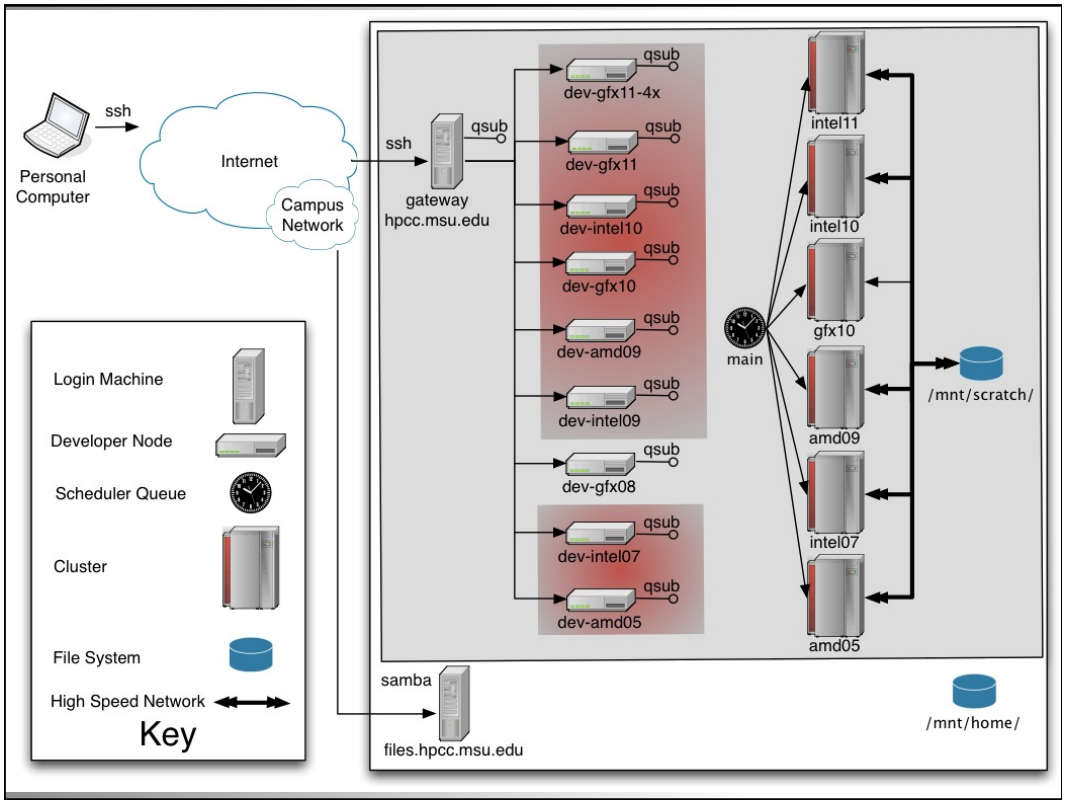
Tips and Tricks for Shared Network Parallelization: MPI on HPCC

Dirk Colbry
colbrydi@msu.edu
Research Specialist
Institute for Cyber-Enabled Research

MICHIGAN STATE UNIVERSITY

ICER



MPI on HPCC

- Two Flavors of MPI
- Switching flavors and compiling
- Running in a script
- Running on the developer nodes

MICHIGAN STATE UNIVERSITY

ICER

Two Flavors of MPI

- **mvapich** vs **openmpi** (default)
- Historically **mvapich** was much faster that **openmpi**
- The newest version of **openmpi** is just as fast as **mvapich**
- I feel that **openmpi** is much easier to use, but either will work on HPCC

MICHIGAN STATE UNIVERSITY

ICER

Switching Flavors

- Use the "module" command to switch between the two versions of mpi
- **Openmpi** module is loaded by default
- To switch to mvapich you first need to unload **openmpi**:

```
> module unload openmpi
```

- Then you need to load **mvapich:**

```
> module load mvapich
```

- You can do both commands in one step by using swap:

```
> module swap openmpi mvapich
```

MICHIGAN STATE UNIVERSITY

ICER

Submission Scripts

openmpi

```
#!/bin/bash -login
#PBS -l nodes=10:ppn=1
cd ${PBS_O_WORKDIR}
mpirun <program_name>
```

mvapich

```
#!/bin/bash -login
#PBS -l nodes=10:ppn=1
cd ${PBS_O_WORKDIR}
module swap openmpi mvapich
mpiexec <program_name>
```

MICHIGAN STATE UNIVERSITY

ICER

Trying out an example

1. Log on to one of the developer nodes
2. Load the powertools module:
   ```
   > module load use.cus powertools
   ```
3. Run the getexample program. This will create a folder called helloMPI:
   ```
   > getexample helloMPI
   ```
4. Change to the helloMPI directory and read the readme files
5. Or just type the following on the command line:
   ```
   > ./README
   ```

MICHIGAN STATE UNIVERSITY

ICER

Testing MPI jobs

- Use mpirun instead of mpiexec
- Need a hostfile

  > echo $HOST >> ./hostfile

  > echo $HOST >> ./hostfile

  > echo $HOST >> ./hostfile

  > echo $HOST >> ./hostfile
- MPIRUN example:

  > **mpirun –np 4 –hostfile ./hostfile helloMPI**

MICHIGAN STATE UNIVERSITY

ICER

Running on the Command Line

- The scheduler automatically knows how many and where to run MPI processes.
- However, on the command line, you need to specify the nodes and processors.
- However **openmpi** and **mvapich** are a little different.

Command Line Differences

- Openmpi
  - **mpirun**
  - Default assumes one process on the current host.
  - You do not even need the **mpirun** command to run the default.
  - Optionally you can use the –n and –hostfile options to change the default
- mvapich
  - **mpirun**
  - Requires both the –np and –machinefile flag to run.

Command line

- mvapich

```
mpirun -np 4 -machinefile machinefile <program_name>
```

- openmpi

```
mpirun -n 4 -hostfile machinefile <program_name>
```

- NOTE: I did a check and either MPI implementation will work with either notation.

MICHIGAN STATE UNIVERSITY

ICER

Which MPI command do you use?

	Command Line	Job Script
openmpi	mpirun	mpirun
mvapich	mpirun	mpiexec

MICHIGAN STATE UNIVERSITY

ICER