ACCELERATED HPC PRODUCTIVITY WITH INTEL CLUSTER READY SOLUTIONS

By Brock A. Taylor
Vishvesh Sahasrabudhe
Onur Celebioglu

The Intel® Cluster Ready program provides a standardized, replicable way to build and run high-performance computing (HPC) clusters, helping simplify cluster deployment and management. By using Intel Cluster Ready–certified Dell™ HPC clusters, organizations can quickly install and configure clusters to begin running registered HPC applications.

Related Categories:

Dell PowerEdge blade servers

High-performance computing (HPC)

Intel

Linux

Visit DELL.COM/PowerSolutions for the complete category index.

High-performance computing (HPC) clusters can be extremely effective for solving large, complex problems in areas such as computational fluid dynamics, weather modeling, bioinformatics, and oil and gas exploration. However, they can also be extremely complex and time-consuming to deploy, manage, and maintain. To help reduce this complexity and the time required to manage cluster resources, Intel—with the help of Dell and other partners—launched the Intel Cluster Ready program. This program is designed to help organizations create compliant cluster software environments to run registered applications, and provide the tools that can help maintain systems at their expected levels of operations. It also aims to speed adoption of clusters by organizations otherwise limited to single workstations or symmetric multiprocessing systems because of the complexities of deploying and operating HPC clusters.[1]

To help demonstrate how this program can help get verified clusters quickly up and running registered applications, a team of Dell and Intel engineers undertook a project to simulate an organization deploying and using an Intel Cluster Ready–certified Dell HPC cluster. The project team deployed a 16-node cluster based on Dell PowerEdge™ M600 blade servers, then ran eight registered applications to showcase the efficiency and effectiveness of Intel Cluster Ready solutions. Using similar clusters in their own environments enables organizations to focus on determining which registered applications to run rather than on handling the complexities of cluster deployment—ultimately helping increase productivity and accelerate the return on their investment in computing resources.

UNDERSTANDING THE INTEL CLUSTER READY PROGRAM

The Intel Cluster Ready program is based on a specification that defines a common interface for HPC applications and incorporates industry-wide practices for building and managing HPC clusters. The defined application-layer interface enables application portability between certified clusters. Compliant clusters are designed to run any registered application, providing a one-to-many cluster-to-application environment. Requiring industry standards and practices helps certified clusters provide the tools and functionality necessary to deploy clusters efficiently and provide the administrative functionality that fosters simplified management. Each certified cluster must be built and verified to be compliant before being released, which shifts the intricacies of designing and building compliant clusters to the vendor, allowing organizations to focus on using the cluster for its intended purpose—running applications and solving problems.

[1] For more information on the Intel Cluster Ready program, visit DELL.COM/Content/Topics/Global.aspx/Sitelets/Solutions/Cluster_Grid/En/US/Clustering_HPCC?c=us&cs=555&l=en&s=biz&~section=008.

As of July 2008, the Intel Cluster Ready program defines a specification for Linux® OS–based clusters; Figure 1 shows the Intel Cluster Ready architecture. Although other operating systems may be considered in the future, the July 2008 version is a specific Linux binding. Example software interface requirements for each cluster node include the following:

- A Linux kernel that conforms to the Linux Standard Base
- A Portable Operating System Interface for UNIX® (POSIX) command system
- Intel 64 binaries for the base libraries, GNU C library (glibc), and X11 libraries
- An implementation of Message Passing Interface (MPI) 1.2 or later
- A set of common runtime components, including Intel compilers and libraries
- The OpenFabrics Enterprise Distribution (OFED) software stack as the network interface for network fabrics beyond Ethernet

Example hardware characteristics and capabilities include the following:

- At least 512 MB of memory per core
- At least 10 GB of storage space available to each compute node
- Identical compute node hardware (only nominal differences allowed)
- Identical file trees across compute nodes
- Node access to a single shared file space for user home directories
- Cluster management network fabric

The Intel Cluster Ready program uses a certification process for HPC clusters and a registration process for applications to enable its one-to-many ecosystem. A cluster recipe is a process or mechanism to build a solution that is compliant with the specification. A reference system built from a recipe is certified to meet the requirements and can provide the software interface to the application layer. Separately, HPC application vendors register that their codes operate on top of the defined minimum interface provided by a compliant cluster. Applications with components not required by the specification must be identified on the Intel Cluster Ready Web site to be a registered code, or those additional components must be bundled with the application itself. Combining these two processes helps provide a one-to-many ecosystem and avoid potential problems when a cluster is built specifically around the requirements of a single application. Compliant clusters are designed to run a single registered application or multiple registered applications without modifications or rebuilding the cluster software stack.

Intel Cluster Checker—a preconfigured tool included in compliant clusters and designed to check system health and performance—is a key component of the program. This tool systematically runs through test modules, checking basic system functionality such as Secure Shell (SSH) capabilities between nodes as well as high-level performance across the cluster—including exhaustive link checking for expected bandwidth and latency performance between each node pair in the cluster. It is designed to support efficient cluster management, reduce the need for downtime for maintenance, and provide quick isolation of problems. Its output provides definitive pass or fail reporting of cluster health, helping confirm that the cluster is operating as expected and helping reduce maintenance time.

DEMONSTRATING CLUSTER SETUP AND USE

In April 2008, a team of Dell and Intel engineers launched a project to showcase the rapid deployment of an Intel Cluster Ready–certified Dell HPC cluster and the

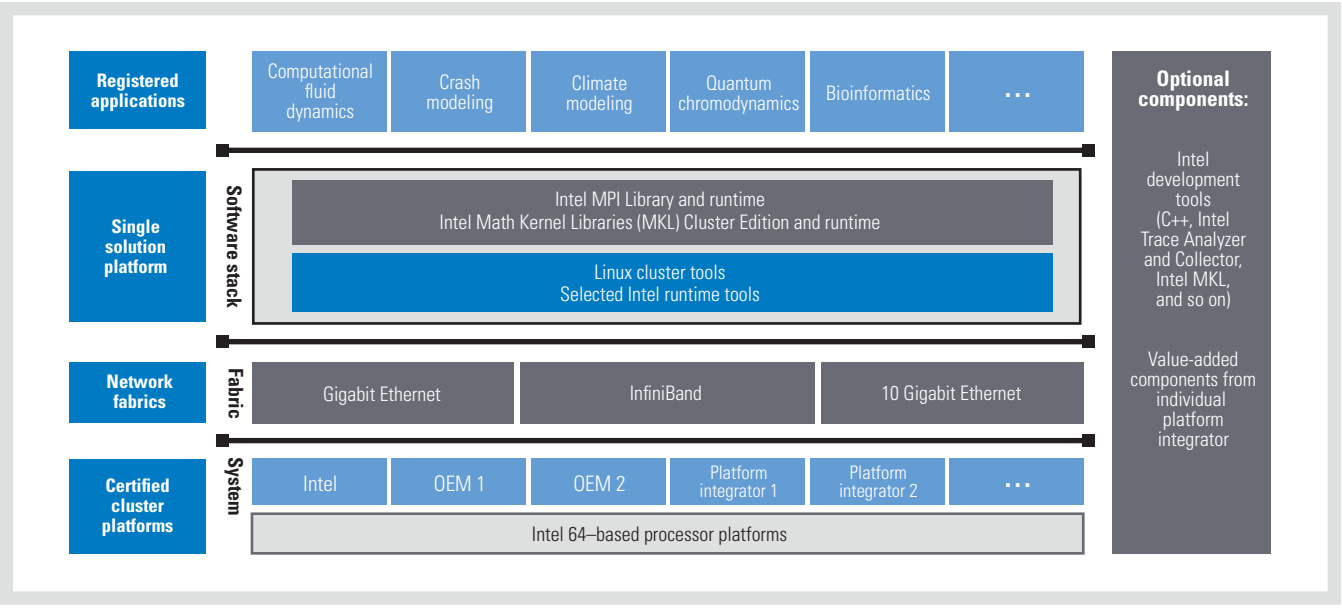

Figure 1. *Intel Cluster Ready architecture*

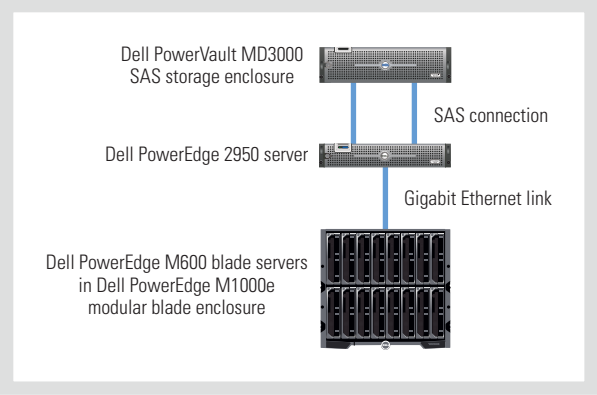

Figure 2. *Cluster configuration in the example environment*

quick, useful results this resource could provide. Dell has invested significant effort and expertise in creating and certifying cluster configurations designed to provide consistently compliant systems. Installation and configuration steps are documented, and automation helps avoid manual steps that may introduce problems in the cluster build. The components are integrated and then verified to help ensure the cluster presents the necessary interface to the applications.

The team deployed a certified cluster of 16 Dell PowerEdge M600 Intel 64–based blade servers in a PowerEdge M1000e modular blade enclosure as compute nodes, with a PowerEdge 2950 server as the master node (see Figure 2). Each PowerEdge M600 contained two quad-core Intel Xeon® E5450 processors at 3.00 GHz and 16 GB of RAM; the PowerEdge 2950 was equipped with two quad-core Intel Xeon E5450 processors at 3.00 GHz and 8 GB of RAM, and was attached to a Dell PowerVault™ MD3000 Serial Attached SCSI (SAS) storage enclosure. The cluster included two network fabrics: Gigabit Ethernet between all nodes and InfiniBand between compute nodes. The InfiniBand fabric included a Cisco M SFS7000E InfiniBand switch module for the enclosure with 16 internal and 8 external 4X, 20 Gbps, double data rate (DDR) InfiniBand ports. Because the InfiniBand fabric was contained in the PowerEdge M1000e enclosure for this cluster, no external InfiniBand cable was needed.

The cluster used Platform Open Cluster Stack 4.5.1 to provision the software stack across the cluster.[2] Each node ran the Red Hat® Enterprise Linux 4 Update 5 OS and required components such as Intel runtime libraries and the Java Runtime Environment (JRE). The software stack also included OFED 1.3 for the InfiniBand fabric. The Intel Cluster Checker tool was installed on the master node along with preconfigured input files specifically designed to check the health of the cluster.

The deployment simulation was carried out in two stages. For the first stage, the project team mimicked the creation of an Intel Cluster Ready–certified Dell HPC cluster by performing actions carried out at the engineering level at Dell. By following the cluster recipe, the team moved the cluster hardware from bare metal to a fully functional system in approximately two hours. The Intel Cluster Checker tool verified that the software stack was a copy of the original certified reference cluster and that the cluster was operating as expected. All nodes were functional, the Gigabit Ethernet and InfiniBand fabrics were operational and performing at the expected levels, and the cluster was ready for use.

After the Intel Cluster Checker tool had verified the cluster, the team began the second stage—simulating an organization using the cluster for real work using registered applications. Eight HPC applications from six independent software vendors (ISVs) were installed on the cluster. One by one, the team launched real workload analysis on ANSYS, FLUENT, STAR-CD, LS-DYNA, Abaqus/Standard, Abaqus/Explicit, NASA Structural Analysis System (NASTRAN), and PAM-CRASH. Each application ran without modifications to the cluster, and each ran to completion. Apart from the effort to bring the cluster to an Intel Cluster Ready–certified stage, no additional time was necessary to customize the cluster configuration to fit the needs of individual applications. This study demonstrates that the Intel Cluster Ready program combined with certified Dell cluster configurations can provide a turnkey HPC platform that end users can take advantage of rapidly and efficiently.

SIMPLIFYING CLUSTER DEPLOYMENT AND MANAGEMENT

The Intel Cluster Ready program is designed to take advantage of the expertise of independent hardware and software vendors to provide a turnkey solution for organizations deploying HPC clusters. Dell and Intel are working closely with ISVs to provide cluster resources that meet organizations' needs without requiring in-depth knowledge of how to make those systems work—helping simplify the installation, configuration, and management of clusters and enabling organizations to focus on their productivity.

Brock A. Taylor is a staff engineer and part of the Software and Services Group at Intel.

Vishvesh Sahasrabudhe is a member of the HPC Engineering Group at Dell.

Onur Celebioglu is an engineering manager in the HPC Engineering Group at Dell.

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Intel Cluster Ready program:
www.intel.com/go/cluster

Dell HPC cluster solutions:
DELL.COM/HPCC

[2] This product includes software developed by the Rocks Cluster Group at the San Diego Supercomputer Center at the University of California, San Diego, and its contributors.