CEPHFS 安装配置 nfs-ganesha

存储系统 孔樱【技术大咖】 2024-03-31 发表

组网及说明

X10000以及OneStor在5.2版本使用了 nfs-ganesha 功能，在原生态 ceph 也部署 nfs-ganesha , 学习底层概念及架构。

配置步骤

CEPHFS 安装配置 nfs-ganesha

所有节点都需要安装 nfs-ganesha

```
[root@ceph-master yum.repos.d]# vim nfs-ganesha.repo
[nfsganesha]
name=nfsganesha
baseurl=https://mirrors.cloud.tencent.com/ceph/nfs-ganesha/rpm-V2.8-stable/nautilus/x86_64/
gpgcheck=0
enable=1

[root@ceph-master yum.repos.d]# yum makecache
Loaded plugins: fastestmirror, langpacks, priorities
Loading mirror speeds from cached hostfile
epel/x86_64/metalink
                | 6.0 kB  00:00:00
 * base: mirrors.jlu.edu.cn
 * epel: mirror.nyist.edu.cn
 * extras: mirrors.jlu.edu.cn
 * updates: mirrors.jlu.edu.cn
Ceph
       | 1.5 kB  00:00:00
Ceph-
noarch
   | 1.5 kB  00:00:00
base
      | 3.6 kB  00:00:00
ceph-
source
   | 1.5 kB  00:00:00
epel
      | 4.7 kB  00:00:00
extras
       | 2.9 kB  00:00:00
nfsganesha
           | 2.9 kB  00:00:00
updates
        | 2.9 kB  00:00:00
(1/8):
epel/x86_64/filelists_db
         |  12 MB  00:00:01
(2/8):
epel/x86_64/updateinfo
           | 1.0 MB  00:00:00
(3/8):
epel/x86_64/prestodelta
           |  576 B  00:00:00
(4/8):
epel/x86_64/primary_db
            | 7.0 MB  00:00:00
(5/8):
epel/x86_64/other_db
           | 3.4 MB  00:00:00
(6/8):
nfsganesha/filelists_db
```

```
         |  13 kB  00:00:00
(7/8): 
nfsganesha/primary_db
            |  19 kB  00:00:00
(8/8): 
nfsganesha/other_db
           | 2.2 kB  00:00:00
Metadata Cache Created
[root@ceph-master yum.repos.d]#
```

所有节点安装软件包

```
[root@ceph-master cephcluster]# yum install -y nfs-ganesha nfs-ganesha-ceph nfs-ganesha-rados-gr
ace nfs-ganesha-rgw nfs-utils rpcbind haproxy keepalived
Loaded plugins: fastestmirror, langpacks, priorities
Loading mirror speeds from cached hostfile
 * base: mirrors.jlu.edu.cn
 * epel: mirror.nyist.edu.cn
 * extras: mirrors.jlu.edu.cn
 * updates: mirrors.jlu.edu.cn
8 packages excluded due to repository priority protections
Package 1:nfs-utils-1.3.0-0.68.el7.2.x86_64 already installed and latest version
Package rpcbind-0.2.0-49.el7.x86_64 already installed and latest version
Package haproxy-1.5.18-9.el7_9.1.x86_64 already installed and latest version
Package keepalived-1.3.5-19.el7.x86_64 already installed and latest version
Resolving Dependencies
--> Running transaction check
---> Package nfs-ganesha.x86_64 0:2.8.1.2-0.1.el7 will be installed
--> Processing Dependency: libntirpc = 1.8.0 for package: nfs-ganesha-2.8.1.2-0.1.el7.x86_64
--> Processing Dependency: libntirpc.so.1.8(NTIRPC_1.8.0)(64bit) for package: nfs-ganesha-2.8.1.2-
0.1.el7.x86_64
--> Processing Dependency: libntirpc.so.1.8()(64bit) for package: nfs-ganesha-2.8.1.2-0.1.el7.x86_64
---> Package nfs-ganesha-ceph.x86_64 0:2.8.1.2-0.1.el7 will be installed
---> Package nfs-ganesha-rados-grace.x86_64 0:2.8.1.2-0.1.el7 will be installed
---> Package nfs-ganesha-rgw.x86_64 0:2.8.1.2-0.1.el7 will be installed
--> Running transaction check
---> Package libntirpc.x86_64 0:1.8.0-0.1.el7 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

================================================================================
================================================================================
==========================================
 Package                              Arch                     Version
Repository                      Size
================================================================================
================================================================================
==========================================
Installing:
 nfs-ganesha                              x86_64                       2.8.1.2-
0.1.el7                 nfsganesha                       680 k
 nfs-ganesha-ceph                         x86_64                       2.8.1.2-
0.1.el7                 nfsganesha                        30 k
 nfs-ganesha-rados-grace                    x86_64                       2.8.1.2-0.1.el7
      nfsganesha                       8.2 k
 nfs-ganesha-rgw                          x86_64                       2.8.1.2-
0.1.el7                 nfsganesha                        21 k
Installing for dependencies:
 libntirpc                             x86_64                       1.8.0-0.1.el7
nfsganesha                       113 k

Transaction Summary
================================================================================
================================================================================
==========================================
Install  4 Packages (+1 Dependent package)

Total download size: 852 k
Installed size: 2.3 M
Downloading packages:
(1/5): nfs-ganesha-2.8.1.2-
0.1.el7.x86_64.rpm                                                                 | 68
0 kB  00:00:01
(2/5): libntirpc-1.8.0-
0.1.el7.x86_64.rpm                                                                   |
```

```
113 kB  00:00:01
(3/5): nfs-ganesha-ceph-2.8.1.2-
0.1.el7.x86_64.rpm                                                                                  |  30 k
B  00:00:00
(4/5): nfs-ganesha-rados-grace-2.8.1.2-
0.1.el7.x86_64.rpm                                                                              | 8.2 kB  0
0:00:00
(5/5): nfs-ganesha-rgw-2.8.1.2-
0.1.el7.x86_64.rpm                                                                                  |  21 k
B  00:00:00
--------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------
Total
324 kB/s | 852 kB  00:00:02
Running transaction check
Running transaction test
Transaction test succeeded
Running transaction
  Installing : libntirpc-1.8.0-
0.1.el7.x86_64
1/5
  Installing : nfs-ganesha-2.8.1.2-
0.1.el7.x86_64
2/5
  Installing : nfs-ganesha-rados-grace-2.8.1.2-
0.1.el7.x86_64                                                                                           3/5
  Installing : nfs-ganesha-rgw-2.8.1.2-
0.1.el7.x86_64                                                                                                4
/5
  Installing : nfs-ganesha-ceph-2.8.1.2-
0.1.el7.x86_64                                                                                               5/
5
  Verifying  : libntirpc-1.8.0-
0.1.el7.x86_64
1/5
  Verifying  : nfs-ganesha-rados-grace-2.8.1.2-
0.1.el7.x86_64                                                                                           2/5
  Verifying  : nfs-ganesha-2.8.1.2-
0.1.el7.x86_64
3/5
  Verifying  : nfs-ganesha-rgw-2.8.1.2-
0.1.el7.x86_64                                                                                                4
/5
  Verifying  : nfs-ganesha-ceph-2.8.1.2-
0.1.el7.x86_64                                                                                               5/
5

Installed:
  nfs-ganesha.x86_64 0:2.8.1.2-0.1.el7      nfs-ganesha-ceph.x86_64 0:2.8.1.2-0.1.el7      nfs-
ganesha-rados-grace.x86_64 0:2.8.1.2-0.1.el7      nfs-ganesha-rgw.x86_64 0:2.8.1.2-0.1.el7

Dependency Installed:
  libntirpc.x86_64 0:1.8.0-0.1.el7

Complete!
[root@ceph-master cephcluster]#
```

在一个节点上创建三个目录作为共享使用

```
[root@ceph-master cephcluster]# mkdir -p /fsdata
```

挂载 cephfs

```
[root@ceph-master cephcluster]# :  q
[root@ceph-master cephcluster]# df -kh
Filesystem                                 Size  Used Avail Use% Mounted on
devtmpfs                                   3.9G     0  3.9G   0% /dev
tmpfs                                      3.9G  4.0K  3.9G   1% /dev/shm
tmpfs                                      3.9G   58M  3.8G   2% /run
tmpfs                                      3.9G     0  3.9G   0% /sys/fs/cgroup
/dev/mapper/centos-root                     90G  6.7G   84G   8% /
/dev/sda1                                 1014M  240M  775M  24% /boot
/dev/mapper/centos-home                    152G   62M  151G   1% /home
tmpfs                                      783M   12K  783M   1% /run/user/42
tmpfs                                      3.9G   52K  3.9G   1% /var/lib/ceph/osd/ceph-0
tmpfs                                      3.9G   52K  3.9G   1% /var/lib/ceph/osd/ceph-1
tmpfs                                      3.9G   52K  3.9G   1% /var/lib/ceph/osd/ceph-2
tmpfs                                      3.9G   52K  3.9G   1% /var/lib/ceph/osd/ceph-3
tmpfs                                      3.9G   52K  3.9G   1% /var/lib/ceph/osd/ceph-4
tmpfs                                      783M     0  783M   0% /run/user/0
10.12.180.122,10.12.180.123,10.12.180.124:6789:/  432G     0  432G   0% /fsdata
[root@ceph-master cephcluster]#
```

```
[root@ceph-master cephcluster]# mkdir -p /fsdata/nfs1      >>> 两个不同用户挂载不同目录
[root@ceph-master cephcluster]# mkdir -p /fsdata/nfs2
[root@ceph-master cephcluster]#

[root@ceph-master fsdata]# ls -al
total 0
drwxr-xr-x   1 root root   2 Mar 21 11:13 .
dr-xr-xr-x. 19 root root 270 Mar 21 09:48 ..
drwxr-xr-x   1 root root   0 Mar 21 11:13 nfs1
drwxr-xr-x   1 root root   0 Mar 21 11:13 nfs2
[root@ceph-master fsdata]#
```

修改 ganesha.conf 配置文件

```
[root@ceph-master cephcluster]# vim /etc/ganesha/ganesha.conf
NFS_CORE_PARAM {
  Enable_NLM = false;
  NFS_Port = 52049;
  Enable_RQUOTA = false;
}
EXPORT_DEFAULTS {
  Access_Type = RW;
#  Anonymous_uid = 65534;
#  Anonymous_gid = 65534;
}
LOG {
  Default_Log_Level = INFO;
#  Facility {
#    name = FILE;
#    description = "/var/log/ganesha/ganesha.log";
#    enable = active;
#  }
}

NFSv4 {
  #DelegatiOns= false;
  #RecoveryBackend = 'rados_cluster';
  #Minor_VersiOns= 1,2
}

EXPORT
{
    Export_Id = 1;
    Path = /nfs1;
    Pseudo =/fsdata;
    Squash = no_root_squash;
    protocols = 3,4;
    transports = "UDP", "TCP";
    Access_Type = RW;
    FSAL {
      secret_access_key = "AQCn9ttlJgcHBxAAbxEhVhixzxII/7zOD0+A3A==";
      user_id = "admin";
      name = "CEPH";
      filesystem = "cephfs";
    }
}
EXPORT
{
    Export_Id = 2;
    Path = /nfs2;
    Pseudo =/fsdata/nfs2;
    Squash = no_root_squash;
    protocols = 3,4;
    transports = "UDP", "TCP";
    Access_Type = RW;
    FSAL {
      secret_access_key = "AQCn9ttlJgcHBxAAbxEhVhixzxII/7zOD0+A3A==";
      user_id = "admin";
      name = "CEPH";
      filesystem = "cephfs";
    }
}

[root@ceph-master ganesha]#
```

在所有节点启动 nfs-ganesha 服务

```
[root@ceph-master ganesha]# systemctl restart nfs-ganesha
[root@ceph-master ganesha]# systemctl status nfs-ganesha
● nfs-ganesha.service - NFS-Ganesha file server
   Loaded: loaded (/usr/lib/systemd/system/nfs-ganesha.service; enabled; vendor preset: disabled)
   Active: active (running) since Wed 2024-03-27 08:53:01 CST; 19s ago
     Docs: http://github.com/nfs-ganesha/nfs-ganesha/wiki
  Process: 385263 ExecStop=/bin/dbus-send --system --dest=org.ganesha.nfsd --type=method_call /o
rg/ganesha/nfsd/admin org.ganesha.nfsd.admin.shutdown (code=exited, status=0/SUCCESS)
  Process: 385826 ExecStartPost=/bin/bash -c /usr/bin/sleep 2 && /bin/dbus-send --system  --dest=or
g.ganesha.nfsd --type=method_call /org/ganesha/nfsd/admin  org.ganesha.nfsd.admin.init_fds_limit (
code=exited, status=0/SUCCESS)
  Process: 385824 ExecStartPost=/bin/bash -c prlimit --pid $MAINPID --nofile=$NOFILE:$NOFILE (co
de=exited, status=0/SUCCESS)
  Process: 385821 ExecStart=/bin/bash -c ${NUMACTL} ${NUMAOPTS} /usr/bin/ganesha.nfsd ${OP
TIONS} ${EPOCH} (code=exited, status=0/SUCCESS)
 Main PID: 385823 (ganesha.nfsd)
    Tasks: 293
   CGroup: /system.slice/nfs-ganesha.service
           └─385823 /usr/bin/ganesha.nfsd -L /var/log/ganesha/ganesha.log -f
/etc/ganesha/ganesha.conf -N NIV_EVENT

Mar 27 08:52:59 ceph-master.h3cu.com systemd[1]: Starting NFS-Ganesha file server...
Mar 27 08:52:59 ceph-master.h3cu.com bash[385821]: libust[385821/385821]: Warning: HOME envir
onment variable not set. Disabling LTTng-UST per-user tracing. (in setup_local_apps()...-comm.c:305
)
Mar 27 08:53:01 ceph-master.h3cu.com systemd[1]: Started NFS-Ganesha file server.
Hint: Some lines were ellipsized, use -l to show in full.
[root@ceph-master ganesha]#
```

因为 nfs-ganesha 启用的是 52049 端口，查看端口是否开启侦听

```
[root@ceph-master ganesha]# ss -tulpn|grep 52049
udp    UNCONN     0      0      [::]:52049              [::]:*                  users:(("ganesha.nfsd",pid=385823,fd
=23))
tcp    LISTEN     0      128    [::]:52049              [::]:*                  users:
(("ganesha.nfsd",pid=385823,fd=24))
[root@ceph-master ganesha]#
```

给 ganesha 配置一个 namespace 命名空间

```
[root@ceph-master ganesha]# ceph dashboard set-ganesha-clusters-rados-pool-namespace cephfs-
ns
Option GANESHA_CLUSTERS_RADOS_POOL_NAMESPACE updated
[root@ceph-master ganesha]#

[root@ceph-master ganesha]# ceph dashboard get-ganesha-clusters-rados-pool-namespace
cephfs-ns
[root@ceph-master ganesha]#
```

Linux 客户端挂载 cephfs nfs 共享目录

```
[root@centos7-c630fc ~]# showmount -e cephdns          >> 通过 dns 负载均衡的域名访问
Export list for cephdns:
/nfs1 (everyone)
/nfs2 (everyone)
[root@centos7-c630fc ~]#

[root@centos7-c630fc ~]# mount -vvv -t nfs cephdns:/nfs1 /nfs1
mount.nfs: timeout set for Thu Mar 28 09:44:14 2024
mount.nfs: trying text-based options 'vers=4.1,addr=10.12.180.124,clientaddr=10.1.1.110'
mount.nfs: mount(2): No such file or directory
mount.nfs: trying text-based options 'addr=10.12.180.124'
mount.nfs: prog 100003, trying vers=3, prot=6
mount.nfs: trying 10.12.180.124 prog 100003 vers 3 prot TCP port 52049
mount.nfs: prog 100005, trying vers=3, prot=17
mount.nfs: trying 10.12.180.124 prog 100005 vers 3 prot UDP port 55389
[root@centos7-c630fc ~]#

[root@centos7-c630fc ~]# df -kh
```

```
Filesystem      Size  Used Avail Use% Mounted on
devtmpfs        3.8G     0  3.8G   0% /dev
tmpfs           3.9G  1.9G  2.0G  48% /dev/shm
tmpfs           3.9G   17M  3.9G   1% /run
tmpfs           3.9G     0  3.9G   0% /sys/fs/cgroup
/dev/vda1       300G   27G  274G   9% /
/dev/vdb        493G   11G  457G   3% /ora11g
tmpfs           783M     0  783M   0% /run/user/0
cephdns:/nfs1   433G   11G  423G   3% /nfs1
[root@centos7-c630fc ~]#

[root@centos7-c630fc ~]# mount
cephdns:/nfs1 on /nfs1 type nfs
(rw,relatime,vers=3,rsize=1048576,wsize=1048576,namlen=255,hard,proto=tcp,port=52049,timeo=600,
retrans=2,sec=sys,mountaddr=10.12.180.124,mountvers=3,mountport=55389,mountproto=udp,local_lo
ck=none,addr=10.12.180.124)
[root@centos7-c630fc ~]#
```

从上面输出来看，只能以 NFS v3 放松挂载

CEPH Dashboard NFS 配置

新建 cephfs-ns 的pool，此 pool专门用来存放一些配置文件，Dashboard管理NFS需要有些配置文件存放在Rados pool中。

```
[root@ceph-master cephcluster]# ceph osd pool create cephfs-ns 16
pool 'cephfs-ns' created
[root@ceph-master cephcluster]#

[root@ceph-master cephcluster]# rados lspools
.rgw.root
default.rgw.control
default.rgw.meta
default.rgw.log
default.rgw.buckets.index
default.rgw.buckets.data
default.rgw.buckets.non-ec
cephfs-metadata
cephfs-data
cephfs-ns                       >>>new created ceph dashboard pool
[root@ceph-master cephcluster]#
```

新建空的daemon.txt文本文件。

```
[root@ceph-master cephcluster]# touch daemon.txt
[root@ceph-master cephcluster]#
```

导入daemon文件到 cephfs-ns pool中。

```
[root@ceph-master cephcluster]# rados -p cephfs-ns put conf-ceph-master.h3cu.com daemon.txt
[root@ceph-master cephcluster]# rados -p cephfs-ns put conf-ceph-node1.h3cu.com daemon.txt
[root@ceph-master cephcluster]# rados -p cephfs-ns put conf-ceph-node2.h3cu.com daemon.txt
[root@ceph-master cephcluster]#

[root@ceph-master cephcluster]# rados -p cephfs-ns ls
conf-ceph-master.h3cu.com
conf-ceph-node1.h3cu.com
conf-ceph-node2.h3cu.com
[root@ceph-master cephcluster]#
```