[HPC Cluster Install: InfiniBand](http://idolinux.blogspot.com/2010/06/hpc-cluster-install-infiniband.html)

Welcome to the fourth installment of [HPC Cluster Install](http://idolinux.blogspot.com/search/label/cluster-install). In this post, we will be covering the configuration of QLogic Infiniband on CentOS / RHEL 5.  
  
Check out the prior three posts for some background on hardware and prerequisites:

* [HPC Cluster Install: Intro](http://idolinux.blogspot.com/2009/10/hpc-cluster-install-intro.html)
* [HPC Cluster Install: Planning and Purchasing](http://idolinux.blogspot.com/2010/04/hpc-cluster-install-planning-and.html)
* [HPC Cluster Install: Admin and Kickstart Server](http://idolinux.blogspot.com/2010/05/hpc-cluster-install-admin-and-kickstart.html)

Let me start by saying that I tried a couple of different approaches until I found one that worked. I tried the stock OFED RPM packages (yum install openib ibutils infiniband-diags opensm), which did not work. The packages included in the 5.4 repo are not new enough for our hardware. After removing that, I tried the latest stable source tarball (OFED-1.5 from openfabrics.org) which seemed to be working except for a few opensm and ibdiagnet errors. In the end, I switched to the the QLogic-provided OFED+ to obtain hardware support, which was timely, through the [QLogic support](http://support.qlogic.com/" \t "_blank) website.

The hardware:

* 1 QLogic InfiniBand Switch 12300 Managed EDGE
* 29 nodes with IB 4X DDR PCIe x8
* no IB card in the head node

The first thing to do is to update the QLogic switch firmware.

* browse to http://admin:adminpass@192.168.100.9/
* select on left Maintenance -> Firmware Update
* upload X.xt3.5.0.5.0.1.pkg provided by support
* reboot switch

Next we install the latest QLogic OFED+ on one of the compute nodes. Note that I was provided with a non-public, brand new build for RHEL 5.4, so I recommend downloading the latest version and/or contacting support.

head# ssh node29

node29# yum -y remove openib ibutils infiniband-diags opensm

node29# yum -y install libsysfs libsysfs-devel pciutils-devel tcl tcl-devel tk tk-devel libstdc++-devel libstdc++-devel gcc-c++ krb5-devel krb5-libs libevent-devel nfs-utils-lib-devel openldap-devel e2fsprogs-devel zlib-devel kernel-devel

node29# cd /usr/global/src/

node29# mkdir openib ; cd openib

5.1.0.2.8 Basic:

node29# wget ftp://ftp.qlogic.com/support/Hidden/code/sti/hca/IFS\_5\_1\_0\_2\_8/OPENIB\_INSTALL\_5\_1\_0\_2\_8/X86\_64\_RHEL5/QLogicIB-Basic.RHEL5-x86\_64.5.1.0.2.8.tgz

node29# wget ftp://ftp.qlogic.com/support/Hidden/code/sti/hca/IFS\_5\_1\_0\_2\_8/OPENIB\_INSTALL\_5\_1\_0\_2\_8/release\_QLogicIB-OFED+HostSW\_51028\_B.pdf

node29# tar xzvf QLogicIB-Basic.RHEL5-x86\_64.5.1.0.2.8.tgz

node29# cd QLogicIB-Basic.RHEL5-x86\_64.5.1.0.2.8

node29# less README

node29# ./INSTALL

press 1 for install

press p for perform uninstall of old and install of new

press return for all defaults ...

x to exit

head# cd /usr/global/src/openib/QLogicIB-Basic.RHEL5-x86\_64.5.1.0.2.8; ./INSTALL -a

head# dsh -acM "cd /usr/global/src/openib/QLogicIB-Basic.RHEL5-x86\_64.5.1.0.2.8; ./INSTALL -a"

nodeXX: A System Reboot is recommended to activate the software changes

nodeXX: Done Installing QLogic IB Software.

head# dsh -acM reboot

head# dsh -acM "cat /sys/class/infiniband/qib0/ports/1/state"

INIT

head# ssh node29

node29# chkconfig opensmd on ; service opensmd start

node29# tail -F /var/log/opensm.log

head# dsh -acM "cat /sys/class/infiniband/qib0/ports/1/state"

ACTIVE

head# dsh -acM "cat /sys/class/infiniband/qib0/ports/1/rate"

20 Gb/sec (4X DDR)

node29# ibdiagnet -pc -c 1000

node29# ibcheckerrors

Now that we have an error-free fabric, we will integrate with the vendor-provided Open MPI SRPM.

head# ssh node29

node29# rpm -ihv /usr/global/src/openib/QLogicIB-Basic.RHEL5-x86\_64.5.1.0.2.8/InfiniPath2.5.1-2010-01-07-RHEL5-x86\_64/OtherMPIs/SRPMS/openmpi\_gcc\_qlc-1.4-1.rhel5.src.rpm

node29# rpmbuild -bb --define 'configure\_options --with-sge --with-contrib-vt-flags=--disable-iotrace' /usr/src/redhat/SPECS/openmpi-1.4.spec

error fixed with above vt option:

vt\_iowrap.c:1242: error: expected declaration specifiers or '...' before numeric constant

vt\_iowrap.c:1243: error: conflicting types for '\_\_fprintf\_chk'

node29# cp /usr/src/redhat/RPMS/x86\_64/openmpi-1.4-1.x86\_64.rpm /usr/global/src/openib/

head# dsh -acM "rpm -i /usr/global/src/openib/openmpi-1.4-1.x86\_64.rpm"

And then validate that Open MPI is functional.

head$ ssh node01

node01$ mpirun --machinefile /etc/machines.list -np 4 hostname

node01

node03

node02

node04

node01$ mpirun --machinefile /etc/machines.list -np 4 ./test

Hello World from Node 0

Hello World from Node 1

Hello World from Node 2

Hello World from Node 3