BGP/GMPLS Optical VPNs

draft-ouldbrahim-bgpgmpls-ovpn-01.txt

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026 [RFC-2026], except that the right to produce derivative works is not granted.

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.

Ould-Brahim, et. al

draft-ouldbrahim-bgpgmpls-ovpn-01.txt  July 2001

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-
Abstract

Consider a service provider network that offers Optical Virtual Private Network (OVPN) service. An important goal in the OVPN service is the ability to support what is known as "single end provisioning", where addition of a new port to a given OVPN would involve configuration/provisioning changes only on the devices connected to that port. Another important goal in the OVPN service is the ability to establish/terminate an optical connection between a pair of (existing) ports within an OVPN without involving configuration/provisioning changes in any of the provider devices.

In this document we describe a set of mechanisms that accomplishes these goals.

Obsoletes draft-fedyk-bgpvpon-auto-00.txt

1. Sub-IP Summary ID

This ID targets the PPVPN working group as it deals with a VPN solution similar to port-based VPNs. It describes an approach to allow service providers to offer optical VPN service. A pair of client devices (a router, a SONET/SDH cross-connect, or an Ethernet switch) could be connected through the service provider network via an optical connection. It is this optical connection that forms the basic unit of service that the service provider network offers.

RELATED DOCUMENTS

draft-ouldbrahim-ovpn-requirements-00.txt. Others can be found in the "references" section.

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

Fits the PPVPN box.

WHY IS IT TARGETED AT THIS WG
This WG is looking at port based VPN over an IP/MPLS infrastructure. This work is exactly a port-based optical VPN using IP related building blocks.

JUSTIFICATION

The current PPVPN chairs have already discussed this work and are considering expanding the PPVPN charter to include OVPN as part of PPVPN mandate.

2. Optical VPN Reference Model

Consider a service provider network that consists of devices such as Optical Network Element (ONE) which may be Optical Cross Connects (OXC's). We partition these devices into P (provider) ONEs and PE (provider edge) ONEs. The P ONEs are connected only to the ONEs within the provider’s network. The PE ONEs are connected to the ONEs within the provider network, as well as to the devices outside of the provider network. We’ll refer to such other devices as Client Edge Devices (CEs). An example of a CE would be a router, or a SONET/SDH cross-connect, or an Ethernet switch.

```
+---+    +---+
| P |....| P |
+---+    +---+

PE /              
+-----+               +-----+    +--+
|     |               |     |----|  |
+--+    |     |               |     |    |CE|
|CE|----+-----+               |     |----|  |
+--+
```

Figure 1 Optical VPN Reference Model

A CE is connected to a PE ONE via one or more links, where each link may consists of one or more channels or sub-channels (e.g., wavelength or wavelength and timeslot respectively). For purpose of this discussion we assume that all the channels
within a given link have shared similar characteristics (e.g., bandwidth, encoding, etc.), and can be interchanged from the CEs point of view. Channels on different links of a CE need not have the same characteristics.

There may be more than one link between a given CE PE ONE pair. A CE may be connected to more than one PE ONE (with at least one port per each PE ONE). And, of course, a PE ONE may have more than one CE connected to it.

If a CE is connected to a PE ONE via multiple links and all these links belong to the same VPN, then for the purpose of OVPN these links could be treated as a single link using the link bundling constructs [LINK-BUNDLING].

In general a link may have only data bearing channels, or only control bearing channels, or both. For the purpose of this discussion we assume that for a given CE - PE ONE pair at least one of the links between them has at least one control bearing channel and at least one data bearing channel.

A link has two end-points - one on CE and one on PE ONE. In the context of this document we’ll refer to the former as "CE port", and to the latter as "PE ONE port". From the above it follows that a CE is connected to a PE ONE via one or more ports, where each port may consists of one or more channels or sub-channels (e.g., wavelength or wavelength and timeslot respectively), and all the channels within a given port have shared similar characteristics (e.g., bandwidth, encoding, etc.), and can be interchanged from the CEs point of view. Channels on different ports of a CE need not have the same characteristics.

Note that in the context of this document both ports and links are logical constructs, and are used to represent grouping of physical resources per OVPN basis that are used to connect a CE to a PE ONE. At any given point in time, a given port on a PE ONE is associated with at most one OVPN. This association is established and maintained by the service provider provisioning system.

A pair of CEs could be connected through the service provider network via an optical connection. It is precisely this optical connection that forms the basic unit of service that the service provider network offers. If a port by which a CE is connected to a PE ONE consists of multiple channels (e.g., multiple wavelengths), the CE could establish optical connection to multiple other CEs over this single port.

An important goal in the OVPN service is the ability to support what is known as "single end provisioning", where addition of a
new port to a given OVPN would involve configuration/provisioning changes only on the PE ONE that has this port and on the CE that is connected to the PE ONE via this port. Another important goal in the OVPN service is the ability to establish/terminate an optical connection between a pair of (existing) ports within an OVPN without involving configuration/provisioning changes in any of the provider’s ONEs. The mechanisms outlined in this document aim at achieving these goals. Specifically, as part of the Optical VPN service offering, these mechanisms (1) enable the service provider to restrict the set of ports that a given port could be connected to, (2) enable the service provider to provide a CE with the information about the ports that the CE could be connected, (3) enable a CE to establish the actual connections to a subset of ports provided by (2). Finally, the mechanisms allow different OVPN topologies to be supported ranging from hub-and-spoke to complete mesh.

The service provider does not initiate the creation of an optical circuit between a pair of PE ONE ports. This is done rather by the CEs, which attach to the ports. However, the SP, by using the mechanisms outlined in this document, restricts the set of other PE ONE ports which may be the remote endpoints of optical circuits that have the given port as the local endpoint. Subject to these restrictions, the CE-to-CE connectivity is under the control of the CEs themselves. In other words, SP allows an OVPN to have a certain set of topologies, and CE-initiated signaling is used to choose a particular topology from that set.

Since this model involves minimal provisioning changes when changing the connectivity among the ports within a OVPN on the providers network and the OVPNs themselves are controlled by the CEs, the tariff structure may be on a port basis or alternatively tariffs could be triggered on the basis of signaling mechanisms.

Finally, it is assumed that CE-to-CE optical connectivity is based on GMPLS [GMPLS].

3. Overview of operations

This document assumes that within a given OVPN each port on a CE that connects the CE to a PE ONE has an identifier that is unique within that OVPN (but need not be unique across several OVPNs). One way to accomplish this is to assign each port an IP address that is unique within a given OVPN, and use this address as a port identifier. Another way to accomplish this is to assign each port an interface index that is unique within
a given CE, assign each CE an IP address that is unique within a given OVPN, and then use a tuple <interface index, CE IP address> as a port identifier.

This document assumes that within a service provider network, each port on a PE ONE has an identifier that is unique within that network. One way to accomplish this would be to assign each port on a PE ONE an interface index, assign each PE ONE an IP address that is unique within the service provider network (in the case of multi-provider operations, the address has to be unique across all the providers involved), and then use a tuple <interface index, PE ONE IP address> as a port identifier within the provider network.

As a result each link connecting the CE to the PE ONE is associated with a CE port that has a unique identifier within a given OVPN, and with a PE port that has a unique identifier within the service provider network. We’ll refer to the former as the customer port identifier (CPI), and to the latter as the provider port identifier (PPI).

This document assumes that in addition to PPI, each port on PE ONE has also an identifier that is unique within the OVPN of the CE connected to that port. One way to accomplish this is to assign each port an IP address that is unique within a given
OVPN, and use this address as a port identifier. Another way to accomplish this is to assign each port an interface index that is unique within a given PE ONE, assign each PE ONE an IP address that is unique within a given OVPN (but need not be unique within the service provider network), and then use a tuple <interface index, PE ONE IP address> acts as a port identifier. We’ll refer to such port identifier as VPN-PPI. Note that PE ONE IP address used for VPN-PPI need not be the same as PE ONE IP address used for PPI, and moreover, a given PE ONE may have multiple PE ONE addresses used for VPN-PPI, one per OVPN. Subject to the constraints outlined in the next paragraph, PPI could be used as VPN-PPI.

For a given link connecting a CE to a PE ONE, if CPI is an IP address, then VPN-PPI has to be an IP address as well. And if CPI is an <interface index, CPI IP address>, then VPN-PPI has to be an <interface index, PE ONE IP address>. However, for a given port on PE ONE, whether VPN-PPI of that port is an IP address or an <interface index, PE ONE IP address> is independent of whether PPI of that port is an IP address or an <interface index, PE ONE IP address>.

Each PE ONE maintains a Port Information Table (PIT) for each OVPN that has at least one port on that PE ONE. A PIT contains a list of <CPI, PPI> tuples for all the ports within its OVPN.

A PIT on a given PE ONE is populated from two sources: the information related to the CEs (optionally received from the CEs) attached to the ports on that PE ONEs, and the information received from other PE ONEs. We’ll refer to the former as the "local" information, and to the latter as the "remote" information.

The local information is propagated to other PE ONEs by using BGP with multi-protocol extensions. To restrict the flow of this information to only the PITs within a given OVPN, we use BGP route filtering based on the Route Target Extended Community [BGP-COMM], as follows.

Each PIT on a PE ONE is configured with one or more Route Target Communities, called "export Route Targets", that are used for tagging the local information when it is exported into provider’s BGP. The granularity of such tagging could be as fine as a single <PPI, CPI> pair. In addition, each PIT on a PE ONE is configured with one or more Route Target Communities, called "import Route Targets", that restrict the set of routes that could be imported from provider’s BGP into the PIT to only the routes that have at least of these Communities.

When a service provider adds a new OVPN port to a particular PE ONE, this port is associated at provisioning time with a PIT on
that PE ONE, and this PIT is associated (again at provisioning time) with that OVPN.

Once a port is configured on the PE ONE, the CE that is attached via this port to the PE ONE MAY pass to the PE ONE the CPI information of that port. This document assumes that this is accomplished by using BGP (however, the document doesn’t preclude the use of other mechanisms).

This information, combined with the PPI information available to the PE ONE, enables the PE ONE to create a tuple \(<\text{CPI}, \text{PPI}>\) for such port, and then use this tuple to populate the PIT of the OVPN associated with that port.

In order to establish an optical connection, a CE needs to identify all other CEs in the CE’s OVPN it wants to connect to. A CE may already have obtained the CE list through configuration or through some other schemes (such schemes are outside the scope of this draft).

It is also desirable, that the service provider, as a value added service, may provide a CE with a list of all other CEs in the CE’s OVPN. This is accomplished by passing the information stored in the PE ONE PITs to the attached CE. This document assumes that this is accomplished by using BGP Multi-protocol extensions (however this draft doesn’t preclude other mechanisms to be used). Although optional, this draft recommends the PE to signal to the attached CEs the remote CIPIs it learnt from the remote CEs part of the same OVPN. A CE may decide to initiate an optical connection request to a remote CE only when it learn the CPI of the remote CE from the PE. This has the benefit to avoid rejecting connection request while the PE is populating the PITs.

Once a CE obtains the information about the CIPIs of other ports within the same OVPN, which we’ll refer to as "target ports", the CE uses a (subset of) GMPLS signaling to request the provider network to establish an optical connection to a target port. Note that this draft assumes that GMPLS is only used to establish optical connections between client devices.

The request originated by the CE contains the CPI of the port on the CE that CE wants to use for the optical connection, and the CPI of the target port. When the PE ONE attached to the CE that originated the request receives the request, the PE ONE identifies the appropriate PIT, and then uses the information in that PIT to find out the PPI associated with the CPI of the target port carried in the request. The PPI should be sufficient for the PE ONE to establish an optical connection. Ultimately the request reaches the CE associated with the target CPI (note that the request still carries the CPI of the CE that originated the request). If the CE associated with the
target CPI accepts the request, the optical connection is established.

Note that a CE need not establish an optical connection to every target port that CE knows about - it is a local to the CE matter to select a subset of target ports to which the CE will try to establish optical connections.

A port, in addition to its CPI and PPI may also have other information associated with it that describes characteristics of the channels within that port, such as encoding supported by the channels, bandwidth of a channel, total unreserved bandwidth within the port, etc. This information could be further augmented with the information about certain capabilities of the Service Provider network (e.g., support RSOH DCC transparency, arbitrary concatenation, etc.). This information is used to ensure that ports at each end of an optical connection have compatible characteristics, and that there are sufficient unallocated resources to establish an optical connection. Distribution of this information (including the mechanisms for distributing this information) is identical to the distribution of the <CPI, PPI> information. Distributing changes to this information due to establishing/terminating of optical connections is identical to the distribution of the <CPI, PPI> information, except that thresholds should be used to contain the volume of control traffic caused by such distribution.

It may happen that for a given pair of ports within an OVPN, each of the CEs connected to these ports would concurrently try to establish an optical connection to the other CE. If having a pair of optical connections between a pair of ports is viewed as undesirable, the a way to resolve this is have CE with the lower value of CPI is required to terminate the optical connection originated by the CE. This option could be controlled by configuration on the CE devices.

4. Encoding

This section specifies encoding of various information defined in this document.

4.1 Encoding of CPI and channel characteristics in GMPLS Signaling

[TBD]

4.2 Encoding of CPI, PPI, and channel characteristics in BGP

4.2.1 Encoding of CPI and PPI information in BGP
The <CPI, PPI> mapping is carried using the Multiprotocol Extensions BGP [RFC2858]. [RFC2858] defines the format of two BGP attributes, MP_REACH_NLRI and MP_UNREACH_NLRI that can be used to announce and withdraw the announcement of reachability information. We introduce a new address family identifier (AFI) for OVPN (to be assigned by the IANA), a new subsequent address family identifier (to be assigned by the IANA), and also a new NLRI format for carrying the CPI and PPI information.

One or more <PPI, CPI> tuples could be carried in the above mentioned BGP attributes.

The format of encoding a single <PPI, CPI> tuple is shown in Figure 3 below:

```
+---------------------------------------+
|    Length (1 octet)                  |
+---------------------------------------+
|    PPI AFI (2 octets)                |
+---------------------------------------+
|    PPI Length (1 octet)              |
+---------------------------------------+
|    PPI (variable)                    |
+---------------------------------------+
|    CPI AFI (2 octets)                |
+---------------------------------------+
|    CPI (length)                      |
+---------------------------------------+
|    CPI (variable)                    |
+---------------------------------------+
```

Figure 3: NLRI BGP encoding

The use and meaning of these fields are as follows:

Length:

A one octet field whose value indicates the length of the <PPI, CPI> Information tuple in octets.

PPI AFI:

A two octets field whose value indicates address family identifier of PPI

PPI Length:

A one octet field whose value indicates the length of the PPI field
PPI field:

A variable length field that contains the value of the PPI (either an address or <interface index, address> tuple.

CPI AFI field:

A two octets field whose value indicates address family of the CPI.

CPI Length:

A once octet field whose value indicates the length of the CPI field.

CPI (variable):

A variable length field that contains the CPI value (either an address or <interface index, address> tuple.

4.2.2 Encoding channel characteristics in BGP

[TBD]

5. Other issues

While the above text assumes that the service provider network consists of ONEs and ports are connected via optical connections, the mechanisms described in this document could be applied in an environment, where the service provider network consists of SONET/SDH cross connects and CE ports being either SONET/SDH or Ethernet.

Since the protocol used to populate a PIT with remote information is BGP, since BGP works across multiple routing domains, and since GMPLS signaling isn’t restricted to a single routing domain, it follows that the mechanisms described in this document could support an environment that consists of multiple routing domains.

The mechanisms described in this document allow for a wide range of choices with respect to addresses used for CPI, PPI, and VPN-PPI. For example, one could use either IPv4 addresses, or IPv6 addresses, or NSAPs. Different OVPN customers of a given service provider may use different types of addresses. Moreover, different OVPNs attaching to the same PE ONE may use different addressing schemes. The types of addresses used for PPIs within a given service provider network are independent from the type of addresses used for CPI and VPN-PPI by the OVPN customers of that provider.
6. Security Considerations

[TBD]

7. References


8. Acknowledgments.

The authors would like to thank Osama Aboul-Magd, Dimitri Papadimitriou, Penno Reinaldo, Erning Ye, and Bryan Gleeson for reviewing the draft and providing comments.
9. Author’s Addresses

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

Don Fedyk
Nortel Networks
600 Technology Park
Billerica, Massachusetts
01821 U.S.A.
Phone: +1 (978) 288 3041
Email: dwfedyk@nortelnetworks.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

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

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

Eric Mannie
Ebone (GTS)
Terhulpsesteenweg 6A
1560 Hoeilaart
Belgium
Phone: +32 2 658 56 52
Email: eric.mannie@gts.com

Luyuan Fang
AT&T
200 Laurel Avenue
Middletown, NJ 07748
Email: Luyuanfang@att.com
Phone: +1 (732) 420 1920

John Drake
Calient Networks
Yong Xue
UNET/WorldCom
Ashburn, Virginia
(703)-886-5358
yxue@uu.net

Riad Hartani
Caspian Networks
170 Baytech Drive
San Jose, CA 95143
Phone: 408 382 5216
Email: riad@caspiannetworks.com