Hadoop Over GlusterFilesystem调研

**1. 背景**

Hadoop Compatibility Filesystem（HCFS），即与Hadoop兼容的文件系统，这类文件系统实现了Hadoop的Filesystem接口，并且提供上层所需要API，可以替代HDFS，成为Hadoop底层的文件系统，GlusterFS便是其中之一。

实现GlusterFS替代HDFS，需要使用GlusterFS Hadoop Plugin （<https://github.com/gluster/glusterfs-hadoop>）。该插件实现了GlusterFilesystem，使Hadoop与GlusterFS可以兼容，使用GlusterFS替代底层的HDFS。

**2. 部署**

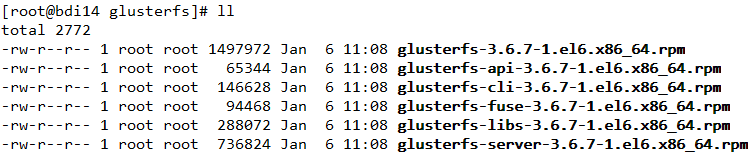
2.1 搭建GlusterFS文件系统

GlusterFS是一个分布式文件系统，所以需要至少2个节点。

我们使用两个节点部署GlusterFS server服务，分别为bdi14.cmss.com和bdi25.cmss.com

首先挂载数据盘，假设我们挂载的数据盘目录为/glusterfs

在gluster server节点，安装gluster相关的rpm包，如下图所示：



启动gluster管理守护进程：

service glusterd start

配置gluster server互信：

在bdi14.cmss.com节点执行gluster peer probe bdi25.cmss.com



在bdi25.cmss.com节点执行gluster peer probe bdi14.cmss.com



安装GlusterFS Volumn

在server1和server2节点上，创建卷目录：

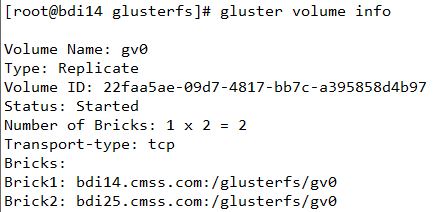
mkdir /glusterfs/gv0

在任意一个server节点上，执行：

gluster volume create gv0 replica 2 bdi14.cmss.com:/glusterfs/gv0 bdi25.cmss.com:/glusterfs/gv0

gluster volume start gv0

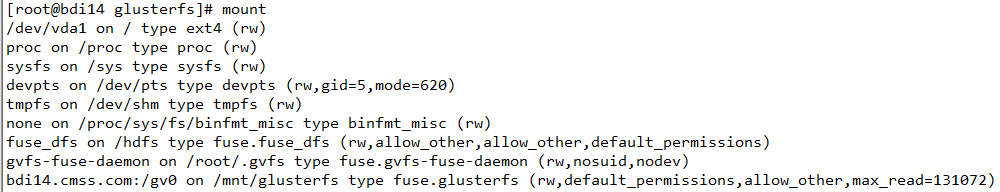
执行gluster volume info查看Vloumn情况：



挂载GlusterFS Volumn：

创建挂载点：mkdir /mnt/glusterfs

mount -t glusterfs bdi14.cmss.com:/gv0 /mnt/glusterfs



此时可以在任意客户端节点，执行上述挂载命令，将GlusterFS挂载到客户端本地。

2.2 下载编译GlusterFS Hadoop Plugin

git clone https://github.com/gluster/glusterfs-hadoop

修改pom文件中的hadoop版本号，并且需要为GlusterFs.java类添加默认构造函数，否则运行会报错。

执行mvn clean install -DskipTests进行编译。

编译完之后，会生成glusterfs-hadoop-2.3.13.jar包，将该包放入$HADOOP\_HOME/share/hadoop/common/lib目录下。

2.3 修改Hadoop相关配置

core-site.xml

<property>

<name>fs.AbstractFileSystem.glusterfs.impl</name>

<value>org.apache.hadoop.fs.local.GlusterFs</value>

</property>

<property>

<name>fs.defaultFS</name>

<value>glusterfs:///bdi14.cmss.com:1234</value>

</property>

<property>

<name>fs.glusterfs.impl</name>

<value>org.apache.hadoop.fs.glusterfs.GlusterFileSystem</value>

</property>

<property>

<name>fs.glusterfs.server</name>

<value> bdi14.cmss.com </value>

</property>

<property>

<name>fs.glusterfs.volume.fuse.gv0</name>

<value>/mnt/glusterfs </value>

</property>

<property>

<name>fs.glusterfs.volumes</name>

<value>gv0</value>

</property>

其中fs.defaultFS配置为glusterfs:///hostname:port格式，hostname为glusterfs server的hostname，port可以随机指定一个，此处的端口只是为了保证Hadoop侧解析，并没有实际的作用，如果不加端口解析会报错。

2.4 客户端挂载glusterfs

Hadoop客户端所在机器如果不是glusterfs集群内的节点，需要安装glusterfs相关的rpm包：

glusterfs-3.6.7-1.el6.x86\_64.rpm

glusterfs-api-3.6.7-1.el6.x86\_64.rpm

glusterfs-cli-3.6.7-1.el6.x86\_64.rpm

glusterfs-fuse-3.6.7-1.el6.x86\_64.rpm

glusterfs-libs-3.6.7-1.el6.x86\_64.rpm

并且执行：

mount -t glusterfs bdi14.cmss.com:/gv0 /mnt/glusterfs

2.5 为所有的hadoop用户添加sudo权限

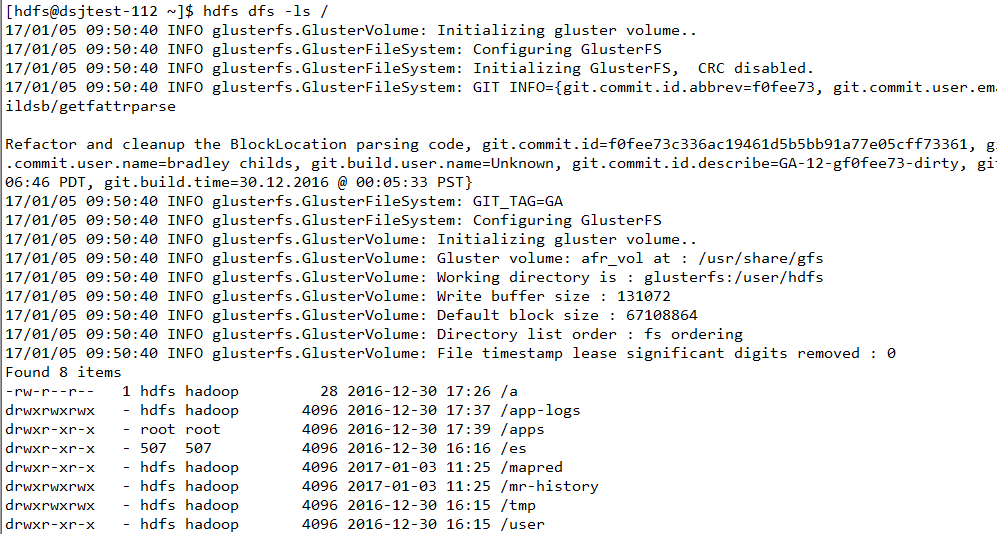
执行visudo，为每个hadoop用户加入下面一行：

hdfs ALL=NOPASSWD:/usr/bin/getfattr -m . -n trusted.glusterfs.pathinfo \*

hbase ALL=NOPASSWD:/usr/bin/getfattr -m . -n trusted.glusterfs.pathinfo \*

…

2.6 执行hadoop shell命令，查看结果是否正常。



2.7 启动hbase

虚拟机环境下，启动后Master initial要等待较长时间。



**3. 总结**

通过以上步骤，我们可以用GlusterFS代替HDFS，作为Hadoop的底层存储。实际上，当我们配置了GlusterFS后，原先HDFS的组件NN、DN等等就已经不再需要了。作为实际的存储节点，GlusterFS server相当于原先HDFS的Datanode，是文件实际落地存储的节点。