CSE 550: Systems for all

Au 2021

Ratul Mahajan

Building Massive Cloud Networks


Microsoft Azure


amazon
web services


Alibaba Cloud


Google Cloud

Image from Microsoft Azure

MICROSOFT | TECH | FACEBOOK

Microsoft and Facebook just laid a 160-terabits-per-second cable 4,100 miles across the Atlantic

47

Enough bandwidth to stream 71 million HD videos at the same time

By Thuy Ong | @ThuyOng | Sep 25, 2017, 7:56am EDT

https://www.nytimes.com/interactive/2019/03/10/technology/internet-cables-oceans.html

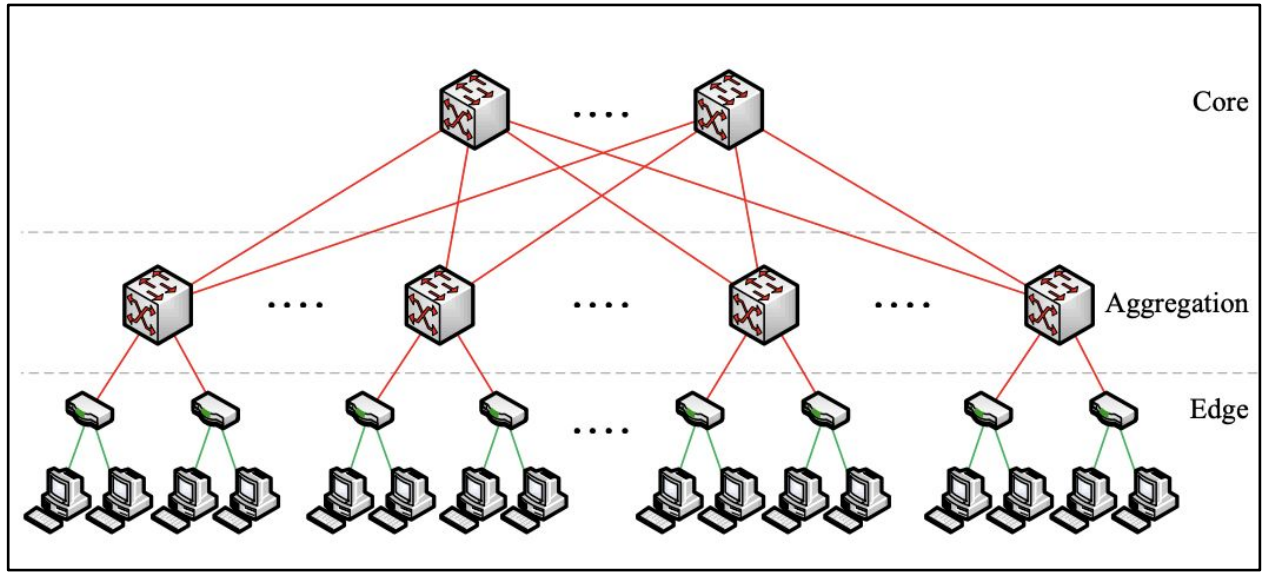
HUGE data center networks (DCN)

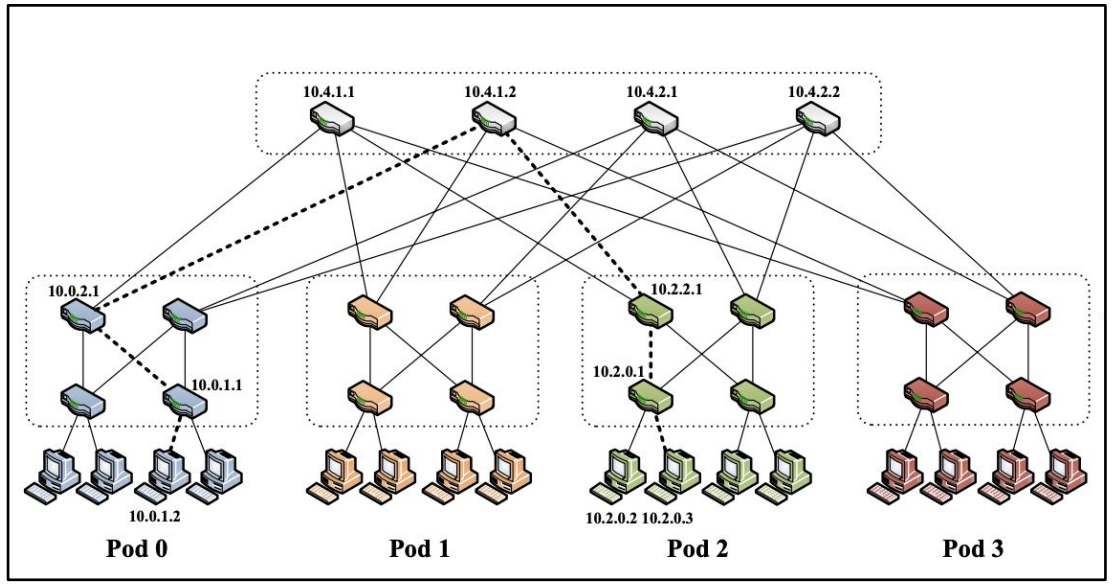
- Thousands of routers
- Hundreds of thousands of servers

Google’s Oregon DC

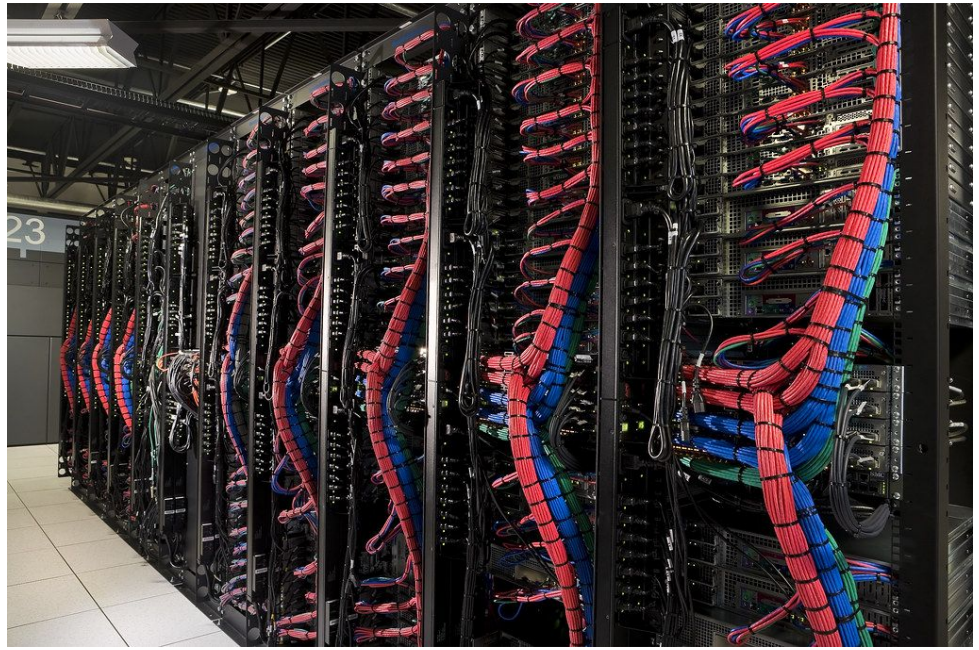
DCN topologies

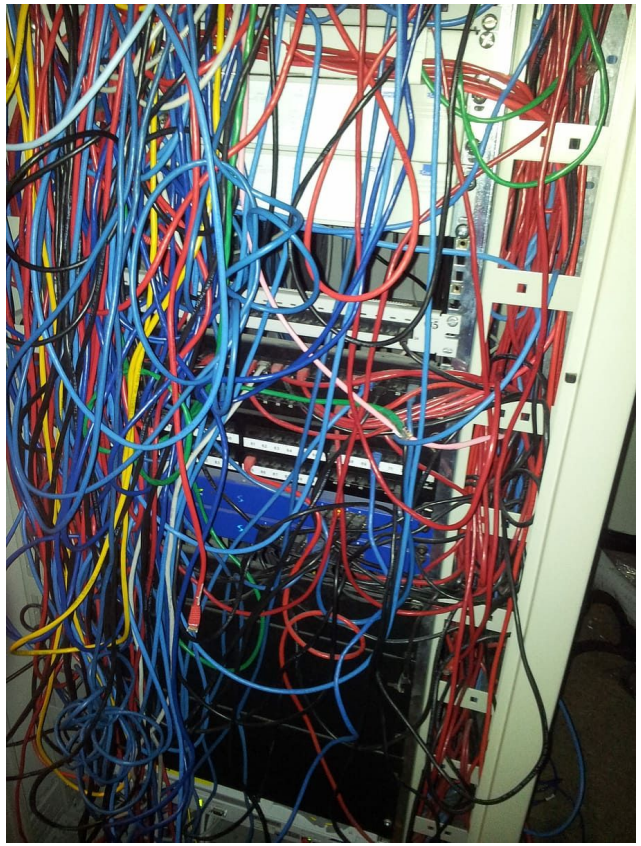
- Big iron Commodity switches





Under the hood





DCN topologies

- Big iron □ Commodity switches
- 1 Gbps □ 10 Gbps □ 40 Gbps □ 100 Gbps (soon)
- Copper □ Fiber

- Often run BGP because of scale and policy controls

Connecting to the cloud

- Public Internet
- VPN from your physical resources to the cloud
- BGP peering
  - E.g., Amazon Direct Connect

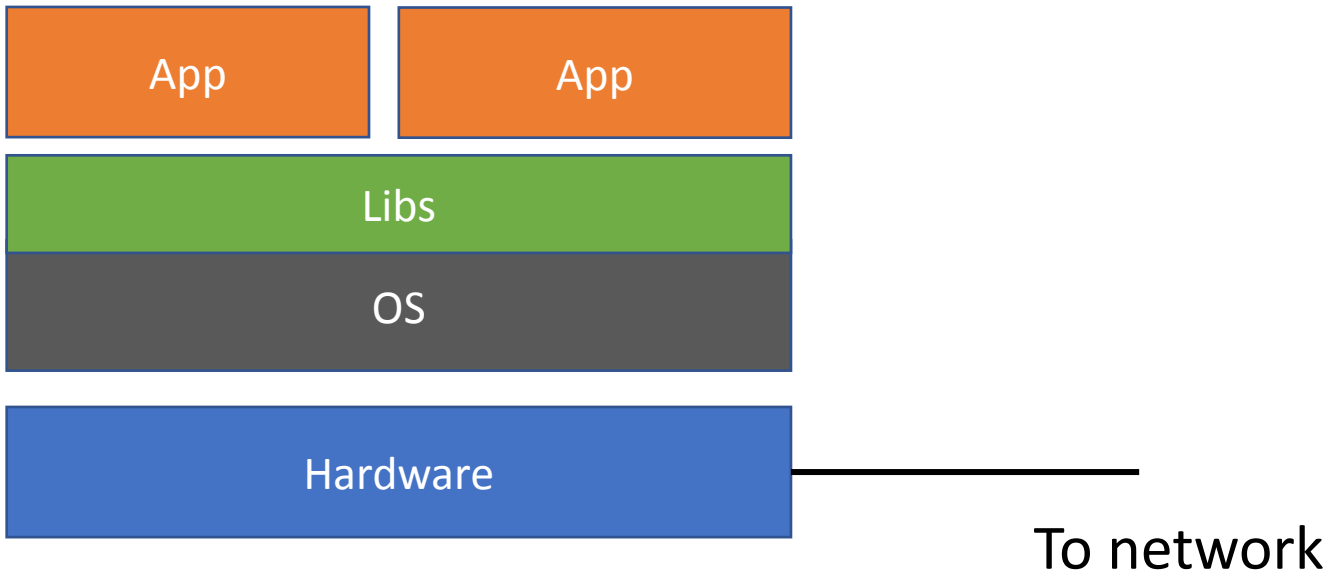
Using the cloud

- SaaS - use a software service (e.g., email)
- PaaS - use application building blocks (e.g., database)
- IaaS - launch VMs
- FaaS – run computations

- Build virtual networks
  - Provides the same abstraction as physical networks but with virtual devices

How about the servers?

Originally



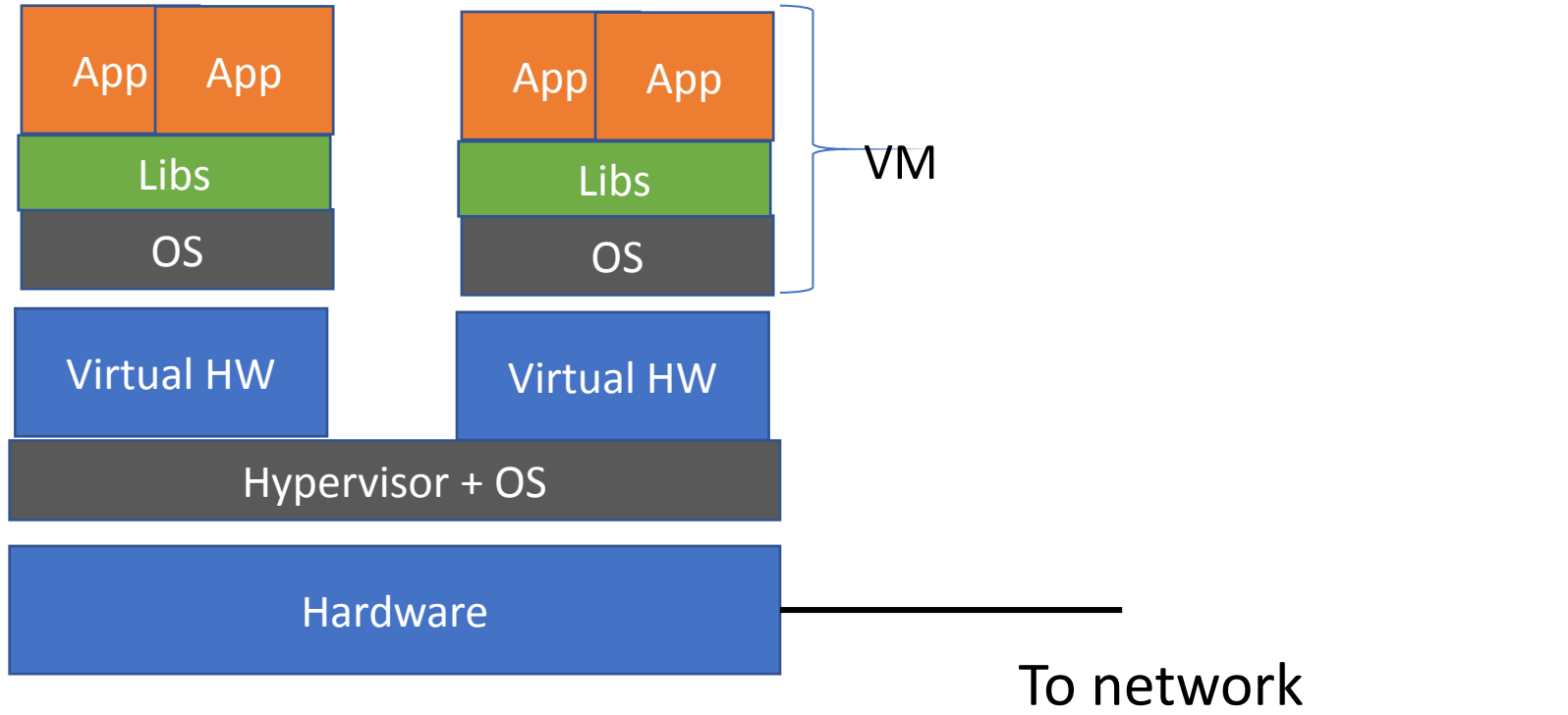
Then came virtual machines (VMs)

HW became too powerful

- Run multiple OSes on the same machine
- Cheaper that way

The hypervisor virtualizes the HW and “fools” the OS

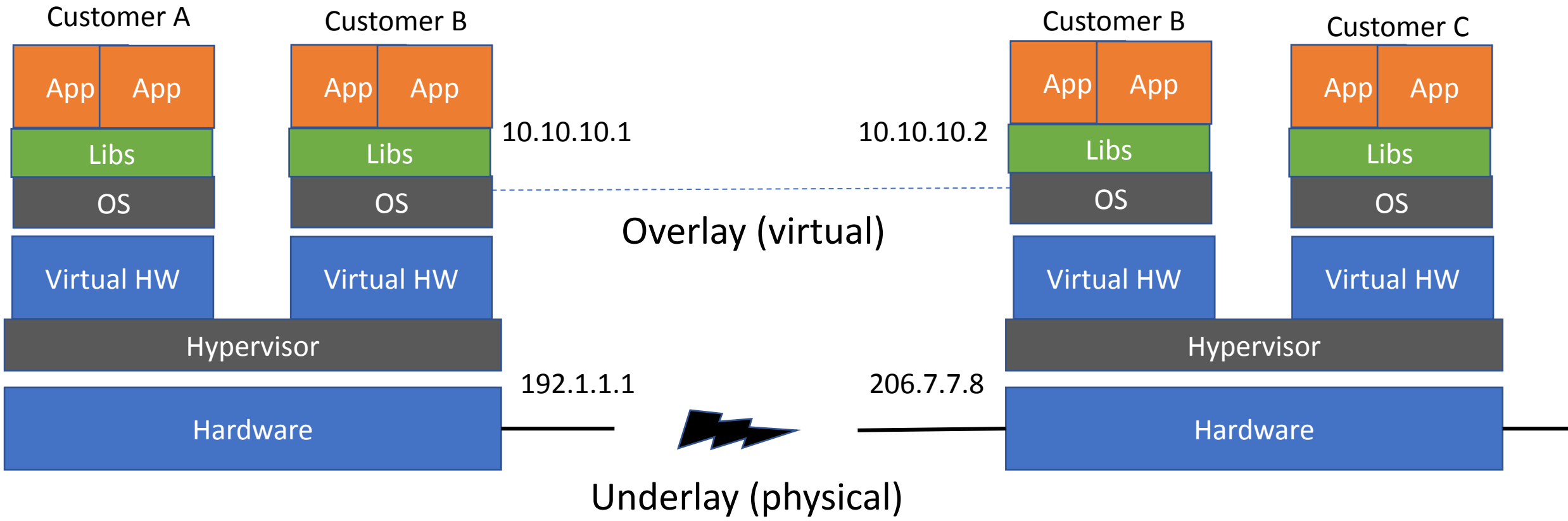
- Provides isolation



The network thinks multiple hosts are connected

The hypervisor acts as a hub for inter-VM traffic

VMs in the cloud

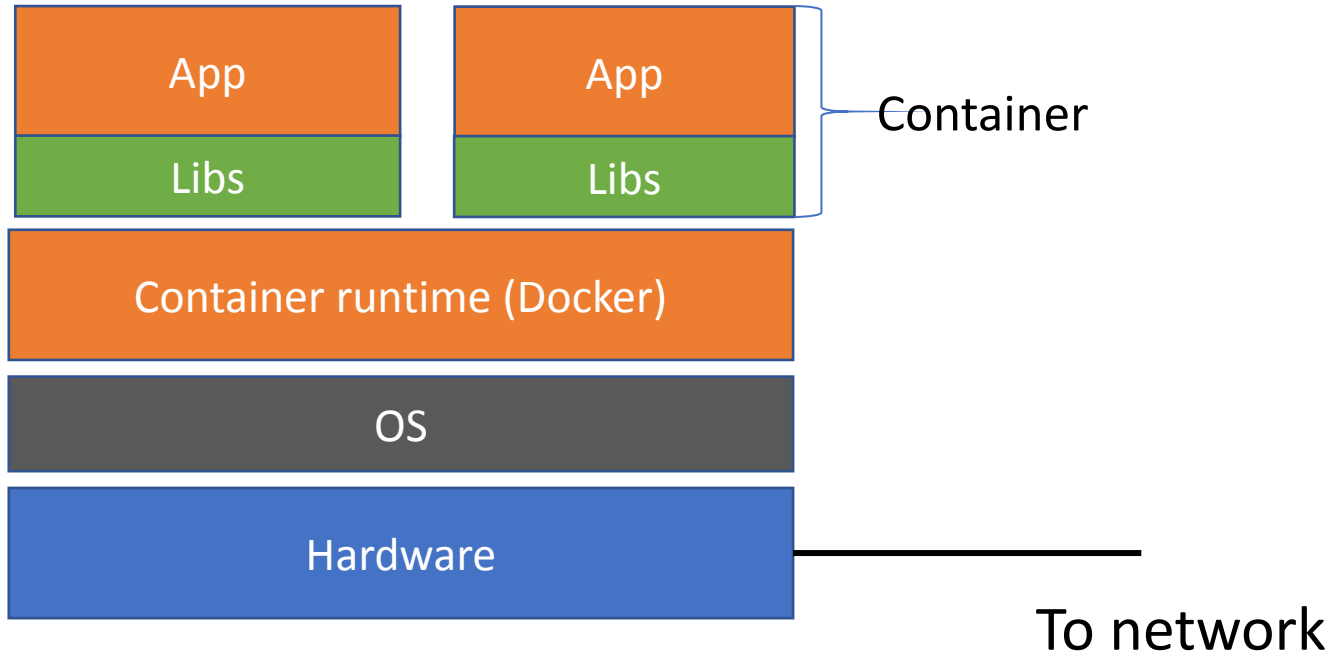


Forwarding between VMs involves a lookup from overlay address to underlay location

Enter containers

Lighter-weight virtualization than VMs
- Libraries, not the full OS

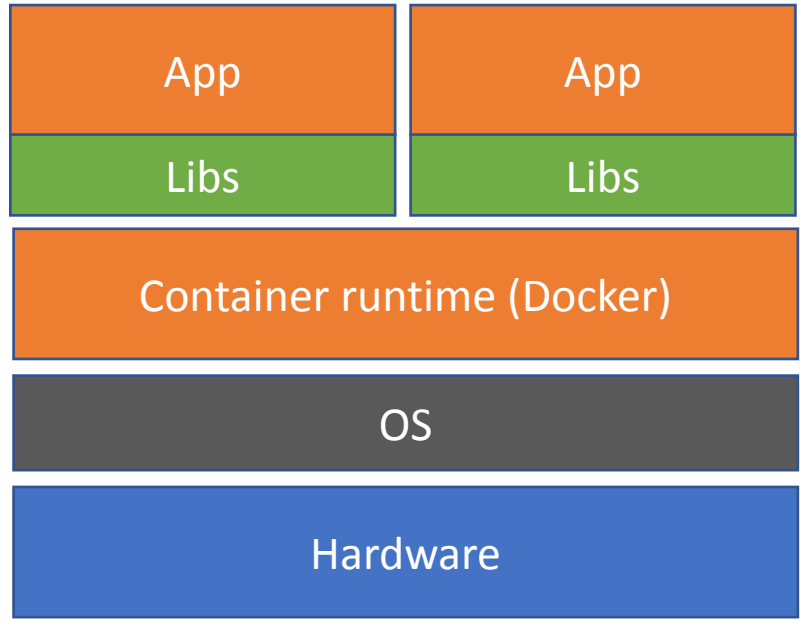
Better isolation and packaging than apps
- Bundle the library versions you need



Container networking

Connect containers to the outside world and to each other

- Port conflicts among containers and other apps running on the same host
- High performance between containers on the same host
- (Virtual) private network between related containers (service mesh)



Container networking: Host

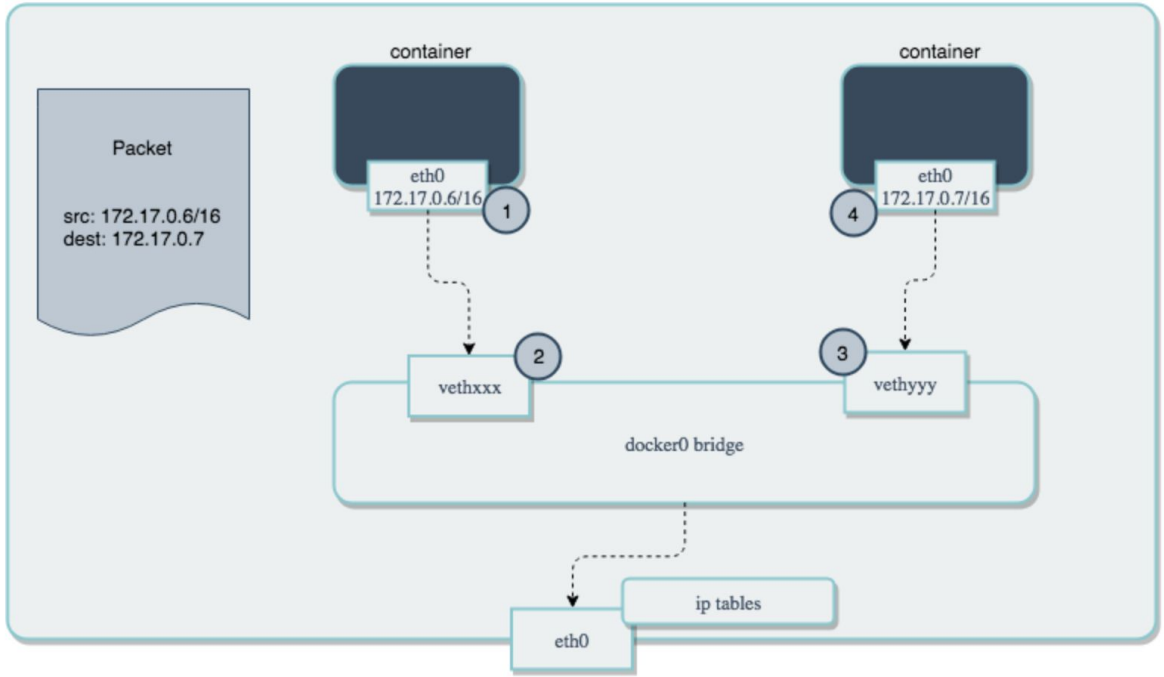
Containers share the IP address (and networking stack) of the host.

- Cannot handle port conflicts
- Minimal overhead

Container networking: Bridge

An internal network for containers on the same host.

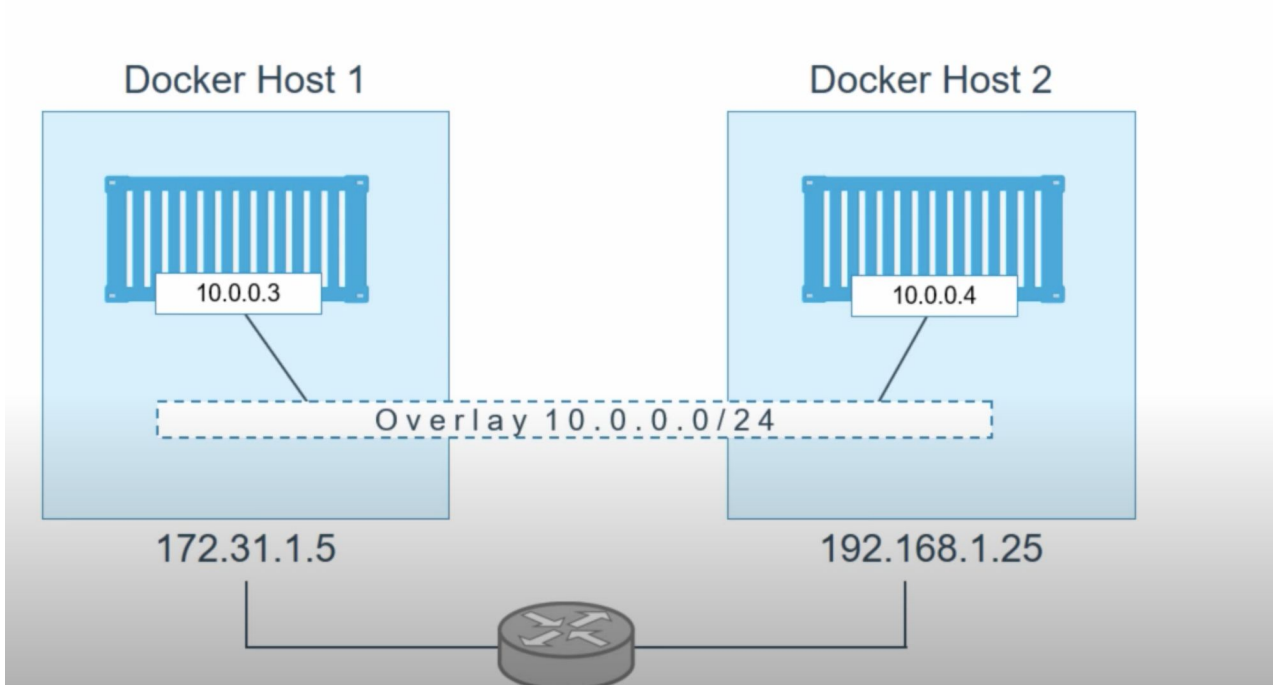
- Use NATs for outside world



Container networking: Overlay

Create a private network across containers on different hosts

- VXLAN is a common way to do that



Current trends

New hardware at the “bottom”

- FPGA, programmable NICs, TPUs, ASICs

New software systems in the “middle”

- Service meshes, ML frameworks

New applications and interfaces

- Serverless computing, edge computing

Over to Wenqing and Yanmeng