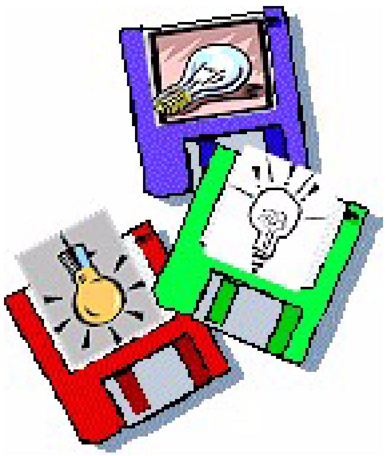
Active Disks For Large-Scale Data Mining

Erik Riedel, Garth Gibson
Parallel Data Laboratory

Andrew Moore, Christos Faloutsos
Center for Automated Learning & Discovery

Carnegie Mellon University
www.pdl.cs.cmu.edu/Active



SIGMOD-DMKD '98 Workshop
June 5, 1998

Outline

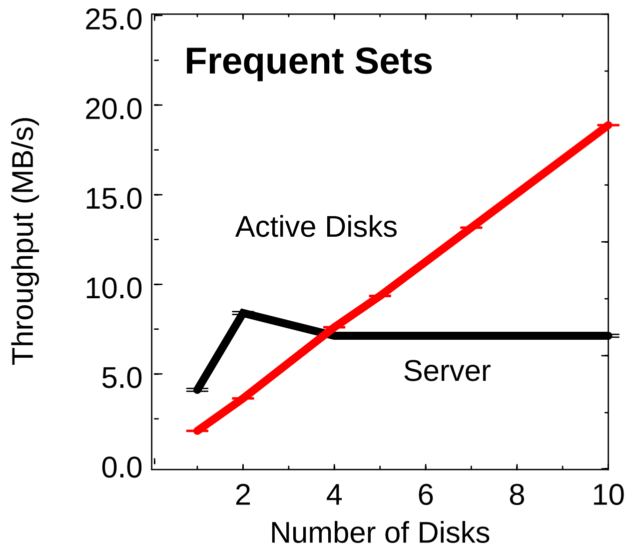
Network-Attached Disks

Industry Trends

Active Disks

Applications

Speedups



Prototype

Motivation

Increasing importance of Data Mining and Multimedia

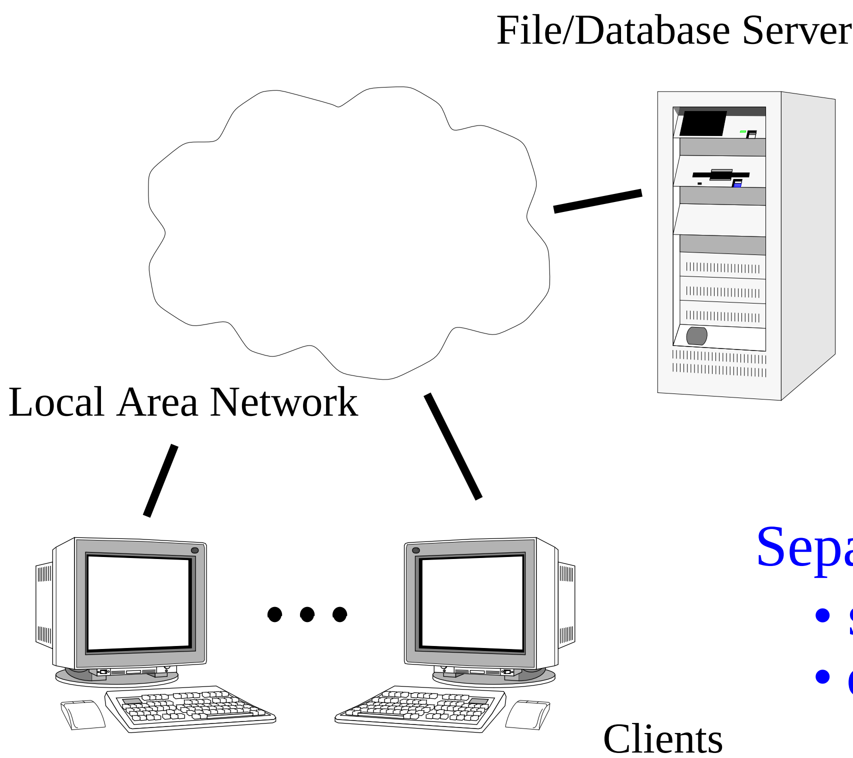
- **Large objects => many disks**
- **High processing rates => high storage bandwidth**
- **No legacy code => applications use stock file systems**

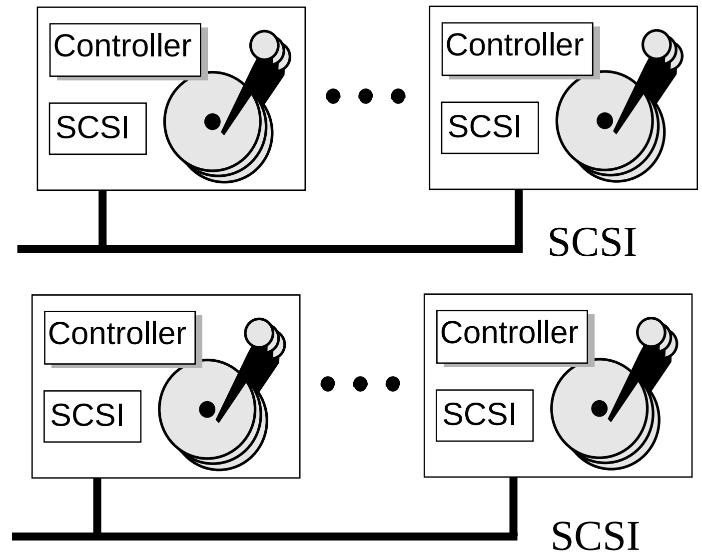
Collaboration

- **Center for Automated Learning & Discovery (Mitchell, Feinberg, Eddy, ...)**
- **Multimedia - Informedia/Digital Library (Wactlar, Reddy, Kanade, ...)**

Today’s Server-Attached Disks

Store-and-forward data copy through server machine



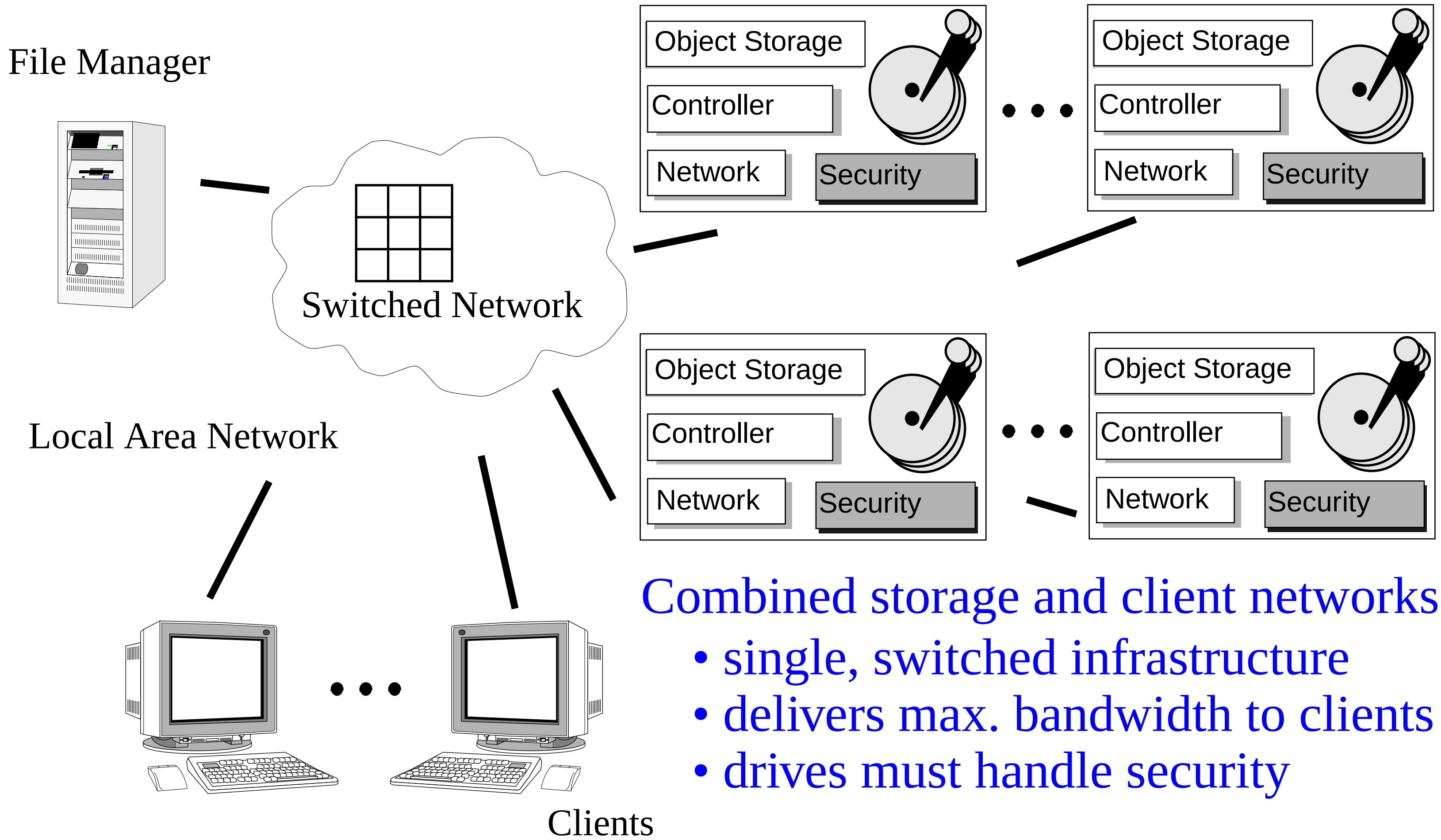


Separate storage and client networks

- storage moving to packetized FC
- clients moving to scalable switches

Network-Attached Secure Disks

Eliminate server bottleneck w/ network-attached



Combined storage and client networks

- single, switched infrastructure
- delivers max. bandwidth to clients
- drives must handle security

Working Group on Network-Attached Storage

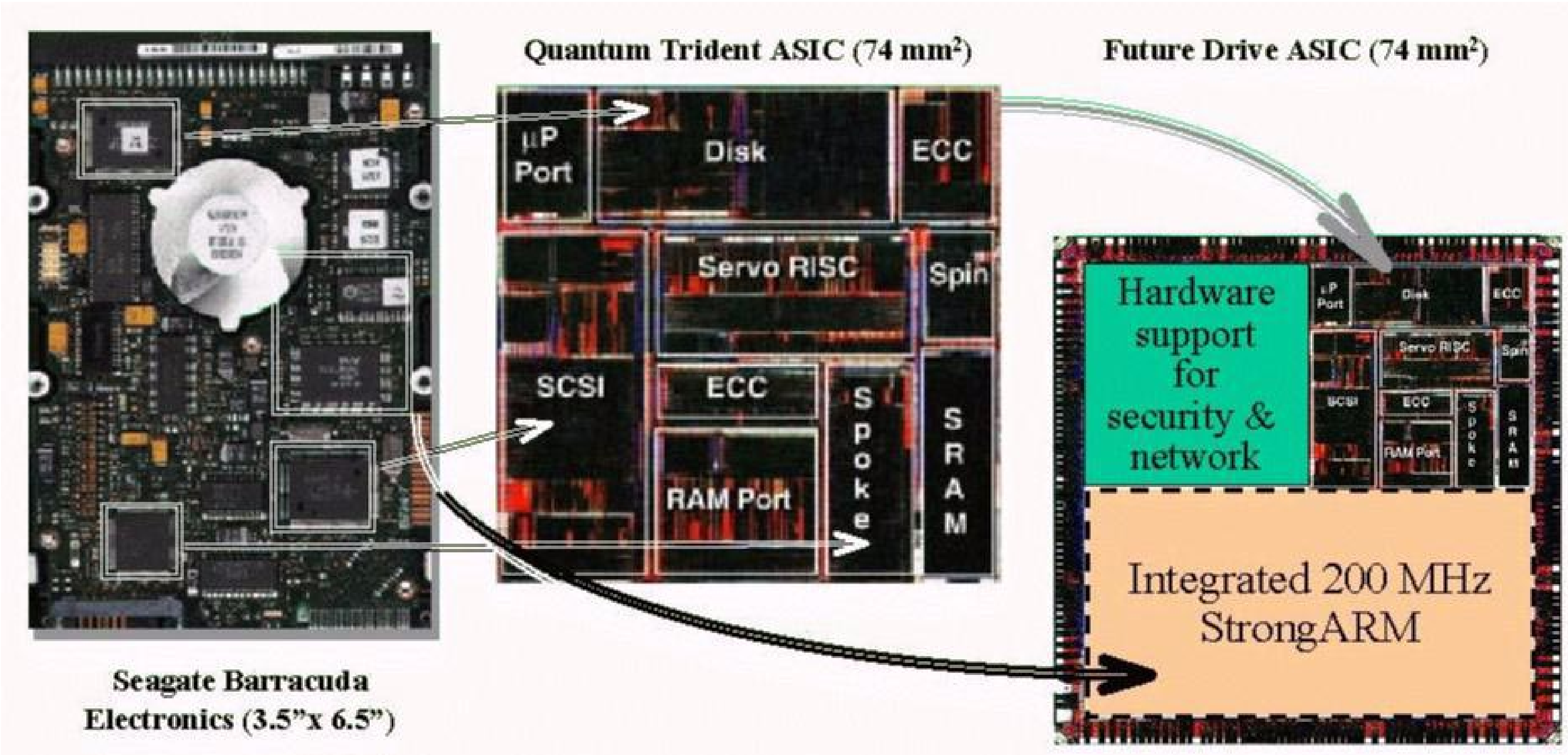
National Storage Industry Consortium (NSIC)

- **launched April 1996 (CMU, HP, IBM, StorageTek)**
  www.nsic.org/nasd
- **signed IP rights sharing agreement January 1997**
  CMU, HP, IBM, StorageTek, Seagate, Quantum
- **participants execute independently funded research, share issues impacting NASD architecture/interfaces**
- **quarterly meetings**
- **public workshop with each meeting**
- **recent SNIA effort to reach larger community (*www.snia.org*)**

Pre-standards recommendations

- **Object-oriented disks (SCSI-4)**
- **Attributes for self-managed storage**

Excess Device Cycles Are Coming



Higher and higher levels of integration in drive electronics

- specialized drive chips combined into single ASIC
- technology trends push toward integrated control processor
- 75 MHz, 32-bit superscalar w/ 2 MB on-chip RAM available in '98

Large database systems - lots of disks, lots of power

System	Processing (MHz)		Data Rate (MB/s)	
	CPU	Disks	I/O Bus	Disks
Compaq TPC-C	4 x 200=**800**	*113* x 75=**8,475**	**133**	**1,130**
Microsoft Terraserver	4 x 400=**1,600**	*320* x 75=**24,000**	**532**	**3,200**
Digital 500 TPC-C	1 x 500=**500**	*61* x 75=**4,575**	**266**	**610**
Digital 4100 TPC-D	4 x 466=**1,864**	*82* x 75=**6,150**	**532**	**820**

- **assume disk offers equivalent of 75 host MHz**
- **assume disk sustained data rate of 10 MB/s**

Lots more cycles and MB/s in disks than in host

Opportunity - Active Disks

Basic advantages of an Active Disks system

- **parallel processing** - lots of disks
- **bandwidth reduction** - filtering operations common
- **scheduling** - little bit of computation can go a long way

Appropriate applications

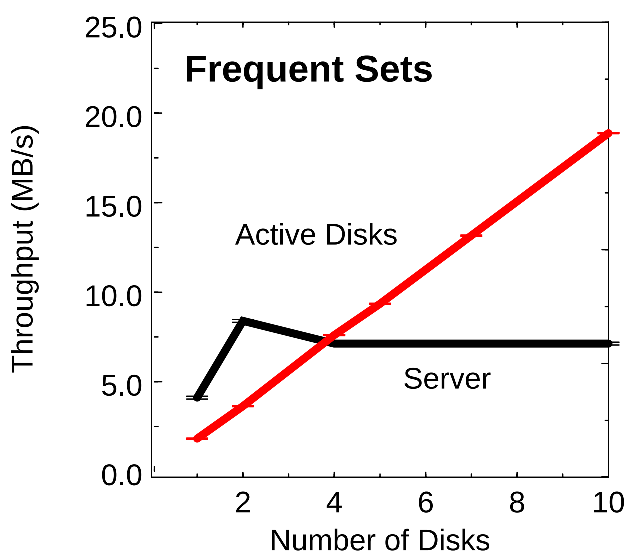
- execution time dominated by data-intensive core
- allows parallel implementation of core
- small memory footprint
- small number of cycles per byte of data processed

Example Application

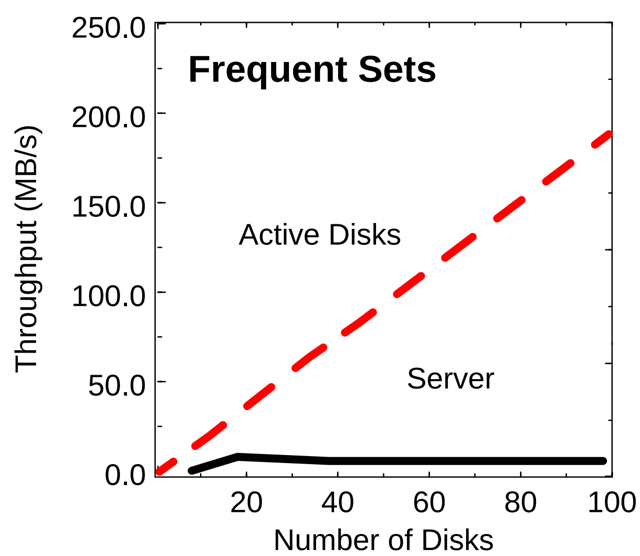
Data mining - association rules [Agrawal95]

- frequent sets summary counts
- count of *1-itemsets* and *2-itemsets*
- milk & bread => cheese
- diapers & beer

Prototype



Scaling Up



Additional Applications

Database - select

- extract records that match a particular predicate

Database - nearest neighbor search

- *k* records closest to input record
- with large number of attributes, reduces to scan

Multimedia - edge detection [Smith95]

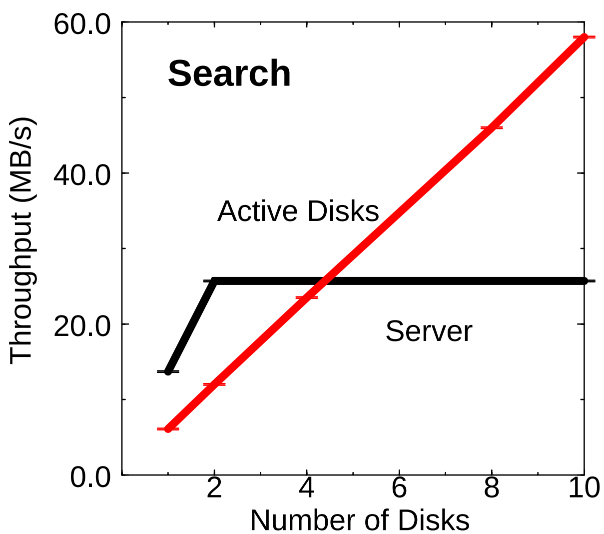
- detect edges in an image

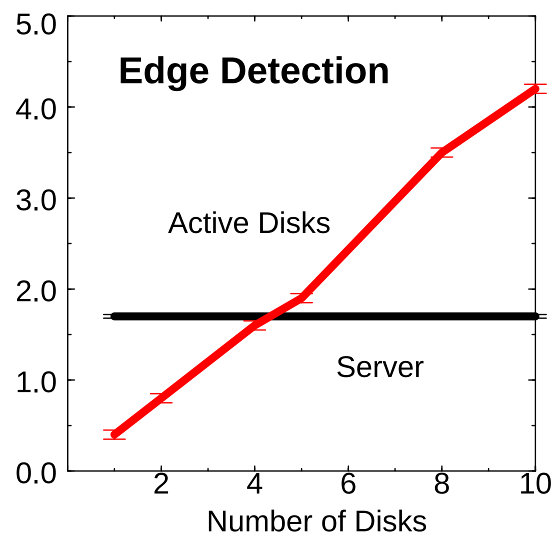
Multimedia - image registration [Welling97]

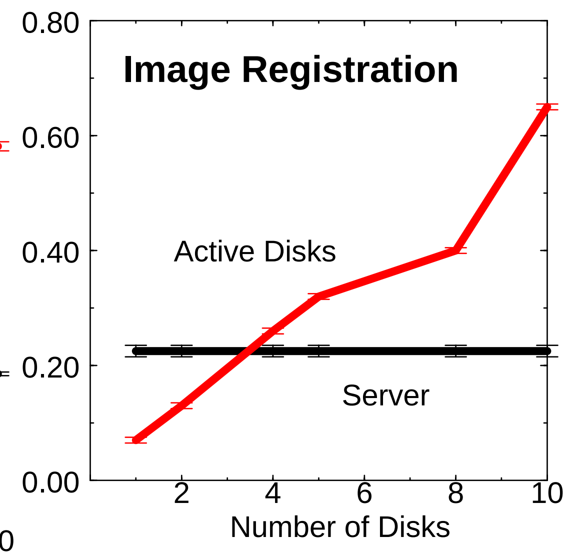
- find rotation and translation from reference image

Performance with Active Disks

application	input	computation (inst/byte)	throughput (MB/s)	memory (KB)	selectivity (factor)	bandwidth (KB/s)
Select	m=1%	7	28.6	-	100	300
Search	k=10	7	28.6	72	80,500	0.1
Frequent Sets	s=0.25%	16	12.5	620	15,000	1
Edge Detection	t=75	303	0.67	1776	110	2
Image Registration	-	4740	0.04	672	150	2







Future Directions

Executables downloaded into drives

- safe, secure, controllable

Applications: schedule, semantic extension

- sort, join, collective I/O, video, web, storage mgmt

Compiler-assisted “Disklet” definition

- library, framework support, automatic partitioning

Active networking for storage

- NASD capabilities extended to network components
- in network: protocol conversion, caching, dynamic routing

Summary

Scalable speedup for Data Mining and Multimedia

- parallel implementations exist
- small footprint, small cycles per byte, data-intensive

Storage industry is listening

- "free" computational power is coming soon
- NSIC/NASD pre-standards group hard at work

Scales down too

- about 4 disks match a host processor (2 VLSI generations)
- factors of 2-3 speedup with "PC" and 10 disks

NSIC/NASD June Meeting on Active Disks

June 8th, 1998 Morning: *Application code in the disk*

8:30 What to do with lots more computing in storage?, Garth Gibson, CMU

9:00 Put EVERYTHING in the Storage Device, Jim Gray, Microsoft Research

9:35 Active Disks for Data Mining and Multimedia, Erik Riedel, CMU

10:25 Intelligent Disks: A New Computing Infrastructure for Decision Support Databases, Kimberly Keeton, UC Berkeley

11:00 Active Disk Architectures for Rapidly Growing Datasets, Anurag Acharya, UC Santa Barbara

11:35 Panel Discussion

June 8th, 1998 Afternoon: *Storage and file systems support in the disk*

1:45 Consideration for smarter storage device, David Anderson, Seagate

2:20 SCSI Disk Requirements for Shared Disk File Systems, Matthew O'Keefe, Univ of Minnesota

3:15 NFS v4 and Compound Requests, Brent Callaghan, Sun Microsystems

3:50 A File system for Intelligent Disks, Randy Wang, UC Berkeley

4:25 Panel Discussion

June 9th, 1998 - Construct white paper outlining opportunities & challenges