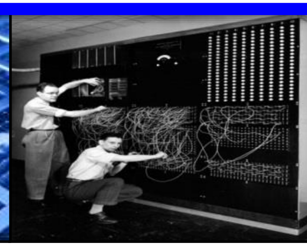
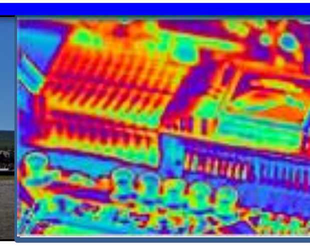
U.S. DEPARTMENT OF ENERGY | Office of Science

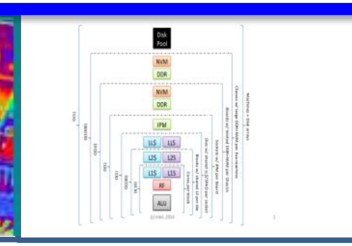

Parallelism


Data Movement


Programmability

Resiliency


Energy Efficiency


Deep Memory Hierarchies

April 2016 X-Stack PI Meeting Update

Vivek Sarkar
Rice University
September 21, 2016

Outline

- <u>Background on X-Stack Program</u>
- X-Stack PI meeting, April 6-7, 2016
- Conclusions

U.S. DEPARTMENT OF ENERGY | Office of Science

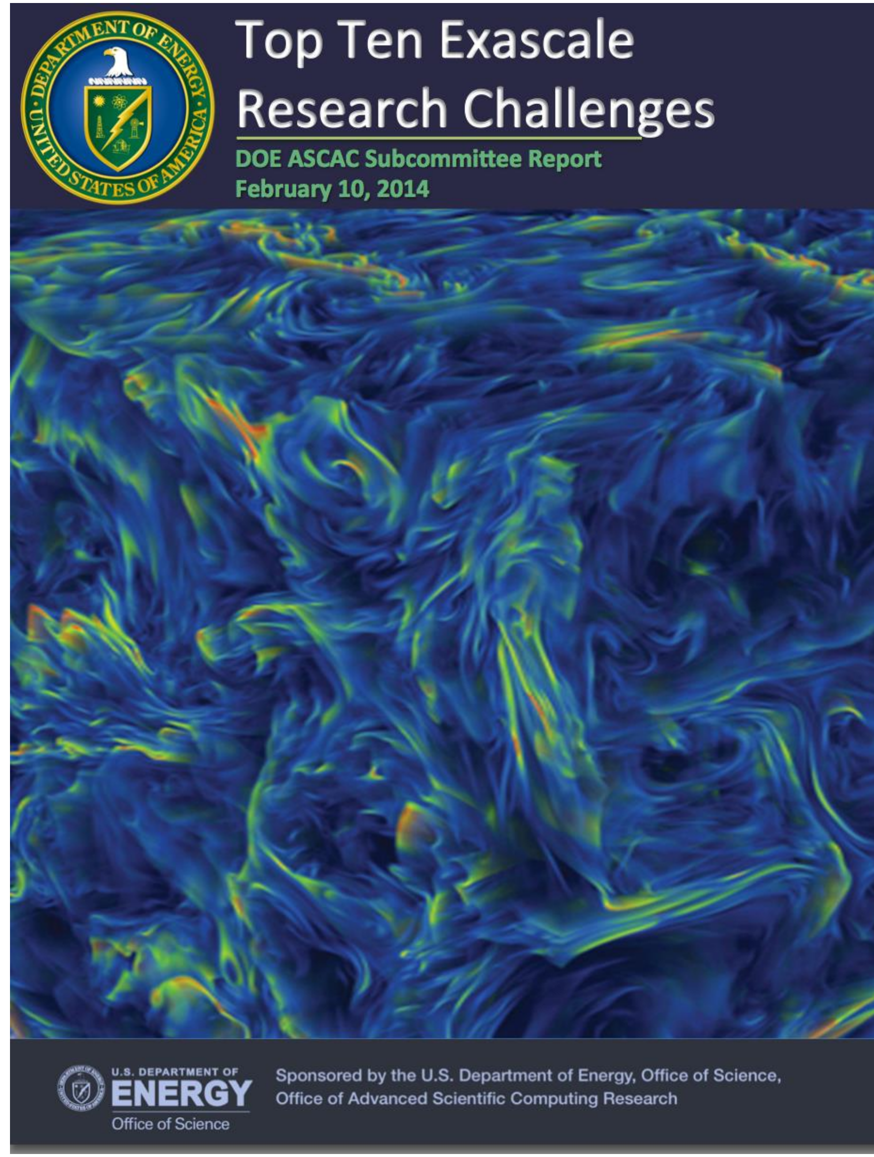
X-Stack Program & Exascale Research Challenges

1. *Energy efficiency*
2. *Interconnect technology*
3. *Memory Technology*
4. ***Scalable System Software***
5. ***Programming systems***
6. Data management
7. Exascale Algorithms
8. Algorithms for discovery, design, and decision
9. *Resilience and correctness*
10. *Scientific productivity*

Red: Central focus

Blue: Significant focus

Green: Coordinated efforts

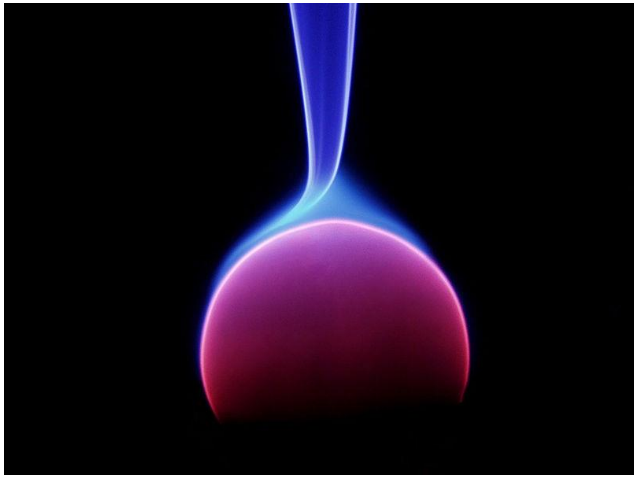


X-Stack Program

- Program Manager: Sonia R. Sachs
- DOE ASCR Research Director: William Harrod
- Program wiki: https://xstackwiki.modelado.org/Extreme_Scale_Software_Stack
- Background
  - Community input collected in two workshops held in 2011 prior to FOA
  - FOA topics: Programming models, languages, compilers, runtime systems, and tools that address identified Exascale challenges (FOA deadline: February 6, 2012)
  - 67 full proposals: 24 Lab proposals, 36 university proposals, 7 industry proposals
- Portfolio
  - Target funding for program: $15M/year for FY12 – FY15, $12M/year for FY16
  - 9 projects were funded for 3 years, September, 2012 - September 2015
    - 7 projects were extended for one additional year until September 2016
  - 2 more projects added for September 2013 – September 2016
  - 1 more project added for June 2014 – September 2016

Goals of X-Stack Research Program

- Programming Systems: new approaches to managing parallelism, locality and data movement through innovations in programming interfaces
- Runtime Systems: new approaches to performance portability, and adaptation to changing application goals and system conditions
- Performance and Debugging tools: enabling end user productivity and performance
- Interoperability: facilitate interoperability across different HPC languages and interfaces, as well as cross-cutting execution models



10 current projects in X-Stack Portfolio

DEGAS (PI: Kathy Yelick)

Hierarchical and resilient programming models, compilers and runtime support.

Traleika Glacier (PI: Shekhar Borkar)

Exascale programming system, execution model and runtime, applications, and architecture explorations, with open and shared simulation infrastructure.

D-TEC (PI: Dan Quinlan and Saman Amarasinghe)

Complete software stack solution, from DSLs to optimized runtime systems code.

XPRESS (PI: Ron Brightwell)

Software architecture and interfaces that exploit the ParalleX execution model, prototyping several of its key components.

X-Tune (PI: Mary Hall)

Unified autotuning framework that integrates programmer-directed and compiler-directed autotuning.

CORVETTE (PI: Koushik Sen)

Automated bug finding methods to make concurrency bugs and floating point behavior reproducible.

SLEEC (PI: Milind Kulkarni)

Semantics-aware, extensible optimizing compiler technologies.

PIPER (PI: Martin Schultz)

Tools for debugging and analysis of performance, power, and energy.

Vancouver 2 (PI: Jeffrey Vetter)

Heterogeneous computing: abstractions, runtime systems, tools and low-level libraries

ARES (Pat McCormick and Jeffrey Vetter)

High-level intermediate representations (HLIR) and optimizations to select architectures

U.S. DEPARTMENT OF ENERGY | Office of Science

X-Stack Workshops & Summits

- ASCR Exascale Programming Challenges Workshop (Marina del Rey, July 2011)
  - http://science.energy.gov/ascr/research/computer-science/programming-challenges-workshop/
- ASCR Exascale Tools Workshop (Annapolis, October 2011)
  - http://science.energy.gov/ascr/research/computer-
- Exascale Research Conference (Anna
  - Included other ASCR exascale research
  - http://exascaleresearch.labworks.org/oc
- Programming Models & Environments 2014)
  - https://xstackwiki.modelado.org/0 - _Programming_Models_%26_Environments_Summit
- Runtime System Summit ashington D.C., April 2014)
  - https://xstackw modelado.org/10.17.2014_Runtime_Systems_Summit
- 2015 Programming Models and Environments Workshop (Rockville, March 2015)
  - http://www.orau.gov/programming2015/
- 2015 Runtime Systems Workshop (Rockville, March 2015)
  - http://www.orau.gov/runtimesys2015/

- NOTE: above web sites also include links to published workshops & summit reports

Presented by Kathy Yelick in July 2015 ASCAC meeting

X-Stack PI Meetings

- X-Stack Kickoff PI meeting:
  - https://xstackwiki.modelado.org/9/17/12:_X-Stack_Portfolio_Kickoff_Meeting
- 2013 X-Stack PI meeting:
  - https://xstackwiki.modelado.org/3/20/13_-_X-Stack_PI_Meeting
- 2014 X-Stack PI meeting:
  - https://xstackwiki.modelado.org/05/[illegible]Stack_PI_Meeting
- 2015 X-Stack PI meeting (combined with OS[illegible] meeting):
  - https://xstackwiki.modelado.org/[illegible]ber_7-8,_2015_X-Stack_and_OS/R_[illegible]meeting
- 2016 X-Stack PI meeting:
  - https://xstackwiki.modelado.org/April_6-7,_2016_-_X-Stack_PI_Meeting

Focus of today's presentation

Marriott North Bethesda, MD

Outline

- Background on X-Stack Program
- <u>X-Stack PI meeting, April 6-7, 2016</u>
- Conclusions

Meeting summary

- Dates: April 6-7, 2016
- Location: LBNL
- Organizing Committee:
  - Vivek Sarkar (Chair)
  - Saman Amarasinghe
  - Dan Quinlan
  - Kathy Yelick
- ~ 109 participants
  - 36 lab participants, 36 industry participants, 37 university participants
- Special thanks to LBNL, NERSC, ASCR, ORISE staff for help and support with all meeting logistics
- Meeting web site:
  - https://xstackwiki.modelado.org/April_6-7,_2016_-_X-Stack_PI_Meeting

NERSC News about X-Stack Software Demonstrations

NERSC

Powering Scientific Discovery Since 1974

Berkeley Lab Hosts Fourth X-Stack PI Meeting

Application code demos utilized more than 1,000 nodes on NERSC supercomputers

APRIL 25, 2016
Contact: Kathy Kincade, kkincade@lbl.gov, +1 510 495 2124

A Technology Marketplace held A 6 during the X-Stack at Berkeley Lab meeting gave application developers the opportunity to demo the software prototypes

Berkeley Lab hosted the fourth annual X-Stack PI meeting April 6 and 7, where X-Stack researchers, the facilities teams, application scientists and developers from national laboratories, universities and industry met to share the latest developments in X-Stack application codes and identify further modifications.

X-Stack was launched in 2012 by the U.S. Department of Energy's Advanced Scientific Computing Research program to support the development of exascale software tools, including programming languages and libraries, compilers and runtime systems, that will help programmers handle massive parallelism, data movement, heterogeneity and failures as the scientific community transitions to the next generation of extreme-scale supercomputers. A total of nine X-Stack programs were designated to develop complete solutions that address multiple components of the system software stack: DEGAS, D-TEC, XPRESS, Traleika, DynAX, XTUNE, GVR, CORVETTE and SLEEC.

During the first three years of the program, these projects have completed research and development of programming models, programming environments and runtime systems for exascale. During the fourth year, which began in September 2015, the development teams are extending their results and developing additional benefits for the application codes.

The goal for this year's X-Stack PI meeting was to demonstrate the latest advances in the codes, with an eye toward delivery in the latter part of 2016. Toward this end, a Technology Marketplace held during the April meeting gave developers the opportunity to demo the software prototypes; a total of 20 demonstrations were given during the two-hour marketplace event, with 15 individual teams enabled by NERSC to show emerging exascale technologies in the development phase.

OLCF News about X-Stack Software Demonstrations

TECHNOLOGY - Written by OLCF Staff Writer on May 31, 2016

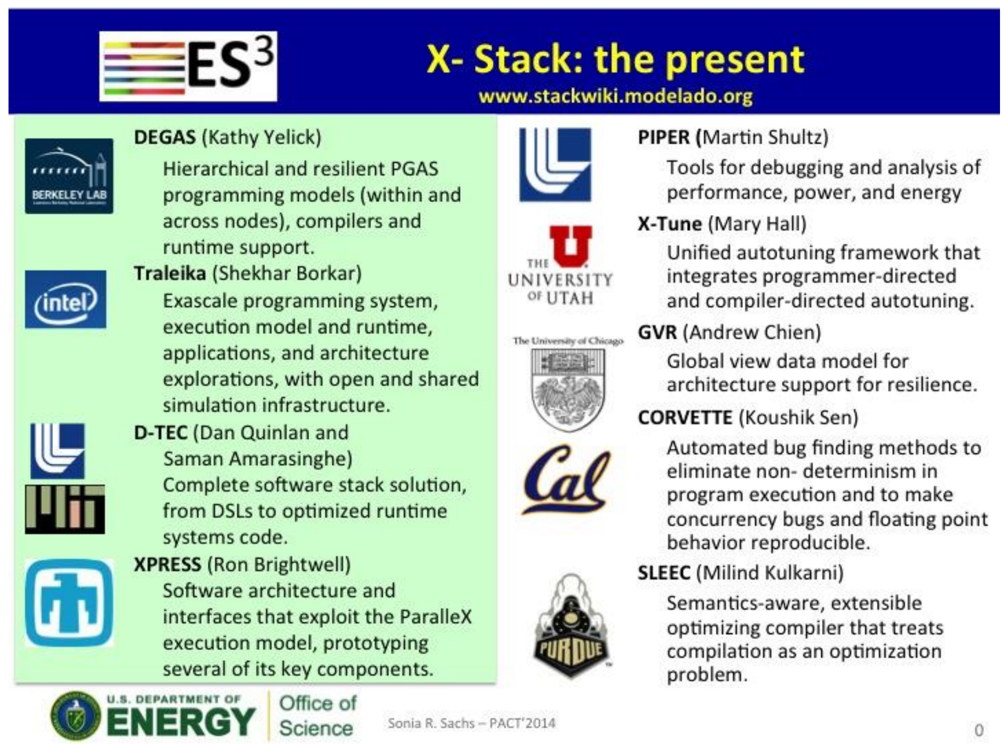
X–Stack Projects Use Titan’s GPUs for Demos

Tags: DSL, GPUs, Titan, X-Stack

SHARE

Project teams demonstrate their research in programming environments for future exascale systems

One of the unique aspects of the Oak Ridge Leadership Computing Facility’s (OLCF’s) Titan supercomputer is its hybrid architecture, consisting of both CPUs and GPUs. Recently, those GPUs played a big role in helping advance US Department of Energy (DOE)-funded scientific discovery.



U.S. DEPARTMENT OF ENERGY | Office of Science

Meeting Agenda: Day 1 (April 6, 2016)

Day 1 Goals

- Review X-Stack Portfolio – share latest research results in presentations and demonstrations
- Understand successes and challenges in adoption of new software technologies in HPC Facilities

Panel (1 hour)

- "Successes & challenges in adopting new software technologies in HPC facilities", panel discussion with Facility representatives

Technology Marketplace (2 hours)

- 18 X-Stack Project Technology Demonstrations

Project presentations (5.5 hours)

- DEGAS (Yili Zheng)
- D-TEC (Dan Quinlan, Saman Amarasinghe)
- Traleika Glacier (Josh Fryman)
- XPRESS (Ron Brightwell)
- PIPER (Todd Gamblin)
- SLEEC (Milind Kulkarni)
- Corvette (Koushik Sen)
- X-TUNE (Mary Hall)
- Vancouver & ARES (Jeff Vetter)

Project Summaries

- DEGAS: Build a PGAS programming environment and toolkit that delivers high performance and productivity to DOE applications on current and future systems.

  Impact: https://xstackwiki.modelado.org/images/0/09/DEGAS-Highlight_Summary.pdf

- D-TEC: Simplify the development of DSLs to enable scalability, migration, resiliency, verification on current and future hardware platforms.

  Impact: https://xstackwiki.modelado.org/images/9/99/D-TEC_Summary_Highlight.pdf

U.S. DEPARTMENT OF ENERGY | Office of Science

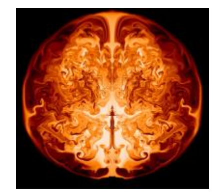
DEGAS Project: Impact of advances in data Structures and Runtime Support for Irregular Data-Intensive Applications

Speedups

- Distributed hash table
  - Applications: HipMer (genomics)

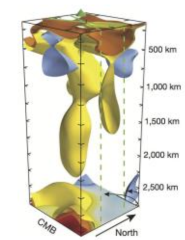
720x

- Irregular data exchange
  - Applications: AMR, HPGMG



1.2x

- Irregular global matrix update
  - Applications: NWChem, seismic tomography

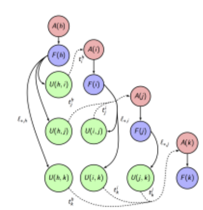


6x

- Distributed work queue
  - Applications: NWChem, Hartree-Fock

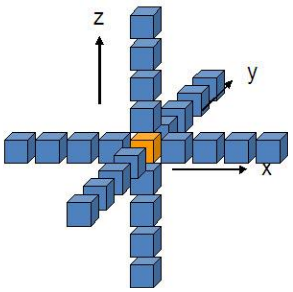
1.2x

- Dynamic task graph
  - Applications: Sparse symmetric matrix solver

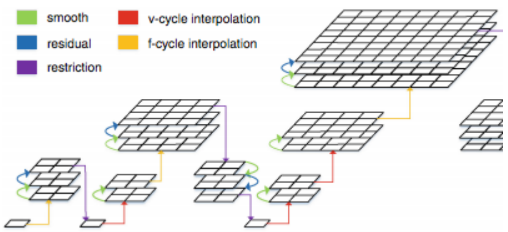


2x

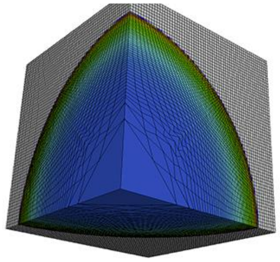
D-TEC Project: Impact of advances on DSL technologies, compilers and runtime systems



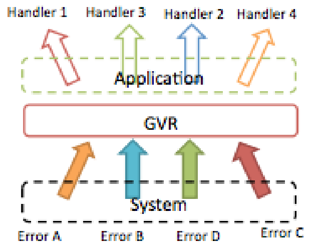
- AMR Shift Calculus DSL with ROSE/Polyopt
  - **7.9x** for 3D 125 pts stencil
  - **7.3x** for 2D 81 pts stencil
- Halide DSL
  - **4.25x** for MiniGMG
  - **1.8x** on GPU for HPGMG



- Bamboo
  - **1.27x** in 32K size for algebraic multigrid
  - **1.29x** with 96K grid cells for 3D stencil



- LULESH with X10
  - **1.12x** better performance
  - **40%** fewer lines of code
- Global-View Resilience (GVR)
  - **85%** parallel efficiency on 1K processes with less than 2% code change



Project Summaries (contd)

- Traleika Glacier: Explore new introspective event-driven execution model embodied in an Open Community Runtime, and evaluate its effectiveness in mapping extreme parallelism on future hardware with open and shared simulation infrastructure.

  Impact: https://xstackwiki.modelado.org/images/2/21/Traleika_Glacier_Impacts.pdf

- XPRESS: Prototype implementation of software stack (OpenX) to support the ParalleX Execution Model with HPX runtime system, LXK lightweight operating system, with support for legacy MPI and OpenMP codes.

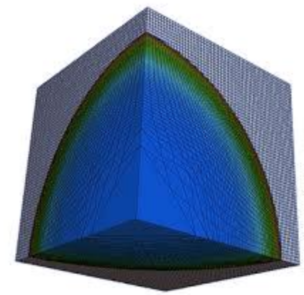
  Impact: https://xstackwiki.modelado.org/images/7/78/XPRESS-Highlights_Summary.pdf

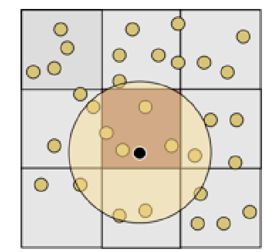
U.S. DEPARTMENT OF ENERGY | Office of Science

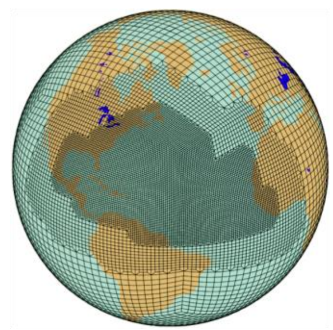
Traleika Glacier Project: The Open Community Runtime Software Suite and its Impact on Applications

- Applications: Full set of DOE proxies and the full app atmospheric circulation, Tempest, Partial list of real applications on OCR
  - Tempest
  - SCF from NWCHEM
  - CoMD
  - HPCC and HPCG kernels
  - Lulesh (multiple versions)
  - miniAMR
  - HPGMG – **11x**
  - **Genomics Smith-Waterman** - **9x**
- Full OCR API supported on real hardware and is exploited by these tool chains
  - C library, C++library
  - CnC on OCR, Hierarchically Tiled Arrays (HTA) on OCR
  - Compiler generation of OCR calls (R-Stream)
  - Habanero-C language on OCR

Open Community Runtime





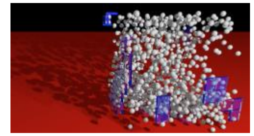


Please see: https://xstack.exascale-tech.com/wiki/index.php/Main_Page#Traleika_Glacier_Research_Products and https://xstack.exascale-tech.com/wiki/index.php/Traleika_Glacier_Software_Releases for details, products and research successes

XPRESS Project: impact of exascale runtime support (HPX)

Comparisons/Results

- N-body Simulation



1.4x over MPI (16,384 cores)

- Mini-Ghost: Boundary Exchange Mini-app

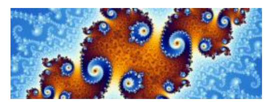
1.13x over MPI+OpenMP (1024 cores)

- Kernels: Stream benchmark, Matrix transposition

1.4x over OpenMP

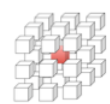
2.5x over MPI+OpenMP

- Distributed GPU work



1.5x over native CUDA on 16 GPUs

- Lulesh: Shock Hydrodynamics



1.2x over MPI on Cori (128 cores)

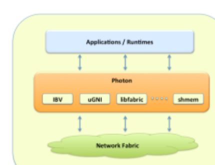
Same as MPI on Cori (4k cores)

- DSEL and MTL for HPX
  - Same Portable code GPU / CPU

DSL for linear algebra through DOE NNSA DE-NA0002377 (PSAAP2)

Same as MPI (256 cores)

- Photon: Integrated Communication Library



1.34x for 16 byte puts

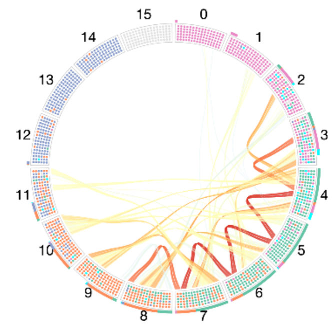
1.37x for 16 byte gets

over MPI-3 one-sided

Project Summaries (contd)

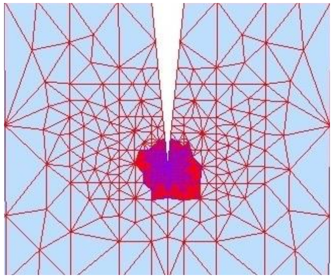
- PIPER: Performance analysis for the X-Stack that provides insights into application performance, enabling users and runtimes to act on these insights.

  Impact: https://xstackwiki.modelado.org/images/b/b8/PIPER-Impact_summary.pdf



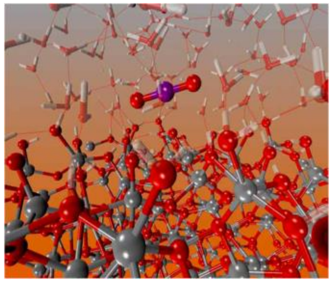
- SLEEC: Improve application performance by leveraging domain knowledge in the form of library annotations, and semantics-aware, extensible optimizing compiler technologies.

  Impact: https://xstackwiki.modelado.org/images/b/b9/SLEEC-Highlight_summary.pdf



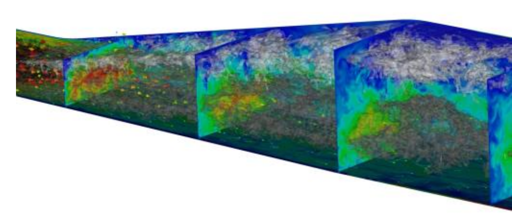
- Corvette: Improved debugging of scientific applications via automated bug finding methods, and making concurrency bugs and floating point behavior reproducible.

  Impact: https://xstackwiki.modelado.org/images/4/4c/Corvette-highlight_summary.pdf

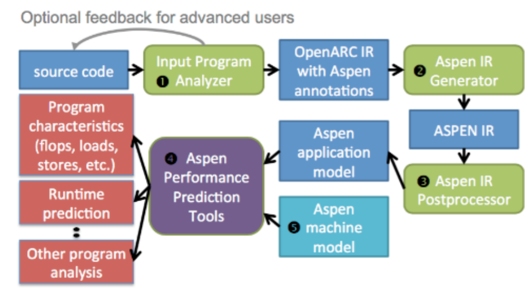


credit:nwchem-sw.org

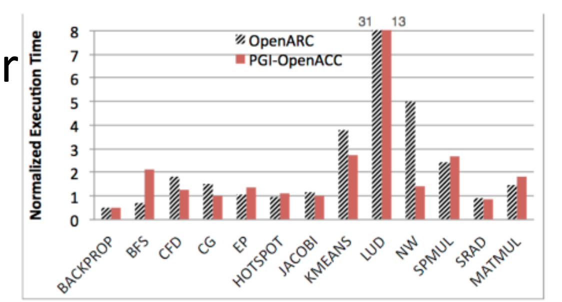
Project Summaries (contd)



- X-TUNE: Use auto-tuning to enable performance and software portability to map simple code specification to optimized implementation for a given platform. **Impact:**https://xstackwiki.modelado.org/images/d/de/Xtune-impact_summary.pdf



- Vancouver 2: Software infrastructure (compilers, libraries, performance tools) for productive heterogeneous exascale computing. **Impact:**https://xstackwiki.modelado.org/images/5/57/Vancouver-highlight_summary-v02.pdf



- ARES: Abstract representations for the extreme-scale stack leveraging the widely adopted LLVM infrastructur **Impact:**https://xstackwiki.modelado.org/images/3/33/Ares-highlight_summary-v02.pdf

https://xstackwiki.modelado.org/images/b/ba/X-Stack_Impact_Summary_-_September_2016.pdf

X-Stack Publications & Software Releases

- Total publications ~ 400
  - Includes top-tier Computer Science venues, e.g.,
  - SC'12, ICS'13, PLDI'13, ASPLOS'14, IPDPS'14, PLDI'14, OOPSLA'14, SC'14, PPoPP'15, CGO'15, ASPLOS'15, IPDPS'15, ICS'15, PLDI'15, IPDPS'16
- Total software products ~ 200
- For details, see:https://xstackwiki.modelado.org/Extreme_Scale_Software_Stack#X-Stack_Program_Software_and_Publications_per_Project

Facilities Panel

- "Successes and Challenges in adopting new software technologies in HPC Facilities"
- Participants
  - Katie Antypas (LBNL)
  - David Bernholdt (ORNL)
  - Hal Finkel (ANL)
  - Gustav Jansen (ORNL)
  - Matthew Norman (ORNL)
  - Prabhat (LBNL)
- Observations
  - Facilities staff install new software all the time, but can only justify doing so when there is a critical mass of users
  - Simplified tool chain (e.g., libraries) and training resources are important enablers for transitioning research software to use in facilities

Meeting Agenda: Day 2 (April 7, 2016)

Day 2 Goals

- Insights on application challenges from keynote speakers
- Five breakout groups to discuss future CS research topics
- Report back from breakout groups + wrap-up

Application presentations (2 hours)

- Next Generation AMR (Ann Almgren)
- Future Programming Challenges for DOE (David Richards)
- The Importance of DSL Technology (Paul Woodward)
- Distributed Learning Dynamics Convergence (Alex Bayen)

Breakout groups & report back (6 hours)

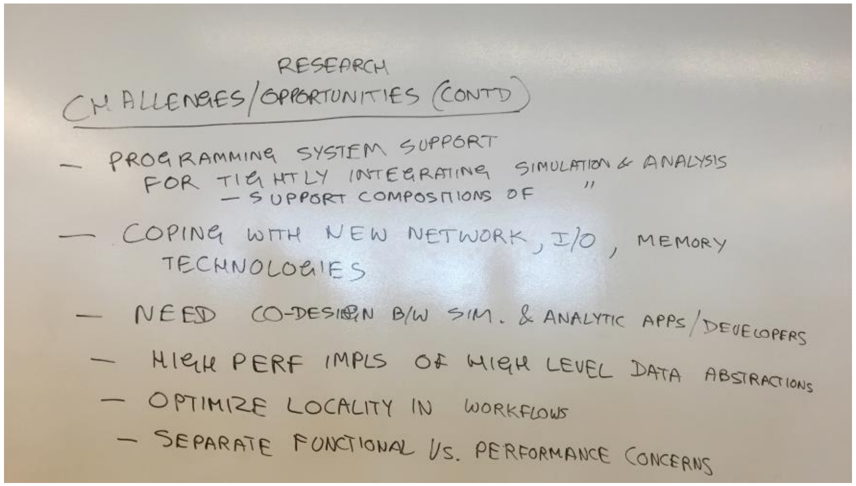
1. Future Synergies in Programming Systems for Data-Intensive and Compute-Intensive Science. Moderator: Vivek Sarkar
2. Programming Systems for Data Centric Computing: Demands from User facilities, experiments and sensor networks. Moderators: Dan Quinlan and John Wu
3. Programming Systems Support for Post Moore's Computing. Moderators: Jeff Vetter and Bob Lucas
4. Ensuring Correctness for Exascale and Beyond. Moderators: Benoit Meister and Koushik Sen
5. High Performance Languages. Moderator: Saman Amarasinghe

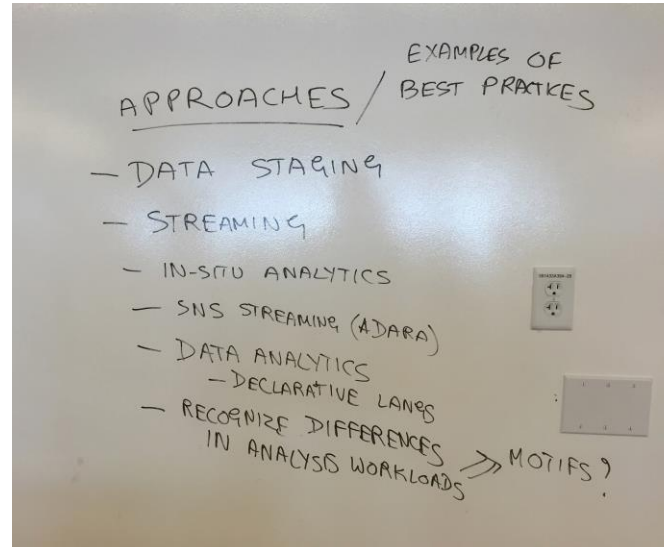
Summary of Breakout 1 on "Future synergies in programming systems for data-intensive and compute-intensive science"

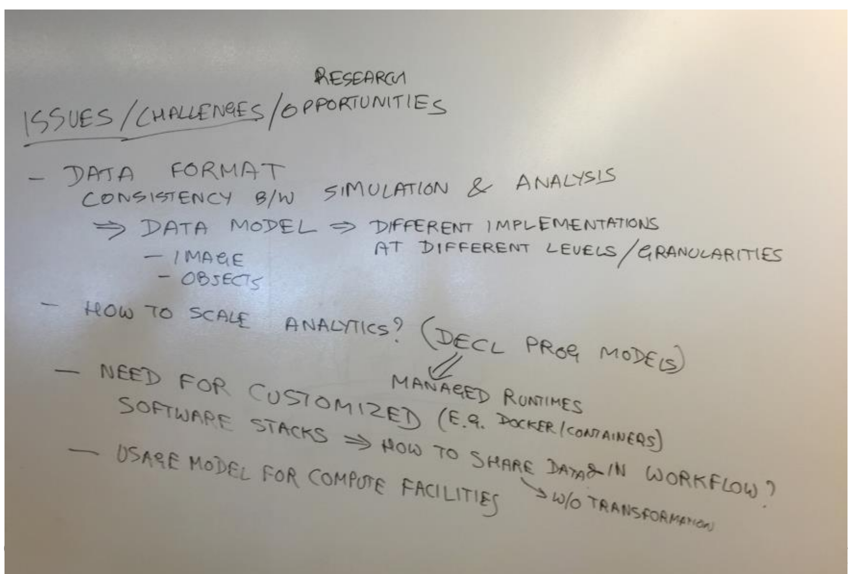
- Exploiting synergies between data-intensive and compute-intensive computations is critical for future HPC applications
- Early experiences point to potential benefits
- Multiple research opportunities, including:
  - HPC data models and implementations
  - Programming systems for integration of simulations and analytics
  - Addressing new scalability and performance bottlenecks
  - Scalable integration of customized software stacks for science workflows
- ~ 20 participants

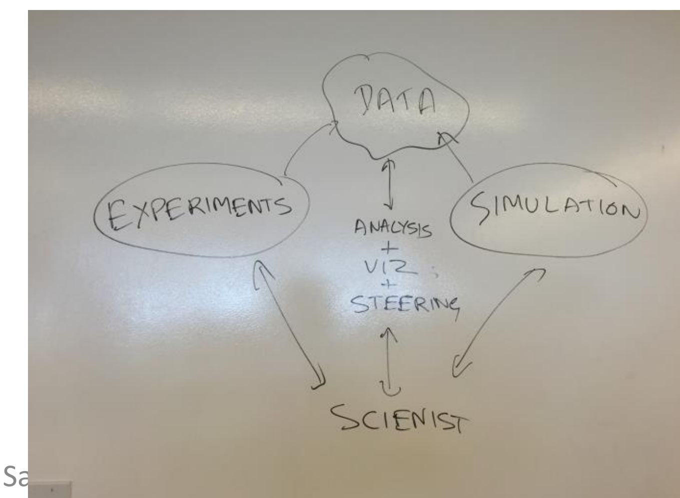
U.S. DEPARTMENT OF ENERGY | Office of Science

Summary of Breakout 1 on “Future synergies in programming systems for data-intensive and compute-intensive science” (contd)





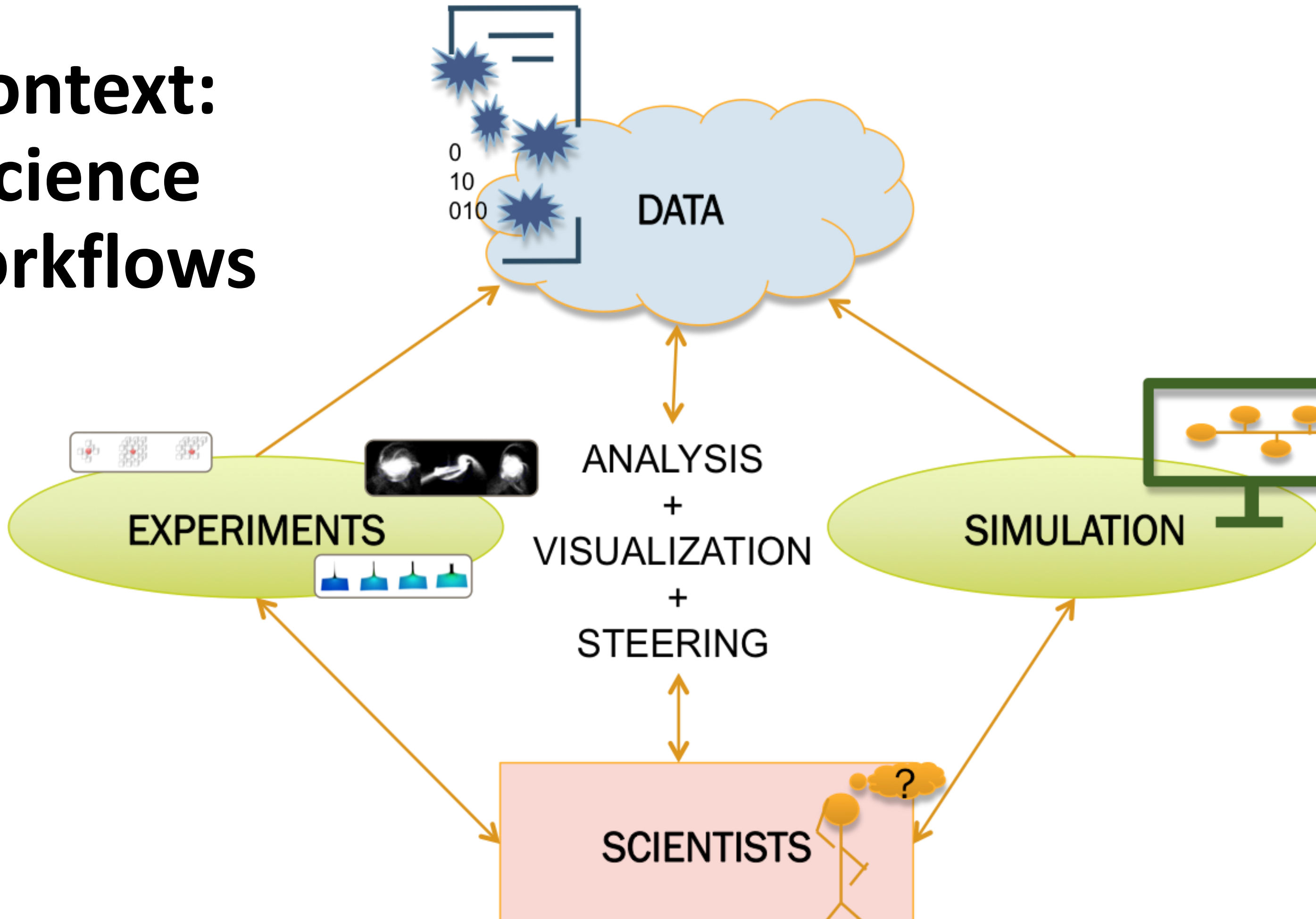




U.S. DEPARTMENT OF ENERGY | Office of Science

Summary of Breakout 1 on "Future synergies in programming systems for data-intensive and compute-intensive science" (contd)

Context: Science Workflows



U.S. DEPARTMENT OF ENERGY | Office of Science

Summary of Breakout 2 on "Programming systems for data centric computing"

- Application examples
  - Smart Grid, Smart Cities, Smart Mobility, BNL nanotube optimization, ,agnetic fusion experimental data management/analysis, Compact Laser Accelerator such as BELLA, NNSA HED physics facility such as NIF, astrophysics simulation, …
- Research opportunities include
  - Programming systems for data-centric applications
  - Schedulers/resource managers that address dynamic behavior of data/systems
  - Application portability across HPC/Cloud platforms
  - Constraint-based software stack: real-time, interactive,
  - Computational Steering (introspect and adapt): terminate experiment, change parameters, feedback
- ~ 15 participants

Summary of Breakout 3 on "Programming Systems Support for Post Moore's Law Computing"

- How can we exploit new technologies for science?
  - QC/QA, FPGA, Optical, Molecular computing, Neuromorphic and brain-inspired computing, Probabilistic and stochastic computing, new memory technologies, …
- How do we incorporate new technologies into DOE's science processes?
- ~ 18 participants

Summary of Breakout 4 on "Ensuring Correctness for Exascale and Beyond"

- Correctness issues: Inaccuracies
  - Inaccuracies from non-determinism ("non-reproducibility")
  - Inaccuracies and Instabilities from using too little FP precision
- Correctness issues: Inefficiencies
  - Inefficiencies from overly conservative data sharing (too many "barriers")
  - Inefficiencies from using high precision floating point everywhere
- Detection Approaches
  - Static analysis (automated, partial guarantees, false positives)
  - Dynamic analysis (automated, partial guarantees, false negatives)
  - Theorem proving (semi-automated, full correctness)
  - New algorithms
- Accuracy & Reproducibility Approaches
  - Static analyisis for worst-case accuracy
  - Adapt computations to improve accuracy, e.g., Herbie
  - Minimize precision while guaranteeing good-enough accuracy, e.g. Precimonious
  - Maximize reproducibility while staying within worst-case accuracy boundaries, e.g., reproBLAS

Summary of Breakout 5 on "High Performance Languages"

- Risk Factors in adopting a New Language
  - New syntax & semantics → new skills to be learned
  - Immature tool chain and developer ecosystem
  - … yet new languages are adopted all the time in commercial/enterprise computing
- How to lower (real/perceived) barrier to entry for adopting new languages in HPC?
  - Hackathons to try out new languages on real-world problems
  - Funding for 3-month mini-projects (micro-interactions)
  - Embed CS graduate students within application teams
  - Embody programming model in library instead of language

Outline

- Background on X-Stack Program
- X-Stack PI meeting, April 6-7, 2016
- <u>Conclusions</u>

Some Observations & Lessons Learned

- Computations and communications will be irregular due to heterogeneous processors characteristics, deep memory hierarchies, and much higher rate of faults than current platforms. Programmability and performance portability become increasingly important and challenging.
- With DSLs, application code size can be reduced by 100x, and the code generated and optimized has about 3x faster execution time.
- Unprecedented parallelism, asynchrony, and irregular computations and communications are well suited to task-oriented dataflow programming models.
- Continued advances in programming systems, runtime systems, and tools are critical for the Moore's Law "end game" (even beyond exascale)

Legend: red = end by Sep'16, black = current

Computer Science Solicitations related to Extreme Scale Computing

- Jan 2010: Advanced Architectures and Critical Technologies for Exascale Computing (FOA and Lab announcement)
  - Program duration: FY10-FY12
  - Renewal: FY13-FY15
- Jan 2010: X-Stack Software Research (FOA and Lab announcement)
  - Program duration: FY10-FY12
- Jan 2010: Scientific Data Management and Analysis at Extreme Scale (FOA and Lab announcement)
  - Program duration: FY10-FY12
- June 2010: Exascale Co-Design Centers (Lab announcement)
  - Program duration: FY11-FY16

Legend: red = end by Sep'16, black = current

Computer Science Solicitations related to Extreme Scale Computing (contd)

- Nov 2011: X-Stack: Programming Challenges, Runtime Systems, and Tools (FOA and Lab announcement)
  - Program duration: FY12 - FY14
  - Renewal: FY15
- Jan 2013: Exascale Operating and Runtime Systems (Lab announcement)
  - Program duration: FY13 – FY15
- March 2014: Scientific Data Management, Analysis and Visualization at Extreme Scale 2 (FOA and Lab announcement)
  - Program duration: FY14 – FY16

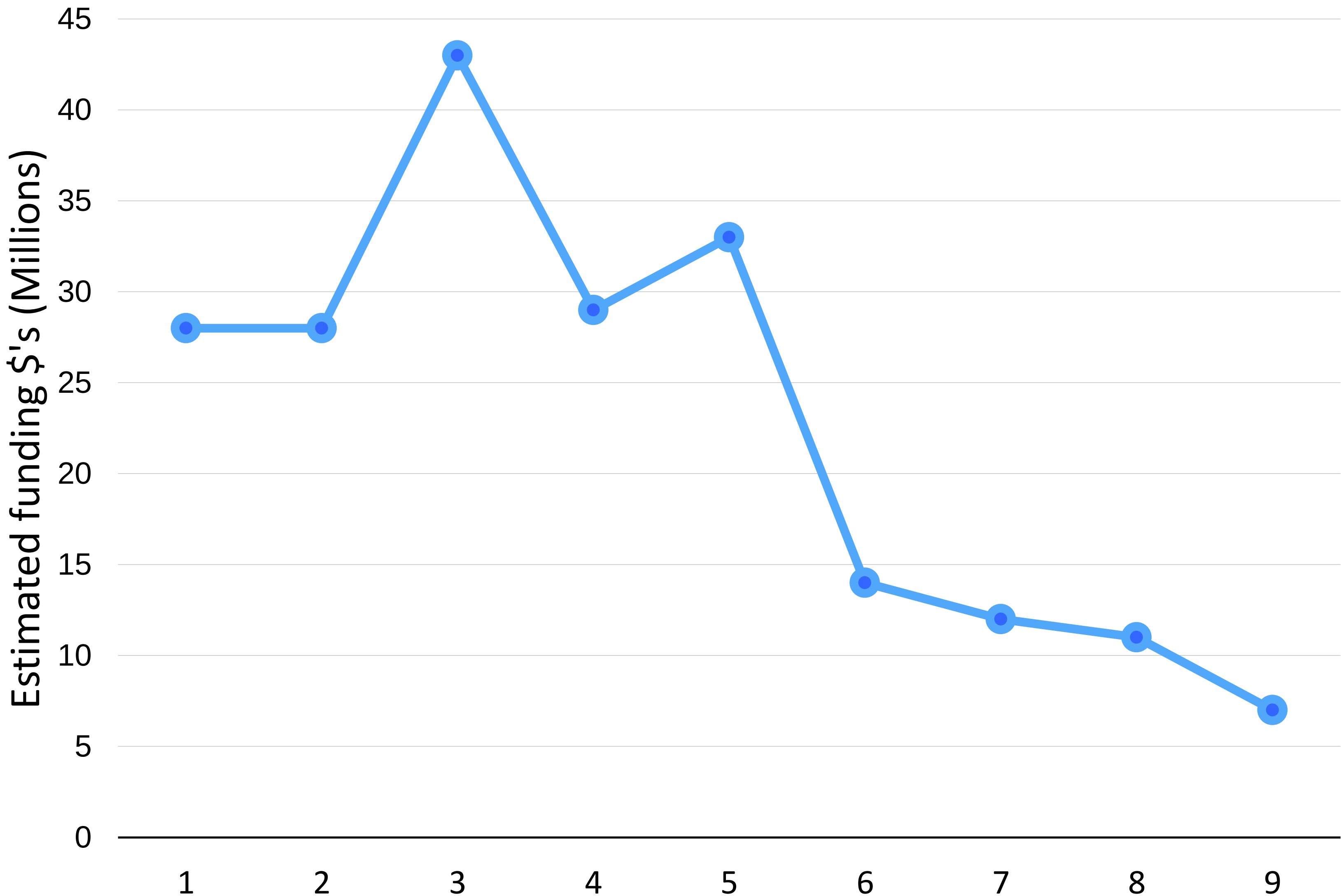
Legend: red = end by Sep'16, black = current

Computer Science Solicitations related to Extreme Scale Computing (contd)

- July 2014: Resilience for Extreme Scale Supercomputing Systems (FOA and Lab announcement)
  - Program duration: FY14 - FY16
- May 2015: Storage Systems and Input/Output for Extreme Scale Science (FOA and Lab announcement, May 2015)
  - Program duration: FY15 – FY17
- April 2016: Machine Learning and Understanding for High Performance Computing Scientific Discovery (FOA and Lab announcement)
  - Program duration: FY16 – FY18

U.S. DEPARTMENT OF ENERGY | Office of Science

CS research programs related to Extreme Scale Computing (Estimates based on target funding $'s in solicitations)



U.S. DEPARTMENT OF ENERGY | Office of Science

Some concluding thoughts

- The 2016 X-Stack PI meeting demonstrated US leadership in cutting-edge software research on programming systems, runtime systems, and tools for current and future HPC systems
  - Given the current funding realities, X-Stack participants in academia have now begun moving to CS research opportunities outside of ASCR and DOE
  - Implications on recruiting & retention in national labs
- NSCI envisions a "coordinated research, development, and deployment strategy" for High-Performance Computing
  - "Executive Order -- Creating a National Strategic Computing Initiative", July 29, 2015
- How do we ensure the continued health of an HPC software research pipeline for exascale systems and beyond?