Internet Engineering Task Force
INTERNET DRAFT
Expires April 2002

Juha Heinanen
Song Networks
November, 2001

Directory/LDP Based Unidirectional Virtual Circuit VPNs
<draft-heinanen-dirldp-uni-vc-vpns-01.txt>

Status of this Memo

Abstract

This memo describes how provider based unidirectional Virtual Circuit VPNs can be implemented using a directory (such as DNS) and LDP for PE discovery and label distribution.

1. Introduction

This memo describes a simple, easy to configure mechanism for implementing provider based unidirectional Virtual Circuit (VC) VPNs using a directory and LDP [1] for PE discovery and label distribution.

Unidirectional VC VPNs are similar to Frame Relay or ATM VPNs, but are based on MPLS technology. CE devices are assumed to be Layer 3 routers that (in this version of the memo) have full mesh VC connectivity between each other. VCs are organized so that each CE has a single inbound VC for receiving packets from all other CEs and

one outbound VC per each other CE for sending packets to the other CEs. The provider network thus merges the VCs bound for the same CE, which greatly reduces signaling traffic and the amount of label state.

LDP is used to distribute label and Layer 3 protocol address information among both CEs and PEs. That makes configuration of CEs and PEs very simple. The only thing that needs to be configured in a CE is a Layer 3 protocol address for each VPN that the CE participates in. In the PEs the only configuration task is assignment of PE ports to VPNs. If Frame Relay LMI or ATM ILMI would be used on the PE-CE links, there would be no way for the a PE to indicate to a CE which VPN an advertised label belongs to. Also, another protocol, such as InARP, would be needed in the CEs for discovery of Layer 3 protocol addresses.

A directory is used to store the IP addresses of PEs of each VPN. In this memo the directory is DNS as proposed in [2]. Advantages of a DNS/LDP based solution include that it doesn't require BGP implementation or configuration complexity in the PE routers and can be easily deployed also in inter-AS cases where the VPN sites are attached to PEs in more than one AS. The choice of DNS for the directory is justified because it is already in wide use and can be deployed without any new standardization effort.

Similar DNS/LDP based solution can also be applied to provider based Ethernet VPNs as described in [3]. It is also possible to use a DNS/LDP based solution for implementing bidirectional VC VPNs. Details of the latter may be described in a later memo.

2. Addition of Sites

2.1 Configuration Actions

DNS/LDP based Ethernet VPNs are very easy to provision. The following three configuration actions are needed when a new site (CE router) is added to a VPN:

(1) The CE is assigned one or more Layer 3 protocol addresses for the VPN. The VPN is identified in the CE using a VPN ID (Route Distinguisher [4]).

(2) If the PE device (PE for short) to which the CE is connected to does not previously have any sites in this VPN, the IP address of the PE is configured to DNS under domain name

vpn-number.as-number.domain

where "vpn-number" and "as-number" are components of the VPN ID "domain" is the domain of the administrative "owner", (e.g., an ISP) of the VPN.

(3) The port of the PE to which the site is connected to is configured to belong to the VPN. This is done by specifying the domain, type, and VPN ID of the VPN.

This document covers the case where the type of the VPN is "Unidirectional VC". Other possibilities include "Ethernet" and "Bidirectional VC". The former has already been described in [3]. If there is sufficient interest, the latter can be later described in in another memo.

Note that also in the case of a multi-provider VPN, the administrative "owner" of the VPN is the single body that operates the master DNS server for the VPN zone. The "owner" of a VPN MAY choose to make all updates to the zone data of the VPN itself or MAY allow other providers to dynamically update the zone data. In the latter case, the use of secure dynamic updates [5] is recommended.

2.2 Protocol Actions

After the above configuration actions, the following protocol actions take place at the PE of the new site:

(1) The PE sends a Label Request Message to the CE of the new site requesting for a label to be used for sending packets from other sites of the VPN to the new site. The CE responds with a Label Mapping Message that, in addition to the label, contains Layer 3 protocol addresses of the CE.

(2) The PE sends a Label Mapping Message to each of the other CEs connected to it that belong to the same VPN as the new site. The Label Mapping Message advertises a label to be used by another CE when it sends packets to the new site. The Label Mapping Message also contains the Layer 3 protocol addresses of the CE of the new site.

(3) The PE maps the labels that it advertised to the other CEs in steps (2) to the label that it got from the CE of the new site in step (1).

(4) If the new site is the first site of the VPN at the PE, the PE queries DNS for IP addresses of the other (remote) PEs of the VPN and establishes an LDP session with each of the remote PEs unless one already exists.

(5) The PE sends a Label Mapping Message to each of the remote PEs that advertises a label to be used when a remote PE sends packets to the new site. The Label Mapping Message also contains Layer 3 protocol address of the CE of the new site.

(6) The PE maps the labels that it advertised to the other PEs in step (5) to the label that it got from the CE of the new site in step (1).

The following protocol actions take place at a PE when it receives a Label Mapping message from another PE:

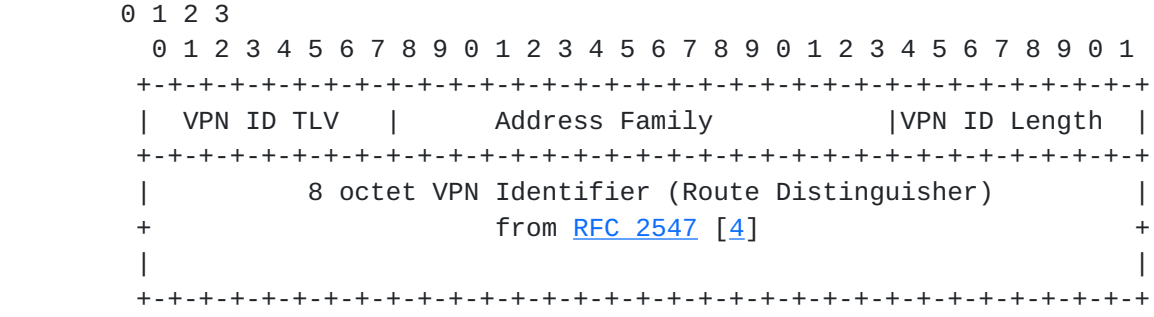
(1) The PE checks from the DNS that the other PE belongs to the VPN of the Label Mapping Message and that it itself has at least one site in that VPN. If not, the PE responds to the Label Mapping Message with a Label Release Message and no other protocol actions take place at the PE.

(2) The PE sends a Label Mapping Message to each CE connected to it that belongs to the VPN of the advertised label. The messages advertise labels to be used by the CEs when they send packets to the site at the other PE. Each Label Mapping Message also contains Layer 3 protocol addresses of the CE of the site at the other PE.

(3) The PE maps the labels that it advertised to its CEs in step (3) to the label that it got from the other PE.

(4) If the label that the PE got from the other PE is the first label from the other PE for this VPN, the PE sends a Label Mapping Message to the other PE for *each* CE that is connected to it for this VPN (unless it has already done so).

The Label Request Messages send from PEs to CEs include the following VPN FEC Element:

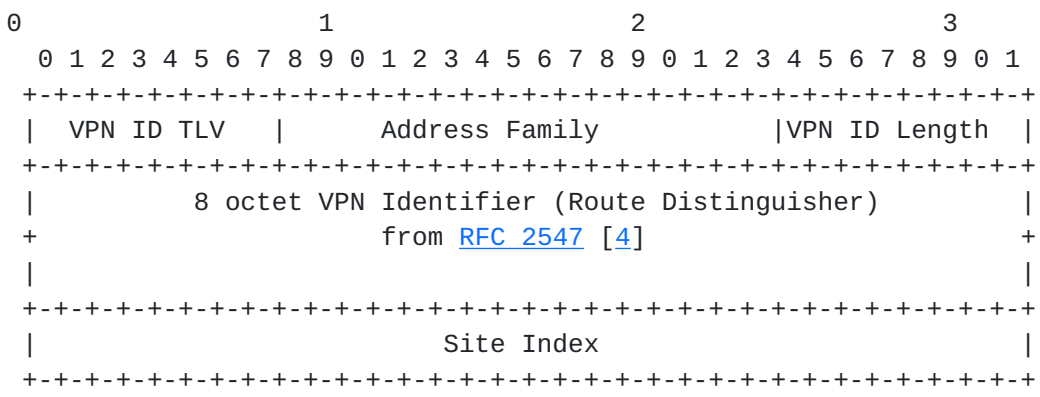


Element type name: VPN
Type: TBD by IANA

Address Family: set to zero
VPN ID Length: 8 octets

The Label Mapping Messages send from CEs to PEs include, in addition to the above VPN FEC Element, a Host Address FEC Element for each Layer 3 protocol address of a CE.

The Label Mapping Messages send between PEs include for the following VPN Site FEC Element:



Element type name: VPN Site
Type: TBD by IANA
Address Family: set to zero
VPN ID Length: 8 octets
Site Index: 4 octets

The Site Index uniquely identifies the site of the VPN at the PE that sends the Label Mapping Message. In addition, the Label Mapping Messages include a Host Address FEC Element for each Layer 3 protocol address of the CE of the site.

3. Addition of Addresses

When new Layer 3 protocol addresses are added to a CE of a site of a VPN, the CE sends to its PE a Label Mapping Message that includes the existing label of the VPN and Host Address FEC Elements containing the new addresses.

When a PE receives from a CE a Label Mapping Message containing new Layer 3 addresses, the PE sends a Label mapping message to each other CE of the VPN at the PE as well as to each other PE of the VPN that includes the existing label of the site of the VPN and Host Address FEC Elements containing the new addresses of the CE.

When a PE receives from another PE a Label Mapping Message containing

new Layer 3 addresses for a CE of an existing VPN site, it sends a Label Mapping message to each CE of the VPN at the PE that includes an existing label of the VPN site and Host Address FEC Elements containing the new addresses of the CE of the site.

4. Removal of Addresses and Sites

When Layer 3 protocol addresses are removed from a CE of a site of a VPN, the CE sends to its PE a Label Withdraw Message that includes the existing label of the VPN and the removed Layer 3 protocol addresses in Host Address FEC Elements.

A whole site can be removed from the VPN either by the CE or by the PE. A CE removes itself from a VPN by sending to its PE a Label Withdraw message that includes the existing label of the VPN and a Wildcard FEC Element. A site is removed by a PE by unconfiguring via a management action the VPN from the port to which the site is connected to. The PE then releases the label it had requested from the removed CE of the VPN by sending the CE a Label Release Message that includes a Wildcard FEC Element.

If the removed site was the last site of the VPN at the PE, the PE is removed from the DNS. This can be done either via a separate network management action or automatically by the PE via a dynamic DNS update.

When a PE receives from a CE a Label Withdraw Message or when a site is removed by the PE, the PE sends a Label Withdraw message to each other CE of the VPN at the PE as well as to each other PE of the VPN. The message includes the existing label of the site of the VPN and either the removed Layer 3 protocol addresses in Host Address FEC Elements or, in case of whole site is to be removed, a Wildcard FEC Element.

When a PE receives from another PE a Label Withdraw Message, it sends a corresponding Label Withdraw Message to each CE of the VPN at the PE. If after receiving a Label Withdraw Message from another PE, there is no remaining need to keep the LDP session up between the PE and the other PE, the PE MAY terminate the LDP session with the other PE.

5. Failure Recovery

If a PE looses its LDP session with another PE having site(s) in a common VPN, the PE releases all labels it had advertised to the other PE for this VPN. The PE then tries to re-establish the LDP session until (a) the session gets established or (b) this PE or the other PE no longer have site(s) in this VPN. Once the LDP session gets established, the PE advertises to the other PE a label for each site of the VPN at this PE as described in section 2.2.

When a PE recovers from a crash, it adds each of the configured VPN site(s) to their respective VPN(s) as described in section 2.2.

6. Exponential Back-off Behavior

If any protocol action does not succeed immediately, the normal behavior is that the PE keeps on trying with exponential back-off until the action succeeds or becomes invalid due to a change in VPN configuration. If the protocol action fails for an implementation specific prolonged period of time, the PE SHOULD notify the VPN operator about the problem via a management action.

7. Data Plane

When a CE needs to send a packet to another CE in the same VPN, it prefixes the packet by a protocol identifier and a label stack entry [6] holding a label that the PE has advertised to it for this VPN.

It is also possible to use some other frame format than the label stack entry, for example, Frame Relay or ATM AAL5, on the CE-PE link. In all cases, the labels used in the frames are those advertised via LDP when a site was added to the VPN. In case of Frame Relay, this means that the advertised label values must fit into the DLCI field of the Frame Relay frame. In case of ATM, the first 4 bits of the advertised label value are used as the VPI value and the remaining 16 bits as the VCI value.

When the PE receives a frame from the CE, it either forwards it directly to another CE at the same PE or uses any available tunnel, such as a HIP, GRE, IPSec, VLAN, or MPLS, to forward the frame to another PE. Before doing so, it replaces the label in the received frame by another label that it had learned from the other CE or another PE for this site. The frame format between PEs is always the label stack entry.

The protocol identifier identifies the protocol of the packet that follows it. What kind of protocol identifier is used, depends on the frame format. The default protocol identifier for label stack entry and Frame Relay is NLPID [7], whereas the default for ATM AAL5 is LLC/SNAP [8].

How a PE decides, which tunneling protocol to use to send labeled packets to another PE, is outside the scope of this memo. Usually the PE would try tunneling protocols in its own preferred order until the tunnel gets established. In most cases the availability of a tunneling protocol can be determined by out-of-band means (e.g., DNS in case of HIP and IPSec, existence of an outer tunnel in case of MPLS, or existence of a shared authentication key in case of GRE).

8. DNS Zone Update Latency

Since the addition of the first site and removal of the last site of a VPN in a particular PE cannot proceed before the change has propagated to all DNS servers serving the zone of the VPN, it is important to try to minimize the latency of VPN zone updates. This can be achieved by turning on DNS NOTIFY [9] in the master server of each VPN zone and by configuring zone refresh times relatively small.

9. DNS Message Size

Correct operation of directory/LDP based VPNs requires that IP addresses of all PE routers of a VPN fit into a single DNS response. In order to be able to support large VPNs with a large number of PEs, the message size requirements of [10] also apply to DNS servers and resolvers used for implementing the mechanism of this memo. Fulfilling those requirements allows provisioning of directory/LDP based VPNs that consist of a few hundred of PEs.

10. Security Considerations

Security of directory/LDP based VPNs depends on security of the directory (DNS), LDP, and the tunneling protocol(s). Security of LDP is covered in the security section of [1]. Also the various tunneling protocol specifications have their own security sections.

Regarding DNS security, the important issues related to this memo are security of zone transfers and integrity and authentication of DNS queries and responses. These two problems are addressed by DNS extensions [11] and [12].

No DNS extensions exist for providing confidentiality for queries or responses. It is thus possible that if a party knows the VPN ID of a VPN and the zone that hosts it, the party can find out the IP addresses of PE routers that connect sites of that domain. Depending on the situation, that may or may not be an acceptable security risk.

In a single-provider VPN, the DNS servers that host the VPN information can be easily fire-walled from all public access. Another way to prevent outside parties from accessing VPN information is to use DNS access lists that VPN zone related queries only from trusted PE routers.

See [2] for additional DNS/VPN related security discussion.

References

[1] Andersson, et al., "LDP Specification". RFC 3036, January 2001.

[2] Luciani et al., "Using DNS for VPN Discovery". draft-luciani-ppvpn-vpn-discovery-00.txt, September 2001.

[3] Heinanen, "Directory/LDP Based Ethernet VPNs". draft-heinanen-dirldp-eth-vpns-01.txt, October 2001.

[4] Rosen and Rekhter, "BGP/MPLS VPNs". RFC 2547, March 1999.

[5] Wellington, "Secure Domain Name System (DNS) Dynamic Update". RFC 3007, November 2000.

[6] Rosen et al., "MPLS Label Stack Encoding". RFC 3032, January 2001.

[7] Brown, "Multiprotocol Interconnect over Frame Relay". RFC 2427, September 1998.

[8] Grossman, "Multiprotocol Encapsulation over ATM Adaptation Layer 5". RFC 2684, September 1999.

[9] Vixie, "A Mechanism for Prompt Notification of Zone Changes (DNS NOTIFY)". RFC 1996, August 1996.

[10] Gudmundsson, "DNSSEC and IPv6 A6 aware server/resolver message size requirements". draft-ietf-dnsext-message-size-04.txt, February 2001.

[11] Vixie, et al., "Secret Key Transaction Authentication for DNS (TSIG)". RFC 2845, May 2000.

[12] Eastlake, "Domain Name System Security Extensions". RFC 2535, March 1999.

Author's Address

Juha Heinanen
Song Networks, Inc.
Hallituskatu 16
33200 Tampere, Finland
Email: jh@song.fi

Full Copyright