Network Working Group L. Dunbar
Internet Draft A. Malis
Intended status: Informational Huawei
Expires: July 2019 C. Jacquenet
Orange
M. Toy
Verizon
January 24, 2019

Seamless Interconnect Underlay to Cloud Overlay Problem Statement
draft-dm-net2cloud-problem-statement-05

Abstract

This document describes the problems of enterprises face today in connecting their branch offices to dynamic workloads in third party data centers (a.k.a. Cloud DCs).

It examines the approach of using SD-WAN to utilize multiple underlay networks by different providers to maximize or optimize interconnection among branch offices, on-premises DCs and Hybrid Cloud DCs.

This document also describes some of the (network) problems that many enterprises face when they have workloads & applications & data split among hybrid data centers, especially for those enterprises with multiple sites that are already interconnected by VPNs (e.g. MPLS L2VPN/L3VPN) and leased lines.

Current operational problems in the field are examined to determine whether there is a need for enhancements to existing protocols or whether a new protocol is necessary to solve them.

Table of Contents

1. Introduction

Cloud applications and services continue to change how businesses of all sizes work and share information. "Cloud applications & workloads" are those that are instantiated in third party DCs that also host services for other customers.

With the advent of widely available third party cloud DCs 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 locations that are geographically closest to their end users. This property aids in improving end-to-end latency and overall user experience. Conversely, an enterprise can easily shutdown applications and workloads when their end users' geographic base changes (therefore needing to change the networking connection to those relocated applications and workloads). In addition, an enterprise may wish to

take advantage of more and more business applications offered by third party private cloud DCs, such as Dropbox, Microsoft365, SAP HANA, Oracle Cloud, Salesforce Cloud, etc.

Most of those enterprise branch offices & on-premises data centers are already connected via VPNs, such as MPLS based l2VPN/L3VPN. Then connecting to the cloud-based resources may not be straightforward if the provider of the VPN service does not have direct connections to the cloud DCs. Under those circumstances, the enterprise can upgrade their existing CPEs to utilize SD-WAN to reach cloud resources (without any assistance from the VPN service provider), or wait for their VPN service provider to make new agreements with data center providers to connect to the Cloud resources. Either way has additional infrastructure and operational costs.

In addition, it is an uptrend with more enterprises instantiating their Apps & workloads in different Cloud DCs to maximize the benefits of geographical convenience, elasticity and special features offered by different Cloud DCs.

2. Definition of terms

Cloud DC: Third party 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 or more sites.

DSVPN: Dynamic Smart Virtual Private Network. DSVPN is a secure network that exchanges data between sites without needing to pass traffic through an organization's headquarter virtual private network (VPN) server or router.

Heterogeneous Cloud: applications & workloads split among Cloud DCs owned & managed by different operators.

Hybrid Cloud: On premise DC plus Cloud DCs owned and managed by different organizations. In this document Hybrid Cloud also include heterogeneous cloud as well.

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), https://www.onug.net/software-defined-wide-area-network-sd-wan/, which is about pooling WAN bandwidth from multiple underlay networks to get better WAN bandwidth management, visibility & control. Some of the underlay networks are private networks over which traffic can traverse without encryption, others requires IPsec tunnels between SD-WAN nodes to carry the traffic.

VPC: Virtual Private Cloud. A service offered by many Cloud DC operators to allocate a logically isolated cloud resources, including compute, networking and storage.

3. Current Practices in Interconnecting Enterprise Sites with Cloud DCs

3.1. Interconnect to Cloud DCs

Most Cloud operators offer some type of network gateway through which an enterprise can reach their workloads hosted in the Cloud DCs. For example, AWS (Amazon Web Services) offers the following options to reach workloads in AWS Cloud DCs:

- Internet gateway for any external entities to reach the workloads hosted in AWS Cloud DC via the internet.
- Virtual gateway (vGW) to which IPsec tunnels [RFC6071] are established between an enterprise's own gateways and AWS vGW, so that the communications between those gateways can be secured from the underlay (which might be the public internet).
- Direct Connect, which allows enterprises to purchase direct connect from network service providers to get a private leased line interconnecting the enterprises gateway(s) and the AWS Direct Connect routers co-located with the network operators.

Via Direct Connect, AWS Transit Gateway can be utilized to interconnect multiple VPCs in different Availability Zones.

CPEs at one Enterprise branch office can have ports facing internet to connect to AWS's vGW via IPsec and other ports connected to AWS DirectConnect via private network (without encryption).

```
     +------------------------+
     |    ,---.         ,---. |
     |   (TN-1 )       ( TN-2)|
     |    `-+-'  +--+   `-+-' |
     |      +----|vR|-----+   |
     |           ++-+         |
     |            |         +-+----+
     |            |        /Internet\ For External
     |            +--------+ Gateway  +----------------------
     |                     \        / to reach via Internet
     |                      +-+----+
     |                        |
     +------------------------+

     +------------------------+
     |    ,---.         ,---. |
     |   (TN-1 )       ( TN-2)|
     |    `-+-'  +--+   `-+-' |
     |      +----|vR|-----+   |
     |           ++-+         |
     |            |         +-+----+
     |            |        / virtual\ For IPsec Tunnel
     |            +--------+ Gateway  +----------------------
     |                     \        /  termination
     |                      +-+----+
     |                        |
     +------------------------+

     +------------------------+
     |    ,---.         ,---. |
     |   (TN-1 )       ( TN-2)|
     |    `-+-'  +--+   `-+-' |
     |      +----|vR|-----+   |
     |           ++-+         |
     |            |         +-+----+              +-------+
     |            |        /        \ For Direct /customer\
     |            +--------+ Gateway  +-----------+ gateway  |
     |                     \        /  Connect    \        /
     |                      +-+----+              +-------+
     |                        |
     +------------------------+
```

Figure 1: Examples of connecting to a Cloud DC

3.2. Interconnect to Hybrid Cloud DCs

According to Gartner, by 2020 "hybrid will be the most common usage of the cloud" as more enterprises see the benefits of integrating public and private cloud infrastructures. However, enabling the growth of hybrid cloud deployments in the enterprise requires fast and safe interconnection between public and private cloud services. The Hybrid Cloud scenario also includes heterogeneous Cloud DCs, meaning Cloud DCs owned and managed by different organizations.

For an enterprise to connect to applications & workloads hosted in multiple Cloud DCs, the enterprise can use IPsec tunnels over internet or/and lease private networks to connect its on-premises gateways to each of the Cloud DC's gateways, virtual routers instantiated in the Cloud DCs, or any other suitable design (including a combination thereof).

Some enterprises prefer to instantiate their own virtual CPEs/routers inside the Cloud DC to connect the workloads within the Cloud DC. Then an overlay path is established between customer gateways to the virtual CPEs/routers for reaching the workloads inside the cloud DC.

3.3. Connecting workloads among hybrid Cloud DCs

There are multiple approaches to interconnect workloads among different Cloud DCs:

- Utilize Cloud DCs provided transit gateways, which usually does not work if Cloud DCs are owned and managed by different Cloud providers.
- Hairpin all the traffic through the customer gateway, which creates additional transmission delay & incurs cost exiting Cloud DCs, or
- Establish direct tunnels among different VPCs (Virtual Private Clouds) via client's own virtual routers instantiated within

Cloud DCs. Using DMVPN (Dynamic Multipoint Virtual Private

Network) or DSVPN (Dynamic Smart VPN) to establish direct Multi-edge tunnels among those client's own virtual routers.

DMVPN & DSVPN use NHRP (Next Hop Resolution Protocol) [RFC2735] so that spoke nodes can register their IP addresses & WAN ports with the hub node. The IETF ION (Internetworking over NBMA (non-broadcast multiple access) WG, standardized NHRP for connection-oriented NBMA network (such as ATM) network address resolution more than two decades ago.

There are many differences between virtual routers in Public Cloud DCs and the nodes in an NBMA network. NHRP & DSVPN are not effective in registering virtual routers in Cloud DCs without major extension. Another option is using other protocols such as BGP approach described in [BGP-SDWAN]. As the result of this evaluation, enhancement or new protocols for distributing edge node/port properties may come out.

4. Desired Properties for Networks that interconnects Hybrid Clouds

The networks that interconnect hybrid Cloud DCs have to enable users to take advantage of Cloud DCs:

- High availability, any time usage for any length of time.
  Many enterprises incorporate Cloud as their disaster recovery strategy, e.g. periodically backup data into the cloud, or running backup applications in the Cloud, etc. Therefore, the connection to the cloud DCs may not be permanent, but rather needs to be on-demand.

- Global accessibility in different geographical zones, thereby facilitating the proximity of applications as a function of the end users' location, for improved latency.
- Elasticity and mobility, to instantiate additional applications at Cloud DCs when end users' usages increase and shut down applications at locations with fewer end users.
  Some enterprises have front-end web portals running in Cloud DCs and Database servers in their on-premises DCs. Those Front-end web portals need to be reachable from the public Internet.

The backend connection to the sensitive data in database servers hosted in the on-premises DCs might need secure connections.

- Scalable security management. IPsec is commonly used to interconnect Cloud GW with enterprises on-premises GWs. For enterprises with large number or branch offices, managing the IPsec's pair-wise security associations among many nodes can be very difficult.

5. Problems with MPLS-based VPNs extending to Hybrid Cloud DCs

Traditional MPLS-based VPNs have been widely deployed as an effective way to support businesses and organizations that require network performance and reliability. MPLS shifted the burden of managing a VPN service from enterprises to service providers. The CPEs attached to MPLS VPN are also simpler and less expensive, since they do not need to manage routes to remote sites; they simply pass all outbound traffic to the MPLS VPN PEs to which the CPEs are attached (albeit multi-homing scenarios require more processing logic on CPEs). MPLS has addressed the problems of scale, availability, and fast recovery from network faults, and incorporated traffic-engineering capabilities.

However, traditional MPLS-based VPN solutions are not optimized for connecting end-users to dynamic workloads/applications in cloud DCs because:

- The Provider Edge (PE) nodes of the enterprise's VPNs might not have direct connection to the third party cloud DCs that are optimal for hosting workloads with the goal of easy access to enterprises' end users.

- It takes a relatively long time to deploy provider edge (PE) routers at new locations. When enterprise's workloads are changed from one cloud DC to another (i.e., removed from one DC and re-instantiated to another location when demand changes), the enterprise branch offices need to be connected to the new cloud DC, but the network service provider might not have PEs located at the new location.

One of the main drivers for moving workloads into the cloud is

the widely available cloud DCs at geographically diverse locations, where apps can be instantiated so that they can be

as close to their end users as possible. When the user base changes, the applications may be moved to a new cloud DC location closest to the new user base.

- Most of the cloud DCs do not expose their internal networks, so the provider MPLS based VPNs cannot reach the workloads natively.

- Many cloud DCs use an overlay to connect their gateways to the workloads inside the DC. There has not been any standard to address the interworking between the Cloud Overlay and the enterprise' existing underlay networks.

Another roadblock is the lack of a standard way to express and enforce consistent security policies to workloads that not only use virtual addresses, but also have a high chance of placement in different locations within the Cloud DC [RFC8192]. The traditional VPN path computation and bandwidth allocation schemes may not be flexible enough to address the need for enterprises to rapidly connect to dynamically instantiated (or removed) workloads and applications regardless of their location/nature (i.e., third party cloud DCs).

6. Problem with using IPsec tunnels to Cloud DCs

As described in the previous section, many Cloud operators expose their gateways for external entities (which can be enterprises themselves) to directly establish IPsec tunnels. Enterprises can also instantiate virtual routers within Cloud DCs to connect to its on-premises devices via IPsec tunnels. If there is only one enterprise location that needs to reach the Cloud DC, an IPsec tunnel is a very convenient solution.

However, many medium-to-large enterprises usually have multiple sites and multiple data centers. For workloads and apps hosted in Cloud DCs, multiple sites need to communicate securely with those Cloud workloads and apps. This section documents some of the issues associated with using IPsec tunnels to connect enterprise' sites with Cloud operator's Gateways.

6.1. Complexity of multi-point any-to-any interconnection

The dynamic workload instantiated in cloud DC needs to communicate with multiple branch offices and on-premises data centers. Most enterprises need multi-point interconnection among multiple locations, as done by MPLS L2/L3 VPNs.

Using IPsec overlay paths to connect all branches & on-premises data centers to cloud DCs requires CPEs to manage routing among Cloud DCs gateways and the CPEs located at other branch locations, which can dramatically increase the complexity of the design, possibly at the cost of jeopardizing the CPE performance.

The complexity of requiring CPEs to maintain routing among other CPEs is one of the reasons why enterprises migrated from Frame Relay based services to MPLS-based VPN services.

MPLS-based VPNs have their PEs directly connected to the CPEs. Therefore, CPEs only need to forward all traffic to the directly attached PEs, which are therefore responsible for enforcing the routing policy within the corresponding VPNs. Even for multi-homed CPEs, the CPEs only need to forward traffic among the directly connected PEs (note: the complexity may vary for IPv6 network). However, when using IPsec tunnels between CPEs and Cloud DCs, the CPEs need to manage the routing for traffic to Cloud DCs, to remote CPEs via VPN, or directly.

6.2. Poor performance over long distance

When enterprise CPEs or gateways are far away from Cloud DC gateways or across country/continent boundaries, performance of IPsec tunnels over the public Internet can be problematic and unpredictable. Even though there are many monitoring tools available to measure delay and various performance characteristics of the network, the measurement for paths over the Internet is passive and past measurements may not represent future performance.

Many cloud providers can replicate workloads in different available zones. An App instantiated in a Cloud DC closest to clients may have to cooperate with another App (or its mirror image) in another region or database server(s) in the on-premises DC. This kind of

coordination requires predicable networking behavior/performance among those locations.

6.3. Scaling Issues with IPsec Tunnels

IPsec can achieve secure overlay connections between two locations over any underlay networks, e.g., between CPEs and Cloud DC Gateways.

If there is only one enterprise location connected to the Cloud gateway, a small number of IPsec tunnels can be configured on-demand between the on-premises DC and the Cloud DC, which is an easy and flexible solution.

However, for multiple enterprise locations to reach workloads hosted in cloud DCs, the Cloud DC gateway needs to maintain multiple IPsec tunnels to all those locations (e.g. hub & spoke topology). For a company with hundreds or thousands of locations, there could be hundreds (or even thousands) of IPsec tunnels terminating at the Cloud DC gateway, which is not only very expensive (because Cloud Operators charge based on connections), but can be very processing intensive for the gateway. Many cloud operators only allow a limited number of IPsec tunnels & bandwidth to each customer. Alternatively, you could use a solution like group encryption where a single IPSec SA is necessary at the GW but the drawback here is key distribution and maintenance of a key server etc.

7. Problems of Using SD-WAN to connect to Cloud DCs

SD-WAN can establish parallel paths over multiple underlay networks between two locations on-demand, for example, two CPEs interconnected by a traditional MPLS VPN ([RFC4364] or [RFC4664]) as well as IPsec [RFC6071] overlay tunnels over Internet.

SD-WAN lets enterprises augment their current VPN network with cost-effective, readily available Broadband Internet connectivity, enabling some traffic offloaded to paths over Internet based on traffic forwarding policy (application-based or otherwise), or when the MPLS VPN connection between the two locations is congested, or otherwise undesirable or unavailable.

7.1. SD-WAN among branch offices vs. interconnect to Cloud DCs

SD-WAN interconnection of branch offices is not as simple as it appears. For an enterprise with multiple sites, using SD-WAN overlay paths among sites requires each CPE to manage all the addresses that local hosts have the potential to reach, i.e. map internal VPN addresses to appropriate SD-WAN paths. This is similar to the

complexity of Frame Relay based VPNs, where each CPE needed to maintain mesh routing for all destinations if they were to avoid an extra hop through a hub router. Even though SD-WAN CPEs can get assistance from a central controller (instead of running a routing protocol) to resolve the mapping between destinations and SD-WAN paths, SD-WAN CPEs are still responsible for routing table maintenance as remote destinations change their attachments, e.g., the dynamic workload in other DCs are de-commissioned or added.

Even though originally envisioned for interconnecting branch offices, SD-WAN offers a very attractive way for enterprises to connect to Cloud DCs.

The SD-WAN for interconnecting branch offices and the SD-WAN for interconnecting to Cloud DCs have some differences:

- SD-WAN for interconnecting branch offices usually have two end-points (e.g. CPEs) controlled by one entity (e.g., a controller or management system operated by the enterprise).
- SD-WAN for interconnecting to Cloud DCs may have CPEs owned or managed by the enterprise and remote end-points being managed or controlled by Cloud DCs (For the ease of description, let's call it asymmetrically managed CPEs).

- Cloud DCs may have different entering points (or devices) with one terminating private direct connect (such as MPLS, or direct line) and other points being the device terminating the IPsec tunnels, as shown in the following diagram.

Therefore, the SD-WAN becomes asymmetric.

```
+------------------------+
|    ,---.         ,---. |
|   (TN-1 )       ( TN-2)|
|    `-+-'  +---+  `-+-' |
|      +----|vR1|----+   |
|           ++--+        |
|            |         +-+-----+
|            |        /Internet\ One path via
|            +-------+ Gateway  +---------------------+
|                     \        /   Internet            \
|                      +-+-----+                        \
+------------------------+                               \
                                                          \
+------------------------+                 native traffic  \
|    ,---.         ,---. |                without encryption|
|   (TN-3 )       ( TN-4)|            over insecure network |
|    `-+-'  +--+   `-+-' |                                  |  +------+
|      +----|vR|-----+   |                                  +--+ CPE  |
|           ++-+         |                                  |  +------+
|            |         +-+-----+                            |
|            |        / virtual\ One path via IPsec Tunnel  |
|            +-------+ Gateway  +-------------------------- +
|                     \        /      Encrypted traffic over|
|                      +-+-----+         public network     |
+------------------------+                                  |
                                                            |
+------------------------+                                  |
```

```
     |    ,---.         ,---. |             Native traffic |
     |   (TN-5 )       ( TN-6)|          without encryption |
     |    `-+-'  +--+   `-+-' |          over secure network|
     |      +----|vR|-----+   |                             |
     |           ++-+         |                             |
     |            |          +-+----+           +-------+   |
     |            |         /        \ Via Direct /customer\  |
     |            +--------+ Gateway  +----------+ gateway  |-----+
     |                      \        /  Connect   \        /
     |                       +-+----+           +-------+
     +-------------------------+
```

Figure 2: Asymmetric Paths SD-WAN

8. End-to-End Security Concerns for Data Flows

When IPsec tunnels from enterprise on-premises CPEs are terminated at the Cloud DC gateway where the workloads or applications are hosted, some enterprises have concerns regarding traffic to/from their workload being exposed to others behind the data center gateway (e.g., exposed to other organizations that have workloads in the same data center).
To ensure that traffic to/from workloads is not exposed to unwanted entities; it is worthwhile to consider having the IPsec tunnels go all the way to the workload (servers, or VMs) within the DC.

9. Requirements for Dynamic Cloud Data Center VPNs

[Editor's note: this section is only a place holder. The requirement listed here are only to stimulate more discussions]

In order to address the aforementioned issues, any solution for enterprise VPNs that includes connectivity to dynamic workloads or applications in cloud data centers should satisfy a set of requirements:

- The solution should allow enterprises to take advantage of the current state-of-the-art in VPN technology, in both traditional MPLS-based VPNs and IPsec-based VPNs (or any combination thereof) that run over-the-top of the public Internet.
- The solution should not require an enterprise to upgrade all their existing CPEs.
- The solution should support scalable IPsec key management among all the nodes.
- The solution needs to support easy and fast VPN connections to dynamic workloads and applications in third party data centers, and easily allow these workloads to migrate both within a data center and between data centers.
- Allow VPNs to provide bandwidth and other performance guarantees.

- Be a cost-effective solution for enterprises to incorporate dynamic cloud-based applications and workloads into their existing VPN environment.

10. Security Considerations

This draft describes the problem space of using SD-WAN to interconnect branch offices with Cloud DCs. As it is a problem statement, the draft itself does not introduce any security concerns. The draft does discuss security requirements as a part of the problem space, particularly in sections 4, 5, and 8.

11. Solution drafts resulting from this work will address particular security concerns inherent in the solution(s), including both protocol aspects and the importance (for example) of securing workloads in cloud DCs and the use of secure interconnection mechanisms.IANA Considerations

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

12. References

12.1. Normative References

12.2. Informative References

[RFC2735] B. Fox, et al "NHRP Support for Virtual Private networks". Dec. 1999.

[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.

[RFC6071] S. Frankel and S. Krishnan, "IP Security (IPsec) and Internet Key Exchange (IKE) Document Roadmap", Feb 2011.

[RFC4364] E. Rosen and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", Feb 2006

[RFC4664] L. Andersson and E. Rosen, "Framework for Layer 2 Virtual Private Networks (L2VPNs)", Sept 2006.

[BGP-SDWAN] L. Dunbar, et al. "BGP Extension for SDWAN Overlay Networks", draft-dunbar-idr-bgp-sdwan-overlay-ext-03, work-in-progress, Nov 2018.

13. Acknowledgments

Many thanks to Ignas Bagdonas, Michael Huang, Liu Yuan Jiao, Katherine Zhao, and Jim Guichard for the discussion and contributions.

Authors' Addresses

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

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

Christian Jacquenet
France Telecom
Rennes, 35000
France
Email: Christian.jacquenet@orange.com

Mehmet Toy
Verizon
One Verizon Way
Basking Ridge, NJ 07920
Email: mehmet.toy@verizon.com