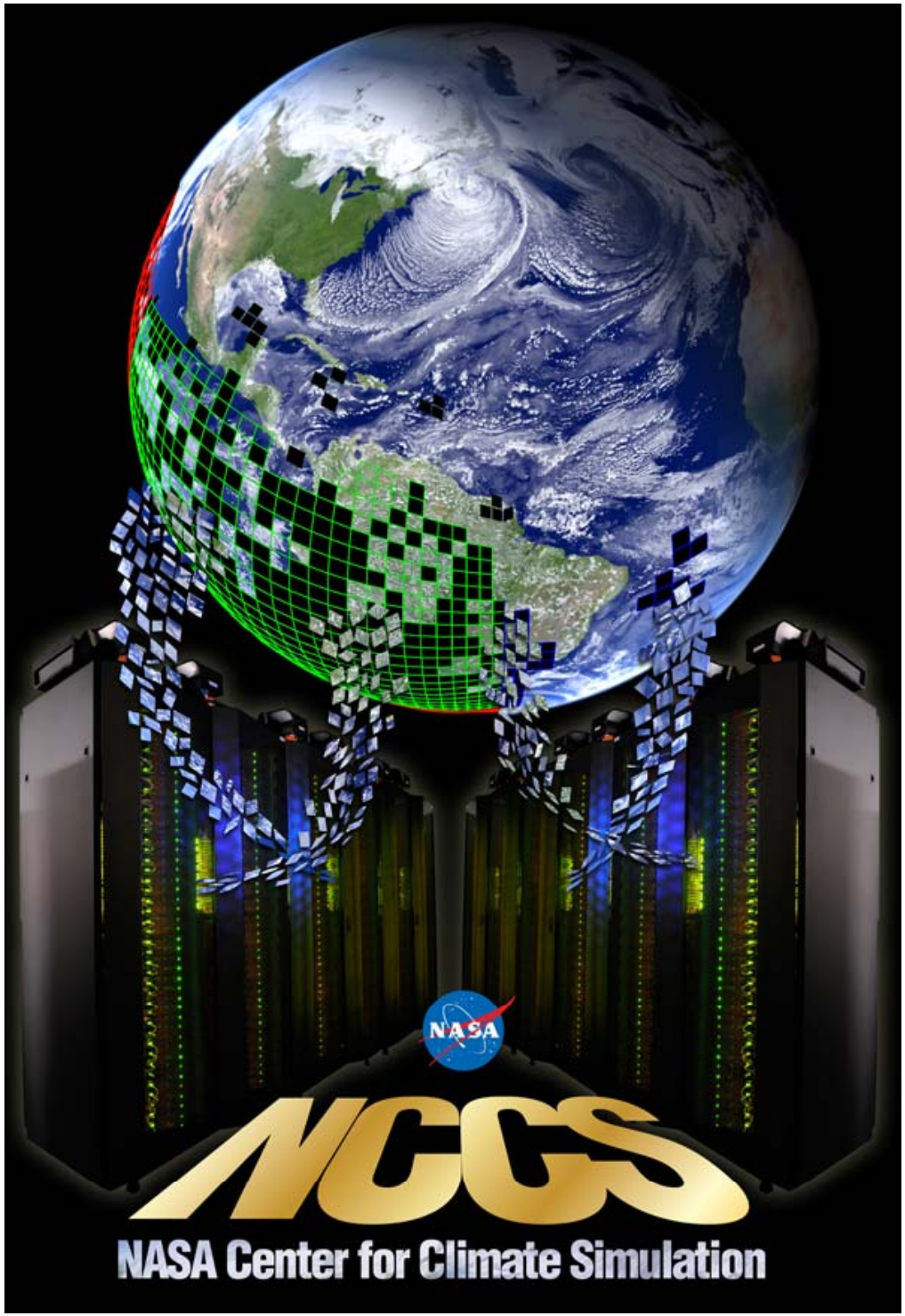


Alternative Architectures for Mass Storage

Dan Duffy
NASA Center for Climate Simulation
Goddard Space Flight Center

2011 IEEE Mass Storage Systems & Technologies
May, 2011

http://nccs.nasa.gov
http://cisto.gsfc.nasa.gov
http://www.hec.nasa.gov

Standard Disclaimers and Legalese Eye Chart

- All Trademarks, logos, or otherwise registered identification markers are owned by their respective parties.
- Disclaimer of Liability: With respect to this presentation, neither the United States Government nor any of its employees, makes any warranty, express or implied, including the warranties of merchantability and fitness for a particular purpose, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.
- Disclaimer of Endorsement: Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government. In addition, NASA does not endorse or sponsor any commercial product, service, or activity.
- The views and opinions of author(s) expressed herein do not necessarily state or reflect those of the United States Government and shall not be used for advertising or product endorsement purposes.
- All errors in this presentation are inadvertent and are the responsibility of the primary author.

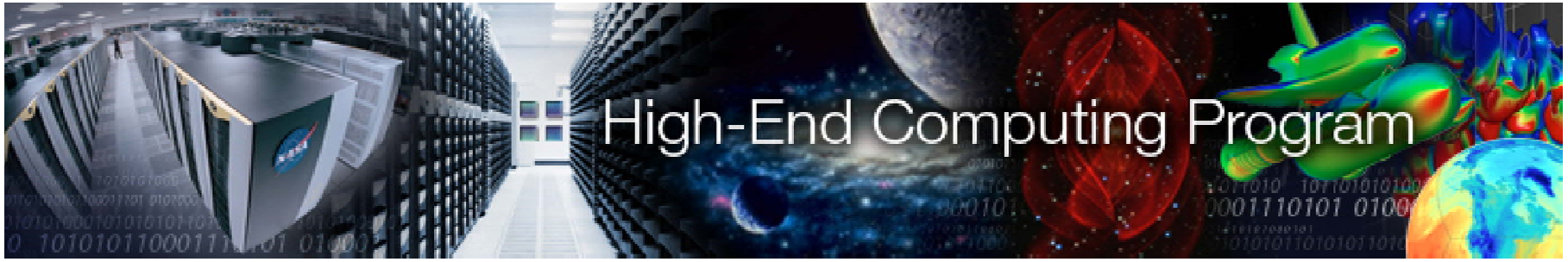
Good News and Bad New

- Good News
  - We are still here, regardless of May 21st, 2011 doomsday predictions
- Bad News
  - I wasn't expecting to have to put together a presentation for this conference
- Good News
  - I was able to pull some slides together
- Bad News
  - We still have a storage problem to deal with; Doomsday did not solve that problem for us; And you have to listen to me!

Agenda

- Quick introduction to the NCCS
- Current NCCS archive architecture
- How much data are we talking about?
- New science goals that are driving changes to the archive
- What is the NCCS doing to work toward these changes?

NASA High-End Computing Program



HEC Program Office
NASA Headquarters
Dr. Mike Little
Scientific Computing Portfolio Manager

High-End Computing Capability (HECC) Project
NASA Advanced Supercomputing (NAS)
NASA Ames
Dr. Rupak Biswas

NASA Center for Climate Simulation
Goddard Space Flight Center
Dr. Phil Webster

Understanding the Science and Computational Requirements

Development of Individual Models

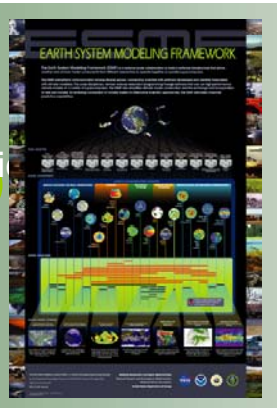
Model Frameworks and Improved Observations

The Predictive Earth System

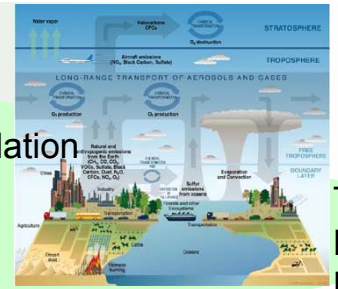
Earth Science

Atmosphere
Sulfate Aerosol
Dynamic Vegetation
Land Surface
Non-Sulfate Aerosol
Atmospheri Chemistry
Ocean and Sea Ice
Carbon Cycle

Mid 1970's → Early 2000's



Data Assimilation

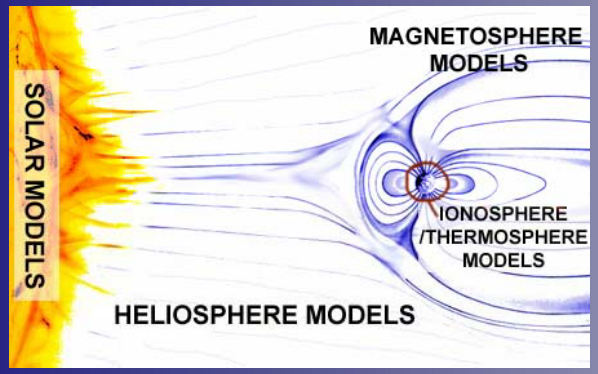


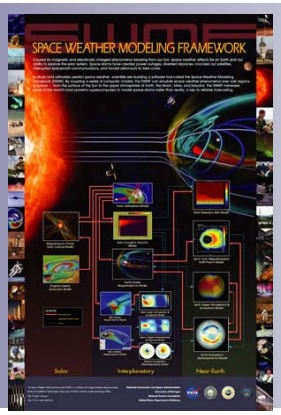
Towards:
Policy
Planning
Education

Climate

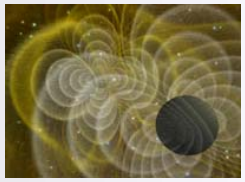
Weather

Space Science



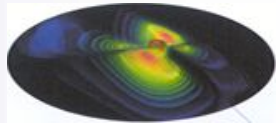


Fundamental Understanding of the Universe



Colliding Black Holes

Orbiting Neutron Stars

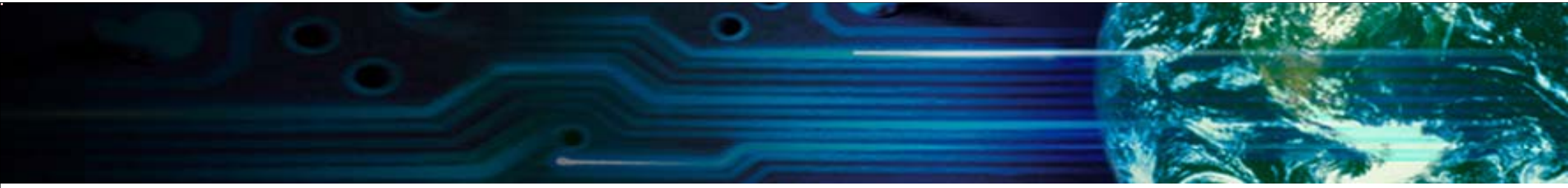


Megaflops → Gigaflops → Teraflops → Petaflops → Exaflops

Megabytes → Gigabytes → Terabytes → Petabytes → Exabytes

References: (1) Strategic Plan for the U.S. Climate Change Science Program – A Report by the Climate Change Science Program and the Subcommittee on Global Change Research, July 2003.

(2) Preview of Our Changing Planet – The U.S. Climate Change Science Program for Fiscal Year 2008, April 2007. (3) Earth Systems Modeling Framework Website, http://www.esmf.ucar.edu/. (4) Community Coordinated Modeling Center, http://ccmc.gsfc.nasa.gov/.



Climate Science Computing

Climate Science is "Data Intensive". . . . . .

- Our understanding of Earth processes is based in the observational data record and is expressed as mathematical models.
- Observation data sets produce large data sets from 100s of terabytes to petabytes and are growing
- Data assimilation combines observational data with model prediction.
- Climate models produce large data sets (100s of terabytes) for the scientific community as well as decision makers.
- Reanalysis with improved models results in vast data sets (100s of terabytes) for the scientific community.

. . . . . And Requires "Data Centric" Computing

- Designed for effective manipulation of large data sets.
- Data that is accessible to a multitude of services with effective data management tools.
- Efficient data analysis needs to have "supercomputing" capability with data sets online.
- Data sets must be made easily accessible to "external users" with analysis and visualization capability.

NCCS Mission

- *Traditional*
  - Enable scientists to increase their understanding of the Earth and the universe by providing state-of-the-art high performance computing, storage, network, and application solutions
  - Provide large-scale compute engines, analytics, data sharing, and high-end computing services
- *Future*
  - Develop a data services capability to better support the climate research communities and prepare the way for technology advances

New Markets for Climate Simulation Data

NASA Scientific Community
- Simulation data consumers
- Advance scientific knowledge
- Direct access to systems
- *Supercomputer* capability required for effective analysis

NASA Modeling Community
- Model development, testing, validation, and execution
- Data creation
- Largest HPC usage
- Requires observational data as input

Observation and Simulation Data

Engineering Community
- Satellite design
- Instrument design

External Scientific Community
- Simulation data consumers
- Advance scientific knowledge
- Web-based access to data

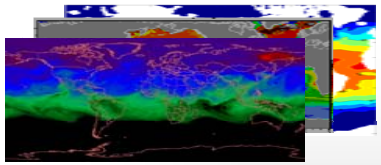
External Applications Community
- Huge opportunity for impact
- GIS Community
- Simulation data consumers
- Limited ES data expertise
- Web-based access to systems

Public/Citizen Scientists
- Web-based access to data an analysis
- No ES expertise

NCCS Data Centric Climate Simulation Environment

Data Sharing and Publication

- Capability to share data & results
- Supports community-based development
- Data distribution and publishing



Code Development*

- Code repository for collaboration
- Environment for code development and test
- Code porting and optimization support
- Web based tools

User Services*

- Help Desk
- Account/Allocation support
- Computational science support
- User teleconferences
- Training & tutorials

DATA Storage & Management

Global file system enables data access for full range of modeling and analysis activities

Analysis & Visualization*

- Interactive analysis environment
- Software tools for image display
- Easy access to data archive
- Specialized visualization support

Data Transfer*

- Internal high speed interconnects for HPC components
- High-bandwidth to NCCS for GSFC users
- Multi-gigabit network supports on-demand data transfers

HPC Computing

- Large scale HPC computing
- Comprehensive toolsets for job scheduling and monitoring

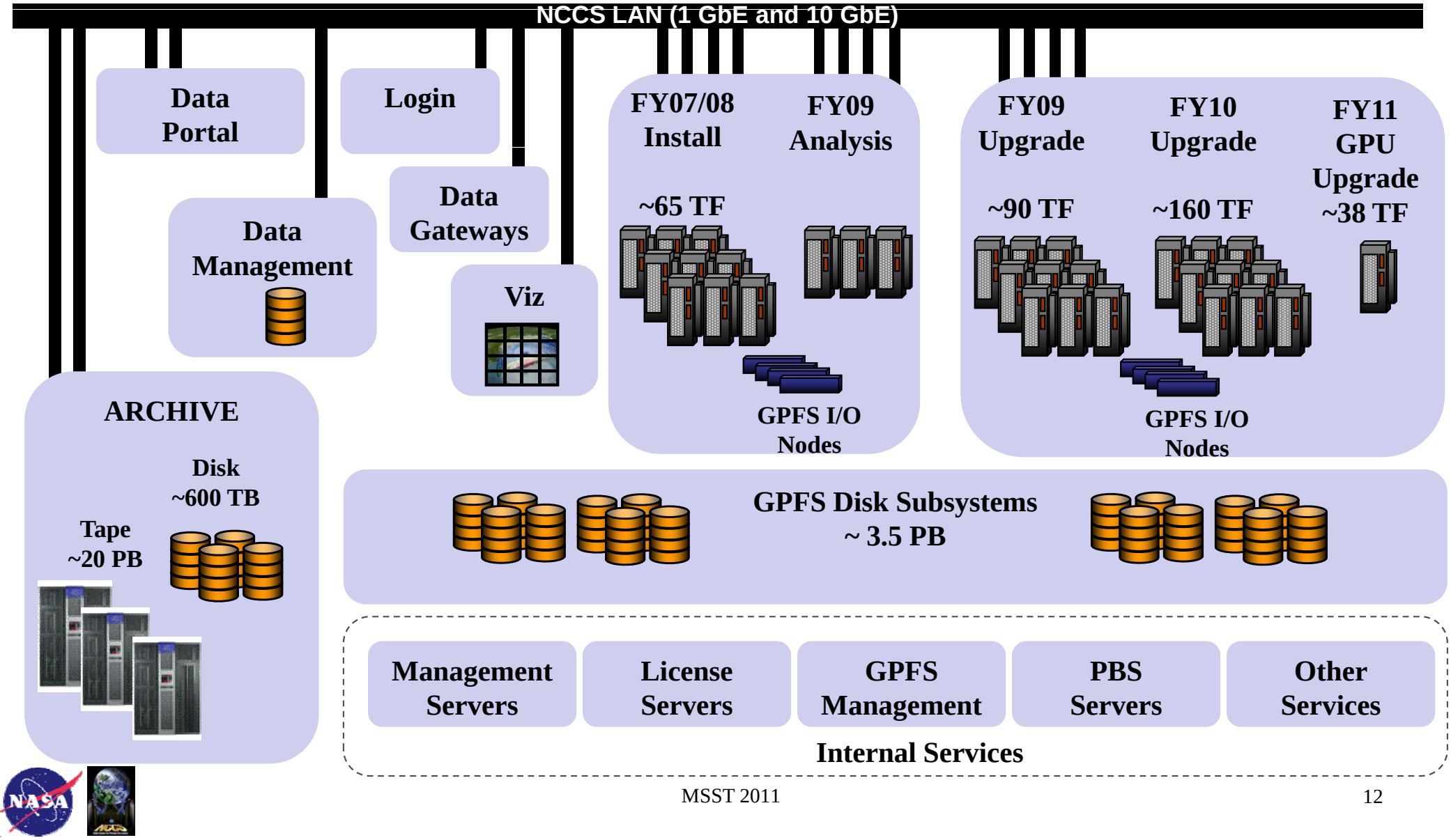
Data Archival and Stewardship

- Large capacity storage
- Tools to manage and protect data
- Data migration support

*Joint effort with SIVO

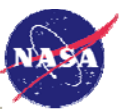
*Joint effort with SEN

Current NCCS Architecture



Current NCCS Archive

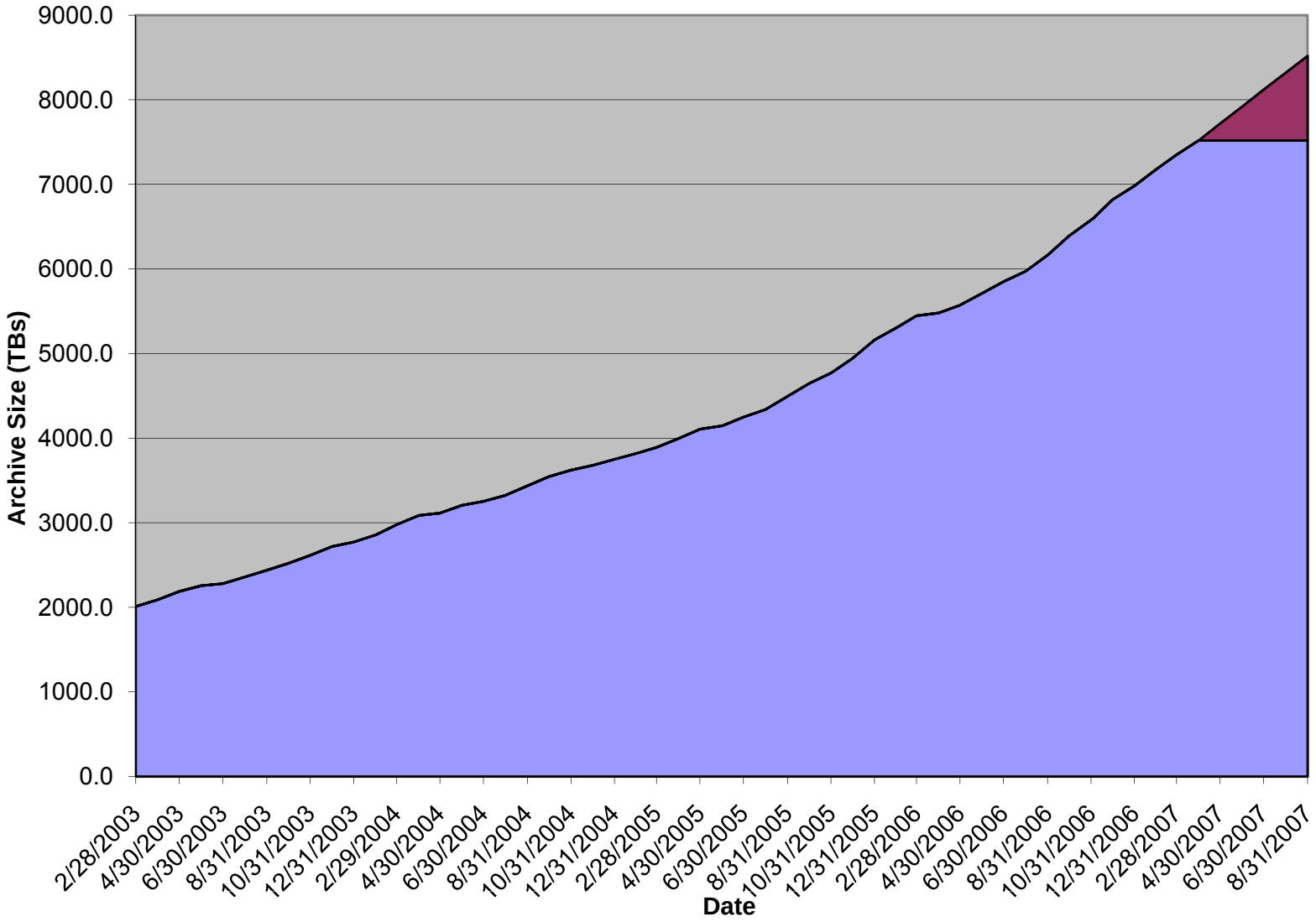
- SGI Parallel DMF
  - Two DMF servers
  - Two NFS edge servers
  - Two login servers
  - Three parallel data movers (pDMOs)
- 600 TB (RAW) of primary disk
- Large number of tape drives (for us anyway)
  - 9940B, T10KA, T10KB
- Tape libraries
  - Combination of 9310s and SL8500s




Archive Upgrades – In Progress and Planned

- Upgrade disk cache – add 600TB
  - Adding capacity to the primary disk cache
- Expand capacity and bandwidth of SL8500 libraries
  - Add sixteen (16) T10000C tape drives plus media
- Removal three 9310 silos
  - Move tape drives around and decommission libraries
- Procure third SL8500 library
  - Needed to address pending capacity shortage for second copies
- Move tape drives around between primary and secondary building

NCCS Archive Data Archive Growth



Growth will reach 8.5 to 9 PB by the end of the year.

And we make 2 copies of all data.

How Much Data?

Site	Data (approximate)	Solution
Yahoo	Many Exabytes	Hybrid solutions: disk, tape, object stores, etc.
Blue Waters	500 PB	Disk and tape
GFDL	100 PB	Disk and tape
NAS	30 PB	Disk and tape
NCCS	17 PB	Disk and tape

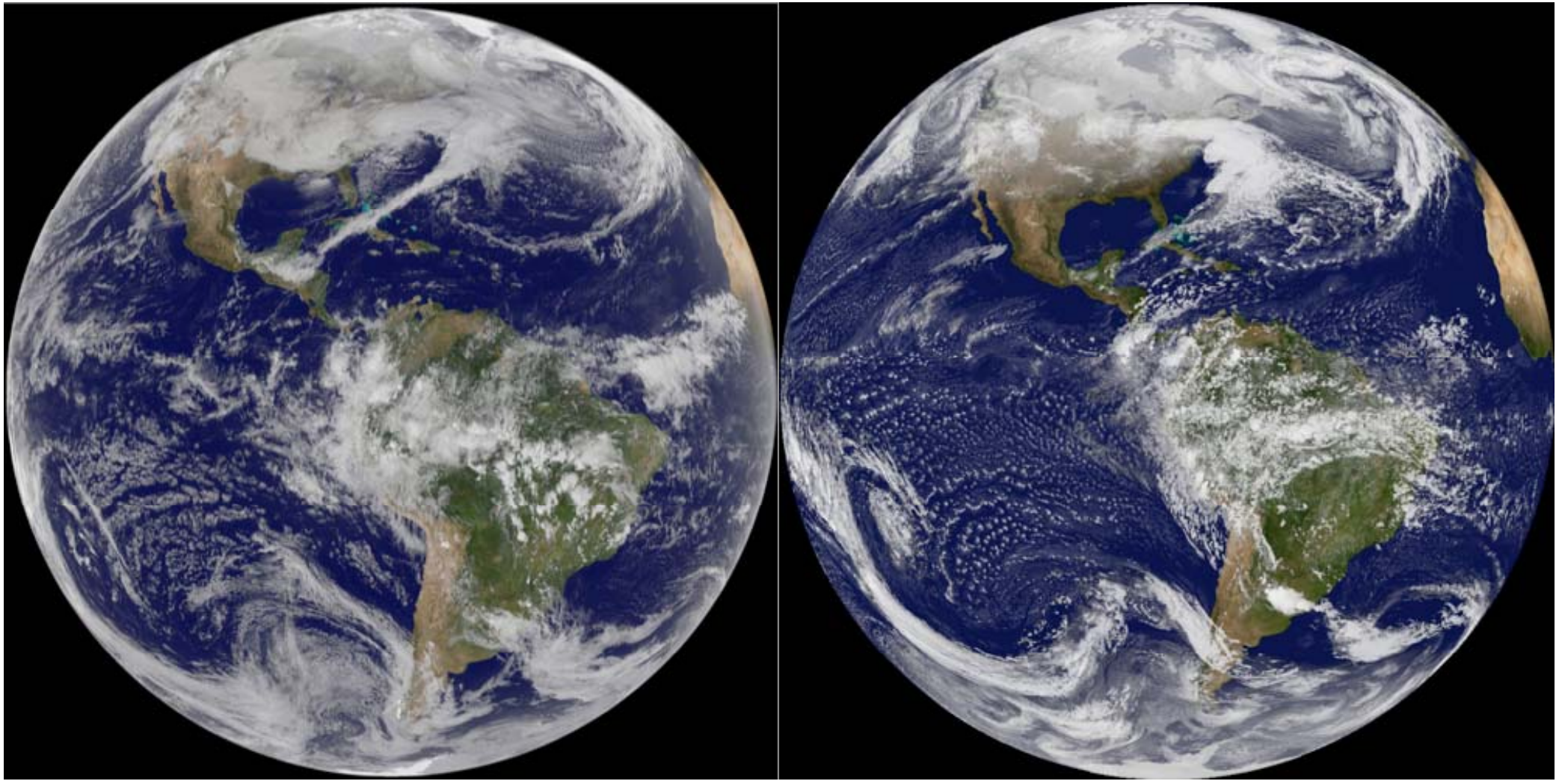
- Came to MSST last year hoping to hear more about how others are tackling the same issues
- Left Tahoe feeling a bit depressed; everyone is building a bigger archive
- Perhaps I am just jealous that our archive is not a big as others, but something does not seem right

New Science Goals

- *Intergovernmental Panel on Climate Change (IPCC) Assessment Report 5 (AR5)*
  - Provide the data management services and analytical tools necessary to support the publication requirements of the IPCC
- *Observation/Simulation Data Integration*
  - Bring the climate modeling and observational communities together to work toward the goal of integrating model outputs and observational data
- *Next Generation HEC Requirements for Modeling and Assimilation*
  - Contribute emerging technologies to address computing the ever increasing requirements for Earth system modeling
  - Continuing to push the resolution of global models to the highest possible resolutions

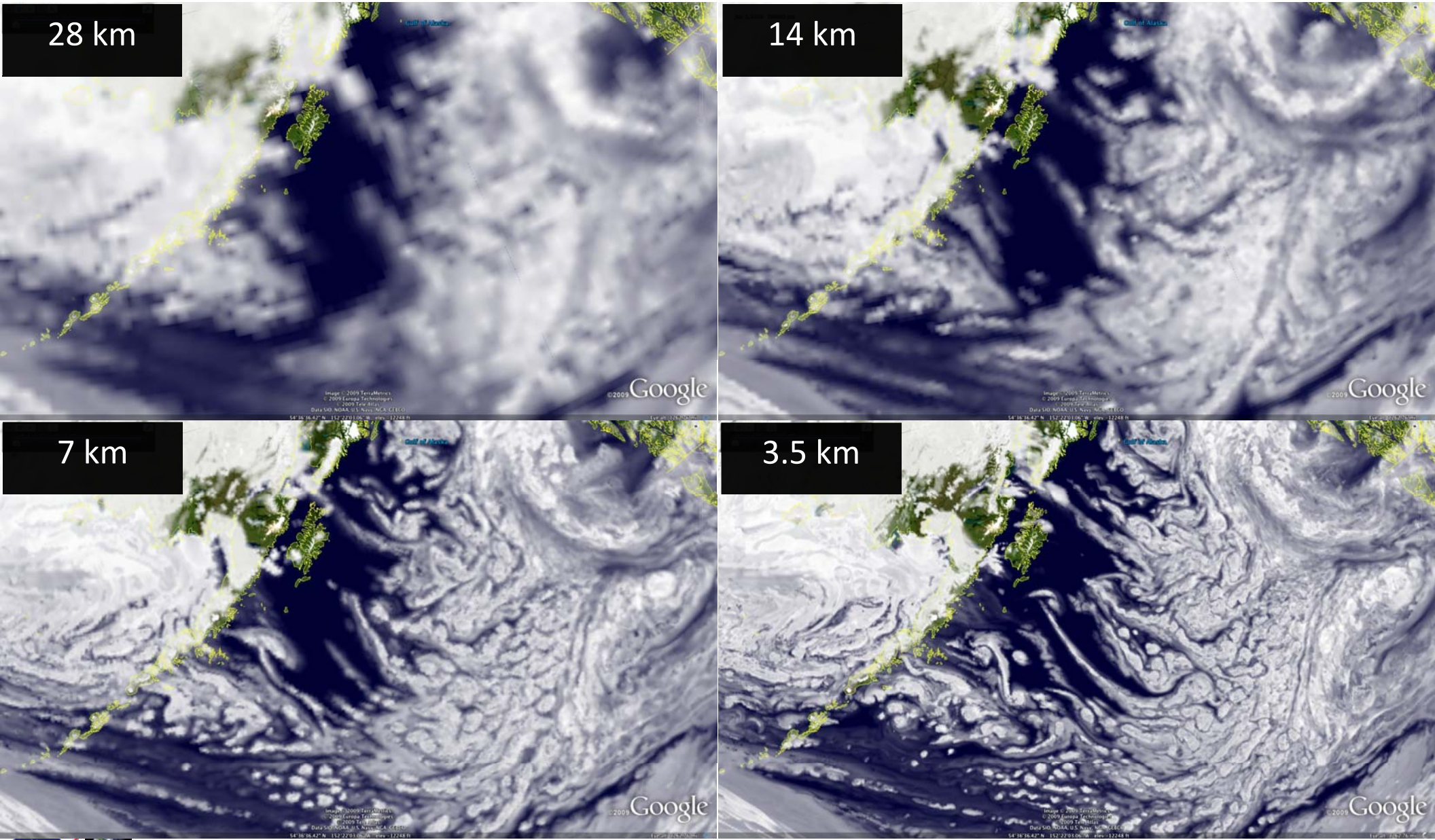
High-Resolution Climate Simulations with GEOS-5 Cubed-Sphere Model

Bill Putman, Max Suarez, NASA Goddard Space Flight Center;
Shian-Jiann Lin, NOAA Geophysical Fluid Dynamics Laboratory



One of these is a satellite image and one is from the high-resolution GEOS-5 climate model. Which one is which?

GEOS-5 – High Resolution of Karman Vortex Sheets



Future Science Questions

- Use high resolution nature climate runs for observing system simulation experiments for the design of new Earth observing instruments
- How can NASA climate model data be used better for the understanding, mitigation, and adaptation due to climate change?
  - Heat waves can kill a significant number of people in urban areas
  - Floods cause huge property loss and displace large amounts of people
  - Will hurricanes become more intense due to climate change?
- How can the US plan their water management resources over the next 50 to 100 years based on climate change?
- What types of crops are going to grow best and where throughout the US?
- What types of grapes should I plant in Napa Valley to harvest certain types of wine in 20 years?
- How can NASA observation and model data be better used for first response and near first response to disasters?

New (and Old) Challenges

- Broader set of services for the archive
- Finding observational and model data for use in climate and weather studies
- Accessing the geographically distributed data
- Managing the massive digital holdings, which are measured in petabytes and hundreds of millions of files
- Maintaining the data, which must often be preserved for decades
- Supporting data sharing, data publication, and data stewardship

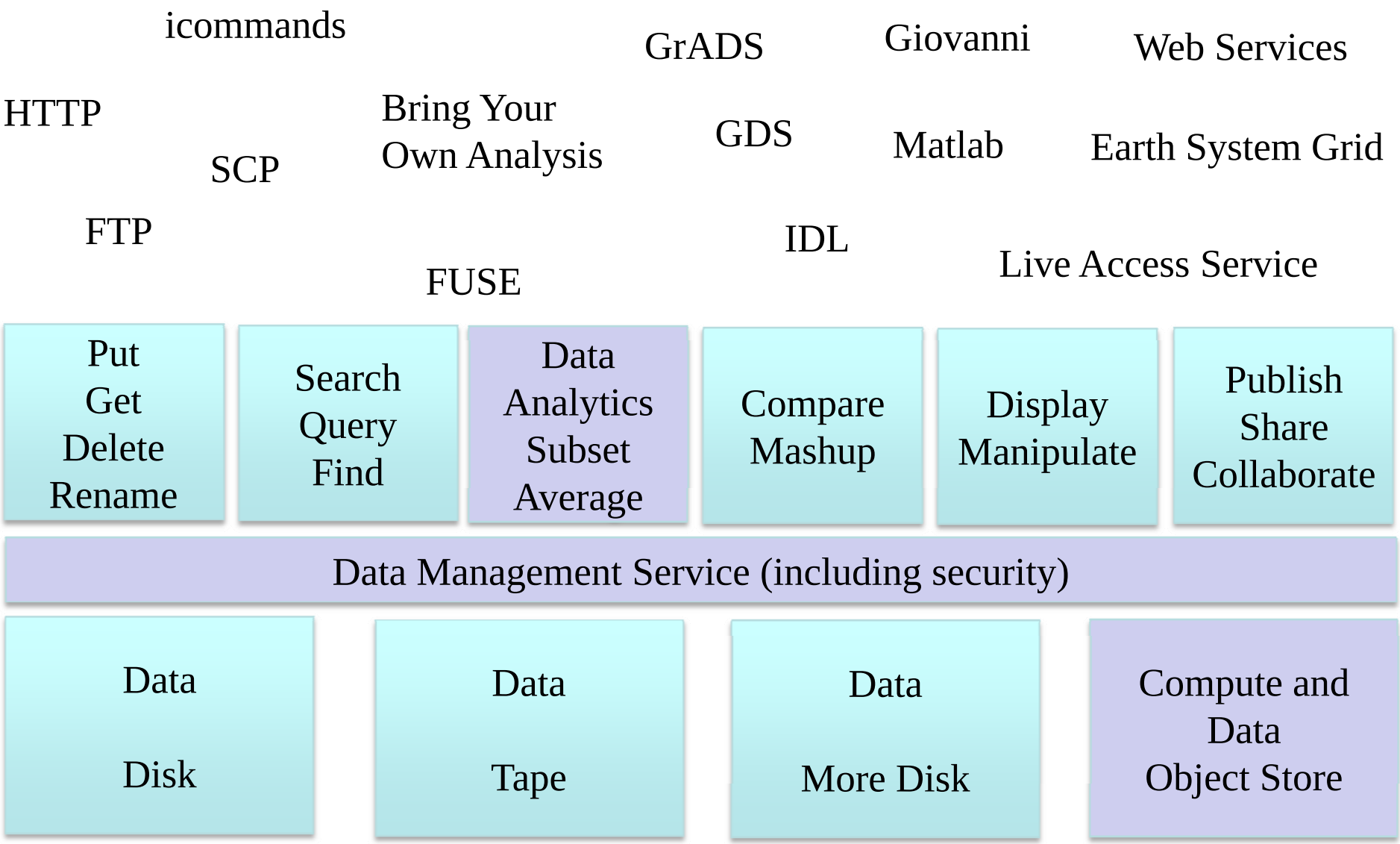
New Data Services for the NCCS Archive

- NASA Modeling and Scientific Community
  - Traditional focus of NCCS (and many scientific centers)
  - Requirements
    - Climate science experts with fairly well known workflows (at least to the modelers)
    - Data management kept in file names, within headers of files, and within people's heads; Growing need for data management
    - Large tape archives
- External Science and Applications Community
  - Add services to focus on these communities
  - Requirements
    - Not climate science experts; unknown workflows – user defined applications
    - Need for self described data sets with robust data management
    - Web services, cloud services, read, search, translate, subset, mashup, etc.

Services in 5 years (some sooner)

- What if you architected an archive based on what services you wanted to provide and not just how much data needed to be stored? What would the archive look like?
- Provide services to publish data quickly to different communities (colleagues, collaborators, science community, public)
- Quickly compare observation and model data
- Robust data analytics
  - Server side computational capacity to subset, average, publish
  - Does not require the users to download the data
- Data management services to allow the federation of NASA data with other repositories
- Combination of different compute and storage services
  - Disk (multi-tiered), Tape, Object storage environment
- Cloud computing environments to enable access to data and for users outside the NASA environment to consume the data where it sits
  - Bring your own app to NASA


Next Generation Archive Services



What are we doing to build this

- Data Management System
  - Conference paper – "The NASA Center for Climate Simulation Data Management System; Toward an iRODS-Based Approach to Scientific Data Services"
- Data Analytics Study
  - Map-reduce project for large scale data analysis project proposed within NASA

iRODS Data Grids

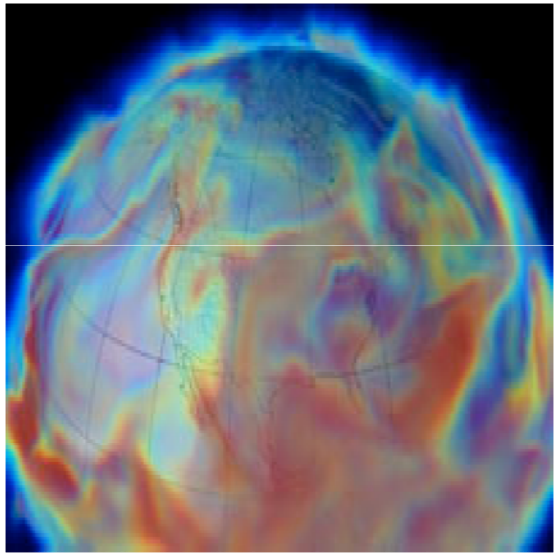
- Observation Data
  - Moderate Resolution Imaging Spectro-radiometer (MODIS) observational data
    - 54 million registered files, 630 TB of data, and over 300 million defined metadata values
  - Small-scale, multi-product, application-specific data service
    - The Invasive Species Data Service (ISDS) manages a collection of MODIS data products for ecological forecasting applications
- Simulation Data
  - Modern Era Retrospective-Analysis for Research and Applications (MERRA) simulation data
    - 360 files, 47 GB of data, and 4000 metadata values
  - Year of Tropical Convection (YOTC) data sets
    - 134,000 files, 12 TB of data, and 400,000 metadata values

Preliminary Tests – Federation

- Tested and evaluated iRODS data federation
  - Federated the YOTC and MODIS grids to simulate the union of observational and simulation data
- Explored the integrated management of observational and simulation data
  - Implemented an interface that enables comingling of remote and local observational and simulation data for advanced scientific study

Preliminary Results

- iRODS is a promising technology for exposing services for data management, publication, and analysis
- The iRODS catalog (ICAT) demonstrated adequate scaling for data registration
  - Optimization desired for searching huge datasets
- Good collaboration with the iRODS development team
- Keep Going
  - Operationalize iRODS within the NCCS

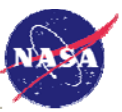


Data Analytics

- Can object stores with MapReduce provide a large scale data analytics environment for climate data?
- Bridge the gap between the archive and large parallel file systems

Proposed NASA Project

- Focus on soil moisture, precipitation, and atmospheric water-vapor, important classes of observation- and simulation-derived data products.
- Data Sets
  - GISS IPCC/AR5 monthly soil frozen water content (MRFSO), 1850 to 2000
  - MERRA monthly precipitation (PREC), 1979 to 2011
  - MODIS Atmospheres, 8-day global water-vapor product, 2000 to 2011
  - SMOS 3-day global soil moisture product, 2011
- Perform qualitative and quantitative evaluations of critical MapReduce attributes in our experimental HDRS, Swift, and AEMR contexts, including assessments of
  - Data preparation, ingest, and space complexities of creating repositories of the canonical data sets
  - Time and space complexities of server-side processing of the canonical operations along scaled ranges of their spatiotemporal parameters.
- Compare to server-side implementations of the canonical operations over the IBM GPFS file system




Data is the Future

- Drive the consumption of NASA data by users and technology
  - Inside the NASA community of modelers and scientists in both traditional and new methods
  - Outside of the traditional NASA scientific community to new communities in new ways
- This in turn will drive the need for more observation and model data from NASA (and other agencies)
- Provides the NCCS will a long term business plan to continue to provide these types of services

Serious Concerns and Questions

- If you sit down with the current NASA stakeholders and ask them what they need, they will say they need more storage.
  - It will be difficult at first to show them the benefit of this approach.
- Can we do this within our budget?
- What partnerships could we explore to work toward this future architecture?
- Can we make an evolutionary change or is this going to be disruptive?
  - Perhaps it is not as important strategically to store as much data as the users want but rather invest in the creation of these types of data services to drive the consumption of data.

Thank You! Questions

- Did you know? Fun Facts
  - NCCS has enough data storage to make an iPod playlist over 34,000 years long!
  - If NCCS printed out all of its data, the stack of paper would reach from the Earth to the moon, and well beyond!
  - Making the paper to print all that data would require about 640 million trees!
  - Using the NCCS network, you could download an HD movie in about 6 seconds!
  - NCCS currently has 1,600 kilowatts of power—the equivalent of 1,200 U.S. households!
  - It would require all 7 billion people in the world to add two numbers together over and over again for more than 13 hours to equal what the Discover supercomputer can perform in a single second!