OpenSM Log Management

OpenFabrics Software User Group Workshop

Hal Rosenstock
Mellanox Technologies

Agenda

- OpenSM Update
- Per Module Logging Feature

OpenSM Releases

- Releases nominally every 6-9 months
  - Independent of OFED
- FDR and FDR-10 support (OpenSM 3.3.11 – Aug 2011)
  - FDR (and EDR) are IBTA standards
  - FDR-10 is MLNX proprietary
- SRIOV support (OpenSM 3.3.14 – May 2012)
  - Additional GUIDs for virtual machines
    - Bug fixes beyond 3.3.14

Quick OpenSM Update

- OpenSM 3.3.17 – Feb 2014
  - Also included in OFED 3.12
- OpenSM 3.3.18 – July 2014
- Latest release is OpenSM 3.3.19 – December 2014
  - Also included in OFED 3.18 which is in process now
- ~25 commits past 3.3.19 right now
- Nightly regression tests against latest master
  - More than 400 test cases used

Additional MLNX OpenSM Features

- Event reporting scalability
- Bad nodes/hardware
- Heldback switches
- Multicast improvements
- Routing chains
- Credit-loop free UC and MC routing for updn/ftree/minhop
- Multithreaded updn/minhop/dor
- PQFT routing engine
  - For quasi and parallel ports generalized fat tree topologies
- Fast multicast routing
- Enhanced EDR support
- Connect-IB SRIOV support

OpenSM Logging Related Command Line Options

- -f, --log_file <file name>
  This option defines the log to be the given file. By default, the log goes to /var/log/opensm.log. For the log to go to standard output use -f stdout.
- -L, --log_limit <size in MB>
  This option defines maximal log file size in MB. When specified the log file will be truncated upon reaching this limit.
- -e, --erase_log_file
  This option will cause deletion of the log file (if it previously exists). By default, the log file is accumulative.

OpenSM Logging Related Config File Options

```
# Force flush of the log file after each log message
force_log_flush FALSE

# Log file to be used
log_file /var/log/opensm.log

# Limit the size of the log file in MB. If overrun, log is restarted
log_max_size 0

# If TRUE will accumulate the log over multiple OpenSM sessions
accum_log_file TRUE

# Per module logging configuration file
# Each line in config file contains <module_name><separator><log_flags>
# where module_name is file name including .c
# separator is either = , space, or tab
# log_flags is the same flags as used in the coarse/overall logging
per_module_logging_file /usr/local/etc/opensm/per-module-logging.conf
```

OpenSM Log Levels

- Overall log verbosity level
  - log_flags config file option
  - Related command line options
    - -D <value>
    - -V
    - -v, --verbose
    - -d, --debug <value>

-D <value>

- This option sets the log verbosity level. A flags field must follow the -D option. A bit set/clear in the flags enables/disables a specific log level as follows:

BIT	LOG LEVEL ENABLED
0x01	ERROR (error messages)
0x02	INFO (basic messages, low volume)
0x04	VERBOSE (interesting stuff, moderate volume)
0x08	DEBUG (diagnostic, high volume)
0x10	FUNCS (function entry/exit, very high volume)
0x20	FRAMES (dumps all SMP and GMP frames)
0x40	ROUTING (dump FDB routing information)
0x80	SYS (syslog at LOG_INFO level in addition to OpenSM logging)

-D <value>

- Without -D, OpenSM defaults to ERROR + INFO (0x3).
- Specifying -D 0 disables all messages.
- Specifying -D 0xFF enables all messages (see -V). High verbosity levels may require increasing the transaction timeout with the -t option.

Other Related Log Level Command Line Options

- -v, --verbose
  This option increases the log verbosity level. The -v option may be specified multiple times to further increase the verbosity level. See the -D option for more information about log verbosity.
- -V
  This option sets the maximum verbosity level and forces log flushing. The -V option is equivalent to -D 0xFF -d 2. See the -D option for more information about log verbosity.
- -d, --debug <value>
  -d2 - Force log flushing after each log message

Issues with OpenSM Logging

- Coarseness of log level
  - One level for all of OpenSM
  - Too many log messages as increase verbosity/log level
- Somewhat “cryptic” nature of messages logged

Per Module Logging (PML)

- Log level per “module”
  - Module is a source code file
- Introduced so can keep “overall” level low but dial up level in specific modules/files
  - Need to have idea of which modules/files to dial up
- PML can change on the “fly” with SIGHUP
- Added to upstream master git tree in June/July 2012
- Part of OpenSM 3.3.15 and beyond

Per Module Logging (PML)

- Enable via per_module_logging_file option in options file – set to PML config file name
  - Disable by setting per_module_logging_file to (null) in options file
- Per module logging config file format

  Set of lines with module name and logging level as follows:
  <module name><separator><logging level>
  where:
  <module name> is the file name including .c
  <separator> is either = , space, or tab
  <logging level> is the same levels as used in the coarse/overall logging
- Module names may vary between releases
  - 3.3.16 and beyond have all modules listed
  - 3.3.15 has one less module (no osm_congestion_control.c)

Per Module Logging (PML) Module Names Based on Latest Upstream Master

From opensm/osm_subnet.c:

```
static const char *module_name_str[] = {
        "main.c",
        "osm_console.c",
        "osm_console_io.c",
        "osm_db_files.c",
        "osm_db_pack.c",
        "osm_drop_mgr.c",
        "osm_dump.c",
        "osm_event_plugin.c",
        "osm_guid_info_rcv.c",
        "osm_guid_mgr.c",
        "osm_helper.c",
        "osm_inform.c",
        "osm_lid_mgr.c",
        "osm_lin_fwd_rcv.c",
        "osm_link_mgr.c",
        "osm_log.c",
        "osm_mad_pool.c",
        "osm_mcast_fwd_rcv.c",
        "osm_mcast_mgr.c",
        "osm_mcast_tbl.c",
        "osm_mcm_port.c",
```

Per Module Logging (PML) Module Names Based on Latest Upstream Master

"osm_mesh.c",
"osm_mlnx_ext_port_info_rcv.c",
"osm_mtree.c",
"osm_multicast.c",
"osm_node.c",
"osm_node_desc_rcv.c",
"osm_node_info_rcv.c",
"osm_opensm.c",
"osm_perfmgr.c",
"osm_perfmgr_db.c",
"osm_pkey.c",
"osm_pkey_mgr.c",
"osm_pkey_rcv.c",
"osm_port.c",
"osm_port_info_rcv.c",
"osm_prtn.c",
"osm_prtn_config.c",
"osm_qos.c",
"osm_qos_parser_l.c",
"osm_qos_parser_y.c",
"osm_qos_policy.c",
"osm_remote_sm.c",
"osm_req.c",
"osm_resp.c",
"osm_router.c",

Per Module Logging (PML) Module Names Based on Latest Upstream Master

"osm_sa.c",
"osm_sa_class_port_info.c",
"osm_sa_guidinfo_record.c",
"osm_sa_informinfo.c",
"osm_sa_lft_record.c",
"osm_sa_link_record.c",
"osm_sa_mad_ctrl.c",
"osm_sa_mcmember_record.c",
"osm_sa_mft_record.c",
"osm_sa_multipath_record.c",
"osm_sa_node_record.c",
"osm_sa_path_record.c",
"osm_sa_pkey_record.c",
"osm_sa_portinfo_record.c",
"osm_sa_service_record.c",
"osm_sa_slvl_record.c",
"osm_sa_sminfo_record.c",
"osm_sa_sw_info_record.c",
"osm_sa_vlarb_record.c",
"osm_service.c",
"osm_slvl_map_rcv.c",
"osm_sm.c",
"osm_sminfo_rcv.c",
"osm_sm_mad_ctrl.c",
"osm_sm_state_mgr.c",
"osm_state_mgr.c",

Per Module Logging (PML) Module Names Based on Latest Upstream Master

```
        "osm_subnet.c",
        "osm_sw_info_rcv.c",
        "osm_switch.c",
        "osm_torus.c",
        "osm_trap_rcv.c",
        "osm_ucast_cache.c",
        "osm_ucast_dnup.c",
        "osm_ucast_file.c",
        "osm_ucast_ftree.c",
        "osm_ucast_lash.c",
        "osm_ucast_mgr.c",
        "osm_ucast_updn.c",
        "osm_vendor_ibumad.c",
        "osm_vl15intf.c",
        "osm_vl_arb_rcv.c",
        "st.c",
        "osm_ucast_dfsssp.c",
        "osm_congestion_control.c",
        /* Add new module names here ... */
        /* FILE_ID define in those modules must be identical to index here */
        /* last FILE_ID is currently 89 */
};
```

OpenSM Log Messages Overview

- Format: date time [thread ID] log level
  Feb 19 12:40:45 897693 [91A48700] 0x01 ->
- ERR number if in message is unique
- Having OpenSM sources helps
  - Tracking error number in source module shows where generated and can read code and comments
- IBA spec knowledge is helpful
  - Primarily volume 1 IB management related chapters

“Cryptic” Log Messages

- Always looking to improve wording of log messages
  - Suggestions are welcome!
- Most common laments about messages are related to SA multicast and SM MAD timeouts/rejections

Thank You

OPENFABRICS ALLIANCE

OpenFabrics Software User Group Workshop