

This morning we receive an email from the ILOM of a Sun Oracle X3-2L server: *The suspect component: /SYS/SP has defect.ilom.fs.full*

A full filesystem on an ILOM? Well I never... So I log into the ILOM to have a quick poke around, both on the CLI and web interface, but there's no directly accessible filesystem and shell to do any changes to the ILO 'filesystem' (without deep dropping into the ILOM's boot sector - but that's a write-up for another day) so I figure it's just some log file that is filling up and hopefully cleaning that up will do the trick. So I cleared the log files and rebooted the ILO. After it came back up all seemed ok - the alert auto-resolved, and all is well. Except I notice in the ILO audit log there's an entry for a login every minute or so. This could be why the log eventually filled up after we patched this ILOM about 4 months ago. At this stage I figured it could just be a little bug in the logging level or something so I pulled a new firmware down and installed it, hoping for a quick fix, and a free patch in the process. Of course, part of the ILO firmware patch is also a BIOS patch, so the server needs a reboot. Luckily this particular server is Well Looked After™ so a quick boot (well, quicker than an HP DL380g9, that's for sure) later the system is back up and running. All good.

Except no - the Infiniband interfaces, which are running IP over Infiniband, are showing as up and have link, but can't send or receive any network traffic.

So I play around with the TCP stack a little - validate configurations have not changed, bring bonded interface up, down, restart network, remove one interface from the bond, try the other, all with no luck. TCP won't go. Technically firmware is a last resort, especially with something that was working beforehand, but the support notes for the ILOM firmware update have planted it in my mind that any HBAs and HCAs on a system getting a BIOS update might need to be patched at the same time. While I'm still in a patchy mood, I verify the firmware levels as per the vendor website to make sure we're on the latest for the IB HCA. Then I started with the Infiniband diagnostic tools. ibdiagnet shows we're pretty much seeing all devices across the IB network. iblinkinfo shows plenty of active links, except the two on this particular problem server. ibswitches confirms switch connectivity.

ibstat is showing LinkUp on both ib0 and ib1, but I'm worried the interfaces are still in "Initializing" state. Also, the Base lid and SM lid still have value 0, which I'm sure should actually be an integer value. Logging into the web management interface on the Infiniband ILO shows nothing wrong, but confirms the missing lid. All the working interfaces have an lid assigned. I confirm this on another server.

I run ibdiagnet one more time and see an error that catches my eye: *Missing master SM in the discover fabric* Now I don't know everything about Infiniband, but I do know from my Exadata patching days that each Infiniband switch has an SM (Subnet Manager) program running, which presumably manages the devices on each Infiniband subnet (you can segment Infiniband switches and devices like is done with VLANs). I recalled from patching the Infiniband switch there were instructions to ensure there is always a SM running on one of the redundant switches, and to make sure there was a master with higher priority.

Long story short, logging into each Infiniband switch and ensuring there was one with a higher SM priority, and then disabling it and re-enabling SM and instantly a lid was assigned for each failing Infiniband interface and IPoIB kicked in and connectivity started working between the servers. Clearly the SM had been hung up and was back in action assigning addresses. It took a few more moments but the Infiniband cards then moved into an "Active" state.

But back to the ILOM filesystem filling up - the ILOM firmware update didn't seem to have reduced the amount of logging - but I was able to identify the source by paying more attention to the processes on the machine. I've been aware for a long time that kipmi0 runs hot on the CPU on this machine,

although with very low priority, but now the repeated ILOM logins suggested that an IPMI-interfacing tool was hanging around and perhaps the IPMI tool was logging in regularly to check for any alerts from the ILOM. On a hunch I shutdown a process I thought it could be, and I was right. *kipmi0* spawns from checks made by the Oracle Hardware Management Pack (hwmgmt). In essence, in this case, the hwmgmt process was the one filling up the logs on the ILOM. The ILOM should handle it better, and hopefully the newer firmware will, but time will tell if I end up asking the vendor for a fix (no newer version for this OS is currently available at this time).

Urgh, what a rabbit hole. I surely could have done that better, but I'm glad I inadvertently discovered the Subnet Manager process on the Infiniband network was failing. That could have led to a whole host of other nightmares when we get round to rebooting the Exadata servers.

On the management controller, type:

```
# setsmpriority priority
```

where priority is 0 (lowest) to 13 (highest). For example:

```
# setsmpriority 3
```

```
-----  
OpenSM 3.2.6_20090717
```

```
  Reading Cached Option File: /etc/opensm/opensm.conf
```

```
  Loading Cached Option:routing_engine = ftree
```

```
  Loading Cached Option:sm_priority = 13
```

```
  Loading Cached Option:sminfo_polling_timeout = 1000
```

```
  Loading Cached Option:polling_retry_number = 3
```

```
Command Line Arguments:
```

```
  Priority = 3
```

```
  Creating config file template '/tmp/osm.conf'.
```

```
  Log File: /var/log/opensm.log  
-----
```

```
#
```

Restart the Subnet Manager:

```
# disablesm
```

```
Stopping IB Subnet Manager.. [ OK ]
```

```
# enablesm
```

```
Starting IB Subnet Manager. [ OK ]
```

```
#
```

From:

<https://wiki.dewberry.co.za/> - Shaun's Wiki

Permanent link:

<https://wiki.dewberry.co.za/doku.php?id=infinibork>

Last update: **2019/09/16 16:09**

