Provider Provisioned VPN WG
Internet Draft
draft-ouldbrahim-bgpvpn-auto-01.txt
Expiration Date: September 2001

Hamid Ould-Brahim
Bryan Gleeson
Peter Ashwood-Smith
Nortel Networks

Eric C. Rosen
Cisco Systems

Yakov Rekhter
Juniper Networks

Luyuan Fang
AT&T

Jeremy De Clercq
Alcatel

March 2001

Using BGP as an Auto-Discovery Mechanism for Network-based VPNs

Status of this Memo

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.

Abstract

In any Network-Based VPN (NBVPN) scheme, the Provider Edge (PE) routers attached to a common VPN must exchange certain information as a prerequisite to establish VPN-specific connectivity. In [RFC2547-bis], VPN-specific routes are exchanged, along with the information needed to enable a PE to determine which routes belong to which VPNs. In [VPN-VR], VR addresses must be exchanged, along with the information needed to enable the PEs to determine which VRs are in the same VPN ("membership"), and which of those VRs are to have VPN connectivity ("topology"). Once the VRs are reachable through the tunnels, routes ("reachability") are then exchanged by running existing routing protocol per VPN basis. The purpose of this draft is to define a common BGP based auto-discovery mechanism used for both the virtual router [VPN-VR] and [RFC2547-bis] architectures. Each scheme uses the mechanism to automatically discover the information needed by that particular scheme. Interworking scenarios between [RFC2547-bis] and the virtual router models are also discussed.

1. Introduction

In any Network-Based VPN (NBVPN) scheme, the Provider Edge (PE) routers attached to a common VPN must exchange certain information as a prerequisite to establish VPN-specific connectivity. In [RFC2547-bis], VPN-specific routes are exchanged, along with the information needed to enable a PE to determine which routes belong to which VPNs. In [VPN-VR], virtual router (VR) addresses must be exchanged, along with the information needed to enable the PEs to determine which VRs are in the same VPN ("membership"), and which of those VRs are to have VPN connectivity ("topology"). Once the VRs are reachable through the tunnels, routes ("reachability") are then exchanged by running existing routing protocols per VPN basis.

The purpose of this draft is to define a common BGP based auto-discovery mechanism used for both the virtual router [VPN-VR] and [RFC2547-bis] architectures. Each scheme uses the mechanism to automatically discover the information needed by that particular scheme. The BGP multiprotocol extension attributes are used to carry either the virtual router or the RFC2547 auto-discovery information. Interworking scenarios between [RFC2547-bis] and the virtual router models are also discussed.

2. Network Based VPNs Reference Model

Both the virtual router and [RFC2547-bis] architectures are using a network reference model as illustrated in figure 1.

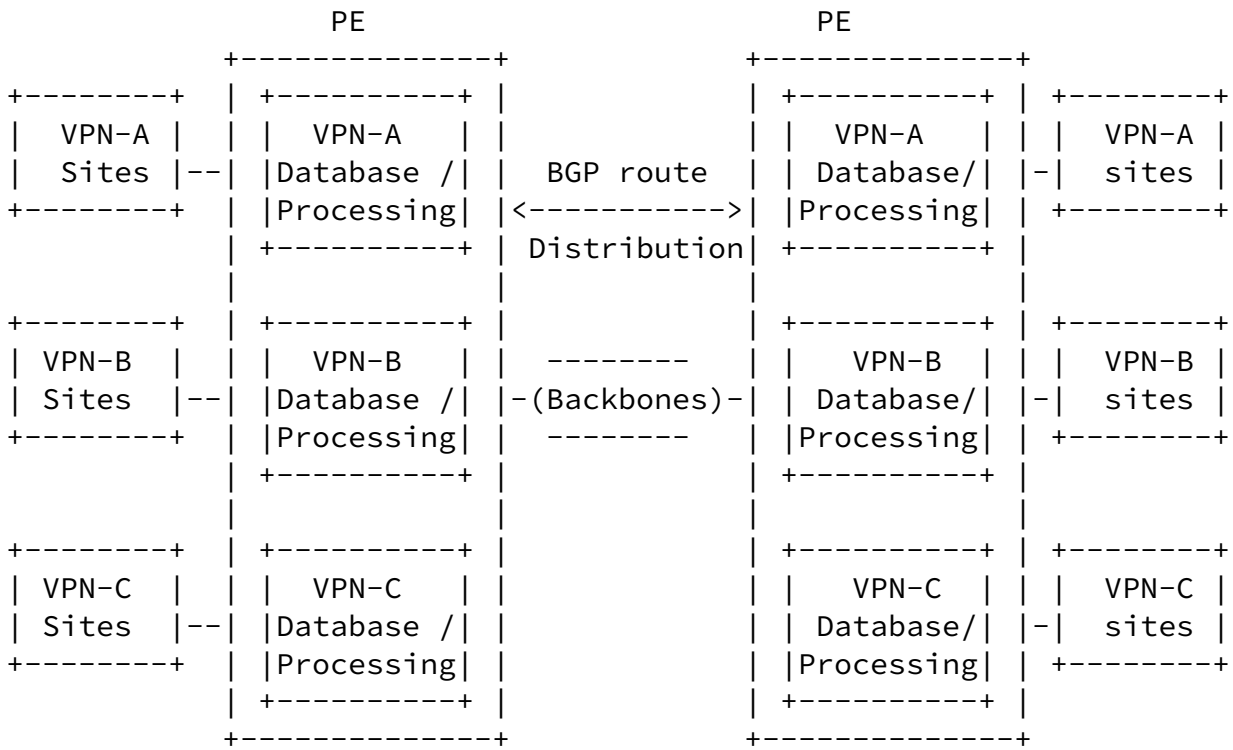

Figure 1: Network based VPN Reference Model

It is assumed that the PE routers can use BGP to distribute information to each other. This may be via direct IBGP peering, via

direct EBGP peering, via multihop BGP peering, through intermediaries such as Route Reflectors, through a chain of intermediate BGP connections, etc. It is assumed also that the PE knows what architecture it is supporting (either the virtual router, or [RFC2547-bis] architectures, or both).

3. Carrying VPN information in BGP Multi-Protocol Extension Attributes

The BGP-4 multiprotocol extensions are used to carry various information about VPNs for both architectures. This is done as follows. The NLRI is a VPN-IP address or a labeled VPN-IP address. VPN-specific information associated with the NLRI is encoded either as attributes of the NLRI, or as part of the NLRI itself, or both.

The address prefix in the NLRI field is ALWAYS within the VPN address space, and therefore MUST be unique within the VPN. The address specified in the BGP next hop attribute, on the other hand, is in the service provider addressing space.

In the case of the virtual router, the NLRI address prefix is an address of one of the virtual routers configured on the PE. Thus this mechanism allows the virtual routers to discover each other, to set up adjacencies and tunnels to each other, etc. In the case of [RFC2547-bis], the NLRI prefix represents a route to an arbitrary system or set of systems within the VPN.

4. Interpretation of VPN Information in the [RFC2547-bis] model

The [RFC2547-bis] model interprets the NLRI reachability information. The BGP attributes (in particular, the Route Target Extended Community) are used by the PE routers to assign the routes to particular VPN database/processing contexts, and hence implicitly determine the topology. The BGP Next Hop attribute specifies the remote end point of the tunnel to be used when sending packets whose destination addresses match the corresponding NLRI. For details, see [RFC2547-bis].

5. Interpretation of VPN Information in the [VPN-VR] model

5.1 Membership Discovery

The VPN-ID format as defined in [RFC-2685] is used to identify a VPN. All virtual routers that are members of a specific VPN share the same VPN-ID. A VPN-ID is carried in the NLRI in order to associate a particular VR address to the VPN.

5.1.1 Encoding of the VPN-ID in the NLRI

For the virtual router model, the VPN-ID is carried within the route distinguisher (RD) field. In order to hold the 7-bytes VPN-ID, the first byte of RD type field is used to indicate the existence of the VPN-ID format. A value of 0x80 in the first byte of RD's type field indicates that the RD field is carrying the VPN-ID format. In this case, the type field range 0x8000-0x80ff will be reserved for the virtual router case.

5.1.2 VPN-ID Extended Community

A new extended community is used to carry the VPN-ID. The first byte of the VPN-ID extended community will be used to indicate the presence of the VPN-ID format. A value of 0x20 indicates that the remaining 7 bytes following the first byte of the type field holds a VPN-ID value. In this case an extended community with type field in the range of 0x2000-0x20ff will be used exclusively for the virtual router case. The BGP UPDATE message will carry information for a single VPN (when used with a VPN-ID extended community). The use of VPN-ID extended community allows a PE to perform route filtering per VPN basis.

5.2 VPN Topology Information

A new extended community is used to indicate different VPN topology values. A type 0x0020 represents the VPN topology field. The first two bytes (of the remaining 6 bytes) are reserved. The actual topology values are carried within the remaining four bytes. The following topology values are defined:

Value	Topology Type
1	"Hub"

2	"Spoke"
3	"Mesh"

Arbitrary values can also be used to allow specific topologies to be constructed. VPN connectivity between two VRs within the same VPN is achieved if and only if at least one of them is a hub (the other is a hub or a spoke), or if both VRs are part of a full mesh VPN topology.

5.3 Tunnel Discovery

Network-based VPNs must be implemented through some form of tunneling mechanism, where the packet formats and/or the addressing used within the VPN can be unrelated to that used to route the tunneled packets across the backbone. There are numerous tunneling mechanisms that can be used by a network based VPN (e.g., IP/IP [RFC-2003], GRE tunnels [RFC-1701], IPSec [RFC-2401], and MPLS tunnels [MPLS-ARCH]). Each of these tunnels allows for opaque transport of frames as packet payload across the backbone, with forwarding disjoint from the address fields of the encapsulated packets. A provider edge router may terminate multiple type of tunnels and forward packets between these tunnels and other network interfaces in different ways.

BGP can be used to carry tunnel endpoint addresses between edge routers. Depending on the type of tunneling mechanism used on a PE, tunnel endpoint addresses can be used to establish either IPSec, IP in IP, or GRE tunnels.

The BGP next hop will carry the service provider tunnel endpoint address. As an example, if IPSec is used as tunneling mechanism, the IPSec tunnel remote address will be discovered through BGP, and the actual tunnel establishment is achieved through IPSec signaling protocol.

When MPLS tunneling is used, a label carried in the NLRI will indicate which VR can be reached through the address carried in the

NLRI. A single service provider tunnel endpoint address can be used to reach multiple VRs (therefore multiple VPNs).

6. Virtual Router and [RFC2547-bis] Interworking Scenarios

Two interwoking scenarios are considered when the network is using both virtual routers and [RFC2547-bis]. The first scenario is a CE-PE relationship between a PE (implementing [RFC2547-bis]), and a VR appearing as a CE to the PE. The connection between the VR, and the PE can be either direct connectivity, or through a tunnel (e.g., IPSec).

The second scenario is when a PE is implementing both architectures. In this particular case, a single BGP session configured on the service provider network can be used to advertise either [RFC2547-bis] VPN information or the virtual router related VPN information. From the VR and the [RFC2547-bis] point of view there is complete separation from data path and addressing schemes. However the PE's interfaces are shared between both architectures.

A PE implementing only [RFC2547-bis] will not import routes from a BGP UPDATE message containing the VPN-ID extended community. On the other hand, a PE implementing the virtual router architecture will not import routes from a BGP UPDATE message containing the route target extended community attribute.

The granularity at which the information is either [RFC2547-bis] related or VR-related is per BGP UPDATE message. Different SAFI numbers are used to indicate that the message carried in BGP multiprotocol extension attributes is to be handled by the VR or [RFC2547-bis] architectures. SAFI number of 128 is used for [RFC2547-bis] related format. A value of 129 for the SAFI number is for the virtual router (where the NLRI are carrying a labeled prefixes), and a SAFI value of 140 is for non labeled addresses.

7. Use of BGP Capability Advertisement

A BGP speaker that uses VPN information as described in this document with multiprotocol extensions should use the Capability Advertisement procedures to determine whether the speaker could use Multiprotocol Extensions with a particular peer. The Capability Code field is set to 1 (which indicates Multiprotocol Extensions capabilities).

8. Security Considerations

This draft does not introduce any new security considerations to either [VPN-VR] or [RFC2547-bis].

9. References

[BGP-COMM] Ramachandra, Tappan, "BGP Extended Communities Attribute", February 2000, work in progress

[BGP-MP] Bates, Chandra, Katz, and Rekhter, "Multiprotocol Extensions for BGP4", February 1998, RFC 2283

[BGP-MPLS] Rekhter Y, Rosen E., "Carrying Label Information in BGP4", January 2000, work in progress

[MPLS-ARCH] Rosen, Viswanathan, and Callon, "Multiprotocol Label Switching Architecture", August 1999, work in progress

[MPLS-ENCAPS] Rosen, Rekhter, Tappan, Farinacci, Fedorkow, Li, and Conta, "MPLS Label Stack Encoding", October 1999,work in progress

[RFC-1701] Hanks, S., Li, T., Farinacci, D. and P. Traina, "Generic Routing Encapsulation (GRE)", RFC 1701, October 1994.

[RFC-2003] Perkins, C., "IP Encapsulation within IP", RFC 2003, October 1996.

[RFC-2026] Bradner, S., "The Internet Standards Process -- Revision 3", RFC2026, October 1996.

[RFC-2401] Kent S., Atkinson R., "Security Architecture for the Internet Protocol", RFC2401, November 1998.

[RFC-2685] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, March 1997.

[RFC2547-bis] Rosen E., et al, "BGP/MPLS VPNs", work in progress.

[RFC-2685] Fox B., et al, "Virtual Private Networks Identifier", RFC 2685, September 1999.

[VPN-VR] Ould-Brahim H., et al., "Network based IP VPN Architecture using Virtual Routers", work in progress.

[VPN-BGP] Ould-Brahim H., et al., "BGP/VPN: VPN Information Discovery for Network-based VPNs", work in progress.

10. Acknowledgments

to be supplied.

11. Author's Addresses

Hamid Ould-Brahim
Nortel Networks
P O Box 3511 Station C
Ottawa, ON K1Y 4H7, Canada
Email: hbrahim@nortelnetworks.com
Phone: +1 613 765 3418

Bryan Gleeson
Nortel Networks
2305 Mission College Blvd
Santa Clara CA 95054
Phone: +1 (505) 565 2625
Email: bgleeson@shastanets.com

Peter Ashwood-Smith
Nortel Networks
P.O. Box 3511 Station C,
Ottawa, ON K1Y 4H7, Canada
Phone: +1 613 763 4534
Email: petera@nortelnetworks.com

Eric C. Rosen
Cisco Systems, Inc.
250 Apollo drive
Chelmsford, MA, 01824
E-mail: erosen@cisco.com

Yakov Rekhter
Juniper Networks
1194 N. Mathilda Avenue
Sunnyvale, CA 94089
Email: yakov@juniper.net

Luyuan Fang
AT&T

200 Laurel Avenue
Middletown, NJ 07748
Email: Luyuanfang@att.com

Phone: +1 (732) 420 1920

Jeremy De Clercq
Alcatel
Francis Wellesplein 1
B-2018 Antwerpen, Belgium
Phone: +32 3 240 47 52
Email: jeremy.de_clercq@alcatel.be

Full Copyright Statement