CS 168
Spring 2025

Datacenters & SDN

Discussion 10

1 True/False

1.1 Achieving low latency for small jobs (mice) is critical for datacenter networks.

True: It is necessary to prevent large slow flows from starving small latency-sensitive tasks.

1.2 Clos topology enables the use of commodity hardware in datacenter networks.

True

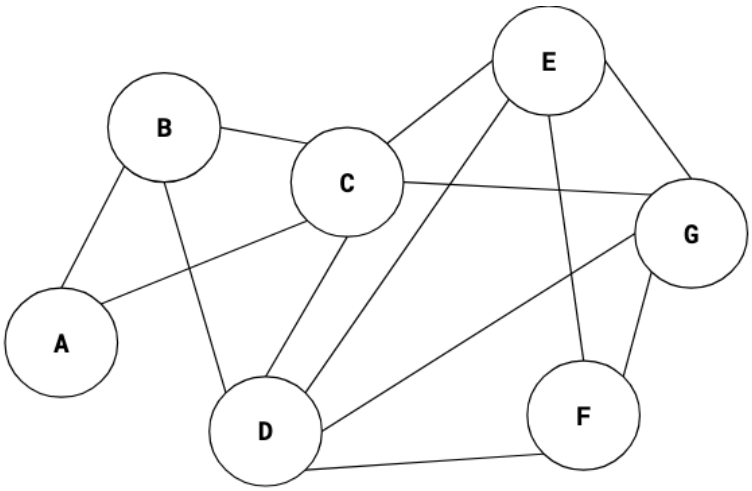
1.3 ECMP leads to reordering of packets in the same flow.

False. Packets from the same flow take the same path.

1.4 Over-subscription is the ratio of worst-case achievable aggregate bandwidth between end-hosts to total bisection bandwidth.

True

2 STP



The diagram above shows a L2 network where all links have equal cost. We want to apply the Spanning Tree Protocol. Choose the switch with the lowest alphabetical ID as the root, and when there are multiple shortest paths to the root, choose the path that uses the neighbor switch with the lowest alphabetical ID.

2.1 Let **XY** denote the link that connects switch("X") and switch("Y"). List all links that are a part of the spanning tree after the STP protocol converges.

AB, AC, BD, CE, CG, DF

2.2 Assume that at some point during STP, switch("F") believes that switch("A") is the root and switch("D") believes that switch("B") is the root. Is it possible that switch("C") also believes that switch("B") is the root?

No. Either switch("G") or switch("E") must have advertised to switch("F") that switch("A") was the root. Since switch("D") still thinks that switch("B") is the root, this advertisement of switch("A") as the root must have come from switch("C").

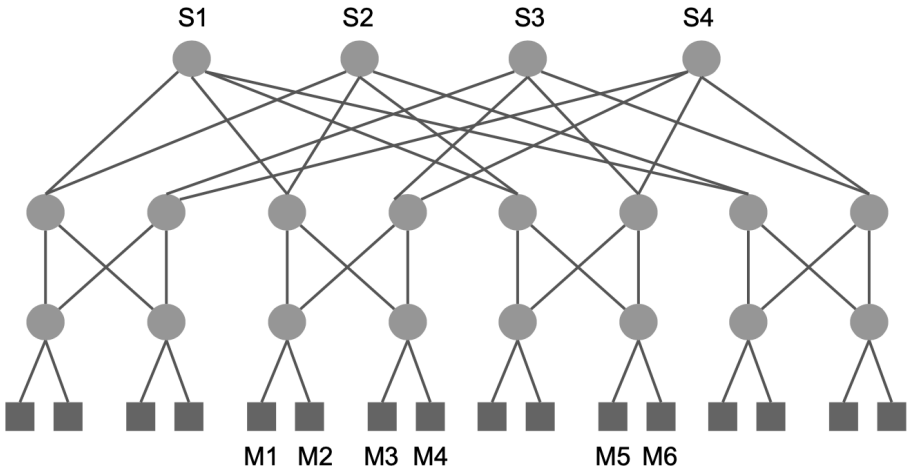
2.3 After the STP protocol has converged, switch("A") and all links directly connected to switch("A") go down. List all links that are a part of the new spanning tree after the STP protocol converges again.

BC, BD, CE, CG, DF

2.4 switch("A") and all links connected to switch("A") finally come back up, and the STP protocol begins to run again. However, switch("F") forgets to send any update messages to its neighbors. After the protocol converges, does the resulting set of links form a valid spanning tree?

Yes. In fact, the resulting set of links is the same as the one from problem 1. This is because switch("F") is not a part of the shortest path to switch("A") from any other switch.

3 Clos-based Topology & ECMP



Given the datacenter network above:

3.1 How many paths from an arbitrary server upward to an arbitrary root switch? Consider only minimal-length paths.

1.

3.2 How many paths can be set up between M1 to M3 in this topology? M1 to M5? Consider only minimal-length paths.

2 ; 4.

3.3 How many of the M1 to M5 paths go through switch S2?

1.

Consider the same network topology in Q3. Assume all links have capacity 10 Gbps and ECMP is used for multipath routing.

3.4 What is the bisection bandwidth of this topology?

80 Gbps.

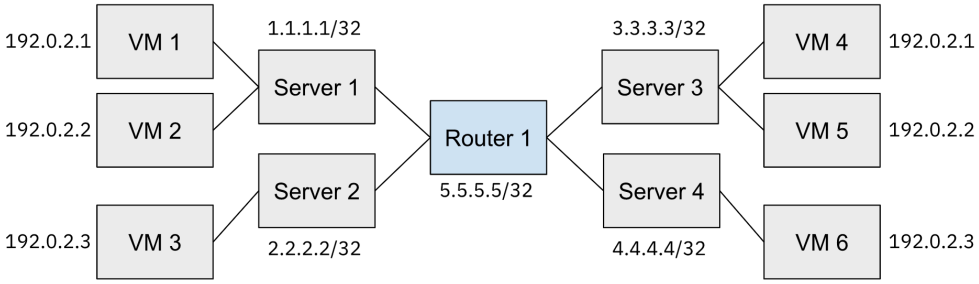
3.5 Assume ECMP makes the best possible hashing/load balancing. If switch S1 and S3 fail, what is the available aggregate bandwidth from each four-server Pod to any other such Pod (a Pod contains four servers closest to each other)?

20 Gbps.

3.6 Assume M1, M2 are sending flows to M5; and M3, M4 are sending flows to M6. TCP makes optimal use of the available bandwidth and ECMP makes the worst possible hashing. What is the approximate average data rate server M1 can send to server M5?

2.5 Gbps.

4 Encapsulation



4.1 If routing a packet from VM 1 to VM 6, what is the expected path?

S1-R1-S4

4.2 What encapsulation and decapsulation should the servers/router on the path from VM 1 to VM 6 perform (if any)? Fill in the following table with the destination address (both overlay and underlay) of the packets at each step.

Server/Router	Adds/Deletes Header	Overlay Address	Underlay Address
Server 1	Adds	192.0.2.3	4.4.4.4/32
Router 1	Neither	192.0.2.3	4.4.4.4/32
Server 4	Deletes	192.0.2.3	4.4.4.4/32

4.3 How many destinations does Router 1 have to keep track of without encapsulation? How about with encapsulation? Consider how this isolation impacts VM scalability in a data center!

Without = all of them = 10, With = just the underlay / physical addresses = 4