Track 7
Clouds, virtualization and containers

D. Dykstra (FNAL)
F. Hernandez (CC-IN2P3)
A. McNab (U. Manchester)
M. Sevior (U. Melbourne)

Some figures

23 oral presentations (half-track)

22 posters

Well-attended sessions

Trends/frequently seen themes

- "Lightweight" sites
- Much more new interest in containers than VMs
  - Docker, singularity, sometimes both
- Using clouds to manage demand variability, and own a smaller baseline of hardware

Lightweight sites

- LW Sites have resources but want to run HEP work without lots of effort
- VMs/Containers are key technologies in this emerging area
- Several solutions discussed: SIMPLE, CloudScheduler, Vac/Vcycle, DODAS, ElasticCluster
- Useful BOF on Thursday to bring people face-to-face
- Ongoing WLCG Working Group

Lightweight site approaches

- Three broad approaches:
  - Make it easier to install a conventional site (eg SIMPLE)
  - “Parachute” appropriate VMs on to local cloud resources
    - (eg CloudScheduler, Vcycle)
  - Radically different ways of running conventional jobs in boxes
    - (eg BOINC, Vac)

Optimisation of cloud platforms

- On-premises OpenStack platforms of significant size used for both science workloads and infrastructure services, e.g. CERN's: 300 K virtual CPU cores, 3K+ users, 9K+ hypervisors, 40K+ VMs
- Work on optimization of OpenStack for reducing virtualization overhead: reached 3%
- Additional OpenStack services being progressively put into production, including provision of baremetal machines, file shares among VMs, container orchestration, etc.

See: B. Moreira, J. Castro Leon

Optimisation of cloud platforms (cont.)

- Improving the mechanisms for sharing the resources (e.g. VMs) allocated to different projects, to increase overall resource utilisation
- Instead of a fixed quota per project, use low and high threshold quotas to allow for sharing otherwise unused resources
- Software agent monitors job queues and send jobs to cloud platforms where resources are available for that project: being used by LHAASO experiment at IHEP

See: H. Li

Large scale cloud usage, cloudbursting

- CERN made use of 10K core baremetal cloud resources from Oracle for a couple of months
  See: K. Wojcik
- Helix-Nebula Science Cloud (plenaries, S. Pardi poster)
- HEPcloud simultaneously used a million cores on an HPC (track 3)
- Belle II experimented with running simulation jobs in containers on Amazon AWS and Microsoft Azure
  - No additional servers needed outside the cloud
  - Small scale but they verified viability

  See: R. Grzymkowski

Integrating HPC resources

- CERN has its own HPC resource and managing it with as many of the common grid/cloud tools as possible
  - Backfill with grid jobs to improve utilization — See: P. Sanmillan
- The University of Torino developed a similar system for managing an HPC they call OCCAM — See: S. Bognasco poster
- The University of Freiburg integrated their HPC system as a full-featured ATLAS Tier2/Tier3 — See: F. Buhrer poster
- Some HPC sites do not allow cvmfs and so “fat” containers need to be built with full software stack — See: W. Yang poster

Backfilling with containerized BOINC

- ATLAS@home developed an option to run “native” jobs on Linux hosts, using singularity containers on non-RHEL hosts
  - Friendly hosts are asked to install cvmfs and singularity
  - Most of them are already grid sites, using BOINC to improve utilization
  - Runs simulation jobs only (which is half of all ATLAS jobs)
- Traditional version uses Virtual-box to run on predominantly Windows PCs
- 85% of ATLAS@Home BOINC uses containerized version
- At times ATLAS@Home provides 3.5% of all CPU resources for ATLAS

See: W. Wu

Opportunistic usage of online farms

- Both ATLAS and CMS working on improving their mechanisms for exploiting the computing capacity of high level trigger farms (~74K CPU cores each) for offline processing during inter-fills and no-beam periods
- Constraints: to be able to switch fast between data taking mode and offline processing mode
- Using OpenStack overlays to build and populate a cluster and launch compute nodes

See: F. Berghaus, D. Da Silva Gomes

A surge in singularity usage

- Singularity is a great lightweight tool for allowing unprivileged users to run containers
  - Especially good for isolation between jobs run by a pilot, and for flexible control of O.S. versions
  - Use cases for it are popping up in many places
- CMS requires all its sites to install it
- ATLAS has put it to some use, has bigger plans See: A. Forti poster
- University of Glasgow created Container Pilot for singularity containers to manage their own lifecycle See: G. Roy poster

Managing VMs and containers

- UK grid sites use a tool called Vac to make it easy to manage VMs, and last year it was extended to also manage docker and singularity containers
  - A mix of all 3 types can be running on the same host
  - < 4k lines of python, available for others to use See: A. McNab
  - Also have a monitoring system called VacMon See: A. McNab poster

Virtualized platforms for infrastructure services

- IHEP has an OpenStack platform for web hosting purposes on VMs, instrumented for collecting network traffic data
  - Purpose: monitoring and detection of abnormal behavior that could be related to security threats or incidents

    See: Y. Tian
- ALICE is using grid services in pre-made containers, "services in a box", run from docker

  See: M. Storetvedt
- LHCb also makes docker containers available as "black boxes" for sites to run DIRAC services

  See: A. McNab

Dynamic services for clouds

- Compute Canada developed an accounting and monitoring system for clouds, used by ATLAS and Belle II
  - Cloud VMs are instantiated by CloudScheduler and are instrumented to upload information to ElasticSearch at CERN
  - Securely transfers authorization secrets to the VMs

See: R Sobie

- They also integrated Dynafed and Rucio for the ATLAS data management system

See: F Berghaus poster

- DESY developed a system for batch system monitoring using singularity containers and Google cAdvisor

See: T Hartmann poster

Apache Spark, SWAN, Up2U

- Apache Spark is an open-source parallel processing framework that allows for sophisticated analytics, real-time streaming and machine learning on large datasets
- CERN has deployed Spark on several clusters
  - 18 to 48 nodes — See: P Kothuri
  - For accelerator logging, computer center monitoring, and CMS “big data”
- Integration of Spark clusters with SWAN (Service for Web-based Analysis) enables interactive data exploration and analysis from jupyter notebook interface — See: E Saavedra
- SWAN used for EU Up2University ed. project — See: E Saavedra poster

FPGAs in the cloud

- Amazon AWS and Microsoft Azure have FPGA cloud resources available
  - Azure's FPGAs support distributed tensorflow models
- There has been some experimentation with FPGA as a Microservice (FaaM) with an FPGA coupled to a CPU

See: R Kunz

Other posters

- Sharing Openstack private clouds
- Openstack accounting management
- Kubernetes for object store stress testing
- Fair share scheduler for OpenNebula
- Oracle Weblogic on Kubernetes
- ComputeOps: container for High Performance Computing
- Extending a site with containers and cloud technology
- A web-based service for parallel applications on the JINR cloud
- Advanced Scheduling in IaaS clouds
- A Software Management and Testing Platform-as-a-Service for Research Software
- OpenShift PaaS as key to the evolution of CERN Web Services
- Collaborative Scientific Authoring at CERN

Overall session link

- Track 7 session