CERN MSS Metrics from the Tier-1 Sites

From: "Olof Barring" <Olof.Barring@cern.ch>
Date: Tue, 28 Apr 2009 13:08:43 +0200
To: "Alberto Aimar" <Alberto.Aimar@cern.ch>, "worldwide-lcg-management-board (LCG Management Board)" <worldwide-lcg-management-board@cern.ch>

```
   Alberto,

At CERN we collect the following metrics in SLS:

DATA READ: data volume read in GB
DATA READ: number of files transferred
DATA READ: average filesize in MB
DATA WRITE: data volume written in GB
DATA WRITE: number of files transferred
DATA WRITE: average filesize in MB
RATE READ: read transfer rate inc drive overhead MB/sec
RATE READ: drive read transfer rate MB/sec
RATE WRITE: write transfer rate inc drive overhead MB/sec
RATE WRITE: drive write transfer rate MB/sec
TAPE MOUNT: successful mounts in last 4hrs
TAPE MOUNT: failed mounts in last 4hrs
TAPE MOUNT: read mounts in last 4hrs
TAPE MOUNT: write mounts in last 4hrs
TAPE MOUNT: average tape mount time in secs
TAPE QUEUES: average wait for read in secs
TAPE QUEUES: average wait for write in secs
FILES PER MOUNT: read average
FILES PER MOUNT: write average
TAPE VOLUME for <VO> in TB
TAPE REPEAT MOUNT: read average last 24hrs
TAPE REPEAT MOUNT: write average last 24hrs
TAPE FRAGMENTATION: Percentage full level for <VO> tapes

The data is collected and can be visualized in SLS.

        Best regards,

                      Olof
```

```
  -----Original Message-----
  From: Alberto Aimar
  Sent: Monday, April 27, 2009 12:02 PM
  To: worldwide-lcg-management-board (LCG Management Board)
  Subject: MSS Metrics from the Tier-1 Sites

  Dear Colleagues
      in tomorrow's MB agenda there is a Sites' round table:

  In practice Sites should specify how they will provide their
  MSS information rates, etc (i.e. any useful metric) for
  STEP09 and further.

  If you send me that information (quite) before the meeting I
  will collect it in a single document for the MB meeting.

  Best regards.
  Alberto Aimar.
```

DE-KIT Metrics from the Tier-1 Sites

From: Andreas Heiss <andreas.heiss@iwr.fzk.de>
Date: Mon, 27 Apr 2009 15:47:39 +0200
To: Alberto Aimar <alberto.aimar@cern.ch>

```
Dear Alberto,
some tape monitoring is available from the general GridKa monitoring
web pages at http://grid.fzk.de/monitoring/
which are public accessible without password.

In the menuframe on the left, there is a link "Tape transfers" and
additional links for each experiment. The graphs show e.g. number of
transferred Gigabytes per hour and average file sizes.
This information comes from the TSM tape management system.

The tape rates are also monitored on the dCache level and this
information is available when you click on "dCache I/O history" in the
menuframe.

We'll try to get additional information gathered and displayed (e.g.
rates per spacetoken) until STEP09 but we cannot promise.


I am not sure if I can attend tomorrows MB phone con. since we have an
important internal meeting.

cheers
 Andreas




Alberto Aimar wrote:
  Dear Colleagues
     in tomorrow's MB agenda there is a Sites' round table:

  In practice Sites should specify how they will provide their MSS
  information
  rates, etc (i.e. any useful metric) for STEP09 and further.

  If you send me that information (quite) before the meeting I will collect
  it in a single document for the MB meeting.

  Best regards.
  Alberto Aimar.


--
  Dr. Andreas Heiss
  Karlsruhe Institute of Technology (KIT)
  Forschungszentrum Karlsruhe GmbH
  Steinbuch Centre for Computing (SCC)
  Abteilung Verteilte Systeme und Grid (VSG)
  Hermann-von-Helmholtz-Platz 1
  D-76344 Eggenstein-Leopoldshafen
  Tel: +49 (0)7247 82 5639, Fax: +49 (0)7247 82 4972
  Email: andreas.heiss@kit.edu
  ------------------------------------------------------------
  Forschungszentrum Karlsruhe GmbH, Weberstraße 5, D-76133 Karlsruhe
  Amtsgericht Mannheim HRB 100302
  Vorsitzende des Aufsichtsrats: MinDir'in Bärbel Brumme-Bothe
  Vorstand (Geschäftsführung): Prof. Dr. Eberhard Umbach (Vorsitzender);
  Dr. Alexander Kurz (stellv. Vorsitzender); Dr.-Ing. Peter Fritz;
  Prof. Dr.-Ing. Detlef Löhe; Prof. Dr. Horst Hippler
```
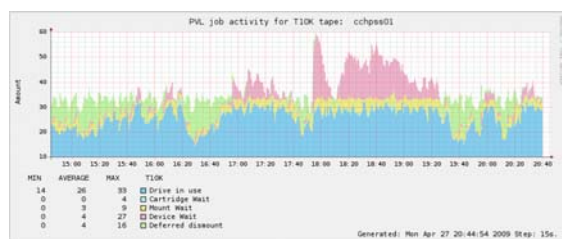
FR-CCIN2P3

Fabio Hernandez
fabio@in2p3.fr

Tape measurements we currently do

- Throughput between dCache and HPSS (as perceived by dCache), for reading & writing data, in MB/sec both agregated for the 4 LHC experiments, and per experiment
- Tape drive usage, per drive type. For LHC experiments, only the usage of STK T10.000 tape drives is relevant. We measure:
  - number of drives in use
  - number of copy requests waiting for a cartridge to be available
  - the number of copy requests waiting for a cartridge to be mounted
  - the number of copyrequests waiting for a tape drive to be free
  - Maximum mount time per drive type (in minutes)
- Network bandwidth used by HPSS tape servers and disk servers
  - as the tape and disk servers are not dedicated per experiment, we currently don't have a means to get this information per experiment
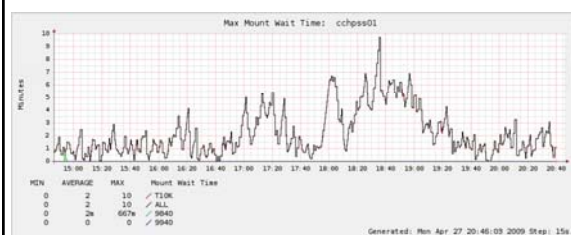
- Some examples follow

Tape drives usage
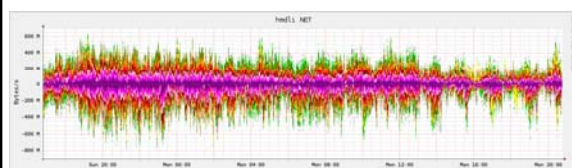### *(STK T10.000 drives used by LHC experiments)*

Tape mount wait time
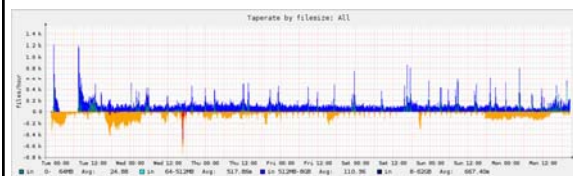### *(Shows potential contention on tape drives)*

Aggregated HPSS disk servers network traffic
(all experiments, including but not limited to LHC ones)

Files copied to/from tape as a function of size
(as seen by dCache, all LHC experiments)

Amount of data sent/received to/from HPSS as seen by dCache
(all LHC experiments)

What measurements we don't currently have

- Distribution of the sizes of files on tape, per experiment
- The number of files read or written per tape mount
- The amount of data read or written per tape mount
- The tape rate per kind of data (e.g. RAW, MC, …) per experiment
  - I think we should be able to extract this information from dCache but I'm not sure – have to check

Final remarks

- Currently this information is available in the intranet

NDGF MSS Metrics from the Tier-1 Sites

From: Oxana Smirnova <oxana.smirnova@hep.lu.se>
Date: Tue, 28 Apr 2009 12:28:36 +0200
To: Alberto Aimar <alberto.aimar@cern.ch>
CC: "worldwide-lcg-management-board (LCG Management Board)"
<worldwide-lcg-management-board@cern.ch>

```
Hi Alberto,

I'm not sure I understand the question "how", and the rate metrics are MB/s, no?

Anyway, here's where NDGF is:

- we can sustain 120MB/s right now over long time to and from tape (~double nominal
rate, IIRC).

- on June first we might be up to 180MB/s, and we can probably come up to 240-300MB/s
sustained, if it is really a high priority. This would involve some changes in internal
priorities and dual tape write&read frontends.

Cheers,
Oxana
```

```
27.04.2009 12:02, Alberto Aimar пишет:
  Dear Colleagues
  in tomorrow's MB agenda there is a Sites' round table:

  In practice Sites should specify how they will provide their MSS
  information
  rates, etc (i.e. any useful metric) for STEP09 and further.

  If you send me that information (quite) before the meeting I will collect
  it in a single document for the MB meeting.

  Best regards.
  Alberto Aimar.
```

NL-T1 tape benchmarking

From: Jeff Templon <templon@nikhef.nl>
Date: Mon, 27 Apr 2009 13:25:28 +0200
To: Alberto Aimar <alberto.aimar@cern.ch>
CC: Ron Trompert <ron@sara.nl>, Mark van de Sanden <sanden@sara.nl>, Maurice Bouwhuis <bouwhuis@sara.nl>

```
Hi Alberto,

here is the information:

http://ui.grid.sara.nl/monitoring/tape_metrics/LHCVOS/LHCVOS_tape_metrics.html
shows how fast we are reading from tape. At
http://ui.grid.sara.nl/monitoring/datatransfer/gridftp_hourly_restore_vo_throughput.html you
see how the rate at which data is copied to the dCache pools.

thanks to Ron for providing it so quickly.

                JT
```

TW-ASGC Metrics from the Tier-1 Sites

From: Jason Shih <hlshih@gate.sinica.edu.tw>
Date: Tue, 28 Apr 2009 04:03:07 +0800
To: Alberto Aimar <alberto.aimar@cern.ch>

```
Dear Alberto,


with tape logging facility (thanks castor dev), we could provide the tape performance
metrics such as:


- availability of tape
- number of data vol have been access (read/write) in certain time period
- number of files being migrated/re-call
- total wall clock of particular vo in certain period of time dealing with data
migration/recall
- total elapse time referring to tape drives
- total mount elapse time, per vo view scope, service classes)
- elapse time for position, transfer, and also umount

monitoring pages available are:

- tape xfer (data migration and recall)
- per vo view scope of data migration/recall to/from tape
- per service class of data migration/recall to/from tape



BR,
J

Alberto Aimar wrote:
  Dear Colleagues
     in tomorrow's MB agenda there is a Sites' round table:

  In practice Sites should specify how they will provide their MSS information
  rates, etc (i.e. any useful metric) for STEP09 and further.

  If you send me that information (quite) before the meeting I will collect
  it in a single document for the MB meeting.

  Best regards.
  Alberto Aimar.


--
Jason Shih
ASGC/OPS
Tel: +886-2-2789-8311
Fax: +886-2-2783-7653
```

UK-T1-RAL MSS Metrics from the Tier-1 Sites

From: "Gordon, JC (John)" <john.gordon@stfc.ac.uk>
Date: Tue, 28 Apr 2009 14:39:30 +0100
To: Alberto Aimar <alberto.aimar@cern.ch>
CC: "Sansum, RA (Andrew)" <andrew.sansum@stfc.ac.uk>, David Britton
<d.britton@physics.gla.ac.uk>

```
For RAL

1) We have the statistical analysis listed at:
https://twiki.cern.ch/twiki/bin/view/LCG/MssEfficiencyUK-T1-RAL

    This is a manually updated set of tables which shows the bandwidth to/from tape for
each VO, irrespective of whether the drives are shared or dedicated. The figures are
weekly averages. We are considering automatic data collection and moving to daily
averages.

2) Internally we have cacti/ganglia plots of network I/O rate aggregated over whole
CASTOR service classes.

John
```

```
  -----Original Message-----
  From: Alberto Aimar [mailto:alberto.aimar@cern.ch]
  Sent: 28 April 2009 13:41
  To: Oxana Smirnova
  Cc: worldwide-lcg-management-board (LCG Management Board)
  Subject: Re: MSS Metrics from the Tier-1 Sites

  Hello Oxana
     I was asking how can the MSS metrics be read, not the rates that are
  planned at the Sites.

  Are there monitoring pages available? URLs?  Nagios channels, ? SLS?
  Like some other sites have sent.

  FYI I attach what I have received until now from the T1 sites.
  And will attach the most recent version to the MB agenda but I hope I
  receive some more information before.

  Cheers.
  Alberto.


  Oxana Smirnova wrote:
    Hi Alberto,

    I'm not sure I understand the question "how", and the rate metrics
  are
    MB/s, no?

    Anyway, here's where NDGF is:

    - we can sustain 120MB/s right now over long time to and from tape
    (~double nominal rate, IIRC).

    - on June first we might be up to 180MB/s, and we can probably come
  up
    to 240-300MB/s sustained, if it is really a high priority. This would
    involve some changes in internal priorities and dual tape write&read
    frontends.

    Cheers,
    Oxana

    27.04.2009 12:02, Alberto Aimar пишет:
      Dear Colleagues
      in tomorrow's MB agenda there is a Sites' round table:
```

US-FNAL-CMS MSS Metrics from the Tier-1 Sites

From: Ian Fisk <ifisk@fnal.gov>
Date: Mon, 27 Apr 2009 12:34:15 -0500
To: Alberto Aimar <alberto.aimar@cern.ch>

Dear Alberto,

We expect to continue to report the MSS metrics on the TWIKI

https://twiki.cern.ch/twiki/bin/view/LCG/MssEfficiency

We've updated these weekly for over a year.   Did I miss a mail to stop providing this information?

Thanks, Ian

On Apr 27, 2009, at 5:02 AM, Alberto Aimar wrote:

> Dear Colleagues
>   in tomorrow's MB agenda there is a Sites' round table:
>
> In practice Sites should specify how they will provide their MSS information
> rates, etc (i.e. any useful metric) for STEP09 and further.
>
> If you send me that information (quite) before the meeting I will collect
> it in a single document for the MB meeting.
>
> Best regards.
> Alberto Aimar.

US-T1-BNL MSS Metrics from the Tier-1 Sites - update

From: "Ernst, Michael" <mernst@bnl.gov>
Date: Mon, 27 Apr 2009 07:52:16 -0400
To: Alberto Aimar <alberto.aimar@cern.ch>

```
I forgot to mention 2 more important things

1) A graph showing the entire data flow performance from user via dCache
via HPSS disk cache to tape (GBytes Flow of
-->dCache-->HPSSdisk-->HPSStape on Hourly Basis)

https://www.racf.bnl.gov/Facility/HPSS/Monitoring/dCacheHPSS/dCacheHPSS.
html

2) A graph showing the data flow from tape to dCache

GBytes Flow of HPSSTape-->HPSSdisk-->dCache on Hourly Basis

https://www.racf.bnl.gov/Facility/HPSS/Monitoring/HPSSdCache/HPSSdCache.
html


The entire menu of all these detailed views can be found at

https://www.racf.bnl.gov/Facility/HPSS/Monitoring/forUsers/atlas_general
stats.html

--
   Michael
-----Original Message-----
From: Ernst, Michael
Sent: Monday, April 27, 2009 7:45 AM
To: Alberto Aimar
Subject: RE: MSS Metrics from the Tier-1 Sites

Alberto,
 At the US ATLAS Tier-1 center we have a variety of MSS related
monitoring pages:

1) A high-level overview of the staging performance, including
information like number of staging requests received and processed,
total number of bytes staged and average file size. The page is publicly
available at

https://www.racf.bnl.gov/Facility/HPSS/Monitoring/summary/summary.htm

2) A comprehensive overview as to migration and retrieval operations
with information on the dataflow performance in both directions (hourly
data rate, presented as graph and as list), e.g.

https://www.racf.bnl.gov/Facility/HPSS/Monitoring/driveFlow/atlasLtoDriv
eFlow.html

for migration presented in form of a list

https://www.racf.bnl.gov/Facility/HPSS/Monitoring/migration2/data/atlas_
list.txt

tape mount count over time, e.g.

https://www.racf.bnl.gov/Facility/HPSS/Monitoring/driveMounts/atlasLtoDr
iveMounts.html

number of drives being used in total over time, e.g.

https://www.racf.bnl.gov/Facility/HPSS/Monitoring/driveUsed/atlasLtoDriv
eUsed.html

number of drives being used for migration versus staging over time, e.g.

https://www.racf.bnl.gov/Facility/HPSS/Monitoring/driveUsed/more_atlasLt
```

oDriveUsed.html

HPSS disk cache utilization over time

https://www.racf.bnl.gov/Facility/HPSS/Monitoring/diskUsage2/atlas_sc59.
html

Total ATLAS archive (total data volume and total number of files
archived)

https://www.racf.bnl.gov/Facility/HPSS/Monitoring/tapeUsage4/atlas_sc34.
monthly.php

All links above are supposed to be publicly viewable. If not, please let
me know.

Regards,
Michael
-----Original Message-----
From: Alberto Aimar [mailto:alberto.aimar@cern.ch]
Sent: Monday, April 27, 2009 6:02 AM
To: worldwide-lcg-management-board (LCG Management Board)
Subject: MSS Metrics from the Tier-1 Sites

Dear Colleagues
    in tomorrow's MB agenda there is a Sites' round table:

In practice Sites should specify how they will provide their MSS
information
rates, etc (i.e. any useful metric) for STEP09 and further.

If you send me that information (quite) before the meeting I will
collect
it in a single document for the MB meeting.

Best regards.
Alberto Aimar.