Secure HPC Cluster Management

with Ninth-Generation Dell PowerEdge Servers

The increasing use of high-performance computing (HPC) clusters for critical or sensitive high-performance computations has created a need for secure node management and monitoring capabilities. This article discusses the available tools for creating a secure HPC environment with the Dell OpenManage™ suite and ninth-generation Dell™ PowerEdge™ servers.

BY ARUN RAJAN, TONG LIU, YUNG-CHIN FANG, AND SAEED IQBAL, PH.D.

High-performance computing (HPC) clusters require secure management to protect against internal and external misuse or abuse of compute resources. In the past, HPC cluster security typically was not a significant problem because clusters were built for private use by a handful of dedicated users. However, many enterprises are now using large-scale clusters that are often shared across departments with easy public access—and increased vulnerability has become an issue. Vulnerable features of HPC clusters include high-bandwidth connections (which can be used to launch denial-of-service attacks on other sites, for example), massive computational power (which can be used to execute parallelized password-cracking tools), and extensive storage (which can be hacked and used to save copyrighted or illegal information).  

Because of this increased vulnerability, substantial efforts have been made to develop HPC management and monitoring tools. One way to provide security in an HPC cluster environment is to enforce secure communication—that is, only nodes and users that are meant to access certain resources should be allowed access. Secure communication can be provided by adhering to secure protocols, providing firewall-like features such as port blocking, allowing role-based authentication for individual users, and providing credible login authentication procedures and sufficient data encryption. Dell HPC cluster environments offer

---

such measures to help provide high levels of security while utilizing a wide range of management components.

**Using ninth-generation Dell PowerEdge servers in HPC clusters**

Dell HPC clusters can comprise both eighth- and ninth-generation Dell PowerEdge servers. Ninth-generation PowerEdge servers introduce the Intel® 5000X and 5000P chipsets with 1,066 MHz and 1,333 MHz dual frontside buses, enhanced processor-memory bandwidth, quad-channel double data rate 2 (DDR2) fully buffered dual in-line memory modules (DIMMs), and PCI Express I/O architecture. These servers use the dual-core Intel Xeon® 5xxx processors, which are based on the Intel Core™ microarchitecture.

The major hardware management components in ninth-generation Dell PowerEdge servers are the Dell Remote Access Controller 5 (DRAC 5) and the baseboard management controller (BMC). The DRAC 5 is used for out-of-band management and provides security features for its command-line interface (CLI) and Web browser–based graphical user interface (GUI) by using Secure Sockets Layer (SSL). The BMC and the associated BMC Management Utility are compliant with the Intelligent Platform Management Interface (IPMI) 2.0 specification with enhanced Remote Management Control Protocol + (RMCP + ) authentication.

Figure 1 shows a typical Dell HPC cluster architecture with public (external) access through network interface card 2 (NIC2), the out-of-band management channel through the DRAC 5 fabric, and the in-band management channel through the cluster and administration fabric over NIC1. NIC1 also supports out-of-band management through the BMC. Figure 2 shows an overview of the security protocols used in the Dell HPC management infrastructure.

**Tools for managing Dell HPC cluster nodes**

The Dell OpenManage suite enables administrators to monitor and manage Dell PowerEdge servers remotely, streamlining node management in HPC clusters. This suite includes several components: Dell OpenManage IT Assistant, Dell OpenManage Server Administrator, Dell OpenManage Storage Services, the BMC Management Utility, the Dell Update Package, the Software Update Utility, the Dell OpenManage Deployment Toolkit, Dell OpenManage Server Assistant, and Dell OpenManage Online Diagnostics. Open source products such as OpenIPMI, IPMItool, Ganglia, and Cluster Monitoring (Clumon) are also available to manage Dell HPC clusters.

The Platform Open Cluster Stack (OCS) software stack—based on the San Diego Supercomputer Center (SDSC) National Partnership for Advanced Computational Infrastructure (NPACI) Rocks—is a comprehensive HPC cluster toolkit, designed to simplify the deployment and management of large-scale Linux® OS–based clusters. Developed by Platform Computing, Platform OCS can be used along with Dell OpenManage to provide comprehensive node management capabilities—for example, to configure cluster nodes for remote management through IPMI and console redirection.

**Implementing secure change management with SNMP**

Simple Network Management Protocol (SNMP) is a systems management standard originally designed for network management. It
is an application-layer protocol that is part of the TCP/IP protocol suite and facilitates the exchange of management information between network devices. SNMP standards are defined by the Internet Engineering Task Force.

SNMP-managed systems provide data to a management system through SNMP agents. At each SNMP agent, a community string is configured that is transmitted as part of all SNMP request messages. Using these community strings means that the requester cannot be verified as part of the specified community, resulting in limited security. Because Get requests are noninvasive, limited security levels can be permitted. However, this limited security is insufficient for Set operations, which can result in intrusive actions such as RAID reconfiguration, system power down, and so on; for this reason, certain vendors do not support SNMP Set operations at all. Dell OpenManage Server Administrator supports SNMP Set operations in a secure manner by forcing its SNMP agents to implement a hash/digest mechanism to help prevent unauthorized operations; however, only Server Administrator SNMP management applications support this mechanism. In an HPC cluster, Server Administrator is usually used to monitor and diagnose remote node health and support OS-level services such as storage management.

The management information base (MIB) is a type of database used to store management data, and is part of a standard that describes the managed objects contained in the MIB. For example, the MIB provides management data, and the Cluster Group defines attributes such as the number of systems in the cluster and the cluster’s capabilities, type, and name. Server Administrator provides various MIBs for different purposes, such as the Instrumentation MIB for instrumentation data, the Remote Access MIB for in-band information on remote hardware, and the Change Management MIB for monitoring the inventory of devices and applications with SNMP management applications. The SNMP master transmits and receives SNMP requests, and the SNMP extension agent registers MIB objects. SNMP traps (asynchronous events generated when some significant action has taken place) are used to preserve network bandwidth, by preventing the need for frequent polling over critical parameters. In an HPC cluster, MIB implementations provide interoperable manageability among various management frameworks.

### Transitioning to IPMI 2.0

IPMI, a joint effort promoted by Dell, Intel, Hewlett-Packard, and NEC, is a hardware-level interface specification that defines a common, abstracted, message-based interface for platform monitoring and control functions. IPMI defines common interfaces to intelligent hardware used to monitor a server’s physical health characteristics, such as temperature, voltage, fans, power supplies, and chassis. Since IPMI 1.0 was introduced in 1998, more than 170 companies across the industry have adopted IPMI, including system and motherboard original equipment manufacturers (OEMs), silicon vendors, and embedded computer manufacturers. IPMI 2.0 contains significant enhancements over IPMI 1.5; it is backward-compatible with previous versions and introduces important features and security enhancements that can be applied to help reduce security concerns for HPC management, including the following:

- **Enhanced authentication**: Extensions to the protocols for IPMI Over IP, collectively referred to as RMCP+, support algorithms that provide robust key-exchange processes for establishing sessions and authenticating users. RMCP+ incorporates authentication based on Secure Hash Algorithm 1 (SHA-1) and supports the Advanced Encryption Standard (AES).
- **Virtual LAN (VLAN)**: VLANs work with VLAN-aware routers and switches to allow a physical network to be partitioned into virtual networks in which a group of devices on different physical LAN segments can communicate with one another as if they were all on the same physical LAN segment. Administrators can use this technology to set up a management VLAN in which only devices that are members of that VLAN receive packets related to management, and conversely these devices are isolated from network traffic for other VLANs.
- **Serial Over LAN (SOL)**: SOL provides a mechanism that enables the serial controller of a managed system to be redirected over an IPMI Over IP session. SOL is implemented as a payload type under the payload capability in RMCP+.
- **Payloads**: RMCP+ enables IPMI Over IP sessions for other types of traffic in addition to IPMI messages, which include both standard payload types defined in the IPMI specification (such as SOL) and OEM value-added payload types.
- **Encryption**: IPMI messages and other payloads carried over RMCP+ can be encrypted, enabling confidential remote configuration of parameters such as user passwords and transfer of sensitive payload data over SOL.
• **Extended user login options:** Extended options support role-only logins for simple environments in which administrators can enable logins according to a given privilege level, without needing to assign or configure usernames. Support for two-key logins, which require both a user-specific and BMC-specific key to connect to a given BMC, enables administrators to configure BMCs for a robust environment.

**Performing secure remote management with RMCP**

LAN interface specifications define how IPMI messages can be sent to and from the BMC encapsulated in RMCP packets as datagrams, a capability also referred to as IPMI Over LAN. IPMI 2.0 defines an extended packet format and capabilities that are collectively referred to as RMCP+. RMCP+ uses the RMCP packet format, but defines extensions to the fields defined under the IPMI message class data carried within the RMCP packet. These extensions support enhanced authentication, encryption, discovery, and the ability to carry additional types of traffic (payloads) in addition to IPMI messages over an IPMI session, whereas IPMI 1.5 supported carrying only IPMI messages.

**Enhancements with RMCP+**

RMCP+ provides support for multiple payload types over an IPMI session, including both standard payloads (such as the payload for the SOL capability defined in this specification) and OEM payloads. It incorporates enhanced user authentication algorithms, including session-setup and key-handling algorithms that are more robust than those for IPMI Over LAN in IPMI 1.5. IPMI messages and other payloads can be encrypted under an IPMI session, enabling confidentiality for remote operations such as setting user passwords and for SOL.

RMCP+ follows many of the packet format and authentication elements defined for RMCP as specified in the Distributed Management Task Force (DMTF) Alert Standard Format (ASF) 2.0 specification. RMCP+ supports encrypted/uncrypted and authenticated/unauthenticated traffic on a single connection; encryption and authentication are handled at the IPMI message class level, which means that encrypted and authenticated sessions can be established on any User Datagram Protocol (UDP) port, including port 26Fh. IPMI allows a BMC to be configured so that authentication and encryption are utilized only when the payload or privilege level of operation requires it, eliminating the need to have all traffic authenticated or encrypted on a connection when only a small portion of the traffic may require that level of security. This configuration can provide a significant performance benefit when using inexpensive microcontrollers for BMCs.

**Secure management using RMCP+**

To utilize the enhanced security of RMCP+, administrators set the encryption key. One way to set the RMCP+ key is to invoke the BIOS utility menu for remote access by pressing Ctrl + E during system startup and entering the key manually in the BMC configuration options. Another way to set the key and assign user privileges is on the managed node through Dell OpenManage Server Administrator using omconfig utility commands (see Figure 3).

Once the RMCP+ key is set along with the user ID and password, accessing the BMC requires users to specify all three parameters, including the hexadecimal RMCP+ encryption key. For example, using the ipmish CLI utility (available with the BMC Management Utility suite) in a Dell HPC cluster with a one-to-one IPMI session to a specific remote server to obtain the system status would require the following command:

```
ipmish -lp IP address -u username -p password
-k hexadecimal encryption key=sysinfo
```

These commands may be scripted (including the RMCP+ key) for easy access to the server BMCs in the HPC cluster.

**Securing out-of-band management with the DRAC 5**

The DRAC 5 provides out-of-band remote management capabilities, crashed system recovery, and power control functions for Dell PowerEdge servers. The DRAC 5 communicates directly with the BMC and can be used to configure e-mail alerts and log event data. The DRAC 5, powered by the system in which it is installed and utilizing its own processor and memory subsystems, can be connected with a LAN cable through its RJ-45 connector to provide an out-of-band fabric for external communication in an HPC environment. It also provides several useful features such as virtual media, console redirection, and a comprehensive GUI to add alerts and modify configuration parameters.
To prevent unauthorized access to remote systems, the DRAC 5 provides IP address filtering, which defines a specific range of IP addresses that can access the DRAC 5, as well as IP address blocking, which limits the number of failed login attempts from a specific IP address. The racadm CLI enables administrators to locally or remotely configure and manage the DRAC 5. This CLI runs on the management station and the managed system and is included on the Dell Systems Management Consoles CD; it is also available through the serial, Telnet, and Secure Shell (SSH) console to the DRAC subsystem using the DRAC IP address. To restrict the login to a single IP addresses (for example, 192.168.0.57), administrators should use the full mask, as shown in the following example commands:

```
racadm config -g cfgRacTuning
   -o cfgRacTune1pRangeAddr 192.168.0.57
racadm config -g cfgRacTuning
   -o cfgRacTune1pRangeMask 255.255.255.255
```

The following example commands prevent a client IP address from establishing a session for five minutes if that client has failed its five login attempts in a one-minute period of time:

```
racadm config -g cfgRacTuning
   -o cfgRacTune1pRangeEnable 1
racadm config -g cfgRacTuning
   -o cfgRacTune1pRangeAddr 192.168.0.57
racadm config -g cfgRacTuning
   -o cfgRacTune1pRangeMask 255.255.255.255
racadm config -g cfgRacTuning
   -o cfgRacTune1pFailCount 5
racadm config -g cfgRacTuning
   -o cfgRacTune1pFailWindows 60
racadm config -g cfgRacTuning
   -o cfgRacTune1pFailPenaltyTime 300
```

The Server Management Working Group’s Server Management Command-Line Protocol (SM CLP) specification is a subcomponent of the overall DMTF Systems Management Architecture for Server Hardware (SMASH) standard. The DRAC 5 provides support for SM CLP, which is hosted from the DRAC 5 controller firmware. SM CLP supports Telnet-, SSH-, and serial-based interfaces. The DRAC 5 SM CLP interface is based on SM CLP Version 1.0, which enables users to perform server power management (powering up, powering down, or rebooting the system) as well as system event log management (displaying or clearing system event log records). The SM CLP interface with the DRAC 5 makes the DRAC hardware independent and provides for secure interoperability. The DRAC 5 also provides the following security features:

- User ID and password configuration through the secure GUI or racadm CLI
- Session time-out configuration (in seconds) through the GUI or racadm CLI
- IP port configuration (where applicable)
- SSL, which uses an encrypted transport layer for enhanced security

The DRAC 5 helps provide a reliable one-to-one out-of-band management fabric for Dell HPC clusters, and the security enhancements can help improve cluster security for critical management tasks under an architecture built on SMASH and SM CLP standards.

**Using Dell OpenManage security features**

The Dell OpenManage suite provides the following security features:

- Role-based authority that allows administrators to configure specific privileges for each user
- User ID and password configuration through the GUI or CLI in most cases
- SSL 128-bit and 40-bit encryption (for countries where 128-bit is not acceptable), except when using Telnet
- Session time-out configuration (in minutes) through the GUI or CLI
- Configuration of many commonly known ports

Features such as role-based access control, authentication, and encryption in both GUIs and CLIs help ensure security through access administration. Security is enforced by restricting the operations that can be executed by users in specific roles, with some users assigned multiple roles; this feature can help enable security administration to resemble any organization’s structure. Group privileges decide access rights for each Dell OpenManage Administrator user. The three levels are User (can view information), Power User (can set warning threshold values, run diagnostic tests, and configure alert actions), and Administrator (can perform power-down actions, configure auto-recovery, clear logs, and send e-mail).

Dell OpenManage software also helps provide security through authentication and encryption.

**Authentication.** The Dell OpenManage Server Administrator authentication scheme validates the context of the current process under which the CLI or GUI is executed, to help ensure that Server Administrator functions are properly authenticated. For authentication mechanisms, Server Administrator uses Integrated Windows Authentication and the Pluggable Authentication Modules (PAM) library on the Microsoft® Windows® and Red Hat® Enterprise Linux operating systems, respectively. PAM is a documented library of functions that allows administrators to determine how individual applications authenticate users.
Encryption. Dell OpenManage Server Administrators uses an HTTP Over SSL (HTTPS) connection to help protect the managed system. The Windows, Red Hat Enterprise Linux, and Novell® SUSE® Linux Enterprise Server operating systems use Java Secure Socket Extension (JSSE) to help protect user credentials and other sensitive data transmitted when a user accesses the Server Administrator GUI over the socket connection.

Features such as authentication, encryption, role-based authority, and port configuration have been incorporated into the Dell OpenManage suite. These features are critical to helping provide a secure operating environment for managing Dell HPC clusters. Administrators can incorporate a hierarchical role distribution (and apply various other security settings) to help prevent unauthorized users from intentionally or unintentionally misusing cluster components and capabilities.

Managing Dell HPC clusters in a secure environment

Dell offers a comprehensive management package for ninth-generation Dell PowerEdge servers through its Dell OpenManage suite, which can be utilized for HPC clusters. In addition to security protocols, the Dell OpenManage suite provides OS-level agents for monitoring software on compute nodes, which use encryption algorithms and authentication mechanisms to enhance cluster security against external attacks. The DRAC 5 subsystem, which provides a hardware- and software-based fabric for out-of-band management, also includes security features such as SSL for the GUI, SSH access to the racadm CLI, configurable RMCP + encryption keys, and role-based authentication—which enables an easily configurable, hierarchical access system. Used together, these components can help enhance security of Dell HPC clusters.

Arun Rajan is a systems engineer in the Scalable Systems Group at Dell. His current interests and responsibilities include HPC cluster management, cluster computing packages, performance benchmarking, and product development. He has a B.E. in Electronics and Communications Engineering from the National Institute of Technology, Tiruchirappalli, in India, and an M.S. in Computer and Information Science from the Ohio State University.

Tong Liu is a systems engineer in the Scalable Systems Group at Dell. His current research interests are HPC cluster management, high-availability HPC clusters, and parallel file systems. Tong serves as a program committee member of several conferences and working groups on cluster computing. Before joining Dell, he was an architect and lead developer of High Availability Open Source Cluster Application Resources (HA-OSCAR). Tong has an M.S. in Computer Science from Louisiana Tech University.

Yung-Chin Fang is a senior consultant in the Scalable Systems Group at Dell. He specializes in HPC systems, advanced HPC architecture, and cyber- infrastructure management. Yung-Chin has published dozens of conference papers and articles on these topics. He also participates in industry-standard organizations, academic conferences, and HPC cluster–related open source communities as a Dell representative.

Saeed Iqbal, Ph.D., is a systems engineer and consultant in the Scalable Systems Group at Dell. He is one of the lead engineers for the Beowulf HPC cluster program based on ninth-generation Dell servers. His current work involves evaluation of resource managers and job schedulers used in HPC clusters in industry and academia. He is also involved in performance analysis of HPC clusters. Saeed has a B.S. in Electrical Engineering and an M.S. in Computer Engineering from the University of Engineering and Technology in Lahore, Pakistan, and a Ph.D. in Computer Engineering from the University of Texas at Austin.

FOR MORE INFORMATION

Dell systems management documentation: support.dell.com/support/systemsinfo/documentation.aspx?cat=68&subcat=111

IPMI 2.0 specification: www.intel.com/design/servers/ipmi


