DESIGNING AND DEPLOYING A HIGH-PERFORMANCE COMPUTING (HPC) CLUSTER CAN BE DIFFICULT FOR EVEN THE MOST EXPERIENCED IT STAFF. TO HELP SIMPLIFY CLUSTER ACQUISITION AND MANAGEMENT, DELL, INTEL, AND CLUSTERCORP HAVE COLLABORATED TO OFFER AN INTEGRATED AND TESTED HPC CLUSTER PLATFORM BASED ON 11TH-GENERATION DELL POWEREDGE SERVERS, INTEL CLUSTER READY CONFIGURATIONS, AND CLUSTERCORP ROCKS+ CLUSTER MANAGEMENT SOFTWARE—THE COMMERCIAL VERSION OF THE AWARD-WINNING ROCKS® SOFTWARE.1 THIS COLLABORATION ENABLES DELL HPC CLUSTERS TO TAKE ON THE FEEL OF A TRADITIONAL SUPERCOMPUTER, AND ENABLES THEM TO BE DELIVERED TO THE DATA CENTER READY TO RUN REGISTERED APPLICATIONS RIGHT OUT OF THE BOX.

DEFINING THE DELL HPC CLUSTER

The clusters resulting from the collaboration between Dell, Intel, and Clustercorp are based on new 11th-generation Dell PowerEdge servers, the Intel Cluster Ready program, and Rocks+ cluster management software.

NEW 11TH-GENERATION DELL POWEREDGE SERVERS

New 11th-generation Dell PowerEdge servers feature performance, energy-efficiency, and systems management enhancements that make them well suited for HPC environments. These servers utilize the quad-core Intel Xeon® processor 5500 series, which are frequency and voltage independent to enable granular control over server performance and power consumption. These processors also incorporate a three-level cache hierarchy with a dedicated level 2 (L2) cache to help reduce replacement-driven cache misses, and a shared L3 cache to help minimize communication between sockets. A three-channel Double Data Rate 3 (DDR3) memory controller integrated directly on each processor helps reduce the latency of local memory accesses. The processor sockets are linked through dedicated communication channels based on Intel QuickPath Interconnect technology and accelerate inter-processor communication for remote memory accesses. When combined, these architectural enhancements help eliminate the memory bottleneck associated with legacy frontside bus platforms. These servers are based on the Intel 5500 and Intel 5520 chipsets, which support registered and unbuffered DDR3 dual in-line memory modules (DIMMs) as well as PCI Express (PCIe) 2.0 expansion cards.

New 11th-generation PowerEdge servers also include additional updates and enhancements such as the Integrated Dell Remote Access Controller 6 (iDRAC 6), which is compliant with the Intelligent Platform Management Interface (IPMI) 2.0 specification. The iDRAC 6 includes an integrated video

1 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.
controller and provides systems management features such as console redirection, remote power control, chassis hardware alert generation, and remote media boot support. The iDRAC Express persistent storage device provides the ability to automatically pull system firmware updates from an external repository, configure the BIOS, and run diagnostics, and includes an on-board repository with critical OS drivers. Figure 1 outlines the key specifications of 11th-generation PowerEdge rack and blade servers.

### Intel Cluster Ready program

Developed in conjunction with hardware and software vendors such as Dell and Clustercorp, the Intel Cluster Ready program focuses on providing cluster solutions that can run HPC applications right out of the box and do not require special cluster expertise to deploy and maintain. Certified clusters provide a standard interface to the application level, and HPC software vendors in turn register applications that execute on top of that interface—helping simplify the purchasing, deployment, and management of HPC clusters.

Each certified cluster includes the Intel Cluster Checker tool, which provides a simple, methodical way to verify the cluster’s operational health. Using a modular framework, Intel Cluster Checker proceeds from basic functionality tests to application-level benchmarks, providing comprehensive feedback on whether the cluster is working properly. This tool helps remove the guesswork from maintaining the system, ultimately contributing to reduced systems management costs. As a result of the collaboration between Dell, Intel, and Clustercorp, specific Dell hardware configurations are Intel Cluster Ready certified, and certification tools such as Intel Cluster Checker have been integrated into the Rocks+ deployment framework.

### Rocks+ cluster management software

Rocks+ provides a robust, comprehensive software stack for HPC clusters. It is designed to provide all the necessary software components for building, deploying, and managing a cluster, including the Red Hat® Enterprise Linux® OS, cluster middleware, tightly integrated compilers, applications, and communication libraries. Rocks+ 5 introduces multiple features and enhancements, including the Rocks command-line interface, support for cluster virtualization, updated OS distribution based on Red Hat Enterprise Linux 5, and the Dell Roll, which was designed specifically for 11th-generation Dell PowerEdge servers.

### DEPLOYING THE CLUSTER WITH ROCKS+ROLLS

The Dell deployment team installs a cluster by combining a carefully selected set of best-of-breed components, which might include low-latency interconnect networks from QLogic or Mellanox, workload management software such as Cluster Resources’ Moab Cluster Suite or Platform Load Sharing Facility (LSF), and open source packages such as the OpenFabrics Enterprise Distribution (OFED). These components are managed and delivered by Rocks+ through modular software components called Rolls. After a cluster has been deployed with Rocks+, the Intel Cluster Checker tool confirms that the cluster is Intel Cluster Ready compliant and functioning correctly. Cluster installation and verification can take place at the deployment site or at the factory through

---


2For more information on the Intel Cluster Ready program, visit www.intel.com/go/cluster.

3Requires separate licenses from Red Hat.

---

**Figure 1.** Key specifications of 11th-generation Dell PowerEdge rack and blade servers

<table>
<thead>
<tr>
<th></th>
<th>PowerEdge R610</th>
<th>PowerEdge R710</th>
<th>PowerEdge M610</th>
<th>PowerEdge M710</th>
</tr>
</thead>
<tbody>
<tr>
<td>Form factor</td>
<td>1U rack server</td>
<td>2U rack server</td>
<td>Half-height blade server</td>
<td>Full-height blade server</td>
</tr>
<tr>
<td>Total number of DIMM slots</td>
<td>12</td>
<td>18</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>Maximum memory capacity</td>
<td>96 GB</td>
<td>144 GB</td>
<td>96 GB</td>
<td>144 GB</td>
</tr>
<tr>
<td>Number of PCIe 2.0 slots</td>
<td>Two x8 slots</td>
<td>Two x4 slots and either two x8 slots or one x16 slot</td>
<td>Two x8 mezzanine card slots</td>
<td>One x4 slot and three x8 mezzanine card slots</td>
</tr>
<tr>
<td>Number of embedded network interfaces</td>
<td>Four</td>
<td>Four</td>
<td>Two</td>
<td>Four</td>
</tr>
<tr>
<td>Maximum number of Serial Attached SCSI (SAS) or Serial ATA (SATA) hard drives</td>
<td>Six 2.5-inch drives</td>
<td>Eight 2.5-inch drives or six 3.5-inch drives</td>
<td>Two 2.5-inch drives</td>
<td>Four 2.5-inch drives</td>
</tr>
<tr>
<td>Systems management</td>
<td>iDRAC 6, iDRAC 6 Express, and optional iDRAC 6 Enterprise</td>
<td>iDRAC 6, iDRAC 6 Express, and iDRAC 6 Enterprise</td>
<td>iDRAC 6, iDRAC 6 Express, and iDRAC 6 Enterprise</td>
<td>iDRAC 6, iDRAC 6 Express, and iDRAC 6 Enterprise</td>
</tr>
</tbody>
</table>

DELL.COM/PowerSolutions 47
Dell Server and Storage Rack and Stack Services, in which the clusters are racked, cabled, installed, and verified at Dell. Organizations can then connect remotely to benchmark the cluster before it is delivered to the final site.¹

The engineering and testing effort behind Dell clusters leaves behind the do-it-yourself, on-the-fly approach from the early days of cluster computing and replaces it with end-to-end services throughout the cluster life cycle. Although this approach is reminiscent of traditional supercomputing, Dell clusters remain both cost-effective and scalable.

**Dell Roll**
The Dell Roll uses the Rocks+ framework to help further simplify cluster deployment and management by automatically configuring cluster hardware and software that is specific to 11th-generation Dell PowerEdge servers. The Dell Roll installs Dell-approved device drivers across the cluster, applies necessary OS fixes and kernel parameters, and modifies BIOS settings for functionality and performance. The Dell Roll can also be used to deploy and configure Dell systems management software and the iDRAC IPMI management interfaces.

**Intel Cluster Ready Roll**
The Intel Cluster Ready Roll incorporates the components required by the Intel Cluster Ready specification and automates the setup process for the Intel Cluster Checker tool and many of its phases of cluster setup.

**InfiniBand Rolls**
Rocks+ can also incorporate additional Rolls that differentiate it from the base Rocks cluster package. The Mellanox OFED Roll was designed by Clustercorp as a core component of Rocks+. It supports OFED 1.4, which includes drivers and firmware for Mellanox InfiniBand host channel adapters (HCAs), fabric management tools, and multiple MPI environments. It also includes comprehensive support for Intel compilers when paired with the Intel Developer Roll from Clustercorp, and helps remove the guesswork and manual steps to create a validated InfiniBand environment.

Clustercorp also offers the QLogic Roll for organizations using QLogic hardware. This Roll—designed by QLogic and tested by Clustercorp—adds support for QLogic OFED+ for TrueScale or InfiniScale HCAs. Integration of QLogic drivers helps provide a standardized deployment strategy that allows organizations to quickly move through the deployment and acceptance phases of cluster setup.

**Moab and LSF Rolls**
The Moab Roll adds the Moab Cluster Suite to Rocks+, providing an automated configuration for powerful workload and resource management that integrates scheduling, management, monitoring, and reporting for cluster workloads. Platform Computing also worked closely with Clustercorp to create the LSF Roll, which adds Platform LSF to Rocks+ and combines intelligent workload management with flexible, simplified cluster management.

**SIMPLIFYING HPC CLUSTERS**
Inspired by the pioneers in traditional supercomputing, when HPC systems were expected to be complete solutions rather than piecemeal projects, Dell, Intel, and Clustercorp have worked together to create HPC clusters designed to be simple, comprehensive, and reliable, helping avoid the inherent complexity of managing disparate components. By combining 11th-generation Dell PowerEdge servers, Intel Cluster Ready configurations, and Clustercorp Rocks+, a cluster can be assembled, customized, and verified on the factory floor—helping make it ready for production use right out of the box and enabling administrators and researchers to focus on running their applications rather than on complex cluster tuning and management.

Garima Kochhar is a systems engineer in the Scalable Systems Group at Dell.

Jacob Liberman is a development engineer in the Scalable Systems Group at Dell.

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

Jason Bishop is vice president of engineering at Clustercorp.

¹For more information, visit DELL.COM/ServerStorageDeployment.