Lustre @ GSI

Hepix Spring Meeting 2011

Thomas Roth

GSI Darmstadt

4. May 2010

Lustre @ GSI

- Lustre @ GSI = success story

Lustre @ GSI

- Lustre @ GSI = success story
- Start with production cluster: October 2008

Lustre @ GSI

- Lustre @ GSI = success story
- Start with production cluster: October 2008
- Center of everything

Lustre @ GSI

- Lustre @ GSI = success story
- Start with production cluster: October 2008
- Center of everything
- Capacity of Cluster now 1.2 PB

Lustre @ GSI

- Lustre @ GSI = success story
- Start with production cluster: October 2008
- Center of everything
- Capacity of Cluster now 1.2 PB

☞ 1 PB used, 10^8 files

Lustre @ GSI

- Lustre @ GSI = success story
- Start with production cluster: October 2008
- Center of everything
- Capacity of Cluster now 1.2 PB

  ☞ 1 PB used, 10^8 files
- Lustre servers: 105 ➡ soon 144

Lustre @ GSI

- Lustre @ GSI = success story
- Start with production cluster: October 2008
- Center of everything
- Capacity of Cluster now 1.2 PB

  ☞ 1 PB used, 10^8 files
- Lustre servers: 105 ➡ soon 144
- Clients connected with 160 Gb/s

Lustre @ GSI

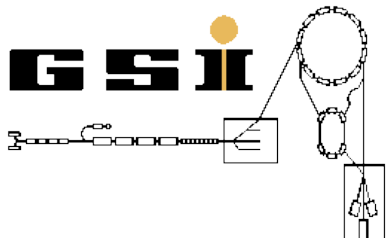
- Lustre @ GSI = success story
- Start with production cluster: October 2008
- Center of everything
- Capacity of Cluster now 1.2 PB

  ☞ 1 PB used, 10^8 files
- Lustre servers: 105 ➡ soon 144
- Clients connected with 160 Gb/s

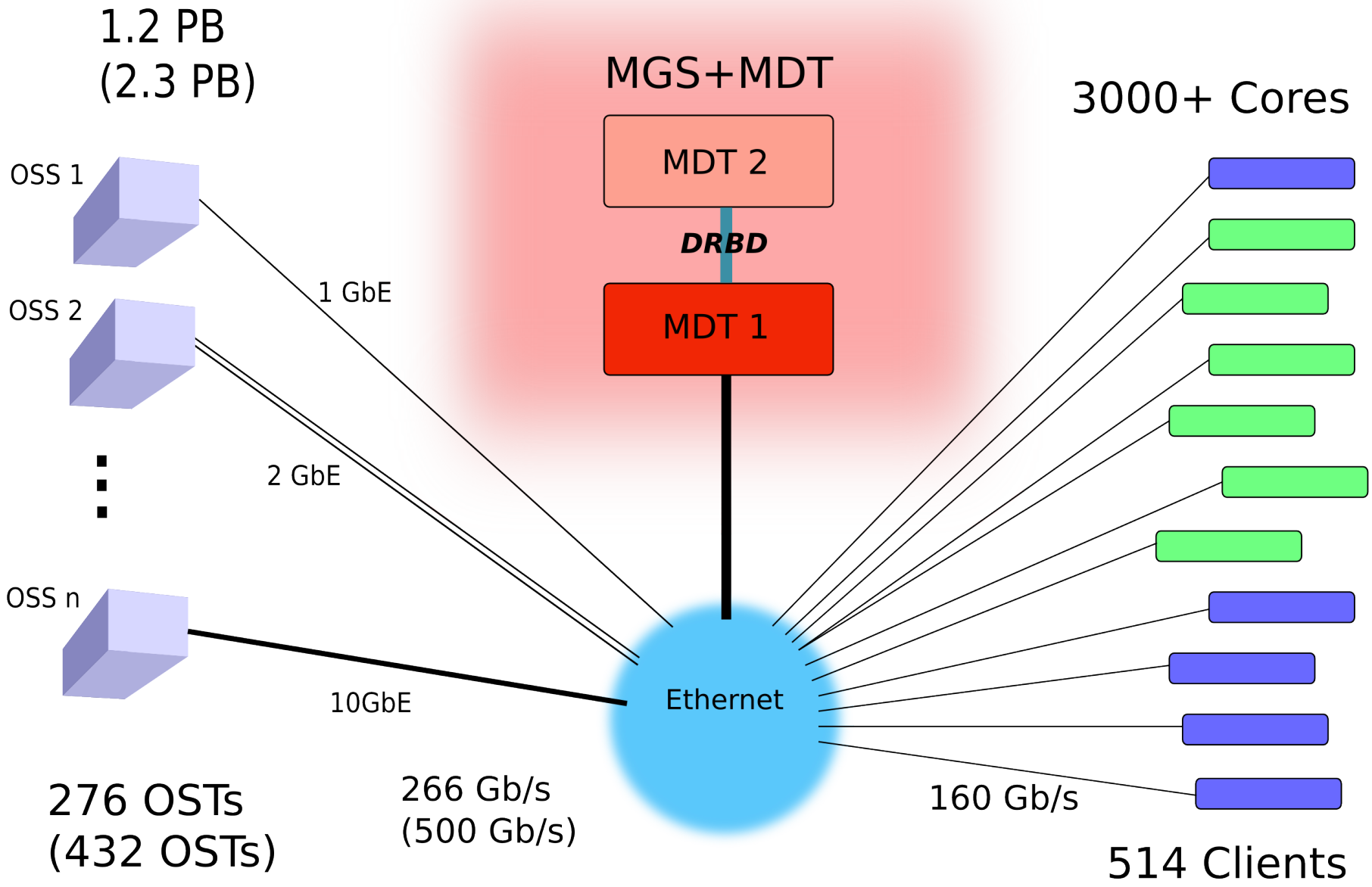
  ☞ seen Alice analysis train @ 90 Gb/s

Outline

- GSI Lustre Setup
- Lustre version upgrade
- Hardware Upgrade of MDS
- Future plans for Lustre

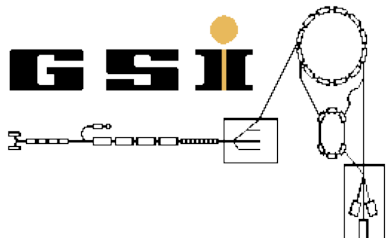


GSI setup


1.2 PB
(2.3 PB)
MGS+MDT
3000+ Cores
MDT 2
DRBD
MDT 1
OSS 1
OSS 2
OSS n
1 GbE
2 GbE
10GbE
Ethernet
276 OSTs
(432 OSTs)
266 Gb/s
(500 Gb/s)
160 Gb/s
514 Clients

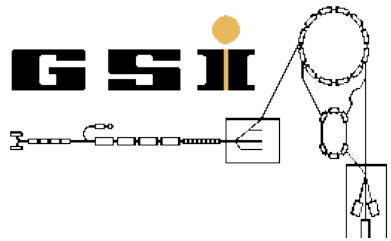
Lustre kernel

- GSI Lustre kernel: port to Debian of vanilla/Suse kernel + Lustre



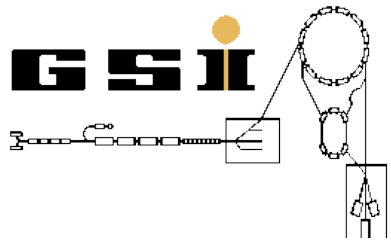
Lustre kernel

- GSI Lustre kernel: port to Debian of vanilla/Suse kernel + Lustre
- provided by software company: Debian kernel hackers, now maintainers of Debian Lustre port



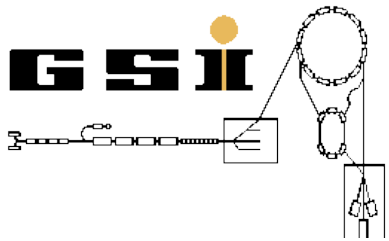
Lustre kernel

- GSI Lustre kernel: port to Debian of vanilla/Suse kernel + Lustre
- provided by software company: Debian kernel hackers, now maintainers of Debian Lustre port
- http://pkg-lustre.alioth.debian.org/

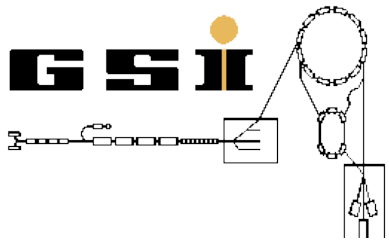


Lustre kernel

- GSI Lustre kernel: port to Debian of vanilla/Suse kernel + Lustre
- provided by software company: Debian kernel hackers, now maintainers of Debian Lustre port
- http://pkg-lustre.alioth.debian.org/
- ☺ GSI supports Open Source!



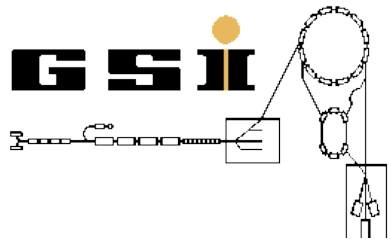
History: Version upgrade



History: Version upgrade

Lustre Crashes

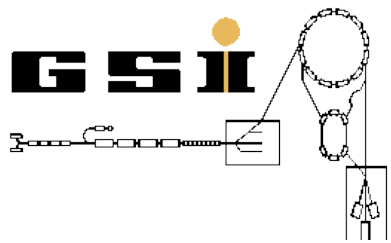
- Until Sep 2010 running Lustre v. 1.6.7.2



History: Version upgrade

Lustre Crashes

- Until Sep 2010 running Lustre v. 1.6.7.2
- Rather frequent LBUGS

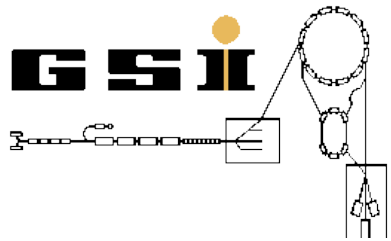


History: Version upgrade

Lustre Crashes

- Until Sep 2010 running Lustre v. 1.6.7.2
- Rather frequent LBUGS

```
(mds_reint.c:1606:mds_orphan_add_link())
ASSERTION(inode->i_nlink = 1) failed:dir link = 0
```



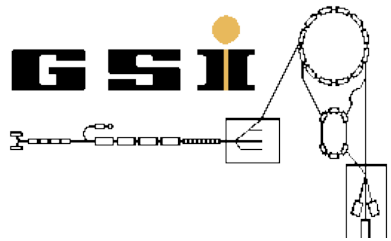
History: Version upgrade

Lustre Crashes

- Until Sep 2010 running Lustre v. 1.6.7.2
- Rather frequent LBUGS

```
(mds_reint.c:1606:mds_orphan_add_link())
ASSERTION(inode->i_nlink = 1) failed:dir link = 0
```

- found set of user scripts that could trigger this LBUG



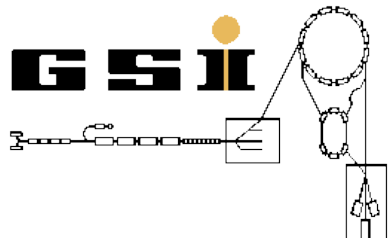
History: Version upgrade

Lustre Crashes

- Until Sep 2010 running Lustre v. 1.6.7.2
- Rather frequent LBUGS

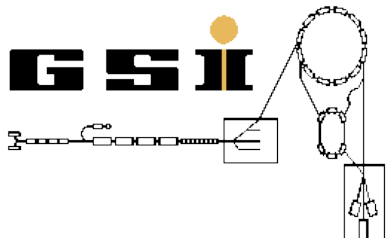
```
(mds_reint.c:1606:mds_orphan_add_link())
ASSERTION(inode->i_nlink = 1) failed:dir link = 0
```

- found set of user scripts that could trigger this LBUG
- shutdown with writeconf necessary April 2010



Upgrade to 1.8.4

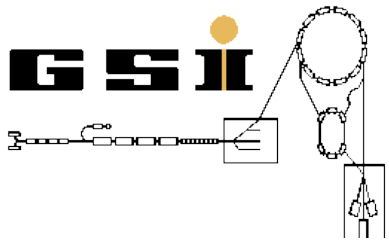
Sep 2010: Upgrade to Lustre v. 1.8.4



Upgrade to 1.8.4

Sep 2010: Upgrade to Lustre v. 1.8.4

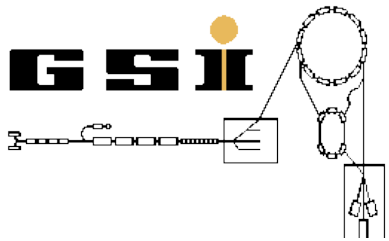
+ upgrade to Debian Lenny



Upgrade to 1.8.4

Sep 2010: Upgrade to Lustre v. 1.8.4

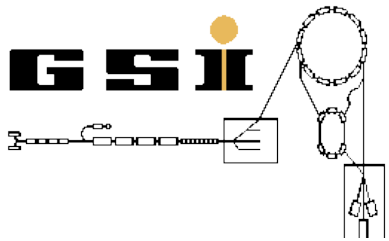
+ upgrade to Debian Lenny
+ FS-check of all OSTs



Upgrade to 1.8.4

Sep 2010: Upgrade to Lustre v. 1.8.4

+ upgrade to Debian Lenny
+ FS-check of all OSTs
+ hardware repairs (controller, BBUs)

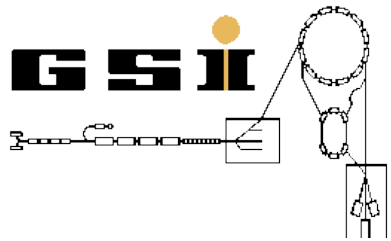


Upgrade to 1.8.4

Sep 2010: Upgrade to Lustre v. 1.8.4

+ upgrade to Debian Lenny
+ FS-check of all OSTs
+ hardware repairs (controller, BBUs)

☺ After upgrade: no more LBUGS ☺



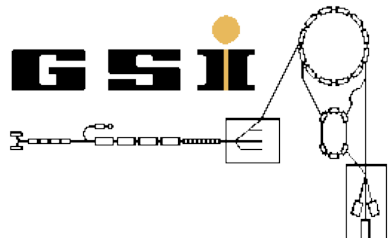
Upgrade to 1.8.4

Sep 2010: Upgrade to Lustre v. 1.8.4

+ upgrade to Debian Lenny
+ FS-check of all OSTs
+ hardware repairs (controller, BBUs)

☺ After upgrade: no more LBUGS ☺

- All subsequent crashes: hardware issues



Upgrade to 1.8.4

Adaptec bashing

Upgrade to 1.8.4

Adaptec bashing

- Adaptec RAID Controller 5401 *booting* longer than entire Linux system

Upgrade to 1.8.4

Adaptec bashing

- Adaptec RAID Controller 5401 *booting* longer than entire Linux system
- nasty features to make admin's life harder (handling of broken disks etc.)

Upgrade to 1.8.4

Adaptec bashing

Adaptec RAID Controller 5401 *booting* longer than entire Linux system

nasty features to make admin's life harder (handling of broken disks etc.)

Controller complaining about high temperatures + missing fans of non-existing enclosures

Upgrade to 1.8.4

Adaptec bashing


Adaptec RAID Controller 5401 *booting* longer than entire Linux system


nasty features to make admin's life harder (handling of broken disks etc.)


Controller complaining about high temperatures + missing fans of non-existing enclosures


System crashes due to

Upgrade to 1.8.4

Adaptec bashing


Adaptec RAID Controller 5401 *booting* longer than entire Linux system


nasty features to make admin's life harder (handling of broken disks etc.)


Controller complaining about high temperatures + missing fans of non-existing enclosures


System crashes due to

```
AAC0: adapter kernel panic'd
```

Upgrade to 1.8.4

Adaptec bashing


Adaptec RAID Controller 5401 *booting* longer than entire Linux system


nasty features to make admin's life harder (handling of broken disks etc.)


Controller complaining about high temperatures + missing fans of non-existing enclosures

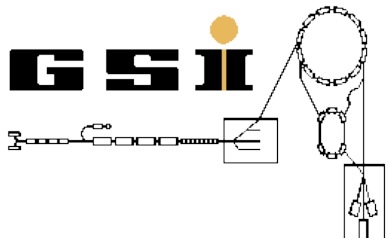

System crashes due to

```
AAC0: adapter kernel panic'd
```

```
aacraid: Host adapter abort request
```

MDS issues

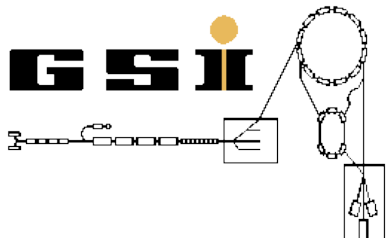
Good ole' MDS hardware



MDS issues

Good ole' MDS hardware

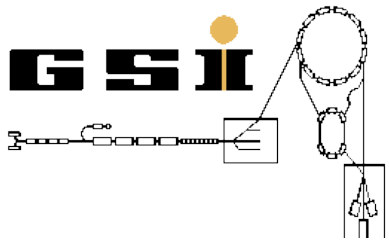
- 3U SuperMicro file server, 8 cores, 32GB RAM, 16 HD Raptor disks



MDS issues

Good ole' MDS hardware

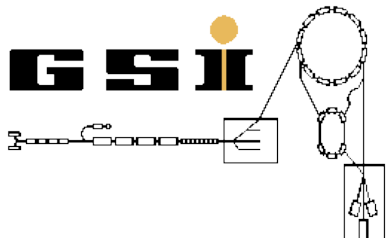
- 3U SuperMicro file server, 8 cores, 32GB RAM, 16 HD Raptor disks
- worn-out hardware



MDS issues

Good ole' MDS hardware

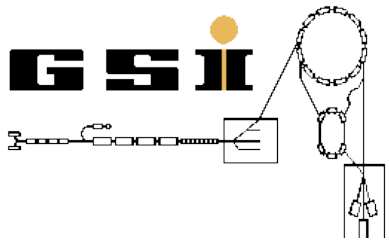
- 3U SuperMicro file server, 8 cores, 32GB RAM, 16 HD Raptor disks
- worn-out hardware
- RAID controller becoming cranky: ejecting several disks at once



MDS issues

Good ole' MDS hardware

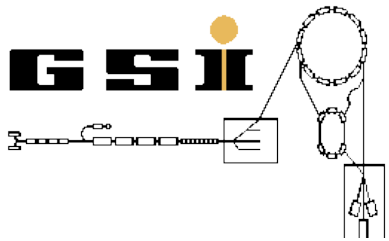
- 3U SuperMicro file server, 8 cores, 32GB RAM, 16 HD Raptor disks
- worn-out hardware
- RAID controller becoming cranky: ejecting several disks at once
- more and more overloaded



MDS issues

Good ole' MDS hardware

- 3U SuperMicro file server, 8 cores, 32GB RAM, 16 HD Raptor disks
- worn-out hardware
- RAID controller becoming cranky: ejecting several disks at once
- more and more overloaded
- too small (2×10^8 inodes only)

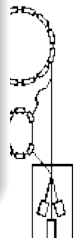


MDS issues

Good ole' MDS hardware

- 3U SuperMicro file server, 8 cores, 32GB RAM, 16 HD Raptor disks
- worn-out hardware
- RAID controller becoming cranky: ejecting several disks at once
- more and more overloaded
- too small (2×10^8 inodes only)

New MDS hardware



MDS issues

Good ole' MDS hardware

- 3U SuperMicro file server, 8 cores, 32GB RAM, 16 HD Raptor disks
- worn-out hardware
- RAID controller becoming cranky: ejecting several disks at once
- more and more overloaded
- too small (2×10^8 inodes only)

New MDS hardware

- 2U SuperMicro box, 48 AMD Cores, 128 GB RAM

MDS issues

Good ole' MDS hardware

- 3U SuperMicro file server, 8 cores, 32GB RAM, 16 HD Raptor disks
- worn-out hardware
- RAID controller becoming cranky: ejecting several disks at once
- more and more overloaded
- too small (2×10^8 inodes only)

New MDS hardware

- 2U SuperMicro box, 48 AMD Cores, 128 GB RAM
- 4U Infortrend Enclosure attached via Fibre Channel

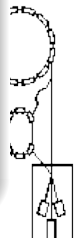
MDS issues

Good ole' MDS hardware

- 3U SuperMicro file server, 8 cores, 32GB RAM, 16 HD Raptor disks
- worn-out hardware
- RAID controller becoming cranky: ejecting several disks at once
- more and more overloaded
- too small (2×10^8 inodes only)

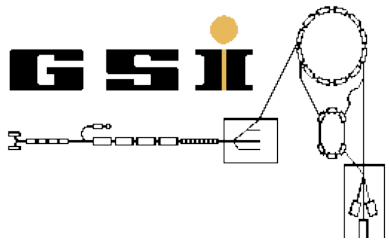
New MDS hardware

- 2U SuperMicro box, 48 AMD Cores, 128 GB RAM
- 4U Infortrend Enclosure attached via Fibre Channel
- 24 Cheetah 15k SAS disks, 300 GB



MDS upgrade

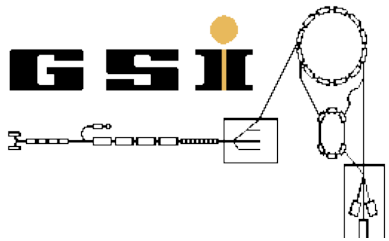
2 issues



MDS upgrade

2 issues

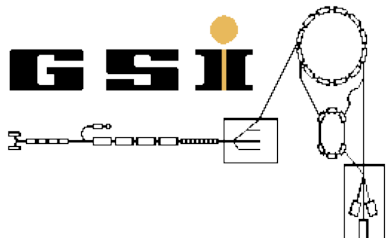
1. Copying the MDT



MDS upgrade

2 issues

1. Copying the MDT
2. Increasing size of MDT

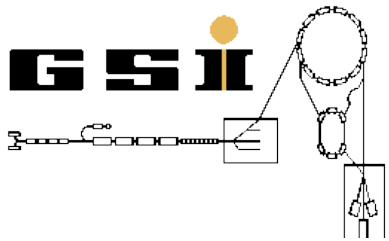


MDS upgrade

2 issues

1. Copying the MDT
2. Increasing size of MDT

- rsync copy: would do (old, small) ➡ (new, large)

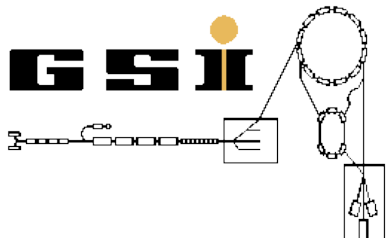


MDS upgrade

2 issues

1. Copying the MDT
2. Increasing size of MDT

- rsync copy: would do (old, small) ➡ (new, large)
- never seems to finish:

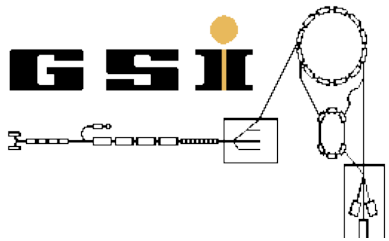


MDS upgrade

2 issues

1. Copying the MDT
2. Increasing size of MDT

- rsync copy: would do (old, small) ➡ (new, large)
- never seems to finish:
  - too many files (10^8) ?

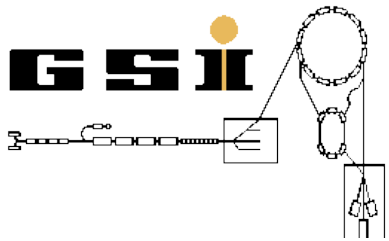


MDS upgrade

2 issues

1. Copying the MDT
2. Increasing size of MDT

- rsync copy: would do (old, small) ➡ (new, large)
- never seems to finish:
  - too many files (10^8) ?
  - too many sparse files ?

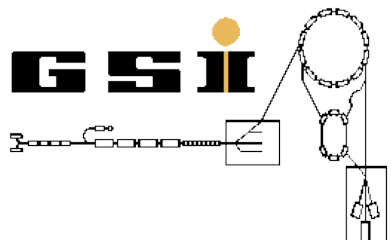


MDS upgrade

2 issues

1. Copying the MDT
2. Increasing size of MDT

- rsync copy: would do (old, small) ➡ (new, large)
- never seems to finish:
  - too many files (10^8) ?
  - too many sparse files ?
- DRBD copy: fast, but copy has original size

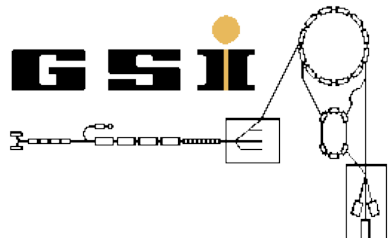


MDS upgrade

2 issues

1. Copying the MDT
2. Increasing size of MDT

- rsync copy: would do (old, small) ➡ (new, large)
- never seems to finish:
  - too many files (10^8) ?
  - too many sparse files ?
- DRBD copy: fast, but copy has original size
- dd: similar to DRBD



MDS upgrade

2 issues

1. Copying the MDT
2. Increasing size of MDT

- rsync copy: would do (old, small) ➡ (new, large)
- never seems to finish:
  - too many files (10^8) ?
  - too many sparse files ?
- DRBD copy: fast, but copy has original size
- dd: similar to DRBD

2 steps

MDS upgrade

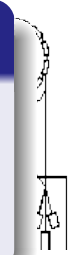
2 issues

1. Copying the MDT
2. Increasing size of MDT

- rsync copy: would do (old, small) ➡ (new, large)
- never seems to finish:
  - too many files (10^8) ?
  - too many sparse files ?
- DRBD copy: fast, but copy has original size
- dd: similar to DRBD

2 steps

1. DRBD copy



MDS upgrade

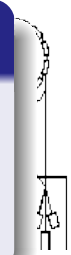
2 issues

1. Copying the MDT
2. Increasing size of MDT

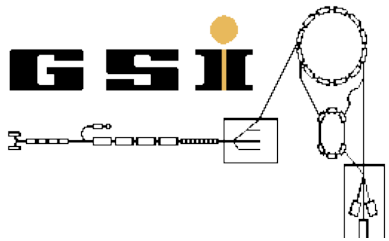
- rsync copy: would do (old, small) ➡ (new, large)
- never seems to finish:
  - too many files (10^8) ?
  - too many sparse files ?
- DRBD copy: fast, but copy has original size
- dd: similar to DRBD

2 steps

1. DRBD copy
2. *resize2fs* to enlarge copy

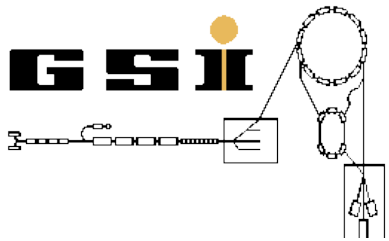


MDS upgrade



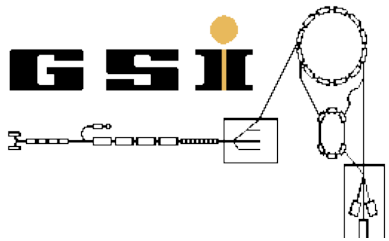
MDS upgrade

- Planned for March 24



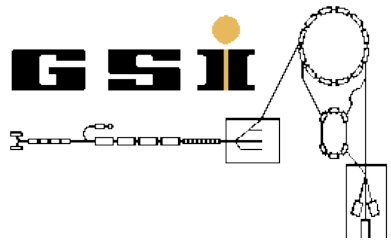
MDS upgrade

- Planned for March 24
- old MDS gave up 36 h early



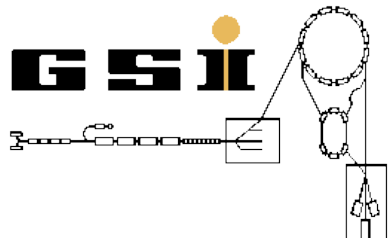
MDS upgrade

- Planned for March 24
- old MDS gave up 36 h early
- DRBD-copy to new machine already in place



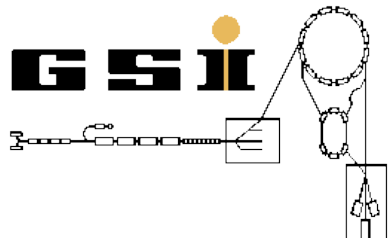
MDS upgrade

- Planned for March 24
- old MDS gave up 36 h early
- DRBD-copy to new machine already in place
- Resizing



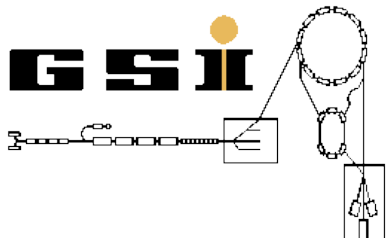
MDS upgrade

- Planned for March 24
- old MDS gave up 36 h early
- DRBD-copy to new machine already in place
- Resizing ✔



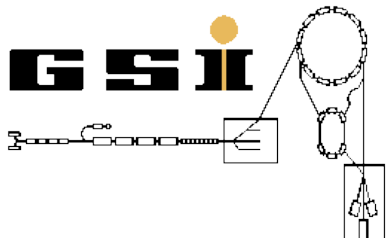
MDS upgrade

- Planned for March 24
- old MDS gave up 36 h early
- DRBD-copy to new machine already in place
  - Resizing ✔
  - Restart + Recovery



MDS upgrade

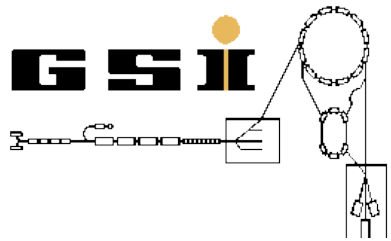
- Planned for March 24
- old MDS gave up 36 h early
- DRBD-copy to new machine already in place
  - Resizing ✔
  - Restart + Recovery ✔



MDS upgrade

- Planned for March 24
- old MDS gave up 36 h early
- DRBD-copy to new machine already in place
  - Resizing ✔
  - Restart + Recovery ✔

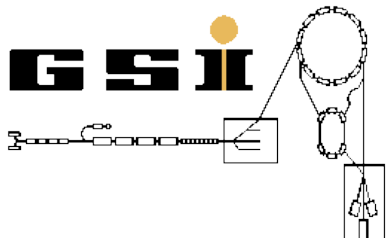
UPS! Lustre extremely slow



MDS upgrade

☹12 days of agony

Searching for errors, trying different hardware, passing weekends and nights at GSI ...

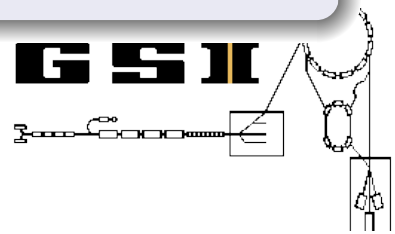


MDS upgrade

☹12 days of agony

Searching for errors, trying different hardware, passing weekends and nights at GSI ...

What has changed?



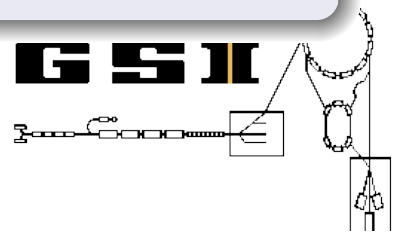
MDS upgrade

☹12 days of agony

Searching for errors, trying different hardware, passing weekends and nights at GSI ...

What has changed?

- new hardware: different CPU architecture



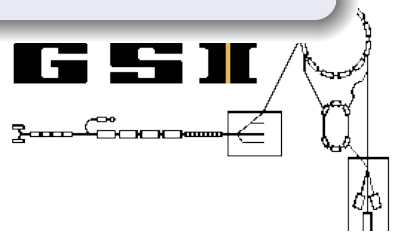
MDS upgrade

☹12 days of agony

Searching for errors, trying different hardware, passing weekends and nights at GSI ...

What has changed?

- new hardware: different CPU architecture
- new hardware: 10 GbE



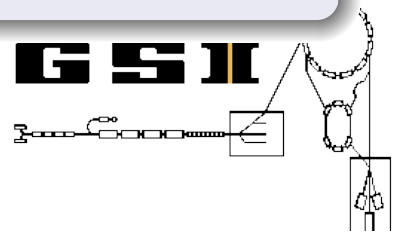
MDS upgrade

☹12 days of agony

Searching for errors, trying different hardware, passing weekends and nights at GSI ...

What has changed?

- new hardware: different CPU architecture
- new hardware: 10 GbE
- new hardware: FC disk enclosure



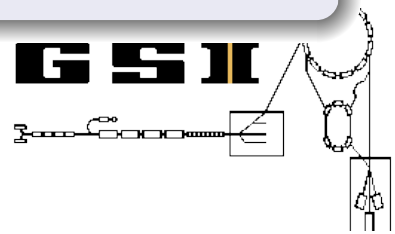
MDS upgrade

☹12 days of agony

Searching for errors, trying different hardware, passing weekends and nights at GSI ...

What has changed?

- new hardware: different CPU architecture
- new hardware: 10 GbE
- new hardware: FC disk enclosure
- MDT was enlarged



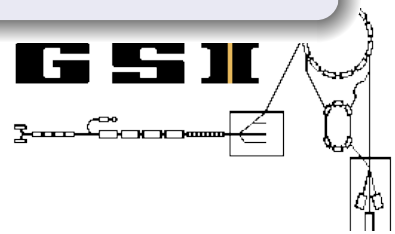
MDS upgrade

☹12 days of agony

Searching for errors, trying different hardware, passing weekends and nights at GSI ...

What has changed?

- new hardware: different CPU architecture
- new hardware: 10 GbE
- new hardware: FC disk enclosure
- MDT was enlarged
- new Alice analysis with new data



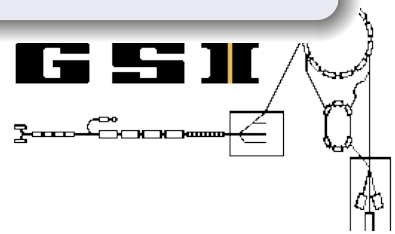
MDS upgrade

☹12 days of agony

Searching for errors, trying different hardware, passing weekends and nights at GSI ...

What has changed?

- new hardware: different CPU architecture
- new hardware: 10 GbE
- new hardware: FC disk enclosure
- MDT was enlarged
- new Alice analysis with new data
  - ↔ old analysis with old data reasonably fast



MDS upgrade

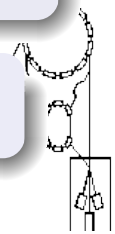
☹12 days of agony

Searching for errors, trying different hardware, passing weekends and nights at GSI ...

What has changed?

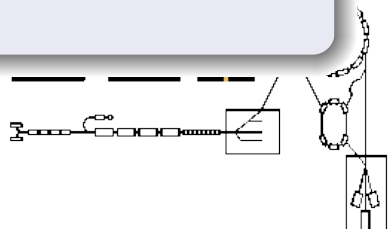
- new hardware: different CPU architecture
- new hardware: 10 GbE
- new hardware: FC disk enclosure
- MDT was enlarged
- new Alice analysis with new data
  - ↔ old analysis with old data reasonably fast

No really satisfactory explanation found



MDS upgrade

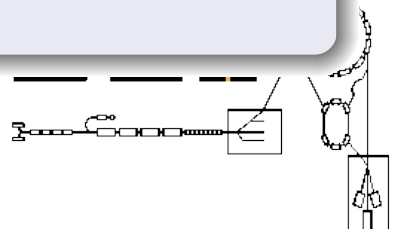
Action



MDS upgrade

Action

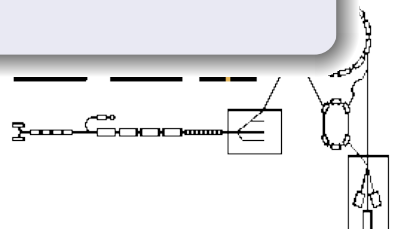
- frantic reconstruction of old hardware by cannibalizing other machines - MDT running there: slightly promising



MDS upgrade

Action

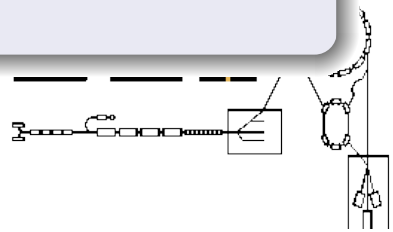
- frantic reconstruction of old hardware by cannibalizing other machines - MDT running there: slightly promising
- complete shutdown of Lustre



MDS upgrade

Action

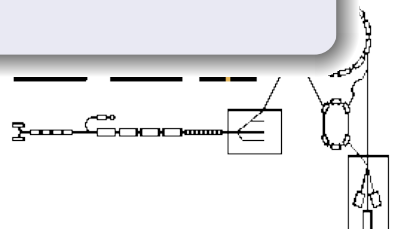
- frantic reconstruction of old hardware by cannibalizing other machines - MDT running there: slightly promising
- complete shutdown of Lustre
  - umount / reset of all clients



MDS upgrade

Action

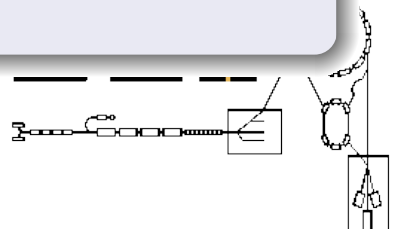
- frantic reconstruction of old hardware by cannibalizing other machines - MDT running there: slightly promising
- complete shutdown of Lustre
  - umount / reset of all clients
  - umount of all OSTs



MDS upgrade

Action

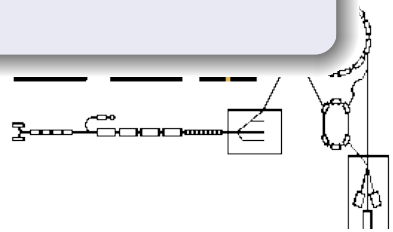
- frantic reconstruction of old hardware by cannibalizing other machines - MDT running there: slightly promising
- complete shutdown of Lustre
  - umount / reset of all clients
  - umount of all OSTs
  - verify there is no more traffic to the MGS



MDS upgrade

Action

- frantic reconstruction of old hardware by cannibalizing other machines - MDT running there: slightly promising
- complete shutdown of Lustre
  - umount / reset of all clients
  - umount of all OSTs
  - verify there is no more traffic to the MGS
- writeconf all servers

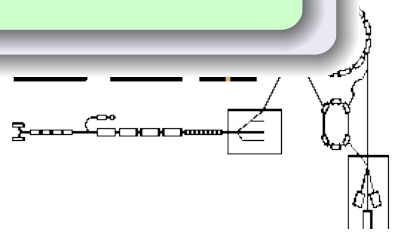


MDS upgrade

Action

- frantic reconstruction of old hardware by cannibalizing other machines - MDT running there: slightly promising
- complete shutdown of Lustre
  - umount / reset of all clients
  - umount of all OSTs
  - verify there is no more traffic to the MGS
- writeconf all servers

☺ Restart of MDT on new hardware ✔



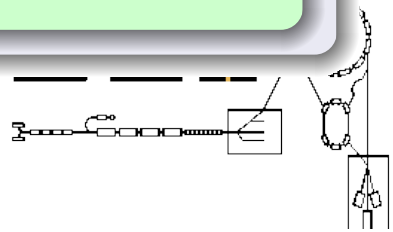
MDS upgrade

Action

- frantic reconstruction of old hardware by cannibalizing other machines - MDT running there: slightly promising
- complete shutdown of Lustre
  - umount / reset of all clients
  - umount of all OSTs
  - verify there is no more traffic to the MGS
- writeconf all servers

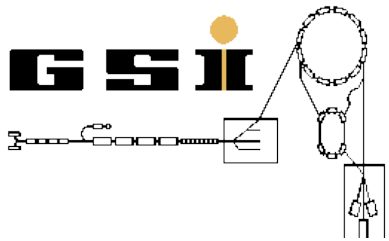
☺ Restart of MDT on new hardware ✔

☺ Speed ✔



MDS upgrade

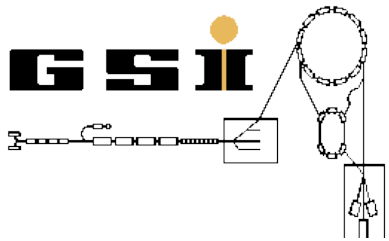
Major issue



MDS upgrade

Major issue

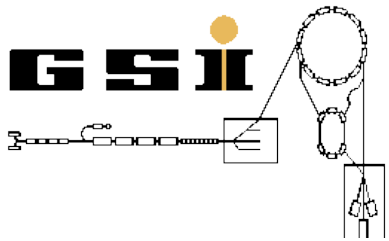
- IP of old MGS ➜ virtual IP on new box



MDS upgrade

Major issue

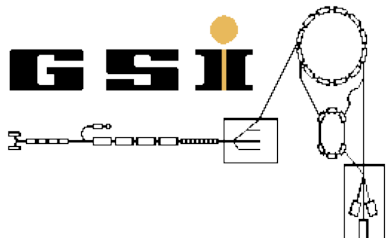
- IP of old MGS ➜ virtual IP on new box
- ☞ Some OSTs: network traffic to machine IP instead of MGS-IP



MDS upgrade

Major issue

- IP of old MGS ➜ virtual IP on new box
- Some OSTs: network traffic to machine IP instead of MGS-IP
- Perhaps misconfigured LNET at first MGS restart ?

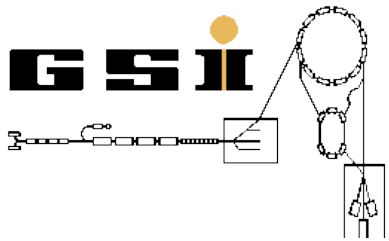


MDS upgrade

Major issue

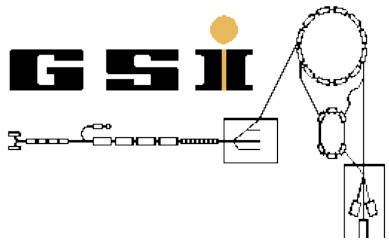
- ▲ IP of old MGS ➜ virtual IP on new box
- ☞ Some OSTs: network traffic to machine IP instead of MGS-IP
- Perhaps misconfigured LNET at first MGS restart ?

- After writeconf and restart no more stray traffic



MDS upgrade

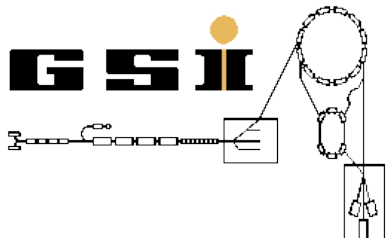
Lessons learned



MDS upgrade

Lessons learned

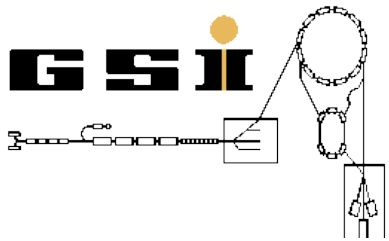
- ✓ Resizing MDT with *resize2fs* works



MDS upgrade

Lessons learned

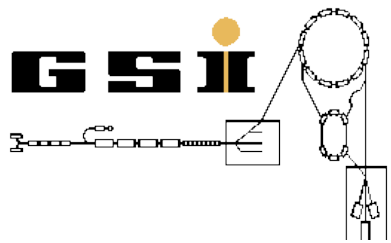
- ✓ Resizing MDT with *resize2fs* works
- ☞ Careful with LNET configuration



MDS upgrade

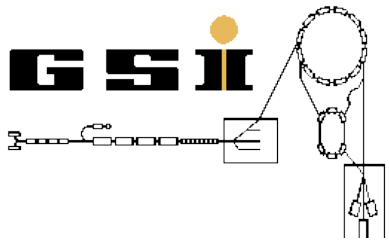
Lessons learned

- ✓ Resizing MDT with *resize2fs* works
- ☞ Careful with LNET configuration
- ☞ Don't do such things on the fly, in a hurry, with the rest of the system under heavy load



Future plans

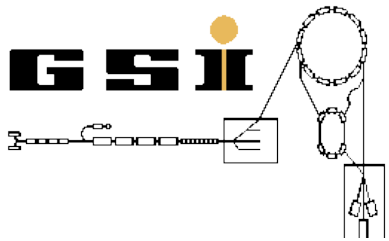
Increase Capacity



Future plans

Increase Capacity

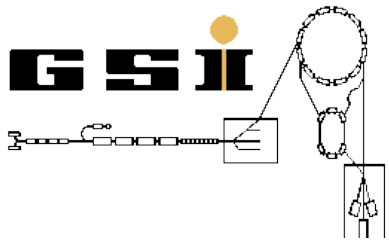
- Container with 10 racks



Future plans

Increase Capacity

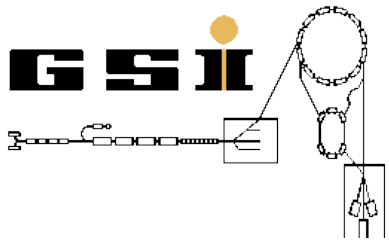
- Container with 10 racks Bouillon Cube



Future plans

Increase Capacity

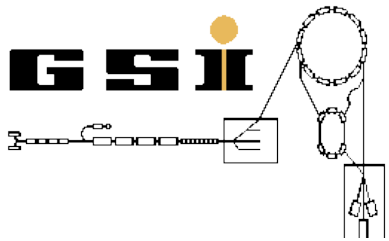
- Container with 10 racks Bouillon Cube
- 5 OSS of 4U + 4U boxes per rack



Future plans

Increase Capacity

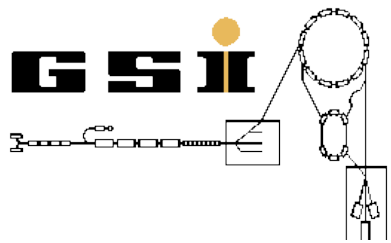
- Container with 10 racks Bouillon Cube
- 5 OSS of 4U + 4U boxes per rack
- 1 TB disks: keep fsck times finite



Future plans

Increase Capacity

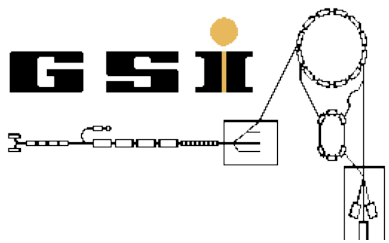
- Container with 10 racks Bouillon Cube
- 5 OSS of 4U + 4U boxes per rack
- 1 TB disks: keep fsck times finite
- Size of OST: keep below 8 TB, in line with Lustre recommendations



Future plans

Increase Capacity

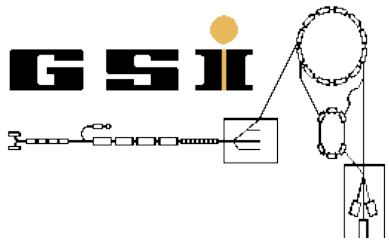
- Container with 10 racks Bouillon Cube
- 5 OSS of 4U + 4U boxes per rack
- 1 TB disks: keep fsck times finite
- Size of OST: keep below 8 TB, in line with Lustre recommendations
- capacity of container 1.3 PB



Future plans

Increase Capacity

- Container with 10 racks Bouillon Cube
- 5 OSS of 4U + 4U boxes per rack
- 1 TB disks: keep fsck times finite
- Size of OST: keep below 8 TB, in line with Lustre recommendations
- capacity of container 1.3 PB ➤ 2.3 PB total



Future plans

Increase Capacity

- Container with 10 racks Bouillon Cube
- 5 OSS of 4U + 4U boxes per rack
- 1 TB disks: keep fsck times finite
- Size of OST: keep below 8 TB, in line with Lustre recommendations
- capacity of container 1.3 PB ➤ 2.3 PB total

Future plans

Increase Capacity

- Container with 10 racks Bouillon Cube
- 5 OSS of 4U + 4U boxes per rack
- 1 TB disks: keep fsck times finite
- Size of OST: keep below 8 TB, in line with Lustre recommendations
- capacity of container 1.3 PB ➤ 2.3 PB total

- one rack connected to Lustre

Future plans

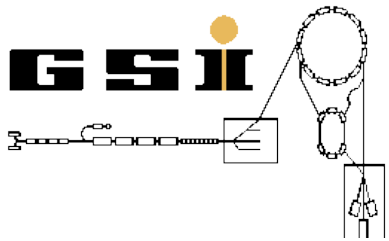
Increase Capacity

- Container with 10 racks Bouillon Cube
- 5 OSS of 4U + 4U boxes per rack
- 1 TB disks: keep fsck times finite
- Size of OST: keep below 8 TB, in line with Lustre recommendations
- capacity of container 1.3 PB ➤ 2.3 PB total

- one rack connected to Lustre
- other racks waiting for metal sheets to provide thermal housing

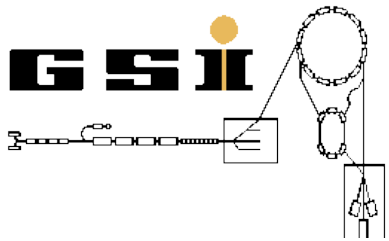
Expansion to testing hall

- Volker Lindenstruth's talk



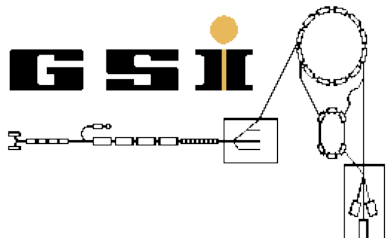
Expansion to testing hall

- ➜ Volker Lindenstruth's talk
  - try to keep future inter-machine communication on Infiniband:



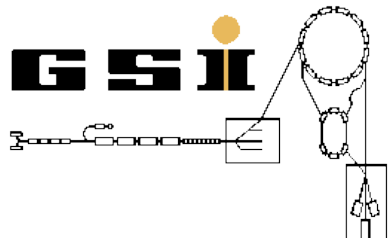
Expansion to testing hall

- ➜ Volker Lindenstruth's talk
  - try to keep future inter-machine communication on Infiniband:
    - perhaps faster



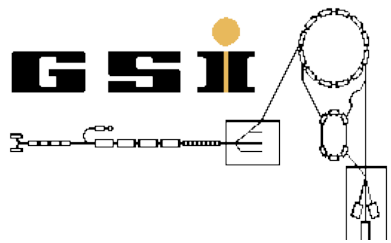
Expansion to testing hall

- ➜ Volker Lindenstruth's talk
  - try to keep future inter-machine communication on Infiniband:
    - perhaps faster
    - more stable (no GSI Ethernet noise)



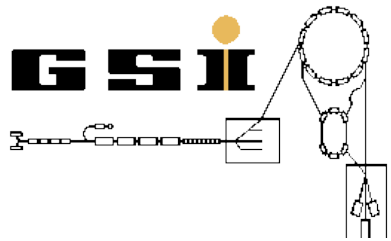
Expansion to testing hall

- ➜ Volker Lindenstruth's talk
  - try to keep future inter-machine communication on Infiniband:
    - perhaps faster
    - more stable (no GSI Ethernet noise)
    - cheaper than 10GbE



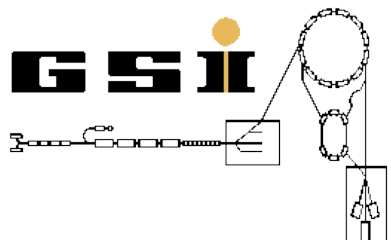
Expansion to testing hall

- ➜ Volker Lindenstruth's talk
  - try to keep future inter-machine communication on Infiniband:
    - perhaps faster
    - more stable (no GSI Ethernet noise)
    - cheaper than 10GbE
  - No expertise / experience @ GSI so far



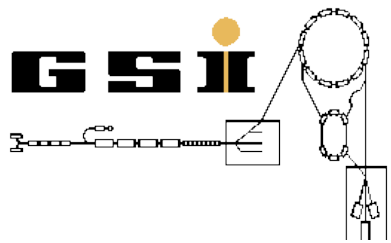
Expansion to testing hall

- ➜ Volker Lindenstruth's talk
- try to keep future inter-machine communication on Infiniband:
  - perhaps faster
  - more stable (no GSI Ethernet noise)
  - cheaper than 10GbE
- No expertise / experience @ GSI so far
- First steps were simple enough, hardware worked out of the box, OFED+Lustre compilation somewhat tedious



Expansion to testing hall

- ➜ Volker Lindenstruth's talk
- try to keep future inter-machine communication on Infiniband:
  - perhaps faster
  - more stable (no GSI Ethernet noise)
  - cheaper than 10GbE
- No expertise / experience @ GSI so far
- First steps were simple enough, hardware worked out of the box, OFED+Lustre compilation somewhat tedious
- Need LNET routers



Lustre @ GSI