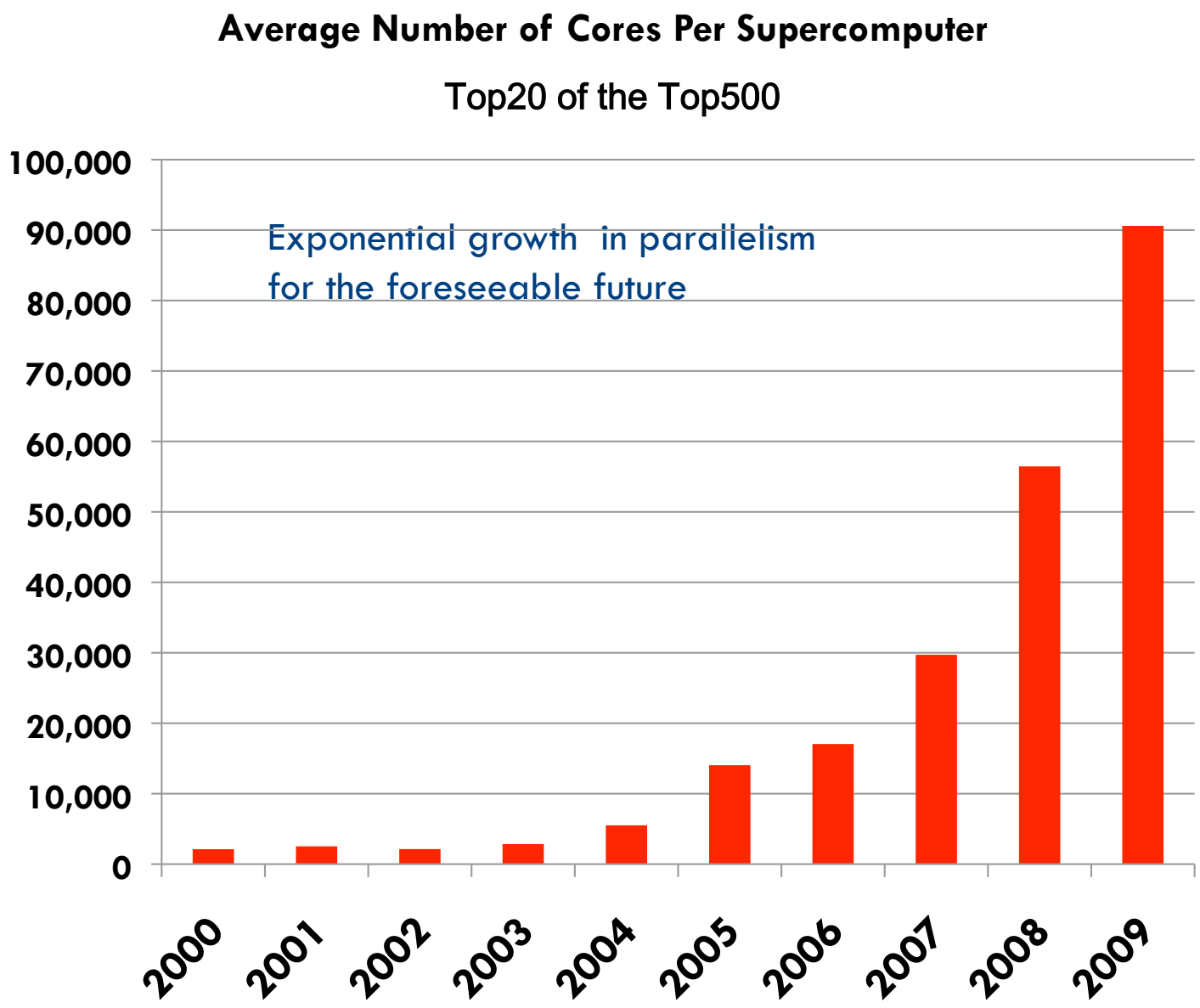
http://www.exascale.org

Pete Beckman & Jack Dongarra


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

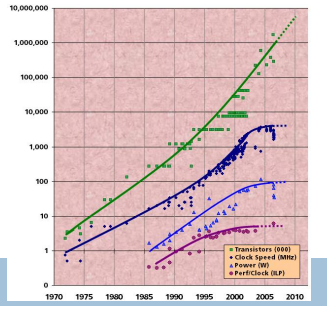
Potential System Architecture

Systems	2009	2018	Difference Today & 2018
System peak	**2 Pflop/s**	**1 Eflop/s**	**O(1000)**
Power	**6 MW**	**~20 MW**	
System memory	0.3 PB	32 - 64 PB [**.03 Bytes/Flop**]	O(100)
Node performance	125 GF	1,2 or 15TF	O(10) – O(100)
Node memory BW	25 GB/s	2 - 4TB/s [**.002 Bytes/Flop**]	O(100)
Node concurrency	12	O(1k) or 10k	O(100) – O(1000)
Total Node Interconnect BW	3.5 GB/s	200-400GB/s (1:4 or 1:8 from memory BW)	O(100)
System size (nodes)	18,700	O(100,000) or O(1M)	O(10) – O(100)
Total concurrency	225,000	O(billion) [O(10) to O(100) for latency hiding]	O(10,000)
Storage	15 PB	500-1000 PB (>10x system memory is min)	O(10) – O(100)
IO	0.2 TB	60 TB/s (how long to drain the machine)	O(100)
MTTI	days	O(1 day)	- O(10)

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**

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 <u>open source software</u> for scientific high-performance computing

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

Apr 2010

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

Roadmap Components

DOE Extreme-Scale Workshops

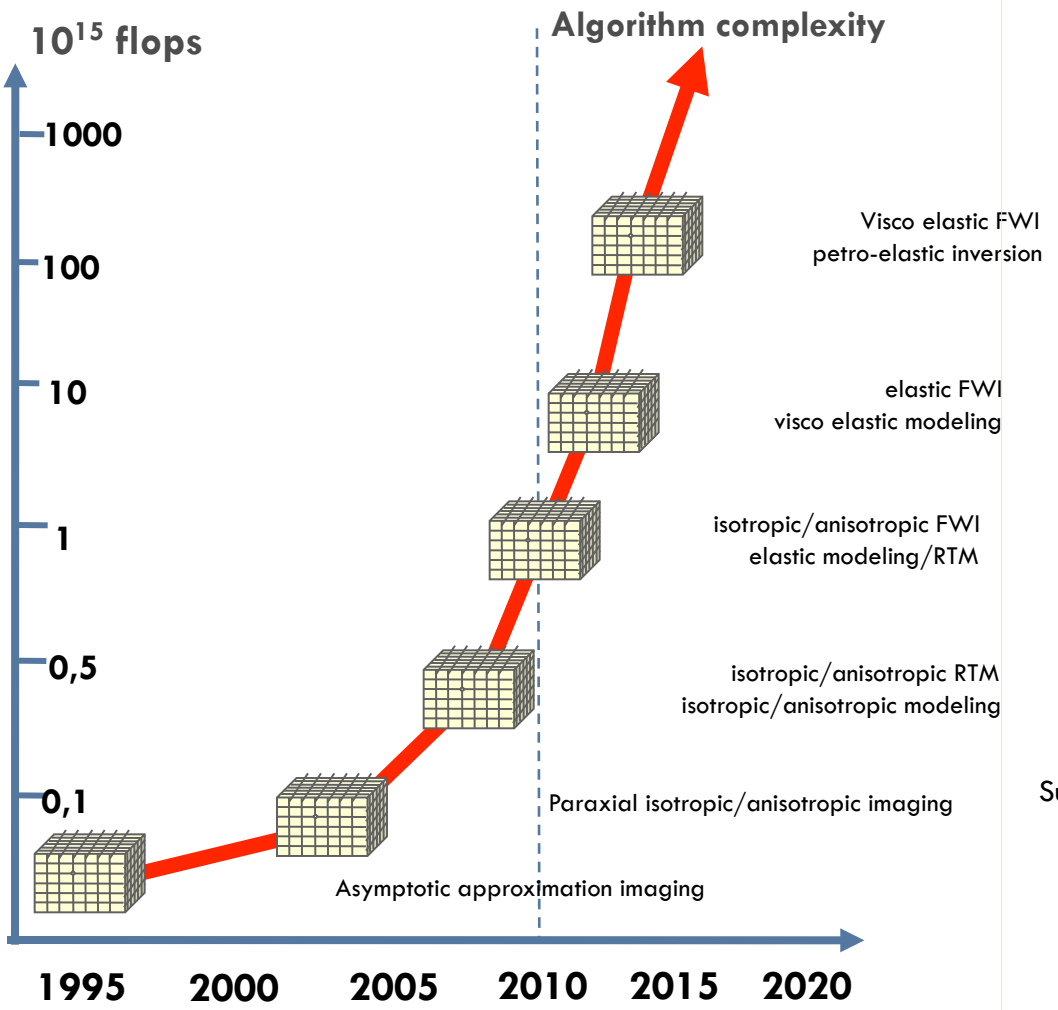
- **Crosscutting Workshop: February 2-4, 2010**
- **Architectures and Technology: December 8-10, 2009**
- **National Security: October 6-8, 2009**
- **Biology: August 17-19, 2009**
- **Basic Energy Sciences: August 13-15, 2009**
- **Nuclear Energy: May 11-12, 2009**
- **Fusion Energy Sciences: March 18-20, 2009**
- **Nuclear Physics: January 26-28, 2009**
- **High Energy Physics: December 9-11, 2008**
- **Climate Science: November 6-7, 2008**

DOE Extreme-Scale Workshops

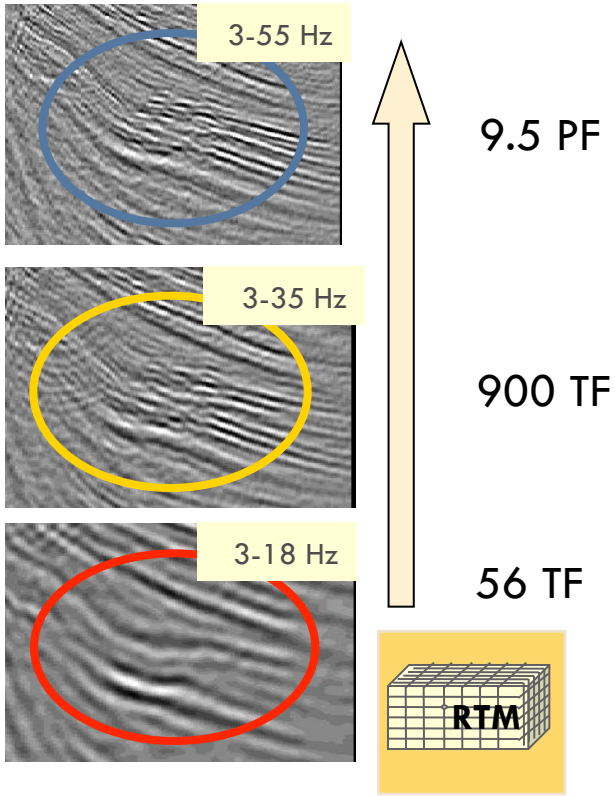
http://extremecomputing.labworks.org/

- To identify grand challenge scientific problems in [research area] that can exploit computing at extreme scales to bring about dramatic progress toward their resolution.
- The goals of the workshops are to
  - identify grand challenge scientific problems [...] that could be aided by computing at the extreme scale over the next decade;
  - identify associated specifics of how and why new high performance computing capability will address issues at the frontiers of [...]; and
  - provide a forum for exchange of ideas among application scientists, computer scientists, and applied mathematicians to maximize the use of extreme scale computing for enabling advances and discovery in [...].

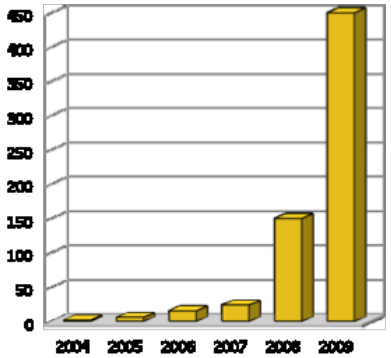
Industrial challenges in the Oil & Gas industry: Depth Imaging roadmap



Algorithmic complexity Vs. corresponding computing power



Substained performance for different frequency content over a 8 day processing duration



HPC Power
PAU (TF)

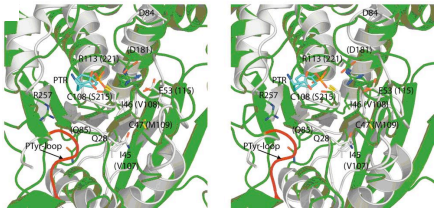
courtesy

From sequences to structures : HPC Roadmap

2009 | **2011** | **2015 and beyond**

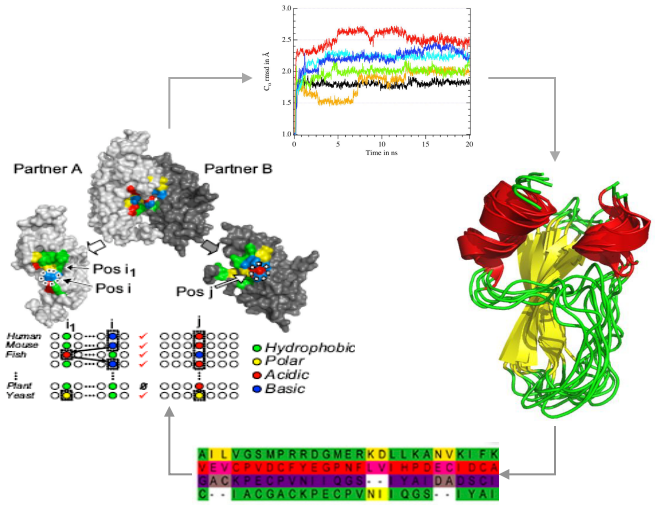
Grand Challenge GENCI/CCRT

```
gi|7380613  (Nm)   105 LAYCRTGT-RCS 115
gi|33571275 (Bp)   94  LAYCRTGT-RCT 104
gi|14027624 (Ml)   91  FAYCRSGA-RCT 101
gi|24461549 (Pa)   91  FAYCRTGT-RSA 101
gi|67672005 (Bp)   91  LAYCRTGM-RSA 101
gi|12734660 (Hs)   137 LIHCYGGLGRSC 148
gi|72014257 (Sp)   152 LVHCFGGIGRSS 163
gi|48856790 (Ch)   405 LIHCVGGLGRSG 416
gi|35211553 (Gv)   142 VIHCKGGLGRTG 153
```

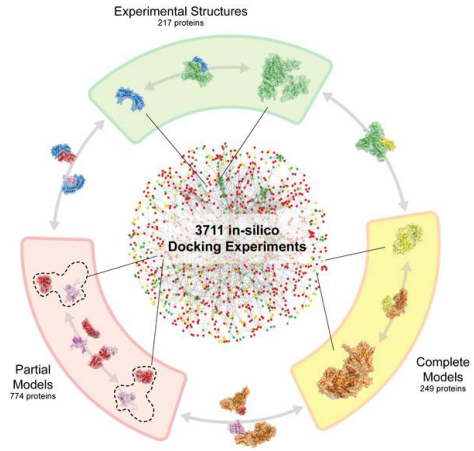


Proteins 69 (2007) 415

Identify all protein sequences using public resources and metagenomics data, and systematic modelling of proteins belonging to the family (Modeller software).



Improving the prediction of protein structure by coupling new bio-informatics algorithm and massive molecular dynamics simulation approaches.



Systematic identification of biological partners of proteins.

Computations using more and more sophisticated bio-informatical and physical modelling approaches ⇒ Identification of protein structure and function

1 family	1 family	1 family
5.10^3 cpu/~week	5.10^4 cpu/~week	~ 10^4*KP cpu/~week
		CSP : proteins structurally characterized ~ 10^4
# 25 Gb of storage	# 5 Tb of storage	# 5*CSP Tb of storage
500 Gb of memory	5 Tb of memory	5*CSP Tb of memory

Applications serve as Co-Design Vehicles

- Technology drivers
  - *Advanced architectures with greater capability but with formidable software development challenges*
- Alternative R&D strategies
  - *Choosing architectural platform(s) capable of addressing PRD's of Co-Design Vehicles on path to exploiting Exascale*
- Recommended research agenda
  - *Effective collaborative alliance between Co-Design Vehicles , CS, and Applied Math with an associated strong V&V effort*
- Crosscutting considerations
  - *Identifying possible common areas of software development need among the Apps that serve as co-design vehicles*
  - *Addressing common need to attract, train, and assimilate young talent into this general research arena*

What Next? (1/3)

Moving from "What to Build" to "How to Build"

- Technology
  - Refining the roadmap for software and algorithms on extreme-scale systems
  - Setting a prioritized list of software components for Exascale computing as outlined in the Roadmap
  - Assessing the short-term, medium-term and long-term software and algorithm needs of applications for peta/exascale systems

What Next? (2/3)

Moving from "What to Build" to "How to Build"

- Organization
  - Developing a governance, management, and organizational structure for the IESP
  - Exploring ways for funding agencies to coordinate their support of IESP-related R&D so that they complement each other
  - Exploring how laboratories, universities, and vendors can work together on coordinated HPC software
  - Creating a plan for working closely with HW vendors and application teams to co-design future architectures

What Next? (3/3)

Moving from "What to Build" to "How to Build"

- Execution
  - Developing a strategic plan for moving forward with the Roadmap
  - Creating a realistic timeline for constructing key organizational structures and achieving initial goals
  - Exploring community development techniques and risk plans to ensure key components are delivered on time
  - Exploring key components of any needed Intellectual Property agreements

Breakout Groups

- Group 1: Technical challenges and needs of academic and industrial software infrastructure research and development
  Chair: Satoshi Matsuoka; Secretary: Michael Heroux
- Group 2: Computational challenges and needs for academic and industrial application communities
  Chair: Jean-Claude Andre; Secretary: Jean-Yves Berthou
- Group 3: Economic and management challenges and needs of computational resource providers and vendors Chair: Peg Williams; Secretary: Bill Kramer
- Group 4: Role and participation of national and international funding agencies
  Chair: Ed Seidel (Rob Pennington); Secretaries: Konstantinos Glinos and Patrick Aerts

Breakout Group "Work Product"

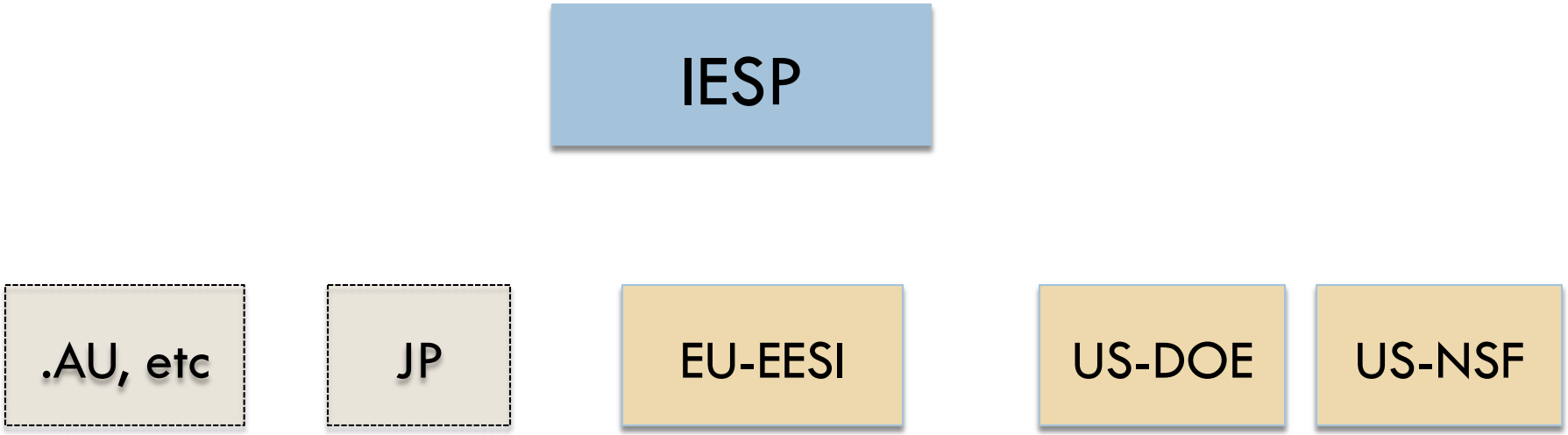
- At Meeting:
- 2-3 slides for each question, summarizing discussion
  - Slides will be presented in 15 min by breakout leads after each session

- After Meeting Synthesis
  - 10 page summary white paper will be written
  - Need volunteer writers / editors

Key Organizational Structure Concepts

- The IESP will focus on software, the hardware strategy will be directly driven by government-industry partnerships
- Each major participant will have their own key Exascale science applications driving development
- The community should help suggest co-design vehicles that represent broad international engagement

Example Organizational Structure: Incubation Period (today):



- IESP provides coordination internationally, while regional groups have well managed R&D plans and milestones

Example Organizational Structure: Maturity: (18 months?):

- LHC-like international organization
- Work is distributed, but a single organization manages the work

Breakouts…