Scalable Data Management, Analysis, and Visualization

THE SDAV INSTITUTE TECHNOLOGIES IN ACTION: ENABLING SCIENCE EXPLORATION

Arie Shoshani

Lawrence Berkeley National Laboratory

2013 SciDAC-3 PI meeting
July 24-26, 2013

Outline

- What do we mean by Big Data in scientific domains
- Overview of SDAV institute technologies
- Examples of applying SDAV technologies to particle-based simulation codes and facilitating highly-efficient data analysis
- Summary

What is “Big Data”?

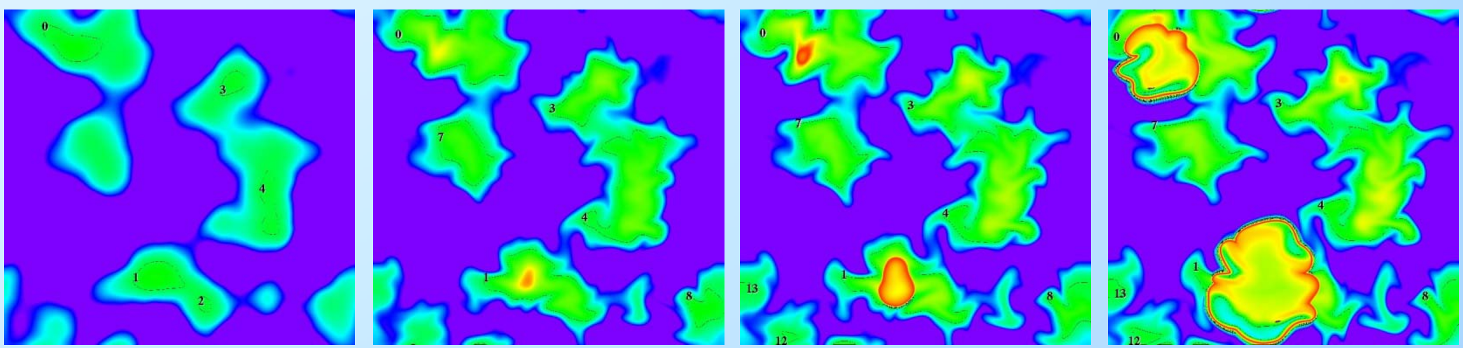
- Definition based on properties of data
- Some combination of the 5 V’s
  - **Volume**- quantity and/or size of individual items
    - E.g. billions of objects, or 100’s PBs
  - **Velocity**- arrival rate
    - E.g. sensor data, satellite data, experimental devices, ...
  - **Variability** (variety) – types of variables/heterogeneous data models
    - E.g. time-series, images, array data, …
  - **Veracity** (verification) – how much do you trust the data
    - accuracy, annotations, noise
    - QA/QC/UQ: Quality Assurance, Quality Control, Uncertainty Quantification
  - **Value** – how valuable is this data
    - E.g. is lots of sensor data be useful to keep?

What is “Big Data” in SDAV?

- Operational definition for scientific domains
  - Data from which you can extract new information, knowledge or insight
  - Applied to large-scale simulation, experimental, and observational data
  - Supported by techniques/technology that give you the opportunity to extract meaning, such as:
    - Machine learning, Data mining
    - Cluster analysis, graph analysis
    - Subset extraction, pattern identification
    - Topology exploration, real-time visual exploration
- Two fundamental aspects to analysis
  - Pattern matching: Perform analysis tasks for finding known or expected patterns
    - can be done in situ
  - Pattern discovery: Iterative exploratory analysis processes of looking for unknown patterns or features in the data
    - In situ preparation of data: summarization, indexing for post-processing

Example of pattern matching in scientific domains

- Finding & tracking of combustion flame fronts
  - Cell identification: Identify all cells that satisfy conditions: “600 < Temperature < 700 AND HO2-concentr. > 10^{-7}”
  - Region growing: connect neighboring cells into regions
  - Region tracking: track the evolution of the features through time

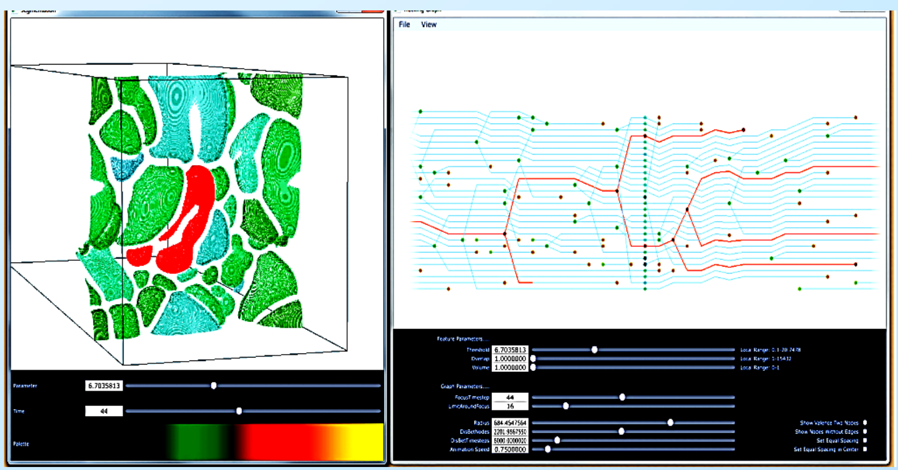


3D simulations: 1000x1000x1000, 1000 time steps = 10^{12} cells, 20 – 50 variables each cell.
400 TBs dataset.

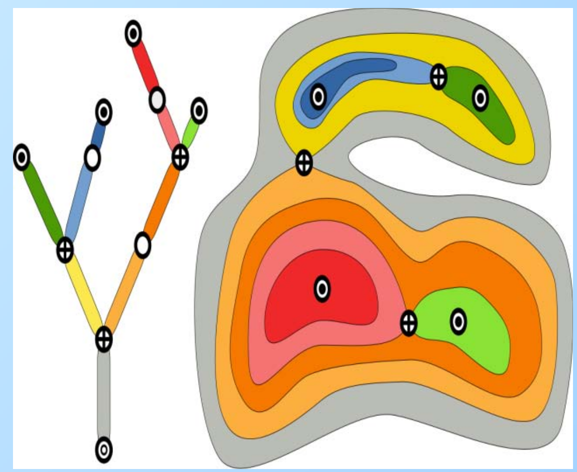
Example of pattern discovery in scientific domains

Time-Varying Data Analysis with Time Activity Curves

- Robust analysis based on topological definitions
- Fast parallel evaluation of dependent statistics
- Compute graph abstractions for multi-resolution feature representation
- Orders of magnitudes in data reduction from raw data to feature space
- Orders of magnitudes speedup reducing batch jobs into interactive processing
- Enable scientists to explore the feature space interactively and understand



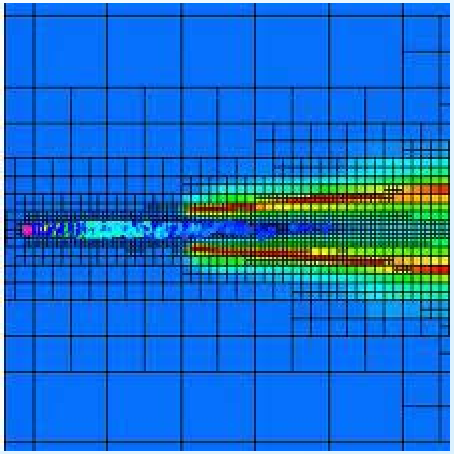
(Left) Presentation of a feature selected in 3D. (Right) Corresponding tracking graph. The color selection (red) used on the feature is used to highlight its time evolution on the graph.



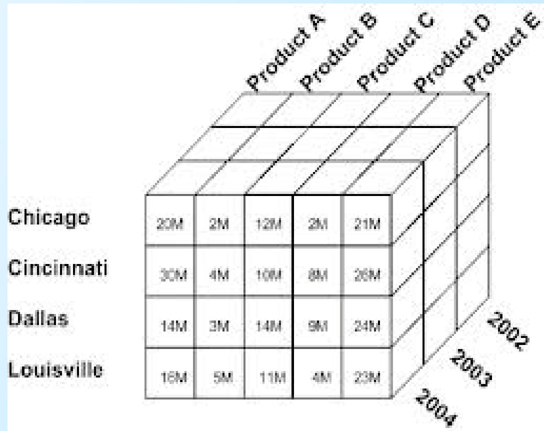
Representing multiple possible nested features of interest

Contact: Valerio Puscucci, UUtah

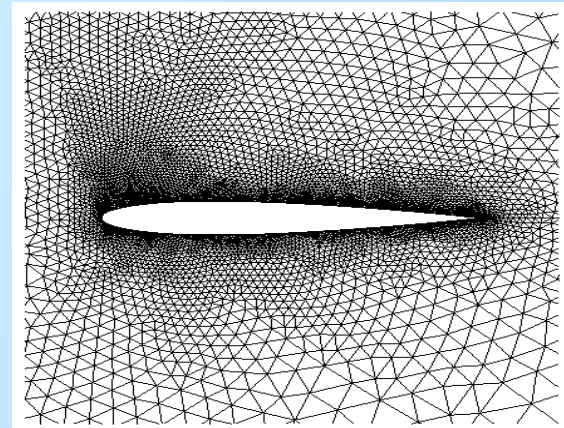
Data Models: logical data structures



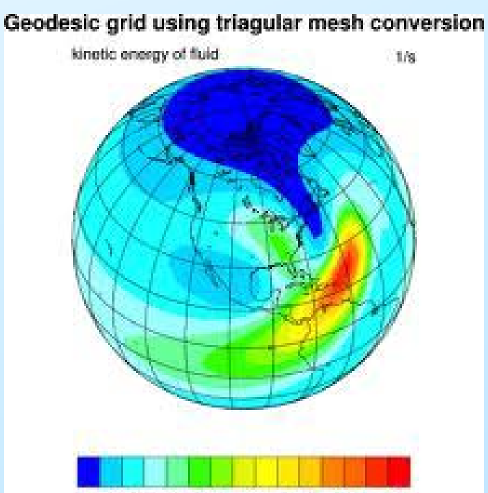
Adaptive Mesh Refinement



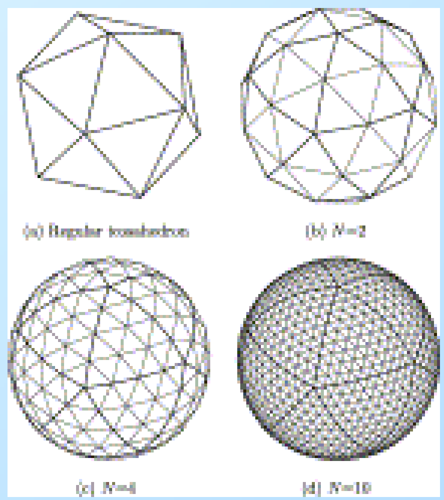
Data Cube



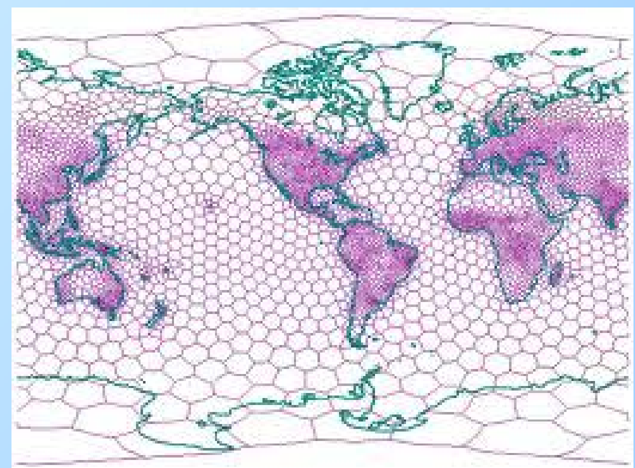
Unstructured triangular grid



Geodesic data model



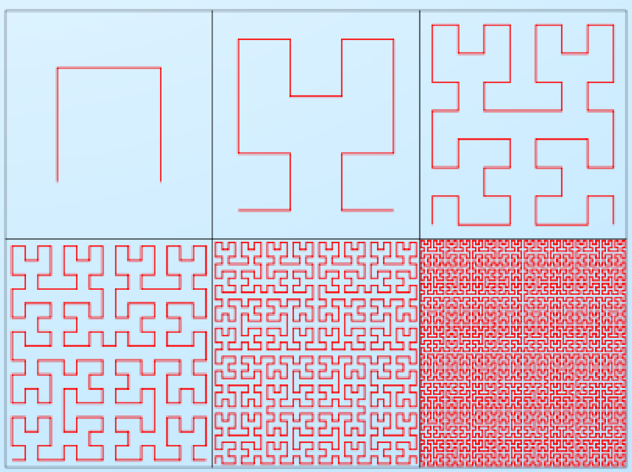
Geodesic triangular data model



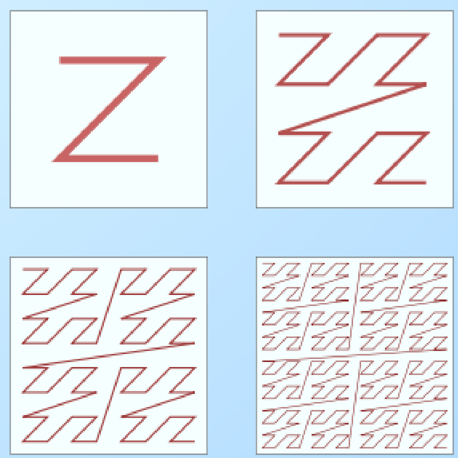
Unstructured grid: Voronoi tesselation

Physical Data Structures

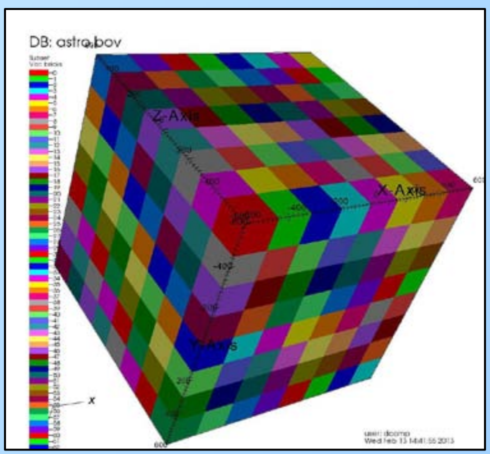
- Linearization of data based on data model
  - By coordinate order based on most prevalent access
  - Hilbert or Z-ordering to support local neighborhood access
- Indexing structures
  - B-tree, Quad-tree, R-tree, Bitmap, Hash
- Partitioning data into blocks for parallel processing
  - Assigning block to different processors
  - Striping blocks on disk



Hilbert linearization order



Z-ordering



512-block dataset colored by thread ID

SDAV technologies address:

- **Data Management**: support of *physical* data structures and optimization of operations over *logical* data structures
- **Data Analysis:** support for manipulations of *logical* data structures to enhance data understanding
- **Visualization:** facilitate real-time visual exploration of space-time data, as well as analysis data structures
- **Specific SDAV focus areas:**
  - Reduce I/O cost on parallel architectures
  - Monitor I/O performance
  - In-situ Analysis and Visualization
  - In-situ Data reduction and Indexing
  - Monitoring simulations progress in real-time
  - Selection of subsets based on variables content
  - reduce in situ data movement
  - Support in situ code coupling
  - Tools for real-time analysis and visualization

The Scalable Data-management, Analysis, and Visualization (SDAV) Institute
2012-2017

Arie Shoshani (PI)

Co-Principal Investigators from:

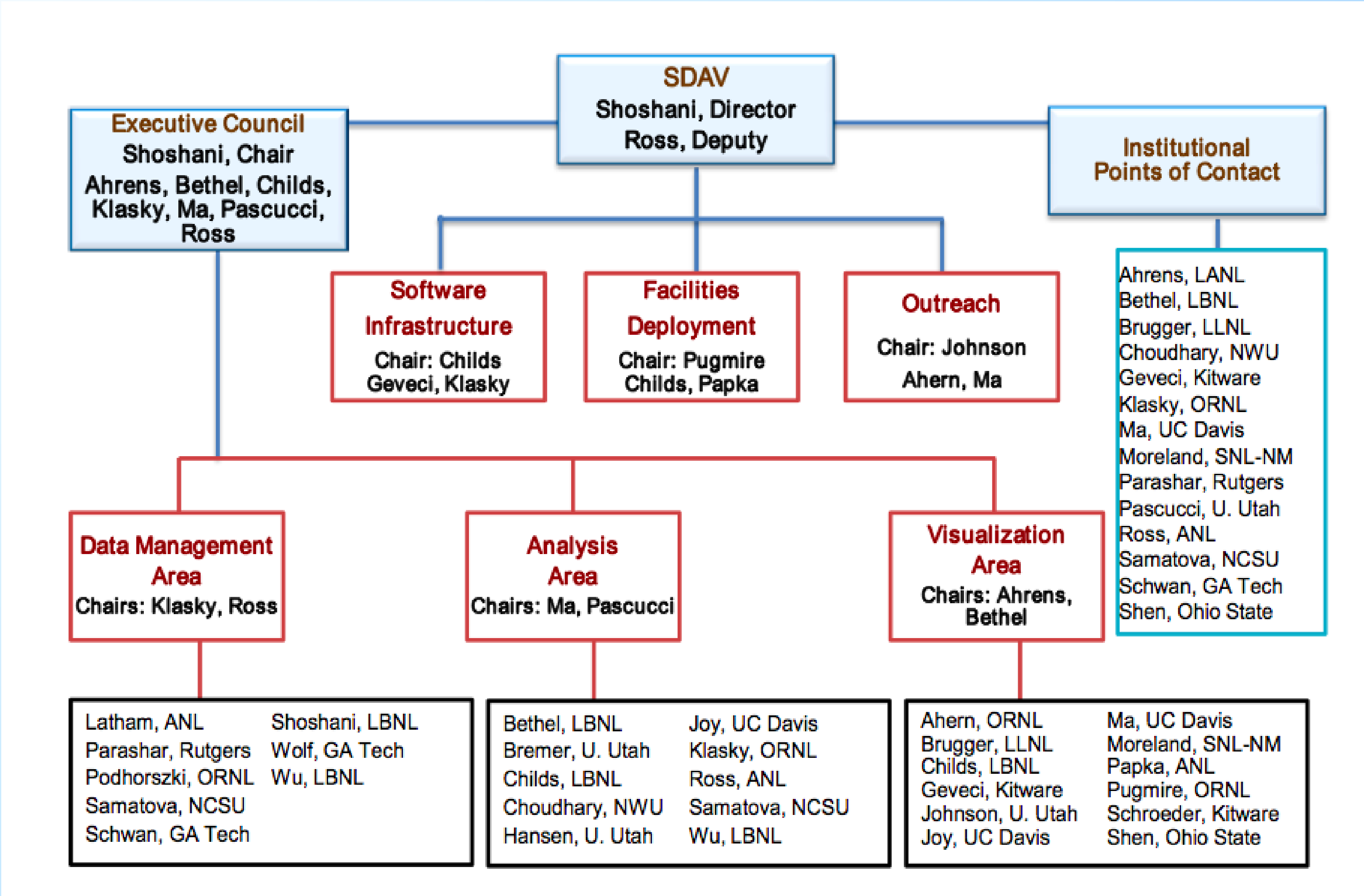
Laboratories	Universities
ANL	GTech
LBNL	NCSU
LLNL	NWU
ORNL	OSU
LANL	UCD
SNL	Rutgers
Kitware (Industry)	UUtah

http://sdav-scidac.org/

SDAV Goals

- The goal of SDAV:
  - 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
  - Current focus is on simulation data
- 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

SDAV Organization



The SDAV institute tools

Data Management tools

- Frameworks for In Situ Processing
  - ADIOS (ORNL)
  - Glean (ANL)
- In situ code coupling and data streaming
  - Data Spaces (Rutgers)
  - FlexPath (Gtech)
- Indexing
  - FastBit (LBNL)
- In Situ Data Compression
  - ISABELLA (NCSU)
- Parallel I/O and File Formats
  - PnetCDF, BP-files, HDF5
- I/O monitoring
  - Darshan (ANL)

Data Analysis tools

- Statistical and Data Mining Techniques
  - NU-Minebench (NWU
- Importance-Driven Analysis Techniques (UCD, Utah)
  - Domain-Knowledge Directed
  - Geometry Based
- Topological Methods (Utah, UCD, LBNL)
  - In Situ Topology (Topologika)
  - Feature-Based Analysis
  - High-Dimensional Analysis

Visualization tools

- Parallel visualization
  - Visit (LBNL, SNL, LLNL)
  - ParaView (LANL)
  - VTK-m framework (Kitware)
- Specialized Vis techniques (UCD, Utah, LBNL, OSU)
  - Flow Visualization Methods
  - Rendering
  - Ensembles, Uncertainty, and Higher-Dimensional Methods

Poster: SDAV's software infrastructure and tools
3 posters: Data Management, Analysis, and Visualization

SDAV uses multiple technologies for various Domain Sciences (see posters)

- Climate
  - UV-CDAT: Community Software Infrastructure for Ultra-scale Climate Data Analysis
  - Statistical Fingerprinting of Climate Data
  - Detecting and Analyzing Atmospheric Rivers, an Extreme Weather Events
  - Accelerating Ocean Climate Analysis
  - Feature Tracking and Visualization of the Madden-Julian Oscillation in Climate Model Output
  - Streaming Computation of Ensembles
- Cosmology
  - Feature Detection & Modeling of Cosmology Objects (Halo Structures)
  - CosmoTools: In Situ Analysis Framework
  - Tess: In Situ Voronoi Tessellation Library
  - Efficient and Effective Visualization of Cosmology Objects
  - Checkpoint and Analysis I/O of Cosmology Codes

SDAV uses multiple technologies for various Domain Sciences (see posters)

- Fusion
  - Large-scale I/O with ADIOS
  - Scibox: Online Sharing of Scientific Data via Cloud
  - FastQuery: Finding Regions of Interest with Complex Mesh
  - Parallel vector field analysis in VisIt
  - Advanced Visualization Techniques for Fusion Data
  - DataSpaces for In Situ / In Transit Data Staging and Analytics
  - Monitoring simulations with eSiMon
  - In transit Analysis and Visualization
  - Data Compression for Fusion Applications
  - Tight-coupling of Fusion Codes for Exascale Computing
- Accelerator design
  - High-performance Query-based Analysis and Visualization of Impact-T, Impact-Z, VPIC codes
  - Automatic Beam Detection for Laser Plasma Accelerator Simulations
  - In Situ Visualization and Analysis using Warp and VisIt

USING SDAV TECHNOLOGIES WITH

PARTICLE-BASED METHODS

ADIOS:

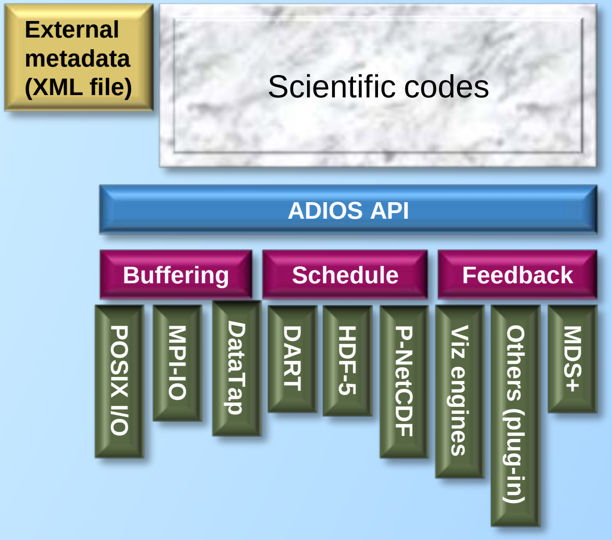
IN SITU PROCESSING FRAMEWORK

(USING I/O ABSTRACTIONS)

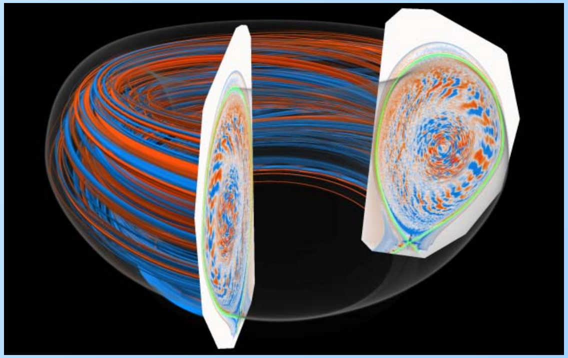
APPLICATION: FUSION

Using ADIOS framework for fusion applications

- ADIOS: Adaptable I/O System
  - Abstracts the API from the method used for I/O
  - service-oriented architecture
  - Simple API, almost as easy as F90 write statement
  - Synchronous and asynchronous transports supported with no code changes
  - Change I/O method by changing XML file only
  - Buffers data to allows in situ optimization of I/O (e.g. Aggregated read/write)



- ADIOS-XCG1 efforts funded by the **EPSI** Fusion SciDAC (C.S. Chang) and **SDAV** SciDAC
  - XGC1 is a gyrokinetic particle turbulence code
  - I/O efficiency: checkpoint writing in XGC-1, GTC, and other large scale codes with low I/O overhead.
  - E.g., XGC-1, **200k** cores on Jaguar, 2TB checkpoint data to write: original parallel I/O >1 hour, ADIOS ~1 min.
  - Embed in ADIOS framework high speed and high fidelity rendering design for visualizing 3D field data from tokamak simulations.



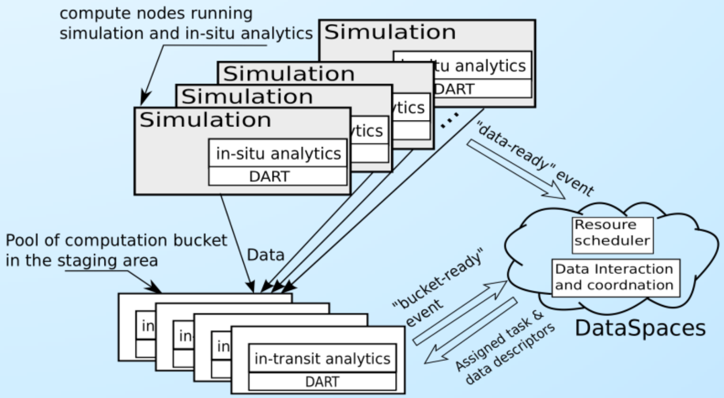
Electrostatic potential perturbation by Ion-Temperature-Gradient (ITG) driven plasma turbulence in a realistic diverted DIII-D tokamak geometry. XGC1 code.

Contact: S. Klasky, N. Podhorszki, ORNL

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



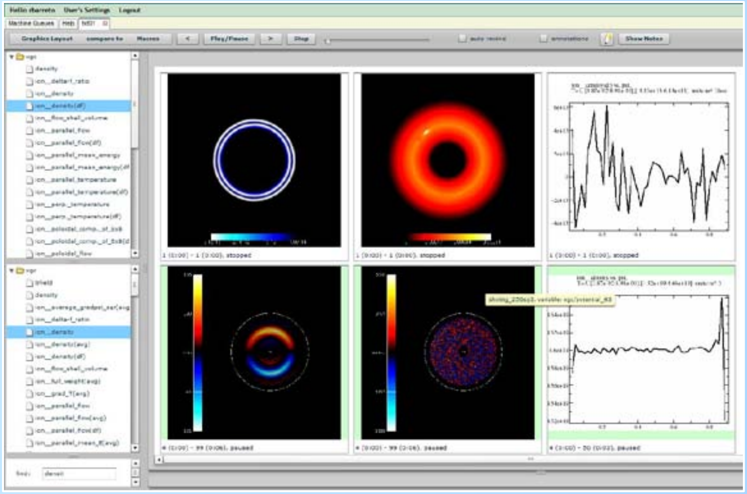
FlexPath data streaming monitoring and control

- An event-based messaging middleware
- Ability to customize data streams
- Enable in situ analytics (histograms, sorting, range queries) for GTC and GTS.

Real-time monitoring on a Dashboard

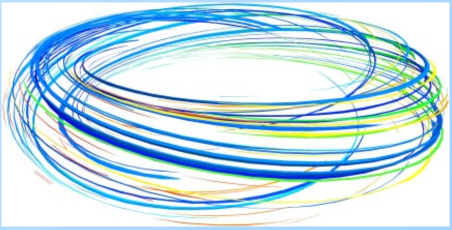
- Time-step data is captured and processed by a workflow system
- Movies generated by a workflow

•Keeps track of provenance information (complete data lineage) which is key to ease of use and efficiency.



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



Contact: M. Parashar, Rutgers, K. Shwan, GT, John Wu, LBNL

GLEAN:

FRAMEWORK FOR IMPROVING I/O AND

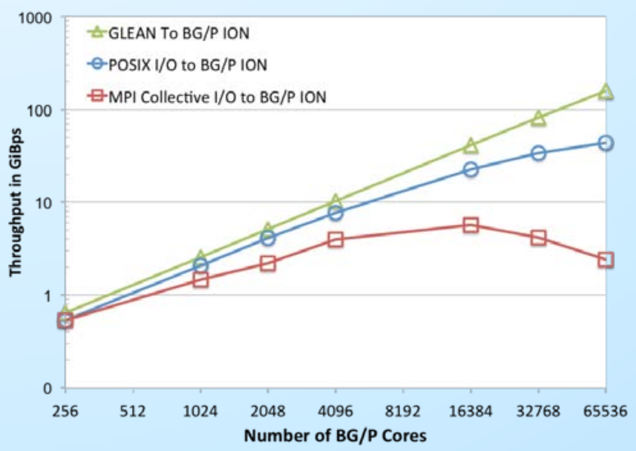
IN SITU ANALYSIS

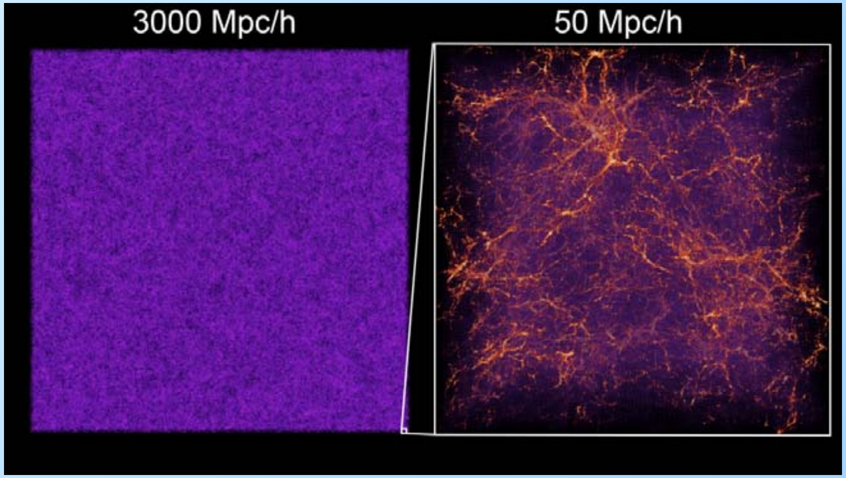
(LEVERAGING HARDWARE TOPOLOGY)

APPLICATION: COSMOLOGY

Accelerating I/O of HACC cosmology with GLEAN

- GLEAN is an I/O acceleration and simulation-time analysis framework
  - Leverages system topology, including 5D torus of BG/Q, for accelerating data movement
  - Features include: asynchronous data staging and sub-filing capabilities to reduce control overheads in parallel file-systems
  - support for various application-specific data models
- GLEAN-HACC I/O efforts funded by the HEP Cosmology SciDAC (S. Habib) and SDAV SciDAC
  - Demonstrated scaling to entire 786,432 cores on Mira for HACC
  - Achieved 160 GB/s (~10X improvement over the previous I/O scheme)
  - Written and read ~10PB of data on Mira, as well as scaled to Hopper





- 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.

- 1.1 trillion particles simulation on Mira
- Visualization scaled to >10000 cores on GPUs
- 10Kx10Kx800 (~9,000 light-years on a side) from a single I/O node (left image)
- 5123 (~150 light-years on a side) from a single process (right image)

Contact: V. Vishwanath, R Ross, ANL

TOPOLOGIKA:

TOPOLOGICAL TECHNIQUES

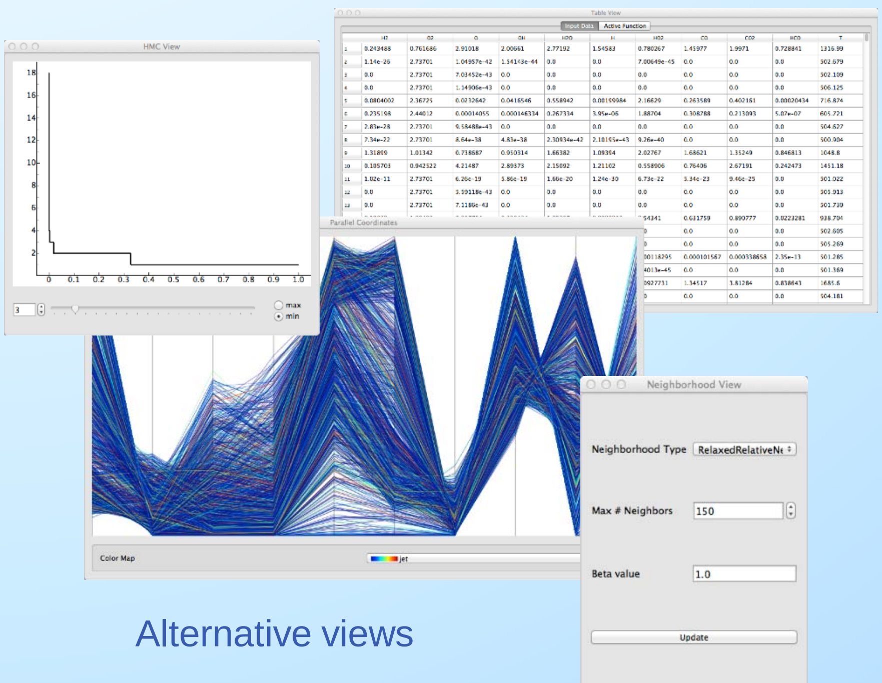
FOR SCIENTIFIC FEATURE EXTRACTION

APPLICATION: COMBUSTION

N-Dimensional Data Analysis and Visualization

Problem

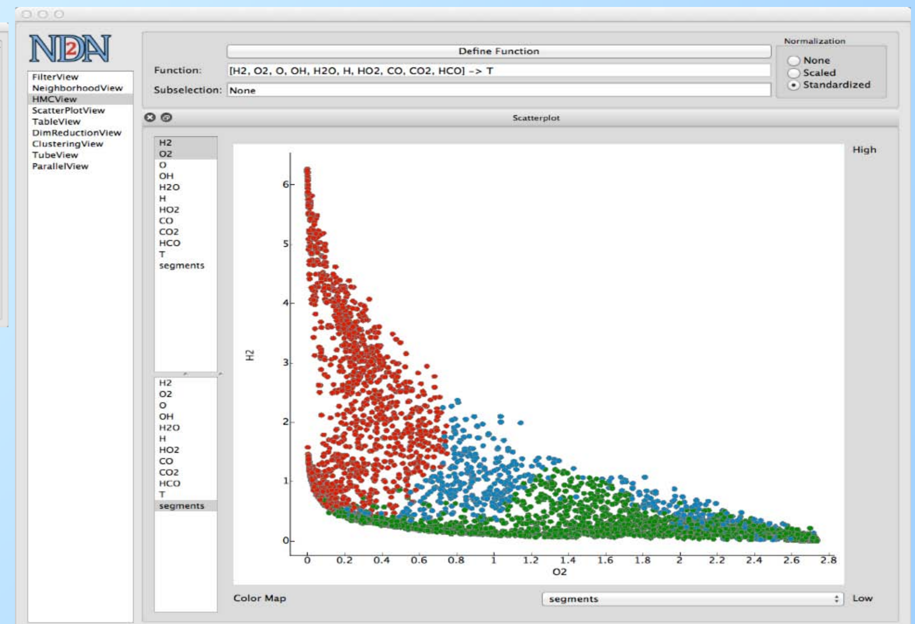
- Understanding complex high-dimensional data is becoming increasingly important in areas such as: UQ, high energy physics, material science, etc.
- At the core we need to analyze and visualize high dimensional functions to extract unknown interactions or relations



Alternative views

Approach:

- Instead of projecting high-dimensional data into low dimensions for analysis we compute topological structures in high-dimensions and project the results
- ND^2AV an interactive analysis framework for high-dimensional data
- Compute high-dimensional topological decomposition to segment data into regions of similar behavior



Fuel vs. oxygen distribution colored according to three minima of heat release. Red / green segments are all fuel / oxidizer and cannot burn. The blue segment seems to correspond to extinction events

Contact: V. Pascucci, U Utah

FastQuery:

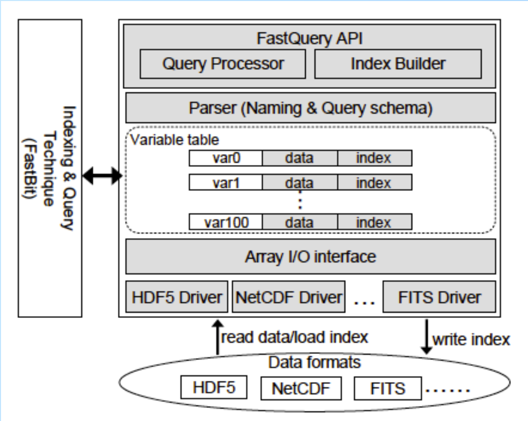
PARALLEL INDEXING AND QUERYING OF MASSIVE DATASETS

APPLICATION: MAGNETIC RECONNECTION

Parallel I/O and Analysis of a Trillion Particles

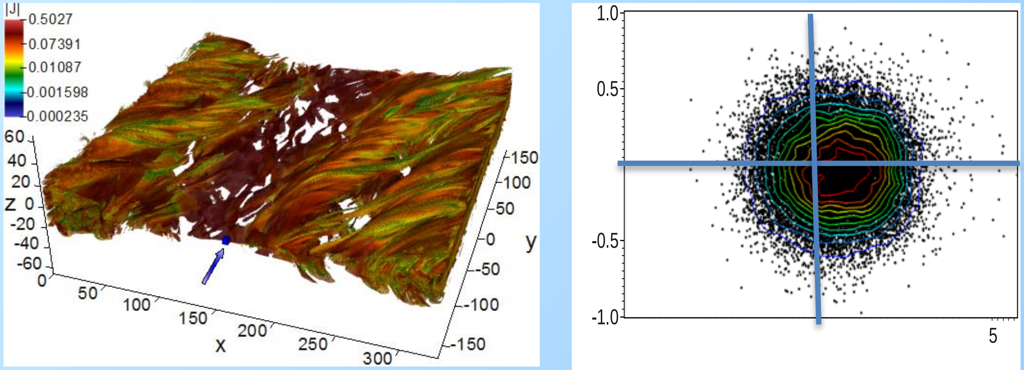
FastQuery: An API to query arbitrary scientific data formats

- Uses FastBit indexing and querying technology
- Parallel implementation previously used MPI-only
- Recent implementation uses hybrid parallelism to utilize multicore CPUs



Vector Particle-in-Cell (VPIC) Simulation of Magnetic Reconnection

- A 3D electromagnetic relativistic PIC plasma physics simulation
- Working with Homa Karimabadi, UCSD
- 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 a 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 FastQuery. Rendering by VisIt.

Contact: S. Byna, Prabhat, John Wu, LBNL

Parallel DBSCAN:

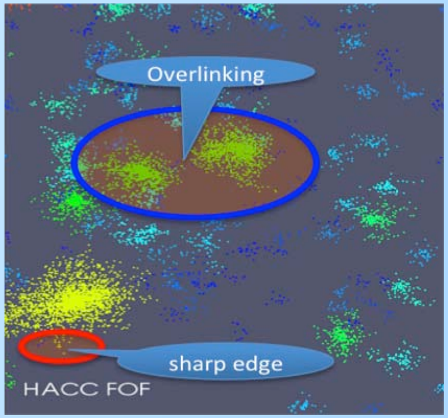
PARALLEL CLUSTERING METHOD

APPLICATION: COSMOLOGY

The Structure of Halos: FOF vs. DBSCAN

Problem Definition

- The dark matter halo mass function is a key repository of cosmological information. N-body simulation shows that Friend-of-Friend (FOF) mass function has a universal form to a surprising level of accuracy.
- However, observed group and cluster masses are usually stated in terms of a spherical over-density (SO) mass, which does not map simply to the FOF mass.
- Two problems: “over-linking” and “sharp edges”
- **Challenge**: Investigating whether over-linking problem of FOF can be mitigated by another clustering method - DBSCAN.
- **SDAV Challenge:** massive parallelism of DBSCAN

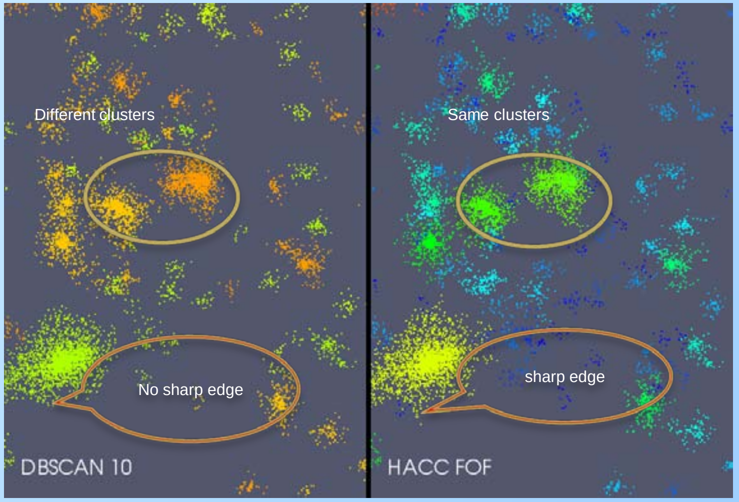


More zoomed in - Halos and subhalos in the astrophysics data (http://arxiv.org/abs/1203.3695)

Collaboration of HEP Cosmology SciDAC (S. Habib) and SDAV SciDAC

Results

- DBSCAN was parallelized and scaled
- Experiment: DBSCAN with 100,000 Monte-Carlo samples, 1,000 particles per sample, concentration
- Can tune minimum number of points
- “over-linking” and “sharp edges” eliminated



Comparing DBSCAN vs. FOF algorithm

Contact: W-K. Liao, A. Choudhari, NWU

INTEGRATION OF TECHNOLOGIES:

HIGH-PERFORMANCE DATA QUERY AND VISUALIZATION FOR INTERACTIVE ANALYSIS

APPLICATION: ACCELERATOR DESIGN

Automatic Beam Detection for Laser Plasma Accelerator Simulations

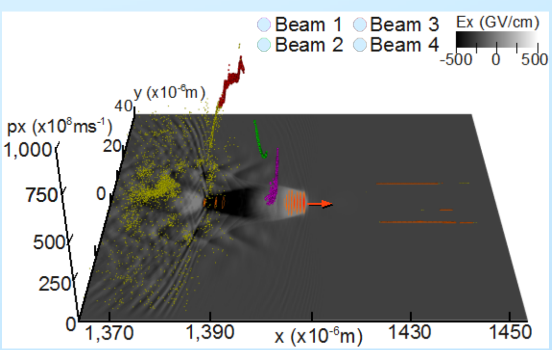
Problem Definition

- Develop new compact, high-quality accelerator technology based on laser-plasma acceleration.
- Datasets are extremely large, heterogeneous and of varying spatial and temporal resolution.
- Particle beams are small (<1% of the data) and beam formation and acceleration are highly complex processes.
- Diverse physical settings, including single laser pulse and multiple colliding laser pulse accelerator designs (see figures)

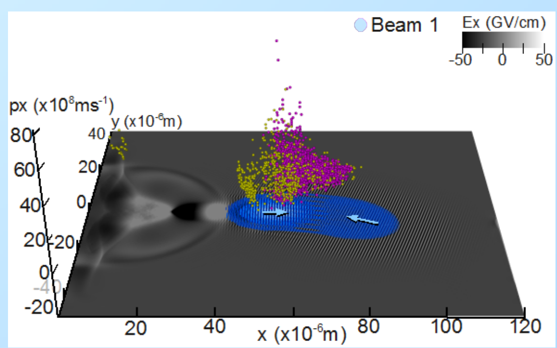
Collaboration of accelerator domain scientists (Robert D. Ryne†, Jean-Luc Vay) using VORPAL code

Results: integration of ViSit and FastBit technologies

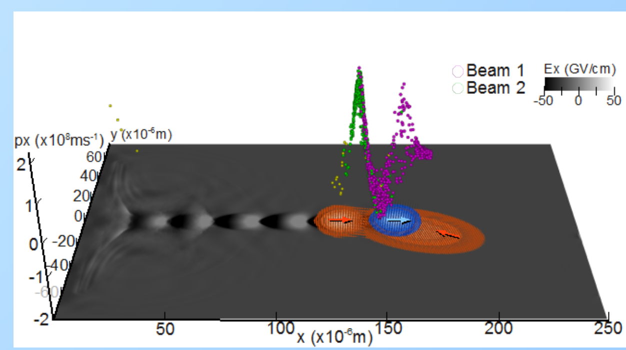
- Fast tracing of particles over time within seconds rather than hours
- Effective exploration of the space of particle features
- Automatic detection of all particle beams and subtle temporal beam substructures
- More accurate and efficient analysis of particle beams and beam substructures then previously possible.



Single laser pulse



Dual colliding laser pulse



Triple colliding laser pulse

Contact: O. Rübel, J. Wu, LBNL

Summary

- SDAV staff have been working closely with application scientists in multiple domain sciences
  - Climate, combustion, fusion, accelerator design, cosmology, biology
  - Lessons learned: such collaboration are necessary and effective
- SDAV emphasizes quality of software
  - Open source, well-documented, regular testing, user support
  - Software infrastructure committee set up guidelines to evaluate products
  - Classified as: Production Software, Emerging Software, Prototype Software
- SDAV outreach
  - 10 tutorials : SC12, LDAV, Ultra-Scale Visualization , CScADS, PDAC-12, …
  - 25 invited presentation, 10 panels.
- SDAV productivity
  - Over 160 papers since start of institute
  - All posted at web site
- Science community engagement
  - Started “invited Scientist” conference calls
  - Identify new problems, forge new collaborations
- We like to work with domain scientist
  - contact Executive Committee members or directly with SDAV staff

THE END