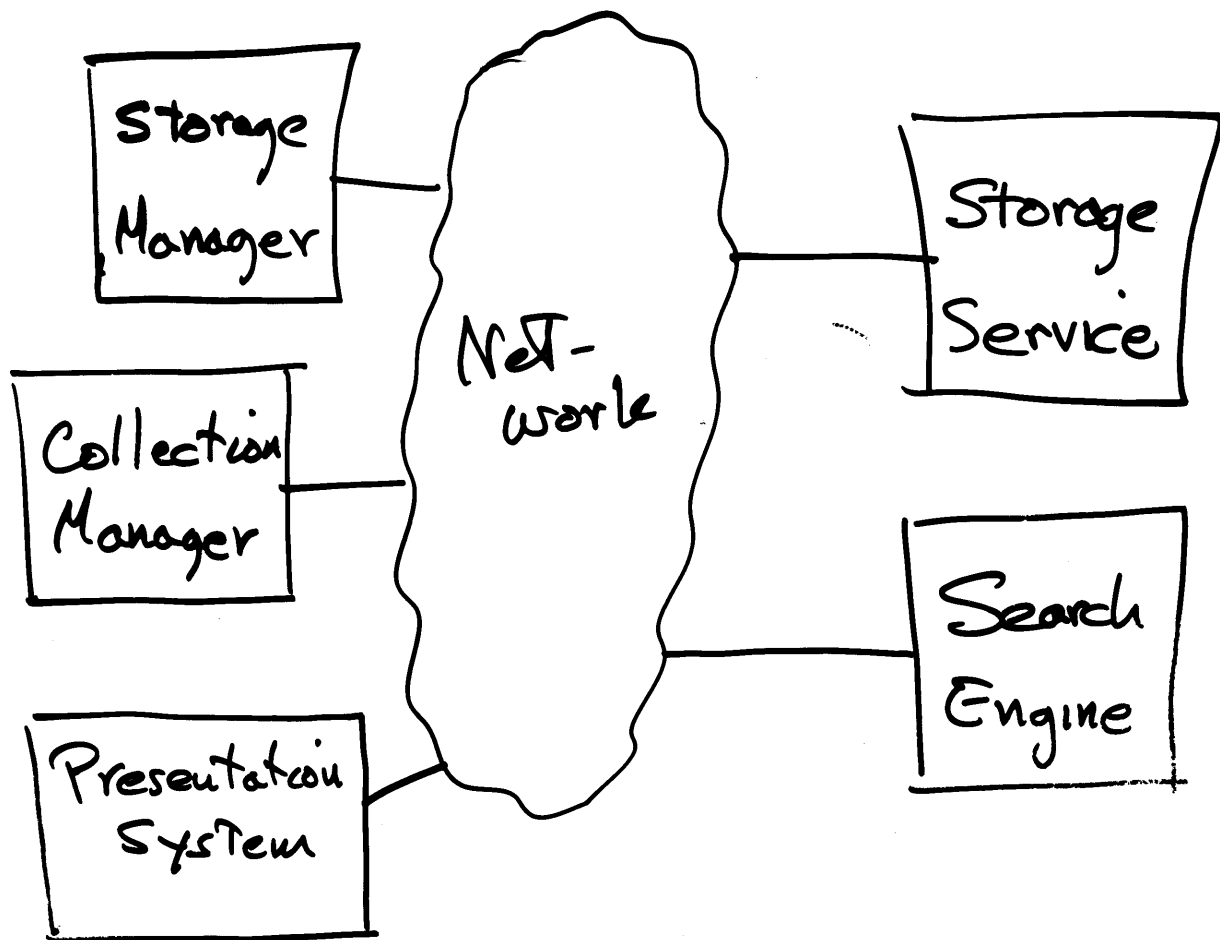


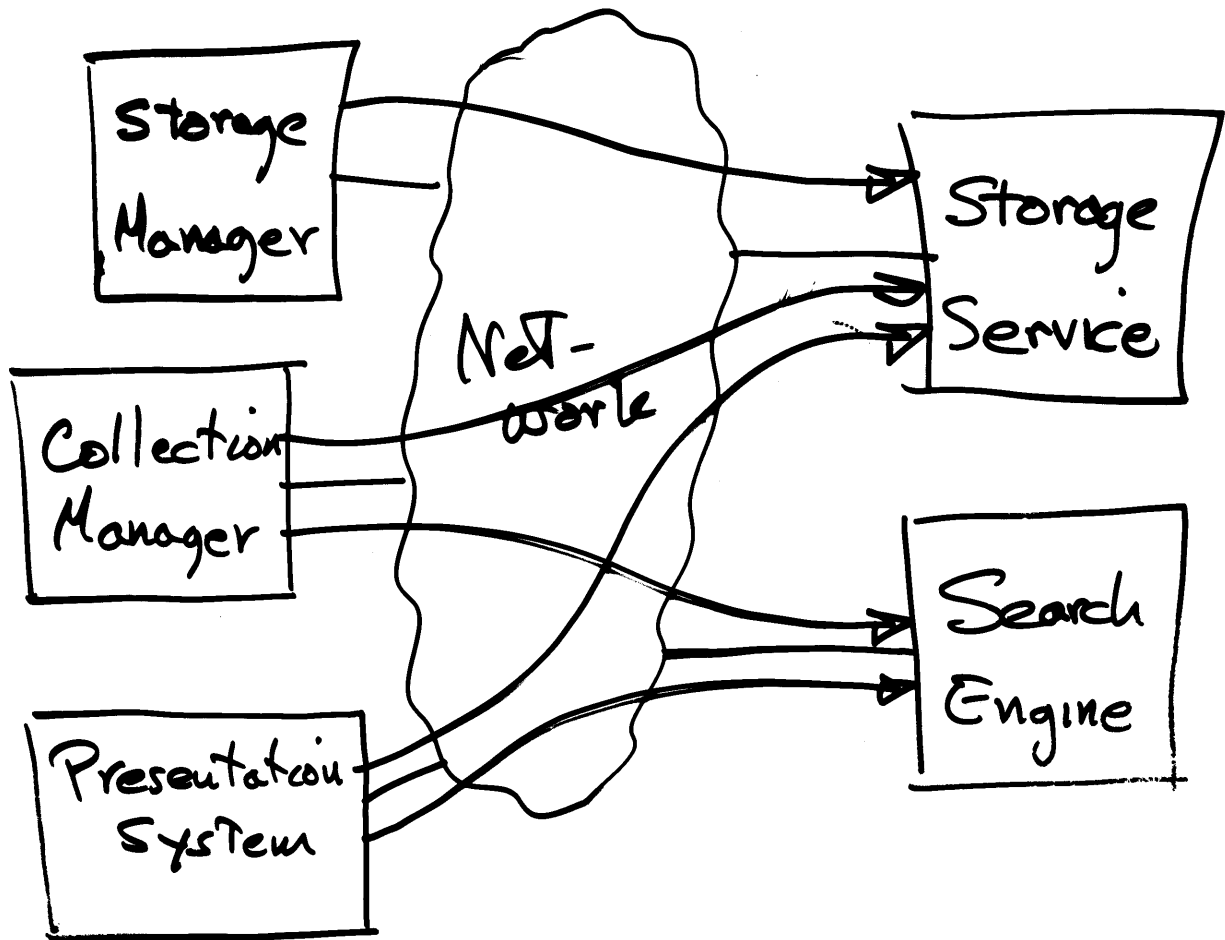
# PROPOSED ARCHITECTURE



## ADVANCED TECHNOLOGY

- RAM INDEXES
- PAGE IMAGES ON MAG. DISK.

# PROPOSED ARCHITECTURE



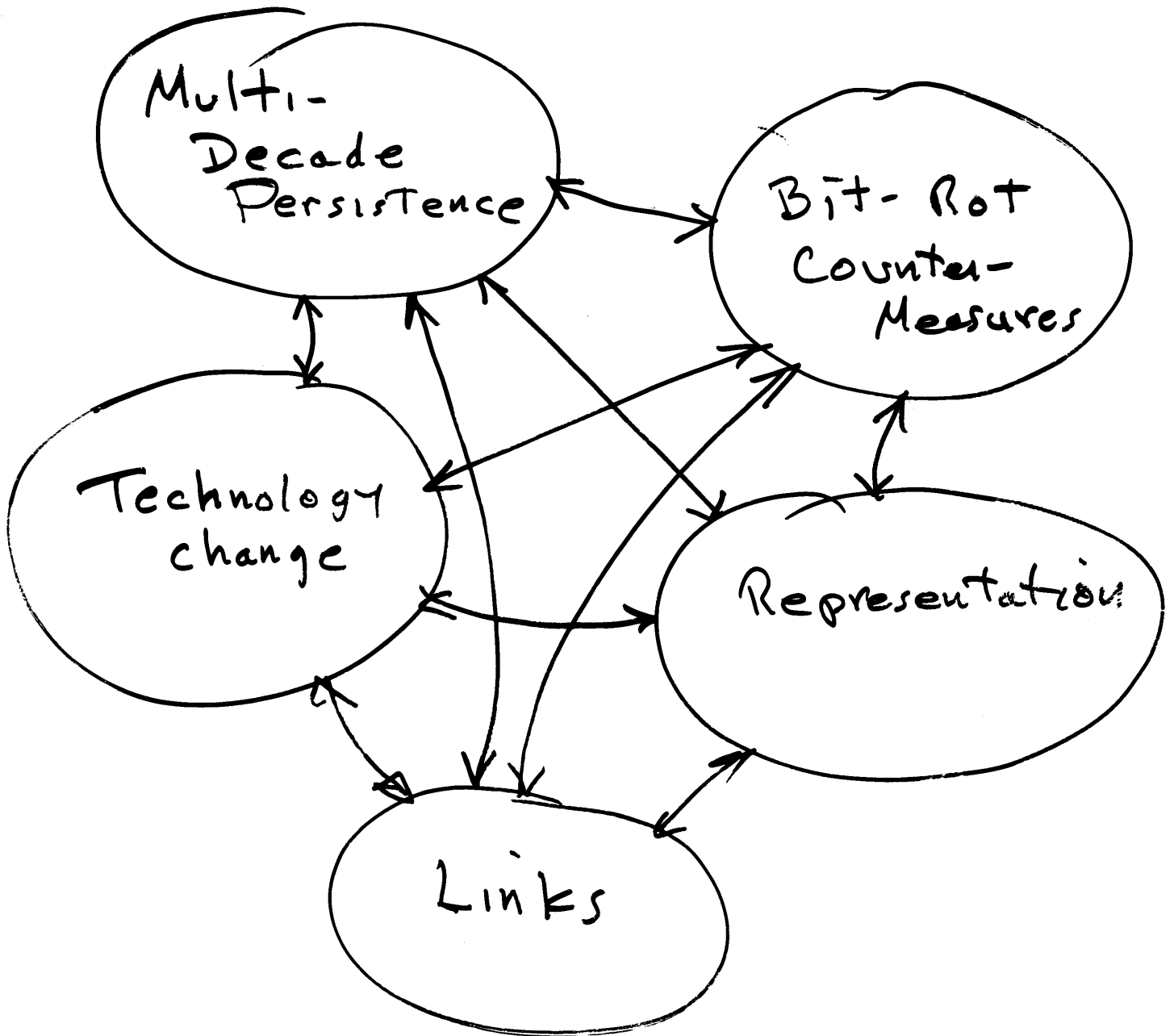
## MACHINE-TO-MACHINE PROTOCOL

Like Z39.50 but

- State less
- Search separate from Storage
- Front-door update

# ENGINEERING

# TRADEOFFS



M.I.T. LIBRARY Ca. 1997

30'

Form: Scanned image, 300 dpi

Medium: Magnetic disk

Access time: 20 ms.

Space: 750 pt<sup>2</sup>

Cost: \$10M

25'

## 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.**

# Observations

*Truth:* Multiple versions

*Indexing:* Provided by  
scholars  
and idiots

*Space:* Severe shortage

*Conservation:* A crisis