MPI Cluster Setup – Multi-node – MPICH2

*** PLEASE **DO NOT TRY** THIS IN LAB 6019. THE CLUSTER IS ALREADY SET UP ***

The file system used for the cluster setup is network file system (NFS). There is a conceptual master storage location, that is mounted over a location in each local node, for providing a view of globally shared data storage.

Here a new user account is created at each node with name **mpiuser, creation of which is discussed later** (not necessary, you may use your existing account).

1. **Defining hostnames**
   Edit /etc/hosts file to provide names for nodes to be connected for ease of use, in all the nodes.
   An example /etc/hosts file may look as follows
   127.0.0.1   localhost
   192.168.133.100   mpimaster
   192.168.133.101   mpi1
   192.168.133.102   mpi2
   192.168.133.103   mpi3

2. **Installing NFS**
   NFS allows to create a folder at notional master and sync it over other nodes
   At master node
           sudo apt-get install nfs-server
   At other nodes
           sudo apt-get install nfs-client

3. **Sharing Master Folder**
   Make a folder in all nodes, where we will store all our data and programs
           sudo mkdir /mirror

   Modify the /etc/exports file in Master node to allow sharing of contents in Master's mirror folder to other nodes
           Way – 1 : From terminal
                   echo "/mirror   *(rw,sync)" | sudo tee -a /etc/exports
           Way – 2 : Open /etc/exports and add the following line at the end
                   /mirror *(rw, sync)
   Now at Master node, restart the NFS kernel server
                   sudo service nfs-kernel-server restart

4. **Mounting mirror folder in nodes**

Add the following line in /etc/fstab file in non-master nodes

       mpimaster:/mirror   /mirror nfs

       assuming hostname of master node given as mpimaster

5. **Define a user for running MPI programs**

Define a new user, say mpiuser, on all nodes with same name and same userid with home directory in /mirror

       sudo useradd -m -d /mirror/mpiuser -s /bin/bash -uid 1001 -G sudo mpiuser

Setup password

       sudo passwd mpiuser

Also give the ownership of /mirror to mpiuser

       sudo chown mpiuser /mirror

6. **Install SSH server**

Install OpenSSH server in all nodes

       sudo apt-get install openssh-server

7. **Setting up passwordless SSH for communication between nodes**

At master node, login to mpiuser

       su -mpiuser

Generate RSA key pair

       ssh-keygen -t rsa

Add this key to authorized keys

       cd .ssh

       cat id_rsa.pub >> authorized_keys

You may check now if passwordless ssh is working by

       ssh non-master_node_hostname

where non-master_node_hostname is the hostname of any of the other node mentioned in /etc/hosts file

8. **Install MPICH2**

Install MPICH2 on all the nodes

       sudo apt-get install mpich2

Alternatively, you may install openmpi in one node in the cluster anywhere in the shared /mirror directory

Test if successful by

       which mpiexec

       which mpirun

9. **Testing**

Write a simple MPI program, say pgm1.c and compile as

       mpicc pgm1.c -o pgm1

Run the program in multiple nodes as

       mpirun -np 10 -host mpimaster,mpi1,mpi2 ./pgm1

where -np 10 says to spawns 10 processes in 3 nodes

Run program in single node as

Mpirun -np 4 ./pgm1

where -np 4 says to spawn 4 processes in the local node

Reference : https://help.ubuntu.com/community/MpichCluster