Notes from HEPiX Fall 2017, KEK

Jérôme Belleman, CERN

16-20 October 2017

HEPiX was never before held at KEK. But it is a key location in the history of the workshop as the proposal which lead to its birth was presented in the 1991 CHEP conference hosted by KEK. 26 years later, HEPiX Fall 2017 becomes an exceptionally rich meeting of 149 participants worldwide, who gave a remarkable 83 presentations. This event was also special in that it was collocated with the LHCOPN/LHCONE and HUF meetings. The organising committee ensured that all events ran extremely smoothly. The seminar room for HEPiX was cosy and outstandingly well equipped, with a power socket and foldable table at each seat, lapel microphones and a spacious stage for speakers.

Coffee and lunches were served in a large area outside the conference room, and the participants could enjoy technical discussions over a large choice of sweets, an enormous selection of different coffees, and delicious meals. The Monday evening social event took place in a luxurious reception room at the Okura Frontier Hotel in Tsukuba where the organisers arranged for astoundingly good Japanese food to be served, after the Kagami Biraki ceremony where large casks of Sakes were opened. Likewise, the traditional Wednesday HEPiX banquet was held at Tsukuba Sansuitei, who served countless superb dishes of all varieties. The restaurant was built in the middle of a small, typical Japanese garden.

The KEK campus was somewhat far from the three hotels which the organisers carefully chose for the attendees. However, they managed an effective bus system to take us to the workshop venue, all three hotels, restaurants and the fascinating facility tour, all at most convenient times.

 DOI: 10.17181/CERN.KGNR.YH1W

Contents

Monday 16 October 2017

Miscellaneous: Common Session of HEPiX, HUF, LHCOPN/LHCONE

Welcome to KEK (Yasuhiro Okada)

Yasuhiro, the executive director of the accelerator division in KEK, bid us welcome and introduced the research overview. KEK is a laboratory working with the Japanese universities. Its activities extend to many fields, serving 8 000 users/year worldwide, making it a leading science pole. There are 2 campuses, one in Tsukuba and one in Tokai (70 km from Tsukuba). Sciences conducted include fundamental research for laws of nature and the origin of materials. There is also a technical division for developing and operating the accelerators. Yasuhiro presented the SuperKEKB accelerator and the Belle II detector. Belle II was installed in April 2017. The preceding Belle experiment ran from 2000 to 2010 and enabled the discovery of CP violation in B decays shortly after 2000. The goal of Belle II is to reach 50 ab^{-1}. He then showed the J-PARC accelerator and the T2K experiment, shooting a high intensity ν_μ beam from J-PARC MR to Super-Kamiokande. There are already hints that the CP is violated in the neutrino sector. The KOTO, COMET and Muon g-2/EDM experiments are international collaborations. The Material and Life Science facility at J-PARC use pulse neutron and muon beams. Light source facilities are located on the Tsukuba campus. They both have been operated for a long time and are the main facilities in Japan. The strength of material science at KEK comes from the four different beams for the probes which can be used, some of them unique, such as the slow positron one.

KEK is a partner to the ATLAS collaboration. KEK also contributes to HL-LHC and developed a prototype for magnets. The Japanese HEP community fosters hosting the International Linear Collider in Japan. There are discussions ongoing for a 250-GeV ILC as Higgs factory. The ICT at KEK includes the Computing Research Centre and the IN Infrastructure and Security Division. All the equipment (CPUs, disks, tapes, network) is replaced every 4-5 years. The Belle II detector becomes more and more demanding – computing and networking are being upgraded.

Logistics information (Tomoaki Nakamura)

The HEPiX workshop will be held at the Kobayashi hall the whole week. The LHCOPN/LHCONE and HUF sessions will be held in other seminar halls. The HEPiX workshop is broadcast on YouTube and IBM Cloud Video. Tomoaki described the social events and emphasised not to take any food before the Kanpai ceremony. The KEK facility tours includes Belle II and KEKCC visits.

The HEPiX workshop series (Helge Meinhard)

This is the first meeting held in Japan. The HEPiX mission is stated on the website, to foster a learning and sharing computing experience between sites. The emphasis is on the site services and we represent sites more than experiments. HEPiX communities are opening up from HEP to other sciences. The workshop is held twice a year, for a duration of 1 week, loosely alternating between Europe, America and Asia. It is a workshop-style event, with only plenary sessions. Abstracts are usually accepted, with no formal Programme Committee. There are no proceedings. There are about 75 contributions per workshop. Having no proceedings, we can afford to have really honest exchanges. People speak about their successes as well as their failures, which is just as valuable.

Helge described the various tracks, including site reports – short summaries of what happened in the site. There are several active working groups, e.g. on the IPv6 migration, benchmarking and many more. There is a HEPiX board comprising representatives of major laboratories and organisers of past meetings. Helge Meinhard and Tony Wong are the 2 co-chairs representing the community. The board meets face-to-face at the workshop. Next workshops will be taking place in May 2018 in Wisconsin (adjacent to the HTCondor Week). The October 2018 event will hosted in Barcelona.

KEK has a very special role because of the 1991 CHEP conference where there was a discussion about coordinating the introduction of UNIX to HEP. A talk by Les Cottrell from SLAC proposed the HEPiX workshop, which is still a successful model 26 years later.

LHCOPN and LHCONE (Edoardo Martelli)

Edoardo presented the origin and milestones of the LHCOPN and LHCONE communities. 2005 was the time to start implementing the LHC computing model. Site operators met at the first LHCOPN meeting at SARA in Amsterdam. LHCOPN is fully operating and dual-stack today. The network was much more reliable than foreseen, allowing for changes to the computing model. Meetings were held since 2010 to define a new model. There are 80 sites connected. It was in 2007 that the LHCOPN became ready. In 2011, the first LHCONE prototype was implemented. This week, there will be common and parallel sessions with HEPiX.

The HPSS User Forum (HUF) (Jim Gerry)

It is a gathering of HPSS customers, developers, engineers and managers discussing the software road maps. They talk about issues, successes and experiences. Somebody in the room asked, what is HPSS? It is a collaboration for high-performance storage. The HPSS collaboration started in 1992. They were addressing terascale operation. Today, it has become exascale. The collaboration meets once a year, the first time in Asia this year. HPSS is an IBM Software-as-a-Service offering based on tape within a large-scale storage deployment. Redundancy and fast failure recoveries are key to the model. They minimise the impact of silent data corruption. People look to HPSS to take the cost of owning data down, moving more data with less hardware. There is an active HPSS community.

Moving a lot of data with less hardware is done by striping data with rotating parity across tapes. Small files are moved fast by aggregation. Tape mounts are subject to a logic that maximises data transfers.

Questions and comments:

- Jérôme Belleman asked, are disk drives not used? They can be used for caching.

Storage & Filesystems: Common Session of HEPiX, HUF, LHCOPN/LHCONE

Particulate magnetic tape for data storage and future technologies (Masahito Oyanagi)

A talk by Fujifilm. It is expected that worldwide data volume will reach 162 ZB by 2025. Masahito showed a plot of the total capacity by year, in compressed PB. Tape is used for data archive because of the cost effectiveness. It is energy efficient and requires no consumption once recorded. It is reliable for long-time storage. It has got a low error rate compared to other solutions. It offers high capacity storage. The structure of a magnetic tape includes a magnetic layer for data recording, and an under layer to control roughness, a substrate and a backcoat to prevent static charge. Magnetic layers are packed at high density. Trends suggest $>$100 TB cartridge capacities by 2025. Tape thickness is decreasing, now of the order of 5 μm. Work is being done to decrease the thickness of all the layers: the magnetic layer, the surface roughness, etc. The goal is to decrease the volume of the metal particles. Masahito showed pictures comparing the metal particles in LTO-1 and LTO-5. BaFe (Barium Ferrite) has become the standard for tape storage. Reducing the particle volume degrades coercivity. This means size constraints involved in data longevity. But the coercivity of BaFe particles is independent from their sizes. Besides, no passivation layer is needed for BaFe particles. The magnetic properties of BaFe particles are not influenced by their shape either.

An important technology involves the perpendicular orientation. Metal particles are oriented longitudinally. BaFe particles are oriented randomly. A highly perpendicular orientation of BaFe particles contributes to increasing the capacity. Another key of increasing capacity is to decrease particle sizes. The ultra fine magnetic particle technology of Strontium Ferrite will allow a particle volume from 1 600 nm^3 to 900 nm^3 (with a capacity to be announced). BaFe can support the next 10 years of the tape road map. SrFe will enable further capacity increases.

Questions and comments:

- What will be the consequence of the acceleration effect? It is a problem because the rate needs to become faster and faster. The transfer rate given by a tape is higher than that of a disk, for instance.

- Helge Meinhard asked about the Strontium Ferrite technology and its impact on hard drives. Yes, hard drive will be able to support it, too.
- You have described sophisticated material science. There is only one vendor for tapes – Fuji. Any possibility for competition? It is difficult to answer.
- Any consideration given to increase the width? No plan to do so, yet.
- What would be the competing technologies? Tape would seem to be the best solution.
- Strontium Ferrite particles: can they be produced at a similar price or will the media become more expensive? The same manufacturing process can be used, so there is probably no risk of a price increase.
- What are the difficulties in using this material? The size of the particles, and being able to displace them.
- Tony Wong asked, the lifetime of the media is 32 years – what about the availability of tape drives? It is of no use if we are not able to to read the data back.

Site Reports

KEK Site Report (Atsushi Manabe)

The Belle II experiment was successfully installed in April 2017. SuperKEKB with Belle II and QCS will start in January 2018. Atsushi described the accelerator complex on an aerial photo. The J-PARC accelerator was commissioned in 2008 and the beam power increased ever since then. SINET is the Science Information NETwork of Japan, SINET5 started in April 2016 with 100 Gb/s. There are direct connections to Europe, America, Singapore, Hong Kong and more. There are plans for more connections. SINET5 provides an LHCONE instance. Since last summer, the KEK data centre was connected to the HLT farm of the Belle II DAQ system.

Last September, the new central computer system was installed, running 10 000 cores and 13 PB connected with InfiniBand. Atsushi showed a usage breakdown, with most of the resources used for Belle II and J-PARC. In the new KEK data centre, 168 MHS06 hour/month are used. Their cloud is deployed with OpenStack. It is managed on a per-group basis. There is a growing volume of targeted-mail attacks in Japan. There are scams seemingly sent from big companies. A firewall blocks malicious URLs and IP addresses. Notification systems were set up and awareness is raised with user training.

Questions and comments:

- Jérôme Belleman wondered, could you say more about the *Resource Connector*? See talk later this week.
- Tony Wong asked about the InfiniBand interconnect: are compute nodes connected to storage? Yes.

Tokyo Tier-2 Site Report (Tomoe Kishimoto)

The site serves the ATLAS experiment, the MEG experiment and R&D for ILC. The ATLAS-Japan group is made of 17 institutes. The ICEPP analysis centre supports only the ATLAS VO as a T2. It offers 10 000 CPU cores, which can be used in single-core or multi-core mode. Many new cores were deployed early 2016. They heavily rely on SINET5. September 2017 saw LHCONE peering with ASGC, KREONET2 and TEIN.

Their system is upgraded every 3 years. Their migration strategy is constrained by hardware leases and physical space. They made a scaled-down system of 512 CPU cores to support the data migration from the old system to the new one. Tomoe showed a few pictures of the computer room during migration. The data migration took place in 32 days. The new system will support data volumes up to 5.8 PB. The transfer is expected to take over a month. They are considering extending the lease and allowing for some overlap with the new system. Discussions with the system provider are ongoing.

Questions and comments:

- During the migration, everything was removed. Can you reuse old racks? Can racks not stay for longer? The type of racks does change, in fact.
- Those 10 000 CPU cores, are they non Hyper-Threading? This is correct.
- How do you calculate the HS06? The numbers are provided by LSF.
- Is leasing cheaper than buying and maintaining your hardware? How much cheaper? The Wisconsin T2 involves two people managing everything in a 12 000 cores farm. Owning the hardware makes migrations easier. ICEPP believe leasing is cheaper. It is not only about the cost, it is also about the limited manpower.

Australia site report (Sean Crosby)

CoEPP is a group of 4 universities: Adelaide, Sydney, Melbourne and Monash. Funding from HEP computing in Australia appears to be uncertain. They are submitting applications to support experiments such as Belle II. HEP computing is being integrated into the University of Melbourne HPC, which is a petascale computing project. A 78 GPU nodes cluster is replacing the Blue Gene cluster. The site is still the number 1 ATLAS T2 site. There were no new CPU purchases. There were some new storage purchases, however. They had the rule of not running out-of-warranty storage. But they will start doing so now for compute.

They are moving to HTCondor with ARC CEs and are running ATLAS test jobs. They need to work out how to do fairshare between T3 and T2 users. How to set up accounting groups for T3 users is another question. The University of Melbourne is involved in the LHCONE network, using VPNs. Sean described their successful move to Puppet 4, although they still run Puppet 3 on SL6 machines. They are working on cleaning up code with Hiera. Their GPU cluster is used for TensorFlow and Pytorch for rare particle identification, N-body and Theano.

Questions and comments:

- You wrote or rewrote modules – why so many? Were there no available modules? A lot of them were adapted, e.g. the `cernops` one on GitHub. Inheriting was not always an option, adapting was easier.
- Are these changes fed back? For some of them yes.

ASGC Site Report (Eric Yen)

Their centre is running a distributed computing infrastructure, focusing on integration, networking and security. They prefer single racks, which allow for different scales. They operate 24 000 cores, now investing into GPUs too, with 230 000 cores. Some work was done on efficiency. ASGC is also operating the Taiwan to Chicago link. They intend to increase their capacity. AMS and ATLAS are the main users at ASGC. More than a million CPU days were spent last year and they are expecting more. They are setting up a storage system based on CERNBox as part of their DiCOS software cloud stack.

To improve the efficiency of big data analysis, they sample nodes every 10 seconds for monitoring. They check both single- and multicore jobs. If a user submits a single-core job to a multicore queue, they need to identify him. They suffered from jobs using no CPU time by occupying the slots for days. They cross-check with the accounting data from PanDA and HTCondor. Another aspect of improving efficiency entails finding idle CPU servers and shutting them down. They estimated the savings to be of the order of 26%. ASGC welcomes us to ISGC in 2018.

Questions and comments:

- Tony Wong asked, what is the source of idle jobs? They come from ATLAS and could be due to job problems. They check the proxy lifetime. If it runs out the job is killed. This shows in PanDA. Some correlation with the HTCondor accounting would be needed.

IHEP Site Report (Xiaowei Jiang)

Xiaowei listed the major experiments at IHEP, located all around China. The data needs to be transferred to IHEP, petabytes every year. They run an 8.5 PB Lustre instance and a 1.34 PB EOS instance. They store 5 PB in CASTOR. They use HTCondor for HTC, SLURM for HPC. HTCondor was upgraded. The new version fixed some schedd response problems. They are expanding their computing room. The data centre is IPv6-ready. They develop a tool to clean up idle jobs. In their SLURM cluster, they have 1 master node, 1 accounting/monitoring node, 16 login nodes and $>$100 worker nodes. They use Malva for monitoring SLURM. ATLAS@home is used to exploit extra CPU, running BOINC native jobs. BOINC acts as a second scheduler. This approach is being tested at IHEP. Their OpenStack installation is used for user self-services, as a virtual computing cluster and a distributed computing system. Distributed computing is based on DIRAC. Their multicore tests shows 70% efficiencies. A new monitoring framework was integrated based on Elasticsearch and InfluxDB.

Questions and comments:

- You mentioned the ATLAS wall time utilisation is 68%. Is it understood why? ATLAS is currently segregated to only run their jobs. BOINC pulls jobs to increase the efficiency.
- Vincent Brillault asked if the jobs pulled by BOINC had an impact on the efficiency of normal jobs. The priority of BOINC jobs is kept at a low level.

KR-KISTI-GSDC-01 Tier-1 Site Reports (Jeongheon Kim)

KISTI, started in 1962, has different departments. The GSDC department is government-funded since 2009 and is used for data-intensive fundamental research. Every 3 years, they replace storage hardware. They ran almost 4 million jobs in the last 6 months. The KISTI ALICE T1 has provided a reliable service. Their storage service will soon expand to 3 000 TB. They suffered disk test failures related to catalogue errors. They store 3 000 TB on tape. They tested OpenSCAP for Foreman with CentOS 7 and use the Katello system installation.

Questions and comments:

- Jérôme Belleman asked, was OpenSCAP used to check security audits from Foreman? Yes.

Security & Networking: Common Session of HEPiX, LHCOPN/LHCONE

The LHCONE network (Vincenzo Capone)

In the beginning of LHC computing, there was MONARC. It worked nicely, it was hierarchical. This model was more or less overcome by the users themselves. A document written by Kors Bos and Ian Fisk described these changes and introduced unstructured traffic. That is where LHCONE started. It allowed to serve any LHC site. Traffic separation between LHC and other applications was important, either physical or logical. LHCONE is a closed network with a trust relationship. When the first prototype of the LHCONE was introduced, it offered support for Layer 3 with routed VPNs. LHCONE had to offer a new model, with segregated routing tables within the router, too. Vincenzo showed a diagram describing how LHCONE has expanded, covering 40 countries.

An integral part of the LHCONE services are the end-to-end monitoring with perfSONAR. There is also a P2P service being developed – not yet in production. An Acceptable Use Policy was established, covering security aspects, as well as operational ones and responsibilities. LHCONE was opened to new players, not just LHC-related experiments. Sites were involved more and more in the collaboration. Pierre Auger, Belle II, Nova and Xenon joined in, too. TIFR connects to both GÉANT and CERN and makes a good use of the 10 GB/s link. Other networks from all over the world use it, too. In Europe, some providers such as DFN and Jisc make great use of LHCONE. Vincenzo showed usage plots for each of them. The overall GÉANT traffic of the last day reached 100 GB.

WLCG/OSG Networking Update (Marian Babik)

Marian brought up the complexity of LHCONE and the need to understand end-to-end performance. Before the working group existed, the only insight were the data management systems. But the data management systems are quite complex themselves. The first activity was the deployment of perfSONAR. Now the working group focuses on running the perfSONAR infrastructure, counting on a response team to solve identified problems in sites. Version 4.0.1 of perfSONAR was released on 16 August. The perfSONAR 4.1 plan is getting into shape. There are new Grafana dashboards. There are currently 207 active perfSONAR instances, 189 of them running the latest version. There are meshes for LHC experiments, and also Belle II. Marian showed LHCONE and LHCOPN latencies and throughput meshes for a selection of sites.

On the perfSONAR road map, 4.0.2 should be released in November. It will fully support CentOS 7, Ubuntu 16 and Debian 9. There are new port opening requirements as well as new SNMP plugins. They will make it possible to collect SNMP data. Some important dates which are coming are 17 October which marks the perfSONAR 3.5 end of life. In Q1 2018, perfSONAR 4.1 will be released. Q3 2018 will see the perfSONAR 4.0 end of life. Pre-emptive scheduling, which had been requested for a while, will be added to perfSONAR 4.1. The current scheduler, bwctl, will be removed and two-way data latency measurements will be introduced. There are already Docker images for perfSONAR and official Docker support will be added soon.

The current perfSONAR network includes the new mesh configuration functionality, an open source project part of the perfSONAR distribution. It feeds 3 different channels: the OSG data store, CERN brokers and OSG brokers. Six networking dashboards based on Grafana were introduced. The traffic as seen from routers is exactly the same as what `http://netstat.cern.ch` shows. Other measurements such as packet loss, throughput and TCP retransmissions are also on display. The round-trip time and maximum number of hops shown by traceroute are visible, too. A dashboard with the interregional activity is available. The data is stored in Elasticsearch at the University of Chicago. The granularity and availability of testing is kept in the central mesh configuration and can be changed on request. The overall network performance dashboard shows overviews of latencies, packet losses, throughput and traffic.

On the improvement of transfer efficiencies, one of the main goals of the working group is to identify and fix network problems. A dedicated support unit is there to help sites. Issue analysis has significantly improved where perfSONAR could be used. Marian presented some recent cases. The next milestone of the working group will be the 4.1 release. The upgrade has been planned, documentation is being written. There is a link in the references. The working group will commission RabbitMQ for the OSG Network Measurement Platform. They plan to improve Grafana dashboards and plugins will send more attributes.

Questions and comments:

- Consider allowing enough available space for perfSONAR. There were cases of running out of space, causing crashes and needing a complete reinstallation. The logrotate is not sufficient. It would be better to allow for more installation space to begin with. Marian answered that there was a partitioning bug that could not be fixed once installed. This bug should have been fixed and should not appear in new installations. Moving everything away from `/home`, increasing the space and bringing the data back is a way to address this with LVM, Shawn explained.
- Error messages are often unclear when setting up perfSONAR with Puppet, for instance when it comes to creating users. Marian agrees that some documentation should be written on that point. Keeping Puppet manifests updated is also a challenge.
- Pete Gronbech asked for a link to Grafana plots. There are links at the end of the presentation, the dashboards should be accessible without any restriction on IP address or username.

The TransPAC project (Andrew Lee)

TransPAC started in 1998, to provide connectivity between the US and the Asia Pacific region. Andrew showed the logical network map. They are in the fourth phase of TransPAC. They wanted to expand the scope, in addition to just providing the infrastructure to improve deliberate support for network research and help more user communities, and working with regional partners and other organisations. IRNC is the

program that funds TransPAC. It is a concerted effort to link American research and the rest of the world. The NEAAR project connects the US to Europe and Africa. They offer training to provide perfSONAR hands-ons.

Some of the research that Andrew mentioned included DDoS detection and mitigation. They work with APAN communities to ensure that traffic is not routed in odd ways. Andrew showed the top-ten sites who use TransPAC, which includes Stanford, CERNET, APAN-JP and NASA. They also do LHC-related traffic.

Guam is central to the Asian region. GOREX is a consortium of organisations headed by the Universities of Hawaii and Guam to establish an open exchange in Guam. Repairing a cable can take 36 hours, a submarine cable can be a matter of months. There are times when not all the network capacity is available. The NEAAR project links New York, London and Paris in a circuit. Some current and planned networks connect African regions.

AutoGOLE bringing high speed data transfers (Joe Mambretti)

AutoGOLE is related to the LHC point-to-point service; it is a distributed fabric put together some years ago. Every open exchange used a different protocol. A working group was set up to develop an architecture as an API. Is is a layer 2 service and an east-west protocol, covering multiple domains. There are today 29 network service agents advertising 30 networks worldwide. They use services for resource discovery. The group are experimenting with path finding. Why not layer 3? If the current data rate remains the same, in not that many years, 100% that all electricity in Japan will have to be devoted to routing. That is why it had to be taken down to layer 2 to save power. The MEICAN pilot is a set of tools.

Joe showed exchange points around the world, a distributed fabric allowing to define programmable networks. The new Global Lambda Integrated Facility (GLIF) map was published. More and more links became 100 Gb/s. Africa is more and more connected. The basic platform is the foundation on which software-defined capabilities are being defined. On-demand services in single-domain abstract out data to stitch resources and segment them. The MEICAN tools are used for the AutoGOLE to provision, change and remove multi-domain services, as a front-end to AutoGOLE. The LHC community is a driver for them. The interface is intuitive and lets you define workflows, policies and see monitoring. MEICAN is the most mature tool for multi-domain provisioning.

The dashboard lets you manage resources through a graphical user interface. The topology is shown on a map. The reservation tool lets you preregister. In a full circuit, implementation now takes minutes, while it used to take months. The path information is available to display. Since 20 September 2016, 143 circuit reservations were made. Some show cases: transferring 74 Gb/s in a DTN-based network, trying software-defined exchanges, transferring airline data. The goal is now to leave demo mode and move some use cases into production.

Questions and comments:

- Shawn asked, if somebody wants to be involved, are there examples on how to set up the DTN, which hardware, etc.? There is no definition of a DTN. Some are GPU-based, some are CPU-based, some are FPGA-based. Dell now has a prototype. You need a high-speed connection and a DTN.

Next Generation Software Defined Services and the Global Research Platform (Joe Mambretti)

A common trend these days is to have everything software-defined. The Global Research Platform (GRP) is a science DMZ. It is a specialised, globally distributed platform which isolates environment from commodity resources. It leverages advanced architecture concepts such as SDN. One-size-fits-all models are not optimal for many applications and services. Joe showed the Pacific Research Platform as an example of where to develop DMZ islands.

It has been proved empirically that SDNs helped improving the reliability of services. How about having a worldwide distributed science fabric? There have been several regional platforms discussed already. Some of the network research topics include transitioning from legacy networks, extremely large capacities, TENET networks, network virtualisation (which is accelerating at a high pace), network programming languages such as P4. Telemetry was made possible to investigate real flows with P4. Jupyter is a scientific

workflow management tool. GENI allows experimenters to implement their own protocols. Springer publishes the *The GENI Book* which describes all its aspects. Joe went through a number of software-defined projects.

Questions and comments:

- Shawn wondered how to glue all these projects to push towards a global research platform. It is a good question, we do not have a service. Specifically, the LHC community might be a good place to start, looking at the services. Edoardo has provided a DTN at CERN.

NetSage, a unified network measurement and visualization service (Andrew Lee)

It is a software project, more than a network infrastructure one. The original charge of the project was to get a better idea of how the infrastructure is being used. The NSF spends millions a year, and understanding how is instrumental. The goal is to project a simple-to-use portal. Other users would be the IRNC NOC and the various network operators to be able to do better planning. Beyond creating a web site people can look at, part of the goal is to create a framework which other projects can expand on when setting up their own measurements. SNMP is the obvious data source, and flow data, Tstat and perfSONAR will be interesting, too.

Andrew showed an example of flow analysis over a week's time. One user had a TCP session with an 8 MB/s transfer for a week. But looking at the flow data, it was not clear what was going on. The data could be used and correlated with TCP data.

The NetSage portal lets you drill down individual circuits showing the bandwidth, also as histograms. You can tell how often a transfer spikes, which can be useful for doing capacity planning. Another visualisation which they developed were heat maps, to cross-compare time of the day and day of the week, for instance. This can help network operators determine when best to run their network maintenance. Beyond just looking at SNMP data, it would be useful to map IP addresses and autonomous system numbers to organisations, science projects and science domains. Flow data is de-identified at the source. Data that is not de-identified never makes it to the archive. Collecting data is an iterative process, where the decision to collect new data is made during visualisation. They use the ELK stack mainly to do prototyping. The Sankey graph is an easy way to visualise flows between countries.

Questions and comments:

- Edoardo Martelli asked if NetSage is open source. Yes, along with a number of other projects shown during the presentation. They provide support, too.
- Stefan Stancu asked about the software tools to make visualisation, especially heat maps. Kibana, mainly.

Tier-1 networking cost and optimizations (Mattias Wadenstein)

NORDUnet is the Nordics T1's network supplier. Mattias described the various connections. NORDUnet have asked for external funding to pay the full cost of networking. But that is only NDGF, because other projects are part of universities. Networking costs are a significant part of a T1, of the order of 20% of the total cost. The LHCOPN is straightforward to pay their fair share. The predictions and technical developments keep ahead of usage needs. The LHCONE and transit costs are rising. The GÉANT IP transit is twice as expensive as the commercial transit, which is surprising. There are unnecessary costs due to some transits and some ways to cut them is to add direct connections. A solution will be needed as data volumes keep increasing.

Questions and comments:

- Many assumptions were made, which do not relate to GÉANT. The so-called GÉANT transit is not a cost that GÉANT impends on users/NORDUnet. There seems to be a disagreement on that point.

Network Functions Virtualisation Working Group Proposal (Marian Babik)

HEP have significantly benefited from LHCOPN/LHCONE projects. There is an increased usage and to continually expand capacity, the working group would like to propose concrete steps. They would have a mid-to-long term scope to cover a specific area of interest for the HEP community. Since the beginning, LHC experiments have been transferring exponentially increasing amounts of data, as we saw today from GÉANT and ESNET numbers. One challenge is the capacity – there is no immediate issue but there cannot be so many HEP-scale science domains to compete for the same resources. Remote data access is proliferating. We integrate commercial clouds. Technology evolution is such that we will need to change our systems along the way, which may be hard.

The first impact is that there is an increasing importance to oversee network capacities. Software will play a major role in networks. Marian then focused on Network Functions Virtualisation (NFV) technologies, which appear to make it to the data centres soon. Software switches include Open vSwitch, an open source solution supporting standard interfaces and protocols. You can turn any network interface into a switch. It was primarily meant for VM to VM communication but has since expanded its usefulness. To achieve this, controllers are needed. Open DayLight implements such a controller. Core use cases include clouds and network resource optimisations. There is a full range of available controllers. Open Virtual Network is one of them, and it comes with Open vSwitch. It provides level 2/level 3 networking. One of the major clients is OpenStack. Solid networking was added to OpenStack at a later stage, especially with Neutron.

An alternative to Open vSwitch is Cumulus Linux. It is a Linux distribution which uses kernel functions and applications to implement network capabilities. It does contain Open vSwitch but does not use it by default. It is certified to run on bare metal. Being a regular Linux distribution, you can make it part of your usual Linux stack.

The motivation of the working group is that massive and large-scale network automation is possible. There is a potential for cost reduction, in avoiding vendor-specific solutions. OpenStack and Docker are leading the way. Some of the T1s have made some exercises with this. There is a number of options to choose from, and hearing the community's experience would be interesting. We should understand together what the requirements are. The proposal would be to start a new working group to engage sites and experiments in testing, deploying and evaluating network virtualisation technologies.

Questions and comments:

- Pepe Flix asked, how do you plan to engage (N)RENs? Initially, it should be focused on sites only, in fact. RENs would come later, especially because they already have projects of their own.
- Michel Jouvin asked who would be creating the working group. Shawn answered that it is something to discuss now. Michel Jouvin added that the HEP Software Foundation might be the right choice, as it publishes the white paper. Ian Collier wonders if it is only sites and experiments. HEPiX is a little too focused on sites for this. Helge Meinhard agrees with Michel it should not be limited to sites and experiments. HEPiX should be the right forum to start. Dave Kelsey added that the experience with the HEPiX IPv6 Working Group could help, too.
- Stefan Stancu thinks there are two aspects to keep in mind: how the technology scales, because everybody talks to everybody in a HEP environment; how do we troubleshoot, bearing in mind that the infrastructure is already complex as it is.
- Will it be for production networking? Yes.

Computing & Batch Services

HEP Community White Paper (Michel Jouvin)

The HEP Software Foundation (HSF) started in 2014. We were facing new challenges, for instance related to HL-LHC and other experiments such as Belle-II generating a lot of data. If everybody tried to reinvent

the wheel, it could not scale. The idea is to coordinate efforts and expertise. The goal of the white paper was for the community to describe a vision for the HL-LHC era. The need is to prioritise activities. Several working groups related to software and computing were active during the community white paper process. Several workshops were held. Each working group produced a white paper chapter. They are organised as Google Docs where we can easily comment.

It is a white paper, where the point is to get feedback from the community. It is a bottom-up process. Everybody is invited to take a look. We are close to the deadline for finishing the global community white paper. The first draft is expected at the end of this week. Some chapters relevant to the network community cover the computing models, distributed computing, data management and data analysis, which describe new models that could impact how resources are used in the future. This white paper is a unique attempt to unite the community to think together. There is still time to provide feedback.

Questions and comments:

- Pepe Flix asked, some of the papers are LaTeX. Are comments supposed to be sent to editors? You can use ShareLaTeX to edit or simply send a mail.

Tuesday 17 October 2017

Site Reports

CERN Site Report (Jérôme Belleman)

After presenting CERN and its IT department, Jérôme went through the latest of the major activities from each team. The AFS phaseout at CERN continues. It appears that users tended to use AFS for purposes it was not designed for. Now is an opportunity to choose the right tool for each purpose: the upcoming EOSFUSEx will be better at performing small writes, copy-in and copy-out should be left to batch systems such as HTCondor and CERNBox is suitable for offline operations.

CERN CentOS 7.4 has become the production version at CERN. SSSD now supports Kerberos authentication with smart cards, Flatpak was added as a new means to deploy software and Btrfs was dropped. A CentOS Dojo day will take place at CERN this week. Windows 10 is the production Windows version running at CERN, Windows 7 support will end in 2020 and Office 2016 is being rolled out. To respond to the WannaCrypt threat, SMB1 is being disabled on file servers.

The CERN cloud continues to expand, with 281 000 cores on 33 000 VMs, running on 8 000 hypervisors across 3 600 projects. More than 90% of compute resources are now in the cloud. There are 120 Magnum container clusters. OpenStack was upgraded to Pike and CentOS 7.4 was deployed. There was the migration to Neutron and work has been done on bare-metal provisioning and File Share as a Service on top of CephFS. Coming up next, containers on bare metal and GPUs in the cloud.

The PuppetDB service received some performance improvements and the team are now focusing on the PuppetDB 4 upgrade. New tools are being looked at for managing secrets and some minor updates were applied to the whole infrastructure. The DNS load balancing software was rewritten in Go to adopt a concurrent approach. Interactive login services received a fairshare setup with cgroups and counters were configured to collect overall usage data helping debug problems and produce accounting. Collectd is being rolled out more and more. The Elasticsearch service covers more use cases and uses ReadonlyREST to manage ACLs.

The migration from analogue to soft phones will be completed end 2018. Clients are available for all major platforms. Useful features include contacts, video calls and instant messaging. CERN now has an Office of Data Privacy Protection, to help the organisation align to best practices. Service owners are to write privacy notices. WLCG adopted the new Policy on the Processing of Personal Data.

- Xiaowei Jiang asked whether all the jobs run in containers. They do not, many run on virtual machines.
- Tony Wong asked about the $>$90% computing resources in the cloud. Indeed, most of the batch cluster, for instance, are virtual worker nodes.

PIC report (Jose Flix Molina)

PIC provide 85 kHS06, 7.5 PB on disk, 23.5 PB on tape. It is the Spanish WLCG T1 centre. Significant CPU and disk purchases are ongoing. About 25% of the compute capacity is running HTCondor as the scheduler and the migration to this batch system is ongoing. IPv6 worker nodes are available for testing. Puppet 4 runs in production. An interesting integration prototype between Barcelona and Madrid was launched.

The HTCondor farm comprises 82 WNs, 1 540 slots, 21 kHS06. There are some problems related to memory leaks and the only solution is currently to restart services. There were several instabilities with dual-stack and IPv6 only. The IPv6 deployment is ongoing. The oil-immersed servers will come last because of some infrastructure work that will need to take place. PIC consider upgrading their router, during LHC LS2. They greatly improved disk pool performance. SAM tests are run in a dedicated dCache pool. They moved dCache metadata to SSDs and introduced tape recall-only pools. Pepe showed considerable data rate improvements

They are deploying Grafana at PIC, building dashboards from metric samples they collect with Ganglia. Pepe praised the speed of Grafana. PIC are working on a more flexible WLCG resource usage by linking

sites in Spain. HTCondor pools are being deployed with flocking. They are integrating gateways for CPU intensive tasks, something they will discuss at the next HTCondor Week.

Pepe presented CosmoHub on Hadoop. It holds the biggest virtual galaxy catalogue to date, 7.4 billion galaxies, 1/8 of the sky. For LHC analysis, work was done to move a real ATLAS analysis program from ROOT/C++ to Python with Hadoop/Spark. The prototype showed that a real analysis program can easily be adapted to Hadoop/Spark. The execution time has been reduced.

Questions and comments:

- Jérôme asked, could Hadoop/Spark replace batch systems in the future? Hard to tell. It is more for quickly testing things out for the moment. People like it because they can use Python.
- Why do you use IPv6-only on worker nodes? For testing.

BNL RACF Site Report (Ofer Rind)

BNL provides full computing services for the 2 RHIC experiments – STAR and PHENIX – and ATLAS as T1. BNL reached an agreement to host a T1 centre for Belle II. They are currently provisioning compute, storage and infrastructure resources for this experiment. Some of the previous stability issues with Knights Landing are now under control.

They purchased new equipment for their Linux farm. They use different storage and network interconnects for the RHIC and ATLAS systems. It will bring their resources to 65 000 logical cores. They already repurposed some servers for Belle II. There are plans to update Linux to SL7. Condor will be upgraded to 8.6 at the same time. Singularity is used to provide SL6 compute containers for ATLAS, Belle II and RHIC.

Earlier this year, the US Atlas T1 facility was migrated outside of the BNL campus network. The migration to the 25/50/100 GbE switching infrastructure has begun with a new ATLAS dCache and T3 GPFS storage. Specifically, a new dedicated GPFS cluster was commissioned for ATLAS T3. The ATLAS T1 recently added some users, now reaching 17.5 PB of unique data. Some of the ATLAS dCache hardware was deployed for Belle II. STAR runs 8 PB of local storage. They operate 3 Ceph clusters: the main one is for BNLBox, the other two are for ATLAS production and testing. Centralised storage is ensured with IBM Spectrum. They are adding a BNL Stratum 0 for CVMFS to provide a repository for local ATLAS production jobs. They have been running HPSS 7.4.3 and just got past 105 PB. The setup is ready to take Belle II data. STAR and ATLAS currently run LTO-7. The migration of LTO-4 to LTO-7 reclaimed 11K cartridge slots.

Questions and comments:

- Michele Michelotto asked about Singularity to run ATLAS jobs, but does ATLAS not run on SL7? No, some certifying is still needed for this, therefore the need to use Singularity.

INFN-T1 site report (Andrea Chierici)

INFN has got a new logo. The INFN T1 WAN upgrade to 100 Gb/s will start at the end of the year. Andrea showed the connections to LHCOPN/LHCONE. IPv6 deployment is ongoing. In particular, dual-stack on CEs is working. They run 23 PB on disks, 56 PB on tapes and CPU services for experiments. The 2016 tender is now in production but the 2017 tender is behind schedule. They had to refurbish in-house 2 old DDN storage appliances.

INFN is running into issues due to the cost increases of software licences. They are trying to find new agreements with IBM, and are investigating alternatives. They reached 220 kHS06. Compared to last tender, they asked for SSD disks, fast network and at least 16 physical cores. There is only one bidder, however. In 2018, many resources will go out of warranty. They will borrow from the CINECA *Marconi* cluster. They suffered from the LSF *Job exit threshold exceeded* issue. They will not invest too much time on this since they are moving to HTCondor. They are investigating grid middleware on CentOS 7. BDII seems ready.

They spent time trying to get the ISO 27001 certification for a subset of racks in the computing room. They will use this certification to host experiments with sensitive data. They are finalising their cloud

deployment. There are doubts as to the responsibility in case a cloud machine is misused. They are trying to start an HTCondor pilot, with HTCondor-CE to follow. They will upgrade from Puppet 3 to 5, thus skipping Puppet 4. They are waiting for Foreman 1.16 to do so.

Questions and comments:

- Jérôme Belleman asked, did the *Job exit threshold exceeded* problem appear only this summer? Yes, it is just a few users, who might have just changed their habits. They had to be banned a number of times.
- Jérôme Belleman asked, does this ISO certification have to do with data privacy? Yes, it is particularly complicated because system administrators have too much access rights.
- Michel Jouvin pointed out GPFS licensing problems. Did BNL see any problem? Not particularly. Mattias Wadenstein described a drastic increase indeed, which then dropped back to more reasonable levels following negotiations.

KIT Site Report (Andreas Petzold)

The T1 batch system now only runs HTCondor with ARC CEs and Andreas showed the ramp-up on charts. The operation was relatively smooth, even though there was a lot to learn. Multi- and single-core scheduling is constantly being optimised. One problem was the empty ATLAS pilots. The aCT (ARC Control Tower) improves the situation. There were stability problems with ARC CEs. They now run 850 nodes, 320 kHS06, 24 000 single-core job slots. They only run Intel CPUs except for a few AMDs for benchmarking.

KIT manage 2 Oracle SL8500 libraries and 1 IBM TS3500. There were several problems with T10KC drives related to *insufficient write efficiency* and a *stuck roller*. These problems are now caught by the latest firmware builds. They plan to migrate from TSM to HPSS for their T1. The T1 network uses 2 border routers. They run several lines to LHCOPN/LHCONE. The new online storage system was installed in later 2016 and runs in production since April 2017, with 20 PB for GridKa and 6 PB for other KIT use cases. Extensions were installed and more are coming. NEC is used as system integrator. The storage system for T1 runs 13 NEC SNA660, NetApp appliances, NSD servers, Mellanox switches and protocol servers for NFS/dCache/XRootD.

On the GPFS front, they keep 4 copies of the metadata. They have 1 or 2 filesystems per VO. It is not one whole filesystem, because they want to segregate VO workloads. Each VO's protocol server performs a remote mount of the filesystem. Their GPFS contract, although not cheap, is a flat-cost contract and does not depend on usage. The ATLAS dCache/XRootD pool has reached 1 PB, which has become challenging – it takes 6 hours before it becomes read/write. The 13.5 PB of old DDN storage will be switched off by the end of 2017.

RAL Site Report (Martin Bly)

RAL reached a capacity of 240 kHS06, 16.5 PB in CASTOR, 13.3 PB in Ceph, 80 PB on tapes. Martin showed the results of some benchmarking from their 2016 procurement. They noticed a few nodes which were running faster, with no explanation yet. There were no significant changes in the T1 WAN/LAN. IPv6 is available on the T1 network, with dual-stack Squids since early August, CVMFS Stratum 1 since mid-August and FTS by the end of the year.

The migration of their batch system to SL7 completed, using the HTCondor Docker universe to run all jobs in containers. The access to the Ceph system happens through a Ceph gateway and proxies each running in containers. CASTOR 2.1.16 was deployed across the site. The Puppet-based configuration management infrastructure was entirely decommissioned in favour of Quattor/SCDB.

They have a Ceph cluster, disk-only storage replacing CASTOR for LHC VOs offering 11.3 PB of raw space. ATLAS have been using it in production since April and CMS are testing it out. Martin showed good-performance plots over 2 weeks, including the balancing. They are moving from Hyper-V to VMware. The testing is nearly complete. They need to migrate off Red Hat 5 for databases. SL6 is pretty much only used for CASTOR.

NDGF Site Report (Mattias Wadenstein)

NDGF is a distributed WLCG T1 site across Denmark, Finland, Norway and Sweden which supports ALICE and ATLAS, getting networking from NORDUnet. CPU is commonly on HPC systems, with ARC as front-end. Storage runs on dCache, tapes are managed by TSM, often collocated with the site backup services. Configuration management is being moved to Ansible.

NSC suffered from an outage reported during HEPiX Spring 2017, caused by a loose nut and bolt for a power rail in the power substation feeding one of the data centres. They are getting more disks for NDGF-T1. An HPC cluster procurement is ongoing for x86-64, Power, ARM, Knights Landing and GPUs. Solar panels were installed, peaking at 225 kW on a sunny summer day and expecting a yearly average of around 25 kW. A cooling outage occurred, caused by the PLC which crashed. Restarting the PLC helped but worker nodes had to be turned off. A couple of servers crashed or shut themselves off, some InfiniBand switches gave errors but hardware otherwise survived.

HPC2N installed a new tape library, using the same TSM tape software. In the future, the hope is to be able to abandon fibre channels. Mattias invited the community to help convincing IBM to improve tape support.

PDSF Site Report and Transition (Tony Quan)

NERSC support experimental and observational science projects since 1996, including ALICE and ATLAS. They are in the process of procuring NERSC 9, their next generation supercomputer, in addition to Edison and Cori. PDSF, the Parallel Distributed Systems Facility offers high serial throughput, using the Univa Grid Engine batch system. ALICE, ATLAS and STAR make 85% of the usage. Tony showed some photos of the site. They developed CHOS to allow multiple Linux environments to run within a single Linux system. The container environment they use is called Shifter. They moved some years ago to SLURM as workload manager. The goal would be to move users to have their own software stack inside containers. They will keep 2-factor authentication.

Questions and comments:

- Tony Wong asked if all the PDSF is physically migrated. Yes, all the PDSF has been migrated to the new building at the lab.
- Tony Wong asked if containers will become part of their general support. Yes.

Swiss National Supercomputing Centre – T2 Site Report (Dario Petrusic)

CSCS moved to a new centre in 2012. The main systems are Piz Daint (a Cray XC40/CX50), Blue Brain 4 (an IBM Blue Gene/Q), Kesch and Escha (two Cray CS-Storm) and Phoenix (an LCG cluster). They use the 9°C Lugano Lake water for cooling. They rely on an IB network. Their Spectrum Scale filesystem is used as scratch area. They run several IBM Spectrum Scale systems, a Lustre filesystem, an IBM TS4500 and dCache. They support 3 VOs, offering ARC CEs interfacing to separate SLURM partitions. They use DataWarp for swap and a CVMFS RAM cache on Piz Daint. They use SAN directors to manage their storage. Dario showed pictures of their installations.

Questions and comments:

- Jérôme Belleman asked, can you say more about DataWarp? It is a Cray caching technology based on SSDs which CSCS use as swap.

AGLT2 Site Report (Shawn McKee)

Shawn showed some numbers since last HEPiX: 6.85 PB of storage and a total of 96.1 kHS06. They have a nice custom front page for monitoring, in front of Ganglia, Elasticsearch, Logstash, Kibana and SRMwatch. They use GLPI to track tickets. They are trying to revive the Dell ATLAS portal. The focus for upcoming purchases moved from storage to computing. They anticipate adding 10 kHS06.

They have been rebuilding VMs to use SL7. dCache was updated to 3.0.28. HTCondor runs 8.4.11. The two major updates are on Lustre and dCache/Open vSwitch. Lustre at AGLT2 is used for local users, about 1 PB of space. The data read rate shows a significant increase of the order of ×2. AGLT2 have been planning to implement Open vSwitch to provide a mechanism to test SDN for real LHC physics tasks. As of the beginning of this month, all the AGLT2 dCache storage servers are running SL7.4 and have their public IPs on Open vSwitch. It has been running for 2 weeks without problem. AGLT2 are now ready to start LHCONE P2P tests.

They will be participating in SC17, experimenting with SDN/Open vSwitch in their T2 as part of the LHCONE P2P test bed. They are working on IPv6 dual-stack for all nodes in the T2 and moving all worker nodes to SL7 in the next 2 months.

US ATLAS SWT2 Site Report (Horst Severini)

The centre is shared between the University of Oklahoma, University of Texas Arlington and Langston University. Their oldest T2 centre is a 72 nodes (844 slots) cluster. The T3 HTCondor cluster runs 275 cores for Athena, ROOT and theory calculations. The biggest T2 SLURM cluster comprises 41 nodes (1 800 slots) with 500 TB of XRootD storage. Lucille is a 30 nodes (960 slots) HTCondor cluster. The UTA T2 runs two PBS/Torque clusters, one of 439 nodes (8 736 slots) and one of 110 nodes (2 360 slots). Horst described the various connections to I2, ESnet, OneNet and LEARN.

Their 10 000 cores Schooner cluster implements fairshare and they expect good usage of the resources. The plan is to dynamically shorten the requested maximum wall time until they start, but the jobs will not be pre-empted. It is a simple loop script which implements this. Their container deployment runs Singularity 2.2.1. They almost successfully ran test jobs, but still need to solve problems related to bind-mount and inaccessible certificates in the HTCondor pool directory.

University of Wisconsin-Madison CMS T2 site report (Ajit Kumar Mohapatra)

There has not been much change in the infrastructure since the last site report. They still use pod-based hot aisles and air coolers using chilled water. They run an SL6 cluster of 11 500 cores (114 kHS06). More cores will come, for which they are hoping to add Intel Skylake-SP chips. They run storage on Hadoop, which they have done for over 6 years. The instance currently holds 6.2 PB of raw data.

Their HTCondor-CEs run both SL6 and CentOS 7. The job batch queue is managed by HTCondor and has been supporting full multicore since 1.5 years. Singularity ensures a better job management. It removes user switching and acts as a replacement for the gLExec-based solution. They have extensively used pool accounts as a key feature of GUMS. They have been using Singularity for the last 6 months without issues. Their Hadoop cluster has been reliably running for one year. The transition from non-high availability to high availability was carried out with extensive testing using Puppet.

IPv4/IPv6 is configured on all nodes. GridFTP, XRootD work with IPv4/IPv6 but Hadoop only supports IPv4. They are connected to the LHCONE network. They are fine-tuning the cluster upgrade to CentOS 7. They manage the clusters with Puppet and Foreman, use Nagios/Ganglia for monitoring and perfSONAR 4.0.1 for investigating network latencies and bandwidth. OSG are planning to retire Bestman2/SRM soon and the LVS/IPVS-balanced GridFTP replaces it. The real server selection is carried out with round-robin. The setup is being exercised with PhEDEx. Benchmarking new hardware is still happening. With some fine-tuning, they now detect drives with errors. OSG BDII was turned off.

Security & Networking

Recent network connectivity around KEK (Soh Suzuki)

The outer connections are SINET5 providing 100 Gb/s, IBBN operated by the Ibaraki prefecture and a commercial ISP. Inside the campus, KEKCC and Eduroam/guestnet are the available networks. The campus network is all about security, because many devices want to connect to it, wireless and wired. The high-performance computing network is geographically compact.

Externally, the LHCONE connection is extended to Asian countries. A new player, JGN, is an NREN in Japan, mostly used for research activities. It is operated by the National Institute of Information and Communications Technology (NICT). As a result, the LHCONE VRFs saw the addition of JGN – not in Japan but Hong-Kong. They requested a pairing to JGN because of the 100 Gb/s bandwidth. Before, they sometimes had traffic problems. They can now enjoy LHCONE connectivity to Korea, China and Taiwan.

Internally, the Belle II DAQ system has intranets under the KEK campus network. These intranets are not directly accessible from the campus network. The KEK campus network and KEKCC have independent filtering policies and firewalls. Soh described the DAQ system and the frequencies of the different event builders. The readout systems have 50 links of 1000T. Event builders reassemble them into HLT farms via 10 GbE. 30 Gb/s is faster than the capacity of the firewall and a dedicated path is necessary. Each HLT unit has a 10G-LR connected to KEKCC. They transfer raw data such that security incident spreads are kept to a minimum. Systems suspected of a security incident are shut down.

Network related updates in IHEP (Zhihui Sun)

Zhihui described the networks and bandwidths from IHEP to the USA, Europe and Asia-Pacific. Last August, the new network architecture finished. A switch connects the campus network, the data centre and the remote sites. IPv6 is now ready. Commercial DNS solutions are being evaluated for DNSv6. They are considering ISC DHCPv6 based on IPDB. Their T2 saw IPv6 reach production. LHCONE is in progress in China. They need more support from the network operators.

They run two perfSONAR hosts for wide area network tests as well as for their local area network. They deployed 13 measurement points in the campus network. They also deployed a monitoring box for data centre network. The latency and packet loss test results looked very positive. They developed a monitoring system for WLAN. Network security operations at IHEP were presented during ISGC 2017. The China Cyber-Security Federation met in June 2016 and the first workshop on cyber-security for HEP in IHEP were helped last week. They are developing a network traffic analysis system. The processing flow includes data acquisition, storage with MySQL (considering MongoDB), analysis and display.

Questions and comments:

- Dave Kelsey remarked upon the increased traffic when they turned on IPv6.

Netbench –- testing network devices with real-life traffic patterns (Stefan Nicolae Stancu)

Netbench is a tool that they used to test devices with real-life traffic. Not many of us in the audience ever had to test a switch against traffic. The performance is key to the services relying on them. When you upgrade your infrastructure adding new devices, you need to ensure the features and performance are there. It is about testing all present and future connections. But plugging all the ports can prove expensive. With commercial testers, you do not exercise buffering. The tests are also performed by the manufacturers themselves and one might be wondering how biased they are. A typical, affordable test is the snake test, involving the use of 2 tester ports. It exercises all ports. But it still does not test buffering. Netbench can do better, using commodity servers. It uses iPerf3 for sending TCP flows. This exercises buffering too, by sending multiple TCP flows. A reasonably-sized test bed becomes affordable.

Servers run a Python agent which the user interacts with. Once data flows run, they are recorded into a PostgreSQL DB and eventually plotted with JavaScript libraries. Graphs plot the difference between expected and observed flows. They can plot the node bandwidth, too. Stefan described the different plots and how to read them. He then explained how to use these results for tendering. A TCP flow fairness evaluation can be useful for this, by running freely a TCP test. In a certain case, they could identify ports being underperforming. Heat maps could show the level of unfairness. Stefan then showed plots and heat maps of what a good device should look like. He presented a comparison matrix of specialised testing hardware, the snake test and Netbench.

Questions and comments:

- Tristan Suerink suggested checking out Warp 17. It makes you test different types of traffic at top speed. You can mix TCP and UDP.

- How do you test many sessions? We do not, with iPerf3 you cannot. The question is, when, in real life, do you get many small sessions? There is a vendor that built 32 ports switches, but can only do 2 million packets per second per port.
- Jérôme Belleman pointed out that other products with that name are visible on the web. The Netbench in question here is currently CERN-internal.

Deployment of IPv6-only CPU on WLCG – an update from the HEPiX IPv6 Working Group (Dave Kelsey)

Between 2011 and 2016, preparatory work took place. In 2011, the IPv4 addressing started to run out. A full analysis of the work had to be done. A worldwide test bed was available early on. By 2012, CERN started to announce that they were running out of IPv4 addresses. This seemed to be happening earlier than later, because of the plan to virtualise aggressively. At the CHEP conference in Amsterdam in 2013, more than 2 PB of data was transferred over IPv6. The most important service to be dual-stack was data. dCache and XRootD for instance went through several iterations before it was decided it was IPv6-ready. In CHEP 2015, we reported that 75% of T1 sites were IPv6-ready. The IPv6-only use case was looked into. By 2016, we entered the production phase. We pushed for the deployment of IPv6, also perfSONAR to involve new sites, as a good way to get experience. It became useful to understand what each experiment's timetable was to define their use cases.

The approved timetable in September 2016 was that April 2017 was the point when IPv6-only CPU should become possible. T1s had to provide dual-stack storage. The CVMFS service at CERN had to be dual-stack. The next point is April 2018, where T1s are required to provide dual-stack storage in production, CVMFS and FTS dual-stack. By the end of Run 2, a large number of sites are required to migrate their storage to IPv6. ALICE in particular needs this before they could use IPv6-only CPU. We are now in a situation where most T1s now have IPv6 peering with LHCOPN. Dual-stack storage exists in some places.

ALICE have got a web site monitoring IPv6 readiness. What is worrying is that the number of SEs and CEs supporting IPv6 does not seem to be changing much. Dave showed a screenshot of the web page. CMS added a DNS test to their production experimental test framework. It is just about looking in DNS whether there is a record or not. They also developed a storage test for XRootD. That would run from the IPv6 instance of ETF. LHCb agreed that they will monitor in the same way as ALICE. ATLAS reported only very few operational problems with the machines allowing jobs to run on IPv6-only worker nodes.

The storage at the T0 – EOS – is an important component. It is based on XRootD. It has successfully been running with IPv6 in production. Enabling it was only a matter of running EOS on an IPv6 node. IPv6-only CPU testing is being done by various members of the working group. For instance, PIC in Barcelona. This is to make sure that everything does work. ATLAS, CMS, LHCb have successfully run tests, even though it is not always stable. QMUL have been running some IPv6-only nodes behind NAT64. It is useful to list which services are contacted by a particular job.

We want to move away from filling in tables and questionnaires asking if we are ready. The FTS dashboards can for instance show whether a transfer was carried out with IPv6 or IPv4. The other monitoring is of course perfSONAR with its dual-stack mesh. There are ETF IPv6 instances which provide IPv6 testing support for SAM, too. Some problems were noted, e.g. some intermittent IPv6 connections which were dropped between SARA and Imperial London. CERN planned to turn on IPv6 on VMs by default but it was delayed (until January 2018) because of a router bug. The question of whether we might hit problems with Docker containers and IPv6-only support arises.

The HEPiX IPv6 Working Group continues to meet. It is a productive environment with participation from all LHC experiments. Everybody are welcome to join the meeting. WLCG T2s need to start planning for IPv6 now.

Questions and comments:

- Is there a recommendation for a given service as to which of IPv4 or IPv6 should be enabled first? The recommendation is to choose IPv6 whenever possible. Mattias Wadenstein commented that this is mostly a choice for software developers. Most sites do not have that choice.
- When do we have to have IPv6 for CPU? It is entirely up to you. If you have clients who do not produce anything for anybody else it does not matter. Mattias Wadenstein commented that by

the time storage becomes IPv6-only it might be required to support IPv6. Ian Collier described a situation where IPv6 is required, when new sites are IPv6-only and VOs require to access their data.

- Tony Cass asked, are you a working group or a standing committee? Dave asked the working group this. It should finish very soon, certainly, but there is still work to do.

Using Configuration Management to deploy and manage network services (Quentin Barrand)

Configuration management had not so far been used in the communication group. The two services to puppetise are the RADIUS services for Wi-Fi, namely eduroam and the visitors Wi-Fi service. Quentin described the eduroam service. Supposing a CERN user and a visiting user, the CERN user will reach out for a Wi-Fi access point, then to the RADIUS servers, while the visiting user will in addition be identified through eduroam. A third case is that of a CERN user abroad contacting the RADIUS service directly.

When visitors register, they send a HTTP request contacting a Wi-Fi controller taking them to the visitors captive portal, which sends an SMS authentication. The RADIUS server somewhere halfway acts as a proxy. The current RADIUS servers were hand-managed. Reinstallations were painful. The visitors portal requires a recent FreeRADIUS with a REST back-end. There has to be a `clients.conf` updater and a monitoring script. In addition, eduroam needs a FreeRADIUS Oracle back-end. That is all there is to it.

The common features are literally identical and can be grouped inside Foreman hostgroups. The visitors and eduroam services will have distinct sub-hostgroups. Installing FreeRADIUS is easy as it comes with a CentOS package. However, it is rather old. Besides, the certificate verification is broken in that old release. The Puppet community have written modules but all relying on the package. So the CERN networking group chose to build the package themselves. Upgrading FreeRADIUS or adding the REST module is a matter of adjusting parameters. Specifying compile-time flags to find C headers can be done by specifying Puppet parameters, too. They use a few management applets as well to update FreeRADIUS `clients.conf` files and test many authentication scenarios. Either they are installed as files, i.e. as Puppet static resources. Or, more agilely, they are installed as virtual environments using the `puppet-python` module.

The new setup will rely on VMs. Each service will run 5 nodes for production, quality assurance and development. They currently run CentOS 7.4 and FreeRADIUS 3.0.14. The visitors servers came into production in July 2017 and eduroam servers will reach production in November 2017. More services will be puppetised in the future.

Questions and comments:

- Michele Michelotto asked, will visiting users and eduroam users have the same privileges? Visiting users and eduroam users will have access to the same services.
- Jérôme Belleman asked, in the agile way to install Python scripts, are the Puppet configuration and the scripts code kept in separate repositories? Yes.
- Does the RADIUS server need to restart with each new user? No, users are kept in a database.

Follow-up about Wi-Fi service enhancement at CERN (Vincent Ducret)

This is a follow up of last year's presentation. The goal is to improve Wi-Fi coverage at CERN, including for roaming between buildings. They have reached the CERN-wide deployment phase. The technical solutions are based on Aruba products. They have got a single point of configuration. It provides a redundant solution. A Wi-Fi control solution comprises access points which contact controllers. When clients connect to controllers, they get IP addresses from the same range, regardless of the building the clients are located in. A mobility master is the controller setup component which centralises the configuration and manages the RF setup. It is made of two nodes for redundancy. The managed devices are configured by the mobility master. Redundancy and traffic load balancing are carried out with the controller cluster. Access points talk to several controllers. That is how it is centralised and fully redundant.

CERN set up three Wi-Fi zones: a lab zone, a production zone and an IT-Pilot zone. Vincent showed the controllers in racks, for production and the pilot, together with the dedicated Brocade MLX. Automation is based on the API available on the mobility master. It simplifies deployment. The advantages of this solution are its scalability, the potential for automation and live upgrades. They suffered from some roughness of version 8.0.0.0. Interaction with the manufacturer raised more than 50 bugs, which was a time-consuming process. Since a few months, there is a regular follow-up with the manufacturer's support. The software version evolution had been a fast one. There were bugs on the routes in charge of all centralised traffic. One major bug on the controllers involved many configuration changes. There were no major issues perceived by the users. Around 30 buildings received the deployment, with more than 1 200 active access points. There are peaks of 3 000 devices connected simultaneously. Vincent showed a plot of the increasing number of access points and the number of connected devices. The performance observed is of the order of 300 Mb/s. The deployment will complete end of 2018 and coverage will extend to outdoor areas.

Questions and comments:

- What is the maximum number of users? One access point can take 50 users. Beyond this number, a performance loss can be felt.
- Michele Michelotto asked, do you need to register your MAC address with eduroam? No, but to reach CERN services you do. If you do not, you will only have access to the internet.

Configuration automation for CERN's new Wi-Fi infrastructure (Quentin Barrand)

Coming back to a question in his previous presentation, Quentin wished to point out that Puppet was in fact used much in the telecom group, unlike what he previously thought.

The production cluster includes mobility masters, around 7 000 controllers and 9 models of access points. Access points belong to access point groups. The group settings are applied to all the access points. But RF settings can also be defined on a per-access point basis. In fact, external antenna gains must be defined on a per-access point basis. The configuration is organised in a hierarchy. Deploying an access point entails including the cluster IP address as DHCP option and adding an entry in the mobility masters specifying the group. However, doing so for 4 000 access points across 3 clusters must be automated.

The CERN network management system – LANDB – is an Oracle DB with a web front-end. It stores information about all the devices. The main management application – CFMGR – uses LANDB data to build the topology. Adding data to LANDB means populating an external antenna table, the cluster table, the access point group table, the RF profile at 2.4 GHz table and the RF profile at 5 GHz table. It was decided to maintain only exceptions. Quentin showed the exception database schema. Deploying an access point is thus made easier – it becomes a matter of just declaring it as a network equipment in LANDB.

The team came up with a new application – NETMGR. It is a Python 3 application. The mobility master provides a REST API. The output is in JSON but feedback is rather unhelpful as the HTTP return code is always 200 and error messages are either missing or meaningless. It is nonetheless more comfortable than working with a CLI.

- Jérôme Belleman wondered, could NETMGR – and therefore LANDB – be used outside CERN? Yes, in principle, if you replicate the schema.

Firewall Load-Balancing solution at CERN (Vincent Ducret)

The traffic between the CERN internal networks (LHC Computing Grid/General Purpose Network) and the public internet is filtered by firewalls. The capacity is currently limited. The Firewall Load-Balancing Solution aims at bringing scalability to the setup. There were four options:

1. To keep the current setup but build a cluster of firewalls.
2. To keep the current setup and use larger firewalls.

3. To review the design completely with specific load balancers.
4. To build an intermediate firewall load balancer solution.

Option 4 was preferred, as it entailed reusing existing and spare firewalls, leaving switches and routers untouched, allowing the team to get familiar with firewall load balancing and keeping the costs down. In this option, the load balancer will receive both LHC Computing Grid and General Purpose Network traffic. This approach will also allow the security team to add intrusion detection. Load balancers have redundancy. If one load balancer fails, everything still has a connection to the second one. If one firewall fails, traffic is redirected to the remaining ones. Load balancing is carried out on source IP. Traffic will always cross the same firewall to avoid connections being dropped.

Vincent showed some configuration listings to configure ACLs and Policy-Based Routing rules. They can be scriptable in Python to ease operation. Some appliances allow running Python.

The current platform only supports IPv4 load balancing. IPv6 is not concerned by this solution. Besides, it is not possible to remove a firewall transparently. The pilot setup is active since mid-July 2017. Half of the devices will have to be moved to the second network hub of CERN. In 2018, they will migrate all traffic and review the whole setup.

Questions and comments:

- Are you planning to look into using OpenFlow? The problem is that the routers we have cannot be used with OpenFlow. The setup will be reviewed anyway in 2018 and maybe there will be a way.
- Do you use stateless ACLs on the routers? Yes, inside the campus network, but it is currently quite limited.
- Is there any indication from the vendors that IPv6 will be supported? Not currently.
- Is the ACL auto-generation carried out with scripts? It is a Cisco `itd` feature.

Wednesday 18 October 2017

Site Reports

DESY site report (Dirk Jahnke-Zumbusch)

The XFEL – European X-ray Free Electron Laser – was inaugurated. Data taking has started. One beam is in operation. Two more are to come. The data comes from two instruments, the FXE – Femtosecond X-ray Experiments – and the SPB – coherent diffraction imaging of Single Particles, clusters and Biomolecules. There is GPFS storage on the XFEL campus. Raw data is copied from the XFEL campus to DESY via InfiniBand. The offline GPFS storage acts as cache. Users see calibrated data from the analysis cluster. Data is valuable because it is unique. It cannot be reproduced. Storage must be scaled massively, a 100 TB online SSD pool. Tests will start this year. They decided to go for LTO-8 for long term storage. Another institute on the DESY campus is CSSB – Centre for Structural Systems Biology. The requirements will be similar to PETRA III. In Zeuthen, they have some involvement with the CTA – Cherenkov Telescope Array. DESY also contributes to Belle II. They provide the Atlassian suite and are involved in Invenio. They host some service VMs.

DESY will run Windows 10 in 2018. There is a long term support channel for the accelerator machines group. DESY heard problems related to UEFI and Dirk asked for advice. They are observing that the vendor does not fix bugs for SL6. Hamburg goes for CentOS 7, Zeuthen for SL7. They work with Docker for their HPC cluster. The June 2017 HTCondor workshop was hosted at DESY. Their pilot will turn into production in 2018. They are subject to an ISO 27001 audit for main IT services. They identified fields of actions, such as two factor authentication, splitting up hypervisor pools to restrict the spread of security breaches and stricter updates. The main computer room capacity will double to 1.5 MVA. They set up a networking switch delegation system with Oracle APEX. They are in the process of working on identity and account management systems.

Questions and comments:

- Tony Wong explained that in BNL they are having discussions about supporting the light source. They do not have single sign-on. Is there a single account per user? Yes, and we have to differentiate between community types. But it is not clear how to achieve this yet. Romain Wartel said, if you design an authentication system, you need to think about outside users. eduGAIN could help there.

GSI Site Report (Jan Trautmann)

They currently use two out of six levels of the GreenITCube. They moved all the HPC systems into it. The so-called *Accelerator IT* uses one whole row. They installed 200 kW in IT load. The PUE stays around 1.07. The MiniCube was decommissioned and Jan showed an image of what that was. They moved from CFEngine to Chef. All new servers are managed by it. Another ongoing migration is from Gitorious to GitLab, which has some integration with Chef. The compute cluster Kronos runs 13 800 cores, with more coming. The GPU part only has 700 GPUs. The Lustre filesystem reached 14 PB. They are working on a HSM-TSM integration.

They are looking at alternatives to Debian, e.g. CentOS. They use a new asset management system. Their old system was developed by GSI and it is no longer maintained. They listed 16 solutions and have evaluated 5 of them in more details. The winning solution is FNT Command which supports many modules. It is a web-based tool. It can show 3D rack views. They had to rework their old database. The May 2017 open house invited many visitors. The FAIR groundbreaking was held in July 2017.

Computing & Batch Services

Integrating HPC and HTC at BNL – a year later (Tony Wong)

The Scientific Data & Computing Center (SDCC) acquired the Institutional Cluster (IC) and the KNL-based cluster for HPC-based projects. Their responsibilities expanded with e.g. Belle II. They had their current resources fully utilised. The worker node cluster for ATLAS was 100% loaded. Next to shared resources, they

run dedicated resources with rigid constraints. They are aiming at enabling resource sharing. Collapsing multiple HTCondor pools into a single pool could be an option and increasing flocking is another one. They developed a mechanism to submit HTCondor jobs to SLURM.

The Institutional Cluster is in production since January 2017. The original cluster ran 108 nodes with CPUs and GPUs. They decided to expand it with GPUs, in production since September. Another 36 machines were ordered. It is available to HPC and HTC users.

The KNL cluster entered production in June 2017, after many issues. It is a 144 nodes machine. It uses the OmniPath interconnect fabric. The cluster is in a useful state, not yet optimised. They decided to stop experimenting with the cluster for the sake of stability. It is used by the LQCD and ATLAS communities. Tony presented the Titan system at Oak Ridge, running 18 688 compute nodes (299 008 logical cores) with GPUs – a Cray XK7. About 2.5% of the total available time on Titan is for running ATLAS jobs.

BNL have plans to buy another cluster to be shared between HPC and HTC. To increase the sharing, they plan to implement HTCondor changes. A likely future direction is a broad effort to encourage the use of Leadership Class Facilities to meet needs. They are still in touch with Amazon and Google for integrating their resources into the cloud.

Questions and comments:

- Are there ways to share resources between SLURM and HTCondor? They have a prototype to submit HTCondor jobs to SLURM. They know it works.
- Helge Meinhard remarked upon the fact that this encouragement to use HPC resources for HTC workload might be a bad deal. The large HTC users are already using HPC to the largest extent possible. They are considered as opportunistic resources. He sees it coming that funding resources consider they should be included in pledges.
- Mattias Wadenstein said it is important to distinguish pledged and opportunistic resources in that opportunistic ones can be interrupted, for instance. If your HTC bursts into volunteered computing, it should be considered as a nice addition, not pledged resources.
- How did you manage to get resources in the leadership clusters? They have more resources than they can use.

Techlab update (Romain Wartel)

It is an IT service, provided across multiple groups, to the benefit of the experiments and SFT. It is about providing hardware to allow performance studies of competing architectures. They try to use software as close as possible to standard production hosts. There is a solid user base in experiments. Everything can been published – no NDA. Hardware is off-the-shelf commercially available; it is chosen according to user community requests. Users can loan hardware like in a library. There may be less interest for low power architectures. A good 80% of recent users are interested in machine learning. They bought more GPUs recently.

The booking system is integrated in the installation of ServiceNow for CERN. They provide monitoring to assist users and are developing a benchmarking platform. Romain showed the variety of hardware that is available. The main challenge is to balance the effort to learn and get access to new platforms. Is it worth managing a given platform with Puppet, for instance? The experiments lack a strategy. It is difficult to move from testing in Techlab to production. Users would need more technical expertise.

They plan to continue providing this service at minimal cost and contribute more to the community. They need to spend more time understanding how users intend to use the hardware, and reduce effort duplication. They expect to report on their work via HEPiX, GDBs, etc. Part of the strategy will be to run questionnaires.

Questions and comments:

- Jérôme Belleman asked, is the benchmarking framework for inventorying results or conveniently running them? It is currently a web portal to classify results and methods. The idea is to work with the Benchmarking Working Group to agree on how to run the tests.

- Do you only focus on CPUs and storage? Or also networking? Mainly CPU, in fact.
- Dennis van Dok asked, do you get special access with vendors? Yes, early in the process, but still off the shelf, not development boards.
- Do people request AMD GPUs? Yes, but rarely, only for comparing.
- Did you get a chance to play with the ARM64 Mellanox CPUs? Not yet but eventually we have got plans to do so.
- How do you help users optimise? We have limited expertise, we can monitor the system e.g. find out when not all cores are used. But we cannot optimise existing code. For GPUs it might be easier. We hope to organise workshops.
- Any request for MPI? Not for a long time.

HEPiX Benchmarking Working Group – Status Report October 2017 (Manfred Alef)

There are 60 people signed up on the mailing list. Biweekly Vidyo meetings take place. Members are site administrators and also experiment representatives. There are some news about the fast benchmark. The DB12 benchmark scales well with ALICE and LHCb experiments. However, it runs for one minute and does not stress many system components.

The purpose of the long-running benchmark is to measure installed capacity for pledges, accounting and procurement. It is not about scaling with each individual experiment workload and job type. It must scale with the job mix. The current HS06 benchmark is based on the SPEC CPU 2006 suite. SPEC CPU 2017 was recently released. Some sites have already purchased it. The working group are packaging ALICE and ATLAS reference workloads in Docker containers. SPEC CPU 2017 looks similar to SPEC CPU 2006 and comprises 43 single benchmarks. As before, speed and rate benchmarks are split. First results are expected to be presented at the next HEPiX.

There are scaling issues with HS06 with respect to HEP applications. Migrating to 64 bit does not fix scaling issues. HS06 was developed by the working group about 10 years ago. At the time, typical chips were quad-core Xeon E53xx or E54xx CPUs. The first servers with Hyper-Threading appeared at the end of the project. Several experiment reports compared different hardware models but few of them took into account individual worker node configurations. Manfred showed a plot where DB12 does not display better performance with multiple benchmark copies, unlike HS06. GridKa reconfigured their compute farm and compared experiment job and benchmark performance. He described performance results from ALICE, ATLAS and LHCb. Comparing all benchmarks, it appears that ALICE DB12 is close to the DB12-at-boot benchmark. The situation is similar with LHCb, with LHCb DB16 close to the DB12-at-boot benchmark. GridKa ran tests of their own for CMS, showing good correlations between ttbar simulations and DB12-at-boot up to 15 parallel processors, but not beyond. DB12-at-boot is not a suitable candidate with multiple copies.

Questions and comments:

- Pete Gronbech asked, did you do tests with other number of cores or did you just draw the lines to connect the points? We did not test with other number of cores because of the complexity to set up. Performance differences are not expected. In addition we are after multiples of 8 (4 if needed) for the purpose of multicore jobs.
- Helge Meinhard pointed out that he finds it surprising that ATLAS has a different scaling behaviour. It could be valuable to get the same kind of information from CMS to cross-check. If the scaling difference is real, there must be a reason for it. It would be valuable if experts could understand where the difference comes from, maybe at the processor or instruction set level. It seems to be hard to perform this analysis because of the code complexity. Michel Jouvin thinks it is on the experiment's side. In CMS, Vincenzo Innocente was the expert and it could be useful to contact him.

- Michele Michelotto had made measurements with HS06 and noticed no oscillations. In slide 21, DB12 has a large error bar, except for the last point. Yes, there should be an error bar.
- Ajit Mohapatra asked, did you run memory-per-core tests? No, but the tests ran more than 2 GB/-core. The more memory you have, the better the file and I/O cache, the smoother the operations.
- If you have 40 job slots/node, it increases cost. But you already have it. Since ATLAS benefit from them, why not use all 40 jobs? It is recommended to use fewer slots to have more memory on the side. It depends if you want to match pledges or optimise jobs.

Running Jobs Everywhere: an Overview of Batch Services at CERN (Jérôme Belleman)

This lightning talk is an update of the batch services at CERN, with an emphasis on the idea of running jobs anywhere. Jérôme started with high throughput computing on HTCondor worker nodes. The capacity reached more than 100 000 cores, almost twice as many as what CERN ever reached with LSF, which is now to be reduced to 20 000 cores by Spring 2018. In fact, the LSF service will stop once the experiments move, which will have to be at the end of LHC Run 2 at the latest. Meanwhile, the team are training new users wishing to migrate.

In addition to running jobs on their worker nodes, they are working on expanding clusters into commercial clouds. The XBatch project was presented before, and the team are working with the Helix Nebula Science Cloud, with a tactical engagement with Oracle Bare Metal Cloud using the Docker Universe.

Running jobs everywhere also means running them on disk servers, an idea presented last HEPiX by Andrea Sciabà and Andrey Kiryanov. Tests are now ongoing. Disk server slots are used as whole batch slots, no lower-SLA. They segregate batch functions from the disk node using cgroups/systemd for the HTCondor service, the Docker Universe – for which they are testing the CPU set – and a dockerised CVMFS.

Back to worker nodes, but this time for high performance computing, the team work on a cluster to provide MPI, shared memory across nodes and InfiniBand. They chose SLURM for these purposes. They are attempting to perform some backfill via a HTCondor/SLURM interface and would require some help with this.

Finally, CERN batch services are working on running jobs on other people's laptops as well as full CPU farms with volunteer computing – LHC@home. This is about running all applications, accelerator ones as well as event generators and detector simulations. SixTrack runs in a native BOINC app, now also on Android. Simulations are run under CernVM and VirtualBox: ATLAS (including Singularity), CMS, LHCb and Theory. The job management back-end is integrated with HTCondor, supporting a wide range of low-IO/high CPU applications. The team are working with the BOINC community to evolve the BOINC software. Jérôme showed a plot describing different aspects of volunteer computing activity, and how the team intend to improve usage.

Questions and comments:

- Pete Gronbech asked, how many of the disk server slots do you use? All of them, with little impact on performance.

Migration from Grid Engine to HTCondor (Thomas Finnern)

BIRD stands for Batch Infrastructure Resource at DESY. They have got a growing HTCondor BIRD pilot with more than 500 cores. They run 1 master server, 2 schedulers, 10 submitter nodes and 500 worker nodes. They are integrating Kerberos and AFS with no better alternative at the moment. Tokens allow for a maximum of 1 week's run time. The current token is repeatedly extended from protected servers, from Kerberos tickets.

They integrated the setup with their user registry, to set the adequate project on the group submit host. A user may switch to another project, but by default it is set to the primary registry group. It is checked against the registry, defines a fairshare quota group which is set on the worker. They set up some automated workflows to disable, drain, reboot or reset nodes. Each user has to define how long his jobs will run.

In the future, they will be working on more grid integration, set up Docker containers for several operating system flavours and backfill HPC resources from HTCondor.

Questions and comments:

- Jérôme Belleman asked, what happens if a user submits a job on behalf of a project he does not belong to? This is currently noticed a little late, and the job will run. Something to work on.

Migrating a WLCG tier-2 to a Cray XC-50 at CSCS-LCG2 (Francesco Giovanni Sciacca)

CSCS hosts Piz Daint, a supercomputer ranking number 3 in the TOP500. They are working on consolidating a project to run LHC jobs on Piz Daint. The goal is to run all VO workloads without changes to the experiment workflows. They will measure performance in a production environment to produce a cost study. They will then choose whether or not to migrate to the Cray or fall back to Phoenix, the previous regular cluster. This would be a 4 years' commitment.

The operating system is a stripped SUSE. The nodes are diskless. Memory management is limited to 2 GB/core. Software provisioning in absence of disk is a problem, too. Francesco described the architecture and interfaces, namely Cray Data Virtualisation Services exposing GPFS and Cray Data Warp Services which are SSD-based. The current cluster is made of 72 HT cores. They run an ARC CE. The SLURM batch system has a single WLCG partition. The operating system is CLE – Cray Linux Environment. Jobs run in Docker containers using Shifter. The image is a full WLCG worker node. Shared filesystems are covered by Cray Data Virtualisation Services and Cray Data Warp Services. Docker images are kept in the Cray Sonexion 1600 Lustre FS.

CVMFS runs natively on the nodes using workspaces and tiered cache, a new feature of CVMFS. However, caching on Data Warp Services resulted in corruptions. They do not support `flock()`, but with the workspace setting it is now possible to set all locks relative to the cache local to the node. Setting `CVMFS_CACHE_hpc_TYPE=tiered` with upper layer in-RAM storage can dramatically increase performance.

Francesco showed a snapshot of the utilisation where we can see a ramp-up. The trend is promising, with gaps caused by various issues. Another view showed unavailable nodes because they were offline by the automatic SLURM health check features. There were issues with the middleware and experiment software. Fairshare tuning is more complex in the Cray environment. Shares are not as well defined as they should be. Some nodes seem to become black holes for no apparent reasons. The filesystem was more challenging, still. Performance seems to degrade quickly with high I/O levels.

They compared the performance of Piz Daint and Phoenix. They evaluated reliability, the produced wall clock, the ratio of job successes and failures and the efficiency. Francesco showed a few resulting plots, per VO. LHCb ran simulations only during this time window, hence the top efficiency. It does not appear that any system performs better. The availability and reliability are dominated with dCache issues harming both systems.

They did manage to address issues and reached a stable situation. Both systems seem to perform well. But the team have concerns with scalability. They hope to be able to perform a test at the 20 000 cores scale in November.

Questions and comments:

- Jérôme Belleman asked, what was the original motivation, especially now that there are indications that both systems preform equally well? Scalability in the longer run.
- Rob Gardner asked, did you have to ask administrators to install FUSE on the compute nodes for CVMFS? The FUSE modules are in the Docker image.

Optimising the resource needs for the LHC computing: ideas for a common approach (Andrea Sciaba)

The motivation is the considerable gap between the best estimates of the required and the available computing resources for Run 4 in particular. Some resource estimates, e.g. from CMS, are to provide 6× CPU

and 4× storage. Algorithms will have to improve and possibly use more specialised hardware. Significant progress already took place. Tools to be able to calculate how much resources and how much they cost will be needed; metrics are required. Also, a model will have to translate physics needs and software performance into resource requirements and costs.

The usual metrics are IOPS, throughput, CPU utilisation and application response time. It is not obvious which metric best characterises our applications and infrastructure. The fact that almost all HEP data processing proceeds on an event by event basis is an important constraint. CPU time/event measures the cost of processing, the CPU/wall clock time ratio measures how well CPUs are utilised and instructions per clock indicate how busy the CPU really is. There are also memory, storage and network metrics.

Ideally, CPU should always be busy, with minimal I/O wait, doing what we need with the least amount of instructions, fully exploiting parallelism. It is unrealistic, but it can be approximated. We need a way to understand when I/O is a bottleneck. The physics program, the detector, the offline software and computing model determine the amount of resources required. Estimating the cost of the infrastructure, including the manpower, is also important. In the real world, there are dozens of parameters to be taken into account. Experiments currently do this independently. Many years ago, there were tools to do that, used by all experiments. There was a spreadsheet – the *megatable* – to do so. The spreadsheet became difficult to maintain. Costs are hard to compute because you need to take into account capital expenditures and operational expenditures.

It is important to compare on-premises resource costs with commercial cloud costs. Commercial clouds are not significantly more expensive, and the cost difference is likely to decrease further. WLCG have been discussing the need of a working group since some time and the decision was taken at the latest WLCG workshop. The goal would be to identify a set of reference workloads and measure them in different environments as input data for the model.

Questions and comments:

- Jérôme Belleman pointed out that models would be useful to implement batch accounting to understand where the numbers come from. Is accounting part of this work? It is an essential part of this work, so everybody can understand the numbers the same way.
- Michael Davis pointed out the trade-off about the storage requirements and CPU requirements, to store data that cannot be analysed immediately.
- Tristan Suerink asked, are there tools for evaluating dual- as well as single-socket systems? There are currently no requirements for this.
- Helge Meinhard pointed out that the discussion shows it is important that the working group clarifies what we want to achieve. We need to be clear on the granularity. It is more important that with a little effort we achieve some result and then agree where to refine further.
- Michel Jouvin followed up, the goal is quite clear. In this sense the model should not be tuned for a single technology. It is a global optimisation at a high scale level. This is connected to the community white paper effort.
- Pepe Flix agreed that instead of attempting to solve all the issues, identifying the topics that need further in-depth investigation would be useful. We will need to track future technologies and what kind of optimisations we can make today.

IT Facilities & Business Continuity

Hot days with no mechanical cooling data center (Cary Whitney)

What do we think about our facilities people? Do we accept our facilities as they are, or do we work with our facilities to match our needs? NERSC's new building has got free air and water cooling all year round. Cary described the path air takes into the building. They told their vendors about the 24°C temperature

intake. The cooling towers provide 18°C water especially for the Crays. They can only take weeks' worth of data because of the high data volume.

One day they got particularly hot weather. They reached a record high temperature (41°C). Operators had to spray water on the chillers. They were evaporating water faster than they took it in. Crays did well under watch. Cary showed plots of the valve positions, pressure differences and consequences on the Cray systems which were pumping as much water as they could. They looked at the blower temperature in the cabinets. The issue was with the wet bulb, part of the heat exchanger. The critical operating temperature is 24°C. They got past this critical temperature often in the past months. They learned that the physical system may be more resilient than expected and that a suitable visualisation of the environmental data could be very helpful. The air quality also has an impact. Cary showed a smoggy day in Berkeley and the air was sucked into the machines. He explained how to filter the particles. The particle count went noticeably up. They tried to improve the situation by closing dampers and manually cycling evaporation cooling. Without the data they collected, they would not have been able to see that the 24°C water temperature was as critical as expected.

Questions and comments:

- Mattias Wadenstein asked, did you see throttling. Just one core. It did not affect job processing.

Grid, Cloud & Virtualisation

Singularity at the RACF/SDCC (Christopher Hollowell)

Singularity is an OS container management system for Linux developed by LBNL. Its objective is mobility of applications rather than isolation. Linux containers all use the same underlying technologies which are namespaces. Only the operating system is abstracted. For HPC and HTC jobs, Singularity allows administrators to use almost any flavour of Linux. Users can create containers fulfilling dependencies. Containers are not guaranteed to run newer versions of the OS, although in practice it is not a problem.

Singularity is lightweight to install, it is a single package. It is secure out of the box. Users are never root in the container. Since recently, you do not even need to have a setuid binary. It also has built-in support for HTCondor, taking advantage of its support for VMs. While starting VMs can take some time, containers are started much more quickly. Containers give you access to the filesystem. Processes also have access directly to the GPUs. In 2015, CERN investigated performance penalties from KVM and noticed a 20% reduction, which was subsequently addressed. With containers, there are no performance penalties.

Singularity supports different container formats. There is no significant difference of start-up times between image files and directories. At the RACF, they created a BNL ATLAS SL6 container image. They preferred creating their own rather than using the official ATLAS one because of some previous problems with OverlayFS. After a successful period of running HammerCloud tests in containers, they opened batch queues to real ATLAS jobs. ATLAS recently validated SL7 worker nodes, but few sites have upgraded. They do not transparently support a mixture of SL6 and SL7. Singularity provides a path for upgrading in a rolling manner.

During testing, they noted that many filesystems contain the + character. Singularity does not permit this character in bind mount path names. Singularity is available for use on their HPC clusters. ATLAS is interested in making opportunistic use of these resources but CVMFS is not available on those systems. A solution is to create a fat ATLAS Singularity container image.

BNL is in the process of becoming the T1 computing centre for Belle II in the US. They created an SL6 container image for their RHIC experiments. Compared to the minimal ATLAS and Belle II images it was more complex to build and larger. Chris showed the SL6 Singularity definition file and a terminal showing Singularity, highlighting that a user could not become root, keeping it secure. They plan on upgrading Singularity and store images into CVMFS. ATLAS is also working on providing Singularity support in Panda.

Questions and comments:

- Is it possible to limit resources in a container? You can use cgroups to do this. There is probably no

mechanism in Singularity itself to do that. Mattias Wadenstein said that the developer scenario is that the batch system would handle that.

Running HEP Payloads on Distributed Clouds (Rolf Seuster)

At the University of Victoria, HTCondor and Cloud Scheduler are among the services they need to run. They run real workloads in a variety of commercial clouds and this is what is meant by *distributed clouds*. But it looks like a regular grid site from the user standpoint. Cloud Scheduler regularly checks jobs in queues and boots VMs on any cloud which has matching resources. Canada is producing roughly 20% or the resources for the Belle II experiment. All Belle II jobs are single core. ATLAS jobs mostly require eight CPU cores. Looking at monitoring, the scale in terms of running jobs seems to be limited by available resources and user needs.

To reduce manual intervention, CERNVM images are configured for ATLAS and Belle II jobs. Accounting is carried out, as well as benchmarks, and application success and failure rates collected. CERNVM gets almost all software via CVMFS. Cloud Scheduler injects contextualisation via cloud-init into vanilla CERNVM images. Rolf presented Shoal – a school of Squids – which is contacted by VMs to find the nearest Squid. It basically runs `curl http://shoal.heprc.uvic.ca/nearest`. All distributed Squids are registered.

DynaFed is a dynamic federation of storage elements, supporting different protocols. It uses GeoIP to work out the closest locations. To answer questions from users on how much resources are used by a given region, they run accounting which reports uptime and user time for VMs, allowing them to identify clouds which provide unreliable resources, requiring manual investigations. To manage application success and failures for Belle II, they store the job ID of all jobs and inquire DIRAC about the status. It is cloud-dependent, allowing them to identify clouds with frequent failures.

They intend to rewrite Cloud Scheduler using unit tests, Jenkins and CI. They wish to overcome a scaling issue at $O(10^4)$ jobs, considering flocking. They would be interested in common interfaces with other provisioning solutions to ease integration.

Using Docker containers for scientific environments – on-premises and in the cloud (Sergey Yakubov)

Earlier on, Thomas Finnern had presented DESY's HTCondor farm and the Maxwell HPC cluster running on SLURM. They wish to separate the IT services from user responsibilities with Docker containers, and provide compute resources dynamically and quickly. `Dockerfiles` are created by group administrators. Puppet creates an image based on changes made to the `Dockerfile`. Compute resources are reserved with SLURM which schedules the job. Users can also submit SLURM job scripts with Docker commands.

A use case is SimEx, a photon science simulation platform. It is a good example because of its many dependencies, which would normally cause headaches for administrators. Thanks to the Docker setup, they could reach a 160× speed-up compared to running a single MPI process in a single thread.

DESY takes part in the Helix Nebula Science Cloud. The group of procurers have committed to provide funds and manpower for testing and evaluation, use cases with applications and data, and in-house IT resources. There are technical challenges on different fronts, such as computing, storage and identity management. The project saw several phases. They are currently in the prototyping one. Each step is competitive as only contractors that successfully complete the previous step can bid for the next one. Only two pilots will remain in the end. The goal is to use the cloud to extend local resources. Sergey showed a hands-on example to define and submit a test. One important next step will be to test performance.

Security & Networking

Low-Power Wide-Area Network (LPWAN) at CERN (Rodrigo Sierra)

IoT networking covers a large number of different specifications depending on the user needs. The different characteristics of the network such as range, data rate, traffic patterns, power consumption, mobility, security, area covered, network price or frequency spectrum need to be adjusted to cope with each case. Services requiring high bandwidth or real-time data (e.g. video-surveillance, smartphones, real-time telemetry, robotics, …) make use of the current proposed wireless technologies such as Wi-Fi and cellular

services (2G, 3G and 4G). Nevertheless, new requirements for low bandwidth, long-life battery-powered devices has raised lately. LPWAN is uplink-centric and is optimised for low data rates within long ranges, long battery life and a massive number of devices at low cost. At CERN, some users started asking for this kind of solutions. The network group studied a number of them and LoRa seemed the most promising one. It is mature and suitable for indoor coverage. The unlicensed spectrum means that CERN will be able to deploy it freely.

It works in a star architecture. There are three classes of efficiency which affect its listening, ranging from every second to every minute. All devices can send information whenever ready while complying with a duty cycle imposed by each country. As a consequence, collision may happen when many devices send data simultaneously so we cannot count on the messages to arrive immediately. Security uses a 128 bit AppKey. The network server contains the intelligence of the system and can be open source or commercial. There is a concern on security in terms of the number of certificates if devices proliferate. The coverage on the CERN campus as well as in tunnels will be another challenge.

The team are at the designing stage. CERN has joined the Proof of Concept of Geneva to gain experience with the technology and the data flow. A PoC within CERN should be proposed in 2018. Rodrigo welcomes discussions on how best to implement all this.

Questions and comments:

- Jérôme Belleman asked, do you have other use cases in difficult environments – you gave the example of ionizing radiation resistance. Some work will be done to harden some components, indeed for instance against radiation.

Fancy Networking (Tristan Suerink)

Nikhef reached the physical limits of their design. They need to replace their equipment to support their new HTC cloud environment and investigate long distance network technology. Tristan described their traditional OPN implementation. Their incentive for cloudification was to cover more HTC use cases, scale dynamically and enable easier multicore.

They came up with a list of requirements ranging from bandwidth to isolation, dynamism to cloud bursting. Possible candidates include products from Arista, Brocade and Juniper. They went for a Juniper QFX10000 in the end. They performed a long distance DWDM test between Amsterdam and Geneva. This was not trivial. Fibres need to be immaculate – any spot will stop them working. In the future, they would like to become flexible with their resources, while keeping their high speed interconnect. So many flaws can be expected in networks that stateless networking is essential.

Questions and comments:

- Stefan Stancu asked, do the backup slides describe plans or things working? A bit of both.
- Shawn McKee asked, what if the QFX dies? They did everything they could to make sure the box does not.

Network Automation for Intrusion Detection System (Adam Lukasz Krajewski)

This project was a collaboration between Brocade, openlab and Extreme Networks. This presentation is a follow up from a related talk at HEPiX Fall 2016. The volume of traffic entering and leaving CERN keeps on growing. The security team have built and deployed an extensible IDS system. The principle is to mirror traffic at network boundaries, aggregate and balance the traffic load across servers and leverage network hardware features. In particular, symmetric load balancing, traffic shunting and selective mirroring are necessary requirements.

The previous architecture involved a Brocade MLXe router, OpenFlow, the Brocade Flow Optimizer and Bro. It was limited in that the configuration of symmetric load balancing was hard to achieve and documentation was lacking – it required much troubleshooting. OpenFlow falls short of features such as traffic shunting and Link Aggregation Groups. The chassis they deployed for the router was 8U and they

would have preferred 1U. But this year, Brocade released the new SLX series with automation readiness, a REST API and virtual machine hosting.

When SDN became popular, the paradigm was to decouple software-implemented control from forwarding. More recently, network automation became a new trend and it can take advantage of configuration management systems such as Puppet, device drivers such as Brocade's PySwitchLib and Juniper's PyEZ, and orchestrators such as StackStorm. StackStorm, recently acquired by Brocade, is a cross-service integration and automation system. It can monitor large traffics and provide a REST API for integrations with higher-level components. Brocade Workflow Composer is the configuration workhorse and Brocade Flow Optimizer is the brain behind, holding the business logic, making decisions. CERN ended up using the Brocade SLX 9540, the Brocade Workflow Composer, the Brocade Flow Optimizer and Bro. Adam described how to perform traffic shunting using these different components. Bro marks traffic as suspicious and triggers actions with the Brocade Workflow Composer and the Brocade Flow Optimizer.

Questions and comments:

- Jérôme Belleman asked, is StackStorm specific to networking? What is the difference with SaltStack?
- Tristan Suerink asked, are you in touch with your firewall colleagues, to use a similar architecture as an alternative to expensive firewalls? Some investigations are ongoing.
- Shawn McKee asked, you are relying on Brocade – any prospect on open source solutions? Yes, we are thinking about just using Brocade Workflow Composer which remains open source, even if owned by Brocade.

Storage & Filesystems

CEPH at RAL (Ian Collier)

Echo is a Ceph cluster with erasure-coded data pools. It is a disk-only storage to replace CASTOR for LHC VOs. DynaFed is designed for federating data sources. In this context, it is not used as federation but for translating protocols. In principles you could add plugins for authentication. The user talks to DynaFed to pass a request and receives a pre-signed URL in return. Since April, ATLAS have been reliably using Echo in production. RAL have been testing CMS activities and workflows. They came up with a migration plan involving a separate XRootD proxy cache cluster for AAA. As a side effect, they upgraded their batch farm to SL7 and they run XRootD in containers. LHCb have been able to transfer data into Echo via FTS. Work with ALICE started just last month. Their use case is similar to CMS but with their own XRootD authentication layer. RAL have been working with LIGO and put together a functioning prototype.

It has not all been smooth although Ceph has generally been reliable. There have been a few XRootD and GridFTP plugin bugs. File overwrite does not work. Memory usage in both plugins is being investigated. But since then day to day running has been smooth, with about one call-out a week and all was kept transparent to VOs. The XRootD and GridFTP Ceph plugin memory usage during reads is proportional to the size of the file being transferred. There was a data loss incident but it was interesting and challenging whence it happened. In early August, a number of OSDs increased to 2 160. An OSD was removed due to read errors. The first 3 OSDs in the set started crashing, flapping and finally died. This resulted in the loss of 23 000 ATLAS files. The team was able to briefly restore the placement group. There turned out to be a bug in the erasure coding backfilling. There was also a separate bug generating unhelpful error messages. The first day of solving this problem was to upgrade the software to solve that bug.

Currently the gateway Ceph plugins do not gracefully handle inactive placement groups. In the future, upgrades fixing the backfilling bug will be deployed and the 2016 generation of disk servers will be installed with BlueStore OSDs. Nonetheless, the first 6 months of Echo in production have been encouraging. Having decided that they needed a new technology for storing Scientific data, developing Echo proved to be a success.

Questions and comments:

- Tristan Suerink asked, do you make use of hardware erasure coding offloading? Ian does not believe so.
- Ofer Rind asked, how many concurrent connections can you handle? Of the order of thousands.

WLCG archival storage group (Helge Meinhard)

Oliver Keeble prepared the slides, calling for input and help. This is about magnetic tapes. With increasing pressure on resources by higher demands, experiments have been asked to use archival project as much as possible. It means that the experiments need to think about how to use tapes effectively, but also for sites to form a pool of knowledge. A CERN e-group was created (`wlcg-data-archival-storage@cern.ch`) to form this group and decide in which form the topic should be followed up. The HEPiX audience know the tape managers and invite them to join the group, bring their ideas and also learn from other people. This is formally a WLCG activity, but it would not be a problem if institutes not involved in the WLCG joined, too. A first meeting would take place end of October 2017.

Questions and comments:

- Pepe Flix said, it is also important that people working in sites have connections with experiments to involve them. We need to understand how they use the system. It is also about how it gets integrated in experiment workflows.

Thursday 19 October 2017

Security & Networking

EGI CSIRT: Keeping EGI Secure (Vincent Brillault)

EGI is made of 300 sites worldwide, with largely varying levels of expertise. The EGI CSIRT is a collaboration of security officers and experts. They focus on prevention, response and training. EGI CSIRT covers various areas of activity: Security Monitoring, Software Vulnerability, Security Drills, Training & Dissemination and Security Policy. Prevention is about keeping sites up to date. From the notification of a vulnerability, they check if it is critical, if the site is up to date, contact and assist data centres. The number of vulnerabilities remains the same, it is a constant effort.

Vulnerabilities are classified – critical means it risks to become an incident in a matter of days, the only level followed up by CSIRT. If escalation is possible, the level is raised from high to critical. Sometimes, vulnerabilities are critical but EGI CSIRT only send a heads-up when no patch is as of yet made available. Every time there is a new vulnerability, the time distribution of vulnerable systems shows a peak at time zero, which goes down as systems are patched over time, with a rather long tail. The longer the time it takes to update all nodes, the longer the tail. Following data centres is time consuming but requires little security expertise and CSIRT are currently evaluating if EGI Operations could be involved.

The Incident Response Task Force coordinates response and helps with forensics analysis. There have been 44 incidents since 2010, sometimes due to repeated mistakes. Vincent gave the example of the VENOM response. They could find the back-door, but because nodes were rebooted, there were no network traces left.

Another task carried out are security challenges to simulate incidents over several sites, testing all links in the chain: communication, setup, response, coordination, help and policies. This year, one such challenge was carried out in a federated cloud. Two sites reported DDoS. Most sites suspended VMs within a day and contributed out of working hours. They organised training on different topics: defensive, offensive, digital forensics and role-play.

One ongoing challenge is federating clouds: system administrators are in charge of everything and the clouds are often directly exposed to the internet. Another challenge is that of federated identity: it is about providing simpler access to the grid. Finally, one last ongoing challenge is that of widening the collaboration. Ties are made with EUDAT and PRACE. There is also the WISE community and collaboration with the US.

Questions and comments:

- Dennis van Dok said, he heard that some sites complain about the number of vulnerabilities forcing them to take action. Is there any way to help more? This is why there are different levels of criticality, reaching critical only if there is a belief that it will result in incidents soon.

Security update (Romain Wartel)

Romain and the CERN security team launched an awareness campaign using the three most common techniques used today: malicious mail (a), water-holing (b), and malicious Word documents (c). (a) Helge Meinhard appeared to have sent a mail linking to a HEPiX Wiki page. Most participants did not notice the mail headers were forged. The mail contained remote contents and read receipts, both of which cannot be disabled in some configuration (CERN is affected by this). This enables detailed profiling of victims to be made (IP address, operating system, etc.) (b) The HEPiX Wiki had areas left unpatched since 2006 and a new page was created containing text encouraging people to download an attached Word document. Because this is on an official website, a number of participants trusted the contents and downloaded the file. (c) The Word document was carefully crafted to include hidden Visual Basic scripting code, which was largely inspired by daily waves of malware received by most organisations. Minor tweaks in the code ensured that almost all antiviruses would say the document is safe. If executed, the HEPiX payload worked on PCs and Macs and fetched a text file from a CERN website, then opened it in Notepad (arbitrary code

execution). All three of these techniques are used constantly against CERN and HEPiX participants by real attackers. FireEye mail appliances (behaviour detection) did however detect that the document was attempting to run a shell.

Romain gave the traditional HEPiX security update. He started by reporting an increase of ransomware attacks. Locky and Trickbot seemed to dominate. There are new waves everyday. The profits are on the rise too, +2 500% in 2017. It is a big market. There is no need to be skilled, the software is easy-to-use. And victims, including universities, sometimes do pay. Software became cheaper, too. ATMs in China, train systems, smart TVs, washing machines, X-ray medical equipment, IoT devices are just examples of affected systems.

Romain showed how it is possible to buy fake credit cards, which sometimes come with a complete set of personal identity on Alphabay. The site, having reached 190 000 members, was taken down in July. The money was invested in minor cryptocurrencies. The founder was caught by leaving his personal e-mail address in the welcome mail header.

There are markets where people sell all they can to monetise data leaks. SlillPP and Paysell sell 6 million accounts at anytime for more than 500 organisations. Selling verified credentials is easy money. The CERN security team sent tens of notifications about data leaks in 2017. There are several tools to perform credentials replay and compromise accounts. You can procure hardware to run the tools.

Social engineering is on the rise, too. CEO frauds are one of them. It is about attackers impersonating a CEO, tell a story involving payment problems in an allegedly top secret case. There was such a case at CERN, which got luckily stopped.

BusyWinman is a malware targeting Firefox. It collects information from Firefox, stored passwords and history. Oilrig fake LinkedIn profiles, sending invitations to conferences. Beijing was accused of an IoT attack targeting a boat. All these attackers use the same techniques and they reuse a lot. Opportunistic, targeted, nation state attacks all use the same ones.

Questions and comments:

- Jérôme Belleman asked, what is this list of antiviruses that you showed? It is the VirusTotal service which spawns VMs to run a variety of antiviruses for free.
- Mattias Wadenstein pointed out that there are sometimes false positives on VirusTotal.
- Dave Kelsey asked, how effective are antiviruses, nowadays? Not very, as this simple case showed. There are more effective behaviour-based solutions, however, detecting that a Word document should not spawn a shell.
- Andrea Chierici asked, does all this work without administrator rights? Yes.

Current Status and Future Directions of KEK Computer Security (Fukuko Yuasa)

Fukuko showed plots of the number of security incidents at KEK year after year. There was a peak in 2005. It has since then overall decreased, with only 2 incidents this year. But is it the silence before the typhoon? There is an increasing number of targeted mail attacks in Japan. Academia has recently become a target in Japan, and KEK are no exception. Fukuko gave the example of a notice mail to researchers about a grant funding from the Japan Society for the Promotion of Science with a malicious ZIP file. The mail traffic is chiefly comprised of spam which is not caught by antiviruses. Many users receive viruses. The Japanese government enacted a law on cybersecurity in 2014. KEK fear the reputation damage and do not want any severe security incidents which would need to be reported to the Ministry of Education, Culture, Sports, Science and Technology.

A new regime for information security started since April 2017 to carry out the plan. They now have a CSIRT at KEK to focus effort on incident prevention, handling, monitoring and logging. On the prevention front, the team run a firewall blocking malicious URLs and IP addresses from various sources such as the Japanese police. They raise awareness of mail attacks and educate their users with face to face courses.

KEK received early on this week a firewall alert on the eduroam network and from their guestnet ISP. This was the CERN security team preparing the HEPiX drill. KEK decided to block the old HEPiX wiki.

NREN is also pressured by law. The National Institute of Informatics operates SINET. They launched in September 2017 NII-SOCS, which monitors packets in a sampling method and notices when suspicious packets are found. KEK have to protect internal network segments where e.g. accelerator control systems run. Wi-Fi VLANs are separated and monitored by the firewall. It is difficult to catch the symptom of malware e.g. when a virus tries to spread over a wide area of the campus VLANs. KEK is planning on a new network environment. Monitoring traffic among VLANs and stronger logging will be required.

Security: case study (Romain Wartel and Vincent Brillault)

This case study is flagged TLP:RED, i.e. confidential, for the audience's eyes and ears only.

Storage & Filesystems

Storage for Science at CERN (Giuseppe Lo Presti)

CASTOR at CERN stores 221 PB of data, for a 800 PB capacity. EOS is their flagship disk-only storage project, storing 160 PB for a capacity of 250 PB on more than 1 000 servers. IPv6 is picking up. They just use CPUs for checksumming, hence the Batch on EOS Extra Resources (BEER) project. EOS traffic is dominated by physics and it has been constant throughout the year except for sometimes significant peaks.

They support Ceph for the OpenStack infrastructure. It has a growing role in physics. It bears many reads and writes and has been working fine for years. A niche application for Ceph is the NFS service. CERN has got some HPC requirements. This year, there has been a new deployment of a CephFS instance to provide storage to 100 HPC nodes. They hope to increase their operational experience on CephFS. They have been performing extreme scale tests on a dedicated Ceph instance, which they do whenever they procure disk servers.

CERNBox is a fast growing service. It has been in production for some years. Individual users are pushed to use it instead of AFS. Project areas have been migrated from AFS to CERNBox, lately. A large fraction of the CERN user base now use it. It is not only about CERNBox as storage, but also the application ecosystem. CERN have a collaboration with ownCloud as users of the system and because they contribute back to the open-source code base, too.

The latest integration was that of Microsoft Office, similar to Office 365. This is carried out with Microsoft's WOPI REST interface, that supports the web-based Office applications with any storage as the back-end. It uses a stateless server to ease scale-out and operations. The architecture decouples the WOPI business logic from the storage access. They noticed that the online version of Office is limited compared to the desktop version. Being new software, there is confidence that it will improve. Lately, Microsoft enhanced concurrent editing, thanks to WOPI's locking mechanism. There are alternatives to Microsoft's solution, such as ONLYOFFICE and a commercial solution for LaTeX. Integration with analytics is picking up speed, too. It has been in production for two years, on top of CERNBox, providing Python and ROOT. Support for Spark is being introduced. The applications have several access methods. There is the XRootD native EOS protocol, the most reliant and performant one. But GridFTP, HTTP, FUSE and SMB are supported, too.

EOS is used in the Joint Research Centre in Ispra (Italy) and in AARNet (Australia). Work has been done to dockerise CERNBox and ease the deployment of the entire system, which is completely decoupled from the whole CERN infrastructure. There are activities which go beyond HEP. Giuseppe concluded presenting the Cloud Services for Synchronisation and Sharing workshops.

Questions and comments:

- Helge Meinhard pointed out ownCloud is used in other institutes, too. What could be the impact of the fork into ownCloud and Nextcloud? There is another company, Nextcloud, providing several services. The clients are not concerned by that split in any major way; for the server, CERN anyway do their own thing. The collaboration with ownCloud will continue.
- Jérôme Belleman asked why the rates on CASTOR were lower than 2016. Because the LHC produced less data.

- Xiaowei Jiang asked why Ceph, not EOS? Because of the need for a real filesystem. This is still a weak point of CERNBox today, which currently only provides FUSE mounts.

Managing cloudy dCache storage pools with Ansible (Ulf Bobson Severin Tigerstedt)

They have 7 sites running their data pools. The systems are managed by local system administrators. dCache pools are normally created once, you do not often update the configuration and you upgrade often. Updating dCache takes a full day, in practice. Cloudifying dCache entails separating out local administrator and dCache tasks. But having root is complicated and systems are wildly different. dCache requires Java, bash, logrotate and tar. The solution is to use Ansible, to manage common configuration and processes, grouped by location and hardware. It will help speed up updates. The result is 3 fairly complex but easy to use Ansible roles:

1. `Config` creates the config files and the basic environment.
2. `Upgrade` updates or installs dCache and Java.
3. `Createpool` creates the actual pool.

But this works only for pools, no other services. Besides, it requires SSH access to be opened. Finally, logs might now go to `/home` instead of `/var` like before. Ulf showed the way to the code in GitHub and documentation will follow.

Questions and comments:

- You could ship your own logger to syslog which requires no privileges and would end up in `/var` anyway.
- Ofer Rind asked, do you still need to coordinate with local system administrators? Only for upgrades.

Cloud storage with the Dynafed data federator (Marcus Ebert)

The redirector for a dynamic data federation was developed by CERN. For data transfers, the client is redirected to a storage element with the data, something which can be done depending on the geographic location, when storage elements closer to the job are preferred. One can federate existing sites without configuration changes at sites.

X.509 and VOMS-based authentication/access authorization can be used with DynaFed and also directly access S3 and Azure-based storage, which is easy to manage and means there is no need for extra manpower to manage a grid storage site. Marcus showed the necessary setup to access the S3-based storage and how to retrieve a file, contacting DynaFed which looks up the file and gets the authorised link, redirecting the client who accesses the file over S3. Authentication and authorisation is possible from a built-in VOMS proxy, a Python module or a gridmap file.

There are several use cases for DynaFed: distributed cloud computing, CEs using HTCondor and Cloud Scheduler, and traditional dCache sites. Marcus described the integration of DynaFed for cloud computing and cloud storage for the grid. There are different installations at CERN, Victoria and TRIUMF, with mostly Ceph and S3-based storage. CERN enabled automated test jobs for ATLAS, which are running well. Victoria have an installation where they enabled multiple VOs, used in production for distributing Belle II input files. TRIUMF are testing ways of building and attaching Ceph clusters to DynaFed.

Client tools can get a new redirect to another site if something happens with an already established connection. Tools based on ROOT can use WebDAV. However, files cannot be renamed. The checksum handling and algorithm for S3 are different to what is used for the grid. Third party copies using WebDAV are poorly supported by sites. While experiments do not appear to support it, gfalFS lets you mount whole federations as regular filesystems. On the experiment side, there are problems with data management systems that were developed without DynaFed in mind, such as double counting of files.

Questions and comments:

- Ofer Rind asked, how many sites are you federating? One at CERN, two in Northern America.
- Did you measure performance, e.g. time to access a given file? Not so far. It does not appear to take longer than waiting for XRootD.

The Outlook for Archival Storage at CERN (Michael Davis)

Up until this point, the tape market has been dominated by the LTO consortium. This is changing in that the market is driven by cloud providers. We are seeing some consolidation. The number of manufacturers – both media and hardware – is shrinking. There is a change of technology ongoing, from Giant MagnetoResistive (GMR) to Tunnel MagnetoResistive (TMR). GMR has reached its density limit. TMR is 6× more sensitive. Michael illustrated the different technologies and there are fewer challenges for tape than for disks to increase capacity. It seems from the road map that the trends are realistic. Last week, IBM announced LTO-8. Oracle have not offered any road map, suggesting they are pulling out of this market. This leaves only one manufacturer putting the R&D into tape drive technologies. Oracle's business was based on reselling their drives to HP. The T10K technology was based on increasing the number of tracks and Oracle might have found that this was not scalable any further. Media is 50% of the total cost of ownership. The cost by TB is declining but not as quickly as predicted. The outlook is very positive in terms of projected improvements in tape capacity but what with only one remaining manufacturer, will this market continue to sustain tape research?

The disk market is also in decline due to increased competition from SSD for enterprise. New technologies are being investigated, such as Shingled Magnetic Recording (SMR), helium-filled disks (allowing for more platters) and Heat-Assisted Magnetic Recording (HAMR). The cost doesn't decline as fast as predicted for disks either. Looking at the road map, all modern disks use Perpendicular Magnetic Recording and there are doubts as to whether the predicted density increase is realistic. The SSD price per TB is expected to remain of the order of 10× more than the HDD price per TB for the foreseeable future.

CERN experimented with Massive Array of Inexpensive Disks (MAID), testing 192 disks on one server, offering up to 1.1 PB. Reliability is a lot lower than tape, which have verify-on-write with two heads. CERN have evaluated different filesystems and redundancy layouts. There is no warranty or certification for the CERN use case. The idea is to compensate lower reliability with higher redundancy. Another option is optical libraries. The Panasonic option is not as sophisticated as the Sony one. The optical technology looks promising but it is still early days. An interesting factor in the Sony solution is that it looks very much like tape libraries. The holographic technology is also early days and is about recording information across media volumes, not just surface. Tape still appears to be the best solution.

Michael showed the architecture of CTA – the CERN Tape Archive. The road map remains unchanged since last HEPiX, leading up to bringing experiments online by Q2 2018 to reach production in Q4 2018. In recent developments releases, recalling files in sequential order is not optimal. It would be more efficient to switch back and forth between tracks. This feature has been implemented in the drives. You provide a list to the drive which can recommend you with the access order. It gives considerable speed-up, up to 5×. It only works on the enterprise drives; LTO drives have not implemented this feature, however. But since cloud providers also use LTOs, they might be pushing for this.

Questions and comments:

- Tony Wong looked at the decrease in the shipment of disk drives. Is this worldwide shipment of drives? Is there more disk capacity than there is demand? Mattias Wadenstein suggested that it is due to laptops using SSDs more and more.
- Mattias Wadenstein asked, have you seen indications that IBM might be offering only LTOs? According to the road map, yes.
- Andreas Petzold asked, if you switch to disks, you will have operational costs in terms of consumption. Would the idea be to turn off disks when unused? Yes, but disks are a last-resort option.

- Giuseppe Lo Presti commented, disks and tapes compete with each other. Tapes just keep slightly ahead of disks.

Basic IT Services

Securing Elasticsearch for free: integration with SSO and Kerberos at CC-IN2P3 (Fabien Wernli)

Elasticsearch's problem has always been the lack of encryption, authentication and authorisation: everyone can by default see everybody else's dashboards. The X-Pack solution is not affordable and provides no SSO integration, no Kibana multi-tenancy and no Kerberos support. Search Guard is an open source alternative. It offers the features that X-Pack was missing, plus a few more. Fabien showed a screenshot of how it is installed, which involves a few `wget` commands and editing an `elasticsearch.yml` file, with a focus on the certificate DNs. Search Guard also takes a number of other configuration files organised in a directory hierarchy. Search Guard adopts a role-based authorisation model. By default, all users can read, not write.

CC-IN2P3 use an Apache Proxy to forward headers to Kibana and Fabien showed the Apache and Kibana configuration to explain how. Search Guard presents you with a screen offering a list of tenants. Fabien showed a playable demo coming with the slides. The `--negotiate` switch passed to curl makes it use Kerberos. Anything that supports SSL contexts will work with Search Guard, as will syslog-ng and language bindings with high-level Elasticsearch clients. Basic authentication, on the other hand, does not. What is more, TLS configuration can be tricky and the curl version available in CentOS is too old to support Kerberos. There is no visible impact on performance. There are different support channels: CC-IN2P3 received much help from `floragunncom`, there is a Search Guard Google group and issues are tracked on GitHub.

On Server Management Interface (BMC) (Alexandru Grigore)

The Baseboard Management Controller – BMC – sits on the server mainboard and is functional even when the server is turned off. It runs several sensors. There are system event logs, iKVM, Serial On LAN (useful to debug boot issues), controls for the fan speed, and so on. IPMI specifies steps to access BMC features. There are multiple vendor-specific implementations. The IPMI specification only provides limited information about the system. Implementations come with custom OEM extensions. Security was largely ignored, authentication is missing and passwords are stored in clear.

To prevent BMCs from being compromised, they separated the networks for management and data traffic. Not all systems provide the management interface, however. A CERN BMC firmware update campaign started with Intel, Supermicro and Quanta. There were bugs in the firmware preventing login. What CERN asked for was to support TLS and a way to upload certificates. They sanitised the situation for most servers. It was a lengthy process, involving the help from other teams at CERN and the manufacturers.

The Redfish API offers a RESTful interface to get information from BMCs. Access is carried out over HTTPS. Alexandru showed some URI examples and a Python script to query them. You can do everything you would with a BMC. OpenBMC is Facebook's approach, running a custom Linux distribution running on the BMC.

Questions and comments:

- Andrea Chierici asked, do you have the problem of having to access the console via a Java applet? Yes, but soon HTML5 implementations will be available.
- Andrea Chierici asked, can OpenBMC be installed in existing machines? Only those with an open firmware.
- Did you record any security incident? No, old BMCs do not log events. Vincent Brillault added that IPMI is kept on a separate network, now. Mattias Wadenstein continued, not all systems have separation of networking. Any comprised host gains access to the BMC. Thomas Bellman added that they had occasions when servers magically reset themselves causing BMCs to appear on normal networks.

- Jérôme Belleman asked, is the Redfish API only for querying or does it also perform operations such as rebooting? It also supports rebooting – in fact, all operations supported by BMC.

riemann: a different stream processor (Fabien Wernli)

"Riemann filters, combines, and acts on flows of events to understand your systems," to quote their website. It is related to Apache Spark streaming, Apache Storm and Kapacitor. Riemann shows a low latency. It is stand-alone. Fabien showed how it is installed, and how collectd, syslog-ng are configured. Riemann is written and configured in Clojure. There are several output storages supported such as Graphite, Librato, even e-mail. An event has three dimensions forming several metrics such as start and end times, TTL, host and metric.

Riemann has a rich API made of stream functions. One can limit the rate, only record events which change according to a parameter, aggregate or split events, ignore flapping, queue, classify or run custom operations on them. The Riemann index keeps track of the most recent events for a given host and service. Fabien showed a real-life example of collecting load average and alerting by e-mail when a threshold is crossed. Another use case CC-IN2P3 have is pre-aggregating data. Riemann comes with a unit test framework. You can use it to restart the Riemann service only if the tests still succeed.

Questions and comments:

- Dennis van Dok asked, can you explain what kind of resources you need? Riemann is quite compact. People reported having millions of events per second. CC-IN2P3 use it on VMs.

Integrated Monitoring results at IHEP (Qingbao Hu)

IHEP's HTCondor cluster runs 13 500 cores and their SLURM cluster runs 2 750 cores. They are a T2 site and use CREAM CEs. They have got 5 PB worth of LTO-4 tapes managed by CASTOR 1, 9 PB in Lustre and 1.2 PB in EOS. The scale of their cluster is expanding and the computing environment is becoming more complex. The various cluster monitoring tools they run are independent and they wish to correlate data among the different monitoring sources and present it in a unified view. They collect data into RRDtool, MySQL, MongoDB and log files, process streams with Kafka and analyse them with InfluxDB, Elasticsearch and HDFS. They use the Logstash plugin to read from multiple sources, and the Logstash filter plugin to pre-process data.

IHEP have put together a flexible configuration framework for data collection. They process data to enrich cluster job information. They collect job attribute information when a job finishes, such as the worker nodes, times and slot IDs. Qingbao showed a snapshot of the batch data they collect from the separate sources, and once combined, including the enriched fields. For Lustre, they also collect data for alarming. In particular, they stumbled upon *lock callback timer expired* errors on which they want to trigger notifications. They use a trainable classifier to choose between raising the alarm or simply collecting. Qingbao showed a number of Grafana screenshots showing the different attributes that were forwarded, including those which resulted from enrichment.

Wigner Datacenter's new software defined datacenter architecture (Zoltan Szeleczky)

The Wigner Datacenter hosts T0 resources for CERN. They want to automate and use open source as much as possible. They actively use GitLab, which they are integrating with Gerrit and Jenkins. The OpenStack deployment is automated with TripleO. It stands for *OpenStack on OpenStack*. There is the concept of an undercloud and an overcloud. They chose to automate adding servers to the overcloud with Katello Discovery, set an IPMI fixed IP from which a JSON file is generated. They use YAML files to describe the environment. Adding more nodes is a matter of changing the counts in the YAML file. For development and release, they devised a 3 steps process: they have a development environment of 3 nodes, a test environment of 9 nodes and a production environment.

The firewall they use is called OPNsense which is problematic for its lack of API support. There are also bugs, such as the port configuration that turns off all the ports. They use two factor authentication with YubiKeys. So far, they have a fully virtualised infrastructure. The 3 different environments allow for

constant testing. However, the overcloud metadata virtual IP (VIP) does not work. Power outages can cause database corruptions. They also have to deal with frequent bugs in the TripleO stable repository, which is why they decided to introduce the 3 steps process. Bugs are normally fixed within a month.

Questions and comments:

- Edoardo Martelli asked, what throughput do you get through the firewalls? We can scale out by adding multiple servers. But we have not made any measurements.
- Pete Gronbech asked, is OPNsense and fork of pfSense? Yes.

CSNS HPC Platform Based on SLURM (Yakang Li)

CSNS is the China Spallation Neutron Source, for studying neutron characteristics and exploring the micro-structure of matter. They need a high performance computing environment. Yakang presented their data flow. They use software such as ORBIT and Geant4, based on MPI. Data reconstructions are also based on MPI and GPUs. They rely on cloud computing, mainly for instrument users. They use SLURM to provide an MPI parallel computing environment. They soon will finish setting up 500 cores for their cloud and 2000 cores for their HPC, together with a few GPU cards. The login farm is made of two parts, the load balancer and the login nodes. They use LDAP to manage users. GlusterFS provides space for experimental data and home directories. Redundancy is ensured with RAID 5. Software is kept on CVMFS. The HPC web user interface shows a status monitor, a feature monitor, job management, etc. Yakang showed a few screenshots. The job feature view lets users see snapshots of the CPU usage. The job management view lists jobs and their status. The supporting systems are based on the ELK stack. CSNS ran some benchmarks with LINPACK. The next phase will be to add GPU computing nodes and large memory computing nodes to the HPC platform. Support for a Docker-based flexible job HPC system is ongoing.

Questions and comments:

- Mattias Wadenstein asked, what do you use as metadata catalogue for long-term storage? ICAT.

Updates from Database Services at CERN (Andrei Dumitru)

Databases are used for administration, accelerators (in particular the logging service) and experiments (for example for online operations), replication for other grid sites, analysis and more. They run Oracle 11g on 100 instances. Depending on the criticality of the service, there are more or fewer replicas. Recently, they introduced a new Oracle database service with the UTF-8 character set. WE8ISO8859P1 was used before. The Oracle Connection Manager is provided by Oracle to proxy access to databases from outside the CERN network. Andrei showed the architecture of this service. There is a tunnel that can be used to access the CERN network. Oracle Tools were moved from AFS to EOS, e.g. the `tnsnames.ora` file which they maintain.

The Database on Demand service offers MySQL, PostgreSQL and InfluxDB. The difference with the Oracle service is that users are their own DBAs. However, they cannot reach the underlying hardware. Nevertheless, they can upload configuration files through a specialised interface. There are 550 instances, some of them replicated. Oracle used to be part of this offering for some use cases, but is no longer available for new DBoD instances. The database team are migrating their infrastructure to CentOS. High availability is implemented with asynchronous replication i.e. hot standby.

Time series database are normally for inserting or appending data workload, where recent timestamps especially matter. Large deletes are expensive. Query operations are different from relational databases. Individual points are not too important, aggregating data and large data sets is a common operation and calculations are time-centric. After some evaluation, InfluxDB looked the most promising one. Several projects use it in production. Most use cases work fine with a single instance. But large needs mean splitting into several instances. The project continues to be very active and is evolving fast. There are constant performance improvements.

The Hadoop service covers the analytics use case. They provide consultancy to users. It is an activity that has been going on for over 2 years. It is made of many different components, ranging from Kafka, to

Zookeeper, Flume, Impala, Spark and Sqoop. There are 3 production and 1 test clusters, the most powerful one of which is intended for the accelerator logging system. The next version will be based on all that infrastructure. It is the largest database – 700 TB, growing by 200 TB every year. It will be based on Kafka, Gobblin, HDFS or HBase, and Spark, which is integrated with SWAN (Service for Web-based ANalysis). It offers Jupyter notebooks for data analysis.

Questions and comments:

- Jérôme Belleman asked, do you have a similar proxy service for accessing Database on Demand databases from outside CERN? No.
- Pepe Flix asked, are you monitoring how many analytics users you have? It is unclear how many of them there are.

Deployment and monitoring for distributed computing sites (Wei Zheng)

IHEP have remote sites of distributed computing, where jobs are dispatched via DIRAC. Due to the restrained manpower, even small errors can have considerable consequences. IHEP need to provide maintenance for those remote sites, as well as monitoring. Their local installation is made of 3 Puppet masters, 1 Puppet CA and 1 mirror host. A software upgrade to 2 000 hosts takes less than 15 minutes. In their remote site, they have 1 Puppet master.

They have been evaluating distributed monitoring tools, namely DNX, check_mk and Nagios with Mod-Gearman, which presented all the necessary features such as easy scaling, proxy support and central configuration. Wei described the architecture around Nagios and Mod-Gearman. If the monitor identifies an error with a HTCondor node, the node will be removed quickly and make it join the cluster again as soon as the issue is solved. This increased the success rate of jobs. They have set up dashboards based on Nagios. Their 5 sites run 2 500 hosts running 30 000 job slots and each host or service is checked every 5 minutes. They added notification integration with WeChat.

The central site is IHEP and online administrators are responsible for the regular operation of all the other sites. Remote sites are monitored by Nagios and IHEP administrators deal with all errors. Three remote sites have joined this infrastructure so far. In the future, the overall availability and reliability of reports from remote sites will be provided. The deployment process will be streamlined, too, and there will be more fine-grained monitoring.

Questions and comments:

- Jérôme Belleman asked, in slide 6 you said the software upgrade to 2 000 hosts takes less than 15 minutes. Is that deploying configuration with Puppet? Xiaowei confirmed this.

Automatic shutdown of servers in case of A/C failure (Peter Gronbech)

When there is a major failure in the chillers the temperature in a 250 kW loaded computer room rockets up in a matter of minutes. Monitoring temperature in different parts of the room is critical, especially out of working hours. Pete showed an incident in 2015 where temperatures hit 50°C. The computing room is used by different groups. Controlling air conditioning can only be carried out by the air conditioning experts. The idea was to measure and publish temperature for everyone to check.

Pete showed pictures of sensors and how they can be connected with a USB cable under the floor grilles. Beyond measuring specific spots, they would like to know the overall temperature. They wrote a Python script using Matplotlib, NumPy and the `logging` module to collect and plot data. You can pass the shutdown temperature to the script. It is a pretty simple system which can let them sleep more soundly.

Questions and comments:

- Martin Bly asked, does the monitoring system tell you when a node has been shut down? Not yet.
- Thomas Bellman asked, do you have something to also power off the room? Not yet. It would be advisable if the room reaches far higher temperatures.
- Xiaowei Jiang asked, how do you define thresholds? By trial and error.

Friday 20 October 2017

End-User IT Services & Operating Systems

Modernising CERN document conversion service (Ruben Domingo Gaspar Aparicio)

The aim of the conversion service is to convert files between different formats. The workload mainly involves Office documents. The conversion is carried out with a REST API. The old service has been in operation for 10 years. It is a Python multithreaded application. It lacks logging and load balancing. In the last 6 months, over 100 000 files were processed, most of them Word documents.

The goal of this overhaul was to keep the same user interface while redesigning the program architecture. Adding conversion should be as simple as possible. The new architecture now involves an OpenShift container, a queue on EOS, statistics on Elasticsearch and logging into PostgreSQL. The new doconverter is an open source project, 50% faster than the previous technology, faster to scale. Ruben showed how to configure the system in an INI file and a list of possible conversions.

The service uses Elasticsearch to run real-time statistics. A Puppet module is used to configure Logstash. There are dashboards showing the number and size of documents, the success rate and more. Using EOS as a back-end allows to keep state among the different players. There were instabilities accessing it via SMB. The team evaluated several candidates for an OCR service and now offer MODI OCR and Tesseract. The resolution of the input document is key.

Questions and comments:

- Jérôme Belleman asked, is it only Windows? For the time being, yes.
- Jérôme Belleman asked, is it the right service for working with LaTeX files? Not easily yet, but maybe when it will run Linux, too.

A user portal at CC-IN2P3 (Renaud Vernet)

CC-IN2P3 is different from IN2P3, Renaud clarified. It is a computer centre running 30 000 slots, 25 PB on disk, 50 PB on tape and various T1-style services for WLCG and other experiments. They serve several physics communities with different needs. Users currently use several monitoring and documentation pages. CC-IN2P3 wish to centralise these resources in a simple-to-use solution which they will be able to release soon. They would like to add alarming, too, for batch jobs failures, password expirations and downtimes, for instance. The intention is also to streamline information in a homogeneous way.

Renaud made a demonstration of the user interface. Once logged in, a portal shows job success rates, static information about the account, storage quota details and a news feed. It is available as a French and English version. There is a notification menu, for quota being full, password expiration, etc. There is a variety of widgets. A page is focused on each service, e.g. batch jobs, where it shows the number of running, pending jobs, available queues, average times, etc. The storage page can show group usage, too.

They contacted 15 beta testers – active users and group managers – and asked whether the testers find it useful. Feedback was overwhelmingly positive. Some users requested more advanced features. Comments were more focused on contents than appearance.

OpenIDM covers identity management. Home-made scripts collect data. The Lavoisier framework aggregates data. Bootstrap handles CSS. It all runs on a single, small VM. Collected data is streamed into Lavoisier and web pages are built into the portal. They will add workflows (e.g. changing the password), implementing single sign-on and integrating their ticketing system. The portal will be moved online in the coming weeks and they keep looking forward to user feedback.

Questions and comments:

- Jérôme Belleman asked, will it be possible to check batch jobs for whole groups, if individual users log in? There does not seem to be a requirement for this at this stage, but they will provide this feature if the users need it.

- Will you stop sending notification e-mails? They will keep them for some time and there are no specific plans to stop them.
- Is there any chance to change password policies such as password validity?
- Tony Wong asked, what kind of hardware will be expected to be required when this becomes production? It is not currently a critical service. It is not clear how to set up high availability yet.
- Helge Meinhard commented, some labs have been suffering from outages such as trees falling on power lines. You should keep in mind that the system will be used to notify of emergency situations.

Continuous Integration for Linux Images (Jérôme Belleman)

Jérôme presented the context, the Linux team being in charge of generating the Linux operating system images used for the majority of servers at CERN. They are built with Koji, a system which is primarily used for making RPMs. Given how widespread they are, it is critical that the images be built quickly, autonomously and that they work flawlessly – which is the reason for this project. Currently, some manual checks are preformed before production release, such as trying to start a VM from a given image. This being a manual and tedious operation, it is instrumental to automate test execution, which will greatly speed up the process and allow to run more comprehensive tests, and will lead to the need of managing them.

This project is leading towards running a greater variety of tests. At its heart, this testing framework uses GitLab CI with a job to build an image from Koji – a prerequisite and a test in its own right. Depending on its success, another job to create a VM from this image will be run. GitLab CI jobs are typically Docker containers which provide a rather bare environment, and it is necessary to set up an OpenStack client environment before being able to use them to spawn a VM. Once successfully created, jobs running specific tests can be run on the VM. CentOS have made available a suite of functional tests covering a wide variety of aspects in the CentOS distribution, from the kernel to the X Window System. By configuring each individual job to SSH into the test VM and running a specific test from this suite, it was easy to get GitLab CI to categorise tests, define dependencies between them and graphically show the progress and results.

The problem of tests setting up their environments such that no other tests can run on the VM was brought up. Certainly, arranging for a fresh VM before running each set of tests seems a good idea. This could lead up to a pool of VMs in a specific OpenStack project and workflows to manage them. Another option would be to configure a GitLab CI Runner as a VM, rather than leaving it to a Docker container to log in into it.

While this project is first meant for the CERN Linux team to streamline their image production, it could be of some use to other communities, too. In fact, some users at CERN already expressed interest. The framework could enable them to run their own tests against their own images. This could involve git workflows allowing users to define their tests in a `.gitlab-ci.yml` file of their own. They could be provided with a CLI to run specific tests against a given image. Or they could send merge requests directly from their git repository. To query the test results, it would be possible to present clear views to the user from the CLI as the GitLab REST API makes it easy to retrieve them. Alternatively, the graphical view that GitLab presents by default in its web application could be sufficient.

Some issues to look out for include GitLab CI jobs which run for too long, the number of jobs run in parallel, especially if many users get to use this framework, and the question of whether or not to adapt some of the CentOS functional tests, especially the ones disruptively adjusting their Linux environments. Using the same framework to test Docker images could be an opportunity, as it is expected that it might come down to the same process.

- Fabien Wernli pointed out, could running Docker inside the test VM be an option to work around the problem of tests modifying their environment? It would be an approach to consider.
- Dennis van Dok asked, what is the capacity of the GitLab cluster? It is not clear, but Jérôme will ask the experts.
- Dennis van Dok asked, is the CERN Koji instance public? No.

- Quentin Barrand asked, how to run CentOS test individually in GitLab CI jobs of their own? By defining a job for each of them in the `.gitlab-ci.yml` file, which will instruct the container to SSH into the test VM to run a specific job. CentOS tests come with the `runtests.sh` script which can run all tests or some specific ones only.
- Dennis van Dok asked, would using snapshots be an alternative to recreating fresh VMs? This is a very good idea.

printing@gsi (Stefan Haller)

There are many different printer manufacturers and models and while printers are mostly managed by the GSI IT department, some devices remain out of their control. There are many printers for Windows, Linux, Macs and guests via CUPS. To improve security, all printers are being segregated into a separate VLAN. Access to printers is only possible via a dual-homed Linux/Windows print server. Linux printing at GSI stopped using LPRng in favour of CUPS.

The printer list is kept in an Excel file in a Windows share. Their move to Chef was an opportunity to programmatically parse the Excel file with a Ruby script which assigns a PPD file to a printer model. Stefan showed a sample JSON file for a specific printer. They use the `cups` Chef cookbook and developed the `gsi_cups` one, which deploys additional PPD files and removes outdated print jobs with cron. There is currently no automatic addition of new printers and data bags for disabled printers are not removed. In the future, they will add Kerberos authentication, accounting and autodiscovery.

There is a small subset of printers on the guest network, managed by the CUPS forwarding server. Public printers are manually chosen and data bags transferred with Chef. To improve their security, they try to check empty, default or easy to guess passwords, publicly accessible shares, outdated firmware versions and disable unused features.

Questions and comments:

- Fabien Wernli asked, why are Linux and Windows separated? This is unclear.
- Is the network separation physical? No, VLANs.

First experience with SELinux (Michel Jouvin)

Who uses and is familiar with SELinux? Not many of us in the audience. There is a tendency to just disable SELinux. Version after version, however, its coverage appears to expand. OpenStack comes with SELinux support out of the box. SELinux helps you set up confinements.

Michel explained the basic concepts. An identifier is attached to every subject and object. They can be files, processes, ports, … The format looks like `user:role:type[:mls]`. Michel gave the example of `system_u:object_r:httpd_sys_content_t`. There are 4 types of access controls for the default so-called *targeted* policy. Only the context type is checked by default. There is Role-Based Access Control (RBAC) and Multi Category Security (MCS).

SELinux implements Mandatory Access Control (MAC), i.e. everything that is not explicitly allowed is forbidden. It is not an alternative to permissions, it supplements them. In essence, SELinux applies policies as exceptions to the *no access* one. They can be defined or supplemented with specific languages. In targeted mode, a policy is required for every process that has a defined SELinux access type.

SELinux presents different modes, in which it can be enabled, enforcing policies or be permissive about them, in which case events are only logged. This is how you can check the potential problems that enforcing a given policy would raise. As a result, permissive is more useful than disabled. The file context assigns labels to files. Contexts can be worked with using the `restorecon` and `chcon` commands. Interestingly, rsync provides the `-X` option to copy SELinux contexts. One can customise or extend SELinux policies. Most of them use booleans. Each application requires you to specify the allowed ports. The `audit2allow` processes errors in the log file to create policies but it is best to use `semanage` and booleans instead. Likewise, it is a good approach to work around problems without completely disabling SELinux by only disabling it for a given service.

With Apache, and possibly other web servers, it is useful to set the appropriate type for each category of files/directories, in particular if you relocate them, and to set the appropriate booleans to enable access to external services. With NFS, SELinux allows to define a security context overriding the file context. NFS enforces an SELinux context at the filesystem level. There is a potential issue with a web server in particular, if you want to support both read-only and read/write areas from the same filesystem. There are many useful links and documentation, but Michel discouraged using `stackoverflow.com` because users provide ready-to-use recipes without explaining them, and when you start getting used to SELinux you will find that they are sometimes inadequate.

Grid, Cloud & Virtualisation

Cloud deployment at KEK (Wataru Takase)

KEK provides a 10 000 cores Linux cluster on SL6. Users log in into an interactive login service and submit jobs to LSF. There is a need for different workloads for different groups, which may be using different operating systems. There is also a desire to improve management efficiency and this lead to cloud computing. CMO stands for IBM Cloud Management with OpenStack and allows simplified OpenStack deployment with Chef. It comes with a service portal summarising all OpenStack projects and lets users manage their nodes. The IBM Platform Resource Scheduler extends a Nova Compute scheduler using a policy-based VM deployment and reallocation system. Chef automates deployment by defining topologies beforehand. In KEK, they use OpenStack Kilo, the latest release supported by CMO.

Their cloud aims at providing batch integration and self-service provisioning. Wataru showed a flowchart where LSF uses a resource connector to launch special instances via CMO. The self-service provisioning service provides a simplified portal to control VMs according to their role. CMO also takes care of launching the instances. They support SL6, CentOS 7 and Ubuntu 16.04. The cloud integrates with LDAP for authentication. They prepared a default Keystone domain (querying a DB) and an LDAP Keystone domain (contacting the LDAP server). GPFS is used for home directories and shared group directories. NFS mounts are made available to compute nodes. Authorisation management is group-based with SSH. OpenStack projects are mapped to UNIX groups. The SSH access control is carried out with cloud-init. They use node aggregation to implement group-based access.

They reached the test phase and in the future they will integrate their batch system with public clouds to temporarily increase their resources. The resource connector will enable this.

Questions and comments:

- Jérôme Belleman asked, will the CMO allow you to implement workflows, e.g. to destroy ill nodes and create new ones? No, there is no orchestration.
- Do you dynamically feed groups with resources? This is under consideration.

Miscellaneous

Workshop wrap-up (Helge Meinhard)

The statistics combine HEPiX and LHCOPN/LHCONE. There were 149 participants for both, but it is expected that about 115 of them took part in HEPiX. There was a healthy mixture of experienced and first-time attendants. There was a very good Asian/Pacific representation. There were 56 different affiliations, including 9 from Asia/Pacific, which is unique. Helge displayed the group photo, showing happy faces who were also happy about the only sunny day of the week.

There had been common sessions with LHCOPN/LHCONE and HUF. While there was little commonality with HUF, there were common interests with LHCOPN/LHCONE. This is an experience to repeat. Most attendants were positive about this.

There were 83 contributions, 28 hours and 45 minutes scheduled, which made for a rich program again. Shorter presentations were introduced, too. The presentations were of high quality. There was a

particularly high number of network talks, especially related to the LHCOPN/LHCONE. Storage talks mentioned Ceph and DynaFed quite often. Cloud presentations covered containers and how to use them in batch systems more and more. In computing, investigations to find a fast benchmark made good progress. HTCondor keeps growing and HPC-HTC integration was mentioned several times.

Only 1 contribution on IT facilities this time, about collecting data and measurements of air quality. Stream processing and time series DBs for basic IT services are ever more in the focus. There was a record number of 20 site reports. Multiple sites have updated their network, look into SDN, GPUs. IPv6 is slowly progressing.

As always, coffee breaks, lunches and social events offered countless hours of valuable discussions. From the board meetings, we are settled for future workshops until 2020 at least. Working groups were reviewed. The Network Function Virtualisation Working Group was established. Wikis need action and the board are checking whether we can move to GitHub Pages. Tomoaki was invited to join the board. It has been decided to extend co-chair terms to 4 years and Tony was re-elected as co-chair until autumn 2021. There were considerations about implying Asia/Pacific next time round. Spring 2018 will be hosted by the University of Wisconsin, Fall 2018 by PIC.

Helge thanked all participants, speakers, track conveners and chairs, sponsors and the local organisers: Takashi Sasaki, Tomoaki Nakamura, Koichi Murakami, Go Iwai, Fukuko Yuasa, Soh Suzuki, Tadashi Murakami, Wataru Takase, Hiroyuki Matsunaga, Noriko Kagaya, and their staff. Helge wished us a good trip back.

The KEK Facility Visit

The tour took place after lunch, after the workshop ended. We took buses for a 5 minutes drive through the KEK campus which showed alternating scenes of industrial-looking buildings and expanses of green grass and trees. We stopped at the entrance of what appeared to be an ordinary white building labelled *Tsukuba*. It hosted the Belle II detector experiment hall.

The Belle II Detector

In the entrance of the building, our guide showed us on a map the location of three other halls around the SuperKEKB accelerator, *Fuji*, *Oho* and *Nikko*. The entrance also proudly kept on display Makoto Kobayashi and Toshihide Maskawa's 6 pages paper on *CP-Violation in the Renormalizable Theory of Weak Interaction*, which lead to the award of the 2008 Nobel Prize in Physics. Opposite the paper, a large poster showed a cross-section of the Belle II detector. The detector is 7 m high. At collision point, its beam is squeezed to 100 nm. The trigger system cost half the price of the whole detector. At its heart, the 8 MP DEPFET – DEpleted P-channel Field Effect Transistor – pixel detector collects data at trigger rate. The wire chamber is located outside of it. The detector is made of 6 layers. Due to their cost, the outermost ones were reused resistive plate chambers which were installed in the Belle detector in 1999.

The rather unique smell and imposing hum of the detector facility greeted us the moment we entered the experiment hall. We were lucky to be able to see Belle II in its final installation state. The calorimeter was still visible and the vertex detector not yet installed. This operation is carried out last, as it is a delicate process. The helium-cooled superconducting final focusing system enclosed in a shiny, bullet-shaped enclosure was also kept outside of the detector for the time being. The digitisers are made of custom, rack-mounted electronic boards on top of the detector and the digital signal extracted through optical fibres. Our guide pointed to the 25 flags representing the 25 member states of the Belle II collaboration. Israel joined last. There were in total 700 collaborators. During some meeting weeks, some 240 of them came to the KEK campus. There are offices located behind glass windows. Staff will still be able to work in this area thanks to the low radiation levels which will emanate from the detector.

Our guide hastily took us across the road to a small museum where relics of old accelerators and detectors were on display. He showed us an old accelerator section made of quadrupole and sextupole magnets, and a dipole corrector. Back in those days, beam pipes used to be made of copper, whereas they nowadays use aluminium. A 2×2 m ARES – Accelerator Resonantly coupled with Energy Storage – sat in a corner of the museum. There were some components of an old linac and superconducting cavities for KEKB. The guide showed us a DEPFET, which he presented as the device which lead to the Nobel Prize. Next to it, a strange, foggy yet solid, sticky, almost invisible, extremely light and friable material called aerogel was kept in little boxes. It provides good time resolution to the photo multiplier. The aerogel is a medium for Cerenkov light; it allows for speed measurements and – if the momentum is known from bending in a magnetic field – particle identification.

The Data Centre

We took our buses back to the conference venue and crossed its building towards the data centre, a few minutes' walking away. In the for once dry weather, we could hear crickets in the large patch of grass we strolled next to. We entered the data centre building and the first of the two computer rooms. Lenovo sponsors were here as well to answer our questions, but the noise was too deafening to chat. It came from our right from 6 dense racks of 358 Lenovo NextScale CPU nodes running 10 024 cores and 55 TB of memory. Their high speed interconnects allow for a 100 GB/s throughput. Behind, racks of 41 Lenovo x3550 M5 servers offer grid services. To store experiment data, the third row of racks are 8 IBM ESS storage servers of a total GPFS disk capacity of 10 PB, with a 600 TB front-end disk rack using software RAID. Mellanox and Cisco units route network traffic to LHCONE and SINET5.

There were grilles at the foot of each racks with fans underneath to cool down the equipment. We could find carefully-oriented office fans here and there on the floor to contribute to cooling the servers.

In fact, the computer room had a lab corner with a whiteboard, which begged the question as to how the staff can possibly have discussions in such a noisy environment.

The next room hosted the HSM equipment. The HPSS disk cache middleware on our right was run by a DDN SFA14K providing a capacity of 3 PB. It uses a HPSS/GPFS interface. Behind, the IBM TS3500/TS1150 tape libraries with space for 70 PB sheltered 2 robots handling IBM enterprise media.