Improving scalability of the AMS-IX network

Stéfan Deelen & Reinier Schoof

February 6, 2008
Improving scalability of the AMS-IX network

Introduction
  Amsterdam Internet Exchange
  AMS-IX Topology
  Scalability definition
  Network efficiency

Cut-through paths

Traffic Engineering
  Multiprotocol Label Switching
  Provider Backbone Bridging

Conclusion
Amsterdam Internet Exchange

World’s biggest IX

293 members (05 Feb 2008)

Peaks of over 400 Gb/s

Copyright (c) 2008 AMS-IX B.V.  [updated: 05-Feb-2008 18:52:41 +0100]
AMS-IX Topology

Improving scalability of the AMS-IX network
AMS-IX fail-over methods

- Completely redundant network
- Virtual Switch Redundancy Protocol
- Fail-over in approx. 300ms
The ultimate level of scalability for the AMS-IX would be to facilitate unlimited traffic exchange for an unlimited amount of members, with the only limits on throughput being either the capacity of the sending or the receiving party.

"AMS-IX is dedicated to offer non-blocking peering services over Ethernet infrastructure."
Bottleneck

75% to 80% of traffic through core
Problem definition

How can the scalability of the AMS-IX network be improved?

1. What other relevant researches have been conducted previously preceded and what is their relevance to the current research project?

2. Which potential solutions can be found to address AMS-IXs problem in scalability and what are their respective cons and pros?

3. Is there a solution which deserves preference?

4. How could this solution be deployed on the AMS-IX network?
Non approaches

Approaches we did not prefer:

- **Up-scaling core switch:**
  - No such hardware
  - Still hardware dependent

- **Applying redundant links:**
  - Loops in network
  - Need for STP
  - No balancing over links

- **’flow based forwarding’**: 
  - Vendor specific feature
  - Not high performance
Full mesh of all customers?

All customers directly connected to each peering partner.

Pros:
- Most efficient offloading
- Fully decentralized

Cons:
- Not scalable
- Not transparent
- High layer 1 costs
- Very high port-cost
Full mesh of all edges?

All edges directly connected to each other

Pros:
- Efficient offloading
- Transparent from customers point of view

Cons:
- High port-cost
- High layer 1 costs
Full mesh
Control Architecture

Provides dynamic network usage

![Diagram of Control Architecture]

- NIKHEF
- euNetworks
- Control Capacity
- Control Server architecture
- SARA
- GlobalSwitch
- Telecity
Increase locally switched traffic
Increase locally switched traffic
Static CAM entry in edge switch

Applying Static CAM entries: Functional, but not a very efficient CTP usage, unless you take congestion risks.
PBB/MPLS Overview
MPLS label

20 Bits

Label

3 Bits

Exp.

1 Bit

S

8 Bits

TTL

Stéfan Deelen & Reinier Schoof

Improving scalability of the AMS-IX network
VPLS label

Payload Ethernet Frame

Payload Original Ethernet Frame

Mpls Label 1 (4 bytes)

Mpls Label 2 (4 bytes)
MPLS/VPLS Label


Compared to a regular Q-tagged Ethernet header of 26 bytes, MPLS/VPLS adds 84% protocol overhead.
Generalized Multiprotocol Label Switching

- MPLS traffic engineering extended with optical cross connect control
- Multi-layer control plane
Provider Backbone Bridging (PBB)

- IEEE 802.1ah, but not standardized
- Encapsulates Ethernet frames in Ethernet headers (MAC-in-MAC)
- Forwarding method untouched for non-PBB devices
- Flow based traffic engineering
PBB Frame

Regular Ethernet frame compared to PBB encapsulated frame
Provider Backbone Briding - Traffic engineering

- IEEE 802.1Qay, but not standardized
- Control plane for PBB
- Suite of several control protocols
Conclusion:

- Load adaption on layer 1
- Traffic engineering on layer 2
- Both PBB and MPLS are solid solutions

Future work:

- Performance comparison PBB vs. MPLS
- Implementation of demand-based CTP preparation
- Implementation of control architecture
Questions?
Full mesh of all customers
Full mesh of all edges
Full mesh of all edges

Vendorprobleem:
stub ziet i in twee vlans

Stéfan Deelen & Reinier Schoof
Improving scalability of the AMS-IX network
Full mesh of all edges

Learning via de core wordt overruled door de statische cam entry in de access node.

Nadeel: 1) spof’s! Toptalkersaannname, 2) beperkt schaalbaar; extra access bridges

Voordeel: beproefde techniek, realistisch
Distributed core