Towards Systematic Design of Enterprise Networks

Geoffrey Xie
Naval Postgraduate School

In collaboration with:
Eric Sung, Xin Sun, and Sanjay Rao (Purdue Univ.)
David Maltz (MSR)

July 7, 2011
Enterprise Network Design

- Ad-hoc and labor-intensive
- Costly
- Complex and error-prone

Ad-hoc, costly, error-prone
Systematic, economical, assured correctness
Solution: Systematic Top-Down Design

Our approach: model network-wide goals of design tasks rather than low-level mechanisms and configurations

Existing Approaches

• Template-driven (e.g., Presto)
• Vendor-neutral config languages (e.g., RPSL and NETCONF)
Contributions

Three-Step Systematic Design Methodology

1. Identify and abstract network-wide requirements (performance, security, resilience, etc) of design tasks
2. Formulate optimization problems
   - Model operator design strategies (e.g. cost functions) explicitly
   - Incorporate additional resource feasibility constraints
3. Analyze and develop solutions to formulated problems

Case study on two key enterprise design tasks
   - Virtual Local Area Network (VLAN) Design
   - Reachability control via packet filter placement

Validation on a large-scale campus network
Why VLAN and Reachability Design?

Design Process of Enterprise Networks
- Plan physical topology and wiring
- Create VLANs and layer-2 topology
- Select and configure routing protocols
- Control reachability with packet filters
What is VLAN?

- Common mechanism used by operators to think about users as collective groups based on user roles
- Allow users in physically separate locations to be treated as one logical subnet – increase manageability
- Hosts (H1~H4), Switches (S, S1~S3), Routers (R1&R2)
Gateway Router and Root Bridge

- Each VLAN is assigned a gateway router (GR) for routing and a root bridge to form spanning tree
  - Data traffic between hosts from 2 distinct VLANs must go through gateway routers of both VLANs
    - Longer delay, harder to debug since data may flow through other buildings, e.g., H1 in VLAN 1 (GR=R1) talks to H2 in VLAN 2 (GR=R2)
  - Broadcast traffic within a VLAN always goes to the root bridge
VLAN Task #1: Host Grouping

- **Goal:** Determine host membership of each VLAN
  - **Correctness criterion:** Separate hosts of different categories in different VLANs, e.g., Sales (H1, H3, other hosts attached to S3), Payroll (H2, H4), IT (other hosts attached to S2), etc.
  - **Operator intent:** (1) Limit the number of VLANs used, e.g., do not split H2 & H4; (2) Keep broadcast traffic (e.g., ARP) cost small in a VLAN, e.g., separate VLAN1 into VLAN 1 and VLAN4
Formulating Host Grouping problem

- **Network-wide Abstraction:** Host-category mapping so that no hosts in two distinct categories belong to the same VLAN
- **Formulation:**

\[
\text{Minimize } [C(x) + \max_{1 \leq i \leq x} \{\text{BroadcastCost}_i\}] \\
\text{subject to correct host-category mapping}
\]

where

\[
C(x) = \begin{cases} 0 & \text{if } x \leq \text{MAX\_VLANs} \\ \infty & \text{otherwise} \end{cases}
\]

and

\[
\text{BroadcastCost}_i = N_i \times B_i \times W_i
\]
Solving Host Grouping problem

• **Complexity theorem**: The Host Grouping problem as formulated is NP-hard with respect to the number of hosts.
  – Reduction from the “3-partition” problem

• **Heuristic solution**: (i) initially create VLANs from host categories, and then (ii) iteratively split largest VLAN into two based on host clustering, until the total number of VLANs reach MAX_VLANs.
VLAN Task #2: GR Selection

- **Goal:** For each VLAN, find the best choice of gateway router such that traffic cost is minimized

- **Considerations:**
  - Need to consider both intra-VLAN and inter-VLAN traffic
  - Exploit common traffic patterns, e.g., if H1 and H2 talk a lot, choose R1 as the gateway router for both VLAN 1 and VLAN 2
Formulating GR Selection problem

- **Network-wide Abstraction**: \((F_{i,j})\) -- Traffic Matrix specifying traffic exchanged between all VLAN pairs
- **Formulation**:

\[
\begin{align*}
\text{Minimize} & \quad \sum_i \sum_j \sum_k \sum_l F_{ij} (d_{ik}^* x_{ik} + d_{kl} x_{ik} x_{jl} + d_{jl}^* x_{jl}) + \\
& \quad \sum_i \sum_k F_{ii} 2d_{ik}^* x_{ik} + \sum_i \sum_k B_{r}(V_i, S_k) x_{ik} \\
\text{subject to:} & \quad \forall i, \sum_k x_{ik} = 1; \quad x_{IG} = 1; \\
& \quad i, j \in \{1, ..., v\} \cup \{I\}; \quad k, l \in \{1, ..., r\} \cup \{G\}
\end{align*}
\]  

(6)

Inter-VLAN traffic cost

Intra-VLAN traffic cost
Solving GR Selection problem

• **Complexity theorem**: The GR Selection problem as formulated is NP-hard with respect to the number of routers to select from.
  – Reduction from Quadratic Semi-Assignment Problem (QSAP)

• **Heuristic solution**: Identify server VLANs and client VLANs. (i) select GR for each of server VLANs while ignoring inter-VLAN traffic cost, and then (ii) select GR for each client VLAN while ignoring traffic cost between client VLANs
Reachability Control

- Control what packets sent by a source can reach a destination, e.g., A1&B1 can talk to C, but A2&B2 cannot talk to C
- Two configuration options:
  - **Control plane solution**: Routing based, e.g., at router X2, traffic from A2 or B2 is redirected to a black hole
  - **Data plane solution**: Access Control Lists (ACL)

<table>
<thead>
<tr>
<th>options</th>
<th>CPU Overhead</th>
<th>Ability to match packets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Routing</td>
<td>low</td>
<td>limited</td>
</tr>
<tr>
<td>ACL</td>
<td>high</td>
<td>flexible</td>
</tr>
</tbody>
</table>

Our Focus
**Design Task: Reachability Control via Placement of ACLs**

- **Goal:** Assume routing domains are fixed, place ACLs to meet desired security goals

- **Considerations:**
  - Security goals met even under failures
  - Limit on total rules that can be processed per router
Formulating ACL Placement problem

- **Network-wide Abstractions**
  - Reachability Set \((i,j)\): packets permitted from VLAN \(i\) to \(j\)
    \[
    RS(i,j) = \{ \text{packet } p \mid f(p) = 1 \}
    \]
  - Reachability Matrix: \(N \times N\) Matrix of *Reachability Sets*
  - Managed Event Set: Set of topology failures under which *Reachability Matrix* must remain invariant

- **Formulation:**
  - Correctness Criterion: *Reachability Matrix* must be upheld under events in *Managed Event Set*
  - Feasibility Criterion: \(b(r) \leq c(r)\), for all routers
    - \(c(r)\): \# of rules that can be configured on a router \(r\)
    - \(b(r)\): \# of rules actually configured on router \(r\)
Four Design Strategies

- Multiple **correct and feasible** placement possible, thereby allowing multiple design strategies, e.g.,
  - **Minimum Rules (MIN):** Minimize total # of rules
    - Minimize $\sum_r b(r)$
  - **Load Balancing (LB):** Minimize max # rules on one router
    - Minimize $\max_r \{b(r)\}$
  - **Capability Based (CB):** Maximize min residual capacity on one router
    - Maximize $\min_r \{c(r) - b(r)\}$
  - **Security Centric (SEC):** Minimize the average number of hops from source gateway to filtering point
    - Minimize $H$
Complexity of ACL Placement problem

• **Complexity theorem:** The ACL placement problem as formulated is NP-complete with respect to the number of ACLs that need to be placed.
  - Reduction from “Bin Packing Problem”
ACL Placement Algorithm

• Key Intuition:
  – To control reachability between VLANs $i$ and $j$, ACLs must be placed on an $(i,j)$ edge-cut-set (ECS)
    • Correctness guaranteed
  – Heuristics based on design strategy which chooses particular edge-cut-set

Strategy A
Strategy B
ECS1
ECS5
ECS2
ECS3
ECS6
ECS4
All $(i,j)$ edge-cut-sets

Strategy C
Strategy D
Validation

• Study on a large-scale campus network
  – 200 routers, 1300 switches, > 100 VLANs, >10,000 hosts
  – Dataset includes configs of all routers & switches.
    Available: http://www.ece.purdue.edu/~isl/network-config

• VLAN are extensively used
  – 10% span 5+ buildings. Largest VLAN spans 60 buildings

• Prominent ACL policies include:
  – Ingress filtering
  – Restricting communication involving (i) dormitory hosts,
    (ii) wireless traffic, and (iii) data centers
Traffic Reduction by Systematic VLAN Design

Systematic grouping
- Measured load: peak broadcast 2.12 pkts/s/source
- Core links: peak broadcast traffic reduced by 2k pkts/s

Systematic Router placement
- Data rate: uniform 10Kbps, or from LBNL enterprise trace
- Core links: data traffic reduced by 2-3 times
Dormitory VLANs

Registration Servers

Core Router

to other parts of the network

- ACL
- Leakage

Avoided by Systematic Placement

Inaccuracies in ACL Placement Today
Customizing ACL placement

- Take largest ACL rule in dataset (~700 rules), examine various systematic placement

<table>
<thead>
<tr>
<th>Metrics</th>
<th>c(r) &lt;= 300</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MIN</td>
</tr>
<tr>
<td>TotRules</td>
<td>1369</td>
</tr>
<tr>
<td>MaxRules</td>
<td>280</td>
</tr>
<tr>
<td>MinRes</td>
<td>20</td>
</tr>
<tr>
<td>ExtraHop</td>
<td>0</td>
</tr>
</tbody>
</table>

Each strategy performs best in its respective metric
Benefit of LB strategy

Old: 700 rules
Conclusion

• **Systematic Enterprise Design methodology**
  1. Identify and abstract network-wide requirements of design tasks
  2. Formulate optimization problems
  3. Analyze and develop solutions for formulated problems

• **Case Study**: VLAN design and ACL placement

• Results confirm benefits of systematic design
  – Correctness always ensured
  – Customizable to operator preferred strategies
Work in Progress

• Systematic Routing Design
  – A tougher nut to crack… so many components and their interactions are complex as we have seen yesterday
  – Starting with a bottom up analysis: quantifying the true complexity of different routing designs
  – We have developed a good abstraction to capture high level routing design objective (policy groups)
More details are in this paper

Thank you!
Analysis of VLAN Host Grouping
Analysis of VLAN GR Selection
Analysis of ACL Placement

\[ c(r_1) = c(r_2) = \ldots = c(r_k) = B \]