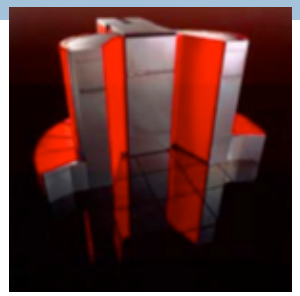
http://www.exascale.org

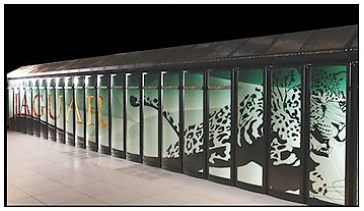
Pete Beckman & Jack Dongarra

Looking at the Gordon Bell Prize
(a slide from Cray :-)



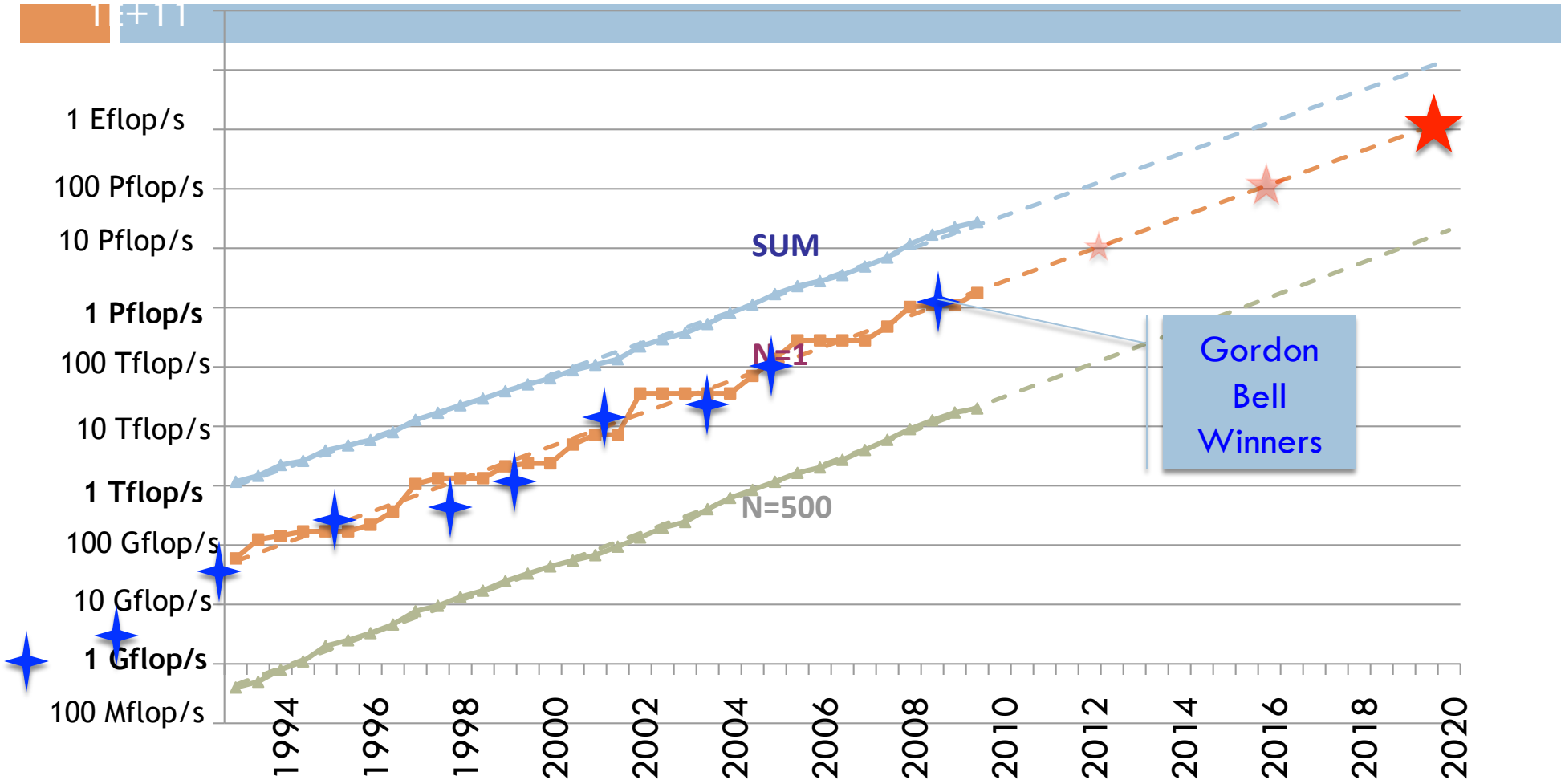
- 1 GFlop/s; 1988; Cray Y-MP; 8 Processors
  - Static finite element analysis

- 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



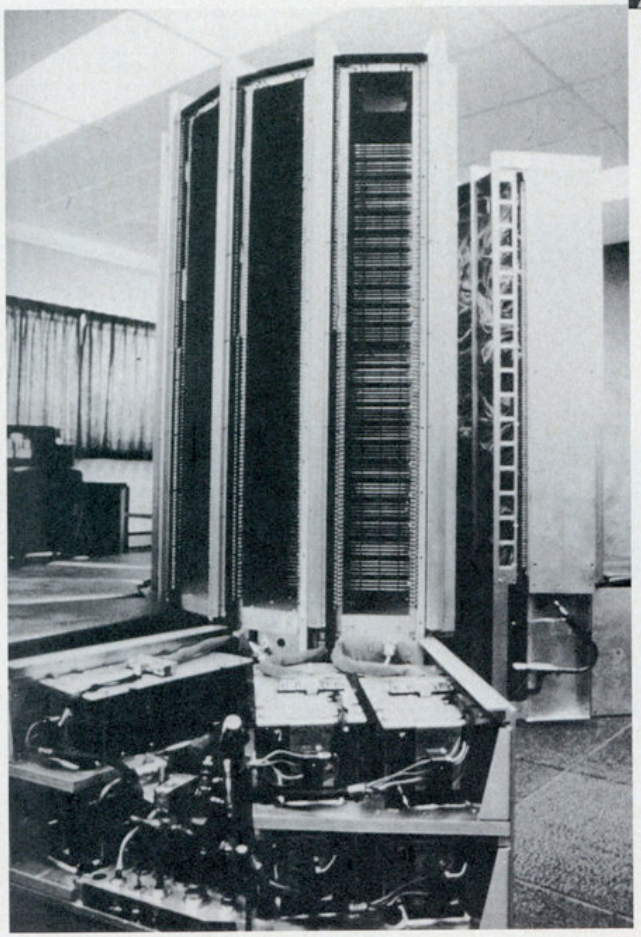
- 1 EFlop/s; ~2018; ?; 1×10^7 Processors (10^9 threads)

Performance Development in Top500



Supercomputer

Incredible Cray-1 cruises at 80 million operations a second



CRAY-1 computer is not much larger than its inventor, Seymour Cray. Outer seats cover the power supply (see below).

It's 10 times faster than the biggest IBM, with six times more memory

By JIM SCHEFTER

"Step into the computer," said my guide.

I did, and felt the chilling sensation of moving into the megabit maw of a machine so advanced in electronic intellect that it can only talk to other machines.

For a moment, I was apprehensive. It was like entering a silicon crypt. The air was significantly colder inside the polyhedral chamber than outside. Yet I knew that I was surrounded by hundreds of thousands of heat-producing electronic circuits, drawing six times as much electricity as any other machine in the room.

This was the CRAY-1, the amazing supercomputer designed by a reclusive Wisconsin genius. It's 10 times faster than the biggest IBM computer on the market. And this particular CRAY-1, installed in a major computer center in Kansas City, was being fed by two giant Control Data computers just to keep it busy.

"You're looking at the architecture of Seymour Cray," said a voice floating over the top of the computer.

The voice belonged to Jack Lorenz, president of United Computer Systems and owner of the first commercially installed CRAY-1 system. I saw what he meant. The CRAY-1 is unique, not only in electronic architecture and performance, but in size and shape as well. It doesn't look like any other computer.

Extreme-Scale Platform Design:

Industrial revolution and globalization has arrived

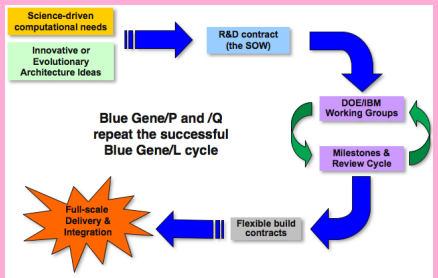
Yesterday **Today** **Tomorrow**

Seymour & team designs and hand builds set of computers

Dozen HPC companies flourish: incompatible OS & components

Commodity components and Open Source move effort to integration

Design-Build partnerships for extreme machines (e.g. LLNL/ANL/IBM)



Japan 10PF: Hitachi, NEC, Riken, Universities

Globally Distributed teams, Diverse technology providers, Open Source Software

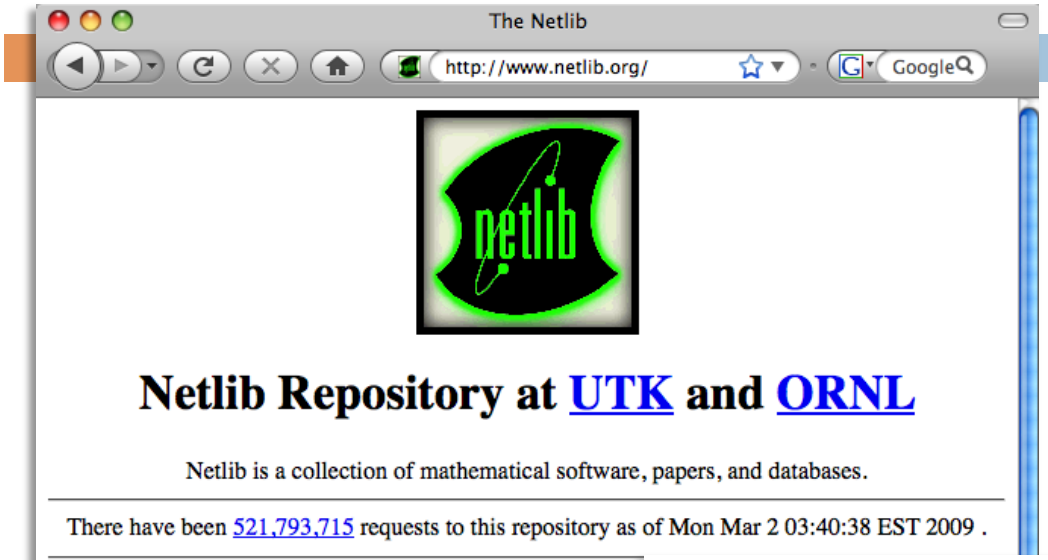
The Community is Diverse and Robust

- Over the last 10 years, the galvanization of the Open Source movement has dramatically improved HPC software

A very small sample:

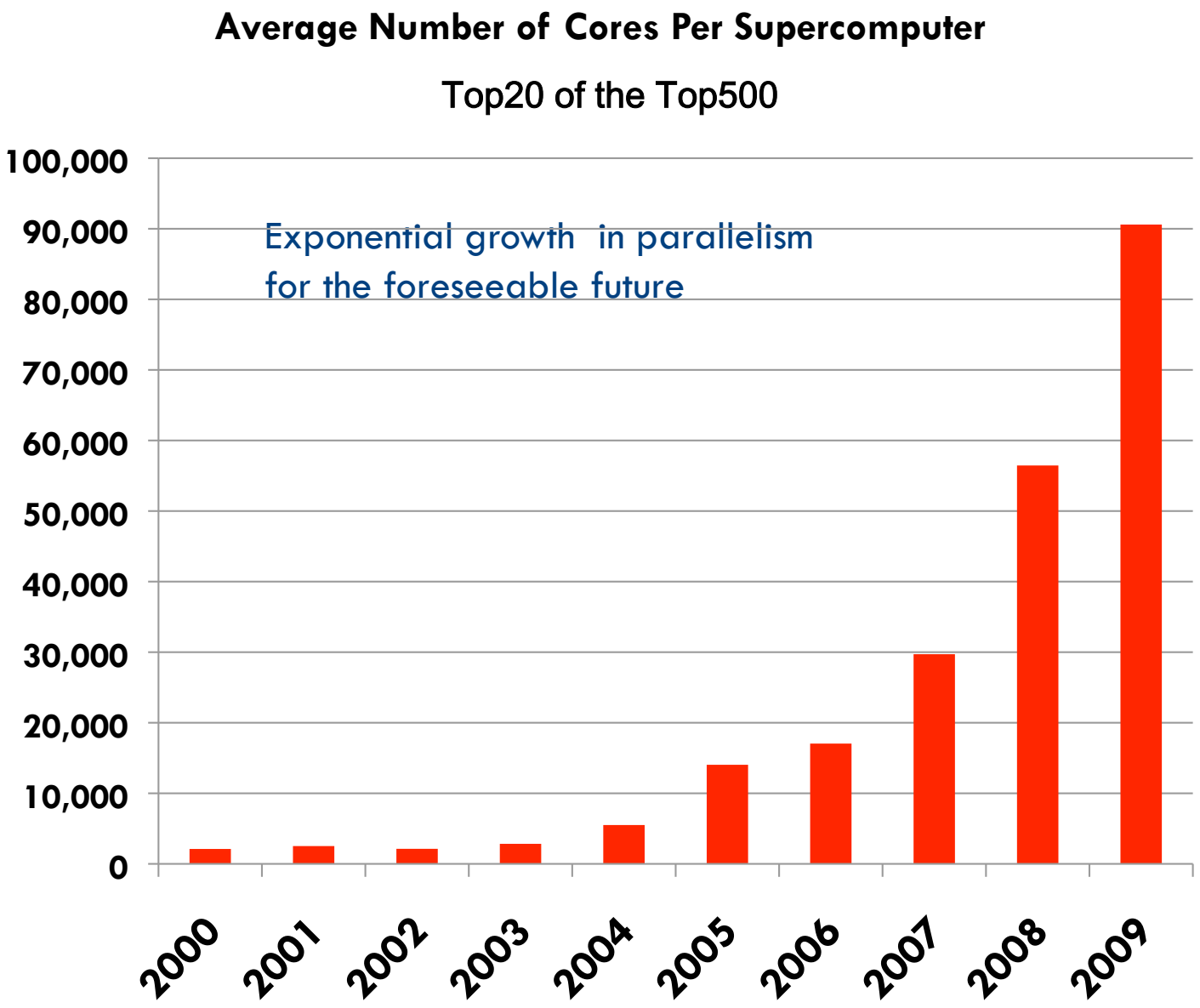
- Linux Operating System, libc
- Python, Perl
- PAPI, TAU, Kojak
- UPC
- MPICH, OpenMPI
- ScaLAPACK
- VisIt
- GASNet, ARMCI/GA
- PVFS
- CFEngine, bconfig
- Ganglia
- SLURM, Cobalt
- Dyninst
- Torque/Moab, OpenPBS
- Charm++
- pNetCDF, HDF5
- GridFTP
- FFTW

A Long History of Collaboration & Sharing



The massive archive site WSMR-SIMTEL20.ARMY.MIL at White Sands Missile Range, New Mexico, USA, which is home to more than 2 gigabytes of files for many computer systems, including MSDOS, Unix, VMS and some mainframes, will be shut down by its operators as of September 20, 1993. Unless a new home is found for the archives, this major archive site will vanish.

The Result….


Average Number of Cores Per Supercomputer
Top20 of the Top500
Exponential growth in parallelism
for the foreseeable future
100,000
90,000
80,000
70,000
60,000
50,000
40,000
30,000
20,000
10,000
0
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009

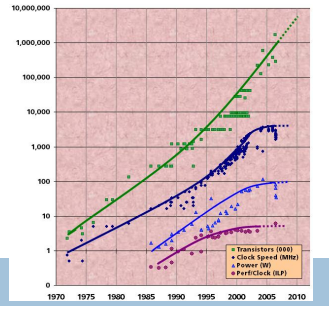
Factors that Necessitate Redesign

- **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**

Potential System Architectures

Systems	2015	2018-2020
System peak	100-200 Pflop/s	1 Eflop/s
System memory	5 PB	10 PB
Node performance	200-400 Gflop/s	1-10 Tflop/s
Node memory bandwidth	100 GB/s	200-400 GB/s
Node concurrency	O(100)	O(1000)
Interconnect bandwidth	25 GB/s	50 GB/s
System size (nodes)	O(100,000)	100,000-1,000,000
Total concurrency	O(50,000,000)	O(1,000,000,000)
Storage	150 PB	300 PB
IO	10 TB/s	20 TB/s
MTTI	days	O(1 day)
Power	~10 MW	~20 MW

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

Four Goals for IESP

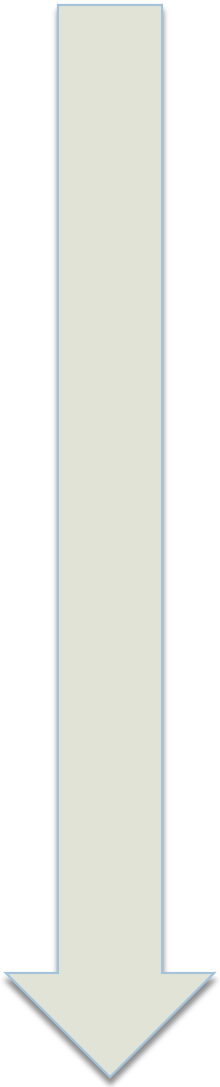
- **Strategy for determining requirements**
  - clarity in scope is the issue
- **Comprehensive software roadmap**
  - goals, challenges, barriers and options
- **Resource estimate and schedule**
  - scale and risk relative to hardware and applications
- **A governance and project coordination model**
  - Is the community ready for a project of this scale, complexity and importance?
  - Can we be trusted to pull this off?

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
- Funding from DOE's Office of Science & NSF Office of Cyberinfratructure
- US meeting (Santa Fe, NM) April 6-8, 2009
  - 65 people
- NSF's Office of Cyberinfrastructure funding
- European meeting (Paris, France) June 28-29, 2009
  - 70 people
  - Outline Report
- Asian meeting (Tsukuba Japan) October 18-20, 2009
  - Draft roadmap
  - Refine Report
- SC09 (Portland OR) BOF to inform others
  - Public Comment
  - Draft Report presented
- Oxford meeting, April 2010



Nov 2008

Apr 2009

Jun 2009

Oct 2009

Nov 2009

Roadmap Purpose

- The IESP software roadmap is a planning instrument designed to enable the international HPC community to improve, coordinate and leverage their collective investments and development efforts.

- After we determine what needs to be accomplished, our task will be to construct the organizational structures suitable to accomplish the work

Key Trends

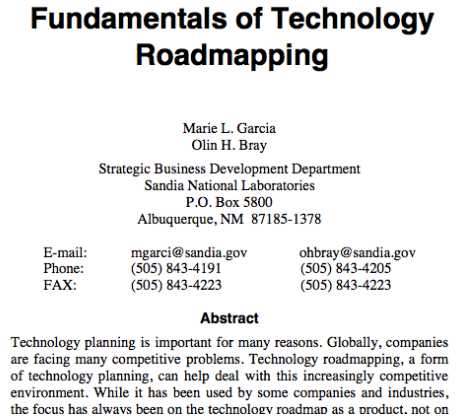
- Increasing Concurrency
- Reliability Challenging
- Power dominating designs
- Heterogeneity in a node
- I/O and Memory: ratios and breakthroughs

Requirements on X-Stack

- Programming models, applications, and tools must address concurrency
- Software and tools must manage power directly
- Software must be resilient
- Software must address change to heterogeneous nodes
- Software must be optimized for new Memory ratios and need to solve parallel I/O bottleneck

Four Goals for IESP

- **Strategy for determining requirements**
  - clarity in scope is the issue
- **Comprehensive software roadmap**
  - goals, challenges, barriers and options
- **Resource estimate and schedule**
  - scale and risk relative to hardware and applications
- **A governance and project coordination model**
  - Is the community ready for a project of this scale, complexity and importance?
  - Can we be trusted to pull this off?



Fundamentals of Technology Roadmapping

Marie L. Garcia
Olin H. Bray

Strategic Business Development Department
Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185-1378

E-mail: mgarci@sandia.gov ohbray@sandia.gov
Phone: (505) 843-4191 (505) 843-4205
FAX: (505) 843-4223 (505) 843-4223

Abstract

Technology planning is important for many reasons. Globally, companies are facing many competitive problems. Technology roadmapping, a form of technology planning, can help deal with this increasingly competitive environment. While it has been used by some companies and industries, the focus has always been on the technology roadmap as a product, not on

Goals for IESP

- **Develop a comprehensive community software roadmap for Exascale systems**
  - Identify those software capabilities that will be needed for fully functional exascale systems, what are the barriers and how can we overcome them
  - Determine which elements will occur naturally and which elements need R+D investment
  - Determine those components that have solid starting points and which that need *ab initio* efforts
  - Determine which components are suitable for an open community development model

Goals for IESP

- **Develop an estimate of the resources required and timeline needed to develop the required software**
  - Need to put the software element of exascale in appropriate budget and schedule context
  - Need to understand the risks (technical, schedule and organizational)
  - Need to distinguish between the applications software efforts and the systems software
  - The software timeline should be aligned with that of the hardware (and precede it where possible)

Roadmap Components

Co-Design Vehicles

- Requirements:
  - Terascale today. Demonstrated need for exascale
  - Can achieve significant scientific impact in an important area such as climate, eng., lifesci, materials, physics
  - A realistic and productive development pathway to exascale can be mapped out over 10 years
  - Community has demonstrated experience in algorithm, software and/or hardware developments and willing to engage in the exascale co-design process

An Example Development Community

The Apache Software Foundation

Meritocracy in Action.

The Apache Software Foundation provides support for the Apache community of open-source software projects. The **Apache projects** are characterized by a collaborative, consensus based development process, an open and pragmatic software license, and a desire to create high quality software that leads the way in its field.

We consider ourselves not simply a group of projects sharing a server, but rather a *community of developers and users*.

This page will give you everything you always wanted to know about the foundation but were afraid to ask. The difference between membership and committership, who decides what, how elections take ... what's the philosophy behind ... nt. Come and see behind the

Project	Sponsor	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
Bluesky	Incubator	2008-01-12	416	True	month	False	True	2008-07-22	224	0,0,8	7	0	True	False	True	True	True	False	False
Cassandra	Incubator	2009-01-01	61	True	month	False	True	2009-01-02	60	1,1,-	9	0	True	True	True	True	False	False	False
Click	Incubator	2008-07-21	225	False	group-3	True	True	2009-02-22	9	2,3,4	7	0	True	True	True	True	True	True	True
Composer	Incubator	2007-11-17	472	False	group-3	True	True	2008-10-09	145	0,0,1	-	0	True	False	True	True	True	False	False
Droids	HC, Lucene	2008-10-09	145	False	group-2	True	True	2008-10-23	131	0,0,3	4	0	True	True	True	True	True	False	False
Empire-db	Incubator	2008-07-08	238	False	group-1	True	True	2009-01-05	57	1,1,4	8	1	True	True	True	True	True	True	True
ESME	Incubator	2008-12-02	91	True	group-3	True	True	2008-12-05	88	0,1,-	10	0	True	True	True	True	True	False	False
Etch	Incubator	2008-09-02	182	False	group-3	True	True	2008-12-08	85	0,3,4	8	0	True	True	True	True	True	False	False
Hama	Incubator	2008-05-20	287	False	group-3	True	True	2008-11-18	105	0,1,3	6	0	True	True	True	True	True	False	False
Imperius	Incubator	2007-11-10	479	False	group-1	True	True	2009-02-05	26	1,1,2	9	1	True	True	True	True	True	False	False
JSecurity	Incubator	2008-05-20	287	False	group-1	True	True	2008-09-28	156	0,0,2	10	0	True	True	True	True	True	False	False
JSPWiki	Incubator	2007-09-17	533	False	group-1	True	True	2008-09-28	156	0,0,1	12	1	True	True	True	True	True	False	False
Kato	? not known		?	True	group-3	True	False		?		9	0	True	True	True	True	True	False	False
Log4php	Logging Services	2004-01-31	1858	False	group-3	True	True	2007-07-18	594	0,0,0	6	0	True	True	True	True	True	False	False
Lokahi	Incubator	2006-03-01	1098	False	group-2	True	True	2006-11-28	826	0,0,0	5	0	True	True	True	True	True	False	False
Lucene.Net	Lucene	2006-03-15	1084	False	group-1	True	True	2006-11-11	843	0,0,0	4	1	True	True	True	True	True	False	False
Olio	Incubator	2008-09-29	155	False	group-1	True	True	2009-02-05	26	1,3,5	12	0	True	True	True	True	True	True	False
OpenWebBeans	Incubator	2008-10-26	128	False	group-3	True	True	2009-01-18	44	1,7,7	6	2	True	True	True	True	True	True	True
PDFBox	Incubator	2008-02-07	390	False	group-2	True	True	2009-01-27	35	3,3,3	8	2	True	True	True	True	True	False	False
PhotArk	Incubator	2008-08-19	196	False	group-2	True	True	2008-10-26	128	0,0,4	6	0	True	True	True	True	True	False	False
Pivot	Incubator	2009-01-26	36	True	group-2	True	True	2009-02-20	11	1,-,-	3	0	False	True	True	True	True	False	False
RAT	Incubator	2008-01-06	422	False	group-3	True	True	2009-02-22	9	3,3,3	8	0	True	True	True	True	True	False	False
RCF	MyFaces	2007-04-06	697	False	group-1	True	True	2008-04-16	321	0,0,0	17	0	True	True	True	True	False	False	False
River	Incubator	2006-12-26	798	False	group-3	True	True	2008-11-10	113	0,1,1	19	0	True	True	True	True	True	True	True
Sanselan	Incubator	2007-09-09	541	False	group-1	True	True	2008-09-28	156	0,0,4	6	0	True	True	True	True	True	True	True

...n incorporated in the United

...ects by supplying hardware,

- create an independent legal entity to which companies and individuals can donate resources and be assured that those resources will be used for the public benefit
- provide a means for individual volunteers to be sheltered from legal suits directed at the

Apache Projects

- HTTP Server
- Abdera
- ActiveMQ
- Ant
- APR
- Archiva
- Beehive
- Camel
- Cayenne
- Cocoon
- Commons
- Continuum
- CouchDB
- CXF
- DB
- Directory
- Excalibur
- Felix
- Forrest
- Geronimo
- Gump
- Hadoop
- Harmony
- HiveMind
- HttpComponents
- iBATIS

Foundation

- FAQ
- Licenses
- News
- Public Record
- Sponsorship
- Donations
- Thanks
- Contact

Foundation Proje

- Conferences
- Infrastructure
- JCP
- Legal Affairs
- Security
- Travel Assistar

How it works

- Introduction
- Meritocracy
- Structure
- Roles
- Collaboration
- Infrastructure
- Incubator
- Other entities

Apache Foundation

- Create a foundation for open, collaborative software development projects by supplying hardware, communication, and business infrastructure
- Incubator projects can become Apache projects
- 800 “committers”
- The ASF Infrastructure is mostly composed of the following services:
  - the web serving environment (web sites and wikis)
  - the code repositories
  - the mail management environment
  - the issue/ bug tracking
  - the distribution mirroring system

Next Steps

- Refine roadmap
- Develop organizational models
- Divide and conquer

http://www.exascale.org