

# *Terabyte IDE RAID-5 Disk Arrays*

David A. Sanders, Lucien M. Cremaldi, Vance Eschenburg, Romulus Godang, Christopher N. Lawrence, Chris Riley, and Donald J. Summers

University of Mississippi

Donald L. Petravick

Fermilab

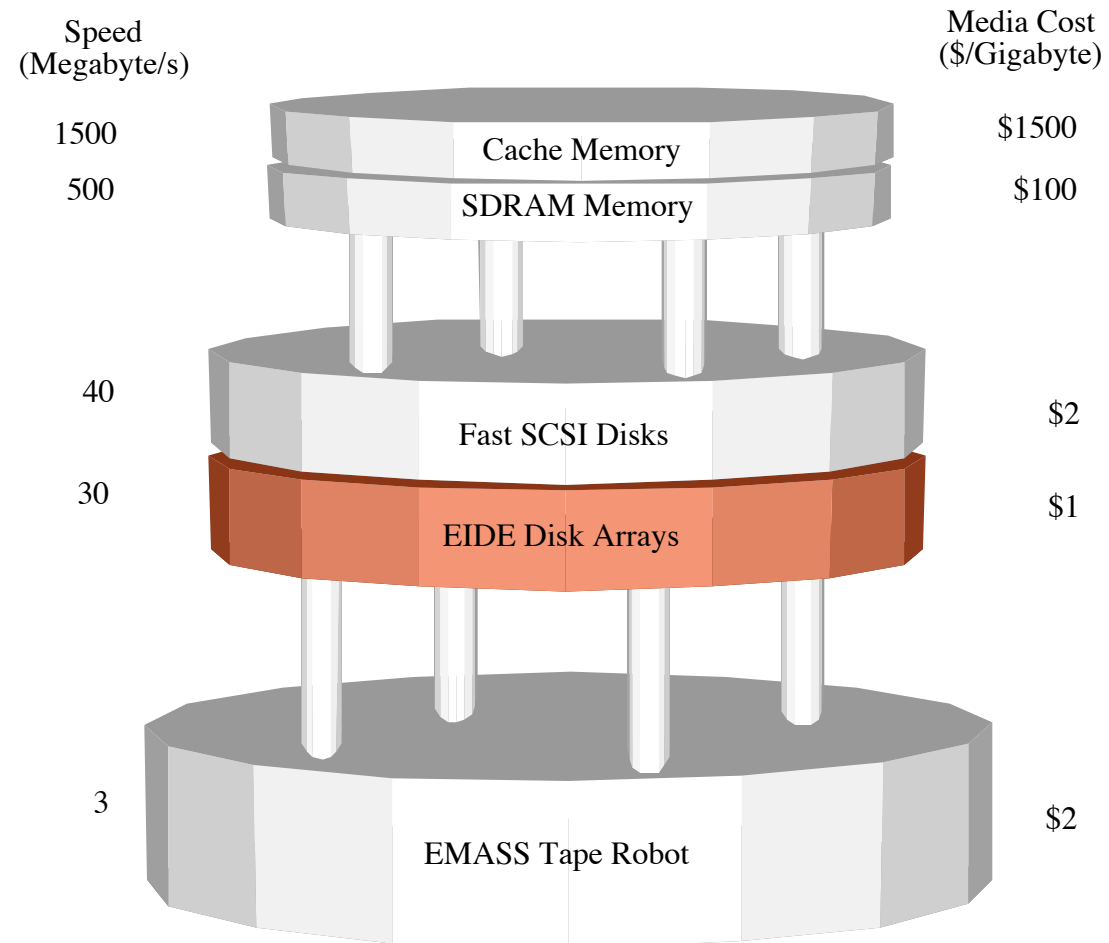


# Introduction

- \$2000 per Terabyte Storage is Available
- Scalable for use at both Small and Large Institutions — From 1 TB to 250 TB, the same as a \$ Million tape silo.
- Fast Access to Data
- Redundant — RAID5
- Commodity Hardware



# Data Storage Cake



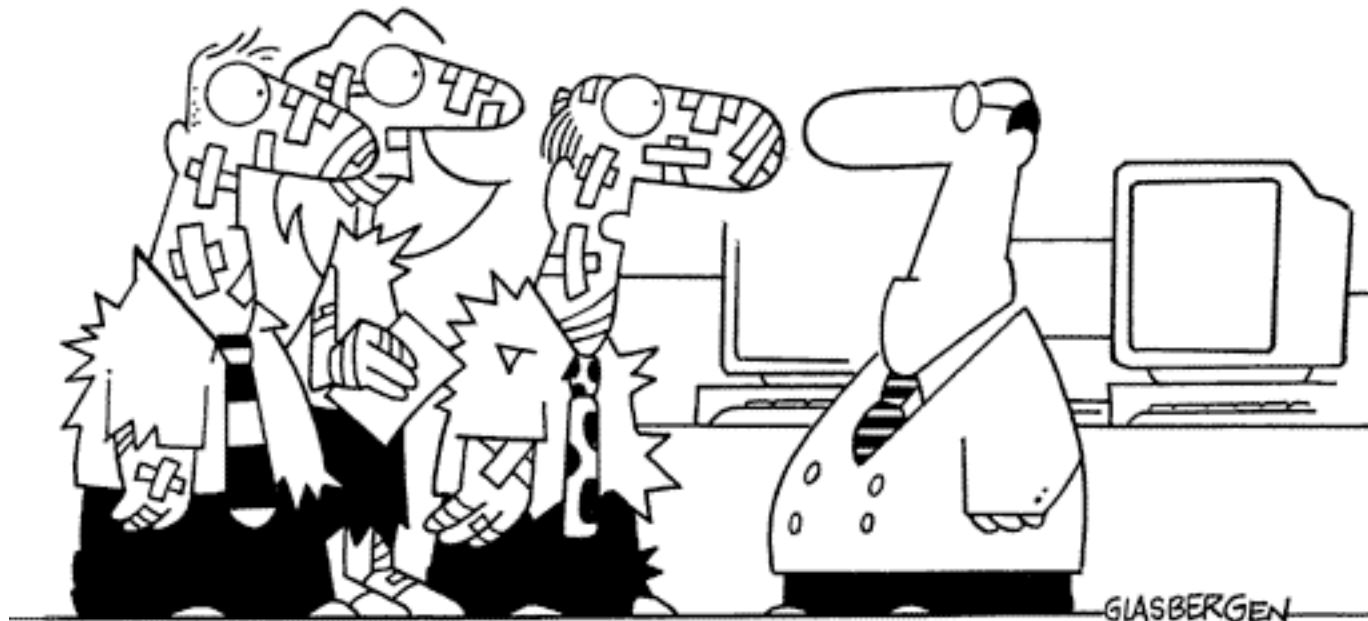
# Definitions

- RAID — Redundant Array of Inexpensive Disks
- RAID level 0 — Concatenation
- RAID level 1 — Mirroring
- RAID level 4 — Parity
- RAID level 5 — Striped-Parity
- EIDE — Enhanced Integrated Drive Electronics



# Why Use Commodity Hardware?

© 1998 Randy Glasbergen. E-mail: [randyg@norwich.net](mailto:randyg@norwich.net) [www.norwich.net/~randyg/toon.html](http://www.norwich.net/~randyg/toon.html)



**“Frankly sir, we’re tired of being  
on the cutting edge of technology.”**

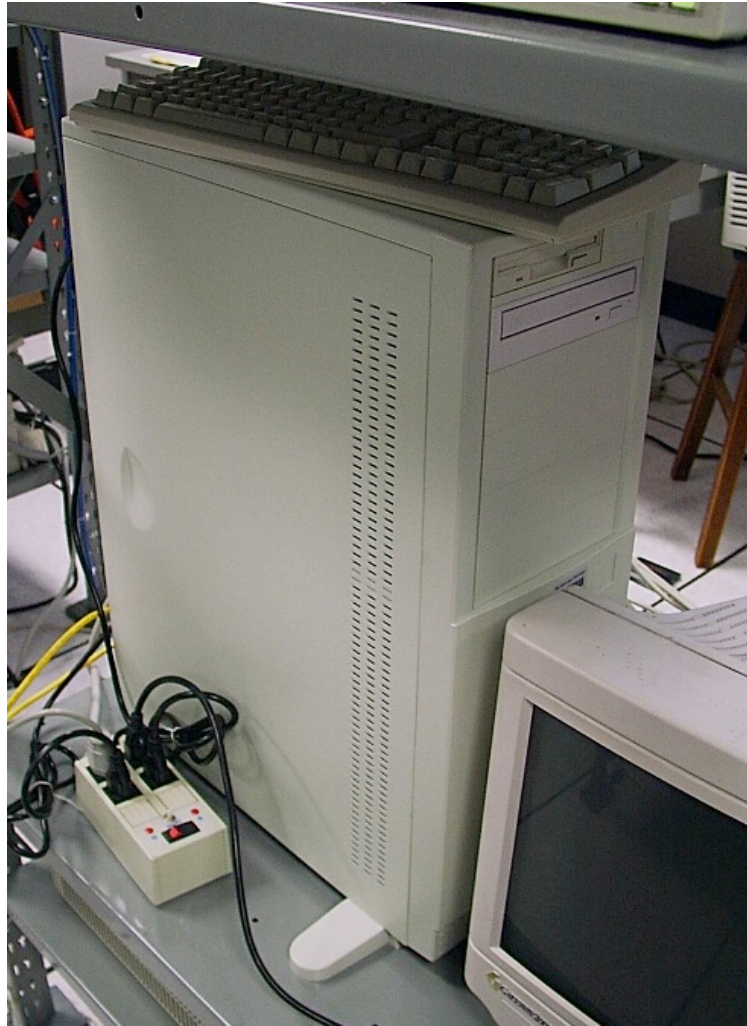


# Hardware

- System Disk — 40 GB Maxtor
- Eight 160 GB Maxtor Disks
- 2 Promise Ultra133 PCI cards
- 24” EIDE Cables
- CPU — 1.6 GHz AMD Athlon
- Motherboard — Asus A7M266
- 512 MB DDR memory
- Second Power Supply (15A at 12V)

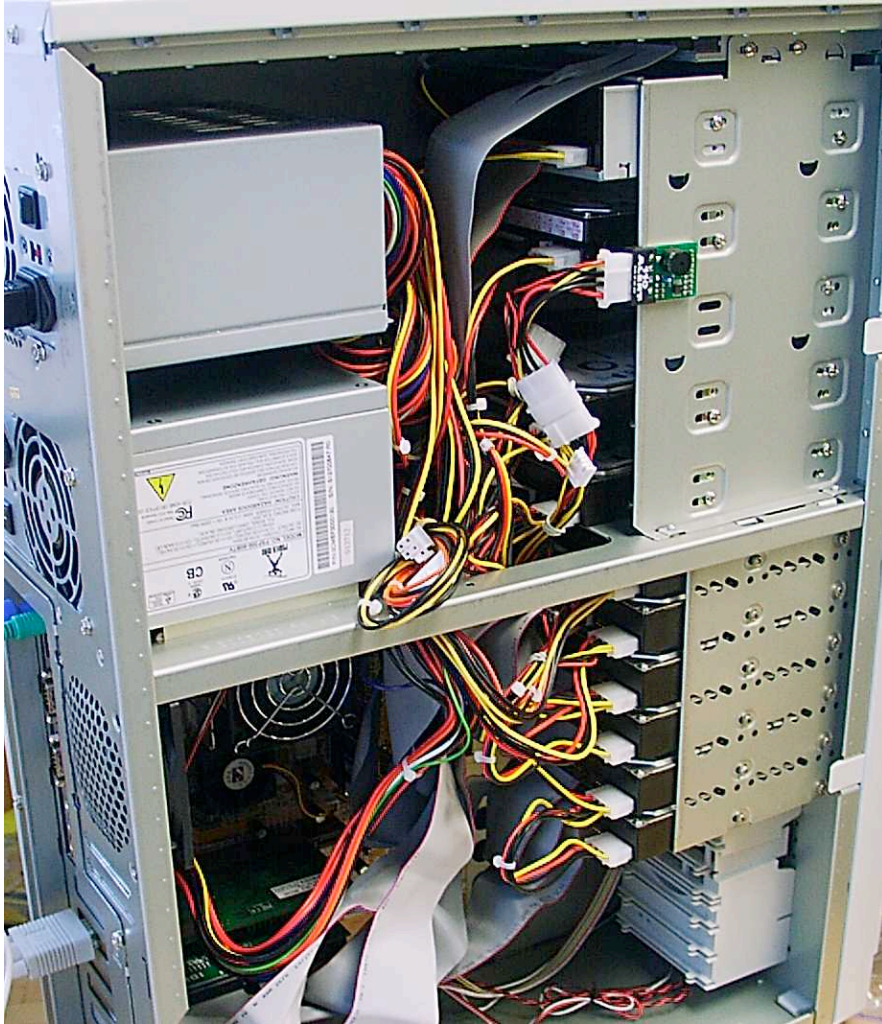


# RAID5 Box for BABAR





# RAID5 Box for CMS





# Disks

Disk	RPM	\$/GB	GB/platter	Amps@12V	Warranty
160 GB Maxtor	5400	1.03	40	1.8	1 Yr.
250 GB Maxtor	5400	1.28	80	~1.1	1 Yr.
200 GB Western Digital (8MB cache)	7200	1.20	66	1.3	3 Yrs.
180 GB Hitachi/IBM	7200	1.34	60	2.0	3 Yrs.



# Software

- Linux 2.4.17 Kernel (with >137 GB patch)  
(the latest stable kernel is the 2.4.20)
- raidtools available with most distributions
- Journaling File systems (ext3)
- NFS to mount on other computers  
(Linux, Sun Solaris, DEC Ultrix, Mac OSX)
- HDPRM speed test (~95 MB/s)
- Simple write test (95 MB/s)

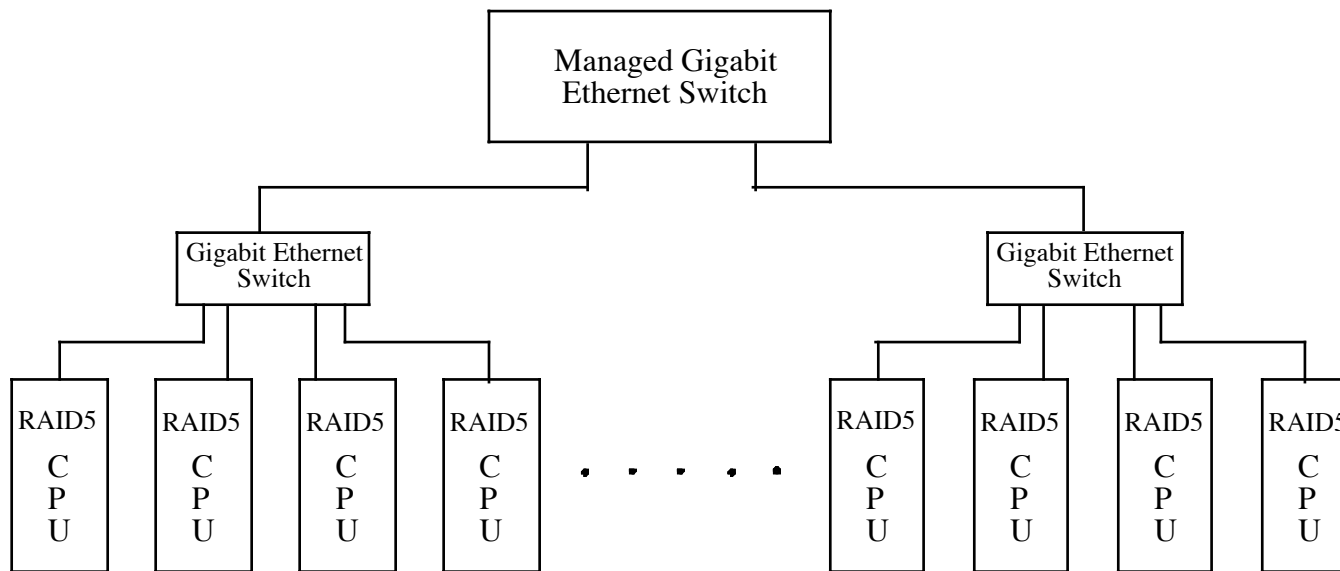


# High Energy Physics Data Analysis Strategy

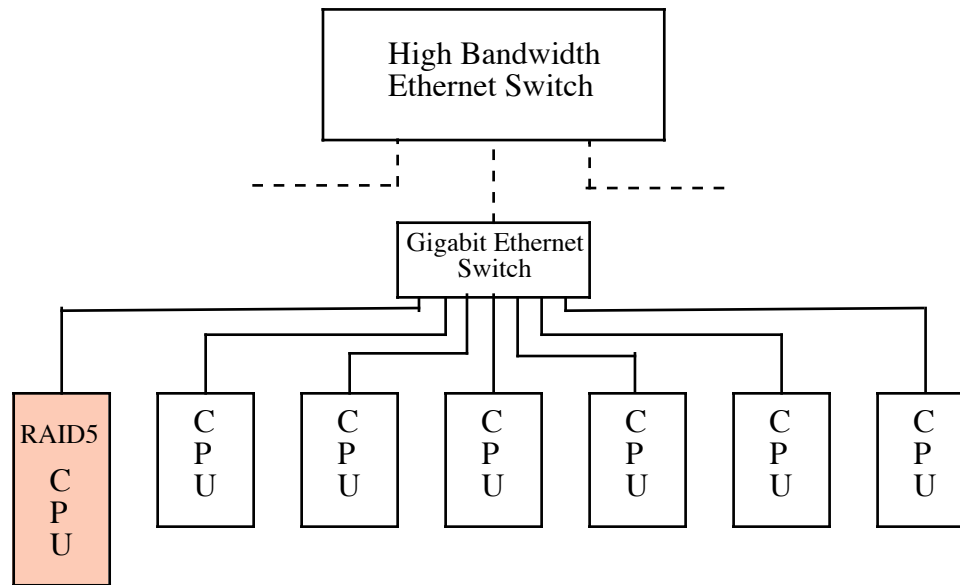
- Use Parallel Processing
- Split data and store on many RAID5 PCs
- Analysis for a subset of data takes place locally on the PC where the data resides
- Network is only used to combine results
- Or use NFS to mount RAID5 array on many PCs (Less efficient due to network overhead)



# High Energy Physics Cluster



# NFS Mounted Cluster





# Future Plans

## Hardware:

- System Disk — 40 GB IBM
- Twelve 250 GB Maxtor Disks
- 3 Promise Ultra133 PCI cards
- CPU — Dual 2.0 GHz AMD Athlon
- Motherboard — Asus A7M266D
- Gigabit Ethernet Card
- Second Power Supply (15A at 12V)

## Software:

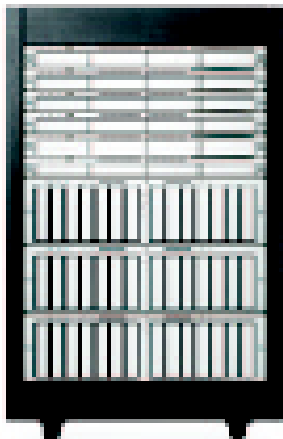
- Try Stock Linux Kernel 2.4.20
- Test other Journaling File systems (ReiserFS, xfs)



# Commercial Systems

	Capacity	Size	Price*	Price/GB
Apple Xserve RAID	2.52 TB	3U	\$10,999	\$4.36
Dell EMC CX200	2.2 TB	3U	\$30,000	\$13.63
HP 7100	2.2 TB	Two 3U	\$109,968	\$50.21
IBM ProFibre DF4000R	2.2 TB	Two 3U	\$43,974	\$20.08
Sun StorEdge T3	2.64 TB	Three 3.5U	\$144,300	\$54.66

\*Based on suggested retail prices on February 7, 2003  
From Apple document L26325A\_XserveRAID\_TO.pdf



# Summary

- \$2000 per Terabyte RAID5 arrays of EIDE Drives tested, without tape backup.
- They are Scalable — Cost less/TB than a tape silo, but scalable down to 1 TB.
- Uses Commodity Hardware.
- Tested with 160 GB hard disks

Supported by the U.S. Department of Energy under DE-FG05-91ER40622 and DE-AC02-76CH03000.

