Architecture of an IP/MPLS Network with Hardened Pipes

Abstract

   This document describes an IP/MPLS network that has an infrastructure
   that can be separated into two or more strata. For the
   implementation described in this document, the infrastructure has
   been separated into two strata: one for the "Hard Pipes", called the
   "Hard Pipe Stratum", and one for the normal IP/MPLS traffic, called
   the "Normal IP/MPLS Stratum".

   This document introduces the concept of a Hard Pipe -- an MPLS Label
   Switched Path (LSP) or a pseudowire (PW) with a bandwidth that is
   guaranteed and can neither be exceeded nor infringed upon.

   The Hard Pipe stratum does not use statistical multiplexing; for the
   LSPs and PWs set up within this stratum, the bandwidth is guaranteed
   end to end.

   The document does not specify any new protocol or procedures. It
   does explain how the MPLS standards implementation has been deployed
   and operated to meet the requirements from operators that offer
   traditional Virtual Leased Line (VLL) services.

Status of This Memo

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

   This is a contribution to the RFC Series, independently of any other
   RFC stream. The RFC Editor has chosen to publish this document at
   its discretion and makes no statement about its value for
   implementation or deployment. Documents approved for publication by
   the RFC Editor are not 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/rfc7625.

Table of Contents

1. Introduction

   IP leased line services, Ethernet Private Line (EPL), and Time-
   Division Multiplexed (TDM) leased line services are commonly offered
   by operators worldwide.

   There are customers, e.g., many enterprises, that insist on TDM
   leased line services. They do so regardless of the fact that the
   same operators often offer IP leased line services and EPL services
   at a lower price and with a guaranteed bandwidth.

   Today we see a trend that TDM (in particular, Synchronous Digital
   Hierarchy / Synchronous Optical Network (SDH/SONET)) networks are
   gradually carrying less and less traffic, and many operators want to
   shut their TDM networks down to reduce costs.

   In light of these trends, vendors and operators have built and
   deployed the Hard Pipe service described in this document. It is a
   way to introduce leased line service with the same characteristics as
   TDM leased line services in IP/MPLS networks.

   Even if leased line has been the initial motivation to define the
   Hard Pipe technology, the Hard Pipe is by no means limited to support
   leased line services. When guaranteed bandwidth is the priority,

Virtual Private Wire Services (VPWS), Virtual Private LAN Services
(VPLS), L3 Virtual Private Networks (L3VPN), and IP-only Private LAN
Services can be mapped to a tunnel in the Hard Pipe stratum.

EPL and Ethernet Private LAN (EPLAN) are out of scope for this
document.

Virtual Leased Line service is used in examples throughout this
document.

The solution soon to be deployed has an Ethernet infrastructure that
has been split into two parallel logical networks -- two parallel
strata. The first stratum -- the Hard Pipe Stratum -- does not use
statistical multiplexing, and bandwidth is guaranteed end to end.
The second stratum -- the Normal IP/MPLS Stratum -- works as a normal
IP/MPLS network. The two strata share the same physical network,
i.e., routers and links, but the resource reserved for the Hard Pipe
stratum will never be preempted by the Normal IP/MPLS stratum.

The routers will handle the traffic belonging to one stratum
differently from how traffic from the other stratum is handled. This
separation in traffic handling is based on support in hardware.

The reader of this document is assumed to be familiar with RFC 3031
[RFC3031] and RFC 5921 [RFC5921].

1.1. Scope

This document has the following purposes:

o to introduce a two strata IP/MPLS network: the purpose of one of
   the strata is to provide capabilities for services that are, from
   a customer's point of view, functionally identical to TDM-like
   leased lines; and

o to indicate how a router differentiates the traffic of the two
   strata.

1.2. Abbreviations

CC: Continuity Check

CV: Connection Verification

L-label: Leased Line label

LSP: Label Switched Path

   LSR: Label Switching Router

   MPLS-TP: MPLS Transport Profile

   NMS: Network Management System

   OAM: Operations, Administration, and Maintenance

   P: Provider Router

   PE: Provider Edge Router

   PW: Pseudowire

   T-label: Tunnel label

   TDM: Time-Division Multiplexing

   tLDP: Targeted LDP

   VLL: Virtual Leased Line

   VPLS: Virtual Private LAN Service

   VPWS: Virtual Private Wire Service

2. The Stratified Network

   The concept of stratified or strata networks has been around for some
   time. It appears to have different meaning in different contexts.
   The way we use the concept is that we logically assign certain
   characteristics to part of the network. The part of the network that
   has the special characteristics form one stratum, and the "remainder"
   forms a second stratum. The network described in this document uses
   a single link-layer technology, Ethernet.

   In many cases, a whole physical interface is assigned to a single
   hard stratum, especially in the scenario where there are many
   physical links between two nodes.

   This document does not address the network configuration
   possibilities for Hard Pipe and IP/MPLS strata in detail. There are
   configuration options, the basic configuration is that one Hard Pipe
   stratum and one IP/MPLS stratum are provisioned.

However, it is also possible to provision more than one Hard Pipe
stratum, e.g., if customers want enhanced separation for their leased
line. Even though the main driver for the Hard Pipe technology is
the leased lines, any service for which an operator does not want to
use statistical multiplexing will benefit from using the Hard Pipes.

2.1. The Physical Network

Consider a network with 10 routers and all the links between are 10G
Ethernet, such as shown in Figure 1. This is the network topology
we've used for this model and also (with topology variations) in our
first deployment.

```
         +---+      10G    +---+    10G     +---+   10G     +---+
     +---| B |-----------| C |-----------| D |----------| E |---+
10G  |   +---+             +---+           +---+           +---+   | 10G
     |     |                 |               |               |     |
   +---+   | 10G    10G     |             10G |           10G |   +---+
 --| F |   |               |               |               |   | G |--
   +---+   |               |               |               |   +---+
     |     |                 |               |               |     |
10G  |   +---+             +---+           +---+           +---+   | 10G
     +---| H |-----------| J |-----------| K |----------| L |---+
         +---+      10G  +---+  10G       +---+   10G    +---+
```
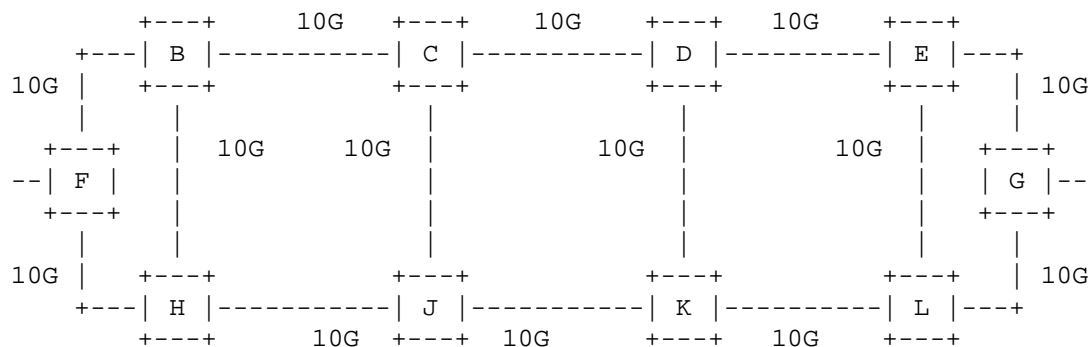
Figure 1

In this document, we use the terms "traffic matrix" or "estimated
traffic matrix" to indicate an estimate of how much traffic will flow
between the ingress and egress (PE) nodes. This may be translated
into how much bandwidth is needed per link in the Hard Pipe stratum.

2.2. The Hard Pipe Stratum

When the intention is to define a Hard Pipe stratum, it is, for
example, possible to start from an estimated traffic matrix to
estimate how much bandwidth to reserve on the links of the Ethernet
link-layer network for the Hard Pipes.

Note that the implication is that the normal traffic gets the
remainder of the available bandwidth. Thus, the link-layer network
will be split into two logical networks, or two strata -- one stratum
for the hardened pipe network and the other for the "normal" IP and
MPLS traffic. This is shown in Figures 2 and 3.

```
       +---+     2G       +---+                              +---+
   +---| B |-----------| C |                              | E |---+
1G |   +---+             +---+                              +---+   |   2G
   |     |                 |                                  |     |
 +---+                    2G |                           1G |   +---+
--| F |                    |                              | | G |--
 +---+                     |                              |   +---+
   |                       |                              |     |
1G |   +---+             +---+           +---+           +---+   |   2G
   +---| H |-----------| J |-----------| K |-----------| L |---+
       +---+     2G     +---+     4G     +---+     4G     +---+
```
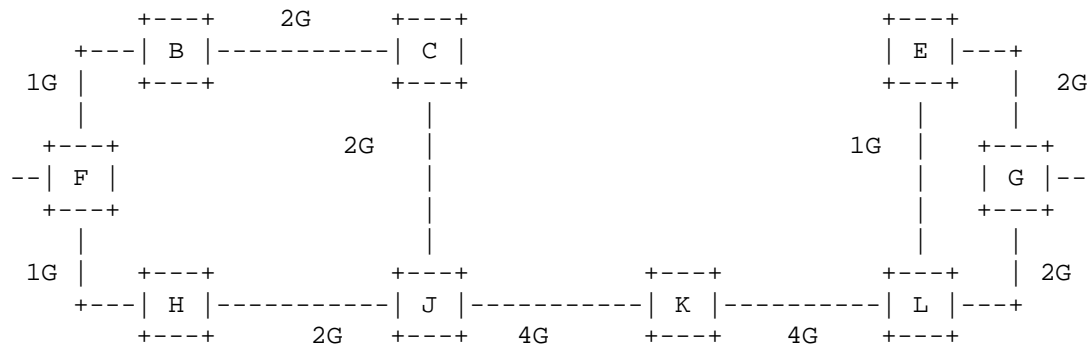
                 Figure 2: The Hard Pipe Stratum

   It is worth noting that even if the figures in this document are
   drawn to indicate "bandwidth on the link", the only bandwidth
   information that the nodes have available is the bandwidth assigned
   to the Hard Pipe stratum and the Normal IP/MPLS stratum. All other
   information is kept on the NMS/Controller. The NMS/Controller keeps
   a global bandwidth resource table for the Hard Pipe stratum.

2.3. The Normal IP/MPLS Stratum

   Given that the starting point is the physical network in Figure 1 and
   the Hard Pipe stratum as defined in Figure 2, the Normal IP/MPLS
   stratum will look as is shown in Figure 3:

```
       +---+     8G      +---+    10G    +---+    10G    +---+
   +---| B |-----------| C |-----------| D |----------| E |---+
9G |   +---+             +---+           +---+           +---+   |   8G
   |     |                 |               |               |     |
 +---+   | 10G      8G |               10G |           9G |   +---+
--| F |  |             |                 |               | | G |--
 +---+   |             |                 |               |   +---+
   |     |             |                 |               |     |
9G |   +---+         +---+             +---+           +---+   |   9G
   +---| H |-----------| J |-----------| K |----------| L |---+
       +---+     8G     +---+     6G     +---+     6G     +---+
```
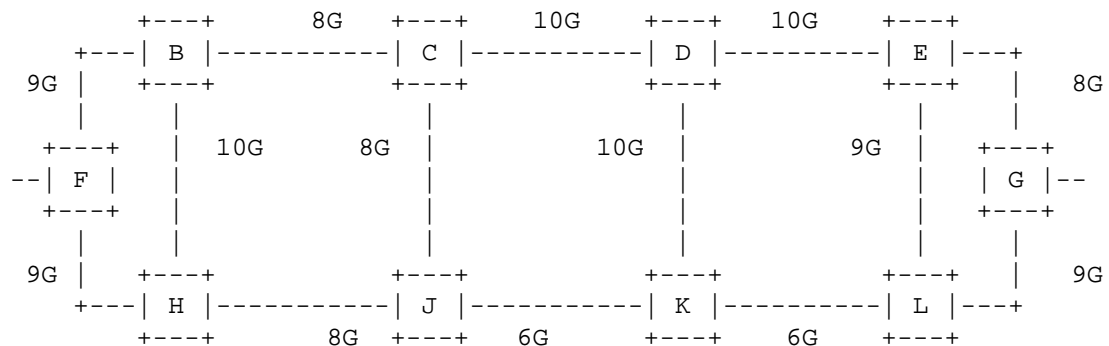
                Figure 3: The Normal IP/MPLS Stratum

2.4. Stratum Networks

   In this document, the concept of stratum network is used to indicate
   basically parallel logical networks with strictly separated
   resources. Traffic sent over one stratum network can not infringe on
   traffic in the other stratum network.

   In the case described here, all the traffic in the Hard Pipe stratum
   is MPLS encapsulated. A number of the labels have been set aside so
   other applications can't allocate them and so the routers recognize
   them as belonging to the Hard Pipe application.

3. Configuring the Leased Lines in the Hard Pipe Stratum

   When the strata are provisioned, the IP/MPLS stratum is set up
   exactly as any other IP/MPLS network. The one small difference
   between provisioning the Hard Pipe stratum and the IP/MPLS stratum is
   that no overbooking is done for the Hard Pipe stratum.

   Overbooking and/or congestion in the IP/MPLS stratum can not affect
   the Hard Pipe stratum.

   All labels used for the Hard Pipe stratum are "Configured Labels",
   i.e., labels that are provisioned and reclaimed by management
   actions. These management actions can be by manual actions or by an
   NMS/Controller or a centralized controller. For the size of network
   being deployed, manual configuration is not practical; we are both
   provisioning and reclaiming a label from an NMS/Controller.

   o If an operator wants to set up a leased line, it is first checked
      if there is a path available in the Hard Pipe stratum that matches
      the criteria (e.g., bandwidth) for the requested leased line.

      * If such a path does exist, it is checked if there is a matching
         MPLS tunnel available over that path.

         + If such a tunnel exists, it is used to establish the leased
            line by adding L-labels forming an LSP that are carried by
            the tunnel. L-labels are known only by the ingress and
            egress LSRs. They are local to the endpoints the same way
            that the label signaled by Targeted LDP (tLDP) is local to
            the endpoints of a targeted session LSP. (Here, "Targeted
            LDP" means LDP as defined in RFC 5036 [RFC5036], using
            Targeted Hello messages.)

            At the same time, the available bandwidth in the Hard Pipe
            stratum is decremented by the bandwidth that is needed for
            the leased line for every hop across this stratum in the
            global resource table (for the Hard Pipe stratum).

         + If such a tunnel does not exist, it can be established so
            that the leased line can be set up as above.

      * If the path does not exist (not enough bandwidth in the Hard
         Pipe stratum for the leased line), available bandwidth on the
         links is checked to see if the stratum can be expanded to
         accommodate such a path.

         + If the Hard Pipe stratum can be expanded, this is done and
            the tunnel for the leased line is established as described
            above.

            It is likely that other modifications of the Hard Pipe
            stratum, e.g., consolidating already set up Hard IP tunnels
            on to existing links so that room for new leased lines are
            created, may have implications that go well outside the
            leased line service, and it is currently not viewed as a
            fully automated operation.

         + If it is not possible to expand the Hard Pipe stratum to
            accommodate the new path, set up of the leased line will
            need to be declined.

   Thus, given the existence of a viable Hard Pipe stratum, leased lines
   are configured in two very simple steps. First, establish a hop-by-
   hop tunnel (T-labels), and second, configure the leased lines
   (L-labels). The T-labels need to be configured on both the PE and P
   routers while L-labels only need to be configured on the PE routers.

   Note that L-labels may be used for normal IP service [RFC3031],
   BGP/MPLS VPNs [RFC4364], or PWs [RFC3985].

4. Efficient State Management

   The system as described here generates a very small amount of state,
   and most of it is kept in the NMS/Controller.

4.1. State in the Forwarding Plane

   The only configured information that is actually kept on the LSRs is

   o the information needed for the label swapping procedures, i.e.,
      incoming label to outgoing label and port, and whether the label
      belongs to the set of labels that are set aside for the Hard Pipe
      stratum tunnels; and

   o the bandwidth available for the Hard Pipe stratum and the Normal
      IP/MPLS stratum.

4.2. State in the NMS/Controller

   The following state needs to be kept in the NMS/Controller:

   o the topology and bandwidth resources available in the Hard Pipe
      network; see Figure 2.

   o the total and available bandwidth per link in the Hard Pipe
      network; see Figure 4.

   o the T-label mappings; see Figure 5.

   o the L-label mappings; see Figure 6.

   o the reserved bandwidth, as well as other constraints and the path
      per L-label.

4.3. Annotations for Configuring Leased Lines

   The annotations given below are neither a programming guideline nor
   an indication how this architecture could be implemented. It is
   rather an indication of how much data needs to be saved for each
   stratum and leased line, as well as where this data could be stored.

   Considering the Hard Pipe stratum as it has been outlined in
   Figure 2, there is actually some additional information related to
   the Hard Pipe stratum that not is shown in the figure.

   Looking explicitly on the link between LSR J and K we find:

```
       +---+             +---+             +---+             +---+
      ---| H |-----------| J |-----------| K |----------| L |---
       +---+             +---+             +---+             +---+
                              [4,0]G
```

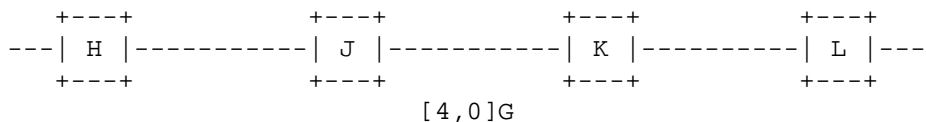   Figure 4

   The annotation [4,0]G means that 4G is allocated to the stratum on
   the link between J and K, and of these, 0G has been allocated to a
   service.

If we were to allocate two tunnels labels from the labels that have
been configured to work within the Hard Pipe stratum, the resource
view would look like this:

```
    +---+              +---+              +---+             +---+
 ---| H |-----------| J |-----------| K |----------| L |---
    +---+              +---+              +---+             +---+
                      [4,0]G T1 ,T2
```
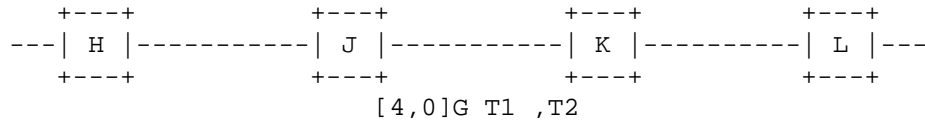
                              Figure 5

Note that allocating the tunnel labels does not reserve bandwidth for
the tunnel from the Hard Pipe stratum.

When the L-labels are assigned, this will consume bandwidth; so we
need to keep track of the bandwidth per leased line and the total of
bandwidth allocated from the Hard Pipe stratum.

The annotation for the link between J and K could look like this:

```
    +---+              +---+              +---+             +---+
 ---| H |-----------| J |-----------| K |----------| L |---
    +---+              +---+              +---+             +---+
         [4,1.5]G, T1, L1 [.5], L2 [.5], T2, L1 [.5]
```
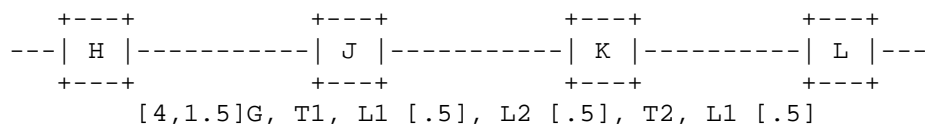
                              Figure 6

The line [4,1.5]G, T1, L1 [.5], L2 [.5], T2, L1 [.5] would be
interpreted as follows:

   The Hard Pipe stratum link between nodes J and K has 4G bandwidth
   allocated; of the total bandwidth, 1.5G is allocated for leased
   lines.

   Tunnel label T1 carries two leased lines, each of 0.5G, and tunnel
   label T2 carries a third leased line of 0.5G.

Note that it is not necessary to keep this information in the nodes;
it is held within the NMS/Controller. Also, it is not necessary to
keep the bandwidth per leased line, but some operations are
simplified (e.g., removing a leased line) if this is done.

5. Setting Up Leased Lines

   Consider the case where an operator wants to set up a leased line of
   0.4G from F to G in the Hard Pipe stratum in Figure 2.

   Since there are no constraints other than bandwidth and ingress and
   egress PEs, the shortest path will be chosen. A tunnel will be
   configured from F to G over the nodes F, H, J, K, L, and G, and a
   Leased Line label (a) will be configured on F and G, and the
   available resources will be recalculated.

   A second leased line of 0.3G between the same PEs is easily
   configured by adding a new Leased Line label (b) at the ingress and
   egress PEs.

   After these operations, a view of the Hard Pipe stratum resources
   (available bandwidth) would look like this:

```
          +---+     2G        +---+                          +---+
       +---| B |-----------| C |                          | E |---+
    1G |    +---+              +---+                          +---+     | 2G
       |                        |                             |     |
     +---+            2G        |                       1G  |    +---+
   --| F |                      |                             |    | G |--
     +---+                      |                             |    +---+
       |                        |                             |     |
   .3G |    +---+              +---+         +---+          +---+   | 1.3G
       +---| H |-----------| J |-----------| K |----------| L |---+
           +---+    1.3G    +---+    3.3G    +---+   3.3G    +---+
```
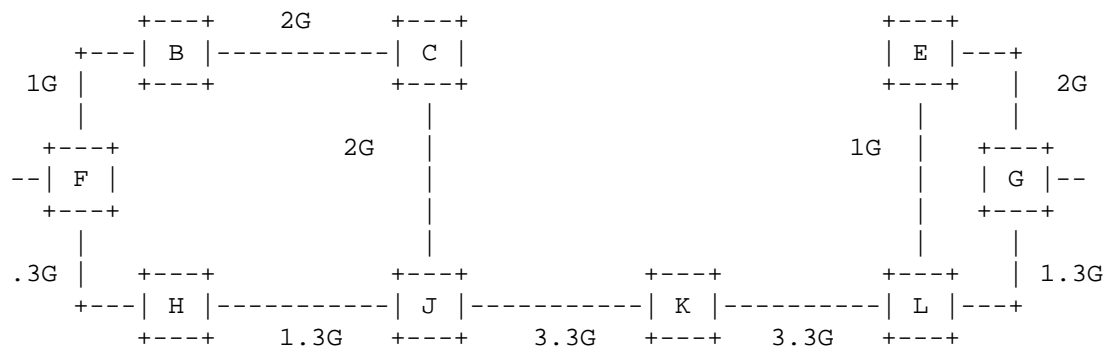
                Figure 7: The Hard Pipe Stratum after Operations

   If the operator now wishes to establish a new leased line with the
   criteria being that it should originate from F and terminate at G,
   have 0.4G bandwidth, and pass through node E, then analysis of the
   Hard Pipe stratum (after establishing the first two listed lines) and
   the criteria for the new leased line would give the following:

   o The existing tunnel cannot be used since it does not pass through
      E; a new tunnel need to be established.

   o The hop from F to H cannot be used since the available bandwidth
      is insufficient.

   o Since no existing tunnels meet the criteria requested, a new
      tunnel will be set up from F, to B, C, J, K, L, E (the criteria to
      pass through E), and to G.

A new L-label (c) to be carried over T2 will be configured on F and
G, and the available resources of the Hard Pipe stratum will be
recalculated.

6. Leased Line Protection

This leased line service uses the MPLS Transport Profile (MPLS-TP)
line protection as it is defined in RFC 6378 [RFC6378] and is updated
as specified in RFC 7271 [RFC7271] and RFC 7324 [RFC7324]

The CV and CC are run over the tunnels between the Maintenance Entity
Group End Points (MEP) at each end, i.e., the entire tunnel is
protected end to end.

In general, all of the MPLS-TP Operations, Administration, and
Maintenance (OAM) as defined in RFC 6371 [RFC6371] is v applicable.

7. Security Considerations

The security considerations as defined in "Security Framework for
MPLS and GMPLS Networks" (RFC 5920 [RFC5920]) and "MPLS Transport
Profile (MPLS-TP) Security Framework" (RFC 6941 [RFC6941]) apply to
this document.

8. Informative References

[RFC3031]   Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol
            Label Switching Architecture", RFC 3031,
            DOI 10.17487/RFC3031, January 2001,
            <http://www.rfc-editor.org/info/rfc3031>.

[RFC3985]   Bryant, S., Ed. and P. Pate, Ed., "Pseudo Wire Emulation
            Edge-to-Edge (PWE3) Architecture", RFC 3985,
            DOI 10.17487/RFC3985, March 2005,
            <http://www.rfc-editor.org/info/rfc3985>.

[RFC4364]   Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private
            Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February
            2006, <http://www.rfc-editor.org/info/rfc4364>.

[RFC5036]   Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed.,
            "LDP Specification", RFC 5036, DOI 10.17487/RFC5036,
            October 2007, <http://www.rfc-editor.org/info/rfc5036>.

[RFC5920]   Fang, L., Ed., "Security Framework for MPLS and GMPLS
            Networks", RFC 5920, DOI 10.17487/RFC5920, July 2010,
            <http://www.rfc-editor.org/info/rfc5920>.

   [RFC5921] Bocci, M., Ed., Bryant, S., Ed., Frost, D., Ed., Levrau,
              L., and L. Berger, "A Framework for MPLS in Transport
              Networks", RFC 5921, DOI 10.17487/RFC5921, July 2010,
              <http://www.rfc-editor.org/info/rfc5921>.

   [RFC6371] Busi, I., Ed. and D. Allan, Ed., "Operations,
              Administration, and Maintenance Framework for MPLS-Based
              Transport Networks", RFC 6371, DOI 10.17487/RFC6371,
              September 2011, <http://www.rfc-editor.org/info/rfc6371>.

   [RFC6378] Weingarten, Y., Ed., Bryant, S., Osborne, E., Sprecher,
              N., and A. Fulignoli, Ed., "MPLS Transport Profile
              (MPLS-TP) Linear Protection", RFC 6378,
              DOI 10.17487/RFC6378, October 2011,
              <http://www.rfc-editor.org/info/rfc6378>.

   [RFC6941] Fang, L., Ed., Niven-Jenkins, B., Ed., Mansfield, S., Ed.,
              and R. Graveman, Ed., "MPLS Transport Profile (MPLS-TP)
              Security Framework", RFC 6941, DOI 10.17487/RFC6941, April
              2013, <http://www.rfc-editor.org/info/rfc6941>.

   [RFC7271] Ryoo, J., Ed., Gray, E., Ed., van Helvoort, H.,
              D'Alessandro, A., Cheung, T., and E. Osborne, "MPLS
              Transport Profile (MPLS-TP) Linear Protection to Match the
              Operational Expectations of Synchronous Digital Hierarchy,
              Optical Transport Network, and Ethernet Transport Network
              Operators", RFC 7271, DOI 10.17487/RFC7271, June 2014,
              <http://www.rfc-editor.org/info/rfc7271>.

   [RFC7324] Osborne, E., "Updates to MPLS Transport Profile Linear
              Protection", RFC 7324, DOI 10.17487/RFC7324, July 2014,
              <http://www.rfc-editor.org/info/rfc7324>.

Acknowledgements

Authors' Addresses

   JiangTao Hao
   Huawei Technologies Co., Ltd
   Q13 Huawei Campus
   No. 156 Beiqing Road
   Hai-dian District
   Beijing 100095
   China
   Email: haojiangtao@huawei.com

   Praveen Maheshwari
   Bharti Airtel, Ltd.
   Plot No. 16, Udyog Bihar,
   Phase IV, Gurgaon - 122015
   Haryana
   India
   Email: Praveen.Maheshwari@in.airtel.com

   River Huang
   Huawei Technologies Co., Ltd
   Q13 Huawei Campus
   No. 156 Beiqing Road
   Hai-dian District
   Beijing 100095
   China
   Email: river.huang@huawei.com

   Loa Andersson
   Huawei Technologies Co., Ltd
   Stockholm
   Sweden
   Email: loa@mail01.huawei.com

   Mach(Guoyi) Chen
   Huawei Technologies Co., Ltd
   Q14 Huawei Campus
   No. 156 Beiqing Road
   Hai-dian District
   Beijing 100095
   China
   Email: mach.chen@huawei.com