



# Avansert CMOS

INF3400  
Del 15  
Øivind Næss

2007



## Oversikt

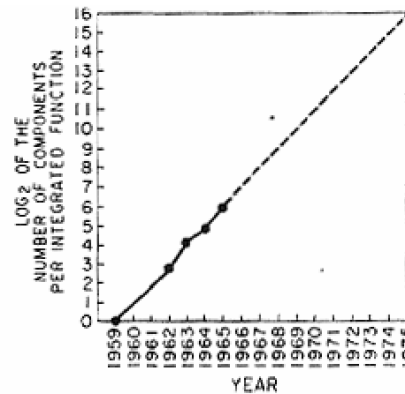
- . 4.9 Skalering
  - . 4.9.1 Transistorskalering
  - . 4.9.2 "Interconnect"-skalering
  - . 4.9.3 Teknologi "roadmap"
  - . 4.9.4 Design-påvirkninger
- . 5.4.1 I-V karakteristikk
- . 5.4.2 Terskelspenning
- . 6.3 Krets-problemer (pitfalls)
  - . 6.3.1 Terskelfall
  - . 6.3.3 Lekkasje
  - . 6.3.4 Ladningsdeling
  - . 6.3.5 Støy i forsyningspenningen
  - . 6.3.6 "Hot Spots"
  - . 6.3.7 Ladningsinjeksjon
  - . 6.3.8 Back-gate
  - . 6.3.9 Følsomhet for inngangsstøy i diffusjon
  - . 6.3.10 Prosessfølsomhet
  - . 6.3.11 Domino støyretningskap

2007



## Innledning: Moores lov

- I 1965 forutså Gordon Moore den eksponentielle veksten i antall transistorer på en IC (brikke)
- Antall transistorer dobles hver 18. måned
- Forutså >65000 transistorer innen 1975
- Vekst begrenset av effektforbruk

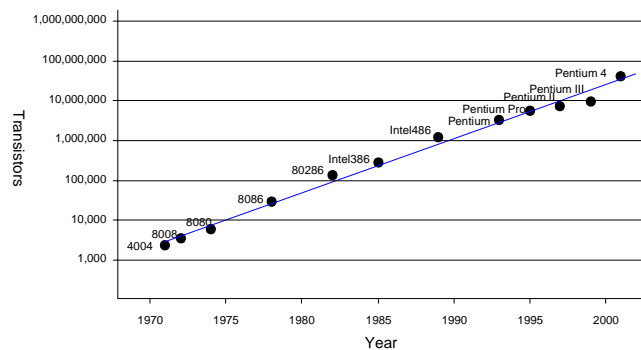


2007



## Mer Moore

- Antall transistorer er doblet hver 26. måned, de tre siste tiårene



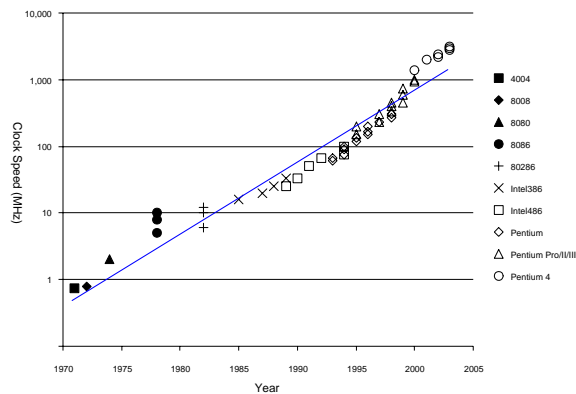
- Dette vil sannsynligvis fortsette hvertfall enda et tiår

2007



## Hastighetsøkning

- . Klokkefrekvensene har også økt eksponentielt
- . Et resultat av Moores lov



2007



## Hvorfor?

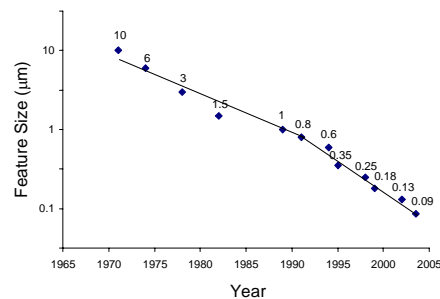
- . Hvorfor flere transistorer per IC
  - . Mindre transistor-størrelser
  - . Større brikker (dice)
- . Hvorfor raskere datamaskiner?
  - . Mindre, raskere transistorer
  - . Bedre mikroarkitekturer (mer IPC)
  - . Lavere port-forsinkelse per sykel

2007



## 4.9 Skalering, s245

- . Den eneste konstanten i VLSI er konstant forandring
- . Enhets-størrelsene reduseres med 30% hvert 2-3 år
  - . Transistorer blir billigere
  - . Transistorer blir raskere
  - . Ledningene forbedrer seg ikke (og kan bli "tregere")
- . Skaleringsfaktor  $S$ 
  - . Typisk  $S = \sqrt{2}$
  - . Teknologi-noder



2007



## Antagelser rundt skalering

- . Hva endrer seg mellom teknologi-nodene?
- . Konstant felt skalering [Dennard74]
  - . Alle dimensjonene skaleres ( $x, y, z \Rightarrow W, L, t_{ox}$ )
  - . Spenning ( $V_{dd}$ )
  - . Doping
- . Lateral skalering
  - . Bare gate-lengden,  $L$ , skaleres
  - . Gjøres ofte som en rask gate-"krymping" ( $S=1.05$ )

2007



## MOS transistor-skalering

**Table 4.15** Influence of scaling on MOS device characteristics

Parameter	Sensitivity	Constant Field	Lateral
Scaling Parameters			
Length: $L$		$1/S$	$1/S$
Width: $W$		$1/S$	1
Gate oxide thickness: $t_{ox}$		$1/S$	1
Supply voltage: $V_{DD}$		$1/S$	1
Threshold voltage: $V_{th}, V_T$		$1/S$	1
Substrate doping: $N_A$		$S$	1
Device Characteristics			
$\beta$	$\frac{W}{L} \frac{1}{t_{ox}}$	$S$	$S$
Current: $I_D$	$\beta(V_{DD} - V_T)^2$	$1/S$	$S$
Resistance: $R$	$\frac{V_{DD}}{I_D}$	1	$1/S$
Gate capacitance: $C$	$\frac{WL}{t_{ox}}$	$1/S$	$1/S$
Gate delay: $\tau$	$RC$	$1/S$	$1/S^2$
Clock frequency: $f$	$1/\tau$	$S$	$S^2$
Dynamic power dissipation (per gate): $P$	$CV^2f$	$1/S^2$	$S$
Chip area: $A$		$1/S^2$	1
Power density	$PA$	1	$S$
Current density	$I_D/A$	$S$	$S$

2007



## Merknader

- gate-kapasitans per  $\mu\text{m}$  er nesten uavhengig av prosessen siden både  $L$ ,  $W$  og  $t_{ox}$  skaleres likt
- 'PA'-motstanden forbedres med prosessen
- Portene blir raskere ved nedskalering
- Dynamisk effekt reduseres ved nedskalering
- Strøm-tetthet øker ved nedskalering
- Lateral skalering tar ikke høyde for hastighetsmetning

2007



## Skaleringsantagelser for ledninger

- . Tykkelse på ledninger
  - . Konstant vs. redusere tykkelsen
- . Lengden på ledninger
  - . Lokal/skalert interkonnekt
    - . Forsinkelsen avtar eller er konstant i nyere prosesser
  - . Global interkonnekt
    - . Brikke- (die-) størrelsen skaleres med  $D_c \approx 1.1$
    - . Øker i nyere prosesser siden brikke-størrelsen øker

2007



## Skalering av interkonnekt

**Table 4.16** Influence of scaling on interconnect characteristics

Parameter	Sensitivity	Reduced Thickness	Constant Thickness
Scaling Parameters			
Width: $w$			$1/S$
Spacing: $s$			$1/S$
Thickness: $t$		$1/S$	1
Interlayer oxide height: $b$		$1/S$	
Characteristics Per Unit Length			
Wire resistance per unit length: $R_w$	$\frac{1}{wt}$	$S^2$	$S$
Fringing capacitance per unit length: $C_{wf}$	$\frac{t}{s}$	1	$S$
Parallel plate capacitance per unit length: $C_{wp}$	$\frac{wt}{b}$	1	1
Total wire capacitance per unit length: $C_w$	$C_{wf} + C_{wp}$	1	between 1, $S$
Unrepeated RC constant per unit length: $t_{wv}$	$R_w C_w$	$S^2$	between $S$ , $S^2$
Repeated wire RC delay per unit length: $t_{wr}$ (assuming constant field scaling of gates in Table 4.15)	$\sqrt{RCR_w C_w}$	$\sqrt{S}$	between 1, $\sqrt{S}$
Crosstalk noise	$\frac{t}{s}$	1	$S$

2007



## Forsinkelse i interkonnekt

**Table 4.16** Influence of scaling on interconnect characteristics

Parameter	Sensitivity	Reduced Thickness	Constant Thickness
Scaling Parameters			
Width: $w$			$1/S$
Spacing: $s$			$1/S$
Thickness: $t$		$1/S$	1
Interlayer oxide height: $b$			$1/S$
Local/Scaled Interconnect Characteristics			
Length: $l$			$1/S$
Unrepeated wire RC delay	$R_{\text{wire}}^2$	1	between $1/S, 1$
Repeated wire delay	$l_{\text{wr}}$	$\sqrt{1/S}$	between $1/S, \sqrt{1/S}$
Global Interconnect Characteristics			
Length: $l$			$D_c$
Unrepeated wire RC delay	$R_{\text{wire}}^2$	$S^2 D_c^2$	between $SD_c^2, S^2 D_c^2$
Repeated wire delay	$l_{\text{wr}}$	$D_c \sqrt{S}$	between $D_c, D_c \sqrt{S}$

2007



## Merknader mht. skalering av interkonnekt

- **Kapasitans per  $\mu\text{m}$  er konstant**
  - Rundt  $0.2\text{fF}/\mu\text{m}$
  - Omtrent  $1/10$  av port-kapasitansen
- **Lokal ruting blir raskere**
  - Følger ikke helt forbedringen til transistorene
  - Ikke et stort problem
- **Global ruting blir tregere**
  - Ikke lenger mulig å krysse en chip i løpet av en klokkesykel

2007



## Teknologi «Roadmap»

(International Technology Roadmap for Semiconductors (ITRS))

- Halvlederindustriens (SIAs) antagelser for veien videre mht. nedskalering i teknologien

**Table 4.17** Predictions from the 2002 ITRS

Year	2001	2004	2007	2010	2013	2016
Feature size (nm)	130	90	65	45	32	22
$V_{DD}$ (V)	1.1–1.2	1–1.2	0.7–1.1	0.6–1.0	0.5–0.9	0.4–0.9
Millions of transistors/die	193	385	773	1564	3092	6184
Wiring levels	8–10	9–13	10–14	10–14	11–15	11–15
Intermediate wire pitch (nm)	450	275	195	135	95	65
Interconnect dielectric constant	3–3.6	2.6–3.1	2.3–2.7	2.1	1.9	1.8
I/O signals	1024	1024	1024	1280	1408	1472
Clock rate (MHz)	1684	3990	6739	11511	19348	28751
FO4 delays/cycle	13.7	8.4	6.8	5.8	4.8	4.7
Maximum power (W)	130	160	190	218	251	288
DRAM capacity (Gbits)	0.5	1	4	8	32	64

2007



## Skaleringsimplikasjoner

- Forbedret ytelse
- Lavere kostnader
- Utfordringer mht. Interkonnekt
- Utfordringer mht. Effektforbruk
- Fysiske begrensninger

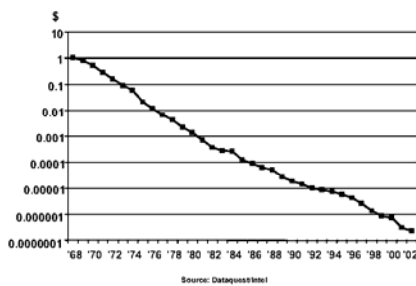
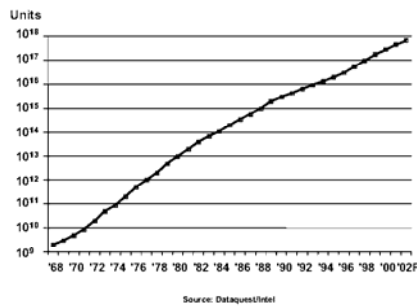
2007





## Lavere kostnader

- . I 2003 kostet 1.000.000 transistorer under 1kr.
- . Moores low gjelder fortsatt

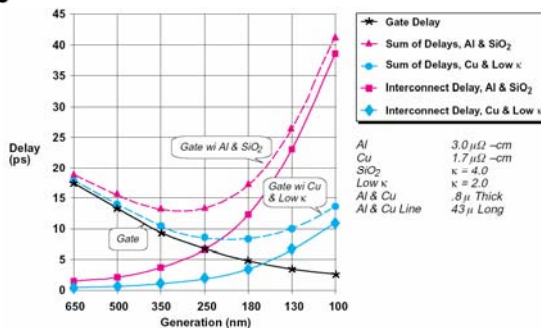


2007



## Utfordringer mht. interkonnekt

- . SIA kom i 1997 med dårlige nyheter
  - . Forsinkelser ville nå minimum ved 250-180nm prosesser, deretter bli større pga. ledningene
- . Men...
  - . Transistor-forsinkelse uten last
  - . Global ruting
- . Kopper-prosess
- . Større systemer burde deles opp i 'mindre' deler på rundt 100.000 porter

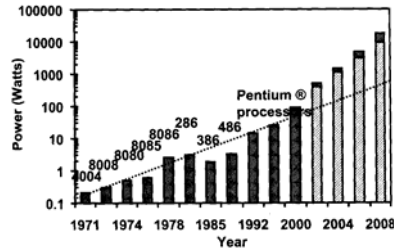


2007



## Dynamisk effektforbruk

- Patrick Gelsinger, VP Intel Corp., (ISSCC 2001):
  - 'Hvis nedskaleringen fortsetter slik som i dag, vil man i 2005 for høyhastighets-prosessorer ha en effekt-tetthet tilsvarende en atom-reaktor, i 2010 lik en romrakett og i 2015 som overflaten til solen'
  - 'Innen 2015 må man finne nye metoder for å designe integrerte kretser'
- Aksjekursen til Intel falt 8% neste dag
- Men fokus rundt effektforbruk ble viktig og er et stadig økende forskningsfelt

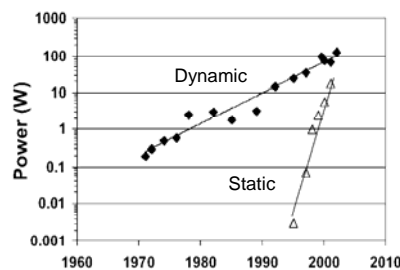


2007



## Statisk effektforbruk

- $V_{dd}$  blir stadig lavere
  - Reduserer dynamisk effektforbruk
  - Beskytter tynt gate-oksidi og korte kanaler
  - Ikke noe poeng i å ha høy  $V_{dd}$  pga. hastighetsmetning
- $V_t$  må også bli lavere for å opprettholde ytelsen
- Men dette forårsaker en eksponentiell økning i lekkasje når transistoren er 'AV'
- Stor utfordring fremover
  - Tidligere:  $V_{dd}=5V_t$
  - Nå:  $V_{dd}=3V_t$



2007



## Fysiske begrensninger

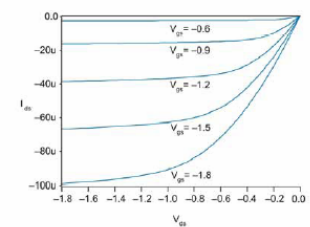
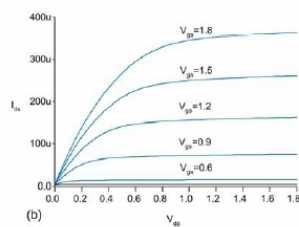
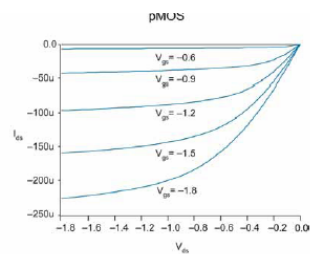
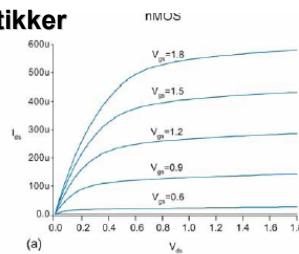
- . Vil Moores lov ta slutt
  - . Kan ikke konstruere transistorer mindre enn et atom
- . Mange faktorer har blitt foreslått som årsak til slutten på nedskaleringen av CMOS
  - . Dynamisk effektforbruk
  - . Subterskel lekkasje, tunnelering
  - . Kort-kanal effekter
  - . Prosesserings-kostnader
  - . Elektronmigrasjon
  - . Forsinkelse i interkonnekt
- . Dommedagsrykter har forøvrig ofte blitt overdrevet

2007



## 5.4 Modellering/simulering s293

- . I-V karakteristikker

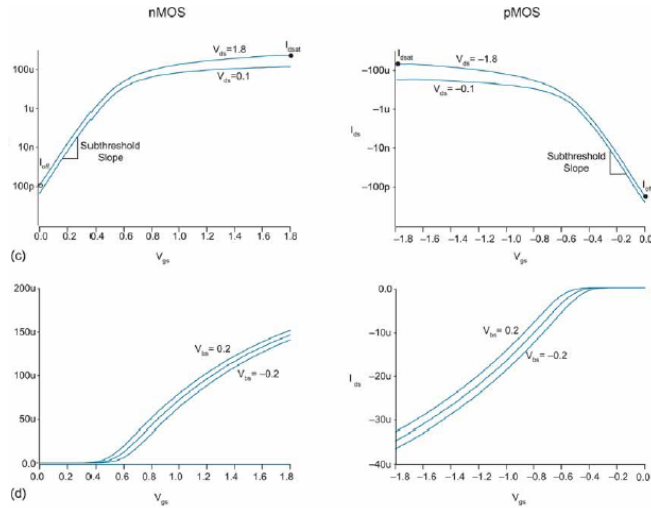


2007



## I-V karakteristikk

• •

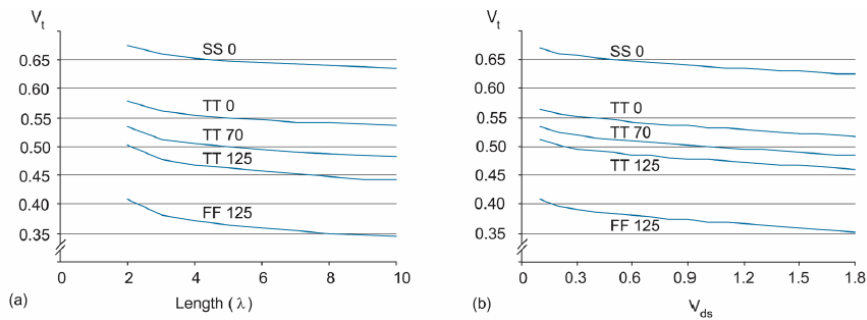


2007



## Terskelspenning

• •



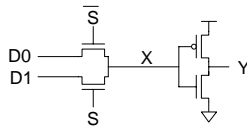
2007

## 6.3 Circuit pitfalls, s.351

### Terskelfall

- **Krets:**

- 2:1 multiplexer



- **Prinsipp: terskelfall**

- X kommer aldri over  $V_{DD} - V_t$
- $V_t$  øker pga body effect
- Terskelfallet blir mer alvorlig når forholdet mellom  $V_t$  og  $V_{DD}$  blir mindre.

- **Løsning: bruk transmisjonsporter og ikke pass-transistorer**

- **Symptom**

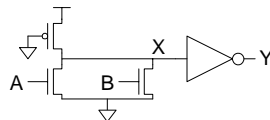
- Mux'en fungerer når D er 0 men ikke når D er 1.
- Eller fungerer ikke ved lav  $V_{DD}$ .
- Eller feiler i SFSF hjørnet.

2007

## Nivåfeil

- **Krets**

- Pseudo-nMOS OR



- **Prinsipp: Nivåfeil**

- nMOS og pMOS 'sloss' mot hverandre.
- Hvis pMOSen er for sterk vil nMOSen ikke trekke X lavt nok.

- **Løsning: Sjekk at forholdet er riktig i alle hjørner**

- **Symptomer**

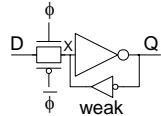
- Når bare en av inngangene er 1,  $Y = 0$ .
- Forekommer muligens bare i SF hjørnet.

2007

## Nivåfeil (2)

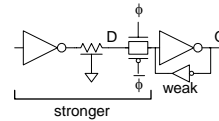
- Krets

- Latch



- Prinsipp: Nivåfeil

- Serie-motstand til D, ledningsmotstanden, og transmisjonsporten må være sterkere enn den svake feedback-inverteren.



- Løsning: sjekk relative 'styrker'

- Unngå ubuffrede diffusjons-innganger hvor det er uklart hvordan de drives

- Symptomer:

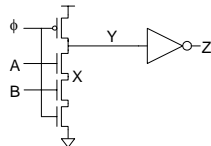
- Q 'låst' til 1.
  - Kan kun forekomme i enkelte latches hvor inngangen drives av en port 'langt' unna.

2007

## Ladningsdeling

- Krets

- Domino AND gate



- Prinsipp: Ladningsdeling

- Hvis X var lav, deler den ladning med Y

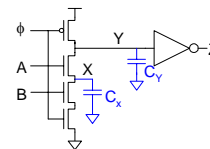
- Løsning: Begrens ladningsdelingen

$$V_x = V_y = \frac{C_y}{C_x + C_y} V_{DD}$$

- Greit hvis  $C_y \gg C_x$

- Symptomer

- Precharge porten mens  $A = B = 0$ , slik at  $Z = 0$
  - Sett  $\phi = 1$
  - A stiger
  - Z viser seg å stige av og til



2007



## Støy i forsyningsspenningen

- $V_{dd}$  og GND er ikke konstant over hele ICen
- Spenningsfall forekommer pga. motstand i forsyningslinjene
- $di/dt$  støy pga. induktanser i spenningsforsyningen og kretser som står inaktive lenge for så å begynne å switche
- Forårsaker ofte degradering av støymarginene og variasjon i forsinkelsene.

2007



## Andre problemer

- Følsomhet for inngangsstøy i diffusjon
  - Kan f.eks heve inngangen over  $V_t$
  - Kan slippe gjennom uønskede signaler i transmisjonporter
- Prosessfølsomhet
- "Hot spots"
  - Ikke-uniform effektforbruk skaper forskjellig temperatur på forskjellige deler av brikken
- Injeksjon av minoritetsbærere
  - Hot Carrier Injection

2007