Internet Draft Document   Olen Stokes
draft-stokes-ppvpn-vpls-discover-00.txt   Kevin Frick
                                                      Extreme Networks

                                                          Giles Heron
                                                   PacketExchange Ltd.

                                                        Vach Kompella
                                                      TiMetra Networks

Expires December 2002   June 2002

          Discovering Nodes and Services in a VPLS Network

Status of this Memo

    This document is an Internet-Draft and is in full conformance with
    all provisions of Section 10 of RFC2026.

    Internet-Drafts are working documents of the Internet Engineering
    Task Force (IETF), its areas, and its working groups. Note that
    other groups may also distribute working documents as Internet-
    Drafts.

    Internet-Drafts are draft documents valid for a maximum of six
    months and may be updated, replaced, or obsoleted by other
    documents at any time. It is inappropriate to use Internet-Drafts
    as reference material or to cite them other than as "work in
    progress."

    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

    This document describes a methodology for allowing the nodes in a
    Virtual Private LAN Service (VPLS) network as described in [VPLS]
    to discover each other via MPLS signaling. It also defines a
    methodology whereby the individual VPLS services are discovered via
    MPLS signaling. The goal is to allow a VPLS network to be created

using MPLS signaling and a minimal amount of configuration.

Conventions

    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 RFC 2119.

Placement of this Memo in Sub-IP Area

RELATED DOCUMENTS

    http://search.ietf.org/internet-drafts/draft-lasserre-vkompella-
    ppvpn-vpls-01.txt

    http://search.ietf.org/internet-drafts/draft-martini-l2circuit-
    trans-mpls-08.txt

    http://search.ietf.org/internet-drafts/draft-martini-ethernet-
    encap-mpls-00.txt

    http://www.ietf.org/rfc/rfc3036.txt

    http://www.ietf.org/rfc/rfc3209.txt

WHERE DOES THIS FIT IN THE PICTURE OF THE SUB-IP WORK

    PPVPN

WHY IS IT TARGETTED AT THIS WG

    The charter of the PPVPN WG includes L2 VPN services and this draft
    specifies a method for establishing Ethernet VPLS services over
    MPLS.

JUSTIFICATION

   There is no Internet document that provides a method for
   establishing a VPLS network using signaling and a minimal amount of
   configuration through the use of MPLS protocols.

Stokes, et al.   [Page 2]

Internet Draft draft-stokes-ppvpn-vpls-discover-00.txt   June 2002

                         Table of Contents

1 Overview

   Service providers are being requested to provide L2 VPN services
   for their customers. Virtual Private LAN Services (VPLS) networks
   as described in [VPLS] are being deployed to provide these
   services. [VPLS] describes the operation of a VPLS network but
   does not specify how the network is configured.

   A provider's network is normally organized with groups of one or
   more edge nodes connecting to a group of core nodes. This
   corresponds to the spoke and core nodes of the "HVPLS" model
   described in [VPLS]. This document describes a methodology for

allowing the core nodes in a VPLS network to discover each other
via MPLS signaling. It also describes how core nodes can discover
the spoke nodes attaching to it via MPLS signaling.

Once the underlying VPLS network is established as described above,
customer sites for individual VPLS services are added to the nodes.
This document also defines a methodology whereby the individual
VPLS services are discovered by the core nodes via MPLS signaling.

1.1 Terminology

The phrase "VPLS network" is used when referring collectively to a
service provider's core and spoke nodes and the targeted LDP
sessions between them.

The phrase "VPLS service" is used when referring to a single VPLS
service emulating a LAN segment.

The phrase "VPLS core node" is used when referring to a core node
in an HVPLS service, or a node in a VPLS service with no HVPLS
"spoke" nodes. Note that all VPLS core nodes for a given VPLS are
meshed.

The phrase "/32 LSP" is used when referring to a LDP-signaled LSP
whose FEC specifies a 32-bit IP address prefix [MPLS-LDP].

2 Node discovery signaling

Each spoke and core node in a VPLS network uses targeted LDP
sessions as defined in [MARTINI-SIG] and [VPLS]. If the underlying
tunnel LSPs are to created via downstream unsolicited LDP [MPLS-
LDP], then each node is configured to advertise a label mapping for
a /32 LSP to its IP address. Normally, this is the only LSP that
is advertised by the node.

These label advertisements are propagated throughout the VPLS
network until all nodes have received one or more label mappings
for a /32 LSP to all other nodes. These label mappings can be used
to signal whether the LSP's egress node is a core node in the VPLS
network.

Internet Draft draft-stokes-ppvpn-vpls-discover-00.txt   June 2002

To signal this Service Capabilities, a new optional TLV is proposed

for use with a LDP Label Mapping message. In fact this TLV may be
used for other services with similar requirements for node
discovery. The new Service Capabilities TLV is shown below.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|U|F|      Type (0x0105)         |             Length           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Service Type          |        Service Instance      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Optional Service Parameters (variable length)         |
~                                                               ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

U bit
     Unknown bit. This bit MUST be set to 1. If the Node Type
format is not understood, the remaining Label Mapping should be
processed, and the Service Capabilities TLV MUST be ignored.

F bit
     Forward bit. This bit MUST be set to 1. Since the network
topology is unknown and there may be non-VPLS nodes receiving the
label message, the Service Capabilities TLV MUST be forwarded.

Type
     Type field. This field MUST be set to 0x0105 (subject to
IANA approval). This will identify the TLV type as a Service
Capabilities TLV.

Length
     Length field. This field specifies the total length of the
Value fields in octets.

Service Type
     Service Type field. This has the following values:
       Value   Node Type
       ----
        0   Undefined
        1   VPLS core node
Note that a node may support multiple service types, in which case
it will advertise multiple Service Capabilities TLVs.

 Service Instance
     Service Instance field. This field contains a 16-bit index
used as an instance identifier for the service. This enables, for
example, one Service Provider to create multiple sets of VPLS nodes
providing different grades of service. Note that a node may have
multiple service instances, in which case it will advertise

multiple Service Capabilities TLVs. Note also that in the case of
VPLS, one service instance may contain multiple VPLS services.

    Service Parameters

        Optional Service Parameters. Some service types may have
    additional parameters that must be discovered for correct operation
    of the service. This field is not required for VPLS. Note that
    for a given service type and instance a node will only advertise
    one Service Capabilities TLV, with one set of Service Parameters.

    If RSVP-TE [MPLS-RSVP] is being used to create the underlying
    tunnel LSPs, then LDP MUST be enabled to advertise a /32 LSP for
    each node. Note also that these additional LSPs do not need to be
    used as part of the underlying tunnel LSPs used to carry the
    customer traffic.

2.1 Service Capabilities TLV procedures

    Each time that a label mapping for a /32 LSP is received from a
    valid next hop, a check is made for the presence of a Service
    Capabilities TLV.

    If no Service Capabilities TLV is present and there were Service
    Capabilities TLVs received previously for the IP address specified
    in the Label Mapping's FEC then any local configuration relating to
    these TLVs MUST be removed. The label mapping MUST be propagated
    upstream.

    If one or more Service Capabilities TLVs are present, but there
    were Service Capabilities TLVs received previously for the IP
    address specified in the Label Mapping's FEC with a (Service Type,
    Service Instance) which is not seen in this message, then any
    configuration relating to these TLVs MUST be removed. The label
    mapping MUST be propagated upstream.

    If one or more Service Capabilities TLVs are present with a
    (Service Type, Service Instance) which has not been received
    previously for the IP address specified in the Label Mapping's FEC,
    then these TLVs will be used locally. The label mapping MUST be
    propagated upstream.

    If one or more Service Capabilities TLVs are present with a
    (Service Type, Service Instance) which has been received previously

for the IP address specified in the Label Mapping's FEC, but with
different Service Parameters to those in the received TLVs, then
these new parameters will be used locally. The label mapping MUST
be propagated upstream.

If one or more Service Capabilities TLVs are present, there are
previously received Service Capabilities TLVs from this same
downstream peer, and any of those match exactly the content of a
Service Capabilities TLV received on the new label mapping, then no
further action is required on behalf of those newly received
Service Capabilities TLVs. Note that there may be other actions
required due to other TLVs in the Label Mapping message.

3 VPLS Core node procedures

Each time a VPLS core node receives information about the presence
of another core node with a shared service instance to one
configured locally, it adds the IP address from the /32 LSP's FEC
field to its list of core nodes. It then establishes a targeted
LDP session to this new core for use in the VPLS network - if one
is not already present. When the node propagates the label mapping
upstream, it leaves the Service Capabilities TLV unchanged. If
there was previously a targeted LDP session to this node (as a
spoke) for this VPLS instance then any VC FECs for this VPLS
instance are first withdrawn on that session to ensure that no
forwarding loops are created at the VPLS level.

If the decision process in Section 2.1 indicates that a remote node
is no longer a core node, then the receiving core node removes the
IP address from the /32 LSP's FEC field from its list of core
nodes. Any VC FECs for this VPLS instance will be withdrawn, and
the targeted LDP session to this node will be dropped if there are
no other VPLS instances or other entities using that session.

4 HVPLS spoke node procedures

If the decision process in Section 2.1 indicates the presence of a
new core node in the network, then a receiving spoke node adds the
IP address from the /32 LSP's FEC field to its list of core nodes.
If the new core node is topologically closer than any other core

node then the spoke node establishes a targeted LDP session to the
new core node (if one is not already present). It also withdraws
any VC FECs from any targeted LDP session to another core node and
drops that session if there are no other VPLS instances or other
entities using it.

If the decision process in Section 2.1 indicates that a node is no
longer a core node then the spoke node removes the IP address from
the /32 LSP's FEC field from its list of core node. Any VC FECs
advertised to that node are withdrawn, and the targeted LDP session
is dropped if there are no more FECs advertised over it. A new
session established to the topologically closest core node.

5 Service discovery signaling

The procedures in Section 2 create a VPLS network with targeted LDP
sessions as required between nodes. Once this is accomplished,
VPLS services can be signaled. When a provider adds a site to a
VPLS service, it requires configuring the node to which the
customer site is added. Using the procedures defined below, no
additional provider configuration is necessary to establish the
VPLS service.

When a customer site is added to a spoke node, it sends a VC FEC
label to its core node as defined in [MARTINI-SIG] and [VPLS]. The
core node stores the VC FEC label mapping in its liberally retained
label database and associates the mapping with the spoke from which
it was received. Note that the VC ID field in the VC FEC
identifies the VPLS service.

Stokes, et al.   [Page 7]

The core node propagates the VC FEC label mapping to all of its
core peer nodes. These core nodes add the VC FEC label mapping in
their liberally retained label database and associate the mapping
with the core node from which it was received.

At this point, if this was the first customer site in the VPLS, the
original spoke node has not received a VC FEC label mapping from
the core node and therefore considers the VPLS service to be non-
operational.

When a customer site is added to a core node, it sends a VC FEC
label to all of its peer core nodes as described above. However as

above this node may not have received a FC FEC label mapping from
any other core node, in which case it will consider the VPLS
service to be non-operational.

When a second customer site is added at a spoke, this node
advertises a VC FEC label to its upstream core node as above.

This core node (whether the second site is attached locally or via
a spoke) recognizes that it has already received a VC FEC label
mapping for this VPLS service by detecting a liberally retained
mapping for this same VC ID. The core node moves both label
mappings to its active label database and, if the second site is
attached at a spoke node, sends a corresponding VC FEC label
mapping to the spoke from which it received the VC FEC label
mapping.

This core node also propagates a VC FEC label mapping to all of its
core peers as normal. The original core node (if the first site
was attached to a different core node) also recognizes that is has
already received a VC FEC label mapping for this VPLS service by
detecting a liberally retained mapping for this same VC ID. The
original core node moves both label mappings to its active label
database and, if the first customer site was attached at a spoke
node, sends a corresponding VC FEC label mapping to its spoke from
which it received the VC FEC label mapping.

Following this, both edge nodes have received VC FEC label mappings
for the VPLS service and consider the VPLS service to be
operational. If either of these nodes are spokes, their upstream
core nodes also consider the service to be operational and have
added their peer core node and their spoke node to the flood list
for this VPLS service.

The process above can be generalized for adding the nth customer
site to the network.

Note that the core nodes that are not participating in this VPLS
service have liberally retained labels but have not dedicated any
operational resources to this VPLS service.

When a customer site attached at a spoke node is removed from an
existing VPLS service, the spoke node sends a label withdrawal to
its core node for the VC FEC. The core node responds with a label
release and removes the label from its active label database. It

also sends a label withdrawal for the VC FEC label mapping that it had sent to the spoke when the VPLS service became operational.

When the last customer site in a VPLS is removed from a core node (either directly or by withdrawal of a label from a spoke node) it sends a label withdrawal for this VC FEC label to all of its peer core nodes. If this is the only core node from which these remote core nodes have received a label mapping for this VPLS service, then they in turn must send a label withdrawal to any spoke nodes participating in the VPLS. They also move the label mapping to the liberally retained database.

Note that when a core node has two or more VC FEC label mappings for the same VPLS service from multiple spoke nodes, or has a locally attached site in the VPLS plus one VC FEC label mapping from a spoke node, the VPLS service remains operational even if no other core nodes have advertised label mappings for this VPLS service.

6 Security Considerations

Security issues resulting from this draft will be discussed in greater depth at a later point. It is recommended in [MPLS-LDP] that LDP security (authentication) methods be applied. This would prevent unauthorized participation by a spoke or core node in a VPLS.

7 Scalability Issues

In [VPLS], targeted LDP sessions are used to distribute VPLS labels. Between core nodes, a complete mesh of targeted LDP sessions is required.

The HVPLS model, where spoke nodes attach to core nodes, and core nodes are meshed, enables the service to scale to a greater number of PE devices.

8 Intellectual Property Considerations

Extreme Networks, Inc. is seeking patent protection on technology described in this Internet-Draft. If technology in this Internet-Draft is adopted as a standard, Extreme Networks agrees to license, on reasonable and non-discriminatory terms, any patent rights it obtains covering such technology to the extent necessary to comply with the standard. This document is being submitted for use in IETF standards discussions.

9 References

[VPLS] "Virtual Private LAN services over MPLS", draft-lasserre-vkompella-ppvpn-vpls-01.txt (Work In Progress)

[MARTINI-SIG] "Transport of Layer 2 Frames Over MPLS", draft-martini-l2circuit-trans-mpls-09.txt, April 2002 (Work In Progress)

[MARTINI-ENC] "Encapsulation Methods for Transport of Ethernet Frames Over IP and MPLS Networks", draft-martini-ethernet-encap-mpls-00.txt, April 2002 (Work In Progress)

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

[MPLS-RSVP] Awduche, et al., "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001

10 Authors' Addresses

Olen Stokes
Extreme Networks
630 Davis Drive, Suite 250
Morrisville, NC 27560
Email: ostokes@extremenetworks.com

Kevin Frick
Extreme Networks
630 Davis Drive, Suite 250
Morrisville, NC 27560
Email: kfrick@extremenetworks.com

Giles Heron
PacketExchange Ltd.
The Truman Brewery
91 Brick Lane
LONDON E1 6QL
United Kingdom
Email: giles@packetexchange.net

Vach Kompella

TiMetra Networks
274 Ferguson Dr.
Mountain View, CA 94043
Email: vkompella@timetra.com