Report from ASCAC Data Subcommittee

ASCAC meeting
October 31, 2012

POC: Vivek Sarkar, vsarkar@rice.edu

Subcommittee Members

Last name	First name	Affiliation
Chen (*)	Jacqueline	Sandia
Choudhary	Alok	Northwestern U.
Feldman	Stuart	Google
Hendrickson	Bruce	Sandia
Johnson	Chris	U. Utah
Mount	Richard	SLAC
Sarkar (**)	Vivek	Rice U.
White (*)	Victoria	FermiLab
Williams (*)	Dean	LLNL

(*) ASCAC member, (**) Subcommittee chair

Outline

1. Our charge
2. Data challenges in Office of Science's mission
3. Next steps

Our Charge

"... In addition to the challenges and opportunities of exascale computing, the Office of Science is facing related challenges from data-intensive research activities, such as the growing volumes of data generated at our next generation scientific user facilities and by the new genomics-based technologies that are enabling a revolution in systems biology research. ... In order to maximize the return on our limited federal resources, we need to understand the similarities among and differences between these data challenges and the potential to leverage research investments to address issues spanning both exascale and data-intensive science.

By this letter, I am charging the ASCAC to assemble a subcommittee to examine the potential synergies between the challenges of data-intensive science and exascale. The subcommittee should take into account the Department's mission needs, which define the Office of Science's unique role in data-intensive science vis-a-vis other agencies. The subcommittee should specifically address what investments are most likely to positively impact both our exascale goals and our data-intensive science research programs, including data management at our next generation facilities.

I would appreciate the committee's preliminary comments by November 2012 and a final report by March 30,2013. ..."

Interpreting the Charge: Timeframe

- The charge did not specify a timeframe for the subcommittee to focus on ...
- ... however, the focus on synergies with exascale computing suggests that a 10-year timeframe should be our primary focus
- Nearer term (5-year timeframe) also deserves attention as migration path to exascale computing
- Likewise, we do not wish to cut off longer-term technology options (20-year timeframe), but they represent a secondary goal for this study

Interpreting the Charge: Technical Scope

- Scope clearly includes experimental facilities and scientific simulations that exemplify the Office of Science's unique role in data-intensive science

- Upon completion of our study, some of our conclusions may be more broadly applicable e.g., to NNSA, DoD. However, it is important to ensure that the scenarios that we study are centered on the Office of Science's mission needs

- Other data-intensive applications (e.g., cyber defense) were ruled to be out of scope for this study

Desired Outcomes

- Recommendations
  - for identifying synergies between data-intensive and exascale computing in the next 10 years
  - within the context of furthering science
  - while identifying investments that are most likely to positively impact both data-intensive science goals and exascale computing goals
- Detailed study of exemplar scenarios from experimental facilities and scientific simulations to provide context for recommendations
- Discussion of cross-cutting concerns
  - Technology roadmaps
  - Analytics and visualization
  - Software challenges

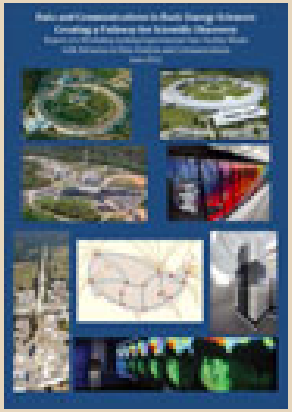
Planned Methodology for Study

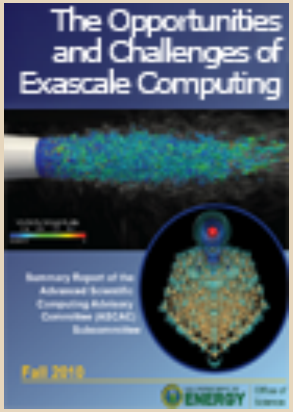
- Build consensus on scope to drive to desired outcomes
- Regular telecons among subcommittee members
  - Already conducted 9 telecons (including calls with Dan Hitchcock and Bill Harrod)
- Division of responsibilities for different aspects of study --- scenarios, roadmaps, cross-cutting concerns
- Detailed discussion of data-intensive scenarios in Office of Science
- Detailed discussion of Office of Science investments in exascale computing and data-intensive computing
- Write-up of draft and final versions of report

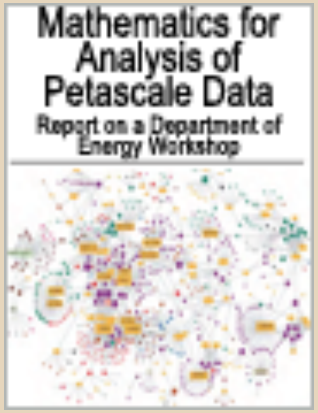
Outline

1. Our charge
2. Data challenges in Office of Science's mission
3. Next steps

Several recent ASCR workshops have focused on exascale & data-intensive computing





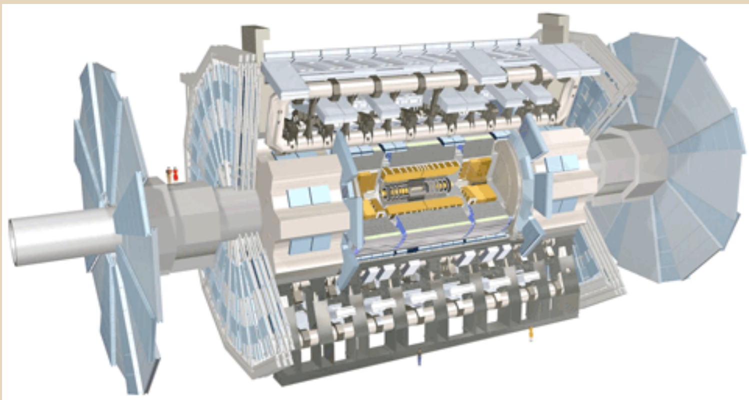


• • •

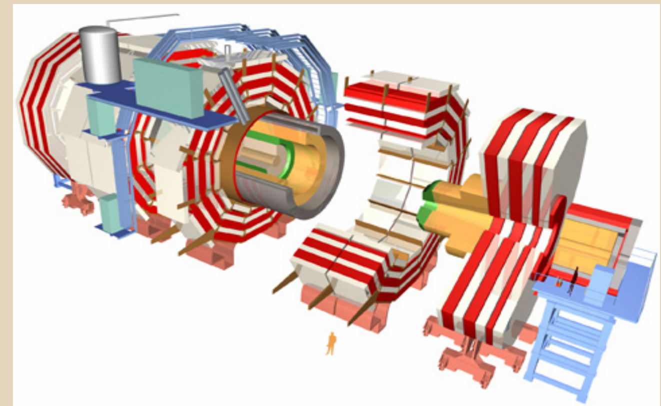
Sources of Data Intensity in Office of Science mission applications

- Large volumes of observational data
  - Necessitates data reduction at data source
- Large volumes of simulation data
  - Necessitates in-situ and in-transit analytics
- Large degrees of data sharing
  - Necessitates coordinated workflows
- We will illustrate the above three categories of data intensity through three *exemplar* use cases
  - High Energy Physics use: LHC experiments
  - Simulation scenario: S3D combustion code
  - Climate scenario: Earth System Grid Federation

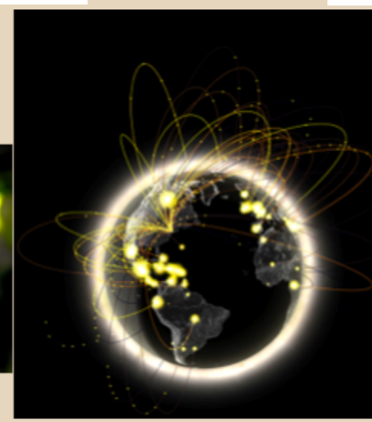
Use Case for Large Volumes of Observational Data: LHC experiments



ATLAS detector



CMS detector



Worldwide LHC Computing Grid (WLCG)

Use Case for Large Volumes of Observational Data (contd)

- ATLAS and CMS detectors generate raw data at rates equivalent to 1PB/s
- These data is reduced in real time to more manageable volumes (~ 1GB/s or 10PB/year)
- Highly efficient networks and Grid technology enables distribution of these data worldwide
- Data volumes are inflated by storing derived data products, replication for safety and efficiency, and the need to store even more simulated data than experimental data. ATLAS currently stores over 100PB, and this is rising rapidly.
- Derived datasets are read-only after creation, thereby offering opportunities for new storage technologies in which reads perform better than writes

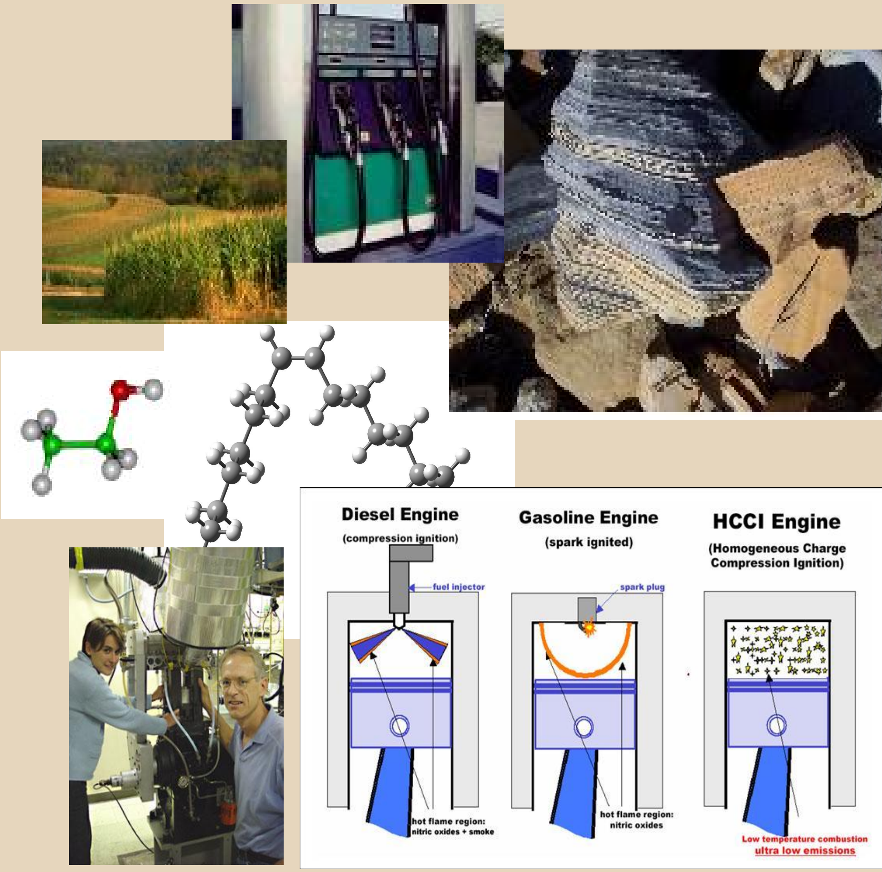
High Energy Physics Exemplar: distribution vs. centralization

- An experimental facility like LHC must be centralized of necessity, but it is feasible, politically essential, and in many ways beneficial, to adopt a distributed approach to computing and storage
  - LHCOPN is an optical private IP network that connects the Tier0 and the Tier1 sites of the LCG
  - Increasingly capable research and education networks worldwide, interconnect distributed LHC computing sites
- Pros of distribution:
  - Access to computing resources around the world.
  - Leveraging worldwide funding and intellect
- Cons of distribution:
  - Data movement and replication overheads
  - Complexity of managing distributed systems

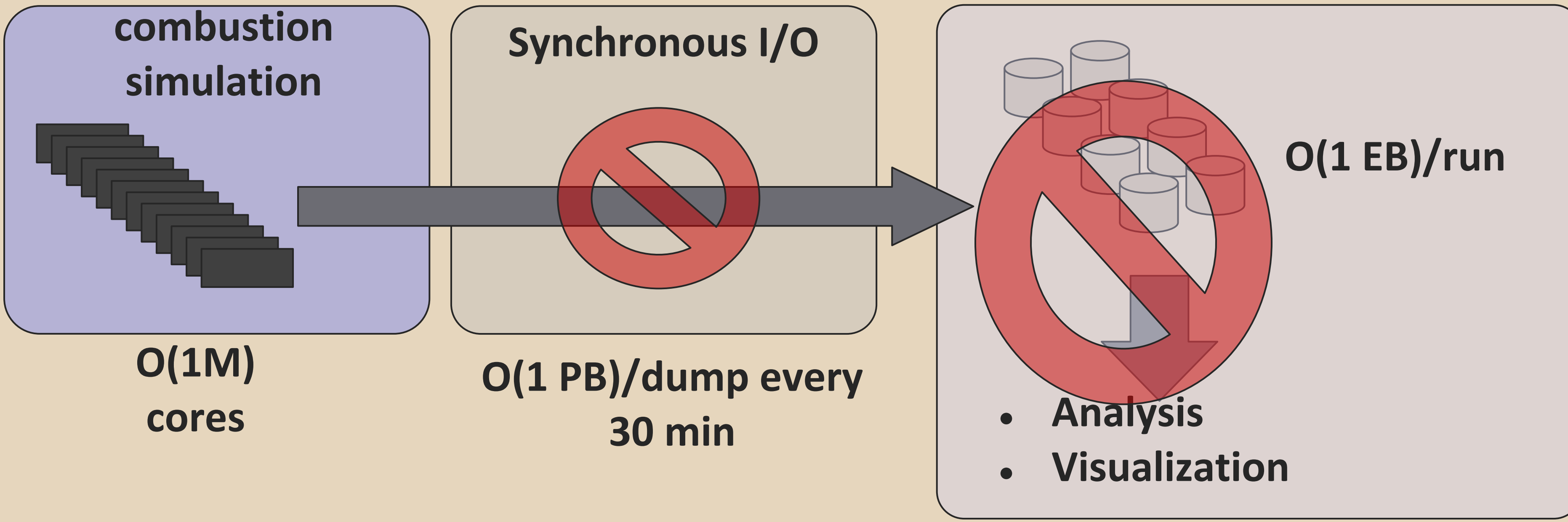
Use Case for Large Volumes of Simulation Data: S3D Combustion code

Goal: simulate turbulence-chemistry interaction at conditions that are representative of realistic systems

- High pressure
- Turbulence intensity
- Turbulent length scales
- Sufficient chemical fidelity to differentiate effects of fuels



Petascale Workflow Model for this application will not work at Exascale

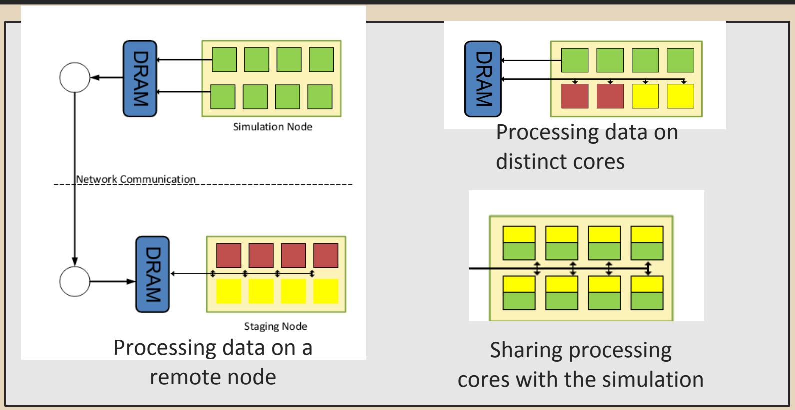


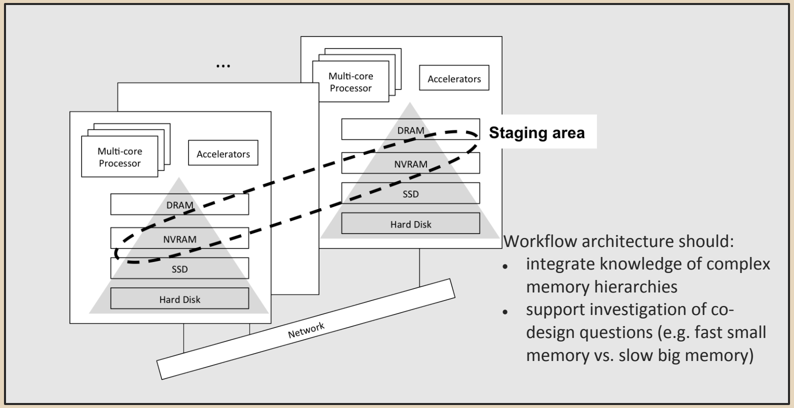
- I/O bandwidth constraints make it infeasible to save all raw simulation data to persistent storage

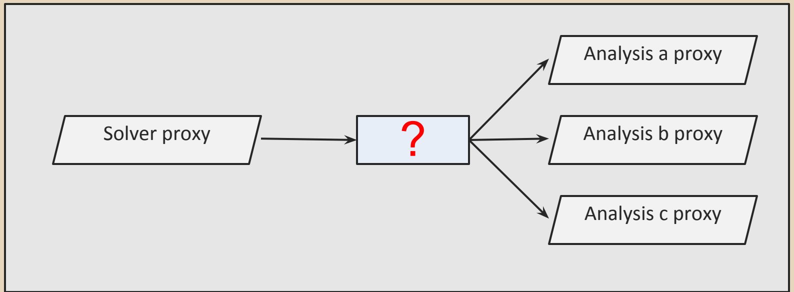
  ==> In-situ and in-transit analyses are a necessity
- Challenge: design a workflow that supports smart placement of analysis and visualization, tracking large graphs, reducing checkpointing size with in-situ analytics

Design Considerations for Data-Intensive Workflows in Simulation Use Case

- Location of the compute resources
  - Same cores as the simulation (in situ)
  - Some smaller number of (dedicated) cores on the same nodes
  - Some dedicated but separate nodes on the machine or an external resource
- Data access, placement, and persistence
  - Direct access to simulation data structures
  - Shared memory access via hand-off / copy
  - Shared memory access via non-volatile near node storage (NVRAM)
  - Data transfer to dedicated nodes or external resources
- Synchronization and scheduling
  - Execute synchronously with simulation every n^{th} simulation time step
  - Execute asynchronously







Use case for Large Degrees of Data Sharing: Bringing Together Large Volumes of Diverse Data at any Velocity to Generate New Insights in Climate Modeling

Data integrating enterprise system

Insight into big data reveals three very significant challenges:

- **Variety: managing complex data**, including storage and retrieval, from multiple regional and non-regional data indices, types and schemas
- **Velocity:** distributing live data streams and **large volume data movement** quickly and efficiently
- **Volume:** analyzing large-volume data (from terabytes to zettabytes) in-place for **big data analytics**

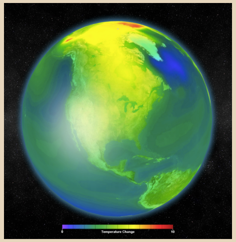
BER and ASCR invests in:

- **Accessing Global Information: Accessing climate data and content information from everywhere via the web, sensors**, and applications in an integrated and federated environment
- **Flexible Infrastructure:** Flexible automated administration, easy-to-use analytics, and virtualization at every level
- **Scalable Framework:** Big data analytics in a scalable environment with efficient parallelism, workload-optimization, and **real-time streaming process**

Simulation

Observation

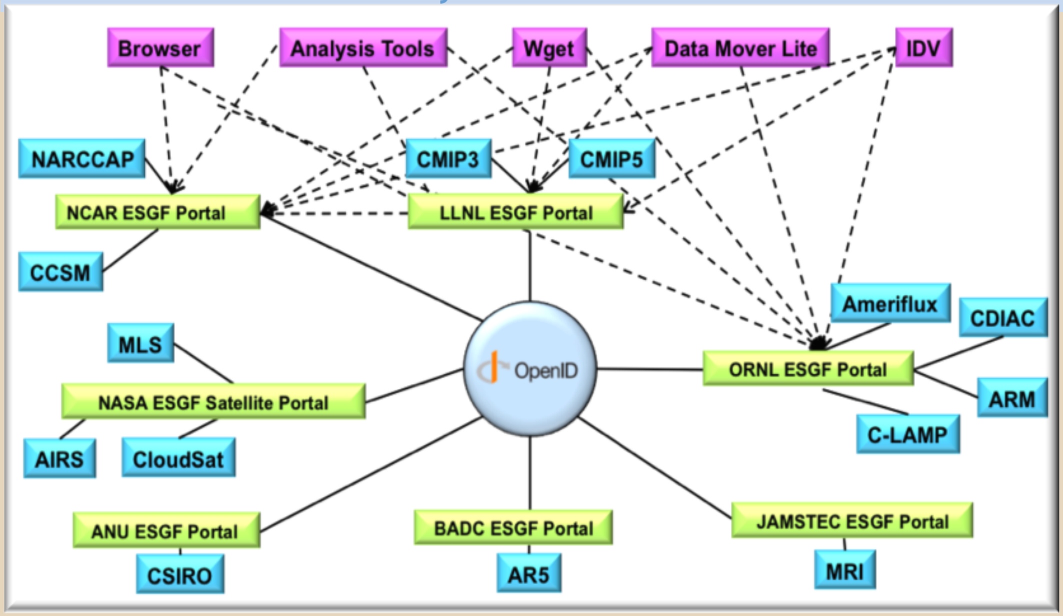
Reanalysis



The ESGF distributed data archival and retrieval system

- Distributed and federated architecture
- Support discipline specific portals
- Support browser-based and direct client access
- Single Sign-on
- Automated script and GUI-based publication tools
- Full support for data aggregations
  - A collection of files, usually ordered by simulation time, that can be treated as a single file for purposes of data access, computation, and visualization
- User notification service
  - Users can choose to be notified when a data set has been modified

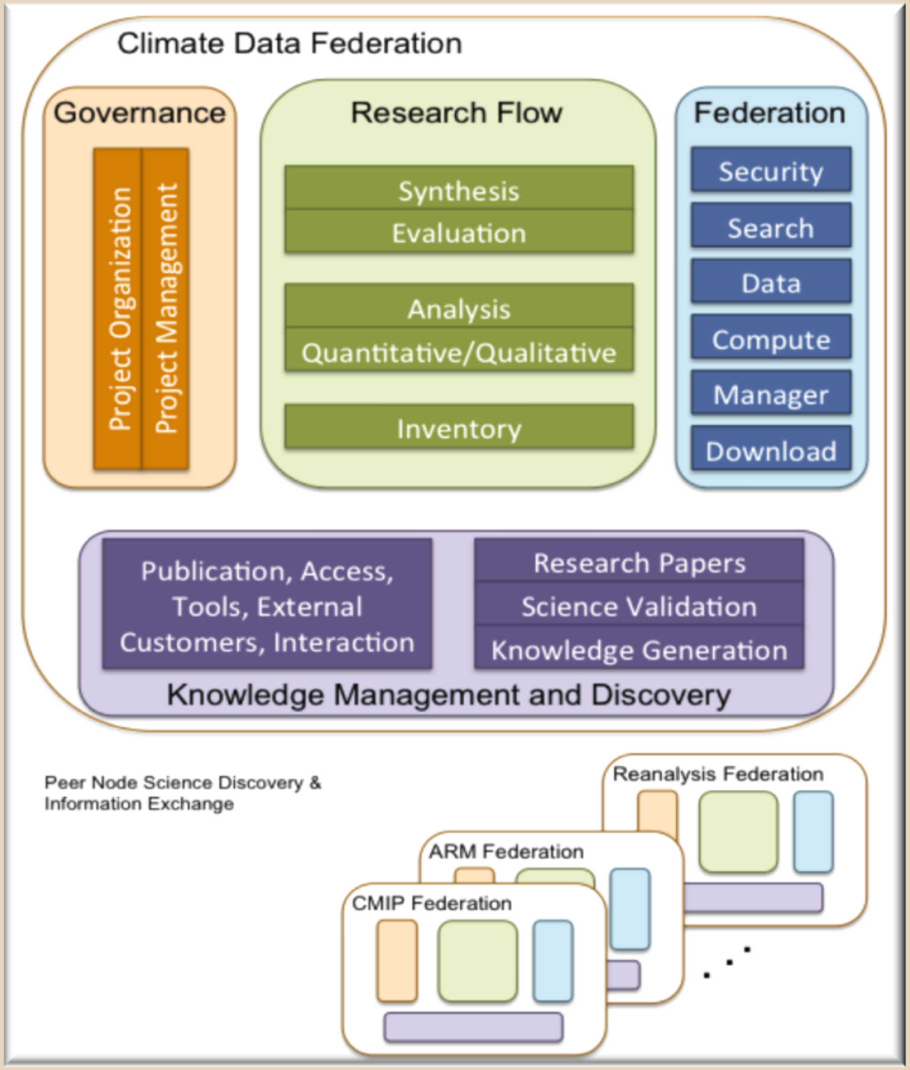
ESGF
Earth System Grid Federation



Large amounts of data movement and management are needed for the complex and diverse climate data community.

ESGF software system integrates data federation services

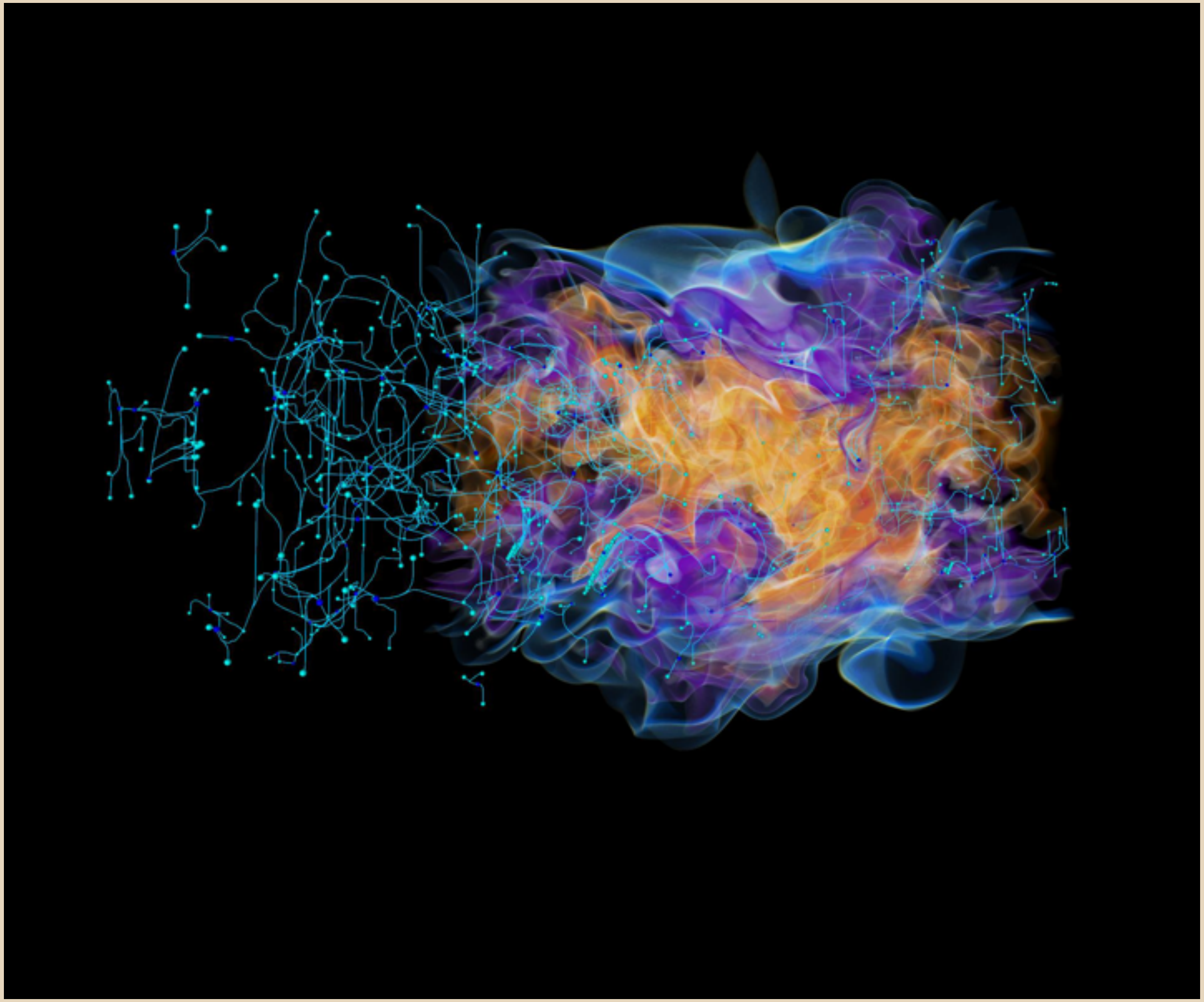
- NetCDF Climate and Forecast (CF) Metadata Convention
  - (LibCF)
  - Mosaic
- Climate Model Output Rewriter 2 (CMOR-2)
- Regriders: GRIDSPEC, SCRIP, & ESMF
- Publishing
- Search & Discovery
- Replication and Transport
  - GridFTP, OPeNDAP, DML, Globus Online, ftp, BeSTMan (HPSS)
  - Networks
- Data Reference Syntax (DRS)
- Common Information Model (CIM)
- Quality Control
  - QC Level 1, QC Level 2, QC Level 3, Digital Object Identifiers (DOIs)
- Websites and Web Portal Development
  - Data, Metadata, Journal Publication Application
- Notifications, Monitoring, Metrics
- Security
- Product Services
  - Live Access Server, UV-CDAT



Cross-cutting Issue: Upcoming Technology Disruptions related to Big Data and HPC

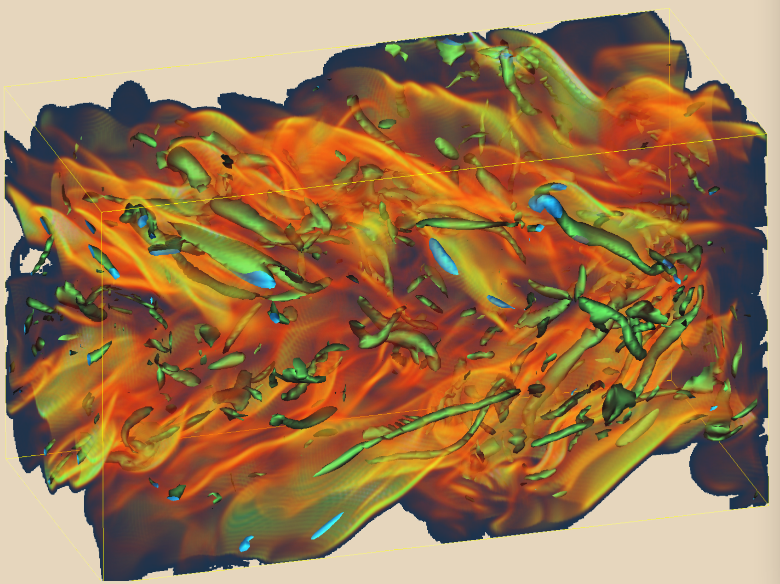
- Solid state memory will scale with Moore’s law and provide higher-performing / lower-energy storage layer than disk
  - Will support new architectural options for data-intensive computing
- Stacked memory will dramatically improve performance and energy utilization
  - Will support new algorithmic paradigms for moving computing to memory
- Silicon photonics will significantly increase bandwidth and reduce power needed to connect to optical networks

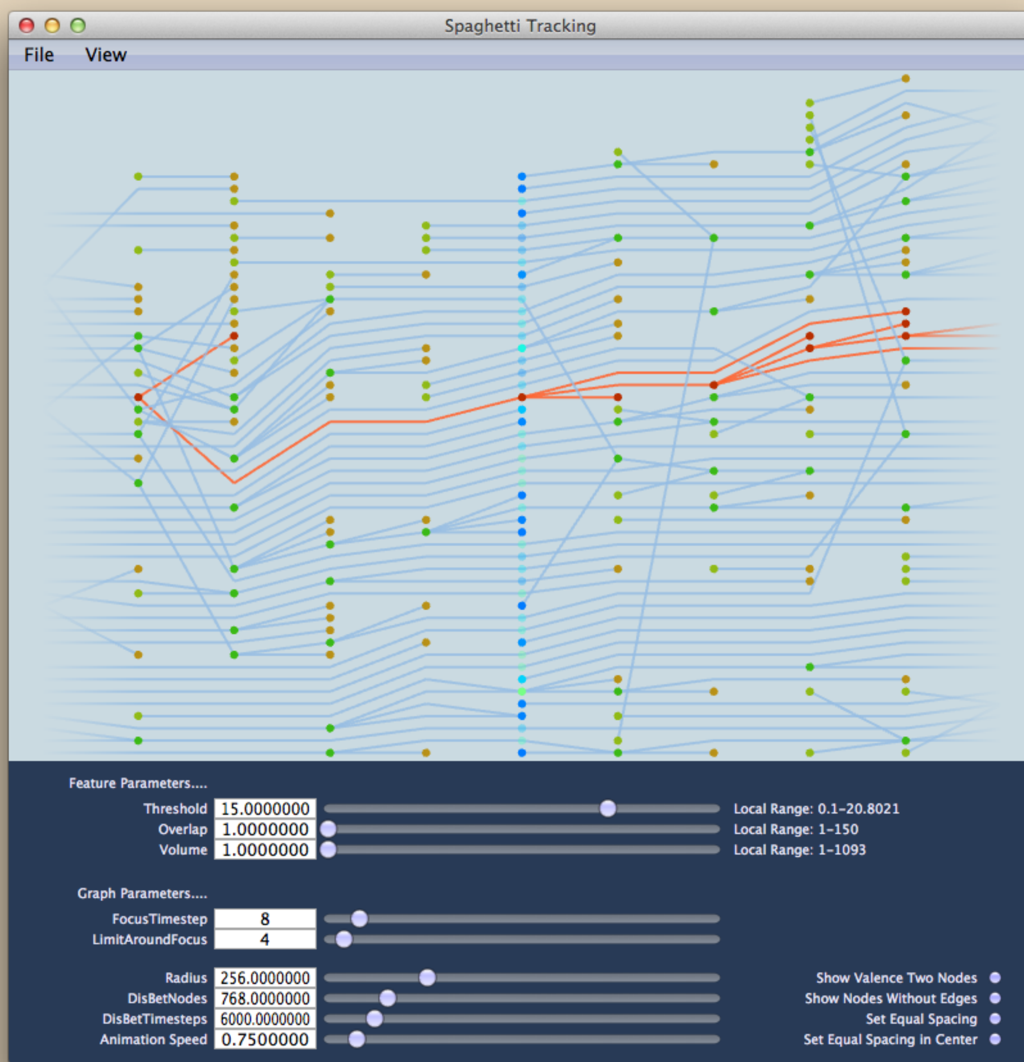
Cross-cutting issue: from Scientific Visualization to Information Visualization



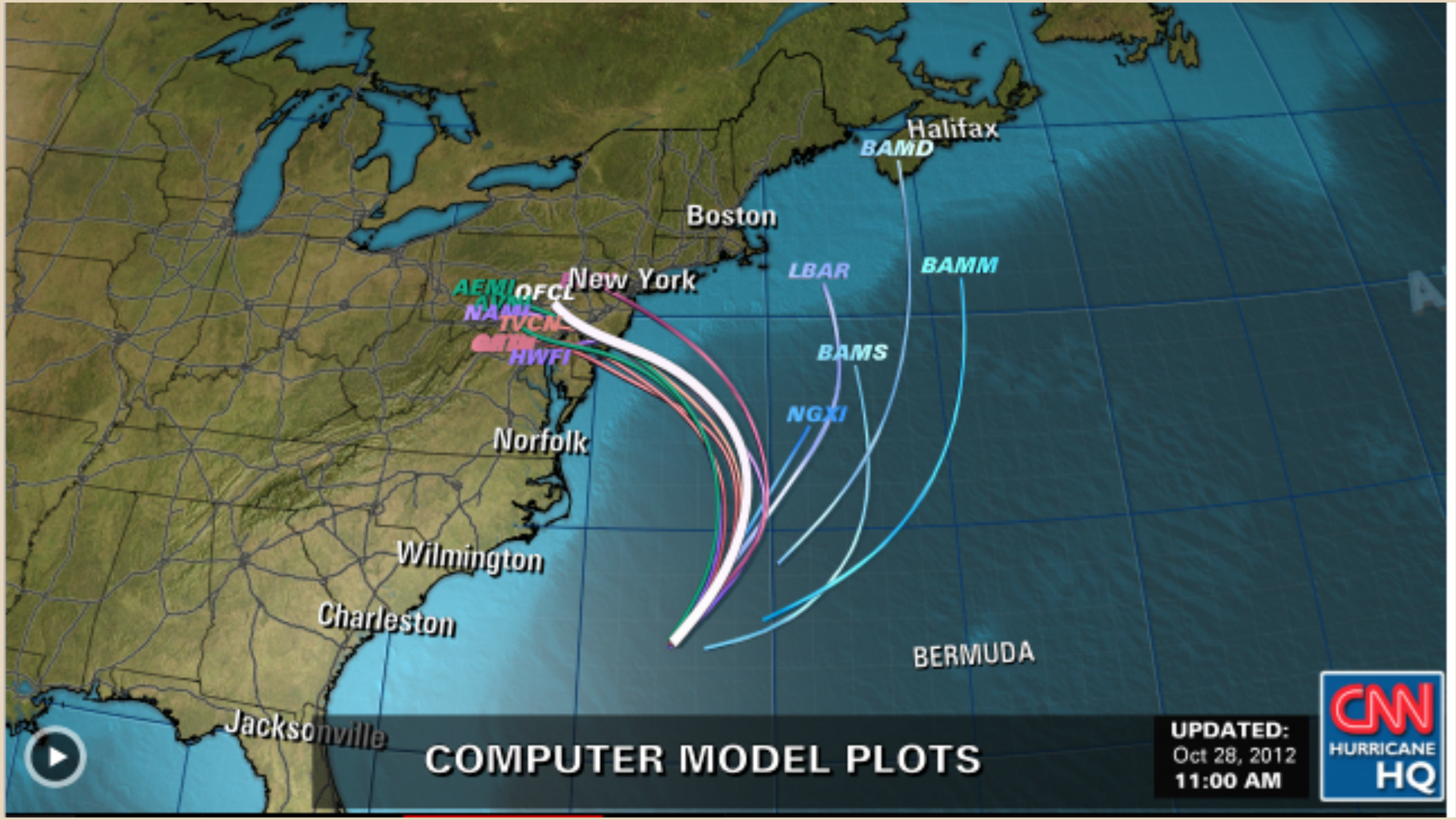
Topological Analysis of Massive Combustion Simulations

- Non-premixed DNS combustion (J. Chen, SNL): Analysis of the time evolution of extinction and reignition regions for the design of better fuels





Impact of Data Capability and Uncertainty Quantification on Predictive Science



Cross-cutting issue: software challenges related to data-intensive science

- Human cost will be a major bottleneck
  - Science codes need to be developed and understood by large numbers of developers, and maintained for ~ 20 years, so frequent rewrites for new data platforms are not an option
- Data reduction and analytics codes are often branch-intensive and do not make good use of GPUs or SIMD vector parallelism
  - Need proxy apps for these codes
  - These codes will also run into severe memory constraints on future extreme scale platforms
- Use of parallel I/O frameworks requires significant expertise
  - State of the art now uses a 3-phase data transfer model

Outline

1. Our charge

2. Data challenges in Office of Science's mission

3. Next steps

Timeline

Past

- Panel formed --- September 2012
- Kick-off telecon --- September 20, 2012
- Additional telecons to date: 9/25, 9/27, 10/4, 10/9, 10/11, 10/23, 10/25
- All-day (virtual) meeting: 10/29

Present

- Interim update to ASCAC: 10/31

Future

- Additional information gathering : Nov-Dec 2012
- Start of report writing: January 2013
- Draft report for feedback from ASCAC: February 2013
- Final report: March 2013

Next Steps

- Additional information gathering
  - Additional scientific domains and scenarios e.g.,
    - Next Generation Light Sources
    - Biology Scientific Grand Challenge Workshop
    - ASCR/BES Data Workshop
  - Supercomputer center costs/requirements for storage
  - Current ASCR investments related to Big Data and Exascale
- Initial draft of report
- Discussion of feedback on report
- Final report

Conclusions

- Many research activities in the Office of Science are or are becoming data-intensive
- Natural interplay between exascale computing and data-intensive computing
  - Synergy among requirements e.g., minimizing data movement
  - Data intensity is changing the nature of exascale applications e.g., in-situ analytics
  - Cross-cutting concerns: technology roadmaps, visualization challenges, software challenges
- Opportunity to make research investments that positively impact both data-intensive science goals and exascale computing goals