# THE STATE OF THE SWARM

## HOW CLOSE TO PRODUCTION READY ARE WE ?

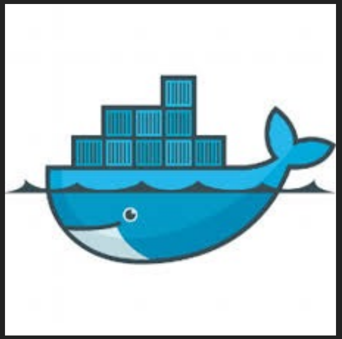
7 / 1 /2016

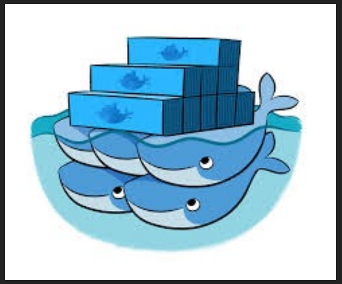
# HI

- Mathieu Buffenoir
- twitter://@MBuffenoir
- mail://mat.buff@gmail.com
  - founder bity.com (running on docker on exoscale)
  - VP swiss bitcoin association
- https://github.com/skippbox/docker-on-cluster-howtos

# WHAT'S NEW IN DOCKER ECOSYSTEM ?

- Docker 1.9
  - Networking in the swarm
- Compose 1.5
  - environnement variable
- Docker-machine
  - Added cloud providers support



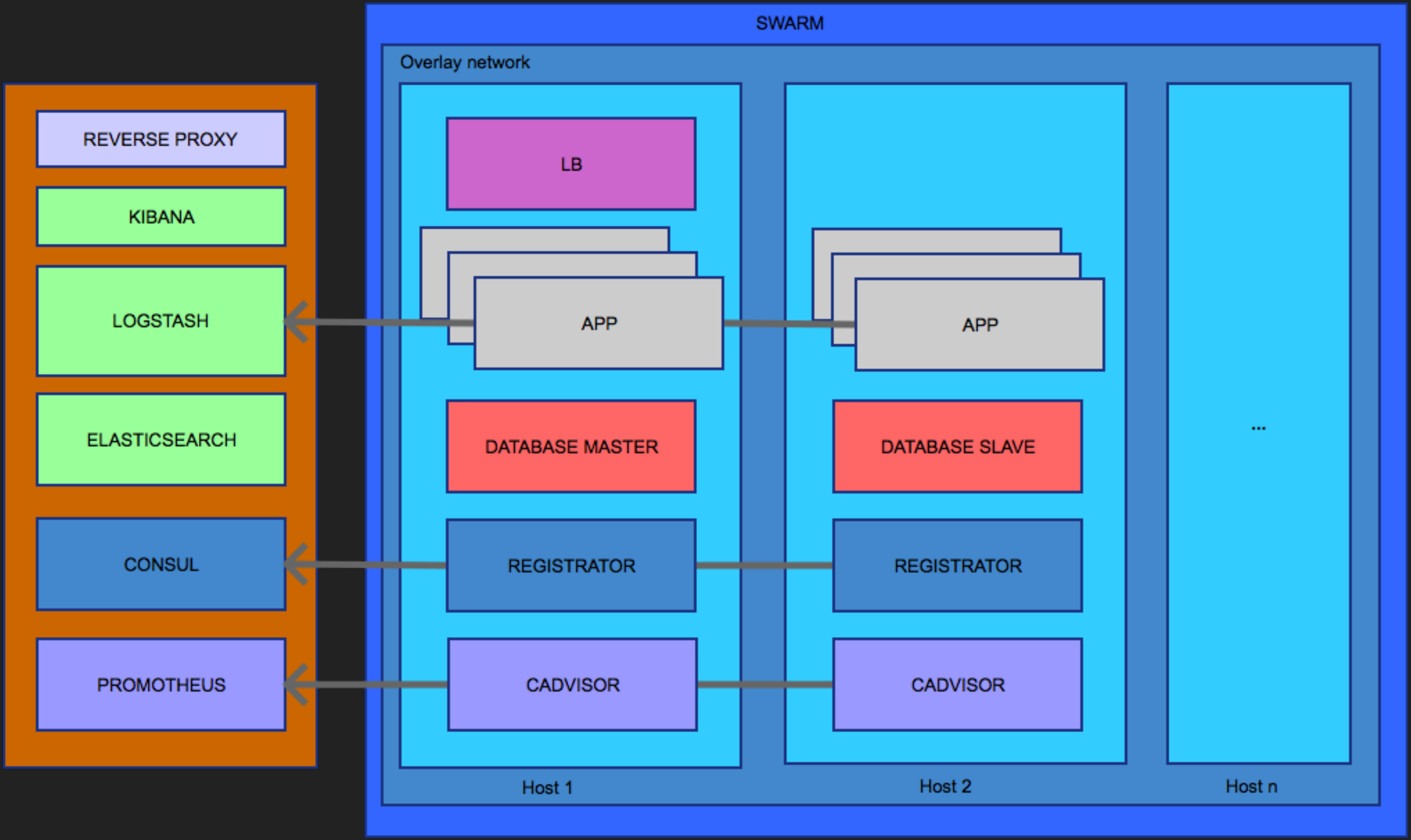


# WHAT DO WE NEED IN PRODUCTION ?

- Provisioning / orchestration (Swarm)
- Service discovery (consul / etc / zookeeper …)
- Logging (ELK, Loggly , syslog …)
- Monitoring (Promotheus, sensu, sysdig …)

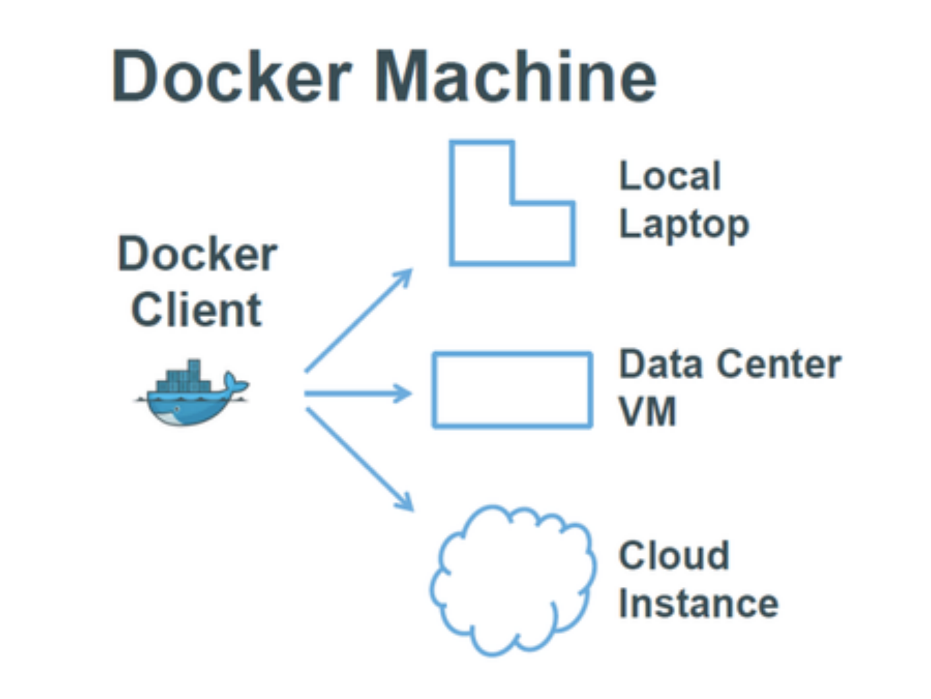
# LET'S CREATE A LITTLE DEMO INFRASTRUCTURE

# DEMO INFRA


REVERSE PROXY
KIBANA
LOGSTASH
ELASTICSEARCH
CONSUL
PROMOTHEUS
SWARM
Overlay network
LB
APP
APP
DATABASE MASTER
DATABASE SLAVE
REGISTRATOR
REGISTRATOR
CADVISOR
CADVISOR
...
Host 1
Host 2
Host n

# DOCKER MACHINE

- cloud provider drivers (12 as of today) or bare metal
- some handy features
  - ssh / scp
- One command to control your node or cluster directly from your shell:
  - eval $(docker-machine env --swarm swarm-master)
  - docker ps

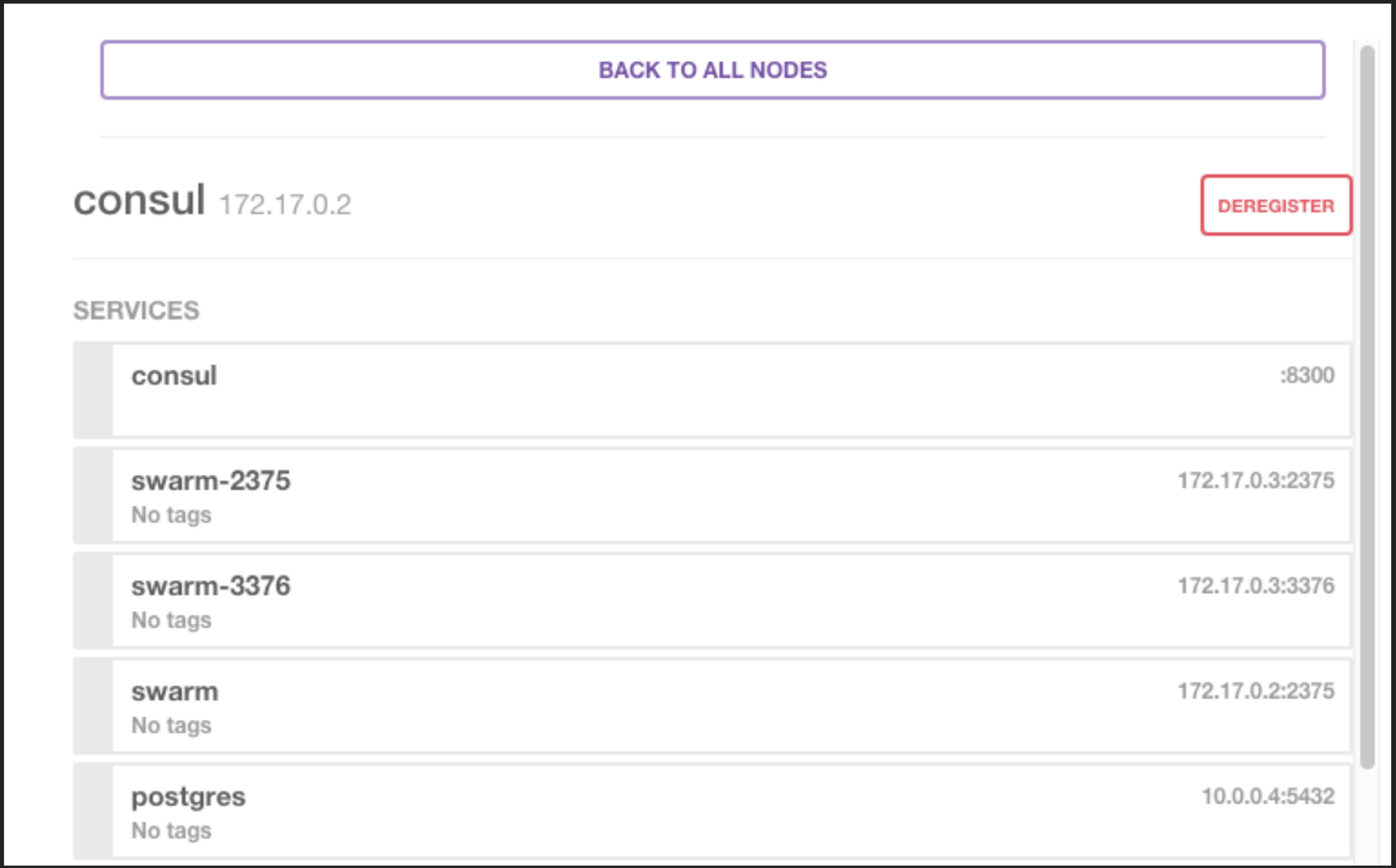


# ONE COMMAND TO CREATE A CLUSTER NODE

```
docker-machine create --driver exoscale \
    --exoscale-api-key $CLOUDSTACK_KEY \
    --exoscale-api-secret-key $CLOUDSTACK_SECRET_KEY \
    --exoscale-instance-profile small \
    --exoscale-disk-size 10 \
    --exoscale-image ubuntu-14.04 \
    --exoscale-security-group swarm \
    --swarm \
    --swarm-master \
    --swarm-discovery="consul://$(docker-machine ip consul):8500" \
    --engine-opt="cluster-store=consul://$(docker-machine ip consul):8500" \
    --engine-opt="cluster-advertise=eth0:2376" \
    --engine-label="apps" \
    swarm-master
```

# KV STORE (CONSUL)

- Consul
  - services
  - nodes
  - key-value
  - multi-datacenter
  - health- check
- REST or DNS api



consul 172.17.0.2

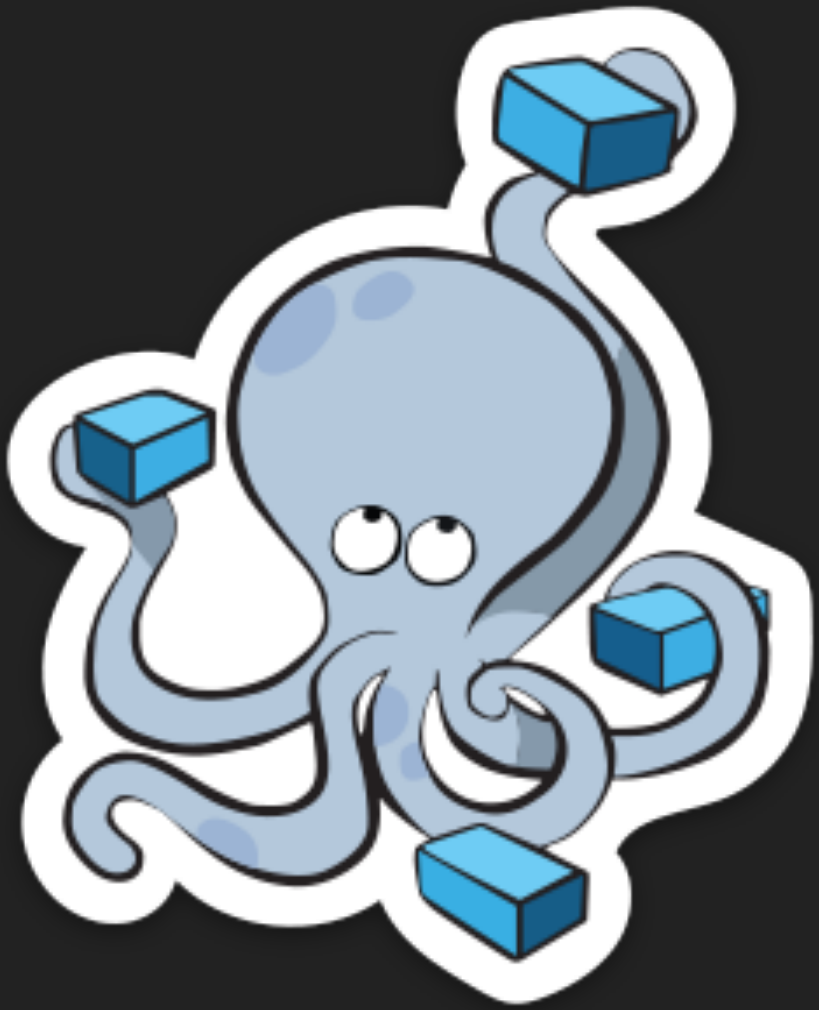
SERVICES

| | |
|---|---|
| consul | :8300 |
| swarm-2375<br>No tags | 172.17.0.3:2375 |
| swarm-3376<br>No tags | 172.17.0.3:3376 |
| swarm<br>No tags | 172.17.0.2:2375 |
| postgres<br>No tags | 10.0.0.4:5432 |

# COMPOSE FEATURES

- control your cluster straight from your shell
- networking support
- environment variables
- support for docker log driver
- scaling
- filters

```
The node filters are:
constraint
health
The container configuration filters are:
affinity
dependency
port
```



# OUR DEMO INFRA COMPOSE FILE

```
ghost:
  image: ghost
  restart: always
  ports:
    - 2368
  volumes:
    - /home/ubuntu/conf-files/
config.js:/var/lib/ghost/config.js
  environment:
    - DB_URI=swarm_db_1
    - NODE_ENV=production
  log_driver: "syslog"
  log_opt:
    syslog-address: "udp://
185.19.29.213:5000"
    syslog-tag: "ghost"
```

```
db:
  image: postgres:9.3
  restart: always
  environment:
    DB_PASSWORD: postgres
    DB_USER: postgres
    DB_NAME: ghost
  ports:
    - 5432

lb:
  image: lalu/haproxy-consul
  restart: always
  volumes:
    - /home/ubuntu/conf-files/
haproxy.ctmpl:/tmp/haproxy.ctmpl
    - /home/ubuntu/conf-files/consule-
template.conf:/tmp/consule-
template.conf
  ports:
    - "80:80"
    - "8001:8001"
  command: -consul 185.19.29.213:8500
```

# COMPOSE COMMANDS

- docker-compose up (-d) <container>
- docker-compose stop / start / restart <container>
- docker-compose ps
- docker-compose logs

# OVERLAY NETWORK

- /etc/hosts
- dns with consul
- Kernel >3.16
  - - udp 4789 Data plane (VXLAN)
  - - tcp/udp 7946 Control plane
- no more links support (use service discovery)
- need to run compose with –x-networking argument

# SERVICE DISCOVERY

- registrator informs consul when services come on/offline
- patch to support overlay network (now merged)
  - currently support only one network

# SCALING

# COMPOSE

- As simple as:
  - docker-compose –x-networking scale app=5

# CONSUL-TEMPLATE

- official haproxy image extended with consul-template

```
#templating system snippet

backend ghost

  option forwardfor # add the X-Forwarded-For header

  http-request set-header X-Forwarded-Port %[dst_port]

  balance roundrobin{{range service "ghost"}}

  server {{.ID}} {{.Address}}:{{.Port}}{{end}}
```

# LOGGING

# EASILY SET UP AN ELK WITH COMPOSE

- docker-compose up -d
- add this in your docker-compose.ml file service definition

```
log_driver: "syslog"
log_opt:
  syslog-address: "udp://185.19.29.213:5000"
  syslog-tag: "ghost"
```

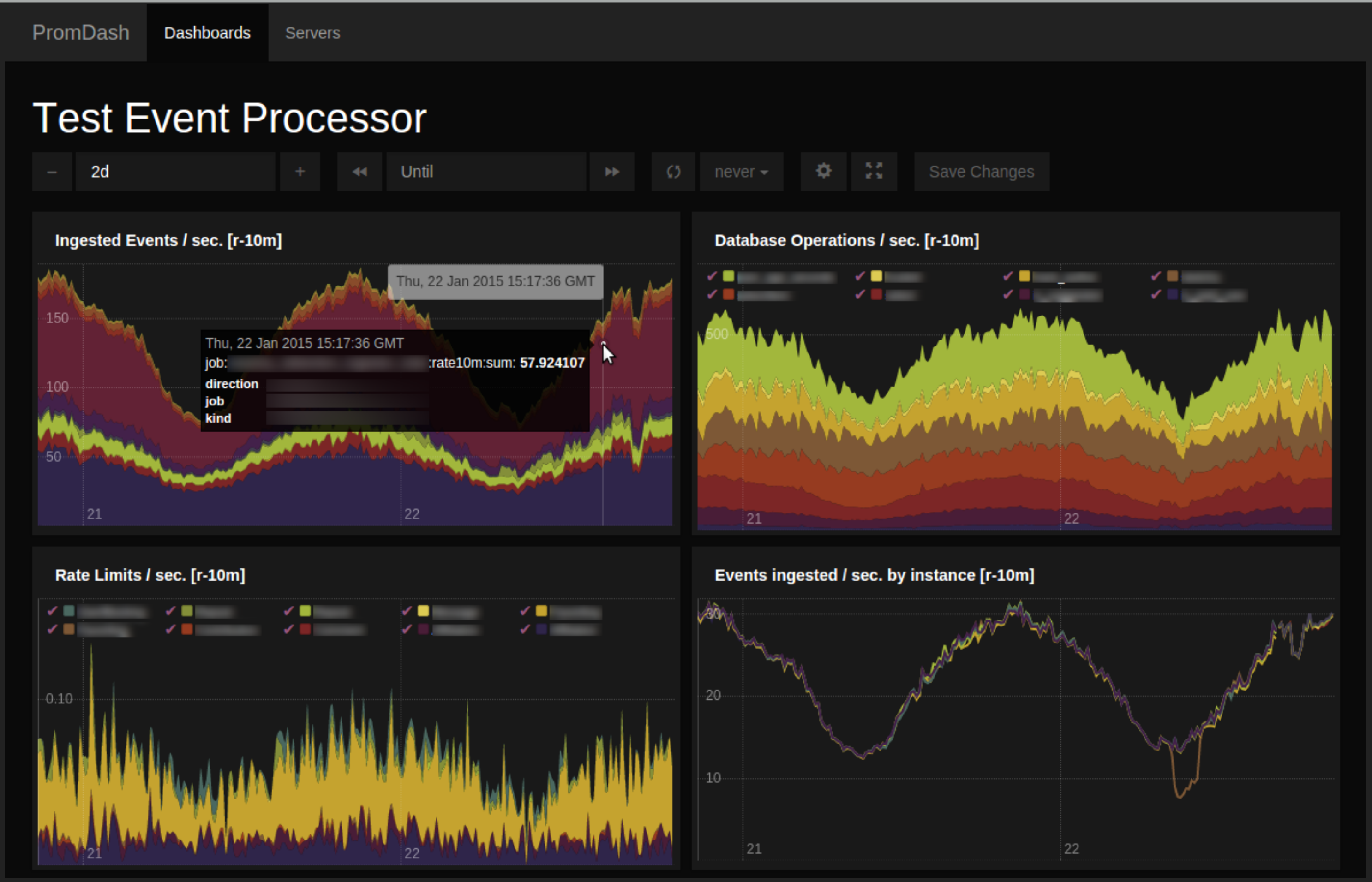
# MONITORING

# CADVISOR

- Collect per host container metrics
- Some visualisations
- not centralised enough

# PROMOTHEUS

- Graphing


PromDash
Dashboards
Servers
Test Event Processor
2d
Until
never
Save Changes
Ingested Events / sec. [r-10m]
Thu, 22 Jan 2015 15:17:36 GMT
Thu, 22 Jan 2015 15:17:36 GMT
job: [redacted]:rate10m:sum: 57.924107
direction
job
kind
150
100
50
21
22
Database Operations / sec. [r-10m]
600
21
22
Rate Limits / sec. [r-10m]
0.10
21
22
Events ingested / sec. by instance [r-10m]
30
20
10
21
22

- Alerting

```
ALERT HighMemoryAlert
  IF container_memory_usage_bytes{image="ubuntu:14.04"} > 1000000000
  FOR 1m
  WITH {}
  SUMMARY "High Memory usage for Ubuntu container"
  DESCRIPTION "High Memory usage for Ubuntu container on {{$labels.instance}} for
container {{$labels.name}} (current value: {{$value}})"
```

# WHAT IS STILL MISSING ?

- secret handling
  - ansible vault
  - hashicorp vault
  - Lots of discussion about this on github
- Discovery service with multiple overlay network support
- support for multiple networks in consul (not sure if it can be achieved with competitors either yet)
- Support in provisioning docker module (Ansible is really good with that)

# THANKS TO

- Exoscale
- hashicorp
- gliderlabs
- sirile
- progrium
- Docker for all the tools

https://github.com/skippbox/docker-on-cluster-howtos

@MBUFFENOIR

# QUESTIONS ?