The Automated Event Service: Efficient and Flexible Searching for Earth Science Phenomena

Tom Clune

Computational & Information Sciences & Technology

NASA Goddard Space Flight Center

AES Extended Team

AES: K.-S. Kuo[1,2], J.A. Rushing[3], R. Ramachandran[4], A. Lin[3], G. Fekete[5,1], K. Doan[9,1], R. Tucker[3]

PROBE: K.-S. Kuo[1,2], M. Bauer[8,1], G. Schmidt[1], A. Oloso[6,1], G. Fekete[5,1]

AWS/MODB: K.-S. Kuo[1,2], M. Schneider[10], R. Linan[7,1], A. Oloso[6,1]

1. NASA GSFC
2. Bayesics, LLC
3. University of Alabama-Huntsville
4. NASA MSFC
5. Computer Science Corporation
6. Science Systems and Applications, Inc.
7. Navteca Inc.
8. Columbia University
9. University of Maryland
10. University of Florida

Everyone talks about Big Data, but no one does anything about it.

– paraphrased from Charles Dudley Warner

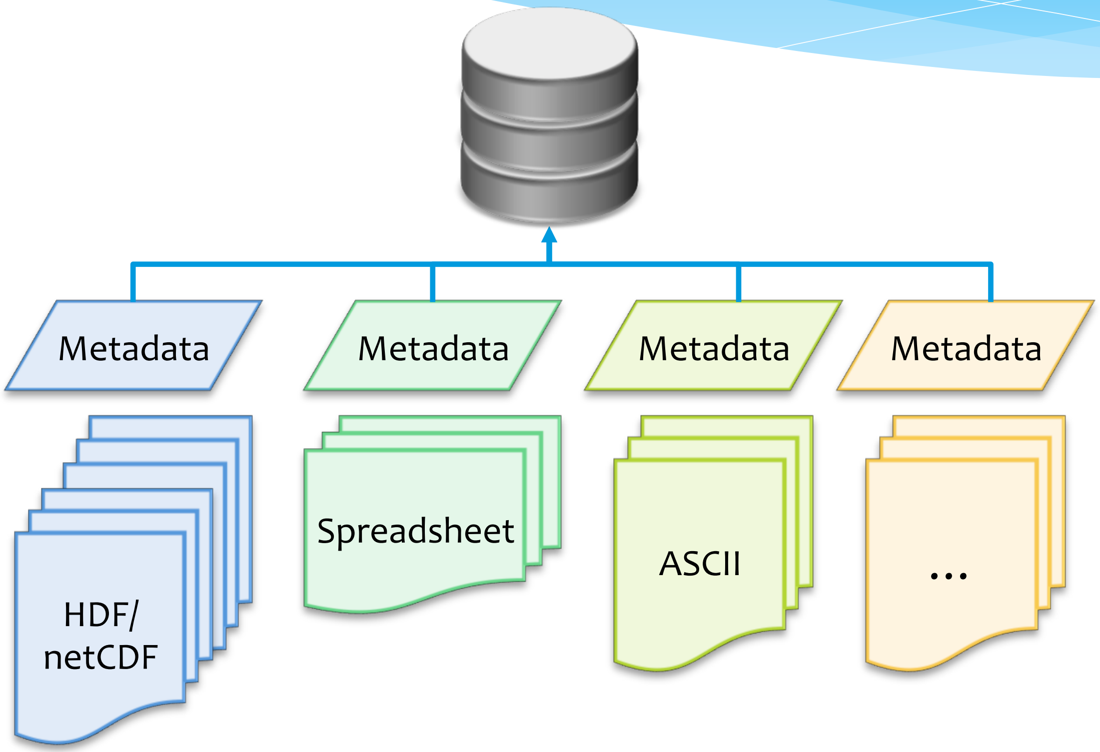
Automated Event[1] Service

ESTO (AIST) project now in 3rd year of funding.

- Enable **systematic** identification of **investigator-defined** Earth science events from reanalysis and satellite data.
  - Addressing a significant portion of ES research;
  - Reducing duplication of effort among research teams;
  - Improving return on investment (ROI) for NASA data and compute resources.
- Provide driver to improve affinity of computing and data resources
  - Move computing to the data rather than data to computing.
- Greatly improve interactive data exploration and analysis.

[1] Events are occurrences of phenomena, usually 4D (space and time) in nature.

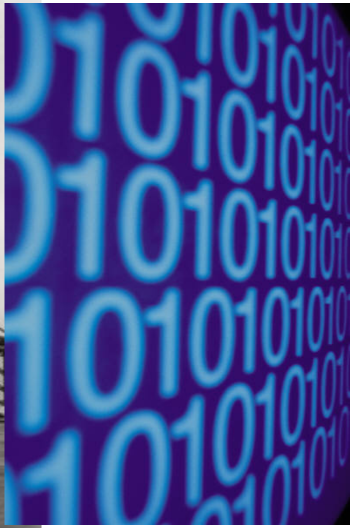
Current Data Archive and Distribution Practices



- File based archive distribution.
- Only metadata are cataloged in databases, thus searchable.

Gap between HPC and Regular User

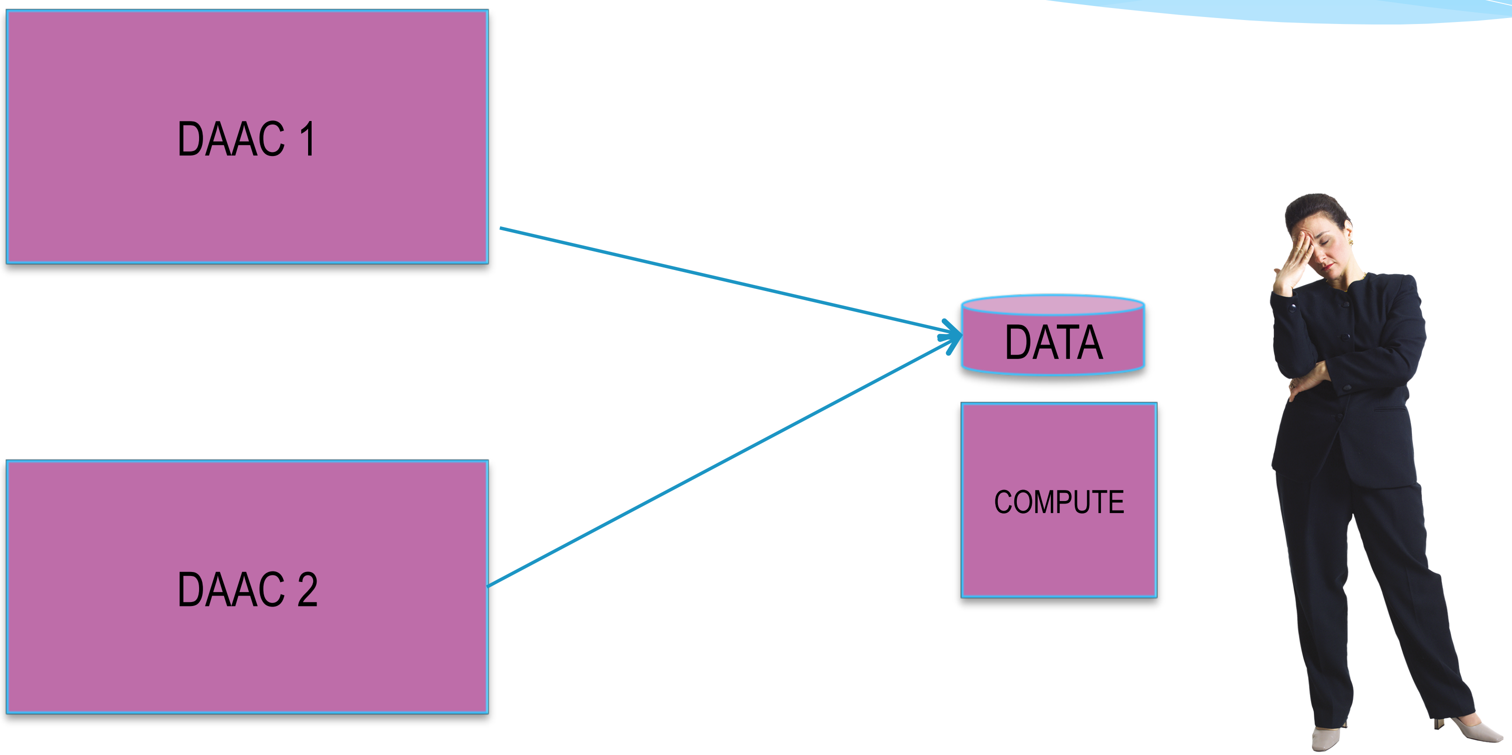
Researcher



High Performance Computing/ High Performance Data

Our solution: Automated Event Services (AES)

What is Happening Now?



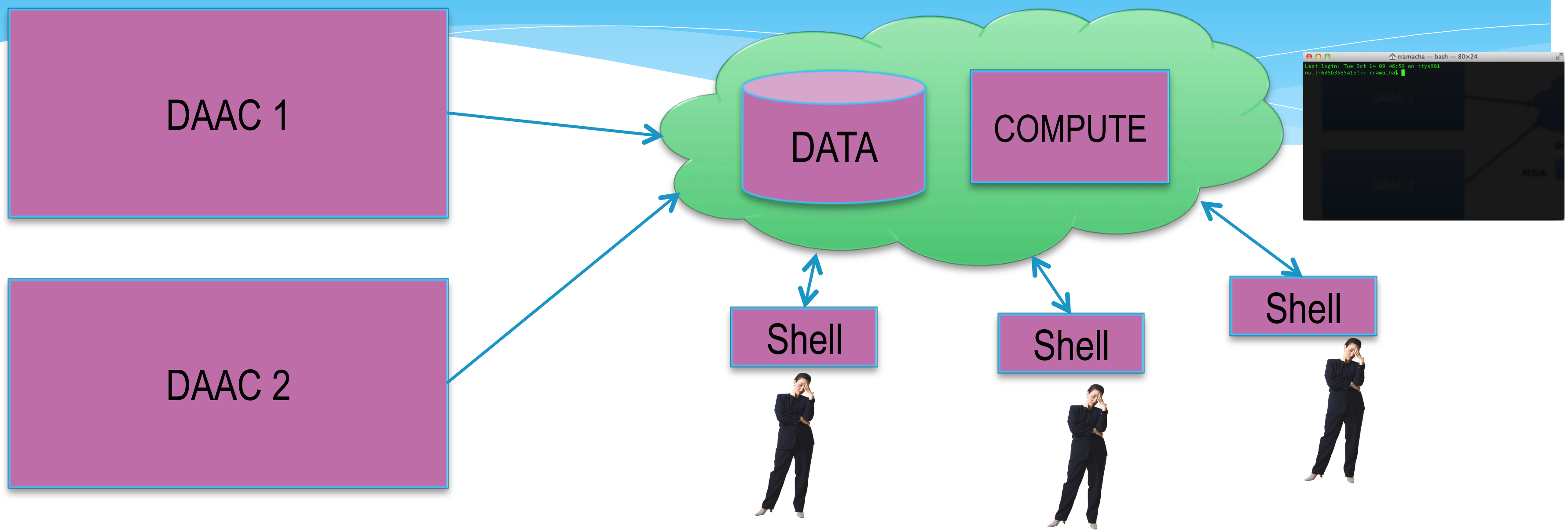
Remote Resources

Local Resources

Results of the Current Paradigm

- File- or granule-level discovery and access:
  - Search metadata store to order,
  - Direct download from URL, or
  - Access within file provided through a data protocol, e.g. OPeNDAP.
- Inherently serial process
- Requiring local storage and compute
  - Researchers must engage in activities unrelated to science:
    - procurement and maintenance of storage and compute resources,
    - data management, i.e. downloading, organizing, backing up…
- Considerable duplication of efforts/resources
- Collaborations difficult.

Standard HPC/Cloud Solution



Advantages

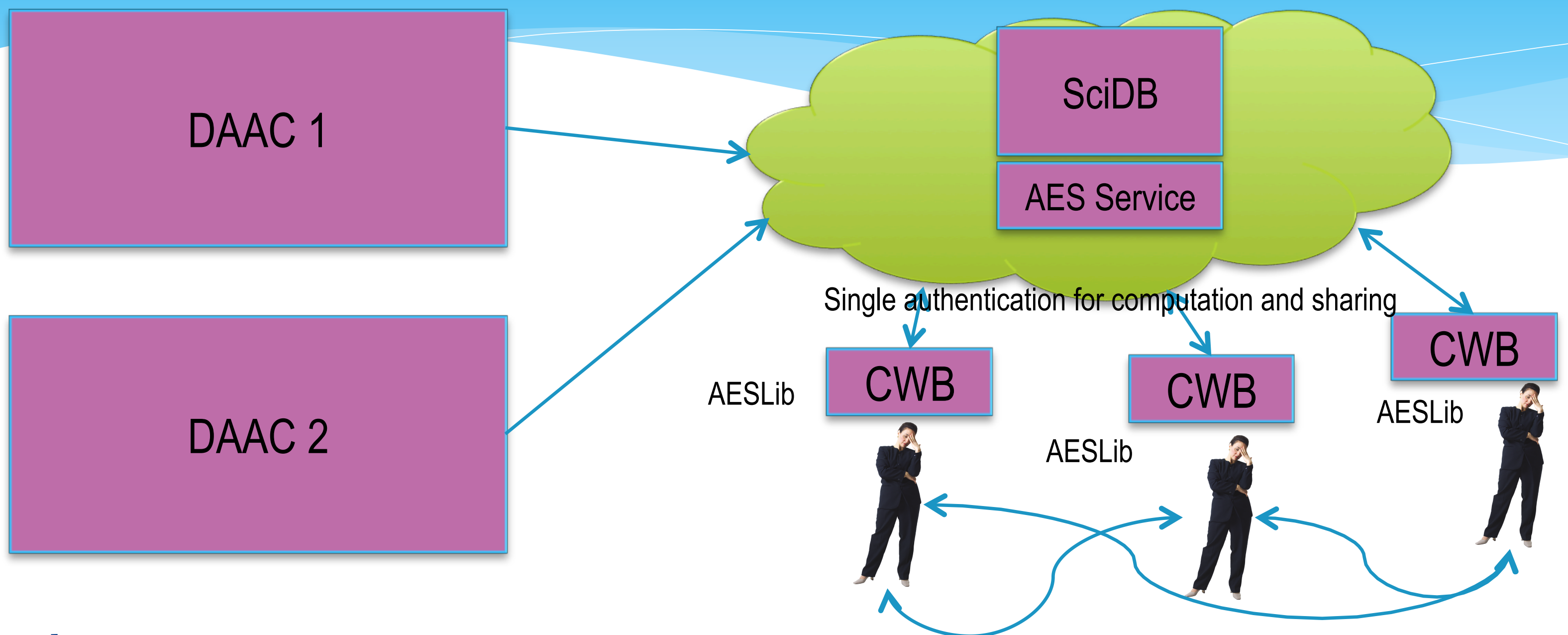
- Solves infrastructure/ redundancy problem
- Data colocated with compute
- Supports a wider range of data analysis problems

Disadvantages

- Analysis limited to data available on the cloud
- *Data is still file-based*
- *Optimization requires data movement and parallel programming*
- No sharing/collaboration

AES Vision

Focus on Event Analytics (Phenomena Detection/Characterization)



Advantages

- AES *middleware* provides parallel optimized algorithms
- Combines server side processing and local analysis providing maximum flexibility
- Collaboration achieved through CWB Integration (authentication, authorization, sharing and collaboration)

Disadvantages

- Data needs to be preloaded
- Scope of analysis *limited* to data loaded on SciDb and the algorithm set available in AES middleware

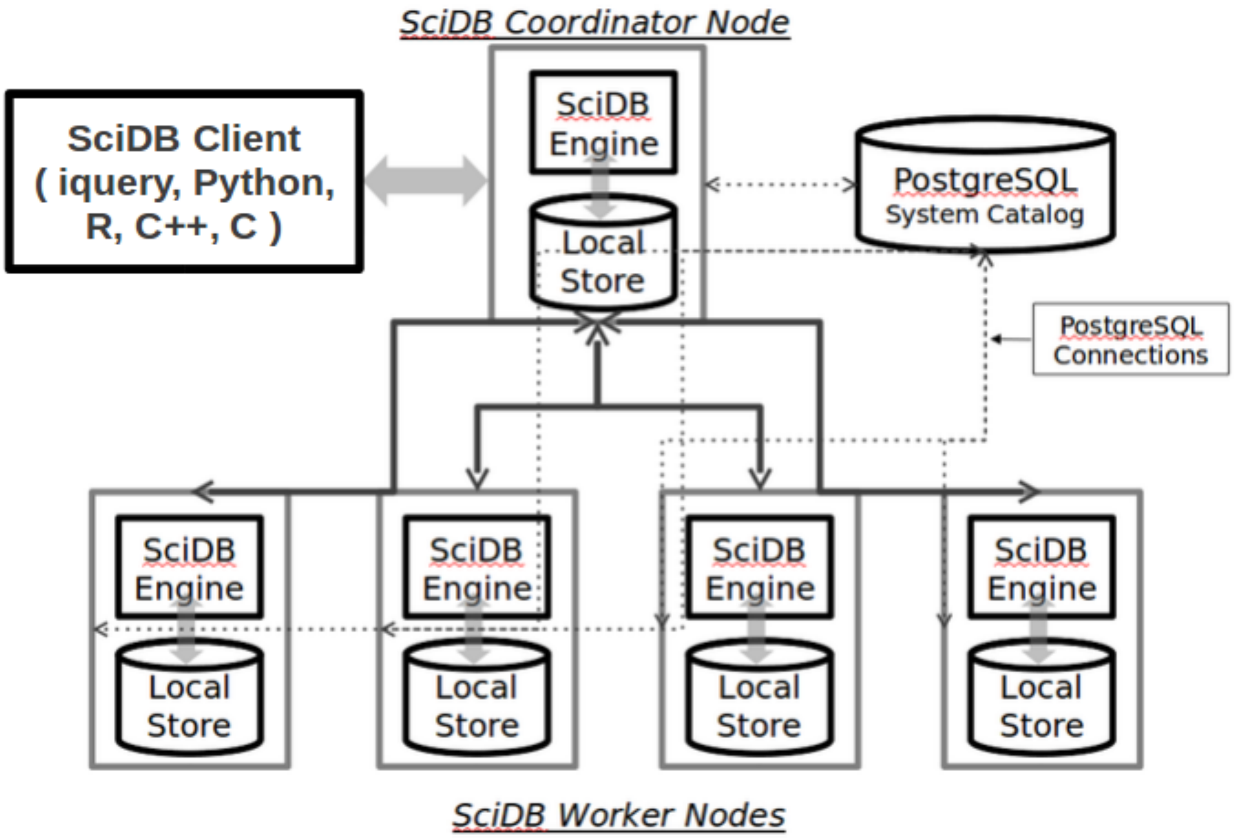
A *Big-Data* Solution: SciDB

An all-in-one ***data management*** and advanced ***analytics platform*** that features:

- Complex analytics inside a next-generation parallel ***array*** database,
  - *i.e.* not *row-based* or *column-based* like RDBMS's based on *table* data model
- Based on the "shared nothing architecture" for data parallelism,
- data versioning and provenance to support science applications, and
- Open source (currently in beta).

A better performer than Hadoop (MapReduce), 2-10 times faster, in almost all benchmarks that we have performed so far.

Basic SciDB Architecture



NCCS[1] "SciCloud" Cluster

- 36 nodes
  - 30 in a main cluster
  - 6 in a test/development cluster
- 2x8 SandyBridge Intel Cores, i.e. 16 cores per node,
- 32 GB memory per node,
- 36 TB local storage per node,
- FDR Infiniband
- Gigabit ethernet

[1] NCCS – NASA Center for Climate Simulation @ GSFC

Blizzard Science Scenario

- **Definition:** According to NWS, a blizzard means that the following conditions are expected to prevail for a period of 3 hours or longer:
  - **Sustained** wind or **frequent** gusts to 15.6 m/s (35 mph) or greater; and
  - **Considerable** falling and/or blowing snow, *i.e.*, reducing visibility **frequently** to less than 400 m (¼ mile)
- The definition contains **imprecise** adjectives/qualifiers.
- Point-based (local and/or instantaneous) definitions do not translate directly to space/time-averaged parameters as in MERRA[1] reanalysis data.
- It is obvious that **visibility** is the crucial criterion in defining blizzard, but MERRA does not yet include visibility observation.
- Visibility is directly related to in-air snow mass concentration and dependent upon snow particle mass-dimension (morphology) property.

[1] MERRA – Modern Era Retrospective analysis for Research and Applications for the satellite era using NASA's GEOS-5 model, focusing on historical analysis of the hydrological cycle. It is composed of multiple regularly gridded data sets, **~100 TB** total.

Additional Considerations

- Most of the ES phenomenon definitions are, like blizzard, not based on space/time-averages.
- It is not possible to define ES phenomena unequivocally (e.g. nothing but grid cells containing blizzard condition) using space/time-averaged data sets such as MERRA re-analysis.
- The goal thus becomes finding **the smallest possible superset** that, for example,
  - captures all blizzard grid cells and
  - minimizes the number of false-positive, non-blizzard grid cells.
- Not all fields may be available. In-air snow mass concentration is contributed primarily by
  - Falling snow – using snow rates in MERRA
  - Blowing snow – using snow accumulation on surface and wind speed at 10 m above surface as proxy
- Events found using MERRA serve as a basis to locate other useful data sets for validation or refinements.

Visibility vs. Snowfall Rate

- Rasmussen et al (1999) plot visibilities and snowfall rates
  - based on theoretical calculations of various snow crystal types, and
  - supported by observations.
- Variations among crystal types are considerable.
- The blue line is described by:

$$\log v = -\log s + 3$$

v: visibility in meter

s: snowfall rate in $\mathrm{mm\ hr^{-1}}$

- The above relation is in turn used to find extinction as a function of snowfall rate.

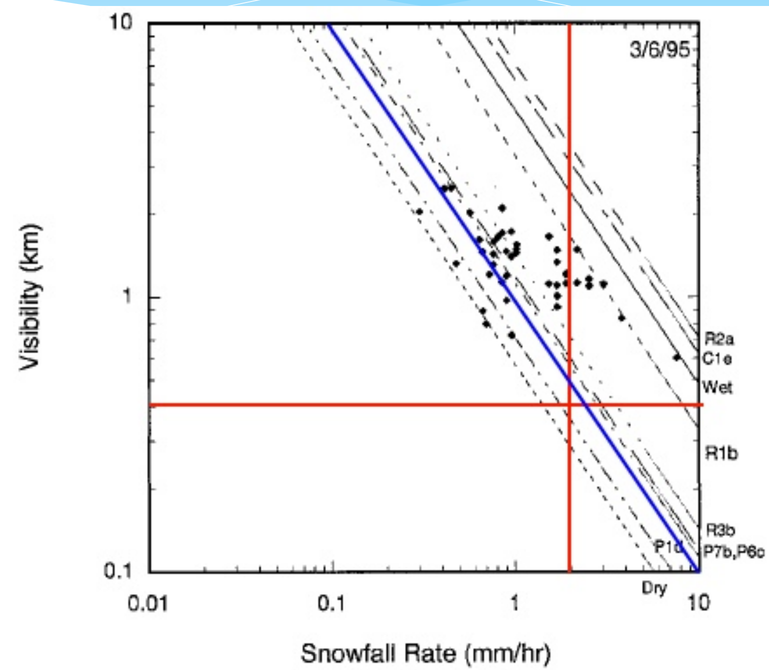


FIG. 16. Theoretical visibility–snowfall relationships from Eq. (19) compared to the observed visibility–snowfall data from 6 Mar 1995. The theoretical curves correspond to the crystal types observed for this event.

Visibility vs. Wind Speed
in Blowing Snow

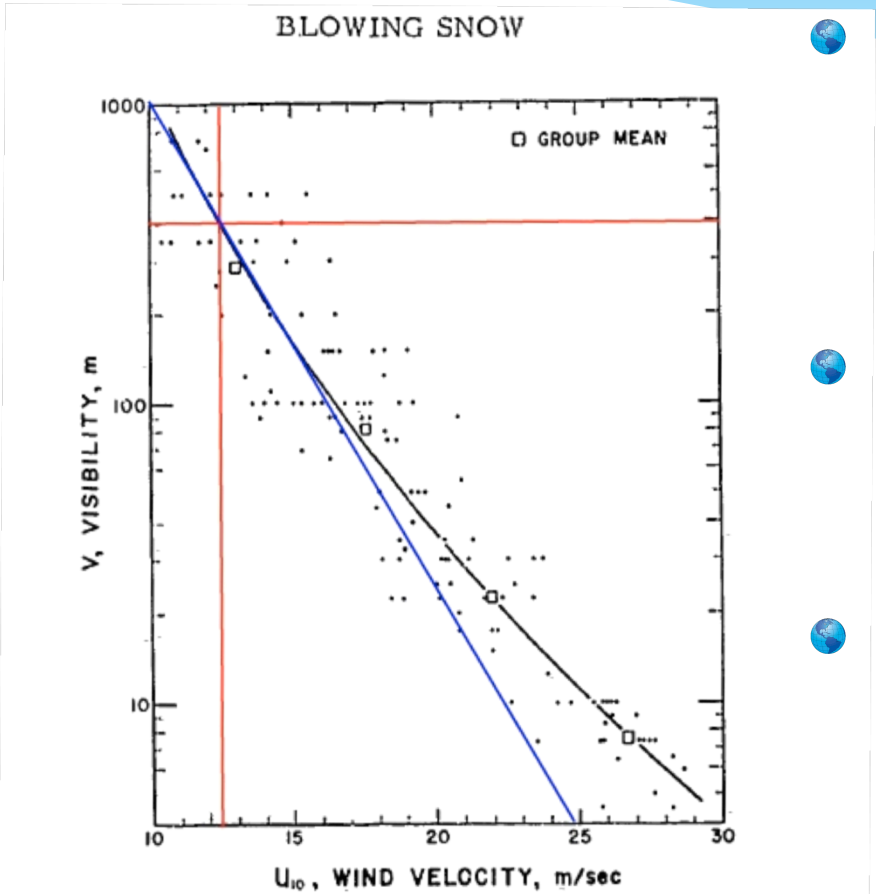


Figure 65. Logarithm of visibility as a function of wind speed at 10-m height for blowing snow conditions. (From Liljequist[53]).

- Most studies relates visibility to in-air snow mass concentration (g m^{-3}).
  - A combination of falling and blowing snow.
- Liljequist (1957) relates visibility directly to wind speed based on typical blowing snow situations in Antarctica.
- The blue line is described by

$$\log v = -0.1592w + 4.5918$$

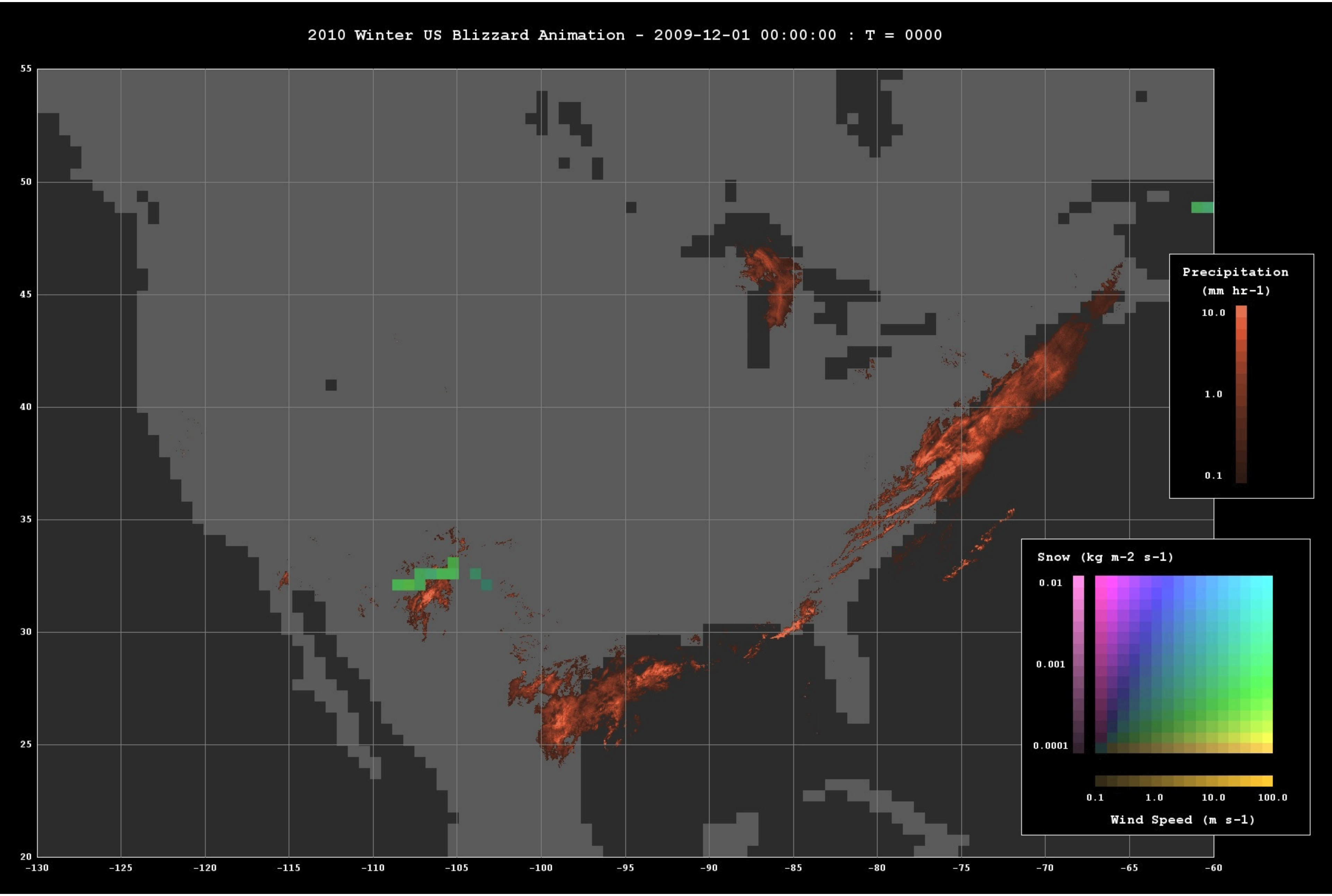
v: visibility in meter

w: wind speed 10-m above surface in m s^{-1}

Steps to Blizzard Definition using MERRA

Using MERRA **hourly** data sets

- Extract 2010 Winter US subset.
- Calculate:
  - wind speed 10-m above surface, w, using the east-west (U10M) and south-north (V10M) components,
  - grid-area-weighted mean (μ_w) and standard deviation (σ_w) of w, and
  - grid-area-weighted mean (μ_s) and standard deviation (σ_s) of **log10** snow rates in **snow-only** grids.
- Conduct trial experiments to define blizzard:
  - Experiment with snow rate threshold defined as rational multiples of σ_s above μ_s, i.e. $\theta_s = \mu_s + q\sigma_s$,
  - Find corresponding wind speed yielding same visibility as snow rate threshold in blowing snow conditions, and
  - Apply wind speed criterion to grids with snow accumulation over 3-cm.
- It is found that $q \approx 1.6$ yields satisfactory results
- For the **global** data sets
  - Find μ_w, σ_w, μ, and σ_s,
  - Use the same q to determine snowfall threshold, and
  - Determine corresponding wind speed criterion.



2010 Winter US Blizzard Animation - 2009-12-01 00:00:00 : T = 0000
55
50
45
40
35
30
25
20
-130
-125
-120
-115
-110
-105
-100
-95
-90
-85
-80
-75
-70
-65
-60
Precipitation
(mm hr-1)
10.0
1.0
0.1
Snow (kg m-2 s-1)
0.01
0.001
0.0001
0.1
1.0
10.0
100.0
Wind Speed (m s-1)

AES Major Features

- Custom user-defined operators (UDOs)
  - Connected Component Labeling (CCL) algorithms implemented, and more coming.
- Event Specification Language (ESL)
  - Allowing scientists to express their using Python rather than low-level SQL.
- Collaboration via the Collaborative Workbench (CWB).
  - Event definitions and search results can be shared and modified.
- Parallel performance.
  - Data-parallelism native to SciDB's shared-nothing distributed architecture.
  - Calculations performed on local data of a to minimize data movements.
- AES provides a web service
  - It not only enables the ESL mentioned above, but also allows AES to be embedded within other applications.

AES Summary

- The *Big Data* analysis solution demonstrated by AES is meant to be an agency-level (or higher) facility solution, similar to Data or HPC centers.
  - i.e. Not a departmental level solution.
  - Cloud-ready!
- Developing User-Defined Operators/Functions (UDO/F's) requires professional software developers.
  - Extensibility suffers but gains better software quality!
- Completed UDO/F's are immediately "reusable".
- Parallelism is built into the shared-nothing architecture and professionally crafted UDO/F's.
  - Scientists can take advantage of parallelism without learning parallel programming!

Related Projects

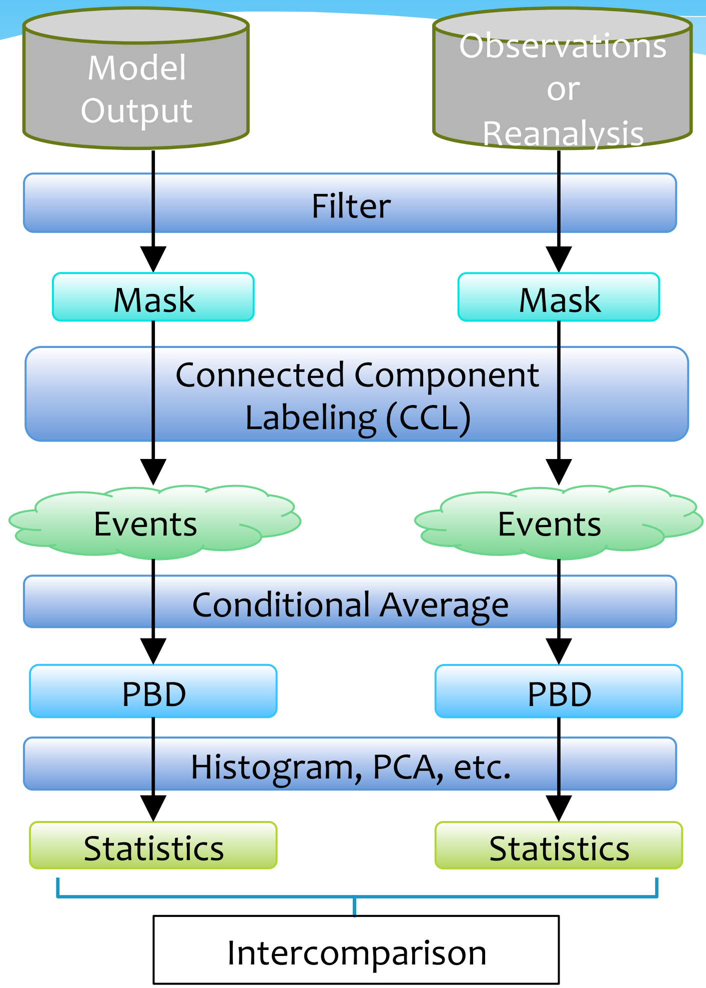
PROBE: Process-Based Diagnostics

Team: G. Schmidt, K. Kuo, M. Bauer, A. Oloso

Enable routine use of process based diagnostics (PBDs) as a means for targeted improvement of weather/climate models.

- Conventional diagnostics are inadequate:
  - Strong coupling - difficult to point finger
  - Signal is diluted by portions of domain where process is inactive
- Builds upon Automated Event Service (AES)
  - Step 1 is to identify regions in which a selected process is active. I.e. "events" by AES definitions.
  - PROBE requires more analysis and customized operators.
- PROBE must also extend AES to support:
  - Ensembles and versioning
  - Routine automations for iterative evaluation

PBD’s



Conditional Average Triggered by anomalies

- Extreme or rare events (heat waves, downpours)
- Introduces context
    - This happens when…

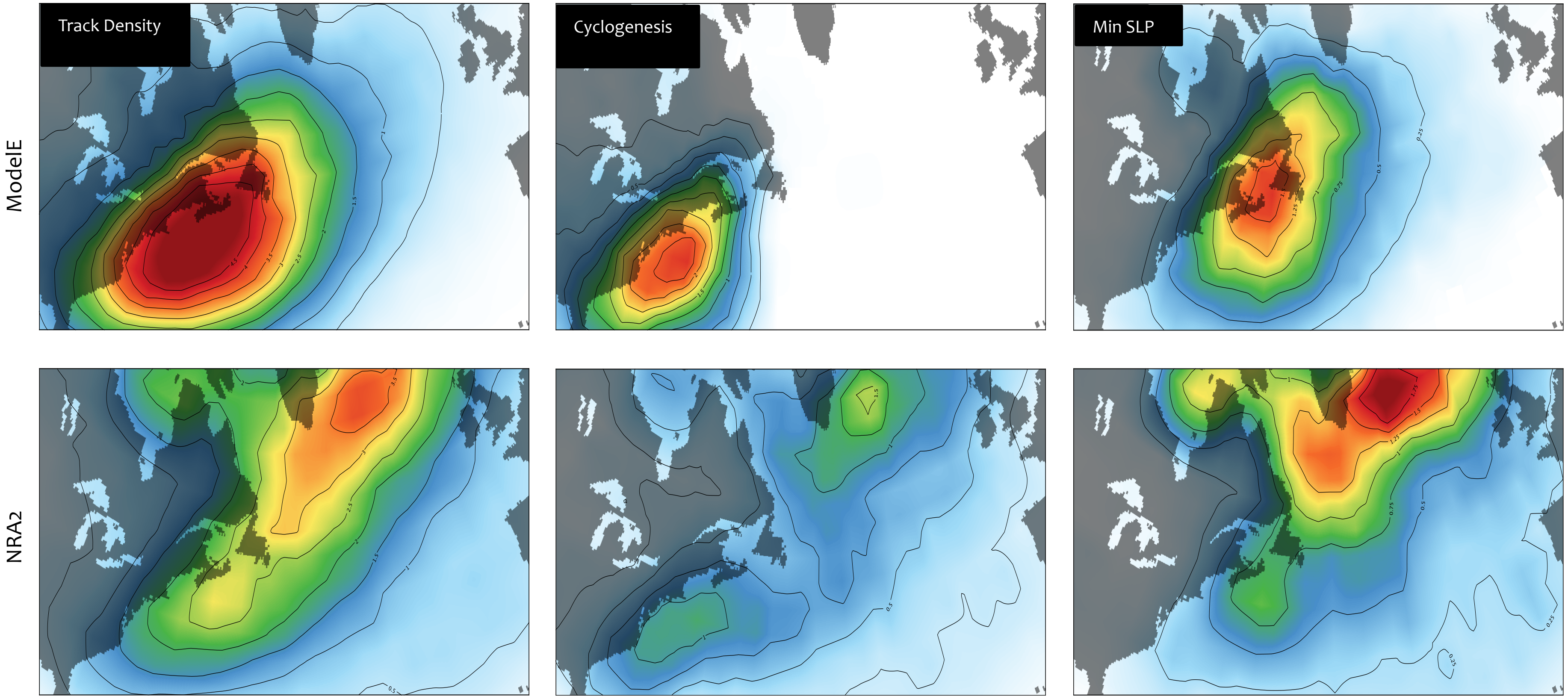
Lagrangian Viewpoint

- Movable frame centered on the object of interest.
- Preserves life-cycle and internal spatial structure.

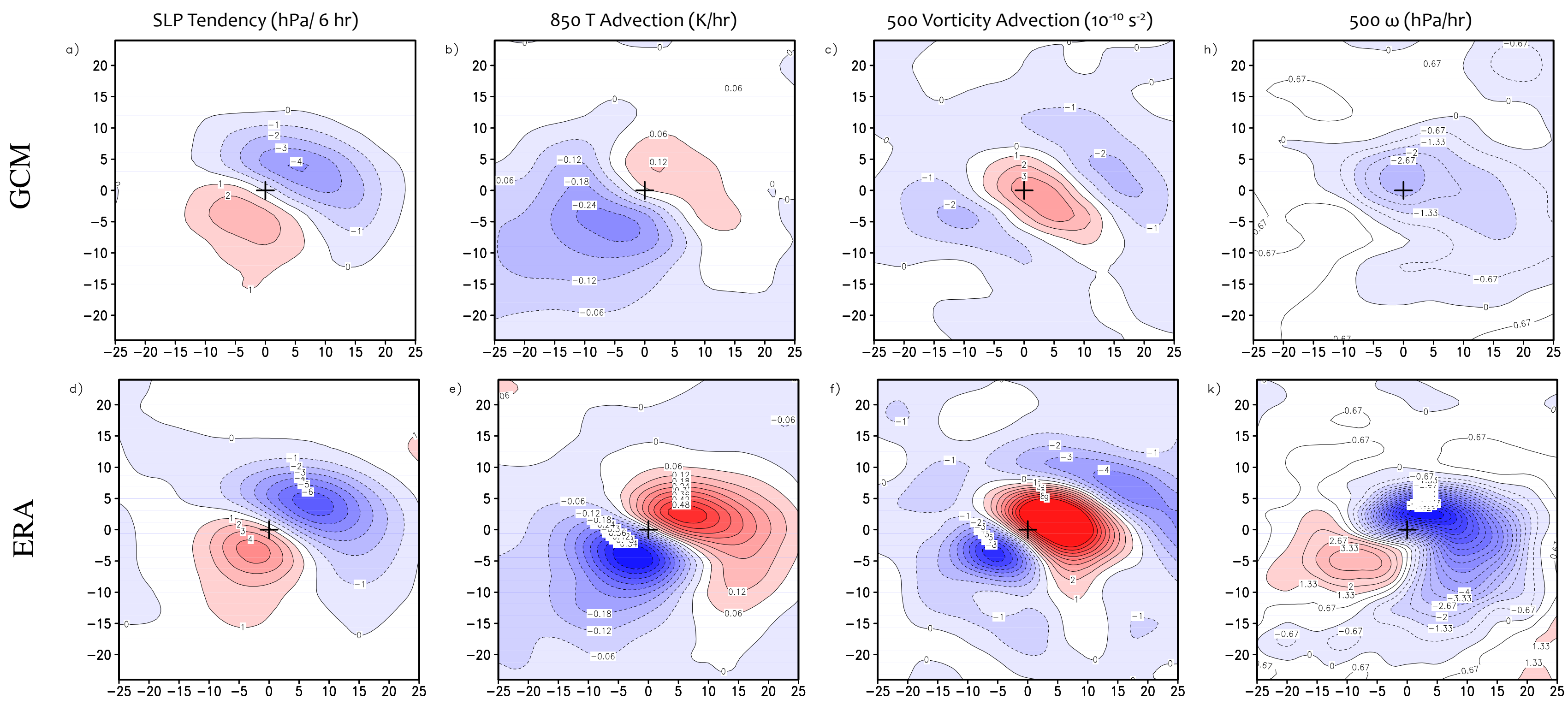
Process-based diagnostics merge these methods.

- Diagnosing issues with development, structure and feedback.
- Closer to elemental processes/procedures.

How are process-based diagnostics useful?



How are process-based diagnostics useful?



Deploying AES in the Cloud

- Goals:
  - Explore feasibility of providing AES as a cloud-service
  - Demonstrate scalability of AES/SciDB on large systems
- Limitations of existing resources:
  - Traditional HPC is not well-suited for interactive exploration and distributed I/O
  - Traditional data centers are not well-suited for custom (user-defined) analysis. Also data is segregated across DAACs.
- <u>Cloud-based computing appears well-suited.</u>

AES in the Cloud

- Procured ~ $180k of Amazon Web Service
  - ~30,000 node-hours
  - 30 TB of storage for 6 months
- Milestones:
  - Replicate AES Blizzard query on 30 nodes
  - Evaluate scalability on 300, and 3000 nodes
- Status?
  - Lots of surprises in various facets of deployment.

AES in cloud

- Lessons learned
  - AWS Service default limits – contact AWS to fix
  - Avoid using AWS Market images in a cluster
  - Starcluster is too aggressive about freeing resources when problems arise
  - Use S3 as primary storage
    - Robust, cost effective, and accessible
    - Good for staging to faster ephemeral (EBS) storage
  - Use EBS storage for application

Feasibility Study: Moving Object Database

- MODB
  - Traditional databases are poorly suited for modeling objects that move and change over time
    - E.g. instrument flies over hurricane between 6 hour snapshots from model output.
  - ESTO funded MODB effort led by M. Schneider to overcome these limitations. Implemented in traditional database.
- Conclusion: array data model is better suited than table model but geometric operations in SciDB still needs extensions

Future Directions

- Extend capabilities:
  - Include support for ungridded/swath data
  - Support regridding to compare across data sets
  - Enable nonlinear dimensionality reduction to find nontrivial correlations in data
  - Enable treatment of moving objects
- **Build user community**

Thanks

- We would like to thank:
  - NASA ESTO/AIST and MAP for supporting this work
  - NASA High End Computing and the NCCS for computing resources

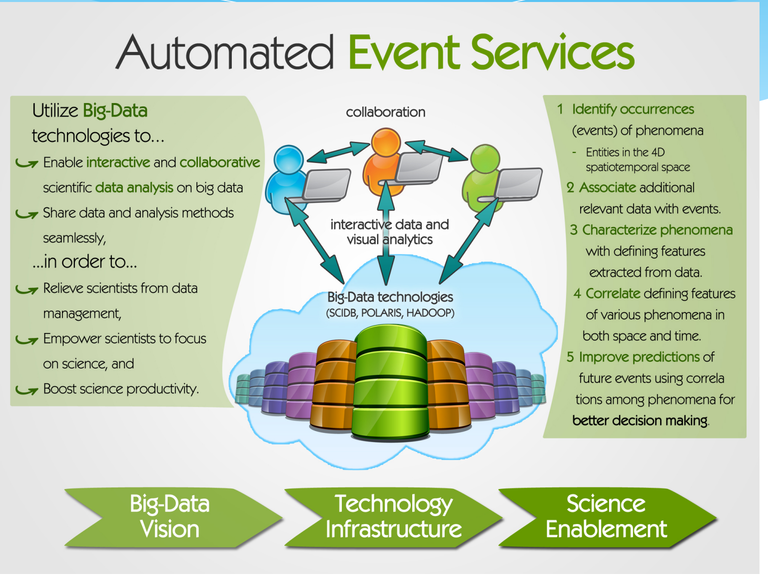
Supplemental Material

Automated Event Service (AES): Efficient and Flexible Searching for Earth Science Phenomena

PI: Tom Clune, NASA GSFC

Develop an Automated Event Service system that:

- Methodically mines custom-defined events (e.g., tornadoes) in the reanalysis data sets of global atmospheric models.
- Enables researchers to specify their custom, numeric event criteria using a user-friendly web interface to search the reanalysis data sets.
- Supports Event Specification Language (ESL) for more flexibility and versatility.
- Contains a social component that enables the dynamic formation of collaboration groups for researchers to cooperate on event definitions of common interest.
- Provides rapid results via high performance computing and advanced search technologies.



- Leverage advances in high-end computing and search technologies to create an efficient mechanism for searching reanalysis data for events.
- Build baseline system by custom integration of mature components: HPC cluster, MapReduce, Hadoop/Hbase.
- Develop ESL via analysis of representative events.
- Adapt advanced tree-based indexing strategies to efficiently support parameter-based event queries
- Apply agile methodology: develop in small increments driven by use cases and synthetic tests.

• Import reanalysis data	10/12
• Implement native indexing	02/13
• Complete event web service	05/13
• Complete basic web portal	11/13
• Complete distributed event database	02/14
• Design review for ESL	05/14
• Demonstrate Blizzard use case on 30-node commercial cloud cluster	07/14
• Complete multifaceted web portal	11/14
• Complete tree-index search capability	03/15
• Complete delivery of ESL	03/15
• Testing and validation	05/15

Co-Is/Partners: Kwo-Sen Kuo, Bayesics; Rahul Ramachandran, NASA MSFC

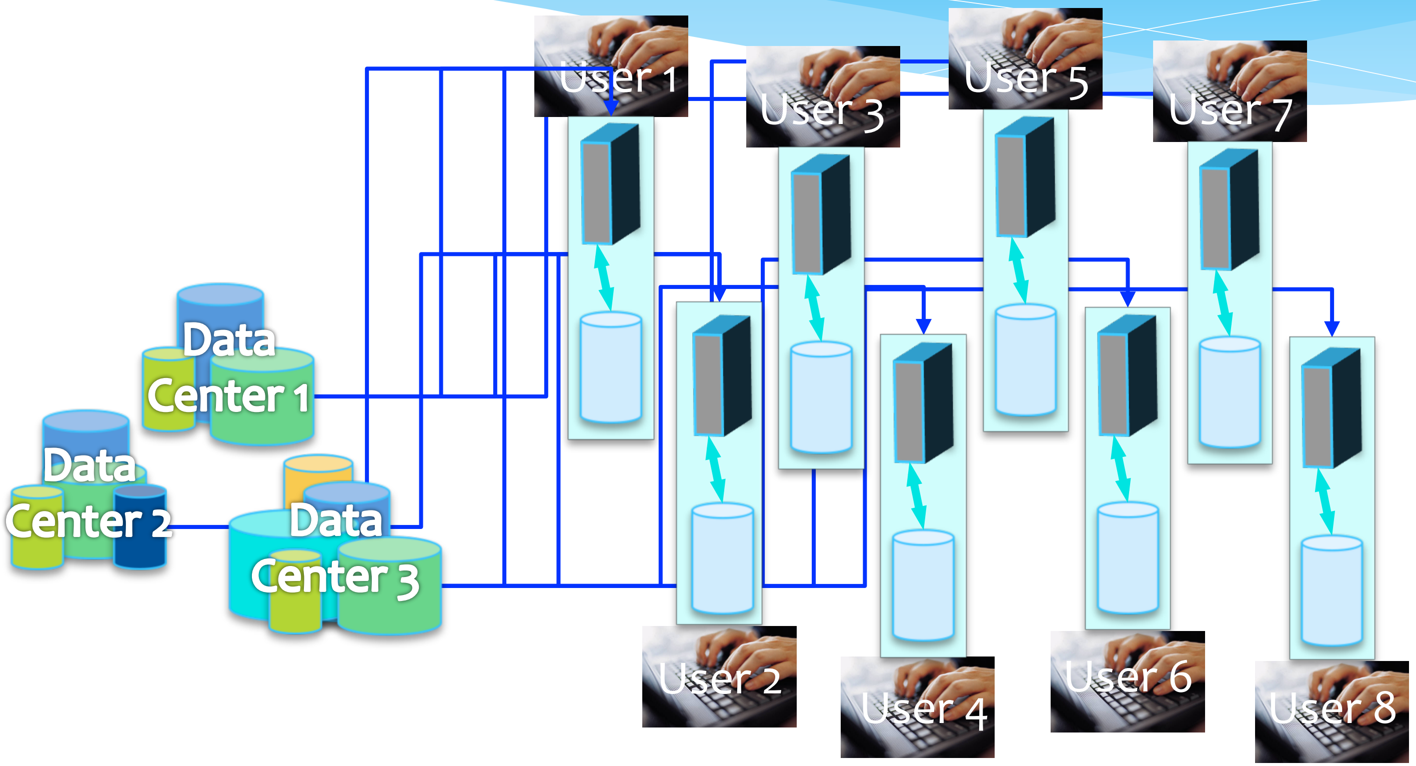
$\mathbf{TRL_{in} = 2}$ $\mathbf{TRL_{current} = 4}$

AIST-11-0008

Current Situation

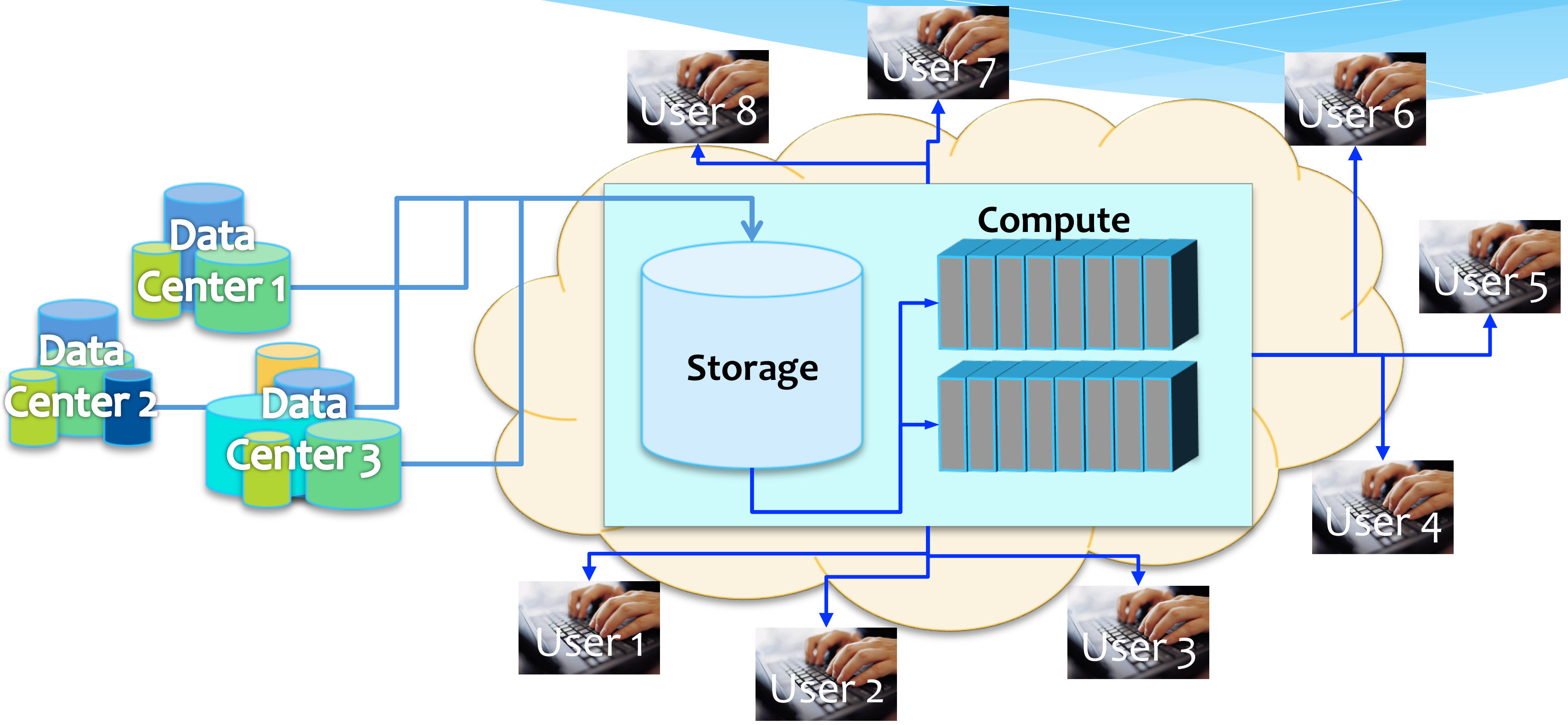
- The majority of NASA Earth science data (~10 PB) are archived and distributed as **<u>files</u>**, …
- Standardized through APIs, such as HDF and netCDF, for access, …
- Only the **<u>metadata</u>** are cataloged into RDBMs and are thus searchable.
- Searching for data not contained in the metadata becomes slow, e.g. precipitation intensity > 0.7 mm hr^{-1}.

Current Data Analysis Practice



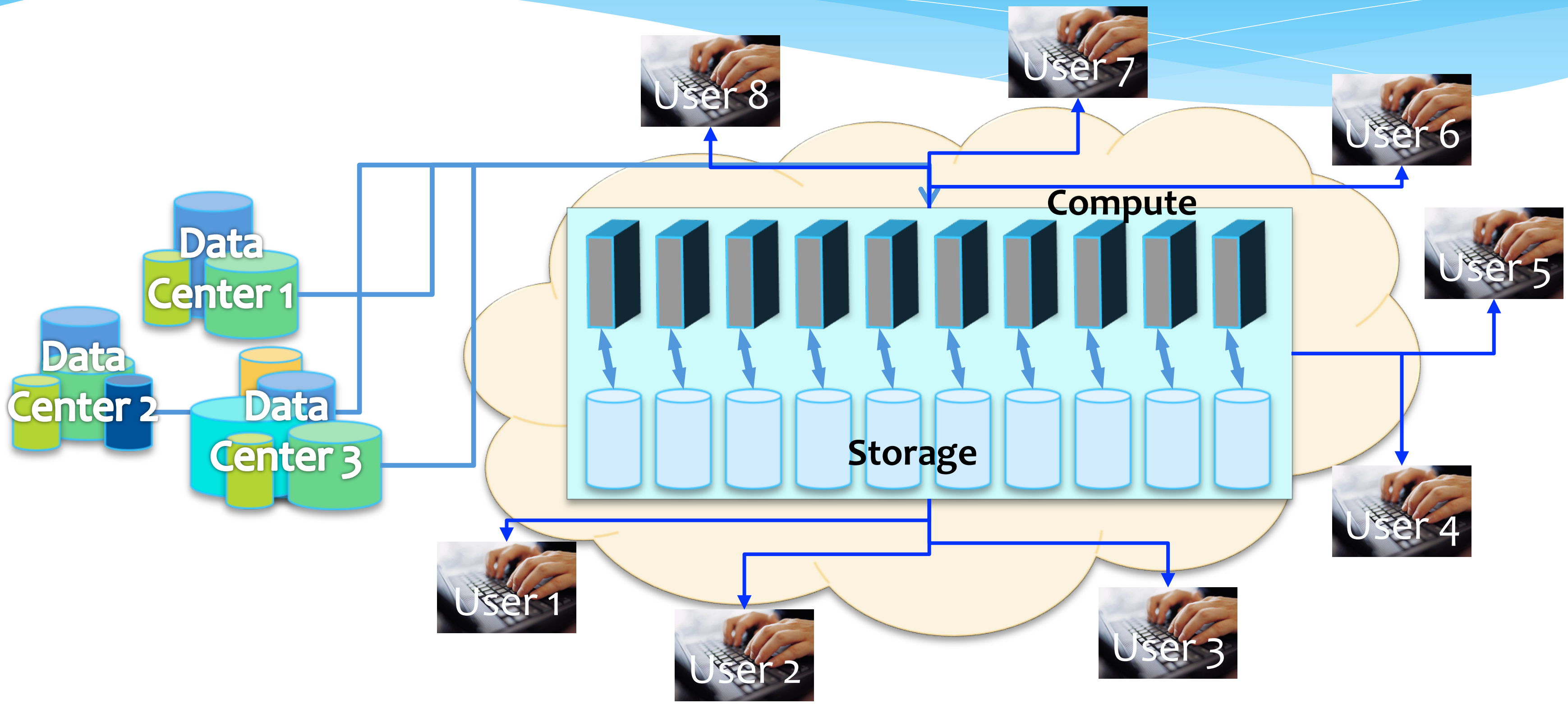
- Everyone downloads needed data from data centers.
- Data analysis is conducted on local resources, mostly serially.

Traditional HPC Architecture



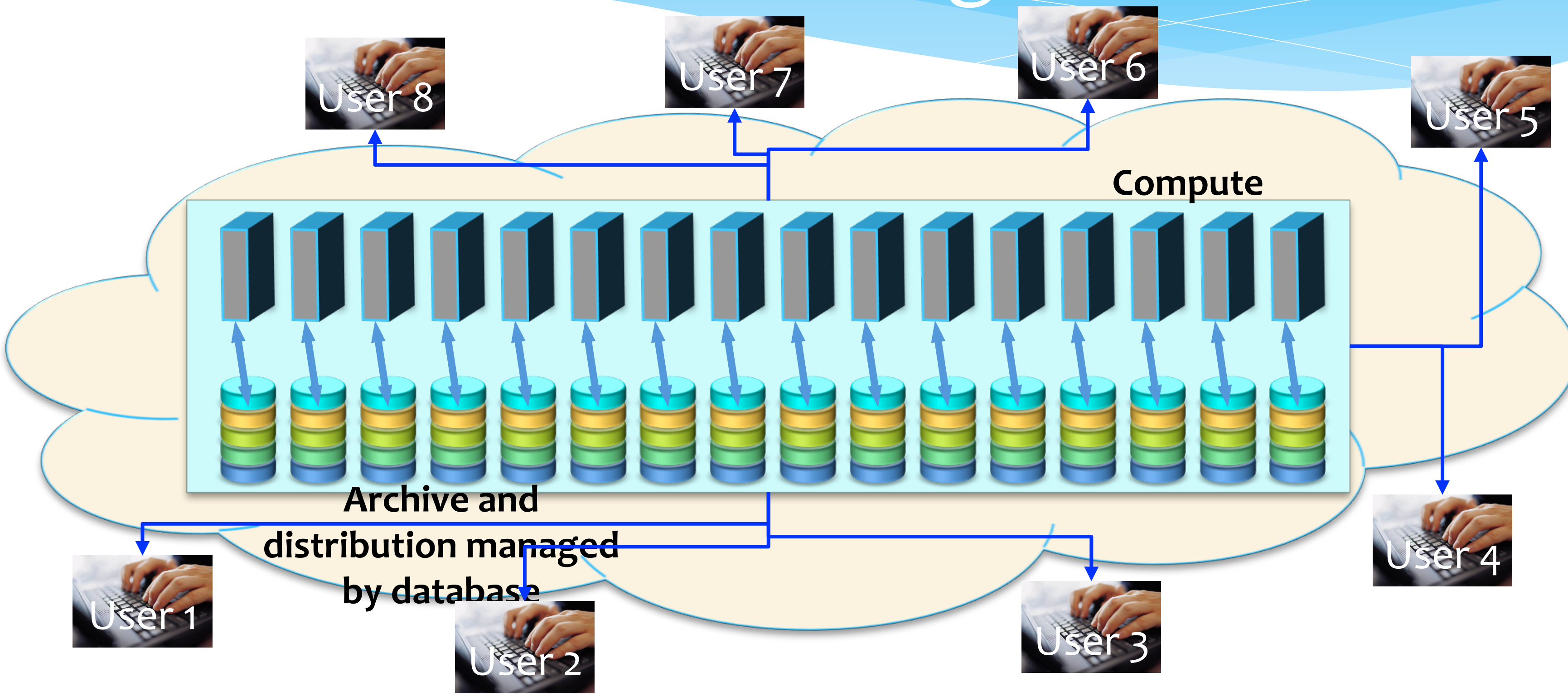
- Fast, central, large storage system.
- Suitable for MPI parallelism, but few scientists can program with it.

Data Intensive Architecture



- Distributed storage with local compute access.
- Loosely coupled parallelism requires little or no inter-process communication

Data Archive, Distribution, and Analytics All Together



- Data directly "archived" into parallel databases for distribution and/or analysis.