

华东师范大学超算中心

MPI 使用情况说明

华东师大超算中心 MPI 使用情况说明

曙光公司为华东师大提供了丰富的 MPI 通讯库。主要有如下几种：

1. INTEL MPI

是 INTEL 公司提供的高性能 MPI，同时支持 Infiniband、TCP 两种网络，安装路径为/data/soft/compiler/mpi/impi/3.2.2.006

2. HP-MPI

目前是 Platform 旗下的 MPI 产品，版本 2.2.7，环境变量设定参考 /data/share/env_hpmpi；

3. MVAICH2

是 MPI 接口在 Infiniband 网路上的 MPI2 实现版本，在 Infiniband 上具有较高的性能。安装路径为

/data/soft/compiler/mpi/mvapich2/1.4rc2/icc.ifort/

4. OpenMPI

是 MPI2 实现的高性能的 MPI，可以运行于任何网络上。安装路径为

/data/soft/compiler/mpi/openmpi/1.3.3/gcc.ifort/	采用 gcc 4.1.2 和 ifort 9.1.041 编译
/data/soft/compiler/mpi/openmpi/1.3.3/gcc.pg90/	采用 gcc 4.1.2 和 pgf90 9.0 编译
/data/soft/compiler/mpi/openmpi/1.3.3/icc.ifort/	采用 icc 11.0 和 ifort 11.0 编译

下面将以 cpi 为例，介绍不同的 MPI 利用作业调度系统运行的方法。

5. 作业调度系统上 MPI 使用举例

5.1 INTEL MPI

环境变量设定为：

```
#  
# for impi  
#  
. /data/soft/compiler/intel/Compiler/11.1/056/bin/iccvars.sh intel64  
. /data/soft/compiler/intel/Compiler/11.1/056/bin/ifortvars.sh intel64  
. /data/soft/compiler/mpi/impi/3.2.2.006/bin64/mpivars.sh
```

测试脚本为：

```
dolphin@CLOUD@ECNU:~/work/mpi/impi> cat 1.job.impi
#!/bin/bash -x

#PBS -N mycpi
#PBS -l nodes=2:ppn=8
#PBS -l walltime=00:08:00
#PBS -j oe
#PBS -q debug

#
#define variables
#
n_proc=$(cat $PBS_NODEFILE | wc -l)

#
#running jobs
#
cd $PBS_O_WORKDIR

#
# Setup the MPI topology
#

time -p /data/soft/compiler/mpi/impi/3.2.2.006/bin64/mpirun --rsh=ssh -env I_MPI_DEVICE
rdma:OpenIB-cma -np ${n_proc} ./cpi

exit 0
```

5.2 HP MPI

环境变量设定为：

```
dolphin@CLOUD@ECNU:~/work/mpi/impi> cat /data/share/env_hpmpi
#!/bin/bash
#
#for hpmpi
#
export MPI_CC=pgcc
export MPI_F77=pgf77
export MPI_F90=pgf90
export MPI_CXX=pgCC
MPI_ROOT=/data/soft/compiler/mpi/hpmpi
PATH=${MPI_ROOT}/bin:$PATH
MANPATH=${MPI_ROOT}/share/man:$MANPATH
LD_LIBRARY_PATH=${MPI_ROOT}/lib/linux_amd64:$LD_LIBRARY_PATH
LM_LICENSE_FILE=$LM_LICENSE_FILE:${MPI_ROOT}/licenses/license.dat
MPI_REMSH=/usr/bin/ssh

export PATH MANPATH LD_LIBRARY_PATH LM_LICENSE_FILE MPI_REMSH
```

测试脚本为：

```
dolphin@CLOUD@ECNU:~/work/mpi/hpmpi> cat 1.job.hpmpi
#!/bin/bash -x

#PBS -N mycpi
#PBS -l nodes=2:ppn=8
#PBS -l walltime=00:08:00
#PBS -j oe
#PBS -q debug

#
#define variables
#
n_proc=$(cat $PBS_NODEFILE | wc -l)

#
#running jobs
#
cd $PBS_O_WORKDIR

#
# Setup the MPI topology
#

time -p /data/soft/compiler/mpi/hpmpi/bin/mpirun -e MPI_IC_ORDER="UDAPL:TCP"
-prot -np ${n_proc} -hostfile $PBS_NODEFILE ./cpi

exit 0
```

5.3 MVAPICH2

环境变量设定为：

```
PATH=/data/soft/compiler/mpi/mvapich2/1.4rc2/icc.ifort/bin:$PATH
```

运行脚本为：

```
dolphin@CLOUD@ECNU:~/work/mpi/mvapich2> cat 1.job.mvapich2
#!/bin/bash -x

#PBS -N mycpi
#PBS -l nodes=2:ppn=8
#PBS -l walltime=00:08:00
#PBS -j oe
#PBS -q debug

#
#define variables
#
n_proc=$(cat $PBS_NODEFILE | wc -l)

#
#running jobs
#
cd $PBS_O_WORKDIR

#
# Setup the MPI topology
#

time -p /data/soft/compiler/mpi/mvapich2/1.4rc2/icc.ifort/bin/mpirun_rsh -ssh -np
${n_proc} -hostfile $PBS_NODEFILE ./cpi

exit 0
```

5.4 OpenMPI

环境变量设定为：

```
PATH=/data/soft/compiler/mpi/openmpi/1.3.3/gcc.pg90/bin:$PATH
```

运行脚本为：

```
#!/bin/bash -x

#PBS -N mycpi
#PBS -l nodes=2:ppn=8
#PBS -l walltime=00:08:00
#PBS -j oe
#PBS -q debug

#
#define variables
#
n_proc=$(cat $PBS_NODEFILE | wc -l)
LD_LIBRARY_PATH=/data/soft/compiler/mpi/openmpi/1.3.3/gcc.pg90/lib:$LD_LIBRARY_PATH
#
#running jobs
#
cd $PBS_O_WORKDIR

#
# Setup the MPI topology
#

time      -p      /data/soft/compiler/mpi/openmpi/1.3.3/gcc.pg90/bin/mpirun      -x
LD_LIBRARY_PATH --mca btl openib,self -np ${n_proc} ./cpi

exit 0
```