PROPOSED ARCHITECTURE

Storage Manager

Collection Manager

Presentation System

Network

Storage Service

Search Engine

ADVANCED TECHNOLOGY

- RAM INDEXES

- PAGE IMAGES ON MAG. DISK.
Proposed Architecture

Storage Manager

Collection Manager

Presentation System

Network

Storage Service

Search Engine

Machine-to-Machine Protocol

Like Z39.50 but
- Stateless
- Search separate from Storage
- Front-end only update
Form: Scanned image, 300 dpi
Hodnium: Magnetic disk
Access time: 20 ms.
Space: 750 ft²
Cost: $10M
RESEARCH PROBLEMS

1. What is in a link?

2. Where to cache what?

3. Connecting scanned images with ASCII

4. Compression that allows low-resolution extraction

5. How to apply client/server architecture?

6. Coping with political/administrative inhibitions

7. Managing very large storage
   - 22 Gbyte disks
   - 70-700 Gbyte optical jukeboxes
MANAGING VERY LARGE STORAGE

1. **Need** more than hierarchy if you have 500,000 files. Property lists? Search on properties? On contents?

2. **Backup:** nothing removable is big enough.

3. **Media obsolescence:** by the time the disk is filled it will be obsolete. Need system plan.

4. **Should** compression be a file system operation?

5. **Does** forward error correction fit in?

6. **Coordination** of variant copies.