



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

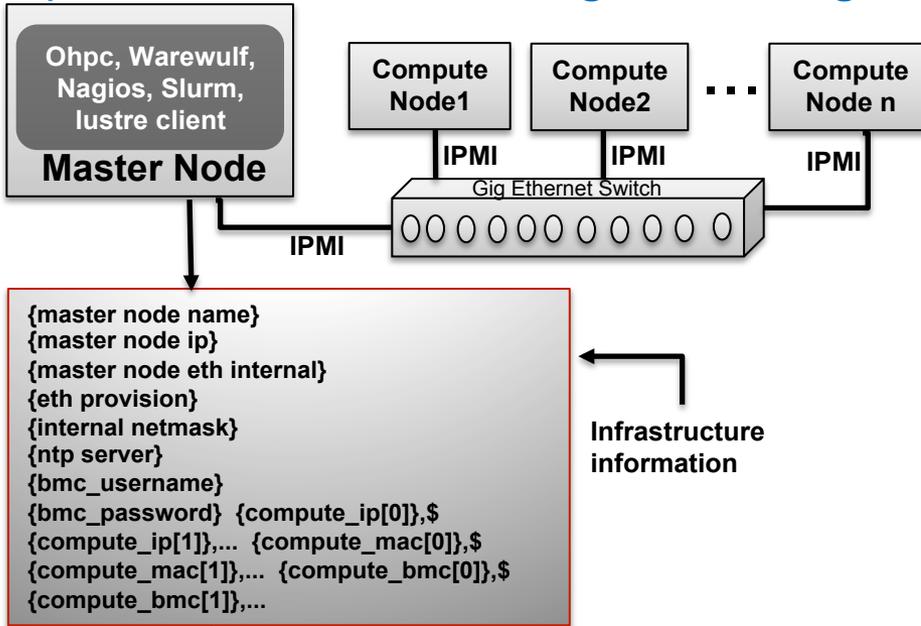
Cloud Autonomic Computing (CAC) NSF, TTU



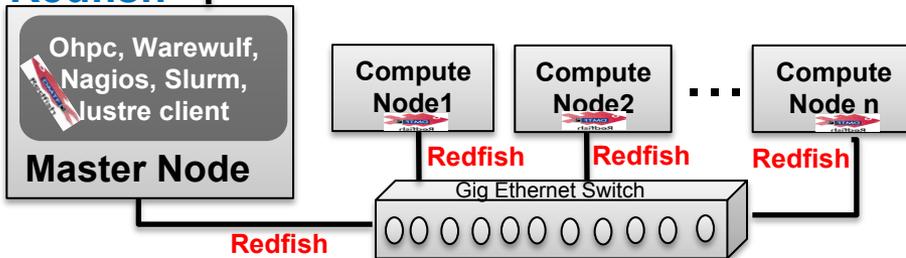
Project Overview



a) HPC infrastructure management using IPMI



b) HPC infrastructure management using Redfish

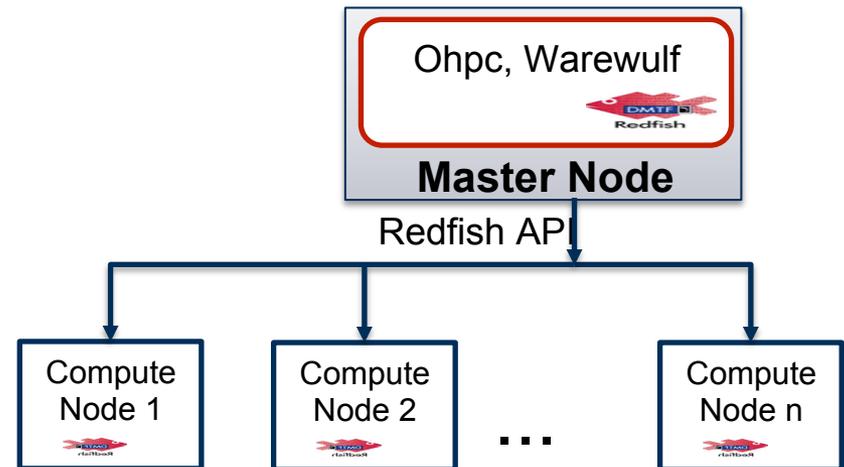


Redfish	IPMI
Out-of-band communication supported	Out-of-band communication supported
Deterministic control (i.e. HTTPS/TCP)	Non-deterministic control (i.e. RCMP/UDP)
Scalable (web-scale architectures, multi-chassis systems)	Non-scalable
Multi-domain communication support (i.e. compute, storage, power,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

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

Warewulf Modules →

RPM Package Name	Version	Info/URL
warewulf-common-ohpc	3.8pre	A suite of tools for clustering. http://warewulf.lbl.gov
warewulf-ipmi-ohpc	3.8pre	IPMI Module for Warewulf. http://warewulf.lbl.gov
warewulf-cluster-node-ohpc	3.8pre	Tools used for clustering with Warewulf. http://warewulf.lbl.gov
warewulf-cluster-ohpc	3.8pre	Tools used for clustering with Warewulf. http://warewulf.lbl.gov
warewulf-vnfs-ohpc	3.8pre	Warewulf VNFS Module. http://warewulf.lbl.gov
warewulf-provision-ohpc	3.8pre	Warewulf - Provisioning Module. http://warewulf.lbl.gov
warewulf-provision-server-ohpc	3.8pre	Warewulf - Provisioning Module - Server. http://warewulf.lbl.gov

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.

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

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

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

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/>



Replacing IPMI with Redfish: Enabling PXE

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

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

Redfish

```
curl -k -H "X-Auth-Token: <authtoken>" -X POST https://<ip:port>/redfish/v1/Systems/
<system-id>/Actions/Reset -d '{"ResetType":"On"}'
```

Feedback



- **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?

