

Internet Draft
Expires: April 30, 2000
<draft-wimer-ospf-subareas-00.txt>

W. Wimer
J. Drake
FORE Systems, Inc.
October 1999

OSPF Sub-Areas

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of [Section 10 of RFC2026](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/ietf/lid-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at
<http://www.ietf.org/shadow.html>.

Abstract

This document outlines a proposal for extending the OSPF intra-domain routing protocol [[1](#)] to support an additional level of hierarchy for traffic engineering purposes. The additional level of hierarchy allows the construction of much larger networks.

With some modification, it should be possible to apply these principles to the IS-IS routing protocol as well.

Contents

	Acknowledgements	2
1	Introduction	2
2	General Approach	3
3	SABR LSA	7
4	Abstract Topology LSA	7
5	Sub-Area Leader Election	8
6	Inside/Outside Link Negotiation	10
7	Handling of Arbitrary Opaque LSAs	10

8	Handling of Explicit Routes for Traffic Engineering	10
9	Open Issues	10
10	Security Considerations	13
11	References	13
12	Authors' Addresses	13

Acknowledgments

The authors would like to thank Rob Rosenberry and Tayfun Gol of FORE Systems for reviewing this draft and providing insightful comments.

[1. Introduction](#)

This document outlines a proposal for extending the OSPF intra-domain routing protocol [[1](#)] to support an additional level of hierarchy for traffic engineering purposes. The described techniques allow existing OSPF areas to be further divided into sub-areas.

Several goals motivate this proposal. An additional level of hierarchy will allow finer-grained control over the size of individual link-state databases, thus allowing the construction of larger areas, and thus larger autonomous systems. Sub-areas provide an additional level of containment; routing problems within a sub-area will tend to be limited to that sub-area. By exposing the topology between sub-areas, this proposal supports traffic engineering across sub-areas.

Furthermore, this proposal includes the explicit goal of minimizing the number of routers that must be upgraded to sub-area functionality. Routers within the interior of sub-areas need not be upgraded. This aids incremental deployment.

With some modification, it should be possible to apply these principles to the IS-IS routing protocol as well.

NOTE: The concepts discussed in this document are at an early stage of development. This document will be expanded and refined as this architecture becomes fully completed.

[2. General Approach](#)

The network administrator divides an existing OSPF area into contiguous regions called sub-areas. Routers at the border of a sub-area are known as sub-area border routers, or SABRs.

Unlike OSPF areas, which cut through the middle of routers, sub-areas

encompass routers and cut through links. Links are manually designated "inside links" if they are inside a sub-area, or "outside links" if they connect to routers in other sub-areas.

Whereas the OSPF area mechanism completely hides the topology of one area from another, this proposal exposes a simplified, abstract topology across sub-areas. In particular, the physical links between sub-areas and the SABRs are exposed across all sub-areas (i.e. across the entire OSPF area). Figure 2 below shows an abstract representation of the network from sub-area 2's perspective: all of sub-area 2's topology is visible, but only simplified versions of sub-areas 1, 3, and 4 are visible. (Note that a non-SABR router within sub-area 2 (e.g. RT13) is unaware of sub-area boundaries (actually, it is unaware of sub-areas entirely!). It just sees what it thinks is a normal OSPF topology, but that topology is simpler than the real topology.)

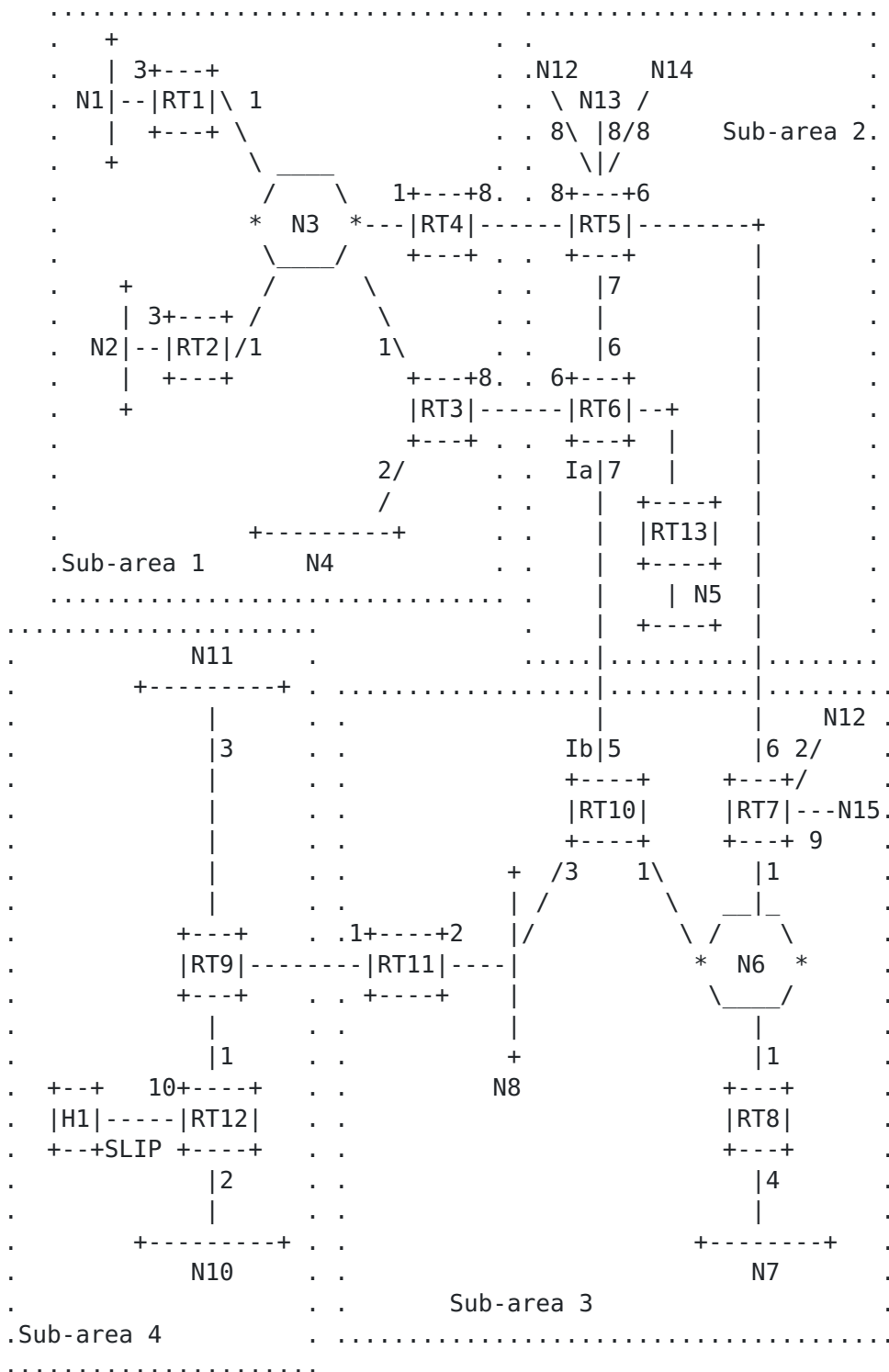


Figure 1: A sample sub-area configuration

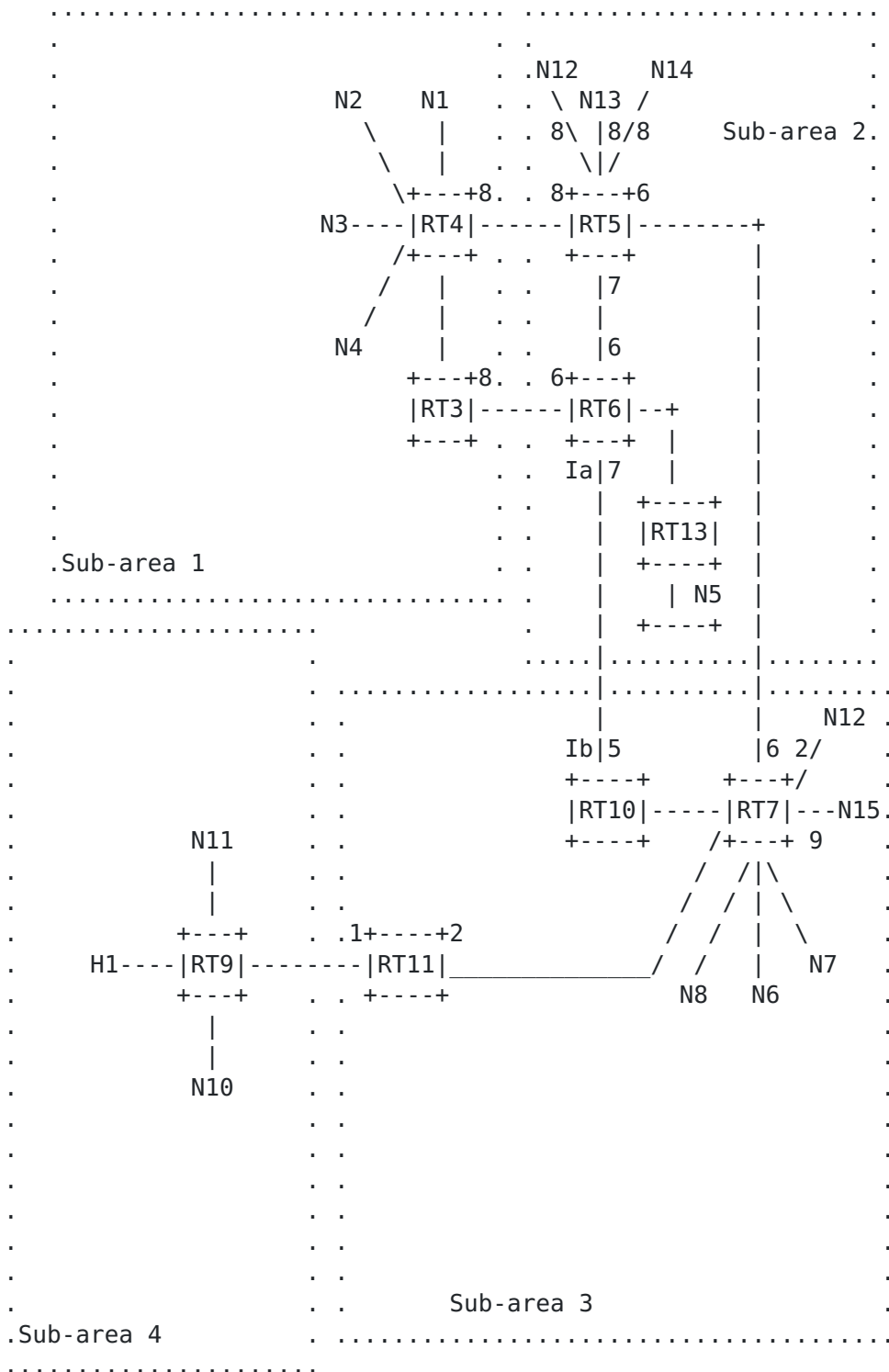


Figure 2: Topology from Sub-area 2's perspective

Outline of approach:

1. Each SABR generates an opaque LSA within its own sub-area that serves to identify itself to other SABRs within the sub-area.
2. Static configuration, or an election process similar to that of the Designated Router election process is used to select a Sub-Area Leader.
3. The Sub-Area Leader generates an Abstract Topology opaque LSA that describes the abstract topology of the sub-area. The topology includes all of the SABRs and shows them being connected to the Sub-Area Leader. In addition, all networks appear as if they were connected directly to the Sub-Area Leader.

DISCUSSION:

Since it is necessary to "filter out" normal LSAs at sub-area boundaries (in order to hide a sub-area's detailed topology), it becomes necessary to somehow distinguish between LSAs that describe the real topology (and thus should be filtered) vs. LSAs that describe the abstract topology. This is the motivation for special opaque LSAs that represent the abstract topology.

Within a sub-area, we want to see:

- o normal LSAs describing all the links inside the sub-area
- o normal LSAs describing the links (and routers) between sub-areas.
- o normal LSAs describing the IP networks that are reachable within all sub-areas (i.e. that are connected to the SABRs in each sub-area).

We don't want to see:

- o LSAs describing routers or links to routers in other sub-areas that are not SABRs.
4. Abstract Topology LSAs are flooded throughout all sub-areas (i.e. throughout the enclosing OSPF area). This allows all SABRs to learn the entire abstract topology.
 5. SABRs do NOT flood normal Router and Network LSAs from non-SABR routers beyond their local sub-area. This keeps the detailed topology information about a sub-area local to the sub-area.
 6. When an SABR receives an Abstract Topology LSA on an outside link, it translates and incorporates the information into the

Router LSAs that it sends on its inside links. This allows the abstract topology information to be disseminated to normal (non-upgraded) OSPF routers within the local sub-area.

3. SABR LSA

Each SABR within an area must originate a SABR LSA to identify itself to all the other SABRs within the area. SABR LSAs are flooded throughout the enclosing area (i.e. across sub-area boundaries). The SABR LSA is used in the Sub-Area Leader election process.

The SABR LSA has the following format:

```

      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   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|Opaque Type TBD|               Opaque ID               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|               Advertising Router               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|               LS Sequence Number               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|               LS checksum               |               Length               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|               HelloInterval               |   Options   | Leader Pri |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|               LeaderDeadInterval               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|               Sub-Area Leader               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|               Backup Sub-Area Leader               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|               |
+               +
|               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

4. Abstract Topology LSA

This LSA is modeled after the normal Router LSA and describes a SABR's abstract links. These include all real links that are outside links plus the abstract link to the Sub-Area Leader. The Sub-Area

Leader describes all of its outside links, all abstract links to other SABRs, and abstract links to all networks that are otherwise internal to the sub-area. Abstract Topology LSAs are flooded throughout the enclosing OSPF area (i.e. across sub-area boundaries).

```

      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      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|Opaque Type TBD|          Opaque ID          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          Advertising Router          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          LS sequence number          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          LS checksum          |          length          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      0      |V|E|B|      0      |          # links          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          Link ID          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          Link Data          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Type      |      # TOS      |          metric          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          ...          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      TOS      |      0      |          TOS metric          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          Link ID          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          Link Data          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          ...          |

```

5. Sub-Area Leader Election

A Sub-Area Leader may be statically configured or dynamically elected. The Sub-Area Leader election process is similar to the OSPF Designated Router election process. This section describes the algorithm used for calculating a Sub-Area Leader and Backup Sub-Area Leader. The initial time a router runs the election algorithm for a sub-area, the Sub-Area Leader and Backup Sub-Area Leader are initialized to 0.0.0.0. This indicates the lack of both a Sub-Area Leader and a Backup Sub-Area Leader.

The Sub-Area Leader election algorithm proceeds as follows: Call the

router doing the calculation Router X. The list of neighbors attached to the sub-area and having established bidirectional communication with Router X is examined. This list is precisely the collection of Router X's neighbors (on this sub-area) whose state is greater than or equal to 2-Way. Router X itself is also considered to be on the list. Discard all routers from the list that are ineligible to become Sub-Area Leader. (Routers having Sub-Area Leader Priority of 0 are ineligible to become Sub-Area Leader.) The following steps are then executed, considering only those routers that remain on the list:

- (1) Note the current values for the sub-area's Sub-Area Leader and Backup Sub-Area Leader. This is used later for comparison purposes.
- (2) Calculate the new Backup Sub-Area Leader for the sub-area as follows. Only those routers on the list that have not declared themselves to be Sub-Area Leader are eligible to become Backup Sub-Area Leader. If one or more of these routers have declared themselves Backup Sub-Area Leader (i.e., they are currently listing themselves as Backup Sub-Area Leader, but not as Sub-Area Leader, in their Hello Packets) the one having highest Sub-Area Leader Priority is declared to be Backup Sub-Area Leader. In case of a tie, the one having the highest Router ID is chosen. If no routers have declared themselves Backup Sub-Area Leader, choose the router having highest Sub-Area Leader Priority, (again excluding those routers who have declared themselves Sub-Area Leader), and again use the Router ID to break ties.
- (3) Calculate the new Sub-Area Leader for the sub-area as follows. If one or more of the routers have declared themselves Sub-Area Leader (i.e., they are currently listing themselves as Sub-Area Leader in their Hello Packets) the one having highest Sub-Area Leader Priority is declared to be Sub-Area Leader. In case of a tie, the one having the highest Router ID is chosen. If no routers have declared themselves Sub-Area Leader, assign the Sub-Area Leader to be the same as the newly elected Backup Sub-Area Leader.
- (4) If Router X is now newly the Sub-Area Leader or newly the Backup Sub-Area Leader, or is now no longer the Sub-Area Leader or no longer the Backup Sub-Area Leader, repeat steps 2 and 3, and then proceed to step 5. For example, if Router X is now the Sub-Area Leader, when step 2 is repeated X will no longer be eligible for Backup Sub-Area Leader election. Among other things, this will ensure that no router will declare itself both Backup Sub-Area Leader and Sub-Area Leader.[5]

- (5) As a result of these calculations, the router itself may now be Sub-Area Leader or Backup Sub-Area Leader. The router's interface state should be set accordingly. If the router itself is now Sub-Area Leader, the new interface state is DR. If the router itself is now Backup Sub-Area Leader, the new interface state is Backup. Otherwise, the new interface state is DR Other.
- (6) If the attached sub-area is an NBMA network, and the router itself has just become either Sub-Area Leader or Backup Sub-Area Leader, it must start sending Hello Packets to those neighbors that are not eligible to become Sub-Area Leader. This is done by invoking the neighbor event Start for each neighbor having a Sub-Area Leader Priority of 0.
- (7) If the above calculations have caused the identity of either the Sub-Area Leader or Backup Sub-Area Leader to change, the set of adjacencies associated with this interface will need to be modified. Some adjacencies may need to be formed, and others may need to be broken. To accomplish this, invoke the event AdjOK? on all neighbors whose state is at least 2-Way. This will cause their eligibility for adjacency to be reexamined.

The reason behind the election algorithm's complexity is the desire for an orderly transition from Backup Sub-Area Leader to Sub-Area Leader, when the current Sub-Area Leader fails. This orderly transition is ensured through the introduction of hysteresis: no new Backup Sub-Area Leader can be chosen until the old Backup accepts its new Sub-Area Leader responsibilities.

The above procedure may elect the same router to be both Sub-Area Leader and Backup Sub-Area Leader, although that router will never be the calculating router (Router X) itself. The elected Sub-Area Leader may not be the router having the highest Sub-Area Leader Priority, nor will the Backup Sub-Area Leader necessarily have the second highest Sub-Area Leader Priority. If Router X is not itself eligible to become Sub-Area Leader, it is possible that neither a Backup Sub-Area Leader nor a Sub-Area Leader will be selected in the above procedure. Note also that if Router X is the only attached router that is eligible to become Sub-Area Leader, it will select itself as Sub-Area Leader and there will be no Backup Sub-Area Leader for the sub-area.

6. Inside/Outside Link Negotiation

A new bit is defined in the OSPF Options field, carried in OSPF Hello packets, that indicates on a link-by-link basis whether the link is an inside link or an outside link.

```

+-----+
| I/O | 0 | DC | EA | N/P | MC | E | * |
+-----+

```

When clear (0), the I/O bit indicates that the link is an inside link. When set (1), the I/O bit indicates that the link is an outside link.

When receiving Hello packets, Hellos in both directions must indicate that the link is an outside link in order for the link to be considered an outside link. If either or both routers believe that the link is an inside link, then both routers must treat the link as an inside link.

7. Handling of Arbitrary Opaque LSAs

Type 10 Opaque LSAs normally have area-local scope. That is, they are flooded throughout an entire area but not beyond. This proposal modifies this flooding behavior by restricting Type 10 Opaque LSAs to their local sub-area unless they are being originated by a SABR. Type 10 Opaque LSAs that are originated by a SABR are flooded throughout the local area.

8. Handling of Explicit Routes for Traffic Engineering

When a SABR receives (on an outside link) an RSVP-MPLS or CR-LDP setup request containing an explicit route object, the SABR must examine the explicit route object. The explicit route object will generally contain a route through the Sub-Area Leader that then continues on to a network within the sub-area or proceeds to another SABR and exits the sub-area. The ingress SABR must replace this path (which was based on abstract topology information) with a real path through the sub-area. The ingress SABR must compute the real path using appropriate traffic engineering information.

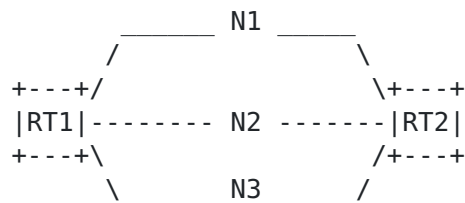
9. Open Issues

The following issues are for further study:

- o The case of a sub-area boundary that cuts between a router and a multiaccess network (e.g. an ethernet) is for further study.
- o Should probably use high metrics for links in abstract sub-areas. This helps encourage the use of intra-sub-area routing over inter-sub-area routing.

- o Abstract Topology LSAs must be translated into normal LSAs so that normal (non-upgraded) routers can learn the abstract topology. This could happen in either of two ways:
 1. A SABR could perform the translation before sending the LSA to a neighbor SABR in another sub-area. (I.e. when going from an inside link to an outside link.)
 2. A SABR could perform the translation when receiving Abstract Topology LSAs from other sub-areas, before injecting them into the local sub-area. (I.e. when going from an outside link to an inside link.)
- o It's not clear exactly how to best represent the abstract topology. In the abstract topology diagram above, a real SABR is elected Sub-Area Leader and all networks and other SABRs within a sub-area are represented as being connected to the Sub-Area Leader.

Another approach would be to show all SABRs, and show links to all networks from all SABRs. For example:



This may become unwieldy as the number of networks and SABRs increases.

Another approach involves the creation of a virtual router in place of the Sub-Area Leader. But this opens issues such as, how do you select an IP address for the virtual router? How do you ensure that all SABRs agree (an election process)?

Yet another approach involves creation of a virtual network to which all SABRs within a sub-area connect.

- o Thought must be given to how sub-areas interact with normal OSPF ABRs and ASBRs.
- o Thought must be given to how default routing should work among sub-areas. It may be possible to make it work for hop-by-hop forwarding, or it may be necessary to introduce MPLS LSP tunnels between sub-areas and use these tunnels for default forwarding.
- o If a dynamic SABR election process is used, is a

Hello/Keepalive mechanism needed between SABRs to detect failure of the Sub-Area Leader?

10. Security Considerations

This proposal introduces no new security concerns for OSPF. The existing OSPF authentication mechanisms should be used to prevent attacks on the protocol.

11. References

- [1] Moy, J. [RFC 2328](#) OSPF Version 2. April 1998. (Also STD0054)
(Status: STANDARD)

12. Authors' Addresses

Walt Wimer
FORE Systems, Inc.
1000 FORE Drive
Warrendale, PA 15086-7502

Phone: 724-742-7324
EMail: wwimer@fore.com

John Drake
FORE Systems, Inc.
1000 FORE Drive
Warrendale, PA 15086-7502

EMail: drake@fore.com