

# Operations and plans – German sites

## ALICE T1T2 Meeting 2019 Bucharest

Max Fischer, Sören Fleischer, Raffaele Grosso, Jan Knedlik,  
Thorsten Kollegger, Paul Kramp, Kilian Schwarz

14.05.2019

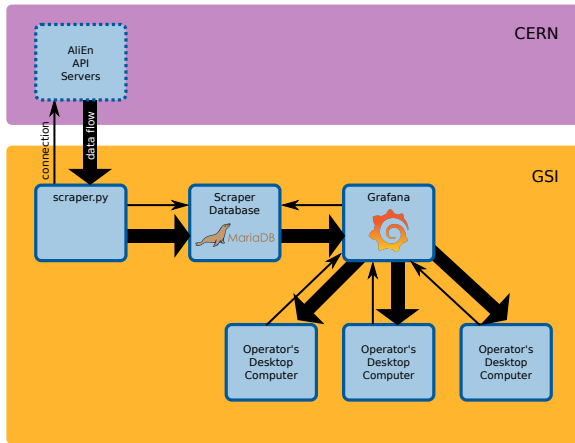
## AliEn Scraper – Motivation

- Problem: ALICE T2 job performance metrics are not obvious for site operators. MonaLisa offers limited information. The raw data for more sophisticated metrics, however, is present in AliEn.
- Solution: A python script scraping the AliEn database and writing the relevant data into a local database
  - Low-level aggregation already taking place in the MariaDB database for faster read access
  - Plotting of the time series data is done further downstream by a Grafana instance

## AliEn Scraper – Motivation

- Problem: ALICE T2 job performance metrics are not obvious for site operators. MonaLisa offers limited information. The raw data for more sophisticated metrics, however, is present in AliEn.
- Solution: A python script scraping the AliEn database and writing the relevant data into a local database
  - Low-level aggregation already taking place in the MariaDB database for faster read access
  - Plotting of the time series data is done further downstream by a Grafana instance

# AliEn Scraper – Block Diagram



## AliEn Scraper – Job Categorization (1)

Low-level metrics for each JobID:

- $ds := \#$  of times a job goes into a **good** final state at **our site**
- $fs := \#$  of times a job goes into an **error** final state at **our site**
- $do := \#$  of times a job goes into a **good** final state **elsewhere**
- $fo := \#$  of times a job goes into an **error** final state **elsewhere**

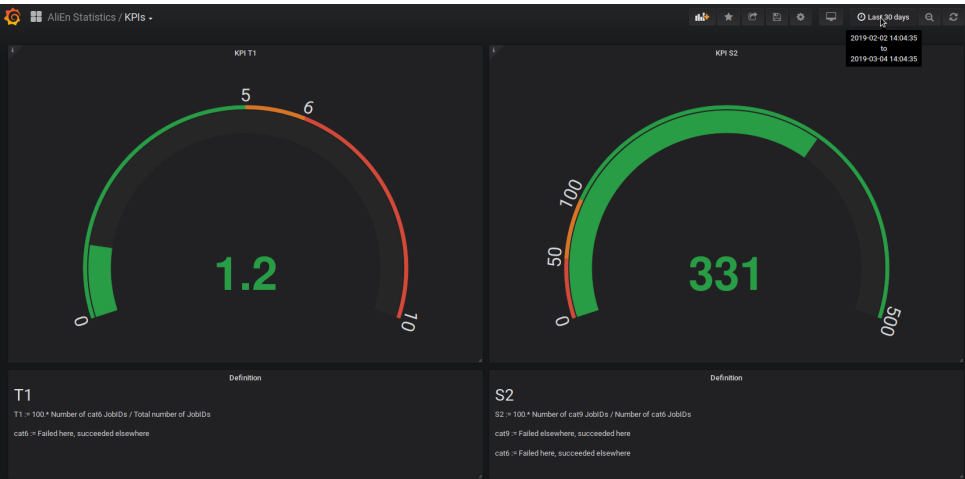
## AliEn Scraper – Job Categorization (2)

These low-level metrics can be used to derive high-level metrics  
**(KPIs)**

(0: 0 times, 1:  $\geq 1$  times)

cat	ds	fs	do	fo	description
0	0	0	0	0	Didn't finish (yet)
1	0	0	0	1	Failed elsewhere
2	0	0	1	0	Succeeded elsewhere
3	0	0	1	1	Failed and succeeded elsewhere
4	0	1	0	0	Failed here
5	0	1	0	1	Failed generally
6	0	1	1	0	Failed here, succeeded elsewhere
7	0	1	1	1	Failed generally, succeeded elsewhere
8	1	0	0	0	Succeeded here
9	1	0	0	1	Failed elsewhere, succeeded here
10	1	0	1	0	Succeeded here and elsewhere
11	1	0	1	1	Failed elsewhere, succeeded generally
12	1	1	0	0	Failed and succeeded here
13	1	1	0	1	Failed generally, succeeded here
14	1	1	1	0	Failed here, succeeded generally
15	1	1	1	1	Failed and succeeded generally

# Grafana – AliEn KPIs

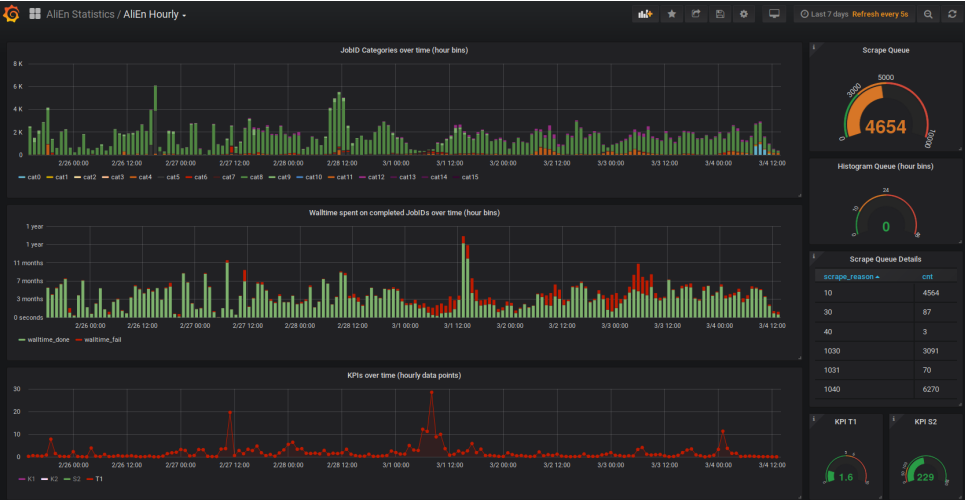


# Grafana – AliEn Daily

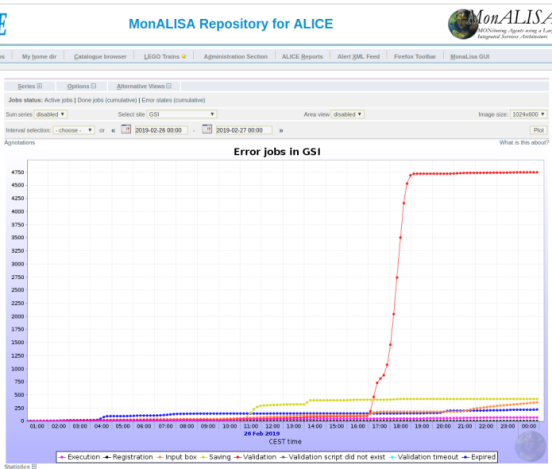




# Grafana – AliEn Hourly



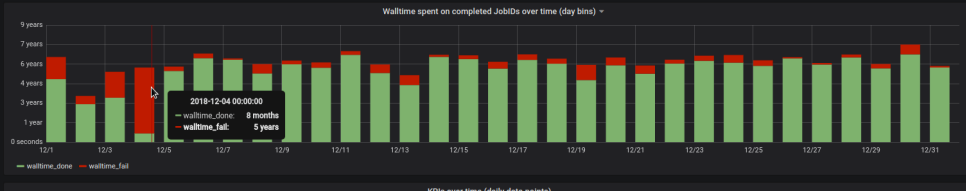
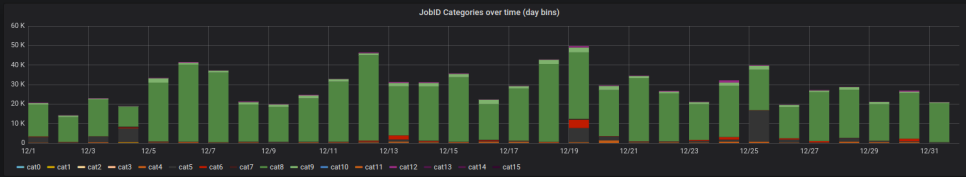
# Scraper Example (1)



# Scraper Example (2)

AliEn Statistics / AliEn Daily -

★ 🔄 🗨️ ⏪ ⌚ Dec 1, 2018



# Scraper git Repository

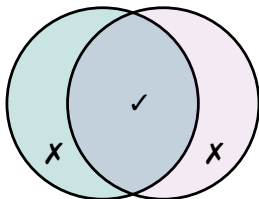
<https://git.gsi.de/dc/common/AliEn-scraper>

## Previous situation / Motivation

- 2 xrootd proxy servers, sharing a hostname with 2 A records
  - High CPU+Memory usage
  - Suspected cause of failing jobs due to failing TCP connections
- Failure behaviour: 50% failed requests per failing server
- CERN had to provide a rewrite mechanism for GSI-outbound requests
  - Some GSI-internal requests were incorrectly rewritten as well
  - Some GSI-outbound requests incorrectly lacked rewriting

## Old rewrite / xrootd proxy mechanism

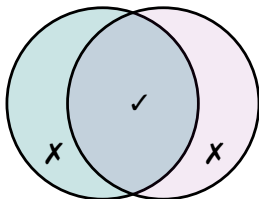
Requests that were supposed  
to get rewritten



Requests that were  
actually rewritten

# Old rewrite / xrootd proxy mechanism

Requests that were supposed to get rewritten

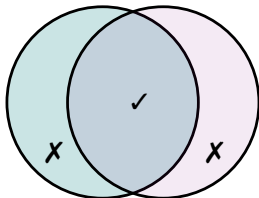


```
..... xrootd-redirector-stats Begin .....  
This redirector has received requests from the following hosts:  
137.138.99.165 73124 ##### pcalienstorage2.cern.ch  
140.181.2.26 39634 ##### lxaliproxy3.gsi.de  
140.181.2.25 39188 ##### lxaliproxy2.gsi.de  
137.138.97.208 8624 # pcalienstorage2.cern.ch  
140.181.2.27 7816 ## lxalir1.gsi.de  
140.181.2.11 4212 # lxalise1.gsi.de  
140.181.2.24 4206 # lxalise4.gsi.de  
140.181.2.23 4206 # lxalise2.gsi.de  
202.13.220.240 1657  
192.168.46.248 1618 fw-nat-inside-outside.gridka.de  
137.138.99.186 1573 pcalienstorage2.cern.ch  
127.0.0.1 1402 localhost  
10.20.2.177 998 lxbk0390.gsi.de  
10.20.2.170 992 lxbk0383.gsi.de  
10.20.3.128 979 lxbk0534.gsi.de  
10.20.3.144 965 lxbk0550.gsi.de
```

Requests that were actually rewritten

# Old rewrite / xrootd proxy mechanism

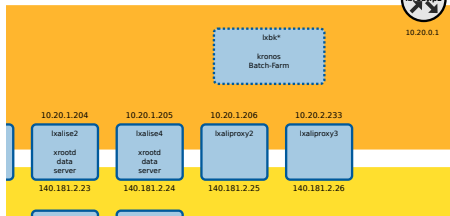
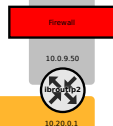
Requests that were supposed to get rewritten



```

..... xrootd-redirector-stats Begin .....
This redirector has received requests from the following hosts:
137.138.99.186  23124 ###### pcalienstorage2.cern.ch
140.181.2.26   39634 ###### lxaliproxy3.gsi.de
140.181.2.25   39188 ###### lxaliproxy2.gsi.de
137.138.97.208 8624 ###### pcalienstorage2.cern.ch
140.181.2.27   7816 ###### lxalise1.gsi.de
140.181.2.11   4212 ###### lxalise1.gsi.de
140.181.2.24   4206 ###### lxalise4.gsi.de
140.181.2.23   4206 ###### lxalise2.gsi.de
282.13.220.240 1657
192.168.46.248 1618 fw-nat-inside-outside.gridka.de
137.138.99.186 1573 pcalienstorage2.cern.ch
127.0.0.1      1402 localnust
10.20.2.177    998 lxbk0390.gsi.de
10.20.2.170    992 lxbk0383.gsi.de
10.20.3.128    979 lxbk0534.gsi.de
10.20.3.144    965 lxbk0550.gsi.de
    
```

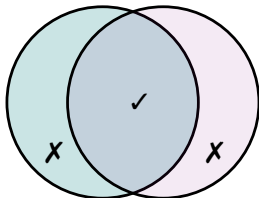
Requests that were actually rewritten





# Old rewrite / xrootd proxy mechanism

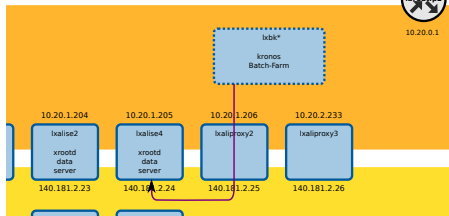
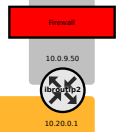
Requests that were supposed to get rewritten



```

..... xrootd-redirector-stats Begin .....
This redirector has received requests from the following hosts:
137.138.99.186  73124 ###### pcalienstorage2.cern.ch
140.181.2.26   39634 ###### lxaliproxy3.gsi.de
140.181.2.25   39188 ###### lxaliproxy2.gsi.de
137.138.97.208 8624 ###### pcalienstorage2.cern.ch
140.181.2.27   7816 ###### lxalir01.gsi.de
140.181.2.11   4212 ###### lxalise1.gsi.de
140.181.2.24   4206 ###### lxalise4.gsi.de
140.181.2.23   4206 ###### lxalise2.gsi.de
282.13.220.240 1657
192.168.46.248 1618 fw-nat-inside-outside.gridka.de
137.138.99.186 1573 pcalienstorage2.cern.ch
127.0.0.1      1402 localnust
10.20.2.177   998 lxbk0390.gsi.de
10.20.2.170   992 lxbk0383.gsi.de
10.20.3.128   979 lxbk0534.gsi.de
10.20.3.144   965 lxbk0550.gsi.de
    
```

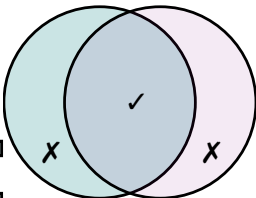
Requests that were actually rewritten



# Old rewrite / xrootd proxy mechanism

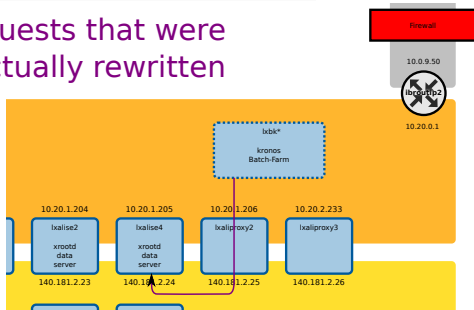
Requests that were supposed to get rewritten

```
{
'10.20.1.221:1094:TCP': 1483,
'137.138.99.175:10000:TCP': 8,
'140.181.2.27:1094:TCP': 148,
'10.20.1.204:1094:TCP': 44476,
'140.181.2.29:1094:TCP': 161,
'140.181.60.25:8084:TCP': 3930,
'137.138.47.221:8080:TCP': 2862,
'137.138.47.221:8083:TCP': 1991,
'137.138.47.216:8080:TCP': 2528,
'137.138.99.180:10000:TCP': 135,
'10.20.1.205:1094:TCP': 25498,
'10.20.1.206:1094:TCP': 11930,
'193.205.66.200:1094:TCP': 40,
'10.20.2.233:1094:TCP': 23851,
'137.138.99.130:8080:TCP': 683,
'137.138.99.140:8080:TCP': 966,
'137.138.99.182:10000:TCP': 105,
'188.184.38.72:1094:TCP': 40,
'137.138.99.167:10000:TCP': 209,
'137.138.99.136:8080:TCP': 179,
'137.138.99.169:10000:TCP': 92,
'137.138.99.170:10000:TCP': 96,
'137.138.99.188:10000:TCP': 64,
'137.138.99.138:10000:TCP': 64,
'137.138.99.145:8080:TCP': 239,
'137.138.99.179:10000:TCP': 32,
'137.138.99.168:10000:TCP': 32
}
```



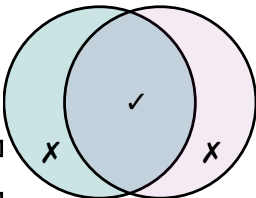
```
----- xrootd-redirector-stats Begin -----
This redirector has received requests from the following hosts:
137.138.99.165 73124 ##### pcaliendbg.cern.ch
140.181.2.26 39634 ##### lxaliproxy3.gsi.de
140.181.2.25 39188 ##### lxaliproxy2.gsi.de
137.138.47.208 8624 ## pcaliendbmonitor.cern.ch
140.181.2.27 7816 ## lxalir01.gsi.de
140.181.2.11 4212 # lxalise1.gsi.de
140.181.2.24 4206 # lxalise4.gsi.de
140.181.2.23 4206 # lxalise2.gsi.de
282.13.230.240 1657
192.168.46.248 1618
137.138.99.186 1573
127.0.0.1 1402
10.20.2.177 998
10.20.2.170 992
10.20.3.128 979
10.20.3.144 965
fw-nat-inside-outside.gridka.de
pcaliensstorage2.cern.ch
localnust
lxbk0390.gsi.de
lxbk0383.gsi.de
lxbk0534.gsi.de
lxbk0550.gsi.de
```

Requests that were actually rewritten



# Old rewrite / xrootd proxy mechanism

Requests that were supposed to get rewritten



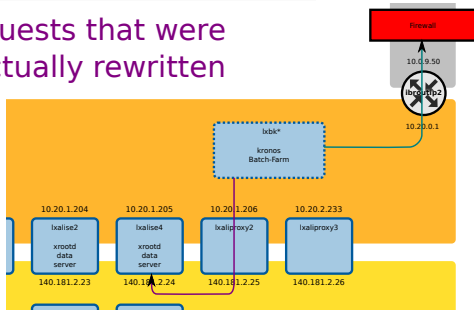
```

'10.20.1.221:1094:TCP': 1483,
'137.138.99.175:10000:TCP': 8,
'140.181.2.27:1094:TCP': 148,
'10.20.1.204:1094:TCP': 44476,
'140.181.2.29:1094:TCP': 161,
'140.181.60.25:8084:TCP': 3930,
'137.138.47.221:8080:TCP': 2862,
'137.138.47.221:8083:TCP': 1991,
'137.138.47.216:8080:TCP': 2528,
'137.138.99.180:10000:TCP': 135,
'10.20.1.205:1094:TCP': 25498,
'10.20.1.206:1094:TCP': 11930,
'193.205.66.200:1094:TCP': 40,
'10.20.2.233:1094:TCP': 23851,
'137.138.99.130:8050:TCP': 683,
'137.138.99.140:8080:TCP': 966,
'137.138.99.182:10000:TCP': 105,
'188.184.38.72:1094:TCP': 40,
'137.138.99.167:10000:TCP': 209,
'137.138.99.136:8080:TCP': 179,
'137.138.99.169:10000:TCP': 92,
'137.138.99.170:10000:TCP': 96,
'137.138.99.188:10000:TCP': 64,
'137.138.99.138:10000:TCP': 64,
'137.138.99.145:8080:TCP': 239,
'137.138.99.179:10000:TCP': 32,
'137.138.99.168:10000:TCP': 32
    
```

```

----- xrootd-redirector-stats Begin -----
This redirector has received requests from the following hosts:
137.138.99.165 73124 ##### pcalienstorage2.cern.ch
140.181.2.26 39634 ##### lxaliproxy3.gsi.de
140.181.2.25 39188 ##### lxaliproxy2.gsi.de
137.138.47.208 8624 ## pcalienmonitor.cern.ch
140.181.2.27 7816 ## lxalir01.gsi.de
140.181.2.11 4212 # lxalise1.gsi.de
140.181.2.24 4206 # lxalise4.gsi.de
140.181.2.23 4206 # lxalise2.gsi.de
282.13.230.240 1657
192.168.46.248 1618
137.138.99.186 1573
127.0.0.1 1402
10.20.2.177 998
10.20.2.170 992
10.20.3.128 979
10.20.3.144 965
    
```

Requests that were actually rewritten

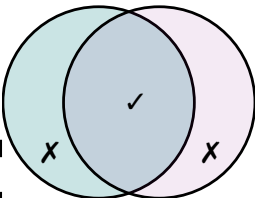


# Old rewrite / xrootd proxy mechanism

Requests that were supposed to get rewritten

```

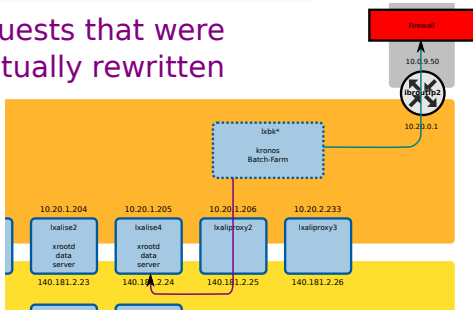
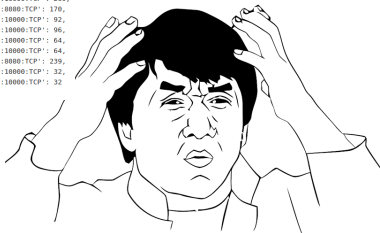
'10.20.1.221:1094:TCP': 1483,
'137.138.99.175:10000:TCP': 8,
'140.181.2.27:1094:TCP': 148,
'10.20.1.204:1094:TCP': 44476,
'140.181.2.29:1094:TCP': 161,
'140.181.60.25:8084:TCP': 3930,
'137.138.47.221:8080:TCP': 2962,
'137.138.47.221:8083:TCP': 1991,
'137.138.47.216:8080:TCP': 2528,
'137.138.99.180:10000:TCP': 135,
'10.20.1.205:1094:TCP': 25498,
'10.20.1.206:1094:TCP': 11930,
'193.205.66.200:1094:TCP': 40,
'10.20.2.233:1094:TCP': 23851,
'137.138.99.130:8050:TCP': 683,
'137.138.99.140:8080:TCP': 966,
'137.138.99.182:10000:TCP': 105,
'188.184.38.72:1094:TCP': 40,
'137.138.99.187:10000:TCP': 209,
'137.138.99.136:8080:TCP': 170,
'137.138.99.169:10000:TCP': 92,
'137.138.99.170:10000:TCP': 96,
'137.138.99.188:10000:TCP': 64,
'137.138.99.138:10000:TCP': 64,
'137.138.99.145:8080:TCP': 239,
'137.138.99.179:10000:TCP': 32,
'137.138.99.168:10000:TCP': 32
    
```



```

----- xrootd-redirector-stats Begin -----
This redirector has received requests from the following hosts:
137.138.99.165 73124 ##### pcaliendbg.cern.ch
140.181.2.26 39634 ##### lxaliproxy3.gsi.de
140.181.2.25 39188 ##### lxaliproxy2.gsi.de
137.138.47.208 8624 # pcaliendbmonitor.cern.ch
140.181.2.27 7816 # lxalir1.gsi.de
140.181.2.11 4212 # lxalise1.gsi.de
140.181.2.24 4206 # lxalise4.gsi.de
140.181.2.23 4206 # lxalise2.gsi.de
282.13.230.240 1657
192.168.46.248 1618 fw-nat-inside-outside.gridka.de
137.138.99.186 1573 pcalienstorage2.cern.ch
127.0.10.1 1402 localhost
10.20.2.177 998 lxbk0390.gsi.de
10.20.2.170 992 lxbk0383.gsi.de
10.20.3.128 979 lxbk0534.gsi.de
10.20.3.144 965 lxbk0550.gsi.de
    
```

Requests that were actually rewritten





**CHALLENGE  
ACCEPTED**

# Solution

TL;DR:

- `+iptables +iproute2 +keepalived`
- `-xrootd`

## Solution – Longer version (1)

On the Worker Nodes...

- We **mark** all packets produced by the users for ALICE grid jobs

```
# iptables -A OUTPUT -t mangle -p tcp -m owner  
--uid-owner 3343 -j MARK --set-mark 1
```

- Also, we create a custom **routing table**

```
# cat /etc/iproute2/rt_tables  
255 local  
254 main  
253 default  
0 unspec  
201 alinat
```

## Solution – Longer version (2)

On the Worker Nodes...

- The new routing table is filled with routes to GSI-internal networks and the **default route** being our proxy server

```
# ip r show table alinat
default via 10.20.3.227 dev ib0
...
140.181.2.0/24 via 10.20.0.1 dev ib0
140.181.60.0/24 via 10.20.0.1 dev ib0
```

- Finally, we assign our marked packets to the new routing table:

```
# ip rule add fwmark 1 table alinat
```



## Solution – Longer version (3)

On the Proxy/NAT machine...

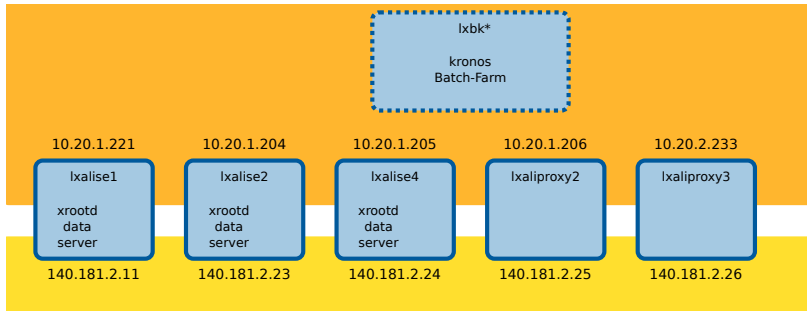
- We enable `/proc/sys/net/ipv4/ip_forward`,
- and forwarding from the infiniband network to the ethernet network
- The packets leaving through the ethernet network are NATted using MASQUERADE so that the packets seem to originate from the NAT machine. The actual source of the packets (private IP) is masked.

# Fault Tolerance

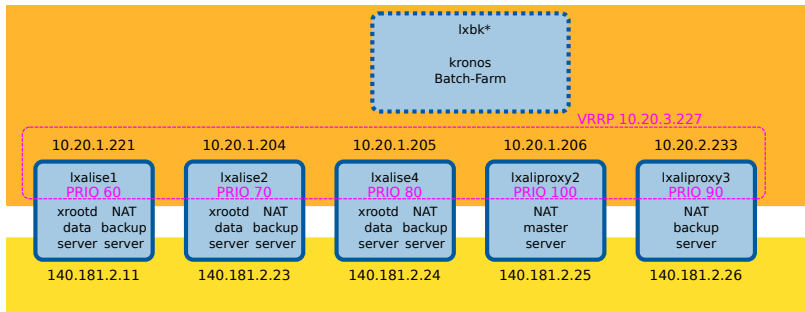
keepalived

- Allows to have multiple servers being in a group
- Each server gets an operator-assigned priority
- In an Active/Standby configuration, exactly one of the servers in the group is the master
- The master status gets elected by the keepalive daemons according to their priorities
- The self-proclaimed master announces that he is now in possession of the virtual IP assigned to the group

# Fault Tolerance



# Fault Tolerance

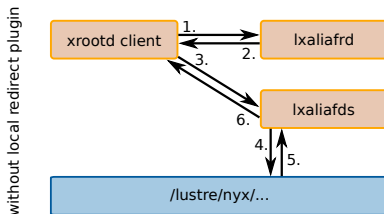


## Result

- Negligible CPU+Memory usage, high performance, lightweight
- Redundancy of 5 (The NAT functionality fails only if both proxy servers and all 3 data servers are down)
- Completely transparent
  - Rewrite mechanism no longer needed
- Minor issues were solved quickly as they appeared
  - Forgot a route to a specific internal network
  - Forgot to use custom routing table for local ALICE users.  
Solved by marking packets of users with primary group ALICE.
  - Some WNs “lost” their routing table on reboot because systemd solved a cyclic dependency by not starting ferm
  - Some WNs tried to start our unit file before the network was up. Race condition solved by adding  
`AFTER=NETWORK-ONLINE.TARGET`

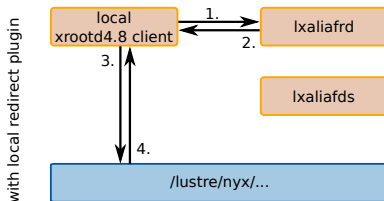
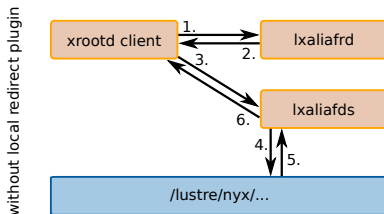
## R&D for an ALICE Analysis Facility

## Local Redirect Plugin: Block Diagram



with local redirect plugin

## Local Redirect Plugin: Block Diagram





## Local Redirect Plugin: Deployment

Needed for use in ALICE T2:

- Migration to xrootd client 4.8
- Resolve remaining issues with TjAliEnFile etc

# xrootd Symlink Plugin (based on work by A. J. Peters)

## PFNs

```
00
|-- 00000
|-- 088d7cc8-da7f-11e7-b758-33a6efa41c83
|-- 08f9f16c-ddd3-11e7-8f50-9f57e7b16c91
|-- 0dc17ac8-e259-11e7-b2cf-1bea57ee5847
|-- 157beb76-ddb9-11e7-bb98-27ed61c99445
|-- 16c63298-ce3c-11e7-8f21-ffc1867ae6dc
|-- 1a799654-ddce-11e7-81f7-57d45aeb36d8
|-- 1dfbb434-ddc8-11e7-b3af-c327ab27e31e
|-- 2371473e-dbda-11e7-a59f-6bd3b799e869
|-- 2ad8a52a-d951-11e7-a5fb-eb58f0ce0e83
|-- 2b3f8ac2-dca0-11e7-ab4d-8fcfef603405
|-- 2e8b0c1e-c798-11e7-8f0d-63092e81fada
|-- 33071f9c-dc03-11e7-bd0c-93f7ffc6b6f30
|-- 3eec59b4-e250-11e7-a31d-3b5de7fdf6a1
|-- 4baa1794-d513-11e7-9ef1-6be888c797fb
|-- 52f08308-dca7-11e7-be2a-6faa7b7c5df2
|-- 588cf82c-d5aa-11e7-8baa-57f00899f828
|-- 50bcde16-db2c-11e7-b8cd-9383431f06ef
|-- 598ccff6-dc51-11e7-a264-1b1fa4f4a5f7
|-- 59a559e0-dc29-11e7-ada6-7fa307fcb939
|-- 59ede29a-ddd3-11e7-8dd1-bbb01635d991
|-- 5e1ce3d6-d54d-11e7-ab6f-4f0fe738722b
|-- 5f9c221a-d559-11e7-a8ce-9b0eabb5cd50
|-- 658a581c-de3d-11e7-880d-f7af987217fa
```

# xrootd Symlink Plugin (based on work by A. J. Peters)

## PFNs

```
00
|-- 00000
|-- 088d7cc8-da7f-11e7-b758-33a6efa41c83
|-- 08f9f16c-ddd3-11e7-8f50-9f57e7b16c91
|-- 0dc17ac8-e259-11e7-b2cf-1bea57ee5847
|-- 157beb76-ddb9-11e7-bb98-27ed61c99445
|-- 16c63298-ce3c-11e7-8f21-ffc1867ae6dc
|-- 1a799654-ddce-11e7-81f7-57d45aeb36dc
|-- 1dfbb434-ddc8-11e7-b3af-c327ab27e31e
|-- 2371473e-dbaa-11e7-a59f-6bd3b799e869
|-- 2ad8a52a-d951-11e7-a5fb-eb58f0ce0e83
|-- 2b3f8ac2-dca0-11e7-ab4d-8fcfef603405
|-- 2e8b0c1e-c798-11e7-8f8d-63092e81fada
|-- 33071f9c-dc03-11e7-bd0c-93f7ffc6b6f30
|-- 3eec59b4-e250-11e7-a31d-3b5de7fdf6a1
|-- 4baa1794-d513-11e7-9ef1-6be888c797fb
|-- 52f08308-dca7-11e7-be2a-6faa7b7c5df2
|-- 588cf82c-d5aa-11e7-8ba4-57f00899f828
|-- 58bcde16-db2c-11e7-b8cd-9383431f06ef
|-- 598ccffe-dc51-11e7-a264-1b1fa4fa5f7
|-- 59a599e0-dc29-11e7-ad46-7fa307fc9393
|-- 59ede29a-ddd3-11e7-8d01-bbb01635d991
|-- 5e1ce3d6-d54d-11e7-ab6f-4f0fe738722b
|-- 5f9c221a-d559-11e7-a8ce-9b0eabb5cd50
|-- 658a581c-de3d-11e7-880d-f7af987217fa
```

## LFNs

```
2015/LHC15o/000245831/pass1/AOD194/3071/root_archive.zip
2015/LHC15o/000246751/pass1/AOD194/PWG/HF/D2H PbPb/1164_20171215-1139_child_1/0363/root_archive.zip
2015/LHC15o/000245353/pass1_pidfix/AOD194/3034/root_archive.zip
2015/LHC15o/000245766/pass1/AOD194/1419/root_archive.zip
2015/LHC15o/000245064/pass3_lowIR_pidfix/AOD194/1447/root_archive.zip
2015/LHC15o/000245409/pass1_pidfix/AOD194/3744/root_archive.zip
2015/LHC15o/000245692/pass1/AOD194/PWG/HF/D2H PbPb/1164_20171215-1139_child_1/0459/root_archive.zip
2015/LHC15o/000245545/pass1_pidfix/AOD194/0158/root_archive.zip
2015/LHC15o/000246984/pass1/AOD194/3672/root_archive.zip
2015/LHC15o/000245752/pass1/AOD194/1392/root_archive.zip
2015/LHC15o/000246844/pass1/AOD194/0311/root_archive.zip
2015/LHC15o/000246989/pass1/AOD194/PWG/HF/D2H PbPb/1173_20171215-1545_child_1/0861/root_archive.zip
2015/LHC15o/000245411/pass1_pidfix/AOD194/4175/root_archive.zip
2015/LHC15o/000246276/pass1/AOD194/PWG/HF/D2H PbPb/1173_20171215-1545_child_1/0034/root_archive.zip
2015/LHC15o/000245683/pass1/AOD194/PWG/HF/D2H PbPb/1172_20171215-1527_child_1/0140/root_archive.zip
2015/LHC15o/000246087/pass1/AOD194/PWG/HF/D2H PbPb/1173_20171215-1545_child_1/3470/root_archive.zip
2015/LHC15o/000245554/pass1_pidfix/AOD194/1394/root_archive.zip
2015/LHC15o/000245146/pass1_pidfix/AOD194/0990/root_archive.zip
2015/LHC15o/000246928/pass1/AOD194/PWG/HF/D2H PbPb/1172_20171215-1527_child_1/0068/root_archive.zip
2015/LHC15o/000246217/pass1/AOD194/1633/root_archive.zip
2015/LHC15o/000245543/pass1_pidfix/AOD194/2045/root_archive.zip
2015/LHC15o/000246275/pass1/AOD194/1203/root_archive.zip
2015/LHC15o/000246048/pass1/AOD194/PWG/HF/D2H PbPb/1173_20171215-1545_child_1/0133/root_archive.zip
2015/LHC15o/000246648/pass1/AOD194/0320/root_archive.zip
2015/LHC15o/000245829/pass1/AOD194/0550/root_archive.zip
2015/LHC15o/000246001/pass1/AOD194/1686/root_archive.zip
2015/LHC15o/000245954/pass1/AOD194/2687/root_archive.zip
```

# xrootd Symlink Plugin (based on work by A. J. Peters)

## PFNs

```
00
|-- 00000
|-- 088d7cc8-da7f-11e7-b758-33a6efa41c83
|-- 08f9f16c-ddd3-11e7-8f50-9f5e7b16c91
|-- 0dc17ac8-e259-11e7-b2cf-1bea57ee5847
|-- 157beb76-ddb9-11e7-bb98-27ed61c99445
|-- 16c32398-ce3c-11e7-8f21-ffc1867ae6dc
|-- 1a799654-ddce-11e7-81f7-57d45aeb36d8
|-- 1dfbb434-ddc8-11e7-b3af-c327ab27e31e
|-- 2371473e-dbda-11e7-a59f-6bd3b799e869
|-- 2ad8a52a-d951-11e7-a5fb-eb58f0ce0e83
|-- 2b3f8ac2-dca0-11e7-ab4d-8fcfef603405
|-- 2e8b0c1e-c798-11e7-8f8d-63092e81fada
|-- 33071f9c-dc03-11e7-bd0c-93f7ffc6bf30
|-- 3eec59b4-e250-11e7-a31d-3b5de7df6a1
|-- 4baa1794-d513-11e7-9ef1-6be888c797fb
|-- 52f08308-dca7-11e7-be2a-6faa7b7c5df2
|-- 588cf82c-d5aa-11e7-8baa-57f00899f828
|-- 58bcde16-db2c-11e7-b8cd-9383431f06ef
|-- 598ccff6-dc51-11e7-a264-1b1fa4fa5f7
|-- 59a59e0b-dc29-11e7-ad46-7fa307fcb939
|-- 59ede29a-ddd3-11e7-8d01-bbb01635d991
|-- 5e1cc3d6-d54d-11e7-ab6f-4f0fe738722b
|-- 5f9c221a-d559-11e7-a8ce-9b0eabb5cd50
|-- 658a581c-de3d-11e7-880d-f7af9872171fa
```



## LFNs

```
2015/LHC15o/000245831/pass1/AOD194/3071/root_archive.zip
2015/LHC15o/000246751/pass1/AOD194/PWG/HF/D2H PbPb/1164_20171215-1139_child_1/0363/root_archive.zip
2015/LHC15o/000245353/pass1_pidfix/AOD194/3034/root_archive.zip
2015/LHC15o/000245766/pass1/AOD194/1419/root_archive.zip
2015/LHC15o/000245064/pass3_lowIR_pidfix/AOD194/1447/root_archive.zip
2015/LHC15o/000245409/pass1_pidfix/AOD194/3744/root_archive.zip
2015/LHC15o/000245692/pass1/AOD194/PWG/HF/D2H PbPb/1164_20171215-1139_child_1/0459/root_archive.zip
2015/LHC15o/000245545/pass1_pidfix/AOD194/0158/root_archive.zip
2015/LHC15o/000246984/pass1/AOD194/3672/root_archive.zip
2015/LHC15o/000245752/pass1/AOD194/1392/root_archive.zip
2015/LHC15o/000246844/pass1/AOD194/0311/root_archive.zip
2015/LHC15o/000246989/pass1/AOD194/PWG/HF/D2H PbPb/1173_20171215-1545_child_1/0861/root_archive.zip
2015/LHC15o/000245411/pass1_pidfix/AOD194/4175/root_archive.zip
2015/LHC15o/000246276/pass1/AOD194/PWG/HF/D2H PbPb/1173_20171215-1545_child_1/0034/root_archive.zip
2015/LHC15o/000245683/pass1/AOD194/PWG/HF/D2H PbPb/1172_20171215-1527_child_1/0140/root_archive.zip
2015/LHC15o/000246087/pass1/AOD194/PWG/HF/D2H PbPb/1173_20171215-1545_child_1/3470/root_archive.zip
2015/LHC15o/000245554/pass1_pidfix/AOD194/1394/root_archive.zip
2015/LHC15o/000245146/pass1_pidfix/AOD194/0990/root_archive.zip
2015/LHC15o/000246928/pass1/AOD194/PWG/HF/D2H PbPb/1172_20171215-1527_child_1/0068/root_archive.zip
2015/LHC15o/000246217/pass1/AOD194/1633/root_archive.zip
2015/LHC15o/000245543/pass1_pidfix/AOD194/2045/root_archive.zip
2015/LHC15o/000246275/pass1/AOD194/1203/root_archive.zip
2015/LHC15o/000246048/pass1/AOD194/PWG/HF/D2H PbPb/1173_20171215-1545_child_1/0133/root_archive.zip
2015/LHC15o/000246648/pass1/AOD194/0320/root_archive.zip
2015/LHC15o/000245829/pass1/AOD194/0550/root_archive.zip
2015/LHC15o/000246001/pass1/AOD194/1686/root_archive.zip
2015/LHC15o/000245954/pass1/AOD194/2687/root_archive.zip
```

# xrootd Symlink Plugin (based on work by A. J. Peters)

## PFNs

```
00
-- 00000
|-- 088d7cc8-da7f-11e7-b758-33a6efa41c83
|-- 08f9f16c-ddd3-11e7-8f50-9f57e7b16c91
|-- 0dc17ac8-e259-11e7-b2cf-1bea57ee5847
|-- 157beb76-ddb9-11e7-bb98-27ed61c99445
|-- 16c63298-ce3c-11e7-8f21-ffc1867ae6dc
|-- 1a799654-ddce-11e7-81f7-57d45aeb36d8
|-- 1dfbb434-ddc8-11e7-b3af-c327ab27e31e
|-- 2371473e-dbda-11e7-a59f-6bd3b799e869
|-- 2ad8a52a-d951-11e7-a5fb-eb58f0ce0e83
|-- 2b3f8ac2-dca0-11e7-ab4d-8fcfef603405
|-- 2e8b0c1e-c798-11e7-8f8d-63092e81fada
|-- 33071f9c-dc03-11e7-bd0c-93f7ffc6b30
|-- 3eec59b4-e250-11e7-a31d-3b5de7fdf6a1
|-- 4baa1794-d513-11e7-9ef1-6be888c797fb
|-- 52f08308-dca7-11e7-be2a-6faa7b7c5df2
|-- 588cf82c-d5aa-11e7-8baa-57f00899f828
|-- 58bcde16-db2c-11e7-b8cd-9383431f06ef
|-- 598ccffe-dc51-11e7-a264-1b1fa4fa5f7
|-- 59a59e0b-dc29-11e7-ad46-7fa307fcb939
|-- 59ede29a-ddd3-11e7-8d01-bbb01635d991
|-- 5e1cc3d6-d54d-11e7-ab6f-4f0fe738722b
|-- 5f9c221a-d559-11e7-a8ce-9b0eabb5cd50
|-- 658a581c-de3d-11e7-880d-f7af9872171fa
```

## LFNs

```
2015/LHC15o/000245831/pass1/AOD194/3071/root_archive.zip
2015/LHC15o/000246751/pass1/AOD194/PWG/HF/D2H PbPb/1164_20171215-1139_child_1/0363/root_archive.zip
2015/LHC15o/000245353/pass1_pidfix/AOD194/3034/root_archive.zip
2015/LHC15o/000245766/pass1/AOD194/1419/root_archive.zip
2015/LHC15o/000245064/pass3_lowIR_pidfix/AOD194/1447/root_archive.zip
2015/LHC15o/000245409/pass1_pidfix/AOD194/3744/root_archive.zip
2015/LHC15o/000245692/pass1/AOD194/PWG/HF/D2H PbPb/1164_20171215-1139_child_1/0459/root_archive.zip
2015/LHC15o/000245545/pass1_pidfix/AOD194/0158/root_archive.zip
2015/LHC15o/000246984/pass1/AOD194/3672/root_archive.zip
2015/LHC15o/000245752/pass1/AOD194/1392/root_archive.zip
2015/LHC15o/000246844/pass1/AOD194/0311/root_archive.zip
2015/LHC15o/000246989/pass1/AOD194/PWG/HF/D2H PbPb/1173_20171215-1545_child_1/0861/root_archive.zip
2015/LHC15o/000245411/pass1_pidfix/AOD194/4175/root_archive.zip
2015/LHC15o/000246276/pass1/AOD194/PWG/HF/D2H PbPb/1173_20171215-1545_child_1/0034/root_archive.zip
2015/LHC15o/000245683/pass1/AOD194/PWG/HF/D2H PbPb/1172_20171215-1527_child_1/0140/root_archive.zip
2015/LHC15o/000246087/pass1/AOD194/PWG/HF/D2H PbPb/1173_20171215-1545_child_1/3470/root_archive.zip
2015/LHC15o/000245554/pass1_pidfix/AOD194/1394/root_archive.zip
2015/LHC15o/000245146/pass1_pidfix/AOD194/0990/root_archive.zip
2015/LHC15o/000246928/pass1/AOD194/PWG/HF/D2H PbPb/1172_20171215-1527_child_1/0068/root_archive.zip
2015/LHC15o/000246217/pass1/AOD194/1633/root_archive.zip
2015/LHC15o/000245543/pass1_pidfix/AOD194/2045/root_archive.zip
2015/LHC15o/000246275/pass1/AOD194/1203/root_archive.zip
2015/LHC15o/000246048/pass1/AOD194/PWG/HF/D2H PbPb/1173_20171215-1545_child_1/0133/root_archive.zip
2015/LHC15o/000246648/pass1/AOD194/0320/root_archive.zip
2015/LHC15o/000245829/pass1/AOD194/0550/root_archive.zip
2015/LHC15o/000246001/pass1/AOD194/1686/root_archive.zip
2015/LHC15o/000245954/pass1/AOD194/2687/root_archive.zip
```



# xrootd Symlink Plugin (based on work by A. J. Peters)

LFN  $\mapsto$  PFN

```
aliafse@lxaliafids1:/lustre/nyx/alice/aliafse/links/alice/data$ tree
```

```
-- 2015
  |-- LHC15o
      |-- 000244917
          |-- lowIR_standaloneITS
              |-- AOD194
                  |-- 0001
                      root_archive.zip -> /lustre/nyx/alice/aliafse/data//12/31062/74606222-b0c1-11e7-bbee-a769c35bc5b3
                  |-- 0002
                      root_archive.zip -> /lustre/nyx/alice/aliafse/data//01/24658/9b6c806c-b0c1-11e7-bd36-bf67323ee43f
                  |-- 0003
                      root_archive.zip -> /lustre/nyx/alice/aliafse/data//02/57812/799d7734-b0c1-11e7-9e85-139fd3e3d5ba
                  |-- 0004
                      root_archive.zip -> /lustre/nyx/alice/aliafse/data//14/47239/a50348a4-b0c1-11e7-b781-07bfa7b0c1a4
                  |-- 0005
                      root_archive.zip -> /lustre/nyx/alice/aliafse/data//10/15913/872c6cf2-b0c1-11e7-b1ba-0ff7c523fa24
                  |-- 0006
                      root_archive.zip -> /lustre/nyx/alice/aliafse/data//07/07668/a9e1555a-b0c1-11e7-ba33-974391c7754e
                  |-- 0007
                      root_archive.zip -> /lustre/nyx/alice/aliafse/data//12/60874/879a97cc-b0c1-11e7-82d8-83ea88562022
                  |-- 0008
                      root_archive.zip -> /lustre/nyx/alice/aliafse/data//13/38259/e3db360e-b0c1-11e7-8630-e3d66a24f2f3
                  |-- 0009
                      root_archive.zip -> /lustre/nyx/alice/aliafse/data//01/63480/9e1899b8-b0c1-11e7-9bc5-5b33fd1dda3e
                  |-- 0010
                      root_archive.zip -> /lustre/nyx/alice/aliafse/data//10/62671/cb210e72-b0c1-11e7-b000-ab29bcd6ae3a
                  |-- 0011
                      root_archive.zip -> /lustre/nyx/alice/aliafse/data//07/47248/c2428722-b0c1-11e7-81bf-e3b556cb45c6
                  |-- 0012
                      root_archive.zip -> /lustre/nyx/alice/aliafse/data//13/25939/e583b8be-b0c1-11e7-a346-475907fe2c8a
                  |-- 0013
                      root_archive.zip -> /lustre/nyx/alice/aliafse/data//05/65220/08e04296-b0c2-11e7-9d73-8f4396de0020
                  |-- 0014
                      root_archive.zip -> /lustre/nyx/alice/aliafse/data//01/26971/0c5f07fe-b0c2-11e7-a18c-f33e60f7a678
                  |-- 0015
```

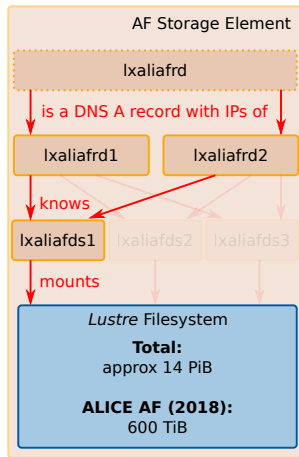
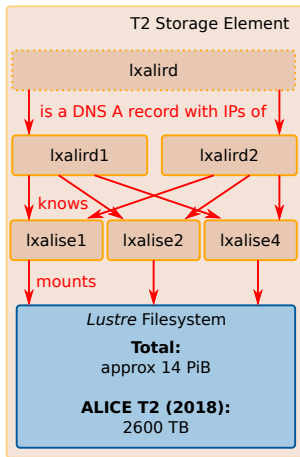
# xrootd Symmlink Plugin (based on work by A. J. Peters)

PFN  $\longleftrightarrow$  LFN

```

-- 0880/cc8-da71-11e7-d038-35a0e1a41c83_symlink -> /2015/LHC150/000245952/pass1/AOD194/3481/root_archive.zip
-- 08f9f16c-ddd3-11e7-8f50-9f57e7b16c91_symlink -> /2015/LHC150/000246980/pass1/AOD194/0899/root_archive.zip
-- 0dc17ac8-e259-11e7-b2cf-1bea57ee5847_symlink -> /2015/LHC150/000246758/pass1/AOD194/PWGHF/D2H_PbPb/1173_20171215-1545_chi
-- 1575eb7c-dbb9-11e7-bb98-27ed61c99445_symlink -> /2015/LHC150/000246805/pass1/AOD194/0360/root_archive.zip
-- 16c63298-ce3c-11e7-8f21-ffc1867ae6dc_symlink -> /2015/LHC150/000245146/pass1_pidfix/AOD194/0524/root_archive.zip
-- 1a799654-ddce-11e7-81f7-57d45aeb36d8_symlink -> /2015/LHC150/000246847/pass1/AOD194/0027/root_archive.zip
-- 1dfbb434-ddc8-11e7-b3af-c327ab27e31e_symlink -> /2015/LHC150/000246847/pass1/AOD194/1258/root_archive.zip
-- 2371473e-dbda-11e7-a59f-6bd3b799e869_symlink -> /2015/LHC150/000245441/pass1_pidfix/AOD194/2099/root_archive.zip
-- 2ad8a52a-d951-11e7-a5fb-eb58f0ce0e83_symlink -> /2015/LHC150/000246153/pass1/AOD194/0353/root_archive.zip
-- 2b3f8ac2-dca0-11e7-ab4d-8fcfef603405_symlink -> /2015/LHC150/000246757/pass1/AOD194/0439/root_archive.zip
-- 2e8b0c1e-c798-11e7-8f8d-63092e81fada_symlink -> /2015/LHC150/000246392/pass3_lowIR_pidfix/AOD194/1307/root_archive.zip
-- 33071f9c-dc03-11e7-bd8c-93f7ffc6b6f30_symlink -> /2015/LHC150/000246271/pass1/AOD194/1291/root_archive.zip
-- 3eec59b4-e250-11e7-a31d-3b5de7fd6fa1_symlink -> /2015/LHC150/000246994/pass1/AOD194/PWGHF/D2H_PbPb/1173_20171215-1545_chi
-- 4baa1794-d513-11e7-9ef1-6be888c797fb_symlink -> /2015/LHC150/000245554/pass1_pidfix/AOD194/3415/root_archive.zip
-- 52f08308-dca7-11e7-be2a-6faa7b7c5df2_symlink -> /2015/LHC150/000246758/pass1/AOD194/0123/root_archive.zip
-- 588cfd2c-d5aa-11e7-8baa-9f700899f828_symlink -> /2015/LHC150/000245700/pass1/AOD194/0129/root_archive.zip
-- 58b8ce16-db2c-11e7-b8cd-9383431f06ef_symlink -> /2015/LHC150/000245963/pass1/AOD194/0581/root_archive.zip
-- 598ccff6-dc51-11e7-a264-1b1fa4f4a5f7_symlink -> /2015/LHC150/000246272/pass1/AOD194/0078/root_archive.zip
-- 59a559e0-dc29-11e7-ada6-7fa307fbc939_symlink -> /2015/LHC150/000246488/pass1/AOD194/2826/root_archive.zip
-- 59ede29a-ddd3-11e7-88d1-bbb01635d991_symlink -> /2015/LHC150/000246805/pass1/AOD194/0010/root_archive.zip
-- 5e1ce3d6-d54d-11e7-ab6f-4f0fe738722b_symlink -> /2015/LHC150/000245505/pass1_pidfix/AOD194/1123/root_archive.zip
-- 5f9c221a-d559-11e7-a8ce-9b0eabbs5cd50_symlink -> /2015/LHC150/000245497/pass1_pidfix/AOD194/1804/root_archive.zip
-- 658a581c-de3d-11e7-880d-f7af987217fa_symlink -> /2015/LHC150/000246984/pass1/AOD194/3387/root_archive.zip
-- 661dd08e-ddba-11e7-a437-ff6felb4f410_symlink -> /2015/LHC150/000246846/pass1/AOD194/0686/root_archive.zip
-- 68d3ee38-dbad-11e7-8ba9-b460bb1eb06c_symlink -> /2015/LHC150/000245439/pass1_pidfix/AOD194/2810/root_archive.zip
-- 6bdf3ab3-d211-11e7-acba-0b3e53fa65c13_symlink -> /2015/LHC150/000246568/pass1/AOD194/0657/root_archive.zip
-- 6e2781f2-dbc2-11e7-9762-9fcdf6bb416_symlink -> /2015/LHC150/000246181/pass1/AOD194/2544/root_archive.zip
-- 784089db-dbac-11e7-a7ff-afb135947ec2_symlink -> /2015/LHC150/000245441/pass1_pidfix/AOD194/1302/root_archive.zip
-- 80d360ac-cd92-11e7-ba32-5fcb1ba7fabb_symlink -> /2015/LHC150/000245232/pass1_pidfix/AOD194/1937/root_archive.zip
-- 812e5666-cd91-11e7-bff1-9b776198df9d_symlink -> /2015/LHC150/000245231/pass1_pidfix/AOD194/0197/root_archive.zip
-- 859ecc90-da42-11e7-8065-af8fa25a47ff_symlink -> /2015/LHC150/000246087/pass1/AOD194/11685/root_archive.zip
    
```

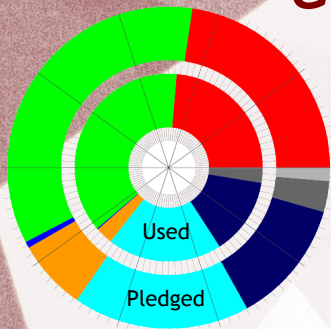
## Redundant xrootd redirectors





# Operations KIT/FZK

# GridKa CPU Resources

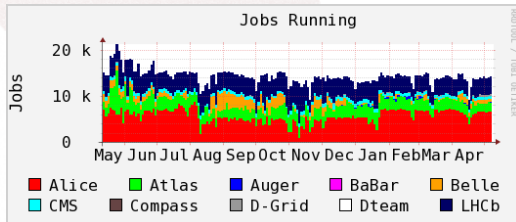


## Cluster Usage 2018

- Support all 4 major LHC collaborations
- 2nd largest pledge for ALICE at 22.65%
- ALICE usage oscillates, but close to pledge (20%-24%)

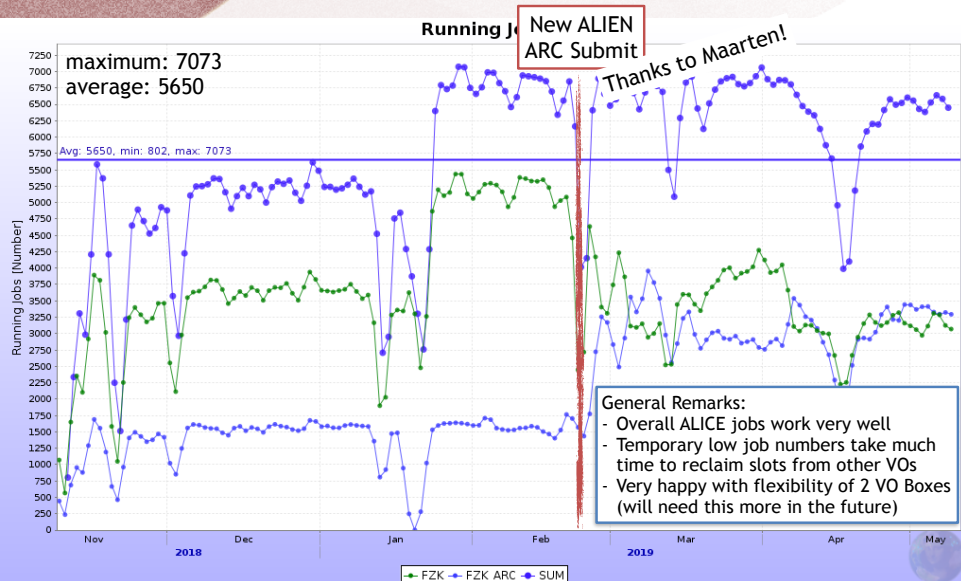
## Batch System Status

- Stable operation with ARC-CE+HTCondor for all VOs
- Ongoing work on fair balancing of Multi-/Single-Core jobs
- Could not match resources to pledge due to legal issues!



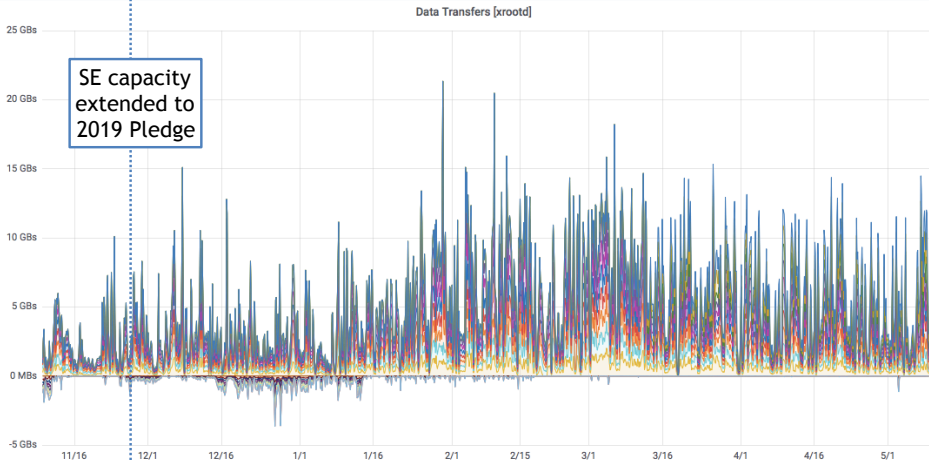
# Jobs at GridKa

(last 6 months)



# XRootD usage@GridKa

- Early addition of 2019 pledge extension in Nov. 2018
- Additional servers to provide stable throughput



# ***GridKa: General Topics***

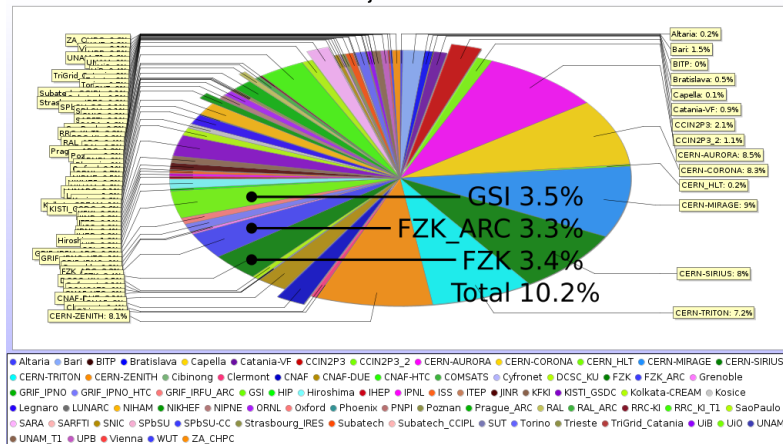
- **Batch System Resource Growth**
  - Procurement for 2019 pledges aborted due to third party legal issues
  - New procurement underway
  - Resources will be made available ASAP
- **Batch System Update**
  - Migrating entire cluster from SL6 to SL7
  - Roll out of IPv6 and Singularity
  - Internal testing done, starting migration middle of May
- **Storage Services Status**
  - Very happy with XRootD on shared GPFS
  - Transparently scaled up File System and added new Servers
  - Ongoing work to improve Tape backend performance

# German Computing Contrib. 2018 by num. DONE jobs

Interval selection: - choose - of 2018-01-01 00:00 - 2018-12-31 23:00 » Plot

What is this about?

Done jobs statistics



# German Storage Contribution 2018

Disk storage elements

GSI,FZK														
AliEn SE			Catalogue statistics					Storage-provided information						
SE Name	AliEn name	Tier	Size	Used	Free	Usage	No. of files	Type	Size	Used	Free	Usage	Version	E
1. FZK - SE	ALICE::FZK::SE	1	7.791 PB	6.904 PB	908.8 TB	88.61%	123,001,705	FILE	7.791 PB	7.269 PB	535.1 TB	93.29%	Xrootd v4.8.0	
2. GSI - AF_SE	ALICE::GSI::AF_SE	2	600 TB	246.1 TB	353.9 TB	41.02%	553,944	FILE	600 TB	245.8 TB	354.2 TB	40.97%	Xrootd v4.8.4	
3. GSI - SE2	ALICE::GSI::SE2	2	2.637 PB	2.432 PB	209.2 TB	92.25%	34,700,341	FILE	2.637 PB	2.433 PB	208.5 TB	92.28%	Xrootd v4.8.4	
<b>Total</b>			<b>11.01 PB</b>	<b>9.577 PB</b>	<b>1.437 PB</b>		<b>158,255,990</b>		<b>11.01 PB</b>	<b>9.942 PB</b>	<b>1.072 PB</b>			

Tape storage elements

GSI,FZK														
AliEn SE			Catalogue statistics					Storage-provided information						
SE Name	AliEn name	Tier	Size	Used	Free	Usage	No. of files	Type	Size	Used	Free	Usage	Version	EOS Versi
1. FZK - TAPE	ALICE::FZK::TAPE	1	601.5 TB	8.348 PB	0	1421%	5,454,282	FILE	601.5 TB	445.1 TB	156.4 TB	74%	Xrootd v4.8.4	
<b>Total</b>			<b>601.5 TB</b>	<b>8.348 PB</b>	<b>0</b>		<b>5,454,282</b>		<b>601.5 TB</b>	<b>445.1 TB</b>	<b>156.4 TB</b>			

## UF: IDS “Arhuaco”

- PhD thesis of Andres Gomez
- Intrusion Prevention and Detection
- Deep Learning
- Grid Computing
- Accuracy > 99%
- False positive rate < 0.07%



## Current challenges / Outlook

- Fulfilling the ALICE pledges for 2020
- Improving our logging, monitoring, and controlling infrastructure

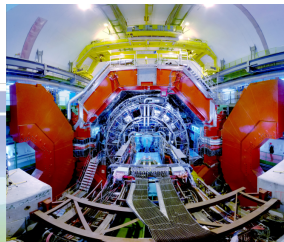
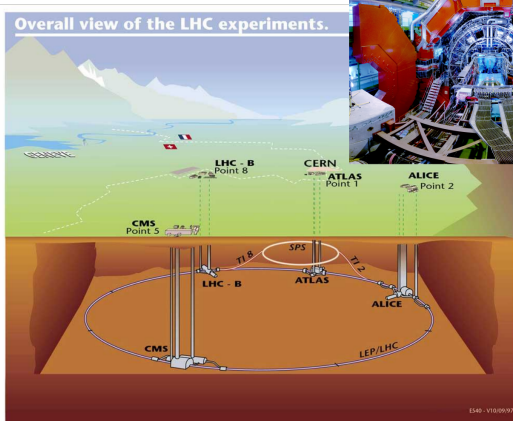
# EOF

Thank you for your attention



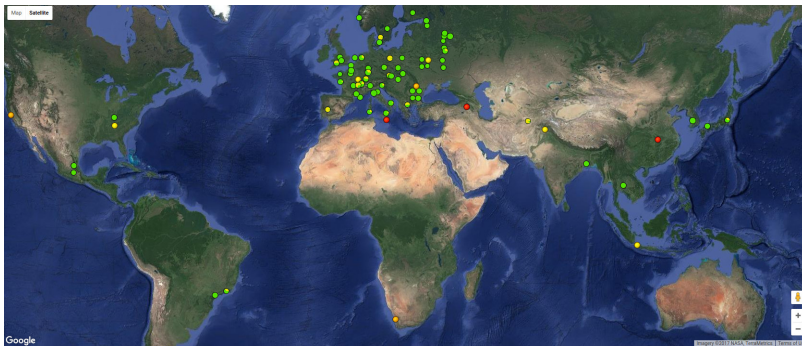
# Backup Slides

# ALICE @ CERN



6 GB/s  
to storage

# ALICE Grid



ALICE sites are distributed worldwide

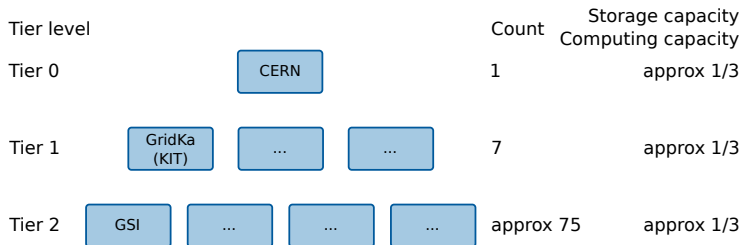
# Hierarchy of ALICE Grid Centers

Tier level		Count	Storage capacity Computing capacity
Tier 0		1	approx 1/3

# Hierarchy of ALICE Grid Centers

Tier level		Count	Storage capacity Computing capacity
Tier 0		1	approx 1/3
Tier 1	  	7	approx 1/3









# Hierarchy of ALICE Grid Centers



- German contribution to ALICE grid:
  - T1 at GridKa:  $\approx 25\%$  of total T1 capacity
  - T2 at GSI:  $\approx 7\%$  of total T2 capacity
  - National analysis facility at GSI ( $2 \times$  GSI T2 capacity)

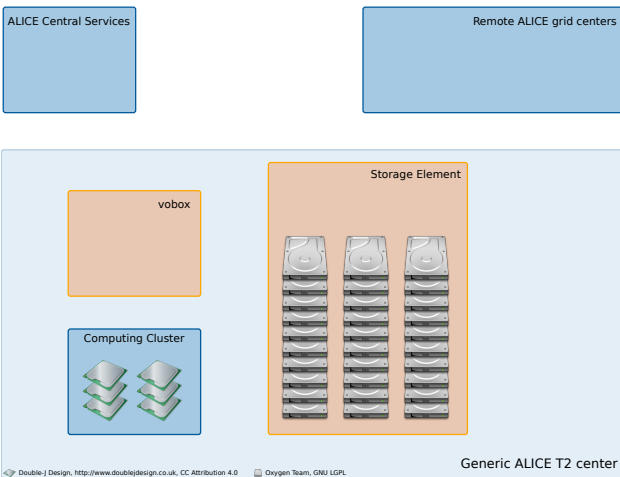


# Hierarchy of ALICE Grid Centers

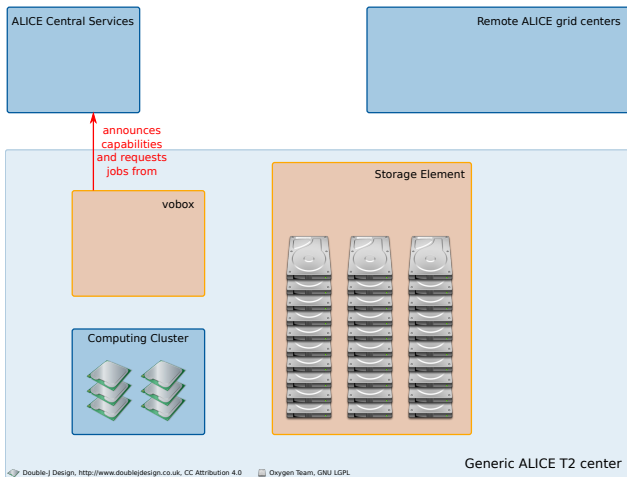
Tier level		Count	Storage capacity Computing capacity
Tier 0		1	approx 1/3
Tier 1	 ...  ... 	7	approx 1/3
Tier 2	 ...  ...  ... 	approx 75	approx 1/3

- German contribution to ALICE grid:
  - T1 at GridKa:  $\approx 25\%$  of total T1 capacity
  - T2 at GSI:  $\approx 7\%$  of total T2 capacity
  - National analysis facility at GSI ( $2 \times$  GSI T2 capacity)

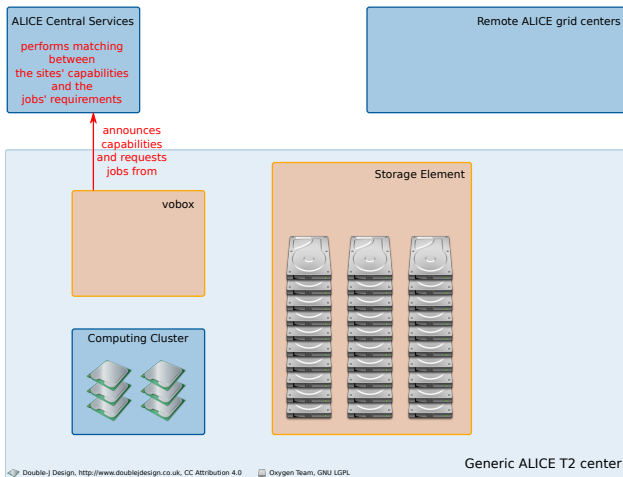
# Generic ALICE Tier 2 Center



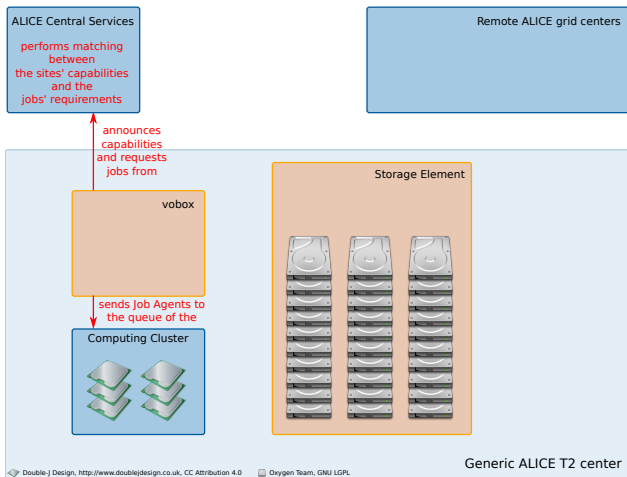
# Generic ALICE Tier 2 Center



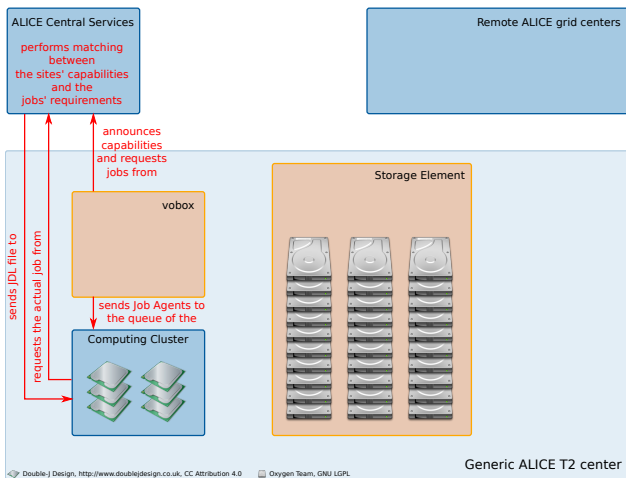
# Generic ALICE Tier 2 Center



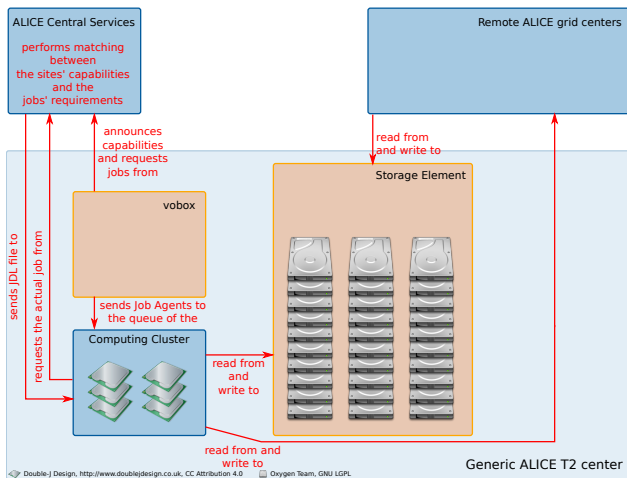
# Generic ALICE Tier 2 Center



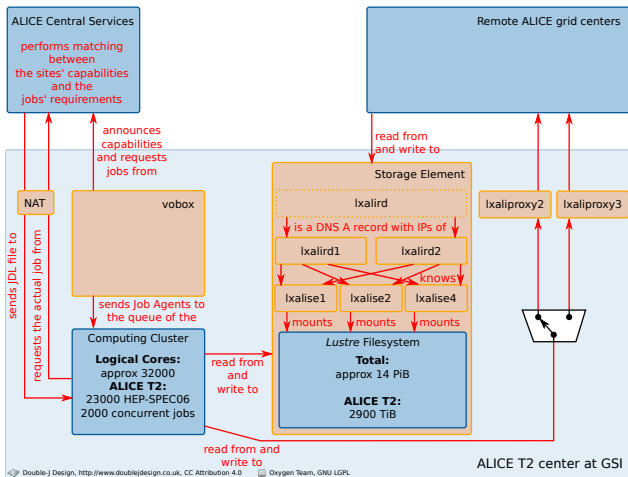
# Generic ALICE Tier 2 Center



# Generic ALICE Tier 2 Center



# ALICE Tier 2 Center at GSI





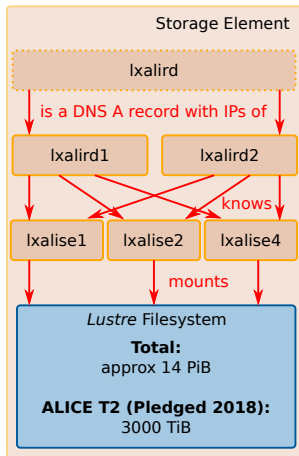
## Redundant xrootd redirectors: Motivation

- Problem: Our xrootd redirector was a single point of failure.
- Solution: Introduce second xrootd redirector
  - Both xrootd redirectors are opaque, with own hostname and IP, but both share the same A record lxalird.gsi.de
  - xrootd clients are capable of dealing with a DNS response with multiple IP addresses

## Redundant xrootd redirectors: Motivation

- Problem: Our xrootd redirector was a single point of failure.
- Solution: Introduce second xrootd redirector
  - Both xrootd redirectors are opaque, with own hostname and IP, but both share the same A record `lxalird.gsi.de`
  - xrootd clients are capable of dealing with a DNS response with multiple IP addresses

## Redundant xrootd redirectors



## Redundant xrootd redirectors

