Introduction

A data center is a centralized repository, either physical or virtual, for the storage, management, and dissemination of data.

The boom of data centers came during the dot-com bubble and with increase in cloud services.

Companies needed fast Internet connectivity and nonstop operation to deploy systems and establish a presence on the Internet.

Problem

Conventional architectures rely on tree-like network configurations built from high-cost hardware.

Limited server-to-server capacity:
Capacity between branches of the tree are over subscribed.

Poor reliability and utilization:
Switch runs with 50% of its maximum utilization to provide fail overs. In layer2 spanning tree protocol only allows single path to be used and in layer 3 ECMP allows multiple paths but the topology offers two paths at layer 3.
Problem

Poor Utilization:
When a service suffers traffic flood, all the servers sharing that particular sub tree suffers from the flood

Fragmentation of resources:
Servers are divided in to VLAN's and assigned IP addresses according to the tree topology. Virtual machines cannot move out of VLAN and can cause configuration burden

Solution Approach (VL2)

Maintain an illusion that a service is assigned all the servers and those servers are connected using a virtual switch. Uniform high capacity, Performance isolation and reconfigurable IP addresses.

Design Decisions

VL2 uses Clos topology (multi-staged switching topology) combined with ECMP (layer 3 multipath routing) to provide extensive path diversity among servers

To cope with high volatility in traffic, VL2 uses Valiant load balancing that picks a random path for traffic flow between one server to another

These two design decisions allow uniform capacity and performance isolation
CLOS Topology in VL2

Initial Data Center Traffic Analysis

- Key Observations:
  - There is four times as data entering the data center as leaving.
  - The demand for bandwidth between servers inside a data center is growing faster than the demand for bandwidth to external hosts.
  - The network is a bottleneck to computation!!!

Initial Data Center Traffic Analysis

- Distribution of Flow sizes:
  - Majority of flows are small (a few KB)
  - For large flows the byte size hovers around 100MB due to data center optimization
  - A server has ten flows 50% of time and 80 flows 5% of time

Initial Data Center Traffic Analysis

- Highly Variable data traffic
- Unstable and poor traffic patterns
- Due to the use of randomness to improve performance of data center applications
- VL2 moves from 1:1 redundancy to n:m redundancy
Digging Deeper VL2

VL2 copes with the high divergence and unpredictability of data-center traffic matrices by using Valiant Load Balancing.

VL2 is based on IP routing and forwarding technologies that are already available in commodity switches: link-state routing, equal-cost multi-path (ECMP) forwarding.

Separating Names from Locators and VL2 agent on the end systems.

VL2 Addressing and Routing

Packet Forwarding – VL2 agent traps and does AA to LA using directory service.

Address Resolution – VL2 agent traps ARP packet and does AA to LA mapping and caches it.

Access control via directory service.

Switch/link failures are handled by assigning same LA for the intermediate switches.

VL2 addressing and routing

VL2 Directory Lookup System

Figure 7: VL2 Directory System Architecture
Evaluation

Aggregate goodput in 2.7 TB shuffle w 75 servers (94% utilization)

Figure 9: Aggregate goodput during a 2.7 TB shuffle among 75 servers.

Fairness

Flows split from aggregate switches to intermediate switches and isolation of services

Figure 10: Fairness measures.

Directory Service Performance

A directory system with three directory servers handles 50K lookups/sec within 10ms (ARP requests)

Up to 40K lookups/sec, the system offers a median response time of < 1ms

Maximum lookup rates increases linearly with the number of directory servers

Observations

Pros
Adaptive routing using Valiant Load Balancing
Fine grain Access Control using Directory Service
End to End argument valid (No modification of switches needed)

Cons
Agent on end systems needs configuration and setup
Higher power consumption than traditional
Centralized lookup and updates even though they are using two tier lookup service