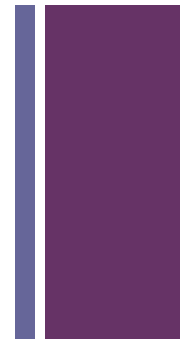




WLCG novelties report at the JUNO
datacenter meeting
November 2020



HSF WLCG Workshop - Virtual



■ 19-24 November

■ <https://indico.cern.ch/event/941278/>

■ Free registration

19-24 November 2020
Europe/Zurich timezone

Today session - Computing: Development in the experiments and implications on storage evolution

< Thu 19/11 Fri 20/11 Sat 21/11 Sun 22/11 Mon 23/11 Tue 24/11 All days >

Print PDF Full screen Detailed view Filter

16:00	Introduction	Simone Campana	16:00 - 16:15
	ALICE	Lachezar Betev	16:15 - 16:35
	ATLAS	Alessandro Di Girolamo et al.	16:35 - 16:55
17:00	CMS	Danilo Piparo et al.	16:55 - 17:15
	Coffee and cake		17:15 - 17:30
	LHCb	Christophe Haen	17:30 - 17:50
18:00	Belle II	Silvio Pardi	17:50 - 18:10
	DUNE	Steven Timm	18:10 - 18:30
	Coffee and cake		18:30 - 18:45
19:00	JUNO and other experiments at IHEP	Lu Wang	18:45 - 19:05
	Discussion		19:05 - 19:25



The second edition of the HSF-WLCG virtual workshop series brings us together again to review progress and discuss plans in the key areas of software and computing for HEP.

This workshop will be run with parallel software and computing tracks.

*Please note that workshop sessions will be **RECORDED** and the video made public afterwards.*



+ GDB and DOMA

- Grid Deployment Board monthly meeting
 - <https://indico.cern.ch/category/6890/>
 - November GDB
 - HEPIX Fall Summary
 - <https://indico.cern.ch/event/898285/timetable/#20201014>
 - XDC Project Results
 - October GDB
 - LHCOPN/LHCONE status report
 - IPv6 WG status report
 - Second Token-Based AuthZ hackathon summary
 - <http://cern.ch/go/DvD9>
 - Call for demonstrators for DOMA QoS
 - September GDB
 - Progresses on TPC (DOMA Third Party Copy)
 - DOMA – Data Organization Management Access
 - <https://indico.cern.ch/category/10829/>
 - TPC – Third Party Copy (SRM-less)
 - QoS – Storage Quality-of-Service
 - Access – Caching systems

+ HEPIX fall Summary



Site reports – general trends

- usage of GPU
- move from almost all intel to AMD
- exploitation of HPC resources
- increasing use of Container, Docker, Singularity
- Jupyter notebooks, and their integration in local batch systems
- network BW improvements
- tape archive migrations - software and tape technologies
- security concerns
- expanding features of institutional "storage boxes"
- implementing 2FA / MFA

- coping with COVID but also using CPU to fight COVID
- Tools for collaboration and communication under COVID and beyond
 - I think the community ever adapted a single tool as universally and quickly as Zoom

7

+ HEPIX fall Summary



Storage & Filesystems

- Alice showed that with latest centos8 it's possible to improve performance close to nominal in transfers to EOS
- Xcache is looking very promising to help storage-less sites as well to access data from remote sites.
- CVMFS is very stable and reliable, roadmap full of new and interesting features (containers)
- FTS is adding new interesting features to support QoS, Authentication and monitoring.
- research institutes in Germany federated through HIFIS
 - relying on FTS for data transfers with several ancillar solutions, with optimal results
- EOS+CTA tape service for atlas in production, replacing CASTOR
 - Proved to be fast and reliable
- HPSS migration to IBM tape technologies at BNL
- scalable High Performance Storage based on Lustre/ZFS over NVMe SSD
- dCache in the cloud environment

© Slides by: Peter van der Reest(DESY) at the
November GDB
<http://cern.ch/go/FwR6>

+ HEPIX fall Summary



Networking & Security


- several activities to meet future requirements of HL-LHC
 - jointly with other Data-Intensive research communities and R&E networks
- developments in perfSONAR and on marking packets to identify research activity
 - all to better understand use of networks
- IPv6 working group investigating move to IPv6-only
- work on CERN DNS load-balancing and NOTED project identifying large transfers in FTS
- traditional updates on security from CERN
- training users to avoid Phishing
- how to best cope with pandemic teleworking and use of virtual conferences
 - personally I would like a training on social engineering...
- work on Identity Management at BNL

© Slides by: Peter van der Reest(DESY) at the
November GDB
<http://cern.ch/go/FwR6>

+ HEPIX fall Summary



Computing & Batch

- Operating a production HTCondor cluster & Seamless automated maintenance,
- Learning-based Approaches to Estimate Job Wait Time in HTC Datacenters
- Status report of Benchmarking Working Group 
 - Building domain specific HEP benchmark from HEP workloads (for experiment but also for type of processing) – using docker and singularity
 - HEP-Score v1.0 to be released soon – candidate to replace HS06 - HEP Benchmark Suite
 - Benchmarking not only HEP/WLCG sites but also HPC resources that are being pledged increasingly
- HTCondor 2020 workshop report
 - first online/virtual HTCondor workshop
 - very successful companion to HTCondor week
 - online workshop organization

© Slides by: Peter van der Reest(DESY) at the
November GDB
<http://cern.ch/go/FwR6>

TPC Working Group in Context

The DOMA TPC effort has been ongoing for 2 years. The aim was to grow alternative ecosystems to GridFTP for TPC and ensure they work in our multifaceted environment!

We quickly settled on two candidates protocols: xrootd and HTTP-TPC

- HTTP-TPC is a set of agreed-upon interpretations of the WebDAV protocol (built on top of HTTPS).
- TPC extensions to Xrootd were already present but required changes in the authentication for X.509 authentication and were not implemented in dCache.

Putting Sites in production

ATLAS and CMS have started to cautiously enable HTTP-TPC at selected sites in production.

Internally, each has a different infrastructure and deployment approach:

- ATLAS uses Rucio; once enabled for a destination, HTTP-TPC transfers start on all configured source sites. Rucio only avoids sites that are not configured in AGIS to serve data via HTTP-TPC.
- CMS uses PhEDEx and enables one single channel at the time. The changeover is managed by site admins.
 - CMS is in the process of changing from PhEDEx to Rucio; once the transition is done, a more aggressive rollout is expected.
 - Currently, a small software patch is required to be a HTTP-TPC destination site.

Both VOs do additional functional testing before putting a site into production with HTTP-TPC.

5

© Slides by: Alessandra Forti (Manchester University) at the September GDB
<http://cern.ch/go/9QZn>

XRootD 5.0

The XRootD 5.x series adds many new features - the most exciting of which is the support for xrootd-over-TLS (xrootd protocol version 4.0).

- This provides support for integrity checking and encryption using the well-understood TLS protocol.
- Critically, encryption permits **use of the bearer tokens for authorization**.

We strongly encourage the rollout of new clients that understand the protocol.

- Integrators are working to validate the server.
- One current blocker bug (fixed, waiting on 5.0.2) prevents HTTP-TPC from working with FTS.

13

© Slides by: Alessandra Forti (Manchester University) at
the September GDB
<http://cern.ch/go/9QZn>

Replacing SRM at T1s

- Outside of operating tape archives, any site can today transition away from SRM.
- We still have no agreed-upon, interoperable protocol for managing archives over the WAN.
 - Although we are relatively close!! CTA has shown a tape archive can be managed completely through the xrootd protocol.
- In the meantime, we are looking at:
 - Splitting up the task in FTS: SRM to stage to disk, HTTP-TPC to transfer.
 - Adding non-X.509 auth methods to SRM. dCache already has a working prototype for this!

© Slides by: Alessandra Forti (Manchester University) at the September GDB
<http://cern.ch/go/9QZn>



eXtreme DataCloud: XDC

- ✘ Three(-ish) year project
 - Primarily focused on **software development**.
- ✘ Successor to INDIGO-DataCloud
- ✘ XDC is a continuation of INDIGO-DataCloud's **storage activity**.
 - Sister project DEEP worked on the more compute.

© Slides by: Paul Millar(DESY) at the November GDB
<http://cern.ch/go/n9vh>



OpenID-Connect and bearer tokens

✘ INDIGO-DataCloud, explored moving **away from X.509**

- ...→ Realisation that X.509 client credentials are a barrier for new communities.
- ...→ Focused on OpenID-Connect, an industry standard.
- ...→ OpenID-Connect support added to StoRM and dCache

✘ In XDC, this continued...

- ...→ with CERN, OIDC support added to servers (**FTS, DynaFed**) and clients (**gfal2, davix**).
- ...→ **Rucio** received funding to implement OIDC support.

✘ WLCG gained a **complete data management stack** with OIDC support.

- ...→ From Rucio client down to data-movement does not require X.509
- ...→ Demonstrated as part of CERN-hosted hackathon events.

© Slides by: Paul Millar(DESY) at the November GDB
<http://cern.ch/go/n9vh>



QoS: exploring something new

- ✗ Based on the idea that **storage** is not all the same:
 - In HEP, we already know this from DISK and TAPE, but it applies more generally
- ✗ Not all **data** is the same:
 - Intrinsic value: raw data vs log files
 - Importance can change over time.
- ✗ Exploration started in INDIGO-DataCloud
- ✗ Continued within XDC as contribution to **DOMA-QoS group**
 - Work continues in connection with EU-funded ESCAPE project.

© Slides by: Paul Millar(DESY) at the November GDB
<http://cern.ch/go/n9vh>



dCache



- ✘ Targeting specific problem: polling
 - ⋯→ Always an **unhappy compromise** between too fast and not fast enough
 - ⋯→ Overheads increase as number of files; particular with increased use of tape storage.
- ✘ Two solutions: site-level (**Kafka**) and global (**inotify-over-SSE**)
 - ⋯→ Both allow near-realtime notification of activity with minimum overhead
- ✘ Allows for powerful innovation: **arbitrary work-flows** driven by storage activity (metadata extract, processing, etc)
- ✘ Deployed within **EU-XFEL** to drive work-flows; support being added or evaluated at various WLCG sites.

© Slides by: Paul Millar(DESY) at the November GDB
<http://cern.ch/go/n9vh>



EOS



✘ **Storage adoption** (S3, WebDAV, xroot).

- ...→ Use external storage as part of EOS.
- ...→ Walk namespace to learn of existing data.

✘ **QoS management** interface

- ...→ Available through the INDIGO-DataCloud standard (CDMI), using the QoS frontend server.
- ...→ Adopting a common approach by implemented (minimal part of) the dCache REST API, so using the same “dCache” plugin for QoS frontend server.

✘ **QoS transition** engine.

- ...→ Existing support allow transition of single file
- ...→ Added engine to schedule transfers, monitor and retry (if needed).

© Slides by: Paul Millar(DESY) at the November GDB
<http://cern.ch/go/n9vh>

✘ Support for **trusted xcache** instances:

- Use SSS as mutual authentication between xcache and EOS
- Trusted xcache passes on details of identity without credentials; no delegation required.

✘ Deploying xcache as a **National cache solution**:

- Takes advantage of NREN high-capacity network links.
- Prototyped at sites in Italy (including CINECA HPC facility).
- Using DynaFed to federate cached content.

NginX: xdc-http-cache

✘ Update existing X.509 support to understand **VOMS proxies**

Allows extraction of information from proxies for authorisation.

✘ NginX may be deployed as a **VO-specific cache**:

- Only members of VO are allowed to use cache
- Allows for safe use of robot certificate for accessing VO data.

✘ Deployed both with and without **DynaFed**:

- Both national and site level caching supported
- DynaFed allows clients to use multiple (geographically distributed) cache instances.