IETF117 Proposal:

Path Selection in Multi-Tunnel SD-WAN

Altanai B (altanai@outlook.com)
Cisco Meraki

Problem statement

Present day traffic streams in a VPN or secure ecosystem, have a pre configured strategy to choose the route, such VPN tunnel, split tunnel, MASQUE tunnel and so on and route is set likewise. At the very best there may be multiplexing, weighted or round-robin load sharing in cases where there are multiple options available.

Keywords : VPN tunnels, MASQUE, dynamic traffic steering

Proposal

By dynamically deciding the tunnel type for a stream or packet, we could avoid the non-performing or counter-productive use-cases such as :

- added latency on real time streaming
- added encryption for already end-to-end encrypted VoIP calls
- NAT traversal nightmare
- nested tunneling and double congestion control
- exhausting limited bandwidth available from VPN providers

The proposal is to standardize an algorithm that computes multiple available options and decides whether, on-demand tunnels are created (via MASQUE, IPSec, SSH, GRE other proprietary protocols such as AutoVPN), an existing set of tunnels be reused or any other route, based on the current network dynamics and vulnerability of the traffic.

Summary

Hybrid work and move towards private access has increased the interest in tunneling traffic between endpoints. However at present, the traffic steering decision is made in a limited scoped or rule based manner which is different for various networks and service providers. Instead an alternative dynamic strategy is proposed which gauges the confidence in the various available options dynamically and may choose to send data directly via edge gateway, use one or more of the available tunnels or create a new on-demand tunnel, leveraging any of the tunneling protocols best suited.

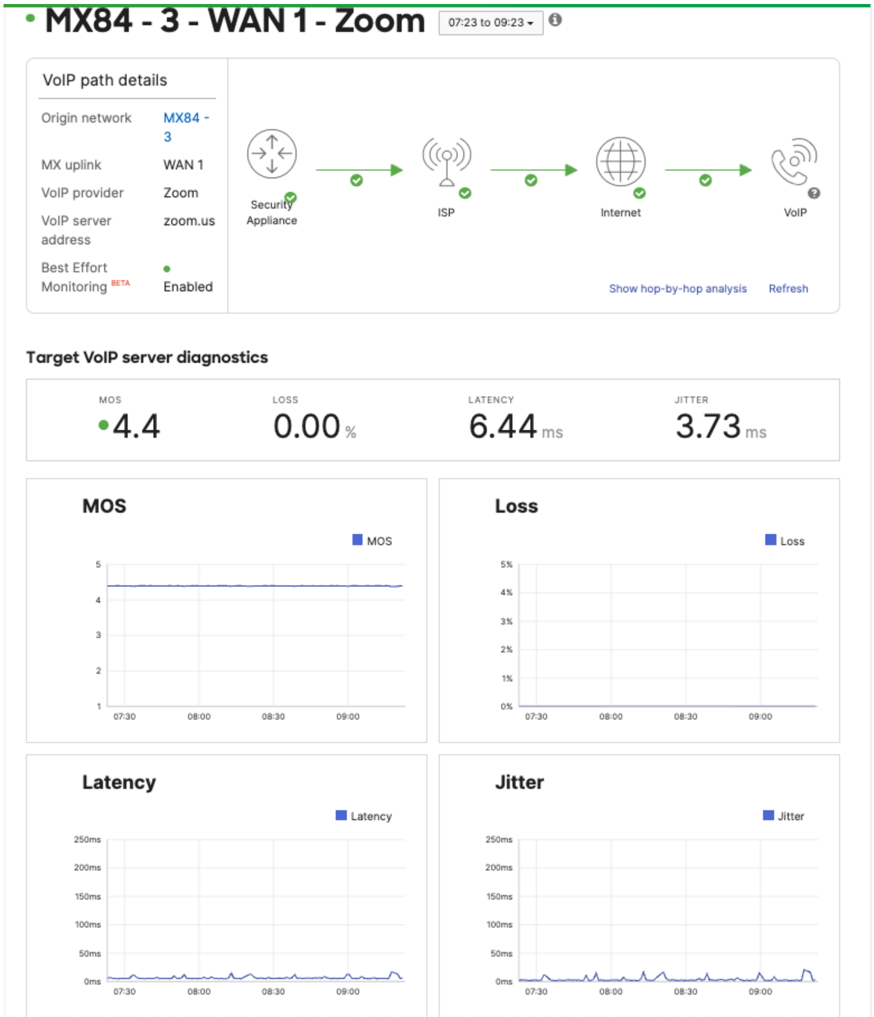
At present, in the case of multiple active uplinks connecting to various ISPs, there are multiple techniques to steer or prioritize traffic across the network, which may include,

- Full or Split tunnel based on DSCP tags(Diffserv). For example in case of dual uplinks connected and two tunnels active, use the one more with better performance for RTP traffic.

Traffic Type	DSCP tag

SIP (Voice)	46 (EF - Expedited Forwarding, Voice)
All Advertising, All Software Updates, All Online Backups	10 (AF11 - High Throughput, Latency Insensitive, Low Drop)
WebEx, Skype	34 (AF41 - Multimedia Conferencing, Low Drop)
All Video & Music	18 (AF21 - Low Latency Data, Low Drop)

- Multiple Active VPN Uplinks used in weighted round robin order or ECMP
- Traffic Shaping generic rules based on
  - QoS (such as MOS, jitter other customized score)



MX84 - 3 - WAN 1 - Zoom

07:23 to 09:23

VoIP path details

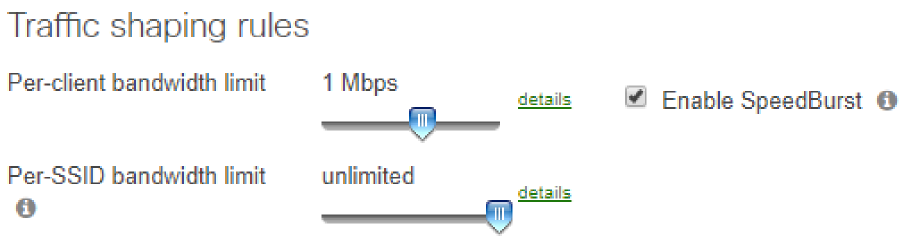
Origin network	MX84 - 3
MX uplink	WAN 1
VoIP provider	Zoom
VoIP server address	zoom.us
Best Effort Monitoring BETA	● Enabled

Target VoIP server diagnostics

MOS	LOSS	LATENCY	JITTER
●4.4	0.00 %	6.44 ms	3.73 ms

Figure : Network metrics to gauge the performance

  - Attributes such as app type or address
  - Client identifier based shaping



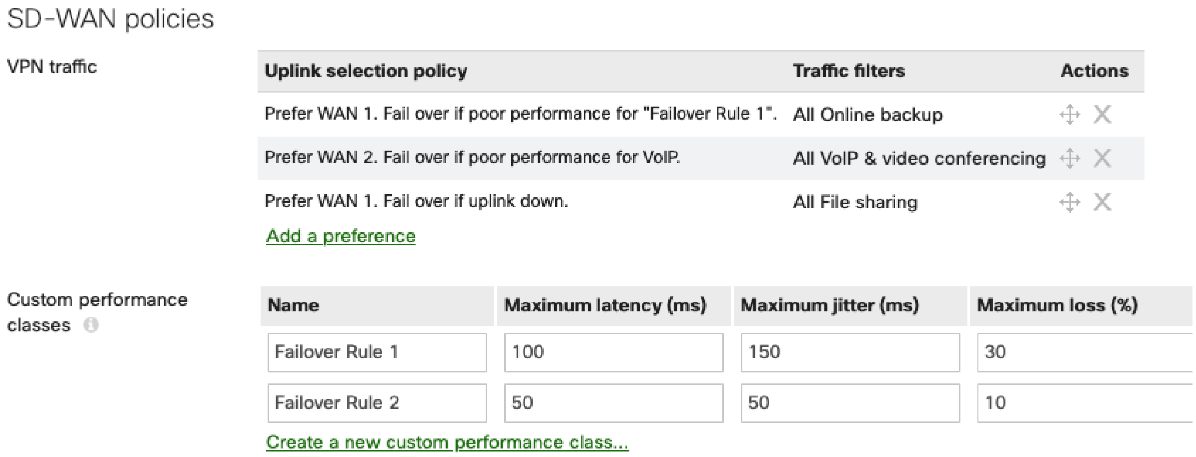
Traffic shaping rules

Per-client bandwidth limit: 1 Mbps details ☑ Enable SpeedBurst

Per-SSID bandwidth limit: unlimited details

Figure : Limiting bandwidth for all clients. Alternatively limits can also be imposed per client by mac addr or IP etc.

- Policy-Based routing that use flow preferences to pin traffic to a particular path. It could also be Geo or proximity based rules.



SD-WAN policies

VPN traffic

Uplink selection policy	Traffic filters	Actions
Prefer WAN 1. Fail over if poor performance for "Failover Rule 1".	All Online backup	✥ ✕
Prefer WAN 2. Fail over if poor performance for VoIP.	All VoIP & video conferencing	✥ ✕
Prefer WAN 1. Fail over if uplink down.	All File sharing	✥ ✕

Add a preference

Custom performance classes

Name	Maximum latency (ms)	Maximum jitter (ms)	Maximum loss (%)
Failover Rule 1	100	150	30
Failover Rule 2	50	50	10

Create a new custom performance class...

Figure : SD-WAN policy based routing in case of dual active tunnels via two uplinks

- Dynamic Path Selection such as Network Based Application Recognition (NBAR) from Cisco
- MASQUE (QUIC multiplexing)

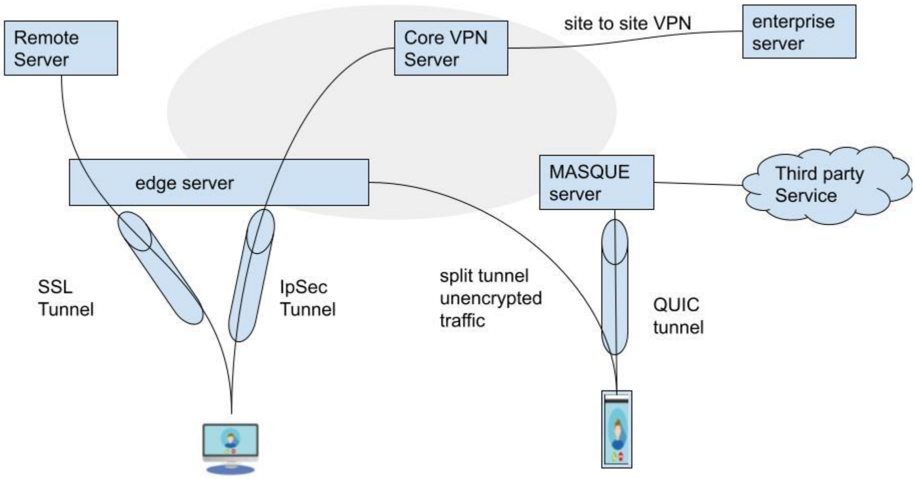


Figure: Multiple tunneling options to steer traffic

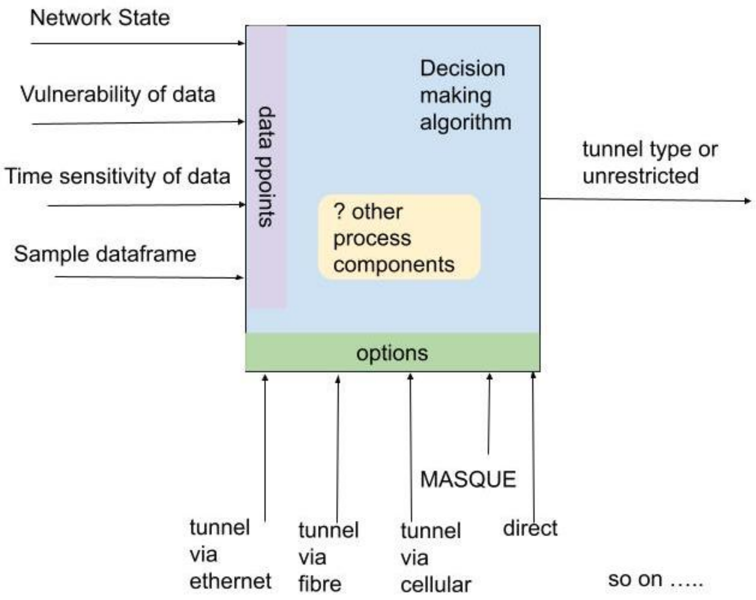


Figure : Suggestive data points for the Decision making algorithm

Example of the decision that may be taken by the standardized algorithm could include;

1. Resource intensive application benefit from direct internet connection such as multiplayer games
2. Tunneling the VoIP traffic via separate routes, for example signaling plane data on VPN tunnel, and media via MOQ.
3. SIP trunk calls may actually benefit from a dedicated IPSec tunnel, pre NATed, pre authenticated and secure, as it would avoid the delay in resetting the path given the volume of calls expected between two endpoints.

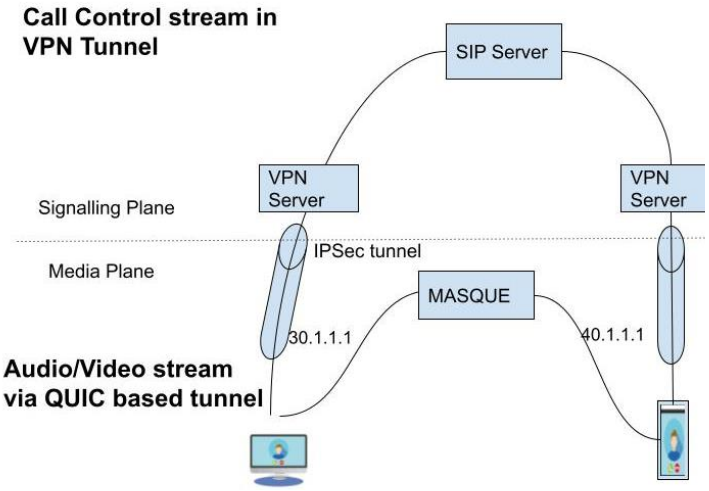


Figure : Send signaling data via tunnel and media via MOQ in VoIP Call

4. Heavy file downloads such as VoD could benefit by load sharing between multiple tunnels.

Other indirect impacts of the standardization of such a tunnel path selection algorithm may also be to overcome strategies which unfairly maximize bandwidth usage in the public internet.

Scope for further work

Mainstream techniques include packet marking and queuing of other non critical traffic to optimize for real time streams is essentially prioritization in practice. However, VPN providers, CSPs and/or ISP may employ polar-opposite algorithms to shape traffic based on their interest which could lead to an overall non-synchronized approach, where it is prioritized in some networks and deprioritized in other networks.
Standardized Path selection decision making making algorithm would ensure same treatment of the stream across heterogeneous network environments.

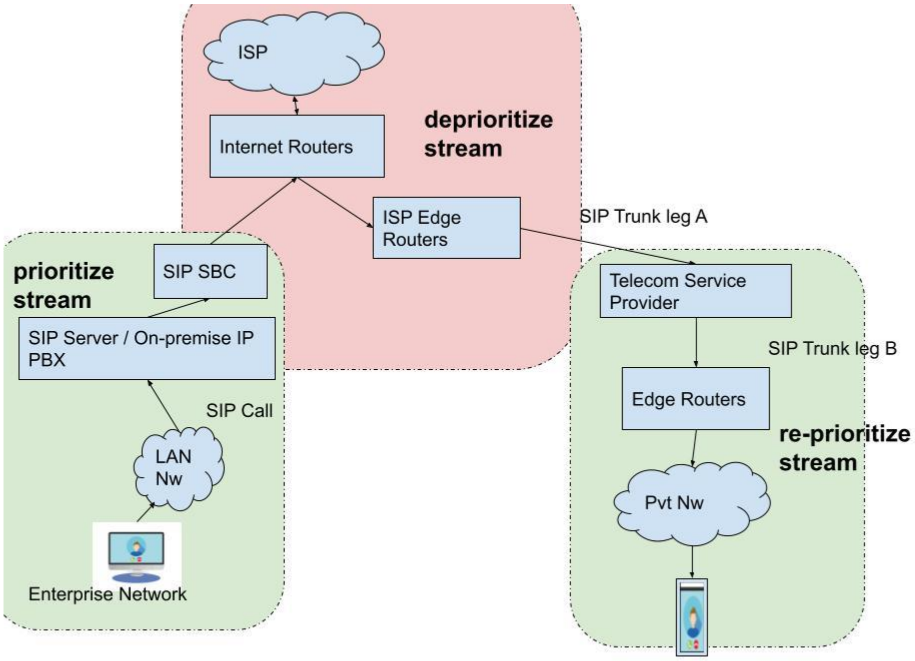


Figure : Various networks encountered in a SIP trunk call

In an heterogeneous network system, an edge gateway could decide the data to be sent via the VPN if it is destined for a specific subnet that is being advertised, or checked against other available routes such as static LAN or third party routes. Alternatively if not matched it could be NATed and sent out unencapsulated.