Investigating Replacing IPMI with Redfish in HPC Bare-Metal Provisioning

Ghazanfar Ali (PhD Student, CS, TTU)
ghazanfar.ali@ttu.edu

Advisors:

Mr. Jon Hass, SW Architect, Dell Inc.
Dr. Alan Sill, Managing Director, HPCC, TTU
Dr. Yong Chen, Associate Professor, CS Dept, TTU

Cloud Autonomic Computing (CAC) NSF, TTU
Project Overview

**a) HPC infrastructure management using IPMI**

- **Master Node**
  - Ohpc, Warewulf, Nagios, Slurm, lustre client

- **Compute Node1**
  - Gig Ethernet Switch
  - IPMI

- **Compute Node2**
  - IPMI

- **Compute Node n**
  - IPMI

- **Gig Ethernet Switch**

- **Infrastructure information**

  - {master node name}
  - {master node ip}
  - {internal netmask}
  - {eth provision}
  - {ntp server}
  - {bmc_username}
  - {bmc_password}
  - {compute_ip[0]}, ...
  - {compute_ip[1]}, ...
  - {compute_mac[0]}, ...
  - {compute_mac[1]}, ...
  - {compute_bmc[0]}, ...
  - {compute_bmc[1]}, ...

**b) HPC infrastructure management using Redfish**

- **Master Node**
  - Ohpc, Warewulf, Nagios, Slurm, lustre client

- **Compute Node1**
  - Gig Ethernet Switch
  - Redfish

- **Compute Node2**
  - Redfish

- **Compute Node n**
  - Redfish

<table>
<thead>
<tr>
<th><strong>Redfish</strong></th>
<th><strong>IPMI</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Out-of-band communication supported</td>
<td>Out-of-band communication supported</td>
</tr>
<tr>
<td>Deterministic control (i.e. HTTPS/TCP)</td>
<td>Non-deterministic control (i.e. RCMP/UDP)</td>
</tr>
<tr>
<td>Scalable (web-scale architectures, multi-chassis systems)</td>
<td>Non-scalable</td>
</tr>
<tr>
<td>Multi-domain communication support (i.e. compute, storage, power, cooling domains)</td>
<td>Communication limited to IT equipment (e.g. Compute)</td>
</tr>
<tr>
<td>Human-readable (JSON); RESTful Architecture; Reliable (HTTP based req/resp) &amp; Secure (HTTPS, session)</td>
<td>Bit-wise protocol; No RESTful support; Unreliable and no credible security mechanisms supported</td>
</tr>
</tbody>
</table>
Overview: Combining Warewulf and Redfish

- Warewulf is the bare-metal provisioning module of OpenHPC.

- Warewulf can leverage Redfish provisioning capabilities (e.g. firmware update, configuring BIOS settings).
Role of IPMI in OpenHPC-based Provisioning

1.2 Requirements/Assumptions
For power management, we assume that the compute node baseboard management controllers (BMCs) are available via IPMI from the chosen master host.

1.3 Inputs
- \${bmc_username} # BMC username for use by IPMI
- \${bmc_password} # BMC password for use by IPMI

3.3 Add provisioning services on master node
Many server BIOS configurations have PXE network booting configured as the primary option in the boot order by default. If your compute nodes have a different device as the first in the sequence, the ipmitool utility can be used to enable PXE.

3.8.4.11 Add ConMan
ConMan is a serial console management program designed to support console device output and connecting to compute node consoles via IPMI

3.10 Boot compute nodes
master server should be able to boot the newly defined compute nodes. Assuming that the compute node BIOS settings are configured to boot over PXE
Replacing IPMI with Redfish: BMC Credentials

Input BMC Credentials for Redfish API

```
${bmc_username}  # BMC username for use by IPMI
${bmc_password}  # BMC password for use by IPMI
```

```
bmc_username="${bmc_username:-Redfish username}"
bmc_password="${bmc_password:-Redfish password}"```

Input.local recipe from OpenHPC installation template @ /opt/ohpc/pub/doc/recipes/vanilla/input.local
http://openhpc.community/downloads/
Changing Boot Sequence on Compute Nodes:
If nodes BIOS configurations do not have PXE network booting configured as the primary option, the following commands enable PXE using ipmitool and Redfish API, respectively.

```
ipmitool -E -I lanplus -H ${bmc_ipaddr} -U root chassis bootdev pxe options=persistent
```

```
$ curl -d '{
    "Boot": {
        "BootSourceOverrideEnabled": "Once",
        "BootSourceOverrideTarget": "Pxe"
    }
}'
-H "Content-Type: application/json"
-X PATCH
http://enclosure-A/redfish/v1/Systems/blade-0
```
Replacing IPMI with Redfish: Boot Nodes

Boot compute nodes:

- Master node boots newly defined compute nodes with the assumption that the compute node BIOS settings are configured to boot over PXE (slide 6)
- The corresponding commands in `ipmitool` and Redfish API are as follows:

**IPMI**

```bash
for ((i=0; i<$num COMPUTES; i++)); do
    ipmitool -E -I lanplus -H ${c_bmc[$i]} -U ${bmc_username} chassis power reset
done
```

**Redfish**

```bash
curl -k -H "X-Auth-Token: <authtoken>" -X POST https://<ip:port>/redfish/v1/Systems/<system-id>/Actions/Reset -d '{"ResetType":"On"}'}
```
- Redfish API integration with OpenHPC doesn’t require modification within OpenHPC provisioning component (i.e. Warewulf, xCAT)

- Redfish API calls can replace corresponding IPMI calls in OpenHPC provisioning template scripts

- OpenHPC can leverage Redfish node discovery capability for better automation provisioning
Questions?