Jupiter Rising: A Decade of Clos Topologies and Centralized Control in Googles Datacenter Network

Paper Note
Junzhi Gong

September 22, 2018

1 Backgrounds

- Datacenter networks are critical to delivering web services, modern storage infrastructure, and are a key enabler for cloud computing. The size of datacenter networks is increasing rapidly, due to the larger dataset sizes, higher quality of Web services, and co-resident applications.
- Previously, datacenter networks have high cost and operational complexity, and they are also built targetting in highest levels of availability.
- Clos topologies can scale to nearly arbitrary size by adding stages to the topology, principally limited by failure domain considerations and control plane scalability.

2 Target

This paper describes our experience with building five generations of custom data center network hardware and software leveraging commodity hardware components, while addressing the control and management requirements introduced by our approach.

3 Solution

- **Use Clos topologies.** To support graceful fault tolerance, increase the scale/bisection of our datacenter networks, and accommodate lower radix switches, we adopted Clos topologies for our datacenters.
- **Use Merchant silicon.** To keep pace with server bandwidth demands which scale with cores per server and Moores Law, we emphasized bandwidth density and frequent refresh cycles, so they use the latest generation of switches.

- **Use centralized control protocols.** To control this complexity, the authors observed that individual datacenter switches played a predetermined forwarding role based on the cluster plan.
- **The five generations of topologies.** Google has experienced five typical generations of topologies, including Firehose 1.0, Firehose 1.1, Watchtower, Saturn, and Jupiter. The second one is the start of Clos topology, and Jupiter is a 40G datacenter-scale fabric.