Table of Contents

WLCG Operations Coordination Minutes, November 3rd 2016

Highlights

- On going work to make a proposal for advanced warning of long site downtimes. ALICE, CMS and LHCb input needed. Please check minutes for more details
- New FTS Configuration Monitoring portal⬚ now in production. See FTS slides for more details
- RFC proxy failures for some services and some CAs (GGUS:124650⬚)
- FTS IPv6 transfers enabled in production at CERN

Agenda

- https://indico.cern.ch/event/540423/⬚

Attendance

- local: Maria Alandes, Alberto Aimar, Vincent Brillault, Maarten Litmaath, Julia Andreeva, Maria Dimou, Marian Babik, Maria Arsuaga, Stephan Lammel, Nurcan Ozturk, Jerome Belleman, Marcelo, Andrea Manzi, David Cameron
- remote: Alejandro Alvarez Ayllon, Alessandra Doria, Catherine Biscarat, Chun-Yu Lin, Dave Dykstra, David Bouvet, Di Qing, Dave Mason, Matt Doidge, Rob Quick, Thomas Hartmann, Tomas Javurek, Ulf Tigerstedt, Vincenzo Spinoso, Xin Zhao, Natalia Ratnikova, Nicolo Magini, Ale Di Girolamo
- Apologies:

Operations News

- The theme of the next meeting on Dec. 1st will be *The results of the Lightweight Site survey* (Agenda⬚). There will be a sub-topic on *HTCondor Accounting*.
- Maarten and Alessandra monitor the above survey participation. Reminders are issued via EGI broadcasting.
- Advance information on the CERN 2016 Year End closing here⬚.

Middleware News

- Useful Links:
    - ♦ Baseline Versions⬚
    - ♦ MW Issues
    - ♦ Storage Deployment
- Baselines/News:
    - ♦ As already reported some meetings ago, dCache 2.10.x support is ending this year. Still 2 T1s are running this version (BNL, RRC-KI-T1), so it would be good to know their upgrade plans to a new dCache golden release.
    - ♦ FNAL is running a really old version of FTS. It's quite important to plan an upgrade of their service to the latest version, because it's required in order to implement the new FTS Configuration Monitoring project (see FTS talk today)
- Issues:
    - ♦ EGI SVG critical advisory (https://wiki.egi.eu/wiki/SVG:Advisory-SVG-CVE-2016-5195⬚) All new kernels are out. Recommendation: schedule kernel updates & reboot WNs. EGI

deadline (for update OR mitigation) was 31/10.
- ♦ ARC-CE job submission broke with latest globus updates. (http://bugzilla.nordugrid.org/show_bug.cgi?id=3604☑). In particular when upgrading globus-gssapi-gsi from 11.22-1 to 12.5-2. Fix released today http://www.nordugrid.org/arc/releases/15.03u10/release_notes_15.03u10.html☑.
- ♦ issue with RFC proxies and Jglobus (GGUS:124650☑). Affecting dCache < 2.14 and Bestman for some user CAs. See RFC TF report.
- ♦ GGUS:123947☑, HTCondor 8.5.6 changed the default condor_q output such that only the current user's jobs are returned. This broke the job monitoring in ARC. setting CONDOR_Q_ONLY_MY_JOBS=false fixes the issue.
- ♦ After the upgrade to GPFS 4.2.1.1 Storm stopped recognising the mounted FSs (GGUS:124293☑). While waiting for a fix a workaround has been documented on the ticket by the devs.
- T0 and T1 services
  - ♦ CERN
    - ◊ see T0 report
    - ◊ FTS upgrade to v. 3.5.7 next week
  - ♦ BNL
    - ◊ upgraded xrootd redirectors to v 4.4.0
    - ◊ FTS upgrade to v 3.5.7 following the CERN upgrade
  - ♦ NL-T1
    - ◊ SURFsara moved to the new datacenter without significant issues
  - ♦ JINR-T1
    - ◊ major dCache upgrade to v 2.13.48
    - ◊ xrootd upgrade to v 4.4.0
  - ♦ KISTI
    - ◊ Update of dCache for ATLAS, CMS and LHCb to 2.13.46.
    - ◊ Updated the CMS xrootd redirector to 4.4.0.

Tier 0 News

- The main LSF batch instance has reached the limit of the supported size; a number of issues have shown up already, which are probably due to the size of the instance. New capacity will hence be added to HTCondor only.
- The submission of local batch jobs to HTCondor now works; consultancy on using and migrating to HTCondor started with a few larger user groups.
- The CPU accounting for CMS has been verified successfully; a similar validation is ongoing for ATLAS.
- During October more than 7.5 PB have been recorded into CASTOR.
- The CASTOR instances have been upgraded to 2.1.16-9. Preparing for the p-Pb run, CASTOR ALICE is now using a Ceph-based pool in addition to the standard disk resources.
- In order to prepare next year's data taking, 3rd-party EOS to CASTOR has been demonstrated that does not use GridFTP proxies. 6Gb/s have been measured (min requirement = 2 GB/s).
- Some instabilities in EOS related to authentication have been observed and have been worked around.
- Following the major, successful network intervention on 03-Nov-2016, a number of issues with the Tier-0 services were identified and quickly resolved.

Tier 1 Feedback

Tier 2 Feedback

Experiments Reports

ALICE

- Normal to high activity on average
- KISTI CVMFS incident
  - ♦ Users reported persistent job failures due to CVMFS errors on Sep 26
  - ♦ The CVMFS Squid servers had `/var` full due to excessive logging
  - ♦ A cleanup was performed, but the services were accidentally left down
  - ♦ Yet more failures were reported on Sep 30
  - ♦ The Squid services were then restarted that evening, thanks!
- EOS crashes at CERN and other sites on Oct 13
  - ♦ Clients unexpectedly used signed URLs instead of encrypted XML tokens
  - ♦ The switch was due to one AliEn DB table being temporarily unavailable
    - ◊ and a wrong default (now fixed)
  - ♦ The EOS devs have been asked to support the new scheme and avoid that unexpected requests crash the service

ATLAS

- Grid is full. Several sites in downtime last week for patching/installing the kernel for linux privilege-escalation bug CVE-2016-5195. All ATLAS central services were also updated last week.
- Network intervention this morning, no major issues so far.
- Very high network usage (~20GB/s) under investigation, mostly caused by input file transfers for the derivation production. Plans for optimizations under discussion.
- Feedback from RRB: Define a common strategy with CMS for LS2. Evaluate measures to mitigate impact on resources. Optimise current Software and workflows (running derivations from tape will be discussed this week).
- EOS to CASTOR xrootd 3rd party test showed a very good rate up to 6GB/s. Next steps to be discussed with DDM.
- FTS:
  - ♦ ATLAS would like to discuss long-term strategies for coordinating ATLAS DDM and FTS.
  - ♦ Addressing questions like what is the intelligence that will be in FTS and what should be on the experiments' DDM? We foresee discussing topics such as FTS awareness of storage protocols, internal knowledge of file structure, SDN integration, cache storage management, etc.

CMS

- proton-proton run of 2016 finished, proton-lead next
- re-reconstruction campaign of early 2016 data almost complete
- preparing Monte Carlo production for spring 2017 analysis round
- Hammer Cloud test not running at many sites for about a week is understood and fixed
- outages last week mainly due to security updates and reboots
- central services of the CMS global pool / workflow management system are running at Fermilab since a few weeks
  - ♦ limits of the virtual machine at CERN reached
  - ♦ discussion and request for more capable machine open with CERN-IT since then
- PhEDEx, DPM, and xrootd upgrade requested at sites
  - ♦ upgrade of PhEDEx and the FTS3 backend mandatory by Dec 31st
- tape repacking ongoing at sites
- we assume WLCG accounting task force will follow up with sites not properly reporting to APEL

- we plan to update the VO-card, adding address space requirement (we exceeded old limits set at sites recently)

CMS requests WLCG to review VO ID card documentation, since it must have been written having in mind single core at the time, and now that multicore is widely used, VO ID cards need some changes and the current documentation is not suitable anymore. An action on WLCG Ops will be defined for this.

LHCb

- Monte Carlo simulation, data reconstruction/stripping and user jobs on the Grid.
- Proton-proton run finished.
- Preparing to proton-lead run.
- Stripping Campaign programmed to end of the year stop. High load expected on Tape system
- Linux kernel exploit "DirtyCOW" fix took down two Tier1 sites for two days.

Ongoing Task Forces and Working Groups

Accounting TF

- Validation of the T0 accounting data generated by the new procedure progressed well. It was crosschecked with the experiment data of ATLAS and CMS. The new accounting reports will start publishing to APEL production instance. First month to be re-published is July. After checking of July data in APEL production instance will switch to the new T0 reporting system.
- Members of the task force agreed on the changes in the accounting reports which are generated by the portal. Proposed changes will be presented at the November MB. We plan to make new reports official by the beginning of the next year.
- Status of APEL accounting for HTCondor will be presented at the December WLCG Operations Coordination meeting

Information System Evolution

- GOCDB developers and EGI contacted to understand how to add extra information associated to service endpoints with extension properties in GOCDB. It is feasible to consider this feature in GOCDB and it is aligned with EGI plans to add more information in GOCDB.
- Next GOCDB release to be released in the next weeks will contain a writeable API. This is an interesting feature to allow sites to publish more information in GOCDB in an automatic way.
- Feedback from GLUE-WG experts to define storage and computing attributes in GLUE 2 needed for CRIC and storage accounting. This list will be used to document the information needed in the different information sources queried by CRIC. Discussions with OSG to make sure they can also provide this list are ongoing.
- Recruitment process to hire a new CRIC developer is ongoing this week and a candidate is expected to be selected very soon.
- Next IS TF meeting⬈ will take place on 10th November. VOfeed structure and integration with CRIC will be discussed.

IPv6 Validation and Deployment TF

Andrea Manzi explains that IPv6 transfers in FTS are now enabled at CERN in two nodes. There's not much traffic right now but the plan is to enable all nodes in one week time if there are no problems in the meantime.

Machine/Job Features TF

Monitoring

Status

- MONIT portal: http://monit.cern.ch (open to all authenticated users).

- Contains all raw FTS, Xrootd, ETF data. Some ATLAS Job Monitoring data.
- Examples of dashboards are available.

- Not production but it is pretty stable

- Discussing on dashboards with the experiments.
    - working groups for ATLAS dashboard
    - meeting with CMS possibly next week
    - reporting to the WLCG in a monitoring WG

- Added a link from the existing FTS Dashboard to the new MONIT portal with FTS dashboards.

MW Readiness WG

The 19th MW Readiness WG meeting took place yesterday Nov. 2nd. Agenda , MInutes. Summary:

- The pakiti client is in cvmfs now. Details here.
- LHCb will participate in the FTS verification effort, a way to avoid, as much as possible, suprises like the checksum problem (GGUS:124136) met on Sept. 28th. They will also participate in the verification of the CE and storage types that they use.
- CMS will discuss internally participation in the EL7 UI bundle/rpm testing.
- The experiment plans around EL7 migration will be discussed in this WG. Today's situation is, mostly, with the exception of an ATLAS update as reported at the dedicated Ops Coord meeting of Sept. 1st.
- The WG Mandate was reviewed and confirmed as still valid.
- The date for the next meeting is not yet defined Please email the e-group of the WG as soon as a vidyo meeting is desirable and to accelerate exchanges in jira. Our tracker is https://its.cern.ch/jira/projects/MWREADY . The jira dashboard view always shows a snapshot of open tickets.
- Please observe the actions and communicate progress to the e-group.

Network and Transfer Metrics WG

- pre-GDB on NETWORKING will take place on 10th of January, participation of experiments and sites is crucial, if you plan to attend PLEASE register at https://indico.cern.ch/event/571501/
- WG results recently reported at various events:
    - Network session at the WLCG workshop
    - CHEP 2016 networking plenary talks
    - CHEP 2016 infrastructure track
    - HEPiX Security and Networking sessions
- Throughput meeting was held on 27th Oct:

- ◆ Focused on the network analytics, see minutes⧉ for details
- perfSONAR 4.0 RC2 was released yesterday, we will intensify validation effort towards final release planned end of November, update campaign to follow once final release is out
- We are now using a new mailing list wlcg-network-throughput-wg@cernNOSPAMPLEASE.ch - joint mailing list for European and NA throughput meetings⧉
- WLCG Network Throughput Support Unit: CERN - RRCKI followed up, see twiki for details.

RFC proxies

- RFC proxy failures for some services and some CAs (GGUS:124650⧉)
    - ◆ certificates of affected CAs have the non-repudiation flag set
        - ◊ so far only GridCanada certificates were seen to be affected
        - ◊ there exist more such CAs
    - ◆ affected services are dCache SRM < 2.14 and BeStMan/EOS SRM
    - ◆ the consensus is that the fault lies in JGlobus, used by both
    - ◆ JGlobus is not officially supported by anyone these days
    - ◆ the dCache team are looking into a "private" build with an easy fix

- VOMS clients 3.0.7 progress
    - ◆ release notes⧉
    - ◆ already present in preview releases for UMD updates this month
- a new version of YAIM core for myproxy-init on the UI is available
    - ◆ temporary location: here⧉
    - ◆ it should soon appear in the UMD as well

Squid Monitoring and HTTP Proxy Discovery TFs

- There was a CHEP talk⧉ on the WLCG Web Proxy Auto Discovery
- CERN IT strongly objected to allowing a http://wpad.cern.ch/wpad.dat⧉, so the HTTP Proxy Discovery TF will be adding http://grid-wpad/wpad.dat⧉ to the beginning of the standard list of proxy auto discovery URLs to try.

Traceability and Isolation WG

- VOs self assessed themselves with regard to traceability (based on ip+timestamp). Results are positive, but some optimisation might be needed:
    - ◆ Alice is able to identify jobs & user within few hours
    - ◆ ATLAS should be able to query BigPanda to get jobs & users
    - ◆ LHCb's shifter can manually go through jobs, experts can query database
    - ◆ CMS didn't have the opportunity to test but is planning to store all jobs records in Kibana
- Containers: RedHat refused to included unprivileged mount namespace support in RedHat 7.3 due to security concerns
- Containers: Brian Bockelman presented a daemonless container solution, 'Singularity'. All present VOs expressed interest.
- Next meeting: 2016-11-22 16:00.
    - ◆ Sites more than welcome to come and express their position

Theme: FTS

Maria Arsuaga presented FTS configuration monitoring portal that is already in production and plans for the next release. There is a discussion on the feature to make user DN sending optional. It is reminded that there was a discussion about this in the past since some national regulations need to be respected and in some cases are quite restricted on this. There is a general agreement to coordinate with the security team for this.

Alejandro explains that the feature is in the roadmap after discussing with the security team.

Maria Alandes asks whether the ATLAS specific view in the configuration monitoring portal can be also done for other VOs. Maria Arsuaga explains that this can be indeed easily done without any problem.

For the FTS specific issues raised by ATLAS in today's report, they will be followed up in the Steering meeting. ATLAS is concerned that not all VOs are attending this meeting very actively. Maarten replies that ALICE do not use the FTS and hence do not attend those meetings. Alessandro replies that issues related to the network are a concern for all VOs. Julia explains that the Data Management Working Group could be also the right forum to discuss some of these issues.

WLCG Operations will discuss internally on whether a pre-GDB or any other sort of meeting would be more suitable to cover aspects that relate also with other areas like networking and affect all VOs.

WLCG Operations will also check the status of the Data Management Working Group, whether it has already some specific actions to work on and whether it plans to give regular updates at the Ops Coord meeting.

For FTS specific issues, it is agreed that the Steering meeting is the right place to discuss.

It is agreed to have a slot for FTS in the Ops Coord meetings for regular reports from now on. Alejandro or Maria will attend regularly and fill in a section in the Ops Coord twiki with relevant news and highlights for WLCG of the FTS world.

Theme: Advanced warning of long site downtimes

ATLAS presents their proposal for advanced warning of long site downtimes where it is suggested to have one month notice for downtimes longer than 5 days. Maarten asks what to do with downtimes that are 2, 3 or 4 days long. There are currently rules for scheduled and unscheduled downtimes (more resp. less than 24h advance warning) that have an impact on site reliability and availability. This has to be reviewed, as availability and reliability targets dictate how long a site could be down. There is a suggestion also for best practices to avoid T1 downtimes overlap. Marcelo also suggests that experiments ask themselves how long in advance they need to know a downtime to have time to react and be prepared for it. It is agreed that the remaining experiments prepare a similar draft, addressing all these issues so that a common draft could be agreed at the next Ops Coord meeting and a final proposal presented at the management board.

Action list

Creation date	Description	Responsible	Status	Comments
01.09.2016	Collect plans from sites to move to EL7	WLCG Operations	On-going	The EL7 WN is ready (see MW report of 29.09.2016). ALICE and LHCb can use it. NDGF will stay on SL6 for now but they plan to go directly to EL7 early 2017. Other ATLAS sites e.g. Triumf are working on a container solution that could mask the EL7 env. for the experiments which can't use it. Maria said that GGUS tickets are a clear way to collect the sites' intentions. Alessandra said we shouldn't ask a vague question. Andrea M. said the UI bundle is also making progress.
03.11.2016	Review VO ID Card documentation and make sure it is suitable	WLCG Operations	NEW	

Creation date	Description	Responsible	Status	Comments
	for multicore			
03.11.2016	Discuss internally on how to follow up long term strategy on experiments data management as raised by ATLAS	WLCG Operations	NEW	
03.11.2016	Check status, action items and reporting channels of the Data Management Working Group	WLCG Operations	NEW	

Specific actions for experiments

Creation date	Description	Affected VO	Affected TF	Comments	Deadline	Completion
29.04.2016	Unify HTCondor CE type name in experiments VOfeeds	all	-	Proposal to use HTCONDOR-CE. Still not done for ALICE. Raja will ask the status for LHCb.		Ongoing
03.11.2016	Proposal for advance warning of long site downtimes	All	-	ALICE, CMS and LHCb to provide a similar proposal like the one presented by ATLAS at the meeting. An agreement after all proposals will be made at the next Ops Coord and a final proposal agreed by all experiments will be presented at the Management Board	1st December 2016	NEW

Specific actions for sites

Creation date	Description	Affected VO	Affected TF	Comments	Deadline	Completion

AOB

-- MariaALANDESPRADILLO - 2016-11-02

This topic: LCG > WLCGOpsMinutes161103
Topic revision: r22 - 2018-02-28 - MaartenLitmaath