
CERN

EOS Citrine status

Elvin Sindrilaru
on behalf of the EOS team

Overview

- Release process & versions
- Current deployment status
- Main features (since the last workshop)
- What to expect next …

Release “flavours”

- **Development releases** - built for every git commit
  - “bleeding-edge” i.e. not necessarily stable
  - YUM repo
    - http://storage-ci.web.cern.ch/storage-ci/eos/citrine/commit/
- **Tag releases** - stable but not necessarily bug free 😁
  - YUM repo
    - http://storage-ci.web.cern.ch/storage-ci/eos/citrine/tag/

Citrine releases

- First release 4.0.0 - 20th March 2015
- 81 releases in the meantime

  ~ a release every two weeks

  ~ 4550 commits
- Current **production release 4.2.12**
- Requires **XRootD >= 4.8.0**

EOS deployment @CERN

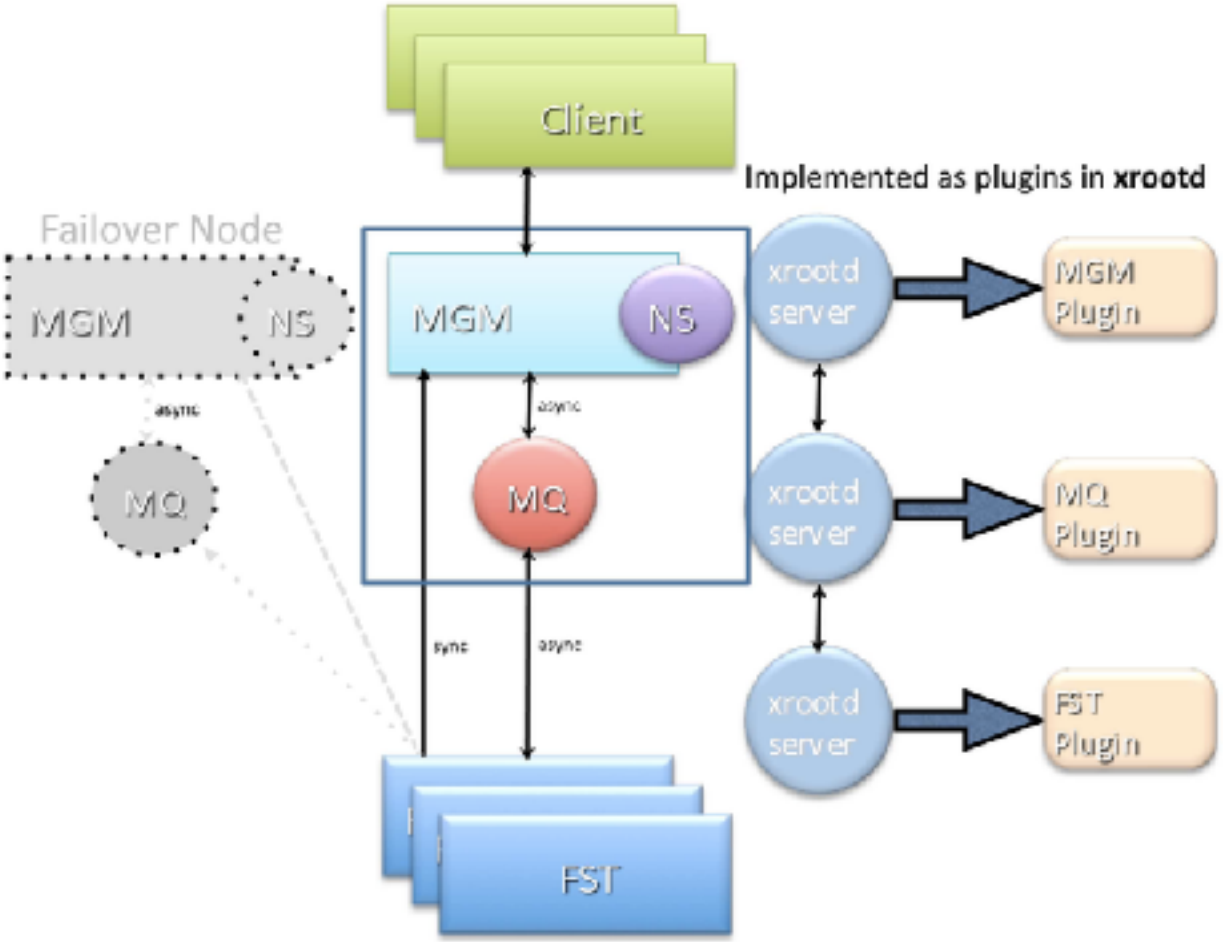
- 7/8 instances running Citrine
- EOS USER running beryl_aquamarine
- Operating System: SLC6/CentOS7

EOS architecture

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

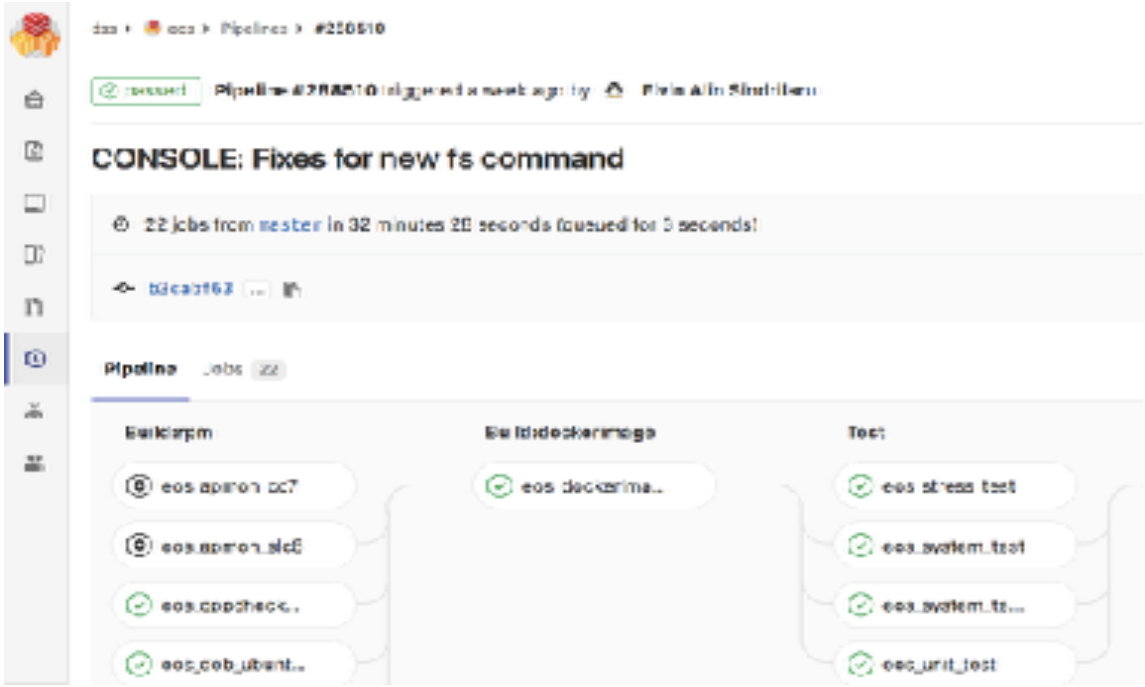
Message Queue
Service State Messages
File Transaction Reports
Shared Objects (queue+hash)

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

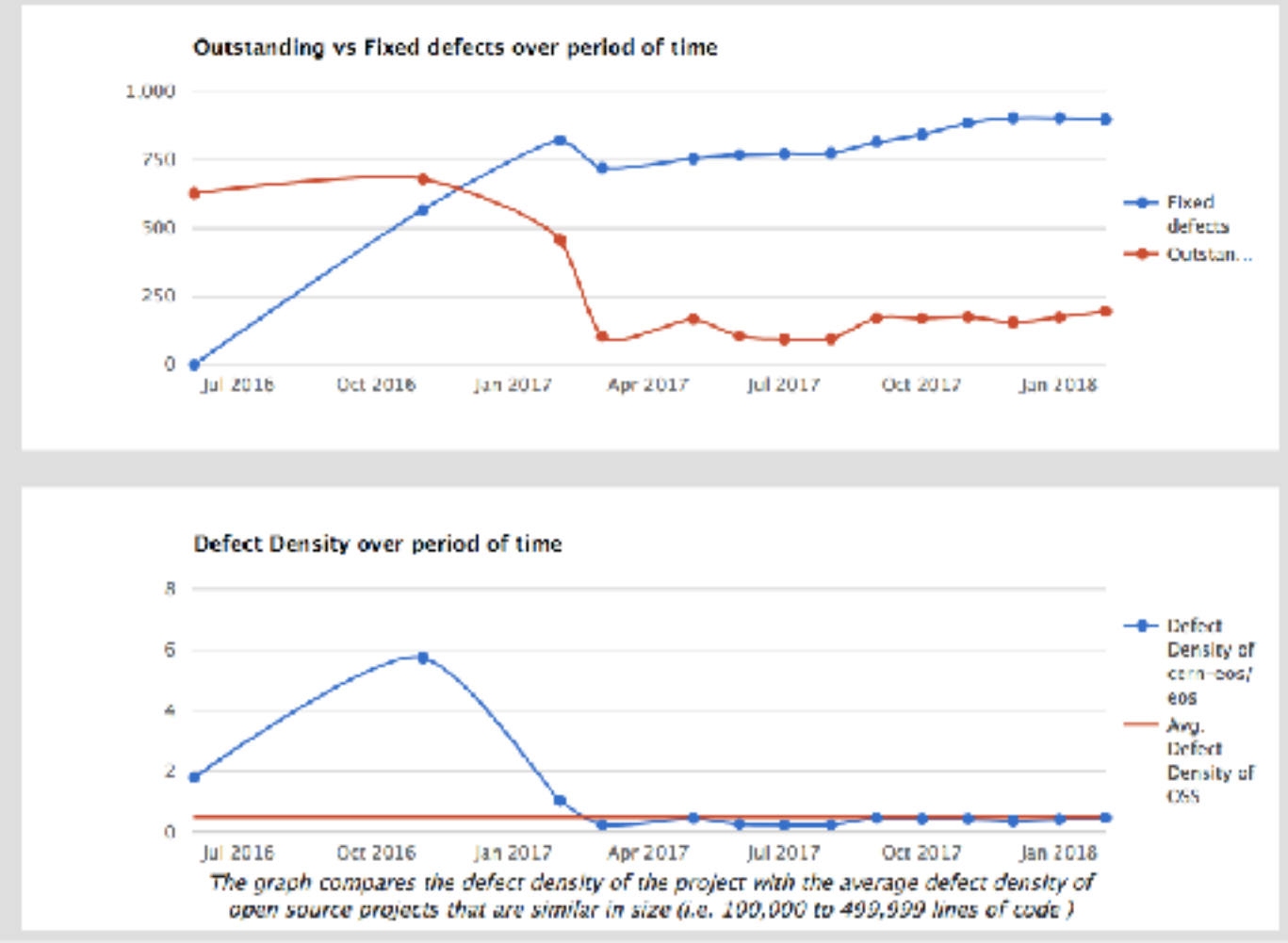


Migration from Jenkins to GitLab

- Extended/added support for new platforms i.e. **MacOS**, **Ubuntu** see Jozsef's talk
- Unit/function/system **tests** integration etc.
- **Koji** integration

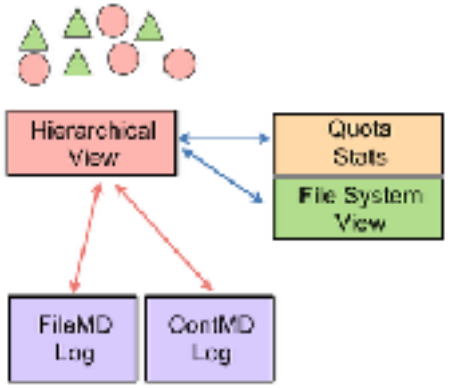


Quality improvements using Coverity



Citrine features & developments

EOS Namespace



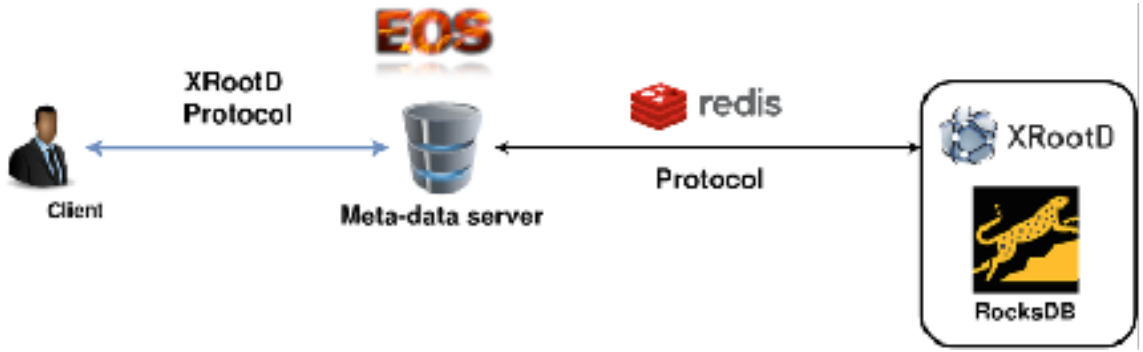
- **In-memory**
  - Speed-up booting process by
    - using **mmap** of the changelog files
    - parallelising the process

      2-6x faster
  - Consolidated the **master-slave failover**

- Fixed “sudden” file **mtime/permission changes**
  - Due to EOS internal housekeeping activities i.e. draining/balancing/group balancing etc.

EOS Namespace (KV)

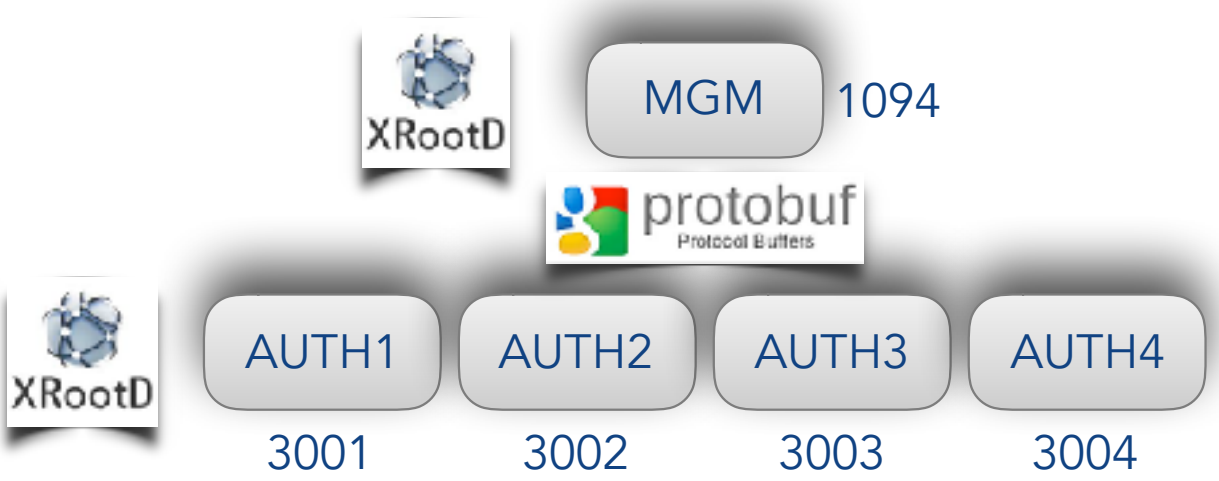
- **Conversion tool** from in-memory to new namespace implementation ~2 hours for 0.5B files
- Support for **CERNBOX specific functionality**
  - mtime propagation and sub-tree size accounting



- Namespace on top of a **key-value store (QuarkDB)**
  - See Georgio's talk https://indico.cern.ch/event/656157/contributions/2866297/

EOS Authentication Proxies

- Bottleneck on number of **XRootD GSI auth handshakes**
  => slow namespace response




```
AUTHPROXY_0  tcp  --  anywhere             anywhere             statistic mode random probability 0.250000 /* 100 Authproxy probability routing */
AUTHPROXY_1  tcp  --  anywhere             anywhere             statistic mode random probability 0.333333 /* 101 Authproxy probability routing */
AUTHPROXY_2  tcp  --  anywhere             anywhere             statistic mode random probability 0.500000 /* 102 Authproxy probability routing */
AUTHPROXY_3  tcp  --  anywhere             anywhere             statistic mode random probability 1.000000 /* 103 Authproxy probability routing */
```

Simple stateful load-balancing 1094

ACL command revisited

- Added support for **recursive & atomic** ACL modifications
- Uses similar syntax as **chmod**

```
[esindril@esdss000 ~]$ sudo eos acl -h
Usage: eos acl [-l|--list] [-R|--recursive][--sys|--user] <rule> <path>

    --help           Print help
-R, --recursive      Apply on directories recursively
-l, --lists          List ACL rules
    --user           Set user.acl rules on directory
    --sys            Set sys.acl rules on directory
<rule> is created based on chmod rules.
Every rule begins with [u|g|egroup] followed with : and identifier.
```

Evolving the CONSOLE - MGM interaction

- **Benefits**
  - avoid double parsing of commands
  - avoid name/string **encoding** “surprises”
  - avoid **snowball effect/replay** for long running requests

- Used only for new commands and some admin commands: acl, ns, fs, **accounting**

Fuse(x) status & dev

- **Fuse - eosd**
  - Only critical bugs fixed
  - Security issues fixed

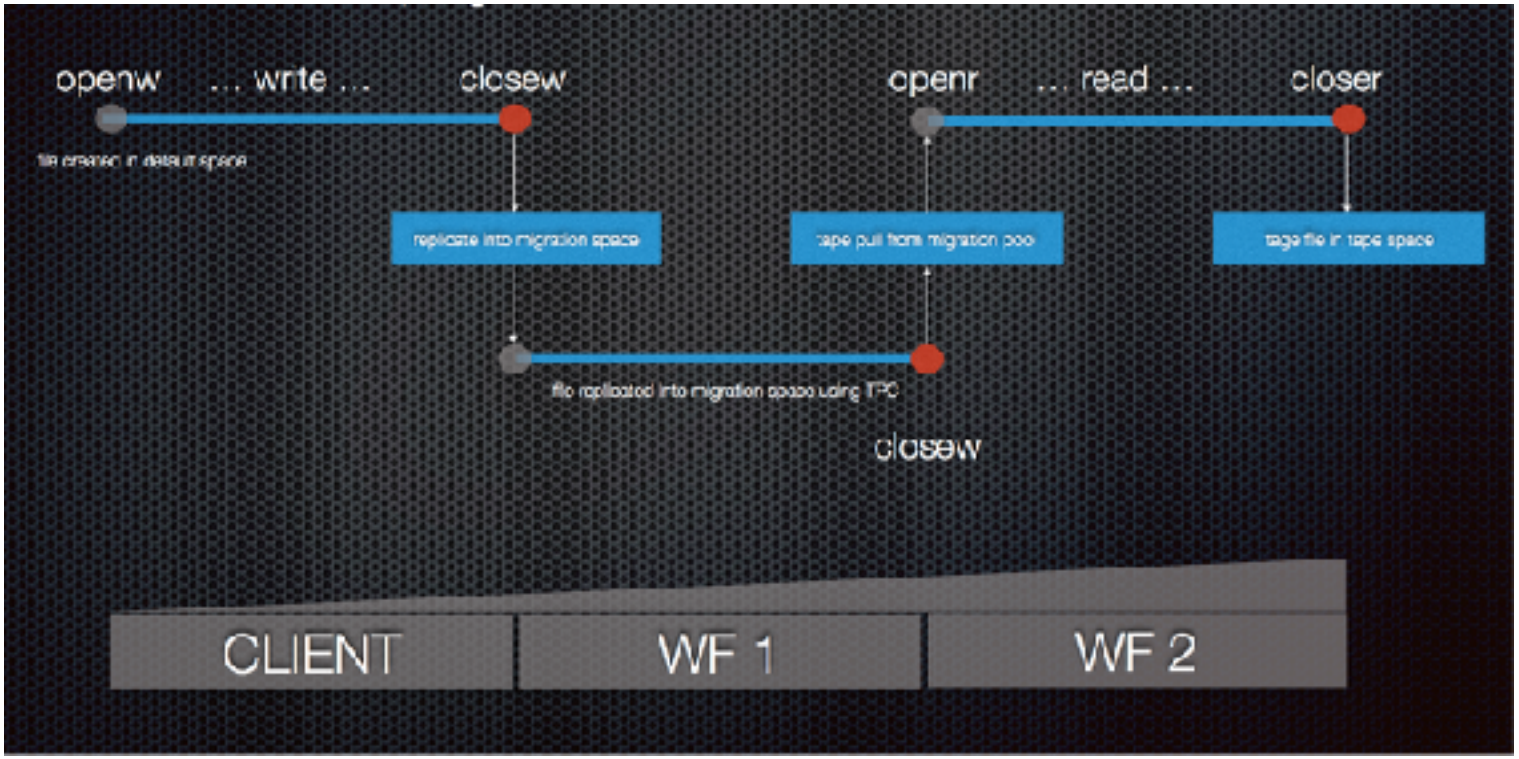
- **Fuse(x) - eosxd**
  - Not almost but just …. awesome 😉
  - See Andreas's talk
    - https://indico.cern.ch/event/656157/contributions/2866291/

Systemd support

- Critical step in migrating to **CentOS7**
- Thanks for the extensive contribution from the **Comtrade** team

- Thanks also for other developments like:
  - Monitoring for full **/var/** partition
  - **zstd** metadata compression
  - Table formatting engine

Extended workflow functionality



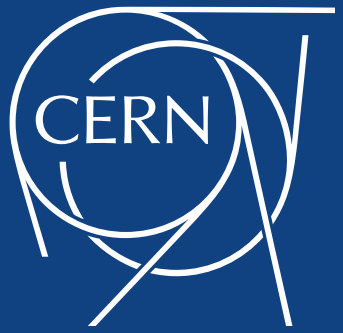
- Better integration with the CERN Tape Archive project

What's next …

- **New namespace consolidation**
  - Adapting internal mechanism like balancing/fsck etc. to the new namespace
  - Revisit the FST booting procedure
- Move **eos configuration** into QuarkDB
  - Bundled with new namespace
- **FUSE(X)**- production ready

Thank you!

Questions & comments?


CERN