Snowmass Computing Frontier I2: Distributed Computing and Facility Infrastructures

Ken Bloom
Richard Gerber
July 31, 2013

Who we are

- Ken Bloom, Associate Professor, Department of Physics and Astronomy, University of Nebraska-Lincoln
  - Co-PI for the Nebraska CMS Tier-2 computing facility
  - Tier-2 program manager and Deputy Manager of Software and Computing for US CMS
  - Tier-2 co-coordinator for CMS
  - Leader of effort to develop and deploy data federations
- Richard Gerber
  - User Services Deputy Group Lead, NERSC (National Energy Research Scientific Computing Center), Berkley Lab
  - NERSC Senior Science Advisor
  - Co-Convener "Large Scale Production Computing and Storage Requirements for Science" series of requirements reviews for DOE.
  - NERSC-7 (Edison) Deputy Project Manager

The Big Picture

- Will adequate computing facilities be available to support HEP science for the next:
  - 5 Years?
  - 10 Years?
  - Beyond?

Our charge

- How do the various computational problems that are posed by HEP science map onto various types of computing facilities?
- Given the computational needs of various HEP efforts in both experimental and theoretical work, will computing facilities of the required size be available over the appropriate timescales without any new targeted efforts?
- Will the existing distributed computing models for particle physics, largely based around grid infrastructures and access to distributed data, scale up to meet future needs without any new targeted efforts?
- Will national computing centers play a larger role in computations for particle physics? Will there be a role for computing on demand, i.e. cloud, facilities?
- What sort of coordination will be required across distributed facilities, and are new models of computing required for it?

Our approach

- Leveraged our work off three forays into the community:
  - NERSC was already scheduled to do an assessment of program needs for high-energy physics [add link], which gave us information about the "facility infrastructures"
    - Requested case studies from experimentalists and theorists on all three "frontiers" that described computing needs of upcoming efforts, both in terms of type and scale
    - Report currently in draft form, expected to be released XXXXX
  - Took advantage of the Open Science Grid All-Hands Meeting in March [add link] to convene a discussion panel on the future of the grid, which gave us information about "distributed computing"
    - Recruited panelists from different parts of the grid world: operations, technology, security, big thinking
    - Snowmass report will summarize the discussion
  - Listened carefully to Tuesday presentations from CpF E,T groups

I: Computing problems and facilities

- The Worldwide LHC Computing Grid (WLCG) efficiently handles data analysis and detector simulation for CMS and ATLAS, the dominant energy-frontier experiments
  - LHC computational problems are well suited to high-throughput computing paradigm
  - WLCG workflow model is working well for experiments, expected to continue to do so in the future
  - The issue is whether there will be adequate resources (compute, storage, network) to support increased data from LHC experiments, both for 2015 run and beyond
    - This could use more quantitative input from Energy Frontier group

I: Computing problems and facilities

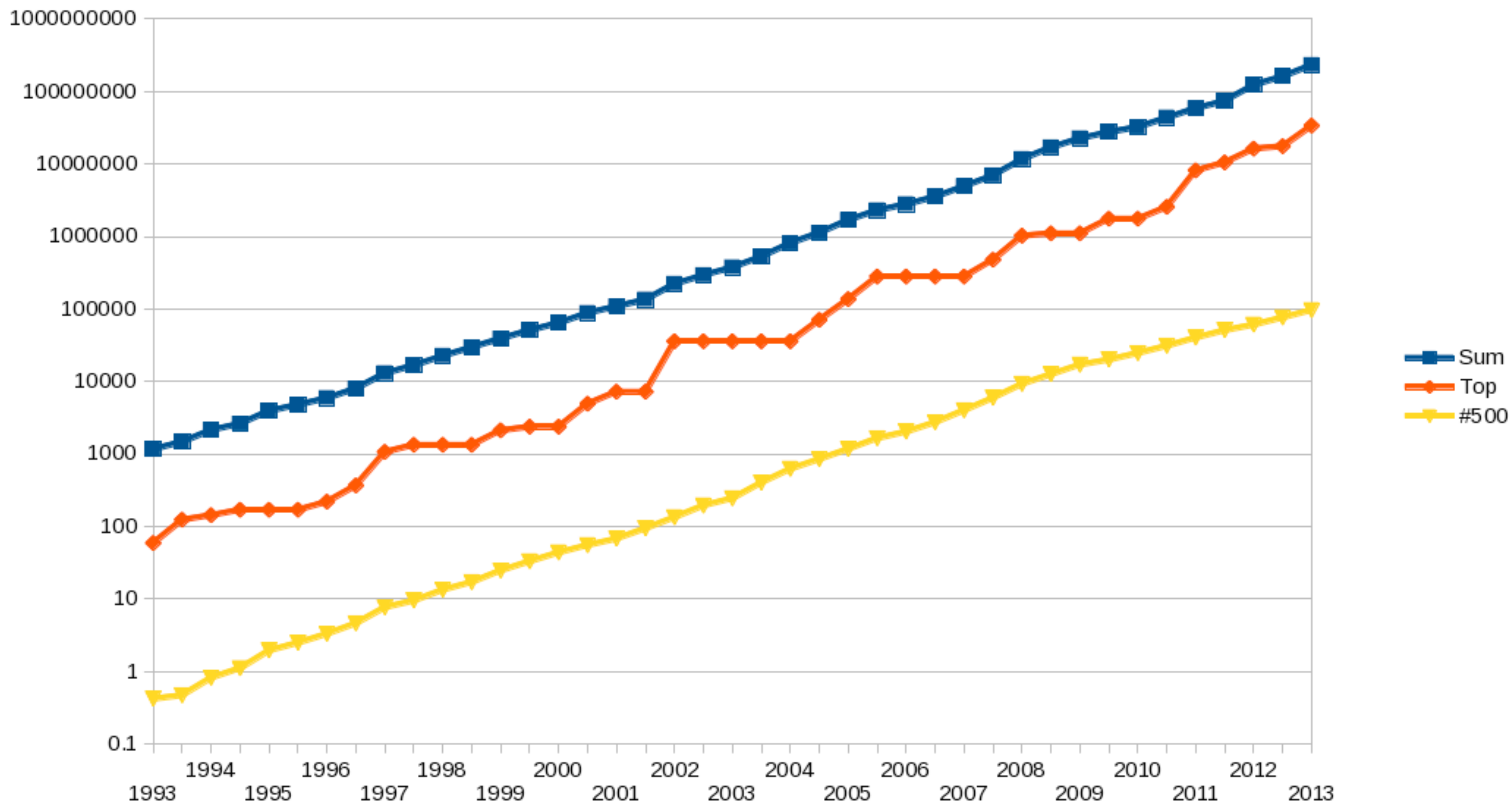
- National High Performance Computing (HPC) centers are used and required by a number of projects
  - Lattice QCD (EF)
  - Accelerator design and R&D (EF and IF)
  - Data analysis and synthetic maps (CF)
  - N-body and hydro-cosmology simulation (CF)
  - Supernova modeling (CF)
  - Efforts underway to perform theory computations (e.g. perturbative QCD) directly related to experiment
- Currently looking for more information on how IF experiments might need and use HPC resources
  - NERSC already hosting (and did host) efforts from Daya Bay (Tier 1), KAMLAND, Ice Cube, BaBar, SNO – experience with data
  - [Might get more of this from Tuesday sessions? Otherwise beg?]

2: Can we meet expected needs?

- [This we are fairly in the dark on -- need to see what the other frontiers say. What can we read up on even in advance of Tuesday?]
- [Do know that EF report starts with the assumption that we'll have to look at some very different paradigms]
- [Also have the rather interesting point that WLCG = Hopper, latter thus can't substitute for former]
- [Can safely assert that shared facilities alone will not satisfy HEP's needs, must be prepared to continue to invest in our own infrastructure?]
- Simulations for cosmology, LQCD and accelerator design need an increase of an order of magnitude over five years, two orders over ten years to keep pace with needs. This exceeds the historical Moore's-Law rate of increase.

Traditional HPC

- The HPC facilities in DOE (NERSC & LCFs) and NSF hope to stay on the Top 500 Moore's Law slope, but it depends on
  - Funding
  - Technology improvements in processors and systems: DOE Fast Forward and Design Forward efforts with vendors

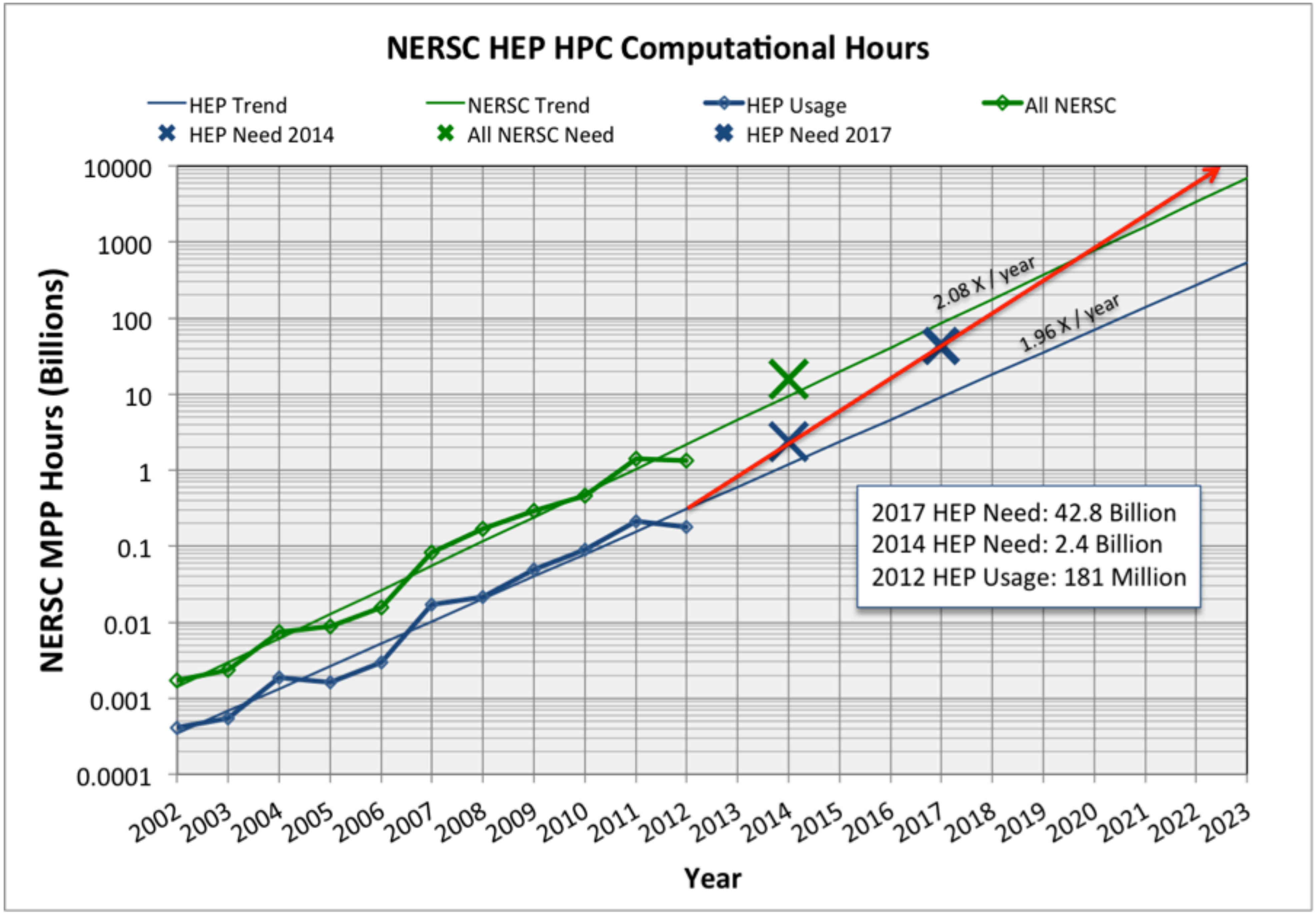


Traditional HPC

- Demand by traditional HEP HPC community will outstrip expected availability by 2017 at NERSC by a factor of 4

Even this is optimistic wrt funding

Driven by LQCD, accelerator, astrophysics



HEP Computing

- 2013 DOE & NSF Allocations for HEP
  - DOE Production (NERSC): 168 M Hours
    - LQCD 50 M (113 M included NP allocation)
    - Cosmology 53 M
    - Accel 23 M (32 M including BES & NP)
  - DOE INCITE (ALCF, OLCF): 820 M Hours
    - LQCD 400 M
    - Supernova 230 M
    - Cosmology 80 M
  - NSF XSEDE: 120 M
    - LQCD: 90 M
    - HEP Theory: 12 M

Distributed and HPC Computing in HEP:

CMS + ATLAS in 2012:
1.4 Billion Hours

National HPC Centers 2013:
1.4 Billion Hours

3: Distributed computing models

- Grid infrastructures well-suited to work done by large collaborations of experimental particle physicists
- Experiments, especially EF, are making good use of the grid, which has been a key technology for physics discovery
- No show-stoppers seen for long-term scaling of high-throughput grid computing, but various developments should be pursued to improve efficiency and ease of use
  - Simplification/scaling of job submission, identity management, streamlined operations, storage management and federated data access, dynamic scheduling, readiness for cloud infrastructures
- HEP is the largest user of the grid and must take a leadership role in its continuing development

4: Role of national facilities

- National HPC centers already play a significant role in some areas.
  - LQCD
  - Cosmology & large scale structure
  - Accelerator research and design
  - Supernova physics
  - Data-driven science in the Intensity and Cosmic Frontiers
- WLCG-based tasks at HPC centers?
  - Opportunities and advantages to take advantage of an additional resource, especially since, HPC resources tend to grow faster than WLCG capability and needs, so worth it to explore more EF-experiment use of HPC

4: Role of national facilities

- Advantages of HPC centers:
  - Facilitate transition to new architectures, strong operations and support, consulting and training, centralized software/data repositories, good growth rate, good for data-intensive projects needing large storage, I/O, world-best networking
- HPC centers have challenges too:
  - Integration with WLCG workflows, job scheduling, designed for parallel rather than serial, virtualization for validated environments, formalities for allocation fo resources, transition to multicore
  - Funding needed to support additional computing

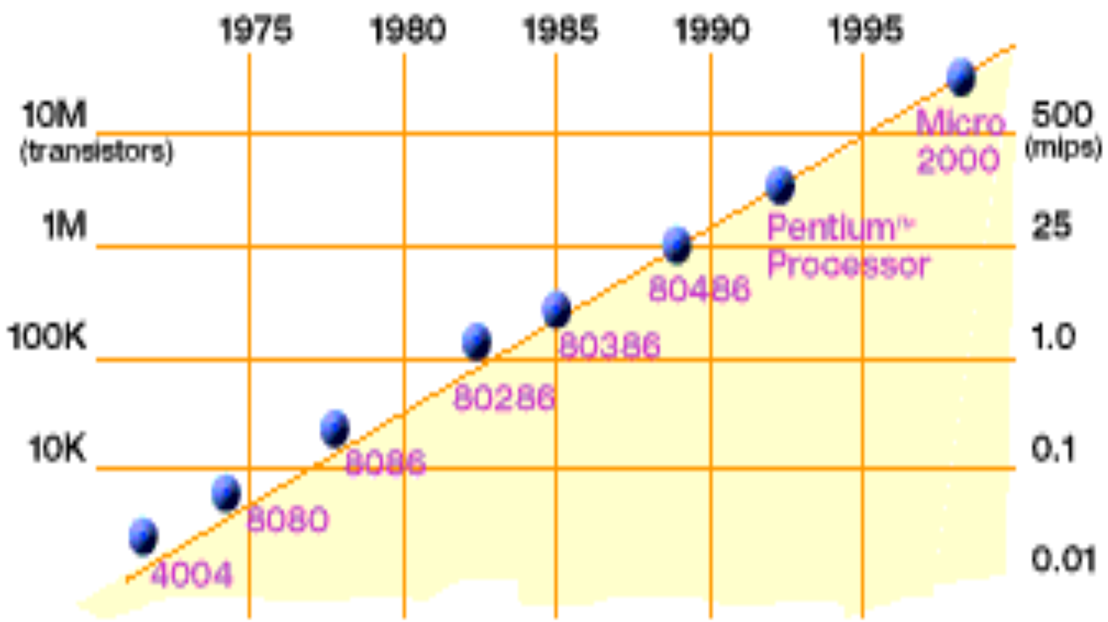
4: Role of cloud computing

- [Commercial] Cloud facilities not currently suited for HEP, mostly due to cost issues at the moment.
- Will access to data limit the usability of clouds?
- Cloud computing provides many advantages, including customized environments that enable users to bring their own software stack.
- Clouds have the ability to quickly surge resources to address larger problems.
- Significant gaps and challenges exist in managing virtual environments, workflows, data, cyber-security, and other areas.
- There are efforts by traditional HPC platforms to combine the flexibility of cloud models with the performance of HPC systems.

5: Coordination and new models

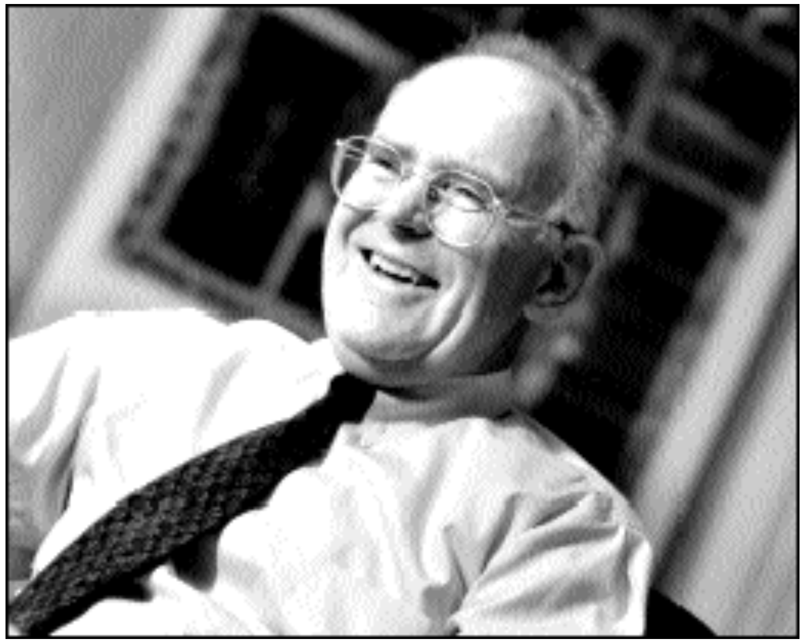
- [We don't have anything on this at the moment!]
- Coordination within DOE programs (Production & INCITE) and NSF on traditional and non-traditional HPC?
- Identity management
- Project management tools

Why You Need Parallel Computing: The End of Moore's Law?



2X transistors/Chip Every 1.5 years

Called "Moore's Law"



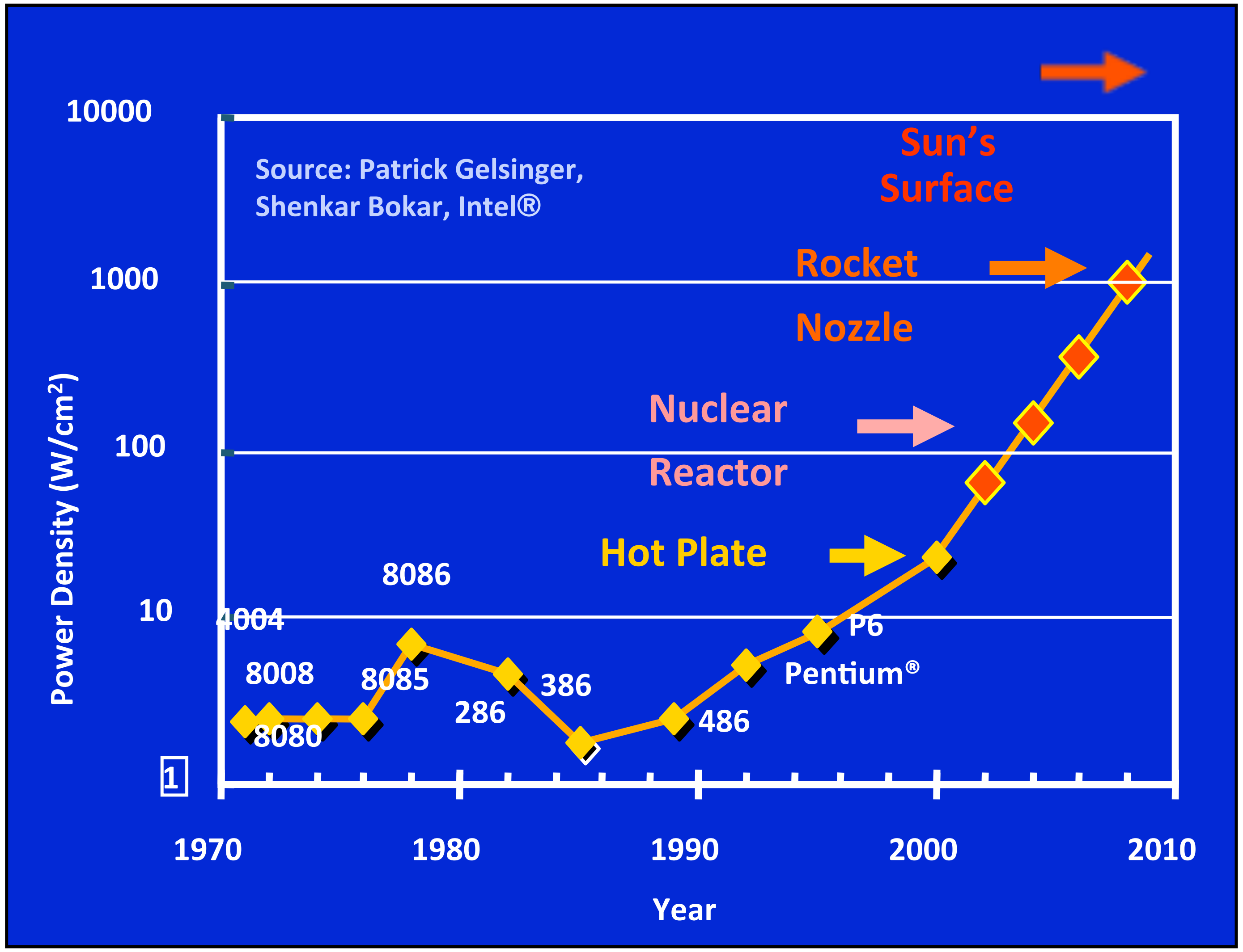
Microprocessors have become smaller, denser, and more powerful.

Gordon Moore (co-founder of Intel) predicted in 1965 that the transistor density of semiconductor chips would double roughly every 18 months.

Slide source: Jack Dongarra

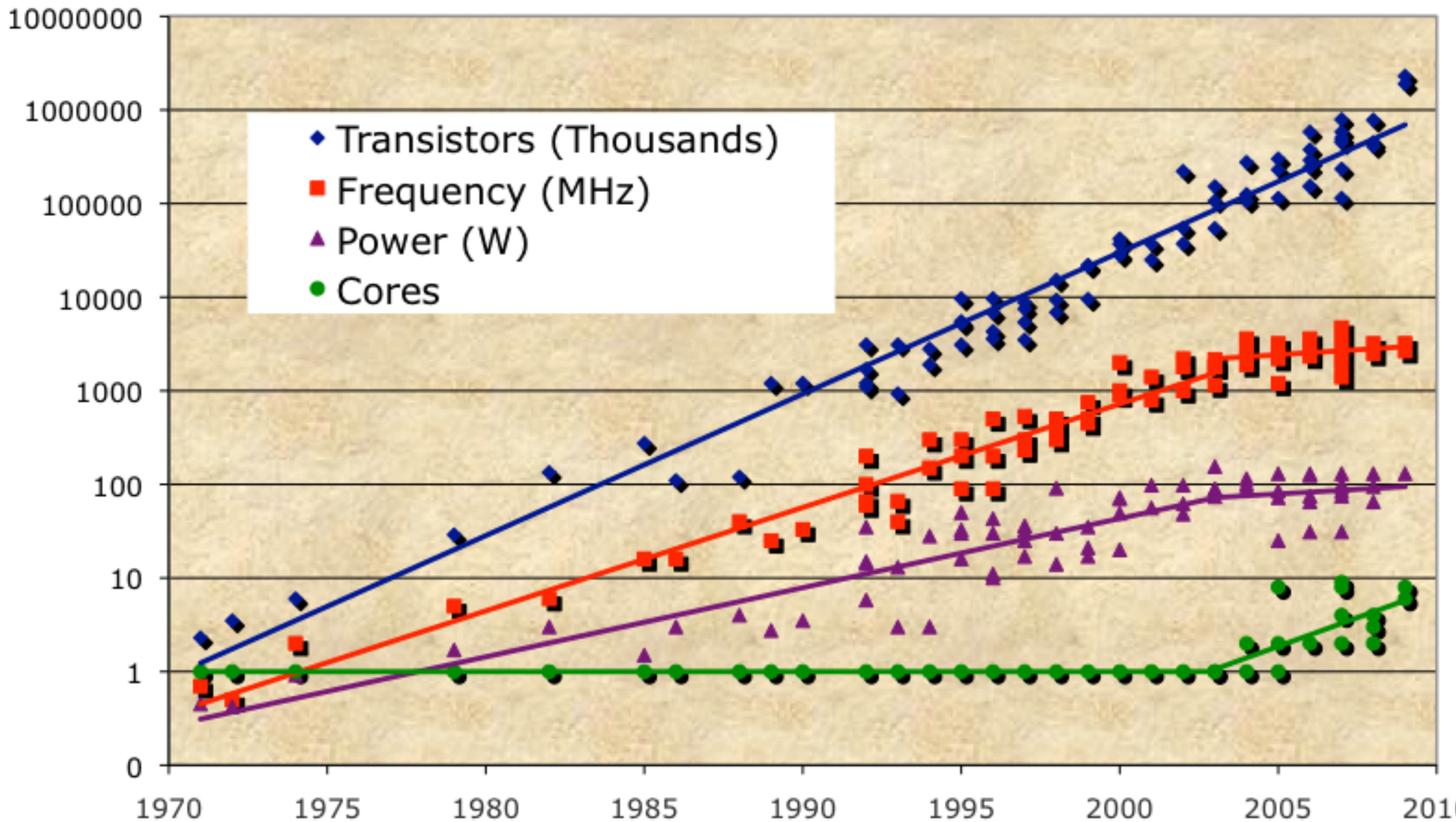
Power Density Limits Serial Performance

- Concurrent systems are more power efficient
  - Dynamic power is proportional to V^2fC
  - Increasing frequency (f) also increases supply voltage (V) → cubic effect
  - Increasing cores increases capacitance (C) but only linearly
  - Save power by lowering clock speed



- High performance serial processors waste power
  - Speculation, dynamic dependence checking, etc. burn power
  - Implicit parallelism discovery
- More transistors, but not faster serial processors

Revolution in Processors

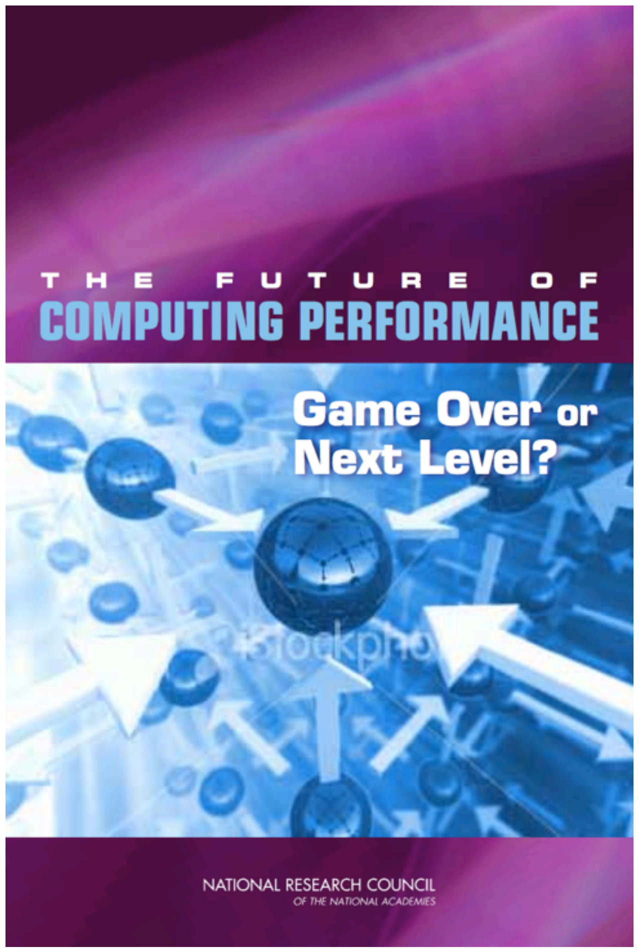


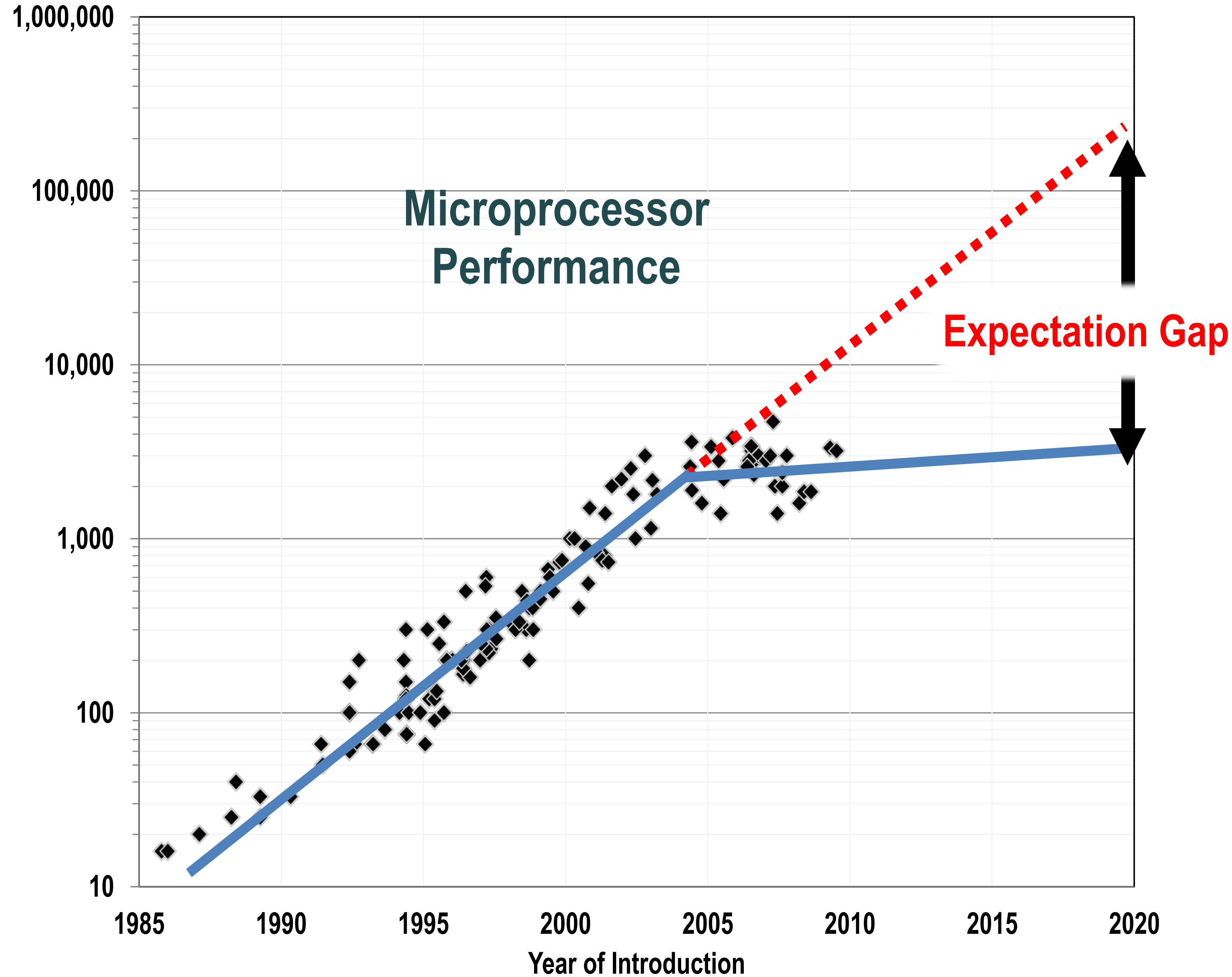
- Chip density is continuing increase ~2x every 2 years
- Clock speed is not
- Number of processor cores may double instead
- Power is under control, no longer growing

Moore's Law Reinterpreted

- Number of cores per chip will increase
- Clock speed will not increase (~~possibly~~ probably decrease)
- Need to deal with systems with millions of concurrent threads
- Need to deal with inter-chip parallelism (OpenMP threads) as well as intra-chip parallelism (MPI)
- Any performance gains are going to be the result of increased parallelism, not faster processors

Serial Processing = Left Behind





Still to do

- Could use more input from the E and T groups on estimates of their needs
  - We are currently not very quantitative
  - At the same time, given the uncertainty of the timelines for so many major projects, how accurately can we predict how much computing will be needed when anyway?
    - NERSC report only tries to go out to 2017, $\simeq$ tomorrow
  - Report outline in decent shape, but more writing needed