**Multi-GPU Computer Cluster**

**2017/07/31**

**Wenming Zhang**

[Forward 2](#_Toc489286536)

[Steps 3](#_Toc489286537)

[Configuration Network 4](#_Toc489286538)

[Install NFS 4](#_Toc489286539)

[Master 4](#_Toc489286540)

[Client 4](#_Toc489286541)

[Install NIS 5](#_Toc489286542)

[Master 5](#_Toc489286543)

[Client 5](#_Toc489286544)

[Install GridEngine 6](#_Toc489286545)

[Master 6](#_Toc489286546)

[Client 7](#_Toc489286547)

[Add new node or queue 7](#_Toc489286548)

[Add new node 7](#_Toc489286549)

[Add new queue 8](#_Toc489286550)

[Using GridEngine 10](#_Toc489286551)

[Useful command 10](#_Toc489286552)

[Server status after installation 12](#_Toc489286553)

# Forward

Motivation

Our target is to set up a multi-GPU computer cluster in our servers (nodes). To achieve this, NFS will be used to realize the data sharing in all servers, NIS (Network Information Service) will be used to make our servers can be freely accessed by uniform authority account, GridEngine will be used to manage all the computer resources in our servers including CPU, GPU and memory.

Servers

We get four computer servers (nodes) in our situation and the detail information is as following:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Server | SCRL-MARD2-DL580 | SCRL-MARD2-Z820-45 | SCRL-MARD2-Z820-46 | PowerEdge-C4130 |
| OS | Ubuntu 14.04.2 | Ubuntu 14.04.2 | Ubuntu 14.04.2 | Ubuntu 14.04.1 |
| IP | 192.168.1.200  43.82.40.44 | 192.168.1.201  43.82.40.45 | 192.168.1.202  43.82.40.46 | 192.168.1.203  43.82.40.47 |
| GPU | GPU 0: GeForce GTX TITAN X  GPU 1: Tesla K40c  GPU 2: Tesla K40c | GPU 0: Tesla K20c  GPU 1: Tesla K20c  GPU 2: Quadro K2000 | GPU 0: Tesla K40c  GPU 1: Tesla K40c  GPU 2: Quadro K600 | GPU 0: Tesla K80  GPU 1: Tesla K80  GPU 2: Tesla K80  GPU 3: Tesla K80  GPU 4: Tesla K80  GPU 5: Tesla K80  GPU 6: Tesla K80  GPU 7: Tesla K80 |
| CPU | 80 \* 2.20GHz | 8 \* 1.80GHz | 8 \* 1.80GHz | 48 \* 2.50GHz |
| memory | 800G | 250G | 250G | 500G |

# Steps

1. Both NIS and GridEngine have the concept of master and client. In our case, we set the SCRL-MARD2-Z820-45 as the “NIS master node” and “GridEngine master node”.
2. To make our servers can reach to each other we need to configure the network mapping file [/etc/hosts](#_Configuration_Network). In our case we use “mard2nis” as the network domain, and also this will be used as the NIS domain name when installing NIS.
3. In the GPU computer cluster, it is better to create a unique user account to using the all the resources in our servers. In our case, “nisuser” will be configured as this account. Using following command to create “nisuser” account in “NIS master node” SCRL-MARD2-Z820-45:

#adduser nisuser -u 51060 -g mard2 -d /home/nisuser

1. Next install NIS, NFS, GridEngine
2. [Install NFS](#_Install_NFS_1)

To make “nisuser” using same home folder in physical disks, we should Sharing the /home/nisuser folder in “NIS master node” SCRL-MARD2-Z820-45, and mount shared /home/nisuser folder in “NIS client node”:

1. In “NIS master node” SCRL-MARD2-Z820-45, add following line to /etc/exports file:

/home/nisuser/ \*(sync,insecure,rw,no\_subtree\_check,no\_root\_squash)

1. In other servers “NIS client node”:

#mkdir /home/nisuser

#mount -t nfs 192.168.1.201: /home/nisuser /home/nisuser

1. [Install NIS](#_Install_NIS)

After this, all servers can use “nisuser” account to loginin, and they have the same home folder in physical disks

1. [Install GridEngine](#_Install_Gridengine)

To make “nisuser” be the GridEngine user and manager, we should run “qconf -am nisuser” command in “GridEngine master node” SCRL-MARD2-Z820-45.

1. By default, all the commands run in root authority and till to now, the multi-GPU computer cluster is ready to be used. Now you can [configure](#_Using_gridengine) your project to use it.
2. In the future, you can and [new computer node](#_Add_new_node) to the multi-GPU computer cluster and also can configure [new queue](#_Add_new_queue) to allocate your computer resources.

# Configuration Network

Configure the network mapping file, Add the computer name and IP mapping information to /etc/hosts file in all servers.

#vim /etc/hosts to add the following lines in the beginning of the hosts file

127.0.0.1 localhost.localdomain localhost

192.168.1.200 SCRL-MARD2-DL580.mard2nis SCRL-MARD2-DL580

192.168.1.201 SCRL-MARD2-Z820-45.mard2nis SCRL-MARD2-Z820-45

192.168.1.202 SCRL-MARD2-Z820-46.mard2nis SCRL-MARD2-Z820-46

* + - 1. PowerEdge-C4130.mard2nis PowerEdge-C4130

# Install NFS

## Master

#apt-get install nfs-kernel-server

Configure the shared folder /home/nisuser:

#vim /etc/exports

/home/nisuser/ \*(sync,insecure,rw,no\_subtree\_check,no\_root\_squash)

Restart RPC service

#service rpcbind restart

Restart NFS service

#service nfs-kernel-server restart

## Client

Create mount folder and change the folder owner

#makedir /home/nisuser

#chown nisuser:mard2 /home/nisuser

Mount the NFS shared folder to local folder(add the command to /etc/rc.local to make it run at start up)

#mount -t nfs 192.168.1.201: /home/nisuser /home/nisuser

# Install NIS

https://www.server-world.info/en/note?os=Ubuntu\_14.04&p=nis&f=1

https://www.server-world.info/en/note?os=Ubuntu\_14.04&p=nis&f=2

## Master

#apt-get install nis

Using mard2nis as domain name when asked.

Configuration master

#vim /etc/default/nis

NISSERVER=master

# vim /etc/ypserv.securenets

# This line gives access to everybody. PLEASE ADJUST!

# comment out

# 0.0.0.0 0.0.0.0

# vim /var/yp/Makefile

# line 52: change

MERGE\_PASSWD=true

# line 56: change

MERGE\_GROUP=true

After you added users in local server, apply them to NIS database.

#cd /var/yp

#make

Start NIS server

#initctl start ypserv

ypserv start/running, process 13543

update NIS database, add all NIS node to it.

#/usr/lib/yp/ypinit -m

## Client

In NIS client node

#apt-get install nis

Using mard2nis as domain name when asked.

In NIS master node

#vim /etc/yp.conf

# add to the end: [domain name] [server] [NIS server's hostname]

domain mard2nis server SCRL-MARD2-Z820-45.mard2nis

#vim /etc/nsswitch.conf

passwd: compat nis

group: compat nis

shadow: compat nis

hosts: files dns nis

#vim /etc/pam.d/common-session

# add to the end

session optional pam\_mkhomedir.so skel=/etc/skel umask=077

Bind NIS client to NIS master

#initctl restart ypbind

ypbind start/running, process 3677

# Install GridEngine

<http://kaldi-asr.org/doc/queue.html>

## Master

#apt-get install gridengine-master gridengine-client

SCRL-MARD2-Z820-45 will be set as master node, set SCRL-MARD2-Z820-45.mard2nis as master node when asked.

**Add “nisuser” account to manager list**

#qconf -am nisuser

In the next, execute node SCRL-MARD2-DL580 will be added to GridEngine

**Add execute node to host list，submit list，execute list**

#qconf -ah SCRL-MARD2-DL580.mard2nis

#qconf -as SCRL-MARD2-DL580.mard2nis

#qconf -ae SCRL-MARD2-DL580.mard2nis

**Configure execute node**

#qconf -me SCRL-MARD2-DL580.mard2nis

complex\_values ram\_free=700G,gpu=3

**add the parallel environment**

#qconf -ap smp

pe\_name smp

slots 9999

**Add queue**

Check execute node cpu number

#grep proc /proc/cpuinfo | wc –l

#qconf -aq all.q

qname all.q

hostlist SCRL-MARD2-DL580.mard2nis

pe\_list make smp

slots 30,[SCRL-MARD2-DL580.mard2nis=80]

**Configure GridEngine resources**

#qconf -mc

update mem\_free and add gpu, ram\_free

mem\_free mf MEMORY <= YES YES 700G 0

gpu g INT <= YES YES 0 10000

ram\_free ram\_free MEMORY <= YES JOB 700G 0

## Client

#apt-get install gridengine-client gridengine-exec

Set SCRL-MARD2-Z820-45.mard2nis as master node when asked.

# Add new node or queue

## Add new node

Adding new node “SCRL-MARD2-Z820-46”

**Hosts:**

Update the new node information to all servers /etc/hosts file

#vim /etc/hosts

127.0.0.1 localhost.localdomain localhost

192.168.1.200 SCRL-MARD2-DL580.mard2nis SCRL-MARD2-DL580

192.168.1.201 SCRL-MARD2-Z820-45.mard2nis SCRL-MARD2-Z820-45

**192.168.1.202 SCRL-MARD2-Z820-46.mard2nis SCRL-MARD2-Z820-46**

* + - 1. PowerEdge-C4130.mard2nis PowerEdge-C4130

**NFS:**

[In client node](#_Client)

#makedir /home/nisuser

#chown nisuser:mard2 /home/nisuser

#mount -t nfs 192.168.1.201: /home/nisuser /home/nisuser

**NIS:**

1. [Install NIS client](#_Client_1).
2. In NIS master update the database，add all NIS node to NIS database

#/usr/lib/yp/ypinit –m

**Gridengine:**

1. [Install GridEngine](#_Client_2).
2. In master node:

Add execute node to host list, submit list, execute list.

Add execute node to all.q

#qconf -mq all.q

qname all.q

hostlist SCRL-MARD2-DL580.mard2nis **SCRL-MARD2-Z820-46.mard2nis**

pe\_list make smp

slots 30,[SCRL-MARD2-DL580.mard2nis=80], [**SCRL-MARD2-Z820-46.mard2nis=8**]

Update GridEngine resources

#qconf -mc

update mem\_free and add gpu, ram\_free

mem\_free mf MEMORY <= YES YES 900G 0

gpu g INT <= YES YES 0 10000

ram\_free ram\_free MEMORY <= YES JOB 900G 0

## Add new queue

Using following command to configure your new queue.

#qconf -aq <queue name>

qname all.q

hostlist SCRL-MARD2-DL580.mard2nis **SCRL-MARD2-Z820-46.mard2nis**

pe\_list make smp

slots 30,[SCRL-MARD2-DL580.mard2nis=80], [**SCRL-MARD2-Z820-46.mard2nis=8**]

# Using GridEngine

Configure the cmd.sh in kaldi project:

export train\_cmd="queue.pl -q K40.q -l arch=\*64 --mem 2G"

export decode\_cmd="queue.pl -q K40.q -l arch=\*64 --mem 2G"

export highmem\_cmd="queue.pl -q K40.q -l arch=\*64 --mem 4G"

export mkgraph\_cmd="queue.pl -q K40.q -l arch=\*64 --mem 4G"

# Useful command

In qconf, option ‘a’ means add, ‘m’ means manage ‘s’ means show, ‘d’ means delete.

**Add GridEngine manager**

sudo qconf -am < user-id>

**ManageGridEngine queue**

sudo qconf -aq <queue-name>

sudo qconf -mq <queue-name>

sudo qconf -sq <queue-name>

sudo qconf -dq <queue-name>

**Manage GridEngine host node**

sudo qconf -ah < node-name>

sudo qconf -sh

sudo qconf -dh < node-name>

**Manage GridEngine submit node**

sudo qconf -as < node-name>

sudo qconf -ss

sudo qconf -ds < node-name>

**Manage GridEngine execute node**

sudo qconf -ae < node-name>

sudo qconf -sel show execute node list

sudo qconf -se < node-name> show execute node detail information

sudo qconf -me < node-name>

sudo qconf -de < node-name>

**Check GridEngine status**

qstat

qstat -f

qstat -j <job-id>

qhost

qhost -q

**Set NIS domain name**

root@SCRL-MARD2-Z820-46:~# nisdomainname

nisdomain

root@SCRL-MARD2-Z820-46:~# nisdomainname mard2nis

root@SCRL-MARD2-Z820-46:~# nisdomainname

mard2nis

root@SCRL-MARD2-Z820-46:~#vim /etc/defaultdomain

mard2nis

**Update GridEngine master name**

vim /var/lib/gridengine/default/common/act\_qmaster

SCRL-MARD2-Z820-45.mard2nis

**GridEngine exec node restart：**

/etc/init.d/gridengine-exec restart

**GridEngine master node restart：**

/etc/init.d/gridengine-master restart

# Server status after installation

Servers

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Server | 44 | 45 | 46 | 47 |
| Name | SCRL-MARD2-DL580 | SCRL-MARD2-Z820-45 | SCRL-MARD2-Z820-46 | PowerEdge-C4130 |
| IP | 192.168.1.200  43.82.40.44 | 192.168.1.201  43.82.40.45 | 192.168.1.202  43.82.40.46 | 192.168.1.203  43.82.40.47 |
| GPU | GPU 0: GeForce GTX TITAN X  GPU 1: Tesla K40c  GPU 2: Tesla K40c | GPU 0: Tesla K20c  GPU 1: Tesla K20c  GPU 2: Quadro K2000 | GPU 0: Tesla K40c  GPU 1: Tesla K40c  GPU 2: Quadro K600 | GPU 0: Tesla K80 GPU 1: Tesla K80  GPU 2: Tesla K80 GPU 3: Tesla K80  GPU 4: Tesla K80 GPU 5: Tesla K80  GPU 6: Tesla K80 GPU 7: Tesla K80 |
| CPU | 80 \* 2.20GHz | 8 \* 1.80GHz | 8 \* 1.80GHz | 48 \* 2.50GHz |
| Mem | 800G | 250G | 250G | 500G |
| NIS role | client | master, client | client | client |
| NFS | mount -t nfs 192.168.1.201:/home/nisuser /home/nisuser  mount -t nfs 192.168.1.201:/disk /disk | #vim /etc/exports  /disk/ \*(sync,insecure,rw,no\_subtree\_check,no\_root\_squash)  /home/nisuser/ \*(sync,insecure,rw,no\_subtree\_check,no\_root\_squash) | mount -t nfs 192.168.1.201:/home/nisuser /home/nisuser  mount -t nfs 192.168.1.201:/disk /disk | mount -t nfs 192.168.1.201:/home/nisuser /home/nisuser  mount -t nfs 192.168.1.201:/disk /disk |
| #qconf -se complex\_values | ram\_free=700G,gpu=3 | ram\_free=250G,gpu=3 | ram\_free=200G,gpu=3 | ram\_free=400G,gpu=8 |
| /etc/hosts | 127.0.0.1 localhost.localdomain localhost  43.82.40.44 SCRL-MARD2-DL580.mard2nis SCRL-MARD2-DL580  43.82.40.45 SCRL-MARD2-Z820-45.mard2nis SCRL-MARD2-Z820-45  43.82.40.46 SCRL-MARD2-Z820-46.mard2nis SCRL-MARD2-Z820-46  43.82.40.47 PowerEdge-C4130.mard2nis PowerEdge-C4130 | | | |
| nis | Nisdomainname: mard2nis | | | |

Queue

|  |  |  |
| --- | --- | --- |
| **qname** | **qconf -sq** | **qconf -sp** |
| all.q | qname all.q  hostlist SCRL-MARD2-DL580.mard2nis SCRL-MARD2-Z820-45.mard2nis \  PowerEdge-C4130.mard2nis SCRL-MARD2-Z820-46.mard2nis  pe\_list make smp  slots 30,[SCRL-MARD2-DL580.mard2nis=80], \  [SCRL-MARD2-Z820-45.mard2nis=8], \  [SCRL-MARD2-Z820-46.mard2nis=8], \  [PowerEdge-C4130.mard2nis=48] | pe\_name smp  slots 9999 |
| K40.q | qname k40.q  hostlist SCRL-MARD2-DL580.mard2nis SCRL-MARD2-Z820-46.mard2nis  pe\_list make smp  slots 1,[SCRL-MARD2-DL580.mard2nis=80], \  [SCRL-MARD2-Z820-46.mard2nis=8] |  |
| K20.q | qname K20.q  hostlist SCRL-MARD2-Z820-45.mard2nis  pe\_list make smp  slots 1,[SCRL-MARD2-Z820-45.mard2nis=8] |  |
| K80.q | qname K80.q  hostlist PowerEdge-C4130.mard2nis  pe\_list make smp  slots 1,[PowerEdge-C4130.mard2nis=48] |  |