

Whitepaper: Data Center Manageability Interface (DCMI) Specification

Authors:

Hari Ramachandran, Manageability Architect, Intel Corporation

Harry Rogers, Principal Hardware Architect, Microsoft Corporation

Date published:

August 2008

© 2008 Intel Corporation. All rights reserved. Portions © Microsoft Corporation.

Information in this document is provided in connection with Intel products. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN THE TERMS AND CONDITIONS OF SALE FOR PRODUCTS, NEITHER INTEL NOR MICROSOFT ASSUMES ANY LIABILITY WHATSOEVER AND INTEL AND MICROSOFT DISCLAIM ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF THEIR RESPECTIVE PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

UNLESS OTHERWISE AGREED IN WRITING BY INTEL, THE INTEL PRODUCTS ARE NOT DESIGNED NOR INTENDED FOR ANY APPLICATION IN WHICH THE FAILURE OF THE INTEL PRODUCT COULD CREATE A SITUATION WHERE PERSONAL INJURY OR DEATH MAY OCCUR.

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined." Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information herein is subject to change without notice. Do not finalize a design with this information.

Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or by visiting Intel's Web Site.

Intel, the Intel logo, Xeon and Xeon Inside are trademarks of Intel Corporation in the U.S. and other countries.

Other names and brands may be claimed as the property of others.

Abstract:

The Data Center Manageability Interface (DCMI) specification explicitly addresses the unique requirements of server platform management within Internet Portal Data Centers (IPDC) and other High Density Data Centers where large numbers (into the ten's of thousands) of servers are deployed. This specification seeks to help High Density Data Centers reduce the cost and complexity of server platform management by defining a concise set of platform monitoring and control capabilities to deliver reliable, stable, interoperable, cross-vendor manageability targeted to the needs of this segment. This paper presents the needs and rationale behind the creation of the DCMI, the main capabilities and benefits of the specification, and example application scenarios illustrating DCMI usage in a High Density Data Center environment.

High Density Data Center Manageability

In the last few years, the computing industry has seen a new trend in Server deployment referred to in this paper as High Density Data Centers. The Data Center as defined by wikipedia.org is a facility used to house computer systems and associated components, such as telecommunications and storage systems. It generally includes redundant or backup power supplies, redundant data communications connections, environmental controls (e.g., air conditioning, fire suppression), and special security devices [http://en.wikipedia.org/wiki/Data_center].

A High Density Data Center is defined as a class of Data Center that incorporates a very large number of servers (starting in the thousands and going up into the ten's of thousands).

The enormous number of servers used in the High Density Data Center presents some unique problems in the areas of server manageability. Different server platform vendors offer a wide and rich variety of platform manageability features and interfaces for managing many different classes of servers in many application environments. The goal of covering a broad range of server applications along with a desire for feature differentiation across vendors has led to an explosion of the manageability features that are available.

Many High Density Data Center customers see this as problematic. The breadth and complexity of the vendor offerings often means that the High Density Data Centers need to invest time developing expertise in understanding and maintaining each different vendor's system management. Different server vendor's solutions are often tied to requiring vendor-specific management console software, agents, drivers, and utilities that are required to provision and utilize the management capabilities. Features were added and changed with each new platform generation, often resulting in a need for software updates. In addition, many of the vendors' management consoles have interactive interfaces

or command line interfaces that are not well suited to the scale of Internet portal data centers or other High Density Data Centers.

The presence of a large number of features that are not needed nor utilized by High Density Data Centers is viewed as adding unnecessary cost and complexity and decreasing the reliability of the solution. When there was a degree of commonality in the features that High Density Data Center customers cared about, those features were often implemented in a non-interoperable manner across vendors. Capabilities would also vary across different platforms from the same vendor.

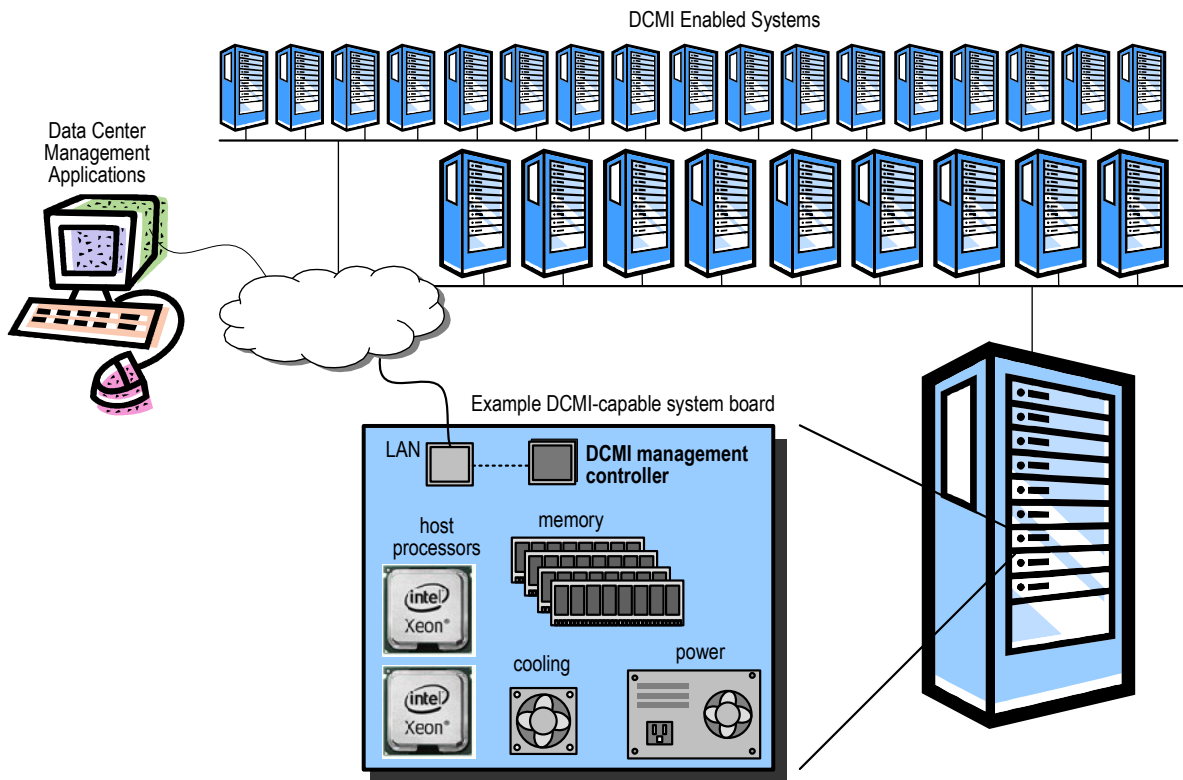


Figure 1: DCMI-based Manageability

Many High Density Data Center customers also desire the capability to readily create and integrate the platform management solutions with their own High Density Data Center management software and architectures. The ability to develop custom software for managing their systems not only enables the High Density Data Center to tailor manageability to their needs, but is often considered an area of expertise and differentiation among competing High Density Data Centers.

However, the interface technologies used across different vendors are also often complex, requiring the High Density Data Center to climb a steep learning curve and develop local experts who understand the protocols, structures, and data models. In some cases the platform management interfaces are

proprietary and accessing them directly requires the High Density Data Center to obtain special licensing or enter confidentiality agreements in order to get the documentation needed to access the capabilities. Often, vendors do not support direct access to their management capabilities, but instead require the Data Center customer to go through an intermediate driver API or console plug-in, or other software from the platform vendor. This has impeded the ability for High Density Data Center customers to develop, deploy, and maintain their High Density Data Center management applications.

The existing solutions have had impact on the server platform vendors as well. The breadth and constant evolution of features have increased the platform cost, only to deliver features that many customers do not need. The lack of a simple and interoperable interfaces and capabilities has required substantial development investments in maintaining and supporting those solutions with their High Density Data Center customers.

The Origins of DDMI

Based on Intel's discussions with several key High Density Data Center and IPDC customers, there a set of common characteristics emerged that High Density Data Center customers like Microsoft were looking for:

- Concise** - Management capabilities are focused on the fundamental platform management needs of High Density Data Centers without unnecessary options or unneeded features.
- Interoperable** - The same capabilities are provided in the same way across multiple platforms from multiple vendors.
- Open** - Specifications available to all platform vendors and High Density Data Center customers for deployment without restrictive licensing or royalties.
- Reliable** - The manageability capabilities must be based on proven technologies and delivered in a dependable manner.
- Simple** - The interfaces, capabilities, and data models are readily comprehensible to any developer that is experienced in basic networking and software, without the need to first become a manageability technology or protocols expert.
- Stable** - The interfaces and capabilities remain stable and compatible across platform generations. Configuration and implementation options are minimized.
- Low Cost** - The implementation of the platform interfaces and capabilities does not require imposing a significant cost on top the basic cost of the platform electronics.

After receiving this input, Intel held additional discussions with a number of key High Density Data Center customers and found that there was also a very high degree of commonality in the types of monitoring, control, and security capabilities they sought.

The next step was to examine existing technologies and specifications for their applicability to the High Density Data Center manageability environment. The IPMI (Intelligent Platform Management Interface) 2.0 specifications were chosen as the starting point. These specifications, first released in February 2004, were designed for providing a simple and low cost implementation of server manageability features and have been deployed by numerous server vendors. As a technology, the IPMI 2.0 specifications already covered many of the desired characteristics for a High Density Data Center server management interface.

However, the IPMI 2.0 Specifications were written to cover many server segments, and therefore had many options that were not applicable or valued for High Density Data Center management. Additionally, the IPMI specifications did not provide a set of mandatory requirements that fit the functions that were identified by High Density Data Centers. Thus, the mandatory requirements of IPMI did not meet the High Density Data Center needs, and the full definition of IPMI 2.0 capabilities exceeded them. As a result, implementations of IPMI fell somewhere in-between and varied significantly across vendors. A number of High Density Data Center customers had already looked at IPMI 2.0 -based systems and rejected it as a general solution because of these problems.

The solution was to create a specification based on the proven technology of IPMI 2.0 that clearly identified the core set of mandatory capabilities and interfaces that are required for High Density Data Centers. Upon additional review, it was recognized that there were some areas where a very small number of High Density Data Center -specific extensions to IPMI 2.0 were required. These were added to the specification and subsequently reviewed by key High Density Data Center customers like Microsoft and server platform vendors. The end result is the DCMI 1.0 Specification.

DCMI (Data Center Manageability Interface) Capabilities

DCMI defines the manageability requirements as a set of standard manageability features and interfaces. By using IPMI 2.0 as the basis, DCMI implementations can take advantage of the existing industry infrastructure of IPMI management controllers, firmware, and expertise. In many cases, the server platform hardware support for DCMI is already present and only the management controller's firmware needs to be changed to support DCMI.

The capabilities of DCMI are available to both local and remote management software. These capabilities include:

Standardized Local and Remote Access Interfaces -	DCMI defines standardized interfaces that enable management applications to access the DCMI platform management subsystem locally (also referred to as "in-band" access) or remotely by direct communication with the platform management subsystem over LAN independent of the system's power state or host software (also referred to "out-of-band" access). The platform management subsystem operates on standby power so that it is always available for remote access.
--	--

Platform Power on/off/reset control -	The ability to control platform power and reset independent of host software. This capability is typically used to support remote provisioning and emergency recovery under situations where this cannot be accomplished via local software.
Baseline Sensors -	DCMI defines a baseline set of required sensors including Inlet Temperature, CPU Temperature, Ambient Temperature and Power Reading.
Platform Identification -	DCMI defines interfaces for retrieving system ID and asset ID to provide a common mechanism for obtaining positive and unique identification of the managed system.
Improved sensor access -	The IPMI definition of sensor access requires large number of command transactions and parsing at the management application to discover and/or read the sensors. DCMI defined a small extension command that greatly reduces the steps for accessing the mandatory sensors.
Authentication, integrity, and confidentiality	Remote session authentication, integrity, and confidentiality are provided using proven industry hashing and encryption algorithms in conjunction with the RMCP+ session establishment protocol defined by [IPMI 2.0].
Platform Event Logging -	Sensor 'out-of-range' events and other platform hardware status events are logged to a non-volatile storage area for local or remote retrieval, enabling 'post mortem' and non- real-time analysis of platform events.

DCMI Relationship to IPMI

The primary difference between DCMI and other specifications is that DCMI was originated and designed expressly for the platform manageability needs of High Density Data Center servers, as specified by High Density Data Center manageability users themselves.

DCMI references the existing IPMI 2.0 specification as much as possible, mainly providing a tighter definition and clarification about which IPMI 2.0 capabilities are mandatory for High Density Data Center servers. In order to improve interoperability, DCMI minimizes the number of optional capabilities in the specification.

In addition, DCMI defines a small number of DCMI-specific functions to address some specific High Density Data Center requirements that were not comprehended in IPMI 2.0. These extensions provide a reduction in the number of commands needed to complete certain manageability operations, such as querying for capabilities, accessing sensor information, etc., and enable management applications to identify DCMI-based systems.

Since DCMI 1.0 is based on the IPMI 2.0 specification, industry support for IPMI 2.0 can be highly leveraged. The format of DCMI commands complies with the IPMI 2.0 specification, enabling existing in-band drivers for IPMI to be re-used for DCMI. Similarly, DCMI utilizes the RMCP+ transport for LAN communications, so most transport infrastructure developed for IPMI can also be re-used. As a result, existing IPMI-based management applications can be updated to support DCMI with minor modifications to scripts and objects. Management Controllers designed for IPMI should also be suitable for DCMI, and platform subsystem hardware implementations required for DCMI should be nearly identical to those used for IPMI 2.0 (and in many cases will be significantly simpler).

DCMI Simplicity and Efficiency

DCMI is highly efficient from a network transaction point-of-view. Establishing a DCMI remote communication session typically requires an exchange of only ten (10) Ethernet packets (five in each direction) between the remote management application and the management controller. This can be reduced to as few as six in many environments.

The message encoding is also highly optimized. For example, the message content for many of the common DCMI commands is ten bytes or fewer. Similarly, logging of DCMI-monitored events and conditions is accomplished with a concise format. This reduces the storage requirements for data center logs both on the managed system and when log information is aggregated by Data Center management software.

DCMI messaging uses a simple request-response protocol and a consistent message format that is easy to comprehend. DCMI is a tightly specified using a command set with defined standardized behavior for each command. Commands are organized according to functional areas, such as System Power and Reset Control, Logging, Identification Support, Sensors, Power Management, and Connectivity, with just four or five commands typically defined for each area.

The efficiency of DCMI provides several benefits including the ability to use less expensive management controllers, shorter network transmissions, and improved performance compared to heavier-weight management protocols and data models. The simplicity of DCMI reduces the time to learn the interfaces and apply them in applications, and eliminates the need for specialized, hard-to-find development expertise. The simplicity also improves robustness by reducing the complexity of the platform management subsystem, and improves interoperability by reducing the number of options in the implementation.

DCMI Interoperability

To help verify interoperability, Intel has created a DCMI Conformance test suite for use as a check by DCMI suppliers and customers. This suite helps to clearly identify any gaps and also reduces the process of checking the conformance for the specification. The conformance test suite can be downloaded from the Intel web site.

Application Scenarios

In-Band Monitoring

The preferred mode of High Density Data Center server monitoring is in-band, that is, with the support of the resources controlled by the managed node's operating system. In-band monitoring typically leverages the same monitoring infrastructure that is used for application and operating system metrics. Therefore, physical metrics, such as those supported by DCMI, can be monitored and correlated with existing consoles and agents by adding a few DCMI-specific components to the managed node.

Figure 2: In-Band Monitoring Application, below, shows a managed node managed by an operations management console. In this example, Microsoft System Center Operations Manager is used as the management console (MC) and a System Center agent runs as a Windows application in the managed server node (MN). The MC communicates with the MN, via TCP/IP, using the Web Services for Management protocol (WS-MAN). The SC agent runs a discovery script to discover DCMI capabilities and populate a Common Information Model (CIM) database. The MC communicates with the MN, via TCP/IP, using the Web Services for Management protocol (WS-MAN). The SC agent runs a discovery script to discover DCMI capabilities and populate a Common Information Model (CIM) database.

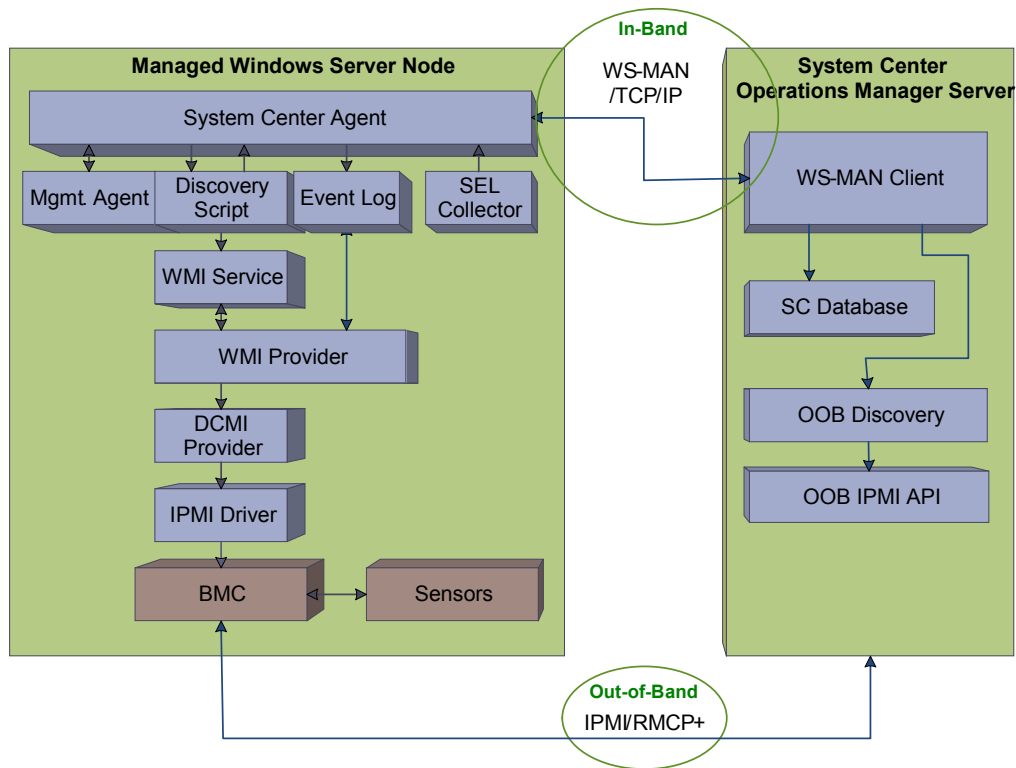


Figure 2: In-Band Monitoring Application

The use of power is a major component both capital and operational High Density Data Center costs. In this scenario, DCMI sensing of server power consumption and temperature can be used to understand how efficient a High Density Data Center and its resident servers are. These costs are best managed by understanding power consumption on a server or application basis. In-band monitoring allows an

automated console to gather power and thermal metrics, application and operating system metrics with a common infrastructure.

DCMI enables this directly by defining explicit sensors, including those available with power supplies instrumented to report server power utilization. Instrumented power supplies report primary power usage on a per-server basis to understand how a given application performs for the power that its host server consumes. The instrumentation within the power supply leverages hardware. With an operation management console, such as Microsoft Systems Center Operations Manager, power metrics can also be aggregated arbitrarily or post-processed to understand power/performance tradeoffs of multi-node applications.

DCMI also specifies explicit temperature sensors that enable fine-grain thermal profiling of a High Density Data Center. For example, measuring the inlet (cold aisle) temperature at each server enables precise mapping of temperature gradients within a rack or within a row of racks. With this data, much of the cost of over-cooling High Density Data Centers can be avoided by cooling to the highest appropriate temperatures.

Out-of-Band Monitoring and Control

In-band monitoring is not always available when there are hardware failures at the managed node, when an operating system is unresponsive or when in-band monitoring is not desirable. Also, managed nodes that are regularly powered down and subsequently powered up remotely must use an out-of-band mode.

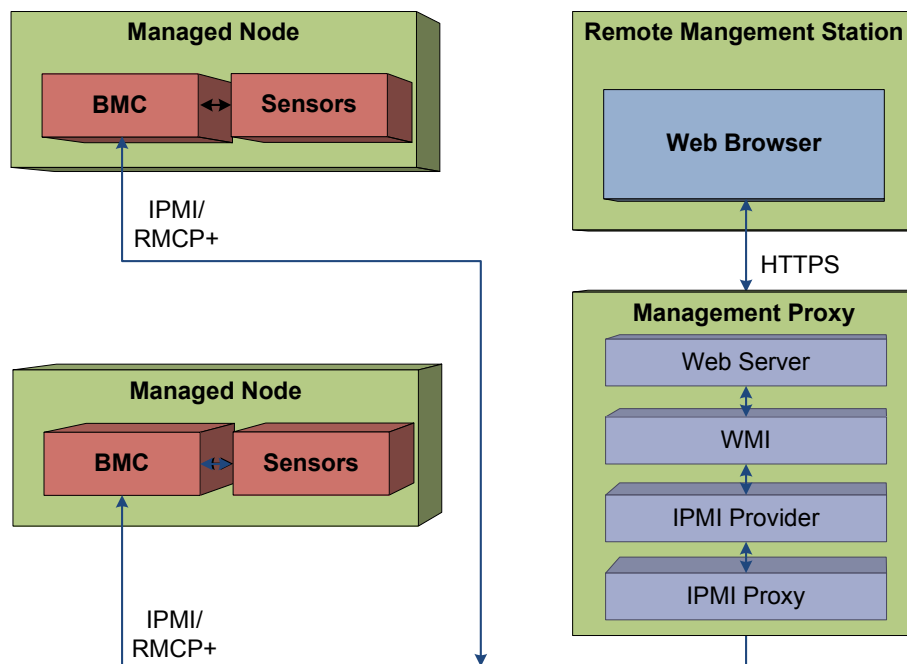


Figure 3: Proxy-Based Management Console

Figure 3 illustrates an out-of-band implementation in which multiple managed nodes are monitored and controlled manually via a remote management station running a common Web browser. The Web browser accesses a Web service that abstracts basic DCMI functions that are required to remotely monitor and control a managed node's hardware. Such functions may include a serial-over-LAN session to the managed node's serial console, power control and power state monitoring.

The Web service accesses IPMI metrics and control via a standard management interface (implemented here as a WMI provider) which, in turn, access and IPMI Provider. The IPMI provider interfaces to an IPMI proxy which can function as a selector of managed nodes and, optionally, can abstract role-based DCMI authentication into a user-based authentication scheme, such as Active Directory. The IPMI Proxy then communicates with the managed nodes' BMCs via RMCP+/IP.

Summary & Conclusion

The DCMI Specifications define a set of standardized interfaces and capabilities that are designed explicitly for the platform management monitoring and control needs of High Density Data Center servers. The specifications were developed with input and review from key High Density Data Center customers and server platform vendors.

DCMI takes advantage of proven technology from the IPMI specifications and leverages IPMI experience among platform vendors, and the ecosystem of firmware, software, management controllers, and expertise to speed the development of DCMI-based servers. A conformance test suite helps assure interoperability between different server vendors' implementations.

DCMI-based servers are expected to deliver reliable, low cost, and interoperable manageability that can be readily used for in-band and out-band Data Center applications. DCMI works with the Data Center applications to improve the efficiency and reduce the cost and overhead of managing server platforms in High Density Data Centers. The DCMI 1.0 specification is available now.

For more information regarding the specification and conformance test suite please go to <http://www.intel.com/technology/product/DCMI>