Redundant Arrays
of IDE Drives

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

University of Mississippi

Donald L. Petravick

Fermilab
Introduction

• $4000 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

- **Speed (Megabyte/s)**
  - 1500
  - 500
  - 40
  - 30
  - 3

- **Media Cost ($/Gigabyte)**
  - Cache Memory: $1500
  - SDRAM Memory: $130
  - Fast SCSI Disks: $6
  - EIDE Disk Arrays: $2
  - EMASS Tape Robot: $2
Definitions

- RAID — Redundant Array of Inexpensive Disks
- RAID level 0 — Concatination
- RAID level 1 — Mirroring
- RAID level 4 — Parity
- RAID level 5 — Striped-Parity
- EIDE — Enhanced Integrated Drive Electronics
Why Use Commodity Hardware?

"Frankly sir, we’re tired of being on the cutting edge of technology."
Hardware

• System Disk — 100 GB Maxtor
• Eight 100 GB Maxtor Disks
• 2 Promise Ultra100 PCI cards
• 24” EIDE Cables
• CPU — 1.4 GHz AMD Athlon
• Motherboard — Asus A7A266
• 512 MB DDR memory
• Second Power Supply (15A at 12V)
RAID5 Box for BABAR
## Disks

<table>
<thead>
<tr>
<th>Disk</th>
<th>RPM</th>
<th>$/GB</th>
<th>GB/platter</th>
<th>Amps@12V</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 GB Maxtor</td>
<td>5400</td>
<td>1.88</td>
<td>20</td>
<td>2.000</td>
</tr>
<tr>
<td>100 GB Maxtor</td>
<td>5400</td>
<td>2.20</td>
<td>33</td>
<td>0.640</td>
</tr>
<tr>
<td>160 GB Maxtor</td>
<td>5400</td>
<td>2.05</td>
<td>40</td>
<td>1.800</td>
</tr>
<tr>
<td>75 GB IBM</td>
<td>7200</td>
<td>3.00</td>
<td>15</td>
<td>2.000</td>
</tr>
</tbody>
</table>
Software

- Linux 2.4.13 Kernel (2.4.5 Tested)
- raidtools available with most distributions
- Journaling File systems (ReiserFS and ext3)
- NFS to mount on other computers (Linux, Sun Solaris, DEC Ultrix, Mac OSX)
- HDPARM speed test (~28 MB/s)
- Simple write test (22 MB/s)
FireWire™

- 63 Disks per chain
- 10 Terabytes per chain using 160 GB disks
- Hot swapable (Using Firewire™ Hubs)
- We tested RAID5 in Linux (Oxford 911 chipset)
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

High Bandwidth Ethernet Switch

Simple Ethernet Switch
- RAID5 CPU
- RAID5 CPU
- RAID5 CPU
- RAID5 CPU

Simple Ethernet Switch
- RAID5 CPU
- RAID5 CPU
- RAID5 CPU
- RAID5 CPU
Moving Data to Institutions

- Internet transfers

- Hot pluggable EIDE disks in Firewire cases

- DVD-R disks — 4.7 GB DVD-R disks are now $6, the DVD writer is only $500, and the DVD-ROM reader is $60.
Summary

• $4000 per Terabyte RAID5 arrays of EIDE Drives tested, without tape backup.
• They are Scalable —Same cost/TB as a tape silo, but scalable down to 1 TB.
• Uses Commodity Hardware.
• FireWire has been tested in Linux

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