Los Alamos
NATIONAL LABORATORY

LA-UR-24-28101

Author: Kabir Vats

Mentors: Rusty Davis, Andres Quan

Title: Effective Database Design for Efficient Workflow Orchestration

Abstract:

High Performance Computing (HPC) resources are valuable, and increasing efficiency by dedicating these resources to essential tasks is a vital goal in workflow orchestration. Build and Execution Environment (BEE) is a workflow orchestration and containerization software that uses a Neo4j graph database to track dependencies between tasks in a workflow. BEE currently launches a containerized instance of Neo4j for each workflow, imposing large resource costs on the system it is run on and significantly increasing the setup time for each submitted workflow. This presentation demonstrates a way to combine these graph databases into a singular database that manages all workflows at once and compares the memory usage and runtime for BEE workflows before and after this change.