Introduction to Azure Networking

Introduction

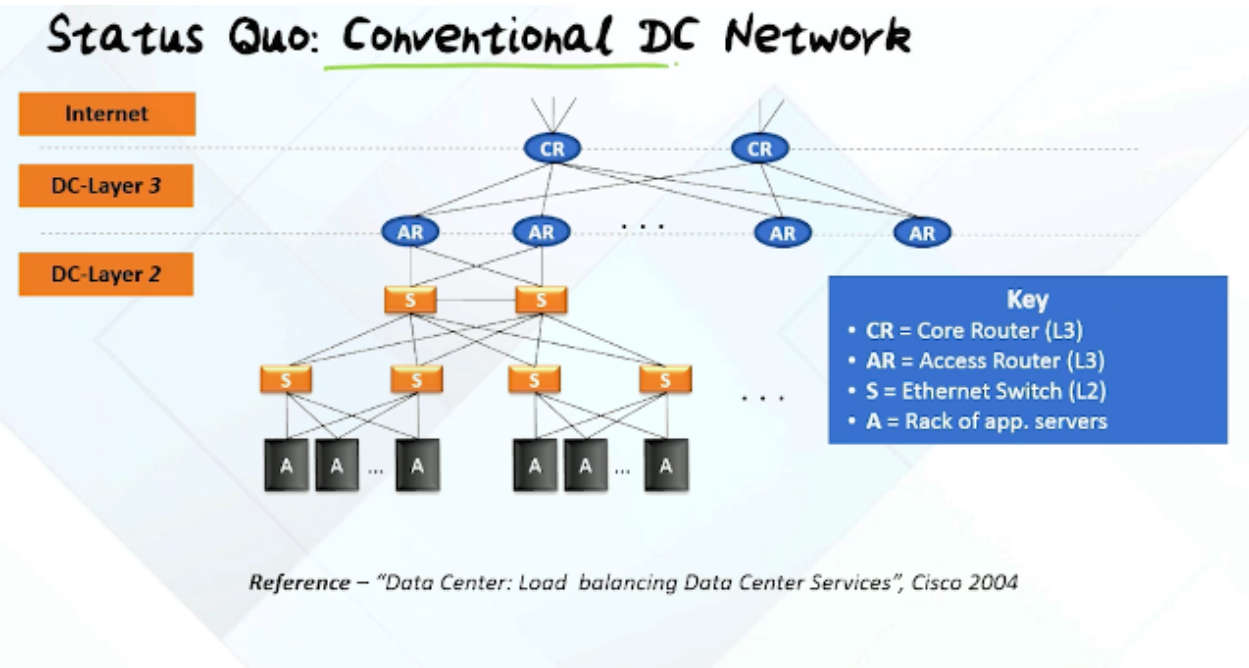
1. Overview of data center networks
2. Details of VL2 switch

Overview of Azure Data Center Networks

1. Applications desire layer-2 semantics for server-server communication
    - Communication latency and throughput bound by the network interface speeds of the source and destination
2. Applications need elasticity of resource allocation
    - Grow and shrink computational resources based on need
    - Do not suffer network performance loss for such flexibility in resource allocation
3. Agility of data center
    - Ability to allocate resources on demand to meet the dynamic application needs
    - Ensure network performance scaling in the presence of such dynamic allocation
4. Limitations to agility in data centers
    - Insufficient network capacity for connecting the servers
    - Conventional network architectures
        - Tree topology using high-cost hardware (links over-subscribed as we reach higher levels of the tree)
        - Fragments server pool (network congestion and server hotspots)
        - Network flooding of one service affects others
        - Prevents easy relocation of services when IP addresses are statically bound to servers
5. VL2 Solution
    - Illusion of Virtual Layer 2
        - Appears as though all servers for a given service connected to one another via non-interfering Ethernet switch
        - Scaling up or down of servers for a service maintains this illusion in tact
6. VL2 Objectives
    - Uniform Capacity
        - Independent of topology, server-server communication limited only by NICs connected to the servers
        - Assigning servers to services independent of network topology
    - Performance Isolation
        - Traffic of one service does not affect others
    - Flexible assignment of IP addresses to Ethernet ports to support server mobility commensurate with service requirements

Azure VL2 Switch

1. Tenets of Cloud-Service Data Center
    - Agility: Assign any servers to any services
        - Boosts cloud utilization
    - Scaling out: Use large pools of commodities
        - Achieves reliability, performance, low cost
2. What is VL2?
    - VL2 is the first data center network that enables agility in a scaled-out fashion
    - Why is agility important?
        - Today's data center network inhibits the deployment of other technical advances toward agility
        - With VL2, cloud data centers can enjoy agility in full
3. Status Quo: Conventional Data Center Network



Conventional Data Center Network

4. Conventional Data Center Network Problems
   - Dependence on high-cost proprietary routers
   - Extremely limited server-to-server capacity
     - Two servers on different subnets will experience significantly worse communication overhead compared to servers that are colocated
     - Applications can interfere with each other due to traffic patterns
   - Resource fragmentation significantly lowers cloud utilization and cost efficiency

Data Center Networks Challenges and Opportunities

1. Challenges
   - Instrumented a large cluster used for data mining and identified distinctive traffic patterns
   - Traffic patterns are highly volatile
     - A large number of distinctive patterns even in a day
   - Traffic patterns are unpredictable
     - Correlation between patterns very weak
     - Optimization should be done frequently and rapidly
2. Opportunities
   - Data center controller knows everything about hosts
   - Host OS's are easily customizable
   - Probabilistic flow distribution would work well enough, because...
     - Flows are numerous and not huge - no elephants
     - Commodity switch-to-switch links are substantially thicker (~10x) than the maximum thickness of a flow
   - Data center network can be made simple

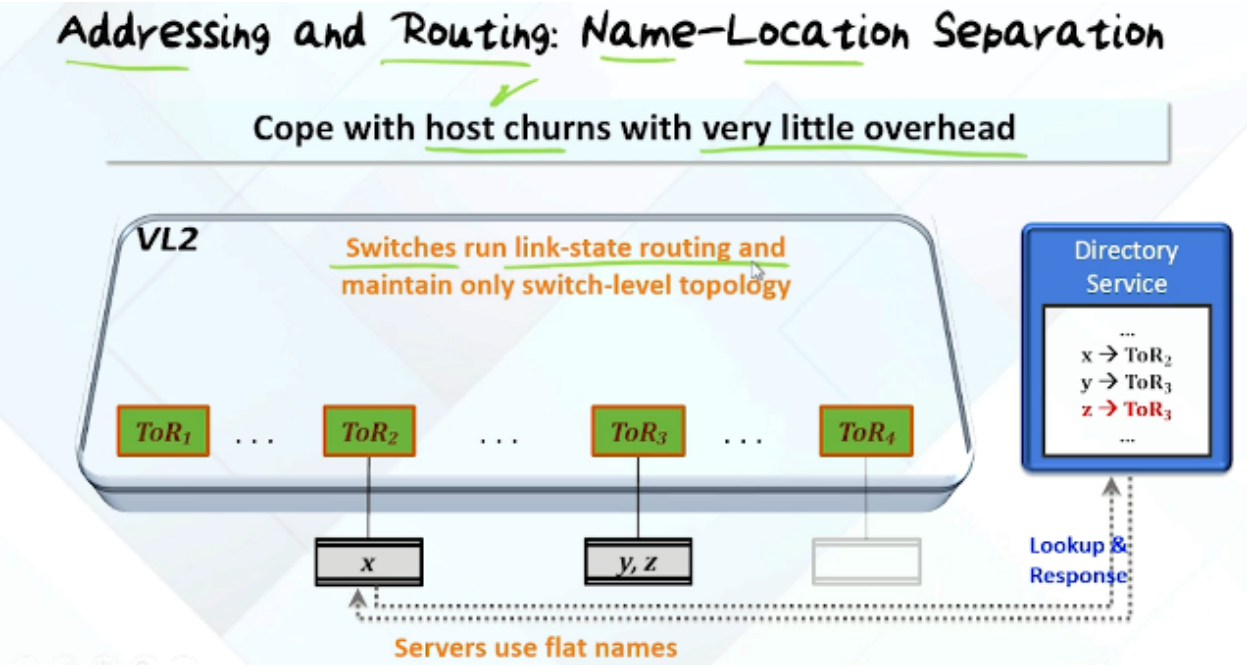
Switch Details

1. All we need is a huge L2 switch, or an abstraction of one
   - Should provide the following:

  - L2 semantics
  - Uniform high capacity
  - Performance isolation
2. Specific Objectives and Solutions
   - L2 semantics
     - Approach: Employ flat addressing
     - Solution: Name-location separation and resolution service
   - Uniform high capacity between servers
     - Approach: Guarantee bandwidth for hose-model traffic
     - Solution: Flow-based random traffic indirection (Valiant load balancing)
   - Performance isolation
     - Approach: Enforce hose model using existing mechanisms only
     - Solution: TCP

VL2 Addressing and Routing

1. Name-Location Separation
   - Cope with host churns with very little overhead
   - Switches run link-state routing and maintain only switch-level topology
     - Allows data centers to use low-cost switches
     - Protects network and hosts from host-state churn
     - Obviates host and switch reconfiguration
   - Directory service maintains server/switch mapping
2. Example Topology: Clos Network
   - Offer huge aggregate capacity and multiple paths at modest cost



Addressing and Routing

VL2 Traffic Forwarding

1. Traffic Forwarding: Random Indirection
   - Cope with arbitrary traffic flows with very little overhead

    - Designate higher-level switches randomly to prevent congestion
  - ECMP + IP Anycast
    - Harness huge bisection bandwidth
    - Obviate esoteric traffic engineering or optimization
    - Ensure robustness to failures
    - Work with switch mechanisms available today
2. Does VL2 Ensure Uniform High Capacity?
  - How "high" and "uniform" can it get?
    - Performed all-to-all data shuffle tests, then measured aggregate and per-flow goodput
    - Goodput efficiency: 94%
    - Fairness between flows: 0.995
3. VL2 Conclusion
  - VL2 achieves agility at scale via
    - L2 semantics
    - Uniform high capacity between servers
    - Performance isolation between services
  - Lessons
    - Randomization can tame volatility
    - Add functionality where you have control

Conclusion

1. VL2 switch design is the foundation for how Azure's data center network is architected today
  - Talk from co-author of paper available as supplemental material