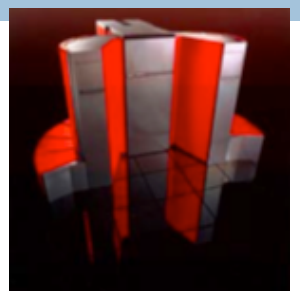
Improving HPC Software

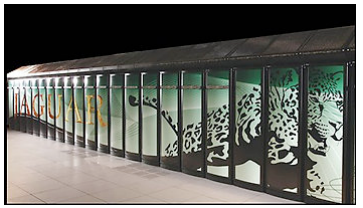
Pete Beckman & Jack Dongarra

Looking at the Gordon Bell Prize



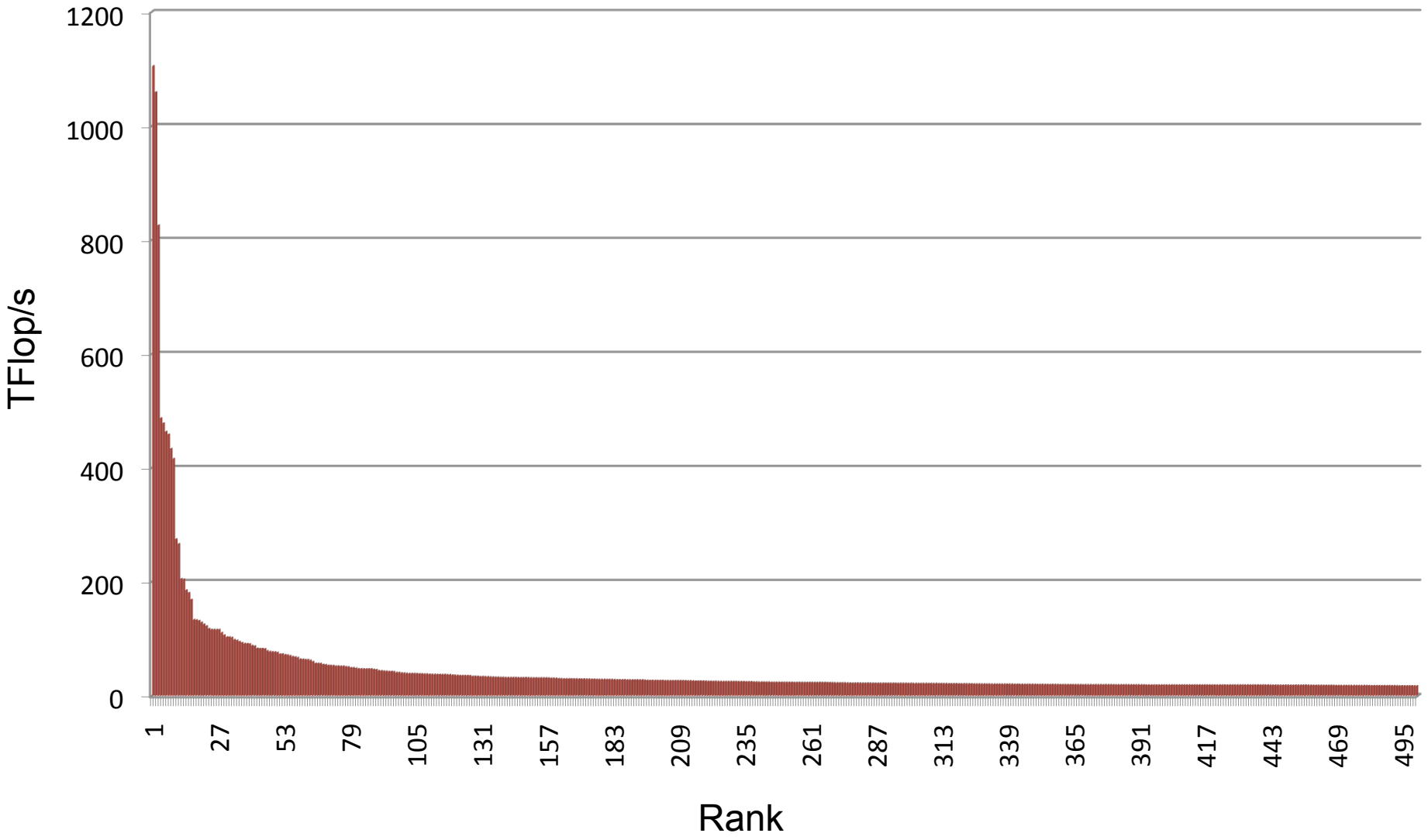
- 1 GFlop/s; 1988; Cray Y-MP; 8 Processors
  - Static finite element analysis

- 1 TFlop/s; 1998; Cray T3E; 1024 Processors
  - Modeling of metallic magnet atoms, using a variation of the locally self-consistent multiple scattering method.
- 1 PFlop/s; 2008; Cray XT5; 1.5×10^5 Processors
  - Superconductive materials

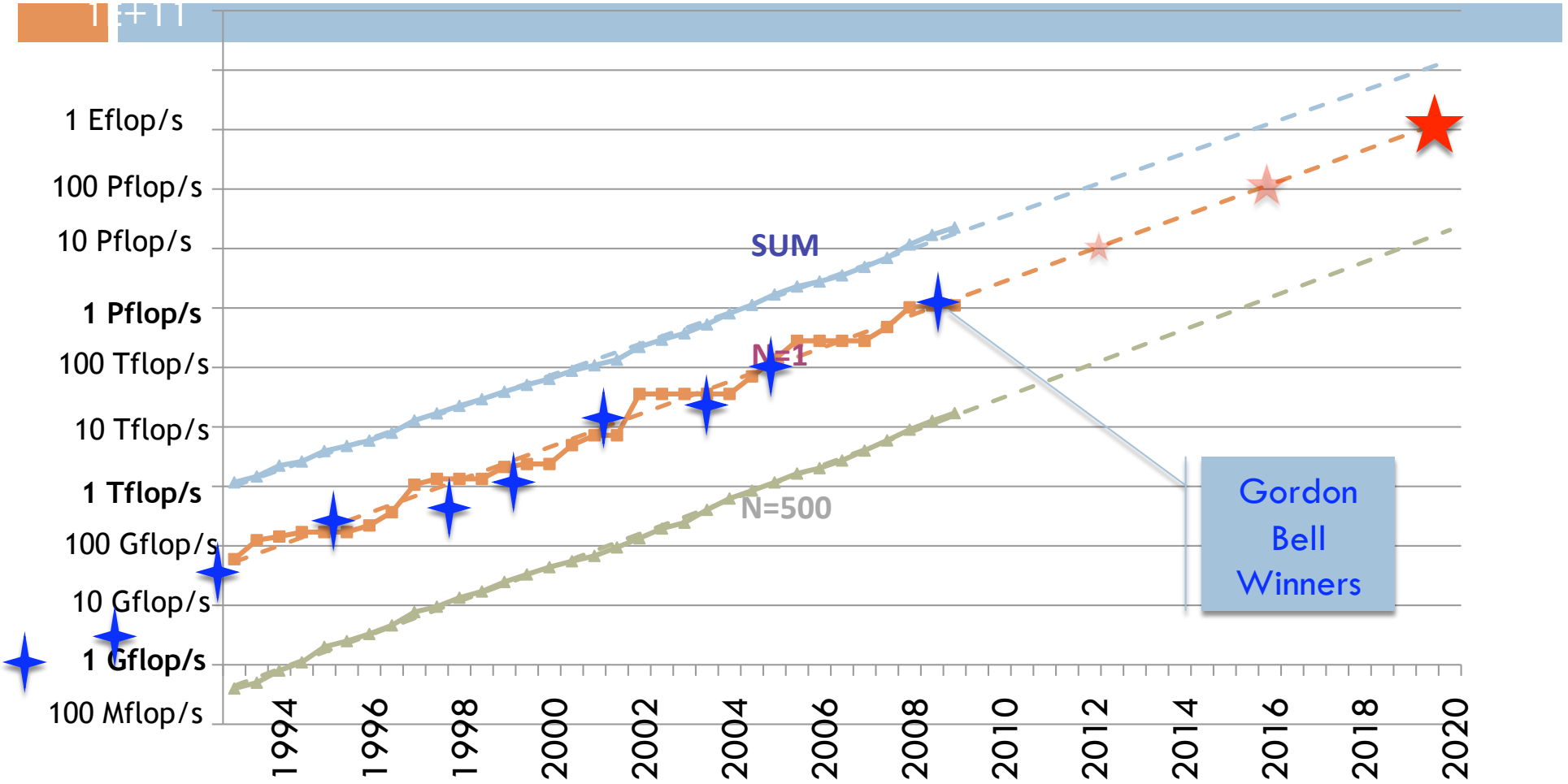


- 1 EFlop/s; ~2018; ?; 1×10^7 Processors (10^9 threads)

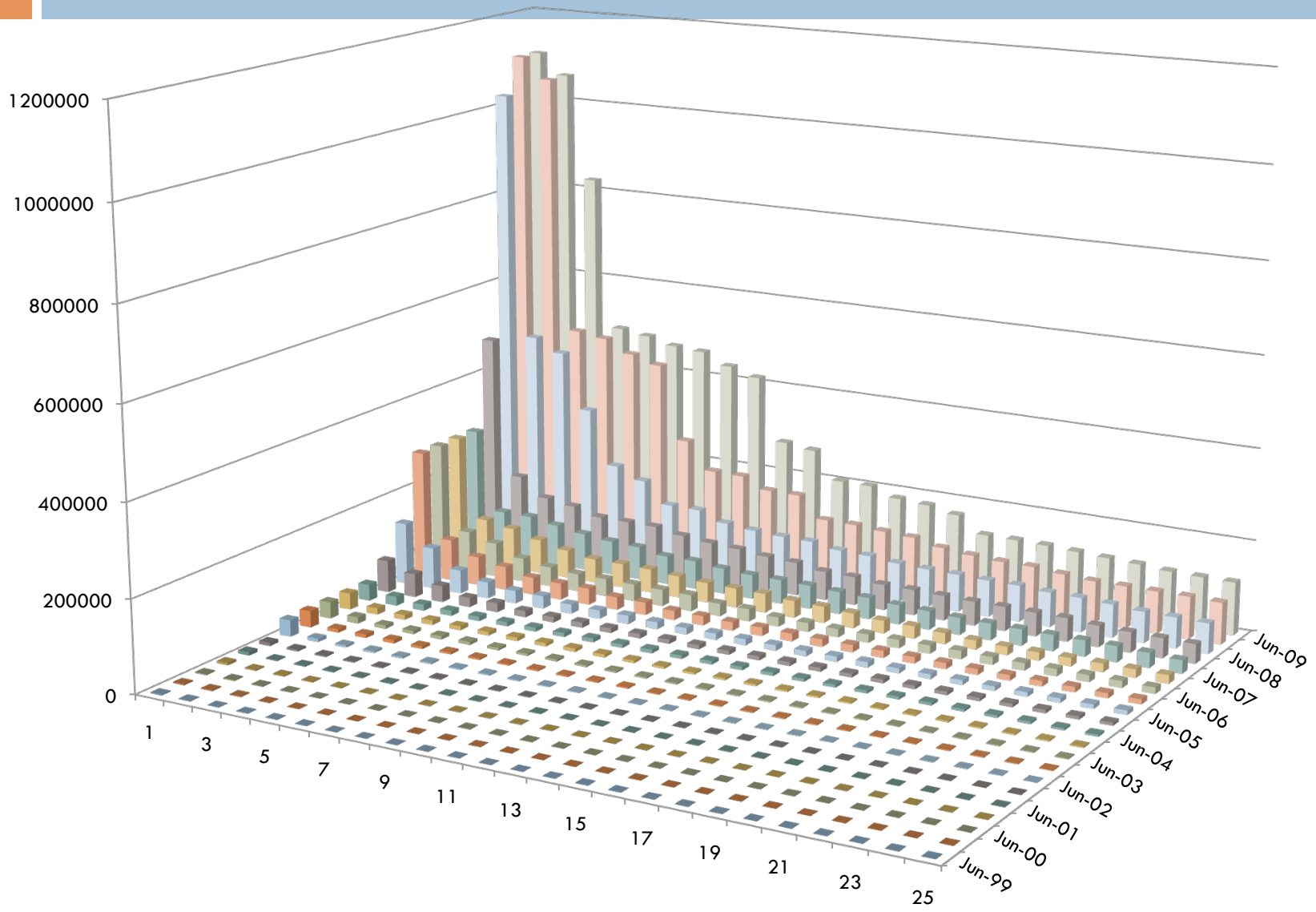
Distribution of the TOP 500



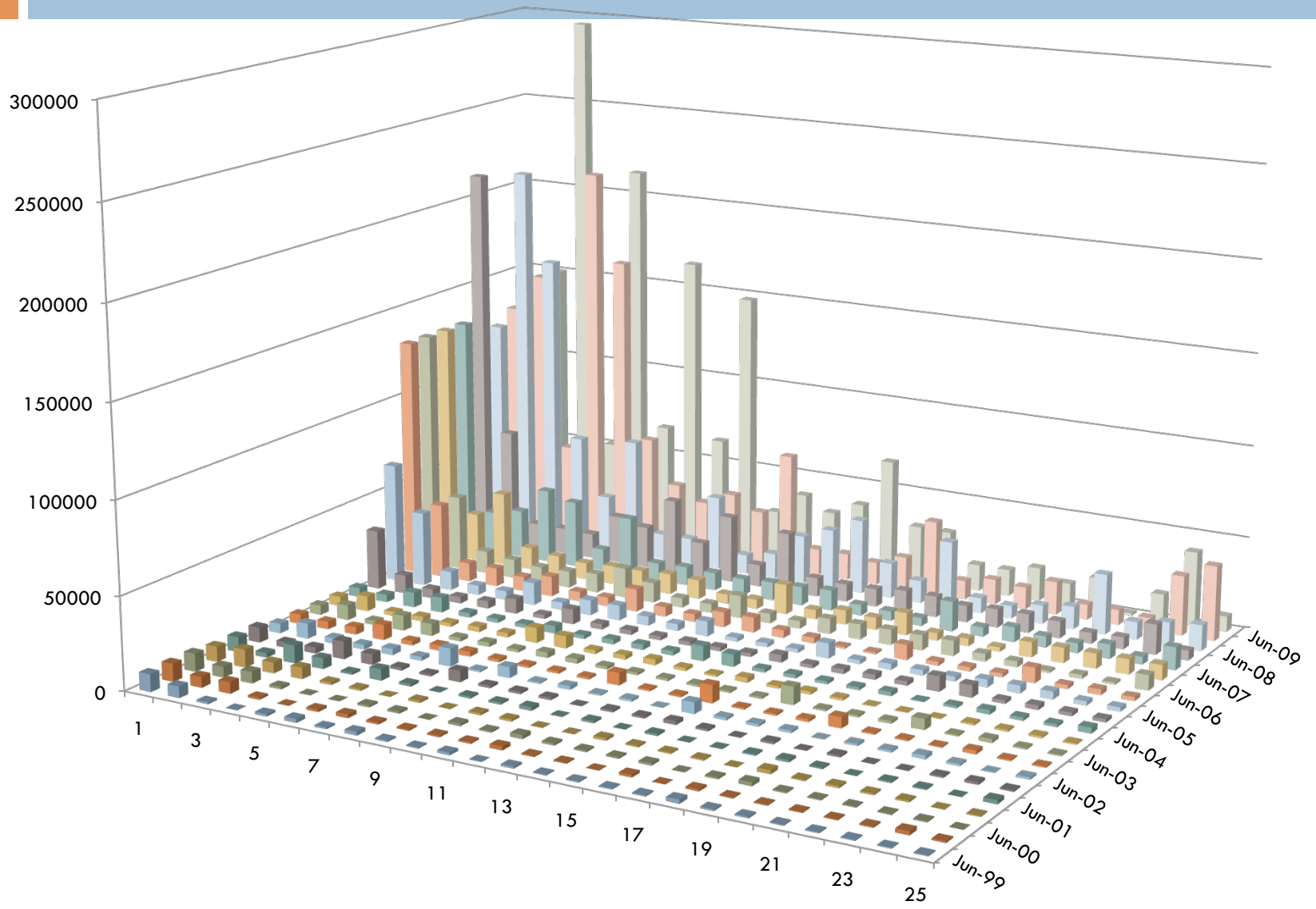
Performance Development in Top500



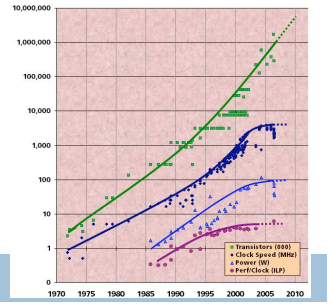
Performance of Top25 Over 10 Years



Cores in the Top25 Over Last 10 Years



A Call to Action



- Hardware has changed dramatically while software ecosystem has remained stagnant
- Previous approaches have not looked at co-design of multiple levels in the system software stack (OS, runtime, compiler, libraries, application frameworks)
- Need to exploit new hardware trends (e.g., manycore, heterogeneity) that cannot be handled by existing software stack, memory per socket trends
- Emerging software technologies exist, but have not been fully integrated with system software, e.g., UPC, Cilk, CUDA, HPCS
- Community codes unprepared for sea change in architectures
- No global evaluation of key missing components

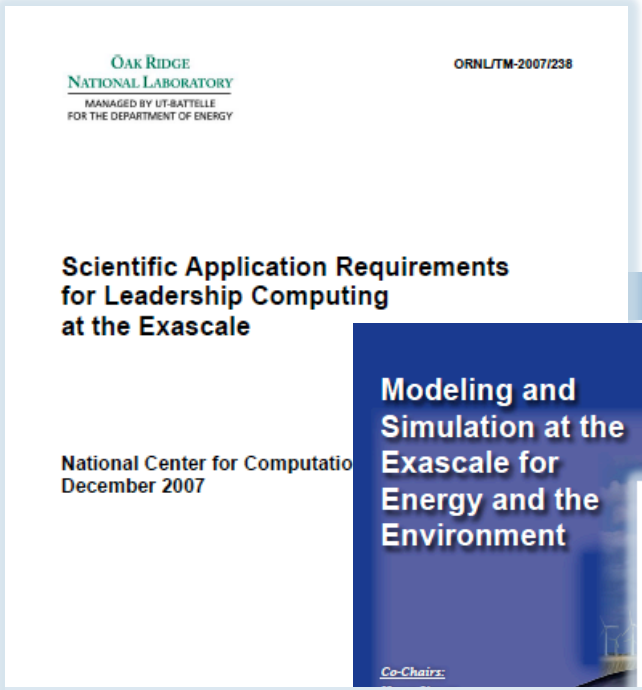
IESP Goal

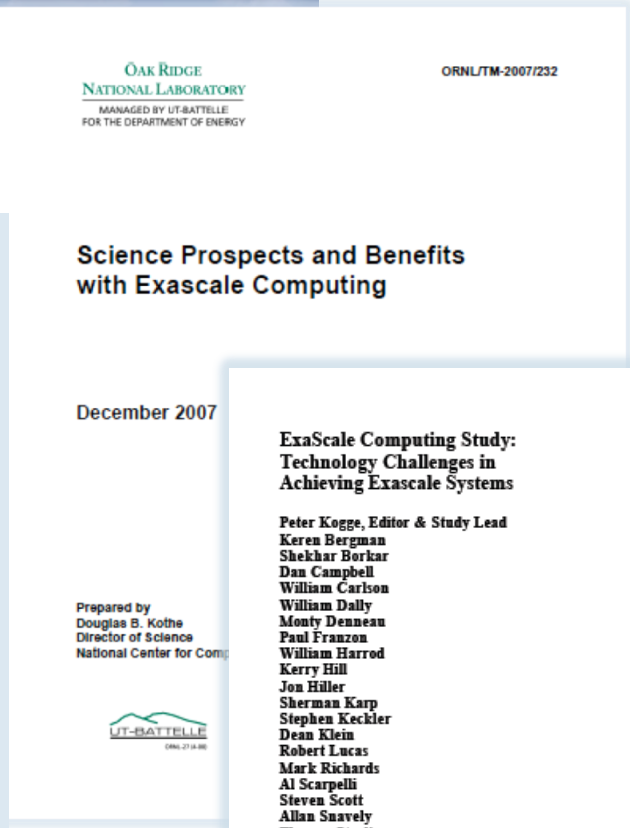
Improve the world's simulation and modeling capability by improving the coordination and development of the HPC software environment

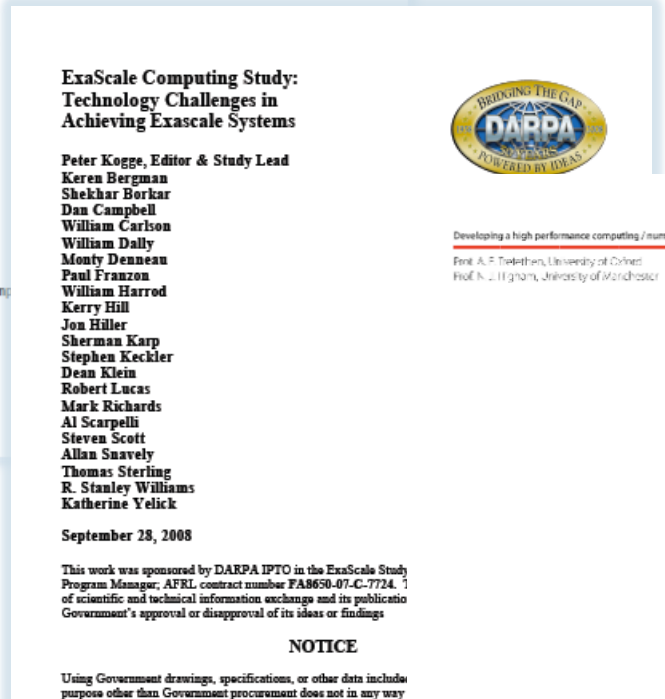
Workshops:

Build an international plan for developing the next generation open source software for scientific high-performance computing

Impact of Exascale







Strong science case for the continued escalation of high-end computing.

Broad consensus necessitate the redesign and replacement of many of the algorithms and software infrastructure that HPC has built on for more than a decade.

Factors that Necessitate Redesign

- **Steepness of the ascent from terascale to petascale to exascale**
- Extreme parallelism and hybrid design
  - Preparing for million/billion way parallelism
- Tightening memory/bandwidth bottleneck
  - Limits on power/clock speed implication on multicore
  - Reducing communication will become much more intense
  - Memory per core changes, byte-to-flop ratio will change
- Necessary Fault Tolerance
  - MTTF will drop
  - Checkpoint/restart has limitations
- **Software infrastructure does not exist today**

International Community Effort

- We believe this needs to be an international collaboration for various reasons including:
  - The scale of investment
  - The need for international input on requirements
  - US, Europeans, Asians, and others are working on their own software that should be part of a larger vision for HPC.
  - No global evaluation of key missing components
  - Hardware features are uncoordinated with software development

Where We Are Today:

- SC08 (Austin TX) meeting to generate interest — **Nov 2008**
- DOE's Office of Science funding
- US meeting April 6-8, 2009 — **Apr 2009**
  - 65 people
- NSF's Office of Cyberinfrastructure funding
- European meeting June 28-29, 2009 — **Jun 2009**
  - 70 people
  - Draft Roadmap
  - Outline Report
- Asian meeting (Tsukuba Japan) October 18-20, 2009 — **Oct 2009**
  - Refine roadmap
  - Refine Report
- SC09 (Portland OR) BOF to inform others — **Nov 2009**
  - Public Comment
  - Draft Report presented

Follow on Workshops

- Spring of 2010
  - Continued coordination
  - Refine Roadmap
  - Implementing the plan

A Running Start - White Papers: www.exascale.org

- *Musings on the Path Toward Exascale*, Robert Lucas – ISI/USC
- *BSC Vision Towards Exascale*, Mateo Valero, BSC
- *Software Challenges of Extreme Scale Computing*, Michael Heroux – Sandia National Laboratory
- *Software and Exascale Computing*, Bill Camp – Intel Corporation
- *Application Analysis and Porting in the PRACE Proj* Michielse – Netherlands National Computing Facili (NCF)
- *The Application Perspective – Seeking Productivity Performance*, David Barkai – Intel Corporation
- *EDF white paper*, J.Y. Berthou and J.F. Hamelin – ED
- *The Biggest Need: A New Model of Computation*, T – Louisiana State University
- *NSF IESP Whitepaper*, Abani Patra, Rob Pennington, Office of Cyberinfrastructure, National Science Fou
- *A Proposal for a Capability Centers Consortium*, Bi Snir – NCSA and the University of Illinois at Urbana
- *Slouching Towards Exascale*, Rusty Lusk, Argonne Laboratory
- *A Collaboration and Commercialization Model for Software Research*, Mark Seager and Brent Gorda, Livermore National Laboratory
- *The Case for A Hierarchal System Model for Linux* Seager and Brent Gorda, Lawrence Livermore Natio
- *IESP Whitepaper: PDE-based applications and solv scale*, DavidKeyes, Columbia University & SciDAC T
- *Developing a high performance computing/numer roadmap*, Ann Trefethen, Nick Higham, Ian Duff, a Coveney

Whitpapers and notes on crosscutting issues for Paris Meeting

- *Performance at Exascale*, Bernd Mohr (Jülich Supercomputing Centre) and Matthias S. Mueller (Wolfgang E. Nagel Center for Information Services and HPC)
- *Resource Management*, Barney McCabe (ORNL) and Hugo Falter (ParTec)
- *Programmability Issues*, Vivek Sarkar (Rice U.), Jesus Labarta (UPC), Mitsuhisa Sato (U. of Tsukuba), Barbara Chapman (U. of Houston)
- *Models of Computation – Enabling Exascale*, Thomas Sterling, Louisiana State University.
- *Major Computer Science Challenges at Exascale*, Al Geist (ORNL) and Robert Lucas (ISI)
- *Towards Exascale File I/0*, Yutaka Ishikawa, University of Tokyo
- *Co-design of Architectures and Algorithms*, Al Geist (ORNL) and Sudip Dosanjh (SNL)
- *IESP Exascale Challenge: Resilience and Fault Tolerance*, Al Geist (ORNL) and Franck Cappello (INRIA)

Workshops and Report

- 3 workshops
  1. Santa Fe, April 7-8
  2. Paris, June 28-29
  3. Tsukuba, October 18-20
- Broad engagement by the community
- Initial reports in fall 2009
- Draft report presented at SC09
- Planning for *IMMEDIATE* payoff
  - Could begin initial components of plan in FY10

June 28, Sun

8:00 – 9:00 am	Bus from Novotel to Teratec
9:00 am	Welcome - **Catherine Riviere & Christian Saguez**
9:15 am	General objectives of IESP - **Jack Dongarra**
9:45 am	Objectives and organization of the meeting – **Franck Cappello & Christian Saguez**
10:00 am	Main results of the Santa Fe Meeting - **Pete Beckman**
10:30 am	Break
11:00 am	Lessons and feedback of the Santa Fe Meeting – **Jean-Yves Berthou**
11:30 am	Synthesis of the white papers - **Thomas Lippert**
12:00	Report on DOE Exascale studies - **Paul Messina**
12:30 – 2:00 pm	Lunch
2:00 pm	Organize to split into subgroups
2:15 – 4:00 pm	Breakout groups
4:00 pm	Break
4:30 – 6:00 pm	Reports from subgroups
Dinner	In Paris on the Seine