MPLS Working Group
Internet-Draft
Intended status: Informational
Expires: April 26, 2012

A. Atlas
J. Drake
D. Ward
Juniper Networks
S. Giacalone
Thomson Reuters
S. Previdi
C. Filsfils
Cisco Systems
October 24, 2011

Performance-based Path Selection for Explicitly Routed LSPs
draft-atlas-mpls-te-express-path-00

Abstract

In certain networks, it is critical to consider network performance criteria when selecting the path for an explicitly routed RSVP-TE LSP. Such performance criteria can include latency, jitter, and loss or other indications such as the conformance to link SLAs and non-RSVP TE traffic load. This specification uses IGP extension data (which is defined outside the scope of this document) to perform such path selections.

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 April 26, 2012.

Copyright Notice

Copyright (c) 2011 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 the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

1. Introduction

In certain networks, such as financial information networks, network performance information is becoming as critical to data path selection as other existing metrics. The ability to distribute network performance information in OSPF [I-D.giacalone-ospf-te-express-path] and in ISIS [I-D.previdi-isis-te-metric-extensions] is being defined (outside the scope of this document). This document describes how to use that information for path selection for explicitly routed LSPs signaled via RSVP-TE [RFC3209].

The path selection mechanisms described in this document apply to paths that are fully computed by the head-end of the LSP and then signaled in an ERO where every sub-object is strict. This allows the head-end to consider IGP-distributed performance data without requiring the ability to signal the performance constraints in an object of the RSVP Path message.

When considering performance-based data, it is obvious that there are additional contributors beyond just the links. Clearly end-to-end latency is a combination of router latency, queuing latency, physical link latency and other factors. However, if application traffic requires paths to be selected based upon latency constraints, the same traffic might be in an Expedited Forwarding Per-Hop-Behavior[RFC3246] with minimal queuing delay or another PHB with known maximal per-hop queuing delay. While traversing a router can cause delay, that can be included in the advertised link delay.

This document does not specify how a router determines what values to advertise by the IGP. However, the end-to-end performance that is computed for an LSP path SHOULD be built from the individual link data. Any end-to-end characterization used to determine an LSP's performance compliance should be fully reflected in the Traffic Engineering Database so that a CSPF calculation can also determine whether a path under consideration would be in compliance.

1.1. Basic Requirements

The following are the requirements that motivate this solution.

1. Select a TE tunnel's path based upon a combination of existing constraints as well as on link-latency, packet loss, jitter, link SLA conformance, and bandwidth consumed by non-RSVP-TE traffic.

2. Ability to define different end-to-end performance requirements for each TE tunnel regardless of common use of resources.

3. Ability to periodically verify that a TE tunnel's current LSP complies with its configured end-to-end perforance requirements.

4. Ability to move tunnels, using make-before-break, based upon computed end-to-end performance complying with configuration

5. Ability to move tunnels away from any link that is violating an underlying SLA

6. Ability to optionally avoid setting up tunnels using any link that is violating an SLA, regardless of whether end-to-end performance would still meet requirements.

7. Ability to revert back to the best path after a configurable period.

2. Using Performance Data Constraints

2.1. End-to-End Constraints

The per-link performance data available in the IGP [I-D.giacalone-ospf-te-express-path] [I-D.previdi-isis-te-metric-extensions] includes: unidirectional link delay, unidirectional delay variation, and link loss. Each (or all) of these parameters can be used to create the path-level link-based parameter.

While it has been possible to compute a CSPF where the link latency values are used instead of TE metrics, this results in ignoring the TE metrics and causing LSPs to prefer the lowest-latency paths. Instead of this approach to minimize path latency, an end-to-end latency bound merely requires that the path computed be no more than that bound without being the minimum. This bound can be used as a constraint in CSPF to prevent exploring links that would create a path over the end-to-end latency bound.

This is illustrated as follows. Let the LSP have an end-to-end latency bound of 20ms. Assume that the path to node X has been minimized and its latency is 12ms. When X's links are to be explored, the link X<->Y has a link latency of 5ms and the link X<->Z has a link latency of 9ms. The path via X to Y along link X<->Y would have a path latency of 12ms + 5ms = 17ms < 20ms; therefore, the link X<->Y can be explored. In contrast, reaching Z via link X<->Z would result in a path latency of 12ms + 9ms = 21ms > 20ms; therefore the link X<->Z would not be explored in the CSPF.

An end-to-end bound on delay variation can be used similarly as a constraint in the CSPF on what links to explore where the path's delay variation is the sum of the used links' delay variations.

For link loss, the path loss is not the sum of the used links' losses. Instead, the path loss percentage is (100 - loss_L1)*(100 - loss_L2)*...*(100 - loss_Ln), where the links along the path are L1 to Ln. The end-to-end link loss bound, computed in this fashion, can also be used as a constraint in the CSPF on what links to explore.

2.2. Link Constraints

In addition to selecting paths that conform to a bound on performance data, it is also useful to avoid using links that do not meet a necessary constraint. Naturally, if such a parameter were a known fixed value, then resource attribute flags could be used to express this behavior. However, when the parameter associated with a link may vary dynamically, there is not currently a configuration-time mechanism to enforce such behavior. An example of this is described in Section 2.3, where links may move in and out of SLA-conformance with regards to latency, delay variation, and link loss.

When doing path selection for TE tunnels, it has not been possible to know how much actual bandwidth is available that inludes the bandwidth used by non-RSVP-TE traffic. In [I-D.giacalone-ospf-te-express-path] [I-D.previdi-isis-te-metric-extensions], the Unidirectional Available Bandwidth is advertised as is the Residual Bandwidth. When computing the path for a TE tunnel, only links with at least a configurable amount of Unidirectional Available Bandwidth might be permitted.

Similarly, only links whose loss is under a configurable value might be acceptable. For these constraints, each link can be tested against the constraint and only explored in the CSPF if the link passes. In essence, a link that fails the constraint test is treated as if it contained a resource attribute in the exclude-any filter.

2.3. Links out of SLA

Link conformance to an SLA can change as a result of rerouting at lower layers. This could be due to optical regrooming or simply rerouting of a FA-LSP. When this occurs, there are three questions to be asked:

a. Should the link be trusted and used for the setup of new LSPs?

b. Should LSPs using this link be immediately verified for continued compliance to their end-to-end constraints?

c. Should LSPs using this link automatically be moved to a secondary path?

2.3.1. Use of Anomalous Links for New Paths

If the answer to (a) is no for latency SLAs, then any link which has the Anomalous bit set in the Unidirectional Link Delay sub-TLV[I-D.giacalone-ospf-te-express-path] [I-D.previdi-isis-te-metric-extensions] should be removed from the topology before a CSPF calculation is used to compute a new path. In essence, the link should be treated exactly as if it fails the exclude-any resource attributes filter.[RFC3209].

Similarly, if the answer to (a) is no for link loss SLAs, then any link which has the Anomalous bit set in the Link Los sub-TLV should be treated as if it fails the exclude-any resource attributes filter. If the answer to (a) is no for jitter SLAs, then any link that has the Anomalous bit set in the Unidirectional Delay Variation sub-TLV[I-D.previdi-isis-te-metric-extensions] should be treated as if it fails the exclude-any resource attributes filter.

2.3.2. Links entering the Anomalous State

When a link enters the Anomalous state with respect to a parameter, this is an indication that LSPs using that link might also no longer be in compliance with their performance bounds. It can also be considered an indication that something is changing that link and so it might no longer be trustworthy to carry performance-critical traffic. Naturally, which performance criteria are important for a particular LSP is dependent upon the LSP's configuration and thus the SLA compliance of a link is indicated per performance criterion.

At the ingress of a TE tunnel, a TE tunnel may be configured to be sensitive to the Anomalous state of links in reference to latency, delay variation, and/or loss. Additionally, such a TE tunnel may be configured to either verify continued compliance, to switch

immediately to a standby LSP, or to move to a different path.

When a sub-TLV is received with the Anomalous bit set when previously it was clear, the list of interested TE tunnels must be scanned. Each such TE tunnel should either have its continued compliance verified, be switched to a hot standby, or do a make-before-break to a secondary path.

2.3.3. Links leaving the Anomalous State

When a link leaves the Anomalous state with respect to a parameter, this can serve as an indication that those TE tunnels, whose LSPs were changed when the link entered the Anomalous state, may want to reoptimize to a better path.

3. IANA Considerations

This document includes no request to IANA.

4. Security Considerations

This document is not currently believed to introduce new security concerns.

5. References

5.1. Normative References

[I-D.giacalone-ospf-te-express-path]
Giacalone, S., Ward, D., Drake, J., Atlas, A., and S. Previdi, "OSPF Traffic Engineering (TE) Express Path", draft-giacalone-ospf-te-express-path-02 (work in progress), September 2011.

[I-D.previdi-isis-te-metric-extensions]
Previdi, S., Giacalone, S., Ward, D., Drake, J., Atlas, A., and C. Filsfils, "IS-IS Traffic Engineering (TE) Metric Extensions", draft-previdi-isis-te-metric-extensions-00 (work in progress), October 2011.

[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.

5.2. Informative References

[RFC3246] 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.

[RFC5420] Farrel, A., Papadimitriou, D., Vasseur, JP., and A. Ayyangarps, "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", RFC 5420, February 2009.

Authors' Addresses

Alia Atlas
Juniper Networks
10 Technology Park Drive
Westford, MA 01886
USA

Email: akatlas@juniper.net

John Drake
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
USA

Email: jdrake@juniper.net

Dave Ward
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
USA

Email: dward@juniper.net

Spencer Giacalone
Thomson Reuters
195 Broadway
New York, NY 10007
USA

Email: Spencer.giacalone@thomsonreuters.com

Stefano Previdi
Cisco Systems
Via Del Serafico 200
Rome 00142
Italy

Email: sprevidi@cisco.com

Clarence Filsfils
Cisco Systems
Brussels
Belgium

Email: cfilsfil@cisco.com