CERN

EOS 101
Hands on Training

Luca Mascetti

CERN - IT Storage

- Introduction
  - Origins of EOS
  - EOS technological foundation
- Description and Architecture
  - Features
  - Components and architecture
  - Deployment view
- Functionality
  - Data placement
  - Balancing
  - Draining
- Creating an EOS Cluster
- Configuration
  - Replication
  - Authentication
  - Pools and scheduling groups
  - Balancing
  - Draining
  - Permission System
  - Quotas
- Node operations
  - Diskserver maintenance
  - How to failover
- Additional Entry-Point

Introduction

Origins of EOS

Development started on may 2010

Still young compared to other distributed storage systems

- AFS ~35y
- CASTOR ~20y
- Lustre/XRootD ~15y
- Ceph ~13y

First target was the physics analysis use cases at CERN

- needed a performant, scalable and reliable storage
- disentangle disk and tape/archive (**hot & cold** storage)

Origins of EOS

What do you mean by Organic Storage?

Similar to the way a living organism is made of cells that constantly and automatically replenish themselves even in a dynamically changing environment, organic storage allows for continuous growth in the face of recurrently evolving storage requirements. An organic storage system is a system that demonstrates many of the features of a living organism: resilience, self-healing, adaptive accommodation, and constant renewal. It accomplishes this by allowing new servers to be added while client applications remain fully functional, and by providing automatic failover strategies when an existing server is retired, disconnected, or simply stops working. When a server dies inadvertently, absolutely no manual intervention is required.

EOS Technological Foundations

Organic* file storage system built on XRootD** framework

- disk-only storage system on top of JBODs
- open source project in CERN IT-Storage group
- released under GPLv3
  - extremely simple architecture
    - one daemon and three plugins written in C++
    - no relational DB backend, no commercial product dependencies
    - built on top of XRootD framework/protocol
  - easily adaptable to requirements

- CERN code ownership
  - agile development style
  - very short development & release cycles

EOS Releases

named after gemstones

Beryl Aquamarine

V 0.3.X

Citrine

V 4.X

XRootD V3 Server

IPV4

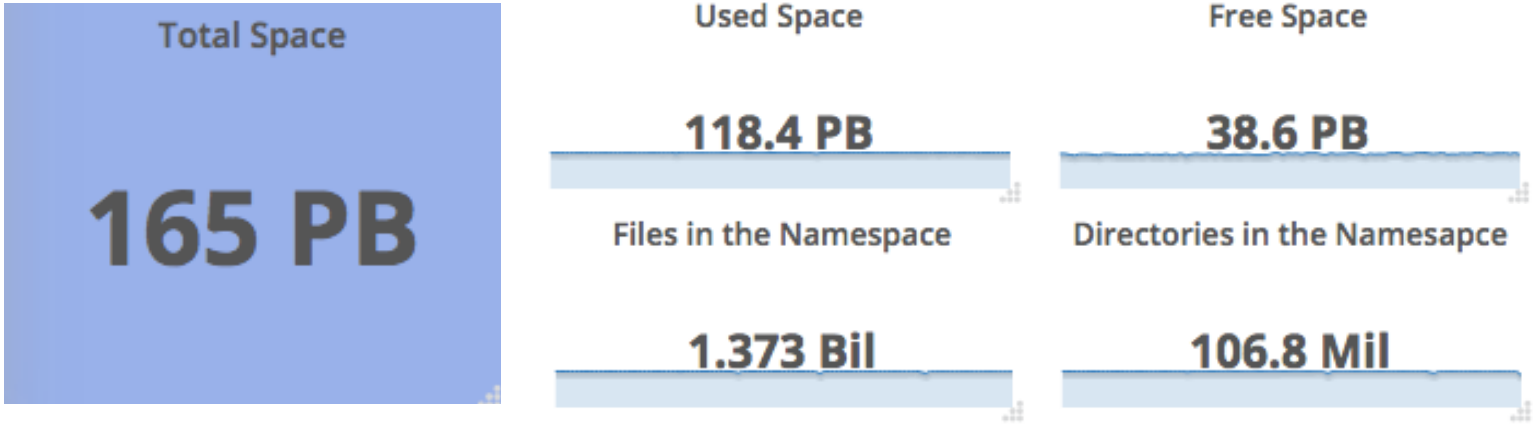
namespace in-memory data on attached disks

XRootD V4 Server

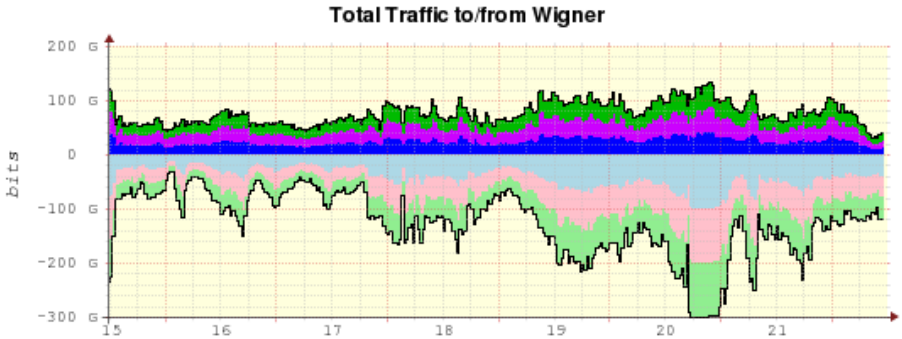
IPV6

plugins for meta data & data persistency

EOS Service @ CERN



- Disk Storage for all LHC and physics data
  - and for CERNBox
- Deployment across two computer centres
  - CERN and Wigner RCP
  - Third link recently added

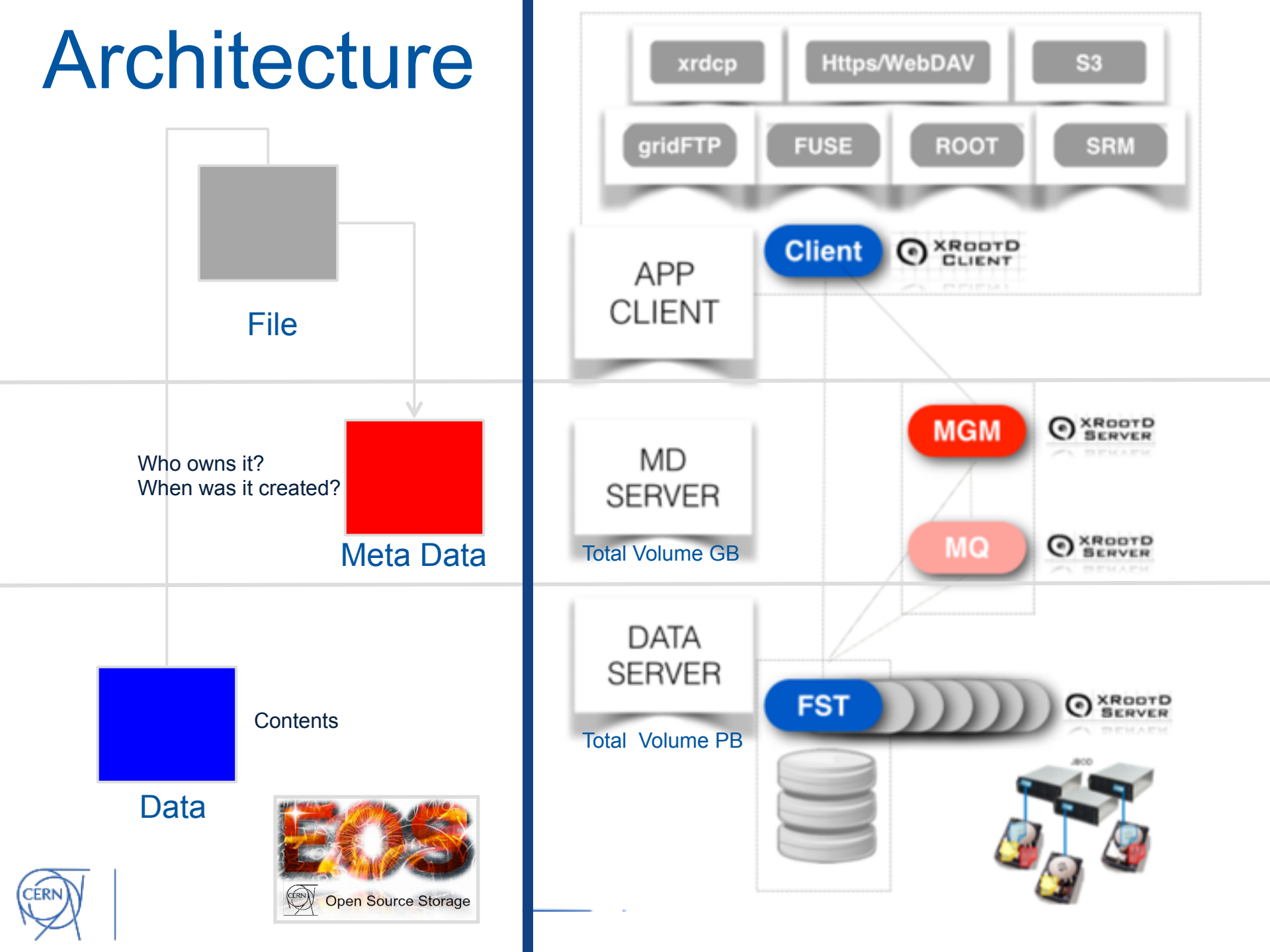


Description and Architecture

A glimpse of EOS features …

- **low latency** due to in-memory namespace (ms)
- **”cheap”** – uses JBODs *just a bunch of disks*
  no RAID controller but software implementation to **replicate** or **erasure encode** files for redundancy
- rich **access control** lists
- user, group & project **quota** system
- easy to operate and deploy
- EOS **server** runs on **Linux** platform
- EOS **client** runs on **Linux**, **OSX** platform
- **SAMBA & WebDav** accessible from **Windows**

Architecture

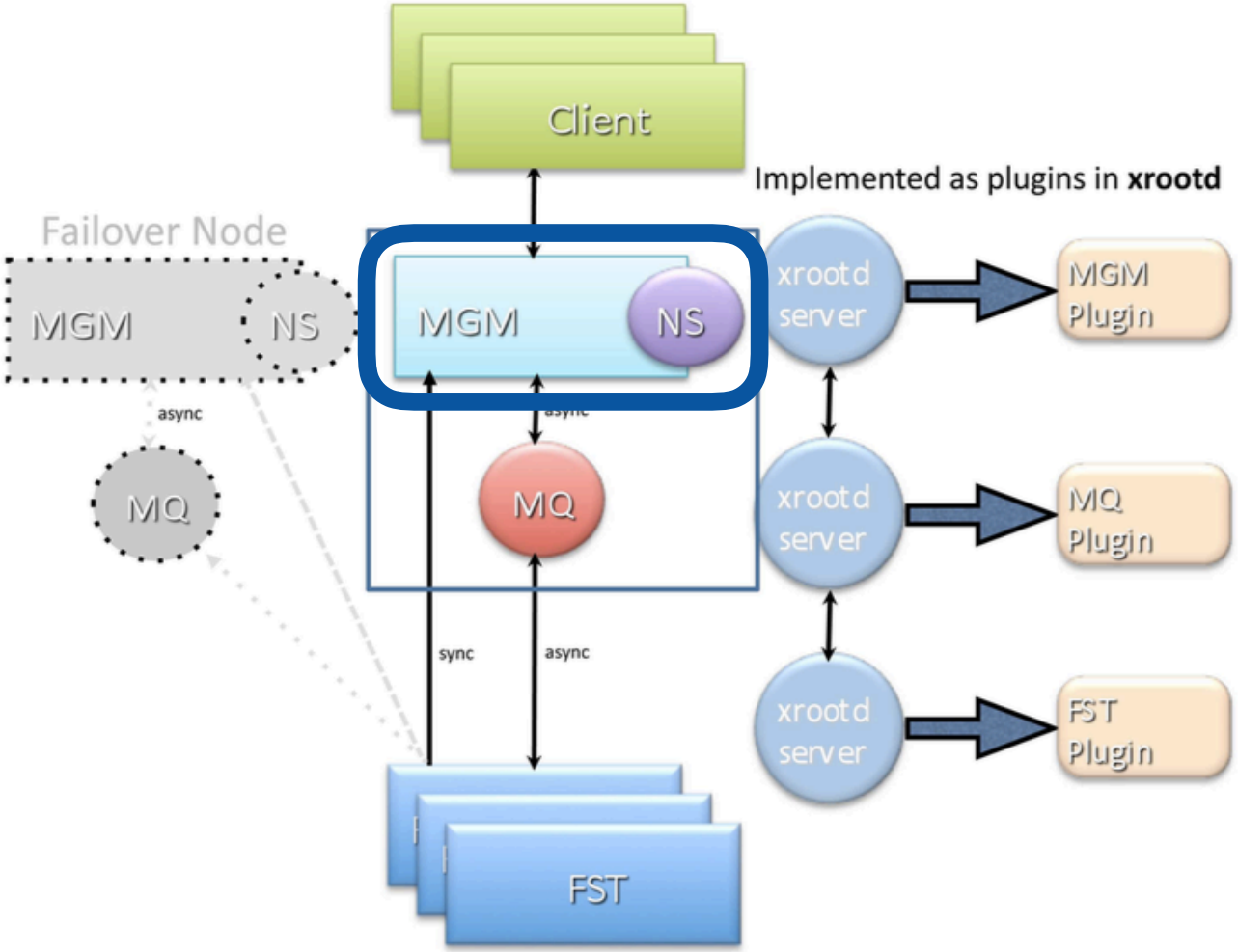


Components

Management Server
Pluggable Namespace, Quota
Strong Authentication
Capability Engine
File Placement
File Location

Message Queue
Service State Messages
File Transaction Reports
Shared Objects (queue+hash)

File Storage
File & File Meta Data Store
Capability Authorization
Check-summing & Verification
Disk Error Detection (Scrubbing)

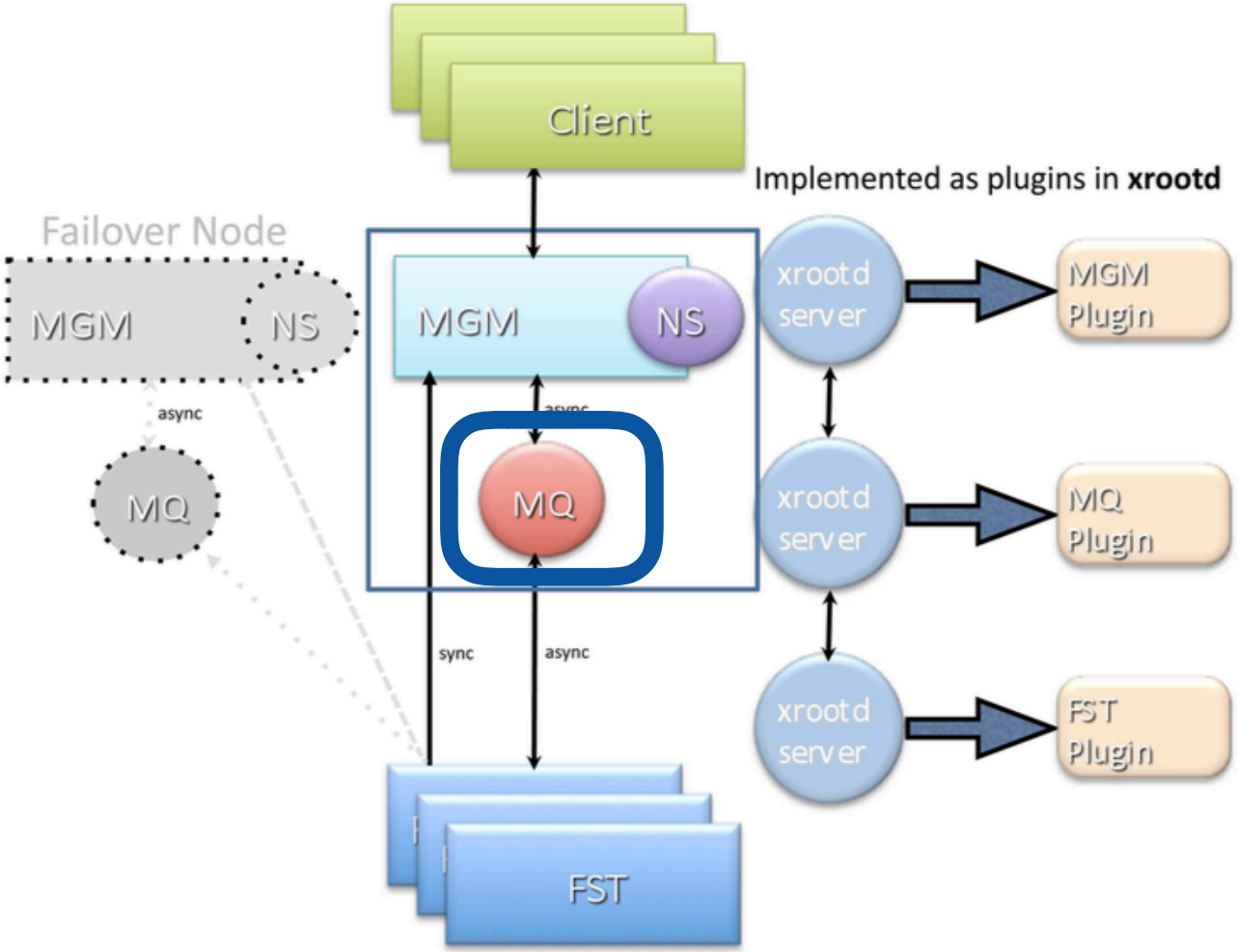


Components

Management Server
Pluggable Namespace, Quota
Strong Authentication
Capability Engine
File Placement
File Location

Message Queue
Service State Messages
File Transaction Reports
Shared Objects (queue+hash)

File Storage
File & File Meta Data Store
Capability Authorization
Check-summing & Verification
Disk Error Detection (Scrubbing)

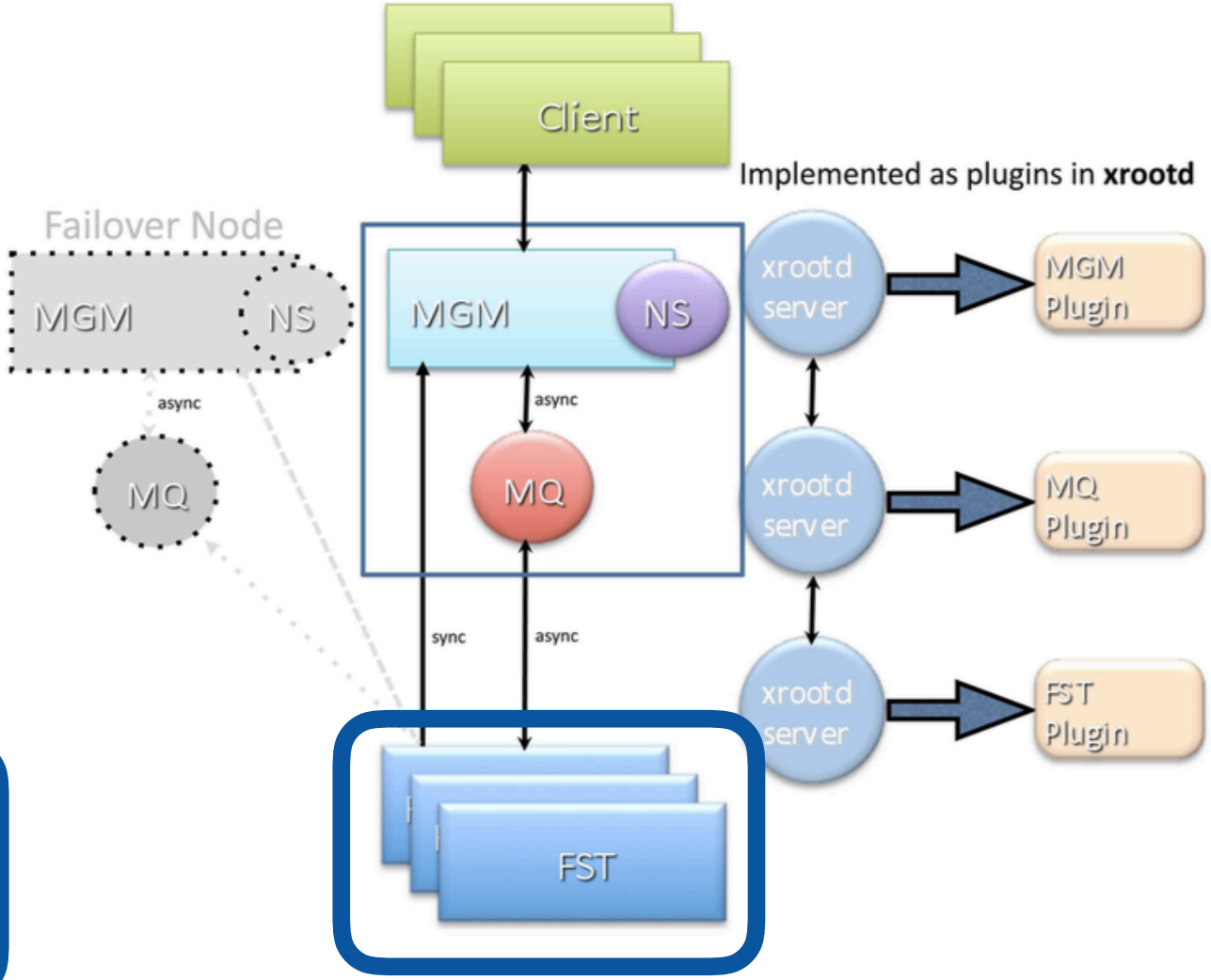


Components

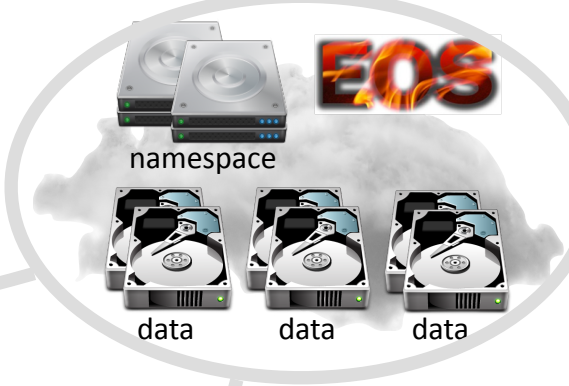
Management Server
Pluggable Namespace, Quota
Strong Authentication
Capability Engine
File Placement
File Location

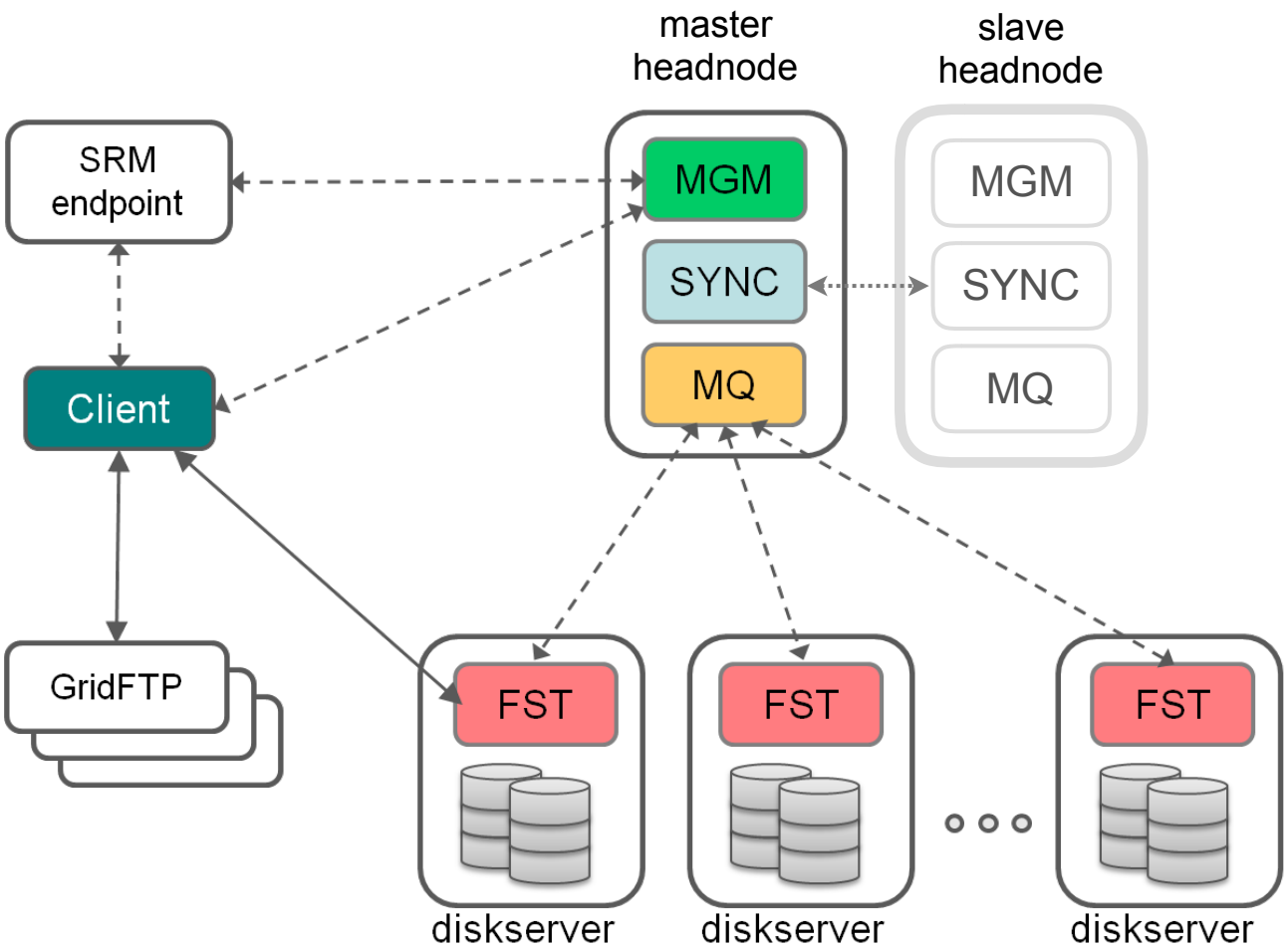
Message Queue
Service State Messages
File Transaction Reports
Shared Objects (queue+hash)

File Storage
File & File Meta Data Store
Capability Authorization
Check-summing & Verification
Disk Error Detection (Scrubbing)



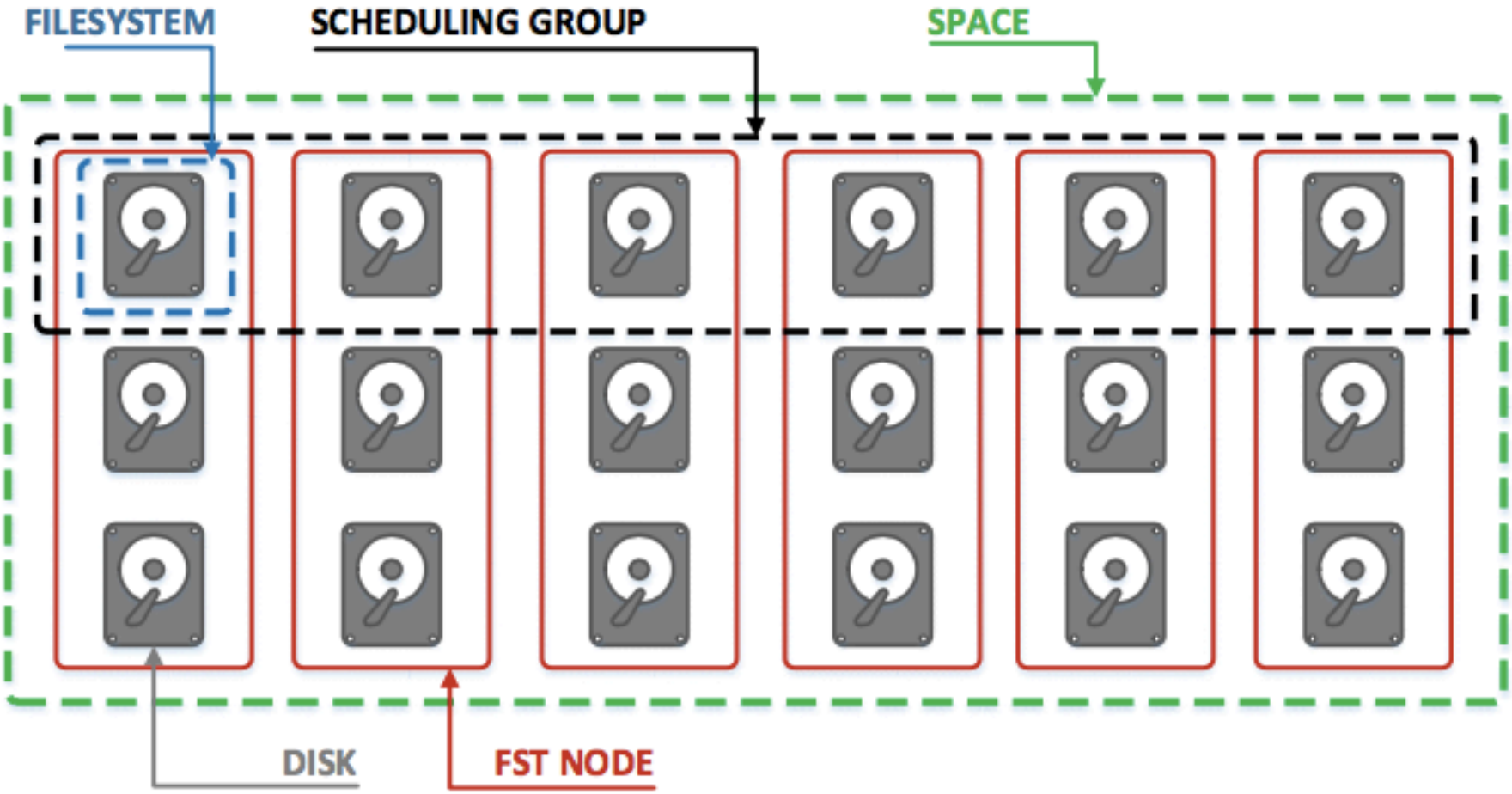
Deployment View



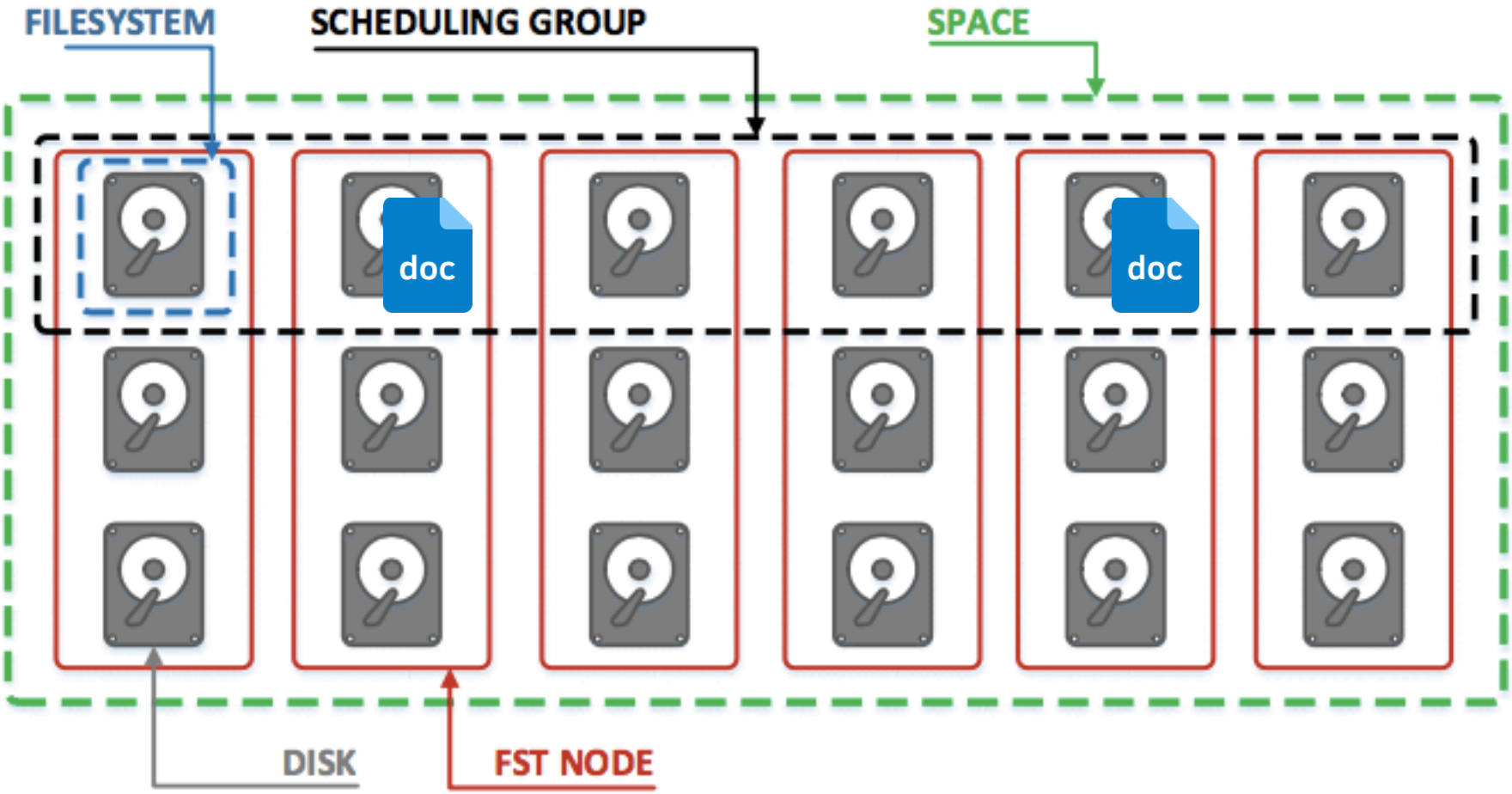


Functionality

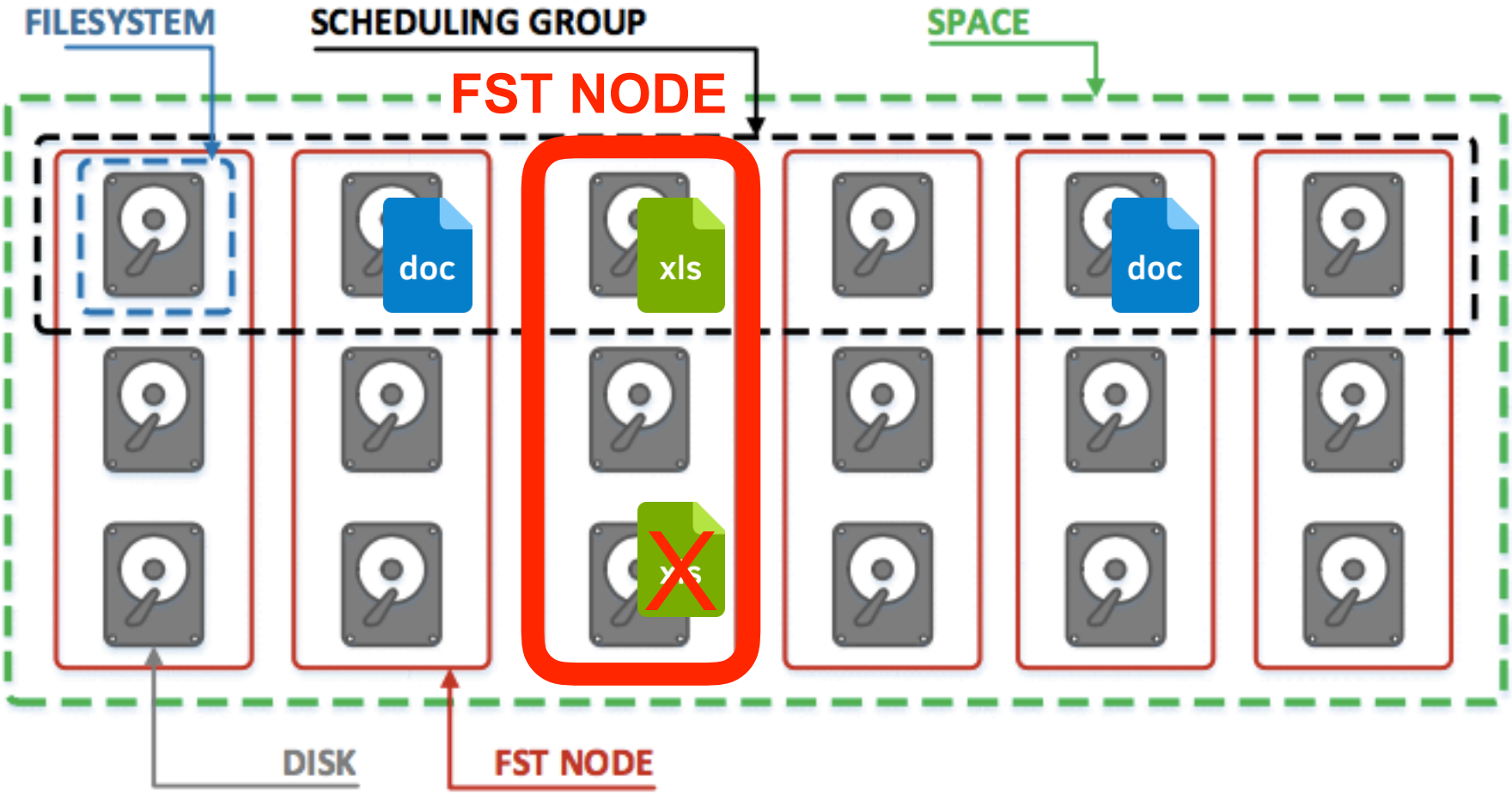
Data Placement



Data Placement



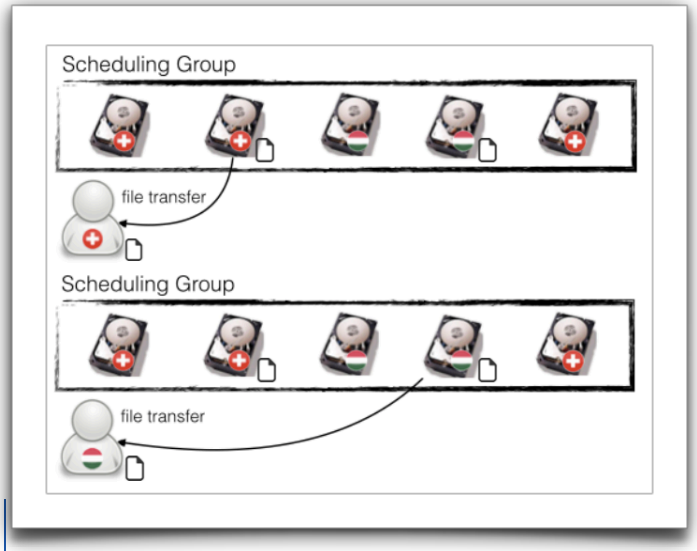
Data Placement

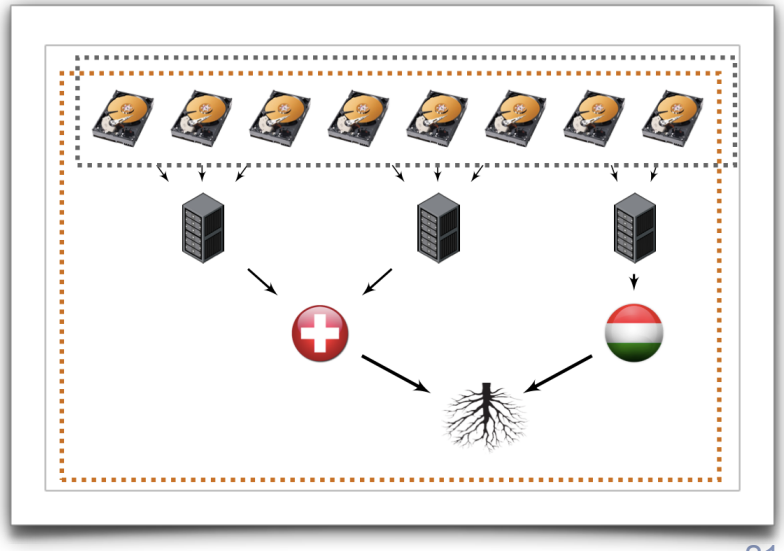


EOS Infrastructure Awareness

- Flat scheduling
  - stateless
  - work only with 2 locations
  - single scheduling policies
  - possible CPOF (e.g. switch)

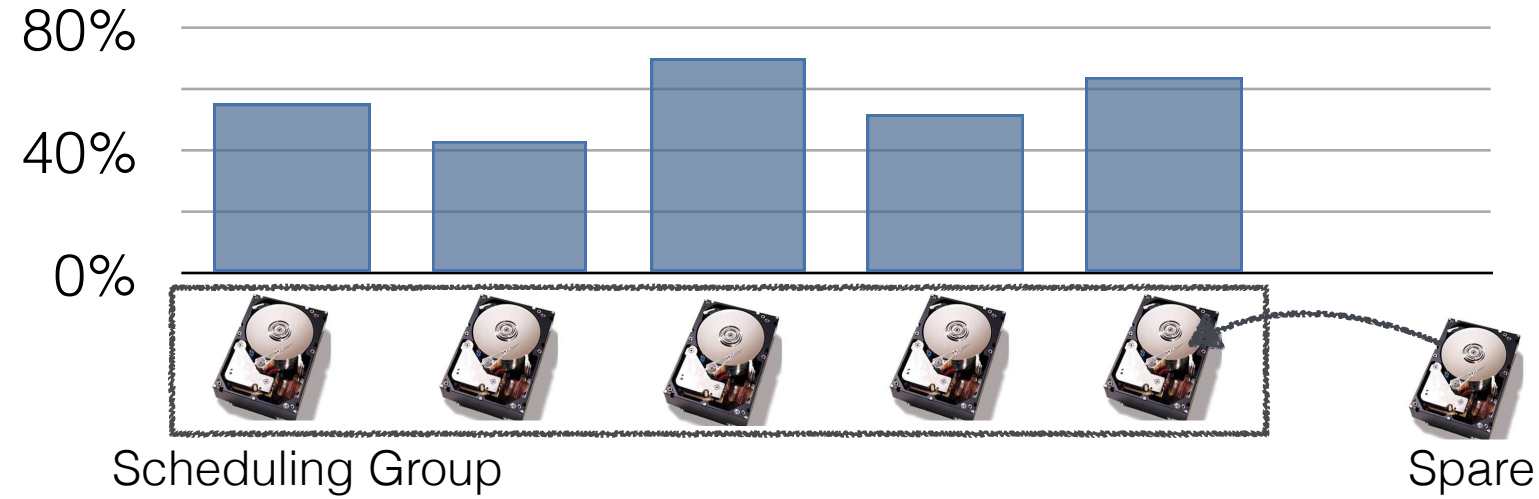
- Tree scheduling
  - stateful (more complex)
  - tree efficient data structures
  - more scheduling policies available
  - avoid CPOF

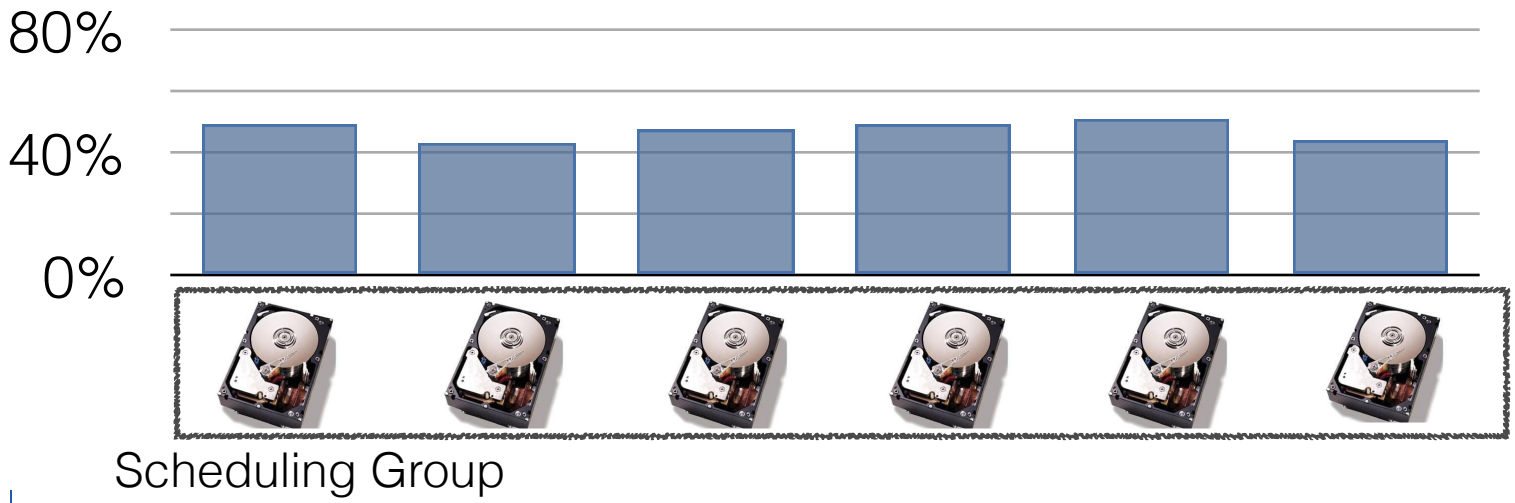




Balancing

```
/> eos space config <space> space.balancer=on
```

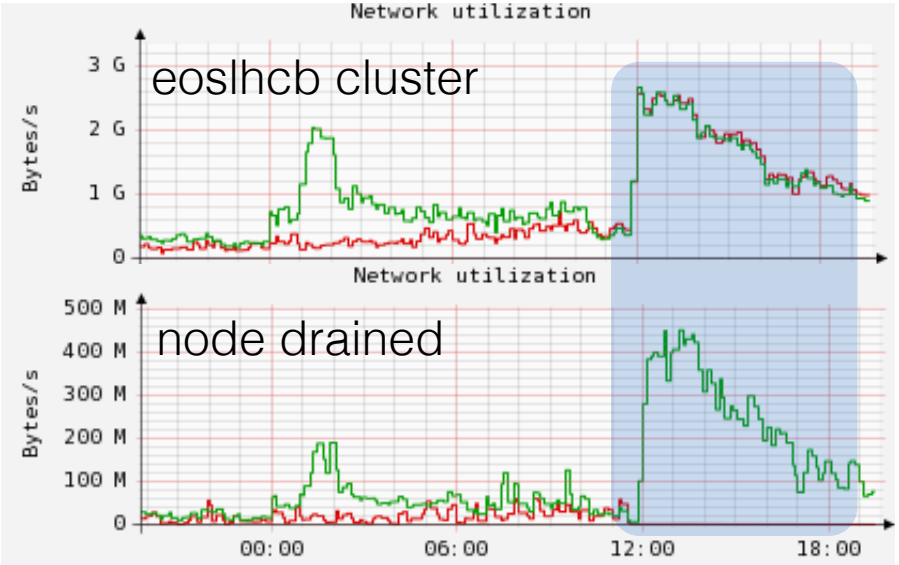




http://eos.readthedocs.io/en/latest/configuration/balancing.html

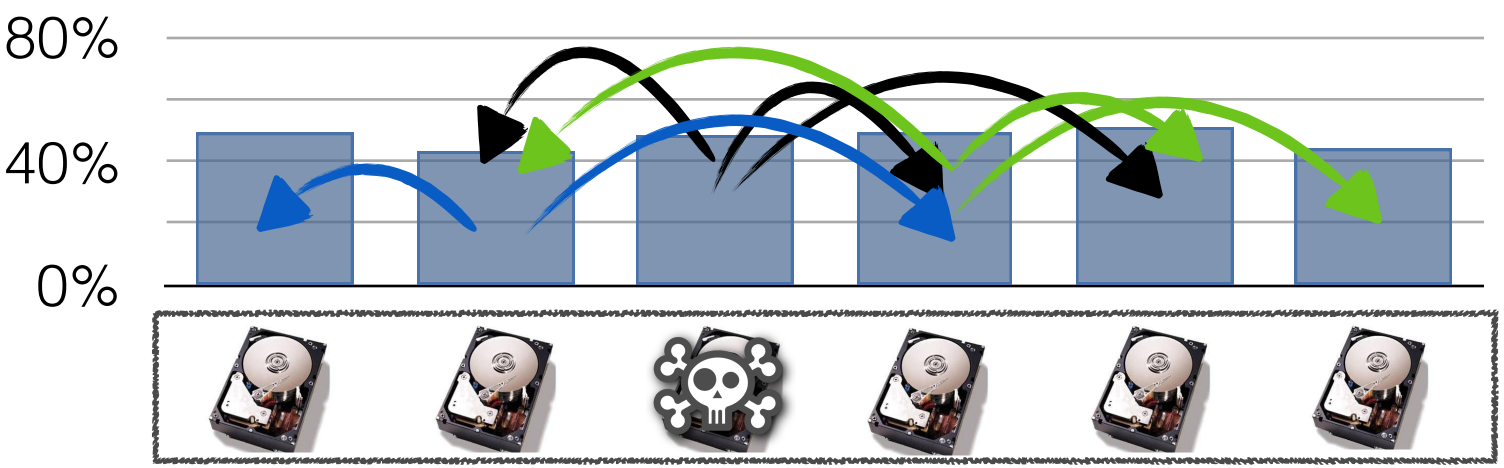
Draining

```
/> eos fs config <fsid> configstatus=drain
/> eos node config <node> configstatus=drain
```




```
[root@eoslhcb-srv-m2 (mgm:master mq:master) ~]$
[root@eoslhcb-srv-m2 (mgm:master mq:master) ~]$ eos fs ls -d

#.....................................................................
#                  host (#...) #   id #          path #      drain #
#.....................................................................
 p05153074058585.cern.ch (1095)   3536          /data16      drained
 p05153074065039.cern.ch (1095)   3621          /data11      drained
 p05153074065039.cern.ch (1095)   3658          /data21      drained
 p05153074065039.cern.ch (1095)   3669          /data25      drained
     lxfsrd46c01.cern.ch (1095)   3858          /data01      drained
     lxfsrd46c01.cern.ch (1095)   3859          /data02      drained
     lxfsrd46c01.cern.ch (1095)   3860          /data03      drained
     lxfsrd46c01.cern.ch (1095)   3861          /data04      drained
     lxfsrd46c01.cern.ch (1095)   3862          /data05      drained
     lxfsrd46c01.cern.ch (1095)   3863          /data06      drained
     lxfsrd46c01.cern.ch (1095)   3864          /data07      drained
     lxfsrd46c01.cern.ch (1095)   3865          /data08      drained
     lxfsrd46c01.cern.ch (1095)   3866          /data09      drained
     lxfsrd46c01.cern.ch (1095)   3867          /data10      drained
     lxfsrd46c01.cern.ch (1095)   3868          /data11      drained
     lxfsrd46c01.cern.ch (1095)   3869          /data12      drained
     lxfsrd46c01.cern.ch (1095)   3870          /data13      drained
     lxfsrd46c01.cern.ch (1095)   3871          /data14      drained
```

http://eos.readthedocs.io/en/latest/configuration/draining.html



Creating an EOS Cluster

EOS Software & Dependencies

Software Repositories

- **Production Branch**
  - Aquamarine https://dss-ci-repo.web.cern.ch/dss-ci-repo/eos/aquamarine/tag/
  - Aquamarine-Dependencies https://dss-ci-repo.web.cern.ch/dss-ci-repo/eos/aquamarine-depend/
  - Citrine https://dss-ci-repo.web.cern.ch/dss-ci-repo/eos/citrine/tag/
  - Citrine-Dependencies https://dss-ci-repo.web.cern.ch/dss-ci-repo/eos/citrine-depend/
- **Development Branch**
  - Aquamarine-Dev https://dss-ci-repo.web.cern.ch/dss-ci-repo/eos/aquamarine/commit/
  - Citrine-Dev https://dss-ci-repo.web.cern.ch/dss-ci-repo/eos/citrine/commit/

Code

- **Gitlab**
  https://gitlab.cern.ch/dss/eos.git
- **Github**
  https://github.com/cern-eos/eos

CERN & COMTRADE Collaboration

- OpenLab Collaboration Project
  - EOS Whitepaper
  - EOS beginner setup

- You can download the material here:

https://cernbox.cern.ch/cernbox/desktop/index.php/s/Nbpi3hGZYYqHN93

Setup MGM

Location	Significance
/etc/xrd.cf.mgm	xrootd server configuration file
/etc/sysconfig/eos	instance configuration file
/var/eos/md/files.<hostname>.mdlog	changelog file cont. file meta data of the name space
/var/eos/md/directories.<hostname>.mdlog	changelog file cont. directory meta data of the name space
/var/eos/config/<hostname>	directory cont. configuration files and configuration changelog
/etc/eos.keytab	keytab file for 'sss' authentication (shared secret)
/var/log/eos/mgm/xrdlog.mgm	eos/xrootd MGM log file
/var/eos/md/so.mgm.dump	Dump of the MGM shared object hash/queues

- How to generate a keytab
  - xrdsssadmin -k <eosinstance> -u daemon -g daemon add /etc/eos.keytab

- RPM needed
  - yum install xrootd-client-libs-3.3.6-4.CERN.slc6.x86_64 xrootd-client-3.3.6-4.CERN.slc6.x86_64 xrootd-debuginfo-3.3.6-4.CERN.slc6.x86_64 xrootd-3.3.6-4.CERN.slc6.x86_64 xrootd-libs-3.3.6-4.CERN.slc6.x86_64 xrootd-server-libs-3.3.6-4.CERN.slc6.x86_64
  - yum install leveldb-1.10.0-0.el6.x86_64 zeromq eos-server eos-client eos-debuginfo libmicrohttpd-0.9.38-eos.wves.slc6.x86_64 libmicrohttpd-debuginfo-0.9.38-eos.wves.slc6.x86_64

Setup MGM

Back them up!!!

Location	Significance
/etc/xrd.cf.mgm	xrootd server configuration file
/etc/sysconfig/eos	instance configuration file
/var/eos/md/files.<hostname>.mdlog	changelog file cont. file meta data of the name space
/var/eos/md/directories.<hostname>.mdlog	changelog file cont. directory meta data of the name space
/var/eos/config/<hostname>	directory cont. configuration files and configuration changelog
/etc/eos.keytab	keytab file for 'sss' authentication (shared secret)
/var/log/eos/mgm/xrdlog.mgm	eos/xrootd MGM log file
/var/eos/md/so.mgm.dump	Dump of the MGM shared object hash/queues

- How to generate a keytab
  - xrdsssadmin -k <eosinstance> -u daemon -g daemon add /etc/eos.keytab
- RPM needed
  - yum install xrootd-client-libs-3.3.6-4.CERN.slc6.x86_64 xrootd-client-3.3.6-4.CERN.slc6.x86_64 xrootd-debuginfo-3.3.6-4.CERN.slc6.x86_64 xrootd-3.3.6-4.CERN.slc6.x86_64 xrootd-libs-3.3.6-4.CERN.slc6.x86_64 xrootd-server-libs-3.3.6-4.CERN.slc6.x86_64
  - yum install leveldb-1.10.0-0.el6.x86_64 zeromq eos-server eos-client eos-debuginfo libmicrohttpd-0.9.38-eos.wves.slc6.x86_64 libmicrohttpd-debuginfo-0.9.38-eos.wves.slc6.x86_64

Setup MQ and Sync

Location	Significance
/etc/xrd.cf.mq	xrootd server configuration file
/etc/sysconfig/eos	instance configuration file
/var/log/eos/mq/xrdlog.mq	eos/xrootd MQ log file
/var/log/eos/mq/proc/stats	current broker statistics file

- Sync Daemon
  - keep in sync the ns files between headnodes
    - /var/eos/md/files.<servername>.mdlog
    - /var/eos/md/directories.<servername>.mdlog
  - keep in sync the configuration of the instance
    - /var/eos/config/<servername>/*.eoscf

Headnode Setup

```
[root@eosuser-srv-b2 (mgm:master mq:master) ~]$ df
Filesystem      1K-blocks      Used  Available Use% Mounted on
/dev/sda2        20642428   2845112   16748740  15% /
tmpfs           264585112         0  264585112   0% /dev/shm
/dev/sda1         1032088    137596     842064  15% /boot
/dev/sda3        20642428    178756   19415096   1% /tmp
/dev/sda4         2064208    129256    1830096   7% /usr/vice/cache
/dev/sda5      2839366856 411696920 2283438412  16% /var
/dev/sdb1       922459044 376630264  498970592  44% /var/eos/md
AFS                9000000         0    9000000   0% /afs
[root@eosuser-srv-b2 (mgm:master mq:master) ~]$
[root@eosuser-srv-b2 (mgm:master mq:master) ~]$ free -m
             total       used       free     shared    buffers     cach
Mem:        516767     508532       8235          4        738     2306
-/+ buffers/cache:     277188     239579
Swap:            0          0          0
[root@eosuser-srv-b2 (mgm:master mq:master) ~]$
[root@eosuser-srv-b2 (mgm:master mq:master) ~]$
```

RAID-1 HDD

RAID-1 SSD

Faster boot in case of crash and the ns files are not in the server cache

RAM provision ~ 1KB per NS entry

- MGM and MQ Daemon
  - First run
    - service eos master mgm
    - service eos master mq
  - Fix /etc/hosts (remove ::1)
  - 127.0.0.1 localhost.localdomain localhost

Headnode Setup

- Headnode Daemons
  - service eos start mgm mq sync
  - service eossync start

- Some IPTABLES ;)

```
-A INPUT -p tcp -m multiport --dports 1094 -m comment --comment "101 allow xroot to EOS-MGM" -j ACCEPT
-A INPUT -p tcp -m multiport --dports 1097 -m comment --comment "101 allow xroot to EOS-MQ" -j ACCEPT
-A INPUT -p tcp -m multiport --dports 1096 -m comment --comment "101 allow xroot to EOS-SYNC" -j ACCEPT
-A INPUT -p tcp -m multiport --dports 1098 -m comment --comment "101.1 allow stuff EOS-FED" -j ACCEPT
-A INPUT -p tcp -m multiport --dports 443 -m comment --comment "102 allow GSI-HTTP to EOS-MGM" -j ACCEPT
-A INPUT -p tcp -m multiport --dports 8000 -m comment --comment "102 allow HTTP to EOS-MGM" -j ACCEPT
-A INPUT -p tcp -m multiport --dports 8443 -m comment --comment "102 allow KRB5-HTTP to EOS-MGM" -j ACCEPT
-A INPUT -p tcp -m multiport --dports 1095 -m comment --comment "101 allow xroot to EOS-FST" -j ACCEPT
-A INPUT -p tcp -m multiport --dports 8001 -m comment --comment "102 allow HTTP to EOS-FST" -j ACCEPT
```

Setup FST

Location	Significance
/etc/xrd.cf.fst	xrootd server configuration file
/etc/sysconfig/eos	instance configuration file
/var/log/eos/fmd.<unix-tst>.<fsid>.mdlog	FST meta data changelog file
/var/log/eos/so.fst.dump	Dump of the FST shared object hash/queues

- FST Daemon
  - service eos start fst

- Other IPTABLES ;)
  -A INPUT -p tcp -m multiport --dports 1095 -m comment --comment "101 allow xroot to EOS-FST" -j ACCEPT
  -A INPUT -p tcp -m multiport --dports 8001 -m comment --comment "102 allow HTTP to EOS-FST" -j ACCEPT

- Register the FST back to the EOS instance
  from the MGM
  eos fs add `uuidgen` <host.domain>:1095 <disk> <pool>
  from the FST
  eosfstregister <disk> <pool>

Configuration

Layout (Replication)

```
....................
....... Layouts ...
....................
- set 2 replica as standard layout ...
     |eos> attr set default=replica  /eos/instance/2-replica
-------------------------------------------------------------
- set RAID-6 4+2 as standard layout ...
     |eos> attr set default=raid6 /eos/instance/raid-6
-------------------------------------------------------------
- set ARCHIVE 5+3 as standard layout ...
     |eos> attr set default=archive /eos/instance/archive
-------------------------------------------------------------
- re-configure a layout for different number of stripes (e.g. 10)
     |eos> attr set sys.forced.stripes=10 /eos/instance/archive
```

```
sys.acl="u:lmascett:rwx!m"
sys.allow.oc.sync="1"
sys.forced.atomic="1"
sys.forced.blockchecksum="crc32c"
sys.forced.blocksize="4k"
sys.forced.checksum="adler"
sys.forced.layout="replica"
sys.forced.maximumsize="10000000000"
sys.forced.maxsize="10000000000"
sys.forced.nstripes="2"
sys.forced.space="default"
sys.mask="700"
sys.mtime.propagation="1"
sys.owner.auth="*"
sys.recycle="/eos/user/proc/recycle/"
sys.versioning="10"
```

- Layout settings are per directories
  - Single files can be changed with a converter
    - eos file convert <path> [<layout>:<stripes>] [targetspace]

Authentication and Vitual IDs

```
[root@eosuser-srv-b2 (mgm:master mq:master) ~]$ eos vid help
usage: vid ls [-u] [-g] [-s] [-U] [-G] [-g] [-a] [-l] [-n]                                  : list configured policies
                                        -u : show only user role mappings
                                        -g : show only group role mappings
                                        -s : show list of sudoers
                                        -U : show user  alias mapping
                                        -G : show group alias mapping
                                        -y : show configured gateways
                                        -a : show authentication
                                        -l : show geo location mapping
                                        -n : show numerical ids instead of user/group names

       vid set membership <uid> -uids [<uid1>,<uid2>,...]
       vid set membership <uid> -gids [<gid1>,<gid2>,...]
       vid rm membership <uid>             : delete the membership entries for <uid>.
       vid set membership <uid> [+|-]sudo
       vid set map -krb5|-gsi|-https|-sss|-unix|-tident|-voms <pattern> [vuid:<uid>] [vgid:<gid>]

                                                                                     -voms <pattern>  : <pattern> is <group>:<role>
       vid set geotag <IP-prefix> <geotag>  : add to all IP's matching the prefix <prefix> the geo location tag <geotag>
                                              N.B. specify the default assumption via 'vid set geotag default <default-tag>'
       vid rm <key>                         : remove configured vid with name key - hint: use config dump to see the key names of vid rules

       vid enable|disable krb5|gsi|sss|unix|https
                                            : enable/disables the default mapping via password database

       vid add|remove gateway <hostname> [krb5|gsi|sss|unix|https]
                                            : adds/removes a host as a (fuse) gateway with 'su' priviledges
                                              [<prot>] restricts the gateway role change to the specified authentication method
[root@eosuser-srv-b2 (mgm:master mq:master) ~]$
```

```
config dump [--fs|-f] [--vid|-v] [--quota|-q] [--policy|-p] [--comment|-c] [--global|-g] [--access|-a] [<name>] [--map|-m]] :
                                                dump current configuration or configuration with name <name>
            -f : dump only file system config
            -v : dump only virtual id config
            -q : dump only quota config
            -p : dump only policy config
            -g : dump only global config
            -a : dump only access config
            -m : dump only mapping config
```

Pools and Scheduling Groups

- eos space define [groupsize] [groupmod]
- eos space ls

- eos group set <group> on/off
- eos group ls

- eos node ls
- eos node ls —sys
- eos node ls —io

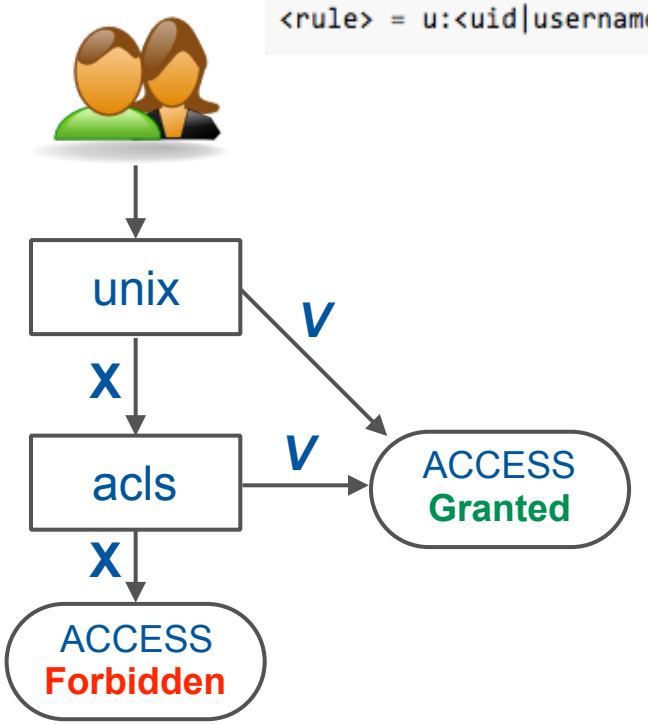
Draining & Balancing

- Settings per space and per storage node

```
[root@eosuser-srv-b2 (mgm:master mq:master) ~]$ eos space help
usage: space ls                                                  : list spaces
usage: space ls [-s] [-m|-l|--io|--fsck] [<space>]                 : list in all spaces or select only <space>. <s
                                                                  -s : silent mode
                                                                  -m : monitoring key=value output format
                                                                  -l : long output - list also file systems after ea
                                                                --io : print IO satistics
                                                              --fsck : print filesystem check statistics
       space config <space-name> space.nominalsize=<value>         : configure the nominal size for this space
       space config <space-name> space.balancer=on|off             : enable/disable the space balancer [default=of
       space config <space-name> space.balancer.threshold=<percent> : configure the used bytes deviation which trig
       space config <space-name> space.balancer.node.rate=<MB/s>   : configure the nominal transfer bandwith per r
       space config <space-name> space.balancer.node.ntx=<#>       : configure the number of parallel balancing tr
       space config <space-name> space.converter=on|off            : enable/disable the space converter [default=o
       space config <space-name> space.converter.ntx=<#>           : configure the number of parallel conversions
       space config <space-name> space.drainer.node.rate=<MB/s >   : configure the nominal transfer bandwith per r
       space config <space-name> space.drainer.node.ntx=<#>        : configure the number of parallel draining tra
       space config <space-name> space.lru=on|off                  : enable/disable the LRU policy engine [default
       space config <space-name> space.lru.interval=<sec>          : configure the default lru scan interval
       space config <space-name> space.headroom=<size>             : configure the default disk headroom if not de
       space config <space-name> space.scaninterval=<sec>          : configure the default scan interval if not de
       space config <space-name> space.drainperiod=<sec>           : configure the default drain  period if not de
       space config <space-name> space.graceperiod=<sec>           : configure the default grace  period if not de
       space config <space-name> space.autorepair=on|off           : enable auto-repair of faulty replica's/files
       space config <space-name> space.geo.access.policy.write.exact=on|off   : if 'on' use exact matching geo repli
       space config <space-name> space.geo.access.policy.read.exact=on|off    : if 'on' use exact matching geo repli

       space config <space-name> fs.<key>=<value>                  : configure file system parameters for each fil

       space define <space-name> [<groupsize> [<groupmod>]]        : define how many filesystems can end up in one

                                                                     => <groupsize>=0 means, that no groups are bu
                                                                     => <groupmod> defines the maximun number of f

       space reset <space-name>  [--egroup|mapping|drain|scheduledrain|schedulebalance|ns|nsfilesystemview|nsfilemap
                                                                   : reset a space e.g. recompute the drain state
       space status <space-name> [-m]                              : print's all defined variables for space
       space set <space-name> on|off                               : enables/disabels all groups under that space
       space rm <space-name>                                       : remove space

       space quota <space-name> on|off                             : enable/disable quota
[root@eosuser-srv-b2 (mgm:master mq:master) ~]$
```

Permission System

- Permission are verified at directory level
  - 2-step permissions:
    - unix bits + additional acls




```
<rule> = u:<uid|username>|g:<gid|groupname>|egroup:<name>:{rwxomqc(!d)(+d)(!u)(+u)}
```

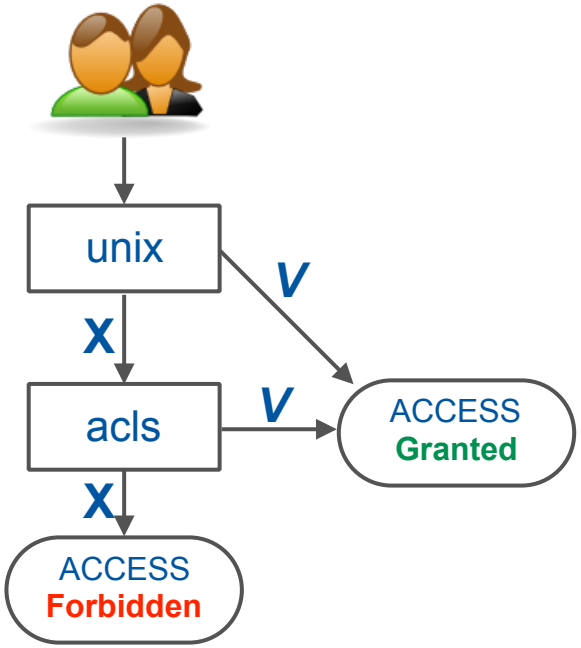
tag	definition
r	grant read permission
w	grant write permissions
x	grant browsing permission
m	grant change mode permission
!m	forbid change mode operations
!d	forbid deletion of files and directories
+d	overwrite a '!d' rule and allow deletion of files and directories
!u	forbid update of files
+u	overwrite a '!u' rule and allow updates for files
q	grant 'set quota' permissions on a quota node
c	grant 'change owner' permission on directory children

Permission System

- Admin and Users ACLs
  - sys.eval.useracl=true
  - sys and user acls are concatenated and evaluated

```
sys.acl=<acllist>

user.acl=<acllist>
```

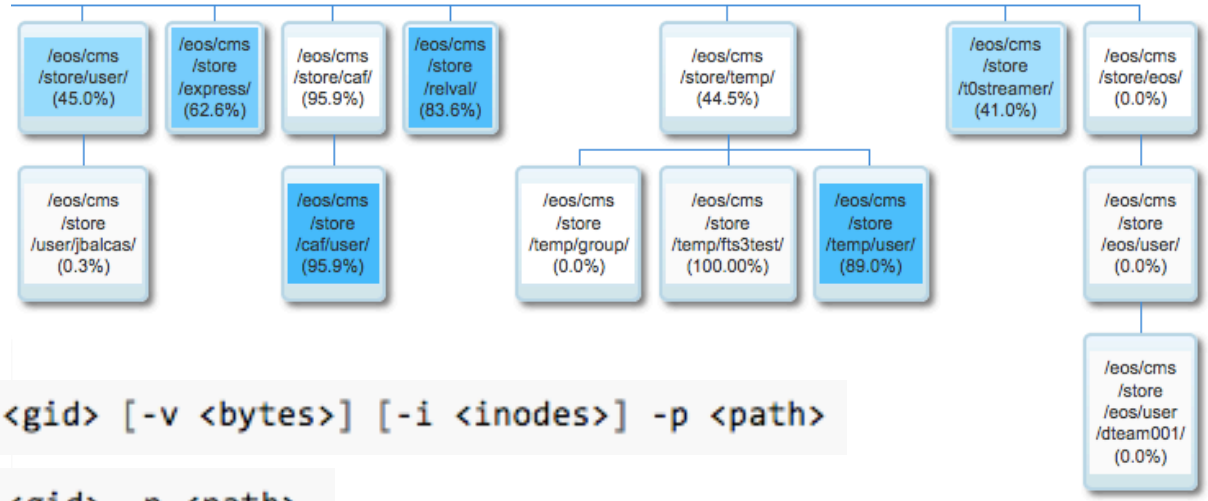
```
<acllist> = <rule1>,<rule2>...<ruleN>
```



http://eos.readthedocs.io/en/latest/configuration/permission.html

Quota system

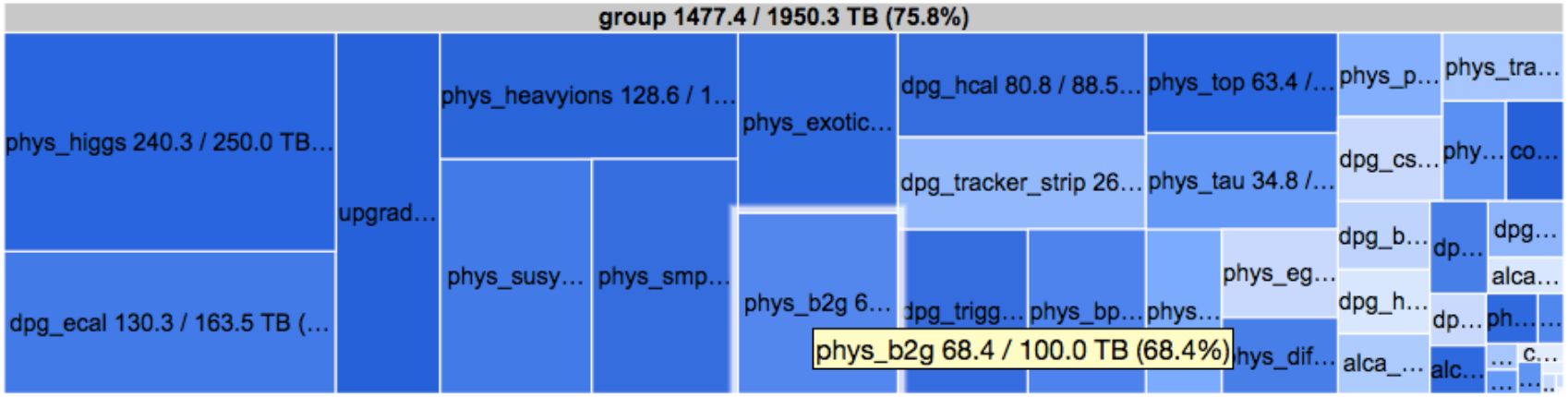
- User
- Group
- Project




```
eos quota set -u <uid>|-g <gid> [-v <bytes>] [-i <inodes>] -p <path>
```

```
eos quota rm -u <uid> |-g <gid> -p <path>
```

```
eos quota rmnode -p <path>
```



http://eos.readthedocs.io/en/latest/configuration/quota.html

Recycle Bin

- FIFO policy for delayed file deletion
  - time and volume based

```
EOS Console [root://localhost] |/eos/> recycle config --lifetime 86400
```

```
EOS Console [root://localhost] |/eos/> recycle config --size 100G
```

```
EOS Console [root://localhost] |/eos/dev/2rep/subnode/> recycle config --add-bin /eos/dev/
```

```
EOS Console [root://localhost] |/eos/> recycle

# ____________________________________________________________________________________________
# used 0.00 B out of 100.00 GB (0.00% volume / 0.00% inodes used) Object-Lifetime 86400 [s]
# ____________________________________________________________________________________________
```

Recycle Bin

Listing recycle entries

```
EOS Console [root://localhost] |/eos/dev/2rep/subnode/> recycle ls
# Deletion Time            UID      GID      TYPE          RESTORE-KEY      RESTORE-PATH
# ==========================================================================================
Thu Mar 21 23:02:22 2013   apeters  z2       recursive-dir 00000000000007cf /eos/dev/2rep/subnode/tree
```

restoring recycle entries

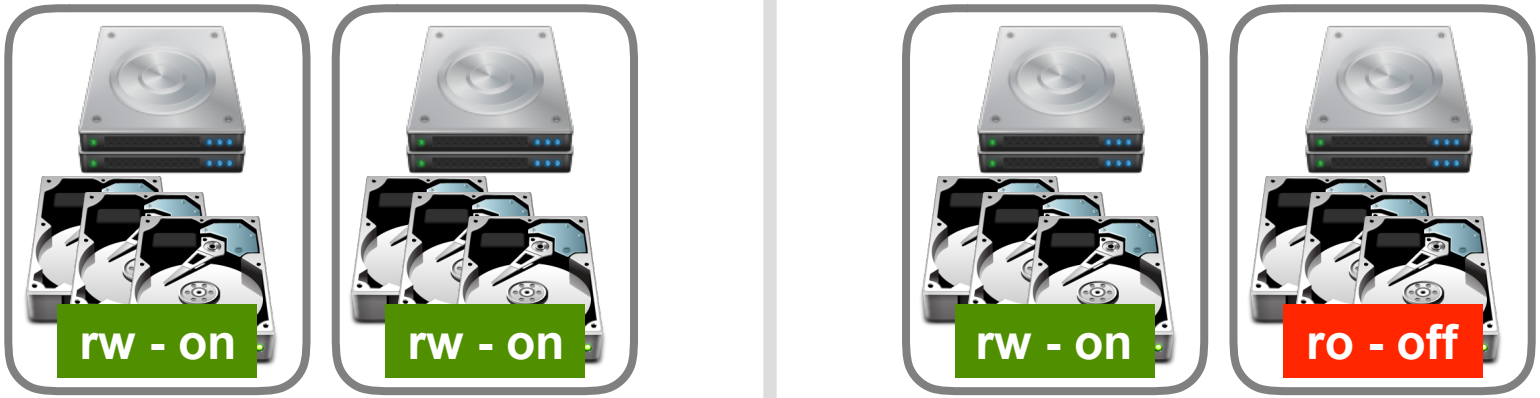
```
EOS Console [root://localhost] |/eos/dev/2rep/subnode/> recycle restore -f 00000000000007cf
warning: renamed restore path=/eos/dev/2rep/subnode/tree to backup-path=/eos/dev/2rep/subnode/tree.000000006
success: restored path=/eos/dev/2rep/subnode/tree
```

emptying completely recycle entries

```
EOS Console [root://localhost] |/eos/dev/2rep/subnode/> recycle purge
success: purged 1 bulk deletions and 0 individual files from the recycle bin!
```

Node Operations

How To: Intervention on Storage

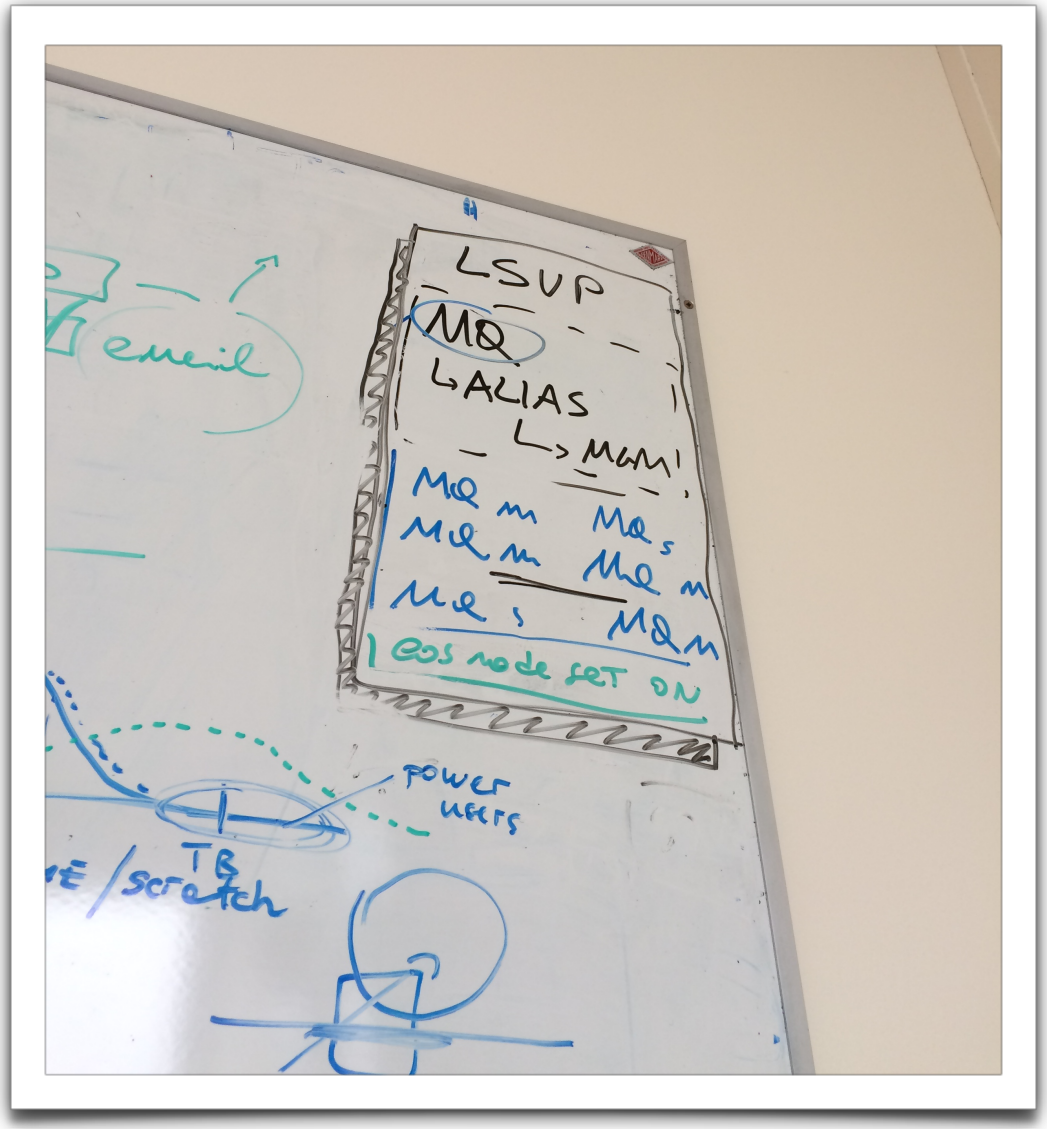


- eos node <node> set off
- eos node <node> config configstatus=ro|off

- wait for clients to go away
  - eos node ls —io

- perform the intervention
- set the machine back online

How To: M/S Failover



How To: M/S Failover

- Most Delicate Operation
  - Make sure everything is correctly sync
  - Failover the MQ
  - Change the alias
  - Failover the MGM

Master/Slave eos shell interface

```
eos -b ns : shows the current state of slave/master MGM/MQ and the current configuration
eos -b ns master: shows the log file of any master/slave transition command including the initial boot
eos -b ns master --log-clear : clean the in-memory logfile
eos -b ns master --disable : disable the heartbeat-supervisor thread modifying redirection/stall variables
eos -b ns master --enable: enable the heartbeat-supervisor thread modifying redirection/stall variables
```

How To: M/S Failover

Bounce the MQ Service

To bounce the MQ service you should make the slave machine also to a master and then declare the other as slave e.g. to move from eosdevsrv1 to eosdevsrv2 you do

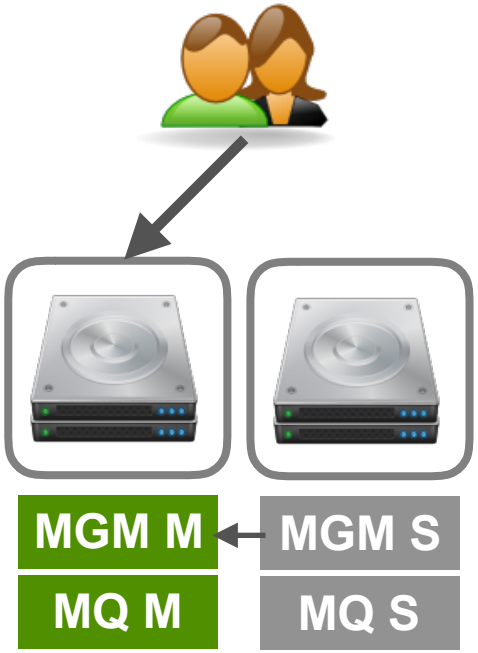
```
eosdevsrv2:# service eos master mq
eosdevsrv1:# service eos slave mq
```

It is important to never declare both machines as slaves at the same time! While it should work well if the broker alias points to both machines it is probably more efficient to use a dedicated alias for the MQ broker and always point only to one box. This has to be tested.

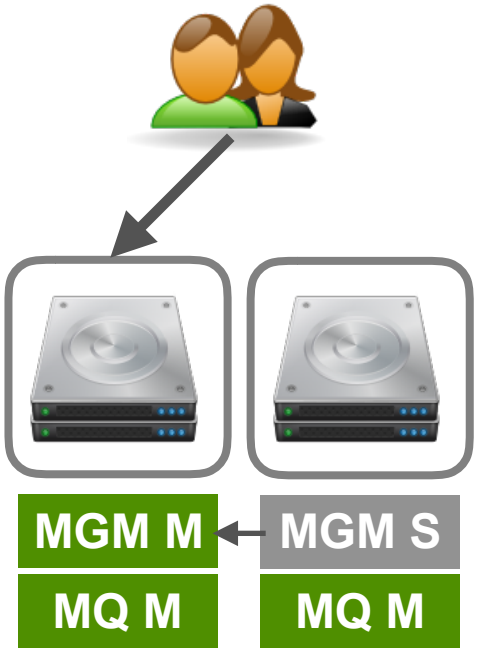
Run a master<=>slave MGM role change procedure

```
# switch the master MGM to RO mode
eosdevsrv1:# eos -b ns master eosdevsrv2.cern.ch

# switch the slave MGM to master mode
eosdevsrv2:# eos -b ns master eosdevsrv2.cern.ch

# switch the RO mode master MGM to slave mode
eosdevsrv1:# eos -b ns master eosdevsrv2.cern.ch
```

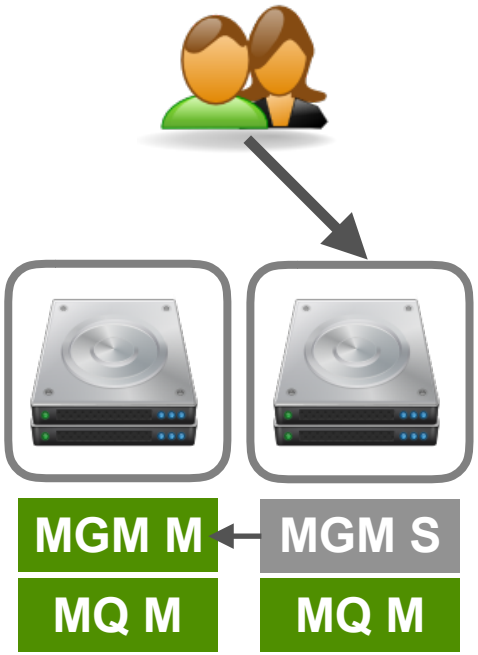
1


MGM M
MGM S
MQ M
MQ S

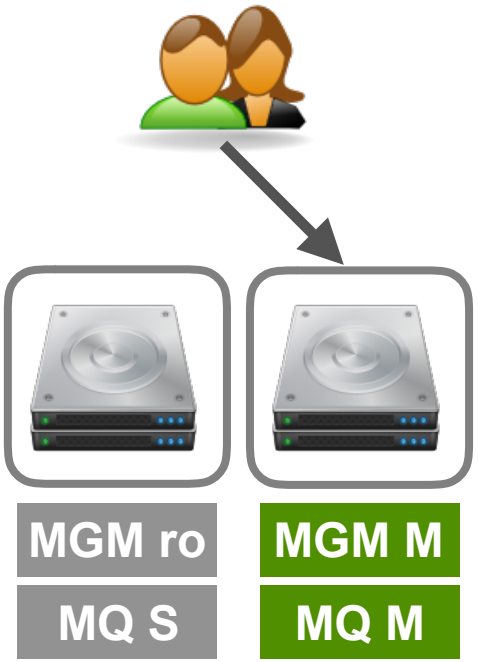
2


MGM M
MGM S
MQ M
MQ M

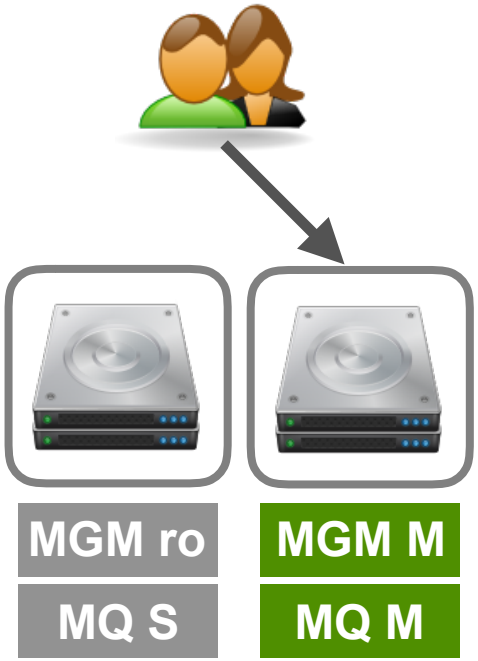
3


MGM M
MGM S
MQ M
MQ M

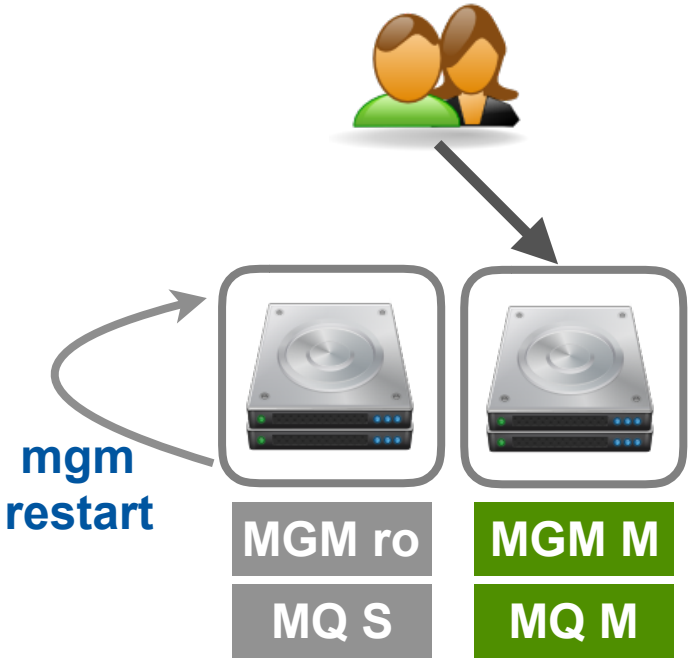
4


MGM ro
MGM M
MQ S
MQ M

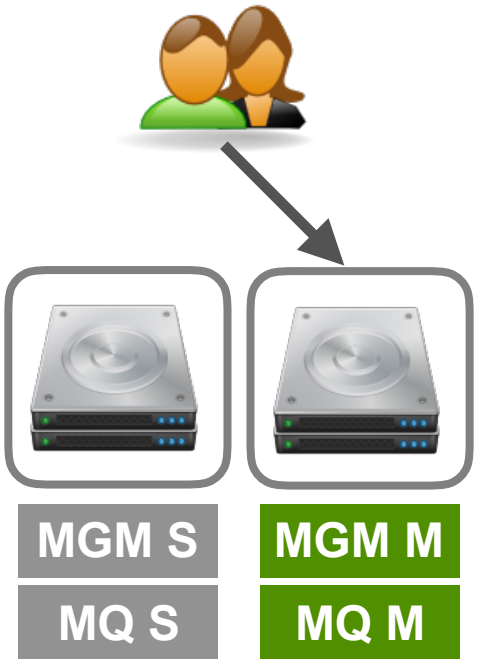
5


MGM ro
MGM M
MQ S
MQ M

6


mgm
restart
MGM ro
MGM M
MQ S
MQ M

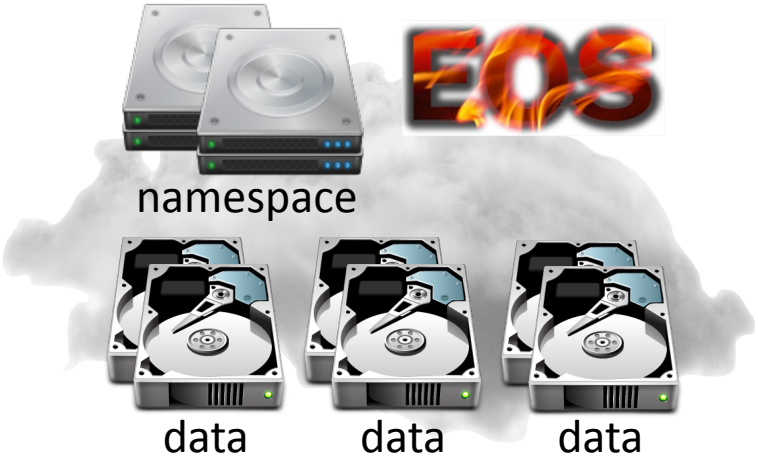
7


MGM S
MGM M
MQ S
MQ M

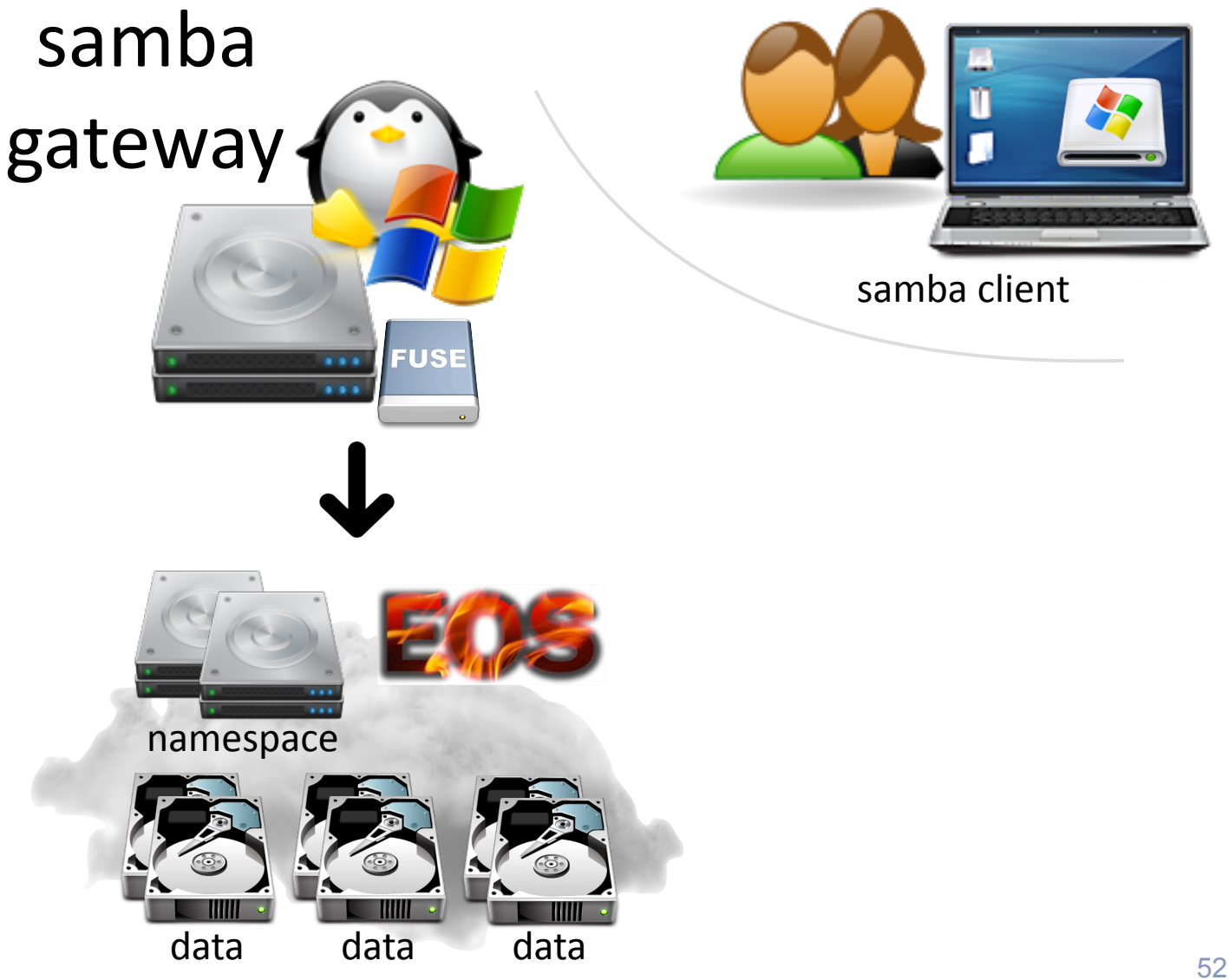
Additional Entry-point

SRM

BeStMan2

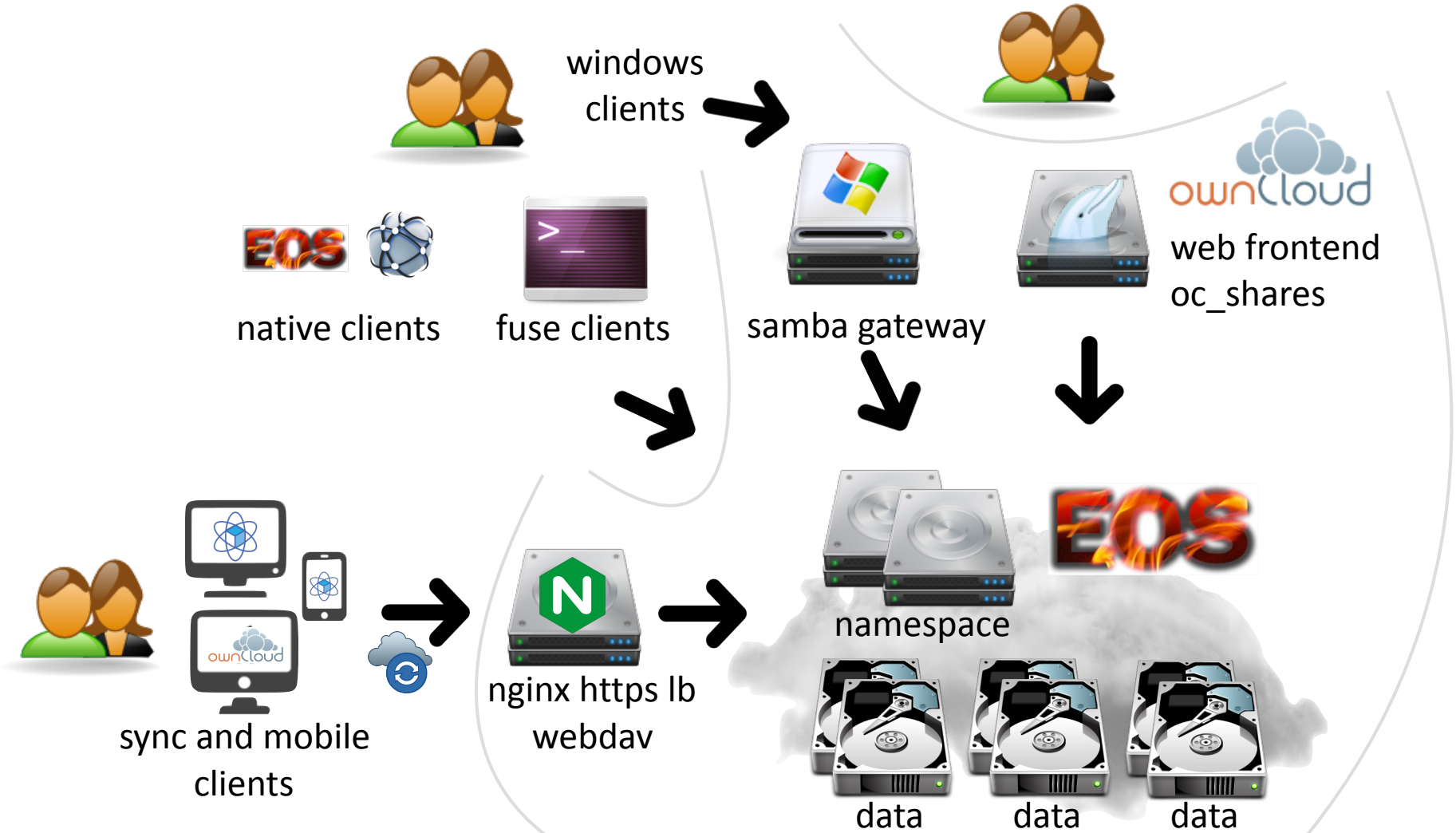


SAMBA


samba
gateway
FUSE
samba client
EOS
namespace
data
data
data

CERNBox Architecture

CERNBox Architecture



and...

rate limiting
geo scheduling policies
id mappings
user/host ban
io monitoring
https/S3 extensions
fsck
workflow engine
sticky ownership
accounting
transfers engine
fuse optimisation
sharing
intergroup balancing
gateways
versioning
and many more…

CERN

www.cern.ch