

# Digital Radiology on NGI WANs

---

Professor Robert Hollebeek

NSCP/NDMA - University of Pennsylvania

US-UK Workshop on Grid Computing

August 4, 2001

San Francisco

# Wide Area Data Infrastructure for storage/retrieval/use of medical records (focus on radiology)

- **Next Generation Internet / Internet 2**
- **National Institutes of Health and the National Library of Medicine**



# Components

---

- **Hospital Portal Systems with security**
- **Hierarchical Storage and Network layout :  
Area and Regional Archives**
- **Fast (OC48) networks and data caching  
and replication**
- **Standards for record encapsulation (XML)**
- **Standards for network transmission (qmp)**

# Three Applications

---

- **Archive Storage and retrieval for clinical use**
- **Training and Teaching for Radiology Departments**
- **Computer Assisted Diagnostics as a service**

# Focus: Digital Radiology

---

mammograms

X-rays

MRI

cat scans

endoscopies

...

## ● Hospital Digital Data

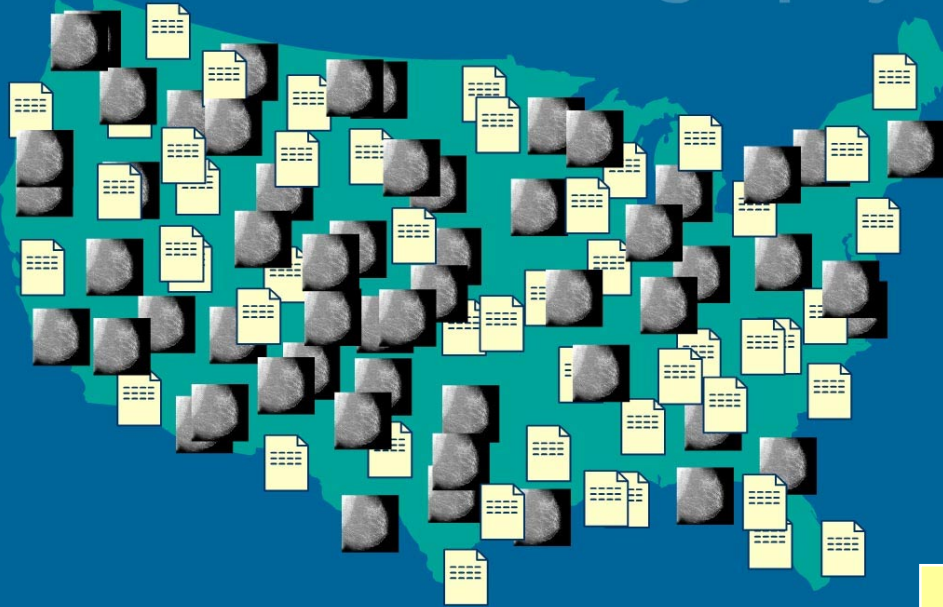
- Very large data sources - great clinical value to digital storage and manipulation and significant cost savings
- 7 Terabytes per hospital per year
- dominated by digital images

## ● Why we chose Mammography

- clinical need for film recall and computer analysis
- large volume ( 4,000 GB/year ) (57% of total)
- storage and records standards exist
- great clinical value to this application

# Managing Large Scale Data

## DIGITAL Mammography



Highly Distributed Source

## Hierarchical Storage and Indexing

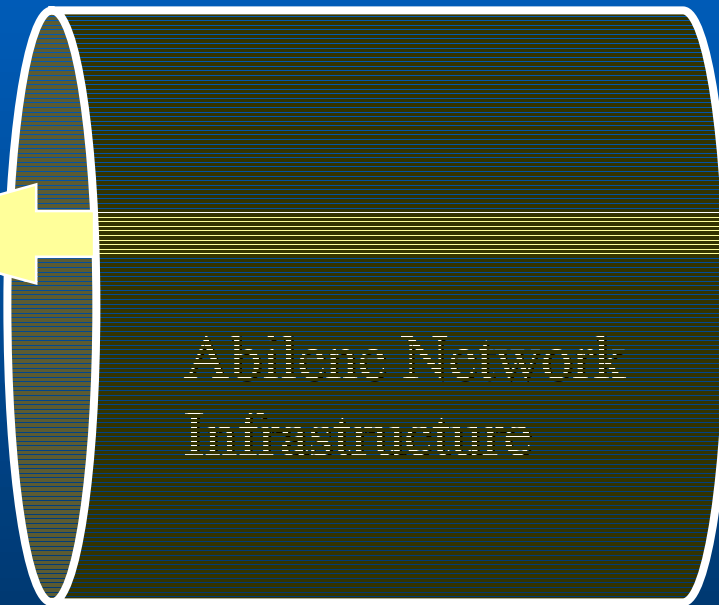


# Major Components

---



Hospital  
Portal  
Systems



**RadAR:  
Large Scale  
Storage**

**Index**

**Audit  
and Log**

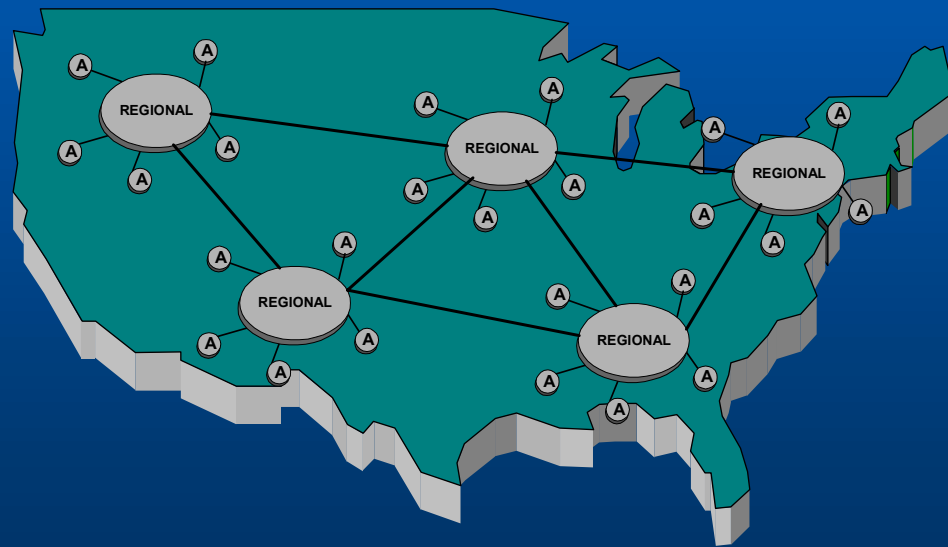
**Search  
and Retrieve**

**Mine**

# Proposed Hierarchical Layout

---

## Regional and Area Archives (A)





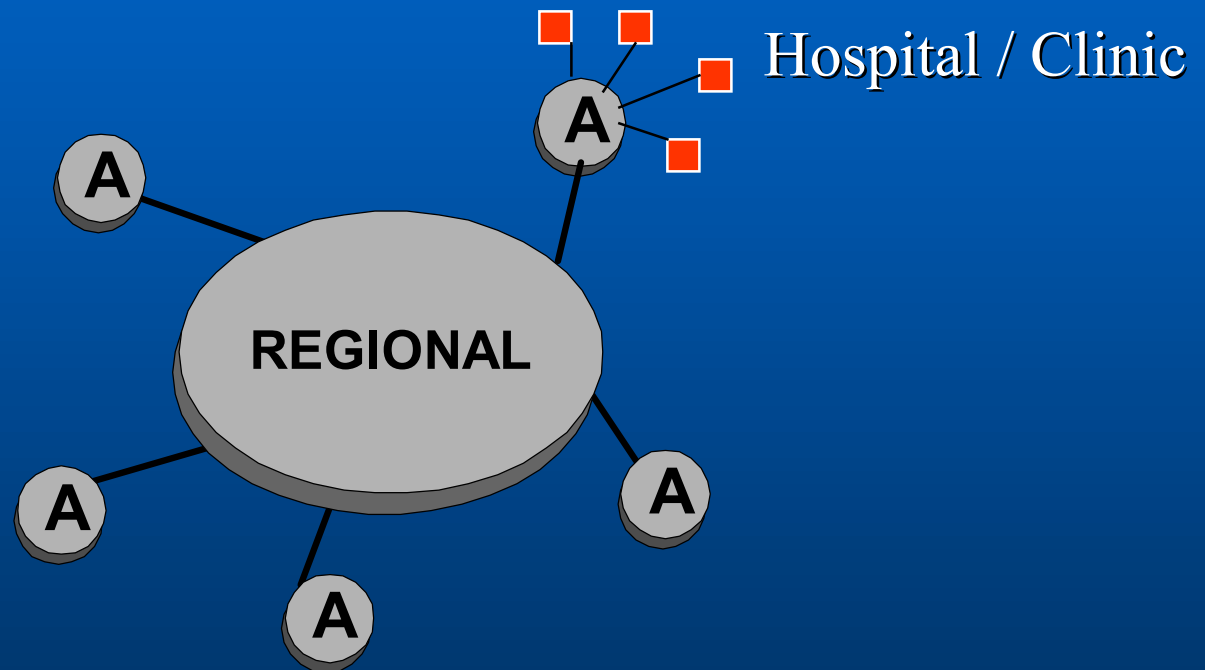
# Storage Hierarchy

---

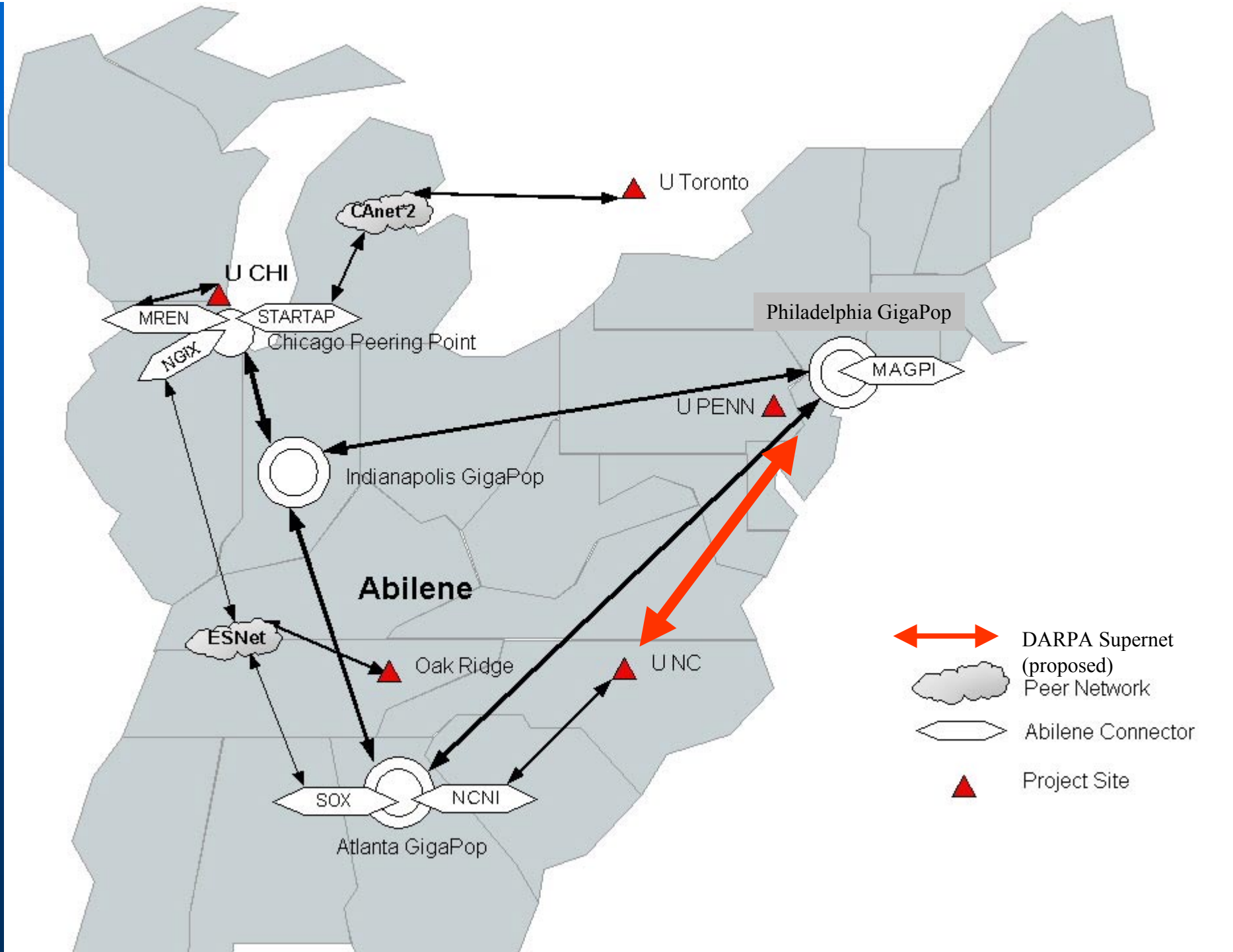
7 R @ 4,000 TB/yr

20 A @ 100 TB/yr

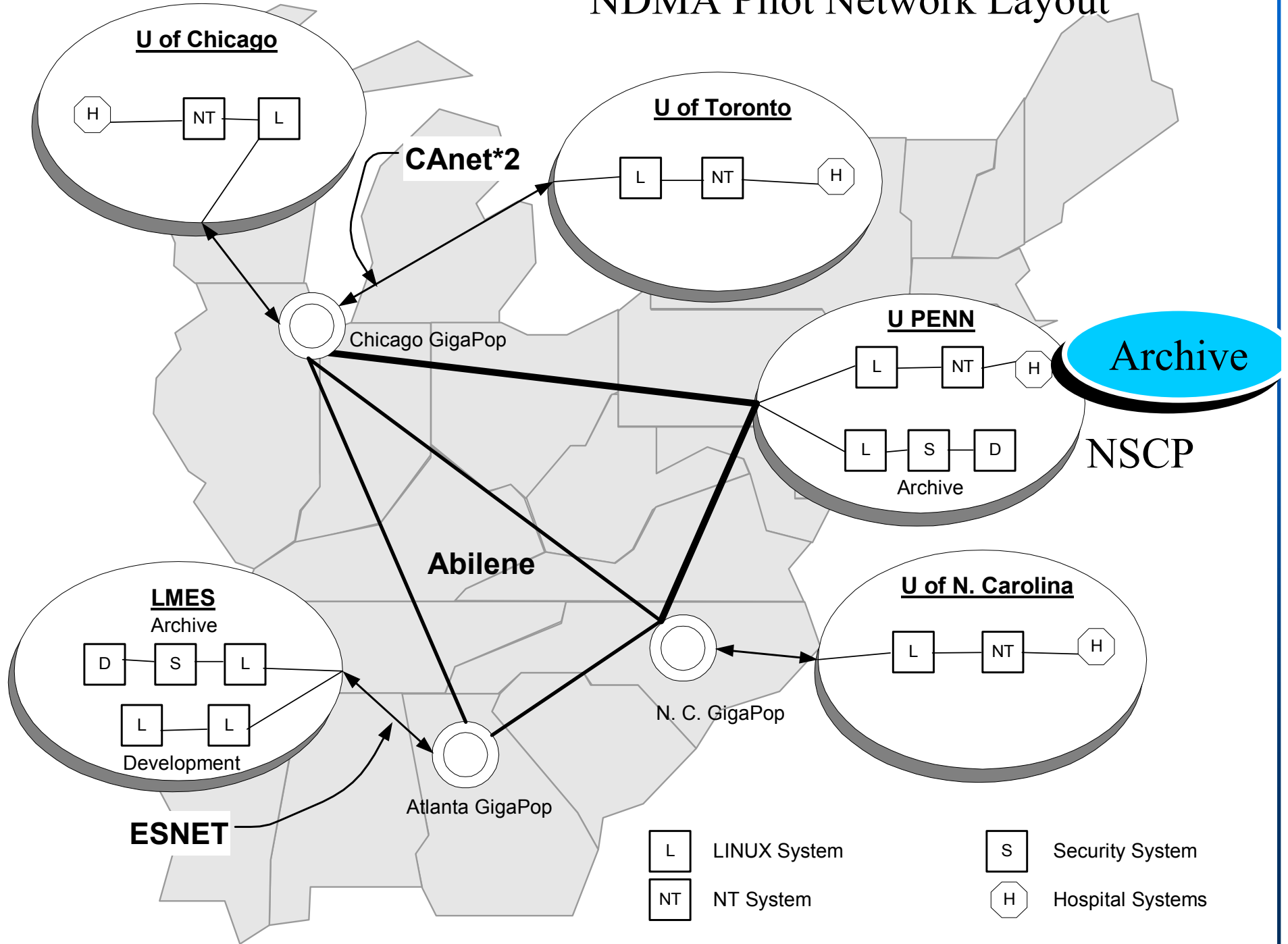
15 H @ 7 TB/yr



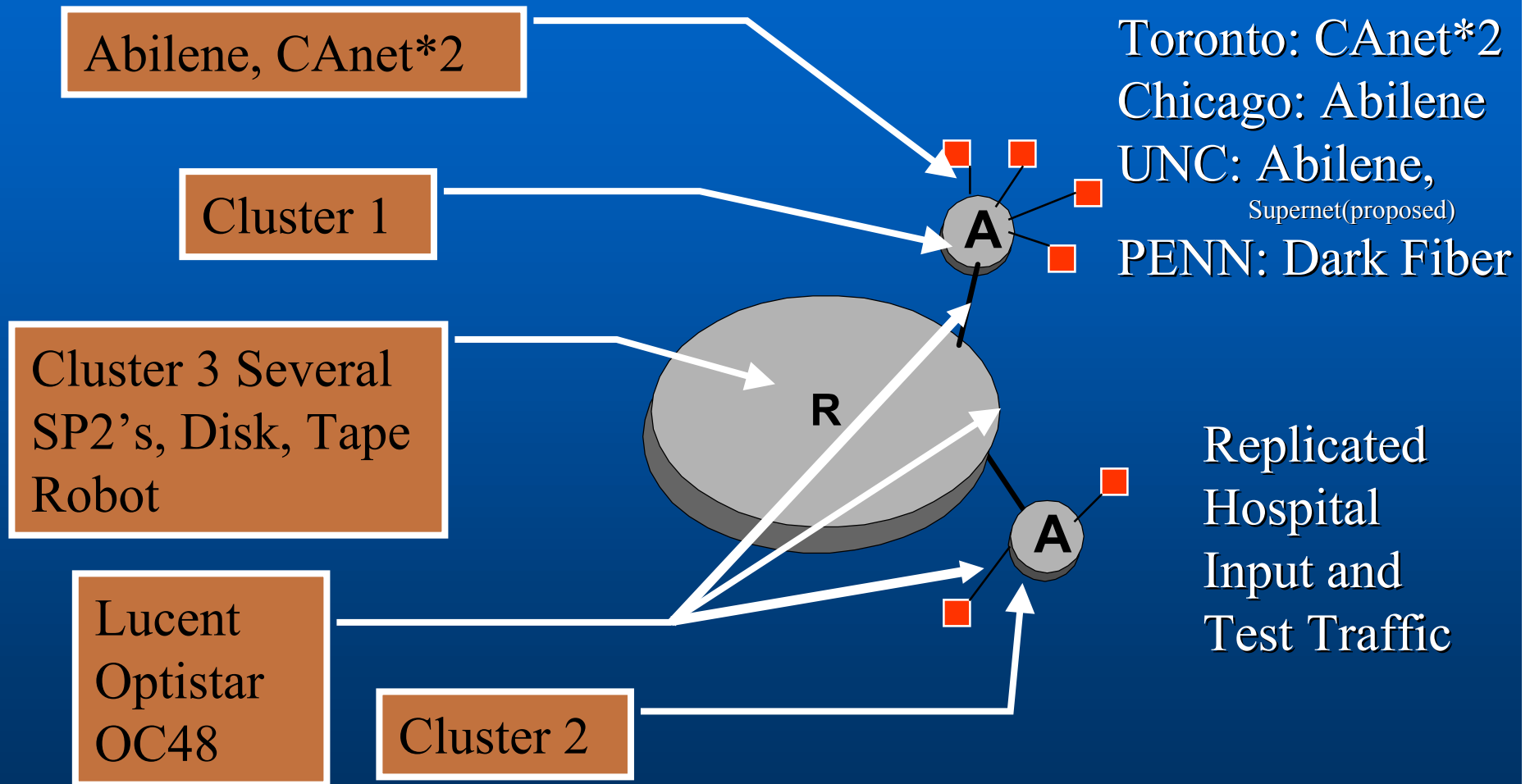
Goal: Distribute Storage Load and Balance Network and Query Loads

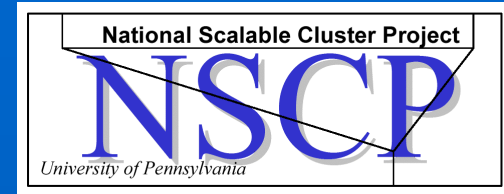


# NDMA Pilot Network Layout



# Archive Testbed

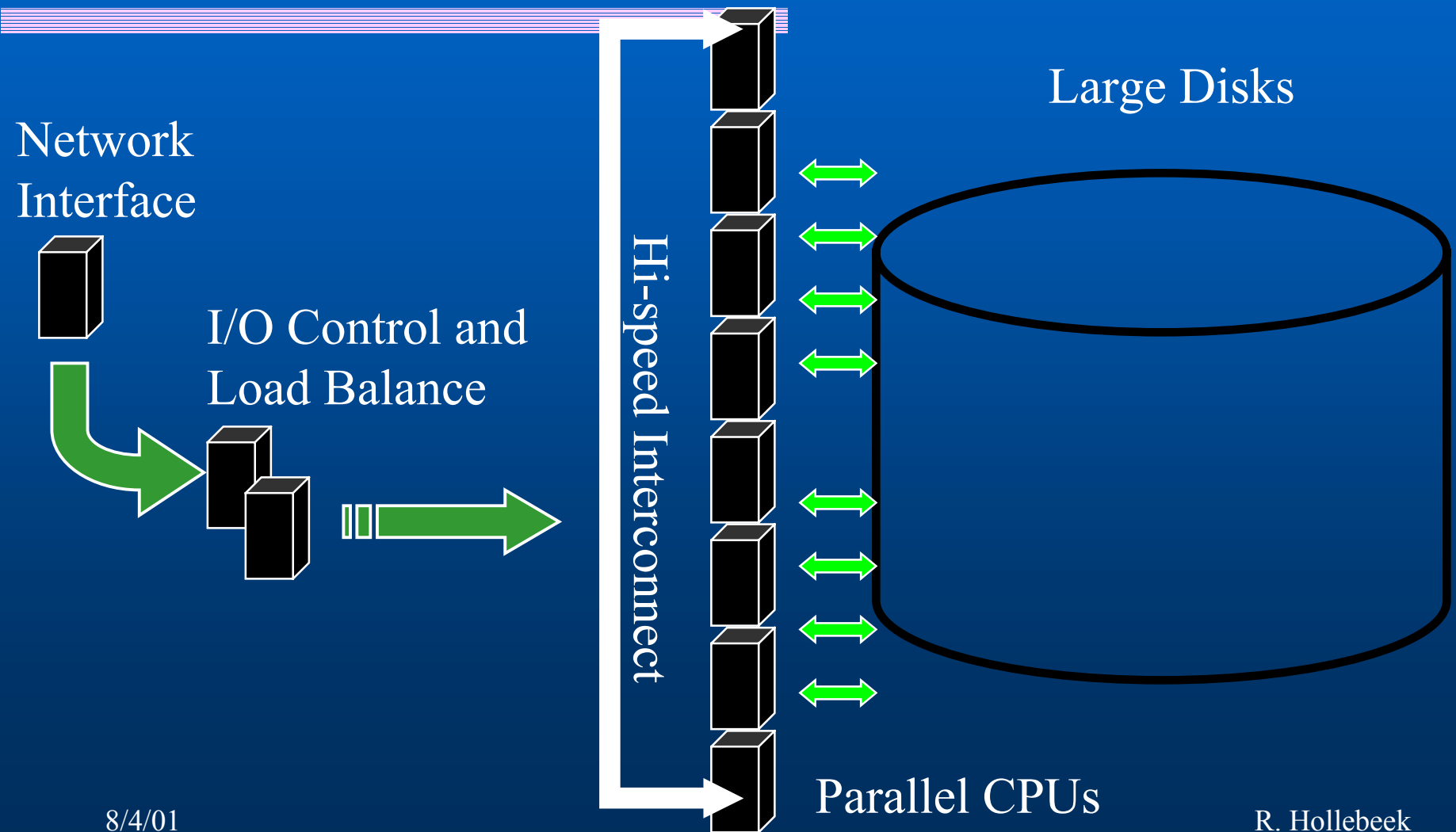




# RadAR :

- **Hierarchical Radiology Storage**
  - Input Load Balancing
  - Mirroring and Caching
  - Short, Medium, and Long term memories
  - Metadata Indices
- **Rad**iology **A**ctive **R**epository

# RadAR Components

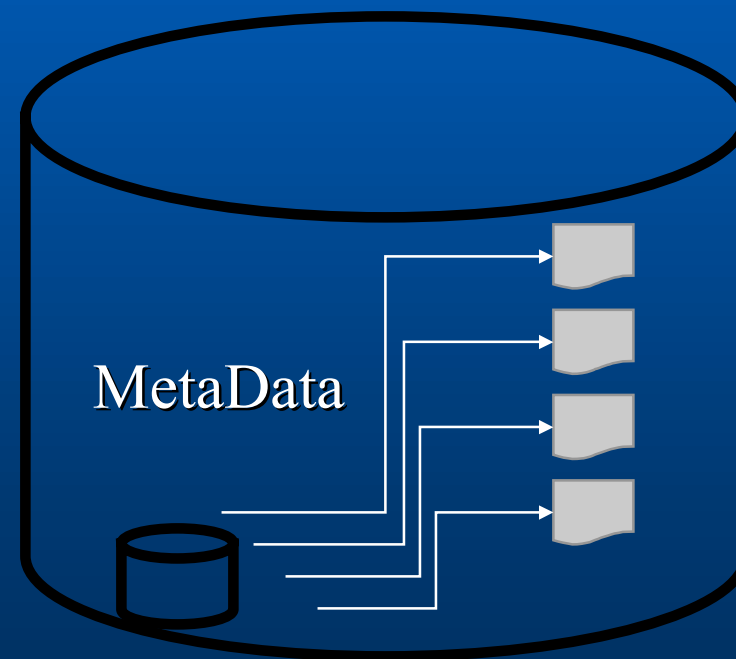


# RadAR MetaData and Indices

---

Large Disks

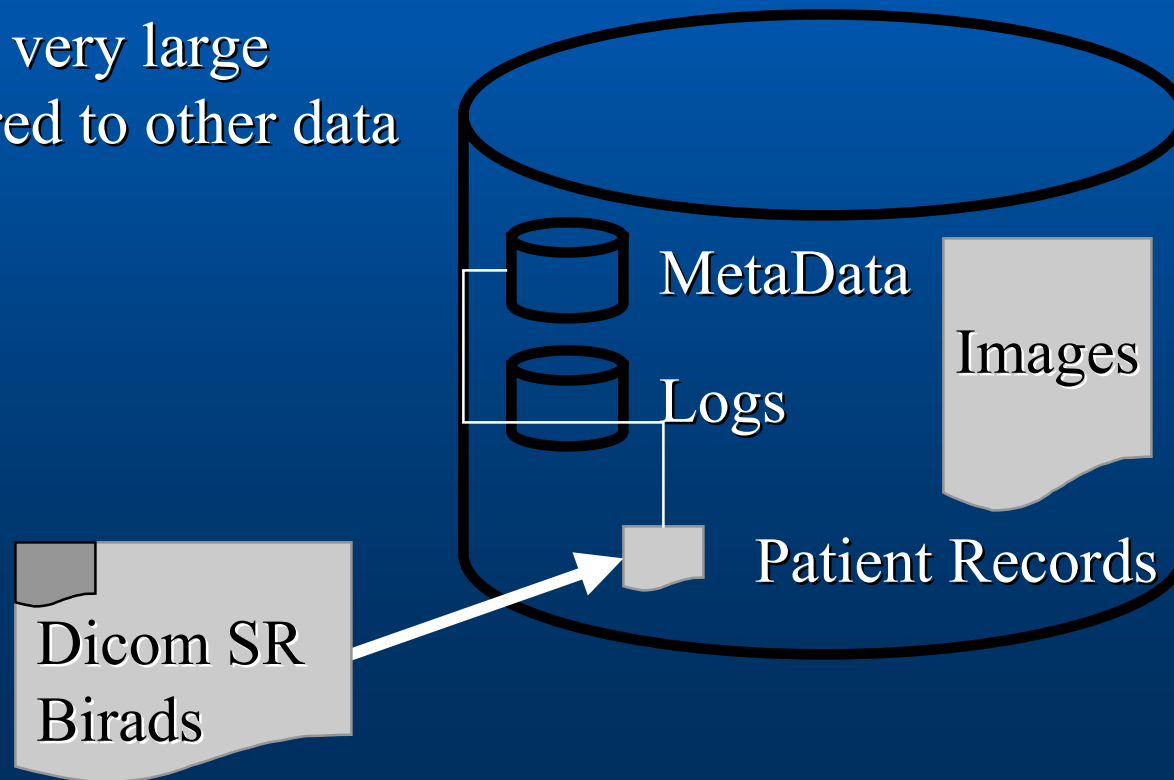
MetaData catalog  
locates data in the  
store



# RadAR Contents

Not to scale  
Images very large  
compared to other data

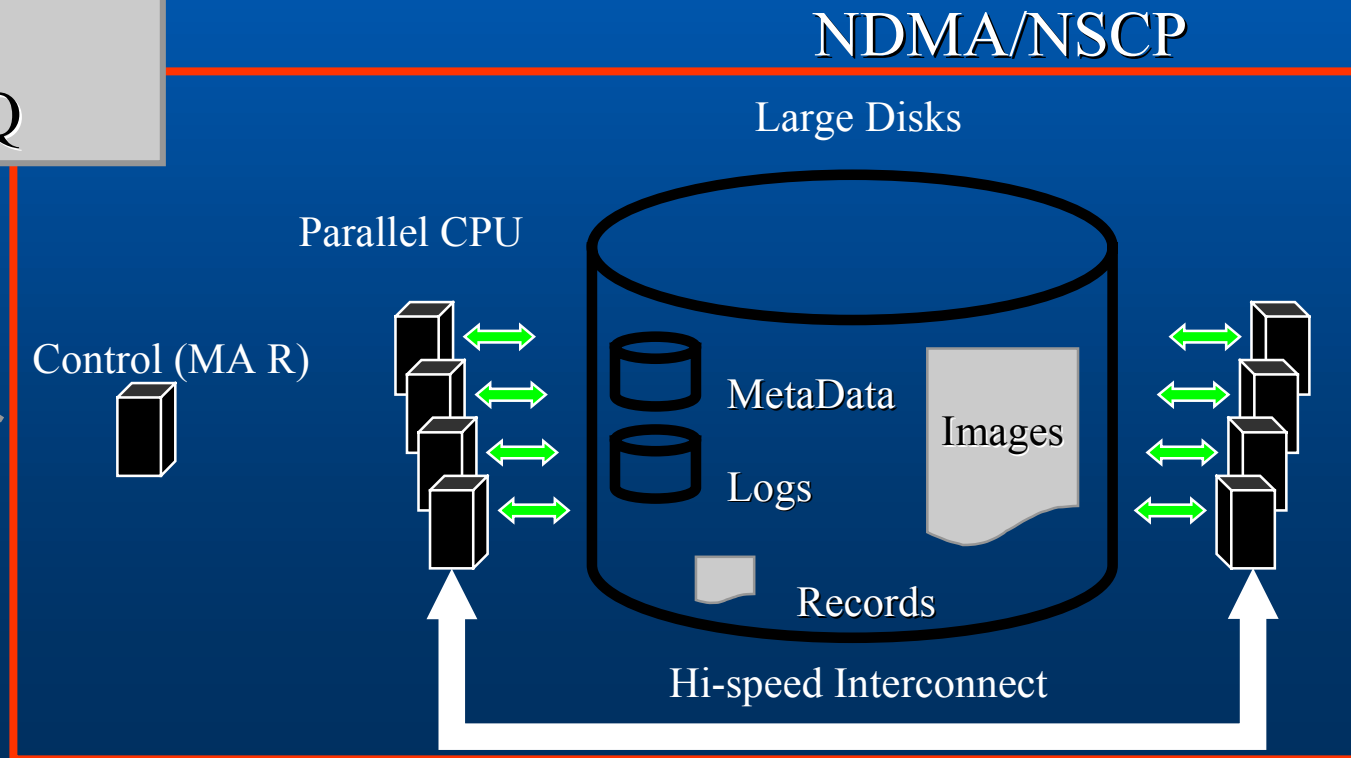
Large Disks



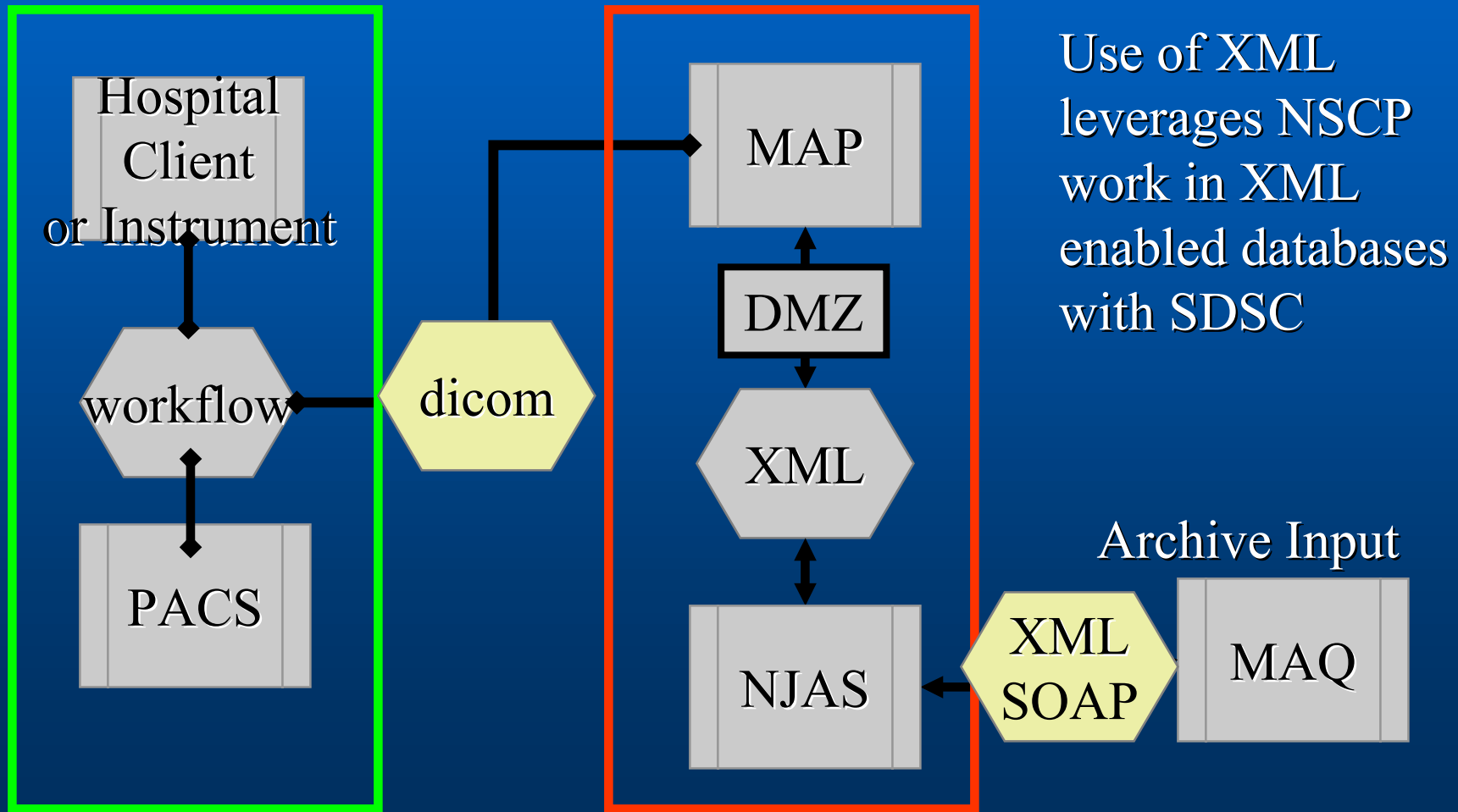


# RadAR + Portals

Portal Systems  
at HUP, UNC,  
UC, SWH  
MAP/MAQ



# Portal to Archive

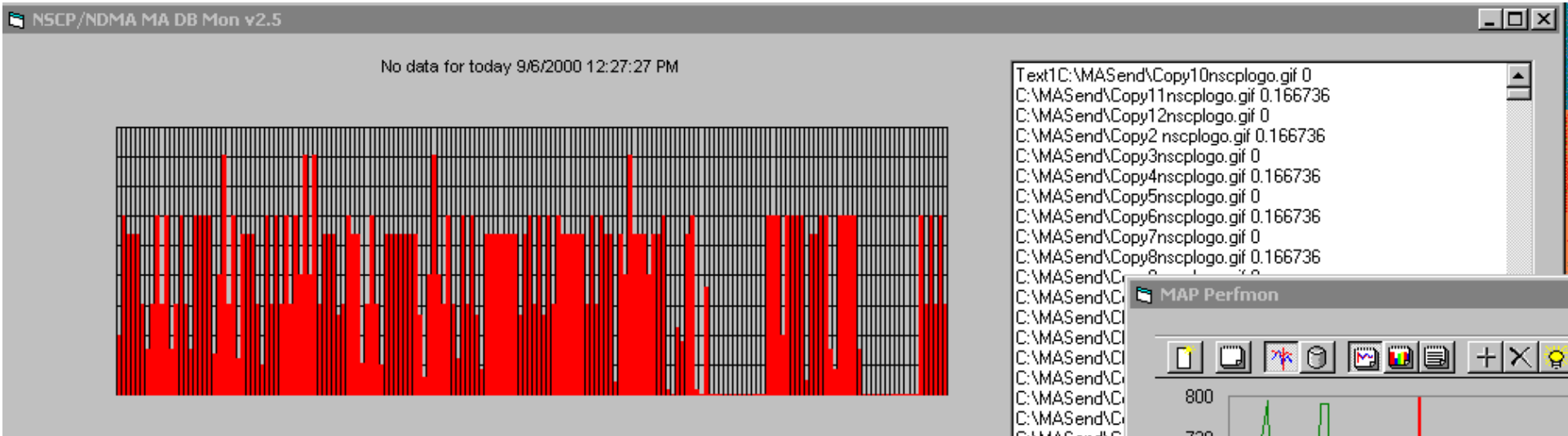


# Map - MA system portal

---



Two Dual Processor IBM/Netfinity 5100 systems  
with Remote Management Processors



NSCP/NDMA MA Portal v 3.31

From Host: 130.91.51.45 ndma04

Files to Transfer...

CRPROPD-A.200002222114f  
 CRPROPD-A.2000022222141f  
 CRPROPD-A.200002222224f  
 CRPROPD-A.200002222224f  
 CRPROPD-A.2000022222341f  
 CRPROPD-A.2000022222341f  
 CRPROPD-A.2000022222370f  
 CRPROPD-A.2000022222370f  
 CRPROPD-A.2000022222370f  
 CRPROPD-A.2000022222370f  
 CRPROPD-A.2000022222560f  
 CRPROPD-A.2000022222560f  
 CRPROPD-A.2000022223132f  
 CRPROPD-A.2000022223132f  
 CRPROPD-A.2000022223132f  
 CRPROPD-A.2000022223192f  
 CRPROPD-A.2000022223201f  
 CRPROPD-A.2000022223212f

AutoSend  Local Simulate

File Transfer 1 In Progress

Transfer Stats

12 sec  
 7.534478 MBytes  
 15.068928 Mbit/sec

Status OK  
 9/6/2000 12:27:42 PM host resolved  
 9/6/2000 12:27:42 PM connecting  
 9/6/2000 12:27:42 PM connected  
 9/6/2000 12:27:42 PM receiving response  
 9/6/2000 12:27:42 PM response received  
 9/6/2000 12:27:42 PM requesting



# Scale of the Problem

---

Recent FDA approval and cost and other advantages of digital devices will encourage digital radiology conversion

- **2000 Hospitals x 7 TB per year x 2**
- **28 PetaBytes per year**
  - (1 Petabyte = 1 Million Gigabytes )
- **Pilot Problem scale in NDMA**
  - $4 \times 7 \times 2 = 56$  Terabytes / year

# Network

---

- **Regional to Regional Links OC48**
- **Time to transmit a study**
  - **1.5 Minutes at DS-3**
  - **2 sec at OC48**

## Portal Systems in the test lab at NSCP/PENN



8/4/01

Hollebeck

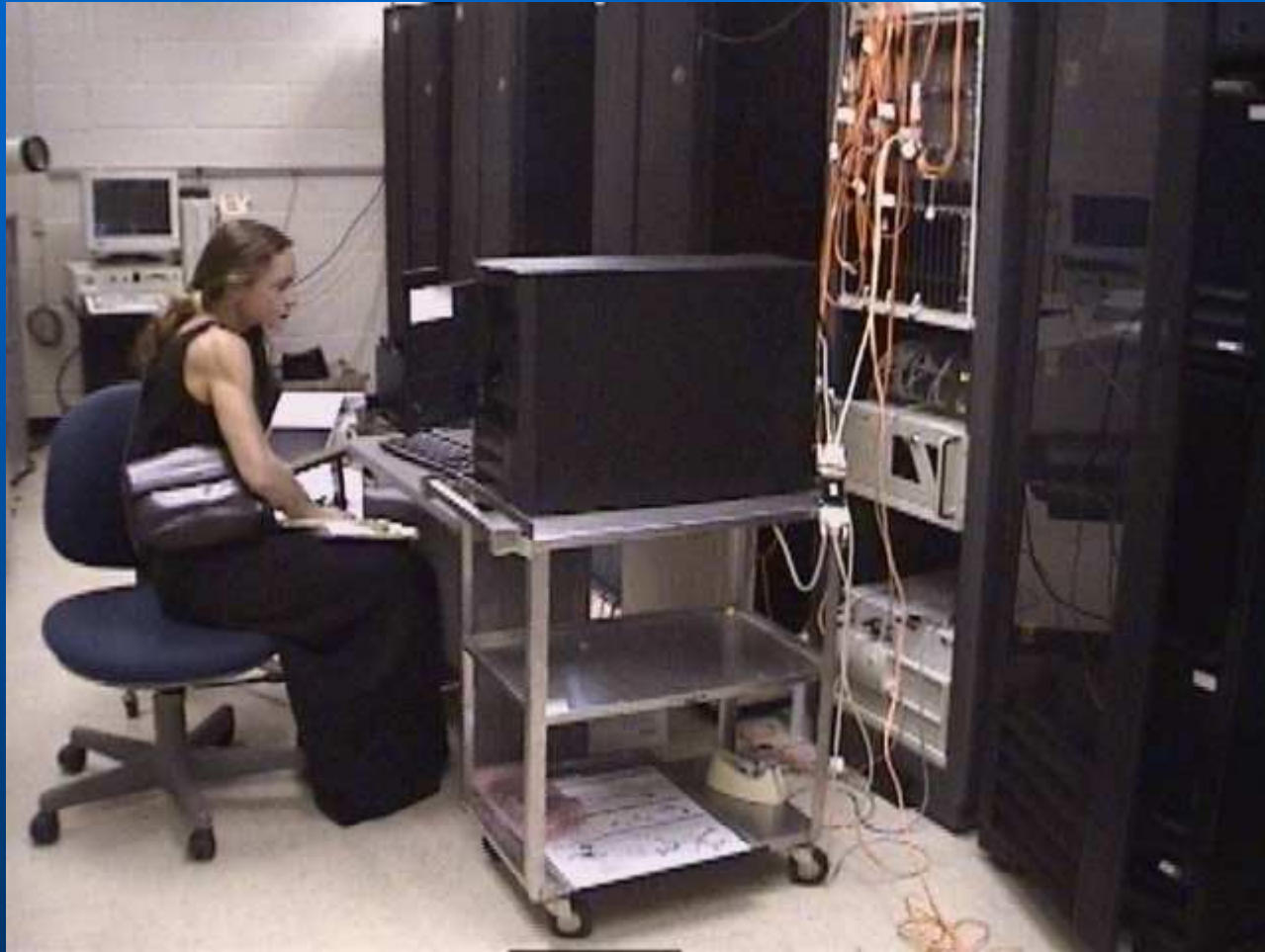


Portal NDMA01 in place in the communications closet

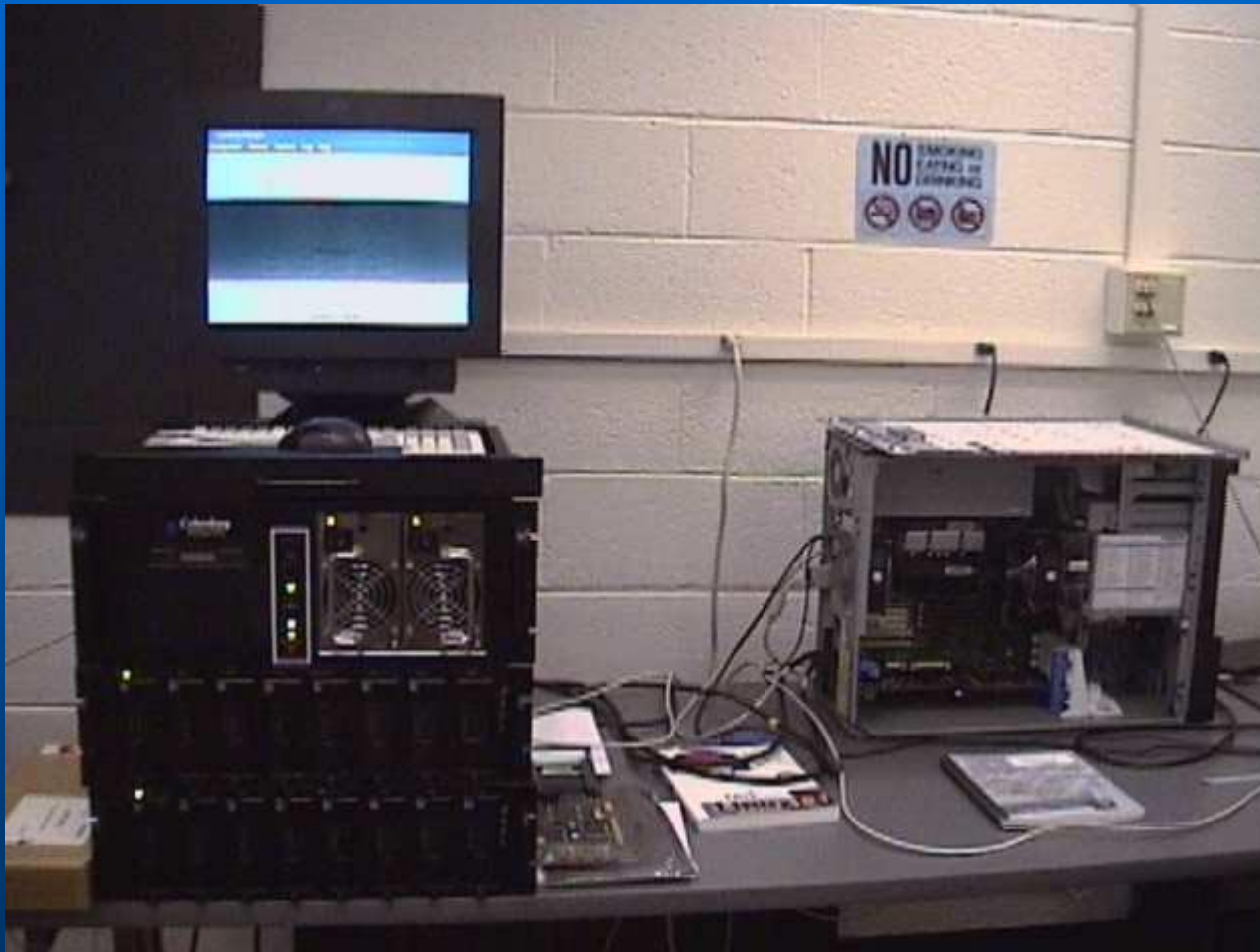




## Systems Undergoing network tests in the server room



1200 Gigabyte OC48 disk under test in a joint program with Lucent and Storage Computers Inc.



# NDMA

---

- **Store Records for retrieval**
  - typical request would retrieve 3-4 yrs
- **Audit and log transmissions**
- **Parse, Index and Store incoming information**
- **Support Computer Assisted Diagnostics**
- **Support Radiologist Training and Evaluation**
- **Support Data Mining**

## Computer-Aided Diagnosis

- Application of algorithms to reduce variability
- Front-end pre-processing using Focus-of- Attention Region algorithms
- Potential centralized service over the network
- Auto-training of CAD tools



## Training, Teaching, Evaluation



# Other Projects

---

- **DARPA Supernet** (proposed) UNC/PENN joint research on NDMA, Radiology consults, and teleimmersion
- **Lucent:** OC48 Optistar switch testbed lab
- **IBM: (proposed)** Regional Archive Pilot
- **Storage Computer:** Terabyte disk tests
- **qmp: queue mover protocol:** development of a protocol for moving identified content on a point to point network with high bandwidth x latency.

# Possible US-UK Projects

---

- **Storage/Retrieval**
- **CAD (Computer Assisted Diagnosis)**
- **Training and Teaching**

# US-UK Project 1: Storage

---

- **Transatlantic records tests**
  - **deploy portal**
    - requires UK site with a digital mammo system
    - requires Hospital Review Board approval
  - **setup VPN**
  - **use developed transmission standards**
  - **store phantom and anonymous records**
  - **request/store**
    - anonymous and/or clinical records
    - **difficult, multi-consult, or teaching cases**

# US-UK Project 2: CAD

---

- **CAD Server:**
  - connect UK site to CADS through VPN
  - Transfer images via net
  - Process
  - Return Annotated Results



# US-UK Project 3: Training and Teaching

---

- **Access Teaching Application and annotated cases through VPN**
  - Request collection
  - Transfer
  - Test/Train
  - Return Results and score

# Conclusions

---

- **Would welcome collaboration with hospitals and other groups in the UK**

**bobh@nscp.upenn.edu**

**NDMA -**

**<http://nscp.upenn.edu/ndma>**

# NDMA Participants

- NSCP@Penn:
  - **Digital Storage, Search and Retrieval**
- Oak Ridge National Lab:
  - **Network (VPN) and Security**
- Hospitals of
  - **University of Pennsylvania**
  - **University of Chicago**
  - **University of North Carolina**
  - **University of Toronto**

**bobh@nscp.upenn.edu**

**Available after 8/7/01**

**<http://nscp.upenn.edu/hollebeek/talks/usuk>**

