上海复旦大学物理楼 324 机房

测试集群概况介绍

一、集群信息介绍

1. 此次测试集群包含 1 台登陆节点、1 台存储节点也是管理节点、10 台刀片计算节点，共计 12 台主机组成。
2. 网络系统包括 ssh 千兆管理网和 InfiniBand 40Gb 存储网。
3. 管理节点作为存储节点、PBS 调度服务端、NIS 用户认证服务端，是真正意义上的管理节点。
4. 集群采用 PBS 作业调试系统。
5. 主机名查看；

```
[root@login ~]# cat /etc/hosts
127.0.0.1 localhost localhost.localdomain localhost4 localhost4.localdomain4
::1 localhost localhost.localdomain localhost6 localhost6.localdomain6
#####################---SSH-MGMT---############################
11.11.11.254 admin node254
11.11.11.253 login node253
11.11.11.1 node1
11.11.11.2 node2
11.11.11.3 node3
11.11.11.4 node4
11.11.11.5 node5
11.11.11.6 node6
11.11.11.7 node7
11.11.11.8 node8
11.11.11.9 node9
11.11.11.10 node10
###################---IB-STORAGE---########################
12.12.12.254 iadmin
12.12.12.253 ilogin
12.12.12.1 inode1
12.12.12.2 inode2
12.12.12.3 inode3
12.12.12.4 inode4
12.12.12.5 inode5
12.12.12.6 inode6
12.12.12.7 inode7
12.12.12.8 inode8
12.12.12.9 inode9
12.12.12.10 inode10
```

二、登陆集群

login 已接入校园网，IP 地址为 10.92.3.111。

```
[root@login ~]# ip a
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
    inet6 ::1/128 scope host
       valid_lft forever preferred_lft forever
2: eth0: <BROADCAST,MULTICAST> mtu 1500 qdisc pfifo_fast state DOWN qlen 1000
    link/ether 00:1b:21:89:0a:a6 brd ff:ff:ff:ff:ff:ff
3: eth1: <BROADCAST,MULTICAST> mtu 1500 qdisc pfifo_fast state DOWN qlen 1000
    link/ether 00:1b:21:89:0a:a7 brd ff:ff:ff:ff:ff:ff
4: eth2: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP qlen 1000
    link/ether 00:25:90:34:67:da brd ff:ff:ff:ff:ff:ff
    inet 11.11.11.253/16 brd 11.11.255.255 scope global eth2
    inet6 fe80::225:90ff:fe34:67da/64 scope link
       valid_lft forever preferred_lft forever
5: eth3: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP qlen 1000
    link/ether 00:25:90:34:67:db brd ff:ff:ff:ff:ff:ff
    inet 10.92.3.111/24 brd 10.92.3.255 scope global eth3
    inet6 2001:da8:8001:9203:225:90ff:fe34:67db/64 scope global dynamic
       valid_lft 2591869sec preferred_lft 2591869sec
    inet6 fe80::225:90ff:fe34:67db/64 scope link
       valid_lft forever preferred_lft forever
6: ib0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 2044 qdisc mq state UP qlen 1024
    link/infiniband a0:00:02:20:fe:80:00:00:00:00:00:00:02:c9:03:00:0f:61:0b brd 00:f
    inet 12.12.12.253/24 brd 12.12.12.255 scope global ib0
    inet6 fe80::202:c903:f:610b/64 scope link
       valid_lft forever preferred_lft forever
```

三、账户管理

1. 集群采用 NIS 进行统一用户认证，创建用户需要以管理员身份 ssh 到 admin 节点操作，示例创建 test 用户；
   useradd test -d /public/home/test
2. 为新用户创建密码
   passwd test
3. 同步用户账户到所有节点（此 3 步操作已完成新增用户操作）；
   make -C /var/yp
4. 用户可自行修改密码，直接在 login 上操作即可；
   yppasswd

四、普通用户测试 PBS 作业系统可用性

1. 调度系统概况查询；

```
[ll@admin ~]$ pestat
   node         state  load     phymem   ncpus   allmem resi    usrs     tasks   jobidlist
   node10       free   2.24     24020       12    56788 790     2/1      2        8 ll
   node1        free   2.33     24020       12    56788 811     2/1      2        8 ll
   node2        free   2.22     24020       12    56788 787     2/1      2        8 ll
   node3        free   2.26     19980       12    52748 741     2/1      2        8 ll
   node4        free   2.03     24020       12    56788 781     2/1      2        8 ll
   node5        free   2.31     24020       12    56788 786     2/1      2        8 ll
   node6        free   2.80*    24020       12    56788 820     3/1      2        8 ll
   node7        free   2.27     24020       12    56788 791     2/1      2        8 ll
   node8        free   2.33     24020       12    56788 794     2/1      2        8 ll
   node9        free   2.19     24020       12    56788 788     2/1      2        8 ll
```

2. 编写作业脚本；

```
[ll@login ~]$ cat iozone.pbs
#! /bin/bash
#PBS -N iozone_storage
#PBS -l nodes=10:ppn=2
#PBS -j oe
#PBS -l walltime=24:00:00
#PBS -q middle

###################################################
cd /public/software/benchmark/iozone/gnu/3.430/
RSH=ssh ./iozone -i 0 -i 1 -r 1M -s 48g -t 20 -+m  /public/home/ll/test/nodelist
```

3. 提交作业 sbatch iozone.pbs。
4. 查询作业正常运行；

```
[ll@admin ~]$ qstat
Job ID                    Name             User            Time Use S Queue
------------------------- ---------------- --------------- -------- - -----
8.admin                    iozone_storage   ll                     0 R middle
[ll@admin ~]$ qstat -n 8

admin:
                                                                                  Req'd    Req'd       Elap
Job ID                  Username    Queue            Jobname          SessID  NDS   TSK   Memory   Time    S   Time
----------------------- ----------- ---------------- ---------------- ------ ----- ------ ------ --------- - ---------
8.admin                 ll          middle           iozone_storage        0    10     20    --   24:00:00 R  00:00:18
   node6/0+node6/1+node9/0+node9/1+node4/0+node4/1+node5/0+node5/1+node2/0
   +node2/1+node7/0+node7/1+node8/0+node8/1+node10/0+node10/1+node1/0+node1/1
   +node3/0+node3/1
[ll@admin ~]$
```

5. 确认计算节点有 iozone 测试进程在运行；

```
[node8]ll         923     1  0 20:06 ?        00:00:11 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node8]ll         924   923  0 20:06 ?        00:00:00 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node8]root       939     2  0 Jan05 ?        00:00:00 [edac-poller]
[node8]root      3550  3548  0 20:48 ?        00:00:00 bash -c  PATH=/sbin:/usr/sbin:/bin:/usr/bin ps -ef | grep ll
[node8]root      3593  3550  0 20:48 ?        00:00:00 grep ll
[node8]root     16370     1  0 Jan06 ?        00:01:04 /opt/gridview/gvsnmp/sbin/snmpd -c /opt/gridview/gvsnmp/conf/snmpd.conf -M /opt/gr
[node10]root      595     2  0 Jan05 ?        00:00:03 [mpt_poll_0]
[node10]ll        935     1  0 20:06 ?        00:00:15 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node10]root      936     2  0 Jan05 ?        00:00:00 [edac-poller]
[node10]ll        937   935  0 20:06 ?        00:00:00 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node10]ll        964     1  0 20:06 ?        00:00:15 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node10]ll        965   964  0 20:06 ?        00:00:00 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node10]root     4028  4025  0 20:48 ?        00:00:00 bash -c  PATH=/sbin:/usr/sbin:/bin:/usr/bin ps -ef | grep ll
[node10]root     4070  4028  0 20:48 ?        00:00:00 grep ll
[node10]root    16391     1  0 Jan06 ?        00:01:04 /opt/gridview/gvsnmp/sbin/snmpd -c /opt/gridview/gvsnmp/conf/snmpd.conf -M /opt/g
d
[node3]root       581     2  0 Jan05 ?        00:00:03 [mpt_poll_0]
[node3]root       869     2  0 Jan05 ?        00:00:00 [edac-poller]
[node3]ll        5398     1  0 20:06 ?        00:00:15 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node3]ll        5399  5398  0 20:06 ?        00:00:00 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node3]ll        5425     1  0 20:06 ?        00:00:15 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node3]ll        5426  5425  0 20:06 ?        00:00:00 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node3]root      8461  8459  0 20:48 ?        00:00:00 bash -c  PATH=/sbin:/usr/sbin:/bin:/usr/bin ps -ef | grep ll
[node3]root      8499  8461  0 20:48 ?        00:00:00 grep ll
[node3]root     11643     1  0 Jan06 ?        00:02:46 /opt/gridview/gvsnmp/sbin/snmpd -c /opt/gridview/gvsnmp/conf/snmpd.conf -M /opt/gr
[node2]root       503     2  0 Jan05 ?        00:00:03 [mpt_poll_0]
[node2]root      1052     2  0 Jan05 ?        00:00:00 [edac-poller]
[node2]ll        5779     1  0 20:06 ?        00:00:14 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node2]ll        5780  5779  0 20:06 ?        00:00:00 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node2]ll        5806     1  0 20:06 ?        00:00:14 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node2]ll        5807  5806  0 20:06 ?        00:00:00 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node2]root      8771  8769  0 20:48 ?        00:00:00 bash -c  PATH=/sbin:/usr/sbin:/bin:/usr/bin ps -ef | grep ll
[node2]root      8814  8771  0 20:48 ?        00:00:00 grep ll
[node2]root     13317     1  0 Jan06 ?        00:02:39 /opt/gridview/gvsnmp/sbin/snmpd -c /opt/gridview/gvsnmp/conf/snmpd.conf -M /opt/gr
[node5]root       580     2  0 Jan05 ?        00:00:03 [mpt_poll_0]
[node5]root       942     2  0 Jan05 ?        00:00:00 [edac-poller]
[node5]ll        5513     1  0 20:06 ?        00:00:15 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node5]ll        5514  5513  0 20:06 ?        00:00:00 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node5]ll        5540     1  0 20:06 ?        00:00:15 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node5]ll        5541  5540  0 20:06 ?        00:00:00 /public/software/benchmark/iozone/gnu/3.430/iozone -+s -t 1 -r 4 -s 4 -+c node6
[node5]root      8515  8513  0 20:48 ?        00:00:00 bash -c  PATH=/sbin:/usr/sbin:/bin:/usr/bin ps -ef | grep ll
[node5]root      8561  8515  0 20:48 ?        00:00:00 grep ll
[node5]root     13444     1  0 Jan06 ?        00:02:49 /opt/gridview/gvsnmp/sbin/snmpd -c /opt/gridview/gvsnmp/conf/snmpd.conf -M /opt/gr
[node7]root       494   492  0 20:48 ?        00:00:00 bash -c  PATH=/sbin:/usr/sbin:/bin:/usr/bin ps -ef | grep ll
[node7]root       547   494  0 20:48 ?        00:00:00 grep ll
[node7]root       580     2  0 Jan05 ?        00:00:03 [mpt_poll_0]
[node7]root       938     2  0 Jan05 ?        00:00:00 [edac-poller]
```

(结尾)