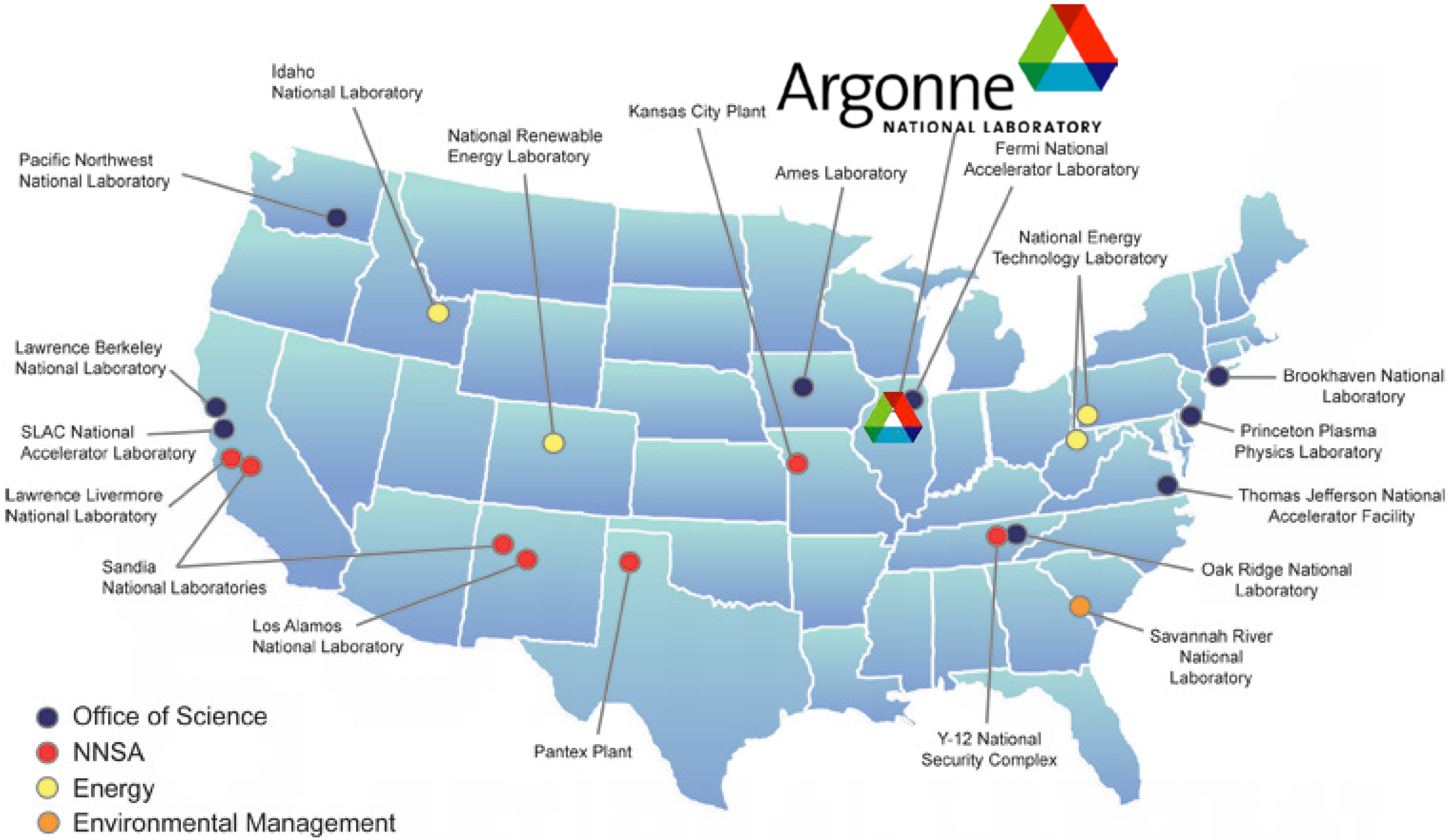
An Introduction to Parallel Supercomputing

Pete Beckman

Argonne National Laboratory

Argonne & the DOE National Laboratory System



Direct descendent of Enrico Fermi's Metallurgical Laboratory

- Opened in Feb 1943 (as new site for Chicago's Metallurgical Laboratory)
- Became Argonne National Laboratory in July 1946 (first national laboratory)

Argonne National Laboratory

- $675M /yr budget
- 3,200 employees
- 1,450 scientists/eng
- 750 Ph.D.s


Argonne
NATIONA LABO

MCS Division meeting c. 1983

- "If our R&D is going to be relevant ten years from now, we need to shift our attention to parallel computer architectures"
- "Los Alamos has a Denelcor HEP: let's experiment with it"

Advanced Computing Research Facility: 1984 -1992

- **The Advanced Computing Research Facility (ACRF) was established in recognition of the role that parallel computers would play in the future of scientific computing.**

- **Principal objectives:**
  - To encourage experimentation on computers with innovative designs
  - To assess the suitability of diverse machines for specific applications
  - To assist research in parallel computation
  - To encourage the incorporation of state-of-the-art computational techniques in research, development and industry.
  - To provide leadership in enhancing computing environments
  - To operate as a national user facility in parallel computing

ACRF Contributed to CRPC's Computational Resources

NSF S&T Center with Ken Kennedy as head

CRPC
CENTER FOR RESEARCH ON PARALLEL COMPUTATION

- **The CRPC has chosen to build a physically distributed, shared computing resource.**
  - Each member institution has research computing facilities with high-performance parallel computers available. The following list includes some of the high-performance computing facilities available to the CRPC:

- **BBN Butterfly GP1000, 96 nodes**
- **BBN Butterfly TC 2000, 45 nodes**
- **CM-2, 8K processors**
- **CM-5, 1024 processors**
- **CM-5, 32 nodes**
- **CRAY T3D, 256 nodes**
- **IBM SP1, 128 nodes**
- **IBM SP1, 8 nodes**
- **Intel Paragon A4, 60 nodes**
- **Intel Paragon L38, 512 nodes**
- **Intel Touchstone Delta, 570 nodes**
- **Intel iPSC/860, 32 nodes**
- **Intel iPSC/860, 64 nodes**
- **Intel iPSC/860, 8 nodes**
- **Kendall Square Research KSR-1**
- **MasPar DEC Mpp 12000 (3)**
- **nCube/2, 64 nodes (2)**
- **SGI 380VGX, 8 processors**
- **Sequent Symmetry, 26 processors**

The ACRF Summer Institutes (Not Dead Yet!)

- Sponsored by NSF
- Held each September, 1987-1989
- Mornings: long lectures by distinguished speakers
- Afternoons: hands-on experience with Argonne software on ACRF machines
- Some speakers: Gordon Bell, Bill Buzbee, Josh Fisher, Dave Kuck, Neil Lincoln, Chuck Seitz, Larry Smarr, Burton Smith, Guy Steele, Don Austin, Mani Chandy, Arvind, Tom DeFanti, David Gelernter, John Gurd, Ken Kennedy, Alex Nicholau, others
- Admission by application and review, about 20 students each summer
- Next Summer Institute: July 29 – August 9, 2013


MIRA

POOMA Project: 1996
John Reynders

Parallel Platform Paradox

"The average time required to implement a moderate-sized application on a parallel computer architecture is equivalent to the half-life of the latest parallel supercomputer."

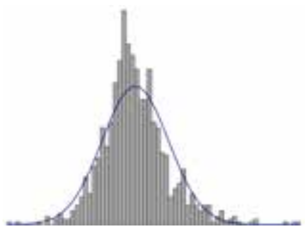
"Although a strict definition of "half-life" could be argued, no computational physicist in the fusion community would dispute the face that most of the time spent implementing parallel simulations was focused on code maintenance, rather than on exploring new physics. Architectures, software environments, and parallel languages came and went, leaving the investment in the new physics code buried with the demise of the latest supercomputer. There had to be a way to preserve that investment."

Pete's Investment Recommendations

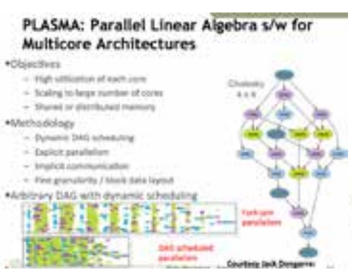
- **Other People's Libraries**
- **Encapsulation**
  - Parallelism & Messaging & I/O
- **Embedded Capabilities**
  - Debugging
  - Performance Monitoring
  - Correctness Detection
  - Resilience
- **The Two Workflow Views**
  - Science: (problem setup, analysis, etc.)
  - Programmer: (mod, testing, document, commit)
- **Automation**
  - A+ Build system, nightly test and build, configuration
  - Embedded versioning and metadata
- **Community: web, tutorial, email, bug tracking, etc**

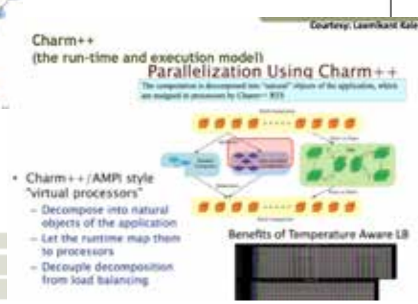
Threads/Tasks: Managing Exploding Parallelism

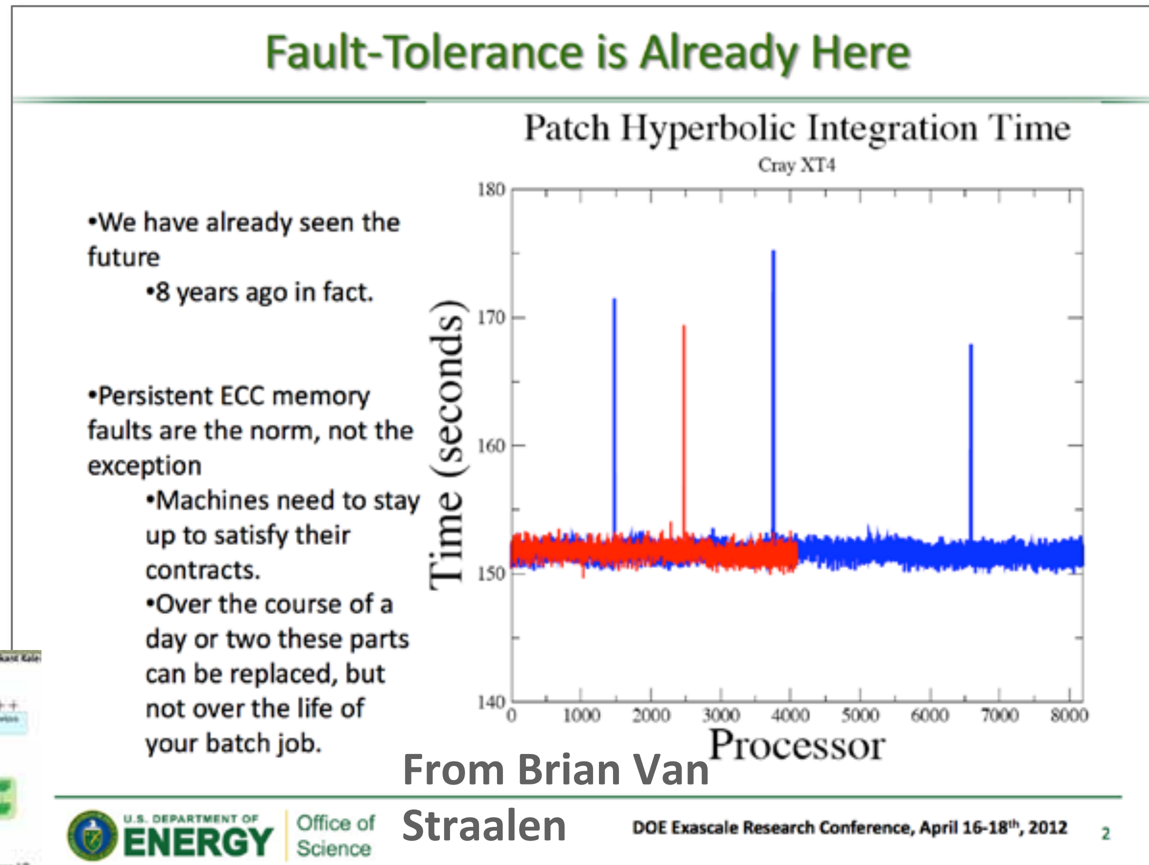
- **Dynamic parallelism and decomposition**
  - Programmer cannot hand-pick granularity / resource mapping
    - (equal work != equal time)



Variability is the new norm:
Power
Resilience
Intranode Contention







From Brian Van Straalen

Future Trends: Invest Wisely

Trending Up	Trending Down
Asynchrony, Latency Hiding	Block synchronous
Over Decomp & Load Balancing	Static partitioning per core
Massive Parallelism	Countable parallelism
Reduced RAM per Flop	Large memory coherence domains
Expensive Data Movement	Expensive flops
Fault / Resilience	Pure checkpoint/restart
Low BW to Storage, in-situ analysis	Save all, let the viz guys sort it out