

高性能计算集群用户操作手册

目录

1. 集群说明.....	2
1.1. 集群硬件配置.....	2
1.2. 编译环境配置.....	2
2. 集群登陆.....	3
2.1. 登录.....	3
2.2. 更改用户自身密码.....	7
2.3. vnc 图形界面登陆（建议图形使用）.....	7
3. 集群上传下载文件.....	9
4. 集群使用.....	10
4.1. 作业调度简介（命令行方式提交作业）.....	10
4.2. PBS 常用 3 条命.....	11
qsub 命令.....	11
qstat 命令.....	11
qdel 命令.....	13
4.3. 作业提交使用示例.....	13
4.4 使用网页界面提交作业.....	15
4.5 集群现有列队配置.....	15
5. date 附录：Linux 基本命令.....	16

集群配备有正版的全套 Intel 编译环境及开源的 GCC 编译器, MPI 支持 intel mpi, openmpi 等, 推荐使用 intel 编译环境。配置环境的脚本在/opt/env/目录下。用户可根据需要将配置文件放置在自己的登录启动设置文件中。

2. 集群登陆

2.1. 登录

集群地址，内网： 10.0.10.119，公网： 159.226.170.100，SSH 端口： 2000.

如果用户工作电脑系统是 LINUX, 则可直接使用 SSH 命令登录，如果是 windows 系统，需要安装 SSH 的软件。Xmanager 是一个 SSH 的登录工具，若需要用图形，请用 VNC 软件进行图形界面操作，VNC 图形界面操作无法上传下载文件，所以 VNC 软件和 Xmanager 软件互补使用，其中，Xmanager 可以用来上传或下载文件，xmanager 登录后，若网络断开，或者电脑关机，xshell 关闭等操作会结束任务，所以需要图形的作业可以用 VNC 登录。

安装 xmanager

点击安装程序后

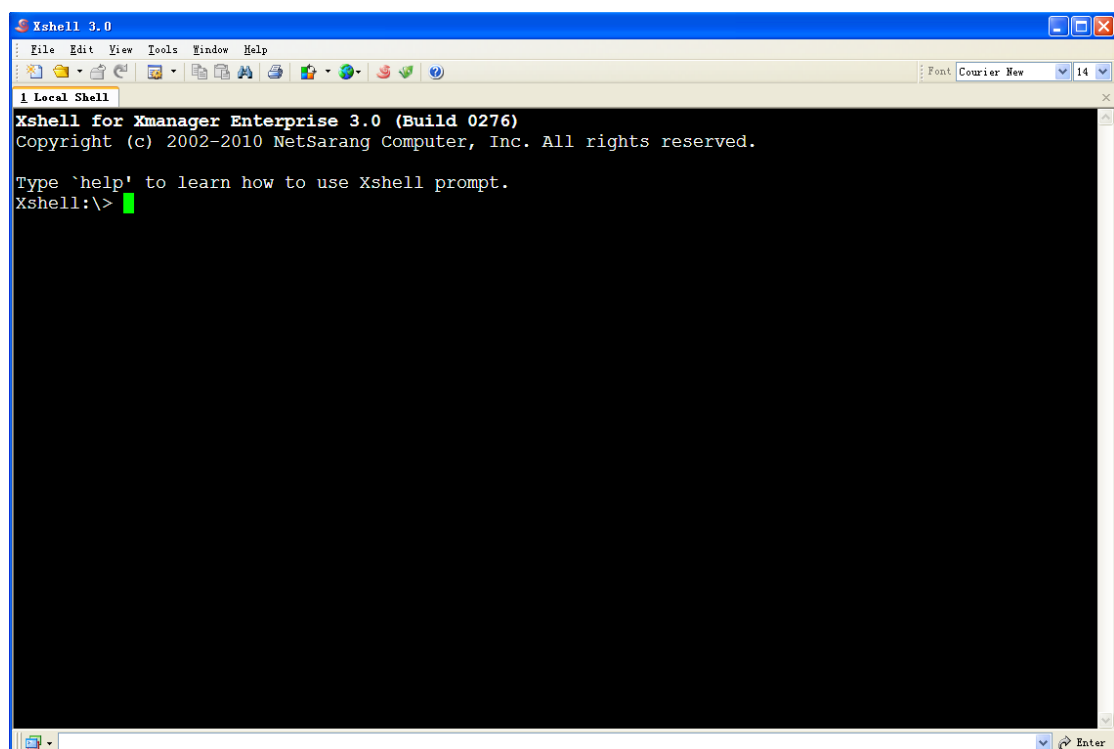


直接默认下一步安装即可

安装完后点击 xshell 选项

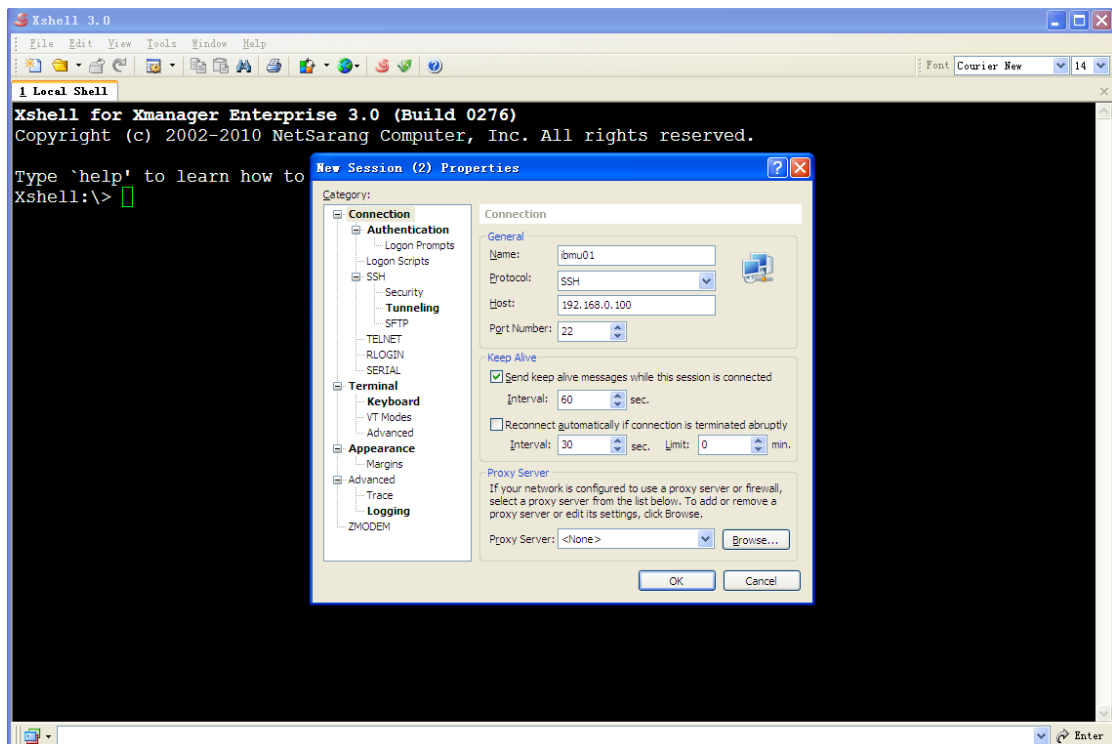


点击后如下界面



添加主机

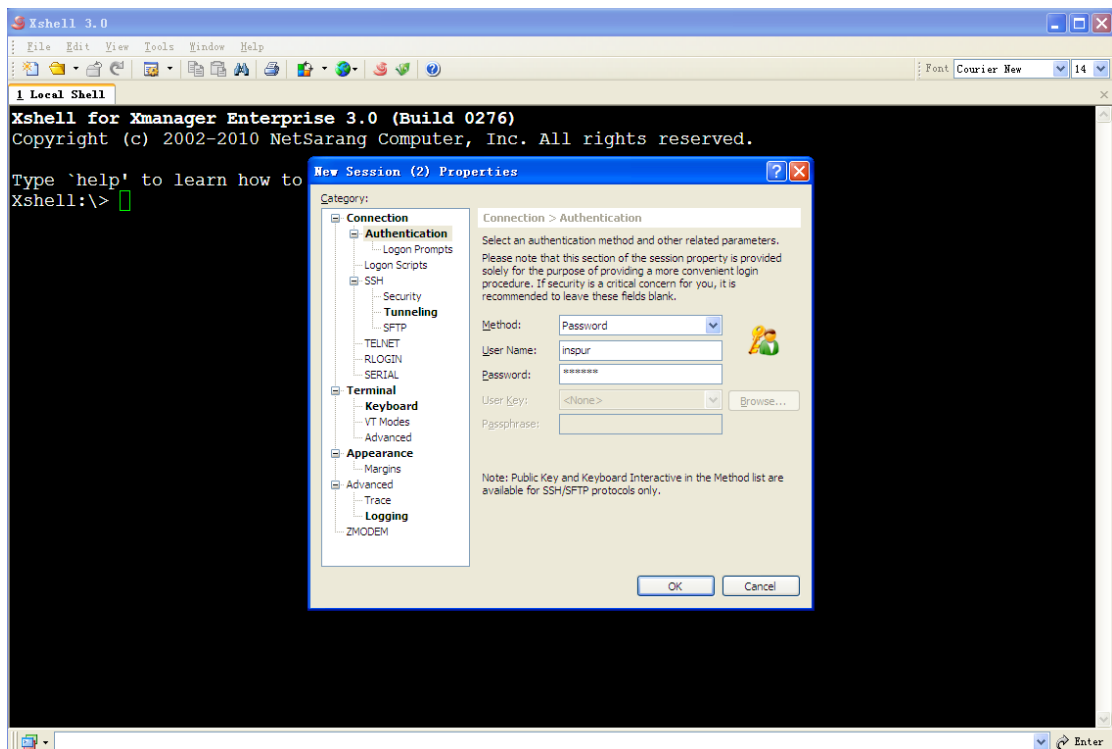
点击 file 里的 new 选项



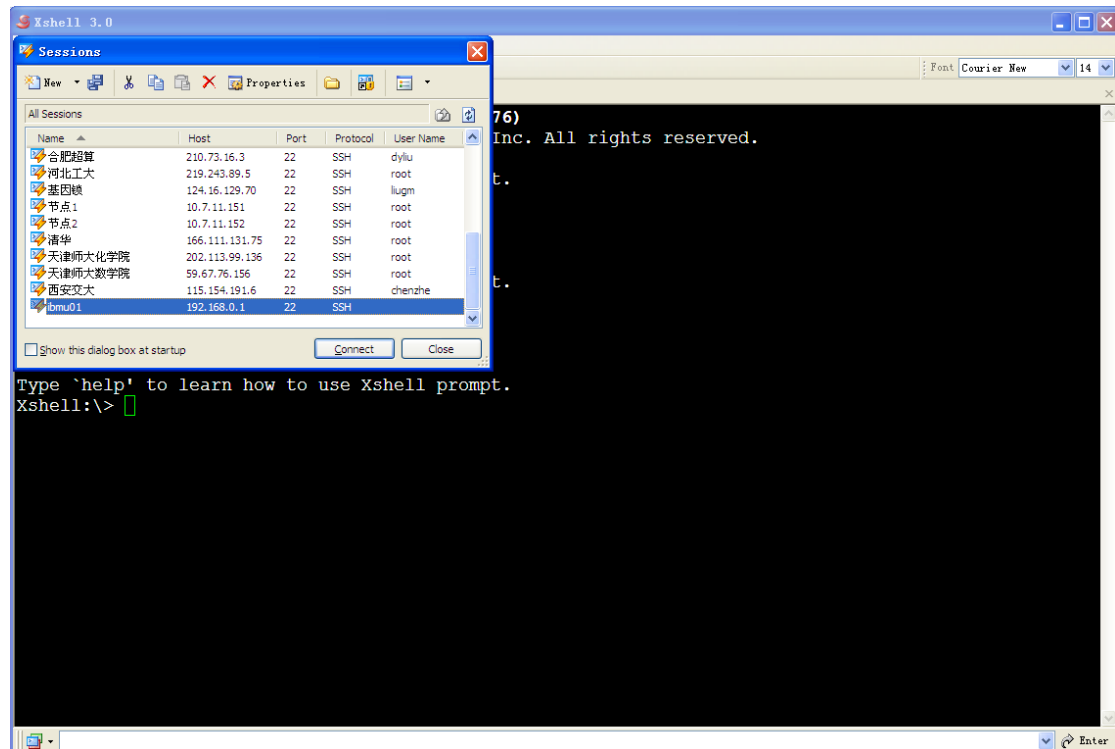
在此界面下，Connection 选项里，Name 里随便填写一个名字用来识别你所添加的机器即可，

Host 选项填写远程主机的 IP 地址

然后点击 Authentication 选项

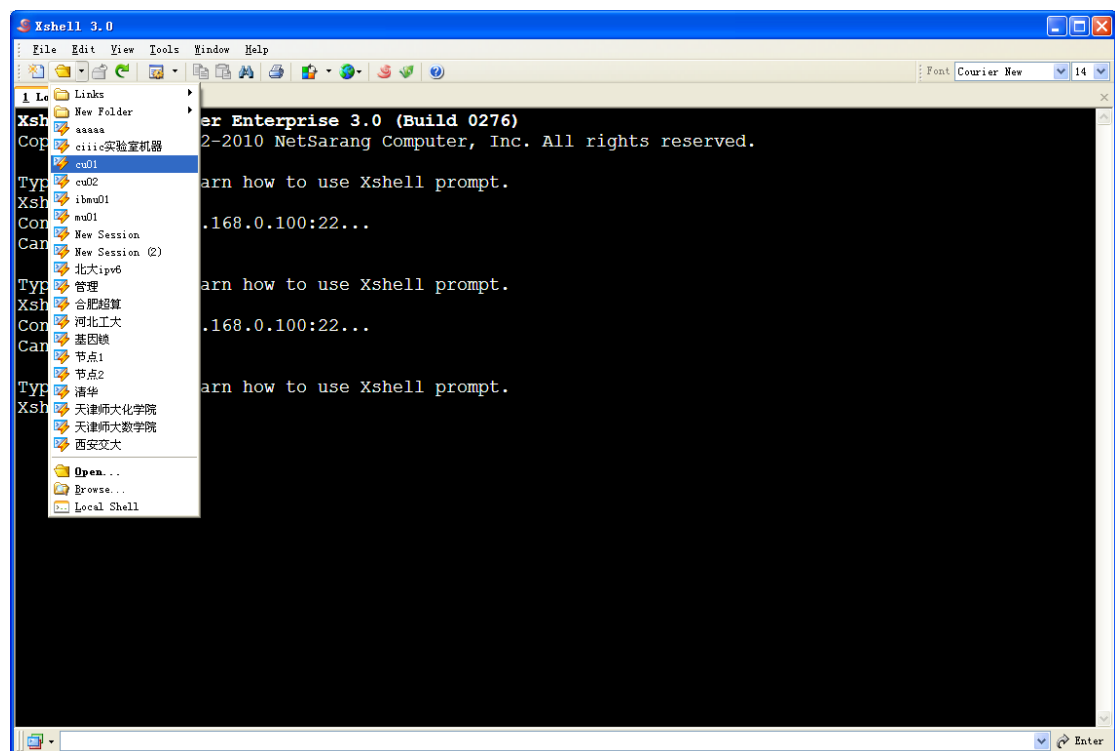


此选项里，user Name 填写登陆用户名，password 填写登陆密码，填完后点击 OK，添加主机完毕



直接点解 connect 即可连上远程主机的 shell 里

以后连接主机，直接点击 open 选项里所保存的主机即可直接登陆



2.2. 更改用户自身密码

第一次登陆后，请更改自己的用户密码

账户密码修改

在 mu02 的节点上运行 yppasswd 命令

```
[zhangsan@mu02 ~] $yppasswd
```

Old password:

New password:

Retype password:

password update Success!

2.3. vnc 图形界面登陆（建议图形使用）

VNC 图形登录后，可以打开图形作业，并且随着 vnc 界面的关闭，相关任务不会退出，下次登录时还是在关闭之前的页面。

Vnc 服务端设置

在第一次用 vnc 图形界面登陆时，先用 xshell 登陆到字符界面下设置本账户的 vnc 图形界面，

Xshell 登陆后先直接运行 vncserver 开启 vnc 功能，此时会提示输入密码，注意此密码是 vnc 的密码不是账户密码。

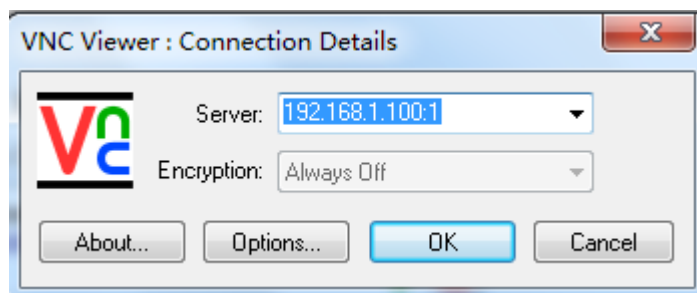
```
[inspur@mu01 ~]$ vncserver
You will require a password to access your desktops.
Password:
Verify:
New 'mu01:1 (inspur)' desktop is mu01:1
Creating default startup script /lustre/home/inspur/.vnc/xstartup
Starting applications specified in /lustre/home/inspur/.vnc/xstartup
Log file is /lustre/home/inspur/.vnc/mu01:1.log
```

用 vncserver -list 查看分配到的 vnc 号

```
[inspur@mu01 ~]$ vncserver -list  
  
TigerVNC server sessions:  
  
X DISPLAY #      PROCESS ID  
:1          53355  
[inspur@mu01 ~]$
```

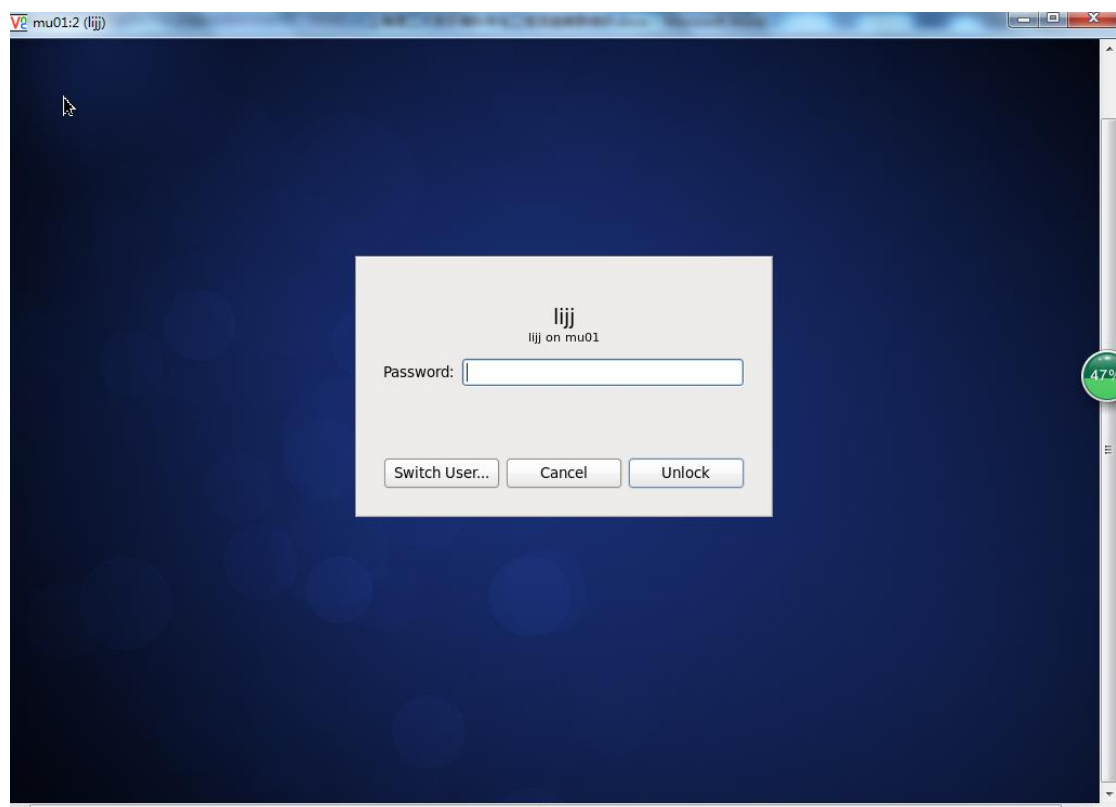
如下记住：1

然后在 PC 端打开 vnc 工具输入 ip 地址加：1 的格式来登陆 vnc 图形



登陆 vnc，输入运算的程序后，可以关闭此窗口，命令不被打断，下次登陆后，还停留在此界面。

若长时间不操作 vnc 图形桌面后，再次操作会锁定屏幕，需要输入账号密码登陆，注意是账号密码不是 vnc 密码了



备注（结束 vnc 功能，`vncserver -kill :1`

vnc 更改 vnc 密码 `vncpasswd`，按提示输入两次密码即可）

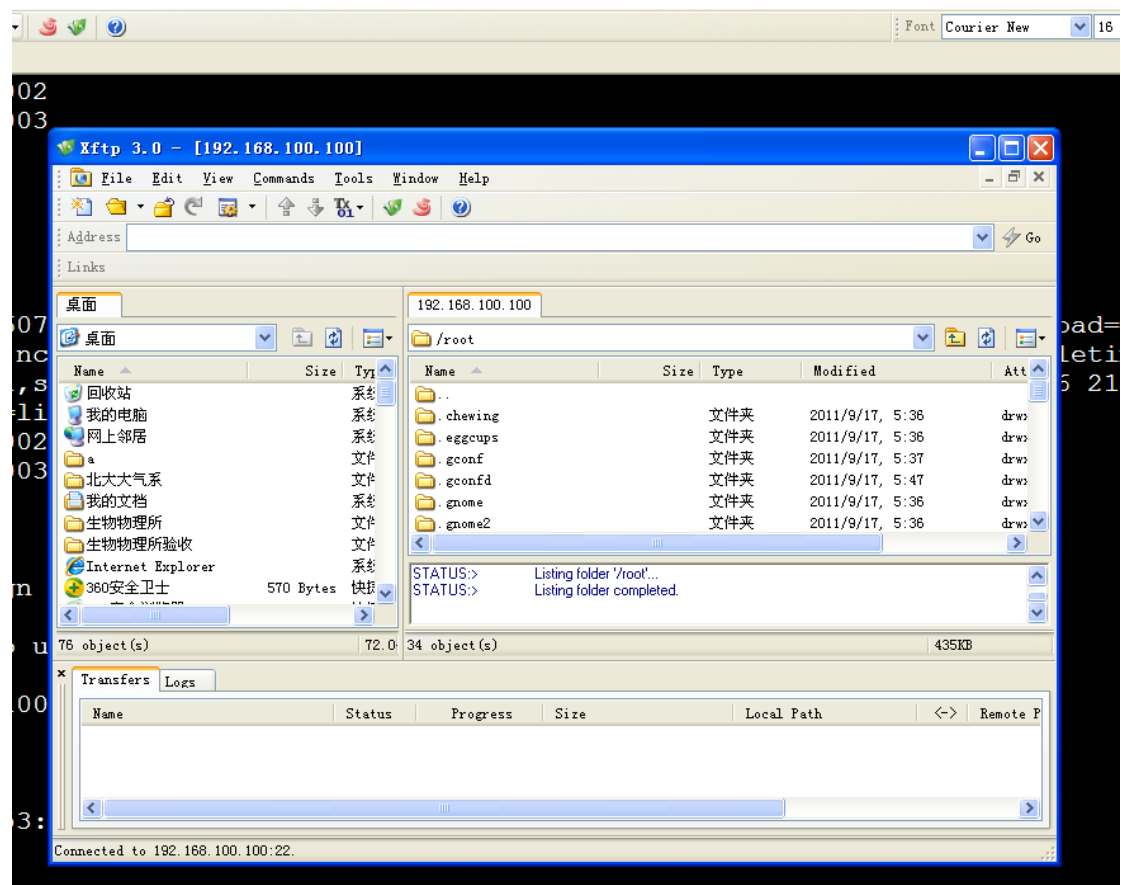
每次要用 vnc 图形登陆，先用 xshell 登陆后，用 `vncserver -list` 查看是否已经开启 vnc 功能，避免占用多个 vnc 号

每运行一次 `vncserver` 命令即可得到一个 vnc 端口号，可以登录不同的端口号，进入多个 vnc 图形窗口。

3. 集群上传下载文件

登陆 xshell 后，

点击绿色的 new file transfer 按钮打开 xftp 工具



右键单击需要上传的文件或者文件夹，选择 `upload` 即可把文件上传到用户指定目录下。

用户 HOME 目录的概念

Linux 系统对用户文件权限分配很严格，普通用户只对自己家目录下的空间有写

权限，普通用户无权查看和删除其余用户的文件，也无法删除系统文件。

集群的 HOME 目录在/public/home 目录下，以自己用户名命名的文件夹，比如 test 用户的家目录就是/public/home/test。

集群中各节点要保持数据和用户的统一，所以集群中的所有节点的/public 目录其实是由 IO 节点搭建的共享并行文件系统。

4. 集群使用

4.1. 作业调度简介（命令行方式提交作业）

PBS(Portable Batch System)最初由 NASA 的 Ames 研究中心开发，主要为了提供一个能满足异构计算网络需要的软件包，用于灵活的批处理，特别是满足高性能计算的需要，如集群系统、超级计算机和大规模并行系统。PBS 的主要特点有：代码开放，免费获取；支持批处理、交互式作业和串行、多种并行作业，PBS 的目前包括 openPBS, PBS Pro 和 Torque 三个主要分支，本集群主要用的 tsce 是 torque+maui。

作业系统有三个主要组成模块，SCHEDULER PBS_SERVER PBS_MOM

作业管理系统通过这三个模块间的相互工作来完成对节点的监控和分配任务。

Scheduler 一直在检测节点状态，主要检测节点的 CPU 占用率,查看节点是否空闲。

并把结果每 10S 传送给 PBS_SERVER,当有一个作业被提交后，

PBS_SERVER 就会根据 Scheduler 返回的结果将作业分配到空闲的节点上，

PBS_mom 接收到运算命令就会告知节点开始计算。如果没有空闲节点，作业就一直在排队

作业管理软件可以自动且合理地分配资源，避免各个用户之间抢占 CPU 资源，造成机器死机。

用 qnodes 命令可以查看通过作业调度使用的集群情况

```

[root@mu01 ~]# qnodes
cu01
state = free
np = 24
ntype = cluster
status = rectime=1438592010, varattr=, jobs=, state=free, size=40482444kb:50367172kb, netload=44968754, gres=, loadave=0.45
, ncpus=24, physmem=132106460kb, availmem=192893368kb, totmem=197642452kb, idletime=0, nusers=1, nsessions=1, sessions=3939, uname
=Linux cu01 2.6.32-431.el6.x86_64 #1 SMP Sun Nov 10 22:19:54 EST 2013 x86_64, opsys=linux
mom_service_port = 15002
mom_manager_port = 15003
gpus = 0

cu02
state = free
np = 24
ntype = cluster
status = rectime=1438592010, varattr=, jobs=, state=free, size=40483148kb:50367172kb, netload=44034836, gres=, loadave=0.17
, ncpus=24, physmem=132106460kb, availmem=192992480kb, totmem=197642452kb, idletime=0, nusers=1, nsessions=1, sessions=3937, uname
=Linux cu02 2.6.32-431.el6.x86_64 #1 SMP Sun Nov 10 22:19:54 EST 2013 x86_64, opsys=linux
mom_service_port = 15002
mom_manager_port = 15003
gpus = 0

```

4.2. PBS 常用 3 条命

qsub 命令

—用于提交作业脚本

命令格式：

```

qsub [-a date_time] [-c interval] [-C directive_prefix]
[-e path] [-l] [-j join] [-k keep] [-l resource_list] [-m mail_options]
[-M user_list][-N name] [-o path] [-p priority] [-q destination] [-r c]
[-S path_list] [-u user_list][-v variable_list] [-V]
[-W additional_attributes] [-z]
[script]

```

参数说明：因为所采用的选项一般放在 **pbs** 脚本中提交，所以具体见 PBS 脚本选项。

例：# qsub vasp.pbs 提交某作业，系统将产生一个作业号

```

[root@mu01 ~]# echo "sleep 300"|qsub
0.mu01

```

qstat 命令

用于查询作业状态信息

命令格式：qstat [-f][-a][-i] [-n][-s] [-R] [-Q][-q][-B][-u]

参数说明：

- f jobid 列出指定作业的信息
- a 列出系统所有作业
- i 列出不在运行的作业
- n 列出分配给此作业的结点
- s 列出队列管理员与 scheduler 所提供的建议
- R 列出磁盘预留信息
- Q 操作符是 destination id, 指明请求的是队列状态
- q 列出队列状态, 并以 alternative 形式显示
- au userid 列出指定用户的所有作业
- B 列出 PBS Server 信息
- r 列出所有正在运行的作业
- Qf queue 列出指定队列的信息
- u 若操作符为作业号, 则列出其状态。

若操作符为 destination id, 则列出运行在其上的属于 user_list 中用户的作业状态。

例: # qstat -f 211 查询作业号为 211 的作业的具体信息。

例如运行 qstat -an

```
[lijj@mu01 fluenttest]$ qstat -an
```

Job ID	Username	Queue	Jobname	SessID	NDS	TSK	Req'd Memory	Req'd Time	Elap Time	S	State
20.mu01	lijj	batch	test	--	8	128	--	7200	--	R	状态

cu08/15+cu08/14+cu08/13+cu08/12+cu08/11+cu08/10+cu08/9+cu08/8+cu08/7+cu08/6+cu08/5+cu08/4+cu08/3+cu08/2+cu08/1+cu08/0+cu07/15+cu07/14+cu07/13+cu07/12+cu07/11+cu07/10+cu07/9+cu07/8+cu07/7+cu07/6+cu07/5+cu07/4+cu07/3+cu07/2+cu07/1+cu07/0+cu06/15+cu06/14+cu06/13+cu06/12+cu06/11+cu06/10+cu06/9+cu06/8+cu06/7+cu06/6+cu06/5+cu06/4+cu06/3+cu06/2+cu06/1+cu06/0+cu05/15+cu05/14+cu05/13+cu05/12+cu05/11+cu05/10+cu05/9+cu05/8+cu05/7+cu05/6+cu05/5+cu05/4+cu05/3+cu05/2+cu05/1+cu05/0+cu04/15+cu04/14+cu04/13+cu04/12+cu04/11+cu04/10+cu04/9+cu04/8+cu04/7+cu04/6+cu04/5+cu04/4+cu04/3+cu04/2+cu04/1+cu04/0+cu03/15+cu03/14+cu03/13+cu03/12+cu03/11+cu03/10+cu03/9+cu03/8+cu03/7+cu03/6+cu03/5+cu03/4+cu03/3+cu03/2+cu03/1+cu03/0+cu02/15+cu02/14+cu02/13+cu02/12+cu02/11+cu02/10+cu02/9+cu02/8+cu02/7+cu02/6+cu02/5+cu02/4+cu02/3+cu02/2+cu02/1+cu02/0+cu01/15+cu01/14+cu01/13+cu01/12+cu01/11+cu01/10+cu01/9+cu01/8+cu01/7+cu01/6+cu01/5+cu01/4+cu01/3+cu01/2+cu01/1+cu01/0

通常作业状态:

- Q 作业排队
- R 作业执行
- C 作业清除
- E 作业退出

- H 作业挂起

qdel 命令

用于删除已提交的作业

命令格式: `qdel [-W 间隔时间] 作业号`

命令行参数:

例: `# qdel 211` 删除作业号为 211 的作业

4.3. 作业提交使用示例

作业运算注意:

1 可以用 `pbsnodes/qnodes` 来查看目前系统资源情况。

```
[inspur@mu01 ~]$ pbsnodes
```

Down 状态是服务器和客户端连接不上,无法提交作业,需要检查计算节点,通过在相应的计算节点上执行 `service pbs_mom restart` (管理员权限)。

Free 状态,可以提交作业的节点

Offline 状态:无法提交作业,通讯正常,需要用 `pbsnodes -c “节点名”` 来使其变为 free 状态,

若不希望某节点能提交作业,可以用 `pbsnodes -o “节点名”`,使其无法提交作业。

PBS 脚本,示例如下。

```
#PBS -N test    (作业的名字)
```

```
#PBS -e /home/inspur/error    ###错误输出路径,若不指定,在当前目录会生成作业名加号的 e 文件
```

```
#PBS -o /home/inspur/work/stand    #####若不指定,在当前目录会生成作业名加号的 o 文件
```

```
#PBS -l nodes=2:ppn=12    (定义启动 2 个节点,每个节点启动 12 个核,必须指
```

定)

#PBS -l walltime=120:00:00 (作业运行最大时间 120 个小时, 估算自己的作业最大运行时间, 这里指定一下。不指定默认最大 1200 小时)

#PBS -q batch (作业提交到 batch 队列, 默认即可)

#PBS -V (定义环境变量范围)

#PBS -S /bin/bash (使用 bash)

cd \$PBS_O_WORKDIR (进入到你的工作目录, 当前作业提交的目录)

EXEC=/lustre/inspur/run (run 文件为你编译好的可执行文件, 注意更改)

#####脚本部分#####

NP=`cat \$PBS_NODEFILE | wc -l` #####计算 cpu 使用核数

NN=`cat \$PBS_NODEFILE | sort | uniq | tee /tmp/nodes.\$\$ | wc -l`#####计算使用节点数

cat \$PBS_NODEFILE > /tmp/nodefile.\$\$

mpirun -genv I_MPI_DEVICE rdssm -machinefile nodefile.\$\$ -n \$NP \$EXEC

rm -rf nodefile.\$\$

以下是一个 ansys 的作业脚本

#!/bin/bash

#PBS -N ansysmpi

#PBS -l nodes=3:ppn=6

#PBS -q batch

export MPI_WORKDIR=\$PBS_O_WORKDIR

export MPI_REMSH=ssh

export MPI_IC_ORDER=IBV:TCP

cd \$PBS_O_WORKDIR

cat \$PBS_NODEFILE > /\$PBS_O_WORKDIR/nodefile

```

machine1=`uniq -c /$PBS_O_WORKDIR/nodefile | awk 'BEGIN
{H=""}{if(H==""){ H=$2":"$1 } else {H=H":"$2":"$1 } } END {print H}'`
echo $machine1 >/$PBS_O_WORKDIR/machine
sed -i 's/c/ibc/g' /$PBS_O_WORKDIR/machine

date

/lustre/soft/ansys/v140/ansys/bin/ansys140 -b -dis -pp -machines `cat
/$PBS_O_WORKDIR/machine` -i /$PBS_O_WORKDIR/vm1.dat -o
/$PBS_O_WORKDIR/vm1.out

```

4.4 使用网页界面提交作业

使用网页界面登录集群的方法，可通过本地计算机，打开浏览器，输入远端节点 ip:8080 打开网页界面，登录帐号密码即系统帐号密码。

4.5 集群现有列队配置

集群目前配置了 big, mid, small,knl 三个作业队列，将根据计算需求作相应调整。

现有队列配置可以通过命令 `qstat -q` 来查看。例如目前情况：

```
[root@mu01 ~]# qstat -q
```

```
server: 192.168.1.100
```

Queue	Memory	CPU Time	Walltime	Node	Run	Que	Lm	State
small	--	--	96:00:00	5	1	0	15	E R
psr	--	--	--	60	0	0	--	E R
cal	--	--	--	--	0	0	--	D R
knl	--	--	--	--	0	0	--	E R
batch	--	--	--	--	0	0	--	E R
big	--	--	72:00:00	32	1	0	50	E R
pmtest	--	--	--	--	0	0	--	E R
mid	--	--	168:00:0	15	4	0	10	E R
					6	0		

Walltime 为执行时间，Node 为是大节点数，Run 当前作业数，Lm 为作业数限制。

集群节点情况可通过执行 `shownodes` 来查看。

5. date 附录：Linux 基本命令

目录操作

名称：cd

语法：cd [directory]

说明：把当前工作目录转到“directory”指定的目录。

实例：进入目录 `/usr/bin/`：

`cd /usr/bin`

名称：ls

语法：ls [options] [pathname-list]

说明：显示目录内的文件名和“pathname-list”中指定的文件名

实例：列出目前工作目录下所有名称是 `s` 开头的文件：

`ls s*`

名称：pwd

语法：pwd

说明：显示当前目录的绝对路径。

名称：mkdir

语法：mkdir [options] dirName

说明：创建名称为 `dirName` 的子目录。

实例：在工作目录下，建立一个名为 `AA` 的子目录：

`mkdir AA`

名称：rmdir

语法：rmdir [-p] dirName

说明：删除空的目录。

实例：将工作目录下，名为 `AA` 的子目录删除：

`rmdir AA`

文件操作

名称: cp

语法: cp [options] file1 file2

说明: 复制文件 file1 到 file2。

常用选项: -r 整个目录复制

实例: 将文件 aaa 复制(已存在), 并命名为 bbb:

```
cp aaa bbb
```

名称: mv

语法: mv [options] source... directory

说明: 重新命名文件, 或将数个文件移至另一目录。

范例: 将文件 aaa 更名为 bbb:

```
mv aaa bbb
```

名称: rm

语法: rm [options] name...

说明: 删除文件及目录。

常用选项: -f 强制删除文件

实例: 删除除后缀名为.c 的文件

```
rm *.c
```

名称: cat

语法: cat[options] [file-list]

说明: 在标准输出上连接、显示文件列表 file-list 里的文件

实例 1: 显示 file1 和 file2 的内容

```
cat file1 file2
```

实例 2: 将 file1 和 file2 合并成 file3

```
cat file1 file2 > file3
```

名称: more

语法: more[options] [file-list]

说明: 在标准输出上连接、分页显示文件列表 file-list 里的文件

实例: 分页显示文件 AAA

```
more AAA
```

名称: head

语法: head[options] [file-list]

说明: 显示文件列表 file-list 中的文件的起始部分, 默认显示 10 行;

实例: 显示文件 AAA 起始部分

```
head AAA
```

名称: tail

语法: tail[options] [file-list]

说明: 显示文件列表 file-list 中的文件的尾部; 默认显示 10 行;

实例: 显示文件 AAA 尾部

```
tail AAA
```

名称: ln

语法: ln[options] existing-file new-file

```
ln[options] existing-file-list directory
```

说明: 为 “existing-file” 创建链接, 命名为 new-file

在 directory 目录, 为 existing-file-list” 中包含的每个文件创建同名链接

常用选项: -f 不管 new-file 是否存在, 都创建链接

-s 创建软链接

实例 1: 建立软连接 temp.soft, 指向 Chapter3

```
ln -s Chapter3 temp.soft
```

实例 2: 为 examples 目录下的所有文件和子目录建立软连接

```
ln -s ~/linuxbook/examples/* /home/fanodelty/linuxbook/examples
```

名称: chmod

语法: chmod [option] mode file-list

说明: 改变或设置参数 file-list 中的读、写或执行权限

实例: 添加文件 job 的可执行权限

```
chmod +x job
```

名称: tar

语法: tar [option] [files]

说明: 备份文件。可用来建立备份文件, 或还原备份文件。

实例 1: 备份 test 目录下的文件，并命名为 test.tar.gz，可执行命令：

```
tar -zcvf test.tar.gz test
```

实例 2: 解压缩相关的 test.tar.gz 文件，可执行命令：

```
tar -zxvf test.tar.gz
```

其他

名称: echo

语法: echo \$variable

说明: 显示变量 variable 的值。

实例 1: 显示当前用户路径 PATH 的值

```
echo $PATH
```

名称: ps

语法: \$ps [options]

说明: 用于查看当前系统中的活跃进程

实例 1: 显示当前所有进程

```
ps -aux
```

名称: kill

语法: \$kill [-signal] pid

说明: 终止指定进程

实例 1: 终止 1511 号进程

```
kill 1511
```

附录 2: Vi 使用

简要使用流程

使用 "vi [选项] [文件 ..]" 命令打开要编辑的文件

使用 "方向键" 浏览文件

按下 "i" 进入编辑模式

编辑

按 "Esc" 键退出编辑模式

输入 ":w" 回车保存，再输入":q" 回车退出。或者直接输入 ":wq" 回车，代表保存并退出

两种操作模式

编辑模式：对文本进行编辑处理

命令模式：接收按键指令执行操作，如复制、粘贴、搜索、替换、保存、另存为等

编辑模式

i: 进入编辑模式

a: 进入编辑模式，将光标向后移动一位

o: 进入编辑模式，在光标处插入一个空行

r: 按下 r 键，再按任意字符键，将光标所在处的字符替换成新输入的字符

Esc: 退出编辑模式

命令模式

移动光标

↑或 k: 把当前光标向上移动一行，保持光标的列位置。

↓或 j: 把当前光标向下移动一行，保持光标的列位置。

→或 l: 把当前光标向右移动一个字符。

←或 h: 把当前光标向左移动一个字符。

\$: 把当前光标移动到该行行末。

^: 把当前光标移动到该行行首。

w: 把当前光标移动到该行的下一个字的首字符上。

b: 把当前光标移动到该行的上一个字的首字符上。

e: 把当前光标移动到该行的该字的末尾字符上。

^F: 向前滚动一整屏正文。

^D: 向下滚动半个屏正文。

^B: 向后滚动一整屏正文。

^U: 向上滚动半个屏正文。

搜索与替换

/word: 从光标处开始，向后搜索文本中出现 word 的字符串

?word: 从光标处开始，向前搜索文本中出现 word 的字符串

:1,\$s/word1/word2/g: 在第 1 行与最后一行之间搜索 word1,并将其替换为 word2

:n1,n2s/word1/word2/g: 在第 n1 行与第 n2 行之间搜索 word1，并将其替换为 word2

删除 (剪切)、复制与粘贴

x,X: x 为向后删除一个字符，X 为向前删除一个字符

dd: 删除光标所在行

yy: 复制光标所在行的内容

nyy: 复制光标到第 1 行的所有内容

y1G: 复制光标到第 1 行的所有内容

yG: 复制光标到最后一行的所有内容

p,P: p 为将复制或剪切的内容粘贴在光标下一行，P 为粘贴在上一行

u: 撤消上一操作

管理命令

:w:保存

:w!: 强制保存

:q:退出 vi 编辑器

:q!:强制退出

:w [文件名]:另存为..

:r[文件名]:读取另一个文档的内容，内容追加到光标所在行之后

:set nu:显示正文的行号。

:set nonu:取消行号。

:[命令]: 暂时离开 vi 编辑器，并在 shell 中执行命令