Foghorns, Lighthouses and the Circuitous, Hazard-laden Path Towards Extreme Scale Data Analysis

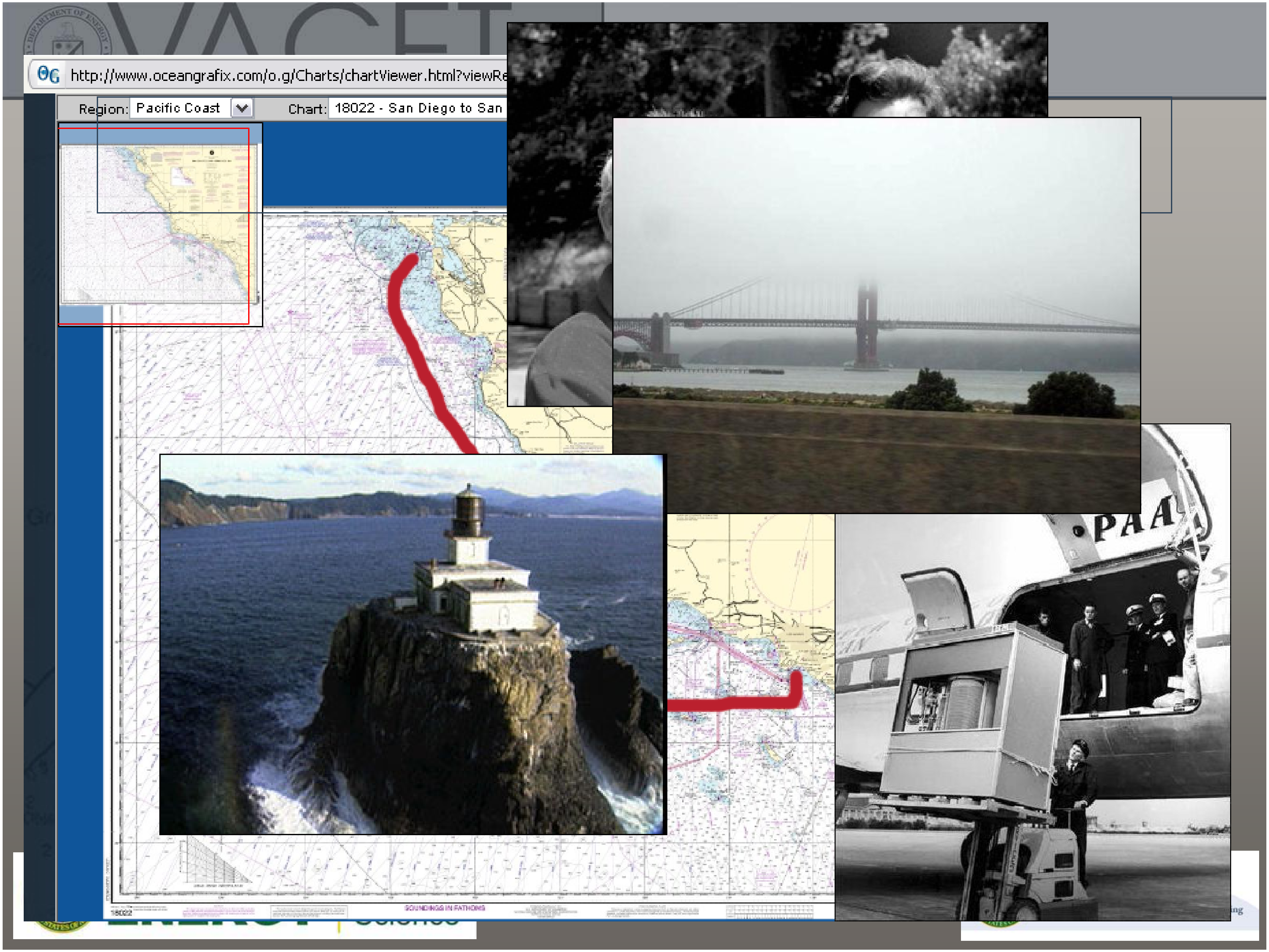
ICAP 2009

4 September 2009

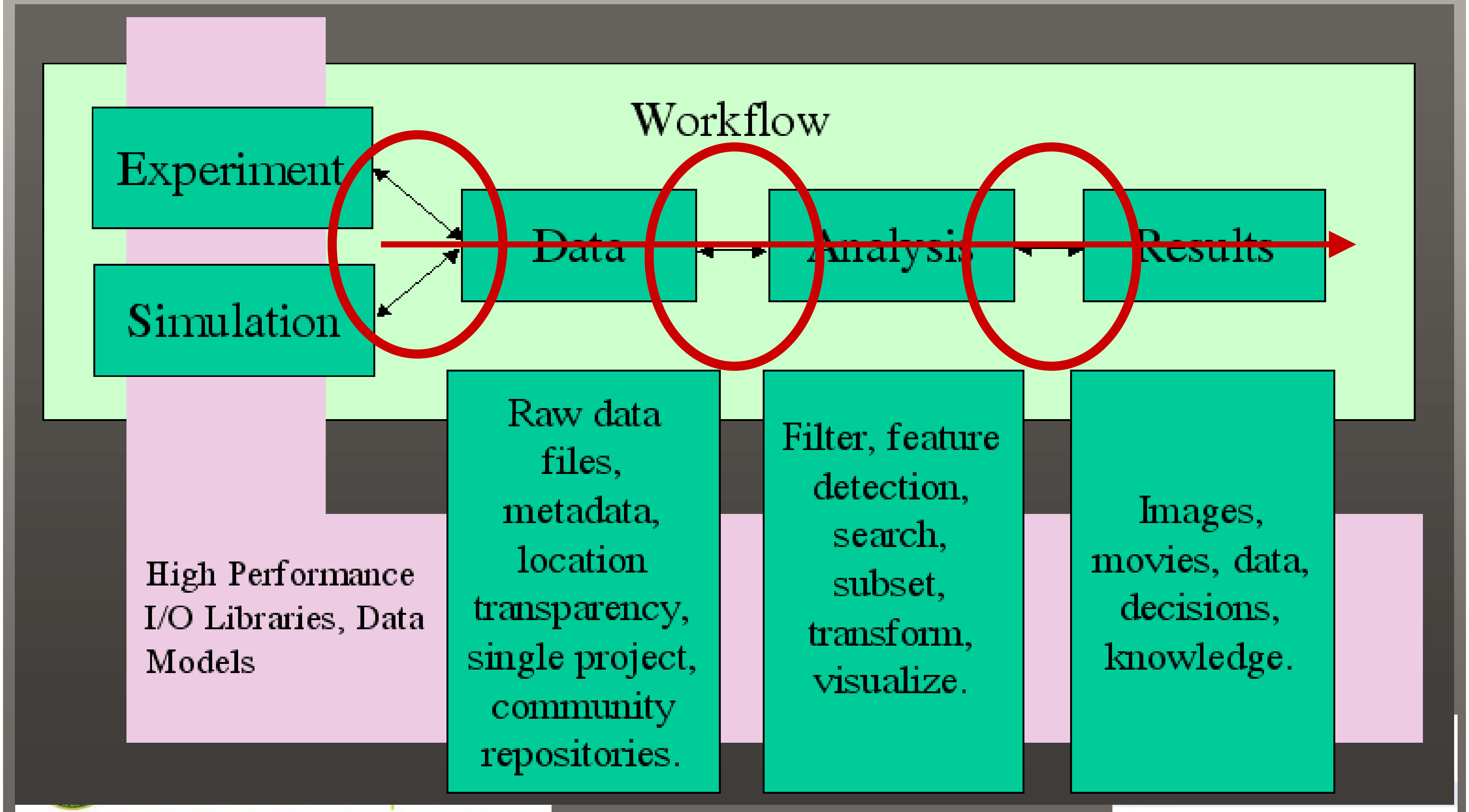
San Francisco, CA

E. Wes Bethel

Lawrence Berkeley National Laboratory


http://www.oceangrafix.com/o.g/Charts/chartViewer.html?viewRe
Region: Pacific Coast
Chart: 18022 - San Diego to San
SOUNDINGS IN FATHOMS
PAA

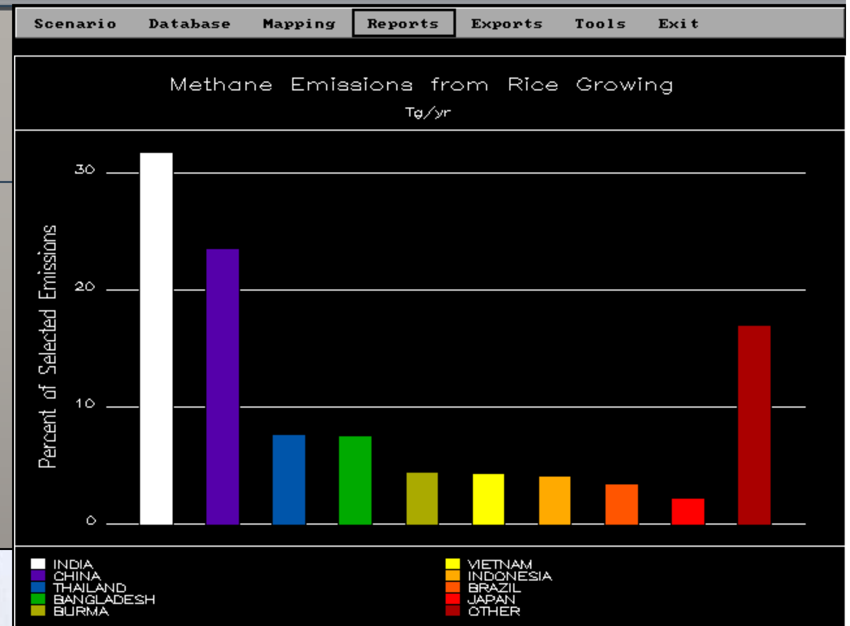
Context


Workflow
Experiment
Simulation
Data
Analysis
Results
High Performance I/O Libraries, Data Models
Raw data files, metadata, location transparency, single project, community repositories.
Filter, feature detection, search, subset, transform, visualize.
Images, movies, data, decisions, knowledge.

Rocks, Shoals, Wrecks, and Other Hazards

- Data: size, complexity, I/O, formats, etc.
  - It takes a long time to read, write big data.
  - Incompatible formats cause big problems.
- Working with big data: visual data analysis.
  - Can you run a 1TB file through gnuplot or IDL?
  - Does gnuplot or IDL really do what you need?

Remember When: 1981



This is no joke!

Data Problems

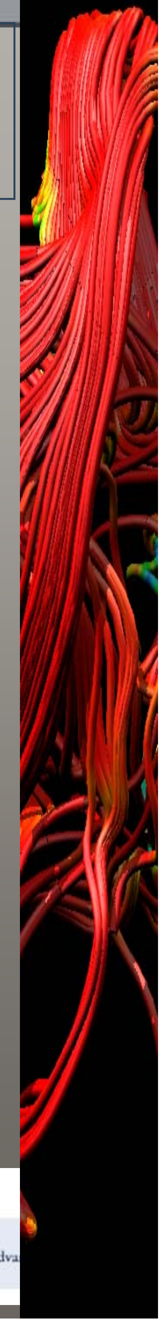
- Serial vs. parallel I/O.
  - One vs. many write streams.
- Formats:
  - How data is written out to disk: what order, storage format, etc.
  - ASCII (ouch) vs. <many options>
  - Want: format compatibility along the tool chain.

Format Propagation Issues

- What happens if each application in a tool chain uses its own unique data model/format?
- What if one or more formats changes during a weekend coding session?
- What if you want to look at results from a few years ago?
- What if you want to share results with your colleagues?

Data Format Solutions

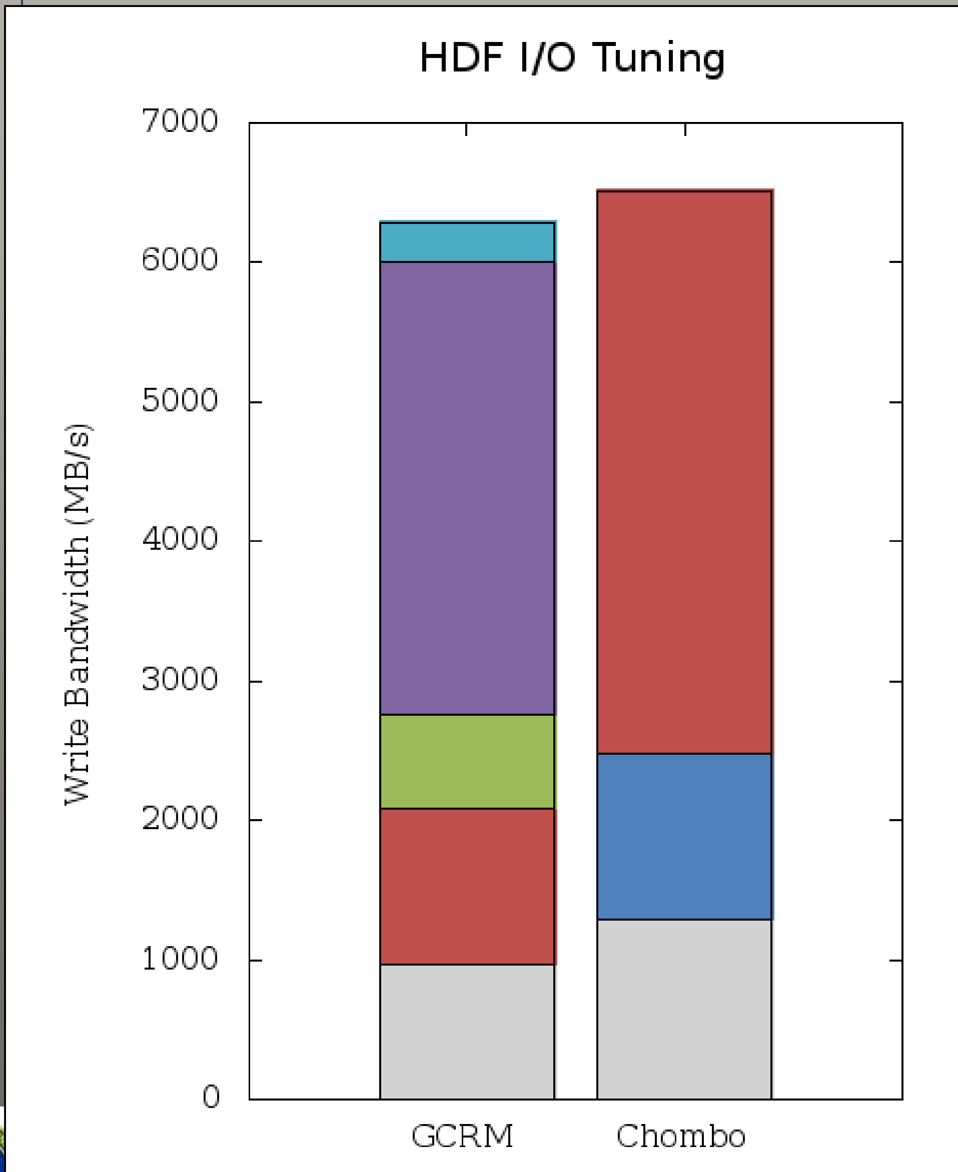
- HDF, netCDF: partial solution (why partial?)
  - Data layout inside HDF5 file: your choice.
  - Data group naming inside HDF5 file: your choice.
- H5part: more complete solution.
  - What is H5part?
    - Veneer API sits atop HDF5 (LBNL+PSI effort)
    - Simplifies use of HDF5.
  - Opaque group naming.
  - Layout defined, managed by H5part.
  - Open Source, see vis.lbl.gov



U.S. DEPARTMENT OF ENERGY | Office of Science

SciDAC Scientific Discovery through Adva

VACET



ENERGY | Science

SciDAC
Scientific Discovery through Advan

VACET

Rocks, Shoals, Wrecks, and Other Hazards

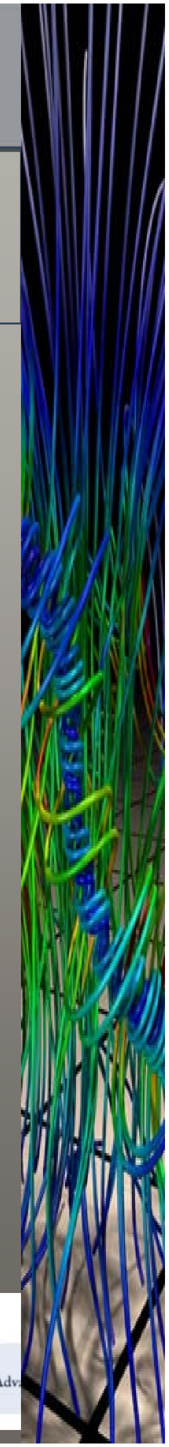
- Data: size, complexity, I/O, formats, etc.
  - It takes a long time to read, write big data.
  - Incompatible formats cause big problems.
- Working with big data: visual data analysis.
  - Can you run a 1TB file through gnuplot or IDL?
  - Does gnuplot or IDL really do what you need?

Big Problem – Information Overload

- Our ability to create and store information exceeds our capacity to understand it.
- Information requires attention to process:
  - "A wealth of information creates a poverty of attention." – Hebert Simon, Nobel Prize, 1971.
- Major challenge: gain insight from data.
  - Visualization, visual data analysis are excellent tools for accomplishing this objective.

Query-Driven Visualization

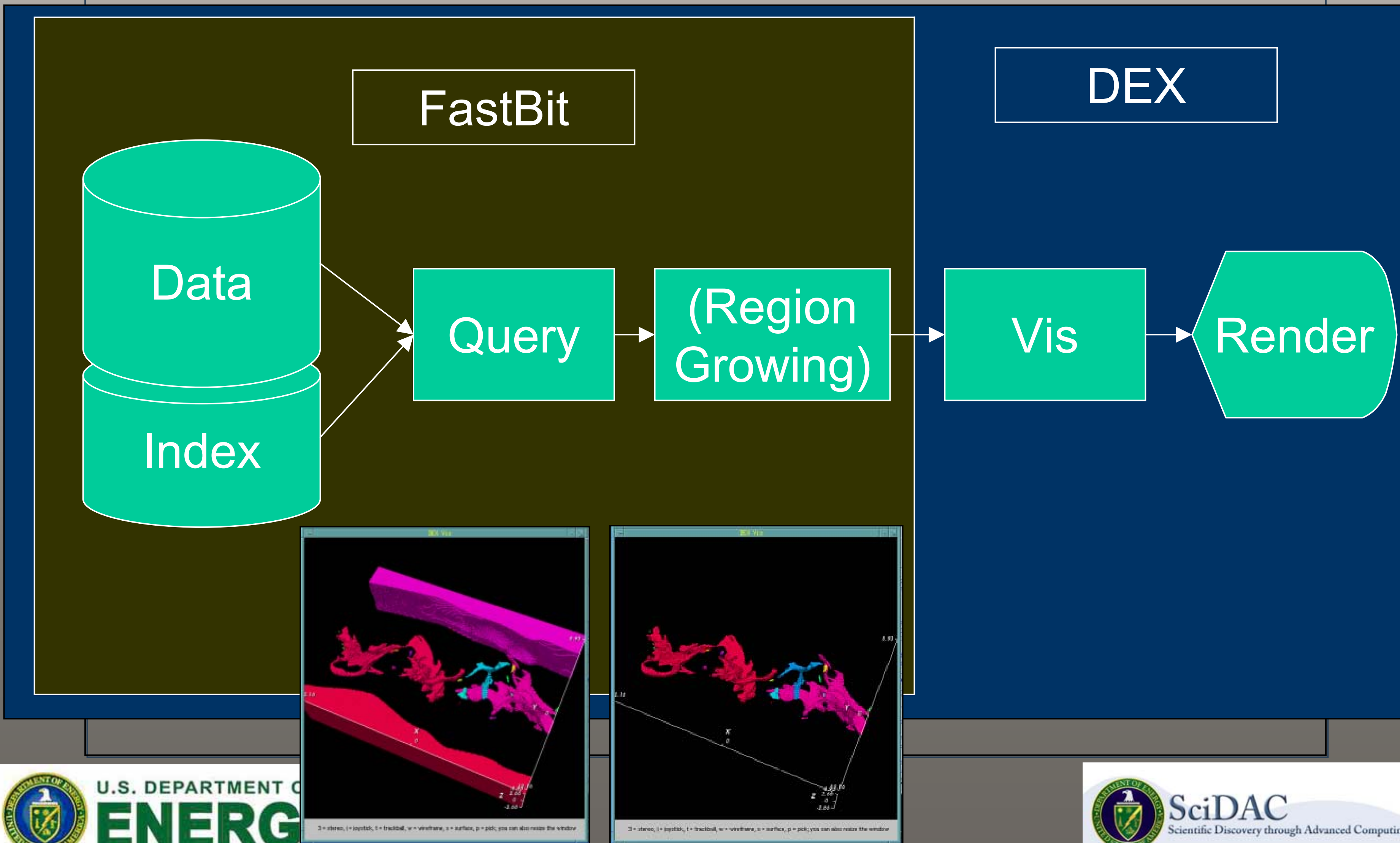
- What is Query-Driven Visualization?
  - Find “interesting data” and limit visualization, analysis, machine and cognitive processing to that subset.
- One way to define “interesting” is with compound boolean range queries.
  - E.g., (CH_4 > 0.1) AND (T_1 < *temp* < T_2)
- Quickly locate those data that are “interesting.”
- Pass results along to visualization and analysis pipeline.
- Another view: “remove the haystack to see needles.”



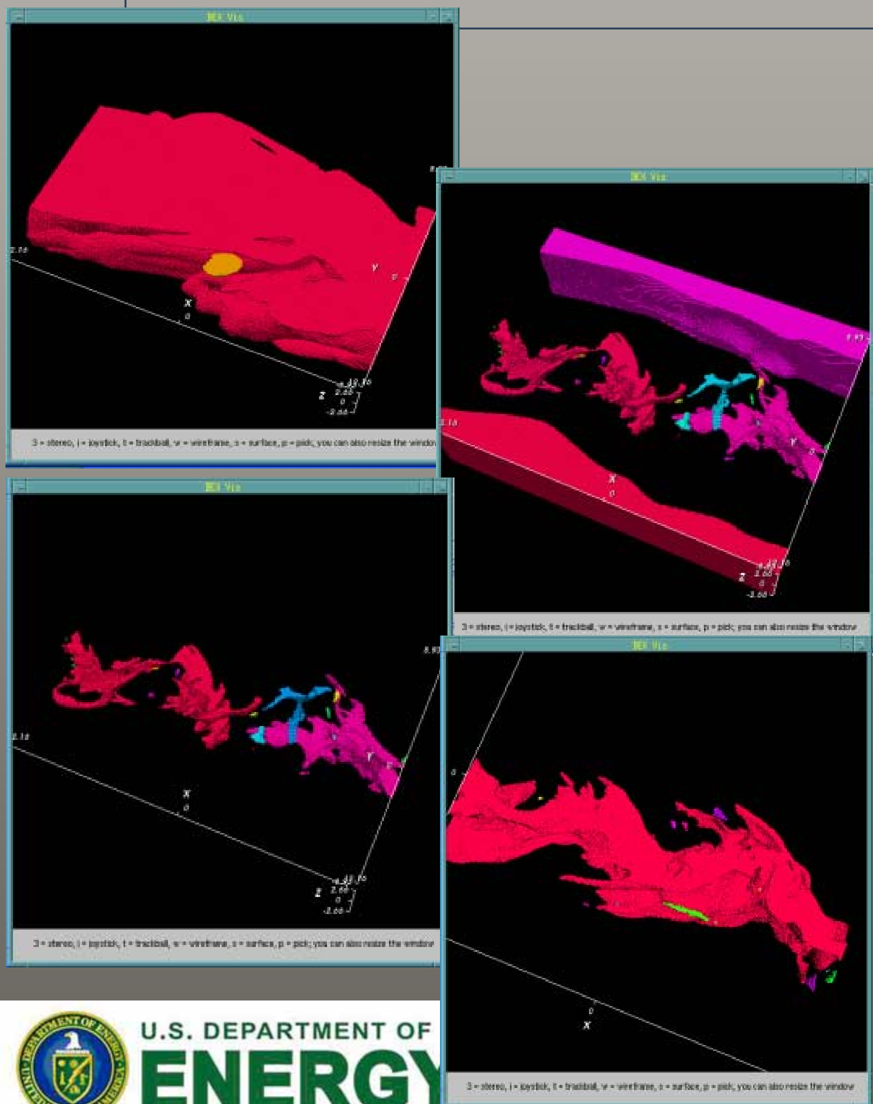
Query-Driven Visualization

The Canonical Visualization Pipeline

Query-Driven Visualization



Query-Driven Visualization



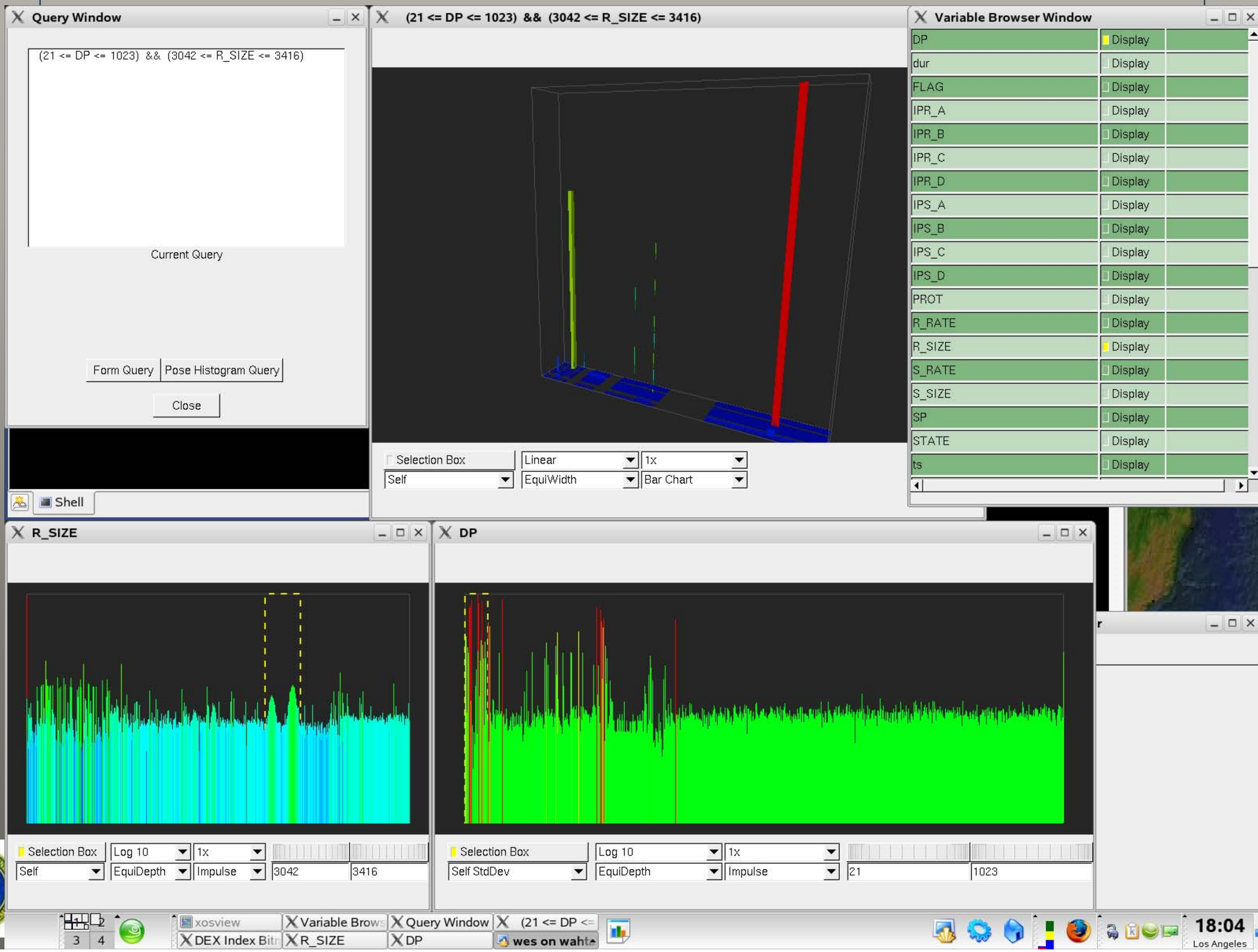
- $CH_4 > 0.3$
- Temp $< T_1$
- $CH_4 > 0.3$ AND temp $< T_1$
- $CH_4 > 0.3$ AND temp $< T_2$
  - $T_1 < T_2$

VACET

Query-Driven Visualization

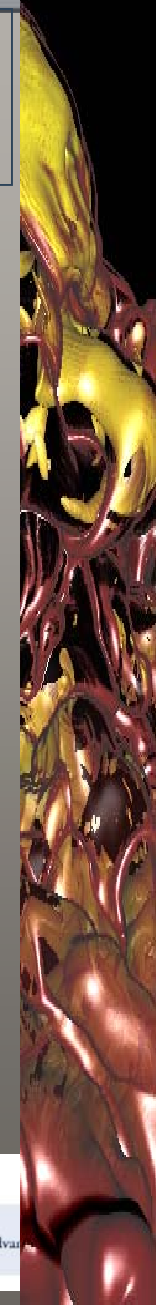
- Compare performance to isocontouring.
- For n data values and k cells intersecting the surface:
    - Marching Cubes: $O(n)$
    - Octtree methods: $O(k + k \log (n/k))$
        - – Acceleration: pruning; sensitive to noisy data
    - Span-space methods:
        - – NOISE: $O(\mathrm{sqrt}(n) + k)$
        - – ISSUE: $O(\log (n/L) + \mathrm{sqrt}(n)/L + k)$
            - » L is a tunable parameter
        - – Interval Tree: $O(\log n + k)$
- FastBit: <u>O(k) – the theoretical optimum.</u>
    - – Profound performance gain for Petascale visualization!
- Our approach supports multidimensional queries
    - – Isocontouring is essentially a 1D query

QDV Interfaces

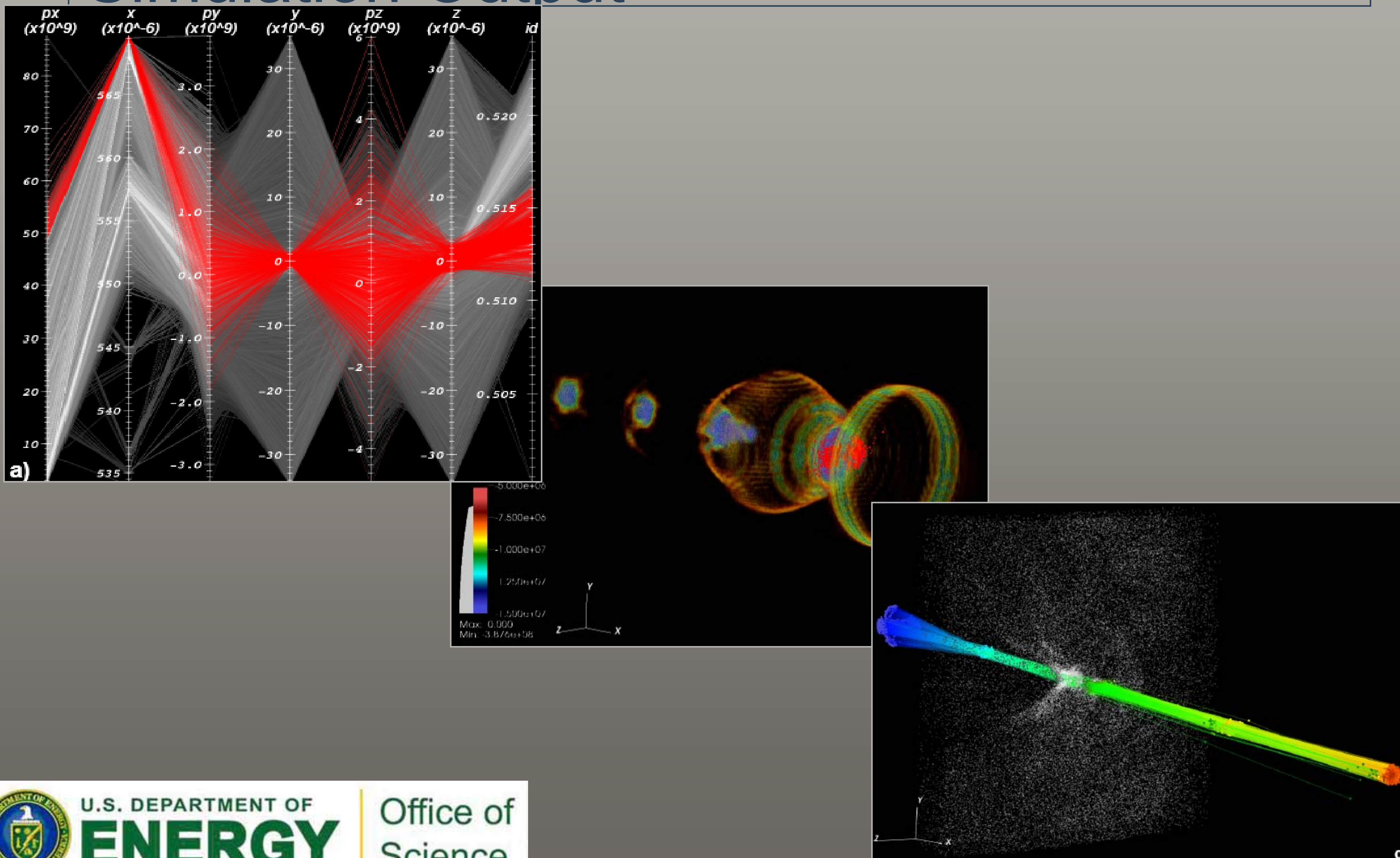


Query-Driven Visual Data Analysis Challenges

- How to define "interesting?"
- Effective interfaces for:
  - Supporting rapid interrogation, propagating query results from step to step in the analysis process.
  - Multivariate visualization
  - Drill-down (mining), linked/correlated views
- Adapting, applying and deploying these principles to many types of scientific data.
- Data file/format challenges.

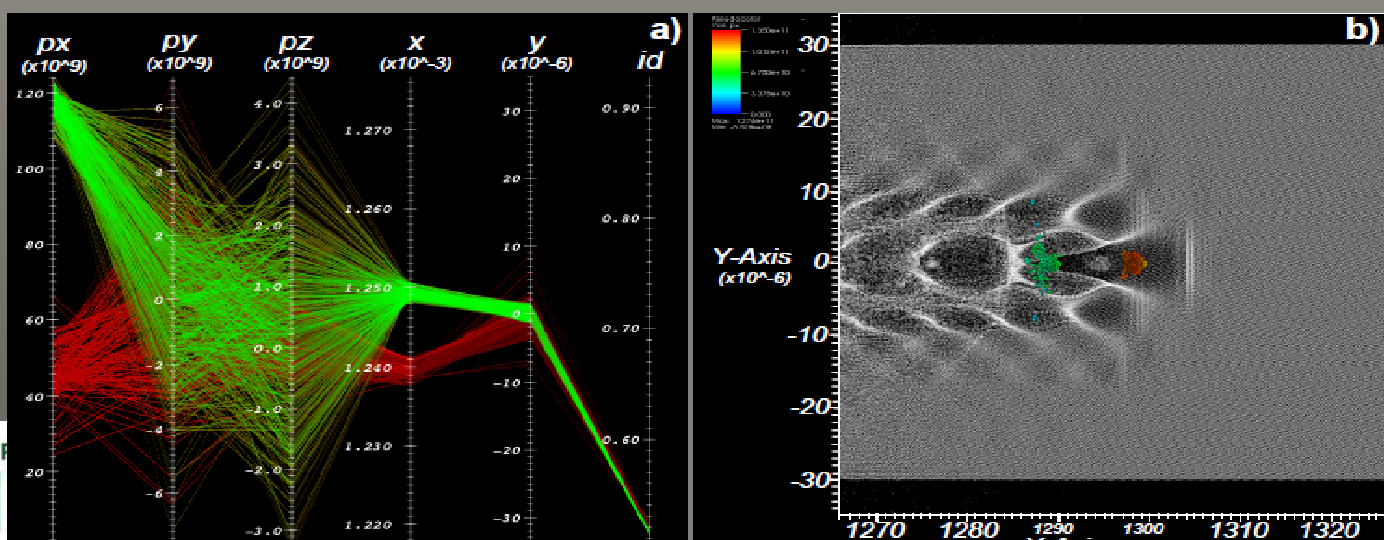


Visual Data Exploration of LWFA Simulation Output



Analysis Task(s)

1. Identify particles that form a beam
    - Interactive visual data exploration
    - Data subsetting: high energy, spatial coherency.
2. Track them over time
    - Given particle ID's from a given time step,
    - Find all those particles in all time steps
    - Subsequent visual data analysis.

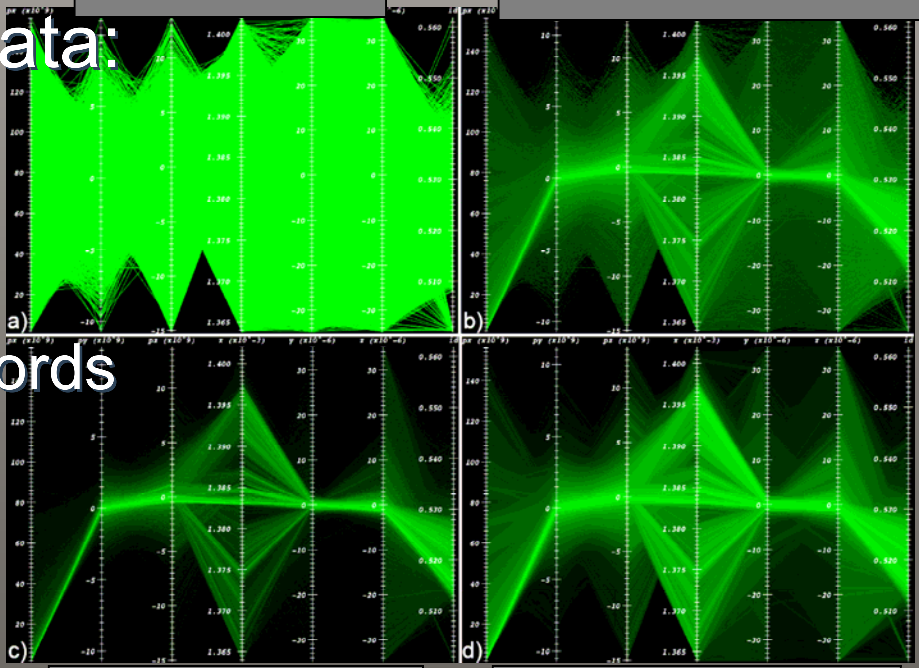


Data Overview

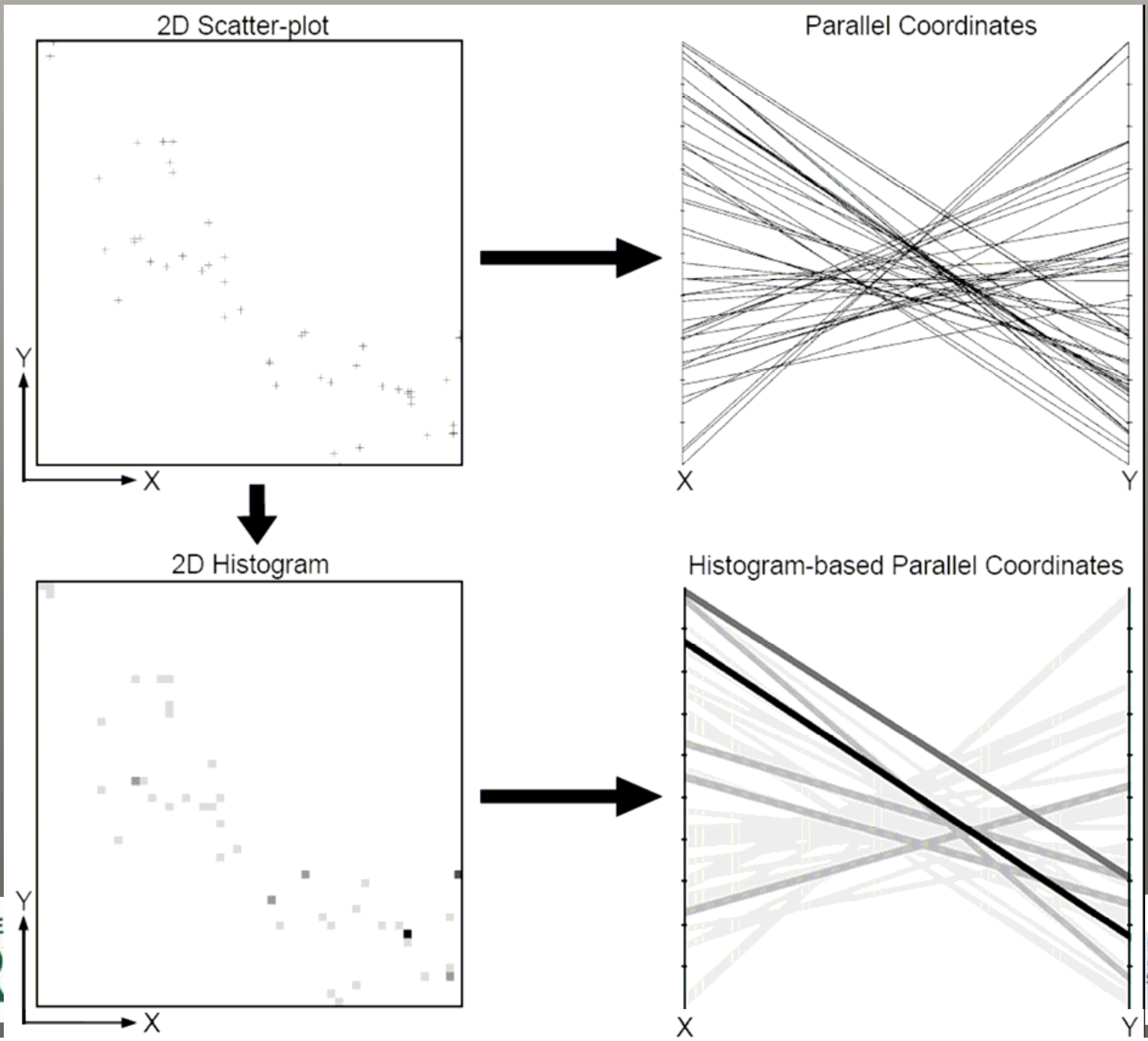
- Simulation: VORPAL, 2D and 3D.
- Particle data:
  - X,y,z (location), px,py,pz (momentum), id.
  - No. of particles per timestep: ~ $0.4*10^6$ – $30*10^6$ (in 2D) and ~$80*10^6$– 200 $*10^6$ (in 3D)
  - Total size: ~1.5GB – >30GB (in 2D) and ~100GB – >1TB (in 3D)
- Field data:
  - Electric, magnetic fields, RhoJ
  - Resolution: Typically ~0.02-0.03µm longitudinally, and ~ 0.1-0.2µm transversely
  - Total size: ~3.5GB - >70GB (in 2D) and ~200GB - >2TB (in 3D)

Fundamental Problem #1 - Interface

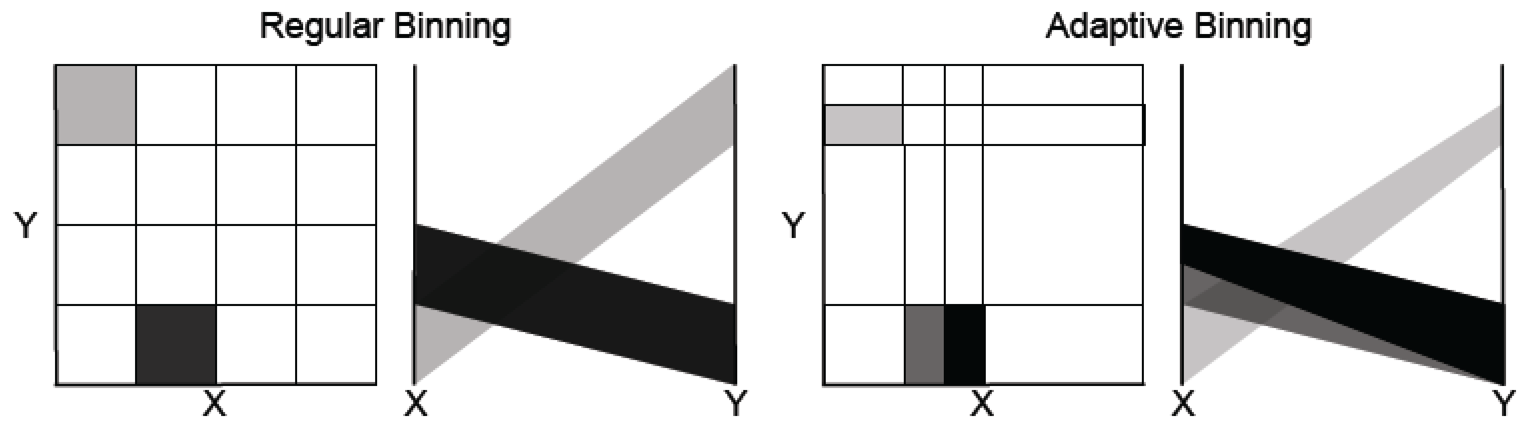
- Parallel coordinates
  - An interface for subset selection.
  - A mechanism for displaying multivariate data.
- Problems with large data:
  - Visual clutter
  - O(n) complexity
- Solution/Approach
  - Histogram-based p-coords

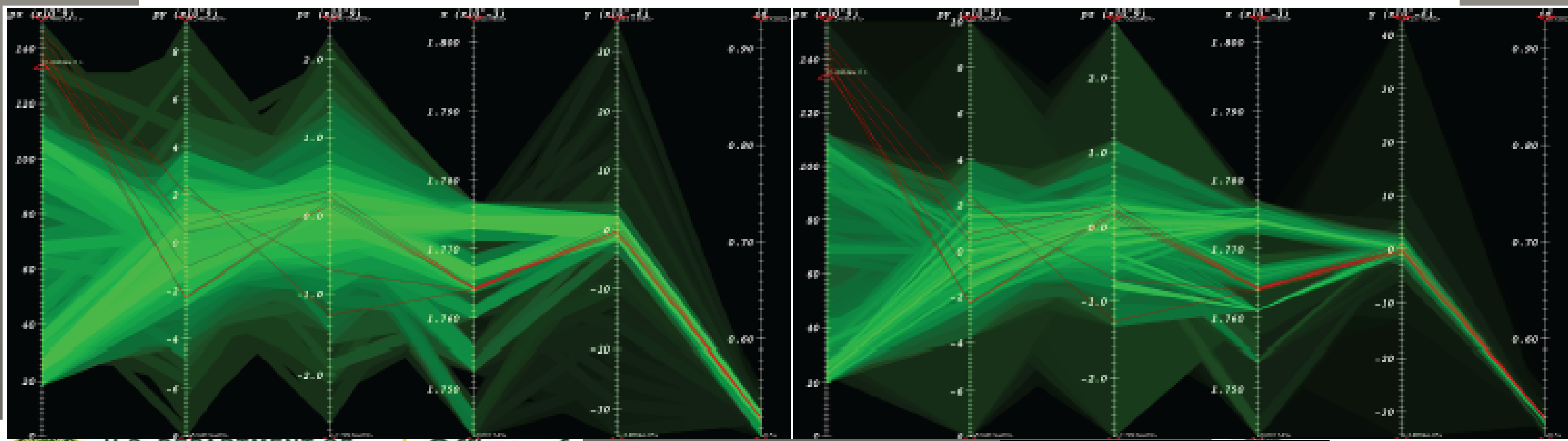


Histogram-Based Parallel Coordinates



Adaptive and Constant-sized Bins

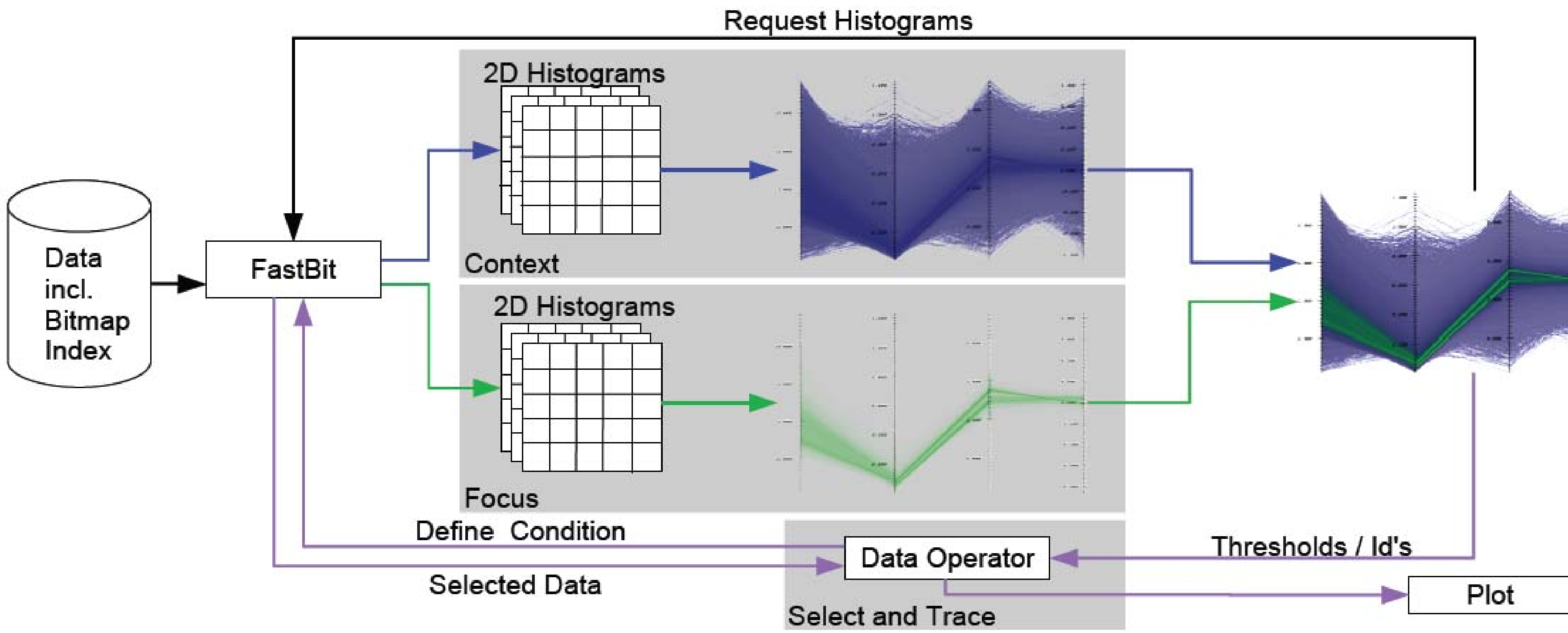




32x32 uniform

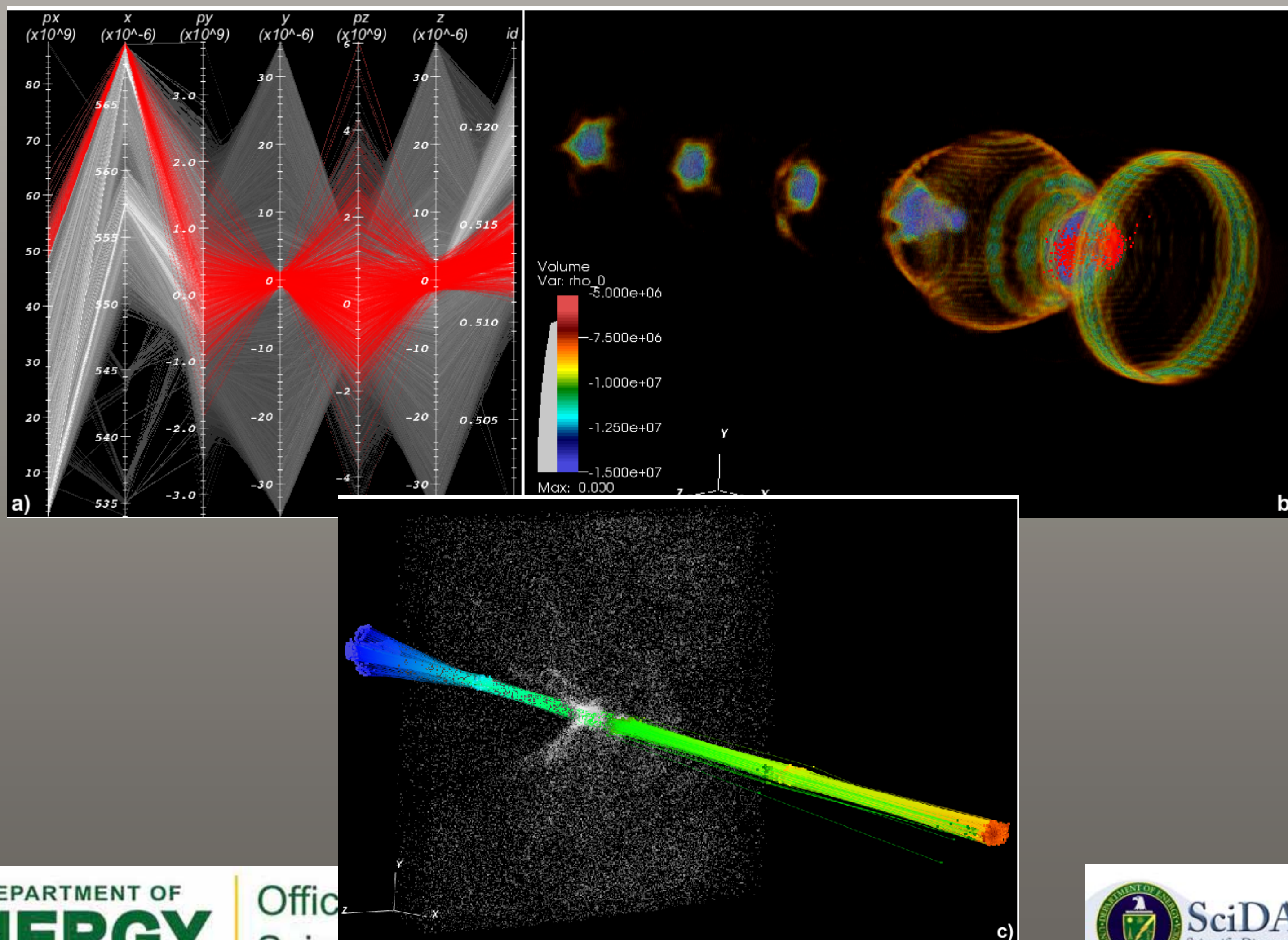
32x32 adaptive

System Overview



VACET

3D Example

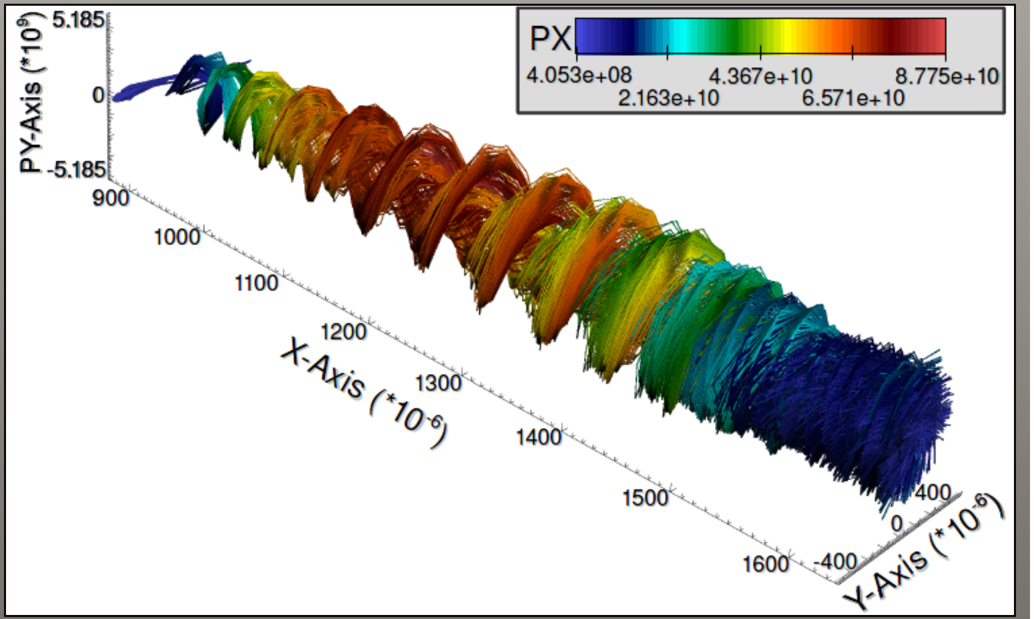


U.S. DEPARTMENT OF ENERGY | Office of Science

SciDAC Scientific Discovery through Advanced Computing

More Recent Results

- Understanding particle behavior over time:
  - After finding interesting particles and tracing them through time,
  - Particles start out slow (blue, left), undergo acceleration (reds), then slow again as the plasma wave outruns them (blue, right).
  - Spiral structure shows particles oscillating transversely in the focusing field.



U.S. DEPARTMENT OF ENERGY | Office of Science

SciDAC Scientific Discovery through Advanced Computing

Fundamental Problem #2 – Performance

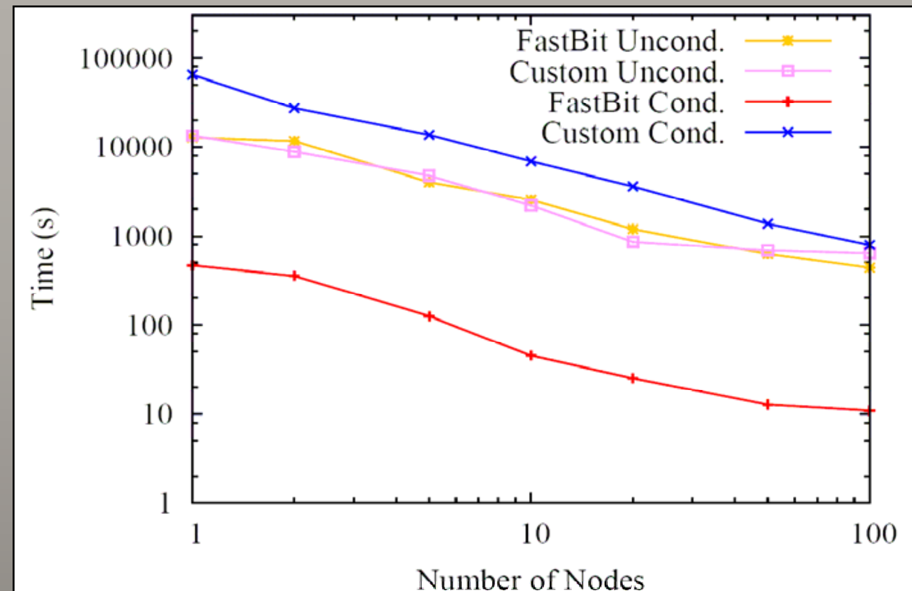
- How to efficiently construct a histogram?
  - Naïve approach: O(n)
  - Better approach: use FastBit
- How to efficiently do particle tracking?
  - Naïve approach: $O(n^2)$
  - Better approach: O(H*t) (use FastBit)

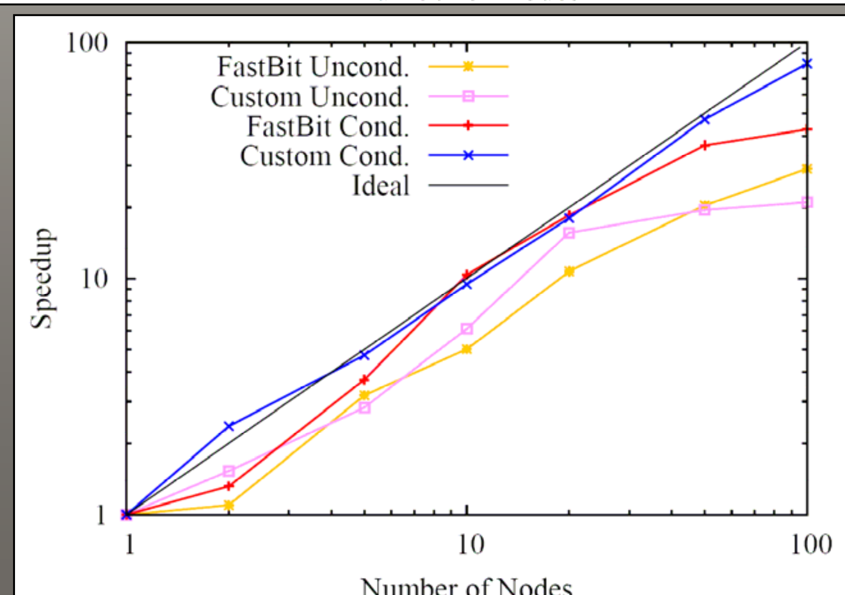
Parallel Performance I: Histograms

Dataset:
- 3D dataset consisting of 100 timesteps
- ~177 million particles per timestep
- ~10 GB per timestep
- ~1TB total size

Test platform: (as of July.2008)
- franklin.nersc.gov
- 9,660 nodes, 19K cores Cray XT4 system
- Filesystem: Lustre Parallel Filesystem
- Each node consists of:
  - CPU: 2.6 GHz, dual-core AMD Opteron
  - Memory: 4GB
  - OS: Compute Node Linux

Test setup:
- Restrict operations to a single core of each node to maximize I/O bandwidth available to each process
- Assign data subsets corresponding to individual timesteps to individual nodes for processing
- Generate five 1024x1024 histograms for position and momentum fields at each timestep
- Conditon: $px>7*10^{10}$
- Levels of parallelism: 1, 2, 5, 10, 20, 50, 100



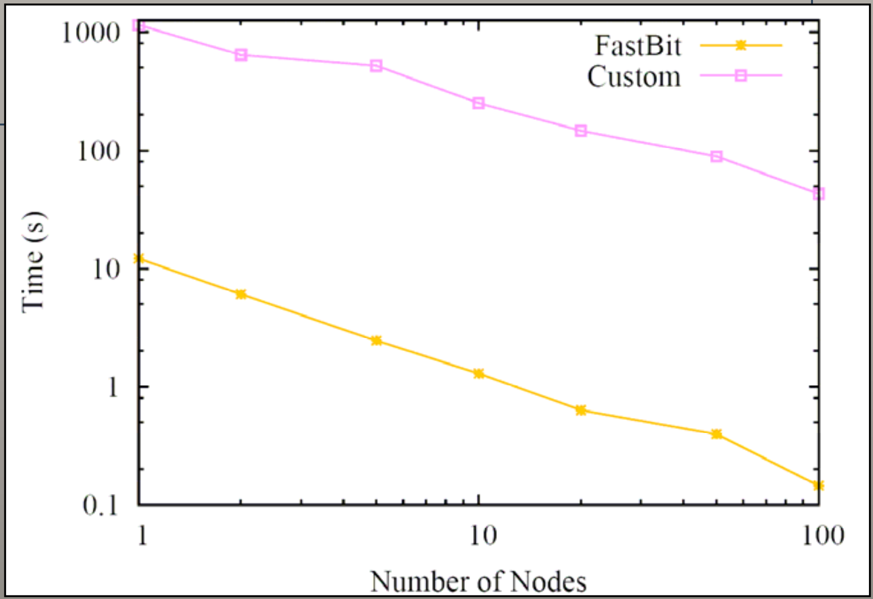


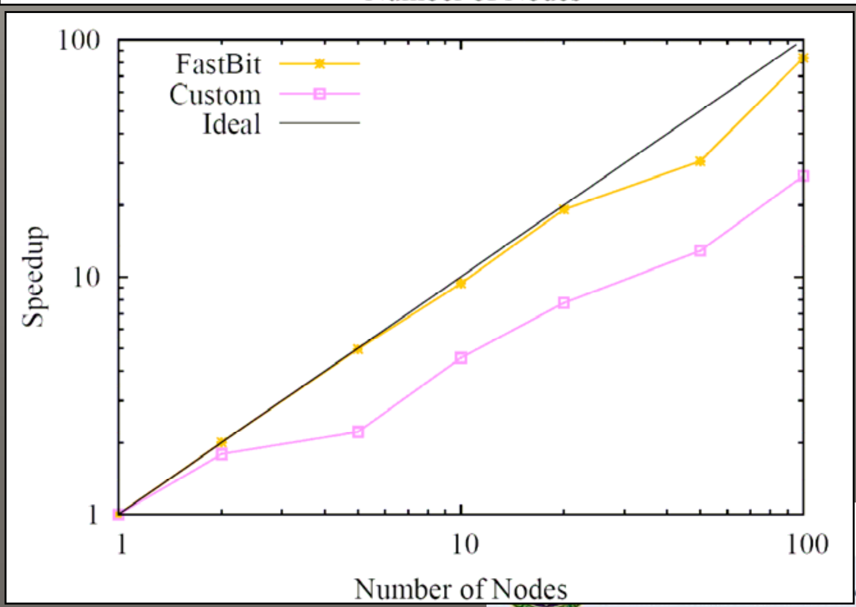
Parallel Performance II: Particle Tracking

Test setup:
- Same as for histogram computation
- Track 500 particles (Condition: $px>10^{11}$) over 100 timesteps

Results:
- FastBit is able to track 500 particles over 1.5TB of data in 0.15 seconds

Performance of original IDL scripts:
- ~2.5 hours to track 250 particles in small 5GB dataset





U.S. DEPARTMENT OF ENERGY | Office of Science

More Than Just a Research Project

- Several technologies from this project have been "productized" in VisIt and are available to "the entire world."
  - Parallel coordinates interface (traditional and histogram-based)
  - H5part, FastBit-enabled file loader to support parallel collective I/O, including index/query.
  - ID-based, or "named" queries.

Concluding Remarks

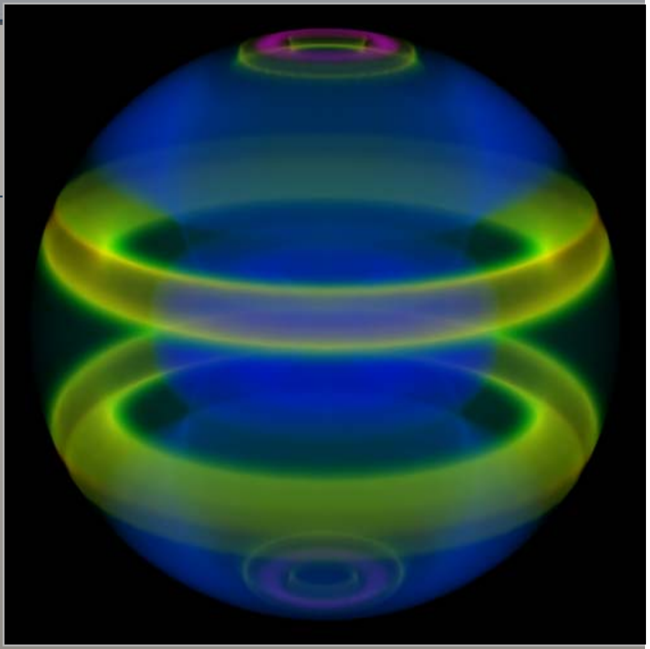
U.S. DEPARTMENT OF ENERGY | Office of Science

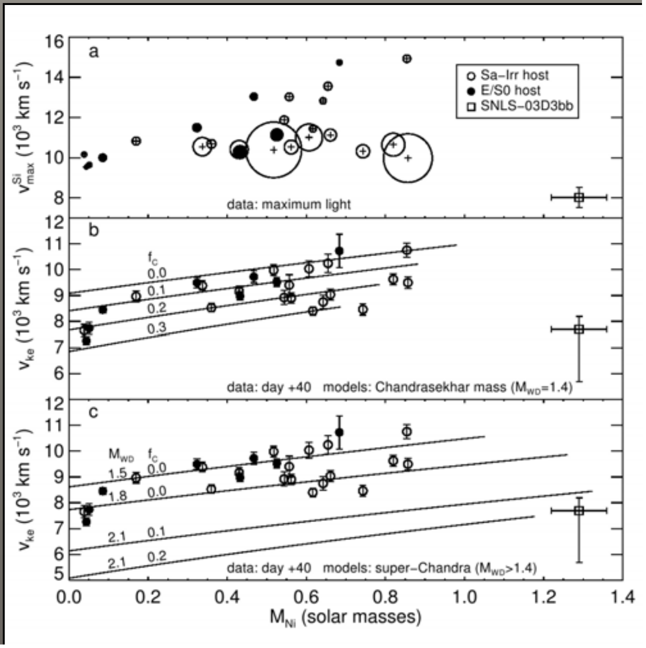
SciDAC Scientific Discovery through Advanced Computing

VACET

Visualization Use Models

- Presentation visualization
  - You know what's there and want to show it to someone else
- Analytical Visualization
  - You know what you are looking for
- Discovery Visualization
  - You have no idea what you're looking for





U.S. DEPARTMENT OF ENERGY | Office of Science

Hazards at PScale and Beyond

- Computing hazards: out of scope for this talk.
  - E.g., solvers, multicore, 10M-100M cores, programming and execution models, etc.
- I/O hazards:
  - Serial vs. parallel I/O
  - Data models and formats.
- Visual data analysis hazards
  - What problem are you trying to solve?
  - Sufficiently capable tools?
  - Effective tools?
  - I/O issues, data duplication?

VACET

The End