

Paper review ^[1]

Yu SU 55323707

1. The Problem

Agility is an important property for the network, because the agility could help data center operator meet the fluctuating demands of individual services from a large shared server pool, which could increase the utilization and reduce the costs of the system. However, the agility is not realized in existed system because of many architecture restrictions. The paper proposed a practical network architecture called VL2, which supports huge data centers with uniform high capacity between servers, performance isolation between services and ethernet layer-2 semantics.

2. Challenge

The main challenge for realizing the agility is to overcome the limitation of the existed network architecture. One of the limitations is the capacity between the servers they interconnect is not enough because of the high cost of the hardware. Second, there is no isolation protection, which means if one service gets in traffic, the around services will be influenced either. Last is the fragmentation mechanism of address space will limit the utilization of the virtual machine. Therefore, there are three problems for the network, uniform high capacity, performance isolation and layer-2 semantics.

3. Key Insight

The paper designed, implemented and evaluated VL2, a network architecture for data centers, which explore a new split in the responsibilities between host and networks. In detail, there are four main contributions in the paper.

First, the paper first proposed the traffic patterns in a production data center and the enormous volatility in the traffic.

Second, the paper designed, built and deployed the VL2 in a 90 servers cluster and verified its capacity and performance isolation.

Third, the paper applied the Valiant Load Balancing in a new context and verified it could smooth utilization while eliminating persistent congestion.

Last, the paper justified the meaning of the VL2 by comparing the cost of it with the equivalent existing designs.

4. Limitation

The first limitation is the optimality of VLB, the evaluations on actual data center workloads show that the simplicity and universality of VLB cost relatively little capacity when compared to much more complex traffic engineering schemes.

The other limitation is the cost and scale. Under different situation, the cost of the VL2 is not stable, which means it could cost more than existing designs in some situations.

5. Future Work

The paper needs to fix the limitation above in the future. And what the paper provided is only a prototype, so there needs more optimization on it and adjustment for the practical network circumstance.

[1] G. Albert, R. James, J. Bavendu, K. Srikanth, K. Changhoon, L. Parantap (2009). VL2: A Scalable and Flexible Data Center Network. *Communications of the ACM* 54(3):95-104. October 2009.