Network Working Group A. D. Zinin
Internet Draft AMT Group
Expiration Date: January 2000 July 1999
File name: draft-ietf-ospf-shortcut-abr-00.txt

OSPF Shortcut ABR
Enhanced OSPF ABR Behavior
draft-ietf-ospf-shortcut-abr-00.txt

Status of this Memo

Abstract

OSPF [Ref1] is a link-state intra-domain routing protocol used for routing in IP networks. Though the definition of the ABR in the current OSPF specification does not require a router with multiple attached areas to have a backbone connection, it is actually necessary to provide successful routing to the inter-area and external destinations. If this requirement is not met all traffic, destined for the areas not connected to such an ABR or out of the OSPF domain, is dropped. The rules of originating and processing Summary-LSAs given in the current OSPF standard [Ref1] can also

result in suboptimal inter-area routing. Though all these problems can be fixed using virtual links, this memo describes an alternative implementation of the OSPF ABR behavior, which allows the administrator to avoid it or, if virtual links are still used, to decrease the number of configured virtual links.

This memo also describes possible situations where the proposed implementation can be used.

Acknowledgements

The author would like to acknowledge Christian Hopps and Peter Psenak for their help in finding weak points in early versions of this document and Thomas M. Thomas for reviewing the preceding version of it.

Special thanks go to John Moy who contributed a lot to this document and provided a simpler algorithm representation, used herein.

Table of Contents

1 Overview

1.1 Introduction

An OSPF routing domain can be split into several subdomains, called areas, which limit the scope of LSA flooding. A router having attachments to multiple areas is called an "area border router" (ABR). The primary function of an ABR is to provide its attached areas with Type-3 and Type-4 LSAs (which are used for describing routes and ASBRs in other areas) as well as to perform actual inter-area routing.

1.2 Motivation

In OSPF flooding of Type-1 and Type-2 LSAs is limited to the area borders, so routers in other areas must somehow know how to reach destinations and ASBRs residing in different areas. OSPF uses Distance-Vector (DV) approach to achieve this goal, i.e., Area Border Routers announce routes and ASBRs internal to directly connected areas in Type-3 and Type-4 Summary-LSAs.

If routers using a DV protocol announce only directly attached networks, they must be fully meshed to provide complete routing information to each other. This condition cannot always be met, so routers also announce the networks they heard about from their neighboring routers. This is the main reason for loops of routing updates in DV protocols, which are solved with such methods as split-horizon, counting-to-infinity, triggered updates, and holddown-timers. Application of these rules to OSPF inter-area routing would make the code very complex, but since areas in OSPF need not be fully meshed, ABRs are allowed to reannounce inter-area routes. In order to prevent loops of summaries in OSPF, ABRs reannounce only those inter-area routes which are associated with the backbone area. Summaries from non-backbone areas are just not considered by ABRs. Because inter-area routes are not reannounced back into the backbone area, the latter functions as a loop-free inter-area routing information repository. In order to achieve normal routing to inter-area and AS-external destinations, all areas in OSPF must be connected to the backbone either physically (via an interface) or logically (via a virtual link). This is to ensure that all areas are provided with inter-area routes from the backbone.

A basic discussion of the disadvantages of the standard inter-area approach are given in [Ref2] and are applicable to this document as well. In addition to that, consider another problem caused by standard OSPF ABR behavior (Figure 1).

```
     .  Area 2         .
     .                 .
     .      +--+       .
      ......|R5|.......
     .      +--+       .
    .       /  \        .
   .       /    \        .
   . BB   /10Mb  \ 2Mb   .
   .     /        \      .
    .    |         \    .
     .  +--+      +--+ .
      ..|R1|......|R2|..
     .  +--+      +--+  .
     .   |  10Mb   |\   .
     .  ------------ |  .
     .               |  .
     .             +--+ .
     .             |R4| .
     . Area 1      +--+ .
      ..................
```

Figure 1. Suboptimal inter-area routing

In this example router R2 has a 2Mb link to R5. At the same time R1 has a better link (10 Mbps), but R2 cannot route traffic going to area 2 through R1. This is because according to [Ref1] R2 is not allowed to consider summary-LSAs from non-backbone areas and, consequently, does not have routes covering destinations in area 2 via R1. The situation looks even more interesting if R4's routing table is considered. Since R2 floods summary-LSAs from R1 to R4, router R4 will have routes to the area 2 via R1 (the best path), expecting traffic to go via 10Mbps links. In reality R2 will not direct traffic to R1, but will forward it via 2Mbps link attached to itself.

The last example shows how the main principle of OSPF---prefer the shortest path---is broken due to distance vector approach used for inter-area routing. Again, the problem can be fixed using the virtual links between R1 and R2 in standard OSPF, but the solution proposed in this document appears to be more elegant and involving no administrative and traffic overhead. More sophisticated examples of how Shortcut ABR approach improves inter-area routing are given in section 6.

2 Description of Shortcut ABR behavior

This section describes an alternative implementation of OSPF ABR behavior, named "Shortcut ABR". It is an improvement on standard ABR behavior, based on relaxation of the restrictions applied to the calculation of the inter-area routes.

ABRs are allowed to consider summary-LSAs from all attached areas (no matter if it is connected to the backbone or not). The routing loop prevention is done by restricting origination of summary-LSAs---inter-area routes are readvertised only if there is a valid summary-LSA for the destination learned from the backbone. Origination of summary-LSAs for intra-area routes is done as in standard OSPF, described in [Ref1].

The relaxation of the routing table calculation allows ABRs without a backbone connection to route traffic between the attached areas, as well as to route traffic destined for the backbone and other areas using the routes derived from the summary-LSAs in each attached area. This approach also enables router R2 in Figure 1 to route inter-area traffic via R1.

Note that the proposed solution does not obviate the need of virtual link configuration in case an area has no physical backbone connection at all. The method described here improves the behavior of a router connecting two or more backbone-attached areas.

Though this document is initially oriented to processing Type-3 LSAs and, consequently, is targeted to improving OSPF inter-area routing, it's acceptable to apply described methods to Type-4 LSAs, which will lead to improvement of external routing in an OSPF domain.

3 Proposed changes to OSPF ABR behavior

This section describes the changes made to the base OSPF described in [Ref1].

3.1 Changes to Router-LSA origination

The algorithm of Type 1 LSA (router-LSA) origination is changed to have the Shortcut ABR announce its Shortcut capability in the Router-LSA as described in A.1. A Shortcut ABR must set the S-bit in the Router-LSA for Area A only if the area A's data structure has ShortcutConfigured bit set to TRUE, i.e., the S-bit directly reflects the state of ShortcutConfigured flag (see section 3.2 for more details). As in [Ref1] Shortcut ABRs identify themselves as ABRs by setting the bit B in their Router-LSAs when they have

more than one attached area.

3.2 Changes to routing table calculation

Shortcut ABRs maintain two additional flags in Area Data Structure for every non-backbone area. The flags are named Shortcut-Configured and ShortcutCapability. The first flag indicates whether the administrator has configured the area to be Shortcut. If so the ABR will set the S-bit in its Router-LSA for this area. The second flag indicates that all other ABRs in the areas are also Shortcut capable. Note that Shortcut ABR is allowed to consider summary-LSAs from a non-backbone area only if ShortcutCapability flag is set to TRUE. ShortcutCapability flag, in turn, can be set to TRUE only if ShortcutConfigured flag is also set to TRUE. This means that the area must be configured as Shortcut on the ABR itself and all other ABRs.

If during the routing table calculation a Shortcut ABR notices that there is a ABR which is not Shortcut-capable in any area, the Shortcut ABR must clear the ShortcutCapability flag for that area, but still announce the ShortcutConfigured flag for the area in the S-bit of the Router-LSA originated for this area.

Should the ABR in question find that there no ABRs in an area, which are not Shortcut-capable, it must set the ShortcutCapability flag for that area.

To implement this algorithm Steps 1 and 2 in section 16.1 of [Ref1] are changed as follows:

Step 1:

"Initialize the algorithm's data structures. Clear the list of candidate vertices. Initialize the shortest-path tree to only the root (which is the router doing the calculation). Set Area A's TransitCapability to FALSE and ShortcutCapability to the value of ShortcutConfigured."

Step 2:

"Call the vertex just added to the tree vertex V. Examine the LSA associated with vertex V. This is a lookup in the Area A's link state database based on the Vertex ID. If this is a router-LSA, and bit V of the router-LSA (see Section A.4.2) is set, set Area A's TransitCapability to TRUE. If this is a router-LSA, and bit B of the router-LSA is set

(the router is an ABR) and bit S of the router-LSA is not set (the ABR is not Shortcut-capable), set Area A's ShortcutCapability to FALSE. In any case, each link described by the LSA gives the cost to an adjacent vertex. For each described link, (say it joins vertex V to vertex W):"

ShortcutCapability flag is used to determine over which areas the ABR can use shortcut paths. Shortcut ABRs are forbidden to consider Summary-LSAs from the areas with ShortcutCapability flag off. This method is introduced to prevent possible routing loops when standard and Shortcut ABRs are simultaneously present in an OSPF domain. The method also allows for gradual enabling shortcutting over specific non-backbone areas when and where it is necessary.

The algorithm of calculating inter-area routes is changed to allow the router to consider the summary-LSAs from attached non-backbone areas that have ShortcutCapability flag set to TRUE. This is achieved by applying section 16.3 of [Ref1] to such areas. The following changes to 16.3 are made.

Paragraph 1 of 16.3 is changed to be as follows:

"This step is only performed by area border routers attached to one or more non-backbone areas that are either capable of carrying transit traffic (i.e., "transit areas", or those areas whose TransitCapability parameter has been set to TRUE in Step 2 of the Dijkstra algorithm (see Section 16.1) or have all ABRs supporting Shortcut feature (i.e., those areas whose ShortcutCapability parameter hasn't been set to FALSE during the Dijkstra algorithm)."

Paragraph 4 of 16.3 is changed to be as follows:

"The calculation proceeds as follows. All summary-LSAs of the areas with TransitCapability or ShortcutCapability parameter set to TRUE are examined in turn. Each such summary-LSA describes a route through a non-backbone area Area A to a Network N (N's address is obtained by masking the LSA's Link State ID with the network/subnet mask contained in the body of the LSA) or in the case of a Type 4 summary-LSA, to an AS boundary router N. Suppose also that the summary-LSA was originated by an area border router BR."

Step (3) of the algorithm in 16.3 is changed to be as follows:

"Look up the routing table entry for N. (If N is an AS boundary router, look up the "router" routing table entry associated with the backbone area). If the route type is other than backbone intra-area or inter-area (associated with any area) then examine the next LSA.

In other words, this calculation updates backbone intra-area routes found in Section 16.1, inter-area routes found in Section 16.2 and installs new inter-area routes if the ABR does not have a backbone connection."

Step (5) of the algorithm in 16.3 is changed to be as follows:

"If this cost is less than the cost occurring in N's routing table entry, overwrite N's list of next hops with those used for BR, and set N's routing table cost to IAC. Else, if IAC is the same as N's current cost, add BR's list of next hops to N's list of next hops. If the area associated with N's routing table entry is the backbone, then the area and the type of the path (either intra-area or inter-area) must remain unchanged. Otherwise (the routing table entry does not exist or the associated area is not the backbone), the type of the route must be set to inter-area and associated area must be set to the area associated with the summary-LSA being processed."

In order to prevent routing loops sections 11.1 and 16.2 of [Ref1] are changed.

Section 11.1 is restricted to require installation of discard routing table entries for each of the router's active area range. So the paragraph 2 of 11.1 should be read as follows:

"Before the lookup begins, "discard" routing table entries MUST be inserted into the routing table for each of the router's active area address ranges (see Section 3.5). (An area range is considered "active" if the range contains one or more networks reachable by intra-area paths.) The destination of a "discard" entry is the set of addresses described by its associated active area address range, and the path type of each "discard" entry is set to "inter-area".[10]"

Step (3) of section 16.2 is changed to instruct the ABRs to ignore summary defaults received from stub areas:

"If it is a Type 3 summary-LSA, and the collection of destinations described by the summary-LSA equals one of the router's configured area address ranges (see Section 3.5), and the particular area address range is active, then the summary-LSA should be ignored. "Active" means that there are one or more reachable (by intra-area paths) networks contained in the area range. The summary-LSA should also be ignored if it is a summary default (Destination ID = DefaultDestination, Address Mask = 0x00000000) and the area it has been received from is a stub area. This is to prevent possible routing loops."

3.3 Changes to Summary-LSA origination

The algorithm of the summary-LSAs origination is changed to include an explicit restriction not to originate summary-LSAs for inter-area routes if the route to the destination is not associated with the backbone.

Note that if there are multiple alternative paths to a destination, some of which are via the backbone and the rest are via non-backbone areas, the area associated with the corresponding routing table entry will remain the backbone area, but the set of next hops will actually direct traffic along the best path even through non-backbone areas.

If the ABR in question has no backbone connection, it will not originate summary-LSA for any inter-area route in any area, because the area associated with the routing table entry will never be the backbone area.

The ABR will also not readvertise an inter-area route from non-backbone area if its backbone link state database does not contain a summary-LSA or router-LSA covering a specific destination.

In order to implement described policy, the paragraph 2 in section 12.4.3 of [Ref1] should be read as follows:

"... Note that only intra-area routes are advertised into the backbone, while both intra-area and inter-area routes are advertised into the other areas. Also, summary-LSAs for inter-area routes are originated if and only if these routes are associated with the backbone area (to prevent loops of summary-LSAs)."

The 6th step of the algorithm given in sections 12.4.3 of [Ref1] must be read as shown below:

"Else, if the destination of this route is an AS boundary router, a summary-LSA should be originated if and only if the routing table entry describes the preferred path to the AS boundary router (see Step 3 of Section 16.4) and it is associated with the backbone area. If so, a Type 4 summary-LSA is originated for the destination, with Link State ID equal to the AS boundary router's Router ID and metric equal to the routing table entry's cost. Note: these LSAs should not be generated if Area A has been configured as a stub area."

The 7th step of the algorithm given in sections 12.4.3 of [Ref1] must be read as shown below:

"Else, the Destination type is network. If this is an inter-area route and it is associated with the backbone area, generate a Type 3 summary-LSA for the destination, with Link State ID equal to the network's address (if necessary, the Link State ID can also have one or more of the network's host bits set; see Appendix E for details) and metric equal to the routing table cost."

Described changes in the ABR behavior allow selection of most optimal paths to inter-area destinations. Note that backbone intra-area routes can be updated with better non-backbone inter-area one, thus directing internal backbone traffic along more optimal paths through other areas.

4 Implementation Details

If the current implementation of OSPF uses the standard described in [Ref1], then support of the proposed Shortcut ABR behavior strategy must be implemented as an implicit configurable option, allowing to set ShortcutConfigured flag for a given area.

Note that the nature of the changes to OSPF presented in this document is so that standard ABR behavior is not altered until at least one area is configured as Shortcut.

5 Compatibility

ABRs following the approach described in this document are required to announce their Shortcut capability for a given area in Router-

LSAs. Since no loops are formed when all ABRs in a given area are Shortcut and Shortcut ABRs do not consider Summary-LSAs from an area when a Shortcut-incompatible ABR in such an area is seen, the approach described in this document is compatible with standard OSPF described in [Ref1].

6 Deployment Considerations

This section discusses the deployment details of Shortcut ABR.

6.1 Necessity of virtual links

First of all it should be repeated that Shortcut ABR behavior does not obviate the need for virtual links in case an area has no physical backbone connection. The difference with standard OSPF is that the administrator does not need to configure virtual links through all areas he or she wants the inter-area traffic to go through. A Shortcut ABR needs a single backbone connection (physical or virtual) to achieve optimal routing, since it considers summary-LSAs from all attached areas.

6.2 Change of traffic patterns

Use of Shortcut ABR can lead to changes in the paths inter-area traffic flows take comparing to those experienced with standard OSPF. This happens because the Shortcut ABR approach allows a router to find paths better than it is possible with the standard OSPF. While standard OSPF tries to forward all inter-area traffic through the backbone area (though it does not guarantee it), the Shortcut ABR finds best routes in the domain even across non-backbone areas. With Shortcut ABR the backbone area is used as a dedicated place of inter-area routing information exchange and inter-area traffic is allowed to cross non-backbone area borders if such a path is really the best.

6.3 Optimized inter-area routing

Use of Shortcut ABR improves inter-area routing in OSPF domains by allowing ABRs to consider summary-LSAs from all attached area and consequently readvertise them into non-backbone areas. Consider an example show in the Figure 2:

```
         .......................
        .       Backbone        .  .........
       .                        . .         .
       .      +--+            +--+          .
       .      |R1|            |R5|--|       .
       .      +--+            +--+ 1|       .
       .     8/ 8\          1/  .   |       .
       .     /    \   --------  .   |       .
        .  8/     8\  /1    1\  .   | Net N .
         .+--+     +--+       +--+ 1|       .
    ......|R2|.....|R3|.......|R4|--|       .
   .      +--+     +--+       +--+          .
   .       .|1    1/  \1    1/ . .   Area 3 .
   .       .--------  -------- .  ..........
   .       .                   .
   .       .                   .
   .Area 1 .    Area 2         .
    ....... ...................
```

Figure 2. Optimized inter-area routing

In case all ABRs use standard OSPF approach, routing to the net N would be as follows:

- o R4 and R5 inject summary-LSAs into the backbone

- o R4 also inject a summary-LSA into area 2

- o R3 is limited to consider summary-LSAs from the backbone only, so it doesn't see the alternative path through area 2 and always routes through the backbone (though parallel paths are available)

- o R3 injects summary-LSA for the inter-area routes derived from the backbone summary-LSAs and received from R4 and R5 into Area 2

- o R2 is not allowed to consider non-backbone summary-LSAs and routes via serial links to R1, though more optimal paths do exist

If R2, R3, and R4 use Shortcut ABR approach inter-area routing is

improved as shown below:

- o R4 and R5 inject summary-LSAs into the backbone

- o R4 also inject a summary-LSA into area 2

- o R3 considers summary-LSAs from both attached areas and installs the route through area 2 (it has three routes in the routing table---via R5, via R4 through the backbone, and via R4 through area 2) and performs traffic sharing between the two ethernet links.

- o R3 injects summary-LSA for the inter-area routes to N (it will be the same as in the previous case, actually)

- o R2 considers summary-LSAs from all attached areas and prefers the route through area 2 rather than the backbone.

6.4 Gradual deployment of Shortcut ABRs

Shortcut ABR behavior is designed in such a way that the administrator can enable shortcutting through non-backbone OSPF areas gradually.

Since Shortcut ABRs are allowed to consider summaries only of those areas that were configured as Shortcut (ShortcutConfigured flag in area data structure is set to TRUE) and whose ShortcutCapability flag is set to TRUE, it is easy to control which areas will accept additional inter-area traffic. For an area to become Shortcut-capable, all ABRs that have links in it must have this area configured as Shortcut. If a single ABR in an area does not announce the S-bit in its Router-LSA for this area, no other Shortcut ABRs connected to this area will direct inter-area traffic through it (except for the cases when standard OSPF behavior leads to it).

The implementers should note that support of a configurable option described in section 4 is very important for traffic control and successful deployment.

7 Security Considerations

Shortcut ABR behavior specified in this document does not raise any security issues that are not already covered in [Ref1].

8 Appendixes

A.1 Router-LSA

An OSPF router originates a router-LSA into each of its attached areas. The router-LSA describes the state and cost of the router's interfaces to the area. The contents of the router-LSA are described in detail in Section A.4.2 of [Ref1]. One more flag has been added to the router-LSA, called bit S below. This flag indicates whether the area has been configured as Shortcut on the ABR. Note that all ABRs in an area must announce the S-bit this area to be used in shortcutting.

```
     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   |       1       |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                        Link State ID                          |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                     Advertising Router                        |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                     LS sequence number                        |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |         LS checksum           |             length            |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |    rtype      |       0       |            # links            |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ +
    |                          Link ID                              | P
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ E
    |                         Link Data                             | R
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |     Type      |     # TOS     |        TOS 0 metric           | #
  + +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ L
  # |      TOS      |       0       |            metric             | I
  T +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ N
  O |                              ...                              | K
  S +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ S
  | |      TOS      |       0       |            metric             | |
  + +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ +
    |                              ...                              |
```

The router LSA

```
+---+---+---+---+---+---+---+---+
| * | * | * | S | W | V | E | B |
+---+---+---+---+---+---+---+-+-+
```

The rtype field

The following defines the flags found in the rtype field. Each flag classifies the router by function:

o bit B. When set, the router is an area border router (B is for border). These routers forward unicast data traffic between OSPF areas.

o bit E. When set, the router is an AS boundary router (E is for external). These routers forward unicast data traffic between Autonomous Systems.

o bit V. When set, the router is an endpoint of an active virtual link (V is for virtual) which uses the described area as its Transit area.

o bit W. Used in MOSPF [Ref3], when set, the router is a wild-card multicast receiver. These routers receive all multicast datagrams, regardless of destination. Inter-area multicast forwarders and inter-AS multicast forwarders are sometimes wild-card multicast receivers (see [Ref3] for more details).

o bit S. When set, the router is a Shortcut-capable ABR and intends to use the area for shortcutting provided that all other ABRs in this area agree on that (also announce the S-bit into this area). See sections 2 and 3 for more details.

9 References

[Ref1] J. Moy. OSPF version 2. Technical Report RFC 2328, Internet Engineering Task Force, 1998. ftp://ftp.isi.edu/in-notes/rfc2328.txt.

[Ref2] Zinin, Lindem, Yeung. Alternative OSPF ABR Implementations. Work in progress, Internet Engineering Task Force. draft-ietf-ospf-abr-alt-00.txt

[Ref3] J. Moy. Multicast Extensions to OSPF. Internet Engineering Task Force, 1998. http://www.ietf.org/internet-drafts/draft-ietf-mospf-mospf-01.txt.

10 Author's Address

Alex D. Zinin
AMT Group
8 Maly Znamensky per., bld. 1
Office 3b
121019 Moscow, Russia
E-mail: zinin@amt.ru