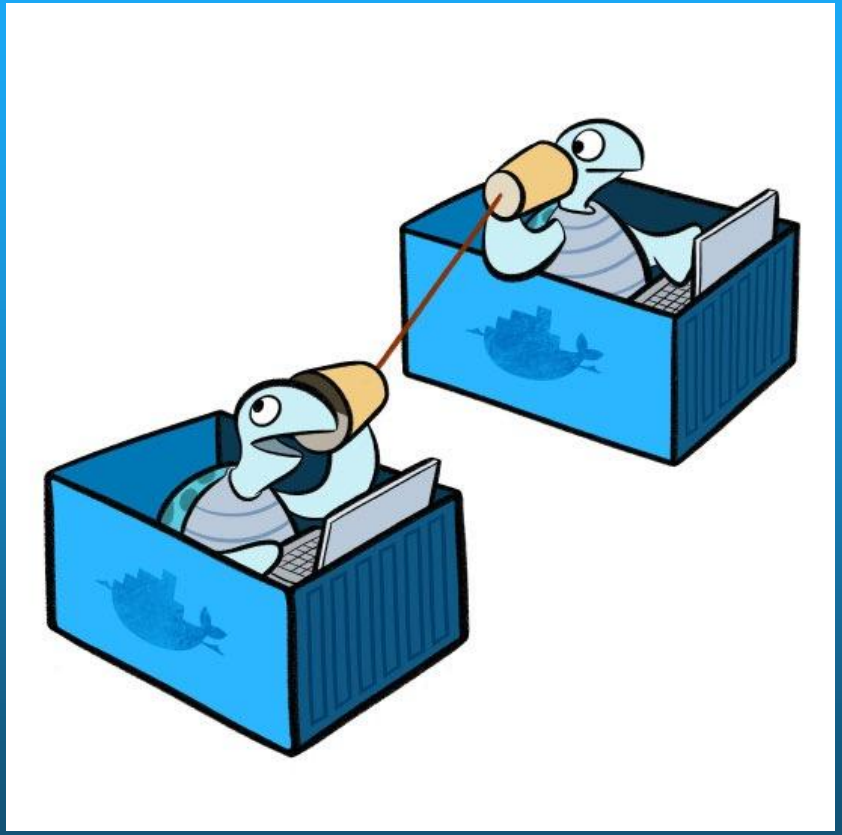


Docker Networking

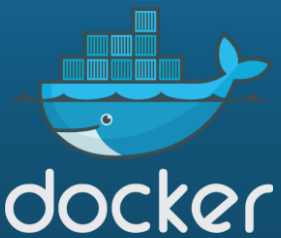
Control-plane & Data-plane

@MadhuVenugopal
@mrjana

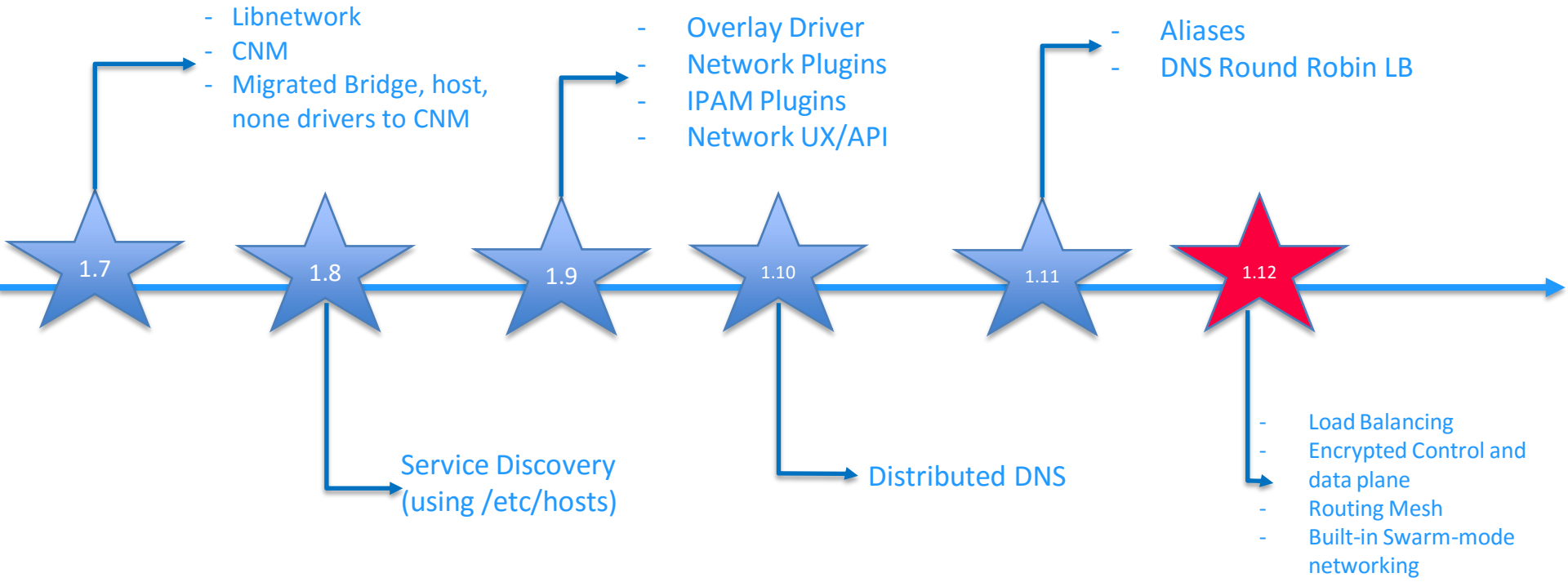
docker

Agenda

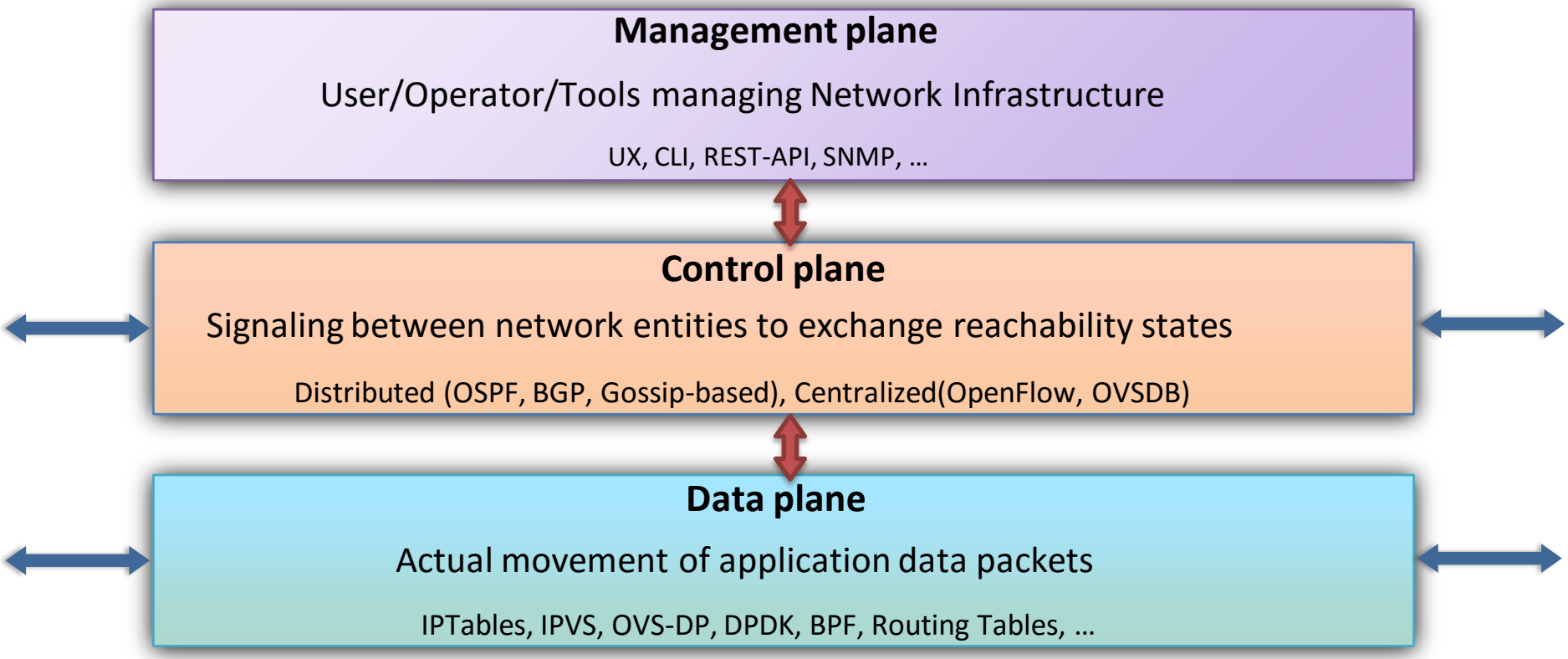
- Docker Networking
  - Features
  - Control plane & Data plane
- Deep Dive
  - Control plane
  - Data plane
- Q & A



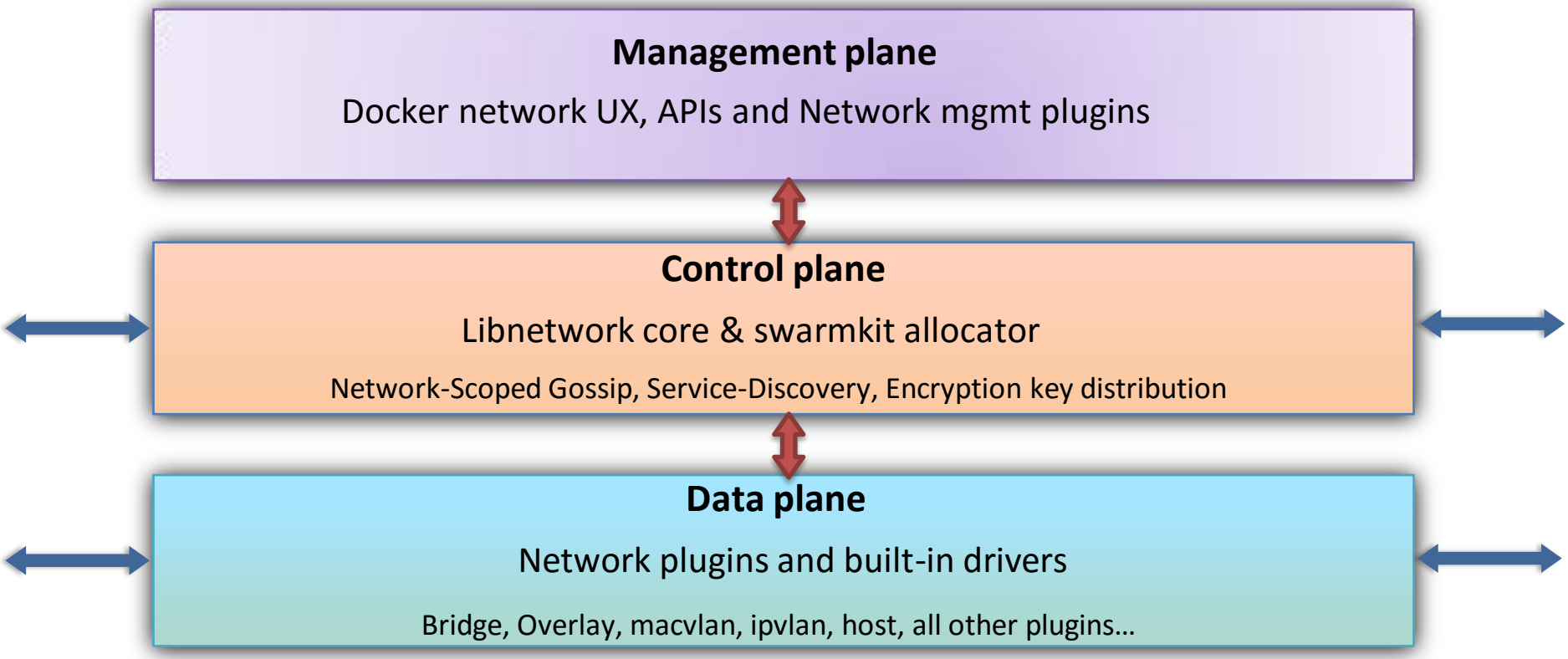
Docker Networking



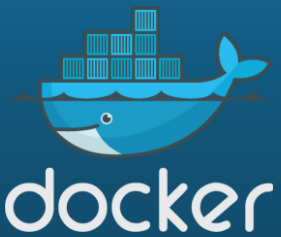
Networking planes



Docker networking planes



Deep Dive - Control Plane

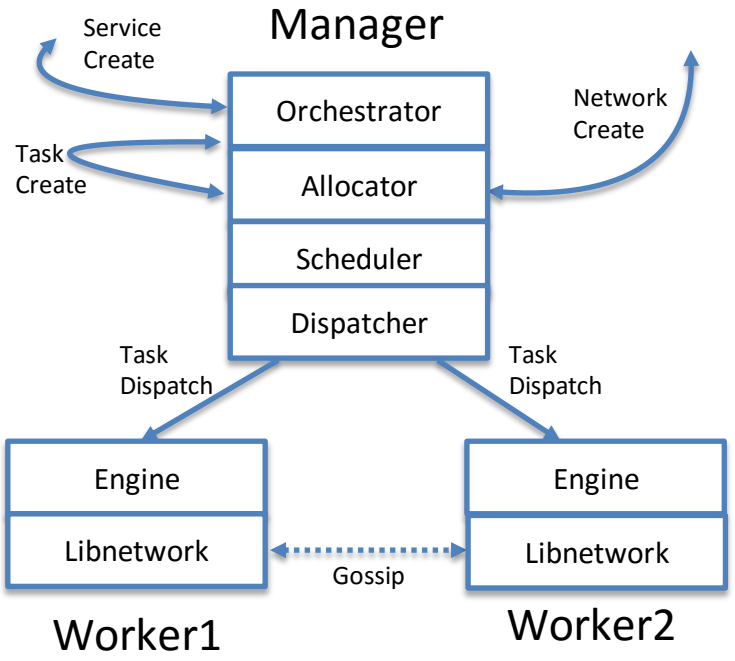


Control plane components

- Centralized resources and policies
- De-centralized events

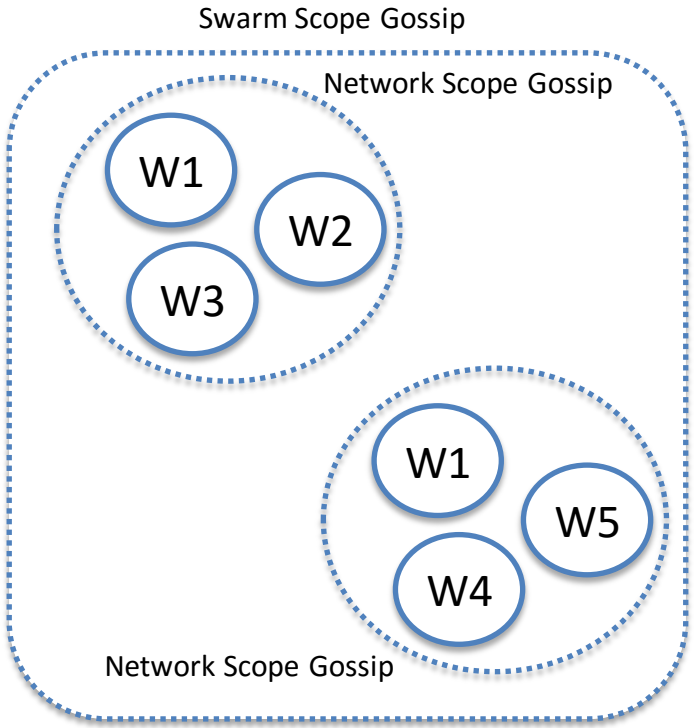
Centralized resources and policies

- Resources and policies are defined centrally
- Networks are a definition of policy
- Central resource allocation (IP Subnets, Addresses, VNIs)
- Can mutate state as long as managers are available



De-centralized events

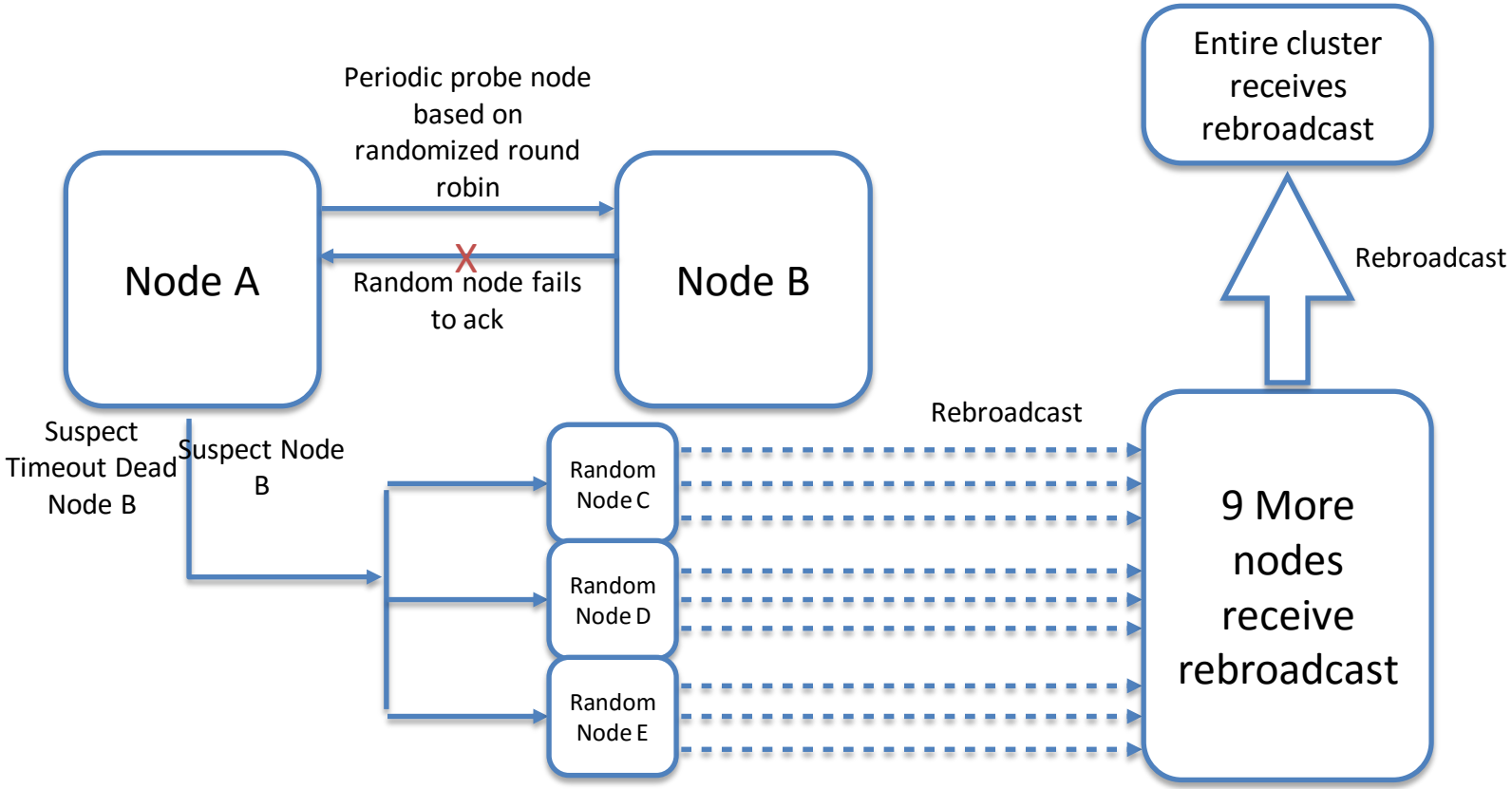
- State is learned through de-centralized dissemination of events
- Gossip based protocol
- Fast convergence
- Highly scalable
- Continues to function even if all managers are Down



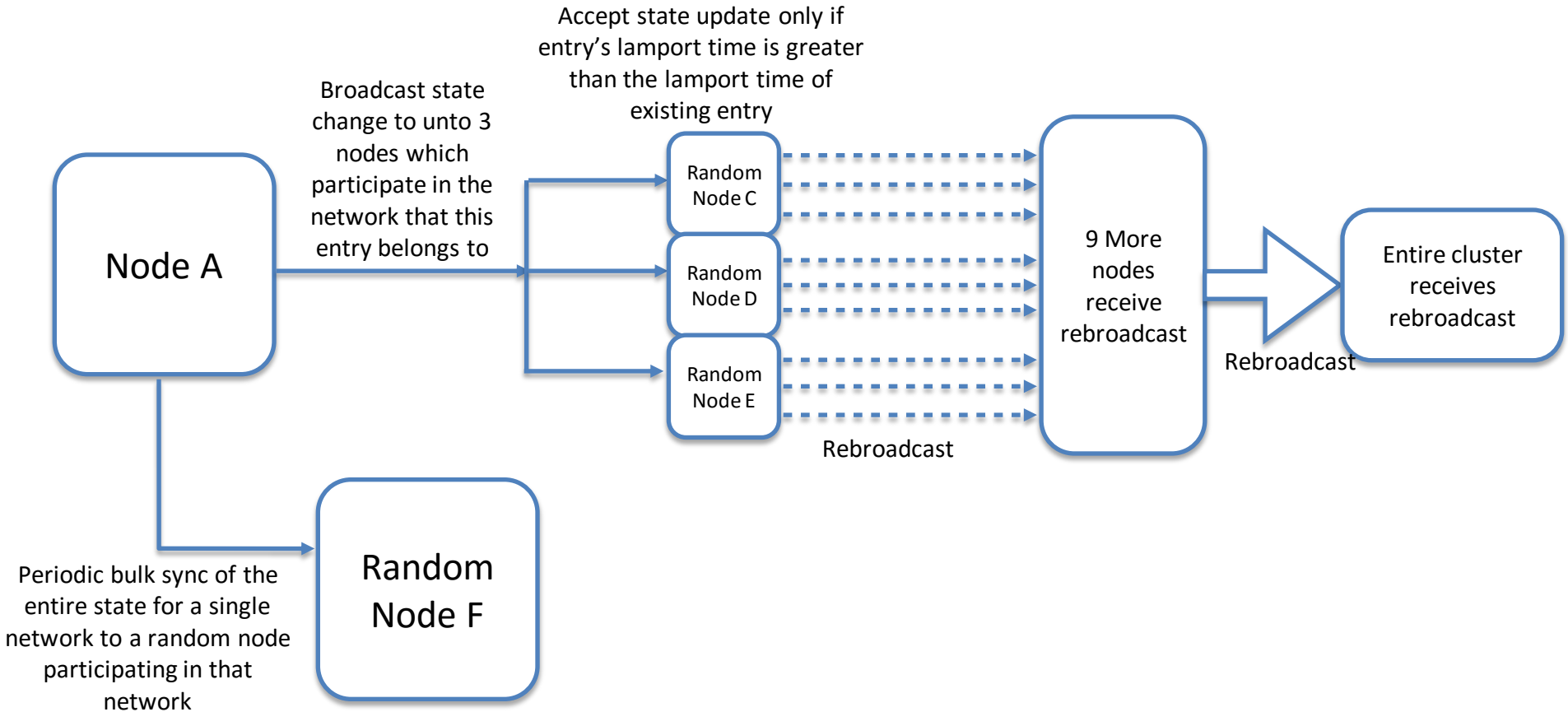
Gossip in detail

- Completely de-centralized discovery of cluster nodes
- Cluster membership is discovered using an implementation of Scalable Weakly-consistent Infection-style Process Group Membership Protocol (SWIM)
- Two kinds of cluster membership:
  - Swarm level
  - Network level
- Sequentially consistent state dissemination ordered by a lamport clock
- Single writer at a record/entry level
- Convergence time roughly has a O(logn) asymptotic time complexity

Failure detection

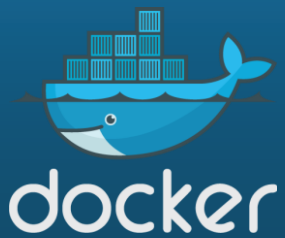


State dissemination



Deep Dive - Data Plane

Overlay driver



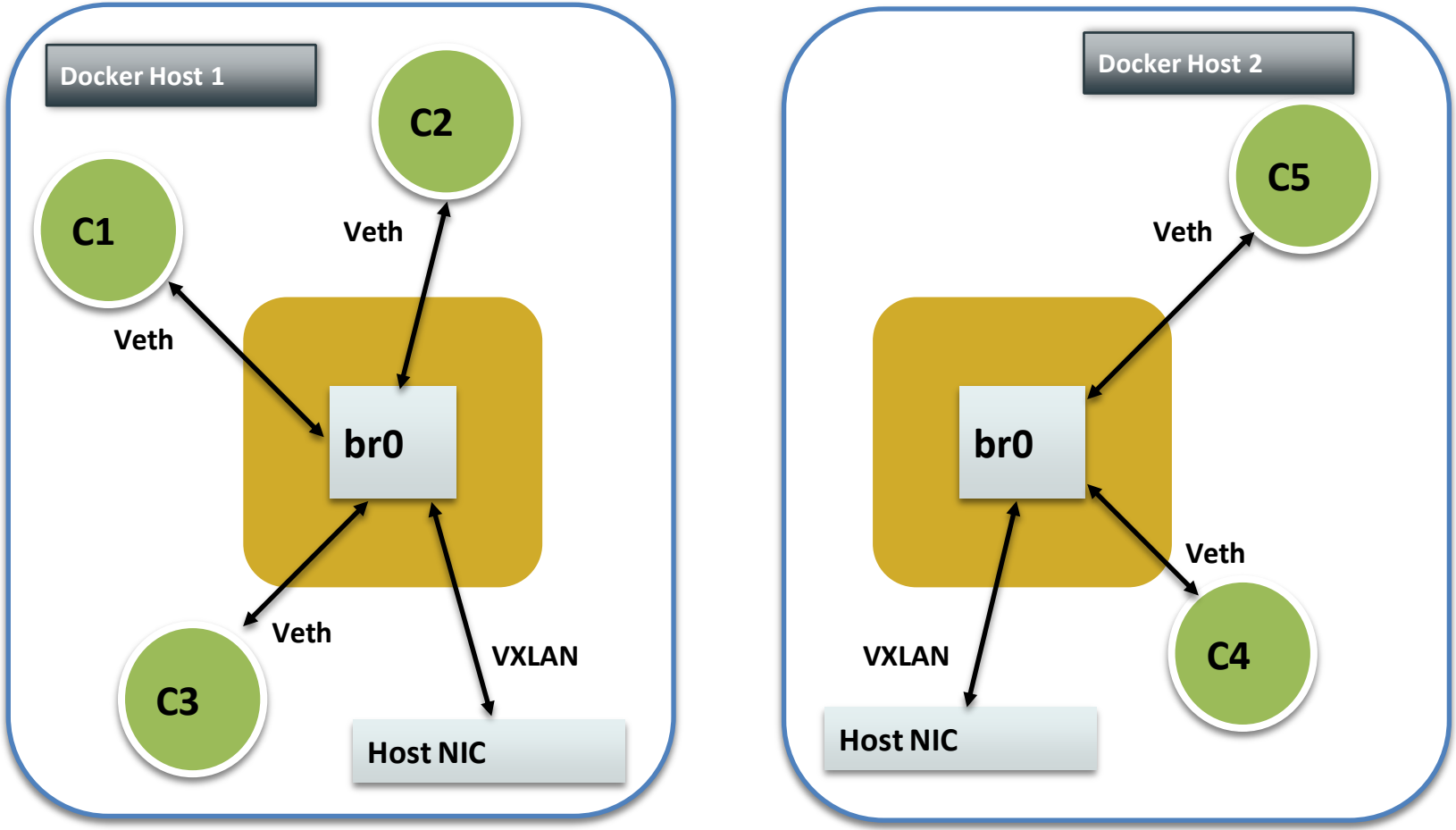
Overlay Networking Under the Hood

- Virtual eXtensible Local Area Network(VXLAN) data transport
- L2 Network over an L3 network (overlay)
- RFC7348
- Host as VXLAN Tunnel End Point (VTEP)
- Point-to-Multi-Point Tunnels
- Proxy-ARP

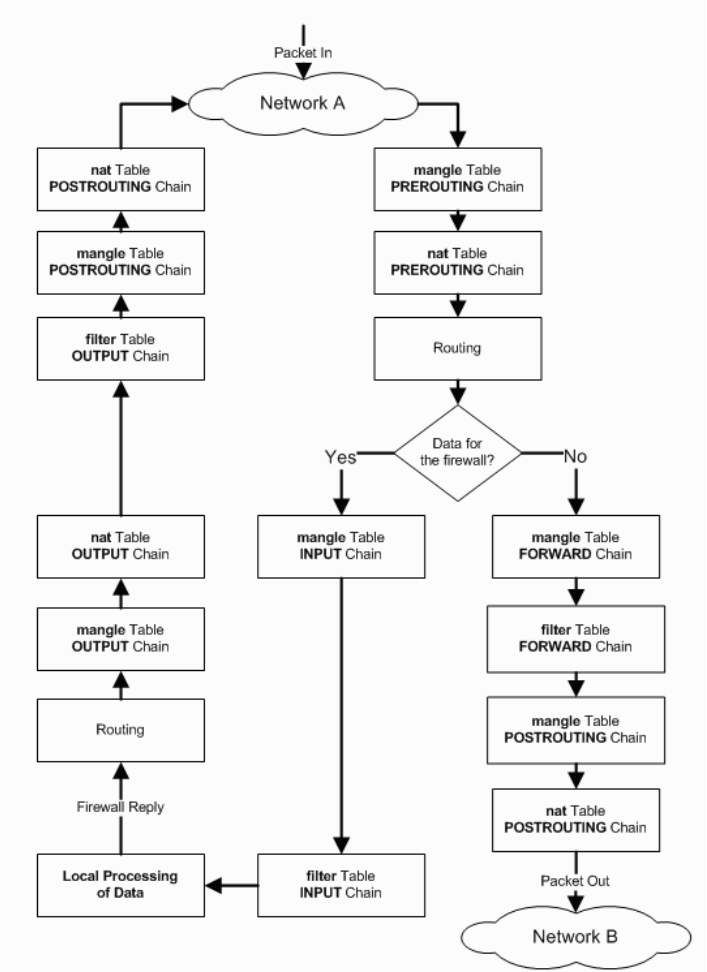
Overlay Networking Under the Hood

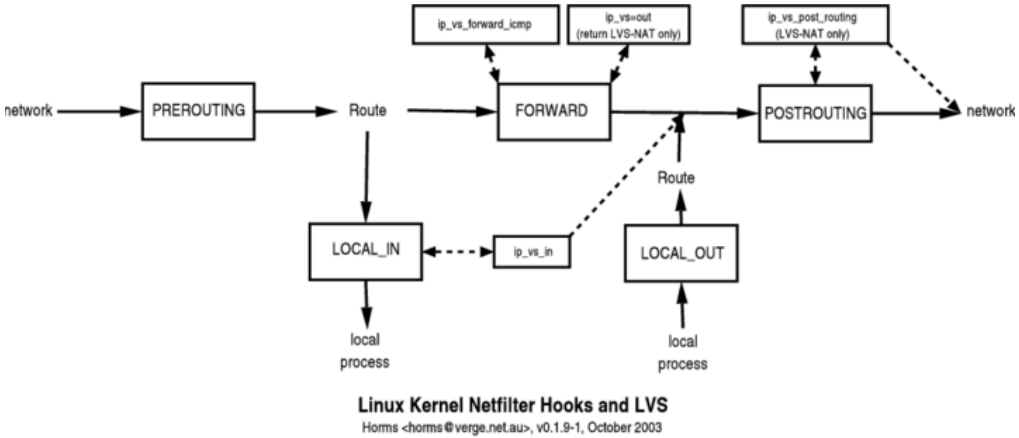
- A Linux Bridge per Subnet per Overlay Network per Host
- A VXLAN interface per Subnet per Overlay Network per Host
- **1** Linux Bridge per Host for default traffic (docker_gwbridge)
- Lazy creation (Only if container is attached to network)

Overlay Networking Under the Hood



Linux Kernel NetFilter dataflow



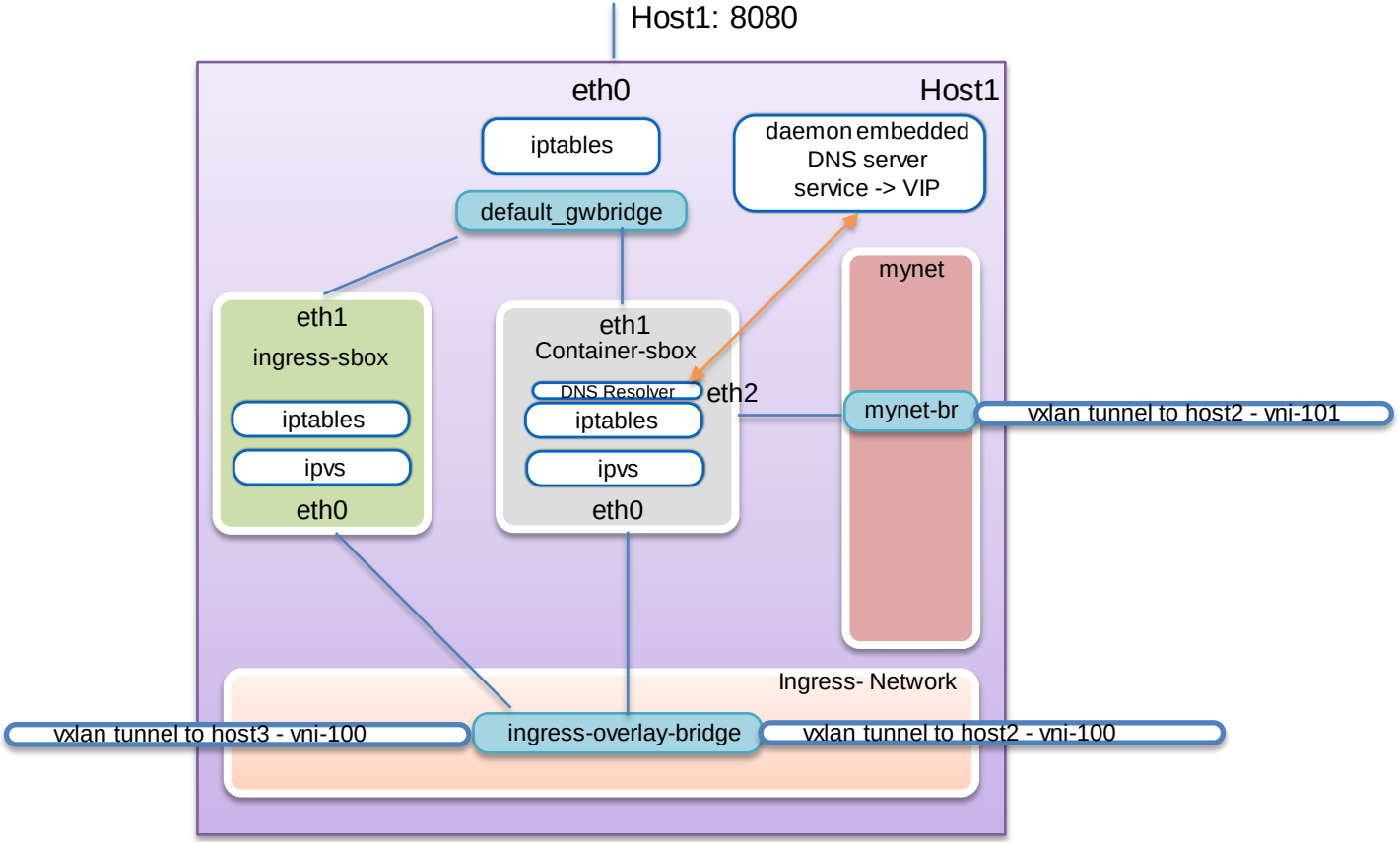


Linux Kernel Netfilter Hooks and LVS

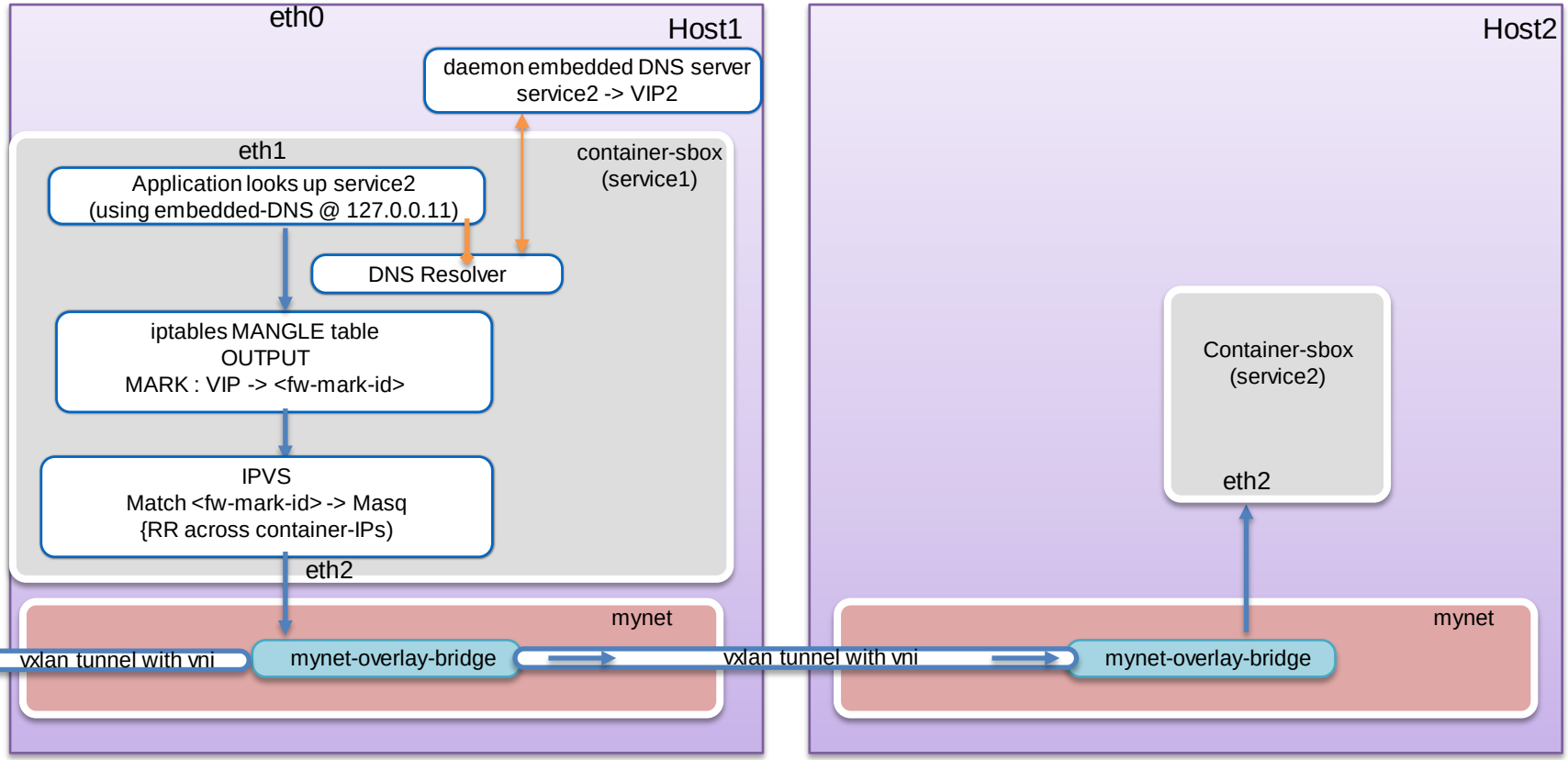
Horms <horms@verge.net.au>, v0.1.9-1, October 2003

Service , Port-Publish & Network

docker service create —name=test —network=mynet -p 8080:80 —replicas=2 xxx

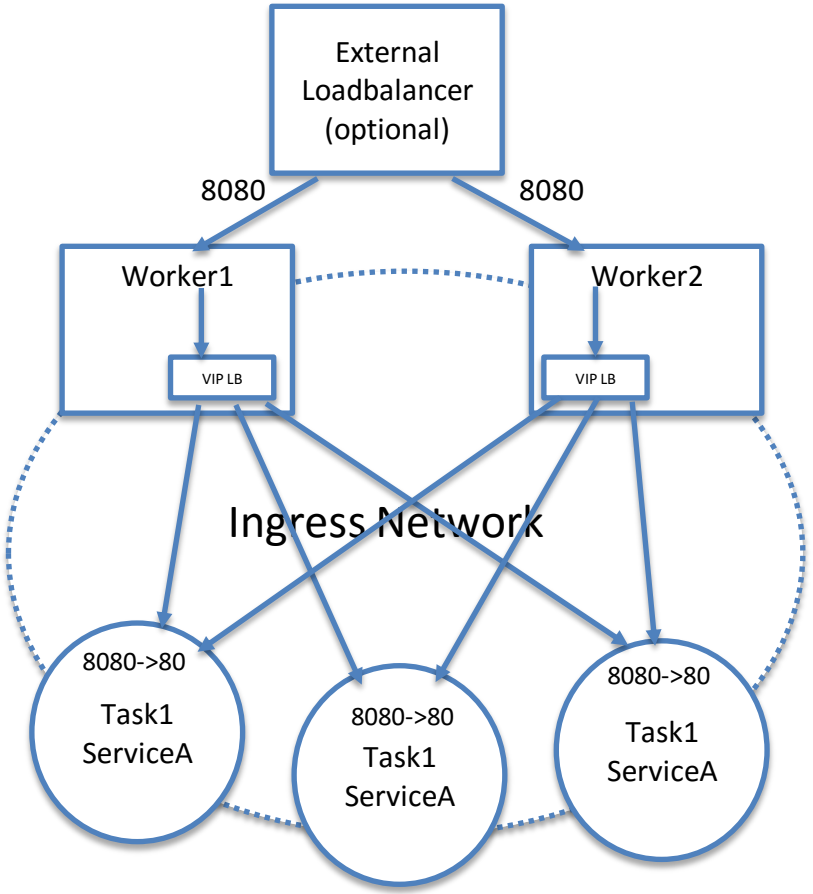


Day in life of a packet - Internal LB

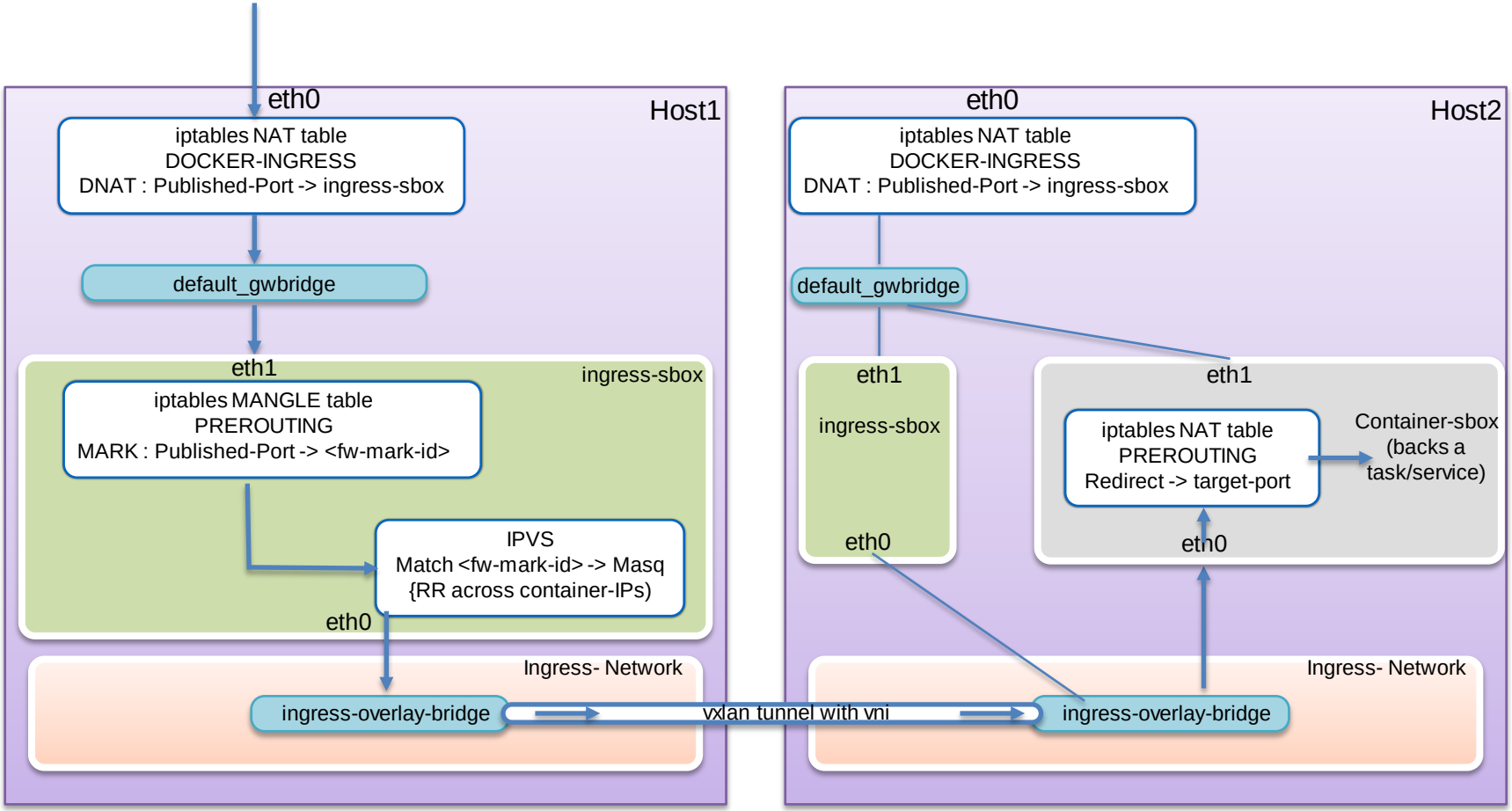


Routing mesh

- Builtin routing mesh for edge routing
- Worker nodes themselves participate in ingress routing mesh
- All worker nodes accept connection requests on PublishedPort
- Port translation happens at the worker node
- Same internal load balancing mechanism used to load balance external requests



Day in life of a packet - Routing Mesh & Ingress LB



Q&A