Exploring Storage Options for Docker Containers

@abrongersma

Sr. Infrastructure Engineer @ Modulus.io

LXC, Docker, AWS, Joyent, Digital Ocean, VMware

MODULUS

A PROGRESS COMPANY

Docker storage is hard

Default Docker storage options

(Batteries Included)

Docker storage drivers

- AUFS
- Device Mapper
- BTRFS
- Overlay
- LVM
- ZFS (New)

Data Volumes

Data volume containers

```
docker create -v /webdata --name webdata ubuntu:14.04 /bin/true
```

Host folders

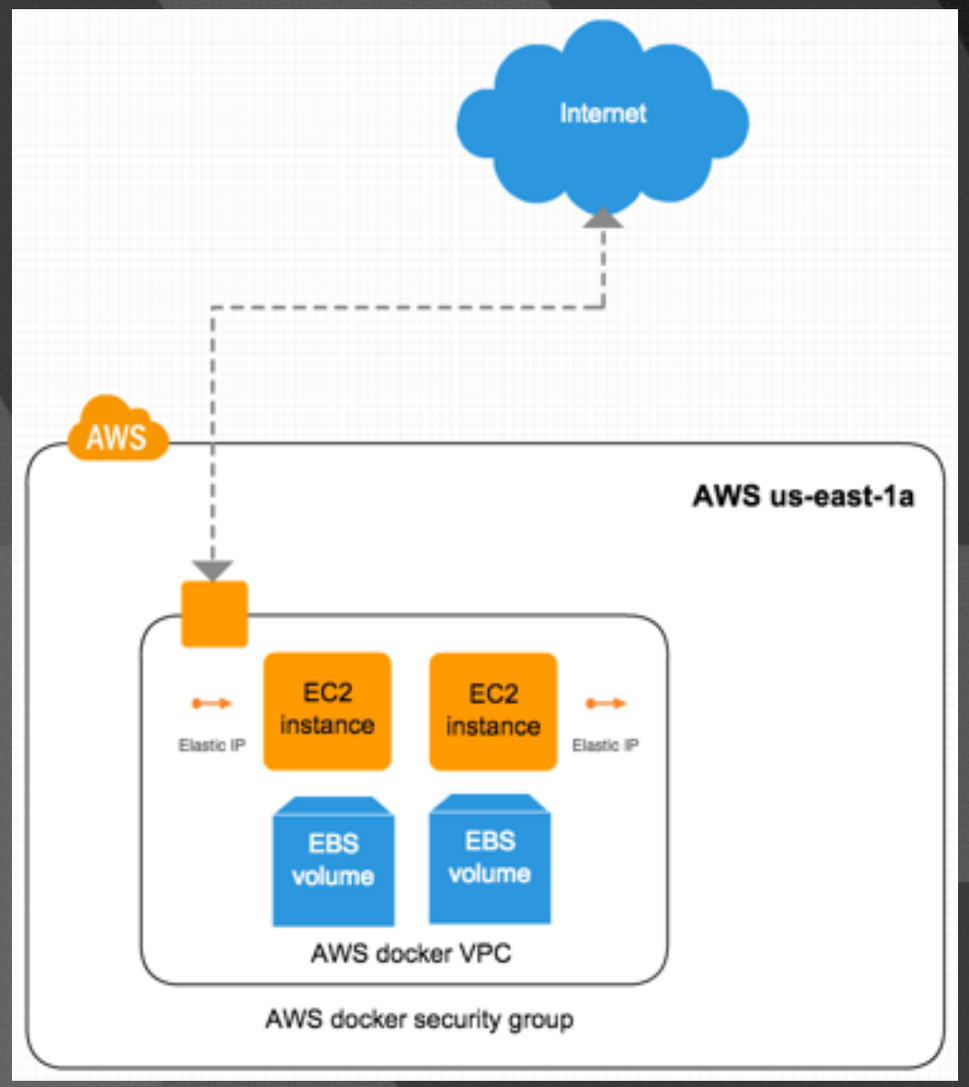
```
docker run -it -v /webdata:/webdata ubuntu:14.04 /bin/bash
```

Contianer runtime storage (AKA mount when the container starts)

- Requires `--privileged` (some excptions)
- S3fs, FUSE, BitTorrent Sync, GlusterFS, CIFS, NFS
- Great for one container, becomes a hassle with lots of little containers
- Privileged mode is scary

The setup

2x AWS C3XL
4 CPU
7.5 GB Memory
2 x 40 SSD
1 x 40 EBS



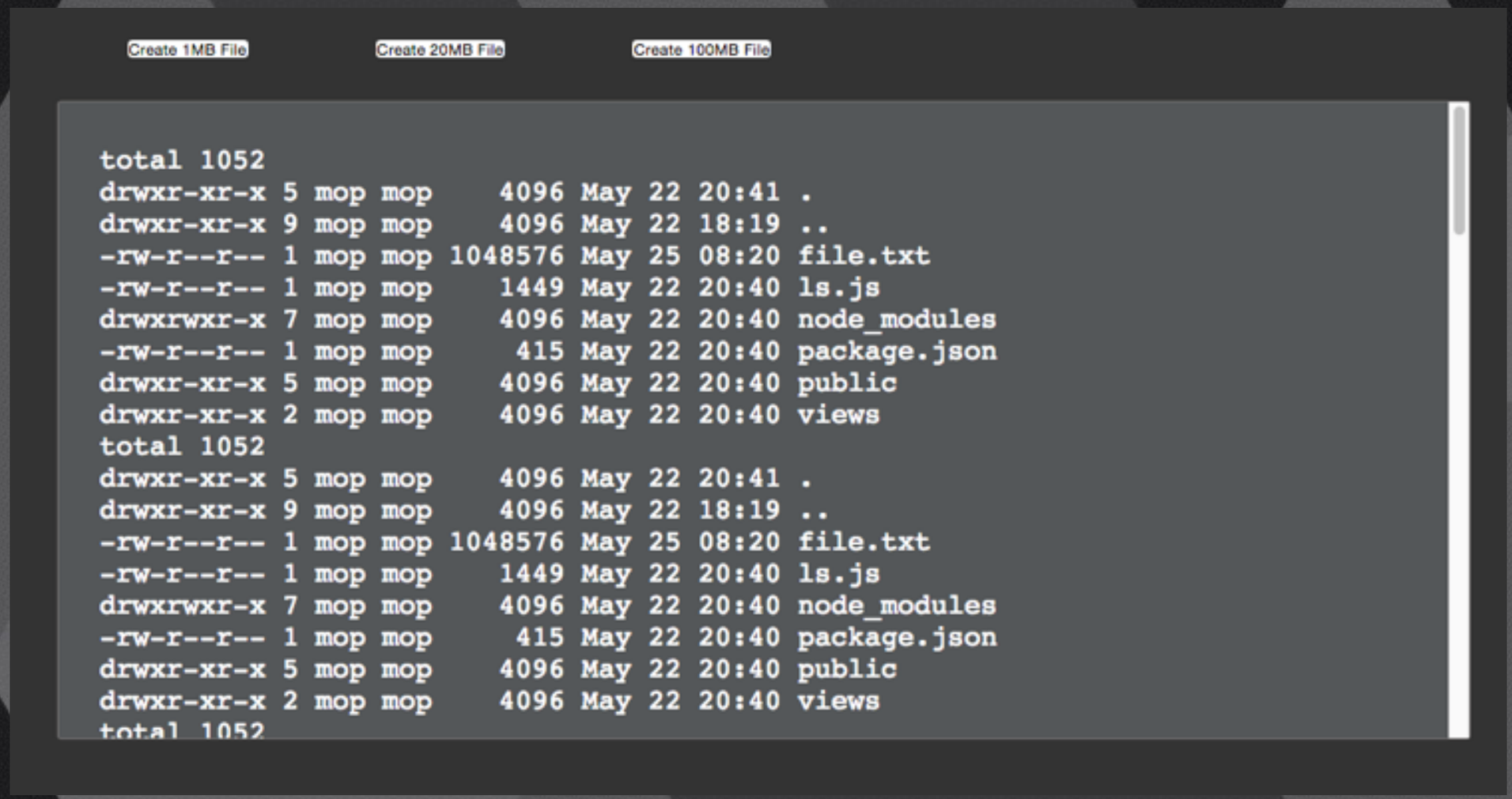
AWS VPC
2376, 80, 21
us-east-1a

Ubuntu 14.04
Docker 1.6.2

Requirements

- Horizontally scalable app
- 2 way replication
- Present a mount point that can be passed to docker with docker run -v
- Very fast eventual consistency

Demo App


Create 1MB File
Create 20MB File
Create 100MB File
total 1052
drwxr-xr-x 5 mop mop 4096 May 22 20:41 .
drwxr-xr-x 9 mop mop 4096 May 22 18:19 ..
-rw-r--r-- 1 mop mop 1048576 May 25 08:20 file.txt
-rw-r--r-- 1 mop mop 1449 May 22 20:40 ls.js
drwxrwxr-x 7 mop mop 4096 May 22 20:40 node_modules
-rw-r--r-- 1 mop mop 415 May 22 20:40 package.json
drwxr-xr-x 5 mop mop 4096 May 22 20:40 public
drwxr-xr-x 2 mop mop 4096 May 22 20:40 views
total 1052
drwxr-xr-x 5 mop mop 4096 May 22 20:41 .
drwxr-xr-x 9 mop mop 4096 May 22 18:19 ..
-rw-r--r-- 1 mop mop 1048576 May 25 08:20 file.txt
-rw-r--r-- 1 mop mop 1449 May 22 20:40 ls.js
drwxrwxr-x 7 mop mop 4096 May 22 20:40 node_modules
-rw-r--r-- 1 mop mop 415 May 22 20:40 package.json
drwxr-xr-x 5 mop mop 4096 May 22 20:40 public
drwxr-xr-x 2 mop mop 4096 May 22 20:40 views
total 1052

Rsync

https://rsync.samba.org/

- Rsync runs once
- Need a tool to trigger it on filesystem events
- https://github.com/axkibe/lsyncd/
- Storage user account
- Generate ssh key pair

```
1  useradd -m -U storage
2  mkdir /mnt/rsync
3  chown storage:storage /mnt/rsync/
4  ssh-keygen -t rsa
5  cat ~/.ssh/id_rsa.pub >>  ~/.ssh/authorized_keys
6  apt-get install lsyncd
7  lsyncd -rsyncssh /mnt/rsync storage@host2 /mnt/rsync
```

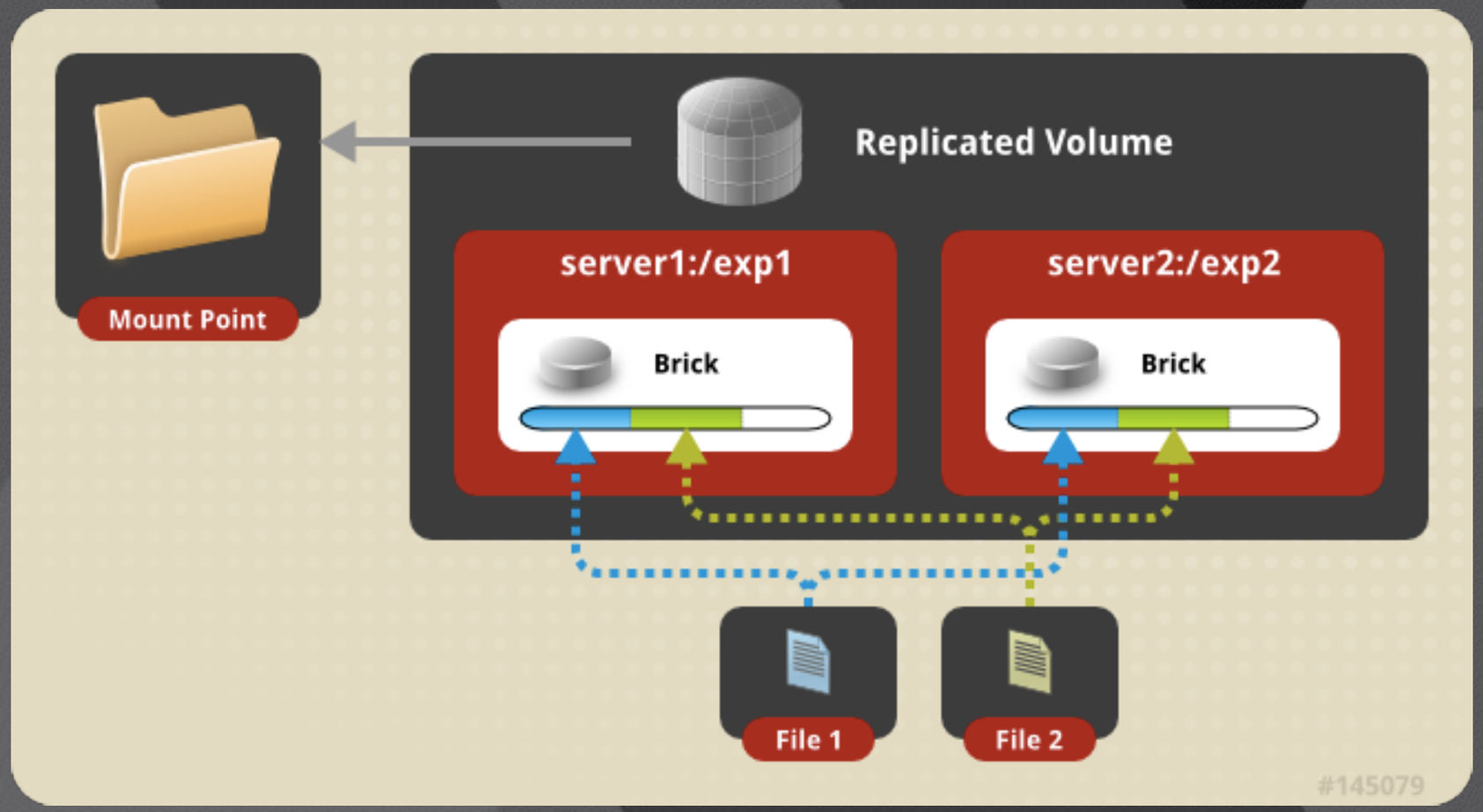
Demo

- 2 party replication is clunky, what about more hosts?
- Large number of lsyncd processes running
- Need a tool to mange scripts, docker pre-run script, post-run script
- Need to know what projects are running on which host
- Single storage user account, hard to cycle keys

GlusterFS

http://www.gluster.org/

- Distributed file system
- Commodity hardware
- NFS or GlusterFS native client
- FUSE(File System in Userspace)


Mount Point
Replicated Volume
server1:/exp1
server2:/exp2
Brick
Brick
File 1
File 2
#145079

```
apt-get install python-software-properties xfsprogs attr
add-apt-repository ppa:gluster/glusterfs-3.5
apt-get update
apt-get install glusterfs-server
gluster peer probe host1
gluster peer probe host2
mkdir -p /export/xvdf1 && mount /dev/xvdf1 /export/xvdf1 && mkdir -p /export/xvdf1/brick
echo "/dev/xvdf1 /export/xvdf1 xfs defaults 0 0"  >> /etc/fstab
gluster volume create gv0 replica 2 host1:/export/xvdf1/brick host2:/export/xvdf1/brick
gluster volume start gv0
```

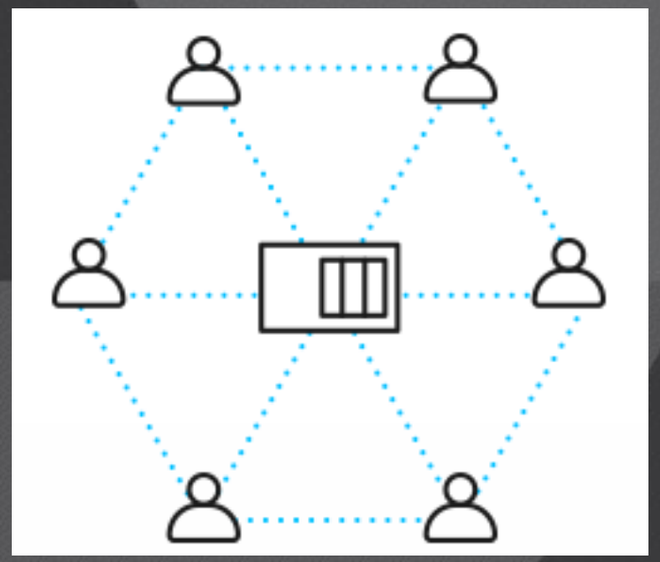
Demo

- Want to have more that 2 hosts
- Not economical to have every docker host as a storage server
- Requires lots of open ports
- Management overhead

BitTorrent Sync

https://www.getsync.com/

- BitTorrent to share files
- Not Open Source, Pro Mode
- Direct file transfers
- Web Gui



```
1  add-apt-repository ppa:tuxpoldo/btsync
2  apt-get update
3  apt-get install btsync
4  sudo dpkg-reconfigure btsync
5  mkdir /storage/btsync
6  chown root:btsync /storage/btsync
7  chmod 2775 /storage/btsync
8  usermod -a -G btsync storage
9  btsync --dump-sample-config
10 btsync --config btsync.config
```

Demo

- The more the merrier
- Bandwidth consumption
- Complex headless install
- Licensing issues?

Ceph

http://ceph.com/

- Object storage (think s3)
- Block storage
- Filestorage (Ceph FS)
- Distributed
- Inside and outside of containers
- https://github.com/ceph/ceph-docker

```
https://ceph.com/git/?p=ceph.git;a=blob_plain;f=keys/release.asc' | sudo apt-key add -
echo deb http://ceph.com/debian-{ceph-stable-release}/ $(lsb_release -sc) main | sudo tee
/etc/apt/sources.list.d/ceph.list
sudo apt-get update && sudo apt-get install ceph-deploy ntp
# Add to sudoers
echo "storage ALL = (root) NOPASSWD:ALL" | sudo tee /etc/sudoers.d/storage
sudo chmod 0440 /etc/sudoers.d/storage
# Open ports
6789 for Ceph Monitors and ports 6800:7300
```

```
ceph-deploy new host1 host2
osd pool default size = 2
public network = {ip-address}/{netmask}
ceph-deploy install admin-node host1 host2
ceph-deploy mon create-initial
ceph health
ssh host1
sudo mkdir /var/local/osd0
exit
ssh host2
sudo mkdir /var/local/osd1
exit
ceph-deploy osd prepare host1:/var/local/osd0 host2:/var/local/osd1
ceph-deploy admin admin-node host1 host2
sudo chmod +r /etc/ceph/ceph.client.admin.keyring
ceph health
```

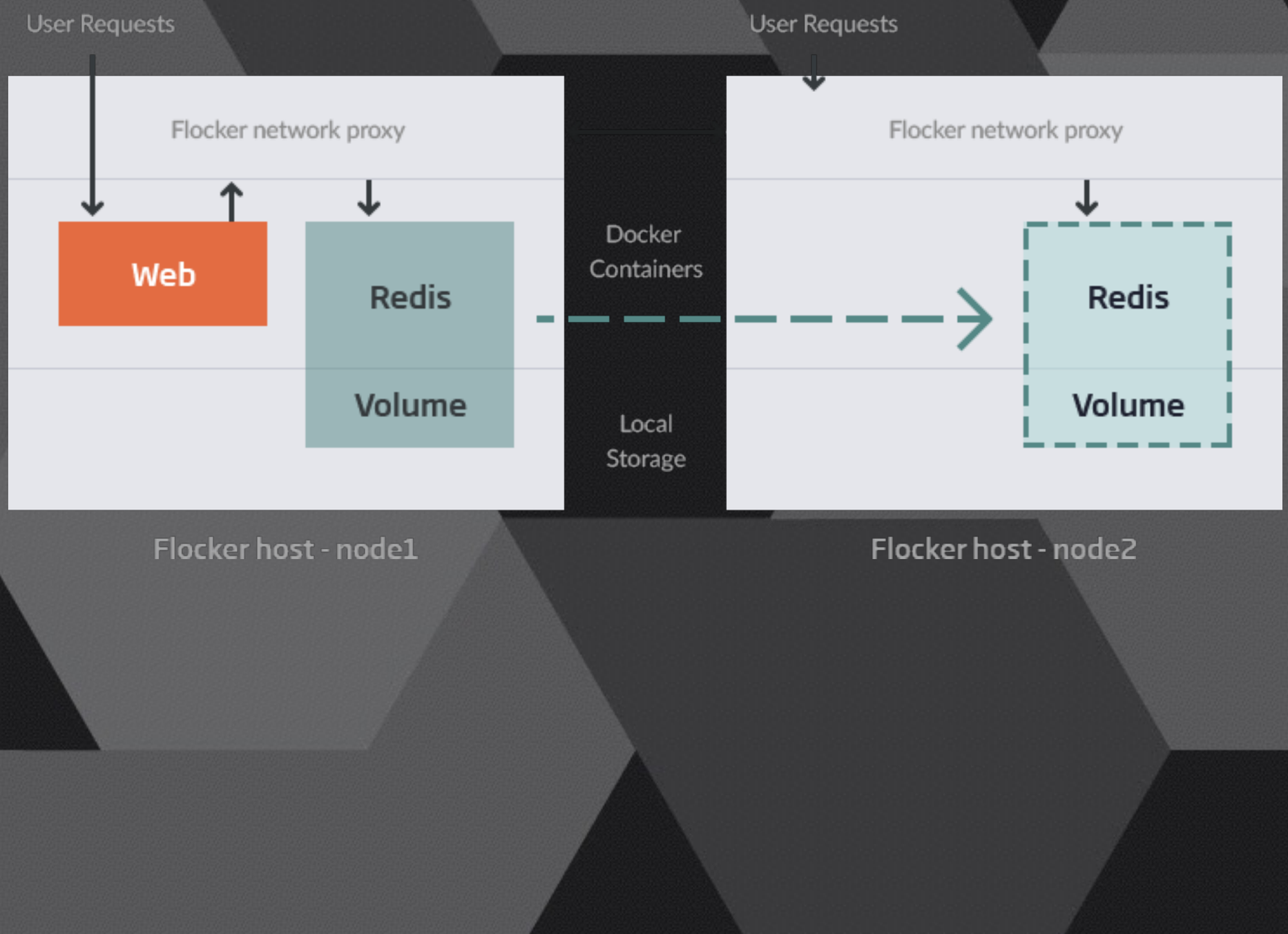
Demo

- Most complex setup of the bunch
- Better if run in a standalone setup
- Reliable, but requires quite a bit of experience to tune

Cluster HQ

http://clusterhq.com

- Flocker
- Client, Node
- Compose / Fig
- Works with local storage


User Requests
Flocker network proxy
Web
Redis
Volume
Docker
Containers
Local
Storage
User Requests
Flocker network proxy
Redis
Volume
Flocker host - node1
Flocker host - node2

Demo


I NOW HAVE
ADDITIONAL QUESTIONS