## VadaTech MicroTCA MCH

# Failover Management

March 12, 2009
Version 1.1

vadatech inc
THE POWER OF VISION

## Notice

While reasonable efforts have been made to assure the accuracy of this document, VadaTech, Inc. assumes no liability resulting from any omissions in this document or from the use of the information obtained herein. VadaTech reserves the right to revise this document and to make changes periodically and the content hereof without obligation of VadaTech to notify any person of such revision or changes.

Electronic versions of this material may be read online, downloaded for personal use, or referenced in another document as a URL to the VadaTech Incorporated Web site. The text itself may not be published commercially in print or electronic form, edited, translated, or otherwise altered without the permission of VadaTech, Inc.

It is possible that this publication may contain reference to or information about VadaTech products (machines and programs), programming, or services that are not available in your country. Such references or information must not be construed to mean that VadaTech intends to announce such products, programming, or services in your country.

# Revision History

| Doc Rev | Description of Change | Revision Date |
|---|---|---|
| 1.0 | Document Created | 02/23/2009 |
| 1.1 | Updated for UTC001 Release 1.3.3 | 03/12/2009 |

# Table of Contents

# Figures

# Tables

# 1 Overview

This document provides a summary of the VadaTech MicroTCA MCH Failover Management.

## 1.1 References

- Intelligent Platform Management Interface Specification Second Generation v2.0
- PICMG® 3.0 Revision 3.0 AdvancedTCA® Base Specification
- PICMG® Specification MTCA.0 R1.0 Base Specification
- VadaTech MicroTCA Shelf Manager Command Line Interface Reference Manual
- VadaTech MicroTCA Carrier Manager Command Line Interface Reference Manual
- VadaTech MicroTCA Shelf Manager SNMP Interface
- VadaTech MicroTCA Carrier Manager SNMP Interface

## 1.2 Acronyms Used in this Document

| Acronym | Description |
|---|---|
| IPMB | Intelligent Platform Management Bus |
| MCH | MicroTCA Carrier Hub |
| SATA | Serial ATA |
| SRIO | Serial Rapid I/O |
| XAUI | 10 Gigabit Attachment Unit Interface |

Table 1: Acronyms

# 2 MCH Redundancy

The VadaTech MicroTCA MCH supports redundancy and failover on a Chassis configured to have two MCHs. When the chassis is first turned on, the MCHs will auto-negotiate to determine which one of the two is healthier. A voting policy ensures one will become active. The active MCH will then host the Carrier Manager and, if the UTC001 is configured to run as a Shelf Manager, the Shelf Manager. The backplane failover interconnections provide the ability to monitor the health status of the controllers and subsequent failover when one of the MCHs becomes unhealthy. When there are two MCHs and the active MCH becomes unhealthy, a failover occurs only if the standby MCH is healthy and is able to take over Carrier and Shelf Management responsibilities. The active MCH will always synchronize its status and data with the passive MCH.

## 2.1 Active LED

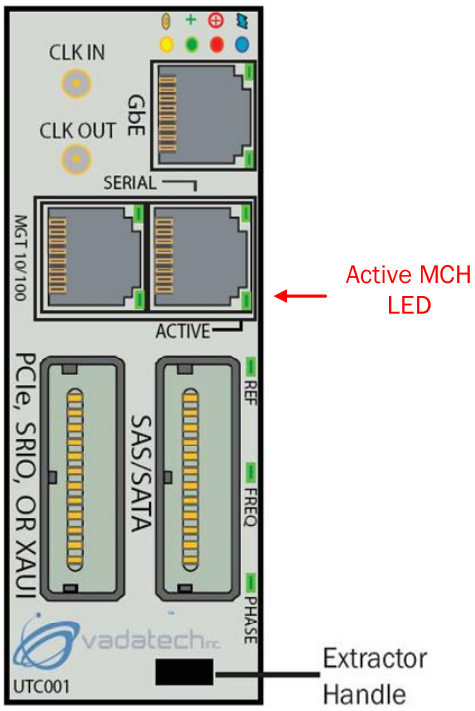


Figure 1: Location of the Active LED

The Active LED on the front panel will turn on indicating the MCH that is currently active between the two redundant MCHs.

## 2.2 MCH Health Monitor

All management applications on the active MCH are monitored by a central heart beat monitoring daemon. If a management application becomes unhealthy, the heartbeat monitor will attempt to restore the health by restarting the unhealthy application. If such an action does not recover the health status of the application, the heartbeat monitor will attempt a failover to the standby MCH. If a standby MCH is not present, then the heartbeat monitor will trigger the watchdog timer to restart the MCH.

## 2.3 Auto-Failover

The MCH Controller core contains an auto-failover negotiation feature that ensures that an MCH failover occurs instantaneously, without any software delay. This is important, since only one of the MCHs may provide the bus pull-ups at a time (Active-Standby).

## 2.4 Synchronization

The active MCH will synchronize all critical files and inventory data to the standby MCH. This is done using the available failover communications interfaces. The redundant failover communications interfaces are discussed in detail in Section 2.5: Failover Synchronization Transport Interface. All the standby files are overwritten by the active MCH if the active MCH's files have been updated following the previous synchronization.

During the initial startup, copies of all the active MCH files are transferred to the standby MCH for synchronization. This ensures that both MCHs are at a known state. If a new MCH is inserted into the Chassis, the active MCH will initiate another data synchronization. Date and time are periodically synchronized between the active and standby MCHs.

The following table describes the functions and files that are associated with failovers, and to which platform(s) they apply:

| File/Function | Shelf | Carrier | Interface | Description |
|---|---|---|---|---|
| Failover IP Address | | | IPMB | Failover IP addresses |
| SEL | ✓ | ✓ | Ethernet | System Event Log Files |
| SDR Repository | ✓ | ✓ | Ethernet | Sensor Data Repository |
| PEF | ✓ | ✓ | Ethernet | Platform Event Filter configuration |
| LAN Configuration | ✓ | ✓ | Ethernet | LAN configuration |
| startup/vtipmi.conf | N/A | N/A | Ethernet | VT IPMI Application configuration |
| UTCC/repository/fru_f0_1 | ✓ | N/A | Ethernet | Shelf FRU-0 Inventory Storage |
| UTCC/repository/fru_d1_60 | N/A | ✓ | Ethernet | Carrier FRU-0 Inventory Storage |
| UTCC/respository/fru_c2_1 | N/A | ✓ | Ethernet | Carrier FRU Inventory Storage |
| UTCC/repository/fru_f2_60 | ✓ | N/A | Ethernet | Shelf FRU Inventory Storage |
| /etc/passwd | N/A | N/A | Ethernet | System password file |
| /etc/shadow | N/A | N/A | Ethernet | System password file |
| SNMP/etc/snmpd.conf | N/A | N/A | Ethernet | SNMP agent configuration file |
| HPI/etc/hpid.conf | N/A | N/A | Ethernet | HPI configuration file |
| /tmp/UTCC/MCState.log | ✓ | N/A | Ethernet | Carrier Manager log file |
| /tmp/UTCSH/MCState.log | N/A | ✓ | Ethernet | Shelf Manager log file |

Table 2: MCH Failover files

## 2.5 Failover Synchronization Transport Interface

File transfer between the active and standby MCH occurs over the Ethernet. The standby MCH is able to communicate with the active MCH if a working Ethernet connection exists between the two. A connection can be established over one of the following redundant interfaces:

- backplane, via the Gigabit Ethernet switch on each of the MCHs
- an external switch from the front Gigabit switch Ethernet port of the MCH pair.
- an external switch from the front management Ethernet port of the MCH pair
- a crossover cable connecting the two front Gigabit switch Ethernet ports of the MCH pair
- a crossover cable connection between the two front management Ethernet ports of the MCH pair

The interfaces are completely redundant. Both MCHs will initiate a file transfer and synchronization on any of the above connections that can successfully establish communication. However, the default interface that each MCH will use is the backplane, via the Ethernet switch.

## 2.6 Failover Initialization

Upon initial power cycle, both of the MCHs will auto-negotiate via the backplane failover interconnections to become the active MCH. The active and the standby MCH will attempt to contact each other via IPMB to exchange failover IP addresses. This is a required step in

order for the active MCH to connect with the standby MCH. Once connected to the standby MCH via Ethernet, data and file synchronization can begin. During this initialization process, a failover may occur and the active snapshot may be lost if the initial synchronization was not completed.

## 2.7 Redundancy Sensor

The MCH Carrier Manager Sensor Data Repository includes a "Redundancy Sensor". This discrete sensor supports the following offsets:

| Sensor Type | Sensor Type Code | Sensor Specific Offset | Event |
|---|---|---|---|
| MCH Redundancy | 0xd8 | 00h<br>01h<br>03h | Fully Redundant<br>Redundancy Lost<br>Non Redundant (Resource not present) |

Table 3: Redundancy Sensor Offsets

On a fully redundant system, with both MCHs at a healthy state, this sensor will report with a "Fully Redundant" sensor event. If one of the MCHs becomes unhealthy, the now active MCH will send out a "Redundancy Lost" event. If the standby or the active MCH is removed, or the standby MCH is not present, then the sensor will report a "Non Redundant" event.

If there is a communication problem with the IPMB or the failover synchronization transport interface, then the sensor will report "Redundancy Lost" event until the communication is restored.

The status of the redundancy sensor can be read using the Shelf or Carrier Manager CLI:

```
# get_sensor_reading –a 0x82 –f 0 –t 0xd8
```

The above command is issued using the Shelf CLI to query the sensor reading for the sensor type 0xd8 at Carrier address 0x82. Refer to the VadaTech MicroTCA Carrier CLI Interferface Reference Manual and VadaTech MicroTCA Shelf CLI Interface Reference Manual for more information on the **`get_sensor_reading`** command.

### 2.7.1 System Event Log

The redundancy sensor will log an event to the Carrier Manager and Shelf Manager System Event Log (SEL) with the current status of the MCH redundancy.

The SEL records can be retrieved from the destination SEL by using the Shelf or Carrier Manager CLI:

```
# get_sel –a 0x82 –s 0xd8
```

The above command is issued using the Shelf CLI to query all the SEL records with sensor type 0xd8 from the System Event Log at the Carrier address 0x82. Refer to the VadaTech MicroTCA Carrier CLI Interface Reference Manual and VadaTech MicroTCA Shelf CLI Interface Reference Manual for more information on the **get_sel** command.

## 2.8 SNMP Trap

When SNMP traps are enabled, the redundancy sensor events are sent out via an SNMP trap.

## 2.9 Forced Failover

Failover from the active to standby MCH can be initiated by any of the following scenarios:

- use the Carrier Manager CLI or Web Interface to issue a failover command. This will result in a temporary redundancy lost reported by the redundancy sensor

  The following CLI Command can be used to initiate a failover:

  ```
  # failover
  ```

  A manual failover can be initiated from both the active and standby MCHs. A failover will only occur if the standby MCH is at least as healthy as the active MCH. Once the command executes, the former standby MCH immediately becomes the active MCH

- remove the active MCH from the chassis. Note, only opening the handle of the MCH will not result in a MCH failover

- reboot the active MCH

## 2.10 Shelf/Carrier IP Connection

The active MCH manages the 'out-of-band' interfaces to the Shelf and Carrier Manager. There is a dedicated out-of-band IP address to the Shelf and Carrier Manager. The out-of-band interfaces can be used to access the Shelf and Carrier Manager Information once a session has been established. When a failover occurs, the control of the out-of-band interfaces for the Shelf and Carrier Managers will change from the active MCH to the standby MCH. An active user session to the management interface may become invalid

when a failover occurs, and a new session will need to be re-established. The Shelf Manager and Carrier Manager IP connection addresses to the out-of-band interface will remain the same, regardless of the MCH hosting the Shelf Manager and Carrier Manager. For more information on the out-of-band interface to an MCH, please refer to the VadaTech MicroTCA Getting Started Guide.

### 2.10.1 MCH Generated Gratuitous ARPs

The active MCH will periodically send out Gratuitous ARP packets. During a failover, control of the out-of-band interfaces for the Shelf and Carrier Managers will change from the active MCH to the standby MCH. The Gratuitous ARP packet will force all other hosts on the network to update their ARP caches with the MCH's current out-of-band IP address and MAC address.