



# Data transfer and data integrity tools

Volodymyr Yurchenko

# Transfers monitoring



## MonALISA Repository for ALICE



My jobs | My home dir | Catalogue browser | LEGO Trains ★ | Administration Section | ALICE Reports | Alert XML Feed | Firefox Toolbar | MonALISA GUI

ALICE Repository

- ALICE Repository
- Google Map
- Shifter's dashboard
- Run Condition Table
- Production Overview
- Production info
- Job Information
- SE Information
- Services
- Network Traffic
- FTD Transfers
  - RAW replication
  - Transfer states
  - Transfer speed
- CAF Monitoring
- SHUTTLE
- Build system
- HepSpec
- Dynamic charts

### Transfer requests (add new request)

ID	Path	Target SE	Status	Progress	Files	Total size	Started	Ended
11211-112			- any -					Filter
11235.	ORNL::EOS Repopulation eos dumpmd files (fsid64)	ALICE::CERN::EOS	Done	<div style="width: 100%;"></div>	2090	4.417 GB	06 Apr 2018 18:11	07 Apr 2018 10:17
11234.	ORNL::EOS Repopulation eos dumpmd files (fsid64)	ALICE::CNAF::SE	Running	<div style="width: 100%;"></div>	941777	21.84 TB	06 Apr 2018 18:11	
11233.	ORNL::EOS Repopulation eos dumpmd files (fsid63)	ALICE::CERN::EOS	Done	<div style="width: 100%;"></div>	1678	3.74 GB	06 Apr 2018 18:09	07 Apr 2018 09:05
11232.	ORNL::EOS Repopulation eos dumpmd files (fsid63)	ALICE::CNAF::SE	Running	<div style="width: 100%;"></div>	1296637	31.2 TB	06 Apr 2018 18:09	ETA: today 19:37
11231.	ORNL::EOS Repopulation eos dumpmd files (fsid41)	ALICE::CERN::EOS	Done	<div style="width: 100%;"></div>	3030	65.5 GB	06 Apr 2018 18:08	07 Apr 2018 10:03
11230.	ORNL::EOS Repopulation eos dumpmd files (fsid41)	ALICE::CNAF::SE	Running	<div style="width: 100%;"></div>	533849	20.14 TB	06 Apr 2018 18:08	ETA: today 01:43
11229.	ORNL::EOS Repopulation eos dumpmd files	ALICE::CERN::EOS	Done	<div style="width: 100%;"></div>	2951	76.82 GB	06 Apr 2018 18:01	07 Apr 2018 07:51
11228.	ORNL::EOS Repopulation eos dumpmd files	ALICE::CNAF::SE	Running	<div style="width: 100%;"></div>	487367	22.17 TB	06 Apr 2018 18:01	ETA: today 02:54
11227.	replicateFiles - ALICE::JINR::SE	ALICE::CERN::EOS	Error	<div style="width: 100%;"></div>	1356719	74.46 TB	29 Mar 2018 11:23	13 Apr 2018 13:44
11226.	replicateFiles - ALICE::JINR::SE	ALICE::JINR::EOS	Error	<div style="width: 100%;"></div>	4450387	168.3 TB	29 Mar 2018 11:23	13 Apr 2018 12:24
11225.	Damaged files from ALICE::ORNL::EOS	ALICE::ORNL::EOS	Error	<div style="width: 100%;"></div>	524785	19.92 TB	16 Mar 2018 15:15	27 Mar 2018 12:08

<http://alimonitor.cern.ch/transfers/>

# Data transfer tasks

- Command-line recovery tool
  - Replicate SE
  - Recover missing/corrupted files in SE
  - Clean up orphaned files
  - Find and handle the „dark data“
  - SE removal
  - Find inconsistencies in LFN-PFN-GUID tables

# SE Migration

- Remount your data servers in read-only mode to avoid new data coming from the Grid
- Define what to copy
  - Prepare the list of GUIDs on server X (for partial migration) OR
  - We make a catalogue dump of all files (PFNs) on SE (call [masterSE](#) method)
- We queue transfers of all known files on SE / data server
  - Ignore dark data, skip non-existent GUIDs
  - Optionally: purge files from storage, i.e. implement „move“ procedure
  - Set an optional second SE to keep constant number of replicas
  - File won't be replicated if already has sufficient number of copies (configurable)
- Shutdown servers for software upgrade (e.g. move to EOS) or decommissioning

# MasterSE example output

```
#PFN,size,MD5,ctime
```

```
root://epgsr5.ph.bham.ac.uk:1094//10/17439/96ade76c-12de-11de-ab90-001e0b4e0440,4265,ebf84b18129fab807601379bc97497fe,1237285808771
```

```
root://epgsr5.ph.bham.ac.uk:1094//11/09198/cf76e0b8-180a-11de-af50-001e0b4e0440,4269,3ec753dc66b95ed972e716d41e127e36,1237854557710
```

```
root://epgsr5.ph.bham.ac.uk:1094//14/05089/e98585f2-155f-11de-af71-001e0b4e0440,4262,c45a631e833f11519172d974ab2a0c83,1237561255161
```

- We expect to have at least GUIDs list as an input for the script

# SE Migration – Time Estimations

- MasterSE dump can take from 2 to 6 hours
- Transfer schedule 1 000 000 entries in 6 hours
- ALICE::Birmingham::SE → ALICE::CERN::EOS
  - 347.2 TB, 8.4 million files, 20 days
- ALICE::JINR::SE → ALICE::JINR::EOS
  - 168.3 TB, 4.4 million files, 14 days

11220.	replicateFiles - ALICE::Birmingham::SE	ALICE::BIRMINGHAM::EOS	Done		194816	26.26 TB	20 Feb 2018 14:31	12 Mar 2018 15:06
11219.	replicateFiles - ALICE::Birmingham::SE	ALICE::CERN::EOS	Error		8428356	347.2 TB	20 Feb 2018 14:31	12 Mar 2018 11:46

# Disk / Server Failure


- Prepare the list of GUIDs, known to be lost (e.g. incorrect symlinks in XRootD namespace)
- Remove local pointers to broken files
- Send us the list (Dropbox / Cernbox / email ...)
- We mirror them back to your SE from other replicas (\*)
  - Old PFNs are deleted from the catalogue
  - If no other copies left – the file loss is unrecoverable, orphan will be removed from the DB
  - Optionally replicate to another SE(s) instead of original one

# Recover Broken files (list unknown)

- We create the catalogue dump and record the timestamp of the beginning of this operation
- Target site should check the sanity of the files in this list (ignoring anything newer than the given timestmap)
- Two diffs should result from this check
  - GUIDs with no replica on the SE (we try to recover)
  - Dark files on storage (delete if older than the timestamp)
- Syncs the catalogue and the storage for a specific point in time



# Recover Damaged Files

11225. Damaged files from ALICE::ORNL::EOS	ALICE::ORNL::EOS	Error		524785	19.92 TB	16 Mar 2018 15:15	27 Mar 2018 12:08
11224. Damaged files from ALICE::CNAF::SE	ALICE::CNAF::SE	Error		22353	808 GB	12 Mar 2018 12:13	13 Mar 2018 18:30
11223. Damaged files from ALICE::CERN::EOS	ALICE::CERN::EOS	Done		177	3.593 GB	02 Mar 2018 15:44	21 Mar 2018 20:29
11222. Damaged files from ALICE::NIPNE::EOS	ALICE::NIPNE::EOS	Error		5869	2.546 TB	23 Feb 2018 14:30	12 Mar 2018 12:07

# Clean Up The Orphan LFNs

- From previous operations we get the list of GUIDs with no copies left
- These files even if listed in the catalogue would not be accessible and cause problems in jobs
- We remove entries from the catalogue
  - Delete GUIDs and LFNs
  - Delete GUIDs, LFNs and PFNs of archive members

# Orphaned Files

- ~1 600 000 in 3 days, 6 hours

AliEn SE		Catalogue statistics						Storage-provided info			
SE Name	AliEn name	Size	Used	Free	Usage	No. of files	Type	Size	Used	Free	Usage
1. Bratislava - SE	ALICE::Bratislava::SE	565.7 TB	587.9 TB	-	103.9%	17,226,032	FILE	565.7 TB	543.9 TB	21.87 TB	96.13%



# Dark data

- Fast: as shown before (with the list of files on storage)
- Slow: `XrdCleanupSingle`
  - `xrdfs ls ...` recursively
  - Compare with the catalogue
  - `rm` orphaned PFNs
  - Estimated time of running: weeks or even months
- Can only check what is extra on the storage, lost files cannot be confirmed (data nodes temporarily offline, commands timeout)

# Dark data

SE Name	AliEn SE	Catalogue statistics						Storage-provided info			
	AliEn name	Size	Used	Free	Usage	No. of files	Type	Size	Used	Free	Usage
1. Bari - SE	ALICE::Bari::SE	1.201 PB	1.019 PB	186.2 TB	84.86%	23,255,101	FILE	1.201 PB	1.18 PB	21.26 TB	98.27%



# SE Removal

- Confirm your SE to be decommissioned
- If the storage is still working do the previous steps (make it all read only, send the list, we move out as much as we can)
- For the remaining files / sudden death
  - Delete all PFNs on this SE
  - If the only copy, clean up orphaned files
    - Otherwise create a replica on another SE
  - Drop the queue for physical file removal

# Catalogue Inconsistencies Clean Up

- We eliminate internal inconsistencies
  - Load big chunks of GUID tables in memory, loop over LFN tables
  - Get list of GUIDs that have no LFNs pointing to them any more, delete them
  - Same procedure for PFNs

# Scripts Repository

- Scripts to prepare lists of local files
- Compare with the catalogue dump for a given timestamp
  - To recover files from other replicas
  - To delete dark data
- Help your colleagues to avoid redundant work for these common cases
- Your help is crucial in populating the repository with the knowledge
- <https://gitlab.cern.ch/data-transfer/SE-scripts>



# Archive Members Clean Up

- JAliEn-based Grid job to remove:
  - TrackRefs.root from root\_archive.zip (for all productions)
  - AliESDs.root from root\_archive.zip (for some productions)

# Archive Members Clean Up

- Will be executed in jobs
- AliEn Broker now can request job token certificates from JAliEn
- Payload script uses identity from token cert to run JAliEn code
- When: May 2018



Thank you!  
Questions?