

KU LEUVEN



MA INGENIEURSWETENSCHAPPEN:  
COMPUTERWETENSCHAPPEN

---

# Digitale Elektronica en Processoren

---

LESNOTA'S

*Author:*  
Helena BREKALO

2015-2016



# Contents

<b>1</b>	<b>Les 1</b>	<b>5</b>
1.1	Slides: 1_Intro . . . . .	5
1.2	Slides: 2_Digitaal_Ontwerp . . . . .	6
<b>2</b>	<b>Les 2</b>	<b>11</b>
2.1	Slides: 2_Digitaal_Ontwerp . . . . .	11
2.2	Sides: 3_Technologie . . . . .	18
<b>3</b>	<b>Les 3</b>	<b>21</b>
3.1	Sides: 3_Technologie . . . . .	21
<b>4</b>	<b>Les 4</b>	<b>31</b>
4.1	Sides: 3_Technologie . . . . .	31
4.2	Slides: 4_Combinatorisch . . . . .	38
<b>5</b>	<b>Les 5</b>	<b>41</b>
5.1	Slides: 4_Combinatorisch . . . . .	41
<b>6</b>	<b>Les 6</b>	<b>49</b>
6.1	Slides: 4_Combinatorisch . . . . .	49



# Chapter 1

## Les 1

### 1.1 Slides: 1\_Intro

**Slide 6:** Die meerdere cores zorgen voor een grotere verwerkingskracht. Voorbeeld van de smartphone: het blokschema is hier van belang. Je hebt een aantal dingen die contact maken met de buitenwereld. In het blauw: de telecommunicatie, het echte gsm-gebeuren. Je hebt ook ergens iets draadloos (WiFi) (bruin). Het is dus heel specifieke hardware die een bepaalde taak vervult, die is geoptimaliseerd om één ding heel goedkoop, eenvoudig en compact te doen. Je hebt ook geheugen en powermanagement. De donkere blokjes in het midden zijn een algemene processor (rechts) die redelijk flexibel is en een specifieke processor (links) die alle signaalverwerking doet.

**Slide 7:** SoC: alles op 1 enkele chip. Heel wat verschillende dingen zijn weer aanwezig. 3 grote stukken: opgelijst met tweekleurige pijltjes:

- Programmeerbare processor: die kan vanalles en nog wat doen.
- Niet-programmeerbare processoren: specifieke processoren voor beeldverwerking bv. Ze zijn meestal beperkt programmeerbaar, om één specifieke familie van problemen op te lossen.
- Controle-eenheden.

Waarom die tweede groep? → Veel goedkoper! Je hebt ongeveer 1000 keer zoveel energie nodig op een programmeerbare processor dan bij een niet-programmeerbare. Bv i.p.v. 1W 1kW.

**Slide 8:** We vertrekken van een doel: we hebben een algoritme of beschrijving op gedragsniveau. Dat is een combinatie van software en hardware. Sommige dingen doe je via software omwille van flexibiliteit, andere dingen doen we via hardware. Dat is allemaal afhankelijk van elkaar. De keuze van welk programma je gaat gebruiken zal impact zal hebben op de hardware en de hardware die je hebt zal bepalen hoe programmeerbaar het is. Men spreekt daarom van hardware-software codesign. hardware gaat ook interageren met wat als algoritme gebruikt kan worden: sommige dingen zijn makkelijk te implementeren, andere moeilijker. Het kan zinvol zijn om de algoritmen eventueel wat aan te

passen.

Bij de hardware heb je een digitaal en een analoog gedeelte. Bij een smart-phone: alle telecommunicatie is analoog: continue signalen. Wat in het midden stond was meer digitaal: meer berekeningen. Beiden moeten geïmplementeerd worden, gewoonlijk op dezelfde componenten, momenteel gebeurt dat met transistoren en dat moet vertaald worden naar een chip (geïntegreerde schakeling). Wij gaan het hier enkel hebben over het linkerdeel. Dit kan nog opgesplitst worden in 3 lagen. Poorten liggen het dichtst aan bij transistoren, die kunnen gebruikt worden om basiscomponenten te maken (bv. basisprocessoren) die samengezet kunnen worden om systemen te bouwen.

**Slide 9:** Het gaat in eerste instantie over het ontwerp van elektronische schakelingen. We gaan daar ook voorbeelden van zien. Hardware kan je beschrijven via een schema, maar ook via een taal: hardware-beschrijvingstalen. We gaan die ook bekijken. We gaan dit ook toepassen in oefeningen en labo's.

**Slide 10:** Drie grote blokken:

1. Basis: principes van digitaal ontwerp en hoe je dat doet met elektronica. Veel van de dingen die we gaan zien kan ook met andere dingen dan elektronica.
2. Dan gaan we zien hoe we daar schakelingen mee kunnen maken. Eerst zonder geheugen en dan met geheugen.
3. Dan gaan ingewikkeldere algoritmen implementeren en dan kijken hoe dat geïmplementeerd kan worden.

**Slide 11:** Vereiste voorkennis.

**Examen:** FSM is schriftelijk, rest is ook mondeling. Eerst wordt de theorie ondervraagd, dan de vertaling van een algoritme (vraag 1).

## 1.2 Slides: 2\_Digitaal\_Ontwerp

**Slide 3:** Documentatie is heel belangrijk: als handleiding en om bij te houden wat geprobeerd is (en wat dus wel en niet werkt).

**Slide 4:** Specificatie: beschrijving, heel dikwijls in natuurlijke taal, van welke berekeningen je wilt/hebt. Specificaties zijn in de meeste gevallen te algemeen en te onvolledig om er een duidelijke implementatie van te maken omdat het in veel gevallen te weinig specificeert. Er kunnen heel veel verschillende algoritmes gebruikt worden om een probleem op te lossen. Langs de andere kant gaat het soms al zaken vastleggen waarvan je niet de bedoeling had dat het vastgelegd werd (bv. vaste of vlottende komma).

**Slide 5:** Het geheel is dikwijls iteratief (vandaar de lus). Synthese: schakeling maken. Het is niet het bouwen van de schakeling maar een iets algemenere beschrijving maken. Normaal doe je geen synthese van een algemene beschrijving om onmiddellijk op een geïntegreerde schakeling uit te komen. Je gaat daarom in niveau's werken: in stapjes oplossen: eerst op het hoogste niveau kijken en een heel algemene beschrijving doen: ik wil een FFT implementeren, wat houdt dat in, wat voor verwerking heb ik nodig?

Vervolgens ga je een niveau lager: welke blokken heb je allemaal nodig? Die ga je verder verfijnen tot op een niveau waarop je blokken krijgt waarvan je ongeveer weet hoe die geïmplementeerd moeten worden. Je gaat dat dan nog verder opsplitsen (in poorten en flip flops). Dan komt er nog een niveau onder: hoe maak ik een poort (met transistoren)?

Door dat in stukjes te doen en elk stukje per keer op te lossen is het doenbaar: je maakt ter hapklare brokken van die je kan verwerken. Wij gaan de twee middenste niveau's vooral bekijken.

**Slide 6:** Belangrijk is het gebruik van bibliotheken: de kennis van anderen, dingen die je zelf niet meer moet doen. We gaan componenten hergebruiken. Als iemand iets nog niet gedaan heeft ga je het zelf maken, het in de bibliotheek steken zodat het hergebruikt kan worden. Die bibliotheken komen op ieder van de niveau's terug. Waarom zo belangrijk? → Wet van Moore.

We krijgen dus heel wat mogelijkheden jaarlijks erbij. Uiteindelijk zal dat wel gaan vastlopen. Wat ontstaat er nu? We krijgen elk jaar meer en meer mogelijkheden. We kunnen dit jaar een twee keer zo complexe schakeling maken dan 2 jaar geleden. Een meer complexe schakeling ontwerpen duurt meer dan twee keer zo lang want je raakt volledig het overzicht kwijt. De processoren aan de rechterkant teken je niet zomaar op 1 blad, dat duurt eindeloos lang. De enige manier om dat doenbaar te krijgen is door zoveel mogelijk te hergebruiken. Dit betekent dat het ontwerpen altijd achterloopt op de mogelijkheden die we hebben.

**Slide 7:** Blauw: wat we als verwerkingskracht kunnen doen, dit is wat ons interesseert. Meer cores op 1 chip: kost qua ontwerpinspanning niet zoveel werk. In de toekomst gaan we dus veel meer in parallel moeten laten gebeuren. Vermogen (Typical Power): is belangrijk want je kan dat niet eindeloos laten toenemen. Het vermogengebruik hangt af van de spanning in het kwadraat ( $V$ ), de grootte van de chip ( $C$ ) en de frequentie ( $f$ ). Hoe hoger de frequentie, hoe meer verwerkingskracht je hebt. Oorspronkelijk zaten we aan minder dan een Watt voor een geïntegreerde schakeling zitten we nu aan 100 Watt.

We willen het vermogen dus constant houden maar dat heeft zijn impact op de frequentie en de verwerkingskracht.

**Slide 8:** Analyse: nagaan of het voldoet aan de specificaties. Dat doe je niet helemaal op het einde: het ontwerp gebeurt in stapjes en ga je telkens testen of dat het wel voldoet aan zijn specificaties op dat niveau. Dat betekent meer dan alleen maar testen of het doet wat het moet doen. Het functionele is eigenlijk maar een aspect ervan, er hoort veel meer bij: het vermogenverbruik bv. mag niet eindeloos oplopen. De snelheid waaraan dat kan werken, de

kostprijs,... Dat soort zaken moet ook nagekeken worden. De testbaarheid is ook een belangrijke factor.

**Slide 10:** Schakelingen gebaseerd op logische werking. We werken met discrete signalen en in de praktijk met bits: de waarden worden enkel 0 of 1.

**Slide 11:** Er is een verband tussen het al dan niet ingedrukt zijn van de knop en de uitgang. Je kan ook andere functies hebben: een complementaire die in rusttoestand een verbinding maakt: in rusttoestand is er een verbinding en brandt de lamp. Druk je de knop in, gaat de lamp uit. Er zijn heel wat verschillende manieren om een logische functie aan te duiden, afhankelijk van welk boek/welke beschrijving je neemt: onderaan: verschillende manieren om NOT te schrijven. Wij gebruiken normaal altijd een accent (').

**Slide 12:** Andere logische functies:

- AND: schakelaars in serie zetten en dan brandt de lamp alleen als beide knoppen ingedrukt zijn.
- OR: schakelaars in parallel zetten: de lamp brandt als een van de twee schakelaars ingedrukt is of ze alletwee ingedrukt zijn.

We kunnen complexere schakelaars maken: je kan een lamp bedienen vanop twee plaatsen, dat is ook een logische functie: XOR. De lamp zal branden als de ene of de andere knop ingedrukt is, maar niet als ze alletwee ingedrukt zijn of alletwee niet ingedrukt zijn. Voor andere logische functies kan je ze herleiden tot de twee eersten (AND en OR): XOR is een parallelschakeling van een serieschakeling: je hebt alleen maar een parallel- en een serieschakeling nodig en de inverter. We hebben dus maar 3 logische schakelingen nodig: het inverse (NOT), een AND en een OR, daarmee kunnen we alle logische functies bepalen.

**Slide 13:** Hoe beschrijven we de functionaliteit van zo'n logische schakeling? → Met een waarheidstabel: een opsomming van alle mogelijke combinaties aan de ingang en daarbij wordt aangegeven wat de uitgang is. Vermits een waarheidstabel eigenlijk de functionaliteit beschrijft zijn twee implementaties met eenzelfde waarheidstabel volledig equivalent, die doen exact hetzelfde. Je kan dus ook zeggen dat als je een waarheidstabel hebt die je op verschillende manieren kan maken, je aan de hand van die waarheidstabellen verschillende schakelingen kan maken. Het is dus niet zo dat bij een functionaliteit altijd maar 1 implementatie mogelijk is, er zijn meerdere mogelijkheden.

**Slide 14:** Logische poorten: meer dan de inverter, de AND en de OR-poort heb je in principe niet nodig, hoewel men dikwijls gebruik maakt van complexe poorten. In CMOS is het bv. goedkoper om met complexe poorten te werken dan met de basispoorten. Een logische schakeling is dan een combinatie van die poorten. Je kan dit met een schema doen of een programma.



**Slide 15:** Als je zo'n schema hebt kan je ook makkelijk de waarheidstabel opstellen. Je gaat gewoon op alle tussenliggende draden kijken wat er gebeurt voor elke combinatie van ingangen aan de uitgangen: dus aan a en b en daarna f en op die manier kan je de functionaliteit gaan beschrijven van wat die schakeling doet.

Dat is een ding: je hebt een schakeling en je weet wat die doet.

Meestal wil men het omgekeerde: men wil het omgekeerde, wat is er de schakeling voor? Maar zelfs dit is een onvolledige beschrijving van wat die schakeling doet want het beschrijft alleen de logische functie en in praktijk is een schakeling meer dan alleen zijn logische functie.

Een schakeling wordt altijd nog aangevuld met een tijdsgedrag. Het gaat er hier niet over hoe traag die poorten werken, het is niet dat als er bij x iets verandert, er onmiddellijk iets bij b verandert, dat gebeurt nooit ogenblikkelijk; ogenblikkelijk bestaat niet: het duurt altijd een zekere tijd. Dat betekent dat als we iets veranderen bij x bv., dat een tijdje later a gaat veranderen. y veranderen gaat een tijdje later b veranderen. Als a en b veranderd zijn, zal f een tijdje later veranderen. Wat je dan ziet verschijnen zijn vertragingen: het ogenblik dat de ingang verandert en dat de uitgang verandert, in elke reële schakeling is dat zo. Die vertraging gaat meespelen in hoe efficiënt de schakeling werkt. Het is dus belangrijk om die te kennen, je gaat dat nooit terugvinden in een logische functie, dat is bijkomende informatie.

Het tijdsgedrag nakijken is ook heel belangrijk om problemen te detecteren, problemen die je niet gaat zien als je alleen maar naar de functionaliteit gaat kijken. Bv. a en b veranderen "tegelijkertijd" (dat bestaat niet! De ene gebeurtenis zal altijd iets voor de andere plaatsvinden.) → dat kan zijn gevolgen hebben: als b iets vroeger verandert dan a, dan krijgen we heel kortstondig een waarde 0, dan zie je aan de uitgang plots een piekje verschijnen, wat er niet zou mogen zijn. Als je gewoon naar de logica kijkt, zou die er niet mogen zijn, maar in praktijk is dat er wel en dit kan voor problemen zorgen. Dat piekje dat er is, dat we niet verwachten, kan opeens gevolgen hebben in de verdere verwerking van de schakeling, dus we moeten ons daar bewust van zijn. We moeten dus niet alleen de logische functies controleren, maar het tijdsgedrag is even belangrijk.

**Slide 16:** Kijken we dan naar het omgekeerde: we hebben een functionaliteit, hoe kunnen we dat implementeren? Er staan twee implementaties, als je die gaat uitrekenen, dan zie je dat die identiek zijn, die doen juist hetzelfde.

Als we dit gedaan hebben kunnen we kiezen voor een van de implementaties en vanuit logisch standpunt maakt het geen enkel verschil. Het is dan niet zo dat je eender welke realisatie zomaar mag kiezen: de linkse heeft meer poorten dan de rechtse: het gaat meer kosten (meer transistoren, een grotere chip, ...). De linkse gaat ook trager werken want als je van x naar de uitgang moet ga je door twee ingangen telkens moeten gaan waar dat rechts niet het geval is.

Als we dan toch kunnen kiezen willen we het zo goedkoop mogelijk en zo snel mogelijk hebben. Hoe sneller die schakeling is, hoe hoger de verwerkingskracht: we kunnen maar aan de volgende bewerking beginnen als de vorige gedaan is. Als die schakeling dus heel snel een resultaat levert kan die heel snel aan een volgende bewerking beginnen. Vandaar dat hoe sneller de schakeling werkt, hoe hoger de verwerkingskracht. Als we dan toch kunnen kiezen willen we die schakeling met de minimale kostprijs en die schakeling met de hoogste verwerk-

ingskracht (die zo snel mogelijk het resultaat geeft). We gaan daarvoor gebruik maken van de formules die erbij staan.

De snelheid is evenredig met het aantal ingangen, dus de linkse zal trager werken dan de rechtse. Op die manier kan men tussen de twee schakelingen kiezen en de beste kiezen.

**Slide 18:** Boole algebra kan gebruikt worden omdat we ook hier met twee (logische) waarden werken. In de Boole algebra vertrekt men van een aantal axioma's. De duale uitdrukkingen zijn hier ook steeds geldig: je mag 0 door 1 vervangen en omgekeerd en  $+$  door  $.$  en omgekeerd.

**Slide 20:** Wij gaan vooral gebruik maken van de wet van De Morgan: de inverse van het product is de som van de inversen en het inverse van de som is het product van de inversen.

In dit vak zal een distributiviteit voor de vermenigvuldiging ten opzichte van de optelling en distributiviteit van de optelling ten opzichte van de vermenigvuldiging gebruikt worden.

# Chapter 2

## Les 2

### 2.1 Slides: 2\_Digitaal\_Ontwerp

**Slide 15:** Je wil weten wat de vertragingen zijn: het duurt een poosje als je iets aan de ingang verandert eer je het resultaat aan de uitgang ziet. Het tijdsgedrag toont ook problemen met de implementatie. Het probleem op de tekening toont dat a en b quasi gelijktijdig veranderen. Als a van 0 naar 1 gaat en b van 1 naar 0 zal de uitgang ook 1 worden dus f zou 1 moeten blijven. Het probleem ontstaat wanneer b iets vroeger verandert dan a: het uiteinde zal even 0 zijn, hoewel dat logisch gezien niet zou mogen. Op het moment dat je het ontwerp maakt is het belangrijk te weten wat zou kunnen optreden als probleem en je moet dat testen aan de hand van de reële implementatie.

**Slide 22:** Wij willen iets doen met de schakeling, hoe maak ik de schakeling zodat die daaraan voldoet? Er zijn 2 manieren om dat te benaderen: een schema tekenen/zoeken dat eraan voldoet. Anderszijds kan je ook een beschrijving in een taal doen.

Vandaag gaan we beiden ingeleid zien.

**Slide 23:** Hoe best het gedrag beschrijven? → Aan de hand van een waarheidstabel. Het is nu de bedoeling om aan de hand van de waarheidstabel een implementatie te maken. De logische functie kun je ook altijd omzetten naar een waarheidstabel. Verschillende logische functies kunnen eenzelfde waarheidstabel geven. De basisfunctionaliteit voor combinatorische schakelingen is dus de waarheidstabel.

In die waarheidstabel staat dat de uitgang 1 is als x en y 0 zijn,...

Je kan dit rechtstreeks implementeren met poorten: zie tekening rechtsonder.

Ofwel zijn beiden nul, ofwel is x 0 en y 1 ofwel zijn ze beiden 1 geeft  $f = 1$ . We kunnen van een beschrijving altijd rechtsreeks overgaan naar een implementatie. Probleem: we zijn niet geïnteresseerd in *een* implementatie maar in de beste, afhankelijk van de factoren die we belangrijk vinden.

Er zijn natuurlijk verschillende oplossingsmogelijkheden. Het komt er dus op neer om de beste te zoeken.

**Slide 24:** Je kan daar naartoe werken door de bovenste uit te schrijven en dat eventueel te vereenvoudigen en daar wiskunding op te beginnen rekenen met de Boole-algebra. Het resultaat op deze slide getoond is veel eenvoudiger dan het voorgaande. Het gaat dus goedkoper zijn en een kleinere vertraging hebben en een kleiner vermogen verbruiken.

Hoe weet je nu dat je die regels moet toepassen? Niet: de prof wist wat hij moest vinden en heeft dan regels gebruikt om ertoe te komen. Je kan dus eindeloos vanalles proberen en als je dat op de zuiver wiskundige manier wil doen moet je heel veel proberen en dan de oplossing behouden tot je het beste resultaat bekomt. Je zou ook nog een regel kunnen toepassen die op zich een minder goede oplossing geeft. Het is dus heel moeilijk om te zien wanneer je moet stoppen met regels toe te passen. In CMOS is het tweede goedkoper dan het eerste: de technologie komt erbovenop om te weten wat het beste is.

Je hebt dus geen stappenplan/algoritme waarmee je er gegarandeerd komt. Hoe komen we dan tot een beter resultaat? → Gaan we later zien, in het volgende hoofdstuk: hoe kunnen we met een minimum aan uitproberen tot een goed resultaat komen?

**Slide 25:** Je kan dat formuleren in functie van min- en maxtermen. Een minterm is een functie zoals beschreven op de slide die alleen waar is voor een van de rijen van de waarheidstabel. Je hebt hier 3 variabelen dus  $2^3$  mogelijkheden. Hier staat geen functie of uitgang bij. Als je de functie erbij betreft spreekt men over 1-minterm: de functie waarvoor de uitgang van de schakeling 1 is. In dit geval zijn dat er 4.

**Slide 26:** Je kan dat als volgt omschrijven: je kan altijd een implementatie van een schakeling doen door de som van zijn 1-mintermen te nemen. Dus voor elke rij hebben we een 1-minterm (?). Men noemt dit de canonieke som van producten oplossing → de voor de hand liggende oplossing, waar je niet over moet nadenken en je zo kan realiseren.

Het mooie aan Boole is dat de duale uitdrukking ook altijd geldig is. Dit kunnen we ook hier gaan toepassen.

**Slide 27:** Maxterm = logische functie die 0 is voor 1 bepaalde rij van de kolom. Weerom kun je de functie er dan bij betrekken. Je krijgt hier geen 1-maxterm maar een 0-maxterm: maxterm waarvoor de functie 0 is en dus de rijen die er bij de 1-minterm niet bijzaten.

**Slide 28:** Als je dit gaat combineren krijg je een product van sommen. Voor elke beschrijving die we hebben weet je dat er 2 implementaties zijn. Dat zijn 2 verschillende implementaties en we gaan sowieso moeten kiezen wat het beste is.

**Slide 29:** Het probleem met de canieke: de duurste oplossing die er is (maar wel makkelijk op te schrijven): elke AND-poort heeft zoveel ingangen als er variabelen zijn. Een OR-poort heeft zoveel ingangen als er rijen zijn waar een 1 in voorkwam op het einde. Dat is dus niet de compactste oplossing. We zijn op zoek naar eenzelfde manier van werken (SOP of POS) maar met poorten met minder ingangen die zo ook minder vermogen verbruiken. We gaan niet

meer voor elke AND-poort alle ingangen proberen gebruiken en niet evenveel ingangen voor de OR-poort als in de waarheidstabellen.

**Slide 30:** De twee canonieke vormen bovenaan. Als je daar een paar regels op toepast kan je beiden vereenvoudigen zoals onderaan getoond. Dit zijn de standaardvormen: die vorm die het minimum aantal poorten en de poorten met de minste ingangen gaat proberen te gebruiken. Als je de bovenste rijen daarmee vergelijkt bevat de bovenste elke term 3 ingangen vs. 2. Dat is al een winst in alle randvoorwaarden die we willen. Daarenboven hebben we vanboven 4 OR-poorten en vanonder maar 2. Ook dit is dus een duidelijke reductie in hardware.

We weten nu wel nog niet hoe we overgaan naar standaardvorm.

**Slide 31:** We gaan dikwijls naar de SOP en POS-implementatie kijken omdat het zo is dat die standaardvorm eigenlijk de goedkoopste implementatie en dikwijls ook de snelste is, zeker wanneer je je beperkt tot 2 lagen. Een laag: bovenste tekening: een deel/laag AND-poorten en een laag OR-poorten (1 poort in dit geval). Die bepaalt de vertraging: je moet door 2 lagen om tot het einde te geraken. Hoe meer lagen, hoe trager het geheel gaat werken. We gaan dus proberen het aantal lagen te beperken. Het minimum dat we in normale omstandigheden nodig hebben zijn 2 lagen. Vandaar dat de getoonde tekening de snelste oplossing is.

We kunnen ook naar meer lagen gaan: onderste tekening. Hier gaat het duidelijk trager zijn: 3 lagen (dit was een examenvraag vorig jaar!). Soms is snelheid niet het belangrijkste: soms wil je een zo goedkoop/compact mogelijke implementatie en dan kan dit wel interessant zijn. Het onderste kan goedkoper zijn dan het bovenste: het verschil zit 'm in het feit dat er bovenaan een poort is met 3 ingangen en je die vanonder niet hebt. Als je dus niet in de snelheid geïnteresseerd bent, kan de onderste dus beter zijn.

Wij gaan meestal met 2 lagen werken omdat dat evidentier is om toe te komen, 3 lagen is minder evident.

**Slide 32:** We gaan nu ook nog kijken naar NAND en NOR.

**Slide 33:** We maken gebruik van 1 enkele soort poort. Dat is mogelijk door inverterende poorten zoals NAND of NOR-poorten te volgen. Het wordt als 1 poort beschouwd (de combinatie van AND en inverse of OR en inverse). In de CMOS-technologie is het makkelijker om een NAND-poort te maken dan een AND-poort:

- Je moet een AND maken door een NAND te gebruiken en er nog een inverter achter te zetten.
- Wanneer je NAND of NOR gebruikt moet je geen 3 basisschakelingen hebben, je hebt er maar 2 nodig. Schema onderaan: illustratie hoe je poorten kan maken met alleen NAND of alleen NOR. Denk tijdens het studeren na welk van de mogelijkheden het interessantst is om te gebruiken.

Ook een NOR-poort kan je maken met een NAND-poort: de wetten van De Morgan: product van inversen is hetzelfde als inverse van een som. Dat zie je in

de omkaderde vakjes. Met een enkele basispoort kan je dus alles maken. Het is nu wel zo dat de technieken die we later gaan zien altijd van deze basispoorten gebruik maken. Die gaan iets leveren dat bestaat uit AND en OR en invertoren. Er zijn geen specifieke technieken om alleen NAND of NOR te gebruiken. Op zich is dat geen probleem omdat een implementatie zoals op **Slide 34** kan omzetten naar alleen NAND of alleen NOR.

**Slide 34:** Maak gebruik van invertoren aan de uitgang en ingang. De drie schakelingen onder elkaar zijn dezelfde logische functies. Hetzelfde met iets alleen NOR-poorten.

Houd er rekening mee dat als je tot een implementatie met NAND-poorten wil komen je eerst een som van producten moet implementeren.

**Slide 35:** Als je gaat kijken naar de ontwerpmiddelen die bestaan, dan gaan die een ontwerp maken en die gaan je erbij helpen. Het grootste werk dat je zal hebben is het ingeven van het ontwerp hetzij via een schema, hetzij via VHDL. Daarna moet je doen wat we in de volgende hoofdstukken gaan bekijken. Die programma's zijn slim genoeg om alles te doen wat wij gaan zien, dus ze gaan in veel gevallen een automatische synthese doen. Op dat moment heb je een logische implementatie en kan je een functionele simulatie doen: testen of het doet wat het moet doen: als je een bepaalde combinatie van input geeft gaat het geven aan de uitgang wat het moet geven.

Het tijdsgedrag is even belangrijk, maar dat kan je op dit moment nog niet testen: je hebt alleen een logische beschrijving. Daarom komt er een stap achter: je moet ook het fysische ontwerp doen en nagaan wat het tijdsgedrag en de elektrische eigenschappen etc. zijn. Je moet een prototype bouwen en daarop metingen uitvoeren. Wanneer dat tot een geïntegreerde schakeling moet leiden is dat een verspilling van tijd en kosten dus normaal gaat de fabrikant van die geïntegreerde schakeling weten wat de eigenschappen gaan zijn van die componenten. → Al die details zijn gekend. Die zal een module aanbieden waarin alle gegevens verwerkt zijn en het fysische ontwerp nabootst. Op die manier kan je metingen doen qua tijd etc. Je kan ook naar alle problemen zoals glitches gaan zoeken. Pas dan ga je de finale stappen uitvoeren: implementeren of de chip laten maken.

**Slide 37:** Andere manier om hardware te omschrijven: hardware-omschrijvingstaal. Wij gaan kijken naar VHDL. Dat is ontstaan uit subsidies van het Amerikaanse ministerie van defensie. De bedoeling was dat men een methode wou hebben om heel complexe ontwerpen te beschrijven en alle gegevens erbij te zetten, alle info om te simuleren en beschrijven en dat systeem op een voldoende hoog niveau te beschrijven.

De taal moet in staat zijn om gelijk welk digitaal systeem te beschrijven en alle aspecten ervan. Het moet het dus volledig kunnen specificeren. Dus niet alleen op poortniveau maar ook op hogere niveau's. Je begint dus op een hoog niveau en je daalt af. De bedoeling is hier dat je het systeem op alle niveau's kan beschrijven:

- Op het hoogst niveau heb je een beschrijving van het gedrag: "Dit blokje moet dat doen" .

- Je moet ook kunnen testen dus tijdsparements erin beschrijven. Simulatie moet hier dus ook mogelijk zijn.
- Men zou ook graag hebben dat er een automatische synthese is: dat je je geen zorgen moet maken om de tussenstappen.
- Ook documentatie is belangrijk.

Het is ook belangrijk dat het gestandaardiseerd is. Dat zorgt ervoor dat verschillende fabrikanten die ieder hun eigen implementaties doen mekaar altijd verstaan: er kan geen verkeerde implementatie mogelijk zijn want in de taal is de implementatie exact beschreven. Nadeel: als er nieuwe ideeën komen/er zijn tekorten, dan kan men het niet meer veranderen → nieuwe versie van de taal/software nodig.

**Slide 38:** VHDL is voor digitale systemen, maar een systeem bestaat meestal uit digitale en analoge componenten. Men heeft het systeem dus uitgebreid naar het combineren met analoge signalen. Om een analoge schakeling op te lossen heb je 200-300 gekoppelde differentiaalvergelijkingen nodig om tot uw oplossing te komen.

**Slide 39:** Andere kandidaat die minstens evenwaardig is: Verilog. Wordt meer gebruikt in de VS. VHDL is populairder in Europa. De concepten die erachter zitten zijn speciaal: hardware beschrijven is niet hetzelfde als een programma schrijven. Er zitten dus speciale concepten achter die je moet begrijpen voor beide talen, alleen de syntax verschilt.

**Slide 40:** Wat is nu het interessantste: schema of hardwaretaal? → Hangt af van wat je ermee moet doen. Schema's zijn interessanter voor kleine ontwerpen, VHDL is gewoonlijk interessanter wanneer het over grotere ontwerpen gaat. Nadelen van VHDL t.o.v. schema's:

- De concepten achter VHDL zijn niet voor de hand liggend en je hebt dus enige tijd nodig om dat te leren beheersen.
- Het is nogal langdradig: om de kleinste component te beschrijven heb je onmiddellijk veel lijnen code nodig. Dit maakt dat het voor mensen moeilijk is om te begrijpen wat er allemaal gebeurt. Een schema is visueler en daarin zijn mensen getrainder.
- VHDL bevat alle mogelijkheden: alle mogelijkheden voor simulaties, maar om te begrijpen wat het doet, om alleen de functionaliteit te kennen heb je dat allemaal niet nodig.

**Slide 41:** Voordelen van VHDL t.o.v. schema's:

- Het is een standaard: je kan het gebruiken om informatie uit te wisselen tussen programma's/tools van verschillende fabrikanten. Het is dus makkelijk overdraagbaar.

- VHDL is goed in alles waar programma's goed in zijn. Wanneer je repetitieve structuren hebben (een aantal cores die allemaal hetzelfde zijn), dat is 1 instructie in VHDL. Als je een schema tekent, moet je dat zoveel keer tekenen. In een programma kan je ook parameters gebruiken. Als je een schema tekent moet je het helemaal uitgommen en opnieuw beginnen indien een parameter verandert. In een programma, als uw parameter bv. niet voldoende groot is, pas je dat gewoon aan.
- In een schema moet je onmiddellijk componenten kiezen om erop te zetten. VHDL heeft het voordeel dat je het gedrag kunt beschrijven, je moet dus niet per sé een implementatie kiezen.

**Slide 42:** Beperkingen:

- VHDL heeft zoveel mogelijkheden dat het enorm veel inspanning vraagt van de fabrikanten om alle aspecten uit de taal om te zetten naar hardware. Men heeft zich dus beperkt tot een subset die makkelijker om te zetten is in hardware. Wanneer je gaat synthetiseren ga je je dus moeten beperken.
- Van een hardwarecompiler verwachten we een enorme intelligentie. In vele gevallen zijn er een tiental mogelijkheden om een implementatie te realiseren. De compiler moet dus raden wat je bedoelt wat je neerschrijft. In de meeste gevallen raadt hij dat zo strict mogelijk. Je moet in de meeste gevallen de compiler dus wat helpen. Je moet ongeveer weten hoe hij interpreteert wat jij hebt neergeschreven. Je moet je dus aanpassen naar de mogelijkheden van de compiler. De hardware die uit een ontwerp van een beginnening komt zal minder optimaal zijn dan die an een meer ervaren ontwerper.

**Slide 43:** Schema tekenen: kan in een keer, maar in de praktijk ga je op de hiërarchische manier werken. Je gebruikt blokjes die je eventueel nog niet definiëert en waarvan je later de ontbrekende blokjes gaat specificeren. Je kan het als een doos bekijken met een aantal in- en uitgangen. Die hebben een aantal karakteristieken en wat er in die doos inzit. Wat zit er in die doos? De bewerkingen die we moeten doen. Stel dat we een component hebben die test of 2 getallen aan elkaar gelijk zijn. In die doos zitten dus componenten en die leveren resultaten af. We hebben ook verbindingen nodig.

We kunnen dit nog niet implementeren omdat we nog niet hebben aangegeven wat de component is. We hebben wel al aangegeven dat het twee keer hetzelfde is. Die component moeten we nu maar 1 keer ontwerpen en die kunnen we 2 keer gebruiken.

Nu de rechtse doos: op een lager niveau specificeren. We kiezen een mogelijke implementatie. Die twee stukken kunnen door 2 verschillende personen gemaakt worden. Men kan eerst het rechtse maken en dan het linkse. In veel gevallen is het rechtse beschikbaar in een bibliotheek.

Wanneer je met VHDL werkt, gaat dat eigenlijk juist hetzelfde zijn.

**Slide 44:** 2 stukken: entiteit die de doos beschrijft. We hebben een comparator comp die twee ingangen heeft: 2 ingangen en 1 uitgang en de ingangen zijn 8 bit en de uitgang 1 bit.



Daaronder staat de architectuur: wat er in de doos zit, wat we daarnet in de doos tekenden.

Het verschil met het schema is dat we nu ook het gedrag kunnen beschrijven zoals gedaan is op de slide: we kunnen schrijven wat de comparator doet. Er staat niet geschreven hoe dit geïmplementeerd moet worden, er is gewoon beschreven wat het doet.

Het is ook zo dat voor een bepaalde component er verschillende implementaties kunnen zijn: gedrag beschrijven, implementatie beschrijven, voor sommige theorieën zal iets goed zijn, in anderen niet. De architecturen krijgen daarom een naam omdat je zo kan kiezen voor een bepaalde implementatie.

**Slide 45:** Kijken we naar het topschema beschreven via de structuur. Een structurele beschrijving is juist hetzelfde als het schema tekenen. Entity Test is weer de doos en de architecture is weer de beschrijving hoe het geïmplementeerd is. We hebben een component Comparator nodig, dat is een virtuele component en we moeten die nog niet meteen beschrijven, gewoon de in- en uitgangen. Bij begin staat er dat je die twee keer gebruikt. Je kan die verschillende namen geven (Comp1 en Comp2) om een onderscheid te kunnen maken. Rechts worden de draden beschreven.

Dit is in woorden neergeschreven wat rechts vanboven met een schema staat. Belangrijk is om in gedachten te houden dat het een beschrijving van hardware is. De volgorde van instructies is hier niet belangrijk omdat het over hardware gaat. Het in een andere volgorde zetten gaat hetzelfde eindresultaat geven: de uiteindelijke manier van werken blijft hetzelfde. Ook de volgorde van de uitdrukkingen doet er dus niet toe: parallele uitdrukkingen. De volgorde is totaal onbelangrijk.

**Slide 46:** Hoe leg je het verband tussen de twee schema's/beschrijvingen? Dat gebeurt heel expliciet: VHDL is een heel stricte taal dus normaal moet je alles heel uitgebreid beschrijven, maar als je uw namen goed schrijft kan hij raden wat je bedoelt. Als de specificaties overeenkomen (naam Comparator met x ingangen en y uitgangen) dan zal die dat kunnen raden. Je kan ook een specifieke mapping maken: we gebruiken de component Comp met architectuur Behav1: we kiezen meteen een specifieke implementatie. Mapping: in het ene geval hebben we X,Y,Z gekozen, in het andere A,B,EQ. je kan dus heel specifiek aangeven hoe je de schema's met elkaar verbindt.

**Slide 47:** VHDL verschilt vrij hard van een gewone programmeertaal, het ziet er alleen uit als een programmeertaal. Vanboven: "gewone" programmeertaal: functie die je definiëert. In VHDL beschrijf je gedrag. Het topniveau is het hoofdprogramma: main() die de functie 2 keer gaat gebruiken. Belangrijk verschil: bij het vorige waren de uitdrukkingen parallel, hier sequentiëel: het ene wordt uitgevoerd na het andere, er is een stricte volgorde en dat kan zijn impact hebben. Bij hardware is dat niet zo: het is niet zo dat de bovenste component eerst werkt en dan de onderste.

Nog een verschil: een softwareprogramma start je op, die voert het programma uit en aan het einde van het programma stopt het. Bij hardware is het juist omgekeerd: hardware stopt niet. Hardware begint te werken zodra je de spanning opzet en die blijft eendeloos altijd maar hetzelfde doen tot je de spanning

afzet. Dat is de enige manier om die tot stilstand te brengen. Software: er is een bug aanwezig als het blijft runnen, hardware: er is een bug aanwezig als het stopt.

**Slide 48:**

- Datatypes: als je een computer gebruikt met bv. 32 bits of 64 en je wilt 2 8-bits optellen, zal dat toch een 32 of 64-bitoptelling zijn. Bij hardware is dat niet zo. Bij hardware ga je niet nodeloos bits gebruiken: als je het met 4 bits kan doen ga je dat doen. Je gaat altijd het minimaal aantal bits gebruiken omdat dat het goedkoopste is en de compactste oplossing is. In VHDL moet je dat daarom ook specificeren: gaat erin als 5 bit en komt eruit als 7 bits.
- Gelijktijdigheid: hardware werkt altijd gelijktijdig, niet de ene na de andere.
- Tijdsconcept: componenten werken continu, hardware stopt nooit. Qua tijdsconcept er ook rekening mee houden: de simulatie is niet de reële tijd. Het kan zijn dat je een minuut moet rekenen om de simulatie 1 nanoseconde vooruit te laten gaan.

## 2.2 Sides: 3\_Technologie

**Slide 3:** Als we elektronische schakelingen willen implementeren, hoe gebeurt dat in de praktijk? Het eerste wat we moeten beslissen is welke fysische waarden we hebben. Als je een logische schakeling maakt, moet je een fysische waarde hebben. We gaan dat doen met fysische bereiken: als de blauwe pijl de spanning voorstelt; L is een lage waarde, H is een hoge waarde. Die twee bereiken raken elkaar niet: daartussen is een stuk waar het noch hoog, noch laag is: ongedefinieerd.

Waarom bereiken? We willen onze schakeling zo robuust mogelijk. We willen niet dat we de schakeling zo nauwkeurig moeten maken zodanig dat er schakelingen nodig zijn waar exact 3V uitkomt en niet 3.1V bv. Als je toelaat dat het 3V, 3.5V of 2.5V kan zijn, laat je veel meer variatie toe. Verschillende componenten hebben allemaal verschillende karakteristieken, er zijn fluctuaties. Het feit van een bereik te gebruiken laat fluctuaties toe. Op die manier kunnen we dat oplossen en kunnen we veel makkelijker ingewikkelde zaken maken, dat hebben we ook als we verschillende transistoren op 1 component willen zetten. Ook het gebruik beïnvloedt dit: elektronica is heel gevoelig aan temperatuur: enorm groot verschil of de schakeling werkt aan 20 graden Celcius of 40 graden Celcius. Hou er rekening mee dat de schakeling zelf ook opwarmt. Je wil dat uw schakeling blijft werken.

Ook de spanning is niet altijd juist dezelfde.

Waarom 2 bits: om het systeem minder gevoelig te maken. Hoe meer mogelijke waarden, hoe kleiner de bereiken zijn.

We hebben het nog niet over 0 of 1 gehad. Het hoeft niet per sé zo te zijn dat 0 laag is en 1 hoog  $\rightarrow$  positieve logica, maar je kan ook negatieve logica hebben waarbij men het omgekeerd doet (0 is hoog en 1 is laag).

**Slide 4:** Hoe die poorten maken? Los van de elektronica: we kunnen dat maken als we beschikken over schakelaars die we kunnen aansturen. We hebben een schakelaar die 2 standen heeft: open of gesloten. We moeten die kunnen aansturen. → 2 soorten schakelaars.

**Slide 5:** NMOS-transistor: poort geïsoleerd van het blokje p. Als de spanning klein is zal er niks gebeuren en zal de halfgeleider zich gedragen als een isolator. Wordt de spanning voldoende groot, zal die elektronen aantrekken om een geleidend pad te creëren en wordt er een verbinding gemaakt. Op die manier kunnen we die gebruiken als schakelaar.

PMOS: hetzelfde prentje waar de p en de n verwisseld zijn en de source en de drain verplaatst zijn. Dat zal ervoor zorgen dat de gesloten schakelaar lage spanning betekent: de negatieve spanning is heel groot.

Op die manier kunnen we die 2 soorten schakelaars maken. We krijgen dit op dezelfde geïntegreerde schakeling.

**Slide 7:** Er poorten mee maken.

1. Enkel gebruik maken van NMOS. Dan krijg je een schakeling zoals op de afbeelding. Stel dat er aan x een lage spanning is, dan zal de transistor (rood) niet geleiden. De uitgang zal via de weerstand verbonden zijn met de voedingsspanning. Als er geen (grote) stromen staan, dan zal de spanning ongeveer gelijk zijn aan de bronspanning. Als er een hoge spanning op zit zal de transistor een lage weerstand hebben (veel lager dan eerst) met als gevolg dat die de uitgangsspanning naar beneden trekt: de weerstand is niet krachtig genoeg om het tegen te houden. Daarom komt er een laag niveau op te staan. Als je dit met positieve logica doet, dan hebben we een inverter gebouwd. Dit is tegenwoordig niet de manier die meestal gebruikt wordt omdat er belangrijke nadelen zijn:

- (a) De weerstand zal nooit helemaal nul zijn. Je krijgt een spanningsdeling over de twee weerstanden. Dat is geen probleem als het maar laag genoeg is, maar het gaat niet echt nul zijn. Dat betekent dat de bovenste weerstand beduidend groter moet zijn want anders krijg je het omgekeerde.
- (b) Dit soort schakeling heeft een statisch vermogenverbruik: we verbruiken ook vermogen op het moment dat er geen veranderingen optreden: alles is constant en toch blijven we vermogen verbruiken. Als er een geleiding is, loopt er een stroom van boven naar beneden. Een stroom door een weerstand betekent vermogenverlies. We kunnen dat alleen maar klein houden als die R zeer hoog is. In het realistische voorbeeld zal het toch 1mW verbruiken, maar een schakeling bestaat niet uit maar 1 inverter/transistor, er zal dus een enorm hoog vermogenverbruik zijn.

Het wordt niet meer gebruikt behalve bij open-drain.

**Slide 8:** De weerstand wordt buiten de verpakking weergegeven (stippellijn). Als de schakeling nog niet gemaakt is, hangt die draad los dus vandaar een open drain. Wat is er hier specifiek aan? Als je 1 inverter hebt: niks. Als

je er meerdere hebt en je hangt de uitgangen aan elkaar, dan krijg je bv. de gegeven waarheidstabel. We hebben zo een NOR-poort gemaakt. We noemen dit een wired-AND functionaliteit: we hadden oorspronkelijk 2 op zichzelf staande invertoren. We hebben die uitgangen aan elkaar gehangen en zo een NOR gemaakt. Die NOR kan ook op rechtsonderstaande manier omschreven worden. In het logische schema zit ook een AND-poort. Dit komt omdat de twee uitgangen aan elkaar hangen. Door een bedrading aan te leggen is er een extra functionaliteit. Het nadeel is dat je met NMOS blijft werken en het statisch vermogenverbruik hebt.

**Slide 9:** Lichtschakelaars in serie of parallel zetten. Als we hier 2 transistoren in serie zetten krijgen we wat op de slide staat. We hebben hier een NAND-poort gemaakt. We kunnen dat ook in parallel zetten, we krijgen dan een NOR: zodra 1 van de twee 1 wordt, wordt de stroom weggeleid.

**Slide 10:** Stel dat we het statisch vermogenverbruik willen vermijden. Dit kan door te zorgen dat er nooit een geleidend pad is. Dit kan door 2 transistoren te voorzien: 1 boven- en 1 onderaan. Vanboven een PMOS en het ander een NMOS. Het een is een pull up, het ander een pull-down. Hier hebben we geen statisch vermogengebruik want in geen van de twee gevallen is er een geleidend pad van de bron naar onder want in elk geval zal een van de twee open zijn.

**Slide 11:** Je mag dit nooit met CMOS doen: die werkt alleen als PUN en PDN complementair zijn: als ze alletwee beginnen geleiden, dan loopt er een stroom die bepaald wordt door de weerstand van een transistor die geleid, wat een zeer lage weerstand is, er zal dus een enorm grote stroom doorlopen. Gevolg: uw schakeling ontploft (wordt getoond in de slides normaal gezien). Je moet dus altijd zorgen dat de twee nooit gelijktijdig geleiden.

Als je de uitgangen aan elkaar hangt kan je dit ook krijgen: ook een grote stroom en minstens een van de twee zal kapot gaan.

CMOS-uitgangen mag je dus nooit zomaar aan elkaar hangen!

**Slide 12:** Als we ze toch aan elkaar moeten hangen, gaan we een speciale oplossing moeten gebruiken. We hebben een inverter gemaakt en een serie- en parallelschakeling. Op deze manier kunen ze nooit gelijktijdig geleiden. Het bovenste PUN is altijd complementair aan het PDN. Ook hier is een NAND-poort gemaakt. Je kan dit ook omkeren: boven serie en vanonder de parallel. Dat leidt tot een NOR-poort. Ook hier zijn de basiscomponenten een NAND en NOR poort.

Interessant hier is dat we daarnet 2 ingangen hadden en 2 transistoren. Hier 2 ingangen en 4 transistoren: per ingang 2 transistoren erbij, dit blijft gelden: per ingang die erbij komt hebben we 2 transistoren nodig. Daaruit volgt dat de kostprijs van een poort bij CMOS evenredig is met het aantal ingangen.

**Slide 13:** Negatieve logica en actief lage signalen.

# Chapter 3

## Les 3

### 3.1 Sides: 3\_Technologie

**Slide 13:** Negatieve logica en actief lage signalen hebben strikt gezien niets met elkaar te maken. Het lage bereik komt overeen met een 1 (negatieve logica).

**Slide 14:** Impact op de implementatie: die implementatie geeft een NAND-poort. De onderste waarheidstabel hoort daar dus bij en niet de rechtsbovenste. De vertaling van een L naar een nul is positieve logica. Werk je met negatieve logica is het natuurlijk omgekeerd. Je krijgt dan een NOR-poort. In de praktijk werken we omgekeerd: we maken een schema en afhankelijk van de technologie kiezen we voor positieve of negatieve logica. Kiezen we voor positieve logica, dan zal het de afgebeelde afbeelding zijn, anders zal het een andere implementatie zijn.

In sommige schema's zie je een klein driehoekje staan: geeft aan dat die draad met negatieve logica werkt. In de praktijk ga je dat vrijwel nooit terugvinden: ofwel werk je met positieve ofwel met negatieve logica. Het is pas wanneer je gaat mengen dat je die driehoekjes nodig hebt.

**Slide 15:** Actief lage signalen: heeft niks met het lage en het hoge te maken, het actief zijn heeft enkel te maken met de interpretatie, niet met het binair signaal dat erop staat.

Voorbeeld: wanneer je een reset-signaal hebt (om een systeem in een initiële toestand te brengen: als je de spanning opzet, komt uw systeem in een random toestand. In praktijk wordt de spanning opgezet en onmiddellijk een reset aangebracht om het toestel in een gekende toestand te brengen). Het resetten van een toestand is een activiteit, je moet dat doen. Iets doen kan met een signaal dat 0 of 1 is. Het is niet de logische spanning die op de lijnen staat, maar wanneer iets wordt gedaan: wanneer het een 0 is wat erop staat of wanneer het een 1 is wat erop staat. Dat heeft dezelfde notatie als het inverse, maar daar heeft het dus niets mee te maken. Wij gaan inverse aangeven met ' en actief signaal met \*.

Als het signaal 0 is, zal de schakeling op de slide gereset worden. Als je poorten/logische functies gaat bekijken treedt hier voor het eerst het verschil op tussen de poort en de functionaliteit. De functionaliteit beschrijft wat er moet gebeuren in welke omstandigheden (bv. uitgang zal actief zijn bij AND

als beide signalen actief zijn). Bekijken we nu de waarheidstabel op de slide verwachten we normaal een AND-functionaliteit, dat is er een voor actief hoge signalen. Als “er gebeurt iets” overeenkomt met een 0, dan is de waarheidstabel een OR-functionaliteit. Dat is en blijft een AND-poort en de positieve of negatieve logica staat hier totaal los van. De interpretatie van “er gebeurt iets” verandert afhankelijk van of het actief hoge of actief lage signalen zijn. Meestal wordt er gekozen om met actief hoge signalen te werken, dat vereenvoudigt de interpretatie.

Toch zijn er veel plaatsen met actief lage signalen. Er zijn typisch 2 plaatsen waar je het frequent tegenkomt: reset en bij afgesloten verbindingen. Wired OR = wired AND (want je kan geen wired OR maken op zich) met actief lage signalen waardoor die AND poort voor een OR-functionaliteit gebruikt kan worden.

**Slide 16:** Bedoeling: OR-functionaliteit nodig en die met wired AND organiseren, of ook: met een schakeling vanonder getoond: met RC-schakeling. De schakeling wordt opgezet en de spanning gaat blijven stijgen. In dat eerste deel kan je niks zeggen over de schakeling, daar gebeurt de rommel, je weet niet wat er gebeurt. Een keer die in werkende toestand is, kan je die resetten. De capaciteit gaat proberen de spanning laag te houden en via de weerstand gaat het opgebouwd worden. Die ingang van de schakeling RST (groene lijn op de grafiek) gaat langzaam opbouwen. De spanning bij RST\* gaat dus vrij lang laag gehouden worden. Om dat soort zaken te doen moet het een actief laag signaal kunnen zijn.

Je wil een RST-knop hebben waar je op drukt en dan moet dat in orde zijn. Het is makkelijker om het met een wired-AND te doen dan met poorten. Het grote voordeel daarvan is dat je dat eindeloos kan uitbreiden: gewoon een draad erbij leggen, geen poort vervangen ofzo.

In de meeste schakelingen is de RST een actief laag signaal.

**Slide 17:** Tweede reden waarom er actief lage signalen gebruikt worden: wanneer dingen afgesloten zijn, zit je meestal in een hogere toestand. Op de slide getoond: langere lijnen, maar het hangt af van de frequenties waar je mee werkt. Verbinding: spanning aan de ene kant en verwachten die onmiddellijk aan de andere kant te zien. Bij een transmissielijn plant die overgang zich voort over de lijn en dat duurt een zekere tijd eer die de andere kant bereikt. Dan gaat niet alleen de spanning verhogen aan de andere kant maar eventueel slechts een deel van de energie opgenomen worden door degene die de spanning al dan niet meet en een deel gaat teruggekaatst worden, het zal daar ook gedeeltelijk geabsorbeerd worden en gedeeltelijk teruggekaatst worden.

Wat als je dat op een oscilloscoop bekijkt? Het rechtse signaal. Hoe meer reflecties er heen en weer gaan, hoe groter de overshoots zijn. Wat we eigenlijk willen is het linkse: je gaat naar een ander niveau en blijft daar. Dat kan je bereiken bij een transmissielijn door ervoor te zorgen dat degene die het signaal moet opnemen dezelfde impedantie heeft als de draden. Het komt erop neer dat je die draden moet afsluiten op de karakteristieke impedantie. Die kan soms redelijk laag zijn.

Hoe sluit je dat af? Zoals linksonder getoond: aan de ontvangtzijde 2 weerstanden zetten die ervoor zorgen dat de karakteristieke impedantie gelijk is aan

de parallelschakeling van die 2 weerstanden. Welke spanning er op de draad hangt wordt bepaald door de twee weerstanden. Als je uw spanning aanzet zal er 5V opstaan. Als die draad niet aangestuurd wordt, als daar niets mee gebeurt, dan staat daar een 1 op. Dat betekent dat we met actief lage signalen moeten werken: we doen niks en er staat een 1 op.

Vandaar, bij plusverbindingen (die tekening) ga je het regelmatig tegenkomen dat er ook daar met actief lage signalen gewerkt wordt.

**Slide 19:** Meer complexe schakelingen maken: bij CMOS moet je ervoor zorgen dat als de bovenste geleidt, de onderste niet mag geleiden en omgekeerd. Je kan ook heel wat andere combinaties van transistoren maken.

Het makkelijkste is te kijken wanneer de onderste tak gaat geleiden: wanneer a & b 1 zijn of wanneer x, y, z 1 zijn, dan zal het signaal nul zijn.

Dit is 1 enkele poort die iets meer doet dan alleen maar AND of OR implementeren: AOI.

Vanonder altijd NMOS, vanboven altijd PMOS.

Het aantal transistoren dat je nodig hebt is twee keer zo groot als het aantal ingangen. De kostprijs van de schakeling is evenredig met het aantal ingangen. Let wel: ook hier gaat het over inverterende poorten. De uitgangen van de twee ANDs of de twee ORs tellen niet mee als ingang!

**Slide 20:** Er zijn een aantal dingen die je niet kan maken met 1 enkele poort: je moet een NAND-poort maken gevolgd door een inverter  $\rightarrow$  impact op kostprijs en vertraging.

Buffer: doet op zichzelf niks, maar je kan meer aansturen.

Als je naar een XOR-poort gaat kijken: die gaat 1 zijn als ingang 1 of 2 1 is maar niet als ze alletwee 1 zijn. XNOR zal 1 zijn als beide ingangen 1 of 0 zijn. Hoe maak je dat? Met AND en OR-poorten, maar het kan efficiënter zoals in het blauw: dat is een OR-AND-inverter. Voordeel: bij gewone poorten heb je er 3 nodig met ieder een kostprijs, bij OAI heb je 1 poort nodig en 2 invertoren.

**Slide 21:** Mogelijkheid: uitgang is zwevend/open/nergens aan verbonden. Zwevende draden zorgen voor problemen dus wil je niet. Omdat ze door niets aangestuurd worden, is er niets dat ervoor zorgt dat die 0 of 1 blijven. Normaal staat daar geen spanning op, maar die reageren als antenne: je steekt de verlichting bv. aan en door inductie verandert de spanning op die draad. Loshangende draden wil je dus niet.

Waarom wil je dat toch kunnen doen (die aan niks laten hangen)? We hebben het al over wired-AND gehad, bij CMOS mag je de uitgangen nooit aan elkaar hangen, maar wat als je de ingangen aan elkaar wilt hangen? De enige oplossing om uitgangen bij CMOS aan elkaar te hangen is door gebruik te maken van een tristate buffer: geeft ingang door aan uitgang zonder functionaliteit, maar die kan de uitgang loskoppelen.

Hoe kan je dat implementeren? Met de getoonde schakeling: geen van beide zal geleidend zijn en dus losgekoppeld van de rest van de schakeling.

**Slide 22:** Bij een normale aansturing heb je 1 element dat aanstuurt en je op de meeste plaatsen gebruikt (denk ik). Als je een bus-verbinding hebt (kan

vanop verschillende plaatsen aangestuurd worden, wel van 1 plaats tegelijkertijd), moet je die vanaf 2 kanten kunnen aansturen, nooit tegelijkertijd (want dan heb je uitgangen die aan elkaar hangen!). Als je nu een intelligente sturing hebt die ervoor zorgt dat maximaal 1 of 2 verbonden kan zijn met de bus, dan kan dat wel op die manier. Je kan uitgangen van CMOS aan elkaar hangen als je gebruik maakt van een tristate buffer en die op een intelligente manier aanstuurt.

**Slide 23:** Praktische aspecten: wat voor soort transistoren we gebruiken, . . .  $\rightarrow$  heeft impact op de karakteristieken van de schakeling (vooral op het tijdsgedrag, niet zozeer op de logische functie).

**Slide 24:** Gevoeligheid voor storingen: men zorgt er altijd voor dat er een ruismarge is: we hebben een hoog bereik en een laag bereik. Normaal verwacht je dat als een poort iets aanstuurt, de uitgang van de poort die aanstuurt moet hetzelfde bereik gebruiken als de poort die zijn uitgang moet bekijken. Die kan dus geen groter bereik hebben bij het produceren van de spanning dan wat gespecificeerd is. Je kan wel een kleiner bereik hebben. Eigenlijk heb je dus een stuk niet nodig: het groene stuk. Het is toch nuttig om het verschil te maken want dat is de ruismarge. Waarvoor gaan we de gebruiken? Stel dat je een spanning genereert aan de uitgang die verbonden wordt aan de ingang van de volgende. Op die draad ontstaan storingen. De spanning die de andere poort aan de ingang ziet kan hoger/lager zijn dan die gegenereerd door de eerste poort. Wanneer dat naar boven gaat is dat niet erg, maar als dat te laag gaat kan de ontvangende poort dat niet meer correct interpreteren. Die ruismarge dient dus om storing op te vangen. Daardoor wordt uw schakeling minder gevoelig voor storingen. Bij CMOS: hoogbereik: tussen 3.5 en 5V. Je hebt 0.9V ruismarge want de meesten gaan tussen 4.4 en 5V genereren. Bij CMOS is dat mooi symmetrisch maar dat hoeft niet zo te zijn. Bij TTL (transistor transistor logic) is dat niet symmetrisch.

**Slide 25:** Niet-ideale gedrag van de schakeling: als je een inverter bekijkt en je gaat naar de spanningen kijken. Horizontaal: ingangsspanning, verticaal: uitgangsspanning. L = laag bereik aan in- en uitgang, H = hoog bereik aan in- en uitgang. Wat in het groen aangegeven is is de normale werking: als er een hoog signaal aan de ingang is, zal er een laag signaal aan de uitgang zijn. Als we schakelen (overgaan van 0 naar 1 of omgekeerd) moeten we overgaan van het ene groene stuk naar het andere. We noemden het stuk ertussen ongedefinieerd, maar die zijn niet ongedefinieerd, die spanningen. De spanning zal niet met een 0 overeenkomen, ook niet met 1, ook niet met tristate, dat heeft logisch gezien geen enkele betekenis. Als je  $V_T$  aanlegt weet je totaal niet wat uw spanning gaat doen, ze gaat in elk geval niet meer digitaal werken, het zal zich analoog gedragen: als een versterker: kleine veranderingen aan de ingang hebben een grote uitgang. Het ideale zou zijn als we geen curve hebben zoals getoond, liever zouden we iets ideaal hebben: hoog tot aan de drempelspanning ( $V_T$ , wanneer schakelen we over van laag naar hoog). Zodra we over die drempelspanning zijn, zullen we van laag naar hoog gaan. In het tussengebied is de uitgang ook niet gedefinieerd. Het probleem van in het ongedefinieerde gebied te zitten: de schakeling werkt daar niet op een digitale manier en we kunnen niet voorspellen



wat de schakeling daar doet. We moeten daar aan heel hoge snelheid doorgaan: we hebben even een probleem, maar we zijn er onmiddellijk voorbij. We weten dat er bij elke overgang heel even in het gebied gezeten zal worden.

Een ander probleem is dat als je kijkt naar CMOS: dynamisch vermogenverbruik: bij tussenin zal niet 1 van de schakelaars open/gesloten zal zijn, ze zullen alletwee een beetje openstaan en er beiden een beetje tussenin zitten: gaat door alle tussenliggende waarden. Dat is het moment dat zowel de bovenste als onderste tak kunnen geleiden, waarbij je vermogenverbruik krijgt. Dus in het rode gedeelte ga je ook bij CMOS vermogenverbruik hebben. Je hebt dit vermogenverbruik ook alleen wanneer je een overgang maakt van de ene naar de andere toestand (vandaar dynamisch). Die curves zijn ook niet constant, die kunnen variëren omwille van het productieproces maar ook door temperatuursvariaties.

**Slide 26:** Het is belangrijk om zo snel mogelijk door dat gebied te gaan. Normaal zorgen we ervoor dat aan de ingang die signalen redelijk snel veranderen, dus we gaan daar maar korte tijd inzitten, in die overgang. Jammer genoeg hebben we dat niet altijd in de hand. Als je het signaal opmeet dat van de buitenwereld komt (je hebt een sensor die iets meet en de snelheid waaraan dat verandert wordt door de omgeving bepaald), dan bepaal je dat niet zelf. Het kan dus zeker zijn dat je traag variërende signalen hebt aan de ingang. Dat soort signalen heeft ook typisch wat ruis/verstoringen. Dat gaat niet mooi van laag naar hoog, daar zit variatie op. Als je het reële signaal in een digitale schakeling binnenbrengt en je steekt dat in een inverter, dan krijg je de tweede bovenste grafiek (de middelste). Rond die drempelspanning zal dat werken als een versterker en kleine veranderingen aan de ingang worden grote veranderingen aan de uitgang. Als je dat signaal verder gaat verwerken, dan gaan bepaalde van die piekjes misschien verdwijnen. Dus in plaats van 1 enkele overgang, wat je normaal wilt, ga je heel wat verschillende overgangen zien. Dat is niet wat we willen: we willen een signaal dat op een bepaald moment van laag naar hoog gaat.

Oplossing: Schmitt-trigger-ingangen: speciale implementaties van een ingang: we hebben niet 1 overgang, maar als we van hoog naar laag gaan volgen we de rechtse curve op de rechtsonderfiguur en als we van laag naar hoog gaan volgen we de linkse curve op de rechtsonder figuur.

**Slide 27:** Het dynamisch gedrag is heel belangrijk want het bepaalt het tijds-gedrag.

**Slide 28:** Wat hier getekend is, is onmiddellijk nadat er overgeschakeld is: de rode geeft geen doorverbinding meer en de bovenste wel. Je zou verwachten dat de spanning onmiddellijk verandert, maar dat gebeurt niet omdat er vertragende elementen in de schakeling zitten. Als je het elektrische equivalent tekent dan is dat het rechtse: je hebt 2 weerstanden (de rechtse: de ingangsimpedantie, wat je aan de ingang ziet en dus niet de weerstand van de rechtsondergroene transistor op de bovenste tekening). Het is de weerstand tussen de gate en de source. Vermits die van elkaar geïsoleerd zijn, verwacht je dat die zeer hoog/oneindig is. Er is ergens een weerstand tussen de gate en de source. Capaciteit: je hebt er een wanneer je 2 geleiders naast mekaar hebt, dus 2 draden boven elkaar (meestal geen grote capaciteit). Dus overal in de schakeling heb je eigenlijk een

capaciteit.

Overall In de schakeling heb je een grond en een verbinding.

Hoe langer uw draad is, hoe groter uw capaciteit is: bij lange draden is die groot, bij korte draden is die capaciteit klein. Bij de onderste overgang heb je iets vergelijkbaar. Deingangsimpedantie blijft altijd.

**Slide 29:** We hebben dit schema, wat we nu gaan doen is een plotse overgang maken van laag naar hoog. Dat heeft tot gevolg dat waar er oorspronkelijk geen spanning op stond, een spanning wordt aangebracht en die schakeling gaat beginnen reageren. Dit is een RC-netwerk. De spanning gaat exponentieel oplopen. De spanning tussen de uitgang van de ene poort naar de ingang van de andere gaat geleidelijk overgaan van een laag niveau naar een hoog niveau. De uiteindelijke spanning is niet de blauwe.

De eindspanning zal dan ongeveer gelijk zijn aan de voedingsspanning uiteindelijk.

Voor de rest zie je nog dat die overgang niet ogenblikkelijk gebeurt: het duurt een poosje voor je van het laag bereik in het hoog bereik terechtgekomen bent. Dat is heel belangrijk: de snelheid waaraan je kan werken hangt af van hoe snel je uw bewerkingen afwerkt. Je kan uw spanning niet laten zakken als ze nog nooit hoog geworden is. Hoe trager die curve dus omhoog gaat, hoe langer je moet wachten. We streven er dus naar dat de groene curve zo snel mogelijk stijgt. Dat hangt af van die tijdsconstante  $\tau$ . Die wordt bepaald door de capaciteit en de twee weerstanden: de parallelschakeling ervan.

Hetzelfde gebeurt wanneer je een  $H \rightarrow L$  overgang hebt.

**Slide 30:** Nu we dat weten, hoe kunnen we er nu voor zorgen dat die overgang zo snel mogelijk gebeurt? Dieingangsimpedantie moet zo hoog mogelijk zijn (groter dan de uitgangsimpedantie). Als uwingangsimpedantie zeer groot is, dan is het statisch vermogenverbruik verwaarloosbaar klein.

Hoe kiezen we onze uitgangsimpedantie? Hier zitten we met een conflict: stel dat deingangsimpedantie zeer hoog is, dan zal de tijdsconstante zich verhouden tot  $R_{OH} * C$ . We willen  $R_{OH}$  dus zo klein mogelijk hebben.

Nadeel aan te kleine weerstand: hoe kleiner die is, hoe groter het ogenblikkelijk vermogenverbruik: ogenblikkelijk grote stromen ( $P(0)$ ). Je hebt dan over de uitgangsimpedantie de volledige voedingsspanning. Dat zorgt voor een groot vermogen. Het is maar ogenblikkelijk, maar als je plots een heel hoge stroom nodig hebt, kan uw voeding in de problemen geraken want die moet die stroom kunnen leveren (voor 1 is dat geen probleem, wel als je er meerdere hebt). Je moet dus dikwijls een compromisoplossing hebben.

Je kan ook proberen van uw  $C$  zo klein mogelijk te maken door korte draden te gebruiken: lange draden zorgen ervoor dat je traag moet beginnen werken.

Uitgangsimpedantie die je zo klein mogelijk wil ( $R_0$ ): de NMOS-poort is niet erg populair want die heeft statisch vermogenverbruik, maar die is ook qua tijdsgedrag enorm slecht als je dat vergelijkt met CMOS. Waarom is die traag? Je kan geen te kleine weerstand zetten want dan heb je een enorm groot statisch vermogenverbruik, je moet dus een grote weerstand hebben, maar dan heb je dus een slechte tijdsconstante.

Je gebruikt NMOS dus niet tenzij je die echt nodig hebt om die open drain te implementeren.

**Slide 31:** Dynamisch gedrag van de poort: hoe snel kan je stijgen en dalen: stijgtijd: tijd om van 0 naar 1 te gaan, daaltijd: tijd nodig om van 1 naar 0 te gaan, daar is een zekere tijd voor nodig. We hebben niet een lage spanning en een hoge spanning maar een laag bereik en een hoog bereik, je kan dus meerdere dingen meten: hoe lang duurt het om van het einde van het laag bereik naar het begin van het hoogbereik te gaan,...

De stijg- en daaltijd is iets wat niet zo specifiek is aan de poort. De poort speelt daar ook in mee omdat die capaciteit deels door de eigenschappen van de poort bepaald wordt, maar dat heeft weinig/niks met de complexiteit van de poort te maken. Dat is niet het enige wat het tijdsgegedrag bepaalt, het bepaalt hoe snel je kan overschakelen van laag naar hoog of omgekeerd, het bepaalt de maximale snelheid waaraan we kunnen werken. Je ziet op die tekening ook dat die stijg- en daaltijd nogal verschillend kunnen zijn, ze wordt bepaald door de weerstand bovenaan (stijgtijd) terwijl de daaltijd bepaald wordt door de transistor (denk ik??). Vertragingstijd: ik zet een signaal op de ingang, hoe lang heeft die poort ervoor nodig om die berekening te doen en dat op de uitgang te zetten? Dat wordt gemeten als wat er op de ingang verschijnt en wat er aan de uitgang uitkomt.

Het probleem is, hoe meet je wanneer de ingang gebruikt wordt en wanneer het resultaat op de uitgang beschikbaar is? Je hebt ook weer 2 soorten vertragingstijd. In vele gevallen is dit ongeveer hetzelfde, in andere gevallen kan dat verschillen. Als je over *de* vertragingstijd spreekt, spreekt men over het gemiddelde.

Strikt gezien zou je kunnen zeggen dat de vertragingstijd gekoppeld is aan het interne van de poort. De stijg- en daaltijd heeft te maken met hoe poorten met elkaar verbonden zijn. Stijg- en daaltijd bepalen hoe snel je kan werken. Wanneer je dat op deze manier gaat meten, in de gemeten vertragingstijd zit de echte vertraging van de poort in plus hoe snel die verandert.

**Slide 32:** Korte stijgtijden links: geeft de tijd aan die nodig is om het resultaat te berekenen. Als je een slechte stijgtijd hebt (rechts), dan komt er nog de helft van de stijgtijd bij.

In de gemeten vertragingstijd zitten dus beide effecten in: complexiteit van de poorten en hoe ze met elkaar verbinden. Vandaar dat ook, als je wil gaan rekenen, kan je de vertragingstijd meten en met beide aspecten gelijktijdig rekening houden. In de gemeten vertragingstijd ga je dus altijd een combinatie van de twee hebben.  $\Rightarrow$  Heel belangrijk!

**Slide 34:** Vermogenverbruik: je moet het vermogen kunnen leveren (dat kost aan energie, batterijen,...) en dat vermogen wordt voor iets gebruikt: wordt gewoonlijk omgezet in andere energie, die gaat niet echt verloren. In ons geval betekent dat dat het omgezet wordt in warmte en die moet afgevoerd worden. Als je die niet snel afvoert, gaat uw schakeling enorm snel opwarmen en smelten. Je moet energie dus niet alleen kunnen leveren, ook wegnemen. Voor grote schakelingen is dit ondertussen een van de belangrijkste problemen geworden. Als je weet dat iets op een bepaald moment niet gebruikt wordt, ga je er gewoon de voeding vanaf nemen, maar als je het dan nodig hebt, heb je tijd nodig om in gang te geraken. Voor de rest zijn er heel wat technieken om schakelingen te maken die heel wat minder vermogen verbruiken.

Je hebt dus statisch vermogen: vermogen dat je constant hebt en dat je zeker wil vermijden (dat heb je bij NMOS), bij CMOS is dat verwaarloosbaar klein (de belangrijkste reden om CMOS te nemen). Ze hebben beiden dynamisch vermogenverbruik: als ze schakelen: die capaciteiten moeten opgeladen worden en ontladen worden.

Je kan nagaan hoeveel energie je nodig hebt om een capaciteit van de ene spanning naar de andere te gangnen. Dat is  $CV^2/2$ . of je dat nu snel doet of traag maakt niet uit, het verschil zit 'm wel in dat als je het snel doet je ogenblikkelijk veel energie nodig hebt op korte tijd. Om die van de ene spanning naar de andere te brengen is  $CV^2/2$ .

De meeste systemen werken op een klok, hoe dikwijls per seconde kunnen we het signaal veranderen? 2 overgangen (van nul naar 1 en omgekeerd):  $CV^2$  aan energie nodig. Voor praktische gevallen kan het dynamisch vermogen beperkt gehouden worden. Hier moet je er ook rekening mee houden dat niet alle poorten tegelijkertijd schakelen, er is geen enkele schakeling waarbij alle poorten tegelijkertijd schakelen.

**Slide 35:** In alle schakelingen moet de PMOS vanboven staan en NMOS vanonder, waarom is dat? Als we die zouden omwisselen, kunnen we niet alleen inverterende schakelingen maken, maar ook een AND poort.

**Slide 36:** Waarom mag dat niet? NMOS-transistoren zijn zeer slechte pull-ups. Een NMOS-transistor is goed om een spanning naar beneden te trekken, PMOS om naar beneden te trekken, daarom staan ze ook in die takken.

Waarom is dat zo? Er zijn verschillende redenen. De meest evidente: stel dat we NMOS bovenaan zouden zetten: wanneer gaat de transistor geleiden bij de middelste tekening? Als de spanning voldoende groot is. Om ervoor te zorgen dat we een hoger niveau krijgen moet er een verschil zijn van de drempelspanning van de transistor tussen de gate en de source. Als je bij V een hoge spanning opzet, zal die nooit boven  $-V_T$  kunnen geraken, moest die hoger geraken zou de transistor niet meer geleiden. In regimetoestand gaat de uitgang van de linkse poort de hoge spanning  $-V_T$  hebben. Wanneer we een bepaalde spanning hierop hebben, wat die ook is, op de uitgang zal altijd een lagere spanning staan. Als je 2 van die poorten na elkaar hebt, dan gaat er hier eerst  $V_T$  af en daarna daar  $V_T$  erbij. Na een tijd ga je een spanning hebben die lager is dan het lage niveau. Is dat hetzelfde probleem als je die beneden zet? Nee, want daar is het niet de spanning tussen de gate en de uitgang maar tussen de gate en de source en de source ligt daar altijd op nul.

Nog een verklaring: die gaat zich zo instellen dat die altijd een beetje blijft geleiden dus gaat altijd ook statisch vermogenverbruik hebben. De boodschap van de dia is dat je nooit NMOS transistoren bovenaan mag zetten, je moet die altijd onderaan zetten. Hetzelfde met PMOS: nooit onderaan.

Daarom kunnen we alleen maar inverterende basispoorten maken.

**Slide 38:** Fan-in: aantal ingangen. Waarom komt dat van pas? Je kan niet een eindeloos aantal ingangen hebben, je kan geen poort maken met 100 ingangen omdat hoe meer ingangen je hebt, hoe slechter de poort gaat werken: de vertragingstijd gaat stijgen: hoe meer ingangen je hebt, hoe meer transistoren je hebt (ingangen  $\sim$  transistoren), ieder van die transistoren heeft een interne

capaciteit,...

Ander probleem: als je kijkt naar de tekening rechts boven: een groot aantal ingangen betekent een groot aantal transistoren in serie. Ieder van die transistoren is niet-ideaal: zelfs als die geleidt zal de spanning over die transistor niet exact 0V zijn. Als je er verschillende boven elkaar zet gaat dat allemaal optellen en kan je in het hoog bereik terechtkomen hoewel dat niet de bedoeling is. We gaan daar rekening mee moeten houden, een oplossing voor mogen zoeken.

Fan-out (!!!):  $\neq$  aantal uitgangen. Fanout = # ingangen dat aan die uitgang verbonden is. Waarom zijn we daarin geïnteresseerd? Ook dat zal onze snelheid gaan bepalen: hoe meer ingangen er aan de uitgang hangen, hoe groter de capaciteit, hoe trager de schakeling werkt. Er is maar een maximale stroom die geleverd kan worden. Dus hoe groter de fanout, hoe trager de schakeling gaat werken. Soms kunnen we daar niet aan ontsnappen: zolang je op een geïntegreerde schakeling werkt, heb je niet veel energie nodig om het te laten omschakelen. Een keer je vanaf een geïntegreerde schakeling naar een andere gaat, heb je een lange verbinding en heb je een veel grotere capaciteit. Wil je dus een hogere snelheid halen op een PCB moet je een andere methode zoeken. Het voordeel van op de chip te werken is dat je aan een grote snelheid kan werken. Gegevens uitwisselen met de buitenwereld (dus buiten de chip) gebeurt niet aan een hoog tempo om deze reden.

Je kan dat wat verhelpen door ervoor te zorgen dat je een schakeling hebt die wel grote stromen kan leveren: een poort die een hoge uitgangsimpedantie heeft (denk ik? Mss kleine!). Een driver/buffer dient om veel stroom te kunnen geven. Die stroom heb je nodig om dingen met een hoge capaciteit toch snel te kunnen laten veranderen. Een keer je vanaf uw schakeling naar uw PCB gaat of van PCB naar randapparatuur heb je buffers nodig die voldoende stroom kunnen leveren om een redelijke snelheid te halen.



# Chapter 4

## Les 4

### 4.1 Sides: 3\_Technologie

**Slide 39:** We hebben besproken hoe poorten gemaakt worden in NMOS en CMOS en welke impact de parameters hebben. Buiten de praktische kanten is er nog een ander aspect: de technologieën: hoe wordt dat geïmplementeerd? Er zijn ook heel wat andere mogelijkheden.

**Slide 40:** Als je dat gaan bekijken zijn er 3 grote categorieën die gebruikt worden als het over implementatie gaat: standaard chips, programmeerbare logica en specifieke chips.

Standaardchips zijn de oudste. De bedoeling was dat je de componenten kocht en die zelf samenzette op een PCB om op die manier een schakeling te maken. Ondertussen is dat wat in onbruik geraakt (die standaardcomponenten worden wel nog gebruikt, voornamelijk als lijm tussen de andere: glue logic). Als je naar een PCB kijkt, dan staan daar een paar belangrijke componenten op die zorgen dat de chip iets specifiek doet. Je hebt een component met verwerkingskracht, uitbreiding, ... Ook een klok nodig die gegenereerd moet worden, buffers om meer stroom te leveren wanneer je naar andere componenten gaat of wanneer je van het bord afgaat. Dat is typisch die glue logic, de rommel die erbij staat die functioneel niet erg belangrijk is maar toch nodig is om het te doen werken. Daar zit niet veel evolutie meer in en we gaan dat verder ook niet meer bespreken.

Specifieke chips: andere kant van het spectrum. Dat is een enkele component die gans het systeem omvat. Het nadeel ervan is dat het erg duur is om te maken. Het ontwerpproces is redelijk lang, vereist heel wat kennis en het maken vergt ook veel kennis en tijd en energie. Daarom ga je dat alleen doen als je die ontwerp-kost kan afschrijven op een heel groot aantal componenten. Bv. als je een microprocessor maakt kan je die voor vanalles en nog wat verkopen.

Je hebt ook een tussencategorie voor de kleine aantallen of voor een beperkt aantal ((tien)duizenden componenten), daarvoor zijn de specifieke chips te duur en de standaardchips zijn te eenvoudig. Daarvoor hebben we programmeerbare logica (is geen microprocessor of waarop je kan programmeren), het is een logica waarop je de functionaliteit kan programmeren. Je kan de functionaliteit van de component wijzigen. Dat heeft het voordeel dat je de component voor heel wat toepassingen kan gebruiken. Door de stijging van het integratieniveau kunnen

we daar heel hoge complexiteiten mee halen.

**Slide 41:** Als we iets specifieker gaan kijken, de specifieke chips, daarbij is maatwerk het meest efficiënte: zo maken dat die optimaal werkt qua snelheid,...

**Slide 42:** Je gaat zelf bepalen waar de transistoren allemaal staan, hoe die verbonden zijn, hoe die poorten gaan vormen,... Dat is heel complex (je gaat de ganse schakeling in een keer maken) maar het is het meest efficiënte. Door het feit dat je geen overzicht meer hebt, is dat niet goed voor grote schakeling. Het kan wel gebruikt worden als onderste laag die de fabrikant meelevert.

Uit de wet van Moore volgt dat er om de 1.5-2 jaar een verdubbeling is van integratie. Als je op het laagste niveau van transistoren gaat werken, ben je verplicht om uw ontwerp om de 2 jaar helemaal opnieuw te maken, anders ben je niet meer mee.

**Slide 43:** Je kan gebruik maken van standaardcellen (het niveau juist erboven, boven de poorten. De poorten kan je uit een bibliotheek nemen). Om het op een structurele manier te regelen wordt het gebruikt in cellen. De hoogte is hetzelfde, de breedte varieert wat er tussen de rijen zitten gaten, daar worden geen transistoren gebruikt. Die dienen om de draden te leggen. Je hebt maar een beperkt aantal metalisatielagen die je boven elkaar gaan leggen, je kan draden niet zomaar boven elkaar leggen/elkaar laten kruisen (zou het veel complexer maken), daarom voorzie je plaats tussenin, dan moet je niet in de hoogte werken. Hoe geberut het ontwerp hier? Je moet eerst een logisch ontwerp maken met alle poorten etc. De volgende stappen zijn:

1. Bepalen waar welke poort ligt: het is niet omdat je 3 AND-poorten hebt en een OR-poort dat je eerst de ANDs moet leggen en dan de OR. De vraag is wat het beste is: hoe kunnen die verbonden worden met elkaar? We hebben bij voorkeur korte draden zodanig dat de schakeling zo snel mogelijk werkt.
2. Het plaatsen van die componenten hangt samen met hoe die draden kunnen lopen. → Bepalen waar de draden liggen.

Die twee dingen liggen samen, ontwerp je samen. Dit is een niveau dat al veel makkelijker is om ontwerpen te maken. Dit wordt heel toegepast op een normaal ontwerp.

**Slide 44:** Je kan nog een stap verder en dat is de gate array. Het gaat er altijd om hoe je het goedkoper en sneller kan laten gebeuren. Je kan een beperking opleggen aan het aantal componenten.

Waarom geen component maken met alleen maar NAND-poorten op? Het voordeel daarvan is dat je dat één keer op voorhand kan maken. Welke schakeling je ook wil hebben, de poorten liggen al vast (vandaar een gate array). Wat moet er dan nog gebeuren? De verbindingen moeten nog gelegd worden. Het grote voordeel is dat de meest dure en complexe stappen al gedaan zijn. De laatste stappen zijn die draadjes leggen bovenaan, dat is goedkoop en eenvoudig. De fabrikant kan een ganse reeks chips maken waar al die poorten opstaan en per gebruiker zal er een andere metalisatie op gelegd worden.

Het legt beperkingen op aan wat je kan gebruiken.



**Slide 45:** Een andere manier van implementeren die redelijk veel gebruikt wordt (en we gaan toepassen in het labo) en dat is die van programmeerbare chips. We gaan hier geen transistoren verschuiven.

**Slide 46:** Wat je wel kan doen is: stel de poorten liggen vast, je kan de verbindingen wijzigen. Je kan zo andere schakelingen maken. Je kan dit ook doen bij een component die helemaal klaar is. Je kan er wel voor zorgen dat verbindingen wel of niet onderbroken zijn (dat is hetzelfde als een verbinding leggen, maar het is soms makkelijker om te onderbreken).

- Zekeringen: voorkomen dat er teveel warmte ontstaat en daarom schade. Een goede zekering is het eerste wat kapot gaat in uw toestel. Als er een te grote stroom door de zekering loopt, dan brandt die door. Je kan een zekering zetten op een draad en als we niet willen dat daar een verbinding is, dan zetten we daar even een grote stroom op zodat die zekering doorbrandt. Het nadeel is dat dit onomkeerbaar is. Als je later vaststelt dat je een fout gemaakt hebt kan je die in principe niet meer wijzigen, tenzij het enige wat moet gebeuren is het onderbreken van extra verbindingen. We veranderen hier dus niet de poorten, enkel de verbindingen.
- Op een reversibele manier: we kunnen een transistor laten geleiden of isoleren. Dat komt op hetzelfde neer als de zekering die wel of niet doorgebrand is. Als je zo een transistor op een draad zet en je stuurt die op een juiste manier, kan je ervoor zorgen dat die verbinding er wel of niet is. Je moet alleen de spanning op de gate veranderen om te zorgen dat de verbinding al dan niet gemaakt is. In de meeste gevallen blijf je wel bij uw keuze, dus ook als je de spanning afzet en terug opzet, wil je dat dat behouden blijft, je wil dat op een niet-vluchtige manier doen. Je kan dezelfde techniek als in flashgeheugen hier toepassen: 2 gates: de bovenste is verbonden met de buitenwereld, de middelste is totaal geïsoleerd (floating gate). Hoe kunnen we die nu gebruiken? Dankzij quantummechanica weten we dat het kan gebeuren dat een elektron op een bepaald moment op een plaats is en op een ander ogeblik ergens anders zonder dat het ergens tussenin geweest is. We kunnen daarvan gebruik maken. Door de juiste spanning op te leggen kunnen we ervoor zorgen dat elektronen die zich op de bovenste poort bevinden zich plots op de onderste poort bevinden. We gaan elektronen overpompen van boven naar die middenste poort. Als je de spanning afzet, kunnen de elektronen niet meer wegtunnelen (dat kan alleen als je de juiste spanning aanlegt). Als je dan de spanning opnieuw aanlegt, dan zullen die elektronen weer geleiden. Je kan dit ook omkeren: die elektronen terug naar de bovenste poort laten gaan. Op die manier kan je een transistor gebruiken om verbindingen te maken en onderbreken. Je kan dit omkeren, maar dat gaat niet zo simpel (vrij traag). Je kan dat ook niet blijven doen, uiteindelijk krijg je daar toch neveneffecten van en dan werkt dat niet zo goed meer, maar normaal wil je dat ook niet.
- Tegenwoordig vind je SRAM het meeste terug: aansturen vanuit een geheugen. We hebben een transistor die voor de verbinding zorgt, een geheugen dat aangeeft of de transistor moet geleiden of niet. Als dit een statisch RAM-geheugen is, gaat de informatie verloren wanneer je de spanning afzet. Je kan ergens bijhouden (in een niet-vluchtig geheugen)

wat wel en niet moet doorlaten. Tijdens het resetten wordt die informatie gewoon overgepompt naar uw SRAM. Het voordeel daarvan is dat het herprogrammeren heel vlot verloopt: gewoon nieuwe informatie in het RAM-geheugen inladen. Dat kan aan hoge snelheid en je kan het vaak veranderen. In extremis kan je die component dat zelf laten doen, maar dat is erg moeilijk want het is moeilijk bij te houden wat er nu juist aan het gebeuren is.

**Slide 47:** Hoe ziet dat eruit? Een laag AND-poorten en een laag OR-poorten. Je hebt dus een laag AND-poorten gevolgd door een laag OR-poorten, daarmee kan je elke mogelijke schakeling maken. Het enige wat je hierbij moet doen is programmeren hoe de ingangen met elkaar verbonden zijn (via de uitgangen). Dat is dan het programmeren en daarmee kan je een schakeling maken.

Als we gaan kijken naar de kostprijs: de poorten gaan niet meer de kostprijs bepalen, wel de plaatsen waar je al dan niet verbindingen maakt. Het wordt dikwijls zo getekend, hoe interpreteer je dat: een beknopte schrijfwijze om de rechtse tekening weer te geven. Tussen ieder van die dingen (bolletjes) zit een zekering. Je kan er sommige dan opblazen en dan krijg je kruisjes en dan weet je wat al dan niet gaat geleiden.

Het is een XNOR-poort op de slide. Kruisje: zekering die is doorgebrand.

Problemen: snelheid en betrouwbaarheid. Hoe groter die matrices worden, hoe langer de afstanden worden en hoe trager het systeem wordt. De snelheid is dus veel lager dan de rechtstreekse implementatie.

Betrouwbaarheid: hoe test je een zekering: opblazen en testen of het goed gebeurd is. Een keer je dat gedaan hebt moet je die eigenlijk weggooien. Je kan dat dus nooit helemaal testen: bij een paar testen en aannemen dat de rest ongeveer ook zo zal zijn.

Die lagere snelheid dan rechtstreekse verbinding en onbetrouwbaarheid zorgt ervoor dat men die volledige flexibiliteit niet altijd zal gebruiken, sommige dingen zal men vastleggen: de OR-poorten onmiddellijk aan de AND-poorten hangen bv.

Dan krijgt dat andere namen: PAL: vaste OR-matrix.

**Slide 48:** Naarmate dat we naar een hogere integratie konden gaan, is men gaan beseffen dat je alle kan implementeren met tweelagen-logica, maar als je bv. een flipflop wil gaan maken, heb je daar 8-10 poorten voor nodig. Als je dat via de trage AND-OR verbindingen gaat doen, werkt dat heel traag. Als je weet dat je toch in 80% van de gevallen een geheugenelement nodig hebt, zet dat er dan gewoon bij. In die overige 20% ga je dat niet gebruiken en staat het wat in de weg, maar voor die 80% verantwoordt het wel die extra meerkost.

Men gaat de dingen waarvan men verwacht dat die regelmatig gaan gebruikt worden, die gaat men erbij zetten. Op de slide: 1-bitgeheugen erbij gezet en een selector die zegt of het geheugenelement gebruikt wordt of niet. Men spreekt dan van macro-cellen: de uitgang wordt complexer en er zitten verschillende functionaliteiten in.

We kunnen alsmaar meer op een chip zetten. Als we die matrix gaan uitbreiden, gaat de chip eerder met de matrix volstaan, niet zozeer met de schakelingen. Je hebt altijd een paar componenten die met elkaar verbonden moeten zijn, maar niet alles moet altijd met alles verbonden zijn. Die heel grote matrix is dus niet

altijd efficiënt.

**Slide 49:** Soms is het interessanter om verschillende kleinere dingen te nemen. In plaats van alles groter te maken, gaan we verschillende kleine matrices gebruiken zodanig dat die lokaal verbonden kunnen worden en weinig verbonden moeten worden met de rest. Sommige dingen gaat men dan onderling wel nog gaan verbinden.

**Slide 50:** Wat we tot hiertoe gedaan hebben is de logica, de poorten die erop staan, die staan vast. Het enige wat we konden programmeren waren de verbindingen. We kunnen nog een stap verder gaan, dat is de meer recente programmeerbare logica. We kunnen niet alleen de verbindingen programmeren, maar ook de poorten. Dan komen we tot de FPGA.

Als je gaat kijken heb je een matrix van logische blokken. Dat is die basiscomponent waarvan je ook de functionaliteit (die nog niet bestaat) kan bepalen.

Het is field programmable: gelijk waar, je moet daarvoor niet naar de fabriek gaan, geen heel speciale apparatuur voor nodig. Het is in dit soort schakelingen dat daarom die transistoren aangestuurd worden vanuit een geheugen.

Wat zit daarin? Je hebt enerzijds de logische blokken en iedere logische blok kan een bepaalde functionaliteit hebben. Je kan er natuurlijk maar iets mee doen als die verbonden zijn met elkaar. Je hebt 2 soorten verbindingen:

- Lange lijnen die lopen over de ganse component want soms kan het gebeuren dat je informatie moet delen met heel wat andere logische blokken en dan ga je typisch een busstructuur gebruiken en die gaat op verschillende plaatsen gebruikt worden. Op die ogenblikken is zo'n lange lijn heel interessant.
- Als je toch snelheid wil hebben, moet je kortere verbindingen hebben. Daarvoor gebruiken we de schakelmatrixes: die gaan u de mogelijkheid geven om elke lijn die erop toekomt te verbinden met elke andere lijn, die kan dat zelfs met meer dan een andere lijn verbinden. De lichtere dingen die er nog tussenstaan (SMC) dienen om de horizontale en verticale lijnen te raken, de verbinding naar de lijnen die naar de schakelmatrixes gaan.

Je hebt bijkomend heel korte verbindingen die rechtstreeks 2 naburige blokken verbinden.

Je moet nog naar de buitenwereld gaan, dus je hebt ook IO-blokken nodig.

Als je dat gaat bekijken binnen een ontwikkelomgeving: de blauwe dingen op slide (via animatie komt die erop te staan): logische blokken. Paarse lijnen zijn de lange lijnen die over de ganse lengte doorlopen. Om een bus aan te sturen met meerdere, dan moet dat met tristate buffers, dat zijn die kleine driehoekjes. Die staan erop, of je die nu gebruikt of niet. De witte lijnen zijn de kortere lijnen. De schakelmatrixes maken dus gewoon verbindingen tussen lijnen. De blauwe puntjes (nauwelijks zichtbaar) maken ook verbindingen.

**Slide 51:** Als je weet dat bepaalde dingen veel gebruikt worden, ga je dat niet met basiscomponenten implementeren, dan zet je dat daar gewoon bij. Gewoonlijk ga je sequentiële schakelingen maken. Wat heb je daarvoor nodig? Geheugen (op die van de labo's zit extra geheugen op: het groene op de tekening). Sequentiële schakelingen werken op een klok, dus dat heb je ook nodig

(zit er automatisch op, dat is dat zwart blokje DCM). Dat kan nog verdergaan: vermenigvuldiger. Een optelling kan je redelijk efficiënt met die logische blokken maken, een vermenigvuldiger niet, daarom zijn er vermenigvuldigers bijgezet. Als je naar complexere toepassingen gaat kijken kan er veel meer opgezet worden: meer geheugen, alle componenten die je denkt nodig te hebben: aansturing van ethernet,...

Wat als je bepaalde componenten niet nodig hebt? Je hebt verschillende soorten FPGA's: met bepaalde componenten of zonder.

**Slide 52:** We hebben 2 manieren om te programmeren: het maken van verbindingen en bijkomend een basiscomponent waarvan we de functionaliteit ook kunnen aanpassen. Dat geeft enorm veel flexibiliteit.

Daar zitten dus geen poorten in waarvan je de schakeling kan wijzigen!

Wat hier is weergegeven: wordt geïmplementeerd met een geheugen: een opzoektabel implementeren. Dat geeft voor elke combinatie van de ingangen (hier 4 bits ingangen dus 16 mogelijkheden) aan of de uitgang 1 of 0 moet zijn. Die lookup table bepaalt dus de functionaliteit.

Wat je nu ziet is de waarheidstabel. Op het moment dat je de waarheidstabel hebt moet je niet meer nadenken wat het wordt in termen van AND en OR, je steekt dat in het geheugen en de component doet wat hij moet doen. Je hebt wel de beperking dat je meestal maar met 4 ingangen kan werken. Elke logische blok kan dus gelijk welke functie van 4 ingangsvariabelen realiseren. Dit is de manier waarop je een programmeerbare component kan maken.

**Slide 53:** Spartan-3: je hebt een logische blok (CLB: configurable logic block) en elke slice is opgedeeld in 2 logische cellen, zie **Slide 54**.

**Slide 54:** Binnen 1 configureerbare logische blok zitten 8 logische cellen in, dus je gaat altijd 8 logische cellen met een keer gebruiken. Als je hiernaar kijkt, kan je denken dat het nodeloos ingewikkeld gemaakt is. Ook hier geldt: wat zetten we erbij? → Dingen die we goed denken te kunnen gebruiken.

Door er een beetje hardware bij te zetten (in het lichtgroen) kan je met een logische cel een optelling maken. Het is de moeite waard om dat te hebben.

Evengoed hebben we dikwijls extra geheugen nodig, dat is in het blauw aangegeven.

Dat rode stukje: die kiest tussen 2 ingangen. Door alleen deze erbij te zetten kan je die 2 logische cellen combineren tot iets nieuw. Als je dus 2 logische cellen hebt, heb je 2 keer een functie van 4 variabelen. Maar heel dikwijls heb je meer variabelen nodig. Door die rode selector erbij te zetten, kan je die 2 samenzetten tot 1 functie van 5 variabelen: je splitst uw waarheidstabel in 2 op: 1 waarbij de meest beduidende variabele 0 is en een waarbij de meest beduidende variabelen 1 is. Op die manier, door het feit dat die selector daar wijzigt, kan je een functie van 5 variabelen realiseren.

**Slide 55:** Schakelmatrices: de verbindingen worden gemaakt/verbroken met transistoren. Hier is het getekend met NMOS-transistoren, in de realiteit is dat ingewikkelder. Je kan het programmeren vanuit een RAM-geheugen dat die poorten aanstuurt.

**Slide 56:** Je hebt een ontwikkelomgeving specifiek voor het soort componenten dat de fabrikant verkoopt.

**Slide 57:** CAD-omgeving waarbinnen je het volledige ontwerp kan doen en dat je zoveel mogelijk zal helpen om alle taken te doen. De basistechnieken zijn al geïmplementeerd in de software en zullen voor u verzorgd worden. Linksboven: project met bestanden. Je kan daar hiërarchisch in werken. Je kan daar verschillende dingen op doen. Je hebt ook een venster met het resultaat. Rechts: overzichtsschema met ofwel schema ofwel resultaat van simulaties.

**Slide 59:** Je bent niet verplicht om gebruik te maken van 1 manier van input ingeven. In het voorbeeld op de slide is alles wat gecombineerd. Je hebt een topschema met een paar dingen die al een beetje uitgewerkt zijn, andere dingen zijn nog gewoon blokjes met een subschema waarop je kan klikken, dan wordt een lager niveau geopend en kan je dat ook in detail zien.

Ook nog andere dingen die afhangen van fabrikant tot fabrikant zoals bv. die logiCORE, soort macro. Om dat allemaal afzonderlijk in de bibliotheek te hebben, ziet men een soort macro. Dan wordt er nog een venster geopend waarin je instellingen kan maken (zie transitie slide). Dit komt een beetje overeen met VHDL. Je moet niet weten wat daar allemaal onderzit, de omgeving zal daar rekening mee houden.

Hetzelfde met de specifieke klokinstellingen.

Als we het over sequentiële schakelingen hebben, gaan we in eerste instantie toestandsdiagrammen gebruiken. Je kan zelf een toestandsdiagram tekenen en die software weet hoe je dat toestandsdiagram efficiënt moet omzetten naar hardware. Alle tussenliggende stappen moet je dus niet doen. Je kan ook een VHDL-beschrijving doen. Soms is dit veel compacter.

Je kan dit allemaal gaan combineren in uw schema.

**Slide 60:** Je kan in die omgeving een logische simulatie doen. Dan krijg je daar de gegeven tekening uit: je kan ingangssignalen aanleggen, zien waar er een reset gebeurt,... Het telt hier op een decimale manier.

Je kan ook elk beetje afzonderlijk bekijken hoe het zich in de tijd gedraagt maar dit is geen tijdsgedrag, alleen een benadering want het toont alleen de logische werking. Er staat geen timinginformatie. Zelfs als hier tijden bij vermeld staan, is het geen echte timinginformatie.

**Slide 61:** Als je daarmee tevreden bent, moet je uw fysisch ontwerp gaan doen. Je moet gaan zeggen hoe je alles in de logische cellen steekt. Dat heeft dus niks met poorten te maken. Wat als je heel complexe berekeningen hebt op 27 bits, wat ga je in welke cel steken, hoe ga je ze verbinden, waar ga je welke cel zetten,...? → Moet allemaal bepaald worden.

**Slide 62:** Dat gebeurt automatisch. Er zijn verschillende stappen die moeten doorlopen worden: alle schema's moeten samengezet worden. Dan moet dat vertaald worden naar logische primitieven, men moet er een logisch ontwerp van maken: poorten en flipflops. Maar er staan geen poorten op uw FPGA. Dan komen de specifieke dingen: wat kan je het best gebruiken, hoe ze je die logische schakeling het best om?

2 stappen: logische schakeling maken en dan die zo goed mogelijk vertalen naar wat er op die FPGA komt. Dan heb je bepaald wat op die cellen moet komen en hoe ze verbonden zijn. Maar dan heb je hetzelfde probleem: welke cel zet je op welke plaats?

Hier heb je eveneens een placement en routing stap. Als dat allemaal gedaan is, is uw fysisch ontwerp af, je hebt dan alle informatie: je weet wat welke cel doet, hoe die met andere cellen verbonden is en als fabrikant ken je de specificaties van die logische verbindingen en kan je het tijdsgedrag gaan voorspellen.

**Slide 63:** Simulatie doen die eruit ziet als getoond, hier is de tijdsinformatie wel de echte tijdsinformatie. De piekjes die je ziet, waren er daarnet niet. Transitie slide: de twee enen moeten nul worden en de tweede nul moet 1 worden. Het tijdsverloop toont wanneer welke bit verandert: er is altijd maar 1 signaal dat tegelijkertijd verandert. Alle piekjes die je zag op het vorige beeld (voor de transitie) zijn ook niet zo erg, de tijdssimulatie zegt u alleen wat zal gebeuren, jij moet dan beslissen of dat OK is of niet.

**Slide 65:** FPGA: omcirkeld (en helemaal zichtbaar). De rest zijn allemaal dingen om te experimenteren, drukknoppen,...

## 4.2 Slides: 4\_Combinatorisch

**Slide 1:** We gaan beginnen ontwerpen met de eenvoudigste schakelingen: combinatorische schakelingen. De ingang verandert en zal impact hebben op de uitgang, maar houdt geen rekening met het verleden: heeft dus geen geheugen.

**Slide 2:** Alle dingen die op schema's zijn uitgelgd, hoe ga je dat in een hardwaretaal beschrijven? We beginnen met de combinatorische schakelingen, we willen die minimaliseren.

**Slide 3:** De eerste vraag die je je kan stellen is: waarom zou je dat willen? Het is omwille van de randvoorwaarden dat je wil minimaliseren. We streven naar een schakeling die zo goedkoop mogelijk is, zo snel mogelijk werkt en zo weinig mogelijk vermogen verbruikt. Het derde is niet vermeld op de slide, dat is evenredig met de capaciteit maal de frequentie maal de spanning in het kwadraat.

Hoe gaan we zorgen dat we zo weinig mogelijk vermogen verbruiken? Als we de frequentie en de spanning constant houden, hangt het dus af van de capaciteit. Die hangt af van de grootte van de component. Als we dus zo compact mogelijk werken, zullen we ook een minimum aan vermogen verbruiken.

Om 2 implementaties te kunnen vergelijken moeten we een maat hebben om de kostprijs en snelheid van een schakeling te kunnen bepalen.

De kostprijs is het makkelijkste: we hebben gezien dat de kostprijs van een poort evenredig is met het aantal ingangen bij CMOS. Het hangt er wel vanaf welke technologie gebruikt wordt. Als het een niet-inverterende poort is, die kost nog eentje extra (alleen voor CMOS). Gaan we naar iets totaal anders kijken zoals de FPGA, daar telt het aantal transistoren niet. Daar hangt de kostprijs er vanaf of je een kleine of een grote FPGA kan gebruiken (met kleine logische cellen of

grote). Het is hier evenredig met het aantal logische cellen. De kostprijs wordt hier dus anders bepaald. Dit is moeilijk omdat het moeilijk is om te schatten hoeveel logische cellen je gaat gebruiken.

In de praktijk moet je je daar niet teveel zorgen om maken. Het eerste (CMOS) is dus meer als je zelf een geïntegreerde schakeling gaat maken.

De kostprijs van de schakeling is de som van de kostprijzen van de afzonderlijke elementen.

**Slide 4:** Vertraging: moeilijker. Vertraging is niet alleen een functie van de complexiteit van de poort, maar ook van de stijg- en daaltijd. Die hebben te maken met hoe die poorten verbonden zijn met elkaar. Dat kan je nooit op een logisch schema zien, dat kan je alleen berekenen indien je een fysisch ontwerp hebt. Om dat toch enigzins te kunnen vergelijken, is er een benadering. Er is een stuk dat evenredig is met het aantal ingangen van de poort en dat is de complexiteit van de poort die je in rekening brengt (hoe meer transistoren, hoe trager het zal zijn). Er staat ook nog een constante bij (0.6 en 1.6) om rekening te houden met de verbindingen. Het probeert beide dingen dus te combineren. Dit geldt weer alleen voor CMOS, niet voor FPGA, daar is het te complex.

De totale vertraging is de som van de vertragingen langs het kritisch pad: dat pad van een ingang naar een uitgang dat voor de grootste vertraging zorgt. Als je dat pad volgt, dat is de traagste manier om van een ingang naar een uitgang te gaan. Het is het traagste pad dat de maximale verwerking bepaalt.

Bv. voor die XOR: donkerblauw is 1 poort (geen 3). Van die één enkele poort kan je de kostprijs en de vertraging berekenen. De poort heeft 4 ingangen en is inverterend, dus  $0.6 + 4 * 0.4$ . De vertraging van de ganse schakeling is die door ook nog de inverter, er zijn hier 2 kritische paden: door de inverters en ze hebben dezelfde vertraging (anders hadden we geen 2 kritische paden), dus nog +1.

**Slide 6:** Karnaughkaarten: wat we tot hertoe gezien hebben is dat je dat kan oplossen door algebraïsch te werken met Boole-algebra. Maar dat is niet echt handig omdat je niet weet welke theorema's je in welke volgorde moet gebruiken om tot het minimum te komen. De waarheidstabel toont het al meteen in sommige gevallen. In het voorbeeld zie je dat de uitgang 1 is in de 5<sup>e</sup> en 6<sup>e</sup> rij. Om die twee gevallen te combineren kan je gewoon zeggen dat  $f=1$  als  $x=1$  en  $y=0$ ,  $z$  doet er niet toe. Je kan dat natuurlijk niet zomaar in alle gevallen zien. Je kan evengoed de groene illustratie zien:  $f=1$  als  $z=1$ . maar dat is minder makkelijk om te zien omdat het niet meer naast elkaar ligt.

**Slide 7:** N-kubus met zoveel dimensies als er variabelen zijn. Als we maar 2 dimensies hebben is het doenbaar, ook bij 3, maar daarna is het niet meer te zien. Visueel ben je dus niets met een N-dimensionale kubus, maar het principe is niet slecht.

**Slide 8:** Oplossing: N-kubus die je platduwt: je maakt die 2-dimensionaal. Op de slide heb je een 3-D kubus die is platgeduwd. 000 ligt naast 001, dat moet in het 2-D geval nog altijd zo zijn! We kunnen hier heel makkelijk tot 4 variabelen weergeven. Als je 2 dimensies hebt en je wil dat uitbreiden, ga je spiegelen rond een bepaalde as.

**Slide 9:** Spiegelen: de spiegeling van 1 is 3.

De mens is heel goed in het herkennen van patronen, maar je moet die wel kunnen herkennen. Je hebt de binaire voorstelling en gaat telkens 1 bit veranderen. Als je 5 bits hebt, heb je 5 mogelijkheden: 5 groene veldjes. 5 variabelen gaat minder makkelijk: wat ligt er allemaal naast 5: alle rode. Dat 21 ernaast ligt is minder evident om dat te herkennen. Die ligt niet alleen in het andere vierkant, maar ook gespiegeld in het andere vierkant. Dat maakt het soms heel moeilijk om het nog visueel te interpreteren.

In plaats van te spiegelen kan je ook gewoon dupliceren. In dat geval moet je wel gewoon op dezelfde plaats kijken en dat is al wat makkelijker.

**Slide 10:** Werkt goed voor een beperkt aantal variabelen, maar eens je meer variabelen hebt dan 4, wordt het moeilijk om het visueel te kunnen oplossen.

Als je een karnaughkaart met 6 variabelen moet oplossen op het examen, kun je met 99.999% zekerheid zeggen dat je een fout hebt gemaakt.

**Slide 11:** Hoe gaan we dat gebruiken? We hebben dat daarnet gemaakt om visueel te zien wat naast mekaar ligt. Als je van een logische functie vetrekt, kan je die gebruiken om uw kaart op te vullen. In zo'n kaart ga je zo'n groot mogelijke gebieden maken die een macht zijn van 2. Hoe groter dat gebied is, hoe meer variabelen je in je poort kan laten vallen. Als je dat op het voorbeeld bekijkt: gebied van 4. Hoe goed je ook zoekt, je vindt geen gebied van 8. Dan zoek je nog naar de anderen: onderaan kan je samennemen en ook verticaal. Er mogen gebieden overlappen! Zo heb je een vorm gevonden die veel compacter is dan uw oorspronkelijke beschrijving.

Karnaughkaarten zijn gebaseerd op het herkennen van patronen. De vraag is of er nog andere patronen te herkennen zijn die nuttig zijn voor een implementatie: dambordpatroon. Dat komt overeen met een XOR.



# Chapter 5

## Les 5

### 5.1 Slides: 4\_Combinatorisch

**Slide 12:** We hadden het over combinatorische schakelingen waarbij de uitgang afhankelijk is van de ingangen, er is geen voorgeschiedenis.

We hebben Karnaughkaarten bekeken, we gaan dat nu iets verder in detail bekijken, want voorlopig hebben we altijd een aantal ingangen en enkele uitgangen gehad, maar in realiteit heb je soms ongespecificeerde uitgangen.

Er zijn altijd onmiddellijk 2 oplossingen: SOP en POS, dus er zijn automatisch 2 mogelijke implementaties. Het minimaliseren ging erom dat we een zo goedkoop mogelijke oplossing, een zo snel mogelijke oplossing en een oplossing die zo weinig mogelijk vermogen verbruikt willen. We willen de implementatie doen met een minimum aan transistoren.

**Slide 13:** Implicanten: producttermen waarvoor de uitgang 1 is. De 1-mintermen behoren daartoe, zoals je op de slide ziet. Dat zijn niet de enige, je kan ook producttermen hebben waarin minder variabelen zitten, zoals op de rode rij. Dat zijn allemaal implicanten die gebruikt kunnen worden om die functie te realiseren wanneer we op zoek gaan om de ene te realiseren.

Priemimplicanten: speciale soort: dat is de compactste manier om iets voor te stellen, er is geen compactere manier en alle manieren die eronder vallen zijn duurder. Het is een implicant die geen onderdeel is van een grotere. De blauwe is een priemimplicant want er is er geen die groter is. We hebben in het voorbeeld maar 2 priemimplicanten.

Wanneer we gaan implementeren is het het interessantst om de priemimplicanten te gebruiken, want als het geen priemimplicant is, is er altijd een efficiëntere manier om het uit te werken.

Dekking: welke implicanten hebben we allemaal nodig om de functie volledig te realiseren? Dat kunnen er verschillende zijn. De oplossing die gebruik maakt van de priemimplicanten is veel efficiënter.

**Slide 14:** Wij gaan vertrekken van een Karnaughkaart, in andere gevallen/oplossingen kan je van iets anders vertrekken (bv. tabellen). Wat daarachter komt is essentieel hetzelfde. We hebben al gezien dat het de priemimplicanten zijn die het interessantst zijn om te gebruiken, dus we gaan die eerst bepalen. Er is

een speciaal soort priemimplicanten: de essentiële priemimplicanten. Dit zijn degenen die je echt nodig hebt om het te implementeren. De essentiële heb je echt nodig en je weet dat je die nodig hebt door het feit dat er minstens 1 1-minterm inzit die in geen enkele andere priemimplicant voorkomt, er is geen enkele andere beschikbaar waarmee je het kan doen, dus die is essentiël. Dan gaan we zien of we nog iets anders nodig hebben, we vullen dat dan aan (liefst met priemimplicanten) tot alle enen bedekt zijn.

**Slide 15:** De eerste stap is het maken van de Karnaughkaart. Je moet enen invullen waar ieder van die termen 1 is. De groene vakjes zijn die waar de stelling waar is. Zo doe je dat term per term. Die enen mogen overlappen, ze mogen alleen niet contradicteren.

**Slide 16:** We hebben nu de Karnaughkaart gemaakt en daarmee gaan we aan de slag. We gaan nu kijken naar een stap-voor-stap oplossing die altijd werkt. Er is ook een snellere manier, maar die is ook gevaarlijker (je ziet makkelijker dingen over het hoofd).

Hoe komen we tot alle priemimplicanten? We gaan die enen een voor een aflopen. Voor iedere 1 gaan we de priemimplicanten zoeken waar die 1 deel van uitmaakt. Als we naar die eerste kijken, hebben we 2 1-mintermen die eraan voldoen. We kunnen er geen van 4 zoeken waarin de linksbovenste 1 inzit. We kunnen geen van 4 vinden, de grootste is dus van 2. De bedoeling is dat we alle priemimplicanten met een grootte van 2 opsommen, in dit geval zijn dat er dus 2. Dat resulteert dus onmiddellijk 2 priemimplicanten. Bij de volgende 1 kunnen we 4 enen samennemen en weerom zijn er 2 mogelijkheden, velden van 4, waar die 1 in voorkomt, dus er komen weer twee priemimplicanten bij. Bij de volgende 1 komt er geen priemimplicant bij (we hadden die al), zo kan je verdergaan en priemimplicanten identificeren. Uiteindelijk heb je de volledige lijst van priemimplicanten.

**Slide 17:** Nu je die priemimplicanten hebt moet je de essentiële gaan zoeken. De essentiële zijn degenen die een 1 bedekken waar geen enkele andere priemimplicant overkomt. Je kan dat zien op de Karnaughkaart. Op veel plaatsen zijn verschillende gebieden die overlappen, bij 2 enen is er maar 1 term die de 1 bedekt. Dat zijn dus termen die priemimplicanten zijn. Door het feit dat je die gekozen hebt, heb je al een zekere dekking. De rechtsbovenimplicant zal ook andere enen bedekken, we hebben zo al 8 enen bedekt. Zo moet je gaan kijken of er nog enen zijn die niet bedekt zijn (zijn er 2) en van die resterende implicanten een liefst zo klein mogelijke priemimplicant zoeken zodanig dat die resterende enen ook bedekt worden. Tot hiertoe kon je met zekerheid zeggen wat erin moest zitten, maar nu wordt het iets moeilijker want je moet het globale overzicht hebben en kijken welke priemimplicanten je gaat gebruiken. In de praktijk is dat vaak moeilijk en gaat men iteratief te werk en gebruikt men een gulzige manier van werken: eerst de priemimplicant zoeken die zoveel mogelijk bedekt, dan de tweede die dat doet, . . . je kan in een lokaal optimum terechtkomen, maar het kan zijn dat het niet de meest minimale oplossing is. In praktijk ga je er meestal wel terechtkomen. We gaan van iedere priemimplicant zoeken hoeveel die bedekt en dan gaan we die nemen die de meeste enen bedekt. Uiteindelijk

kom je op de oplossing getoond op de slide, dat is een minimale standaardoplossing.

**Slide 20:** Je kan ook kijken naar de nullen in plaats van naar de enen. Je krijgt dan de kaart zoals getoond. Ook hier ga je op dezelfde manier tewerk: zo groot mogelijke gebieden zoeken die een nul realiseren. Je zoekt de essentiële, die minimale gebruiken en dan zoeken wat je nog nodig hebt. Op zich is dat niet zo erg verschillend, maar er is een groot gevaar: het is hier iets minder makkelijk om die term te omschrijven die die nullen bedekt. Als je concreet kijkt naar de rood omcirkelde, hoe beschrijf je dat? Moesten daar enen staan, zou je zeggen  $zw'y'$ . Maar er staan geen enen, er staan nullen, dus je moet er het complement van nemen, de beschrijving die dat beschrijft is  $(zw'y')$ . Met de wet van De Morgan kan je dat omzetten naar  $w+y+z'$  want nu ben je niet op zoek naar een SOP maar naar een POS. je kan dat via de omweg doen zoals net uitgelegd, je kan ook rechtstreeks naar de kaart kijken en bv. zeggen: er staan 2 nullen, dus op alle andere plaatsen staan enen. Hoe kan je die gaan beschrijven? Alle enen onderaan is  $w$ , alle enen rechts is  $y$  en alle zijkanten is  $z'$  en dat is meteen een OR-verband.

Hier zie je die 5 priemimplicanten. Dan moet je de essentiële gaan zoeken en weerom, over twee nullen komt er maar 1 priemimplicant. Op dit moment staat er is zo'n vakje een 1 of een 0, die zijn complementair ten opzichte van elkaar.

**Slide 21:** We gaan gebruik maken van volledig gespecificeerde functies. Dat betekent dat we niet voor iedere combinatie van ingangen een zinvolle uitgang hebben.

**Slide 22:** Soms kan het ons niet schelen wat er op een uitgang komt voor die ingangscombinatie. Een reden daarvoor is dat het niet kan voorkomen dat die combinatie aan de ingang gaat staan. Dan zegt men dat die uitgang "don't care" is. We hebben nu al 4 mogelijkheden om een signaal aan te geven: 0, 1, Z & don't care. De mogelijkheden om een binair signaal te beschrijven nemen alsmaar toe. Een voorbeeld waarbij je dat makkelijk ziet waarom dit optreedt is een 7-segmentdisplay van een BCD-getal (binary-coded decimal). Je kan dat op zo'n display weergeven. Je hebt 4 ingangen en 7 uitgangen. Als je een 0 wilt tonen, moeten alle uitgangen branden buiten het middelste stuk, ... De tabel is nog niet helemaal opgevuld eens we 9 hebben voorgesteld. Je kan zeggen dat het u niet kan schelen wat eruitkomt. We geven dat aan met een streepje. Er zijn verschillende manieren om dat te doen (in het boek: d, vaak ook X). Wanneer we dat ontwerp maken, verdwijnt die don't care en wordt die omgezet naar een 0 of een 1.

**Slide 23:** Hoe maak je daar gebruik van? Wat is het voordeel van een don't care vs. allemaal nullen? De don't care heeft het voordeel dat je die zowel voor een 1 als een 0 kan gebruiken zoals het u uitkomt. Je kan die gebruiken om gebieden groter te maken als dat interessant is. Je hebt bv. de kaart op de slide (links), als je die kaart hebt en je gaat daar een SOP van maken, als je daar zo groot mogelijke gebieden zoekt, kan je don't cares gebruiken om die te zien als een 1 en kan je grotere gebieden maken, zo kan je een ingang minder gebruiken aan uw poort. Zo'n don't care geeft u dus de mogelijkheid om uw

implementatie efficiënter te maken: je kan kiezen wat u het beste uitkomt. Als je gaat kijken, wat is er nu met die don't care gebeurd eens je dat gemaakt hebt? Als je de blauwe gebieden als priemimplicanten gebruikt voor de implementatie, de onderste don't care is dan een 1 geworden want je hebt de onderste rij samengenomen (anders mocht je dat niet samennemen). Evengoed zijn de twee bovenste don't cares enen geworden. De tweede onderste is een nul geworden al je heb die niet gebruikt om er een 1 van te maken. Na de implementatie bestaan er dus geen don't cares meer, je gebruikt dat enkel tijdens het ontwerp, maar eens je de implementatie gemaakt hebt, bestaat dat niet meer. Hier is het nu ook zo dat als je de nullen zou implementeren, dat dat niet automatisch een complementaire implementatie is: rechtse tekening: je neemt de twee bovenste erbij om een groter gebied van nullen te maken en de tweede onderste gebruik je ook als nul, de onderste is een 1 geworden. Het is dus niet zomaar het complement van de linkse tekening! Het is natuurlijk niet zo dat dat de ene keer 0 en de andere keer 1 is, het hangt er vanaf, van uw implementatie. Indien het mogelijk is, ga je zoveel mogelijk don't cares in uw Karnaughkaarten proberen in te schrijven.

**Slide 24:** Een normale schakeling heeft meer dan 1 uitgang. Op het eerste gezicht lijkt dat niet zo erg want 1 schakeling met 4 ingangen en 7 uitgangen zou je kunnen bekijken als 7 schakelingen met ieder 4 ingangen en 1 uitgang. Je kan dat zo implementeren, maar dat geeft meestal geen efficiënte implementatie.

**Slide 25:** Je kan poorten die je in de ene implementatie gebruikt ook voor de andere gebruiken. Het voordeel is dat die maar 1 keer kosten. Zoveel mogelijk poorten hergebruiken is waar we naar op zoek zijn, daarom mag je dat niet onafhankelijk van elkaar maken. Je moet alle kaarten bijeen bekijken en dan als 1 probleem oplossen. Stel dat we enkel priemimplicanten gebruiken. Er is een speciaal soort priemimplicanten, die essentiële heb je zeker nodig. Wat je wel kan doen is die 7 kaarten opstellen en in ieder van die kaarten naar de essentiële zoeken. Je weet dat je die nodig hebt om die geïmplementeerd te krijgen. Je kan geluk hebben dat er priemimplicanten overeenkomen. Je gaat in iedere kaart de essentiële priemimplicanten zoeken en die al als basisoplossing gebruiken. In het tweede deel, waarin je de resterende enen gaat proberen oplossen, mag je het niet onafhankelijk van elkaar bekijken want ondertussen heb je al poorten ter beschikking en het zou het interessantst zijn om poorten te hergebruiken. Wat in het rood aangegeven is, is in iedere kaart de essentiële priemimplicant. In de eerste twee kom je toe met de ene essentiële die er is, maar bij de derde zijn er nog 2 enen die niet bedekt zijn door de essentiële. De vraag is nu hoe je die gaat realiseren. Nu gaan we ook naar de andere kaarten kijken, want het zou kunnen dat die al in een andere kaart geïmplementeerd zijn. Dat is al zo: die twee blauwe zijn geen essentiële voor c, maar degene die de middelste implementeert is een essentiële in a en b en de blauwe die de rechtse kolom samenneemt is een essentiële in a. in dit geval bestaan ze beiden al. Als er al meerdere geïmplementeerd zijn, hoe kies je dan? Qua kostprijs maakt dat op zich niets uit. Je kan kijken naar andere randvoorwaarden zoals bv. een zo snel mogelijke schakeling. Die hangt af (bij CMOS) van de fan-out. Dus, wat ga je hier doen? De schakeling zal zo snel mogelijk zijn als je de fan-out van alle poorten zo laag mogelijk kan houden. Dat betekent dat ze allemaal

zo weinig mogelijk gebruikt moeten worden. Je gaat ze gebruiken omwille van de kostprijs, maar je gaat een poort die al veel gebruikt is niet hergebruiken als er een is die weinig gebruikt is. Die van b is al 2 keer gebruikt, die van a (de rechtse kolom) is nog maar 1 keer gebruikt, dus we gaan die gebruiken. Belangrijk is dus dat je dat niet onafhankelijk van elkaar mag bekijken. De uiteindelijke implementatie van de drie uitgangen staat onderaan. Je ziet dat er dingen hergebruikt worden.

**Slide 26:** Wat het moeilijker maakt is wanneer je meerdere kaarten hebt die je samen moet bekijken. Het kan hierbij gebeuren dat, als je meerdere uitgangen hebt, het interessant is om niet-essentiële priemimplicanten te gebruiken, of zelfs implicanten te gebruiken die geen priemimplicant zijn. Normaal is dat geen goed idee, maar dat kan een goed idee zijn als je er meerdere hebt. We hebben hier twee kaarten/uitgangen voor 4 ingangen. We gaan dat implementeren op de manier juist uitgelegd en dan krijg je kostprijs 32. Dit is evenwel niet de meest compacte implementatie. Je kan gebruik maken van niet-essentiële priemimplicanten, maar alleen als ze al ergens anders voor gebruikt werden, alleen als je daardoor hergebruik kan stimuleren (alleen als je meerdere uitgangen hebt, nooit bij 1 uitgang). Bij hergebruik stijgt de kostprijs niet: het maakt dan niet uit als je 1 poort hebt die duurder is maar wel meer gebruikt kan worden. Vanonder is men niet op zoek gegaan naar essentiële priemimplicanten, maar is men gaan kijken hoe zoveel mogelijk hergebruikt kon worden. Wat in het groen weergegeven is en de 1 onderaan kan je 2 keer gebruiken. Je kan zeggen dat de 1 bij a niet nuttig is omdat je die kan samennemen met de enen erboven, maar dat zou enkel zijn wanneer je maar 1 kaart hebt, nu gaat dat omdat je meerdere kaarten hebt: het is toch gratis om ze een tweede keer te gebruiken. Bij de twee enen in de derde kolom is geen enkele kaart een priemimplicant, maar je kan het hergebruiken dus uw kostprijs daalt. Het probleem hierbij is dat er geen mooi stappenplan is om dat te vinden. Je moet die kaarten naast elkaar leggen en ernaar beginnen staren en hopen dat er dingen naar boven komen, je gaat misschien dingen moeten proberen. Het is dus trial-and-error. Het probleem is daarmee nog niet helemaal opgelost: alles wat we tot nu toe gezien hebben geldt alleen als je tot een tweelagenimplementatie komt. Daar ben je dikwijls naar opzoek want het is de snelste implementatie, maar in vele gevallen kan het interessant zijn om naar meerdere lagen te zoeken, maar daar heb je ook niet echt mogelijkheden.

**Slide 27:** Dit soort technieken is in alle CAD-omgevingen geïmplementeerd, maar niet met Karnaughkaarten (is een visuele manier van werken). Er is dus een tegenhanger van: Quine-McCluskey. Die maakt gebruik van tabellen. Is niet echt interessant om dat met potlood en papier te doen, zo kan die tot een goede oplossing komen. De technieken die daarachter zitten zoals essentiële priemimplicanten zoeken,...komen daar ook aan bod. We gaan dat niet nog eens zien omdat het niet nuttig is in dit geval. Het staat in het boek in detail beschreven als het u interesseert. Je moet gewoon weten dat het een tegenhanger is van Karnaugh die hetzelfde doet.

**Slide 30:** Het kan zijn dat nadat je je specificaties geanalyseerd hebt, je dat als logische bewerking op de manier na de slide beschrijft. We kunnen geen

poorten maken met een fan-in van 100, dat is niet mogelijk om technologische redenen. Als je dus met die technieken van Karnaughkaarten uitkomt op een poort met 100 ingangen, dan kan je die niet maken. In die technieken van daarnet kan je dat niet inbrengen dat er een beperkte fan-in is. Hoe doe je dat dan? Je lost het probleem in stukjes op: eerst hou je er geen rekening mee, dan ga je kijken hoe je uw poorten kunt realiseren met minder ingangen dan je bent uitgekomen. Het komt erop neer dat je haakjes gaat introduceren: de onderste formuleset is eerst een OR-poort met 4 ingangen en dan reduceer je die naar 3 OR-poorten met 4 ingangen. Je kan uw poorten omzetten naar meerdere lagen van OR-poorten. Je kan de haakjes ook anders zetten en dan krijg je ook poorten met 4 ingangen.

Op zichzelf is die laatste lijn niet slechter dan de voorlaatste (zelfde kostprijs), maar qua vertraging is dat niet hetzelfde: in de onderste heb je 3 lagen en die gaat trager zijn. Als we dan toch kunnen kiezen, kiezen we natuurlijk de snelste oplossing. Je gaat altijd een boomstructuur toepassen, dat is de snelste manier van werken. Dat is hier geïllustreerd voor OR-poorten, dat geldt evengoed voor AND-poorten en XOR-poorten. Zo kan je tegemoet komen aan de beperkte fan-in die we soms hebben.

**Slide 31:** Functionele ontbinding: de functie ontbinden in verschillende delen die je ieder afzonderlijk gaat implementeren. Bij voorkeur ga je die delen zo kiezen dat de uitgang van het ene deel op meerdere plaatsen kan gebruikt worden, dan heb je iets nuttig apart geïmplementeerd. Je gaat nadenken van wat erin zit, waaruit bestaat het probleem dat je wil oplossen en wat zijn de subproblemen? Elk van die subproblemen ga je apart implementeren. Dat is typisch wat je doet als je hiërarchisch gaat oplossen. Ieder van de subschema's ga je laten invullen. Zelfs als je de subschema's optimaal maakt, gaat het totaal geen tweelagenlogica zijn.

Heel dikwijls, zonder het zelf te beseffen, kom je tot meerlagenlogica. We gaan dat straks toepassen. Je komt dan tot een andere (goedkopere) implementatie en gedeeltelijk snellere implementatie. Het probleem hier weer is dat het ook hier niet evident is wat je juist moet gebruiken, wat een goed resultaat zal opleveren. Er is dus heel veel trial-and-error en ervaring begint hier mee te spelen.

**Slide 32:** Alles wat we tot nu toe gezien hebben gaat er vanuit dat we alleen beschikking hebben over basispoorten in een CMOS-technologie. Maar dat is niet het geval, we kunnen ook AND-OR-invert gebruiken. Als je een FPGA gebruikt, doet het er eigenlijk niet toe wat voor poorten je hebt want 1 blokje in een FPGA kan elk van die functies gaan uitvoeren. Hoe los je het probleem nu op? Door technology mapping.

**Slide 33:** Je kapt uw probleem in zoveel kleine stukjes die je wel kan oplossen en zet die allemaal samen. We gaan dat hier ook zo doen. Zowel voor FPGA als voor gatearrays. We gaan in 2 stappen werken:

1. We hebben een beschrijving, we gaan die proberen zo compact mogelijk voor te stellen. We gaan doen alsof we het met basispoorten gaan implementeren, dan hebben we de meest compacte beschrijving van de functionaliteit die we moeten implementeren.

2. Een keer we dat hebben gaan we dat implementeren, rekening houdend met de beperkingen die er zijn.

In praktijk worden we daar meestal bij geholpen want stap 2 is zeer sterk afhankelijk van de implementatie: komt het in een FPGA, in een gate-array, ... Degene bij wie we dat gaan implementeren weet dat we met dat probleem zitten en zal daar gewoonlijk software voor leveren om die stap te doen want het is te specifiek voor de implementatie.

**Slide 34:** (EXAMEN: je mag alleen NAND en/of NOR poorten gebruiken met 2 ingangen, niet met meer dan 2 ingangen, ook geen AND en OR). Hoe ga je dan te werk? Je gaat eerst de Karnaugh-technieken toepassen en dan kom je bv. tot de tekening links. Je hebt dan de compactste oplossing gezocht, die moet je gaan omzetten naar een implementatie die met die beperking rekening houdt. Het eerste wat je dan doet is die poorten opsplitsen. Daarmee heb je de eerste beperking opgelost: geen enkele poort die meer dan 2 ingangen heeft. Dan moet je dat omzetten naar NAND en NOR. Op die manier heb je een implementatie die aan alle beperkingen voldoet.

**Slide 35:** We hebben nu gezien dat er heel veel verschillende implementaties zijn, welke is nu de goeie? We kunnen alleen zoveel mogelijk implementaties bedenken en er dan de beste uitkiezen.

**Slide 36:** Je kan de canonieke oplossing gebruiken. De canonieke POS is duurder want je hebt meer nullen dan enen.

**Slide 37:** Karnaugh toepassen om de minimale SOP te zoeken. Vertraging: kritisch pad, alle poorten zijn hetzelfde en soms moet je door een inverter.

**Slide 40:** Met meer lagen werken, complexe poorten gebruiken, ... wat in het blauw gekleurd is is 1 poort: OR-AND-invertpoort. Dat heeft dus 1 kostprijs van 6 en een vertraging van 3.

**Slide 41:** In de praktijk is het niet mogelijk om alles te gaan bekijken, er wordt dan meestal gevraagd wat volgens u de beste oplossing zou zijn en te vermelden waarom. Alles nog eens vergeleken: illustreren dat vertraging en kostprijs niet altijd hand in hand gaan. De goedkoopste oplossing is hier niet de snelste oplossing. Het is niet automatisch zo dat de meest compacte oplossing ook de snelste is. Wanneer je het dan moet implementeren moet je weten of snelheid of kostprijs het belangrijkste is voor jou. Afhankelijk daarvan zal je voor 1 van de 2 oplossingen kiezen. Die tabel is geen algemene tabel! Die is anders voor elk probleem dat je moet oplossen, het dient gewoon om te illustreren voor 1 specifiek probleem.

**Slide 42:** We weten hoe we combinatorische schakelingen maken en wat de mogelijkheden zijn. Nu moeten we dat op een hoger niveau gaan toepassen. Bij de combinatorische schakelingen heb je niet zoveel mogelijkheden: je hebt een aantal bouwblokken die nodig zijn om berekeningen te doen en een beperkt aantal basisschakelingen die bepaalde dingen doen (bv. kiezen uit 10 mogelijkheden). We gaan eerst kijken naar de rekenkundige basisschakelingen.

**Slide 43:** Eerst moeten we overeenkomen hoe we getallen gaan voorstellen. Heel formeel beschreven wat die cijfers te doen hebben met het getal. De plaats van het cijfers bepaalt de radix (met welke macht van 10 je werkt).

**Slide 44:** Wij rekenen normaal niet met het decimale stelsel. Het idee blijft wel hetzelfde. We gaan heel dikwijls niet binair werken omdat het misnte grote getal dat we willen voorstellen een enorme opeenvolging van enen en nullen is, dat kunnen wij niet makkelijk onthouden. Men is op zoek gegaan om een aantal bits samen te nemen en dat als 1 cijfer voor te stellen. De grootste macht van 2 die je voor kan stellen met 1 cijfer is een  $8 \rightarrow$  octaal stelsel. Dit zijn 3 bits, is niet handig om neer te schrijven. Hexadeximaal is veel makkelijker. Daarmee kan je met een beperkt aantal cijfers redelijk grote getallen weergeven.

**Slide 45:** Het is ook niet moeilijk om van het ene naar het andere over te gaan, zolang ze eenzelfde basis hebben.

**Slide 46:** We gaan daar nu mee rekenen.

**Slide 47:** Herhaling van het eerste leerjaar.

**Slide 48:** Halve opteller: kolom  $s =$  “uitkomst” voor die rij/kolom,  $c$  is de carry-over.  $c_{i+1}$  is de AND van  $x_i$  &  $y_i$ ,  $s_i$  is het dambordpatroon. Het is een halve opteller omdat het maar 2 bits kan optellen. Als je overdracht hebt, moet je 3 bits optellen, dus dat hebben we ook nog nodig.

**Slide 49:** De Karnaughkaart wordt hier iets moeilijker.

**Slide 50:** We moeten naar kostprijzen gaan kijken. Als je meerdere cijfers moet optellen wordt de snelheid niet bepaald door van de ingang naar de carry te gaan, maar door van de inkomende carry naar de uitgaande carry te gaan. Van  $c_i$  naar  $c_{i+1}$  is de vertraging van een AND naar een OR. Dit is beter want op de vorige slide was het een AND en een OR-poort met 3 ingangen.

**Slide 51:** Waarom ben je geïnteresseerd in de inkomende carry naar de uitgaande carry? Als je die cijfers naast elkaar zet krijg je wat te zien is op de slide. Waar wordt hier nu de snelheid door bepaald? Door de langste weg die je hebt: als de carry verandert kan de volgende carry ook veranderen, . . . Je moet dus van helemaal rechts naar helemaal links geraken en da's het slechtste geval. Vandaar ripple-carry: rimpelt door van helemaal rechts naar helemaal links. Hoe meer carries je hebt, hoe langer het duurt. Als je dus een 4-bits opteller wilt hebben, gaat het een poosje duren voor het meest beduidende resultaat zichtbaar gaat zijn. De vertraging van die opteller is dus evenredig met het aantal bits. We gaan nu 1 grote opteller maken van 8 bits met 17 ingangen en 19 uitgangen.



# Chapter 6

## Les 6

### 6.1 Slides: 4\_Combinatorisch

**Slide 50:** We hebben gekeken hoe we bouwblokken van combinatorische schakelingen kunnen maken. De eerste die we aan het bekijken waren was een opteller. Uiteindelijk waren we ertoe gekomen dat we dat kunnen opdelen in onderdeeljes: 1 opteller per cijfer en dan kregen we een dergelijke implementatie en die kunnen we combineren naar meerdere cijfers. En dan krijg je wat op **Slide 51** staat.

**Slide 51:** Die werken niet onafhankelijk van elkaar. Via de carry gebeurt de overdracht van de ene naar de andere. Dit zorgt voor vertragingen want wanneer er iets gebeurt moet je wachten tot het doorgerimpeld is voor je zeker kan zijn dat het juiste cijfer bekomen is. Een 32-bits opteller is 2 keer zo traag als een 16-bits opteller en we willen dat nu sneller maken. Dat kan door dat niet mooi gestructureerd zoals hier te implementeren, maar door dat als een grote schakeling te zien, bv. met 9 ingangen en 5 uitgangen. Het nadeel is dat we niet meer iets kunnen hergebruiken. De vraag is dan ook of er een tussenoplossing is waarbij we wat kunnen hergebruiken maar dat sneller werkt dan de ripple-carry opteller. Er zijn verschillende mogelijkheden.

**Slide 52:** Meest courante: carry-lookahead opteller. Wat het meest vertragende was (carry doorgeven) gaan we proberen versnellen. Je kan de dia van **Slide 50** hertekenen en dan krijg je wat op deze slide staat. Als je die signalen ( $p_i$  en  $g_i$ ) hebt, kan je de werking herschrijven zoals getoond op de slide.  $p$  en  $g$  zijn de signalen voor propagate en generate, met de betekenis op een carry: ik ga in dit cijfer een carry genereren. Wanneer genereer je zeker een overdracht? Wanneer je 1+1 doet, onafhankelijk of er een gepropageerd gaat worden of niet. Wanneer ga je er een voortplanten? Als er 1 aan de ingang is, dan ook 1 aan de uitgang. 1 van de ingangsbits moet 1 zijn (staat in  $p_i$ ). Je kan dat vereenvoudigen, maar nu kunnen we hetzelfde schema behouden. Je kan het verband tussen de uitgaande en binnenkomende carry schrijven als  $c_{i+1}$ . De carry hangt enkel onrechtstreeks af van  $x_i$  en  $y_i$  (denk ik). Uitgaande van de 3 ingangssignalen berekenen we de volgende carry die gebruik maakt van die ingangen en  $p$  en  $g$ . Wanneer we meerdere cijfers hebben gaan we dit gedeelte dat niet in het

hokje zit in tweelagen logica implementeren. We zorgen dat de overdracht van de carry in naar de carry out zeer snel gebeurt.

**Slide 53:** Die blokjes op de dia genereert de p en de g en de som daarvan. Dat gebeurt nog altijd per cijfer, maar daarnaast heb je nog het doorgeven van de carry. Dat gebeurt in het gele blokje en dat is voor alle cijfers samen. Wat is het verschil? Stel dat 1/meer bits veranderen, dan zullen die blauw/groene blokjes ongeveer tegelijkertijd een p en een g kunnen genereren, daar is 1 vertraging onafhankelijk van welke bit je veranderd hebt. Dan heb je van die p en die g die veranderd is via een tweelagenlogica naar alle carries die er zijn. Die worden allemaal onmiddellijk gegenereerd. Die carry kan dan gebruikt worden om de juiste som te genereren. De totale vertraging heb je dus: door het groen daar propagate en generate, naar de carry en die carry wordt dan nog eens gebruikt om de som te genereren en daar wordt dan de som van genomen. De vertraging hangt nu dus niet meer af van het aantal cijfers. Doordat je aan p en g een betekenis toekent kan je neerschrijven wat er binnenin zit. De vertraging is nu onafhankelijk geworden van het aantal cijfers, maar niet helemaal. Hoe meer cijfers je hebt, hoe groter die producten worden. Je hebt wel tweelagen logica, maar je hebt altijd maar grotere poorten: zoveel ingangen als er cijfers zijn. De vertraging neemt ook toe met het aantal ingangen, dus ingangen blijven toevoegen is ook geen optie.

**Slide 54:** Uiterekend wat de vertraging is: het aantal ingangen zit er nog steeds in, zij het wel met een kleinere factor. Het grootste probleem is dat als je heel veel bits hebt, je dat niet met 1 poort kan realiseren. Dan zal je een boomstructuur van poorten moeten gebruiken. Nog een probleem dat we niet opgelost hebben: de CLA-generator is afhankelijk van het aantal bits (2x zoveel). Wat daarin zit is ook verschillend voor het aantal bits. Als je dat dus wil hebben voor 4 of 5 bits moet je dat anders ontwerpen. Voor elke hoeveelheid cijfers moet je een nieuw ontwerp maken. We kunnen dit oplossen: ervoor zorgen dat we een oplossing hebben waar we redelijk wat kunnen hergebruiken en we toch niet die lineaire toename hebben met het aantal bits. Dat is de schakeling op de slide: elk geel blokje is een CLA zoals op de vorige slide. We gaan die 1 keer ontwerpen voor een vast aantal cijfers, in dit geval voor 4 cijfers. Er komen 4 p's en 4 g's binnen en een carry. Voor 8 bits kunnen we die 2 keer gebruiken, voor 16 bits kunnen we die 4 keer gebruiken. Het is niet zo dat dat onafhankelijk is van elkaar: je moet dat doen zoals op de slide. Eerst worden die samengezet in groepjes van 4, die op hun beurt kunnen in eenzelfde component gestoken worden die de informatie van 16 generates en propagates combineert. Kom je niet toe met 16, leg je nog een laag bij, zo kan je tot 64 gaan. Elke keer je het aantal ingangen verdubbelt, komt er 1 laag bij. We hebben dus een component waarmee we alles kunnen doen. Het is meer hardware dan oorspronkelijk en op zich hebben ze elk geen fan-in probleem, maar er is meer hardware, dus we gaan er meer voor moeten betalen. Wat is de vertraging van dit? We gaan kijken wat het langste pad is: we moeten een aantal keer door de lagen door eer alle carries gegenereerd zijn. Waar zit de vertraging? Die is evenredig met het aantal lagen en dat neemt logaritmisches toe met het aantal bits. Het heeft dus geen constante vertraging maar een logaritmisches stijgende vertraging. Wanneer je dus naar snelheid gaat kijken is dit meestal de oplossing die gekozen wordt.

Zelfs voor een opteller bestaan er verschillende ontwerpen. Afhankelijk van wat je belangrijk vindt ga je een ander ontwerp kiezen: als kostprijs belangrijk is kies je bv. voor de ripple-carry, als snelheid belangrijk is ga je eerder kiezen voor de carry-lookahead.

**Slide 55:** Als we nu getallen van elkaar willen aftrekken, kunnen we de opteller gebruiken waarbij we negatieve getallen voorstellen.

**Slide 56:** De meest gekende is de sign-magnitude: we hebben een grootte en een teken. Bij conventie is een 0 een + en een 1 een -.

**Slide 57:** Je hebt ook complementsvoorstelling: je gaat van ieder cijfer het complement nemen. Als je in het decimaal stelsel werkt, neem je het 9-complement. Bij binair is dat het 1-complement.

Radix-complement: als we  $n$  cijfers hebben is het radixcomplement  $r^n - D$  (oorspronkelijk getal). Als je die twee vergelijkt zit er altijd 1 verschil tussen (tussen cijferscomplement en radixcomplement). Dat is geen toeval. Dat heeft ermee te maken dat bij het cijfercomplement van ieder cijfer van 9 afgetrokken werd. Je kan evengoed het volledig getal van 999 aftrekken (want dan zal je nooit moeten gaan lenen). Dat is wat er staat met die  $D$ . Het is belangrijk omdat we dat gaan gebruiken om het radixcomplement te berekenen, dat is veel makkelijker.

**Slide 58:** Mooie aan het tweecomplement: in het blauw: we kunnen dat beschouwen alsof het de negatieve voorstelling is. Waarom is dat zo? Als je die definitie boven het blauw hebt, kan je geen onderscheid maken tussen  $r^n$  en 0 want  $r^n$  is een 1 met  $n$  nullen achter, maar als je maar  $n$  cijfers hebt, zie je die 1 niet meer, dan is dat hetzelfde want je kan het verschil niet meer maken tussen  $r^n$  en 0. Je kan geen onderscheid meer maken tussen het radixcomplement en het negatieve getal. Je kan dat dus evengoed gebruiken als voorstelling van het negatieve getal. Eens je dat doet kan je dat optellen en dan optellen met een negatief getal.

**Slide 58:** Wat als je het omgekeerde nu hebt, hoe weet je dan of dat bv. -3 is of 13? Aan de bitcombinatie kan je dat niet zien: die bitcombinatie kan verschillende dingen voorstellen.

**Slide 59:** Bitcombinaties: de interpretatie is verschillend van afhankelijk van welke voorstelling je veronderstelt. Een bitcombinatie op zichzelf zegt niks, je weet niet over welk getal het gaat: unsigned, sign-magnitude, ... In veel gevallen gaat dat er niet bijstaan, dan betekent het gewoonlijk het 2-complement, maar het kan er evengoed ergens bijstaan.

Probleem op het examen regelmatig: er is een verschil tussen een 2-complementvoorstelling (hoe je die bitcombinatie moet interpreteren: 1010 komt overeen met -6) en het 2-complement nemen (je doet een bewerking) dat is iets anders: bij 2-complementvoorstelling komt geen hardware overeen. Bij het nemen van het 2-complement heb je hardware nodig. Je moet dus geen operator er gaan tussenzetten om het 2-complement te berekenen als er staat dat het in 2-complement staat.

Als het sign-magnitude is, is de eerste bit het teken. Bij de complementsvoorstelling

is het anders: de eerste bit geeft het teken aan, maar om de grootte te kennen moet je het complement nemen: 1000: negatief getal, hoe bereken je de grootte? Door van dat getal het 1-complement te nemen:  $0111 = 111 - 7$ . Bij het 2-complement: 1110: je moet het 2-complement nemen: eerst het 1-complement en er dan 1 bij optellen: 0001 en dan +1: 0010. Bij -8: je komt opnieuw 8 uit en je weet dat het negatief is, dus -8. Bij het 2-complement is er maar 1 voorstelling voor 0, bij het 1-complement zijn er 2 voorstellingen voor 0.

**Slide 60:** Ander probleem dat veel voorkomt: als je een voorstelling hebt in 4 bits en je moet dat eerst omzetten naar een voorstelling in 8 bits, hoe doe je dat? Voor unsigned en sign-magnitude is dat makkelijk: gewoon uitbreiden met nullen vooraan. Bij de complementsvoorstellingen is het iets complexer: als het getal positief is, moet je gewoon nullen toevoegen. Als het negatief is, moet je de radix-1 toevoegen. In het geval van bits is dat een 1 toevoegen (denk ik??). Als je een getal wil uitbreiden waarvan je weet dat het in 2-complementsvoorstelling staat, dan moet je gewoon de tekenbit er extra aan toevoegen aan de meest beduidende bits. Als die 0 was: nullen ervoor, als het 1 was: 1 ervoor.

**Slide 61:** Als we 2 getallen hebben waarvan er 1/beide negatief kunnen zijn, gaan we eerst kijken naar de tekens. Als die tekens gelijk zijn, is dat het teken van het resultaat. Als de tekens niet aan elkaar gelijk zijn, kan het zijn dat je moet testen op + of - 0. Als ze een verschillend teken hebben, moet je kijken welke van de twee de grootste grootte heeft. Dat is dus heel wat nadenken, niet gewoon een berekening doen. Afhankelijk van de verschillende gevallen telkens iets anders doen.

Een aftrekking doen vergt alleen dat je het teken moet veranderen.

Dat neemt niet weg dat je dit normaal niet gaat gebruiken. Als je dit wil implementeren in hardware heb je hardware nodig voor ieder van die testen en voor ieder van die blokjes, ook een controle die afhankelijk van de testen bepaalt welk stukje in gang schiet. Daarom gebruiken we dit normaal niet in computers.

**Slide 62:** Als je kijkt naar de ééncomplementsvoorstelling: je telt de getallen op en moet een carry-aanpassing doen. Als er een carry uitkomt, tel je die nog eens bij het resultaat op. Als je negatieve getallen hebt, moet je niet het teken veranderen, maar gewoon het 1-complement nemen. Dat is veel makkelijker met hardware.

Nadeel: je hebt 2 stappen: je moet eerst een optelling doen (met ripple carry: vertraging), dan komt daar een carry uit en moet je nog eens een optelling doen → dubbele vertraging ten opzichte van gewone optelling.

**Slide 63:** Optelling met 2-complement: door de voorstellingswijze is het optellen van een positief of negatief getal hetzelfde. Je neemt gewoon de som en dat blijft hetzelfde. Dit is dus het eenvoudigste. Het nadeel is dat de negatieve waarde berekenen vrij ingewikkeld zou kunnen zijn. Je kan doen wat op de slide staat en dan een 1 bij optellen. Dat is niet zo.

**Slide 64:** Illustratie aan de hand van voorbeelden: bovenaan staan de carries. De zwarten zijn normaal. Groene: carry-in van de minst-beduidende opteller. Als je een halve opteller gebruikt, hoeft je die niet toe te voegen. Grijs: de

carry die eruitkomt. Als je je beperkt tot 4 bits heb je die niet meer nodig en mag je die weggooien. Wat we nu gaan doen om die +1 erbij op te tellen is die bijkomende carry te gebruiken. Die +1 Doen we door de carry-ingang te gebruiken. Zo gaan we met 1 opteller een optelling doen en tegelijk +1 zonder dat we daar extra hardware voor nodig hebben. Dus zowel voor optelling als aftrekking in 2-complement heb je alleen de traditionele manier nodig zoals uitgelegd. Daarom gebruikt men meestal 2-complement.

Als je niet het aantal bits verandert, kan je in de problemen geraken (in alle gevallen). Als je de twee grootste getallen met elkaar optelt, krijg je een getal dat je niet meer kan voorstellen:  $7+6$ =niet-voorstelbaar in 2-complement met 4 bits. 13 is dus niet voor te stellen, je hebt een extra bit nodig. Als je dat toch met 4 bits wil doen, hou je dat aantal bits vast. Moesten er toch 2 grote getallen opgeteld worden, wil je kunnen detecteren dat het fout is. Dat kan je doen door dat af te schatten wanneer het resultaat boven de 7 of onder de -8 gaat zijn. Eenvoudige manier: kijk naar de carries. De laatste gebruikte carry en degene die eruit komt die je normaal nooit gebruikt: altijd hetzelfde normaal: altijd 00 of 11. Hier is het verschillend. Het feit dat die verschillend zijn toont dat je een overflow hebt. Je moet dus alleen die 2 bits vergelijken. Dat geldt ook voor negatieve getallen.

**Slide 65:** Er is heel veel gemeenschappelijk tussen optelling en aftrekking bij hardware. Je gaat dus veel implementaties hebben waarbij we soms een optelling moeten doen en soms een aftrekking. Je kan dan quasi dezelfde hardware gebruiken. De instructiebit (s) gaat bepalen of er een optelling of een aftrekking gaat gebeuren. Bij optelling ga je optellen zoals de traditionele optelling, aftrekking is door eerst de radix te gebruiken (denk ik?).

XOR: als er een 0 opstaat kan je bepalen dat het de x gewoon doorgeeft, als het 1 is, gaat dat x' doorgeven. Je moet die s hangen aan de onderste carry. Als je de overflow wil berekenen, komt er een XOR bij (test of 2 bits verschillend zijn van elkaar of niet).

Je kan ook nog een stap verder gaan: nog meer dingen combineren. In plaats van een afzonderlijke optelling en aftrekking te nemen, wil je een stuk hardware dat de ene keer als opteller en een andere keer als aftrekker gebruikt kan worden, dat drukt de kostprijs. Kunnen we dat nog uitbreiden?

**Slide 66:** ALU: 1 component/schakeling dat verschillende dingen kan doen: optellingen, aftrekkingen, logische functies of gelijk wat je beslist dat die allemaal moet kunnen.

Voorbeeld van een die optellingen en aftrekkingen kan doen. Logische bewerkingen: kan je bit bij bit doen, de bits beïnvloeden elkaar niet. Bij rekenkundige berekeningen beïnvloeden de bits elkaar.

Vastgelegd en nu kan je weerom zeggen dat je de specificaties hebt en dat met Karnaughkaarten oplost. Bij complexe dingen is het soms interessant om eerst na te denken en te kijken welke kennis je al hebt. Je kan eerst een uitbreiding doen van het vorige. Hier gaan we dat ook doen. We hebben hier 2 soorten bewerkingen: rekenkundige en logische. Voor de rekenkundige ga je de opteller voorzien. Je kan daar iets bovenzetten: ALE.

**Slide 67:** Dan krijg je dit schema. Waarom zo getekend? Die ALE moet zo zijn dat het voor alle cijfers hetzelfde is: je moet het maar 1 keer maken en je kan het  $n$  keer gebruiken. Eens je dat vastgelegd heb kan je aangeven of het over een rekenkundige of logische bewerking gaat, ... Dan moet je vastleggen wat er gaat uitkomen, dat is die F-kolom. Omdat het we het bovenstaande schema al kennen, kunnen we zeggen wat X en Y moeten zijn. Alle logische bewerkingen gaan in de extender gebeuren en alle rekenkundige bewerkingen, daarvoor gaan we zorgen dat die op de juiste manier gebeuren zoals gedefinieerd. De extender gaan we wel met Karnaughkaarten doen.

**Slide 68:** Elke rij in de tabel komt overeen met een kolom bij de Karnaughkaarten. (denk ik). Je moet dan de (essentiële (priem))implicanten gaan zoeken. Snellere manier hierop toepassen: eerst de enen op geïsoleerde plaatsen zoeken. Bv. de rood omcirkelde: valt niet met iets anders te combineren. De eentjes boven het vierkant rechtsonder ook. Je veronderstelt dan dat dit de enige goede oplossing is. Let er wel op dat je altijd de grootst mogelijk oplossing moet zoeken. -1 betekent een 1 voor elk cijfer.

**Slide 69:** We gaan kijken hoe we een vermenigvuldiger kunnen implementeren in hardware.

**Slide 70:** De manier van werken is hetzelfde. We gaan elk cijfer in de tweede rij vermenigvuldigen met het cijfer erboven, dat genereert een partiële product en zo voort en dat tellen we dna op. Dat is een iteratief proces, ook bij hoe wij dat doen. Je kan dat in hardware ook op een iteratieve manier implementeren. Je moet lang wachten tot je uw resultaten hebt. De vraag is of dat sneller kan. Vermenigvuldigen met  $2^n$  is er  $n$  nullen achterzetten, gewoon opschuiven: heel eenvoudig te doen.

Ook het meer algemene geval is nog te doen: we moeten die partiële producten berekenen en dan optellen. Die partiële producten zijn onafhankelijk van elkaar. Je kan die dus in parallel berekenen, je moet er alleen veel hardware tegenaan gooien.

**Slide 71:** Wat betekent dat? Tafels van vermenigvuldiging voor het binaire geval. Als we meer-bits getallen hebben, bv. 2-bits:  $a_0b_0$  en dat geeft het partiële product. Het volgende is juist hetzelfde. Dan moeten we dat nog optellen. Je zou denken dat we een 4-bitsopteller nodig hebben, maar dat is niet zo want die eerste bit moet nergens bij opgeteld worden, pas vanaf de tweede bit. En de laatste is de carry die erbij opgeteld wordt. We moeten dat wel nog aanvullen met een 0.

**Slide 72:** Meer algemeen, bv. 4x3-bit: de grootte van de schakeling/het aantal transistoren dat je al nodig hebt is een functie van het product van de bits van de getallen, bij 4x3 is dat dus 12. Als je naar grote getallen gaat worden die vermenigvuldigers zeer snel zeer groot.

**Slide 73:** Voor andere voorstelling zoals 2-complementsvoorstelling: de meest linkse bit is de tekenbit. Als je 1111 schrijft als machten van 2 is de meest linkse bit -8 en niet 8. Dat laatste is niet gewoon uitbreiden met tekenbits want

dat zou  $8x-1$  geven. Je moet gaan kijken wat je moet doen. Het is dat soort zaken dat ervoor zorgt dat een vermenigvuldiger maken voor een 2-complement maken moeilijker is dan voor natuurlijke getallen. Dat is ingewikkelder. In sommige gevallen levert het op om eerst van 2-complement naar sign-magnitude te gaan en daarmee te rekenen en dan terug te gaan.

**Slide 75:** Delen: het grote verschil is dat hier geen partiële producten zijn die onafhankelijk zijn van elkaar, hier moet je eerst kijken of de deler in de eerste  $X$  getallen van het deeltal geraakt. Bijna de enige manier om vermenigvuldigingen te doen is op een iteratieve manier. De meesten werken ook iteratief. Zelfs als je er heel veel hardware tegenaan gooit, zelfs dan zijn ze nog traag. Daarom willen we delingen zoveel mogelijk vermijden in de hardware, tenzij het een deling door een macht van 2 is want dan moet je gewoon opschuiven: minst beduidende bit weglaten. Alle anderen zijn ingewikkeld en ga je vermijden. Als je dat toch probeert op het examen, moet je dat ook kunnen verantwoorden op het examen.

**Slide 76:** Modulo rekenen: rest bij deling: eerst delen en dan rest bepalen. Dat is nog ingewikkelder dan een deling, dus dat ga je ook niet in hardware doen, behalve wanneer de deling makkelijk is zoals bij 2. De modulo van een macht van 2 is ook makkelijk. Bij VHDL heb je er 2 soorten: het bovenste en de remainder.

Wij zijn gewoonlijk geïnteresseerd in modulo en niet in de remainder.

Als we modulo van een macht van 2 doen, dan is dat gewoon zoveel bits overhouden. Als je modulo 8 ( $2^3$ ) doet, moet je gewoon de 3 minst beduidende bits overhouden. Je moet zelfs niet weten of dat 1- of 2-complement is of sign-magnitude. Dit kost dus qua hardware niks, je verbindt gewoon de  $n$  minst beduidende bits door.

**Slide 77:** We hebben basisbewerkingen gehad die we makkelijk in hardware kunnen uitvoeren. Nu gaan we komma's introduceren. Normaal spreken de meeste mensen direct over de vlottende kommavoorstelling.

**Slide 78:** In hardware met vaste komma werken is hetzelfde als met gehele getallen werken omdat een komma met niks in hardware overeenkomt: in hardware is de komma-interpretatie iets imaginair. Je doet dat in het normale rekenen eigenlijk ook zo: 2 kommagetallen met elkaar vermenigvuldigen: eerst doen alsof die er niet stonden, met elkaar vermenigvuldigen en dan de komma plaatsen. Hetzelfde gebeurt in hardware met vaste komma. Je moet gewoon bijhouden waar de komma terecht gaat komen. Voor de hardware zelf maakt dat totaal niks uit.

Wat wel belangrijk is: als je 2 kommagetallen met elkaar optelt/vermenigvuldigt, in wat resulteert dat: hoeveel bits en waar staat de komma? Je moet in gedachten houden hoe je 2 kommagetallen optelt: zo boven elkaar zetten dat de komma's boven elkaar staan en dan doe je de optelling. Je moet 1 van de 2 dus misschien wat verschuiven: nullen toevoegen aan de achterkant, dat is wat op de slide staat. Die voor de komma, daar komt 1 bit bij, dat is nodig om te zorgen dat je in geen enkel geval zorgen moet maken over overflow.

Tussendoor: hoeveel bits heb je nodig om dat resultaat correct te kunnen

voorstellen? 8 bits (lijkt in tegenspraak met dat van daarnet): de regels gaan eravuit dat alle mogelijke waarden kunnen optreden. Om echt te weten wat je moet doen, moet je kijken waartussen  $\alpha$  kan variëren: tussen 0 en 15.  $17 \cdot \alpha$  kan variëren tussen 0 en 17 keer 15. Dat is dus allemaal voorstelbaar met 8 bits. Stel dat je 4 8-bits getallen moet optellen. Je zou denken dat er met elke optelling 1 bijkomt, maar dat is fout: zie Figuur 6.1.

$$\begin{array}{r}
 8 + 8 \\
 \hline
 9 \\
 \hline
 9 + 8 \\
 \hline
 10
 \end{array}$$

Je hebt genoeg aan 10 bits. Waarom klopte dat van die 11 niet: zie Figuur 6.1.

$$\begin{array}{r}
 8 + 8 \\
 \hline
 9 + 8 \\
 \hline
 10 + 8 \\
 \hline
 10 \rightarrow \text{geen } M!
 \end{array}$$

**Slide 80:** Vlottende komma: totaal anders. Tekenbit, exponent en mantisse. Waarom in die volgorde? Als we hardware kunnen hergebruiken, doen we dat graag. Stel dat je hardware hebt om gehele getallen te behandelen en je wil die gebruiken om ook vlottende kommagetallen met elkaar te vergelijken. Die moeten er dan een beetje hetzelfde uitzien. Dat betekent dat de tekenbit helemaal links moet staan want is bij gehele getallen ook zo. Eens je meer naar rechts gaat, hoe minder beduidend uw cijfer is, hoe minder belangrijk het is. Als je twee getallen vergelijkt ga je eerst kijken of de hoogste cijfers verschillend zijn of niet en als die verschillend zijn ga je bepalen welk van de twee het grootste is. Wat het meeste impact heeft op de grootte staat rechts, dus dat moeten we hier ook doen want de exponent is belangrijker dan de mantisse.

**Slide 81:** NaN is het enige getal dat niet gelijk is aan zichzelf. Wanneer je naar hardware-implementaties kijkt gebruikt men meestal het IEEE-formaat, maar alleen de genormaliseerde voorstellingen erav. Er zijn implementaties die naar niet mee overweg kunnen. De exponent kan zowel positief als negatief zijn. Hier wordt een heel speciale voorstelling gebruikt: excess-formaat: je gaat bij de exponent een getal bijtellen zodanig dat het altijd positief is. Waarom niet als bv. 2-complement? Zou makkelijker zijn tot je de hardware wil gebruiken om te vergelijken. Als we er altijd iets bij optellen zodanig dat het zuiver positief is, geldt het weer want dan hebben we geen tekenbit meer bij wijze van spreken. Daarom wordt er zoiets speciaals gebruikt: je telt het op op het moment dat je stockeert, wanneer je het gaat interpreteren moet je dat er terug van aftrekken.

**Slide 82:** Optellen: eerst de exponenten gelijk maken, dan optellen, dan normaliseren. Dat zijn dus heel wat bewerkingen. Dat betekent dat dit ook



nog meer hardware vergt voor vlottende komma dan voor gewone bewerkingen (gehele getallen of vaste komma). Je gebruikt dus vaste komma indien mogelijk. Vermenigvuldiging is makkelijker qua stappen dan optellen omdat je de mantisses gewoon met elkaar kan vermenigvuldigen en de exponenten met elkaar kan optellen.

**Slide 83:** Speciale gevallen: als je geen getal wil voorstellen, maar bv. symbolen.

**Slide 86:** ASCII: hoe voorstellen?

**Slide 84:** BCD: voor elk decimaal cijfer voorzie je 4 bits. Je gebruikt dus maar 10 van de 16 mogelijkheden. Die anderen kunnen en mogen nooit voorkomen. Je moet niet veel bewerkingen doen op getallen: je hebt een decimale invoer en je wil 2 getallen optellen: van decimaal naar binair en terug is moeilijker. Je kan een opteller ontwerpen die BCD-getallen optelt.

**Slide 85:** Je kan het complement berekenen, dat is dan niet het 2-complement maar het 10-complement.

Hoe begin je hieraan? Niet met een Karnaughkaart, je begint met na te denken. Als we 2 cijfers willen optellen: een gewone optelling. 10: niet meer voor te stellen met 1 cijfer, ook  $7+6$ . Je moet terug van het cijfer dat binair 13 voorstelt moet je 10 aftrekken en een carry genereren naar de volgende. Wat nu in woorden gezegd is geweest kan direct getekend worden zoals te zien is op de slide.