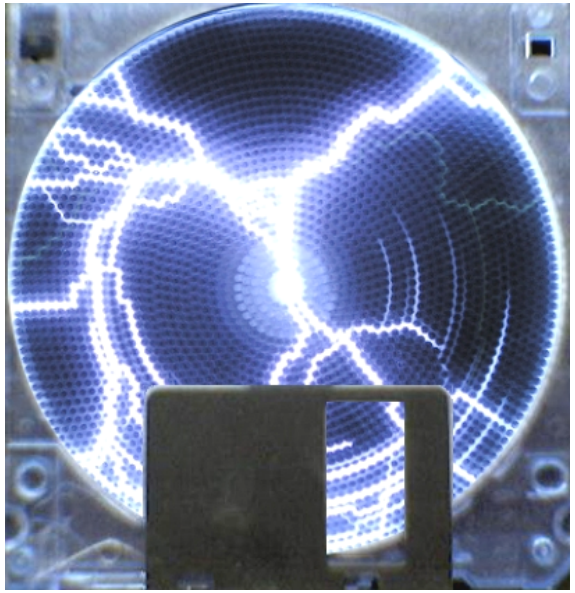
Indiana University's Lustre WAN: The TeraGrid and Beyond



Stephen C. Simms
Manager, Data Capacitor Project
TeraGrid Site Lead, Indiana University
ssimms@indiana.edu

Lustre User Group Meeting
April 17, 2009
Sausalito, California

The Data Capacitor Project

NSF Funded in 2005

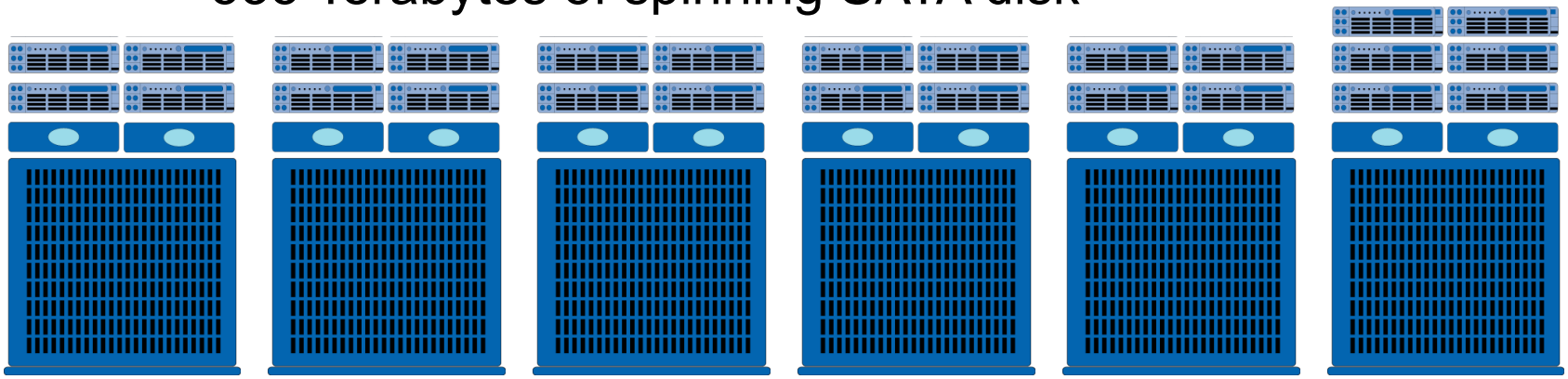
535 Terabytes Lustre storage

14.5 GB/s aggregate write

Short term storage

What's A Data Capacitor?

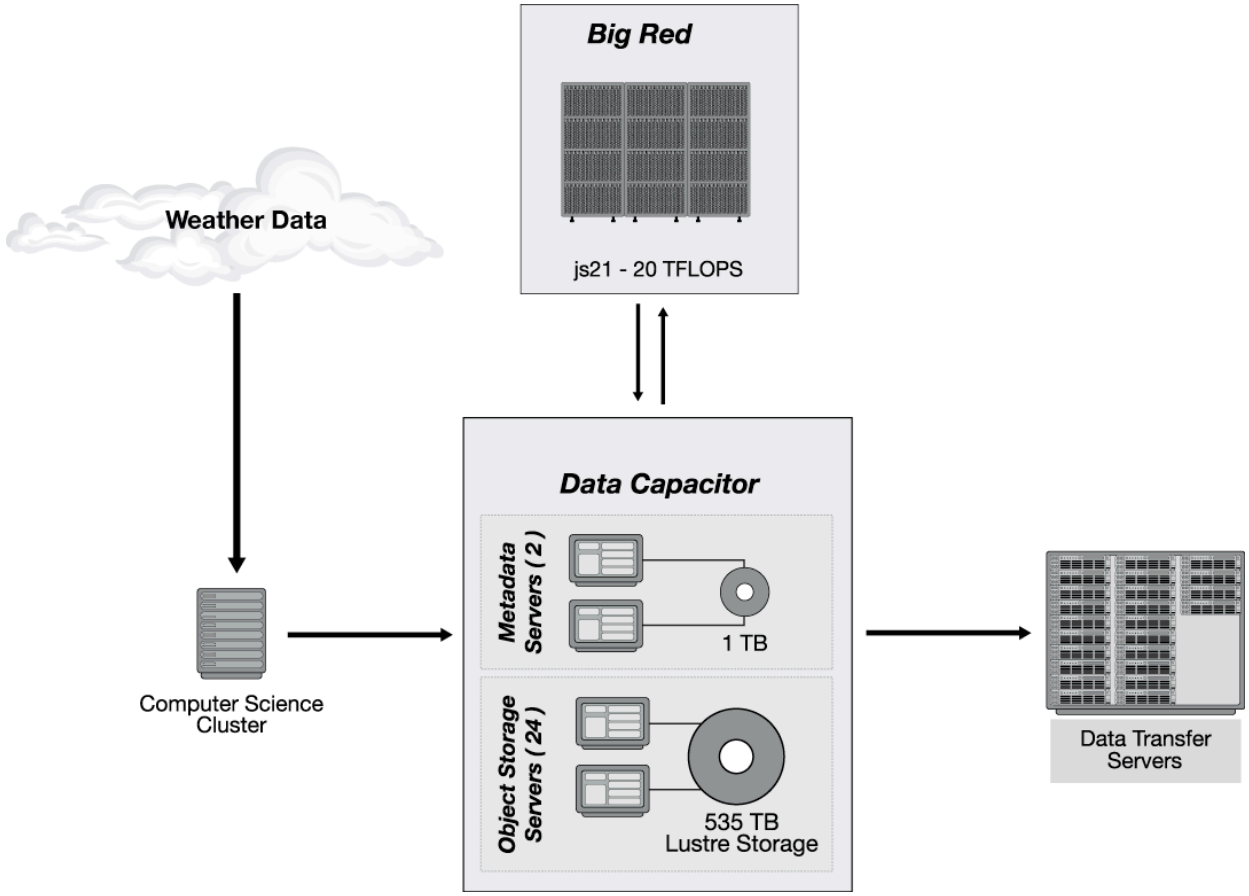
- 12 pairs Dell PowerEdge 2950
  - 2 x 3.0 GHz Dual Core Xeon
  - Myrinet 10G Ethernet
  - Dual port Qlogic 2432 HBA (4 x FC)
  - 2.6 Kernel (RHEL 4)
- 6 DDN S2A9550 Controllers
  - Over 2.4 GB/sec measured throughput each
  - 535 Terabytes of spinning SATA disk



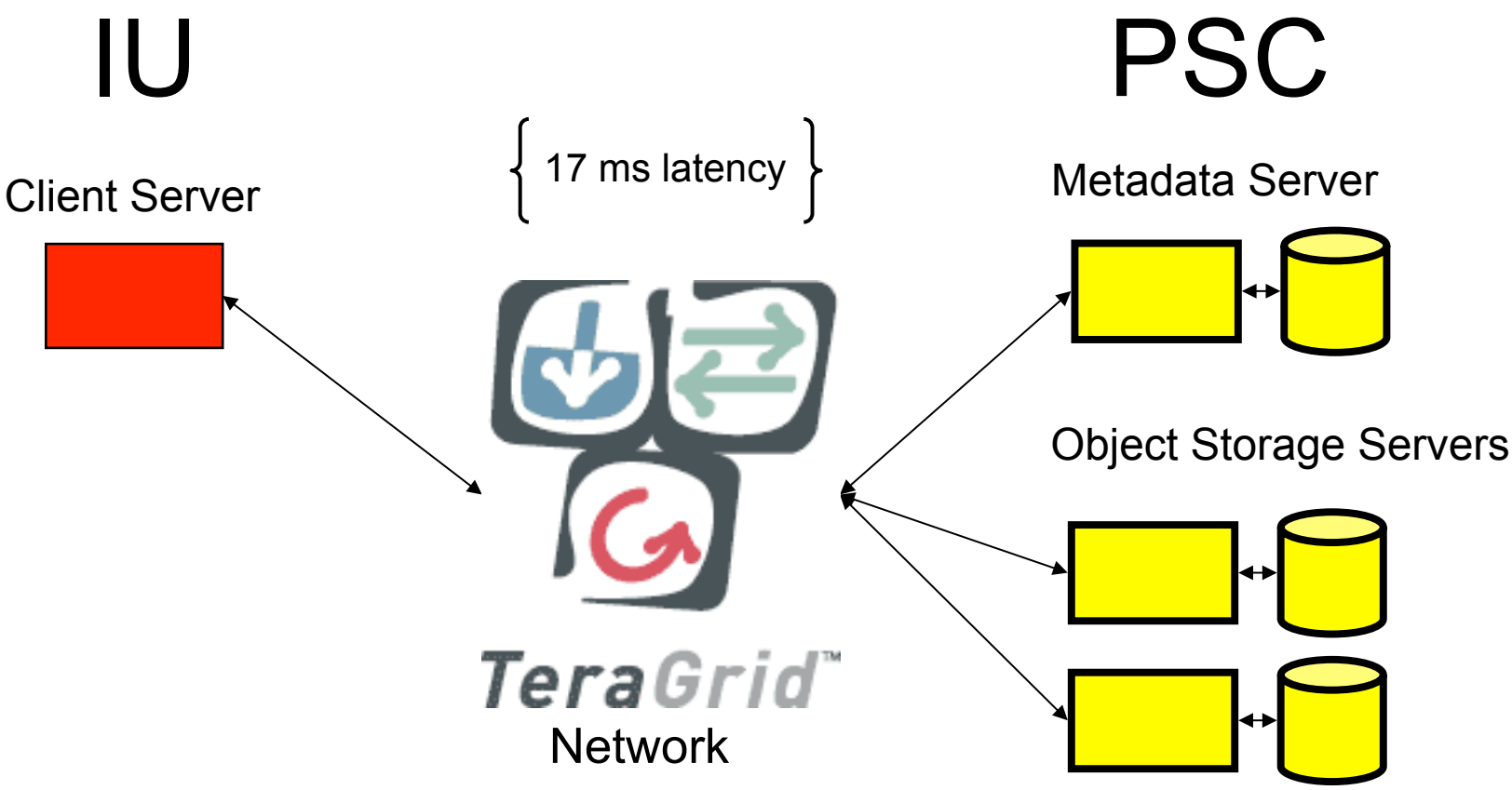
Maximize Utility Through One Stop Shopping

- Data Capacitor
  - Sits at the center of IU's cyberinfrastructure
  - Opens the door to new workflow opportunities
  - Provides the user with:
    - Fast paths to archive storage
    - A shared filesystem for
      - Computation
      - Visualization
      - Instruments
      - Potentially desktops

LEAD Workflow Example



Lustre Across the WAN



10 GigE Single Client Tests

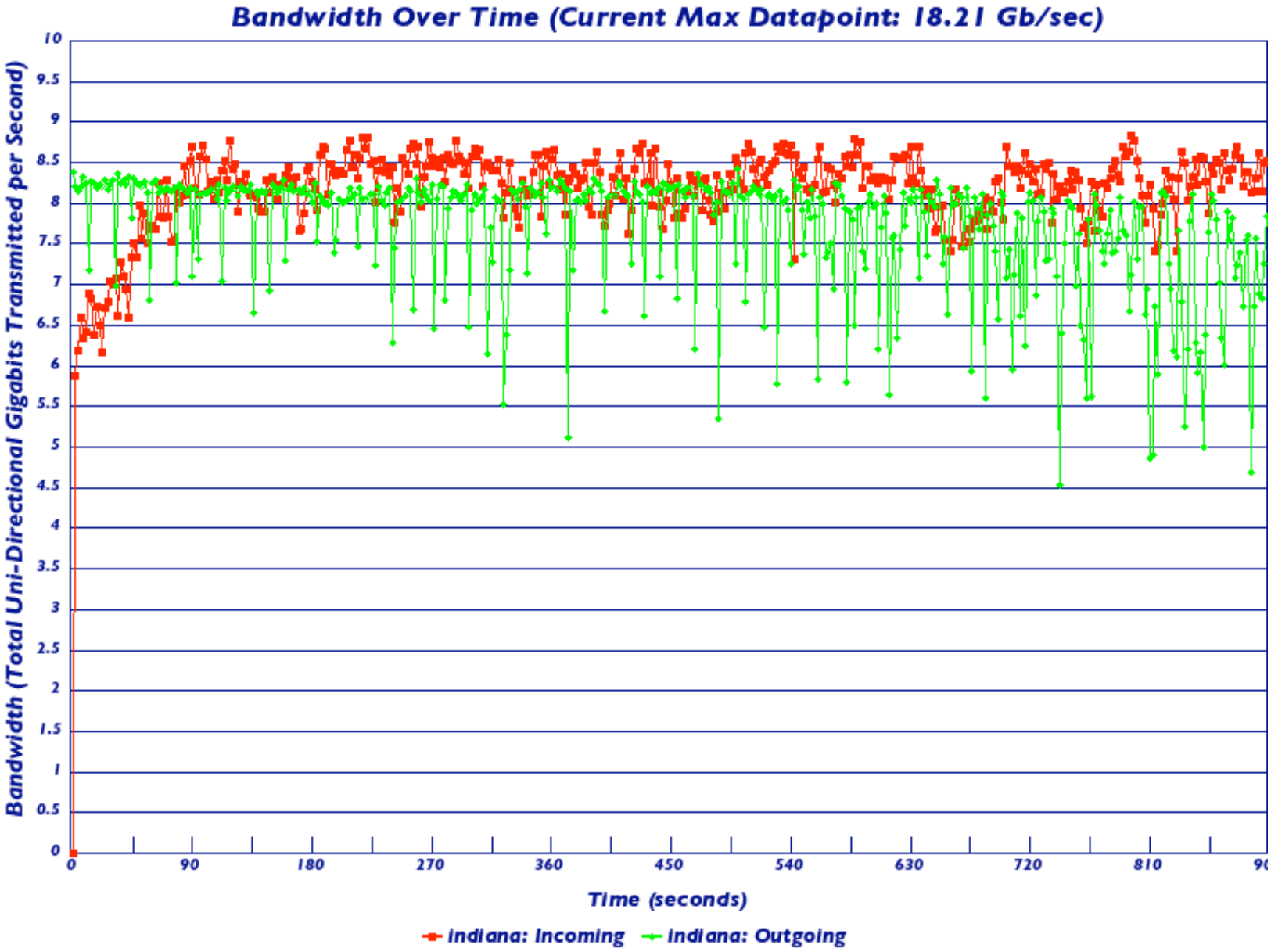
977 MB/s between ORNL and IU
Using 10Gb TeraGrid connection
Identical Dell 2950 clients
- 2x dual 3.0 GHz Xeon
- 4GB RAM
- 1 Myricom Myri10G card
  - Ethernet mode
- Lustre 1.4.7.1

16 of 24 Data Capacitor (DC) OSSs

2007 Bandwidth Challenge: Five Applications Simultaneously

- Acquisition and Visualization
  - Live Instrument Data
    - Chemistry
  - Rare Archival Material
    - Humanities
- Acquisition, Analysis, and Visualization
  - Trace Data
    - Computer Science
  - Simulation Data
    - Life Science
    - High Energy Physics

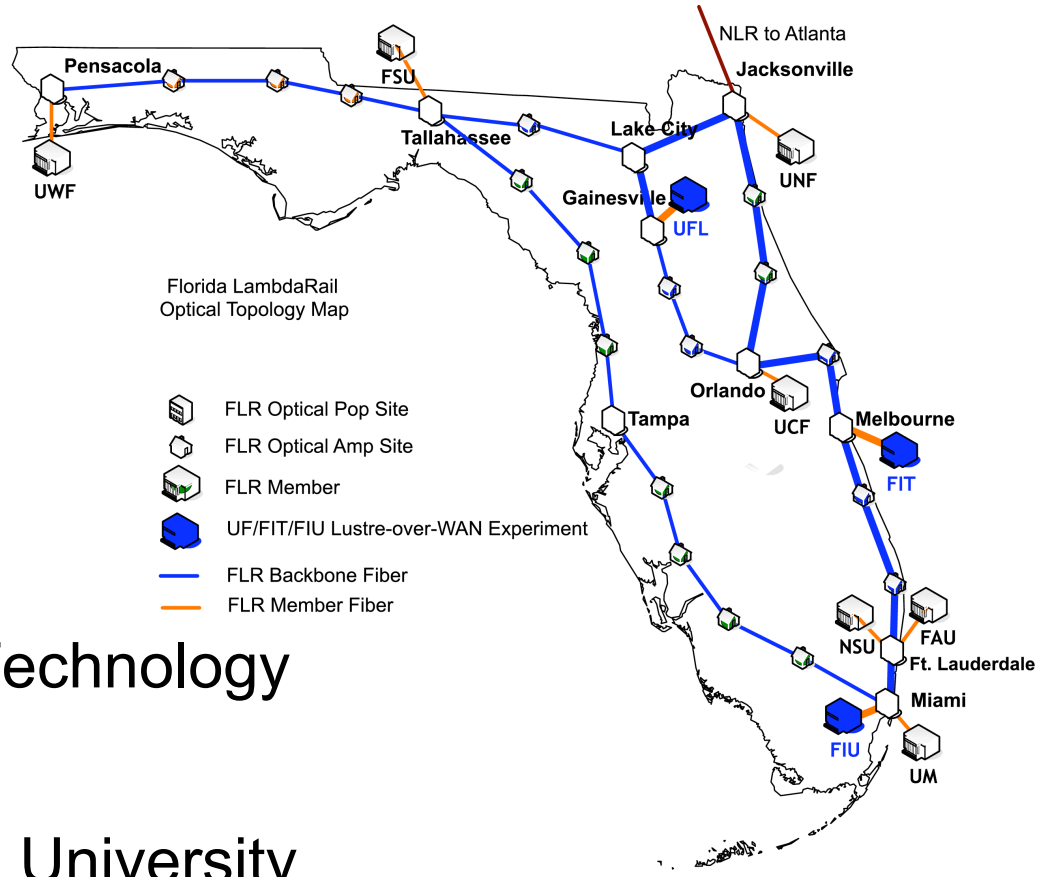
Challenge Results



Beyond a Demo

Florida Lambda Rail

- In 2008
  - UFL
    - University of Florida
  - FIT
    - Florida Institute of Technology
  - FIU
    - Florida International University



IU UID Mapping

- Lightweight
  - Not everyone needs / wants kerberos
  - Not everyone needs / wants encryption
- Only change MDS code
  - Want to maximize clients we can serve
- Simple enough to port forward

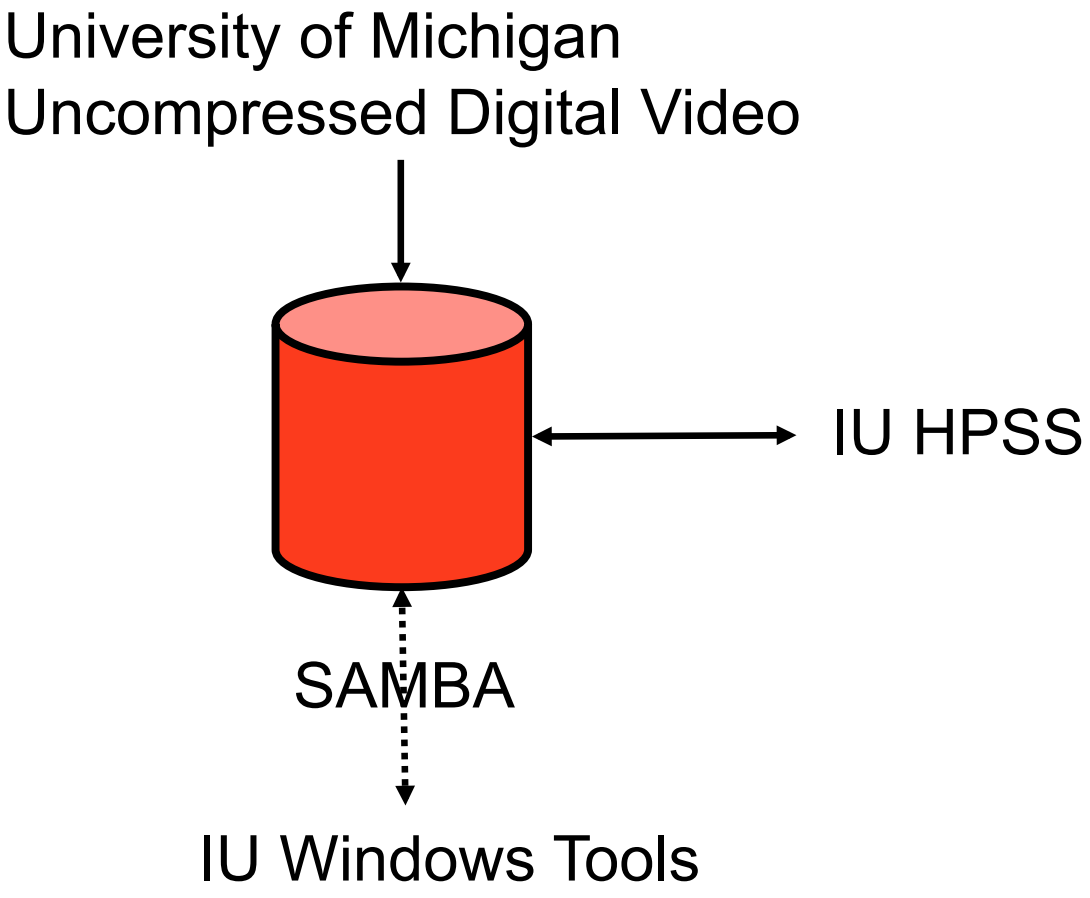
UID Mapping

- Lookup calls pluggable kernel module
  - Binary tree stored in memory
  - Based on NID or NID range
  - Remote UID mapped to Effective UID
- Other lookup schemes are possible

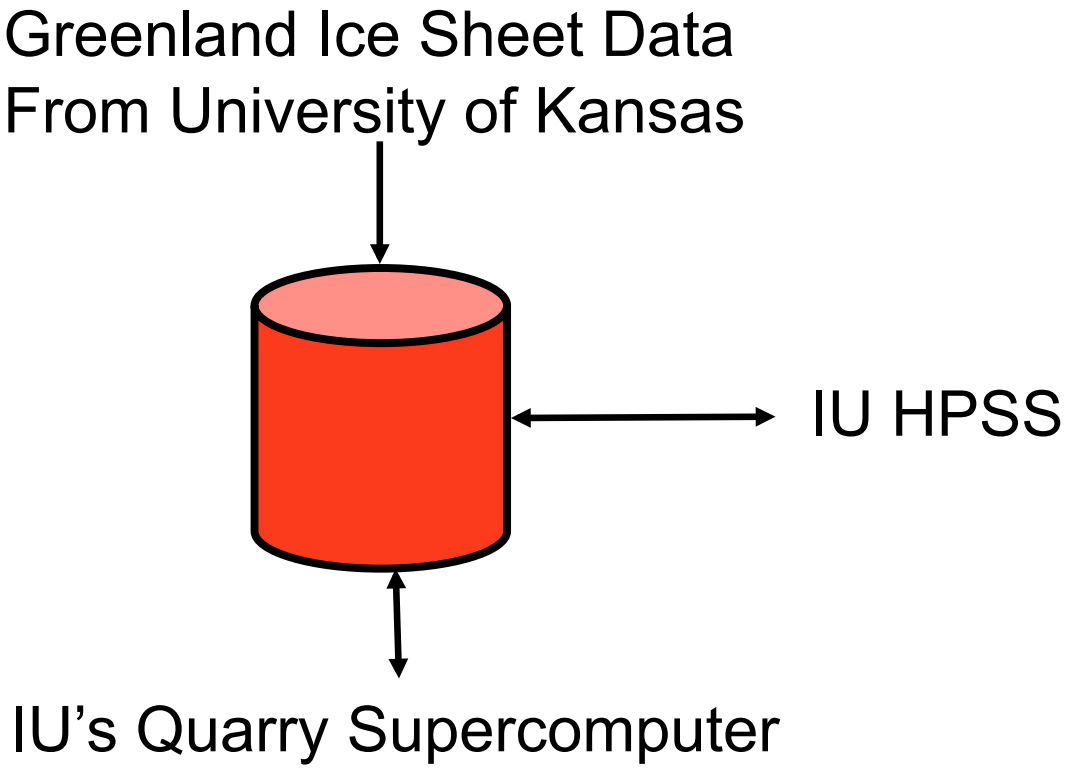
Announced One Year Ago

- IU Lustre WAN Service
- Dedicated system of 360 TB for research
- Serve on a per project basis
- Single DDN S2A9950 Controller
- 6 Dell 2950 Servers
  - 4 Object Storage Servers
  - 1 pair of Metadata Servers for failover

EVIA – University of Michigan



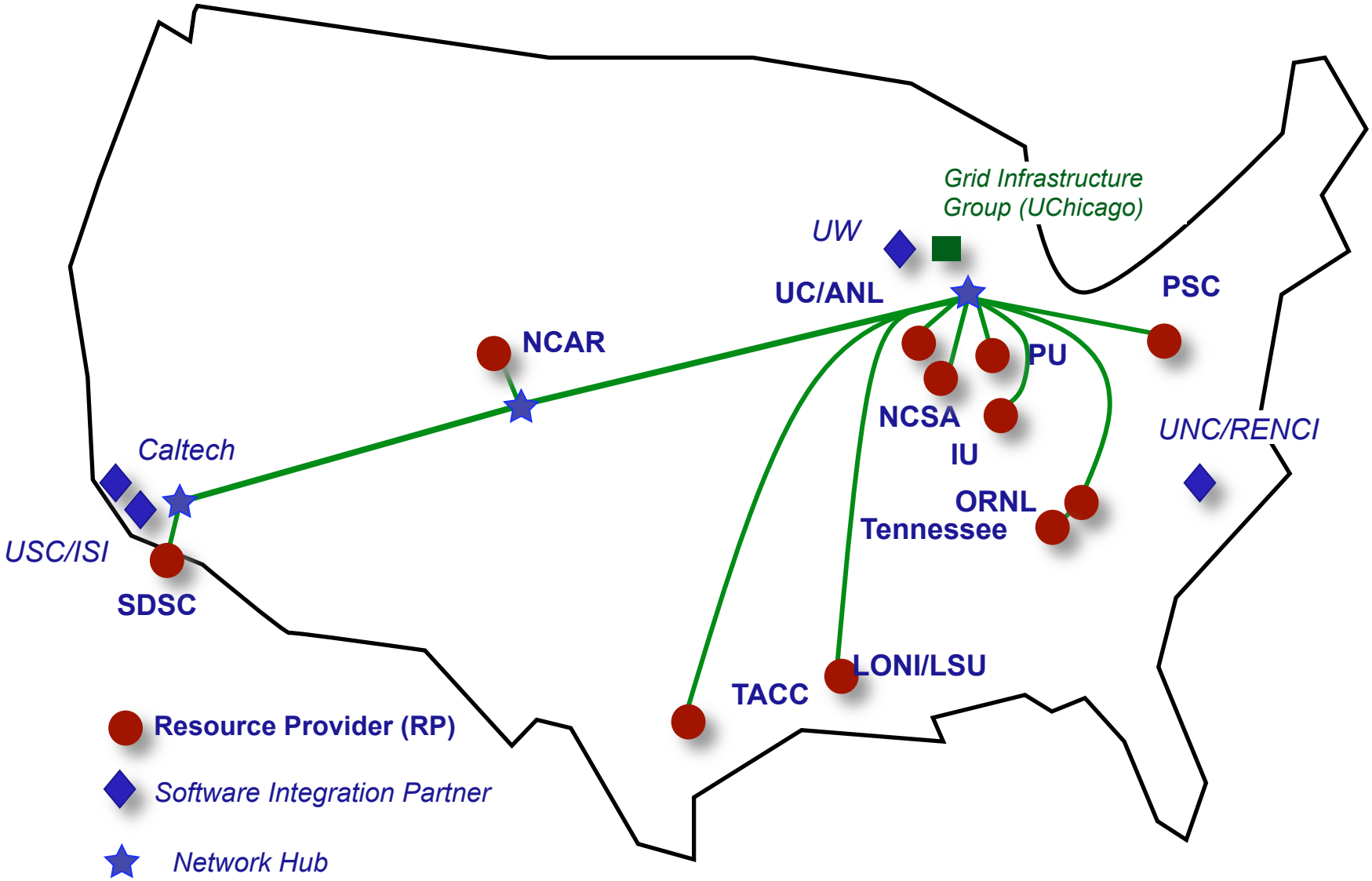
CReSIS – University of Kansas



The TeraGrid

- TeraGrid is an open scientific discovery infrastructure combining leadership class resources at eleven partner sites to create an integrated, persistent computational resource.
- ANL, IU, Louisiana Optical Network Initiative (LONI), NCSA, NICS, ORNL, PSC, Purdue, SDSC,TACC, NCAR

The TeraGrid Map



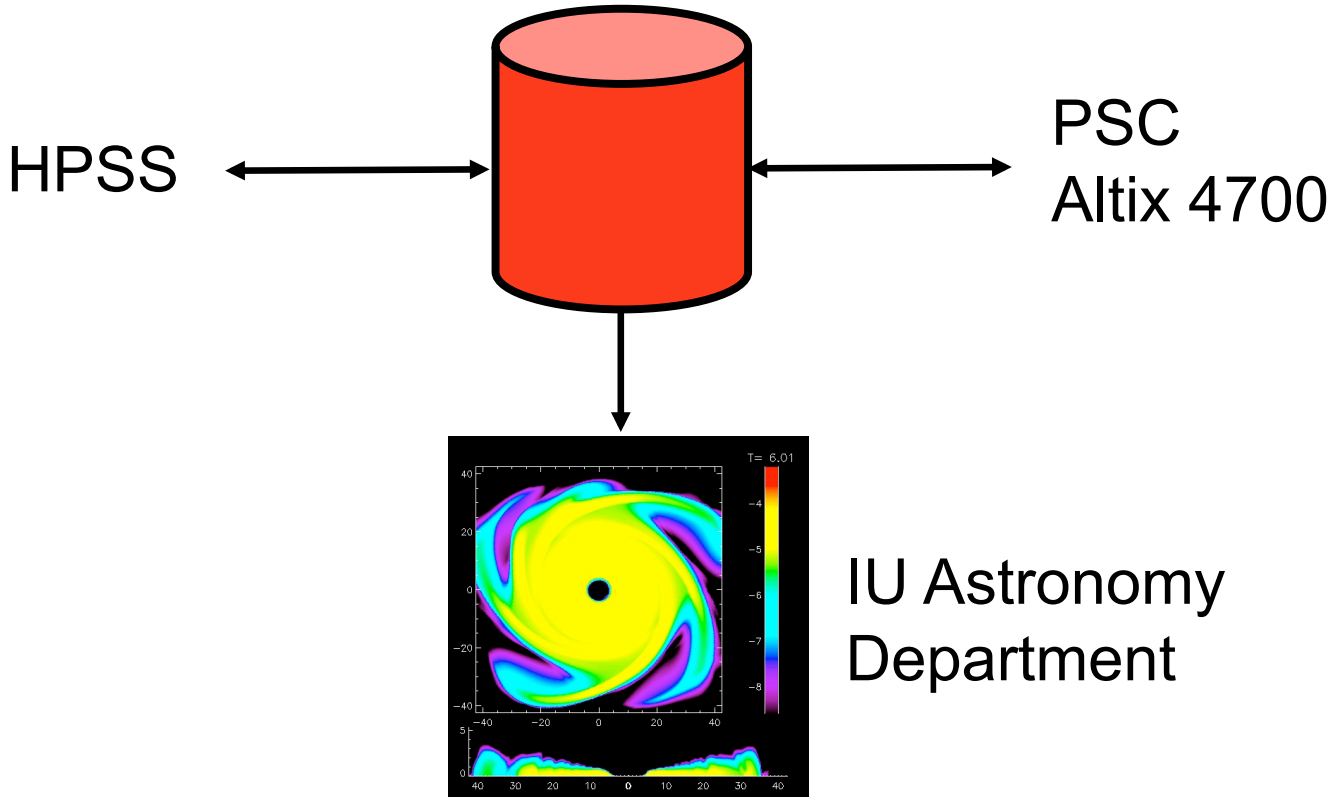
Lustre on the TeraGrid

- IU
- LONI
- NCSA
- NICS
- PSC
- TACC

IU's Lustre WAN on TeraGrid

- Production mount at PSC
  - Altix 4700 and Login Nodes
  - GridFTP and Gatekeeper nodes
- Test mounts at LSU, TACC, SDSC
  - Login nodes
  - LSU and SDSC ready to deploy on some compute nodes
- Future plans for NCSA, ORNL, and Purdue

3D Hydrodynamic Workflow

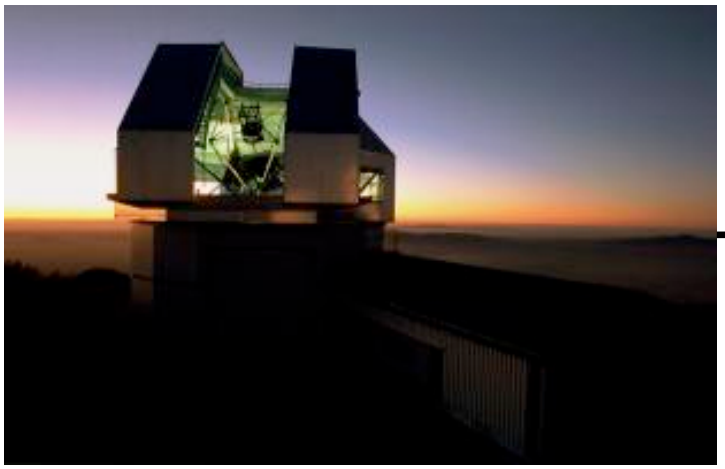


J-WAN

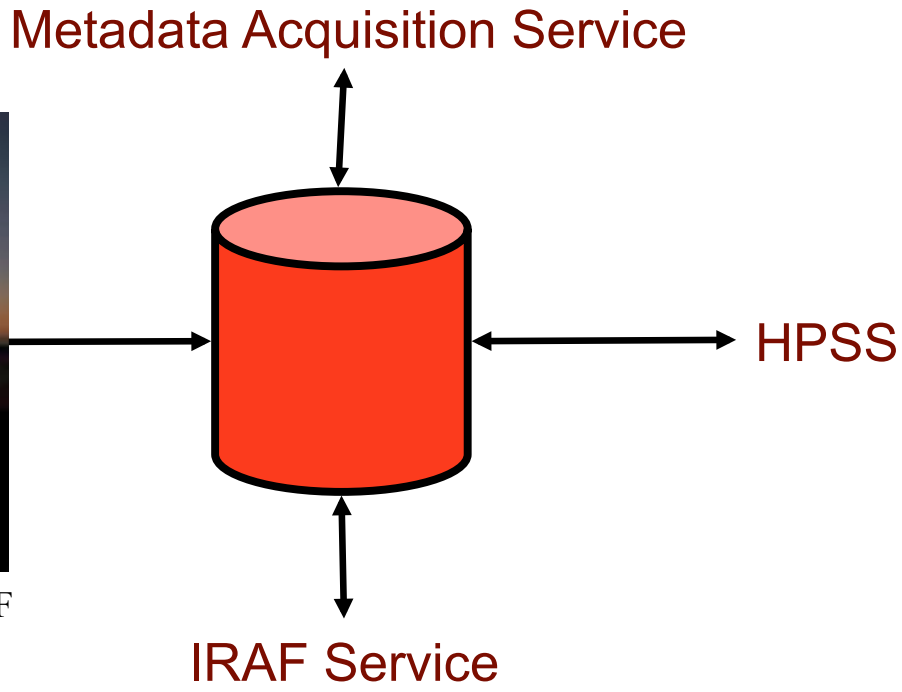
- Josephine Palencia at PSC
- Lustre Advanced Features Testing
  - Distributed OSTs
  - Kerberos
  - Eventually CMD
- Successful mount and testing with SDSC

Beyond

WIYN Telescope at Kitt Peak



NOAO/AURA/NSF



Many Thanks

- Josh Walgenbach, Justin Miller, Nathan Heald, James McGookey,, Scott Michael, Matt Link (IU)
- Kit Westneat (DDN)
- Sun/CFS support and engineering
- Michael Kluge, Guido Juckeland , Robert Henschel, Matthias Mueller (ZIH, Dresden)
- Thorbjorn Axellson (CReSIS)
- Craig Prescott (UFL)
- Ariel Martinez (LONI)
- Greg Pike and ORNL
- Doug Balog, Josephine Palencia, and PSC

Support for this work provided by the National Science Foundation is gratefully acknowledged and appreciated (CNS-0521433)

Thank you!

Questions?

ssimms@indiana.edu

dc-team-l@indiana.edu

http://datacapacitor.iu.edu