_w5.png _w5.bb _w5.png height

Contribution ID: **550** Type: **Talk**

Secure HPC and AI Inference Workloads

Monday, 1 December 2025 13:50 (20 minutes)

Traditional HPC systems offer various levels of job isolation, including secure RDMA enclaves, but often assume a 'friendly neighbor' shared batch system environment outside of that. Creating end-to-end attested workflows on them is still a novel development. Trusted research environments instead are often built as Kubernetes clusters, on the other hand, and offer more isolated execution environments, but network separation typically ends at the VLAN level.

We present a method to execute workloads in an attested environment using RDMA and IP network separation at the linux namespace level on HPE Slingshot in a setup where K8s is used to run elastic inference workloads."

Presenting Author

Email

Student or Postdoc?

Institute

Registered for the conference?

CHPC User

CHPC Research Programme

Primary author: Dr DYKES, Tim

Presenter: Dr DYKES, Tim

Session Classification: HPC Technology

Track Classification: HPC Technology