HEPIX Spring 2021—Wrap-up

Tony Wong

Statistics

- 510 registered participants
  - Varying participation but observed a peak of 200+ at instances
- 54 presentations covering all 8 standard HEPIX tracks
  - 11 in Site Report track
  - 6 in End-User IT & Operating Systems track
  - 8 in Networking and Security track
  - 6 in Computing & Batch Services track
  - 10 in Storage & Filesystems track
  - 4 in IT Facilities & Business Continuity track
  - 5 in Grid, Cloud & Virtualization track
  - 4 in Basic IT Services track
- CodiMD
  - in use for comments and Q&A—lots of content added
  - ~60 users online
  - To be published (and announced) after workshop ends
- All sessions (including BoF's) were recorded. Announcement about availability will be made after workshop ends
- The following is a brief summary of presentations. For full details, please refer to the Indico agenda for this workshop https://indico.cern.ch/event/995485/timetable/#all.detailed

Site Reports

- Completed replacement and full operational availability of KEK CC resources in Sept. 2020
- Several sites described integration of Cloud-like services (ASGC, INFN, RAL)
- Migration from CASTOR to CTA underway (CERN) or planned (RAL)
- Discussions about long-term plan to replace AFS at CERN (effort needed, organization, etc)
- Provisioning of HPC resources (in some cases, for the benefit of LHC experiments) at several sites and (IHEP, PIC, BNL)
- Expansion and/or construction of new data centers at several sites (CERN, GSI, BNL, TRIUMF, IHEP)
- Support for Photon Sciences at various sites (Diamond, BNL, IHEP)
- Reliance on various collaborative tools (MatterMost, Zoom, etc) throughout the community as a result of operational changes due to the pandemic

End-User IT Services & Operating Systems

- IHEP described the activities of a new IT group (30+ members) formed to support the IT and software needs of HEPS (new light source due to come online in 2025)
- CERN presented on productivity tools OnlyOffice and Collabora Online as alternatives to MS Office as part of the MALT project timelines.
- Proposal for a change in focus for desktop service at CERN—from computer to user-centric
  - Addresses office desktops and personal devices (smartphones, laptops, etc)
  - Better aligned with current licensing models—per user, rather than device
- Technical presentations on latest Linux developments, in response to recent unfavorable IBM/RedHat announcement regarding CentOS availability and support
  - CERN presented on its current short-term plans and long-term choices
  - RedHat held a Q&A session to answer technical questions on CentOS and CentOS Stream
- CERN also presented on status of migration of mail service from MS Exchange to Dovecot (with some words about the Kopano pilot testbed)

Networking & Security

- CERN presented on the status of migrating ISC to Kea DHCP due to software end of life
- Security update of recent incidents and uncovered vulnerabilities was provided by the CERN cyber-security team
- IPv6 WG reported on the progress of migration to IPv6 among WLCG sites – steady, but with room for improvement (see WG slides for details)
- Authorization WG report – progress on activities to enable migration from X.509 to tokens
- WLCG/OSG Network Activities report – current focus on analytical studies and user interface development for a production-level network alarming service
- Research Network Technical WG update – progress on activities about packet marking to identify and shape traffic flows
- Framework for Cybersecurity Programs – presentation by Bob Cowles (Trusted CI)
- CERN Proximeter—device to alert and track people within 2 meters of each other (compliance with COVID-19 prevention protocols)

Computing & Batch Services

- Univ. of Bonn—pilot operation of Jupyter notebooks running on worker nodes without inbound (but with outbound) connectivity
- Univ. of Bonn—presentation on concept (COBaID/TARDIS) for opportunistic access to non-dedicated resources across 5 sites in Germany. Works well, looking to expand.
- INFN provided a migration status from LSF to HTCondor and a brief description of resource usage for COVID-19 R&D
- Report by WLCG HEP-SCORE DeploymentTask Force—steady progress on a complex task
- Report by HEP CPU Benchmarking WG—benchmark suite and HEP-SCORE ready and need volunteer sites to participate in testing
- BNL presentation on the replacement process of institutional (HPC) resources—aim to have new resources available in early 2022

Linux BoF

- Announcement by IBM/RedHat changed availability of a reliable Linux OS to community
- Lively discussion attended by 150+ attendees
  - Affects many communities beyond HEP
  - Potential path(s) forward
    - Streams model – short (5-yr) support not aligned with community needs
    - Paid subscription – RH not only choice
    - Alternative Open-Source distributions
      - Revive Scientific Linux?
      - Rocky Linux?
      - Others?
- Discussions about community-driven effort and common solution(s)
  - Several comments about not relying on an external (ie, Rocky Linux) organization. How do we know it won't be monetized and put us in the same situation a few years from now?
  - Encourage large institutions to make a commitment to lead and organize effort
- Schedule follow-up discussions
  - Organize discussion forums at HEPIX and elsewhere (GitHub?)

Storage & Filesystems

- IHEP presentation on activities and progress on designing Data Management System for HEPS

- CERN presentation on migration to CTA in preparation to Run 3

- CERN described the increasing support for containers deployment across a diverse number of usage scenarios (service deployment, large-scale analysis, secure container image distribution, etc)

- Ceph is being increasingly adopted at RAL (with and without erasure coding), with a "self-healing" feature to increase data reliability and integrity

- Tape infrastructure at CERN readied for Run 3. Flexibility to adjust to integration of new technologies (ie, LTO-9 drives and media)

Storage & Filesystems

- BNL presentation the future of tape for mass storage—cost effectiveness of deployment dependent on optimizing usage patterns and shared capital investments
- DESY presented on the developments on small-file packing for dCache tape interface, based on experiences at Petra-III—observed significant packing and archiving throughput increase
- CERN presentation on release of xRootD5, including new features such as encryption to allow for transfer of confidential data and handling of authorization token (ie, SciToken), etc
- Erasure Coding (EC)—new Working Group (WG) in HEPIX to study EC usefulness to help cope with data storage requirements for the LHC-HL era.
  - Welcome feedback from community
  - Periodic, detailed discussions with updates presented at existing venues (ie, pre-GDB, HEPIX, EOS workshop, etc)
- Presentation on the CERN SOLID project (collaboration with IT Univ. of Copenhagen)—de-centralize Web and return users' control of their own data

IT Facilities & Business Continuity

- CERN presentation on HARRY—improve resource acquisition and allocation process by extracting usage data from existing monitoring tools

- RAL described an on-going migration to new monitoring tools, such as Nagios → Icinga, RT → JIRA, etc

- BNL will occupy a new data center in 2021—presentation described plans to gradually migrate resources to new facility to insure operational continuity

- BNL's SDCC began comprehensive support for Photon Sciences program (NSLS-II)—expect swift ramp-up in the medium-term

DPM BoF

- Attended by ~80 participants
- ~85 sites with storage capacities varying from ~100 TB to ~10 PB, providing a total of ~100 PB (~40% of WLCG Tier'2 resources)
- Decreasing manpower in support of DPM presents challenges for continued support and development
- DPM continued to be used for Run 3, but migration to other storage systems must occur prior to Run 4.
- Discussions on potential replacements among sites. Some are evaluating EOS, dCache, xRootd+Ceph, etc
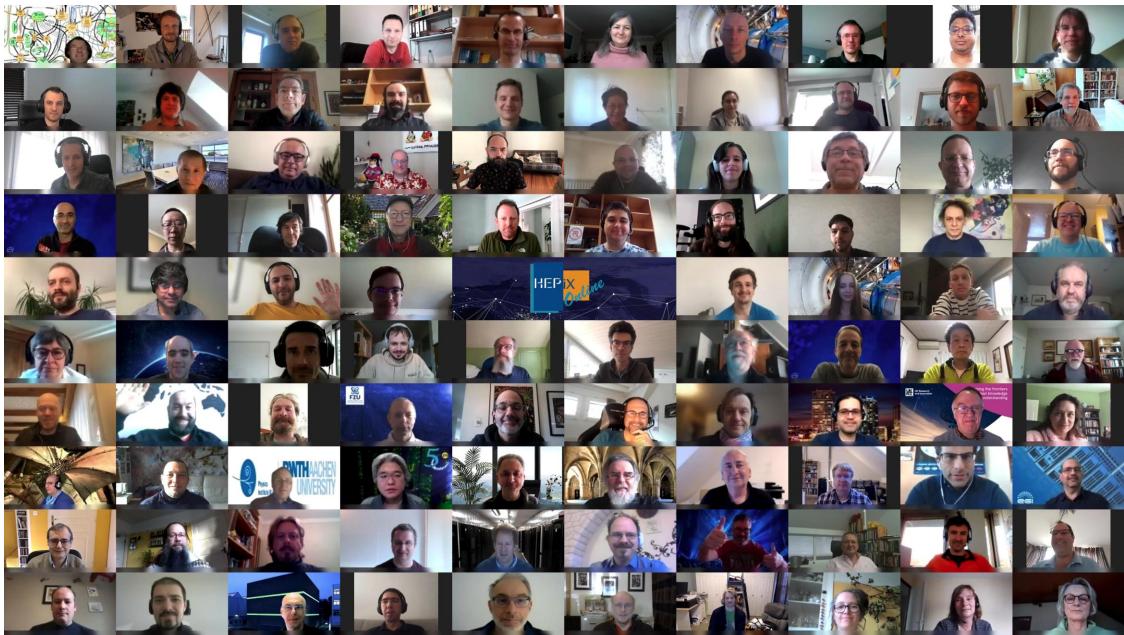- Interest in organizing regular meetings to follow-up BoF discussion

Grids, Clouds & Virtualization

- ITMO Univ. and ZEDEDA described the EVE project to develop and provision an open-source lightweight, secure virtualization engine for IoT edge gateways and servers
- CERN provided an update of its cloud infrastructure
  - OpenStack updates and availability of gpu's
  - Integration with CERN ID management
  - SDN and LBaaS services
- CERN presentation on the design of a Anomaly Detection system for its cloud infrastructure—rely on monitoring data and the development of several algorithms
- FTS3 prototype installation at FNAL deployed on OKD-managed containers—tests for DUNE in the near-term plans
- Shoal—Univ. of Victoria presentation on updates to squid cache publishing and advertising tool in support of grid computing on distributed clouds

Basic IT Services

- Univ. of Bonn presentation on usage of PGpool-II cluster (sits between PostgreSQL DB servers and clients) for monitoring purposes (server room, desktops, environmental, etc) with Zabbix
- Status update on CERN's migration progress away from physical phones to softphone based solution (CERNphone)
- CERN presentation on latest AAI developments
  - SSO based on KeyCloak
  - LDAP/Kerberos service based on FreeIPA (pilot phase)
  - Integration with EduGAIN
- PrivacyIDEA deployment at BNL to help end-user simplify OTP token management for services requiring MFA

Group Photo

Plans for a Fall 2021 workshop

Early Fall timeframe—announcement in due time

We wish you a healthy and safe year, wherever you are, and look forward to meeting you again soon!