简单集群搭建和性能测试

一、虚拟机的创建

1、Hypervisor 和Linux 映像文件的下载

- **Hypervisor的下载**

  本实验中选用**VMware Workstation Pro**

  > 官网下载**VMware Workstation Pro**
  >
  > 
  >

- **Linux**镜像文件的下载

  本实验中选择**Ubuntu 22.04 LTS**系统

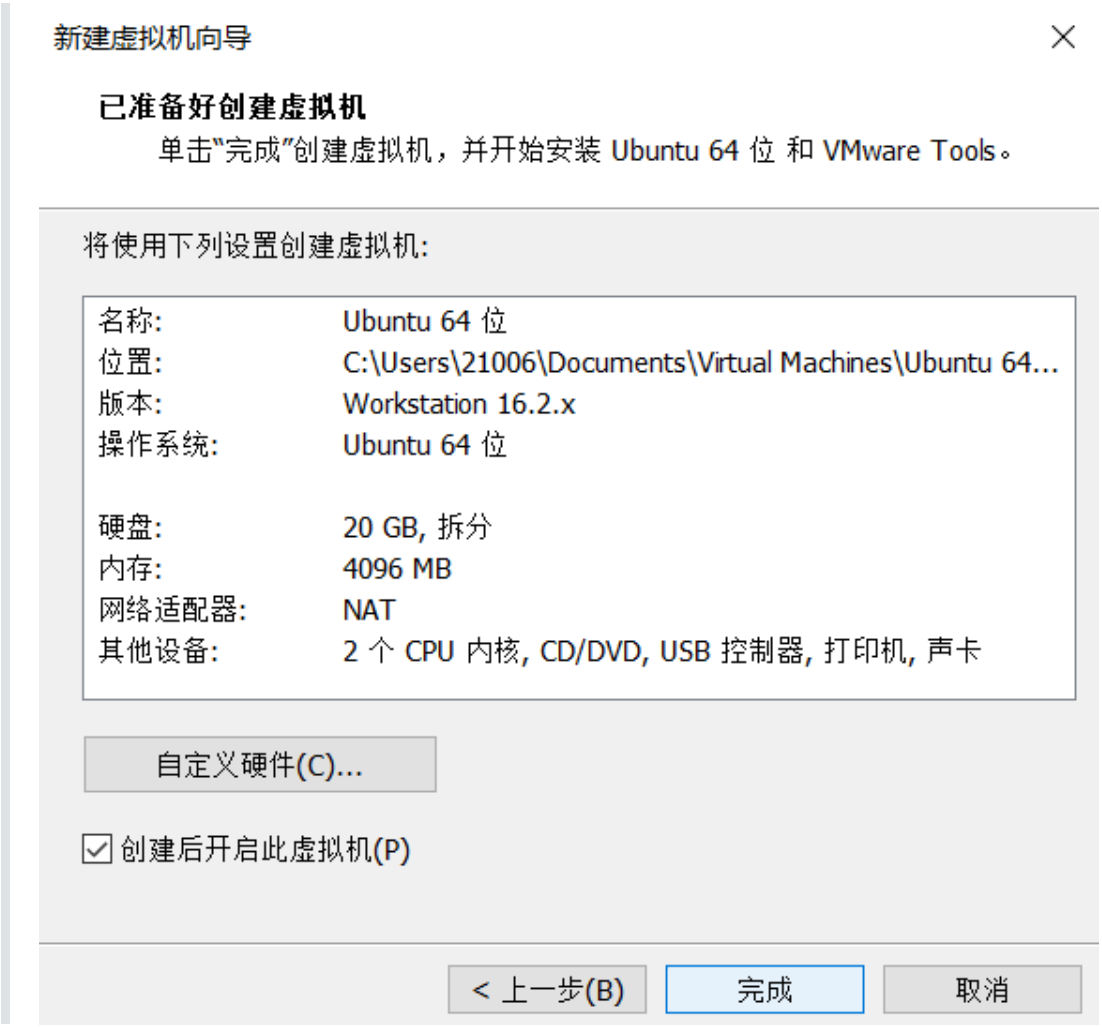
为了解决下载速度过慢的问题，选择**Ubuntu中国**

2、创建虚拟机

- 使用自定义方式创建虚拟机

- 选择镜像文件

- 为虚拟机分配**4096MB**内存，**20.0GB**硬盘容量，**2**核
- 采用**NET**网络类型，使用**LSI Logic**控制器



- 虚拟机创建成功

二、HPL环境搭建

HPL是**Linpack**测试的一种，需要依赖**OpenMPI**和**OpenBLAS**来实现

1、虚拟机环境配置

> 安装相关编译器
>
> 安装make工具和网络工具

- **安装gcc、g++、python3以及gfortran编译器**

  `sudo apt install gcc`

  `sudo apt install g++`

  `sudo apt install  python3`

  `sudo apt install gfortran`

- **检查相关编译器是否安装成功**

  `gcc --version`

  ...

- **安装make工具、网络工具和SSH Server**

  `sudo apt install make`

  `sudo apt install net-tools`

`sudo apt install openssh server`

- **验证网络工具和SSH Server**

  `sudo systemctl status ssh`

  `ifconfig`

- **UFW打开SSH端口**

  `sudo ufw allow ssh`

2、Openmpi安装

- **下载并解压Openmpi**

  下载**Openmpi-4.1.4**

> | Release | File names | Size | Date | Checksums (GNU md5sum and sha1sum v5.2.1) |
> |---|---|---|---|---|
> | 4.1.4 SRPM notes | openmpi-4.1.4-1.src.rpm | 16.63 MiB | May 26, 2022 | MD5: ecf2593c494d71c8d923f4efa41ca7e6
SHA1: 0a2cb4a2b70b42b2fa94c3d9a961f69fc141604f
SHA256: 3186b4d7e1c67e1bae78f6a45aa04d9b30a7801c8aa24435b47f96b48971e116 |
> | | openmpi-4.1.4.tar.bz2 | 9.58 MiB | May 26, 2022 | MD5: f057e12aabaf7dd5a6a658180fca404e
SHA1: 357c61a8e06e103d987c0e4a054e8780a034c8b1
SHA256: 92912e175fd1234368c8730c03f4996fe5942e7479bb1d10059405e7f2b3930d |
> | | openmpi-4.1.4.tar.gz | 16.97 MiB | May 26, 2022 | MD5: 926aa26390091fb0c51629170b1acb01
SHA1: ec49e93db63e937a0de6b1b6885f57017e977970
SHA256: e166dbe876e13a50c2882e11193fecbc4362e89e6e7b6deeb69bf095c0f4fc4c |

  解压缩到当前文件夹

  `tar zxvf openmpi-4.1.1.tar.gz`

- **安装openmpi**

```
cd openmpi-4.1.4
./configure
sudo make all install
```

- **添加路径**

  `sudo vi /etc/profile`

  在**profile**文件末尾添加

  `export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/usr/local/lib`

  执行以下语句使配置生效

  `source /etc/profile`

- **测试环境**

```
cd Downloads/openmpi-4.1.4/examples
make mpirun -np 2 hello_c
```

  测试结果

```
make[2]: Leaving directory '/home/test1/Downloads/openmpi-4.1.4/examples'
make[1]: Leaving directory '/home/test1/Downloads/openmpi-4.1.4/examples'
Hello, world, I am 0 of 2, (Open MPI v4.1.4, package: Open MPI test1@test1-virtu
al-machine Distribution, ident: 4.1.4, repo rev: v4.1.4, May 26, 2022, 122)
Hello, world, I am 1 of 2, (Open MPI v4.1.4, package: Open MPI test1@test1-virtu
al-machine Distribution, ident: 4.1.4, repo rev: v4.1.4, May 26, 2022, 122)
```

3、openblas安装

`sudo apt-get install libopenblas-dev`

安装路径：

`/usr/lib/x86_64-linux-gnu/openblas-pthread`

三、HPL的安装和环境配置

- **下载并解压** `hpl-2.3.tar.gz`

  `tar zxvf hpl-2.3.tar.gz`

- **复制粘贴Make文件**

```
cp Downloads/hpl-2.3/setup/Make.Linux_PII_CBLAS Downloads/hpl-2.3/Make.TEST
mv hpl-2.3 hpl
```

- **修改Make中的路径**

```
RCH = TEST
TOPdir = $(HOME)/Downloads/hpl
MPdir = /usr/local/lib/openmpi
MPinc =
MPlib = /usr/local/lib/libmpi.so
LAdir = /usr/lib/x86_64-linux-gnu/openblas-pthread
LAinc =
LAlib = $(LAdir)/libblas.a $(LAdir)/libblas.so
HPL_OPTS =
CC = /usr/local/bin/mpicc
CCFLAGS = $(HPL_DEFS) -fomit-frame-pointer -O3 -funroll-loops -W -Wall -
pthread
LINKER = $(CC)
arch = TEST
```

- **编译**

  `make arch = TEST`

- **检验编译是否成功**

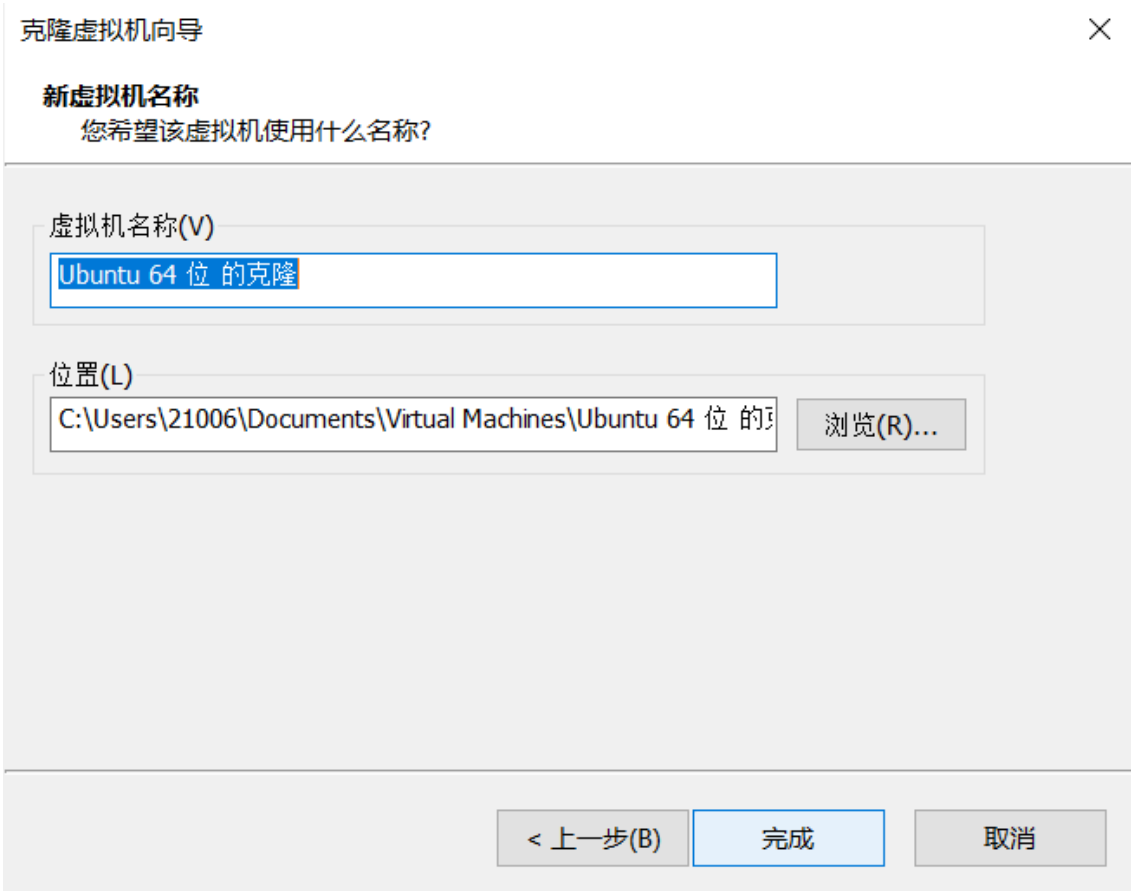
```
cd hpl/bin/TEST
dir
HPL.dat  xhpl
```

  证明HPL安装成功

四、克隆节点

在VMware中，克隆已经配置完成的节点，成为集群中的其他三个节点

- **创建完整克隆**



五、测试集群

1、ping通

- **获得每台虚拟机的IP地址**

  `ifconfig`

```
test1@test1-virtual-machine:~/Downloads/openmpi-4.1.4/examples$ ifconfig
ens33: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 1500
        inet 192.168.233.130  netmask 255.255.255.0  broadcast 192.168.233.255
        inet6 fe80::ede1:6a42:6c2f:b723  prefixlen 64  scopeid 0x20<link>
        ether 00:0c:29:dc:f6:d7  txqueuelen 1000  (Ethernet)
        RX packets 104423  bytes 154061703 (154.0 MB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 15329  bytes 1183887 (1.1 MB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
```

  获得的虚拟机的地址为

```
192.168.233.128
192.168.233.129
192.168.233.130
192.168.233.131
```

- **确认能互相ping通**

  `ping 192.168.233.129`

  ...

```
test1@test1-virtual-machine:~/Downloads/openmpi-4.1.4/examples$ ping 192.168.233.130
PING 192.168.233.130 (192.168.233.130) 56(84) bytes of data.
64 bytes from 192.168.233.130: icmp_seq=1 ttl=64 time=0.025 ms
64 bytes from 192.168.233.130: icmp_seq=2 ttl=64 time=0.058 ms
64 bytes from 192.168.233.130: icmp_seq=3 ttl=64 time=0.062 ms
64 bytes from 192.168.233.130: icmp_seq=4 ttl=64 time=0.072 ms
64 bytes from 192.168.233.130: icmp_seq=5 ttl=64 time=0.062 ms
^C
--- 192.168.233.130 ping statistics ---
5 packets transmitted, 5 received, 0% packet loss, time 4083ms
rtt min/avg/max/mdev = 0.025/0.055/0.072/0.016 ms
```

2、配置SSH

- **生成公钥私钥对**

  `ssh-kegen -t rsa`

```
test1@test1-virtual-machine:~$ ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (/home/test1/.ssh/id_rsa):
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /home/test1/.ssh/id_rsa
Your public key has been saved in /home/test1/.ssh/id_rsa.pub
The key fingerprint is:
SHA256:xtXZRYKqQP/sTejVmi/ZnAMYAZsbuwYU3KFCF0AQD7Y test1@test1-virtual-machine
The key's randomart image is:
+---[RSA 3072]----+
|  =++o+ooo   ...o|
| . = .oo.o...o o |
|  E o.o.+ .oo .  |
|     o...=o      |
|      ..S+ + .   |
|       o..= + .  |
|        oo + B . |
|       .  o * =  |
|             o.. |
+----[SHA256]-----+
```

- **将主机上的公钥拷贝到另外三台虚拟机目录下**

  `ssh-copy-id  -i ~/.ssh/id_rsa.pub test1@192.168.233.129`

```
test1@test1-virtual-machine:~/Downloads/hpl/bin/TEST$ ssh-copy-id -i ~/.ssh/id_rsa.pub test1@192.1
68.233.130
/usr/bin/ssh-copy-id: INFO: Source of key(s) to be installed: "/home/test1/.ssh/id_rsa.pub"
The authenticity of host '192.168.233.130 (192.168.233.130)' can't be established.
ED25519 key fingerprint is SHA256:lSUefknpBUO5VPJgwA5A4csNUkY/i6KHgmoQXD1Jlos.
This host key is known by the following other names/addresses:
    ~/.ssh/known_hosts:1: [hashed name]
    ~/.ssh/known_hosts:2: [hashed name]
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
/usr/bin/ssh-copy-id: INFO: attempting to log in with the new key(s), to filter out any that are a
lready installed
/usr/bin/ssh-copy-id: INFO: 1 key(s) remain to be installed -- if you are prompted now it is to in
stall the new keys
test1@192.168.233.130's password:

Number of key(s) added: 1

Now try logging into the machine, with:   "ssh 'test1@192.168.233.130'"
and check to make sure that only the key(s) you wanted were added.
```

- **远程免密访问测试**

  `ssh test1@192.168.233.129`

3、mpirun尝试

- **创建主机中的hostfile文件**

```
localhost slots=2
192.168.233.128 slots=2
192.168.233.129 slots=2
192.168.233.130 slots=2
192.168.233.131 slots=2
```

- **查看每个节点的上线时间**

  `mpirun --hostfile myhostfile uptime`

```
test1@test1-virtual-machine:~/Downloads/hpl/bin/TEST$ mpirun --hostfile myhostfile uptime
 21:31:32 up 17 min,  1 user,  load average: 0.01, 0.06, 0.07
 21:31:32 up 27 min,  1 user,  load average: 0.03, 0.10, 0.07
 21:31:32 up 20 min,  1 user,  load average: 0.02, 0.06, 0.05
 21:31:32 up 25 min,  1 user,  load average: 0.07, 0.09, 0.11
 21:31:32 up 27 min,  1 user,  load average: 0.03, 0.10, 0.07
 21:31:32 up 17 min,  1 user,  load average: 0.01, 0.06, 0.07
 21:31:32 up 20 min,  1 user,  load average: 0.02, 0.06, 0.05
 21:31:32 up 25 min,  1 user,  load average: 0.07, 0.09, 0.11
test1@test1-virtual-machine:~/Downloads/hpl/bin/TEST$
```

- **运行HPL**

  `mpirun --hostfile myhostfile ./xhpl`

  - **开始运行**

```
test1@test1-virtual-machine:~/Downloads/hpl/bin/TEST$ mpirun --hostfile myhostfile ./xhpl
================================================================================
HPLinpack 2.3  --  High-Performance Linpack benchmark  --   December 2, 2018
Written by A. Petitet and R. Clint Whaley,  Innovative Computing Laboratory, UTK
Modified by Piotr Luszczek, Innovative Computing Laboratory, UTK
Modified by Julien Langou, University of Colorado Denver
================================================================================

An explanation of the input/output parameters follows:
T/V    : Wall time / encoded variant.
N      : The order of the coefficient matrix A.
NB     : The partitioning blocking factor.
P      : The number of process rows.
Q      : The number of process columns.
Time   : Time in seconds to solve the linear system.
Gflops : Rate of execution for solving the linear system.

The following parameter values will be used:

N      :      29       30       34       35
NB     :       1        2        3        4
PMAP   : Row-major process mapping
P      :       2        1        4
Q      :       2        4        1
PFACT  :    Left    Crout    Right
NBMIN  :       2        4
NDIV   :       2
RFACT  :    Left    Crout    Right
BCAST  :   1ring
DEPTH  :       0
SWAP   : Mix (threshold = 64)
L1     : transposed form
U      : transposed form
EQUIL  : yes
ALIGN  : 8 double precision words

--------------------------------------------------------------------------------
```

  - **运行结束**

```
--------------------------------------------------------------------------------
||Ax-b||_oo/(eps*(||A||_oo*||x||_oo+||b||_oo)*N)=   2.47303684e-02 ...... PASSED
================================================================================
T/V                N    NB     P     Q               Time                 Gflops
--------------------------------------------------------------------------------
WR00R2C4          35     4     4     1               0.01             2.2796e-03
HPL_pdgesv() start time Fri May 27 21:28:12 2022

HPL_pdgesv() end time   Fri May 27 21:28:12 2022

--------------------------------------------------------------------------------
||Ax-b||_oo/(eps*(||A||_oo*||x||_oo+||b||_oo)*N)=   2.87441979e-02 ...... PASSED
================================================================================
T/V                N    NB     P     Q               Time                 Gflops
--------------------------------------------------------------------------------
WR00R2R2          35     4     4     1               0.03             9.0753e-04
HPL_pdgesv() start time Fri May 27 21:28:12 2022

HPL_pdgesv() end time   Fri May 27 21:28:12 2022

--------------------------------------------------------------------------------
||Ax-b||_oo/(eps*(||A||_oo*||x||_oo+||b||_oo)*N)=   2.47303684e-02 ...... PASSED
================================================================================
T/V                N    NB     P     Q               Time                 Gflops
--------------------------------------------------------------------------------
WR00R2R4          35     4     4     1               0.01             3.8489e-03
HPL_pdgesv() start time Fri May 27 21:28:12 2022

HPL_pdgesv() end time   Fri May 27 21:28:12 2022

--------------------------------------------------------------------------------
||Ax-b||_oo/(eps*(||A||_oo*||x||_oo+||b||_oo)*N)=   1.99396688e-02 ...... PASSED
================================================================================

Finished    864 tests with the following results:
            864 tests completed and passed residual checks,
              0 tests completed and failed residual checks,
              0 tests skipped because of illegal input values.
--------------------------------------------------------------------------------

End of Tests.
================================================================================
```