Introduction

The top-level instructions include information for the installation and functional validation of OpenMPI, with installation on the head and compute nodes.

```
$> dnf -y install openmpi openmpi-devel gcc-c++
$> dnf -y --installroot=$CHROOT install openmpi openmpi-devel gcc-c++
$> packimage centos8-x86_64-netboot-compute
$> pdsh -w c[1-2] reboot
```

Successful installation is confirmed with functional validation testing. There were basic connectivity tests by pinging to ensure the compute nodes can reach the head node, and using a sample MPI script to demonstrate that the compute nodes can communicate with each other. In this document we will move beyond functional testing and validate network performance to ensure it meets expectations. Here, performance validation is used to determine the speed of network communications over a given interface. In other words, this document will validate that the communication network(s) are performing at an acceptable level. Performance will be measured using the using the OSU Micro-Benchmark suite, executed with the open-source OpenMPI, in addition to the IMB benchmarks from Intel OneAPI.

Ethernet: OpenMPI

This section will validate the performance of the standard Ethernet connection using OpenMPI. It assumes a head node connected to two compute nodes - c1 and c2 - in a minimal cluster setup using the instructions in the top-level cluster_setup documentation, with no other connections on the cluster such as Infiniband or high-speed Ethernet connection. Using root is not recommended unless otherwise noted.

Install the OSU Micro-Benchmarks that will be used for performance testing. The following commands install the OSU benchmarks in the user's home directory.

```
$> wget https://mvapich.cse.ohio-state.edu/download/mvapich/osu-micro-benchmarks-5.8.tgz
$> tar -xzf osu-micro-benchmarks-5.8.tgz
$> cd osu-micro-benchmarks-5.8/
$> ./configure CC=/usr/lib64/openmpi/bin/mpicc CXX=/usr/lib64/openmpi/bin/mpicxx
$> make
$> make install exec_prefix=~/osu_benchmarks_openmpi
```

We will validate performance by using one-sided (RMA) communication for lower overhead. The osu_put_bw test is adequate for this purpose. Use the following sample script to execute the osu_put_bw test on two compute nodes:

```
#!/bin/bash -l
#SBATCH -N 2
#SBATCH -J perf_test
#SBATCH -p normal
#SBATCH -t 20
#SBATCH -o osu_perf_test.out
#SBATCH -e osu_perf_test.err

export PATH=/usr/lib64/openmpi/bin:$PATH
mpirun -n 2 -N 1 -mca btl self,tcp
~/osu_benchmarks/libexec/osu-micro-benchmarks/mpi/one-sided/osu_put_bw
```

Submit the script on the head node with sbatch. A sample output from the osu_put_bw benchmark is included below:

```
# OSU MPI_Put Bandwidth Test v5.8
# Window creation: MPI_Win_allocate
# Synchronization: MPI_Win_flush
# Size      Bandwidth (MB/s)
1                 0.20
2                 0.40
4                 0.81
8                 1.58
16                2.89
32                5.67
64               10.22
128              21.14
256              40.65
512              65.06
1024             84.70
2048             98.46
4096            106.42
8192            110.92
16384           113.80
32768           115.32
65536           116.06
131072          116.53
262144          116.76
524288          116.88
1048576         116.92
2097152         116.95
4194304         116.96
```

The output from the benchmark is listed in MB/s. For easier comparison we will convert to Gb/s. Using the highest listed bandwidth output, convert as followed:

$$\frac{116.96\,MB}{1\,s} \times \frac{1\,GB}{1000\,MB} \times \frac{8\,Gb}{1\,GB} = \frac{.9357\,Gb}{1\,s}$$

The calculated .93Gb/s is close to the theoreitcal peak rate of 1Gb/s for the ethernet connection.