CS236635
Network Functions Virtualization (NFV)

Class 5

Danny Raz
Last week

Cloud Computing and Data Centers
This week

• Networking
• Data Center Networks
• SDN – Software defined Networks
• Actual DC networks
Just to make sure we all remember?
A networking scenario

- **Web browsing**
  - A user is connected from home via her WiFi
  - Opens a browser
  - Goes to cnn.com and checks on Bin-Ladin

**2013 Top 10 Searches**

1. iPhone
2. Casey Anthony
3. Kim Kardashian
4. Katy Perry
5. Jennifer Lopez
6. Lindsay Lohan
7. "American Idol"
8. Jennifer Aniston
9. Japan Earthquake
10. Osama bin Laden

**2012 Top 10 Searches**

1. Whitney Houston
2. Gangnam Style
3. Hurricane Sandy
4. iPad 3
5. Diablo 3
6. Kate Middleton
7. Olympics 2012
8. Amanda Todd
9. Michael Clarke Duncan
10. BBB 12

**2013 Top 10 Searches for Sunday Oct 30 2016 in ISRAEL**


© ראובן כהן, הפקולטה למדעי המחשב, טכניון
What is going on?

- User is connected to the WiFi Access Point
- Wireless router is connecting to the service provider via ADSL
- Connection to remote server
- Downloading the file
What is really going on?

- **WiFi – 802.11 network**
- **ADSL connection**
- **Connection to remote server**
- **Downloading the file**

The application layer of the laptop computer senses the existence of the WiFi network and connects (or asks the user) to the network by setting the “right” parameters to the **physical** and the **MAC** layers.
What is really going on (2)?

- WiFi – 802.11 network
- ADSL connection
- Connection to remote server
- Downloading the file

The ADSL router connects over the PTN to the ISPs Point of Present at the Central Office of the Public Telephony Network Infrastructure Provider (BEZEK) and from there to the ISP’s Internet service provider RAP Remote Access Point and establishes a PPP Point to Point connection.
What is really going on (3)?

- Dial up networking
- User login
- Connection to remote server
- Downloading the file

A DHCP Dynamic Host Configuration Protocol message is sent to the DHCP server at the ISP, resulting in an assignment of a valid IP address to the ADSL router. Then an ARP Address Resolution Protocol is used to update the machine networking stack.
What is really going on (4)?

- Dial up networking
- User login
- Connection to remote server
- Downloading the file

A DNS *Domain Name Server* message is sent to the DNS server who may need update from other DNS servers, resulting in the translation of the server name to a valid IP address. A TCP *Transport Control Protocol* connection is established between the host and the remote server.
What is really going on (5)?

- Dial up networking
- User login
- Connection to remote server
- Downloading the file

An HTTP *Hypertext Text Transfer Protocol* message is sent to the remote server with the name of the wanted file *Get Bin-Laden.html* and the data in the file is sent back through the TCP connection to the host.
What is **really** going on (6)?

- Dial up networking
- User login
- Connection to remote server
- Downloading the file

What you see is **NOT** what you get. NAT *(Network Address Translation)* is used to change the IP address, firewalls are used to check the content of the packets, and proxy servers or even CDN *(Content Distribution)* are used as the source of the data.
Summary: Network layers

Layers:
- Application (7)
- Transport (4)
- Network (3)
- Link (2)
- Physical (1)

Presentation 6
Session 5
Summary: Internet Architecture

- Packet-switched datagram network
- IP is the glue
- Hourglass architecture
  - all hosts and routers run IP
- Stateless architecture
  - no per flow state inside network
A note on the use of these ppt slides:

We’re making these slides freely available to all (faculty, students, readers). They’re in PowerPoint form so you see the animations; and can add, modify, and delete slides (including this one) and slide content to suit your needs. They obviously represent a lot of work on our part. In return for use, we only ask the following:

- If you use these slides (e.g., in a class) that you mention their source (after all, we’d like people to use our book!)
- If you post any slides on a www site, that you note that they are adapted from (or perhaps identical to) our slides, and note our copyright of this material.

Thanks and enjoy! JFK/KWR

© All material copyright 1996-2012
J.F Kurose and K.W. Ross, All Rights Reserved
Four sources of packet delay

\[ d_{\text{nodal}} = d_{\text{proc}} + d_{\text{queue}} + d_{\text{trans}} + d_{\text{prop}} \]

\( d_{\text{proc}} \): nodal processing
- check bit errors
- determine output link
- typically < msec

\( d_{\text{queue}} \): queueing delay
- time waiting at output link for transmission
- depends on congestion level of router
Four sources of packet delay

- Transmission delay ($d_{trans}$):
  - $L$: packet length (bits)
  - $R$: link bandwidth (bps)
  - $d_{trans} = L/R$

- Propagation delay ($d_{prop}$):
  - $d$: length of physical link
  - $s$: propagation speed in medium (~$2 \times 10^8$ m/sec)
  - $d_{prop} = d/s$

$$d_{nodal} = d_{proc} + d_{queue} + d_{trans} + d_{prop}$$
Packet loss

- queue (aka buffer) preceding link in buffer has finite capacity
- packet arriving to full queue dropped (aka lost)
- lost packet may be retransmitted by previous node, by source end system, or not at all
Throughput: Internet scenario

- per-connection end-end throughput: \( \min(R_c, R_s, R/10) \)
- in practice: \( R_c \) or \( R_s \) is often bottleneck

10 connections (fairly) share backbone bottleneck link \( R \) bits/sec
Data center networks

- 10’s to 100’s of thousands of hosts, often closely coupled, in close proximity:
  - e-business (e.g. Amazon)
  - content-servers (e.g., YouTube, Akamai, Apple, Microsoft)
  - search engines, data mining (e.g., Google)

- challenges:
  - multiple applications, each serving massive numbers of clients
  - managing/balancing load, avoiding processing, networking, data bottlenecks

Inside a 40-ft Microsoft container, Chicago data center
Data center networks

**load balancer: application-layer routing**
- receives external client requests
- directs workload within data center
- returns results to external client (hiding data center internals from client)

---

Intro to NFV - Dr. Elisha Rosensweig
Data center networks

- rich interconnection among switches, racks:
  - increased throughput between racks (multiple routing paths possible)
  - increased reliability via redundancy
Traffic Categories

- North-South Traffic
- East-West Traffic
Traffic Categories

• In “classic” Datacenters, traffic is mostly N-S
  – E.g. hosting many web servers
• In emerging Datacenters, traffic is increasingly E-W
  – Large applications
  – Big data analysis
• Impact on placement: inter-VM distance becoming a factor
  – For N-S traffic, all VM locations are identical in principle
  – For E-W traffic, need to place VMs close together
    • “Close” = network distances \( \rightarrow \) depend on datacenter topology
Network Topologies

- Classic tree-based topology
Network Topologies

• Classic tree-based topology

• Pros
  – Simple design
  – Commodity Off the Shelf (COTS) switches

• Cons
  – No robustness
  – Inter-Rack distances – variable and therefore significant
  – ...

Intro to NFV - Dr. Elisha Rosensweig
Network Topologies

• Clos Topology
  – Named after a paper by Charles Clos on a better design for telephone switching boards
• Design properties
  – Every switch connected to all switches above and below it.
  – High robustness to link and switch failure
Network Topologies

• Fat-Tree concept
  – Links become thicker (more ports) as we go up the tree
  – Capacity up = capacity down
• We want this, but with using
• The same HW at each level...
Network Topologies

- Fat-Tree in practice (variation of Clos)
  - Each layer uses the same switch HW
  - Incoming BW per layer = outgoing BW per layer
Network Topologies

• Creating a Fat Tree with K-port switches
  – K pods, each containing two layers of $K/2$ switches
  – Each switch in lower layer connected to $K/2$ hosts
  – Remainder of ports – connected to $K/2$ switches in upper layer

• Variations exist with relaxed requirements
  – Up:down link ratio determines performance
  – 1:1 ratio considered *non-blocking*
  – Cost can be reduced at price of performance
Network Topologies in Network Virt.

- Fat trees are used in several network virtualization proposals
  - E.g. VL2 – see soon
- Leverage multiple paths between end-points
  - Robustness, of course, but more importantly...
  - ... multi-path routing
Multi-path routing

• In principle, can reach completely uniform link loads
  – Rely on Fat-Tree 1:1 ratio
  – In practice – more complex to achieve

• Graceful performance decrease
  – If a single link fails in a layer with N links, available rates decrease proportionately
    • 1/N less resources
A real example: Google’s Jupiter

SDN – Software defined networks (2012)

http://www.cs.princeton.edu/courses/archive/spr12/cos461/docs/lec23-measurement.ppt
A real example in WAN: Google’s B4

Questions?