##### 8. 系统扩展维护

地域复制（geo-replication）：

# gluster volume geo-replication <MASTER> <SLAVE> start | status | stop

IO信息查看：

# gluster volume profile <VOLNAME> start | info | stop

Top监控：

Top命令允许你查看Brick的性能，例如：read,write, file open calls, file read calls, file write calls, directory opencalls, and directory real calls。所有的查看都可以设置 top数，默认100。

查看打开的 fd：

# gluster volume top <VOLNAME> open[brick <BRICK>] [list-cnt <COUNT>]

其中，open可以替换为read, write, opendir, readdir等。

查看每个 Brick 的读性能：

# gluster volume top <VOLNAME> read-perf [bs <BLOCK-SIZE> count <COUNT>] [brick <BRICK>] [list-cnt <COUNT>]

其中，read-perf可以替换为write-perf等。

##### 3、从挂载点（client）修复脑裂

提供了一组 getfattr 和 setfattr 命令来检测文件的数据和元数据裂脑状态，并从挂载点解析裂脑（如果有）。

# gluster v heal test info split-brain

Brick test-host:/test/b0/

/file100

/dir

Number of entries in split-brain: 2

Brick test-host:/test/b1/

/file100

/dir

Number of entries in split-brain: 2

Brick test-host:/test/b2/

/file99

<gfid:5399a8d1-aee9-4653-bb7f-606df02b3696>

Number of entries in split-brain: 2

Brick test-host:/test/b3/

<gfid:05c4b283-af58-48ed-999e-4d706c7b97d5>

<gfid:5399a8d1-aee9-4653-bb7f-606df02b3696>

Number of entries in split-brain: 2

使用如下命令了解处于哪种脑裂状态

getfattr -n replica.split-brain-status <path-to-file>

eg:

处于元数据脑裂

# getfattr -n replica.split-brain-status file100

file: file100

replica.split-brain-status="data-split-brain:no metadata-split-brain:yes Choices:test-client-0,test-client-1"

处于数据脑裂

# getfattr -n replica.split-brain-status file99

file: file99

replica.split-brain-status="data-split-brain:yes metadata-split-brain:yes Choices:test-client-2,test-client-3"

元数据和数据都处于脑裂中

# getfattr -n replica.split-brain-status file99

file: file99

replica.split-brain-status="data-split-brain:yes metadata-split-brain:yes Choices:test-client-2,test-client-3"

dir不在数据或者元数据脑裂下，

# getfattr -n replica.split-brain-status dir

file: dir

replica.split-brain-status="The file is not under data or metadata split-brain"

###### 从client解决数据和元数据脑裂

尝试在挂载点上对脑裂的文件进行操作（例如 cat、getfattr 等），会出现输入/输出错误。为了使用户能够分析此类文件，提供了一个 setfattr 命令。

# setfattr -n replica.split-brain-choice -v "choiceX" <path-to-file>

使用此命令，可以选择一个特定的块来访问裂脑中的文件。

eg：

1、“file1”位于数据脑裂中。尝试从文件中读取会产生输入/输出错误。

# cat file1

cat: file1: Input/output error

为file1提供的服务端选择有test-cilent-1和test-client-2

将test-client-2设置为file1的脑裂选择源，将从test-client-2的brick中读取文件

# setfattr -n replica.split-brain-choice -v test-client-2 file1

然后再cat文件

cat file1

xyz123

要撤销已设置的脑裂选择，可以使用“none”作为setfattr的扩展属性的值

# setfattr -n replica.split-brain-choice -v none file1

再cat文件，就会和以前一样，输出Input/output error

# cat file

cat: file1: Input/output error

如果想解决脑裂问题，应该设置源brick

# setfattr -n replica.split-brain-heal-finalize -v <heal-choice> <path-to-file>

eg：

# setfattr -n replica.split-brain-heal-finalize -v test-client-2 file1

上述过程可以解决所有文件的数据/元数据脑裂

###### 注意：

1、如果禁用了“fopen-keep-cache”保险丝装载选项，则每次在选择新副本之前都需要使 inode 失效。拆分大脑选择检查文件。这可以通过使用来完成：

# sefattr -n inode-invalidate -v 0 <path-to-file>

2、上面从client修复脑裂的过程不适合nfs挂载，因为nfs不提供xattrs支持