... for a brighter future

UChicago ▶
Argonne LLC

A U.S. Department of Energy laboratory
managed by UChicago Argonne, LLC

Real Data Grids: Three Pragmatic Case Studies

Jennifer M. Schopf

Argonne National Laboratory
University of Chicago

What is a Grid? (or Cyberinfrastructure or eScience or …)

- Resource sharing
  - Computers, storage, sensors, networks, …
  - Sharing always conditional: issues of trust, policy, negotiation, payment, …
- Coordinated problem solving
  - Beyond client-server: distributed data analysis, computation, collaboration, …
- Dynamic, multi-institutional virtual orgs
  - Community overlays on classic org structures
  - Large or small, static or dynamic

Applications and Grids

- Why is this hard?
  - Lack of central control
  - Shared resources
  - Communication and coordination
- So why do it?
  - Computations that need to be done with a time limit
  - Data that can't fit on one site
  - Data owned by multiple sites
- Applications that need to be run bigger, faster, more

Why is this more common now?

- Geographically distributed user communities
  - Numerous labs, universities, industry
- Integration with other national resources
  - Inevitably multi-agency, multi-disciplinary
- Extremely large quantities of data
  - Petabytes, with complex access patterns
- Challenging interaction modalities
  - Interactive/collaborative/distributed analysis
  - Coupling of computation and instrumentation
  - Extreme performance requirements

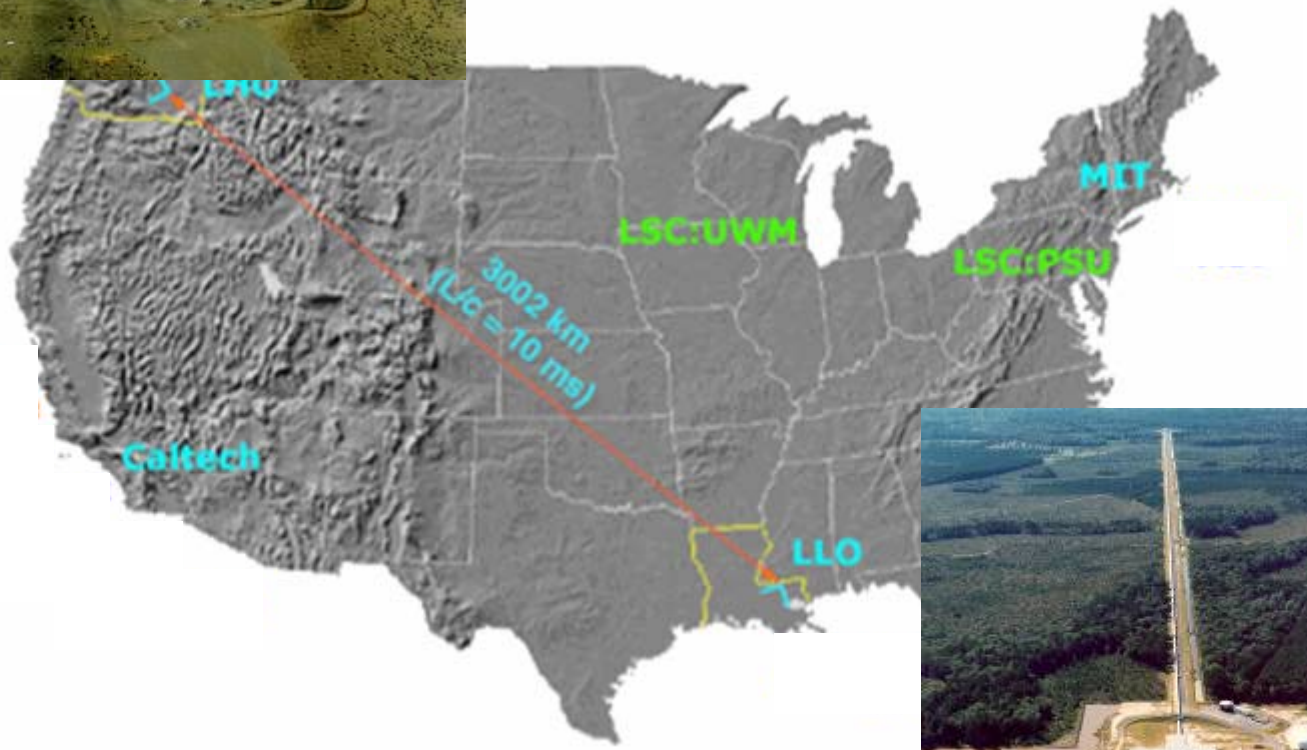
Grids and Data

- Distributed file management and policy
- Long term storage and curation issues
  - Policies for sharing and access
- Metadata
  - Need for automation and standards
- Easier access to own data
  - From laptop to local analysis work to collaboration-wide data storage
- Scalability when serving data to a large number of end users
- etc.

Outline

- Three applications and how they are currently working with their own growing distributed data concerns
- Laser Interferometer Gravitational wave Observatory (LIGO)
  - Replication
- Earth Systems Grid (ESG)
  - Ease of use for non-technical users
  - Security
  - Reliability
- Cancer Bioinformatics Grid (caBIG)
  - Interoperability across subfields
- Guidelines

LIGO Gravitational Wave Observatory



Some material compliments of Ann Chervenak (ISI), Scott Koranda (UWM)

LIGO: The Science

- Laser Interferometer Gravitational wave Observatory
- Measure time for light to pass between 2 mirrors
  - Space-time ripples cause the distance measured by a light beam to change as the gravitational wave passes by
- Two of the three interferometers operated in unison must agree in order to rule out false signals and confirm that a gravitational wave has passed through the earth

LIGO Data Needs in a Nutshell

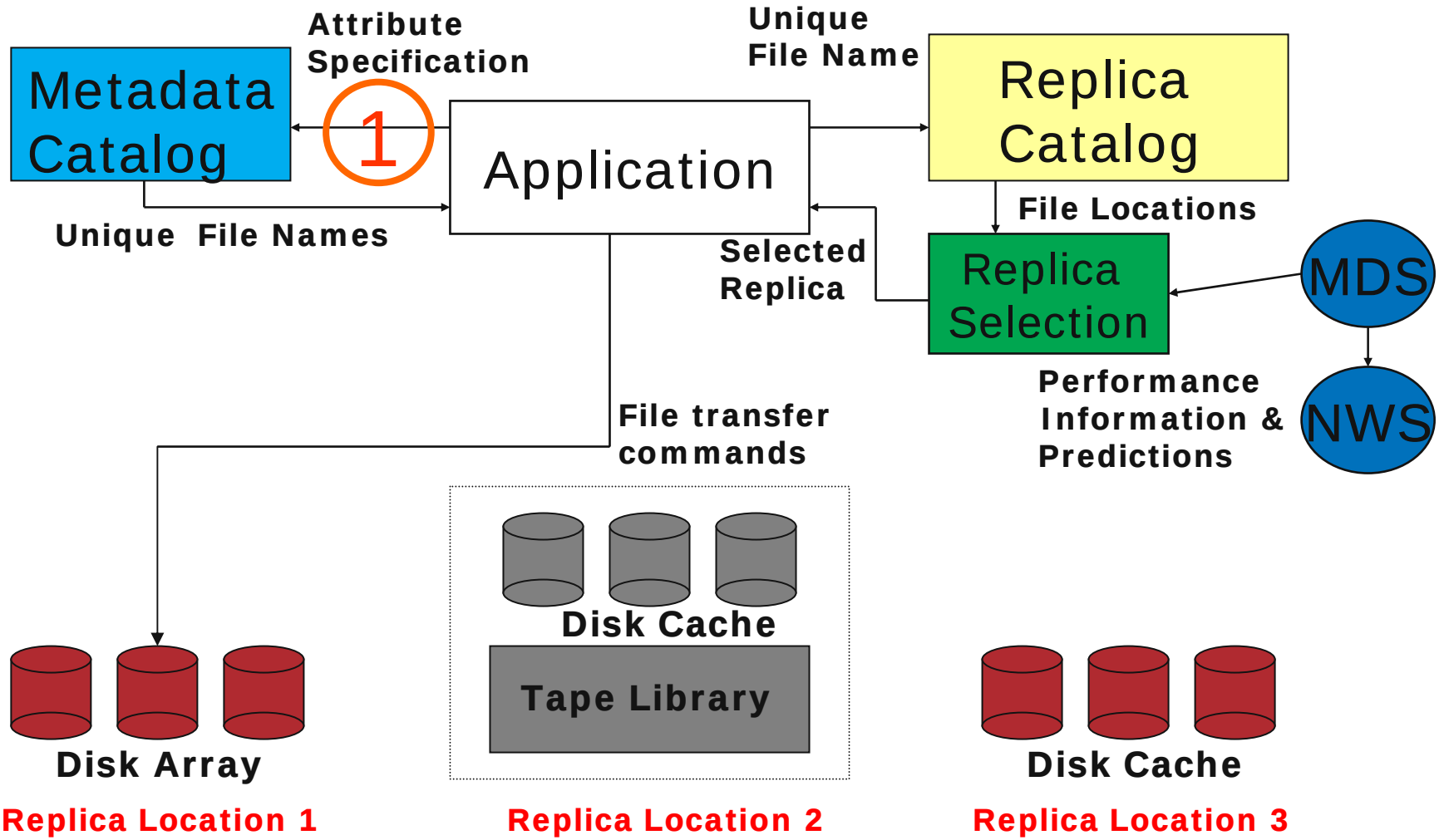
- Current Data
  - 25 Million unique files, with replications resulting in over 125 million actual files across 10 sites
  - Detector gathers data every 16 seconds
    - *1 to 100 Meg files, slightly less than a terabyte/day*
    - *Full data stored on tape at CalTech*
  - *Derived data sets* from users
    - *About 1/3 of the data being stored, but growing*
- End Users
  - Scientists studying gravitational waves
- Primary Technological Concern
  - Replication for faster access and reliability

The Challenge

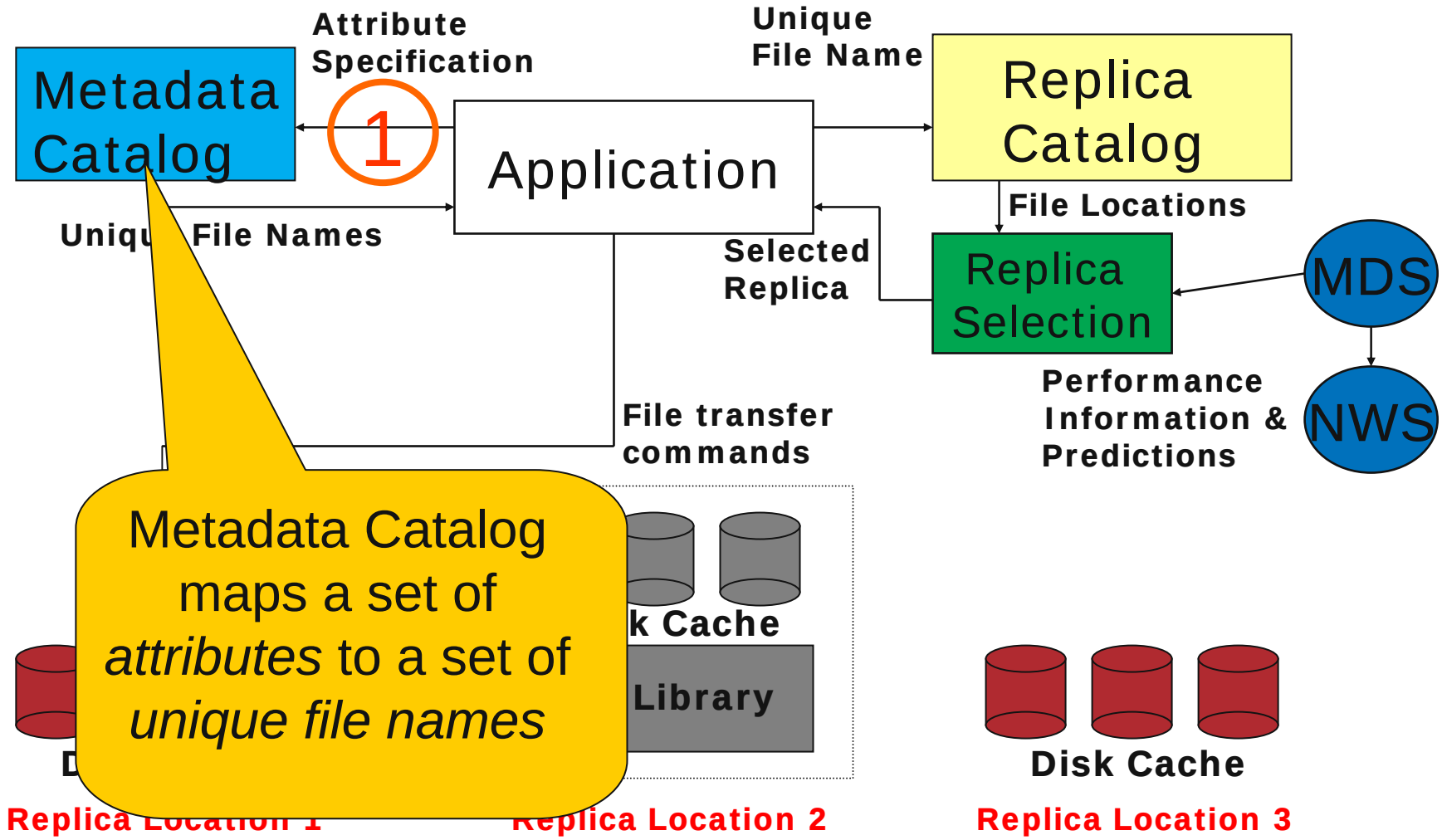
Replicate 1 TB/day of data to 10+ international sites.

- Provide scientists with the means to specify and discover data based on application criteria (metadata)
- Provide scientists with the means to locate copies of data

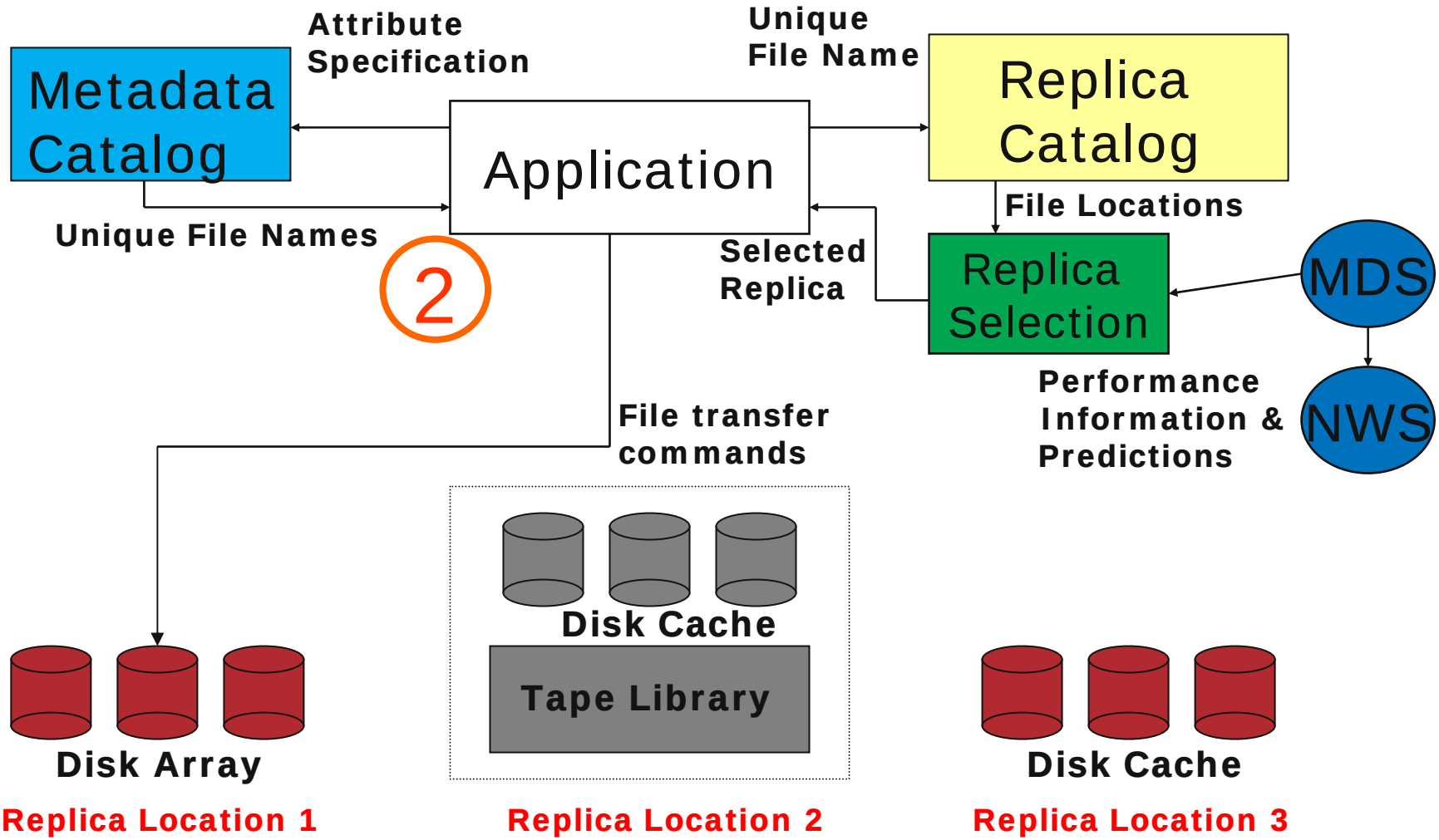
Basic Replication Selection Architecture



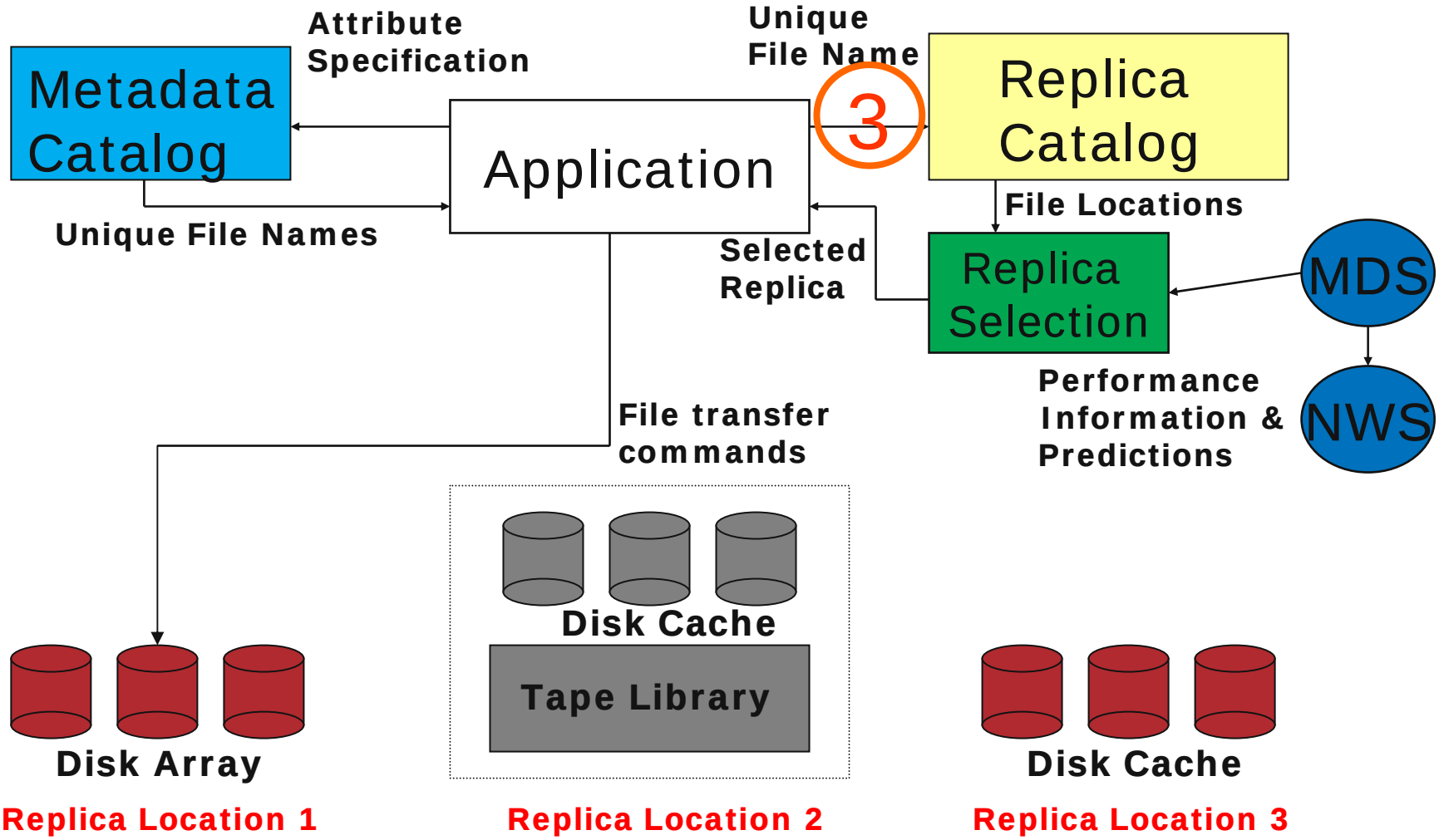
Basic Replication Selection Architecture



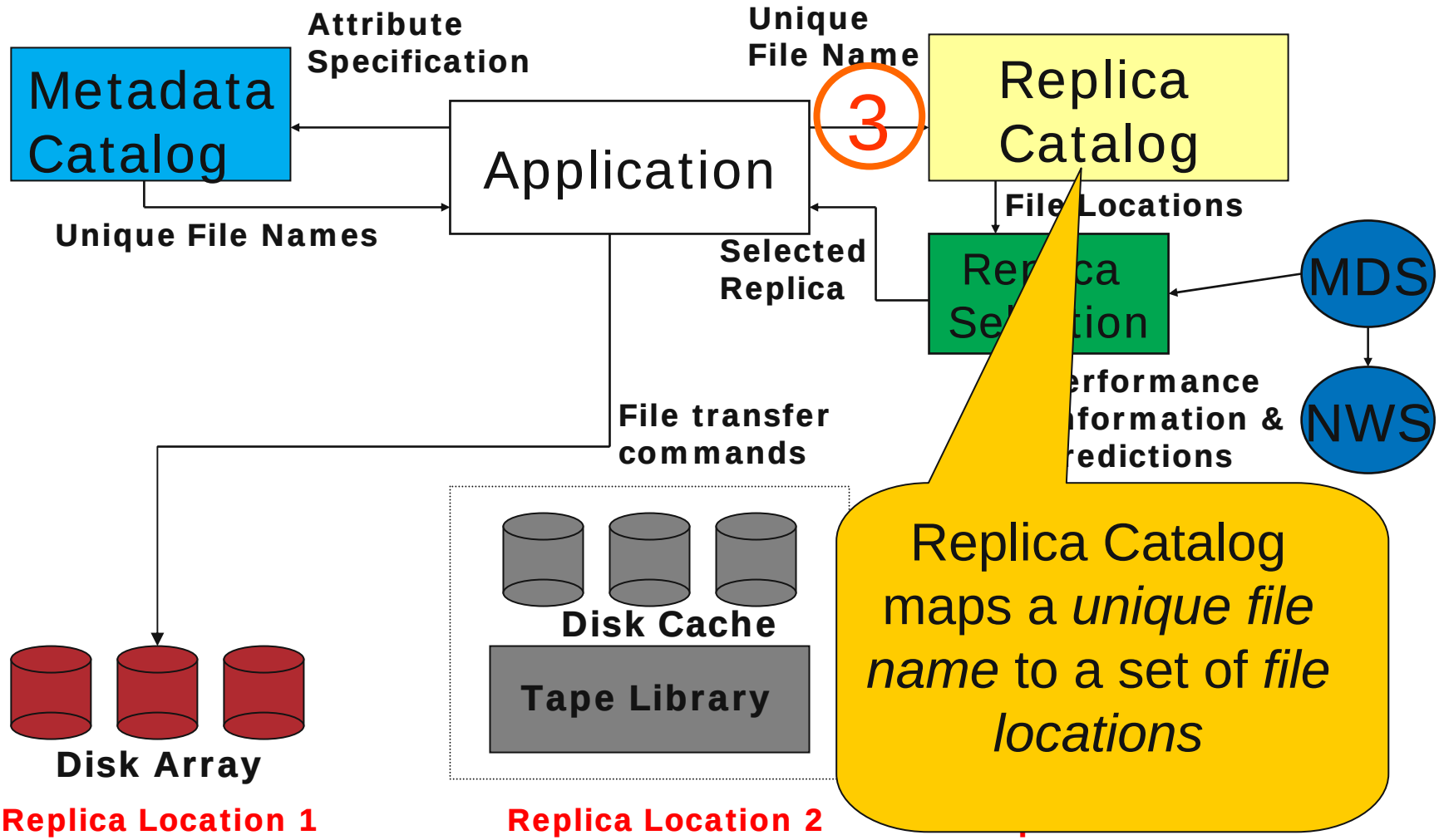
Basic Replication Selection Architecture



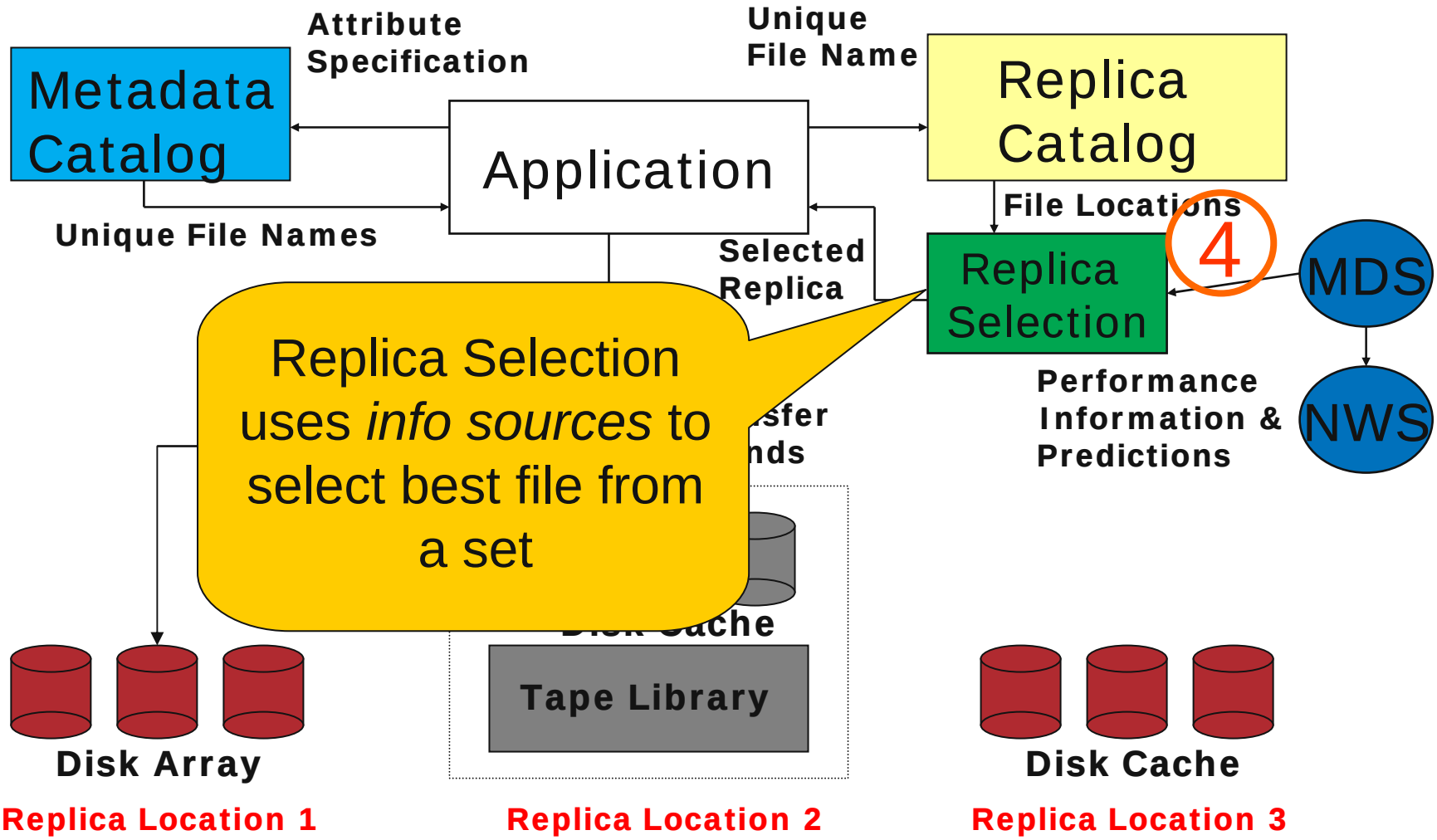
Basic Replication Selection Architecture



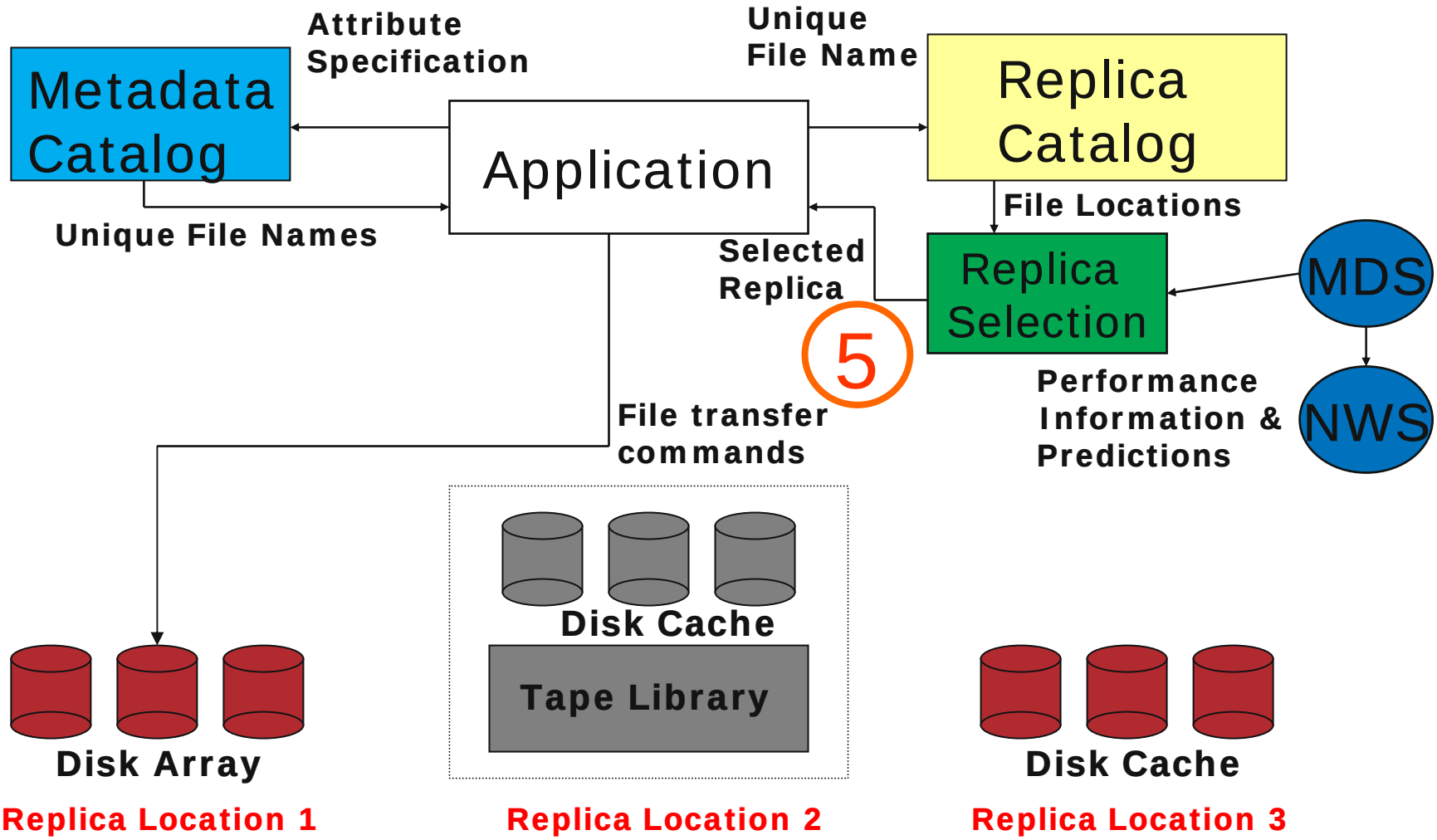
Basic Replication Selection Architecture



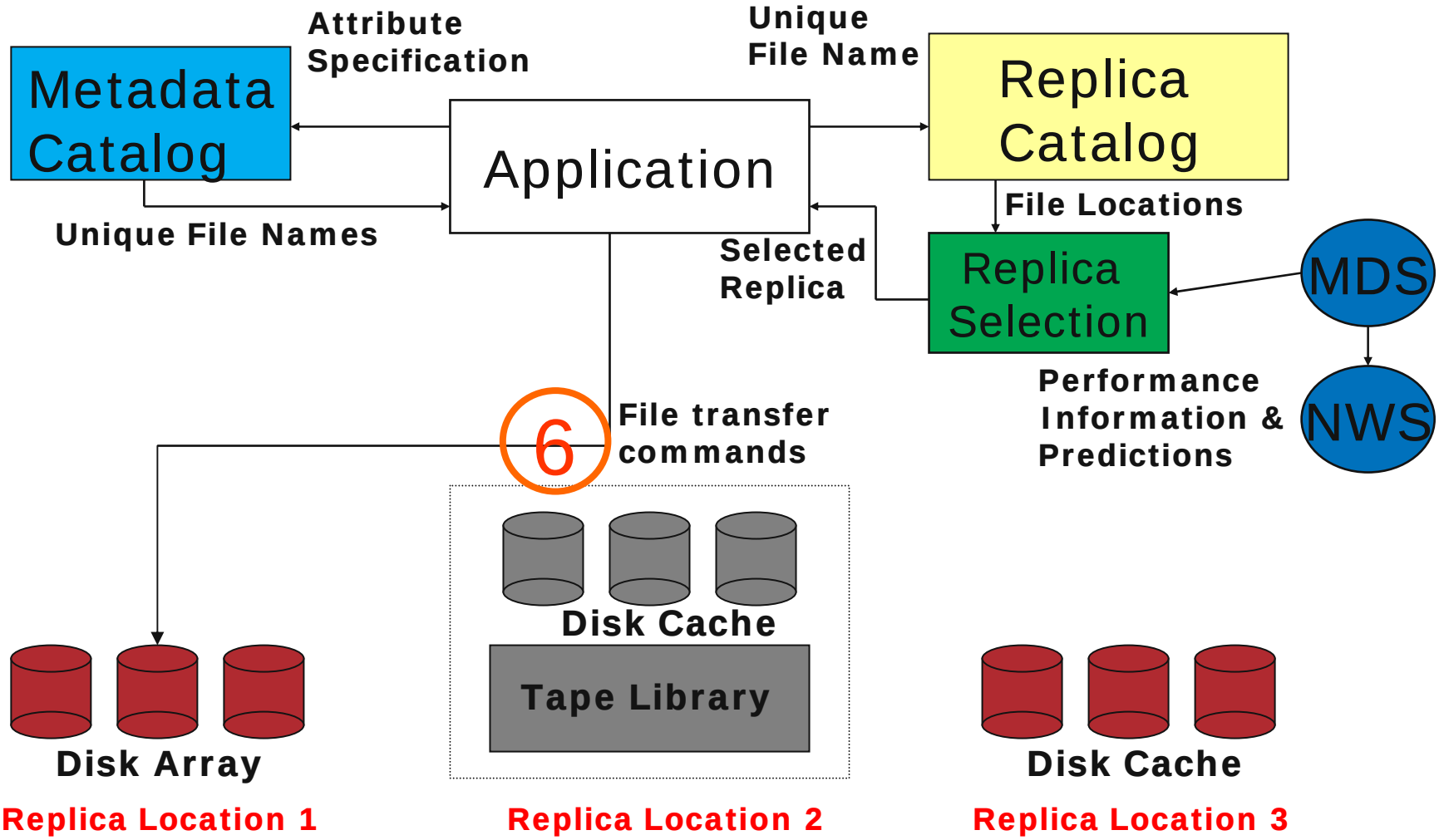
Basic Replication Selection Architecture



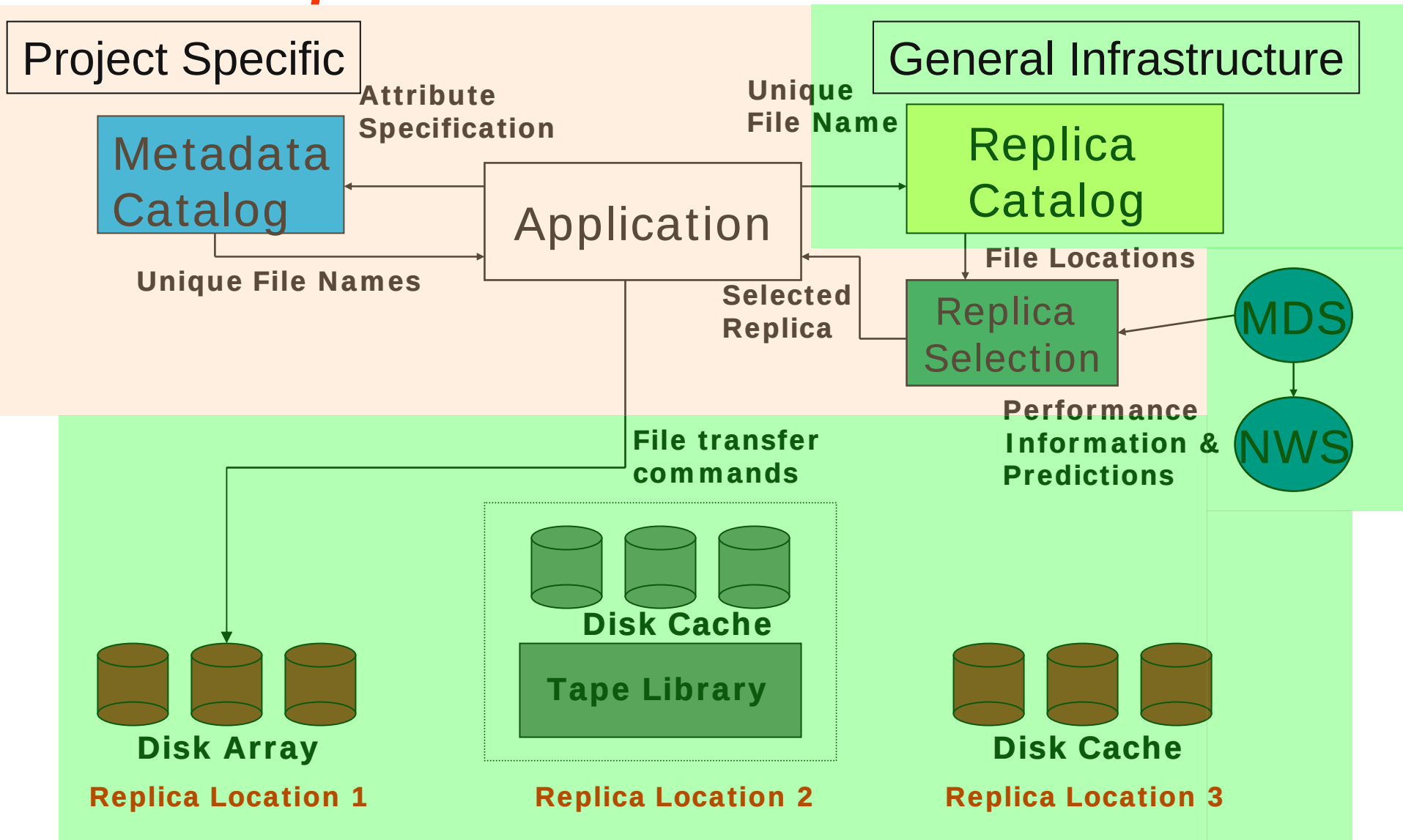
Basic Replication Selection Architecture



Basic Replication Selection Architecture



Basic Replication Selection Architecture



LIGO Data Grid: Before & After

Before:

- Data replication via "FedEx" Grid
- Ad-hoc site-by-site idioms for finding data in storage
- Ad-hoc error prone mapping from metadata to file names
- Workflow limited to a single compute resource site

After:

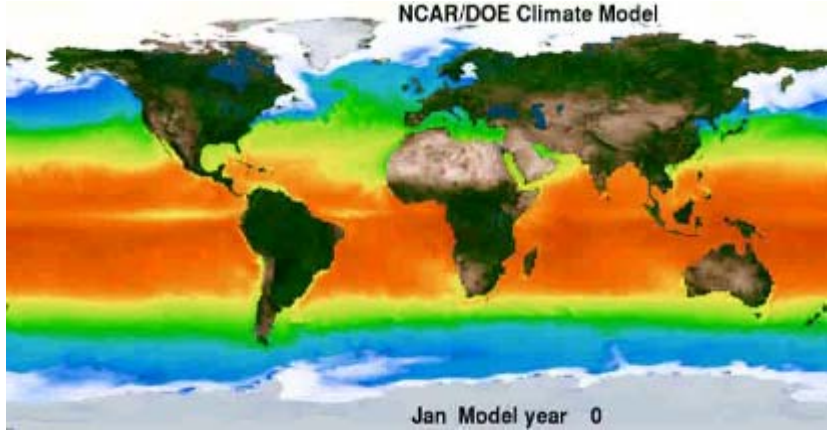
- 24 x 7 x 365 continuous fault tolerant data streaming
- Single client tool for scientists and applications to find data
- Scientists concentrate on metadata and not file names
- Multi-site planning of workflows across LIGO Data Grid

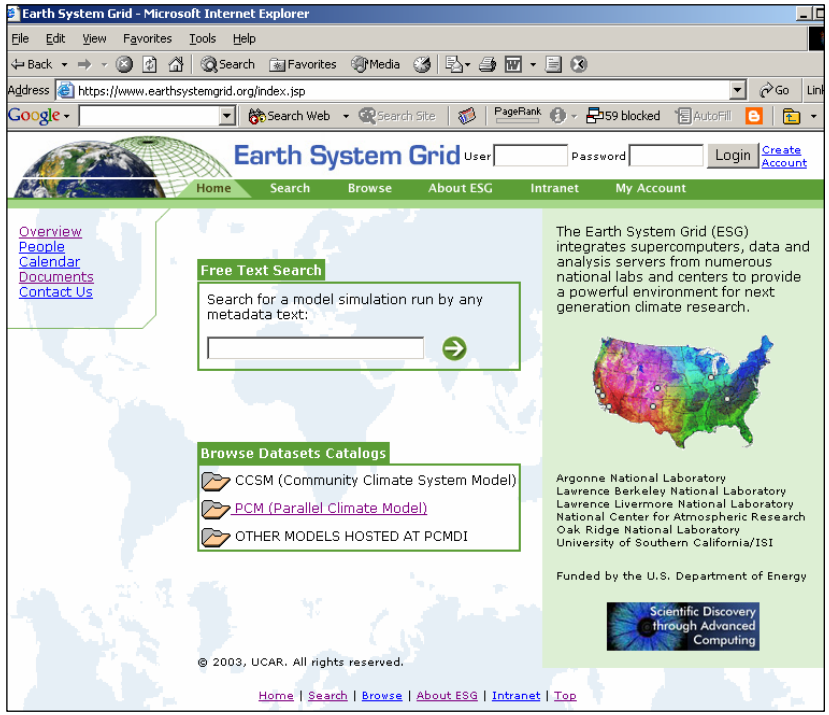
LIGO scientists searching for signals from neutron stars and black holes run **more jobs** across **more resources** and access **more data** using the LIGO Data Grid**.**

Papers are published faster due to the LIGO Data Grid.

Earth System Grid

- On-demand access to climate simulation data
  - Multiple archives
  - Interactive query
  - Per-collection control
  - Server-side processing
- Crucial technology to securely access, monitor, catalog, transport, and distribute climate data





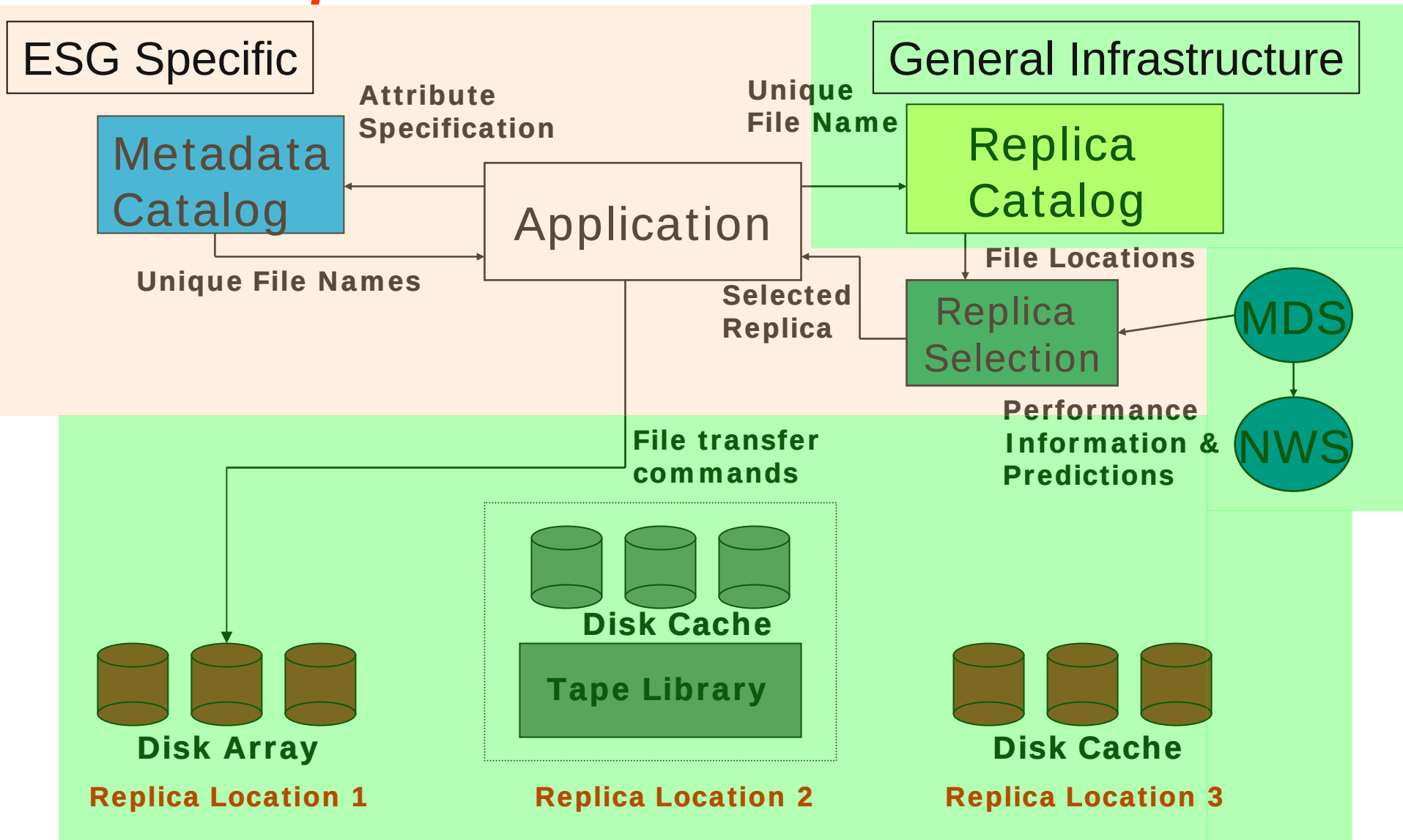
Some material compliments of D. Berholdt (ORNL), D. Middleton (NCAR)

ESG In a Nutshell

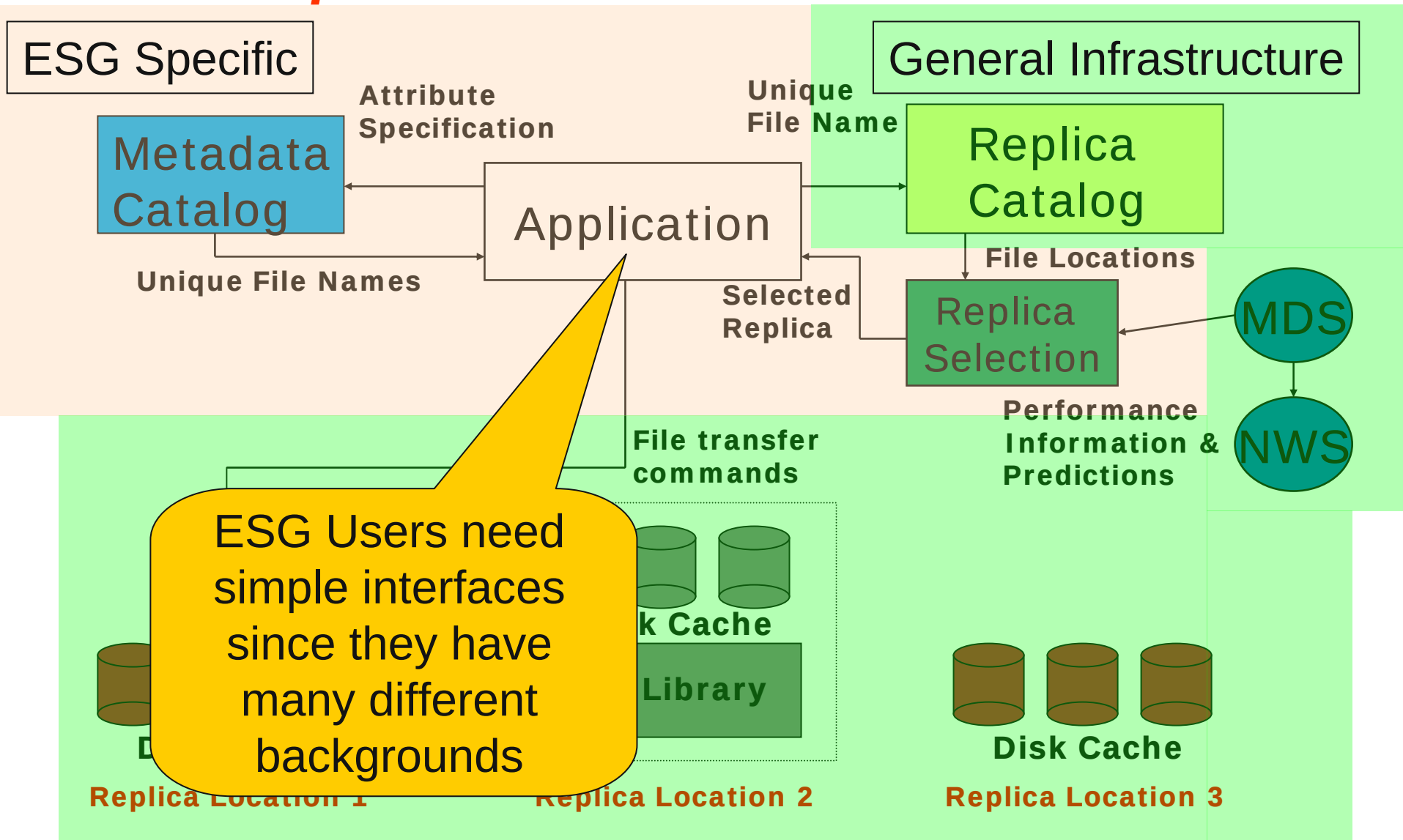
- Users
  - Scientists studying climate changes
  - Educational use
  - Governmental studies

- Primary concerns:
  - 24/7 delivery of data to a wide set of end users
  - Portal-based access
  - Troubleshooting infrastructure

Basic Replication Selection Architecture



Basic Replication Selection Architecture



Edit View History Bookmarks Tools Help

http://www.earthsystemgrid.org/

Google

ug - incubator list | Bug List- jms | edinburgh movies | esnet meetings | Movie Showtimes - G... | TV listings- edinburgh

gle Search Bookmarks PageRank Check AutoFill Send to Set

Earth System Grid (ESG)

Earth System Grid

Home | Data | About ESG | Login

ESG News

ew Multifile Download Options: ESG ow offers two ways of downloading ultiple files at once to the user achine: Data Mover Light (see DML structions) and the standard Unix mmand wget (see wget instructions). oth DML and wget are desktop clients at can be used to download large umber of files from the ESG system ithout clicking on multiple hyperlinks.

ew Browsing Options: now the ESG owsing interface allows to specify an otional string to match filenames, and select many files with one single click.

ata Subsetting: OPeNDAP-G offers gh-performance subsetting capabilities h a number of virtual aggregated atasets - see list of available ggregations

egistration is required to download me of the data: please request an count. Please send us comments or edback.

ew: IPCC Working Group 1 data ailable.

e NCAR MSS is scheduled for owntime each Sunday morning from

Data Search

Search Dataset metadata for:

Search

Examples: c02, B06.77

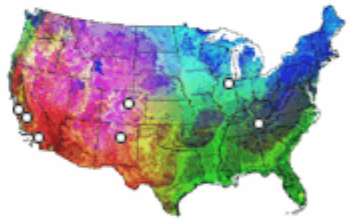
Browse Dataset Catalogs

- CCSM (Community Climate System Model)
- CCSM POP (modified version of Parallel Ocean Program)
- CSIM (CCSM Sea Ice Model)
- CLM (CCSM Community Land Model)
- PCM (Parallel Climate Model)
- POP (Parallel Ocean Program)
- Scientific Data Processing and Visualization Software

Shortcuts menu: - jump to -

Welcome to ESG

The Earth System Grid (ESG) integrates supercomputers with large-scale data and analysis servers located at numerous national labs and research centers to create a powerful environment for next generation climate research. This portal is the primary point of entry into the ESG.



ESG Collaborators

- Argonne National Laboratory
- Lawrence Berkeley National Laboratory
- Lawrence Livermore National Laboratory
- Los Alamos National Laboratory
- National Center for Atmospheric Research
- Oak Ridge National Laboratory
- University of Southern California/Information Sciences Institute


http://www.earthsystemgrid.org/metadata/searchSimulation.do
Google
Edit View Favorites Tools Help
earthsystemgrid.org
Go
Bookmarks
PageRank
3 blocked
Check
AutoLink
AutoFill
T-Mobile - HotSpot
Bug processed
T-Mobile - HotSpot
Earth System...
http://users.sdsc...
Page
Tools
Earth System Grid
Home
Data
About ESG
Login
Subset | Search | Browse | My Data Cart | My Data Requests
Data Search
Search Dataset metadata for:
c02, B06.77
Search
Examples: c02, B06.77
QUERY "c02, B06.77" RESULTS 1 - 3 of 3
1. PCM (Parallel Climate Model) run B06.77
description: 1890-2000 volcano only run
View metadata
Get data
2. PCM (Parallel Climate Model) run B07.21
description: Solar only with fixed 1890 ozone
View metadata
Get data
3. PCM (Parallel Climate Model) run B07.31
description: Solar only with fixed 1890 ozone
View metadata
Get data
User: guest | ESG Home | Help Page | Contact Us | My Data Cart
Portal Software version 4.0 © UCAR, all rights reserved.
Earth System Grid, funded by the U.S. Department of Energy
Internet
100%

Simulation

id: ucar.cgd.pcm.B06.77
description: 1890-2000 volcano only run
date (end_generated): 2002-04-01 (encoding:)
date (start_generated): 2002-04-01 (encoding:)
note: GHGs: fixed at 1890 levels Sulfates: fixed at 1890 levels Ozone: cycle thru 1890 ozone data (new file: noaao3.1890.093098.nc) Solar: fixed at 1367 GHGs and sulfates fixed by code mod to ramp.F, where the CO2, N2O, CH4, F11, F12, and sulfate scale factor were hardwired to 1890 conditions.
note: Repeat B06.60 and B06.66, branch from case B05.14, 1890-0101, but volcano only.
note: Integration run as follows: bettge@ucar.edu model time 1890-01 to 1999-12
note: Documented at relevant pages under http://www.cgd.ucar.edu/pcm.
note: All components (atmosphere, ocean, sea ice, land surface) active.
isPartOf: ucar.cgd.pcm
participant: ucar.cgd.pcm.TomBettge (Submitter)
simulationInput (data): pcm_volcanic_1890-2000.T42.nc, noaao3.1890.093098.nc, surfdat.T42.0596
simulationHardware: IBM cheetah cheetah.ccs.ornl.gov

PCM run B06.77 monthly atmosphere data

ESG metadata: http://www.earthsystemgrid.org/metadata/displayMetadata.htm?id=ucar.cgd.pcm.B06.77.atm

+ Metadata

File-level access

You may either download a file by clicking on it (hyperlink in the first column, if available), or add files to your **Data Cart** to prepare a multi-files request.

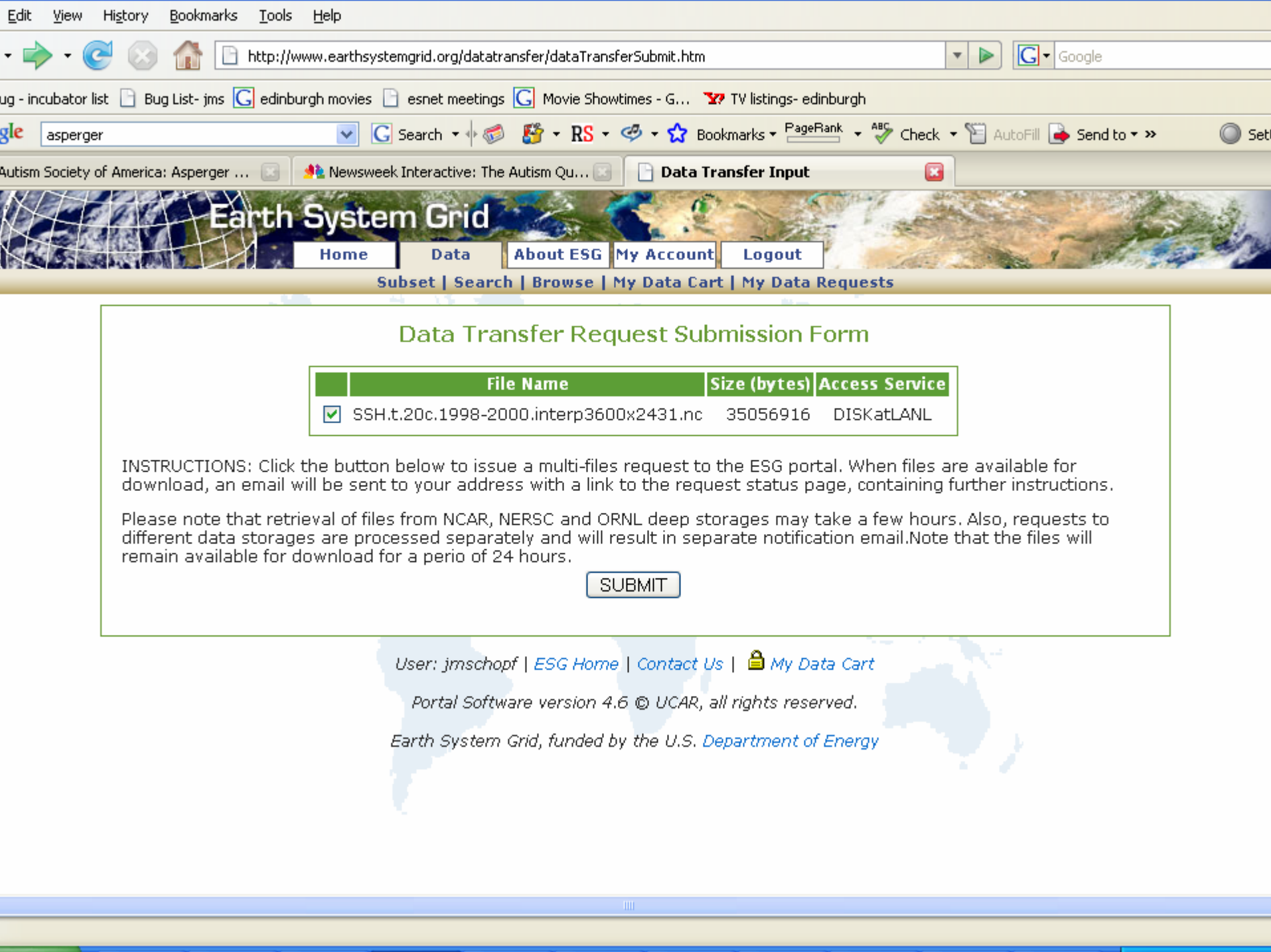
1-20 of 50 datafiles | 21-40 | 41-50

start from file #: 0 +1 and display 20 files per page (max:100) (NEW: optional filename match:) go

NEW: Multifile Download via Desktop Client (DISK data only) Generate DML input file (see DML instructions) | Generate *wget* script (see wget instructions)

My Data Cart: Add selected files to Data Cart | Empty Data Cart | Go to Data Cart .

File	Metadata	Format	Type	Size	Add to Data Cart	OPeNDAP
1. B06.77.atm.1890.nc				327575484	☐ ORNL HPSS	
2. B06.77.atm.1891.nc				327575484	☐ ORNL HPSS	
3. B06.77.atm.1892.nc				327575484	☐ ORNL HPSS	
4. B06.77.atm.1893.nc				327575484	☐ ORNL HPSS	
5. B06.77.atm.1894.nc				327575484	☐ ORNL HPSS	
6. B06.77.atm.1895.nc				327575484	☐ ORNL HPSS	
7. B06.77.atm.1896.nc				327575484	☐ ORNL HPSS	


Edit View History Bookmarks Tools Help
http://www.earthsystemgrid.org/datatransfer/dataTransferSubmit.htm
Google
ug - incubator list
Bug List- jms
edinburgh movies
esnet meetings
Movie Showtimes - G...
TV listings- edinburgh
asperger
Search
Bookmarks
PageRank
Check
AutoFill
Send to
Autism Society of America: Asperger ...
Newsweek Interactive: The Autism Qu...
Data Transfer Input
Earth System Grid
Home
Data
About ESG
My Account
Logout
Subset | Search | Browse | My Data Cart | My Data Requests
Data Transfer Request Submission Form
File Name
Size (bytes)
Access Service
SSH.t.20c.1998-2000.interp3600x2431.nc
35056916
DISKatLANL
INSTRUCTIONS: Click the button below to issue a multi-files request to the ESG portal. When files are available for download, an email will be sent to your address with a link to the request status page, containing further instructions.
Please note that retrieval of files from NCAR, NERSC and ORNL deep storages may take a few hours. Also, requests to different data storages are processed separately and will result in separate notification email.Note that the files will remain available for download for a perio of 24 hours.
SUBMIT
User: jmschopf | ESG Home | Contact Us | My Data Cart
Portal Software version 4.6 © UCAR, all rights reserved.
Earth System Grid, funded by the U.S. Department of Energy

Data Transfer Request Monitoring

Data Request:

Name: ESG jmschopf DISKatLANL 20070825 104713
ID: ESG_jmschopf_DISKatLANL_20070825_104713
Status: SUCCESS
Error Message:
Submitted: 2007-08-25 10:47:13
Last Update: 2007-08-25 10:47:13

Download Instructions:

After a request has completed, datafiles can be downloaded in one of the following ways:

1. Click on each file *download* hyperlink
2. Generate XML file to be used as input to Data Mover Light (see DML instructions)
3. NEW: Generate wget script (*wget* is a program available only on Unix systems - see wget instructions)

Data Files:

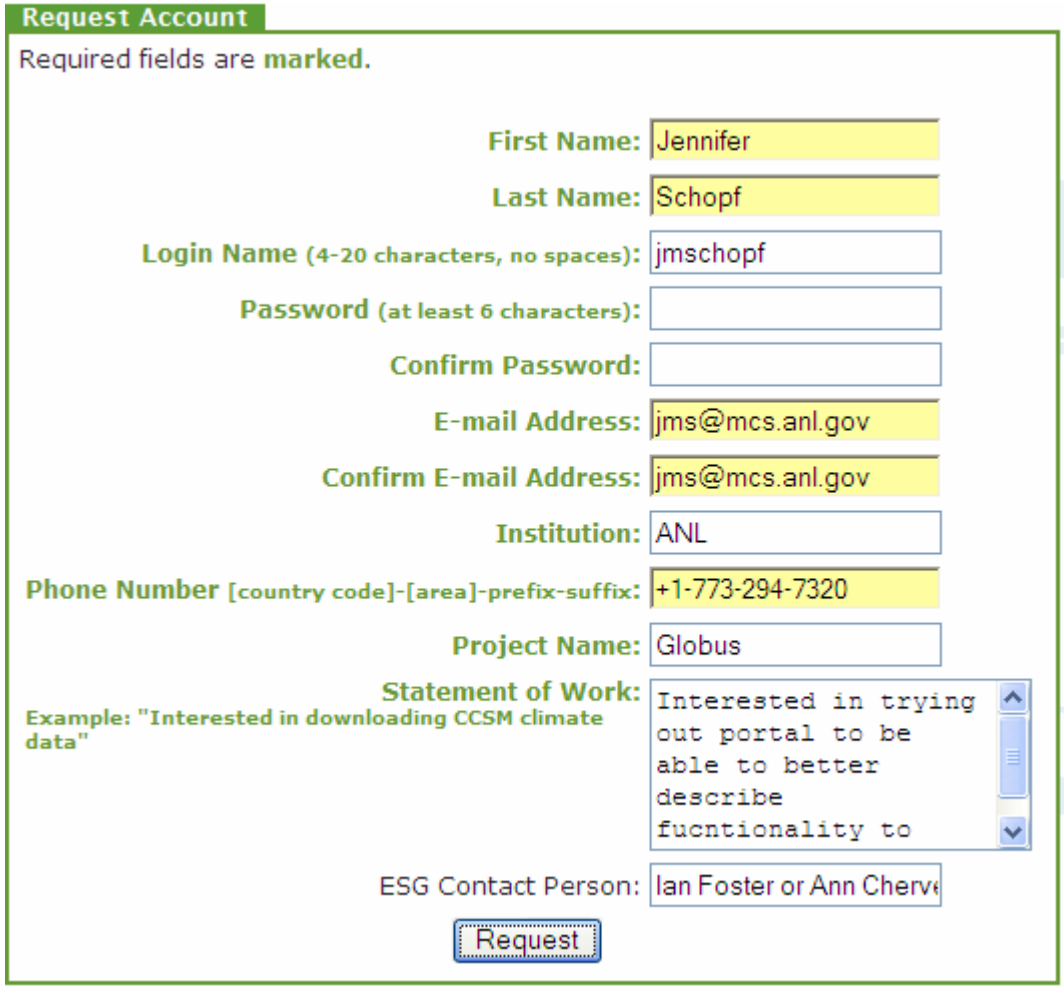
Name	Status	Submitted	Completed	Expiration	Size	HyperLink	Error Message
SSH.t.20c.1998-2000.interp3600x2431.nc	SUCCESS	2007-08-25 10:47:13	2007-08-25 10:47:13		35056916	download	

But What About Security?

- Security needed so that ESG software can act on users behalf as well
- Even if data is public, data access needs to be tracked
- In the Grid space, this is generally done using x509 electronic certificates
  - Digital document that is part of a public key private key infrastructure
  - Signed by 3rd party – the Certificate Authority (CA)
  - Often not easy for users to interact with
  - Can be heavy weight for and administrators as well
- ESG uses a system called PURSE: Portal-Based User Registration Service

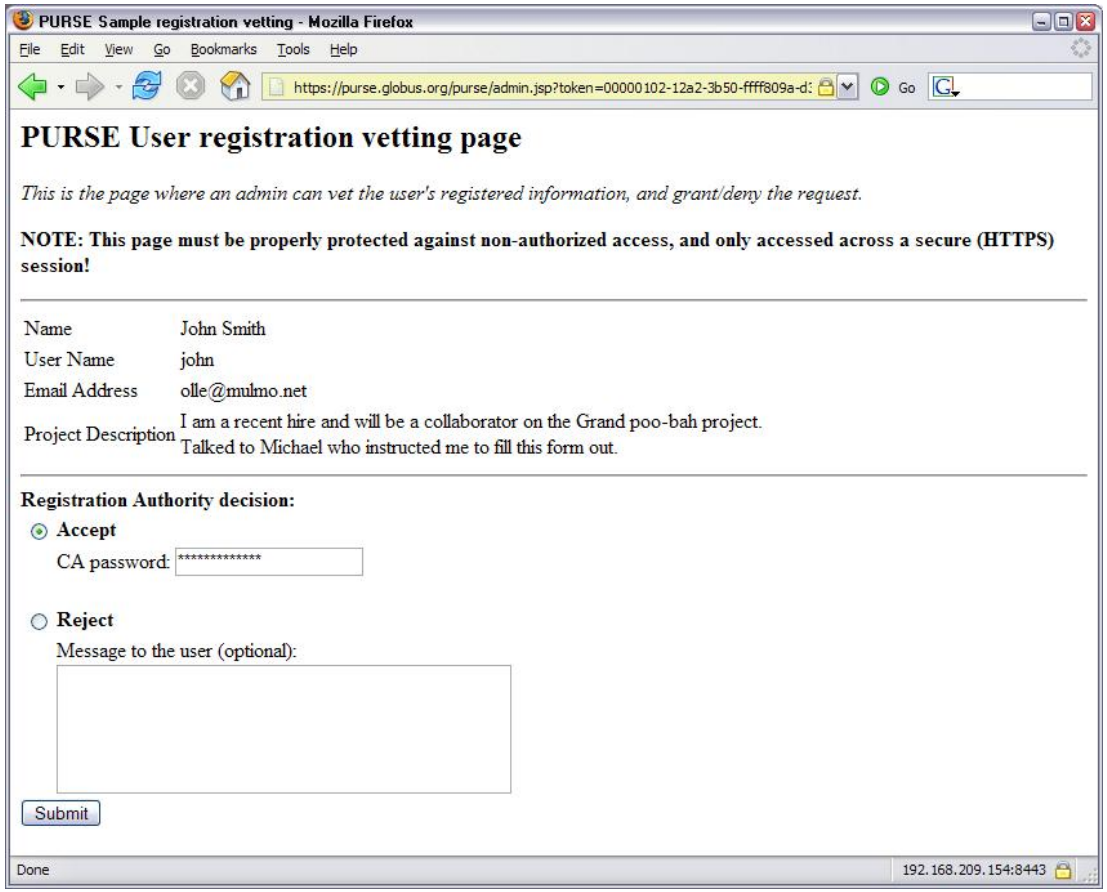
User Registration

- The user fills out the registration Web form
  - Establishes an ID/password
  - Information is stored in PURSE database
- The administrator is sent email



Administrator Approval

- Administrator visits the registration Web page and retrieves the registration data
- If the administrator approves the request, system generates a certificate and stores it in a proxy repository
- The user is sent email



User Login

- The user logs into the application Web site using the ID/password established during registration
- The application obtains a proxy using from the proxy repository
- The application uses the proxy to authenticate to Grid services and allows other services to act on behalf of the user

- ESG manages to track their data use
- Services run as known users (safely with Grid Security Infrastructure)
- Users have very easy access

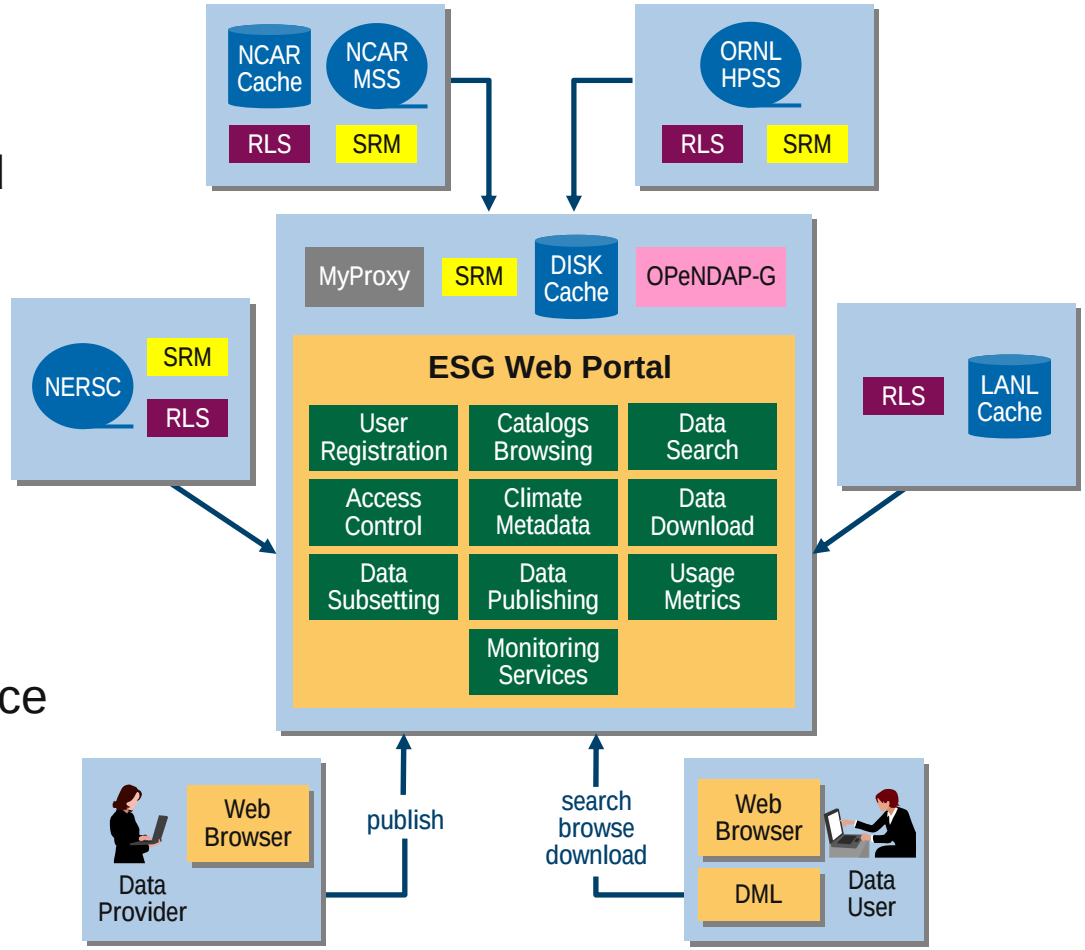
Reliability

- ESG users are 24 x 7
- ESG sys admins are not!

- Distributed systems have many components, and the more pieces added into the mix, the more likely a failure is to occur

ESG Technologies

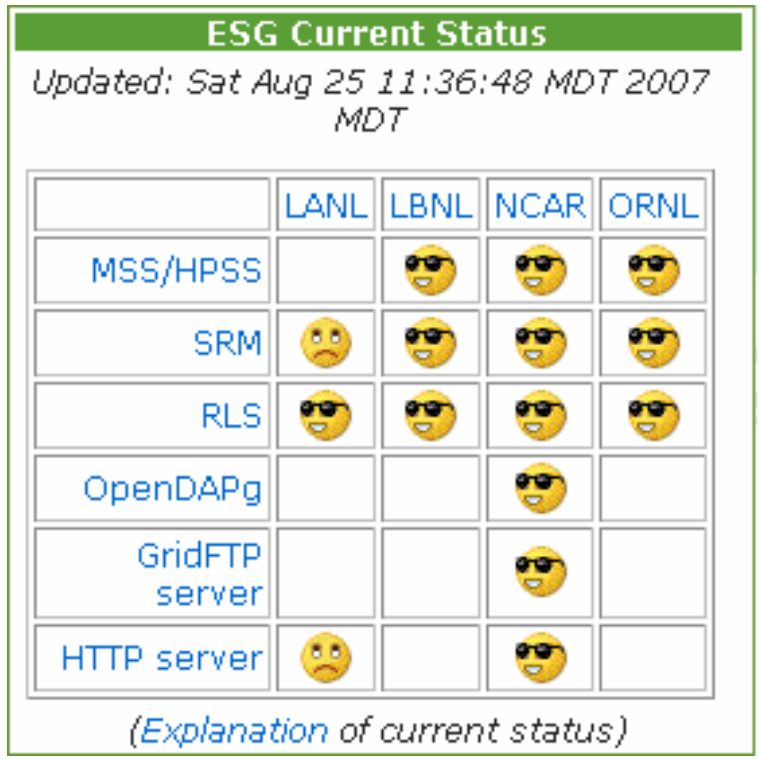
- Climate data
  - Metadata catalog
  - OPenDAP-G (aggregation and subsetting)
- Data management
  - Data Mover Lite
  - Storage Resource Manager
  - Globus Security Infrastructure
  - GridFTP
  - Globus Replica Location Service
- Security services
  - Access control
  - MyProxy
  - PURSE User registration



MSS, HPSS: **Tertiary data storage systems**

Monitoring Overall System Status

- Deployed Globus Monitoring and Discovery System (MDS4)
- Information providers check resource status every 10 minutes
- Report status to Index Service
  - Collection point for data
  - One place to query for overall status
- Information in Index Service is queried by ESG Web portal
  - Used to generate overall picture of state of ESG resources



MDS4 Trigger Service

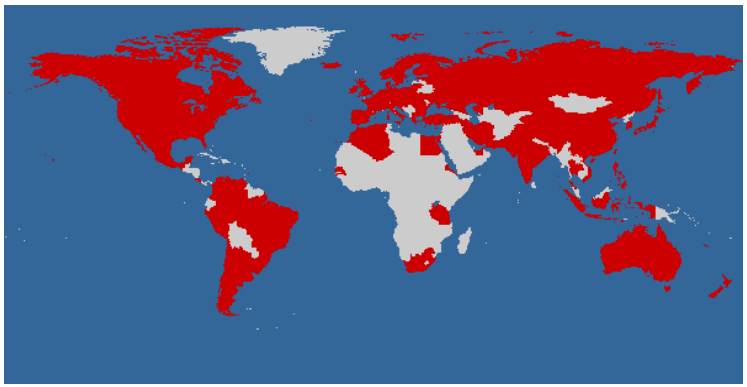
- Use same information providers to feed data to a warning system
- Evaluate that data against a set of pre-configured conditions (triggers)
- When a condition matches, action occurs
  - Email is sent to pre-defined address
  - Web site updated

ESG: Warning on Errors Sample

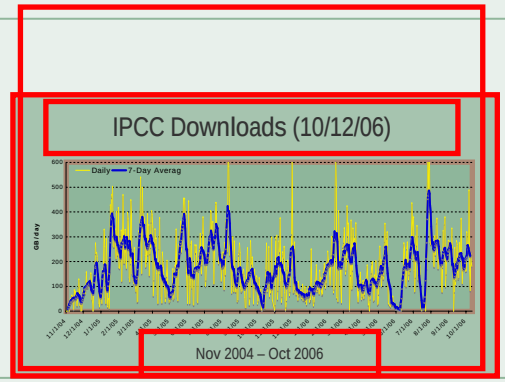
Total error messages for one month (May 2006)	**47**
Messages related to certificate and configuration problems at LANL	38
Failure messages due to brief interruption in network service at ORNL on 5/13	2
HTTP data server failure at NCAR 5/17	1
RLS failure at LLNL 5/22	1
Simultaneous error messages for SRM services at NCAR, ORNL, LBNL on 5/23	3
RLS failure at ORNL 5/24	1
RLS failure at LBNL 5/31	1

ESG facts and figures

Main ESG Portal	IPCC AR4 ESG Portal
130 TB of data at four locations ■ 840,331 files ■ Includes the past 6 years of joint DOE/NSF climate modeling experiments	35 TB of data at one location ■ 77,400 files ■ Generated by a modeling campaign coordinated by the Intergovernmental Panel on Climate Change ■ Model data from 11 countries
3,200 registered users	1,245 registered analysis projects
Downloads to date ■ 25 TB ■ 91,000 files	Downloads to date ■ 245 TB ■ 914,400 files ■ 500 GB/day (average)

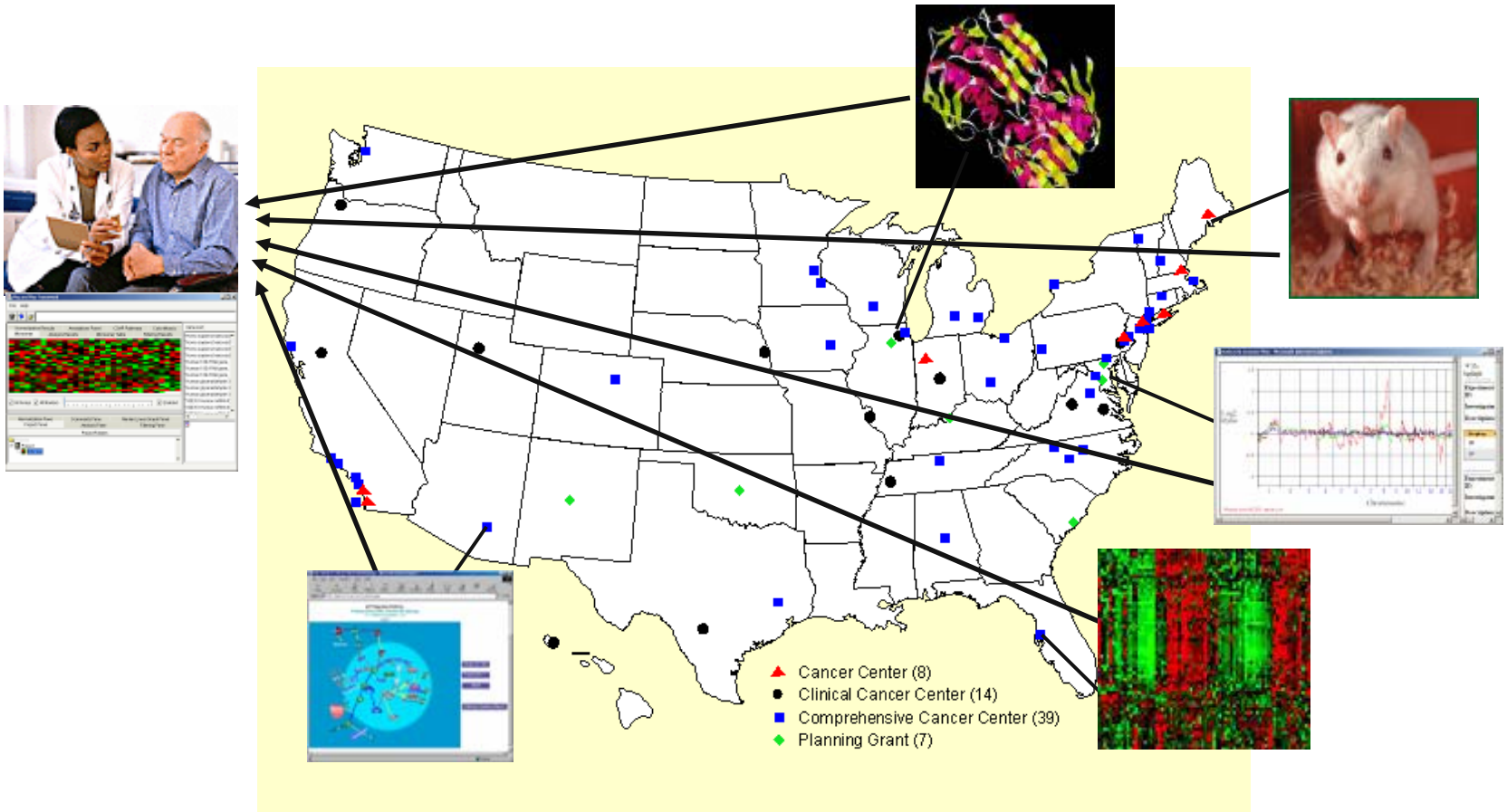


Worldwide ESG user base



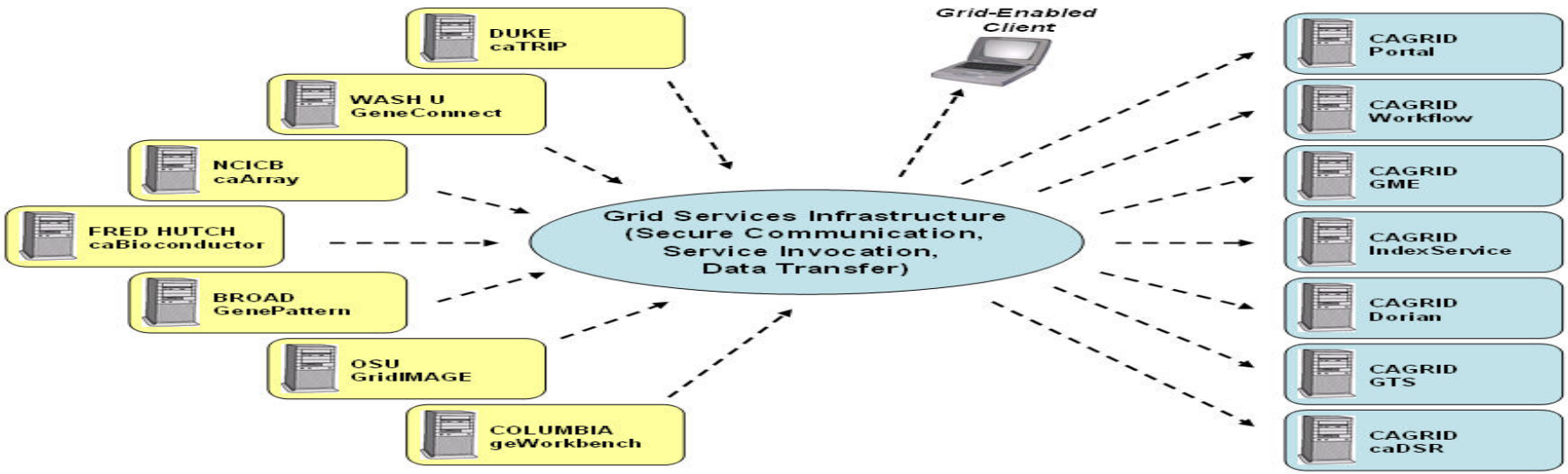
> 300 scientific papers published to date based on analysis of IPCC AR4 data

caBig: Cancer Biomedical Informatics Grid



caBIG Overview

- caBIG Principles
  - **Open Source, Open Access, Open Development, Federated, Syntactic and Semantic Interoperability**
- Driven by cancer research community requirements
- Services-Oriented Architecture
- Metadata driven and implements Virtualization
- Standards based



caBIG In a Nutshell

- Users
  - Scientists doing data analysis for cancer studies
- Primary concern:
  - Interoperability of services and data sets

Interoperability

- Functional (syntactic) interoperability
  - The ability to exchange information
- Semantic interoperability
  - The ability to use the information that has been exchanged
- Standards are frequently thought of as focusing primarily on functional interoperability
- Critical for health care is semantic interoperability
  - If you don't know exactly what the information you received means, having received it is useless, or worse

Functional Interoperability

- Well documented public API access to data
  - Use standard tools
  - Define common security infrastructure
  - Standardize service and metadata advertisement mechanisms
- Data Services
  - Expose data to the Grid in a unified way
- Analytical Services
  - Expose analytical operations to the Grid
- Compatibility testing
  - Four levels: legacy, bronze, silver, gold
  - Reflect a tool's adherence to guidelines

Semantic Interoperability

- Metadata is how we describe our data
  - What the data is, how it was collected, accuracy statements, etc
- caBIG meta data issues
  - Different dialects present for different subtopics
  - Need for cross-area collaboration has grown
- caBIG approach
  - No particular technology or tool specified
  - Use standards whenever possible
  - Be able to “unambiguously” and programmatically determine the meaning of data

How to do this?

- Create an explicit, global shared context
  - Terminology – shared terms, codes, and precise definitions
  - Information model – data elements, relationships and definitional (terminological) links
- Map existing information structures onto this space
  - Translation
  - Sandboxing
    - *Wrappers around current metadata to adapt to community standard*
- Define future information systems requirements in terms of a shared information model and terminology

caBIG process

- Develop caBIG compatibility guidelines in the areas of information modeling, metadata and vocabularies
- Provide mentors to caBIG-funded development projects
- Facilitate development of caBIG CDE and Vocabulary Standards
- Perform compatibility reviews to ensure compliance with compatibility guidelines for semantics

Metadata is an Ongoing Problem

- Communities that share data need metadata standards
  - This is often overlooked in project planning, but it is essential and non-trivial
- Metadata is as much a social challenge as a technical one
  - Requires different expertise, including a heavy dose of expertise and experience in the problem domain
  - Can’t expect an IT team to solve this on their own
- Data curation is a job description, not a software feature

caBig Summary

- Since February 2004, that community has swelled to over 600 researchers
  - Of NCI's 61 chosen cancer institutes, 50 are now participating
- caGrid test bed started 2005, production in March '07
  - 10 nodes
  - 42 applications and datasets
- "Another important result [is the] dramatic shift in the research mind-set away from individual analysis and toward more open communication and cooperation among physicians and scientists."
  - Ken Buetow, director Center for Bioinformatics, NCI

Common Issues

- Need to tune your approach to your end users
  - Both data users and admins
  - No two projects are the same, but many have common components
- Replication is the basic approach to dealing with
  - Basic widespread data access
  - Reliability
- Ease of use is a concern
  - Need to know users comfort space
- Metadata is the key to sharing, and a process that continues

How to start out

- Talk to your users- find out what's needed
  - How can you make it easier for them to do science
  - Don't solve the wrong problem
  - Listen carefully to requirements vs wish lists
- Do something real
  - If it's a stunt no one will really care
- Start small
  - Work in as controlled an environment as possible
  - Start with experienced users, move to simpler interfaces as a second stage

How to start out (2)

- Close interactions
  - Need a team that's in it as a team
  - Embed the IT folks
  - Iterative process
- Think long term
  - Data *must* be tagged with metadata to be useful to others – do what you can so this is automatic
  - Talk to others who might eventually be interested in your data sets to include what they need as well

How to start out (3)

- Leverage what you can
  - One-off solutions may seem best in the short term but rarely are in the long term
  - Look at what other people are doing and steal what you can – imitation is the sincerest form of flattery
- Establish policies
  - How long do you need to store data?
  - Who should have access and when?
- Tell people about it!
  - Show them you have something that works
  - Let other groups know what's feasible

More Information

- Jennifer M. Schopf
  - jms@mcs.anl.gov
  - http://www.mcs.anl.gov/~jms
- LIGO
  - http://www.ligo.caltech.edu/
- Earth Science Grid
  - http://www.earthsystemgrid.org/
- caBIG
  - https://cabig.nci.nih.gov/