StackHPC and King's College London collaborate on extending the reach of OpenStack Kayobe to Ubuntu

Bristol, England StackHPC Ltd. are pleased to announce today, the culmination of a major collaboration with King's College London on the provision of a Research Computing environment built upon StackHPC's preferred OpenStack distribution environment, OpenStack Kayobe, with significant input from King's e-Research staff in making this environment available on Ubuntu Focal 20.04. Previously, OpenStack Kayobe was only supported on CentOS and latterly CentOS Stream (see the following for a list of supported options) .

The new OpenStack deployment forms the core of the newly launched King's Computational Research Engineering and Technology Environment (CREATE). CREATE provides the following services:

- CREATE Cloud: a private cloud platform to provide flexible, scalable development environments and allow users greater control over their own research computing resources using virtual machines;
- CREATE HPC (High Performance Computing): a compute cluster with CPU and GPU nodes, fast network interconnects and shared storage, for large scale simulations and data analytics;
- CREATE RDS (Research Data Storage): a very large, highly resilient storage area for longer term curation of research data;
- CREATE TRE (Trusted Research Environment): tightly controlled project areas making use of Cloud and HPC resources to process sensitive datasets (e.g. clinical PIID) complying with NHS Digital audit standards (DSPT);
- CREATE Web: a self-service web hosting platform for static content (HTML/CSS/JS) and WordPress sites;

The previous Research Computing infrastructure included both an OpenStack cloud and a batch-queued Slurm HPC infrastructure. The HPC infrastructure had been procured over various generations, and several generations of hardware will remain active in the CREATE production environment. The new OpenStack cloud has been provisioned with Kayobe using Ubuntu, with considerable joint development between King's e-Research team and StackHPC in supporting this distribution, which has been upstreamed to the Open Infrastructure Foundation. The HPC and Cloud services now share the same switching fabric, such that resources can be re-deployed flexibly between the two use cases and bare metal HPC nodes can be provisioned within individual OpenStack projects.

The previous infrastructure already used several generations of RDMA fabric networking. The new phase hardware will include 100G Ethernet (with RoCE RDMA support) and HDR-100 Infiniband supporting both the HPC batch and virtualised computational resources. A new performance-focused Ceph cluster including an NVMe SSD tier will provide storage for both Cloud and HPC services.

Use Cases

The CREATE infrastructure is optimised for flexibility in order to support the broad range of use cases that exist across King's College London's diverse research community.

Simple Self-Service: "Get me a VM" A simplified Infrastructure-as-a-Service is used to provide a number of self-service users with basic VM environments that are self-managed.

Managed Compute Platforms Managed platforms are deployed and operated by the King's e-Research team on a project's behalf.

Instrument Pipelines OpenStack hosted web applications interacting with scientific instrument pipelines to move and curate generated data (e.g. electron microscopy) between RDS and HPC file systems for various stages of processing, analysis and archival.

Trusted Research Environments (TRE) TRE or Data Safe Haven projects are already active within King's College London, and the next-generation OpenStack infrastructure provides an opportunity to host those projects upon infrastructure designed with appropriate security measures for this purpose. Secure Research Computing Platforms are expected to implement tier 2 of the ATI's Data Safe Havens classifications. The platform is expecting to host pseudonymised datasets, and potentially personally-identifiable data.

Bare Metal HPC An HPC-focused bare metal batch processing cluster will be deployed alongside the virtualised OpenStack infrastructure. The two systems share a single 100GbE switching fabric allowing the boundary between the two to be flexible, enabling the transfer of servers between bare metal and virtualised roles in an automated process. A common CephFS filesystem will be available from both deployment types.

Multi-Site Deployment

A key aspect of the deployment is that the new hardware procurement will be largely deployed at a second data centre, approximately 20km distant from the first site, and will be connected with multiple dedicated fibre links.

We are immensely proud of the collaboration with King's. The initial joint design of the infrastructure was a particularly fruitful and enjoyable process as the skills and experience brought together an optimum solution which led to a very efficient and cost-effective procurement. From the design phase we worked very closely with King's e-Research team to support Ubuntu and are now extremely pleased to see CREATE moving forward to a fully flexible Research Computing cloud service supporting the outstanding research output of the university.