Setting up an HPC lab from scratch

with Mr-Provisioner, Jenkins, OpenHPC and Ansible

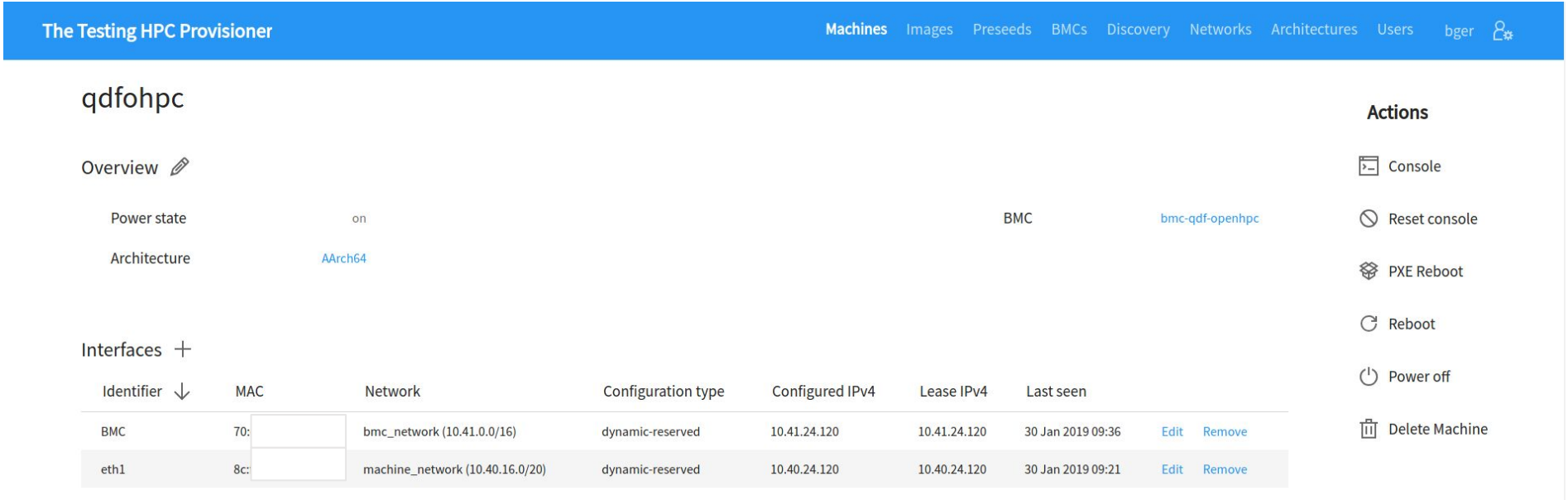
FOSDEM 2019 - Feb 3rd - HPC devroom

Baptiste Gerondeau
Renato Golin

Setting Up

- Lab bring up
  - KVM (libvirt) bringup: provisioner, jenkins, file server
  - Network setup: static IPs, split BMC / provisioning networks
  - Ansible used to install Jenkins, provisioner, SFTP, apt-cache
  - Private Git repository
- Mr-Provisioner: https://github.com/mr-provisioner/mr-provisioner
  - Bare metal provisioning: BMC control, PXE, boot, console
  - DHCP/TFTP service: dynamic images per MAC address (via Grub scripts)
  - Machine discovery, user management, preseed/image upload, JSON automation
- Jenkins: https://github.com/Linaro/hpc_lab_setup
  - Standard install with LDAP + Jenkins Job Builder
  - Install / update jobs via Ansible
  - Jobs checkout same repo for scripts & dynamic execution
- File Server
  - SFTP for benchmark / jobs results (user control), apt-cache, temp web for toolchains

Provisioning



Provisioning

Linaro

Provisioning

Subarchitecture	GrubWithOptionEfiboot
Kernel	bger/vmlinuz1.d058e26c (CentOS 7.5)
Kernel cmdline	ip=dhcp text inst.stage2=http://10.40.0.13/mirror.centos.org/altarch/7/os/aarch64/ inst.repo=http://10.40.0.13/mirror.centos.org/altarch/7/os/aarch64/ inst.ks=file:/ks.cfg earlycon console=ttyAMA0,115200
Initrd	bger/initrd1.img.f7025fca (CentOS 7.5)
Preseed	CentOS.Upstream (CentOS Upsteam)
Netboot enabled	

Assignees

User	Reason	Assignment date	
bger		19 Apr 2018 20:11	Edit Remove
jenkins		20 Jun 2018 11:26	Edit Remove
rengolin		19 Apr 2018 20:11	Edit Remove

Events

Show all

▲ Console accessed by **bger**	29 Jan 2019 16:38:56
● **DHCPREQUEST** packet seen	29 Jan 2019 16:35:10
● **DHCPREQUEST** packet seen	29 Jan 2019 16:20:36

Provisioning

Linaro

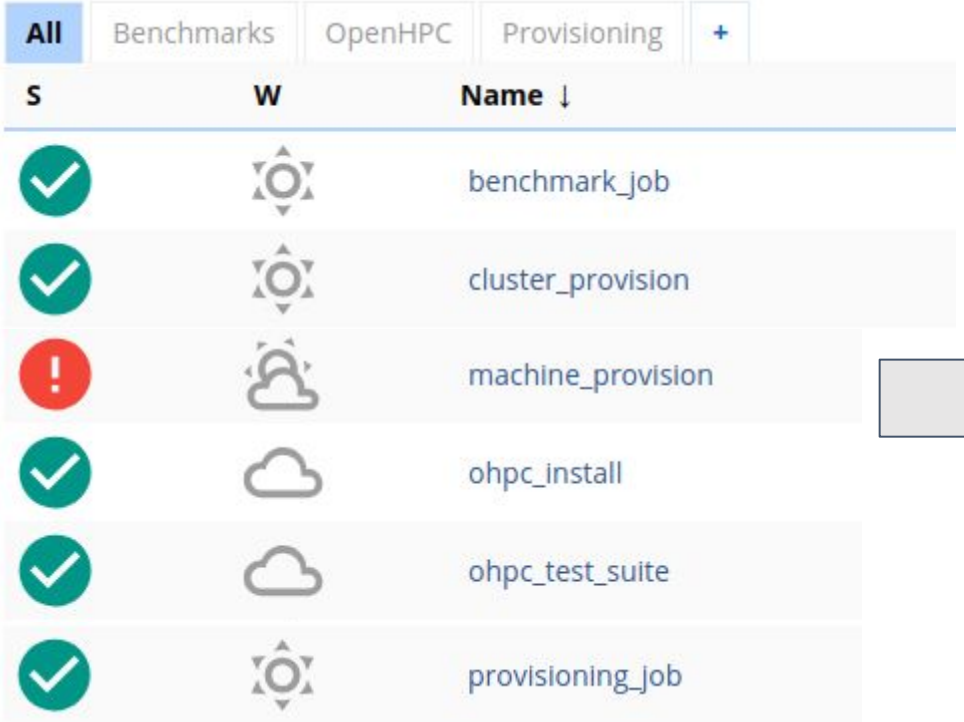
Console for qdfohpc

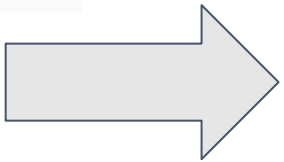
```
[  OK  ] Stopped target Initrd Root File System.
[  OK  ] Listening on udev Control Socket.
         Mounting Debug File System...
         Starting Collect Read-Ahead Data...
         Mounting Huge Pages File System...
[  OK  ] Stopped target Initrd File Systems.
[  OK  ] Reached target Slices.
[  OK  ] Listening on /dev/initctl Compatibility Named Pipe.
[  OK  ] Created slice system-serial\x2dgetty.slice.
[  OK  ] Created slice system-selinux\x2dpol...grate\x2dlocal\x2dchanges.slice.
[  OK  ] Started Collect Read-Ahead Data.
         Starting Remount Root and Kernel File Systems...
         Starting Apply Kernel Variables...
[  OK  ] Started Apply Kernel Variables.
[  OK  ] Started Create list of required sta...ce nodes for the current kernel.
         Starting Create Static Device Nodes in /dev...
[  OK  ] Started Journal Service.
[  OK  ] Mounted POSIX Message Queue File System.
[  OK  ] Mounted Huge Pages File System.
[  OK  ] Mounted Debug File System.
[   12.112922] EXT4-fs (sda3): re-mounted. Opts: (null)
[  OK  ] Started Read and set NIS domainname from /etc/sysconfig/network.
[  OK  ] Started Remount Root and Kernel File Systems.
         Starting Rebuild Hardware Database...
```

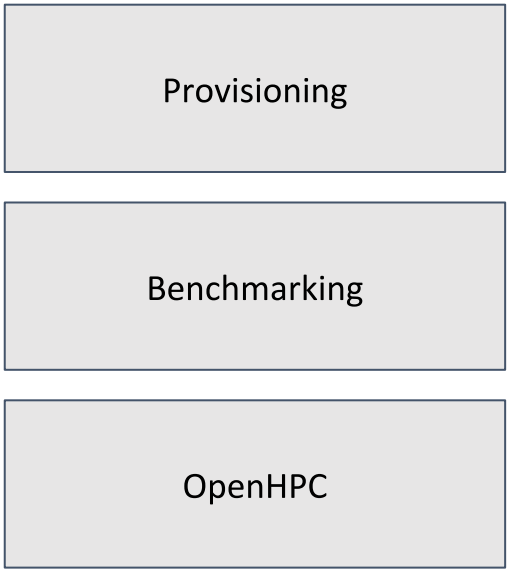
Actions

Download log

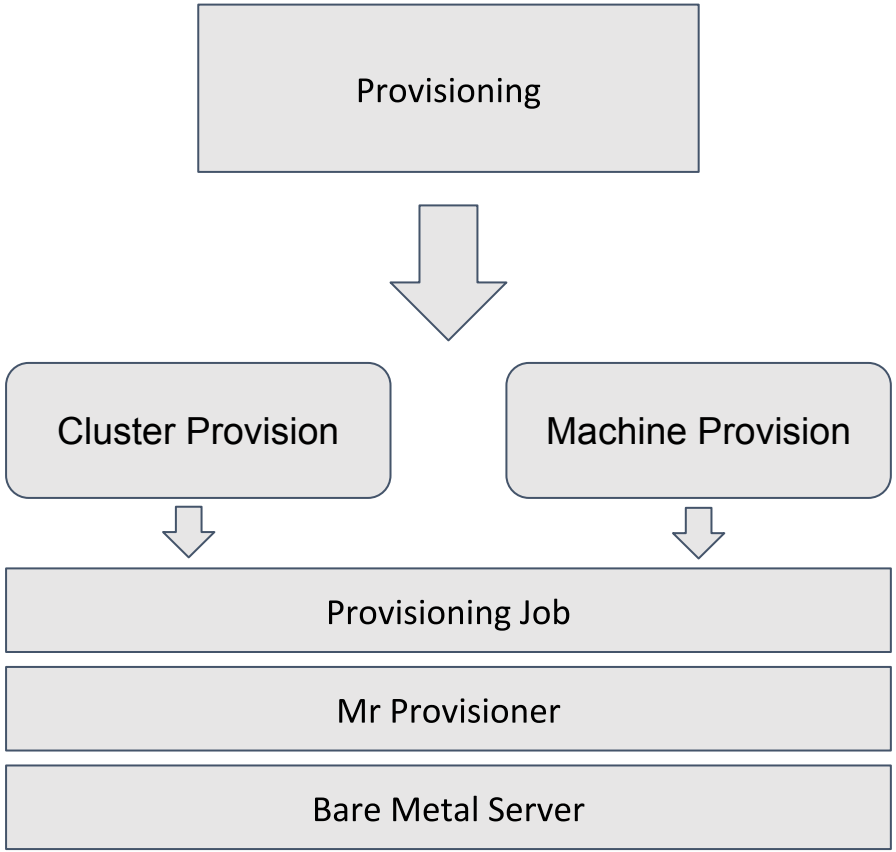
Jenkins Jobs







Jenkins Jobs



Project cluster_provision

This build requires parameters:

machine_type: d05
The cluster of machines to be provisioned

client_branch: master
The branch of the mrp client to use

automation_branch: master
The Ansible logic branch to use

Build

Project machine_provision

This build requires parameters:

machine_type: d03
Micro-architecture of the machine to be provisioned

job_type: bench
The job type to be executed on the machine

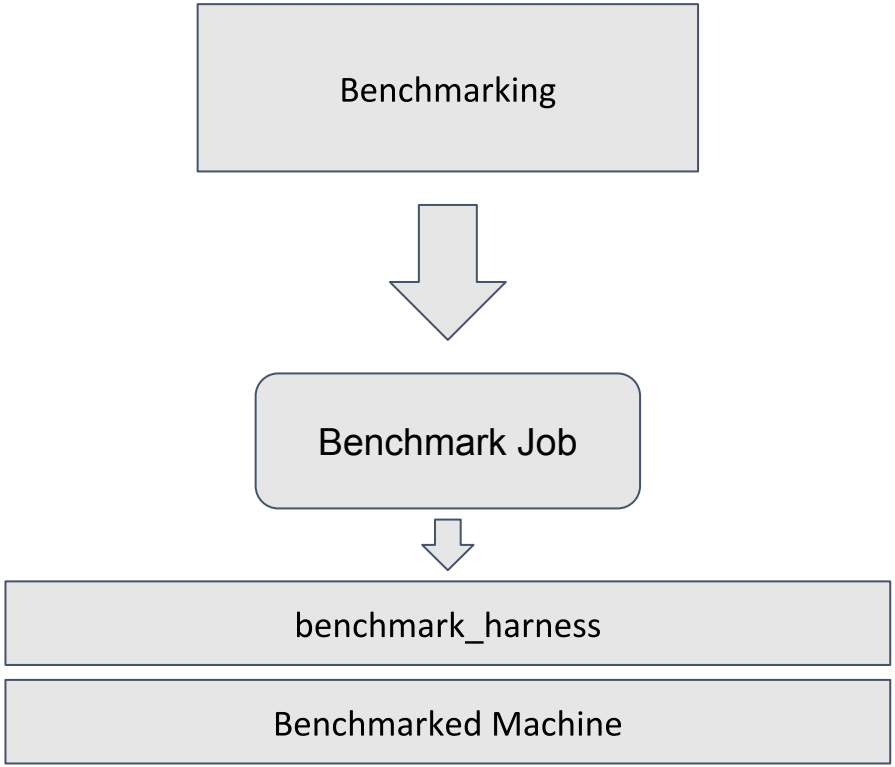
os_type: default
The OS distribution. Default: Debian for benchmark, CentOS for openhpc

client_branch: master
The branch of the MrP client to use

automation_branch: master
The Ansible logic branch to use

Build

Jenkins Jobs



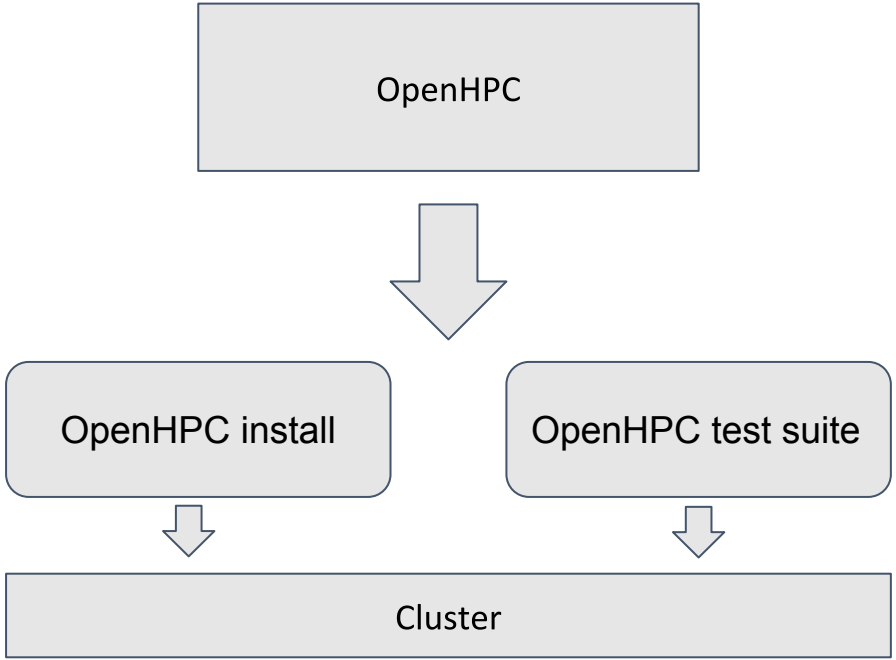
Project benchmark_job

This build requires parameters:

node	d03bench d05bench qdfbench tx2bench Node with which to run the benchmark on the machine
branch	master The benchmark_harness branch to use
automation_branch	master The Ansible logic branch to use
benchmark	lulesh The benchmark to run on the machine
size	3 The benchmark specific size parameter. 1 is small, 3 is large.
iterations	1 The number of times to run the binaries (statistical purposes)
compiler	gcc The compiler (url or system binary) with which to compile the benchmark
compiler_flags	The extra compiler flags to compiler with
linker_flags	The extra linker flags to link with
run_flags	The extra execution flags to run the benchmark
harness_options	Additional harnesss options

Build

Jenkins Jobs



Project ohpc_install

This build requires parameters:

node	d05ohpc qdfohpc tx2ohpc Node on which to execute this job
method	stateful The type of OHPC install to do
client_branch	master Branch name of the mr-provisioner-client to use
automation_branch	linaro Branch name of the ohpc install recipe

Build

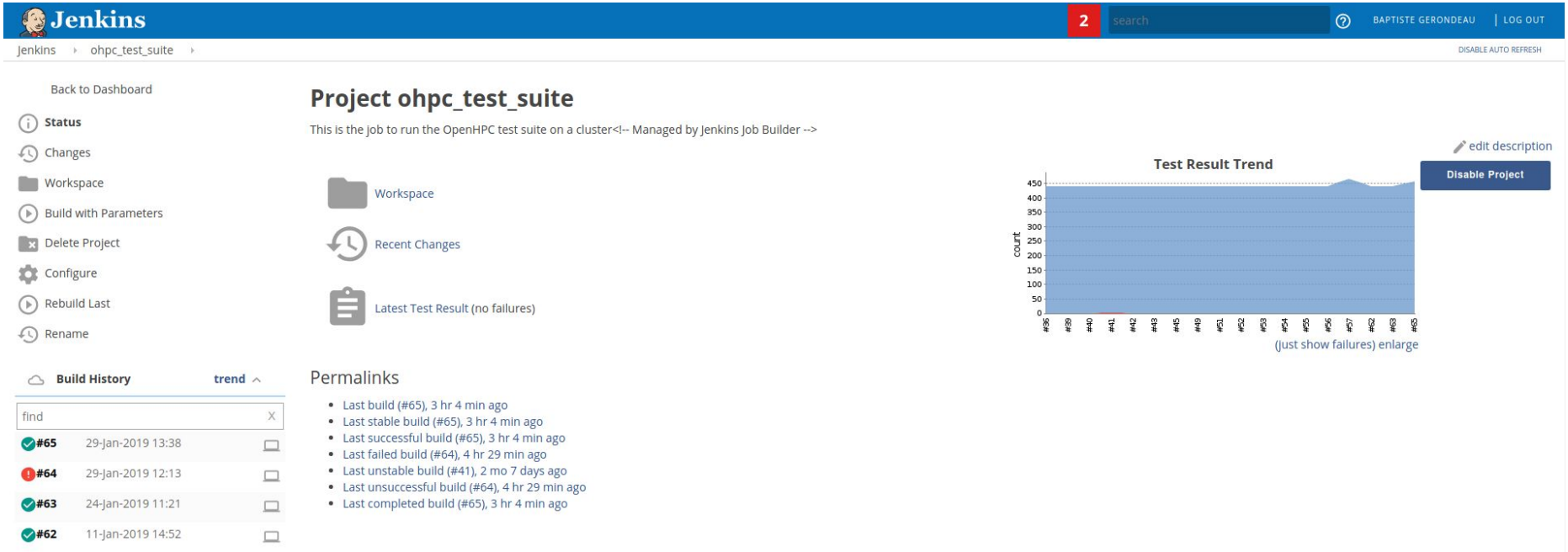
Project ohpc_test_suite

This build requires parameters:

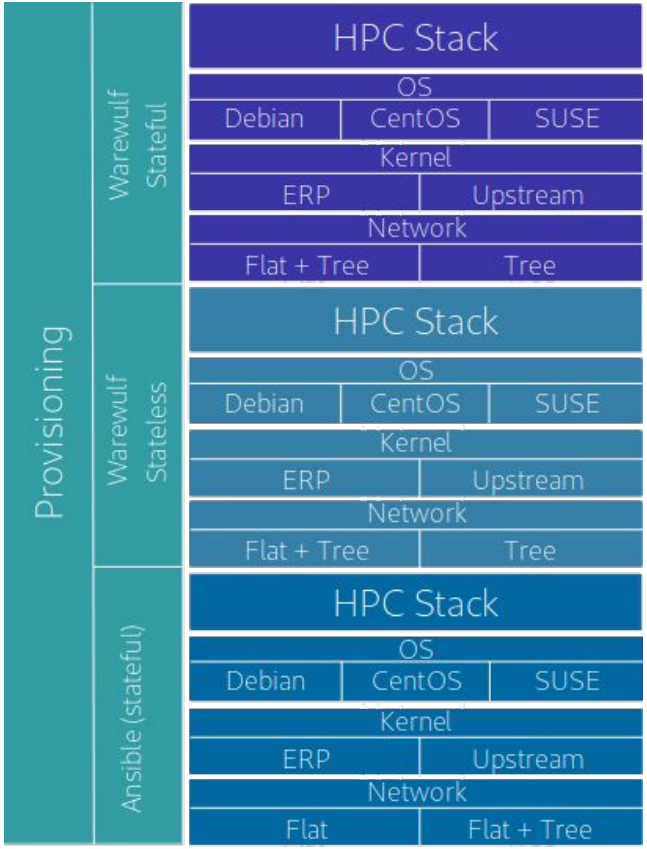
node	d05ohpc qdfohpc tx2ohpc Node on which to execute this job
method	stateful The type of OHPC install to do
test_type	long The length of the tests
client_branch	master Branch name of the ohpc install recipe
automation_branch	linaro Branch name of the ohpc install recipe

Build

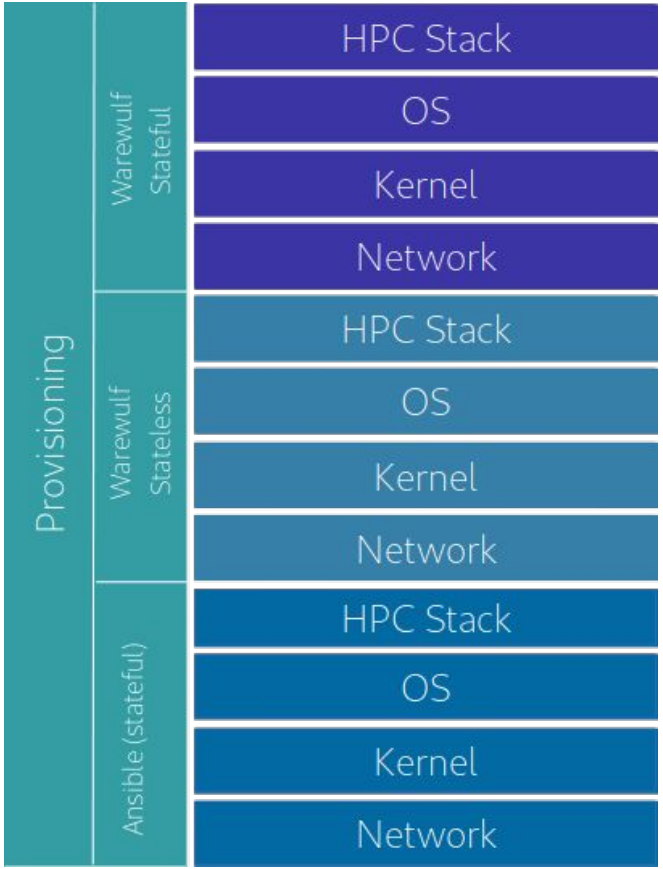
Jenkins Jobs



OpenHPC


Provisioning
Warewulf Stateful
HPC Stack
OS
Debian
CentOS
SUSE
Kernel
ERP
Upstream
Network
Flat + Tree
Tree
Warewulf Stateless
HPC Stack
OS
Debian
CentOS
SUSE
Kernel
ERP
Upstream
Network
Flat + Tree
Tree
Ansible (stateful)
HPC Stack
OS
Debian
CentOS
SUSE
Kernel
ERP
Upstream
Network
Flat
Flat + Tree

OpenHPC


Provisioning
Warewulf Stateful
Warewulf Stateless
Ansible (stateful)
HPC Stack
OS
Kernel
Network
HPC Stack
OS
Kernel
Network
HPC Stack
OS
Kernel
Network

Conclusion

- Modularity
  - Iterative development leading to modular design
- Multiple configurations for each machine/cluster
  - Support for OpenHPC installation on OpenSUSE coming very soon
- Low Maintenance
  - Automation, versioning for the SDI and Jenkins Jobs
- Jenkins Job Builder, Ansible and Python
  - Well known, well used, accessible technologies
- Multiple provisioning options in OpenHPC
  - Ansible stateful only a “bonus” to accomodate for (any) network structure
  - Ansible easily runnable in degraded mode (ignoring certain machines in the cluster)
  - Warewulf Stateless coming very soon
  - Warewulf Stateful coming

Thanks!