

How to install and use GlusterFS

(การติดตั้งและใช้งาน GlusterFS)

(Ubuntu 16.04 TLS)

1. บทนำ

GlusterFS คือระบบไฟล์แบบเครือข่ายที่สามารถขยายตัวได้ (scalable network filesystem) ที่ทำงานอยู่บนระบบปฏิบัติการ Linux เป็นระบบที่ถูกออกแบบและพัฒนาให้สามารถใช้งานกับฮาร์ดแวร์ปกติทั่วไปหรือที่มีใช้งานอยู่แล้ว สามารถสร้างไฟล์ที่มีขนาดใหญ่ กระจายการจัดเก็บไฟล์ไปยังหลายเครื่อง สามารถนำไปใช้กับงานในกลุ่มวิเคราะห์ข้อมูล โดย GlusterFS เป็นซอฟต์แวร์แบบเปิดเผยแพร่รหัส (open source software) เว็บไซต์ (<https://www.gluster.org>)

GlusterFS มีโครงสร้างระบบไฟล์ 2 ส่วนหลัก ประกอบด้วย

- ก้อนข้อมูล (brick) เป็นไดเรกทอรีบนระบบไฟล์จริงบนแต่ละเครื่องที่เป็นส่วนประกอบ การสร้างก้อนข้อมูลนั้นสามารถแยกพาร์ติชัน (partition) จากระบบไฟล์ปกติ หรือจะเป็นไดเรกทอรีในระบบไฟล์ที่มีอยู่เดิมก็ได้ โดยระบบไฟล์ที่เหมาะสมกับการทำก้อนข้อมูลที่แยกเป็นพาร์ติชันต่างหากนั้น คือระบบไฟล์แบบ XFS
- โวลุ่ม (volume) เป็นชุดของก้อนข้อมูลที่เป็นจุดให้บริการ หรือระบบไฟล์ที่โปรแกรมหรือระบบปฏิบัติการใช้เข้าถึงหรือกระทำไฟล์ใน GlusterFS

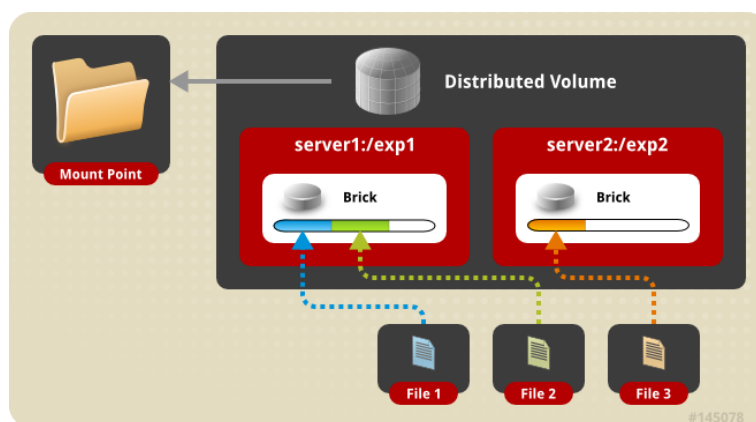
ประเภทของโวลุ่ม (Volume Types)

โวลุ่ม (volume) คือหนึ่งระบบไฟล์ที่เกิดจากชุดของก้อนข้อมูล (bricks)

เป็นส่วนที่โปรแกรมหรือระบบปฏิบัติการใช้อ้างอิงถึงไฟล์ โดย GlusterFS มีความสามารถในการรองรับการสร้างโวลุ่มได้หลายแบบ ตามความต้องการของการใช้งาน บางแบบเหมาะสำหรับงานเกี่ยวกับการทำสำเนาข้อมูล บางแบบเอกลักษณ์เฉพาะในการขยายขนาด บางแบบจะมีความสามารถเฉพาะด้านประสิทธิภาพ รวมถึงบางแบบจะมีหลายความสามารถรวมกัน แต่ละแบบมีรายการและแสดงด้วยภาพดังนี้

(<http://gluster.readthedocs.io/en/latest/Quick-Start-Guide/Architecture/>)

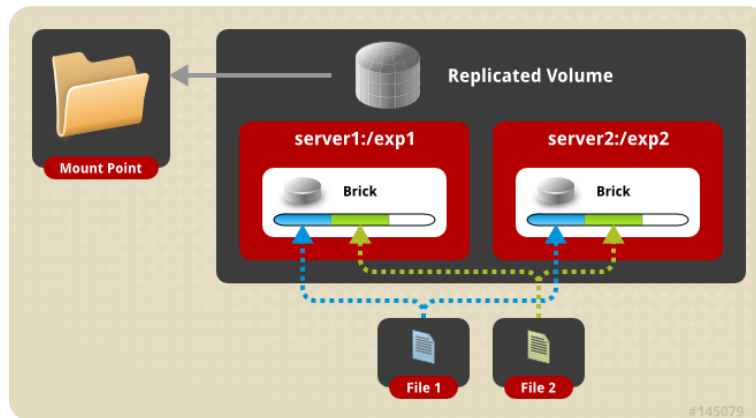
- Distributed volume



คำสั่งสำหรับการสร้างโวลุ่ม ไม่มีการกำหนดด้วยข้อกำหนด replica หรือ strip
`gluster volume create NEW-VOLNAME NEW-BRICK ...`

เช่น การสร้างโวลุ่มกระจายไปยัง 2 brick
`gluster volume create gtfs_volume server1:/exp1 server2:/exp2`

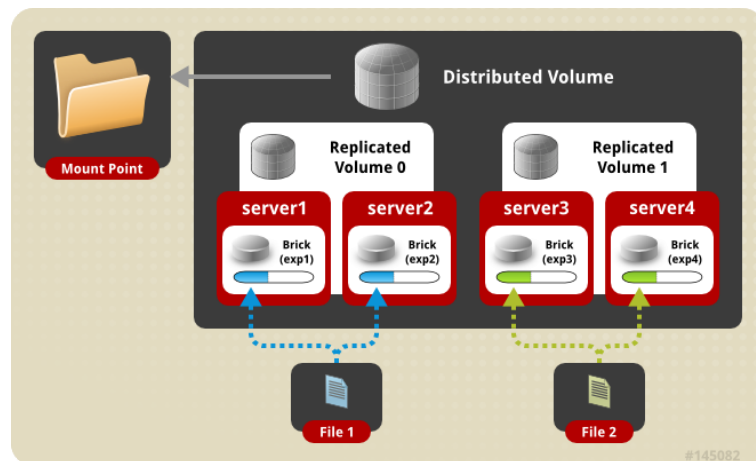
- Replicated volume



คำสั่งสำหรับการสร้างโวลุ่ม โดยการใช้ข้อกำหนด replica เท่ากับจำนวน brick
`gluster volume create NEW-VOLNAME [replica COUNT] NEW-BRICK ...`

เช่น การสร้างโวลุ่มกระจายไปยัง 2 brick
`gluster volume create gtfs_volume replica 2 server1:/exp1 \`
`server2:/exp2`

- Distributed Replicated volume

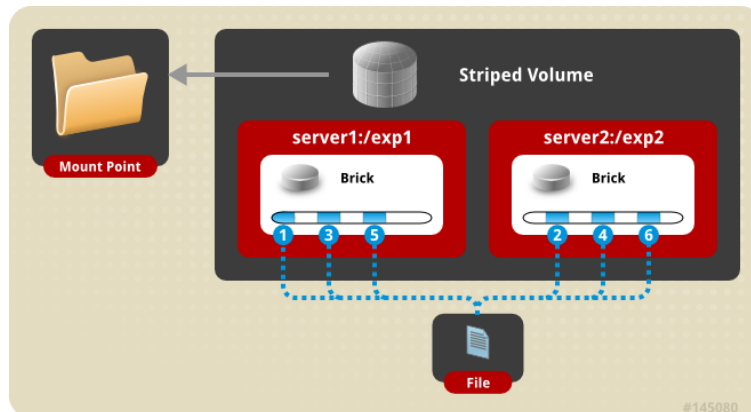


คำสั่งสำหรับการสร้างโวลุ่ม โดยการใช้ข้อกำหนด replica น้อยกว่าจำนวน brick
`gluster volume create NEW-VOLNAME [replica COUNT] NEW-BRICK ...`

เช่น การสร้างโวลุ่มกระจายไปยัง 2 brick

```
gluster volume create gtfs_volume replica 2 server1:/exp1 \
    server2:/exp2 server3:/exp3 server4:/exp4
```

- Striped volume



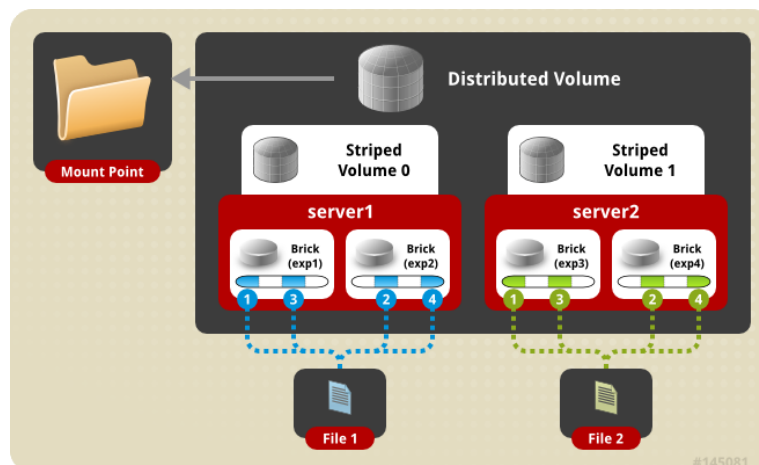
คำสั่งสำหรับการสร้างโวลุ่ม โดยการใช้ข้อกำหนด strip เท่ากับจำนวน brick

```
gluster volume create NEW-VOLNAME [strip COUNT] NEW-BRICK ...
```

เช่น การสร้างโวลุ่มกระจายไปยัง 2 brick

```
gluster volume create gtfs_volume strip 2 server1:/exp1 \
    server2:/exp2
```

- Distributed Striped volume



คำสั่งสำหรับการสร้างโวลุ่ม โดยการใช้ข้อกำหนด replica และ strip ร่วมกัน

```
gluster volume create NEW-VOLNAME [replica COUNT] [strip COUNT] \
    NEW-BRICK ...
```

เช่น การสร้างโวลุ่มกระจายไปยัง 4 brick (2x2)

```
gluster volume create gtfs_volume replica 2 strip 2 server1:/exp1 \
server2:/exp2 server3:/exp3 server4:/exp4
```

2. ฟังการเชื่อมต่อเครือข่าย

Block0	Block1	Block2
OS/Application	OS/Application	OS/Application
GlusterFS Client	GlusterFS Client	GlusterFS Client
mount:/gtfs-vms	mount:/gtfs-vms	mount:/gtfs-vms
GlusterFS Server	GlusterFS Server	GlusterFS Server
	brick:/gtfs/vms/brick	brick:/gtfs/vms/brick
enp0s3 192.168.3.200	enp0s3 192.168.3.241	enp0s3 192.168.3.242
----+-- Storage Network	-----+-----10.0.2.0/24	-----+-----
enp0s9 192.168.56.10	enp0s9 192.168.56.11	enp0s9 192.168.56.12
--+---VirtualBox Management	---+-----192.168.56.0/24	---+-----

3. เตรียมคุณสมบัติด้านเครือข่าย

Edit file `/etc/network/interfaces`

```
# ...
# VirtualBox connect to internet via NAT
auto enp0s3
iface enp0s3 inet dhcp

# VirtualBox Bridge to Host Ethernet Adapter
auto enp0s8
iface enp0s8 inet static
    address 192.168.3.241 #241 for block1, 242 for block2
    netmask 255.255.255.0
```

Restart networking

```
blockX# service networking restart
```

Edit file `/etc/hosts`

```
# ...
192.168.3.200    block0
192.168.3.241    block1
192.168.3.242    block2
```

4. การเตรียมพื้นที่และตำแหน่งไฟล์สำหรับก้อนข้อมูล (bricks)

4.1 สร้างจุดสำหรับการ mount

```
block0# mkdir -p /gtfs/vms
block1# mkdir -p /gtfs/vms
block2# mkdir -p /gtfs/vms
```

4.2.1 สร้างและฟอร์แมตพาร์ติชัน (partition)

```
block1# fdisk /dev/sdb
block1# mkfs.xfs -f -i size=512 /dev/sdb1
```

```
block2# fdisk /dev/sdb
block2# mkfs.xfs -f -i size=512 /dev/sdb1
```

4.2.2 เพิ่มรายการพาร์ติชันเข้าไปในระบบไฟล์ของทุกเครื่องเครื่องที่ไฟล์ `/etc/fstab`

Node: block1, block2

```
#...
/dev/sdb1 /gtfs/vms    xfs    defaults    0    0
```

6.2.3 ดำเนินการเมาท์พาร์ติชันใหม่ (แบบไม่ต้องปิด/เปิดเครื่อง)

```
block1# mount -arw
block2# mount -arw
```

5. การติดตั้งแพ็คเกจและทดลองคำสั่ง

5.1 ติดตั้งแพ็คเกจ

```
block0# apt install glusterfs-client glusterfs-common \
          glusterfs-server attr
block1# apt install glusterfs-client glusterfs-common \
          glusterfs-server attr
block2# apt install glusterfs-client glusterfs-common \
          glusterfs-server attr
```

5.2 ทดสอบการทำงานของโปรแกรม

```
blockX# gluster help
blockX# gluster volume info
```

6. บริหารจัดการรายการเครื่อง (nodes or peers)

6.1 Probe และเพิ่ม node เข้าไปในรายการ peers ของแต่ละ node

```
block0# gluster peer probe block1
block0# gluster peer probe block2

block1# gluster peer probe block0
block1# gluster peer probe block2

block2# gluster peer probe block0
block2# gluster peer probe block1
```

6.2 แสดงสถานะของเครื่อง

```
block0# gluster peer status
```

```
Number of Peers: 2
```

```
Hostname: block1
```

```
Uuid: 6f7bd949-1a8b-47ce-ae4e-a54686f49df3
```

```
State: Peer in Cluster (Connected)
```

```
Hostname: block2
```

```
Uuid: 3fd1c610-69b3-4ead-99f0-15ab8546c0fd
```

```
State: Peer in Cluster (Connected)
```

```
block0# gluster pool list
```

UUID	Hostname	State
11c0dbf4-3dd6-4023-832f-d3065135fd87	localhost	Connected
6f7bd949-1a8b-47ce-ae4e-a54686f49df3	block1	Connected
3fd1c610-69b3-4ead-99f0-15ab8546c0fd	block2	Connected

6.3 ปลดเครื่องออกจากรายการ

```
block0# gluster peer detach <node> force
```

7. สร้างโวลุ่มของ Glusterfs

7.1 สร้างไดเรกทอรีสำหรับเป็นก้อนข้อมูล

```
block1# mkdir -p /gtfs/vms/brick
```

```
block2# mkdir -p /gtfs/vms/brick
```

7.2 สร้างโวลุ่มแบบ Replicated type -- 3 เครื่อง

```
blockX# gluster volume create gluster_vms_volume replica 2 \
        transport tcp \
        block1:/gtfs/vms/brick \
        block2:/gtfs/vms/brick \
        force
```

7.3 เริ่มการทำงานของโวลุ่ม

```
blockX# gluster volume start gluster_vms_volume
```

7.4 แสดงรายการโวลุ่ม

```
blockX# gluster volume list
```

```
gluster_vms_volume
```

7.5 แสดงสถานะของโวลุ่ม

```
blockX# gluster volume status
```

```
Status of volume: gluster_vms_volume
```

Gluster process	TCP Port	RDMA Port	Online	Pid
Brick block0:/gtfs/vms/brick	49152	0	Y	8233
Brick block1:/gtfs/vms/brick	49152	0	Y	8110
Brick block2:/gtfs/vms/brick	49152	0	Y	8039

```
Task Status of Volume gluster_vms_volume
```

```
-----
There are no active volume tasks
```

8. การเมาท์และใช้งานโวลุ่ม

8.1 สร้างไดเรกทอรีสำหรับเป็นจุดเมาท์ของโวลุ่ม

```
block0# mkdir -p /gtfs-vms
```

```
block1# mkdir -p /gtfs-vms
```

```
block2# mkdir -p /gtfs-vms
```

8.2 ดำเนินการเมาท์โวลุ่มไปยังจุดเมาท์

```
block0# mount -t glusterfs localhost:/gluster_vms_volume /gtfs-vms
```

```
block1# mount -t glusterfs localhost:/gluster_vms_volume /gtfs-vms
```

```
block2# mount -t glusterfs localhost:/gluster_vms_volume /gtfs-vms
```

8.3 ทดสอบใช้งาน -- โดยการสร้างไฟล์จากเครื่องใดเครื่องหนึ่ง และแสดงรายการไฟล์จากเครื่องอื่น

```
block0# touch /gtfs-vms/sample.txt
```

```
block0# ls -l /gtfs-vms/
```

```
-rw-r--r-- 1 root root 0 Mar 22 23:06 sample.txt
```

```
block1# ls -l /gtfs-vms/
```

```
-rw-r--r-- 1 root root 0 Mar 22 23:06 sample.txt
```

```
block2# ls -l /gtfs-vms/
```

```
-rw-r--r-- 1 root root 0 Mar 22 23:06 sample.txt
```

8.4 กำหนดให้มีการเมาท์โวลุ่มในขณะบู๊ทเครื่อง โดยแก้ไขไฟล์ /etc/fstab

Host: block0

```
localhost:/gluster_vms_volume /gtfs-vms glusterfs \
defaults,netdev,direct-io-mode=disable,fetch-attempts=10 0 0
```

Host: block1

```
localhost:/gluster_vms_volume /gtfs-vms glusterfs \
defaults,netdev,direct-io-mode=disable,fetch-attempts=10 0 0
```

Host: block2

```
localhost:/gluster_vms_volume /gtfs-vms glusterfs \
defaults,netdev,direct-io-mode=disable,fetch-attempts=10 0 0
```

9. การจัดการโวลุ่ม (optional)

9.1 การขยายเครื่องให้กับโวลุ่ม -- ด้วยการเพิ่มเครื่องเข้าไปในโวลุ่ม

```
blockX# gluster peer probe <new-node>
```

```
blockX# gluster volume add-brick gluster_vms_volume <new-node>:<brick>
```

OR

```
blockX# gluster volume add-brick gluster_vms_volume replica <count> \
<new-node>:<brick> force
```


9.2 การลบก้อนข้อมูลออกจากโวลุ่ม

```
blockX# gluster peer probe <new-node>
```

```
blockX# gluster volume remove-brick gluster_vms_volume \  
          <new-node>:<brick> force
```

OR

```
blockX# gluster volume remove-brick gluster_vms_volume replica <count> \  
          <new-node>:<brick> force
```

```
blockX# gluster volume rebalance gluster_vms_volume start
```

9.3 หยุดการทำงานและลบโวลุ่ม

```
block0# umount /gtfs-vms
```

```
block1# umount /gtfs-vms
```

```
block2# umount /gtfs-vms
```

```
blockX# gluster volume stop gluster_vms_volume
```

```
blockX# gluster volume delete gluster_vms_volume
```

10. การตั้งค่าอื่น (optional)

การกำหนดให้ยอมรับการเชื่อมต่อหรือสื่อสารกับตัวบริการโวลุ่มแบบไม่ปลอดภัย

```
blockX# gluster volume set gluster_vms_volume server.allow-insecure on
```