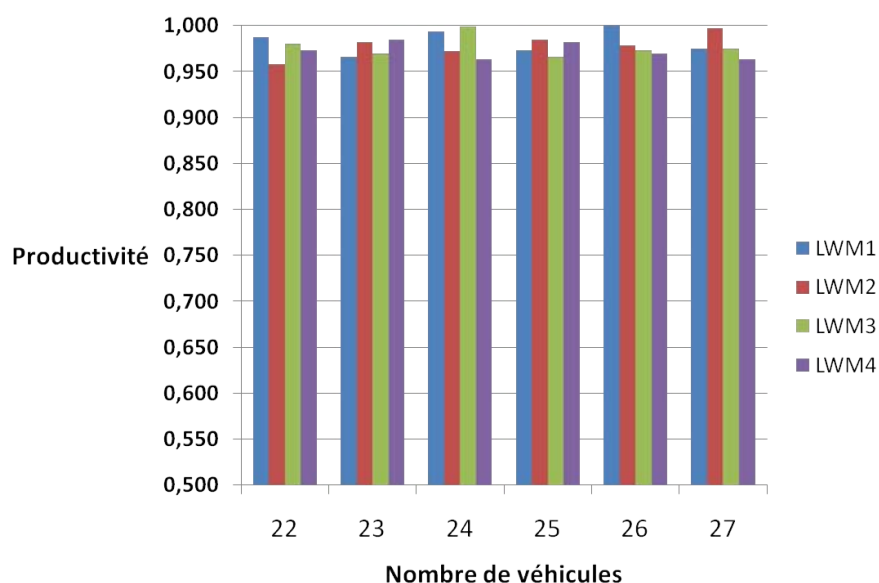


FIGURE 6.4 – Productivité en fonction du nombre de véhicules.

Plus précisément, les résultats numériques sont donnés par le tableau 6.4.

À *LWM constant*, l'écart maximum moyen est de 3%. Il y a très peu de variation lorsque le nombre de véhicules varie car le temps de cycle est beaucoup plus long. La nouveauté de ces résultats est que *l'augmentation de la productivité d'une unité de fabrication n'est pas directement liée à celle du nombre de véhicules*.

À **nombre de véhicules constant**, l'écart maximum moyen est également d'environ 3%. Ici également, la variation de la productivité est en dents de scie. Ce résultat nous montre que la modification du paramètre LWM seul ne suffit pas à l'augmentation de la productivité. Soulignons que *le cas LWM=3 donne le meilleur rapport temps de cycle - productivité*.

LWM	Nombre de véhicules					
	22	23	24	25	26	27
1	0,987	0,966	0,994	0,973	1,000	0,975
2	0,958	0,982	0,972	0,985	0,978	0,997
3	0,980	0,970	0,999	<i>0,966</i>	0,973	0,975
4	0,973	0,985	<i>0,963</i>	0,982	0,970	0,963

TABLE 6.4 – **Productivité** en fonction des véhicules.

Remarque : d'un point de vue financier l'écart maximum observé n'est pas négligeable. En effet, 3% de 20.000 plaques représente 600 plaques. A raison d'environ \$ 4000 la plaque, cela représente un gain d'environ \$ 2.400.000.

6.3.1.2 Carrier Exchange Time en fonction du nombre de véhicules

Nous nous sommes intéressés à l'impact global de la variation du nombre de véhicules. Nous allons à présent étudier l'impact de cette variation au niveau local. L'indicateur utilisé est le *Carrier Exchange Time* (CET) ou temps de changement d'un lot sur un équipement. Nous distinguerons deux types d'équipements : les équipements internes à la zone et ceux externes de la zone.

Cas des équipements internes. Les résultats numériques du *Carrier Exchange Time* sont présentés dans le tableau 6.5. Ils sont exprimés en secondes mais normalisés.

À LWM constant, l'écart maximum moyen varie entre 5 et 16%. Le *Carrier Exchange Time* décroît lorsque le nombre de véhicules augmente. Cette augmentation permet au système des appels de véhicules qui s'ajoutent à la zone d'étude pour combler sa demande importante. Donc en local, le système est sensible à la variation du nombre de véhicules. Ce qui est nouveau dans ce résultat c'est qu'*au niveau local la politique de gestion par « service minimum » est sensible à l'augmentation ou à la réduction du nombre de véhicules.*

À nombre de véhicules constant, l'écart maximum moyen varie entre 8 et 18%. Le *Carrier Exchange Time* décroît lorsque le LWM croit. Ce résultat confirme

l'importance du paramètre LWM en local. En effet, les résultats confirment le fait que le LWM est essentiel pour améliorer l'enchaînement de tâches sur les équipements de production.

	LWM	Nombre de véhicules					
		22	23	24	25	26	27
Equipement 1	1	0,911	0,984	0,932	0,911	0,858	<i>0,847</i>
	2	0,937	0,911	0,842	0,853	0,826	<i>0,805</i>
	3	0,853	0,842	0,779	0,779	0,753	0,768
	4	0,816	0,811	0,795	0,795	0,753	0,768
Equipement 2	1	0,932	1,000	0,942	0,926	0,884	<i>0,874</i>
	2	0,963	0,932	0,863	0,868	0,842	<i>0,826</i>
	3	0,879	0,874	0,805	0,805	0,779	0,795
	4	0,832	0,837	0,821	0,821	0,779	0,795
Equipement 3	1	0,689	0,737	0,689	0,674	0,653	<i>0,637</i>
	2	0,711	0,684	0,637	0,632	0,616	<i>0,600</i>
	3	0,647	0,632	0,595	0,589	<i>0,579</i>	0,589
	4	0,621	0,605	0,589	0,589	0,574	0,579
Equipement 4	1	0,895	0,953	0,884	0,868	0,837	<i>0,821</i>
	2	0,905	0,874	0,816	0,811	0,784	<i>0,768</i>
	3	0,821	0,805	0,768	0,758	<i>0,742</i>	0,753
	4	0,805	0,784	0,763	0,747	0,737	0,742
Equipement 5	1	0,911	0,968	0,895	0,879	0,832	<i>0,811</i>
	2	0,889	0,853	0,784	0,784	0,753	<i>0,726</i>
	3	0,774	0,763	0,726	0,726	<i>0,695</i>	0,705
	4	0,779	0,753	0,721	0,684	0,663	0,679
Equipement 6	1	0,905	0,905	0,879	0,858	0,858	<i>0,853</i>
	2	0,837	0,837	0,805	0,800	<i>0,779</i>	<i>0,779</i>
	3	0,747	0,768	0,789	0,789	0,779	<i>0,763</i>
	4	0,816	0,821	0,768	0,716	0,716	0,716
Equipement 7	1	0,863	0,853	0,826	0,816	<i>0,795</i>	<i>0,795</i>
	2	0,779	0,768	0,747	0,737	0,732	<i>0,711</i>
	3	0,684	0,716	0,732	0,726	0,721	<i>0,700</i>
	4	0,779	0,758	0,711	0,647	0,647	0,653

TABLE 6.5 – *Carrier Exchange Time* moyen en fonction des véhicules [équipements internes].

Cas des équipements externes. Les résultats numériques du *Carrier Exchange Time* sont présentés dans le tableau 6.6.

À **LWM constant**, l'écart maximum moyen varie entre 15% et 39%. Les plus fortes variations correspondent celles des équipements 8 et 9 qui échangent directement les lots avec la zone d'étude. Cette variation est plus petite, entre 1 et 5%, avec les équipements numérotés de 10 à 13 qui n'échangent pas directement de lots avec la zone d'étude. Le CET pour les équipements externes est décroissant lorsque le nombre de véhicules augmente.

À **nombre de véhicules constant**, hormis les équipements 8 et 9 directement liés à la zone d'étude, la variation du LWM de la zone a peu d'impact sur le CET des équipements externes, $\approx 1\%$. Ces résultats nous donnent une information précieuse. *La détermination des valeurs du LWM d'une zone influe directement sur les équipements qui échangent des lots avec cette zone.* Par contre, pour les autres équipements, il y a peu d'impact.

Conclusion : *le LWM n'est pas un paramètre sensible aux petites variations du nombre de véhicules (1 ou 2 véhicules). La politique de « service minimum » présente une très faible variabilité pour le temps moyen de cycle des lots, lorsque le nombre de véhicules varie.*

Les écarts types des temps moyens de cycle sont donnés dans l'annexe A, autant pour les équipements internes qu'externes. Nous pouvons à présent répondre à la question suivante : **comment faire un maximum de service avec un minimum de ressources (véhicules) ?** La politique de gestion par « service minimum » détermine les conditions minimales par zone (nombre de véhicules), nécessaires pour mettre en adéquation ressources et besoins en transport. La politique de repositionnement des véhicules cherche à reconstituer en temps réel les conditions de service prédéfinies dans chaque zone. La réponse à la question est la suivante : *les résultats obtenus suggèrent de déterminer une valeur du paramètre LWM la plus proche du besoin de la zone, et de fournir le nombre minimal de véhicules nécessaires.*

	LWM	Nombre de véhicules					
		22	23	24	25	26	27
Equipement 8	1	0,832	0,948	0,718	0,653	0,680	0,570
	2	0,959	0,876	0,821	0,698	0,619	0,574
	3	0,821	0,790	0,540	0,515	0,478	0,546
	4	0,763	0,608	0,619	0,615	0,667	0,622
Equipement 9	1	0,955	0,935	0,753	0,715	0,715	0,595
	2	1,000	0,979	0,856	0,766	0,653	0,608
	3	0,849	0,784	0,495	0,536	0,502	0,564
	4	0,804	0,615	0,643	0,639	0,646	0,708
Equipement 10	1	0,447	0,447	0,436	0,436	0,436	0,433
	2	0,450	0,447	0,433	0,436	0,433	0,433
	3	0,430	0,436	0,433	0,433	0,433	0,426
	4	0,440	0,440	0,440	0,426	0,423	0,423
Equipement 11	1	0,526	0,498	0,488	0,488	0,485	0,481
	2	0,526	0,512	0,495	0,495	0,488	0,485
	3	0,533	0,515	0,502	0,502	0,498	0,481
	4	0,533	0,526	0,505	0,502	0,491	0,481
Equipement 12	1	0,381	0,385	0,375	0,375	0,368	0,371
	2	0,388	0,385	0,375	0,375	0,371	0,375
	3	0,357	0,381	0,371	0,371	0,368	0,354
	4	0,375	0,378	0,378	0,357	0,354	0,357
Equipement 13	1	0,364	0,340	0,323	0,326	0,320	0,316
	2	0,357	0,344	0,333	0,326	0,326	0,323
	3	0,337	0,347	0,326	0,323	0,326	0,309
	4	0,340	0,340	0,337	0,320	0,320	0,316

TABLE 6.6 – *Carrier Exchange Time* moyen en fonction des véhicules [équipements externes].

6.3.2 Impacts du nombre de lots

Dans la section 6.3.1, nous avons étudié la sensibilité du système en faisant varier le nombre de véhicules. Dans cette section, nous allons étudier un autre type de sensibilité liée à la charge de l'unité de fabrication (nombre de lots chargés par semaine). On souhaite ainsi déterminer, sous la politique de « service minimum », à partir de quelle charge le système de production devient un goulet d'étranglement.

Pour des raisons de confidentialité, il n'est pas possible de donner les valeurs de la charge de l'unité de fabrication. Nous noterons donc X le nombre de lots lancés par semaine dans l'unité de fabrication. Puis, nous donnerons la valeur de l'augmentation de X (en pourcentage) en fonction des différents scénarios.

6.3.2.1 Temps de cycle et productivité

Les résultats de ces tests sont illustrés par les figures 6.5 pour le temps de cycle et 6.6 pour l'écart type. Les tendances de la figure 6.5 illustrant le temps de cycle montrent qu'à LWM constant, le temps moyen de cycle augmente quand la charge des lots en entrée augmente. Tandis qu'à charge constante, le temps de cycle *ne semble pas sensible* aux variations du LWM. On observe également une faible variabilité du temps moyen de cycle autant à LWM constant qu'à charge constante (figure 6.6). Nous rappelons que le nombre total de véhicules est constant et vaut 26. Les résultats en *italiques* dans tous les tableaux de cette section représentent la meilleure valeur à LWM constant, c'est-à-dire en ligne. Tandis que les résultats en **gras** représentent la meilleure valeur à nombre constant de lots à charger, c'est-à-dire en colonne. Les résultats numériques de la figure 6.6 donnés par le tableau 6.7 étayent nos précédentes observations.

À LWM constant, l'écart maximum du temps moyen de cycle est de 16% pour toutes les valeurs du LWM. Plus on augmente le nombre de lots en entrée, plus le temps de cycle des lots augmente. Le nombre de lots en entrée a un impact *très significatif* sur le temps de cycle des lots. La loi de little [61] est vérifiée. Le temps moyen de cycle a une *très faible variabilité* en fonction du nombre de lots quand on est à LWM constant (voir tableau 6.8).

À nombre de lots en entrée constant, l'écart maximum du temps cycle moyen est de 1% quelque soit le nombre de lots en entrée. Ainsi la variation du LWM a *très peu d'impact* sur le temps de cycle des lots, lorsque la charge des lots entrée est constante.

La figure 6.7 illustre la variation de la productivité. Elle montre que la productivité est une fonction décroissante du nombre de lots en entrée. Par contre, à nombre de lots en entrée constant, les résultats suggèrent que la meilleur productivité est obtenue lorsque pour la valeur 3 du LWM.

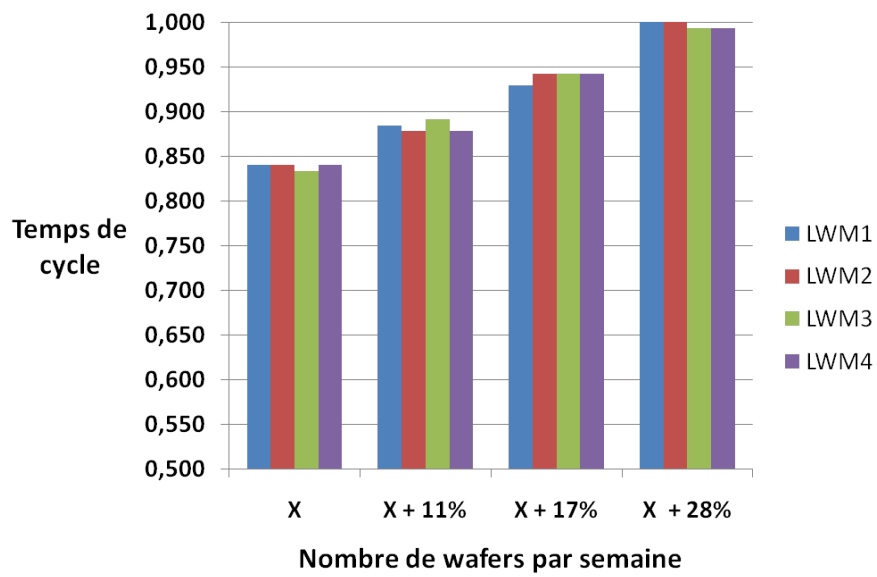


FIGURE 6.5 – temps moyen de cycle.

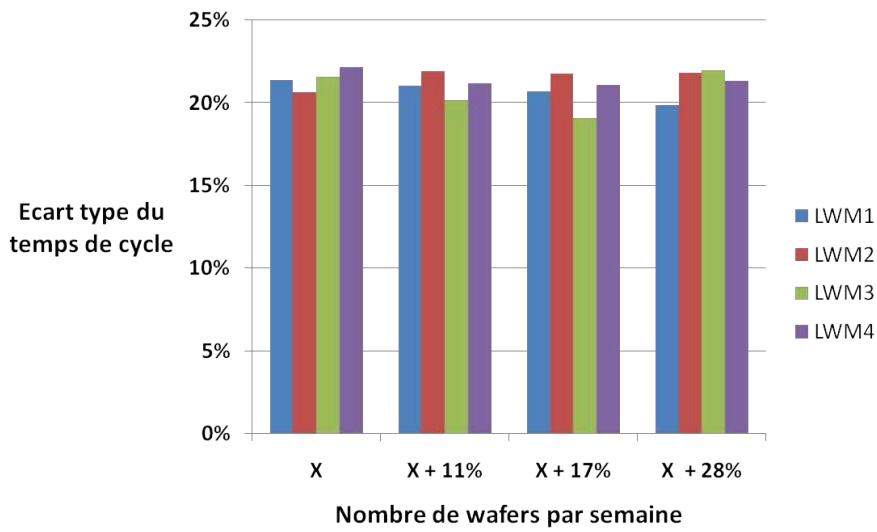


FIGURE 6.6 – Écart type du temps moyen de cycle.

Ces résultats montrent que *la productivité est inversement proportionnelle au*

	LWM	Nombres de wafers par semaine			
		X	X + 11%	X + 17%	X + 28%
Temps de cycle	1	0,840	0,885	0,929	1,000
	2	0,840	0,878	0,942	1,000
	3	0,833	0,891	0,942	0,994
	4	0,840	0,878	0,942	0,994

TABLE 6.7 – temps moyen de cycle en fonction de la charge des wafers en entrée.

	LWM	Nombres de wafers par semaine			
		X	X + 11%	X + 17%	X + 28%
Écart type	1	21,30	21,01	20,69	19,87
	2	20,61	21,90	21,77	21,79
	3	21,54	20,14	19,05	21,94
	4	22,14	21,17	21,09	21,29

TABLE 6.8 – Écart type du temps de cycle en fonction de la charge des wafers en entrée.

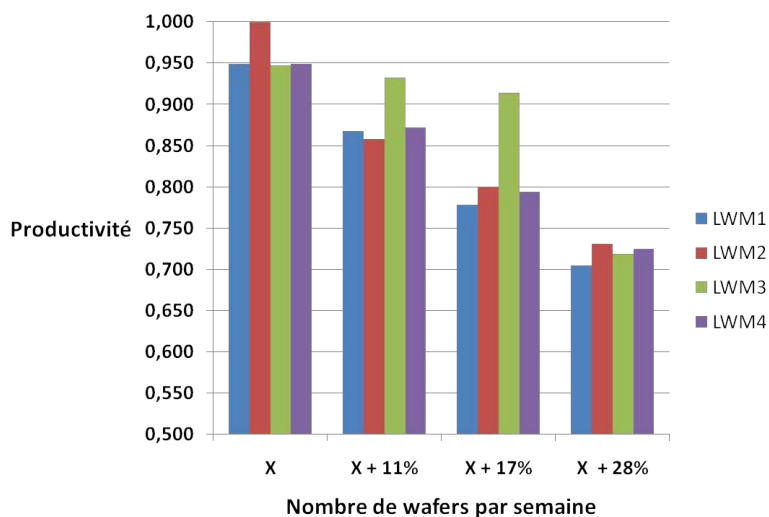


FIGURE 6.7 – Productivité en fonction de la charge des wafers en entrée.

	LWM	Nombres de wafers par semaine			
		X	X + 11%	X + 17%	X + 28%
Productivité	1	0,949	0,868	0,778	0,705
	2	1,000	0,858	0,800	0,731
	3	0,947	0,933	0,914	0,718
	4	0,949	0,871	0,794	0,725

TABLE 6.9 – Productivité en fonction de la charge des wafers en entrée.

nombre de lots en entrée.

6.3.2.2 Carrier Exchange Time en fonction du nombre de lots en entrée

Nous allons à présent observer l'impact de la variation du nombre de lots en entrée à un niveau local c'est-à-dire celui de la zone. L'indicateur utilisé est toujours le *Carrier Exchange Time*. Les résultats concernant les équipements internes sont donnés par le tableau 6.10 et ceux pour les équipements externes par le tableau 6.11.

À **LWM constant**, l'écart maximum moyen est de 2%. Le temps mis pour changer de lot sur un équipement *varie très peu* lorsque le nombre de lots en entrée augmente. Un écart de 2% dans le calcul du CET n'est pas significatif, car la variabilité du CET est du même ordre de grandeur (voir Annexe A).

À *charge constante*, le temps mis pour changer un lot sur le port de chargement d'un équipement *décroit significativement* lorsque le LWM augmente pour les équipements internes. L'écart maximum est d'environ 10%. Ces résultats (niveau local) confirment ceux obtenus précédents (au niveau global). En effet, plus le CET est petit, plus la cadence des lots sur les équipements de production augmente. On a donc une réduction du temps de cycle global des lots, donc augmentation de la productivité à charge constante.

A présent regardons les résultats pour les équipements externes présentés par le tableau 6.11.

A LWM constant, ce sont les équipements 8 et 9 qui échangent directement les lots avec la zone d'étude, qui présentent le plus grand écart maximal moyen, soit $\tilde{9}\%$. Les autres équipements, de 10 à 13, présentent un écart maximal de 1%. Ces

	LWM	Nombres de wafers par semaine			
		X	X + 11%	X + 17%	X + 28%
Equipement 1	1	<i>0,943</i>	0,966	0,943	0,977
	2	<i>0,891</i>	0,897	0,909	0,891
	3	0,869	0,886	0,874	0,886
	4	<i>0,869</i>	0,874	0,880	0,880
Equipement 2	1	0,971	0,989	<i>0,960</i>	1,000
	2	0,920	<i>0,914</i>	0,931	0,914
	3	<i>0,903</i>	0,903	0,909	0,920
	4	0,897	0,903	0,909	0,909
Equipement 3	1	<i>0,709</i>	0,714	0,714	0,726
	2	<i>0,663</i>	0,663	0,669	0,674
	3	0,646	0,651	0,651	0,657
	4	<i>0,646</i>	0,646	0,651	0,651
Equipement 4	1	<i>0,914</i>	0,926	0,920	0,931
	2	0,857	<i>0,851</i>	0,863	0,869
	3	<i>0,829</i>	0,840	0,834	0,840
	4	0,823	0,823	0,834	0,829
Equipement 5	1	<i>0,909</i>	0,937	0,931	0,943
	2	0,817	<i>0,811</i>	0,823	0,834
	3	<i>0,766</i>	0,789	0,777	0,789
	4	0,749	0,749	0,771	0,760
Equipement 6	1	<i>0,920</i>	0,954	0,966	0,954
	2	<i>0,857</i>	0,863	0,869	0,886
	3	<i>0,823</i>	0,834	0,840	0,851
	4	0,800	0,811	0,823	0,823
Equipement 7	1	<i>0,857</i>	0,891	0,886	0,903
	2	<i>0,794</i>	0,800	0,823	0,817
	3	<i>0,743</i>	0,766	0,760	0,771
	4	0,726	0,731	0,754	0,749

TABLE 6.10 – *Carrier Exchange Time* en fonction de la charge des wafers en entrée [équipements **internes**].

résultats nous apprennent que la variation de la charge des lots en entrée impact beaucoup plus des équipements critiques externes (8 et 9).

À nombre de lots en entrée constant, l'écart maximum du temps moyen de

cycle est à peine de 1%, donc pas du tout significatif, pour les équipements de 10 à 13. Par contre pour les équipements échangeant directement avec la zone d'étude (équipement 8 et 9), l'écart maximal est plus grand. En particulier lorsque le nombre de lots en entrée est de $X + 17\%$ et $X + 28\%$, dans ce cas l'écart maximum est $\tilde{11}\%$.

	LWM	Nombres de wafers par semaine			
		X	X + 11%	X + 17%	X + 28%
Equipement 8	1	0,815	0,823	0,914	0,806
	2	0,802	0,806	0,797	0,909
	3	0,776	0,823	0,802	0,832
	4	0,819	0,828	0,836	<i>0,815</i>
Equipement 9	1	0,810	0,862	1,000	0,845
	2	0,836	0,832	0,871	0,966
	3	<i>0,823</i>	0,849	0,849	0,888
	4	<i>0,853</i>	0,866	0,884	0,879
Equipement 10	1	<i>0,547</i>	0,552	0,552	0,556
	2	0,543	0,547	0,552	0,552
	3	<i>0,543</i>	0,547	0,552	0,552
	4	<i>0,543</i>	0,547	0,552	0,552
Equipement 11	1	0,608	0,608	0,616	0,621
	2	<i>0,612</i>	0,616	0,621	0,621
	3	<i>0,616</i>	0,616	0,625	0,629
	4	<i>0,621</i>	0,625	0,629	0,629
Equipement 12	1	0,466	0,474	0,474	0,483
	2	<i>0,466</i>	0,470	0,478	0,483
	3	<i>0,470</i>	0,474	0,474	0,483
	4	<i>0,470</i>	0,474	0,483	0,483
Equipement 13	1	0,401	0,401	0,401	0,401
	2	0,409	0,405	<i>0,401</i>	0,401
	3	0,414	0,418	<i>0,405</i>	0,409
	4	0,422	0,418	0,414	<i>0,409</i>

TABLE 6.11 – *Carrier Exchange Time* en fonction de la charge des wafers en entrée [équipements **externes**].

Conclusion : si le nombre de lots en entrée de l'unité de fabrication augmente et si on fixe le nombre de véhicules, ce sont les équipements de production qui deviennent

un goulet d'étranglement. Ces résultats montrent également que le choix de la valeur du LWM est très important pour la gestion de la capacité du système de transport. Par exemple, pour un LWM=3 et pour un nombre de lots qui passe de X+17% à X+28% lots, la productivité présente une différence de près de 20% de différence.

Quand les équipements de production deviennent-ils un goulet d'étranglement ? Les résultats suggèrent que ce sont les équipements critiques (dans notre exemple équipements 8 et 9) qui déterminent le niveau d'encombrement. Ainsi, lorsque le nombre de lots en entrée sature ces équipements, c'est la production qui devient un goulet d'étranglement.

6.4 Application Industrielle

« On peut toujours améliorer un record. On peut toujours faire mieux. »

Bruno Peyron

6.4.1 Introduction

Lorsqu'un lot a besoin d'être transporté, il fait appel à un véhicule. S'il n'y a pas de véhicule disponible, le temps d'attente du lot augmente, de même que le temps de non production de l'équipement où le lot doit aller. L'amortissement du coût de l'équipement est alors plus long et les lots en attente représentent de la trésorerie non disponible pour l'entreprise.

De manière générale, ce problème est celui de la garantie d'un service minimum dans des lieux à forte demande. Par exemple, un client qui attend un taxi à une gare, un malade qui attend d'être pris en charge aux urgences, etc. La difficulté est donc de pouvoir maintenir un niveau de service dans tout le réseau. Plus précisément, notre problème concerne l'interruption momentanée d'une ligne de production sans que la cause ne soit la panne de l'équipement. Nous cherchons à réduire les situations où c'est le transport qui est un goulet d'étranglement pour la production.

Si le système est automatisé alors cela signifie que c'est le transport qui devient un goulet d'étranglement pour la production. Cette situation est inacceptable car le transport est un service pour la production.

Nous avons souligné dans le chapitre plan d'expérience 5.5 une limitation de la politique du « service minimum » : la zone vide. C'est *la possibilité pour une zone de se « vider »*, c'est-à-dire de ne plus avoir de véhicules dans la zone. Or, dans le cas de Crolles 300, on ne peut assigner de requêtes de transport à un lot que si le lot est à l'intérieur d'une zone. Si la zone ne contient pas de véhicules, il y a augmentation du temps d'assignation de la requête de transport, donc allongement du temps de cycle du lot.

Pour y remédier, nous proposons de dédier des véhicules à une zone donnée. Nous espérons ainsi réduire (voire supprimer) les temps d'interruption des équipements

de production dus au système de transport.

Nous commencerons par définir ce nous appelons « véhicule dédié » dans la section 6.4.2. Ensuite, nous décrirons brièvement les conditions de l'expérience et nous discuterons des résultats obtenus à la section 6.4.3.

6.4.2 Les véhicules dédiés

Définition : un véhicule dédié est un véhicule qui est affecté à une et une seule zone. Cela signifie que le véhicule ne peut répondre qu'aux requêtes de transport interne de la zone, et qu'il ne peut en aucun cas sortir de la zone (voir figure 6.8).

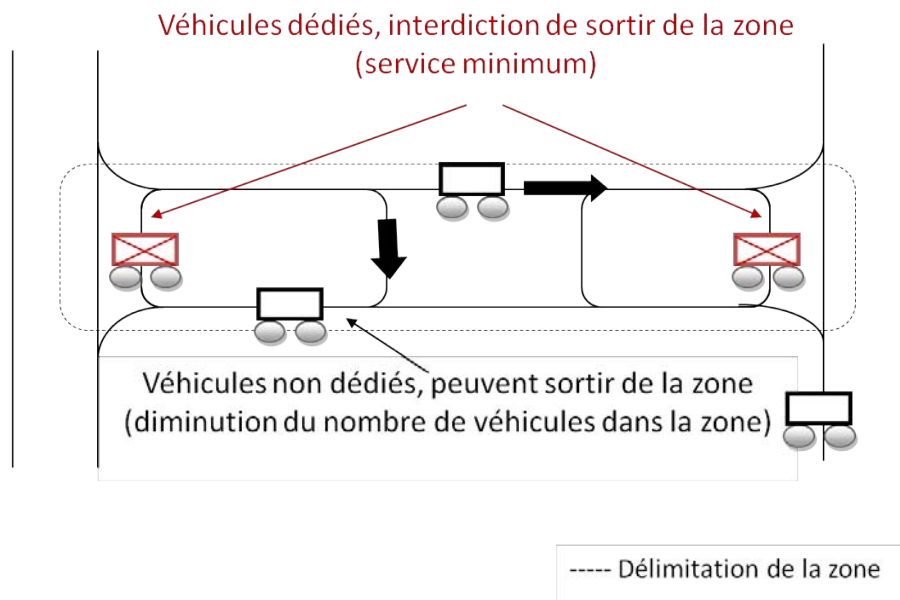


FIGURE 6.8 – Illustration de véhicules dédiés.

Dans la littérature, on utilise le plus souvent « les véhicules partagés » pour améliorer les capacités et la performance des systèmes de transport [14], [41], [70]. Un « véhicule partagé » est un véhicule qui n'est pas dédié. Le véhicule peut donc répondre à n'importe quelle requête provenant d'une zone. Rappelons que ce n'est pas possible dans l'unité de fabrication de Crolles 300mm. Les « véhicules partagés » ne sont pas adaptés à notre situation.

Pour pallier à cette contrainte, nous allons remplacer les « véhicules partagés » par des « véhicules dédiés » définis précédemment. Soulignons que le système de transport de Crolles 300mm utilise la politique de gestion par « service minimum ». Pour rappel, elle consiste à essayer de maintenir dans une zone un nombre minimum de véhicules et à « pousser » hors de la zone les véhicules en « surplus ». Le nombre minimum de véhicules correspond à la valeur du *Low Water Mark* (LWM), tandis que la limite à partir de laquelle le système doit commencer à « pousser » les véhicules hors de la zone, correspond au *High Water Mark* (HWM).

Le tableau 6.12 décrit notre plan d'expérience sur les véhicules dédiés. Nous étudierons le comportement de deux véhicules dédiés dans la zone d'étude. Nous utiliserons deux autres véhicules pour assurer l'échange avec l'extérieur. Nous ferons varier le nombre de véhicules afin de déterminer l'impact des véhicules dédiés sur la productivité de l'unité de fabrication.

Nombre de véhicules dans la zone critique	Caractéristique
2 véhicules dédiés (LWM=4)	restent dans la zone

TABLE 6.12 – Plan d'expérience pour les véhicules dédiés.

Nous rappelons que nous utilisons une approche par simulation à événements discrets. Nous avons modélisé finement l'unité de fabrication de Crolles 300mm en utilisant le logiciel AutoMod. Les indicateurs étudiés sont le temps de cycle et la productivité des lots. On considère trois types de produits. Les caractéristiques de la simulation sont identiques à celles de la section 6.3.

6.4.3 Résultats sur les véhicules dédiés

Avant de présenter les résultats, rappelons que la zone étudiée est la plus dense¹ en termes de trafic. On y relève 4 à 5 déposes/enlèvements par minute et le transport interne y est le plus élevé. En effet, c'est une zone où se concentrent les équipements de mesure dont les différentes étapes (probables) sont illustrées par la figure 6.9. Par exemple, si après l'inspection macroscopique automatique, le lot présente peu

1. Densité = nombre de lots sortis de l'équipement / longueur de rail.

de défauts, il peut directement aller sur une étape de mesure overlay (flèche en pointillé). S'il y a des doutes, il passe par quatre étapes de mesure successives, qui se trouvent dans la zone étudiée. Cela justifie la pertinence d'introduire des véhicules dédiés dans cette zone.

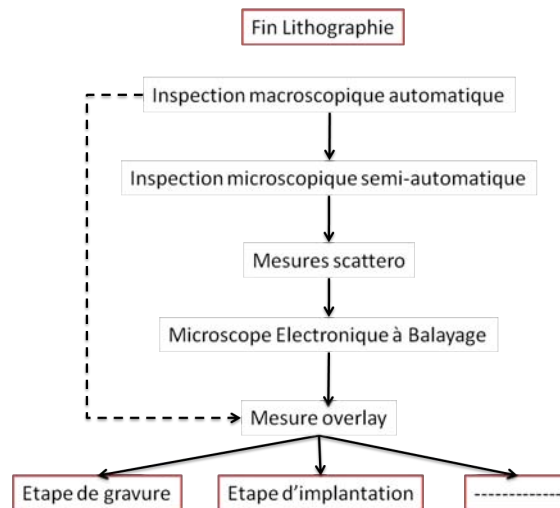


FIGURE 6.9 – Enchaînements possibles de tâches dans la zone étudiée.

Deux véhicules dédiés dans la zone critique

Nous présentons à présent les résultats obtenus lorsque l'on affecte deux véhicules dédiés à la zone cible. Le choix de deux véhicules dédiés se justifie par la nécessité d'avoir un véhicule pour le chargement et un autre pour le déchargement. Les résultats obtenus dans la section 6.3 (gestion de la politique de service minimum) sont comparés avec ceux obtenus en utilisant les véhicules dédiés.

26 véhicules au total dont 2 véhicules dédiés. On observe une amélioration de 30% sur la productivité globale, comme l'illustre la figures 6.10, et une réduction de près de 31% sur le temps moyen de cycle, illustrée par la figure 6.11. Les résultats numériques sont présentés dans le tableau 6.13. On peut observer en effet qu'à **nombre total de véhicules constant**, soit 26 véhicules, la productivité est

meilleure avec les véhicules dédiés quelque soit la valeur du *Low Water Mark* que sans véhicule dédié. On observe le même résultat pour le temps moyen de cycle. Cependant, la variabilité du temps moyen de cycle est une fois et demi plus élevée avec les véhicules dédiés. Ceci peut s'expliquer par le fait que le système ne peut garantir la présence quasi-certaine de deux véhicules pour assurer les échanges avec l'extérieur (LWM=4 auquel on retranche 2 véhicules dédiés à la zone).

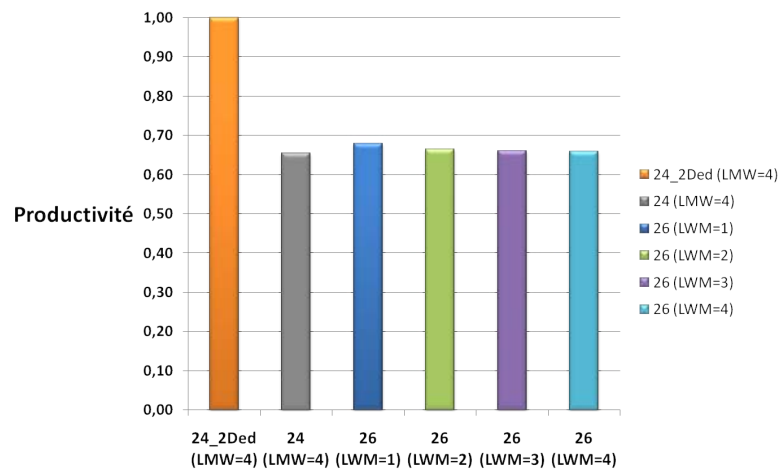


FIGURE 6.10 – Productivité : comparaison entre avec et sans véhicules dédiés.

	Nombre de véhicules (LWM)					
	24 - 2 Ded.	24 (4)	26 (1)	26 (2)	26 (3)	26 (4)
Temps moyen de cycle	0,630	1,000	0,949	0,942	0,935	0,942
Ecart type	30%	19%	21%	21%	21%	21%
Productivité	1,000	0,654	0,679	0,665	0,661	0,659

TABLE 6.13 – Temps moyen de cycle et productivité avec et sans véhicules dédiés.

Ces résultats montrent que, dans une gestion du transport d'une unité de fabrication par zones, les véhicules dédiés à une zone sont une bonne alternative pour améliorer la cadence d'enchaînement des tâches (*continuous processing*) sur les équipements de production. Ce résultat est nouveau car il fournit une approche qui, combinée à la gestion du transport par la politique du « service minimum » ,

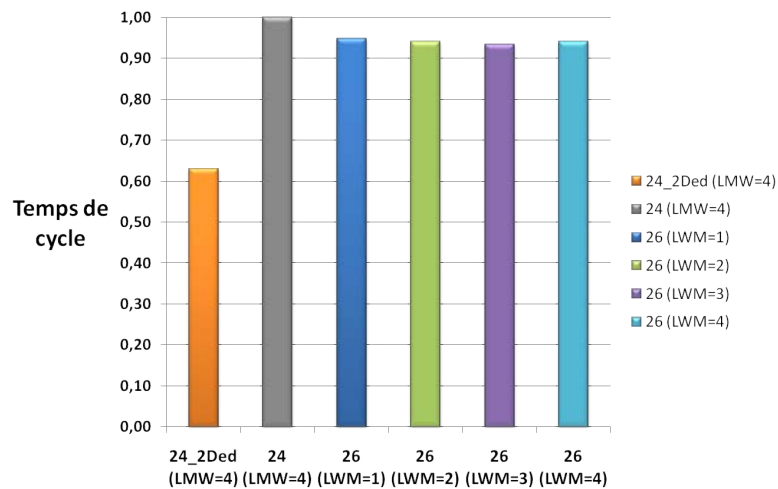


FIGURE 6.11 – Temps moyen de cycle avec et sans véhicules dédiés.

permet d’optimiser le système de transport dans les *grandes* unités de fabrication de semi-conducteurs.

En conséquence, les véhicules dédiés permettent une non interruption de la ligne de production, car ils garantissent qu’il y a toujours au moins un véhicule dans la zone. La probabilité pour que la zone se vide en nombre de véhicules est nulle. D’un point de vue scientifique, cela contribue à réduire la variabilité de la cadence des lots sur les équipements de production, qui sont plus réguliers. On a ainsi augmenté le niveau d’adéquation entre requêtes de transport et capacité de transport de la zone. Puisque c’est l’une des zones les plus denses de l’unité de fabrication, l’impact est ressenti sur toute l’unité de fabrication.

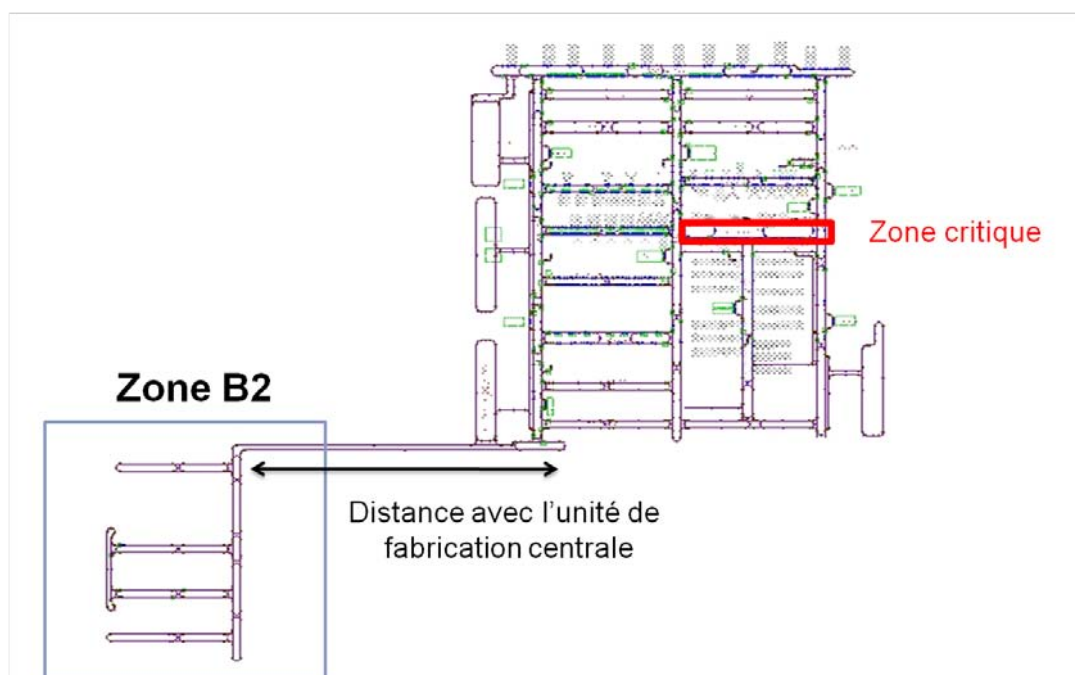
Application

Le projet B2. C’est un projet qui vise à agrandir l’unité de fabrication de Crolles 300mm. Les deux principales raisons sont :

- la réduction de la densité du trafic dans la zone critique étudiée,

- l'augmentation de la capacité de production de l'unité de fabrication.

La problématique. Pour des raisons de disponibilité d'espace, la zone B2 qui va servir à agrandir l'unité de fabrication actuelle sera relativement éloignée de l'unité de fabrication actuelle (temps de trajet estimé à une minute et demi environ). La conséquence est l'allongement des temps de transport. Cela va impacter le temps moyen de cycle et donc diminuer la productivité. La figure 6.12 illustre l'extension des nouveaux rails dans l'unité de fabrication actuelle.



6

FIGURE 6.12 – Illustration des futurs rails.

Nous avons expliqué dans la section 5.5 une limitation majeure de la gestion des véhicules par la politique de « service minimum ». Cette limitation est la possibilité pour une zone d'être vidée de ses véhicules, c'est-à-dire qu'il n'y a plus aucun véhicule dans la zone. Ainsi, il y a un risque que la zone B2 (figure 6.12) se vide. Au vue du temps de trajet estimé, il faudrait un peu plus de 3 minutes pour venir prendre un

lot et l'amener dans l'unité de fabrication centrale. Ceci est inacceptable car c'est le système de transport qui devient un goulet d'étranglement.

Une solution : les véhicules dédiés. Les résultats obtenus par simulation suggèrent l'utilisation de véhicules dédiés comme une excellente solution pour éviter que la zone ne se vide. De plus, cette future zone aura un transport interne très important du fait de la distance, il faut maximiser l'enchaînement des tâches dans la zone B2 avant de retourner dans l'unité de fabrication réelle.

La demande de modification sera négociée très prochainement chez le constructeur MURATEC.

6.5 Conclusion

Dans ce chapitre nous avons présenté une approche pour optimiser la productivité du système. Notre approche comprend trois étapes :

1. *la détermination de valeurs pas trop mauvaises du Low Water Mark.* Nous avons mis en place un prototype avec une interface graphique. Notre approche est exportable pour n'importe quelle politique de gestion de transport par zones,
2. *l'amélioration de ces valeurs* en faisant varier d'une part le nombre de lots en entrée, et d'autre part le nombre total de véhicules. On a notamment pu déterminer la limite du nombre de lots en entrée qui fait que ce sont les équipements de production qui deviennent limitatifs,
3. *la consolidation de la politique de gestion par « service minimum ».*

Nous avons répondu aux deux questions suivantes :

- *quelle est la valeur optimale du Low Water Mark pour notre configuration ?.* Nous avons montré que la valeur 3 présentait les meilleurs résultats dans tous les cas de figures.
- *Quelle est la solution pour la limitation de la politique de gestion par « service minimum » ?* Nous avons proposé les véhicules dédiés à une zone. Ils assurent la présence permanente de véhicules dans la zone. Leurs présences améliorent l'enchaînement des tâches entre les équipements de production de la zone.

Les véhicules dédiés permettent de conserver la qualité d'un service de transport. Ces résultats nous interpellent sur l'existence d'une borne inférieure et d'une borne supérieure. A partir de quel niveau de densité de transport interne doit-on commencer à mettre des véhicules dédiés (borne inférieure) ? Si on doit mettre des véhicules dédiés, combien sont suffisants afin ne pas ralentir le trafic interne (borne supérieure) ?

CONCLUSION GÉNÉRALE

6.5.1 Conclusion

Dans cette thèse nous avons recherché de nouvelles conditions nécessaires à la mise en adéquation entre ressources de transport et besoins de la production. Notre motivation vient du constat que le système de production est « naturellement » dés-équilibré, c'est-à-dire qu'il y a moins de ressources que de tâches à effectuer.

La complexité du problème (beaucoup de variables, de données, etc.) fait que nous avons utilisé une approche par simulation car cela est mieux adapté qu'une formule. Á cela s'ajoutent les particularités propres à chaque industrie. Dans notre cas, l'unité de fabrication utilise une politique de gestion des véhicules que nous avons nommé « service minimum ». Elle consiste à essayer de toujours conserver un nombre prédéfini de véhicules dans une zone donnée.

Un **prototype** de simulation très précis de l'unité de fabrication a été développé, et validé industriellement. L'écart global toléré entre simulation et réalité, décidé en accord avec la manager est de $\pm 15\%$. Ce prototype a ensuite été utilisé pour la **détermination des valeurs des paramètres critiques** de la zone la plus chargée. L'optimisation de ces paramètres a été effectuée en fonction de deux critères (le nombre de véhicules et le nombre de lots lancés au départ) et de trois indicateurs (temps de cycle, productivité et temps de changement d'un lot sur un équipement). Il en ressort que lorsque l'on réquisitionne 3 véhicules (*Low Water Mark* = 3) dans la zone d'étude, on obtient le meilleur compromis, en fonction des indicateurs considérés. Ce constat est valable autant à l'intérieur qu'à l'extérieur de la zone d'étude. En amont, un **prototype de pré-calcul des paramètres critiques**, basé sur le nombre de transports internes et externes, permet de déterminer des premières

solution acceptables. Nous avons montré que la politique de gestion par « service minimum » est meilleure que l'une des politiques de transport les plus utilisées dans la littérature que nous avons nommé « service classique ». Dans cette politique les véhicules peuvent être appelés n'importe où dans l'unité de fabrication et utilise la politique de dispatching FCFS. Nous avons enfin montré que, dédier des véhicules uniquement à une zone en utilisant la politique de « service minimum » permettait d'obtenir une meilleure productivité (réduction de près de 30%).

L'originalité de cette thèse tient en une modélisation fine de l'unité de fabrication qui prend en compte le transport, la production et le stockage. De plus le système de stockage est modélisé avec les contraintes réelles de capacité de stockage. Soulignons que l'unité de fabrication de Crolles 300 est en pleine extension et fonctionne en mode hybride, c'est-à-dire en coexistence homme-machine.

Le prototype est déjà utilisé par l'industriel à travers différentes études comme le positionnement de nouveaux équipements, l'évaluation de nouvelles capacités de transport et de stockage, etc. À court et moyen termes il servira pour l'étude de l'effet des lots de petites tailles, l'influence des équipements à *batches* sur le système de transport, etc. Aux académiciens, ce prototype permettra de mieux comprendre les réalités industrielles par la prise en compte des contraintes de production. Il permettra aussi d'étayer (ou pas) les hypothèses en fournissant des résultats probants. Enfin, nos résultats ont permis un meilleur retour sur investissement par une augmentation de la productivité, une réduction du temps d'attente des lots par la diminution du temps de cycle ce qui traduit également une meilleure allocation des ressources de transport.

6.5.2 Perspectives

Ce travail de thèse de par son originalité, c'est-à-dire la simulation fine du transport, de la production et du stockage, permet d'élargir le champ de recherche dans le milieu de la fabrication des semi-conducteurs. En effet, les trois thèmes de recherche suivants méritent de notre point de vue un approfondissement comme futurs axes de recherche.

Le premier axe concerne l'**intégration du processus de calcul de priorité des lots**. Des auteurs comme [24], [55] ont tenté de le faire mais avec des logiciels différents et en deux temps. Les résultats n'ont pas été très encourageants car les temps de calculs étaient très longs. Nous pensons qu'il est possible d'intégrer suffisamment de détails, mais pas trop, dans un modèle unique prenant en compte le calcul de la priorité des lots. Nous proposons par exemple d'intégrer dans un premier temps, le temps de « séjour » du lot devant une famille de station appelées *Day At Operation*. L'intérêt d'une telle approche est double : pour les académiciens c'est de trouver le meilleur compromis entre niveau de détails et temps de calcul, pour les industriels c'est d'augmenter le niveau de confiance aux modèles de simulation.

Le second axe concerne **la gestion de la densité du trafic**. La densité du trafic représente le nombre de véhicules sur une portion de rails. Elle est donnée par la formule suivante : densité = Nombre de track out/longueur rail. Le nombre de *track out* représente le nombre de lots traités par un équipement. Cette formule montre que plus la longueur est petite plus la densité est élevée. L'intérêt d'étudier la densité est motivée par deux points majeurs. Le regroupement des lots en batchs et la diminution de la taille des lots. Dans [51], l'auteur a montré que l'utilisation de lots de petites tailles, inférieurs à 25 wafers, augmentera le nombre de transports (augmentation du nombre de *tracks out*). De même, à la fin du traitement d'un ou de plusieurs batchs de lots, énormément de ressources de transport seront sollicitées. Le problème de densité est et sera un problème à gérer pour les unités de fabrication utilisant les véhicules unitaires.

Le troisième et dernier axe de recherche que nous proposons concerne la **gestion du regroupement d'emplacements unitaires de stockage**. De par leur temps d'accès rapide, près de 6 fois plus rapides qu'un stockeur, l'utilisation des stockeurs unitaires permet de réduire considérablement le temps d'attente des lots. Deux points importants sont à prendre en compte. Il y a d'abord d'un problème de capacité. Il s'agit de déterminer s'il y a suffisamment d'emplacements unitaires de stockage dans le groupe. Ensuite, il y a un problème de positionnement géographique : il s'agit de déterminer à quel endroit choisir le groupe afin qu'il soit le plus proche possible de

tous les équipements auxquels il est dédié.

Annexe A

A.1 Écart type des équipements internes en fonction des véhicules.

Le tableau [A.1](#) représente les écarts-types du CET moyen exprimés en pourcentage pour les équipements internes.

À LWM constant, l'écart maximum varie entre 4 à 16%. L'écart type du CET moyen décroît lorsque l'on augmente le nombre de véhicules. Il y a moins de dispersion car plus de véhicules disponible pour venir en support dans la zone. Les scénarios de chargement et de déchargement des lots sont de mieux en mieux assurés.

À nombre de véhicules constant, l'écart maximum varie entre 8 à 18%. L'écart type du CET moyen décroît lorsque le Low Water Mark augmente. L'augmentation du LWM réduit la variabilité du CET en réquisitionnant les nombres nécessaire de véhicules dont la zone à besoin à l'intérieur de la zone.

	LWM	Nombre de véhicules					
		22	23	24	25	26	27
Equipement 1	1	36	42	37	34	29	29
	2	40	36	29	31	27	25
	3	32	31	24	23	20	22
	4	28	27	25	25	21	22
Equipement 2	1	35	42	36	34	31	29
	2	40	37	29	30	27	25
	3	33	31	24	24	21	22
	4	27	27	26	25	21	23
Equipement 3	1	29	35	28	27	24	22
	2	36	32	25	24	22	19
	3	28	26	19	17	16	18
	4	23	21	20	20	17	19
Equipement 4	1	36	42	35	33	30	27
	2	41	36	29	29	26	24
	3	33	30	23	22	20	21
	4	28	25	25	23	21	22
Equipement 5	1	46	51	43	42	37	36
	2	45	42	34	33	30	27
	3	35	33	27	27	23	24
	4	34	31	28	24	21	23
Equipement 6	1	35	38	33	29	30	29
	2	32	30	25	26	22	22
	3	21	22	22	22	21	19
	4	26	28	21	16	14	14
Equipement 7	1	38	36	33	32	29	29
	2	31	28	26	25	23	20
	3	19	23	24	22	22	18
	4	29	26	22	15	14	14

TABLE A.1 – Écart type *Carrier Exchange Time* moyen en fonction des véhicules [équipements internes].

A.2 Écart type des équipements externes en fonction des véhicules.

A.3. ÉCART TYPE DES ÉQUIPEMENTS INTERNES EN FONCTION DE LA CHARGE DES LOTS

	LWM	Nombre de véhicules					
		22	23	24	25	26	27
Equipement 8	1	80	71	47	35	49	29
	2	69	69	71	40	32	25
	3	56	50	27	22	19	30
	4	66	34	34	31	41	35
Equipement 9	1	91	74	54	42	48	27
	2	78	76	65	52	37	30
	3	60	56	28	23	22	28
	4	71	34	38	32	37	47
Equipement 10	1	12	12	10	10	9	9
	2	14	12	10	10	9	9
	3	10	11	9	9	8	7
	4	11	11	10	8	7	7
Equipement 11	1	14	12	10	10	10	9
	2	17	14	11	11	10	10
	3	19	14	12	12	11	10
	4	17	16	13	13	11	10
Equipement 12	1	12	12	10	10	9	9
	2	13	12	10	10	9	9
	3	9	11	9	9	8	7
	4	11	11	10	8	7	7
Equipement 13	1	13	11	9	9	8	8
	2	14	12	11	9	9	9
	3	12	12	9	9	9	7
	4	11	11	12	8	8	8

TABLE A.2 – Écart type du *Carrier Exchange Time* moyen en fonction des véhicules [équipements **externes**].

A.3 Écart type des équipements *internes* en fonction de la charge des lots.

Le tableau A.1 représente les écarts-types du CET moyen exprimés en pourcentage pour les équipements internes.

À LWM constant, l'écart maximum varie entre 4 à 16%. L'écart type du

CET moyen décroît lorsque l'on augmente le nombre de véhicules. Il y a moins de dispersion car plus de véhicules disponibles pour venir en support dans la zone. Les scénarios de chargement et de déchargement des lots sont de mieux en mieux assurés.

À nombre de véhicules constant, l'écart maximum varie entre 8 à 18%. L'écart type du CET moyen décroît lorsque le Low Water Mark augmente. L'augmentation du LWM réduit la variabilité du CET en réquisitionnant les nombres nécessaire de véhicules dont la zone à besoin à l'intérieur de la zone.

A.3. ÉCART TYPE DES ÉQUIPEMENTS INTERNES EN FONCTION DE LA CHARGE DES L

	LWM	Nombres de wafers par semaine			
		X	X + 11%	X + 17%	X + 28%
Equipement 1	1	36	36	36	37
	2	33	33	35	33
	3	31	33	33	32
	4	32	31	32	32
Equipement 2	1	35	36	35	37
	2	32	32	33	31
	3	30	32	31	32
	4	31	31	31	31
Equipement 3	1	37	38	38	39
	2	34	34	35	36
	3	34	35	34	36
	4	35	35	36	36
Equipement 4	1	36	37	37	37
	2	33	32	32	34
	3	31	33	32	33
	4	32	32	32	32
Equipement 5	1	45	48	47	47
	2	39	39	40	41
	3	37	39	38	38
	4	37	37	38	37
Equipement 6	1	34	38	37	36
	2	28	29	29	30
	3	26	27	27	28
	4	24	25	26	24
Equipement 7	1	36	40	39	38
	2	32	31	32	32
	3	28	31	30	30
	4	26	27	30	27

TABLE A.3 – Écart type du *Carrier Exchange Time* en fonction de la charge des wafers en entrée [équipements **internes**].

A.4 Écart type des équipements *externes* en fonction de la charge des lots.

	LWM	Nombres de wafers par semaine			
		X	X + 11%	X + 17%	X + 28%
Equipement 8	1	51	55	79	56
	2	54	52	51	71
	3	49	53	52	56
	4	52	54	54	60
Equipement 9	1	54	61	75	60
	2	59	57	59	74
	3	54	59	59	60
	4	58	58	60	56
Equipement 10	1	20	22	22	22
	2	21	22	22	23
	3	21	22	23	23
	4	21	22	23	23
Equipement 11	1	27	21	21	21
	2	21	21	22	21
	3	22	22	22	23
	4	24	23	24	25
Equipement 12	1	24	25	26	27
	2	25	26	27	26
	3	27	27	27	28
	4	27	29	28	29
Equipement 13	1	26	28	27	28
	2	27	29	28	30
	3	29	31	30	31
	4	32	33	31	32

TABLE A.4 – Écart type du *Carrier Exchange Time* moyen en fonction de la charge des wafers en entrée [équipements **externes**].

Bibliographie

- [1] ADAMS, W. P., AND JOHNSON, T. A. Improved linear programming-based lower bounds for the quadratic assignment problem. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science 16* (1994), 43–75.
- [2] ADUSUMILLI, K. M., AND WRIGHT, R. L. Comparative factory analysis of standard FOUP capacities. In *Proceedings of the 2004 Winter Simulation Conference* (2007), vol. 2, pp. 1930–1934.
- [3] AGRAWAL, G. K., AND HERAGU, S. S. A survey of automated material handling systems in 300-mm semiconductor fabs. *IEEE Transactions on Semiconductor Manufacturing 19*, 1 (2006), 112–120.
- [4] ALBERT, P. *Pilotage de véhicules autoguidés dans un système automatisé de production*. PhD thesis, Université de Nantes, 1998.
- [5] ANH, T. L. *Intelligent Control of Vehicle-Based Internal Transport*. PhD thesis, University of Rotterdam, 2005.
- [6] AUTOMOD, VERSION 11. Getting Started With AutoMod. *Applied Materials, Inc* (2004).
- [7] BAHRI, N., REISS, J., AND DOHERTY, B. A Comparison of Unified vs. Segregated Automated Material Handling Systems for 300mm Fabs. In *Proceedings of the 2001 IEEE International Symposium on Semiconductor Manufacturing* (2001), pp. 3–6.
- [8] BANKS, J., AND NORMAN, N. Justifying Simulation Today’s Manufacturing Environment. *IIE Solutions, November* (1995).

- [9] BARKETAU, M. S., CHENG, T. C. E., NG, C. T., KOTOV, V., AND KOVALYOV, M. Batch Scheduling of Step Deteriorating Jobs. *Journal of Scheduling* 11, 1 (2008), 17–28.
- [10] BESCHORNER, A., AND GLUER, D. MaxFlow theory for availability calculation of automated material handling systems. *Robotics Computer Integrated Manufacturing* 19, 1-2 (2003), 141–145.
- [11] BIOSIM. La modélisation au service de la vie, de la science et de la technologie. Biosim [Online] <http://www.biosim.be/alternative/fr/modeles.htm>, 2005.
- [12] BLACKSTONE JR, J. H., PHILLIPS, D. T., AND HOGG, D. L. A state-of-the-art survey of dispatching rules for manufacturing jobshop operations. *International Journal of Production and Research* 25 (1982), 1143–1156.
- [13] BRAIN, M. D. An automation route to upgrading fabs towards 300mm. *Solide State Technology* (2000), 69–76.
- [14] BRAIN, M. D., AND LIN, P. Impact of speed on the performance of vehicle-free transport. In *Proceedings of the 2007 International Symposium on Semiconductor Manufacturing (ISSM)* (2007), pp. 1–3.
- [15] CAMPBELL, L. P., ROHAN, D., AND MACNAIR, A. E. A model of a 300mm wafer fabrication line. In *Proceedings of the 1999 Winter Simulation Conference* (1999), pp. 909–911.
- [16] CAMPEN, V. *Design of a multi-process multi-product wafer fab*. PhD thesis, Technische Universiteit Eindhoven, The Netherlands, 2001.
- [17] CARDARELLI, G., AND PELAGAGGE, P. M. Simulation Tool for Design and Management Optimization of Automated Interbay Material Handling and Storage Systems for Large Wafer. *IEEE Transactions On Semiconductor Manufacturing* 8, 1 (1995), 44–49.
- [18] CHASE, M. 300 mm begins full-fab automation challenges. *Semiconduct. Int.* (2000), 152.
- [19] CHRISOS, J., AND NESTEL-PATT, J. Integration Risks In 300-mm Wafer Fab Automation. *Solid State Technology* 41, 7 (1998), 196–202.

- [20] DA-YIN, L., AND HSIEN-SHENG, F. Dynamic OHT Allocation And Dispatching In Large-Scaled 300mm AMHS Management. *Robotics And Automation* 4 (2002), 3630–3635.
- [21] DA-YIN, L., AND HSIEN-SHENG, F. Speedy Delivery - Dynamic OHT allocation And Dispatching In Large-Scale 300mm AMHS Management. *Robotics And Automation* 11, 3 (2004), 22–23.
- [22] DONG-YEH, J., CHEN-CHUNG, Y., YEAUN-JYH, S., AND JORSUN, Y. A new approach to improve transporting efficiency of amhs. *Semiconductor Manufacturing Technology workshop* (2002), 115–118.
- [23] FUKUNARI, M., RAJANNA, S., J., G. R., AND M., S. E. Data-Based Node Penalties In A Path-Finding Algorithm In An Automated Material Handling system. In *Proceedings of the 2002 Winter Simulation Conference* (2002), vol. 2, pp. 1383–1386.
- [24] GAN, B. P., CHAN, L. P., AND TURNER, S. J. Interoperating simulation of automatic material handling systems and manufacturing process. In *Proceedings of the 2006 Winter Simulation Conference* (2006), pp. 1129–1135.
- [25] GARTLAND, K., AND SHERWOOD, E. Automation needs reach new levels. *Semiconductor International* (2000), 157–163.
- [26] GAXIOLA, G., AND HENNESSY, L. Evaluation of advantages of integrating 300mm AMHS fab layout in the photo area. *IEEE International Symposium of Semiconductor Manufacturing, ISSM* (2001), 373–376.
- [27] GERSHWIN, S. B. *Manufacturing Systems Engineering*. Prentice Hall, 1993.
- [28] GHISELLI, J. 2001 :the make or break year for 300mmm wafer transition goals. *Solid State Technology* 44, 5 (2001), 173–174.
- [29] GILMORE, P. C. Optimal and suboptimal algorithms for the quadratic assignment problem. *Journal of the Society for Industrial and Applied Mathematics* 10, 2 (1962), 305–313.
- [30] GREENBERG, A. L. Next generation factory challenges and opportunities. In *Proceedings of 2007 Advanced Semiconductor Manufacturing Conference* (2007). Keynote Presentation.

- [31] GROOVER, M. P. *Fundamentals of modern manufacturing : materials, process, and systems*. 2006. 3rd Edition.
- [32] HERAGU, S. S. *Facilities Design*. PWS Publishing Company, 1997.
- [33] HIDETO, G. Flexible manufacturing in next generation factory. NEC Electronics Confidential [Online] <http://www.sematech.org/videos/SemiconWest06/p039160.pdf>, 2006.
- [34] HOCHBAUM, D. S., AND B., S. D. A best possible heuristic for the k-center problem. *Mathematic of Operations Research* 10, 2 (1985), 180–184.
- [35] IKEDA, T., HSU, M., AND IMAI, H. A fast algorithm for finding better routes by ai search techniques. In *Proceedings of 1994 Vehicle Navigation & Information Systems Conference* (1994), pp. 291–296.
- [36] JANG, Y. J., CHOI, G. H., AND KIM, S. I. Modeling and analysis of stocker system in semiconductor and LCD fab. *International Symposium on Semiconductor Manufacturing* (2005), 273–276.
- [37] JIMENEZ, J. A., MACKULAK, G. T., AND FOWLER, J. W. Levels of Capacity and Material Handling System Modeling for Factory Integration Decision Making in Semiconductor Wafer Fabs. *IEE Transactions on semiconductor manufacturing* 21, 4 (2008), 600–613.
- [38] KANUNGO, T., MOUNT, D., NETANYAHU, N. S., PIATKO, C. D., SILVERMAN, R., AND WU, A. Y. An efficient k-means clustering algorithm : Analysis and implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 7 (2002), 881–892.
- [39] KIBA, J.-E., DAUZERE-PERES, S., YUGMA, C., AND LAMIABLE, G. Simulation of a full 300mm semiconductor manufacturing plant with material handling constraints. In *Proceedings of the 2009 International Conference on Modeling and Analysis of Semiconductor Manufacturing (MASM)* (2009), pp. 1601–1609.
- [40] KIBA, J.-E., DAUZERE-PERES, S., YUGMA, C., AND LAMIABLE, G. Comparing transport policy in a full-scaled 300mm wafers manufacturing facility. *Progress In Material Handling Research* (2010), 1–12.
- [41] KIBA, J.-E., DAUZERE-PERES, S., YUGMA, C., AND VERMARIEN, L. Gestion optimisée des véhicules dans un large système de transport unifié pour

- la fabrication des semi-conducteurs. In *Proceedings of the 7ième Conférence Internationale de Modélisation et Simulation (MOSIM 08)* (2008).
- [42] KIBA, J.-E., DAUZERE-PERES, S., YUGMA, C., AND VERMARIEN, L. Improve vehicle management for large unified AMHS in semiconductor manufacturing facilities. *Progress In Material Handling Research* (2008), 745–750.
- [43] KIM, B., OH, S., SHIN, J., JUNG, M., CHAE, J., AND LEE, S. Effectiveness of vehicle reassignment in a large scale overhead hoist transport system. *International Journal of Production Research* 45, 4 (2007), 789–802.
- [44] KIM, B., SHIN, J., JEONG, S., AND KOO, J. Effective overhead hoist transport dispatching based on the hungarian algorithm for a large semiconductor fab. *International Journal of Production Research* 47, 10 (2008), 2823–2834.
- [45] KIM, B. I., KIM, S., AND SAHOO, S. Waste collection vehicle routing problem with time windows. *Computers and Operations Research* 33, 12 (2006), 3624–3642.
- [46] KIM, B. I., AND PARK, J. Idle vehicle circulation policies in a semiconductor FAB. *Journal of Intelligent Manufacturing* 20, 6 (2009), 709–717.
- [47] KIM, B.-I., SHIN, J., AND CHAE, J. Simple blocking prevention for bay type path-based automated material handling systems. *International Journal of Advanced Manufacturing and Technology* 44 (2009), 809–816.
- [48] KIM, D. I. The Evolution of Automated Material Handling Systems (AMHS) in s. In *In Proceedings of 6th IEEE International Conference on Industrial Informatics*.
- [49] KIM, Y.-D. A comparison of dispatching rules for job shops with multiple identical jobs and alternative routings. *International Journal of Production and Research* 28 (1990), 953–962.
- [50] KIM, Y.-D., KIM, J.-G., CHOI, B., AND KIM, H.-U. Production scheduling in a semiconductor wafer fabrication facility producing multiple product types with distinct due dates. *IEEE Transaction On Robotics And Automation* 17, 5 (2001), 589–598.

- [51] KONDO, H. Impacts of Small Lot Manufacturing on AMHS. In *Proceedings of the 2008 IEEE International Symposium on Semiconductor Manufacturing* (2008).
- [52] KONDO, H. Requirement for AMHS in 450mm era. In *Proceedings of the 2008 IEEE International Symposium on Semiconductor Manufacturing* (2008), pp. 189–192.
- [53] KONDO, H., AND HARADA, M. Study for realizing effective direct tool-to-tool delivery. In *Proceedings of the 2005 IEEE International Symposium of Semi-conductor Manufacturing* (2005), pp. 21–24.
- [54] KONDO, H., AND HONOLD, A. Timed move command for improved on-time AMHS delivery. In *Proceedings of the 2006 IEEE International Symposium on Semiconducteur Manufacturing* (2006).
- [55] KONG, S. H. Two-step simulation method for automatic material handling system of semiconductor fab. *Robotics and Computer-Integrated Manufacturing* 23, 4 (2007), 409–420.
- [56] KOOPMANS, T. C., AND BECKMANN, M. J. Assignment problems and the location of economic activities. *Econometrica* 25, 1 (1957), 53–76.
- [57] LAMIABLE, G. AMHS transport management. *STMicroelectronics, Crolles* (2008), 1–12. Présentation interne STMicroelectronics.
- [58] LAW, A. M. How to build valid and credible simulation models. In *Proceedings of the 2009 Winter Simulation Conference* (2009), pp. 24–33.
- [59] LI, Z., UZAM, M., AND ZHOU, M. Deadlock control of concurrent manufacturing processes sharing finite resources. *International Journal of Advanced Manufacturing Technology* 38 (2008), 787–800.
- [60] LIN, J. T., WANG, F. K., AND WU, C. K. Connecting transport AMHS in a wafer fab. *International Journal of Production Research* 41, 3 (2003), 529–544.
- [61] LITTLE, J. D. C. A proof for the queuing formula : $l = \lambda w$. *Operation Research* 9, 3 (1961), 383–387.
- [62] LUO, M., ZHANG, J. B., WONG, M. M., ZHUANG, L. Q., NG, K. C., AND AUNG, S. W. A heuristic approach to auto-recovery in the supervisory control

- of heterogeneous automatic equipment for material handling. In *Proceedings of the 31th Annual Conference of the IEEE Industrial Electronics Society (IECON'05)* (2005), pp. 2325–2330.
- [63] MACKULAK, G. T., LAWRENCE, F. P., AND J., R. Simulation analysis of 300mm intrabay automation vehicle capacity alternatives. In *Proceedings of the 1998 IEEE/SEMI Advanced Semiconductor Manufacturing Conference* (1998), pp. 445–450.
- [64] MACKULAK, G. T., AND SAVORY, P. A simulation-based experiment for comparing amhs performance in a semiconductor fabrication facility. *IEEE Transaction on Semiconductor Manufacturing* 14, 3 (2001), 180–206.
- [65] MACQUEEN, J. B. Some methods for classification and analysis of multivariate observations. *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability* (1967), 281–297.
- [66] MARTIN, D. P. Total operational efficiency (TOE) : The determination of 2 capacity and cycle time components and their relationship to productivity improvements in a semiconductor manufacturing line. In *Proceedings of the 1999 IEEE/SEMI Advanced Semiconductor Manufacturing Conference* (1999), pp. 37–41.
- [67] MEYERSDORF, D., AND TAGHIZADEH, A. Fab layout design methodology : Case of the 300mm fabs. *Semiconductor International* 21, 8 (1998), 187–196.
- [68] MONCH, L., BALASUBRAMANIAN, H., FOWLER, W., AND PFUND, E. Heuristic scheduling of jobs on parallel batch machines with incompatible job families and unequal ready times. *Computers & Operations Research* 32, 11 (2005), 2731–2750.
- [69] MONTOYA-TORRES, J. R. A literature survey on the design approaches and operational issues of automated wafer-transport systems for wafer fabs. *Production Planning & Control* 17, 7 (2005), 648–663.
- [70] MONTOYA-TORRES, J. R. *Transport Automatisé dans les systèmes de fabrication de semi-conducteurs*. PhD thesis, Ecole Nationale Supérieure des Mines de Saint-Etienne, 2005.

- [71] MONTOYA-TORRES, J. R. Manufacturing performance evaluation in wafer semiconductor factories. *International Journal of Productivity and Performance Management* 55, 3/4 (2006), 300–310.
- [72] MONTOYA-TORRES, J. R., DAUZÈRE-PÉRÈS, S., AND VERMARIEN, L. A consistent approach for vehicle planning and control in large unified automated material handling systems. *Progress In Material Handling Research* (2006).
- [73] MOORE, G. Cramming more components onto integrated circuits. *Electronics* 38, 8 (1965), 114–117.
- [74] NAZZAL, D., AND MCGINNIS, L. F. Expected response times for closed-loop multivehicle amhs. *IEEE Transactions on Automation Science and Engineering* 4, 4 (2007), 533–542.
- [75] PANWALKAR, S. S., AND ISKANDER, W. A survey of scheduling rules. *Operational Research* 25, 1 (1977), 45–61.
- [76] PAPROTNY, I., SHIAU, J.-Y., HUH, Y., AND MACKULAK, G. Simulation based comparison of semiconductor AMHS alternatives : Continuous flow vs. overhead monorail. In *Proceedings of the 2000 Winter Simulation Conference* (2000), pp. 1333–1338.
- [77] PILLAI, D. Material handling automation for wafer fabrication facilities. In *Proceedings of the ninth IEEE/CHMT International Electronic Manufacturing Technology Symposium* (1990), pp. 277–286.
- [78] PILLAI, D., EDWARD, L. B., DEMPSEY, J. C., AND YELLIG, E. J. 300-mm full-factory simulation for 90- and 65-nm IC manufacturing. *IEEE Transaction on semiconductor manufacturing* 17, 3 (2004), 292–298.
- [79] PILLAI, D., QUINN, T., KRYDER, K., AND CHARLSON, D. Integration of 300mm fab layouts and material handling automation. In *Proceedings of the 1999 IEEE International Symposium Semiconductor Manufacturing Conference* (1999), pp. 23–26.
- [80] PUZAILOV, E. P., GOLANY, B., AND GUREVICH, A. Developing a 3D layout for wafer fabrication plants. In *Proceedings of the 2005 Semiconductor Manufacturing Conference (ISSM)* (2005), pp. 281–284.

- [81] ROBINSON, J., FOWLER, J., AND NEACY, E. Capacity loss factors in semiconductor manufacturing. Working paper [online] www.fabtime.com/files/CapPlan.pdf, 2003.
- [82] SAHNI, S., AND GOZALEZ, T. P-Complete approximation problems. *Journal of the Association of Computing Machinery* 23, 3 (1976), 555–565.
- [83] SAUNIER, C. L'industrie de la microélectronique : reprendre l'offensive. Rapport n°147.
- [84] SCHMIDT, K., AND ROSE, O. Simulation analysis of semiconductor manufacturing with small lot size and batch tool replacements. In *Proceedings of the 2008 Winter Simulation Conference* (2008), pp. 2118–2126.
- [85] SEMICON, E. *Geneva* (1993).
- [86] SHAH, V., ENGLHARDT, E., KOSHTI, S., AND ARMER, H. Metrology delay time reduction in lithography with an enhanced AMHS using local FOUP buffering. In *Proceedings of the 2006 IEEE Advanced Semiconductor Manufacturing Conference* (2006), pp. 120–125.
- [87] STEEL, J., AND BISWAS, T. Development of a shock & vibration spec for 300mm wafer AMHS handling. In *Proceedings of the 2006 IEEE Advanced Semiconductor Manufacturing Conference* (2006), pp. 245–250.
- [88] TAUSCH, F., AND HENNESSEY, L. Evaluation and comparison of a car-based vs a CFT material handling system for a 300mm fab. In *Proceedings of the 2002 International Conference on Modeling and Analysis of Semiconductor Manufacturing (MASM)* (2002), pp. 235–239.
- [89] TREBILCOCK, B. Overhead handling. [Online] <http://www.mmh.com>, 2004.
- [90] TYAN, J. C., DU, T. C., CHEN, J. C., AND CHANG, I. H. Multiple response optimization in a fully automated FAB : an integrated tool and vehicle dispatching strategy. *Computers and Industrial Engineering* 46, 1 (2004), 121–139.
- [91] UZAM, M., LI, Z., AND ZHOO, M. Identification and elimination of redundant control places in petri net based liveness enforcing supervisors. *International Journal of Advanced Manufacturing Technology* 35, 1-2 (2007), 150–168.

- [92] WAKABAYASHI, T., WATANABE, S., KOBAYASHI, Y., OKABE, T., AND KOIKE, A. High-speed AMHS and its operation methods for 300mm QTAT fab. *IEEE Transactions on Semiconductor Manufacturing* 17, 3 (2004), 317–323.
- [93] WIETHOFF, T., AND SWEARINGEN, C. AMHS software solutions to increase manufacturing system performance. In *Proceedings of the 2006 IEEE/SEMI Advanced Semiconductor Manufacturing Conference* (2006).
- [94] WRIGHT, D., AND NORMA, R. Commons AMHS myths. In *Proceedings of the 2005 IEEE International Symposium on Semiconductor Manufacturing* (2005), pp. 277–280.
- [95] ZHAO, H., HU, F., AND JUN LI, G. Batch scheduling with due window on a single machine. *Lecture notes in computer science* 4223 (2006), 641–645.
- [96] ZIMMERSACKL, O., ROTHE, J., SCHMIDT, K., MARSHALL, L., AND HONOLD, A. The effects of small lot manufacturing on amhs operation and equipment front-end design. In *Proceedings of the 2007 international Symposium on Semiconductor Manufacturing* (2007), pp. 1–5.
- [97] ZIMMERMANN, J., MONCH, L., MASON, S., AND FOWLER, J. W. Determining an appropriate number of froups in semiconductor wafer fabrication facilities. In *Proceedings of the 2008 Winter Simulation Conference* (2008), pp. 2164–2170.

**École Nationale Supérieure des Mines
de Saint-Étienne**

N° d'ordre : **2010 EMSE 0582**

Téwendé Jean-Etienne KIBA

Dissertation title

Simulating and Optimizing Automated Transportation in Semiconductor Manufacturing

Major

Industrial Engineering

Keywords

Automated Material Handling System (AMHS), Modeling, Discrete-Event Simulation, Semiconductor Manufacturing

Abstract

Essential components of all computers, semiconductors, support of Integrated Circuits (ICs), are widely used in many sectors (telecommunication, automotive, etc.). Quick advances in these technologies force a challenging rhythm to manufacturers. Following the Moore's Law, chip capacity doubles approximately every two years and prices are divided by two. Thus, the increase of the diameter of wafers to 300 mm makes them heavier to transport in lots. Automated Material Handling Systems (AMHS) are an excellent alternative to tackle this problem. Because of the high price of production equipment, the goal is to use production capacity as much as possible while avoiding to keep too much inventory. To reach this goal, it is important that the transportation and storage system ensures that machines do not wait for available lots, and also that lots do not wait for available machines.

In the literature, due to the problem complexity (up to 700 steps, re-entrant flows, etc.), the modeling of transportation is made with strong assumptions on the production or vice-versa. However, transport is a service for production. This thesis aims at providing a modeling approach that allows the integration in details transport, production and storage constraints, in order to improve production key indicators. A detailed understanding of the system is first required. This allows us to build of a discrete event simulation model which is, in our point of view, the best compromise between the necessary level of details and computational times. Finally, through industrial validations, the complete study of the model allows critical parameters of transport management to be analyzed. The results help to get a better understanding of the system and open interesting research perspectives.

**École Nationale Supérieure des Mines
de Saint-Étienne**

N° d'ordre : **2010 EMSE 0582**

Téwendé Jean-Etienne KIBA

Titre de la thèse

Simulation et optimisation du transport automatisé dans la fabrication de semi-conducteurs

Spécialité

Génie Industriel

Mots clefs

Système automatisé de transport (AMHS), modélisation, simulation à événements discrets, fabrication de semi-conducteurs.

Résumé

Composants essentiels de tout ordinateur, les semi-conducteurs, supports de fabrication des circuits intégrés, sont utilisés dans de nombreux secteurs (télécommunications, automobile, etc.). Les percées technologiques dans ce domaine imposent un rythme vertigineux aux industriels. Tous les deux ans environ, la capacité des puces est doublée et leur prix est divisé par deux : c'est la loi de Moore. Le diamètre des plaquettes de silicium augmente et, regroupées en lots, les plaquettes sont plus lourdes à transporter. Les systèmes automatiques de transport (AMHS) se présentent comme une excellente alternative. Le prix très élevé des équipements de production fait que l'objectif est de saturer les capacités de production. Pour y parvenir, il est important que le système de transport et de stockage garantisse que les machines n'attendent pas des lots disponibles, et aussi que les lots n'attendent pas une machine disponible.

Dans la littérature, la complexité du problème (jusqu'à 700 étapes de fabrication, flux réentrants, etc.) fait que les études de modélisation du transport se font avec de fortes hypothèses sur la production ou inversement. Pourtant, le transport est un service pour la production. Cette thèse propose une approche de modélisation permettant d'intégrer le plus fidèlement possible les contraintes de transport, production et stockage, afin d'améliorer les indicateurs clés de la production. Une analyse détaillée du système a été nécessaire, afin de construire un modèle de simulation à événements discrets, car nous pensons qu'il propose le meilleur compromis entre niveau de détails du modèle et temps de calcul. Enfin, après une validation industrielle, l'étude complète du modèle a permis d'analyser les paramètres critiques de gestion du transport. Les résultats permettent une meilleure compréhension du système et mettent en exergue d'intéressantes perspectives de recherche.