Weave Net
Five years with no central control.

FOSDEM 2020

Bryan Boreham
https://weave.works
@bboreham
@weaveworks
Bryan Boreham

Lead on Weave Net since 2015.
Project member of Kubernetes, CNI, Cortex, Scope, ...
Not a networking expert.
Weave Net

- Open Source container network
- Easy to install; runs anywhere*
- No “Enterprise Version”
What is a “container network”?
“There’s no such thing as container networking”

It’s just networking. Anyone who tells you differently is either trying to sell you something or doesn’t understand how networking works.

Justin Garrison
Sep 26, 2016

https://medium.com/@rothgar/no-sdn-kubernetes-5a0cb32070dd
Justin Garrison
@rothgar

Replying to @capileigh and @resouer

I use weave (via kubeadm) in my home metal cluster. I don't have a ton of experience debugging it but that's because it has just worked for me.

11:42 PM · May 20, 2018

https://twitter.com/rothgar/status/998333265739042816
What is a “container network”

Containers give you isolation.

- Each container runs in its own network namespace.

How do these network namespaces talk to each other?
- That’s a container network.
Let’s look at how it works
Container network model
Matthew Sackman

Ex-RabbitMQ, Erlang expert.

Wrote the first version of Weave Net.

3,400 lines of Go
Containers with bridges
Weave Net 1.0

veth

pcap

UDP
Distributed Ethernet Switch*

Weave Net daemon learns where MACs come from
- when it sees the first packet from that MAC.

Thus, it knows where to send each packet**.

If it doesn’t know where a MAC comes from?
- send it everywhere!
In our new world of containers Docker, many old problems have been rediscovered. Thankfully, the fact that these problems were solved decades ago has not stopped people from coming up with their own solutions, and we now all get to witness the resulting disasters.

The particular problem I’ll talk about today is IP level overlay networks. The basic problem these overlay networks try to solve is “I have 1 IP per machine, but I need a subnet multiple IPs per machine”. This was originally relevant because you might want have a few networks and want to do some funky networking, then became relevant because you might have a few VMs and want to do some funky networking, and now, in 2015, it is relevant because you might have a few containers and want to do some funky networking. Obviously, these use cases are distinct enough to require their own implementations and protocols, as we will see.

The way you solve this problem is usually via some sort of IP encapsulation, though the specific implementation will vary wildly. The IP encapsulation RFC talks about a structure that would look like
Weave Net 1.2 “Fast Data Path”

[Diagram showing the flow of data through OVS Datapath and VXLAN using UDP]
David Wragg

Ex-Pivotal

Implementer of the “fast data path”

Now at Cloudflare

https://github.com/weaveworks/go-odp/
How to set up all the devices?
Software-Defined Networking tools for LXC (LinuX Containers)

Jérôme Petazzoni

Pipework

Software-Defined Networking for Linux Containers

Pipework lets you connect together containers in arbitrarily complex scenarios. Pipework uses cgroups and namespacpace and works with "plain" LXC containers (created with lxc-start), and with the awesome Docker.
The weave script
Encryption

https://github.com/weaveworks/weave/blob/master/docs/fastdp-crypto.md
Martynas Pumputis

Implementer of Weave Net XFRM encryption.

Kernel fixes for conntrack race conditions, etc.

Now at Isovalent (Cillium)

https://www.weave.works/blog/racy-conntrack-and-dns-lookup-timeouts
Multicast

Weave Net handles multicast

- via the “send the packet everywhere” logic.
Peers and Topology
Peers and topology
IP Address Management

Gossip
Community
Weave Net installs per week
Lots of requests, very few PRs

Please make NO_MASQ_LOCAL=1 the default setting.

Lost 2 days on this stuff. Indeed, please make this the default.

I'm wondering, is there any specific reason this isn't the default? Does this being on break something else, or cause something to behave unexpectedly?
Mostly paid contributors
Kubernetes
Kubernetes

Mandates NAT-free network between “pods”.

3rd-party pod networks.

Rkt, from CoreOS, has a simple ‘exec’ model to add a network.
CNI - the Container Network Interface

```
{  
    "cniVersion": "0.3.0",  
    "name": "mynet",  
    "type": "my-plugin",  
    "ipam": {  
        "type": "host-local",  
        "subnet": "10.4.0.0/24",  
    }  
}
```
Installing via DaemonSet

DaemonSet runs on every node

Pod mounts host directory and copies plugin at startup
End of main content
Things I didn’t cover

Kubernetes Network Policy

Launch modes

Scalability

Service Management / Service Discovery

Bug bounty programme
Questions?

Bryan Boreham
https://weave.works
@bboreham
@weaveworks