Advanced Computer Networks

Assignment 4: Datacenter routing

Submitted by Jinank Jain

Solution 1

Connecting tens of thousands of servers in a data center with a traditional Ethernet domain does not work because of the following issues:

- MAC Addressing: ?? What is the difference between this and ARP Request?

- Spanning trees: Problem with this approach is because we are using just one route and others routes which have the same cost are neglected which might cause congestion in the network which can be avoided by using other routes. Otherwise we are just wasting resources.

- ARP Request: This process involve a lot of broadcasting messages which is sort of unnecessary traffic which could be avoided. Another problem with increasing number of servers ARP cache size increases and it is not possible to cache all the results.

Solution 2

TRILL stands for TRansparent Interconnection of Lots of Links. TRILL applies network layer routing protocols to the link layer and – with knowledge of the entire network – uses that information to support Layer 2 multipathing. TRILL was introduced in the hope to overcome the shortcoming Spanning Tree Protocol (STP). STP, which was created to prevent bridge loops, only allows one path between network switches or ports. When a network segment goes down, an alternate path is chosen and this process can cause unacceptable delays in a data center network. TRILL is designed to address this problem by applying the Intermediate System-to-Intermediate System protocol (IS-IS) Layer 3 routing protocol to Layer 2 devices. This essentially allows Layer 2 devices to route Ethernet frame. With coming up of TRILL enterprises will be better able to migrate virtual machines (VMs) across the data center network. There will also be more bandwidth available for intensive applications like real-time communications (RTC) and for the transport of storage traffic across the Ethernet network. TRILL will make switches more cost-effective because it will allow enterprises to use more links in their data center network designs. It will also allow switches to load balance traffic over multiple Layer 2 links.

Solution 3

VLANs
VLANs stands for virtual LANs. A VLAN is a group of devices on one or more LANs that are configured to communicate as if they were attached to the same wire, when in fact they are located on a number of different LAN segments.

Why use VLANs?

- **Performance:** VLAN's can reduce the need to send such traffic to unnecessary destinations. For example, in a broadcast domain consisting of 10 users, if the broadcast traffic is intended only for 5 of the users, then placing those 5 users on a separate VLAN can reduce traffic.

  Compared to switches, routers require more processing of incoming traffic. As the volume of traffic passing through the routers increases, so does the latency in the routers, which results in reduced performance. The use of VLAN's reduces the number of routers needed, since VLAN's create broadcast domains using switches instead of routers.

- **Reduced Cost:** VLAN's can be used to create broadcast domains which eliminate the need for expensive routers.

- **Security:** Periodically, sensitive data may be broadcast on a network. In such cases, placing only those users who can have access to that data on a VLAN can reduce the chances of an outsider gaining access to the data. VLAN's can also be used to control broadcast domains, set up firewalls, restrict access, and inform the network manager of an intrusion

Problems that arise due to a large number of VLANs in a data center are as follows:
The routing design in conventional networks achieves scale by assigning servers topologically signicant IP addresses and dividing servers among VLANs. Such fragmentation of the address space limits the utility of virtual machines, which cannot migrate out of their original VLAN while keeping the same IP address

Solution 4

Part a:
In VL2, servers in each service are configured to believe that they all belong to the same IP subnet. Hence, when an application tries to communicate with another application, the networking stack on the host generated a broadcast ARP request for the destination host. The VL2 agent (some sort of hypervisor) running on top of these hosts intercepts this ARP requests and convert this to a unicast query to the VL2 directory system. The directory systems answers the query and this response is cached by the VL2 agent similar to host's ARP cache.

Part b:
VL2 uses ECMP and Valiant Load Balancing (VLB) to reduce the creation of hot-spots in the network. VL2 agent uses encapsulation to implement VLB by sending traffic through a randomly-chosen Intermediate switch. The packet is first delivered to one of the Intermediate switches, decapsulated by the switch, delivered to the ToR's LA, decapsulated again, and finally sent to the destination.

But there is one problem with approach so whenever an intermediate switch's availability changes due to switch or link failure a large number VL2 agents needs to get an update about this information. In order to solve this problem VL2 assigns the same LA address to all the intermediate switches and directory returns this anycast address to agents upon lookup.

Solution 5

Part a:

ECMP stands for equal-cost multipath. ECMP arises when the routing table contains multiple next-hop addresses for the same destination with equal cost. (Routes of equal cost have the same preference and metric values.) It can substantially increase bandwidth by load-balancing traffic over multiple paths. Thus this approach will try to reduce the congestion by utilizing the resources upto maximum extend.

Part b:
ECMP does not do packet-level randomization because it would packet out of order of transmission which is not good for protocols like TCP.

Solution 6

In ECMP one strategy that people follow is hashing the flow parameter to decide which path to take and if it following this approach then Flow A and Flow B would not be mapped to same path until or unless there is some hash collision or we are following some different strategy.