

# Planning a Globally Accessible Archive of MODIS Data

*Micah Beck*

*Director & Associate Professor*

Logistical Computing and Internetworking Lab  
Computer Science Department  
University of Tennessee

Digital Archives PI Mtg

5/17/2005



**LOCI**

LOGISTICAL COMPUTING AND  
INTERNETWORKING LAB



UNIVERSITY OF TENNESSEE

# Planning Participants & Collaborators

- » Terry Moore, Scott Atchley, Murray Browne  
Computer Science, Univ. of Tennessee
- » Suzie Allard, Peiling Wang  
Information Science, Univ. of Tennessee
- » John Dwyer  
USGS, EROS Data Center
- » Bill Hargrove, Bob Cook  
DOE, Oak Ridge National Laboratory
- » P.R. Blackwell  
Stephen F. Austin State Univ. (AmericaView)

## Challenge: A Massive, Accessible Archive

- » Data from earth observing satellite instruments has uniquely high continuing research value
  - Moderate Resolution Imaging Spectroradiometer (MODIS)
- » The flow is continuous and massive
  - MODIS data is 1TB/day
- » ... must be processed and archived
  - The responsibility of the NASA-USGS Land Processes Distributed Active Archive Center (LP DAAC) at the EROS Data Center
- » ... must be accessible to the global community of research scientists and policy planners.

# MODIS Archive Access Platform (MAAP)

» **MAAP** is a

- distributed service platform
- based on *Logistical Networking*
- for providing a global community with fast, flexible, and efficient access
- to all MODIS archive data

# Logistical Networking in the Design of the MAAP

- » Apply *Logistical Networking (LN)* Technology to the problem of accessing and distributing data to communities on a global scale
- » Storage resources are provisioned for shared use within distributed communities.
- » Issues relating to the placement and movement of data are factored out from higher level concerns
  - Definition of metadata and directory structures
  - Resource allocation between competing users
- » Build a common platform based on functionality that is sufficient to support future development

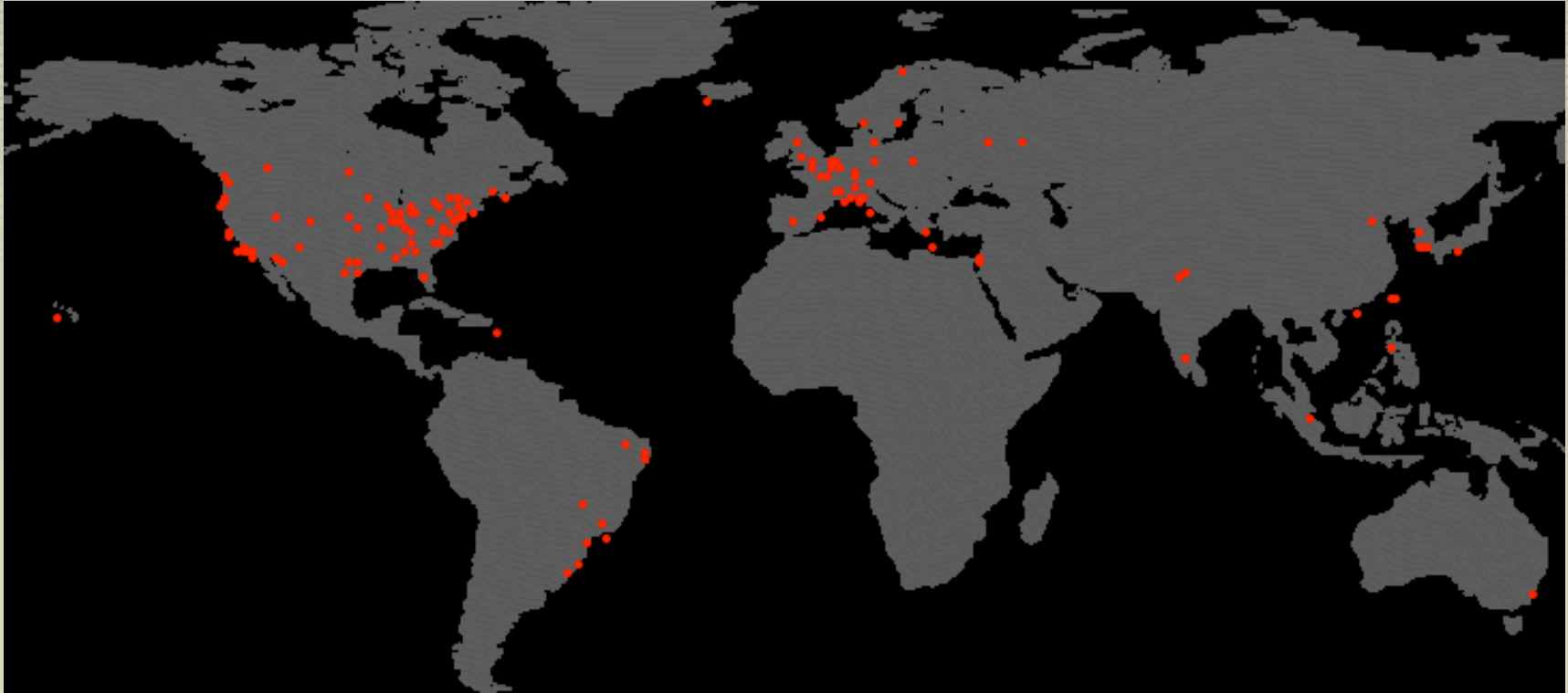
# What is Logistical Networking?

- » A common, generic buffer service deployed as part of the shared network fabric
- » A best effort service that takes advantage of the continuing exponential increase in storage density
- » Exposed replication and geographical RAID
- » Content distribution, prestaging and caching
- » A lightweight way to manage transient storage
- » A means to achieve very high data movement performance through control over buffering
- » A means to connect switched optical (>1Gbps) and routed IP (<1Gbps) networks

# Areas of Current LN Application

- » Physics
  - Fusion Energy Simulation
  - Astrophysics Simulation
  - High Energy Particle Physics
- » Genomic Databases
  - Caching
  - Data Mining
- » Grid Computing
  - Flexible data movement
- » Scientific Visualization
  - Caching/prestaging
  - Computation as a Network Service
- » Multimedia
  - High Definition Video on Demand
  - Video Postprocessing & Transcoding
  - TV-over-IP
- » Content Distribution
  - High demand software

# Current Infrastructure Deployment



The public deployment includes 500+ IBP depots in 30 countries serving 30+ TB storage (leverages PlanetLab). Private deployments for DOE, Brazilian and Czech backbones.





## Why a Planning Grant for MAAP?

- » Responsibility for the archive lies with the DAACs
  - Tape is well understood as an archiving technology; network storage is not
- » Any proposal for an accessibility platform must connect with the operator and user communities
- » An operational archive would support the missions of non-NSF agencies (eg NASA, NOAA, USGS, DOE)
  - Funding vehicle must be found for full project
- » More complete service must be defined through collaboration with Information Science
- » Need for discussions, outreach, technology demo

# First Steps in the Planning Process

- » Engaging with collaborative teams
  - NASA-USGS LP DAAC at EROS
  - UT Information Science colleagues
- » Establish Web outreach tools
  - Project Web site
  - Distribution of high demand MODIS products
- » Begin creation of outputs
  - MAAP planning document
  - Presentation on LN for MODIS community
- » Important Meetings
  - Earth Science Data System Working Group
  - American Geographic Union (Science Teams)

**Thank you!**

**mbeck@cs.utk.edu**

**To learn more...**

**<http://loci.cs.utk.edu>**

