U.S. DEPARTMENT OF ENERGY | Office of Science

ASCR High End Computing Requirements Gathering

Briefing to MAGIC

September, 6 2016

Prepared by: Carolyn Lauzon, Advanced Scientific Computing Research

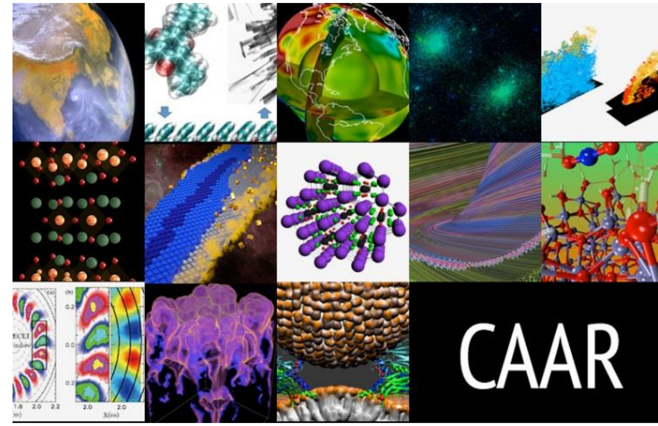
Disclaimer Notes

- **Requirements Gathering is currently ongoing**
- **Conclusions and Summaries <u>herein represent my own views</u> to an on-going process and will be updated when requirements process is complete**

ASCR Facilities and Mission: Accelerate Scientific Discovery

Argonne **Leadership Computing** Facility

Oak Ridge **Leadership Computing** Facility



National Energy Research **Scientific Computing**

Energy Sciences **Network**

1986–2016
ESnet
30 YEARS OF NETWORKING
AT THE SPEED OF SCIENCE

ASCR Computing Upgrades At a Glance

System attributes	NERSC Now	OLCF Now	ALCF Now	NERSC Upgrade	OLCF Upgrade	ALCF Upgrades	
Name Planned Installation	**Edison**	**TITAN**	**MIRA**	**Cori** **2016**	**Summit** **2017-2018**	**Theta** **2016**	**Aurora** **2018-2019**
System peak (PF)	2.6	27	10	> 30	150	>8.5	180
Peak Power (MW)	2	9	4.8	< 3.7	10	1.7	13
Total system memory	357 TB	710TB	768TB	~1 PB DDR4 + High Bandwidth Memory (HBM)+1.5PB persistent memory	> 1.74 PB DDR4 + HBM + 2.8 PB persistent memory	>480 TB DDR4 + High Bandwidth Memory (HBM)	> 7 PB High Bandwidth On-Package Memory Local Memory and Persistent Memory
Node performance (TF)	0.460	1.452	0.204	> 3	> 40	> 3	> 17 times Mira
Node processors	Intel Ivy Bridge	AMD Opteron Nvidia Kepler	64-bit PowerPC A2	Intel Knights Landing many core CPUs Intel Haswell CPU in data partition	Multiple IBM Power9 CPUs & multiple Nvidia Voltas GPUS	Intel Knights Landing Xeon Phi many core CPUs	Knights Hill Xeon Phi many core CPUs
System size (nodes)	5,600 nodes	18,688 nodes	49,152	9,300 nodes 1,900 nodes in data partition	~3,500 nodes	>2,500 nodes	>50,000 nodes
System Interconnect	Aries	Gemini	5D Torus	Aries	Dual Rail EDR-IB	Aries	2nd Generation Intel Omni-Path Architecture
File System	7.6 PB 168 GB/s, Lustre®	32 PB 1 TB/s, Lustre®	26 PB 300 GB/s GPFS™	28 PB 744 GB/s Lustre®	120 PB 1 TB/s GPFS™	10PB, 210 GB/s Lustre initial	150 PB 1 TB/s Lustre®

Science Requirements for HPC Acquisitions

What HPC systems and ecosystem will best advance science in the 2020 - 2025?

Ongoing Activities to Collect Science Drivers and Barriers to Exascale

ASCR Exascale Requirements Reviews

– ASCR Facilities Partner with DOE Office of Science

Request for Information

– ASCR solicit from DOE national labs

– NIH led with DOE and NSF, solicit broader community

NSF

NIH National Institutes of Health
Turning Discovery Into Health

Ongoing Activities to Collect Science Drivers and Barriers to Exascale

ASCR Exascale Requirements Reviews

– ASCR Facilities Partner with DOE Office of Science

Request for Information

– ASCR solicit from DOE national labs

– NIH led with DOE and NSF, solicit broader community

Previous Requirements Gathering Efforts: "Lead with the Science"

Value of Approach

- **Review meetings establish requirements, capabilities, services**
- **Scientists, programs offices, and facilities have the same conversation**
- **Provides a solid, fact-based foundation for service and capability investments**
- **Addresses DOE mission goals by ensuring DOE science is effectively supported**

Objectives of Exascale Requirements Review (RR)

Goal: Ensure the ability of ASCR facilities to support SC mission science in the exascale regime (2020-2025 timeframe).

Partner Office

- Identify key computational science objectives that push exascale
- Describe the HPC ecosystem needed to successfully accomplish your science goals
  - HPC machine and related resources

ASCR

- Communicate to DOE SC scientists the known/fixed characteristics of upcoming compute system
- Ask the computational scientists for feedback on proposed architectures.

Strengthen and inform interactions between HPC facility experts and scientists

Implementation of Exascale Requirements Review (RR)

Series of workshops, one per SC Office Website: http://www.orau.gov/exascaleage/

- **Location:** Washington DC area
- **Attendance**: ~100 attendees
  - Domain Scientists
  - ASCR Facilities Staff
  - DOE Headquarters (ASCR and Partner Office)

- **Pre-meeting homework**: Templates will be developed and provided to chairs and attendees of breakout session for discussing and documenting case studies
- **Output**: Summary workshop report written for each workshop.

Schedule

June 10-12,2015	High Energy Physics
November 3-5 2015	Basic Energy Sciences
January 2015	Fusion Energy Sciences
April/March 2016	Biological and Environmental Research
June 2016	Nuclear Physics
September 2016	ASCR

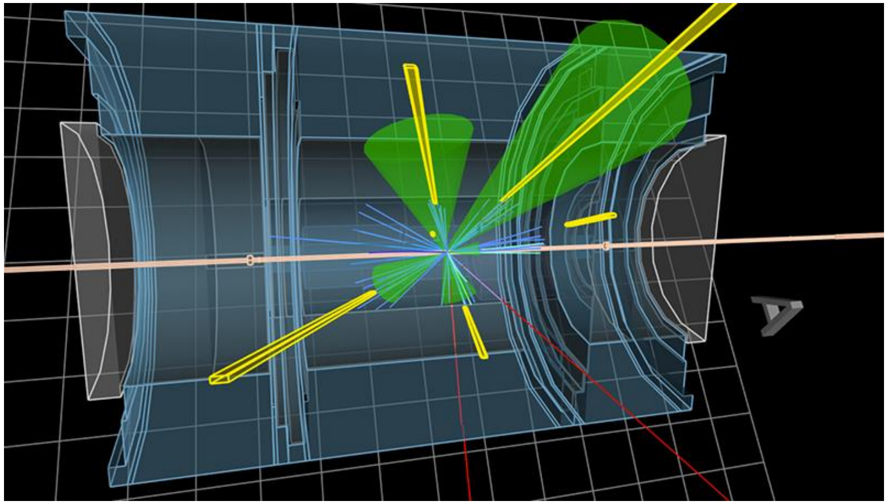
High Energy Physics Requirements Review

Stats:

- Held at Bethesda Hyatt on June 10-12, 2015
- Headquarter’s POC: Lali Chaterjee (HEP), Carolyn Lauzon (ASCR)
- Conference Co-Chairs: Salman Habib (ANL); Rob Roser (Fermi Lab)
- 71 attendees: 34 application scientists
- 9 white papers and 9 case studies

Breakout Sessions:

- Compute Intensive Modeling and Simulation
- Data Focused Analysis and Workflows



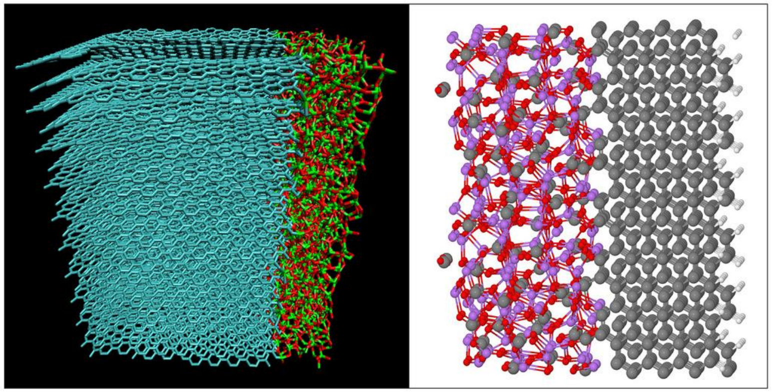
Basic Energy Sciences Requirements Review

Stats:

- Held at Rockville Hilton on November 3-5, 2015
- Headquarter's POC: Jim Davenport, Mark Pederson, Eliane Lessner (Basic Energy Sciences), Carolyn Lauzon (ASCR)
- Conference Co-Chairs: Michael Banda (LBNL); Tom Devereaux (Stanford/SLAC), Theresa Windus (ISU, Ames Lab)
- 99 attendees, 60 domain scientists, 36 observers
- 50 of 60 domain scientists submitted white papers

Breakout Sessions:

- Science Areas
  - Quantum Materials, Heavy Element Chemistry, Exotic States, Emergence
  - Catalysis, Photosynthesis, Light-harvesting, Combustion
  - Materials and Chemical Discovery
  - Soft Matter, Biochemistry, Bioinspired Materials
- Ecosystem Challenges
  - Computing and Data Challenges at BES Facilities
  - Next Generation Programming
  - Advances in Quantum Algorithms
  - Math and Computer Science



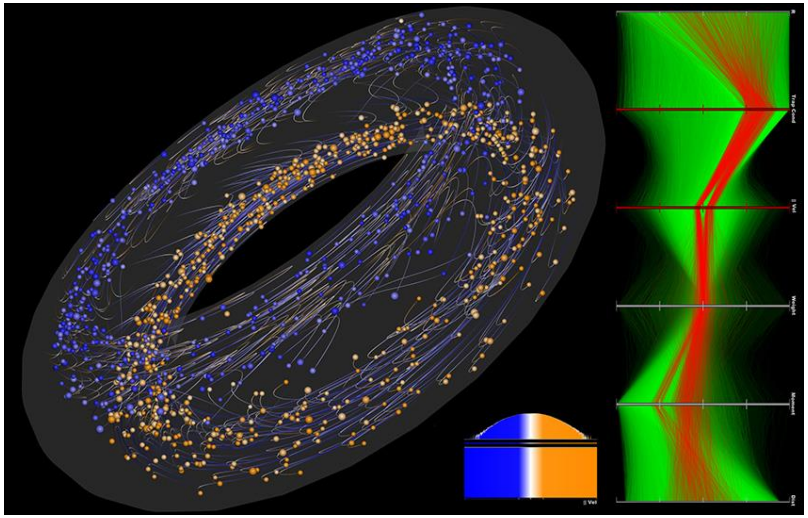
Fusion Energy Sciences Requirements Review

Stats:

- Held at Gaithersburg Hilton on January 27-29, 2016
- Headquarter's POC: John Mandrekas (Fusion Energy Sciences), Carolyn Lauzon (ASCR)
- Conference Co-Chairs: C.S. Chang (PPPL), Martin Greenwald (MIT)

Breakout Sessions:

- Magnetic Fusion Energy Sciences
  - Turbulence and Transport
  - MagnetoHydroDynamics & Energetic Particles
  - RF Heating & Current Drive
  - Whole Device Modeling (cross-cut)
  - Validation and Verification (cross-cut)
- Materials Science
  - Plasma Surface Interactions
  - Structural Materials
- Discovery Plasma Science
  - General Plasma Science
  - High Energy Density Plasmas
  - Low Temperature Plasmas



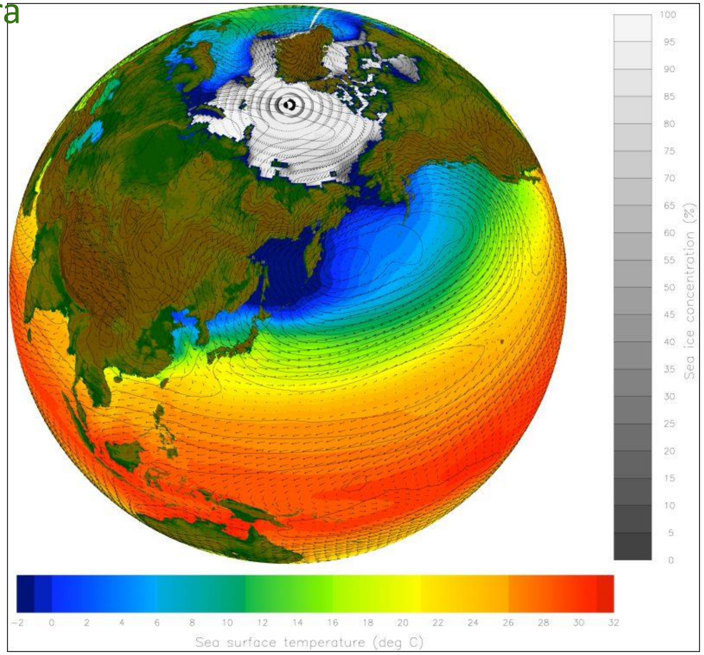
BER Requirements Reviews

Stats:

- Held at Rockville Hilton on March 29-31, 2016
- Headquarter's POC: Doroth Koch (BER-CESD), Ramana Madupu (BER-BSSD), Carolyn Lauzon (ASCR)
- Conference Co-Chairs: David Bader (LLNL), Adam Arkin (LBL)

Breakout Sessions:

- Climate
  - Climate and the Environmental Sciences in the Exascale Era
  - Component Specific Challenges: Atmosphere, Terrestrial, Oceans and cryosphere
  - Coupled Systems Integration
  - Modeling climate and environment
  - Unlocking scientific knowledge with Exabytes of Data
  - Data-Model Fusion
- Biology
  - Sequence to molecular structure and function
  - Integrative modeling of cells and networks
  - Big Data for Knowledge Discovery in Biology
  - Microbes and the Environment
  - Multiscale simulation of Biological Processes



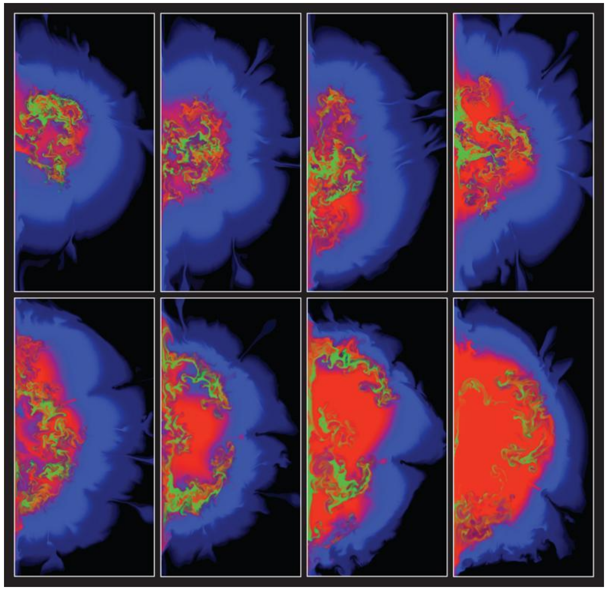
Nuclear Physics Requirements Reviews

Stats:

- Held at Gaithersburg Hilton on June 15-17, 2016
- Headquarter's POC: Ted Barnes (NP), Carolyn Lauzon (ASCR)
- Conference Co-Chairs: Martin Savage (U. Washington), Joe Carlson (LANL)
- 104 attendees; 51 Domain Scientists

Breakout Sessions:

- Astrophysics
- Cold QCD
- Hot QCD
- Nuclear Structure
- Data/Experiment



Chemical structure of the ejected debris 100 s after ignition for a subset of SNe Ia explosion models with different ignition and detonation conditions. (D Kasen *et al. Nature* **460**, 869-872 (2009) doi:10.1038/nature08256)

ASCR Requirements Reviews

Stats:

- To be Held at Rockville Hilton on September 27-28, 2016
- Headquarter's POC: Ceren Bennett, Lucy Nowell, Betsy Riley, Carolyn Lauzon (ASCR)
- Conference Co-Chairs: Jeffrey Vetter (ORNL), Ann Almgren (LBNL), Phil DeMar (FNAL)
- TBD but we are booked (100+ pre-registered)

Breakout Sessions:

TBD

Data Intensive and Experimentally Driven Computing

- **4 of 5 partner offices had specific data/experiment breakout group**
  - Need for rapid data movement between compute facility and experimental facility growing
- **FES, with no data/experiment breakout group, had very active discussion about 'can we use HPC for experiments?'**
- **Emergence of new use models**
  - HEP with ATLAS event generation
  - BES compute demand at beamlines is breaking current model (scientist process on site or takes data home)
- **ESnet representatives present at meeting and participating in summary analysis by ASCR HPC facilities**

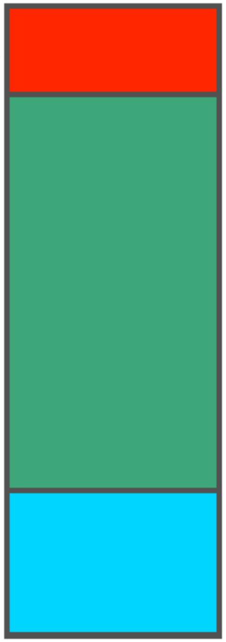
Example from HEP

Slides from Tom LeCompte (ANL) presentation at HEP Exascale Requirements Review. Available Here: http://www.nersc.gov/science/hpc-requirements-reviews/exascale/HEP/hep-exascale-review-presentations/

Excerpt from HEP RR Presentation (Tom LeCompte, ANL)

Computing to Reach The Science Goals

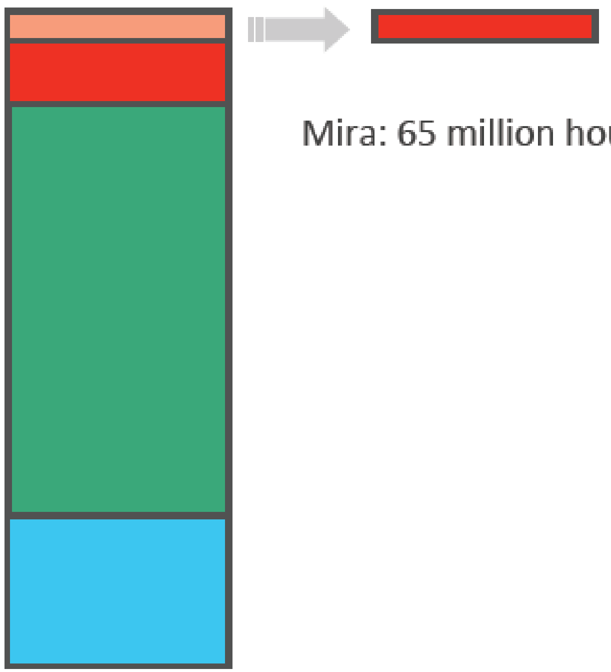
- ATLAS uses about a billion CPU-hours per year on the Grid
  - This does not include the cycles spent calibrating or reconstructing the data; the problem is defined as what happens after this point
- Event Generation
  - Simulate the physics process of interest: produces lists of particles and their momenta
- Simulation
  - Simulate the interaction of these particles with the detector
  - You may have heard the term “Geant” or “G4”. Geant4 is the toolkit by which we do this.
- Reconstruction and Analysis
  - Treat the simulated data as real, reconstruct the particles, and do the final analysis



5

Excerpt from HEP RR Presentation (Tom LeCompte, ANL)

ALCC-2014 Use



Mira: 65 million hours

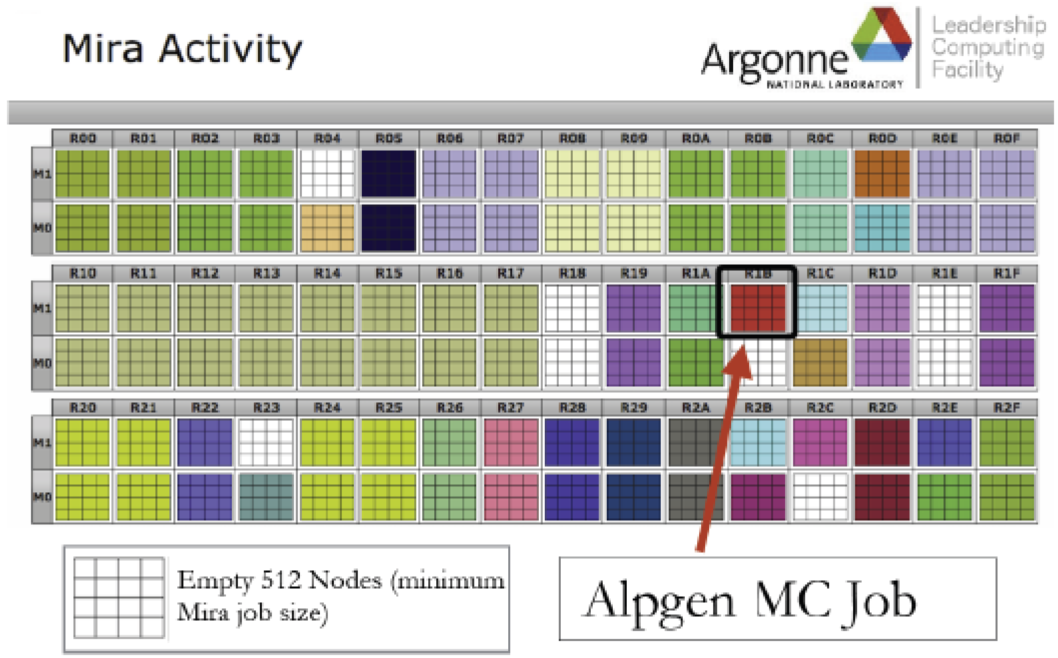
Grid: One billion+ hours

- 70 million Grid-equivalent cpu-hours of event generation have been run on Mira and thereby offloaded from the Grid
  - 2 FTE worth of effort
  - If we were a country, we would be the 7th largest provider of cycles to ATLAS
- This is only event generation (red), but the recovered cycles can be used for whatever the experiment wants
  - This is dominated by one event generator (called Alpgen)

Excerpt from HEP RR Presentation (Tom LeCompte, ANL)

Where we were 1 year ago

- We could run in the minimum Mira partition (and only in the minimum efficiently)
- Event generation rate was 1/15 of a Grid node
  - ALCF suggests a nominal of 1/10

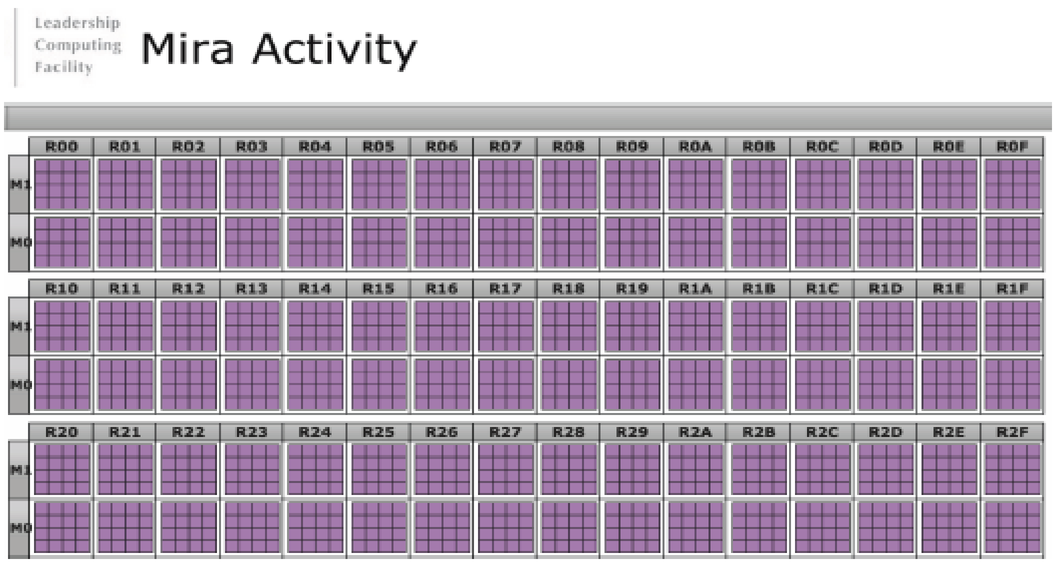


At this time we were running, but limited by I/O.

Excerpt from HEP RR Presentation (Tom LeCompte, ANL)

Where we are today

- We can run using the entire machine
  - For throughput reasons, we normally limit ourselves to 1/3 of the machine: a million parallel processes
- Per core event generation rate is ~1.5x a Grid core
  - Speedup of x23 over the year
  - Mira has a lot of cores! (768432)



While this job was running, Mira was producing the equivalent computing as 5 or 6 ATLAS Grids.

On our best days, we provide the equivalent computing capacity of the whole ATLAS Grid.

One person's perspective on themes (workshop reports TBD)

- **Integration of Data Intensive Research/Experiments**
  - Realtime remote access to resources such as databases
  - Realtime data analysis
  - Machine learning from multiple data sources – pattern recognition
  - Data visualization tools
  - I/O bandwidth challenges, scheduling, queuing policy

- **Usability of Exascale Systems**
  - Libraries: scalable, accelerator enabled, and hide hardware details
  - Parallel primitives/ general purpose mid to low level tools that can be replaced with hand tuned code
  - Method for managing memory hierarchies
  - PORTABILITY
    - Data layout for memory hierarchy, DSL, common low level tools
  - Concern regarding Memory limitations , reduced I/O, and interconnect for communication

One person's perspective on themes (workshop reports TBD)

- **Implementation of Physics**
  - Mulitscale and multiphysics – how pass information from one scale to another in mathematically rigorous way?
  - Mathematics + CS (BES mentioned specifics: e.g. FFT on non-equi-spaced data, methods to sample rare configurations, scalable asynchronous runtime systems)
  - VV- Integrate theory and experiment to develop algorithms, library and codes for exascale
  - UQ for model reduction, assessing predictive power of models and transfer of basic research to engineering applicaitons
- **Partnerships and Workforce Development**
  - More SciDAC!
  - New ASCR-SC partnership model
    - Integrate applied mathematicians, computer scientists and application developers
  - Workforce development
    - Next generation of computational scientists that know science and know computational issues

Ongoing Activities to Collect Science Drivers and Barriers to Exascale

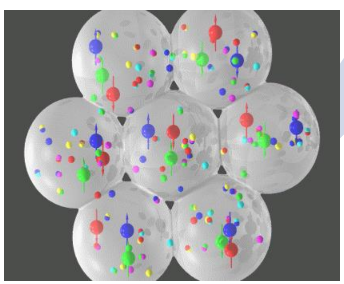
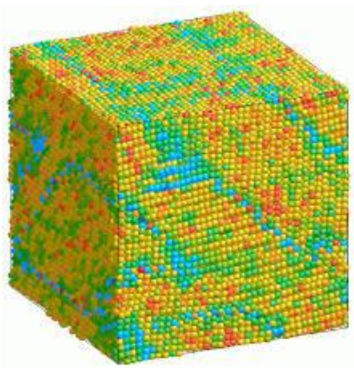
ASCR Exascale Requirements Reviews

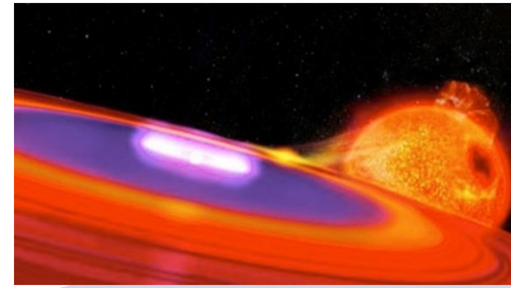
– ASCR Facilities Partner with DOE Office of Science

Request for Information

– ASCR solicit from DOE national labs
– NIH led with DOE and NSF, solicit broader community

DOE National Laboratories Responses: 135
Academic Responses: 94
Industry Responses: 8
Foreign Response: 2
Others: 5

Nuclear, Particle & Plasma Physics

Chemical and Material Sciences

Biological Sciences & Public Health

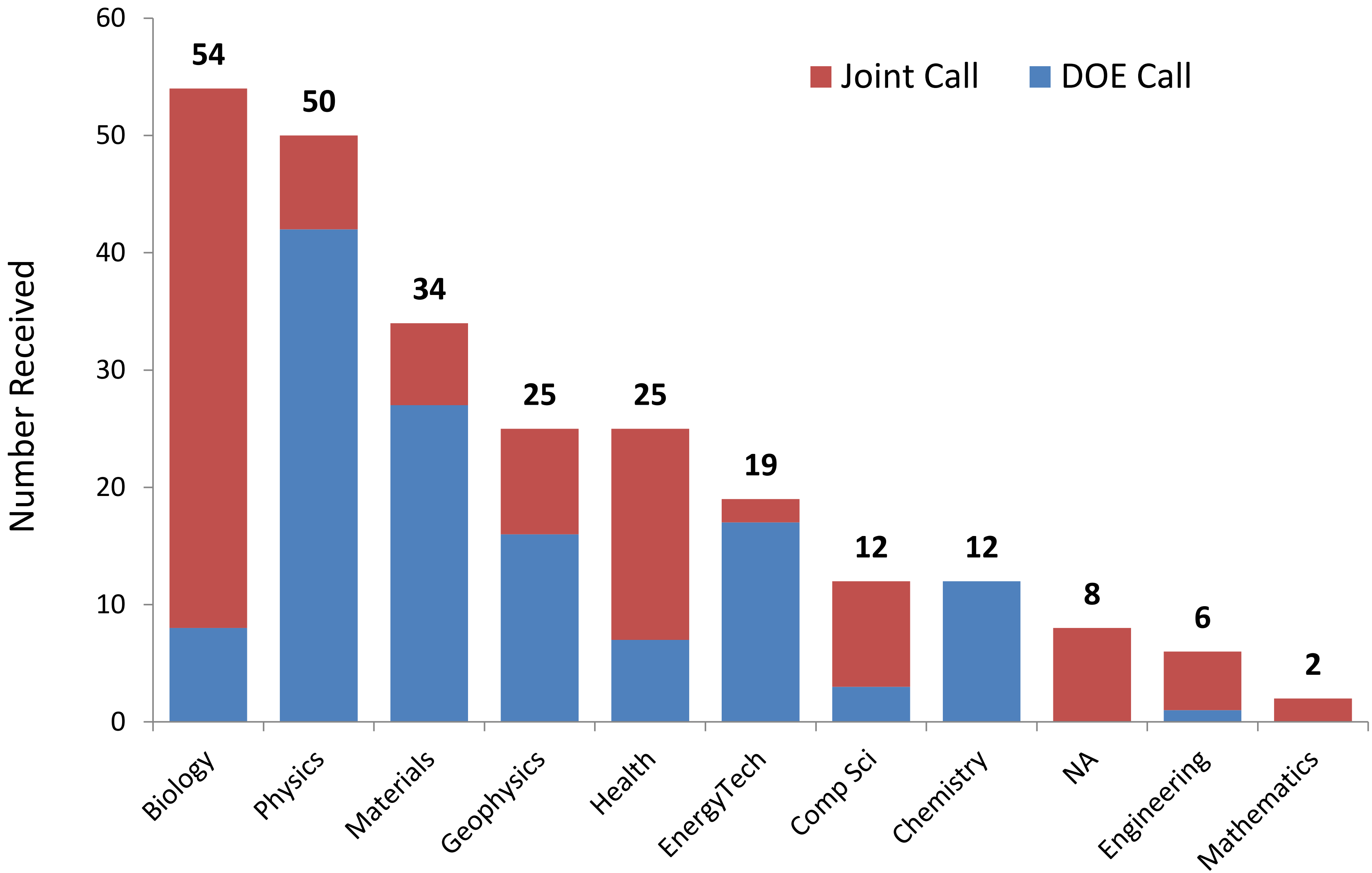
Geoscience & Atmospheric Sciences

Astrophysics & Cosmology

Impacts Highlighted in Responses:

- Assuring National CyberSecurity
- Self Assembly based Nano-manufacturing
- Galaxy formation and extreme gravitational fields
- Optimizing the Power Grid
- Realistic Hypersonic flow for Flight Vehicles
- High Resolution Atmospheric & Climate Models
- N-by-N Comparison of All Patients in US
- Time-sensitive simulations of global scale epidemics

Pre-lim Analysis: Responders Represent Broad Science Spectrum

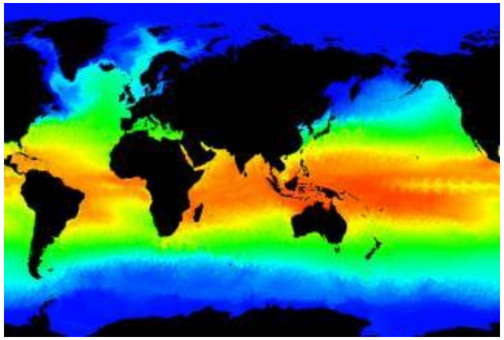


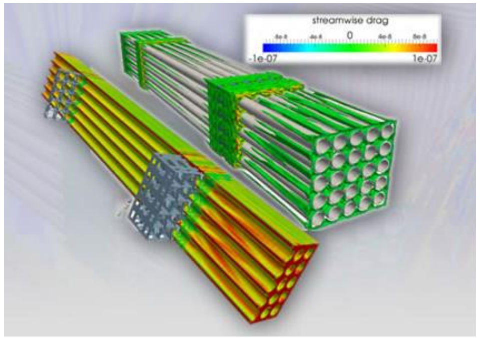
Responses Provide Compelling Science Drivers for Exascale

Climate Challenge Problems: 1km resolution in global simulations, Fully coupled atmosphere-ocean-land-ice models

Impacts:

- Forecast water resources with increased confidence
- Project future changes in severe weather,
- Address food supply change with defensible estimations,
- Inform decision on the resilience of energy and public health infrastructure





Nuclear Energy Challenge Problems: Simulations over full 18-month fuel cycle, LES of reactor core with 10B elements, Explicitly resolve 51K fuel rods in 3D

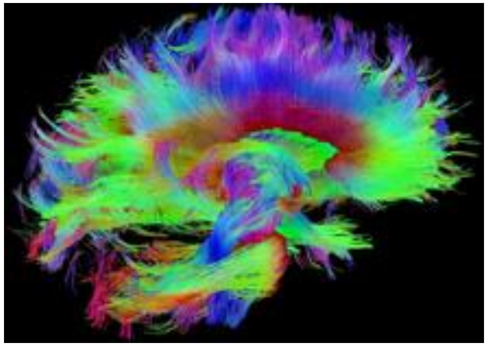
Impacts:

- Predictive capabilities for untested new designs
- Inform engineering design in operational nuclear reactors
- Improve radiation therapy treatments for clinical use

Health Science Challenge Problems: Integrated imaging , genomic and medical data; High-resolution large-scale multi-factor simulations of infectious disease

Impacts:

- Transform patient care with early and accurate diagnosis
- Inform decisions on public health policy and response
- Economic benefit of cost-effective public and personal health



Prelim Anlaysis: Responses Identified Common Barriers and Challenges

Barriers to Exascale

- **Applications**: scaling, using accelerators, lack of HPC ready codes, lack of software containers
- **Workforce development-** scientific professionals who can use highly parallel systems
- **Data workflows -** memory and I/O footprint on future architectures, data protection, data access from multiple sources, data sharing
- **Ease of resource access**

Exascale Data Challenges

- **52 of the 248** responses emphasized data as important component
- **56% of the 51 in Biology and Health**: mostly "Omics" and large scale clinical data analysis

 # Responses Emphasizing Data

Biology (18 of 54)	Geophysics (3 of 25)
Health Science (11 of 25)	NA (3 of 8)
CIS (8 of 12)	Mathematics (1 of 2)
Physics (3 of 50)	Social Sciences (1 of 1)

My Own Summary with Eye to Networking

- **Emerging use models of HPC put demand on network and associated tools**
  - Realtime HPC computing with experiment is growing
  - Incorporation of HPC facility to offload from grid/ grid computing
  - Demand for co-location of large scale experimental and simulated data with compute
  - Demand for tools to make geographically dispersed data usable in large scale studies
    - Finding data in dispersed data repositories
    - Analyzing and co-processing data from dispersed repositories
- **Biology and health sciences represents a burgeoning user community with greater emphasis on particular networking and computing requirements**
  - Cybersecurity for patient data, and human clinical research data
  - Large demand for tools to make geographically dispersed data usable in large scale studies
    - Precision medicine with impacts in important initiatives
      - BRAIN
      - Cancer Moonshot

EXTRA