Quick troubleshooting guide

Creation Date
Lior Ofer, 9/23/2007 4:27 AM
Last Modified
Lior Ofer, 9/23/2007 4:27 AM

Sometimes a little helpful direction can be just what you need to get through a tough installation problem. This Quick Troubleshooting Guides below provide some helpful instructions how to isolate problems.

You will find yourself trying to troubleshoot a network related problem which usually appears in few forms. The main reasons are MPI problem, performance problem or complete lack of connectivity. These symptoms can be caused by:

- Bad cabling
- HCA problem
- IPOIB Interface problem
- Missing Configuration
- SM problem

Let's start by checking the basics:

1. Be sure your switch and hosts are powered on
2. Be sure cables are plugged in properly.
3. Check that the SM is running
    a. Login to the master Switch CLI
    b. Run the command sm-info show and make sure that sm mode is enabled and sm state is master
    c. Run the command sm-info show few times , make sure sm activity counter is progressing
    d. In case the sm mode is disabled, enable it (CLI ◊enable ◊ config ◊ sm ◊ sm-info mode set enable
    e. . In case the sm state is not master it means that other switch or node in the fabric is running another SM that may be the master

For more information on Testing you HCA please refer to FAQ-F10018
http://www.voltaire.com/ftp/support-products/source/FAQ/FAQ-F10018.pdf

4. Please follow the procedure below to make sure all of your HCA is working
    a. Run lspci check that you see mellanox HCA is identifiede on the PCI bus

b. If not reseat HCA or the raiser card -

c. Replace HCA with another

5. Check that the Host links are active

a. on Gridstack 4.x run ibstat or ibv_devinfo and check that PORT IS ACTIVE

b. On Gridstack 3.x or windows run vstat

Note that for Gridstack 4.x drivers (OFED based), HCA port 1 is "hard coded" to ib0 interface, port 2 to ib1, etc.

6. Check that the IPOIB interface is up

a. Run ifconfig –a to view all network interface it might be that the ib0 or ib1 is there but not activated

b. Run ifconfig make sure your ipoib interface is configured (in case its not use ib-config to configure it <ib-config –h for help>)

c. In not ACTIVE check you have a green and Amber LED's on the HCA

d. Check that SM in running

7. Check that you can ping between nodes on IPOIB

a. Run the command ifconfig and make sure the following line appears exactly at your IB interface:

UP BROADCAST RUNNING MULTICAS

If the RUNNING The IPOIB host is not joined to the IB multicast group. In this case check the SM health.

b. Check for IP problems such as duplicate IP, wrong routing table or wrong destination address

c. If not check to see you have latest firmware on the switches and HCA ASIC

8. Run Cable and Link Tests

Your servers won't be able to communicate with any other nodes on your network unless the HCA physical and logical" link" light is on. This indicates that the connection between your server and the switch is functioning correctly. In most cases a lack of link is due to link failure include: The cables are bad. The cables aren't plugged in properly. Please follow the procedure below to make sure all you fabric cable are ok

a. Login to the switch CLI ◊ enable ◊ utilities

b. Run port-verify –z to clear performance monitoring counters

c. Run traffic on your cluster. Try to use an application that can exercise as many ports as possible. Such an application can be MPI benchmark as pallas

d. Run port-verify –be to get a list of ports with errors

e. Run port-verify –bw to get a list of ports with 1x reduced link width

f. Bellow you may find the list of counters and what is the recommendations

g. If you high amount and rate of LinkErrorRecovery , LinkDowned , PortRcvErrors reseat or replace cable

a. Reset (“zero”) the counters

b. Run traffic

c. Check again with port-verify -eb

h. If you find 1x links

a. To reset all 1x ports back to 4x Login to CLI and go to utilities and Run the command netdiscover –r r1x

b. Run port-verify –bw again

c. If you still see 1x links reseat or replace the cable

i. Repeat step b-i till fabric is error free

for more information please refer to FAQ-F10042 -fabric diagnostics
http://www.voltaire.com/ftp/support-products/source/FAQ/FAQ-F10042.pdf

Counter name	Meaning
Width	All links should be running at 4X. Any links reporting 1X are bad and need to be repaired
SymbolError	Can increase without a significant problem present
LinkErrorRecovery	Increasing LinkErrorRecovery errors may indicate a bad link
LinkDowned	Indicates number of times the port has gone down (usually for valid reasons)
PortRcvErrors	This counter should not be increasing. Increasing number indicates a bad link
PortRcvRemotePhysicalErrors	This indicates that a problem is occurring ELSEWHERE in the fabric and that this port received a packet that was intentionally corrupted by another switch in the fabric
PortRcvSwitchRelayErrors	Does not indicate a problem
PortXmitDiscards	May indicate HOQ or other parameter should be tweaked
PortXmitConstraintErrors	May indicate that a parameter should be tweaked
PortRcvConstraintErrors	May indicate that a parameter should be tweaked
LocalLinkIntegrityErrors	Counter should not be increasing. Increasing number

	indicates a bad link
ExcessiveBufferOverrunErrors	May indicate that a parameter should be tweaked
Status	If the value reads “IB-Timeout” this means that the link is in Initialize mode and failed to negotiate a logical link
VL15Dropped	This counter increasing in small increments is not seen as a problem.

9. If you MPI application is not working

    1. Run one of the precompiled applications that are supplied with the Gridstack package under the MPI bin directory (cpi, mpi_bandwidth, etc)

    2. Run the mpi with 2 ranks and increase until you hit the problem (2, 4, 500, 1000, 2000 ,...)

    3. Run the mpi with Test one rank per node

    4. Run the mpi with hello_world program (supplied with Gridstack package)

    5. Run the mpi with low memory consumption parameters.

    6. Run ldd on the application: make sure all dynamic loaded libraries links are OK.

    7. Understand system: mount, nfs, diskless, /tmp/ directory memory size (especially on disc-less nodes), dmesg, ibstat.

    8. Run ssh hostname
    Change mpirun_ssh timeout 300

    9. Run mpirun_ssh show – in order to make sure that the command parameters are correct.

    10. Run the mpi with -V 5 look for higher verbosity. Check that all ranks re-connect

    11. Run debug application: mpirun_test (not IB application: test only TCP job start).

    12. Disable shared mem

    13. Disable RDMA collective

    14. Changes to TCP configuration for job start (mainly needed in the head server):
    echo 5000 > /proc/sys/net/ipv4/tcp_max_syn_backlog
    echo 5000 > /proc/sys/net/core/netdev_max_backlog
    echo 5000 > /proc/sys/net/core/somaxconn
    /proc/sys/net/ipv4/tcp_mem, which is by default a fairly modest value:12288 16384 24576 By bumping up the limits in tcp_mem to 12288 65536 98304
    Run the mpi over IpoIB or GB-Eh

For more details and specific directions to knowledge base please review the followingFAQ (Common MPI problems):

http://www.voltaire.com/ftp/support-products/source/FAQ/FAQ-F10134.pdf