Hardware Technology Trends and Database Opportunities

David A. Patterson and Kimberly K. Keeton

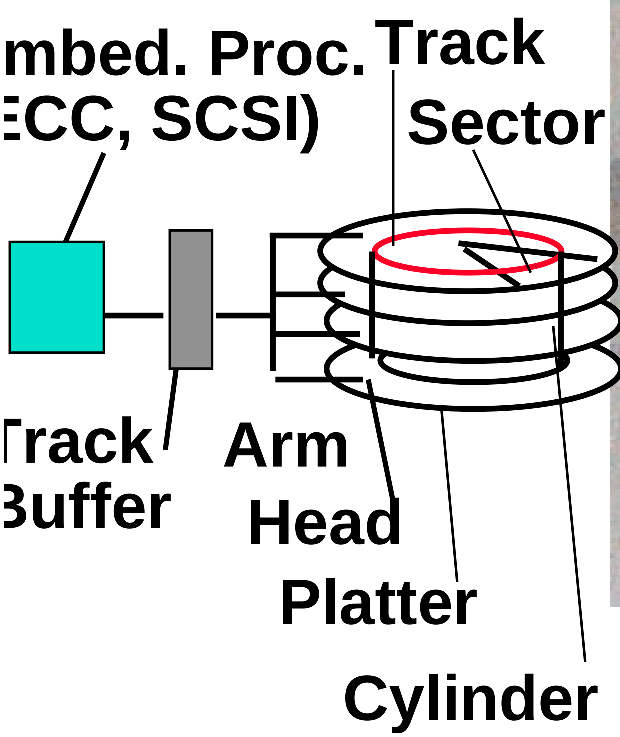
http://cs.berkeley.edu/~patterson/talks

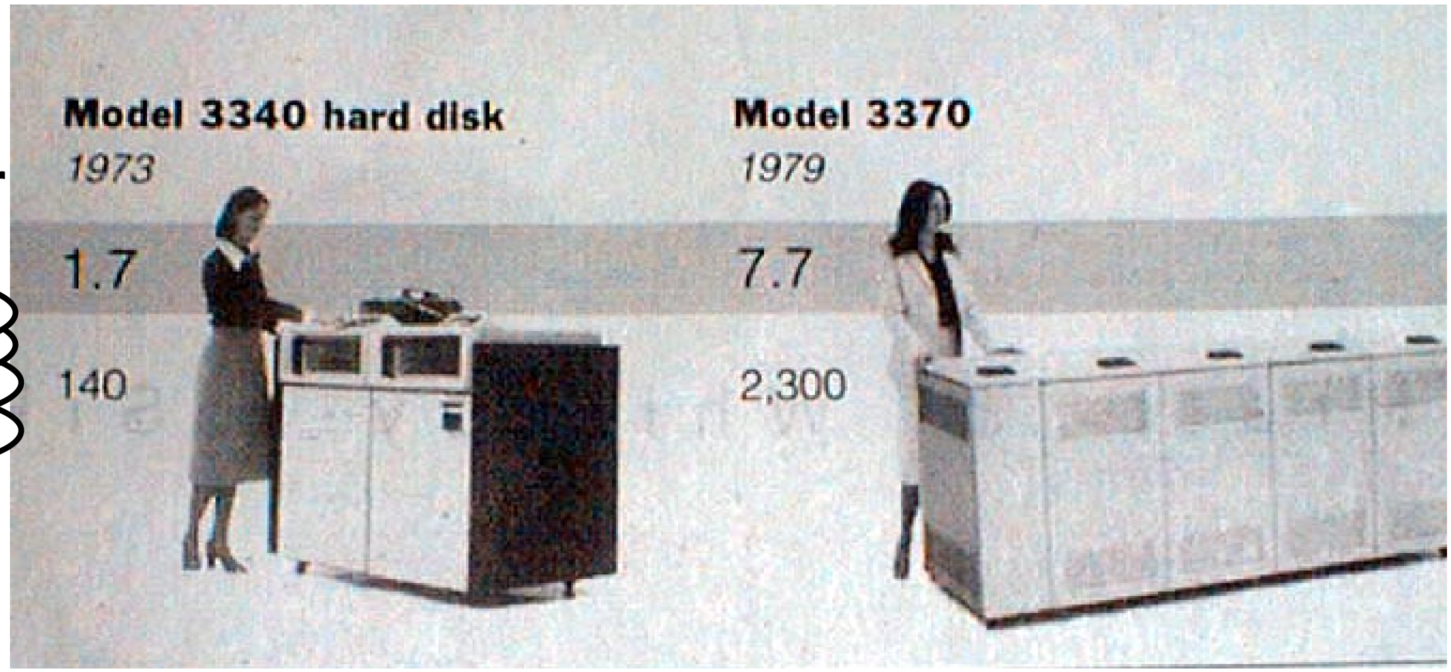
{patterson,kkeeton}@cs.berkeley.edu
EECS, University of California
Berkeley, CA 94720-1776

Outline

- Review of Five Technologies:
  Disk, Network, Memory, Processor, Systems
  - Description / History / Performance Model
  - State of the Art / Trends / Limits / Innovation
  - Following precedent: 2 Digressions
- Common Themes across Technologies
  - Perform.: per access (latency) + per byte (bandwidth)
  - Fast: Capacity, BW, Cost; Slow: Latency, Interfaces
  - Moore's Law affecting **all** chips in system
- Technologies leading to Database Opportunity?
  - Hardware & Software Alternative to Today
  - Back-of-the-envelope comparison: scan, sort, hash-join

Disk Description / History



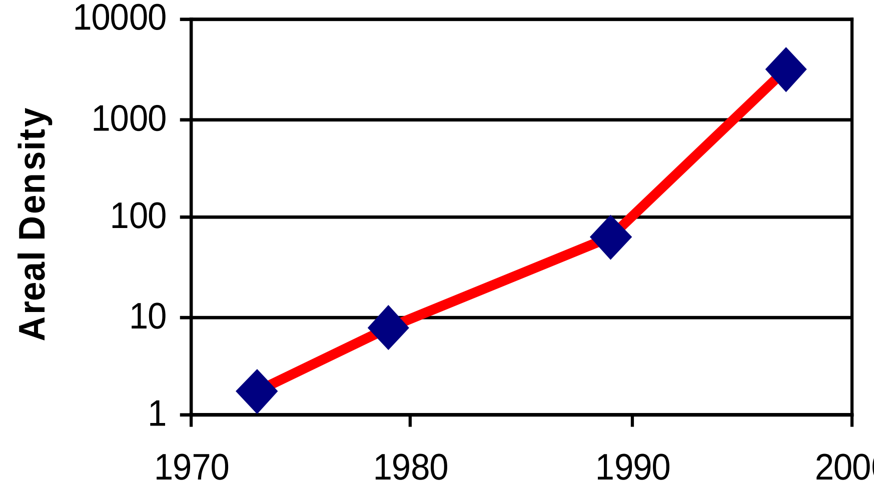


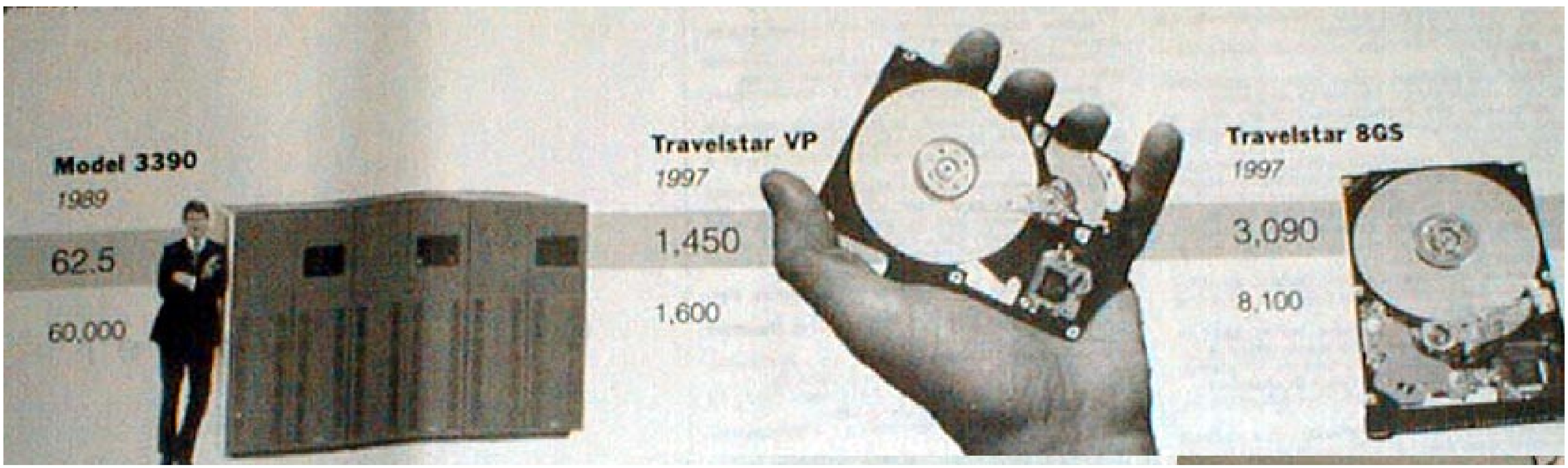
1973:
1. 7 Mbit/sq. in
140 MBytes

1979:
7. 7 Mbit/sq. in
2,300 MBytes

source: New York Times, 2/23/98, page C3,
"Makers of disk drives crowd even more data into even smaller spaces"

Disk History





1989:
63 Mbit/sq. in
60,000 MBytes

1997:
1450 Mbit/sq. in
2300 MBytes

1997:
3090 Mbit/sq. in
8100 MBytes

ource: New York Times, 2/23/98, page C3

Performance Model / Trends

Latency =

er access { Queuing Time + Controller time + Seek Time + Rotation Time

+

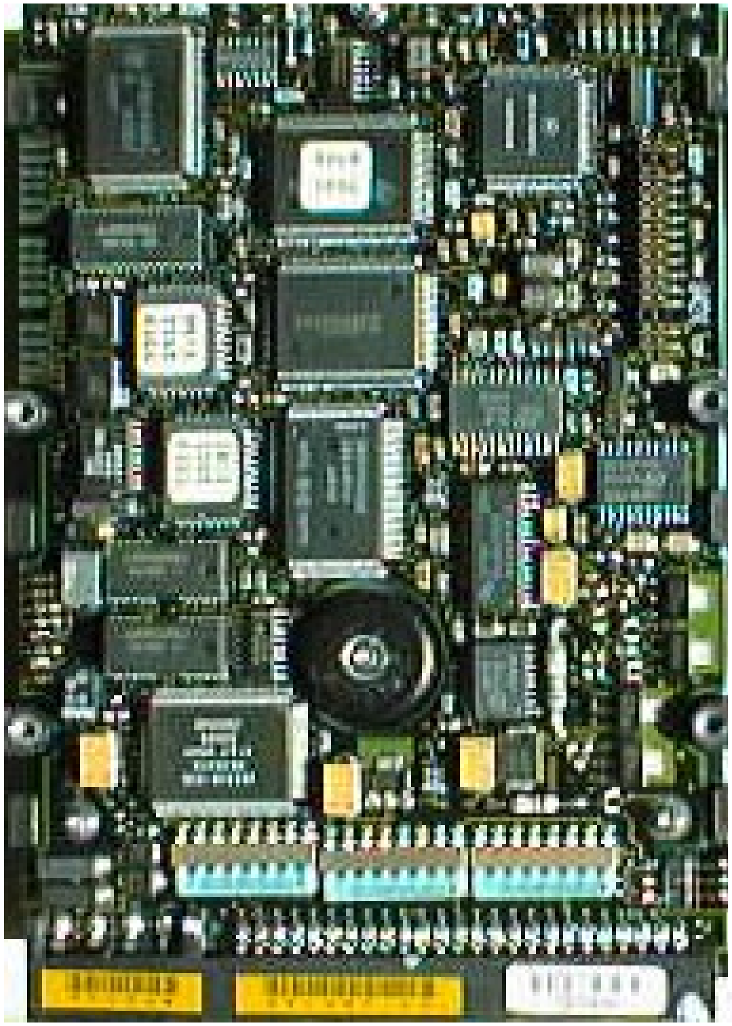
per byte + Size / Bandwidth

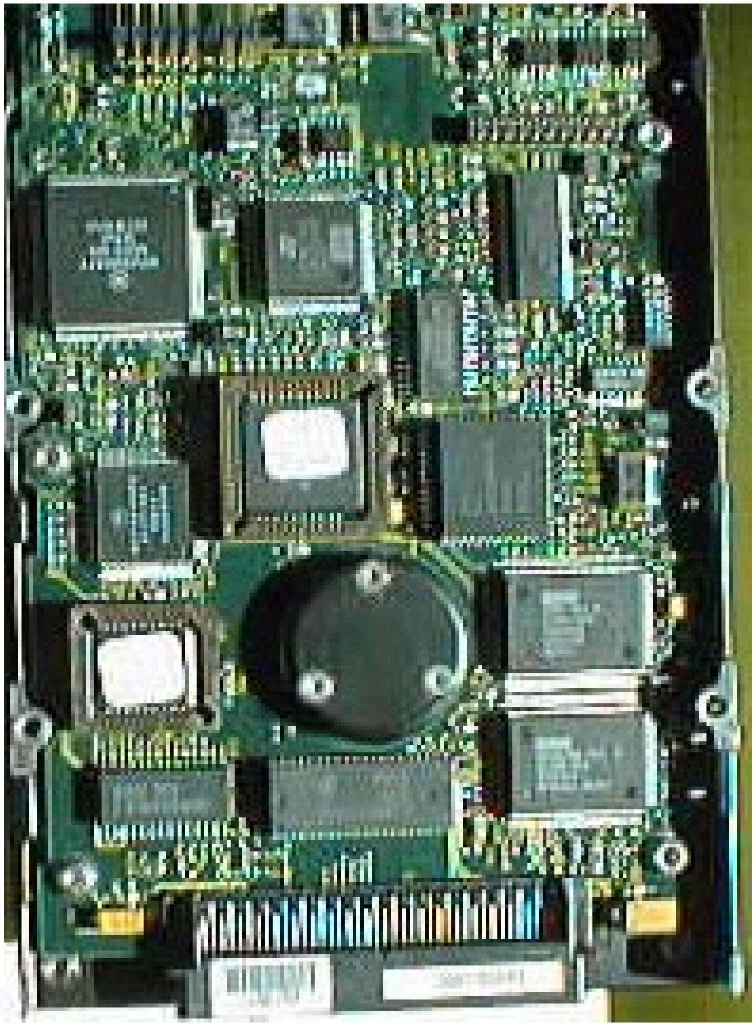
- Capacity
  - + 60%/year (2X / 1.5 yrs)
- Transfer rate (BW)
  - + 40%/year (2X / 2.0 yrs)
- Rotation + Seek time
  - – 8%/ year (1/2 in 10 yrs)
- MB/$
  - > 60%/year (2X / <1.5 yrs)
  - Fewer chips + areal densit

ource: Ed Grochowski, 1996,
BM leadership in disk drive technology";
ww.storage.ibm.com/storage/technolo/grochows/grocho01.htm,

Chips / 3.5 inch Disk: 1993 v. 1994

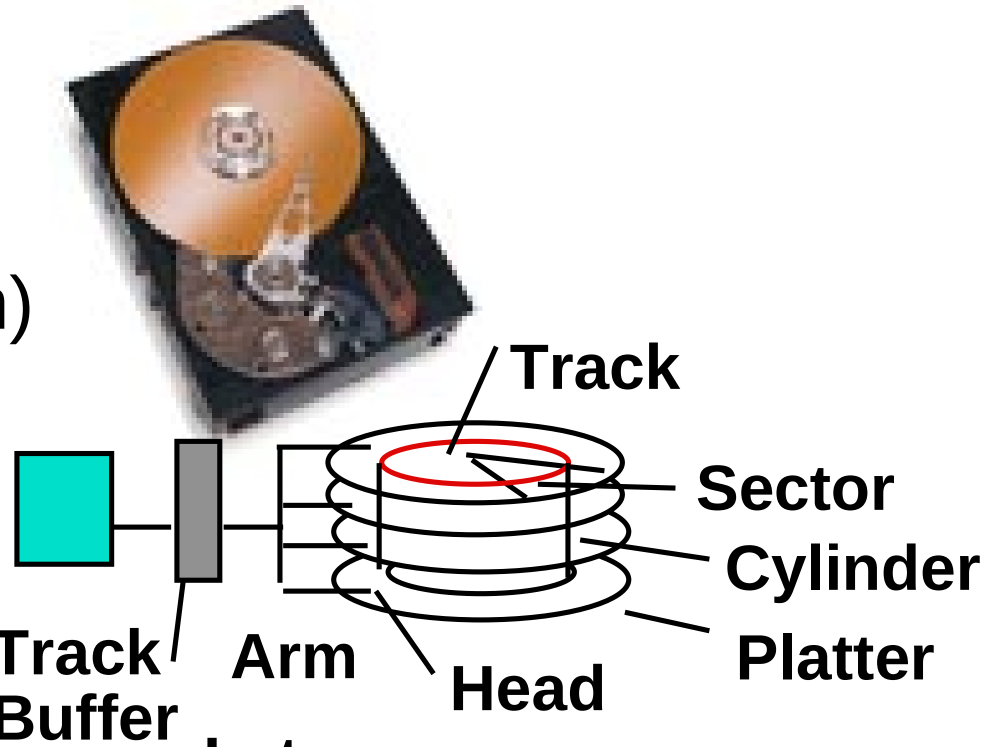
15 vs. 12 chips; 2 chips (head, misc) in 200x?





State of the Art: Seagate Cheetah 18

- 6962 cylinders, 12 platters
- 18.2 GB, 3.5 inch disk
- 1MB track buffer
  (+ 4MB optional expansion)
- 19 watts
- 0.15 ms controller time
- avg. seek = 6 ms
  (seek 1 track = 1 ms)
- 1/2 rotation = 3 ms
- 21 to 15 MB/s media
  (=> 16 to 11 MB/s)
  - » deliver 75% (ECC, gaps...)
- $1647 or 11MB/$ (9¢/MB)



Latency =

per access
Queuing Time +
Controller time +
Seek Time +
Rotation Time

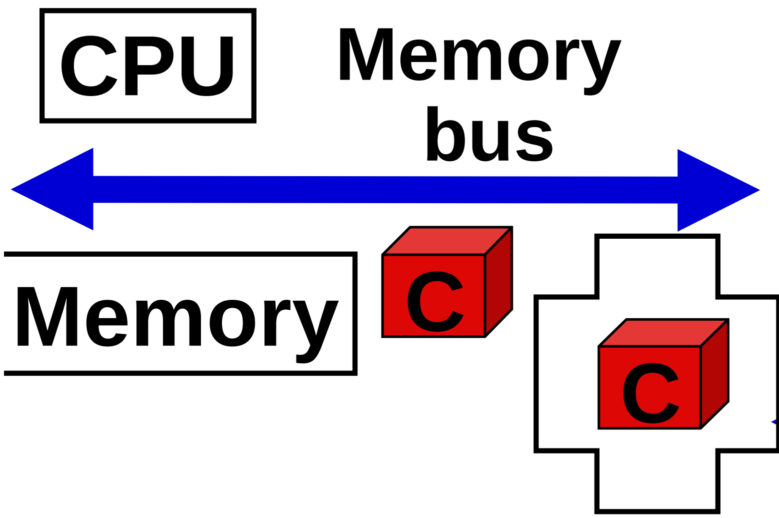
+

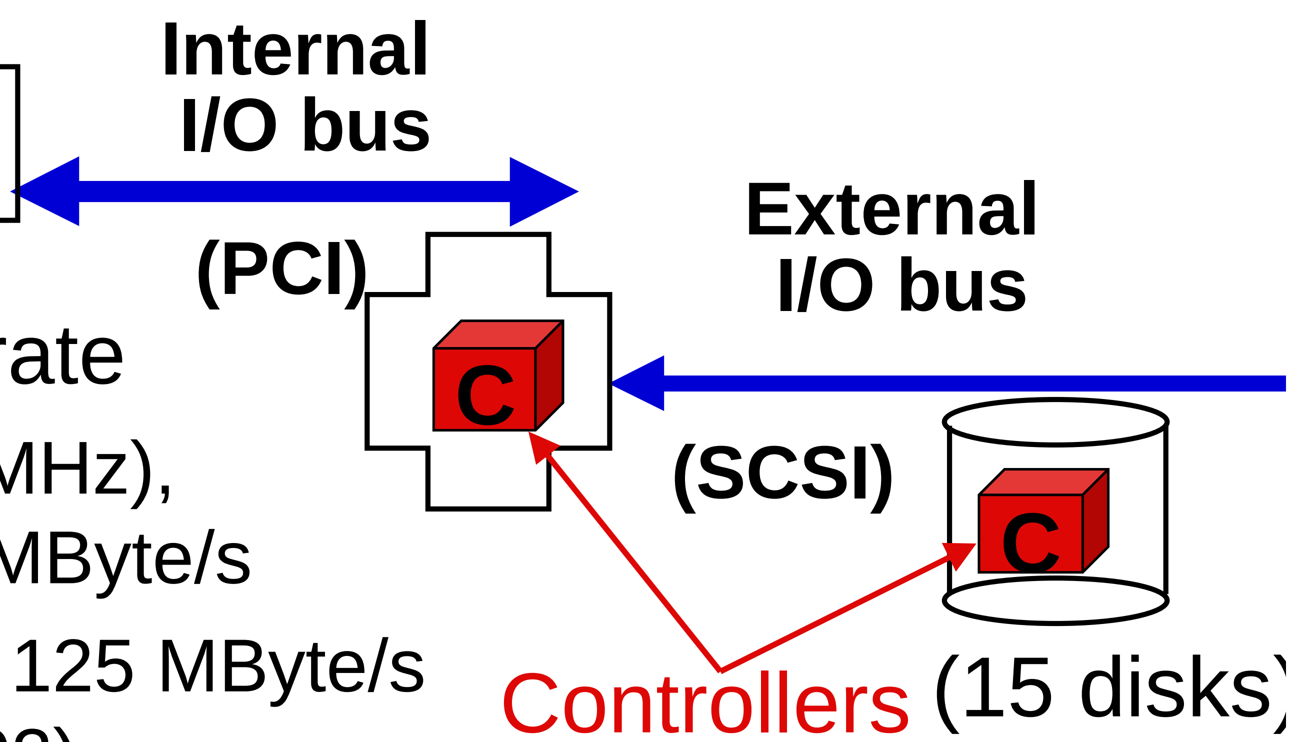
per byte
+ Size / Bandwidth

source: www.seagate.com;
www.pricewatch.com; 5/21/98

Disk Limit: I/O Buses

- Multiple copies of data, SW layers
- Cannot use 100% of bus
  - Queuing Theory (< 70%)
  - Command overhead (Effective size = size x 1.2)





- Bus rate vs. Disk rate
  - SCSI: Ultra2 (40 MHz), Wide (16 bit): 80 MByte/s
  - FC-AL: 1 Gbit/s = 125 MByte/s (single disk in 2002)

Disk Challenges / Innovations

- Cost SCSI v. EIDE:
  - $275: IBM 4.3 GB, UltraWide SCSI (40MB/s) 16MB/$
  - $176: IBM 4.3 GB, DMA/EIDE (17MB/s) 24MB/$
  - Competition, interface cost, manufact. learning curve?
- Rising Disk Intelligence
  - SCSI3, SSA, FC-AL, SMART
  - Moore's Law for embedded processors, too

source: www.research.digital.com/SRC/articles/199701/petal.html; www.pricewatch.com

Disk Limit

- Continued advance in capacity (60%/yr) and bandwidth (40%/yr.)
- Slow improvement in seek, rotation (8%/yr)
- Time to read whole disk

Year	Sequentially	Randomly
1990	4 minutes	6 hours
2000	12 minutes	1 week

- Dynamically change data layout to reduce seek, rotation delay? Leverage space vs. spindles?

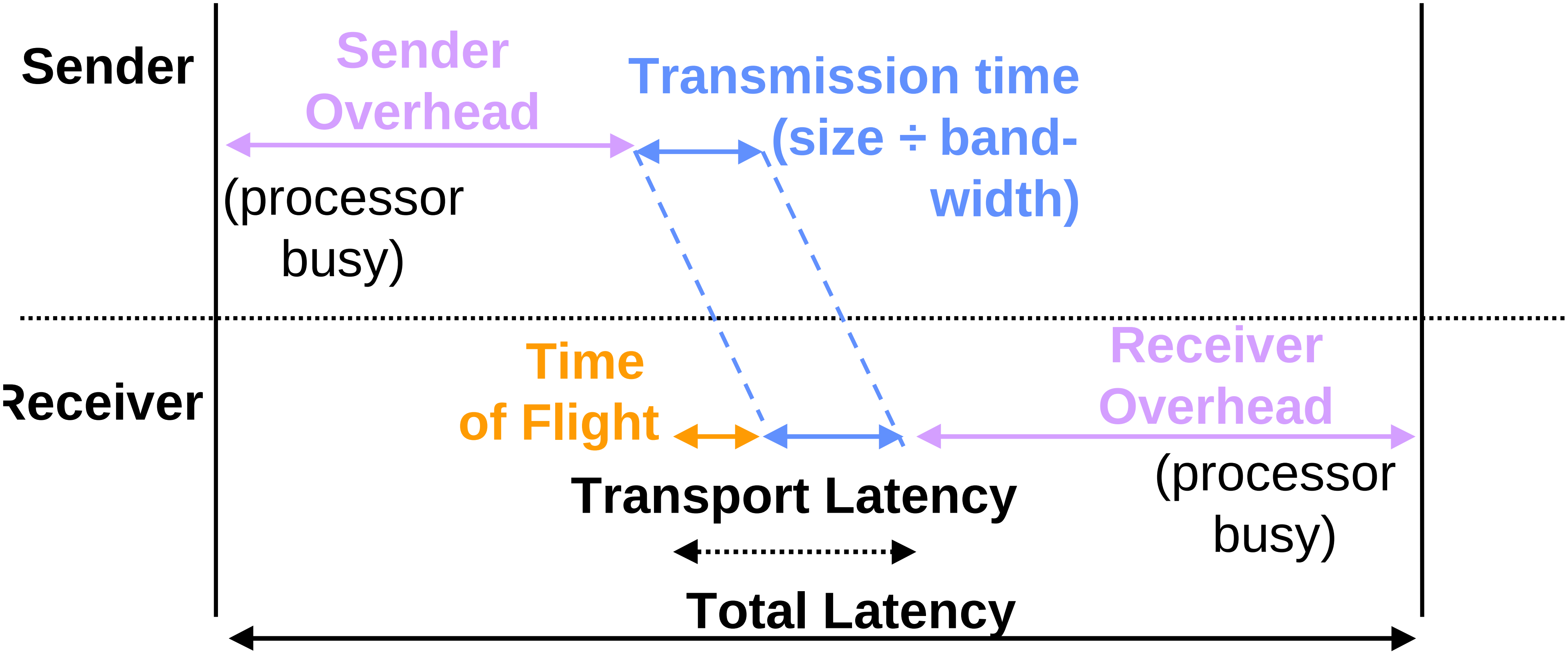
Disk Summary

- Continued advance in capacity, cost/bit, BW; slow improvement in seek, rotation
- External I/O bus bottleneck to transfer rate, cost? => move to fast serial lines (FC-AL)?
- What to do with increasing speed of embedded processor inside disk?

Network Description/Innovations

- Shared Media vs. Switched: pairs communicate at same time
- Aggregate BW in switched network is many times shared
  - point-to-point faster only single destination, simpler interface
  - Serial line: 1 – 5 Gbit/sec
- Moore's Law for switches, too
  - 1 chip: 32 x 32 switch, 1.5 Gbit/sec links, $396 48 Gbit/sec aggregate bandwidth (AMCC S2025)

Network Performance Model



Total Latency = per access + Size x per byte

per access = Sender + Receiver Overhead + Time of Flight

+ (5 to 200 µsec + 5 to 200 µsec + 0.1 µsec)

per byte + Size ÷ 100 MByte/s

Network History/Limits

- TCP/UDP/IP protocols for WAN/LAN in 1980s
- Lightweight protocols for LAN in 1990s
- Limit is standards and efficient SW protocols

10 Mbit Ethernet in 1978 (shared)

100 Mbit Ethernet in 1995 (shared, switched)

1000 Mbit Ethernet in 1998 (switched)

  - FDDI; ATM Forum for scalable LAN (still meeting)

- Internal I/O bus limits delivered BW
  - 32-bit, 33 MHz PCI bus = 1 Gbit/sec
  - future: 64-bit, 66 MHz PCI bus = 4 Gbit/sec

Network Summary

- Fast serial lines, switches offer high bandwidth, low latency over reasonable distances
- Protocol software development and standards committee bandwidth limit innovation rate
  - Ethernet forever?
- Internal I/O bus interface to network is bottleneck to delivered bandwidth, latency

Memory History/Trends/State of Art

- DRAM: main memory of all computers
  - Commodity chip industry: no company >20% share
  - Packaged in SIMM or DIMM (e.g.,16 DRAMs/SIMM)
- State of the Art: $152, 128 MB DIMM (16 64-Mbit DRAMs),10 ns x 64b (800MB/sec)
- Capacity: 4X/3 yrs (60%/yr..)
  - Moore's Law
- MB/$: + 25%/yr.
- Latency: – 7%/year, Bandwidth: + 20%/yr. (so far)

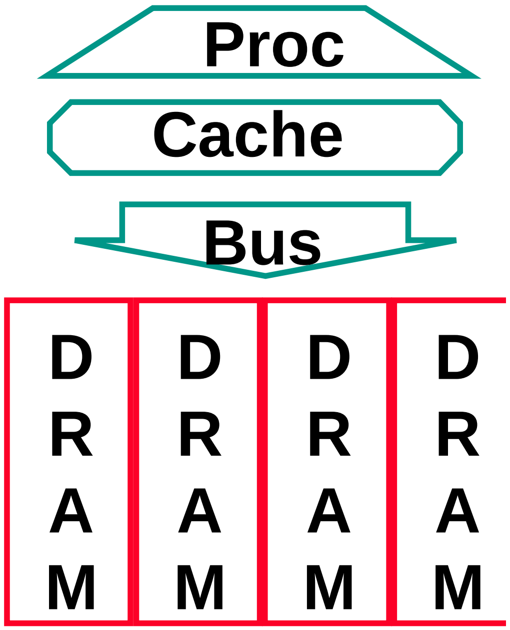
source: www.pricewatch.com, 5/21/98

Memory Innovations/Limits

- High Bandwidth Interfaces, Packages
  - RAMBUS DRAM: 800 – 1600 MByte/sec per chip
- Latency limited by memory controller, bus, multiple chips, driving pins
- More Application Bandwidth
  => More Cache misses
  = per access + block size x per byte

  Memory latency
  + Size / (DRAM BW x width)

  = 150 ns + 30 ns
  - Called Amdahl's Law: Law of diminishing returns



Memory Summary

- DRAM rapid improvements in capacity, MB/$, bandwidth; slow improvement in latency
- Processor-memory interface (cache+memory bus) is bottleneck to delivered bandwidth
  - Like network, memory “protocol” is major overhead

Processor Trends / History

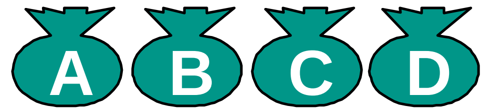
- Microprocessor: main CPU of “all” computers
  - < 1986, +35%/ yr. performance increase (2X/2.3yr)
  - >1987 (RISC), +60%/ yr. performance increase (2X/1.5yr
- Cost fixed at $500/chip, power whatever can cool

$$\text{CPU time} = \frac{\text{Seconds}}{\text{Program}} = \frac{\text{Instructions}}{\text{Program}} \times \frac{\text{Clocks}}{\text{Instruction}} \times \frac{\text{Seconds}}{\text{Clock}}$$

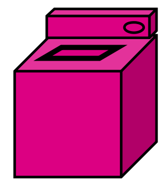
- History of innovations to 2X / 1.5 yr (Works on TPC?)
  - Multilevel Caches (helps clocks / instruction)
  - Pipelining (helps seconds / clock, or clock rate)
  - Out-of-Order Execution (helps clocks / instruction)
  - Superscalar (helps clocks / instruction)

Pipelining is Natural!

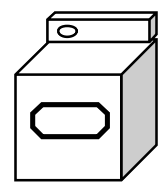
- **Laundry Example**



- **Ann, Brian, Cathy, Dave each have one load of clothes to wash, dry, fold, and put away**

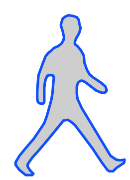


- **Washer takes 30 minutes**



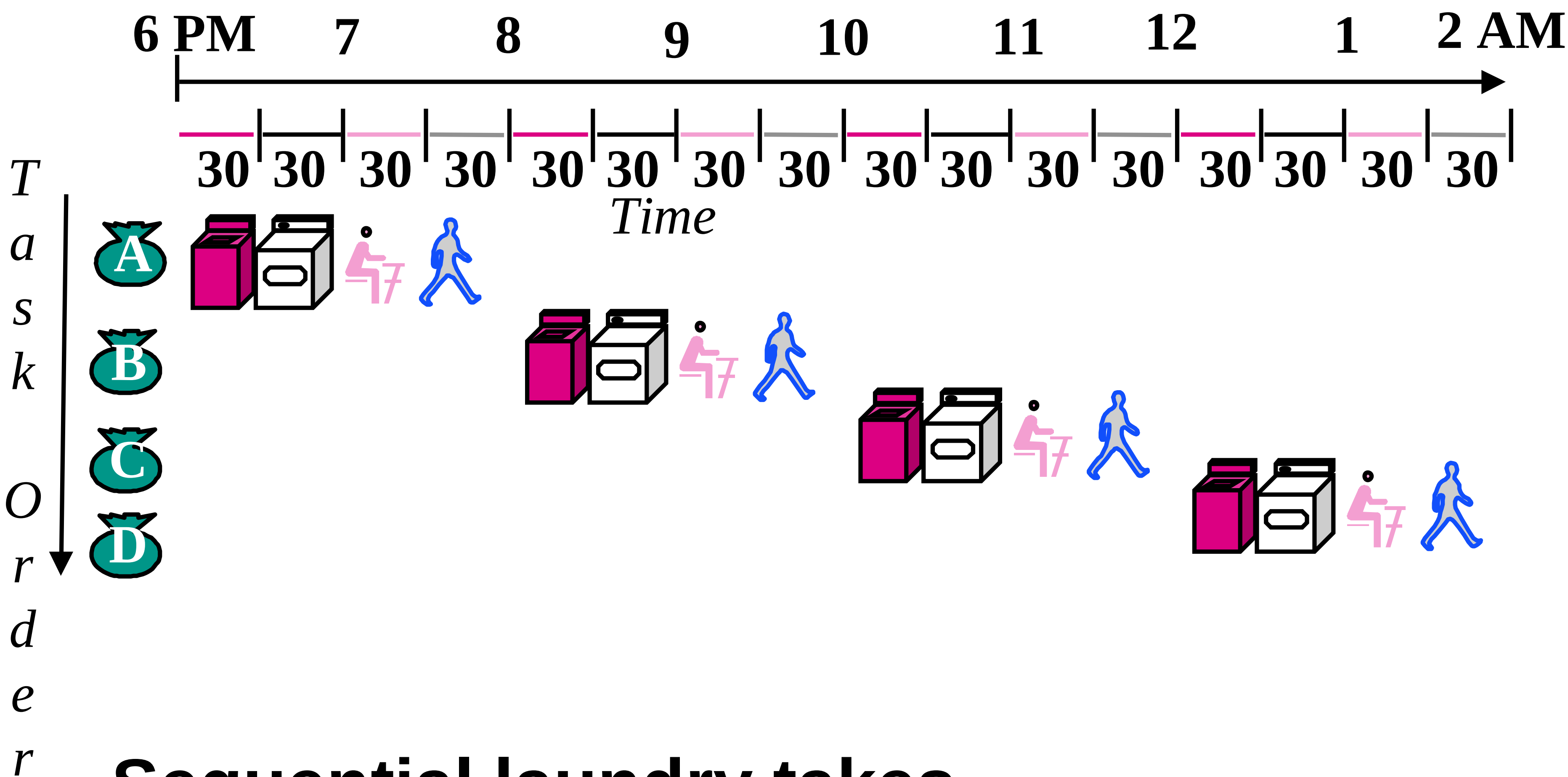
- **Dryer takes 30 minutes**

- **“Folder” takes 30 minutes**



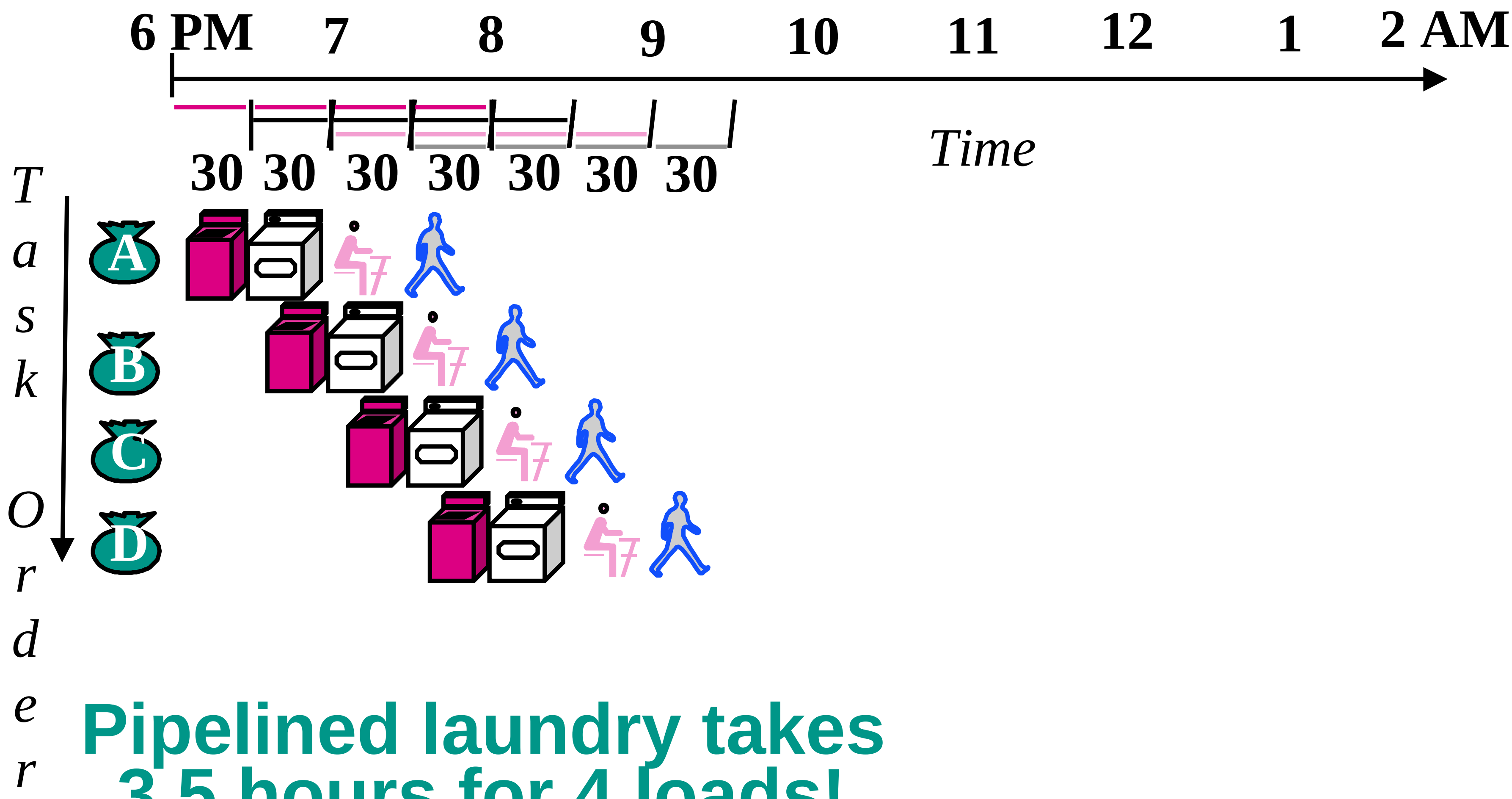
- **“Stasher” takes 30 minutes to put clothes into drawers**

Sequential Laundry



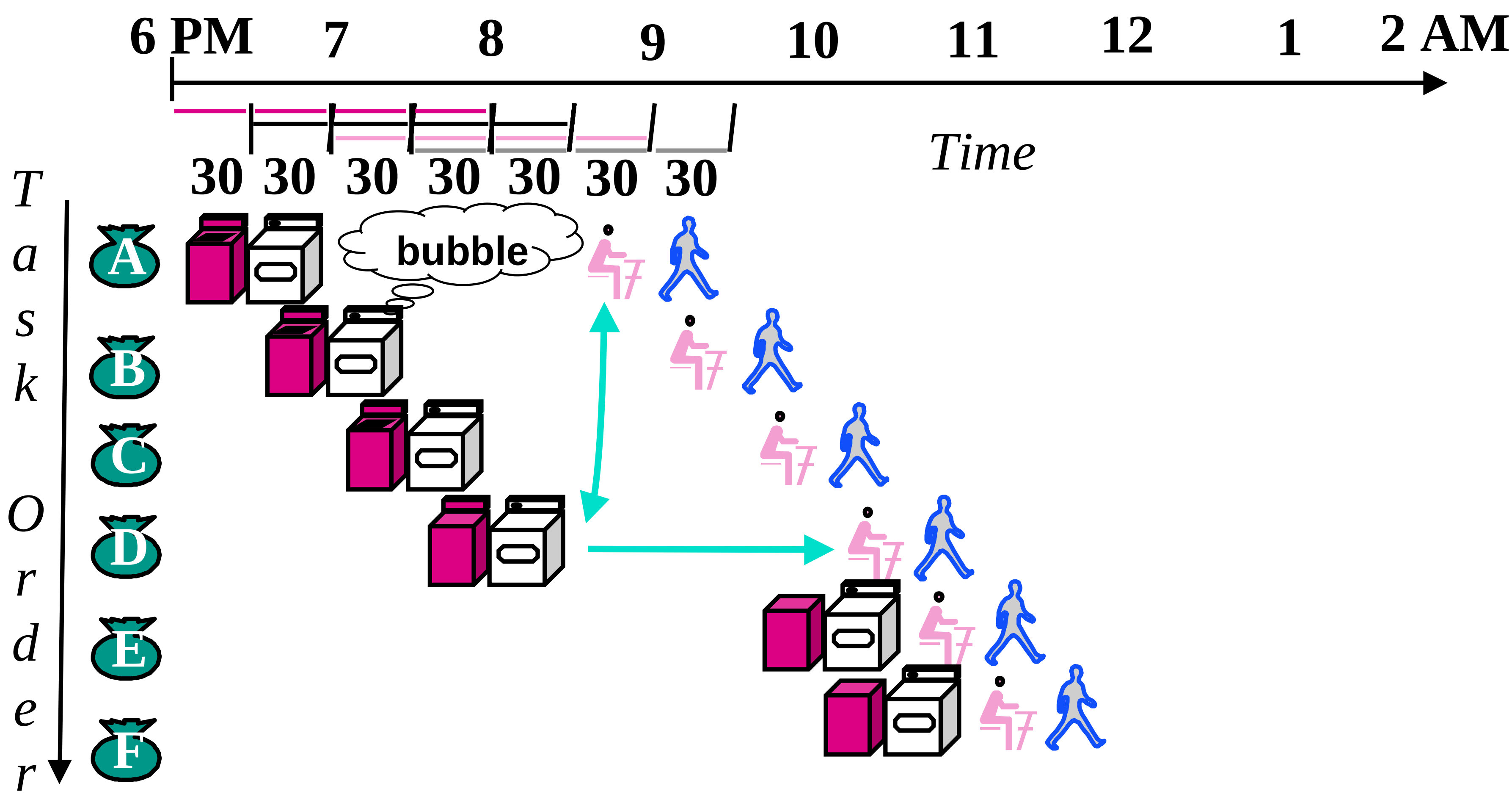
Sequential laundry takes 8 hours for 4 loads

Pipelined Laundry: Start work ASAP



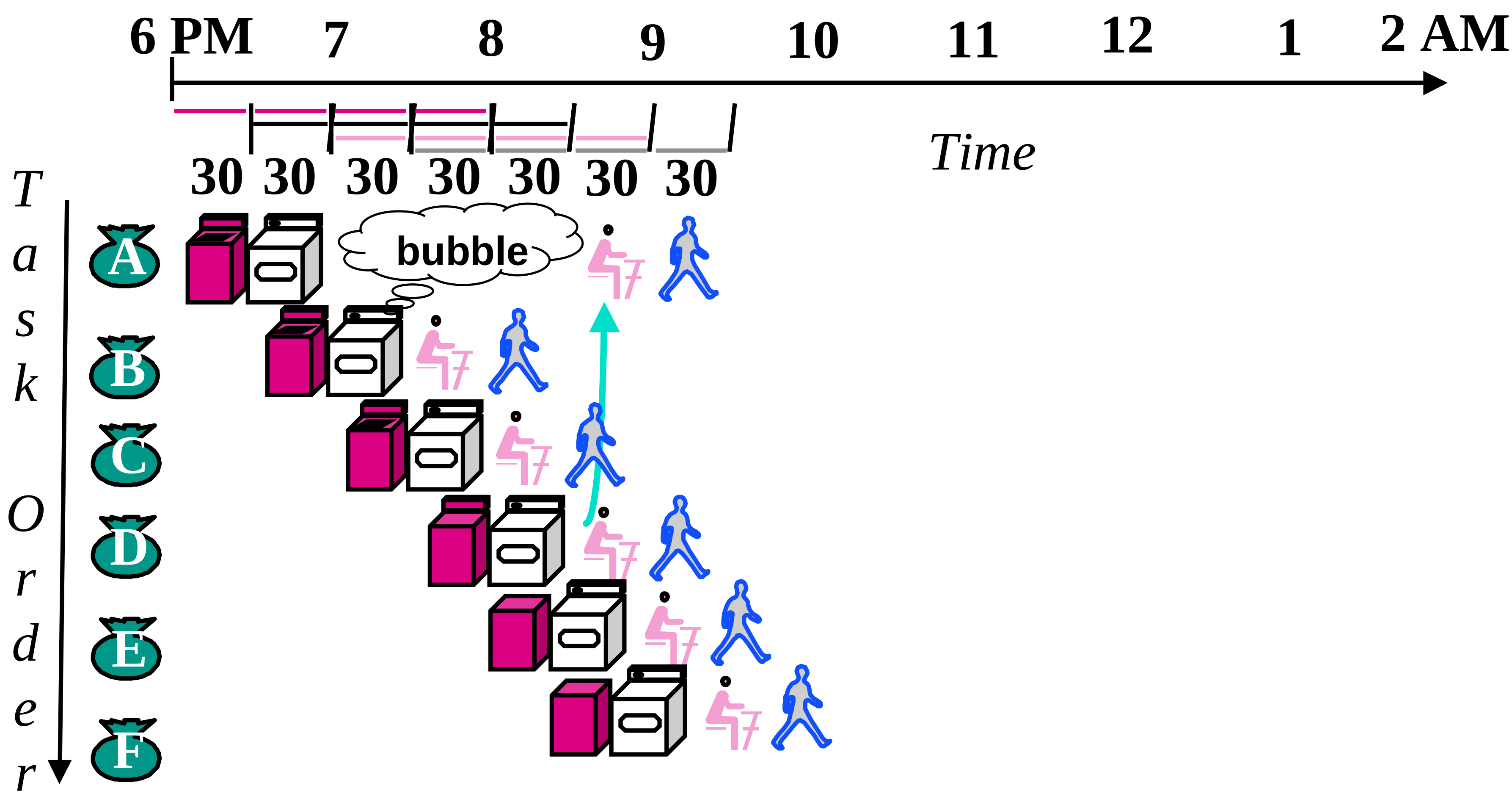
Pipelined laundry takes 3.5 hours for 4 loads!

Pipeline Hazard: Stall



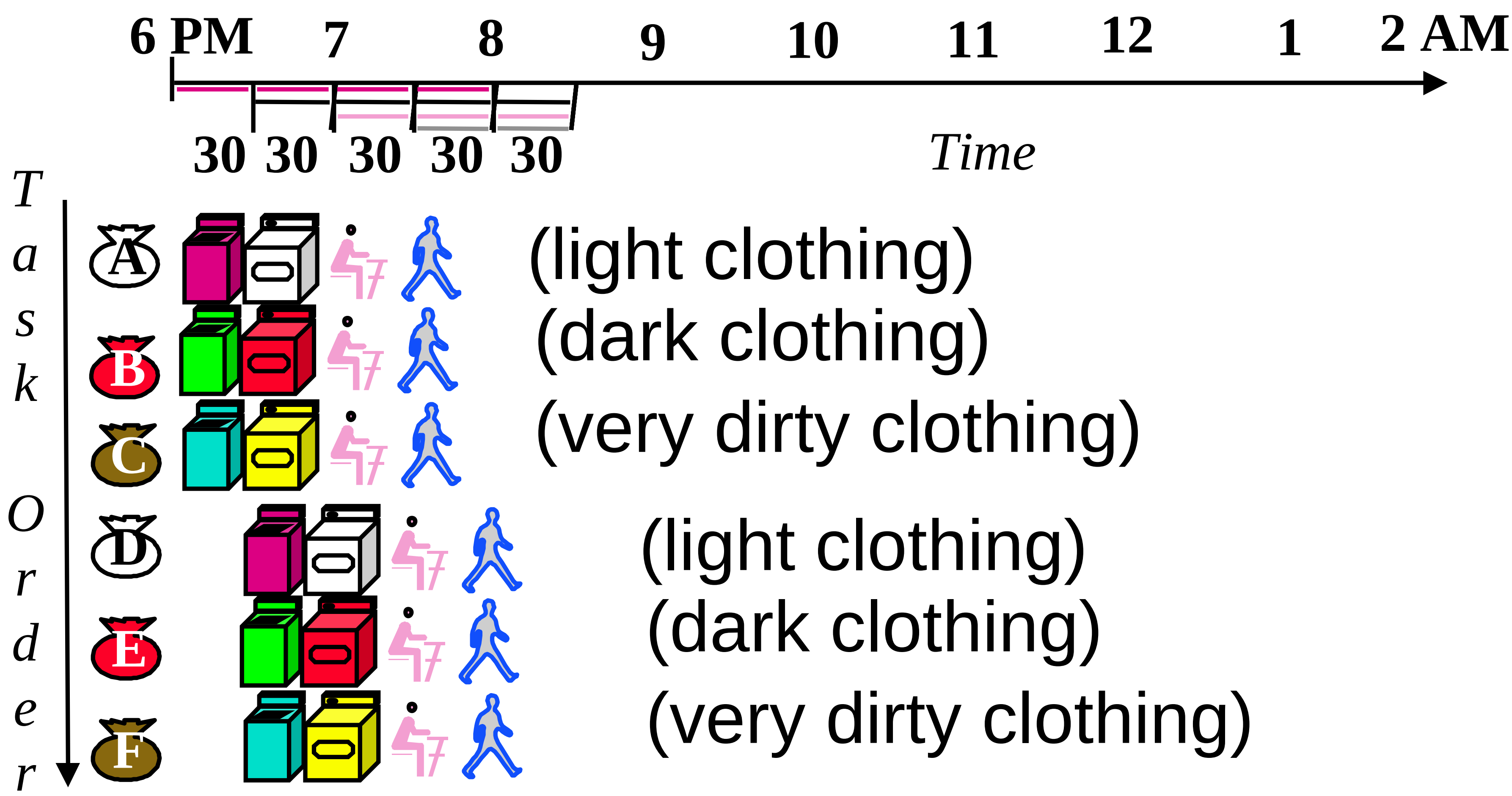
A depends on D; stall since folder tied up

Out-of-Order Laundry: Don't Wait



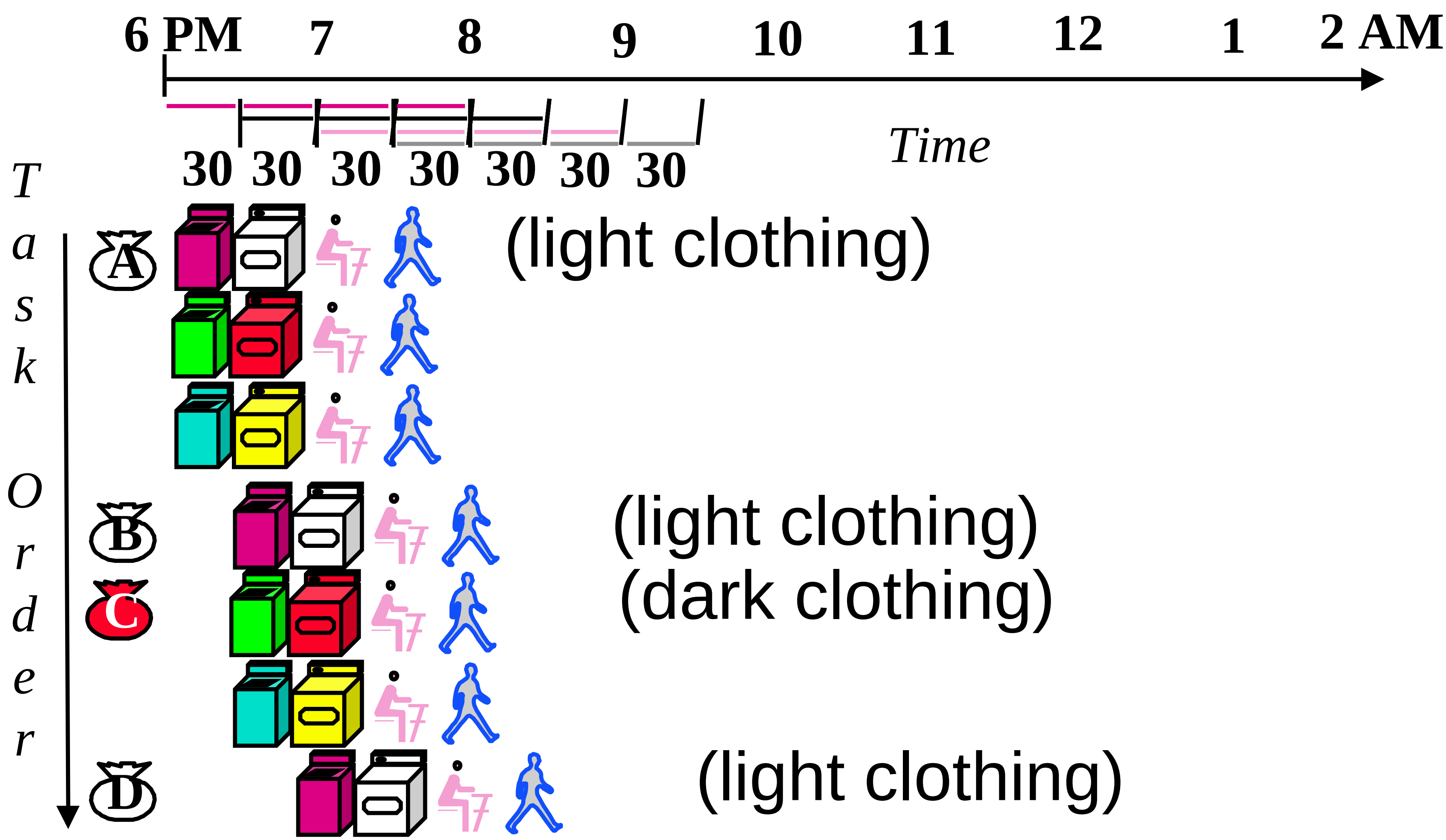
A depends on D; rest continue; need more resources to allow out-of-order

Superscalar Laundry: Parallel per stage



More resources, HW match mix of parallel tasks?

Superscalar Laundry: Mismatch Mix

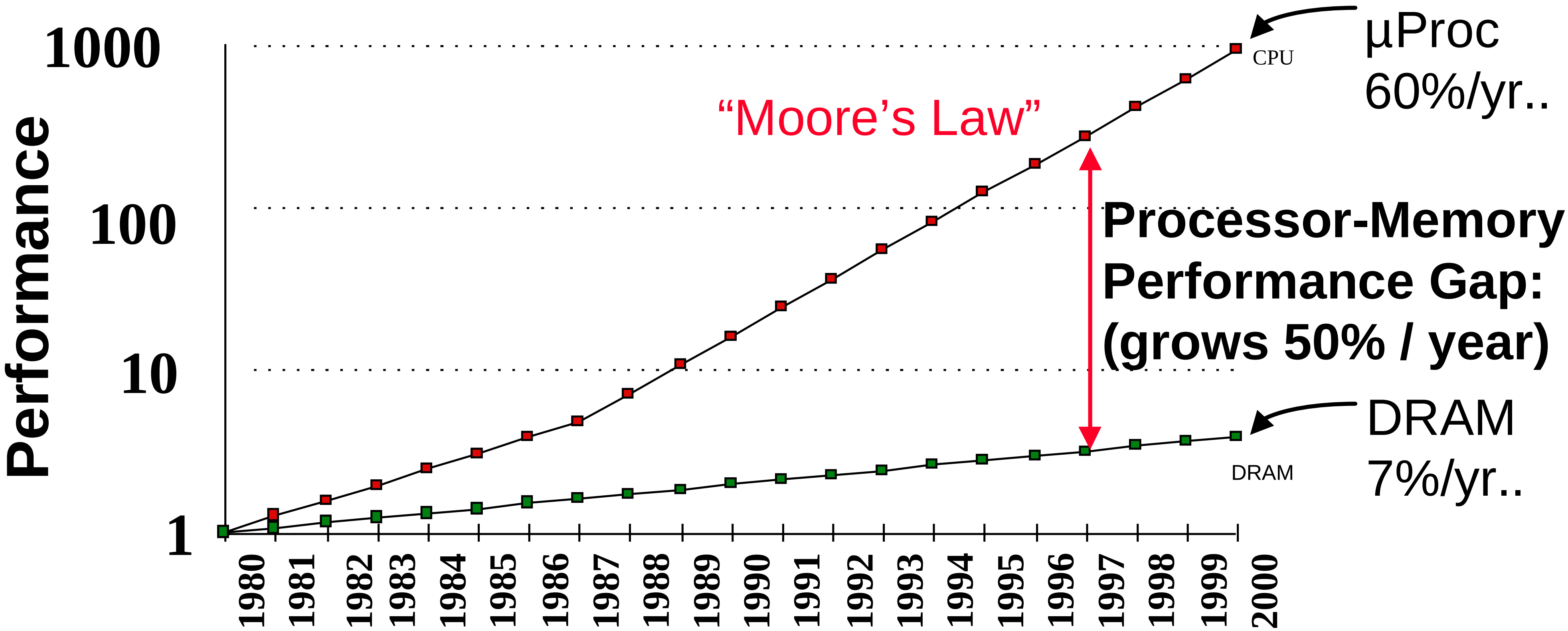


Task mix underutilizes extra resources

State of the Art: Alpha 21264

- 15M transistors
- 2 64KB caches on chip; 16MB L2 cache off chip
- Clock <1.7 nsec, or >600 MHz
  (Fastest Cray Supercomputer: T90 2.2 nsec)
- 90 watts
- Superscalar: fetch up to 6 instructions/clock cycle
  retires up to 4 instruction/clock cycle
- Execution out-of-order

Processor Limit: DRAM Gap



Alpha 21264 full cache miss in instructions executed:
180 ns/1.7 ns =108 clks x 4 or 432 instructions

Caches in Pentium Pro: 64% area, 88% transistors

Processor Limits for TPC-C

Pentium Pro	SPEC-int95	TPC-C
❍ Multilevel Caches: Miss rate 1MB L2 cache	0.5%	5%
❍ Superscalar (2-3 instr. retired/clock): % clks	40%	10%
❍ Out-of-Order Execution speedup	2.0X	1.4X
❍ Clocks per Instruction	0.8	3.4
■ % Peak performance	40%	10%

source: Kim Keeton, Dave Patterson, Y. Q. He, R. C. Raphael, and Walter Baker. "Performanc Characterization of a Quad Pentium Pro SMP Using OLTP Workloads," Proc. 25th Int'l. Symp. on Computer Architecture, June 1998. (www.cs.berkeley.edu/~kkeeton/Papers/papers.html) Bhandarkar, D.; Ding, J. "Performance characterization of the Pentium Pro processor." Proc. 3rd Int'l. Symp. on High-Performance Computer Architecture, Feb 1997. p. 288-97.

Processor Innovations/Limits

- Low cost , low power embedded processors
  - Lots of competition, innovation
  - Integer perf. embedded proc. ~ 1/2 desktop processor
  - Strong ARM 110: 233 MHz, 268 MIPS, 0.36W typ., $49
- Very Long Instruction Word (Intel,HP IA-64/Merced)
  - multiple ops/ instruction, compiler controls parallelism
- Consolidation of desktop industry? Innovation?

PA-RISC MIPS PowerPC Alpha **SPARC** IA-64 **x86**

Processor Summary

- SPEC performance doubling / 18 months
  - Growing CPU-DRAM performance gap & tax
  - Running out of ideas, competition? Back to 2X / 2.3 yrs?
- Processor tricks not as useful for transactions?
  - Clock rate increase compensated by CPI increase?
  - When > 100 MIPS on TPC-C?
- Cost fixed at ~$500/chip, power whatever can cool
- Embedded processors promising
  - 1/10 cost, 1/100 power, 1/2 integer performance?

Systems: History, Trends, Innovations

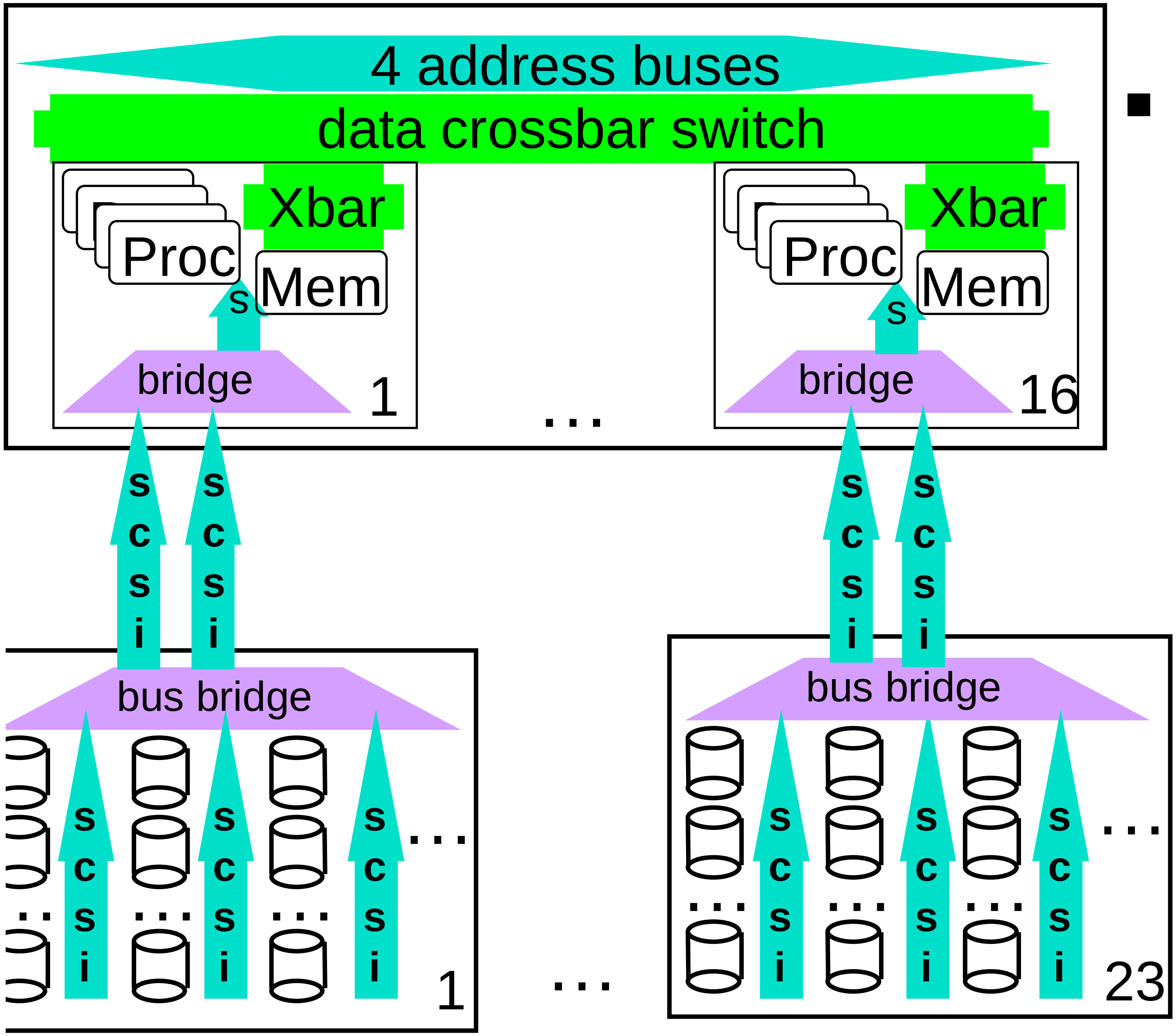
- Cost/Performance leaders from PC industry
- Transaction processing, file service based on Symmetric Multiprocessor (SMP)servers
  - 4 - 64 processors
  - Shared memory addressing
- Decision support based on SMP and Cluster (Shared Nothing)
- Clusters of low cost, small SMPs getting popular

State of the Art System: PC

- $1140 OEM
- 1 266 MHz Pentium II
- 64 MB DRAM
- 2 UltraDMA EIDE disks, 3.1 GB each
- 100 Mbit Ethernet Interface
- (PennySort winner)

source: www.research.microsoft.com/research/barc/SortBenchmark/PennySort.ps

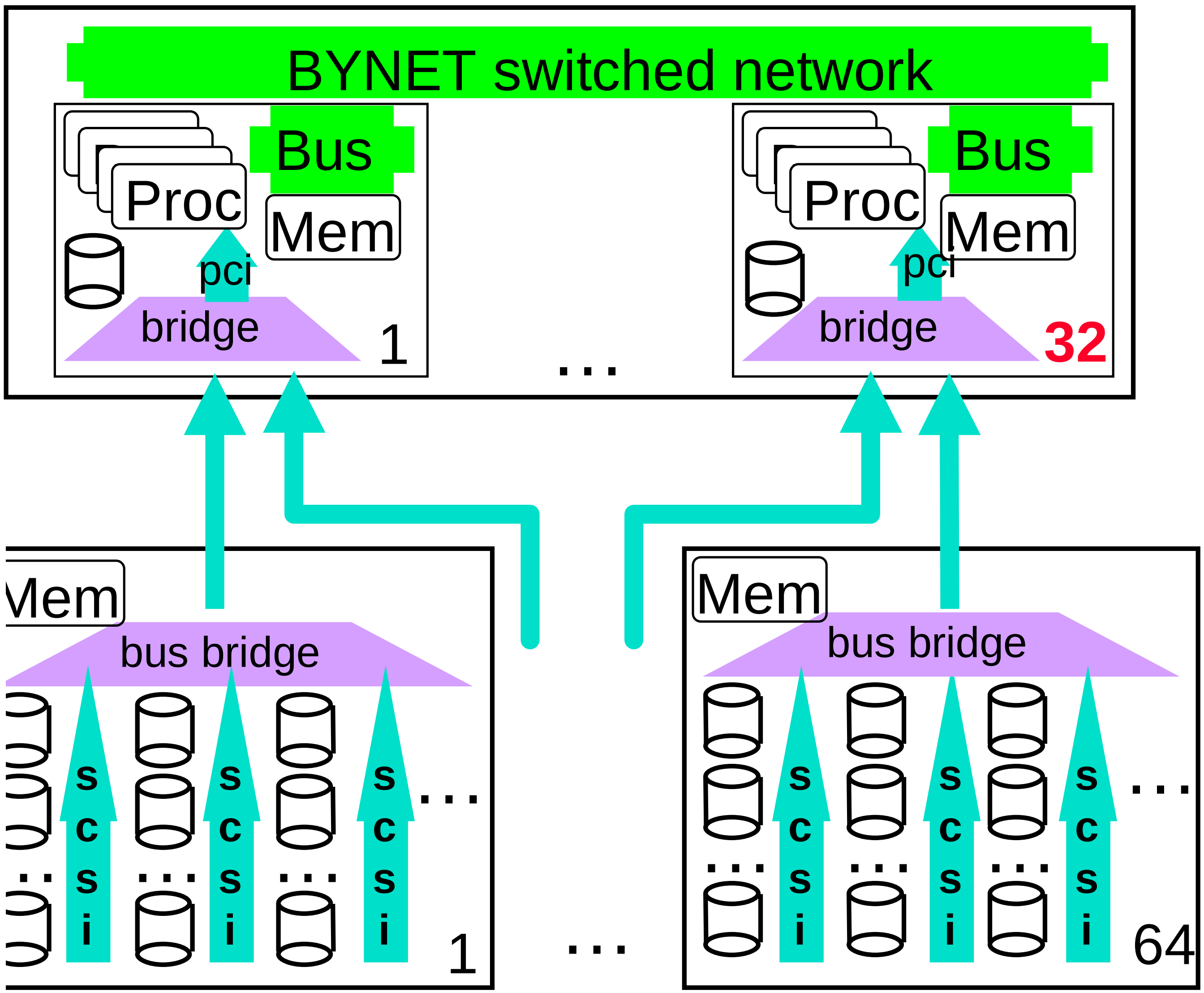
State of the Art SMP: Sun E10000



- TPC-D,Oracle 8, 3/98
  - SMP 64 336 MHz CPUs, 64GB dram, 668 disks (5.5TB)
  - Disks,shelf $2,128k
  - Boards,encl. $1,187k
  - CPUs $912k
  - DRAM $768k
  - Power $96k
  - Cables,I/O $69k
  - HW total $5,161k

source: www.tpc.org

State of the art Cluster: NCR WorldMark



- TPC-D, TD V2, 10/97
  - 32 nodes x
    4 200 MHz CPUs,
    1 GB DRAM, 41 disks
    (128 cpus, 32 GB,
    1312 disks, 5.4 TB)
  - CPUs, DRAM, encl., boards, power $5,360k
  - Disks+cntlr $2,164k
  - Disk shelves $674k
  - Cables $126k
  - Console $16k
  - HW total $8,340k

source: www.tpc.org

State of the Art Cluster: Tandem/Compaq SMP

- ServerNet switched network
- Rack mounted equipment
- SMP: 4-PPro, 3GB dram, 3 disks (6/rack)
- 10 Disk shelves/rack @ 7 disks/shelf
- Total: 6 SMPs (24 CPUs, 18 GB DRAM), 402 disks (2.7 TB)

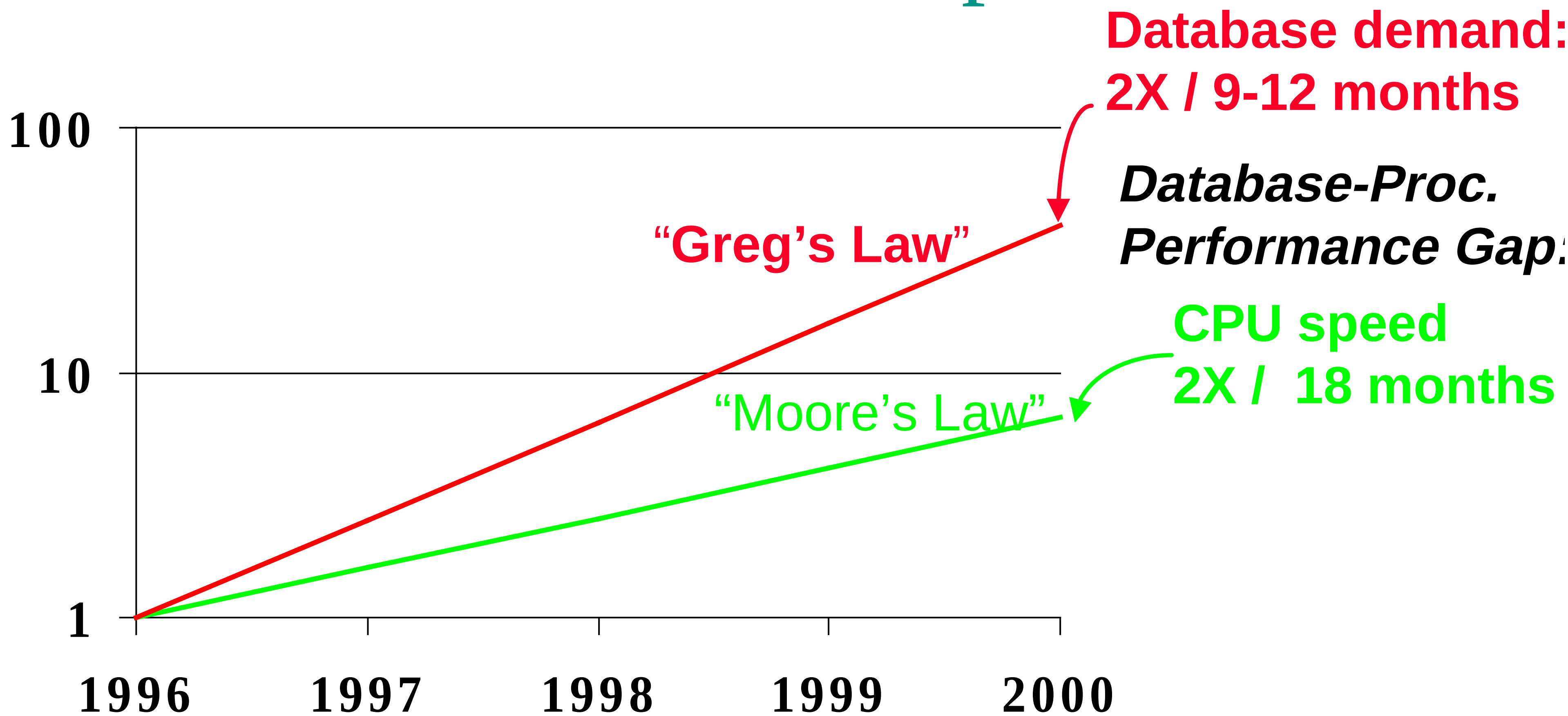
- TPC-C, Oracle 8, 4/98

CPUs	$191k
DRAM,	$122k
Disks+cntlr	$425k
Disk shelves	$94k
Networking	$76k
Racks	$15k
HW total	$926k

Berkeley Cluster: Zoom Project

- 3 TB storage system
  - 370 8 GB disks,
    20 200 MHz PPro PCs,
    100Mbit Switched Ethernet
  - System cost small delta (~30%) over raw disk cost
- Application: San Francisco Fine Arts Museum Server
  - 70,000 art images online
  - Zoom in 32X; try it yourself!
  - www.Thinker.org (statue)

User Decision Support Demand vs. Processor speed



Outline

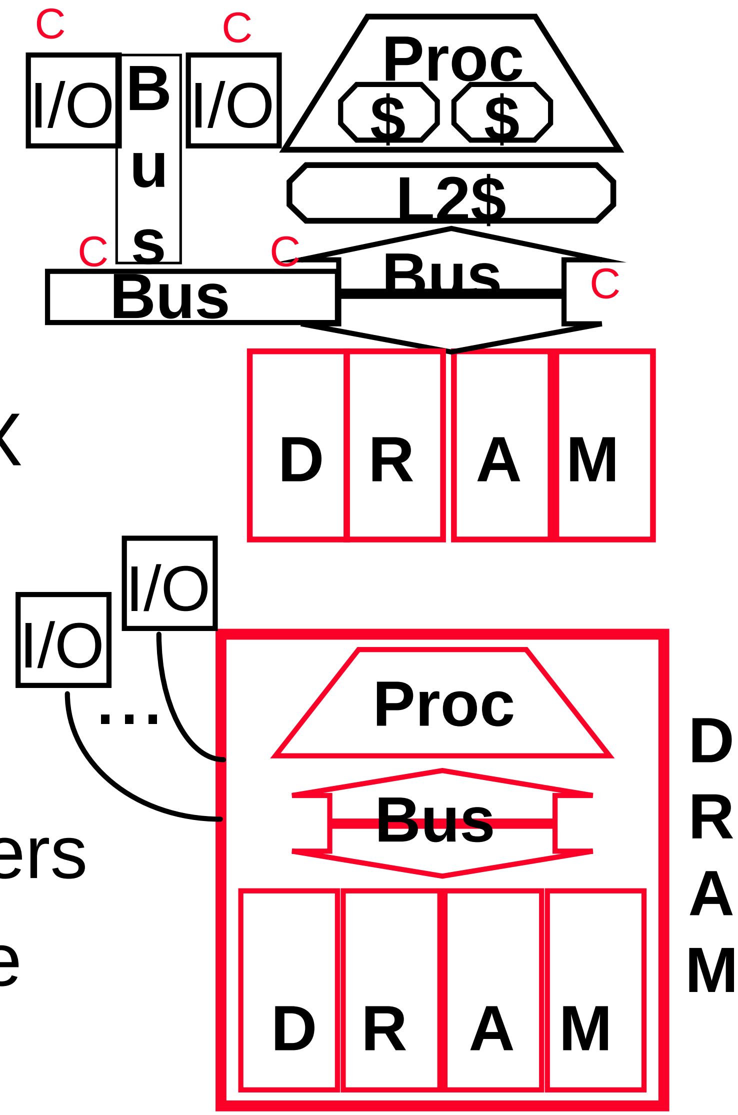
- Technology: Disk, Network, Memory, Processor, Systems
  - Description/Performance Models
  - History/State of the Art/ Trends
  - Limits/Innovations
- Technology leading to a New Database Opportunity?
  - Common Themes across 5 Technologies
  - Hardware & Software Alternative to Today
  - Benchmarks

Review technology trends to help?

- Desktop Processor:
  - \+ SPEC performance
  - – TPC-C performance, – CPU-Memory perf. gap
- Embedded Processor: + Cost/Perf, + inside disk
  - – controllers everywhere

	Disk	Memory	Network
Capacity	+	+	…
Bandwidth	+	+	+
Latency	–	–	–
Interface	–	–	–

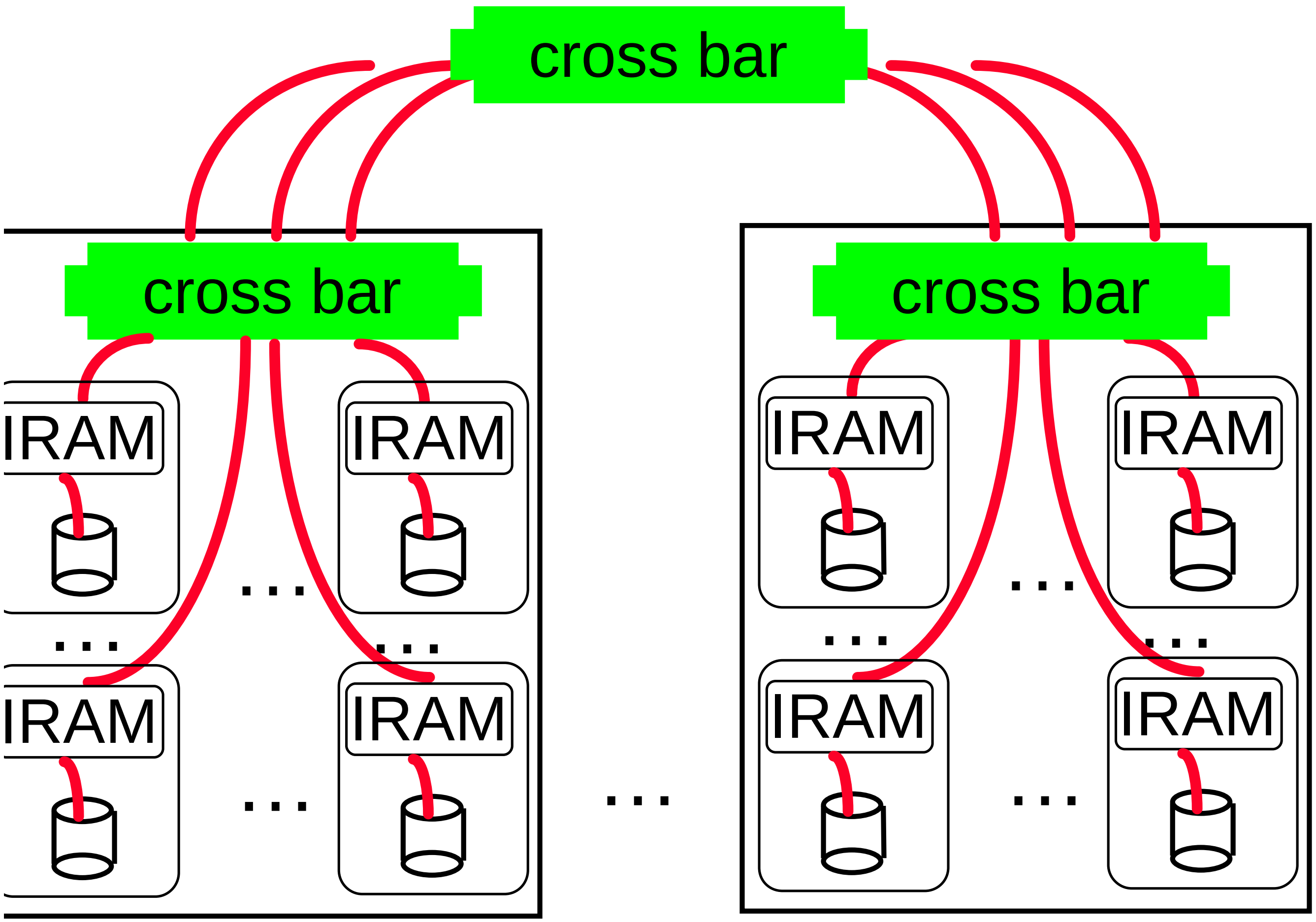
IRAM: "Intelligent RAM"



Microprocessor & DRAM on a single chip:

- on-chip memory latency 5-10X, bandwidth 50-100X
- serial I/O 5-10X v. buses
- improve energy efficiency 2X-4X (no off-chip bus)
- reduce number of controllers
- smaller board area/volume

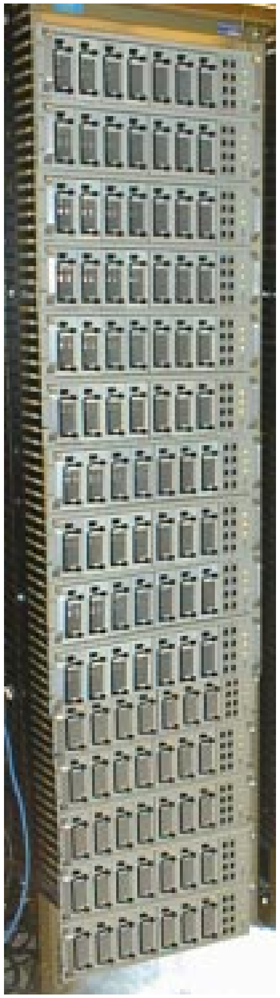
"Intelligent Disk"(IDISK): Scalable Decision Support?



- Low cost, low power processor & memory included in disk at little extra cost (e.g., Seagate optional track buffer)
- Scaleable processing AND communication as increase disks

IDISK Cluster

- 8 disks, 8 CPUs, DRAM /shelf
- 15 shelves /rack = 120 disks/rack
- 1312 disks / 120 = 11 racks
- Connect 4 disks / ring
- 1312 / 4 = 328 1.5 Gbit links
- 328 / 16 => 36 32x32 switch



- HW, assembly cost: ~$1.5 M

Cluster IDISK Software Models

1) Shared Nothing Database:
(e.g., IBM, Informix, NCR TeraData, Tandem)

2) Hybrid SMP Database:
Front end running query optimizer,
applets downloaded into IDISKs

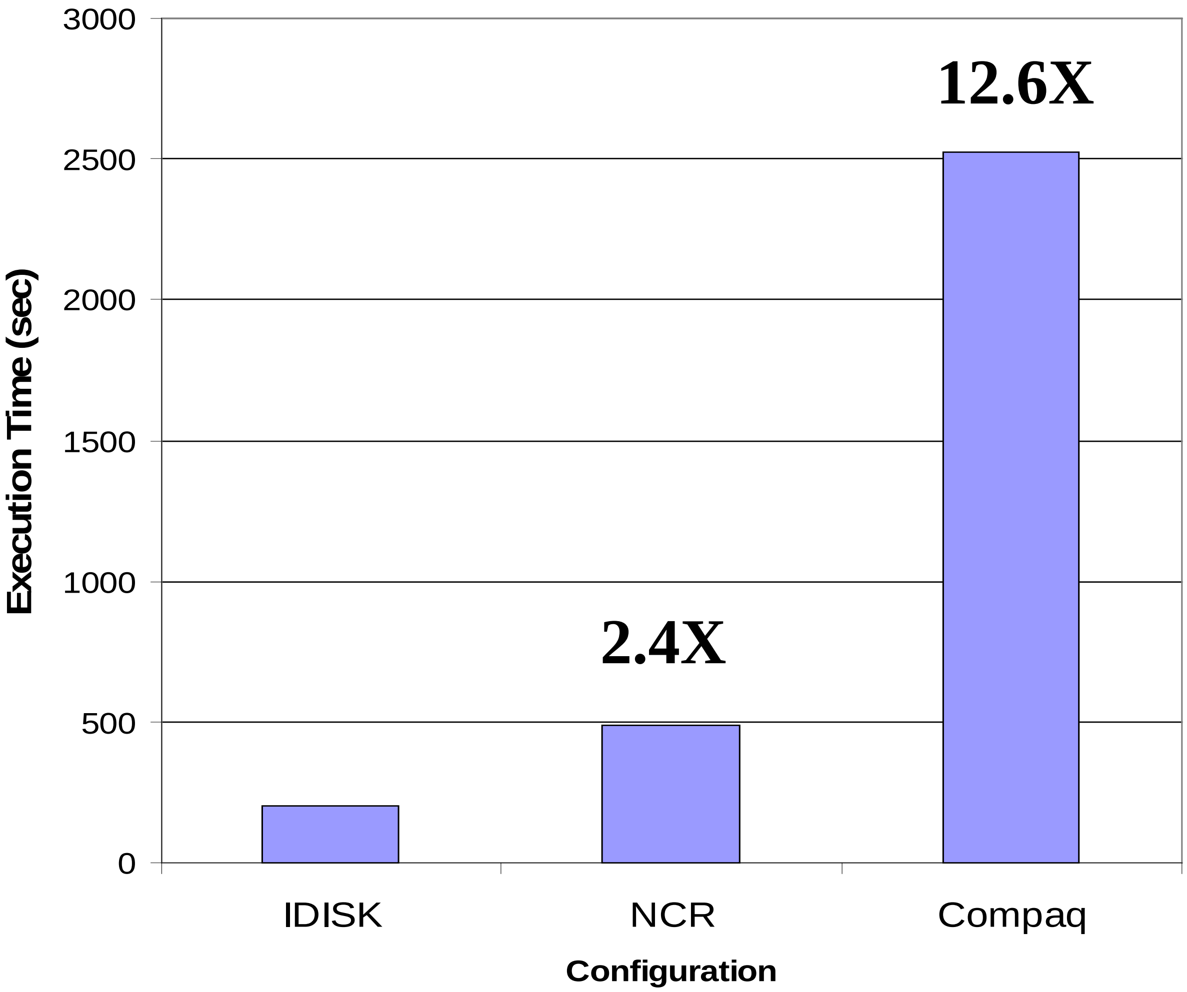
3) Start with Personal Database code developed for portable PCs, PDAs (e.g., M/S Access, M/S SQLserver, Oracle Lite, Sybase SQL Anywhere) then augment with
new communication software

Back of the Envelope Benchmarks

Characteristic	**IDISK**	**“NCR”**	**“Compaq”**
Processors per node	1 * 500 MHz	4 * 500 MHz	4 * 500 MHz
Nodes	300	32	6
Total processors	**300**	**128**	**24**
Memory capacity per node, total	**32 MB, 9.6 GB**	**4096 MB, 128 GB**	**6144 MB, 144 GB**
Disk capacity per node	1 * 10.75 GB	10 * 10.75 GB	50 * 10.75 GB
Interconnect B/W	**300*2 GB/s**	**32*125 MB/s**	**6*125 MB/s**
Disk transfer rate	29 MB/s	29 MB/s	29 MB/s
Relative Cost	**1**	**10**	**2**

- All configurations have ~300 disks
- Equivalent speeds for central and disk procs.
- Benchmarks: Scan, Sort, Hash-Join

Scan



- Scan 6 billion 145 B rows
  - TPC-D lineitem table
- Embarrassingly parallel task; limited by number processors
- IDISK Speedup:
  - NCR: 2.4X
  - Compaq: 12.6X

MinuteSort

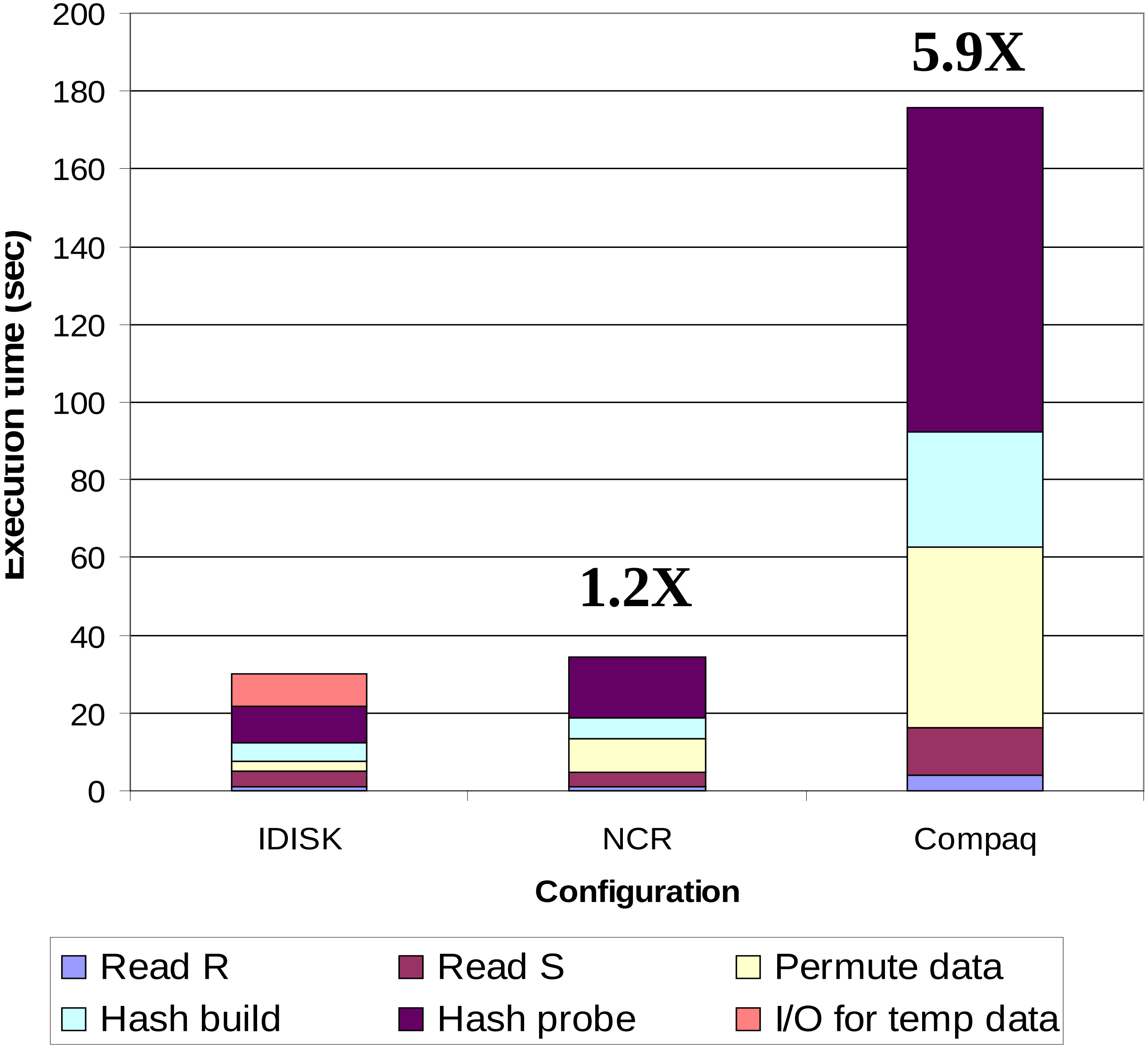
- External sorting: data starts and ends on disk
- MinuteSort: how much can we sort in a minute?
  - Benchmark designed by Nyberg, et al., SIGMOD '94
  - Current record: 8.4 GB on 95 UltraSPARC I's w/ Myrinet [NOWSort:Arpaci-Dusseau97]
- Sorting Review:
  - One-pass sort: data sorted = memory size
  - Two-pass sort:
    - Data sorted proportional to sq.rt. (memory size)
    - Disk I/O requirements: 2x that of one-pass sort

MinuteSort

	IDISK	**NCR**	**Compaq**
Algorithm	2-pass	1-pass	1-pass
Memory capacity	300*24 MB = 7 GB	32*4 GB = 128 GB	6*6 GB = 36 GB
Disk B/W	**0.03 GB/s**	0.05 GB/s	0.05 GB/s
Comm. B/W	0.06 GB/s	**0.10 GB/s**	**0.10 GB/s**
Memory B/W	0.25 GB/s	0.45 GB/s	0.45 GB/s
MinuteSort Amount	**124 GB**	**48 GB**	**9 GB**

- IDISK sorts 2.5X - 13X more than clusters
- IDISK sort limited by disk B/W
- Cluster sorts limited by network B/W

Hash-Join



- Hybrid hash join
  - R: 71k rows x 145 B
  - S: 200k rows x 165 B
  - TPC-D lineitem, part
- Clusters benefit from one-pass algorithms
- IDISK benefits from more processors, faster network
- IDISK Speedups:
  - NCR: 1.2X
  - Compaq: 5.9X

Other Uses for IDISK

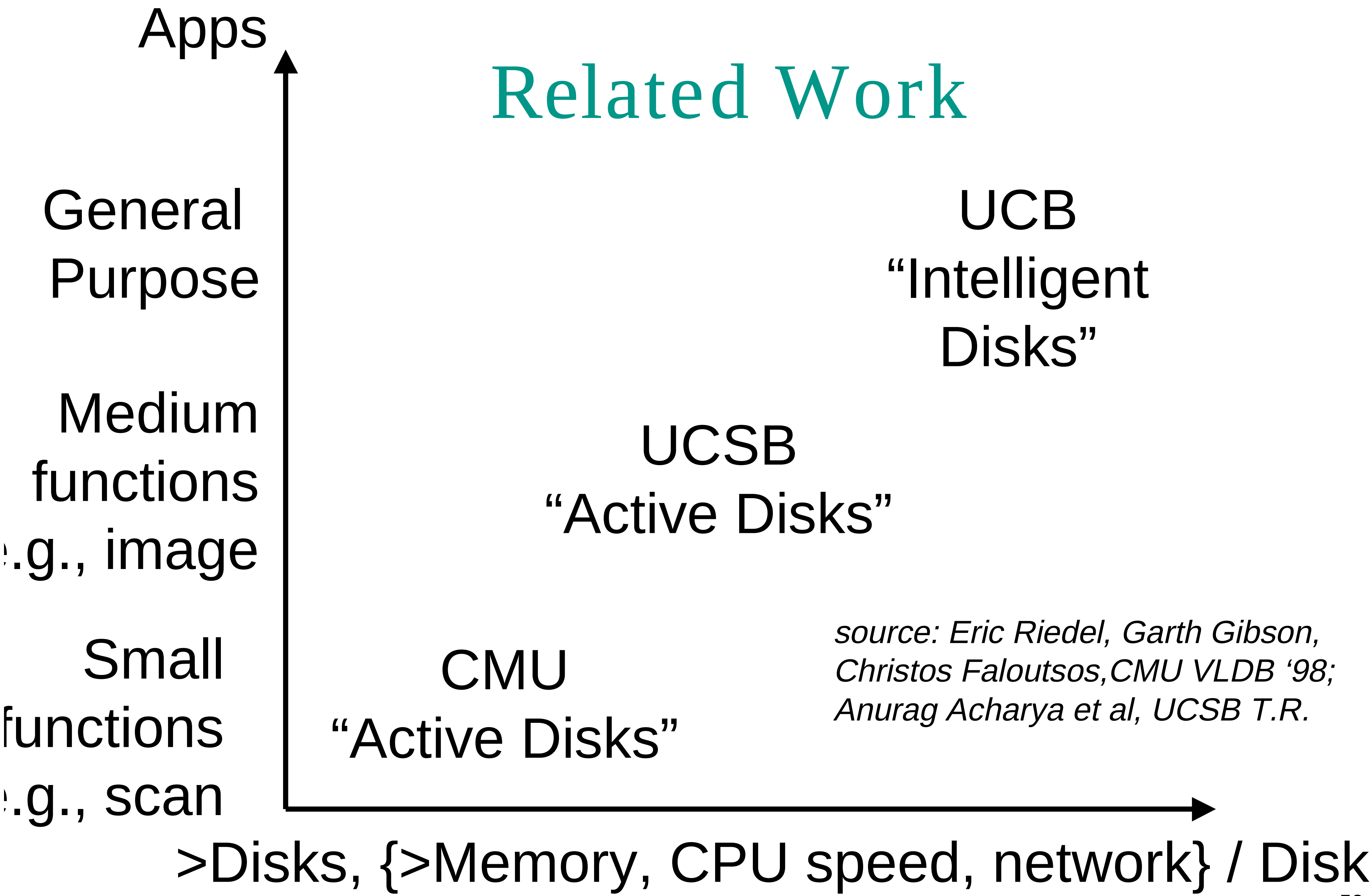
- Software RAID
- Backup accelerator
  - High speed network connecting to tapes
  - Compression to reduce data sent, saved
- Performance Monitor
  - Seek analysis, related accesses, hot data
- Disk Data Movement accelerator
  - Optimize layout without using CPU, buses

IDISK App: Network attach web, files

- Snap!Server:
  Plug in Ethernet 10/100
  & power cable, turn on
- 32-bit CPU, flash memory, compact multitasking OS, SW update from Web
- Network protocols: TCP/IP, IPX, NetBEUI, and HTTP (Unix, Novell, M/S, Web)
- 1 or 2 EIDE disks
- 6GB $950,12GB $1727 (7MB/$, 14¢/MB)

source: www.snapserver.com, www.cdw.com

Related Work



source: Eric Riedel, Garth Gibson, Christos Faloutsos,CMU VLDB ‘98; Anurag Acharya et al, UCSB T.R.

IDISK Summary

- IDISK less expensive by 10X to 2X, faster by 2X to 12X?
  - Need more realistic simulation, experiments
- IDISK scales better as number of disks increase, as needed by Greg's Law
- Fewer layers of firmware and buses, less controller overhead between processor and data
- IDISK not limited to database apps: RAID, backup, Network Attached Storage, ...
- Near a strategic inflection point?

Messages from Architect to Database Community

- Architects want to study databases; why ignored?
  - Need company OK before publish! (“DeWitt” Clause)
  - DB industry, researchers fix if want better processors
  - SIGMOD/PODS join FCRC?
- Disk performance opportunity: minimize seek, rotational latency, ultilize space v. spindles
- Think about smaller footprint databases: PDAs, IDISKs, ...
  - Legacy code a reason to avoid virtually all innovations???
  - Need more flexible/new code base?

Acknowledgments

- Thanks for feedback on talk from M/S BARC (Jim Gray, Catharine van Ingen, Tom Barclay, Joe Barrera, Gordon Bell, Jim Gemmell, Don Slutz) and IRAM Group (Krste Asanovic, James Beck, Aaron Brown, Ben Gribstad, Richard Fromm, Joe Gebis, Jason Golbus, Kimberly Keeton, Christoforos Kozyrakis, John Kubiatowicz, David Martin, David Oppenheimer, Stelianos Perissakis, Steve Pope, Randi Thomas, Noah Treuhaft, and Katherine Yelick
- Thanks for research support: DARPA, California MICRO, Hitachi, IBM, Intel, LG Semicon, Microsoft, Neomagic, SGI/Cray, Sun Microsystems, TI

Questions?

Contact us if you’re interested:

email: patterson@cs.berkeley.edu

http://iram.cs.berkeley.edu/

1970s != 1990s

- Scan Only
- Limited communication between disks
- Custom Hardware
- Custom OS
- Invent new algorithms
- Only for for databases

- Whole database code
- High speed communication between disks
- Optional intelligence added to standard disk (e.g., Cheetah track buffer)
- Commodity OS
- 20 years of development
- Useful for WWW, File Servers, backup

Stonebraker's Warning

"The history of DBMS research is littered with innumerable proposals to construct hardware database machines to provide high performance operations. In general these have been proposed by hardware types with a clever solution in search of a problem on which it might work."

Readings in Database Systems (second edition), edited by Michael Stonebraker, p.603

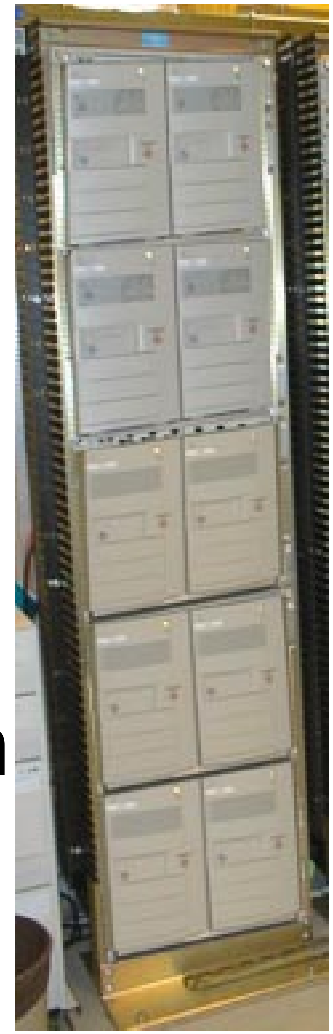
Grove's Warning

"...a strategic inflection point is a time in the life of a business when its fundamentals are about to change. ... Let's not mince words: A strategic inflection point can be deadly when unattended to. Companies that begin a decline as a result of its changes rarely recover their previous greatness."

Only the Paranoid Survive, Andrew S. Grove, 1996

Clusters of PCs?

- 10 PCs/rack = 20 disks/rack
- 1312 disks / 20 = 66 racks, 660 PCs
- 660 /16 = 42 100 Mbit Ethernet Switches
  + 9 1Gbit Switches
- 72 racks / 4 = 18 UPS
- Floor space: aisles between racks to access, repair PCs
  72 / 8 x 120 = 1100 sq. ft.



- HW, assembly cost: ~$2M
- Quality of Equipment?
- Repair?
- System Admin.?

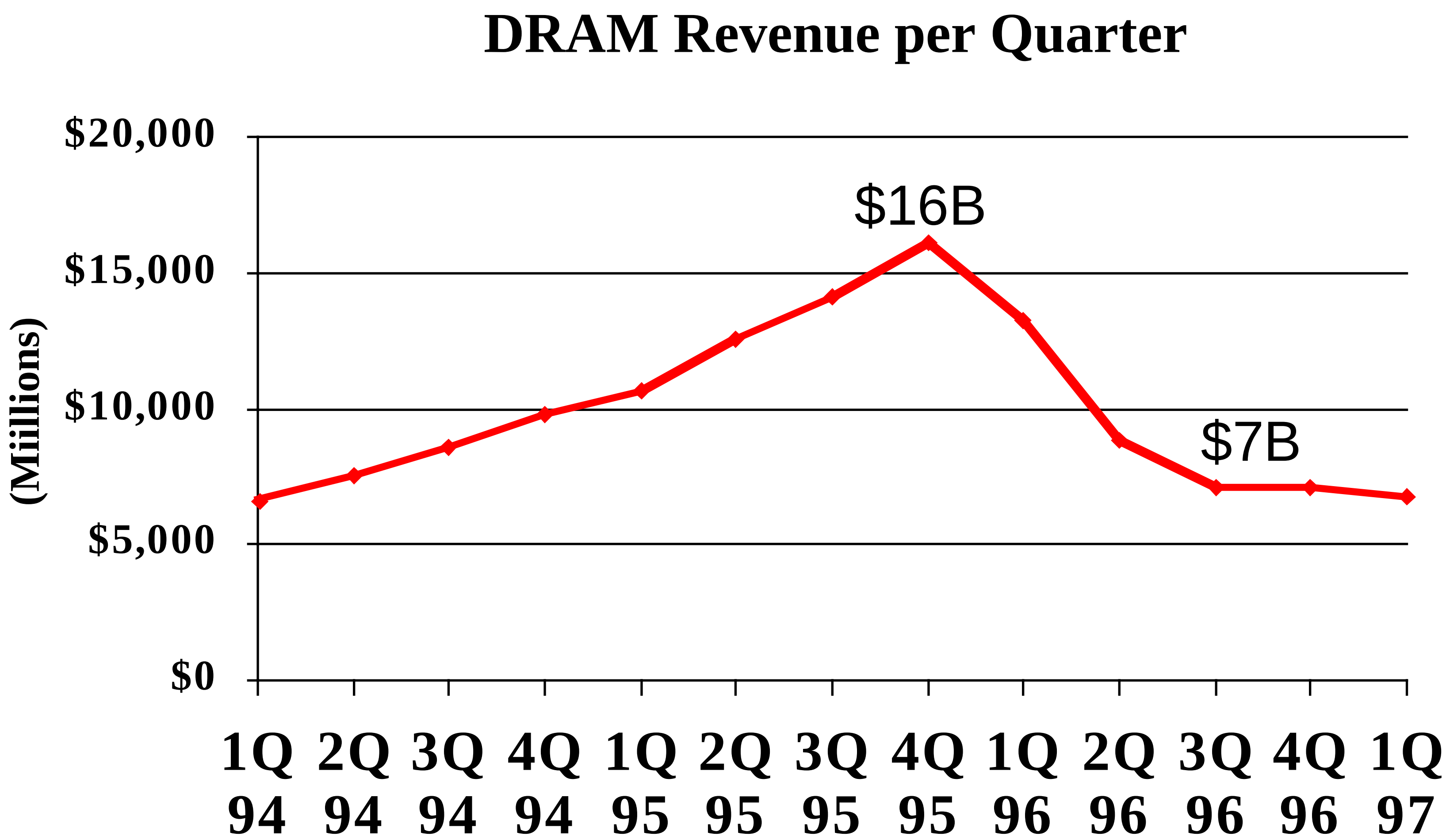
Today's Situation: Microprocessor

MIPS MPUs	R5000	R10000	10k/5k
■ Clock Rate	200 MHz	195 MHz	1.0x
■ On-Chip Caches	32K/32K	32K/32K	1.0x
■ Instructions/Cycle	1(+ FP)	4	4.0x
■ Pipe stages	5	5-7	1.2x
■ Model	In-order	Out-of-order	---
■ Die Size (mm^2)	84	298	3.5x
❍ without cache, TLB	32	205	6.3x
■ Development (man yr..)	60	300	5.0x
■ SPECint_base95	5.7	8.8	1.6x

Potential Energy Efficiency: 2X-4X

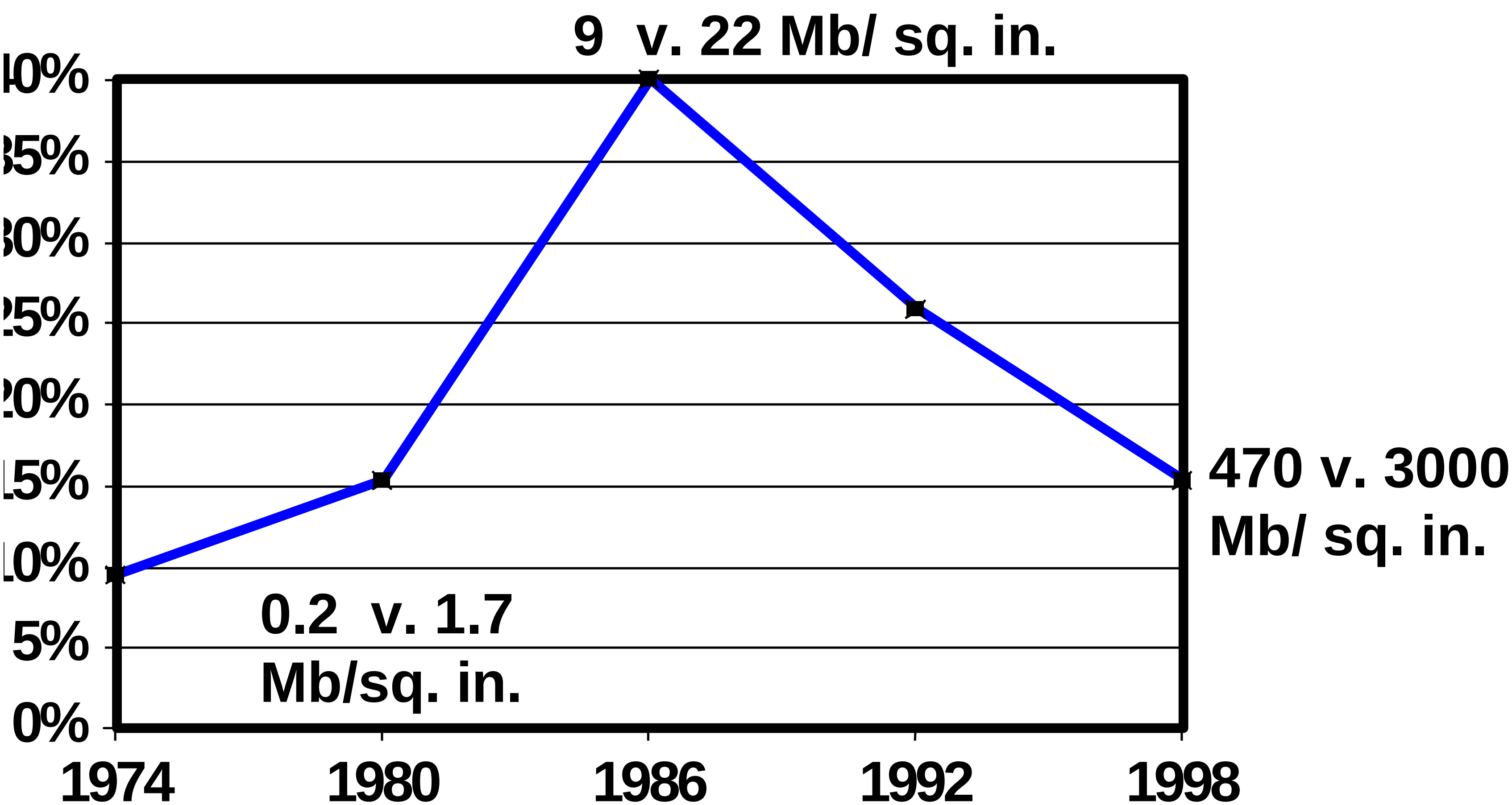
- Case study of StrongARM memory hierarchy vs. IRAM memory hierarchy
  - cell size advantages ⇒ much larger cache
    ⇒ fewer off-chip references
    ⇒ up to 2X-4X energy efficiency for memory
  - less energy per bit access for DRAM
- Memory cell area ratio/process: P6, α '164,SArm
  cache/logic : SRAM/SRAM : DRAM/DRAM
  20-50 : 8-11 : 1

Today’s Situation: DRAM



- Intel: 30%/year since 1987; 1/3 income profit

MBits per square inch: DRAM as % of Disk over time



source: New York Times, 2/23/98, page C3, "Makers of disk drives crowd even more data into even smaller spaces"

What about I/O?

- Current system architectures have limitations
- I/O bus performance lags other components
- Parallel I/O bus performance scaled by increasing clock speed and/or bus width
  - E.g.. 32-bit PCI: ~50 pins; 64-bit PCI: ~90 pins
  - Greater number of pins $\Rightarrow$ greater packaging costs
- Are there alternatives to parallel I/O buses for IRAM?

Serial I/O and IRAM

- Communication advances: fast (Gbit/s) serial I/O lines [YankHorowitz96], [DallyPoulton96]
  - Serial lines require 1-2 pins per unidirectional link
  - Access to standardized I/O devices
    - » Fiber Channel-Arbitrated Loop (FC-AL) disks
    - » Gbit/s Ethernet networks
- Serial I/O lines a natural match for IRAM
- Benefits
  - Serial lines provide high I/O bandwidth for I/O-intensive applications
  - I/O BW incrementally scalable by adding more lines
    - » Number of pins required still lower than parallel bus