Internet Engineering Task Force (IETF)   A. Malis, Ed.
Internet-Draft   Huawei Technologies
Intended status: Informational   R. Skoog
Expires: August 17, 2014   H. Kobrinski
                                        Applied Communication Sciences
                                                              G. Clapp
                                                     AT&T Labs Research
                                                              J. Drake
                                                               Juniper
                                                             V. Shukla
                                               Verizon Communications
                                                     February 13, 2014

                 Requirements for Very Fast Setup of GMPLS LSPs
                        draft-malis-ccamp-fast-lsps-01

Abstract

   The Defense Advanced Research Projects Agency (DARPA) Core Optical
   Networks (CORONET) program has laid out a vision for the next
   evolution of IP and optical commercial and government networks, with
   a focus on highly dynamic and resilient multi-terabit core networks.
   It anticipates the need for rapid (sub-second) setup and SONET/SDH-
   like restoration times for high-churn (up to tens of requests per
   second network-wide and one second to one minute holding times) on-
   demand wavelength, sub-wavelength and packet services for a variety
   of applications (e.g., grid computing, cloud computing, data
   visualization, fast data transfer, etc.). This must be done while
   meeting stringent call blocking requirements, and while minimizing
   the use of resources such as time slots, switch ports, wavelength
   conversion and wavelength-km.

   This document discusses the requirements for extensions to
   Generalized Multi-Protocol Label Switching (GMPLS) signaling for
   expediting the control of Label Switched Paths (LSPs), including sub-
   wavelengths (e.g., OTN ODUs) and full wavelengths, in order to
   satisfy application requirements laid out in this program.

Status of This Memo

   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."

   This Internet-Draft will expire on August 17, 2014.

Copyright Notice

Table of Contents

1. Introduction

   The Defense Advanced Research Projects Agency (DARPA) Core Optical
   Networks (CORONET) program [Chiu] has laid out a vision for the next
   evolution of IP and optical commercial and government networks, with
   a focus on highly dynamic and resilient multi-terabit core networks.
   The program anticipates an environment where there are multiple
   Bandwidth-on-Demand service requests per second, such as might arise
   as cloud services proliferate. It includes dynamic services with
   connection setup requirements that are two to three orders of
   magnitude faster than possible with current connection setup

protocols. The aggregate traffic demand, which is composed of both
packet (IP) and circuit (wavelength and sub-wavelength) services,
represents a five to twenty-fold increase over today's traffic levels
for the largest of any individual carrier. It is the desired goal of
the program to achieve transition of these advances to commercial and
government networks in the next few years. Thus, the aggressive
requirements must be met with solutions that are scalable, cost
effective, and power efficient, while providing the desired quality
of service (QoS).

Thus, CORONET anticipates the need for rapid (sub-second) setup and
restoration times for high-churn (up to tens of requests per second
network-wide and one second to one minute holding times) on-demand
wavelength, sub-wavelength and packet services for a variety of
applications (e.g., grid computing, cloud computing, data
visualization, fast data transfer, etc.). This must be done while
meeting stringent call blocking requirements, and while minimizing
the use of resources such as time slots, switch ports, wavelength
conversion and wavelength-km.

GMPLS protocols and procedures have been developed to enable
automated control of Label Switched Paths (LSPs), including setup,
teardown, modification, and restoration, for switching technologies
extending from layer 2 and layer 3 packets, to time division
multiplexing, to wavelength, and to fiber.

However, while the current GMPLS constituent protocols are geared for
a wide scope of applications and robust performance, they have not
specifically addressed the more aggressive characteristics envisioned
here, e.g., applications requiring low connection setup times while
maintaining a high success ratio (i.e., low blocking) in a high-churn
environment. For example, in Internet2, a network which provides
CORONET-like high bandwidth circuit services for the Research &
Education community, a circuit is currently established, on average,
roughly at a rate of one per hour. In contrast, the CORONET vision
is a churn rate of up to tens of circuits per second, over four
orders of magnitude greater.

Furthermore, scenarios with highly dynamic connection request
activity, where the connection request arrival rate is higher than
the TE update rate allowed by OSPF-TE, could lead to unacceptable
blocking ratios or low resource utilization. The purpose of this
draft is to determine the requirements to augment the GMPLS framework
to allow specific applications, or users, to rapidly set up
connections over GMPLS networks with minimal delays and a high
probability of success.

Preliminary simulations and analyses of national and global scale
networks, both WSON and sub-wavelength OTN, have shown that using
current GMPLS protocols and procedures does not meet the CORONET
performance targets with respect to blocking, setup delays, and
resource utilization. These simulations have also indicated limited
scalability of current protocols to increasing loads and churn beyond
the baseline design. Some of the factors affecting these results in
a highly dynamic network include:

1. Stale TE information when the connection request rate exceeds TE
    information update rate based on OSPF-TE LS updates. This leads
    to increased blocking and indirectly to longer setup delays.

2. Real-time path computation and PCE communication, i.e., following
    connection request, thus increasing setup delays.

3. Cross-connection procedures resulting in accumulating cross-
    connection delays when cross-connection must be completed before
    the Resv signaling message is propagated upstream. This
    contribution may be significant in WSON but less so with TDM or
    L2 switching.

4. Crankbacks.

2. Scope and Motivation

   [RFC6163] provides the framework, basic elements, and terminology of
   wavelength switched optical networks (WSON) and wavelength-based
   LSPs. These basic elements generally apply to other GMPLS
   technologies as well, e.g., spectral switching (SSON), sub-wavelength
   TDM, and L2 LSPs. This draft refers to the same general framework
   and technologies, but addresses an extension of the general problem
   space addressed in [RFC6163]. Specifically, this draft addresses the
   requirements of expediting LSP setup, under heavy connection churn
   scenarios, while achieving low blocking, under an overall distributed
   control plane. Once there is agreement on the requirements, further
   drafts will describe the procedures and signaling contents required
   to meet the requirements (potentially more than one if separate
   standard track drafts are found necessary for wavelength and sub-
   wavelength LSPs). Both single-domain and multi-domain network
   scenarios are addressed. A connection setup delay is defined here as
   the time between the arrival of a connection request at an ingress
   edge switch - or more generally a Label Switch Router (LSR) - and the
   time at which information can start flowing from that ingress switch
   over that connection. Note that this definition is more inclusive
   than the LSP setup time defined in [RFC5814] and [RFC6777], which do
   not include PCE path computation delays.

The motivation for GMPLS extensions as described here is thus two-fold:

1. The anticipated need for rapid setup while maintaining low
    blocking, on-demand, of large bandwidth connections (in the form
    of sub-wavelengths, e.g., OTN ODUx, and wavelengths, e.g., OTN
    OCh) for a variety of applications including grid computing,
    cloud computing, data visualization, and intra- and inter-
    datacenter communications.

2. The performance of current GMPLS protocols and procedures in
    networks with the above characteristics.

The ability to setup circuit-like LSPs for large bandwidth flows and
with low setup delays provides an alternative to packet-based
solutions implemented over static circuits that may require tying up
more expensive and power-consuming resources (e.g., router ports).
Reducing the LSP setup delay will reduce the minimum bandwidth
threshold at which a GMPLS approach is preferred over a layer 3
(e.g., IP) approach. Dynamic circuit and virtual circuit switching
intrinsically provide guaranteed bandwidth, guaranteed low-latency
and jitter, and faster restoration, all of which are very hard to
provide in a packet-only networks. Again, a key element in achieving
these benefits is enabling the fastest possible circuit setup times.

Future applications are expected to require setup times as fast as
100 ms in highly dynamic, national-scale network environments while
meeting stringent blocking requirements and minimizing the use of
resources such as switch ports, wavelength converters/regenerators,
wavelength-km, and other network design parameters. Of course, the
benefits of low setup delay diminish for connections with long
holding times.

The need for rapid setup for specific applications may override and
thus get traded off against some other features currently provided in
GMPLS, e.g., robustness against setup errors.

With the advent of data centers, cloud computing, video, gaming,
mobile and other broadband applications, it is anticipated that
connection request rates may increase, even for connections with
longer holding times, either during limited time periods (such as
during the restoration from a data center failure) or over the longer
term, to the point where the current GMPLS maximum frequency of TE
information updates is not sufficient to provide adequate path
computation and resource allocation, as network conditions and
resource attributes may be changing faster than can be reflected in
OSPF-TE updates.

Thus, GMPLS and routing protocol traffic engineering (e.g. OSPF-TE) extensions are also needed to address heavy churn of connection requests (i.e., high connection request arrival rate) in networks with high traffic loads, even for connections with relatively longer holding times.

3. Requirements for Very Fast Setup of GMPLS LSPs

   This section lists the requirements for very fast setup of GMPLS LSPs in order to provide the services described in the previous sections. They will be the basis for future standards-track drafts to satisfy these requirements. Some of these requirements may be implementation-dependent to some extent, but they may also have LSP signaling protocol dependencies as well. Protocols that satisfy these requirements can be further compared based on other important factors such as resource efficiency, and implementation complexity.

   The requirements are divided in two general categories - control plane requirements and network requirements. Note that network requirements essentially reflect DARPA CORONET program requirements, but anticipate cloud and other emerging application requirements. The networks considered in the CORONET program are primarily long haul national and global networks. The model for a national network is that of the continental US with up to 100 nodes and LSPs distances up to ~3000 km and up to 15 hops.

3.1. Control Plane Requirements

   R1 Protocol extensions must be backward compatible with existing GMPLS control plane protocols.

   R2 Use of GMPLS protocol extensions for this application must be selectable by provisioning or configuration.

   R3 Must support the use of PCE for path computation, and in particular the PCE-based approach for multi-domain LSPs in [RFC5441].

3.2. Network Requirements

   R4 Must have an LSP setup time less than or equal to 100 ms for intra-continental LSPs, and less than or equal to 250 ms for transcontinental LSPs, including PCE path computation delays.

   R5 Must support LSP holding times of one second to one minute.

   R6 While there are implementation-dependent aspects of supporting high LSP setup rates, the protocol aspects of LSP signaling must

not preclude LSP request rates of tens per second. A possible
example of a protocol aspect is the ability to update the IGP TE
database to accurately reflect resource availability at all
times. Note that LSP request rates may be dependent on LSP
bandwidth, where very high bandwidth LSPs (such as for an entire
wavelength) could be less frequent than lower-rate LSPs (such as
an ODUx connection).

R7 Must support restoration for all cases of single node or link
failures.

R8 At most one blocked LSP setup request per 1000 requests. LSP
setup blocking depends on network variables (topology, available
resources) and on the setup protocol. The choice of selected
protocol is primarily determined by the level of resource
utilization.

4. IANA Considerations

This memo includes no request to IANA.

5. Security Considerations

Being able to support very fast setup and a high churn rate of GMPLS
LSPs is not expected to adversely affect the underlying security
issues associated with existing GMPLS signaling, and potentially
could improve GMPLS' resistance against denial of service attacks
that attempt to deny service through the use of a high frequency of
GMPLS LSP setup requests.

6. Acknowledgements

The authors would like to thank Ann Von Lehmen, Joe Gannett, and
Brian Wilson of Applied Communication Sciences for their comments and
assistance on this document.

7. References

7.1. Normative References

[RFC5441] Vasseur, JP., Zhang, R., Bitar, N., and JL. Le Roux, "A
           Backward-Recursive PCE-Based Computation (BRPC) Procedure
           to Compute Shortest Constrained Inter-Domain Traffic
           Engineering Label Switched Paths", RFC 5441, April 2009.

[RFC5814] Sun, W. and G. Zhang, "Label Switched Path (LSP) Dynamic
           Provisioning Performance Metrics in Generalized MPLS
           Networks", RFC 5814, March 2010.

   [RFC6163] Lee, Y., Bernstein, G., and W. Imajuku, "Framework for
              GMPLS and Path Computation Element (PCE) Control of
              Wavelength Switched Optical Networks (WSONs)", RFC 6163,
              April 2011.

   [RFC6777] Sun, W., Zhang, G., Gao, J., Xie, G., and R. Papneja,
              "Label Switched Path (LSP) Data Path Delay Metrics in
              Generalized MPLS and MPLS Traffic Engineering (MPLS-TE)
              Networks", RFC 6777, November 2012.

7.2. Informative References

   [Chiu]   A. Chiu, et al, "Architectures and Protocols for Capacity
              Efficient, Highly Dynamic and Highly Resilient Core
              Networks", Journal of Optical Communications and
              Networking vol. 4, No. 1, pp. 1-14, January 2012,
              <http://dx.doi.org/10.1364/JOCN.4.000001>.

Authors' Addresses

   Andrew G. Malis (editor)
   Huawei Technologies

   Email: agmalis@gmail.com

   Ronald A. Skoog
   Applied Communication Sciences

   Email: rskoog@appcomsci.com

   Haim Kobrinski
   Applied Communication Sciences

   Email: hkobrinski@appcomsci.com

   George Clapp
   AT&T Labs Research

   Email: clapp@research.att.com

   John E. Drake
   Juniper

   Email: jdrake@juniper.net

Vishnu Shukla
Verizon Communications

Email: vishnu.shukla@verizon.com