Cloud and containers

Ratul Mahajan
CSE 461
Image from Microsoft Azure
HUGE data centers (DCN)

- Thousands of routers
- Hundreds of thousands of servers

Connected by massive pipes

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

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

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

Google’s Oregon DC
DCN topologies

• Big iron $\rightarrow$ Commodity switches
Reality may look like either of these
DCN topologies

• Big iron $\rightarrow$ Commodity switches
• 1 Gbps $\rightarrow$ 10 Gbps $\rightarrow$ 40 Gbps $\rightarrow$ 100 Gbps $\rightarrow$ 400 Gbps (soon)
• Copper $\rightarrow$ Fiber
Oversubscription ratio

• Ratio of bisection bandwidth across layers of hierarchy
• Key design parameter that trades-off cost and performance
  • Higher oversubscription = lower cost but higher chance of congestion
DCN routing

• Spanning tree (L2) → OSPF/ISIS → BGP

• Each router acts as its own autonomous system (AS)
Backbone

• Provides global connectivity to DCs
Backbone

• Provides global connectivity to DCs

• May also have two backbones
  • A “public” backbone to connect to the outside world
  • A ”private” backbone for inter-DC connectivity

• Uses transcontinental and transoceanic fiber cables

• Routing: Distributed routing → SDN-based traffic engineering
SDN – Software Defined Networking

Decouple control plane (routing) and data plane (forwarding)

Control plane separation opens up lots of new opportunities
  • Traffic engineering in backbones (next)
  • Network virtualization (later)
What is in this box?
Originally

- Hardware
- OS
- Libs
- App
- App

To network
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
Container orchestration (Kubernetes)

Containers are wrapped in Pods which are run on a Cluster of Nodes

Pods implement a service

https://sensu.io/blog/how-kubernetes-works