Automatic Cut-Through Paths

System and Network Engineering
Research Project 2
Class 2005 - 2006

Lourens Bordewijk
René Jorissen
Agenda

- AMS-IX network
- Problem definition
- Cut-through path
- RBridges
- Additional solution
- Conclusion
AMS-IX network (2)

- VLANs
  - Internet, multicast...
  - Quarantine

- Virtual Switch Redundancy Protocol
  - Foundry Networks proprietary
AMS-IX network (3)

› Customer statistics
  - Number of customers: 240
  - Number of routers: 390

› Traffic statistics
  - Average load: 90 Gb/s
  - Peak load: 150 Gb/s
Problem definition (1)

- Cut-through switching
- Layer two network
  - Loops
  - Broadcast
  - Spanning tree

Avoidance of Network Loops
Problem definition (2)

- Management
  - Thresholds
  - Sampling
  - Computation
  - Configuration
Cut-through path (1)

Why

- Lessen load on core switches
- Lessen traffic congestion
- Involves less jitter
- More bandwidth capacity
- More efficient traffic streams
Cut-through path (2)

- How
  - Sampling process
  - Filtering process
  - Trigger
  - Control server architecture
sFlow (1)

- Packet-based sampling technology
- From layers two till seven
- Provide information about switch ports, MAC addresses, VLANs, IP addresses and ICMP/TCP/UDP/AS-based information
sFlow (2)

- Supported by the Foundry switches
- Inspecting all packets costs extensive CPU power
- Can handle volume of high speed backbone links
- Provides a result with quantifiable accuracy
Resource information

- SNMP
  - Data transfer
  - CPU utilization, memory utilization
  - CAM statistics and process utilization
  - Logging
Sampling process

- When
  - A load of more than 90% for 30 minutes on a certain switch port
  - A constant data flow of more than 4 Gb/s for 30 minutes on a certain switch port
  - Determine the exact values after further research
Filtering process (1)

- How

- Starts when first sFlow data from a switch is collected

<table>
<thead>
<tr>
<th>SSwitch</th>
<th>DSwitch</th>
<th>VLAN</th>
<th>SPort</th>
<th>DPort</th>
<th>SMAC</th>
<th>DMAC</th>
<th>Count</th>
<th>Priority</th>
<th>STime</th>
<th>TTL</th>
</tr>
</thead>
</table>

Filtering process (2)

- Sort flows based on priority and packet count
  - Per DSwitch, than per SPort & SSwitch and than per VLAN
  - “Priority & packet count” must reach threshold before the TTL ends, (decisions taken after TTL period)
Filtering process (3)

- Combine the total flows per SPorts from the SSwitch
  - Calculate average
Filtering process (4)

Example:
Bandwidth prediction

- Traffic cycle

- Several algorithms for bandwidth prediction

- Forecast traffic flows with long lifetime

  - Use for setting priority
Cut-through creation (1)

- Huge amount of traffic is flowing between two customers
- Flow triggers cut-through path creation
- Create a new VLAN
- Photonic switch connects two edges
Cut-through creation (2)

- Create MAC filter based on destination MAC addresses

- Configure an egress filter on switch port
  - Encapsulate Internet VLAN tagged frames with the new VLAN tag
  - 802.1ad (Provider Bridges)
Cut-through creation (3)

How (3)
Control server architecture (1)

- Why
  - To collect data
  - Consider the priorities
  - Makes calculations
  - Automatically configures a dynamic cut-through path
  - To manage all resources
Control server architecture (2)

How

- Separate networks, one private

- Control process must be physically separated from the filtering process

- Validate all configuration steps (roll back)

- Control server should be redundantly for failover in the event of a system failure
Control server architecture (3)
RBRidges (1)

- Transparent Interconnection of Lots of Links (TRILL)

- Inefficient paths
- Convergence
- Backup paths
- Ethernet extensions

- Services
- Loop mitigation
- VLAN
- Security
RBRidges (2)

- Advantages of routers and bridges
- “Routing” on layer two
- Full mesh possible
- Ethernet frame encapsulation
- Hardware or firmware
- Approximately 2 years
R Bridges (3)

- General operations
  - Peer and topology discovery
  - Designated R Bridge election
  - Ingress R Bridge Tree computation
  - Link-state routing
  - Advertisements
RBriges (4)

- Ingress / Egress RBridge
  - Encapsulation
  - Decapsulation

<table>
<thead>
<tr>
<th>outer header</th>
<th>shim header</th>
<th>original frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>DestAddr RBr.</td>
<td>TTL</td>
<td>egress/ingress RBr.</td>
</tr>
<tr>
<td>SrcAddr RBr.</td>
<td>egress/ingress RBr.</td>
<td>original frame</td>
</tr>
</tbody>
</table>
RBriges (5)

- Hop-by-hop vs. edge-by-edge
  - Different headers

- Forwarding
  - Unicast
  - Broadcast
  - Multicast
Additional solution (1)

- Two uplinks
  - Secondary path
  - Adding customer routing tables
Conclusion (1)

- Capacity problem (approx. in 1,5 year), best solution?
  - 100 Gb/s capable switch ports

- RBridges
  - Full mesh layer two topology
  - Uses all paths efficiently
  - No STP and VSRP needed
  - 1 to 2 years
Conclusion (2)

- Interim solution could be the use of VLANs
  - Automatically configured cut-through VLANs, when specific traffic flow reaches threshold
  - Control architecture takes care of the sampling, filtering, computation and triggering process
Future

- Further research to determine thresholds
- Development software
- Build test environment
- Other technologies
  - GMPLS
    - Looks like a solution
    - No hardware support
Questions

Thanks for the attention