CBeST Overview

PSSC Labs continues to develop the most full featured cluster management software in the industry. CBeST is a proprietary cluster management package comprised of kernel optimizations, custom scripts and open source tools that are elegantly integrated, optimized and customized for your specific PowerWulf HPC Cluster hardware and applications. PSSC Labs provides support for CBeST at NO additional cost. Below are the components included with CBeST.

Operating System

The Operating System (OS) is the most central CBeST component. Selecting the correct OS depends on many factors including application compatibility, support level and organizational standards. Once an OS has been selected, PSSC Labs Cluster Technicians work closely with you to properly configure the operating system as the foundation for your PowerWulf HPC Cluster. CBEST supports several Operating Systems including: CentOS Linux and Red Hat Enterprise Linux.

- CentOS (Community ENTerprise Operating System) is a Linux distribution derived from sources freely provided to the public by Red Hat. CentOS strives to be 100% binary compatible with Red Hat. One of the great advantages of CentOS is its stability in releases. CentOS version releases follow those of commercially available operating systems such as Red Hat Enterprise Linux.

- Red Hat Enterprise Linux is available for HPC Clusters, Workstations, Servers and Storage devices. Red Hat offers a wide variety of commercial support options to satisfy university users, corporate environments and everyone in between.

Message Passing Interface (MPI)

MPI is what turns a bunch of networked servers into a parallel supercomputer. The main goals of MPI are high performance, scalability and portability. PowerWulf Clusters can standard with MPI interface configured and optimized for the specific hardware included. You can select from a variety of different MPI versions however Open MPI is the default package included.

- The Open MPI Project is an open source [Message Passing Interface](#) implementation that is developed and maintained by a consortium of academic, research, and industry partners. Open MPI is therefore able to combine the expertise, technologies, and resources from all across the High Performance Computing community in order to build the best MPI library available. Open MPI offers advantages for system and software vendors, application developers and computer science researchers. [1]

Portable Batch Schedulers

Every PowerWulf Cluster includes a portable batch scheduler which allocates resources and schedules jobs according to established priority levels. Having this portable batch scheduler is critical for any size cluster that may be accessed by a small to large group of users. PSSC Labs supports several different batch schedulers. The standard installed batch scheduler is Torque.

PSSC Labs is also a proud partner of Cluster Resources and can include their MOAB management middleware as part of the PowerWulf Cluster

- *TORQUE (Tera-scale Open-source Resource and QUEue manager)* is a resource manager providing control over batch jobs and distributed compute nodes. It is a community effort based on the original OpenPBS project and has incorporated significant advances in the areas of scalability and fault tolerance.

- *MOAB* is an intelligent management middleware that provides simple Web-based job management, graphical cluster administration and management reporting tools. Organizations will benefit from the ability to provide guaranteed service levels to users and organizations, higher resource utilization rates, and the ability to get more jobs processed with the same resources, resulting in an improved ROI.

Network File Systems
Typically a PowerWulf Cluster will include some amount of data storage that is configured internal to the head node or as a separate network attached storage device. Using a Network File System (NFS) users have access to local file systems to clients over a standard TCP / IP network. PSSC Labs Cluster Engineers can configure the following file systems on the attached PowerWulf Cluster storage.

- Ext4 for small to mid range strage capacities below 10 TBs
- XFS for large storage capacities exceeding 10 TBs

PowerWulf Clusters can also access very large storage arrays exceeding several PetaBytes (PBs). Parallel file systems including Gluster and BeeGFS can be configured as either a component of the PowerWulf Cluster or a separate storage cluster.

Compilers
Included with each PowerWulf Cluster are at least two sets of comilers: the basic GNU devleopment tools (GCC, make, etc) and the compilers that are associated with the installed MPI versions. PSSC Labs Cluster Engineers will ensure that the different versions of MPI are properly compiled prior to shipping. In addition, commercial compilers from Intel are also available. Intel offers a variety of editions depending on the number of users and floating licenses. To see the different Intel compiler options visit: http://software.intel.com/en-us/intel-compilers

Hardware Monitoring Utilities
Immediate status of the node and cluster is vital for any user. PowerWulf Clusters include Ganglia, a widely utilized GUI based remote monitoring tool. Ganglia allows user to quickly and easily monitor critical componentes including: cpu usage, cpu temperatures, chassis and cpu fan speeds, hard drive temperatures, memory utilization, hard drive swap space and much more. Ganglia requires very level resource overhead and scales to thousands of compute nodes.

Node Imaging and Replication Utility

PowerWulf Cluster users may face many issues including maintaining a consistent OS version across nodes, repairing a corrupted file system, replacing a failed hard drive as well as adding new nodes. Your PowerWulf Cluster is equipped with a useful utility to easily add new nodes, update images on existing nodes & even recover a node image in the event of a hard drive failure. This utility is called Clonezilla. Clonezilla is an open source disk cloning utility that allows saving and imaging of hard drives via CD, USB or Ethernet connection. Clonezilla enables you to manually create an entire compute node kernel image on the head node and then push this image to every compute node. Users can update the compute nodes by syncing them to an updated image on the head node. These updates are extremely fast because only the configuration portions that have changed will be pushed to the compute nodes.

Remote Management

It is critial for any user to be able to access the PowerWulf Cluster remotely. All systems include a separate remote management interface via IPMI (Intel Platform Managmeent Interface). This remote management interface allows users to interact with the system as if they were standing right in front of it. Now you have the ability to remotely power cycle a single node, a subset of nodes or the entire cluster with simple commaned. A few examples of these remote management scripts are listed below:

> node-wakeup
> This script uses WOL (Wake-On-Lan) to remotely turn on nodes. It uses the MAC addresses found in /etc/dhcpd.conf. WOL is not a reliable service. It uses Ganglia to build a node list of active nodes. Arguments: all or short node name.
>
> Examples:
> :: node-wakeup all
> :: node-wakeup n12
>
> node-reboot
> This script reboot all the currently running compute nodes. It uses gsh or rsh/ssh for remote command execution. Arguments: all or short node name.
>
> Examples:
> :: node-reboot all
> :: node-reboot n12
>
> node-poweroff
> This script halts (and powers off when supported) all the running compute nodes. It uses gsh or rsh/ssh for remote command execution. Arguments: all or short node name.
>
> Examples:
> :: node-poweroff all
> :: node-poweroff n12

Security

The PowerWulf Cluster is NOT a security-hardened system and must not be connected to a public / unsecure network. The system must be installed on a private network behind a secure firewall. The PowerWulf Cluster Head Node is configured with a basic iptables firewall that allows incoming traffic only on port 22 on its Internet-facing interface. Failure to install the system on a secure network will increase the chance of unwanted intrusion and may ultimately lead to compromise of both the system software and data. You must change the user password & firewall settings prior to connecting PowerWulf Cluster to external network.

CBeST User Manual

The CBeST User Manual has been called "exactly what I want to see" by all types of PowerWulf Cluster users from basic users to the most experienced system administrators. Detailed information on CBeST tools, installation, customization, operation and troubleshooting are included. The CBeST user manual is your ultimate resource. From the minute your PowerWulf Cluster arrives at your door, the CBeST User Manual will guide you through the initial installation all the way to troubleshooting MPI and PBS jobs. Developed by the same PSSC Labs HPC Cluster Technicians that installed the software on your PowerWulf Cluster, the user manual is designed to assist and educate any and all levels of users. If you are a research scientist, focusing on your own work and not administering a supercomputer, CBeST User Manual is the perfect first resource guide. If you are a full time system administrator, this manual will walk you through all aspects of the CBeST installation and give you plenty of options to make your own changes.

References:
1 - https://www.open-mpi.org//