Network Working Group
Internet Draft
Intended status: Informational
Expires: January 2018

L. Dunbar
A. Malis
Huawei

October 30, 2017

Gap Analysis of VPN Extension to Dynamic Cloud Data Center
draft-dm-vpn-ext-to-cloud-dc-gap-analysis-00

Abstract

This document analyzes the technological gaps necessary to enable existing VPN to securely connect to dynamic workloads hosted in cloud data centers when the cloud DC doesn't have the VPN PEs co-located [dynamic-cloudDC].

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79. This document may not be modified, and derivative works of it may not be created, except to publish it as an RFC and to translate it into languages other than English.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

This Internet-Draft will expire on April 30, 2009.

Table of Contents

1. Introduction

[dynamic-cloudDC] describes the problems of today's State-of-Art VPN technologies in connecting enterprise branch offices to dynamic workloads in Cloud DC. This document analyzes the technological gaps necessary to enable existing VPN to securely connect to dynamic workloads hosted in cloud data centers when the cloud DC does not have the VPN PEs co-located.

2. Conventions used in this document

Cloud DC: Off-Premise Data Centers that usually host applications and workload owned by different organizations or tenants.

Controller: Used interchangeably with SD-WAN controller to manage SD-WAN overlay path creation/deletion and monitoring the path conditions between two sites.

OnPrem: On Premises data centers and branch offices

SD-WAN: Software Defined Wide Area Network, which can mean many different things. In this document, "SD-WAN" refers to the solutions specified by ONUG (Open Network User Group), which build point-to-point IPsec overlay paths between two end-points (or branch offices) that need to intercommunicate.

3. Connect OnPrem DCs & branches with dynamic workloads in Cloud DC

With the advent of widely available third party cloud data centers in diverse geographic locations and the advancement of tools for monitoring and predicting application behaviors, it is technically feasible for enterprises to instantiate applications and workloads

in Cloud DCs that are geographically closest to their end users. This property can improve overall end user experience.

However, those Cloud DCs might not have the co-located PEs for the commonly deployed VPNs (e.g. L2VPN, L3VPN) that interconnect enterprises' branch offices and on-premise data centers.

SD-WAN, conceived in ONUG (Open Network User Group) a few years ago, has emerged as an on-demand technology to securely interconnect any two locations, which theatrically can connect the OnPrem branches with the workloads instantiated in Cloud DCs that do not have MPLS VPN PE co-located. However, to use the SD-WAN to connect the enterprise existing sites with the workloads in Cloud DC, the enterprise existing sites' CPEs have to be upgraded to support SD-WAN. If the workloads in Cloud DC need to be connected to many sites, the upgrade process can be very expensive.

[dynamic-cloudDC] describes a hybrid network approach, (a.k.a. VPN extension to Dynamic Cloud DC throughout the document), that integrates SD-WAN with traditional MPLS-based VPNs, to connect OnPrem locations with Cloud DC Workloads with minimum changes to existing CPEs.

The VPN Extension to dynamic workload in Cloud DC has the assumption that the workloads in Cloud DC can be temporary or may be migrated to different DCs over time, therefore, cannot justify the cost of adding new PEs to the existing MPLS VPN in order to reach the Cloud DC.

To extend the existing MPLS VPN to Cloud DC over the access paths that are not under the VPN provider control, a small number of the PEs of the MPLS VPN can be designated to connect to the remote workloads via SD-WAN secure IPsec tunnels. Those designated PEs are shown as fPE (floating PE or smart PE) in Figure 1 below. Once the secure IPsec tunnels are established, the workloads in Cloud DC can be reached by the enterprise's VPN without upgrading all of the enterprise's existing CPEs. The only CPE that needs to support SD-WAN would be a virtualized CPE instantiated within the cloud DC.

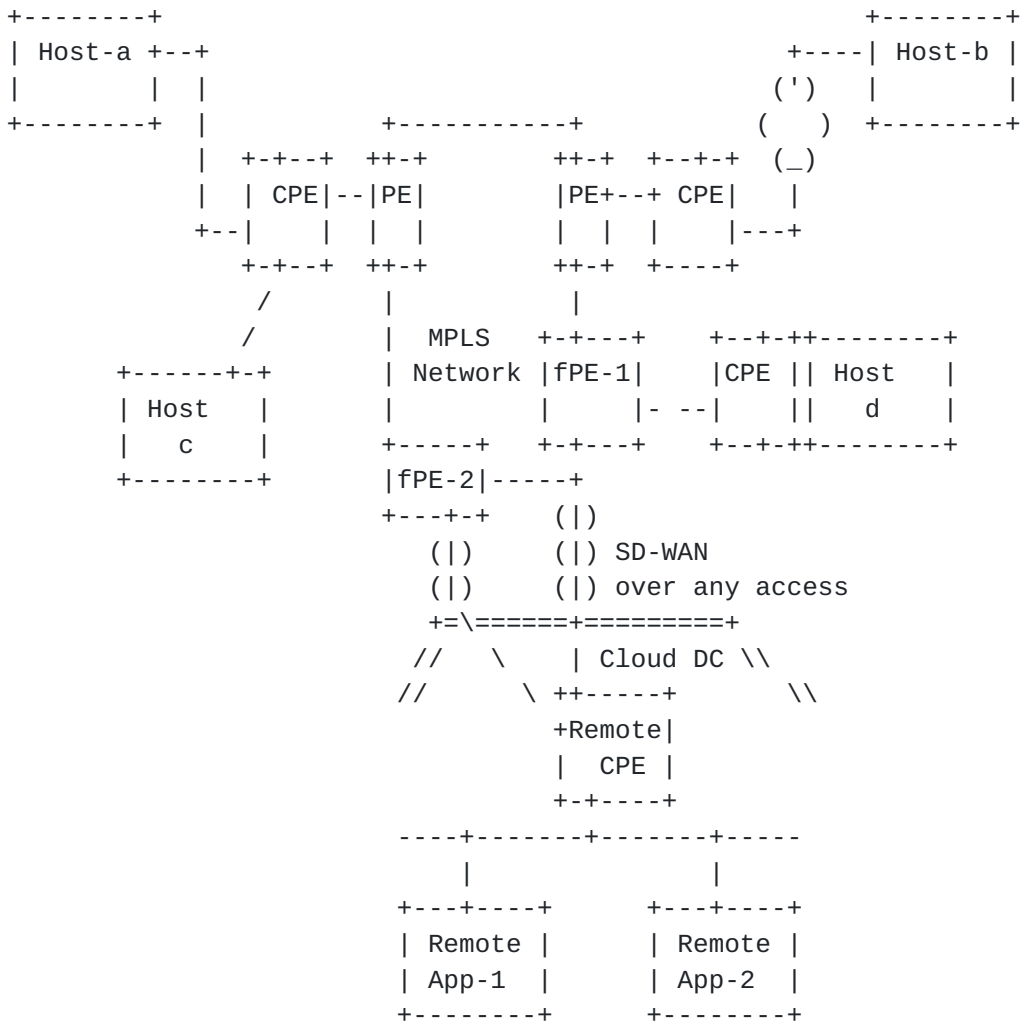

Figure 1: VPN Extension to Cloud DC

In Figure 1 above, the optimal Cloud DC to host the workloads (due to proximity, capacity, pricing, or other criteria chosen by the enterprises) does not happen to have a direct connection to the PEs of the MPLS VPN that interconnects the enterprise's existing sites.

4. Gap Analysis

4.1. Floating PEs to connect to Remote Edges

When an Enterprise's MPLS VPN does not have PEs co-located with the Cloud DC that is the optimal location to host workloads, a small set of PEs can be designated as the "floating PEs (fPE)" to connect to the (virtualized) CPEs in the Cloud DC via SD-WAN IPsec tunnels over the any access paths, such as public Internet, LTE, or others.

As long as PEs have the following property, the SD-WAN IPsec tunnels can be established:

- Be able to support IPsec tunnel termination
- The performance measurements between the PE and the remote CPE (or the virtualized CPE in Cloud DC) can be measured, such as round time delay, two way active measurement protocol (TWAMP) [RFC5357], etc., so that more intelligent selection can be made if there are multiple PEs available for connection.
- Have sufficient capacity to route traffic to/from remote CPEs in Cloud DC.

Gap:

Even though a set of PEs can be manually selected to act as the floating PE for a specific cloud data center, there are no standard protocols for those PEs to interact with the remote CPEs (most likely virtualized) instantiated in the third party cloud data centers (such as exchanging performance information or route information).

Some SD-WAN networks use the NHRP protocol [RFC2332] to register SD-WAN endpoints with an NHRP server, which then has the ability to map a private VPN address to a public IP address of the destination node (i.e. PE). However, not all CPEs in cloud data center support NHRP registration for the set of private addresses of workload instantiated in the data center, and does not have ways to be automatically configured with the address of the NHRP server. Without proper address of the CPE in Cloud DC, it is difficult for an "optimal" fPE to act as the SD-WAN conduit to the DC.

When there is more than one fPE available for use (as there should be for resiliency or the ability to support multiple cloud DCs scattered geographically), multi-homing from the remote CPE in cloud has issues to VPN has unresolved issues.

4.2. Need Secure Channel into the Cloud DC

Today's common network connection to Cloud DC is via IPsec tunnel terminated at the Cloud DC Gateway, and depends on Cloud DC network to connect to the leased compute & storage resources , or virtual private cloud within the Cloud DC.

Some enterprises prefer to have secure tunnels all the way to their own workloads hosted in the cloud to increase its own security control to its workloads. Since the OnPrem workloads or application might not have the application layer secure layer, the end to end secure path would be from either OnPrem CPE or PE into the virtual CPEs in the Cloud DC.

4.3. Need to virtual networks differentiation on the IPsec tunnel

When there are multiple virtual networks in Cloud DC to be connected to enterprise's existing VPN, it is desirable to have traffic from those virtual networks sharing the same IPsec tunnel between PEs and the Cloud DC Gateway. Therefore, It is necessary to differentiate traffic belong to different virtual networks within one IPsec tunnel.

4.4. NAT Traversing

Most cloud DCs only assign private IP addresses to the workloads instantiated. Therefore, the traffic to/from the workload usually need to traverse NAT.

4.5. Complication of use BGP between PE and remote CPEs via Internet

Even though EBGP (external BGP) Multihop method can be used to connect peers that are not directly connected to each other, there are still some complications/gaps in extending BGP from MPLS VPN PEs to remote CPEs via any access paths (e.g. internet):

EBGP Multi-hop scheme requires static configuration on both peers. To use EBGP between a PE and remote CPEs, the PE has to be statically configured with "next-hop" to the IP addresses of the CPEs. When remote CPEs, especially remote virtualized CPEs dynamically instantiated or removed, the configuration on the PE Multi-Hop EBGP has to be changed accordingly.

Gap:

Egress peering engineering (EPE) is not enough. Running BGP on virtualized CPE in Cloud DC requires GRE tunnels being established first, which requires address and key management for the remote CPEs. RFC 7024 (Virtual Hub & Spoke) and Hierarchical VPN is not enough

Also need a method to automatically trigger configuration changes on PE when remote CPEs' are instantiated or moved (IP address change) or deleted.

EBGP Multi-hop scheme does not have embedded security mechanism. The PE and remote CPEs needs secure communication channel when connected via public internet.

Remote CPEs, if instantiated in Cloud DC, might have to traverse NAT to reach PE. It is not clear how BGP can be used between devices outside the NAT and the entities behind the NAT. It is not clear how to configure the Next Hop on the PEs to reach private addresses.

4.6. Controller Facilitated Route Distribution

Some remote applications & workloads hosted in third party Cloud DCs may only need to communicate with a small number of subnets (or Virtual Networks) at a limited number of an enterprise's VPN sites. Running an IGP among the remote (virtual) CPE in the Cloud DC and all of the VPN sites to establish a full mesh routing table for every site could be overkill.

Instead of running IGP with all other sites, the remote CPEs can register its attached hosts to the controller via NHRP, which in

turn passes the addresses attached to the remote edges to the relevant PEs/CPEs that need to communicate with the remote edges.

Gap:

A complicating issue is that the remote CPEs are not directly connected to any of the PEs of the MPLS VPN. This may make it difficult to use either an IGP or BGP as a method distribute routes within the VPN to reach a particular private address within a data center. However, route distribution may be possible once an IPSec tunnel has been established. This needs to be investigated.

4.7. Designated Forwarder to the remote edges

Among multiple floating PEs available for a remote CPE, multicast traffic from the remote CPE towards the MPLS VPN can be broadcasted back to the remote CPE due to the PE receiving the broadcast data frame forwarding the multicast/broadcast frame to other PEs that in turn send to all attached CPEs. This process may cause a traffic loop.

Therefore, it is necessary to designate one floating PE as the CPE's Designated Forwarder, similar to TRILL's Appointed Forwarders [RFC6325].

Gap: the MPLS VPN does not have features like TRILL's Appointed Forwarders.

4.8. Traffic Path Management

When there are multiple floating PEs that have established IPsec tunnels to the remote CPE, the remote CPE can forward the outbound traffic to the Designated Forwarder PE, which in turn forwards the traffic to egress PEs to the destinations. However, it is not straightforward for the egress PE to send back the return traffic to the Designated Forwarder PE.

Example of Return Path management using Figure 1 above.

- fPE-1 is desired for communication between App-1 <-> Host-a due to

latency, pricing or other criteria.
- fPE-2 is desired for communication between App-1 <-> Host-b.

4.9. Smart PE

A Smart PE is the PE that can interact with remote CPE (or the Controller) to learn the communication peer and pattern of services hosted in third party DC. With that learned information, the Smart PE can intelligently manage transport paths within the MPLS-based VPN (for example, choosing the optimized egress PEs) based on the delay and QoS measurement among different PEs (in order to support high SLA requests from the CPE).

Gap: There needs to be a protocol to select Smart PEs.

5. Manageability Considerations

TBD

6. Security Considerations

TBD.

7. IANA Considerations

This document requires no IANA actions. RFC Editor: Please remove this section before publication.

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

8.2. Informative References

[RFC8192] S. Hares, et al "Interface to Network Security Functions (I2NSF) Problem Statement and Use Cases", July 2017

[ITU-T-X1036] ITU-T Recommendation X.1036, "Framework for creation, storage, distribution and enforcement of policies for network security", Nov 2007.

[Dynamic-CloudDC] L.Dunbar and A. Malis, "Dynamic Cloud Data Center VPN Problem Statement", Nov 2017

9. Acknowledgments

Acknowledgements to xxx for his review and contributions.

This document was prepared using 2-Word-v2.0.template.dot.

Authors' Addresses

Linda Dunbar
Huawei
Email: Linda.Dunbar@huawei.com

Andrew G. Malis
Huawei
Email: agmalis@gmail.com