

Networking Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 2, 2012

P. Jain
K. Singh
S. Venkataraman
Alcatel-Lucent, Inc.
March 1, 2012

RSVP-TE Extensions for ECMP Tie-Breaking of Inter-Domain TE LSPs
draft-mplste-intedomain-ecmp-tie-breaking-00

Abstract

This document specifies RSVP-TE extensions that will communicate to all the ABRs that they need to explicitly perform a tie-breaking process to select a particular path if path calculation results in multiple equal-cost paths across different TE-Domains. This will help is efficient utilization of end-to-end network resources for Inter-Domain TE LSPs

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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 September 2, 2012.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in [Section 4.e](#) of

Internet-Draft

Inter-Domain TE-LSP ECMP Tie Breaking

March 2012

the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	4
2.	Terminology	5
3.	Establishment and Inability of End-to-End Load Balancing of Inter-Domain Contiguous TE LSP	6
4.	RSVP-TE Extensions	8
5.	Signaling Procedures	10
6.	Management Considerations	11
7.	Security Considerations	12
8.	Acknowledgements	13
9.	IANA Considerations	14
9.1.	IANA Considerations for RSVP-TE	14
10.	References	15
10.1.	Normative References	15
10.2.	Informative References	15
	Authors' Addresses	16

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 [RFC2119](#) [[RFC2119](#)].

When used in lower case, these words convey their typical use in common language, and are not to be interpreted as described in [RFC2119](#) [[RFC2119](#)].

1. Introduction

If multiple equal-cost paths exists for a constraint TE-LSP, then we typically use tie-breaking process to select a particular path. This tie-breaking process is executed at the Head-End node where Constrained Shortest Path First (CSPF) computation is exercised.

In case of RSVP Inter-Domain TE-LSPs of type Contiguous LSP [RFC4726](#) [RFC4726] and [RFC5151](#) [RFC5151] each ABR is specified as loose hop and each ABR along the path whose next hop is specified as a loose hop triggers a path computation (also referred to as an ERO expansion), before forwarding the RSVP Path message downstream. Thus, each ABR is responsible for calculating path within its TE-Domain.

Every node that triggers path computation for its TE-Domain can have multiple equal-cost paths and has to chose one of them. Since there are multiple nodes that perform path computation (w.r.t. their respective TE-Domains), hence there is no common selection criteria for tie-breaking to be followed by each node performing ERO expansion.

This document specifies RSVP-TE extensions that will communicate to all the nodes doing ERO expansion that they need to explicitly perform common tie-breaking process to select a particular path if path calculation results in multiple equal-cost paths across its TE-Domain. By doing this we achieve uniform path selection criteria for Inter-Domain TE LSPs.

[2.](#) Terminology

Terminology used in this document:

CSPF: Constrained Shortest Path First.

TE LSP: Traffic Engineering Label Switched Path.

ERO: Explicit Route Object.

TE: Traffic Engineering.

IGP: Interior Gateway Protocol.

OSPF: Open Shortest Path First.

IS-IS: Intermediate System-to-Intermediate System.

LSR: Label Switching Router.

TE LSP head-end: head/source of the TE LSP.

ABR: Area Border Router.

Intra-domain TE LSP: A TE LSP whose path does not transit across areas.

Inter-domain TE LSP: A TE LSP whose path transits across at least two different IGP areas.

3. Establishment and Inability of End-to-End Load Balancing of Inter-Domain Contiguous TE LSP

The aim of this section is purely to summarize the mechanisms involved in the establishment of a loosely routed TE LSP, based on RSVP as specified in [RFC3209](#) [[RFC3209](#)], [RFC4726](#) [[RFC4726](#)] and [RFC5151](#) [[RFC5151](#)].

In the context of this document, a loosely routed Inter-Domain TE LSP is defined as one that does not contain a full, explicit route identifying each LSR along the path of the LSP at the time it is signaled by the ingress LSR. Such an LSP is signaled with an ERO that contains at least one loose hop that identifies ABR node. Each LSR along the path whose next hop is specified as a loose hop triggers a path computation (also referred to as an ERO expansion),

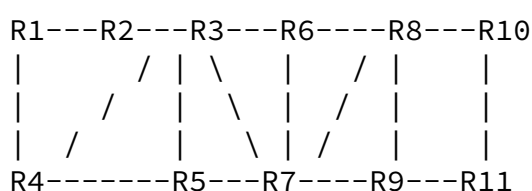
before forwarding the RSVP Path message downstream. The computed path may be either partial (up to the next loose hop) or complete (set of strict hops up to the TE LSP destination).

Note that although the examples in the rest of this document are provided in the context of MPLS Inter-Domain TE, the proposed mechanism applies equally to loosely routed paths within a single routing domain and across multiple Autonomous Systems. The examples below are provided with OSPF as the IGP, but the described set of mechanisms similarly apply to IS-IS.

An example of an explicit loosely routed TE LSP signaling follows.

Assumptions.

<-area 1-><-area 0-><-area 2->



- o R3, R5, R8, and R9 are ABRs.
- o The path of an Inter-Domain TE LSP T1 from R1 (head-end LSR) to R11 (tail-end LSR) is defined on R1 as the following loosely routed path: R1-R3(loose)-R8(loose)-R11(loose). R3, R8, and R11 are defined as loose hops.
- o Step 1: R1 determines that the next hop (R3) is a loose hop (not directly connected to R1) and then performs an ERO expansion operation to reach the next loose hops R3. The new ERO could

become: R2(S)-R3(S)-R8(L)-R11(L) or R4(S)-R3(S)-R8(L)-R11(L), where S is a strict hop (L=0) and L is a loose hop (L=1). Both the paths R1-R2-R3 and R1-R4-R3 are equal-cost paths and satisfies T1's set of constraints.

- o Step 2: The RSVP Path message is then forwarded by R1 following the path specified in the ERO object and reaches R3 with the following content: R8(L)-R11(L).

- o Step 3: R3 determines that the next hop (R8) is a loose hop (not directly connected to R3) and then performs an ERO expansion operation to reach the next loose hops R8. The new ERO could become: R6(S)-R8(S)-R11(L) or R7(S)-R8(S)-R11(L). Both paths R3-R6-R8 and R3-R7-R8 are equal-cost paths and satisfies T1's set of constraints.
- o Step 4: The same procedure is repeated by R8 to reach T1's destination (R11) and which may also lead to two equal-cost paths R8-R10-R11 and R8-R9-R11.

In the above example we see that there are two equal-cost paths that results from path computation done by ingress node (R1) and subsequent ABRs (R3 and R8) for their TE-Domain. In order of efficiently using the links as desired by network administrator only the ingress node can know about the tie-breaking process to be used as part of the TE-LSP configuration, which may lead to efficient use of the links in area-1 (or domain 1) only. However there is no way to communicate that subsequent ABRs (R3 and R8) also need to use common tie-breaking process so as to efficiently use the link in their subsequent domains.

This document defines mechanisms that would allow each ABR to know what tie-breaking process to use among equal-cost paths when doing path computation for their respective TE-domain.

There are following well known and understood techniques followed by various routing vendors used for tie-breaking in case multiple equal-cost paths exists.

- o Least-Fill.
- o Most-Fill.
- o Random.

Most known and deployed implementations employ a consistent definition of above three Path Selection tie-breaking schemes, and the specifications of each tie-breaking load balancing algorithms is outside the scope of this document. This document only lays down the mechanism for signaling what type of tie-breaking process is intended to be used for a given Inter-Domain TE LSP. Such that all the nodes along the Inter-Domain TE LSP which does path computations can factor such constraints and help in signaling end-to-end TE-LSP which is efficiently using links as desired by network administrator spanning across multiple TE-Domains.

In order to indicate nodes (e.g ABRs) that are participating ERO expansion for Inter-Domain Contiguous TE LSP to exercise either Least-Fill or Most-Fill type of tie-breaking procedure for equal-cost paths, this document defines new set of flag values in the Attribute Flags TLV, which are carried in LSP_ATTRIBUTES Object [RFC5420](#):.

- o Bit Number (to be assigned by IANA, recommended bit one) : Least-Fill Bit.
- o Bit Number (to be assigned by IANA, recommended bit two) : Most-Fill Bit.

Both the bits would together can be referred as ECMP tie-breaking bits. Both the proposed bits would have meaning only in Path message.

If Least-Fill Bit is set, then the node responsible for Path expansion MUST use the Least-Fill process for tie-breaking of ECMP paths.

If Most-Fill Bit is set, then the node responsible for Path expansion MUST use the Most-Fill process for tie-breaking of ECMP paths.

If none of the Bit (Least-Fill Bit or Most-Fill Bit) is set, then the node responsible for Path expansion can use the default process for

tie-breaking of ECMP paths. This could be Random or whatever is configured as default tie-breaking process on the node.

The rules of the processing of the Attribute Flags TLV are not changed.

5. Signaling Procedures

If it is desired to perform uniform way of tie-breaking process (with could be Least-Fill, Most-Fill or Random) across all domains for Inter-Domain TE-LSP, then head-end node is responsible for setting Least-Fill or Most-Fill Bit in the Attribute Flags TLV which is carried in LSP_ATTRIBUTES Object.

When a node receives a Path message which carries an LSP_ATTRIBUTES Object with either Least-Fill or Most-Fill Bit set in Attribute Flags TLV, then :-.

If the node is going to perform the ERO expansion (e.g ABR node), then it MUST use the appropriate method as what is been signaled (Least-Fill or Most-Fill) to perform tie-breaker of ECMP paths. If both Least-Fill and Most-Fill bits are unset, then node can use the default process for breaking the tie-breaker of ECMP paths. This could be Random or whatever is configured as default tie-breaking process on the node.

If the node does not need to perform the path expansion (e.g non-ABR node), it should do regular Path message processing as defined in [RFC2205](#) [[RFC2205](#)] and [RFC3209](#) [[RFC3209](#)].

If the node does not support Least-Fill or Most-Fill process of tie-breaking of ECMP paths, then it MUST send PathErr to reject the Path message with the defined error code 'Policy Control Failure'/'Inter- domain policy failure'.

Jain, et al.

Expires September 2, 2012

[Page 10]

Internet-Draft

Inter-Domain TE-LSP ECMP Tie Breaking

March 2012

[6.](#) Management Considerations

None

[7.](#) Security Considerations

The function described in this document does not create any new security issues for RSVP-TE protocol.

[8.](#) Acknowledgements

The editors gratefully acknowledge the useful inputs of Mustapha Aissaoui of Alcatel-Lucent, Inc

[9.](#) IANA Considerations

[9.1.](#) IANA Considerations for RSVP-TE

IANA will assign two Bits in Attribute Flags TLV, which is carried in an LSP_ATTRIBUTES Object [RFC5420](#) [[RFC5420](#)]. First bit would communicate Least-Fill, and the second bit would communicate Most-Fill type of tie-breaking procedure for equal-cost paths.

Following Bit values for Attribute Flags TLV are suggested:-

Bit	Tie-Breaking Method	Reference
1 (suggested)	Least-Fill	(this document)
2 (suggested)	Most-Fill	(this document)

[10.](#) References

[10.1.](#) Normative References

- [RFC4726] Farrel, A., Vasseur, J., and A. Ayyangar, "A Framework for Inter-Domain Multiprotocol Label Switching Traffic Engineering", [RFC 4726](#), November 2006.
- [RFC5151] Farrel, A., Ayyangar, A., and JP. Vasseur, "Inter-Domain MPLS and GMPLS Traffic Engineering -- Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Extensions", [RFC 5151](#), February 2008.
- [RFC5420] Farrel, A., Papadimitriou, D., Vasseur, JP., and A. Ayyangar, "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", [RFC 5420](#), February 2009.

[10.2.](#) Informative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC2205] Braden, B., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", [RFC 2205](#), September 1997.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", [RFC 3209](#), December 2001.
- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", [RFC 3471](#), January 2003.
- [RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", [RFC 3473](#), January 2003.
- [RFC3784] Smit, H. and T. Li, "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)", [RFC 3784](#), June 2004.

Authors' Addresses

Pradeep Jain
Alcatel-Lucent, Inc.
701 E Middlefield Rd
Mountain View, CA 94040
USA

Email: pradeep.jain@alcatel-lucent.com

Kanwar Singh
Alcatel-Lucent, Inc.
701 E Middlefield Rd
Mountain View, CA 94040
USA

Email: kanwar.singh@alcatel-lucent.com

Srikrishnan Venkataraman
Alcatel-Lucent, Inc.
701 E Middlefield Rd
Mountain View, CA 94040
USA

Email: Srikrishnan.Venkataraman@alcatel-lucent.com

