# GlusterFS+Heketi部署手册

## 背景

解决NFS存储存在的单点，不支持横向扩展，不支持分布式，在程序数据访问量大时，容易引起局部数据热点问题。以及数据单副本，存储服务单节点问题。

解决kubernetes底层的分布式存储供应问题。

鉴于存储服务的特殊行和重要性，采用独立部署存储服务的方式，存储相关服务以二进制形式运行。

理想的部署方式为，K8S占用的机器与GlusterFS占用的机器分开部署。

如果GlusterFS部署在k8s集群的节点上，申请的主机需要满足以下条件

条件：

* 建议每个主机，最少拥有两块磁盘（k8s的CRI目录占一个，gluster占一个）
* 其中一块盘可以正常格式化挂载使用，另一块要保持裸盘，不做任何操作

**推荐分开部署，以下以分开部署为例。**

## 部署（独立）

前提：

* 机器须有至少1块磁盘是裸盘。不需要挂载和格式化之类的操作
* 集群节点数需>=3

部署过程：

以下操作，如无特殊说明，则需在每个节点上都进行操作

cat > /etc/yum.repos.d/gfs.repo  << EOF

# CentOS-Gluster-9.repo

#

# Please see http://wiki.centos.org/SpecialInterestGroup/Storage for more

# information

[centos-gluster9]

name=CentOS-$releasever - Gluster 9

mirrorlist=http://mirrorlist.centos.org?arch=$basearch&release=$releasever&repo=storage-gluster-9

#baseurl=http://mirror.centos.org/$contentdir/$releasever/storage/$basearch/gluster-9/

gpgcheck=0

enabled=1

gpgkey=file:///etc/pki/rpm-gpg/RPM-GPG-KEY-CentOS-SIG-Storage

[centos-gluster9-test]

name=CentOS-$releasever - Gluster 9 Testing

baseurl=http://buildlogs.centos.org/centos/$releasever/storage/$basearch/gluster-9/

gpgcheck=0

enabled=0

gpgkey=file:///etc/pki/rpm-gpg/RPM-GPG-KEY-CentOS-SIG-Storage

EOF

curl -o /etc/yum.repos.d/CentOS-Base.repo http://mirrors.aliyun.com/repo/Centos-7.repo

wget -O /etc/yum.repos.d/epel.repo http://mirrors.aliyun.com/repo/epel-7.repo

sed -i.ori 's/SELINUX=enforcing/SELINUX=disabled/' /etc/selinux/config

systemctl stop firewalld

systemctl disable firewalld

下一步根据需要为每台机器自行设置主机名

hostnamectl set-hostname  xxxxx

yum install -y centos-release-gluster  glusterfs glusterfs-server glusterfs-fuse glusterfs-rdma glusterfs-cli glusterfs-geo-replication  vim net-tools   rpcbind

systemctl start glusterd

systemctl enable glusterd

## 安装Heketi

**Heketi可单独另选一台机器安装，也可以在前提中提到的最少三台机器中某一台安装，一般在第一台**

yum install -y heketi heketi-client 或者自行下载heketi 9.0

vim /etc/heketi/heketi.json 内容如下：

{

  "\_port\_comment": "Heketi Server Port Number",

  "port": "8080",

  "\_use\_auth": "Enable JWT authorization. Please enable for deployment",

  "use\_auth": true,

  "\_jwt": "Private keys for access",

  "jwt": {

    "\_admin": "Admin has access to all APIs",

    "admin": {

      "key": "admin@123"

    },

    "\_user": "User only has access to /volumes endpoint",

    "user": {

      "key": "user@123"

    }

  },

  "\_glusterfs\_comment": "GlusterFS Configuration",

  "glusterfs": {

    "\_executor\_comment": [

      "Execute plugin. Possible choices: mock, ssh",

      "mock: This setting is used for testing and development.",

      "      It will not send commands to any node.",

      "ssh:  This setting will notify Heketi to ssh to the nodes.",

      "      It will need the values in sshexec to be configured.",

      "kubernetes: Communicate with GlusterFS containers over",

      "            Kubernetes exec api."

    ],

    "executor": "ssh",

    "\_sshexec\_comment": "SSH username and private key file information",

    "sshexec": {

      "keyfile": "/etc/heketi/heketi\_key",

      "user": "root",

      "port": "22",

      "fstab": "/etc/fstab"

    },

    "\_kubeexec\_comment": "Kubernetes configuration",

    "kubeexec": {

      "host" :"https://kubernetes.host:8443",

      "cert" : "/path/to/crt.file",

      "insecure": false,

      "user": "kubernetes username",

      "password": "password for kubernetes user",

      "namespace": "OpenShift project or Kubernetes namespace",

      "fstab": "Optional: Specify fstab file on node.  Default is /etc/fstab"

    },

    "\_db\_comment": "Database file name",

    "db": "/var/lib/heketi/heketi.db",

    "\_loglevel\_comment": [

      "Set log level. Choices are:",

      "  none, critical, error, warning, info, debug",

      "Default is warning"

    ],

    "loglevel" : "info"

  }

}

设置heketi免密

ssh-keygen -m PEM -t rsa -f /etc/heketi/heketi\_key -b 4096 -q -N ''

chown heketi:heketi /etc/heketi/heketi\_key

ssh-copy-id -i /etc/heketi/heketi\_key.pub root@node01

ssh-copy-id -i /etc/heketi/heketi\_key.pub root@node02

ssh-copy-id -i /etc/heketi/heketi\_key.pub root@node03

systemctl enable heketi

systemctl start heketi

systemctl status heketi

验证

curl <http://127.0.0.1:8080/hello>

在/etc/heketi/下配置topology.json文件，以下是一个文件包含3个节点信息的例子

vi topology.json 例子：

{

  "clusters": [

    {

      "nodes": [

        {

          "node": {

            "hostnames": {

              "manage": [

                "gfs-test01"

              ],

              "storage": [

                "10.60.2.62"

              ]

            },

            "zone": 1

          },

          "devices": [

            "/dev/sdb"

          ]

        },

        {

          "node": {

            "hostnames": {

              "manage": [

                "gfs-test02"

              ],

              "storage": [

                "10.60.2.63"

              ]

            },

            "zone": 2

          },

          "devices": [

            "/dev/sdb"

          ]

        },

        {

          "node": {

            "hostnames": {

              "manage": [

                "gfs-test03"

              ],

              "storage": [

                "10.60.2.64"

              ]

            },

            "zone": 3

          },

          "devices": [

            "/dev/sdb"

          ]

        }

      ]

    }

  ]

}

导入gluster拓扑信息

heketi-cli --server http://localhost:8080 --user admin --secret admin@123 topology load --json=/etc/heketi/topology.json

查看添加后的拓扑信息

heketi-cli --user admin --secret admin@123 topology info --server <http://localhost:8080>

## K8S节点机器安装GFS客户端支持

yum install -y glusterfs glusterfs-fuse -y

## K8S集群添加配置

echo -n "admin@123"|base64 # [这里的admin@123是heketi.json](mailto:这里的admin@123是heketi.json)中配置的密码

vim heketi-secret.yaml # 部署所需的secret

apiVersion: v1

kind: Secret

metadata:

name: heketi-secret

namespace: default

data:

# base64 encoded password. E.g.: echo -n "mypassword" | base64

key: YWRtaW5AMTIz

type: kubernetes.io/glusterfs

添加storageclass

vim gluster-heketi-storageclass.yaml

apiVersion: storage.k8s.io/v1beta1

kind: StorageClass

metadata:

name: gluster-heketi-storageclass

provisioner: kubernetes.io/glusterfs

reclaimPolicy: Delete

parameters:

resturl: "http://node01-IP:8080"

restauthenabled: "true"

restuser: "admin"

secretNamespace: "default"

secretName: "heketi-secret"

volumetype: "replicate:2" # 这里指明底层存储的副本数，建议大于2，设置为2或者3

以上两个文件分别apply

Kubectl apply -f heketi-secret.yaml

Kubectl apply -f gluster-heketi-storageclass.yaml

以下是一个使用heketi来创建PVC的例子，按举例内容，可完成日常的PVC创建与使用

vim gluster-heketi-pvc.yaml

kind: PersistentVolumeClaim

apiVersion: v1

metadata:

name: gluster-heketi-pvc

namespace: default

#annotations:

# volume.beta.kubernetes.io/storage-class: "glusterfs"

spec:

#与storageclass名字对应

storageClassName: gluster-heketi-storageclass

accessModes:

- ReadWriteMany

resources:

requests:

storage: 2Gi

其余的，即可按正常的PVC使用流程，用于服务的部署yaml文件中