

# InfiniBand Practical Monitoring

**Susan Coulter**  
**Los Alamos National Laboratory**

High Performance Computing Division  
HPC-3 Production Systems  
skc@lanl.gov

April 19, 2013

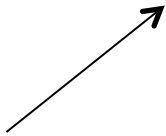
## So many choices ... so much data ...

---

- ◆ **ibdiagnet**
- ◆ **ibcheckfabric**
- ◆ **ibqueryerrors**
- ◆ **ibtrackerrors**
- ◆ **saquery**
- ◆ **smpquery**
- ◆ **perfquery**
- ◆ **ibtraceroute**
- ◆ **lbnetworkdiscover**
- ◆ **OpenSM Console** ← future development for LANL

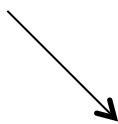
# Current LANL implementation

RPM



```
ibmon2-1.0.0-9.noarch
[root@mu-master infiniband]# rpm -ql ibmon2
/etc/cron.d/zz_run_ibmon
/etc/ibmon
/etc/ibmon/README
/etc/ibmon/data
/etc/ibmon/data/grid.dest
/etc/ibmon/grid.pl
/etc/ibmon/grid.sh
/etc/ibmon/ib_cookietrail.sh
/etc/ibmon/ib_ct_helper.pl
/etc/ibmon/ib_links.pl
/etc/ibmon/ib_rosetta.pl
/etc/ibmon/ibmon.sh
/etc/ibmon/ibmon_cleaner.pl
```

cronjob



```
[root@mu-master infiniband]# cat /etc/cron.d/zz_run_ibmon
# $Header:
03,33 * * * * root /etc/ibmon/ibmon.sh &> /dev/null
09,39 * * * * root /etc/ibmon/grid.sh
07 7 * * 0 root /etc/ibmon/ibmon_cleaner.pl
```

# Under the covers of ibmon.sh

---

```
#!/bin/bash

# new ibnet map
/usr/sbin/ibnetdiscover -g --node-name-map /etc/opensm/ib-node-name-map &> /etc/ibmon/data/ibnet_map
/bin/cp /etc/ibmon/data/ibnet_map /etc/ibmon/data/ibnet_map.`date +%Y%m%d%H%M`

# HCA list with lids
/usr/sbin/ibnetdiscover | egrep "^\[.*H-" | sort -k 4 | cut -d " " -f 1-6 > /etc/ibmon/data/hca_list

# grab counters then reset
/usr/sbin/ibqueryerrors -c -s PortXmitWait | grep -v "##" | grep -v ALL | sed 's/.*GUID .* port/ Port/'
| /etc/ibmon/ib_rosetta.pl >& /dev/null

/usr/sbin/ibclearerrors >& /dev/null

# check links
grep -i sdr /etc/ibmon/data/ibnet_map | sed 's/#// ' | sed 's/ lid.*// ' | /etc/ibmon/ib_links.pl SDR >& /dev/null
grep -i 1x /etc/ibmon/data/ibnet_map | sed 's/#// ' | sed 's/ lid.*// ' | /etc/ibmon/ib_links.pl 1X >& /dev/null
~
```

# Syslog output

---

```
Apr 2 10:33:04 mu-master ibmon2[12939]: mu1352 Port 1: [PortRcvRemotePhysicalErrors == 1] - (muib77 port 8)
Apr 2 10:33:04 mu-master ibmon2[12939]: muib77 Port 25: [PortRcvRemotePhysicalErrors == 1] - (muibcore3 Line 25 Port 5)
Apr 2 10:33:04 mu-master ibmon2[12939]: muib76 Port 26: [SymbolErrorCounter == 1] [PortRcvErrors == 1] - (mu1343 Port 1)
Apr 2 10:33:04 mu-master ibmon2[12939]: muibcore3 Line 25 Port 4: [PortRcvRemotePhysicalErrors == 1] - (muib76 Port 25)
Apr 2 10:33:23 mu-master ibmon2[14871]: No InfiniBand SDR Link Problems this run
Apr 2 10:33:23 mu-master ibmon2[14875]: No InfiniBand 1X Link Problems this run
Apr 2 11:03:04 mu-master ibmon2[23509]: muib41 Port 16: [SymbolErrorCounter == 1] [PortRcvErrors == 1] - (mu0716 Port 1)
Apr 2 11:03:23 mu-master ibmon2[25946]: No InfiniBand SDR Link Problems this run
Apr 2 11:03:23 mu-master ibmon2[25950]: No InfiniBand 1X Link Problems this run
Apr 2 11:33:04 mu-master ibmon2[9193]: No InfiniBand Errors this run
Apr 2 11:33:23 mu-master ibmon2[11080]: No InfiniBand SDR Link Problems this run
Apr 2 11:33:23 mu-master ibmon2[11084]: No InfiniBand 1X Link Problems this run
Apr 2 12:03:04 mu-master ibmon2[24356]: No InfiniBand Errors this run
Apr 2 12:03:23 mu-master ibmon2[26790]: No InfiniBand SDR Link Problems this run
Apr 2 12:03:23 mu-master ibmon2[26794]: No InfiniBand 1X Link Problems this run
```

# Interface with Zenoss

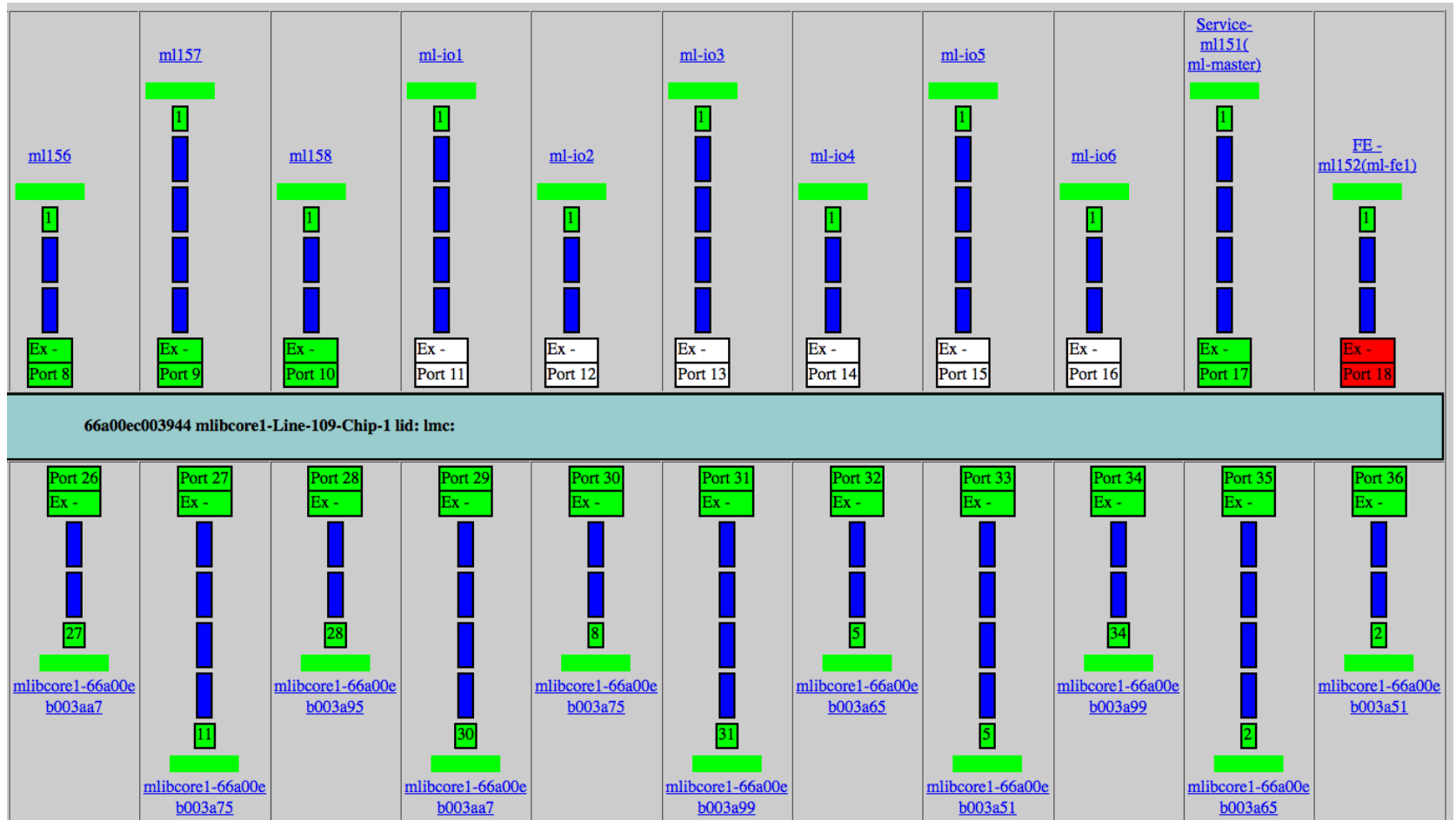
The screenshot shows the Zenoss CORE interface. The top header displays 'Zenoss™ | CORE' and the current path '/Master/moonlight-mlGB-ml'. A 'UNCLASSIFIED' label is in the top right. Below the header is a 'Welcome to the Grid!' message. A control bar includes a '30' input field, a 'Save' button, and a timer '00:00:04 until page refresh!'. The main content area features a grid of server status cards. Each card shows a host name, its status percentage, and an information icon. Below each card is a table with 'CVLAN' and 'IB' columns. The 'IB' column for 'moonlight' is highlighted in red, indicating a problem.

Host	Status (%)	CVLAN	IB
cerrillos	idle (96.26%)	CVLAN	IB
conejo	(88.39%)	CVLAN	IB
lobo	(0.74%)	CVLAN	IB
mapache	(0.00%)	CVLAN	IB
moonlight	(0.00%)	-N/A-	IB
mustang	(96.31%)	-N/A-	IB
panasas		-N/A-	-N/A-
pinto	(98.70%)	-N/A-	IB

# Grid details

mllibcore1			
Spine		Line	
Spine-201A-Chip-1	<a href="#">66a00eb003a51</a>	Line-101-Chip-1	<a href="#">66a02e800132f</a>
Spine-201B-Chip-1	<a href="#">66a00eb003a95</a>	Line-102-Chip-1	<a href="#">66a01e800132f</a>
Spine-203A-Chip-1	<a href="#">66a00eb003a55</a>	Line-103-Chip-1	<a href="#">66a00ec003943</a>
Spine-203B-Chip-1	<a href="#">66a00eb003a99</a>	Line-104-Chip-1	<a href="#">66a00ec00392b</a>
Spine-205A-Chip-1	<a href="#">66a00eb003a75</a>	Line-105-Chip-1	<a href="#">66a00ec00391e</a>
Spine-205B-Chip-1	<a href="#">66a00eb003ab7</a>	Line-106-Chip-1	<a href="#">66a00ec003947</a>
Spine-207A-Chip-1	<a href="#">66a00eb003a65</a>	Line-107-Chip-1	<a href="#">66a00ec003932</a>
Spine-207B-Chip-1	<a href="#">66a00eb003aa7</a>	Line-108-Chip-1	<a href="#">66a00ec00394e</a>
Spine-209A-Chip-1	<a href="#">66a00eb00381f</a>	Line-109-Chip-1	<a href="#">66a00ec003944</a>
		Line-110-Chip-1	<a href="#">66a00ec003919</a>
		Line-111-Chip-1	<a href="#">66a00ec00391f</a>
		Line-112-Chip-1	<a href="#">66a00ec003921</a>
		Line-113-Chip-1	<a href="#">66a00ec003909</a>
		Line-114-Chip-1	<a href="#">66a00ec00393b</a>
		Line-115-Chip-1	<a href="#">66a00ec003945</a>
		Line-116-Chip-1	<a href="#">66a00ec003925</a>
		Line-117-Chip-1	<a href="#">66a00ec003934</a>
		Line-118-Chip-1	<a href="#">66a00ec003940</a>

# More grid details





# Message as displayed by Zenoss

Sev All State Acknowledged Stop 60 66a00ec003944

Select: All None Acknowledged Unacknowledged 1-6 of 6

eventT	device	comp	eventCl	summary	firstTime	lastTime	count		
<input type="checkbox"/>	None	66a00ec003944	gridState	/infiniband/ibmon2	mllibcore1 Line 109 Port 18: [SymbolErrorCounter > 100] - (ml194 Port 1)	2013/04/02 08:03:02.00	2013/04/02 08:03:02.00	1	
<input type="checkbox"/>	None	66a00ec003944	ibmon2	/infiniband/ibmon2	mllibcore1 Line 109 Port 17: [VL15Droppe d == 2] - (ml-master Port 1)	2013/04/01 23:03:02.00	2013/04/02 14:03:02.00	13	
<input type="checkbox"/>	None	66a00ec003944	ibmon2	/infiniband/ibmon2	mllibcore1 Line 109 Port 17: [VL15Droppe d == 1] - (ml-master Port 1)	2013/03/31 16:03:02.00	2013/04/02 13:03:02.00	35	
<input type="checkbox"/>	None	66a00ec003944	ibmon2	/infiniband/ibmon2	mllibcore1 Line 109 Port 17: [VL15Droppe d == 3] - (ml-master Port 1)	2013/04/02 06:03:02.00	2013/04/02 11:03:02.00	3	
<input type="checkbox"/>	None	66a00ec003944	ibmon2	/infiniband/ibmon2	mllibcore1 Line 109 Port 17: [VL15Droppe d == 4] - (ml-master Port 1)	2013/04/02 10:33:02.00	2013/04/02 10:33:02.00	1	

Fields	Details	Log	Issue	Rollup
Field	Value			
explanation	If SymbolErrorCounter > 100 Then email system oncall and hpcnet-day-oncall, turn grid red. Regex catches numbers >= 100. Trigger command(resendMessageForError) to resend a message to trigger filter to increment count for this local device(catchLocalDevi			
LLABEL	mllibcore1 Line 109			
LLOC	18			
LOCAL_DEVICE	mllibcore1 Line 109 Port 18			
MESG	[SymbolErrorCounter == 65535] [LinkDownedCounter == 1]			
NAME	SymbolErrorCounter			
originalTime	Apr 2 08:03:02			
pid	89835			
REMOTE_DEVICE	ml194 Port 1			
SEC	65535			

# Wiki pages / documentation ...

## AWOL Link

This is not a counter, but an error as identified by LANL processes. When ibnetdiscover is run, any link that is *live* but not responding throws an error. ibmon and logged to syslog.

### Example:

```
Aug 9 12:33:15 mu-master ibmon2[32424]: AWOL Link: (DR path slid 0; dlid 0; 0,1,1,20,25,16,34 Attr 0x11:0)
```

```
[root@mu-master ~]# smpquery portinfo -D 0,1,1,20,25,16,34
```

This should result in an error. If it does not you will see information like what is in the next example. If this works - it means the node/port was probably coming up but not yet able to respond to a MAD packet. If it fails, remove the last port number and run again, grep'ing for the LID.

```
[root@mu-master ~]# smpquery pi -D 0,1,1,20,25,16,34 | grep -i lid
# Port info: DR path slid 65535; dlid 65535; 0,1,1,20,25,16,34 port 0
Lid:.....1095
SMLid:.....250
```

## -- SymbolErrorCounter

Mmm Dd Hh:Mm:Ss	Host	Class	Local Side:	Error	- Remote Side
May 25 08:03:03	mu-master	ibmon2[15906]:	mul246	Port 1: [SymbolErrorCounter == 1] [PortRcvErrors == 1]	- (ib71 port 4)

**Action:** If `SymbolErrorCounter > 100` Then email system oncall and hpcnet-day-oncall, turn grid red. Trigger command(`resendMessageForError`) to resend a message to trigger filter to increment count for this local device(`catchLocalDevice`)

**Regex:** `(.*) .*[SymbolErrorCounter == (.*)] .* - \((.*)\)` Where \$1 is local device, \$2 is SymbolErrorCounter value, \$3 is remote device

**Regex(Python):** `(?P<LOCAL_DEVICE>(P<LLABEL>[A-Za-z0-9]+)s*(?P<LLOC>.*)):(?P<MESG>.*\[(?P<NAME>SymbolErrorCounter) == (?P<SEC>\d{3,})\].*) - \((?P<REMOTE_DEVICE>.*))\)`

# End

## Questions?