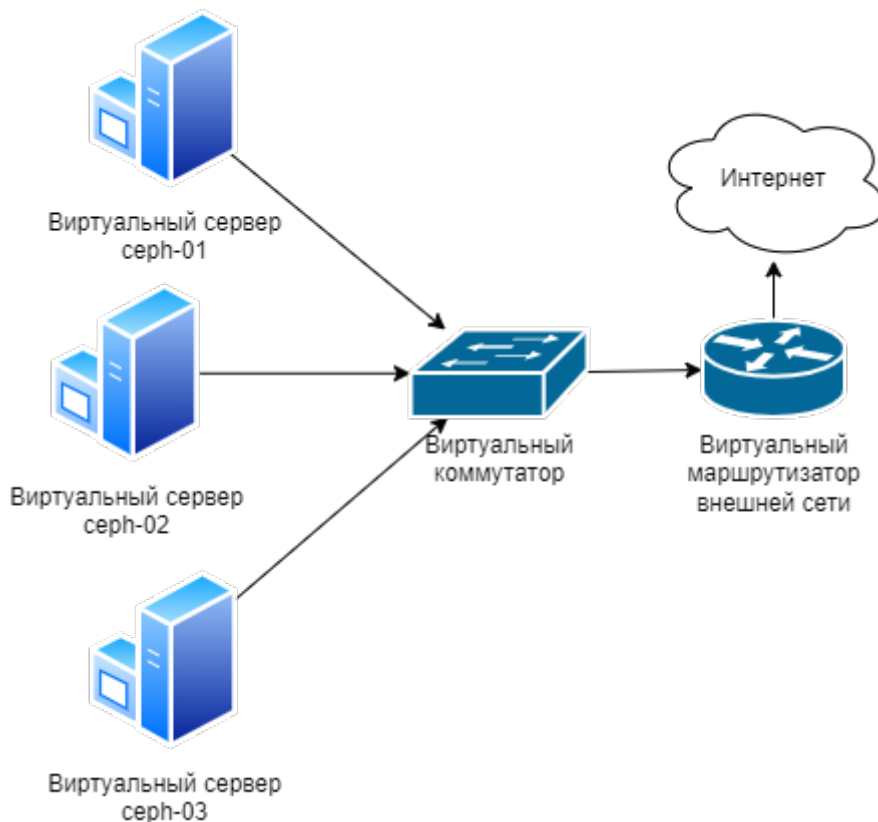


# Практическое задание №5-

## 1. Запуск кластера Ceph

Схема виртуального лабораторного стенда



Для работы с облачной платформой необходимо прочитать [инструкцию](#).

Переключиться на проект [GROUP]:[team]-lab:sandbox.

## Подготовка системы

### Запуск сети

Создание и настройка маршрутизатора

1. Необходимо перейти в раздел **Сеть**, в пункте **Сети**, у вас должна быть доступна сеть external-net, вторую сеть необходимо создать самостоятельно. Название для сети выбираем произвольно, например **my-internal-net**.

Проект / Сеть / Сети								
Сети								
Отображено 6 значений								
<input type="checkbox"/>	Name	Subnets Associated	Shared	External	Status	Admin State	Availability Zones	Actions
<input type="checkbox"/>	external-net	172.17.32 0/19	Да	Да	Активна	UP	-	

2. Создаем подсеть в этой сети аналогично скриншоту ниже:

- В Деталях подсети отключаем DHCP

## Создать сеть

Сеть

Подсеть

Детали подсети

Название подсети

ceph-cluster

Сетевой Адрес ⓘ

192.168.0.0/24

☒ Запретить шлюз

Создает подсеть, связанную с сетью. Необходимо указать правильные "Сетевой адрес" и "IP-адрес шлюза". Если не указан IP-адрес шлюза, то по умолчанию присваивается первое значение сети. того чтобы не использовать шлюз, включите опцию "Выключить шлюз". Дополнительные параметры доступны на вкладке "Сведения о подсети".

Отмена

« Назад

Следующий

Создание и настройка портов для инстансов

1. Заходим в созданную ранее сеть -> Порты :

- Нажимаем Создать порт
- Заполняем поля формы:
  - Имя: ceph-port-1
  - IP адрес или подсеть: Фиксированный IP-адрес
  - Фиксированный IP-адрес: 192.168.0.21

- Выключить Безопасность порта

## Создать порт

### Информация

Имя

ceph-port-1

☒ Разрешить Admin State ⓘ

ID устройства ⓘ

Владелец устройства ⓘ

Укажите IP адрес или подсеть ⓘ

Фиксированный IP-адрес ▼

Фиксированный IP-адрес\* ⓘ

192.168.0.21

MAC адрес ⓘ

☐ Безопасность порта ⓘ

### Описание:

Вы можете создать порт сети. Если вы укажет устройства, то это устройство будет подключе созданному порту.

- Аналогично создаём порты ceph-port-2 и ceph-port-3 с адресами 192.168.0.22 и 192.168.0.23 соответственно

## Создание инстансов

1. Запустить 3 инстанса с характеристиками:
  - Имена инстансов: **ceph-01, ceph-02, ceph-03**;
  - Источник: Образ **Ubuntu-server-20.04**;
  - Тип инстанса: **small**;
  - Сети: **external\_net**;
  - Сетевые порты: **port-ceph-1, port-ceph-2, port-ceph-3** для каждого инстанса соответственно.
2. Создать 3 диска по 2 Гб, подключить по одному диску на каждый инстанс. Список подключенных дисков может иметь следующий вид:

## Диски

<div> <div>Фильтр</div> <div>+</div> <div>Создать диск (Превышена квота)</div> <div>≡</div> <div>Принять передачу (Превышена квота)</div> <div>Удалить д</div> </div>											
Отображено 6 значений											
<input type="checkbox"/>	Name	Description	Size	Status	Group	Type	Attached To	Availability Zone	Bootable	Encrypted	Actions
<input type="checkbox"/>	ceph-storage-03	-	2 ГиБ	Используется	-	Ceph RBD	/dev/vdb в ceph-03	nova	Нет	Нет	Редактировать диск
<input type="checkbox"/>	ceph-storage-02	-	2 ГиБ	Используется	-	Ceph RBD	/dev/vdb в ceph-02	nova	Нет	Нет	Редактировать диск
<input type="checkbox"/>	ceph-storage-01	-	2 ГиБ	Используется	-	Ceph RBD	/dev/vdb в ceph-01	nova	Нет	Нет	Редактировать диск

3. На каждом инстансе добавить записи в файл `/etc/hosts` (после подстановки адресов, кавычек быть не должно):

```
192.168.0.21 ceph-01
192.168.0.22 ceph-02
192.168.0.23 ceph-03

"внешний адрес узла #1" ext-ceph-01
"внешний адрес узла #2" ext-ceph-02
"внешний адрес узла #3" ext-ceph-03
```

4. Создать пару ключей на ceph-01 для соединения между инстансами(вводить никакие значения не требуется, на все вопросы оставить значение по умолчанию и нажать Enter)

```
ssh-keygen
```

Далее задать пароли для пользователя cloudadmin на узлах ceph-01,ceph-02,ceph-03

```
sudo passwd cloudadmin
```

“ пароль может быть любой

и скопировать публичный ключ ceph-01 на каждый инстанс(в том числи и на сам ceph-01 для автоматизации):

```
ssh-copy-id ceph-01
```

на второй:

```
ssh-copy-id ceph-02
```

И на третий:

```
ssh-copy-id ceph-03
```

## Запуск кластера Ceph

### Установка Ceph

На данном этапе необходимо установить пакет `ceph` на каждый узел.

Для этого нужно предварительно добавить ключ безопасности репозитория в список доверенных и добавить репозиторий для скачивания указанных пакетов.

1. Для этого на каждом узле выполнить следующие команды:

```
sudo apt update
sudo apt install -y ca-certificates
wget -q -O- 'https://download.ceph.com/keys/release.asc' | sudo apt-key add -
sudo apt-add-repository 'deb https://download.ceph.com/debian-quincy/ focal main'
sudo apt update && sudo apt install ceph ceph-mds -y
```

- Либо использовать скрипт:

#### Скрипт для запуска на ceph-01

```
#!/bin/bash

for NODE in ceph-01 ceph-02 ceph-03
do
    ssh $NODE \
        "sudo apt update; \
        sudo apt install -y ca-certificates; \
        wget -q -O- 'https://download.ceph.com/keys/release.asc' | sudo apt-key add -; \
        sudo apt-add-repository 'deb https://download.ceph.com/debian-quincy/ focal main'; \
        sudo apt update && sudo apt install ceph ceph-mds -y"
done
```

2. Проверить статус установки можно с помощью команды:

```
sudo ceph -v
```

## Запуск демонов ceph-mon, ceph-mgr, ceph-osd, ceph-mds

Подготовка к запуску демонов

На узле ceph-01 необходимо выполнить следующие действия:

1. Получить уникальный идентификатор для кластера (можно с помощью команды `uuidgen`):  
`ad56ab6d-7f7a-4ee4-8b02-9c7ef1ddb438` - используется в качестве примера
2. Создать файл конфигурации для кластера по пути `/etc/ceph/ceph.conf`. Необходимо вставить свои значения параметров: `fsid`, адреса интерфейсов внешней сети.  
`vi /etc/ceph/ceph.conf`:

```
[global]
fsid = ad56ab6d-7f7a-4ee4-8b02-9c7ef1ddb438
mon initial members = ceph-01, ceph-02, ceph-03
mon host = ext-ceph-01, ext-ceph-02, ext-ceph-03
cluster network = 192.168.0.0/24
public network = 172.17.32.0/19
auth cluster required = cephx
auth service required = cephx
auth client required = cephx
osd pool default size = 3
osd pool default min size = 2
osd pool default pg num = 64
osd pool default pgp num = 64
osd crush chooseleaf type = 1
```

- Разберём параметры конфигурации:

- `fsid` - идентификатор кластера
- `mon initial members` - имена узлов стартовых мониторов
- `mon host` - адреса стартовых мониторов
- `cluster network` - сеть кластера в формате CIDR, в которой будут передаваться heartbeat-сигналы, реплики данных между демонами OSD
- `public network` - внешняя сеть для взаимодействия с клиентами кластера
- `auth cluster/service/client required` - включение авторизации между демонами кластера и подключаемыми клиентами
- `osd pool default size` - количество реплик по умолчанию
- `osd pool default min size` - минимальное количество записанных реплик для объектов, чтобы подтвердить операцию ввода-вывода клиенту
- `osd pool default pg num` - количество групп размещения по умолчанию для пула
- `osd pool default pgp num` - количество групп размещения по умолчанию для размещения пула. PG и PGP должны быть равны
- `osd crush chooseleaf type` - максимальный уровень размещения кластера (при значении больше 1 имеется возможность создания групп узлов, таким образом можно разделить иерархию узлов на стойки, ЦОДы и другие физически разнесённые группы)

3. Создать связку ключей для своего кластера и сгенерировать секретный ключ монитора. Также создать связку ключей администратора, `client.admin` пользователя и добавить пользователя в связку ключей. Создать связку ключей `bootstrap-osd`, `client.bootstrap-osd` пользователя и добавить пользователя в связку ключей:

```
sudo ceph-authtool --create-keyring /tmp/ceph.mon.keyring --gen-key -n mon. --cap mon 'allow *'
sudo ceph-authtool --create-keyring /etc/ceph/ceph.client.admin.keyring \
--gen-key -n client.admin --cap mon 'allow *' --cap osd 'allow *' \
--cap mds 'allow *' --cap mgr 'allow *'
sudo ceph-authtool --create-keyring /var/lib/ceph/bootstrap-osd/ceph.keyring \
--gen-key -n client.bootstrap-osd --cap mon 'profile bootstrap-osd' \
--cap mgr 'allow r'
```

4. Добавить сгенерированные ключи в связку `ceph.mon.keyring`:

```
sudo ceph-authtool /tmp/ceph.mon.keyring --import-keyring /etc/ceph/ceph.client.admin.keyring
sudo ceph-authtool /tmp/ceph.mon.keyring --import-keyring /var/lib/ceph/bootstrap-osd/ceph.keyring
```

5. Сгенерировать карту мониторов, используя имена узлов, внешние адреса (записи из файла `/etc/hosts` здесь не работают) и FSID. Сохранить как `/tmp/monmap`:

```
monmaptool --create --fsid ad56ab6d-7f7a-4ee4-8b02-9c7ef1ddb438 \
--add ceph-01 "внешний адрес узла #1" --add ceph-02 "внешний адрес узла #2" \
--add ceph-03 "внешний адрес узла #3" /tmp/monmap
```

#### Запуск ceph-mon

1. Необходимо скопировать файлы `/etc/ceph/ceph.conf`, `/etc/ceph/ceph.client.admin.keyring`, `/tmp/monmap`, `/tmp/ceph.mon.keyring`, `/var/lib/ceph/bootstrap-osd/ceph.keyring` с `ceph-01` на `ceph-02` и `ceph-03`.

- Примеры выполнения данного пункта:

#### Пошаговые действия с описанием

1. Создадим папку `/tmp/ceph_files` и скопируем в неё все необходимые файлы:

```
mkdir /tmp/ceph_files
sudo cp /etc/ceph/ceph.conf /tmp/ceph_files/ceph.conf
sudo cp /etc/ceph/ceph.client.admin.keyring
/tmp/ceph_files/ceph.client.admin.keyring
sudo cp /tmp/monmap /tmp/ceph_files/monmap
sudo cp /tmp/ceph.mon.keyring /tmp/ceph_files/ceph.mon.keyring
```

```
sudo cp /var/lib/ceph/bootstrap-osd/ceph.keyring /tmp/ceph_files/ceph.keyring
```

2. Далее добавим общие права на чтение для всех файлов в папке и отправим эти файлы на ceph-02 и ceph-03. Удалим созданную ранее папку:

```
sudo chmod 644 /tmp/ceph_files/*  
scp -r /tmp/ceph_files/ ceph-02:~  
scp -r /tmp/ceph_files/ ceph-03:~  
sudo rm -rf /tmp/ceph_files
```

3. На ceph-02 и ceph-03 выполняем следующие команды (возвращаем файлам их исходные права, копируем их по нужным путям и удаляем полученные файлы:

```
cd ~/ceph_files  
sudo chmod 600 ceph.client.admin.keyring ceph.mon.keyring ceph.keyring  
sudo cp ceph.conf /etc/ceph/ceph.conf  
sudo cp ceph.client.admin.keyring /etc/ceph/ceph.client.admin.keyring  
sudo cp monmap /tmp/monmap  
sudo cp ceph.mon.keyring /tmp/ceph.mon.keyring  
sudo cp ceph.keyring /var/lib/ceph/bootstrap-osd/ceph.keyring  
cd ..  
rm -rf ceph_files
```

### Скрипт для запуска на ceph-01

```
#!/bin/bash  
  
mkdir /tmp/ceph_files  
sudo cp /etc/ceph/ceph.conf /tmp/ceph_files/ceph.conf  
sudo cp /etc/ceph/ceph.client.admin.keyring /tmp/ceph_files/ceph.client.admin.keyring  
sudo cp /tmp/monmap /tmp/ceph_files/monmap  
sudo cp /tmp/ceph.mon.keyring /tmp/ceph_files/ceph.mon.keyring  
sudo cp /var/lib/ceph/bootstrap-osd/ceph.keyring /tmp/ceph_files/ceph.keyring  
  
sudo chmod 644 /tmp/ceph_files/*  
scp -r /tmp/ceph_files/ ceph-02:~
```

```

scp -r /tmp/ceph_files ceph-03:~
sudo rm -rf /tmp/ceph_files

for NODE in ceph-02 ceph-03
do
ssh $NODE \
  "cd ~/ceph_files; \
  sudo chmod 600 ceph.client.admin.keyring ceph.mon.keyring ceph.keyring; \
  sudo cp ceph.conf /etc/ceph/ceph.conf; \
  sudo cp ceph.client.admin.keyring /etc/ceph/ceph.client.admin.keyring; \
  sudo cp monmap /tmp/monmap; \
  sudo cp ceph.mon.keyring /tmp/ceph.mon.keyring; \
  sudo cp ceph.keyring /var/lib/ceph/bootstrap-osd/ceph.keyring; \
  cd ..; \
  rm -rf ceph_files"
done

```

2. Далее создадим каталог для данных мониторов на узлах. Заполним демоны монитора картой мониторов и набором ключей. Это можно сделать с помощью следующих команд на каждом хосте:

```

sudo chown ceph:ceph /tmp/ceph.mon.keyring
sudo -u ceph mkdir /var/lib/ceph/mon/ceph-$HOSTNAME
sudo -u ceph ceph-mon --mkfs -i $HOSTNAME \
  --monmap /tmp/monmap --keyring /tmp/ceph.mon.keyring
sudo systemctl enable ceph-mon@$HOSTNAME --now

```

- Либо использовать скрипт:

#### Скрипт для запуска на ceph-01

```

#!/bin/bash

for NODE in ceph-01 ceph-02 ceph-03
do
ssh $NODE \
  [ "sudo chown ceph:ceph /tmp/ceph.mon.keyring; \
  [ sudo -u ceph mkdir /var/lib/ceph/mon/ceph-$NODE; \
  [ sudo -u ceph ceph-mon --mkfs -i $NODE \

```

```
❏ --monmap /tmp/monmap --keyring /tmp/ceph.mon.keyring; \  
❏ sudo systemctl enable ceph-mon@$NODE --now"  
done
```

3. Включим протокол `MESSENGER V2`, отключим небезопасное использование клиентами глобальных идентификаторов и установим вывод предупреждений о малом свободном пространстве на основном диске хоста при остатке в 10% (иначе без этих параметров кластер может иметь статус `HEALTH_WARN`). Также установим параметр, позволяющий удалять пулы, в значение `true`:

```
sudo ceph mon enable-msgr2  
sudo ceph config set mon auth_allow_insecure_global_id_reclaim false  
sudo ceph config set mon mon_data_avail_warn 10  
sudo ceph config set mon mon_allow_pool_delete true
```

4. Так как ещё не добавлено никаких устройств хранения, то в выводе состояния нас интересует только образовавшийся кворум мониторов:

```
sudo ceph -s
```

#### Запуск ceph-mgr

На каждом узле, где вы запускаете демон `ceph-mon`, вы также должны настроить демон `ceph-mgr`. Сначала создадим ключ аутентификации для демона, поместим этот ключ в путь `/var/lib/ceph/mgr/ceph-$(HOSTNAME)/keyring` и запустим демон `ceph-mgr`.

1. Для этого необходимо выполнить данные действия на каждом хосте:

```
sudo ceph auth get-or-create mgr.$HOSTNAME mon 'allow profile mgr' osd 'allow *' mds 'allow *'  
sudo ceph auth get-or-create mgr.$HOSTNAME > /tmp/mgr.admin.keyring  
sudo -u ceph mkdir /var/lib/ceph/mgr/ceph-$(HOSTNAME)  
sudo cp /tmp/mgr.admin.keyring /var/lib/ceph/mgr/ceph-$(HOSTNAME)/keyring  
sudo chown -R ceph. /var/lib/ceph/mgr/ceph-$(HOSTNAME)  
sudo systemctl enable --now ceph-mgr@$HOSTNAME
```

- Либо использовать скрипт:

#### Скрипт для запуска на ceph-01

```
#!/bin/bash  
  
for NODE in ceph-01 ceph-02 ceph-03
```

```
do
    MGR_PATH=/var/lib/ceph/mgr/ceph-$NODE;
    ssh $NODE \
    [ "sudo ceph auth get-or-create mgr.$NODE \
    [ mon 'allow profile mgr' osd 'allow *' mds 'allow *'; \
    [ sudo ceph auth get-or-create mgr.$NODE > /tmp/mgr.admin.keyring; \
    [ sudo -u ceph mkdir $MGR_PATH; \
    [ sudo cp /tmp/mgr.admin.keyring $MGR_PATH/keyring; \
    [ sudo chown -R ceph. $MGR_PATH; \
    [ sudo systemctl enable --now ceph-mgr@$NODE"
done
```

2. На данном этапе мы можем увидеть статус запуска демонов менеджеров:

```
sudo ceph -s
```

Запуск ceph-osd

Серв предоставляет `ceph-volume` утилиту, которая может подготовить логический том, диск или раздел для использования с Серв. Утилита `ceph-volume` создает идентификатор OSD путем увеличения индекса. Кроме того, `ceph-volume` добавит новый OSD в карту CRUSH хоста.

1. На каждом хосте выполнить:

```
sudo ceph-volume lvm create --data /dev/vdb
```

- Либо использовать скрипт:

#### Скрипт для запуска на ceph-01

```
#!/bin/bash

for NODE in ceph-01 ceph-02 ceph-03
do
    ssh $NODE "sudo ceph-volume lvm create --data /dev/vdb"
done
```

2. На данном этапе кластер должен иметь статус HEALTH\_OK:

```
sudo ceph -s
```

3. Также подключенные OSD можно получить с помощью команды:

```
sudo ceph osd tree
```

## Запуск услуг кластера

### Создание пула и блочного устройства

1. Создадим пул `rbd` и инициализируем его:

```
sudo ceph osd pool create rbd 32  
sudo rbd pool init rbd
```

2. Получить список созданных пулов с помощью команды:

```
sudo ceph osd pool ls
```

3. Создать блочное устройство в новом пуле:

```
sudo rbd create --size 10G --pool rbd rbd01
```

#### Проверка

1. Подключаем созданное устройство к `ceph-01`:

```
sudo rbd map rbd01
```

2. Получить список подключённых устройств можно с помощью команды:

```
sudo rbd showmapped
```

3. Установим файловую систему на подключенное устройство и примонтируем его к каталогу `/mnt/ceph_rbd`:

```
sudo mkfs.xfs /dev/rbd0  
sudo mkdir /mnt/ceph_rbd  
sudo mount /dev/rbd0 /mnt/ceph_rbd/
```

4. На `ceph-01` создать файл с произвольным текстом и поместить его в каталог `/mnt/ceph_rbd`. Затем отмонтировать блочное устройство:

```
echo "my text for ceph rbd" | sudo tee /mnt/ceph_rbd/rbd.txt  
sudo umount /dev/rbd0  
sudo rbd unmap /dev/rbd/rbd/rbd01
```

5. На `ceph-02` подключить использованное ранее блочное устройство и примонтировать его к каталогу `/mnt/ceph_rbd`:

```
sudo rbd map rbd01  
sudo mkdir /mnt/ceph_rbd  
sudo mount /dev/rbd0 /mnt/ceph_rbd/
```

6. Проверить содержимое подключенного блочного устройства:

```
ls /mnt/ceph_rbd/  
cat /mnt/ceph_rbd/rbd.txt
```

7. Отмонтировать блочное устройство и удалить его, удалить пул:

```
sudo umount /dev/rbd0  
sudo rbd unmap /dev/rbd/rbd/rbd01  
sudo rbd rm rbd01 -p rbd  
sudo ceph osd pool delete rbd rbd --yes-i-really-really-mean-it
```

---

Версия #13

Артем Швидкий создал 3 ноября 2022 16:34:41

Тарабанов Илья Федорович обновил 12 декабря 2023 13:45:57