



– DDMI RR–

Data Center Manageability
Interface

Reliability & Resilience
Guidelines

v1.0

Revision 1.0

Intel Corporation

THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITH NO WARRANTIES WHATSOEVER, INCLUDING ANY WARRANTY OF MERCHANTABILITY, NONINFRINGEMENT, FITNESS FOR ANY PARTICULAR PURPOSE, OR ANY WARRANTY OTHERWISE ARISING OUT OF ANY PROPOSAL, SPECIFICATION OR SAMPLE. Intel disclaims all liability, including liability for infringement of any proprietary rights, relating to the use of any information in this document. No license, express or implied, by estoppel or otherwise, to any intellectual property rights is granted herein, except that a license is hereby granted to copy and reproduce the information in this document for internal use only.

Intel retains the right to make changes to this document at any time, without notice. Intel makes no warranty for the use of this document or the information herein and assumes no responsibility for any error which may appear in the document, nor does it make a commitment to update the information contained herein.

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.

Intel and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

Copyright 2007-2009 Intel Corporation. All Rights Reserved.

Revision History

Date	Version	Rev	Modifications
01/27/2009	1.0	1.0	1.0 release

Table of Contents

1. INTRODUCTION	5
1.1. SCOPE	5
1.2. AUDIENCE	5
1.3. DOCUMENT ORGANIZATION	6
1.4. REFERENCE DOCUMENTS	6
1.5. CONVENTIONS AND TERMINOLOGY	7
2. DDMI OVERVIEW	8
2.1. DATA CENTER SERVER MANAGEMENT	8
3. RELIABILITY AND RESILIENCE GUIDELINES	9
A.1 SUGGESTED RELIABILITY STATEMENTS FOR IPDCS	12

List of Tables

Table 1, Glossary	7
Table 2, Reliability and Resilience Guidelines.....	9

1. Introduction

This document presents the reliability and resilience guidelines for Internet Portal servers, which are conformant to the Data Center Manageability Interface (DCMI) specification.

Reliability is defined in this context as the predictable and consistent behavior of the DCMI manageability controller within the server and across the Data Centers.

Resilience is defined in this context as the ability of the DCMI manageability controller to recover from or adjust to changes in the external factors of the data centers and continue to provide availability of DCMI capabilities.

The key areas of reliability and resilience are influenced by the Data Center networks, the server operating systems and the behavior of the remote agent.

The DCMI specification defines standardized, abstracted interfaces to the server management subsystem specific to Data Center Servers. This specification is built upon the *Intelligent Platform Management Interface (IPMI) 2.0* specification. The DCMI specification specifies the interfaces and capabilities, but does not specify reliability and resilience for the implementation of those interfaces and capabilities, which is covered in this document.

Data Centers require consistency in the reliability and resilience of the implementations of the DCMI specification in order to yield the most benefit from systems that support DCMI.

The purpose of this document is to act as a guide to help Data Center architects and OEM(s) to provide a baseline for reliability and resilience requirements that can be applied to suppliers of DCMI conformant systems. For example, Data Centers can use this document to define Service Level Agreements (SLA) with their suppliers. Similarly, suppliers can use this document as part of their validation and quality assurance processes for their DCMI implementations.

1.1. Scope

This document defines a baseline set of reliability and resilience requirement targets for DCMI specification conformant platform management subsystems used in Data Center servers.

This document serves as additional guidelines for DCMI specification compliance by covering reliability and resilience characteristics of the platform management subsystem's operation..

1.2. Audience

This document is written for engineers, system integrators and software developers involved in the designing of or interfacing to Data Center Server management hardware. Familiarity with microcontrollers, software programming, and server architecture is assumed. For additional information, refer to the appropriate reference documents.

1.3. Document Organization

Chapters 1 to 2	Provides overview and intent of the guideline
Chapters 3	Describes the reliability and resilient factors
Appendix 1	Suggests Reliability Requirements statements for IPDCs

1.4. Reference Documents

The following documents are companion and supporting specifications for this guideline:

[DCMI] *Data Center Manageability Interface Specification, Version 1.0*, © 2008, Intel Corporation.
www.intel.com/technology/product/DCMI or www.intel.com/go/DCMI

[IPMI 2.0] *Intelligent Platform Management Interface Specification, Version 2.0*, © 2006, Intel Corporation, Hewlett-Packard Company, NEC Corporation, Dell Inc.
www.intel.com/design/servers/ipmi

[RFC 2119] *Key words for use in RFCs to Indicate Requirement Levels*, S. Bradner, Harvard University, March 1997

1.5. Conventions and Terminology

If not explicitly indicated, bits in figures are numbered with the most significant bit on the left and the least significant bit on the right. Also, unless otherwise indicated byte order, command notations, and syntax follow the conventions used in [DCMI].

Refer to [RFC 2119] for terminology definition of shall, should and may.

This document uses the following terms and abbreviations:

Table 1, Glossary

Term	Definition
BMC	Baseboard Management Controller.
DCMI	Data Center Management Interface.
in-band	Refers to functions that are accessed locally by software via an IPMI System (host) interface to BMC.
IPMI	Intelligent Platform Management Interface
IPDC	Internet Portal Data Center such as MS-Live†, Amazon†, Yahoo† etc.
OOB	Out-of-band. Refers to platform management subsystem capabilities that operate independently of the host processors. The term out-of-band is also used to refer to direct remote communication with the platform management subsystem that operates independently of the host processors, such as over LAN.

2.DCMI Overview

2.1. Data Center Server Management

The term Data Center Server Management refers to autonomous monitoring and recovery features implemented directly in server management hardware and firmware. The key characteristic of Intelligent Platform Management is that inventory, monitoring, logging, and recovery control functions are available independent of the main processors, BIOS, and operating system. Platform management functions can also be made available when the system is in a powered down state.

The DCMI specifications defines a set of key components of IPMI that is suited for the Data Centers and delivers interoperability across DCMI implementations on different systems and from different vendors.

3. Reliability and Resilience Guidelines

The following guidelines are a set of reliability and resilience requirements which Data Center consumers may use to augment requirements for servers that incorporate a DCMI conformant platform management subsystem.

Where explicit values are not specified in this section, the requirement levels are set by each individual Data Center based on their reliability and resilience expectations.

The guidelines consist of two parts: a description of the reliability or resilience expectation and the "Measurable Primitives" that provide information on how the requirements can be measured or tested.

Please note these requirements only provide the necessary condition to comply with the requirement, but the sufficient condition needs to be derived by the Data Centers with the help of the OEM(s). The unit of "Measurable Primitives" is defined by the Data Center unless it is explicitly stated in the requirements.

All requirements apply for both deployment and operation phases of the server.

Table 2, Reliability and Resilience Guidelines

Ref #	Details
RR.010	<p>Manageability Controller shall not cause any unexpected traffic in the network.</p> <p>Note: Unexpected can be defined as any network traffic generated autonomously by the management controller that is outside compliance with the [DCMI] OOB (remote) communication specifications.</p> <p><u>Measurable primitives:</u></p> <ol style="list-style-type: none"> 1. Network bandwidth usage is consistent with expected values derived from the DCMI specification. 2. No autonomously generated packets are emitted from the platform management subsystem that are not covered by the DCMI specification. This requirement can be verified with a network packet analyzer.
RR.020	<p>Manageability Controller shall not have significant performance impact on the network behavior of the host for both dedicated and shared network solutions.</p> <p><u>Measurable primitives:</u></p> <ol style="list-style-type: none"> 1. Quantification of the percentage of Host network bandwidth impact due to DCMI out-of-band communications.
RR.030	<p>Manageability Controller shall perform power on/off/reset in a reliable and repeatable manner through out-of-band or in-band interfaces.</p> <p>A DCMI platform subsystem implementation is allowed to work in conjunction with platform vendor-specific features that temporarily block DCMI power on/off/reset control if the platform is in the middle of an operation where the power control change could cause a platform corruption that could not be remotely recovered, such as a BIOS or management controller firmware update. (Whether such behavior or capability exists is outside the scope of the DCMI Specifications). The following requirements apply when the platform is not in the middle of such an operating mode.</p> <p><u>Measurable primitives:</u></p> <ol style="list-style-type: none"> 1. The time between power operations such as power on, off and reset Under normal conditions, new power and reset requests shall be accepted within 30 seconds of initial platform power up (Power enabled from the main switchboard/circuit) and within 30

	<p>seconds of any prior power or reset request. The time interval is measured at the interface to the management controller rather than at the remote console.</p> <ol style="list-style-type: none"> The latency between the power transition request received by the manageability controller and physically observed power transition shall be less than 30 sec. The availability of a power control test plan from the system supplier that provides a description of the methodology used to test for reliable power control operation. <p><u>Additional guidelines</u></p> <ol style="list-style-type: none"> A power control request that is received when no other power control requests are in process should result in the requested power state. Note that a power control request that is received while another power control request is in process may either be rejected by the implementation or accepted as a pending request. If the request is accepted, the [IPMI 2.0] specification does not indicate whether the implementation should abort the operation that is in process or should execute the requests sequentially. In either case, however, the power state should go to the state provided in the last accepted request. A reset or power cycle request that is received while the platform is powered down shall be rejected.
RR.040	<p>Manageability Controller shall support persistent out-of-band session across power operations through out-of-band or in-band interfaces. Periods of non-responsiveness from the management controller due to platform power transitions should be minimal.</p> <p><u>Measurable primitives:</u></p> <ol style="list-style-type: none"> Non-responsive intervals during platform power transitions shall be less than 60 seconds. Note that this interval is measured at the interface to the management controller rather than at the remote console. The number of lost sessions due to power transitions.
RR.050	<p>Manageability Controller shall support Serial Over LAN (SOL) with no data corruption (assumes that hardware handshake is supported by the serial software on the managed system).</p> <p><u>Measurable primitives:</u></p> <ol style="list-style-type: none"> Quantification of data corruption within the serial data transferred via SOL.
RR.060	<p>Manageability Controller shall not be significantly impacted by network traffic surge or cause unexpected network activity.</p> <p>A network surge for DCMI would typically be comprised of an unusual high amount of broadcast traffic or directed traffic that gets routed to the Management Controller, such as ARP requests or multiple back-to-back UDP packets to the RMCP port number that could occur due to an errant remote management software stack or because of a coincidence of accesses to the DCMI platform management subsystem by multiple remote management software stacks.</p> <p>Note if a session loss occurs due to an unusual network traffic surge, session restoration should be possible after the default [IPMI 2.0] Session Inactivity Timeout interval has expired and the surge has ceased.</p> <p><u>Measurable primitives:</u></p> <ol style="list-style-type: none"> Level of network surge(% of switch traffic) Length of network surge Ability to maintain session during network surges.
RR.070	<p>In-band KCS driver / software failures shall not affect DCMI out-of-band communication.</p> <p>DCMI out-of-band communication shall operate independently of the presence or health of an in-band driver. In-band driver or operating system hangs shall not affect the ability to connect out-of-band to the DCMI platform management subsystem. Prematurely terminated or incomplete in-band DCMI message transactions shall not affect out-of-band communication or DCMI operation.</p> <p>Note: This requirement only covers operation using DCMI-specified commands. Since DCMI includes configuration commands that affect out-of-band operation, this requirement does not apply to driver or software errors that cause legitimate DCMI commands to be sent and change the configuration in a way that affects or disables out-of-band communication.</p> <p><u>Measurable primitives:</u></p> <ol style="list-style-type: none"> DCMI out-of-band communication operates without requiring an in-band driver. DCMI out-of-band communication remains available after an OS 'panic' or 'blue screen'

	<p>event.</p> <p>3. <u>DMCI out-of-band communication is unaffected by aborting the in-band driver process.</u></p>
RR.080	<p>Incomplete in-band DCMI message transfers shall not hang the in-band interface.</p> <p>Normal in-band messaging operation shall be able to be restored by initiating a new in-band message transfer from the beginning. The new in-band message does not need to be the same message as the previous transaction.</p> <p><u>Measurable primitives:</u></p> <p>1. Ability to recover from an incomplete in-band message transfer.</p>

A.1 Suggested Reliability Statements for IPDCs

The following statements summarize the reliability guidelines in Section 3 in a format suitable for IPDCs to generate a RFP for their own requirements. It is anticipated that each IPDC will work with their OEMs to bring out the level of reliability required and how to enforce the reliability in the products.

High availability of a feature is calculated over a period of time and the acceptable downtime of the feature. For example, a 99% High-Availability for a 3 year period, will constitute a downtime of approximately 11 days or equivalent in time units.

1. Manageability component of the Server will not cause any traffic outside the permitted network protocols in the IPDC's network. The network bandwidth usage of Server Manageability component will not exceed a measurable and acceptable IPDC-OEM specified unit of the available network bandwidth.
2. Manageability component of the Server will not impact the network behavior of the IPDC application including the underlying operating system running on the processors, which includes any manageability level network traffic and protocol handshaking such as DHCP and link negotiations.
3. Manageability component of the Server will provide a measurable and acceptable IPDC-OEM specified unit of High-Availability
 - a. For chassis power on/off/reset through out-of-band from IPDC remote agents through the sideband LAN connections and chassis power off/reset through in-band from host application.
 - b. For persistent out-of-band using the sideband LAN connection service to the IPDC remote agents across chassis power operations such as power on/off/reset.
4. Manageability component of the Server will have less than a measurable and acceptable IPDC-OEM specified unit (character distortion per # of characters transmitted) distortion in user experience on the Console redirection for both BIOS and OS using Serial Over LAN capability.
5. Manageability component of the Server will provide high availability at a measurable and acceptable IPDC-OEM specified unit for the IPDC remote agents communicating through sideband LAN connections, during abnormal network surges in the IPDC network environment due to other IPDC components outside the influence of the Server.

[The IPDC remote-agents will specify the acceptable blackout periods or network performance impact during unexpected network traffic surge greater than specified level of switch traffic.]
6. Server host based in-band communication to the Manageability component will not impact the sideband LAN connection service to the IPDC remote agents, which includes any host based monitoring and control but excludes any provisioning expectations.
7. Manageability component of the Server will provide high availability at a measurable and acceptable IPDC-OEM specified unit for host based in-band communication and also will provide communications restoration mechanisms built in to manage any broken connections between Host and Manageability component.