

ALICE Status

XRootD workshop - Lyon CC-IN2P3 - June 2019

Nikola Hardi - nhardi@cern.ch

ALICE Status

Current storage status - quick facts

Disk storage	Tape storage
Total number of files: 1.5 B	Total number of files: 60 M
Total size: ~ 70 PB	Total size: ~ 70 PB
Largest disk SE: ALICE::CERN::EOS <ul style="list-style-type: none">- 0.5 B files- 20 PB storage	Largest tape SE: ALICE::CERN::T0ALICE <ul style="list-style-type: none">- 37 B files- 40 PB storage
CC-IN2P3: ALICE::CCIN2P3::SE <ul style="list-style-type: none">- 126 M files- 4 PB storage	CC-IN2P3: ALICE::CCIN2P3::TAPE <ul style="list-style-type: none">- 2.1 M files- 3 PB storage

File sizes and data rates

- Jobs read 50 GB of data on average
- High CPU efficiency, 5 MB per CPU core per second
- Read 1 EB of data in 2018
- Currently raw data on tapes, 2 replicas, 70 PB in total
- For run 3 we will store just a single replica of raw data on tape
- We expect additional 60 PB installation (by the end of this year?)
- We will have 60 PB of new data per year
- And additional 15 PB of disk next year
- This storage will be served by XRootD

Job scheduling and WAN traffic

- Jobs placed close to data: over 90% (used to be 98%)
- Some very well connected sites where we relax this policy
 - Examples: Romania, Italy
- This results in good CPU efficiency
- No need for caching, but we follow DOMA and XCache activities

ALICE Storage Elements

- 75 Storage Elements
 - 64 disk storage
 - 11 tape storage
- 41 XRootD instances (30 PB on disk / 65 PB on tape)
- 21 EOS instance (36.5 PB on disk / no tape)
- XRootD versions:
 - Few XRootD v3.x sites
 - Most of the sites at XRootD v4.7.x +
- Good backwards compatibility and interoperability

ALICE TPC Status

- Most of our XRootD (and EOS) backed sites have working TPC
- The ones that don't are:
 - Not yet configured, just the matter of communicating with site administrators
 - Use too old XRootD version, to be decommissioned
- 57 PB / 70 PB of disk storage available over TPC
- Mostly older and smaller sites don't have working TPC

ALICE XRootD clients

- From ROOT: TAlienFile / TJAlienFile
- From JAliEn shell - calling xrdcp as a subprocess
- From JAliEn-py shell - thanks to the xrootd python bindings!
 - Python 2 bindings only in PIP? - comment from T1/T2 workshop last month
- Very good backwards compatibility, no issues in general

ALICE SE Security

ALICE SE Authentication and Security

- Federated, single, distributed namespace
- Files available anywhere for anybody
- Single central file catalogue
 - Currently backed by MySQL
 - Working prototype developed based on ScyllaDB / Cassandra
- Mapping from LFNs to PFNs
 - LFN maps to one or more physical replicas (PFNs)
- Access envelopes issued by the central services
- Requests are verified on storage elements directly, without calling back central
- Authorization done by Alice token authorization XRootD plugin
- Single user owns all files on storage elements

Access envelope contents

-----BEGIN ENVELOPE-----

CREATOR: AuthenX
UNIXTIME: 1560175951
DATE: **Mon Jun 10 16:12:31 2019**

....

CERTIFICATE: none

-----BEGIN ENVELOPE BODY-----

```
<access>write-once</access>
<url>root://prod-se-01.ct.infn.it:1094//09/24875/c900bd00-8b89-11e9-8818-0242ca9cd510</url>
</fn>/NOLFN</fn>
<size>10459447</size>
<guid>C900BD00-8B89-11E9-8818-0242CA9CD510</guid>
<md5>cb1e998da55ad99f0d8979a36efc2e59</md5>
<pf>/09/24875/c900bd00-8b89-11e9-8818-0242ca9cd510</pf>
<se>ALICE::CATANIA::SE</se>
</file>
</authz>
-----END ENVELOPE BODY-----
-----END ENVELOPE-----
```

-----BEGIN SEALED CIPHER-----

```
0b8VJKR2ae00Q95rhmgPppE8lUSp1y4qMkUjKcPgcWzJs5p1zQLFRYiySHR7PQRB8A7GfPzK0WNI
UY50XMOrisw1uUZf8RqCooy3041XqbPkr4R6LUatoRiajf7grXq+F5+tNE3PvBmLKwEi06j/j3at
yECfBu2BMAzrugAe3RE=
-----END SEALED CIPHER-----
```

-----BEGIN SEALED ENVELOPE-----

```
AAAAGFfvYV0wE0yoiFYinJBxE7GxTTbL35zwdWYWI2iIozVgYc4n0UcJ2JHq/00mDn12vdm+ayU+
wbT2UeMzBbYb16vKWRUFQ295u+90AUwkqv9N1NDbYM/kdJXQ/42NQGP4EgcoGFEGtCG585bK5IQ3A
ZvLkt0Thlq6KvizaiZwUS0aTQUQp1NugSj0saLCJIjy0WwKb3seRI4ygMN4MqHWIt4u6MmUZL6n
cWJjeUMuFNiMbUTAQwsk1EluEc9Ftej153lRFtbywVwAQZBHXZbty+dygc8h3Bc2Q9J1YFb9mL94
/yNsNx2TkH1mR7LAFjN7LMjV3NNEEnMBzwo/8FQ0v1j2gRvZnR0sdmALITr2ibWjWr0IotabupRa1
Bc48Vh8GgKtrODQ/Yk/5zI30icELgCXzzIFEEIEXhmvusG/+wirNQ096Kgz1G5bwa0RZppp5uFa6
dy6SQUGzC8nB1Vp/WI1FUPDpbaE/TNwT06ev22eVXJlcyHYna16TF6XPFdG6GohFjm8Px2UwVACJ
6vtYtTrAUx6eKBvWkgfXueki5339VvXBuFcWUp/F1w8gjt5W4dkduFsdHc2CmKVWjWGcPaYnuC1m
aQ/zuDWkT1PUjmbVZfuFncn19wM35MYfxY6ldHwsz4Po9pJCiiousgue6/7CY6EoevdE25y58KZQ
YHdGf8M/in1yB181m1B+1oTNI1g72Gp8VG1dV8Ingu6P1KLuz1tncTwxZxk/Tr3TeB5nMdW8Y4WD
NqV2oQImA7yI1Ey0ayD21WL5QprxDfQ5P8NIUkkidQRG1kE2SvD02A+latcsDfllcP1/ijpCIN
0WmL2556WD41wDmZXA+08SvISSh+6hpD9XjUuhZyhIT3UB0Ze/nGR3h01YeLNOp+a1hTn+uWitIj
fe+4/CTeRgDi001WjxeLAU43iTasrChVPb6hXQMDjbnIB5kF+L4cACcgAiiX6NKxyfP/+fvcXO1U
3s1YU+mWj7MF9sZ9sVvUxKLa6IhtC6aPn6YS8dUZtjYuwRN4zHuUNwc5k0tg7p4=
-----END SEALED ENVELOPE-----
```

DECRYPTED & DECODED - **PLAINTEXT**

RSA ENCRYPTED & B64 - **CIPHER**

Blowfish ENCRYPTED & B64 - **ENCRYPTED ENVELOPE**

XRootD token authz plugin

- Implemented as XrdAcc plugin
- In use since 2005
- Had several revisions since then
- The plugin is distributed to all our storage elements

ALICE Signed URLs

- Alternative to the access envelopes
- Working prototype available since 2016
- However, we still rely on the access envelopes
- Looking forward to alternatives - such as macarons?
- ALICE is following DOMA activities

?turl=root://t2-xrdrd.lnl.infn.it:1094//08/41595/f10139b0-c9ae-11e5-a508-2c44fd849358&access=read&lfn=/NOLFN&guid=f10139b0-c9ae-11e5-a508-2c44fd849358&size=479&md5=3a0ff51c8b46f853e22f084f4e50a0e2&se=ALICE::LEGNARO::SE&hashord=turl-access-lfn-guid-size-md5-se-hashord-issuer-issued-expires&issuer=jsh_aliendb4&issued=1476363783&expires=1476450183&signature=h3BQgXm53Yy7TWq1HHUUthyEPyRW3PpqfgUI86ubtcvm9rUVEpgdzHjt2i2jdditP6M9jiDu3PPzmLBoe7Cv8pZEao4/37P+N1WXWs1Vem76nZSlernIT95QGzKaGDJdfSTpMoGwM2HYL0e8IN0ZPnPuWcklxaPRybCyKoXzc+U=

Conclusion

Conclusion

- ALICE stores 70 PB data on disk + 70 PB on tape
- ALICE expects 60 PB per year in run 3
- We rely heavily on XRootD, you are doing excellent job!
- ALICE Grid jobs read 1 EB of data in 2018, and this will keep increasing
- ALICE maintains XRootD authorization plugin
- Macaroons are considered as an alternative
- The bindings are important to us - native Java bindings?
- More than 80% of ALICE data on disk is reachable over TPC

Thank you!

Questions?

ALICE Status

XRootD workshop - Lyon CC-IN2P3 - June 2019

Nikola Hardi - nhardi@cern.ch