IST instability due to tMainTask failing to release CPU in high CPU utilization conditions

BULLETIN ID:	2010010309, Rev 1
PUBLISHED:	2010-10-07
STATUS:	Active
REGION:	APAC
	CALA
	EMEA
	GC
	NA
PRIORITY:	Major
TYPE:	Bulletin

Background:

In large campus networks with SMLT topologies where multicast routing protocols (such as PIM) have been provisioned and scaled to large amounts of multicast senders and receivers, it has been observed that high CPU utilization (sometimes combined with high CPU buffer utilization) leading to IST instability may occur during re-convergence of the multicast routing protocols after failures.

Analysis:

The IST instability manifested in this scenario is due to delays in IST protocol message processing. The IST messaging delay is attributable to an OS task "tMainTask" process not providing enough CPU time to process IST messages during reconvergence of the mullticast protocols due to high CPU and/or CPU buffer utilization.

The following output illustrates the footprint of the issue:

- CPU utilization snapshot taken during the issue:

*EAMC Core 2 Switch .4:5# show sys perf

```
            CpuUtil: 81% ? High CPU utilization
      SwitchFabricUtil: 0%
   OtherSwitchFabricUtil: 0%
          BufferUtil: 99% ? high CPU buffer utilization
           DramSize: 256 M
           DramUsed: 59 %
           DramFree: 107306 K
*EAMC Core 2 Switch .4:5#
```

- Log file extracts:

```
MLT Task=tMainTask CPU5 [01/18/10 19:52:35] MLT INFO SMLT 34 DOWN, vlanId 8192
 MLT Task=smltSlave CPU5 [01/18/10 19:55:10] MLT INFO smltProcessMsgs: Time (ms) taken to Receive
"SMLT_MSG_SMLT_DOWN" Pkt = 0,Process = 155792 ? protracted processing of IST/SMLT messages
 MLT Task=tMainTask CPU5 [01/18/10 19:55:10] MLT INFO SMLT 34 UP
 MLT Task=smltSlave CPU5 [01/18/10 19:57:45] MLT INFO smltProcessMsgs: Time (ms) taken to Receive
"SMLT_MSG_SMLT_UP" Pkt = 0,Process = 156295 ? protracted processing of IST/SMLT messages
```

 MLT Task=smltSlave CPU5 [01/18/10 19:57:46] MLT ERROR smltProcessMsgs: Problem while reading msg hdr from socket 15, rcvLen -1
 MLT Task=smltSlave CPU5 [01/18/10 19:57:46] MLT INFO smltSlave: socket error - closing socket: 15
 MLT Task=tMainTask CPU5 [01/18/10 19:57:47] MLT ERROR smltSendHelloMsgs: Failed to send Hello msg! Counter at 80545
  [01/18/10 19:57:48] The previous message repeated 1 time(s).
 MLT Task=tMainTask CPU5 [01/18/10 19:57:48] MLT INFO smltIstSessionDown: Socket error ? IST session going down
 MLT Task=tMainTask CPU5 [01/18/10 19:58:14] MLT INFO All the SMLTs are down

- MainTask task trace sample obtained during the issue manifesting itself:

Trace back 15 Duration 0x00208bcaa Sample 207
<snip>
count   8 next 00000000 hash 0x7fcb5ff4 budget 00000000 file - line 0 taskid 0x7b19ea8 (tMainTask) state 1 (interrupt) ? tMaintask busy processing multicast specific function calls
     018cc25c _ipmPepStrWriteToPid+0x38c
     018cbe78 ipmPepStrWriteToPid+0x38
     018cf244 ipmPepStrUpdateSmltPorts+0x434
     01810e00 igmpUpdateSmltPortInfo+0x110
     019ae67c smltProcSmltUpMsgBody+0x53c
     016553e8 syncEventSink+0x140
     01663bbc cppScheduleBody+0x85c
     01664558 cppMainTask+0x15c
     010bdd60 vxTaskEntry+0x68
count   7 next 00000000 hash 0xd3a4240 budget 00000000 file - line 0 taskid 0x7b19ea8 (tMainTask) state 1 (interrupt) ? tMaintask busy processing multicast specific function calls
     018cc08c _ipmPepStrWriteToPid+0x1bc
     018cbe78 ipmPepStrWriteToPid+0x38
     018cf244 ipmPepStrUpdateSmltPorts+0x434
     01810e00 igmpUpdateSmltPortInfo+0x110
     019ae67c smltProcSmltUpMsgBody+0x53c
     016553e8 syncEventSink+0x140
     01663bbc cppScheduleBody+0x85c
     01664558 cppMainTask+0x15c
     010bdd60 vxTaskEntry+0x68
*EAMC Core 2 Switch .4:5#

Recommendations:
Release 5.1.3.0 has been modified with changes that were originally introduced in release 7.0.0.0. These changes allow IST protocol messages to be processed even under high CPU utilization. This is achieved by checking to see if IST control messages are queued up (but not yet processed) before deciding that the IST session has timed out and needs to be brought down. Each line card recognizes and counts IST control messages when they arrive and before they are sent to the CP, and the IST message processing logic on the CP will check for outstanding IST control messages before deciding the IST needs to be brought down due to inactivity.

Required Actions:
Upgrade to GA release 5.1.3.0 or later.

Attachments:
There are no attachments for this bulletin

Products and Releases:
The information in this bulletin is intended to be used with the following products and associated releases:

PRODUCT	RELEASE
Ethernet Rtng Switch-Ethrnt Rtng Swt 8600-Ethernet Rtng Switch 8600	

To view the most recent version of this bulletin, access technical documentation, search our knowledge base, or to contact a Technical Support Representative, please visit Nortel Technical Support on the web at: http://support.nortel.com/. You may also sign up to receive automatic email alerts when new bulletins are published.

REFERENCE:
PRE-REQUIRED PATCH:
PATCH ID:
FIXED RELEASE: 5.1.3.0