Scalable Firewalls for Publicly-Routable Cloud Tenants

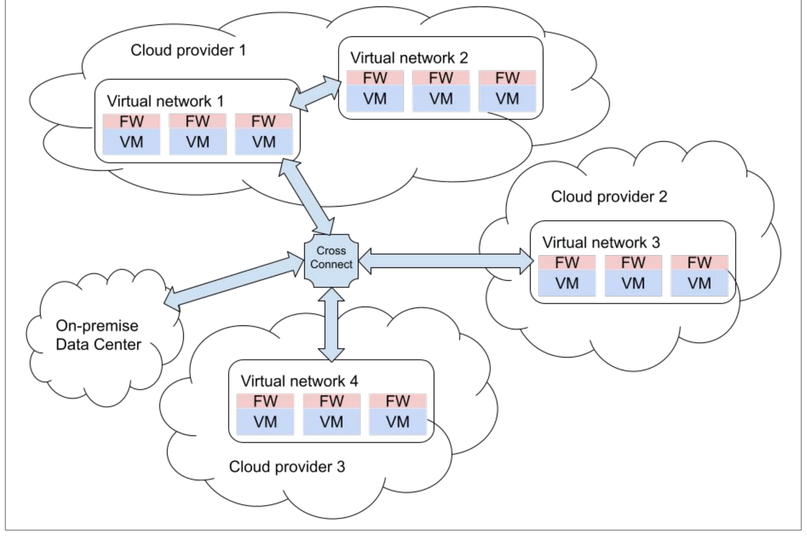
Project #3: Emily Marx, Tenzin Ukyab

Motivation

- Cloud tenants use workloads that span multiple cloud providers and on-premise data centers. Setting up the virtual network to connect all regions is a complex and expensive task.
  - A recent survey reports that over 88% of surveyed enterprises use two or more cloud providers and over 92% have both public and private cloud deployments.
- McClure *et al.* suggests solving this problem by abstracting away the details of virtual networks and only presenting cloud tenants with a clean API to set up their networks.
- This solution requires all endpoints be ***publicly-routable but default-off***: All endpoints should be public but accessible to only the permitted. This introduces scaling and security issues.
- In order to sell the network abstraction to cloud providers, we need to design a firewall system that protects publicly-routable endpoints ***scalably***, *i.e.*:
  - It does not need a drastic amount of CPU usage or congest network bandwidth.
  - It has reasonable latency performance.

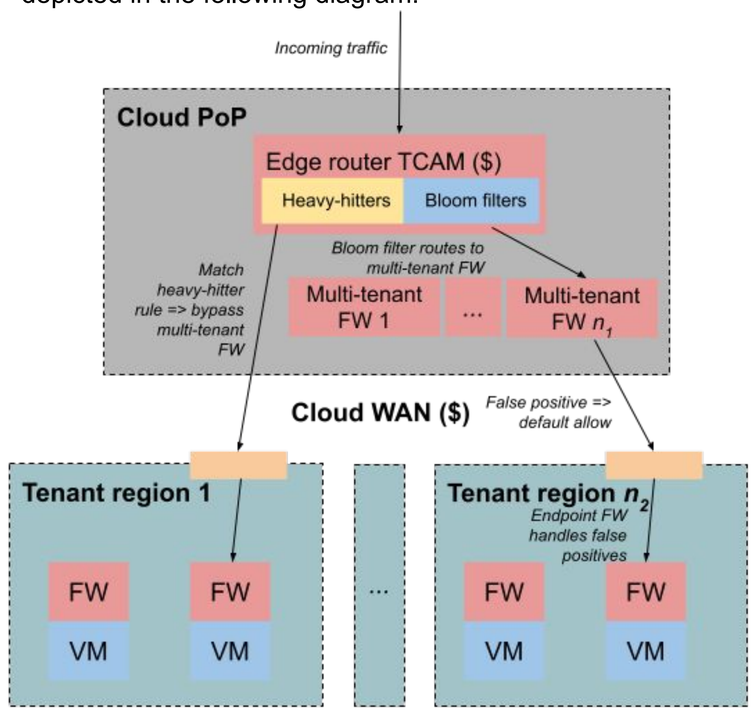
Background

- Currently, this problem does not exist because cloud do not make all endpoints publicly routable by default. Endpoints that are made public are protected by access lists using firewalls at the endpoints.
- However, with publicly routable endpoints, we cannot solely rely on per-VM firewalls because:
  - Cloud network's WAN resources are expensive and vulnerable to resource-exhaustion attacks.
  - Cloud firewalls will not scale with the increase of the number of ACL rules.

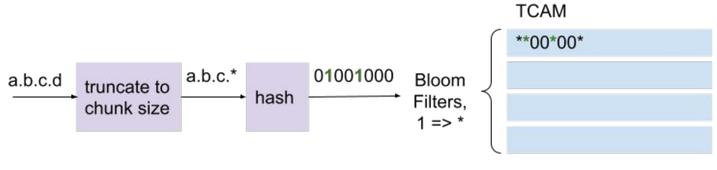


Design

- In order to preserve cloud WAN bandwidth and reduce the risk of DDoS attacks on endpoints, our design must prevent most disallowed traffic from leaving the PoP. However, this cannot be done entirely at the PoP edge router for two reasons:
  - TCAM is expensive, and existing edge routers are already near 100% TCAM utilization.
  - Even in SRAM, the number of access rules across all endpoints is too large to fit in a ***single*** firewall with acceptable lookup latency.
- Our design addresses these challenges in three main ways, depicted in the following diagram:



  - 1. Repurpose the cloud PoP edge router's TCAM to represent multiple tenants per TCAM line. Each TCAM line is a Bloom filter containing the *destination* IPs of multiple tenants, represented as fixed-size blocks, with 1 bits replaced by *:



  - 2. In the extra TCAM space generated by #1, add the most heavily-used complete rules (destination *and* source IP).
  - 3. Add NetFilter firewalls sharded by tenant to the cloud PoP. If the destination and source IP of an incoming packet matches one of the heavy-hitter rules in TCAM, route it over the WAN. Else, the TCAM action matching the destination routes to a multi-tenant firewall.

Evaluation

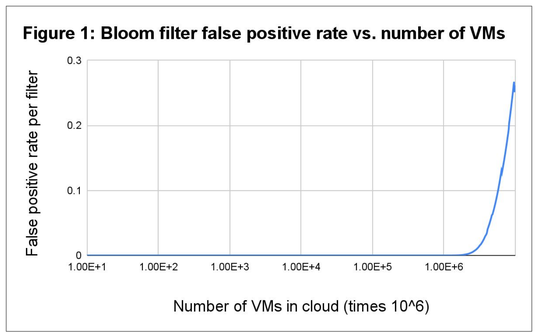
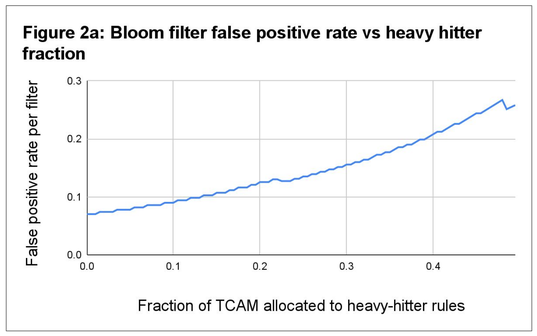


Figure 1: As the number of tenant endpoints increases, the TCAM Bloom filters are more likely to have false positives.



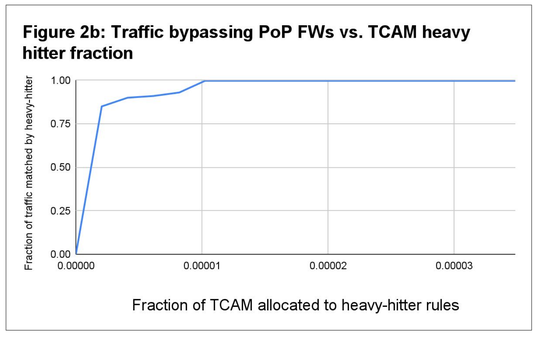


Figure 2a and 2b: Increasing the fraction of TCAM allocated to heavy-hitter rules increases Bloom filter false positive rate, but allows more traffic to bypass the PoP's multi-tenant firewalls.

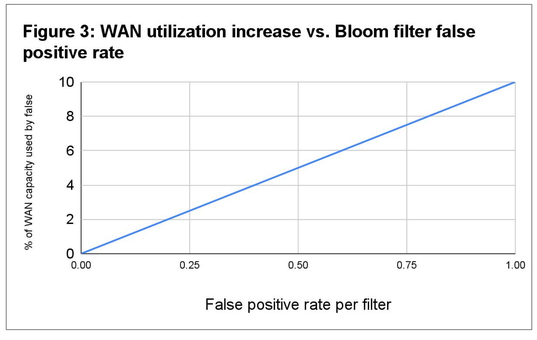


Figure 3: Bloom filter false positives consume cloud WAN bandwidth.

Summary of CPU utilization results: The status quo uses 0.5 cores for VPN encryption per tunnel. Our design eliminates these cores, but adds cores for the multi-tenant firewalls. In these firewalls, CPU utilization is nearly constant with the number of rules.

References

- McClure, S., Ratnasamy, S., Bansal, D., Padhye, J. Rethinking networking abstractions for cloud tenants. In Proceedings of the Workshop on Hot Topics in Operating Systems(New York, NY, USA, June 2021), HotOS '21, Association for Computing Machinery, pp. 41–48.