

---

# HPSS User Forum - Site Report

**Bill Weeks**

**Stanford Linear Accelerator Center**

Produced under contract DE-AC03-76SF00515 between Stanford University and the Department of Energy



# BaBar & The B-Factory

---

- Use big-bang energies to create B meson particles
  - ◆ Look at collision decay products
  - ◆ Answer the question “where did all the anti-matter go?”
- 600 physicists collaborating from >75 sites in 10 countries
  - ◆ USA, Canada, China, France, Germany, Italy, Norway, Russia, UK, Taiwan
- The experiment produces large quantities of data
  - ◆ 200 - 400 TBs/year for 10 years
  - ◆ Data analysis converting from Objectivity to ROOT
- Heavy computational load
  - ◆ ~202,000+ SpecInt95's available in compute farm
  - ◆ 2201 compute nodes with 3577 cpus



# BaBar HPSS Software Levels

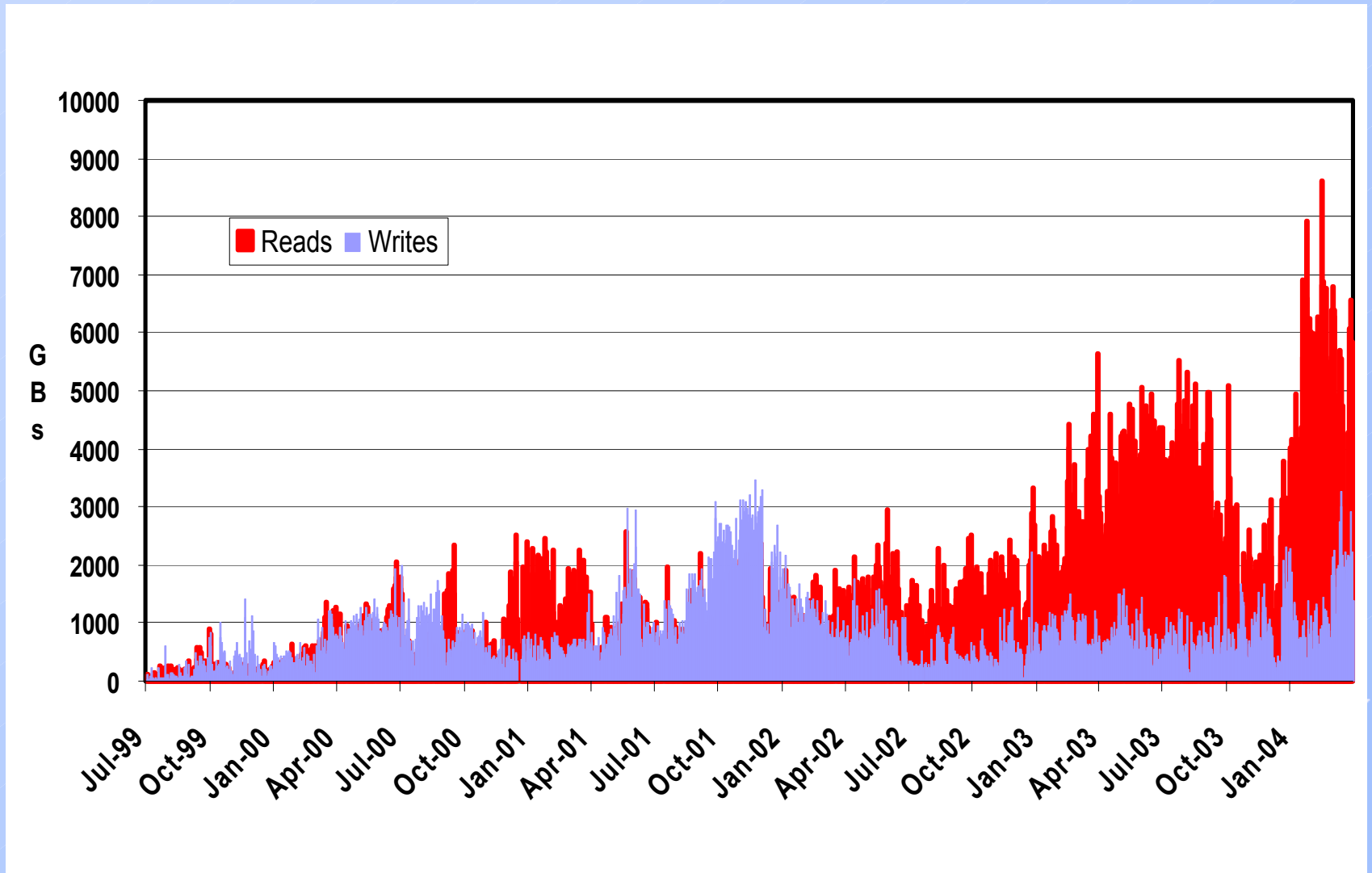
---

<u>Date</u>	<u>Machine</u>	<u>OS</u>	<u>DCE</u>	<u>Encina</u>	<u>HPSS</u>
Jul 98	F40	4.1.5	2.1	2.1	3.2
Sep 99	F50	4.2.1	2.1	2.5	4.1
Jan 01	F50	4.3.3	2.2.0.8	4.2.0.14	4.1.1.4
Aug 03	480R	8	3.2.0.3	5.0	4.5.0.1

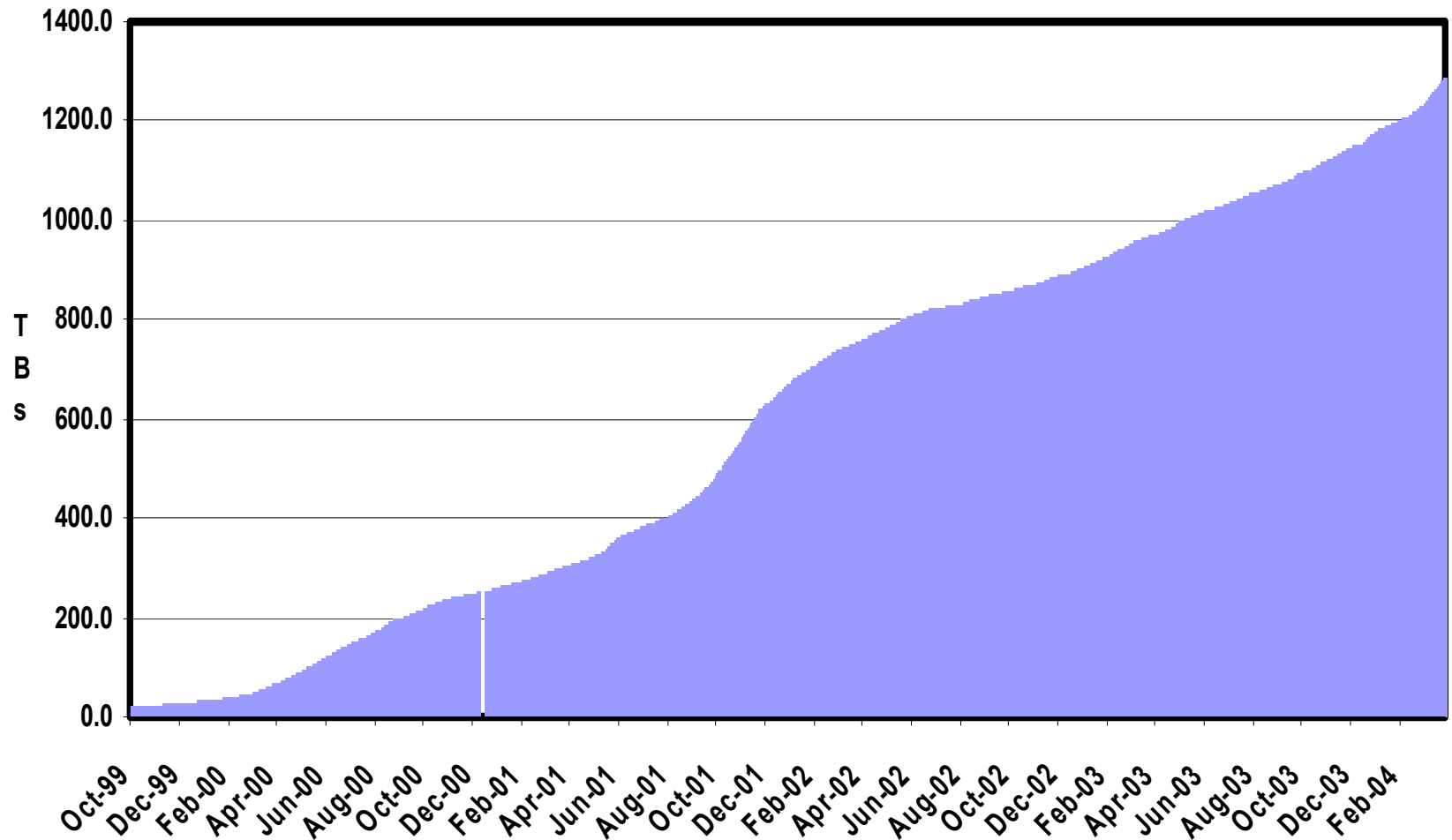
- System converted from AIX to Solaris and upgraded to 4.5.0.1
- BaBar operation makes major software upgrades difficult



# BaBar HPSS Bytes Processed per Day



# BaBar HPSS Total Space Used



# Public HPSS Software Levels

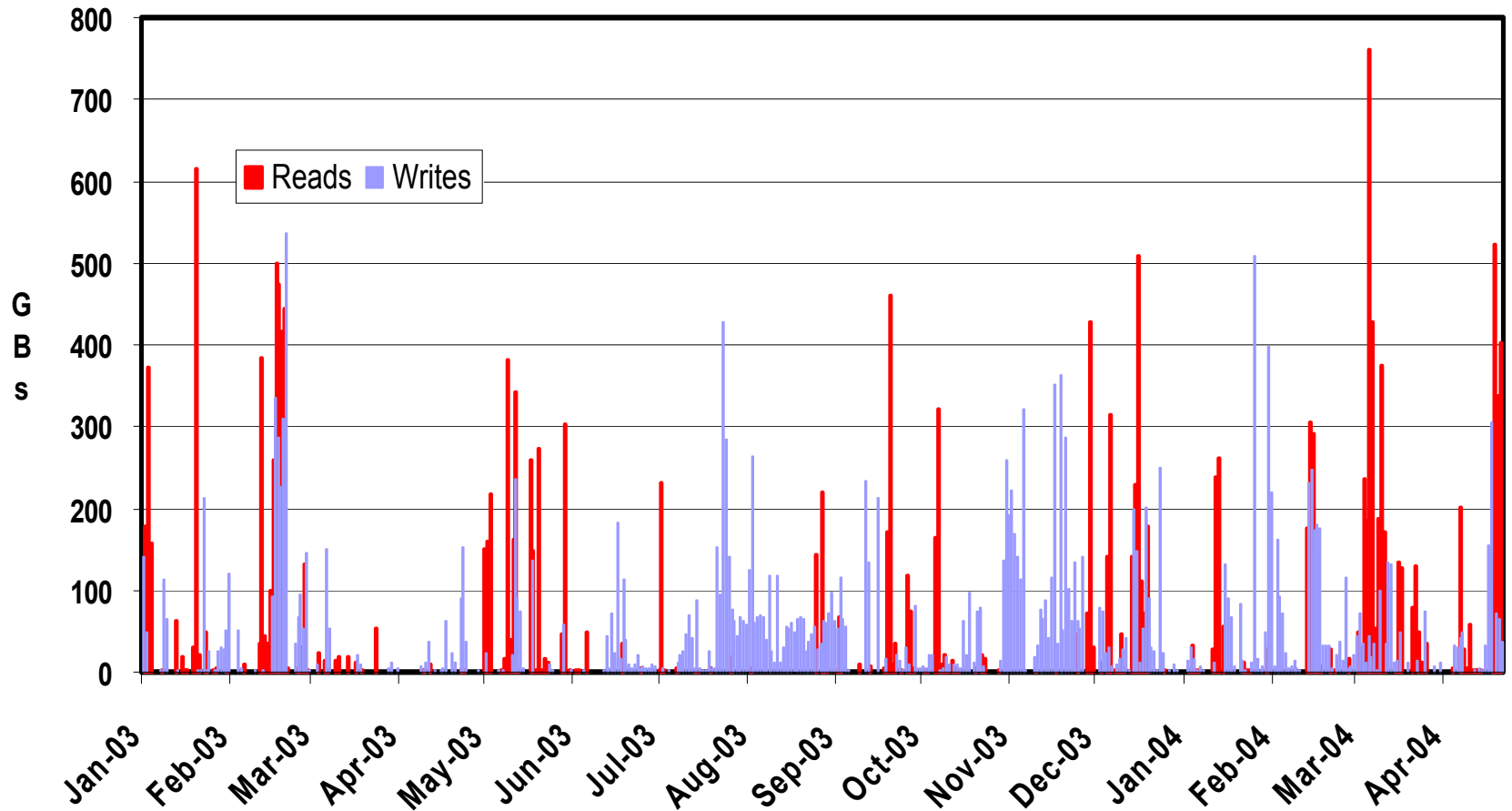
---

<u>Date</u>	<u>Machine</u>	<u>OS</u>	<u>DCE</u>	<u>Encina</u>	<u>HPSS</u>
May 01	420R	7	3.1	4.3.0.3	4.2
Dec 01	420R	8	3.1	4.3.0.4	4.3
??????	420R	8 or 9	3.2.0.3	5.0	4.5.0.2

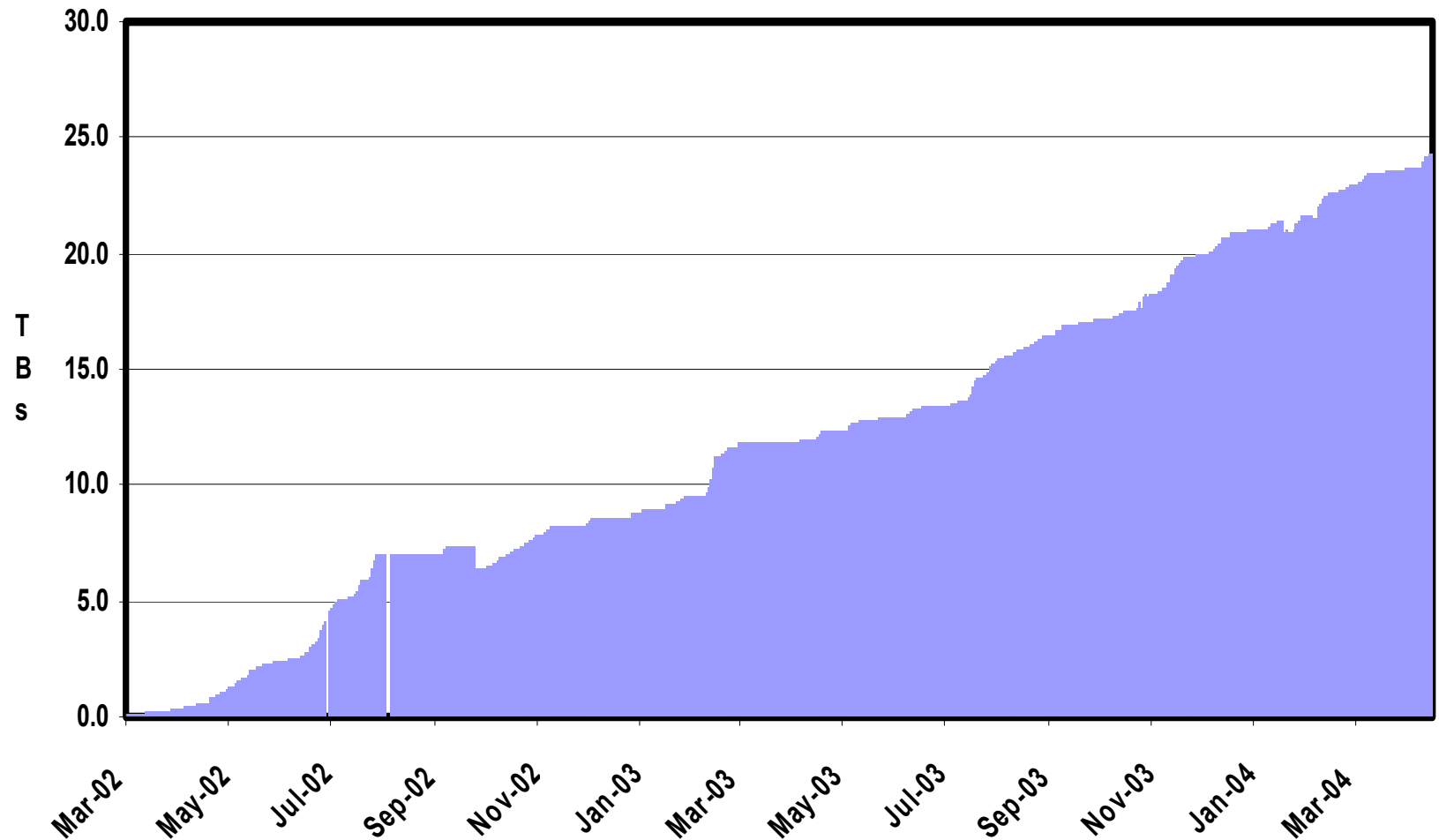
- No changes for 2.5 years
- Modest activity level



# Public HPSS Bytes Processed per Day



# Public HPSS Total Space Used



# BaBar HPSS Reliability

---

## ■ Service outages since June 2003

- ◆ 20 Separate outages, total down time 123 hours
- ◆ 13 Unscheduled outages (90 hours)
  - ◆ 7 HPSS software problems (45 hours)
  - ◆ 4 STK silo problems (39 hours)
  - ◆ 2 Power outages (6 hours)
- ◆ 7 Scheduled outages (33 hours)
  - ◆ 4 Software upgrade (20 hours)
  - ◆ 2 Hardware upgrade (10 hours)
  - ◆ 1 Security patches (3 hours)

## ■ System available 98.4% of real time



# Conversion from AIX to Solaris, HPSS 4.1.1.4 to 4.5

---

- Built HPSS system on Solaris machine
  - ◆ Installed DCE, SFS, HPSS 4.5
  - ◆ Ran mkhpss except for creating SFS files
  - ◆ Copy CDS security objects from old system
    - ◆ Thanks to Mike Gleicher for scripts to do this
- Copy SFS files from old system
  - ◆ sfsadmin export file
  - ◆ sfsadmin import file
  - ◆ Copy ACLs for SFS files



## Conversion from AIX to Solaris, HPSS 4.1.1.4 to 4.5 (2)

---

- Convert SFS files from 4.1.1.4 to 4.5
  - ◆ hpss\_convert\_4v3
  - ◆ hpss\_convert\_4v5
- Bring up SSM and modify configuration using hpssadm
  - ◆ Change hostname
  - ◆ Convert movers to non-DCE
    - ✦ Delete startup daemons and log clients
    - ✦ Change mover's Execute Hostname, TCP Path Name, Encryption Key
  - ◆ Make a few changes with GUI that hpssadm cannot do
    - ✦ Set root Name Server in global configuration
    - ✦ Set default COS in storage subsystem configuration
- Propagate encryption keys to movers



## Conversion from AIX to Solaris, HPSS 4.1.1.4 to 4.5 (3)

---

- Run in read-only mode for first day
  - ◆ Use wrapper for hpss\_pftpd
  - ◆ Modify hpss\_pftpd to support “-r” read-only flag
  - ◆ Wrapper adds “-r” flag based on existence of a file



# Conversion conclusions

---

- Advantages of creating new system on separate hardware
  - ◆ Old system never modified, easy to fallback if necessary
  - ◆ Able to do several dry runs
- Using non-DCE movers simplified conversion
  - ◆ Movers could be used by both old and new systems
  - ◆ No issues with connecting mover to new DCE cell
- Actual conversion completed in ~4 hours
- hpss\_pftpd wrapper allowed production read-only testing



## Conversion of 9940A drives to 9940B

---

- Swap out all 30 9940A drives and add 10 additional 9940B
- 40 tape drives attached to 20 movers
- Use hpssadm scripts to automate conversion
  - ◆ Delete 9940A tapes
  - ◆ Add new non-DCE movers
  - ◆ Add 9940B tapes
    - ✦ Actually 9940 since hpssadm did not support 9940B
  - ◆ Propagate encryption keys to movers
  - ◆ Delete old movers
- Export and re-label all 9940A free tapes



## 9940B read compatibility problem

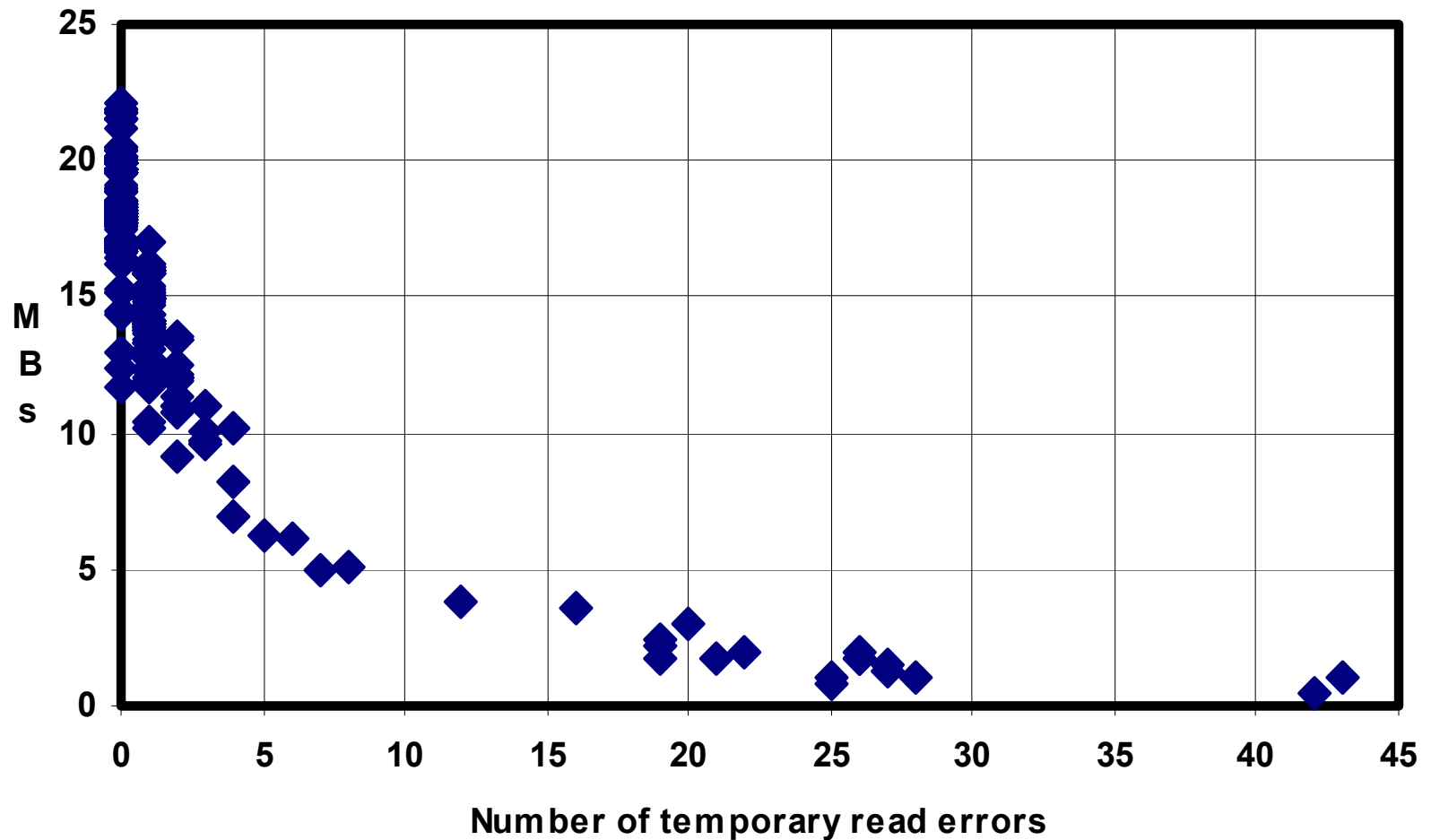
---

- Permanent read errors on 9940B drives reading 9940A tapes
- Problem mostly caused by one 9940A drive
  - ◆ Problem drive in service for 10 months
  - ◆ 2363 tapes written during 10 month period
- Add 9940A drives controlled by second PVR

When written	Tapes mounted	Perm Read Errs	Error rate
Before	437	5	1.14%
10 Month	1312	65	4.95%
After	420	4	0.95%



# Effect of temporary read errors on 9940B performance



# Monthly average file transfer performance

Month	KB/s Read	TB	KB/s Write	TB
05/03	4657	103	6914	31
06/03	3179	121	5633	24
07/03	3195	121	5135	20
08/03	3802	119	5498	21
09/03	3831	115	6270	20
10/03	5488	76	8752	28
11/03	6869	51	11163	27
12/03	6868	53	12529	27
01/04	5727	94	12890	38
02/04	5024	166	11477	26
03/04	6424	169	12498	43
04/04	7980	116	13659	71

9  
9  
4  
0  
A

9  
9  
4  
0  
B



# Media conversion from 9940A to 9940B

---

- Alleviate read compatibility problem
- Reduce media costs
- Automation difficult because lack of utilities
  - ◆ Move Cartridge (GUI only)
  - ◆ repack
  - ◆ retire
  - ◆ remove
  - ◆ Import Volumes (GUI only)
  - ◆ Create SS Resources (GUI only)



# Tape Statistics

---

- 1300TB stored on 12712 tapes
- Tapes are 95% full

Tape type	TB stored	# of tapes	Percent full
9840	242	5934	93.4%
9940A	527	4685	91.6%
9940B	531	2093	99.3%



# HPSS Problems

---

- PVL ss\_MountCallback hangs (4.1.1.4)
  - ◆ Increase connection and thread count for Storage Server
  - ◆ Increase PVL binding table size from 32 to 150
- hpssadm problems
  - ◆ Leading 0's suppressed for mover encryption key
    - ✦ Fixed by PTR 3473
  - ◆ Does not support 9940B tape drive
    - ✦ Fixed by PTR 3533



## HPSS Problems (2)

---

- Tape mounting strategy when system over-committed
  - ◆ Problems with drive selection, unnecessary passthru operations
    - ✦ Past two months 102,000 of 230,000 mounts = 44% passthru
  - ◆ Tapes on deferred dismount queue not dismounted first



# HPSS Wish List

---

- Utilities to manage media migration
- Error statistics in mover
- Improvements in tape mounting strategy
- Dynamic device reconfiguration
  - ◆ PVL, PVR outages too disruptive



# Summary

---

- HPSS is reliable
- HPSS tape handling is a concern
- Software upgrades will always be a problem

