

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 7, 2012

P. Pillay-Esnault
B. Pithawala
D. Yeung
Cisco Systems
March 6, 2012

Service Distribution using OSPF
draft-pillay-esnault-ospf-service-distribution-00

Abstract

The Open Shortest Path First (OSPF) protocol is used to carry data on behalf of other services using the Opaque Link State Advertisements. The protocol's flooding mechanism is well suited to cover the data propagation requirements of services such as Traffic Engineering. The current mechanism cannot scale for a large number of services nor satisfy some of their new set of requirements. This document describes a new mechanism in OSPF to support service and data distribution for a large number of services, computation of preferred service access points and a controlled service data exchange.

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 7, 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

Internet-Draft

Service Distribution using OSPF

March 2012

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.

Table of Contents

1.	Introduction	3
2.	Specification of Requirements	3
3.	Requirements for service data propagation	3
4.	Typical Scenario for Services Distribution Router	4
5.	OSPF Service Distribution Router	5
6.	Storage Of Service Data	6
7.	Mechanics of the OSPF Service Information Distribution Implementation	7
7.1.	Advertising and Signaling of SDR Capability	7
7.2.	Advertising the Service Distribution Router and its address mapping	8
7.3.	Advertising the Directory of Producers and Consumers . . .	9
7.4.	Service Routing Capable Router Operations	12
7.4.1.	Operation due to Producer changes	13
7.4.2.	Operation due to Consumer changes	13
8.	Calculation of Optimal Producer	14
9.	Service Router Data Operations	14
9.1.	Implementation of SDDA	15
9.2.	SDR Message Formats	15
9.3.	SDR Message Header Format	16
9.4.	OPEN Message Format	16
9.5.	CONFIRM Message Format	18
9.6.	UPDATE Message Format	18
9.7.	NOTIFICATION Message Format	20
9.7.1.	Message Header Error	21
9.7.2.	OPEN Message Error	21
9.7.3.	CONFIRM Message Error	22
9.7.4.	UPDATE Message Error	22
9.7.5.	Cease	22
10.	Security Considerations	22
11.	IANA Considerations	23
12.	Contributors	23
13.	Acknowledgments	23
14.	Normative References	23

Internet-Draft

Service Distribution using OSPF

March 2012

1. Introduction

Originally, routing protocols were designed to propagate routing related information only. With the advent of Traffic Engineering, the IGPs started to be used as a transport mechanism. Most of the applications using IGPs as transport are still very much limited and confined to routing applications with similar requirements.

Today, OSPF can carry data for applications using Opaque LSAs. These Opaque LSAs are an integral part of the OSPF database and will be flooded, synchronized and updated just as any other LSA. However, they do not contribute directly to any routes or trigger an SPF.

Opaque LSAs will need to be flooded across all the OSPF area or domain and neighbor adjacencies to FULL state will depend on successful exchange of these LSAs.

The Link State IGPs are limited on the size of payload information they can carry as it will be flooded and stored in every single router all across their areas or domain regardless whether it is of interest or not.

This document describes a new mechanism in OSPF to support service and data distribution for a large number of services, computation of preferred service access points and controlled distribution of service data.

We presuppose familiarity with the contents of [[RFC4970](#)], [[RFC5250](#)], [[RFC2328](#)] and [[RFC5340](#)] .

2. Specification of Requirements

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](#)].

[3.](#) Requirements for service data propagation

Services requirements differ from the traditional routing information dissemination model. The service data may be unrelated to routing or be of interest only to some routers. The new set of requirements for using OSPF as Service Distribution Router (SDR) is as follows

Scale to a large number of services

No assumption regarding size, format or nature of the data

No assumption regarding topology

Routing and service data separated and independent

Must support cases where minimal number of routers only may be upgraded

Must support dynamic events

Routers only store and process data of interest

Ability to compute the shortest path to a producer or consumer of a service per IGP metrics or service metrics.

There is no assumptions regarding producers and consumers of services, their location or uniqueness.

Secured data may reside only on some routers.

In addition the routing requirements for OSPF as Service data distribution are

Backward compatible with Open Standard OSPF

Minimal/No impact on routing convergence and performance

[4.](#) Typical Scenario for Services Distribution Router

A SDR is typically reachable by multiple consumers or producers of

data. The router itself may not be connected directly to any other router with Service Distribution Capability. The intermediate routers may have limited storage capability or cannot store the data for security reasons.

The SDR is aware of the topological information of the other service routers and can compute paths to the preferred Producer SDR (PSDR) or the Consumers SDR (CSDR) of a service. The SDR will implement tables of producers and consumers for services.

The SDR ensures that interested subscribers to a service are notified with the latest updates.

Producers or consumers can join or leave a service at any time using APIs. The SDR receiving the notification of "registration" or "de-registration" flood the change of state to all the known SDRs in its topology. Therefore, all SDR have the same view of the producers/consumers topology.

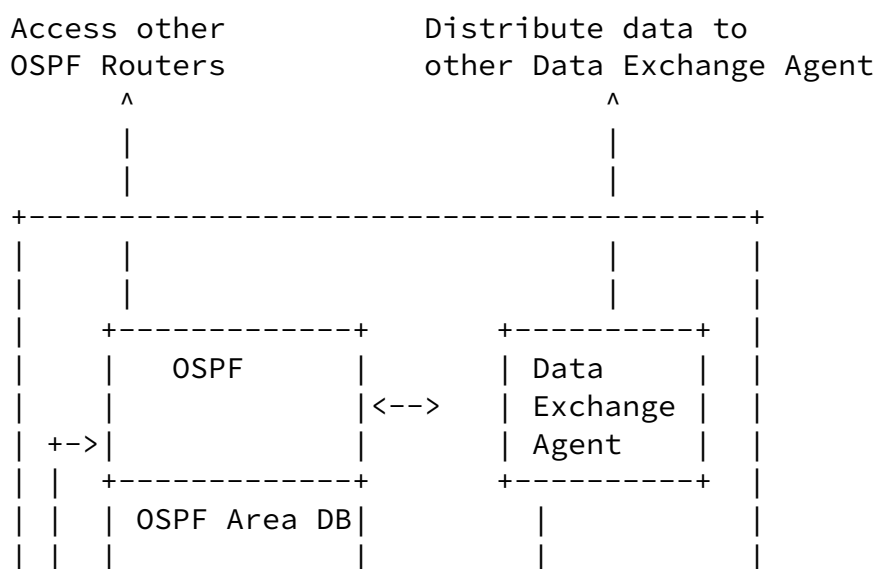
[5.](#) OSPF Service Distribution Router

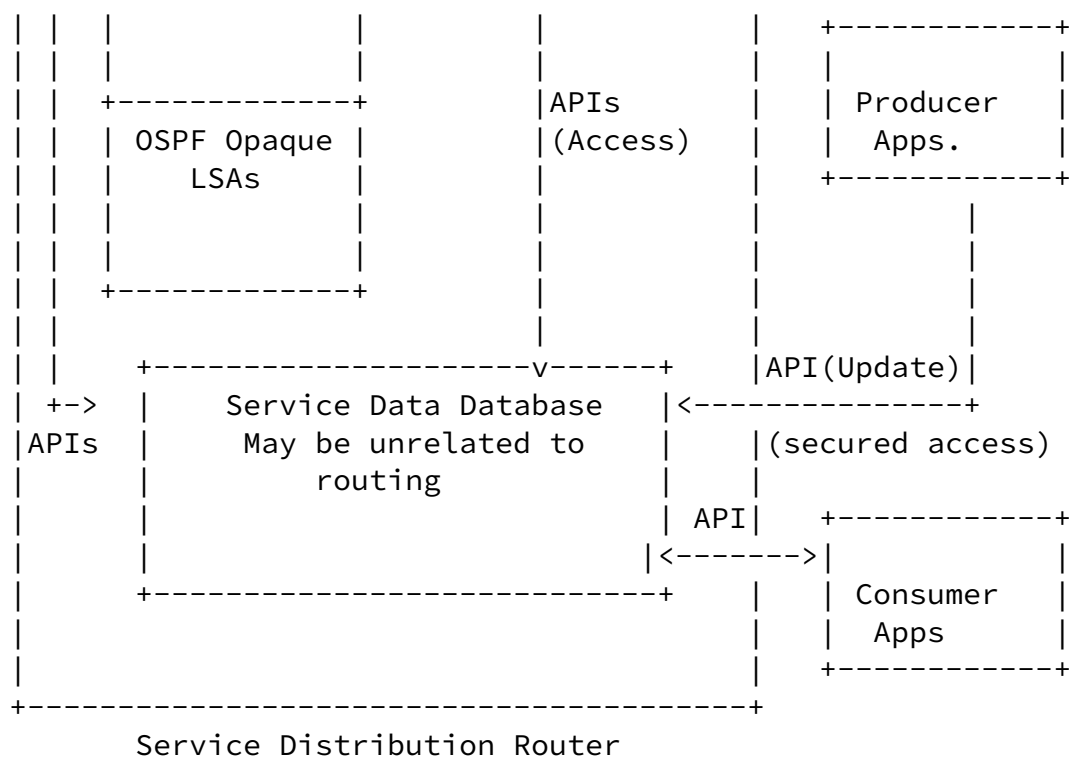
A SDR leverages OSPF's capability to store and flood the topology and other attributes of SDR capable routers. SDRs form an overlay and do not require to be directly connected to each other. SDRs do not need to maintain adjacency between them other than the normal OSPF adjacency for routing purposes. The SDRs rely on the OSPF underlying network for reachability to other SDR routers.

SDRs advertise a directory of producers and consumers of services and are capable to compute preferred producers. The SDRs delegate data exchange processing to remote SDRs to an external agent. This agent is described in detail in [section 9](#) of this document.

The OSPF Opaque LSAs is used to carry relevant and interesting information for reachability and nature of SDR capable routers.

In order to limit the service data dissemination costs (storage, bandwidth, security, ..), SDRs may store only data of interest.





The OSPF Service Distribution Router

The following sections describe the extensions in OSPF protocol to support this capability.

6. Storage Of Service Data

The service data can be stored in an independent Service Data Database(SDD). There is no assumption made here on the size, format or nature of data. The data can even be stored on the disk of the router and accessible by APIs to OSPF and other applications for

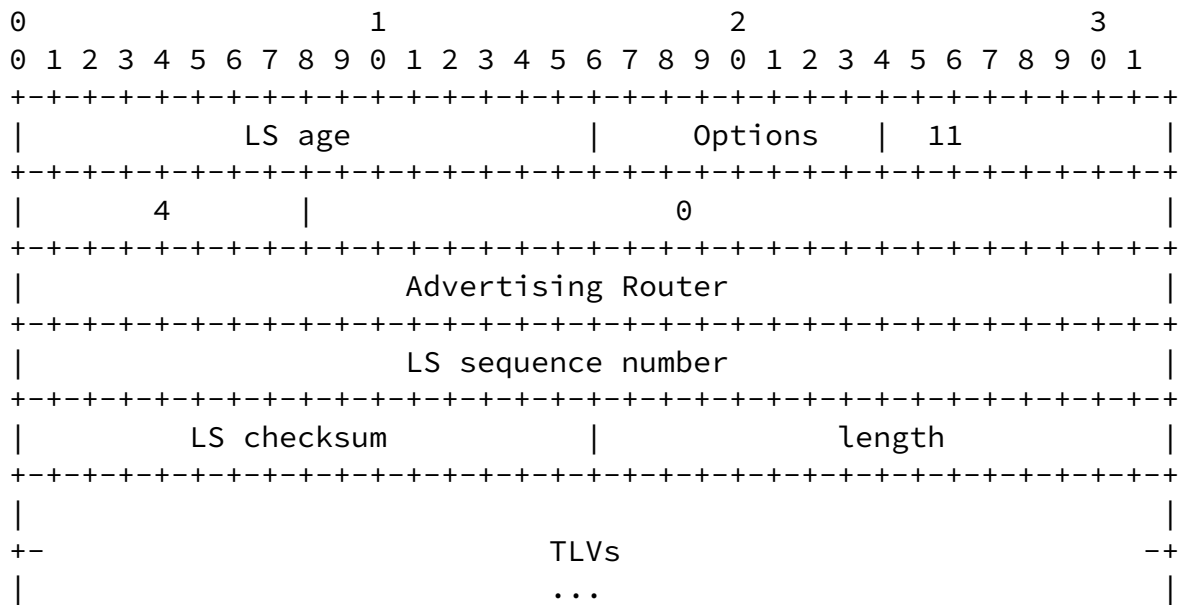
query and update. A SDD is not part of OSPF and does not participate in the bringing up of adjacencies.

It is desirable that the service data database have a very flexible format to cater for a broad range of applications. A possible solution is that the database records be defined as container objects which themselves contain service metadata.

7. Mechanics of the OSPF Service Information Distribution Implementation

7.1. Advertising and Signaling of SDR Capability

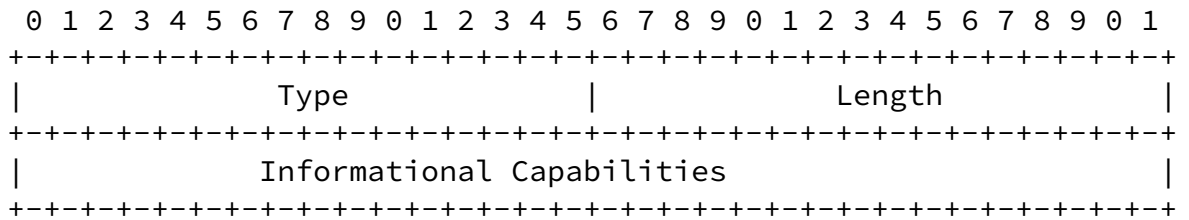
The OSPF SDR router will identify itself to the rest of the domain by advertising its capability and a routable ip address. For example, this address MAY be a loopback interface configured to uniquely identify an OSPF SDR router. A new bit for SDR capability is reserved in the Router Information Capabilities TLV of the Router Information LSA, as defined in [section 2.1 of \[RFC4970\]](#).



The OSPF Router Information LSA

Flooding scope for AS 11

The format of the Router Information Capabilities TLV is defined in 2.3 and 2.4 of [\[RFC4970\]](#)



The Router Informational Capabilities TLV

A new informational capability bit is defined for Service Distribution Routers

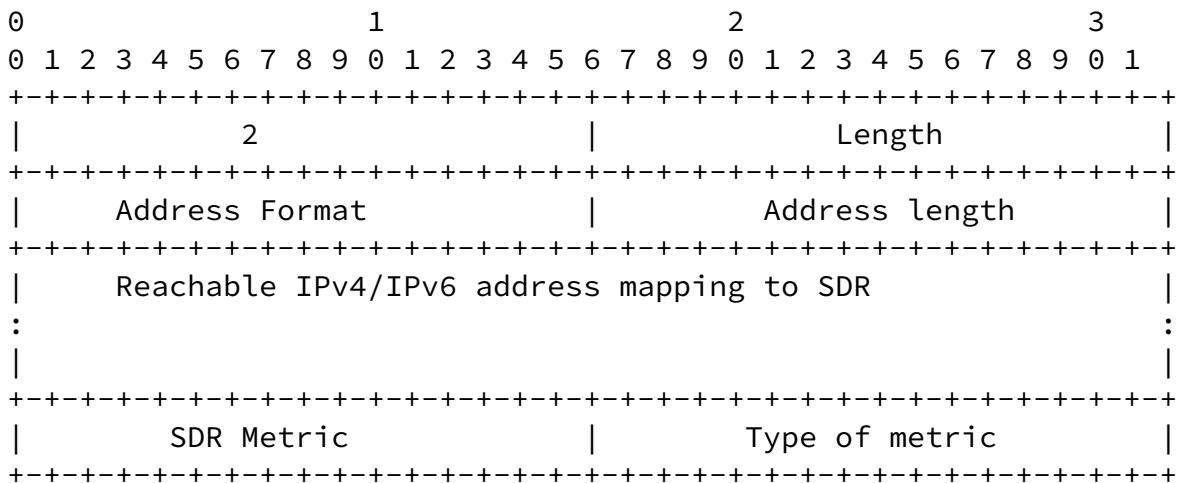
Bit Capabilities

6 Service Distribution Router Capability

[7.2.](#) Advertising the Service Distribution Router and its address mapping

A new TLV is defined in the Router Information LSA is used to advertise a routable address to reach the router.

TLV



Service router TLV and address Mapping

Type: A 16-bit field set to 2 representing the Service Distribution Router Address Mapping This TLV is applicable both to OSPFv2 and OSPFv3.

Length: A 16-bit field that indicates the length of the value portion in octets.

Address Format: A 16-bit field that indicates the length of the value portion in octets.

Possible values

1 : IPv4 Address

2 : IPv6 Address

Address Length: A 16-bit field that indicates the length of the value portion in octets.

Address : Routable IPv4/IPv6 address mapped to SDR

SDR metric: A 16-bit field that indicates SDR metric greater than 0.

Type of metric:

0 : None defined - Ignore SDR Metric

1 : SDR metric overrides the IGP metric

2 : Computed metric is composite of IGP metric + SDR metric

[7.3.](#) Advertising the Directory of Producers and Consumers

Opaque LSAs with autonomous system flooding scope as described in [[RFC5250](#)], are used to describe the services reachable through this router using TLVs.

Definition of TLV

Internet-Draft

Service Distribution using OSPF

March 2012

```

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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          3          |          Length          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Producer          | Number of services(subTLVs) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          Sub TLV Description Serv n          :
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
:          .          :
~+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+~
|          Sub TLV Description Serv m          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Subscriber          | Number of services of interest (SubTLVs) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          Sub TLV Subscribe Serv x          :
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
:          .          :
~+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+~
|          Sub TLV Subscribe Serv y          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
:          .          :

```

Example of TLV for Directory of Producers and Subscribers

Type: A 16-bit field set to 3 representing the Directory of the Service Distribution Router. This TLV is applicable both to OSPFv2 and OSPFv3.

Length: A 16-bit field that indicates the length of the value portion in octets.

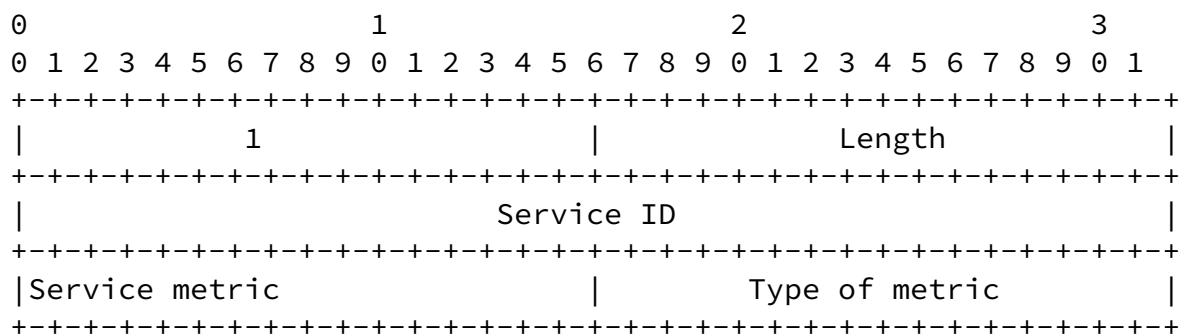
Services are described in a unique sub-TLV. The sub-TLV should contain a Service Identifier which uniquely identifies the service with network wide significance. The sub-TLV format should be flexible and it MAY be used to advertise a preference metric for the service.

Internet-Draft

Service Distribution using OSPF

March 2012

TLV



Service Description Sub-TLV

Type: A 16-bit field set to 1 representing the Service Description Sub-TLV This TLV is applicable both to OSPFv2 and OSPFv3.

Length: A 16-bit field that indicates the length of the value portion in octets.

Service ID: A 32-bit field representing the Service Identifier. This TLV is applicable both to OSPFv2 and OSPFv3.

Service Metric: A 16-bit field that indicates the metric associated with the service. A metric of 0 would represent undefined. An unreachable or oversubscribed service has a metric of 0xFFFFFFFF.

Type of metric:

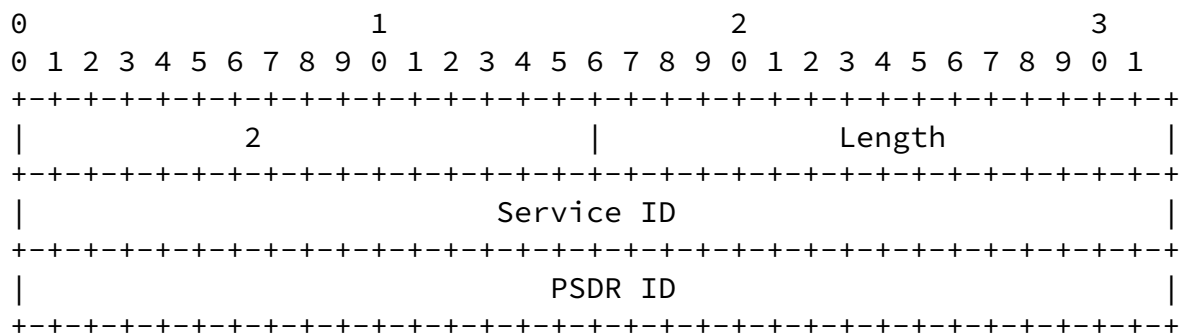
0 : None defined - Ignore Service Metric

1 : Service metric overrides the IGP/SDR metric

2 : Computed metric is composite of IGP metric + SDR metric + Service metric

The Services of interest (Consumers exist) are described in a unique sub-TLV. The sub-TLV should contain a Service Identifier which uniquely identifies the service with network wide significance. The sub-TLV format should be flexible and MUST contained the preferred SDR ID. If no producer exists yet for the service then the Preferred SDR ID should be set to 0.

TLV



Service Subscription Sub-TLV

Type: A 16-bit field set to 2 representing the Service Subscription Sub-TLV This TLV is applicable both to OSPFv2 and OSPFv3.

Length: A 16-bit field that indicates the length of the value portion in octets.

Service ID: A 32-bit field representing the Service Identifier. A Service Identifier may only be defined in a unique sub-tlv. This TLV is applicable both to OSPFv2 and OSPFv3.

PSDR ID: A 32-bit field that indicates the PSDR for data exchange.

Set to 0 if there is no producer for the service.

The topological view and characteristics of the OSPF Overlay Service Distribution Routers can be used to compute preferred producer independent of IGP metrics. It is possible to have multiple LSAs for large directories however a service must be described in a unique sub-tlv for the SDR.

[7.4.](#) Service Routing Capable Router Operations

The additional requirements are

- No assumption on topology

- Multiple producers may exists

- Multiple consumers can all have different service interest

- Producers/Consumers may join and leave at anytime

- Consumers and Producers have access to the Service Data database

The SDR capable routers advertise the consumers who subscribe to a

service. The producers may connect to the SDR router to update the services/data in the Service Data database. The SDR router then builds the Opaque LSA describing the producer services which are reachable through it as well as the services its consumers are interested in.

When the router has a full neighbor relationship, it now has the topological view of all SDR capable routers in the domain as well as the services they offer and are interested in.

Leveraging the fact that the OSPF has already run its SPF, the reachability of overlay SDR capable routers and services offered. It is possible to calculate the preferred Producer SDR for a service by using a composite of the IGP metric, the SDR metric and the service metric. The list of preferred producers for a service can then be evaluated at each SDR.

The list of Consumer SDRs interested in service can also easily be

computed from the directory of consumers.

[7.4.1.](#) Operation due to Producer changes

The producer service operations are

New producer advertises a service

Existing Producer start advertising a new service

Existing Producer stops advertising a service.

The router will be notified by the application regarding the new producer and the services offered. The router will then either update or create an Opaque LSA to advertise this new information and flood it to all SR routers.

Upon receiving this information, remote SDR routers can recalculate the preferred PSDR. It may also need to perform some operations if it have consumers for this new service.

[7.4.2.](#) Operation due to Consumer changes

The consumer service operations are

A new consumer join and add subscription

An existing consumer stops subscriptions

An existing consumer adds subscriptions

The router will be notified by the application regarding the new consumer and the services it is interested in. The router will then either update or create an Opaque LSA to advertise this new information and flood it to all SDR routers.

Upon receipt of the new Opaque LSA the remote SDR routers can then update the list of CSDRs interested in their services per latest information.

[8.](#) Calculation of Optimal Producer

Leveraging OSPF capability to store and compute paths on a topology, the same mechanisms can be used to compute the Optimal PSDRs using the SPF for SDR reachable address using IGP metrics, SDR metric and the service metric. The Optimal PSDR is used in the consumer subtlv.

[9.](#) Service Router Data Operations

OSPF SDR delegates the task of SDD distribution to the Data Exchange Agent. This text defines an implementation of such an agent and named it the Service Data Distribution Agent (SDDA). OSPF SDR provides SDDA information about which Consumer SDR is interested which service provided by this OSPF (Producer) SDR. The SDDA makes use of such information to setup distribution channel for SDD distribution from this OSPF Producer SDR to other OSPF Consumer SDRs.

For each OSPF Consumer SDR which subscribes to at least one service provided by this OSPF Producer SDR, there will be a different distribution channel created.

The distribution channel is setup when the OSPF Consumer SDR has subscribed to its first service provided by this OSPF Producer SDR. When the OSPF Consumer SDR subscribes to additional service provided by this OSPF Producer SDR, service data for the new service will be carried over the existing distribution channel. In other words, the same distribution channel can carry service data for different services. The services carried are said to be bound to the distribution channel.

When a distribution channel is first setup for a service or a new service is bound to the channel, the SDDA will notify the SDD. In turn, the SDD will send the latest data for that service to the SDDA for distribution over that channel.

On the other hand, whenever the SDD has new version of data for a service, the SDD will send those data to the SDDA, which will

distribute the new data to all the distribution channels which carry the service.

[9.1.](#) Implementation of SDDA

The SDDA uses TCP [[RFC793](#)] as its transport protocol for service data distribution. This eliminates the need to implement explicit update fragmentation, retransmission, acknowledgement, and sequencing. SDDA listens on TCP Port 1001 . The error notification mechanism used in SDDA assumes that TCP supports a "graceful" close (i.e., that all outstanding data will be delivered before the connection is closed).

The SDDA MUST listen on TCP Port 1001 for incoming connections.

Connection is unidirectional in respect of service data distribution. At most two connections exists between any two SDRs, one for each direction. The same connection is used to distribute data of multiple services from the PSDR to the CSDR.

Connection is initiated and teardown by the PSDR. The connection is referred as the outgoing connection on the PSDR and as incoming connection on the CSDR.

Connection is initiated when the PSDR discover, through the producer and consumer directory, that the CSDR has chosen it as the provider of a service. If PSDR already have an outgoing connection to the same CSDR, no new connection is made and the existing outgoing connection will be used to distribute data for the service.

Connection is teardown when the PSDR discover, through the producer and consumer directory, the CSDR no longer use it as provider for any services.

Connection is teardown when the CSDR discover, through the producer and consumer directory, the PSDR no longer support any services it is interested in.

Connection is also teardown when error occurs, as indicated by the NOTIFICATION message.

[9.2.](#) SDR Message Formats

This section describes message formats used by SDDA over the TCP connection.

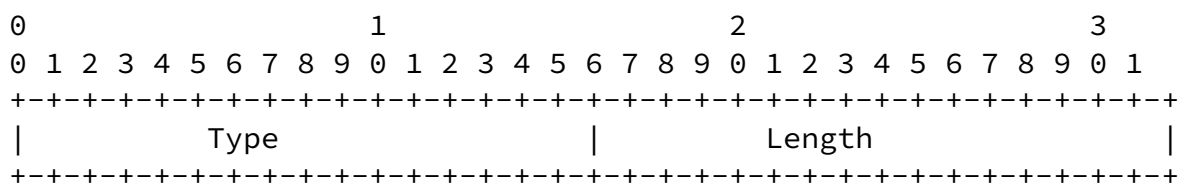
SDR messages are sent over TCP connections. A message is processed only after it is entirely received. The maximum message size is 65535, including the message header.

All multi-octet fields are in network byte order.

TCP A0 [[RFC5925](#)] is used for authentication.

9.3. SDR Message Header Format

Each message has a fixed-size header. There may or may not be a data portion following the header, depending on the message type. The layout of these fields is shown below:



SDR Message Header Format

Type : This 2-octet unsigned integer indicates the type code of the message. This document defines the following type codes:

- 1 - OPEN
- 2 - CONFIRM
- 3 - UPDATE
- 4 - NOTIFICATION

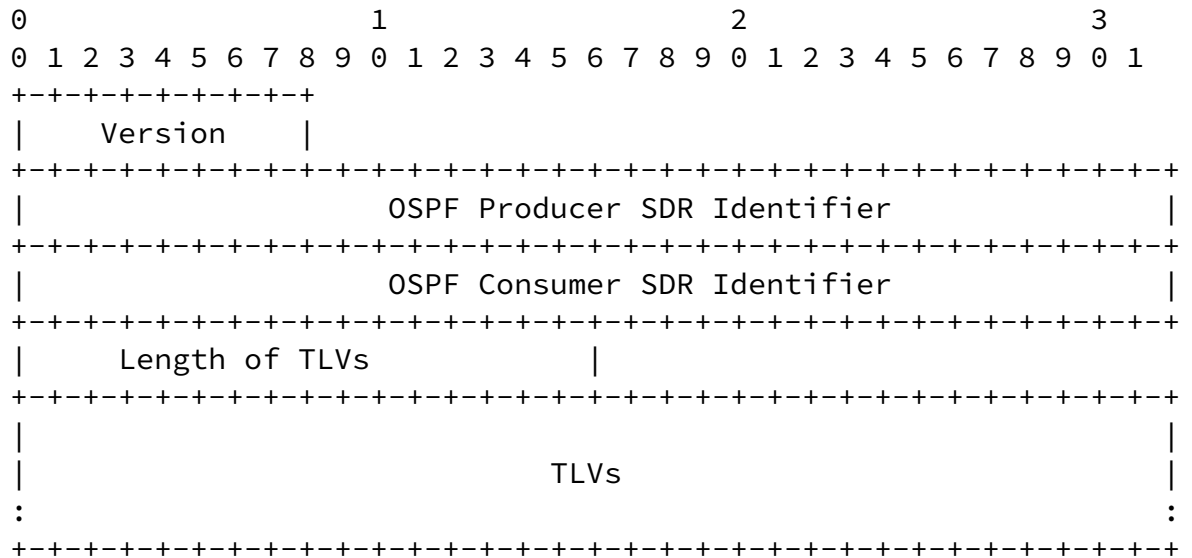
Length: This 2-octet unsigned integer indicates the total length of the message, including the header in octets. The maximum message length is 65535, and MAY be further constrained, depending on the message type. The value must be at least 4.

9.4. OPEN Message Format

After a TCP connection is established, the first message sent by both side is an OPEN message. If the OPEN message is accepted and processed successfully, a CONFIRM message confirming the OPEN is sent back.

Following the message header, the OPEN message contains the following fields:

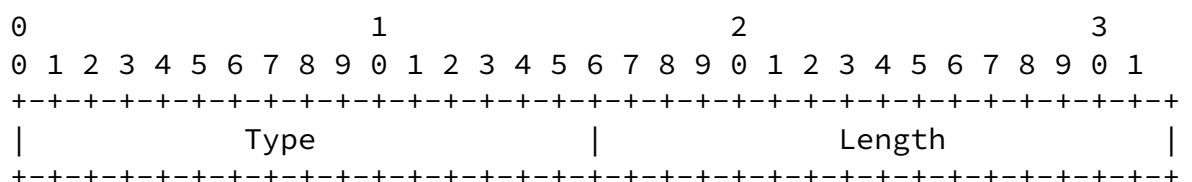
March 2012



Version: This 1-octet unsigned integer indicates the protocol version number of the message. The current OSPF SDR version number is 1.

OSPF Consumer SDR Identifier: This 4-octet unsigned integer indicates the OSPF router ID of the consumer SDR of the service.

TLVs: Zero or more Type-Length-Value tuple of the following format.



Value
<pre> +--+ </pre>

No TLV for OPEN message is defined at the moment.

If the OSPF Producer SDR Identifier or the OSPF Consumer SDR

Identifier does not match information learned through the producer and consumer directory, the NOTIFICATION is sent back to reject the connection.

The subTLVs could define further constraint that must be met before the connection is OPEN message is considered succeed.

The NOTIFICATION message will be sent to indicate if the OPEN is accept or not.

If OPEN is rejected, the TCP connection MUST be teardown after sending the NOTIFICATION. The TCP connection should be retried if the condition that cause the failure is corrected.

[9.5.](#) CONFIRM Message Format

CONFIRM messages are used to confirm the OPEN message is accepted.

It has no additional data following the header.

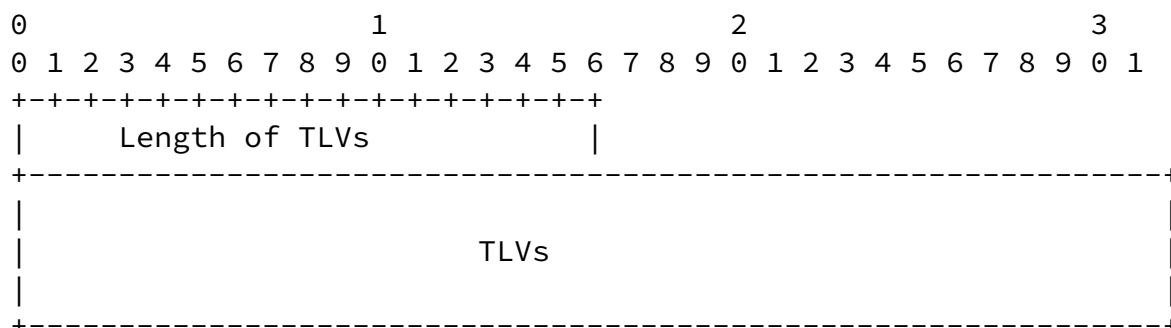
The OSPF Producer SDR MUST NOT send UPDATE message until a CONFIRM message is received.

The OSPF Consumer SDR MUST NOT accept UPDATE message until a CONFIRM message is received. If the OSPF Consumer receive UPDATE message before the CONFIRM message, it should send a NOTIFICATION and close the connection.

[9.6.](#) UPDATE Message Format

UPDATE messages are used to push the latest service data from SDD from OSPF producer SDR to OSPF consumer SDR.

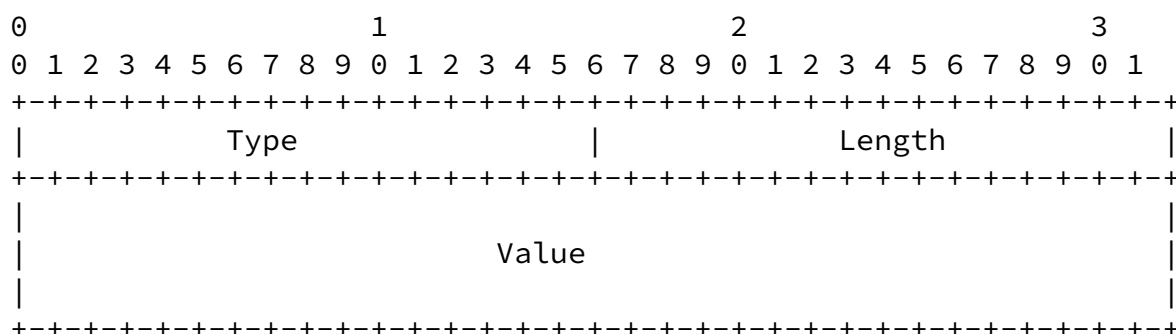
Following the message header, the UPDATE message contains the following fields:



UPDATE Message Format

Length of TLVs: This 2-octets unsigned integer indicates the total length of TLV information contained in the message.

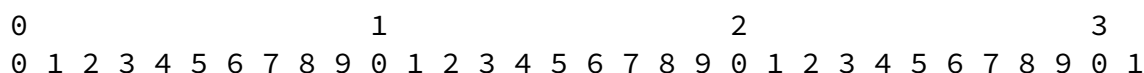
TLVs: Zero or more Type-Length-Value tuple of the following format.

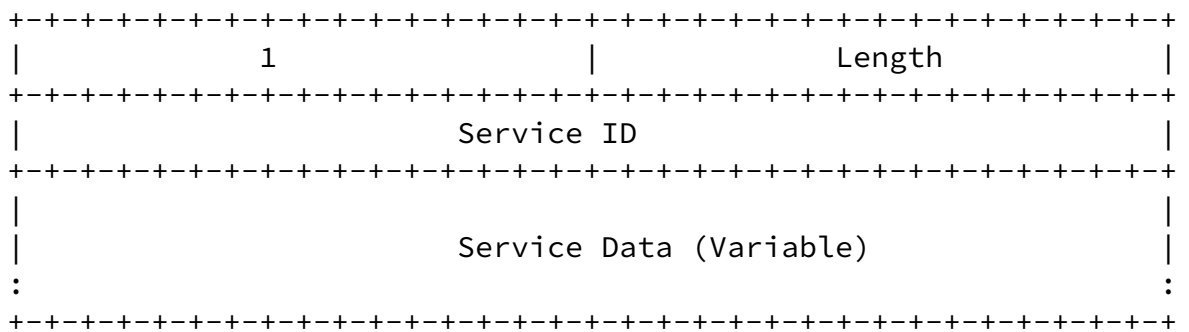


TLV Format

The following TLV is defined for the UPDATE message.

Services Update TLV:





Service Update TLV Format

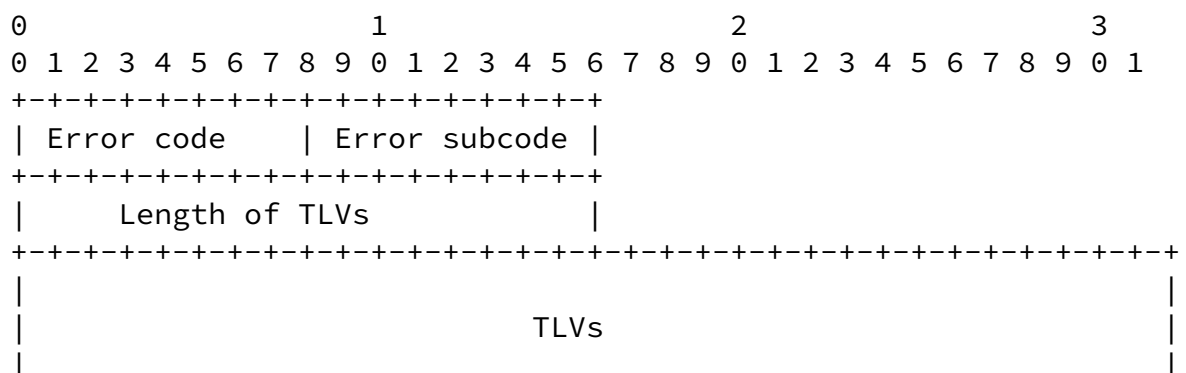
Service ID: The 4-Octets indicates the service this update is originated from.

Service Data: Various length of service data. Its format is opaque to OSPF SDR and is defined by the service itself.

[9.7.](#) NOTIFICATION Message Format

A NOTIFICATION message is sent when an error condition is detected. The TCP connection is closed immediately after it is sent.

Following the message header, the NOTIFICATION message contains the following fields:



[9.7.3.](#) CONFIRM Message Error

The error code 3 is used for error detected while processing the UPDATE message.

Error subcodes:

- 0 - Unknown
- 1 - Bad Message Sequence

[9.7.4.](#) UPDATE Message Error

The error code 4 is used for error detected while processing the UPDATE message.

Error subcodes:

- 0 - Unknown
- 1 - Bad Message Sequence
- 2 - Malformed TLV

[9.7.5.](#) Cease

The error code 5 is used when the OSPF SDR choose to close the connection due to reason other than message processing.

Error subcodes:

- 0 - Unknown
- 1 - No Producer
- 2 - No Consumer

[10.](#) Security Considerations

The new extensions defined in this document do not introduce any new security concerns other than those already defined in [Section 6 of \[RFC2328\]](#) and [\[RFC5340\]](#).

11. IANA Considerations

IANA is requested to allocate the value 1001 to "Service Data Distribution " within the "Port Numbers" registry.

12. Contributors

The authors would like to acknowledge the contributions of

Mike Dubroskiy

Rashmi Shrivastava

Jean-Michel Esnault

13. Acknowledgments

The authors would like to thank Les Ginsberg, Keyur Patel and many others who participated in numerous discussions.

This document was produced using Marshall Rose's xml2rfc tool.

14. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC2328] Moy, J., "OSPF Version 2", [RFC 2328](#), April 1998.
- [RFC4970] Lindem, A., Shen, N., Vasseur, JP., Aggarwal, R., and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities", [RFC 4970](#), July 2007.
- [RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", [RFC 5250](#), July 2008.
- [RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", [RFC 5340](#), July 2008.
- [RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", [RFC 5925](#), June 2010.
- [RFC793] Postel, J., "Transmission Control Protocol", [RFC 793](#), September 1981.

Internet-Draft

Service Distribution using OSPF

March 2012

Authors' Addresses

Padma Pillay-Esnault
Cisco Systems
510 McCarty Blvd
Milpitas, CA 95035
USA

EMail: ppe@cisco.com

Burjiz Pithawala
Cisco Systems
510 McCarty Blvd
Milpitas, CA 95035
USA

EMail: bpithaw@cisco.com

Derek Yeung
Cisco Systems
510 McCarty Blvd
Milpitas, CA 95035
USA

EMail: myeung@cisco.com

