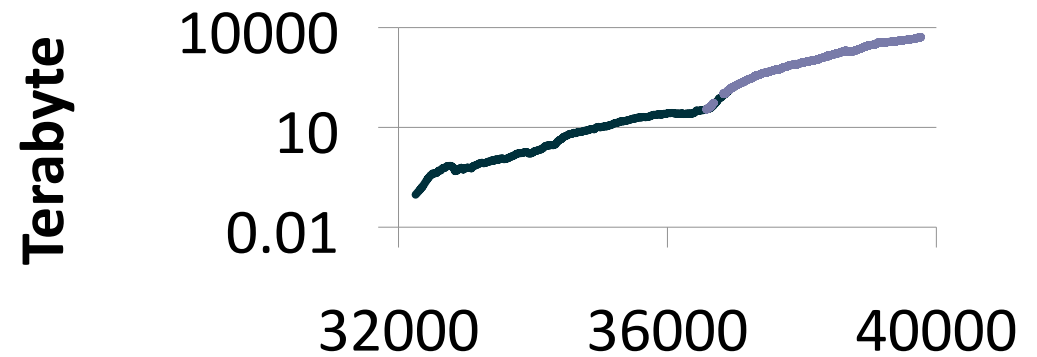


## CSIRO ASC Data Store Total holdings (log scale)



# CSIRO ASC Monitoring and reporting on DMF Usage

Robert C. Bell | Technical Services Manager  
4 December 2012

# CSIRO ASC

- DMF, including data protection by backup
  - Monitoring tools
  - Reports
  - Graphs
- What the users can see
- What systems staff can see
- What is reported to a wider audience, e.g. management
- Case studies – anomalies detected through the monitoring

# CSIRO ASC

- Host cherax – UV 1000
- DMF-managed file systems:
  - /datastore
    - User access: \$HOME
    - 29 Tbyte (hybrid)
  - /backup1, /backup2, /backup3
    - No user access: – targets for backups
    - $\frac{3}{4}$  Tbyte each
- Up to 4 levels: primary, DMF cache, MAID, tapes (2 copies)



# DMF Reporting – for users

- HPCbull items
- Userguides
- Traffic lights
- quota
- dmget --list
- WWW pages: holdings
- WWW pages: backup status



# DMF Reporting – HPCbull items

## 1. CSIRO ASC - Our new MAID system

On 11th September, we brought a MAID (Massive Array of Idle Discs) into operation on the ASC Data Store.

This provides an extra layer of storage between the primary disc area and the tapes, providing faster recalls than from tape. Typical access times for the first file from a disc set are about 20 seconds, with subsequent recalls from the same disc set starting within milliseconds. This contrasts with tapes, where the first access typically takes a minute, and subsequent accesses half a minute.

The MAID has 870 Tbyte of capacity.

CSIRO staff can see the Monday Mail article at <http://goo.gl/WLiHH> which also highlights the 'green' aspects of the MAID in the Data Store hierarchy.



# DMF Reporting – Userguides: Data Handling

Added by [Williams, Gareth \(CSIRO IM&T, Docklands\)](#), last edited by [Ludbey-Bruhwei, Joel \(CSIRO IM&T, Docklands\)](#) on Jun 18, 2012 ([view change](#))

- Introduction
- Motivation: Why should I care about data handling anyway?
  - Data I/O performance and filesystem choice.
  - Filesystems management for performance and integrity
    - Backup
    - Quotas.
    - Flushing of 'scratch' storage
  - Using multiple filesystems
    - Job or Session temporary storage
    - Medium term temporary storage
    - Long term storage.
- File transfer
- Sharing and access control
  - Replication and synchronization
- Data file formats















# DMF Reporting – Userguides: Data Store

- Extract from contents

- ▣ Getting the most from the migrating file system
  - ▣ DMF - hierarchies
  - ▣ DMF - migration
  - ▣ DMF - recall
  - ▣ DMF - dm commands
  - ▣ DMF - dmIs command - list files
  - ▣ DMF - dmget command - recall files
  - ▣ DMF - dmput command - migrate files
  - ▣ DMF - dmfind command - find files
  - ▣ DMF - dmcopY command - copy (part of) migrated file
  - ▣ DMF - dmattr command - list file attributes
  - ▣ DMF - partial-state files
  - ▣ DMF - status



# DMF Reporting – Traffic lights

|              |   |  |   |
|--------------|---|--|---|
| cherax       |    | <br>incident<br><br>down     | 512 cores available on login node.  |
| cherax-batch |    | <br>incident<br><br>down     | Jobs: 28 running (357 cores), 4 queued (64 cores), 34 total (518 cores)   |
| cherax-tapes |    | <br>incident<br><br>down    | Drives: 16 configured, 16 up, 0 down, 2 in use.   |
| DMF          |  | <br>incident<br><br>down | At 16:03, Recalls: 10 totalling 38642.0 Mbyte.<br>Migrates: 3 totalling 11693.4 Mbyte.<br>At 16:00:00, longest recall wait . 5054 Tbyte total stored. |





# DMF Reporting – Traffic lights - more

DMF Status: Thursday 2012-11-01 16:16

| VolGrp | Current |            | Reads |           | Today    |       | Hit Rate |  |
|--------|---------|------------|-------|-----------|----------|-------|----------|--|
|        | Queued  | MiB-Queued | Total | MiB-Total | %Recalls | %Data |          |  |
| cache  | 0       | 0.0        | 683   | 585914.7  | 12       | 7     |          |  |
| maid   | 0       | 0.0        | 1452  | 1152210.8 | 26       | 13    |          |  |
| se2    | 0       | 0.0        | 706   | 3350957.4 | 13       | 39    |          |  |
| se3    | 0       | 0.0        | 22    | 624885.8  | 0        | 7     |          |  |
| te2    | 13      | 50335.5    | 2681  | 2926772.5 | 48       | 34    |          |  |
| te3    | 0       | 0.0        | 0     | 0.0       | 0        | 0     |          |  |
| Total  | 13      | 50335.5    | 5544  | 8640741.2 |          |       |          |  |



# DMF Reporting – Traffic lights – even more

Requests at 2012/11/01 16:30:00

| Req-type | Req-num | Time     | Elapsed | Owner  | MSP/VG:vsn | Filename                 |
|----------|---------|----------|---------|--------|------------|--------------------------|
| move     | 585124  | 16:17:54 | 0:12:21 | lio009 | se3,te3    | /datastore/cmisis/lio009 |
| move     | 585161  | 16:24:14 | 0:06:01 | ngu038 | te3        | /datastore/cmar/ngu038   |
| move     | 585162  | 16:24:14 | 0:06:01 | ngu038 | te3        | /datastore/cmar/ngu038   |
| move     | 585163  | 16:24:14 | 0:06:01 | ngu038 | te3        | /datastore/cmar/ngu038   |
| move     | 585164  | 16:24:14 | 0:06:01 | ngu038 | te3        | /datastore/cmar/ngu038   |
| move     | 585165  | 16:24:14 | 0:06:01 | ngu038 | te3        | /datastore/cmar/ngu038   |
| move     | 585166  | 16:24:14 | 0:06:01 | ngu038 | te3        | /datastore/cmar/ngu038   |
| move     | 585167  | 16:24:14 | 0:06:01 | ngu038 | te2:G62249 | /datastore/cmar/ngu038   |



# DMF Reporting – quota

quota

Disk quotas for user abc123 (uid 12345): (blocks in kiB unless "-s" used)

| Filesystem  | blocks  | quota      | limit      | grace | files | quota  | limit  | grace |
|-------------|---------|------------|------------|-------|-------|--------|--------|-------|
| /data/flush | 3605548 | 4194304000 | 4613734400 | -     | 5182  | 200000 | 220000 | -     |
| /datastore  | 957112  | ++++       | ++++       | -     | 16952 | 150000 | 165000 | -     |
| /var        | 5016    | 20480000   | 22528000   | -     | 97    | 10000  | 11000  | -     |

++++ Space quotas only apply to online files and are normally irrelevant for /datastore



# DMF Reporting – dmget --list

- CSIRO enhanced dmget

```
>> dmget --list monitor111*  
monitor1115.out  
monitor1116.out  
monitor1117.out  
You would be recalling 5 of the 8 files  
specified.  
5 tape mounts likely to be required.  
monitor1113.out  
monitor1114.out  
monitor1110.out  
monitor1112.out  
monitor1111.out
```



# DMF Reporting – dmget use

- CSIRO enhanced dmget

```
>> dmget monitor111* &
```

```
[1] 337368
```

```
cherax monitor.archived/2012> You are recalling 5  
of the 8 files specified.
```

```
The oldest currently queued recall request has  
been waiting for 0h 7m 11s.
```

```
5 tape mounts likely to be required.
```



# DMF Reporting – WWW pages: holdings (monthly updating)

## Data Storage User Report

Users in group "csssg" sorted by total usage at 2012-10-01

| User   | online only GB | online only files | dual-state GB | dual-state files | partial-state GB | partial-state files | offline only GB | offline only files | Total GB | Total files | MB average |
|--------|----------------|-------------------|---------------|------------------|------------------|---------------------|-----------------|--------------------|----------|-------------|------------|
| ho033  | 0.080          | 23225             | 0.256         | 30               | 221.719          | 401                 | 1928.416        | 8414               | 2150.470 | 32070       | 68.665     |
| wil240 | 0.108          | 27823             | 3.014         | 277              | 1042.636         | 1005                | 16.508          | 5684               | 1062.266 | 34789       | 31.267     |
| gio011 | 0.063          | 9957              | 0.003         | 65               | 0.285            | 102                 | 988.139         | 3486               | 988.490  | 13610       | 74.373     |



# DMF Reporting – WWW pages: backup status

## User Filesystem Backup Status

Last updated Fri 2012-11-02 12:32:01 +1100

Error codes 23 and 24 are soft errors (files disappeared or open) - these can be ignored.

| Host       | Filesystem     | Last Backup          | Status | Files Transferred | Total Files | Bytes Transferred | Total Bytes | Transfer Rate (bytes/s) | Elapsed Time (s) |
|------------|----------------|----------------------|--------|-------------------|-------------|-------------------|-------------|-------------------------|------------------|
| CHERAX     | /datastore_0   | Sun 2012-10-28 06:47 | 0      | 23M               | 24M         | 55G               | 24T         | 2.3M                    | 24214            |
| tardis     | /cs/home       | Thu 2012-11-01 19:15 | 0      | 12K               | 2M          | 1.4G              | 609G        | 309K                    | 4429             |
| bragg-user | /gpugwy02/home | Thu 2012-11-01 19:16 | 0      | 4K                | 3.6M        | 2.5G              | 467G        | 561K                    | 4471             |
| burnet-srv | /home          | Thu 2012-11-01 20:54 | 0      | 16K               | 4.4M        | 37G               | 1.5T        | 4.5M                    | 8188             |
| CHERAX     | /datastore_9   | Fri 2012-11-02 00:31 | 0      | 299K              | -1          | 1.3G              | 27T         | 1.4M                    | 873              |
| Total: 5   |                |                      |        | 24M               | 34M         | 97G               | 54T         | 9.1M                    | 42175            |

(In the above table, backups shown with a blue/aqua background are older than 24 hours).



# DMF Reporting

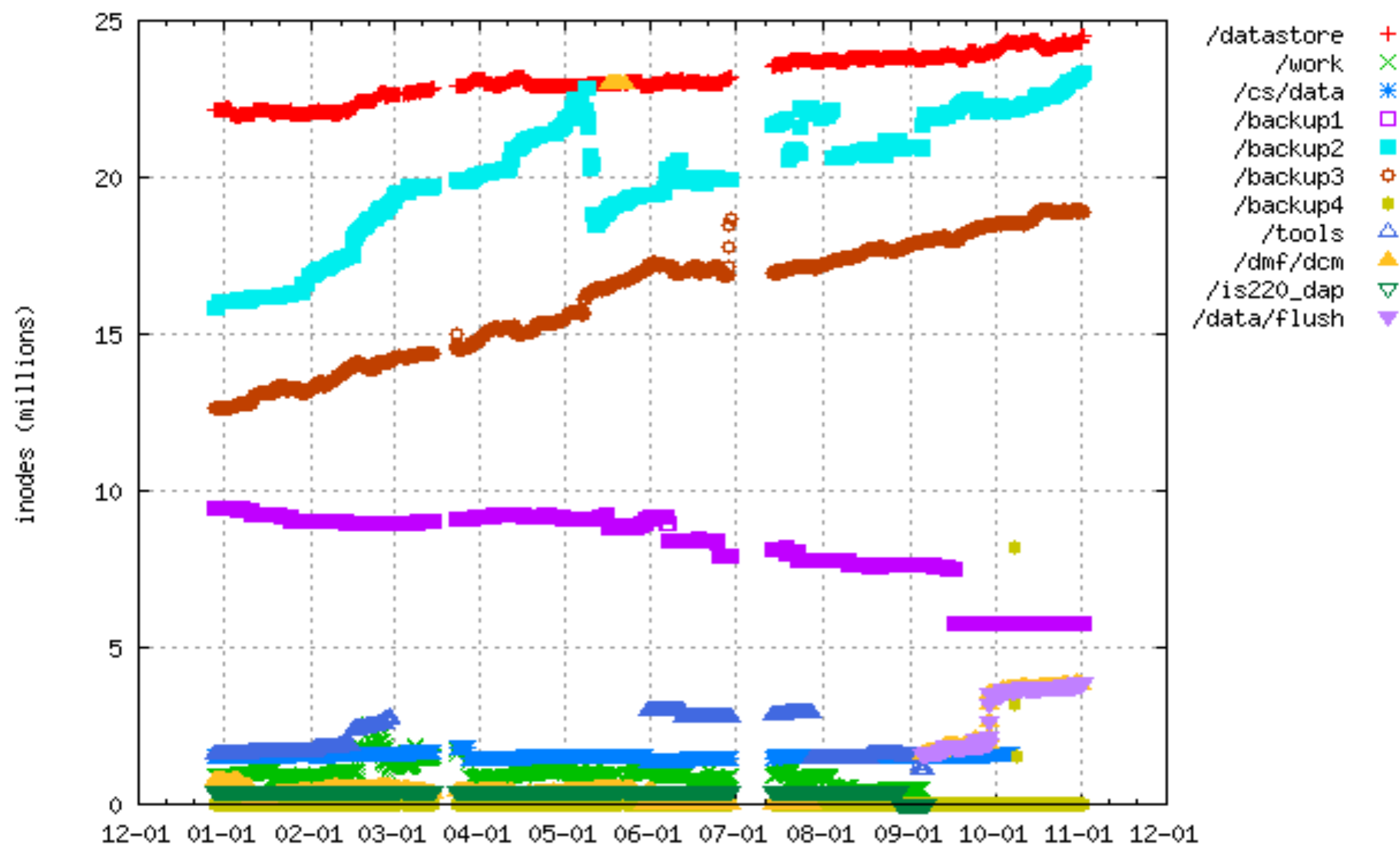
- DMF system
  - Logs
  - e-mails
  - Collected statistics

*See some of Peter's talk*
- Trend reporting
  - File systems
  - Holdings
  - Cache usage
  - Recall statistics
  - Dumps and backups

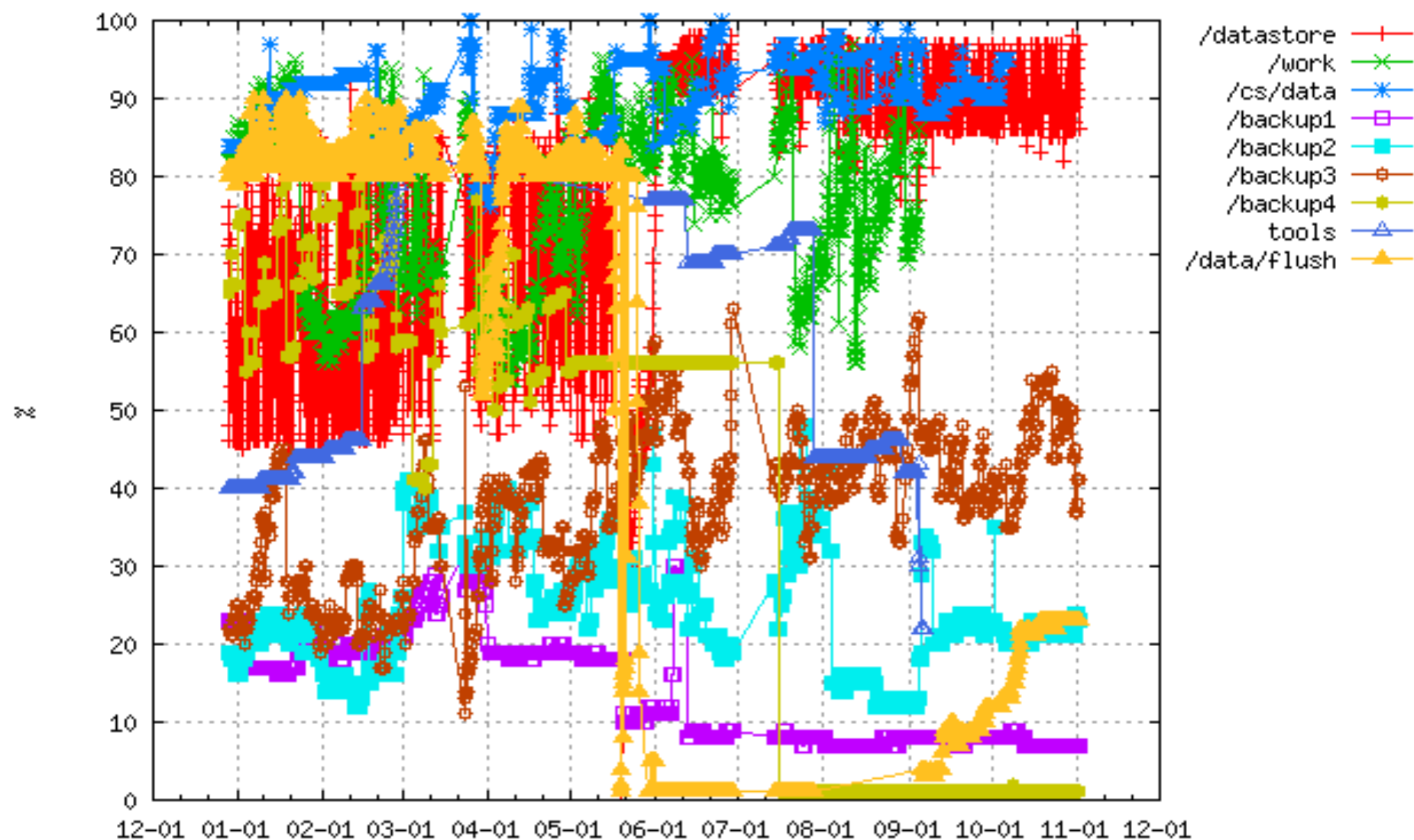




CSIRO ASC cherax - File system inodes to 2012-11-02

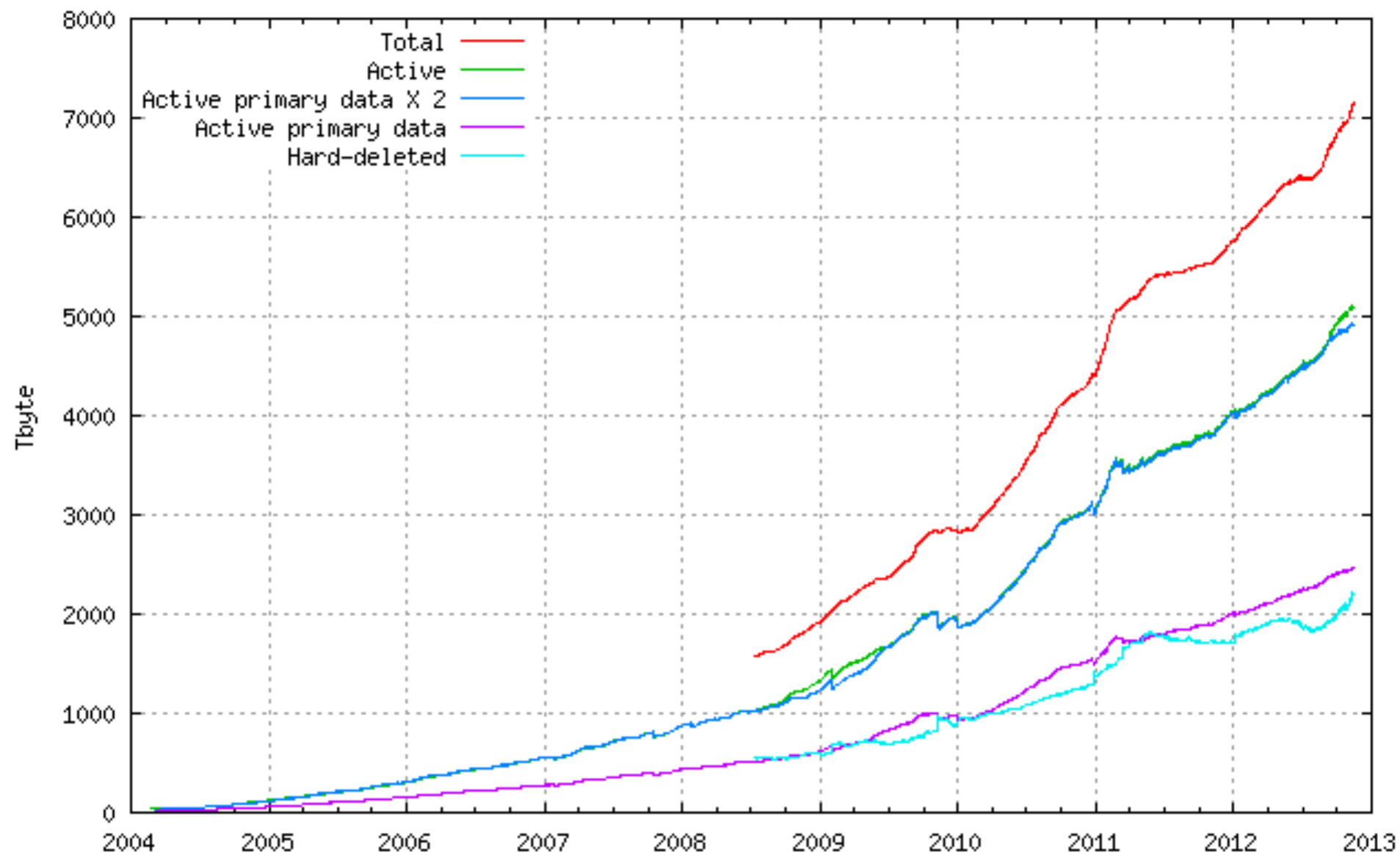


CSIRO ASC cherax - File system space to 2012-11-02





CSIRO ASC Data Store - Total Holdings to 2012-11-14



# DMF Reporting: dumps and backups

| Host    | Filesystem                                | Last Backup ▲        | Status | Files Transferred | Total Files | Bytes Transferred | Total Bytes | Transfer Rate (bytes/s) | Elapsed Time (s) |
|---------|---|----------------------|--------|-------------------|-------------|-------------------|-------------|-------------------------|------------------|
| CHERAX  | /data/flush/int_backup/backup_stage/daily | Fri 2012-11-02 03:21 | 23     | 3                 | 5           | 1.7G              | 1.7G        | 11M                     | 140              |
| CHERAX  | DUMP_TAPE_CHECK                           | Fri 2012-11-02 03:18 | 0      | -1                | -1          | -1                | -1          | -1                      | -1               |
| CHERAX  | dump-tape-write                           | Fri 2012-11-02 03:06 | 0      | -1                | -1          | -1                | -1          | -1                      | -1               |
| CHERAX  | /backup2_0                                | Fri 2012-11-02 02:44 | 0      | 14M               | 23M         | 11G               | 168G        | 1.2M                    | 8883             |
| CHERAX  | /datastore_9                              | Fri 2012-11-02 00:31 | 0      | 299K              | -1          | 1.3G              | 27T         | 1.4M                    | 873              |
| CHERAX  | /data/flush_0                             | Fri 2012-11-02 00:28 | 0      | 5.4K              | -1          | 82G               | 22T         | 113M                    | 730              |
| CHERAX  | dmsnap                                    | Fri 2012-11-02 00:16 | 0      | 17                | 17          | 31G               | 31G         | 38M                     | 836              |
| CHERAX  | bootstrap                                 | Fri 2012-11-02 00:16 | 0      | 3.6K              | 3.6K        | 185M              | 185M        | 15M                     | 12               |
| cherax  | housekeeping                              | Thu 2012-11-01 21:37 | 0      | -1                | -1          | -1                | -1          | -1                      | 1404             |
| stream1 | /var                                      | Thu 2012-11-01 21:12 | 0      | 117               | 4.6K        | 188M              | 587M        | 27M                     | 7                |
| stream1 | /usr                                      | Thu 2012-11-01 21:12 | 0      | 0                 | 121K        | 0                 | 861M        | 0                       | 69               |



# DMF Reporting – DMF reports

## Volume Occupancy:

|         | MB Active      | MB Hard-deleted | % Active | Approx MB |
|---------|----------------|-----------------|----------|-----------|
| bc3     | 10060524.701   | 453191.149      | 96       | ??        |
| bk2     | 10061833.820   | 19862357.132    | 34       | 1391178   |
| se2     | 2216543458.993 | 1146628810.791  | 66       | 1992974   |
| se3     | 230277965.266  | 9338141.859     | 96       | 13324084  |
| te2     | 957460094.730  | 641116551.527   | 60       | 2925018   |
| te3     | 1496301414.631 | 277212648.303   | 84       | 5395094   |
| testvg1 | 5.833          | 0.000           | 100      | ??        |
| Totals  | 4920705297.974 | 2094611700.761  | 70       | 25028348  |



# DMF Reporting – DMF reports

Data movement for the day:

| VG      | MB migrated  | MB recalled |
|---------|--------------|-------------|
| bc3     | 22432.702    | 0.000       |
| bk2     | 22432.702    | 0.000       |
| se2     | 4124555.228  | 3833371.817 |
| se3     | 4319517.558  | 1479020.870 |
| te2     | 0.000        | 4414108.359 |
| te3     | 9765819.636  | 0.000       |
| testvg1 | 5.833        | 0.000       |
| Totals  | 18254763.659 | 9726501.046 |



# DMF Reporting – for Staff

- Nagios





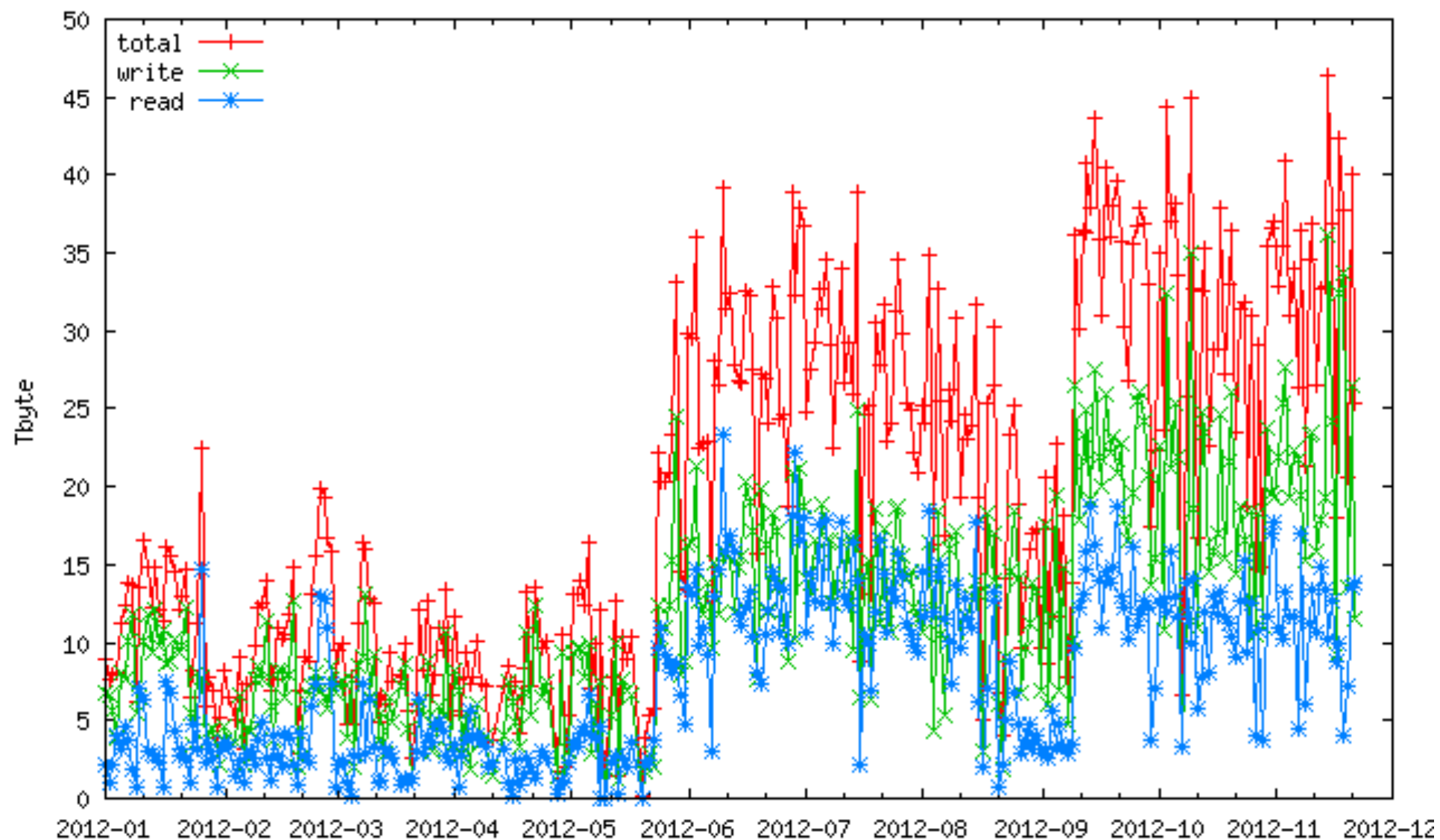
| Host                | Service                              | Status              | Last Check          | Duration       | Attempt                             | Status Information                                     |
|---------------------|--------------------------------------|---------------------|---------------------|----------------|-------------------------------------|--|
| cherax              | <a href="#">Backups - external</a>   | OK                  | 2012-11-02 14:58:11 | 0d 17h 22m 23s | 1/4                                 | Nonproduction external backup is running               |
|                     | <a href="#">Backups - iVec</a>       | OK                  | 2012-11-02 14:57:12 | 0d 11h 38m 22s | 1/4                                 | iVec backup - not running                              |
|                     | <a href="#">Backups - internal</a>   | OK                  | 2012-11-02 14:57:46 | 5d 3h 22m 48s  | 1/4                                 | Internal backup - not running                          |
|                     | <a href="#">Cherax-Cortex</a>        | OK                  | 2012-11-02 14:57:16 | 5d 3h 23m 20s  | 1/4                                 | OK   |
|                     | <a href="#">Cherax-Vayu</a>          | OK                  | 2012-11-02 14:58:03 | 5d 3h 22m 33s  | 1/4                                 | OK   |
|                     | <a href="#">DMF</a>                  | OK                  | 2012-11-02 14:55:36 | 0d 16h 9m 58s  | 1/4                                 | OK   |
|                     | <a href="#">DMFMAN</a>               | OK                  | 2012-11-02 14:59:20 | 0d 4h 51m 14s  | 1/4                                 | OK   |
|                     | <a href="#">FC disk - IS220</a>      | OK                  | 2012-11-02 14:58:00 | 5d 3h 22m 34s  | 1/4                                 | PING OK - Packet loss = 0%, RTA = 0.72 ms              |
|                     | <a href="#">FC disk - IS4600</a>     | OK                  | 2012-11-02 14:55:29 | 5d 3h 20m 5s   | 1/4                                 | PING OK - Packet loss = 0%, RTA = 0.84 ms              |
|                     | <a href="#">FC disk - IS5500</a>     | OK                  | 2012-11-02 14:56:33 | 5d 3h 24m 1s   | 1/4                                 | PING OK - Packet loss = 0%, RTA = 0.67 ms              |
|                     | <a href="#">FC disk agent - SMEE</a> | OK                  | 2012-11-02 15:00:18 | 5d 3h 20m 17s  | 1/4                                 | SMEE Agent problem - PMUNIXLauncher not running        |
|                     | <a href="#">Free memory</a>          | OK                  | 2012-11-02 14:57:59 | 5d 3h 22m 35s  | 1/4                                 | 3856 GB remains available                              |
|                     | <a href="#">HTTP</a>                 | OK                  | 2012-11-02 14:58:17 | 5d 3h 7m 17s   | 1/4                                 | HTTP OK HTTP/1.1 200 OK - 13661 bytes in 0.020 seconds |
|                     | <a href="#">Load</a>                 | OK                  | 2012-11-02 14:56:36 | 5d 3h 19m 0s   | 1/4                                 | OK - load average: 314.24, 276.25, 260.98              |
|                     | <a href="#">Local filesystems</a>    | OK                  | 2012-11-02 14:57:57 | 1d 19h 57m 37s | 1/4                                 | All 14 filesystems OK                                  |
|                     | <a href="#">Misc Procs</a>           | OK                  | 2012-11-02 14:56:37 | 5d 3h 19m 0s   | 1/4                                 | Miscellaneous processes all running                    |
|                     | <a href="#">NFS Daemons</a>          | OK                  | 2012-11-02 14:57:34 | 5d 3h 8m 10s   | 1/4                                 | NFS daemons are now OK                                 |
|                     | <a href="#">OpenVault</a>            | OK                  | 2012-11-02 14:55:17 | 3d 3h 5m 17s   | 1/4                                 | OK   |
|                     | <a href="#">PCP</a>                  | OK                  | 2012-11-02 14:58:29 | 0d 4h 52m 5s   | 1/4                                 | PCP is running   |
|                     | <a href="#">PING</a>                 | OK                  | 2012-11-02 14:58:06 | 5d 3h 22m 30s  | 1/4                                 | PING OK - Packet loss = 0%, RTA = 0.77 ms              |
|                     | <a href="#">TMF/ACSLs</a>            | OK                  | 2012-11-02 14:59:17 | 4d 20h 11m 17s | 1/4                                 | OK   |
|                     | <a href="#">Tape Numbers</a>         | OK                  | 2012-11-02 14:20:14 | 0d 14h 40m 20s | 1/4                                 | OK   |
|                     | <a href="#">Tape drives</a>          | OK                  | 2012-11-02 14:56:13 | 5d 3h 4m 28s   | 1/4                                 | TMF drives seem OK                                     |
|                     | <a href="#">Tape hammer</a>          | OK                  | 2012-11-02 14:55:11 | 0d 12h 55m 23s | 1/4                                 | Tape usage seems OK                                    |
|                     | <a href="#">Torque Batch System</a>  | OK                  | 2012-11-02 14:58:34 | 5d 3h 7m 7s    | 1/4                                 | OK   |
|                     | <a href="#">hpn-ssh</a>              | OK                  | 2012-11-02 15:00:17 | 5d 3h 20m 17s  | 1/4                                 | SSH OK - OpenSSH_5.8p1-hpn13v11 (protocol 2.0)         |
| <a href="#">ssh</a> | OK                                   | 2012-11-02 14:59:12 | 5d 3h 21m 33s       | 1/4            | SSH OK - OpenSSH_5.1 (protocol 2.0) |  |

# DMF Reporting – for staff

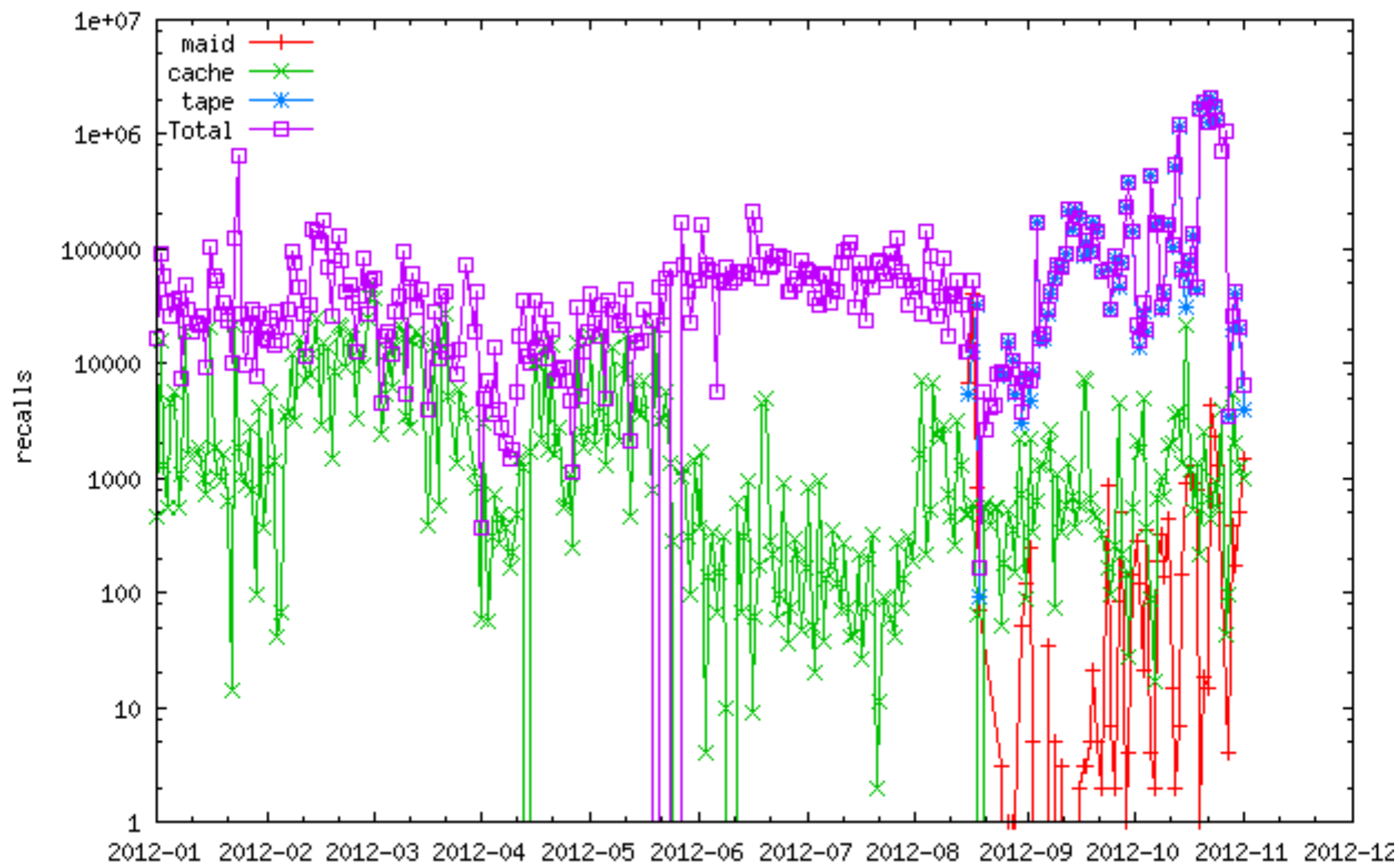
- Trend reporting (some of this for Management too)
  - Traffic
  - DCM usage
  - Recall statistics
  - Performance



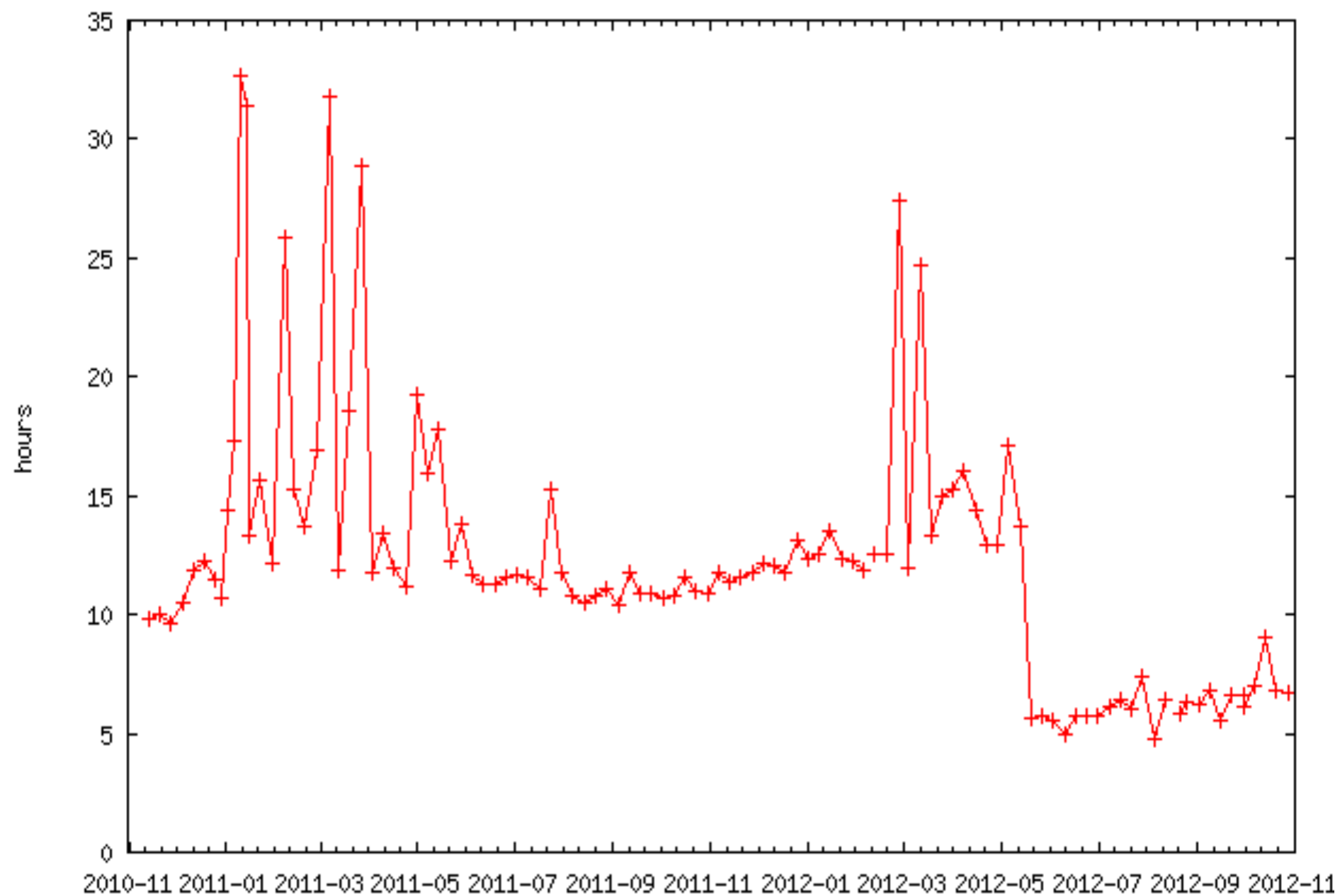
CSIRO ASC DMF traffic to 2012-11-22



CSIRO ASC DMF daily file recalls



CSIRO ASC - Backup elapsed time, -datastore\_0



# DMF Reporting – for staff

- Can plot for dumps and backups:
  - Finishing times
  - Elapsed times
  - Counts of files
  - Data available and transferred
  - Transfer rates
  - Churn rate
  - Progress (dumps)

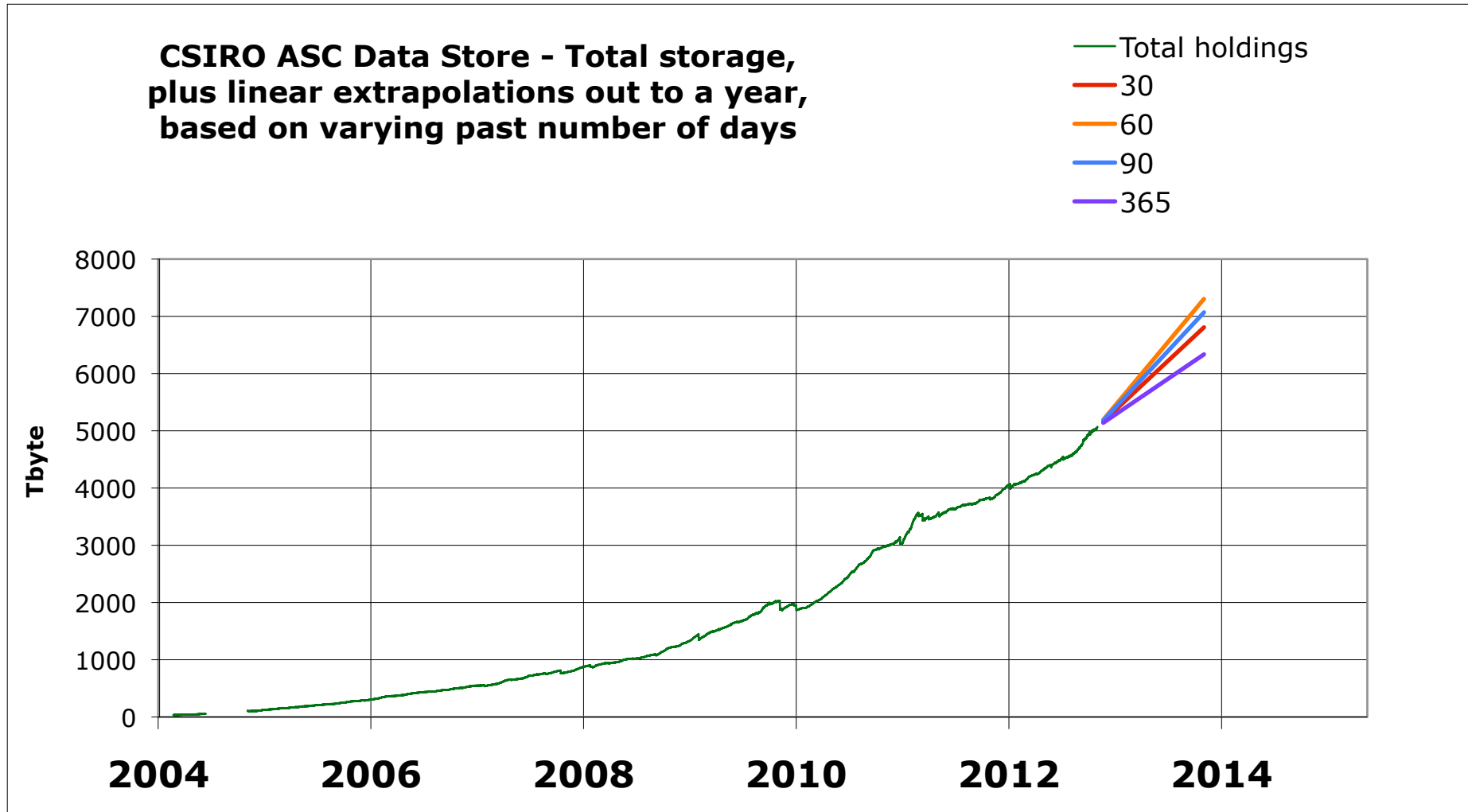


# DMF Reporting – for Management

- DMF trends
- Highlights from other reports



# DMF Reporting – for Management

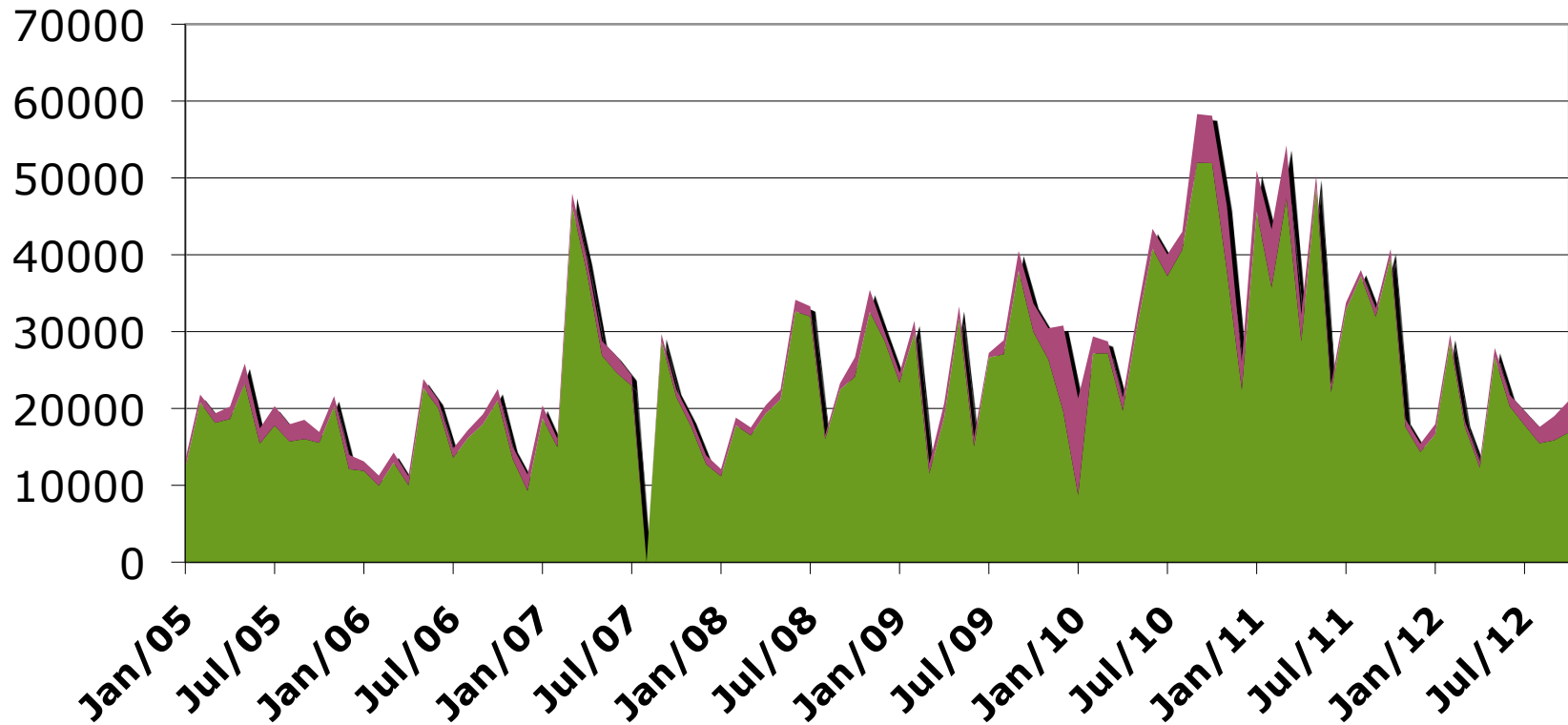




# CSIRO ASC Data Store - monthly tape mounts

(data incomplete, Jul-Oct 2008, Jun 2009)

- Tape mounts - W
- Tape mounts - R

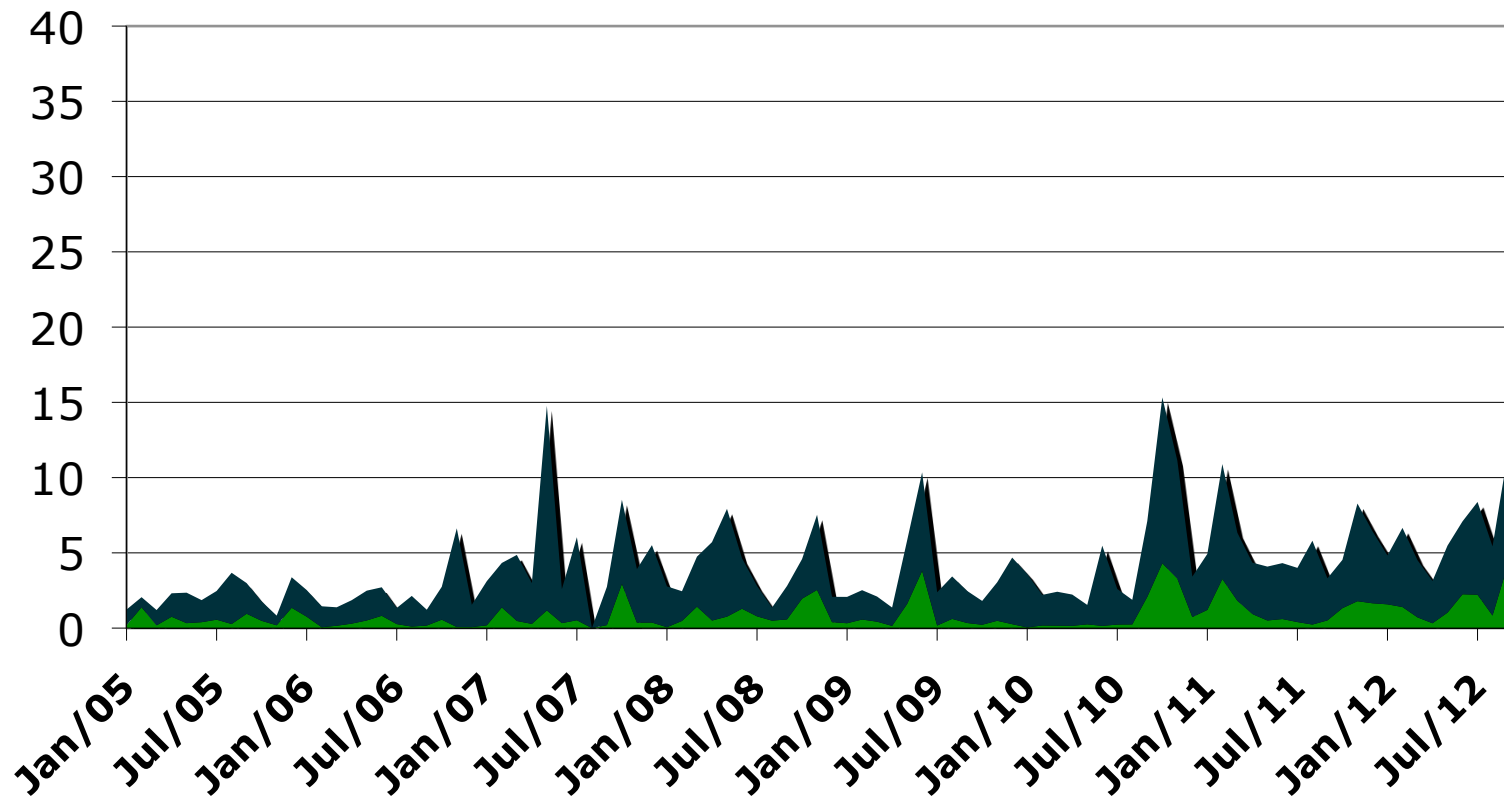


# CSIRO ASC Data Store - monthly files moved

(data incomplete, Jul-Oct 2008, Jun 2009)

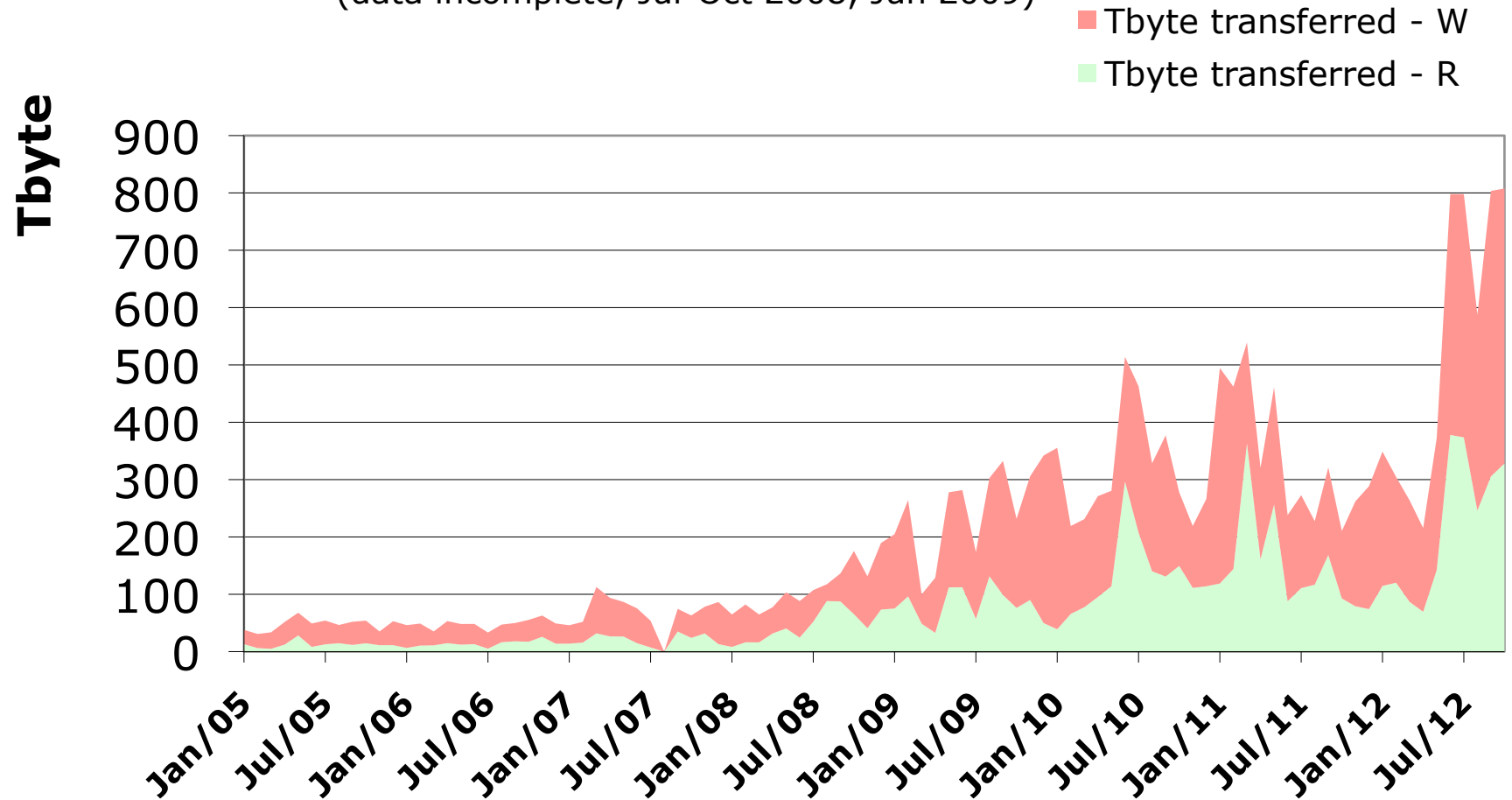
Millions

- Files moved - W
- Files moved - R



# CSIRO ASC Data Store - monthly data moved

(data incomplete, Jul-Oct 2008, Jun 2009)

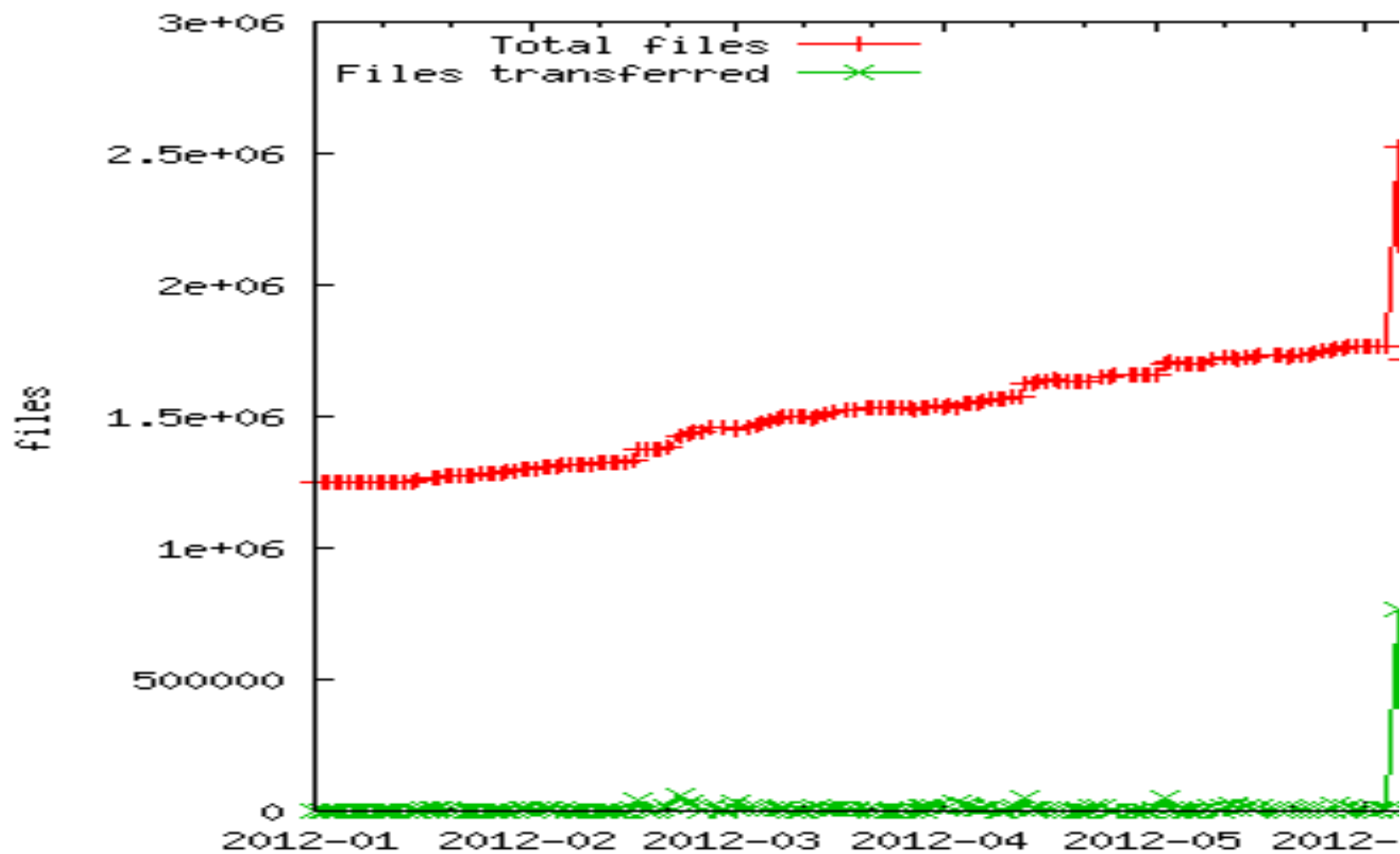


# DMF Reporting – diagnostics

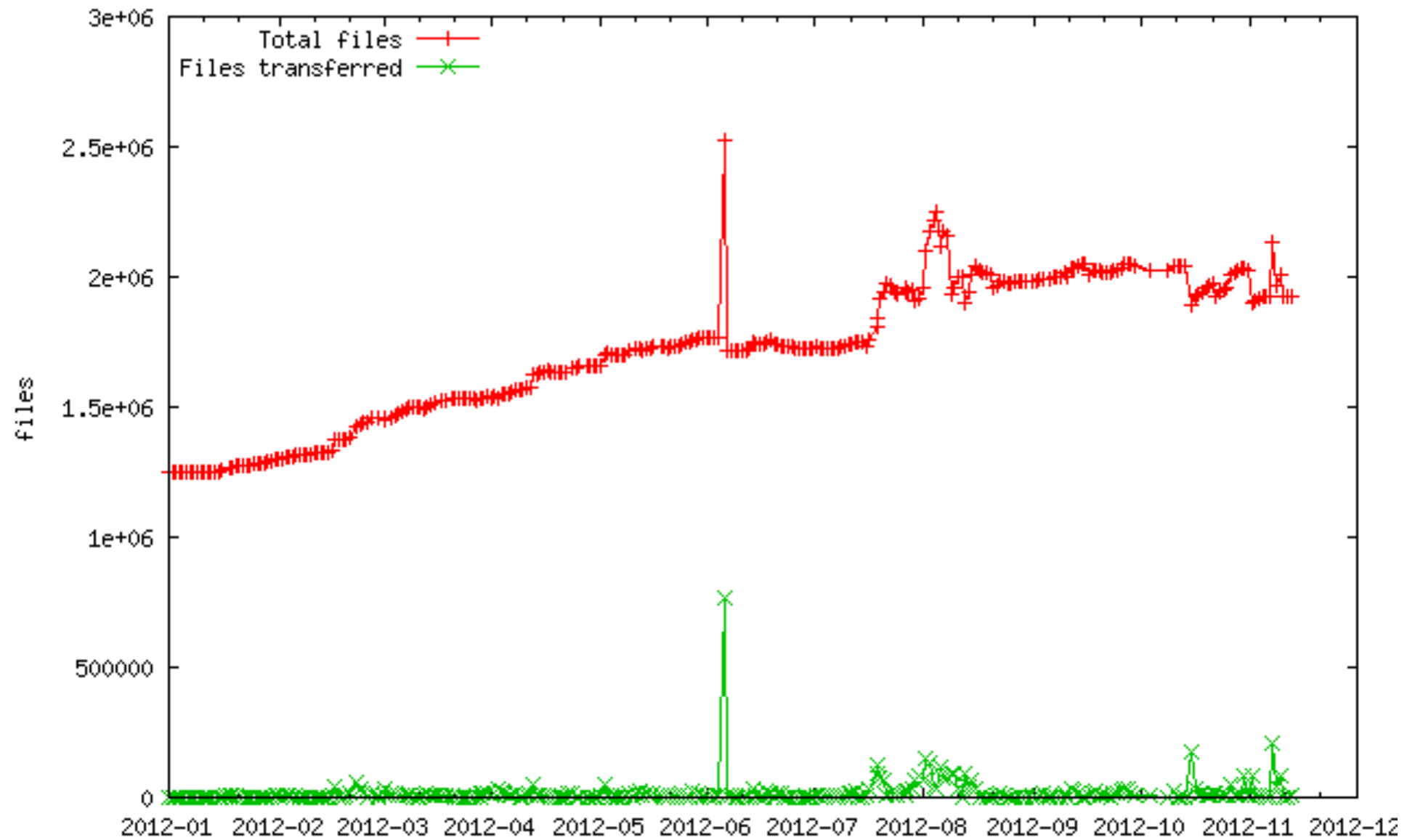
- Look for exceptions, and changes in trends
  - Diagnosis of system problems
  - Diagnosis of aberrant user behaviour
- Example 1: spike in numbers of inodes
  - Backups slower than normal
  - (from WWW page or graph)
  - Looked at graph of inodes for all FSes
  - Found jump in inodes into /backup2
  - Analysis from WWW table of backups – looked for large numbers of files transferred



CSIRO ASC - Backup file c



CSIRO ASC - Backup file counts , tardis.-cs-home.2012-

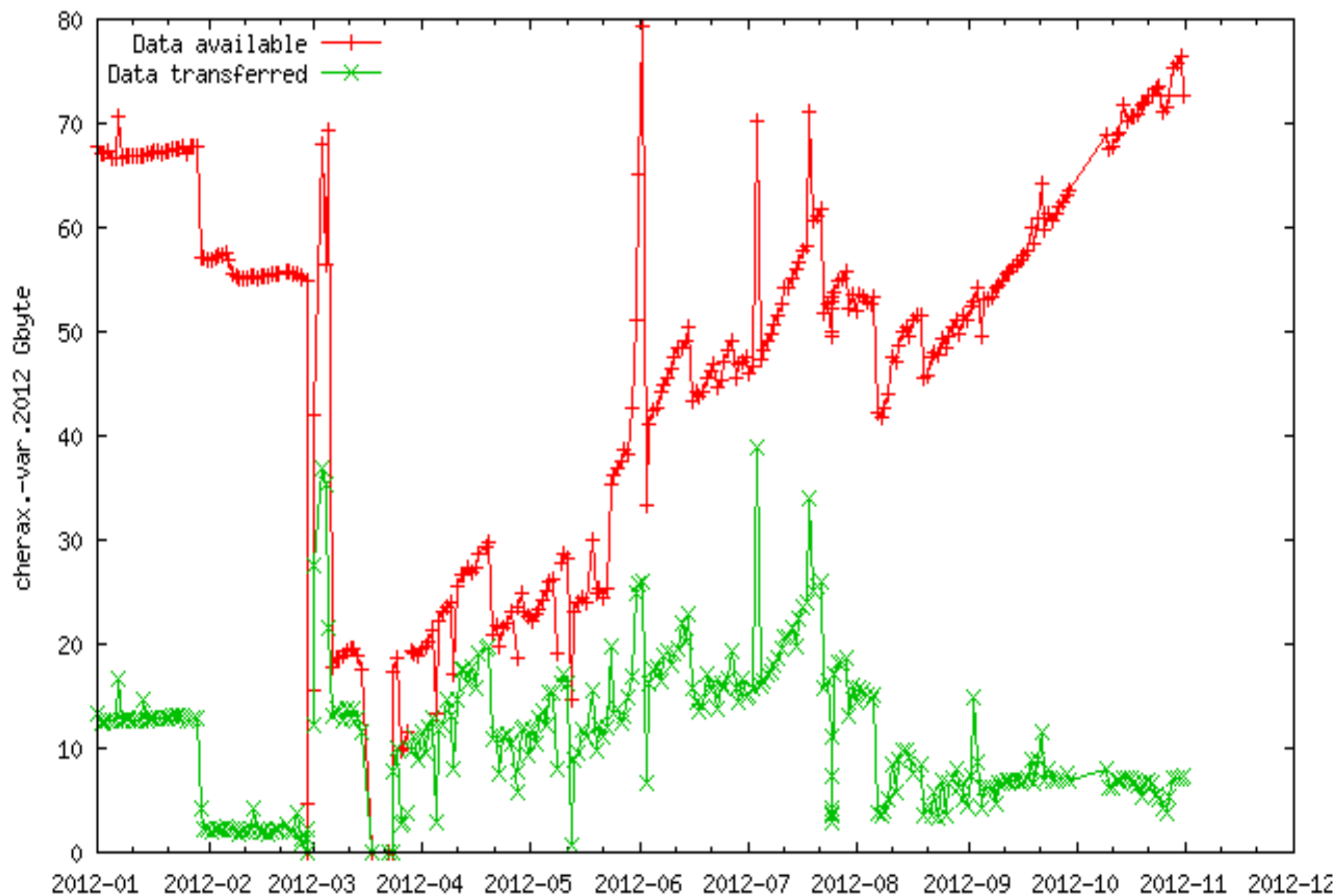


# DMF Reporting – diagnostics

- Example 1: user was creating a million temporary files in a backed-up area
    - Fixed by counselling
- 
- Example 2: churn in /var
    - Large proportion of data being transferred each day

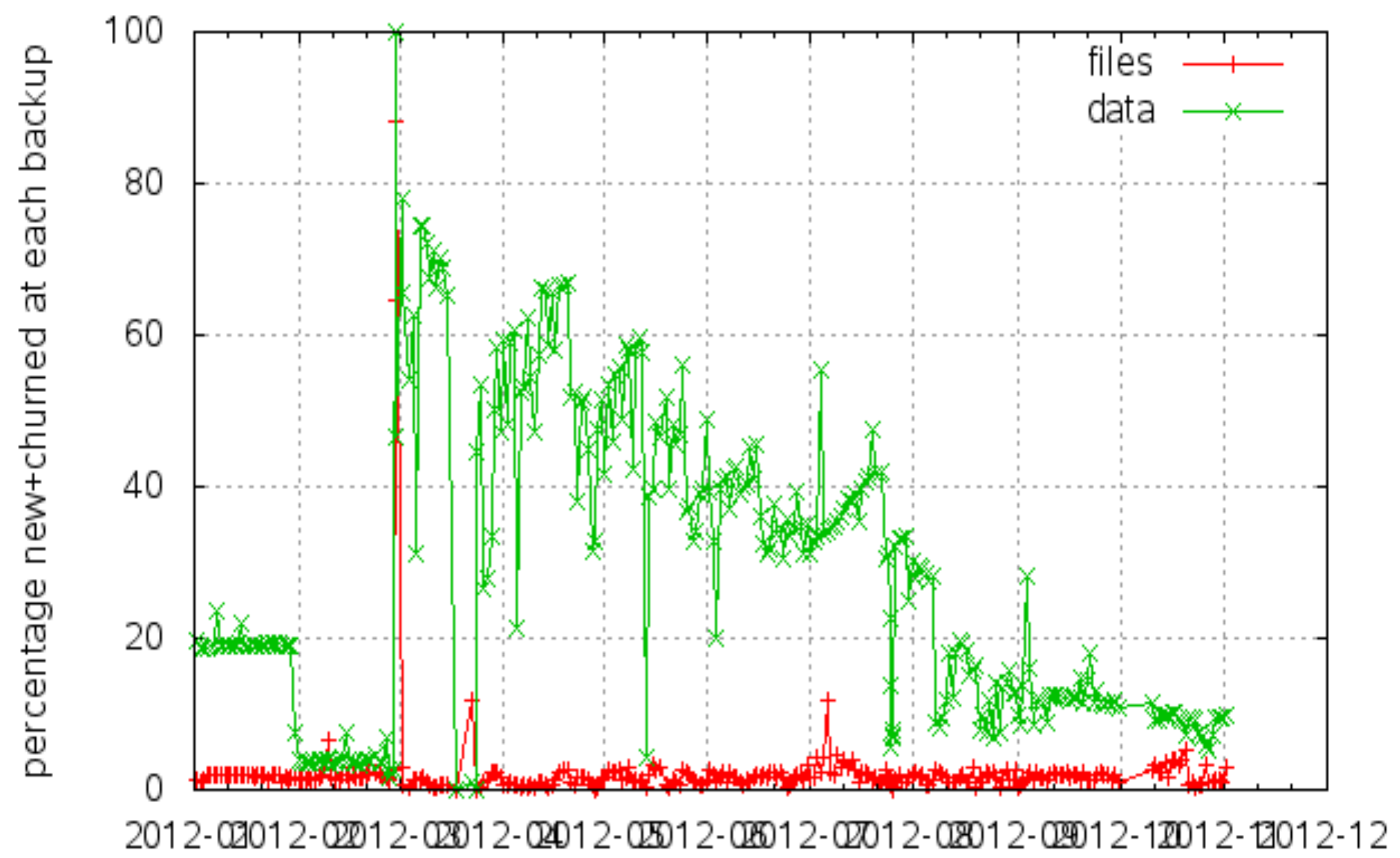


CSIRO ASC - Backup source data and amount transferred. cherax.-var.2012





CSIRO ASC - backup new+churn rate for 2012 cherax-var  
Average: files 2.191 %, data 27.747 %



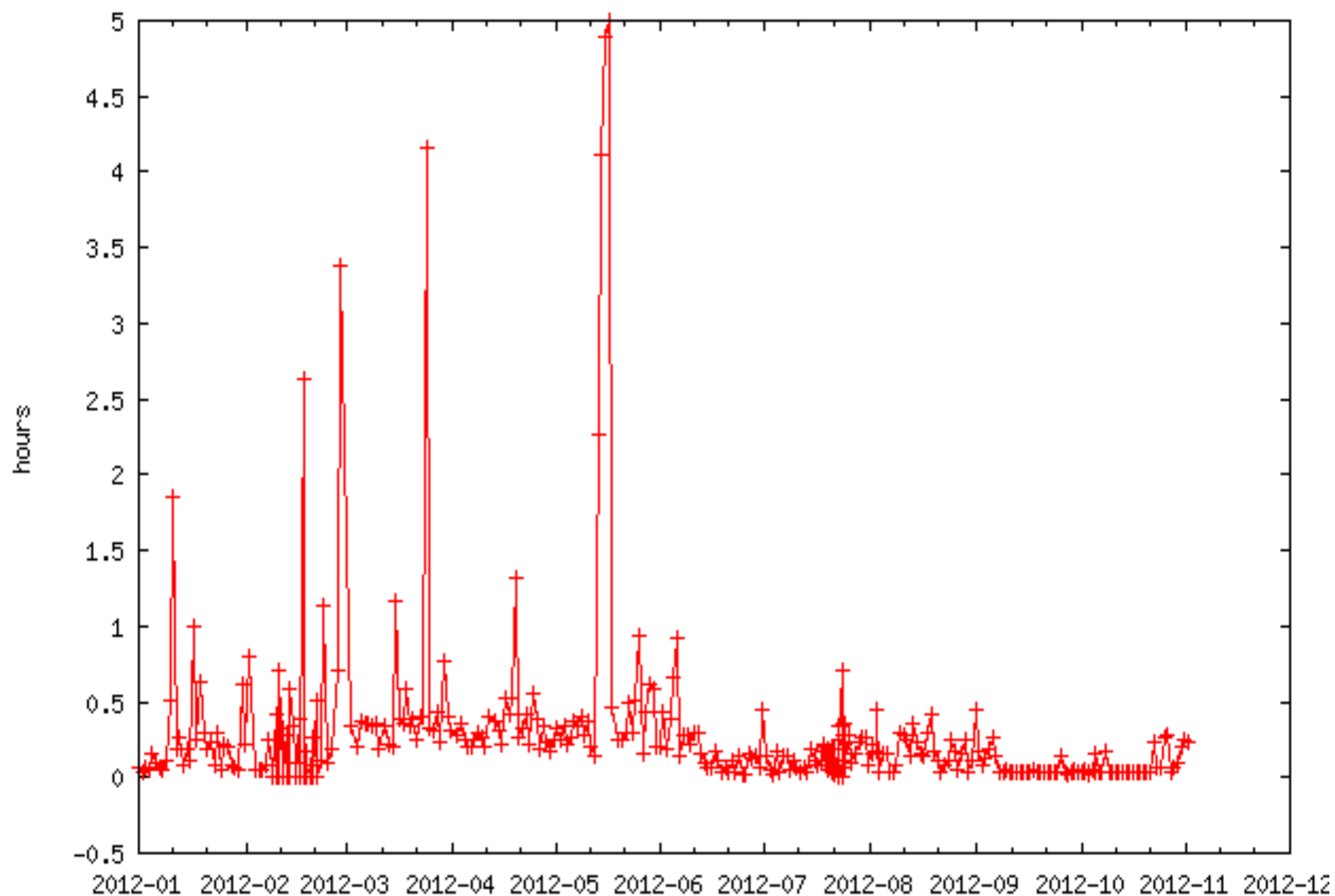
# DMF Reporting – diagnostics

- Example 2: churn in /var
  - Fixed with rolling over of large logs

- 
- Example 3: time to do DMF snaps
    - Highly variable



CSIRO ASC - Backup elapsed time, snap.2012-



# DMF Reporting – diagnostics

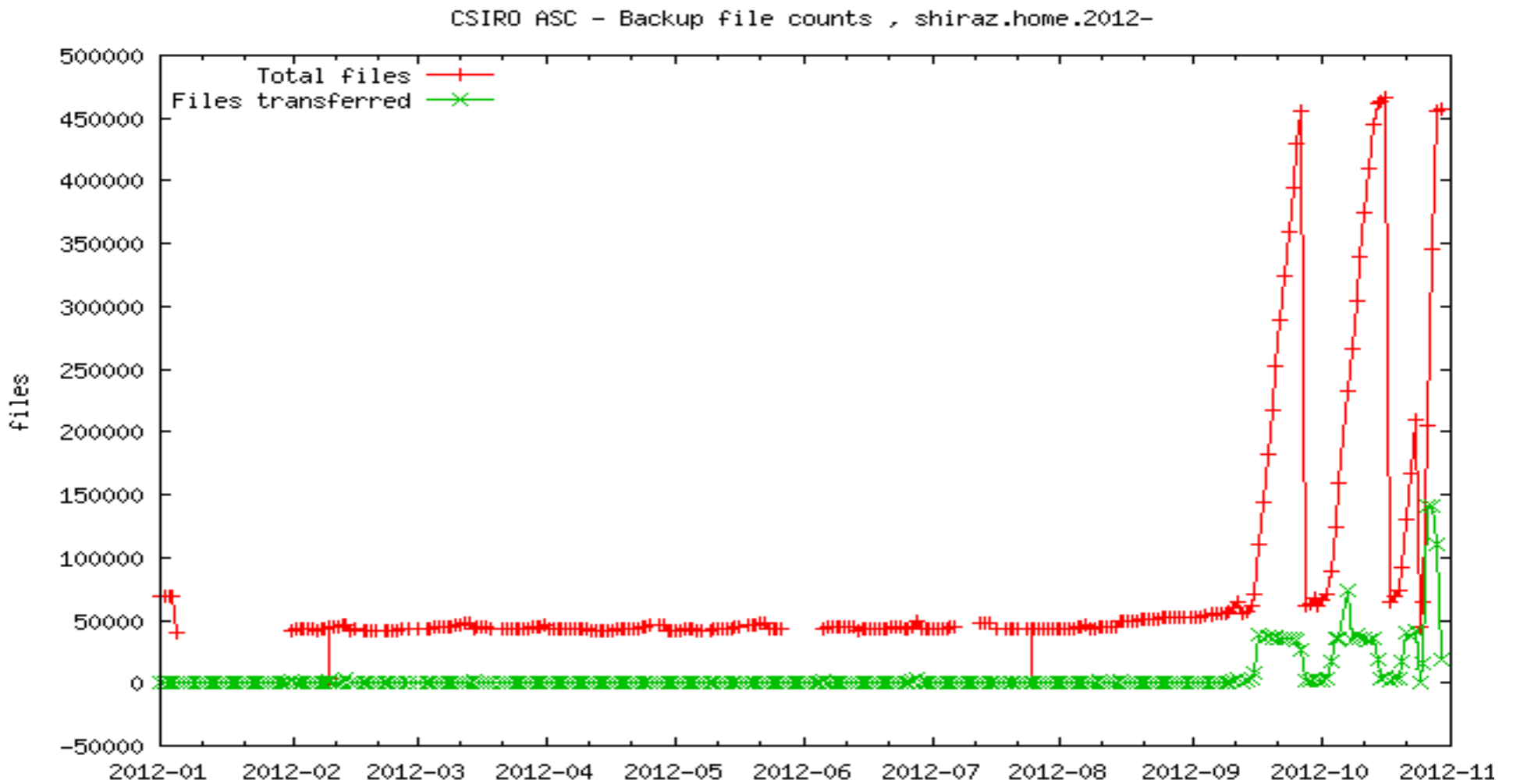
- Example 3: time to do snaps
  - Fixed by new host, speeding up other processes, moving the target to new disc area (plus databases on better Fses)

---

- Example 4: jump in numbers of files in backup
  - Seen first in increase in inode count
  - Used WWW page of logs to find increased numbers of files transferred,
    - 40,000 new files every day over 10 days
  - Identified as coming from backup of user workstation



# DMF Reporting – diagnostic example



# DMF Reporting – diagnostics

- Example 4: spike in numbers of files being backed-up
  - Solution: identified source:
    - firefox cache on workstation
  - Fix is to use `rsync --exclude` to stop backing up firefox cache



# DMF Reporting – Conclusion

“First get your facts right  
– then distort as you wish” – CSIRO colleague

- Important to collect data
- Important to monitor for:
  - Exceptions
  - Trends
  - Reporting to management
  - Capacity planning
  - ‘Helping’ users



# Thank you

**CSIRO IM&T Advanced Scientific Computing**

Robert C. Bell

Technical Services Manager

**t** +61 3 9669 8102

**e** [Robert.Bell@csiro.au](mailto:Robert.Bell@csiro.au)

**w** [www.hpsc.csiro.au/](http://www.hpsc.csiro.au/)

**CSIRO IM&T ADVANCED SCIENTIFIC COMPUTING**

[www.csiro.au](http://www.csiro.au)





[www.csiro.au](http://www.csiro.au)



# Section 1:

## Level 2

Level 3

Section divider slide with 3 text levels