



HEPiX Autumn 2022 summary

Peter van der Reest
European co-chair

2022-12-14

including many contributions from the HEPiX board and track conveners



HEPiX in general

- for around 30 years now, IT professionals from various research centers worldwide have come together to discuss common themes
 - evolved from HEPvm to cover Unices
 - joined with HEPNT in the early 2000s, so from then all computing environments covered
- in the last 5-10 years we have encouraged IT groups from other disciplines to join: Genetics, Photon & Neutron Science, Geo-Sciences
 - many of the storage and compute challenges are very similar
- usually 2 meetings every year
 - hosted in Europe, North America or Asia by local organisers
 - coordination by two co-chairs

HEPiX meetings



- for the last two years, we have held online meetings only
- this years' conference in Umeå, Sweden was the first full hybrid meeting
- most of the usual tracks we covered – we'll come to that in a second





HEPiX Autumn 2022 Closing Remarks

Peter van der Reest
European co-chair

2022-11-03

including many contributions from the board and track conveners

welcome back to HEPiX

welcome to Umeå



- Mattias and his colleagues did a great job of setting up this first hybrid meeting in little time
- and they managed to create a great experience here:
- brand new venue with great conference facilities
- very nice lunches and coffee breaks
 - with great opportunities to talk and exchange ideas
 - helped a lot to quickly re-establish the HEPiX atmosphere
- a very good conference dinner
 - with walk along the river under a beautiful rising moon
- only downside: no Aurora Borealis
 - we need to return for that



first real hybrid meeting

- new format – more effort to manage meeting than in fully online format (my opinion)
- not only remote presenters, but also remote conveners and moderators
- but importantly re-enabling in-person interaction while offering remote access for those who cannot travel yet



Participants

- Last week, around 120 registrants
- During the week steady increase of people joining us
- Currently: 185 registered participants, of whom 45 attended in person
- all in all, the meeting ran very smoothly
video conferencing etiquette was well observed!
- and the local participants behaved reasonably well also
- Thank you all!



Part

- Last
- Dur
- Curr
- all in
- and
- Than



ed in person

observed!



Programme

- Most regular tracks were present
- Workshop running fully in local time zone from Monday through Thursday
 - between 10:00 and 18:00h CET
- Most sessions were attended by 70-90 participants in total
 - about the same as in the on-line HEPiXes
- Most sessions were recorded and will be made available online
- Addition of one new track: “Show us your Toolbox”
 - fast paced look into daily ops
 - the tools and procedures used by participants



Site reports I

- HTCondor and Slurm (almost) everywhere
- Many sites manage several HPC and HTC clusters.
Some provide part of the WLCG pledges in HPC environment
- Kubernetes/OKD everywhere
- CTA transition finished at former CASTOR sites.
New dCache+CTA setup to be in production at DESY in 2023
- Networking:
- 800Gb capable network HW already available at NIKHEF.
- >1Tb WAN ROADM-based connection CERN-CNAF via GARR

Site reports II



- Still not clear preference between the RedHat clones, sites using either RHEL (with site licence), AlmaLinux or RockLinux. Lots of sites using more than 1 Linux flavours
- Jupyter is being widely used as the main analysis tool
- new facilities being built or just built:
 - technopolo (INFN), PCC (CERN), BNL CFR (BNL), SDMC (DESY), VIL3 (CC-IN2P3) and STFC-RAL
- 2FA appearing (TOTP and/or YubiKeys)
- Many open questions related to token transition. (CERN, Light Diamond Source, NSC)
- Quantum Technologies arise in IHEP & DESY

Thanks to Michele Michelotto for summary

End user services & operating systems



- Dennis van Dok showed how to create (even globally distributed) highly-available services by the use of ip anycast. In his example it was implemented for the PKIX translation service RCAUTH.eu.
- Alexandre Lossent presented how CERN is going to unify and consolidate its portfolio of webhosting platforms on top of the new OpenShift Kubernetes Distribution platform ODK4
- Alex Iribarren presented CERN's experience with CS8 which was chosen as successor for CERN-CentOS7. CERN feels unhappy with that choice and is collecting experience from other sites.
- There was a longer discussion on this during a BoF session on Monday DESY provided a talk regarding their experience with AlmaLinux-8.

Computing & Batch Services I



- BigHPC consortium presented by LIP. They are using GitOPS to improve the reliability of OPS of the HPC center with a code repository on Gitlab and CI/CD pipeline
- KubeFlow at CERN to manage the flow of machine learning application
KubeFlow is already in production in several companies, some very big like Bloomberg, Google
So they started experimenting with Jupyter notebooks and then automated ML pipelines. There is an active community world wide
- Automation of remediation procedures for batch service at CERN. Huge number of batch queues, some problems can be automatically corrected without human intervention. State machine daemon controls the WN state and may decide to drain queues or reboot multiple machines. This is very successful, the incident count has improved by an order of magnitude (10 times less)

Computing & Batch Services II



- Very interesting summary talk of the HTCondor European Workshop. very comprehensive overview of all that happened that week
- HEPscore is very close to the final definition. The HEPiX WG has evaluated the alternatives and designed the new benchmark. The WLCG Task Force is recommending the final HEPscore composition before end of the year and defining the strategy for the transition and overlap phase
A 2 day workshop in September at CERN has permitted to finalize many points. The final HEPscore will be based on 6 or 7 workloads from LHC run3 plus Belle2. After the WS the decision to try a benchmark that runs also on ARM and to start add power consumption measurements
It is important to address the few open questions on short term
- Final talk was a detailed review of HTC scheduling strategies in GridKA at KIT



Grids, Clouds & Virtualisation

- one contribution and excellent talk
- Cloud Infrastructure Update at CERN
 - Production services have been started since 2013
 - Most of the Components are provided by OpenStack Xena
 - More than 1300 VM are running on 9000 physical servers
 - VM migration campaign for 4000 VM was completed at the background
 - ARM server can be supported by Ironic (bare metal provisioning)
 - A lot of types of GPUs are available
 - Introducing Software Defined Network



Networking & Security I

- Networking (4 talks)
 - Transferring data from ALICE to the CERN Computer Center
 - Network setup and design of the connection between CERN computer center and ALICE container hosting Infiniband setup. The infrastructure provides a total of 2.4Tbps capacity by using two DWDM (Dense Wavelength-Division Multiplexing) lines.
 - Update from the HEPiX IPv6 working group
 - WLCG Tier2 storage now ~90% IPv6-capable; FTS Monitoring needs work to enable tracking of use of IPv6; packet marking in IPv6 Header for network analysis requires IPv6; working group is analysing and fixing cases where IPv4 is still used for data transfers.
 - Update on the Global perfSONAR Network Monitoring and Analytics Framework
 - Included the status/plans for, the evolution of the WLCG/OSG perfSONAR infrastructure, as well as the new, associated applications that analyze and alert upon the metrics that are being gathered
 - Research Networking Technical WG Status and Plans
 - an update on the working group activities and challenges. This included flow labeling and packet marking technologies (scitags), tools and approaches that have been identified as important first steps for the group



Networking & Security II

- Security (2 talks)
 - Enforcing Two-Factor Authentication at CERN
 - CERN moving to a new SSO using Keycloak. A good opportunity to better integrate 2FA and require 2FA for access to critical services
- Computer Security Landscape Update
 - The traditional HEPiX update on all related to IT security. The current threat landscape and appropriate mitigations were presented in relation to security attacks including data breaches, DDoS, ransomware and compromise of credentials

IT Facilities and Business Continuity I



- Saving Energy at DESY -Yves Kemp

Yves describe several approaches to saving energy firstly by recuperating energy from the cryogenics plant in place since 2017, but identified several more as yet unused possibilities. When looking specifically at computing essentially you turn-off machine or throttle them. DESY prefers switching off to throttling. Discussions with HT Condor developers leading to ideas to limit the maximum power per cluster. Other longer-terms ideas are also being investigated such as reduction of UPS coverage, changing to a water-cooled cooling system and purchasing more efficient systems in the future and improving the workflows and tools.

- CERN IT Hardware Procurement – Luca Atzori

Luca described the CERN procurement process including optimising the technical specification. He then summarised the main tenders in 2020-2022 for storage and CPU. He then detailed some technical considerations CERN is seeing such as increased power consumption, large number of cores and sockets, the fact that AMD has again become competitive and the emergence of M.2 SSD/NVMe on the motherboard which can be interesting for some cases. He finally addressed issues with the supply chain which have impacted CERN's planning.

IT Facilities and Business Continuity II



- Managing CERN Data Centres with OpenDCIM – Jarek Polok

Jarek summarised the scale and complexity of managing CERN's IT infrastructure and tools that exists to ease this Data Centre Infrastructure Management (DCIM) of which there are many both commercial and open source. He noted that after evaluation CERN selected OpenDCIM but quickly took the decision to fork from upstream to speed up adding new functionality that CERN needed. OpenDCIM has been integrated with many existing CERN tools which he explained. He then addressed whether CERN's OpenDCIM implementation was now CERN specific, which he explained that it isn't except for the asset information display but there are plans to make this more open too. Although the tool is operational and heavily used further enhancements are planned which Jarek highlighted.

- High Density Computing Room, 15 years of experience – Mattias Wadenstein

Mattias first revisited a HEPiX presentation from 2009 describing the setup of 28 racks at 25kW per rack with air cooling. The experience has generally been good and the setup working as expected but some fixes were nonetheless necessary during the operational period such as changing the cooling regulation from temperature to pressure based and introducing sound dampening to improve the work environment. He described an upgrade from 400kW to 650kW and the experience with this. To be able to handle more power in the DC without increasing the air flow the inlet temperature was increased by 4 degrees with 50% Rh. He described some challenges and mishaps as well as experience with a UPS-less DC. He finished by concluding that the experience was positive.



Show us your toolkit

very dynamic 80 minutes of quick demos and background infos

- creating visualisations on command line from HTCondor operational data
- GreenBone Security Assistant / Greenbone Vulnerability Monitor to scan systems
- Collaboration in a sysadmin group: Flight rules
 - short descriptions and recipes for handling standard problems in daily ops.
- using emacs for everything - from notetaking to creating presentation slides
- git for quality control & assurance
- tools for multi OS creation of RPMs with rpmsci
- Tools for controlled updates of HTCondor farms
- Tools to check for quality of server certificates



Storage & Filesystems I

- DPM storage migration to dCache and EOL of DPM - Petr Vokac
 - Sites should migrate away from DPM by Summer 23 as EGI support is ending
 - Procedures for smooth-as-possible migration into dCache were presented
 - Stressing the importance of using supported and evolving software for important infrastructure
- Evolving storage services at INFN-T1 - Andrea Rendina
 - Overview of technologies in use
 - XrootD and StoRM (with its StoRM WebDAV service) and the way services are evolving in the context of new protocols and authorization approaches for bulk data transfers between WLCG sites
- CERNBox : sync, share and science - Diogo Castro
 - aims to bring the ease of a file sync and share service to scientific data processing at CERN
 - showed the developments around CERNbox and the surrounding ecosystem



Storage & Filesystems II

- Storage for the LHC: operations during Run 3 - Cedric Caffy
 - in 2022, more than 380PB have been written by the experiments and 2.4EB have been read
 - decommissioning of CASTOR service and successful replacement by the CERN Tape Archive (CTA)
 - data storage strategies in EOS
- Lessons learned from an ATLAS data movement stress test for the Lancaster WLCG Tier-2's new XROOTD/CEPHFS storage - Matthew Steven Doidge
 - analysis of unexpected bottlenecks and problems after stress test of the new XRootD fronted CEPHFS storage element
 - results have changed the site design to horizontally expand our xrootd infrastructure
- Multi-experiment Storage service at BNL - Carlos Fernando Gamboa
 - overview of the BNL storage service archive, challenges and recent developments at the Scientific Data and Computing Center (SDCC) at BNL
 - diverse range of HEP scientific experiments, e.g, LHC-ATLAS, Belle2, DUNE with 122M files and data storage of 76PB

Storage & Filesystems III



- Using dCache Storage Events in NextCloud and On-Site Experiments - Christian Voss
 - using dCache provided storage events to register 'dark data' files with NextCloud
 - this is important for data ingest into Nextcloud served space via dCache directly
 - adapting this workflow to suit the needs of other experiments undertaken at DESY
- A Scalable and Efficient Staging System between Dcache and HPSS at BNL - Jane Liu
 - detailing a new efficient staging system between dCache and the backend HPSS tape storage
 - using three modules: the Endit Provider Plugin originated from NDGF, the ENDIT HPSSRetriever developed by BNL, and the optimized BNL's HPSS Batch (ERADAT) system performance bottlenecks in staging were identified during the WLCG tape challenges
 - system significantly improves overall tape staging performance through its asynchronous nature



Comments and Discussions

- being back in an in-person setting revived offline discussions and brief 'coffee break' exchanges
- these were not possible with the remote participants
 - HedgeDoc shared document was used, but not as intensively as before
 - still available for background questions to and discussion with speakers
 - document will be closed some time next week, so you can still add info



Board meeting - I

- saying goodbye to long time co-chair Tony Wong
- and welcoming him as an honorary board member on unanimous invitation by the board
- being updated on the election of a new non-European co-chair
 - committee with David Kelsey (chair), Gang Chen, Josep Flix and Ofer Rind
 - one round of voting has delivered a new co-chair:

Tomoaki Nakamura (KEK) has been elected and he has accepted the position

- Tomoaki sends his regards, but had to leave early today to catch his flight



Board meeting - II

- for some time there has been a running discussion on the topic of adding a third co-chair to the board
 - to share the load on the co-chairs
 - to more equally represent the 3 major global regions that HEPiX participants come from
- the board has extensively discussed this in its meeting yesterday evening and has decided to proceed with the extension
 - but in a dedicated procedure to find the 3rd co-chair
 - taking place in 2023 (probably Autumn) and then searching for a non-European, non-Asian candidate



Board meeting III

- review of current meeting
 - tracks, ideas for new themes, break-out rooms during coffee break
- planning ahead – physical meetings as our main target
 - board is in favour of in-person meetings with best-effort streaming for the future
 - hybrid meetings make for a lot more effort and provide less personal interaction
 - may do a hybrid meeting for the upcoming Spring meeting
 - we have been invited to Taipei and will be doing a back-to-back meeting with [International Symposium on Grids & Clouds \(ISGC\) 2023](#)



Board meeting IV

- meetings after that will be coordinated in the coming months
- board is looking at shifting back to European meetings in Spring, North-American and Asian meetings in Autumn – all dependent on local boundary conditions of course
- issue of energy conservation and handling shortages are important issues
 - following the various discussions already taking place
 - considering a BoF on experiences so far in Spring '23 as well as planned future measures
 - feel free to discuss on hepux-users@hepux.org until then!
- rethinking the tracks and restructuring a bit



Drawing to a close

- thank you again for making HEPiX
the valuable resource it is!

- Share your thoughts and comments with us, via email

hepix-board (at) hepix.org

Finally



Hope to see you again in Spring 2023



Academia Sinica – ASGC
Taipei
27 – 31 March 2023