



the globus project  
www.globus.org

# The Data Grid:

## An Architecture for Distributed Management of Large Scientific Data Sets

Ann Chervenak  
Carl Kesselman

Ian Foster  
Chuck Salisbury  
Steve Tuecke

Information Sciences Institute,  
University of Southern California

Argonne National Laboratory



# Overview

- Target Environment
- Design Principles
- Grid Services
  - ◆ Storage systems
  - ◆ Metadata
  - ◆ Management of replicated files
- Implementation



# Data Grid Environment

- Scientific applications
  - ◆ Global climate change, High energy physics
- Computationally demanding
- Large data sets and archives
  - ◆ Terabytes, eventually petabytes
  - ◆ Raw and derived data
- Geographically dispersed users and resources
  - ◆ Data replication for enhanced performance
- Broad range of capabilities and resources
  - ◆ Networks, systems, storage, and applications



# Building a Data Grid: Building Blocks

Ingest/  
catalog  
service

pftp  
GASS

MCAT,  
SRB

STACS,  
Condor  
others

Globus toolkit: security, information,  
fault detection, resource management,  
communication, etc.

MPI-IO

Netlogger

Condor

Akenti

Autopilot

...

HPSS  
Archival, multi-PB  
Access > 100 MB/s  
No QoS

QoS

DPSS  
Fast disk cache  
No QoS

QoS

Computers  
Preliminary QoS  
work (e.g., DSRT)  
XFS: QoS for disk

QoS

ESnet, MREN, B/s  
NTON  
QoS: e.g., diffserv



# Data Grid Objectives

- Integrate heterogeneous data archives into a distributed data management “grid”
- Identify services for high performance, distributed, data intensive computing

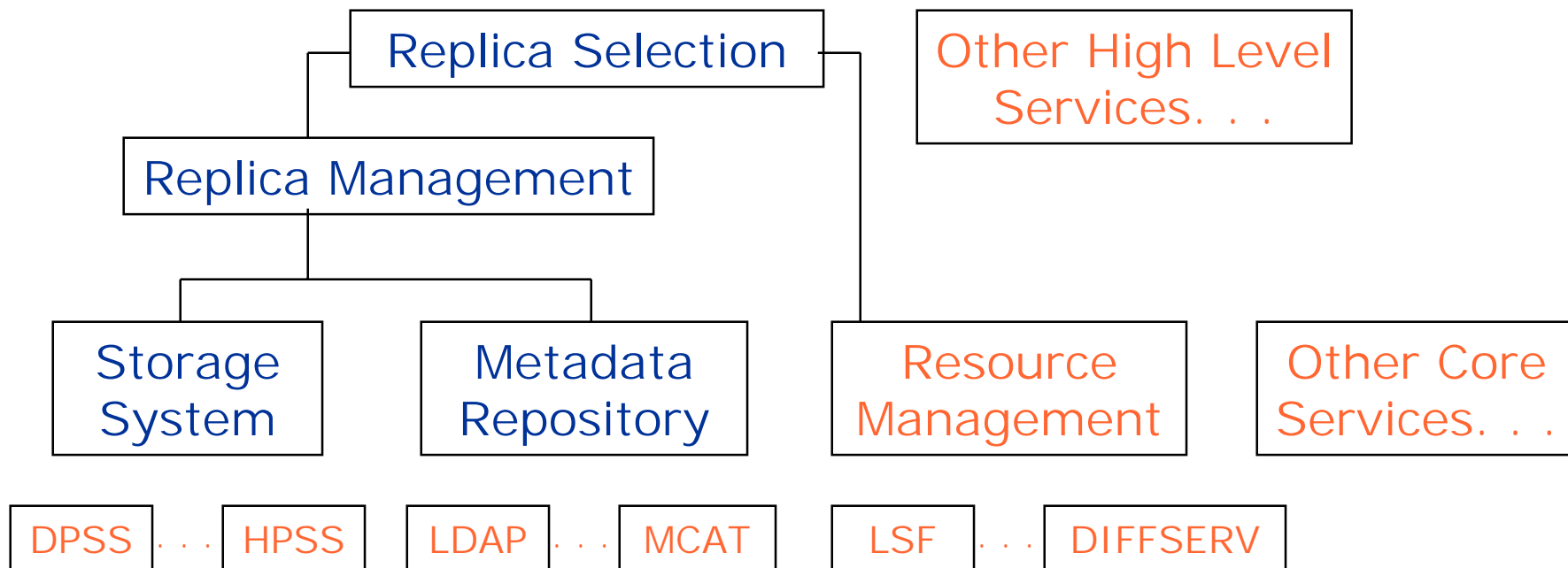


# Design Principles

- Mechanism Neutrality
  - ◆ Support heterogeneous systems
- Policy Neutrality
  - ◆ User / local decision making and control
- Compatibility with Computational Grid
  - ◆ Integration of storage and computation
- Uniformity of Information Infrastructure
  - ◆ Data model and interface for metadata



# Data Grid Services





## Data Access Service

- Uniform access to heterogeneous systems
  - ◆ remote: *e.g.* DPSS, HTTP, FTP, HPSS
  - ◆ local: *e.g.* UNIX
- High performance data movement over WANs
  - ◆ Third party transfer
- Data extraction and filtering functions
- Access to data is subject to global and local policy constraints





# Metadata Access Service

- Uniform treatment for all metadata
  - ◆ Grid components
  - ◆ Application-related metadata
  - ◆ Storage system characteristics
  - ◆ Relationships between data items
- Uniform access to metadata
  - ◆ LDAP protocol
- Uniform storage structure
  - ◆ LDAP hierarchical structure for distribution, replication, referral services



# Replica Management

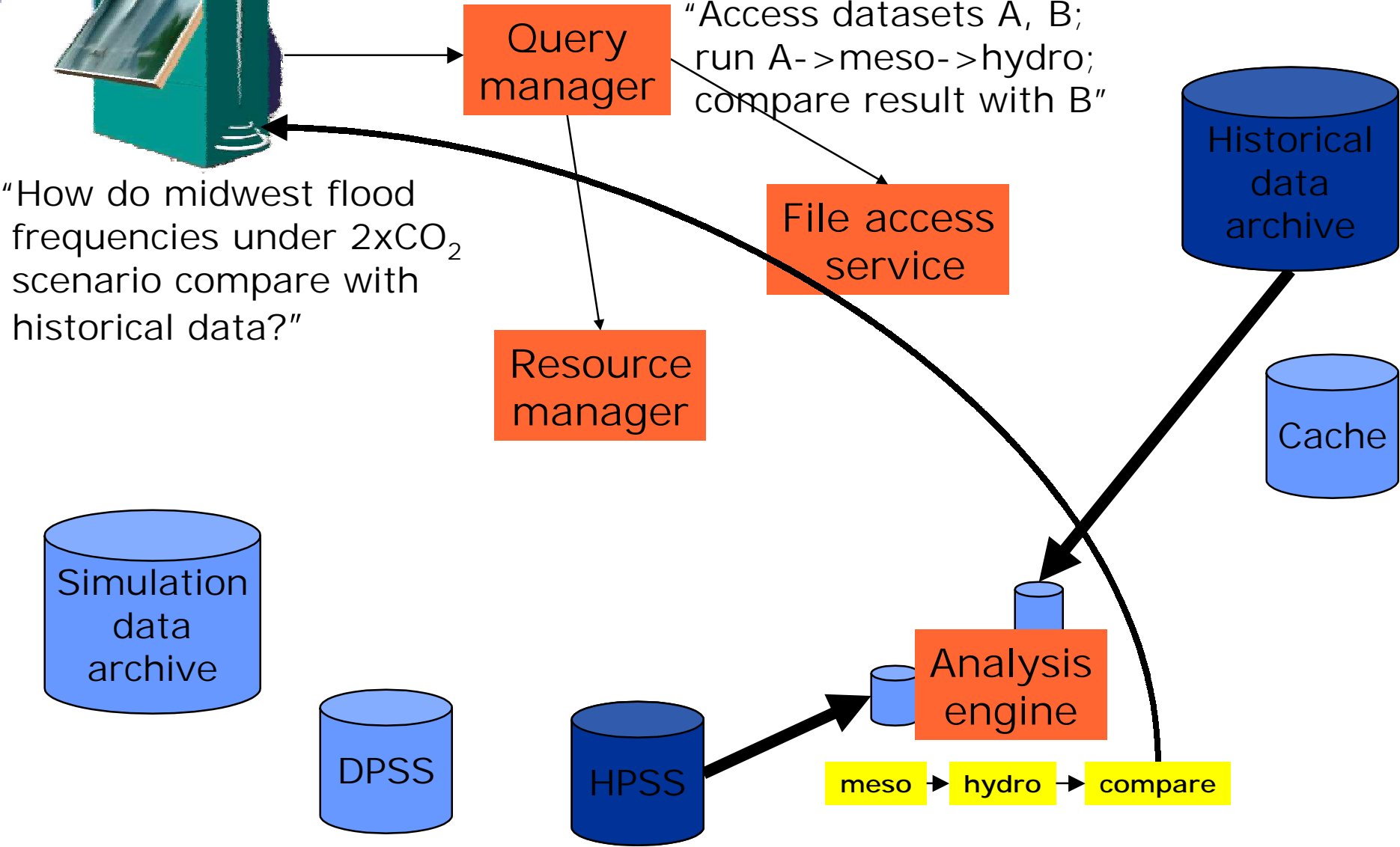
- Collections contain related files
- Logical files describe replicated physical files
- Services for managing replicated file instances
  - ◆ Create / delete
  - ◆ Schedule / manage data transfer
  - ◆ Register in the replica catalog
  - ◆ Metadata display



# Replica Selection

- User can optimize access characteristics
  - ◆ Grid structure and performance
  - ◆ Storage system and file characteristics
- Intelligent scheduling to determine appropriate replica, site for (re)computation, etc.

# Climate Data Scenario





# Current Activity

- Ongoing collaborations
  - ◆ Climate
  - ◆ High Energy Physics
- Storage API for uniform access to data
  - ◆ API specification document
  - ◆ Prototype code for HTTP, FTP, DPSS
- Replica management
  - ◆ Replica catalog based on LDAP
  - ◆ API and GUI tools for catalog access
- Quality of Service implementation



# Replica Management

The screenshot displays the Data Grid Browser interface. On the left, a file tree shows a hierarchy starting with 'o=globus, c=us' and 'gn=Replica Catalog'. A table in the center lists storage systems with columns for Host, Type, Performance, Bandwidth, and Latency. A 'Create Replica' dialog box is open in the foreground, showing a similar table and a 'Storage Systems Tree' on the left. The dialog also includes a 'Destination Storage System' section with fields for URL, Type, Performance, Host, and Description, and an 'Instance Info' section with a 'Destination url' field.

Host	Type	Performance	Bandwidth	Latency
hps5.lbl.gov	HPSS	56 megabytes...	N/A	N/A
crater.isi.edu	DPSS	30 megabytes...	59.511452	1.091323
modi4.ncsa.uiuc.edu	DPSS	100 megabyte...	8.852283	44.922228

Host	Type	Performance	Bandwidth	Latency
hps5.lbl.gov	HPSS	56 megabytes/second	N/A	N/A
modi4.ncsa.uiuc.edu	DPSS	100 megabytes/se...	7.467607	40.715717
crater.isi.edu	DPSS	30 megabytes/second	59.522268	0.526674

**Storage Systems Tree**

- o=The University of Texas at Austin
- o=Columbia University
- o=Argonne National Laboratory
  - ou=Mathematics and Computer Science
    - adsm
- o=University of Lecce
- o=Systran
- o=The Royal Institute of Technology
- o=Konrad-Zuse-Zentrum für Informationst...

**Destination Storage System**

**Storage System**

URL: N/A  
Type: HPSS  
Performance: 80 megabytes/second  
Host: lemon.mcs.anl.gov  
Description: Maximum Strategy disk array

**Instance Info**

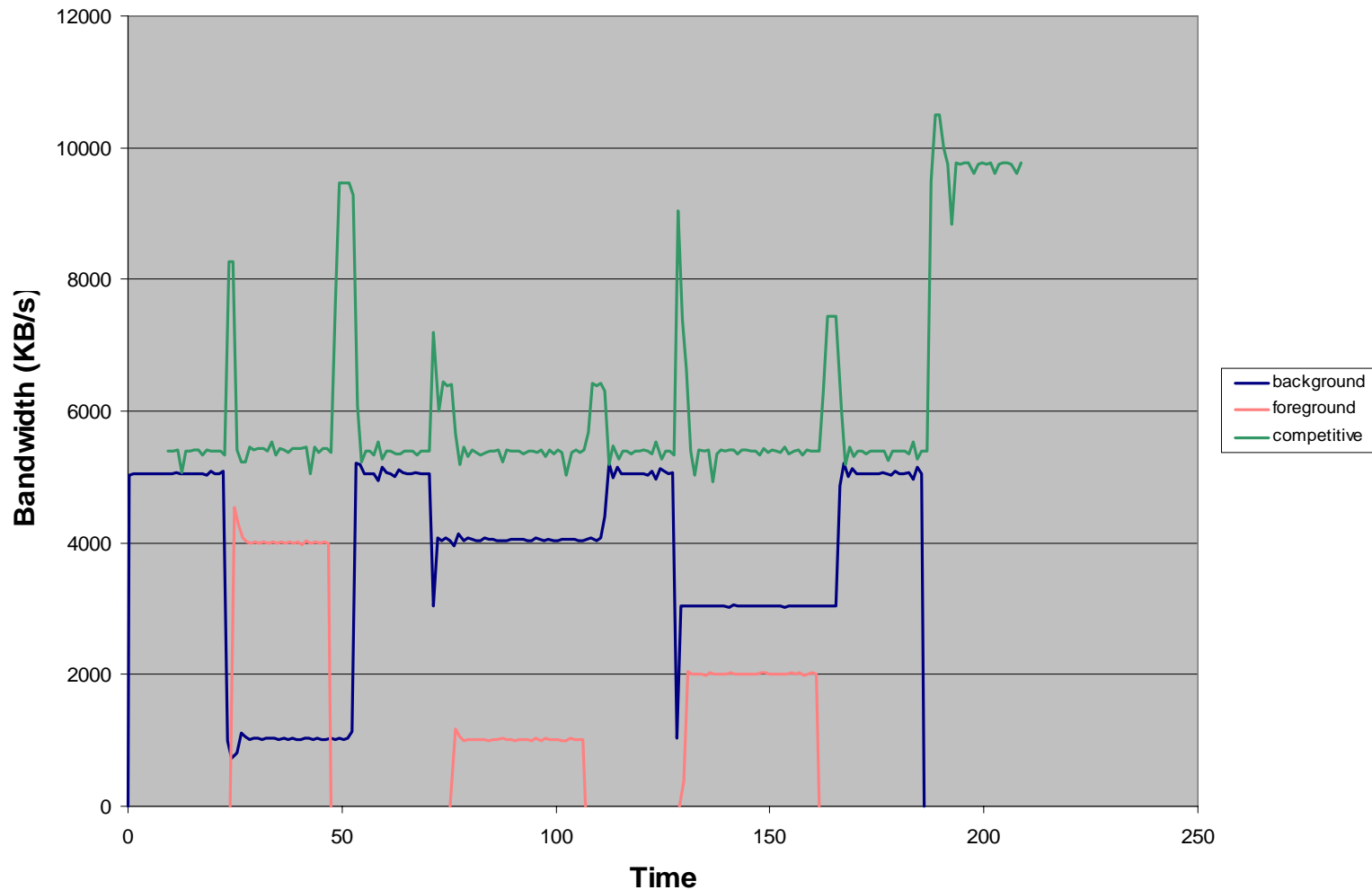
Destination url:

Create Cancel



# Quality of Service

## Bulk Transfer support in GARA





# Planned Activity

- Data Access
  - ◆ Integrated quality of service, security
  - ◆ Performance enhancements for networking
- Performance guarantees for the Data Grid
- Automatic operation of the Data Grid
  - ◆ Agent technologies used for distributed data replication, selection, and analysis
- Integrated CPU scheduling
  - ◆ Server-side data reduction, affinity scheduling



