Lecture 1

Design and Technology Trends

R. Saleh
Dept. of ECE
University of British Columbia
res@ece.ubc.ca

Recently Designed Chips

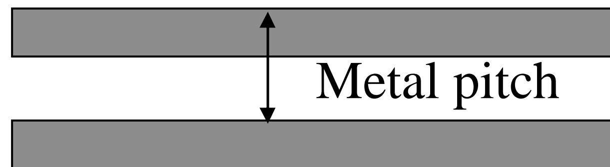
- Itanium chip (Intel), 2B tx, 700mm^2 , 8 layer 65nm CMOS (4 processors)
- TILE64 Processor, 64-Core SoC with Mesh NoC Interconnect, 90nm CMOS
- 153Mb-SRAM (Intel), 45nm, high-k metal-gate CMOS
- FPGAs recently fabricated in 45nm

- What are the major technology and design issues that are driving the IC industry?

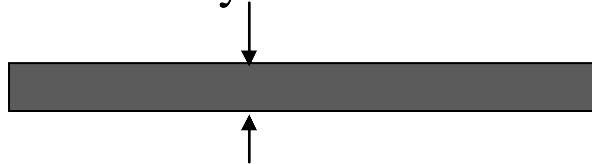
  Let's start from the simple rules of MOS scaling…

MOS Transistor Scaling
(1974 to present)

Scaling factor **s**=0.7 per node (0.5x per 2 nodes)



Technology Node
set by 1/2 pitch
(interconnect)

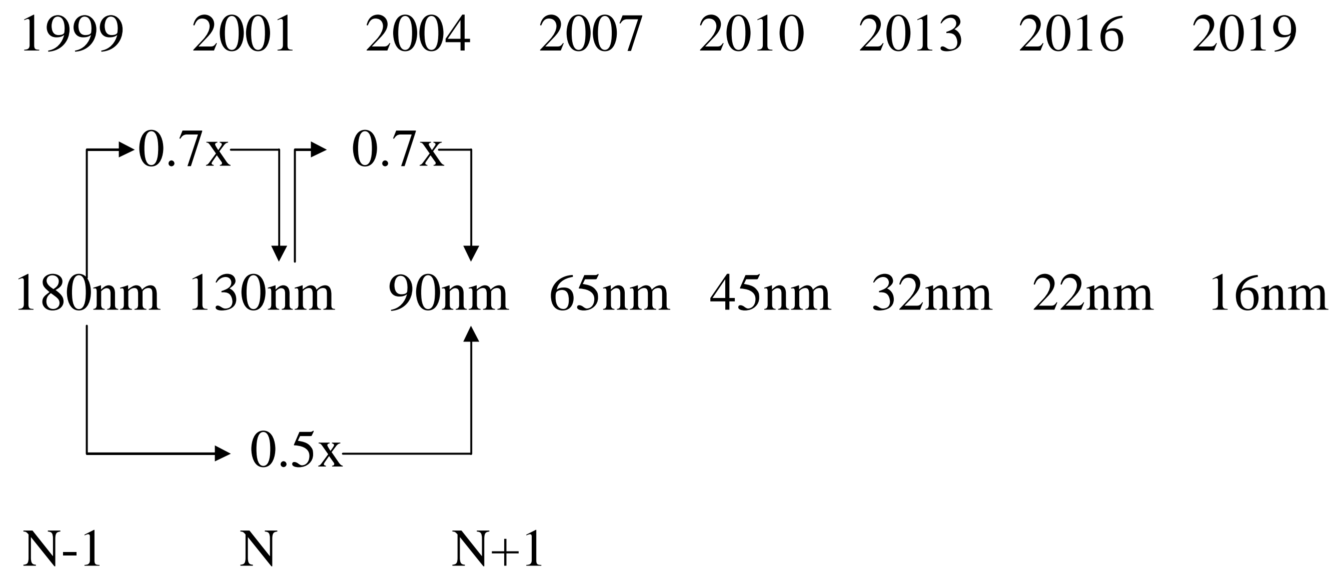


Gate length
(transistor)

Ideal Technology Scaling (constant field)

Quantity	Before Scaling	After Scaling
Channel Length	L	$L' = L * s$
Channel Width	W	$W' = W * s$
Gate Oxide thickness	t_{ox}	$t'_{ox} = t_{ox} * s$
Junction depth	x_j	$x'_j = x_j * s$
Power Supply	V_{dd}	$V_{dd}' = Vdd * s$
Threshold Voltage	V_{th}	$V'_{th} = V_{th} * s$
Doping Density, p n+	N_A N_D	$N_A' = N_A / s$ $N_D' = N_D / s$

Technology Nodes 1999-2019



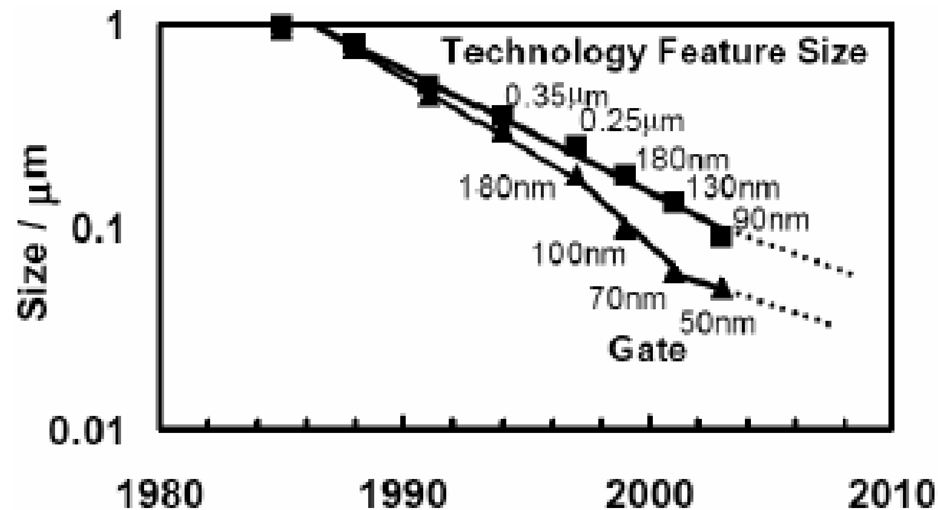
Two year cycle between nodes until 2001, then 3 year cycle begins.

Forecast Technology Parameters

Year	**Technology Node(nm)**	**Physical Gate(nm)**	**tox (nm)**	**Dielec-tric K**	**Vdd (V)**	**Vth (V)**	**Na (/cm^3)**	**Nd (/cm^3)**	**xj (nm)**
2001	130	90	3.0	3.7	1.2	0.34	1.0e16	1.0e19	67.5
2004	90	53	2.4	3.0	1.1	0.32	1.4e16	1.4e19	46.7
2007	65	32	1.7	2.5	0.9	0.29	2.0e16	2.0e19	33.8
2010	45	22	1.5	2.0	0.8	0.29	2.9e16	2.9e19	23.4
2013	32	16	1.4	1.9	0.7	0.25	4.0e16	4.0e19	16.6
2016	22	11	1.3	1.7	0.6	0.22	5.9e16	5.9e19	11.4

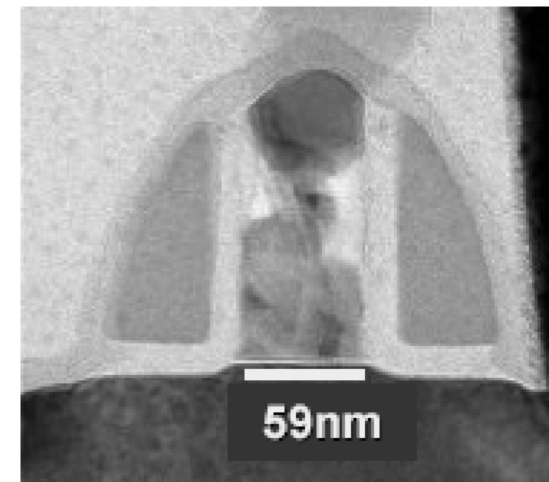
Where are we now?

- 130nm and 90nm CMOS volume production
- Early production of 65nm, Leading-edge use of 45nm



Source: Thompson *et al.*, Intel (2002)

90nm Technology



Source: Wu *et al.*, TSMC (2002)

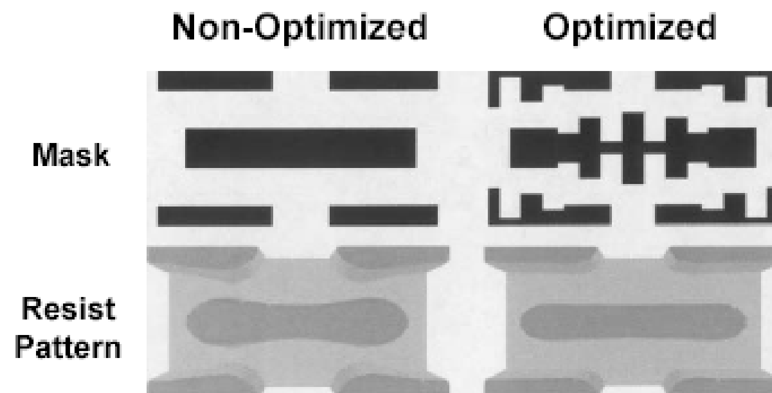
- Scaling of gate is leading scaling of wire
- Scaling is driven by DIGITAL design needs

Making Photolithograph Work

- Extensive use of OPC and PSM in 90nm and below:

Optical Proximity Correction (OPC)

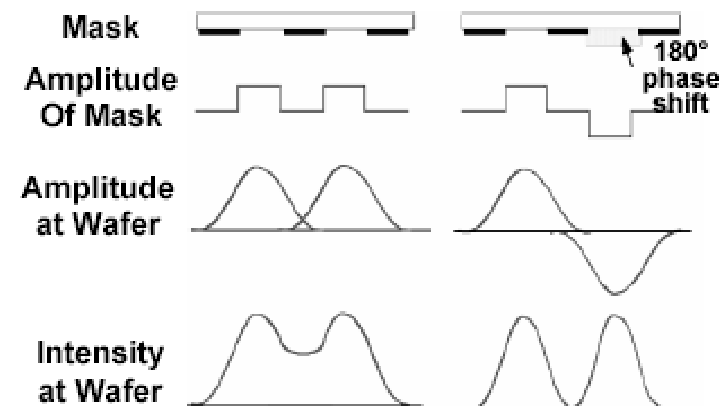
- **Add scattering features to sharpen corners**
- **Used extensively for poly gate definition**



Source: Socha, ASML (2004)

Phase Shift Masking (PSM)

- **Modulate optical path through mask**
- **Used extensively for contacts & vias**
- **Complicated for irregular patterns**



Source: Plummer, Stanford (2004)

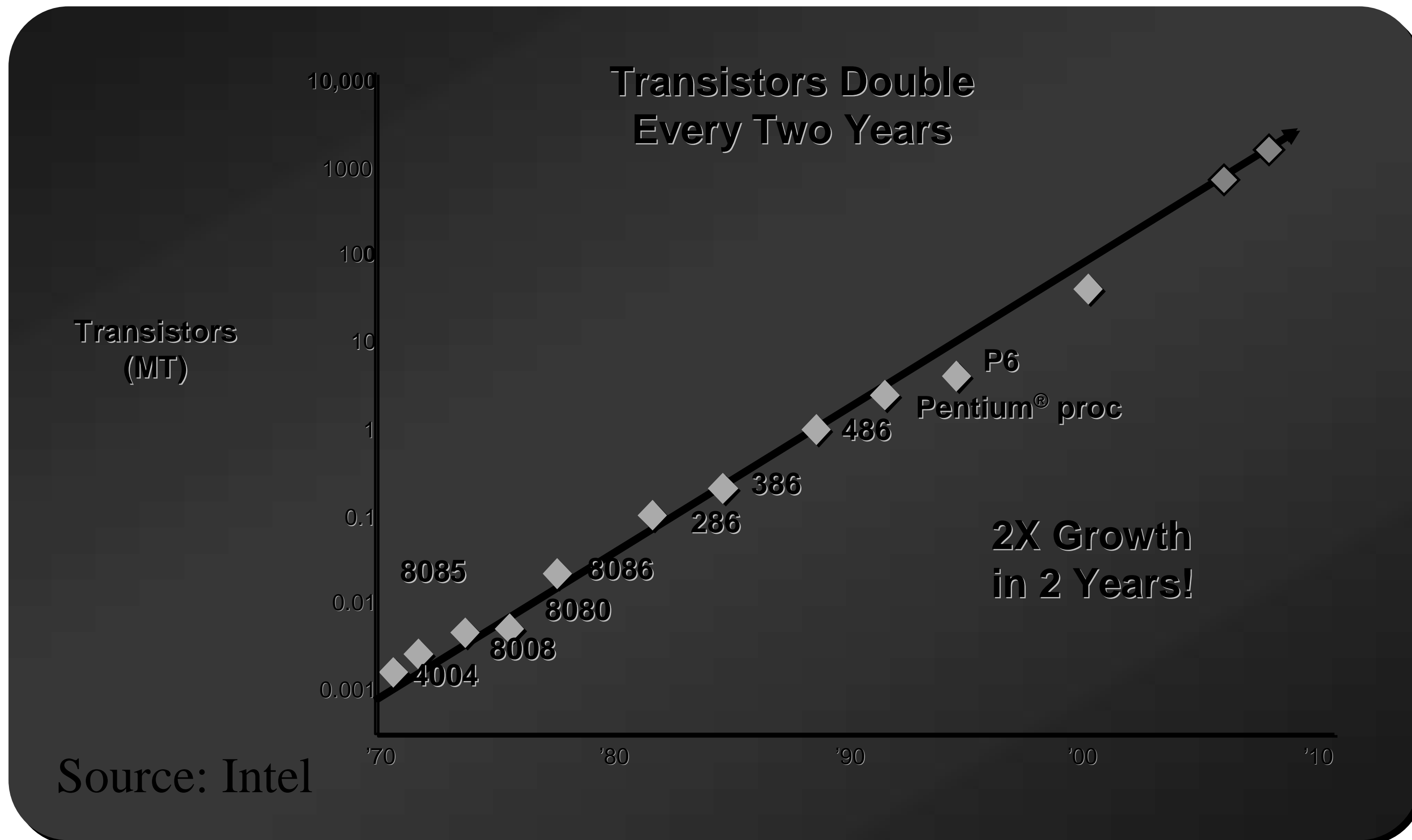
Deep Submicron Technology Generations

Table 1: Time overlap of semiconductor generations

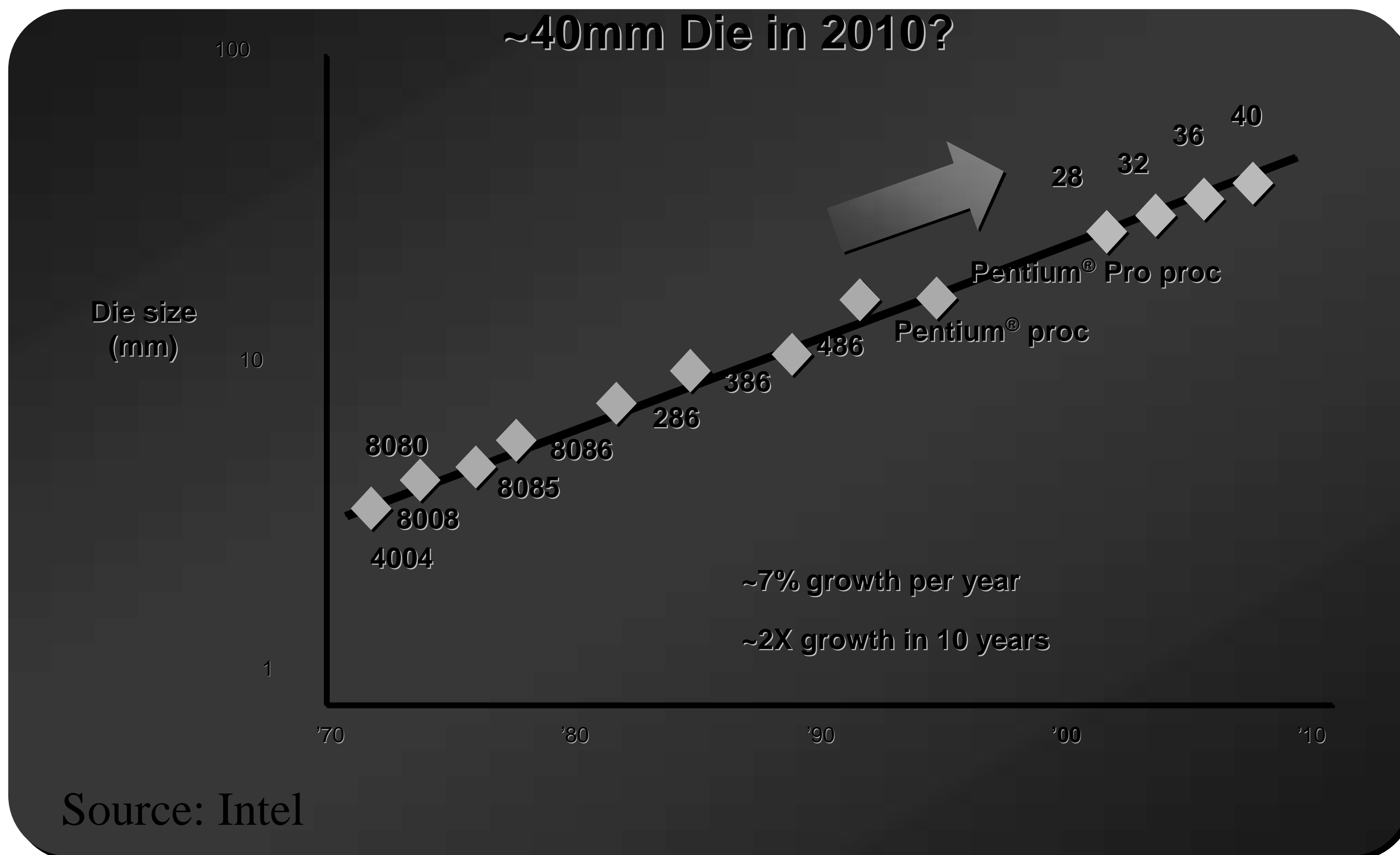
95	96	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12
350 nm	1	2	3	4	5												
-2	-1	250 nm	1	2	3	4	5										
-4	-3	-2	-1	180 nm	1	2	3	4	5	6	7	8					
-6	-5	-4	-3	-2	-1	130 nm	1	2	3	4	5	6					
-9	-8	-7	-6	-5	-4	-3	-2	-1	90 nm	1	2	3	4	5			
	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	65 nm	1	2	3	4	5
				-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	45 nm	1	2
							-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
										-11	-10	-9	-8	-7	-6	-5	-4

Each generation spans ~17 years…we are unlikely to be totally suprised

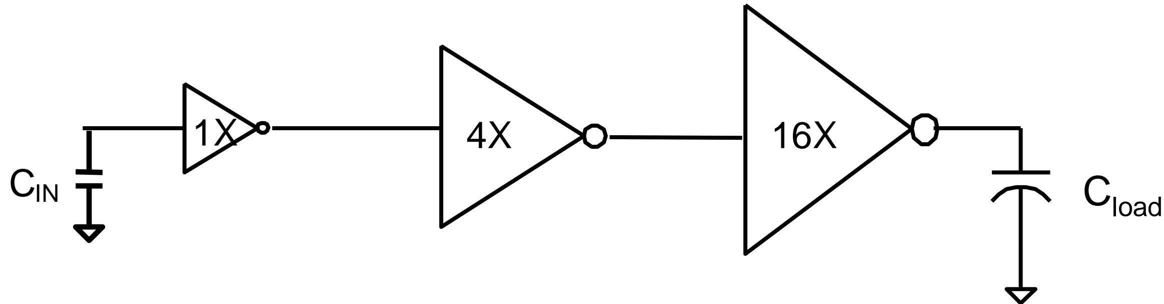
MPU Trends - Moore's Law



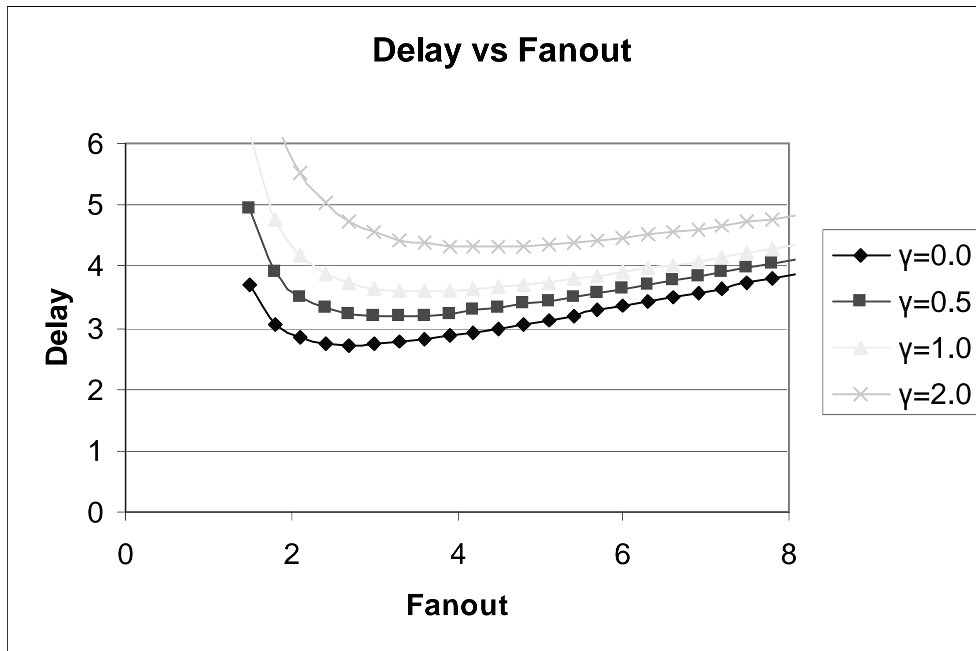
More MPU Trends



Delay Metric - FO4 Concept



Use FO4 delay as optimal delay



where γ **is ratio of Parasitic output Capacitance to gate capacitance**

FO4 INV Delay Scaling

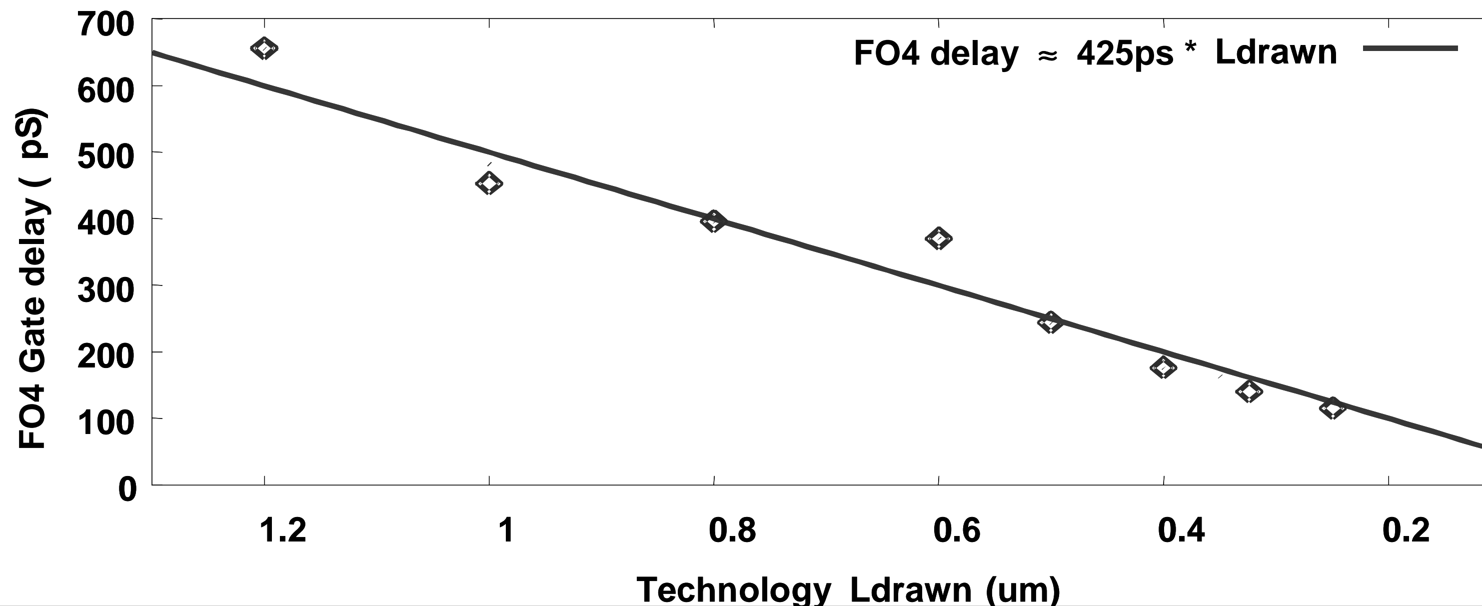
For scaling purposes, the alpha-power model is very useful:

$I_{dsat} = K\,W\,L_{eff}^{-0.5} T_{ox}^{-0.8} (V_{gs} - V_{th})^{1.25}$

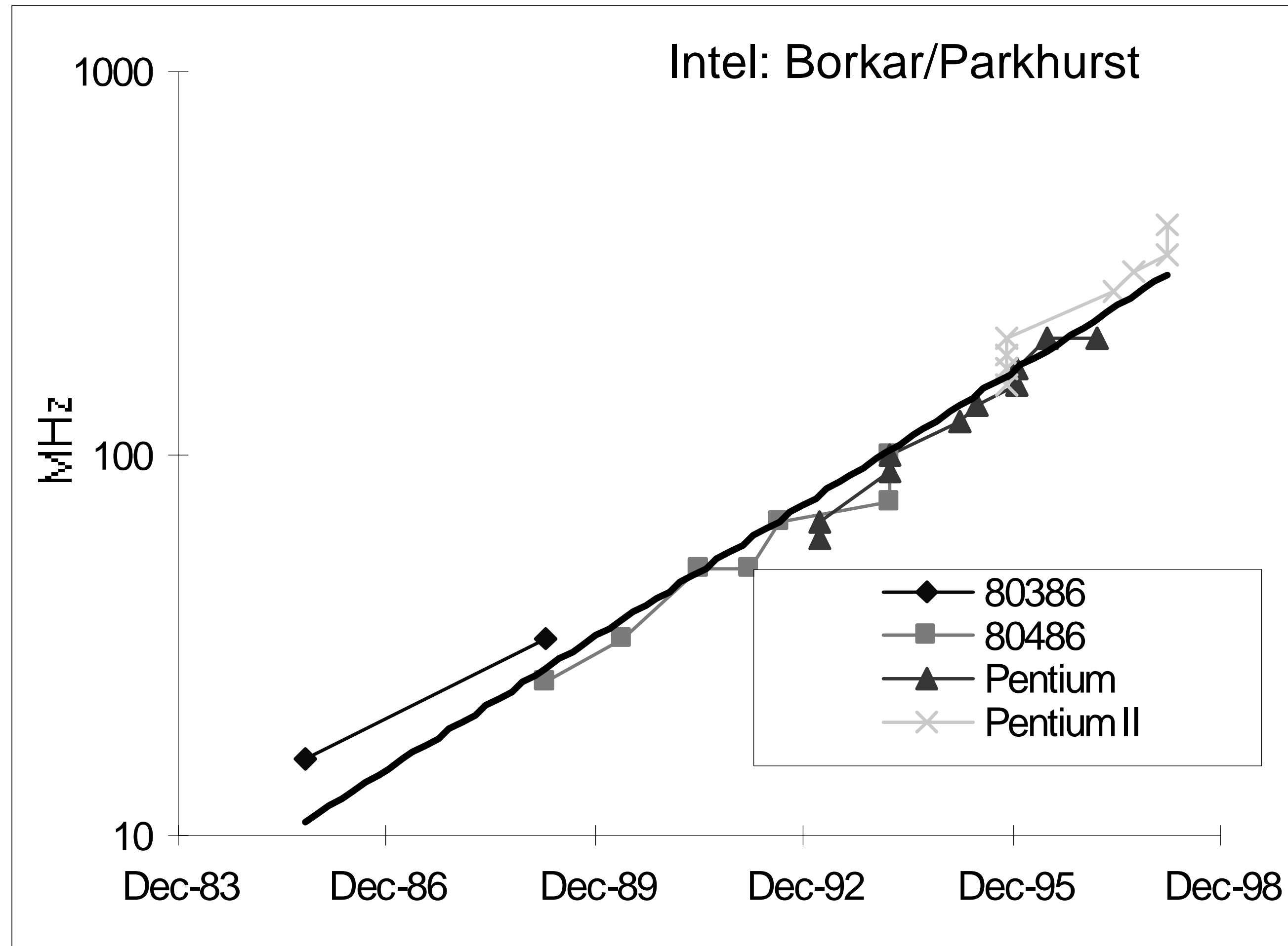
If L, T_{ox} V all scale (note V scaling will be limited by V_{th} scaling),

Current should remains constant per micron of width (approx. 600 to 800uA/μm)

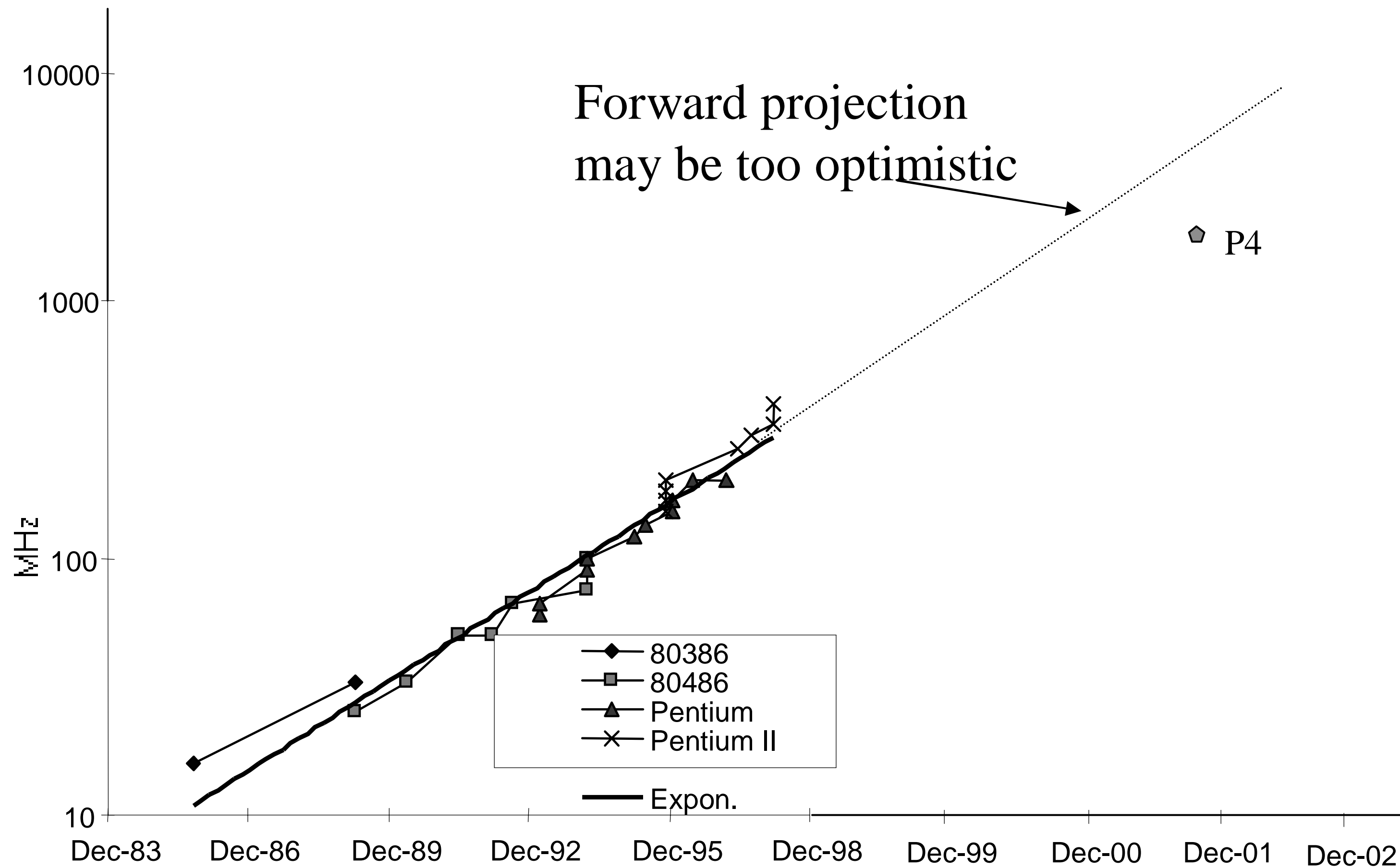
$\Delta t' = CV/i = s\Delta t$ since C, V, i all scale down by s



MPU Clock Frequency Trend

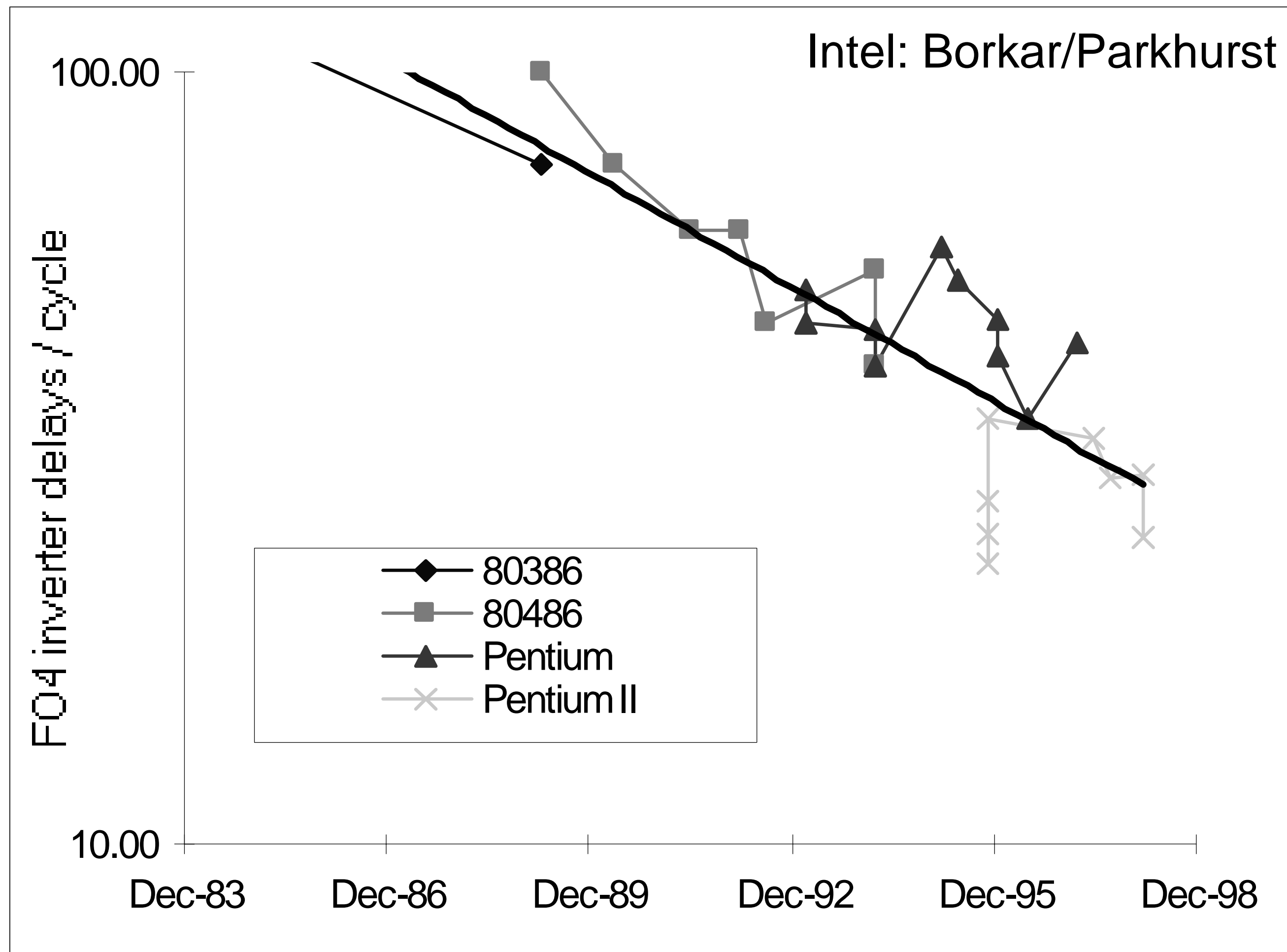


MPU Clock Frequency Trend

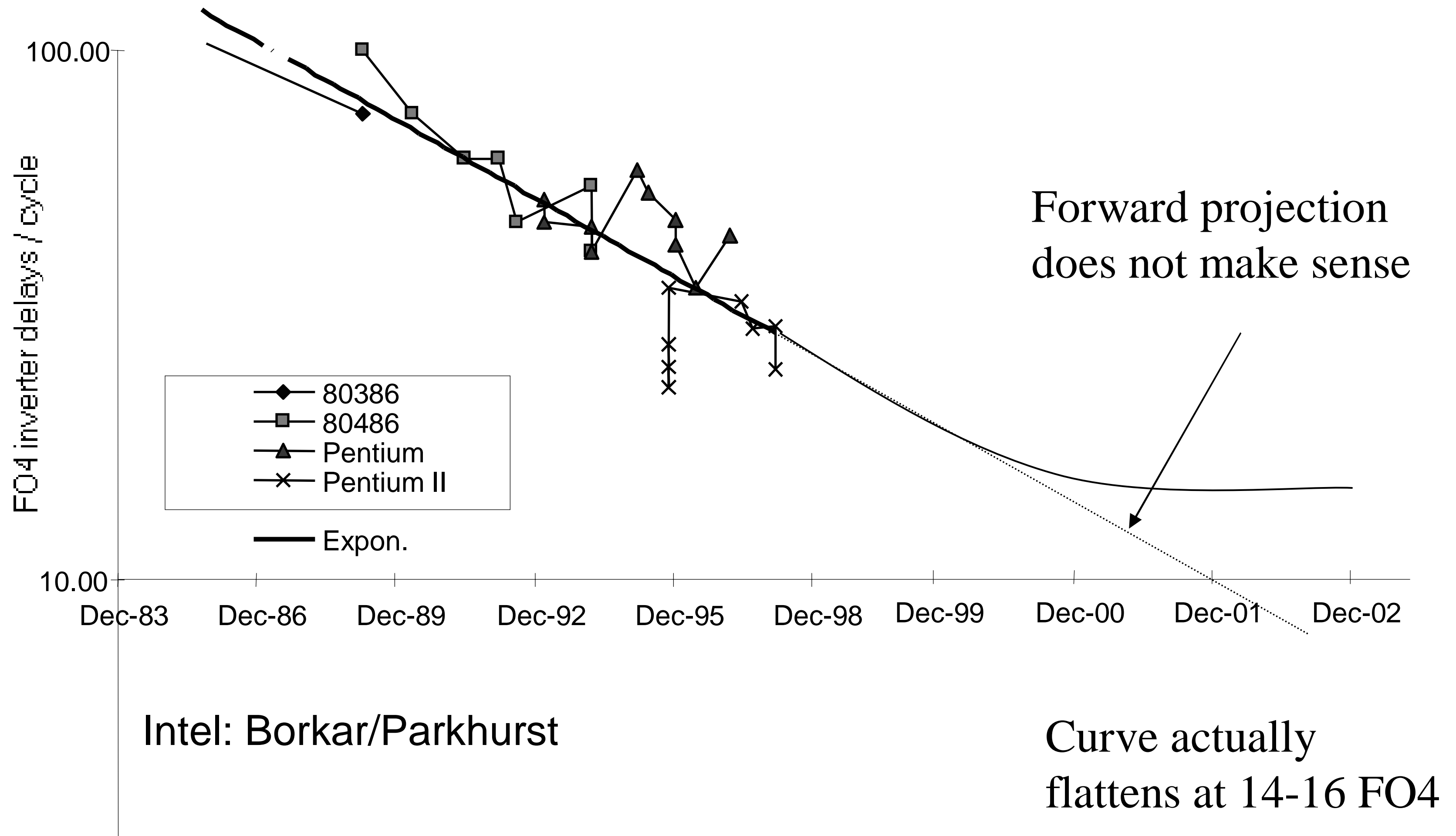


Intel: Borkar/Parkhurst

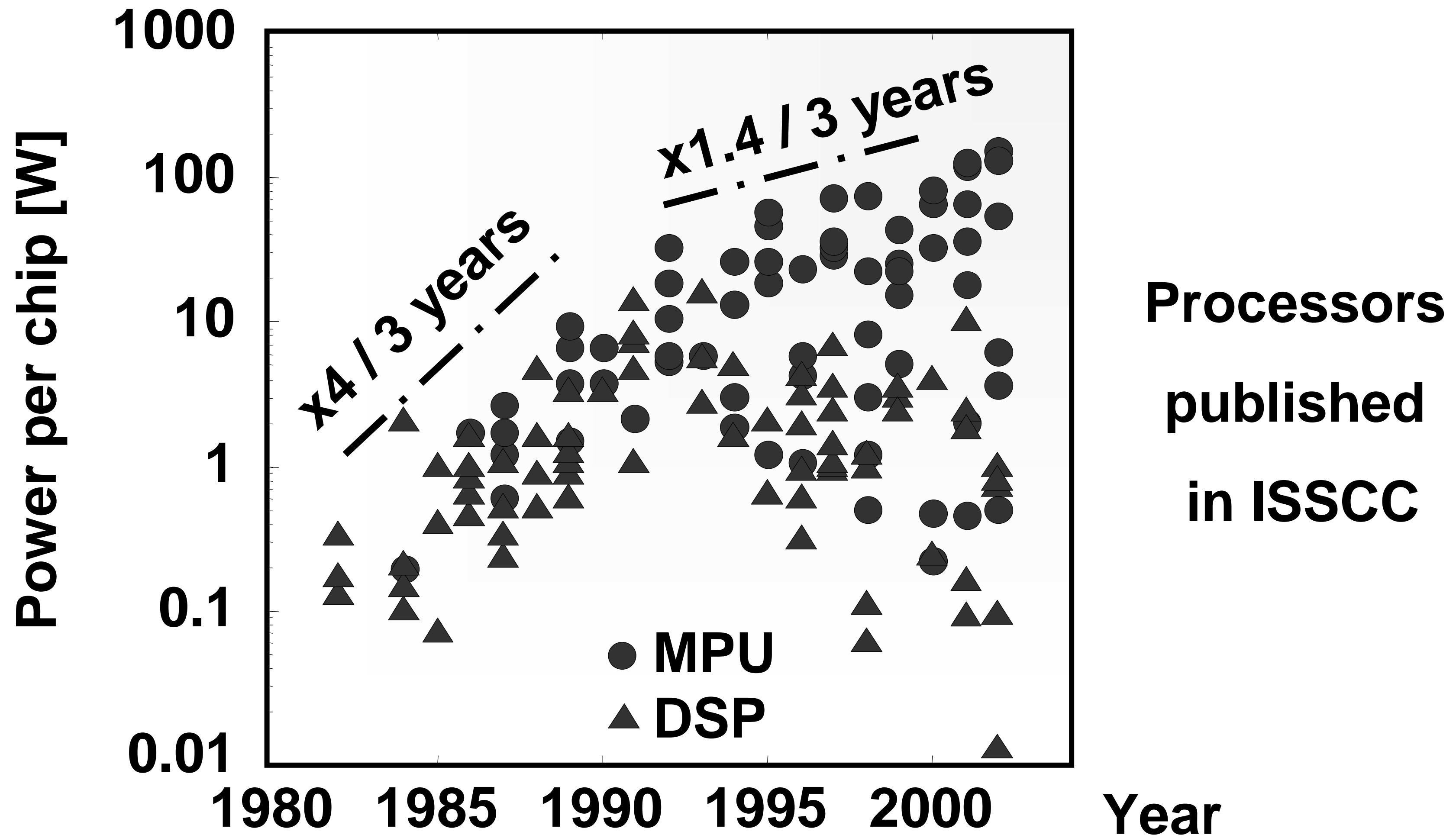
MPU Clock Cycle Trend (FO4 Delays)



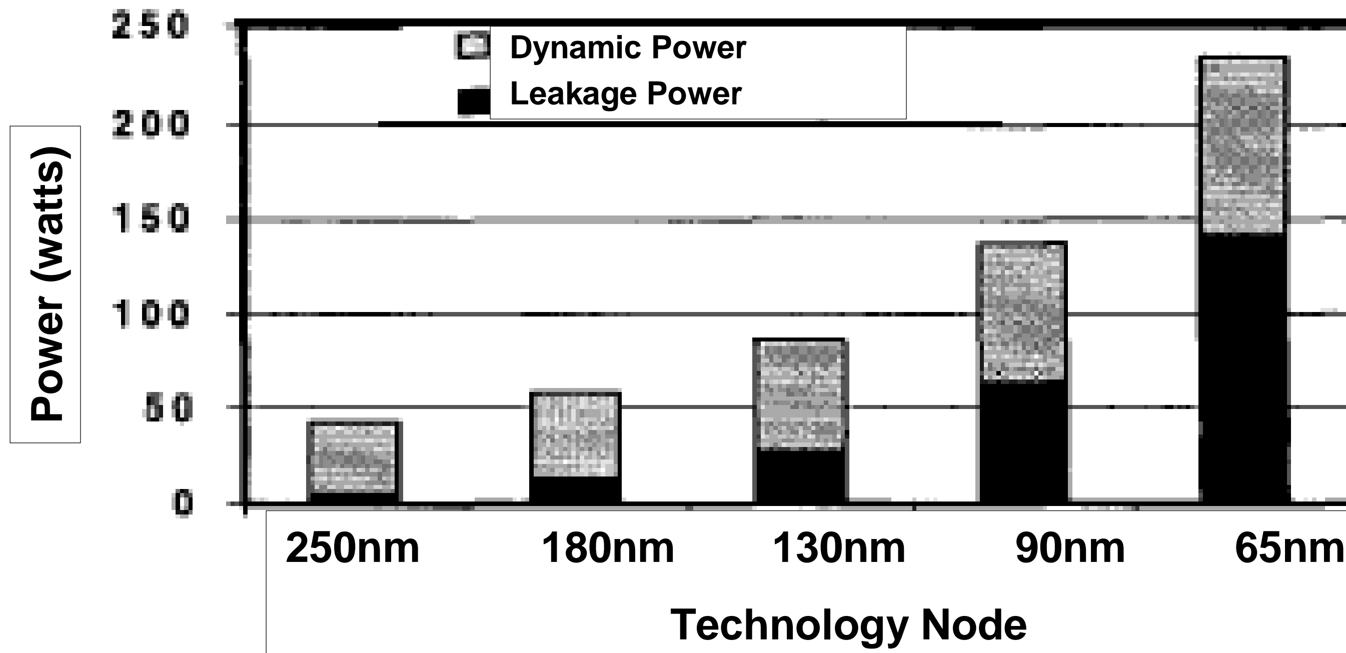
MPU Clock Cycle Trend (FO4 Delays)



Power Trend - Ever Increasing

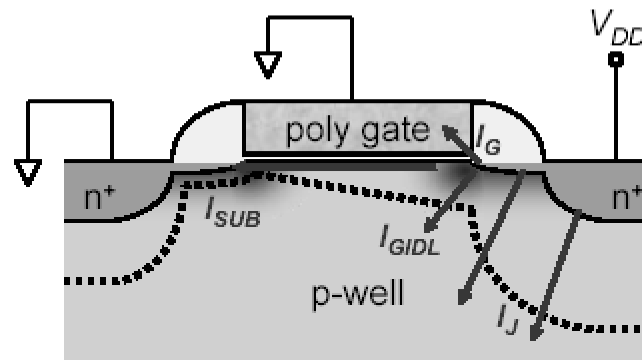


Dynamic vs. Leakage Power

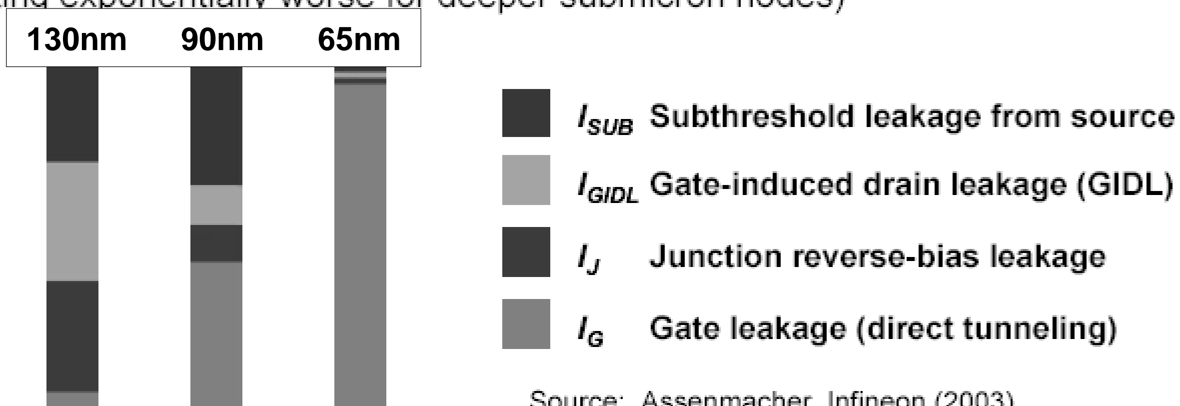


Krishnamurthy, et al., CICC 2002

Leakage Current Contributions



- Relative contributions of OFF-state leakage (but magnitude of total leakage getting exponentially worse for deeper submicron nodes)



MPU Diminishing Returns

- Power knob running out
  - Speed == Power
  - 10W/cm^2 limit for convection cooling, 50W/cm^2 limit for forced-air cooling
  - Large currents, large power surges on wakeup
  - Cf. 125A supply current, 150W total power at 1.2V Vdd for EV8 (Compaq)
  - die size will not continue to increase unless more memory is used to occupy the additional area
  - additional power dissipation coming from subthreshold leakage

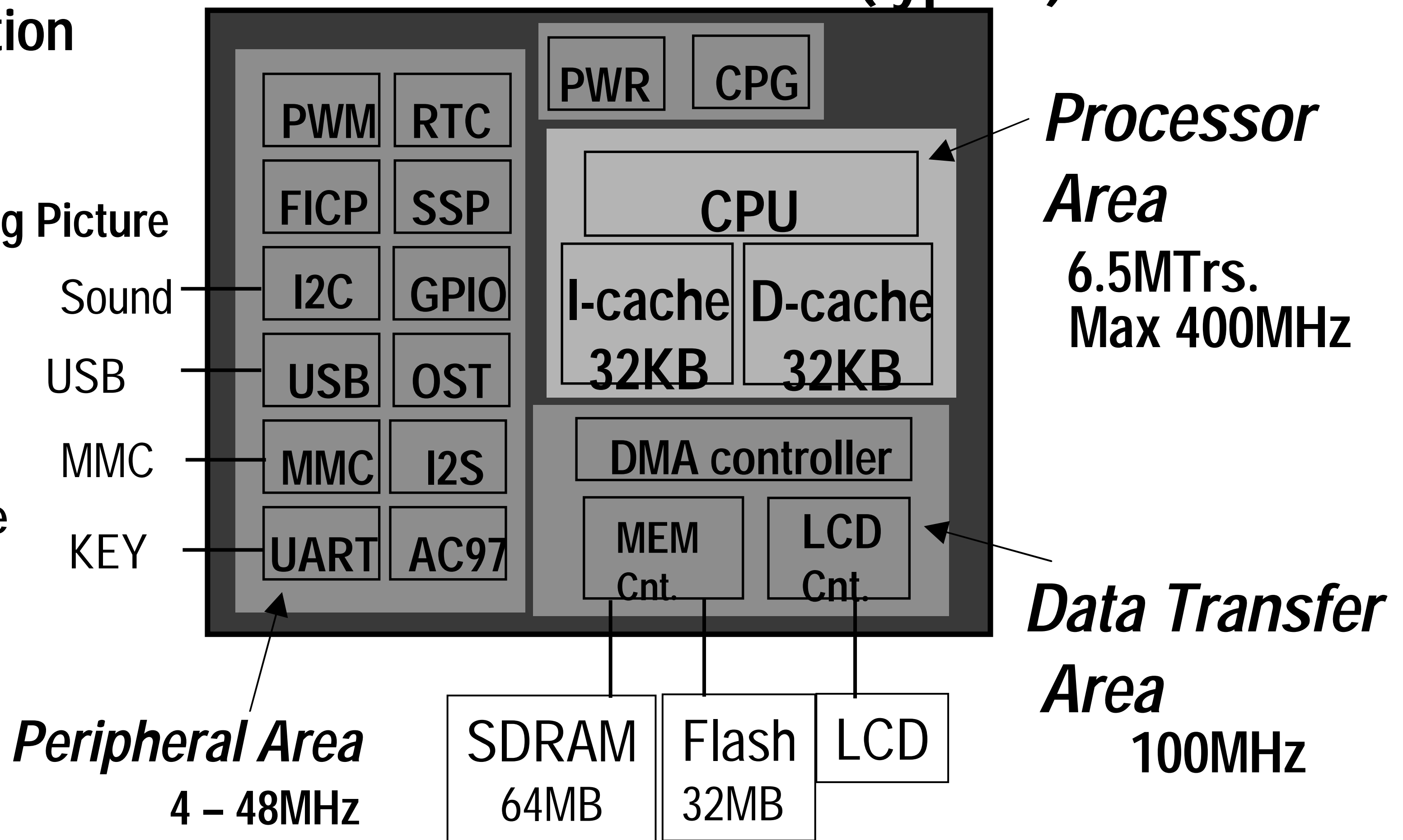
- Speed knob running out
  - Historically, 2x clock frequency every process generation
    - 1.4x from device scaling
    - 1.4x from pipelining, hence fewer logic stages (from 40-100 down to around 16 FO4 INV delays)
  - Clocks cannot be generated with period < 6-8 FO4 INV delays
  - Around 14-16 FO4 INV delays is limit for clock period

Unrealistic to continue 2x frequency trend!

Low-Power Design Techniques

- Supply Voltage Scaling
- Frequency Scaling
- Multiple Supply Voltages (Voltage Islands)
- Clock Gating
- Power Gating
- Multiple Threshold Voltages: LVT, SVT, HVT
- Substrate Biasing
- Power Shut Off
- HW/SW Power Management

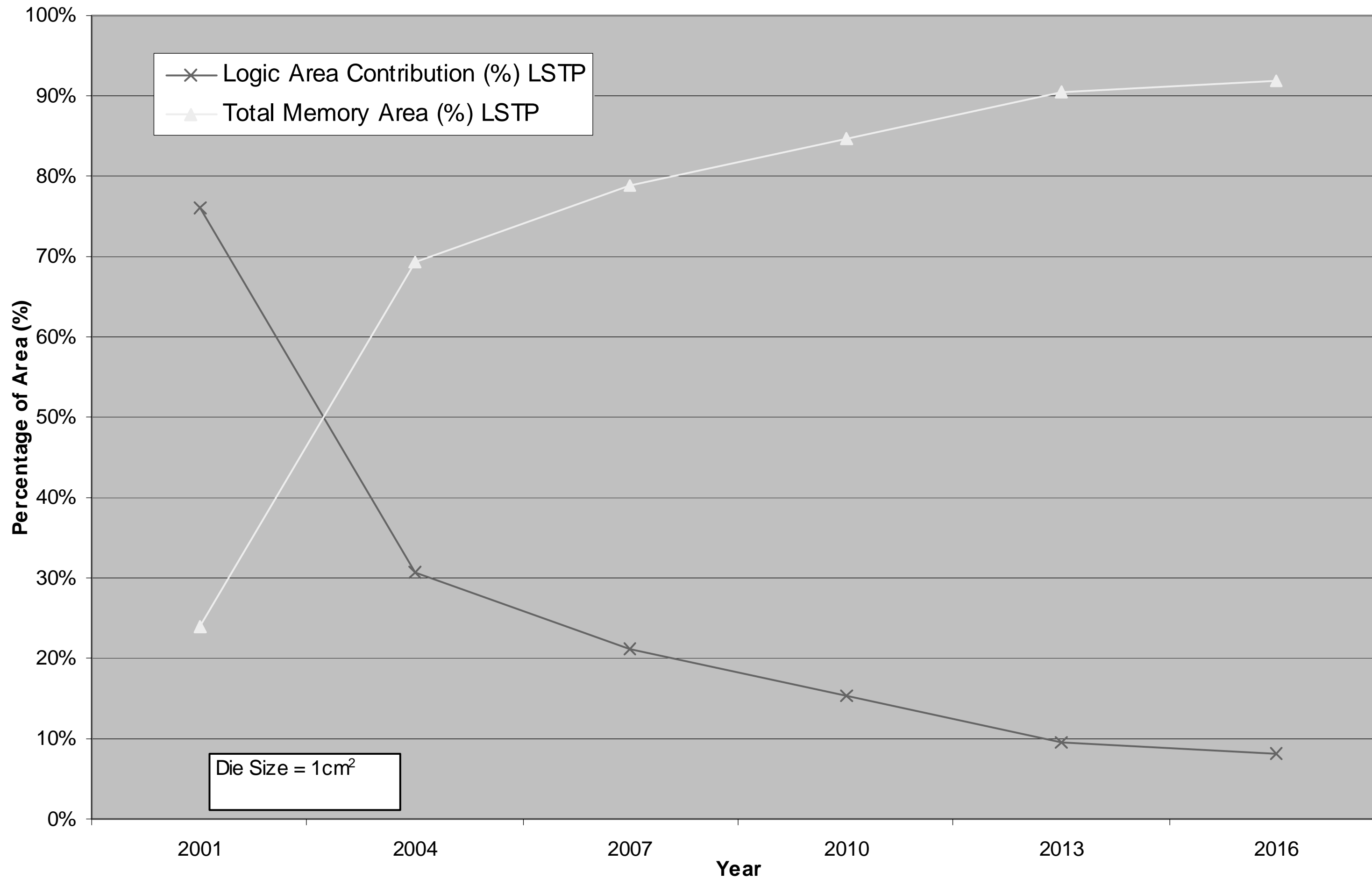
Low-Power Application: PDA



Trends in Low-Power Design Content

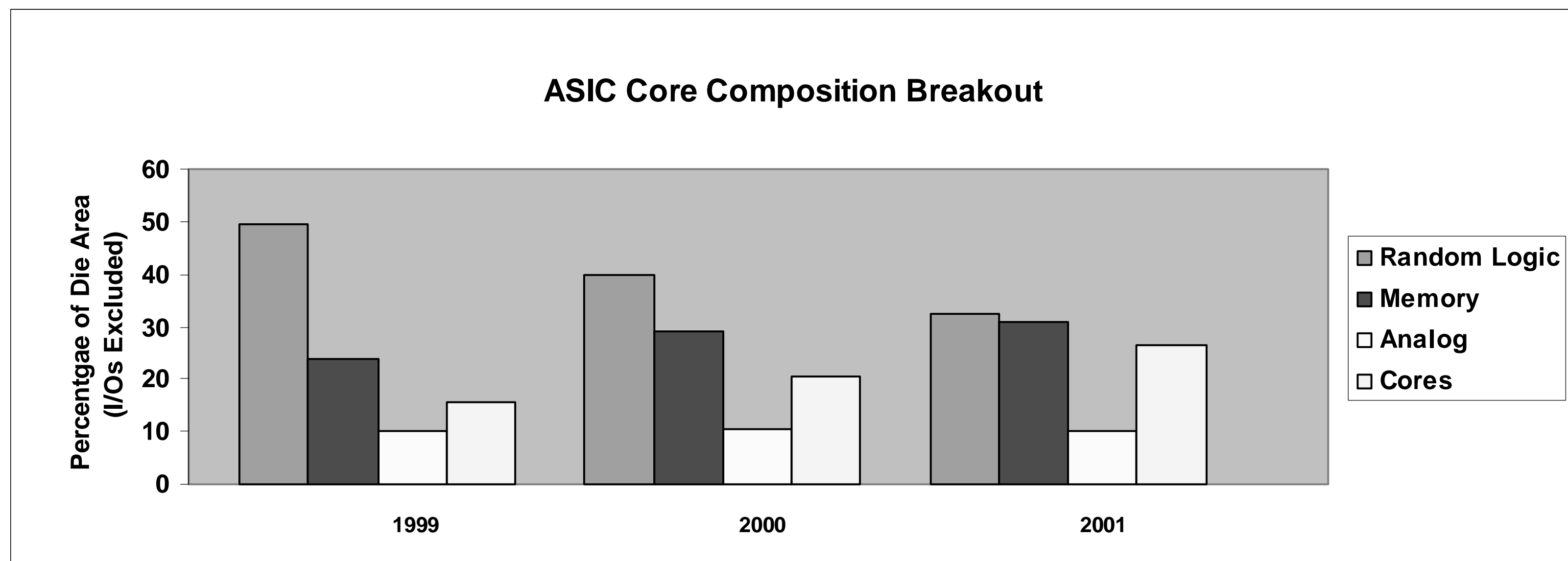
- Today, such designs contain embedded processing engines such as CPU and DSP, and memory blocks such as SRAM and embedded DRAM
- As we scale technology and keep power constant how does the amount of logic vs. memory change?
- Consider the following assumptions to develop trends for on-chip logic/memory percentages
- Die size is $100mm^2$
- Clock frequency starts at 150MHz increases by about 40% per technology node
- Average power dissipation in limited to 100mW at 100^oC
- Initial condition at Year 2001: area percentage 75% logic, 25% memory

Logic/Memory Content Trend

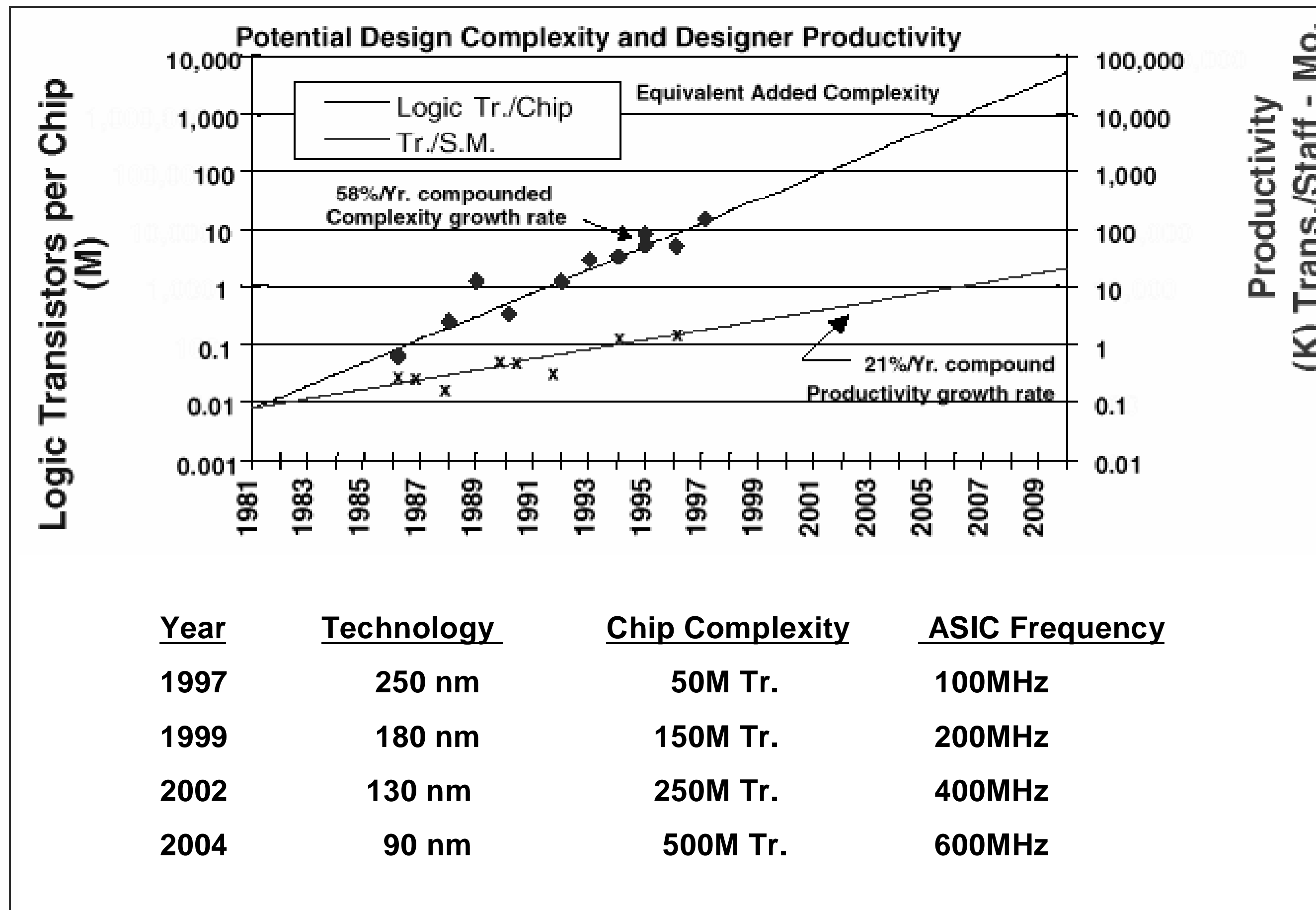


ASIC Logic/Memory Content Trends

- Source: Dataquest (2001)



Design Trend: Productivity Gap



Year	Technology	Chip Complexity	ASIC Frequency
1997	250 nm	50M Tr.	100MHz
1999	180 nm	150M Tr.	200MHz
2002	130 nm	250M Tr.	400MHz
2004	90 nm	500M Tr.	600MHz

Designing a 50M Transistor IC

• Gates Required	~12.5M
• Gates/Day (Verified)	1K (including memory)
• Total Eng. Days	12,500
• Total Eng. Years	35
• Cost/Eng./Year	$200K
• Total People Cost	$7M
• Other costs (masks, tools, etc.)	$8M

Actual Cost is $10-15M to get actual prototypes after fabrication.

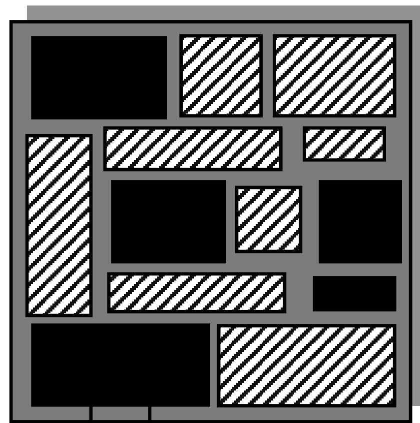
Productivity Gap

- Deep submicron (DSM) technology allows hundreds of millions of transistors to be integrated on a single chip
- Number of transistors that a designer can design per day (~1000 gates/day) is not going up significantly
- New design methodologies are needed to address the integration/productivity issues

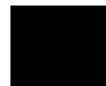
⇒ "System on a chip" Design with reusable IP

- new design methodology, IP development
- new HW/SW design and verification issues
- new test issues

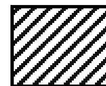
SoC Design Hierarchy



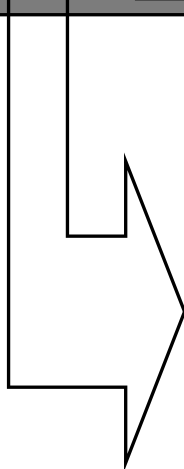
SOC consists of new logic blocks and existing IP

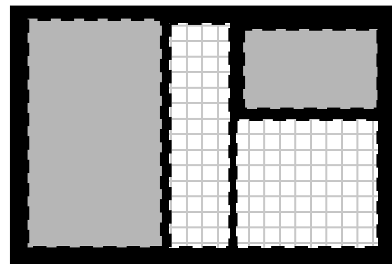


New Logic blocks

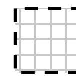


Existing IP including memory





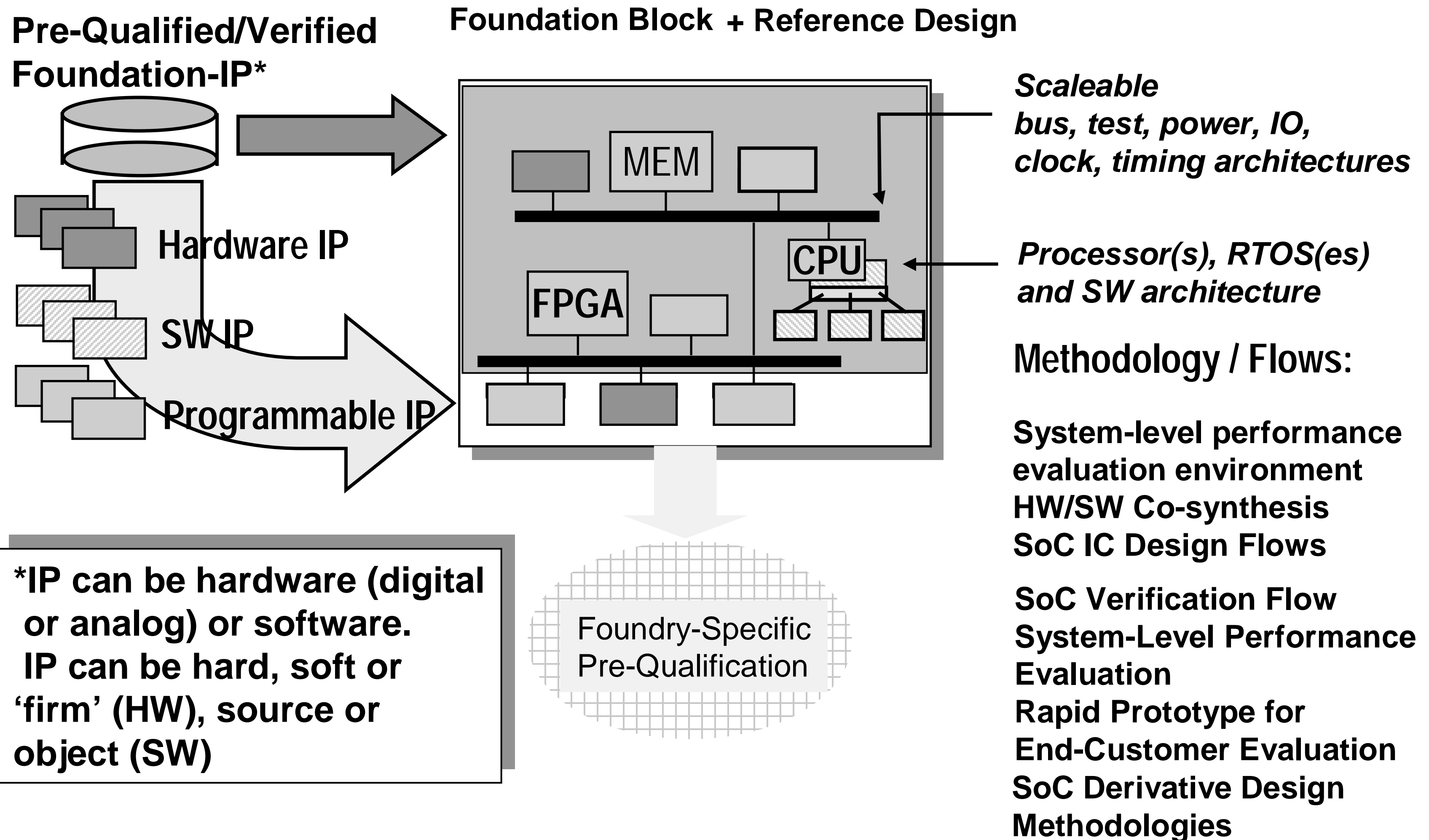
Each logic block can be implemented by newly designed portion and a re-use portion based on IPs



Newly designed portion

Re-use portion including memory

SoC Platform Design Concept



Purpose of this Course

- This course addresses SoC/IP design in DSM technologies
- It is a very broad subject, one that industry is grappling with on a daily basis – one course cannot address all the issue properly
- The goal is to present an overview of the various issues from "Systems to Silicon" to provide a perspective on what is happening in technology and design.
- We will begin with the Systems Level and work our way down to the Silicon Level
- The projects, presentations, and assignments will provide in-depth analysis of the subjects that are of interest to you