

Unchaining JupyterHub: Running notebooks on resources without inbound connectivity

Tuesday, 16 March 2021 09:35 (25 minutes)

JupyterLab has become an increasingly popular platform for rapid prototyping, teaching algorithms or sharing small analyses in a self-documenting manner.

However, it is commonly operated using dedicated cloud-like infrastructures (e.g. Kubernetes) which often need to be maintained in addition to existing HTC systems. Furthermore, federation of resources or opportunistic usage are not possible due to a requirement of direct inbound connectivity to the execute nodes.

This talk presents a new, open development in the context of the JupyterHub batchspawner: Extending the existing functionality to leverage the connection broker of the HTCondor batch system, the requirement for inbound connectivity to the execute nodes can be dropped, and only outbound connectivity to the Hub is needed.

Combined with a container runtime leveraging user namespaces, unprivileged CVMFS and the HTCondor file transfer mechanism, notebooks can not only be executed directly on existing local HTC systems, but also on opportunistically usable resources such as HPC centres or clouds via an overlay batch system.

The presented prototype paves the way towards a federation of heterogeneous and distributed resources behind a single point of entry.

Speaker release

Yes

Desired slot length

20

Primary authors: Dr FREYERMUTH, Oliver (University of Bonn (DE)); WIENEMANN, Peter (University of Bonn (DE)); KOHL, Katrin (Physikalisches Institut, Universität Bonn)

Presenter: Dr FREYERMUTH, Oliver (University of Bonn (DE))

Session Classification: Computing & Batch Services

Track Classification: Computing & Batch Services