



Contribution ID: 105

Type: Poster

# Optimizing Resource Provisioning Across Diverse Computing Facilities with Virtual Kubelet Integration

*Wednesday, 13 March 2024 16:15 (30 minutes)*

The integration of geographically diverse computing facilities involves dynamically allocating unused resources, relocating workflows, and addressing challenges in heterogeneous, distributed, and opportunistic compute provisioning. Key hurdles include effective resource management, scheduling, data transfer optimization, latency reduction, and ensuring security and privacy. Our proposed solution, part of the “JLAB Integrating Research Infrastructure Across Facilities (JRIAF)” project, leverages the Kubernetes framework within userspace. It utilizes Virtual Kubelet implementation to overcome high-level permission limitations on worker nodes, connecting Kubernetes with arbitrary APIs. This implementation enables Virtual Kubelet deployment in userspace for executing shell commands, resulting in an elastic and cross-site Kubernetes cluster that offers enhanced flexibility and resource utilization.

## Significance

This solution leverages the widely recognized Kubernetes (K8s) framework. By integrating Virtual Kubelet (VK), it effectively addresses challenges in diverse, geographically distributed computing environments. This approach enhances resource allocation and scheduling, enabling the deployment of an elastic, cross-site Kubernetes cluster. With its utilization of the well-known K8s framework, this advancement is poised to significantly optimize workflows and efficiently utilize resources across varied computing facilities.

## References

<https://indico.jlab.org/event/459/contributions/11501/>

## Experiment context, if any

The conducted experiment involved implementing a streaming data workflow from ESnet to NERSC, with the deployment of Virtual Kubelet (VK) at NERSC. The key observation from this experiment is that users can submit jobs to the local control-plane, and Kubernetes (K8s) efficiently distributes these jobs to remote sites based on available resources. This practical application demonstrates the effectiveness of the proposed solution in optimizing resource utilization and streamlining job distribution across geographically distributed computing facilities.

**Primary author:** TSAI, Jeng-Yuan

**Co-authors:** LARRIEU, Christopher (Thomas Jefferson National Accelerator Facility); LAWRENCE, David; HEYES, Graham (Jefferson Lab); Dr GYURJYAN, Vardan

**Presenter:** TSAI, Jeng-Yuan

**Session Classification:** Poster session with coffee break

**Track Classification:** Track 1: Computing Technology for Physics Research