SDAV
Scalable Data Management, Analysis, and Visualization

The Scalable Data Management, Analysis, and Visualization Institute (SDAV)

Arie Shoshani, institute director
Lawrence Berkeley National Laboratory

http://sdav-scidac.org/

SciDAC-3 PI meeting
September 2012

SDAV partners

Laboratories	Universities
ANL	GTech
LBNL	NCSU
LLNL	NWU
ORNL	OSU
LANL	UCD
SNL	Rutgers
Kitware (Industry)	UUtah

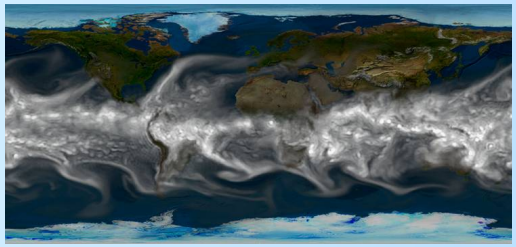
- Constituents:
  - SciDAC (1, 2) Scientific Data Management Center (Shoshani)
  - (2) Visualization and Analytics Center for Enabling Technologies (Bethel)
  - SciDAC (2) Institute for Ultrascale SciDAC Visualization (Ma)
  - New: LANL, Kitware, …
- Portfolio: ongoing relationships with science teams, technologies, expertise

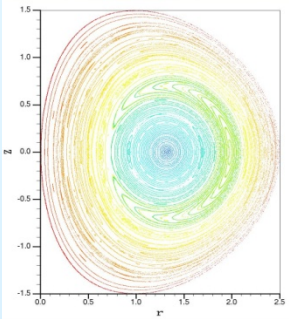
Outline

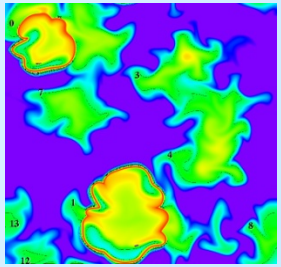
- Emerging challenges with Big Data in scientific domains
- Examples of current approaches and solutions
- Brief discussion on handling of experimental data (future challenges)

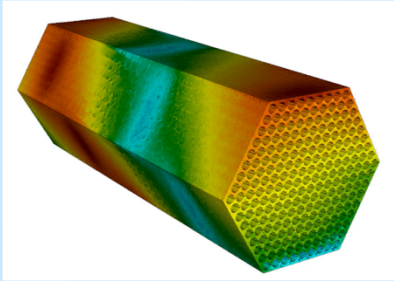
Scientific Data Management, Analysis, and Visualization means:

- Target Scientific Applications
  - Climate
  - Combustion
  - Fusion
  - Nuclear
  - Accelerator design
  - Cosmology







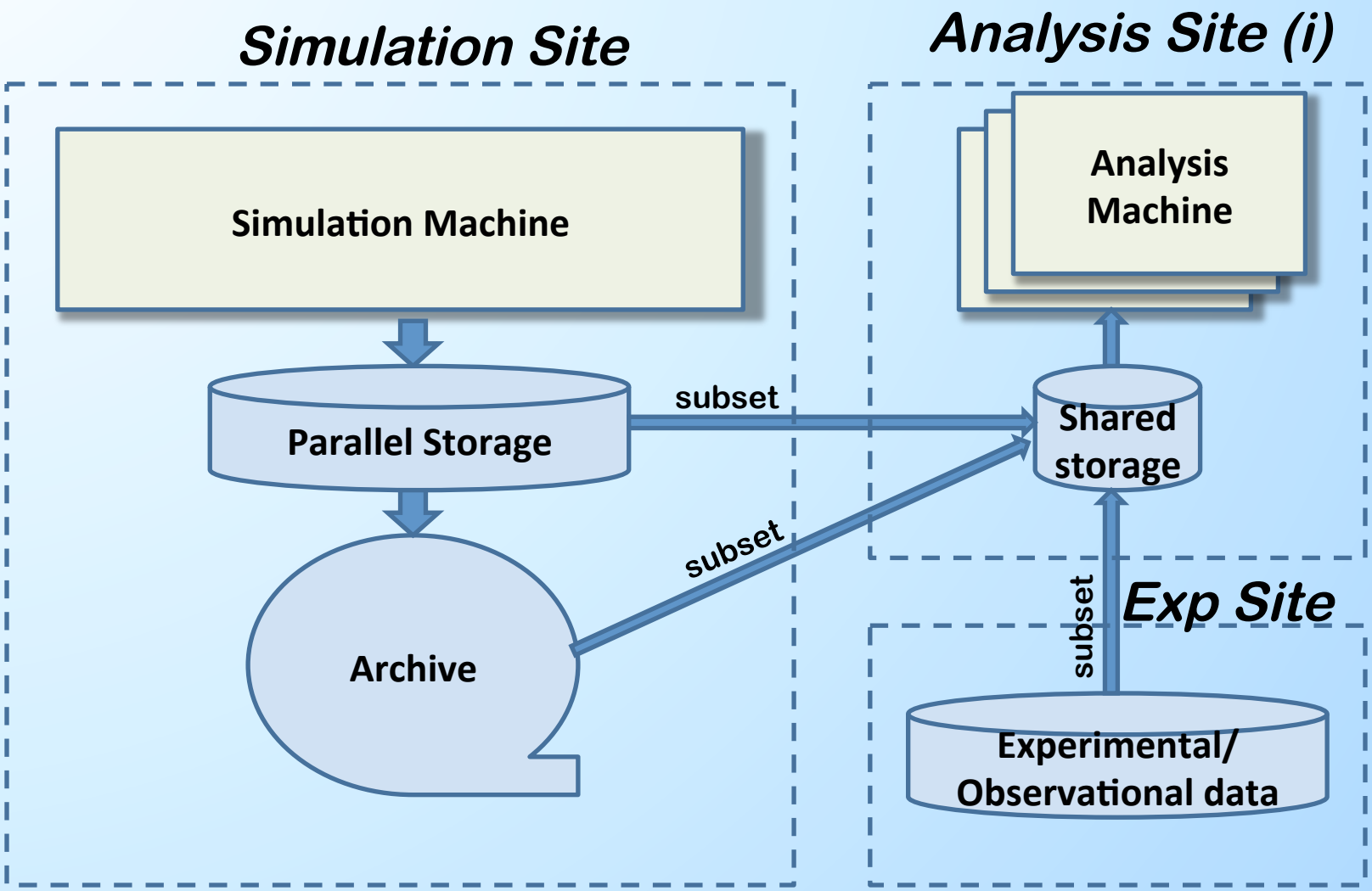


- Algorithms, techniques, and software
  - Representing scientific data – data models, metadata
  - Managing I/O – methods for removing I/O bottleneck
  - Accelerating efficiency of access – data structures, indexing
  - Facilitating data analysis – data manipulations for finding patterns and meaning in the data
  - Visual analytics – help understand data visually

A Typical Scientific Investigation Process

- Current practice – data intensive tasks
  - Runs large-scale simulations on large supercomputers
  - Dump data on parallel disk systems
  - Export some of the data to archives
  - Move data to users' sites – usually selected subsets
  - Perform data manipulations and analysis on mid-size clusters
  - Collect experimental / observational data
  - Move experimental / observational data to analysis sites
  - Perform comparison of experimental/observational to validate simulations
  - Iterate

A typical Scientific Investigation
lots of Data Movement (GBs – TBs)



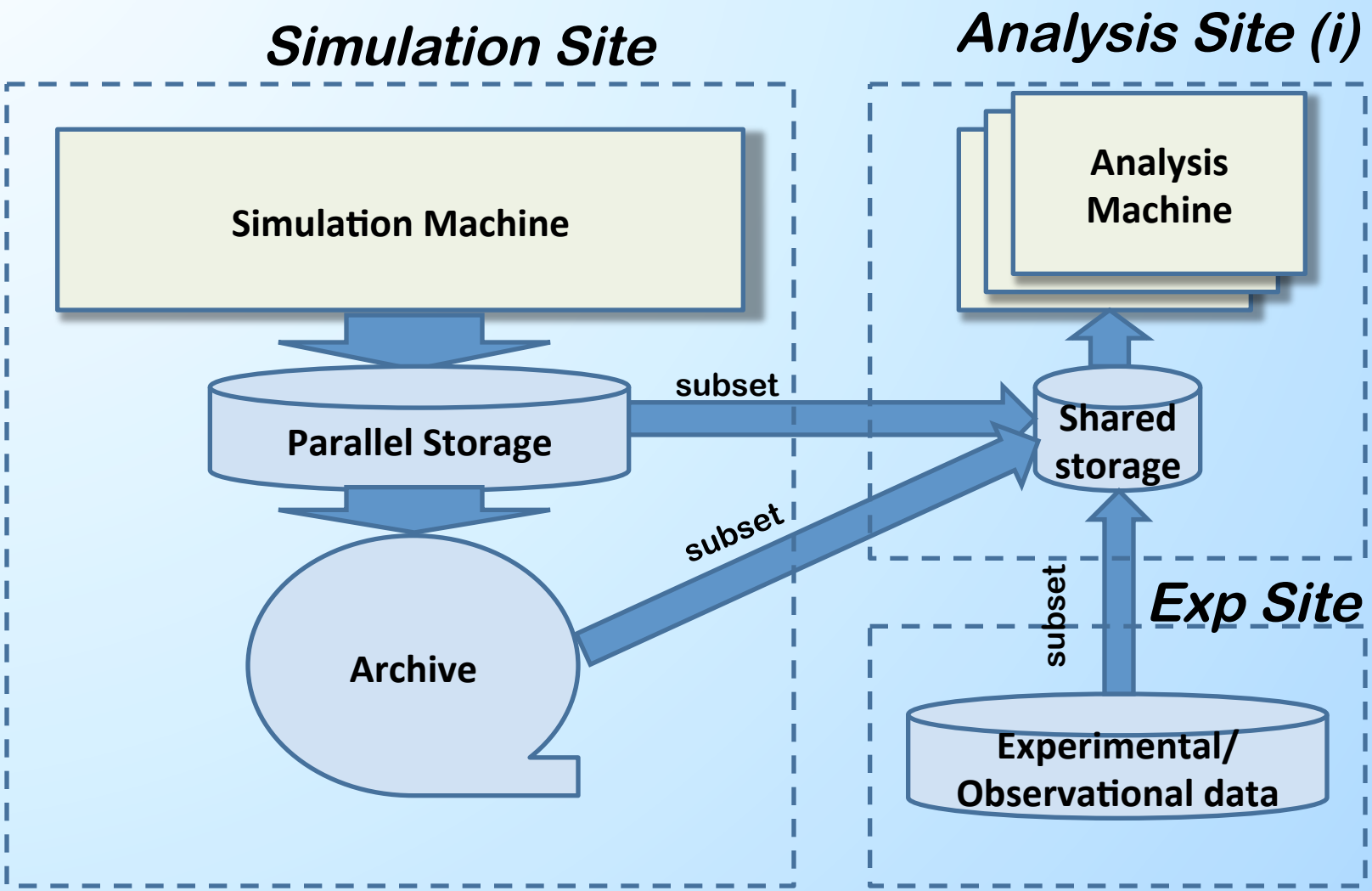
SDAV
Scalable Data Management, Analysis, and Visualization

Exascale Systems: Potential Architecture

Systems	2009	2018	Difference
System Peak	2 Pflop/sec	1 Eflop/sec	O(1000)
Power	6 Mwatt	20 Mwatt	
System Memory	0.3 Pbytes	32-64 Pbytes	O(100)
Node Compute	125 Gflop/sec	1-15 Tflop/sec	O(10-100)
Node Memory BW	25 Gbytes/sec	2-4 Tbytes/sec	O(100)
Node Concurrency	12	O(1-10K)	O(100-1000)
Total Node Interconnect BW	3.5 Gbytes/sec	200-400 Gbytes/sec	O(100)
System Size (Nodes)	18,700	O(100,000-1M)	O(10-100)
Total Concurrency	225,000	O(1 billion)	O(10,000)
Storage	**15 Pbytes**	**500-1000 Pbytes**	**O(10-100)**
I/O	**0.2 Tbytes/sec**	**60 Tbytes/sec**	**O(100)**
MTTI	Days	O(1 day)	

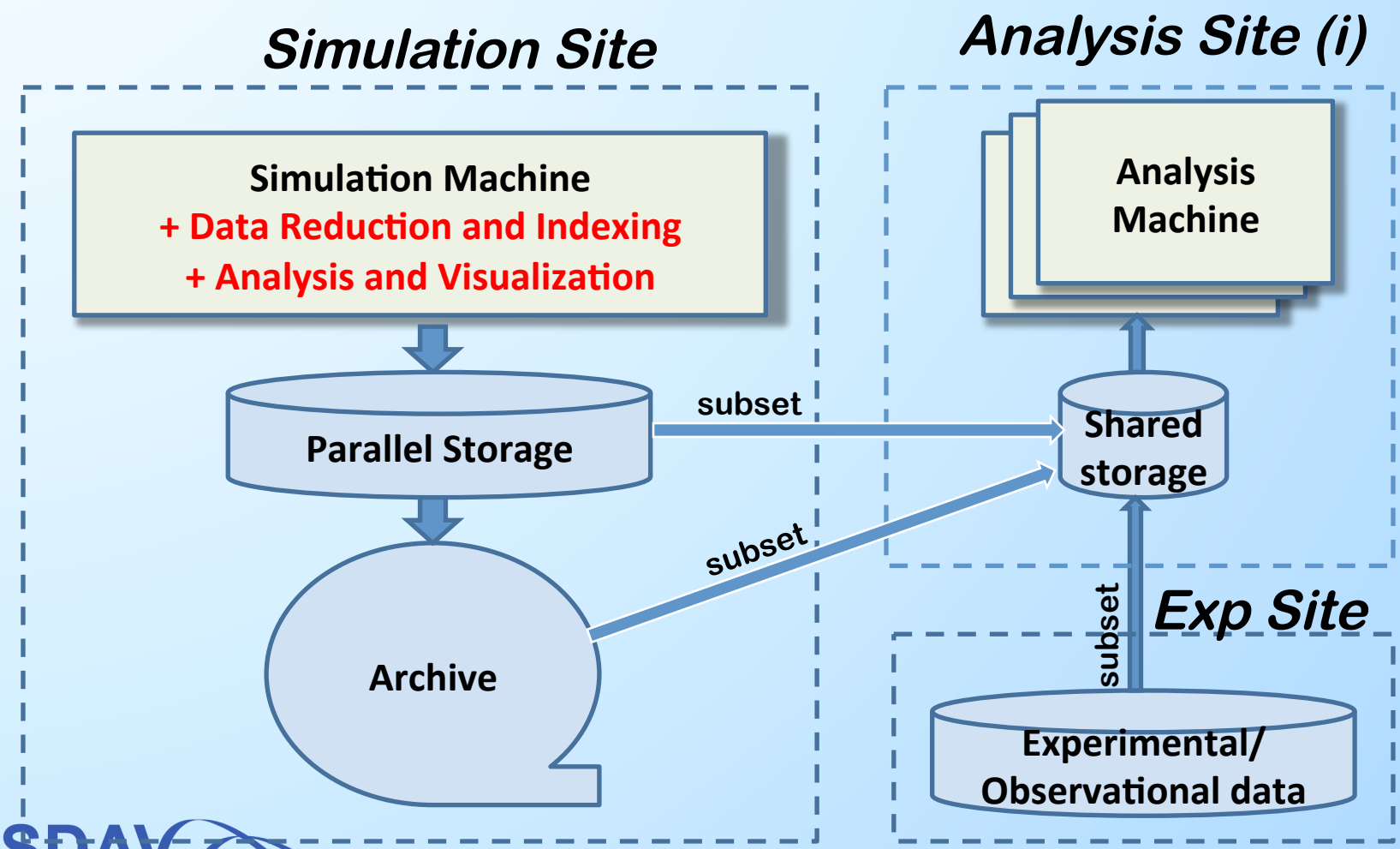
From J. Dongarra, "Impact of Architecture and Technology for Extreme Scale on Software and Algorithm Design," Cross-cutting Technologies for Computing at the Exascale, February 2-5, 2010.

Scaling simulations generates a data volume challenge (PBs)

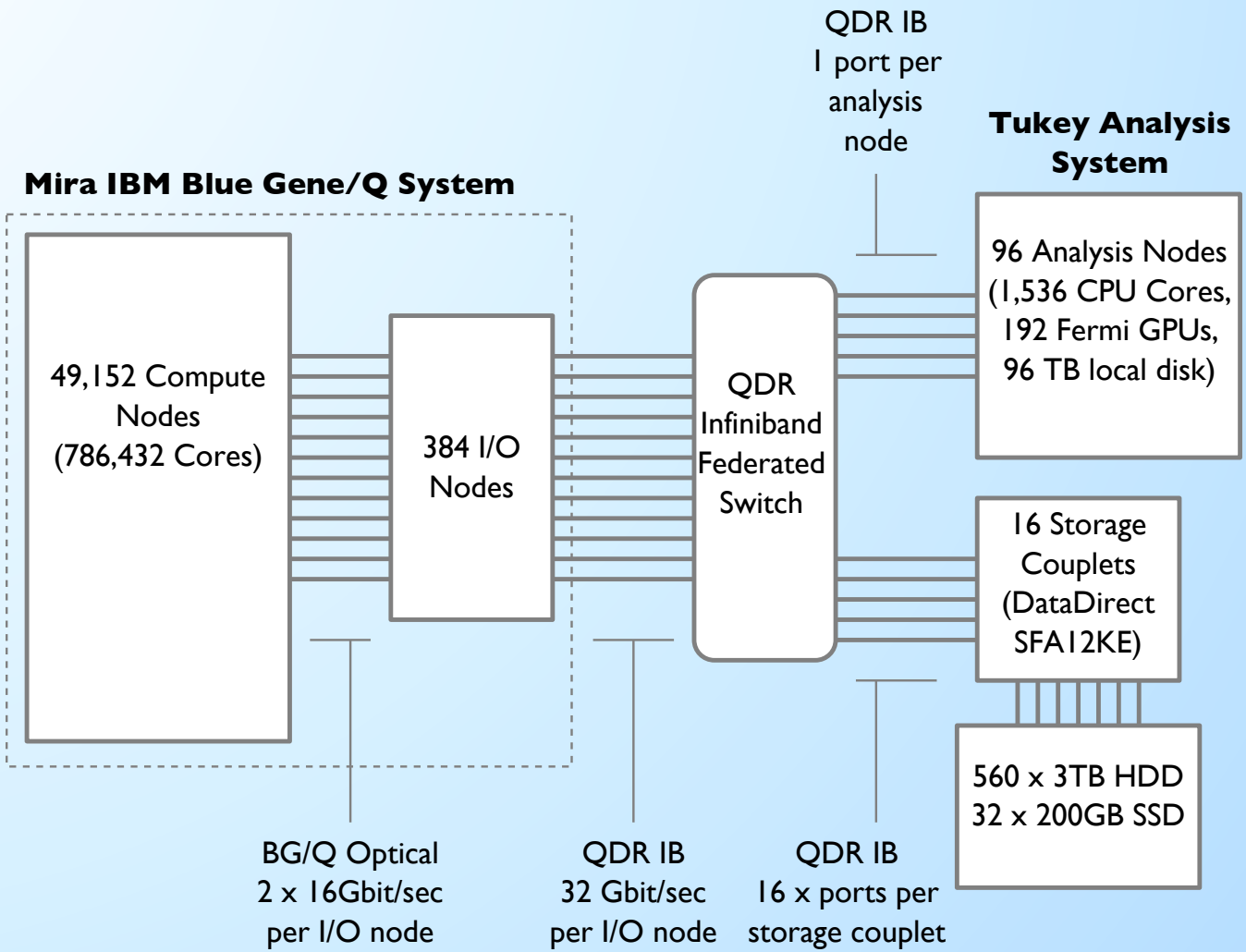


What Can be Done?

- Perform some data analysis and visualization on simulation machine (in-situ)
- Reduce Data and prepare data for further analysis (in-situ)

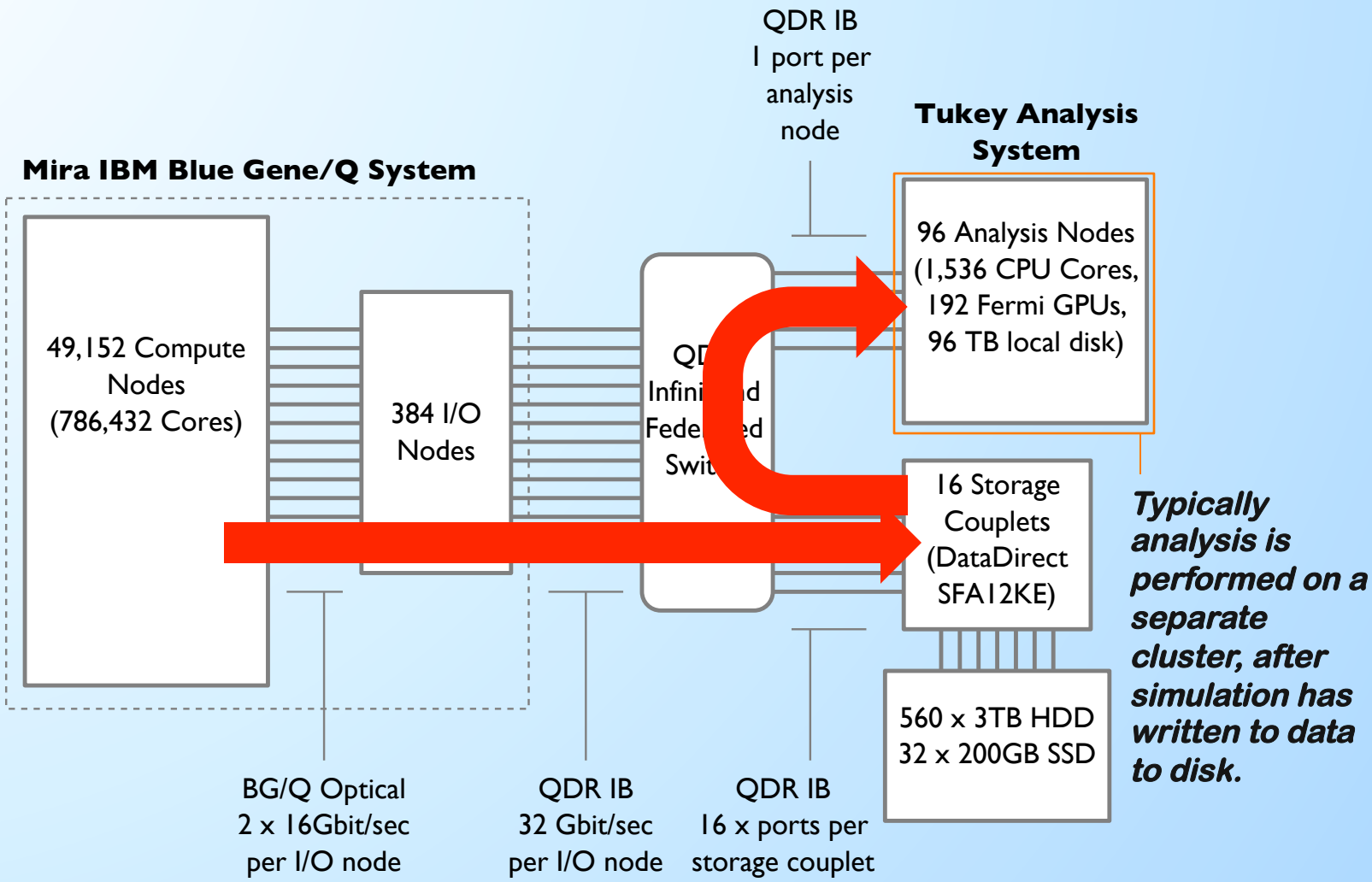


Leadership System Architectures



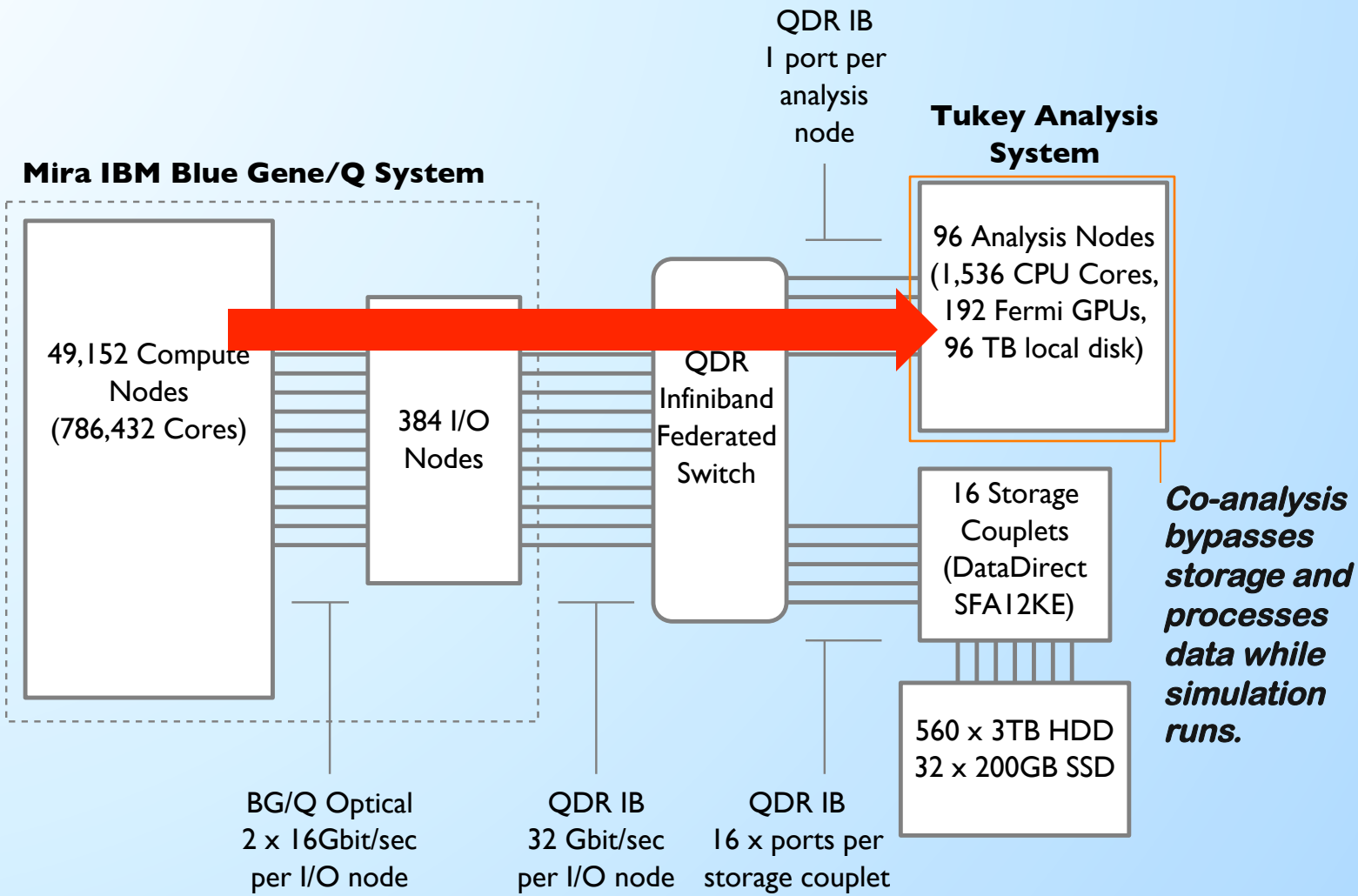
High-level diagram of 10 Pflop IBM Blue Gene/Q system at Argonne Leadership Computing Facility

Analyzing Data: Traditional Post-Processing



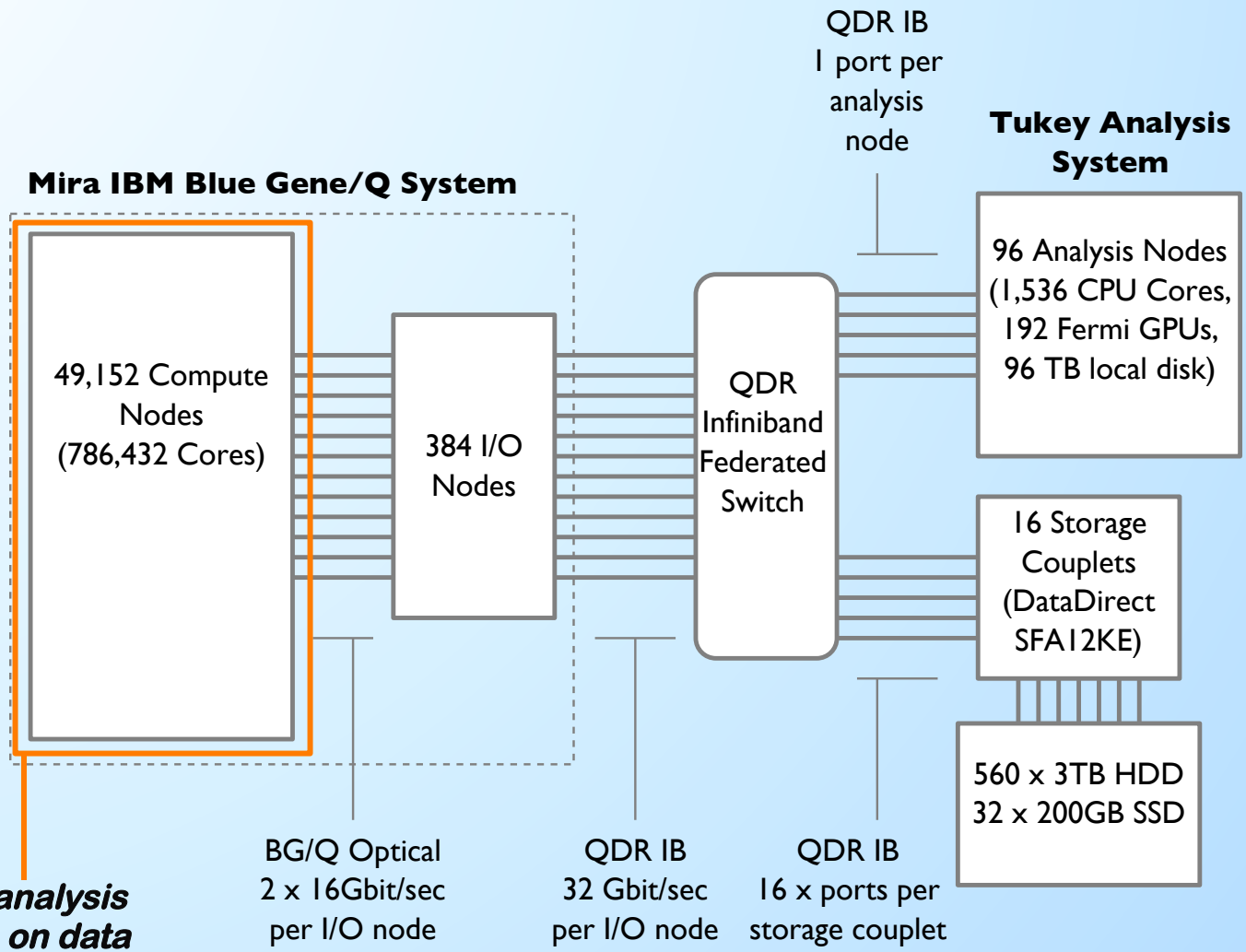
Typically analysis is performed on a separate cluster, after simulation has written to data to disk.

Analyzing Data: Co-Analysis



Co-analysis bypasses storage and processes data while simulation runs.

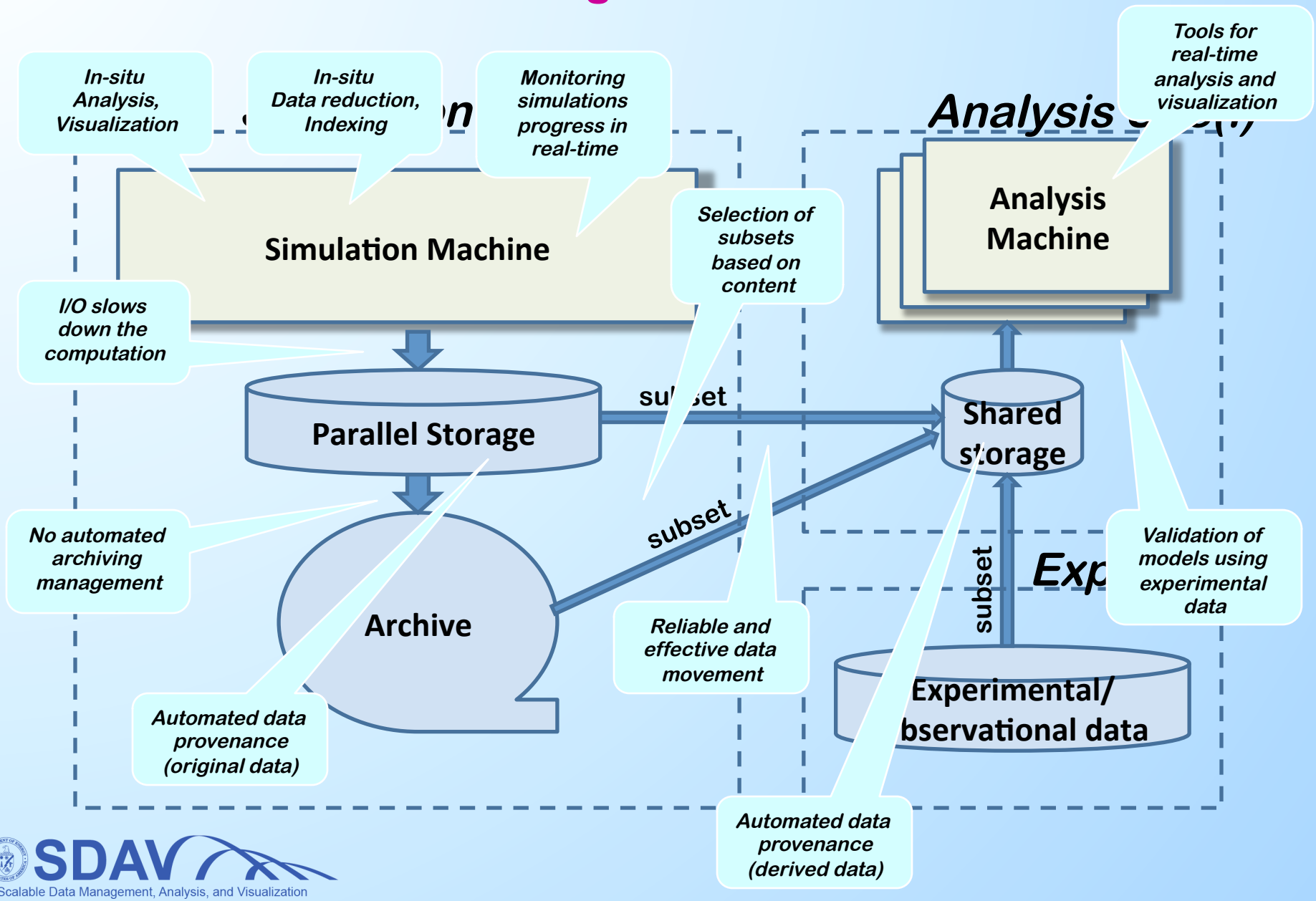
Analyzing Data: In Situ Analysis



"In situ" analysis operates on data before it leaves the compute nodes.

SDAV
Scalable Data Management, Analysis, and Visualization

Bottlenecks/Problems for handling Big Data



SDAV
Scalable Data Management, Analysis, and Visualization

Some Solutions are Emerging

(1) In-situ Analysis and Visualization

(2) In-situ Analysis Data reduction and Indexing

(3) Monitoring simulations progress in real-time

(4) I/O slows down the computation

(5) Automated archival management

(6) Selection of subsets based on content

(7) Selective, reliable, and effective data movement

(8) Tools for real-time analysis and visualization

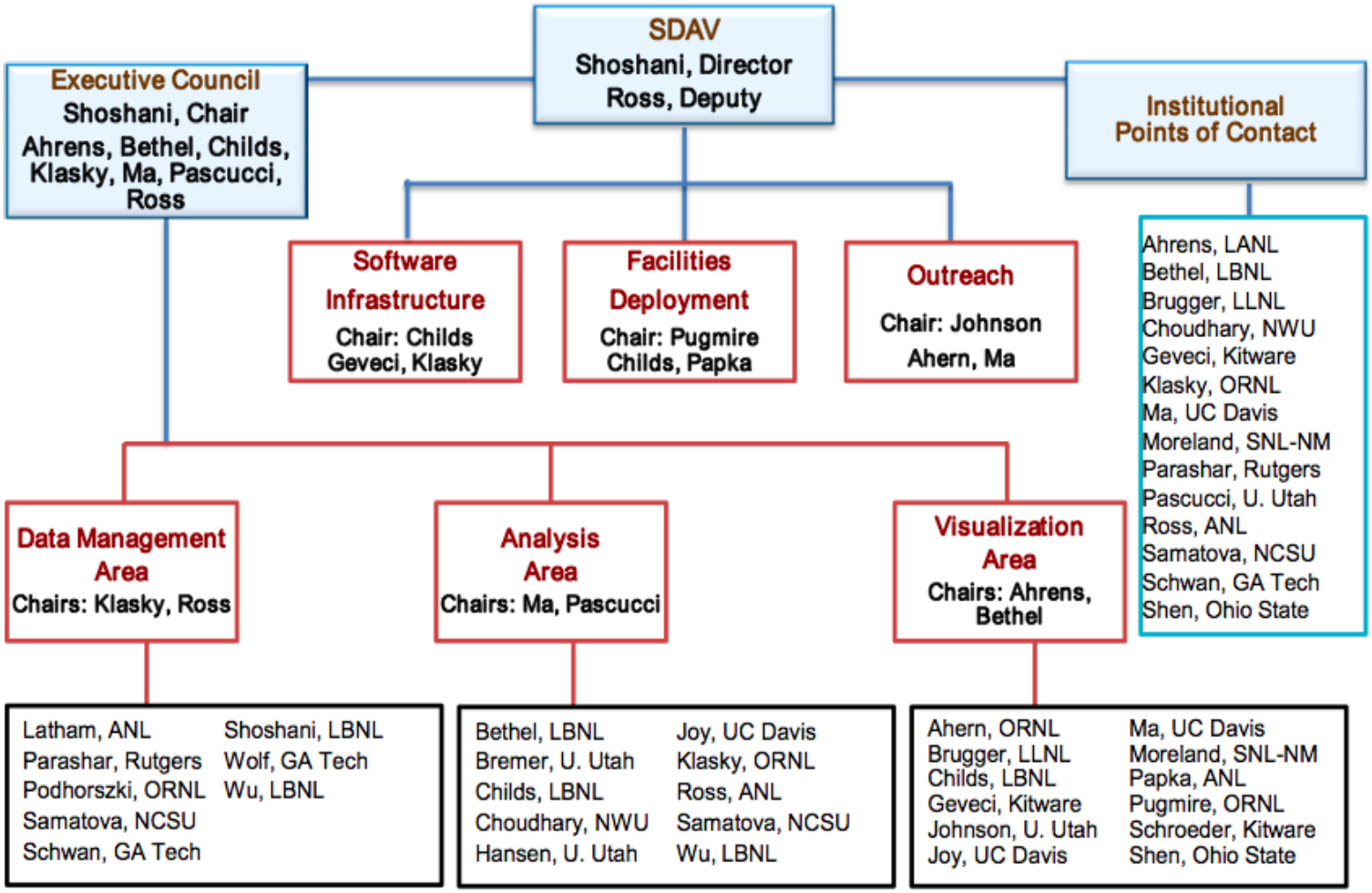
(9) Validation of models using experimental data (some)

(10) Automated data provenance (some)

Data Analysis

- Two fundamental aspects
  - Pattern matching : Perform analysis tasks for finding known or expected patterns
  - Pattern discovery: Iterative exploratory analysis processes of looking for unknown patterns or features in the data
- Ideas for the exascale
  - Perform pattern matching tasks in the simulation machine
    - “In situ” analysis
  - Prepare data for pattern discovery on the simulation machine, and perform analysis on mid-size analysis machine
    - “In-transit” data preparation
    - “Off-line” data analysis

SDAV Organization



The SDAV institute tools

SDAV Goals:

- to actively work with application teams to assist them in achieving breakthrough science
- to provide technical solutions in the data management, analysis, and visualization regimes that are broadly used by the computational science community running on Leadership Class Machines
- To use existing robust tools to the extent possible and develop/adapt tools on an as-needed basis

Data Management tools

- In Situ Processing and Code Coupling
  - ADIOS
  - Glean
- Indexing
  - FastBit
- In Situ Data Compression
  - ISABELLA
- Parallel I/O and File Formats
  - PnetCDF
  - BP-files
  - HDF5

Data Analysis tools

- Statistical and Data Mining Techniques
  - NU-Minebench
- Importance-Driven Analysis Techniques
  - Domain-Knowledge Directed
  - Geometry Based
- Topological Methods
  - In Situ Topology
  - Feature-Based Analysis
  - High-Dimensional Analysis

Visualization tools

- Parallel visualization
  - Visit
  - ParaView
- VTK-m framework
- Flow Visualization Methods
- Rendering
- Ensembles, Uncertainty, and Higher-Dimensional Methods

SDAV Technology Use in Leadership Applications

Application	Code	Contact	Allocation (M node hours)	SDAV Technologies
Astrophysics	Chimera	T. Mezzacappa	60	ADIOS, VisIt, Ultravis-V
Astrophysics	FLASH	D. Lamb	80	PnetCDF, GLEAN, ROMIO, VisIt, VTK
Astrophysics	Maestro	J. Bell	50	VisIt
Astrophysics	Enzo	M. Norman	35	ParaView, VisIt
Biology	Nektar	G. Karniadakis	50	ParaView
Climate	POP	P. Jones	110	PnetCDF, ParaView, ROMIO
Combustion	S3D	J. Chen	60	ADIOS, Dataspaces, Ultravis-V, Ultravis-P, ViSUS IDX, Topologika
Combustion	Boxlib	J. Bell	60	VisIt, ADIOS, Topologika
Combustion	Nek5000	C. Frouzakis	150	VisIt
Cosmology	HACC	S. Habib	150	ParaView, ROMIO, Ultravis-P
Fusion	GTC	Z. Lin	35	ADIOS, DataTap, FastBit, Ultravis-V
Fusion	XGC	C.S. Chang	50	ADIOS, Dataspaces, FastBit, Ultravis-V, VTK
Fusion	GTC-P	W. Tang	58	ADIOS, Ultravis-V, Ultravis-P
Plasma	VPIC	B. Daughton	30	PnetCDF, ParaView, ROMIO
Nuclear	Nek5000	P. Fischer	25	ROMIO, VisIt

SDAV: 8 posters

5 APPLICATION POSTERS

Title: (Accelerator design application)

Scalable Data Management, Analysis and Visualization of Particle Accelerator Simulation Data

Presenter/POC: Oliver Ruebel, LBNL

Title: (Climate application)

Recent Achievements in Visual Data Exploration and Analysis for Climate Science

Presenter/POC: Wes Bethel, LBNL

Title: (Combustion application)

SDAV Support to Combustion Research with Visualization, Analysis, and Data Movements

Presenter/POC: Attila Gyulassy, University of Utah

Title: (Cosmology application)

Visual and Numerical Data Analysis of Large-Scale Structures in Cosmological Simulations

Presenter/POC: Tom Peterka, ANL

Title: (Fusion application)

SDAV technologies for the next generation fusion techniques

Presenter/POC: Scott Klasky, ORNL

3 TECHNOLOGY POSTERS

Title: (Data Management area)

Technologies for Extreme-Scale Data Management

Presenter/POC: Rob Ross, ANL

Title: (Analysis area)

Advanced Data Analysis Techniques for Science Discovery

Presenter/POC: Valerio Pascucci, University of Utah

Title: (Visualization area)

Technologies for Scientific Visualization

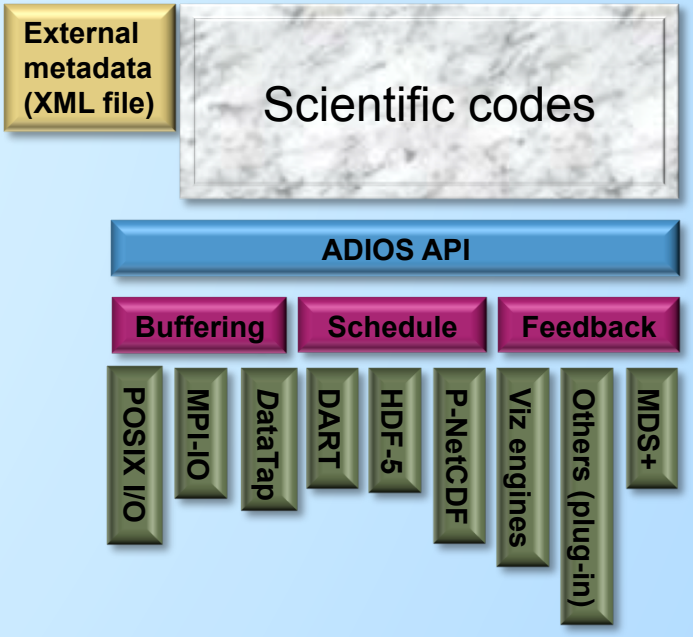
Presenter/POC: James Ahrens, LANL

EXAMPLE:

CAPTURING I/O FUNCTIONS

ADIOS: Adaptable I/O System

- **ADIOS**: adaptable I/O framework
- Provides portable, fast, scalable, easy-to-use, metadata rich output with a simple API
- Allows plug-ins for different I/O implementations
- Abstracts the API from the method used for I/O
- Synchronous and asynchronous transports
- Insulates codes from underlying I/O system



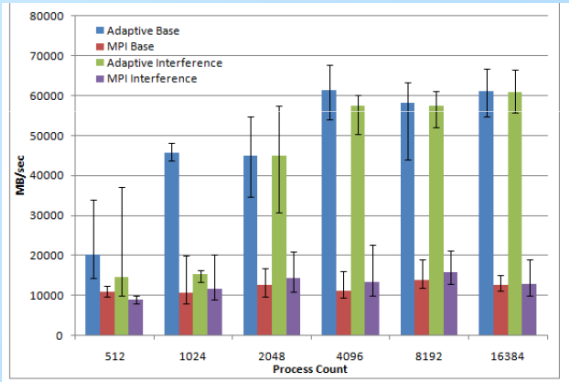
Reducing I/O variability

Objectives:

- Utilize the maximum number of storage targets.
- Schedule the processes to avoid internal interference.
- Shift work from slower to faster storage targets for external interference.

Impact:

- Improved I/O overhead in all simulation codes



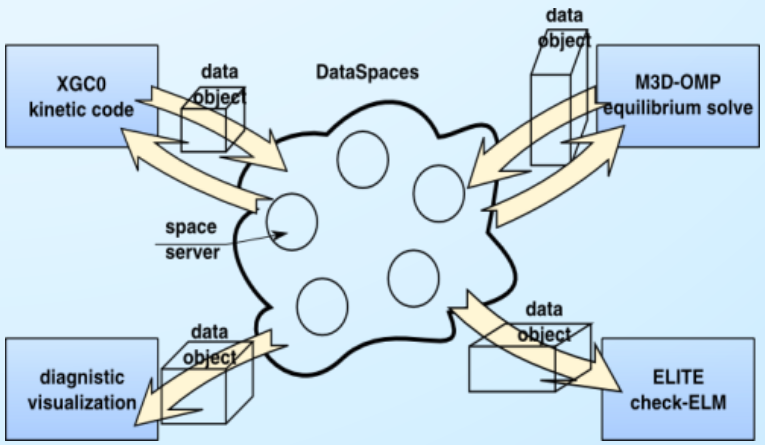
SDAV
Scalable Data Management, Analysis, and Visualization

Contact: Scott Klasky

What else can be done when capturing I/O

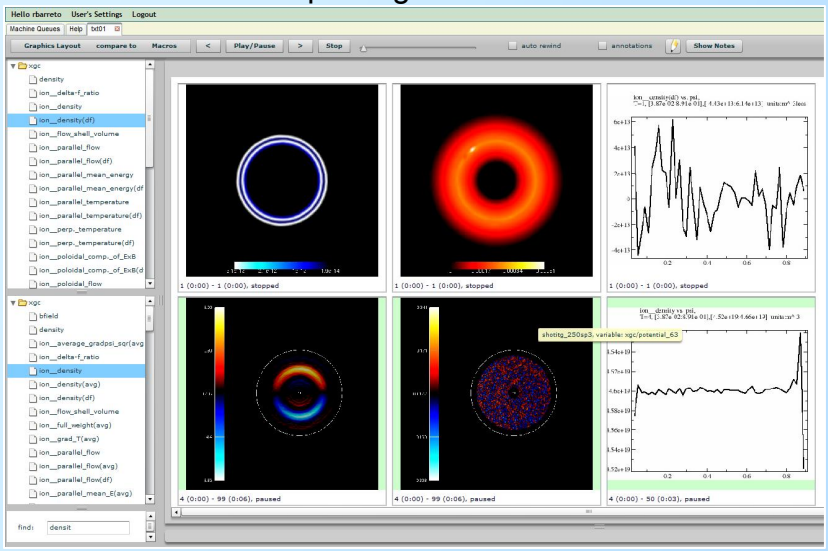
Code coupling with DataSpaces

- Virtual shared space
- Constructed on-the-fly on staging nodes
  - Indexes data for quick access and retrieval
  - Provides asynchronous coordination and interaction and realizes the shared-space abstraction
- In-space (online) data transformation and manipulations
- Decentralized data analysis in-the-space



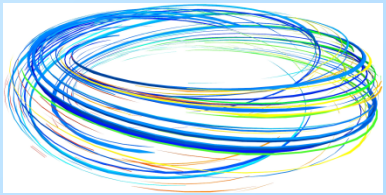
Real-time monitoring on a Dashboard

- Time-step data is captured and processed by a workflow system
- Movies generated by a workflow
- Interactive vector plotting



Imbed FastBit indexing to find regions of interest

- Set of regions with high electromagnetic potential in a torus modeled by GTC
- Achieved speed up of 500-900 fold on magnetic coordinates



EVPath data streaming for monitoring and control

- Multilevel management hierarchy
- Runtime resource management for I/O pipelines

SDAV
Scalable Data Management, Analysis, and Visualization

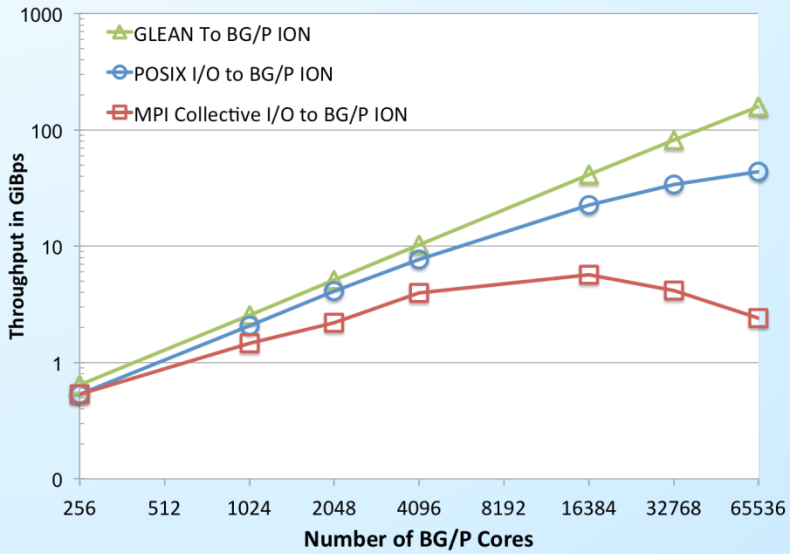
Contact: Scott Klasky

EXAMPLE:

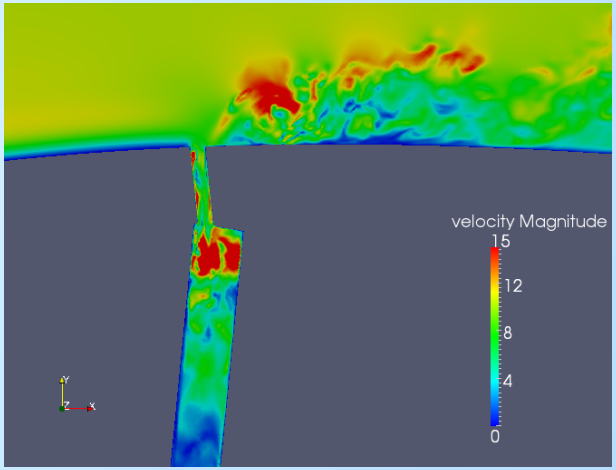
FRAMEWORK FOR IMPROVING I/O AND IN SITU VISUALIZATION

GLEAN

- GLEAN is a flexible and extensible framework to facilitate simulation-time data analysis and I/O acceleration.
- Features include: topology-aware data movement, asynchronous data staging and burst buffering, leverages application data models, scalable analysis algorithms and infrastructure (in situ, co processing, in flight).
- Scaled to entire ALCF infrastructure (160K BG/P Intrepid cores), achieved multi-fold I/O improvement for FLASH, and demonstrated in situ analysis.



Strong scaling performance for 1GB data movement from ALCF Intrepid Blue Gene/P. Strong scaling is critical as we move towards systems with increased core counts.



Co-visualization of a 3.3 billion element PHASTA simulation of an aircraft wing running on 160K cores of ALCF Intrepid Blue Gene/P using ParaView on 100 nodes of ALCF Eureka analysis cluster enabled by GLEAN.

[Redesigning the vertical tail of a commercial jet could reduce jet fuel use by 0.5%, resulting in annual savings of $300 million]

SDAV
Scalable Data Management, Analysis, and Visualization

Contact: Rob Ross, V. Vishwanath

EXAMPLE:

TOPOLOGICAL TECHNIQUES

FOR SCIENTIFIC FEATURE EXTRACTION

Topological and Statistical Analytics of Turbulent Combustion

Application:

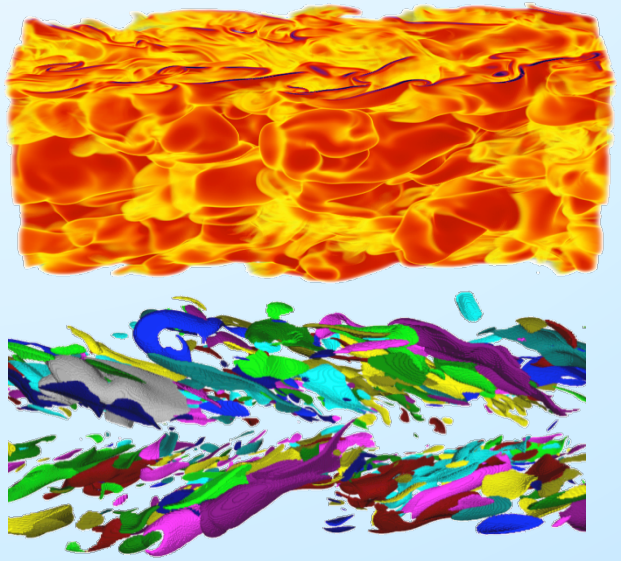
- *Turbulent Combustion*

Goal: understanding turbulent combustion

- Combustion provides 85% of the US energy needs
- Low emission/temperature engines will operate at combustion regimes currently poorly understood
- Need new understanding of turbulence-chemistry to predict efficiency and pollutant emissions

Requirement: exploration of flexible Feature definition, analysis and tracking

- Develop robust feature definition
- Allow user to change parameters interactively and evaluate consequences on results
- Track features over time based on variable correspondence parameters
- Concurrent analysis of multiple fields restricted to the features of interest



Challenges

- Accurate combustions simulation generate massive datasets that are hard to manage in postprocessing
- The data needs to be inspected by the user since selection parameters are not predetermined
- For each parameter selection
  - the user needs to get access the shape characterizations
  - understand the consequence on events like merging or splitting of features of time

(Upper left): Simulation of turbulent combustion
(Lower left): Segmentation of the data based on a particular feature definition

SDAV
Scalable Data Management, Analysis, and Visualization

Contact: Valerio Puscucci

TOPOLOGIKA: Time-Varying Data Analysis with Time Activity Curves

Technology

Robust analysis based on topological definitions

Fast parallel evaluation of dependent statistics

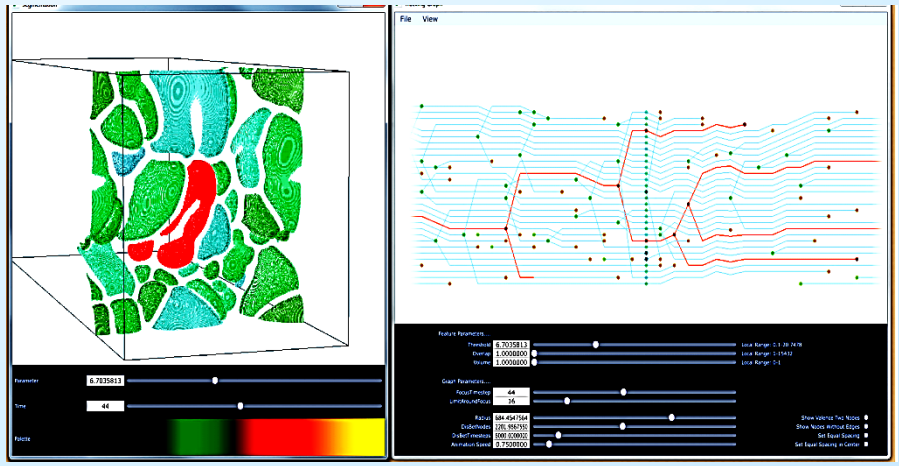
Compute graph abstractions for multi-resolution feature representation

Computer graph models of concurrent tracking of features at multiple scales

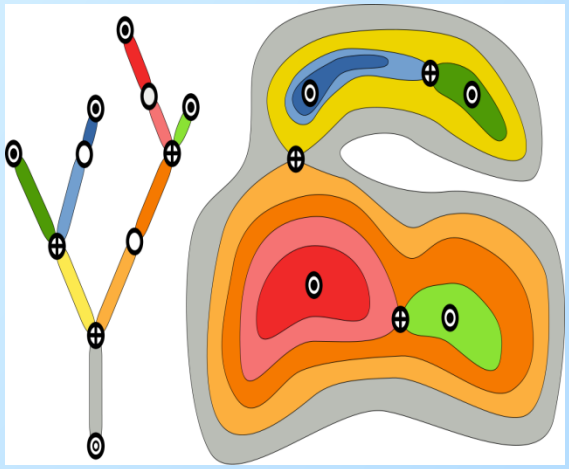
Fast access and computation of dependent statistics

Result/Impact

- *Orders of magnitude in data reduction from raw data to unrestricted feature space*
- *Orders of magnitudes speedup reducing batch jobs into interactive processing*
- *Enable scientists to explore the feature space interactively and understand*
- *Allow data-driven characterization of turbulence from burning cell analysis*



(Left) Presentation of a feature selected in 3D. (Right) Corresponding tracking graph. The color selection (red) used on the feature is used to highlight its time evolution on the graph.



Topological model (left) representing multiple possible nested features of interest (right).

SDAV
Scalable Data Management, Analysis, and Visualization

Contact: Valerio Puscucci

EXAMPLE:

USE OF PARALLEL VISUALIZATION FOR REMOTE MULTI-DISCIPLINE SCIENTISTS (VIRTUAL REACTOR)

Analysis and Visualization of Light Water Reactors

Application:

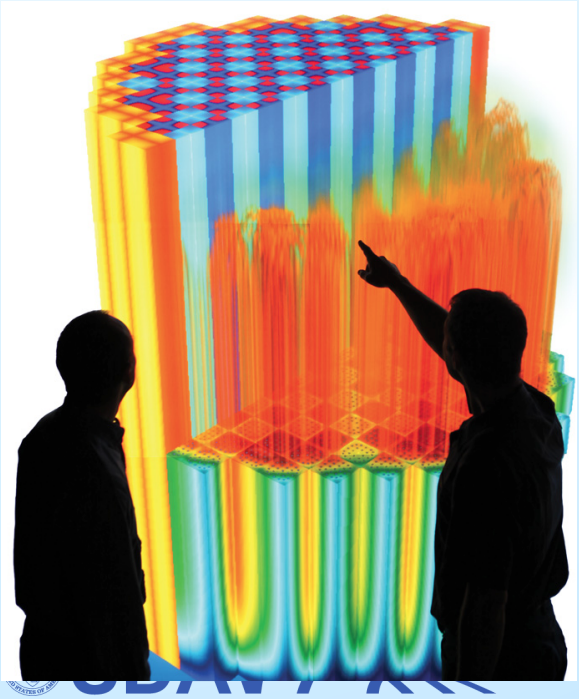
- CASL: Consortium for Advanced Simulation of Light Water Reactors

Simulation Goal:

- Use leadership-class computing for engineering design and analysis of next generation reactors, life extension, and higher fuel burnup ("virtual reactor")

Requirement: Remote Interactive Analysis

- Interactive 3D visualization of disparate simulation data
- Support for large numbers of remote, multi-discipline scientists and engineers
- Multi-physics, multi-material simulations across a broad range of both spatial and temporal scales



Challenges

- Large data generated by multiple, high resolution simulation codes
  - High core count simulations
  - Multi-physics, multi-materials
  - Whole reactor core
- Size of data requires multi-site, remote access

Contact: Dave Pugmire

Analysis and Visualization with VisIt

Technology

VisIt: Turnkey, general purpose tool for large data analysis and visualization

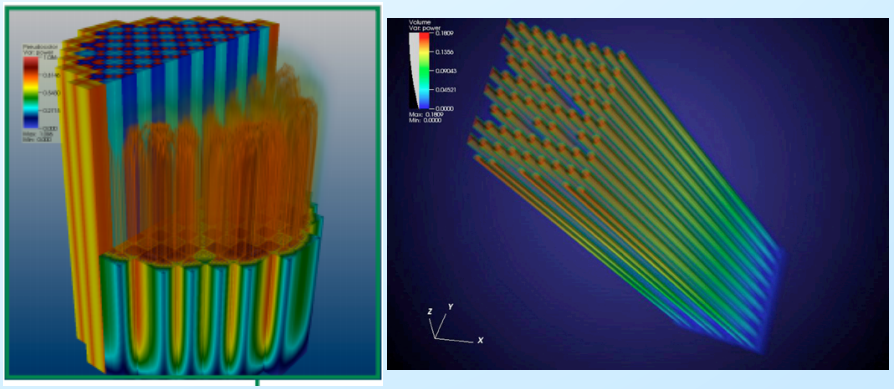
Client/server architecture for remote visualization on analysis clusters and supercomputers

Demonstrated scaling to > 100K cores

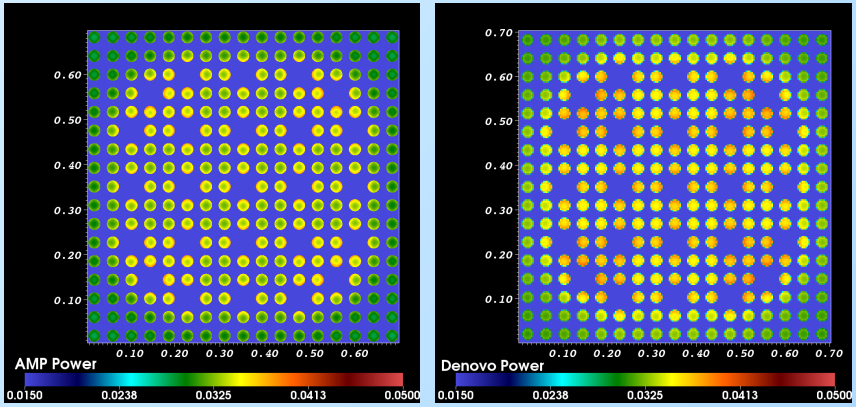
Plugin architecture for flexibility, and extensibility

Result/Impact

- Client-server architecture allows scientists to easily perform analysis and visualization remotely using analysis clusters, or supercomputers.
- Allows scientists to explore simulation results at a variety of scales, including: full reactor core, materials, and structural components.
- Scripted, comparative visualizations were critical in validating new computational techniques for power distribution.



Full core, and pin assembly transport simulations rendered in VisIt using the parallel volume renderer.



Validation study comparing assembly reactor core power calculations from two different codes

SDAV
Scalable Data Management, Analysis, and Visualization

Contact: Dave Pugmire

EXAMPLE:

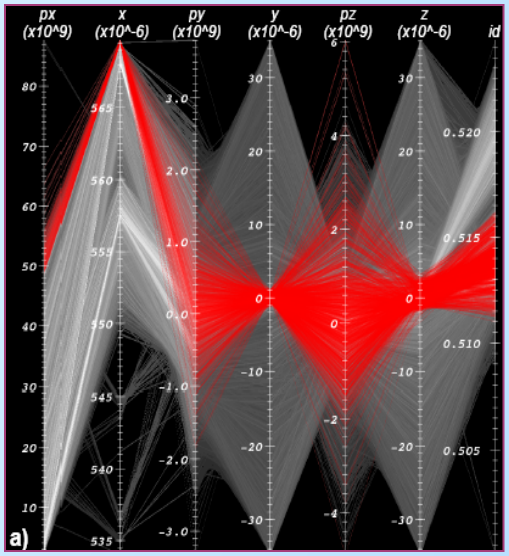
QUERY-DRIVEN VISUALIZATION

Application: Laser Plasma Accelerator Modeling

Beam Selection

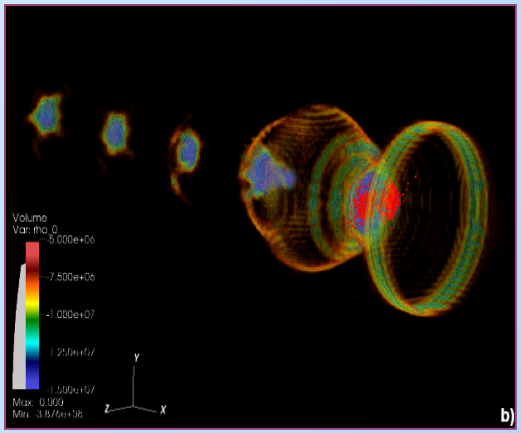
Parallel coordinates view of t = 12

- Grey particles represent initial selection ($px > 2*10^9$)
- Red particles represent "focus particles" in first wake period following pulse ($px > 4.856*10^{10}$) && ($x > 5.649*10^{-4}$)



Volume rendering of plasma density with focus particles included in red (t = 12)

- Helps locate beam within wake



Approach/Impact

- Integration of the parallel index/query system FastQuery with VisIt, enabling high-performance query-driven data analysis
- Enable for the first time detailed analysis of massive particle dataset of unprecedented size in real-time

SDAV
Scalable Data Management, Analysis, and Visualization

Contacts: Wes Bethel, John Wu

EXAMPLE:

SIDE-BY-SIDE VISUALIZATION

Side-by-Side Volume Visualizations of Time Dependent Behavior - Laser Back Scatter

Technology

Ability to view data using a client/server architecture using **VisIt**

Ability to quickly generate side-by-side animations of key physics quantities to understand time dependent behavior

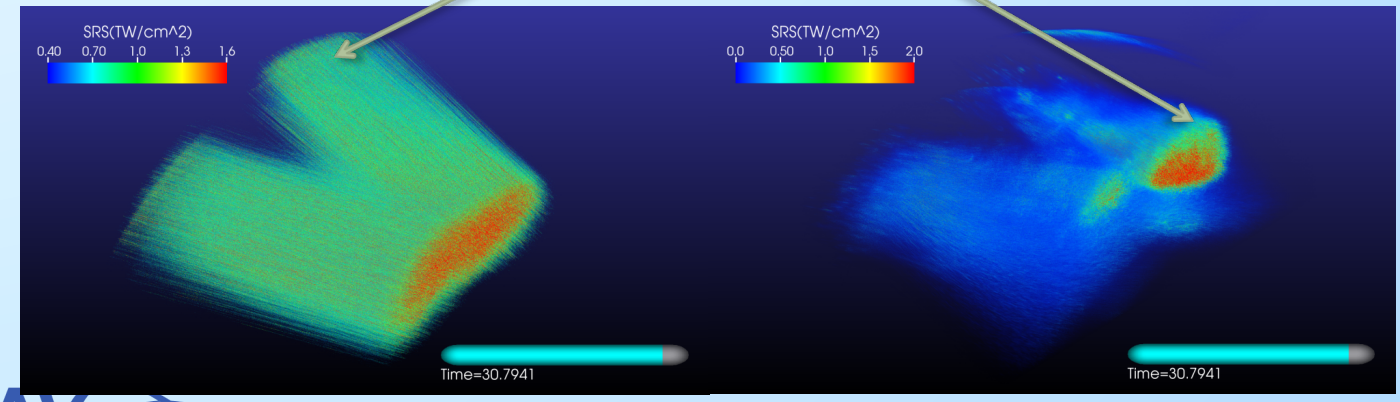
Automated movie generation makes it easy for the user to see time evolution of data with different views and transfer functions.

Extreme data size generated by high resolution simulation (220 billion cells)

Result/Impact

- Client/server architecture allows the user to view his data on his desktop without moving the data
- Ability to interactively set transfer function to bring out features of interest at key points in the simulation
- This is the first time scientists could see how the backscatter was forming. Orthogonal slices was inadequate, since this was a 3d phenomenon. This helps them design the target and laser pulse timing.

Reduced input beam corresponds to high back scatter



Input beam

Back scatter

SDAV
Scalable Data Management, Analysis, and Visualization

Contact: Eric Brugger

EXAMPLE:

IN SITU DATA REDUCTION

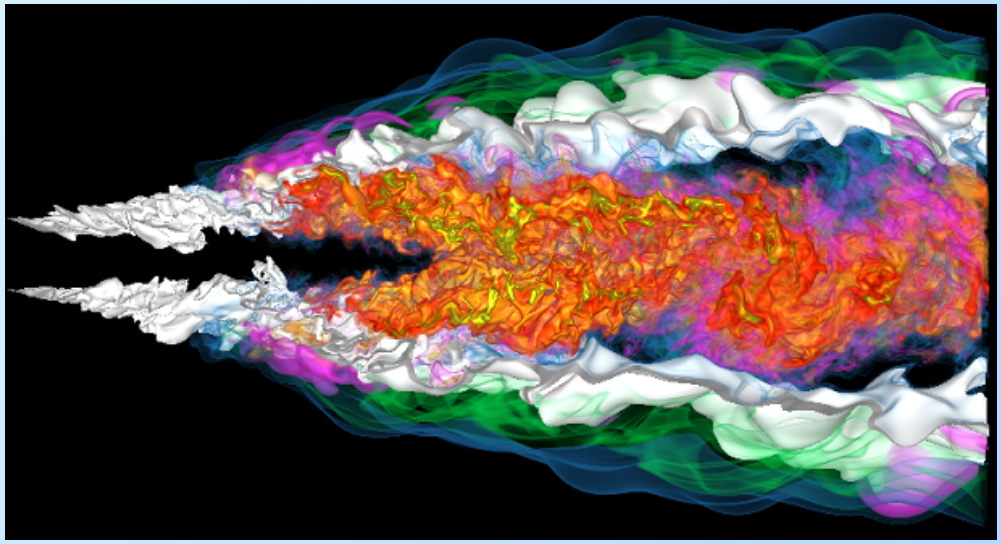
Promising Ideas for larger scale data: and User-Assisted Data Reduction

Importance-Driven Analysis

- In situ analysis incorporates analysis routines into the simulation code. This technique allows analysis routines to operate on data while it is still in memory.
- One way to take advantage of in situ techniques is to perform initial analysis for the purpose of data reduction. With help from the application scientist to identify features of interest, we can compress data of less interest to the scientist, reducing I/O demands during simulation and further analysis steps.

The feature of interest in this case is the mixture fraction (white surface). Colored regions are a volume rendering of the HO2 variable (data courtesy J. Chen (SNL)).

By compressing data more aggressively the further it is from this surface, we can attain a compression ratio of 20-30x while still retaining full fidelity in the vicinity of the surface.



Contact: Kwan-Liu Ma

EXAMPLE:

PARALLELIZATION OF LARGE-SCALE 3D MOVIES

Analysis and Visualization of Magnetic Reconnection

Application:

- Simulation of magnetic reconnection by Bill Daughton (LANL) and Homa Karimabadi (UCSD)

Simulation Goal:

- To understand the 3D evolution of tearing modes – a type of plasma instability that spontaneously produces magnetic reconnection while giving rise to topological changes in the magnetic field.

Requirement: Remote Interactive Analysis

- Interactive 3D visualization of simulation data
  - Particle data
  - Mesh data
- Comparison with theoretical expectations
- Rapid exploration due to limited availability of supercomputer to run large simulations

Challenges

- Large data size generated by high resolution simulation
  - Simulation on 98304 cores
  - 6.4 billion cells
  - 1.5 trillion particles
  - 57 TB data
- Only remote access to the supercomputer
- Lack of dedicated visualization resources

SDAV
Scalable Data Management, Analysis, and Visualization

Contacts: Berk Geveci, Jim Ahrens

Remote Visualization with ParaView

Technology

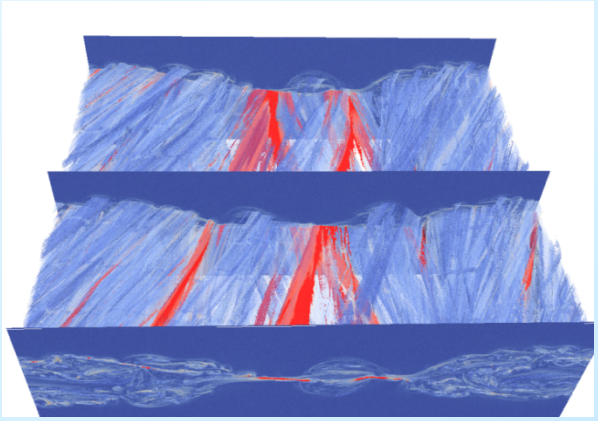
ParaView: general purpose data analysis and visualization tool focused on large data

Client/server architecture for remote visualization on supercomputers

Designed to be extensible: I/O routines and domain-specific algorithms developed by science teams

Result/Impact

- Allowed scientists to remotely analyze and visualize their data when it is not possible to copy locally
- Allowed scientists "to rapidly explore the grid data to understand the 3D evolution of magnetic reconnection"
- As expected, a spectrum of tearing instabilities develops which interact, forming new current sheets and triggering secondary tearing instabilities



Two isosurfaces showing the structure of particle density (blue) and current density (red).

SDAV
Scalable Data Management, Analysis, and Visualization

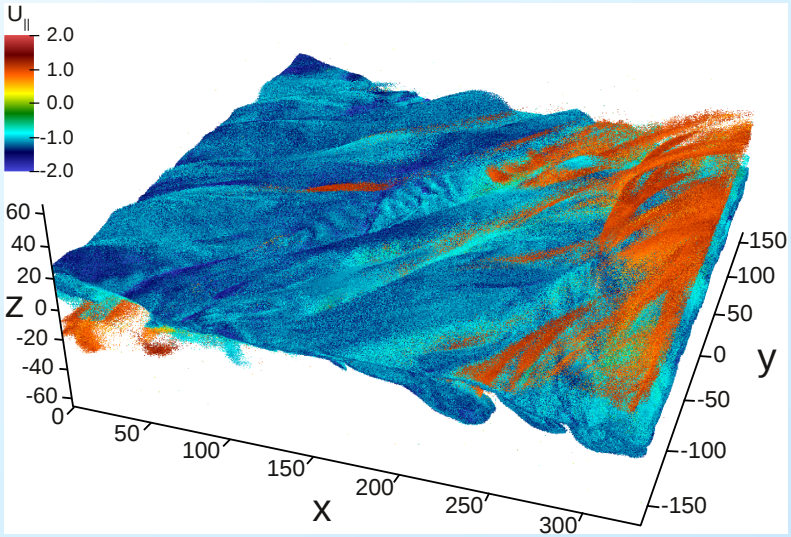
Contacts: Berk Geveci, Jim Ahrens

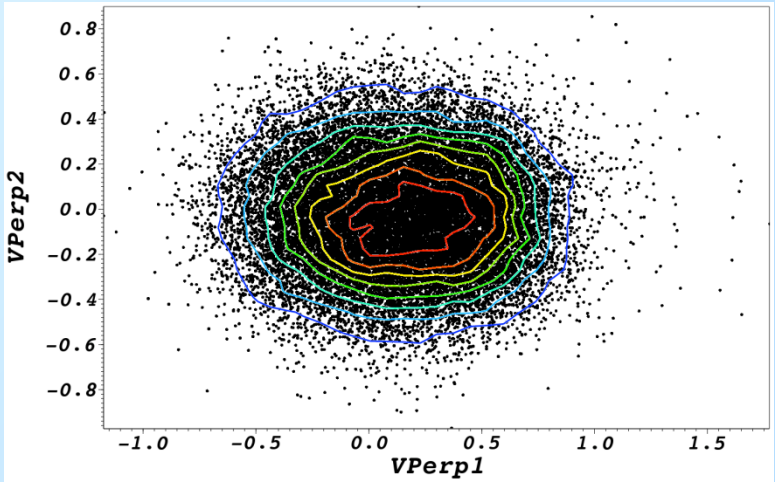
EXAMPLE:

SCALING TO TRILLION PARTICLES

Parallel I/O and Analysis of a Trillion Particles

- Trillion particle plasma physics simulation conducted on 120,000 cores @NERSC
- Enhanced Parallel HDF5 obtained peak 35GB/s, and 80% sustained I/O rate
- FastBit was used to index 30TB timestep in 10 minutes and query in 3 seconds
- Software enabled scientists to examine and gain insights from the trillion particle dataset for the first time:
  - Confinement of energetic particles by the flux ropes
  - Asymmetric distribution of particles near the reconnection hot-spot





Magnetic reconnection from a plasma physics simulation (Left). Scientists were able to query and find an asymmetric distribution of particles near the reconnection event (Right) using our software tools. Rendering by VisIt.

SDAV
Scalable Data Management, Analysis, and Visualization

Contacts: Prabhat, John Wu

Summary of Plan for Integration of Visualization Accelerator Technology

Scientists

use general-purpose visualization tools such as:

VisIT & ParaView

these tools are built on the:

Visualization ToolKit Library

which will integrate these accelerator technology R&D efforts:

DAX
Focus:
Visualization Developers

EAVL
Focus:
Advanced Data Models

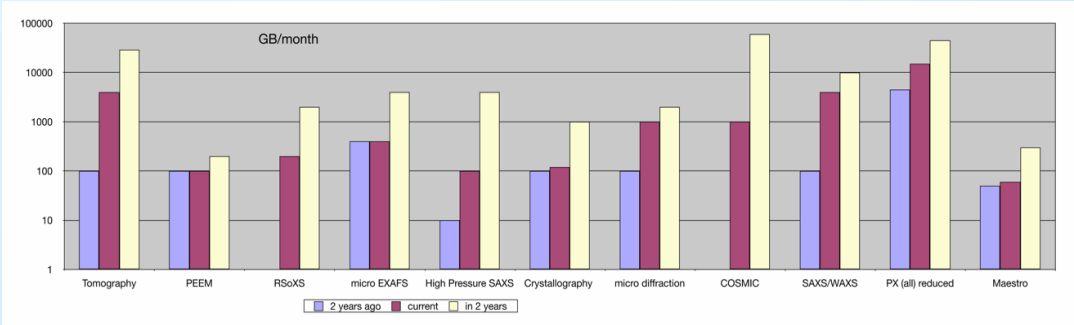
PISTON
Focus:
Portability

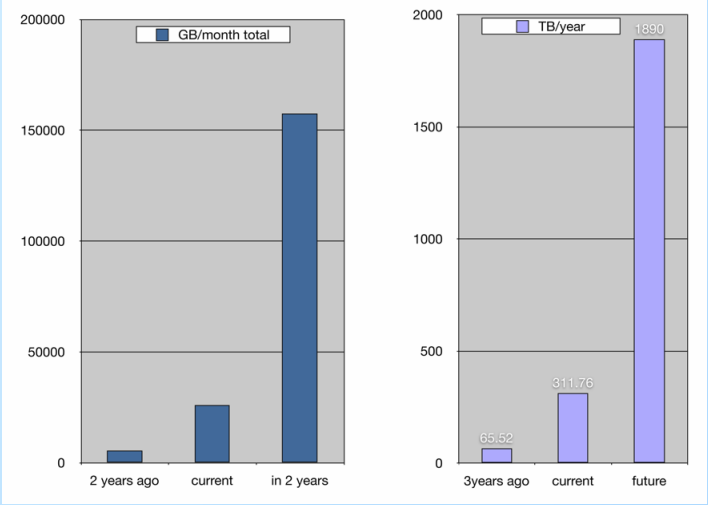
VTK-Threading
Focus:
Evolutionary performance improvement

EXPERIMENTAL DATA

Experimental data – e.g. ALS

- Domain scientists are routinely able to generate10,000's to 100,000's of images in a few days of running.
- Such data volumes cannot be analyzed individually, but rather must rely upon automated methods that translate Materials Science descriptions to input for modeling and simulation.
- The imaging data analysis involves reconstruction of large 3D images, segmentation of the images into sub-regions, discrimination of multiphase solid materials, identification of microstructures, and calculation of statistical correlation functions (e.g. surface, pore-size).

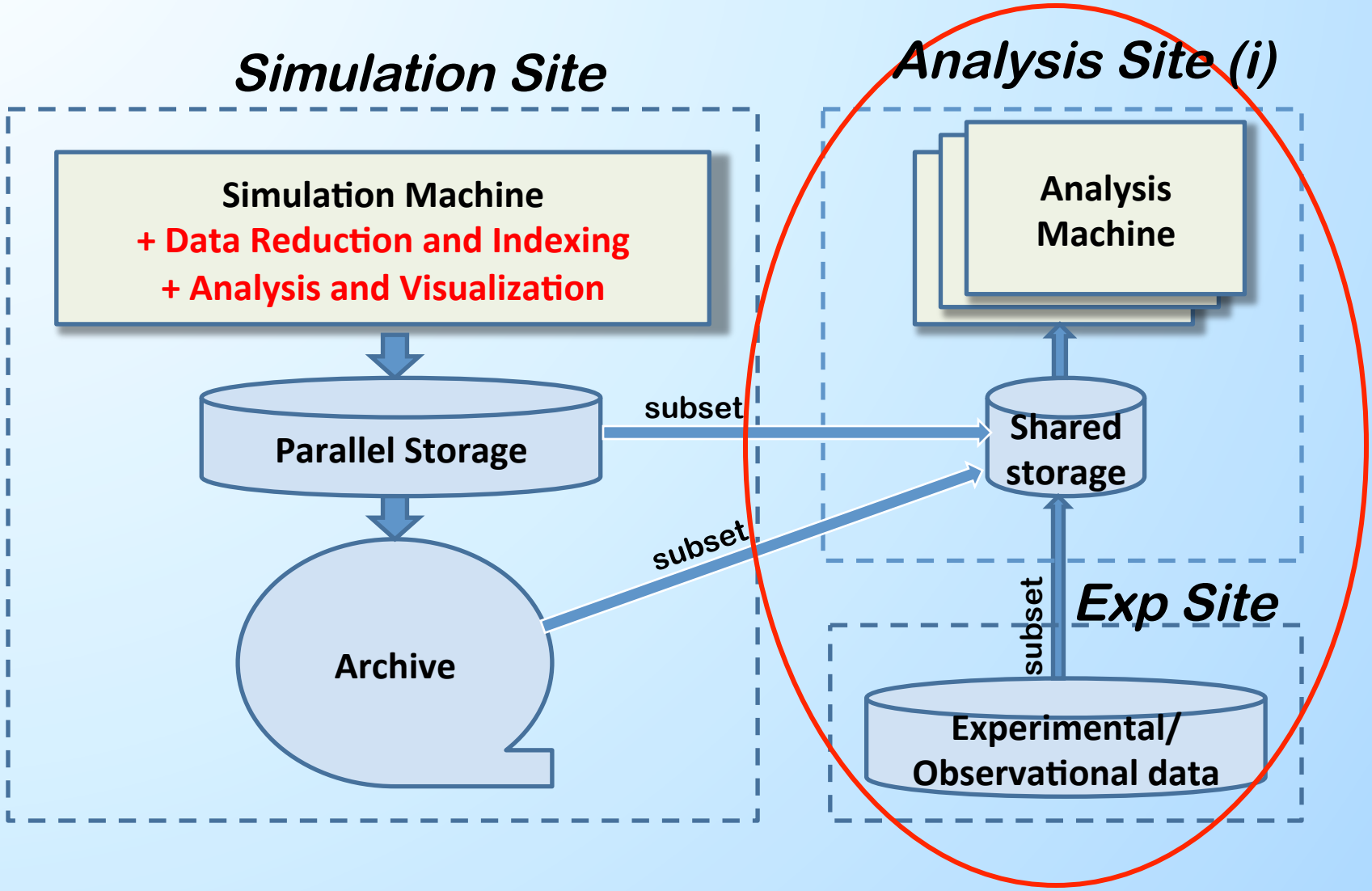




Courtesy of Craig Tull, LBNL

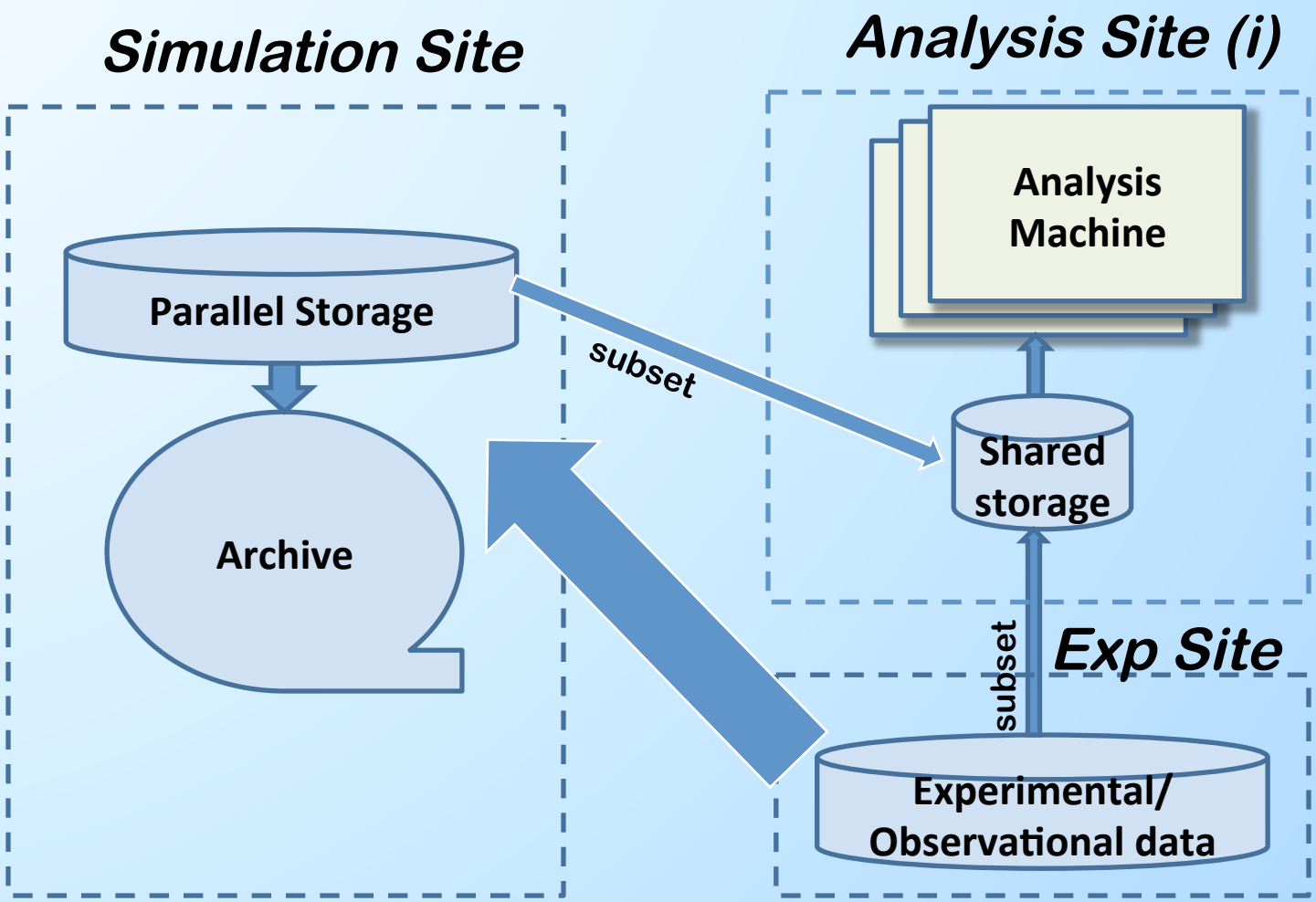
SDAV
Scalable Data Management, Analysis, and Visualization

What about experimental data?



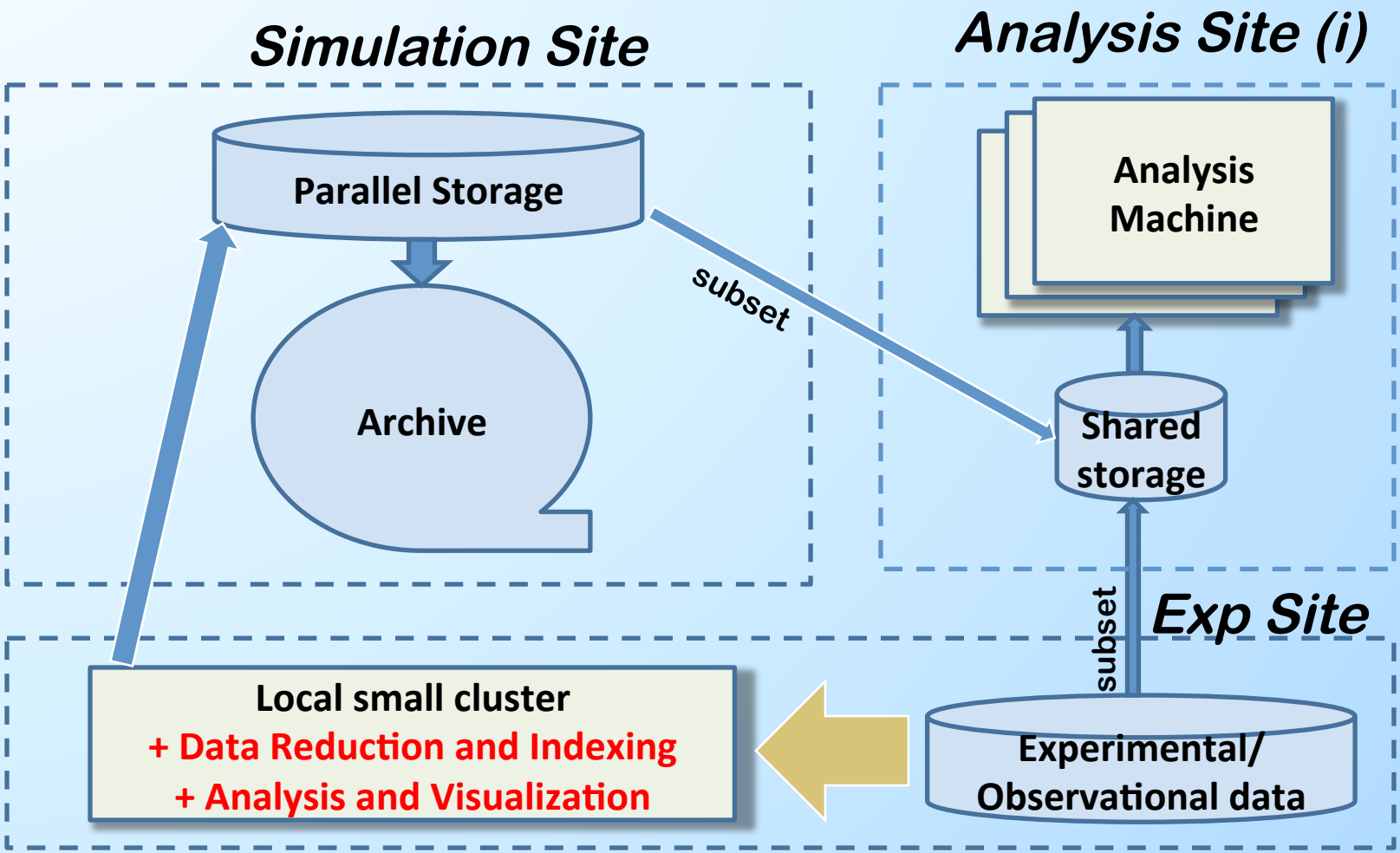
Option 1

- Move all data to a data center that keep tracks and archives data?



Option 2

- Perform analysis, data reduction, and visualization (in-situ)
- Then move data to a shared facility for management and distribution of reduced experimental data



Summary

- Many of the tools in the SDAV institute have been developed over many years and are robust and well-documented
- Some of the tools have been designed to take advantage of high level parallelism
- Such tools have been used for multiple scientific domains, but often require collaboration between application scientists and tool experts; that is the role of SDAV
- For anticipated future needs, tools are being enhance in several ways:
  - Scale tools for high parallelization levels
  - Adapt tools to take advantage of new hybrid hardware (CPUs + GPUs), and hybrid storage (disk and SSDs)
  - minimize data movement between nodes
  - Adapt tools for in situ processing and analysis to provide early insight of the generated data
  - Compress and index data in situ for both in situ and post-processing analysis

THE END