Link Tuning and Fabric Consistency Checks in PowerEdge M1000e Chassis Management Controller

This technical brief explains link tuning and fabric consistency check between I/O modules and modular server mezzanine cards.

Dell PG
February 2014

Author(s)

Aftab Hassan,
Saritha KG,
Rama Bisa

A Dell Best Practices
Revisions

<table>
<thead>
<tr>
<th>Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feb 2014</td>
<td>Initial Release</td>
</tr>
</tbody>
</table>

THIS WHITE PAPER IS FOR INFORMATIONAL PURPOSES ONLY, AND MAY CONTAIN TYPOGRAPHICAL ERRORS AND TECHNICAL INACCURACIES. THE CONTENT IS PROVIDED AS IS, WITHOUT EXPRESS OR IMPLIED WARRANTIES OF ANY KIND.

© 2014 Dell Inc. All rights reserved. Reproduction of this material in any manner whatsoever without the express written permission of Dell Inc. is strictly forbidden. For more information, contact Dell.

PRODUCT WARRANTIES APPLICABLE TO THE DELL PRODUCTS DESCRIBED IN THIS DOCUMENT MAY BE FOUND AT: http://www.dell.com/learn/us/en/19/terms-of-sale-commercial-and-public-sector Performance of network reference architectures discussed in this document may vary with differing deployment conditions, network loads, and the like. Third party products may be included in reference architectures for the convenience of the reader. Inclusion of such third party products does not necessarily constitute Dell’s recommendation of those products. Please consult your Dell representative for additional information.

Trademarks used in this text:
Dell™, the Dell logo, Dell Boomi™, Dell Precision™, OptiPlex™, Latitude™, PowerEdge™, PowerVault™, PowerConnect™, OpenManage™, EqualLogic™, Compellent™, KACE™, FlexAddress™, Force10™ and Vostro™ are trademarks of Dell Inc. Other Dell trademarks may be used in this document. Cisco Nexus®, Cisco MDS®, Cisco NX-OS®, and other Cisco Catalyst® are registered trademarks of Cisco System Inc. EMC VNX®, and EMC Unisphere® are registered trademarks of EMC Corporation. Intel®, Pentium®, Xeon®, Core® and Celeron® are registered trademarks of Intel Corporation in the U.S. and other countries. AMD® is a registered trademark and AMD Opteron™, AMD Phenom™ and AMD Sempron™ are trademarks of Advanced Micro Devices, Inc. Microsoft®, Windows®, Windows Server®, Internet Explorer®, MS-DOS®, Windows Vista® and Active Directory® are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. Red Hat® and Red Hat® Enterprise Linux® are registered trademarks of Red Hat, Inc. in the United States and/or other countries. Novell® and SUSE® are registered trademarks of Novell Inc. in the United States and other countries. Oracle® is a registered trademark of
# Table of contents

Executive Summary: ................................................................................................................................. 6
Introduction ................................................................................................................................................ 6
Combination of I/O Modules and Modular Servers .................................................................................. 6
Fabric Consistency Check ......................................................................................................................... 10
Trouble Shooting Fabric Mismatch .......................................................................................................... 14
Link Tuning ................................................................................................................................................ 16
Trouble Shooting Link Tuning error or warning ...................................................................................... 16
References ............................................................................................................................................... 18
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig 1</td>
<td>PowerEdge M1000e Chassis Fabrics and Capabilities</td>
<td>07</td>
</tr>
<tr>
<td>Fig 2</td>
<td>I/O Fabric Architecture with Quarter Height Blades</td>
<td>08</td>
</tr>
<tr>
<td>Fig 3</td>
<td>I/O Fabric Architecture with Quarter Height Blades</td>
<td>09</td>
</tr>
<tr>
<td>Fig 4</td>
<td>I/O Fabric Architecture with Quarter Height Blades</td>
<td>11</td>
</tr>
<tr>
<td>Fig 5</td>
<td>PowerEdge M1000e InfiniBand I/O Interoperability</td>
<td>12</td>
</tr>
<tr>
<td>Fig 6</td>
<td>PowerEdge M1000e Fibre Channel I/O Interoperability</td>
<td>13</td>
</tr>
<tr>
<td>Fig 7</td>
<td>PowerEdge M1000e 10Gb Ethernet I/O Interoperability</td>
<td>14</td>
</tr>
<tr>
<td>Fig 8</td>
<td>PowerEdge M1000e 1Gb Ethernet I/O Interoperability</td>
<td>15</td>
</tr>
<tr>
<td>Fig 9</td>
<td>Racadm getdcinfo output</td>
<td>16</td>
</tr>
<tr>
<td>Fig 10</td>
<td>Health page of Chassis Management Controller 4.5 GUI</td>
<td>16</td>
</tr>
<tr>
<td>Fig 11</td>
<td>CMC Health Page showing link tuning warning</td>
<td>17</td>
</tr>
<tr>
<td>Fig 12</td>
<td>Device option ROM on mezzanine card C failed to support Link Tuning or FlexAddress</td>
<td>18</td>
</tr>
</tbody>
</table>
Executive Summary:

This document explains the following:

- Combination of I/O modules and modular servers.
- Fabric Consistency Checks
- Link Tuning
- Troubleshooting link tuning errors or warnings and Fabric mismatch conditions

Introduction

The Dell PowerEdge M1000e Chassis is enabled to contain up to six I/O modules, which are classified into 3 separate fabric types – Ethernet, Fibre Channel, Infiniband. The fabrics can be used independently of each other and are user configurable.

As these fabrics are user configurable i.e. depending on the speed requirement of the user, there is a need to:

- Tune the speed of operation of the I/O ports, which is called link tuning.
- Check the fabric consistency between the I/O module and the server fabric present on the chassis.

These operations are explained in detail in this white paper. Also, mechanisms to troubleshoot certain scenarios related to link tuning and fabric consistency are also explained.

Combination of I/O Modules and Modular Servers

Based on the speed requirements of the user, different combinations of I/O modules and mezzanine cards (HBA) can be selected. The following sections provide more details on the mezzanine cards and I/O module combinations that can be used.

PowerEdge M1000e chassis supports 3 types of fabrics:

- Ethernet
- Fibre Channel
- Infiniband
A fabric is defined as a method of encoding, transporting, and synchronizing data between devices. Examples of fabrics are Gigabit Ethernet (GE), Fibre Channel (FC) or InfiniBand (IB). Fabrics are carried inside the PowerEdge M1000e system, between server module and I/O Modules through the midplane. They are also carried to the outside world through the physical copper or optical interfaces on the I/O modules.

Mezzanine cards that are placed on the server are designed to connect via 8 lane PCIe to the server module’s chipset. For PCIe Gen1 this provides up to 16 Gbps of data bandwidth per mezzanine card. Mezzanine cards may have either one dual port ASIC with 4 or 8 lane PCIe interfaces or dual ASICs, each with 4 lane PCIe interfaces.

Fabric A supports only Ethernet type, while Fabric B and Fabric C support all the 3 types of interfaces and can be independently configured by the user.

---

![PowerEdge M1000e Chassis Fabrics and Capabilities](image)

**Fig 1. PowerEdge M1000e Chassis Fabrics and Capabilities**
I/O Fabric Architecture for Half-Height Blades

Fig 2. I/O Fabric Architecture with Half-Height Blades
I/O Fabric Architecture for Full-Height Blades

Fig 3. I/O Fabric Architecture with Full-Height Blades
Fabric Consistency Check

As Fabric B and Fabric C can be flexibly configured by the user, PowerEdge M1000e checks the fabric consistency. Fabric Consistency Checks (FCC) allow compatible and prevent incompatible devices to be powered up on the same fabric. Compatibility of given I/O device with other fabric types has been specified in the below table. After the Fabric is designated by the first device ("device" refers to server or I/O Modules) to boot (typically the IOM). Once designated, any mismatched device does not receive power.
### PowerEdge M1000e InfiniBand I/O Interoperability

<table>
<thead>
<tr>
<th>Mezzanine Cards</th>
<th>M2401G Mellanox DDR</th>
<th>M3601Q Mellanox QDR</th>
<th>M4001Q Mellanox DDR</th>
<th>M4001T Mellanox FDR10</th>
<th>M4001F Mellanox FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mellanox DDR Connect-X</td>
<td>✓ DDR</td>
<td>✓ DDR</td>
<td>Not Supported</td>
<td>Not Supported</td>
<td>Not Supported</td>
</tr>
<tr>
<td>Mellanox QDR Connect-X2</td>
<td>✓ DDR</td>
<td>✓ QDR</td>
<td>✓ QDR</td>
<td>✓ QDR</td>
<td>✓ QDR</td>
</tr>
<tr>
<td>Mellanox QDR Connect-X3</td>
<td>Not Supported</td>
<td>✓ QDR</td>
<td>✓ QDR</td>
<td>✓ QDR*</td>
<td>✓ QDR</td>
</tr>
<tr>
<td>Mellanox FDR10 Connect-X</td>
<td>Not Supported</td>
<td>✓ QDR</td>
<td>✓ QDR</td>
<td>✓ FDR10</td>
<td>✓ FDR10</td>
</tr>
<tr>
<td>Mellanox FDR Connect-X3</td>
<td>Not Supported</td>
<td>✓ QDR</td>
<td>✓ QDR</td>
<td>✓ FDR10</td>
<td>✓ FDR**</td>
</tr>
</tbody>
</table>

✓ QDR*: Requires switch firmware version “fw-sx_0JF9G6_9.1.0562” and adapter version “fw-ConnectX3_rel_0065Y0_B1_2_11_0560_Flexboot-3_4_000.bin”. Customers with this combination can call Dell Support if they would like it to function on the M420 or M820.

✓ FDR**: Not supported with original mid-plane (L0).

---

Fig 5. PowerEdge M1000e InfiniBand I/O Interoperability
### PowerEdge M1000e Fibre Channel I/O Interoperability

<table>
<thead>
<tr>
<th>Mezzanine Cards</th>
<th>I/O Modules</th>
<th>FC4 Passthrough</th>
<th>M4424 Brocade FC4</th>
<th>FC8 Passthrough</th>
<th>Dell 8/4Gbps FC SAN Module</th>
<th>M5424 Brocade FC8</th>
<th>M6506 Brocade FC16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emulex FC4</td>
<td>✓ FC4</td>
<td>✓ FC4</td>
<td>✓ FC4</td>
<td>✓ FC4</td>
<td>✓ FC4</td>
<td>Not compatible</td>
<td></td>
</tr>
<tr>
<td>QLogic FC4</td>
<td>✓ FC4</td>
<td>✓ FC4</td>
<td>✓ FC4</td>
<td>✓ FC4</td>
<td>✓ FC4</td>
<td>Not compatible</td>
<td></td>
</tr>
<tr>
<td>Emulex LPe1205-M FC8 (for 11G and 12G servers)</td>
<td>✓ FC4</td>
<td>✓ FC4</td>
<td>✓ FC8</td>
<td>✓ FC8</td>
<td>✓ FC8</td>
<td>✓ FC8</td>
<td>✓ FC8</td>
</tr>
<tr>
<td>QLogic QME2572 FC8 (for 11G and 12G servers)</td>
<td>✓ FC4</td>
<td>✓ FC4</td>
<td>✓ FC8</td>
<td>✓ FC8</td>
<td>✓ FC8</td>
<td>✓ FC8</td>
<td>✓ FC8</td>
</tr>
<tr>
<td>Emulex LPM16002 FC16</td>
<td>Not compatible</td>
<td>Not compatible</td>
<td>✓ FC8</td>
<td>✓ FC8</td>
<td>✓ FC8</td>
<td>✓ FC8</td>
<td>✓ FC16*</td>
</tr>
<tr>
<td>QLogic QME2662 FC16</td>
<td>Not compatible</td>
<td>Not compatible</td>
<td>✓ FC8</td>
<td>✓ FC8</td>
<td>✓ FC8</td>
<td>✓ FC8</td>
<td>✓ FC16*</td>
</tr>
</tbody>
</table>

✓ FC16*: 16Gbps speeds require enhanced midplane (1.1). Auto-negotiates to FC8 with original mid-plane (1.0)

Fig 6. PowerEdge M1000e Fibre Channel I/O Interoperability
### PowerEdge M1000e 10Gb Ethernet I/O Interoperability

<table>
<thead>
<tr>
<th>Adapters</th>
<th>MXL</th>
<th>PowerEdge M I/O Aggregator</th>
<th>M8024-k</th>
<th>M8024</th>
<th>M8428-k</th>
<th>10Gb Pass-Through (original model)</th>
<th>10Gb Pass-Through II</th>
<th>10Gb Pass-Through III</th>
<th>B22DILL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broadcom 57710 Mezz</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
</tr>
<tr>
<td>Broadcom 57711 Mezz</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
</tr>
<tr>
<td>Emulex OCm10102-f-m Mezz</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
</tr>
<tr>
<td>QLogic QME8142 Mezz</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
</tr>
<tr>
<td>Intel X520 Mezz</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
</tr>
<tr>
<td>Intel X520-k/k Mezz (for 11G)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>QLogic QME8242-k Mezz</td>
<td>✓*</td>
<td>✓*</td>
<td>✓*</td>
<td>✓*</td>
<td>✓</td>
<td>✓</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>✓*</td>
</tr>
<tr>
<td>Brocade BR1741M-k Mezz</td>
<td>✓*</td>
<td>✓*</td>
<td>✓*</td>
<td>✓*</td>
<td>✓</td>
<td>✓</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>✓*</td>
</tr>
<tr>
<td>Broadcom 57712-k NDC</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>✓</td>
</tr>
<tr>
<td>Broadcom 57810-k NDC</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>N/A</td>
<td>N/A</td>
<td>✓</td>
</tr>
<tr>
<td>Intel X520-k NDC</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>✓</td>
</tr>
<tr>
<td>QLogic QMD8262-k NDC</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>N/A</td>
<td>N/A</td>
<td>✓</td>
</tr>
<tr>
<td>Broadcom 57810-k Mezz</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>N/A</td>
<td>N/A</td>
<td>✓</td>
</tr>
<tr>
<td>Intel X520-k Mezz (for 12G)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>N/A</td>
<td>N/A</td>
<td>✓</td>
</tr>
<tr>
<td>QLogic QME8262-k Mezz</td>
<td>✓*</td>
<td>✓*</td>
<td>✓*</td>
<td>✓*</td>
<td>✓</td>
<td>✓</td>
<td>Not Compatible</td>
<td>Not Compatible</td>
<td>✓*</td>
</tr>
</tbody>
</table>

10GbE on fabric ‘A’ with original mid-plane (1.0) will shift down to 1Gb. Note: fabrics B & C remain 10Gb with original mid-plane (1.0)

N/A: This combination is not possible
Not Compatible: This combination will not link
✓*: In Fabric ‘A’ with original mid-plane (1.0), this combination will not link

Fig 7. PowerEdge M1000e 10Gb Ethernet I/O Interoperability
**Troubleshooting Fabric Mismatch**

If there is a fabric consistency mismatch condition, it could prevent blade server power up. To detect fabric mismatch, use the following RACADM CLI command:

```
racadm getdcinfo
```

The DC1 State and DC2 State fields indicate if a fabric mismatch has occurred. The values “OK” or “N/A” indicate that there is no fabric mismatch while the value “Invalid” indicates a fabric mismatch condition.

In the example below, server-9 (extension slot of server-1) has mismatched fabric with switch-3.
On Chassis Management Controller 4.5 GUI, this is reported on the health page (shown below):
Link Tuning

The information required to tune electrical parameters such as drive strength and pre-emphasis on the SERDES (Serializer/Deserializer) interface in IOMs for certain mezzanine cards is contained in a configuration file maintained by the Dell PowerEdge M1000e Chassis Management Controller.

Whenever a modular server is inserted into the chassis, the Dell PowerEdge M1000e Chassis Management Controller checks if these electrical parameters are required to be tuned with the existing I/O modules present on it, to achieve the correct speed of operation. This check is also performed whenever an IOM is inserted into the chassis, to ensure right speed of operation of the IOM ports. The decision to perform link tuning depends on the combination of I/O module and the mezzanine card present on the blade server.

The interfaces that use link tuning are:

- Fibre channel
- 10G XAUI

Trouble Shooting Link Tuning error or warning

Use Case 1: “A link tuning failure detected on I/O module C1”.

This message appears on the CMC Chassis Health page and in the CMC logs.
To detect the cause of the issue, when a link tuning warning is displayed, use the following RACADM CLI command:

```
racadm gettracelog
```

The racadm output indicates which server and mezzanine is causing the issue.

Some excerpts from `racadm gettracelog` output below:

```
Jan  3 02:39:10 CMC-JV1J62S iomd[942]: link tuning: unable to tune switch-6 to server-12

Resolution Suggested: This indicates which combination of server-switch link tuning has failed. The reason for failure of link tuning could be because the mezzanine card used on the server is not allowed with the switch, the interoperability tables shown can be referred to check the combinations allowed.

Jan  3 03:47:11 CMC-JV1J62S bmd[940]: linkt_search: blade(114) information not found

Resolution Suggested: This indicates the presence of unrecognized hardware. In other words, the link tuning table that is part of the Chassis Management Controller firmware does not support the modular server with form factor 114. So it indicates that a Chassis Management Controller firmware upgrade is required to support this modular server.

Jan  3 03:47:10 CMC-JV1J62S bmd[940]: linkt_search: IOM(M8428-k 10GbEE-FCoE/8Gb FC SW ) info not found.

Resolution Suggested: Confirm that this I/O Module is allowed to be used with the mezzanine card from the interoperatability table in Fig. If it is not compatible replace the I/O Module or the mezzanine with a compatible one.

Use Case 2: Device option ROM on mezzanine card C failed to support Link Tuning or FlexAddress.

The following message is displayed in the CMC Chassis health page

```
Critical Alerts

Server-5  Device option ROM on mezzanine card B1 failed to support Link Tuning or FlexAddress.
```

Fig 12. Device option ROM on mezzanine card C failed to support Link Tuning or FlexAddress.

In the iDRAC logs, the message is displayed as follows as:
System Software event: Link Tuning sensor: device option ROM failed to support link tuning or flex address (Mezz C) was asserted

To resolve this issue:

1. Re-flash the Host Bus Adaptor (HBA) and correct the NOVRAM settings.
2. Record the serial number and WWN directly from the label on the HBA.
3. Download the current BIOS and flash utilities and run the update.bat batch file. If prompted for the serial number, enter it.
4. Allow the flash utility to run, make note of any error messages encountered as these may help in further analysis.
5. Reboot the server and observe if any symptoms have changed or are corrected.

References

1. You can download Chassis Management Controller firmware version 4.5 from this link.
2. Dell PowerEdge M1000e Chassis Management Controller version 4.5 User's Guide on Dell.com
3. For additional info on CMC, visit Chassis Management Controllers on www.DellTechCenter.com/CMC.
4. I/O Connectivity Options for the Dell PowerEdge M1000e Blade Enclosure, PowerEdge M-Series Blades I/O Guide