Gordon Bell Prize, Three Decades: Motivating and measuring HPC progress

The Gordon Bell Prizes chronicles the important innovations and transitions of HPC beginning with the first award in 1987 that demonstrated Amdahl's Law was not impenetrable. For example, in 1987 MPI provided both a model and standard for rapid adoption . Almost every gain in parallelism has been recognized, from widely distributed workstations to 10 million core processors. The overlap from the Seymour Cray recipe of a shared memory, multi-vector processor (aka Fortran Computer) to today's multicomputer turned out to show up and be incredibly fast. Bell would like to believe the prize helps in rapid adoptions of new paradigms e.g. GPUs and possibly FPGAs. The prize also recognized the value of specialized hardware. Finally, the prize recognizes the tremendous effort required to exploit the increasingly parallelism afforded by the computers as seen in the increase in number of authors.

Gordon Bell

Microsoft Research

-

Supercomputing Frontiers 2017, 13-17 March

Singapore

Outline: 30 Years of the GB Prize

- Origin and motivation for the Gordon Bell Prize
- 1965-1995: From mono-memory computers to multicomputers
  - Bell religion: Scaleup, Giveup, …Scaleout, new hope for Scaleup
  - Ancient history of HPC Cray Recipes for the best FORTRAN computer
- 1984-1994 Transition from mono-memory to clustered computers
- 1987-Present Bell Prize and 1992 Linpack chronicle progress
  - Work, teams and difficulty of experiments, e.g. building a detector
  - Winner organizations, teams, and applications
  - Number of participants for prize, number winners.
  - Platforms and special hardware
  - Significant work by the judges… selecting the best work

Bell Prize for Parallelism, July 1987… CISE parallelism fosus, Karp-Knuth-Thompson,

Hardware:multicore/thread, stimulate, not one shot , reward what was difficult now and foreseeable future… 10 year prize, move computational science

Karp challenge:

Alan Karp: $100 for a program with 200 X parallelism by 1995.

Fujitsu, Num. Wind Tunnel Pv:128

My own 1987 predictions:

10 X by 1992… ***Cray *MP likely***

Intel Delta 512, TMC 1024

100 X by 1997

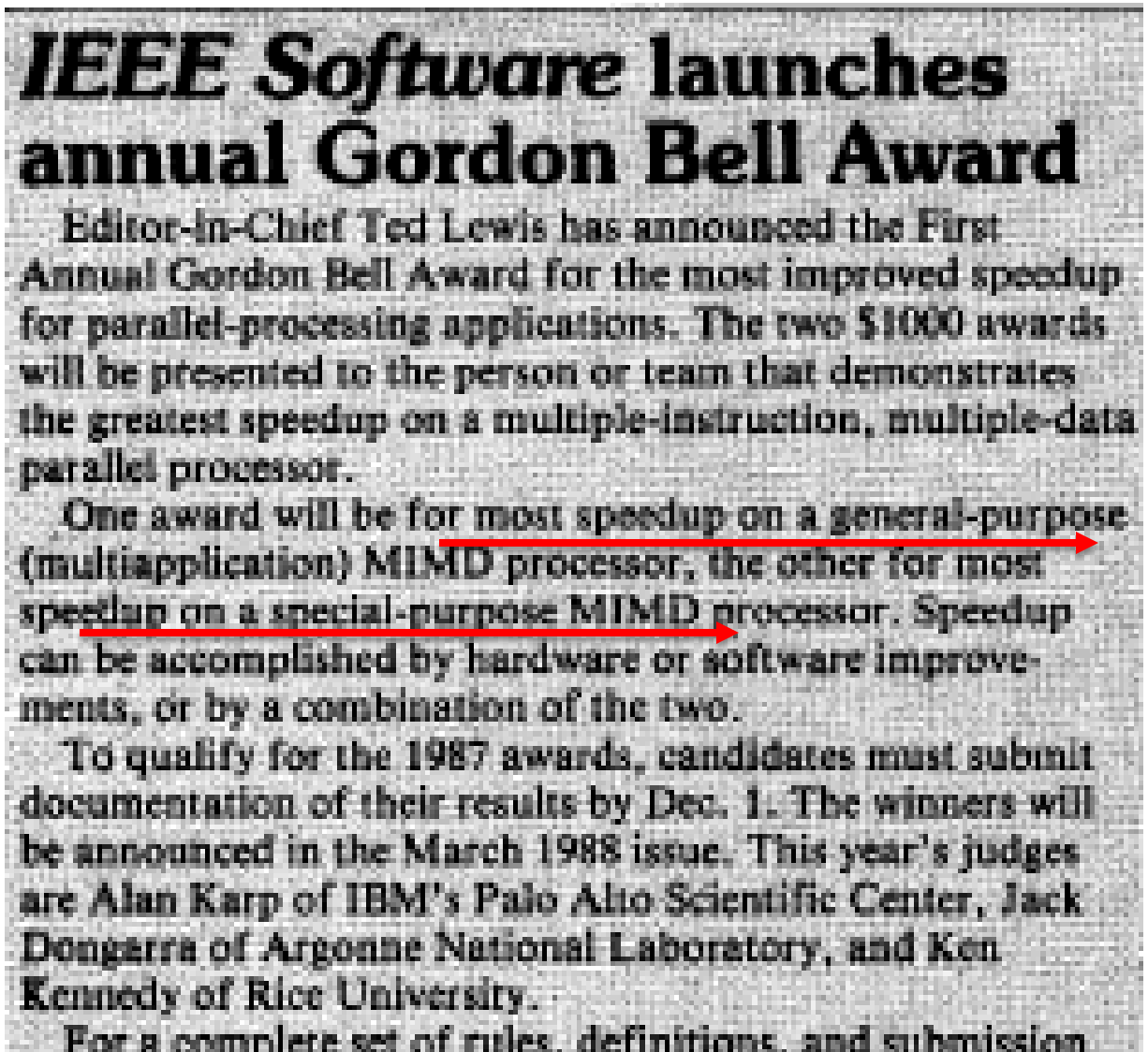
ASCI Red 9,152 computers

SIMD would allow 1 million

Research community claims:

1 million X by 2002 ***Sequoia in 2013***

Earth Simulator 5.120 Pvector



IEEE Software launches annual Gordon Bell Award

Editor-in-Chief Ted Lewis has announced the First Annual Gordon Bell Award for the most improved speedup for parallel-processing applications. The two $1000 awards will be presented to the person or team that demonstrates the greatest speedup on a multiple-instruction, multiple-data parallel processor.

One award will be for most speedup on a general-purpose (multiapplication) MIMD processor, the other for most speedup on a special-purpose MIMD processor. Speedup can be accomplished by hardware or software improvements, or by a combination of the two.

To qualify for the 1987 awards, candidates must submit documentation of their results by Dec. 1. The winners will be announced in the March 1988 issue. This year's judges are Alan Karp of IBM's Palo Alto Scientific Center, Jack Dongarra of Argonne National Laboratory, and Ken Kennedy of Rice University.

For a complete set of rules, definitions, and submission

Prize evolution

Scaling, **HM (1987**), Peak Perf (1988), Price-Perf (1988),

Compiler Parallelization (1989), Speedup (1992),

Special Purpose Machine (1995),
Special Award for Language (2002),
Special achievement (lifetime, 2003),

Algorithm Innovation (2008),

Sustained Performance (2011),
Scalability and time to solution (2011),

Achievement in scalability (2015)

Wikipedia Prize Criteria, Aug 2015

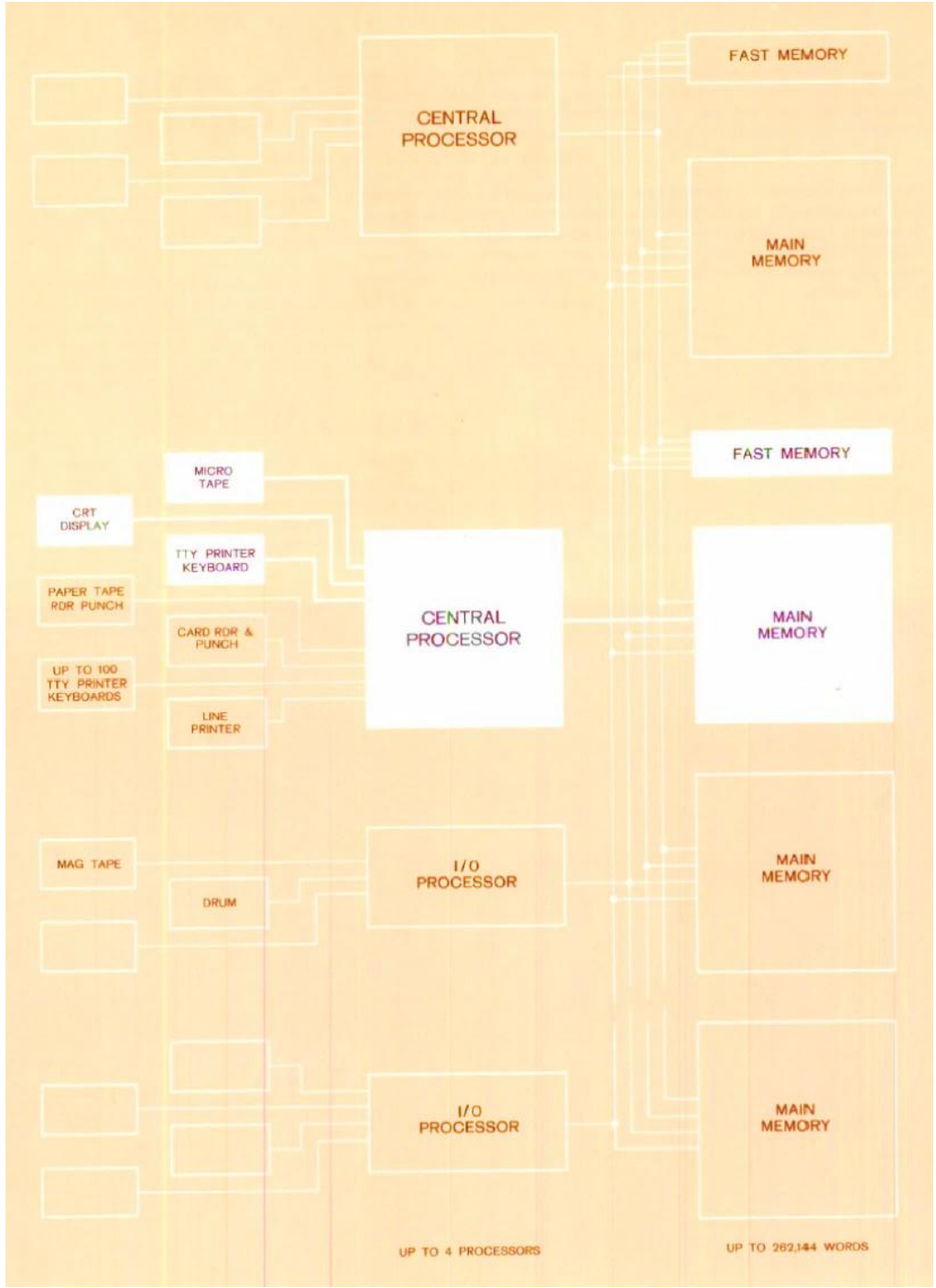
The ACM Gordon Bell Prize is primarily intended to recognize performance achievements that demonstrate:

- evidence of important algorithmic and/or implementation innovations
- clear improvement over the previous state-of-the-art
- solutions that don't depend on one-of-a-kind architectures (systems that can only be used to address a narrow range of problems, or that can't be replicated by others)
- performance measurements that have been characterized in terms of scalability (strong as well as weak scaling), time to solution, efficiency (in using bottleneck resources, such as memory size or bandwidth, communications bandwidth, I/O), and/or peak performance
- achievements that are generalizable, in the sense that other people can learn and benefit from the innovations

In earlier years, multiple prizes were sometimes awarded to reflect different types of achievements.

According to current policies, the Prize can be awarded in one or more of the following categories, depending on the entries received in a given year:

- **Peak Performance**: If the entry demonstrates outstanding performance in terms of floating point operations per second on an important science/engineering problem; the efficiency of the application in using bottleneck resources (such as memory size or bandwidth) is also taken into consideration.
- **Special Achievement in Scalability, Special Achievement in Time to Solution**: If the entry demonstrates exceptional Scalability, in terms of both strong and weak scaling, and/or total time to solve an important science/engineering problem.



Bell Architecture beliefs

1965-1994 Shared memory mP…

Scaleup

- 1965 PDP-6 multiprocessor
- 1972 C.mmp 16 processor
- ***1982 VAX Clusters, scaleout, shared storage for capacity***
- 1985 Encore “multi” 16 P
  1987 Encore DARPA 1K mP
- 1988 Titan graphics super (4 mPv)
- 1990 KSR multiprocessor (1K mP)

1994 SNAP (with Jim Gray) Scaleout

2017 Scaleup on scaleout computers?

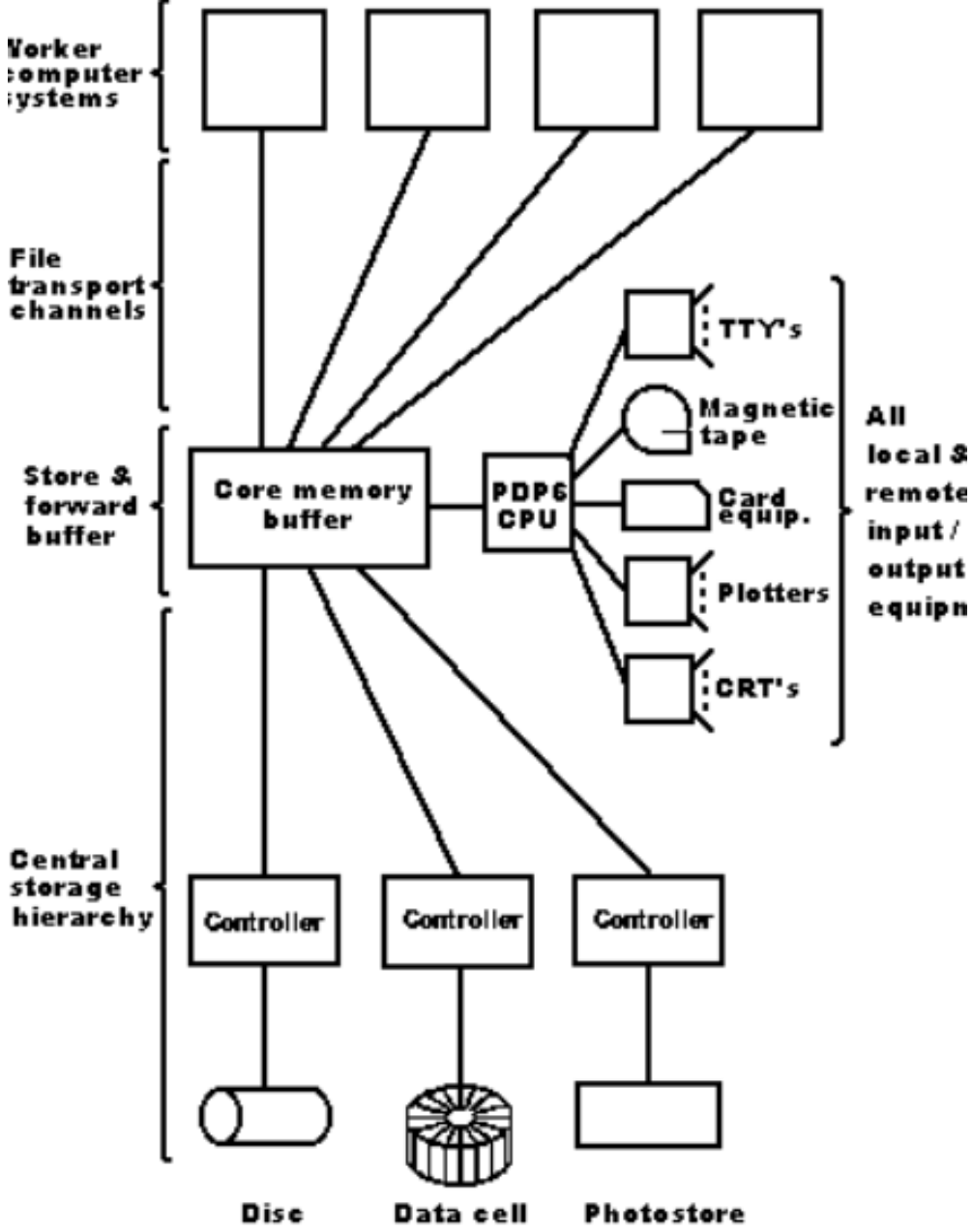


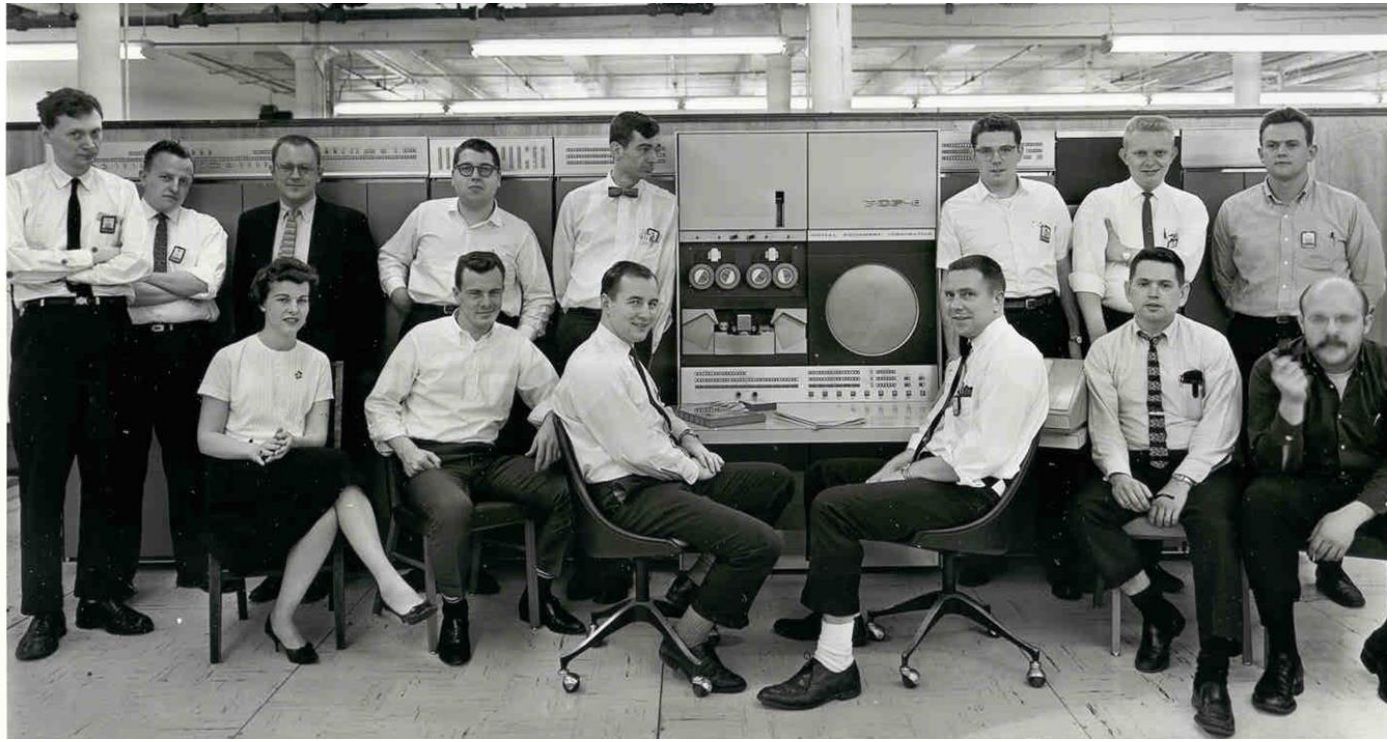
Figure 1

The Classical Octobus Concept

1965: LLNL Octopus hub

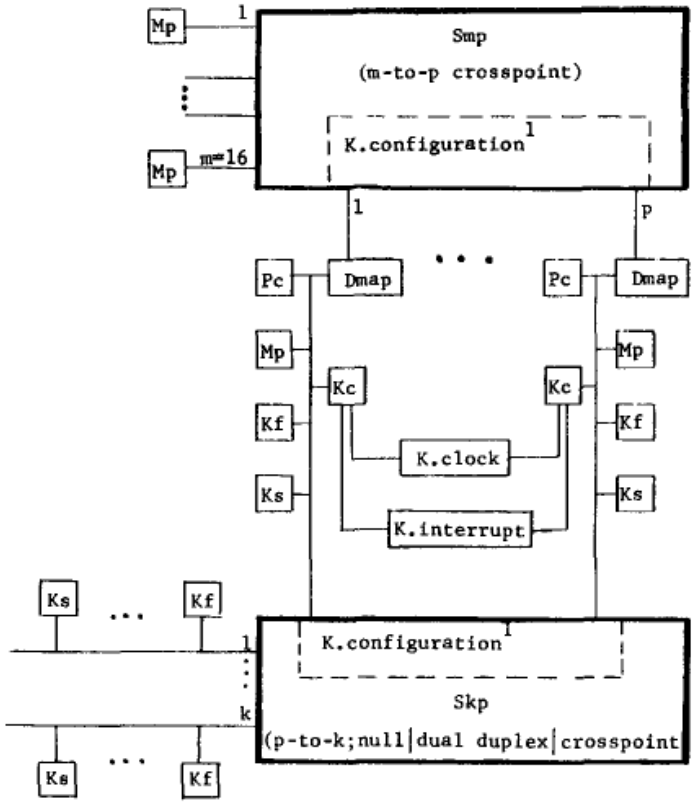
Digital Equipment PDP-6 Multiprocessor

256Kword, 36 bits/word



2 x 10 x 25 =500 modules; 5000 transistors?1900 watts

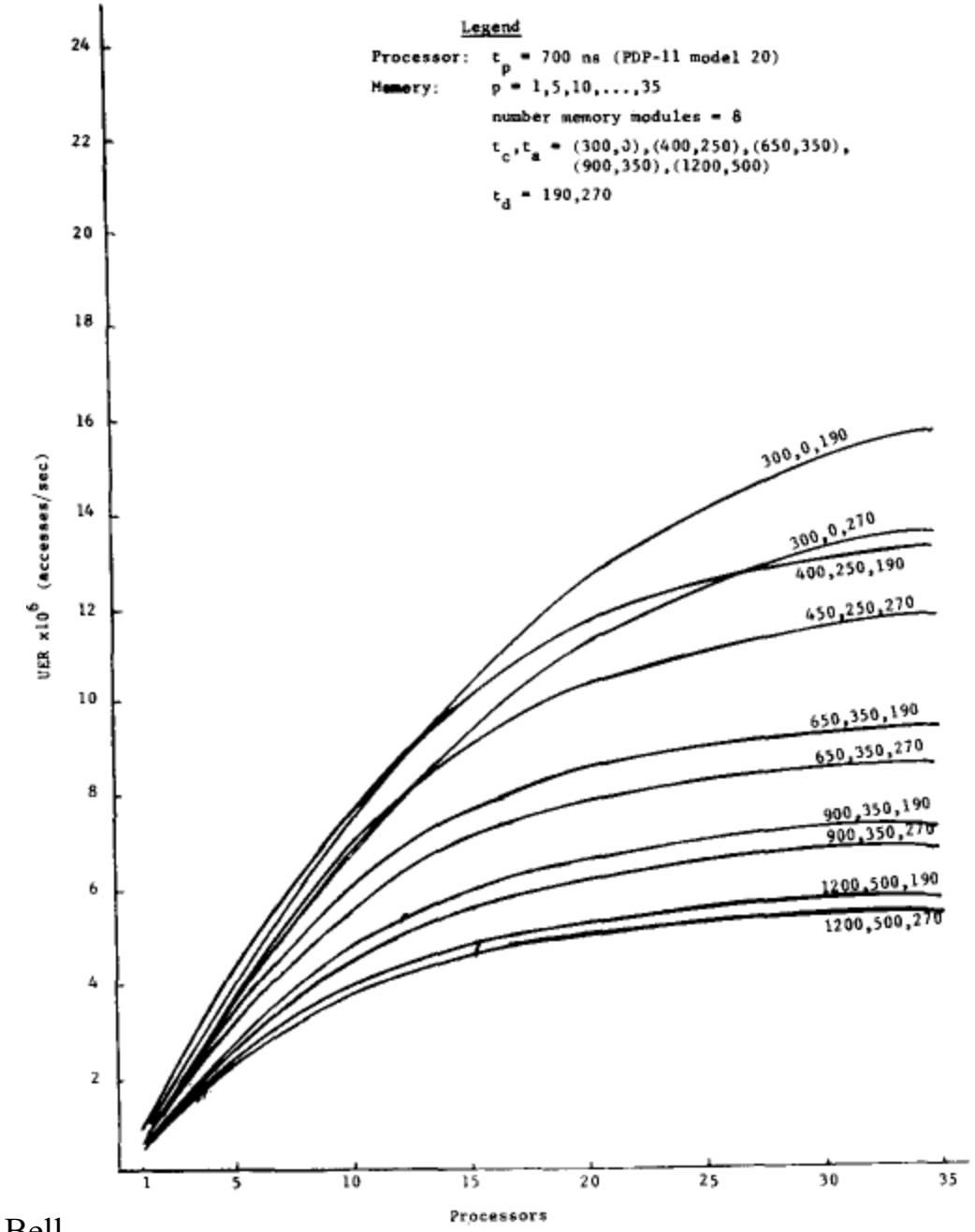
1972-1978 CMU C.mmp Discovered the need for multiple locks for scaleup …



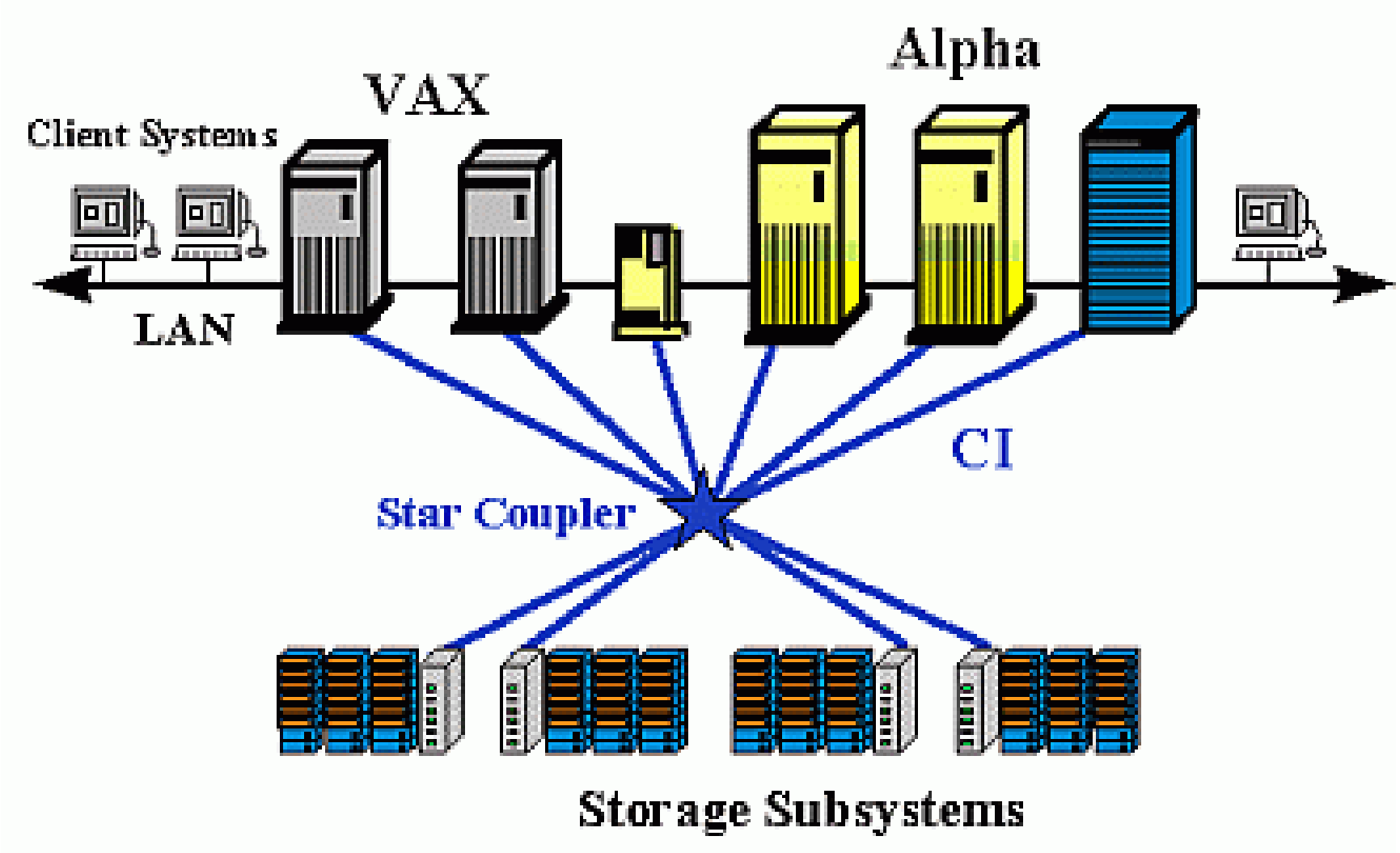
where: Pc/central processor; Mp/primary memory; T/terminals;
Ks/slow device control (e.g., for Teletype);
Kf/fast device control (e.g., for disk);
Kc/control for clock, timer, interprocessor communication

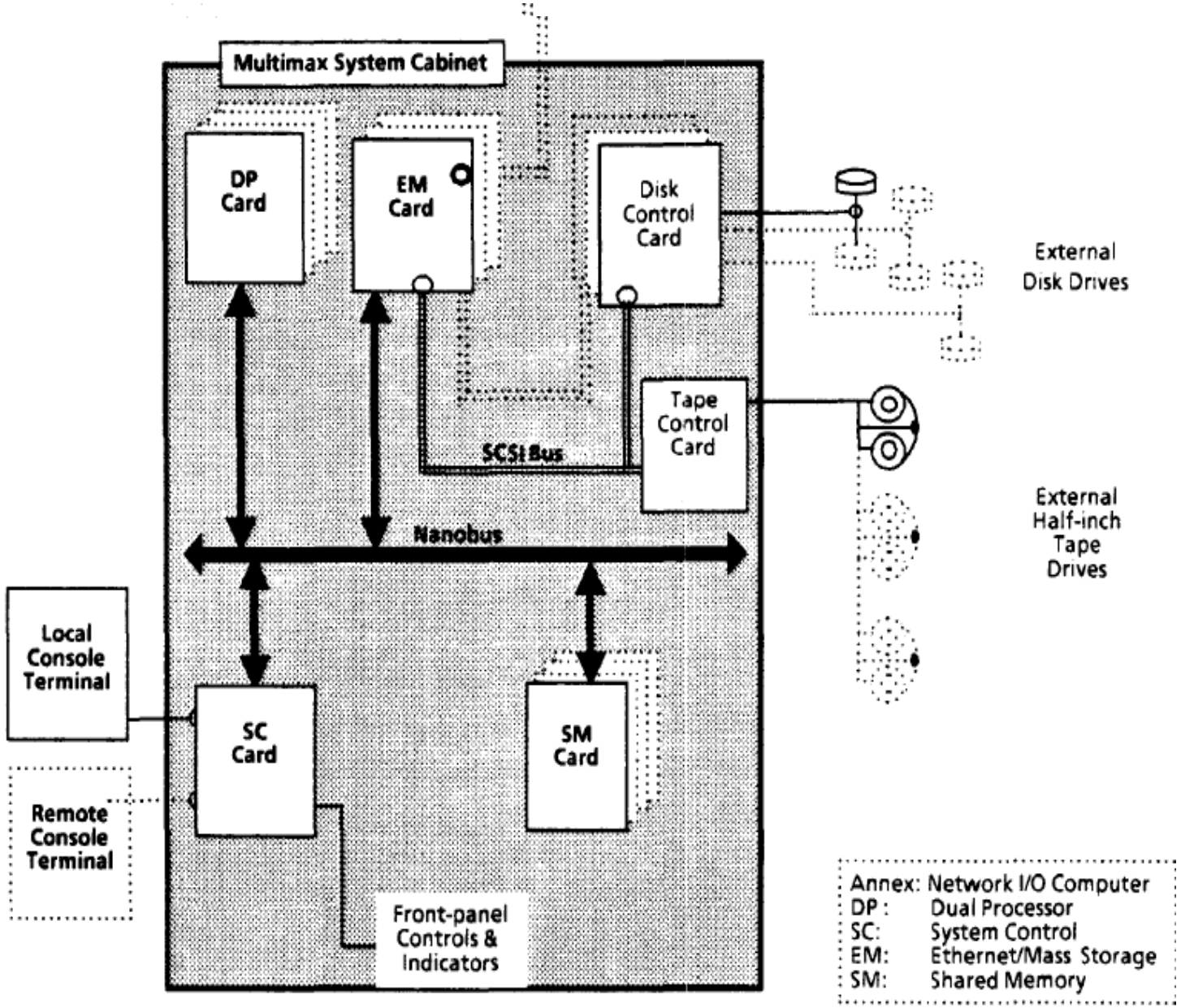
[1] Both switches have static configuration control by manual and program control

Figure 1—Proposed CMU multiminiprocessor computer/C.mmp



VAX Cluster c1983



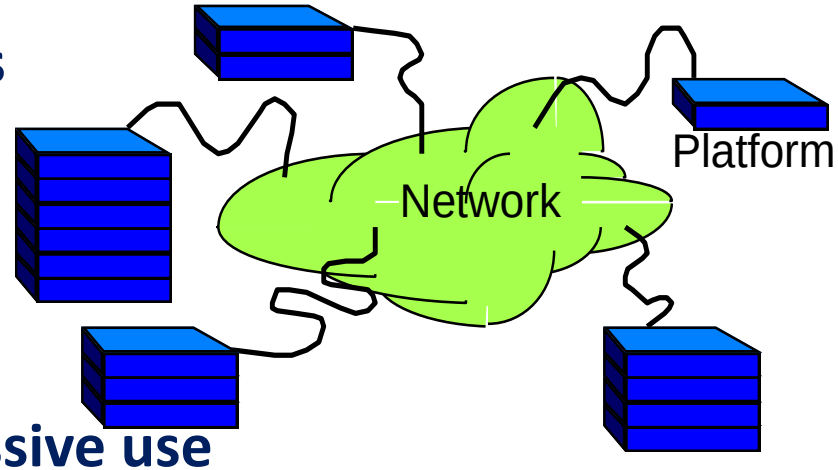


Encore Multimax, 1986, 15 years later "multi":
a shared memory multiprocessor with a single bus for memory and cached processors.

Figure 2–3

1994: Computers will All be Scalable

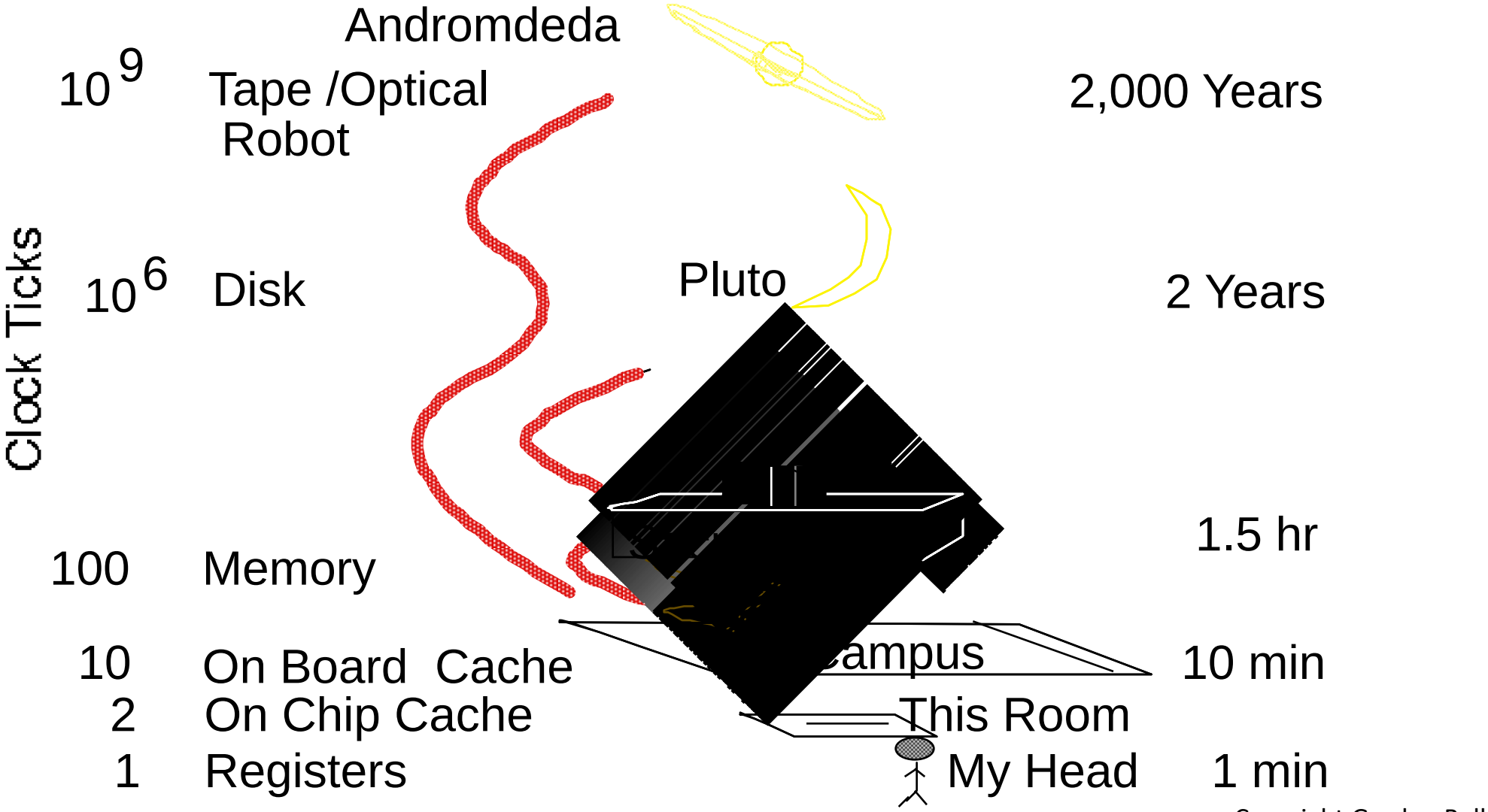
- Thesis: SNAP: Scalable Networks as Platforms
  - upsize from desktop to world-scale computer
  - based on a few standard components
- Because:
  - Moore's law: exponential progress
  - standards & commodities
  - stratification and competition
- When: Sooner than you think!
  - massive standardization gives massive use
  - economic forces are enormous



2017 A new approach: scaleup on scaleout

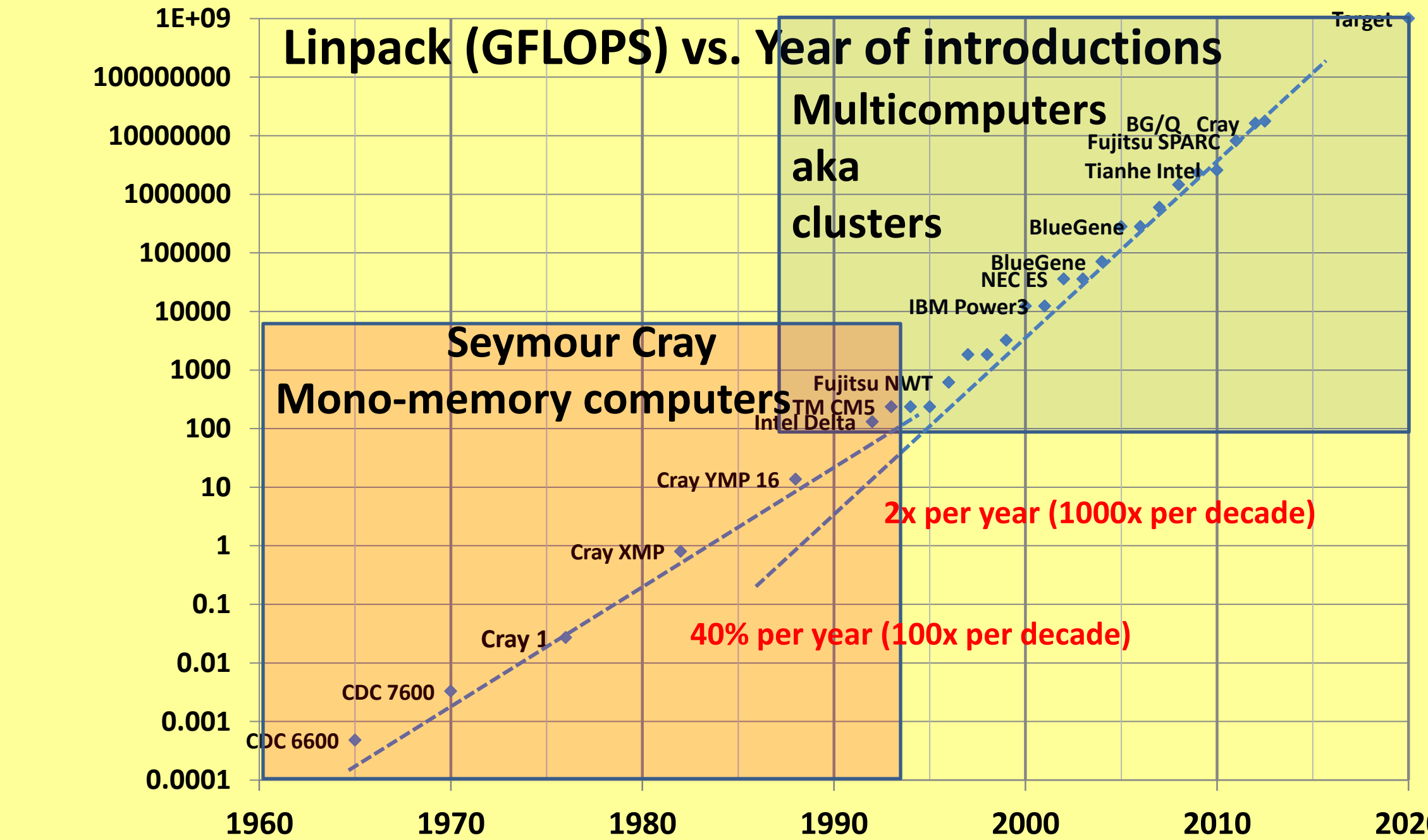
TidalScale: a single memory space aka mP on a multicomputer

Latency: How Far Away is the Data?



HPC aka Supercomputing ideas & machines: It's about speed, parallelism, standards, & cost (COTS)

1. The Cray Era (1965-1995): Mono-memory computer
   - Increased clock: high Khz => Mhz => low Ghz (10,000x)
   - Intra-processor parallelism techniques (10x > 100x vector)
   - Shared memory, multi-processor parallelism (1>32)
2. *"Killer Micros" transition (1984-1994)*
   - *Searching for the way… scalability, CMOS micros, standards, failures*
   - *Similar to search for the first supers, and for xxMD way.*
   - **1987… Prize offer to acknowledge parallelism**
3. The Multicomputer aka Clusters era (1984-present)
   - Parallelism: <10 => 1000x => 10,000 => 100,000 => million… billion
   - Now it is up to programmers to exploit parallelism



Top 20+ HPC seminal events: recipes, standards, plans, and prizes

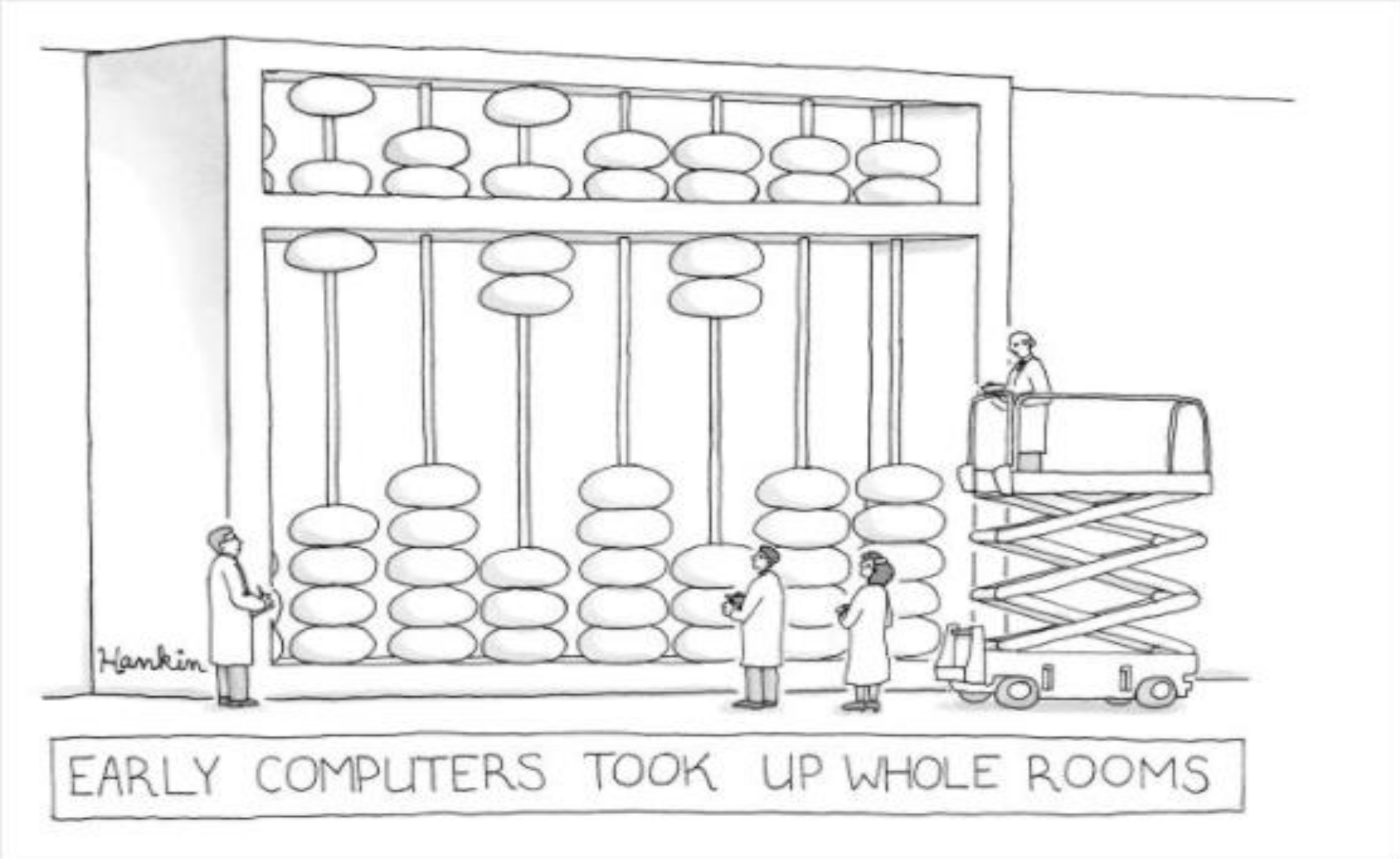
1. *1945,6 EDVAC Recipe and IAS Architecture. Over a dozen IACS were built e.g. Silliac*
2. 1957 *Fortran* first delivery establishes standard for scientific computing… FORTRAN 2008
3. 1960 LARC and 1961 Stretch—response to customer demand; 1962 Atlas commissioned
4. 1964 CDC 6600 (.48 MF) introduces parallel function units. Seymour establishes 30 year reign
5. 1965 Amdahl's Law: single processor vs multi-P's or vectors
6. 1976 Cray 1 (26 MF) introduces first, practical vector processor architecture.
7. 1982 Cray XMP… (PK: 1 GF) Intro of mP. The end of the beginning of mono-memory

8. 1982, 83 Caltech Cosmic Cube demo of multicomputer with 8, 64 computers. New beginning.
9. 1987 nCUBE (1K computers) 400-600 speedup, Sandia wins first *Bell Prize*.
   1988 Gustafson's Law as Amdahl's Law Corollary (Simon #1)
10. 1992 Intel Touchstone Delta at Sandia Reaches 100 GF
11. 1993 CM5 (60 GF Bell Prize) 1024 Sparc computers. Cray C90 was 16! No way or plan to compete
12. 1993 *Top500* established at prize using *LINPACK Benchmark. (Simon #10)* *Begin multicomputer era*
13. 1994 *Beowulf* kit and recipe for multicomputers and *MPI-1* Standard established

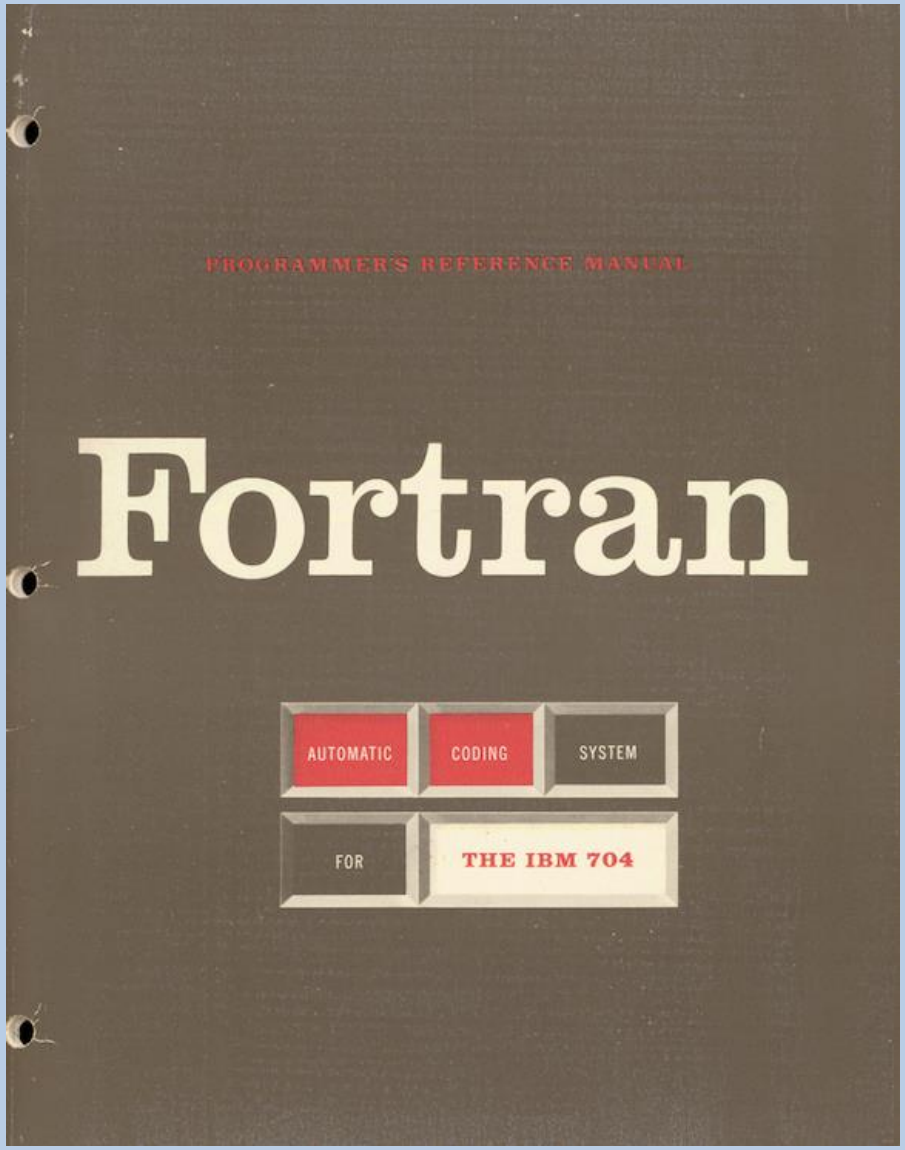
14. 1995 ASCI > Advanced Simulation and Computing (ASC) Program
15. 1996 Seymour R. Cray is killed in a car accident. Building a shared memory computer using itanium
16. 1997 ASCI Red (1 TF) at Sandia
17. 1999 The Grid: Blueprint for a New Computing Infrastructure (Simon #8)
18. 2008 IBM BlueGene (1.5 PF)
19. 2012 Cray Titan (17.6) GPU and CUDA
20. Tiahne-2 at NUDT, 2016 Sunlight achieves 93 PF with > 10M cores

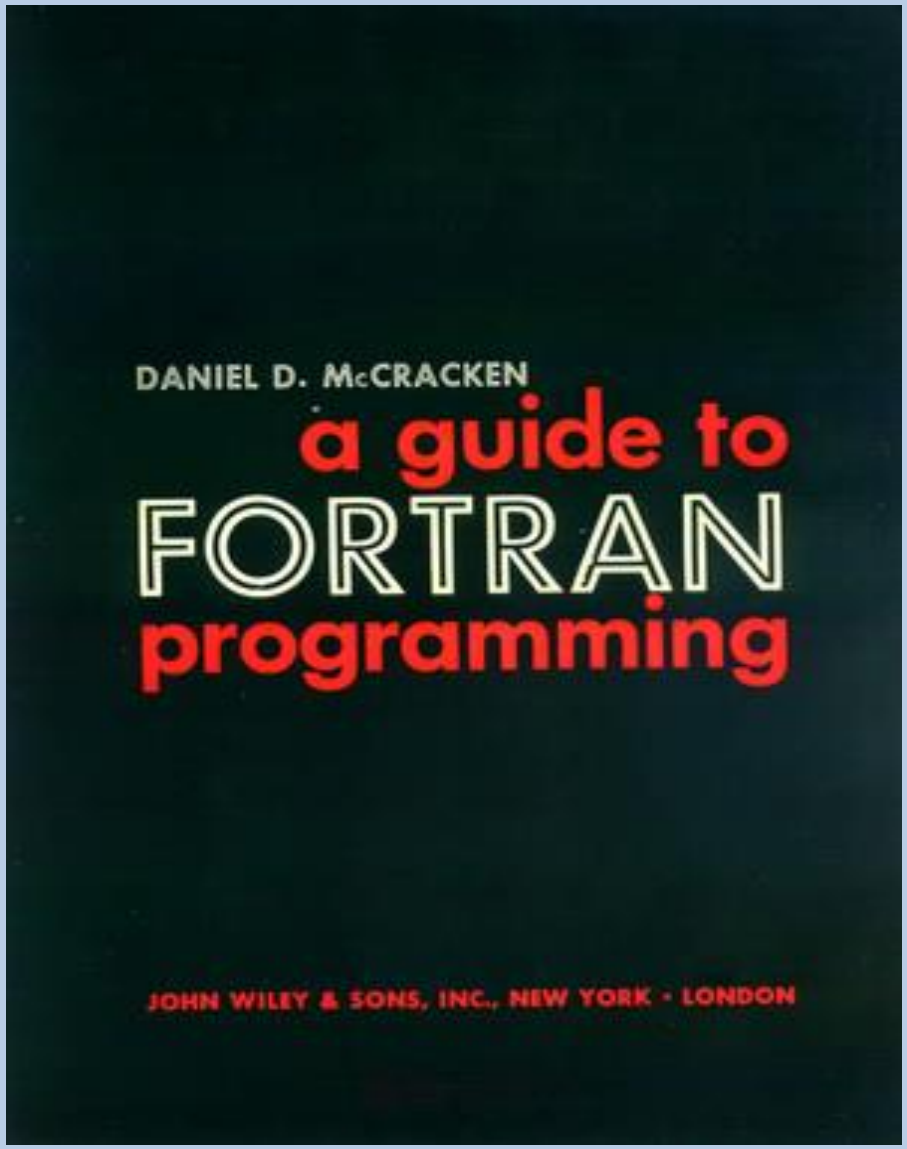
Ancient history, and how we got here

The history… and we started calling them supercomputers



Fortran 1957, ’60, … ‘08





CDC 6600 #1 Console & frame c1964 LLNL

First Supercomputer?

1970s search for parallelism constrained by Amdahl's Law

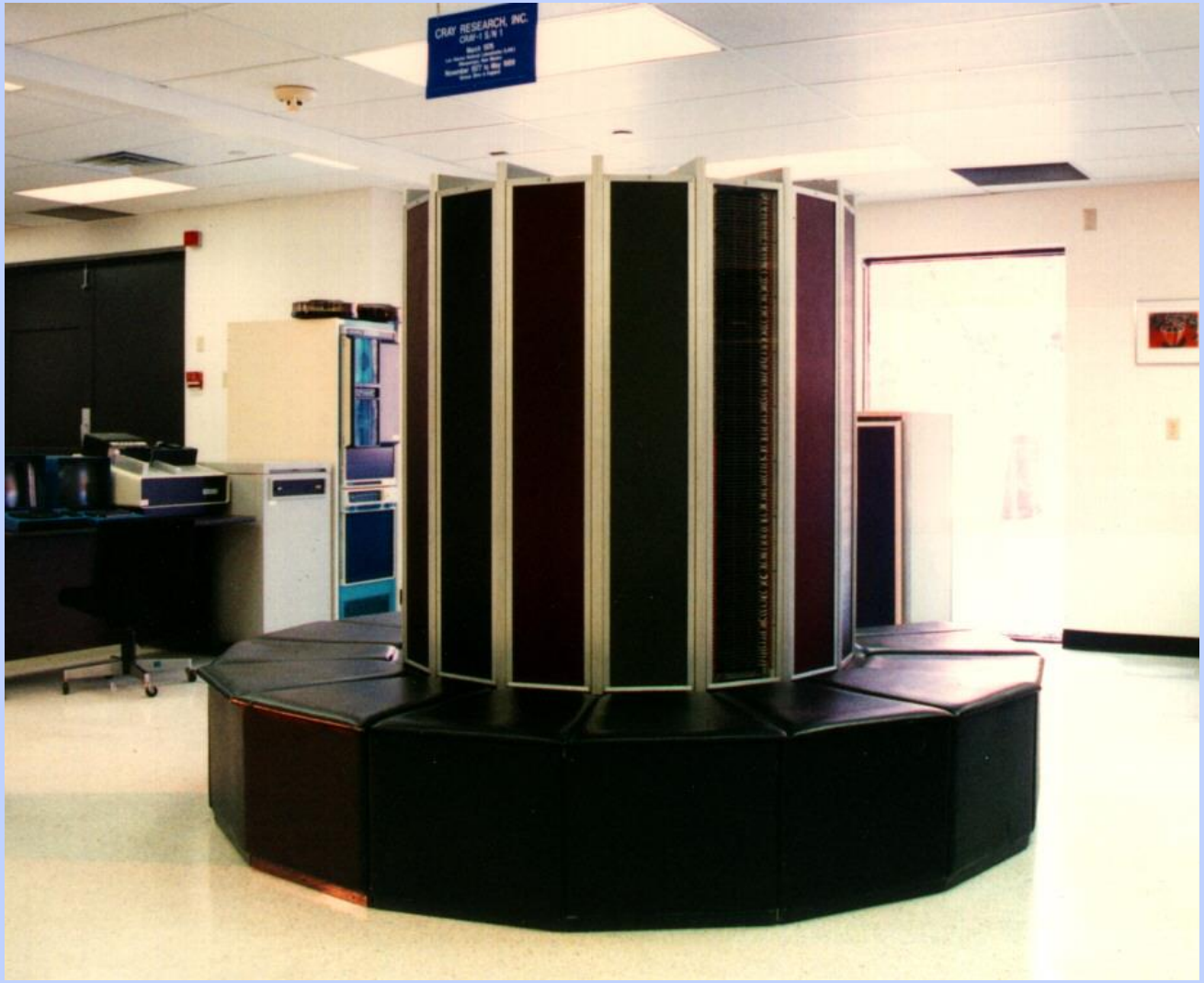
Three approaches… that didn't work

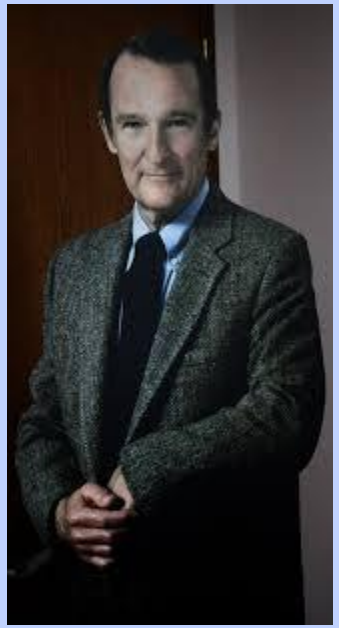
- **Illiac IV (SIMD) m=64**
- **CDC STAR and ETA10 vectors in memory**
- **TI ASC (Vector arch; Cray 1's 5x clock)**

Then success

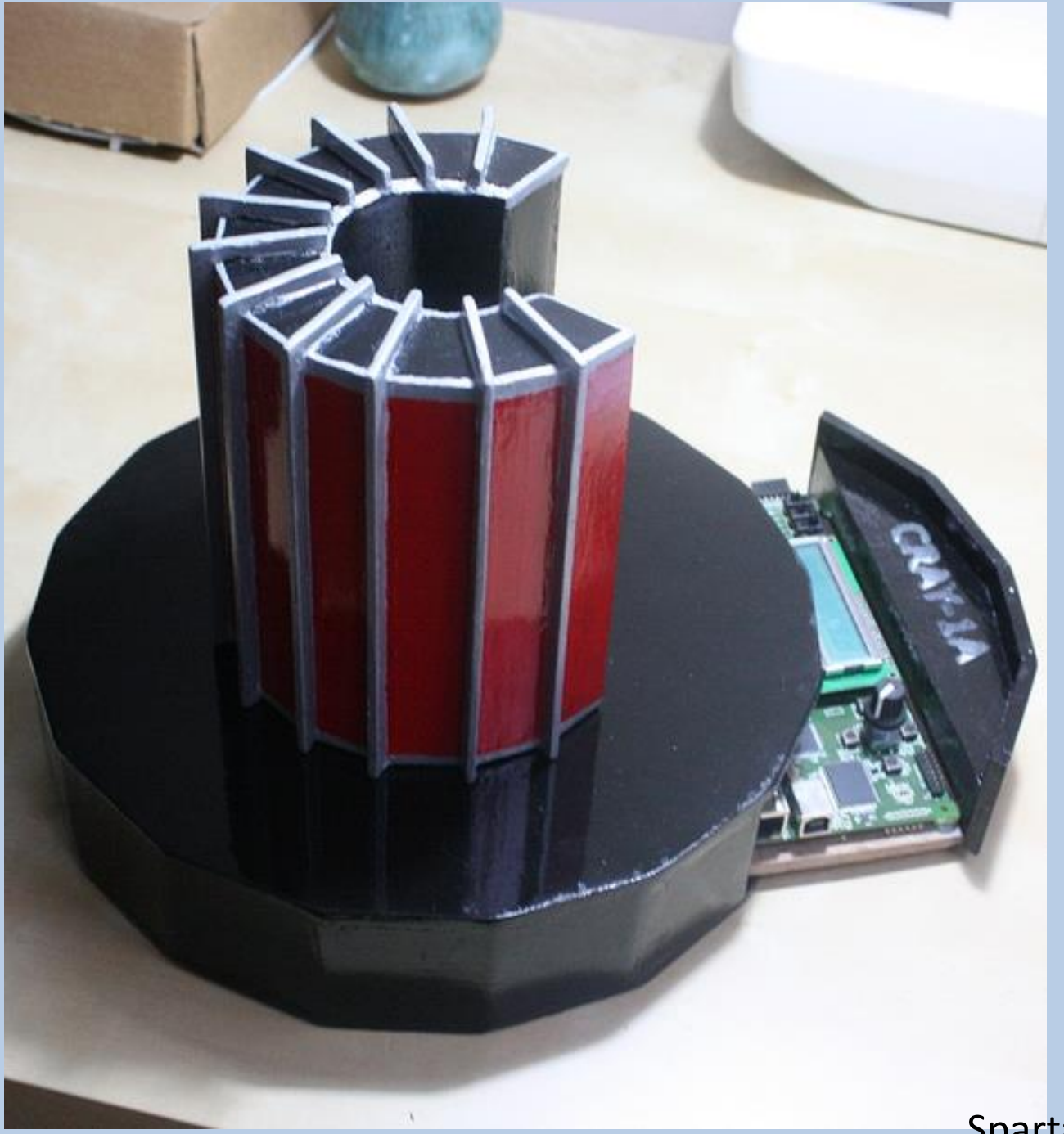
- **Cray 1 vector architecture**

Cray-1 c1976: Supercomputer … vector era

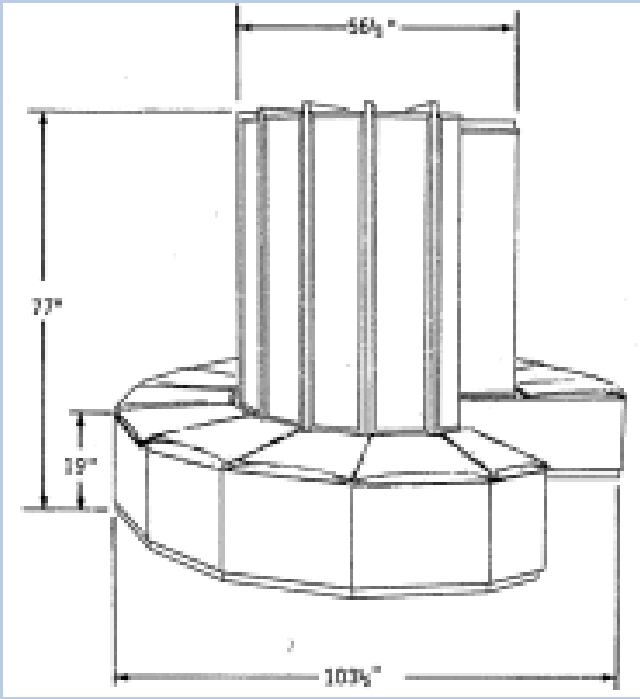




Courtesy of Burton Smith, Microsoft



Fenton-Tantos
Desktop Cray 1
or XMP
at 0.1 size.



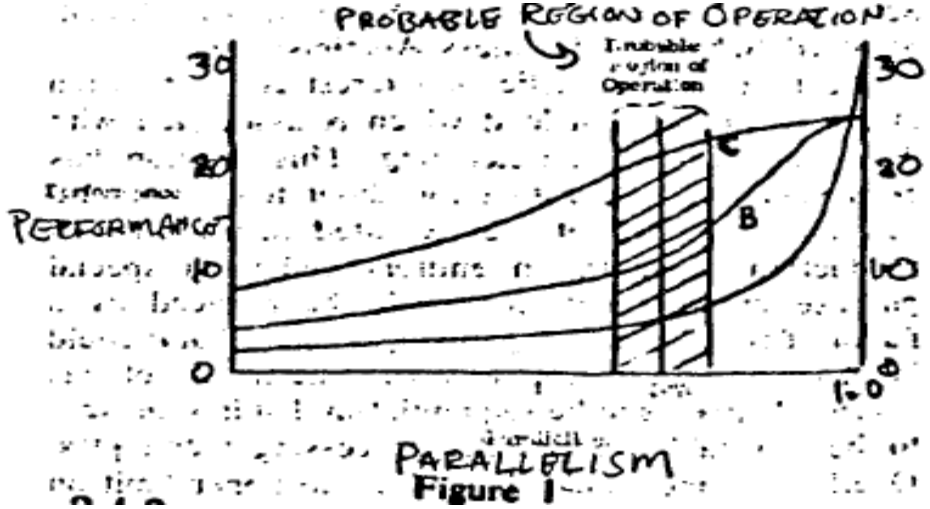
Spartan-3E-1600-Development-Board

1983-1993 transition to multicomputers: The search for the next supercomputer recipe

- **Break from single memory computers**
- **Factors that enabled the transition**
  - **Japanese 5th Generation (AI) threat**
  - **DARPA Response stimulates many startups**
  - **Powerful inexpensive "killer" CMOS micros cross-over performance of TTL and ECL proprietary processors**
  - **Clustered micros declared mono-memory alternative**
  - **Existence proof that "they work"**
  - **Incentive & Measurement: Bell Prize '87, Top500 '93**
  - **Standards to build i.e. Beowulf and program i.e. MPI**
  - **"Grand challenge" problems to justify funding for NBC**

Amdahl's law… the limit of parallelism

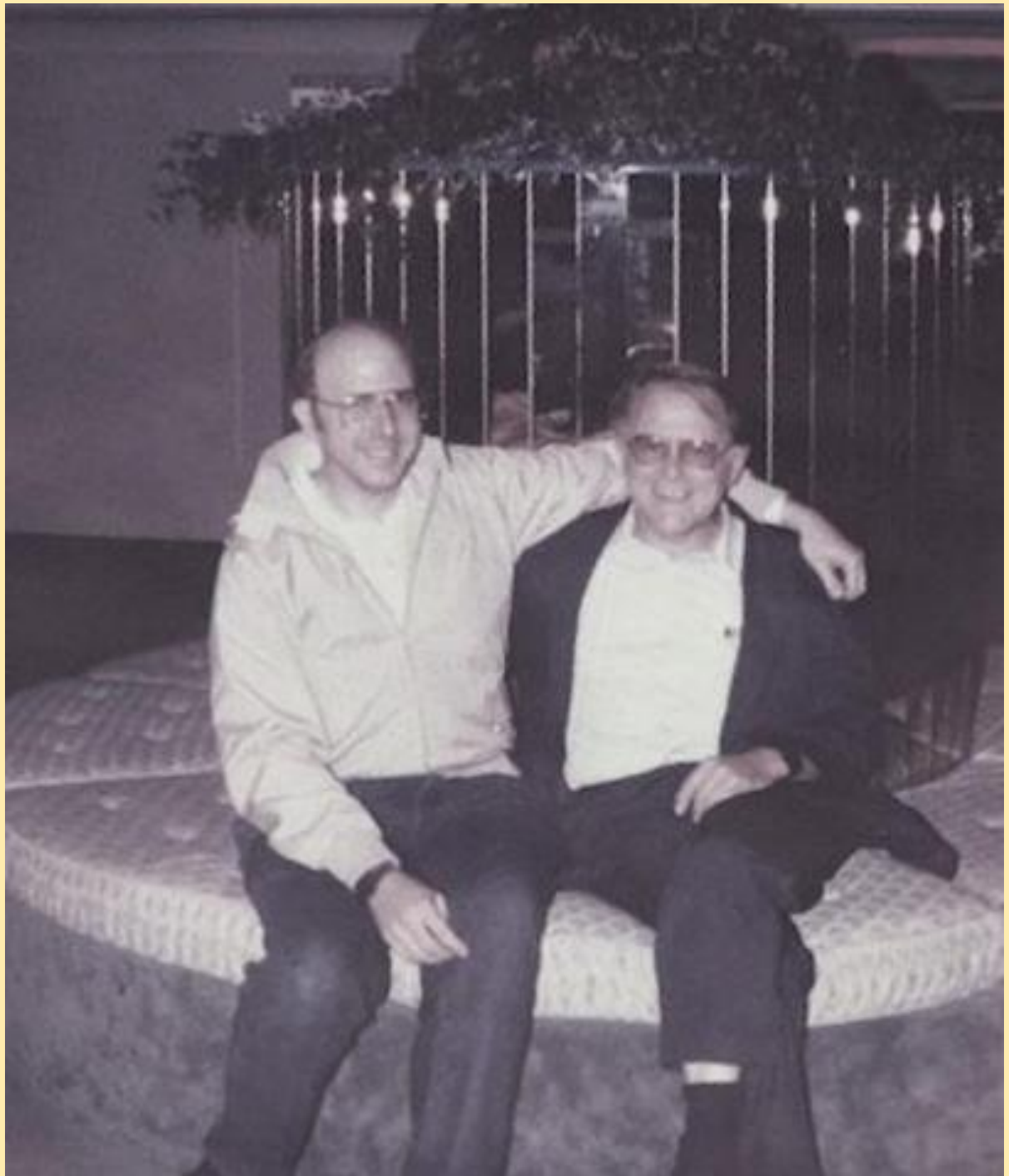
- If w_1 work is done at speed s_1 and w_2 at speed s_2, the average speed s is $(w_1+w_2)/(w_1/s_1 + w_2/s_2)$
  - This is just the total work divided by the total time
- For example, if w_1= 9, w_2= 1, s_1= 100, and s_2= 1 then s = 10/1.09 ≅ 9 (speed)



Amdahl, Gene M, "Validity of the single processor approach to achieving large scale computing capabilities", Proc. SJCC, AFIPS Press, 1967

Courtesy of Burton Smith, Microsoft

1983 Caltech Cosmic Cube
8 node prototype (‘82) & 64 node ‘83
Intel iPSC 64 Personal Supercomputer ‘85



Threat Response: 1984 Kickoff of DARPA's SCI Strategic Computing Initiative…

Steve Squires, DARPA & Gordon Bell, Encore seated at our "Cray".

<u>10+ years later:</u>
<u>"Killer micros" Clusters become standard</u>

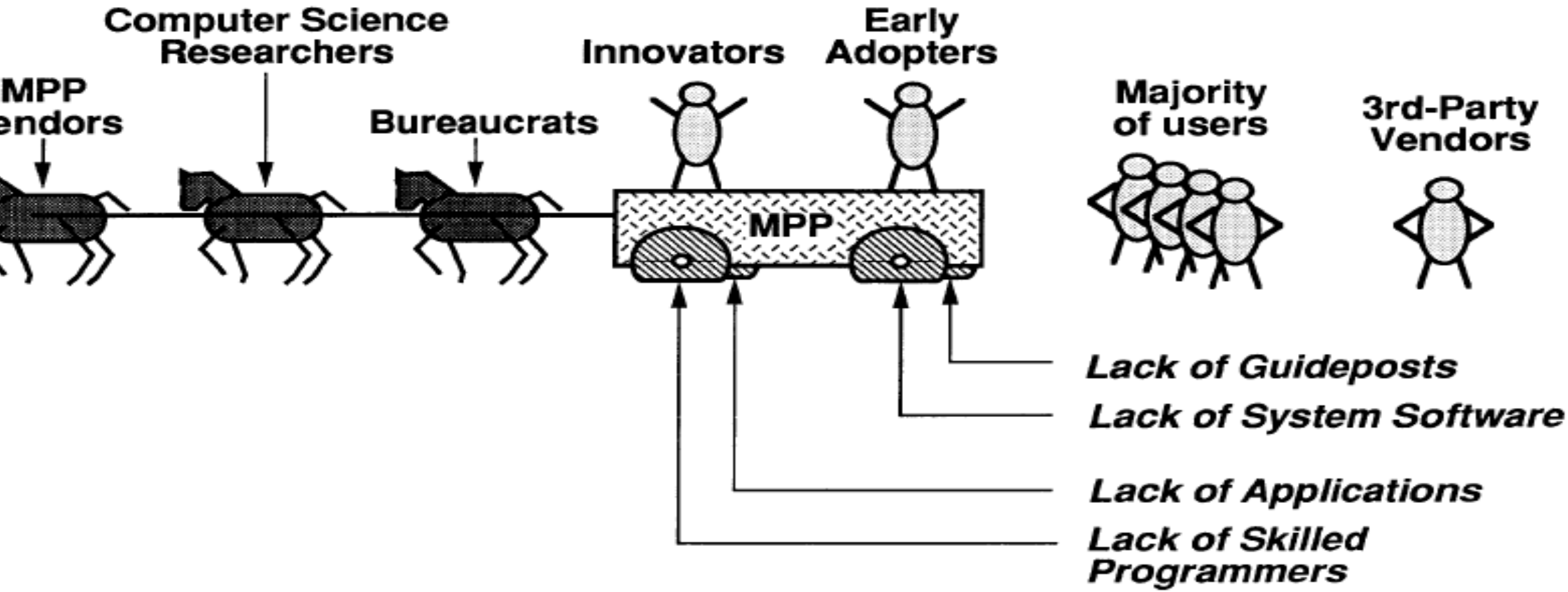
1989: The “killer micros” –Eugene Brooks, LLNL

Challenge: how do you utilize (program)
a large number of interconnected,

independent computers?

THE MPP BANDWAGON

Jack Worlton view c 1991



IEEE Software launches annual Gordon Bell Award

Editor-in-Chief Ted Lewis has announced the First Annual Gordon Bell Award for the most improved speedup for parallel-processing applications. The two $1000 awards will be presented to the person or team that demonstrates the greatest speedup on a multiple-instruction, multiple-data parallel processor.

One award will be for most speedup on a general-purpose (multiapplication) MIMD processor, the other for most speedup on a special-purpose MIMD processor. Speedup can be accomplished by hardware or software improvements, or by a combination of the two.

To qualify for the 1987 awards, candidates must submit documentation of their results by Dec. 1. The winners will be announced in the March 1988 issue. This year's judges are Alan Karp of IBM's Palo Alto Scientific Center, Jack Dongarra of Argonne National Laboratory, and Ken Kennedy of Rice University.

Bell Prize for Parallelism, July 1987

Benner, Gustafson, and Montry of Sandia
for
beam stress,
surface wave simulation,
unstable fluid flow

Development of Parallel Methods For a 1024-Processor Hypercube

John L. GUSTAFSON, Gary R. MONTRY, and Robert E. BENNER
Sandia National Laboratories, Albuquerque, New Mexico

March 1988

As printed in SIAM Journal on Scientific and Statistical Computing
Vol. 9, No. 4, July 1988, pp. 609–638.
(Minor revisions have been made for the Web page presentation of this paper. JLG 1995)

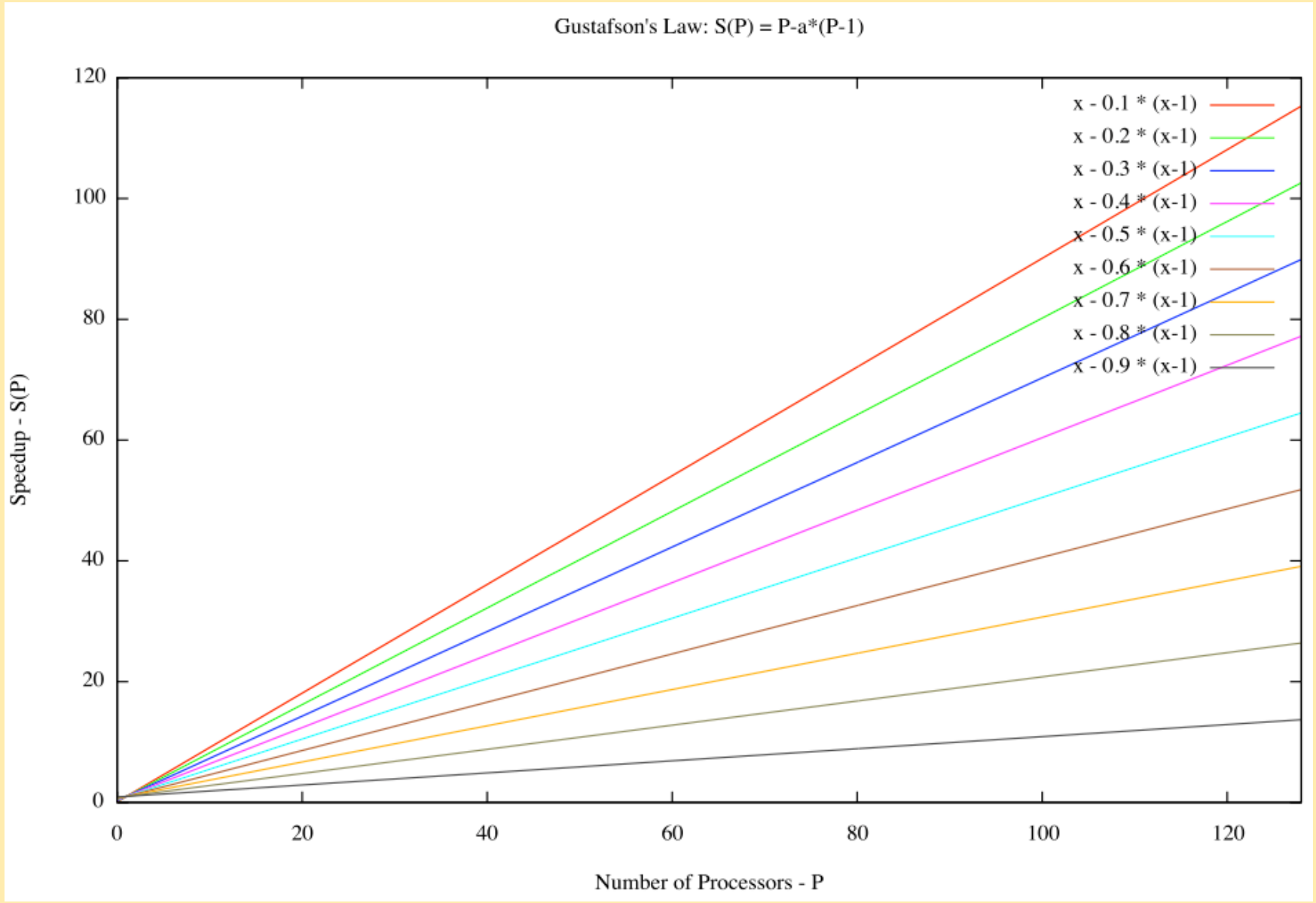
EDITOR'S NOTE

[This paper] reports on the research that was recognized by two awards, the Gordon Bell Award and the Karp Prize, at IEEE's COMPCON 1988 meeting in San Francisco on March 2.

The Gordon Bell Award recognizes the best contributions to parallel processing, either speedup or throughput, for practical, full-scale problems. Two awards were proposed by Dr. Bell: one for the best speedup on a general-purpose computer and a second for the best speedup on a special-purpose architecture. This year the two awards were restructured into first through fourth place awards because of the nature of the eleven December 1987 submissions. Bell presented the first place award of $1,000 to the authors of [this paper].

1988 Gustafson's Law

Benner, Gustafson, Montry winners of first Gordon Bell Prize



S(P) = P –α x (P-1)

P, **number of processors,**

S, **SPEEDUP, and**

α, non-parallelizable fraction of parallel process

1994: MPI 1.0 Massage Passing Interface

MPI: A Message-Passing Interface Standard
Version 3.0

Message Passing Interface Forum

September 21, 2012

Version 3.0: September 21, 2012. Coincident with the development of MPI-2.2, the MPI Forum began discussions of a major extension to MPI. This document contains the MPI-3 Standard. This draft version of the MPI-3 standard contains significant extensions to MPI functionality, including nonblocking collectives, new one-sided communication operations, and Fortran 2008 bindings. Unlike MPI-2.2, this standard is considered a major update to the MPI standard. As with previous versions, new features have been adopted only when there were compelling needs for the users. Some features, however, may have more than a minor impact on existing MPI implementations.

Version 2.2: September 4, 2009. This document contains mostly corrections and clarifications to the MPI-2.1 document. A few extensions have been added; however all correct MPI-2.1 programs are correct MPI-2.2 programs. New features were adopted only when there were compelling needs for users, open source implementations, and minor impact on existing MPI implementations.

Version 2.1: June 23, 2008. This document combines the previous documents MPI-1.3 (May 30, 2008) and MPI-2.0 (July 18, 1997). Certain parts of MPI-2.0, such as some sections of Chapter 4, Miscellany, and Chapter 7, Extended Collective Operations, have been merged into the Chapters of MPI-1.3. Additional errata and clarifications collected by the MPI Forum are also included in this document.

Version 1.3: May 30, 2008. This document combines the previous documents MPI-1.1 (June 12, 1995) and the MPI-1.2 Chapter in MPI-2 (July 18, 1997). Additional errata collected by the MPI Forum referring to MPI-1.1 and MPI-1.2 are also included in this document.

Version 2.0: July 18, 1997. Beginning after the release of MPI-1.1, the MPI Forum began meeting to consider corrections and extensions. MPI-2 has been focused on process creation and management, one-sided communications, extended collective communications, external interfaces and parallel I/O. A miscellany chapter discusses items that do not fit elsewhere, in particular language interoperability.

Version 1.2: July 18, 1997. The MPI-2 Forum introduced MPI-1.2 as Chapter 3 in the standard "MPI-2: Extensions to the Message-Passing Interface", July 18, 1997. This section contains clarifications and minor corrections to Version 1.1 of the MPI Standard. The only new function in MPI-1.2 is one for identifying to which version of the MPI Standard the implementation conforms. There are small differences between MPI-1 and MPI-1.1. There are very few differences between MPI-1.1 and MPI-1.2, but large differences between MPI-1.2 and MPI-2.

Version 1.1: June, 1995. Beginning in March, 1995, the Message-Passing Interface Forum reconvened to correct errors and make clarifications in the MPI document of May 5, 1994, referred to below as Version 1.0. These discussions resulted in Version 1.1. The changes from Version 1.0 are minor. A version of this document with all changes marked is available.

Version 1.0: May, 1994. The Message-Passing Interface Forum (MPIF), with participation from over 40 organizations, has been meeting since January 1993 to discuss and define a set

Beowulf: Computer Cluster by Don Becker & Tom Sterling, NASA 1994

BSD, LINUX, Solaris, and Windows Support for MPI and PVM

Lessons from Beowulf

- An experiment in parallel computing systems ‘92
- Established vision- low cost high end computing
- Demonstrated effectiveness of PC clusters for some (not all) classes of applications
- Provided networking software
- Provided cluster management tools
- Conveyed findings to broad community
- Tutorials and the book
- Provided design standard to rally community!
- Standards beget: books, trained people, software … virtuous cycle that allowed apps to form
- Industry began to form beyond a research project

Courtesy, Thomas Sterling, Caltech.

Not so grand challenge

1991 Bet
"By 1996 supercomputing will be done predominately with >1000 processors"
–D. Hillis,

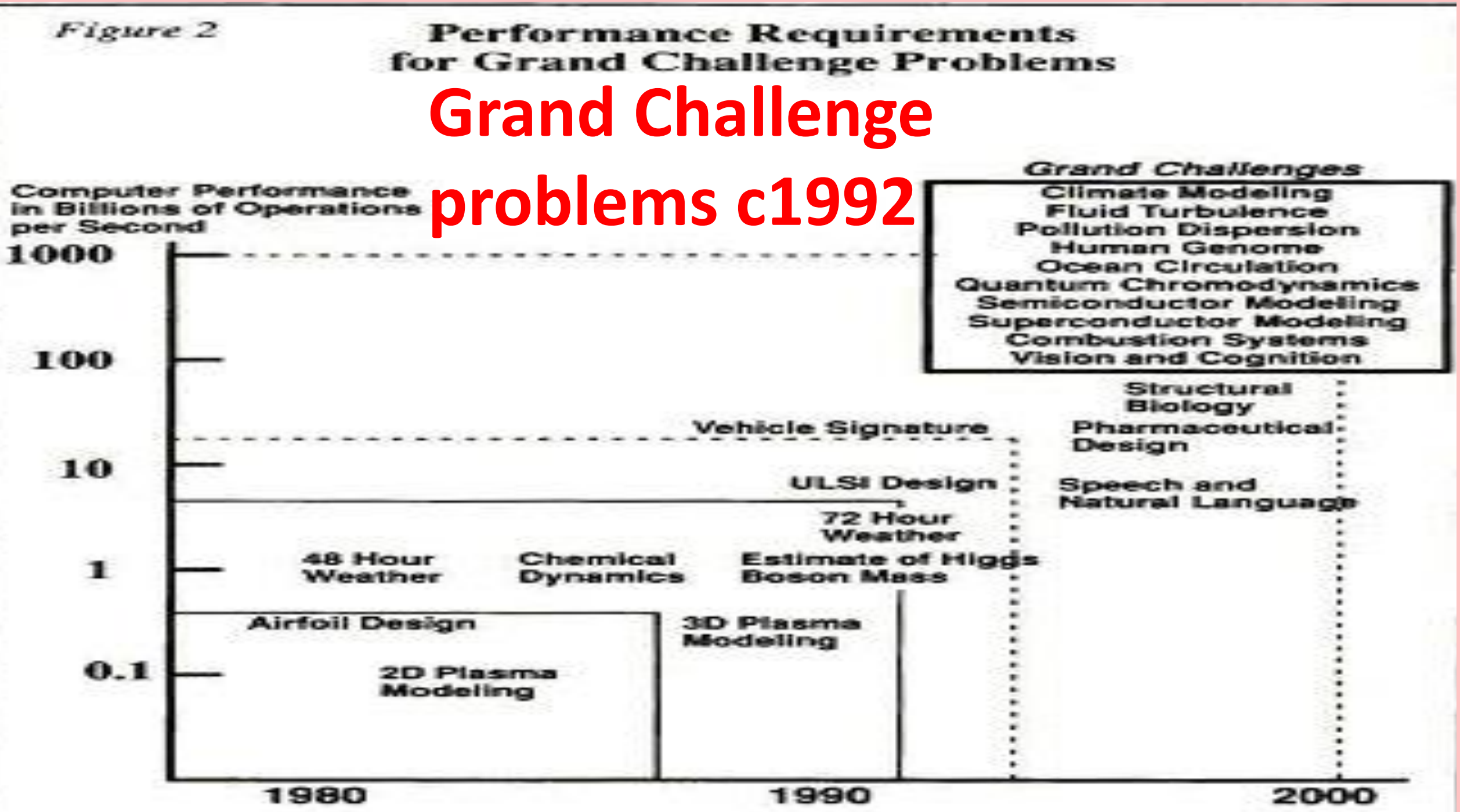
SUPERCOMPUTERS
Danny Hillis Settles Bet With Gordon Bell

Lost: The search for parallelism c1983-1997

DOE and DARPA Adv. Sci Comp. Initiative

- ACRI *French-Italian program*
- Alliant *Proprietary Crayette*
- American Supercomputer
- Ametek
- Applied Dynamics
- Astronautics
- BBN
- CDC >ETA *ECL transition*
- Cogent
- Convex > HP
- Cray Computer > SRC *GaAs flaw*
- Cray Research > SGI > Cray *Manage*
- Culler-Harris
- Culler Scientific *Vapor…*
- Cydrome *VLIW*
- Dana/Ardent/Stellar/Stardent
- Denelcor
- Encore
- Elexsi
- ETA Systems *aka CDC;Amdahl flaw*
- Evans and Sutherland Computer
- Exa
- Flexible
- Floating Point Systems *SUN savior*
- Galaxy YH-1
- Goodyear Aerospace MPP *SIMD*
- Gould NPL
- Guiltech
- Intel Scientific Computers
- International Parallel Machines
- Kendall Square Research
- Key Computer Laboratories *searching again*
- MasPar
- Meiko
- Multiflow
- Myrias
- Numerix
- Pixar
- Parsytec
- nCUBE
- Prisma
- Pyramid *Early RISC*
- Ridge
- Saxpy
- Scientific Computer Systems (SCS)
- Soviet Supercomputers
- Supertek
- Supercomputer Systems
- Suprenum
- Tera > Cray Company
- Thinking Machines
- Vitesse Electronics
- Wavetracer *SIMD*

Grand Challenge problems c1992



Evolution of Bell Prize winner applications

- ‘87 Beam Stress Analysis,
- Surface Wave Simulation,
- Unstable fluid flow model
- Global Ocean model
- ‘88 QCD
- Circuit Simulation
- Nonlinear network optimization
- Fluid flow (spectral method)
- Financial simulation
- Seismic data processing
- Oil reservoir modelling
- DNA Sequence matching
- Structure, high temp., superconductor
- Grid generation for PDE solution
- Parallelizing Pascal code
- ‘92 Simulating 9 M gravitation stars
- Grape: Motion of 100 K, 780 K stars
- Ultrascalable implicit finite element analyses in solid mechanics with over half a billion degrees of freedom
- K computer astrophysical N-body simulation of the gravitational trillion-body problem
- 42 TFlops hierarchical N-body simulations on GPUs in both astrophysics and turbulence
- Molecular dynamics including Anton 1, 2
- Shock front wave simulation
- Air flow
- Bio fluidics. Blood flow of complex structure
- An extreme-scale implicit solver for complex PDEs: highly heterogeneous flow in earth’s mantle
- Cactus and Globus execution in heterogeneous distributed computing environments

Bell Prize winner orgs.

7 *Caltech*
7 IBM
6 LANL
6 LLNL
6 ORNL
6 Sandia
5 Earth Sim. Ctr.

- 3 Japan AEC
- 3 NEC
- 3 Cray
- *3 ETH*
- 3 Japan Marine
- 3 NAL
- *3 NYU*
- *3 U Tokyo*
- *3 Yale*

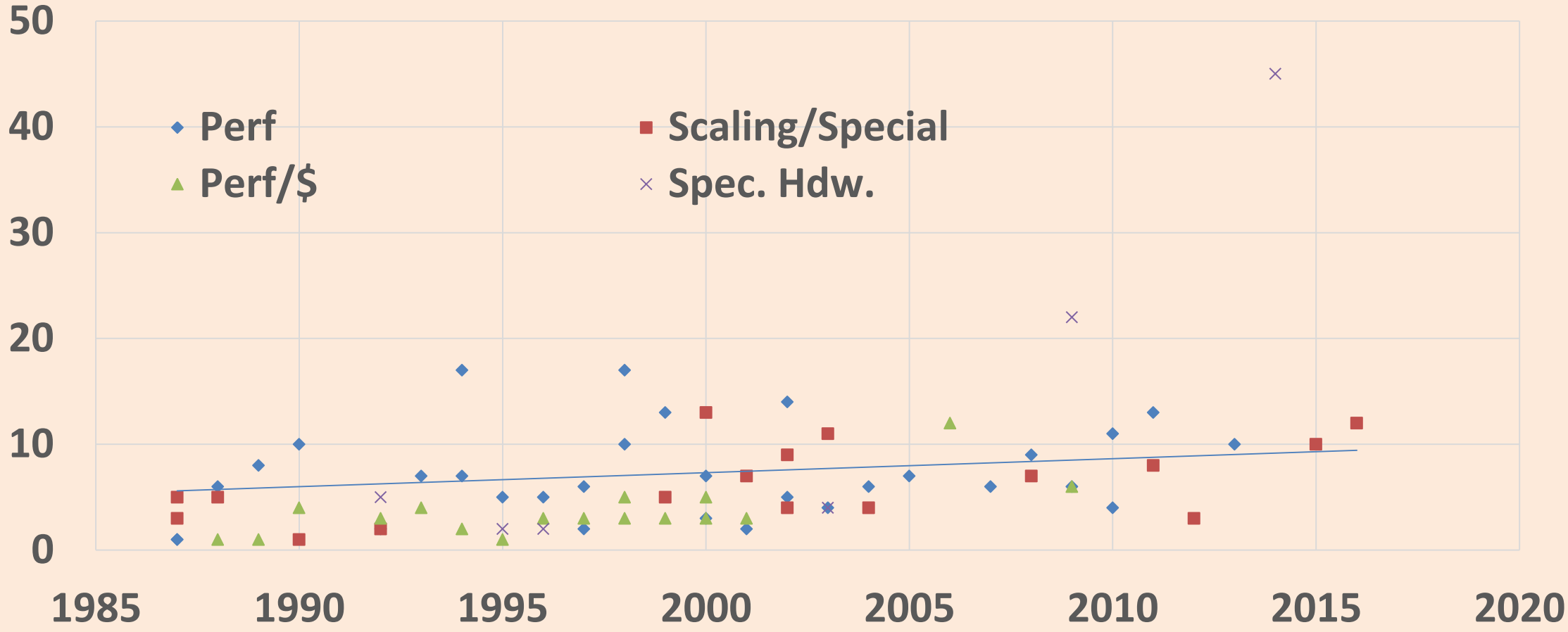
- 4 Argonne
- 4 Intel
- 4 Riken

- *2 CMU*
- *2 Cornell*
- 2 D E Shaw
- *2 FSU*
- 2 Fujitsu
- 2 HIT
- 2 JAMSTEC
- 2 Max Planck Inst
- 2 Mobil
- *2 Nagoya U*
- 2 Pittsburg SC
- 2 THC
- 2 TMC
- *2 Tokyo Inst of Tech*
- *2 Tokyo U*
- *2 Tskuba U*
- *2 U Colorado*
- *2 U MN*
- *2 U TX*
- *2 UC/Berkeley*

- 118

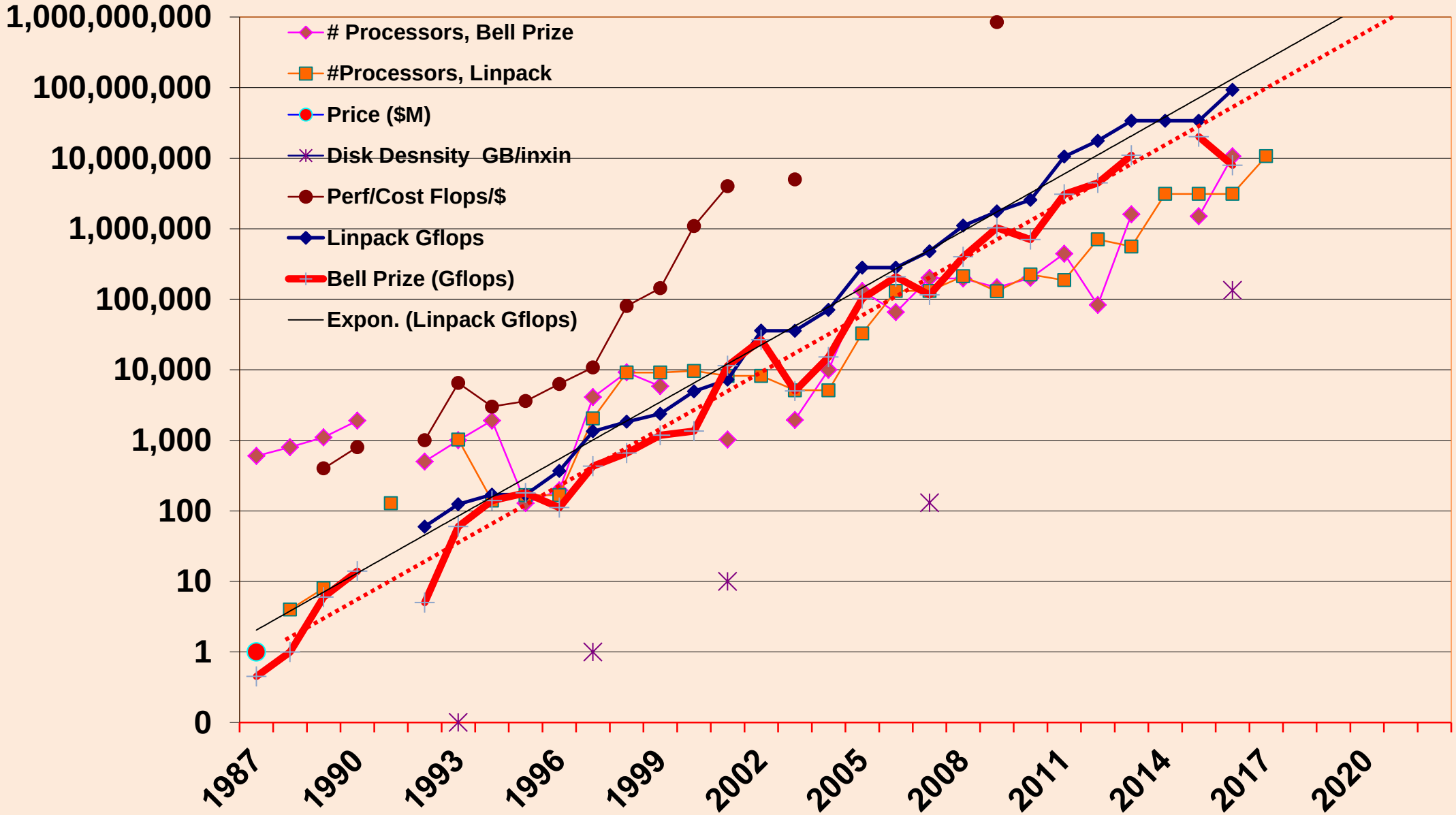
Abuques
Ansoft
Bejing Normal U
Brown
BTL
Center of Earth System
Columbia
Emory
Fermilab
Found MST
GA Tech
Hiroshima U
HNC
IDA
Inst fir Frontier Res
Japan NAL
Keio U
LBNL
MIT
Munich Tech U
Nagoya U
NAO Japan
NASA Ames
NASA Goddard
NASA Langley
Nat Space Dev
NCAR
Next Gen SC
NRL
NSC
Ohio State
Old Dominion U
Penn State
Purdue
Rutgers
Sandia
Tel Aviv U
THC
Traco
Tsingua U
U Chicago
U de Louviain
U IL
U Messina
U MI
U Milano
U NM
U of Bristol
U of Chinese Acad
U of Electro-communication
U of IL
U of TN
U Penn
U Sydney
UC/Davis
United Tech
Vienna U of Tech
Wills Phy Lab
Yamagata U

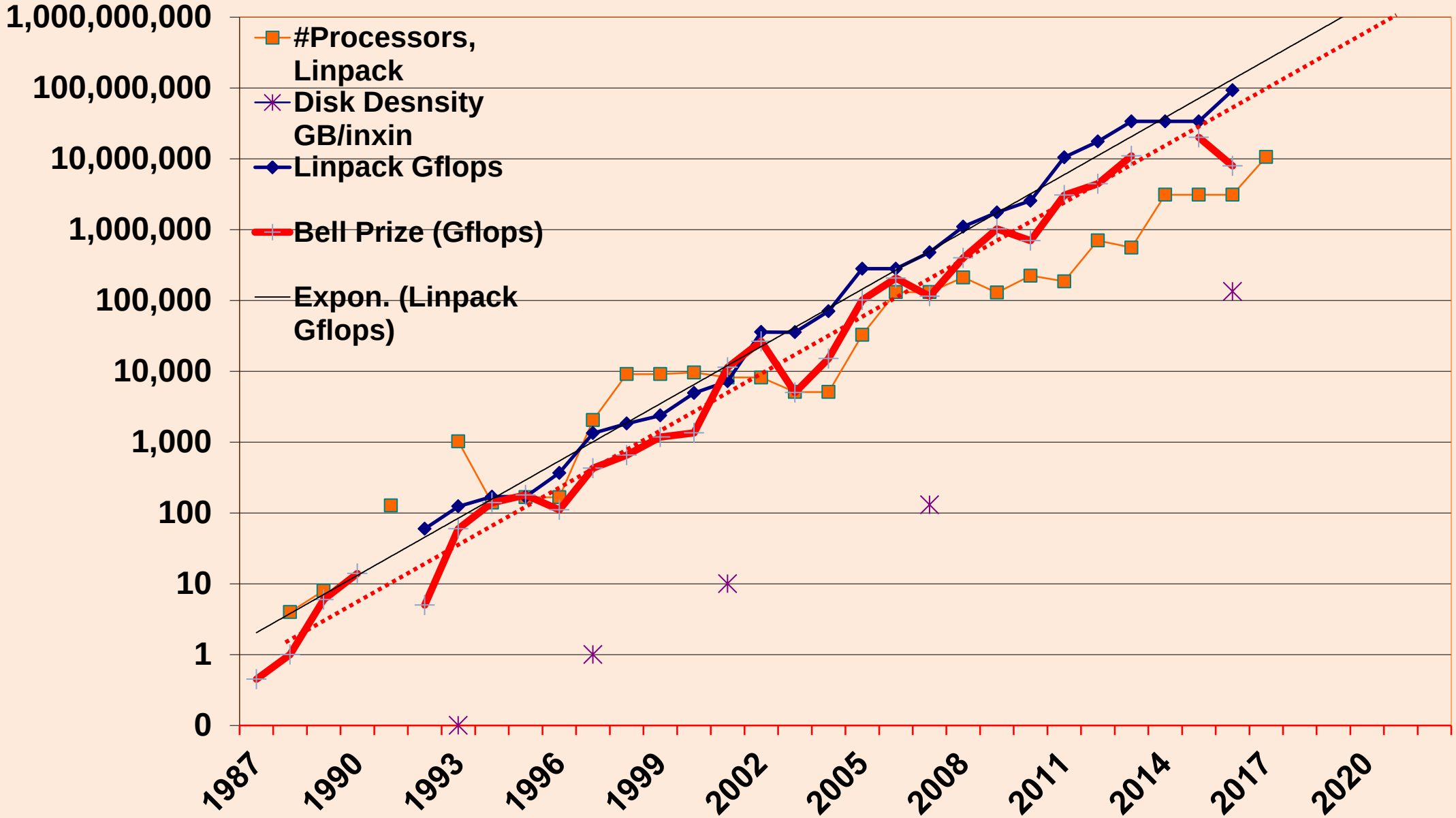
GBELL PRIZE WINNER TEAM SIZE

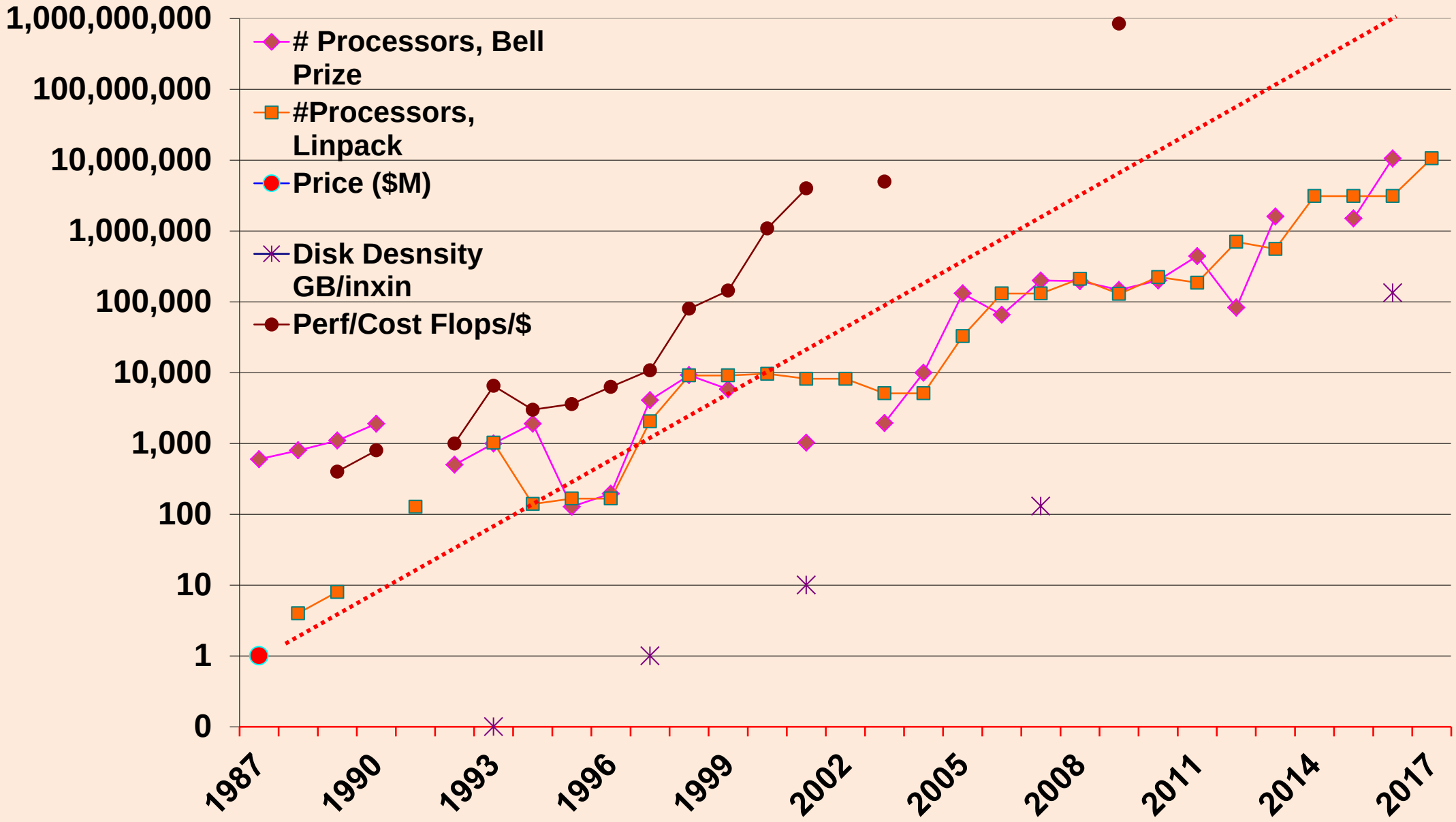


GBell Prize Platforms 1987-2016

Year	Parallelism	M 1	M 2
1987 Mf	600	Ncube 1K	Cray XMP
1988 Gf	800	Cray YMP	Ncube, iPSC
1989	1100	*TM CM2, 1K Pe*	PLA
1990	16,000	*TM CM2, 16K*	iPSC 860
1991	-		-
1992	500	Intel Delta	Dist. WS
1993	1000	TM CM5	SNAP
1994	1904	Intel Paragon	*Cluster WS*
1995	128	Fujitsu NWT	*Grape P:288*
1996	196	Fujitsu NWT	*Grape 1296*
1997	4096	ASCI Red	Cluster Alpha
1998	9200	Cray T3E/1024PE	ASCI Red
1999 Tf	5832	Blue Pacific	*Grape 5*
2000		*Grape 6*	Cluster WS
2001	1024	*IBM 16 mP cluster*	*Distributed Cs*
2002		NEC ES	ASCI White
2003	1944	NEC ES	
2004	4096	NEC ES	*1.34TF Grape*
2005	131,072	BlueGene/L LLNL	
2006	65,536	BlueGene	*MD Grape*
2007	200,000	BlueGene	
2008	196,000	Jaguar, Oak Ridge	-
2009 Pf	147,464	Jaguar, Oak Ridge	*Anton 1; GPUs*
2010	200,000	Jaguar, Oak Ridge	
2011	442,368	Fujitsu K	CPU:16K/GPU:4K
2012	82,944	Fujitsu K	
2013	1,600,000	Sequoia	
2014		Anton 2	
2015	1,500,000	Sequoia	
2016	10,600,000	Sunway TaihuLight	







Hardware awards

1987-	**CM-2**	Thinking Machines SIMD, 1K-4K…65K Processing elements.
1989	**PLA**	Building and using a highly parallel PLA for DNA sequence matching
1993	**SNAP (32 P SIMD)**	Image analysis using the bispectrum analysis algorithm
1996	**Grape 4 P: 1269**	Simulation of the motion of 780,000 stars
1998	**TI Signal processors**	Lattice QCD
1999	**Grape 5, P:32**	Astrophysical n-body simulation
2003	**Grape 6 40 TF**	Performance Evaluation and Tuning of GRAPE-6—Towards 40 ‘Real’ Tflop/s
2006		A 185 Tflops Simulation of Amyloid-forming Peptides from Yeast Prion Sup35 with the Special-Purpose Computer System MD-GRAPE3
2006	**FPGA cards**	$158/Gflops Astrophysical *N*-Body Simulation with Reconfigurable FPGA
2009	**CPU/GPU**	42 TFlops hierarchical N-body simulations on GPUs with applications to astrophysics and turbulence
2009	**Anton**	Millisecond-scale molecular dynamics simulations on Anton
2014	**Anton 2**	performance and programmability in a special-purpose molecular dynamics
2016	*FPGA*	*Evaluating and Optimizing OpenCL Kernels for High Performance Computing with FPGAs*

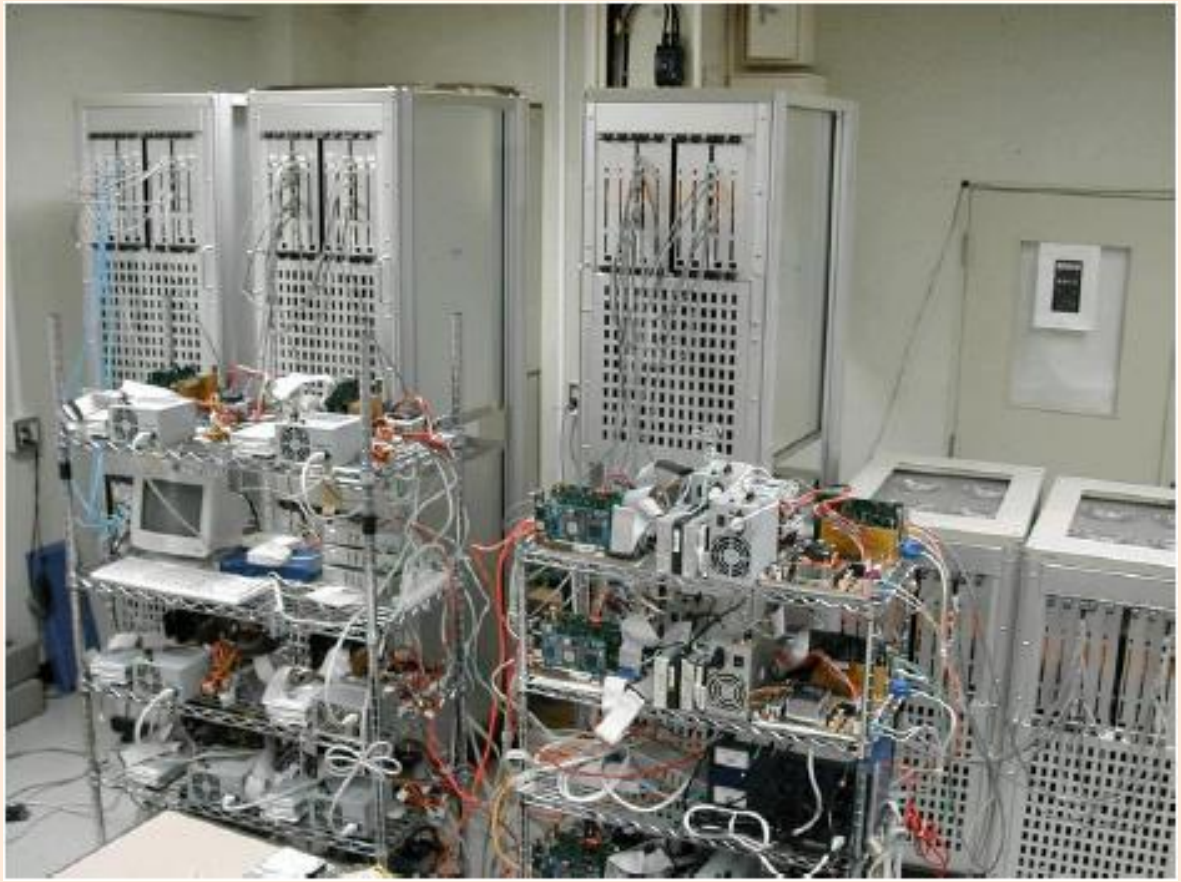
Thinking Machines SIMD CM2 c1988-1990

- QCD and Circuit simulation
- Seismic data processing
- Oil reservoir
- Grid generation for PDE

	CM-2a	CM-2
Clock cycle	142 ns	142 ns
Theor. peak performance:		
Per processor	0.007 Gflop/s	0.007 Gflop/s
Maximal	1.792 Gflop/s	14.34 Gflop/s
Main memory	˙ 8 GB	˙ 8 GB
Memory bandwidth:		
NEWS grid	˙ 310 MB/s	˙ 2.5 GB/s
General routing	˙ 31 MB/s	˙ 0.25 GB/s
No. of processors	4,096–8,192	16384–65,536

Grape 6 for astrophysical N-body calculations (2002)

Junichiro Makino, Eiichiro Kokubo, Toshiyuki Fukushige,and Hiroshi Daisaka



- Peak 63.4 Tflops, 2,048 Pipelined processing chips. 16 computers hosting 4 boards with 32 processors
- 29.5 Tflops simulation of planetesimals in Uranus-Neptune
- 1.8 million planetesimals and two massive protoplanets
- 2002 Earth Simulator, 5K processors, $200 M, 35 GFlops

42 TFlops Hierarchical *N*-body Simulations on cluster of 128 CPU/*256 GPU* for Applications in both Astrophysics and Turbulence (2009)

Tsuyoshi Hamada, Rio Yokota, Keigo Nitadori, Tetsu Narumi, Kenji Yasuoka, and Makoto Taiji

Anton 2 (2014): Raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer

	Anton 1	Anton 2
Process technology	90 nm	40 nm
Clock speed (GC/PPIM)	485/970 MHz	1.65/1.65 GHz
# of general-purpose processor cores	13	66
# of PPIMs	32	76
Total SRAM + data cache	384 KB	5,280 KB
HTIS memory capacity (atoms)	6,144	32,768
Total data bandwidth to torus neighbors	221 Gb/s	1,075 Gb/s

- 512 nodes. 66 cpus feed pipelines
- Anton 2, first to achieve simulation rates of multiple microseconds of physical time per day for systems with millions of atoms.
- simulates standard 23,558-atom benchmark at a rate of 85 µs/day—180 times faster than supers

Hardware…nice surprises, more than I thought

- Long history of special hardware e.g. QCD
- Cost effective … depends on utilization and generality
- SIMD works… highly special. Late 80s before multicomputers
- First use of PLA (1989) at IDA's Supercomputer Res. Center
- Tokyo Inst. Of Tech. introduces GPU as a component
- SC 2016 FPGA study paper. Microsoft introduces FPGAs into datacenters
- Anton 1, 2 are production machines that serve MD community

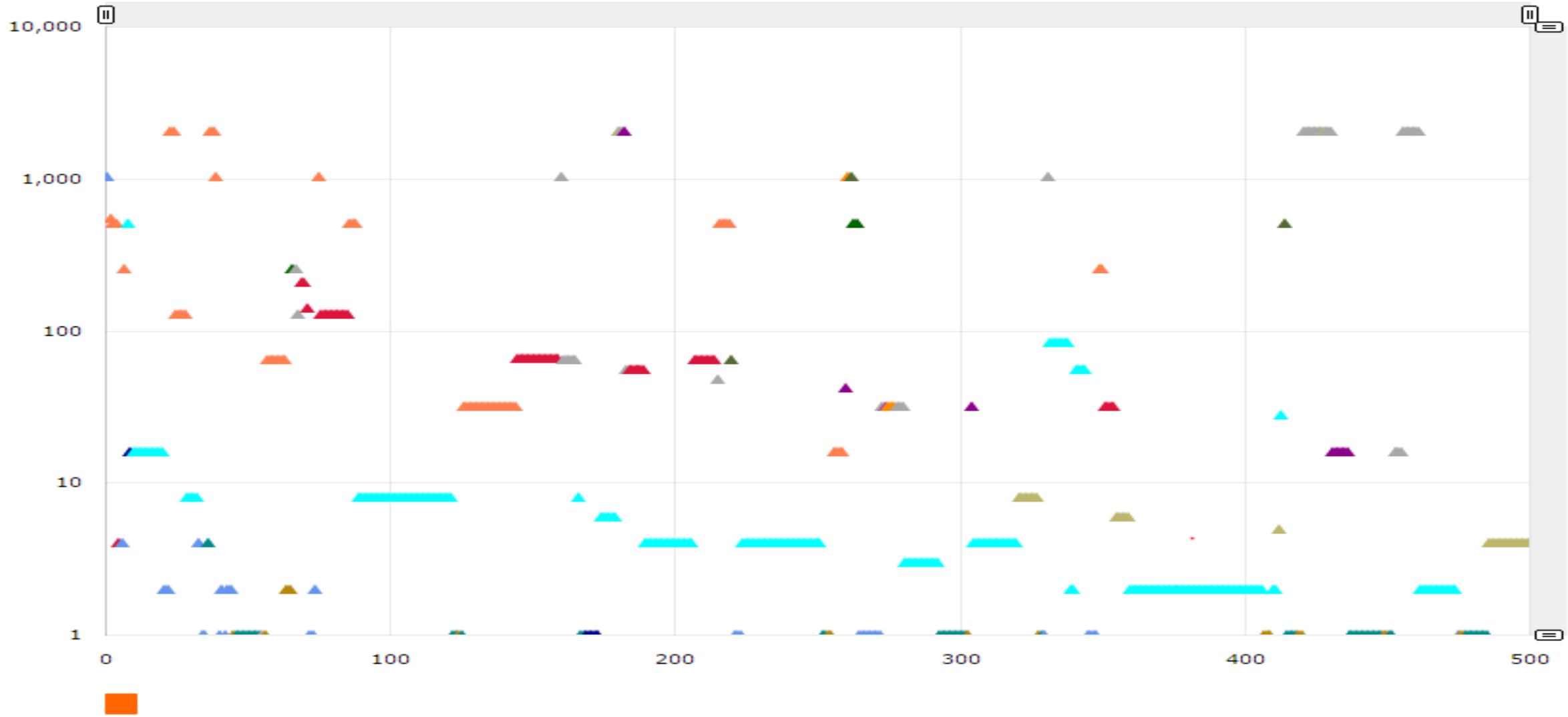
Surprises and confirmation along the way? Why does it take so long? Or so quickly?

1. Multicomputer vs. multiprocessors. MPP, cluster, constellations, cores, confuse Transition from mono memory, aka Fortran computers to clusters Ncube, 87…demo critical to any transition.
    1. 1982,83 Caltech Cosmic Cube demonstrated, stimulates Intel and nCube spinoff
    2. 1987 Bell Prize with Gustafson et al
    3. 1992 community accept this as crossover CM5 1K exceeds a new Cray with 64 processors
    4. 1994 Beowulf kit and multicomputer recipe followed by MPI-1 Standard
    5. 1995 ASCI program at DOE was critical to focus on software
2. VASTNESS: SCALE to fully distributed, multi cabs and vast no of threads and/or cores, after 1994.
3. GPU and many processing elements CUDA has enabled
4. and potentially FPGA
5. Both Hardware and especially Anton as a very special purpose computer was demonstrated
6. NOT
    1. SIMD
    2. Globus and more distributed apps across a very large number,
    3. Large WS clusters NOW ala distributed COW
7. TidalScale surprise “shared memory” on “multicomputer”… due to differences in memory speeds

2015: 100 petaflops; 2018: 1 exaflops

- NUDT (China) Tianhe-2
  - 100 petaflops supercomputer for completion in 2015.
  - 1 exaflop supercomputer online by 2018 using the Intel MIC multi-core processor architecture
- Given the current speed of progress … supercomputers will reach one exaflops (10^{18}) (one quintillion FLOPS) by 2018-20.
- SGI plans to achieve a 500 fold increase in performance by 2018, and achieve one exaflops.

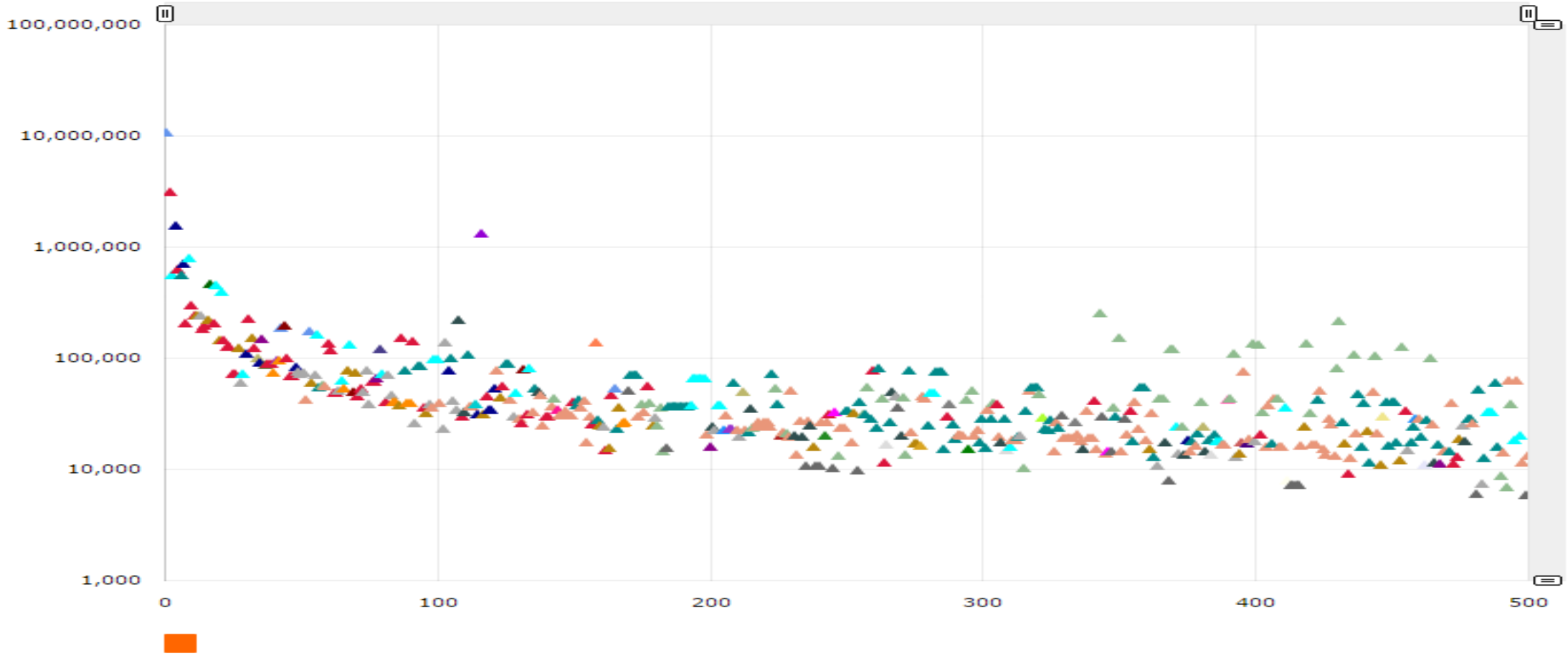
Top500 cores per system June 1993



Legend:

Intel, Meiko, KSR, NEC, KSR/SNI, Parsytec, Hitachi, Fujitsu/SNI, TMC, Cray Inc., Fujitsu, nCube, MasPar, HPE,

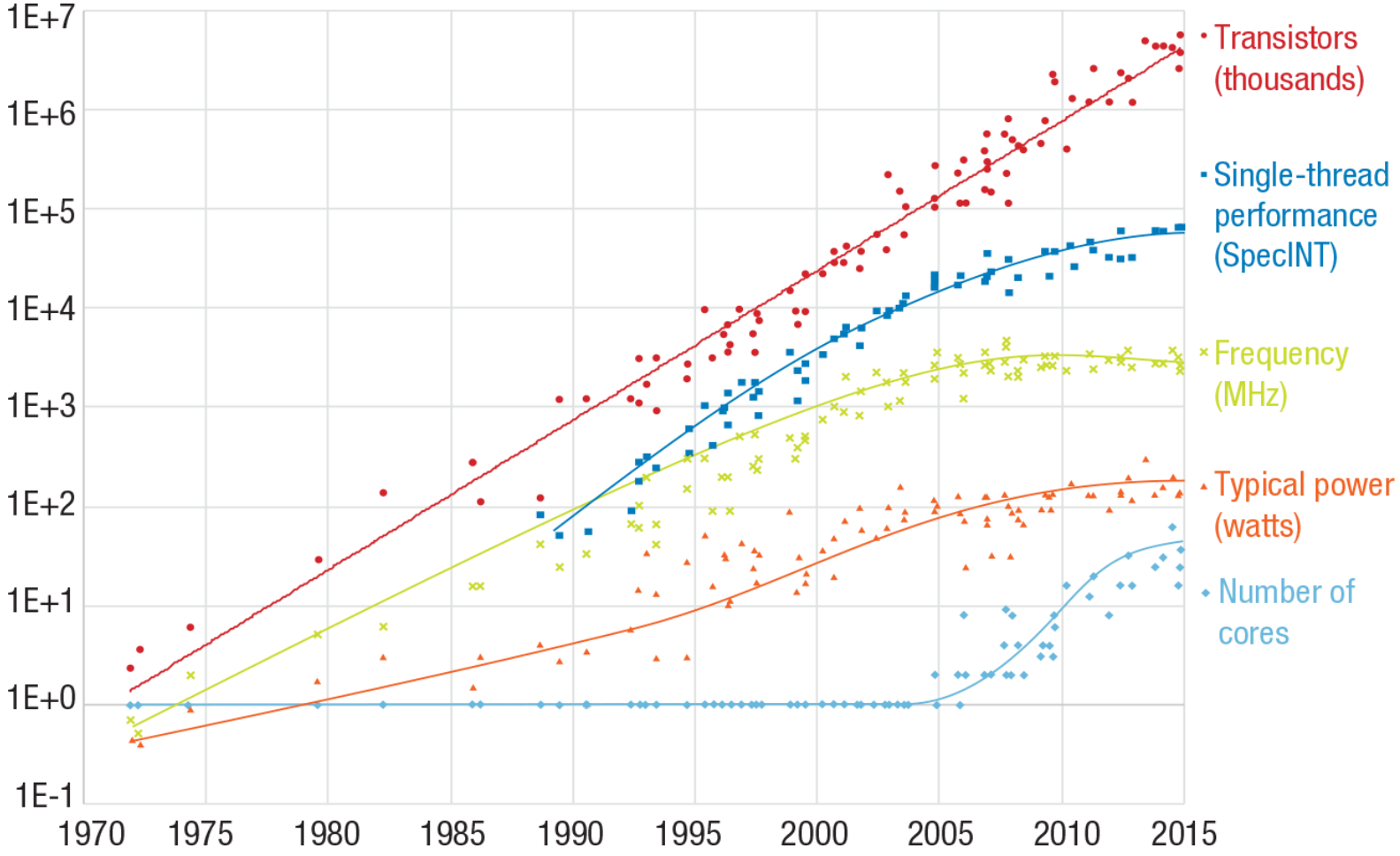
Top500 cores per system, Nov. 2016



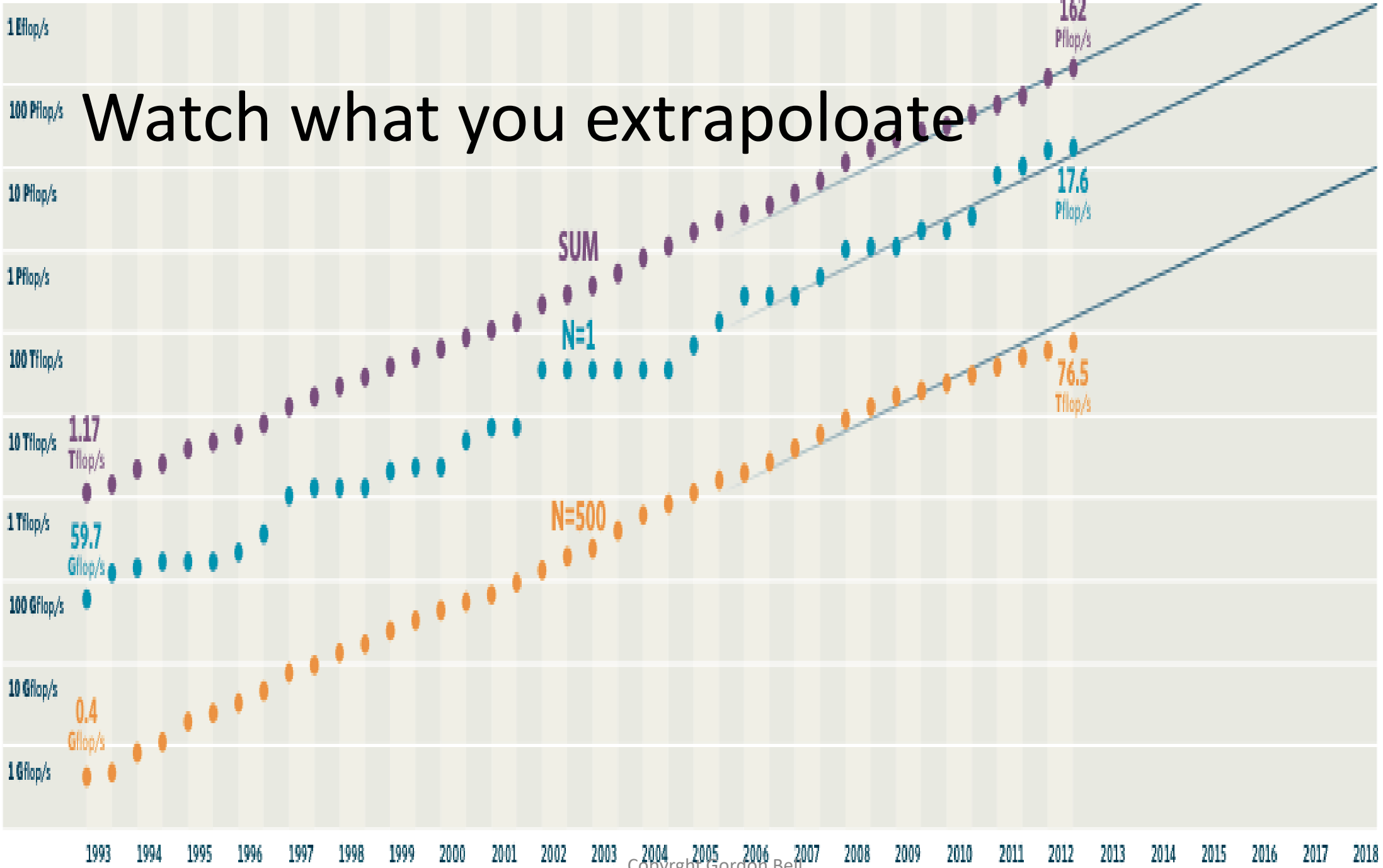
Legend:

IBM, Hitachi/Fujitsu, HPE/SGI, Nvidia, NEC/HPE, Inspur, Huawei, T-Platforms, Sugon, ClusterVision, Koi Computers, Adtech, Dell / Intel, IPE, Nvidia, Tyan, Dell, SuperMicro/Mellanox, Cray Inc., Bull, ACTION, NUDT, Lenovo, NEC, Intel, DALCO, MEGWARE, Atipa, Self-made, IBM/Lenovo, Fujitsu, Penguin Computing, PEZY Computing / Exascaler Inc., Network Technologies, HPE, Megatel/Action, Dell/Sugon, EXXACT, Supermicro, Lenovo/IBM, Oracle, NRCPC, RSC Group,

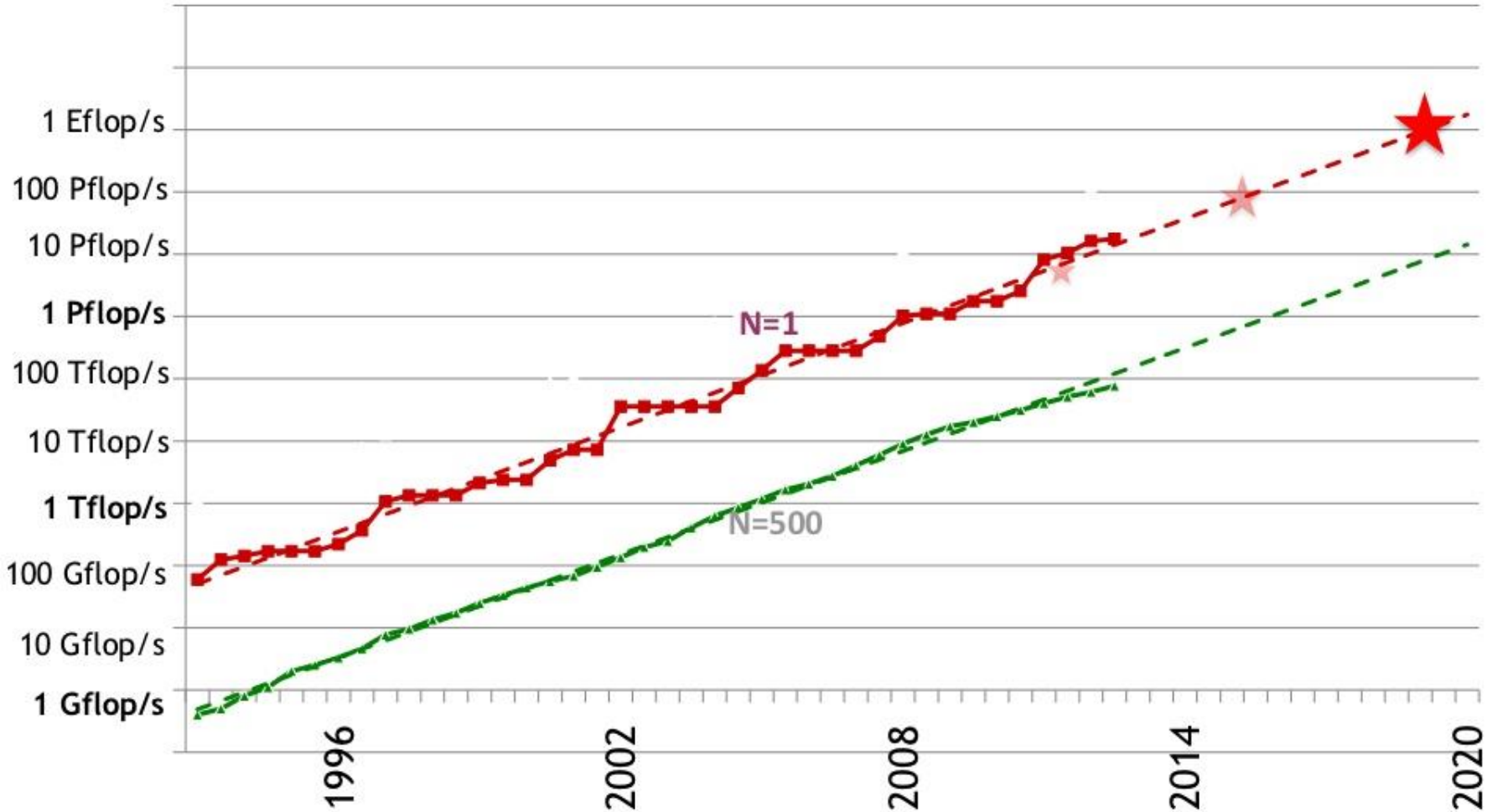
End of historical scaling of clock, cores, SpecNT required new architectures =>> new programming

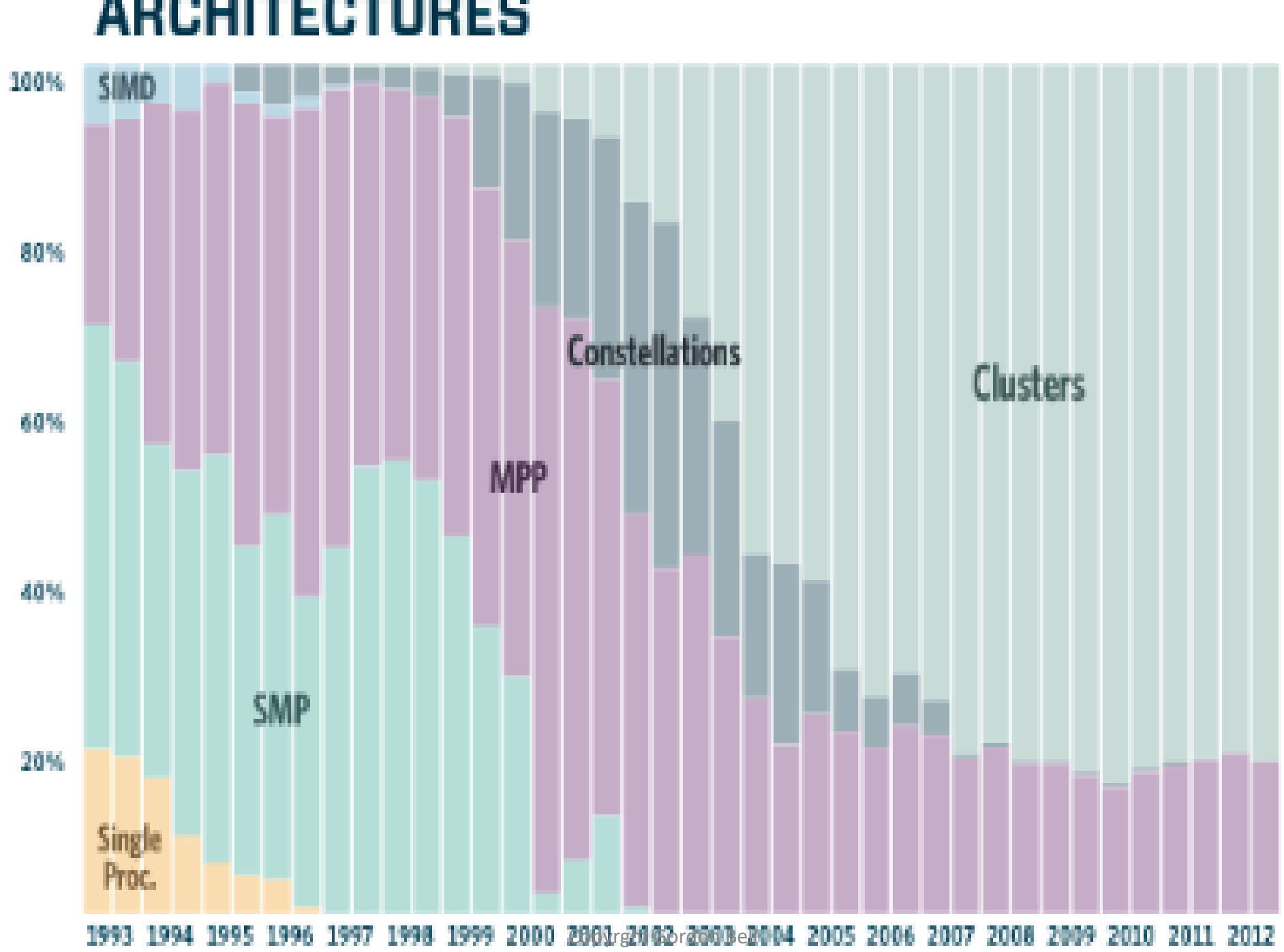


Watch what you extrapoloate



Performance Development in Top500




ARCHITECTURES
100%
80%
60%
40%
20%
SIMD
Constellations
Clusters
MPP
SMP
Single Proc.
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012

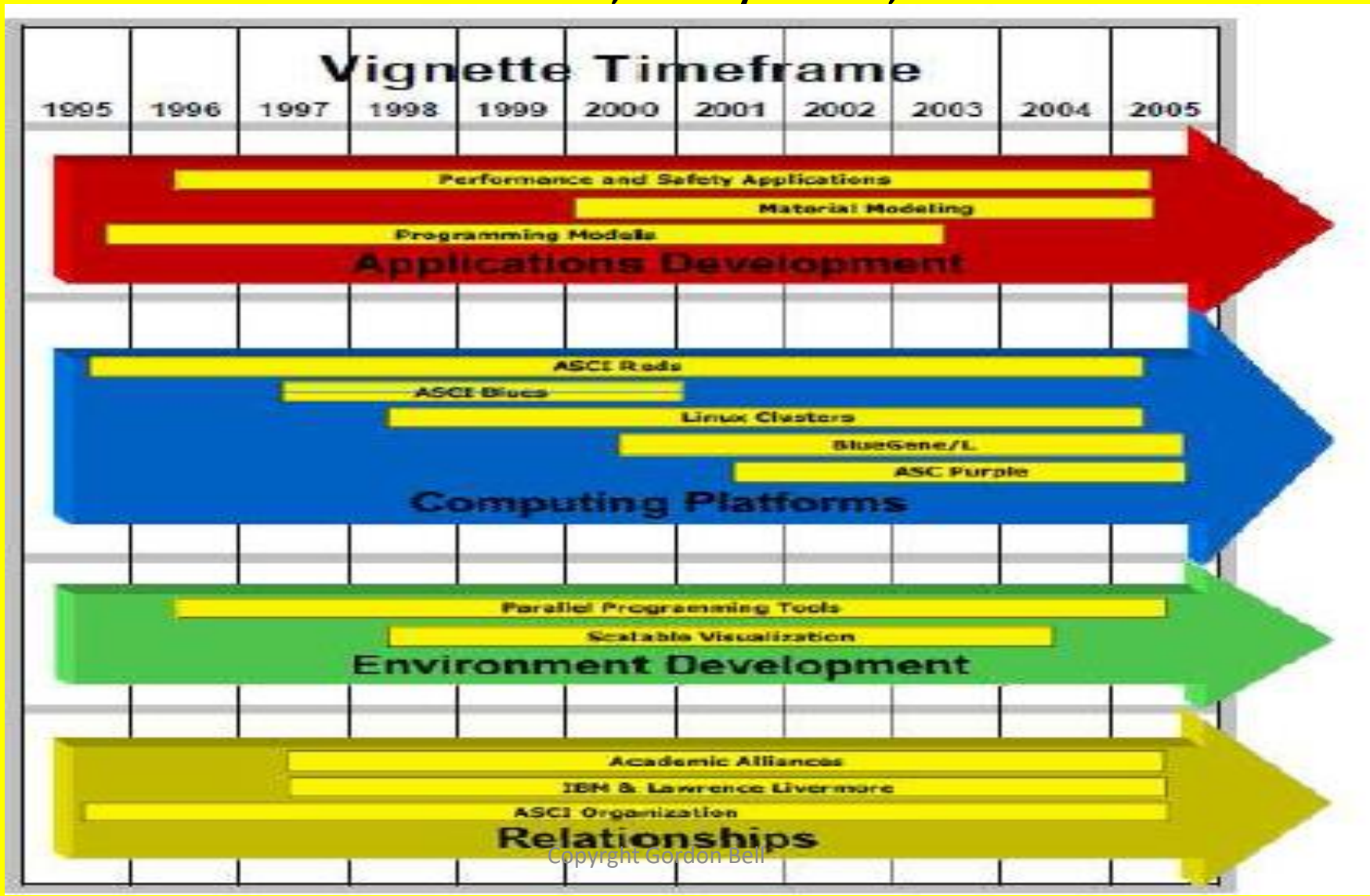
Challenge i.e. change in HPC to come

- 20 years of stability since 1994…10^p p= 3,…6, … 9?
- Jan. 2007 Jim Gray posits Fourth Paradigm of Science
  - Analysis of experimental data by computer to "do science"
  - Analysis of data coming from models (3rd Paradigm data)
- Merger of data science and computational science architectures at machine and application program
- Cloud service evolves to support HPC…
  - Cloud provides HPC and database services… architecture and application

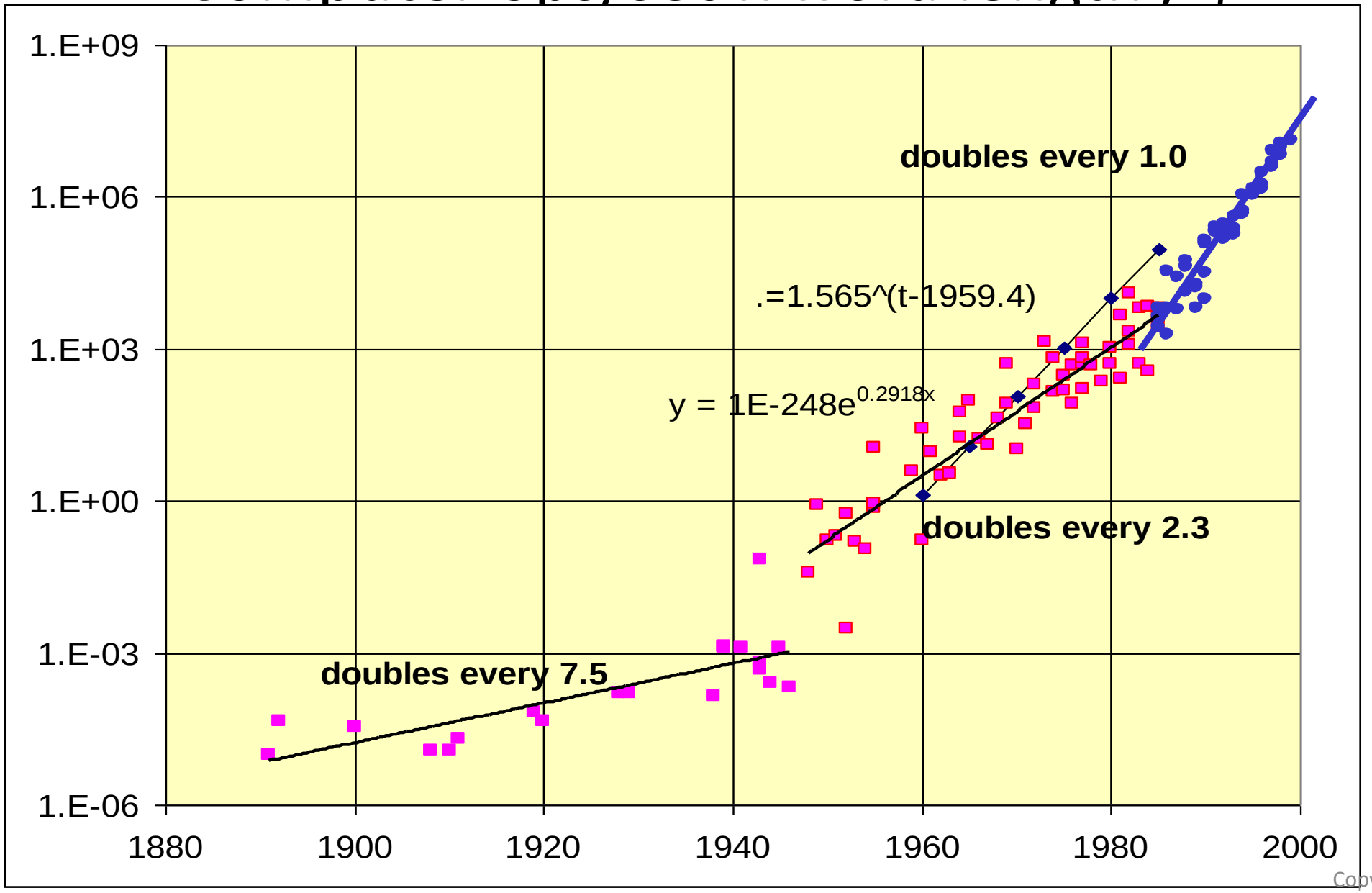
The End

ASCI: Accelerated Strategic Computing Initiative =>ASC: Advanced Simulation and Computing

from Alex R. Larzelere II, History of ASCI, 1995-2005



Computer ops/sec x word length / $



What's in a Data Center?

The Virtuous Economic Cycle drives the PC industry… & Beowulf