Project Summary

1. Overview

A shared heterogenous distributed research computing environment has been the hallmark of the University of Wisconsin—Madison campus for almost two decades. What started with a 2003 MRI acquisition award for the Grid Laboratory of Wisconsin (GLOW) is today an effective source of heterogeneous computing capacity that delivers close to 1M core hours daily to researchers across campus and the nation. The proposed **Heterogeneous Accelerator Laboratory (HAL)** adds a new dimension to this environment. It brings much-needed accelerator computing to an ever-growing community of researchers who require accelerator-based capacity for scientific discovery and education.

Building on the experience of GLOW and recognizing the rapidly evolving national cyberinfrastructure (CI) and commercial cloud accelerator offerings, HAL is designed to meet the evolving needs of an ever-growing portfolio of science drivers. Ranging from neutrino physics to digital agriculture and from machine learning to molecular dynamics, researchers will interact with the powerful and diverse accelerators in HAL through a uniform and familiar interface.

HAL will also continue the operational model established through the GLOW award. Collaborators share resources and consume opportunistically from the whole pool. This pool is supported by a team of experienced research CI specialists, facilitators, and software engineers spread across teams on campus, including the American Family Insurance Data Science Institute, Center for High Throughput Computing, Division of Information Technology, Morgridge Institute of Research, and Wisconsin IceCube Particle Astrophysics Center.

2. Intellectual Merit

Scientific discovery is driven by the capacity to analyze and understand ever-increasing amounts of data. The closer compute resources are to the data and the easier they are to access, the greater the speed of scientific innovation. Bringing the novel computing resources of HAL to University of Wisconsin—Madison will enable myriad science drivers to accelerate their scientific discovery and explore novel ways to perform data-intensive research, such as continuously training complex neural networks on Graphical Processing Units, while concurrently running the inference on real-time data streams on Field Programmable Gate Arrays. The team with nearly two decades of experience in designing, operating, and managing a similar set of resources will allow the team to effectively translate the raw capacity of the proposed heterogenous collection of accelerators into a platform that advances scientific discovery and education.

3. Broader Impacts

Because, data-driven research and education have become tightly linked, students need to learn how to make data-informed decisions to be productive in the workforce. Placing HAL at students' fingertips will provide a novel educational experience. Offering HAL's capabilities to diverse science drivers will empower researchers to develop and offer public outreach materials, conduct research that advances drug discovery, enhance national security, improve understanding of catastrophic weather phenomena, and support NASA missions to the Moon. HAL will be integrated into the regional and national CI ecosystem. Regionally, HAL will coordinate with the UW system to enable researchers and students at UW System campuses to harness the diverse accelerator capacity of HAL as a research and teaching resource. Nationally, HAL will be integrated into the Open Science Grid computing and data federation; making HAL one of the largest sources of accelerator capacity offered on the Open Science Pool.