Technology Trends and Developments

ECE 697J

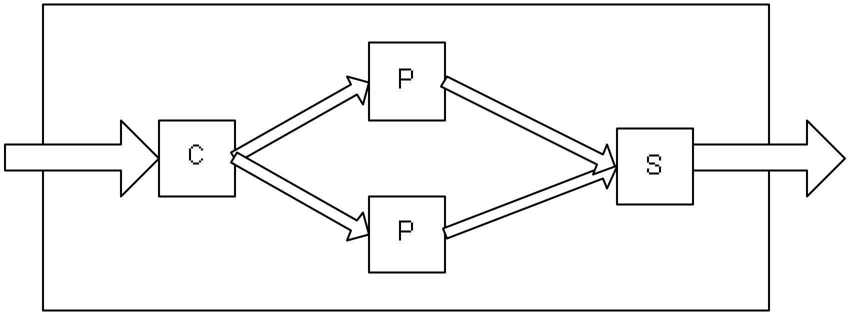
November 7^{th}, 2002

Parallelism is Key

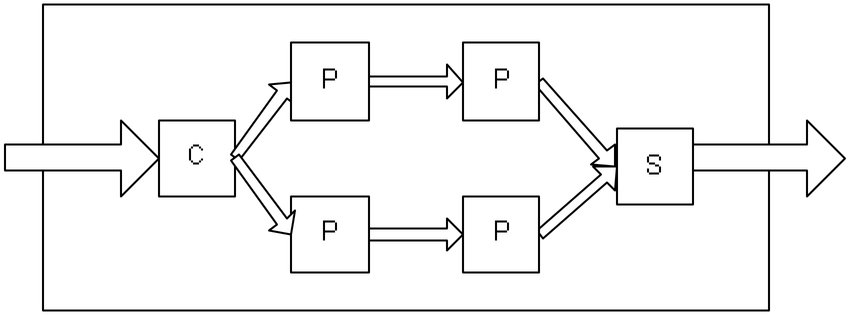
- Workstation processor workloads:
  - Few complex tasks
  - Hard to parallelize (ILP)
  - Limited speedup possible
- Network processor workloads:
  - Many parallel processing tasks
  - More opportunities for parallelism (CMP, multithreading, ILP)
  - Much speedup possible
- Why?
  - Packets can be processed independently (because of IP)
  - Maybe some intra-flow dependencies

Exploiting Parallelism

- Multiple processors in parallel:



- Plus pipelining:



Architecture Implications

- Many parallel processors
- Many processors for functional pipelining
- What are the limitations?

Limitations

- Number of parallel flows
- Overhead for communication between processors
  - Interconnect that distributes packets
- Centralized components
  - Packet classifier
  - Queuing system
- Shared data structures
  - Routing tables
- Chip size
  - Cost
  - Power consumption
- Technical feasibility

Technology Trends

- Relevant technologies for network processors
  - Link speed
  - CMOS feature size (density)
  - Maximum chip size
  - Clock speed
  - Memory technologies
  - Application complexity
- Moore’s Law:
  - “Number of components on chip doubles every 18 months”

Moore’s Law 1965

The experts look ahead

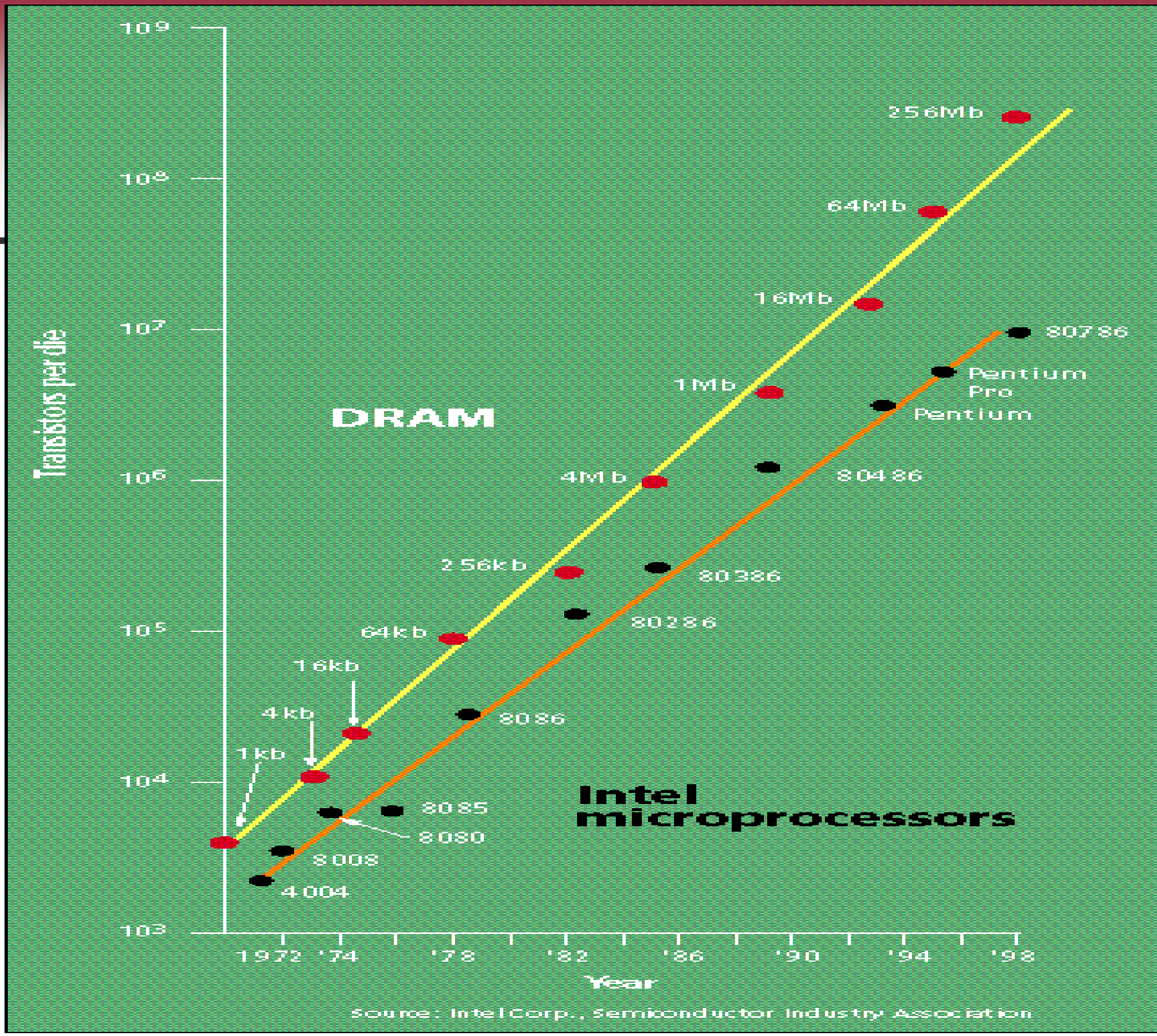
Cramming more components onto integrated circuits

With unit cost falling as the number of components per circuit rises, by 1975 economics may dictate squeezing as many as 65,000 components on a single silicon chip

By Gordon E. Moore

Director, Research and Development Laboratories, Fairchild Semiconductor division of Fairchild Camera and Instrument Corp.

Humor in Moore's Paper



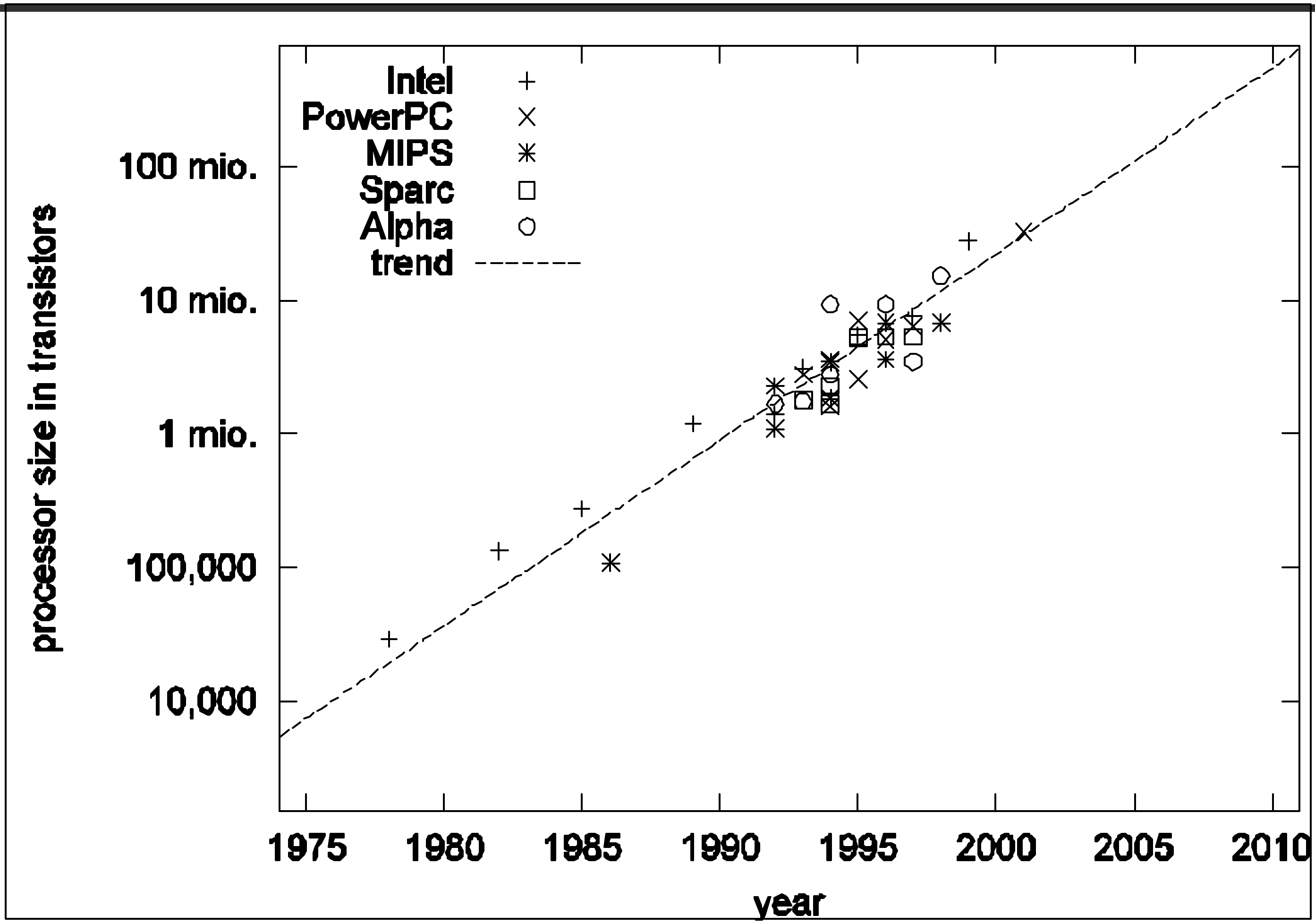
IEEE Spectrum, 1997

University of Massachusetts Amherst

Long-Term Trends

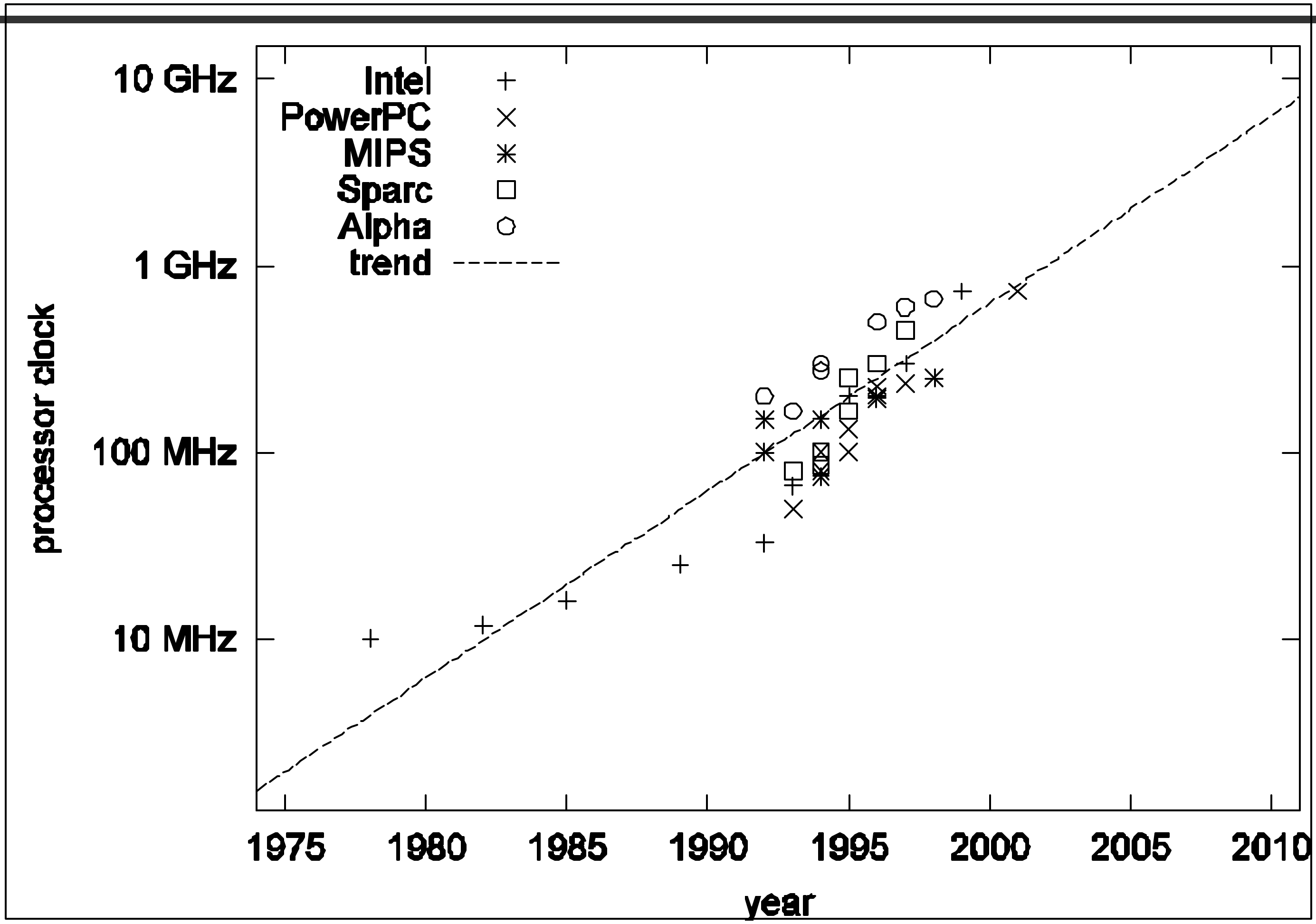
- Moore's Law is pretty "stable"
  - Of course it's not a "law"
- Will probably continue until end of decade
  - Semiconductor Industry Association's roadmap
- Let's look at individual metrics

Workstation Processor Size

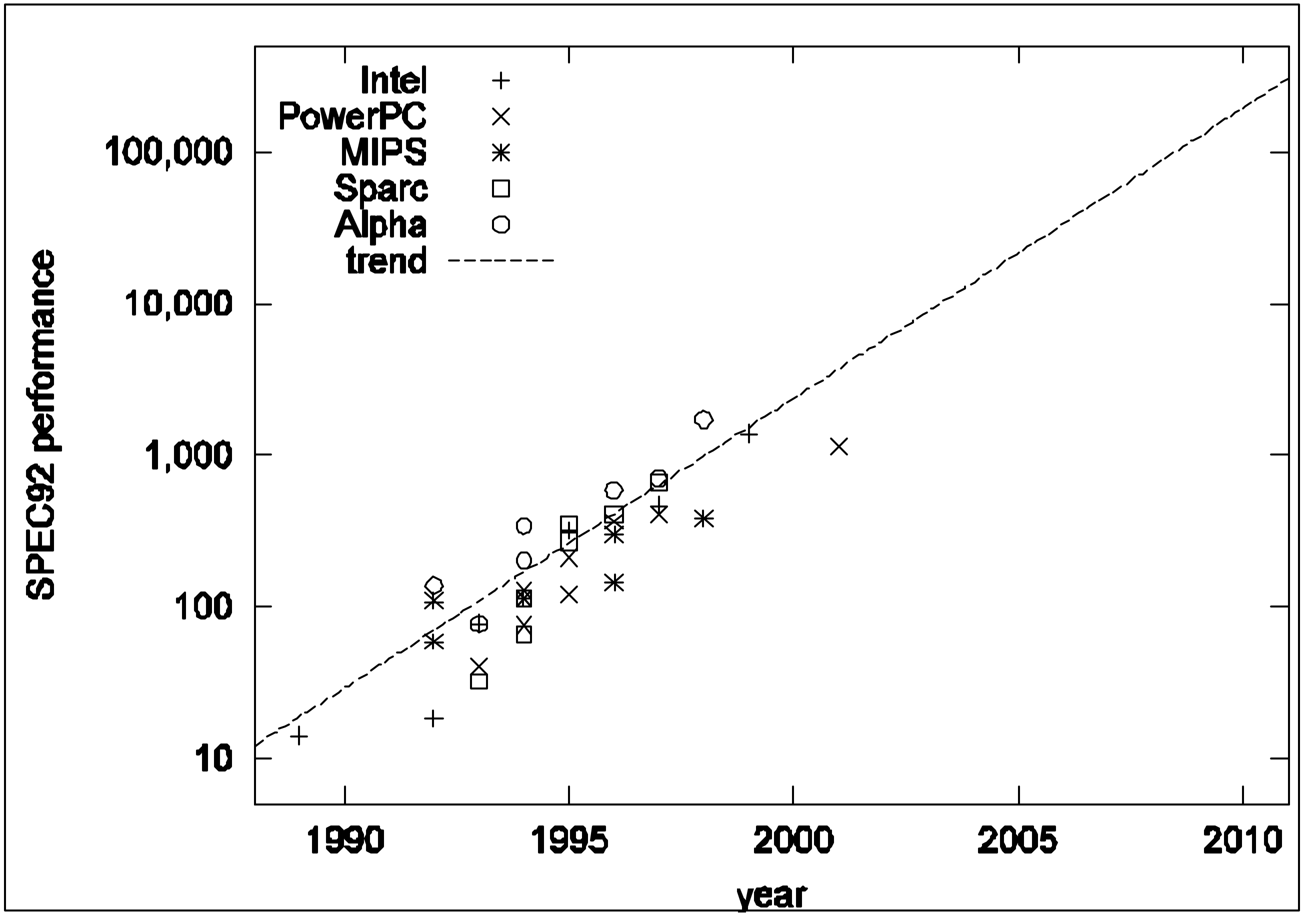


University of Massachusetts Amherst

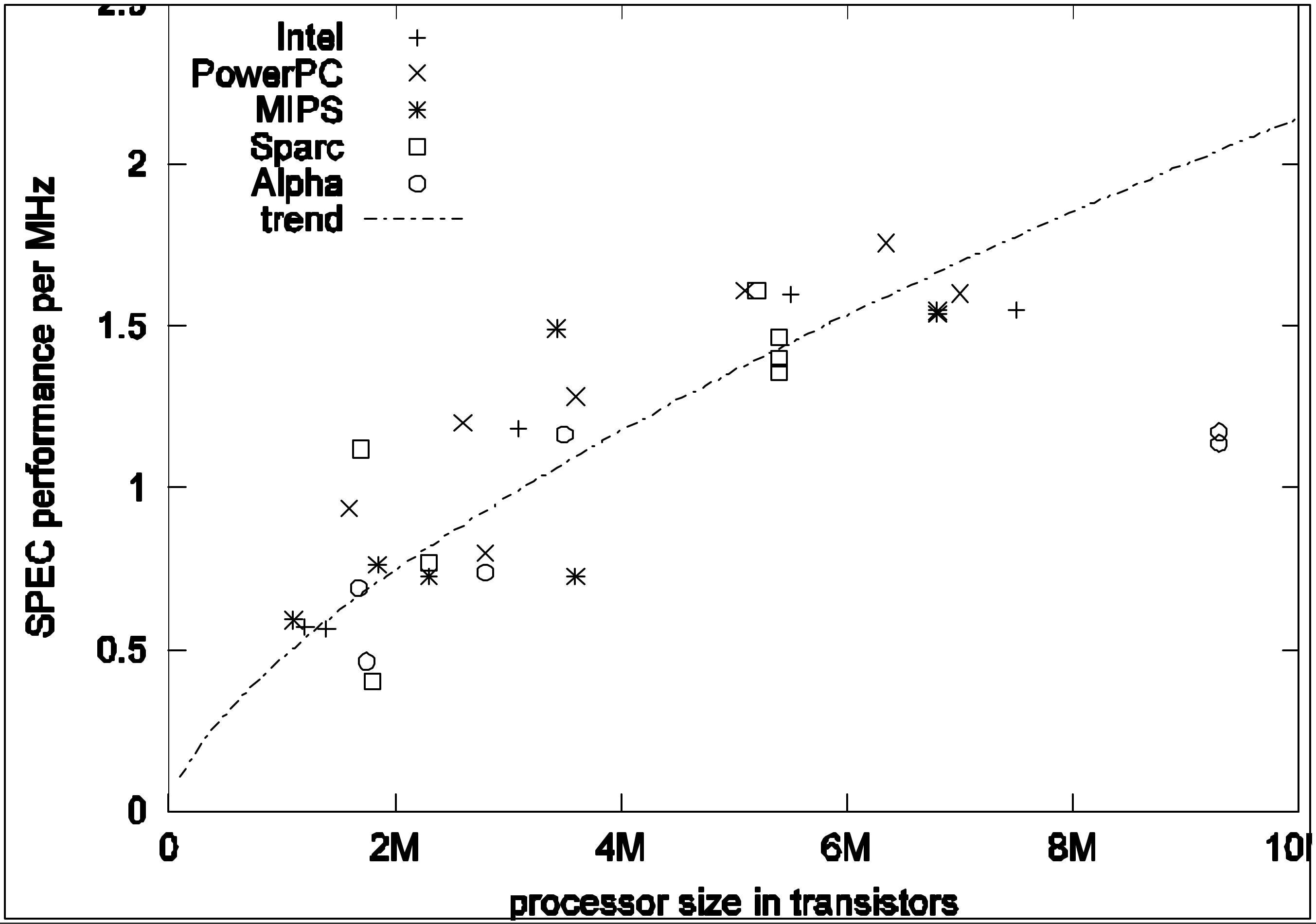
Processor Clock Rate



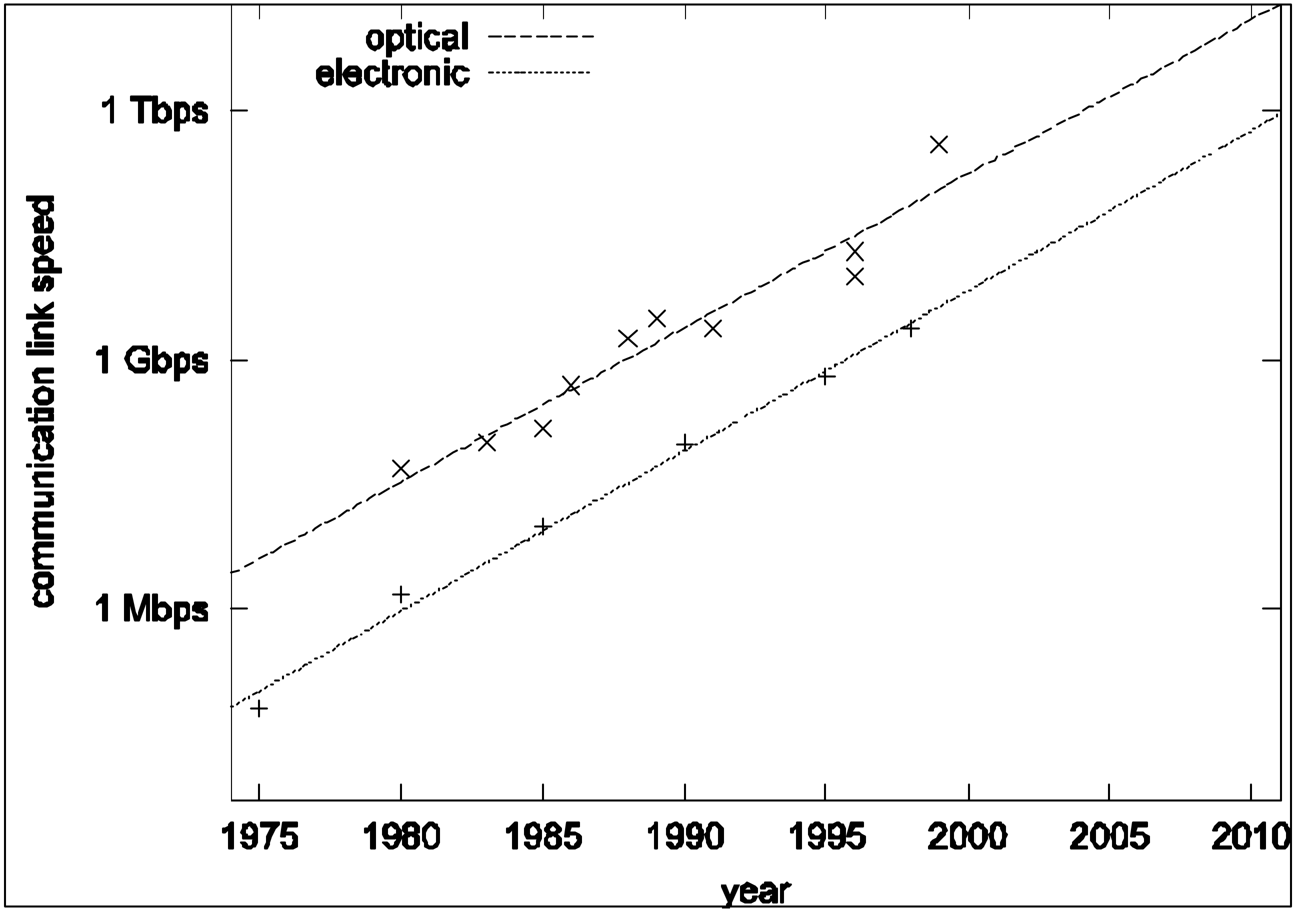
SPEC Performance



Performance vs. Size



Link Speed


optical
electronic
1 Tbps
1 Gbps
1 Mbps
communication link speed
1975
1980
1985
1990
1995
2000
2005
2010
year

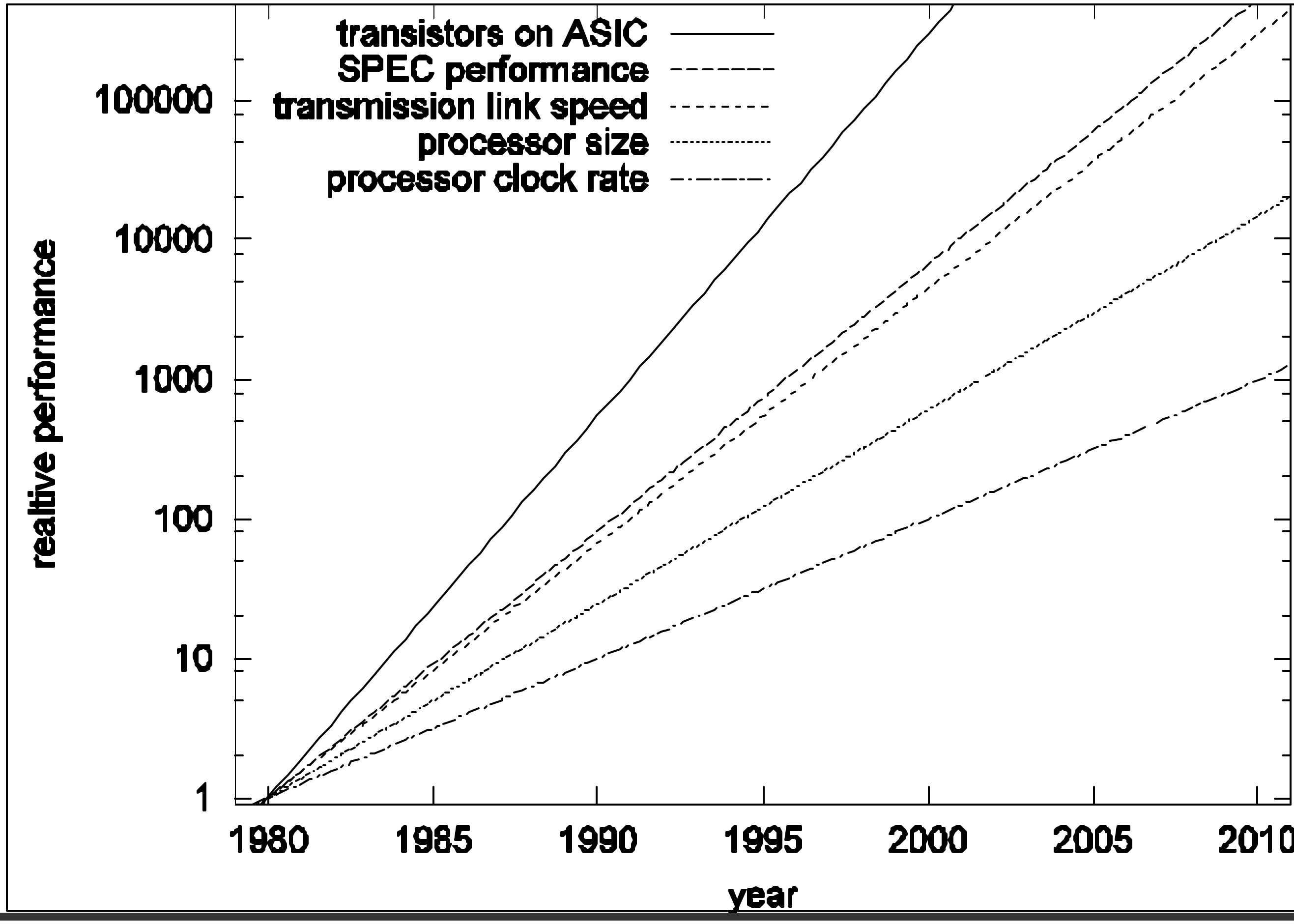
Comparison of Trends

$$\text{performance in year } x = a_{t_0} \cdot e^{b \cdot (x - t_0)}$$

Table 2.1: Growth Parameters of Key Technologies. The values for a are normalized to the year 2000. Sizes are given in million transistors (Mtx) and million gates (Mgates) (1 gate $\approx$ 4 transistors).

Technology		a	b	Time to double
Communication	electronic links	6.8 Gbps	0.44	18 months
	optical links	175 Gbps	0.42	19 months
Processor	SPEC performance	2400	0.44	18 months
	clock	630 MHz	0.23	36 months
	size	22 Mtx	0.32	26 months
ASIC	size	55 Mgates	0.63	8 months

Comparison of Trends



What about Network Processors?

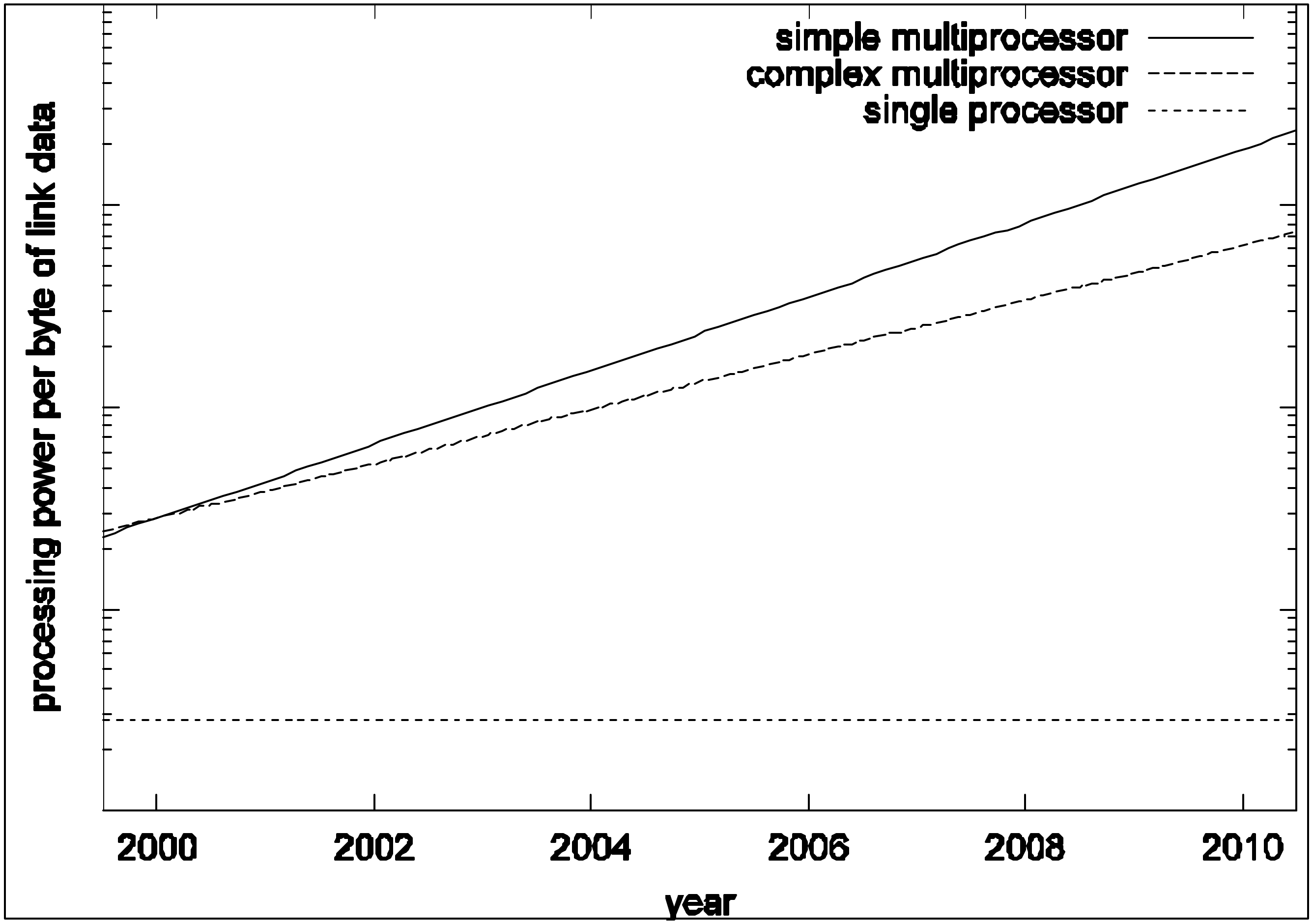
- How can we use these trends for scalable designs?

Possible Architectures

- Arch 1: single CPU
- Arch 2: CMP with high-performance processors
- Arch 3: CMP with low-performance processors

- What is the performance criteria?
  - How much processing for each packet
  - Measured in SPEC per byte of link data

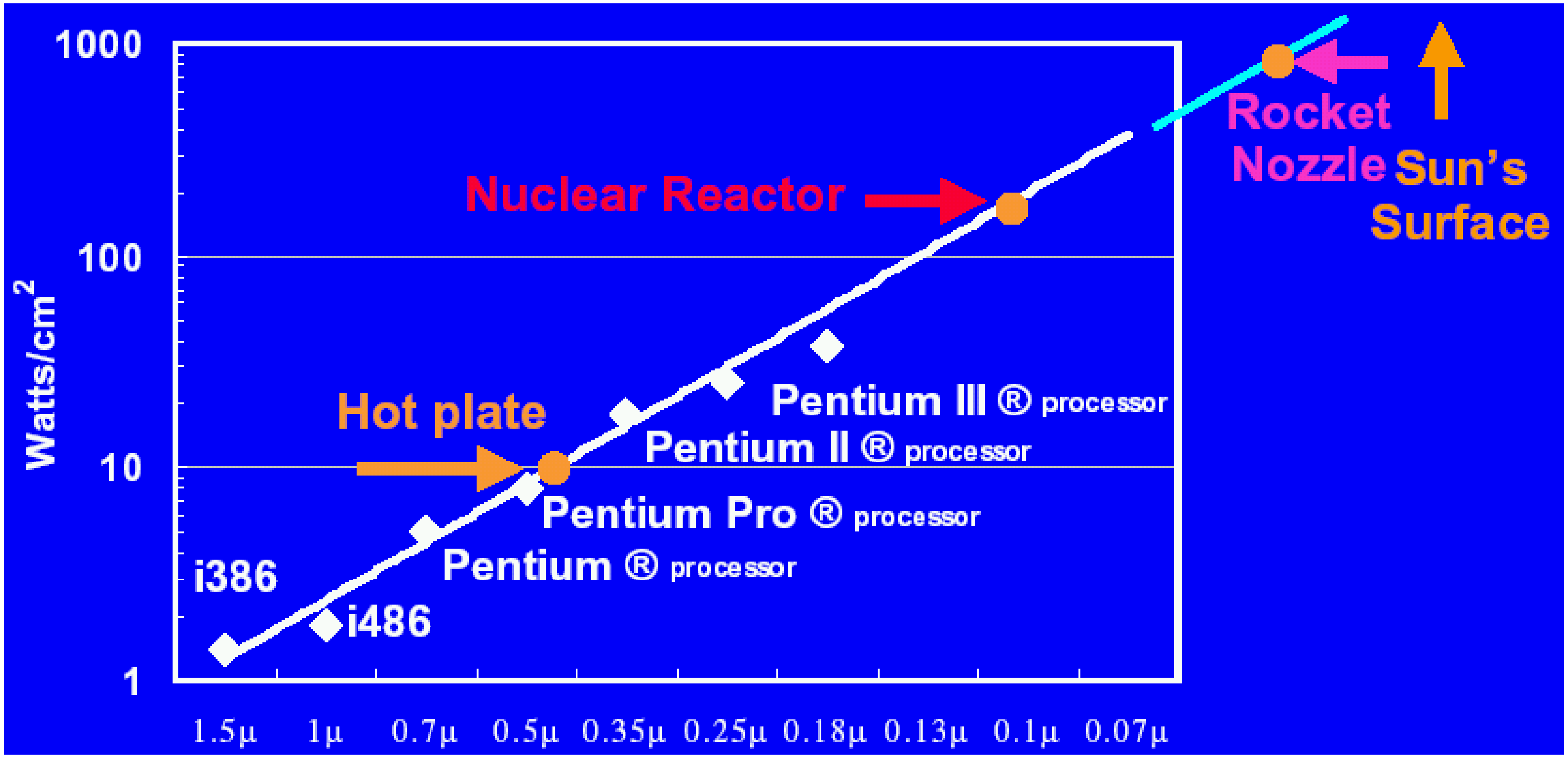
Architecture Performance Trends



Limitations

- What are the limits to these trends?
  - Bottleneck in centralized components
  - Memory gap
  - Power consumption
  - Power density

Power Density



by Fred Pollack

Summary

- Networking workloads are highly parallelizable
- NPs can leverage technology trends
  – CMP with simple RISC cores
- Challenge:
  – Design architecture that can scale to use parallelism

Next Lecture

- Network processors
  - Intel IXP1200
  - IBM PowerNP
- Change class organization
  - Shorter presentations (20-25 minutes)
  - More discussion
  - What to do to ensure people are prepared?
    - Write essay on papers
    - Write down three discussion questions