Моделирование и анализ информационных систем. Т. 26, № 2 (2019), с. 203–212
Modeling and Analysis of Information Systems. Vol. 26, No 2 (2019), pp. 203–212

DOI: 10.18255/1818-1015-2019-2-203-212

UDC 517.9

On the Way to SD-WAN Solution

Korsakov S., Sokolov V.

Received April 29, 2018
Revised May 21, 2019
Accepted May 23, 2019

Abstract. The article describes a background and some steps in the implementation of an industrial solution to build manageable mesh overlay network on top of a complete or partially non-manageable underlay network. The overlay network has (or may have) some features from the software defined networks world. We call this solution as SD-WAN Lite to highlight that this solution is not (and will not be in a visible future) a complete SD-WAN solution.

Keywords: software defined networks, mesh networks, routing protocol, network node, encrypted data connection, generic router encapsulation protocol, multi-point GRE tunnel

For citation: Korsakov S., Sokolov V., "On the Way to SD-WAN Solution", *Modeling and Analysis of Information Systems*, **26**:2 (2019), 203–212.

On the authors:
Stanislav Valentinovich Korsakov, orcid.org/0000-0003-2349-7980, CEO,
NETSHe Lab Ltd,
Belinskogo str., 28-75, Yaroslavl, 150047, Russia,
Assistant Professor,
P.G. Demidov Yaroslavl State University,
14 Sovetskaya str., Yaroslavl, 150003, Russia e-mail: sta@stasoft.net

Valery Anatolyevich Sokolov, orcid.org/0000-0003-1427-4937, Doctor, Professor,
P.G. Demidov Yaroslavl State University,
14 Sovetskaya str., Yaroslavl, 150003, Russia, e-mail: valery-sokolov@yandex.ru

Acknowledgments:
This work was supported by RFBR under the Grant №17-07-00823-a.

Introduction

In this article, we use the following agreement on terms:
BGP – Border gateway protocol. The routing protocol and software suite;
OSPF – Open Shortest Path First.The routing protocol and software suite;
HUB – a network node which accepts initial data connections and plays the major role in the routing process;
SPOKE – a network node which initiates data connections;
CONTROLLER – a network node which authorizes other nodes in the network, forms and deploys the node configuration;
IPSec – a method and a protocol set to establish encrypted data connections;
GRE – a generic router encapsulation protocol;

mGRE – a multi-point GRE tunnel;
DM VPN – a method to establish an overlay network with the use of mGRE tunnels, IPSec and some routing engines on the top of mGRE tunnels. DM VPN operates with at least one HUB and one or more SPOKEs;
SDN – software defined networks;
SD-WAN – a kind of methods to establish an overlay network through classical networks (non SDN) with some features taken from SDN;
TE – traffic engineering. A set of methods to manage traffic flows in the network;
VPN – Virtual private network;
NHRP – Next hop resolution protocol;
NAT – Native address translation;
NBMA – Non broadcasting multiply access networks.

The article describes a concept of the solution of the task how to establish a manageable overlay network on the top of non-manageable public IP networks and methods to provide security and TE in such networks.

1. Common Requirements for SD-WAN Lite Solution and Design Stages

At the first stage of the project, we analyzed some trends, existing and announced solutions in the target area. As a result, we have formed some mandatory requirements for implementation. The solution shall:

- be simple in configuration;
- provide a zero touch (or near to zero touch) deploy;
- provide a fault tolerant overlay network configuration;
- establish a mesh network (as HUB-SPOKE connections as well as SPOKE-SPOKE connections) without manual intervention;
- have automatic exchange of routing information;
- support one or two links to the controller and (or) HUBs. The configuration with two links must support a hot link backup;
- support secured data connections and transfer only;
- be able to manage traffic flows between nodes (to find best path, to have backups, to be able to do traffic engineering).

The second stage of the project included search and analysis for existing technologies and methods which can be used in the solution as a proof of concepts, partially or completely. Priority was provided for standard and widely used technologies which have open source implementations at this step. We assumed that the last thing can reduce time and required resources to implement the final solution.

During the second stage, we selected two technologies to use in the final solution:

– DM VPN as a primary method to establish the overlay network and to secure data [3].

– OSPF as the primary routing method because it has traffic engineering features inside.

The selected things match common requirements and allow to establish the secured mesh overlay network on top of the common IP network.

As we may remark, such a secured mesh network is the same as VPN.

The selected things defined a typical solution topology. It includes at least one controller, two or more HUBs and at least one SPOKE. The controller processes SPOKE and HUB requests, authorizes requests, helps to select a reliable access method and data, verify the right LAN network assignment and deploy the result settings to SPOKEs and HUBs. HUB acts as a key point in a virtual private network organization and routing and as a part of traffic engineering process in the future. SPOKE initiates VPN organization, acts as a source for routing and traffic engineering data.

2. DM VPN Overview

Why is DM VPN selected as a basement for SD-WAN Lite solution?

DM VPN is the technology to establish secured VPN over public IP networks such as the Internet.

DM VPN relies on three proven, widely used and standard technologies:

1. The Next Hop Resolution Protocol (NHRP); it creates a distributed (NHRP) mapping database of all the spoke tunnels to real (public interface) addresses [1, 2].

2. The multi-point GRE Tunnel Interface – a single GRE interface to support multiple GRE tunnels.

3. IPSec; it secures data through GRE tunnels.

To simplify and automate the routing management, DM VPN also uses a dynamic routing service like OSPF or BGP.

DM VPN operates with two types of nodes:

– HUB, which represents a central office or a cloud service and controls all VPN connections;

– SPOKE, which represents a branch office and establishes a connection with the central office and (may be) another branches.

DM VPN can implement as star (all traffic in VPN goes through HUB) as well as full mesh (the traffic goes between branches) topologies.

It is possible to prohibit spoke-to-spoke communications by firewall or advanced routing rules in the star topology. This case may be useful to meet different enterprise requirements.

Mesh topology allows to unload hub channels, to reduce the hub load and to reduce the spoke-to-spoke delay.

Spokes may have statically or dynamically assigned IP address, may be behind NAT [1].

The hub must have a reachable IP address and must not be behind NAT [2].

The hub may be addressed through FQDN (Full Qualified Domain Name). In case of FQDN addressing, the use of dynamically assigned IP addresses is allowed with understanding and agreement about service interruption when Hub address changes.

2.1. DM VPN Components. NHRP

NHRP provides registration, resolution and redirect services [1, 2].

NHRP registration:

– a spoke dynamically registers its mapping with NHS;

– supports spokes with dynamic NBMA addresses or NAT.

NHRP resolutions and redirects:

– support building dynamic spoke-to-hub and spoke-to-spoke tunnels;

– create star or full-mesh overlay network topology.

2.2. DM VPN Components. Multi-point GRE Tunnel

It provides single tunnel interface and NHRP source. Single tunnel interface (multi-point):

– a Non-Broadcast Multi-Access (NBMA) network;

– a smaller hub configuration;

– a multicast and broadcast support;

– a dynamic tunnel destination.

Next Hop Resolution Protocol (NHRP):

– VPN IP-to-NBMA IP address mapping;

– short-cut forwarding;

– direct support for dynamic addresses and NAT.

[1] Some types of NAT or NAT settings may prohibit DM VPN Spoke functionality.

[2] The hub may be behind properly configured NAT with multi-point GRE and IPSec bypassing

2.3. DM VPN Components. IPSec

DM VPN builds out a dynamic tunnel overlay network.

IPSec is triggered through "tunnel protection" and works together with NHRP:

- NHRP triggers IPsec before installing new mappings;
- IPSec notifies NHRP when the encryption is ready;
- NHRP installs mappings and sends registration if needed;
- NHRP and IPSec notify each other when a mapping or service assurance is cleared.

2.4. DM VPN Components. Dynamic Routing

The dynamic routing service such as BGP or OSPF is used to establish automatically right routing rules to Hub and every Spoke.

Each available routing service has its own advantages:

- OSPF requires a bit less configuration in simple cases. It allows to use traffic engineering features.
- BGP provides less delay for the overlay network reconfiguration and routing exchange.

2.5. Major Features

DM VPN offers a configuration reduction and no-touch deployment:

- it supports IP Unicast, IP Multicast and dynamic routing protocols;
- it supports remote peers with dynamically assigned addresses;
- it supports spoke routers behind the dynamic NAT and hub routers behind the static NAT.

It has open source NHRP implementation for Linux-based system (OpenNHRP and port of OpenNHRP for Quagga) [3].

Dynamic spoke-to-spoke tunnels for scaling partial- or full-mesh VPNs use IPsec encryption to secure data.

2.6. Typical DM VPN Use Cases [3]

Controlled corporate extranet network.
Star topology. Spoke-to-spoke communications are prohibited (or controlled) by firewall or routing rules.

Meshed corporate network.
Meshed topology. Spoke-to-spoke communications are allowed to reduce latency and hub load.

DM VPN as the backup for a MPLS network.
Meshed topology. Spoke-to-spoke communications are allowed to reduce latency and hub load. DM VPN is used only when the primary MPLS network is down.

[3]Both implementations do not support authentication and NBMA.

3. Traffic Engineering Elements in the Solution

BGP does not have TE elements and was rejected as routing engine for solution. In case of OSPF, some trivial traffic engineering elements exist 'out of box':

- best path selection;
- route injection;
- interface cost and priority [7];
- route (HUB) metrics;
- different areas.

Also, OSPF has more TE elements like bandwidth and delay based to implement more reliable traffic management in the future [4]-[6]. Such a development must be based on a non-manageable entity of an underlay network and on a big delays/slow feedback between nodes in an overlay network. We think that this area mostly requires statistical methods and is perspective for further study.

As a result of development, we have got solution (software) which is based on Linux, implements DM VPN as primary overlay method and OSPF as primary routing method.

3.1. DM VPN Implementation Details and Limitations in SD WAN Lite

Current solution provides:

- simplifield (few steps) configuration;
- Hub and Spoke functionality out of the box;
- PSK protected IPSec tunnels;
- OSPF or BGP routing over the overlay network. Routing does not require any special configuration.

At the same time, we have some limitations:

- we support only dual hub configuration for fault-tolerance;
- we have compatibility with CISCO Hub and Spoke only. Another vendor Hub and Spoke compatibility with OSPF has not been tested yet.

3.2. SD WAN Lite Implementation Background

To get the solution, we used some existing software like Linux kernel as OS basement, Quagga as BGP and OSPF routing daemon, OpenNHRP port to Quagga as NHRP daemon; StrongSWAN as IPSec management daemon.

In the development process, we encountered some issues in Linux-based DM VPN implementation:

1. Multicast packets could not forward through the mGRE tunnel in Linux. It disallows to run OSPF through mGRE tunnels.

2. NHRP implementation does not support Cisco-style authentication.

3. NHRP implementation works only with host networks (/32 or /128).

4. OSPF implementation in Quagga cannot run over unconnected networks (Through /32 networks. According to the RFCs). It disallows the distribution of OSPF hello and update packets through mGRE tunnels.

We have made some patches for Quagga OSPF and NHRP daemons to resolve this issues.

For example, we implemented Cisco-style authentication for Quagga NHRP (issue 2) and made patch to Quagga OSPF to allow working through interfaces with /32 (for Ipv4) and /128 (for Ipv6) masks (issue 4).

To complete OSPF support in DM VPN, we rewrote tunnel support in Linux kernel to fix multicast packet transfer through multipoint tunnels (issue 1).

The patch code does lookup in routing and neighbor tables to find valid outer address for every multi-point tunnel endpoints and to send a fixed copy of the original packet with a discovered outer address to every endpoint.

3.3. DM VPN Cons as a Method to Build an Overlay Network

As DM VPN has its own pros, as well as its own cons. And the main con is a trouble (up to full impossibility) to traverse some NAT and firewalls.

E.g. it will not work in case when IPSec is prohibited by carrier or will have troubles in mesh topology for SPOKE with low traffic.

To overcome this issue, we must add a component which meets the next requirements:

– it is able to traverse most of NAT types and firewalls;

– it is able to provide similar security for data transfer as IPSec;

– it is able to run the same routing engine as DM VPN part.

3.4. Alternative to DM VPN Method to Build Overlay Networks

We have found that OpenVPN in TCP mode can be successfully used to traverse most of firewalls and NATs which are not traversal by DM VPN (especially with using port 443).

OpenVPN provides strong encrypted connections and security for data transfer through these connections is similar to IPSec.

OpenVPN can be used with OSPF or BGP as an overlay routing engine.

Depending on these facts, OpenVPN can be used as the second method to establish an overlay network in SD-WAN Lite solution in addition to DM VPN.

The OpenVPN place in the solution:

– when SPOKE sends a registration request to the controller, it reports properties of SPOKE Internet connections;

– the controller checks which connections do not have a real IP address and (or) exist behind the NAT;

– after checking, the controller assigns to use DM VPN and starts to measure connection time. If SPOKE does not establish DM VPN connection in some time, it requests HUB again and gets settings to establish OpenVPN connection as the last resort.

3.5. A Big Fat Con Against OpenVPN Method

The use of OpenVPN implements the standard star-like topology between HUB and SPOKE and reduces an area for any kinds of traffic engineering.

We have spoke about CONTROLLER above. Let us explain place and role for CONTROLLER in the solution below.

CONTROLLER has relations with HUB(s) as well as with SPOKE(s). Such relations may be represented as charts.

3.5.1. HUB-Controller Interactions Chart

– HUB sends registration requests to the controller;

– the controller validates requests, forms and deploys HUB configuration;

– HUB reports traffic and routing properties to the controller;

– the controller forms management actions for HUB (affects QoS and routing).

3.5.2. SPOKE-Controller Iterations Chart

– SPOKE sends registration requests to the controller;

– the controller validates requests, forms and deploys SPOKE configuration;

– SPOKE reports link, traffic and routing properties to the controller;

– the controller forms management actions for SPOKE (affects QoS and routing. For example, through the interface cost and (or) TE metrics);

– the controller forms management actions for HUB (affects routing).

3.5.3. Base Solution Security

Initial transactions between a SPOKE candidate and the controller and a HUB candidate and the controller go encrypted over HTTPS with the mandatory certificate validation.

All transactions and data transfer between HUB and SPOKE and SPOKE and SPOKE go over IPSec encrypted tunnels or OpenVPN encrypted tunnels.

Security key points are:

– a key to access the controller;

– passphrases and (or) certificates which are deployed from the controller;

– NHRP secrets which are deployed from the controller;

– a root certificate which is used to sign the controller certificate and which must be kept in every HUB and SPOKE.

Physical access to any device provides ability to stole the root certificate, implements MITM attack and obtain any credentials to connect a rogue device in the network and provides some destructive actions in a future.

3.6. Security Enhancement

To improve security, the controller will change and redistribute secrets, passphrases and (or) node certificates by some algorithm (for example, by schedule).

To minimize losses in case of a rogue device, the solution must have:

– a feature to divide SPOKEs to "trusted" and "untrusted". Only trusted SPOKEs can establish meshed connections, can have full routing updates, can submit traffic engineering data, can be involved in the certificate and credentials exchanging process. Untrusted SPOKEs can only establish SPOKE-to-HUB connections, can have aggregated routes only to minimize a potential destructive effect. Their data will be ignored in traffic engineering and management processes. SPOKE state exchange is performed manually;

– a feature to upload a new root certificate to all nodes and to regenerate any inherited certificates;

– a set of features to detect possible rogue SPOKEs;

– a feature to allow console/WebUI access to SPOKE after the central user authentication (for example, against RADIUS server in the controller). Some failed logins will mark SPOKE as a possible rogue device;

– route announce validation. We assume that non-rogue SPOKE will announce non-overlapped networks and this network set will be stable in the time, will not try to forward any kind of traffic through the device, will not try to establish 'strange' routes.

4. Conclusion

When writing the article, we have got a proof of the concept for SD-WAN Lite solution where:

– the complete Cisco-style DM VPN solution is working and well tested;

– the Open VPN TCP mode part is working and well tested;

– OSPF-based routing with some trivial traffic engineering elements is working;

– the proof of the concept for the solution controller is obtained.

References

[1] Luciani J., Katz D., Piscitello D., Cole B., Doraswamy N., "RFC 2332: NBMA Next Hop Resolution Protocol (NHRP)", *IETF, April*, 1998, https://tools.ietf.org/html/rfc2332.

[2] Fox B., Petri B., "RFC 2735: NHRP Support for Virtual Private Networks", *Request for Comments,'Online*, 1999, https://tools.ietf.org/html/rfc2735.

[3] *Cisco. Dynamic Multipoint VPN Configuration guide.*, https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/sec_conn_dmvpn/configuration/xe-16/sec-conn-dmvpn-xe-16-book.html.

[4] Katz D.,Kompella K., Yeung D., "RFC 3630: Traffic Engineering (TE) Extensions to OSPF Version 2", *Internet RFCs*, **1** (2003), 1, https://tools.ietf.org/html/rfc3630.

[5] Giacalone S., Ward D., Drake J., Atlas A., Previdi S., "RFC-7471. OSPF Traffic Engineering (TE) Metric Extensions", 2015, https://tools.ietf.org/html/rfc7471.

[6] Retvari G., Cinkler T., "Practical OSPF traffic engineering", *IEEE Communications Letters*, **8**:11 (2004), 689-691.

[7] Nemeth K., Korosi A., Retvari G., "Optimal OSPF traffic engineering using legacy Equal Cost Multipath load balancing", *2013 IFIP Networking Conference. – IEEE*, 2013, 1-9, https://www.researchgate.net/publication/261208279_Optimal_OSPF_traffic_engineering_using_legacy_Equal_Cost_Multipath_load_balancing.

Корсаков С. В., Соколов В. А., "На пути к SD-WAN решению", *Моделирование и анализ информационных систем*, **26**:2 (2019), 203–212.

DOI: 10.18255/1818-1015-2019-2-203-212

Аннотация. В статье описываются предпосылки и некоторые этапы реализации промышленного решения по построению управляемой ячеистой оверлейной сети поверх полностью или частично неуправляемой опорной сети. Оверлейная сеть имеет (или может иметь) некоторые функции из мира программно-определяемых сетей. Мы называем это решение SD-WAN Lite, чтобы подчеркнуть, что это решение не является (и не будет в обозримом будущем) полным SD-WAN решением.

Ключевые слова: программно-определяемые сети, ячеистые сети, протокол маршрутизации, сетевой узел, зашифрованное соединение данных, общий протокол инкапсуляции маршрутизатора, многоточечный GRE туннель

Об авторах:
Станислав Валентинович Корсаков, orcid.org/0000-0003-2349-7980, директор,
ООО «Нетше лаб»,
ул. Белинского, 28-75, г. Ярославль, 15004,7 Россия,
ассистент кафедры теоретической информатики,
Ярославский государственный университет им. П.Г. Демидова,
ул. Советская, 14, г. Ярославль, 150003, Россия, e-mail: sta@stasoft.net

Валерий Анатольевич Соколов, orcid.org/0000-0003-1427-4937, докт. физ.-мат. наук., профессор, заведующий кафедрой теоретической информатики,
Ярославский государственный университет им. П.Г. Демидова, ул. Советская, 14, г. Ярославль, 150003, Россия, e-mail: valery-sokolov@yandex.ru

Благодарности:
Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта №17-07-00823-а.