Network Working Group Dave Katz
Internet Draft Juniper Networks, Inc.
Expiration Date: April 2002 Derek Yeung
Procket Networks, Inc.
Kireeti Kompella
Juniper Networks, Inc.

Traffic Engineering Extensions to OSPF

draft-katz-yeung-ospf-traffic-06.txt

Status

Abstract

This document describes extensions to the OSPF protocol to support Traffic Engineering, using opaque LSAs.

1. Introduction

This document specifies a method of adding traffic engineering capabilities to OSPF [1]. The architecture of traffic engineering is described in [2]. The semantic content of the extensions is essentially identical to the corresponding extensions to IS-IS [3]. It is expected that the traffic engineering extensions to OSPF will continue to mirror those in IS-IS.

The extensions provide a way of describing the traffic engineering topology (including bandwidth and administrative constraints). This topology does not necessarily match the regular routed topology, though this proposal depends on Network LSAs to describe multiaccess links.

1.1. Applicability

Many of the extensions specified in this document are in response to the requirements stated in [2], and thus are referred to as "traffic engineering extensions", and are also commonly associated with MPLS Traffic Engineering. A more accurate (albeit bland) designation is "extended link attributes", as what is proposed is simply to add more attributes to links in OSPF advertisements.

The information made available by these extensions can be used to build an extended link state database just as router LSAs are used to build a "regular" link state database; the difference is that the extended link state database (referred to below as the traffic engineering database) has additional link attributes. Uses of the traffic engineering database include:

- o monitoring the extended link attributes;
- o local constraint-based source routing; and
- o global traffic engineering.

For example, an OSPF-speaking device can participate in an OSPF area, build a traffic engineering database, and thereby report on the reservation state of links in that area.

In "local constraint-based source routing", a router R can compute a path from a source node A to a destination node B; typically, A is R itself, and B is specified by a "router address" (see below). This path may be subject to various constraints on the attributes of the links and nodes that the path traverses, e.g., use green links that have unreserved bandwidth of at least 10Mbps. This path could then be used to carry some subset of the traffic from A to B, forming a simple but effective means of traffic engineering. How the subset of traffic is determined, and how the path is instantiated is beyond the scope of this document; suffice it to say that one means of defining the subset of traffic is "those packets whose IP destinations were learned from B", and one means of instantiating paths is using MPLS tunnels. As an aside, note that constraint-based routing can be NP-hard, or even unsolvable, depending on the nature of the attributes and constraints and thus many implementations will use heuristics. Consequently, we don't attempt to sketch an algorithm here.

Finally, for "global traffic engineering", a device can build a traffic engineering database, input a traffic matrix and an optimization function, crunch on the information, and thus compute optimal or near-optimal routing for the entire network. The device can subsequently monitor the traffic engineering topology and react to changes by recomputing the optimal routes.

1.2. Limitations

The extensions specified in this document capture the reservation state of point-to-point links. The reservation state of multiaccess links is not accurately reflected, except in the special case that there are only two devices in the multiaccess subnetwork.

This document also does not support unnumbered links. This deficiency is addressed in [4]; see also [5] and [6].

2. LSA Format

2.1. LSA type

This extension makes use of the Opaque LSA [7].

Three types of Opaque LSAs exist, each of which has different flooding scope. This proposal uses only Type 10 LSAs, which have area flooding scope.

One new LSA is defined, the Traffic Engineering LSA. This LSA describes routers, point-to-point links, and connections to multiaccess networks (similar to a Router LSA). For traffic engineering purposes, the existing Network LSA suffices for describing multiaccess links, so no additional LSA is defined for this purpose.

2.2. LSA ID

The LSA ID of an Opaque LSA is defined as having eight bits of type and 24 bits of type-specific data. The Traffic Engineering LSA uses type 1. The remaining 24 bits are broken up into eight bits of reserved space (which must be zero) and sixteen bits of instance:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       1       |    Reserved   |           Instance            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The Instance field is an arbitrary value used to maintain multiple Traffic Engineering LSAs. A maximum of 65536 Traffic Engineering LSAs may be sourced by a single system. The LSA ID has no topological significance.

2.3. LSA Format Overview

2.3.1. LSA Header

The Traffic Engineering LSA starts with the standard LSA header:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            LS age             |    Options    |      10       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       1       |    Reserved   |           Instance            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Advertising Router                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     LS sequence number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         LS checksum           |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

2.3.2. TLV Header

The LSA payload consists of one or more nested Type/Length/Value (TLV) triplets for extensibility. The format of each TLV is:

```
      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            |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                            Value...                           |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The Length field defines the length of the value portion in octets (thus a TLV with no value portion would have a length of zero). The TLV is padded to four-octet alignment; padding is not included in the length field (so a three octet value would have a length of three, but the total size of the TLV would be eight octets). Nested TLVs are also 32-bit aligned. Unrecognized types are ignored. All types between 32768 and 65535 are reserved for vendor-specific extensions. All other undefined type codes are reserved for future assignment by IANA.

2.4. LSA payload details

An LSA contains one top-level TLV.

There are two top-level TLVs defined:

1 - Router Address
2 - Link

2.4.1. Router Address TLV

The Router Address TLV specifies a stable IP address of the advertising router that is always reachable if there is any connectivity to it. This is typically implemented as a "loopback address"; the key attribute is that the address does not become unusable if an interface is down. In other protocols this is known as the "router ID," but for obvious reasons this nomenclature is avoided here.

If IS-IS is also active in the domain, this address can also be used to compute the mapping between the OSPF and IS-IS topologies. For example, suppose a router R is advertising both IS-IS and OSPF Traffic Engineering LSAs, and suppose further that some router S is building a single Traffic Engineering Database (TED) based on both IS-IS and OSPF TE information. R may then appear as two separate nodes in S's TED; however, if both the IS-IS and OSPF LSAs generated by R contain the same Router Address, then S can determine that the IS-IS TE LSA and the OSPF TE LSA from R are indeed from a single router.

The router address TLV is type 1, and has a length of 4, and the value is the four octet IP address. It must appear in exactly one Traffic Engineering LSA originated by a router.

2.4.2. Link TLV

The Link TLV describes a single link. It is constructed of a set of sub-TLVs. There are no ordering requirements for the sub-TLVs.

Only one Link TLV shall be carried in each LSA, allowing for fine granularity changes in topology.

The Link TLV is type 2, and the length is variable.

The following sub-TLVs are defined:

1 - Link type (1 octet)
2 - Link ID (4 octets)
3 - Local interface IP address (4 octets)
4 - Remote interface IP address (4 octets)
5 - Traffic engineering metric (4 octets)
6 - Maximum bandwidth (4 octets)
7 - Maximum reservable bandwidth (4 octets)
8 - Unreserved bandwidth (32 octets)
9 - Resource class/color (4 octets)

32768-32772 - Reserved for Cisco-specific extensions

The Link Type and Link ID sub-TLVs are mandatory, i.e., must appear exactly once. All other sub-TLVs defined here may occur at most once. These restrictions need not apply to future sub-TLVs. Unrecognized sub-TLVs are ignored.

2.5. Sub-TLV Details

2.5.1. Link Type

The Link Type sub-TLV defines the type of the link:

1 - Point-to-point
2 - Multiaccess

The Link Type sub-TLV is TLV type 1, and is one octet in length.

2.5.2. Link ID

The Link ID sub-TLV identifies the other end of the link. For point-to-point links, this is the Router ID of the neighbor. For multiaccess links, this is the interface address of the designated router. The Link ID is identical to the contents of the Link ID field in the Router LSA for these link types.

The Link ID sub-TLV is TLV type 2, and is four octets in length.

2.5.3. Local Interface IP Address

The Local Interface IP Address sub-TLV specifies the IP address(es) of the interface corresponding to this link. If there are multiple local addresses on the link, they are all listed in this sub-TLV.

The Local Interface IP Address sub-TLV is TLV type 3, and is 4N octets in length, where N is the number of local addresses.

2.5.4. Remote Interface IP Address

The Remote Interface IP Address sub-TLV specifies the IP address(es) of the neighbor's interface corresponding to this link. This and the local address are used to discern multiple parallel links between systems.

The Remote Interface IP Address sub-TLV is TLV type 4, and is 4N octets in length, where N is the number of neighbor addresses.

2.5.5. Traffic Engineering Metric

The Traffic Engineering Metric sub-TLV specifies the link metric for traffic engineering purposes. This metric may be different than the standard OSPF link metric.

The Traffic Engineering Metric sub-TLV is TLV type 5, and is four octets in length.

2.5.6. Maximum Bandwidth

The Maximum Bandwidth sub-TLV specifies the maximum bandwidth that can be used on this link in this direction (from the system originating the LSA to its neighbor), in IEEE floating point format. This is the true link capacity. The units are bytes per second.

The Maximum Bandwidth sub-TLV is TLV type 6, and is four octets in length.

2.5.7. Maximum Reservable Bandwidth

The Maximum Reservable Bandwidth sub-TLV specifies the maximum bandwidth that may be reserved on this link in this direction, in IEEE floating point format. Note that this may be greater than the maximum bandwidth (in which case the link may be oversubscribed). This SHOULD be user-configurable; the default value should be the Maximum Bandwidth. The units are bytes per second.

The Maximum Reservable Bandwidth sub-TLV is TLV type 7, and is four octets in length.

2.5.8. Unreserved Bandwidth

The Unreserved Bandwidth sub-TLV specifies the amount of bandwidth not yet reserved at each of the eight priority levels, in IEEE floating point format. The values correspond to the bandwidth that can be reserved with a setup priority of 0 through 7, arranged in increasing order with priority 0 occurring at the start of the sub-TLV, and priority 7 at the end of the sub-TLV. The initial values (before any bandwidth is reserved) are all set to the Maximum Reservable Bandwidth. Each value will be less than or equal to the Maximum Reservable Bandwidth. The units are bytes per second.

The Unreserved Bandwidth sub-TLV is TLV type 8, and is 32 octets in length.

2.5.9. Resource Class/Color

The Resource Class/Color sub-TLV specifies administrative group membership for this link, in terms of a bit mask. A link that is a member of multiple groups will have multiple bits set.

The Resource Class/Color sub-TLV is TLV type 9, and is four octets in length.

3. Elements of Procedure

Routers shall originate Traffic Engineering LSAs whenever the LSA contents change, and whenever otherwise required by OSPF (an LSA refresh, for example).

Upon receipt of a changed Traffic Engineering LSA or Network LSA (since these are used in traffic engineering calculations), the router should update its traffic engineering database. No SPF or other route calculations are necessary.

4. Compatibility Issues

There should be no interoperability issues with routers that do not implement these extensions, as the Opaque LSAs will be silently ignored.

The result of having routers that do not implement these extensions is that the traffic engineering topology will be missing pieces; however, if the topology is connected, TE paths can still be calculated and ought to work.

5. Security Considerations

This document raises no new security issues for OSPF.

6. References

[1] Moy, J., "OSPF Version 2", RFC 2328, April 1998.

[2] Awduche, D., et al, "Requirements for Traffic Engineering Over MPLS," RFC 2702, September 1999.

[3] Smit, H. and T. Li, "ISIS Extensions for Traffic Engineering," work in progress.

[4] Kompella, K., Rekhter, Y., et al, "OSPF Extensions in Support of Generalized MPLS," work in progress.

[5] Kompella, K., Rekhter, Y., and Kullberg, A., "Signalling Unnumbered Links in CR-LDP," work in progress.

[6] Kompella, K., and Rekhter, Y., "Signalling Unnumbered Links in RSVP-TE," work in progress.

[7] Coltun, R., "The OSPF Opaque LSA Option," RFC 2370, July 1998.

7. Authors' Addresses

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

Phone: +1 408 745 2000
Email: dkatz@juniper.net

Derek M. Yeung
Procket Networks, Inc.
1100 Cadillac Court
Milpitas, CA 95035 USA

Phone: +1 408 635-7900
Fax: +1 408 987-6166

Kireeti Kompella
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089 USA

Phone: +1 408 745 2000
Email: kireeti@juniper.net