                 Aggregation of Diffserv Service Classes

Status of This Memo

   This memo provides information for the Internet community. It does
   not specify an Internet standard of any kind. Distribution of this
   memo is unlimited.

Abstract

   In the core of a high-capacity network, service differentiation may
   still be needed to support applications' utilization of the network.
   Applications with similar traffic characteristics and performance
   requirements are mapped into Diffserv service classes based on end-
   to-end behavior requirements of the applications. However, some
   network segments may be configured in such a way that a single
   forwarding treatment may satisfy the traffic characteristics and
   performance requirements of two or more service classes. In these
   cases, it may be desirable to aggregate two or more Diffserv service
   classes into a single forwarding treatment. This document provides
   guidelines for the aggregation of Diffserv service classes into
   forwarding treatments.

Table of Contents

1. Introduction

   In the core of a high capacity network, it is common for the network
   to be engineered in such a way that a major link, switch, or router
   can fail, and the result will be a routed network that still meets
   ambient Service Level Agreements (SLAs). The implications are that
   there is sufficient capacity on any given link such that all SLAs
   sold can be simultaneously supported at their respective maximum
   rates, and that this remains true after re-routing (either IP re-
   routing or Multiprotocol Label Switching (MPLS) protection-mode
   switching) has occurred.

   Over-provisioning is generally considered to meet the requirements of
   all traffic without further quality of service (QoS) treatment, and
   in the general case, that is true in high-capacity backbones.
   However, as the process of network convergence continues, and with
   the increasing speed of the access networks, certain services may
   still have issues. Delay, jitter, and occasional loss are perfectly
   acceptable for elastic applications. However, sub-second surges that
   occur in the best-designed of networks [12] affect real-time
   applications. Moreover, denial of service (DoS) loads, worms, and
   network disruptions such as that of 11 September 2001 affect routing
   [13]. Our objective is to prevent disruption to routing (which in
   turn affects all services) and to protect real-time jitter-sensitive
   services, while minimizing loss and delay of sensitive elastic
   traffic.

   RFC 4594 [3] defines a set of basic Diffserv classes from the points
   of view of the application requiring specific end-to-end behaviors
   from the network. The service classes are differentiated based on
   the application payload's tolerance to packet loss, delay, and delay
   variation (jitter). Different degrees of these criteria form the
   foundation for supporting the needs of real-time and elastic traffic.
   RFC 4594 [3] also provides recommendations for the treatment method
   of these service classes. But, at some network segments of the end-
   to-end path, the number of levels of network treatment
   differentiation may be less than the number of service classes that
   the network segment needs to support. In such a situation, that
   network segment may use the same treatment to support more than one
   service class. In this document, we provide guidelines on how
   multiple service classes may be aggregated into a forwarding
   treatment aggregate. This entails having the IP traffic belonging to
   service classes, expressed using the DSCP (Differentiated Services
   Code Point), as described by RFC 4594 [3]. Note that in a given
   domain, we may recommend that the supported service classes be
   aggregated into forwarding treatment aggregates; however, this does
   not mean all service classes need to be supported, and hence not all
   forwarding treatment aggregates need to be supported. A domain may

support a fewer or greater number of forwarding treatment aggregates
than recommended by this document. Which service classes and which
forwarding treatment aggregates are supported by a domain is up to
the domain administration and may be influenced by business reasons
or other reasons (e.g., operational considerations).

In this document, we've provided:

o definitions for terminology we use in this document,

o requirements for performing this aggregation,

o an example of performing the aggregation when four treatment
   aggregates are used, and

o an example (in the appendix) of performing this aggregation over
   MPLS using E-LSP, EXP Inferred PHB Scheduling Class (PSC) Label
   Switched Path (LSP).

The treatment aggregate recommendations are designed to aggregate the
service classes [3] in such a manner as to protect real-time traffic
and routing, on the assumption that real-time sessions are protected
from each other by admission at the edge. The recommendation given
is one possible way of performing the aggregation; there may be other
ways of aggregation, for example, into fewer treatment aggregates or
more treatment aggregates.

In the appendix, an example of aggregation over MPLS networks using
E-LSP to realize the treatment aggregates is provided. Note that the
MPLS E-LSP is just an example; this document does not exclude the use
of other methods. This example only considers aggregation of IP
traffic into E-LSP. The use of E-LSP by non-IP traffic is not
discussed.

1.1. Requirements Notation

   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 [1].

2. Terminology

   This document assumes the reader is familiar with the terms used in
   differentiated services. This document provides the definitions for
   new terms introduced by this document and references information
   defined in RFCs for existing terms not commonly used in
   differentiated services.

For new terms introduced by this document, we provide the definition here:

o Treatment Aggregate. This term is defined as the aggregate of Diffserv service classes [3]. A treatment aggregate is concerned only with the forwarding treatment of the aggregated traffic, which may be marked with multiple DSCPs. A treatment aggregate differs from Behavior Aggregate [2] and Traffic Aggregate [14], each of which indicate the aggregated traffic having a single Diffserv codepoint and utilizing a single Per Hop Behavior (PHB).

For terms from existing RFCs, we provide the reference to the appropriate section of the relevant RFC that contain the definition:

o Real-Time and Elastic Applications and their traffic. Section 3.1 of RFC 1633 [4].

o Diffserv Service Class. Section 1.3 of RFC 4594 [3].

o MPLS E-LSP, EXP Inferred PHB Scheduling Class (PSC) Label Switched Path (LSP). Section 1.2 of RFC 3270 [6].

o MPLS L-LSP, Label Only Inferred PHB Scheduling Class (PSC) Label Switched Path (LSP). Section 1.3 of RFC 3270 [6].

3. Overview of Service Class Aggregation

In Diffserv domains where less fine-grained traffic treatment differentiation is provided, aggregation of the different service classes [3] may be required.

These aggregations have the following requirements:

1. The end-to-end network performance characteristic required by the application MUST be supported. This performance characteristic is represented by the use of Diffserv service classes [3].

2. The treatment aggregate MUST meet the strictest requirements of its member service classes.

3. The treatment aggregate SHOULD only contain member service classes with similar traffic characteristic and performance requirements.

4. The notion of the individual end-to-end service classes MUST NOT be destroyed when aggregation is performed. Each domain along the end-to-end path may perform aggregation differently, based on the original end-to-end service classes. We recommend an easy

way to accomplish this by not altering the DSCP used to indicate
the end-to-end service class. But some administrative domains
may require the use of their own marking; when this is needed,
the original end-to-end service class indication must be restored
upon exiting such administrative domains. One possible way of
achieving this is with the use of tunnels to encapsulate the end-
to-end traffic.

5. Each treatment aggregate has limited resources; hence, traffic
    conditioning and/or admission control SHOULD be performed for
    each service class aggregated into the treatment aggregate.
    Additional admission control and policing may be used on the sum
    of all traffic aggregated into the treatment aggregate.

In addition to the above requirements, we have the following
suggestions:

1. The treatment aggregate and assigned resources may consider
    historical traffic patterns and the variability of these
    patterns. For example, a point-point service (e.g., pseudowire)
    may have a very predictable pattern, while a multipoint service
    (e.g., VPLS, Virtual Private LAN Service) may have a much less
    predictable pattern.

2. In addition to Diffserv, other controls are available to
    influence the traffic level offered to a particular traffic
    aggregate. These include adjustment of routing metrics, and
    usage of MPLS-based traffic engineering techniques.

This document only describes the aggregation of IP traffic based on
the use of Diffserv service classes [3].

4. Service Classes to Treatment Aggregate Mapping

The service class and DSCP selection in RFC 4594 [3] has been defined
to allow, in many instances, mapping of two or possibly more service
classes into a single forwarding treatment aggregate. Notice that
there is a relationship/trade-off between link speed, queue depth,
delay, and jitter. The degree of aggregation and hence the number of
treatment aggregates will depend on the aggregation's impacts on
loss, delay, and jitter. This depends on whether the speed of the
links and scheduler behavior, being used to implement the
aggregation, can minimize the effects of mixing traffic with
different packet sizes and transmit rates on queue depth. A general
rule-of-thumb is that higher link speeds allow for more aggregation/
smaller number of treatment aggregates, assuming link utilization is
within the engineered level.

4.1. Mapping Service Classes into Four Treatment Aggregates

   This section provides an example of mapping all the service classes
   defined in RFC 4594 [3] into four treatment aggregates. The use of
   four treatment aggregates assumes that the resources allocated to
   each treatment aggregate are sufficient to honor the required
   behavior of each service class [3]. We use the performance
   requirement (tolerance to loss, delay, and jitter) from the
   application/end-user as a guide on how to map the service classes
   into treatment aggregates. We have also used section 3.1 of RFC 1633
   [4] to provide us with guidance on the definition of Real-Time and
   Elastic applications. An overview of the mapping between service
   classes and the four treatment aggregates is provided by Figure 1,
   with the mapping being based on performance requirements. In Figure
   1, the right side columns of "Service Class" and "Tolerance to Loss/
   Delay/Jitter" are from Figure 2 of RFC 4594 [3].

   It is recommended that certain service classes be mapped into
   specific treatment aggregates. But this does not mean that all the
   service classes recommended for that treatment aggregate need to be
   supported. Hence, for a given domain, a treatment aggregate may
   contain only a subset of the service classes recommended in this
   document, i.e., the service classes supported by that domain. A
   domain's treatment of non-supported service classes should be based
   on the domain's local policy. This local policy may be influenced by
   its agreement with its customers. Such treatment may use the Elastic
   Treatment Aggregate, dropping the packets, or some other
   arrangements.

   Our example of four treatment aggregates is based on the basic
   differences in performance requirement from the application/end-user
   perspective. A domain may choose to support more or fewer treatment
   aggregates than the four recommended. For example, a domain may
   support only three treatment aggregates and map any network control
   traffic into the Assured Elastic treatment aggregate. This is a
   choice the administrative domain has. Hence, this example of four
   treatment aggregates does not represent a minimum required set of
   treatment aggregates one must implement; nor does it represent the
   maximum set of treatment aggregates one can implement.

Treatment Aggregate	Tolerance to				Service Class	Tolerance to		
	Loss	Delay	Jitter			Loss	Delay	Jitter
Network Control	Low	Low	Yes		Network Control	Low	Low	Yes
Real-Time	Very Low	Very Low	Very Low		Telephony	VLow	VLow	VLow
					Signaling	Low	Low	Yes
					Multimedia Conferencing	Low - Medium	Very Low	Low
					Real-time Interactive	Low	Very Low	Low
					Broadcast Video	Very Low	Medium	Low
Assured Elastic	Low	Low - Medium	Yes		Multimedia Streaming	Low - Medium	Medium	Yes
					Low-Latency Data	Low	Low - Medium	Yes
					OAM	Low	Medium	Yes
					High-Throughput Data	Low	Medium - High	Yes
Elastic	Not Specified				Standard	Not Specified		
					Low-Priority Data	High	High	Yes

Figure 1: Treatment Aggregate and Service Class Performance
Requirements

As we are recommending to preserve the notion of the individual end-
to-end service classes, we also recommend that the original DSCP
field marking not be changed when treatment aggregates are used.
Instead, classifiers that select packets based on the contents of the
DSCP field should be used to direct packets from the member Diffserv
service classes into the queue that handles each of the treatment
aggregates, without remarking the DSCP field of the packets. This is

summarized in Figure 2, which shows the behavior each treatment
aggregate should have, and the DSCP field marking of the packets that
should be classified into each of the treatment aggregates.

```
 --------------------------------------------------------------
|Treatment  |Treatment  || DSCP                                |
|Aggregate  |Aggregate  ||                                     |
|           |Behavior   ||                                     |
|=========+=========++====================================|
| Network   | CS        || CS6                                 |
| Control   |(RFC 2474) ||                                     |
|=========+=========++====================================|
| Real-     | EF        || EF, CS5, AF41, AF42, AF43, CS4, CS3 |
| Time      |(RFC 3246) ||                                     |
|=========+=========++====================================|
| Assured   | AF        || CS2, AF31, AF21, AF11               |
| Elastic   |(RFC 2597) ||------------------------------------|
|           |           || AF32, AF22, AF12                    |
|           |           ||------------------------------------|
|           |           || AF33, AF23, AF13                    |
|=========+=========++====================================|
| Elastic   | Default   || Default, (CS0)                      |
|           |(RFC 2474) ||------------------------------------|
|           |           || CS1                                 |
 --------------------------------------------------------------
```

Figure 2: Treatment Aggregate Behavior

Notes for Figure 2: For Assured Elastic and Elastic Treatment
Aggregates, please see sections 4.1.3 and 4.1.4, respectively, for
details on additional priority within the treatment aggregate.

4.1.1. Network Control Treatment Aggregate

The Network Control Treatment Aggregate aggregates all service
classes that are functionally necessary for the survival of a network
during a DoS attack or other high-traffic load interval. The theory
is that whatever else is true, the network must protect itself. This
includes the traffic that RFC 4594 [3] characterizes as being
included in the Network Control service class.

Traffic in the Network Control Treatment Aggregate should be carried
in a common queue or class with a PHB as described in RFC 2474 [2],
section 4.2.2.2 for Class Selector (CS). This treatment aggregate
should have a lower probability of packet loss and bear a relatively
deep target mean queue depth (min-threshold if RED (Random Early
Detection) is being used).

Please notice this Network Control Treatment Aggregate is meant to be
used for the customer's network control traffic. The provider may
choose to treat its own network control traffic differently, perhaps
in its own service class that is not aggregated with the customer's
network control traffic.

4.1.2. Real-Time Treatment Aggregate

The Real-Time Treatment Aggregate aggregates all real-time
(inelastic) service classes. The theory is that real-time traffic is
admitted under some model and controlled by an SLA managed at the
edge of the network prior to aggregation. As such, there is a
predictable and enforceable upper bound on the traffic that can enter
such a queue, and to provide predictable variation in delay it must
be protected from bursts of elastic traffic. The predictability of
traffic level may be based upon admission control for a well-known
community of interest (e.g., a point-point service) and/or based upon
historical measurements.

This treatment aggregate may include the following service classes
from the Diffserv service classes [3], in addition to other locally
defined classes: Telephony, Signaling, Multimedia Conferencing, Real-
time Interactive, and Broadcast Video.

Traffic in each service class that is going to be aggregated into the
treatment aggregate should be conditioned prior to aggregation. It
is recommended that per-service-class admission control procedures be
used, followed by per-service-class policing so that any individual
service class does not generate more than what it is allowed.
Furthermore, additional admission control and policing may be used on
the sum of all traffic aggregated into this treatment aggregate.

Traffic in the Real-Time Treatment Aggregate should be carried in a
common queue or class with a PHB (Per Hop Behavior) as described in
RFC 3246 [9] and RFC 3247 [10].

4.1.3. Assured Elastic Treatment Aggregate

The Assured Elastic Treatment Aggregate aggregates all elastic
traffic that uses the Assured Forwarding model as described in RFC
2597 [8]. The premise of such a service is that an SLA that is
negotiated includes a "committed rate" and the ability to exceed that
rate (and perhaps a second "excess rate") in exchange for a higher
probability of loss using Active Queue Management (AQM) [7] or
Explicit Congestion Notification (ECN) marking [11] for the portion
of traffic deemed to be in excess.

This treatment aggregate may include the following service classes
from the Diffserv service classes [3], in addition to other locally
defined classes: Multimedia Streaming, Low Latency Data, OAM, and
High-Throughput Data.

The DSCP values belonging to the Assured Forwarding (AF) PHB group
and class selector of the original service classes remain an
important consideration and should be preserved during aggregation.
This treatment aggregate should maintain the AF PHB group marking of
the original packet. For example, AF3x marked packets should remain
AF3x marked within this treatment aggregate. In addition, the class
selector DSCP value should not be changed. Traffic bearing these
DSCPs is carried in a common queue or class with a PHB as described
in RFC 2597 [8]. In effect, appropriate target rate thresholds have
been applied at the edge, dividing traffic into AFn1 (committed, for
any value of n), AFn2, and AFn3 (excess). The service should be
engineered so that AFn1 and CS2 marked packet flows have sufficient
bandwidth in the network to provide high assurance of delivery.
Since the traffic is elastic and responds dynamically to packet loss,
Active Queue Management [7] should be used primarily to reduce the
forwarding rate to the minimum assured rate at congestion points.
The probability of loss of AFn1 and CS2 traffic must not exceed the
probability of loss of AFn2 traffic, which in turn must not exceed
the probability of loss of AFn3 traffic.

If RED [7] is used as an AQM algorithm, the min-threshold specifies a
target queue depth for each of AFn1+CS2, AFn2, and AFn3, and the max-
threshold specifies the queue depth above which all traffic with such
a DSCP is dropped or ECN marked. Thus, in this treatment aggregate,
the following inequalities SHOULD hold in queue configurations:

o min-threshold AFn3 < max-threshold AFn3

o max-threshold AFn3 <= min-threshold AFn2

o min-threshold AFn2 < max-threshold AFn2

o max-threshold AFn2 <= min-threshold AFn1+CS2

o min-threshold AFn1+CS2 < max-threshold AFn1+CS2

o max-threshold AFn1+CS2 <= memory assigned to the queue

Note: This configuration tends to drop AFn3 traffic before AFn2, and
AFn2 before AFn1 and CS2. Many other AQM algorithms exist and are
used; they should be configured to achieve a similar result.

4.1.4. Elastic Treatment Aggregate

   The Elastic Treatment Aggregate aggregates all remaining elastic
   traffic. The premise of such a service is that there is no intrinsic
   SLA differentiation of traffic, but that AQM [7] or ECN flagging [11]
   is appropriate for such traffic.

   This treatment aggregate may include the following service classes
   from the Diffserv service classes [3], in addition to other locally
   defined classes: Standard and Low-Priority Data.

   Treatment aggregates should be well specified, each indicating the
   service classes it will handle. But in cases where unspecified or
   unknown service classes are encountered, they may be dropped or be
   treated using the Elastic Treatment Aggregate. The choice of how to
   treat unspecified service classes should be well defined, based on
   some agreements.

   Traffic in the Elastic Treatment Aggregate should be carried in a
   common queue or class with a PHB as described in RFC 2474 [2],
   section 4.1, "A Default PHB". The AQM thresholds for Elastic traffic
   MAY be separately set, so that Low Priority Data traffic is dropped
   before Standard traffic, but this is not a requirement.

5. Treatment Aggregates and Inter-Provider Relationships

   When treatment aggregates are used at provider boundaries, we
   recommend that the inter-provider relationship be based on Diffserv
   service classes [3]. This allows the admission control into each
   treatment aggregate of a provider domain to be based on the admission
   control of traffic into the supported service classes, as indicated
   by the discussion in section 4 of this document.

   If the inter-provider relationship needs to be based on treatment
   aggregates specified by this document, then the exact treatment
   aggregate content and representation must be agreed to by the peering
   providers.

   Some additional work on inter-provider relationships is provided by
   inter-provider QoS [15], where details on supporting real-time
   services between service providers are discussed. Some related work
   in ITU-T provided by Appendix VI of Y.1541 [16] may also help with
   inter-provider relationships, especially with international
   providers.

6. Security Considerations

   This document discusses the policy of using Differentiated Services
   and its service classes. If implemented as described, it should
   require that the network do nothing that the network has not already
   allowed. If that is the case, no new security issues should arise
   from the use of such a policy.

   As this document is based on RFC 4594 [3], the Security Consideration
   discussion of no new security issues indicated by RFC 4594 [3] also
   applies to treatment aggregates of this document.

7. Acknowledgements

   This document has benefited from discussions with numerous people,
   especially Shane Amante, Brian Carpenter, and Dave McDysan. It has
   also benefited from detailed reviews by David Black, Marvin Krym,
   Bruce Davie, Fil Dickinson, and Julie Ann Connary.

8. References

8.1. Normative References

   [1]   Bradner, S., "Key words for use in RFCs to Indicate Requirement
          Levels", BCP 14, RFC 2119, March 1997.

   [2]   Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of
          the Differentiated Services Field (DS Field) in the IPv4 and
          IPv6 Headers", RFC 2474, December 1998.

   [3]   Babiarz, J., Chan, K., and F. Baker, "Configuration Guidelines
          for DiffServ Service Classes", RFC 4594, August 2006.

   [4]   Braden, B., Clark, D., and S. Shenker, "Integrated Services in
          the Internet Architecture: an Overview", RFC 1633, June 1994.

   [5]   Black, D., "Differentiated Services and Tunnels", RFC 2983,
          October 2000.

   [6]   Le Faucheur, F., Wu, L., Davie, B., Davari, S., Vaananen, P.,
          Krishnan, R., Cheval, P., and J. Heinanen, "Multi-Protocol
          Label Switching (MPLS) Support of Differentiated Services",
          RFC 3270, May 2002.

   [7]   Braden, B., Clark, D., Crowcroft, J., Davie, B., Deering, S.,
         Estrin, D., Floyd, S., Jacobson, V., Minshall, G., Partridge,
         C., Peterson, L., Ramakrishnan, K., Shenker, S., Wroclawski,
         J., and L. Zhang, "Recommendations on Queue Management and
         Congestion Avoidance in the Internet", RFC 2309, April 1998.

   [8]   Heinanen, J., Baker, F., Weiss, W., and J. Wroclawski, "Assured
         Forwarding PHB Group", RFC 2597, June 1999.

   [9]   Davie, B., Charny, A., Bennet, J., Benson, K., Le Boudec, J.,
         Courtney, W., Davari, S., Firoiu, V., and D. Stiliadis, "An
         Expedited Forwarding PHB (Per-Hop Behavior)", RFC 3246,
         March 2002.

   [10] Charny, A., Bennet, J., Benson, K., Boudec, J., Chiu, A.,
         Courtney, W., Davari, S., Firoiu, V., Kalmanek, C., and K.
         Ramakrishnan, "Supplemental Information for the New Definition
         of the EF PHB (Expedited Forwarding Per-Hop Behavior)",
         RFC 3247, March 2002.

   [11] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of
         Explicit Congestion Notification (ECN) to IP", RFC 3168,
         September 2001.

8.2. Informative References

   [12] Choi, B., Moon, S., Zhang, Z., Papagiannaki, K., and C. Diot,
         "Analysis of Point-To-Point Packet Delay in an Operational
         Network", INFOCOMM 2004, March 2004,
         <http://www.ieee-infocom.org/2004/Papers/37_4.PDF>.

   [13] Ogielski, A. and J. Cowie, "Internet Routing Behavior on 9/11",
         March 2002, <http://www.renesys.com/tech/presentations/pdf/
         renesys-030502-NRC-911.pdf>.

   [14] Nichols, K. and B. Carpenter, "Definition of Differentiated
         Services Per Domain Behaviors and Rules for their
         Specification", RFC 3086, April 2001.

   [15] MIT Communications Futures Program, "Inter-provider Quality of
         Service", November 2006, <
         http://cfp.mit.edu/resources/papers/Interprovider QoS
         MIT_CFP_WP_9_14_06.pdf>.

   [16] International Telecommunications Union, "Network Performance
         Objectives for IP-Based Services", Recommendation Y.1541,
         February 2006.

Appendix A. Using MPLS for Treatment Aggregates

   RFC 2983 on Diffserv and Tunnels [5] and RFC 3270 on MPLS Support of
   Diffserv [6] provide a very good background on this topic. This
   document provides an example of using the E-LSP, EXP Inferred PHB
   Scheduled Class (PSC) Label Switched Path (LSP), defined by MPLS
   Support of Diffserv [6] for realizing the Treatment Aggregates.

   When treatment aggregates are represented in MPLS using EXP Inferred
   PSC LSP, we recommend the following usage of the MPLS EXP field for
   treatment aggregates.

```
  -----------------------------------------------
 |Treatment  || MPLS  ||  DSCP   |    DSCP       |
 |Aggregate  || EXP   ||  name   |   value       |
 |=========++======+==========|=============|
 | Network   || 110   ||  CS6    |   110000      |
 | Control   ||       ||         |               |
 |=========++======+==========|=============|
 | Real-     || 100   ||   EF    |   101110      |
 | Time      ||       ||---------|-------------|
 |           ||       ||  CS5    |   101000      |
 |           ||       ||---------|-------------|
 |           ||       ||AF41,AF42|100010,100100|
 |           ||       ||  AF43   |   100110      |
 |           ||       ||---------|-------------|
 |           ||       ||  CS4    |   100000      |
 |           ||       ||---------|-------------|
 |           ||       ||  CS3    |   011000      |
 |=========++======+==========|=============|
 | Assured   || 010*  ||  CS2    |   010000      |
 | Elastic   ||       ||  AF31   |   011010      |
 |           ||       ||  AF21   |   010010      |
 |           ||       ||  AF11   |   001010      |
 |           ||------||---------|-------------|
 |           || 011*  ||  AF32   |   011100      |
 |           ||       ||  AF22   |   010100      |
 |           ||       ||  AF12   |   001100      |
 |           ||       ||  AF33   |   011110      |
 |           ||       ||  AF23   |   010110      |
 |           ||       ||  AF13   |   001110      |
 |=========++======+==========|=============|
 | Elastic   || 000*  || Default |   000000      |
 |           ||       || (CS0)   |               |
 |           ||------||---------|-------------|
 |           || 001*  ||  CS1    |   001000      |
  -----------------------------------------------
```

               Figure 3: Treatment Aggregate and MPLS EXP Field Usage

   * Note: For Assured Elastic (and Elastic) Treatment Aggregate, the
   usage of 010 or 011 (000 or 001) as EXP field value depends on the
   drop probability. Packets in the LSP with EXP field of 011 (001)
   have a higher probability of being dropped than packets with an
   EXP field of 010 (000).

The above table indicates the recommended usage of EXP fields for
treatment aggregates. Because many deployments of MPLS are on a per-
domain basis, each domain has total control of its EXP usage and each
domain may use a different EXP field allocation for the domain's
supported treatment aggregates.

A.1. Network Control Treatment Aggregate with E-LSP

   The usage of E-LSP for Network Control Treatment Aggregate needs to
   adhere to the recommendations indicated in section 4.1.1 of this
   document and section 3.2 of RFC 4594 [3]. Reinforcing these
   recommendations, there should be no drop precedence associated with
   the MPLS PSC used for Network Control Treatment Aggregate because
   dropping of Network Control Treatment Aggregate traffic should be
   prevented.

A.2. Real-Time Treatment Aggregate with E-LSP

   In addition to the recommendations provided in section 4.1.2 of this
   document and in member service classes' sections of RFC 4594 [3], we
   want to indicate that Real-Time Treatment Aggregate traffic should
   not be dropped, as some of the applications whose traffic is carried
   in the Real-Time Treatment Aggregate do not react well to dropped
   packets. As indicated in section 4.1.2 of this document, admission
   control should be performed on each service class contributing to the
   Real-Time Treatment Aggregate to prevent packet loss due to
   insufficient resources allocated to Real-Time Treatment Aggregate.
   Further, admission control and policing may also be applied on the
   sum of all traffic aggregated into this treatment aggregate.

A.3. Assured Elastic Treatment Aggregate with E-LSP

   EXP field markings of 010 and 011 are used for the Assured Elastic
   Treatment Aggregate. The two encodings are used to provide two
   levels of drop precedence indications, with 010 encoded traffic
   having a lower probability of being dropped than 011 encoded traffic.
   This provides for the mapping of CS2, AF31, AF21, and AF11 into EXP
   010; and AF32, AF22, AF12 and AF33, AF23, AF13 into EXP 011. If the
   domain chooses to support only one drop precedence for this treatment
   aggregate, we recommend the use of 010 for EXP field marking.

A.4. Elastic Treatment Aggregate with E-LSP

   EXP field markings of 000 and 001 are used for the Elastic Treatment
   Aggregate. The two encodings are used to provide two levels of drop
   precedence indications, with 000 encoded traffic having a lower
   probability of being dropped than 001 encoded traffic. This provides
   for the mapping of Default/CS0 into 000; and CS1 into 001. Notice

that with this mapping, during congestion, CS1-marked traffic may be
starved. If the domain chooses to support only one drop precedence
for this treatment aggregate, we recommend the use of 000 for EXP
field marking.

A.5. Treatment Aggregates and L-LSP

Because L-LSP (Label Only Inferred PSC LSP) supports a single PSC per
LSP, the support of each treatment aggregate is on a per-LSP basis.
This document does not further specify any additional recommendation
(beyond what has been indicated in section 4 of this document) for
treatment aggregate to L-LSP mapping, leaving this to each individual
MPLS domain administration.

Authors' Addresses

   Kwok Ho Chan
   Nortel
   600 Technology Park Drive
   Billerica, MA 01821
   US

   Phone: +1-978-288-8175
   Fax: +1-978-288-8700
   EMail: khchan@nortel.com

   Jozef Z. Babiarz
   Nortel
   3500 Carling Avenue
   Ottawa, Ont. K2H 8E9
   Canada

   Phone: +1-613-763-6098
   Fax: +1-613-768-2231
   EMail: babiarz@nortel.com

   Fred Baker
   Cisco Systems
   1121 Via Del Rey
   Santa Barbara, CA 93117
   US

   Phone: +1-408-526-4257
   Fax: +1-413-473-2403
   EMail: fred@cisco.com