# CoreOS Flannel for container Networking


Héctor Fernández
Software Engineer
@hectorj2f

# Avoid the complexity

# Deploy independently

# Overview

- Intro CoreOS
- Etcd
- Flannel
- Demo
- Other projects

# What is CoreOS ?

# Linux distro

Based on a **lightweight** Gentoo distribution

Scalable clustering system

Immutable system

- RO rootfs
- Writeable /etc

B-tree filesystem (btfs)

systemd

# Linux basis

## Features

- Automatic and atomic updates
- Cluster management
- All services are in containers (docker, rkt, ...)
- Service discovery

Tools: **etcd**, locksmith, cloud-config, fleet, rkt, **flannel...**

# etcd

## a distributed key-value store

# Etcd

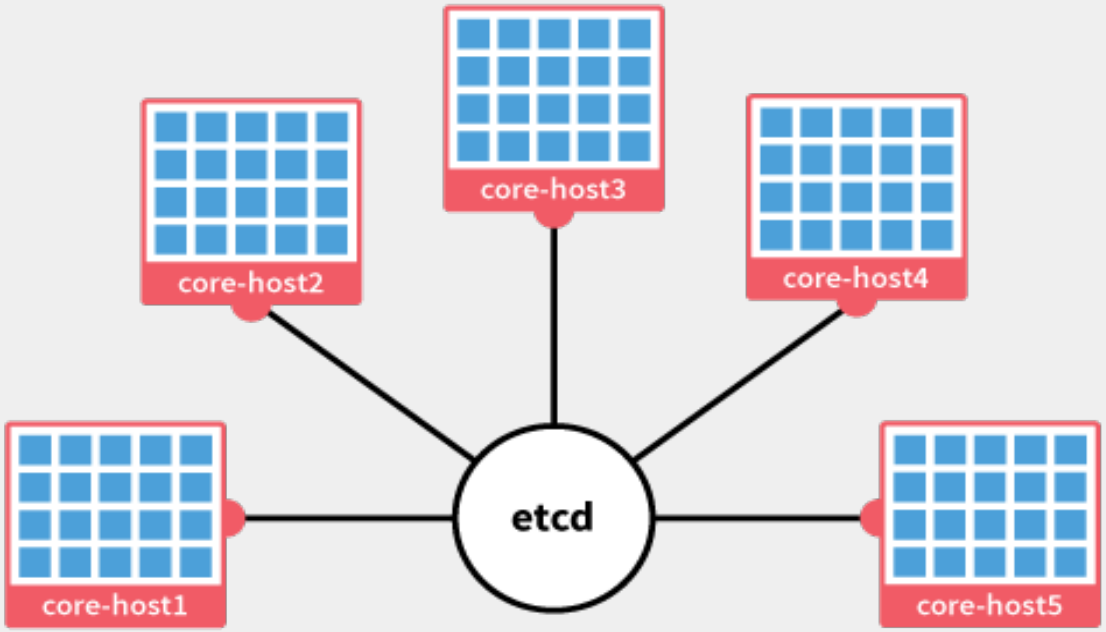
A highly-available key-value store

- Shared configuration
- Service discovery

REST API

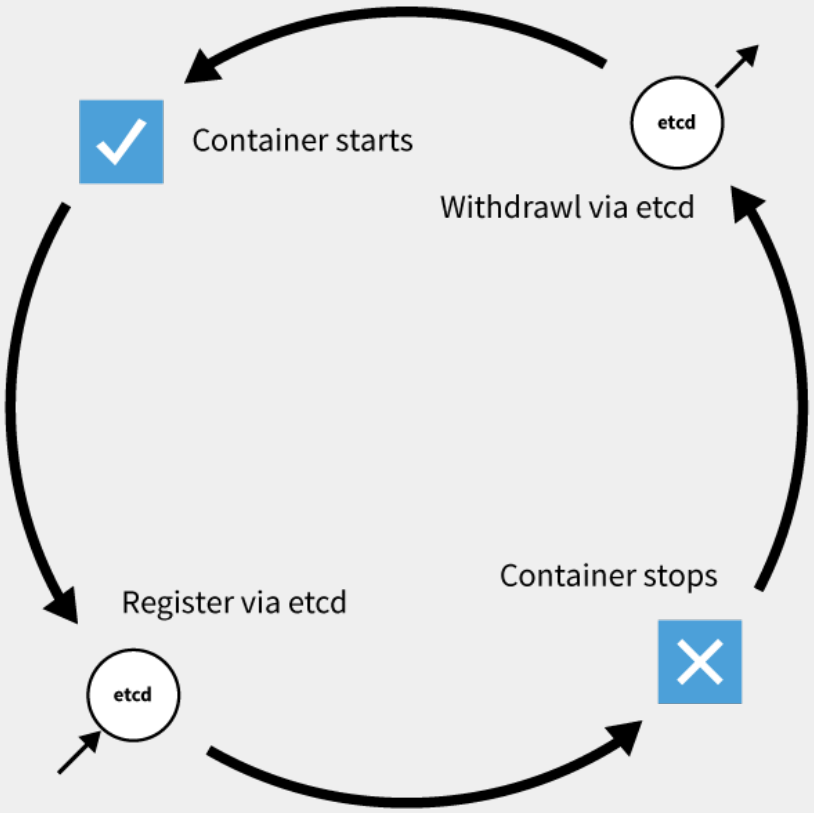
Raft consensus

# Etcd cluster




Giant Swarm

# Shared configuration




Giant Swarm

# flannel

## a network fabric layer for container networking

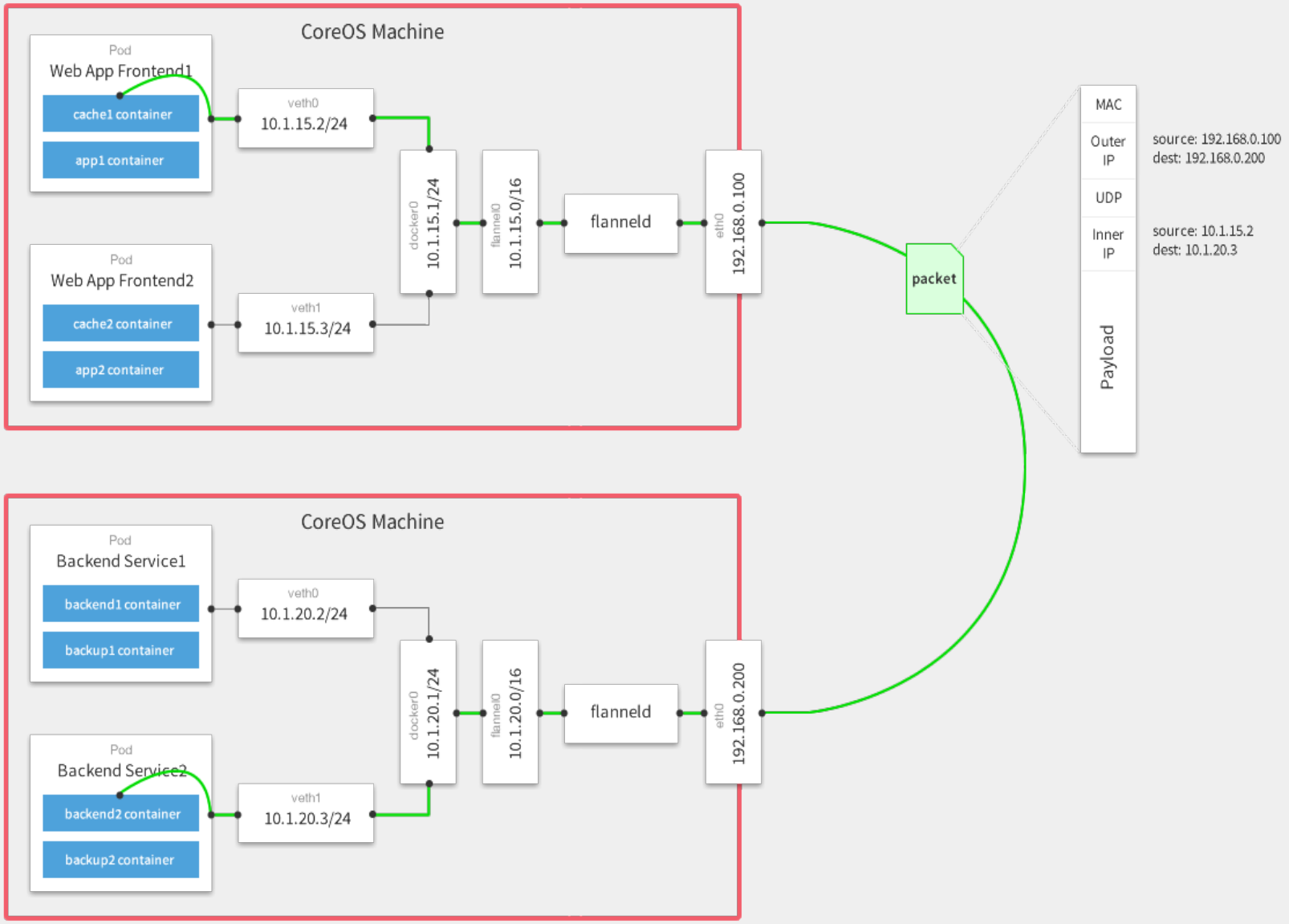
# Flannel

Cloud overlay network (L3)

- One CIDR subnet per host (Kubernetes)
  - Host A: 11.0.47.1/24
  - Host B: 11.0.87.1/24
- Not anymore docker port-based mapping
- Peer network configuration is backed by etcd
- Container talking via IP addresses

Encapsulate packets using UDP and VxLAN

**Backends:** VxLAN, UDP, alloc, host-gw, aws-vpc


CoreOS Machine
Pod
Web App Frontend1
cache1 container
app1 container
veth0
10.1.15.2/24
Pod
Web App Frontend2
cache2 container
app2 container
veth1
10.1.15.3/24
docker0
10.1.15.1/24
flannel0
10.1.15.0/16
flanneld
eth0
192.168.0.100
packet
MAC
Outer IP
source: 192.168.0.100
dest: 192.168.0.200
UDP
Inner IP
source: 10.1.15.2
dest: 10.1.20.3
Payload
CoreOS Machine
Pod
Backend Service1
backend1 container
backup1 container
veth0
10.1.20.2/24
Pod
Backend Service2
backend2 container
backup2 container
veth1
10.1.20.3/24
docker0
10.1.20.1/24
flannel0
10.1.20.0/16
flanneld
eth0
192.168.0.200

Giant Swarm

# Flannel

Limited performance footprint (AWS m3.medium)

| | **Without Flannel** | **With Flannel** |
|---|---|---|
| **UDP** Latency | 133us | 201us |
| **TCP** Bandwidth | 47.8 MB/sec | 47.2 MB/sec |

## On-going work:

- IPSEC as encapsulation protocol (not trusted networks)
- ...

Giant Swarm

[1] Introducing flannel: An etcd backed overlay network for containers

# Limitations

IP address overlap is not possible

Not ideal for multi-tenancy networking #50

Not support for broadcast ARP

- Flannel does proxy ARP

User-space encapsulation and forwarding

- Performance penalty

# Time for a Demo

# Demo

Cassandra database container


Kong API server (http://getkong.org/)

# Setup flannel with user-data

```
# cloud-config
coreos:
     etcd:
          ...
     flannel:
          interface: $public_ipv4
     fleet:
          ...
units:
 - name: flanneld.service
   drop-ins:
     - name: meetup-network-config.conf
       content: |
       [Service]
       ExecStartPre=/usr/bin/etcdctl set /coreos.com/network/config
'{"Network":"10.1.0.0/16", "Backend":{"type":"vxlan",...}}'
   command: start

  ...
```

Giant Swarm

# Manual Setup of Flannel

```
$ git clone https://github.com/coreos/flannel.git

$ docker run -v $SRC:/opt/flannel -i -t google/golang /bin/bash -c "cd

/opt/flannel && ./build"

$ etcdctl --rm /coreos.com/network/ --recursive

$ etcdctl mk /coreos.com/network/config '{"Network":"12.0.0.0/16"}'

$ etcdctl get /coreos.com/network/config

$ docker -d --bip=${FLANNEL_SUBNET} --mtu=${FLANNEL_MTU}

$ sudo /home/core/flannel/bin/flanneld -iface=eth1 &
```

# Other container networking projects

## Project Calico (L3) http://www.projectcalico.org/

- Works like Internet
- ACLs
- Raw network performance

## Weave (L2) https://zettio.github.io/weave/

- Multicast communication
- Work across datacenters
- Container mobility
- DNS

# Giant Swarm

**Simple Microservice Infrastructure, deploy your containers in seconds.**

18 terrific folks, and **hiring**!

http://giantswarm.io/

# Thanks for listening!

**Reach out:**

Héctor Fernández
@hectorj2f
@giantswarm

# Who am I ?

- @hectorj2f
- Software Engineer
- Dislikes: repetitive tasks and top as monitoring tool
- Likes: scalability, performance, go and shell  scripts

Photo References:

https://raw.githubusercontent.com/coreos/flannel/master/packet-01.png

https://alankent.wordpress.com/2014/09/04/cluster-control-with-etcd-and-docker/