

ECE 486/586

Computer Architecture

Prof. Mark G. Faust

Maseeh College of Engineering
and Computer Science

**PORTLAND STATE
UNIVERSITY**

Outline

- Technology trends
 - Technology
 - Power and energy
 - Cost
 - Effect of time, volume, commodification
 - IC costs: die area, yield, packaging, test
 - Price vs. cost
- Reliability

Milestone	1	2	3	4	5	6
Microprocessor	16-bit address/bus, microcoded	32-bit address/bus, microcoded	5-stage pipeline, on-chip I & D caches, FPU	2-way superscalar, 64-bit bus	Out-of-Order, 3-way superscalar	Superpipelined, on-chip L2 cache
Product	Intel 80286	Intel 80386	Intel 80486	Intel Pentium	Intel Pentium Pro	Intel Pentium 4
Year	1982	1985	1989	1993	1997	2001
Die size (mm²)	47	43	81	90	308	217
Transistors	134,000	275,000	1,200,000	3,100,000	5,500,000	42,000,000
Pins	68	132	168	273	387	423
Latency (clocks)	6	5	5	5	10	22
Bus width (bits)	16 bits	32 bits	32 bits	64 bits	64 bits	64 bits
Clock rate (MHz)	12.5	16	25	66	200	1500
Bandwidth (MIPS)	2	6	25	132	600	4500
Latency (nsec)	320	313	200	76	50	15
Memory Module	DRAM	Page Mode DRAM	Fast Page Mode DRAM	Fast Page Mode DRAM	Synchronous DRAM	Double Data Rate SDRAM
Module width	16 bits	16 bits	32 bits	64 bits	64 bits	64 bits
Year	1980	1983	1986	1993	1997	2000
Mbits/DRAM chip	0.06	0.25	1	16	64	256
Die size (mm²)	35	45	70	130	170	204
Pins/DRAM chip	16	16	18	20	54	66
Bandwidth (MB/s)	13	40	160	267	640	1,600
Latency (nsec)	225	170	125	75	62	52
Local Area Network	Ethernet	Fast Ethernet	Gigabit Ethernet	10 Gigabit Ethernet		
IEEE Standard	802.3	802.3u	802.3ab	802.3ae		
Year	1978	1995	1999	2003		
Bandwidth (Mb/s)	10	100	1000	10000		
Latency (msec)	3000	500	340	190		
Hard Disk	3600 RPM	5400 RPM	7200 RPM	10000 RPM	15000 RPM	
Product	CDC Wrenl 94145-36	Seagate ST41600	Seagate ST15150	Seagate ST39102	Seagate ST373453	
Year	1983	1990	1994	1998	2003	
Capacity	0.03 Gbytes	1.4 Gbytes	4.3 Gbytes	9.1 Gbytes	73.4 Gbytes	
Disk form factor	5.25 inch	5.25 inch	3.5 inch	3.5 inch	3.5 inch	
Media diameter	5.25 inch	5.25 inch	3.5 inch	3.0 inch	2.5 inch	
Interface	ST-412	SCSI	SCSI	SCSI	SCSI	
Bandwidth (MB/s)	0.6	4	9	24	86	
Latency (msec)	48.3	17.1	12.7	8.8	5.7	

Computer architect/designer must be aware of major technology trends

- Process technology

- Transistor density increases by ~ 35%/year
- Die size increasing by ~10-20%/year
- Wafer size increasing (step : 8" → 12")

- DRAM density

- ~40-60% density increase per year
- Cycle time reduced by 1/3 every 10 years
- Bandwidth increases at 2x rate of latency decreasing
- Evidence rate of capacity increase slowing

- Magnetic disk technology

- >100% density increase annually since 1990
- Access time reduced 30% in 10 years

- Networking technology

- Bandwidth and latency (bandwidth primary focus)
- Accelerated in recent years (1Gb Ethernet vs 100 Mb availability)

Process Technology Scaling

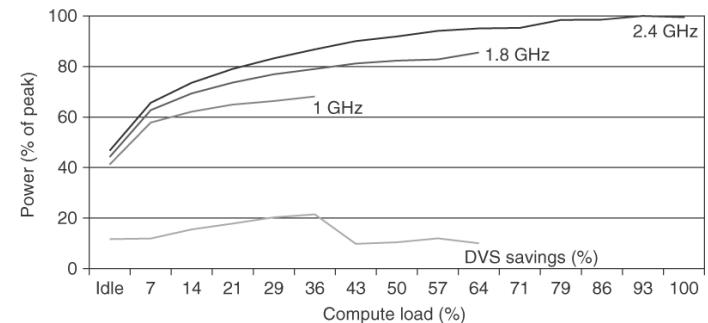
- Scaling refers to reduction in integrated circuit feature size (transistor length, minimum wire width, wire spacing)
 - 1971: 10μ ($\mu = 10^{-6}$ meters)
 - 2001: 0.18μ
 - 2011: 32nm, 22nm
 - Early 1980s: 32-bit microprocessor on single chip
 - Late 1980s: Integrated L1 cache onto chip
 - Late 1990s: Integrated L2 cache onto chip
 - Early 2000s: Integrate L3 cache onto chip, multicore

Process Technology Scaling

- Transistor performance typically increases linearly with decreasing feature size (complex for variety of reasons including decrease in supply voltages)
- Transistor count increases quadratically with decreasing feature size
- Wire delay more complex, doesn't scale well
 - Delay proportional to Resistance x Capacitance
 - Wire length decreases with feature size
 - Resistance and Capacitance /Unit Length increase
- Power
 - Proportional to: # Transistors x C_L x frequency x V^2
 - 1st microprocessor: 0.1W, Intel P4 >200W

Power and Energy

- Power, energy important in all computing segments
 - Maximum power → determines power supply
 - Sustained power → determines cooling requirements
 - Impact on cost
- Techniques
 - DVFS: Dynamic Voltage-Frequency Scaling
 - Adjust clock rate and voltage for workload
 - Do nothing well
 - Turn off clock to inactive modules (e.g. floating point unit, idle cores, caches)
 - Design for typical case
 - Low power modes require time to emerge from (slower spinning disk, low power DRAM states) for even low levels of activity
 - Instead design for typical case, throttle back if temperature increases
 - Overclocking
 - Intel "Turbo" mode
 - 3.3 GHz Core i7 can run at 3.6 GHz for short periods of time



Power Consumption

- Limits to power/heat dissipation
 - Laptops: heat, battery life
 - Desktops: cost, noise of fans
 - Servers: significant cost of ownership

$$\text{Power}_{\text{dynamic}} \text{ (watts)} = \frac{\text{Capacitive Load} \times \text{Voltage}^2 \times \text{Switching Frequency}}{2}$$

more relevant for battery-powered devices...

$$\text{Energy}_{\text{dynamic}} \text{ (joules)} = \text{Capacitive Load} \times \text{Voltage}^2$$

Power & Energy: An Example

Some microprocessors today are designed to have adjustable voltage, so that a 15% reduction in voltage may result in a 15% reduction in frequency. What would be the impact on energy and dynamic power?

$$\text{Power}_{\text{dynamic}} \text{ (watts)} = \frac{\text{Capacitive Load} \times \text{Voltage}^2 \times \text{Switching Frequency}}{2}$$

$$\frac{\text{Power}_{\text{new}}}{\text{Power}_{\text{old}}} = \frac{\frac{\text{Capacitive Load} \times (0.85 \times \text{Voltage})^2 \times (0.85 \times \text{Switching Frequency})}{2}}{\frac{\text{Capacitive Load} \times \text{Voltage}^2 \times \text{Switching Frequency}}{2}}$$
$$= 0.85^3$$
$$= 0.61$$

$$\text{Energy} = \text{Capacitive Load} \times \text{Voltage}^2$$

$$\frac{\text{Energy}_{\text{new}}}{\text{Energy}_{\text{old}}} = \frac{\text{Capacitive Load} \times (0.85 \times \text{Voltage})^2}{\text{Capacitive Load} \times \text{Voltage}^2}$$
$$= 0.85^2$$
$$= 0.72$$

Power or Energy?

The energy to execute a task (or workload) is the product of the average power times the execution time. So to optimize "efficiency", compare energy consumption for the same task.

Processor A has 20% higher average power consumption than Processor B but executes the task in only 70% of the time required for Processor B.

Energy consumption of Processor A on the task will be $1.2 \times 0.7 = 0.84$ of energy consumption of Processor B.

Increasing importance of static power

- Dynamic power consumption primary source of power dissipation in CMOS

$$\text{Power}_{\text{static}} \text{ (watts)} = \text{Current}_{\text{static}} \times \text{Voltage}$$

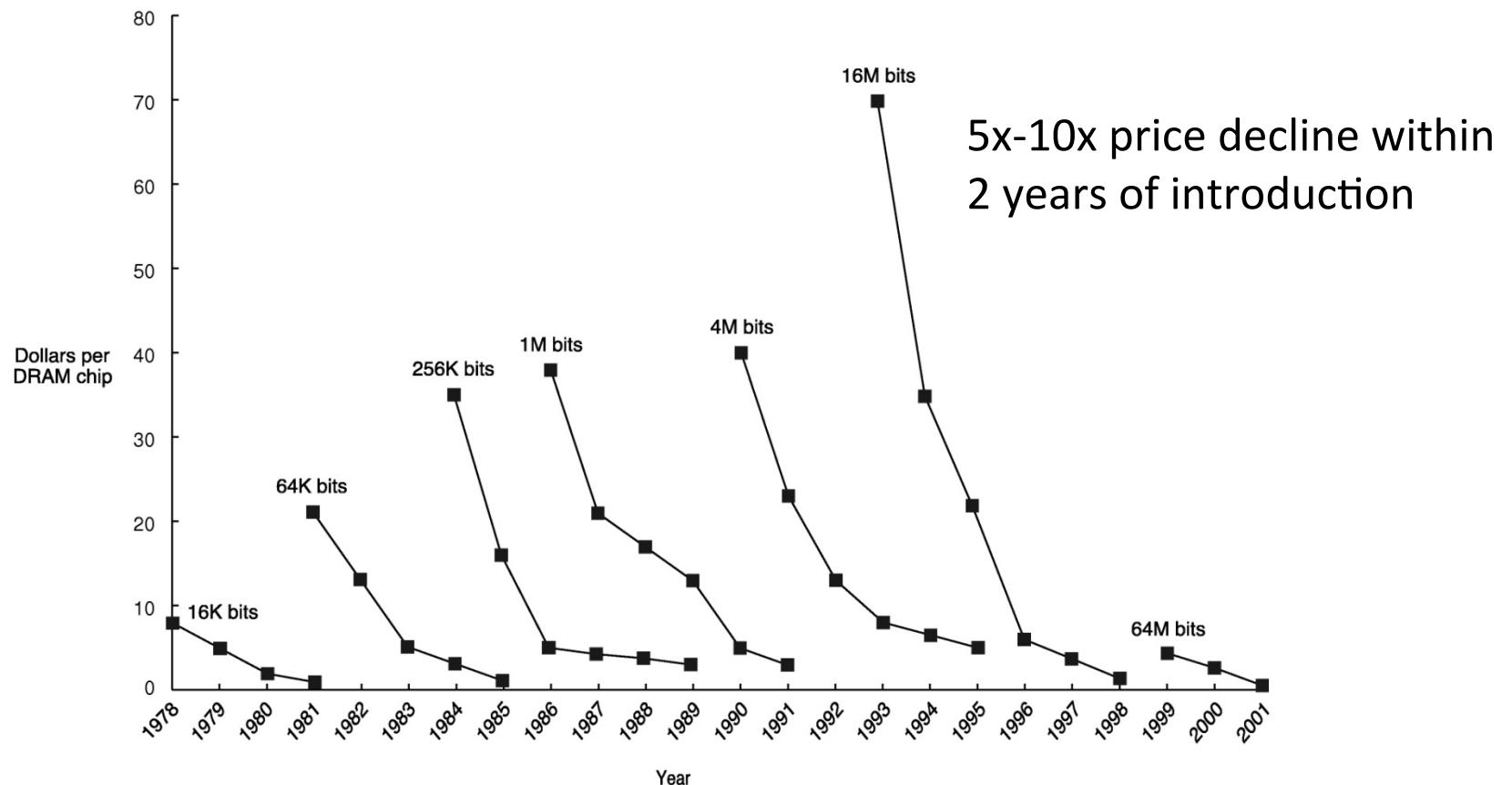
- Leakage current flows even when transistor is off
- Increases with smaller geometry and number of transistors
- Goal of < 25% of total power consumption
- Can exceed 50% of total power consumption
 - Large on-chip SRAM caches
- Gating clocks for inactive modules → gating power to inactive modules

Cost and Price Trends

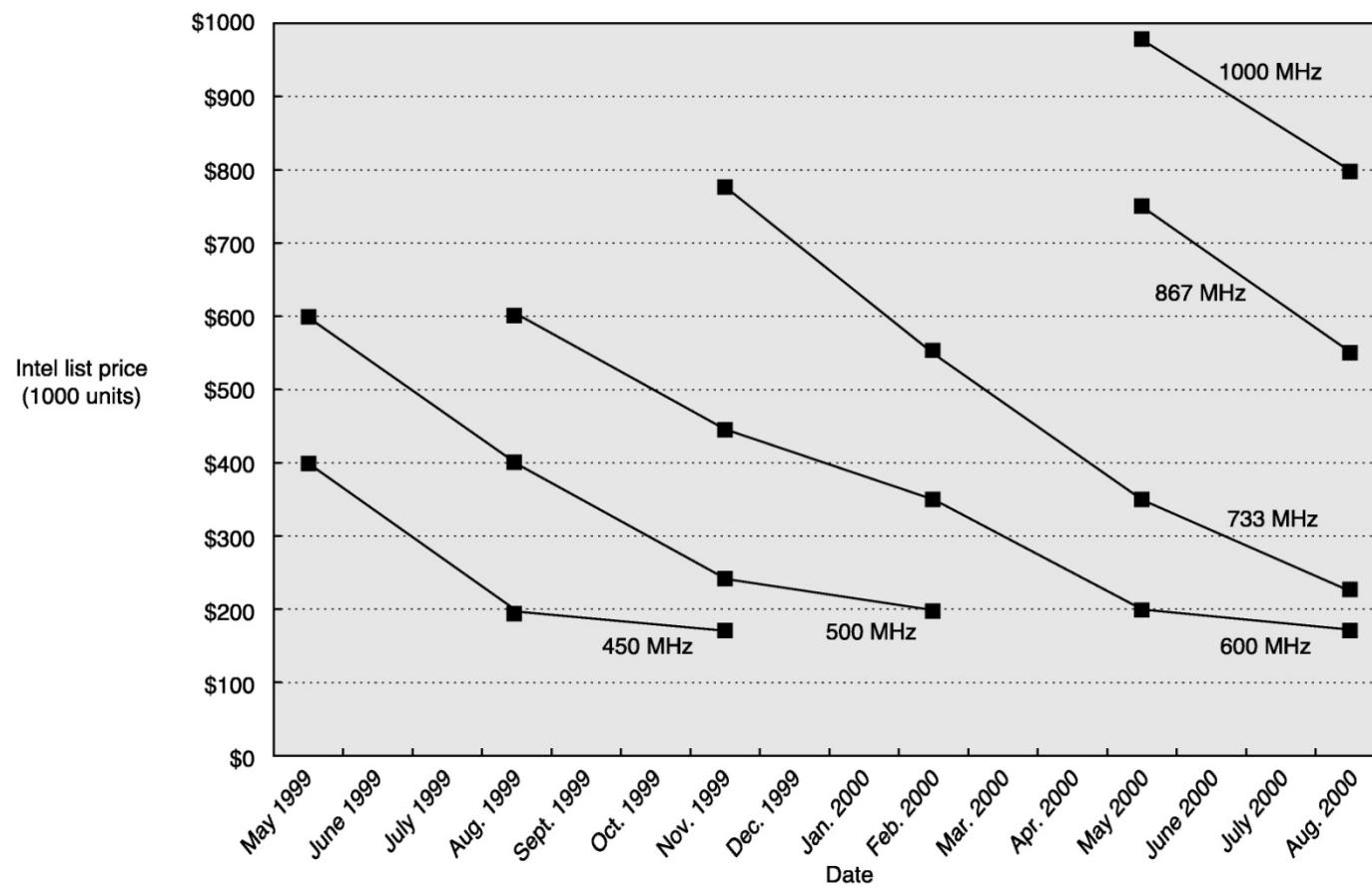
- Impacted by time, volume, and commodification
 - Cost decreases with time due to
 - Learning curve resulting in improved yields
 - Recognized opportunities for cost reductions
 - Cost decreases with volume due to
 - Learning curve
 - Purchasing and manufacturing efficiency
 - Amortized per unit cost of R&D decreases
 - Commodification
 - Multiple vendors producing large volume of essentially identical products leads to competition and price reduction
 - Intense competition results in lower profit margins and reduced prices

>50% of PCs sold for < \$500

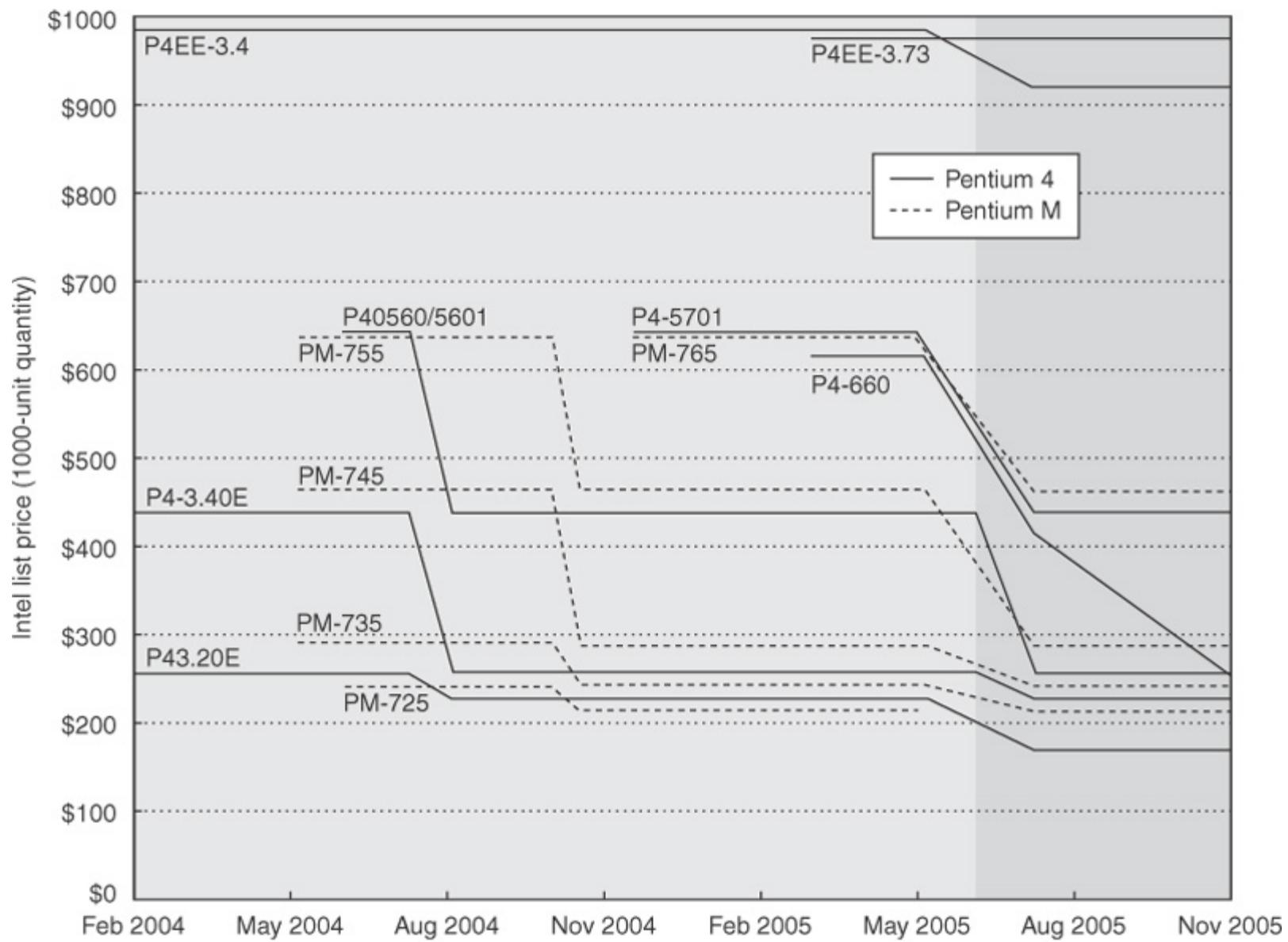
Historical DRAM Price Reductions



Intel Pentium III Pricing



Intel Pentium 4 and Pentium M Pricing



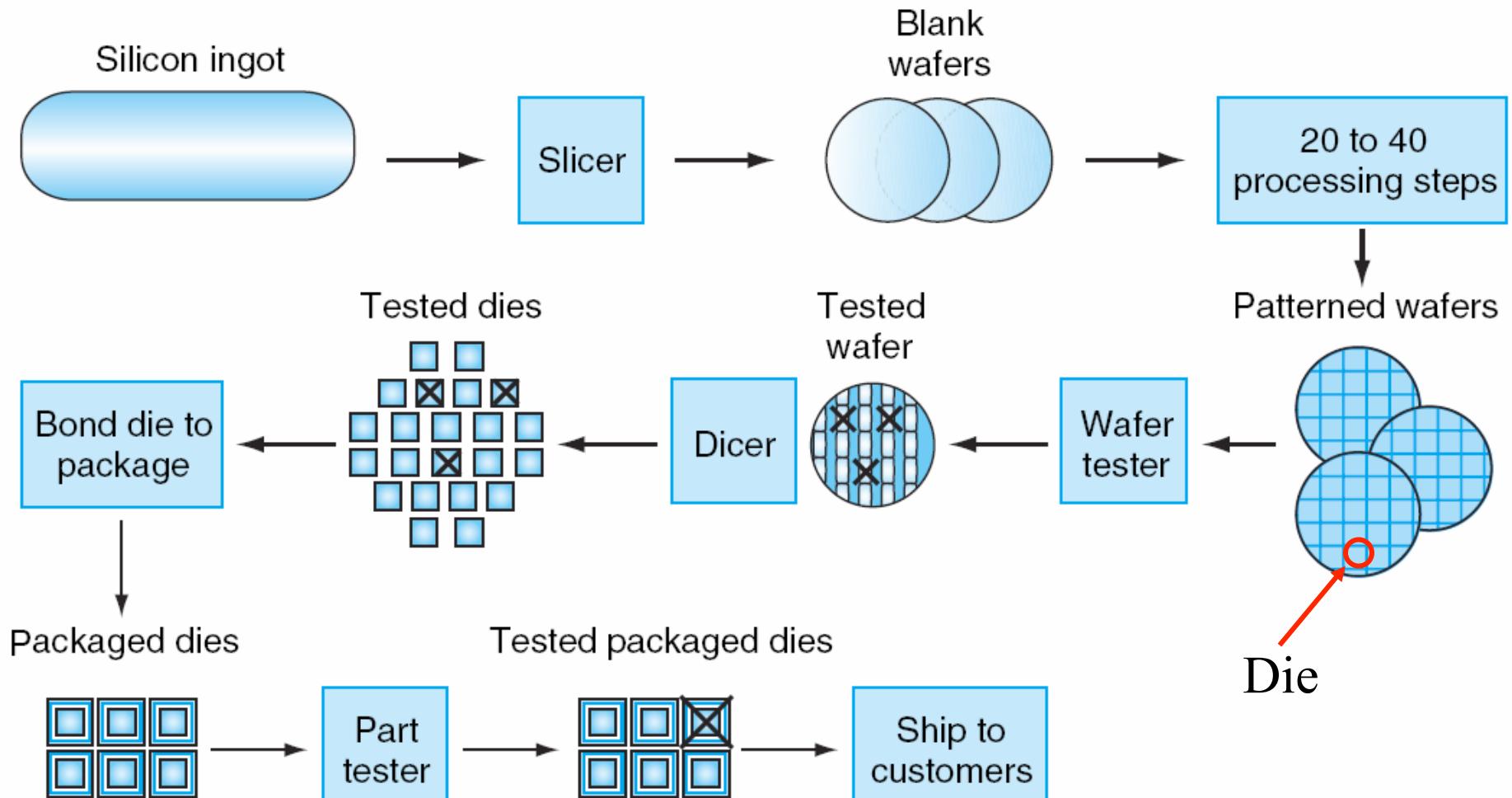
Why is all this important?

- You need to be aware of these trends to make the right design decisions and trade-offs
- You need to design for where technology will be when your product comes to market (two years from now?)
 - IC technology (density, price)
 - DRAM
 - Disk

Why is all this important?

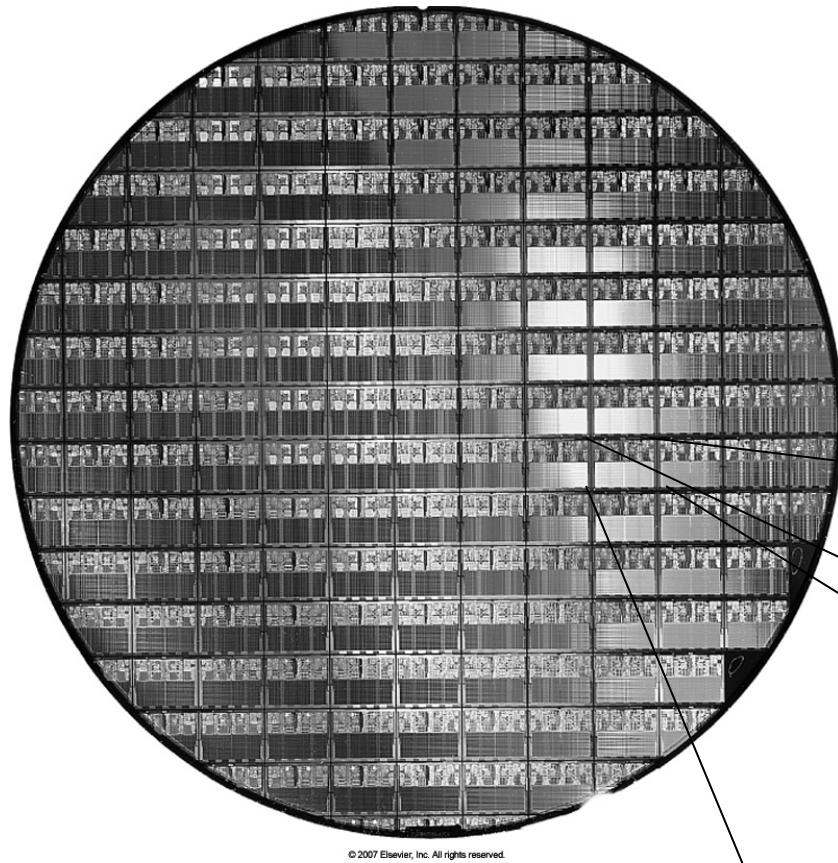
- The RAMBUS lesson
 - Commodification of DRAM put pressure on DRAM makers' profit margins so reluctant to pay RAMBUS royalties
 - Prices dropping rapidly enough that within two years plain DRAM price/performance more attractive for all but most performance hungry applications
- Opportunities
 - Patterson embarked on research which led to RAID after noticing failure of disk technology to keep pace with processor speed improvements

Integrated Circuit Costs

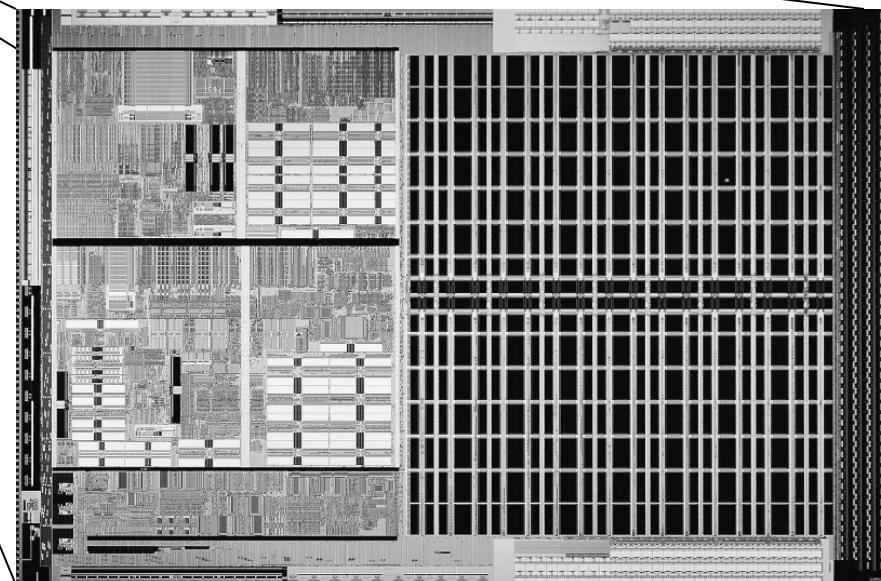


$$\text{Cost} = \frac{\text{Cost of Die} + \text{Cost of Testing Die} + \text{Cost of Packaging and Final Test}}{\text{Final Test Yield}}$$

AMD Opteron



117 die/300mm wafer
90nm process

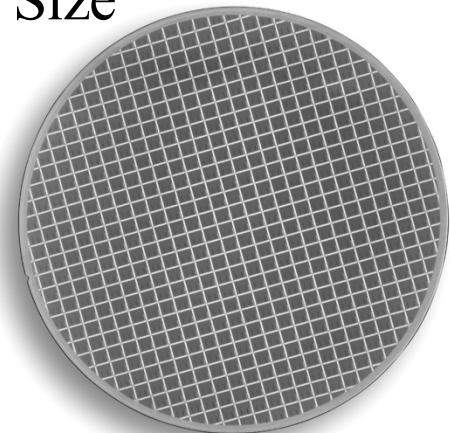
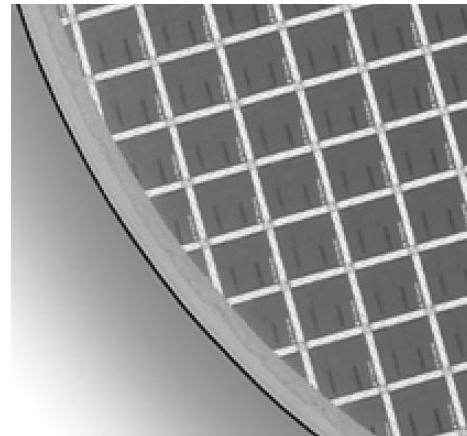


Cost of an Integrated Circuit

$$\text{Cost} = \frac{\text{Cost of Die} + \text{Cost of Testing Die} + \text{Cost of Packaging and Final Test}}{\text{Final Test Yield}}$$

$$\text{Cost of Die} = \frac{\text{Cost of Wafer}}{\# \text{ of Die/Wafer} \times \text{Die Yield}}$$

$$\text{Die/Wafer} = \frac{\pi \times (\text{Wafer Diameter}/2)^2}{\text{Die Size}} - \frac{\pi \times \text{Wafer Diameter}}{\sqrt{2} \times \text{Die Size}}$$



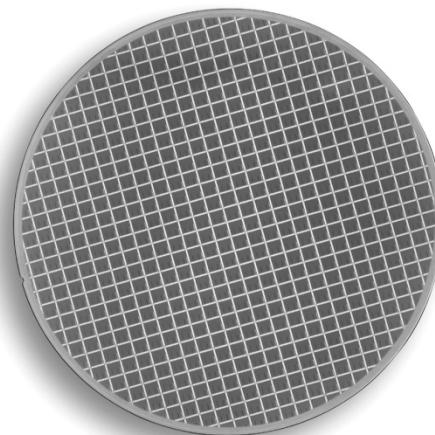
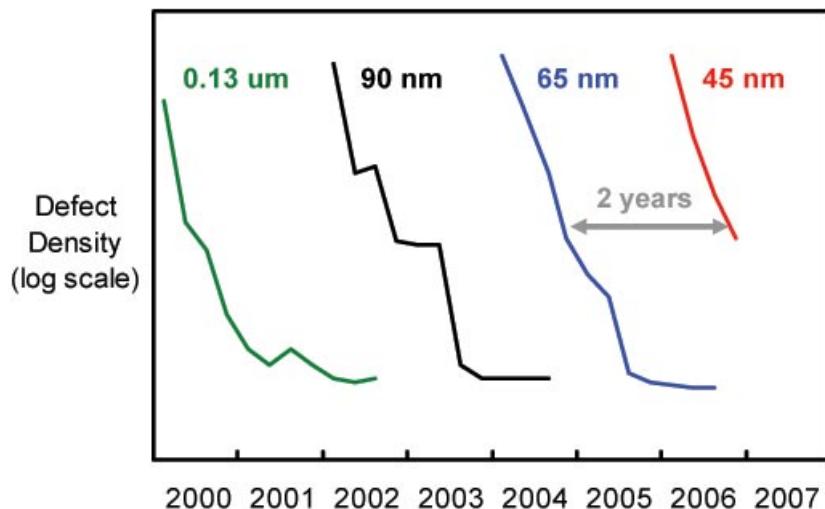
Cost of an Integrated Circuit

Bose-Einstein formula

$$\text{Die Yield} = \text{Wafer Yield} \times 1 / [1 + \text{Defects/Unit Area} \times \text{Die Area}]^N$$

$$\text{Defects/Unit Area} = [0.016 - 0.057/\text{cm}^2] \text{ (in 2010, 40nm process)}$$

$$N = [11.5 - 15.5] \text{ (related to process complexity)}$$



D1D
Oregon

Ramp in 2H '07

Mask costs $\approx \$1,000,000$

300mm wafer $\approx \$3,000 - \$5,000$ (28nm process, 300mm wafer)

New 300mm fab cost $\approx \$3$ billion

Wafers/month $\approx 20,000 - 40,000$

Yield Example

$$\text{Die/Wafer} = \frac{\pi \times (\text{Wafer Diameter}/2)^2}{\text{Die Size}} - \frac{\pi \times \text{Wafer Diameter}}{\sqrt{2} \times \text{Die Size}}$$

$$\text{Die Yield} = \text{Wafer Yield} \times 1/\left[1 + \text{Defects}/\text{Unit Area} \times \text{Die Area}\right]^N$$

Die 1: 1.5cm x 1.5cm

$$\text{Die Yield} = 100\% \times 1/\left[1 + 0.031/\text{cm}^2 \times (1.5\text{cm})^2\right]^{13.5} = 0.40$$

$$\text{Dies per (300 mm) Wafer} = \frac{(30 \text{ cm}/2)^2 \pi}{(1.5\text{cm})^2} - \frac{(30 \text{ cm}) \pi}{\sqrt{2} \times (1.5\text{cm})^2} = 270$$

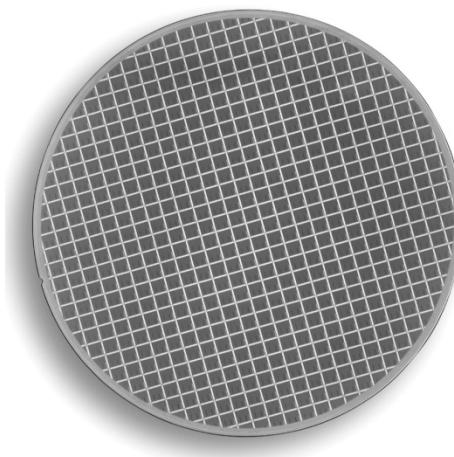
$$\text{Good Die per Wafer} = 40\% \times 270 = 109$$

Die 2: 1.0cm x 1.0cm

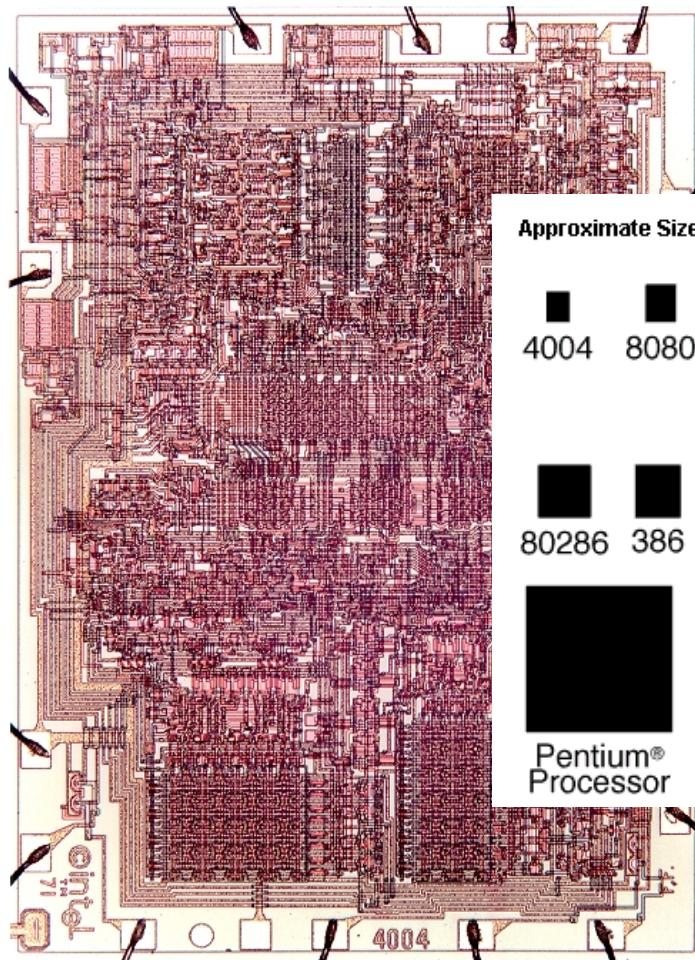
$$\text{Die Yield} = 100\% \times 1/\left[1 + 0.031/\text{cm}^2 \times (1\text{cm})^2\right]^{13.5} = 0.66$$

$$\text{Dies per (300 mm) Wafer} = \frac{(30 \text{ cm}/2)^2 \pi}{(1 \text{ cm})^2} - \frac{(30 \text{ cm}) \pi}{\sqrt{2} \times (1 \text{ cm})^2} = 640$$

$$\text{Good Die per Wafer} = 66\% \times 640 = 422$$

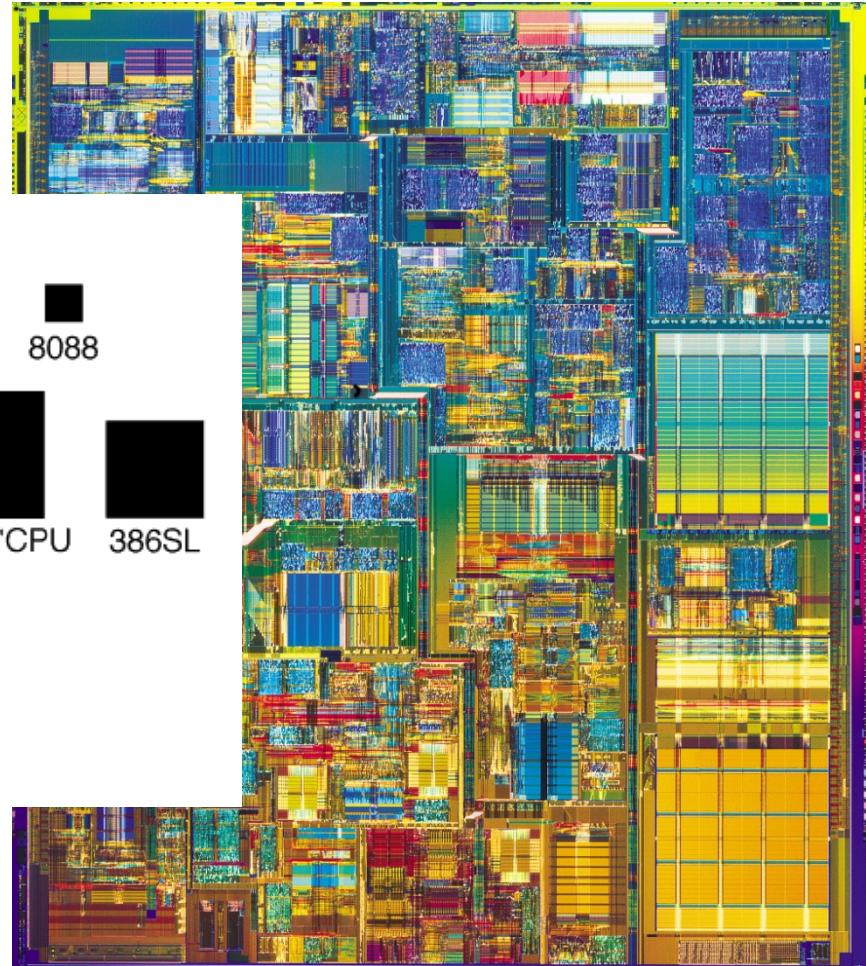
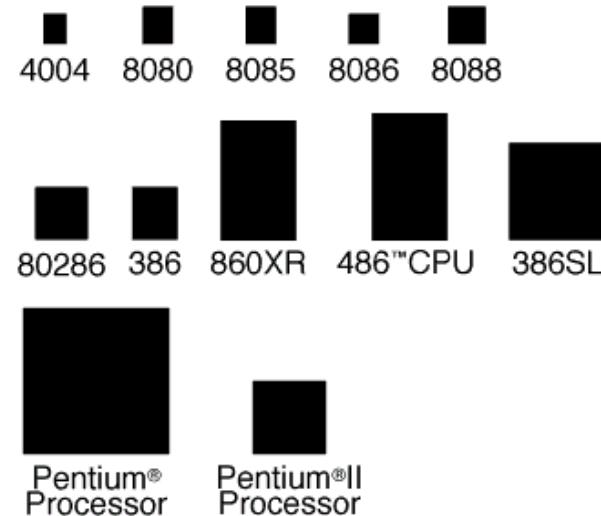


IC Density and Die Size Evolution



Intel 4004 (November 1971)
[2,300 transistors]

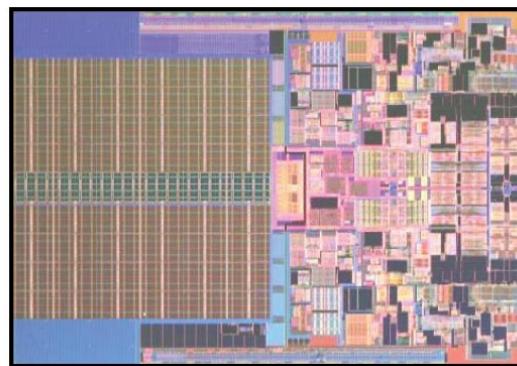
Approximate Size Relationship



Intel Pentium P4
[55,000,000 transistors]

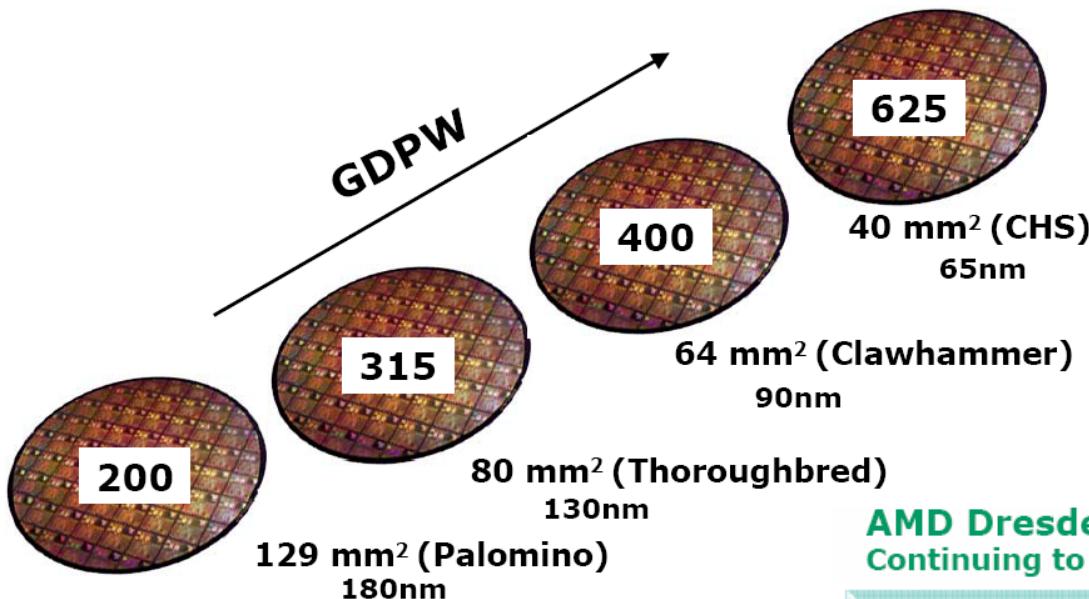
8086: 29,000
Xeon: 286,000,000

45nm Intel Penryn die



- Core 2 Duo (dual core)
- 410 million transistors
- 6 MB L2 cache
 - 2/3 of transistors
 - $\frac{1}{2}$ of die area
- SSE4 instructions
- 107 mm² die size

Impact of Die Size



November 8, 2001

AMD Analyst Conference

AMD Dresden
Continuing to Set the Standard



Fab 36

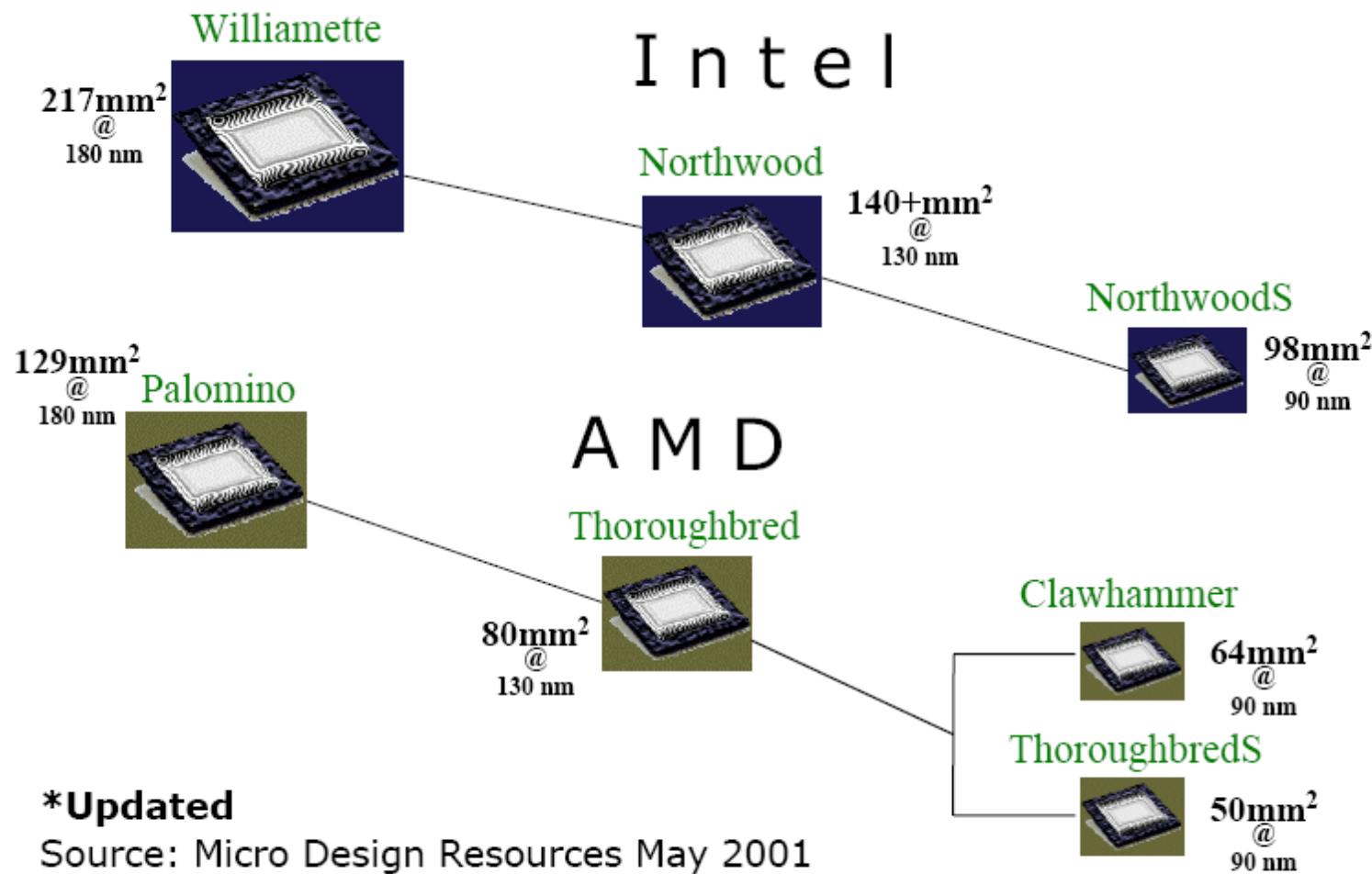
- 300mm microprocessor Fab
- Output continues to increase
- Currently transitioning to 65nm
- Expected to reach full 65nm conversion by mid-2007
- Ramped 65nm at mature yields with extremely low defect densities
- First production wafers left Fab36 in October 2006

Fab 30/Fab38

- 200mm microprocessor Fab with 300mm transition to Fab38 starting in 1H07
- New "Bump and Test" facility to be completed Q107



Die Size Transition



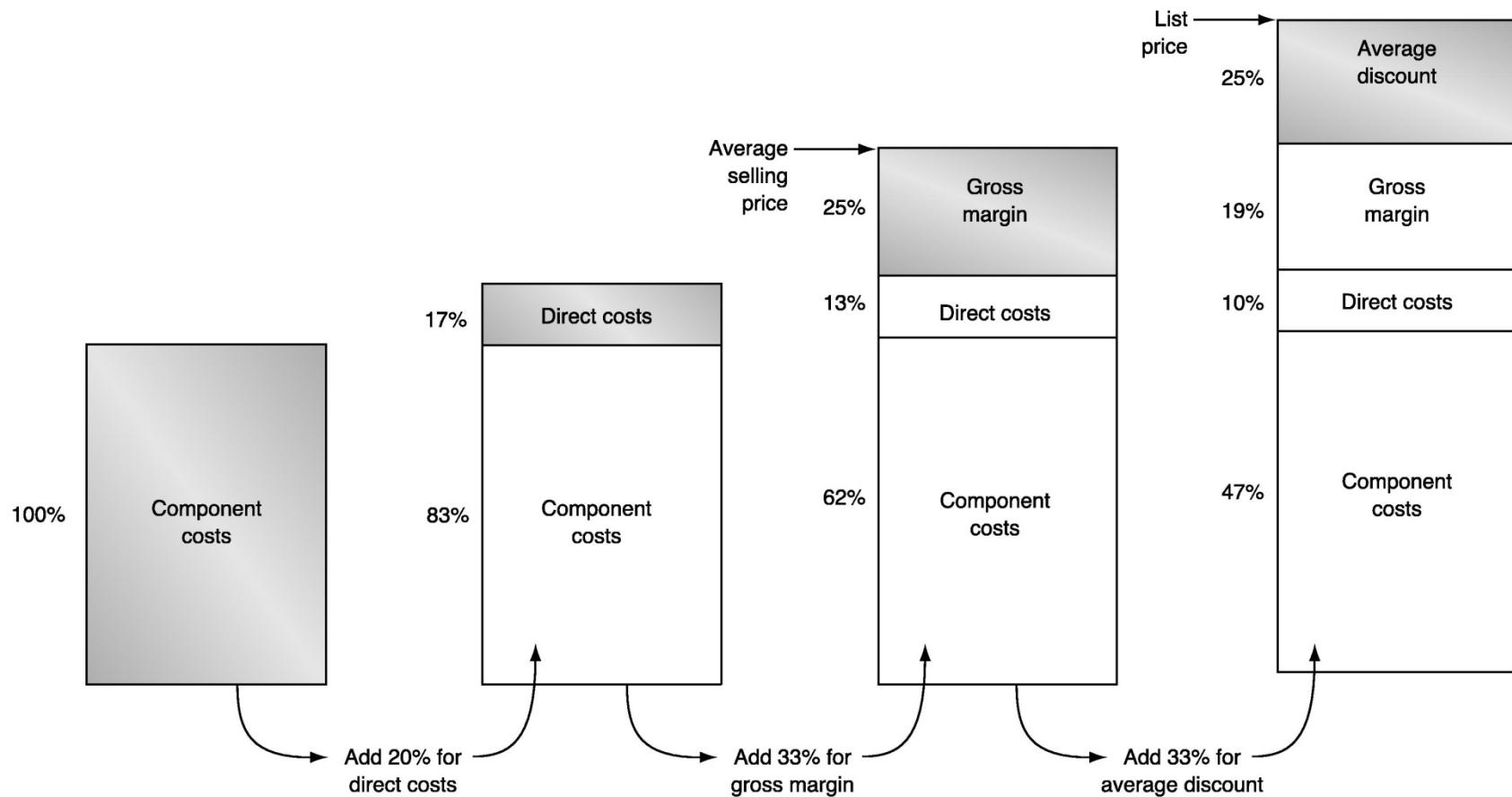
Why Does the Computer Architect Care?

- Manufacturing process dictates
 - Wafer cost
 - Wafer yield
 - Defects/Die Area
- Architect/Design controls
 - Die Size
 - On-chip vs. Off-chip functionality
 - Package Pins
 - I/Os

Distribution of Cost in a System

Component	System 1		System 2		System 3	
		Cost (% Cost)		Cost (% Cost)		Cost (% Cost)
Base server	PowerEdge R710	\$653 (7%)	PowerEdge R815	\$1437 (15%)	PowerEdge R815	\$1437 (11%)
Power supply	570 W		1100 W		1100 W	
Processor	Xeon X5670	\$3738 (40%)	Opteron 6174	\$2679 (29%)	Opteron 6174	\$5358 (42%)
Clock rate	2.93 GHz		2.20 GHz		2.20 GHz	
Total cores	12		24		48	
Sockets	2		2		4	
Cores/socket	6		12		12	
DRAM	12 GB	\$484 (5%)	16 GB	\$693 (7%)	32 GB	\$1386 (11%)
Ethernet Inter.	Dual 1-Gbit	\$199 (2%)	Dual 1-Gbit	\$199 (2%)	Dual 1-Gbit	\$199 (2%)
Disk	50 GB SSD	\$1279 (14%)	50 GB SSD	\$1279 (14%)	50 GB SSD	\$1279 (10%)
Windows OS		\$2999 (32%)		\$2999 (33%)		\$2999 (24%)
Total		\$9352 (100%)		\$9286 (100%)		\$12,658 (100%)
Max ssj ops	910,978		926,676		1,840,450	
Max ssj_ops/\$	97		100		145	

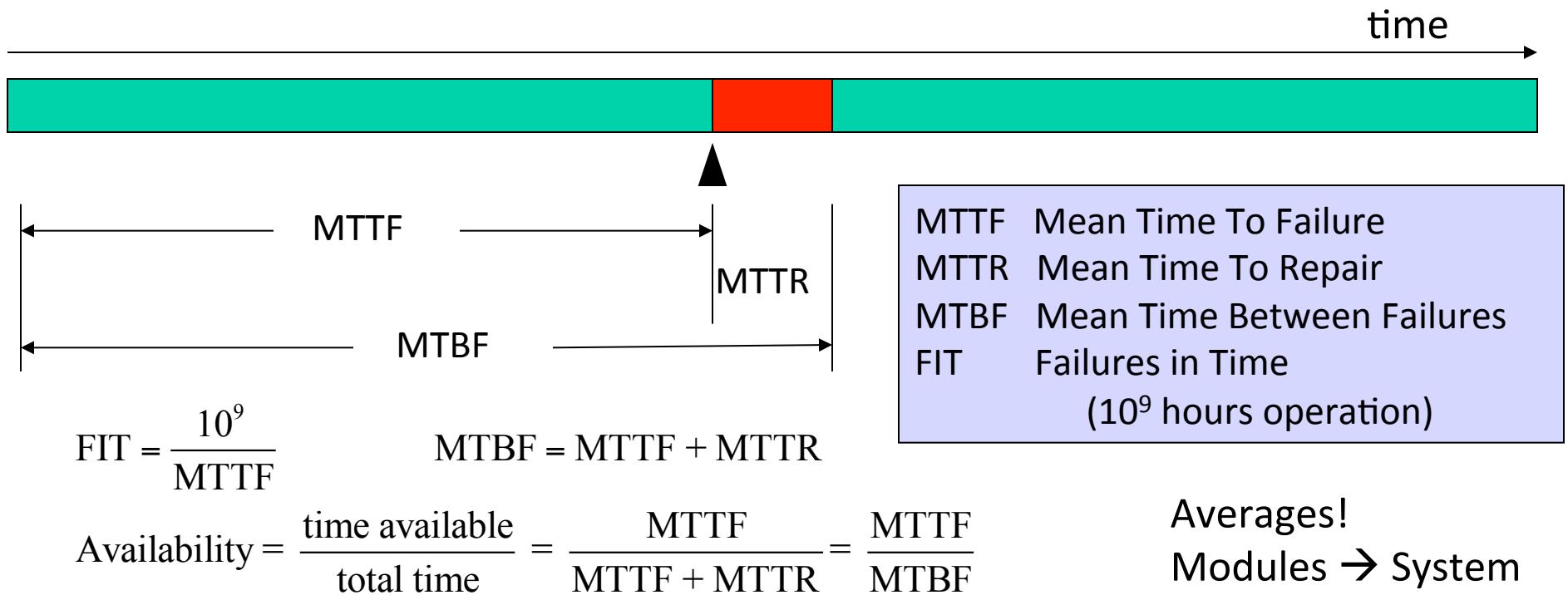
Cost and Price



Reliability

- Service accomplishment vs. Service interruption

- System operating/available according to service level agreement (SLA)
- System unavailable or performance not as agreed
- ▲ Failure which causes disruption of service



Reliability: An Example

Assume a disk subsystem with the following components and MTTF:

10 disks, each rated at 1,000,000-hour MTTF

1 SCSI controller, 500,000-hour MTTF

1 power supply, 200,000-hour MTTF

1 fan, 200,000-hour MTTF

1 SCSI cable, 1,000,000-hour MTTF

Component lifetimes are exponentially distributed, failures independent, compute MTTF of the system as a whole

$$\begin{aligned}\text{Failure rate}_{\text{system}} &= 10 \times \frac{1}{1,000,000} + \frac{1}{500,000} + \frac{1}{200,000} + \frac{1}{200,000} + \frac{1}{1,000,000} \\ &= \frac{10 + 2 + 5 + 5 + 1}{1,000,000} = \frac{23}{1,000,000} = \frac{23,000}{1,000,000,000}\end{aligned}$$

$$\text{MTTF} = \frac{1}{\text{Failure rate}} = \frac{1,000,000,000}{23,000} = 43,500 \text{ hours } (< 5 \text{ years})$$

Reliability – Exploiting Redundancy

Assume example from before but we add a second (redundant) power supply.
Maintain assumption about independence of failures!

MTTF for redundant supplies is mean time until one power supply fails divided by the chance that the second fails before the first is replaced.

$$\text{MTTF of single supply failure} = \frac{\text{MTTF}_{\text{power supply}}}{2}$$

$$\text{Probability of a second failure before first is repaired} = \frac{\text{MTTR}_{\text{power supply}}}{\text{MTTF}_{\text{power supply}}}$$

$$\text{MTTF}_{\text{power supply pair}} = \frac{\frac{\text{MTTF}_{\text{power supply}}}{2}}{\frac{\text{MTTR}_{\text{power supply}}}{\text{MTTF}_{\text{power supply}}}} = \frac{\text{MTTF}_{\text{power supply}}^2 / 2}{\text{MTTR}_{\text{power supply}}} = \frac{\text{MTTF}_{\text{power supply}}^2}{2 \times \text{MTTR}_{\text{power supply}}}$$

assuming it takes 24 hours to detect and replace failed power supply...

$$= \frac{200,000^2}{2 \times 24} \cong 830,000,000$$

making the pair about 4150 times more reliable than a single power supply