

Lustre @ GSI

Hepix Fall Meeting 2012

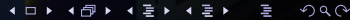
Thomas Roth

GSI Darmstadt

16. October 2010

Lustre @ GSI

- Lustre @ GSI = success story



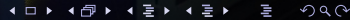
Lustre @ GSI

- Lustre @ GSI = success story
- Old cluster: October 2008



Lustre @ GSI

- Lustre @ GSI = success story
- Old cluster: October 2008
- Capacity 2.3 PB



Lustre @ GSI

- Lustre @ GSI = success story
- Old cluster: October 2008
- Capacity 2.3 PB
- filled up, > 250 M files



Lustre @ GSI

- Lustre @ GSI = success story
- Old cluster: October 2008
- Capacity 2.3 PB
- filled up, > 250 M files
- 147 OSSs



Lustre @ GSI

- Lustre @ GSI = success story
- Old cluster: October 2008
- Capacity 2.3 PB
- filled up, > 250 M files
- 147 OSSs
- 440 OSTs



Lustre @ GSI

- Lustre @ GSI = success story
- Old cluster: October 2008
- Capacity 2.3 PB
- filled up, > 250 M files
- 147 OSSs
- 440 OSTs
- Ethernet 500 Gb/s



Lustre @ GSI

- Lustre @ GSI = success story
- Old cluster: October 2008
- Capacity 2.3 PB
- filled up, > 250 M files
- 147 OSSs
- 440 OSTs
- Ethernet 500 Gb/s
- softly phased out ...



Lustre @ GSI

- Lustre @ GSI = success story
- Old cluster: October 2008
- Capacity 2.3 PB
- filled up, > 250 M files
- 147 OSSs
- 440 OSTs
- Ethernet 500 Gb/s
- softly phased out ...
- Minicube cluster: March 2012



Lustre @ GSI

- Lustre @ GSI = success story
- Old cluster: October 2008
- Capacity 2.3 PB
- filled up, > 250 M files
- 147 OSSs
- 440 OSTs
- Ethernet 500 Gb/s
- softly phased out ...
- Minicube cluster: March 2012
- Capacity: 1.4 PB



Lustre @ GSI

- Lustre @ GSI = success story
- Old cluster: October 2008
- Capacity 2.3 PB
- filled up, > 250 M files
- 147 OSSs
- 440 OSTs
- Ethernet 500 Gb/s
- softly phased out ...
- Minicube cluster: March 2012
- Capacity: 1.4 PB

👉 1.1 PB used, 110 M files



Lustre @ GSI

- Lustre @ GSI = success story
- Old cluster: October 2008
- Capacity 2.3 PB
- filled up, > 250 M files
- 147 OSSs
- 440 OSTs
- Ethernet 500 Gb/s
- softly phased out ...
- Minicube cluster: March 2012
- Capacity: 1.4 PB
- 1.1 PB used, 110 M files
- 50 OSSs



Lustre @ GSI

- Lustre @ GSI = success story
- Old cluster: October 2008
- Capacity 2.3 PB
- filled up, > 250 M files
- 147 OSSs
- 440 OSTs
- Ethernet 500 Gb/s
- softly phased out ...
- Minicube cluster: March 2012
- Capacity: 1.4 PB
- 1.1 PB used, 110 M files
- 50 OSSs
- 200 OSTs



Lustre @ GSI

- Lustre @ GSI = success story
- Old cluster: October 2008
- Capacity 2.3 PB
- filled up, > 250 M files
- 147 OSSs
- 440 OSTs
- Ethernet 500 Gb/s
- softly phased out ...
- Minicube cluster: March 2012
- Capacity: 1.4 PB
- 1.1 PB used, 110 M files
- 50 OSSs
- 200 OSTs
- QDR Infiniband



Lustre @ GSI

- Lustre @ GSI = success story
- Old cluster: October 2008
- Capacity 2.3 PB
- filled up, > 250 M files
- 147 OSSs
- 440 OSTs
- Ethernet 500 Gb/s
- softly phased out ...
- Minicube cluster: March 2012
- Capacity: 1.4 PB
- 1.1 PB used, 110 M files
- 50 OSSs
- 200 OSTs
- QDR Infiniband
- Hera



Lustre @ GSI

- Lustre @ GSI = success story
- Old cluster: October 2008
- Capacity 2.3 PB
- filled up, > 250 M files
- 147 OSSs
- 440 OSTs
- Ethernet 500 Gb/s
- softly phased out ...
- Minicube cluster: March 2012
- Capacity: 1.4 PB
- 1.1 PB used, 110 M files
- 50 OSSs
- 200 OSTs
- QDR Infiniband
- Hera

Migrating Data + Hardware → Hera

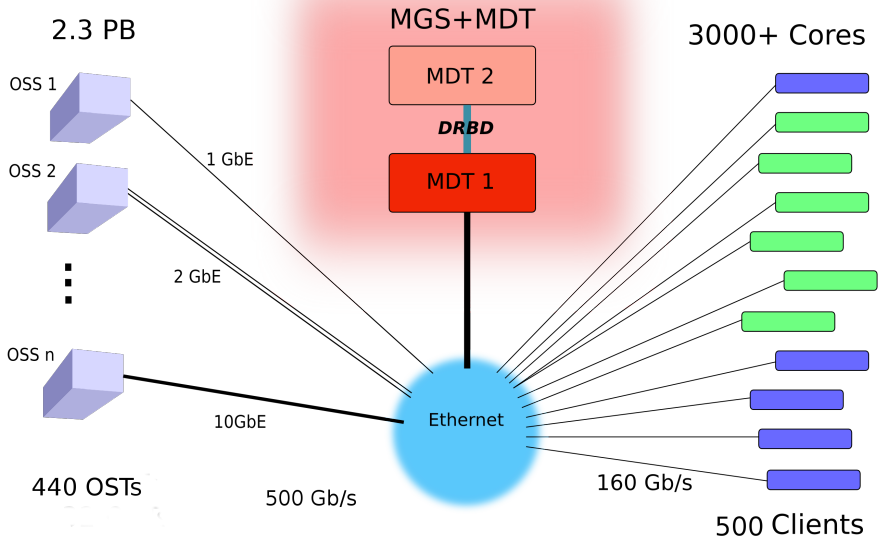


Outline

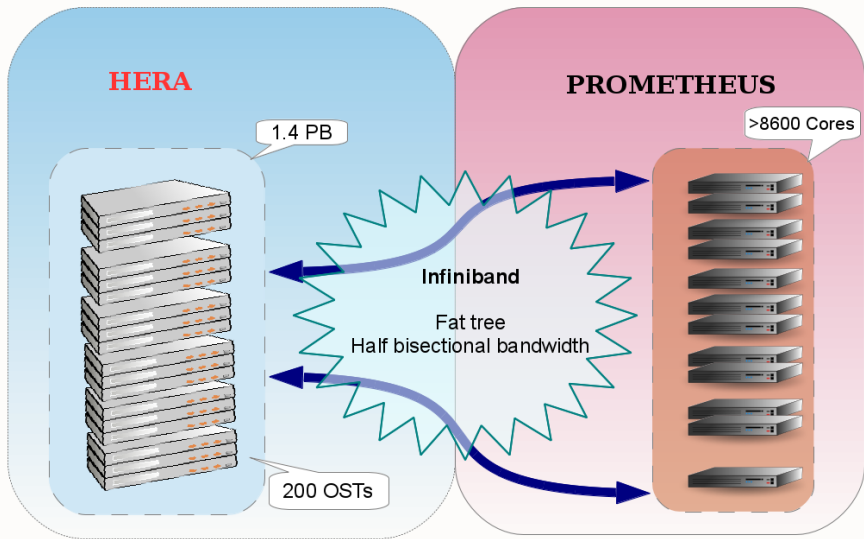
- GSI Lustre Setup
- Lustre Stability
- Migrating from Lustre to Lustre
- Growing beyond Minicube



Old cluster

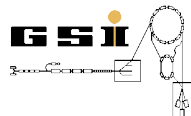


New cluster: Minicube



Minicube

- Infiniband fabric



Minicube

- Infiniband fabric
- administrative network (IPMI), no ethernet



Minicube

- Infiniband fabric
- administrative network (IPMI), no ethernet

I/O of minicube



Minicube

- Infiniband fabric
- administrative network (IPMI), no ethernet

I/O of minicube

😎 Lustre traffic: 4 LNET routers (10GE)



Minicube

- Infiniband fabric
- administrative network (IPMI), no ethernet

I/O of minicube

- 🕶️ Lustre traffic: 4 LNET routers (10GE)
 - ➡️ 40 Gb/s for clients outside



Minicube

- Infiniband fabric
- administrative network (IPMI), no ethernet

I/O of minicube

- 🕶️ Lustre traffic: 4 LNET routers (10GE)
 - ➡️ 40 Gb/s for clients outside
- *IP* traffic: Eth-IB gateway box with 1GE only 🕶️

Minicube

- Infiniband fabric
- administrative network (IPMI), no ethernet

I/O of minicube

- 🕶️ Lustre traffic: 4 LNET routers (10GE)
 - ➡️ 40 Gb/s for clients outside
- IP traffic: Eth-IB gateway box with 1GE only 🏆
 - 👉 Used also for PXE booting, DHCP, installation over Infiniband



Minicube

- Infiniband fabric
- administrative network (IPMI), no ethernet

I/O of minicube

- 🕶️ Lustre traffic: 4 LNET routers (10GE)
 - ➡️ 40 Gb/s for clients outside
- *IP* traffic: Eth-IB gateway box with 1GE only 🏠
 - 👉 Used also for PXE booting, DHCP, installation over Infiniband
- But: no NFS, no home directories on compute nodes



Minicube

- Infiniband fabric
- administrative network (IPMI), no ethernet

I/O of minicube

- 🕶️ Lustre traffic: 4 LNET routers (10GE)
 - ➡️ 40 Gb/s for clients outside
- IP traffic: Eth-IB gateway box with 1GE only 🏆
 - 👉 Used also for PXE booting, DHCP, installation over Infiniband
- But: no NFS, no home directories on compute nodes
- 🙄 Software installation/compilation on Lustre: bad choice

Minicube

- Infiniband fabric
- administrative network (IPMI), no ethernet

I/O of minicube

- 🧐 Lustre traffic: 4 LNET routers (10GE)
 - ➡ 40 Gb/s for clients outside
- IP traffic: Eth-IB gateway box with 1GE only 🏆
 - 👉 Used also for PXE booting, DHCP, installation over Infiniband
- But: no NFS, no home directories on compute nodes
- 🧐 Software installation/compilation on Lustre: bad choice
- 🧐 Put software on CVMFS !

Minicube

- Infiniband fabric
- administrative network (IPMI), no ethernet

I/O of minicube

- 🕶️ Lustre traffic: 4 LNET routers (10GE)
 - ➡️ 40 Gb/s for clients outside
- IP traffic: Eth-IB gateway box with 1GE only 🏆
 - 👉 Used also for PXE booting, DHCP, installation over Infiniband
- But: no NFS, no home directories on compute nodes
- 🙄 Software installation/compilation on Lustre: bad choice
- 🕶️ Put software on CVMFS !
 - Great success



Minicube

- Infiniband fabric
- administrative network (IPMI), no ethernet

I/O of minicube

- 🕶️ Lustre traffic: 4 LNET routers (10GE)
 - ➡️ 40 Gb/s for clients outside
- IP traffic: Eth-IB gateway box with 1GE only 🏆
 - 👉 Used also for PXE booting, DHCP, installation over Infiniband
- But: no NFS, no home directories on compute nodes
- 🙄 Software installation/compilation on Lustre: bad choice
- 🕶️ Put software on CVMFS !
 - Great success
 - Now testing proxy on LNET router



Lustre Software

- Old Lustre: 2.6.27 Lustre 1.8.4



Lustre Software

- Old Lustre: 2.6.27 Lustre 1.8.4
- **Hera**: 2.6.32 Lustre 1.8.7-wc, no additional patches



Lustre Software

- Old Lustre: 2.6.27 Lustre 1.8.4
- **Hera**: 2.6.32 Lustre 1.8.7-wc, no additional patches
- Soon: test system running Lustre 2.3



Lustre Software

- Old Lustre: 2.6.27 Lustre 1.8.4
- **Hera**: 2.6.32 Lustre 1.8.7-wc, no additional patches
- Soon: test system running Lustre 2.3
- **Hera**: Robinhood 2.3.3



Lustre Software

- Old Lustre: 2.6.27 Lustre 1.8.4
- **Hera**: 2.6.32 Lustre 1.8.7-wc, no additional patches
- Soon: test system running Lustre 2.3
- **Hera**: Robinhood 2.3.3
 - **Hera** Scan: 24 hours



Lustre Software

- Old Lustre: 2.6.27 Lustre 1.8.4
- **Hera**: 2.6.32 Lustre 1.8.7-wc, no additional patches
- Soon: test system running Lustre 2.3
- **Hera**: Robinhood 2.3.3
 - **Hera** Scan: 24 hours
 - 🤖 Very helpful: Tell users where their Terabytes are



Lustre Software

- Old Lustre: 2.6.27 Lustre 1.8.4
- **Hera**: 2.6.32 Lustre 1.8.7-wc, no additional patches
- Soon: test system running Lustre 2.3
- **Hera**: Robinhood 2.3.3
 - **Hera** Scan: 24 hours
 - 🤔 Very helpful: Tell users where their Terabytes are
 - 🤖 Dangerous to MDS!



Lustre Stability



Lustre Stability

Lustre Crashes

- ▶ Both old system and new installation very stable

Lustre Stability

Lustre Crashes

- ▶ Both old system and new installation very stable
- 🤔 One LBUG recurring

Lustre Stability

Lustre Crashes

- ▶ Both old system and new installation very stable

🤖 One LBUG recurring

```
(ldlm_lock.c:165:ldlm_lock_put())  
ASSERTION(atomic_read(&(lock->l_export)->exp_refcount)  
< 0x5a5a5a) failed
```

Lustre Stability

Lustre Crashes

- ▶ Both old system and new installation very stable

🤖 One LBUG recurring

```
(ldlm_lock.c:165:ldlm_lock_put())  
ASSERTION(atomic_read(&(lock->l_export)->exp_refcount)  
< 0x5a5a5a) failed
```

- ➔ LU-919 (<http://jira.whamcloud.com/browse/LU-919>),
unresolved in 1.8

Lustre Stability

Lustre Crashes

- ▶ Both old system and new installation very stable

😓 One LBUG recurring

```
(ldlm_lock.c:165:ldlm_lock_put())  
ASSERTION(atomic_read(&(lock->l_export)->exp_refcount)  
< 0x5a5a5a) failed
```

- ➔ LU-919 (<http://jira.whamcloud.com/browse/LU-919>), unresolved in 1.8
- 😓 Trigger: high load by users + Robinhood running

Lustre Stability

Lustre Crashes

- ▶ Both old system and new installation very stable

🤖 One LBUG recurring

```
(ldlm_lock.c:165:ldlm_lock_put())  
ASSERTION(atomic_read(&(lock->l_export)->exp_refcount)  
< 0x5a5a5a) failed
```

- ➔ LU-919 (<http://jira.whamcloud.com/browse/LU-919>), unresolved in 1.8

🤖 Trigger: high load by users + Robinhood running

- **Hera**: switch to DRDB-partner (manual, no HA)

Lustre Stability

Lustre Crashes

- ▶ Both old system and new installation very stable

😓 One LBUG recurring

```
(ldlm_lock.c:165:ldlm_lock_put())  
ASSERTION(atomic_read(&(lock->l_export)->exp_refcount)  
< 0x5a5a5a) failed
```

- ➔ LU-919 (<http://jira.whamcloud.com/browse/LU-919>), unresolved in 1.8

😓 Trigger: high load by users + Robinhood running

- **Hera**: switch to DRDB-partner (manual, no HA)
- **Hera**: ext4-based ldiskfs: fsck times down to < 20 min

Hardware/Data migration

Data Migration

- Old Lustre → Hera



Hardware/Data migration

Data Migration

- Old Lustre → **Hera**
- done by the users



Hardware/Data migration

Data Migration

- Old Lustre → **Hera**
- done by the users filesystem summary: 1.4P 1.1P 275.8T 80%
/hera



Hardware/Data migration

Data Migration

- Old Lustre → **Hera**
- done by the users filesystem summary: 1.4P 1.1P 275.8T 80%
/hera
- smaller groups still working on old cluster

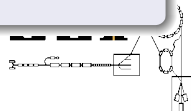


Hardware/Data migration

Data Migration

- Old Lustre → **Hera**
- done by the users filesystem summary: 1.4P 1.1P 275.8T 80%
/hera
- smaller groups still working on old cluster

Hardware Migration



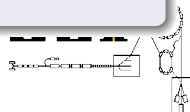
Hardware/Data migration

Data Migration

- Old Lustre → **Hera**
- done by the users filesystem summary: 1.4P 1.1P 275.8T 80%
/hera
- smaller groups still working on old cluster

Hardware Migration

- Container with 10 racks



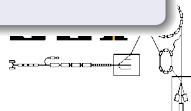
Hardware/Data migration

Data Migration

- Old Lustre → **Hera**
- done by the users filesystem summary: 1.4P 1.1P 275.8T 80%
/hera
- smaller groups still working on old cluster

Hardware Migration

- Container with 10 racks **Bouillon Cube**



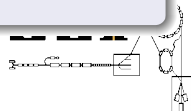
Hardware/Data migration

Data Migration

- Old Lustre → **Hera**
- done by the users filesystem summary: 1.4P 1.1P 275.8T 80%
/hera
- smaller groups still working on old cluster

Hardware Migration

- Container with 10 racks **Bouillon Cube**
- 44 OSS, 1.3 PB, less than 3 years old



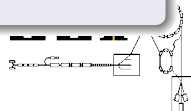
Hardware/Data migration

Data Migration

- Old Lustre → **Hera**
- done by the users filesystem summary: 1.4P 1.1P 275.8T 80%
/hera
- smaller groups still working on old cluster

Hardware Migration

- Container with 10 racks **Bouillon Cube**
- 44 OSS, 1.3 PB, less than 3 years old
- ▶ Move them to Minicube



Hardware/Data migration

Hardware Migration: Move OSS to Minicube



Hardware/Data migration

Hardware Migration: Move OSS to Minicube

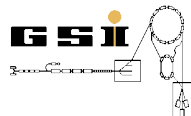
→ Exceedingly tedious task:



Hardware/Data migration

Hardware Migration: Move OSS to Minicube

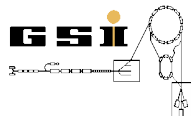
- Exceedingly tedious task:
 - 1 Set OST015c read-only on MDT



Hardware/Data migration

Hardware Migration: Move OSS to Minicube

- Exceedingly tedious task:
- 1 Set OST015c read-only on MDT
 - 2 Find files on OST015c

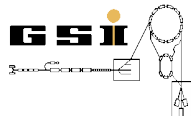


Hardware/Data migration

Hardware Migration: Move OSS to Minicube

→ Exceedingly tedious task:

- 1 Set OST015c read-only on MDT
- 2 Find files on OST015c 🤖 Scan of entire Lustre fs



Hardware/Data migration

Hardware Migration: Move OSS to Minicube

→ Exceedingly tedious task:

- 1 Set OST015c read-only on MDT
- 2 Find files on OST015c 🤖 Scan of entire Lustre fs
 - 👉 Do it for many directories in parallel



Hardware/Data migration

Hardware Migration: Move OSS to Minicube

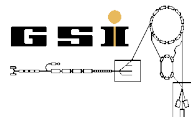
→ Exceedingly tedious task:

① Set OST015c read-only on MDT

② Find files on OST015c 🤖 Scan of entire Lustre fs

👉 Do it for many directories in parallel

⚡ Don't kill MDT!



Hardware/Data migration

Hardware Migration: Move OSS to Minicube

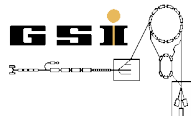
→ Exceedingly tedious task:

① Set OST015c read-only on MDT

② Find files on OST015c 🤖 Scan of entire Lustre fs

👉 Do it for many directories in parallel ⚡ Don't kill MDT!

③ Migrate OST015c-files to other OSTs



Hardware/Data migration

Hardware Migration: Move OSS to Minicube

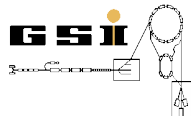
→ Exceedingly tedious task:

① Set OST015c read-only on MDT

② Find files on OST015c 🤖 Scan of entire Lustre fs

👉 Do it for many directories in parallel ⚡ Don't kill MDT!

③ Migrate OST015c-files to other OSTs ⚡ Don't kill that OSS!



Hardware/Data migration

Hardware Migration: Move OSS to Minicube

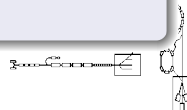
→ Exceedingly tedious task:

① Set OST015c read-only on MDT

② Find files on OST015c 🤖 Scan of entire Lustre fs

👉 Do it for many directories in parallel ⚡ Don't kill MDT!

③ Migrate OST015c-files to other OSTs ⚡ Don't kill that OSS!



Hardware/Data migration

Hardware Migration: Move OSS to Minicube

→ Exceedingly tedious task:

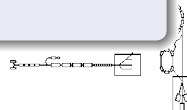
① Set OST015c read-only on MDT

② Find files on OST015c 🤖 Scan of entire Lustre fs

👉 Do it for many directories in parallel ⚡ Don't kill MDT!

③ Migrate OST015c-files to other OSTs ⚡ Don't kill that OSS!

• Procedure needs enough space on remaining OSTs



Hardware/Data migration

Hardware Migration: Move OSS to Minicube

→ Exceedingly tedious task:

① Set OST015c read-only on MDT

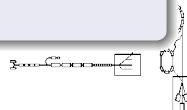
② Find files on OST015c 🤖 Scan of entire Lustre fs

👉 Do it for many directories in parallel ⚡ Don't kill MDT!

③ Migrate OST015c-files to other OSTs ⚡ Don't kill that OSS!

• Procedure needs enough space on remaining OSTs

➡ Depends on users: move their important data to **Hera**, delete *old* stuff



Hardware/Data migration

Hardware Migration: Move OSS to Minicube

→ Exceedingly tedious task:

① Set OST015c read-only on MDT

② Find files on OST015c 🤖 Scan of entire Lustre fs

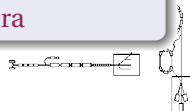
👉 Do it for many directories in parallel ⚡ Don't kill MDT!

③ Migrate OST015c-files to other OSTs ⚡ Don't kill that OSS!

- Procedure needs enough space on remaining OSTs

- ➡ Depends on users: move their important data to **Hera**, delete *old* stuff

- Half of container done, 0.5 PB to be added to **Hera**



Hardware/Data migration

Hardware Migration: Move OSS to Minicube

→ Exceedingly tedious task:

① Set OST015c read-only on MDT

② Find files on OST015c 🤖 Scan of entire Lustre fs

👉 Do it for many directories in parallel ⚡ Don't kill MDT!

③ Migrate OST015c-files to other OSTs ⚡ Don't kill that OSS!

- Procedure needs enough space on remaining OSTs

- ➡ Depends on users: move their important data to **Hera**, delete *old* stuff

- Half of container done, 0.5 PB to be added to **Hera**



Hardware/Data migration

Hardware Migration: Move OSS to Minicube

→ Exceedingly tedious task:

① Set OST015c read-only on MDT

② Find files on OST015c 🤖 Scan of entire Lustre fs

👉 Do it for many directories in parallel ⚡ Don't kill MDT!

③ Migrate OST015c-files to other OSTs ⚡ Don't kill that OSS!

- Procedure needs enough space on remaining OSTs

- ➡ Depends on users: move their important data to **Hera**, delete *old* stuff

- Half of container done, 0.5 PB to be added to **Hera**

- Old Lustre as of 15. Oktober 2012: Size down to 1.6 PB, 1 PB used

Growing beyond Minicube



Growing beyond Minicube

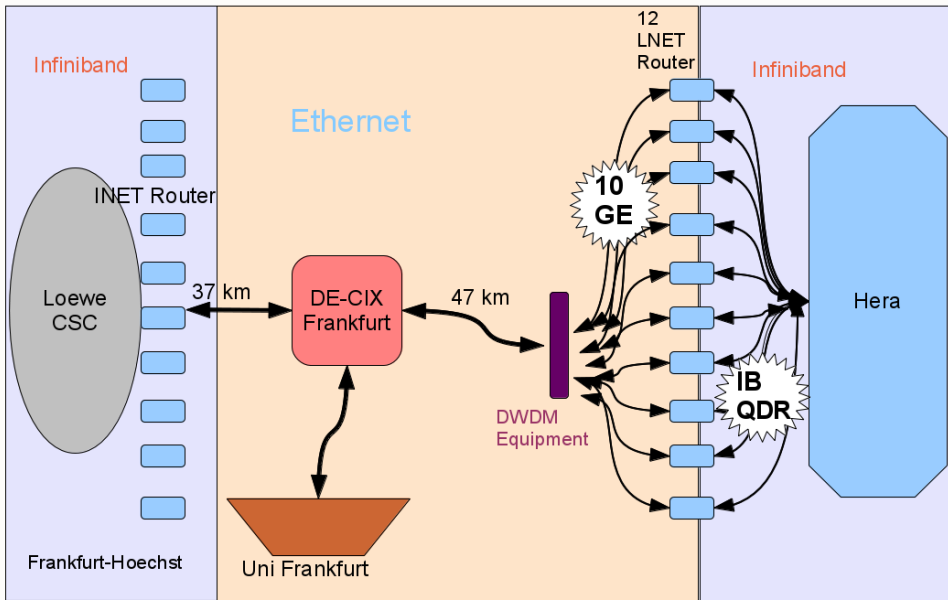
TeraLink Project

- Connect institutes in Rhein-Main-area to storage @ GSI



TeraLink: Connect outside institutes to storage @ GSI

First step: Connecting Loewe with Hera



TeraLink

- First step: Loewe CSC to **Hera**



TeraLink

- First step: Loewe CSC to **Hera**
- 12 batch machines, stuffed w 10GE cards, as LNET router



TeraLink

- First step: Loewe CSC to **Hera**
- 12 batch machines, stuffed w 10GE cards, as LNET router
- Second route added to Hera servers' *modprobe.d/lustre.conf*



TeraLink

- First step: Loewe CSC to **Hera**
- 12 batch machines, stuffed w 10GE cards, as LNET router
- Second route added to Hera servers' *modprobe.d/lustre.conf*

```
lnet routes='tcp1 10.16.1.[199-202]@o2ib0; tcp2 10.16.1.[220-231]@o2ib0'
```



TeraLink

- First step: Loewe CSC to **Hera**
- 12 batch machines, stuffed w 10GE cards, as LNET router
- Second route added to Hera servers' *modprobe.d/lustre.conf*

```
lnet routes='tcp1 10.16.1.[199-202]@o2ib0; tcp2 10.16.1.[220-231]@o2ib0"
```

- Add route to GSI-LNET-Routers onto Loewe clients



TeraLink

- First step: Loewe CSC to **Hera**
- 12 batch machines, stuffed w 10GE cards, as LNET router
- Second route added to Hera servers' *modprobe.d/lustre.conf*

```
lnet routes="'tcp1 10.16.1.[199-202]@o2ib0; tcp2 10.16.1.[220-231]@o2ib0"
```

- Add route to GSI-LNET-Routers onto Loewe clients

```
lnet networks=tcp2(ib0)
```

```
lnet routes="'o2ib 172.16.0.[1,5,9,13,17,21,25,29,33,37,41,45]@tcp2"
```



TeraLink

- First step: Loewe CSC to **Hera**
- 12 batch machines, stuffed w 10GE cards, as LNET router
- Second route added to Hera servers' *modprobe.d/lustre.conf*

```
lnet routes="'tcp1 10.16.1.[199-202]@o2ib0; tcp2 10.16.1.[220-231]@o2ib0"
```

- Add route to GSI-LNET-Routers onto Loewe clients

```
lnet networks=tcp2(ib0)
```

```
lnet routes="'o2ib 172.16.0.[1,5,9,13,17,21,25,29,33,37,41,45]@tcp2"
```

lozone from Loewe

TeraLink

- First step: Loewe CSC to **Hera**
- 12 batch machines, stuffed w 10GE cards, as LNET router
- Second route added to Hera servers' *modprobe.d/lustre.conf*

```
lnet routes="'tcp1 10.16.1.[199-202]@o2ib0; tcp2 10.16.1.[220-231]@o2ib0"
```

- Add route to GSI-LNET-Routers onto Loewe clients

```
lnet networks=tcp2(ib0)
```

```
lnet routes="'o2ib 172.16.0.[1,5,9,13,17,21,25,29,33,37,41,45]@tcp2"
```

lozone from Loewe

- 1 client: 730 MB/s

TeraLink

- First step: Loewe CSC to **Hera**
- 12 batch machines, stuffed w 10GE cards, as LNET router
- Second route added to Hera servers' *modprobe.d/lustre.conf*

```
lnet routes="'tcp1 10.16.1.[199-202]@o2ib0; tcp2 10.16.1.[220-231]@o2ib0"
```

- Add route to GSI-LNET-Routers onto Loewe clients

```
lnet networks=tcp2(ib0)
```

```
lnet routes="'o2ib 172.16.0.[1,5,9,13,17,21,25,29,33,37,41,45]@tcp2"
```

lozone from Loewe

- 1 client: 730 MB/s
- 12 clients: 4.3 GB/s

TeraLink

- First step: Loewe CSC to **Hera**
- 12 batch machines, stuffed w 10GE cards, as LNET router
- Second route added to Hera servers' *modprobe.d/lustre.conf*

```
lnet routes="'tcp1 10.16.1.[199-202]@o2ib0; tcp2 10.16.1.[220-231]@o2ib0"
```

- Add route to GSI-LNET-Routers onto Loewe clients

```
lnet networks=tcp2(ib0)
```

```
lnet routes="'o2ib 172.16.0.[1,5,9,13,17,21,25,29,33,37,41,45]@tcp2"
```

lozone from Loewe

- 1 client: 730 MB/s
- 12 clients: 4.3 GB/s

➡ More investigation necessary

Future plans

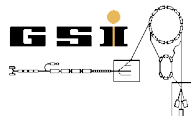
TeraLink Project



Future plans

TeraLink Project

- Expand Lustre link to other institutes



Future plans

TeraLink Project

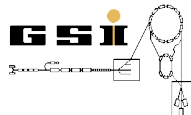
- Expand Lustre link to other institutes
- Mapping of *alien* UIDs and GIDs into GSI user space



Future plans

TeraLink Project

- Expand Lustre link to other institutes
- Mapping of *alien* UIDs and GIDs into GSI user space
- Progress on Lustre 2.3x with Kerberos @ GSI



Future plans

TeraLink Project

- Expand Lustre link to other institutes
- Mapping of *alien* UIDs and GIDs into GSI user space
- Progress on Lustre 2.3x with Kerberos @ GSI
 - MDS ↔ OSS works, clients still uncooperative



Future plans

TeraLink Project

- Expand Lustre link to other institutes
- Mapping of *alien* UIDs and GIDs into GSI user space
- Progress on Lustre 2.3x with Kerberos @ GSI
 - MDS ↔ OSS works, clients still uncooperative
 - Testing various Kerberos modi on performance impact



Future plans

TeraLink Project

- Expand Lustre link to other institutes
- Mapping of *alien* UIDs and GIDs into GSI user space
- Progress on Lustre 2.3x with Kerberos @ GSI
 - MDS ↔ OSS works, clients still uncooperative
 - Testing various Kerberos modi on performance impact

More space, more, more



Future plans

TeraLink Project

- Expand Lustre link to other institutes
- Mapping of *alien* UIDs and GIDs into GSI user space
- Progress on Lustre 2.3x with Kerberos @ GSI
 - MDS ↔ OSS works, clients still uncooperative
 - Testing various Kerberos modi on performance impact

More space, more, more

- **Hera** already full



Future plans

TeraLink Project

- Expand Lustre link to other institutes
- Mapping of *alien* UIDs and GIDs into GSI user space
- Progress on Lustre 2.3x with Kerberos @ GSI
 - MDS ↔ OSS works, clients still uncooperative
 - Testing various Kerberos modi on performance impact

More space, more, more

- **Hera** already full
- + 1.3 PB from container servers



Future plans

TeraLink Project

- Expand Lustre link to other institutes
- Mapping of *alien* UIDs and GIDs into GSI user space
- Progress on Lustre 2.3x with Kerberos @ GSI
 - MDS ↔ OSS works, clients still uncooperative
 - Testing various Kerberos modi on performance impact

More space, more, more

- **Hera** already full
- + 1.3 PB from container servers
- + 3 PB new hardware



Lustre @ GSI



Questions ?

