

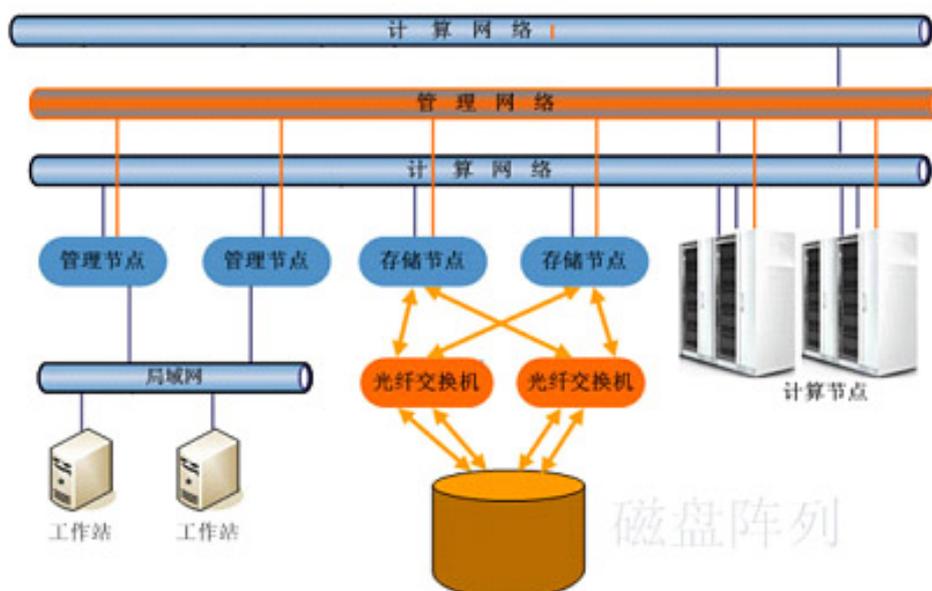
CPU-GPU 并行计算机群并行运算指南

登陆：只能通过校内IP登陆到CPU-GPU服务器,登陆方式为：

`ssh -X UserName@222.195.76.97` (启用X11图形模型)

`ssh UserName@222.195.76.97` (不启用X11图形模型)

本机不是一个普通的服务器，它是由独立的内部的计算网络和管理网络，联合高性能服务器以及专业的网络云服务平台，构筑的并行服务器阵列。



希望用户在使用时请特别注意：

为防止错误使用造成的堵塞，不可以把计算放到登录节点上。登录节点（管理节点）运算能力很差，然而网络交换却很频繁。请用户将运算准确地展开到自己指定的计算节点上（不可只使用 **-np** 还应联合使用 **-machinefile ma**）！使用前应学习一些并行计算的知识、以及高性能计算平台的使用方法。

废进程可能产生堆积，有时严重占用内存，降低**Cluster**的使用效率。用户应经常查找自己僵死掉的进程(**ps -aux**),并**kill**掉！

机器的故障是无法避免的，重要数据、中间结果.....用户自己要随时在本机以外的设备上备份，自动保护措施也会因机器的故障而失去作用！

另外：

本曙光服务器集群有**27**个节点，其中**25**个节点(**node1~node25**)

是计算节点、 **Node99** 是存储节点、 **node100**是管理节点，也是用户登录的节点。计算节点用的是 **2 CPU**（每CPU **8核**）服务器、每服务器 **32G**内存，服务器内有壹块**PCI-E X16 GPU**卡（ **Nvidia Tesla™ C2050 3GB** 和 **Nvidia Tesla™C2075 6GB Fermi CUDA**架构 **448**个**CUDA**核心），是带**GPU**的计算节点。本集群,既可以按普通集群做高性能计算，也可以使用**KUDA**编程做包含**GPU**的高性能计算。

管理网使用了快速千兆以太交换网络、数据计算网使用了高速通信交换网络（**Infiniband 40Gb/s**）

- 1、本集群的工作 **Home** 在/**public**下。
- 2、本集群的高性能计算软件 在/**public/software**下。
- 3、环境变量脚本在相应的/**public/software/mpi**中。
如发现环境变量有缺失，请运行自己要用的 **.sh** 文件（**openmpi** 或 **mpich2**），以构筑自己运行时的环境变量。或参考：**.sh** 文件，设置长期使用环境变量。
- 4、如要使用 **pbs**，设置环境变量时，请注意 **pbs** 的位置：
/opt/gridview/pbs/dispatcher/bin

/opt/gridview/pbs/dispatcher-sched/bin

在用 **qsub** 提交任务时，如提示非法，请通知我将其加入队列。

- 5、磁盘阵列存储空间由/**public1** 和/**public**组成，
/public1 也可创建目录，存放大数据。

- 6、**CUDA toolkit** 开发工具包安装在/**usr/local/cuda**下。

使用开发工具包的用户在~/.**bashrc**里，应添加环境变量：

```
#####CUDA PATH#####  
export PATH= /usr/local/cuda/bin:$PATH
```

```
#####CUDA LIB#####  
export LD_LIBRARY_PATH= /usr/local/cuda/  
lib64:$LD_LIBRARY_PATH
```

- 7、用户既可以做普通的并行计算,也可以使用**GPU**计算,还可以做使用**GPU**的并行计算

- 8、有问题请联系：**3601095** lsc@ustc.edu.cn

下面也许是你要注意的：

MPICH1 已於2006年停止更新，对于现代多核CPU（例如：至强12核CPU），已不能适应。

请使用MPICH2，MPICH2支持最新的**MPI-2**接口标准，与**Mpich1**相比**MPICH2**具备更加严谨和合理的结构，可移植性和效率更好，支持**C/C++**和**Fortran**。

MPICH2的主页是：

<http://www-unix.mcs.anl.gov/mpi/mpich2/index.htm>。

在<ftp://ftp.mcs.anl.gov/pub/mpi>下有**MPICH2**的

安装和使用指南文档，分别是**User's Guide**，**Installer's Guide**

MPICH2

特点

支持mpi2.0,功能强大,效率高,使用rsh或ssh可以自由切换,缺点是仅支持以太网。

使用

设置环境变量,确认用 mpich2

运行 /public/software/mpi/ mpich2.sh

设置临时环境变量,确认使用 mpich2

或参考: /public/software/mpi/ mpich2.sh的内容,

设置环境变量,确认长期使用 mpich2

```
( #export PATH=/public/software/mpi/mpich2/bin:$PATH )
```

编译MPI程序

```
#mpicc -o hello hello.c
```

```
#mpicc -show 可以看到实际调用编译器的情况
```

手动运行

```
vim ma
```

```
node1:8
```

```
node2:8
```

```
mpdboot -f ma -n 2 -r rsh
```

```
-n node number(节点数)
```

```
mpdtrace
```

```
mpirun -machinefile ma -np 16 ./cpi
```

```
mpdallexit
```

使用pbs

脚本内容如下, vi mpich2.pbs

```
#PBS -N mpi
```

```
#PBS -l nodes=2:ppn=8
```

```
#PBS -j oe
```

```
#PBS -l walltime=1000:00:00
```

```
cd $PBS_O_WORKDIR
```

```
NP=`cat $PBS_NODEFILE|wc -l`
```

```
cat $PBS_NODEFILE | uniq -c | awk '{ printf("%s:%s\n", $2, $1); }' >ma
```

```
NNODE=`cat hostfile|wc -l`
```

```
mpdboot -f ma -n $NNODE -r rsh
```

```
mpirun -machinefile ma -np $NP mpiprogram
```

```
mpdallexit
```

提交作业

```
qsub mpich2.pbs
```

对于CPU核心较多的集群,为了快速高效地计算,推荐使用 **openmpi**

openmpi

特点

支持mpi2.0,功能强大,灵活,支持各种网络,包括以太网和infiniband,效

率高。使用rsh或ssh可以自由切换，路径可以自己标志，编译器也可以改，一个版本支持多种通讯方式。

设置环境变量

运行 /public/software/mpi/ openmpi.sh

设置临时环境变量，确认使用 OPENMPI

或参考： /public/software/mpi/ openmpi.sh的内容，

设置环境变量，确认长期使用 OPENMPI

(export PATH=/public/software/mpi/openmpi/bin:\$PATH

export LD_LIBRARY_PATH=/public/software/mpi/openmpi/lib:\$LD_LIBRARY_PATH

因为默认的openmpi编译出来的库为动态库，所有要设置LD_LIBRARY_PATH变量。如果想要不设，在configure openmpi时加上 --disable-shared --enable-static 选项)

编译MPI程序

```
#mpicc -o hello hello.c
```

```
#mpicc -show 可以看到实际调用编译器的情况
```

手动运行：

-machinefile ma 文件写出在哪些节点上执行该命令，格式如下

```
#vim ma
```

```
node1 slots=8
```

```
node2 slots=8
```

(也可将node1和node2分别写8行)

slots=8代表每个结点上有8个cpu

```
mpirun -np 16 -machinefile ma --prefix /data/software/mpi/openmpi
```

```
--mca pls_rsh_agent rsh --mca btl self,tcp yourprogram
```

```
-np 16 起16个进程
```

```
--prefix /data/software/mpi/openmpi
```

如果有两个以上节点，告诉远端机器openmpi的安装路径

如果configure openmpi 时加 --enable-mpirun-prefix-by-default 则可不加--prefix

```
--mca pls_rsh_agent rsh 告诉节点间通讯用rsh
```

```
--mca btl self,tcp 使用以太网tcp/IP通讯
```

self,sm 当单节点运行的时候，使用内存通讯，效率高

self,openib 有infiniband设备时，使用IB通讯

```
--mac btl_tcp_if_include eth0 以太网通讯时用eth0通讯
```

使用pbs

脚本内容如下，vi mpich.pbs

```
#PBS -N mpi
```

```
#PBS -l nodes=2:ppn=8
```

```
#PBS -j oe
```

```
#PBS -l walltime=1000:00:00
```

```
cd $PBS_O_WORKDIR
```

```
NP=`cat $PBS_NODEFILE|wc -l`
```

```
mpirun -machinefile $PBS_NODEFILE -np $NP --prefix  
/data/software/mpi/openmpi --mca pls_rsh_agent rsh --mca btl  
self,tcp ./yourmpiprogram
```

注意：叙述中忽略了文件和目录的具体版本号，应在实际使用的环境中确认。

软件程序：

其他一些常用的软件和程序安装在 **/pub/software/**目录下，请先查看此目录中的应用程序再决定是否请管理员或自己安装其它应用程序。