Internet Engineering Task Force (IETF)
Request for Comments: 7018
Category: Informational
ISSN: 2070-1721

V. Manral
HP
S. Hanna
Juniper

September 2013

Auto-Discovery VPN Problem Statement and Requirements

Abstract

This document describes the problem of enabling a large number of systems to communicate directly using IPsec to protect the traffic between them. It then expands on the requirements for such a solution.

Manual configuration of all possible tunnels is too cumbersome in many such cases. In other cases, the IP addresses of endpoints change, or the endpoints may be behind NAT gateways, making static configuration impossible. The Auto-Discovery VPN solution will address these requirements.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are a candidate for any level of Internet Standard; see Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7018.

Table of Contents

1. Introduction

IPsec [RFC4301] is used in several different cases, including tunnel-mode site-to-site VPNs and remote access VPNs. Both tunneling modes for IPsec gateways and host-to-host transport mode are supported in this document.

The subject of this document is the problem presented by large-scale deployments of IPsec and the requirements on a solution to address the problem. These may be a large collection of VPN gateways connecting various sites, a large number of remote endpoints connecting to a number of gateways or to each other, or a mix of the two. The gateways and endpoints may belong to a single administrative domain or several domains with a trust relationship.

Section 4.4 of RFC 4301 describes the major IPsec databases needed for IPsec processing. It requires extensive configuration for each tunnel, so manually configuring a system of many gateways and endpoints becomes infeasible and inflexible.

The difficulty is that a lot of configuration mentioned in RFC 4301 is required to set up a Security Association. The Internet Key Exchange Protocol (IKE) implementations need to know the identity and credentials of all possible peer systems, as well as the addresses of hosts and/or networks behind them. A simplified mechanism for dynamically establishing point-to-point tunnels is needed. Section 2 contains several use cases that motivate this effort.

1.1. Terminology

Auto-Discovery Virtual Private Network (ADVPN) - A VPN solution that enables a large number of systems to communicate directly, with minimal configuration and operator intervention, using IPsec to protect communication between them.

Endpoint - A device that implements IPsec for its own traffic but does not act as a gateway.

Gateway - A network device that implements IPsec to protect traffic flowing through the device.

Point-to-Point - Communication between two parties without active participation (e.g., encryption or decryption) by any other parties.

Hub - The central point in a star topology/dynamic full-mesh topology, or one of the central points in the full-mesh style VPN, i.e., a gateway to which multiple other hubs or spokes connect. The hubs usually forward traffic coming from encrypted links to other encrypted links, i.e., there are no devices connected to them in the clear.

Spoke - The endpoint in a star topology/dynamic full-mesh topology or gateway that forwards traffic from multiple cleartext devices to other hubs or spokes, and some of those other devices are connected to it in the clear (i.e., it encrypts data coming from cleartext devices and forwards it to the ADVPN).

ADVPN Peer - Any member of an ADVPN, including gateways, endpoints, hubs, or spokes.

Star Topology - Topology in which there is direct connectivity only between the hub and spoke, and where communication between the 2 spokes happens through the hub.

Allied and Federated Environments - Environments where we have multiple different organizations that have close associations and need to connect to each other.

Full-Mesh Topology - Topology in which there is direct connectivity between every spoke to every other spoke, without the traffic between the spokes having to be redirected through an intermediate hub device.

Dynamic Full-Mesh Topology - Topology in which direct connections exist in a hub-and-spoke manner but dynamic connections are created/removed between the spokes on an as-needed basis.

Security Association (SA) - Defined in [RFC4301].

1.2. Conventions Used in This Document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

2. Use Cases

This section presents the key use cases for large-scale point-to-point VPNs.

In all of these use cases, the participants (endpoints and gateways) may be from a single organization (administrative domain) or from multiple organizations with an established trust relationship. When multiple organizations are involved, products from multiple vendors are employed, so open standards are needed to provide interoperability. Establishing communications between participants with no established trust relationship is out of scope for this effort.

2.1. Use Case 1: Endpoint-to-Endpoint VPN

Two endpoints wish to communicate securely via a point-to-point SA.

The need for secure endpoint-to-endpoint communications is often driven by a need to employ high-bandwidth, low-latency local connectivity instead of using slow, expensive links to remote gateways. For example, two users in close proximity may wish to place a direct, secure video or voice call without needing to send

the call through remote gateways, as the remote gateways would add latency to the call, consume precious remote bandwidth, and increase overall costs. Such a use case also enables connectivity when both users are behind NAT gateways. Such a use case ought to allow for seamless connectivity even as endpoints roam and even if they are moving out from behind a NAT gateway, from behind one NAT gateway to behind another, or from a standalone position to behind a NAT gateway.

In a star topology, when two endpoints communicate, they need a mechanism for authentication such that they do not expose themselves to impersonation by the other spoke endpoint.

2.2. Use Case 2: Gateway-to-Gateway VPN

A typical Enterprise traffic model follows a star topology, with the gateways connecting to each other using IPsec tunnels.

However, for voice and other rich media traffic that require a lot of bandwidth or is performance sensitive, the traffic tromboning (taking a suboptimal path) to the hub can create traffic bottlenecks on the hub and can lead to an increase in cost. A fully meshed solution would make best use of the available network capacity and performance, but the deployment of a fully meshed solution involves considerable configuration, especially when a large number of nodes are involved. It is for this purpose that spoke-to-spoke tunnels are dynamically created and torn down. For the reasons of cost and manual error reduction, it is desired that there be minimal configuration on each gateway.

The solution ought to work in cases where the endpoints are in different administrative domains that have an existing trust relationship (for example, two organizations that are collaborating on a project may wish to join their networks while retaining independent control over configuration). It is highly desirable that the solution works for the star, full-mesh, and dynamic full-mesh topologies.

The solution ought to also address the case where gateways use dynamic IP addresses.

Additionally, the routing implications of gateway-to-gateway communication need to be addressed. In the simple case, selectors provide sufficient information for a gateway to forward traffic appropriately. In other cases, additional tunneling (e.g., Generic Routing Encapsulation (GRE)) and routing (e.g., Open Shortest Path First (OSPF)) protocols are run over IPsec tunnels, and the configuration impact on those protocols needs to be considered.

There is also the case where Layer 3 Virtual Private Networks (L3VPNs) operate over IPsec tunnels.

When two gateways communicate, they need to use a mechanism for authentication such that they do not expose themselves to the risk of impersonation by the other entities.

2.3. Use Case 3: Endpoint-to-Gateway VPN

A mobile endpoint ought to be able to use the most efficient gateway as it roams in the Internet.

A mobile user roaming on the Internet may connect to a gateway that, because of roaming, is no longer the most efficient gateway to use (reasons could be cost, efficiency, latency, or some other factor). The mobile user ought to be able to discover and then connect to the current, most efficient gateway in a seamless way without having to bring down the connection.

3. Inadequacy of Existing Solutions

Several solutions exist for the problems described above. However, none of these solutions is adequate, as described here.

3.1. Exhaustive Configuration

One simple solution is to configure all gateways and endpoints in advance with all the information needed to determine which gateway or endpoint is optimal and to establish an SA with that gateway or endpoint. However, this solution does not scale in a large network with hundreds of thousands of gateways and endpoints, especially when multiple administrative domains are involved and things are rapidly changing (e.g., mobile endpoints). Such a solution is also limited by the smallest endpoint/gateway, as the same exhaustive configuration is to be applied on all endpoints/gateways. A more dynamic, secure, and scalable system for establishing SAs between gateways is needed.

3.2. Star Topology

The most common way to address a part of this problem today is to use what has been termed a "star topology". In this case, one or a few gateways are defined as "hub gateways", while the rest of the systems (whether endpoints or gateways) are defined as "spokes". The spokes never connect to other spokes. They only open tunnels with the hub gateways. Also, for a large number of gateways in one administrative domain, one gateway may be defined as the hub, and the rest of the gateways and remote access clients connect only to that gateway.

This solution, however, is complicated by the case where the spokes use dynamic IP addresses and DNS with dynamic updates needs to be used. It is also desired that there is minimal to no configuration on the hub as the number of spokes increases and new spokes are added and deleted randomly.

Another problem with the star topology is that it creates a high load on the hub gateways, as well as on the connection between the spokes and the hub. This load impacts both processing power and network bandwidth. A single packet in the hub-and-spoke scenario can be encrypted and decrypted multiple times. It would be much preferable if these gateways and clients could initiate tunnels between them, bypassing the hub gateways. Additionally, the path bandwidth to these hub gateways may be lower than that of the path between the spokes. For example, two remote access users may be in the same building with high-speed WiFi (for example, at an IETF meeting). Channeling their conversation through the hub gateways of their respective employers seems extremely wasteful, given that a more optimal direct path exists.

The challenge is to build large-scale IPsec-protected networks that can dynamically change with minimal administrative overhead.

3.3. Proprietary Approaches

Several vendors offer proprietary solutions to these problems. However, these solutions offer no interoperability between equipment from one vendor and another. This means that they are generally restricted to use within one organization, and it is harder to move away from such solutions, as the features are not standardized. Besides, multiple organizations cannot be expected to all choose the same equipment vendor.

4. Requirements

This section defines the requirements on which the solution will be based.

4.1. Gateway and Endpoint Requirements

1. For any network topology (star, full mesh, and dynamic full mesh), when a new gateway or endpoint is added, removed, or changed, configuration changes are minimized as follows. Adding or removing a spoke in the topology MUST NOT require configuration changes to hubs other than where the spoke was connected and SHOULD NOT require configuration changes to the hub to which the spoke was connected. The changes also MUST NOT require configuration changes in other spokes.

Specifically, when evaluating potential proposals, we will compare them by looking at how many endpoints or gateways must be reconfigured when a new gateway or endpoint is added, removed, or changed and how substantial this reconfiguration is, in addition to the amount of static configuration required.

This requirement is driven by use cases 1 and 2 and by the scaling limitations pointed out in Section 3.1.

2. ADVPN Peers MUST allow IPsec tunnels to be set up with other members of the ADVPN without any configuration changes, even when peer addresses get updated every time the device comes up. This implies that Security Policy Database (SPD) entries or other configuration based on a peer IP address will need to be automatically updated, avoided, or handled in some manner to avoid a need to manually update policy whenever an address changes.

3. In many cases, additional tunneling protocols (e.g., GRE) or routing protocols (e.g., OSPF) are run over the IPsec tunnels. Gateways MUST allow for the operation of tunneling and routing protocols operating over spoke-to-spoke IPsec tunnels with minimal or no configuration impact. The ADVPN solution SHOULD NOT increase the amount of information required to configure protocols running over IPsec tunnels.

4. In the full-mesh and dynamic full-mesh topologies, spokes MUST allow for direct communication with other spoke gateways and endpoints. In the star topology mode, direct communication between spokes MUST be disallowed.

   This requirement is driven by use cases 1 and 2 and by the limitations of a star topology as pointed out in Section 3.2.

5. ADVPN Peers MUST NOT have a way to get the long-term authentication credentials for any other ADVPN Peers. The compromise of an endpoint MUST NOT affect the security of communications between other ADVPN Peers. The compromise of a gateway SHOULD NOT affect the security of the communications between ADVPN Peers not associated with that gateway.

   This requirement is driven by use case 1. ADVPN Peers (especially spokes) become compromised fairly often. The compromise of one ADVPN Peer SHOULD NOT affect the security of other unrelated ADVPN Peers.

6. Gateways SHOULD allow for seamless handoff of sessions in cases where endpoints are roaming, even if they cross policy boundaries. This would mean the data traffic is minimally affected even as the handoff happens. External factors like firewalls and NAT boxes that will be part of the overall solution when ADVPN is deployed will not be considered part of this solution.

   Such endpoint roaming may affect not only the endpoint-to-endpoint SA but also the relationship between the endpoints and gateways (such as when an endpoint roams to a new network that is handled by a different gateway).

   This requirement is driven by use case 1. Today's endpoints are mobile and transition often between different networks (from 4G to WiFi and among various WiFi networks).

7. Gateways SHOULD allow for easy handoff of a session to another gateway, to optimize latency, bandwidth, load balancing, availability, or other factors, based on policy.

   This ability to migrate traffic from one gateway to another applies regardless of whether the gateways in question are hubs or spokes. It even applies in the case where a gateway (hub or spoke) moves in the network, as may happen with a vehicle-based network.

   This requirement is driven by use case 3.

8. Gateways and endpoints MUST have the capability to participate in an ADVPN even when they are located behind NAT boxes. However, in some cases they may be deployed in such a way that they will not be fully reachable behind a NAT box. It is especially difficult to handle cases where the hub is behind a NAT box. When the two endpoints are both behind separate NATs, communication between these spokes SHOULD be supported using workarounds such as port forwarding by the NAT or detecting when two spokes are behind uncooperative NATs, and using a hub in that case.

   This requirement is driven by use cases 1 and 2. Endpoints are often behind NATs, and gateways sometimes are. IPsec SHOULD continue to work seamlessly regardless, using ADVPN techniques whenever possible and providing graceful fallback to hub-and-spoke techniques as needed.

9. Changes such as establishing a new IPsec SA SHOULD be reportable and manageable. However, creating a MIB or other management technique is not within scope for this effort.

   This requirement is driven by manageability concerns for all the use cases, especially use case 2. As IPsec networks become more dynamic, management tools become more essential.

10. To support allied and federated environments, endpoints and gateways from different organizations SHOULD be able to connect to each other.

    This requirement is driven by demand for all the use cases in federated and allied environments.

11. The administrator of the ADVPN SHOULD allow for the configuration of a star, full-mesh, or partial full-mesh topology, based on which tunnels are allowed to be set up.

    This requirement is driven by demand for all the use cases in federated and allied environments.

12. The ADVPN solution SHOULD be able to scale for multicast traffic.

    This requirement is driven by use case 2, where the amount of rich media multicast traffic is increasing.

13. The ADVPN solution SHOULD allow for easy monitoring, logging, and reporting of the dynamic changes to help with troubleshooting such environments.

    This requirement is driven by demand for all the use cases in federated and allied environments.

14. There is also the case where L3VPNs operate over IPsec tunnels, for example, Provider-Edge-based VPNs. An ADVPN MUST support L3VPNs as applications protected by the IPsec tunnels.

    This requirement is driven by demand for all the use cases in federated and allied environments.

15. The ADVPN solution SHOULD allow the enforcement of per-peer QoS in both the star and full-mesh topologies.

    This requirement is driven by demand for all the use cases in federated and allied environments.

16. The ADVPN solution SHOULD take care of not letting the hub be a single point of failure.

    This requirement is driven by demand for all the use cases in federated and allied environments.

5. Security Considerations

This is a problem statement and requirements document for the ADVPN solution and in itself does not introduce any new security concerns. The solution to the problems presented in this document may involve dynamic updates to databases defined by RFC 4301, such as the Security Policy Database (SPD) or the Peer Authorization Database (PAD).

RFC 4301 is silent about the way these databases are populated, and it is implied that these databases are static and preconfigured by a human. Allowing dynamic updates to these databases must be thought out carefully because it allows the protocol to alter the security policy that the IPsec endpoints implement.

One obvious attack to watch out for is stealing traffic to a particular site. The IP address for www.example.com is 192.0.2.10. If we add an entry to an IPsec endpoint's SPD that says that traffic to 192.0.2.10 is protected through peer Gw-Mallory, then this allows Gw-Mallory to either pretend to be www.example.com or proxy and read all traffic to that site. Updates to this database require a clear trust model.

Hubs can be a single point of failure that can cause loss of connectivity of the entire system; this can be a big security issue. Any ADVPN solution design should take care of these concerns.

6. Acknowledgements

Many people have contributed to the development of this problem statement. While we cannot thank all contributors, some have played an especially prominent role. Yoav Nir, Yaron Sheffer, Jorge Coronel Mendoza, Chris Ulliott, and John Veizades wrote the document upon which this specification was based. Geoffrey Huang, Toby Mao, Suresh Melam, Praveen Sathyanarayan, Andreas Steffen, Brian Weis, Lou Berger, and Tero Kivinen provided essential input.

7. Normative References

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

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, December 2005.

Authors' Addresses

Vishwas Manral
Hewlett-Packard Co.
3000 Hanover St.
Palo Alto, CA 94304
USA

EMail: vishwas.manral@hp.com

Steve Hanna
Juniper Networks, Inc.
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
USA

EMail: shanna@juniper.net