Los Alamos
NATIONAL LABORATORY

Effective Power Management for the OpenCHAMI Cluster Management Stack

Lucas Ritzdorf — Mentors: David Allen, Nicholas Jones

LANL High Performance Computing Mini Showcase
August 7, 2025

LA-UR-25-28205

Background

- High performance computing (HPC) clusters have traditionally relied on **proprietary provisioning and management infrastructure**. This can be problematic, especially with regard to ongoing security and maintenance for vendored systems.
- The Los Alamos National Laboratory (LANL) leads development of the Open Composable Heterogeneous Application Management Infrastructure (OpenCHAMI) stack, which provides **size- and platform-independent cluster management tools**.
- OpenCHAMI includes a Redfish baseboard management controller (BMC) discovery tool, `magellan`, which would benefit from the addition of **power management functionality**

Goals

The power-management system should:

- Provide easy access to **query and update** compute node power states, supporting all Redfish power options exposed by the BMC
- **Integrate** with the OpenCHAMI SMD inventory service, or with third-party services implementing the same standard interfaces
- Use **few resources** to complete actions, and do so in a **reasonable amount of time** (in contrast to existing vendor solutions)

Testing Environments

The power-management system is being developed and tested on two LANL-affiliated clusters:

- The New Mexico Consortium’s Redondo cluster (one head node, five compute nodes)
- LANL’s Razorback cluster, which serves as a testing platform for the flagship Crossroads production cluster (management plane, 60 compute nodes)

Redondo already utilizes the OpenCHAMI management stack, while Razorback relies on vendor tooling. These mixed environments are useful for testing the new power-management functionality alongside other ecosystem tools, and in isolation.

Managing Power

The new power-management functionality includes the following operations:

- **Query state:** `magellan power <node>`
  - Retrieves compute nodes' current power states by communicating with their BMC(s)
- **Query reset types:** `magellan power <node> -l`
  - Retrieves the list of supported Redfish reset types, e.g. `[On ForceOff GracefulShutdown ForceRestart Nmi PushPowerButton GracefulRestart]`
- **Apply reset:** `magellan power <node> -r <type>`
  - Commands the appropriate BMC(s) to reset the target compute nodes
  - For example, the `GracefulShutdown` reset type sends an ACPI signal instructing the host OS to power off

Challenges and Ongoing Work

- Internal configuration consistency, via both command-line flags and external files
- Subscription-based, rather than polling-based, power state monitoring via the Redfish eventing framework
  - OEM Redfish implementations make this challenging, as API levels and support vary dramatically
- Project-wide `magellan` daemonization support, to better operate in long-running environments like containers
  - Avoids spinning up a new container for each user interaction, as some existing vendor tools do
- Authentication with, and relaying of power-state information to, the centralized OpenCHAMI SMD inventory service

Questions?

Los Alamos NATIONAL LABORATORY