U.S. DEPARTMENT OF ENERGY

Update on Exascale Research

Karen Pao
Program Manager
Advanced Scientific Computing Research
23 March 2011

U.S. DEPARTMENT OF ENERGY

Office of Science

ASCR Exascale Research Kick-off Meeting

- March 7-11, 2011 in San Diego, CA
- PI meeting for
  - Advanced Architecture, X-Stack, Scientific Data Management and Analysis at Extreme Scales awardees
  - Xstack planning team: An Open Integrated Software Stack for Extreme Scale Computing *(PIs: Pete Beckman (ANL), Kathy Yelick (LBNL), Al Geist (ORNL), Mike Heroux (SNL), Jack Dongarra (UTK))*
  - Seven Co-Design proposal teams
- Expected Outcomes:
  - Awareness within each solicitation communities and ASCR what members are doing and areas where they can leverage and supplement their work
  - Awareness across solicitation communities of what is going on and where each project fits in relation to the broad spectrum
  - Identification of gaps in ASCR exascale research portfolio
  - Lay groundwork for collaboration/cooperation with NNSA Exascale Roadmapping activities

Meeting material at
http://exascaleresearch.labworks.org/ascr2011

U.S. DEPARTMENT OF ENERGY

Office of Science

The PI meeting was facilitated & choreographed

- **Day 1: ASCR presentations (starting at 1:30 pm)**
- Day 2 – Day 3 morning: 3 poster sessions
  - each project representative made a 5-minute intro to the poster prior to poster session; project teams were to use poster sessions for networking
- Day 3 afternoon: working groups to identify crosscutting topics + outbrief
  - Each attendee was assigned to one of 8 Working Groups;
  - each WG was moderated by an NNSA facilitator, with two (ASCR PIs) reporters to present working group outbrief and a scribe (a student) to take notes
  - Day 3 dinner: ASCR & NNSA program offices identify five main topics
- Day 4: further discussions of crosscutting topics + outbrief
  - Morning: definition & interrelations (in same working group)
  - Afternoon: issues (in self-identified working groups)
- Day 5: gap analysis
  - Each attendee was assigned to a different WG
- In addition, there were homework assignments (due before 1st working session)
  - Exascale Research Coverage Assessment
  - Dance cards
- Results are being tabulated & analyzed

U.S. DEPARTMENT OF ENERGY

- Secretary: Steven Chu
  Deputy Secretary: Daniel B. Poneman
- Advanced Research Projects Agency – Energy: Arun Majumdar

- Under Secretary for Nuclear Security/Administrator for National Nuclear Security Administration: Thomas P. D'Agostino
  - Defense Nuclear Nonproliferation
  - Defense Programs
  - Naval Reactors
  - Counter-terrorism
  - Defense Nuclear Security
  - Emergency Operations
- Under Secretary for Science: Steven E. Koonin
  - Office of Science: William Brinkman, Patricia Dehmer
    - Basic Energy Sciences: Harriet Kung
    - High Energy Physics: Mike Procario(A)
    - Advanced Scientific Computing Research: Daniel Hitchcock (A)
    - Nuclear Physics: Tim Hallman
    - Biological & Environmental Research: Sharlene Weatherwax (A)
    - Fusion Energy Sciences: Ed Synakowski
    - SBIR/STTR: Manny Oliver
    - Workforce Develop. for Teachers & Scientists: Bill Valdez
- Under Secretary: Cathy Zoi (A)
  - Energy Efficiency & Renewable Energy: Cathy Zoi
  - Fossil Energy: Victor Der (A)
  - Nuclear Energy: Pete Lyons (A)
  - Electricity Delivery & Energy Reliability: Pat Hoffman

Office of Science

Science to Meet the Nation’s Challenges Today and into the 21st Century

The frontiers of science

- Supporting research that led to over 100 Nobel Prizes during the past 6 decades—22 in the past decade alone
- Providing 45% of Federal support in the physical sciences
- Supporting over 27,000 Ph.D.s, graduate students, undergraduates, engineers, and support staff at more than 300 institutions

Energy and environmental science and technology

- Fostering research integration and supporting specialized facilities for collaborative studies

21st century tools of science

- Providing the world’s largest collection of scientific user facilities to over 26,000 users each year

U.S. DEPARTMENT OF ENERGY

Office of Science

ASCR Mission

The mission of the Advanced Scientific Computing Research (ASCR) program is to *discover, develop, and deploy the computational and networking capabilities* that enable researchers to analyze, model, simulate, and predict complex phenomena important to the Department of Energy.

A particular challenge of this program is *fulfilling the science potential of emerging multi-core computing systems and other novel "extreme-scale" computing architectures,* which will require significant modifications to today's tools and techniques.

U.S. DEPARTMENT OF ENERGY

Office of Science

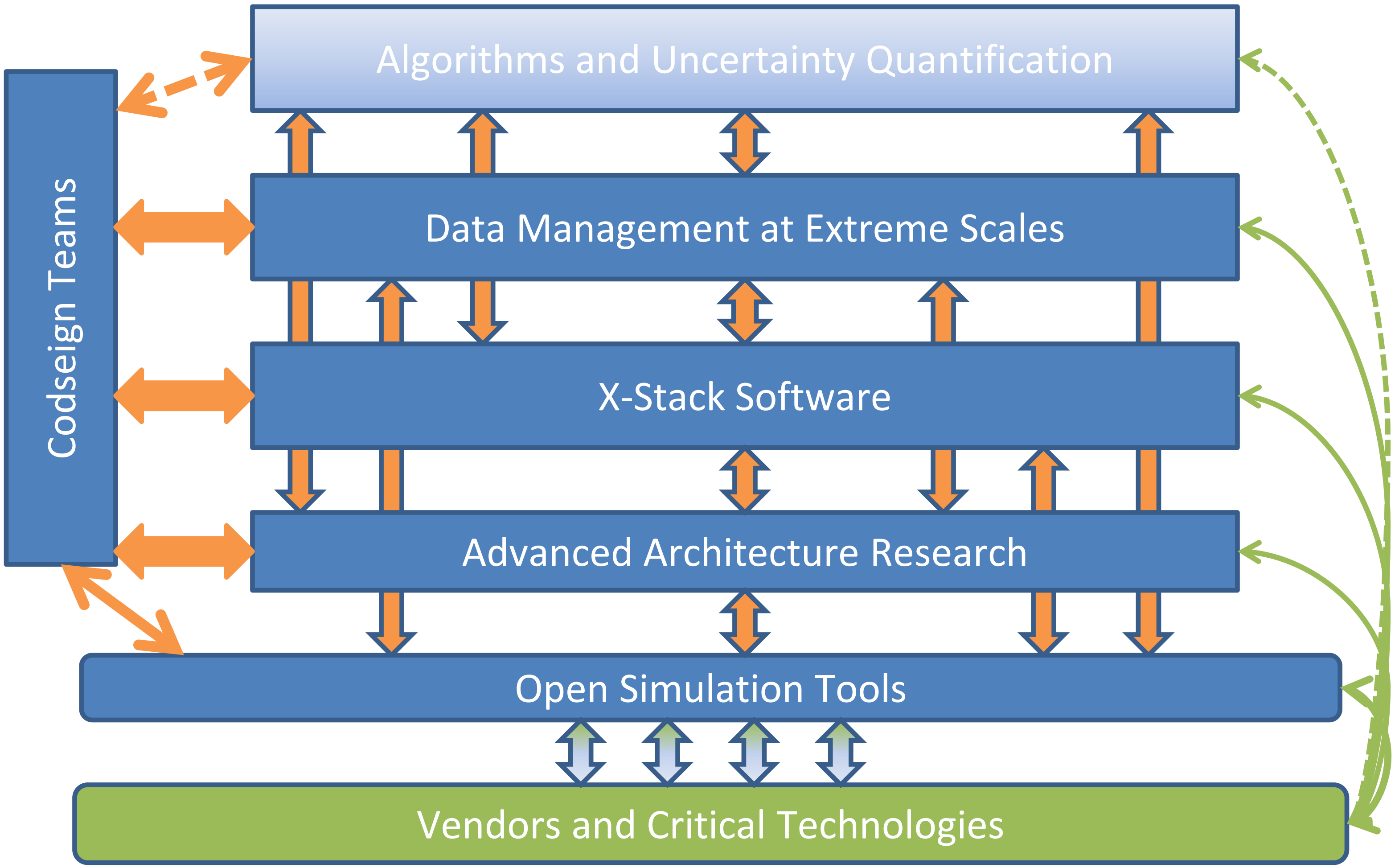
DOE Progress toward Exascale

- Proposals processed in Exascale related topic areas:
  - Applied Math: Uncertainty Quantification (90 proposals requesting ~$45M/year; 6 funded at $3M/yr)
  - Computer Science: Advanced Architectures(28 proposals requesting ~$28M/year, 6 funded at $5M/yr)
  - Computer Science: X-Stack (55 proposals requesting ~$40M/year, 11 funded at $8.5M/yr)
  - Computer Science: Scientific Data Management and Analysis (37 proposals requesting ~$22M/year; 11 projects funded at $5M/year)
  - Computational Partnerships: Co-Design (21 Proposals requesting ~ $160M/year)
- Exascale Coordination meetings with other Federal Departments and Agencies
- Formalizing Partnership with National Nuclear Security Administration (NNSA) within DOE

U.S. DEPARTMENT OF ENERGY

Office of Science

Exascale in Practice



U.S. DEPARTMENT OF ENERGY

Office of Science

Why Exascale is not just 1,000 X Petascale

- End of increases in clock speed means that all increases in performance come through concurrency (billion-way);
- Floating point operations cost less power than data movement on and off chip, therefore:
  - Remap multiphysics to put as much work per location on same chip;
  - Include embedded UQ to increase concurrency;
  - Reformulate to trade flops for memory use.
  - Weak scaling approaches (constant memory/flop) probably will not work;
- I/O to disk will be relatively slower than it is today, therefore:
  - Part of the file system may be on the node;
  - Include in-line data analysis if you can for more concurrency
  - Trigger output to only move important data off machine;
- There will be silent errors as well as lower mean time to interrupt;
  - Uncertainty Quantification including hardware variability;
  - New approaches to checkpointing and resilience.

Why Exascale is not just 1,000 X Petascale Power Challenges

- Between now and 2018 the energy required on industry roadmaps for:
  - 1 double precision flop decreases from ~100 picoJoules to ~25 picoJoules; but
  - Moving 1 bit from cpu to memory stays constant at ~30pJ/bit.
- Therefore a 1EF system with a 20MW power envelope can only have ~.1 Bytes per flop!
- To maintain ~.3 Bytes per flop need to decrease memory energy per bit ~4X from current commercial roadmap!

U.S. DEPARTMENT OF ENERGY

Office of Science

Why Exascale is not just 1,000 X Petascale
Potential System Architectures

Systems	2009	2015 +1/-0	2018 +1/-0
System peak	**2 Peta**	**100-300 Peta**	**1 Exa**
Power	**6 MW**	**~15 MW**	**~20 MW**
System memory	0.3 PB	5 PB	64 PB (+)
Node performance	125 GF	0.5 TF or 7 TF	1-2 or 10TF
Node memory BW	25 GB/s	1-2TB/s	2-4TB/s
Node concurrency	12	O(100)	O(1k) or 10k
Total Node Interconnect BW	3.5 GB/s	100-200 GB/s 10:1 vs memory bandwidth 2:1 alternative	200-400GB/s (1:4 or 1:8 from memory BW)
System size (nodes)	18,700	50,000 or 500,000	O(100,000) or O(1M)
Total concurrency	225,000	O(100,000,000) *O(10)-O(50) to hide latency	O(billion) * O(10) to O(100) for latency hiding
Storage	15 PB	150 PB	500-1000 PB (>10x system memory is min)
IO	0.2 TB	10 TB/s	60 TB/s (how long to drain the machine)
MTTI	days	O(1day)	O(1 day)

U.S. DEPARTMENT OF ENERGY

Office of Science

Aspects of Codesign

- Need to allocate complexity between hardware, systems software, libraries, and applications;
- Potential for applications to modify designs at all levels;
- Need to consider reformulating as well as re-implementing;
- Need to include uncertainty quantification, in line data analysis, and resilience in applications;
- Co-adapt applications to new programming models and perhaps languages;
- Impact of massive multithreaded nodes and new ultra-lightweight operating systems.

The PI meeting was facilitated & choreographed

- Day 1: ASCR presentations (starting at 1:30 pm)
- **Day 2 – Day 3 morning: 3 poster sessions**
  - each project representative made a 5-minute intro to the poster prior to poster session; project teams were to use poster sessions for networking
- **Day 3 afternoon: working groups to identify crosscutting topics + outbrief**
  - Each attendee was assigned to one of 8 Working Groups;
  - each WG was moderated by an NNSA facilitator, with two (ASCR PIs) reporters to present working group outbrief and a scribe (a student) to take notes
  - Day 3 dinner: ASCR & NNSA program offices identify five main topics
- **Day 4: further discussions of crosscutting topics + outbrief**
  - Morning: definition & interrelations (in same working group)
  - Afternoon: issues (in self-identified working groups)
- **Day 5: gap analysis**
  - Each attendee was assigned to a different WG
- **In addition, there were homework assignments (due before 1st working session)**
  - Exascale Research Coverage Assessment
  - Dance cards
- Results are being tabulated & analyzed

Exascale Research Coverage Assessment

- Complete throughout the week (press firmly)
- Turn in top copy by end of lunch on Wednesday

Exascale Research Coverage Assessment Name: ____________________

Issues	Importance (Hi/Med/Lo)	Coverage in ASCR FOAs (Yes / No)	Coverage in ASCR Exascale Projects (Adequate/Inadequate)	Coverage by Other Agencies' Projects (Agency & Project Title)
Hardware Architecture and I/O				
Fault detection and resilience				
File systems and I/O				
Memory systems				
Power management				
Simulators				
Synchronization				
Testbeds				
Other:				

U.S. DEPARTMENT OF ENERGY

Office of Science

Research Collaboration "Dance Cards"

1. ASCR Researchers
   - 1 per program affiliation (extras at registration desk)
   - Use for Wednesday lunch

2. ASCR Project Lead PIs
   - Project Summary version (complete during Wednesday lunch with project team)
   - Turn in top copy on Thursday morning (press firmly)

3. Government/Other Invitees
   - Turn in top copy on Thursday morning (press firmly)

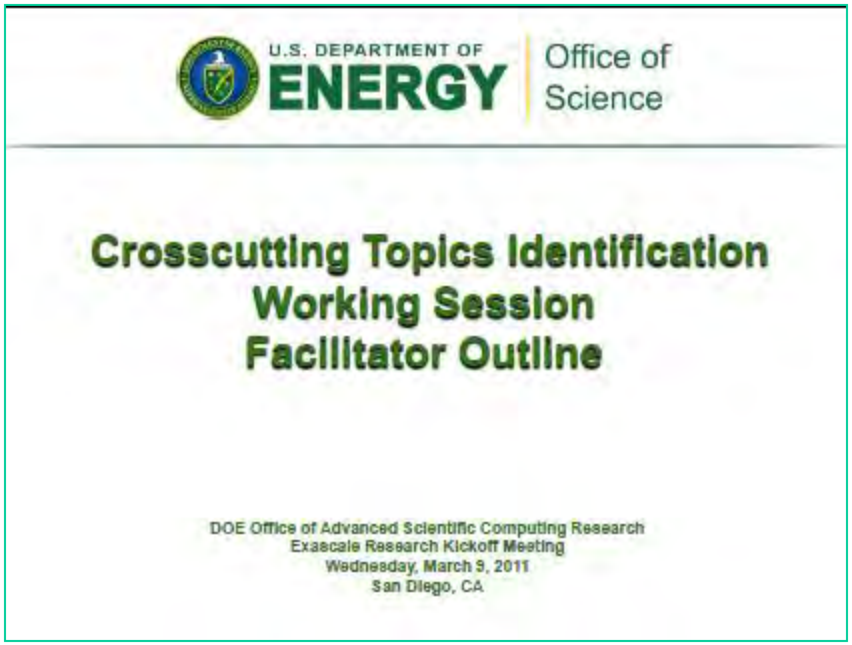
ASCR Project / Lead PI	Poster Session	Priority (Hi/Med/Lo)	Likely Nature of Collaboration and Timeframe
Advanced Architectures and Critical Technologies for Exascale Computing			
Blackcomb (Vetter)	2 Tues pm		
CODES (Ross)	3 Wed am		
CoDEx (Shalf)	1 Tues am		
DMD (Rodrigues)	3 Wed am		
NoLoSS (Ross)	2 Tues pm		
Thrifty (Torrellas)	1 Tues am		

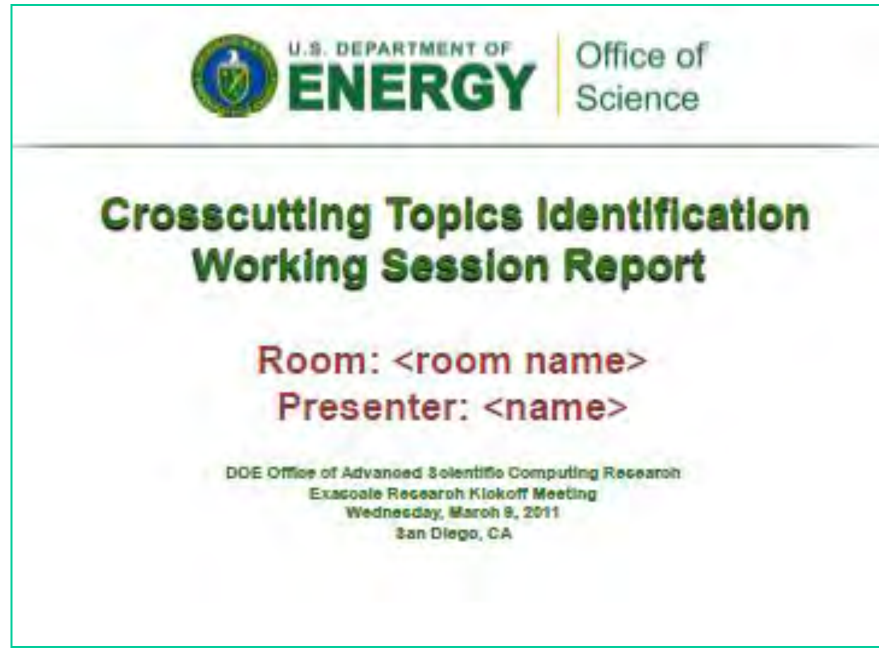
ASCR Project / Lead PI	Poster Session	Priority (Hi/Med/Lo)	Suggested Collaborations (identify agency, program/project title and contact if not ASCR funded)
Advanced Architectures and Critical Technologies for Exascale Computing			
Blackcomb (Vetter)	2 Tues pm		
CODES (Ross)	3 Wed am		
CoDEx (Shalf)	1 Tues am		
DMD (Rodrigues)	3 Wed am		
NoLoSS (Ross)	2 Tues pm		
Thrifty (Torrellas)	1 Tues am		

U.S. DEPARTMENT OF ENERGY

Office of Science

Identification of crosscutting topics & outbrief





In each working group the facilitator, reporters, and the scribe are given instructions on Day 1. For each working session, there is a template to facilitate and capture the discussions, and another for the outbrief.

U.S. DEPARTMENT OF ENERGY

Office of Science

Identification of crosscutting topics

Seed Questions

1. What topics are crosscutting, foundational, or important across programs and/or projects to collectively address the challenges of exascale?

2. What topics exist in the "white space" between programs and projects that must be addressed in order to support science at the exascale?
   - Topics that are independent of individual FOAs and projects

For the identification of crosscutting topics, each working group is given these seed questions (developed by ASCR).

A responsible person writes the topics on flip charts. The participants in the room get to decide the level of granularity and wording of these topics

Participants vote on topics

Voting Instructions

- **Each person gets 6 colored dots for voting**
  - 3 blue
  - 3 red
- **Use the dots to vote for your top three in each category**
  - Blue dot = what is **most important** to your project(s) over the next year
  - Red dot = what you **worry about** the most
- **You cannot put more than one dot of the same color on an item**
- **You do not have to use all your dots**

Exascale Research Kickoff Meeting March 9, 2011

Crosscutting Topics Identified – Top Five

Topic	Most Important	Worry About Most	Total Votes
			0
			0
			0
			0
			0

Exascale Research Kickoff Meeting March 9, 2011

d – Next Five

	Most Important	Worry About Most	Total Votes
			0
			0
			0
			0
			0

Exascale Research Kickoff Meeting March 9, 2011

Each participant gets 3 blue and 3 red sticky dots to put next to flip charts.

Votes are tallied (red + blue) and entered into the outbrief presentation

Participants get to wax eloquent

Observations

1. What observations do you have in general about the priorities?
2. Were there any surprises regarding the items with the highest numbers of votes?
3. Were there any significant differences between the placement of the different colored dots – i.e., mostly one color vs. a combination of both colors?
4. Why do participants worry about the topics that drew the most red dots?

Two common observations by participants:

- blue = what I'm working on (therefore important);
- red = what nobody is working on (therefore worrisome)

Five crosscutting topics are identified based on researcher input

1. Programming Models
2. Resilience and Fault Tolerance
3. Data Management & I/O Architecture
4. Memory / Data Movement and Energy Efficiency
5. Application Architecture

Participants are then asked to define these crosscutting topics

Defining Crosscutting Topics Working Session

- **Crosscutting topic priorities identified**
  1. Programming Models
  2. Resilience and Fault Tolerance
  3. Data Management & I/O Architecture
  4. Memory / Data Movement and Energy Efficiency
  5. Application Architecture

Timing guidance:
•~25 minutes per topic
•15-minute break after 2nd topic (~10:00 am)
•Remainder for pair-wise question

- **For each crosscutting topic above:**
  - Define the problem space
  - Characterize the major challenges
- **Identify which two pair-wise interdependencies among these crosscutting topics you believe to be most important, and why**
  - Explain your answers in a sentence or two

Same room assignments as on Wednesday for everyone

After an outbrief, the participants work in self-selected groups for further discussions

- Only first 4 topics are discussed (to force codesign application scientists to spread out)
- No working group assignments were made (thus those who wish to only participate in discussions of their areas of specialty may do so)
- Five questions are used to guide the discussions (next slide)

U.S. DEPARTMENT OF ENERGY

Office of Science

Discussion Questions

1. Dependencies, Interactions, and Bottlenecks
   - A. What **functional dependencies** have you discovered?
     - What are many people depending on that might require further resources in order to prevent bottlenecks?
   - B. What **new interactions** are needed (e.g., hardware/software, people)?
   - C. (Optional) What have you learned that you may need to **provide to other people**?
2. Emerging Needs
   - A. Based on what you have learned so far this week, what has **caught you by surprise**?
   - B. What **new requirements** for hardware, software, or both have you identified?
3. What aspects of this topic do you **need to understand** more completely in order for your project to be a success?
4. Have you identified any solutions as a result of this working session? If so, what are they?
5. Path Forward
   - A. Does this **working group** need to continue; if so, how?
   - B. What else should ASCR do to help **facilitate linkages** and **address outstanding questions**?

U.S. DEPARTMENT OF ENERGY

Office of Science

The participants work on gap analysis on the last day

Discussion Questions

1. **Gap Analysis**
   A. What did you expect to hear that wasn't covered in any of this week's discussions?
   B. What computer science challenges for exascale are not being addressed (adequately) in efforts currently being funded?
   C. What are the major risks you see?

2. **Schedule and Timelines**
   A. Should a shared timeline be established that defines what research teams plan to accomplish at specific times so that people adopting the technology can plan?
      - If so, when and how?
   B. Given scope/schedule/resources tradeoffs, where is there slack? What can be deferred?

3. **What should ASCR do to help these diverse projects build towards an integrated Exascale Initiative?**

Separately, the Research Assessment form has been tallied and the result is being analyzed.

It was a good start

ASCR POC: Lucy Nowell

ASCR CS Program Managers: Lucy Nowell, Sonia Sachs

Event Facilitator & Organizer: Lee Ann Dudney, PNNL, assisted by Kris Cook, PNNL

Administrative Support: Trina Pitcher & Kelly Sandretto

Meeting material at
http://exascaleresearch.labworks.org/ascr2011

U.S. DEPARTMENT OF ENERGY

Office of Science

Backup

Seven Co-Design proposal teams are here this week

TITLE	LEAD PI
Center for Exascale Simulation of Advanced Reactors (CESAR)	Rosner (ANL)
FLASH High Energy Density Physics Exascale Codesign Center	Lamb (ANL)
The CERF Center: Co-design for Exascale Research in Fusion	Koniges (LBNL)
Exascale Co-Design Center for Materials in Extreme Environments: Engineering-Scale Predictions	Germann (LANL)
Chemistry Exascale Co-Design Center	Harrison (ORNL)
Combustion Exascale Co-Design Center	Chen (SNL)
Exascale Performance Research for Earth System Simulation (EXPRESS)	Jones (LANL)

(Exa)Scale Changes Everything

	2010	2018	Factor Change
System peak	2 Pf/s	1 Ef/s	500
Power	6 MW	20 MW	3
System Memory	0.3 PB	10 PB	33
Node Performance	0.125 Gf/s	10 Tf/s	80
Node Memory BW	25 GB/s	400 GB/s	16
Node Concurrency	12 cpus	1,000 cpus	83
Interconnect BW	1.5 GB/s	50 GB/s	33
System Size (nodes)	20 K nodes	1 M nodes	50
Total Concurrency	225 K	1 B	4,444
Storage	15 PB	300 PB	20
Input/Output bandwidth	0.2 TB/s	20 TB/s	100

DOE Exascale Initiative Roadmap, Architecture and Technology Workshop, San Diego, December, 2009.

FY10 Computer Science Funding Opportunity Announcements (FOAs)

- **Advanced Architectures and Critical Technologies for Exascale Computing -- ~$5M / year for 3 years to fund 5 awards, Closed March 26, 2010**
  - Design of energy-efficient, resilient hardware and software architectures and technology for high performance computing systems at Exascale.

- **X-Stack Software Research -- ~$8.5M / year for 3 years to fund 11 awards, Closed April 2, 2010**
  - Development of a scientific software stack that supports extreme scale scientific computing, from operating systems to development environments.

- **Scientific Data Management and Analysis at the Extreme Scale -- ~$5M / year for 3 years for 10 awards, Closed March 18, 2010**
  - Management and analysis of extreme-scale scientific data in the context of Petascale computers and/or Exascale computers with heterogeneous multi-core architectures.

Advanced Architectures Topics of Interest

- Approaches for reducing and/or managing power requirements for high performance computing systems, including the memory and storage hierarchy
- Approaches for reducing and/or managing heat in high performance computing systems
- Methods for improving system resilience and managing the component failure rate, including approaches for shared information and responsibility among the OS, runtime system, and applications
- Co-design of systems that support advanced computational science at the extreme scale
- Scalable I/O systems, which may include alternatives to file systems
- Approaches to information hiding that reduce the need for users to be aware of system complexity, including heterogeneous cores, the memory hierarchy, etc.

Advanced Architectures Awards

(28 proposals requesting ~$28M/yr, 6 projects funded at ~$5M/yr)

- **Blackcomb: Hardware-Software Co-design for Non-Volatile Memory in Exascale Systems**
  - Jeffrey S. Vetter (Oak Ridge National Laboratory), Robert Schreiber (Hewlett Packard Labs), Trevor Mudge (University of Michigan), Yuan Xie (Penn State University)
  - http://exascaleresearch.labworks.org/uploads/dataforms/A_OPH_ORNL_Blackcomb_110217.pdf

- **Enabling Co-Design of Multi-Layer Exascale Storage Architectures**
  - Robert B. Ross, William Allcock, and Samuel Lang (Argonne National Laboratory), and Christopher D. Carothers (Rensselaer Polytechnic Institute)
  - http://exascaleresearch.labworks.org/uploads/dataforms/A_OPH_ANL_CODES_110215.pdf

- **NoLoSS: Enabling Exascale Science through Node Local Storage Systems**
  - Robert B. Ross , Darius Buntinas, and Kamil Iskra (Argonne National Laboratory), Bronis R. de Supinski and Maya B. Gokhale (Lawrence Livermore National Laboratory)
  - http://exascaleresearch.labworks.org/uploads/dataforms/A_OPH_ANL_NoLoSS_110214.pdf

Advanced Architectures Awards

(28 proposals requesting ~$28M/yr, 6 projects funded at ~$5M/yr)

- **CoDEx: A Hardware/Software Co-Design Environment for the Exascale Era**
  - John Shalf, David Donofrio, Alice Koniges, Leonid Oliker, Samuel Williams (Lawrence Berkeley National Laboratory), Curtis Janssen, Helgi Adalsteinsson, Damian Dechev, Ali Pinar (Sandia National Laboratories, Dan Quinlan, Chunhua Liao, Thomas Panas (Lawrence Livermore National Laboratory), Sudhakar Yalamanchili (Georgia Institute of Technology)
  - http://exascaleresearch.labworks.org/uploads/dataforms/A_OPH_LBNL_CoDEx_110214.pdf
- **Data Movement Dominates: Advanced Memory Technology to Address the Real Exascale Power Problem**
  - Arun Rodrigues, Richard C. Murphy , Brian Barrett, and Scott Hemmert (Sandia National Laboratories), Paul Hargrove and John Shalf (Lawrence Berkeley National Laboratory), David Resnick (Micron Technologies), Keren Bergman (Columbia University), and Bruce Jacob (University of Maryland)
  - http://exascaleresearch.labworks.org/uploads/dataforms/A_OPH_SNL_DMD_110214.pdf
- **Thrifty: An Exascale Architecture for Energy-Proportional Computing**
  - Josep Torrellas (UIUC), Daniel Quinlan (LLNL), Allan Snavely (UCSD), and Wilfred Pinfold (Intel)
  - http://exascaleresearch.labworks.org/uploads/dataforms/A_OPH_UIUC_Thrifty_110215.pdf

X- Stack Topics of Interest

- **System software, including operating systems, runtime systems, adaptable operating and runtime systems, I/O systems, system management/administration, resource management and means of exposing resources, and external environments;**
- **Fault management, both by the operating and runtime systems and by applications;**
- **Development environments, including programming models, frameworks, compiler, & debugging tools;**
- **Application frameworks;**
- **Cross-cutting dimensions, including resilience, power management, performance optimization, and programmability; and**
- **Design and/or development of high-performance scientific workflow systems that incorporate data management and analysis capabilities**

X-Stack Awards

(55 proposals requesting ~$40M/year, 11 projects funded at ~8.5M/year)

- **A Fault-oblivious Extreme-scale Execution Environment**
  - Ronald G Minnich, Curtis L. Janssen (Sandia National Laboratories), Sriram Krishnamoorthy and Andres Marquez (Pacific Northwest National Laboratory), Maya Gokhale (Lawrence Livermore National Laboratory), P. Sadayappan (Ohio State University), Eric Van Hensbergen (IBM Research), Jim McKie (Alcatel-Lucent Bell Labs), and Jonathan Appavoo (Boston University)
  - http://exascaleresearch.labworks.org/uploads/dataforms/X_OPH_SNL_FOX_110214.pdf

- **Auto-tuning for Performance and Productivity on Extreme-scale Computations**
  - Samuel W. Williams and Leonid Oliker (Lawrence Berkeley National Laboratory), John Gilbert (University of California, Santa Barbara), Stephane Ethier (Princeton Plasma Physics Laboratory)
  - http://exascaleresearch.labworks.org/uploads/dataforms/X_OPH_LBNL_Auto-tuning_110214.pdf

- **Software Synthesis for High Productivity ExaScale Computing**
  - Armando Solar-Lezama (MIT) and Rastislav Bodik and James Demmel (UC Berkeley)
  - http://exascaleresearch.labworks.org/uploads/dataforms/X_OPH_MIT_Synthesis_110225.pdf

X-Stack Awards

(55 proposals requesting ~$40M/year, 11 projects funded at ~8.5M/year)

- **ExM: System support for extreme-scale, many-task applications**
  - Michael Wilde, Ewing Lusk, and Ian Foster (Argonne National Laboratory), Daniel S. Katz (University of Chicago), and Matei Ripeanu (University of British Columbia)
  - http://exascaleresearch.labworks.org/uploads/dataforms/X_OPH_ANL_ExM_110303.pdf
- **COMPOSE-HPC: Software Composition for Extreme Scale Computational Science and Engineering**
  - David E. Bernholdt (Oak Ridge National Laboratory), Matt Sottile (Galois, Inc.), Tom Epperly (Livermore National Laboratory), Manojkumar Krishnan (Pacific Northwest National Laboratory), and Rob Armstrong (Sandia National Laboratories)
  - http://exascaleresearch.labworks.org/uploads/dataforms/X_OPH_ORNL_COMPOSE_110217.pdf
- **Damsel: A Data Model Storage Library for Exascale Science**
  - Alok Choudhary and Wei-keng Liao (Northwestern University), Robert Latham, Robert Ross, and Timothy Tautges (Argonne National Laboratory), Nagiza Samatova (North Carolina State University and Oak Ridge National Laboratory), and Quincey Koziol (The HDF Group)
  - http://exascaleresearch.labworks.org/uploads/dataforms/X_OPH_NU_Damsel_110221.pdf

X-Stack Awards

(55 proposals requesting ~$40M/year, 11 projects funded at ~8.5M/year)

- **Vancouver: Designing a Next-Generation Software Infrastructure for Productive Heterogeneous Exascale Computing**
  - Jeffrey S. Vetter (Oak Ridge National Laboratory), Wen-mei Hwu (Unviersity of Illinois at Urbana Champaign), Allen D. Malony (University of Oregon), and Rich Vuduc, (Georgia Institute of Technology)
  - http://exascaleresearch.labworks.org/uploads/dataforms/X_OPH_ORNL_Vancouver_110217.pdf
- **Enabling Exascale Hardware and Software Design through Scalable System Virtualization**
  - Patrick G. Bridges (University of New Mexico), Peter Dinda (Northwestern University), Kevin Pedretti (Sandia National Laboratories), and Stephen Scott (Oak Ridge National Laboratory)
  - http://exascaleresearch.labworks.org/uploads/dataforms/X_OPH_UNM_Virtualization_110218.pdf

X-Stack Awards

(55 proposals requesting ~$40M/year, 11 projects funded at ~8.5M/year)

- **ZettaBricks: A Language, Compiler and Runtime Environment for Anyscale Computing**
  - Saman Amarasinghe, Alan Edelman, Una-May O'Reilly, Martin Rinard, and Chris Hill (MIT)
  - http://exascaleresearch.labworks.org/uploads/dataforms/X_OPH_MIT_ZettaBricks_110222.pdf
- **Compiled MPI: Cost-Effective Exascale Application Development**
  - Grigory Bronevetsky and Daniel Quinlan (Lawrence Livermore National Lab), Andrew Lumsdaine (Indiana University) and Torsten Hoefler (University of Illinois at Urbana-Campaign)
  - http://exascaleresearch.labworks.org/uploads/dataforms/X_OPH_UIUC_CoMPI_110221.pdf
- **An Open Integrated Software Stack for Extreme Scale Computing**
  - Pete Beckman (Argonne National Laboratories) in collaboration with several PIs at UTK, UCB, UIUC, University of Houston, LBNL, Rice University, University of Oregon, and Sandia National Laboratories
  - http://exascaleresearch.labworks.org/uploads/dataforms/X_OPH_ANL_ESC_110304.pdf

Data Management and Analysis at Extreme Scale Topics of Interest

- **Extreme-scale data storage and access systems** that are functional at the granularity and massive scale of scientific endeavors, that support data models that are consistent with scientists' view of their data, and that minimize the need for scientists to have detailed knowledge of system hardware and operating systems;
- **Scalable data triage, summarization, and analysis** methods and tools, including adaptive, power-aware algorithms and software for in-situ data reduction and/or analysis of massive multivariate data sets;
- **Semantic integration of heterogeneous scientific data**, including support for integrated analysis of multi-scale, multi-physics data from multiple types of devices and experiments, and/or simulations and spanning multiple scientific disciplines;
- **Knowledge representation and automated machine reasoning and/or data mining** methods and tools that support automated analysis and integration of large scientific data sets, especially those that include tensor flow fields, and/or representation of and reasoning about uncertainty and sources of uncertainty; and
- **Visual analysis of extreme-scale scientific data**, including multi-user visual analysis methods and tools for scientist users who may or may not be co-located, and interactive visual steering of computational processes.

Data Mgmt & Analysis at Extreme Scale Awards

(37 proposals requesting ~$22/yr, 10 projects funded at ~5M)

- **Dynamic Non-Hierarchical File Systems for Exascale Storage**
  - Darrell Long and Ethan Miller (University of California, Santa Cruz)
  - http://exascaleresearch.labworks.org/uploads/dataforms/D_OPH_UCSC_NHFS_110214.pdf
- **Runtime System for I/O Staging in Support of In-Situ Processing of Extreme Scale Data**
  - Scott Klasky, Nagiza Samatova, Matthew Wolf, and Norbert Podhorszki (Oak Ridge National Laboratory), Arie Shoshani and John Wu (Lawrence Berkeley National Laboratory), and Karsten Schwan and Greg Eisenhauer (George Tech)
  - http://exascaleresearch.labworks.org/uploads/dataforms/D_OPH_ORNL_SciDEx_110214.docx
- **ExaHDF5: An I/O Platform for Exascale Data Models, Analysis and Performance**
  - Prabhat (Lawrence Berkeley National Lab), Karen Schuchardt (Pacific Northwest National Lab) and the HDF Group
  - http://exascaleresearch.labworks.org/uploads/dataforms/D_OPH_LBNL_ExaHDF5_110215.docx
- **Adding Data Management Services to Parallel File Systems**
  - Scott A. Brandt (University of California, Santa Cruz) and Maya Gokhale (Lawrence Livermore National Laboratory)
  - http://exascaleresearch.labworks.org/uploads/dataforms/D_OPH_UCSC_Damasc_110215.pdf

Data Mgmt & Analysis at Extreme Scale Awards

(37 proposals requesting ~$22/yr, 10 projects funded at ~5M)

- **Scalable and Power Efficient Data Analytics for Hybrid Exascale Systems**
  - Alok Choudhary and Wei-keng Liao (Northwestern University) , Nagiza Samatova (Oakridge National Laboratory), and Kesheng (John) Wu (Lawrence Berkley National Laboratory)
  - http://exascaleresearch.labworks.org/uploads/dataforms/D_OPH_NU_Xanalytics_110211.pdf
- **A Pervasive Parallel Processing Framework for Data Visualization & Analysis at Extreme Scale**
  - Kenneth Moreland (Sandia National Laboratories), Kwan-Liu Ma (University of California, Davis) and Berk Geveci (Kitware, Inc.)
  - http://exascaleresearch.labworks.org/uploads/dataforms/D_OPH_SNL_DAX_110214.pdf
- **An Information-Theoretic Framework for Enabling Extreme-Scale Science Discovery**
  - Han-Wei Shen (Ohio State University), Robert Ross and Thomas Peterka (Argonne National Laboratory) and Yi-Jen Chiang (Polytecnic Institute of New York University)
  - http://exascaleresearch.labworks.org/uploads/dataforms/D_OPH_OSU_InfoTheory_110211.pdf

Data Mgmt & Analysis at Extreme Scale Awards

(37 proposals requesting ~$22/yr, 10 projects funded at ~5M)

- **Enabling Scientific Discovery in Exascale Simulations**
  - Dr. Chandrika Kamath (Lawrence Livermore National Laboratory) and Prof. George Karypis (University of Minnesota)
  - http://exascaleresearch.labworks.org/uploads/dataforms/D_OPH_LLNL_Exa-DM_110214.pdf
- **Graph-Based 3D Flow Field Visual Analysis**
  - Pak Chung Wong (Pacific Northwest National Laboartory) , Han-Wei Shen (Ohio State University) and Daniel Chavarría (Pacific Northwest National Laboratory)
  - http://exascaleresearch.labworks.org/uploads/dataforms/D_OPH_PNNL_FlowGraphAnalysis_110215.pdf
- **Topology-Based Visualization & Analysis of Multidimensional Data & Time-Varying Data at the Extreme Scale**
  - Gunther Weber (Lawrence Berkeley National Laboratory)
  - http://exascaleresearch.labworks.org/uploads/dataforms/D_OPH_LBNL_ExScaTopoAna_110214.pdf