Network Working Group   Ina Minei
Internet Draft   Der-Hwa Gan
Expiration Date: December 2006   Kireeti Kompella
Category: Informational   Juniper Networks

                                                     Xiaoming Li
                                                    China Unicom

                                                       June 2006

  Extensions for Differentiated Services-aware Traffic Engineered LSPs

                   draft-minei-diffserv-te-multi-class-02.txt

Status of this Memo

Abstract

   This document specifies the extensions necessary to set up multi-
   class diffserv-aware traffic engineered label switched paths (LSPs).
   A multi-class diffserv-te LSP carries traffic from multiple traffic
   classes and is set up along a path that satisfies the bandwidth
   constraints defined for each of the classes it carries.

Conventions used in this document

   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 [RFC2119].

Table of Contents

1. **Definitions**

   For readability a number of definitions from [RFC3564] are repeated
   here:

   Traffic Trunk: an aggregation of traffic flows of the same class
   [i.e. which are to be treated equivalently from the DS-TE perspec-
   tive] which are placed inside a Label Switched Path.

   Class-Type (CT): the set of Traffic Trunks crossing a link that is
   governed by a specific set of Bandwidth constraints. CT is used for
   the purposes of link bandwidth allocation, constraint based routing
   and admission control. A given Traffic Trunk belongs to the same CT
   on all links.

   TE-Class: A pair of:
      1. a Class-Type
      2. a preemption priority allowed for that Class-Type. This means
         that an LSP transporting a Traffic Trunk from that Class-Type
         can use that preemption priority as the set-up priority, as the
         holding priority or both.

2. **Introduction**

   [RFC4124] defines the protocol extensions for setting up diffserv-
   aware traffic engineered LSPs. An LSP set up according to [RFC4124]
   carries traffic from a single diffserv class and is set up along a
   path that satisfies the bandwidth constraints specified for this
   class. In this case, a single Class-Type is configured per LSP. In
   the rest of this document such an LSP is called single-class diff-
   serv-TE LSP. Note that from a diffserv point of view, a single-class
   LSP can be either an L-LSP or an E-LSP (as defined in [RFC3270]).

   In some scenarios it is required to allow traffic with different
   diffserv behaviors to be mapped to the same LSP and to traffic engi-
   neer the LSP according to the bandwidth constraints for each one of
   these classes. In this case, multiple Class-Types are configured per
   LSP. In the rest of the document such an LSP is called a multi-class
   diffserv-TE LSP. From a diffserv point of view, a multi-class LSP is
   always an E-LSP.

   An example scenario for multi-class LSPs comes in the context of ATM
   trunk emulation using MPLS LSPs. In this case, it is preferable to
   have all the traffic classes following the same path and exhibit the
   same behavior in case of failure.

   Another application of multi-class diffserv-TE LSPs is for reducing

the number of LSPs in a network by setting up reservations for sev-
eral classes in one LSP rather than one reservation per class. With
single class LSPs, the total number of LSPs in the network is equal
to the number of classes times the number of LSPs in the mesh. With
multi-class LSPs, the total number of LSPs is equal to the size of
the LSP mesh. The reduction in the number of LSPs is important from a
scaling and manageability point of view.

Here are a few observations regarding single-class and multi-class
LSPs:
  o An LSP that is signaled using the extensions defined in this doc-
    ument is not required to request bandwidth from multiple Class-
    Types. As a special case, when the request is from a single
    Class-Type, the LSP behaves as a single-class LSP and from a
    diffserv point of view can be either an L-LSP or an E-LSP. There-
    fore, conceptually, multi-class LSPs are a superset of single-
    class LSPs.

  o In order to achieve bandwidth reservation from several traffic
    classes, two approaches can be taken: 1) create one multi-class
    LSP or 2) create several single-class LSPs. Note however that the
    two approaches have different properties in terms of: the number
    of LSPs established, the path that is taken by traffic belonging
    to different classes and the fate sharing between the different
    classes. The choice of single-class or multi-class LSP is made
    based on the requirements of the application and on the network
    design.

  o Both single-class and multi-class LSPs must integrate smoothly
    with non-diffserv-te LSPs.

3. Setup and holding preemption priorities

As per existing TE, multi-class LSPs can be configured with a setup
and holding priority, each with a value between 0 and 7. The combi-
nation of the setup priority and each of the Class-Types for which
the multi-class LSP is set up and of the hold priority and each of
the Class-Types for which the multi-class LSP is set up, MUST be such
that together they form one of the (up to) 8 TE-Classes configured in
the TE-Class Mapping. This requirement is similar to the requirement
spelled out in [RFC4124] regarding the combination of priority and
class-type for single-class LSPs.

4. **Setting up multi-class RSVP-signaled LSPs**

4.1. **The extended classtype object**

   To request LSP setup along a path with bandwidth constraints from
   more than one Class-Type, a new object is defined, the extended-
   classtype-object. This object has a class_num that ensures that a
   node that doesn't recognize it will reject it and return an "Unknown
   Object Num" error. According to the guidelines in [RFC3936], the val-
   ues used are class_num 124 and "class class_type" 1.

   The contents are a set of subobjects, each encoded as a TLV.

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                    Subobjects                                 |
   .                     ...                                       .
   .                     ...                                       .
   .                     ...                                       .
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

   A TLV has the following format:

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   | Type          | Length                        | MBZ   | CTi |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |     BW requested for CTi (32-bit IEEE floating point number)  |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```


   Length: 16 bits
         The length contains the total length of the subobject in
         bytes, including the type, length, mbz, CTi fields.
         The length MUST be at least 8, and MUST be a multiple of 4.

   Type : 8 bits
         The type of the contents of the subobject. Currently defined
         value is
         1 - bandwidth

   MBZ: 3 bits
         This field is reserved. It must be set to zero on transmission
         and must be ignored on receipt.

   CT : 3 bits

        Indicates the Class-Type. Values currently allowed are 0, 1,
   2, ... , 7. Note that this is different from the definition
   given in [RFC4124] where the value 0 is not allowed.

   BW requested : 32 bits
        Indicates the number of bytes (not bits) per second that need
   to be reserved for the CT indicated.

4.2. RSVP signaling

   The extended-classtype object is signaled in the PATH message, and
   saved in the PSB at every hop. The extended-classtype object MUST NOT
   be included in a RESV message.

   The format of the Path message is as follows:

     <Path Message> ::=   <Common Header> [<INTEGRITY>]
                               <SESSION> <RSVP_HOP> <TIME_VALUES> [
   <EXPLICIT_ROUTE>]
                               <LABEL_REQUEST> [<SESSION_ATTRIBUTE>]
                               [<DIFFSERV>]
                               [<EXTENDED_CLASSTYPE>]
                               [<POLICY_DATA> ...]
                               [<sender descriptor>]

     <sender descriptor> ::= <SENDER_TEMPLATE> [<SENDER_TSPEC>]
                               [<ADSPEC>]
                               [<RECORD_ROUTE>]

   When the classtype object is present, the values in the tspec are
   ignored for admission control purposes and SHOULD be 0.

   Admission control is performed for the requirements specified in each
   of the (CT, BW) pairs in the extended-classtype object. Admission
   control succeeds only if the requirements for all the (CT, BW) pairs
   are satisfied.

   Merging MUST NOT be performed among reservations that belong to LSPs
   for which the extended-classtype object was signaled.

   If the extended-classtype object is not present in the PATH message,
   the LSR associates the LSP with CT0 and performs admission control
   from the bandwidth/priority values for CT0. In the case where the LSP
   is from a single class, and the class is CT0, the extended-classtype

   object MAY be included in the PATH message.

4.3. RSVP error processing

   If the extended-classtype object is not understood, an "Unknown
   Object Num' error is returned.

   An LSR that recognizes the extended-classtype object and that
   receives a path message which contains the extended-classtype object
   but which does not contain a LABEL_REQUEST object or which does not
   have a session type of LSP_Tunnel_IPv4, MUST send a PathErr towards
   the sender with the error code 'extended-classtype Error' and an
   error value of 'Unexpected extended-classtype object'. These are
   defined below.

   An LSR receiving a Path message with the extended-classtype object,
   which recognizes the object but does not support the particular
   Class-Type, MUST send a PathErr towards the sender with the error
   code 'extended-classtype Error' and an error value of 'Unsupported
   Class-Type'.

   An LSR receiving a Path message with the extended-classtype object,
   which:
     - recognizes the object,
     - supports the particular Class-Type, but
     - determines that the tuple formed by (i)this Class-Type and (ii)
       the holding priority signaled in the same Path message, is not
       one of the eight TE-classes configured in the TE-class mapping,
   MUST send a PathErr towards the sender with an error value of 'CT and
   holding priority do not form a configured TE-Class'. For setup prior-
   ity, return 'CT and setup priority do not form a configured TE-class'

   The error checking ensures that multi-class LSPs get established only
   through LSRs that support multi-class diffserv-te LSPs and have con-
   sistent TE-mappings with the ingress.

   Errors related to the processing of the extended-classtype-object are
   reported using the vendor-specific Error-Code "extended-classtype-
   error" 252. The first four octets following the Error Value are 2636
   - the vendor's SMI enterprise code in network octet order (as
   required by [RFC3936])

   The error values are the same as defined in [RFC4124] for the
   classtype object. The values 3, 6, 7 are not used with multi-class-
   LSPs.

[5](5). The extended-MAM bandwidth constraints model

   A new BC model-id is defined, the extended-mam bandwidth model, with
   value 254. The extended-mam model has the same behavior as the MAM
   model [[RFC4125](RFC4125)].

[6](6). IGP extensions

   [RFC4124] defines the bandwidth-constraint (BC) sub-tlv. The BC sub-
   tlv includes the bandwidth model and a list of bandwidth-constraints.
   In [[RFC4124](RFC4124)], the semantic of the "Unreserved Bandwidth" sub-TLV is
   changed such that the eight bandwidth values advertised correspond to
   the unreserved bandwidth for each of the TE-Class (instead of for
   each preemption priority as per existing TE). This approach has two
   drawbacks:
     o It creates a flag day in the network with regards to the existing
       TE LSPs set up without any diffserv properties, as explained in
       [[RFC4124](RFC4124)].
     o It requires that LSPs from CT0 be set up at priorities such that
       the combination of (CT0, priority) is one of the configured TE-
       classes. Thus, the priorities asssigned to pre-existing CT0 LSPs
       use up some of the TE-classes.

   In the method documented here, when the bandwidth-model is extended-
   mam, the semantic of the "Unreserved Bandwidth" sub-TLV is the same
   as in existing TE and the bandwidth available per TE-class is adver-
   tised in the BC sub-tlv. This approach achieves the following:
     o There is no interference between the diffserv-aware LSPs and the
       plain TE LSPs
     o There are no migration issues when deploying this feature in a
       network where TE LSPs are deployed.
     o A total of 16 TE-classes are effectively created.

[7](7). Migration issues

   Since the semantics of the "Unreserved Bandwidth" sub-TLV are main-
   tained unchanged, there are no migration issues when deploying multi-
   class LSPs in a network where traffic engineering is already in use.

   Single-class and multi-class LSPs cannot co-exist in the same net-
   work. When both the multi-class and the single-class functionality
   is required, the semantics of single-class LSPs can be provided by
   the multi-class LSPs.

8. Security considerations

   The same security considerations apply as for single-class diffserv-
   TE LSPs [RFC4124].

9. Intellectual Property Statement

   The IETF takes no position regarding the validity or scope of any
   Intellectual Property Rights or other rights that might be claimed to
   pertain to the implementation or use of the technology described in
   this document or the extent to which any license under such rights
   might or might not be available; nor does it represent that it has
   made any independent effort to identify any such rights. Information
   on the procedures with respect to rights in RFC documents can be
   found in BCP 78 and BCP 79.

   Copies of IPR disclosures made to the IETF Secretariat and any assur-
   ances of licenses to be made available, or the result of an attempt
   made to obtain a general license or permission for the use of such
   proprietary rights by implementers or users of this specification can
   be obtained from the IETF on-line IPR repository at
   http://www.ietf.org/ipr.

   The IETF invites any interested party to bring to its attention any
   copyrights, patents or patent applications, or other proprietary
   rights that may cover technology that may be required to implement
   this standard. Please address the information to the IETF at ietf-
   ipr@ietf.org.

10. Acknowledgments

   This work is the outcome of discussions with Arthi Ayyangar and Chai-
   tanya Kodeboyina. The authors would like to thank them for their
   suggestions and insight. The authors would like to thank Yakov
   Rekhter and Peter Busschbach for their review.

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
Requirement Levels", RFC 2119.

[RFC3270] Le Faucheur et al, "MPLS Support of Diff-Serv", RFC3270.

[RFC3564] Le Faucheur et al, "Requirements for support of Diff-Serv-
aware MPLS Traffic Engineering", RFC3564.

[RFC3936] Kompella, K., Lang, J., "Procedures for Modifying the
Resource reSerVation Protocol (RSVP)", RFC3936.

[RFC4125] Le Faucheur, Lai, "Maximum Allocation Bandwidth Constraints
Model for DS-TE", RFC4125.

[RFC4124] Le Faucheur et al, "Protocol extensions for support of
Diff-Serv-aware MPLS Traffic Engineering", RFC4124.

12. Authors' Addresses

Ina Minei
Juniper Networks
1194 N.Mathilda Ave
Sunnyvale, CA 94089
e-mail: ina@juniper.net

Der-Hwa Gan
Juniper Networks
1194 N.Mathilda Ave
Sunnyvale, CA 94089
e-mail: dhg@juniper.net

Kireeti Kompella
Juniper Networks
1194 N.Mathilda Ave
Sunnyvale, CA 94089
e-mail: kireeti@juniper.net

Xiaoming Li
China United Telecommunications Corporation
No.133A, Xidan North Street,
Xicheng District, Beijing 100032, China
email: lixm@chinaunicom.com.cn