

Building a Massive, Distributed Storage Infrastructure at Indiana University

Anurag Shankar & Gerry Bernbom

University Information Technology Services
Indiana University

<http://storage.iu.edu/presentations/ieee02.ppt>

Contributors: Past and Present

- Dave Vernon
- Gustav Meglicki
- Leigh Grundhoefer
- Jeff Russ
- Chris Hemmerich
- Haichuan Yang
- Chris Garrison

Recentralization of Storage

- Personal, desktop computer hard disks at tens of GB. Why recentralize?
- Answer: Problems with long term data integrity (lack of backups, unreliable hardware, etc.)

Why Recentralize?

“An organization’s computer data is its intellectual property.”

Our vision: To implement a bottomless, deep store with ubiquitous, native, and secure access to data, excellent file sharing mechanisms, interoperability with current and future IT services, and a long term view of IU’s storage needs.

Candidates for a Massive Data Storage System at IU

- Research data - astronomy, physics, chemistry, geology, genomics, etc.
- Administrative data – student records, scanned documents, HR data, etc.
- Digital libraries – large digitized collections of images, audio, video, scanned books, etc.
- General user data – email, personal work files, etc.

Infrastructure Choices

- Tape to disk ratio at the high-end favors tape over disk. Ergo the largest data repositories today are built with hierarchical storage management (HSM) software
- Data migrates from fast disk caches to massive tape libraries seamlessly.
- Retrievals from tape incur a 30-90 second penalty per file if data is not found in the disk cache.

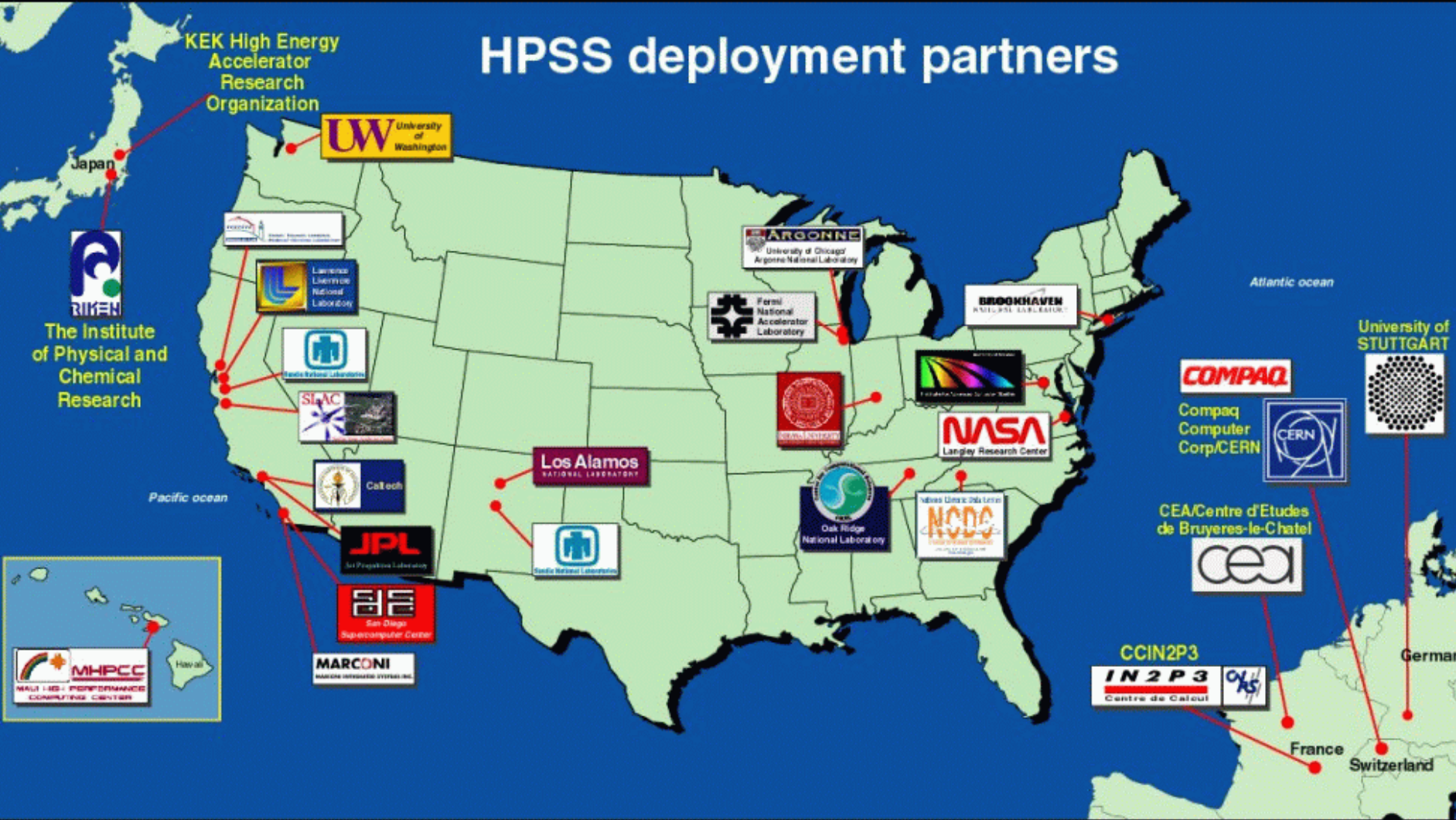
Brief History

- IU started looking at storage trends in 1998 and went through a RFP process in 1999 to build a massive data storage system.
 - Main HSM software choices included HPSS (High Performance Storage System), LSC's SamFS, and UniTree.
 - HPSS chosen since it is the most likely technology to survive long term (it's a collaboratively developed, not a vendor supplied solution). Collaboration consists of Los Alamos/ Livermore/Sandia/Oak Ridge labs, NASA, and IBM Global Services.

Brief History...

- IU joined the HPSS collaboration in early 1999. IBM chosen as the hardware vendor. Base system in production in June 1999.
- The migrating-DFS front-end to HPSS in production in August 1999.
- IUPUI remote mover in production in October 2000.
- Static-DFS based Common File System augments HPSS in October 2000.
- Gatewayed CFS for the masses in Sept. 2001.

HPSS deployment partners



The HPSS collaboration

HPSS software development partners

IBM Global Services - Federal
Lawrence Berkeley National Laboratory
Lawrence Livermore National Laboratory
Los Alamos National Laboratory
NASA Langley Research Center
Oak Ridge National Laboratory
Sandia National Laboratories

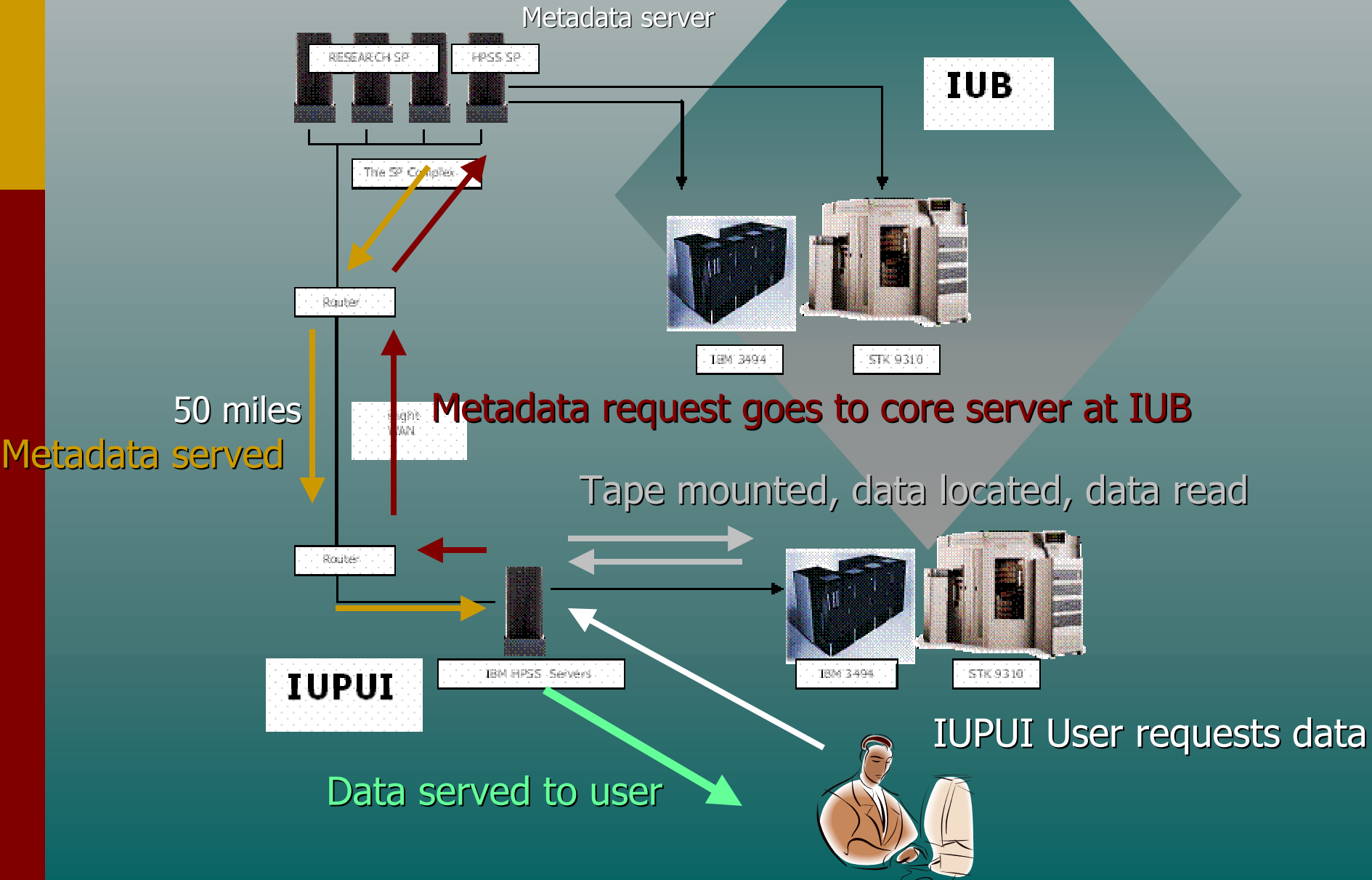
HPSS deployment partners

Brookhaven National Laboratory
Caltech and Jet Propulsion Laboratory
CCIN2P3-Centre de Calcul de l'Institut National de Physique Nucleaire at de Physique des Particules
CEA-Centre d'Etudes de Bruyeres-le-Chatel
Compaq Computer Corp/CERN
Fermi National Accelerator Laboratory
Indiana University
Lawrence Berkeley National Laboratory
Lawrence Livermore National Laboratory
Los Alamos National Laboratory
Maui High Performance Computer Center
Marconi Integrated Systems
NASA Langley Research Center
NOAA National Climatic Data Center
Oak Ridge National Laboratory
Sandia National Laboratories
San Diego Supercomputer Center
Stanford Linear Accelerator Center
RIKEN-The Institute of Physical and Chemical Research
University of Chicago/ANL
University of Maryland
University of Washington
University of Stuttgart

Industry participants

Compaq
IBM
Kinesix
Objectivity
StorageTek
Sun Microsystems

Distributed HPSS between Bloomington and Indianapolis



Modes of Accessing HPSS at IU

- FTP
- Parallel FTP (from Unix only)
- DFS (via clients on Unix/Windows NT/2000 desktops)
- HSI (from Unix)
- WWW
- HPSS API

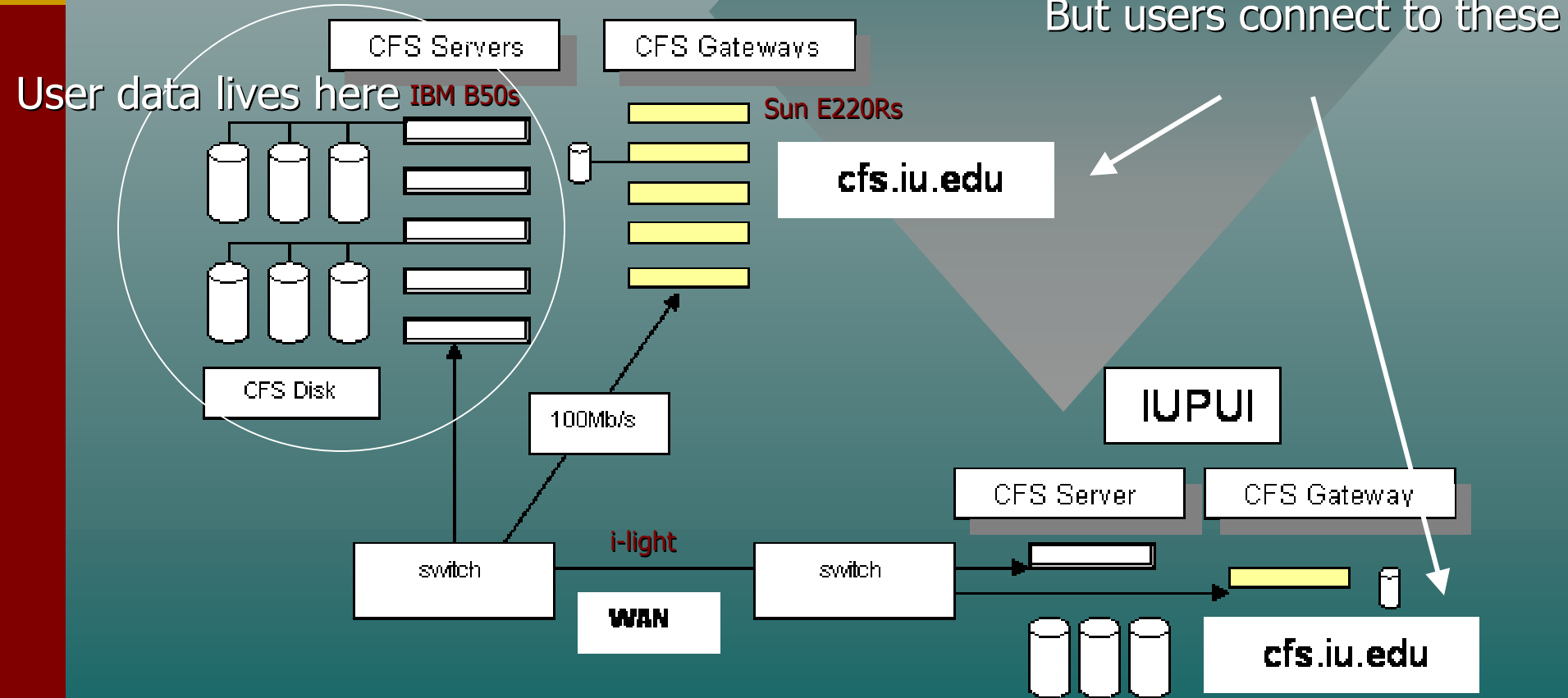
The Common File System: Disk Storage for the Masses

- HPSS provides the ability to store massive amounts of data, however it is biased toward large files (since tapes perform best when streaming).
- The CFS addresses the small file & frequent access issues, and provides conventional, disk-based storage for the masses.
- The CFS is based on DFS, providing a seamless extension to HPSS via the migrating DFS front-end.

Distributed, Gatewayed CFS Design

HIGHLY SCALABLE

But users connect to these



IU's Gatewayed CFS Service

- Accessible without a client as a mapped drive letter under Windows (9x, NT4, 2000, XP).
 - \\cfs.iu.edu\<username>
- Accessible without a client as an Appleshare IP share under MacOS (8, 9, X).
 - cfs.iu.edu
- Accessible via smbclient/smbfs from Unix/Linux.
- Secure web access from any browser.
 - <http://cfs.iu.edu/<username>>

For More Information

- Please come by the poster.
- <http://storage.iu.edu/>.