

# Virtualization & Cloud Computing

## Progetto finale

Terracciano Vincenzo

Usai Giacomo

# INTRODUZIONE

## -OBIETTIVO

Costruire un'architettura distribuita costituita da tre macchine virtuali operanti su VMWare. Esse saranno inserite in un Docker Swarm e condivideranno un file system grazie alla tecnologia di GlusterFS. Lo Swarm ospiterà diversi container, tra cui:

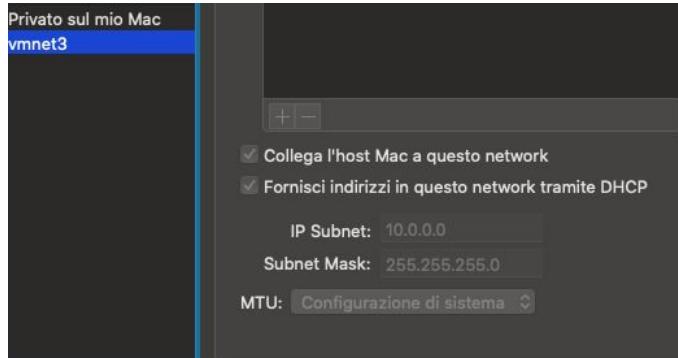


# Cos'è Docker?

Docker fornisce una modalità standard per eseguire il tuo codice. Si tratta di un sistema operativo per container. Così come la macchina virtuale virtualizza i server hardware (ovvero elimina la necessità di gestirli direttamente), i container virtualizzano il sistema operativo di un server. Docker è installato su ogni server e fornisce semplici comandi con cui creare, avviare o interrompere i container.

# Configurazione schede di rete

VMnet3



VMnet8

(di default su MacOS)

```
vmnet3: flags=8863<UP,BROADCAST,SMART,RUNNING,SIMPLEX,MULTICAST> mtu 1500
        ether 00:50:56:c0:00:03
        inet 10.0.0.1 netmask 0xffffffff broadcast 10.0.0.255
vmnet8: flags=8863<UP,BROADCAST,SMART,RUNNING,SIMPLEX,MULTICAST> mtu 1500
        ether 00:50:56:c0:00:08
        inet 192.168.75.1 netmask 0xffffffff broadcast 192.168.75.255
```

# Configurazione VM

Mostra tutte Vm1: Processori e memoria Aggiungi dispositivo...

**Processori**  
1 core processore

**Memoria**  
2048 MB (consigliato)

Mostra tutte Vm1: Impostazioni Aggiungi dispositivo...

**Impostazioni sistema**

- Opzioni avanzate
- Generale
- Congiungere
- Tastiera e Mouse
- Processori e memoria
- Monitor

**Dispositivi rimovibili**

- Adattatore network
- Adattatore network 2
- Disco rigido (SCSI)
- Disco rigido 2 (SCSI)
- CD/DVD (SATA)
- Scheda audio
- USB e Bluetooth

**Stampante**

**Altro**

- Disco di avvio
- Criptatura e Restrizioni
- Compatibilità
- Isolamento
- Avanzate

Mostra tutte Vm1: Disco rigido 2 (SCSI) Aggiungi dispositivo...

Nome file: Disco virtuale-cl3.vmdk Dimensioni disco: 5,00 GB

Opzioni avanzate

Tipo di bus: SCSI

- Preallocare spazio su disco
- Dividi in più file

Ripristina Applica

Mostra tutte Vm1: Disco rigido (SCSI) Aggiungi dispositivo...

Nome file: Disco virtuale-cl2.vmdk Dimensioni disco: 20,00 GB

Opzioni avanzate

Tipo di bus: SCSI

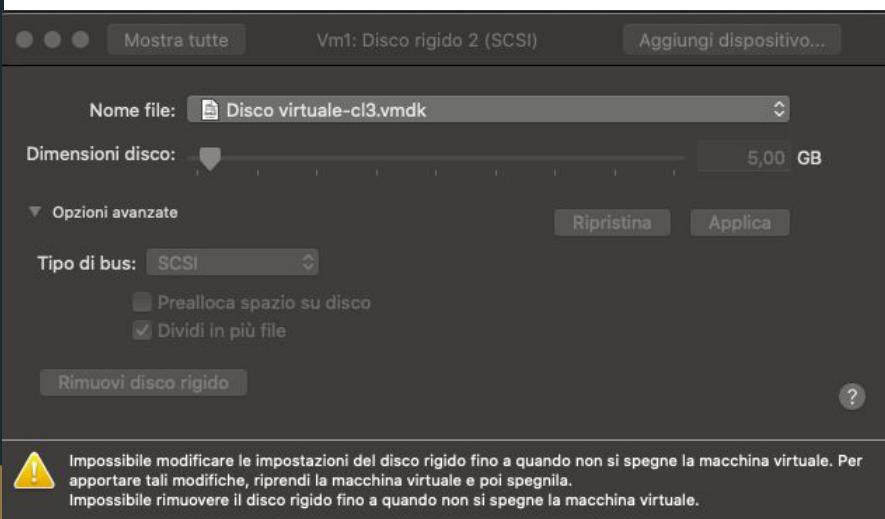
- Preallocare spazio su disco
- Dividi in più file

Rimuovi disco rigido

?

Impossibile modificare le impostazioni del disco rigido fino a quando non si spegne la macchina virtuale. Per apportare tali modifiche, riprendi la macchina virtuale e poi spegnila.  
Impossibile rimuovere il disco rigido fino a quando non si spegne la macchina virtuale.

# Hard Disk



## Formattazione

```
enz93@vm1:~$ sudo fsdik -l
[sudo] password di enz93:
sudo: fsdik: comando non trovato
enz93@vm1:~$ sudo fdisk -l
Disk /dev/sda: 20 GiB, 21474836480 bytes, 41943040 sectors
Units: sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disklabel type: dos
Disk identifier: 0xff710148

Dispositivo Avvio Start Fine Settori Size Id Tipo
/dev/sda1 * 2048 39942143 39940096 19G 83 Linux
/dev/sda2 39944190 41940991 1996802 975M 5 Esteso
/dev/sda5 39944192 41940991 1996800 975M 82 Linux swap / Solaris

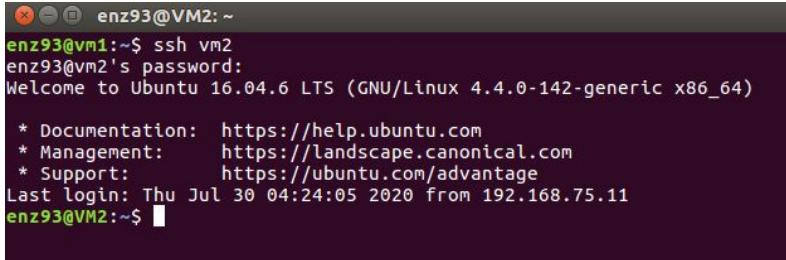
Disk /dev/sdb: 5 GiB, 5368709120 bytes, 10485760 sectors
Units: sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes

enz93@vm1:~$ sudo mkfs.ext4 /dev/sdb
```

# Setup VM

## SSH

```
sudo apt -y install openssh-server
```



```
enz93@VM2:~$ ssh vm2
enz93@vm2's password:
Welcome to Ubuntu 16.04.6 LTS (GNU/Linux 4.4.0-142-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/advantage
Last login: Thu Jul 30 04:24:05 2020 from 192.168.75.11
enz93@VM2:~$
```

## DOCKER

### Installing Docker (Ubuntu)

- add repository (`sudo curl -fsSL https://download.docker.com/linux/ubuntu/gpg | sudo apt-key add -; sudo add-apt-repository "deb [arch=amd64] https://download.docker.com/linux/ubuntu $(lsb_release -cs) stable"`)
- install Docker engine (`sudo apt-get update; sudo apt-get install docker-ce docker-ce-cli containerd.io`)
- add user to docker group (`sudo usermod -aG docker [user]; sudo systemctl restart docker`)

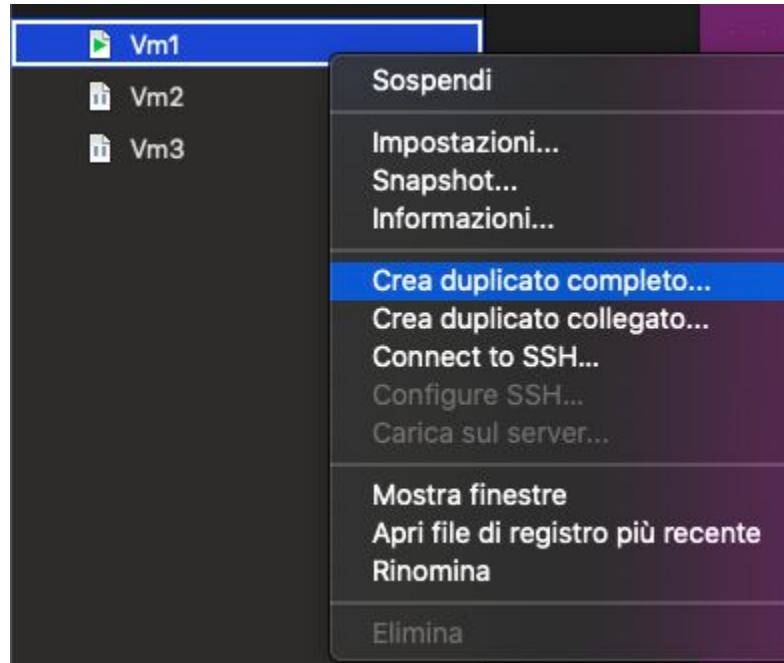
\*[user] == enz93

## DOCKER-COMPOSE

```
sudo curl -L "https://github.com/docker/compose/releases/download/1.24.1/docker-compose-$(uname -s)-$(uname -m)" \
-o /usr/local/bin/docker-compose

sudo chmod +x /usr/local/bin/docker-compose
```

# Clonazione



# Setup Network e Gluster Network

## Settare la rete di VM1

```
GNU nano 2.5.3          File: /etc/network/interfaces
# interfaces(5) file used by ifup(8) and ifdown(8)
auto lo
iface lo inet loopback

auto ens33
iface ens33 inet dhcp
iface ens33 inet static
    address 192.168.75.11
    netmask 255.255.255.0
    network 192.168.75.0
    gateway 192.168.75.2
    dns-nameservers 8.8.8.8

auto ens38
iface ens38 inet static
    address 10.0.0.10
    netmask 255.255.255.0
```

Per VM2 sostituire 192.168.75.12

Per VM3 sostituire 192.168.75.13

Per vm2g sostituire 10.0.0.20

Per vm3g sostituire 10.0.0.30

## Settaggio nome macchina

```
GNU nano 2.5.3          File: /etc/hostname
vm1
```

## Assegnare un nome all'indirizzo IP

```
GNU nano 2.5.3          File: /etc/hosts
127.0.0.1      localhost
127.0.1.1      ubuntu
127.0.1.1      vm1

192.168.75.11   vm1
192.168.75.12   vm2
192.168.75.13   vm3

10.0.0.10  vm1g
10.0.0.20  vm2g
10.0.0.30  vm3g
```

# Test di verifica

Reboot

Testando connettività con indirizzo IP e nome

```
enz93@vm1:~$ ping 192.168.75.12
PING 192.168.75.12 (192.168.75.12) 56(84) bytes of data.
64 bytes from 192.168.75.12: icmp_seq=1 ttl=64 time=0.565 ms
64 bytes from 192.168.75.12: icmp_seq=2 ttl=64 time=0.661 ms
^C
--- 192.168.75.12 ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, time 1015ms
rtt min/avg/max/mdev = 0.565/0.613/0.661/0.048 ms
enz93@vm1:~$ ping vm2
PING vm2 (192.168.75.12) 56(84) bytes of data.
64 bytes from vm2 (192.168.75.12): icmp_seq=1 ttl=64 time=0.846 ms
64 bytes from vm2 (192.168.75.12): icmp_seq=2 ttl=64 time=0.647 ms
64 bytes from vm2 (192.168.75.12): icmp_seq=3 ttl=64 time=0.795 ms
^C
--- vm2 ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2009ms
rtt min/avg/max/mdev = 0.647/0.762/0.846/0.090 ms
enz93@vm1:~$ ping vm2g
PING vm2g (10.0.0.20) 56(84) bytes of data.
64 bytes from vm2g (10.0.0.20): icmp_seq=1 ttl=64 time=0.695 ms
64 bytes from vm2g (10.0.0.20): icmp_seq=2 ttl=64 time=0.726 ms
^C
--- vm2g ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, time 1005ms
rtt min/avg/max/mdev = 0.695/0.710/0.726/0.030 ms
enz93@vm1:~$
```

# Installazione GlusterFS

- install packages: `sudo apt update; sudo apt install glusterfs-server glusterfs-client`
- start server service: `sudo service glusterfs-server start`
- each server needs to know about the others. Peer with other nodes (example from KVM1):

```
sudo gluster peer probe kvm2g
sudo gluster peer probe vmstorageg
(peer probe: success)
```

- check status: `sudo gluster peer status`

```
enz93@vm1:~$ sudo gluster peer status
Number of Peers: 2

Hostname: vm3g
Uuid: cd964b10-a056-40e1-9133-1b59c93a007f
State: Peer in Cluster (Connected)

Hostname: vm2g
Uuid: 46f92cc2-af48-49d5-8031-a99b31374199
State: Peer in Cluster (Connected)
```

## Settare le cartelle dove risiede il brick

- setup the GlusterFS directories where the bricks will reside
  - `mkdir -p /gluster/bricks/1 (kvm1)`
  - `mkdir -p /gluster/bricks/2 (kvm2)`
  - `mkdir -p /gluster/bricks/3 (vmstorage)`
- and add an entry to `/etc/fstab` on each VM (brick gets mounted when the operating system boots)
  - `echo '/dev/sdb /gluster/bricks/1 ext4 defaults 0 0' >> /etc/fstab (kvm1)`
  - `echo '/dev/sdb /gluster/bricks/2 ext4 defaults 0 0' >> /etc/fstab (kvm2)`
  - `echo '/dev/sde /gluster/bricks/3 ext4 defaults 0 0' >> /etc/fstab (vmstorage)`
- mount filesystems: `mount -a`

# I volumi replicati di GlusterFS

## Creare un nuovo volume gfs

- create a new gfs volume: sudo gluster volume create GlusterFS \

```
replica 3 \
kvm1g:/gluster/bricks/1/brick \
kvm2g:/gluster/bricks/2/brick \
vmstorageg:/gluster/bricks/3/brick
```

- start volume: sudo gluster volume start GlusterFS
- show info/status: sudo gluster volume info/status GlusterFS

```
enz93@vm1:~$ sudo gluster volume status GlusterFs
Status of volume: GlusterFs
Gluster process          TCP Port  RDMA Port  Online  Pid
-----
Brick vm1g:/gluster/bricks/1/brick    49152    0        Y      4605
Brick vm2g:/gluster/bricks/2/brick    49152    0        Y      1143
Brick vm3g:/gluster/bricks/3/brick    49152    0        Y      1162
NFS Server on localhost              N/A       N/A       N      N/A
Self-heal Daemon on localhost        N/A       N/A       Y      4603
NFS Server on vm3g                 N/A       N/A       N      N/A
Self-heal Daemon on vm3g            N/A       N/A       Y      1152
NFS Server on vm2g                 N/A       N/A       N      N/A
Self-heal Daemon on vm2g            N/A       N/A       Y      1135

Task Status of Volume GlusterFs
-----
There are no active volume tasks
```

```
Volume Name: GlusterFs
Type: Replicate
Volume ID: e2215510-ff6c-4eba-9683-86cc2e7d463c
Status: Started
Number of Bricks: 1 x 3 = 3
Transport-type: tcp
Bricks:
Brick1: vm1g:/gluster/bricks/1/brick
Brick2: vm2g:/gluster/bricks/2/brick
Brick3: vm3g:/gluster/bricks/3/brick
Options Reconfigured:
performance.readdir-ahead: on
auth.allow: 10.0.0.30,10.0.0.10,10.0.0.20
nfs.disable: Off
```

# Montaggio del volume

- on each node, create a mountpoint: `sudo mkdir /GlusterFs`
- add a fstab entry: `echo 'localhost:/GlusterFs /GlusterFs glusterfs defaults,_netdev,backupvolfile-server=localhost 0 0' >> /etc/fstab`
- mount volume: `mount -a`

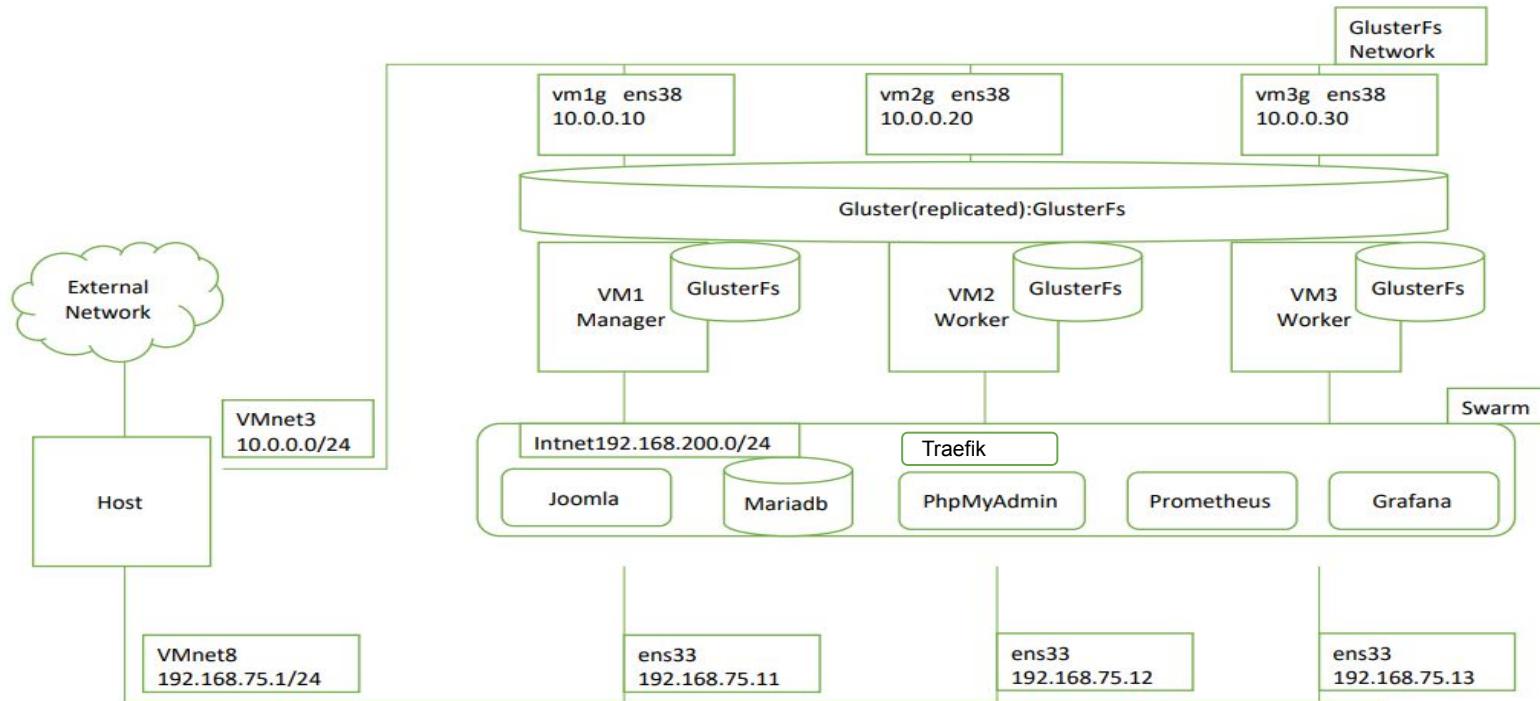
Creare le cartelle dei servizi

```
mkdir /GlusterFs/joomla  
mkdir /GlusterFs/mysql  
mkdir /GlusterFs/phpmyadmin  
mkdir /GlusterFs/grafana  
chmod -R a=rwx /GlusterFs/grafana
```

## Test

- show space of mounted volumes: `df -h`
- The total space available on the volume will be 5 Gb even though we have 3 disks of 5 Gb each connected to the GlusterFS
- create file (KVM1): `echo "Hello world" > /GlusterFs/hello.txt`
- test replication (KVM2, vmstorage): `cat /GlusterFs/hello.txt`

# Soluzione



# Docker Build

Build delle immagini di joomla e phpmyadmin create da noi e da installare in ogni vm

```
cd /GlusterFs/joomla_img
```

```
docker build . -t joomla_img
```

```
cd ..
```

```
cd pmy_img
```

```
docker build . -t pmy_img
```

# Docker Stack

```
enz93@vm1:~/Desktop/3172020$ docker stack deploy service -c docker-compose.yml
Creating network service_intnet
Creating network service_extnet
Creating service service_pmy
Creating service service_db
Creating service service_prometheus
Creating service service_node-exporter
Creating service service_grafana
Creating service service_cadvisor
Creating service service_traefik
Creating service service_joomla
enz93@vm1:~/Desktop/3172020$ docker stack services service
ID          NAME      MODE      REPLICAS  IMAGE
48yskj2z04g0  service_prometheus  replicated  1/1      prom/prometheus:latest
k0siykme7t5w  service_joomla     replicated  1/1      joomla_img:latest
kswqmcpaaty3  service_pmy        replicated  1/1      pmy_img:latest
lwtj99l7qom0   service_node-exporter  global    3/3      prom/node-exporter:latest
t62yzyi8ti8n  service_cadvisor   global    3/3      google/cadvisor:latest
wjuapkcbx748  service_grafana    replicated  1/1      grafana/grafana:latest
yuai6kw9yxd9  service_db         replicated  1/1      mariadb:latest
zz98sx1lp18w  service_traefik   replicated  1/1      traefik:v2.0.0-rc3
PORTS
*:9090->9090/tcp
*:8081->80/tcp
*:8082->80/tcp
*:9100->9100/tcp
*:8084->8080/tcp
*:3000->3000/tcp
*:3306->3306/tcp
*:80->80/tcp, *:8080->8080/tcp
```

# Joomla!

## Home

You are here: Home

### Main Menu

Home

### Login Form

Username

Password

Remember Me

Log in

[Forgot your username?](#)

[Forgot your password?](#)

managedb.joomla.local

Apple iCloud Yahoo Google Wikipedia Facebook Twitter LinkedIn Il Meteo Yelp TripAdvisor Bing Google

Home managedb.joomla.local / db / joomla / prfx\_users | phpMyAdmin 4.9.1

**phpMyAdmin**

Recente Preferiti

information\_schema  
joomla

Inserisci stringa di ricerca, oppure E X

Nuova

prfx\_assets  
prfx\_associations  
prfx\_banners  
prfx\_banner\_clients  
prfx\_banner\_tracks  
prfx\_categories  
prfx\_contact\_details  
prfx\_content  
prfx\_contentitem\_tag\_map  
prfx\_content\_frontpage  
prfx\_content\_rating  
prfx\_content\_types  
prfx\_core\_log\_searches  
prfx\_extensions  
prfx\_finder\_filters  
prfx\_finder\_links  
prfx\_finder\_links\_terms0  
prfx\_finder\_links\_terms1  
prfx\_finder\_links\_terms2  
prfx\_finder\_links\_terms3  
prfx\_finder\_links\_terms4  
prfx\_finder\_links\_terms5  
prfx\_finder\_links\_terms6  
prfx\_finder\_links\_terms7  
prfx\_finder\_links\_terms8

Server: db » Database: joomla » Tabella: prfx\_users

Mostra Struttura SQL Cerca Inserisci Esporta Importa Operazioni Trigger

✓ Mostrò le righe 0 - 0 (1 del totale, La query ha impiegato 0.0010 secondi.)

SELECT \* FROM `prfx\_users`

Profiling [Modifica inline] [ Modifica ] [ Spiega SQL ] [ Crea il codice PHP ] [ Aggiorna]

Mostra tutti | Numero di righe: 25 Filtra righe: Cerca nella tabella

+ Opzioni

	id	name	username	email	password	block	sendEmail	registerDate	lastvisitDate	activ
<input type="checkbox"/>	21	Me	asd	asd@localhost.com	81dc9bdb52d04dc20036dbd8313ed055	0	0	0000-00-00 00:00:00	0000-00-00 00:00:00	

Selezione tutto Se selezionati:  Modifica  Copia  Elimina  Esporta

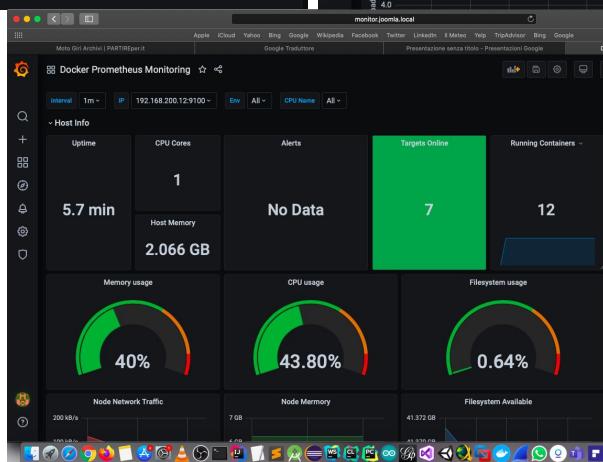
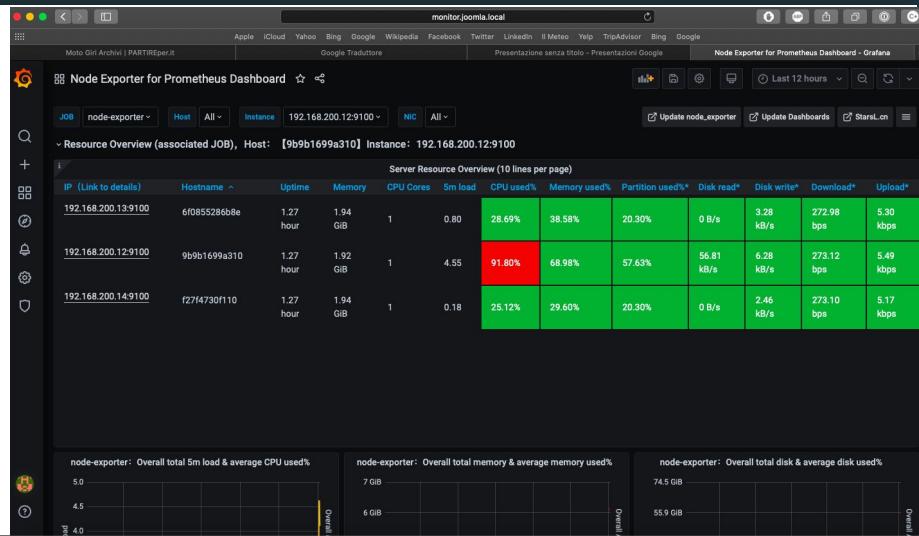
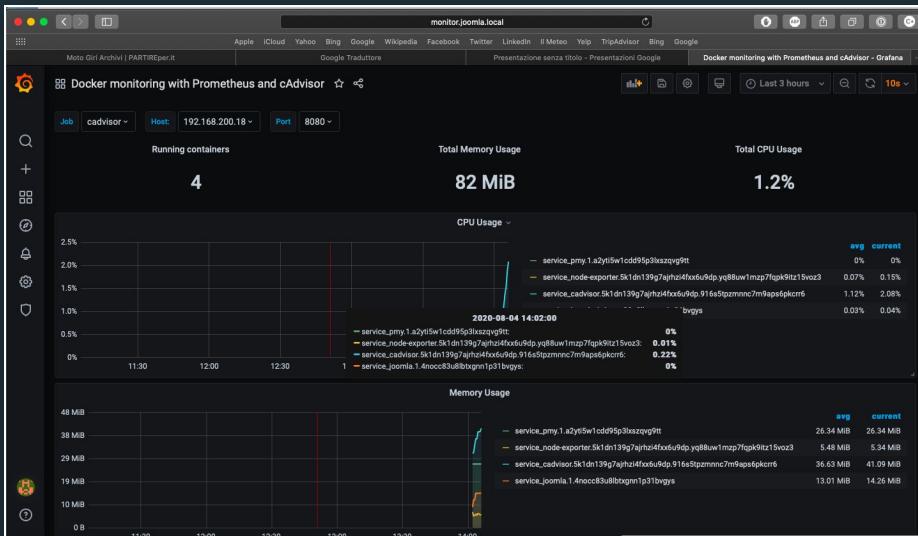
Mostra tutti | Numero di righe: 25 Filtra righe: Cerca nella tabella

Operazioni sui risultati della query

Stampa  Copia nella clipboard  Esporta  Mostra diagramma  Crea vista

Apri "managedb.joomla.local/sql.php?db=joomla&table=prfx\_users&pos=0" in un nuovo pannello dietro quello attuale





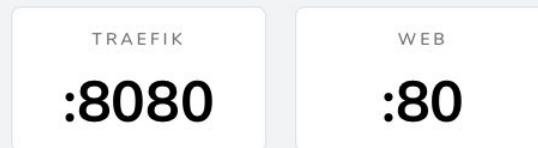
Apple iCloud Yahoo Bing Google Wikipedia Facebook Twitter LinkedIn Il Meteo Yelp TripAdvisor Bing Google

Moto Giri Archivi | PARTIREper.it Google Traduttore Presentazione senza titolo - Presentazioni Goo... Probabile errore nella sicurezza, mi serve aiut... Traefik

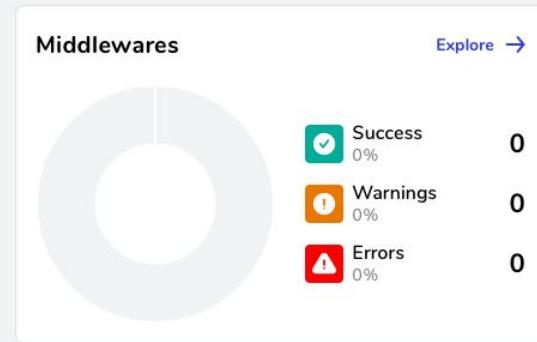
træfik [Dashboard](#) [HTTP](#) [TCP](#)

Documentation Traefik 2.0.0-rc3

→] Entrypoints



HTTP



TCP

# Split-brains Condition

Usando GlusterFS può succedere che a causa di **network partitioning** o di **crash** di alcuni nodi, alcuni dei file presenti nei volumi glusterfs non siano più accessibili. Ciò accade quando uno o più dei file (dati o metadati) di replica cominciano a **divergere** tra loro.

GlusterFS non è quindi più in grado di risolvere automaticamente e di avviare un processo di self heal perchè non ci sono abbastanza informazioni per scegliere in modo autorevole una copia come originale (curando così le “bad copies”), nonostante tutti i bricks siano attivi e online

## Soluzioni

**con “quorum”**

**troubleshooting**

# Soluzioni con “quorum”

- Replica Volume 3**(nostro caso)
- Arbiter Volume:** replica volume 2 con un terzo usato come “**arbitro**” che mantiene solo nomi dei file e metadata

**Replica Volume 3** offre **High Availability** rispetto a **Arbiter Volume** ma occupa più risorse

**Entrambi** attenuano il problema ma non lo risolvono

# Troubleshooting

Mostra lista dei file in condizione di “**split-brains**”:

```
gluster volume heal <miovolume> info split-brain
```

Individuati i file in Split-brains condition, utilizzo come **sorgente**:

- il file più **grande**;
- il file con l'**ultima modifica** più recente;
- uno dei “**brick**” nella **replica**.

Fortunatamente dalla versione 3.8 hanno introdotto un parametro di configurazione all'interno delle opzioni dei volumi:

**cluster.favorite-child-policy**

Questa opzione può assumere i seguenti valori:

- **none**: Valore di default, in questo caso si deve procedere manualmente.
- **ctime**: Seleziona il ed esegue un heal del volume utilizzando il file con il ctime maggiore.
- **mtime**: Seleziona il ed esegue un heal del volume utilizzando il file con il mtime maggiore.
- **size**: Seleziona il ed esegue un heal del volume utilizzando il file con la dimensione maggiore.
- **majority**: Seleziona il file con mtime e size identico in più della metà dei brick nella replica (nel caso di volumi con 3 o più repliche).

Quando è settata ad una delle policy disponibili, risolve in **automatico** lo **Split-brains**.

# Sessioni Persistenti

Condividendo su GlusterFs i volumi dei servizi

# Load Balancing

## Traefik

- "traefik.http.services.monitor.loadbalancer.server.port=...."

Traefik è in grado di fornire, nativamente, servizi di **load balancing** verso le repliche **Swarm**

# Miglioramenti

## Single point of failure

**Traefick** è eseguito sullo Swarm Manager, ciò presenta la problematica di “*Single point of failure*”.

### Soluzione:

Eseguire traefik su un Worker e, in ogni caso, utilizzare un numero maggiore di Manager, possibilmente in numero dispari.

Uno Swarm di tre manager tollera una perdita massima di un manager.

Uno Swarm di cinque manager tollera una perdita massima simultanea di due nodi manager.

Un cluster N manager tollera la perdita di al massimo  $(N-1) / 2$  manager.

Docker consiglia un massimo di sette nodi manager per uno Swarm.