

Visualiseren van de value en optimaliseren van de flow in een wafer productieomgeving

Thomas Bergé

Promotoren: prof. dr. ir. Dirk Van Goubergen, dr. Rita Vercaemst
Begeleider: ir. Joris April

Masterproef ingediend tot het behalen van de academische graad van

Master in de ingenieurswetenschappen:
bedrijfskundige systeemtechnieken en operationeel onderzoek

Vakgroep Technische bedrijfsvoering
Voorzitter: prof. dr. ir. Hendrik Van Landeghem
Faculteit Ingenieurswetenschappen
Academiejaar 2009-2010



Visualiseren van de value en optimaliseren van de flow in een wafer productieomgeving

Thomas Bergé

Promotoren: prof. dr. ir. Dirk Van Goubergen, dr. Rita Vercaemst
Begeleider: ir. Joris April

Masterproef ingediend tot het behalen van de academische graad van

Master in de ingenieurswetenschappen:
bedrijfskundige systeemtechnieken en operationeel onderzoek

Vakgroep Technische bedrijfsvoering
Voorzitter: prof. dr. ir. Hendrik Van Landeghem
Faculteit Ingenieurswetenschappen
Academiejaar 2009-2010



Voorwoord

Op het einde van het eerste masterjaar hadden we precies gezien wat de naam van de richting aangeeft: enkele technieken uit de bedrijfswereld en een inleiding tot operationeel onderzoek. De nieuwsgierigheid naar de toepasbaarheid hiervan in de praktijk deed mij kiezen voor een masterproef in een bedrijf. ON Semiconductor in Oudenaarde leek ook ideaal om iets bij te leren over de mij onbekende halgeleiderindustrie.

De oorspronkelijke titel luidde 'Het invoeren van Lean in een wafer productieomgeving'. Vanuit ON kwam echter de vraag om mee te werken aan de nieuwe dispatchstrategie die ontwikkeld zou worden. Het was even zoeken om het uiteindelijk onderwerp van de masterproef duidelijk te omlijnen, maar uiteindelijk werd de knoop doorgehakt en een nieuwe titel gekozen. Vooral het tweede deel over de dispatchstrategie was erg interessant en ik ben dan ook blij met mijn keuze voor deze ervaring in ON Semiconductor. Ik wil graag de mensen bedanken die geholpen hebben bij het tot stand komen er van:

prof. dr. ir. Dirk Van Goubergen, mijn promotor bij Universiteit Gent,
ir. Joris April, mijn begeleider bij Universiteit Gent,
dr. Rita Vercaemst, mijn promotor en begeleider bij ON Semiconductor,
Wim Van Rentergem en Johan Vastmans, ingenieurs bij ON Semiconductor,
Geert Raman, teamleader bij ON Semiconductor.

Toelating tot bruikleen

De auteur geeft de toelating deze masterproef voor consultatie beschikbaar te stellen en delen van de masterproef te kopiëren voor persoonlijk gebruik. Elk ander gebruik valt onder de beperkingen van het auteursrecht, in het bijzonder met betrekking tot de verplichting de bron uitdrukkelijk te vermelden bij het aanhalen van resultaten uit deze masterproef.

Datum: 30 mei 2010

Handtekening:

Visualiseren van de value en optimaliseren van de flow in een wafer productieomgeving

Thomas Bergé

Promotoren: prof. dr. ir. Dirk Van Goubergen, dr. Rita Vercaemst
Begeleider: ir. Joris April

Masterproef ingediend tot het behalen van de academische graad van

Master in de ingenieurswetenschappen:
bedrijfskundige systeemtechnieken en operationeel onderzoek

Vakgroep Technische bedrijfsvoering
Voorzitter: prof. dr. ir. Hendrik Van Landeghem
Faculteit Ingenieurswetenschappen
Academiejaar 2009-2010

Samenvatting

Deze masterproef ontwikkelt een onderdeel van de nieuwe dispatchstrategie in ON Semiconductor in Oudenaarde. Eerst wordt value stream mapping toegepast op de grootste familie in de productieruimte. Deze voorstelling wordt met het oog op het tweede deel aangevuld met een nieuwe techniek die de complexe verbindingen tussen de verschillende machinegroepen in kaart brengt: de Equipment Group Map. In het tweede deel wordt de sputtergroep bestudeerd en een dispatchsysteem ontworpen dat de capaciteit van deze bottleneckgroep optimaal benut. Dit systeem is gebaseerd op twee principes. Het eerste zegt dat elke operatie moet toegewezen worden aan een welbepaalde sputter. Het tweede principe bepaalt dat de volgorde van loten op een sputter zodanig georganiseerd moet worden dat noch preventief onderhoud op de sputterkamers, noch onnodige setups de productie onderbreken.

Trefwoorden

Value stream mapping, sputter dispatch, capaciteitsoptimalisatie, wafer productieomgeving

Visualization of the Value and Optimization of the Flow in a wafer production environment

Thomas Bergé

Promotors: prof. dr. ir. Dirk Van Goubergen, dr. Rita Vercaemst
Supervisor: ir. Joris April

Abstract This article gives an overview of research performed in ON Semiconductor in Oudenaarde. It covers Value Stream Mapping in the cleanroom and proposes a new dispatch strategy for the sputter equipment group.

Keywords Value stream mapping, sputter dispatch, capacity optimization, wafer fabrication

I. INTRODUCTION

Developing a new dispatch strategy for a wafer fab is a complex task. The huge amount of re-entrant product flows through the many shared resources make it very difficult to get a clear view of the value stream.

Therefore, in a first stage of the research, a Value Stream Map is created for the biggest product family. This map however doesn't display the interrelationships between the equipment groups, nor shows the detailed loading of the sputters. To meet these shortcomings, a new technique is proposed; The Equipment Group Map.

In a second stage, different sputter issues that can improve the (use of) capacity are discussed. Based on this list, the dispatch strategy is developed around two principles. The first one says that all the operations should be assigned to a well defined sputter, depending on the characteristics of the operations and the configuration of the machine. The second principle involves the sequence of processing on each sputter. The wafers are sputtered in groups that alternate in a way that nor the maintenance of the machine chambers, nor avoidable setups interrupt production.

II. LITERATURE REVIEW

A. Lean Thinking in the semiconductor industry

Since the publication of 'The machine that changed the world' by Womack et al [1], based on the Toyota Production System from around '50 in Japan, Lean Thinking has become a popular subject in literature. Identifying value for the customer, creating flow by establishing a pull-system and the continued/continuous pursuit of perfection at all levels in the company, should enable a company to produce more with less time, less space, less inventory, less money and less errors [1].

In the semiconductor industry, some concepts of Lean Thinking are hot topics while others seem to have difficulties to be brought in practice. The common strategy to make profit is increasing throughput and hence lowering the cost per wafer. Eliminating waste helps to increase capacity and is one of the two focuses in today's fabs [2]. The second is balancing the WIP in

the line by doing the right wafer moves in the right amount.

Although both issues are also important in Lean, big changeovers to a Lean approach in all levels of the company can be considered as long-distance goals.

No literature or cases are found concerning value stream mapping in the semiconductor industry, nor are other attempts to visualize the product flows.

B. Job shop planning in the semiconductor industry

Recently, much research has been set up involving local search algorithms and artificial intelligence techniques to develop a schedule for the wafer fab. Those methods are tested in small and medium size simulation models, but have not found their way to the work floor yet [3].

Nowadays mathematical programming and mainly dispatching rules are used to define the sequence of lots. There is a tendency to prefer combinations of simple local rules such as SPT and critical ratio that are good for fab cycle time and on time delivery, with more complex global rules. Global rules often focus on sending WIP to downstream processes in a way in which all are equally loaded [4].

III. MAPPING THE VALUE STREAM

Although methods exist to combine two or three product families and their corresponding value streams in one map, value stream mapping in a cleanroom restricts the focus to only a very small part of the product mix. Drawing the inventory of the other families before the shared resources meets this problem insufficiently.

Furthermore, creating the current state VSM is not easily done by walking the flow from the finishing sinter process, upstream to the release of new lots in the fab. Although the re-entrant nature of the product routes, the operations need to be drawn in a linear way. The result is a 200-steps long chain of process boxes and inventory triangles. It is hard to see possible improvements and indicate kaizen bursts in this map.

In function of the dispatch strategy to be developed, it is decided to introduce a new way to visualize the flow. The Equipment Group Map shows the different product families moving around one central machine group. It also defines the relationships between the central group and the upstream and downstream equipment groups.

In Figure 1, the dark boxes represent the upstream equipment groups from where wafers flow to the centred machines. The light coloured boxes represent the downstream groups. If only bottleneck stations are displayed in the boxes and the steps in between are represented on the arrows, flow between bottlenecks can

be managed. In the detailed version, each arrow displays also the number of moves, the operation numbers at the centred machine group where it arrives and which product families take that path (including their percentage).

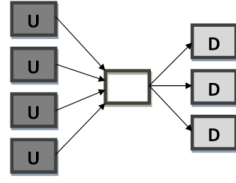


Figure 1: Simplified Equipment Group Map.

IV. DEVELOPING A DISPATCH STRATEGY

The dispatch strategy in the cleanroom in ON Semiconductor Oudenaarde is different, depending on the calculated load for every machine group. The dispatch of equipment groups with smaller utilization time is based on critical ratio and line balancing. The so-called bottlenecks, which are very heavily loaded, focus on optimal capacity utilization. In this article, the sputters, an equipment group of the second kind is investigated.

A. List of measures

To counter loss of capacity, a number of measures can be taken. All have to be taken into account in the dispatch strategy. They are classified in four categories:

- measures that reduce process times of the wafers
- measures that avoid unnecessary setups
- measures that avoid a sputter to run out of work
- measures that organize maintenance to achieve a higher uptime.

B. Principle 1

The first principle says that all operations should be assigned to a well defined sputter based on

- the chamber configuration of the sputter
- the characteristics of the operation

In this way a wafer will always be sputtered on the machine with the lowest process time. Furthermore, wafers with the same setup requirements can be assigned to the same sputter which leads to less setups. A third advantage is that it is possible to assign wafers with complementary chamber combinations. This opens up the possibility to give maintenance to certain chambers while production can continue on the other ones.

C. Operation groups

To simplify the assignment of the near sixty different operations on the four available sputters in ON, they are divided in twelve operation groups. This splitting up is based on three criteria. All operations must have

- the same setup requirements
- the same chamber combinations
- (possible) process speed advantages on the same sputter

Principle one says now that all twelve operation groups should be assigned to one of the four sputters. Figure 2 provides an example of the loading for each sputter, after assigning the operation groups of a random WIP-status. Each blocks represents the sum of the run times of the

wafers that are waiting for the operations in the same operation group.

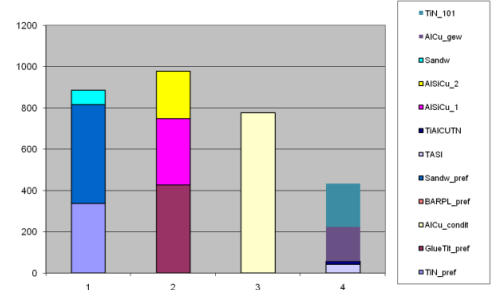


Figure 2: Operation groups, each represented by the run time of the current inventory in the group, assigned to the different sputters.

D. Principle 2

Once the operation groups are assigned to the four sputters, the sequence of the wafers on each sputter has to be decided. Principle two says that wafers that use the same chamber combinations, should be grouped. This means that operation groups should alternate with intervals that are large enough to execute preventive maintenance on chambers that are not used by the currently sputtered group(s).

Since the wafers that require a setup are now combined in the same interval and there is no interruption in the production process during that interval, only one setup is needed for all the wafers in that interval.

At last, the SPC-tests can also be planned now at the beginning of each interval. In that way, the monitor wafers can be placed in the same wafer box as the setup wafer, so time is saved again.

E. Flexibility

Although it is possible to assign operation groups to the sputters in a way that the total amount of process time will be equal for all four of them, variation in the arriving WIP can cause sputters running out of WIP while others are overloaded. To counter this, some well-chosen operation groups are flexible. This means that if a sputter runs low on WIP, an operation group from another machine can be reassigned to the almost empty sputter.

F. Conclusion

By grouping and assigning operations to well defined sputters and processing the wafers in blocks of similar chamber combinations, the lowest possible process time can be achieved and preventive maintenance of the chambers can be scheduled in a way that production is not interrupted for this reason.

REFERENCES

- [1] J.P. Womack, D.T. Jones en D.Roos, *The Machine That Changed the World*, Harper Perrenial, New York, 1991
- [2] C.M. Weber en A.Fayed, *Optimizing your position on the operating curve: How can a fab truly maximize its performance?* IEEE Trans. on Semic. Manuf., 2010, Vol. 23, 1
- [3] A.K. Gupta en A.I. Sivakumar, *Job shop scheduling techniques in semiconductor manufacturing*, Intern. J. of Adv. Manuf. Techn., 2006, Vol.27, 1163–1169
- [4] R.M. Dabbas en J.W. Fowler, *A New Scheduling Approach Using Combined: Dispatching Criteria in Wafer Fabs*, IEEE Trans. on Semic. Manuf., 2003, Vol. 16, 3, 501

INHOUDSTAFEL

HOOFDSTUK 1: VISUALISEREN VAN DE FLOW	1
1. LEAN	1
1.1. HET CONCEPT LEAN	1
1.2. EVOLUTIE VAN LEAN	5
1.3. KRITIEK OP LEAN	6
1.4. LEAN IN DE HALFGELEIDERINDUSTRIE	6
2. VALUE STREAM MAP (VSM)	8
2.1. VSM, EEN TOOL BINNEN LEAN	8
2.2. EVOLUTIE VAN VSM	8
2.3. VSM IN DE SEMICONDUCTORINDUSTRIE	9
3. VALUE STREAM MAP IN FABII BIJ ON SEMICONDUCTOR IN OUDENAARDE	10
3.1. FABII IN ON SEMICONDUCTOR	10
3.2. WAARDE VOOR DE KLANT	14
3.3. INDELING IN PRODUCTFAMILIES	14
3.4. DE CURRENT STATE MAP	18
3.5. UITLEG BIJ DE CURRENT STATE MAP	23
3.5.1. Value-added en not value-added	23
3.5.2. Cycle time en process time	24
3.5.3. Setup / Uptime / Utilization Time / Scrap	24
3.5.4. Voorraad	25
3.5.5. Push arrows en de informatiestroom	27
3.6. CONCLUSIES VSM BACKEND C07I	28
3.6.1. Figuren	28
3.6.2. Lead time en process time	28
3.6.3. Evaluatie	29
4. EQUIPMENT GROUP MAP (EGM)	31

HOOFDSTUK 2: OPTIMALISEREN VAN DE FLOW

35

1. PLANNEN VAN EEN JOB SHOP-OMGEVING	36
1.1. HET PLANNINGSPROBLEEM	36
1.2. KORTE HISTORIEK VAN HET ONDERZOEK NAAR PLANNINGSMETHODES	37
1.3. CLASSIFICATIE VAN PLANNINGSMETHODES	38
1.4. OPLOSSINGSMETHODES EN TECHNIKEN VOOR HET PLANNINGSPROBLEEM	40
2. EVALUEREN VAN DE SPITTERS	46
2.1. DE SPITTERGROEP	46
2.2. UPTIME, BEREKEND DOOR ON	49
2.3. WORK SAMPLING	51
3. OPTIMALISEREN VAN DE SPITTERGROEP	59
3.1. HUIDIG SYSTEEM	59
3.1.1. Technologisch	59
3.1.2. Dispatch	60
3.2. BELANGRIJKE PUNTEN VOOR HET NIEUWE SYSTEEM	62
3.2.1. Maatregelen die de procestijd verkorten.	62
3.2.2. Maatregelen die de tijd voor conditioneren verkorten.	62
3.2.3. Maatregelen die leegloop vermijden.	63
3.2.4. Maatregelen die in verband met onderhoud en technologische beperkingen.	63
3.3. VOORSTEL NIEUW SYSTEEM	64
3.3.1. Operatiegroepen	64
3.3.2. Operatiegroepen toewijzen aan spitters.	68
3.3.3. Plannen van onderhoud, conditionering en monitoring.	71
3.3.4. Dediceren versus flexibiliteit	72
3.3.5. Tijdsgebaseerd plannen	73
3.3.6. Samengevat	74

HOOFDSTUK 3: SAMENVATTING

75

BIJLAGEN	77
REFERENTIES	78

Tabel van afkortingen

APF	Applied Productivity Family (softwarepakket)
BS	Batch Size (batchgrootte)
CT	Cycle Time (cyclustijd)
CO	Change Over (omschakeling)
EG	EquipmentGroup (machinegroep)
EGM	Equipment Group Map
FAB	Fabricatieruimte (productieruimte voor wafers)
FIFO	First-In First-Out
NVA	Not Value Added (processtap die geen waarde toevoegt aan het product)
ON	ON Semiconductor (Oudenaarde)
SPC	Statistical Process Control
UPT	Uptime (tijd beschikbaar voor productie, als percentage van de totale tijd)
UT	Utilization Time (effectieve productietijd, als percentage van de totale tijd)
VSM	Value Stream Map
WIP	Work In Process (gedeeltelijk afgewerkte producten)

Hoofdstuk 1: VISUALISEREN VAN DE FLOW

Het eerste hoofdstuk van deze thesis start met een literatuurstudie. Na een overzicht van het ontstaan en de evolutie van Lean, wordt de toepassing in de halfgeleiderindustrie besproken. Ook al staat deze mijlenver van de automobielsector waarbinnen Lean ontstond, toch zijn hier en daar succesverhalen te vinden. Daarna wordt ingezoomd op de Leantool Value Stream Mapping. In het derde deel van het hoofdstuk wordt deze techniek ook toegepast op de grootste productfamilie van ON Semiconductor in Oudenaarde. Tot slot wordt ook een Equipment Group Map voorgesteld die in hoofdstuk twee verder van nut kan zijn.

1. LEAN

1.1. Het concept Lean

Lean Thinking is tegenwoordig in een heel aantal sectoren doorgedrongen en veel mensen zijn er dan ook al op één of andere manier mee in contact gekomen. Toch is het een relatief recent begrip. De meeste auteurs verwijzen voor de basis van Lean naar het Toyota productiesysteem dat ontstond in de jaren '40 tot '70 in Japan. Sommigen gaan nog een stapje terug en beschouwen Henri Ford met de continue productie van zijn Model T als de eerste Lean denker [1]. Deze productieomgeving was echter eerder een uitzondering; enorm hoog volume, nul variëteit, een eeuwig blijvend model en een vraag die veel groter was dan de capaciteit. Een andere aanpak drong zich op in Japan. Sommige auteurs geven als voornaamste redenen het tekort aan grondstoffen en de sterke concurrentie na de recessie door WO II [1] en [2]. Anderen wijzen als belangrijkste oorzaak op de kleine en sterk uiteenlopende automarkt in Japan alsook het gebrek aan gastarbeiders [3].

In ieder geval waren Kiichiro Toyoda, Taiichi Ohno en anderen bij Toyota van mening dat enkele eenvoudige vernieuwingen voldoende waren om tegelijk een continue productstroom te garanderen alsook een grote variëteit aan producten. In Europa was men ondertussen namelijk overgeschakeld van flow zonder variatie naar variatie zonder flow, met steeds snellere, grotere en vooral duurdere machines. Gelijksortige processen werden apart in grote afdelingen samen gezet waarbinnen dan verschillende varianten mogelijk waren [4]. Maar door het verlies aan continuïteit (flow) en het extra transport explodeerden de lead times en waren grote voorraden nodig. Vele auteurs verwijzen voor deze periode naar General Motors onder leiding van Alfred Sloan.

Bij Toyota was de nieuwe strategie gekend onder de naam *Toyota Production System*. De term Lean Manufacturing werd voor het eerst gebruikt door James Womack, Daniel Jones en Daniel Roos in hun boek *The machine that changed the world*. Deze drie managers van het International Motor Vehicle Program definiëren Lean door het te vergelijken met zijn voorgangers: stukwerk (craftmanship) en massaproductie. Stukwerk wordt uitgevoerd door goed opgeleide specialisten met eenvoudige maar flexibele gereedschappen, om exact te maken waar de klant om vraagt. Dit geeft dus de grootst denkbare variëteit en kwaliteit, maar aan een enorm hoge kost. Massaproductie werkt met dure machines die slechts een beperkt aantal producten kunnen maken en een eenvoudige bediening vergen. De machines produceren grote volumes standaardproducten. Om het productieproces goedkoop (= vlot) te laten verlopen, moeten voorraadbuffers toegevoegd worden en moet het design zo lang mogelijk in productie blijven. Lage prijzen ten koste van de variëteit dus. Lean echter gebruikt flexibele werknemers die eveneens flexibele en sterk geautomatiseerde machines bedienen om grote volumes producten in een enorme variëteit te fabriceren. Het is een strategie die de operationele activiteiten van een bedrijf organiseert en verbetert [5].

Womack et al. omschrijven Lean als “de helft van de tijd, de helft van de ruimte, de helft van de ingenieursinspanning, de helft van de fouten en één tiende van de voorraad” [3]. Veel andere auteurs, zoals Melton [6], definiëren Lean als het constant zoeken naar en elimineren van *waste* (“*the pursuit of perfection*”), waarbij ze *waste* dan weer definiëren als alle processen, inspanningen en grondstoffen die wel kosten maar geen waarde bijbrengen aan het eindproduct. Toyota onderscheidt 7 soorten verspilling (*muda*’s); overproductie, voorraad, beweging, wachten, transport, onnodige processtappen en *scrap* of *rework*. Als achtste verspilling wordt vaak de verspilling van niet gebruikt talent bij de werknemers gezien. Om te bepalen wat waste is en wat niet, is een grondige analyse nodig van wat waarde nu precies is en dit vanuit het perspectief van de klant die het eindproduct koopt. Eens dit vastgelegd, kan een productieproces beschouwd worden als een voertuig dat die waarde naar de klant brengt. Overal waar het voertuig komt, moet waste verwijderd worden [6].

Womack en Jones resumeren de leanfilosofie in hun later gepubliceerde boek “*Lean Thinking: banish waste and create wealth in your organization*” [7] samen in vijf punten die nu algemeen aanvaard en verspreid zijn als een vijfstappenplan om Lean te benaderen. [5],[7] en [8]

1. Definieer wat waarde is vanuit het standpunt van de klant en dit voor elke productfamilie apart. Waarde is door Womack en Jones gedefinieerd als “*capability provided to customer at the right time and at an appropriate price, as defined in each case by the customer*”. En met customer



Figuur 1: De vijf stappen om Lean na te streven [8].

wordt hier de klant van het eindproduct bedoeld en niet de interne klanten zoals marketing, sales, distribution, suppliers, het volgende proces, ...

2. Ga na wat de stappen zijn in de *value stream* voor elke productfamilie en elimineer alle stappen die geen waarde creëren. De value stream of waardeestroom zijn alle stappen nodig om een product te ontwerpen, produceren en leveren van het ontwerp via de grondstoffen tot in de handen van de klant. In de value stream bevinden zich drie soorten activiteiten: activiteiten die waarde toevoegen aan het eindproduct (value-added), activiteiten die geen waarde toevoegen maar noodzakelijk zijn om het product te vervaardigen (type 1 verspilling) en activiteiten die overbodig zijn (type 2 verspilling).
3. Laat dan de stappen die waarde toevoegen kort en vlot op elkaar volgen zodat de producten zonder tussenstoppen, zonder fouten en zonder loops naar de handen van de klanten stromen. Dit in tegenstelling tot een systeem waar tussen alle processen tussentijdse voorraad ligt.
4. Laat klanten de producten (waarde) trekken uit stroomopwaarts gelegen processen. Dat wil zeggen dat geen enkele activiteit gestart kan worden zonder dat een proces dichterbij de klant (downstream) daarom gevraagd heeft (*pull*-systeem).
5. Als laatste Lean-principe komt het streven naar perfectie. Het implementeren van Lean is nooit voltooid en heel het bedrijf, van de top tot de werkvloer moet blijven zoeken naar waste in de value stream. De bedoeling is het begrip value concreet te definiëren voor iedereen binnen het bedrijf zodat elke werknemer zijn steentje kan bijdragen in het streven naar verbetering.

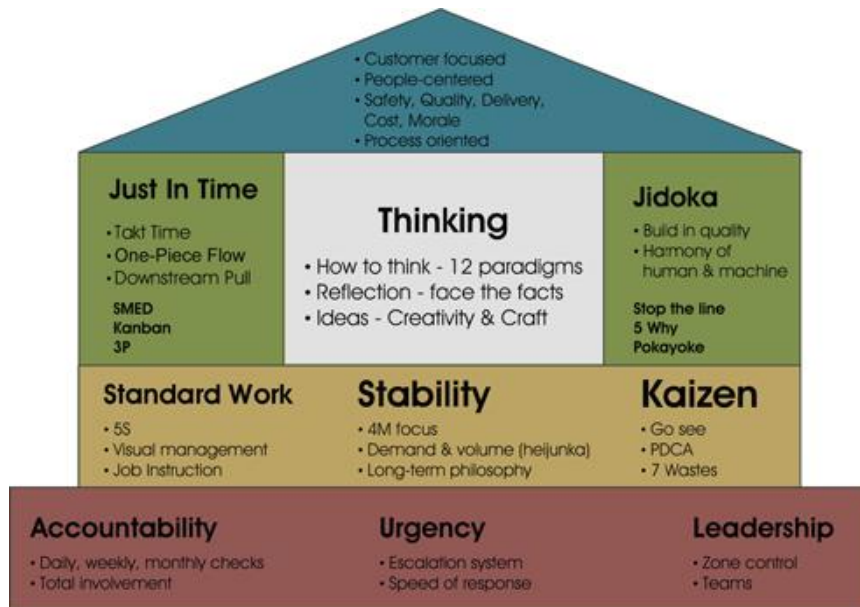
Lean wordt ook vaak voorgesteld met behulp van de “14 principles of the Toyota way” van Liker [1]:

1. Denk als management op lange termijn, ook al gaat dat mogelijk ten koste van korte termijn cijfers.
2. Zorg voor een continue productiestroom zodat problemen direct opgelost moeten worden.
3. Gebruik pull-systemen om overproductie te vermijden.
4. Balanceer de lijn, verdeel de werkbelasting gelijkmatig.
5. Probeer niet steeds achter de feiten aan te lopen, zorg dat de kwaliteit onmiddellijk goed is.
6. De basis van degelijk werk ligt in het standaardiseren van de jobs en de processen.
7. Visualiseer zoveel mogelijk om problemen snel aan het licht te brengen.
8. Zorg voor betrouwbare machines en technologieën.
9. Zorg voor continuïteit in de bedrijfsfilosofie en geef kennis van de processen en de filosofie erachter door aan nieuwe generaties.
10. Tracht een team van werknemers te verzamelen dat achter de bedrijfsfilosofie staat.
11. Denk verder dan je eigen onderneming en werk samen met leveranciers en afnemers. Werk samen om het geheel te verbeteren.
12. Blijf in contact met de werkvloer, weet waarover je spreekt en verifieer je gegevens zelf.

13. Neem de tijd voor beslissingen en onderzoek alle opties. Voer daarna snel en correct uit.

14. Blijf continu leren van fouten door na te denken over waarmee je bezig bent.

Deze 14 principes en een hele lijst aan tools om ze te realiseren en onderhouden worden vaak voorgesteld in een huisvorm, the Toyota Production System House (Figuur 2).



Figuur 2: The TPS House. bron: Gemba Research

Rond het centrale Lean denken dat door de 14 principes wordt weergegeven, staan de pijlers *Just In Time* en *Jidoka*. *Just in Time* staat voor een continue flow van producten en informatie, het verminderen van de WIP en het pull principe. De bekendste technieken zijn hier het *kanbansysteem* dat slechts WIP toelaat indien expliciet gevraagd door een proces downstream en SMED of Single Minute Exchange Die, een techniek die gebruikt wordt om de set-up tijd voor processen zo kort mogelijk te houden. *Jidoka* staat voor ingebouwde kwaliteit. Het systeem wordt zodanig ontworpen dat fouten bijna onmogelijk zijn (het Poka Yoke systeem van Shigeo Shingo) en zowel de werknemer als de machine hebben de autonomie om het proces stil te leggen als er toch één voorkomt. Verder is de betrouwbaarheid van de machine van hoog belang; preventieve maatregelen staan centraal. [8] en [9]

De pijlers zijn gebouwd op gebalanceerd werken en stabiele productie. Verder is het Toyota systeem ook gebaseerd op gestandaardiseerde werkplaatsen die voldoen aan 5S – *Seiri* (organisatie), *Seiton* (ordelijkheid), *Seiso* (netheid), *Seiketsu* (standaardisatie) en *Shitsuke* (discipline). Voor die standaardisatie zijn duidelijke werkomschrijvingen nodig en Visual Management maakt productie en organisatie overzichtelijk. Kaizen ten slotte duidt op constante verbetering en het actief op zoek blijven naar waste in de productieomgeving.

Als concrete en meetbare veranderingen worden vaak volgende zaken genoemd [6]

- Verminderde voorraad
- Minder rework
- Verminderde lead time
- Minder verspilling in de processen
- Beter begrip van de processen en hun aaneenschakeling
- Financiële voordelen

1.2. Evolutie van Lean

Na zijn eerste toepassingen in de automobielandustrie heeft de Leanfilosofie zich niet stormenderhand verspreid over andere sectoren en continenten. Ondanks de vele literatuur en case studies alsook het succesverhaal van Toyota, is het voor een bedrijf namelijk vaak niet evident om de stap te zetten. Onder andere Melton [6] geeft hiervoor twee belangrijke redenen aan. Ten eerste wordt het directe tastbare bewijs niet altijd onderkend en ten tweede verkeren veel bedrijven in de waan dat hun bedrijfsstructuur en –processen al optimaal zijn.

Toch verspreidt Lean zich ondertussen over de hele wereld en de klassieke technieken en principes worden aangepast aan een waaier van sectoren: in de dienstensector [5], in logistiek en distributie [10], in retail [11], in de zorgomgeving [12], in de bouwsector [13 en 14], bij de overheid [15 en 16] en in onderhoud [17]. Volgens het Lean Enterprise Institute zijn in de meeste van deze sectoren de Leanprincipes nog maar net doorgedrongen en is er nog een lange weg te gaan [18]. Volgens Womack is dit voor een groot deel te wijten aan het feit dat we hier geen te heropbouwen industrieën hebben zoals destijds in Japan, maar moeten opboksen tegen een lange traditie van productiefilosofieën [1]. Eén van de bekendste bronnen voor succesverhalen over Lean is het boek *“Becoming Lean”* van Jeffrey Liker. Hij beschrijft een groot aantal cases van Lean-pioniers in de Verenigde Staten en laat ook de managers zelf aan het woord over de eerste stappen van hun implementaties in de meest uiteenlopende sectoren[1].

Behalve Lean zijn er nog andere filosofieën die de laatste jaren onder de aandacht kwamen van het management. Six Sigma heeft tot doel productieprocessen onder controle te houden wat moet leiden tot minder producten met fouten (of kortweg een betere kwaliteit) [19]. Continuous Improvement (CI) is vergelijkbaar met het streven naar perfectie bij Lean, namelijk het idee om continu te leren van gemaakte fouten [20]. Total Quality Management (TQM) tracht dan weer alle werknemers te betrekken bij het verbeteren van de werkomgeving en het oplossen van problemen. Al deze systemen staan duidelijk niet los van Lean en zijn zelfs gemakkelijk assimileerbaar binnen het grote Lean-kader. [20]

1.3. Kritiek op Lean

Kritiek zal er altijd zijn, ook voor het Leansysteem dat door velen geprofileerd wordt als ideaal. Ze kan opgedeeld worden in enkele grote hoofdlijnen.

Eerst en vooral: wordt Lean terecht als beter beschouwd dan andere systemen? Philip Jennings, ex-secretaris-generaal van FIET¹, beweert dat niet noodzakelijk het toepassen van Lean op andere sectoren de moeilijkheid is, maar wel de situatie van waaruit vertrokken wordt. Verder verwijst hij ook naar onderzoek van Berggren [22] dat kritiek levert op het '*job enrichment*' begrip (het systeem waarbij werknemers flexibel zijn doordat ze verschillende taken kunnen uitvoeren). [21]

Verder is er ook kritiek op de werkomstandigheden in de Japanse cultuur. Darius Mehri [23] zegt hierover dat de werknemer inderdaad sterk betrokken wordt bij het bedrijfsgebeuren, maar dat hij daar wel een grote tol voor betaalt. Enorm lange werkuren en grote druk om de hoge snelheid van de machines bij te kunnen houden zouden typisch zijn in bedrijven die volgens de Toyotacultuur proberen te werken.

Als derde en vierde punt worden ook vaak de moeilijke omgang met variabiliteit en lage productvolumes en het gebrek aan strategische perspectief genoemd (Lean zou slechts een opsomming van technieken zijn). [24]

Over zowat al deze punten bestaat er opnieuw literatuur die deze weerlegt. Vooral het laatste punt wordt juist in veel literatuur omschreven als één van de belangrijkste zaken bij het implementeren van Lean. Namelijk dat Lean niet beschouwd mag worden als een aantal tools waarmee je een bedrijf kan optimaliseren, maar wel als een manier van denken. Liker [1] zegt hierover dat er inderdaad niemand zal betwisten dat een slordige, ongeordende werkplaats minder goed functioneert dan een propere en goed gestructureerde. Werken aan kwaliteit, machine betrouwbaarheid en 5S is echter niet hetzelfde als flow en pull installeren en streven naar perfectie. De moeilijkheid is in te zien waarom er soms – voor massaproductie tegennatuurlijke – beslissingen genomen moeten worden zoals het verkleinen van de veiligheidsvoorraad of het kopen van 5 kleine aan één lijn toegewezen machines in plaats van één grote gedeelde machine (shared resource).

1.4. Lean in de halfgeleiderindustrie

Eén van de kenmerken van de halfgeleiderindustrie is het enorm snel dalen in waarde (verkoopprijs) van de producten. Het is daarom belangrijk om de eerste te zijn om nieuwe technologieën op de markt te brengen. Dit vereist behalve efficiënt onderzoek ook een zo kort mogelijke FAB² cycle time³. Hoe korter de cycle time en dus

¹ The international trade union organization for *white collar* and professional workers.

² FAB komt van fabricatieruimte en slaat op de ruimte waarin de wafers gemaakt worden.

³ De tijd die een wafer nodig heeft om de volledige productieomgeving te doorlopen

ook de cycle of learning⁴, hoe sneller een nieuwe technologie kan ontwikkeld worden. Historisch gezien was (en dit is vaak nog steeds het geval) de algemeen aanvaarde manier om winst te maken het maximaliseren van de throughput van de FAB. Met andere woorden het produceren van zoveel mogelijk wafers per tijdseenheid. De FAB cycle time verkorten en de FAB throughput vergroten werkt elkaar echter tegen. De cycle time stijgt namelijk buiten proportie als de throughput zijn theoretisch maximum benadert [25]. [26]

De afstand tussen de huidige halfgeleiderindustrie en een volledig Lean-denkende versie ervan lijkt enorm. Toch worden op bepaalde vlakken inspanningen in die richting gedaan. Zo wordt het elimineren van waste in alle stappen van het proces genoemd als één van de punten die bijdragen aan zowel het verkorten van de cycle time als het besparen van kosten. Dit elimineren van waste focust op [26]

- kwaliteit van de producten (Total Quality Management) met voornamelijk snellere feedback bij kwaliteitsproblemen bij de machines
- het verhogen van de opbrengst door het verminderen van variabiliteit in procesfactoren (noise) [27]
- het verhogen van de beschikbaarheid van machines
- het optimaliseren van het gebruik van bottleneck machines
- het balanceren van de WIP in de FAB (fabricatieruimte voor wafers)

Het grootste probleem om Lean te implementeren in de halfgeleiderindustrie is de moeilijkheid om flow te creëren. Dit komt enerzijds door de complexiteit van de productroutes en anderzijds door de beperkingen op de FAB layout als gevolg van de strenge stofnormen en het gebrek aan verplaatsbaarheid van de machines.

⁴ De tijd die een nieuwe technologie nodig heeft om test te lopen door de FAB.

2. Value Stream Map (VSM)

2.1. VSM, een tool binnen Lean

In de vijf stappen om Lean te benaderen (**Error! Reference source not found.**), kwam het begrip Value Stream al ter sprake. De *value stream* is het geheel van acties (zowel *value added* als *non value added*) die nodig zijn om van basisgrondstoffen tot eindproduct te evolueren. Ook werd reeds gesproken over de sterke focus op visualisatie. De Value Stream in kaart brengen, Value Stream Mapping, is dan ook geen verrassende tool. De VSM legt de link tussen Lean en de supply chain. Het overschrijdt de grenzen van de eigen fabriek en betreft zowel de leveranciers (upstream) als de klanten (downstream).

In een productieomgeving zijn er drie stromen. Een materiaalstroom, een informatiestroom en een stroom van mensen (of processen) [28]. De value stream map visualiseert de eerste twee. Het is een “plan” dat de huidige en toekomstige staat van een productiesysteem beschrijft en toont waar je staat en welke acties je kunt ondernemen om *waste* te elimineren. Het is met andere woorden een ideale basis voor het implementeren van lean principes en technieken [29]. Concreet betekent het dat je het productiepad van één specifiek item volgt van klant tot leverancier en een visuele voorstelling maakt van elke proces in de materiaal- en informatiestroom [28].

Het hele principe van Value Stream Mapping kan beschreven worden in de volgende vijf stappen [28] en [30]:

1. Selecteer een productfamilie.
2. Stel een current-state VSM op.
3. Creëer een future-state VSM met behulp van Lean technieken.
4. Stel een plan op om deze future-state te implementeren. Verwezenlijk dit plan op een continue en gestructureerde manier.

2.2. Evolutie van VSM

Rother en Shook passen Value Stream Mapping in hun boek “*Learning to see*” [28] toe op een eenvoudig voorbeeld. Dit is interessant om kennis te maken met de principes van Lean en Value Stream Mapping. Het lijkt echter absoluut niet op een reële bedrijfssituatie en gaat niet erg in detail. Duggan [31] schrijft hierop een soort vervolg “*Creating mixed model value streams*” waarbij hij een situatie met een grote productmix en complexere materiaalstroom onder de loep neemt. Dit is echter niet voldoende om de kritiek te smoren dat Value Stream Mapping enkel effectief toegepast kan worden op systemen met lineaire productroutes en dat de value stream

best éénrichtingsverkeer is met een minimum aan backtracking en kruisingen tussen de verschillende productfamilies [32].

Naast onderzoek naar en publicaties over specifieke industrieën, zijn er ook algemene aanpassingen en uitbreidingen voorgesteld op de VSM zoals ze door Rother en Shook beschreven wordt. McDonald et al. [33] paste VSM toe op drie parallelle lijnen die een assemblageproces voorstelden. Hij gebruikte discrete-event simulation om de sleutelprocessen in het productieproces te ontdekken. Lian en Van Landeghem [34] bouwden ook simulatiemodellen voor een systeem met twee parallelle lijnen waarmee ze lead times, throughput rates en value-added ratios vergeleken. Kashwala en Irani [35] introduceerden Value Network Mapping, een techniek die VSM combineert met PFAST (Production Flow Analysis and Simplification Toolkit). Braglia et al. [32] stelden een nieuwe structuur op voor systemen met niet-lineaire productroutes. Het achterliggende idee is het op voorhand uitvoeren van een analyse die het kritische productiepad aan het licht moet brengen. Dit pad zal dan het eerst in kaart gebracht worden, waarna de andere paden er op aangesloten kunnen worden. De klassieke 5 stappen worden nu 7 stappen. Onder andere wordt er ook aandacht besteed aan het feit dat machines die door verschillende productfamilies gebruikt worden (shared resources), een beperking vormen voor elke productfamilie apart. [32]

2.3. VSM in de Semiconductorindustrie

Een zoektocht naar het gebruik van de tool VSM in de halfgeleiderindustrie leverde absoluut geen resultaten op. De auteur van deze masterproef vond geen publicaties hieromtrent en heeft toen verschillende halfgeleiderbedrijven gecontacteerd. Geen van de antwoorden was bevestigend op de vraag of er ooit één of andere vorm van een Value Stream Map werd opgesteld voor een product(familie) van wafers. Wél werd VSM gebruikt voor randactiviteiten zoals bijvoorbeeld de verwerking van afval en reinigingsmateriaal, dit is echter absoluut niet relevant voor deze masterproef.

3. Value Stream Map in FABII bij ON Semiconductor in Oudenaarde

Vooraleer een productfamilie te kiezen en de techniek toe te passen, wordt in een eerste deel een inleiding gegeven op de productieruimte (FAB) in ON Semiconductor. Enkele voor value stream mapping relevante kenmerken van het waferproductieproces worden vergeleken met de automobiellindustrie, waar Lean zijn oorsprong vond.

3.1. FABII in ON Semiconductor

In ON worden Geïntegreerde Schakelingen (Internal Circuits - IC's) gemaakt. De basis hiervoor is een silicium schijf waarop verschillende lagen boven elkaar worden aangebracht (tot 19). Voor elk van deze lagen moet de wafer een heel aantal processtappen doorlopen. Eens afgewerkt, wordt ze versneden tot individuele chips (IC's). Een IC bevat honderdduizenden componenten zoals transistoren, capaciteiten en weerstanden.

Omdat er op heel kleine dimensies (micron) gewerkt wordt, is stof de grootste vijand die de werking van de IC kan verstoren. Daarom wordt er in een stofarme ruimte gewerkt die ook wel cleanroom genoemd wordt. Een cleanroom is een geconditioneerde ruimte die aan een aantal normen inzake particles (stofdeeltjes), temperatuur (21°C) en vochtigheidsgraad (40%) moet voldoen. Een cleanroom wordt ook FAB genoemd en omgekeerd. [36] Terwijl de site in Oudenaarde vroeger bestond uit twee FAB's, blijft er nu nog één over. In deze masterproef wordt vaak verwezen naar FABII, waarmee dus de cleanroom van ON in Oudenaarde wordt bedoeld.

De vele machines in de cleanroom onder de strenge stofnormen laten draaien, is heel moeilijk. Daarom worden de bevuilende gedeeltes van de machines afgescheiden van het deel waar de wafers in- en uitgeladen worden. Dit gebeurt volgens het 'through the wall'-principe. Een muur scheidt de eigenlijke machine van de ruimte waar de wafers zich bevinden. De wafers worden door de operator in- en uit de machine geladen via één of meerdere laadblokken die door de muur komen.

De lay-out van de FAB is omwille van dit through-the-wall-principe zodanig ontworpen dat er zoveel mogelijk muren zijn. Concreet is het zo dat er een hoofdgang is met een aantal zijgangen die productiebays genoemd worden. Aangezien in deze aaneengesloten ruimte de wafers blootgesteld worden aan de omgeving, heeft ze de strengste stofnormen. Tussen die productiebays liggen dan de dienstgangen, servicebays genoemd. Deze zijn afgesloten van de hoofdgang. Hierin bevinden zich de bevuilende gedeeltes van de machines en gelden minder strenge stofnormen. Er zijn strenge regels om van de ene naar de andere gang te gaan zodat stofdeeltjes te allen

tijde geweerd wordt uit de productiegangen. Een grondplan met aanduiding van de twee types gangen is te zien in Figuur 3. Op Foto 1 zijn productiebay 1 (links) en service bay 1 (rechts) te zien.

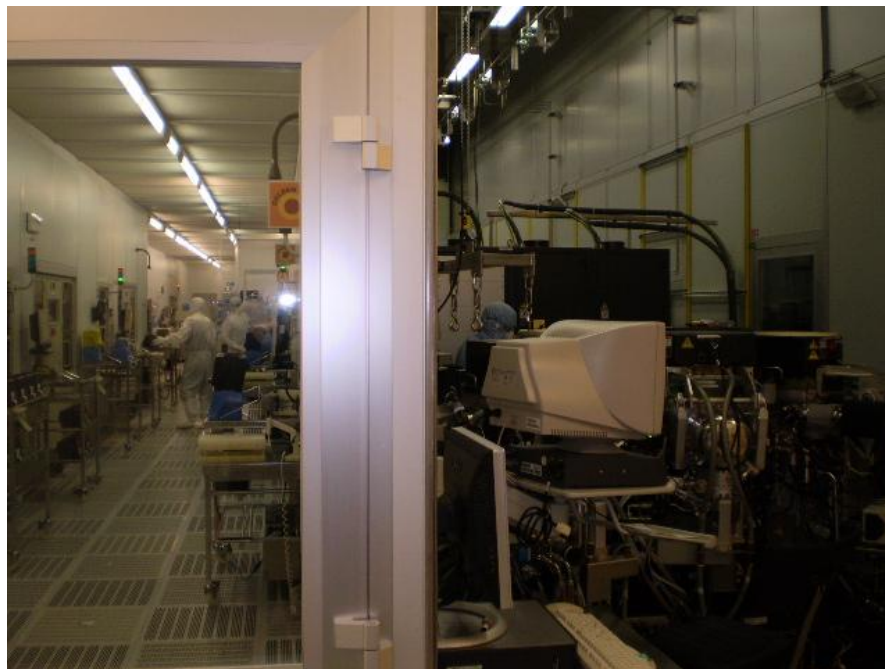


Foto 1: Links productiegang 1 (productiebay) en rechts diensgang 1 (servicebay).



Figuur 3: Grondplan van de FAB met aanduiding productiebay en servicebay

Andere typische kenmerken van een wafer productieomgeving worden samengevat in Tabel 1. Lean en dus ook Value Stream Mapping is ontstaan met het Toyota Productie Systeem. Door de grote verschillen met de halfgeleiderindustrie vereisen de technieken, zoals onder andere Value Stream Mapping, soms enige aanpassing. Daarom is het interessant enkele, voor Value Stream Mapping relevante, kenmerken in deze inleiding al tegenover elkaar te zetten.

	Automobieliindustrie	Halfgeleiderindustrie
Productievorm	Assemblagelijijn met continue productie.	Er ligt voorraad voor elke machinegroep die na verwerking naar de volgende machinegroep wordt geduwd.
Productieomgeving	Geen speciale eisen	Strenge stofnormen.
Lay-out productieomgeving	Alle werkposten gecentreerd rond de centrale band.	De machines staan in productiebays (gangen) die aansluiten op één hoofdgang. Equipment met een gelijkaardige functie (bvb de verschillende ovens) en ook de tools die de wafers daarvoor voorbereiden, staan meestal gegroepeerd in eenzelfde bay en vormen zo eigenlijk job shops.
Machines	Elke tool staat min of meer vast langs de assemblagelijijn en wordt gebruikt voor één specifieke stap in het assemblageproces.	Alle machines zijn 'shared resources' waar de wafers verscheidene keren op bewerkt worden, vaak met een (licht) verschillende operatie.
Batchgrootte	Eén auto	Loten van 25 wafers (mogelijk minder door rework en scrap, technologiелотен ⁵ hebben ook vaak minder wafers). Soms ook batches van verschillende loten tegelijk (bij de ovens bijvoorbeeld)

⁵ Technologiелотен of 'engineeringloten' zijn loten die niet bestemd zijn voor productie, maar om bepaalde nieuwe methodes of technologieën te testen.

Productroute	Lineair: elke auto komt één maal voorbij elke werkpost.	Loops: een wafer maakt verschillende lussen en komt vaak meerdere malen langs dezelfde machines. ⁶
Productdifferentiatie	Pas op het einde van de assemblage.	Al vanaf het opstarten van de blanco wafers.
Procescontrole	Voornamelijk visuele controle, ingebouwd error proof systeem.	Machines en producten worden stap voor stap gecontroleerd met testwafers en SPC-testen.
Operatoren	Een operator is verantwoordelijk voor een bepaalde werkpost op één bepaald punt aan de lijn.	Een operator is toegewezen aan een volledige productiebay, dus aan een heel aantal machines. Het is onmogelijk alle machines altijd direct te bedienen.
Capaciteit	De capaciteit van de fabriek wordt bepaald door de snelheid van de lijn.	De capaciteit van de FAB wordt bepaald door de capaciteit van enkele machinegroepen.

Tabel 1: Automobielerindustrie vs halfgeleiderindustrie

⁶ Op wafers worden verschillende lagen aangebracht. Binnen elke laag wordt vaak een gelijkaardige serie stappen uitgevoerd (chemisch reinigen, fotolithografie, etsen, implant en metaaldepositie). De wafers maken als het ware telkens lussen in hun route en komen dus vaak op dezelfde machines terug. Dit wordt een re-entrant fabrication line genoemd.

3.2. Waarde voor de klant

Volgens Rother en Shook [28] is het specificeren van waarde voor de klant een extreem belangrijk punt alvorens je een value stream map begint te tekenen. Indien deze stap overgeslagen wordt, is de kans reëel dat de eindgebruiker een product in handen krijgt dat niet overeenstemt met zijn wensen. Het in kaart brengen van de waardestream begint dus met de eisen van de klant.

De klant bij ON wil de bestelde wafers geleverd zien op het afgesproken moment, beantwoordend aan zijn kwaliteitseisen en tegen de beste prijs.

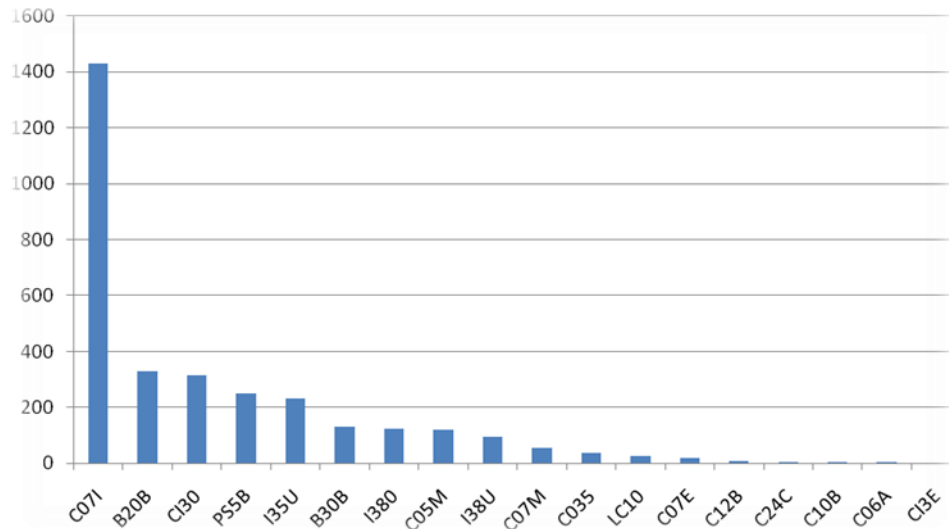
3.3. Indeling in productfamilies

Zoals al blijkt uit voorgaande passages, is de grote mix van eindproducten een typisch kenmerk van een wafer productieomgeving. In de FABII in Oudenaarde, kan ON Semiconductor duizenden verschillende microprocessors produceren. In de praktijk zijn dat er elke maand ongeveer 250, door het bedrijf reeds onderverdeeld in een 20-tal technologieën.

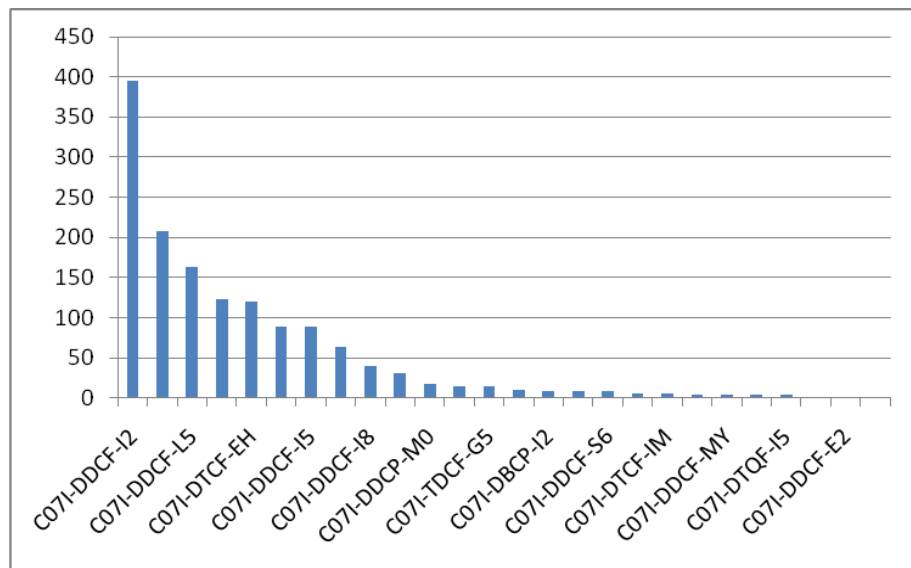
Een productfamilie is een groep producten die gelijke processtappen ondergaan en waarvan de flow dus ook langs dezelfde machines passeert in de downstream processen [28]. Het is echter moeilijk te spreken van upstream en downstream processen in een wafer productieomgeving. Downstream processen zijn namelijk die processen die in de flow na de shared resources komen. Shared resources worden gedefinieerd als machines waarop meer dan één productfamilie geproduceerd wordt [31]. In FABII (de productieruimte van ON Semiconductor in Oudenaarde) zijn echter alle machines shared resources en kunnen we de definitie van een productfamilie niet letterlijk toepassen.

In de halfgeleiderindustrie spreekt men wel van de front end en back end van de productflow. De front end bevat ongeveer ¼ van de operaties en eindigt voor de machinegroep APCVD (Watkins Johnson). Daar start het maken van de connecties tussen de componenten van de chip en de buitenwereld. De back end van de verschillende producten zijn over het algemeen iets minder verschillend van elkaar dan de front ends.

Vaak wordt eerst een Product Quantity Analysis opgesteld [32]. Deze kan voorgesteld worden met behulp van een Pareto diagram. Figuur 4 stelt het Pareto diagram van alle technologieën voor met het aantal waferstarts per week in oktober 2009. Figuur 5 is het Pareto diagram van de grootste technologie C071, opnieuw met het aantal waferstarts per week in de maand oktober 2009.



Figuur 4: De productfamilies (horizontale as) met aantal waferstarts per week in oktober 2009 (verticale as).



Figuur 5: De producten in C07I (horizontale as) met aantal waferstarts per week in oktober 2009 (verticale as).

De klassieke methode om productfamilies samen te stellen, is de productfiliematrix. Dit is een rooster dat een lijst met processen bevat in de kolommen en een lijst met producten in de rijen [31]. In het geval van de wafers is dit echter absoluut niet de eenvoudig visualiserende tool die het zou moeten zijn. Als je rekent dat elk product

ongeveer 200 operaties (processen) vergt en er enkele honderden producten zijn, dan wordt de lijst al snel heel erg onoverzichtelijk.

De verschillende benamingen (Tabel 2) binnen het bedrijf bieden echter een voor de hand liggende opdeling. Deze duiden namelijk op de technologie waarmee de microprocessor gemaakt wordt. Per definitie is het zo dat de verschillende producten binnen een technologie een redelijk gelijklopende sequentie van processtappen doorlopen, zowel in de front end als in de back end. Vaak liggen de verschillen binnen een technologie in bijvoorbeeld een lus (loop) van een aantal stappen die één of twee keer meer of minder uitgevoerd worden, of een bepaalde laag die net iets anders is. Het lijkt dus logisch om de productfamilies gelijk te stellen aan de verschillende technologieën. Op deze manier zouden we voor de FABII in Oudenaarde dus 20 (actieve) productfamilies bekommen.

B20B	C07E	C12B	CI3E	I380
B30B	C07I	C12C	HFVR	I38U
C035	C07M	C24C	HVFT	LC10
C05M	C10B	CI30	I35U	PS5B

Tabel 2: De productfamilies die in november 2009 in productie waren.

Uit de benamingen van de technologieën blijkt reeds dat er onderlinge gelijkenissen bestaan. ON zelf noemt het geheel van C07I, CI3E en CI30 bijvoorbeeld I2T, het geheel van I35U, I380, I38U wordt I3T genoemd en het geheel van B20B en B30B ook wel BIMOS. Het is dus interessant te onderzoeken of het misschien aangewezen is er bepaalde families samen te voegen tot één productfamilie.

Een productflow bestaat uit een opeenvolging van operaties. Elke operatie heeft een uniek nummer van 4 cijfers. Het is natuurlijk niet zo dat elke tool slechts één operatie kan uitvoeren. Een machinegroep (van gelijke machines) heeft dus een lijst van in zekere zin gelijkaardige operaties die er op uitgevoerd kunnen worden.

Figuur 6 toont een onderdeel van de productfiliematrix voor drie producten van BIMOS. Twee van B20B (B20B TDCB F1 en B20B TDCB F3) en één van B30B (B30B TDCB S2). In de kolommen staan de operatienummers van de eerste 25 stappen van het productieproces..

	8430	8415	6201	7611	7045	2028	2628	3006	3065	3701	4314	7611	7603	7269	1116	1118	7613	7612	7034	7047	4919	7659	7705	7713	7611
B20B TDCB F1	x		x	x	x	x	x	x	x	x	x	x		x	x		x		x		x	x	x		x
B20B TDCB F3		x	x	x	x	x	x	x	x	x	x		x	x		x		x		x	x	x	x		x
B30B RDCB S2	x		x		x	x	x	x	x	x	x	x		x		x	x		x		x	x		x	x

Figuur 6: Deel van de productfamilymatrix op basis van operatienummers.

Er bestaan echter zodanig veel operaties dat producten echt zeer grote overeenkomsten moeten vertonen om gelijkenissen te kunnen vinden in de productfamilymatrix. Sommige operaties lijken nochtans heel goed op elkaar. Zo kunnen twee operaties een in alle opzichten gelijke metaallaag op een wafer leggen, maar toch een verschillend operatienummer hebben omdat de ene laag een fractie dikker is dan de andere.

Een betere manier van voorstellen is daarom equipmentgroepen⁷ te gebruiken in plaats van operatienummers. Uiteindelijk is het de bedoeling om producten samen te beschouwen die een min of meer gelijk traject afleggen langs de machines in de FAB. Figuur 7 toont dezelfde eerste 25 stappen van dezelfde drie producten.

B20B TDCB F1	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x	x
B20B TDCB F3	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x	x	x
B30B RDCB S2	x	x		x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x	x
	INSPECTIE4	WAFERMARK	CLEAN NC FSI	LI DRIVE	CLUSTER	INSPECTIE1	STRIP RF	OX WETCH	RINSE&DRY	FSI-STRIP	CLEAN NC FSI	CLEAN CRIT	ELLIPSONOMETER	HC9500	CLEAN NC FSI	HN DRIVE	SEZ	CLEAN CRIT	EPI	CLEAN NC FSI	LI DRIVE	CLUSTER	INSPECTIE1	STRIP RF	OX WETCH

Figuur 7: Deel van de productfamilymatrix op basis van equipmentgroepen.

In het geval van I2T blijken de verschillen tussen C07I enerzijds en CI3E en CI30 anderzijds toch duidelijk groter dan de verschillen in flows van de producten binnen beide categorieën, het zijn vooral de back ends die gelijk lopen. Hetzelfde geldt voor I3T. Het is dus beter deze apart te beschouwen. In het geval van B20B en B30B is het niet steeds duidelijk. Uit de productfamiilymatrix, waarvan het eerste stuk weergegeven is in Figuur 7**Error! Reference source not found.**, kan echter berekend worden dat de eerste 100 stappen van de productflows van

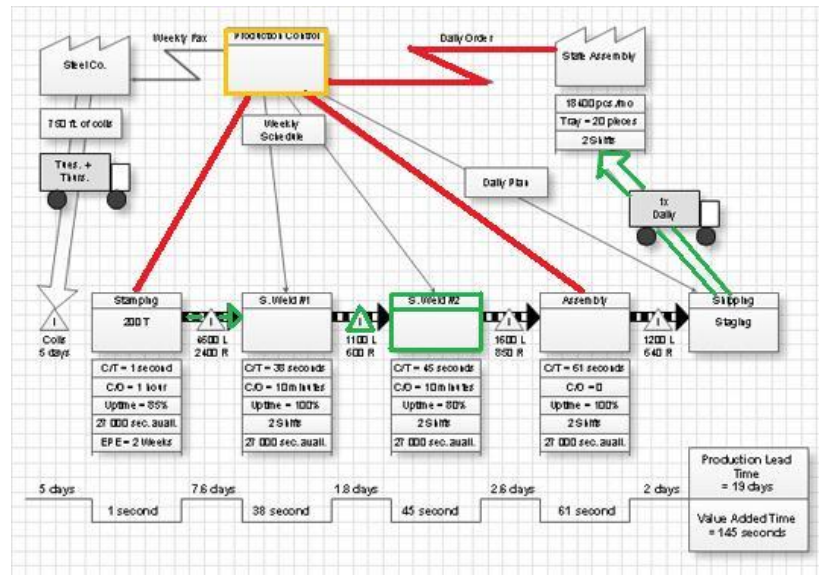
⁷ In deze masterproef worden de termen machinegroep, toolgroup en equipmentgroep door elkaar gebruikt. Ze doelen allemaal op hetzelfde, namelijk een aantal gelijke of bijna gelijke machines die vaak dezelfde operaties kunnen uitvoeren. Niet altijd exact dezelfde operaties, gezien de configuraties van de machines bijvoorbeeld anders kunnen zijn. Dit komt uitvoerig aan bod bij het bespreken van de sputtergroep in hoofdstuk 3 van deze masterproef.

B20B TDCB F1 en B20B TDCB F3 voor 90% gelijk lopen. De productflows van B20B TDCB F1 en B30B TDCB S2 lopen slechts voor 64% gelijk. Hieruit zou afgeleid kunnen worden dat B20B en B30B inderdaad meer van elkaar verschillen dan de twee producten binnen B20B. Dit geldt echter niet voor alle producten binnen die families. Omdat BIMOS (B20B en B30B) eigenlijk het enige geval is waar twijfel over samenvoegen mogelijk is, lijkt het beter voor de eenvoud de huidige indeling van het bedrijf te volgen en een onderscheid te maken tussen BIMOS2 en BIMOS3. Van de 20 technologieën uit Tabel 2 worden dus geen combinaties gemaakt en we blijven bijgevolg in deze masterproef uitgaan van 20 productfamilies.

De mix van technologieën binnen ON is bekend en het valt op dat C07I veruit de grootste is; ze vertegenwoordigde in het vierde kwartaal van 2009 45% van de totale productie. Dit lijkt dus de meest aangewezen familie om een eerste Value Stream Map voor op te stellen.

3.4. De current state map

Om een Value Stream Map te maken, tekent men een serie iconen die elk een verschillende stap voorstellen. Er zijn drie types iconen: voor de materiaalstroom, de informatiestroom en algemene iconen. Materiaalstroom iconen (groen in voorbeeld, **Error! Reference source not found.**) stellen haltes voor in het productieproces (machines, tools, ...), externe resources (fabrieken), voorraden, push arrows, verzendingen naar de customers, ... Deze materiaalstroomiconen worden vaak vergezeld van databoxen (blauw in voorbeeld, Figuur 8) met de belangrijkste gegevens over die stap [29]. De informatiestromen (rood in voorbeeld, **Error! Reference source not found.**) leggen de link tussen verschillende stappen in het productieproces waartussen informatie uitgewisseld wordt. Tot slot zijn er de overige iconen (oranje in voorbeeld, Figuur 8), die bijvoorbeeld de lead-time en value added time aangeven.



Figuur 8: Voorbeeld van een VSM. bron [37]

Er bestaan talrijke tips over het construeren van een VSM. Rother en Shook [28] vatten ze als volgt samen.

1. Verzamel informatie door effectief langs de materiaal- en informatiestromen te lopen.
2. Start met het snel doorlopen van de volledige stroom om een eerste idee te krijgen van het verloop van de processen. Loop daarna in de andere richting terug en verzamel informatie over elk proces.
3. Begin aan het einde van de productflow. Op die manier begin je met de processen die het dichtst bij de klant staan, wat je een idee zou moeten geven van het ritme van de processen meer upstream.
4. Meet zelf tijden op en vertrouw niet op bestaande informatie.
5. Bewaar het overzicht door de hele VSM zelf in kaart te brengen.
6. Gebruik steeds potlood en papier om te starten. Later kan eventueel alles ingebracht worden in de computer.

Value Stream Mapping is ontstaan in een omgeving die sterk afwijkt van de wafer-productieomgeving die hier besproken wordt, zie Tabel 1. Bovenstaande richtlijnen volgen was dan ook vaak moeilijk tot onmogelijk.

De vraag van ON Semiconductor in Oudenaarde was om specifiek de waardeestroom in de FAB in kaart te brengen. De levering van de blanco silicium schijven en het versnijden van de wafers tot IC's vallen hier dus buiten.

Het even snel doorlopen van het pad dat een product uit de familie I2T aflegt, is absoluut geen evidentie. Dat pad bevat namelijk 234 operaties, verspreid over een 50tal machines. Ter illustratie in Figuur 9 de locatie van de machines voor de laatste 20 operaties van C071-DDCF-I2 en het pad daartussen in vogelvlucht.



Figuur 9: In vogelvlucht langs de 20 laatste operaties van C07I in de FAB

Het leek interessanter om de rondleiding in de FAB te volgen die alle nieuwe operatoren krijgen en verder informatie te vergaren over de verschillende equipmentgroepen, de layout van de FAB en de manier waarop de WIP zich voortplant en hoe dit alles bijgehouden wordt. De belangrijkste zaken die bij deze eerste (informatie)ronde opvallen zijn:

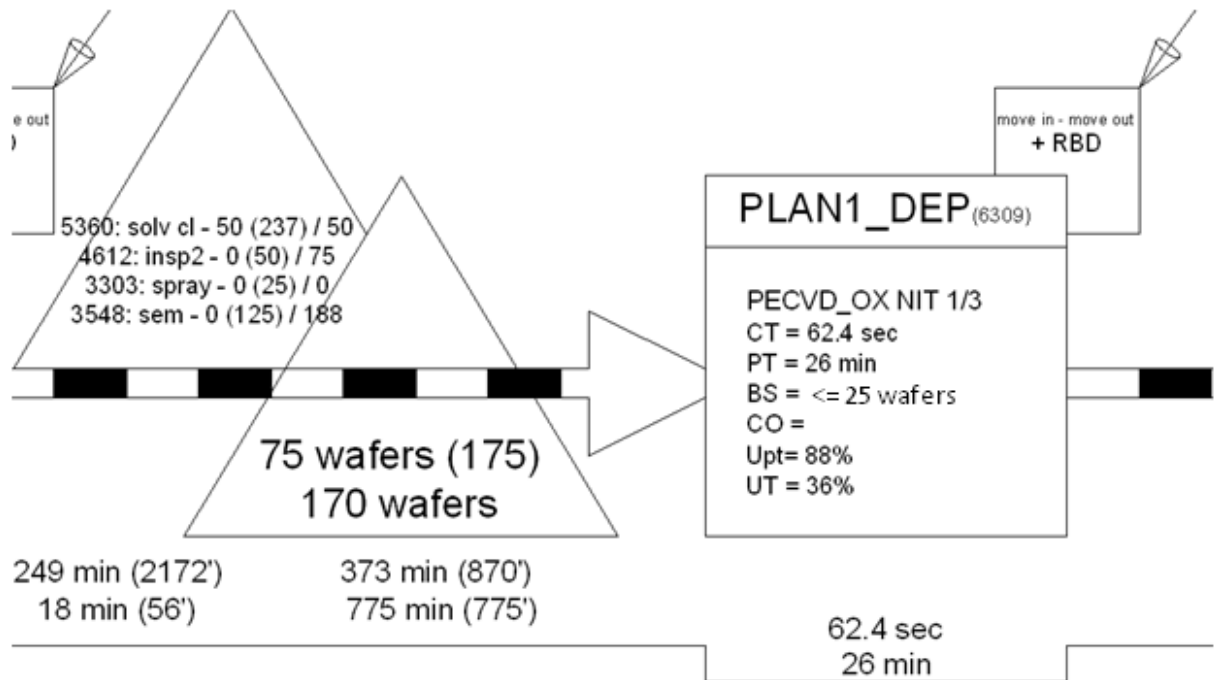
1. Slechts een beperkt deel van de FAB is toegankelijk voor operatoren, de zogenaamde production bays. Van daaruit zijn de machines te bedienen en daarbinnen staat ook alle WIP. Daarnaast zijn er ook service bays. Deze bevatten de eigenlijke machines en zijn enkel toegankelijk voor onderhoud. Het productiegedeelte bestaat uit een hoofdgang waarop 2 keer 7 zijgangen uitkomen.
2. De machines zijn verbonden met de verdieping boven én onder het productiegebied. Dat maakt het zodanig duur om ze te verplaatsen, dat dit zelfs niet in overweging mag genomen worden.
3. De keuze voor de plaatsing van de machines is op basis van job shops, maar met een heel aantal uitzonderingen.
4. WIP staat op rekken verspreid over de hoofdgang in de FAB. Meestal is ze gegroepeerd per equipmentgroup (doch de verschillende machines binnen één equipmentgroup kunnen dan weer verspreid zijn). Eens een operator beslist een bepaald lot op te starten, moet hij een 'move in' uitvoeren. Dit betekent dat hij of zij het lot van het voorraadrek in de hoofdgang haalt en scant vooraleer het de

juiste bay binnenkomt. Als het lot klaar is, wordt het lot opnieuw gescand en brengt de operator het naar het rek van het volgende proces.

5. Als een lot de FAB voor het eerst binnen komt, is de route langs alle operaties gekend. Telkens wanneer een lot gescand wordt na een bepaalde operatie, weet een programma (Workstream) wat de volgende operatie moet zijn en op welke machines die allemaal uitgevoerd kan worden. Door dit scannen (in- en uitmoven genoemd op de werkvloer) wordt het correcte recept opgeroepen op het toestel, wordt bijgehouden waar de loten zich bevinden in hun productroute, weet de operator waar het lot naartoe moet indien een proces voltooid is en kan er bekeken worden hoeveel WIP er bijvoorbeeld staat voor een bepaalde machinegroep. Via de software en een dubbelchecknummer op de wafers wordt ook telkens geverifieerd op het juiste lot in de juiste box zit.
6. De meeste machines werken per lot. Dat wil zeggen dat er telkens een lot in het toestel verdwijnt en daaruit de wafers één voor één (maar wel automatisch) behandeld worden. Een aantal echter werkt met grotere batches, zoals bijvoorbeeld de ovens.
7. Sommige operaties hebben conditionering nodig, andere niet.

Een volgende stap in het opstellen van de Value Stream Map start met het bekomen van de verschillende routes binnen de productfamilie. Zoals eerder aangehaald, focussen we op C07I-DDCF-I2. Uit de flowdefinities kan een lijst gehaald worden met in volgorde alle operatienummers, opgedeeld in routes van 1 tot 5 operaties. Met behulp van tabellen uit andere bestanden kan aan elk van deze operatienummers een equipmentgroep gelinkt worden.

Waar het normaal de bedoeling is de VSM met potlood en papier en op de werkvloer te tekenen, is hier uit noodzaak gekozen om dit te doen aan de hand van de flow uit Excel. Met deze schets was het wel mogelijk (doch enorm omslachtig) de weg van een product te volgen door de FAB. De volgende stap is er alweer één die niet volgens de voorgeschreven regels kan verlopen. Het opmeten van tijden voor alle 234 operaties + hun conditioneringen en de uptimes/availabilities is een werk dat onmogelijk uit te voeren is met de beschikbare tijd en middelen. Te meer omdat alle machines shared resources zijn en je dus bij elke operatie mogelijk een lange wachttijd hebt vooraleer je de gewenste tijden zelfs maar kan beginnen meten. In 3.5 wordt uitgelegd welke tijden dan wel gebruikt worden voor de VSM.



Figuur 10: Detail uit de VSM van C07I.

Op dit kleine stukje uit de Value Stream Map zijn reeds een heel aantal elementen te onderscheiden. Hieronder worden al deze elementen benoemd. In de volgende sectie (3.5), wordt alles verklaard en beargumenteerd.

In de databox (de rechthoek) van operatie 6309 is te zien:

- Operatie 6309 met als naam PLAN1_DEP, deze bevindt zich naar het einde toe van de productroute van C07I – DDCF – I2.
- Deze operatie wordt uitgevoerd op equipmentgroep PECVD_OX_NIT (ze legt een dunne isolerende film)
- Het is de eerste keer van drie dat de flow aan deze equipmentgroep passeert
- De machinale CT voor een wafer is 62.4 seconden.
- De tijd nodig om een lot te processen is 26 minuten (25 wafers aan 62.4 seconden). Deze machinale lot cycle time wordt door ON procestime genoemd.
- De batchgrootte is eigenlijk 1 lot, ofwel maximum 25 wafers. Het lot wordt ingeladen waarna (automatisch) elke wafer door de machine gaat. Eens allemaal klaar, kan het lot opnieuw worden uitgeladen (handmatig).
- De set-up tijd voor deze operatie is niet bekend.
- De uptime van de equipmentgroep is 88%
- De utilization time van PECVD_OX_NIT is 36%.

Links van de databox is te zien:

- Een driehoek die de actuele voorraad weergeeft die vóór de equipmentgroup PECVD_OX_NIT ligt
- Er liggen 175 wafers (normaal gezien 7 loten) van de productfamilie C07I vóór de equipmentgroup, waarvan 75 voor operatie 6309. Verder liggen er ook nog 170 wafers van andere productfamilies te wachten voor de betreffende toolgroep (zie ook 3.5.4.).
- Die 175 wafers worden omgerekend met de taktijd van C07I (4,97 minuten) tot 870 minuten, waarvan 373 minuten op rekening van die 75 wafers voor operatie 6309 komen (zie ook 3.5.4.).
- De tweede driehoek is de voorstelling van de door ON gedefinieerde non-value-added processtappen. In dit geval betreft het een solvent clean (cleanen⁸ van de wafer), twee soorten inspecties (insp2 en spray) en een sem (meetoperatie) met de bijbehorende operatienummers ervoor. De getallen ernaast volgen dezelfde redenering als de echte voorraad: bij solvent clean liggen er 50 wafers voor operatie 5360, verder nog 187 wafers van C07I voor andere operaties op solvent clean en tot slot nog 50 wafers van andere productfamilies. De tijden (in minuten) eronder worden op dezelfde manier berekend.

3.5. Uitleg bij de current state map

3.5.1. Value-added en not value-added

ON maakte reeds een opdeling tussen value-added operaties (VA) en not-value-added operaties (NVA). Men verzocht om de Value Stream Map te vereenvoudigen en enkel de VA op te nemen. Omdat daardoor eigenlijk voorbij gegaan wordt aan het doel van de current state VSM (namelijk het in kaart brengen van alle stappen die het product doorloopt), lijkt het beter om de NVA stappen toch op te nemen, maar ze voor te stellen zoals in **Error! Reference source not found..** Op die manier wordt tegemoet gekomen aan het verzoek van ON om enkel de VA stappen op te nemen in de Value Stream, maar zijn ook de andere voorgesteld op de 'kaart'. Ze vormen als het ware een soort extra inventory voor de VA-stappen. Dit met het oog op één van de toekomstige projecten in ON om die NVA-stappen in een soort snelle FIFO-lijn te steken die aangestuurd kan worden naar gelang van wat nodig is in de VA-stappen. Een dergelijk project zou eenvoudiger te realiseren zijn als alle NVA-stappen, die vaak voorbereidingen zijn op de volgende VA stap, zich ook in dezelfde bay als die VA stap zouden bevinden. Samen zouden die dan complete jobshops vormen. Vaak bevinden deze machines zich echter niet in dezelfde bay.

⁸ De Nederlandse vertaling 'schoonmaken' past niet in dit kader. In deze masterproef wordt dan ook de in ON gebruikte term 'cleanen' overgenomen.

3.5.2. Cycle time en process time

In de VSM zijn door ON (historisch) geregistreerde tijden gebruikt wat betreft de CT. Eén van de moeilijkheden was om een zo correct mogelijke tijd te kiezen uit de beschikbare data. Jaren geleden werd er opgemeten volgens volgend systeem:

Er werd zowel een throughput-getal (THP) als een efficiency-getal (EFF) genoteerd. De zogenaamde THP is de runtijd van een lot, gedeeld door 25 wafers. Dit is dus eigenlijk de machinale cycle time van één wafer. Het EFF-getal is de tijd die erbij komt om dat lot te conditioneren (indien nodig), in te laden, op te starten en uit te laden, opnieuw gedeeld door 25 wafers. Bij sommige machines (bijvoorbeeld de sputters) is het zo dat de loten zodanig opgestart kunnen worden dat de eerste wafer van een lot direct aansluit op de laatste wafer van het vorige lot. Er gaat dus geen tijd verloren met het uitladen van het eerste en inladen van het tweede lot. In dit geval wordt het THP-getal berekend als de runtijd van het gemiddeld aantal loten dat na elkaar opgestart wordt zonder onderbreken, gedeeld door dat aantal maal 25. Ook het EFF-getal wordt op die manier berekend.

De meest gebruikte definitie voor cycle time is diegene die rekent van start tot start, dat wil zeggen de tijd tussen het opstarten van een lot en het opstarten van het volgende. Met andere woorden om de hoeveel tijd dat er een lot klaar is. Geen van de twee geregistreerde tijden voldoet hieraan. Het THP-getal is de machinale cycle time en dus een onderschatting. Het THP-getal en EFF-getal samen als cycle time nemen zou een overschatting zijn, aangezien het EFF-getal ook de setup tijd en mogelijk zelfs SPC-testen bevat. Deze overschatting (setup en spc) lijkt in de meeste gevallen groter te zijn dan de onderschatting (in- en uitladen) die, naargelang van de toolgroep, tussen de 1 en de 12 min bedraagt. Er is dus gekozen om het THP-getal weer te geven, maar dan met de belangrijke opmerking dat dit de machinale cycle time is!

Verder moet ook nog opgemerkt worden dat het in- en uitladen waarvan hier sprake is ook puur machinaal gebeurt, met andere woorden dat dit gebeurt IN de machine. Dat wil zeggen dat de handelingen van de operator niet inbegrepen zijn in het EFF-getal. Meestal hebben deze echter reeds plaats terwijl het vorige lot nog in de machine zit.

Er werd dus gekozen om het THP-getal gedeeld door 25 te nemen als CT per wafer. Dit is een onderschatting door het ontbreken van de tijd voor in- en uitladen. Deze is echter klein (samen tussen de 2 en de 12 min, afhankelijk van de toolgroep) in vergelijking met de runtijden voor een volledig lot.

3.5.3. Setup / Uptime / Utilization Time / Scrap

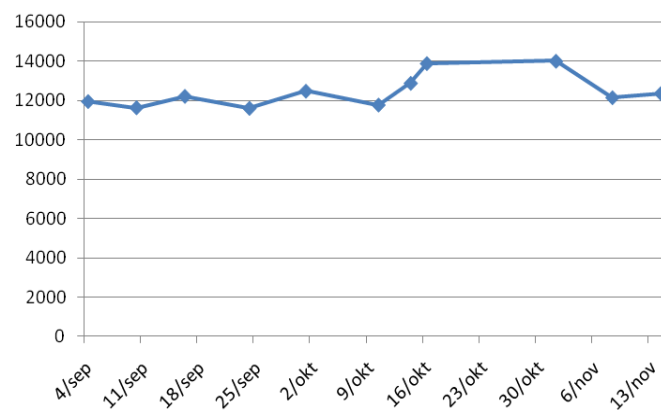
De set-up tijden voor de verschillende operaties zijn niet beschikbaar in ON en dus blanco gelaten. De sputtergroep, die in hoofdstuk 2 van deze masterproef uitgebreid behandeld wordt, is hierop een uitzondering.

Daarvoor is met de ingenieurs die voor die machinegroep verantwoordelijk zijn, een gemiddelde afgesproken. De beschikbaarheid van de machines is opnieuw gebaseerd op historische percentages van ON. Het leidt te ver om voor elke machinegroep de telkens andere manier van berekenen uit te leggen. Voor de sputtergroep is dit echter opnieuw uitgelegd in hoofdstuk 2 (paragraaf 3.2). De bezettingsgraad wordt door ON berekend aan de hand van de planning (aantal geplande moves⁹ x tijd per move / totaal beschikbare tijd). Dit kan bizarre resultaten opleveren. Zo kan de bezettingsgraad (utilization time) groter zijn dan 100%! Scrap per equipmentgroup wordt niet geregistreerd, line yield (percentage goede wafers op het totaal aantal wafers) wordt wel verrekend in de bezettingsgraad..

3.5.4. Voorraad

Het registreren van de voorraad bleek geen sinecure. Aangezien de WIP in de FAB behoorlijks fluctueert, is het concept van de momentopname waarmee een VSM werkt, mogelijk niet erg representatief.

De locaties van de WIP-ophopingen binnen de FAB zijn ook erg variabel en afhankelijk van allerlei invloeden. Het is echter vooral van belang de totale WIP te kennen en minder waar die precies ligt. Daaruit valt bij een momentopname immers toch weinig af te leiden. Een Value Stream Map heeft niet direct tot doel de extreme situaties voor te stellen. Voor ON is het ook interessanter een voorstelling te krijgen van een gemiddelde situatie. Om een dergelijk moment te kunnen kiezen, zijn de WIP-statussen binnen gehaald op 11 tijdstippen in de periode september tot november 2009, zie Figuur 11.



Figuur 11: Voorraad (in aantal wafers) in de volledige FAB op 11 tijdstippen in de periode september tot november 2009.

⁹ Er wordt hier over moves gesproken en niet over wafers, omdat één wafer verschillende keer op een zelfde machine kan langskomen. Elke keer dat een wafer in en uit een machine gaat, is een move.

Het gemiddelde van de totale voorraad¹⁰ in de FAB is hier 12 449. De voorraadhoeveelheid van 1 oktober (6u 's morgens) ligt hier het dichtst bij. Er is dus gekozen om dat tijdstip te kiezen voor de momentopname.¹¹

Het doel van het meerekenen van voorraad in een VSM is om enerzijds een idee te krijgen van de hoeveelheid voorraad, in functie van de klantvraag. Op die manier wordt duidelijk hoelang eigenlijk nog aan de vraag voldaan kan worden met wat er nu in voorraad ligt. Of anders gezegd, voor hoe ver in de toekomst heb je al voorraad liggen? Anderzijds wordt dan ook duidelijk welk percentage van de lead time waarde toevoegt aan het product.

We zetten de fysieke voorraad, met andere woorden het aantal wafers voor een operatie, om naar een tijdseenheid met behulp van de takt time (taktijd)¹² voor de productfamilie. In een productielijn met aan die lijn toegewezen tools (dedicated), spreekt dat voor zich. Elk product in elke voorraadstapel draagt expliciet bij tot de volledige

voorraad voor de productfamilie op die lijn. Alle machines in de FAB zijn echter shared resources (machines die door verscheidene families gebruikt worden) waarvoor deze redenering niet op gaat. Voor een shared resource ligt over het algemeen voorraad voor alle productfamilies die langs die machine(groep)¹³ komen, dus niet enkel van C07I in dit specifieke geval. Het zou niet correct zijn de fysieke voorraad van die andere productfamilies naar tijdseenheden om te zetten met de taktijd van de productfamilie C07I. Voor elke productfamilie zijn eigen taktijd of een gemiddelde van alle taktijden gebruiken, geeft onrealistische resultaten. Sommige productfamilies zijn

4/sep	11950	499
10/sep	11633	816
16/sep	12216	233
24/sep	11602	847
1/okt	12479	30
10/okt	11771	678
14/okt	12880	431
16/okt	13891	1442
1/nov	14009	1560
8/nov	12147	302
14/nov	12358	91
	12449	gemiddeld
	dus	1/10/2009

Figuur 12: De 11 tijdstippen met het aantal wafers als voorraad in de rekken en hun afwijking van het gemiddelde.

¹⁰ Dit is niet te verwarren met het totaal aantal wafers in de FAB. Dat ligt namelijk nog een stuk hoger. Wafers kunnen zich op drie manieren in de FAB bevinden. In voorraad, in een machine, of al klaarstaand vóór een machine. Er is dus een groot verschil tussen het totale aantal wafers in de FAB en de voorraadgroottes waarmee hier gewerkt wordt (20 980 tegenover 12 479). Voor het doel – een gemiddelde situatie vinden – maakt het niet uit welk getal precies gekozen wordt.

¹¹ Er moet opgemerkt worden dat de waferproductie in het 4^{de} kwartaal lager lag dan bijvoorbeeld nu het geval is. De zomer van 2009 was een dieptepunt en de productiehoeveelheden waren nog maar kort weer aan het stijgen.

¹² De takt time is uitgedrukt in minuten en duidt aan om de hoeveel tijd een product afgewerkt moet zijn op basis van de vraag. Deze vraag is bij ON eigenlijk een combinatie van vraag en forecasting. Er wordt namelijk zowel make – to – stock als make – to – order geproduceerd. De combinatie van beiden is echter terug te vinden in de planning en bedraagt voor het derde kwartaal van 2009, de periode waarbinnen de VSM werd opgemaakt, 26359 wafers. De takt time wordt dan berekend door het aantal minuten te delen door het aantal wafers, dat is 4,97 minuten.

¹³ Er ligt in de FAB nooit voorraad voor een bepaalde machine, maar steeds voor een machinegroep. Er is namelijk geen vaste machine uit een dergelijke groep toegewezen aan een operatie.

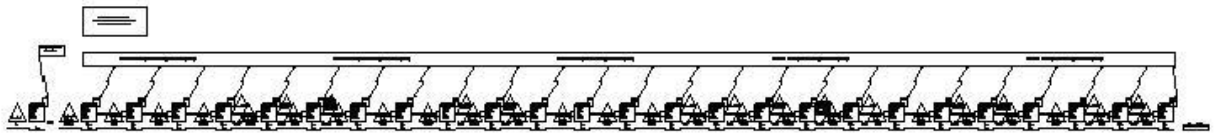
namelijk zeer klein (voor C10B bvb 17 wafers in het vierde kwartaal van 2009). C10B heeft dan een taktijd van 130 uur. Die 17 wafers worden wel allemaal samen gestart. Als die dan toevallig in de FAB liggen op het moment dat de VSM gemaakt wordt, lijkt de voorraad enorm. In samenspraak met professor Van Goubergen en ir. Joris April is dan besloten om enkel voor de voorraad van productfamilie C07I de taktijd te gebruiken. De andere voorraden mogen eigenlijk meer als een hindernis gezien worden in de VSM voor C07I. Deze voorraden moeten eerst weggewerkt worden voor C07I gestart kan worden. Daarom wordt voor deze families gerekend met de cycle times. Op die manier ligt er voor elke machinegroep een hoeveelheid voorraad C07I, omgerekend met de taktijd, die verwerkt kan worden na een zekere tijdsspanne, namelijk de tijd die het kost om alle andere families weg te werken. Het is belangrijk dit verschil in acht te nemen bij het optellen van beide om tot een Lead Time te komen (zie 3.6).

3.5.5. Push arrows en de informatiestroom

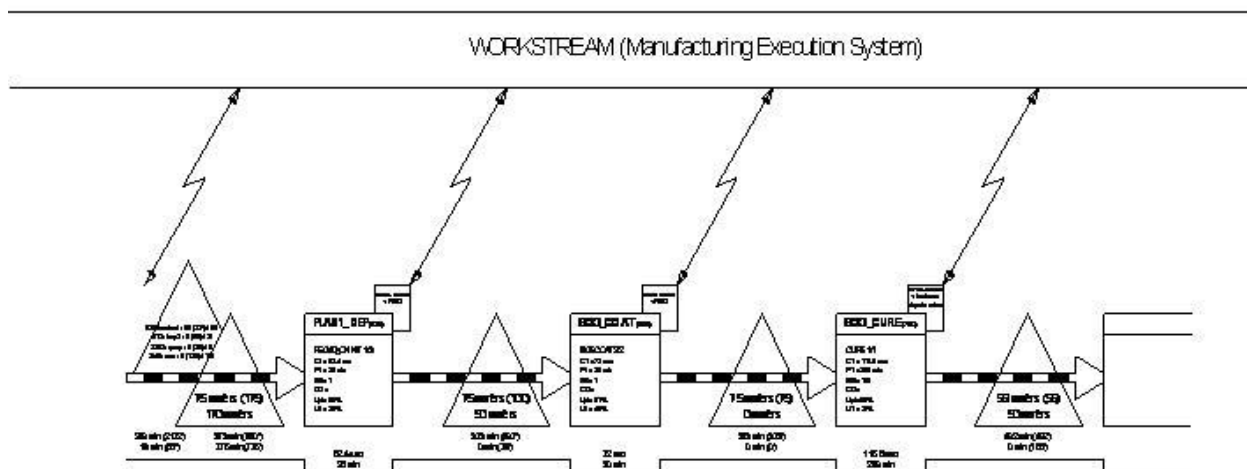
Het doorgeven van de loten tussen de processen wordt gevisualiseerd met een push arrow. Het is namelijk zo dat de wafers volgens een bepaald plan worden opgestart bij de eerste operatie en vervolgens operatie na operatie door de FAB geduwd worden. Een operator neemt telkens een lot uit de voorraad die voor de equipmentgroep staat en plaatst het nadien op het rek voor de volgende operatie. De volgorde van operaties zit bepaald in een computerprogramma (Workstream). Als een lot na een bepaalde operatie gescand wordt als 'klaar', weet het programma dat dit lot nu voorraad is voor de volgende operatie. Het lot verschijnt dan als zodanig op het scherm van de volgende equipmentgroep. In de loop van de jaren zijn voor alle equipmentgroepen richtlijnen ontstaan die aangeven in welke volgorde de WIP best verwerkt wordt. Die regels houden slechts met een beperkt aantal zaken rekening, zijn niet erg eenduidig en laten nog (te) veel vrijheid aan de operatoren die niet altijd de nodige informatie hebben om een goede beslissing te nemen. Voor enkele toolgroepen wordt ook een extra programma gebruikt, Business Object. Dit houdt rekening met de specifieke batchvoorwaarden van de betreffende machines (ovens en lithografie). Hier kan het voorkomen dat enkele specifieke loten uit voorgaande processen voorrang krijgen om een volledige batch te kunnen vullen met wafers die een gelijke behandeling krijgen. Ook al lijkt dit op het eerste gezicht een soort pull-systeem dat loten trekt uit een stroomopwaarts gelegen proces, push arrows zijn hier nog steeds gerechtvaardigd. Het 'trekken' gebeurt namelijk niet op basis van de vraag maar op basis van capaciteitsgebruik. Het kan eigenlijk meer als een bevoorrecht pushen beschouwd worden.

3.6. Conclusies VSM backend C07I

3.6.1. Figuren



Figuur 13: Overzichtsfiguur van de VSM (backend) van C07I – zie bijlage A.



Figuur 14: Ingezoomd op de backend van de VSM van C07I

3.6.2. Lead time en process time

Op verzoek van ON werd gestart met de back end. Als we deze bekijken

- is de som van de procestijden (ofwel lot cycle times) 31.5 uur
- is de som van de voorraden C07I 248,0 uur
- is de som van de overige voorraden voor de toolgroepen waar C07I passeert 130,0 uur

Het is echter moeilijk om een correct percentage te geven van de Value Added Time ten opzichte van de volledige Lead Time. Zoals eerder vermeld in 3.4.4 worden de verschillende voorraden immers verschillend omgerekend en is het ook onrealistisch te stellen dat C07I altijd zal moeten wachten tot de voorraad van andere families verwerkt

is. Dit laatste kan wel verholpen worden door een ondergrens en een bovengrens te formuleren. Namelijk de situatie waarin C07I elke keer zou moeten wachten tot de laatste en de situatie waarin de voorraad altijd alleen uit C07I bestaat. Om een realistischer beeld te scheppen wordt de 130 uur nog vermenigvuldigd met een factor 1,1. De 130 uur is namelijk een som van machine cycle times en dus korter dan de effectieve duur die de andere families nodig hebben om verwerkt te worden. Het EFF-getal besproken in 3.4 is hiervoor een goede maat. Onderstaande breuk is de gewenste factor per operatie. Een gewogen gemiddelde over alle operaties geeft 1,1.

$$\frac{\text{machine cycle time} + \text{EFF} - \text{getal}}{\text{machine cycle time}}$$

Bovengrens: de Lead Time in de back end is $248 + 31.5 = 279.5$ uur (11.6 d), waarvan 31.5u VA, dat is 11.3%.

Ondergrens: de Lead Time in de back end is $248 + 130 \times 1.1 + 31.5 = 409.5$ u (17d), waarvan 31.5u VA, dat is 7.4 %.

Om een (zeer ruw) idee te hebben van de volledige Lead Time (front end en back end) kan het totaal aantal wafers in de FAB gedeeld worden door het aantal wafers dat elke week opgestart word. Op 1 oktober 6uur was dat totaal aantal 20 980, terwijl in die periode 2600 wafers per week werden opgestart. Een ruwe schatting van de Lead Time geeft dan iets meer dan 8 weken.

3.6.3. Evaluatie

De Value Stream Map voor de back end van C07I was voor ON redelijk teleurstellend. De bedoeling was om op de VSM de bottlenecks te kunnen situeren¹⁴. Gezien het dynamisch karakter van de WIP in de FAB, kan daar aan de hand van de voorraadopstapeling niet veel over gezegd worden. Potentiële bottlenecks kunnen echter gemakkelijk herkend worden aan de bezettingsgraad. Zo zijn er in het derde kwartaal van 2009 negen machinegroepen die een geplande belasting hebben van meer dan 80% (waarvan vier zelfs boven de 95%).

Een tweede bedoeling van ON was om met behulp van de Value Stream Map inzicht te krijgen in hoe de equipmentgroepen en productstromen interageren binnen de FAB. Anders gezegd een beeld te krijgen van de flow tussen de verschillende bottleneckgroepen. Hiervoor is de VSM van één productfamilie echter niet voldoende.

¹⁴ In theorie kan er slechts één bottleneck zijn. In de halfgeleiderindustrie is die door de niet-lineariteit van de productroutes echter moeilijk te bepalen. Alle zwaar belaste toolgroepen waar soms grote WIP-ophopingen ontstaan, worden bottlenecks genoemd. Op elk moment is er één van de groepen de 'golden tool', de meest kritische equipment group. Heel vaak is dat de sputter. De variatie in de mix en de vele random breakdowns kunnen er voor zorgen dat de twijfelachtige eer om golden tool genoemd te worden, doorgegeven wordt tussen een aantal machinegroepen. De bordjes waarmee de golden tool aangeduid wordt, zijn te zien in Foto 1.

Na het opstellen van de current state map, is het normaal ook de bedoeling te kijken welke aanpassingen nodig zijn om flow te creëren, een pull systeem te integreren en verspilling te elimineren. De grote verspilling (waste) in de value stream van C07I is de voorraad (en zeker ook de voorraad van de andere productfamilies). Deze voorraad is echter een noodzakelijk kwaad in een job shop-systeem waarbij een zo hoog mogelijke throughput het doel is. De verspilling die ontstaat door de interacties tussen de verschillende productfamilies kan niet weergegeven worden met Kaizen bursts op de value stream map van C07I, een future state map voor deze ene familie is dan ook geen meerwaarde.

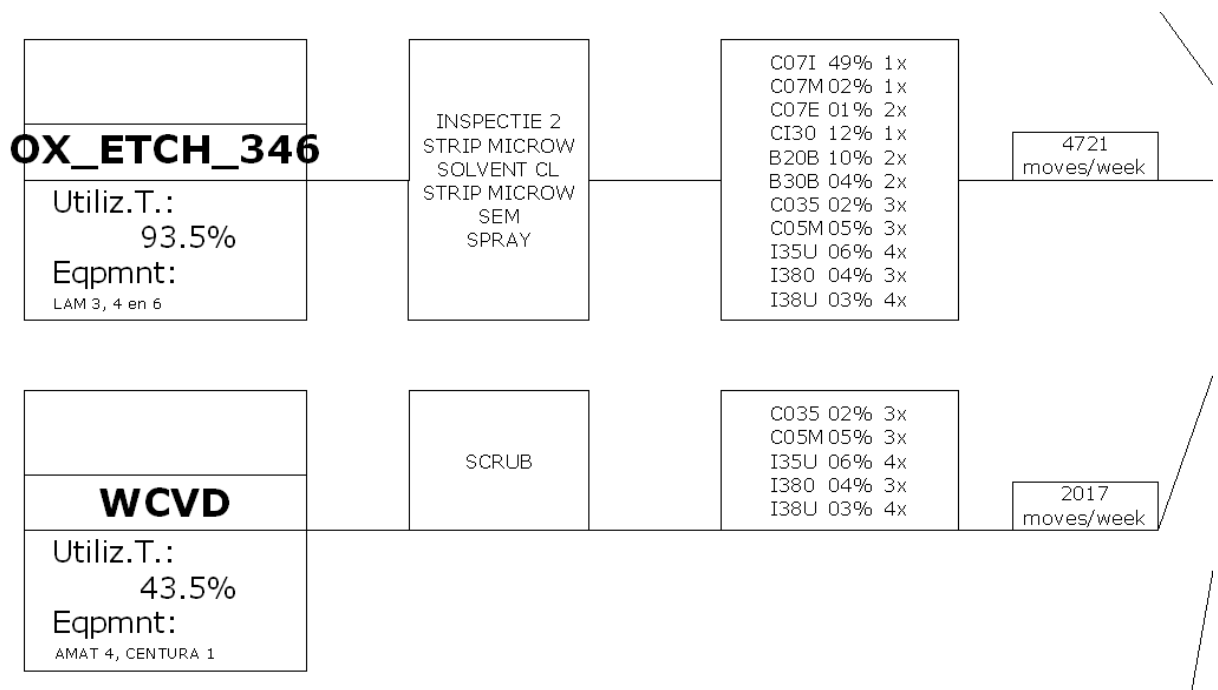
Verder is ook een pull-systeem en continue flow door de hele FAB niet haalbaar. De drie belangrijkste redenen daarvoor zijn

1. De productroutes kruisen elkaar én zichzelf te vaak (niet lineaire productroutes).
2. Als je het geplande aantal producten beschouwt en aan de hand van hun routes het aantal moves (één operatie op één machine) bekijkt per equipmentgroep, dan zijn er een heel aantal tools zeer zwaar belast (> 80%). Zolang er voor die hoog beladen groepen geen extra tools bijkomen, moeten de huidige gepland worden met het oog op het optimaal gebruik van de capaciteit. Dit is vaak tegenstrijdig met de snelle omschakelingen (kleine intervallen) die Lean (en meer specifiek pull) vraagt.
3. De huidige lay-out (productiegangen aangesloten op de hoofdgang met de vervuilende gedeeltes in servicegangen) is grotendeels bepaald door de strenge (stof)normen die gelden voor de productieomgeving, zoals besproken in 3.1. Zelfs als het volgens die normen mogelijk zou zijn een lay-out op te stellen die (al was het maar gedeeltelijk) flow in de hand werkt, dan is de omschakeling die daarvoor nodig is extravagant duur.

Wat wel mogelijk is en waar veel halfgeleiderbedrijven dan ook aan werken, is de WIP in de lijn zoveel mogelijk te balanceren door per equipmentgroep te kijken hoe deze functioneert binnen de gehele FAB en aan de hand daarvan regels te ontwerpen. Daarvoor is het interessant de flows op een andere manier te visualiseren.

4. EQUIPMENT GROUP MAP (EGM)

Een Value Stream Map heeft de beperking slechts één productfamilie in kaart te brengen (zie 3.6.3). In de literatuur worden wel gevallen besproken met dubbele en zelfs driedubbele lijnen (zie literatuurstudie van dit hoofdstuk, paragraaf 2.2), hier gaat het echter om te veel verschillende flows. Aanvullend aan de VSM is het dus interessant om een soort “equipment group map” of EGM op te stellen. Deze visualiseert de stroom naar en van één bepaald machinegroep. Samen met een lijst van de zwaar beladen machinegroepen, is dit een interessant vertrekpunt om bijvoorbeeld een dispatchstrategie op te stellen. In de EGM zijn de relaties tussen de verschillende machinegroepen te zien waardoor ze op elkaar kunnen afgestemd worden. Als voorbeeld in Figuur 16 een stuk EGM waar de sputtergroep centraal staat, met een detail er uit in Figuur 15.



Figuur 15: Detail uit de EGM voor de sputters.

Op de figuren (detail en overzicht) valt het volgende te lezen:

1. De equipmentgroep (EG) bevat 4 aparte machines, in dit geval sputter 1 t.e.m. 4
2. De belasting van de EG sputter is 73%.
3. De wafers komen bij de sputters aan vanuit vanaf 5 verschillende andere EG. PECVD_OX_NIT, HC9500, OXIDE ETCH, WCVD en RTP. Ze vertrekken vanaf de sputters naar RTP, PECVD_OX_NIT, CLUSTER, en WCVD. Daaruit blijkt inderdaad reeds dat er veel lussen inzitten. Er is ook nog een directe lus; er gaan 1828 wafers per week van de sputtergroep rechtstreeks nog eens naar de sputtergroep.
4. In de gedetailleerde figuur zie je dat na oxide-ets er een aantal NVA stappen volgen voordat de wafers bij de sputters aankomen. Dit zijn voornamelijk clean-stappen. De wafers van WCVD passeren maar één stap (de EG SCRUB) en de wafers naar RTP gaan rechtstreeks.
5. Het kader met de verschillende productfamilies bevat ook een aantal interessante gegevens. Zo staat ten eerste per familie het percentage van de productmix weergegeven. Verder staat er ook hoe vaak die familie langs deze weg naar sputter gaat. De wafers van I38U bijvoorbeeld gaan 4 keer via WCVD (en SCRUB) naar sputter, maar ook 4 keer via OX_ETCH en één keer via RTP! Inderdaad, elke I38U wafer moet 9 operaties op sputter doorlopen tijdens zijn productie.
6. Het laatste kadertje rekent uit hoeveel wafers er, volgens de planning van november, via die lijn aankomen op sputter. Voor WCVD zijn dat er bijvoorbeeld 2017 per week.
7. De routes waarlangs de wafers de sputter verlaten aan de rechterkant, zijn op dezelfde manier voorgesteld. Alleen is het daar zo dat de NVA stappen om aan een VA equipmentgroep te komen, niet altijd gelijk zijn. Daarom zijn de kaders soms opgesplitst. Het totaal aantal wafers dat de sputter verlaat is dan (volgens de planning! – geen scrap meegerekend dus) noodzakelijkerwijs gelijk aan het aantal dat aan de sputter toe komt.

Opmerkingen:

- De EGM geldt zolang er geen productfamilies bijkomen of verdwijnen. De percentages en het aantal moves variëren echter naar gelang van de productmix. Deze zijn in het voorbeeld van de sputters afkomstig uit de planning van de maand oktober 2009, dezelfde als waarin de VSM werd opgesteld. De productie lag toen een heel stuk lager, net zoals ook de mix verschillend was, dan tijdens het voorjaar 2010 bijvoorbeeld. Dit vertolkt zich onder andere in het lagere bezettingspercentage van de sputter en het aantal moves per week!
- De productfamilies die minder dan 1% van de mix voorstellen in november zijn weggelaten. De 11 opgenomen families representeren echter samen 99.6% van de mix.
- Het verschil met de bezettingsgraden uit de VSM is te wijten aan de vereenvoudiging van de equipmentgroep sputters. Het is zo dat ON de 4 sputters eigenlijk als 8 equipmentgroepen gebruikt.

Namelijk voor elke technologie één groep: TiN (Titanium Nitride), AlCu (Aluminium Koper), AlSiCu (Aluminium, Silicium, Koper), ... Bij de berekening van de bezettingsgraad en beschikbaarheid krijgt iedere technologie apart een aantal minuten toegewezen. Dat leidt tot bezettingspercentages die sterk variëren. Zo was in november de TiN technologie 91% bezet, de SANDWICH 84% en de BARPL10 9%. In de EGM wordt echter de volledige tijd van de sputtergroep bekeken en daarvoor de beschikbaarheid berekend. Dit sluit ook nauwer aan bij de realiteit, aangezien in de praktijk de technologieën (nu nog) door elkaar geprocessed worden.

Tot zover is er dus het één en ander gevisualiseerd.

1. Om een eerste idee te krijgen over de productstromen in ON, zijn deze met stift uitgezet op papier. Zowel de front end als de back end van de 10 grootste families zijn opgetekend. Het resultaat is toegevoegd als bijlage C.
2. Alle processtappen van de 10 grootste families zijn structureel onder elkaar gezet in Excel om ze met elkaar te kunnen vergelijken. Het resultaat is toegevoegd als bijlage D.
3. De flow van de productfamilie C07I (back end) is met behulp van een VSM in kaart gebracht. Het resultaat is toegevoegd als bijlage A.
4. Een nieuwe methode werd voorgesteld om de verschillende productflows rond één bepaalde equipmentgroep (sputters) te visualiseren. Het geheel hiervan dient als input voor komende projecten rond flowoptimalisatie in de FAB. Eén van deze projecten is het onderwerp van deel 2 van deze masterproef. Het resultaat is toegevoegd als bijlage B.

Hoofdstuk 2: OPTIMALISEREN VAN DE FLOW

Aan de hand van de VSM werd in hoofdstuk 1 besproken dat er momenteel in de FAB gewerkt wordt met het programma Workstream. Daarop draait RBD (Rule Based Dispatching) om de volgorde te bepalen waarin loten opgestart worden op een machinegroep. Voor sommige tools (onder andere de ovens) wordt ook Business Objects gebruikt, een programma dat toelaat iets meer geavanceerde regels aan te wenden. De voornaamste toepassing hiervan is om de maximale batchgrootte in de ovens volledig te benutten, rekening houdend met de WIP die binnen een bepaalde tijdsspanne zal aankomen.

ON kocht recent een nieuw softwarepakket voor FABII in Oudenaarde. APF (Applied Productivity Family) moet dienen om een complexere dispatchstrategie te ontwikkelen die tot doel heeft de lead-times te verminderen en de WIP te balanceren. Daartoe worden de machinegroepen onderverdeeld in enerzijds de zwaar beladen en anderzijds de licht beladen groepen. Deze laatste worden gepland op basis van line balancing (het juiste aantal moves op de juiste operaties¹⁵) en op basis van critical ratio. De equipmentgroups die door ON als bottlenecks worden aangeduid, krijgen een planning opgelegd die de capaciteit zo optimaal mogelijk benut, met andere woorden een zo hoog mogelijke throughput realiseert.

Na het implementeren en live brengen van de strippers, een niet-bottleneckgroep die als eenvoudig beschouwd wordt, volgen de sputters. Dit is wel een bottleneckgroep en hij wordt als complex beschouwd. De opdracht in deel twee van deze thesis is dus een dispatchsysteem te ontwerpen dat de capaciteit van de vier sputters optimaal benut en dat in de mate van het mogelijke ook nog rekening houdt met secundaire voordelen zoals de bottleneckgroepen verderop.

In een eerste deel wordt een literatuurstudie uitgevoerd over het plannen van een job shop-omgeving. Speciale aandacht wordt besteed aan onderzoek naar dispatchregels in de halfgeleiderindustrie. In een tweede deel wordt de sputtergroep bestudeerd en geëvalueerd. In het derde deel wordt het voor de sputters ontwikkelde dispatchsysteem uitgebreid voorgesteld.

¹⁵ Aan de hand van de geplande productmix wordt berekend hoeveel moves er ideaal per dag moeten gebeuren op welke operaties om het geplande werk af te ronden.

1. Plannen van een job shop-omgeving

In een eerste punt komt aan bod waarom planning noodzakelijk is, hoe het planningsprobleem gedefinieerd wordt en welke soorten er zijn. Dan wordt heel kort de historiek van het onderzoek besproken, waarna in een derde punt een aantal verschillende classificaties uit de literatuur behandeld worden. Het vierde en laatste punt beschrijft dan de verschillende oplossingsmethodes die in de halfgeleiderindustrie voorkomen.

1.1. Het planningsprobleem

De halfgeleiderindustrie is kapitaalintensief. Om de productiekosten per wafer te kunnen drukken, moet de capaciteit van de duurdere machines optimaal benut worden. Verder zijn korte lead times noodzakelijk om snel te kunnen reageren op de markt. De prijzen van chips dalen namelijk zeer snel [25]. Tot slot zijn de half afgewerkte silicium wafers in de FAB een heel kapitaal waard en is het dus ook de bedoeling de WIP zo laag mogelijk te houden. Om aan deze drie punten te voldoen, is een goede planning uitermate belangrijk.

Plannen of *schedulen* betekent een optimalisatieprobleem oplossen met beperkingen, ofwel een Constraint Optimisation Problem. In een productieomgeving komt het erop neer om (meestal sterk beperkende) resources toe te wijzen aan jobs volgens een bepaalde strategie (volgorde, batchgrootte, interval) die een vooropgesteld doel probeert te benaderen. Dit doel is vaak lead time reductie, minimaliseren te late leveringen, gemiddelde flowtime, ... of een combinatie van verschillende [38].

Een planning zorgt voor de verdeling van de capaciteit van de resources, de release van nieuwe producten en het aanwezig zijn van het nodige materiaal. Verder kan ze preventief onderhoud incalculeren, een datum voorspellen waarop het product afgewerkt zal zijn en geeft het de operators duidelijk aan wat er wanneer gedaan moet worden om de productie zo vlot mogelijk te laten verlopen. Rescheduling of herplannen is het aanpassen van het bestaande plan aan de huidige verstoorde situatie. Vaak voorkomende oorzaken zijn het uitvallen van een machine, het voorschuiwen van een 'hot lot', het wegvallen van een lot (scrap bvb), verandering van due date, vertraging in beschikbaarheid van het benodigde materiaal of de benodigde operators, verandering van prioriteit en kwaliteitsproblemen. [39]

Het planningsprobleem wordt als volgt gedefinieerd: een set van n jobs (productie-eenheden, patiënten, klanten) moet uitgevoerd worden op m machines (processors, work stations, loketten). Elke job moet in een (eventueel individuele) vaste volgorde (van taken) langs één of meer machines en er kan slechts één taak tegelijk gebeuren per machine, net zoals één taak slechts door één machine tegelijkertijd kan geprocessed worden. Verder moet een begonnen taak ook in één geheel afgerond worden alvorens de volgende kan starten. Het doel is de totale tijd

(makespan), nodig om alle jobs op de machines af te werken, te minimaliseren, rekening houdend met de capaciteitsbeperkingen en taakvolgordes. [38]

Als alle jobs gekend zijn vooraleer de planning gemaakt wordt, dan spreekt men van een statisch probleem. Is het aantal taken echter onbeperkt of de samenstelling niet op voorhand gekend, dan is het dynamisch [39]. Voor dynamische stochastische job-shops is het noch theoretisch, noch praktisch mogelijk om op voorhand tot een optimale planning te komen. Men kan enkel proberen bij het beschikbaar komen van een machine te bepalen wat een goede keuze zou zijn om als volgende job te starten. Deze beslissing wordt typisch gemaakt aan de hand van dispatchingregels. De uitdaging is snel te reageren op moeilijk voorspelbare en/of onverwachte gebeurtenissen. Voorbeelden hiervan zijn ophoping van WIP of juist een tekort aan een bepaald soort WIP. Dit veroorzaakt grote verstoringen in de actuele planning. Een systeem dat de planning aanpast naargelang van de situatie is dan noodzakelijk. In de praktijk wordt een systeem ook aangepast naar gelang van de WIP-status [39]. In de loop der jaren zijn vele dispatchregels ontworpen, allemaal goed om te optimaliseren naar een zeker criterium. Geen enkel(e) (stel) regel(s) voldoet echter om voor alle belangrijke productieparameters een optimum te bereiken. [39]

Waar in een deterministisch probleem alle parameters met (bijna) zekerheid gekend zijn, is dit voor stochastische problemen niet het geval. Minstens één van de parameters is daar probabilistisch [38]. De meeste praktijkproblemen behoren tot deze laatste categorie. Het tijdstip van machinepannes (breakdowns) bijvoorbeeld is bijna steeds onvoorspelbaar, net zoals het vaak ook onzeker is hoe lang het zal duren vooraleer de productie herstart kan worden. Door de zeer variërende beschikbaarheid van equipment en de bijna random machine breakdowns is het planningsprobleem in de halfgeleiderindustrie dan ook stochastisch bij uitstek.

1.2. Korte historie van het onderzoek naar planningsmethodes

Jain en Meeran [38] geven een duidelijk overzicht van de evolutie binnen het onderzoek naar schedulingproblemen: alles startte met Johnson [40] die een optimaal algoritme ontwikkelde voor een job-shop met twee machines. Daarop volgden een heel aantal effectieve heuristieken die de basis leggen voor de huidige planningstheorieën. In de jaren '60 werd meer gezocht naar exacte oplossingen met vooral enumeratieve algoritmes – het zoeken in een stelselmatig gegenereerde lijst van mogelijke oplossingen. Het belangrijkste algoritme in dit kader is zeker en vast Branch and Bound. Met de oplossingen in een boomstructuur gegoten, wordt hierin door slim schrappen van en tijdig stoppen in bepaalde takken gepoogd om snel tot een exacte oplossing te komen. Veel van deze enumeratieve technieken worden momenteel gebruikt om lower bounds te bepalen.

In de jaren '70 en begin '80 ging de aandacht vooral naar benaderende algoritmes. Men beseftte namelijk dat het planningsprobleem NP-hard¹⁶ is. Dat heeft als concreet gevolg dat er geen algoritmes voor bestaan die niet een lijst van alle oplossingen overlopen. [41] De exacte algoritmes bleken dus niet toepasbaar op grote problemen uit de praktijk. De eerste benaderingsalgoritmes waren priority dispatch rules. Deze zeer eenvoudig implementeerbare algoritmes kenden elke operatie een zekere prioriteit toe en kozen bij het vrijkomen van een machine dan voor de operatie met de hoogste prioriteit.

Halfweg de jaren '80 groeide de interesse in allerlei varianten en uitbreidingen op deze dispatchregels onder impuls van Fisher en Rinnooy Kan [42]. Rodammer en White [43] benadrukken de flexibiliteit en de kracht van goede benaderingsalgoritmes en passen ze ook succesvol toe op grote real-world problemen. Een techniek waarvan veel andere afgeleid werden, is the Shifting Bottleneck Procedure (SBP) [44]. Deze deelt het probleem op in m single-machineproblemen en lost deze dan één voor één op. Door vergelijking wordt dan één machine aangeduid als de bottleneck. Ze wordt optimaal gepland zonder rekening te houden met de andere machines. In een volgende stap wordt dan opnieuw gezocht naar de nieuwe bottleneck enz. Hoewel SBP ook enkele nadelen heeft, wordt ze toegepast in onder andere project scheduling, semiconductor manufacturing en toepassingen met assemblagelijnen. In deze periode van enorme groei ging nog steeds een klein deel van de aandacht naar exacte algoritmes en Branch and Bound werd succesvol toegepast op enkele belangrijke problemen [38].

1.3. Classificatie van planningsmethodes

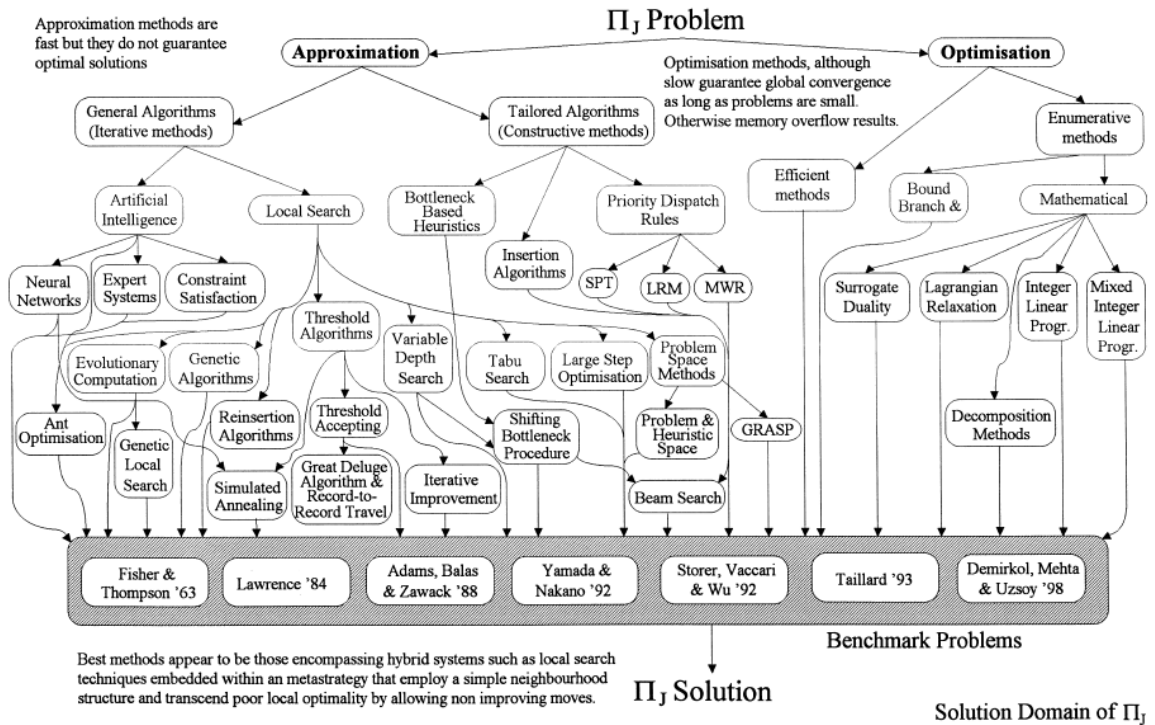
In de literatuur zijn verscheidene classificaties te vinden voor oplossingsmethodes voor het planningsprobleem. Sellers [45] deelt bijvoorbeeld slechts op in heuristieken, klassieke optimalisatieproblemen en benaderingen met neurale netwerken. De meeste auteurs maken ergens ook een onderscheid tussen exacte methodes die effectief tot de optimale oplossing komen enerzijds en benaderingsalgoritmes anderzijds. Blazewick [46] bijvoorbeeld start daarmee en deelt de twee categorieën dan verder op. Tot de exacte methodes behoren in zijn classificatie Branch and Bound en allerlei verschillende technieken om goede lower bounds te bekomen. De categorie benaderingsalgoritmes bestaat bij hem uit voorrangregels (priority rules), the shifting bottleneck heuristiek en

¹⁶ Een optimale oplossing voor het job shop planningsprobleem kan in polynomiale tijd gevonden worden als

- het aantal jobs gelijk is aan 2 of
- het aantal machines gelijk is aan 2 en alle jobs 1 of 2 operaties hebben of
- het aantal machines gelijk is aan 2 en alle operaties een duur hebben van 1 eenheid

Alle andere gevallen die verkregen worden door het verhogen van het aantal machines, jobs, operaties of duur met minstens één eenheid, zijn NP-hard. NP staat voor niet-deterministisch polynomiaal-tijd en oplossingen voor NP problemen kunnen in polynomiale tijd gecontroleerd worden. Het is nog onbekend of ze sowieso onoplosbaar zijn in polynomiale tijd. NP-hard-problemen zijn die problemen die 'ten minste zo moeilijk zijn als de moeilijkste problemen in de NP-klasse' [47]

opportunistisch plannen (tabu search, simulated annealing, genetische algoritmes). Jain en Meeran [38] hebben hun meer gebreide classificatie visueel voorgesteld, zie Figuur 17.



Figuur 17: Overzicht van oplossingstechnieken voor het planningsprobleem [38].

Jonas en Rabelo [48] delen in hun survey alle methodes op in 6 grote groepen: wiskundige algoritmes, dispatchregels, artificiële intelligentie, neurale netwerken, genetische algoritmes en zelflerende technieken. Gupta en Sivakumar [49] tenslotte hebben een classificatie gemaakt voor alle technieken die specifiek in de halfgeleiderindustrie voorkomen. De 4 categorieën hier zijn dispatch heuristieken en prioriteitsregels, wiskundige algoritmes, *neighbourhood search methods* en artificiële intelligentie. De verdere onderverdeling is gemaakt in Figuur 18. Deze laatste opdeling wordt in de volgende sectie gebruikt om de verschillende methodes te bespreken.

Table 1. A list of scheduling techniques in semiconductor manufacturing

Scheduling techniques	
1. Dispatching heuristics/priority rule	
2. Mathematical programming techniques	(a) Branch and bound method (b) Lagrangian relaxation-based approaches (c) Filtered beam search (d) Decomposition methods (e) Queuing network models
3. Neighbourhood search methods	(a) Tabu search (b) Simulated annealing (c) Genetic algorithm
4. Artificial intelligence (AI) techniques	(a) Expert-/knowledge-based systems (b) Artificial neural networks (c) Fuzzy logic (d) Petri net based approaches

Figuur 18: Overzicht van planningstechnieken in de halfgeleiderindustrie.

1.4. Oplossingsmethodes en technieken voor het planningsprobleem

Achtereenvolgens worden dispatchregels, wiskundig programmeren, zoekmethodes en technieken op basis van artificiële intelligentie besproken. Bij elk van de categorieën wordt uitgelegd wat de techniek inhoudt en wat de toepassingen zijn in de halfgeleiderindustrie. De eerste groep, dispatchregels, krijgt extra aandacht omwille van het verdere verloop van de masterproef.

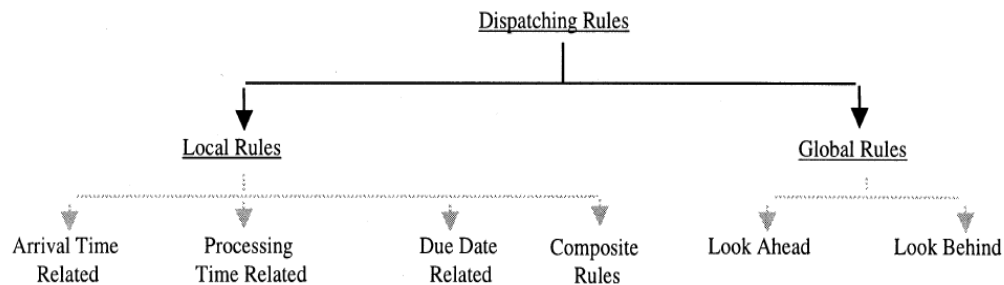
1.4.1. Dispatchregels

Scheduleregels, dispatchregels, prioriteitsregels en heuristieken worden in de literatuur vaak door elkaar gebruikt. Gere [50] maakt echter een onderscheid. Prioriteitsregels kennen volgens een zekere methode eenvoudigweg een waarde toe aan wachtende jobs en de job met de laagste waarde wordt eerst gekozen. Een heuristiek duidt hij eerder als een soort vuistregel en planningsregels zijn combinaties van prioriteitsregels en heuristieken. Dispatchregels is het geheel van deze drie. [51]

Onderzoek naar dispatchregels bestaat grotendeels uit het beantwoorden van de vraag ‘welke regel te kiezen voor het optimaliseren van een bepaald prestatie criterium?’. In het begin werd vooral gefocust op de kortste procestijd-regel (Shortest Proces Time, SPT) [51]. Het eerste onderzoek naar dispatchregels in de

halfgeleiderindustrie werd gedaan door Glassey en Resende [52], Glassey en Lozinski [53] en Wein [54]. Hun doel was bottleneck starvation te vermijden.

Verder zijn er ook weer zowel statische als dynamische regels. Statische regels zijn die waar de toegekende prioriteiten niet wijzigen in de tijd, dynamische prioriteiten doen dit wel. [55] Een tweede opdeling is die tussen lokale en globale regels (Figuur 19). Lokale regels houden enkel rekening met de informatie over de machines waarop de planning moet gebeuren en de taken die ervoor te wachten staan. Globale regels kijken ook naar de WIP en machinestatussen op andere machines [51]. In de halfgeleiderindustrie hebben lokale regels het nadeel dat ze geen rekening houden met de loops die wafers maken in de FAB en zo de lijn output moeilijk kunnen balanceren. Globale regels worden vaak nog eens opgesplitst in look ahead (bvb minimum inventory variability scheduler) en look behind (bvb de dynamic batch heuristic). [56]



Figuur 19: Classificatie van dispatchregels [56].

De meest gangbare classificatie van dispatchregels is echter een opdeling in drie categorieën. De eerste klasse bestaat uit eenvoudige prioriteitsregels die enkel en alleen gebaseerd zijn op gegevens over de jobs in de wachtlijn. Dat kan procestijd zijn, leverdatum, slack, aankomsttijd,... [66] Een aantal van de meest voorkomende regels is opgenomen in Tabel 3.

Procestijd	SPT / LPT	Shortest / Longest Processing Time
	SSPT / LSPT	Shortest / Longest Setup plus Processing Time
	SRP / LRP	Shortest / Longest Remaining Processing Time
	LSPON	Longest Subsequent Operation – langste er op volgende operatie
Setup tijd	NSUT	No SetUpTime – selecteer de job die geen setup vereist
	NB	kleinste setup tijd, rekening houdend met laatst afgewerkte job
Leverdatum	Job-EDD	Earliest Job Due Date
	Op-EDD	Earliest Operation Due Date (voor elke operatie in de productroute)

		wordt apart een due date berekend)
Kost	COVERT	Cost Over Time – de kost van te laat leveren
	VALUE	select the job with the highest “dollar” VALUE
Slack	MINSLACK	Minimum Slack
	S2	Static Slack: kleinste (leverdatum – aankomsttijd)
Aantal operaties	FOPNR / MOPNR	Fewest / Most Operations Remaining
Aankomsttijd	FIFO	First In First Out
	LIFO	Last In First Out

Tabel 3: Klasse 1 regels. Bronnen: Uzsoy et al. [57], Mittler & Schoemig [58], Jones & Rabelo [48], Panwalkar & Iskander [51]

Deze eenvoudige regels zijn praktisch allemaal onderzocht voor een wafer productieomgeving. Zoals in de inleiding van dit hoofdstuk al vermeld, moeten in de halfgeleiderindustrie de machinecapaciteiten optimaal gebruikt worden om de kost per wafer te drukken. Er moet dus absoluut voorkomen worden dat bepaalde machines stil staan door een tijdelijk tekort aan voorraad. De meeste classificaties definiëren klasse 1 als regels die afhankelijk zijn van informatie over de wachtende jobs. Voor de waferfabrieken is het interessant deze klasse uit te breiden met regels gebaseerd op machinebelastingen. Een voorbeeld van een dergelijke regel is dat de job geselecteerd wordt die als volgende stap terecht komt op de equipmentgroep die het minst belast is. Deze en andere gelijkaardige regels zullen op zich niet voldoende zijn. Ze kunnen wel als onderdeel gebruikt worden in de complexere klassen 2 en 3.

Veel gebruikt in de halfgeleiderindustrie zijn de prioriteitsregels op basis van productfamilies. Die horen in Tabel 3 eigenlijk onder ‘setup tijd’. Ze werden uitgebreid onderzocht door Chern en Liu [59]. Zij beweren dat die voor de typische job re-entrance manufacturing systemen (zie loops) grote cycle time reducties met zich mee brengen. Een dergelijke ‘family based rule’ neemt taken van dezelfde familie voor op andere waardoor het aantal setups verkleint. Missbauer [60] bewijst dit voor een systeem met één machine, Cern en Liu [59] passen het principe toe op een multi-server systeem.

De tweede klasse bestaat uit combinaties van regels uit de eerste klasse. Op deze manier kan een regel gekozen worden in functie van de verschillende scenario’s die zich voordoen op de werkvloer. Zo kan SPT gebruikt worden zolang de voorraad onder een bepaald niveau blijft, maar overgeschakeld worden naar FIFO eens erboven. Zo kan bijvoorbeeld vermeden worden dat taken met een lange procestijd te lang in de wachtlijn blijven hangen.[48]

Andere voorbeelden zijn: [51]

- SMOVE: selecteer de job die naar een volgende machine gaat waar het minst procestijd voor staat, indien alle (volgende) wachtlijnen boven een veiligheidsniveau zitten: gebruik FIFO

- 3CL-FIFO: verdeel de jobs in drie prioriteitsklassen en gebruik dan FIFO
- OPN-DDP: gebruik SPT tenzij een job zich al na de operation due time bevindt

Een variatie hierop die speciaal ontworpen is voor de halfgeleiderindustrie wordt beschreven door Johri [61]. Op basis van de locatie en de volgende stappen in de route van elk lot wordt berekend hoeveel WIP er vóór elke belangrijke equipmentgroep komt te staan. In combinatie met een lijst van machines die op dat moment beschikbaar zijn, wordt berekend hoeveel idle time er zal zijn tijdens de komende periode. Het doel is die idle time te minimaliseren. Dit wordt gedaan door voorrang te geven aan loten die de kritische equipmentgroepen naderen. Indien er geen wafers zijn die deze speciale voorrang genieten, wordt met een andere regel gewerkt.

Een ander voorbeeld is er één van Duwayri et al [62]. Ze ontwierpen een regel die de operator moet helpen te beslissen of hij moet omschakelen of niet om de FAB cycle time te verminderen. De heuristiek werkt als volgt: een bottleneck wordt gezien als een punt waar verschillende loops vertrekken en terug aankomen. De bedoeling is om telkens de taken samen te nemen die tot eenzelfde loop behoren. Elke lus die start en eindigt bij de machinegroep heeft een gewenste waferbezetting. Dat wil zeggen dat er omgeschakeld wordt naar wafers voor de volgende lus vanaf het moment dat een bepaalde loop te veel of te weinig wafers bevat.

In klasse 3 worden de regels uit de eerste klasse werkelijk samengevoegd. Elke aparte regel optimaliseert een prestatiecriteria en naargelang van het belang hiervan wordt een gewicht toegekend. De gewichten kunnen variëren in de tijd, om net zoals bij regels uit de tweede klasse in te spelen op verschillende scenario's. [48]

Voorbeelden zijn: [51]

- P+WKR: selecteer de job met de kleinste gewogen som van de kleinste volgende procestijd en het aantal operaties dat nog moet komen
- P/OPNR: selecteer de job met de kleinste verhouding tussen de procestijd en het gewogen aantal nog komende operaties

Dit soort regels wordt in de halfgeleiderindustrie vaak gebruikt. Dabbas en Fowler [56] hebben hiervan een goed voorbeeld ontwikkeld. Ze combineren lokale en globale dispatchcriteria tot één enkele regel. De 4 prestatiekenmerken waarop ze focussen zijn on time delivery (OTD), de variantie van het te laat zijn, de gemiddelde cycle time en de variantie op die cycle time. Daartoe gebruiken ze vooral SPT (shortest processing time) die een goede FAB cycle time geeft en CR (critical ratio) met een goede on time delivery als lokale regels. De twee lokale regels worden gecombineerd met een nieuwe globale regel LB (line balancing). Deze zorgt ervoor dat de wafers die het meest afwijken van de WIP-doelen de hoogste prioriteit krijgen (de WIP-doelen worden berekend op basis van de productmix die zich in de FAB bevindt). De gewichten worden bepaald met Design of Experiments en de doelfunctie is dan een lineaire combinatie van de verschillende regels met hun gewichten.

Deze combinatieregels is in volgende onderzoeken [63] grondig getest op zowel mini FAB-modellen als in de praktijk met zeer positieve resultaten.

Samenvattend kan men stellen dat de halfgeleiderindustrie vooral regels nodig heeft die de machinebezetting en WIP-verdeling in de lijn optimaliseren. Gewogen combinaties van regels die beide doelstellingen beogen worden dan ook het meest gebruikt. Hierbij wordt rekening gehouden met loops in de productroutes, de tools die werken met grote batches en de talrijke machinepannes. Vaak worden de dispatchregels binnen een FAB gecombineerd met een systeem dat het inbrengen van nieuwe wafers (release) beheert. De bedoeling is dan om de WIP in de FAB op een zeker niveau stabiel te houden en zo de FAB cycle time onder controle te hebben. [64]

1.4.2. Wiskundig Modelleren

Aangezien het job shop planningsprobleem behoort tot de klasse van NP-moeilijke problemen, zijn er geen gekende algoritmes die een gegarandeerd optimale oplossing leveren in polynomiale tijd. Dat wil zeggen dat dynamisch of geheeltalig programmeren enkel toegepast kan worden op kleine schaal (bijvoorbeeld één of twee machinegroepen). Problemen op grote schaal kunnen worden opgelost met wiskundige heuristieken zoals simulated annealing, tabu search en genetische algoritmes [48]. Hoewel deze ook ondergebracht kunnen worden in die tweede categorie, volgen we de voorgestelde opdeling en worden ze besproken in categorie 3.

Er bestaan echter ook een aantal technieken die ertoe kunnen bijdragen om de rekentijd te verminderen. Davis en Jones onderzochten de toepassing van decompositiemethodes voor een controlesysteem van de werkvloer waarin de totale makespan geminimaliseerd werd [65]. Luh and Hoitomt gebruikten naast decompositie ook Lagrangian relaxatie om een FAB-schedulingprobleem te ontmantelen tot jobs op operatieniveau [66]. Lagrangian relaxatie laat de constraints voor geheeltaligheid vallen en integreert deze als een extra kost in de doelfunctie.

Een andere, meer populaire, techniek om problemen met geheeltalig programmeren op te lossen, is de enumeratieve methode Branch and Bound. Het probleem wordt hier voorgesteld in een boomstructuur. De oplossingsverzameling wordt in elke knoop van de structuur verder opgedeeld zodat de bladeren onderaan de boom uiteindelijk al de oplossingen bevatten. Door met behulp van onder andere lower bounds bepaalde takken van de boom te schrappen, kunnen relatief grote problemen opgelost worden. Toch vergt dit nog veel rekenkracht (en –tijd) [48].

Een toepassing op relatief kleine schaal werd onderzocht door Sung and Choung [67]. Zij gebruikten een Branch and Bound algoritme voor het dynamisch plannen van de batches voor één oven in een FAB, met als doel de totale tijd nodig om alle jobs te processen (makespan) te minimaliseren. Toepassingen op grote schaal zijn wel onderzocht in de literatuur, doch niet terug te vinden op de werkvloer [49].

1.4.3. Local Search Verbeteralgoritmes

Deze algoritmes kunnen planningsproblemen lokaal optimaliseren. Door voorgedefinieerde kleine wijzigingen aan een initiële oplossing aan te brengen, creëren ze buuroplossingen (*neighbors*) waarbinnen ze de beste oplossing selecteren. Eens de oplossing niet meer verbetert, stopt het algoritme. Het gevaar hierbij is dat er een lokaal optimum gevonden is, dat wil zeggen dat er in de omgeving van bereikbare burens geen betere oplossing bestaat maar dat niet de volledige oplossingsruimte doorzocht is. Procedures gebaseerd op lokaal zoeken, garanderen dan ook geen optimale oplossing. Tabu search, simulated annealing en genetische algoritmes zijn de bekendste voorbeelden van local search verbeteralgoritmes. Elk heeft zijn eigen techniek om buuroplossingen te genereren, een eigen stopregel en een eigen methode om lokale optima te vermijden. [68]

Tabu doorloopt mogelijke oplossingen voor het planningsprobleem in een welbepaalde volgorde. Bij het springen naar een volgende kandidaat, worden op een lijst van mogelijke sprongen sommige als verboden (taboe) gemarkeerd [49] en [69]. Geiger ontwikkelde een dergelijke heuristiek voor het nat etsen in een halfgeleiderfabriek [70]. Simulated Annealing is vergelijkbaar, alleen kiest het in plaats van goede burens de beste volgende startpositie [71]. Yim en Lee [72] pasten dit toe op een clustergroep in een FAB. Genetische Algoritmes werken naar analogie met de evolutietheorie van Darwin. Er wordt gestart vanuit een willekeurige populatie oplossingen die elk een beoordeling (fitnesswaarde) meekrijgen. De populatie wordt telkens uitgedund (waarbij de betere oplossingen meer overlevingskansen hebben) en opnieuw aangevuld (door middel van tevens voorgedefinieerde wijzigingen) [73].

1.4.4. Artificiële Intelligentietechnieken

Door combinatie van zowel kwalitatieve als kwantitatieve types kennis, kunnen heuristieken gegenereerd worden die een stuk complexer zijn dan de eenvoudige prioriteitsregels. Artificiële Neuronale Netwerken bijvoorbeeld werken met *learning rules* die pogen het leerproces van mensen te imiteren. De techniek bestaat eruit dat patronen die als training werden aangegeven, herkend worden in de structuur van een nieuw netwerk. Hoewel elders in job shop scheduling al toegepast en in de halfgeleiderindustrie reeds onderzocht, is het waarschijnlijk nog lang wachten op praktijkvoorbeelden. [49]

2. Evalueren van de sputters

In hoofdstuk 2 van deze masterpoef worden de vier zwaar belaste sputters nader bekeken. Deze machinegroep wordt door ON als één van de grootste bottlenecks beschouwd. Dit tweede deel van hoofdstuk 2 start met meer uitleg over de sputtergroep. Technische details worden vermeden, tenzij relevant voor het uiteindelijke optimaliseren van het capaciteitsgebruik. In een tweede punt wordt de methode van ON toegelicht die gebruikt wordt om de uptime van de sputters te berekenen. Tot slot wordt deze berekening geverifieerd door een soort work sampling uit te voeren op sputter 1. De resultaten hiervan geven extra inzicht in de sputters, wat van nut kan zijn bij het verhogen van de capaciteit.

2.1. De sputtergroep

2.1.1. Algemeen

De sputtergroep bestaat uit 4 individuele sputters waarvan 3 in productiebay 1 en 1 in bay 5. In de sputtertoestellen worden dunne filmen onder een hoog vacuüm op de wafers aangebracht. Dit zijn onder andere de metaallagen die moeten instaan voor de interne verbindingen tussen de transistorcomponenten en de link tussen de componenten en de buitenwereld (bondpad). De sputters bevinden zich in de tweede helft (voornamelijk back end) van de productroute en alle productfamilies moeten er minstens 3 maal bewerkt worden.

Voor het sputteren van de wafers moet een native oxide verwijderd worden. Bij een groot deel van de wafers wordt dit in de sputter zelf gedaan, een ander deel moet echter eerst door Wetlijn 1. Deze bevindt zich naast sputter 1. Het cleanen gebeurt vaak pas vlak voordat een lot de sputter binnen gaat. Er is namelijk een maximum tijdsinterval tussen beide operaties.

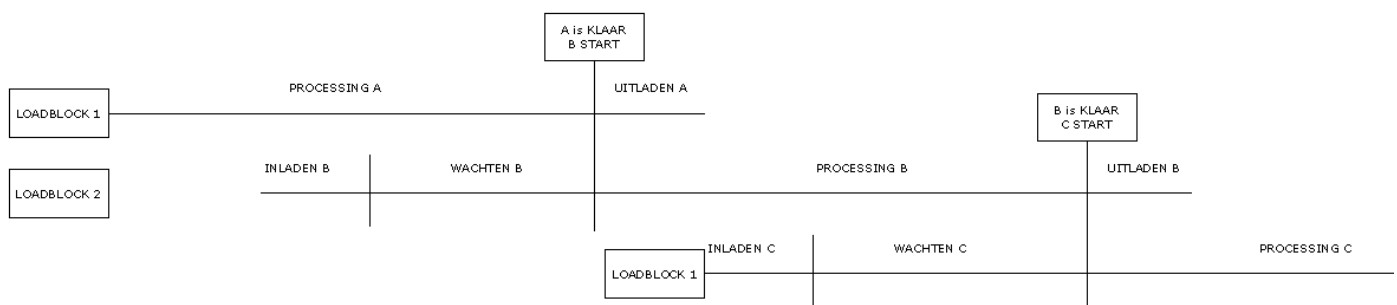
Het mainframe van de sputter (zie Foto 2) bestaat uit een aantal aparte kamers (A, D, 1, 2, 3 en 4). De wafers dienen verschillende van die kamers te doorlopen, afhankelijk van welke laag erop aangebracht moet worden (TitaniumNitride, AluminiumKoper, AluminiumSiliciumKoper, ...). Op elke sputter kan dus een kamerconfiguratie gekozen worden die toelaat bepaalde lagen wel en bepaalde lagen niet te sputteren op die machine. Het is zo dat alle wafers door kamer A moeten, het is een soort van doorgeefkamer. Bij sommige operaties zal de A-kamer ook effectief sputteren, bij de meeste niet.



Foto 2: Zicht op het mainframe van een sputter met verschillende kamers (service bay).

2.1.2. De loadlocks¹⁷

De mainframes bevinden zich zoals alle machines in de dienstgangen. In de productiegangen bevinden zich de *loadlocks*, twee per toestel. Het hele sputterproces gebeurt in een vacuümomgeving. Nadat de operator een lot in een *loadlock* geïntroduceerd heeft, is er dus nog een zekere tijd nodig voor het vacuüm trekken voordat de eerste wafer effectief de eerste kamer (A) binnen gaat. Dit zou heel wat tijd verspillen, moesten er geen twee laadstations zijn. Terwijl de wafers in laadblok 1 nog geprocesseerd worden, kan een lot in 2 al voorbereid worden. Zo kan de eerste wafer uit het lot in laadblok 2 kamer A binnen gaan, zodra de laatste wafer uit laadblok 1 die kamer verlaten heeft. In theorie kan een sputter dus zonder stoppen en zonder NVA time blijven processen met dit systeem (zie Figuur 20). In de praktijk wordt die continue productie echter verstoord door SPC-testen, conditioneringswafers, onderhoudsscenario's en dergelijke meer.



Figuur 20: Systeem met twee laadstations bij de sputters (groter in bijlage D).



Foto 3: Zicht op de sputter vanuit de productiebay met één gesloten en één open load block.

2.1.3. Setups

Op de 4 sputters kunnen momenteel in totaal 58 operaties¹⁸ gebeuren. Sommige daarvan hebben een setup nodig, sommige niet. Een setup betekent dat een zogenaamde conditioneringswafer eerst door de kamers moet, alvorens de productiewafers dat mogen. De setup blijft geldig zolang er geen half uur verstrijkt tussen het sputteren van twee opeenvolgende wafers van die operatie (of andere operaties met gelijke setup-eisen). Als er een andere operatie tussen komt, of wanneer om andere redenen een half uur verstreken is, moet een nieuwe conditioneringswafer door de sputter. Dat kost behalve de effectieve runtijd van die wafer, normaal ook in- en uitlaadtijd.

2.1.4. Onderhoud en monitoring

Zowel het mainframe van de sputter als de kamers apart worden regelmatig op HOLD gezet (buiten gebruik) om preventief onderhoud uit te voeren. Dit duurt tussen de 4 en de 12u. In die tijd zit begrepen de nodige tijd om na het onderhoud releasetesten uit te voeren vooraleer de productie opnieuw opgestart kan worden. Dit preventief onderhoud kan bijvoorbeeld het metaal dat op de wafers gesputterd wordt, vernieuwen.

¹⁸ Een wafer die de machine binnengaat, wordt gesputterd volgens een bepaald recept. Elk operatienummer heeft dus telkens een uniek recept. Behalve de 58 operaties om productiewafers te behandelen, zijn er ook nog recepten voor setups, testwafers, rework, ...

Een sputter wordt niet op hold gezet voor het regelmatig controleren van de kamers. Telkens wanneer 300 wafers door dezelfde kamercombinatie gegaan zijn, moet een monitorwafer tussengevoegd worden. Deze gaat (normaal) juist zoals de conditioneringswafer in een waferbox apart door de sputter en vergt dus opnieuw in- en uitlaadtijd. Deze monitorwafer moet nadien enkele SPC-testen (metingen) ondergaan door de operator. Het aantal wafers wordt bijgehouden door zogenaamde counters. Ook op tijdsbasis worden de kamers gecontroleerd d.m.v. monitorwafers. Al deze testen hebben als hoofdbedoeling de kamers te controleren op partikels.

2.2. Uptime, berekend door ON

Het is absoluut niet evident om de uptime van de sputtergroep te berekenen. Als de volledige sputter onbeschikbaar wordt, dan is het vrij logisch dat dit als downtime beschouwd wordt. Maar als één van de kamers op hold komt te staan (dat is: down, om welke reden dan ook), dan kunnen enkel die operaties die de betreffende kamer nodig hebben niet meer uitgevoerd worden. Er zijn echter nog andere operaties die geen gebruik maken van die kamer, waarvoor er geen problemen zijn. Als er verschillende kamers down zijn, wordt het aantal nog mogelijke kamercombinaties - en dus operaties - wel klein tot absoluut nul. Opgemerkt moet worden dat kamer A opnieuw een speciaal geval is. Als er een defect is op kamer A is het mogelijk dat heel de sputter op hold komt te staan, alle wafers moeten namelijk door A (uitzondering: bij een klein aantal specifieke problemen met kamer A, kan ze eventueel wel nog werken als doorgeefkamer).

ON hangt voor het berekenen van de uptime een gewichtsfactor aan elke kamer. Deze duidt aan hoe groot de invloed op de volledige sputter is wanneer er iets mee mis gaat. Voor kamer A rekenen ze dus 100% (wat door de besproken uitzondering eigenlijk niet volledig klopt). Een kamer die in zowat alle mogelijke combinaties zit, zal een grote invloed krijgen; een kamer die weinig gebruikt wordt, heeft een lage gewichtsfactor. In Tabel 4 wordt een overzicht gegeven van sputter 1 met zijn kamers.

	UP kamer	Gewichtsfactor	Down (1 – UP)	UP SPUTTER = 1 - down x invloed
<i>SPUTTER1A</i>	95,0%	100%	5,0%	95,0%
<i>SPUTTER1D</i>	100,0%	15%	0,0%	100,0%
<i>SPUTTER11</i>	77,5%	15%	22,5%	96,6%
<i>SPUTTER12</i>	85,0%	55%	15,0%	91,8%
<i>SPUTTER13</i>	86,1%	0%	13,9%	100,0%
<i>SPUTTER14</i>	100,0%	10%	0,0%	100,0%

Tabel 4: De kamers van sputter 1 met gewichtsfactoren en uptime.

Kamer 2 heeft een redelijk hoge gewichtsfactor omdat deze in praktisch alle mogelijke combinaties gebruikt wordt op sputter 1. Bij het op 'hold' komen van kamer 3, kunnen enkel de 4 GLUE operaties niet meer doorgaan.

De getallen in de rechterkolom worden vermenigvuldigd tot 84,2%. Als uiteindelijke uptime (per week) wordt dan het minimum genomen van die 84,2% en de tijd dat de volledige sputter apart op hold is gezet (voor bvb onderhoud aan iets anders dan de kamers). Tabel 5 geeft een overzicht voor de 4 weken van februari.

	Week 1	Week 2	Week 3	Week 4	Gemiddeld
Kamers sp1	84,2 %	80,1 %	62,9 %	91,5 %	
Sputter 1	92,6 %	89,1 %	98,3 %	98,2 %	
Min	84,2 %	80,1 %	62,9 %	91,5 %	79,7 %

Tabel 5: Uptime van sputter 1 over de 4 weken van februari 2010.

De uptime is in dit geval wel erg laag. De andere sputters deden het die maand een stuk minder slecht. Het gemiddelde voor de 4 sputters was in de maand februari voor de opeenvolgende weken 85,5%, 82,7%, 86,6% en 91,7%. Gemiddeld 86,6% dus volgens de berekeningsmethode van ON zelf.

2.3. Work Sampling

Volgens de ingenieurs van ON sluiten hun berekeningen behoorlijk aan bij de werkelijkheid, ook al werd dat eigenlijk nooit extensief gecontroleerd. Als enige minpunt melden ze dat de gewichtsfactoren misschien wel eens een revisie zouden mogen krijgen. Deze veranderen namelijk bij grote verschuivingen in de productmix en ook bij nieuwe operaties.

Er zijn twee manieren om de uptime van machines te controleren en te weten te komen welk percentage van de tijd besteed wordt aan processen, conditioneren en onderhouden, namelijk op directe en op indirecte wijze. Direct wil zeggen dat er continu of discreet geregistreerd wordt wat de machine aan het doen is. Indirect is gebaseerd op de productie-elementen zelf (aantal geproduceerde items, standaard werktijden, ...) [74]. In ON Semiconductor worden een groot aantal zaken geregistreerd. Bijvoorbeeld wanneer welk lot geladen wordt in de machine, wanneer het klaar is en van wanneer tot wanneer de sputters op HOLD staan (dat wil zeggen onbeschikbaar zijn). Op het eerste gezicht lijkt het dus evident de verschillende percentages te berekenen aan de hand van de continu geregistreerde data. Om enigszins relevant te zijn zou dat moeten gebeuren over een voldoende lange periode. De registratie van al die tijdstippen gebeurt echter vrij onduidelijk. Er kunnen op een aantal plaatsen verschillende log-files opgevraagd worden die telkens andere elementen bijhouden. Het kost echter teveel werk om voor bijvoorbeeld de tijdspanne van een maand uit te zoeken wat een machine elke seconde uitvoert. Iemand gedurende een volledige maand met een chronometer naast de sputters zetten is om voor de hand liggende redenen ook geen optie. Daarom is voor een aangepaste vorm van 'work sampling' gekozen, de populairste methode om direct en discreet (discontinu) machineperformanties te meten. Deze aangepaste vorm houdt in dat er observaties uitgevoerd worden in de geregistreerde gegevens in plaats van op de werkvloer. Hoewel dit laatste te verkiezen is, was dit praktisch niet haalbaar.

De techniek werd voor het eerst toegepast in 1927. Leonard Tippett bestudeerde er de up- en downtime van weefgetouwen in Engeland mee. Morrow introduceerde ze in de Verenigde Staten (1941) onder de naam "ratio delay". De huidige naam work sampling werd voor het eerst gebruikt door Brisley en Wadell in 1952. [75] Verder verwijzen ook de termen 'activity sampling', 'snap readings' en 'occurrence sampling' naar dezelfde techniek [74].

Work Sampling betekent dat een groot aantal random observaties uitgevoerd wordt. Tijdens elke observatie wordt de activiteit van de sputter op dat moment geregistreerd. Het is gebaseerd op kansverdeling en meer bepaald op de binomiale verdeling. Deze laatste wordt benaderd door de normaalverdeling in het geval van een groot aantal observaties. Om statistisch aanvaardbaar te zijn, moet geobserveerd worden op statistisch willekeurige momenten. Deze tijden moeten dus at random, onbevooroordeeld en onafhankelijk gekozen worden. Om dat te bereiken, bestaan er een heel aantal methodes, waarvan een lijst met random nummers, computergegenereerde tijdstippen en manueel getrokken nummers de populairste zijn [74].

De uitvoering van Work Sampling wordt opgedeeld in de volgende stappen [74] en [76]:

1. vastleggen van het doel
2. vastleggen van de verschillende elementen binnen de studie
3. schatten van de percentages tijd die aan deze elementen worden besteed
4. bepalen van de gewenste nauwkeurigheid en betrouwbaarheidsintervallen
5. bepalen van het nodige aantal observaties
6. vastleggen van de tijden en routes
7. ontwerpen van het observatieformulier
8. verzamelen van de sample data
9. de resultaten neerschrijven

2.3.1. Vastleggen van het doel

De groep sputters wordt door ON als een bottleneck beschouwd. Dit oordeel is bepaald op enerzijds de ervaring dat er daar vaak grote hoeveelheden WIP samenkomen en blijven staan, anderzijds op berekeningen, gebaseerd op historische downtime en huidige productieplanning. Met Work Sampling kan nagegaan worden of

- deze toolgroep inderdaad een hoge bezettingsgraad heeft
- hoeveel tijd naar preventief (en in mindere mate correctief) onderhoud gaat
- of de tijd dat er effectief geprocesst wordt, overeenstemt met de tijden waarmee ON werkt
- hoeveel tijd de machine beschikbaar is maar niet gebruikt wordt en waarom niet (idle time)

2.3.2. Vastleggen van de verschillende elementen binnen de studie

Als de machine in de toestand VVP (vrij voor productie) staat, dan is ze dus UP. Deze UPTIME kunnen we in verschillende situaties opdelen. De sputter kan aan het processen zijn (voor productie of voor engineering¹⁹), ze kan aan het conditioneren zijn (om te veranderen van operatie) en ze kan idle zijn, dat wil zeggen niet gebruikt worden terwijl dat wel mogelijk is.

Als de machine in de toestand HOLD staat, dan is ze niet beschikbaar. Dat kan omwille van drie verschillende redenen zijn: preventief onderhoud (na een vastgelegde tijd én na een vastgelegd aantal wafers), correctief onderhoud en releases (het testen van de machine voordat productie opnieuw mogelijk is).

¹⁹ Dit zijn bijvoorbeeld nieuwe technologieën of nieuwe methodes die getest worden.

Tot slot is er nog specifiek voor de sputtergroep dat wanneer een bepaalde kamer op HOLD staat om één of andere reden, de machine daarom niet volledig onbeschikbaar is (zie 2.1). Idle heeft hier dus wel een kanttekening nodig. Als een sputter UP is maar bijvoorbeeld 2 van zijn kamers niet, zodat slechts een beperkt aantal operaties kan doorgaan, dan is het goed mogelijk dat er nu juist geen voorraad staat voor die operatienummers. De sputter is dus eigenlijk idle, terwijl twee van zijn kamers down zijn. Dat zou dan ook apart bekeken moeten worden.

Dit samengevat geeft de mogelijke toestanden van de machines weer bij een work sampling observatie. (Tabel 6)

UPTIME	processing voor productie
	processing voor engineering
	processing terwijl andere kamers op hold staan
	conditionering
	idle time
DOWNTIME	preventief onderhoud – counterbased
	preventief onderhoud – timebased
	correctieve actie

Tabel 6: De elementen voor observatie.

2.3.3. Schatten van de percentages tijd die aan deze elementen worden besteed

Om in 2.3.4 het aantal nodige observaties te kunnen bepalen in functie van de gewenste nauwkeurigheid, is voor alle elementen een schatting nodig van het percentage. Deze schatting kan indirect gebeuren, dat wil zeggen op basis van de standaard werktijden, productiehoeveelheden en onderhoudsplannen.

Om te schatten hoeveel tijd naar het processen van productiewafers gaat, kan het aantal moves op de sputter per maand vermenigvuldigd worden met de gemiddelde tijd voor die move. De som hiervan over alle operaties, geeft dat de totale procestijd. Dat blijkt 29125 minuten in de maand februari te zijn, ofwel 72 % van de totale tijd. Zie Figuur 21.

Hierin zit echter geen handling én er is van uit gegaan dat alle operaties uitgevoerd werden op de sputter met de beste tijd daarvoor. In de praktijk zal dit soms niet gebeuren. In samenspraak met de ingenieurs in ON wordt geschat dat 3% extra tijd ten opzichte van het gekende getal voldoende is om rekening te houden met handling en langere procestijden. Het geschatte tijdspercentage voor het bewerken van productiewafers is dus 75%.

operatienummer	1401	1402	1403	1405	1406	1407	1410	1411	1412	1414	1415	1417	1418	1419	1428	1429	1430	1431	1432		
# moves in feb	1347	3	1	6	5	36	36	139	461	25	411	25	42	42	1	5	2280	139	139		
beste tijd	2,6	2,6	1,6	2,3	1,6	1,9	2,8	1,8	2,6	1,8	1,8	2,6	1,8	1,8	1,9	1,9	2,1	2,5	2,5		
totale tijd	3529	8	2	14	8	68	100	256	1208	44	740	64	76	76	2	10	4765	346	346		
	1433	1434	1435	1436	1437	1438	1441	1442	1443	1444	1445	1447	1448	1451	1452	1453	1454	1459	1460	1461	
	139	25	139	552	552	480	378	60	253	1808	458	480	1146	61	390	50	350	13	378	376	
	2,5	2,5	1,8	1,8	1,8	1,8	1,8	1,8	1,8	2,1	2,5	2,5	2,5	2,6	2,2	2,2	2,2	4,2	2,5	2,5	
	346	61	250	994	994	864	680	108	455	3706	1127	1195	2854	160	870	112	781	54	930	925	29125
																					72%

Figuur 21: Schatting van het percentage dat een sputter effectief productieloten sputtert.

Engineeringwafers maken typisch ongeveer 5% uit van het totaal aantal wafers. Dat is 3.7% van de totale tijd.

Voor het inschatten van de tijd die besteed wordt aan conditioneren en de idle time zijn er geen gegevens beschikbaar. Als van de door ON berekende uptime (89%) de schattingen voor de totale procestijd (75% plus 3.7%) afgetrokken worden, blijkt het verschil 10.3% te zijn. Voor de idle time zal daarom een schatting van 5% gebruikt worden en voor de conditioneringstijd 5.3%.

De verschillende elementen bij downtime worden geschat op 8% voor preventief onderhoud (5% time based en 3% counterbased) en 3% voor onverwachte stilstand. Tabel 7 geeft een samenvatting.

UP	processing voor productie	75 %
	processing voor engineering	3.7 %
	conditionering	5.3 %
	idle time	5 %
DOWN	preventief onderhoud – counterbased	5 %
	preventief onderhoud – timebased	3 %
	Onverwachte stilstand (correctieve actie)	3 %

Tabel 7: De elementen voor observatie met geschatte percentages.

2.3.4. Bepalen van gewenste nauwkeurigheid + aantal observaties

				n(z=1,96)	vaste n	A (z=1,96)
UP	processing voor productie	75%	0,030	800	800	3,0%
	processing voor engineering	4%	0,015	608	800	1,3%
	conditionering	5%	0,015	857	800	1,6%
	idle time	5%	0,015	811	800	1,5%
DOWN	preventief onderhoud – counterbased	5%	0,015	811	800	1,5%
	preventief onderhoud – timebased	3%	0,015	497	800	1,2%
	correctieve actie	3%	0,015	497	800	1,2%

Tabel 8: Bepalen van de nauwkeurigheid bij 800 observaties

Tabel 8 bevat in de eerste twee kolommen opnieuw de verschillende toestanden waarin een sputter zich kan bevinden. In de 3^{de} kolom staan de in 2.3.3 geschatte percentages. De vierde kolom geeft de halve breedte van het betrouwbaarheidsinterval aan. Deze breedtes zijn gekozen naargelang van de grootte van het geschatte percentage. Het 95%-betrouwbaarheidsinterval is dan [p-a, p+a]. Dat wil zeggen dat de (in de Work Sampling) bekomen percentages met 95% zekerheid tussen de grenzen van dat interval liggen.

$$a = z * \sigma_p = z * \sqrt{\frac{p * (1 - p)}{n}}$$

Met behulp van bovenstaande formule wordt de vijfde kolom berekend. Dat zijn met andere woorden het aantal observaties dat nodig is om de vooropgestelde betrouwbaarheidsintervallen te halen. Met 800 observaties blijken alle elementen ongeveer de gewenste nauwkeurigheid te halen. Door het vastleggen van het aantal observaties (kolom 6), zal de breedte van het betrouwbaarheidsinterval voor sommige elementen wijzigen (kolom 7).

2.3.5. Vastleggen van tijden en routes

De observaties dienen over een voldoende lange periode gespreid te zijn. Bij preventief onderhoud van kamer A, kan de volledige sputter al snel 10 uur onbeschikbaar zijn. Van dag tot dag de uptime berekenen zou dus totaal niet relevant zijn. Er is gekozen om de 800 observaties te verspreiden over een volledige maand. Februari lijkt daarvoor ideaal. Als een dag ingedeeld wordt in tijdsintervallen van 10 minuten, dan zijn er 24uur x 6intervallen/uur mogelijke slots per dag, ofwel 28 dagen x 24 uur/dag x 6 slots/uur = 4032 slots. Als daarbinnen 800 observaties gedaan worden, is dat dus het random kiezen van 1 op 5 loten. Dit is uitgevoerd in Excel met de functie “rand”.

2.3.6. De resultaten

Tijdens de observaties bleek dat twee extra categorieën nodig waren voor de situatie waar één of twee kamers niet beschikbaar waren en

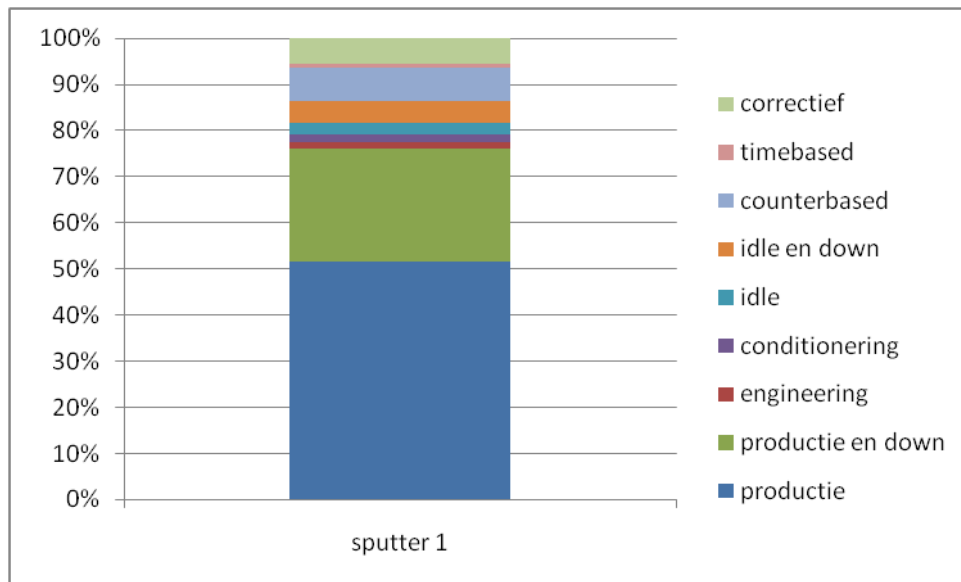
- de sputter toch aan het werk was
- de sputter idle was

Verder zijn de twee soorten preventief onderhoud samengevoegd tot één, aangezien het onderscheid soms moeilijk te maken was. De resultaten zijn samengevat in Tabel 9 en visueel voorgesteld in Figuur 22.

		p	a	p - a	p + a
UPTIME	productie	51,6%	3,5%	48,2%	55,1%
	productie + kamers down	24,4%	3,0%	21,4%	27,4%
	engineering	1,5%	0,8%	0,7%	2,3%
	conditionering	1,5%	0,8%	0,7%	2,3%
	idle + kamers down	4,8%	1,5%	3,3%	6,2%
	idle	2,5%	1,1%	1,4%	3,6%
	TOTAAL	86,3%	2,4%	83,9%	88,6%
DOWNTIME	preventief onderhoud	8,1%	1,9%	6,2%	10,0%
	Onverwachte breakdown	5,5%	1,6%	3,9%	7,1%
	TOTAAL	13,6%	2,4%	11,2%	16,0%

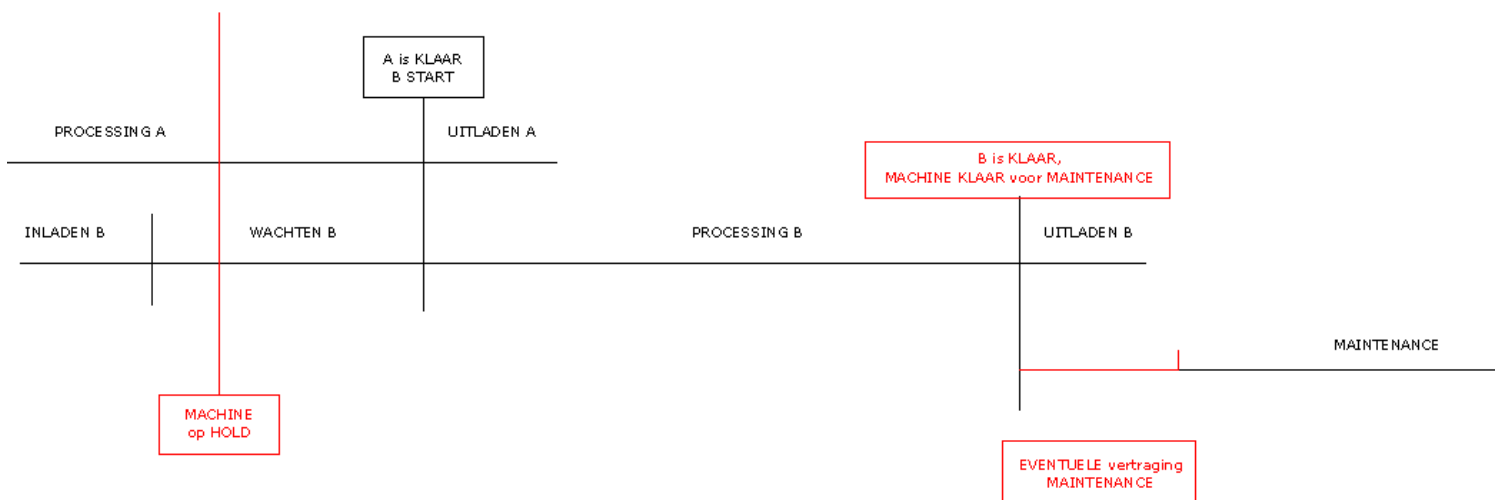
Tabel 9: Resultaten van de work sampling op sputter 1.

De totale uptime ligt met 95% zekerheid tussen 83.9% en 88.6% en de downtime tussen 11.2% en 16,0%. Wat echter onmiddellijk opvalt is dat minder dan 60% van de tijd de sputter volledig gebruikt kan worden. In de andere kleine 30% dat ze beschikbaar is, is dit niet voor alle operaties het geval. Verder blijkt het aandeel engineering kleiner dan verwacht. Ook conditionering lijkt op het eerste gezicht niet zo veel tijd in beslag te nemen.



Figuur 22: Tijdsbesteding sputter 1.

Deze manier van Work Sampling heeft een klein nadeel. De machine wordt soms al op HOLD gezet terwijl ze nog de laatste loten aan het processen is, gewoon om te voorkomen dat er nog een nieuw lot opgestart wordt. Hiermee werd rekening gehouden in de analyse (de eindtijd van elk lot was bekend), er werd dus niet te veel downtime aangerekend. Wel is het zo dat het onderhoud niet steeds onmiddellijk na het laatste lot gestart wordt. De sputter staat dan al op HOLD maar eigenlijk is het idle time. Figuur 23 stelt dit visueel voor. Als er effectief een observatie van de machine zou plaats vinden op de werkvloer, zou dat probleem zich niet voordoen.



Figuur 23: Mogelijke verborgen idle time (grotere figuur in bijlage E).

Afsluitend is het interessant om de uptime uit de Work Sampling - 86.3% voor het gemak - te vergelijken met de door ON berekende uptime, 79.7%. Toch wel een merkelijk verschil. Dat valt deels te verklaren doordat de berekening van ON de machine snel als onbeschikbaar zal beschouwen wanneer een belangrijke kamer uitvalt. In de Work Sampling is een aparte categorie voorzien voor die situatie "idle en maintenance". Deze valt onder uptime. Dat is een bewuste keuze aangezien het vermeden zou moeten worden dat een sputter zonder WIP valt, zelfs in geval van panne. Maar zelfs als die categorie "idle maar niet alle operaties beschikbaar" deels onder downtime zou vallen, is er nog een verschil dat groter is dan een half betrouwbaarheidsinterval.

Mogelijk moeten de gewichtsfactoren in de berekening door ON inderdaad geüpdate worden. Maar nog beter zou zijn enkele extra zaken te registreren (zoals bijvoorbeeld start en eind van setupwafers en SPC-wafers) en dan een systeem te ontwerpen dat alle verschillende log-files samenvoegt zodat de proporties van de verschillende machinetoestanden bijna live weergegeven kunnen worden.

3. Optimaliseren van de sputtergroep

In dit vierde deel van het tweede hoofdstuk wordt eerst de huidige werkwijze op de sputters besproken. Daarna volgt een overzicht van zaken die in de nieuwe dispatchstrategie opgenomen zouden moeten worden. Als laatste deel komt het voorstel voor deze nieuwe strategie, die de capaciteit van de sputters optimaal moet benutten en eventueel zelfs verhogen.

3.1. Huidig systeem

3.1.1. Technologisch

technologie	# operaties	sputter 1	sputter 2	sputter 3	sputter 4	moves maart
TiN	4					2190
AlSiCu	11					2608
Sandwich	6					2307
Barpl10	2					9
Glue	9					2965
Titaan	1					121
AlCu	20					3079
M2.5	1					510
TiAlCuTiN	2					0
Tasi	2					25

Figuur 24: De 10 processen die kunnen gesputterd worden met aantal operaties, op welke sputter ze kunnen en het aantal moves per week voor de maaand maart 2010.

Zoals eerder vermeld, heeft een sputter verschillende kamers en zijn de technologieën die op een sputter geprocesseerd kunnen worden afhankelijk van de kamercombinatie. In Figuur 24Error! Reference source not found. staan de 10 technologieën bij elkaar. De tweede kolom duidt het aantal operaties aan die met die bepaalde technologie werken. De kolommen met de kleuren tonen aan op welke sputters de technologieën mogelijk zijn; rood staat voor onmogelijk, geel is gedeeltelijk mogelijk (sommige operaties wel, andere niet) en

groen betekent volledig mogelijk. De laatste kolom geeft (met als voorbeeld de maand maart) een idee over het aantal moves dat er voor elke technologie moet gebeuren en vooral de onderlinge verhouding.

3.1.2. Dispatch

De operatoren hebben in het oude systeem twee bronnen als input.

- Enerzijds is er het overzicht van loten in het programma Workstream. Vereenvoudigd komt het erop neer dat bovenaan de *super hot* loten staan. Daaronder staan de *hot*, *zeer dringende* en *vlugge* loten, vanaf dat ze langer dan respectievelijk 4, 6 en 8 uur aan de sputters wachten. Daarna hangt de volgorde af van de toestand (setup) waarin de sputter staat. Als de machine een lot aan het sputteren is met Titanium Nitride (TiN) bijvoorbeeld, dan zullen alle loten die klaar staan en ook een TiN-laag nodig hebben, naar boven schuiven.
- De tweede input voor de operatoren is de Excel-tool WP02 (Werkpost 2). Deze toont een groot aantal extra zaken waarmee de operatoren al dan niet rekening kunnen houden (vrijblijvend): de status van de counters (zie 2.1.3), de status van de paste counters (extra counters voor een beperkt aantal operaties), welke kamers UP en welke DOWN zijn, hoeveel WIP er staat per operatie en per technologie en welk preventief onderhoud er gepland staat.

Dit systeem laat de operatoren een enorme vrijheid. Wie het goed door heeft, kan aan de hand van de Excel-tool relatief goede keuzes maken over wanneer welke loten te processen. Slechts een deel van de operatoren gebruikt de tool echter en zelfs als ze gebruikt wordt, dan worden soms andere doeleinden nagestreefd dan optimaal capaciteitsgebruik. Men zal bijvoorbeeld zo weinig mogelijk SPC-testen uitvoeren door te veranderen van operatie na bvb 299 wafers (net geen 300).

Links in Figuur 25 staan in het geel alle kamers (6 per sputter) met hun status “up” of “down”. Het middenveld van de figuur geeft per operatie weer wat voor gevolgen de kamerstatussen hebben. Voor elke operatie kan de operator zien op welke sputter ze (momenteel) kan uitgevoerd worden en hoeveel wafers voorraad er staan (in de kolom WIP). Rechts in de figuur staan de counters voor SPC-testen.

UPDATE			SPUTTER1		SPUTTER2		SPUTTER3		SPUTTER4		WIP	COUNTER	STATUS	WIP	PASTE	PASTE	MAK	MAK	MAK
Run op de knop hierboven voor refresh			Materiaal	Operator	combinatie	VVP?	combinatie	VVP?	combinatie	VVP?	combinatie	VVP?	10-41						
Alu sputter 1A daunt u kan er niet op voort varen!!! Indien ja 1A de daarvoor kamers gebruiken. T bij daarvoor (y-ruimte - daarvoor) opgelat mechanische partijt moet ok zijn!!			TiH	1403	2	OK	2	OK	4	OK			0	SPUTTER11	UP	0			
				1406	2	OK	2	OK	4	OK			0	SPUTTER112	UP	228			
				1420	2	OK	2	OK	4	OK			397	SPUTTER12	UP	50			
				1440	2	OK	2	OK	4	OK			0	SPUTTER121	UP	212			
			AlSiO ₂	1401	D412	OK	D412	OK	D414	OK			100	SPUTTER14	UP	0	75	98	
				1402	D412	OK	D412	OK	D414	OK			0	SPUTTER142	UP	299			
				1405	D412	OK	D412	OK	D414	OK			0	SPUTTER143	UP	0			
				1407	D412	OK	D412	OK	D414	OK			0	SPUTTER22	UP	204			
				1412	D412	OK	D412	OK	D414	OK			0	SPUTTER223	UP	239			
				1425	D412	OK	D412	OK	D414	OK			0	SPUTTER23	UP	237			
				1426	D412	OK	D412	OK	D414	OK			0	SPUTTER232	UP	273			
				1427	D412	OK	D412	OK	D414	OK			0	SPUTTER24	UP	50	150	140	
				1460	D41	OK	D41	OK	D41	OK			50	SPUTTER241	UP	326			50
				1461	D41	OK	D41	OK	D41	OK			123	SPUTTER242	UP	247			
				1462	3	HEEN	1	OK			4	OK	0	SPUTTER31	UP	98			
			Sandwich	1408	D412	OK	D412	OK					0	SPUTTER312	UP	112			
				1409	D412	OK	D412	OK					0	SPUTTER32	UP	149			
				1410	D412	OK	D412	OK					50	SPUTTER334	UP	50			
				1416	D412	OK	D412	OK					0	SPUTTER34	UP	200			
				1444	D412	OK	D412	OK					100	SPUTTER42	UP	125	148	140	
				1445	D412	OK	D412	OK					99	SPUTTER4234	UP	133			
			Borpl10	1428	21	OK	21	OK					0	SPUTTER424	UP	229			179
				1429	21	OK	21	OK					0	SPUTTER434	UP	108			
			Gluc	1431	D412	HEEN	C411-D411	OK			D424	OK	0	SPUTTER44	UP	104			
				1432	D412	HEEN	C411-D411	OK			D424	OK	0	SPUTTER441	UP	299			
				1433	D412	HEEN	C411-D411	OK			D424	OK	0						
				1434	D412	OK	C411-D411	OK			D424	OK	0						
				1438	41	OK	41	OK			24	OK	93						
				1439	41	OK	41	OK			24	OK	0						
				1447	C411-D411	OK	C411-D411	OK			D424	OK	0						
				1448	C411-D411	OK	C411-D411	OK			D424	OK	0						
				1457	C411-D411	OK	C411-D411	OK			D424	OK	0						
			Titaan	1411	4	OK	4	OK			2	OK	0	SPUTTER1A	SWI3573	SWI3593	1431 (POLY2_OC_ON_POLY1)	118W	118W
			Alcu	1413	12	OK	12	OK			34	OK	0	SPUTTER2	1408,1409	SWI3593	1408,1409	218W	118W
				1415	12	OK	12	OK			34	OK	0	SPUTTER4	1452/PROC/ERRORS	1456/PROC/H3T50-DRYSPOT	118W	118W	118W
				1417	12	OK	12	OK			34	OK	0	WETLJN1	0	0	0	0	0
				1418	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1419	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1420	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1421	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1425	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1426	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1427	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1441	12	OK	12	OK			34	OK	25	0	0	0	0	0	0
				1442	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1443	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1451	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1453	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1454	12	OK	12	OK			34	OK	4	0	0	0	0	0	0
				1455	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1456	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1458	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1459	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
			M2.5	1452	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
			Ti-AlCu-TiN	1422	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1450	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
			Tari	1404	12	OK	12	OK			34	OK	0	0	0	0	0	0	0
				1414	12	OK	12	OK			34	OK	0	0	0	0	0	0	0

Figuur 25: Excel-tool WP02.

3.1.3 Onderhoud

Preventief onderhoud wordt op dit moment gepland op vaste tijdstippen na bepaalde ingrepen. Als er bijvoorbeeld een cryoregeneratie gebeurt, dan zal exact 7 dagen later een onderhoud gepland worden. Er wordt met andere woorden geen enkele rekening gehouden met de productieplanning bij het inplannen van preventief onderhoud.

3.2. Belangrijke punten voor het nieuwe systeem

Om de capaciteit optimaal te benutten, zijn er een groot aantal zaken die in aanmerking genomen kunnen en moeten worden. Deze zijn opgedeeld in enkele categorieën: maatregelen die de procestijd verkorten, maatregelen die de tijd voor conditioneren verkorten, maatregelen die zorgen dat er geen enkele sputter zonder werk komt te staan en maatregelen die de beschikbare tijd vergroten door betere organisatie.

3.2.1. Maatregelen die de procestijd verkorten.

1. Toewijzen van de loten aan de sputters die deze het snelst kunnen processen.

Doordat de sputters allemaal andere kamercombinaties hebben, kunnen sommige operaties sneller op de ene sputter gebeuren dan op de andere. Het is evident de snellere sputter voorkeur te geven.

2. Veranderen van bepaalde kamers zodanig dat sommige operaties sneller gaan.

Een wafer moet in elke kamer een bepaalde tijd verblijven, afhankelijk van de dikte van de laag die gesputterd wordt. Als de wafer beduidend langere tijd in één kamer moet blijven dan in de andere, is het misschien interessant om van die soort kamer er twee te voorzien. De helft van de wafers kan dan naar de ene en de andere helft naar de andere kamer. Hierdoor wordt de procestijd sterk verkort. Dat gaat wel ten koste van een andere kamer natuurlijk, waardoor een aantal operaties niet meer uitgevoerd kan worden. De flexibiliteit verlaagt dus, om de snelheid te verhogen. Dit proces noemt men het ontdubbelen van kamers.

3. Idle time minimaliseren.

Uit de Work Sampling bleek dat er toch meer dan 7% idle time was, waarvan 2,5% dat alle kamers beschikbaar waren. 7% is op 1 week bijna 12 uur en dus 1,5 shift. In die tijd kan heel wat werk verzet worden.

3.2.2 Maatregelen die de tijd voor conditioneren verkorten.

4. Vermijden van te veel setups.

In het huidige systeem wordt soms bijna willekeurig omgeschakeld. Het aantal setups verminderen, betekent meer tijd dat de machines effectief kunnen sputteren. Het is echter niet de bedoeling van de batches enorm groot te maken en zo setups te vermijden. Dat zou bottlenecks verderop in de lijn in de problemen kunnen brengen. Lang dezelfde operaties uitvoeren, betekent ook lang andere operaties niet uitvoeren terwijl een machine stroomafwaarts aan het leeglopen is.

5. Tijd van conditioneringen naar beneden brengen (SMED).

Omschakelen of conditioneren doet de operator door een conditioneringswafer in een waferbox te plaatsen en deze net zoals een gewoon lot in een loadblock te steken. Daar komt dus een inlaadtijd (vacuümpompen), procestijd en uitlaadtijd bij kijken. Het eventuele inkorten van het recept waarmee de wafer wordt behandeld, valt buiten het bestek van deze thesis.

6. Combineren van setups met SPC's.

Zoals besproken in (2.1.3.) zijn er SPC-testen nodig telkens er 300 wafers door een bepaalde kamercombinatie gaan. Als deze onderling gecombineerd kunnen worden en met de conditioneringswafer, betekent dat een vermindering van het aantal waferboxen die slechts één wafer bevatten. Dat bespaart dan ook veel in- en uitlaadtijd.

3.2.3. Maatregelen die leegloop vermijden.

7. Balanceren van de belasting op de verschillende sputters.
8. Zorgen dat er steeds de juiste voorraad staat.

Punt 7 en 8 hebben tot doel te vermijden dat één sputter zonder werk staat terwijl de andere nog voorraad hebben. Met het huidige systeem kan het gebeuren dat een sputter bijna zonder werk zit, zonder dat een maatregel verhindert dat de andere sputters de laatste voorraad van de leeglopende sputter wegwerken. De WP02 Excel-tool komt aan dat probleem tegemoet, maar wordt niet (altijd) goed gebruikt.

3.2.4. Maatregelen die in verband met onderhoud en technologische beperkingen.

9. Toewijzen van operaties met complementaire kamercombinaties aan één sputter.

In het geval (bepaalde) operaties aan een vaste sputter gedediceerd worden, is het interessant om ervoor te zorgen dat deze operaties complementaire kamers gebruiken. Als er dan een kamer uitvalt, is het nog altijd mogelijk verder te processen met de andere gedediceerde operaties.

10. Preventief onderhoud plannen als er niet veel WIP staat.

Als een bepaalde sputter ondanks alle voorzorgen uitzonderlijk toch (bijna) geen voorraad meer heeft, kan er bekeken worden welk onderhoud in de nabije toekomst voorzien was dat vervroegd kan worden.

11. Mainframe onderhoud combineren met kameronderhoud.

Als het mainframe of kamer A in onderhoud is, wil dat zeggen dat de hele sputter buiten gebruik is. Het is ook hier interessant om te bekijken of het onderhoud van één of meer kamers niet wat vervroegd kan worden zodat dit samen kan gebeuren.

3.3. Voorstel nieuw systeem

3.3.1. Operartiegroepen

De bedoeling is om de 58 operaties op te delen in 12 operatiegroepen die dezelfde mogelijkheden, voorkeuren en behandeling vergen. Concreet wil dat zeggen dat binnen één operatiegroep alle operaties

1. uitgevoerd kunnen worden op dezelfde sputters en dus ook op dezelfde niet mogelijk zijn.
2. dezelfde voorkeur hebben voor bepaalde sputters omwille van processnelheid (sommige operaties kunnen sneller uitgevoerd worden op één specifieke sputter, omwille van de kamerconfiguraties).
3. dezelfde conditioneringseisen hebben (zodat geen setups nodig zijn zolang loten uit dezelfde groep gesputterd worden)

Hier en daar wordt een uitzondering toegelaten. Deze zijn goedgekeurd door de procesingenieurs, teamleaders en mevrouw Vercaemst.

Dit opdelen in operatiegroepen heeft verschillende voordelen op het gebruik van de losstaande operaties.

1. Het systeem oogt en werkt veel eenvoudiger met 12 groepen in plaats van 58 operaties. Het is voor alle betrokkenen gemakkelijk te volgen wat er gebeurt.
2. Als er een nieuwe operatie bijkomt, is het eenvoudiger die toe te voegen aan een operatiegroep waarrond het systeem gebouwd is, dan dat heel het systeem aangepast moet worden.
3. De operaties met gelijke conditioneringseisen worden op deze manier samen gehouden, wat het aantal setups verkleint. Ook het samenhouden van gelijke operaties zonder conditioneringseisen heeft voordelen; SPC-testen en preventief onderhoud zullen regelmatig en volgens een planning kunnen gebeuren (dit wordt in 3.3.3. verder uitgelegd)

Deze 12 operatiegroepen zijn eigenlijk een uitdieping en herschikking van de reeds bestaande groepen op basis van proces (TiN, AlCu, ...). Bij het overlopen van de nieuwe operatiegroepen wordt ook verwezen naar die oude groepen, ter verduidelijking voor de lezer die er voorheen mee werkte.

TiN_pref is een groep met drie operaties die elk gelijke conditioneringseisen hebben en die mogelijk zijn op sputter 1, 2 en 3. Ze gebeuren echter significant sneller op sputter 1. Operatie 1440, nochtans ook van de TiN-technologie hoort op vlak van conditionering in de groep GlueTit_pref. Op vlak van technologische beperking komt ze niet volledig overeen met GlueTit_pref (de operatie kan op 2 en 3, terwijl de andere operaties van GlueTit_pref op 2 en 4 kunnen). Na overleg met de procesingenieurs, teamleaders en mevrouw Vercaemst wordt de uitzondering echter toegestaan. De GlueTit_pref wordt namelijk gedicteerd aan sputter 2, wat ook een mogelijkheid is voor operatie 1440.

Group	Operation	WIP+coming	1	2	3	4	1	2	3	4
TiN_pref	1403	0	1,06	1,6	2,62		0	0	0	0
TiN_pref	1406	0	1,06	1,6	1,6		0	0	0	0
TiN_pref	1430	100	1,38	2,09	2,09		138	209	209	0
GlueTit_pref	1431	10		2,49		3,74	0	25	0	37
GlueTit_pref	1432	49		2,49		3,74	0	122	0	183
GlueTit_pref	1433	25		2,49		3,74	0	62	0	94
GlueTit_pref	1434	25		2,45		2,45	0	61	0	61
GlueTit_pref	1447	75		2,49		2,49	0	187	0	187
GlueTit_pref	1448	197		2,49		2,49	0	491	0	491
GlueTit_pref	1457	0		2,49		3,74	0	0	0	0
GlueTit_pref	1411	0		1,84		1,84	0	0	0	0
GlueTit_pref	1440	0		1,9	1,9		0	0	0	0
GlueTit_pref	1438	0		1,8		1,8	0	0	0	0
GlueTit_pref	1439	0		1,9		1,9	0	0	0	0

Figuur 26: De operatiegroepen TIN_pref en GlueTit_pref met de tijd per wafer per sputter en de runtijd van de WIP per operatie.

GlueTit_pref bevat operaties die hetzelfde conditioneringsrecept vereisen. Het is een samenvoegsel van de oorspronkelijke groepen Glue (10 operaties) en Titaan (1 operatie: 1411) met operatie 1440. Operatie 1411 (of de Titaantechnologie) is hierbij toegevoegd omwille van gelijke conditionering, dat wil zeggen dat er niet opnieuw geconditioneerd moet worden om operatie 1411 uit te voeren nadat één van de andere operaties uit de groep geprocessed werd. Een aantal operaties binnen de groep kunnen sneller uitgevoerd worden op sputter 2 dan op sputter 4. De tijden in het geel zijn recent aangepast (kleiner geworden) door het aanpassen van de kamerconfiguratie van sputter 1.

De groep AlCu condit bevat de operaties van de vroegere groep AlCu die conditioneringseisen hebben. Deze groep bevat geen uitzonderingen. Het lijkt dat de meeste operaties sneller gaan op sputter 3 dan op 4. Bij het voorstellen van deze groepen in ON Semiconductor, werd echter ontdekt dat deze tijden fout geregistreerd zijn. Nieuwe metingen zijn hier nodig, het is evenwel zeker dat wafers uit de operatiegroep even snel uitgevoerd kunnen worden op beide sputters gezien de gelijke kamerconfiguratie. De groep AlCu_gew bevat de overige operaties uit de vroegere AlCu-technologiegroep. Deze zijn ook even snel op sputter 3 en 4 en vereisen geen conditionering.

AlCu_condit	1413	0			2,03	2,03	0	0	0	0
AlCu_condit	1415	75			2,03	2,03	0	0	152	152
AlCu_condit	1418	0			2,03	2,03	0	0	0	0
AlCu_condit	1419	0			2,03	2,18	0	0	0	0
AlCu_condit	1420	0			2,18	2,18	0	0	0	0
AlCu_condit	1435	0			2,43	2,43	0	0	0	0
AlCu_condit	1436	0			1,88	1,88	0	0	0	0
AlCu_condit	1437	25			2,03	2,03	0	0	51	51
AlCu_condit	1441	194			2,03	2,03	0	0	394	394
AlCu_condit	1442	0			2,03	2,03	0	0	0	0
AlCu_condit	1443	0			2,43	2,43	0	0	0	0
AlCu_condit	1451	0			2,72	2,72	0	0	0	0
AlCu_condit	1456	48			3,74	3,74	0	0	180	180
AlCu_condit	1458	0			1,88	1,88	0	0	0	0
AlCu_condit	1452	0			2,23	2,23	0	0	0	0

Figuur 27: Operatiegroep AlCu_condit met de tijd per wafer per sputter en de runtijd van de WIP per operatie.

AlCu_gew	1421	0			1,9	1,90	0	0	0	0
AlCu_gew	1417	0			4,16	4,16	0	0	0	0
AlCu_gew	1453	0			2,23	2,23	0	0	0	0
AlCu_gew	1454	74			2,23	2,23	0	0	165	165
AlCu_gew	1455	0			3,74	3,74	0	0	0	0

Figuur 28: Operatiegroep AlCu_gew met de tijd per wafer per sputter en de runtijd van de WIP per operatie.

TiAlCuTiN is een groep die slechts 2 operaties bevat. Ze kan enkel op sputter 4 geprocessed worden. TASI kan ook enkel op sputter 4 geprocessed worden, toch zijn het aparte groepen aangezien TiAlCuTiN conditioneringseisen stelt en TASI niet. Die eisen zijn echter dezelfde als die van AlCu. Samenvoegen is echter geen optie gezien de grote AlCu-groep zowel op sputter 3 als 4 kan en TiAlCuTiN enkel op sputter 4.

TiAlCuTiN	1422	0				1,88	0	0	0	0
TiAlCuTiN	1450	4				3,74	0	0	0	15

Figuur 29: Operatiegroep TiAlCuTiN met de tijd per wafer per sputter en de runtijd van de WIP per operatie.

TASI	1459	25				4,16	0	0	0	104
TASI	1404	24				1,77	0	0	0	42
TASI	1414	0				1,77	0	0	0	0

Figuur 30: Operatiegroep TASI met de tijd per wafer per sputter en de runtijd van de WIP per operatie.

De vroegere op proces gebaseerde groep AlSiCu is opgedeeld in AlSiCu 1 en AlSiCu 2. Deze groep heeft geen conditioneringseisen en gaat overal even snel. AlSiCu 2 kan echter niet op sputter 1 geprocessed worden, waar AlSiCu 1 dat wel kan. Er wordt hier geen uitzondering gemaakt. AlSiCu is namelijk één van de flexibele groepen die gemakkelijk verschoven kan worden naar een sputter die te weinig WIP heeft staan. Als van die ene groep twee

kleinere groepen gemaakt worden, biedt dat nog meer mogelijkheden tot schuiven en komt dat de flexibiliteit dus ten goede.

AlSiCu_1	1401	399	2,62	2,62	2,62		1045	1045	1045	0
AlSiCu_1	1405	0	2,25	2,25	2,25		0	0	0	0
AlSiCu_1	1425	0	2,1	2,1	2,1		0	0	0	0
AlSiCu_1	1426	0	2,1	2,1	2,1		0	0	0	0
AlSiCu_1	1427	0	2,3	2,3	2,3		0	0	0	0
AlSiCu_1	1460	0	2,46	2,46	2,46		0	0	0	0
AlSiCu_1	1461	0	2,46	2,46	2,46		0	0	0	0
AlSiCu_2	1402	0		2,62	2,62		0	0	0	0
AlSiCu_2	1407	0		1,9	1,9		0	0	0	0
AlSiCu_2	1412	1		2,62	2,62		0	3	3	0

Figuur 31: Operatiegroepen AlSiCu 1 en AlSiCu 2 met de tijd per wafer per sputter en de runtijd van de WIP per operatie.

De vroegere op technologie gebaseerde groep Sandwich is opgedeeld in Sandw en Sandw_pref. Geen van beide heeft conditioneringseisen, doch Sandw_pref wordt best toegewezen aan sputter 1, gezien de significant hogere snelheid.

Sandw	1408	0	2,78	2,78			0	0	0	0
Sandw	1409	0	2,78	2,78			0	0	0	0
Sandw	1410	0	2,78	2,78			0	0	0	0
Sandw	1416	0	2,25	2,25			0	0	0	0

Figuur 32: Operatiegroep Sandwich met de tijd per wafer per sputter en de runtijd van de WIP per operatie.

Sandw_pref	1445	0	1,95	2,46			0	0	0	0
Sandw_pref	1444	355	1,35	2,05			479	728	0	0

Figuur 33: Operatiegroep Sandwich_pref met de tijd per wafer per sputter en de runtijd van de WIP per operatie.

Tot slot blijven de twee groepen BARPL_pref en TiN_101 over. BARPL_pref wordt beter toegewezen aan sputter 1 omwille van de snelheid en heeft ook conditioneringseisen, jammer genoeg andere dan TiN die ook aan sputter 1 is toegewezen. TiN_101 is een uitzondering, deze operatie heeft een ander conditioneringsrecept dan alle andere operaties. Ze vormt dus een groep apart.

BARPL_pref	1428	0	1,9	2,71			0	0	0	0
BARPL_pref	1429	0	1,9	2,71			0	0	0	0
TiN_101	1462	78		2,71		2,71	0	211	0	211

Figuur 34: Operatiegroepen BARPL_pref en TiN_101 met de tijd per wafer per sputter en de runtijd van de WIP per operatie.

6 van deze operatiegroepen (TiN, GlueTit_pref, AlCu_condit, TiAlCuTiN, TiN_101 en Barpl10) bevatten operaties waarvoor een sputter geconditioneerd moet worden. De andere 6 bevatten operaties waarvoor dat niet nodig is.

Verder zijn er 5 van de 12 operatiegroepen die omwille van een significant hogere processnelheid op één van de sputters beter daaraan toegewezen worden. 2 van de 12 kunnen met de huidige kamerconfiguraties slechts op één bepaalde sputter uitgevoerd worden (sputter 4). De overige 5 operatiegroepen hebben keuze tussen 2 of 3 machines zonder speciale voorkeur.

3.3.2. Operatiegroepen toewijzen aan sputters.

De operatiegroepen worden toegekend (gedediceerd) aan bepaalde sputters. Voor 7 groepen (5 uit voorkeur en 2 uit noodzaak) is het onmiddellijk duidelijk welke sputter gekozen wordt (zie Figuur 35). Ook de overige operatiegroepen worden toegewezen aan een sputter, maar kunnen wanneer nodig verschuiven zonder dat dit nadelig is voor de processnelheid of om conditioneringseisen. Ze worden daarom flexibele groepen genoemd (zie Figuur 36). Aan welke sputter die 5 flexibele operatiegroepen toegewezen worden en wanneer ze moeten verschuiven, wordt bepaald door de WIP die voor de sputters (en voorafgaande cleaningoperaties) klaar staat. De dispatchregel dient de flexibele operatiegroepen toe te wijzen aan de sputters zodanig dat de WIP goed verdeeld is over de 4 machines maar vooral zodanig dat geen van de 4 in de risicozone komt voor leegloop (het verplicht idle zijn door gebrek aan WIP).

Het stopt uiteraard niet na het eenmalig toewijzen van de operatiegroepen aan sputters, gebaseerd op een bepaalde WIP-status. Er komen constant nieuwe loten aan en wafers die geprocesseerd worden, verdwijnen natuurlijk ook uit de sputter-WIP. Er is dus nood aan een herplanningsmechanisme; een systeem dat er voor zorgt dat de toewijzingen regelmatig opnieuw bekeken worden. Dit mechanisme is zowel tijds- als eventgestuurd.

Tijdstgestuurd: elk uur wordt een nieuwe WIP-status binnengehaald om de toewijzingen te evalueren.

Eventgestuurd: als een kamer (of volledige sputter) gedurende lange tijd onbeschikbaar zal zijn, kunnen bepaalde wafers niet meer uitgevoerd worden. Het mechanisme dient deze dan aan andere sputters toe te wijzen.

Deze WIP-status die door het softwarepakket APF wordt binnengehaald, is eigenlijk een lijst van de operatiegroepen met elk een runtijd. Deze is berekend volgens onderstaande formule:

$$\text{runtijd per operatiegroep} = \sum_{\substack{\text{alle operaties} \\ \text{in de operatiegroep}}} \text{aantal wafers in voorraad} \times \text{cycle time}$$

Let op: het is interessant om verder te kijken dan welke loten er effectief klaar staan voor de sputters, meer bepaald naar de loten voor de cleanoperatie die vooraf gaat aan het sputteren. Cleanoperaties duren niet lang en

worden vaak pas uitgevoerd als de operator al bepaald heeft wanneer ze op de sputter komen. Er bestaat immers een limiet op de tijd tussen cleanen en sputteren (vaak 8u, soms korter).

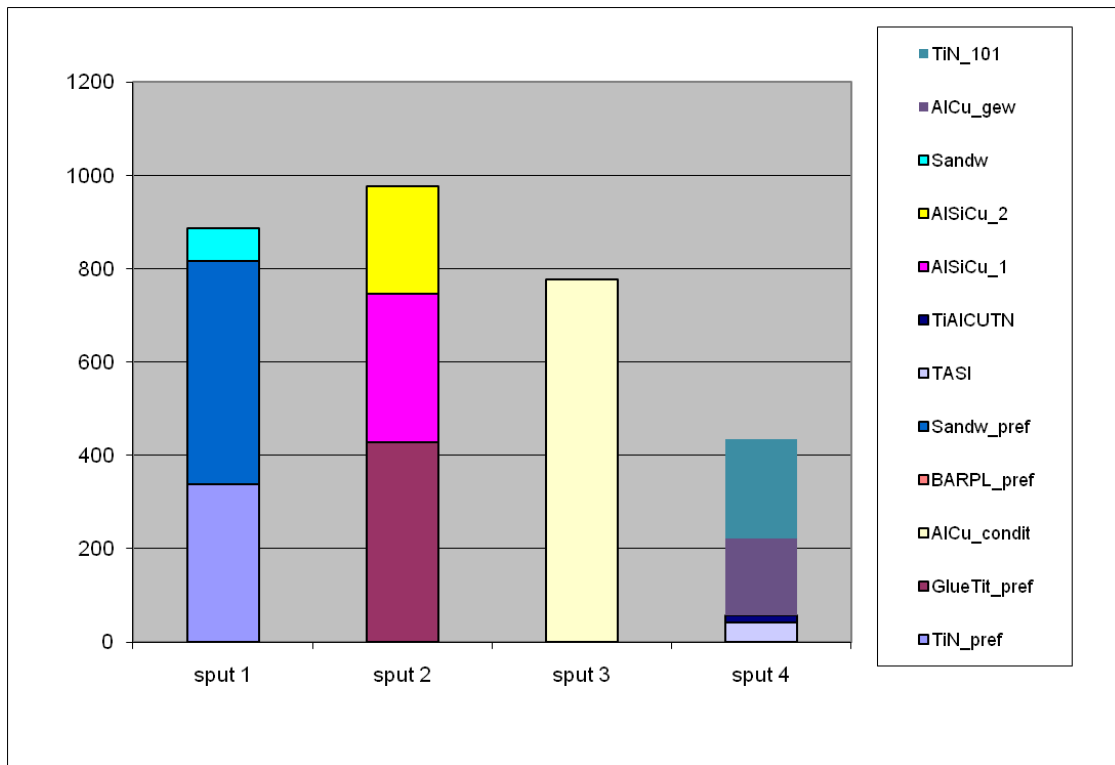
toewijzing aan	sput 1	sput 2	sput 3	sput 4	sput 1	sput 2	sput 3	sput 4		
van	(binair)				(run time)					
TiN_pref	1	0	0	0	337	0	0	0		vast door voorkeur
GlueTit_pref	0	1	0	0	0	427	0	0		kan ook
AlCu_condit	0	0	1	0	0	0	776	0		vast uit noodzaak
BARPL_pref	1	0	0	0	0	0	0	0		geen voorkeur
Sandw_pref	1	0	0	0	479	0	0	0		conditionering nodig
TASI	0	0	0	1	0	0	0	42		
TiAlCUTN	0	0	0	1	0	0	0	15		

Figuur 35: De 7 operatiegroepen die (al dan niet noodzakelijk) bij voorkeur op één bepaalde sputter uitgevoerd worden.

toewijzing aan	sput 1	sput 2	sput 3	sput 4	sput 1	sput 2	sput 3	sput 4		
van	(binair)				(run time)					
AlSiCu_1	0	1	0	1	0	320	0	0		vast door voorkeur
AlSiCu_2	0	1	0	0	0	231	0	0		kan ook
Sandw	1	0	0	0	70	0	0	0		vast uit noodzaak
AlCu_gew	0	0	0	1	0	0	0	165		geen voorkeur
TiN_101	0	0	0	1	0	0	0	211		conditionering nodig

Figuur 36: De 5 operatiegroepen waarvoor geen sputtervoorkeur bestaat.

Figuur 35 zijn de 7 groepen die gedicteerd worden (5 omwille van significant snellere procestijd en 2 omwille van technologische beperkingen). Figuur 36 zijn de 5 flexibele groepen die aan 2 of 3 sputters toegewezen kunnen worden en indien nodig verschuiven. De 4 kolommen met nullen en énen zijn binaire variabelen die een groep toewijzen (één) aan een sputter of niet (nul). De 4 kolommen rechts daarvan bevatten de runtijd per operatiegroep, berekend volgens de voorgestelde formule. Figuur 37 stelt deze runtijden per sputter visueel voor. De WIP-status die in de figuren getoond wordt, is van een willekeurig gekozen moment in april.



Figuur 37: Voorstelling van de runtijden van de operatiegroepen per sputter.

Een belangrijke opmerking over het verschuiven is dat dit enkel de toewijzing betreft. De fysieke loten worden niet van plaats verschoven bij een veranderde toewijzing. De voorraad voor de sputters staat namelijk op een gemeenschappelijk rek.

Hoe wordt nu beslist of een flexibele operatiegroep moet verschuiven naar een andere sputter of niet? Er zijn twee redenen die dit veroorzaken:

- als de flexibele groepen verschoven kunnen worden, zodat de verschillen in belasting tussen de sputters kleiner wordt dan voor het verschuiven.
- als de runtime voor een bepaalde sputter kleiner wordt dan een zekere grens.

De eerste reden zou nagegaan kunnen worden op verschillende manieren. Een set regels die bepalen dat onder bepaalde voorwaarden een bepaald blokje mag verschuiven bijvoorbeeld. Lineair programmeren is zeker en vast ook een optie. Bij ON wil men echter alles zoveel mogelijk automatiseren in APF, waarvoor het gemakkelijker is om gewoon de verschillende scenario's uit te rekenen. 4 flexibele operatiegroepen kunnen op 2 mogelijke sputters, 1 flexibele operatiegroep kan op 3 mogelijke sputters. Dat geeft 48 scenario's die de flexibiliteit van het systeem voorstellen.

De tweede reden die impliceert dat er geschoven moet worden met operatiegroepen is de minimum gewenste belasting per sputter. Om leegloop van de machine te vermijden, moet steeds een minimale hoeveelheid runtijd per sputter toegewezen zijn. Deze grens is voorlopig vastgelegd op 6 uur, ofwel 360 minuten. Eens het systeem in de praktijk gebruikt wordt, kan deze natuurlijk aangepast worden als dat nodig blijkt.

In Tabel 10 wordt per sputter een overzicht gegeven van de gedediceerde groepen, de flexibele groepen en tot slot ook de ‘mogelijke maar te vermijden’ operatiegroepen. Het is namelijk zo dat indien het toegewezen aantal minuten runtijd op een sputter onder de minimumgrens van 360 minuten komt, het mogelijk is dat geen enkele flexibele operatiegroep verschoven kan worden (bij zeer specifieke en uitzonderlijke situaties van de WIP-status). In dat geval zal om leegloop te vermijden, één serie van een gedediceerde operatiegroep ook kunnen verschuiven. Die zal dan bijvoorbeeld minder snel verwerkt worden, maar dat is uiteraard minder erg dan een machine die bij gebrek aan voorraad stil valt. Deze noodoplossing wordt in de onderste rij van Tabel 10 gevisualiseerd.

	SPUTTER 1	SPUTTER 2	SPUTTER 3	SPUTTER 4
GEDEDICEERD	TiN_pref	GlueTit_pref	AlCu_condit	TASI
	Barpl_pref			AlCuTiN
	Sandw_pref			
FLEXIBEL	AlSiCu_1	AlSiCu_1	AlSiCu_1	AlCu_gew
	Sandw	AlSiCu_2	AlSiCu_2	TiN_101
		Sandw	AlCu_gew	
		TiN_101		
MOGELIJK maar NADELIG		TiN_pref	TiN_pref	GlueTit_pref
		Barpl_pref		
		Sandw_pref		

Tabel 10: De sputters en hun mogelijke operatiegroepen.

3.3.3. Plannen van onderhoud, conditionering en monitoring.

Eén van de doelstellingen die de dispatchstrategie wil bereiken, is er voor zorgen dat het preventief onderhoud op kamers niet de hele sputter onbeschikbaar maakt. Daartoe moet dit goed afgestemd worden op de productie. De maximale duur van het preventief onderhoud is 12 uur. De bedoeling is dus om steeds gedurende minstens 12 uur dezelfde kamers te gebruiken en dus ook dezelfde kamers niet. Op die laatste kan dan onderhoud uitgevoerd worden zonder dat de productie daar problemen van ondervindt. Kamercombinaties om de 12 uur afwisselen, wil zoveel zeggen als operatiegroepen (met complementaire kamers) in blokken van een vast aantal loten laten sputteren. Het onderhoud van de kamers wordt dus gepland op basis van wat er geprocessed wordt terwijl dat in het oude systeem willekeurig was en ook vaak slecht uit kwam.

Een tweede doelstelling is het vermijden van te veel setups. Ook hier is het dus interessant om loten van een operatiegroep met conditioneringseisen, in blok na elkaar te plannen. Zo moet er slechts één keer geconditioneerd worden per serie loten van (minstens) 12 uur.

Een derde doelstelling is om de monitorwafers (waarop SPC-testen gebeuren) de productie niet continu te laten onderbreken. Als in een bloksysteem wordt gewerkt (zoals hier voorgesteld), dan is het een goed idee om alle monitorwafers samen door de sputter te laten lopen met de conditionering in één waferbox.

Dit bloksysteem wordt voorgesteld in Figuur 38, toegepast op sputter 2 met WIP-status voorgesteld in Figuur 37.

Box met wafers voor - Conditionering - SPC vr GlueTit - SPC vr eerdere kamercombo's	Serie van loten uit de operatiegroep GlueTit_pref.	Operatiegroepen toegewezen aan sputter 2, zonder conditionering.
	Start preventief onderhoud aan kamers die GlueTit_pref niet gebruikt.	Start preventief onderhoud aan kamers gebruikt door GlueTit_pref

Figuur 38: Tijdlijn van een sputter die toont hoe een serie gelijke loten kan samenvallen met onderhoud op de complementaire kamers.

In Figuur 37 is te zien dat GlueTit_pref ‘vast’ gedediceerd is aan sputter 2 en de flexibele groepen AlSiCu_2 en AlSiCu_1 eveneens toegewezen zijn. Wafers uit de operatiegroep GlueTit_pref gebruiken kamers 1 en 4. Wafers uit de operatiegroepen AlSiCu_1 en AlSiCu_2 gebruiken kamers 2 en 3. De dispatchregel zal er nu voor zorgen dat de AlSiCugroepen enerzijds en GlueTit_pref anderzijds elkaar afwisselen en dit in tijdsblokken van 12 uur, groot genoeg om preventief onderhoud uit te voeren op de niet gebruikte kamers (dit komt overeen met 12 loten). AlSiCu_1 noch AlSiCu_2 hebben conditioneringseisen; loten uit deze operatiegroepen moeten dus niet apart gehouden worden en mogen elkaar afwisselen. Met dit bloksysteem is ook ineens bereikt dat er slechts éénmaal per 12 loten GlueTit_pref een conditioneringswafer nodig is.

3.3.4. Dediceren versus flexibiliteit

Onder bepaalde omstandigheden zou het een goed idee zijn om met (al dan niet lineair) programmeren alle operatiegroepen toe te wijzen aan één vaste sputter. Een optimaal schema voor productie, onderhoud en monitoring kan dan opgesteld worden zodat geen enkele beschikbare minuut van de sputters verspild wordt. Verder zouden ook een aantal kamers op bepaalde sputters mogen verdwijnen, waardoor andere in de plaats kunnen komen die de processen versnellen (zie 3.2 over ontdubbelen van kamers). Hiervoor moet voldaan zijn aan volgende voorwaarden:

- de voorraad voor de sputters is voor elke operatie oneindig of de instroom van loten voor elke operatie is stabiel en gekend
- de machines zijn 100% betrouwbaar, dat wil zeggen geen (onvoorziene) breakdowns

De WIP is echter dynamisch en de breakdowns stochastisch. Als alle operatie(s)(groepen) gedediceerd zijn, dan zal de WIP ophopen bij de ene sputter terwijl een andere stil staat bij gebrek aan loten. Dat is een gevolg van de variatie in enerzijds die WIP en anderzijds de tijdstippen en duur van de breakdowns. Daarom wordt hier gekozen voor een systeem dat enkel de operatiegroepen dediceert die daardoor een voordeel krijgen en dus ook een groot stuk flexibiliteit bewaart.

3.3.5. Tijdsgebaseerd plannen

Het herplannen zou zowel continu als discreet kunnen gebeuren. Discreet lijkt hier echter een betere optie. De snelheid waarmee wafers aankomen is namelijk niet zo hoog. Voor de maand april 2010 bedraagt die 13814 moves per week, ofwel gemiddeld 82 wafers (3 à 4 loten) per uur voor de 4 sputters samen. Om een idee te krijgen van de impact van de nieuwe loten op de totale WIP-status is Tabel 11 opgesteld. Uitgaand van een gemiddeld aantal wafers per uur dat aankomt en een gemiddelde WIP die voor de sputters staat, wordt de impact berekend van de nieuw aangekomen wafers gedurende verschillende intervallen. Als er om de 30 minuten gecontroleerd zou worden, zal 2% van de wafers nieuw zijn. Dat is natuurlijk benaderend. Als de WIP die voor de sputters staat kleiner wordt, zal de impact van de aankomende wafers natuurlijk stijgen. Aan de hand van tabel 10 lijkt elk uur herplannen meer dan voldoende.

ON stelde voor in plaats van elk uur en bij machinepannes te herplannen, de operator te laten beslissen wanneer herplanning nodig is. Deze zou namelijk kunnen inschatten hoe lang een machinepanne zal duren en beoordelen of het nodig is om de wafers die niet meer verwerkt kunnen worden, te verdelen over de andere sputters. Dat heeft tot gevolg dat de strategie opnieuw niet meer noodzakelijk gebeurt op basis van vooropgestelde regels. Operators kunnen dan (al dan niet bewust) te lang wachten met het herplanningsalgoritme uit te voeren en zo capaciteit verloren laten gaan.

aankomst	82	wafers/uur								
gem. WIP	1200	wafers								
plus 6uur	492	wafers								
herplanningsinterval (min)			1	10	30	60	90	120	180	240
nieuw (%)			0%	1%	2%	5%	7%	10%	15%	19%

Tabel 11: Impact van de intervalgrootte op de WIP.

3.3.6. Samengevat

Door gelijkaardige operaties in operatiegroepen te verzamelen en het introduceren van een bloksysteem, is voldaan aan zo goed als alle punten die in 3.2 genoemd zijn. Het sleutelement is dat met dit bloksysteem telkens voor een zekere periode vast staat welke loten verwerkt zullen worden en welke kamers dus beschikbaar zijn of niet. Op basis daarvan kan dan zowel onderhoud als monitoring ingepland worden zonder verlies van capaciteit op de sputter.

.

Hoofdstuk 3: Samenvatting

Hoofdstuk 1 van deze masterproef startte met de literatuur over zowel Lean in het algemeen als Value Stream Mapping, een veel gebruikte techniek daarbinnen. Een korte inleiding over FABII van ON Semiconductor en enkele typische kenmerken van waferproductie maakten duidelijk dat de cleanroom in Oudenaarde in vele opzichten afwijkt van een doorsnee productieomgeving.

Uit de 20 productfamilies werd C07I gekozen om een Value Stream Map op te stellen. Om praktische redenen werd gewerkt met de data die ON ter beschikking had. Het zoeken hierin naar de meest correcte weergave van de situatie, nam enige tijd in beslag. De echte moeilijkheid lag echter enerzijds in het feit dat geen enkele machine toegewezen was aan de gekozen productfamilie en anderzijds in de vele gelijkaardige maar toch allemaal licht verschillende loops die wafers maken in de FAB. Voorraad in de value stream is de grootste vorm van verspilling (waste). Deze voorraad kan echter niet los gezien worden van de andere productfamilies en een future state map met Kaizen bursts voor de C07I-familie was daardoor dus geen meerwaarde. Gezien het ontbreken van de complexe relaties tussen de verschillende machinegroepen op de VSM en het beperkte zicht op de stroom van de WIP, was het resultaat ook van weinig nut voor deel twee van de masterproef.

Om aan deze tekortkomingen te verhelpen, werd onder andere gewerkt aan een verbeterde en gedigitaliseerde versie van het 'stift en papierwerk' in bijlage C. Uiteindelijk was het echter de Equipment Group Map (EGM) die, ook met het oog op hoofdstuk 2, uitgewerkt werd voor de sputtergroep en zijn relaties. Zowel de stroomopwaarts als stroomafwaarts gelegen machinegroepen worden voorgesteld, met daartussen de productstromen per familie.

In hoofdstuk twee wordt eerst een literatuurstudie uitgevoerd naar het plannen van een job shop omgeving. Speciale aandacht wordt besteed aan dispatchregels in de halfgeleiderindustrie. In dit tweede deel van de masterproef wordt namelijk mee gewerkt aan de ontwikkeling van een nieuwe dispatchstrategie in ON. Hierin krijgen de zwaarbelaste machinegroepen (bottlenecks genoemd) een aparte behandeling. Het doel was om een systeem te bedenken dat de capaciteit van de complexe sputtergroep optimaal benut.

In het voorgestelde systeem staan twee principes centraal. Het eerste zegt dat de 58 mogelijke operaties toegewezen worden aan welbepaalde sputters. Dat moet gebeuren op basis van de eigenschappen van de operaties enerzijds en de specifieke kamerconfiguraties van de verschillende sputters anderzijds. Om dit dediceren te vereenvoudigen, worden de operaties opgedeeld in 12 operatiegroepen waarbinnen voldaan is aan bepaalde criteria. Het eerste principe wijst dus ook de wafers toe aan één van de sputters, afhankelijk van de operatie die ze op dat moment nodig hebben. De volgorde waarin een sputter de loten in zijn wachtlijn

behandelt, wordt bepaald door het tweede principe. Dat zegt dat de operatiegroepen zodanig afgewisseld moeten worden, dat de productie noch door kameronderhoud noch door te vermijden conditioneringen onderbroken wordt. Enkele goed gekozen operatiegroepen worden als flexibel beschouwd en de toewijzing ervan kan wanneer nodig veranderen. Zo wordt de variatie in de toestromende WIP opgevangen en leegloop van een bepaalde sputter vermeden.

Half mei startte ON met de APF-implementatie van de dispatchstrategie voor de sputtergroep. Hoewel hier en daar met een kleine aanpassing, werd de in deze masterproef voorgestelde strategie tot op het moment van indienen nauwgezet gevolgd.

BIJLAGEN

A. De Value Stream Map van de back end van C07I

Deze is op geen enkele manier duidelijk af te drukken en bevindt zich dan ook op de CD-ROM die is toegevoegd aan deze masterproef.

B. De Equipment Group Map rond de sputters

Deze is te gedetailleerd om op een A4-formaat in te voegen en bevindt zich op de CD-ROM die is toegevoegd aan deze masterproef.

C. Voorstelling van de verschillende productfamilies en hun routes (back end en front end)

De verschillende flows met stift en papier waren te groot om toe te voegen. Foto's ervan bevinden zich op de CD-ROM die is toegevoegd aan deze masterproef.

D. Voorstelling van de verschillende productfamilies en hun productroutes (front end en back end)

De Excel-bladen waren niet duidelijk ingevoegd. Het bestand bevindt zich op de CD-ROM die is toegevoegd aan deze masterproef.

E. Voorstelling van het laadbloksysteem van de sputters

De figuur was te groot om duidelijk ingevoegd te kunnen worden. Ze bevindt zich op de CD-ROM die is toegevoegd aan deze masterproef.

F. Voorstelling van de verloren tijd na onderhoud

De figuur was te groot om duidelijk ingevoegd te kunnen worden. Ze bevindt zich op de CD-ROM die is toegevoegd aan deze masterproef.

REFERENTIES

- [1] J.K. LIKER, *Becoming Lean*, Productivity Press, New York, 2004
- [2] K. SHIMOKAWA en T. FUJIMOTO, *The Birth of Lean*, Lean Enterprise Institute Inc., 2009
- [3] J.P. WOMACK, D.T. JONES en D. ROOS, *The Machine That Changed the World*, Harper Perrenial, New York, 1991
- [4] A.P. SLOAN, *My years with General Motors*, Doubleday, New York, 1990
- [5] D. VAN GOUBERGEN, *Design of Manufacturing and Service Operations*, cursus bij het gelijknamige vak
- [6] T. MELTON, *The Benefits of Lean Manufacturing*, Chemical Engineering Research and Design, June 2005, 83(A6), 662-673
- [7] J.P. WOMACK en D.T. JONES, *Lean Thinking: Banish Waste and Create Wealth in your Corporation*, paperback edition, Simon and Chuster UK Ltd, London, 2003
- [8] Website Toyota Motor Corporation, http://www2.toyota.co.jp/en/vision/production_system/
- [9] SHIGEO SHINGO, *Zero Quality Control: Source Inspection and the Poka-Yoke System*, Productivity Press, 1986
- [10] D.T. HINES, P. HINES en N. RICH, *Lean Logistics*, International Journal of Physical Distribution and Logistics Management, 1997, Vol. 27, Issue 3/4, 153-173
- [11] F.H. ABERNATHY, J.T. DUNLOP, J.H. HAMMOND en D. WEIL, *A Stitch in Time: Lean Retailing and the Transformation of Manufacturing*, Oxford University Press, New York, 1999
- [12] R. CHALICE, *Improving Healthcare using Toyota Lean Production Methods*, Quality Press, Milwaukee, 2007
- [13] G. A. HOWELL, *What is Lean Construction?*, Proceedings IGLC-7 26-28 July 1999, University of California, USA
- [14] L. KOSKELA, G. HOWELL, G. BALLARD en I. TOMMELEIN, *The foundation of Lean Construcion*, Chapter 14 of *Design and Construction: building and value*, Elsevier Science Ltd, Oxford, 2002
- [15] N. BHATIA en J. DREW, *Applying Lean Production to the Public Sector*, The McKinsey Quarterly, Web Exclusive, June 2006
- [16] Z. RADNOR en P. WALLEY, *Learning to Walk Before We Try to Run: Adapting Lean for the Public Sector*, *Public Money en Management*, February 2008, Vol. 28, Issue 1, 13-20
- [17] R. SMITH en B. HAWKINS, *Lean maintenance; reduce costs, improve quality and increase market share*, Elsevier, Oxford, 2004
- [18] Website Lean Enterprise Institute, <http://www.lean.org>, *What is Lean?*
- [19] R. ANDERSSON, H. ERIKSSON en H. TORSTENSSON, *Similarities and differences between TQM, six sigma and lean*, The TQM Magazine, 2006, Vol. 18, Issue 3, 282-296

- [20] P. LINDBERG en A. BERGER, *Continuous Improvement: design, organization and management*, International Journal of Technology Management, 1997, Vol. 14, number 1, 86-101
- [21] C. BERGGREN, *Lean production – the end of history?* Work, Employment & Society, 1993, Vol. 7, Number 2, 163-188
- [22] FIET (the international trade union for white collar and professional workers), *Lean production and lean management: impact on white-collar workers, working conditions, collective bargaining and trade union strength*.
- [23] D. MEHRI, *Notes from Toyotaland*, Cornell University Press, New York, 2005
- [24] P.A. HINES, N. RICH en M. HOLWEG, *Learning to Evolve: A Review of Contemporary Lean Thinking*. International Journal of Operations & Production Management, 2004, Vol. 24 (10), 994-1011
- [25] S.S.AURAND en P.J. MILLER, *The operating curve: A method to measure and benchmark manufacturing line productivity*, Proceedings IEEE/SEMI-ASMC Advanced Semiconductor Manufacturing Conference Workshop, 1997, September 10–12, 391–397.
- [26] C.M. WEBER en A. FAYED, *Optimizing your position on the operating curve: How can a fab truly maximize its performance?*, IEEE Transactions on Semiconductor Manufacturing, February 2010, Vol. 23, Number 1
- [27] R.E. BOHN, *Noise and learning in semiconductor manufacturing*, Management Sciences, 1995, Vol. 41, number 1, 31–42
- [28] M. ROTHER en J. SHOOK, *Learning to See: Value Stream Mapping to Add Value and Eliminate Muda*, The Lean Enterprise Institute: Brookline, MA, 1999
- [29] J. LOVELLE, *Mapping The Value Stream*, IIE Solutions, February 2006, 26-33
- [30] D. TAPPING, T. LUYSTER en T. SHUKER, *Value Stream Management*, Productivity Press, New York, 2002
- [31] K.J. DUGGAN, *Creating Mixed Model Value Streams*, Productivity Press, New York, 2002
- [32] M. BRAGLIA, G. CARMIGNANI en F. ZAMMORI, *A new value stream mapping approach for complex production systems*, International Journal of Production Research, 15 September–1 October 2006, Vol. 44, 18–19, 3929–3952
- [33] T. McDONALD, E.M. VAN AKEN en A.F. RENTES, *Utilizing simulation to enhance value stream mapping: a manufacturing case application*. International Journal of Logistics, 2002, Vol. 5, 213–232
- [34] Y. LIAN en H. VAN LANDEGHEM, *An application of simulation and value stream mapping in lean manufacturing*, 14th European Simulation Symposium, 2002
- [35] Z.N. KHASWALA en S.A. IRANI, *Value Network Mapping (VNM): visualization and analysis of multiple flows in Value Stream Maps*, 2000
- [36] D. DE SPLENTER, *Introductiepakket - Trainingsdocumenten ON Semiconductor*, April 2008
- [37] Website GRT Consulting and Training, <http://grt77.com>
- [38] A.S. JAIN en S. MEERAN, *Deterministic job-shop scheduling: Past, present and future*, European Journal of Operational Research, 1999, Vol.113, 9390-9434

- [39] G.E. VIEIRA, J.W. HERRMANN en E. LIN, *Rescheduling manufacturing systems: a framework of strategies, policies and methods*, Journal of scheduling, 2003, Vol. 6, 39-62
- [40] S.M. JOHNSON, *Optimal two- and three-stage production schedules with setup times included*, Naval Research Logistics, 1954, Quarter I, 61-68
- [41] M.R. GAREY, D.S. JOHNSON en R. SEHTI, *The Complexity of Flowshop and Jobshop Scheduling*, Mathematics of operations research, Vol. 1, 2, May 1976, 117-129
- [42] M.L. FISHER en A.H.G. RINNOOY KAN, *The design, analysis and implementation of heuristics* (special issue). Management Science, 1988, Vol.34 (3), 263
- [43] F.A. RODAMMER en K.P. WHITE, *A recent survey of production scheduling*. IEEE Transactions of Systems, Management and Cybernetics, 1988, Vol. 18 (6), 841-851.
- [44] J. ADAMS, E. BALAS en D.ZAWACK, *The shifting bottleneck procedure for job-shop scheduling*. Management Science, 1988, 34 (3), 391-401
- [45] D.W. SELLERS, *Survey of approaches to the job shop scheduling problem*, April 1996, 28th Southeastern Symposium on System Theory, SSST; Baton Rouge, LA; USA; 396-400. 1996
- [46] J. BLAZEWICZ, W. DOMSCHKE en E. PESH, *The job shop scheduling problem: conventional and new solution techniques*, European Journal of Operational Research, 1996, Vol. 93, 1-33
- [47] J. RINTANEN, *AI Planning*, University of Freiburg, Duitsland
- [48] A.JONES en L.RABELO, *Survey of Job Shop Scheduling Techniques*, National Institute of Standards and Technology, Gaithersburg, MD, 1998
- [49] A.K. GUPTA en A.I. SIVAKUMAR, *Job shop scheduling techniques in semiconductor manufacturing*, International Journal of Advanced Manufacturing Technology, 2006, Vol.27, 1163–1169
- [50] W. S. GERE, *Heuristics in Job Shop Scheduling*, Management Sciences, 1996, Vol. 13, 167-190
- [51] S.S. PANWALKAR en W. ISKANDER, *A Survey of Scheduling Rules*, Operations Research, 1977, Vol. 25, 45-61
- [52] C. R. GLASSEY en M.G.C. RESENDE, *A Scheduling Rule for Job Release in Semiconductor Fabrication*, Operations Research Letters, 1988, Vol. 7, 213-217
- [53] C. LOZINSKI, C.R. GLASSEY, *Bottleneck Starvation Indicators for Shop Floor Control*, IEEE Transactions on Semiconductor Manufacturing, 1988, Vol. 1, 147-153
- [54] L. M. Wein, *Scheduling semiconductor wafer fabrication*, IEEE Transactions on Semiconductor Manufacturing, September 1988. Vol. 1, 202–215
- [55] J. R. JACKSON, *Simulation Research on Job Shop Production*, Naval Research in Logistics, 1957, Quart. 4, 287-295
- [56] R.M. DABBAS en J.W. FOWLER, *A New Scheduling Approach Using Combined: Dispatching Criteria in Wafer Fabs*, IEEE Transactions on semiconductor manufacturing, August 2003, Vol. 16, 3, 501

- [57] R. UZSOY, C. LEE en L. MARTIN-VEGA, *A review of production planning and scheduling models in the semiconductor industry, Part I: System characteristics, performance evaluation and production planning*, IIE Transactions on Scheduling and Logistics, 1992, Vol. 24, 47-61
- [58] M. MITTLER en A.K. SCHOEMIG, *Comparison of dispatching rules for semiconductor manufacturing using large facility models*, Proceedings of the 1999 Winter Simulation Conference
- [59] C.C. CERN en Y.L. LIU, *Family-Based Scheduling Rules of a Sequence Dependent Wafer Fabrication System*, IEEE Transactions on semiconductor manufacturing, February 2003, Vol.16, 1
- [60] H. MISSBAUER, *Order release and sequence-dependent setup times*, International Journal of Production and Economics, 1997, Vol. 49, 131-143
- [61] P.K. JOHRI, *Dispatching in an integrated circuit wafer fabrication line*, Proceedings of the 1989 Winter Simulation Conference, 1989, 918-921
- [62] Z. DUWAYRI, M. MOLLAGHASEMI, D. NAZZAL en G. RABADI, *Scheduling setup changes at bottleneck stations in semiconductor manufacturing*, Production Planning & Control, October 2006, Vol. 17, nr. 7, 717-727
- [63] R.M. DABBAS, H.N. CHEN, J.W.FOWLER en D.SHUNK, *A combined dispatching criteria approach to scheduling semiconductor manufacturing systems*, Computers and Industrial Engineering, 2001, Vol.39, 307 – 324
- [64] D.Y. SHA, S.Y. HSU, Z.H. CHE en C.H. CHEN, *A dispatching rule for photolithography scheduling with online rework strategy*, Computers & Industrial Engineering, 2006, Vol. 50, 233-247
- [65] W.DAVIS en A.JONES, *A real-time production scheduler for a stochastic manufacturing environment*, International Journal of Computer Integrated Manufacturing, Vol. 1 (2), 101-112.
- [66] P.B. LUH en D.J. HOITOMT, *Scheduling of manufacturing systems using the Lagrangian relaxation technique*, IEEE Transactions on Semiconductor Manufacturing, 1993, Vol. 38(7), 1066-1079
- [67] C.S. SUNG en Y.L. CHOUNG, *Minimizing makespan on a single burn-in oven in semiconductor manufacturing*, European Journal of Operations Research, 2000, Vol. 120, 559-574
- [68] R. VAESSENS, E.AARTS en J. LENSTRA, *Job Shop Scheduling by Local Search*, Journal on Computing, 1996, Vol. 8, 302-317
- [69] S. G. PONNAMBALLAM, P. ARAVINDAN, en S.V. RAJESH, *A tabu search algorithm for job shop scheduling*, International Journal of Advanced Manufacturing Technology, 2000, Vol. 16, 765-771
- [70] C.D. GEIGER, K.G. KEMPF KG en R. UZSOY, *A Tabu search approach to scheduling an automated wet etch station*, Journal of Manufacturing Systems, 1997, Vol. 16(2), 102-116
- [71] P. VAN LAARHOVEN, E. AARTS en J. K. LENSTRA, *Job shop scheduling by simulated annealing*, Operations Research, 1992, Vol. 40, 113-125
- [72] S.J. YIM, D.Y. LEE, *Scheduling cluster tools in wafer fabrication using candidate list and simulated annealing*, Journal of Intelligent Manufacturing, 1999, Vol. 10(6), 531-540
- [73] F. D. CROCE, R. TADEI en G. VOLTA, *A genetic algorithm for the job shop problem*, Computers and Operations Research, 1995, Vol. 22, 15-24

- [74] P. RAY, *Handbook of Industrial and Systems engineering*, University of Alabama, Taylor and Francisgroup, 2006
- [75] M.P. GROOVER, *Work systems and the methods, measurement and management of work*, Pearson Education, 2007
- [76] D. VAN GOUBERGEN, *Work Measurement and Method Engineering*, cursus bij het gelijknamige vak, UGent