

Задание номер 2: «Построение кластерной файловой системы GFS2».
Выполнил слушатель курса «Linux. Advanced»
Бурнейка Артурас Витаутасович.

С помощью terraform создаются четыре виртуальные машины:

- first, second, third (инициаторы)
- storage (target)

Все машины развернуты на базе дистрибутива Ubuntu 18.04.

Настройка провайдера Yandex.Cloud для terraform:

```
arturas@ginzzu: ~/Documents/otus/task-02/terraform
arturas@ginzzu:~/Documents/otus/task-02/terraform$ cat ~/.terraformrc
provider_installation {
  network_mirror {
    url = "https://terraform-mirror.yandexcloud.net/"
    include = ["registry.terraform.io/*/*"]
  }
  direct {
    exclude = ["registry.terraform.io/*/*"]
  }
}
arturas@ginzzu:~/Documents/otus/task-02/terraform$
```

Web-консоль Yandex Cloud:

Виртуальные машины														
Фильтры по имени: Все статусы: Все платформы: Все зоны доступности:														
<input type="checkbox"/> Имя	Статус	ОС	Платформа	vCPU	Доля vCPU	RAM	Прерываемая	Размер дисков	Зона доступности	Внутренний IPv4	Публичный IPv4	Дата создания	Идентификатор	Метки
<input type="checkbox"/> third	Running		Intel Cascade Lake	2	5%	1 GB	Нет	20 GB	ru-central1-a	10.128.0.5	62.84.124.105	14.12.2023, в 14:40	fw1sv174df94e86qnv6	—
<input type="checkbox"/> storage	Running		Intel Cascade Lake	2	5%	1 GB	Нет	20 GB	ru-central1-a	10.128.0.20	84.201.126.208	14.12.2023, в 16:12	fw21s235a13aphn2g89	—
<input type="checkbox"/> first	Running		Intel Cascade Lake	2	5%	1 GB	Нет	20 GB	ru-central1-a	10.128.0.33	51.250.61.19	14.12.2023, в 14:40	fw1nd5gg4jksut7481	—
<input type="checkbox"/> second	Running		Intel Cascade Lake	2	5%	1 GB	Нет	20 GB	ru-central1-a	10.128.0.17	178.154.204.7	14.12.2023, в 14:40	fw1sv1g3781on21feu	—

С помощью Ansible (роль target_initializer) настраивается iSCSI таргет.
Результат после применения указанной роли (сервер storage):

```
ubuntu@storage:~$ sudo targetcli ls
o- / ..... [..]
o- backstores ..... [..]
o- fileio ..... [Storage Objects: 0]
| o- cluster-vhd ..... [/mnt/cluster-vhd.img (1.0GiB) write-back activated] [Storage Objects: 1]
| o- pscsi ..... [Storage Objects: 0]
| o- randisk ..... [Storage Objects: 0]
o- iscsi ..... [Targets: 1]
o- iqn.2023-12.ru.arturas:storage ..... [TPGs: 1]
| o- tpg1 ..... [no-gen-acls, no-auth]
| o- acls ..... [ACLs: 3]
| | o- iqn.2023-12.ru.arturas:first ..... [Mapped LUNs: 1]
| | | o- mapped_lun0 ..... [lun0 fileio/cluster-vhd (rw)] [Mapped LUNs: 1]
| | o- iqn.2023-12.ru.arturas:second ..... [Mapped LUNs: 1]
| | | o- mapped_lun0 ..... [lun0 fileio/cluster-vhd (rw)] [Mapped LUNs: 1]
| | o- iqn.2023-12.ru.arturas:third ..... [Mapped LUNs: 1]
| | | o- mapped_lun0 ..... [lun0 fileio/cluster-vhd (rw)] [Mapped LUNs: 1]
| o- luns ..... [LUNs: 1]
| | o- lun0 ..... [fileio/cluster-vhd (/mnt/cluster-vhd.img)] [LUNs: 1]
| o- portals ..... [Portals: 1]
o- 0.0.0.0:3260 ..... [OK]
o- loopback ..... [Targets: 0]
o- vhost ..... [Targets: 0]
ubuntu@storage:~$
```

В результате:

- настроено блочное устройство (cluster-vhd -> /mnt/cluster-vhd.img).
- настроен iSCSI Target (iqn.2023-12.ru.arturas:storage).
- настроен Portal (10.128.0.14:3260).
- настроен LUN (lun0).
- настроен список Initiators, которые могут подключиться к общему блочному устройству:
 - iqn.2023-12.ru.arturas:first
 - iqn.2023-12.ru.arturas:second
 - iqn.2023-12.ru.arturas:third

С помощью роли pacemaker_initializer устанавливаются все необходимые пакеты для работы с HA (corosync, pacemaker, pcs), ресурсные агенты (fence_agents), устанавливается пароль для пользователя hacluster, копируются предварительно подготовленные файлы (hosts, corosync.conf), перезагружаются соответствующие службы.

С помощью роли clvm_initializer устанавливается пакет clvm и его зависимости (нам важны dlm_controld и lvm2).

С помощью роли initiators_initializer устанавливается клиентский пакет для работы по протоколу iSCSI (open-iscsi), всем инициаторам назначаются IQN, происходит поиск и подключение к порту, настроенному ранее.

С помощью роли cluster_manager происходит авторизация всех узлов кластера, создание, запуск кластера, установка свойств кластера (stonith-enabled, no-quorum-policy), создание ресурсов типа systemd:dlm и osf:heartbeat:clvm, оба – clone, установка порядка запуска ресурсов (сначала dlm, затем clvm).

В результате кластер успешно создан и запущен на всех узлах:

```
ubuntu@first:~$ sudo pcs cluster setup --name debian first.ru-central1.internal second.ru-central1.internal third.ru-central1.internal --force
Destroying cluster on nodes: first.ru-central1.internal, second.ru-central1.internal, third.ru-central1.internal...
first.ru-central1.internal: Stopping Cluster (pacemaker)...
second.ru-central1.internal: Stopping Cluster (pacemaker)...
third.ru-central1.internal: Stopping Cluster (pacemaker)...
third.ru-central1.internal: Successfully destroyed cluster
second.ru-central1.internal: Successfully destroyed cluster
first.ru-central1.internal: Successfully destroyed cluster

Sending 'pacemaker_remote authkey' to 'first.ru-central1.internal', 'second.ru-central1.internal', 'third.ru-central1.internal'
first.ru-central1.internal: successful distribution of the file 'pacemaker_remote authkey'
second.ru-central1.internal: successful distribution of the file 'pacemaker_remote authkey'
third.ru-central1.internal: successful distribution of the file 'pacemaker_remote authkey'
Sending cluster config files to the nodes...
first.ru-central1.internal: Succeeded
second.ru-central1.internal: Succeeded
third.ru-central1.internal: Succeeded

Synchronizing pcsd certificates on nodes first.ru-central1.internal, second.ru-central1.internal, third.ru-central1.internal...
first.ru-central1.internal: Success
second.ru-central1.internal: Success
third.ru-central1.internal: Success
Restarting pcsd on the nodes in order to reload the certificates...
first.ru-central1.internal: Success
second.ru-central1.internal: Success
third.ru-central1.internal: Success
```

Статус кластера:

```
ubuntu@first:~$ sudo pcs status
Cluster name: debian
WARNING: no stonith devices and stonith-enabled is not false
Stack: corosync
Current DC: second.ru-central1.internal (version 1.1.18-2b07d5c5a9) - partition with quorum
Last updated: Thu Dec 14 12:05:01 2023
Last change: Thu Dec 14 11:50:45 2023 by hacluster via crmd on second.ru-central1.internal

3 nodes configured
0 resources configured

Online: [ first.ru-central1.internal second.ru-central1.internal third.ru-central1.internal ]

No resources

Daemon Status:
corosync: active/enabled
pacemaker: active/enabled
pcsd: active/enabled
ubuntu@first:~$ sudo pcs status corosync

Membership information
-----
Nodeid      Votes Name
3           1 third.ru-central1.internal
2           1 second.ru-central1.internal
1           1 first.ru-central1.internal (local)
```

Свойства кластера:

```
ubuntu@first:~$ sudo pcs property
Cluster Properties:
cluster-infrastructure: corosync
cluster-name: debian
dc-version: 1.1.18-2b07d5c5a9
have-watchdog: false
no-quorum-policy: freeze
stonith-enabled: false
```

Ресурсы dlm, clvm успешно запущены на всех узлах кластера:

```
ubuntu@first:~$ sudo pcs status
Cluster name: debian
Stack: corosync
Current DC: first.ru-central1.internal (version 1.1.18-2b07d5c5a9) - partition with quorum
Last updated: Thu Dec 14 12:38:56 2023
Last change: Thu Dec 14 12:37:17 2023 by root via cibadmin on first.ru-central1.internal

3 nodes configured
6 resources configured

Online: [ first.ru-central1.internal second.ru-central1.internal third.ru-central1.internal ]

Full list of resources:

Clone Set: dlm-clone [dlm]
Started: [ first.ru-central1.internal second.ru-central1.internal third.ru-central1.internal ]
Clone Set: clvm-clone [clvm]
Started: [ first.ru-central1.internal second.ru-central1.internal third.ru-central1.internal ]

Daemon Status:
corosync: active/enabled
pacemaker: active/enabled
pcsd: active/enabled
```

Все три узла (first, second, third) подключены к порталу:

```
ubuntu@first:~$ sudo iscsiadm -m session
tcp: [1] 10.128.0.20:3260,1 tqn.2023-12.ru.arturas:storage (non-flash)
```

На всех трех узлах (first, second, third) доступно блочное устройство размером 1 ГБ:

```
ubuntu@first:~$ lsblk
NAME MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
sda 8:0 0 1G 0 disk
vda 252:0 0 20G 0 disk
└─vda1 252:1 0 1M 0 part
└─vda2 252:2 0 20G 0 part /
```

С помощью роли `clusterfs_initializer` устанавливается пакет `gfs2-utils`, изменяется тип блокировки (`locking_type = 3`), создаются физический раздел (`/dev/sda`) группа разделов (`()`) и логический раздел. Физический раздел, группа кластерного раздела (`ClusterVG`) и логический кластерный раздел (`/dev/ClusterVG/ClusterLV`).

Результат представлен на рисунке:

```
ubuntu@first:~$ sudo pvscan
WARNING: Not using lvm2 because locking_type is 3 (clustered).
PV /dev/sda VG ClusterVG lvm2 [1016.00 MiB / 0 free]
Total: 1 [1016.00 MiB] / in use: 1 [1016.00 MiB] / in no VG: 0 [0 ]
ubuntu@first:~$ sudo vgscan
WARNING: Not using lvm2 because locking_type is 3 (clustered).
Reading all physical volumes. This may take a while...
Found volume group "ClusterVG" using metadata type lvm2
ubuntu@first:~$ sudo lvscan
sudo: lvscan: command not found
ubuntu@first:~$ sudo lvscan
WARNING: Not using lvm2 because locking_type is 3 (clustered).
ACTIVE '/dev/ClusterVG/ClusterLV' [1016.00 MiB] inherit
ubuntu@first:~$
```

Форматирование логического раздела в GFS2 (произведено на узле first вручную):

```
Are you sure you want to proceed? [y/n] y
ubuntu@first:~$ sudo mkfs.gfs2 -j 3 -p lock_dlm -t debian:gfs2 /dev/ClusterVG/ClusterLV
/dev/ClusterVG/ClusterLV is a symbolic link to /dev/dm-0
This will destroy any data on /dev/dm-0
Are you sure you want to proceed? [y/n] y
Discarding device contents (may take a while on large devices): Done
Adding journals: Done
Building resource groups: Done
Creating quota file: Done
Writing superblock and syncing: Done
Device: /dev/ClusterVG/ClusterLV
Block size: 4096
Device size: 0.99 GB (260096 blocks)
Filesystem size: 0.99 GB (260094 blocks)
Journals: 3
Resource groups: 6
Locking protocol: "lock_dlm"
Lock table: "debian:gfs2"
UUID: 37a5cb2f-4d22-e95f-8d59-78993e3b5fc3
ubuntu@first:~$
```

Ресурс кластерной ФС также создан вручную.

Статус расемакер после создания ресурса для кластерной файловой системы:

```
pcsd: active/enabled
ubuntu@first:~$ sudo pcs status
Cluster name: debian
Stack: corosync
Current DC: third.ru-central1.internal (version 1.1.18-2b07d5c5a9) - partition with quorum
Last updated: Thu Dec 14 18:22:26 2023
Last change: Thu Dec 14 18:20:07 2023 by root via cibadmin on first.ru-central1.internal

3 nodes configured
9 resources configured

Online: [ first.ru-central1.internal second.ru-central1.internal third.ru-central1.internal ]

Full list of resources:

Clone Set: dlm-clone [dlm]
Started: [ first.ru-central1.internal second.ru-central1.internal third.ru-central1.internal ]
Clone Set: clvm-clone [clvm]
Started: [ first.ru-central1.internal second.ru-central1.internal third.ru-central1.internal ]
Clone Set: ClusterFS-clone [ClusterFS]
Started: [ first.ru-central1.internal second.ru-central1.internal third.ru-central1.internal ]

Daemon Status:
corosync: active/enabled
pacemaker: active/enabled
pcsd: active/enabled
```

Проверка работы кластерной файловой системы:

На узле first создан файл и в него добавлена строка.

На узле second созданный файл успешно прочитан.

```
ubuntu@first:/mnt/gfs2$ sudo chmod -R 777 /mnt/gfs2/
ubuntu@first:/mnt/gfs2$ ls -la
total 8
drwxrwxrwx 2 root root 3864 Dec 14 18:27
drwxr-xr-x 3 root root 4096 Dec 14 18:20 ..
ubuntu@first:/mnt/gfs2$ touch test.txt
ubuntu@first:/mnt/gfs2$ ls -la
total 12
drwxrwxrwx 2 root root 3864 Dec 14 18:27
drwxr-xr-x 3 root root 4096 Dec 14 18:20 ..
-rw-rw-r-- 1 ubuntu ubuntu 0 Dec 14 18:27 test.txt
ubuntu@first:/mnt/gfs2$ echo "hi" >> test.txt
ubuntu@first:/mnt/gfs2$
```

```
ubuntu@second:/mnt/gfs2$ cat test.txt
hi
ubuntu@second:/mnt/gfs2$
```

Проверка завершена. Задание выполнено! Спасибо за новый навык!