

Castor Monitoring

Castor is:

- **Nameserver(s) (3 in load balanced service + 1 DB)**
- **Diskservers ($\sim 10^2$)**
- **Stagers (~ 50 – up from ~ 5 when monitoring started)**
- **Tape Robots (4)**
- **Tapeservers ($\sim 10^2$)**
- **Tapedrives ($\sim 10^2$)**
- **Tapes ($\sim 3 \cdot 10^4$)**

All of which can go wrong and some of which can be easily abused/overwhelmed by users.

User abuse

- Almost always occurs because internally Castor is not protected against 'unreasonable' behavior.
- Because Castor appears to have infinite capacity – no feedback from the system to users if they do silly things (and a lot of requests come from batch anyway)
- Lack of user information/education
- Because they can/because it's there

(Peer pressure is one effective way to control user excesses '(Please) don't do that again'
cc: your_experiment.support@cern.ch)

User abuse

- **Nameserver: (nsls of large directories or nslstape of many tapes) (up to $\sim 2 \cdot 10^5$ files in a directory, up to $1.4 \cdot 10^5$ files on one tape..)**
- **Too many stager requests: (no scheduling, all requests started until the stager grinds to a halt)**
- **Too many stager requests/second: requests submitted $> \sim 5$ /second will reliably crash a stager**
- **Too many rfio requests to same filesystems/diskserver (no scheduling..) can kill a diskserver (very high load)**
- **Too many tape requests (read) can cause vdqm meltdown (solved by use of multfile stageing)**
- **Too many files in a stager catalogue can cause staggers to grind to a halt**
- **Too many files can cause migrations to last days..(blocking TapeDrives) average Castor file size ~ 105 MB; many < 10 KB.**

Reasons for monitoring (1)

- **First and foremost: to understand the beast (which came with little or no documentation - though many man pages). In the early days 'nasties' lurked in most nooks and crannies.. A fair number bugs were detected..**
- **To ensure similar objects (TapeServers, DiskServers, Stagers, FileClasses, TapePools etc) are correctly configured**
- **To add additional functionality (e.g. control of stage catalogues, prevention of needless tape mounting of problematic tapes)**
- **To impose useful conventions (i.e. run a 'tighter ship' than Castor itself imposed)**
- **To record (administrative actions) who did what,why and when. e.g. change of a Tape status.**

Reasons for monitoring (2)

- To provide a hierarchical overview of the highly distributed system involving hundreds of boxes, hundreds of Castor objects (FileClasses, TapePools) and tens of thousands of tape volumes.
- To ensure the correct running of ~50 stagers:
 - compaction of Cdb when file > ~.5GB (gain~factor 5 in performance for catalogue changes such as file delete)
 - Adequate staged files lifetimes in disk pools
 - To 'clean up' after Castor (infamous PUT_FAILED files)
- To detect user abuse (and take action).
- To understand the long term evolution of resource usage by the experiments (TapePools, diskserver filespace, stagers)

Reasons for monitoring (3)

- To enable power users (admins) in the experiments to understand their data management resource usage.
- To enable end-users to understand why (perhaps) they cannot access their data (Tape problems, DiskServer problems).
- To inform end-users of 'their (user)' errors (e.g. cannot use a FileClass/TapePool, or write to a specific Castor directory)

Basics

- Put it all on the web (comparing 200 objects is very simple – anomalies highlighted by colour coding)
- Warn of serious issues by email (Castor operations, Tape operations, Computer operations)
- Web pages use a drill down approach. One page shows a summary of the entire system with colour codes showing possible problems at lower levels.
- In *conjunction with* Lemon monitoring. Warn FIO admins of serious individual machine state transitions - daemons dead, no_contact - for sub-groups of a large number of machines ~850. Even though sysadmins may fix many problems we need to know that they occur.

Monitoring methods and tools (1)

- **Analysis model is (mostly) centralized analysis accessing distributed data via rfc.**
- **Perl scripts gather data via cron/acrontabs with varying sampling periods – from 5 minutes to 2 hours – depending on resource consumption; results glued together on the web.**
- **Batch jobs once a week to check ALL files in ALL stager diskpools (-> orphaned files)**
- **Make use of existing Castor accounting data (stager stats, RTCOPY errors, TapeMounting stats)**

Monitoring aids and tools (2)

- Enable any new problem to be quickly added to list of monitored issues (shutting the barn door after... is useful)
- Parse log files (incrementally) – errors tolerated up to predefined limits. Also used for extraction of performance statistics.
- Use existing Castor tools:
 - stageqry --format x,y,z.... considered v. powerful!
 - showqueues (Tape queue analysis, TapeServer and TapeDrives status)
 - vmgrlistpool (TapePool evolution, media management)
- Deduce syntactic rules for different *types* and *sets* of configuration files and regularly scan for inconsistencies.

Top level page (mother..)

The screenshot shows a Microsoft Internet Explorer browser window displaying the 'CASTOR Mother of all Status displays' page. The page title is 'CASTOR Mother of all Status displays' and the address bar shows 'http://castor.web.cern.ch/castor/MONITOR/DIR_SURVEY/moab.html'. The page content includes a legend for error status (green for 'No errors', yellow for 'Relatively minor problem(s)', orange for 'More serious level of problems', and red for 'Serious trouble(s)') and several sections of status reports. Annotations in white ovals with black text and arrows point to specific items: 'All stagers ok' points to the 'Stager Survey' link; 'Non-uniform configuration of castorgrid machines' points to the 'Castorgrid grid-mapfile and stagemap.conf survey' link; 'Some trivial RTCOPY errors' points to the 'RTCOPY errors, drive and volume error DB, Tape repairs DB, Tape/Segment status changes' link; and 'Too many DISABLED tapes' points to the 'Fileclass, Tape and TapePool Information - including TapePool refilling and DISABLED tapes' link.

CASTOR Mother of all Status displays - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://castor.web.cern.ch/castor/MONITOR/DIR_SURVEY/moab.html

Google rome tennis final

CASTOR Mother of all Status displays

Last refresh: Sun May 14 13:06:02 CEST 2006

Legend: ■ No errors ■ Relatively minor problem(s).. ■ More serious level of problems ■ Serious trouble(s)... You really should follow the link to see what is wrong..

High Volume Central Data Recording (CDR)

- [2004: compass_na48 and ntof](#)
- [2005: ntof](#)

Key Castor Components

- [Castor NameServer and Database checks](#) (Sun May 14 13:06:07 CEST 2006) ■
- [VDOM queue: Analysis and Plots of Queues \(per Device Group Name\)](#) (Sun May 14 13:05:26 2006) ■ The size of the vdom queue is: 102. (It was 93 at 13:00:25) ■
- [Stager Survey](#) (Sun May 14 13:01:58 2006) ■
- [TapeServer Survey](#) (Sun May 14 13:01:23 2006) ■ [Tape Drive Dedications/Reservations](#) ■
- [Castorgrid grid-mapfile and stagemap.conf survey](#) (Sun May 14 12:11:05 2006) ■

Basic Infrastructure Checking

- [Context list of ALL FIO/DS Machines \(Lemon/SMS status,Cdb info\)](#)
- [Cdb/Lemon info on ALL FIO/DS Machines \(breakdown by Cdb cluster/functionality/model\)](#) (Sun May 14 13:02:06 2006) ■
- MSA alarm status:**
 - [NameServers](#) (Sun May 14 13:00:42 2006) ■
 - [Stagers](#) (Sun May 14 13:00:56 2006) ■
 - [Production Diskservers \(+ rfdir check/Castor version\)](#) (Sun May 14 13:01:37 2006) ■
 - [Production Service TapeServers with drives](#) (Sun May 14 13:01:23 2006) ■ (2 alarms/problems)
- [SUNstik machines ping survey](#) (Sun May 14 13:03:11 2006) ■
- [rfiod and stager log size scan](#) (Sun May 14 12:30:58 2006) ■

Basic Infrastructure Information

- [RTCOPY errors, drive and volume error DB, Tape repairs DB, Tape/Segment status changes](#) (Sun May 14 13:06:07 CEST 2006) ■
- [Fileclass, Tape and TapePool Information - including TapePool refilling and DISABLED tapes](#) (Sun May 14 13:06:07 CEST 2006) ■
- Tape Mount Information**
 - [Analysis of Tape Mounting data per week](#) (breakdowns by dgn, Experiment, TapeServer, TapePool and VID)
 - [Breakdown by vmgr status and No. of rtcopy errors](#)
 - [Castor Tape mount distributions, most active and least active tapes:# Files/tape volume.](#)
 - [Top 30 most busy \(most mounted\) tapes per week](#)
- Tape Robot Information:**
 - [Silo Inventory comparison with VMGR: Most mounted tape volumes \(since time began..\)](#) (Sun May 14 00:25:21 2006) (currently 33 differences) ■
 - [SL3500_D/sunstk51 error scan](#)
 - [ACSL34/sunstk52 error scan](#)
 - [ACSL35/sunstk62 error scan](#)
- [Useful Documentation](#)

Internet

Tape Queues (vdqm) Source: showqueues

CASTOR showqueues survey

Date of last check: Sun May 14 15:35:25 CEST 2006 See related [\(Tape/Drive/Server Survey\)](#)

Available drives = total drives - (dedicated/in_dev + UNKNOWN/DOWN)

dgn	Free drives (of available)	Running Tape jobs	Total requests in queue	No. different VIDs in queue	No. requests for unmounted VIDs	Ave wait time (unmounted)	Max wait time (unmounted)	Ave wait time (all)	Max wait time (all)
3592A0	4 (of 4)	0	0	0	0	0	0	0	0
3592B1	11 (of 23)	12	0	0	0	0	0	0	0
3592B2	13 (of 13)	0	0	0	0	0	0	0	0
994BR0	2 (of 10)	5	1	1	0	0	0	57	57
994BR4	7 (of 12)	3	1	1	0	0	0	30	30
994BR5	0 (of 21)	15	85	23	16	90	187	253	1191
LTO3R0	0 (of 6)	6	5	5	5	3733	10865	3733	10865
T10KR0	8 (of 24)	16	0	0	0	0	0	0	0
T10KR1	14 (of 14)	0	0	0	0	0	0	0	0

Average time to mount a tape not already mounted

Time to 'remount' an already mounted tape is often >> time to mount a 'new' tape

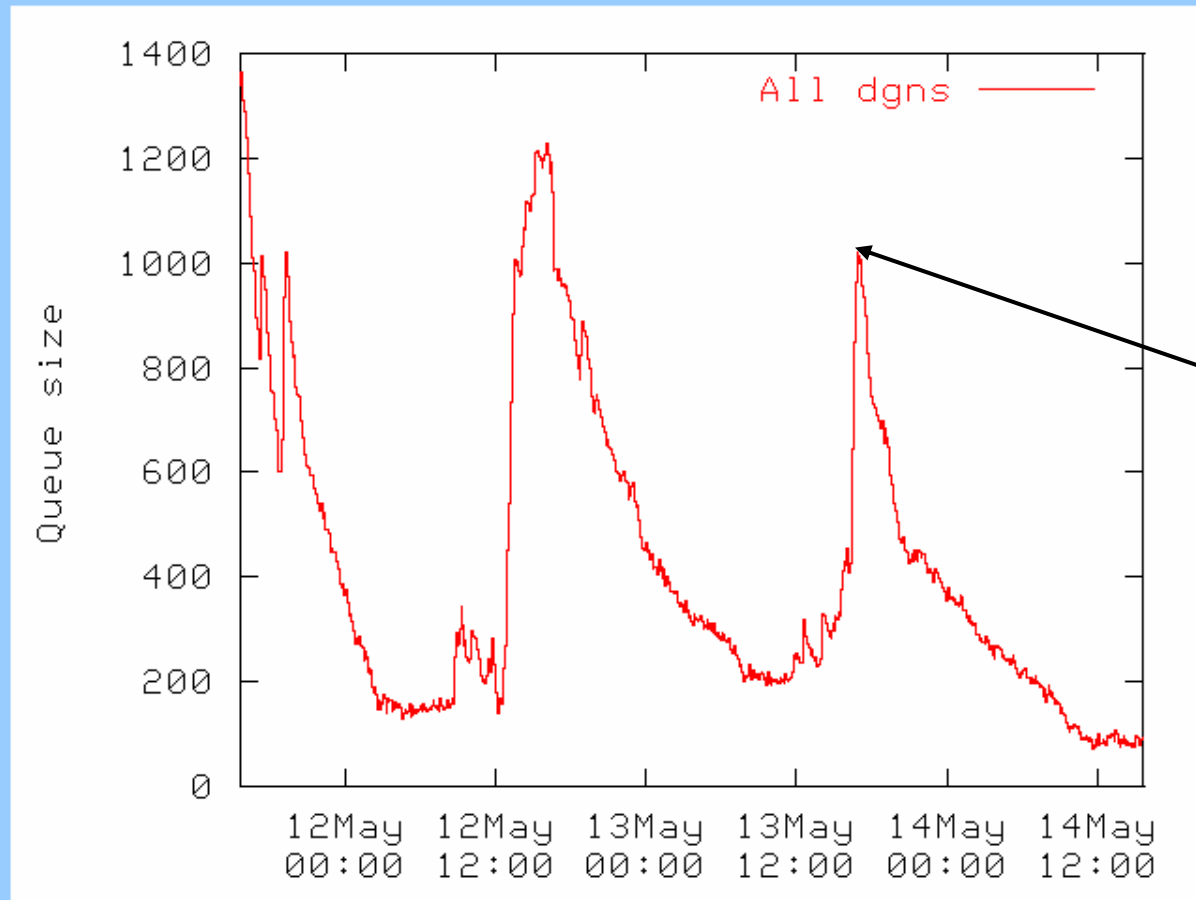
Overall breakdown by user of vdqm queue (top 10 users)

User	Number of queue entries
(hternier,z5)@stage009.cern.ch	39
(rolandc,vp)@stage010.cern.ch	27
(lhcb012,z5)@stage004.cern.ch	11
(stage,st)@castor2srv02.cern.ch	5
(stage,st)@c2atlasrv02.cern.ch	4
(stage,st)@c2cmsrv02.cern.ch	4
(na48mon,v1)@stage005.cern.ch	1
(isolcdr,z6)@stage004.cern.ch	1

userid@stager requesting most tape mounts

Tape Queues (vdqm)

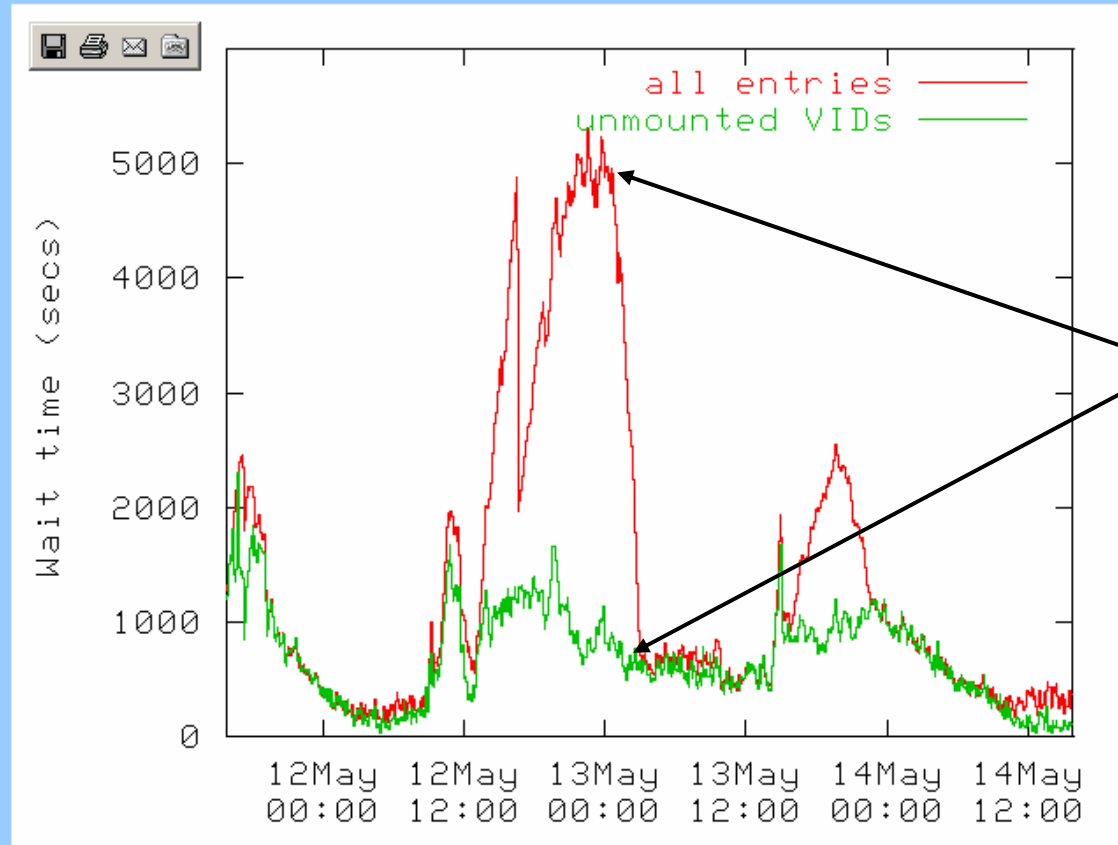
Fig 1. vdqm queue size as function of time for last 3 days; current queue size is 92 (previously it was 90).



Total tape queue must be kept < 5000 or system freezes and all drives must be manually reset (tapes dismounted and daemons restarted) E-mail alert system used (with precise instructions)

Tape activity per DeviceGroupName (dgn)

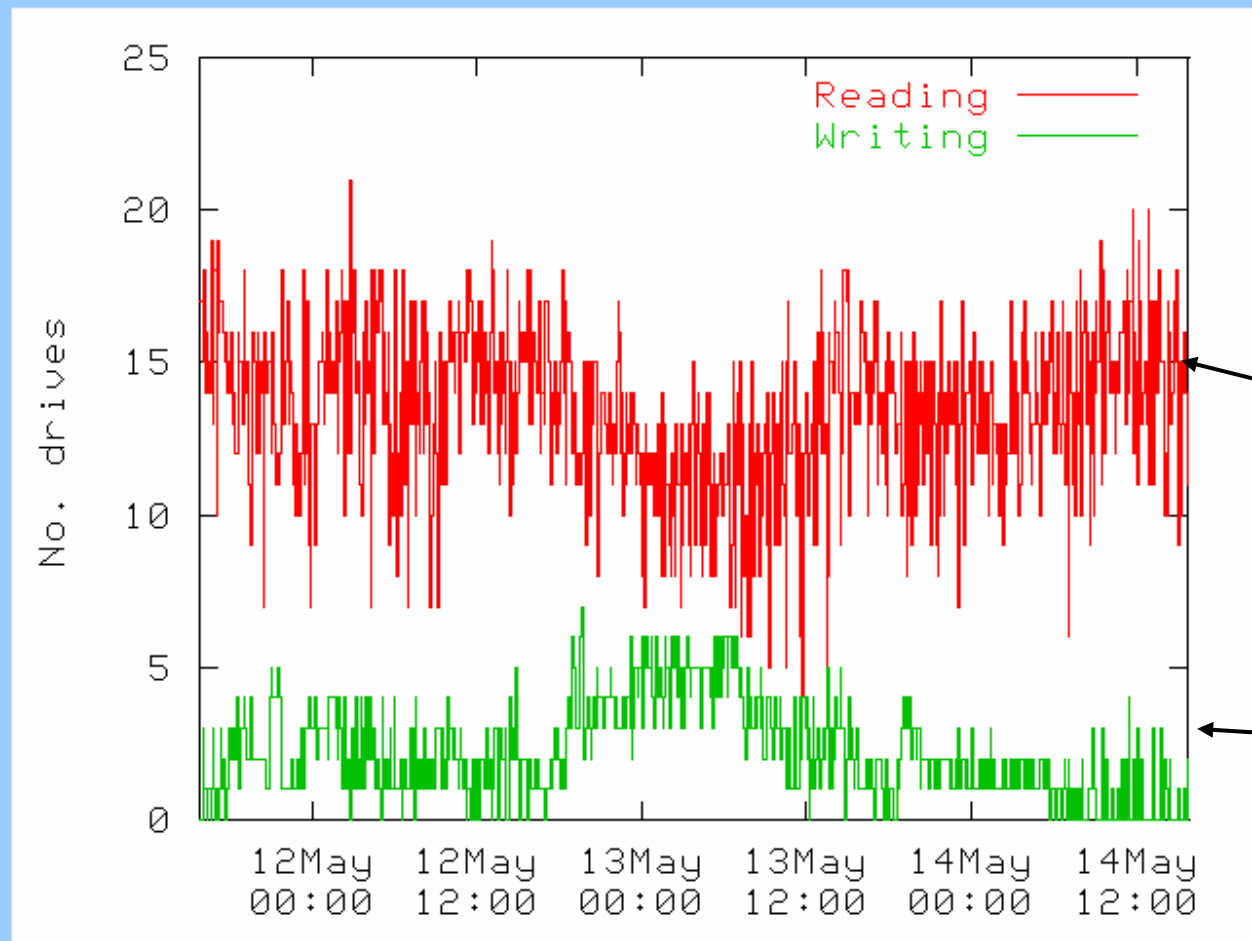
Fig 6. (994BR5) Average queue wait times; all VIDs, unmounted VIDs



Reading through a 200 GB tape at ~25 MB/sec takes a long time (>2 hours). This issue does not go away with newer tape technologies

Tape activity per DeviceGroupName (dgn)

Fig 2. (994BR5) Number of READING/WRITING drives

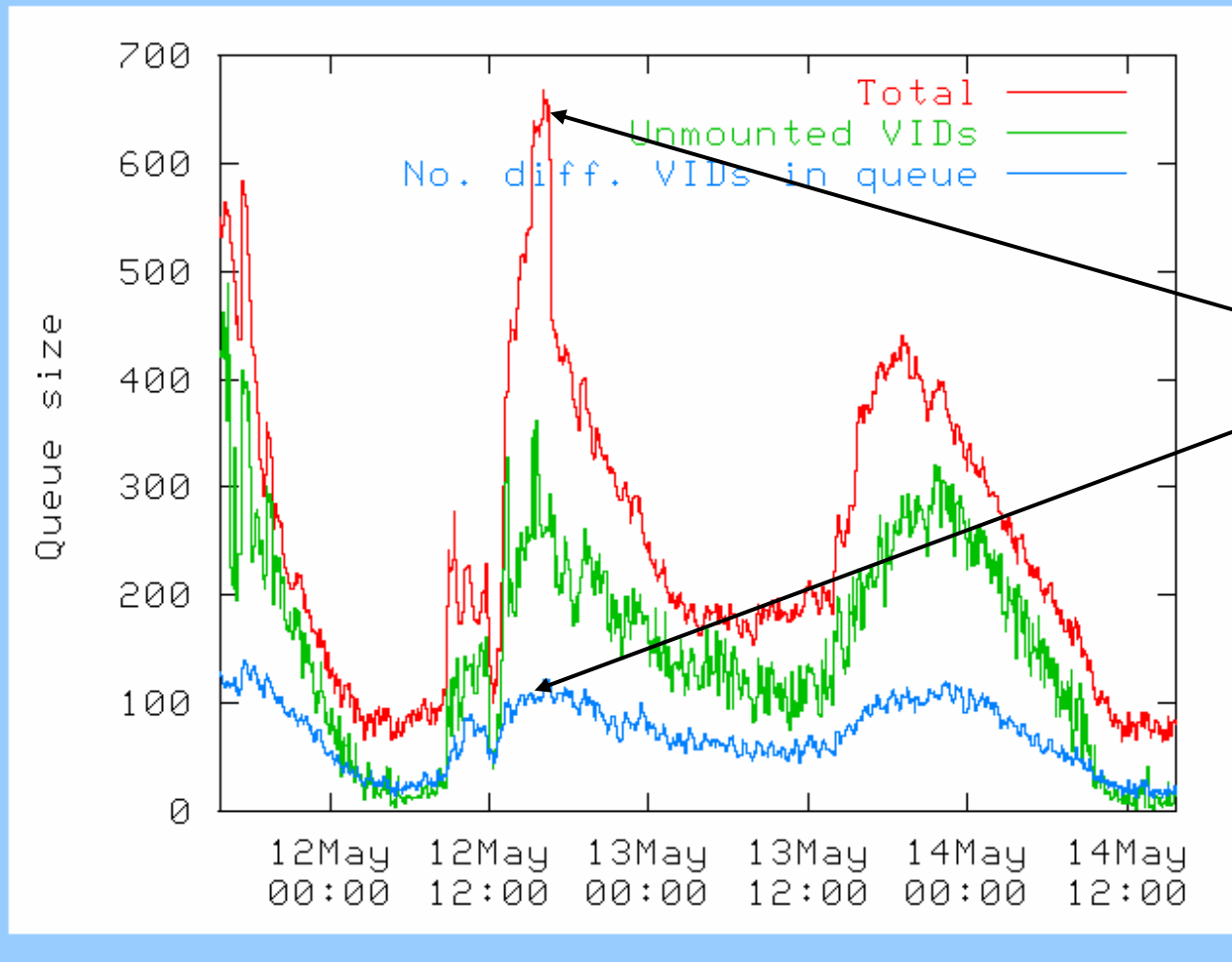


Typically many mounts each to read a few (~2) files - **INEFFICIENT** (not streaming)

Typically few writes to output many files (stager migrations) **EFFICIENT** (streaming)

Tape activity per DeviceGroupName (dgn)

Fig 3. (994BR5) Queue Sizes: Total, unmounted VIDs, No. different VIDs



Each tape requested ~6 times on average

Tape activity per DeviceGroupName (dgn)

Top VID's requested (at most the top 5 are displayed)

Queue analysis for DGN: 994BR5

Number of different VID's requested is 16

VID requests per tapepool

Tape pool	No. different vid requests
lhcb	<u>10</u>
atlas2	<u>1</u>
isolde2	<u>1</u>
default	<u>1</u>
largeuser	<u>1</u>
na49	<u>1</u>
delphi	<u>1</u>
Total	16

Number of different tapes (VIDs) per TapePool

VID	TapePool	No. of queued entries
P08857	na49	26
P12079	lhcb	8
P14397	lhcb	7
P00137	lhcb	6
P00353	lhcb	5

Most solicited tapes (VIDs)

Top TapePool's requested (at most the top 5 are displayed)

Number of different TapePools's requested is 7.

TapePool	No. of queued entries
lhcb	49
na49	26
atlas2	1
isolde2	1
default	1

Most solicited TapePools

TapeDrive status: overview

Breakdown of 994B drives; also see [3592 LTO3 T10K](#)

dgn	Specified in Config file	Seen by showqueues	In Development (and/or Repair)	Reserved (Dedicated)	non-dedicated (available to the public)
994BR0	10	10	0	0	10
994BR4	12	12	0	0	12
994BR5	22	22	1	0	21

Dedication used as an administrative tool - write precedence (over read) used for stager migrations and CDR

994B **non-dedicated** drives by dgn and status (details of device groups: [994BR4](#) [994BR5](#) [994BR0](#))

Status/dgn	994BR0	994BR4	994BR5
FREE	0	9	0
RELEASE	4	3	7
RUNNING	6	0	14
Total	10	12	21
Total AVAILABLE	10	12	21

Overview by drive type, drive status and dgn

Breakdown of 3592 drives; also see [994B LTO3 T10K](#)

dgn	Specified in Config file	Seen by showqueues	In Development (and/or Repair)	Reserved (Dedicated)	non-dedicated (available to the public)
3592A0	3	3	0	0	3
3592B1	25	25	2	0	23
3592B2	14	14	1	0	13

3592 **non-dedicated** drives by dgn and status (details of device groups: [3592B1](#) [3592A0](#) [3592B2](#))

Status/dgn	3592A0	3592B1	3592B2
FREE	3	11	8
RELEASE	0	0	5
RUNNING	0	12	0
Total	3	23	13
Total AVAILABLE	3	23	13

TapeDrive status: dgn 994BR0

CASTOR TapeServer survey

Data refreshed Sun May 14 15:46:04 CEST 2006 .

Details of dgn: 994BR0

NB: Warning E-mails are only generated for tapeservers in SMS state 'production'

MSA OK + background means /home/operator/nomorealarms file exists

Links to recent
RTCOPY errors

Would show tapes
'trapped' on drives that are
DOWN for a long period

Would show requests
that are 'stuck'

Tape Server	prod or dev	MSA alarms (Lemon)	SMS state	Recent rtbody errors in last 4hrs:day:3days	Castor Version	No. 4MB buffers	Mode	Drive [Dedication](Time in State) {VID (used by)}	rtcopy.log Date / size (rflir/activity test)
tprsv110	prod	OK	production	0:0:1	2.0.1.5	160		994B0214 [Dedn: None] (Release for 59 secs) {VID: P11576}	May 14 15:45 (243986 B)
tprsv114	prod	OK	production	0:0:0	2.0.1.5	160	write	994B021C [Dedn: None] (In Use for 91 secs) {VID: P12398, (stage,st)@stage005}	May 14 15:44 (439809 B)
tprsv116	prod	OK	production	0:0:0	2.0.1.5	160	read	994B0218 [Dedn: None] (In Use for 40 secs) {VID: P09176, (rolandc,vp)@stage010}	May 14 15:45 (638641 B)
tprsv117	prod	OK	production	0:0:0	2.0.1.5	160		994B061C [Dedn: None] (Release for 81 secs) {VID: P00270}	May 14 15:44 (426976 B)
tprsv118	prod	OK	production	0:0:0	2.0.1.5	160	write	994B0210 [Dedn: None] (In Use for 1514 secs) {VID: P08928, (stage,st)@stage010}	May 14 15:46 (566388 B)
tprsv119	prod	OK	production	0:1:1	2.0.1.5	160		994B061D [Dedn: None] (Release for 202 secs) {VID: P08580}	May 14 15:42 (394766 B)
tprsv120	prod	OK	production	0:0:0	2.0.1.5	160	read	994B0610 [Dedn: None] (In Use for 5 secs) {VID: P03264, (rolandc,vp)@stage010}	May 14 15:46 (189472 B)
tprsv121	prod	OK	production	0:0:0	2.0.1.5	160	read	994B0614 [Dedn: None] (In Use for 76 secs) {VID: P08582, (stage,st)@c2cmssrv02}	May 14 15:46 (327346 B)
tprsv123	prod	OK	production	0:0:0	2.0.1.5	160	read	994B0618 [Dedn: None] (In Use for 69 secs) {VID: P08374, (stage,st)@c2cmssrv02}	May 14 15:44 (383508 B)
tprsv127	prod	OK	production	0:0:0	2.0.1.5	160		994B021D [Dedn: None] (Release for 111 secs) {VID: P05570}	May 14 15:44 (236751 B)

TapeDrives with too many errors

I/O and Medium ERRORS for last 14 days (from Sun May 14 15:50:39); a total of 25 errors have been observed.

- Click [here](#) to see the 6 errors for previous 1 days.
- Click [here](#) to see the 9 errors for previous 2 days.
- Click [here](#) to see the 14 errors for previous 4 days.
- Click [here](#) to see the 18 errors for previous 7 days.
- Click [here](#) to see the 32 errors for previous 30 days.

The following tape servers are currently not in service. They are thus shown as in the table below:

- **in_dev:** tpsrv055 tpsrv206 tpsrv221 tpsrv227 tpsrv233 tpsrv611 tpsrv640

Frequency of errors per TapeServer; number of different VID's affected on each TapeServer

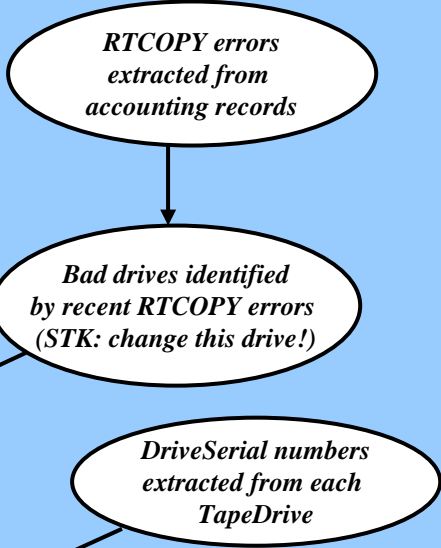
NB: Only cases where > 1 VID is affected are shown in the following table.

- Tapes in erase pools or ARCHIVED status are ignored.
- The following VID's have been ignored; errors on too many tapeservers in last 30 days: P11550 (3) P12853 (15) P12352 (5)

Tape Server	Number of Different VID's	Number of Errors
tpsrv033	4	4
tpsrv602	2	2
tpsrv614	2	2
tpsrv619	2	2

Data is ordered by TapeServer and Date of Error. means TapeServer currently in_dev or not in showqueues (not in production).

VID	Tape Server	Tape Drive	Drive SerialNo	Date of Error	Error Message
P00707	tpsrv018	994B5491	479000014134	12 May 16:32:17	CFTPDISK ! I/O ERROR READING FROM TAPE: Medium error ASC=11 ASCQ=1 (BLOCK # 162)
P06143	tpsrv034	994B5497	479000001686	08 May 12:15:41	CPDSKTP ! I/O ERROR WRITING ON TAPE: wrttpmk: TP042 - PATH14436 : ioctl error : Medium error ASC=C ASCQ=0 (BLOCK # 800) CPDSKTP ! TAPE IS NOW INCORRECTLY TERMINATED
P15621	tpsrv033	994B53A0	479000025348	04 May 15:48:35	CFTPDISK ! I/O ERROR READING FROM TAPE: Medium error ASC=11 ASCQ=1 (BLOCK # 26698)
P08429	tpsrv033	994B53A0	479000025348	06 May 23:14:56	CPDSKTP ! I/O ERROR WRITING ON TAPE: wrttpmk: TP042 - PATH29131 : ioctl error : Medium error ASC=C ASCQ=0 (BLOCK # 3747) CPDSKTP ! TAPE IS NOW INCORRECTLY TERMINATED
P10029	tpsrv033	994B53A0	479000025348	07 May 13:19:37	CPDSKTP ! I/O ERROR WRITING ON TAPE: wrttpmk: TP042 - PATH30580 : ioctl error : Medium error ASC=C ASCQ=0 (BLOCK # 694) CPDSKTP ! TAPE IS NOW INCORRECTLY TERMINATED
P10395	tpsrv033	994B53A0	479000025348	09 May 14:46:51	CPDSKTP ! I/O ERROR WRITING ON TAPE: wrttpmk: TP042 - PATH7186 : ioctl error : Medium error ASC=C ASCQ=0 (BLOCK # 1860) CPDSKTP ! TAPE IS NOW INCORRECTLY TERMINATED
P00707	tpsrv036	994B53A.A	479000011230	14 May 03:52:36	CFTPDISK ! I/O ERROR READING FROM TAPE: Medium error ASC=11 ASCQ=1 (BLOCK # 484)
P10193	tpsrv057	994B549A	479000001823	05 May 12:52:07	CFTPDISK ! I/O ERROR READING FROM TAPE: Medium error ASC=11 ASCQ=1 (BLOCK # 323)
P06646	tpsrv036	994B53A.A	479000011230	01 May	CFTPDISK ! I/O ERROR READING FROM TAPE: Medium error ASC=11 ASCQ=1 (BLOCK # 1030)



TapeDrives with too many errors

CASTOR last read/write TapeServers for Disabled Tapes

Last read and write tapeserver associations for DISABLED tapes. Also shown for tapes with BAD MIRs, where relevant and data is available.

Data extracted on Sun May 14 17:18:16 CEST 2006. Last 30 days of information is shown in reverse time order.

- **Found discovered:** date,time on which cron job found this tape to have been DISABLED. Such cron jobs run at ~ 15 minutes past the hour.
- **Last Read/Write TapeServer and Dates:** taken from vmgr
- **BAD MIR TapeServer and Date:** only meaningful if tape has a BAD MIR; data only available if operators have followed the correct procedure to cancel the request after the BAD MIR has been found
- This page shows *HISTORY* data starting on June 24 2005. Data is ordered by date when the tapes were discovered to be DISABLED.
- The aim of this information is to be able to spot TapeServers that have recently caused multiple DISABLED tapes (i.e. those requiring a TapeDrive change).

in VID field means the tape is no longer DISABLED

VID	Discovered Disabled on	Last Read TapeServer	Last Read date	Last Write TapeServer	Last Write date	TapeServer on which BAD MIR found	When BAD MIR found
P00707	14-May-2006 04:15	tpsrv056	Sun May 14 03:50:07 2006	tpsrv055	Mon Oct 4 01:31:53 2004		
P00707	12-May-2006 17:16	tpsrv022	Fri May 12 17:07:09 2006	tpsrv055	Mon Oct 4 01:31:53 2004		
P16928	12-May-2006 17:16	tpsrv121	Fri May 12 17:09:43 2006	tpsrv108	Wed Feb 1 05:29:45 2006		
T01165	12-May-2006 11:15	tpsrv608	Fri May 12 10:16:47 2006	tpsrv603	Tue May 2 13:11:55 2006		
P19224	12-May-2006 10:16	tpsrv055	Fri May 12 08:58:49 2006	tpsrv055	Tue Aug 24 02:18:24 2004	tpsrv055	12-May-2006 09:29
T00335	12-May-2006 07:15	N/A	Thu Jan 1 01:00:00 1970	tpsrv622	Fri May 12 06:54:45 2006		
T01163	12-May-2006 06:15	N/A	Thu Jan 1 01:00:00 1970	tpsrv623	Fri May 12 00:53:21 2006		
T00337	12-May-2006 05:15	N/A	Thu Jan 1 01:00:00 1970	tpsrv239	Mon Mar 6 14:43:22 2006		
P03033	24-Apr-2006 20:15	tpsrv018	Mon Apr 24 19:44:53 2006	tpsrv022	Sun May 9 11:35:50 2004		
P22244	24-Apr-2006 12:15	tpsrv118	Mon Apr 24 10:25:59 2006	tpsrv123	Fri Jul 4 20:50:40 2003		
P13417	24-Apr-2006 10:06	tpsrv023	Sun Apr 23 11:39:01 2006	tpsrv028	Tue Apr 18 15:55:27 2006		
P22244	23-Apr-2006 14:15	tpsrv119	Sun Apr 23 12:13:08 2006	tpsrv123	Fri Jul 4 20:50:40 2003		
P22243	23-Apr-2006 11:15	tpsrv103	Sun Apr 23 10:16:03 2006	tpsrv110	Fri Jul 4 05:25:43 2003		
P17234	23-Apr-2006 09:16	tpsrv103	Sun Apr 23 08:05:39 2006	tpsrv104	Fri Mar 31 22:53:09 2006		
P17236	23-Apr-2006 06:15	tpsrv103	Sun Apr 23 05:26:57 2006	tpsrv100	Tue Apr 4 02:54:53 2006		
P17230	23-Apr-2006 06:15	tpsrv103	Sun Apr 23 05:16:36 2006	tpsrv103	Wed Apr 5 09:57:13 2006		
P11033	23-Apr-2006 04:16	tpsrv103	Sun Apr 23 03:46:57 2006	tpsrv102	Mon Sep 20 21:59:43 2004		
P17230	23-Apr-2006 04:16	tpsrv103	Sun Apr 23 03:18:21 2006	tpsrv115	Fri Mar 31 13:23:57 2006		
P03240	22-Apr-2006 21:15	tpsrv103	Sat Apr 22 20:43:03 2006	tpsrv124	Thu Aug 26 18:27:16 2004		
P08438	22-Apr-2006 08:15	tpsrv103	Sat Apr 22 07:43:47 2006	tpsrv119	Wed Feb 22 16:02:50 2006		
P10103	22-Apr-2006 00:15	tpsrv054	Fri Apr 21 23:24:01 2006	tpsrv020	Fri Apr 21 07:08:49 2006	tpsrv057	21-Apr-2006 23:49
P09137	21-Apr-2006 12:15	tpsrv053	Fri Apr 21 12:06:48 2006	tpsrv020	Sun May 22 06:18:46 2005		
T00120	21-Apr-2006 11:15	N/A	Thu Jan 1 01:00:00 1970	tpsrv211	Fri Apr 21 09:35:14 2006		
P09173	21-Apr-2006 08:31	tpsrv018	Fri Apr 21 09:17:25 2006	tpsrv059	Sun Mar 19 13:04:58 2006		
T01199	21-Apr-2006 05:16	N/A	Thu Jan 1 01:00:00 1970	tpsrv622	Fri Apr 21 02:31:33 2006		

7 tapes DISABLED
by same TapeDrive
Source: vmgr

Bad tapes also identified

Castor Drive and Tape volume errors

Page extracted on 14-MAY-2006 16:09:35

vmgr information about P17457

- Library: STK_ACS5.... TapePool: atlas_new2
- Density: 200GC.... Status: RONLY
- Manufacturer: IMATION.... Serial No.: Unknown.... Label Type: aul
- Entered into vmgr on: 27-May-2003.... Read count: 2.... Write count: 7
- Last read date: 15-Jun-2004 09:21 on host: tpsrv010
- Last write date: 28-Apr-2006 11:22 on host: tpsrv026
- Estimated free space: 200GB.... No. files: 0

Same tape has errors on multiple TapeDrives within a short time frame. (tape needs repair/copying)

List of ALL RTCOPY errors for volume P17457

Data are sorted by time

Category	Date:time	userid	gg	fseq	drive	mode	tapeserver	diskserver	clienthost	code	error message
Tape	28/04/2006 12:37:56	stage	st	1	994B5496	w	tpsrv023	lxfsra1007	c2atlasrv02.cern.ch	222	readlbl: TP042 - /var/tmp/RTCOPY/PATH29240 : read error : Medium error ASC=11 ASCQ=1
Tape	28/04/2006 12:54:28	stage	st	1	994B53A7	w	tpsrv059	lxfsrk6102	c2atlasrv02.cern.ch	222	readlbl: TP042 - /var/tmp/RTCOPY/PATH2339 : read error : Medium error ASC=11 ASCQ=1
Tape	28/04/2006 13:12:20	stage	st	1	994B53A3	w	tpsrv021	lxfsra1007	c2atlasrv02.cern.ch	222	readlbl: TP042 - /var/tmp/RTCOPY/PATH24904 : read error : Medium error ASC=11 ASCQ=1
Tape	28/04/2006 13:20:59	stage	st	1	994B5491	w	tpsrv018	lxfsra1008	c2atlasrv02.cern.ch	222	readlbl: TP042 - /var/tmp/RTCOPY/PATH834 : read error : Medium error ASC=11 ASCQ=1
Tape	28/04/2006 13:26:28	stage	st	1	994B53A5	w	tpsrv026	lxfsra1007	c2atlasrv02.cern.ch	222	readlbl: TP042 - /var/tmp/RTCOPY/PATH7805 : read error : Medium error ASC=11 ASCQ=1

[Change Repair data for a tape \(mark as repaired\)](#)

TapeRepairs: drive/volume/errDB

Castor Drive and Tape volume errors

Tape Repair and Disabled Segment management

On this page you may select a vid that is not yet repaired in order to update the data

If you wish to set a *list of tapes as repaired*, it may be easier to use the script `/afs/cern.ch/project/castor/TOOLS/.../tape_list_repaired`. Type the script name for help.

Select a tape volume (from the list of volumes that have not yet been repaired)

To retrieve the information

Current list of vid's repaired ordered by vid and date of repair

Id	vid	Date of entry	Date/time of last update	Repair Time Estimate	Problem Diagnosis	Work log
8161	P00707	14-May-2006 04:15:13	14-May-2006 04:15:13	Unknown	Unknown	14-May-2006 04:15:13:[Castor] Disabled Tape
8039	P09085	04-Apr-2006 20:15:25	11-Apr-2006 14:22:04	Unknown	Unknown	16-Feb-2005 19:36:03:[JFL]Tape found stuck in tape drive 994B5497 due to a demount failed. PROCESS_FAILURE. MIR should be invalid. Needs to be repacked. 05-Apr-2006 10:27:32:[JFL] Tape has a bad MIR. Trying to rebuild it but the tape drive gave a CHK A737 error on LCD display during the rebuilding. Repack is in progress (1486 files). 11-Apr-2006 14:22:04:[JFL] Unable to do anything from file 1371 to end of tape on fseq 1486 in normal stagein or in mt/dd mode (got input/output error). Tape will be sent to US for these last data recovery (116 files).
8083	P16236	18-Apr-2006 19:15:56	25-Apr-2006 12:37:51	Unknown	Unknown	31-Jan-2006 14:38:58:[JFL]Needs to be checked due to a twice demount failed on tape drives 994B43A2 & 994B42A8. 19-Apr-2006 10:19:25:[JFL] Tape DISABLED in CASTOR due to a demount failed on tape drive 994B021C but found broken in drive in the middle of data. A lot of files should be damaged but when it will be repaired I'm trying a stagein on it... wait and see.... 25-Apr-2006 12:37:51:[JFL] A lot of files have been recovered in mt/dd mode. Now the tape will be sent to US for the remaining of files not recovered at CERN. Set FULL ARCHIVED in CASTOR & in library P_B186.
8137	P23181	08-May-2006 17:15:10	12-May-2006 13:26:19	Unknown	Unknown	08-May-2006 17:15:10:[Castor] Disabled Tape 12-May-2006 13:26:19:[JFL] Tape dedicated to EHM Felix for castor_tapetest.
8144	P23183	09-May-2006 18:15:43	12-May-2006 13:27:00	Unknown	Unknown	09-May-2006 18:15:43:[Castor] Disabled Tape 12-May-2006 13:27:00:[JFL] Tape dedicated to EHM Felix for castor_tapetest.
8138	P23187	08-May-2006 17:15:10	12-May-2006 13:27:41	Unknown	Unknown	08-May-2006 17:15:10:[Castor] Disabled Tape 12-May-2006 13:27:41:[JFL] Tape dedicated to EHM Felix for castor_tapetest.

Select a DISABLED tape

Simple work log:
who did what, when and why

TapeRepairs: drive/volume/errDB

Castor Drive and Tape volume errors

Tape repair information for tapes in status "to be repaired"

Here you may modify the existing data for vid **P09085** which has status: **FULL|ARCHIVED**

[Click here to see ALL rtbody errors for this volume](#)

Id	vid	Current state of repair	Date/time of entry	Date/time of last update	Repair Time Estimate	Problem Diagnosis	Work log
8039	P09085	To be repaired	04-Apr-2006 20:15:25	11-Apr-2006 14:22:04	Unknown	Unknown	16-Feb-2005 10:36:03:[JFL] Tape found stuck in tape drive 994B5497 due to a demount failed : PROCESS_FAILURE. MIR should be invalid. Needs to be repacked. 05-Apr-2006 10:27:32:[JFL] Tape has a bad MIR. Trying to rebuild it but the tape drive gave a CHK A737 error on LCD display during the rebuilding. Repack is in progress (1486 files). 11-Apr-2006 14:22:04:[JFL] Unable to do anything from file 1371 to end of tape on fseq 1486 in normal stagein or in mt/dd mode (got input/output error). Tape will be sent to US for these last data recovery (116 files).

(Optionally): change the problem diagnosis by choosing a new value:

(Optionally): change the time estimate for repair by choosing a new value:

(Optionally): change the repair status by choosing "repaired":

- To be repaired
- Repaired:

Mandatory: Work log for repair. (You must enter at least 10 characters)

NB: Text wraps automatically.

Enter your initials please:

*Re-enable a tape after repair
(log of who did what, when
and why)*

Drive/volume/errDB

Castor Drive and Tape volume errors

Page extracted on 14-MAY-2006 17:40:40

Links are dynamic; response will vary with number of items selected

RTCOPY errors by category for varying time periods

User, Network and Operational category errors older than 30 days are cleared from the database

Clicking on the numbers in the following table will show error messages **ordered by time**

Clicking on a category type will show a plot of errors over time

Type	# Errors for last 4 hours	# Errors for last 24 hours	# Errors for last 7 days	# Errors for last 30 days	# Errors for last 100 days
User	29	205	1665	5791	5931
Operational	0	13	144	299	299
Tape	2	51	83	2845	3105
Unknown	0	2	29	302	942
Network	0	4	110	1512	1514
System	0	1	23	67	1089
Robotics	0	0	2	6	32
Total	31	276	2056	10822	12912

Oracle start and stop times for this page:14-May-2006 17:40:40 and 14-May-2006 17:40:41

Clicking on the numbers in the following table will show error messages **ordered by drive and time**

Clicking on a category type will show a plot of errors with time

Type	# Errors for last 4 hours	# Errors for last 24 hours	# Errors for last 7 days	# Errors for last 30 days	# Errors for last 100 days
User	29	205	1665	5791	5931
Operational	0	13	144	299	299
Tape	2	51	83	2845	3105
Unknown	0	2	29	302	942
Network	0	4	110	1512	1514
System	0	1	23	67	1089
Robotics	0	0	2	6	32

Overview of error rate by category

User errors are usually not important (and not usually caused by the user..)

'Tape' and 'Unknown' (unclassified) are the most significant errors – and even these need to be filtered to spot bad VID's and/or bad TapeDrives

Drive/volume/errDB

Castor Drive and Tape volume errors

Page extracted on 14-MAY-2006 17:42:15

Links are dynamic; response will vary with number of items selected

RTCOPY (tape + unknown) errors by volume for type 9940

Volume (# Errors)	Volume (# Errors)	Volume (# Errors)	Volume (# Errors)	Volume (# Errors)	Volume (# Errors)	Volume (# Errors)	Volume (# Errors)
P00002 (1)	P00003 (2)	P00004 (1)	P00006 (1)	P00008 (2)	P00014 (1)	P00018 (1)	P00019 (1)
P00023 (1)	P00025 (9)	P00027 (18)	P00030 (5)	P00036 (1)	P00039 (2)	P00040 (1)	P00041 (1)
P00043 (1)	P00044 (2)	P00045 (1)	P00046 (1)	P00050 (1)	P00063 (1)	P00065 (1)	P00066 (1)
P00068 (1)	P00076 (2)	P00077 (1)	P00078 (1)	P00080 (30)	P00082 (2)	P00083 (4)	P00084 (2)
P00088 (5)	P00090 (1)	P00092 (1)	P00093 (2)	P00094 (2)	P00096 (1)	P00097 (3)	P00101 (1)
P00102 (1)	P00103 (3)	P00105 (1)	P00115 (10)	P00124 (1)	P00125 (1)	P00133 (1)	P00138 (1)
P00140 (1)	P00141 (1)	P00145 (1)	P00158 (2)	P00164 (6)	P00169 (1)	P00172 (1)	P00176 (7)
P00178 (2)	P00179 (32)	P00180 (1)	P00188 (1)	P00192 (1)	P00200 (8)	P00203 (3)	P00208 (1)
P00210 (3)	P00214 (1)	P00218 (4)	P00223 (1)	P00225 (1)	P00228 (39)	P00231 (4)	P00239 (1)
P00241 (1)	P00245 (1)	P00248 (1)	P00249 (1)	P00251 (1)	P00278 (1)	P00280 (38)	P00282 (1)
P00283 (1)	P00286 (1)	P00287 (2)	P00289 (5)	P00291 (1)	P00293 (2)	P00297 (1)	P00298 (1)
P00301 (2)	P00302 (1)	P00303 (1)	P00305 (1)	P00310 (1)	P00311 (1)	P00312 (3)	P00313 (2)
P00314 (2)	P00317 (1)	P00318 (1)	P00319 (2)	P00320 (2)	P00322 (13)	P00324 (2)	P00335 (2)
P00337 (3)	P00339 (1)	P00341 (1)	P00344 (17)	P00345 (1)	P00351 (2)	P00352 (1)	P00354 (1)
P00356 (1)	P00357 (2)	P00360 (2)	P00368 (84)	P00373 (2)	P00375 (1)	P00377 (1)	P00382 (2)
P00384 (1)	P00386 (1)	P00394 (1)	P00396 (2)	P00398 (45)	P00400 (1)	P00401 (1)	P00402 (1)
P00407 (2)	P00408 (40)	P00409 (2)	P00414 (32)	P00416 (1)	P00420 (1)	P00421 (1)	P00426 (2)
P00438 (3)	P00442 (1)	P00449 (2)	P00454 (16)	P00455 (1)	P00456 (20)	P00458 (13)	P00462 (1)
P00463 (2)	P00464 (1)	P00470 (16)	P00480 (2)	P00482 (26)	P00483 (1)	P00493 (1)	P00499 (1)
P00501 (2)	P00514 (1)	P00530 (2)	P00532 (41)	P00537 (1)	P00548 (1)	P00553 (2)	P00557 (1)
P00565 (1)	P00567 (1)	P00571 (3)	P00573 (1)	P00576 (1)	P00577 (3)	P00581 (1)	P00584 (1)
P00593 (3)	P00598 (1)	P00600 (5)	P00601 (1)	P00604 (1)	P00605 (33)	P00607 (1)	P00612 (1)
P00615 (2)	P00616 (2)	P00619 (1)	P00620 (2)	P00621 (4)	P00622 (1)	P00623 (2)	P00626 (1)
P00627 (3)	P00628 (1)	P00629 (3)	P00631 (2)	P00632 (1)	P00636 (1)	P00637 (1)	P00640 (1)
P00654 (1)	P00661 (1)	P00664 (5)	P00670 (1)	P00674 (1)	P00687 (1)	P00707 (3)	P00717 (1)

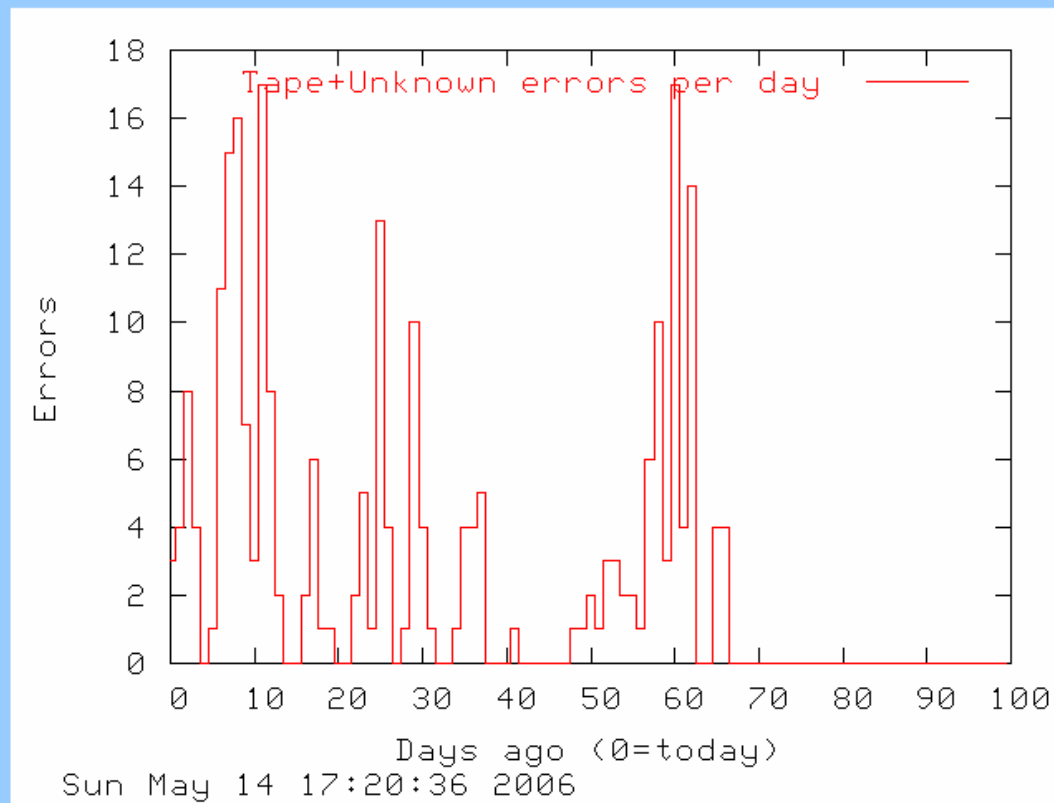
*Serious errors by VID
Likely solution is to
copy/retire this VID*

Drive/volume/errDB

Castor Drive and Tape volume errors

Page extracted on Sun May 14 17:20:36 2006

Errors (Tape + Unknown) per calendar day for drive type T10K



*New TapeDrive/media
not without problems!*

Drive/volume/errDB

Oracle forms query front-end allows to query for specific error strings, TapeDrives, VIDs

Castor Drive and Tape volume errors

On this page you may query the RTCOPY error messages for 'tape' and 'unknown' errors

Query options # 1: 'Tape' and 'Unknown' errors that have occurred in last N days (you specify N))

select a drive (or not) (from the list of drives that have had errors)
select a tape volume (or not) (from the list of volumes that have had errors)
enter at least part of an error message text (or nothing) (Spaces are significant, query is case independent)
enter the time cut -- units are days (default = 1.0)

Query option # 2: all 'Tape' and 'Unknown' errors in last 7 days

Type any character here to query on ALL 'Tape' and 'Unknown' errors for last 7 days :

NB: checking this option means any query data specified above will be ignored

Query option # 3: all 'Tape' and 'Unknown' errors in last 7 days for a given clienthost

select a clienthost (or not) (from the list of clienthosts that have had errors in last 7 days)

NB: checking this option means any query data specified above will be ignored

Query option # 4: all 'Tape' and 'Unknown' errors in last 7 days for a given diskserver

select a diskserver (or not) (from the list of diskservers that have had errors in last 7 days)

NB: checking this option means any query data specified above will be ignored

Submit query

Reset fields

End-user pages to explain data inaccessibility

CASTOR Disabled and Archived Tapes

This page was refreshed on Sun May 14 17:15:50 CEST 2006 and is remade 15 mins past each hour.

There are 6 DISABLED tapes (0 of which are also be ARCHIVED).

CASTOR Tape Pools may contain a variety of tapes. Below you see the tapes currently in CASTOR Tape Pools which are in DISABLED status. This means they cannot be accessed, and there may be a problem with them (stuck in drive?) or suspected damage to the tape. Usually such a tape has not in fact been damaged and can be returned to normal status once checked. Less often a recovery action is required, and this can take a day or more and may not always be completely successful.

A list of unavailable tapes may be found in /afs/cern.ch/project/cndoc/wwwds/HSM/CASTOR/DOCUMENTATION/END_USERS/unavailable.tapes

vid	vmgr status	TapePool	No. Castor files	When detected Disabled	Reason
P23181	DISABLED	castor_tapetest	0	08 May 2006 17:15	Castor software operation
P23189	DISABLED	castor_tapetest	0	08 May 2006 17:15	Castor software operation
P23191	DISABLED	castor_tapetest	0	08 May 2006 17:15	Castor software operation
P23187	DISABLED	castor_tapetest	0	08 May 2006 17:15	Castor software operation
P23183	DISABLED	castor_tapetest	0	09 May 2006 18:15	Castor software operation
P00707	FULLDISABLED	largeuser	6131	14 May 2006 04:15	Castor software operation

Similar pages for disabled tape segments (files)

Castor does not reveal why it does what it does..

Files still on these tapes may be lost forever

There are 33 ARCHIVED tapes.

When tapes are badly damaged, they may be logically and/or physically ejected from the tape robots either to send them to a recovery or because nothing more can be done.

vid	vmgr status	TapePool	No. Castor files	Last Action	Reason
I00075	ARCHIVED	itdc_3592B1_new	0	24 Mar 2006 15:28	Tape does not exist in the IBM robot. Set ARCHIVED in CASTOR.
P00027	FULLARCHIVED	default	243	07 Mar 2006 14:40	STK was unable to recover any more files; tape was too badly damaged
P00605	FULLRDONLYARCHIVED	na482	498	15 Feb 2005 10:17	All files have been recovered as far as possible
P00621	FULLARCHIVED	na482	2	07 Mar 2006 14:41	STK was unable to recover any more files; tape was too badly damaged
P00852	FULLRDONLYARCHIVED	alicedc04	9942	15 Feb 2005 10:18	All files have been recovered as far as possible
P01197	FULLRDONLYARCHIVED	l3c	411	Unknown (before July 2005)	Unknown
P01223	FULLARCHIVED	cms	43	26 Oct 2005 10:33	Tape repacked as far as possible (mt/dd mode), now ejected from ACS5 and stored in a rack in B513.
P01288	FULLARCHIVED	cms	39	07 Mar 2006 14:41	STK was unable to recover any more files; tape was too badly damaged
P01425	ARCHIVED	erase_994BR5	0	30 Sep 2005 11:20	Sent to US for all data recovery (23 files should be repaired). (22/9/05)
P02571	FULLARCHIVED	default	1465	17 Nov 2004 13:11	Tape is damaged, bad file structure around file 2804. Files recovered with b1p stagein and file identification as far as possible to /castor/cern.ch/c3/9940/P02572/user/
P02850	FULLARCHIVED	lhcb	30	16 Feb 2006 11:46	Tape Repaired
P03008	ARCHIVED	lhcb	0	11 Oct 2005 09:35	Tape repacked except the files 34 & 422 (Medium error ASC=11 ASCQ=1). into Pool lhcb
P03555	FULLARCHIVED	na49	41	14 Dec 2005 11:39	Tape (broken) came back from US after repair & copied as far as possible in mt/dd mode. A lot of files are now renamed

Tapes that have been repaired as much as possible but still contain (inaccessible) files are archived. Users may (and do) try to access these files years later.

Stager survey

CASTOR Stager survey

Page Generated on Sun May 14 16:18:16 CEST 2006

For all stage-pingable/stage-queryable stagers with Castor version \geq 1.7.1.5

This script took 12 seconds to run

Stager response timeout rates (tested twice each 15 mins)

stager	% failure rate during last 24 hours	% failure rate during last 2 hours
stagera49	0.52	0.00
stagenomad	6.01	0.00

Information pages for all stagers:

- [PUT_FAILED statistics](#) (on 14 May 2006 14:31:51 there were a total of 7 files in PUT_FAILED status)
- [Number of links managed by the stagers](#)
- [Castor filenames with imbedded white space in stager catalogues](#)
- [Files resident on stager filesystems yet not in the stager catalogue](#)
- [Number of STAGEALLOC files not accessed for > 1 month](#)
- [Number of STAGEOUT files not accessed \(incremented\) for > 30 days](#)
- [Number of STAGEIN requests in that state for > 2 days](#)
- [Number of files with internal paths not defined in /etc/STGCONFIG](#)
- [Castor version, O/S \(kernel\), stager process and opened file descriptor limits](#)
- [Castor stager database \(Cdb\) file size, last modification date](#)
- [Survey of migration speeds for all stagers for today/last 3 days](#)

Indicates basic stager response to simple stageqry

A set of problems/parameters for all stagers – colour coded to indicate if action is required

rfdir of every DiskPool FileSystem using weekly batch jobs

Cdb filesize (influences speed of change of catalogue) - see next slide

Stager survey

Castor stager internal database (Cdb) file size (/usr/spo

Data extracted on **Sun May 14 15:45:01 CEST 2006**. Script `dbsize_s`

Maximum db file size is 2 GB. The database should be exported and i

Stager	DB file size (GB)	Last Modification date
stagealeph	0.00	Mar 30 17:20:00 2006
stagecdr	0.05	May 14 15:06:00 2006
stagechorus	0.00	May 12 18:36:00 2006
stagecompass	0.24	May 14 15:44:00 2006
stagedelphi	0.04	May 14 15:43:00 2006
stagedirac	0.11	May 13 17:01:00 2006
stageharp	0.03	May 14 04:28:00 2006
stagei3	0.11	May 14 15:40:00 2006
stageihcb	0.23	May 14 15:44:00 2006
stagen45	0.03	May 13 08:25:00 2006
stagen48	0.07	May 14 15:45:00 2006
stagen49	0.28	May 14 15:45:00 2006
stagen60	0.01	May 13 18:47:00 2006
stagenomad	0.10	May 14 10:16:00 2006
stagentof	0.04	May 7 09:56:00 2006
stageopal	0.02	May 13 09:05:00 2006
stagepublic	0.27	May 14 15:44:00 2006
stageslap	0.17	May 12 00:21:00 2006

Aleph stager hardly used these days

*Keep catalogue to < 0.5GB
(Size significantly affects speed of catalogue operations – notably stageclrs)*

Stager survey

stager	disk_pool (# files) (Ave. Lifetime Days)	disk_pool capacity	% free space	GC?	GC % thresholds	last GC started	last GC duration	migrator? (files to migr.)	mig start thresh	mig data thresh	last migr started	last mig duration
stagera49	na49_dst (178034) (32.14 days)	7.35T	61.2%	GC	15 .. 20	May 12 12:09:23	0:20:35	na49migr (889)	0	0	May 14 15:17:47	active
	na49_keep (31948) (155.37 days)	2.91T	45.1%	GC	15 .. 20		has not run	na49migr (889)	0	0	May 14 15:17:47	active
	na49_stage (0)	894.08G	99.9%	GC				na49migr (889)	0	0	May 14 15:17:47	active
stagera60	na60 (19913) (84.56 days)	4.44T	46.4%	GC	10 .. 13		has not run	na60 (0)	0	20.00M	Apr 28 13:15:05	0:01:57
stagenomad	nomad_ana (10694) (161.11 days)	1.64T	61.6%	GC	10 .. 15				0			has not run
	nomad_pro (32967) (33.36 days)	1.96T	54.0%	GC	04 .. 08							has not run
	nomad_ref (9621) (216.89 days)	1.53T	80.0%	GC	04 .. 08							has not run
stagentof	ntof_dst (120135) (418.01 days)	4.22T	37.9%	GC	12 .. 14		has not run	ntofmigr (0)	0	20.00G		has not run
	ntof_proc (2455) (415.62 days)	4.29T	16.9%	GC	12 .. 14		has not run	ntofmigr (0)	0	20.00G		has not run
	ntof_stage (688) (169.76 days)	1.09T	28.6%	GC	15 .. 25		has not run	ntofmigr (0)	0	20.00G		has not run
stageopal	opalstage (18653) (141.86 days)	3.56T	13.1%	GC	10 .. 15	Apr 14 21:31:33	0:01:01	opalmigr (0)	0	0		has not run
stagepublic	public (147325) (54.36 days)	14.84T	38.1%	GC	20 .. 24	May 9 17:48:57	2:47:19	public_migr (18)	0	10.00G	May 14 15:57:59	0:03:20
stageslap	colldata (0)						has not run	slapmigr (0)	0	10.00G	May 14 16:07:48	0:01:54
	slapdata (98...) (201.27 days)						has not run	slapmigr (0)	0	10.00G	May 14 16:07:48	0:01:54

Current migration under way

Basic information about configuration and performance of stagers and diskpools. Many fields change colour in case of problems. Source : stageqry. Admins notified of serious conditions by e-mail

Colour warning if e.g. DiskServer(s) in trouble or catalogue too big

Stager survey

CASTOR Stager

Page extracted on Sun May 14 16:18:25 CEST 2006

[Previous weeks activity plots for stagena49](#)

[Daily Evolution of stager catalogue for stagena49.](#)

[Activity plots for stagena49 so far today.](#)

Stager parameters for [stagena49](#) :: result of stageping -v

```
Stager daemon on stage010.cern.ch - CASTOR 1.7.1.5
Generated May 13 2005 around 13:49:10
Running since Apr 11 09:25:05, pid=4652
Maximum/reserved/available number of opened file descriptors: 4096/48/4048
Maximum number of processes: { cur=7168, max=7168 }
```

Diskservers for stager [stagena49](#):

MSA information extracted on Sun May 14 16:15:43 2006

Diskserver	rfio test status	Lemon MSA status	Hardware model	O/S (Kernel)	SMS state
bxfs5031	OK	OK	e2	slc3 (2.4.21-40.EL.cernsmp)	production
bxfs5047	OK	OK	e2	slc3 (2.4.21-40.EL.cernsmp)	production
bxfsrk5103	OK	OK	s0	slc3 (2.4.21-40.EL.cernsmp)	production
bxfsrk5503	OK	OK	s0	slc3 (2.4.21-40.EL.cernsmp)	production
bxfsrk6105	OK	OK	s0	slc3 (2.4.21-40.EL.cernsmp)	production
na49008d	OK	OK	a1	slc3 (2.4.21-40.EL.cernsmp)	production
na49009d	OK	OK	a1	slc3 (2.4.21-40.EL.cernsmp)	production

Last few year's weekly statistics of basic stager operations

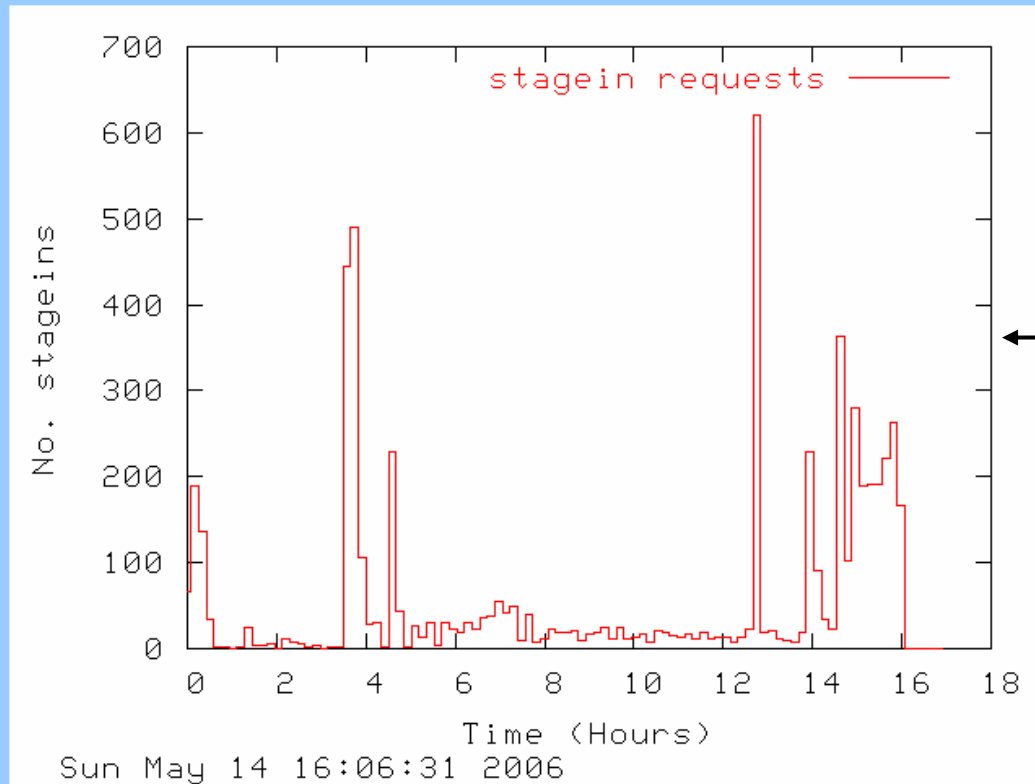
Evolution of overall stager catalogue, files sizes, lifetimes etc over last months

The story so far today (basic stager operations)

Overview of DiskServers status, tests and links to Lemon

Stager activity so far today

There have been 5847 stagein requests. Time distribution since 00:00 hrs:



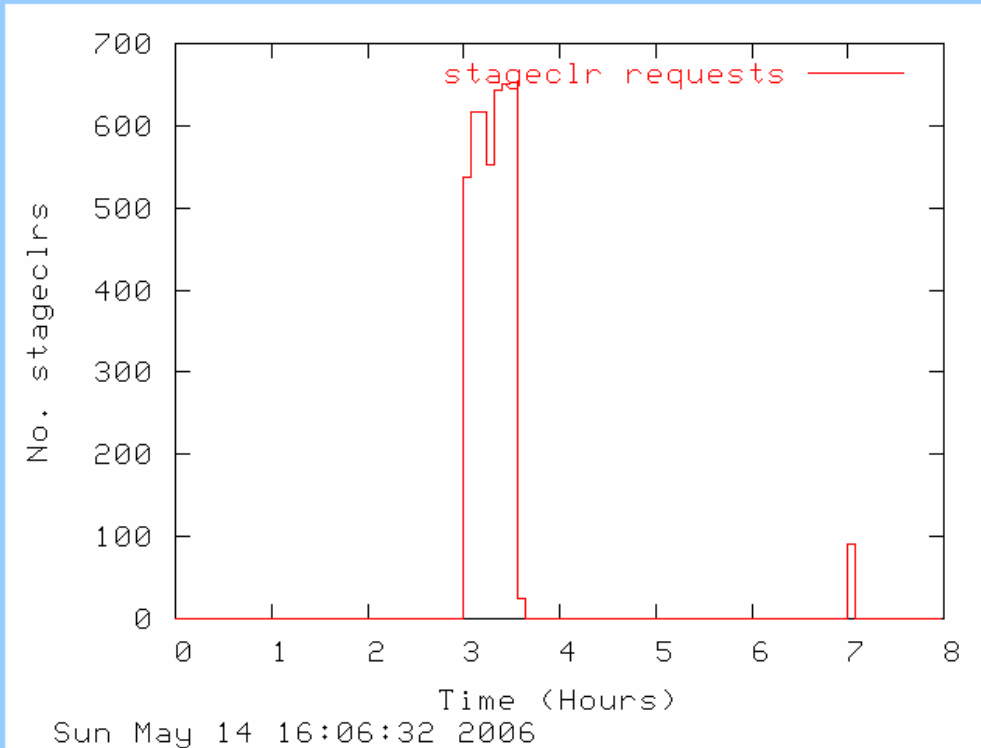
Top 5 stagein users so far today

userid	gg	# stagein requests
cblume	vp	1932
rolandc	vp	1343
pchung	vp	1332
laszloa	vp	1210
mmitrov	vp	21

Top users opening files to read today

Stager activity so far today

There have been 4384 stageclr requests. Time distribution since 00:00 hrs:



*Time distribution of stageclrs
(garbage collection for space or
catalogue reduction -too many files)*

Top 5 stageclr users so far today

userid	gg	# stageclr requests
castorc3	c3	4384

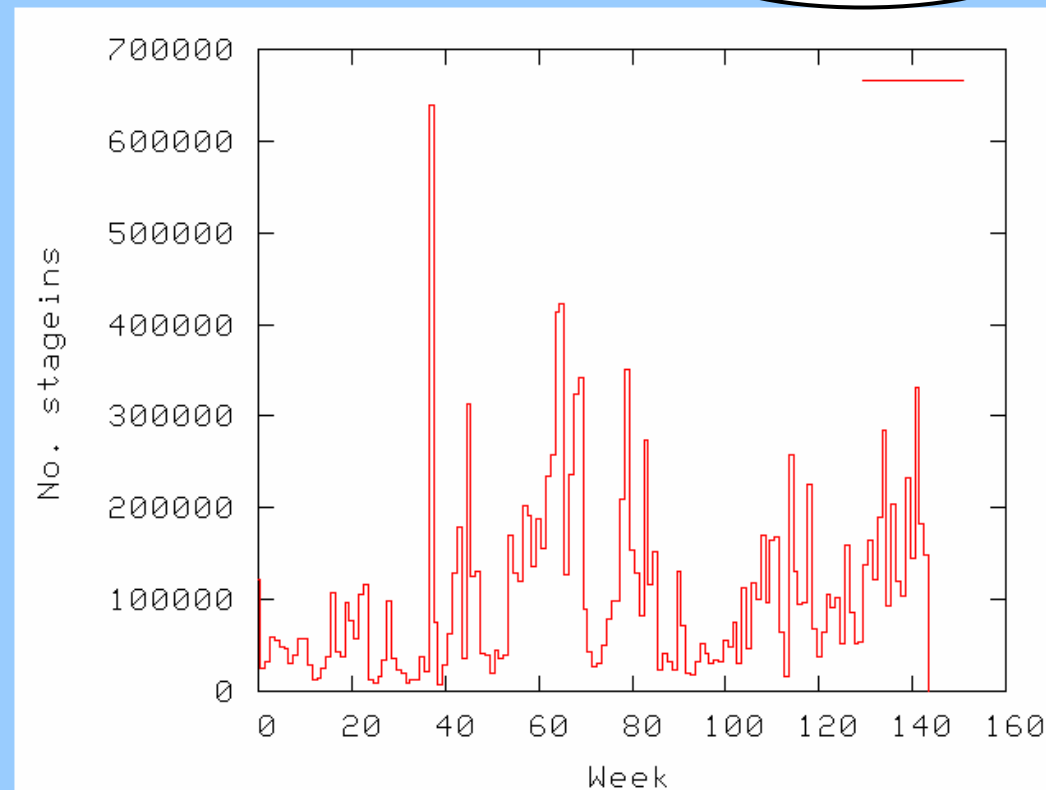
*castorc3 is the userid used for
catalogue control. (root performs
garbage collections)*

Evolution of stager (past few years)

Weekly activity for stager stagera49

Page extracted on Mon May 8 01:17:00 CEST 2006 . Data points every week

Fig 1. Number of stager's per week (from week 0330 to 0618) *March 2003-> May 2006*



First week (0) == week no. 0330, last week (143) == week no. 0618 in above plot.

*Source: Castor weekly
accounting records*

Evolution of stager (past few months)

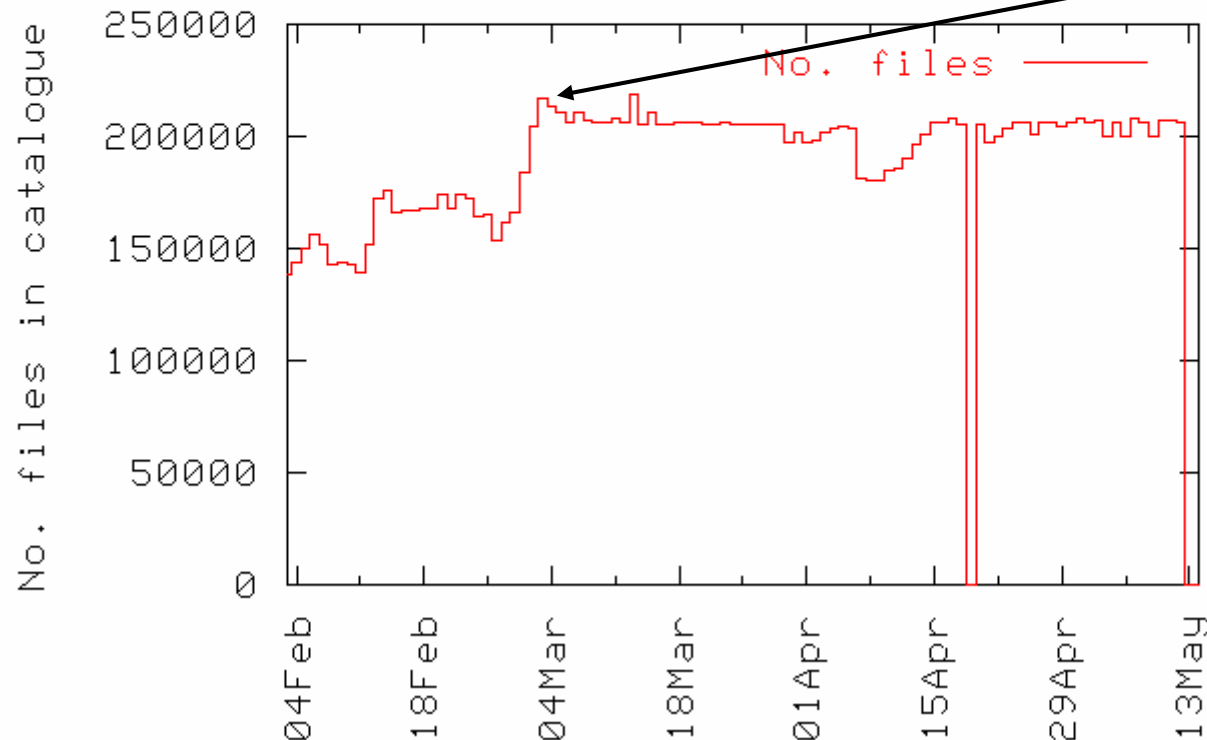
CASTOR Stager: Evolution of the stager catalogue

Evolution of the catalogue for stager [stagera49](#)

Page extracted on Sun May 14 00:03:01 CEST 2006.

Data limited to last 100 days

Fig 1: Total Number of files in Catalogue for stagera49 (snapshot each midnight)



*Catalogue size grows beyond
Castor's ability to manage
(on this hardware and with this castor version)
The catalogue starts to be limited
by external scripts tailored
to each stager instance*

Stager DiskPool statistics

CASTOR Stager survey

Scan of data for pool: na49_dst of stager: stagena49

Page extracted on Sun May 14 16:31:34 CEST 2006

- [STGCONFIG details](#)
- [PUT_FAILED analysis](#)

Data extracted on Sat May 13 14:10:54 2006. <<< Please note this date/time

Total files in catalogue: 176458, of these 176458 are HSM files and 0 are non-HSM files

File distribution by status

Status	No. of files	Total Size (GB)
PUT_FAILED	17	0.1
STAGEOUT	5	0.0
STAGED	176433	3962.3
STAGEIN	3	0.0
All	176458	3962.4

There are NO migration candidates currently

Number of Garbage Collectable files for this pool is 176433 with total data volume of 3962.3 GB. Number of migratable files for this pool is 0 with total data volume of 0.0 GB.

Distribution by status for zero length files

Status	No. of files
STAGEOUT	1
STAGED	133
STAGEIN	3

Statistics on files currently in the stager catalog

- o Total number of STAGED files is: 176433 with total volume 3962.3 GB.
- o Average lifetime of STAGED files is: 32.14 days.
- o Average number of accesses for STAGED files: 7.51.
- o No. files accessed in last 24 hours: 9079.
- o No. files written in last 24 hours: 4656.
- o Average size of STAGED files is: 23.0 MB.
- o Last file accessed: 0.00 days ago.

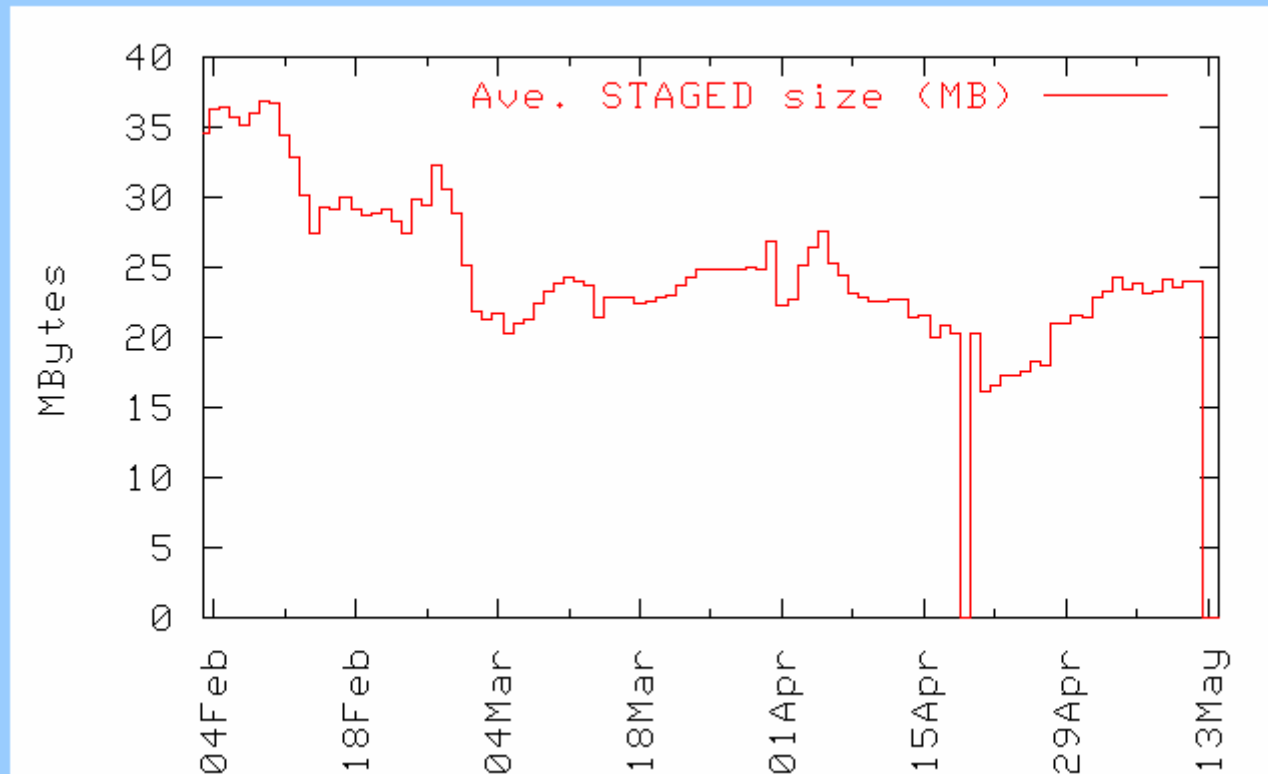
*Files in this DiskPool
by status. Source: stageqry*

Basic file statistics

23 MB!

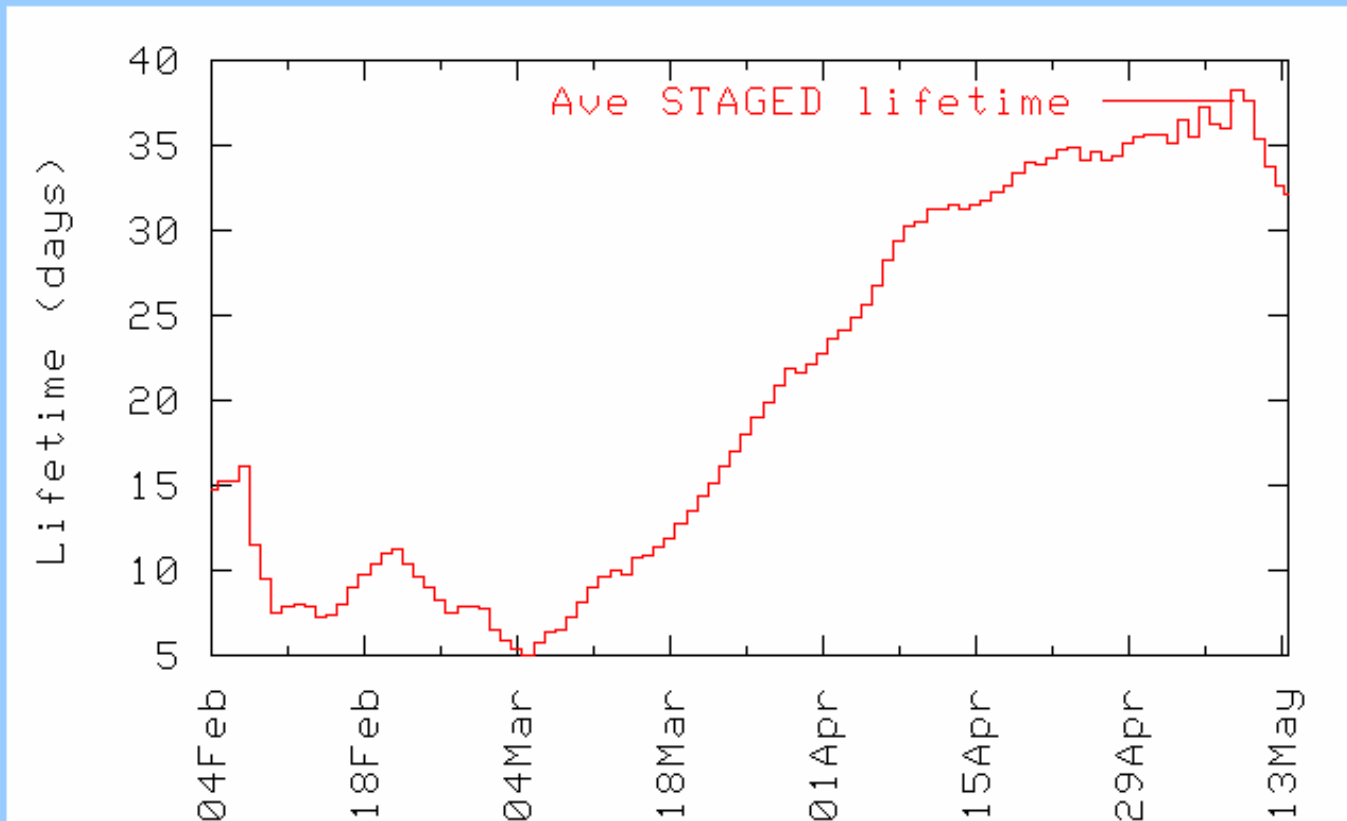
Stager DiskPool: time evolution

Fig 3. Daily evolution of average STAGED file size (MB) (snapshot each midnight)



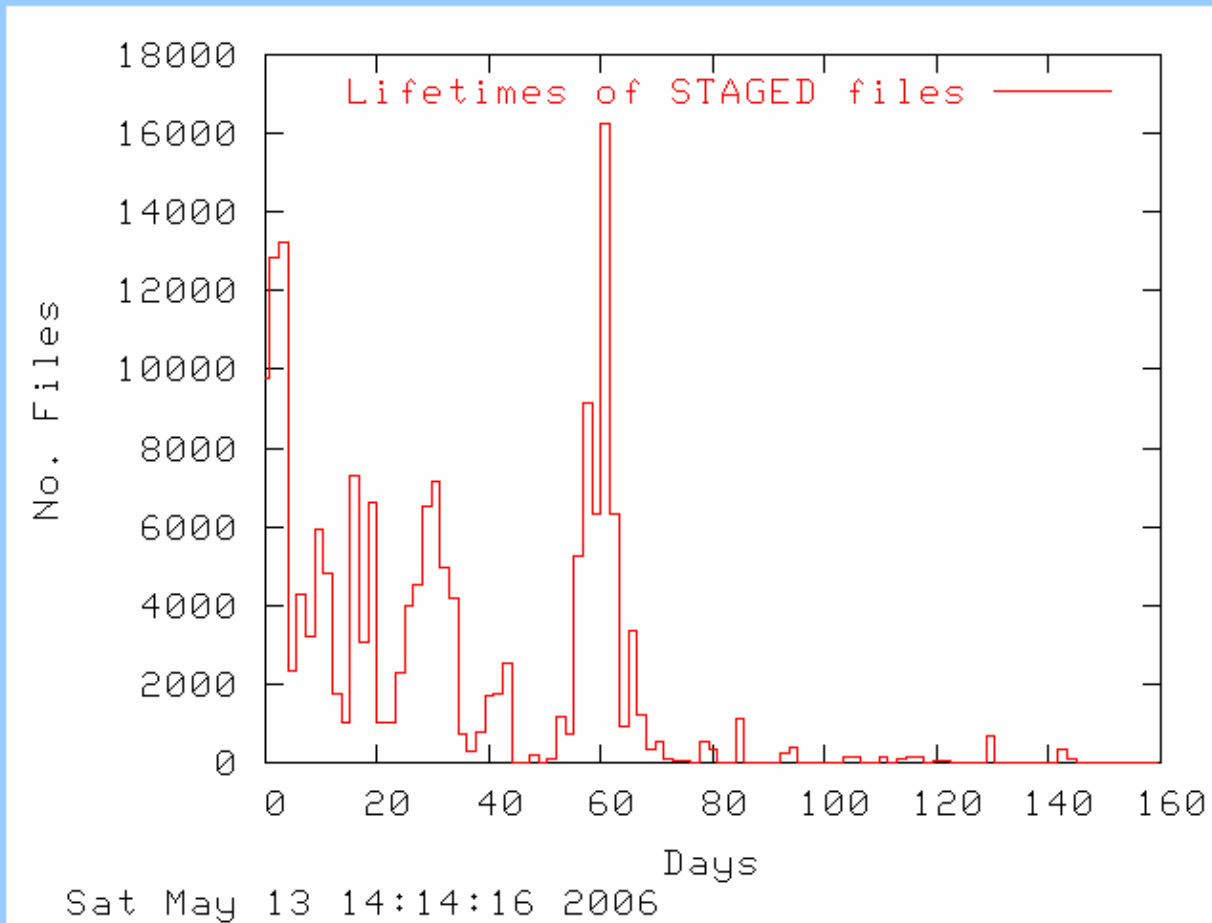
Stager DiskPool: time evolution

Fig 4. Daily evolution of average Lifetime of STAGED files (snapshot each midnight)



Stager DiskPool: current distribution

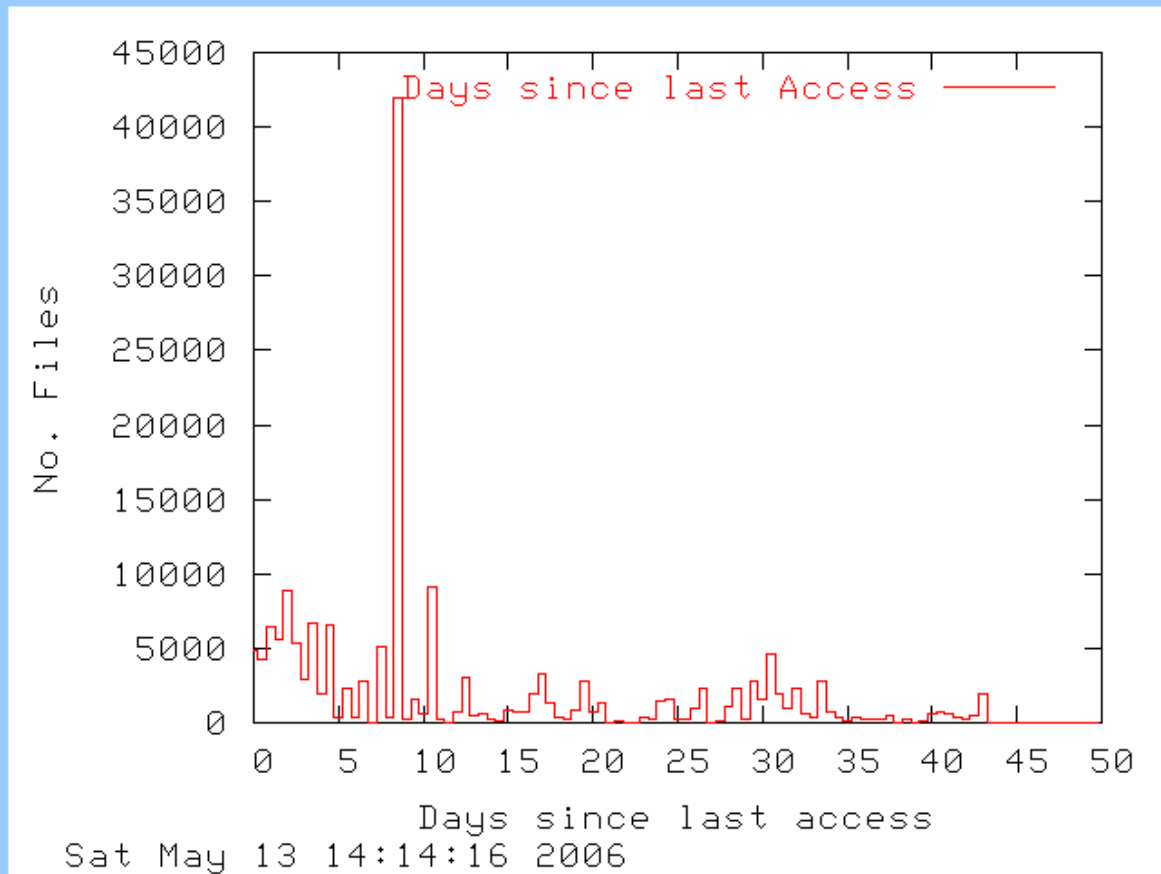
Fig 5. Current lifetime distribution of STAGED files



File lifetimes: minimum: 0.00, maximum: 144.69 and average: 32.14 (days)

Stager DiskPool: current distribution

Fig 6. Days since last Access of STAGED files



Total contents of above histogram is : 176433.00

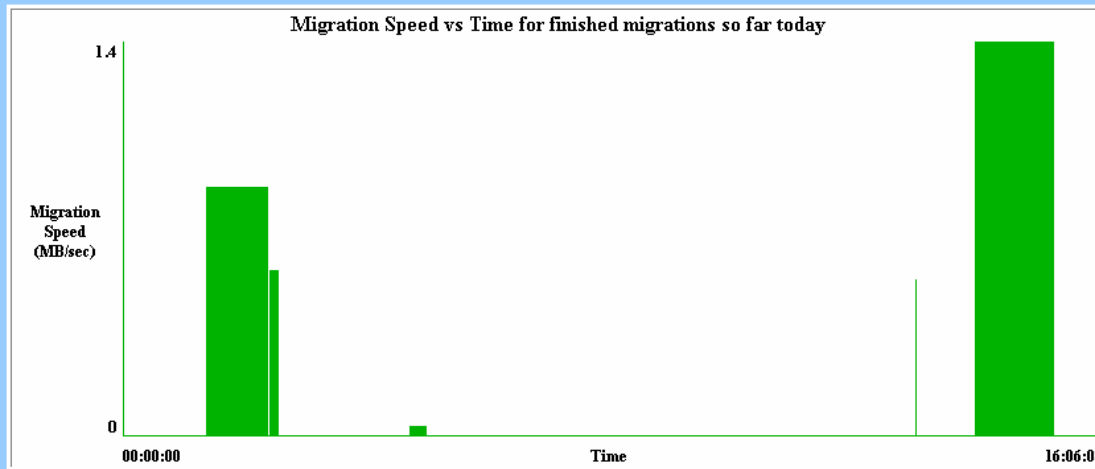
Days since last access: minimum: 0.00, maximum: 43.51 and average: 12.67

Migrator information

CASTOR Stager survey

Results for stager: [stagen49](#) migrator: [na49migr](#) so far today

Data is extracted every 30 mins, last on Sun May 14 16:06:08 CEST 2006 . <<< Please note this date/time



Migrator speeds/data volumes

If files ~> 1GB are used then we get close to the TapeDrive streaming speed (~ 30MB/sec); otherwise....

The details of the individual migrations are given in the table below. Migration speed = (data volume migrated/migration time) where the latter includes drive wait time, tape mount time etc. times between disk and tape are a function of file size.

Click [here](#) to see the evolution with time of a) the number of migration candidates and b) the free space of the diskpools that this migrator services.

Details of individual migrations. N.B.:

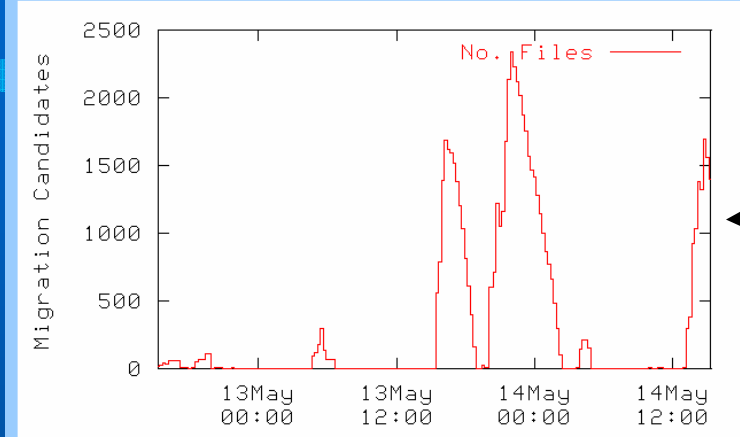
- Stream speed includes wait time for tape server availability and tape mount.
- Migrator speed includes wait time for tape server availability and tape mount(s) for each stream.

Migrator pid	Migrator Start/Stop	Time (s)	No. of streams	Stream data	No. files	Data Vol. GB	Migrator Speed (MB/s)	File class Breakdown		
								File class	No. files	Data Vol. MB
14465	05/14 01:20:03 05/14 02:21:35	3692	1	#1 3691s 742 files 3298 M 0.9 M/s last write 05/14 02:21:34 (file 742) P08928@tpsrv121	742	3.22	0.9	95	742	3297.3
14730	05/14 02:21:35 05/14 02:31:36	601	1	#1 599s 78 files 356 M 0.6 M/s last write 05/14 02:31:34 (file 78) P08928@tpsrv110	78	0.35	0.6	95	78	356.1
15060	05/14 03:40:06 05/14 03:41:50	104	1	#1 102s 1 files 0 M 0.0 M/s last write 05/14 03:41:48 (file 1) P08928@tpsrv116	1	0.00	0.0	95	1	0.0

Detailed information on streams and tapes

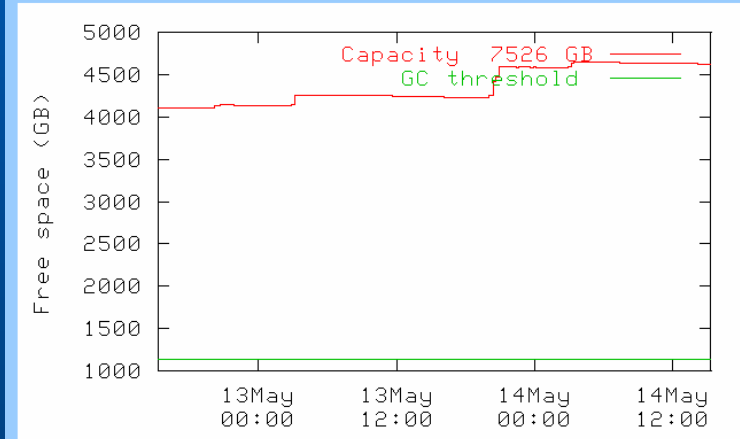
Migration candidates: FreeSpace Evolution

No. of migration candidates for migrator `na49migr`. Data points from May 12 17:02 to May 14 16:31. (Every 15 minutes)



Number of files to be migrated. (been as high as 150k!)

FreeSpace for diskpool `na49_dsl`. Data points from May 12 17:02 to May 14 16:31. (Every 15 minutes)

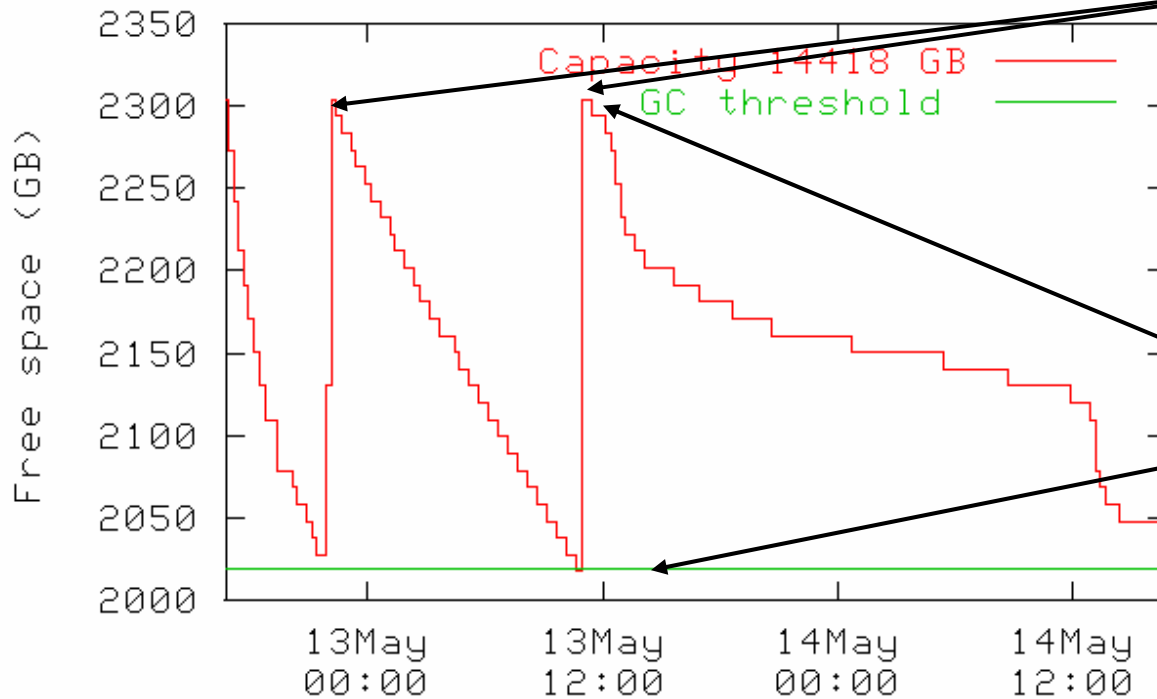


FreeSpace evolution (no GC ran in this time period)

FreeSpace for diskpool `na49_keep`. Data points from May 12 17:02 to May 14 16:31. (Every 15 minutes)

Migration candidates: FreeSpace Evolution

FreeSpace for diskpool [na48](#). Data points from May 12 16:46 to May 14 16:46. (Every 15 minutes)



Garbage collector starts

Garbage collection limits

Machine management – by groups

Machine breakdown by 'refined' CDB cluster

- The 'castor' cluster has been sub-divided into: 'production service' (in /etc/STGCONFIGs), 'fio service' (test and migr clusters) and 'other' i.e. the rest.
- The 'tapeserver' cluster has been sub-divided into: 'production service' (according to /afs/cern.ch/project/castor/TOOLS/drive.tapeserver.list SMS state is 'production'), 'tape (test service)' and 'tapeserver (spares)' (no drives).
- The 'fileserver' cluster has been sub-divided into: 'spare', 'spare-castor2', 'spare-castor1', 'zombie' and 'other' i.e. the rest.
- The 'castorgid' cluster has been sub-divided into: 'production service' (function = gridftp+sm+thtar) and other

Cluster	Total number of machines	Number of machines NOT of production quality
c2alice	17	0
c2cms	27	2
c2lhcb	20	2
c2sc4	48	5
castor (fio production service)	22	1
castor (other)	9	5
castor (production service)	139	2
castor2	56	54
castor2_test	1	0
castorgid (production service)	8	0
dbserver	108	30
fileserver (other)	49	16
fileserver (spare)	25	24
fileserver (spare-castor1)	9	8
fileserver (spare-castor2)	55	29
fileserver (zombie)	2	N/A
gridfts	14	6
gridlfc	18	3
itrac	47	8
sm	21	12
tape (production service)	131	0
tape (test service)	5	5
tapeserver (spares)	19	14
Total	850	226

*High level view of monitoring and configuration information for ~850 machines
Data sources: Cdb, Castor confign. (stageqry), Lemon, Landb, SMS and local db for state transition detection. The overall consistency between the sources is continuously checked.*

All castor1 machines OK (just 2 minor problems)

Transitions between significant states (no contact, high-load, daemons wrong, machine added/removed to/from each Cdb cluster,) are also signalled by e-mail to person(s) responsible for that cluster

These are in use now in castor1 staggers. Group castor (other): in standby/repair

Machine management – by groups

FIO/DS machines by cluster

Machines in cluster: castor

Data extracted on Sun May 14 17:00:51 2006 (every 15 mins). Page refreshed on Sun May 14 17:00:54 CEST 2006

 + OK in MSA field means /home/operator/nomorealarms file exists
 + OK in MSA field means at least one MSA sensor has died
 in the **model** field means /etc/nospma file exists
 in the **sub_cluster** field means information came from ADC/DES database
 in the **Warranty Expiry** field means machine was tampered with warranty thus invalid
 in the **O/S (kernel)** field means Lemon and CDB values of kernel disagree.
 in the **HostName** field means part of the production service.

This page shows all machines in cluster 'castor'. Click [here](#) to see ONLY machines NOT of production quality in this cluster.

HostName	sub_cluster	IPService	Contract::Importance	Function	CDB O/S (kernel)	SMS state	Model	MSA status (Lemon)	function comment	Warranty Expiry
bfsrk4508	bfsrk4508	SS13-C-IP66	D: 70	stager+fileserver	slc3 (2.4.21-40.EL.cernsm)	production	e5	OK		25-FEB-2008
castorsrv1	server	SS13-V-IP49	E: 70	nameserver	slc3 (2.4.21-40.EL.cernsm)	production	SEIL_2.4GHz	OK		10-APR-2006
castorsrv2	server	SS13-V-IP48	E: 70	nameserver	slc3 (2.4.21-40.EL.cernsm)	production	SEIL_2.4GHz	OK		10-APR-2006
castorsrv3	server	SS13-V-IP48	E: 70	nameserver	slc3 (2.4.21-40.EL.cernsm)	production	SEIL_2.4GHz	OK		26-NOV-2006
castorsrv4	server	SS13-V-IP49	E: 70	nameserver	slc3 (2.4.21-40.EL.cernsm)	production	SEIL_2.4GHz	OK		10-APR-2006
aleph005d	stagealeph	SS13-C-IP67	D: 50	fileserver	slc3 (2.4.21-40.EL.cernsm)	production	a1	OK		05-JUL-2002
aleph007d	stagealeph	SS13-C-IP73	D: 50	fileserver	slc3 (2.4.21-40.EL.cernsm)	production	a1	OK		05-JUL-2002
aleph008d	stagealeph	SS13-C-IP61	D: 50	fileserver	slc3 (2.4.21-40.EL.cernsm)	production	e1	OK		14-OCT-2007
lfs5501	stagealeph	SS13-V-IP25	D: 50	fileserver	slc3 (2.4.21-40.EL.cernsm)	production	b1	OK		01-FEB-2002
lfs5624	stagealeph	SS13-V-IP25	D: 50	fileserver	slc3 (2.4.21-40.EL.cernsm)	production	e3	OK		01-JUL-2008
stage007	stagealeph	SS13-C-IP4	D: 70	stager	slc3 (2.4.21-40.EL.cernsm)	production	stager-1200	OK		04-NOV-2007
lfs6103	stageatlas	SS13-V-IP43	D: 50	fileserver	slc3 (2.4.21-40.EL.cernsm)	standby	e4	MIRROR_BROKEN	fs needed	01-FEB-2009
stage015	stageatlas	SS13-V-IP36	D: 70	stager	slc3 (2.4.21-40.EL.cernsm)	production	stager-2600	OK		01-JAN-2008
lfs6004	stagecdr	SS13-V-IP44	D: 50	fileserver	slc3 (2.4.21-40.EL.cernsm)	production	e4	OK	fs needed	01-FEB-2009
lfs6022	stagecdr	SS13-V-IP44	D: 50	fileserver	slc3 (2.4.21-40.EL.cernsm)	production	e4	OK	fs needed	01-FEB-2009
stage003	stagecdr	SS13-C-IP4	D: 70	stager	slc3 (2.4.21-40.EL.cernsm)	production	stager-1200	OK		04-NOV-2007
lfs5522	stagechorus	SS13-V-IP30	D: 50	fileserver	slc3 (2.4.21-40.EL.cernsm)	production	e3	OK		01-JUL-2008
stage012	stagechorus	SS13-C-IP4	D: 70	stager	slc3 (2.4.21-40.EL.cernsm)	production	stager-1200	OK		04-NOV-2007

View only problem cases

Data organized by Cdb cluster and sub_cluster

Colour coding indicates problems/issues

Links to Lemon/ITCM Remedy feed

Not in production

Link to a comparison of all machines of this model

Link to see all machines on a given switch

vmgr and Robotic library comparisons

CASTOR Robot Inventory comparison with VMGR

- Pages refreshed on Sun May 14 00:25:13 CEST 2006 (daily at this time).
- Silo data is extracted (by Charles) once per day at about 00:18.
- VMGR information is extracted when the script runs.
- Drive information is taken from config file /afs/cern.ch/project/castor/TOOLS/drive.tapeserver.list).
- The Robot inventories for ACS4 and ACS5 are saved weekly into the Castor directory /castor/cern.ch/c3/silo_inventories.
- Tape volumes starting with CL ST CD are IGNORED both in VMGR and in the robot inventories (cleaning or other special tapes).

Definition of Tape Volume Classes (as given by /afs/cern.ch/project/castor/TOOLS

Leading Character(s)	Class
A	BACKUP
P	9940
I	3592B
L	LTO3
T	T10K

Tape 'type': VID definition (convention)

Robot information: distribution of tape volumes and tape drives by silo in ACS4/6

Date of Robot inventory data (/afs/cern.ch/project/castor/TOOLS/output/STK_ACS4.inventory.data):

Class\Silo	0	1	2	3	4	Total
9940	139	0	1976	3778	0	5893
BACKUP	2	0	768	0	0	770
All Vols	141	0	2744	3778	0	6663
9940 drives			1	11		

Robot information: distribution of tape volumes and tape drives by silo in ACS5/7

Date of Robot inventory data (/afs/cern.ch/project/castor/TOOLS/output/STK_ACS5.inventory.data):

Class\Silo	0	1	2	3	4	Total
9940	0	0	60	4556	5381	9997
BACKUP	0	0	841	0	0	841
All Vols	0	0	901	4556	5381	10838
9940 drives				11	11	

Robot information: distribution of tape volumes and tape drives by silo in SL8500_0 (total of 6745 volumes)

Date of Robot inventory data (/afs/cern.ch/project/castor/TOOLS/output/SL8500_0.inventory.data) is Sun May 14 00:17:13 2006.

Class\Silo	0	1	2	3	4	5	6	7	Total
9940	909	987	1645	270	305	250	356	272	4994
LTO3	0	1	8	160	0	0	0	81	250
T10K	726	621	0	0	56	87	5	6	1501
All Vols	1635	1609	1653	430	361	337	361	359	6745
LTO3 drives				5				3	
T10K drives	10	10			10	10			
9940 drives			5				5		

Layout of drives and tape media within Robot silos

Robot information: distribution of tape volumes and tape drives by frame in IBM_LIB1 (total of 1552 volumes)

Date of Robot inventory data (/afs/cern.ch/project/castor/TOOLS/output/IBM_LIB1.inventory.data) is Sun May 14 00:19:13 2006.

Class\Fram	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Total
3592B	0	173	387	400	400	191	0	0	0	1	0	0	0	0	0	0	0	0	1552
All Vols	0	173	387	400	400	191	0	0	0	1	0	0	0	0	0	0	0	0	1552
3592 drives					14			10		10			10						

Overview of Robot Inventory/VMGR comparison

Data source\Library	ACS4/6	ACS5/7	IBM_LIB1	ACS8500
Robot Inventory	6663	10838	1552	6745
VMGR	6662	10839	1580	6744
Common volumes	6661	10838	1552	6744

Agreement between vmgr and Robot libraries inventories

Data volumes per Experiment

CASTOR files; Numbers/Data Volumes of Files and Directories per Experiment

Click [here](#) to see breakdown by fileclass.
 Click [here](#) to see fileclass definitions.

Script runs daily at this time; last update on Sun May 14 02:18:03 CEST 2006.

Experiment	Fileclass(es)	Number of Directories	Number of Files	Total File Size (GB)	Average File Size (MB)
General Expt Copy	1	130218	3571545	0.496 PB	149
Temporary Files	58	870	2302416	34321	15
afsbac	104	49193	1301544	96351	76
aleph	3 12 13 14 61 62	5049	261686	42210	165
alice	7 23 24 78 117 121 126	282771	2783509	296932	109
ams	82	1851	201543	19088	97
atlas	33 34 59 115 138 139	48848	5263229	226184	44
cast	71 72	478	167767	2053	13
na45	54	4956	285832	25778	92
chorus	5 75 83	1598	136362	16358	123
cms	8 30 31 32 51 108 136	28661	3159265	326923	106
compass	6 74 85 86 87 88 89 90 91 92 93 94	3266	6052471	1.188 PB	211
delphi	4 67 76	3137	314549	20818	68
dirac	41	216	80506	7540	96
engineering	118	1578	165303	16290	101
fio	26 27 68 106 113 114 116 123 128 131 135	17707	659633	56042	87
grid	65	1168	308430	61196	203
harp	11 16 17 18 19 20 21 45 46 64 96	134	104826	54255	530
Experiment	Fileclass(es)	Number of Directories	Number of Files	Total File Size (GB)	Average File Size (MB)
isolde	80 81	121	8391	792	97
l3	43 52 56 57	12055	217062	19685	93
lhcb	9	15706	1720041	209196	125
misc	44 53	1391	70019	6587	96
na48	35 38 40 70 97 98 99 100	11265	4546375	0.651 PB	154
na49	10	352	42589	8440	203
na60	48 49 109 110 111 137	1287	77952	13065	172
nap	102 125	2804434	2387942	14999	6.4
named	77 120	162	108540	5572	52

~50,000,000 files in ~5 PetaBytes
 (~100 MB average filesize)

These files are from stageslap - which once had 150k migration candidates

VIDs in TapePools

Link to evolution of TapePool size with time

CASTOR Tape Pools survey

TapePool [cmsfamily_new1](#) breakdown by date entered and library. Ordering is by vmgr status, then by vid.

- Date of refresh: Sun May 14 16:55:27 CEST 2006 (every 30 mins).
- Only 9940, 3592B, LTO3 and T10K tapes are considered here.
- Number of 9940, 3592B, LTO3 and T10K tapes in pool is 49 (33 9940s, 0 3592B's, 0 LTO3's and 16 T10K's)
- **User restriction for this pool is: stage group restriction is: st.**
- Pool capacity is: 14.26TB. Free space is 2.05TB
- EOT file count is the last file number on the tape. This is >= the number of active files (segments).
- No. active files was correct on Sun May 14 02:33:05 2006
- Click on VID to see VMGR information + RTCOPY errors.
- VID link shows ALL RTCOPY errors whereas number of RTCOPY errors quoted in column 11 is 'filtered' (those associated with the tape vol)

LHC experiment castor2 pools recently moved to new media

Links to vmgr information and RTCOPY errors

VID	Library	Date Entered in vmgr	Last Write date	Write count	Last Read date	Read count	No. active files	EOT file Count	Free space at eot (GB)	No. RTCOPY errors (filtered)	vmgr status
T00175	SL8500_1	10:33 28/02/06	15:14 03/03/06	1	unknown	0	0	0	500.00	0	AVAILABLE
T00176	SL8500_1	10:33 28/02/06	15:16 03/03/06	1	unknown	0	0	0	500.00	0	AVAILABLE
T00177	SL8500_1	10:33 28/02/06	15:18 03/03/06	1	unknown	0	0	0	500.00	0	AVAILABLE
T00952	SL8500_0	10:34 28/02/06	16:51 14/05/06	14	unknown	0	0	85	95.44	0	AVAILABLE
T00953	SL8500_0	10:34 28/02/06	18:03 06/03/06	1	unknown	0	0	0	500.00	0	AVAILABLE
P06885	STK_ACS5	unknown	19:37 19/03/06	16	22:02 21/04/06	306	1251	1251	0.09	0	FULL
P06889	STK_ACS5	unknown	00:51 22/03/06	18	08:20 22/04/06	378	1261	1261	0.07	0	FULL
P07127	STK_ACS5	unknown	22:00 26/03/06	35	14:06 22/04/06	218	1484	1484	0.18	0	FULL
P07753	STK_ACS5	unknown	04:15 27/03/06	127	12:44 09/04/06	651	1700	1700	0.04	0	FULL
P08044	STK_ACS4	09:52 27/05/03	16:53 11/03/06	299	13:16 27/04/06	497	1487	1488	0.07	0	FULL
P08054	STK_ACS4	09:52 27/05/03	10:51 19/03/06	57	13:25 27/04/06	830	1309	1309	0.02	0	FULL
P08080	STK_ACS4	09:52 27/05/03	15:25 27/02/06	3	06:26 03/04/06	110	1698	1698	0.06	0	FULL
P08226	SL8500_0	unknown	20:25 23/01/06	483	13:58 14/05/06	121	4850	4861	0.35	0	FULL
P08334	STK_ACS4	09:25 24/05/03	20:57 08/02/06	351	13:14 22/04/06	585	1515	1527	0.04	0	FULL
P08358	STK_ACS4	09:25 24/05/03	18:59 14/02/06	339	13:04 22/04/06	664	1512	1533	0.06	0	FULL
P08405	STK_ACS5	10:52 27/05/03	23:26 22/03/06	14	11:29 22/04/06	1204	1431	1431	0.00	0	FULL
P08412	STK_ACS4	09:25 24/05/03	18:19 17/02/06	249	17:07 22/04/06	487	1591	1592	0.19	0	FULL
P08423	STK_ACS5	10:52 27/05/03	03:31 29/04/06	490	09:30 26/04/06	1346	4160	4234	1.22	0	FULL
P08432	STK_ACS4	09:25 24/05/03	05:53 18/02/06	260	15:42 22/04/06	1299	1571	1571	0.15	0	FULL
P08438	STK_ACS4	09:25 24/05/03	16:02 22/02/06	300	11:43 26/04/06	548	1589	1589	0.00	1	FULL
VID	Library	Date Entered in vmgr	Last Write date	Write count	Last Read date	Read count	No. active files	EOT file Count	Free space at eot (GB)	No. RTCOPY errors (filtered)	vmgr status
P08441	STK_ACS5	10:52 27/05/03	15:25 27/02/06	1	05:21 05/04/06	587	1699	1699	0.18	0	FULL
P08449	STK_ACS5	10:52 27/05/03	03:05 03/03/06	93	15:54 10/05/06	1033	3647	3647	0.14	0	FULL
P08654	STK_ACS4	unknown	05:56 17/03/06	395	13:57 14/05/06	1120	3615	3622	0.11	0	FULL

Supply TapePools

Castor TapePools fed from supply_pools via specified filling policies

CASTOR Tape Supply Pool Survey

Survey of tape supply pools on Sun May 14 16:00:12 CEST 2006

Scans are done at 0800 and 1600 daily

A log of tapes allocated from the supply pools to Castor TapePools is available [here](#)

Summary of the overall situation:

Supply Pool	No. Volumes remaining	No. of days supply left (Based on last month's consumption)
supply_994BR0	215	in 102 days (i.e. at the rate of 2.1 per day, about Aug 24 2006)
supply_994BR5	464	in 71 days (i.e. at the rate of 6.5 per day, about Jul 24 2006)
supply_994BR4	227	in 50 days (i.e. at the rate of 4.5 per day, about Jul 3 2006)
supply_3592B1	1011	
supply_3592B2	134	in 89 days (i.e. at the rate of 1.5 per day, about Aug 11 2006)
supply_LTO3R0	22	
supply_LTO3R1	0	
supply_T10KR0	473	
supply_T10KR1	117	in 24 days (i.e. at the rate of 4.8 per day, about Jun 7 2006)

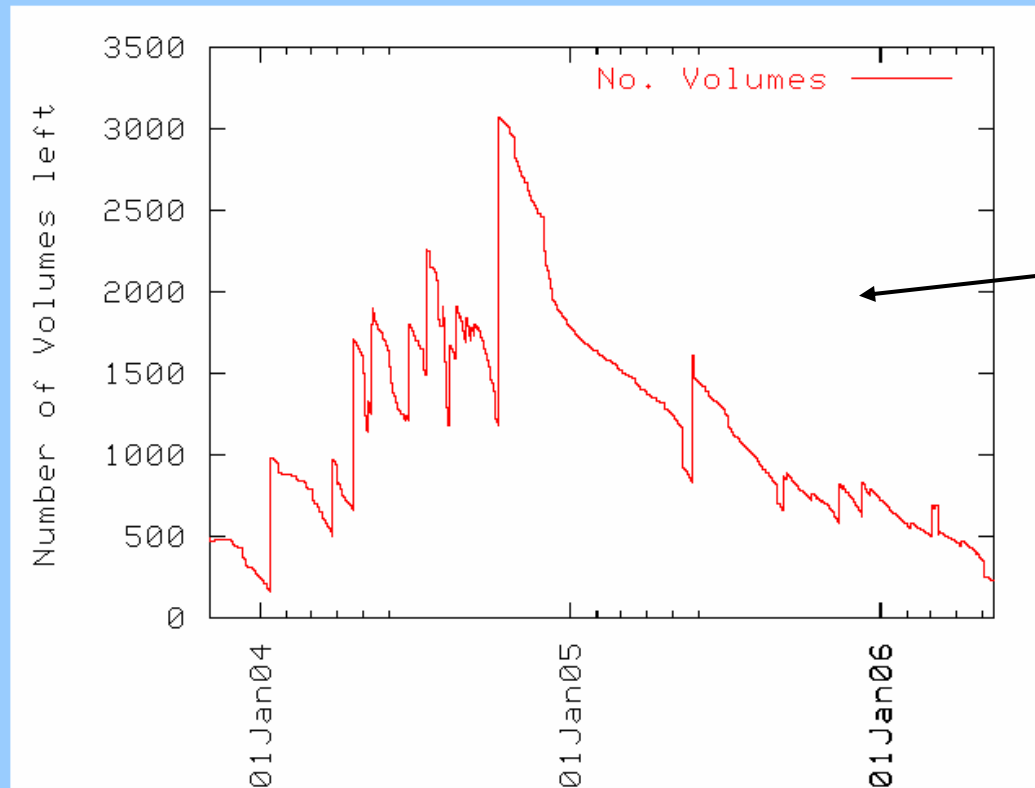
Prediction of lifetimes of current supply pools based on last month's consumption

SupplyPool monitoring

Fig 3. Pool supply_994BR4 currently has 227 volumes.

Time Evolution of number of *available* volumes left in supply pool: supply_994BR4

Data from Mon Mar 20 16:00:21 2006 to Sun May 14 16:00:12 2006



*Tape purchases, repacking exercises,
CDR and normal tape consumption
(typical background rate
of ~1-2 TeraBytes/day)*

Largest tape consumers

Top 10 T10K consuming TapePools

Data extracted on Sun May 14 17:00:28 CEST 2006

T10K Transfers made in last 24 hours

A total of 64 tapes were transferred in this time period

Tape Pool	Vid's transferred	% of total vids transferred
lhcb_new	52	81
cmsfamily_new1	12	19

T10K Transfers made in last 168 hours

A total of 212 tapes were transferred in this time period

Tape Pool	Vid's transferred	% of total vids transferred
lhcb_new	196	92
cmsfamily_new1	12	6
cmsfamily_new4	4	2

T10K Transfers made in last 31 days

A total of 324 tapes were transferred in this time period

Tape Pool	Vid's transferred	% of total vids transferred
lhcb_new	288	89
cmsfamily_new1	20	6
cmsfamily_new4	8	2
cmsfamily_new2	4	1
cmsfamily_new3	4	1

In few weeks lhcb has written 150 TB of data onto T10K media ()*

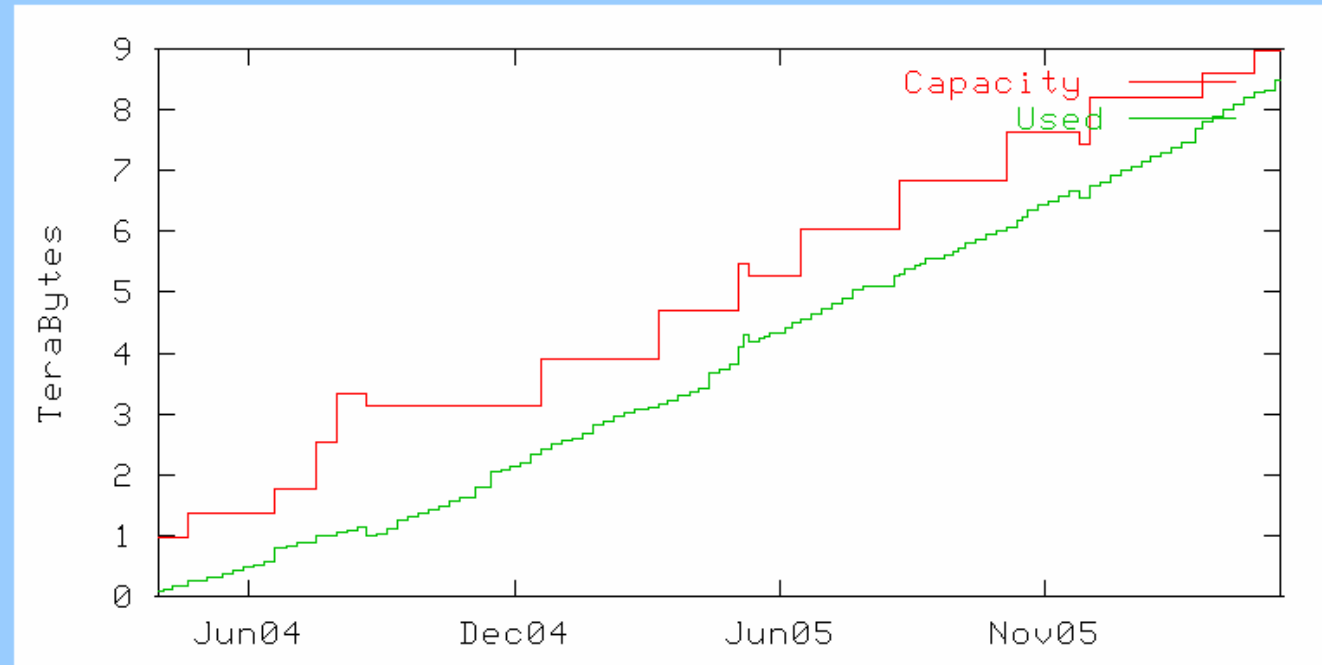
•In fact this was not the case. We had problems with new media processing by castor.

Evolution of tape consumption by Engineers

CASTOR Tape Pool Capacity/used space evolution

Evolution with time of capacity and used space for Castor TapePool: [engineering1](#)

- Oldest data from 2003, last extract on Mon May 8 11:31:14 CEST 2006. **Last data from Mon May 8 11:30:21 2006.**
- Weekly vmgrlistpool data is stored in /afs/cern.ch/project/castor/TOOLS/output/vmgrlistpool.
- Castor filecopy in /castor/cern.ch/c3/vmgrlistpool. Files are named ~ Mon_dd_yyyy.
- Data sampled weekly.



Backup/archives of mechanical and electrical engineering data

TapeMount Statistics

CASTOR TapeMount distributions for week 0618

Page extracted on Mon May 8 02:58:25 CEST 2006

Statistics per dgn for week 0618

dgn	Total Mounts	Read Mounts	Write Mounts	Total vol (TB)	Read vol (TB)	Write vol (TB)	Total Files	Total Files read	Total Files written	Total drive hours	Total read hours	Total write hours	# files per read mount	# files per write mount	Ave. read file size (MB)	Ave. write file size (MB)
3592B2	784	189	595	2.18	0.23	1.94	16141	249	15892	46	6	40	1.3	26.7	985.7	128.2
994BR0	11828	10156	1672	14.12	10.74	3.38	89815	43760	46055	777	620	158	4.3	27.5	257.4	77.0
994BR4	16488	14609	1879	17.12	11.73	5.38	163504	40697	122807	1009	738	271	2.8	65.4	302.3	46.0
994BR5	33234	26541	6693	33.82	21.66	12.16	341189	78894	262295	1989	1308	681	3.0	39.2	287.9	48.6
LTO3R0	776	605	171	40.14	1.82	38.32	28335	1958	26377	392	39	353	3.2	154.3	973.3	1523.4
T10KR0	1704	153	1551	109.20	16.90	92.30	134651	13080	121571	862	104	758	85.5	78.4	1354.9	796.1
Total	64814	52253	12561	216.57	63.09	153.49	773635	178638	594997	5076	2816	2260	3.4	47.4	370.3	270.5

Statistics per TapePool (>= 100 mounts)

N.B. Links in 'No. Mounts' column are invalidated after 1 year.

TapePool	No. Mounts	No. read Mounts	No. write Mounts	No. files read	No. files written	Total Mount Time (Hours)	Total Volume (GB)	# files per read mount	# files per write mount	Ave. read file size (MB)	Ave. write file size (MB)
lhcb	16532	16387	145	42134	45216	1010	14039	2.6	311.8	227.5	105.9
largeuser	10586	9826	760	42249	66187	640	11708	4.3	87.1	231.9	33.1
user_new	6891	2066	4825	3745	195185	534	7215	1.8	40.5	261.0	32.8
atlas2	3863	3829	34	6541	665	135	732	1.7	19.6	103.7	106.8
atlas1	3847	3799	48	6863	838	138	827	1.8	17.5	106.8	135.7
compass2	3746	3741	5	6574	34	187	6008	1.8	6.8	931.2	883.7
compass	3618	3613	5	6499	35	178	5936	1.8	7.0	930.4	920.0
default	2541	2532	9	3251	21	91	467	1.3	2.3	144.6	406.3
atlas_new2	2024	222	1802	319	25587	121	2674	1.4	14.2	186.8	104.7
atlas_new1	1971	238	1733	326	22017	115	2305	1.4	12.7	167.2	104.7
cms	1466	1405	61	21884	1036	146	5083	15.6	17.0	234.7	65.9
lhcb_new	1337	2	1335	1378	67855	199	8470	689.0	50.8	138.8	125.0
grid1	992	992	0	2534	0	53	544	2.6	0	219.7	0
compassdst1	798	719	79	10880	17562	159	3033	15.1	222.3	232.6	32.7
itdc_LTO3_new	718	587	131	587	22917	347	37943	1.0	174.9	1662.4	1652.8
cmsfamily_new1	685	27	658	27	4914	31	320	1.0	7.5	223.6	65.4
compassdst2	481	399	82	2010	17572	57	1048	5.0	214.3	254.9	31.9

Source: weekly accounting

TapeMount Statistics

Mounts per experiment/affiliation for week 0618 (>= 100 mounts)

Experiment (Affiliation)	No. Mounts	No. read Mounts	No. write Mounts	No. files read	No. files written	Total Mount Time (Hours)	Total Volume (GB)	# files per read mount	# files per write mount	Ave. read file size (MB)	Ave. write file size (MB)
lhcb	17869	16389	1480	43512	113071	1209	22508	2.7	76.4	224.7	117.4
atlas	11705	8088	3617	14049	49107	509	6538	1.7	13.6	108.6	105.3
large user files (largeuser1/largeuser2)	10586	9826	760	42249	66187	640	11708	4.3	87.1	231.9	33.1
compass	8643	8472	171	25963	35203	581	16025	3.1	205.9	585.9	34.0
others	7311	2286	5025	8183	247767	710	18090	3.6	49.3	452.2	59.8
user files (default/default2)	2541	2532	9	3251	21	91	467	1.3	2.3	144.6	406.3
cms	2212	1468	744	21988	6130	180	5447	15.0	8.2	235.1	66.4
itdc	1181	724	457	10471	74794	994	137195	14.5	163.7	1631.9	1649.9
grid	992	992	0	2534	0	53	544	2.6	0	219.7	0
harp	489	464	25	533	964	22	634	1.1	28.6	1150.5	37.6
na48	356	356	0	2541	0	32	1562	7.1	0	629.4	0
delphi	272	269	3	1392	4	14	75	5.2	1.3	52.6	799.0
sd3images	153	152	1	1409	1	14	346	9.3	1.0	251.6	3.0
cast	152	31	121	81	280	8	248	2.6	2.3	1624.9	435.3
l3	121	121	0	121	0	3	22	1.0	0	186.8	0
Total	64814	52253	12561	178638	594997	5076	221771 216.6 (TB)	3.4	47.4	370.3	270.5

NB: totals include all experiments (and affiliations)

Don't be misled by itdc - this is 'us'

We can do much more than this per week with the currently installed equipment

Few files per read mount

Many files per write (stager migration)

Summary

- **Highly visual overview, plus by email notifications, allowed management of ~50 castor1 stagers, 200 disk servers, ~200 TapePools, ~200 Fileclasses, ~100 TapeServers, ~10**3 VIDs**
- **A lot (most?) work went into the nitty gritty of Tape media and TapeDrive management..**
- **~250 scripts of 50k lines of perl/sql.**
- **~1.5 man-years dev+maintenance over 3 years.**

Acknowledgements

- **Charles Curran:** for provision of tape related information (serial numbers, Silo inventories, current state of media on robotic drives) etc.
- **Jean-Damien Durand:** for patient answers to endless questions concerning stager workings.