MPI Overview

- Communication Protocol
- Language Independent
- Specification of what MPI library should do
- Portable (Implemented on almost every distributed architecture)
- Available implementations are openMPI, LAM/MPI, MPICH etc

Why MPI?

- Threads are used for intra node communication
- Program can be made to run on multiple processors

When to use MPI?

- Small memory requirement, then go for threading
- Data size is moderate and intensive computation, go for MPI
- Data size is large and computation is moderate, go for MapReduce

Cons

- Designing a parallel algorithm is hectic
- Harder to debug

Installation

- Download source code from http://www.mpich.org/downloads/
- Step 1 : Configure installation
  - sudo ./configure –disable fortran
  - It should say Configuration completed
- Step 2: Build and install mpich2
  - sudo make and sudo make install

Installation

- To install over LAN
  - Need machine communicating over network via ssh
  - Also need to share the data using NFS

Compilation

- To compile use mpicc
- mpicc
  - Program help to use standard C language compiler together with MPI library
- Also other option available are mpig++, mpif77

Execution

- To execute use mpirun or mpiexec
- Options used with mpirun
  - np : specifies number of processes
  - host : Host file with IP addresses (if multiple node are present)

Execution

- Initializes and gets arguments
- Gets node list
- Allocate node to no of process
- Launch processes
- Waits for completion
- Checks exit status and prints error message if any