

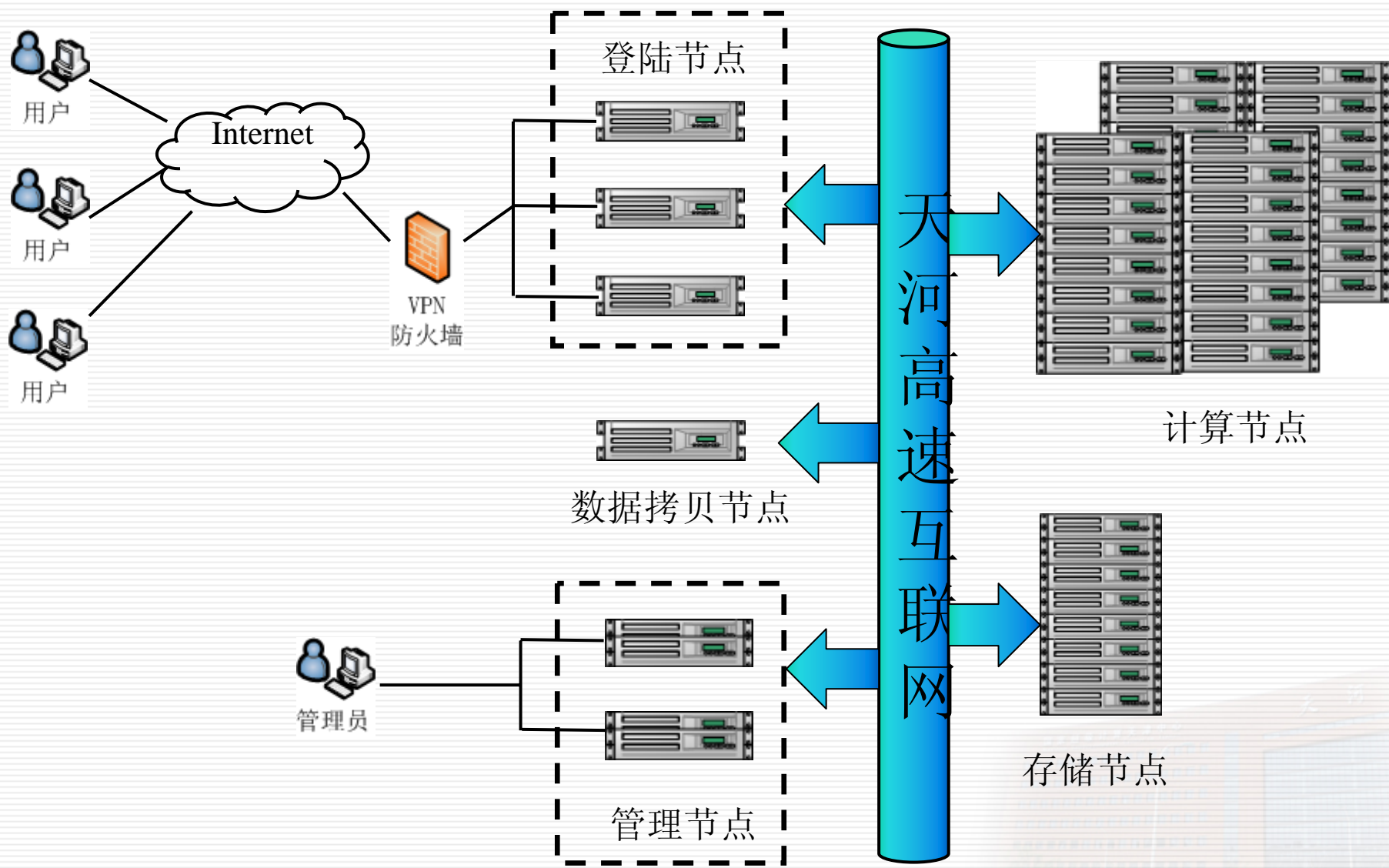
# 天河系统架构及用户使用方法

菅晓东

[jianxd@nsc-tj.gov.cn](mailto:jianxd@nsc-tj.gov.cn)

[www.nsc-tj.gov.cn](http://www.nsc-tj.gov.cn)

- 系统平台架构
- 系统及编译环境
- 系统登陆
- 作业管理
- 天河开源科学计算软件部署情况介绍
- 典型科学计算软件提交方法
- 常见问题





## ◆ 登陆节点

登陆节点为用户提供一个登陆系统的平台，用户可以通过internet网络登录VPN，然后通过ssh终端登录到登陆节点上；

麒麟操作系统，具备软件编译环境，用户可以登录的是LN1-LN3, LN8, LN9；

LN1, LN2, LN8作为windows与linux用户的登陆节点；

LN3, LN9作为windows与linux用户的数据传输节点；

**允许操作：**在登陆节点上用户可以进行软件编译与调试, 环境变量配置，

作业提交，文件编辑，结果查看等操作。

**禁止操作：**禁止用户直接在登陆节点上运行计算程序。

## ◆ 数据拷贝节点

NAS存储服务器，支持EXT3文件系统SATA硬盘直接拷贝；



## ◆ 计算节点

计算节点本身没有本地硬盘，采用ramdisk精简内核系统，挂载共享存储，具备软件运行所需的运行环境；

计算节点采用CPU+GPU的架构，其具体配置为：两个Intel Xeon 5670 六核CPU，运行频率2.93GHz，一个Nvidia M2050的GPU，24GB内存，无本地磁盘；

## ◆ 存储节点

采用分布式存储文件系统，该文件系统由多个存储节点构成，对外提供一个统一的大分区，供所有登陆节点与计算节点进行挂载。

## ◆ 文件系统

集中式管理的分布式lustre文件系统；

## ◆ 互连网络

天河高速互连网络，高通信效率，低通信延迟；

## ◆ 作业管理系统

基于开源slurm开发

- 系统平台架构
- 系统及编译环境
- 系统登陆
- 作业管理
- 天河开源科学计算软件部署情况介绍
- 典型科学计算软件提交方法
- 常见问题



◆ 共享目录

`/vol-th、 /vol6、 /WORK`

◆ 用户根目录

`/vol-th/home/、 /vol6/home、 /WORK /home`

◆ 常用软件安装目录

`/vol-th/software、 /vol6/software`

◆ 常用动态链接库目录

`/vol-th/lib、 /vol-th/lib/mk1em64t、 /vol6/lib、  
/vol6/lib/mk1em64t`



- 在TH-1A大系统的登陆节点中，目前安装了Intel编译器和GCC编译器。用户可根据程序需求，选择相应的编译器进行编译和应用程序开发，由于TH-1A大系统广泛采用了Intel的CPU，因此在编译中除特定需要，建议用户首选Intel编译器。另外，在TH-1A大系统的登陆节点上还提供了MPI并行编译环境，以及针对GPU的CUDA编译环境。下面将分别具体介绍各编译器及编译环境。
- ◆ Intel编译器:用户在登陆节点上使用Intel编译器进行程序编译时需添加如下环境变量声明

### 11.1.059

```
source  
/opt/intel/Compiler/11.1/059/bin/intel64/iccvars_intel64.sh  
  
source  
/opt/intel/Compiler/11.1/059/bin/intel64/fortvars_intel64.sh
```

### Intel2013

```
source /opt/intel/composer_xe_2013.0.079/bin/iccvars.sh intel64
```





- 用户在计算节点上提交作业运行时，如需要调用Intel编译器的动态库，则需要添加如下环境变量声明：
- `export LD_LIBRARY_PATH=/vol-th/lib:$LD_LIBRARY_PATH`
- Intel 11.1对应的mk1安装路径为/opt/intel/Compiler/11.1/059/mkl，用户可以使用该目录下的lib/em64t的mkl库。
- 用户在使用mkl库进行程序编译时需添加如下环境变量声明：
- `source /opt/intel/Compiler/11.1/059/mkl/tools/environment/mklvarsem64t.sh`
- 用户在在计算节点上提交作业运行时，如需要调用mkl动态库，则需要添加如下环境变量声明：
- `export LD_LIBRARY_PATH=/vol-th/lib/mklem64t/:$LD_LIBRARY_PATH`

## ◆ GCC编译器

4.4.7            系统自带

4.6.1            /vol-th/software/gcc-4.6.1/

```
export PATH=/vol-th/software/gcc-4.6.1/bin:$PATH
```

```
export LD_LIBRARY_PATH=
```

```
                    /vol-th/software/gcc-
```

```
4.6.1/lib64:$LD_LIBRARY_PATH
```

## ◆ CUDA编译器

CUDA编译环境包含三个部分，编译器、SDK和设备驱动；

目前计算节点CUDA编译环境和设备驱动，均已经更新至CUDA5.0；

用户可以选择相应的编译器，路径在 /vol-th/software/cuda；

用户使用CUDA进行程序编译时需添加如下环境变量声明：（以cuda5.0为例

```
export PATH=/vol-th/software/cuda/cuda-5.0/bin:$PATH
```

```
export LD_LIBRARY_PATH=/vol-th/software/cuda/cuda-5.0/lib64:$LD_LIBRARY_PATH
```



## ◆ mpi编译器

TH-1A大系统包括两种基本编译环境，Intel编译器和gcc编译环境，因此为了适应用户需要，系统的mpi编译环境包括两部分，即底层分别用Intel编译器和gcc编译器编译的mpi版本。

TH-1A采用了自主互连的高速网络，因此底层mpi为自主实现，分别基于Intel和GCC编译器进行编译。用户使用天河系统提供的MPI进行并行编译可以充分发挥天河高速网的性能，提供并行效率，这里给用户推荐使用基于intel编译器编译的mpi（如mpi\_1.2.1\_intel\_11.1, mpi\_1.4.1\_intel\_11.1）。

如果用户的程序有特定的mpi版本需求，用户也可以在自己的根目录下安装所需要的mpi。

天河系统提供的mpi编译器在/vol-th/software/mpi/目录下，目录命名格式为mpi\_aaa\_BBB\_ccc\_ddd，其中aaa为使用的mpi版本，BBB为底层采用的Intel编译器或者GCC编译器，ccc为相应编译器的版本，ddd为编译选项标注。



- 用户使用mpi编译器进行程序编译时需添加如下环境变量声明：（以mpi\_1.2.1\_intel\_11.1为例）

```
source /opt/intel/Compiler/11.1/059/bin/intel64/iccvars_intel64.sh
source /opt/intel/Compiler/11.1/059/bin/intel64/ifortvars_intel64.sh
export PATH=/vol-th/software/mpi/mpi_1.2.1_intel_11.1/bin:$PATH
export LD_LIBRARY_PATH=/vol-th/software/mpi/mpi_1.2.1_intel_11.1/lib:
$LD_LIBRARY_PATH
```

- 并行mpi编译环境使用注意事项：

TH-1A大系统安装了两种底层编译环境的mpi，程序如无特殊需要，推荐使用/vol-th/software/mpi/目录下的mpi。且该mpi的库均为静态库，用户不用担心动态链接库问题；

TH-1A大系统具备自主高速互连网络，并提供MPI编程环境，如用户必须使用其他版本mpi，比如openmpi1.4.8，mpich2-1.3.1等，也可以自己安装并部署；

用该mpi编译的程序，同样可以利用高速互连网络的虚拟以太网运算任务，但性能会较TH-1A自主MPI低很多。

- 系统平台架构
- 系统及编译环境
- 系统登陆
- 作业管理
- 天河开源科学计算软件部署情况介绍
- 典型科学计算软件提交方法
- 常见问题



## ◆ 前提条件

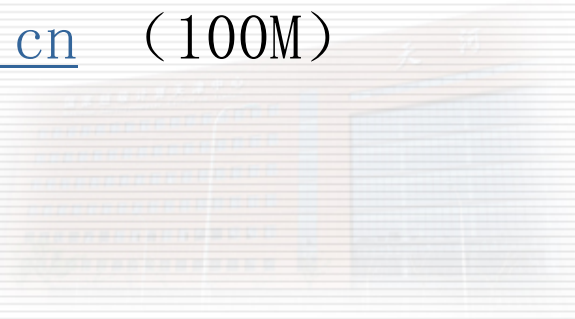
1. 拥有VPN账号和密码;
2. 拥有系统账号和密码;
3. windows系统登陆终端工具;

Putty, secureCRT, Xmanage, SSHSecureShell等等。

4. 确认用户当前网络:

联通: <https://vpn.nscg-tj.cn> (200M)

电信: <https://vpn1.nscg-tj.cn> (100M)







### ◆ Windows系统终端登陆

用户成功登陆VPN后，可以看到自己允许使用的资源。

用户可以使用ssh客户端软件（如SSH Secure Shell Client, SecureCRT, Putty）来登录系统。SSH Secure Shell Client, SecureCRT, Putty等均为免费软件，网络上均有下载。

登录时，Host Name项填写 TH-1A-LN2，以Secure Shell Client为例，登陆方式如下：







登陆之后，输入相应的用户名，密码，即可登陆至TH-1A大系统的LN2节点。

登陆LN2后，会收到如下提示：

1. Welcome to TH-1A System of NSCC-TJ.

2. If you have any problem, you can send mail to support@nscc-tj.gov.cn

之后您即可以开始编译、提交任务等操作。

## ◆文件传输

建议用户使用WinScp的sftp数据传输软件（免费软件，网络容易下载，且该软件支持断点续传）





## ◆ 用户帐号密码修改

目前系统采用LDAP进行用户管理，新创建的用户第一次登陆节点时会创建相应的工作目录。用户可以通过passwd命令修改用户密码，以ncps用户为例，举例说明如下：

```
[ncps@ln2 ~]$ passwd
Changing password for user ncps.
Enter login(LDAP) password:
New password:
Re-enter new password:
LDAP password information changed for ncps
passwd: all authentication tokens updated
successfully.
```

首先需要输入中心给分配的账户密码，之后再输入新的密码，重复输入一次后，就会显示密码更新成功。



◆ VPN登陆注意事项如下:

- 如果您安装有360安全卫士等防护软件，当有提示时，请点允许，来保证VPN的正常使用。
- VPN用户默认限制，同时使用上限为5人。（一台PC机只允许登录一个VPN账号）如果您有特殊需求，需要更多的人同时使用，请告知中心。
- 在登陆VPN后，通过软件登陆系统，进行编译、提交任务等操作时，请不要关闭浏览器或退出VPN，否则会断开连接。
- 系统默认用户登陆VPN后，如不进行任何操作（登录系统等操作），30分钟后会自动下线。（如果您通过ssh软件连接天河登陆节点则不会自动下线）
- 如果您需要退出VPN时，请点击VPN右上角的“退出”，而不要直接关闭浏览器。（关闭浏览器系统会保留该用户5分钟，如果在该段时间内超过VPN用户限制5，您将短时无法登陆）
- 您在使用中遇到任何问题可以email或电话咨询

- 系统平台架构
- 系统及编译环境
- 系统登陆
- 作业管理
- 天河开源科学计算软件部署情况介绍
- 典型科学计算软件提交方法
- 常见问题



- TH-1A大系统上的作业管理系统以CPU核作为并行作业的资源分配单位，实现并行作业的调度运行。
- TH-1A大系统中，所有在计算结点中运行的串行或并行应用程序，都必须通过资源管理系统来提交运行。
- 资源管理系统首先将用户提交的应用程序构造成作业进行排队处理，然后根据TH-1A大系统的实时运行资源状态，决定何时以及在哪些计算结点中加载应用程序的运行。
- 不同的应用程序之间不存在资源的竞争冲突，用户也可以通过作业管理系统来监控应用程序的运行。



## ◆ 分区限制

用户可以使用yhi命令查看节点的使用情况

PARTITION表示分区，TIMELIMIT表示该分区的时间限制，NODES表示节点数，STATE表示节点运行状态其中 down表示未启动，idle表示启动后处于空闲状态，alloc表示节点已经分配了作业，NODELIST为节点

```
[zhangjian@ln2%tianhe ~]$ yhi
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
TH_NEW    up 10-00:00:0    1 drain* cn11
TH_NEW    up 10-00:00:0    1  drng  cn3869
TH_NEW    up 10-00:00:0    6  drain cn[16,34,48,147,153,191]
TH_NEW    up 10-00:00:0   369  alloc cn[0-10,12-15,17-30,35-47,49-146,148-151]
TH_NEW    up 10-00:00:0    35  idle  cn[31-33,309-339,3732]
gpu_test  up 2-00:00:00    1 drain* cn3430
gpu_test  up 2-00:00:00    3  drain cn[3408-3409,3431]
gpu_test  up 2-00:00:00   38  alloc cn[3392-3394,3397-3400,3402,3404-3407,3409]
gpu_test  up 2-00:00:00   22  idle  cn[3395-3396,3401,3403,3412-3419,3422-3423]
TH_NET    up 2-00:00:00    1 drain* cn582
TH_NET    up 2-00:00:00    2  down* cn[3915,4181]
TH_NET    up 2-00:00:00    1  drng  cn3869
TH_NET    up 2-00:00:00   71  drain cn[442,583,3508-3509,3562-3563,3592-3593,3838-3841,3868,3887,3912-3913,3930-3931,3934-3935,3992-3993,4020-4021,4038-4039]
TH_NET    up 2-00:00:00  1193 alloc cn[384-435,437-441,443-445,447-581,584-585,3758-3775,3778-3783,3787-3791,3794-3813,3816-3827,3830-3837,3842-3867,3870-3871,4104-4104,4106-4113,4115-4170,4172-4178,4180,4182-4227,4229-4249,4251-4293,4295-4300]
TH_NET    up 2-00:00:00    12  idle  cn[436,446,3608,3732,3932-3933,4016,4080]
```



用户可以使用yhq查看自己提交的作业，为了保证用户的数据安全，普通用户通过yhq只能看到自己提交的作业。

```
[root@ln2~tianhe ~]# yhq
JOBID PARTITION NAME USER ST TIME NODES NODELIST(REASON)
68386 TH_NET DI-r2-26 jlou PD 0:00 25 (AssociationResourceLimit)
68389 TH_NET IV-r2-24 jlou PD 0:00 25 (AssociationResourceLimit)
68626 TH_NET rWT jlou PD 0:00 10 (AssociationResourceLimit)
68646 TH_NET rlinAB jlou PD 0:00 10 (AssociationResourceLimit)
68647 TH_NET rAB10 jlou PD 0:00 10 (AssociationResourceLimit)
68650 TH_NET rAB40 jlou PD 0:00 10 (AssociationResourceLimit)
68651 TH_NET rA1_10 jlou PD 0:00 10 (AssociationResourceLimit)
68652 TH_NET rA10 jlou PD 0:00 10 (AssociationResourceLimit)
68653 TH_NET rA_40 jlou PD 0:00 10 (AssociationResourceLimit)
68737 TH_NET H-5th-3 jlou PD 0:00 10 (AssociationResourceLimit)
68754 TH_NET rA_70 jlou PD 0:00 10 (AssociationResourceLimit)
69001 TH_NET A-4th-4 jlou PD 0:00 10 (AssociationResourceLimit)
69216 TH_NET zip_y33. yyq PD 0:00 1 (AssociationResourceLimit)
69174 TH_NET run.sh weili PD 0:00 3 (AssociationResourceLimit)
69357 TH_NET qszhang simm PD 0:00 38 (AssociationResourceLimit)
69359 TH_NET sub.sh zlin PD 0:00 32 (Resources)
69363 TH_NET mhci7-2 liutao PD 0:00 20 (Priority)
69362 TH_NET run.sh weili PD 0:00 3 (Priority)
69364 TH_NET run.sh weili PD 0:00 3 (Priority)
69366 TH_NET MAOXIONG pkucheml PD 0:00 27 (Priority)
67918 TH_NET ABF-11th jlou R 1-12:50:28 25 cn[3966-3990]
67909 TH_NET H-4th-2 jlou R 1-19:46:57 10 cn[3676-3685]
```

JOBID表示任务ID，Name表示任务名称，USER为用户，TIME为已运行时间，NODES表示占用节点数，NODELIST为任务运行的节点列表。

## 注意：

- 由于大型集群系统具备一定故障率，TH-1A大系统系统十分庞大，为了保证系统稳定性，分区中有限定任务执行时间的限制，因此建议用户为程序设立“断点”从而保证任务由于意外中断后，可以继续运算。
- 如果您的程序没有办法“续算”，而且运行时间超过2\*24小时，请联系中心技术人员。
- debug是用户调试分区，每个用户都可以使用最大2个节点24核的资源，作业时间限制为30分钟。





## ◆ 用户限制

除了上述的分区限制，目前还根据用户的申请情况，针对用户做了一定的限制，该限制主要基于用户和中心签订合同的规模。包括，最多可以使用的节点数、最多可以使用的核数、单个任务最多可以使用的节点数、单个任务最多可以使用的核数等。

用户在使用过程中，如果有超出自己合同范围内的计算规模的计算需求，请基于自己的需求，向中心提出申请，中心会根据用户需要审查后，进行一定的修改。

为了保证系统和用户数据的安全，目前普通用户不能在还没有申请资源时，就ssh链接到计算节点，只有分配了相应的计算节点资源后，才能ssh到指定计算节点。

## ◆ 磁盘配额限制

为了合理利用有限的存储资源，目前中心对用户默认进行存储软限制500G，存储硬限制1T，文件数软限制100万，文件数硬限制200万的磁盘配额限制。



## ◆ 作业提交

三种方式: yhrun      yhbatch      yhalloc

yhrun      交互作业提交方式

用户的作业命令在提交节点上运行，关闭当前登陆终端会导致程序退出

yhbatch      批处理作业提交方式

用户的作业脚本在分配的第一个节点上运行，关闭当前登陆终端不会影

响作业运行

yhalloc      分配模式

常用来先获取分配的作业节点，再进行任务提交，用户的作业脚本直接

在提交节点上运行，关闭当前登陆终端会导致程序退出



## ◆ 作业提交

yhrun 交互作业提交方式

```
yhrun -n 240 -N 20 -p TH_NET ./a.out
```

-n, 指定要运行的进程数

-N, 指定要运行的节点数

-p, 指定使用的分区

-w, 请求指定列表中的结点 -w cn[32-51]

-x, 排除指定列表中的结点 -x cn[32-51]

-J, 指定作业的名字



## ◆ 作业提交

yhbatch 批处理作业提交方式

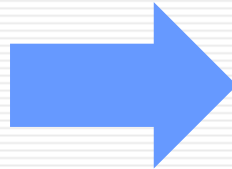
```
yhbatch -n 240 -N 20 -p TH_NET ./mpibatch.sh
```

脚本第一行务必为 “#!/bin/bash”

```
yhbatch ./test.sh
```

PBS script

```
#!/bin/bash
#PBS -l nodes=2:ppn=12
#PBS -q TH_NET
#PBS -j test
./a.out
```



Slurm

```
script
#!/bin/bash
#SBATCH -n 24
#SBATCH -N 2
#SBATCH -p TH_NET
# SBATCH -J test
./a.out
```



## ◆ 作业提交

yhalloc 分配模式

```
yhalloc -N 10 -p TH_NET
```

退出时需在当前终端下输入 `exit` 。

## ◆ 取消已提交的作业

通过 `yhq` 命令查看作业队列，获取作业id号

```
[mahy@ln2%tianhe ~]$ yhq
JOBID PARTITION  NAME      USER  ST      TIME  NODES  NODELIST(REASON)
69331  gpu_test    12-2-9    mahy  PD      0:00   4      (Priority)
69344  gpu_test    11-2-16    mahy  PD      0:00   4      (Priority)
69325  gpu_test    9-2-10     mahy  R       4:07:10  2      cn[3410-3411]
69310  gpu_test    11-2-17     mahy  R       4:35:08  2      cn[3420-3421]
```

```
yhcancel 69325
```

- 系统平台架构
- 系统及编译环境
- 系统登陆
- 作业管理
- 天河开源科学计算软件部署情况介绍
- 典型科学计算软件提交方法
- 常见问题

◆ TH-1A大系统在共享目录下部署了常用的科学计算开源软件

- 生物分子动力学模拟软件namd

/vol-th/software/namd2

/vol-th/software/namd2.9

- 生物分子动力学模拟软件gromacs

/vol-th/software/gmx455

/vol-th/software/gmx455\_float

- 材料分子动力学模拟软件 lammps

/vol-th/software/lmp\_mkl

- 材料模拟软件pwscf

/vol-th/software/pwscf

- 电子结构计算软件siesta

/vol-th/software/siesta-3.2

- 计算流体力学OPENFOAM

/vol-th/software/OpenFOAM

/vol-

th/software/OpenFOAM\_ThirdParty

- 系统平台架构
- 系统及编译环境
- 系统登陆
- 作业管理
- 天河开源科学计算软件部署情况介绍
- 典型科学计算软件提交方法
- 常见问题



- ◆ namd提交任务脚本如下:

```
#!/bin/sh  
yhrun -pTH_NET -n12 /vol-th/software/namd2 input
```

- ◆ gromacs提交任务脚本如下:

```
#!/bin/sh  
yhrun -pTH_NET -n12 /vol-th/software/gmx455/bin/mdrun_d -deffnm
```

- ◆ lammps提交任务脚本如下:

```
#!/bin/sh  
yhrun -pTH_NET -n12 /vol-th/software/lmp_mkl -in input
```

- ◆ Pwscf提交任务脚本如下:

```
#!/bin/sh  
yhrun -pTH_NET -n12 /vol-th/software/pwscf/pw.x -npool 12-ntg 1 -  
input <YMnCo0.in> YMnCo0.log
```

- 系统平台架构
- 系统及编译环境
- 系统登陆
- 作业管理
- 天河开源科学计算软件部署情况介绍
- 典型科学计算软件提交方法
- 常见问题



## ◆ VPN登陆问题

● Q: Windows系统登陆VPN无法加载插件或报错

● A: 按照用户手册确认是否打开了浏览器的允许加载插件选项，并确保您机器的安全软件、杀毒软件等不会阻止插件的正常运行。

● Q: Windows系统登陆VPN成功，但仍连接不上服务器

● A: 请确认您的桌面任务栏的通知区域里是否有“SSL VPN Client”图标。如果没有，则说明您没有登录成功，插件没有正常运行，请您参考手册中登陆部分重新进行设置确认；如果有，则需要进一步去确认您系统的主机名解析文件C:\WINDOWS\system32\drivers\etc\hosts内是否有VPN页面上的机器标识，如果没有，则说明您的该文件被保护或您当前的用户对该文件没有写权限，请您用系统管理员用户登录，以及关闭安全软件及杀毒软件对该文件的控制。

● Q: 登录VPN提示“超出人数限制”

● A: VPN用户默认最多允许5个人同时登陆，超过这个限制就会报错。但如果用户登录VPN后没有点击“退出”按钮正常退出，而是直接关掉页面，则会保持该次连接会话直到超时，此时也会占用一个登录连接数。

- Q: Linux系统登陆VPN不成功
- A: linux用户登录VPN需要使用root账户，并且同时关闭自己机器的ssh服务，常用版本的关闭方法为“/etc/init.d/sshd stop”。具体请参考用户手册或者VPN登陆下面的“SSL VPN设置向导”。linux系统的登陆与windows大致相同，但其需要java的支持，其对应的主机名解析文件为/etc/hosts。
- Q: Mac系统如何在1个IP下实现多台Mac系统接入“TH系统”？
- A: 目前MAC系统登录的VPN认证是基于IP建立的，也就是说同一个IP地址只能建立起一个连接，一旦连接建立，使用同一个IP访问的客户端就都可以连接到系统上来了，其他客户端不需要再登录这个VPN进行认证了，但是一旦登录VPN的客户端退出，那么整个网络也就会断开了。
- ◆ 系统登陆问题
- Q: 登陆节点home目录下看不到原有的用户文件
- A: 这是由于登录节点启动后还没有挂载相应的共享存储，请用户先退出系统，稍等待后再重新进行登录。
- Q: 在登陆节点执行“ls; cd”这样的基本命令会卡死
- A: 这有可能是由于该登陆节点负载过大造成的，此时用户可以尝试更换其他登陆节点。

- Q: 登陆节点无法连通
- A: 这有可能是用户在登陆节点上运行非法程序导致节点宕机，我们会实时对系统进行监控，出现这种情况请用户更换其他登陆节点。建议用户不要在登陆节点上运行任何计算，一旦查到并影响到其他人的使用，则会进行警告，屡次不改者可能会被封号。
- ◆ 作业运行问题
- Q: 作业在某一个时间点后无输出
- A: 导致这种现象的原因很多，需要具体问题具体分析，其中一种可能是由于用户磁盘配额已满，无法写入数据造成的，因此需要用户及时清理自己的数据，如果并非这种情况请您邮件或者电话与我们的系统管理人员取得联系，我们将进行进一步查看。
- Q: 作业断开，slurm日志中出现“DUE TO TIME LIMIT”报错信息
- A: 这是因为作业运行时间超过队列最大运行时间限制，请注意您所在队列的运行时间限制以及您作业已运行时间。

- Q: 作业断开, slurm日志中出现“Not enough endpoint resources”报错信息
- A: 这是由于上一个作业结束时出现异常, 节点endpoint没有正常释放。用户提交可以加-x剔除问题节点, 然后联系管理员进行解决。
- Q: 作业断开, slurm日志中出现“Group ID not found on host”报错信息
- A: 这是由于计算节点的passwd和group没有与管理节点同步导致。用户提交可以加-x剔除问题节点, 然后联系管理员进行解决。
- Q: 作业断开, slurm日志中出现“No such file or directory: going to /tmp instead”报错信息
- A: 这是由于计算节点没有挂载共享存储。用户提交可以加-x剔除问题节点, 然后联系管理员进行解决。
- Q: 作业断开, slurm日志中出现“Job credential expired”报错信息
- A: 这是由于计算节点时间没有与管理节点同步。用户提交可以加-x剔除问题节点, 然后联系管理员进行解决。

- Q: 作业断开, slurm日志中出现“bus error”报错信息
- A: 导致“bus error”的报错原因很多, 具体问题需要使用工具排查。用户提交可以加-x剔除问题节点, 然后联系管理员进行解决。
- Q: 提交的作业总是被自动退出
- A: 您如果是使用yhrun提交任务, 那么终端关闭、脚本终止都会导致任务被杀掉。建议用户使用yhbatch的提交方式, yhbatch提交的任务, 终端关闭甚至登陆节点宕机都不会对已提交的作业有影响。另外, 还有可能是您提交的作业所分配的计算节点有问题导致自动退出, 请您仔细查看产生的日志文件的报错信息, 是否属于以上问题中的一种, 并采取相应的处理。
- Q: 作业状态“S; CG”分别表示什么含义
- A: “S”表示管理员将用户作业挂起以进行故障检测或故障处理, 处理完后会将该作业恢复, 不会对作业产生任何影响; “CG”是为作业结束或取消后的“退出”标记, 不会对用户作业造成影响, 请正常使用, 如作业长时间处理“CG”状态, 管理员会对其进行恢复处理。

- Q: 作业断开, slurm日志提示找不到某个动态链接库
- A: 需要用户将动态链接库的路径添加到自己运行的环境变量中, 假设缺少x库, 先“locate x”找到该链接库的地址\$DIR, 然后编辑用户目录下的配置文件.bashrc, 添加“export LD\_LIBRARY\_PATH=\$DIR:\$LD\_LIBRARY\_PATH”。
- Q: 查看有可用节点, 但作业却一直处于PD状态
- A: TH系统的资源管理器采用“先进先出”的作业调度方式, 作业处于PD状态说明在用户前面有其他用户先提交了作业, 并且之前的用户作业超出了目前的可用资源总数, 请用户耐心等待。根据用户资源需求, 系统管理人员也会定期进行资源调整, 降低作业排队时间。
- ◆ 存储问题
- Q: 登陆系统时提示“Some errors happened when getting quota info”
- A: 这是由于在对系统进行调整时登陆节点quota服务没有启用导致, 对用户本身的操作和作业不会有影响, 管理员会定时对此进行调整, 请您放心使用。



- Q: 默认的磁盘配额是多少？磁盘配额的含义是什么？
- A: 为了合理利用有限的存储资源，目前中心对用户默认进行存储软限制500G，存储硬限制1T，文件数软限制100万，文件数硬限制200万的磁盘配额限制。以磁盘存储为例说明软、硬限制的含义，文件数软、硬限制的含义与其一样。用户使用存储低于500G时，存储状态正常；当用户使用存储介于500G和1T之间时，用户配额异常，通过“`lfs quota -g username /vol-th`”查看账号配额会看到已使用存储的数字旁边有一个“\*”号，状态“4w1d23h59m57s”表示一个月的倒计时，如果用户在倒计时结束前将使用存储清理到500G以下，则存储状态恢复正常，否则，用户存储无法写入；如果用户使用存储大于1T，用户会无法写入。
- Q: 磁盘无法写入，报“quota error”错误
- A: 这是由于用户使用存储或文件数超过配额设定，需要用户对数据进行清理到磁盘配额软限制一下方可继续使用。



感谢您的支持与合作

