Using SGI® MEMlog™ Software

This document describes the SGI MEMlog (Memory Error Manager and Logger) software, which provides advanced handling of memory errors in all SGI servers.

Contents

Overview

MEMlog software is a privileged application that monitors and logs corrected memory errors, as well as other correctable or recoverable errors, including L3 cache errors and link errors (does not include PCI). MEMlog generates a detailed description of the DIMM configuration and maintains a separate history of corrected errors that are tracked by DIMM serial number for further analysis. If the page retirement kernel module is loaded, MEMlog passes corrected memory error addresses to the page retirement interface to remove pages from use when a DIMM failure is determined to be of a severe nature.

Corrected memory errors are initially recorded by the operating system kernel in the machine check exception handler. The kernel exception handler reads the machine check bank registers, and makes the register data available via the `/dev/mcelog` interface. MEMlog monitors this interface to collect and log corrected errors.

On older systems using the Intel 5100/5200/5300/5400 processor family, machine check exception does not support corrected memory errors. For these type of processors, MEMlog directly monitors the memory controller error registers for corrected errors.

MEMlog runs as a daemon (`memlogd`) that automatically starts when a system or node boots. The MEMlog RPM is included in the SGI Foundation Software (SFS) suite. When the MEMlog RPM is installed, it will `chkconfig` the `memlog` service on and install the `/etc/init.d/memlog` control script. (You can edit the `/etc/sysconfig/memlog` configuration file to modify the default MEMlog options.)

Error Logging

MEMlog software provides detailed memory error logging and tracks DIMMs by serial number. It reports corrected errors to the system log file and also logs additional information to a more detailed `memlog.log` file. This log file typically contains start up information, internal MEMlog error data, and unexpected results. This file is set up with the following defaults:

- By default, the `memlog.log` files is located in `/var/log/memlog`. You can change the location to a different path using the `-f` option.
- By default, the `memlog.log` file does not record routine corrected memory errors. You can change this by adding the `-m` option to log corrected error messages to `memlog.log` as well as the system log.

The system log file (and the `memlog.log` file, if configured with the `-m` option) includes comprehensive detail for each corrected memory error, including DIMM location, DRAM, rank, bank, row and column addresses, and failing bit(s). The bits are labeled using the DQ naming scheme instead of bit position, to follow JEDEC standard DIMM bit identification. Because of the potential to create very large log files, there is an option to stop logging a repeated error after a threshold error count.

The option to stop logging errors is tracked at the DIMM rank level. For example, if Rank 1 of a DIMM is producing errors and reaches the logging limit, subsequent unrelated errors from Rank 0 are logged. This occurs until Rank 0 error counts reach the logging limit. Regardless of whether the logging limit is reached or not, MEMlog continues to receive corrected error information from the kernel and will continually update the failure history of the DIMM.

Note: The default number of log entries per rank is 10.

The kernel continues to make system log file entries similar to the following:

```
Mar 27 12:53:56 n001 kernel: [522123.510343] [Hardware Error]: Machine
check events logged
```

Until the logging limit is reached, or the offending DRAM of the DIMM is mapped out using Device Tagging, MEMlog will match the kernel message with its detailed system log message.

```
Read ECC P1-DIMM2A Rank 0 Bank 5 Row 0x54d3 Col 0x248 DRAM U0 DQ3 Temp =
31.7C transient
```

In addition to error logging, a record of the failure history is maintained internally for each DIMM. The failure history is used to determine if a DIMM (more precisely, a DRAM) failure is transient due to alpha particle flips, row hammering, or other phenomena that are not considered stable circuit failures. The history analysis determines when a repeated failure is persistent, and if it is a single storage cell failure or a more widespread failure affecting multiple bits or multiple addresses. If a pattern of persistent failures emerge that affects multiple addresses, MEMlog issues a log entry notification that the DIMM has reached a critical failure threshold and should be replaced.

Configuring MEMlog

You can configure MEMlog by using the `memlogd` command-line options and by modifying the configuration values in the `/etc/sysconfig/memlog` file. The `memlogd` command-line options are described in the `memlogd` man page and in the `memlogd -h` output.

The following example demonstrates changing the logging limit reporting to a maximum of 40 errors per DRAM, sends corrected error log entries to `memlog.log` (in addition to the system log), and directs the error-logging output and log files to the `/var/log/memlog` directory:

```
MEMLOGD_EXTRA="-m -l 40 -f /var/log/memlog"
```

Note: A `-l 0` setting results in unlimited error logging.

For more information about `memlogd` operation and configuration options, refer to the `memlog` man page.

MEMlog Usage Examples

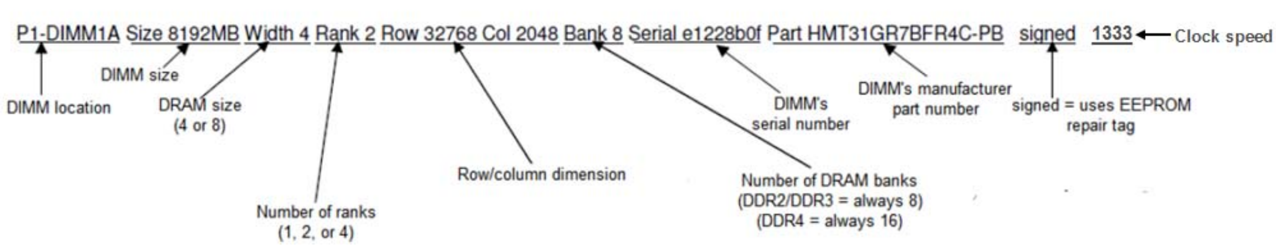
When the `memlogd` daemon is running, MEMlog can provide a listing of the DIMM inventory and any associated DIMM repair tags. Use the `memlogd -c` command to generate this list. This command sends a signal to the daemon, opens a named pipe, and gets the daemon's in-memory data of the inventory and repair tags.

For example:

```
harp34-smn:~ # memlogd -c
 P1-DimmA1 Size 4096MB Width 4 Rank 1 Row 15 Col 11 Bank 8 Serial 489B52C5 Part HMT351R7CFR4C-PB
 Tue Sep 18 14:14:12 2012 Rank 0 Dram U2 Bank 4 Row 0x2a0 Col 0x0 single  Temp = 23C hits 16
 P1-DimmB1 Size 4096MB Width 4 Rank 1 Row 15 Col 11 Bank 8 Serial 488B52C1 Part HMT351R7CFR4C-PB
 P1-DimmC1 Size 4096MB Width 4 Rank 1 Row 15 Col 11 Bank 8 Serial 484B52BD Part HMT351R7CFR4C-PB
 P1-DimmD1 Size 4096MB Width 4 Rank 1 Row 15 Col 11 Bank 8 Serial 488B52BE Part HMT351R7CFR4C-PB
 P2-DimmE1 Size 4096MB Width 4 Rank 1 Row 15 Col 11 Bank 8 Serial 48BB5304 Part HMT351R7CFR4C-PB
 P2-DimmF1 Size 4096MB Width 4 Rank 1 Row 15 Col 11 Bank 8 Serial 482B52FF Part HMT351R7CFR4C-PB
 P2-DimmG1 Size 4096MB Width 4 Rank 1 Row 15 Col 11 Bank 8 Serial 482B5305 Part HMT351R7CFR4C-PB
 P2-DimmH1 Size 4096MB Width 4 Rank 1 Row 15 Col 11 Bank 8 Serial 483B5305 Part HMT351R7CFR4C-PB
```

Note: This listing has a repair tag for the DIMM in location P1-DIMMA1; the repair tag follows the DIMM description. (Refer to See “MEMlog Repair Tags” on page 5 for more information about the repair tag.)

The DIMM inventory description includes the following components:



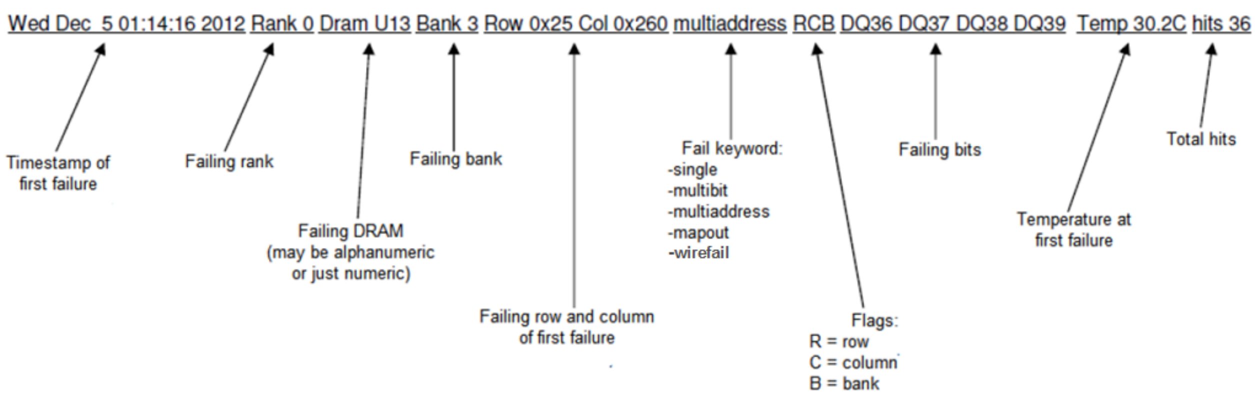
MEMlog Repair Tags

MEMlog generates repair tags for failing DIMMs. A DIMM's repair tag is generated when the first corrected error for a DRAM is reported. The failure data usually includes enough detail for MEMlog to decode the failing address to the DIMM, rank, bank, row, and column. If sufficient detail is available, MEMlog can also decode the DRAM value and DQ bits.

When additional corrected errors occur for this DRAM, MEMlog compares the new bank, row, and column values to the original failure to determine if the failure is taking place at the same DRAM location (listed as: `single` in the repair tag) or if the row, column, or bank vary, which implies a more complex and widespread failure (listed as: `multiaddress` in the repair tag).

Another possible keyword in the repair tag is `multibit`, which appears when the address of the failure is unchanged but multiple DQ bits appear as failed. On some processor types, the keyword `mapout` may appear, which indicates that the processor memory controller has determined a DRAM has failed and stopped logging corrected errors from that DRAM entirely.

Example repair tag:



The location of a repair tag depends on the hardware being used:

- For all SGI UV, ICE 8200/8400, and Altix XE systems as well as the older Woodcrest, Harpertown, Nehalem, and all AMD processor families, MEMlog stores the repair tag in the DIMM EEPROM. This is enables the repair tag to travel in the DIMM when the DIMM returns to the factory for repair.
- For SGI ICE X, ICE XA, and Rackable systems as well as Sandybridge, Ivybridge, and all future processor families, MEMlog stores the repair tags in separate files instead of the DIMM EEPROM because of technical restrictions imposed by these processor families.

  The file naming convention is:

  `.`<*DIMM_serial_number*>`.dat`, which is the hidden binary repair tag.
  <*DIMM_location*>`.tag`, which is the ASCII readable version of the file.

  These files are stored in the same directory as the `memlog.log` file (Default: `/var/log/memlog/memlog.log`).

  Be sure to print the ASCII version of the repair tag and return it with the failing DIMM to SGI Logistics. You should also copy and paste the repair tag into the support case.

MEMlog DIMM Replacement Criteria

MEMlog messages are very explicit about when to replace a DIMM. Transient errors are logged but are not cause for replacement. Persistent errors are tracked in the MEMlog repair tags; however, the type of failure determines if a DIMM should be replaced. A single storage cell failure on a single rank and DRAM of a DIMM does not require replacement. This is one bit at one address out of a vast range of addresses. Additionally, MEMlog attempts a page migration for persistent failures. This further reduces exposure. Broader failure modes such as multiaddress, wirefail, or non-single failure mode merit DIMM replacement, because this failure mode affects a notable span of addresses, or possibly the entire DIMM.

The error event that triggers failure mode analysis resulting in a DIMM replacement sends a message to the system log file stating explicitly that the DIMM must be replaced. In situations where system log files are lost, the DIMM repair tag contains the failure mode (single, multiaddress, or wirefail) to guide a DIMM replacement decision.

NB/SB Lane CRC Error Example

Note: This example only applies to SGI UV 1000 systems.

Each channel has 14 Northbound (DIMM->MCH) and 10 Southbound (MCH->DIMM) lanes between the FB-DIMMs and the MCH. When CRC errors occur, the MCH resets and retrains the lanes and then reissues the affected memory requests.

The first FB-DIMM in the channel typically caused the error:

```
Apr 4 11:27:41 cl1n001
MEMLOG: Northbound CRC on Read on Memory Branch 1
Channel 3: Packet data cffbc0141f90acd2 080000001
ECC: 6067 CRC: e38 SLOT: 4A 4B
```

Southbound errors are reported as “FBD Alert or Fast Reset Timeout” on the Intel 5000P and 5000X chipsets, and “Memory Write Error on First Attempt” on the Intel 5400 chipset.

Notes and Tips

- You can use the `memlog -V` command to determine the version of `memlog` that is running. This can be useful to determine if MEMlog problems are being caused by an older version of MEMlog being run on a newer motherboard.
- The following is an example of MEMlog generating an error message that is actually a correctable error on the QPI path, where no action is required:

  ```
  Jul 29 01:02:50 afispb05 MEMLOG[63701]: r001i23b07 socket 1 cpu 376 Rbox:
  Flit CRC error at input port at QPI port 4 to Skt 0 port 0
  ```

Document Revision History

001 - November 2015 - Original publication as SGI part number 007-6457-001. The content for this document was taken from the *memlog Reference Guide*, SGI part number 108-0594-003. This document contains information for the updated version of MEMlog software, its improved error reporting, and more specific DIMM replacement criteria.

[11.12.2015