Similar Yet Different: Protocol Design Choices in IS-IS and OSPF

Yu Zhang
Harbin Institute of Technology, China
yuzhang@hit.edu.cn

Lan Wang
University of Memphis
lanwang@memphis.edu

Alexander Afanasyev
Florida International University
aa@cs.fiu.edu

Lixia Zhang
University of California, Los Angeles
lixia@cs.ucla.edu

ABSTRACT

This paper summarizes our comparative study on the design choices of two classic link-state routing protocols: IS-IS and OSPF. Although both protocols are based on the same algorithmic foundations for computing best paths across large networks, they made different choices at various aspects in the protocol designs. We selected ten major design differences between the two to understand the reasoning behind their choices. We hope that this comparative study helps shed new light on the design space of link-state routing protocols and prove useful to future routing protocol design efforts.

CCS CONCEPTS

• **Networks → Network protocol design**; **Routing protocols**;

KEYWORDS

Link-state routing, OSPF, IS-IS

ACM Reference Format:
Yu Zhang, Lan Wang, Alexander Afanasyev, and Lixia Zhang. 2017. Similar Yet Different: Protocol Design Choices in IS-IS and OSPF. In *Proceedings of The 13th Asian Internet Engineering Conference (AINTEC'17).* ACM, New York, NY, USA, Article 4, 7 pages. https://doi.org/10.475/123_4

1 INTRODUCTION

The intra-domain routing in the Internet today is mainly supported by two link-state routing protocols: Intermediate System to Intermediate System (IS-IS) [1, 6] and Open Shortest Path First (OSPF) [4, 9, 10]. Both protocols use Dijkstra's shortest path computation algorithm, and share the same root of the early ARPRnet's Shortest Path First (SPF) routing protocol developed by McQuillan et al. [8] in 1970's. Both protocols have been successfully developed and widely deployed, and continue to evolve to support new features demanded by the ever evolving global Internet, such as IPv6, MPLS and multi-topology routing.

Although IS-IS and OSPF were designed with similar requirements and algorithms, they differ in many aspects in their designs and operations. For readers interested in the history and details of the differences between IS-IS and OSPF, our recommendation is to look into Perlman's early article on the subject [11], Bhatia et al.'s discussions [2], IESG's recommendation regarding the choice of IGP for the IP Internet [7], and Doyle's book [5]. Those references suggest that the dissimilarity of the two protocols' designs and the differences between their supported functions could be attributed to a few factors. First, the two protocols originated from different communities, one was designed to support OSI protocols and the other to support IP. Second, various protocol design decisions were made according to different engineering tradeoffs as perceived by the designers [11]. Although no single answer may explain the differences comprehensively, the second one has been widely accepted as the major factor by the community.

The purpose of this paper is neither to re-evaluate the merits of the two protocols nor to give yet another recommendation on how to choose in between. Rather, our goal is to assess the the lessons learned from the past and use them as one of the inputs to inform the design of a new link-state routing protocols for a non-IP Internet architecture, the Named Data Networking (NDN) [12]. Thus, we are most interested in gaining insights into the reasoning behind IS-IS and OSFP's different design choices. Different from previous comparative studies, we evaluate the two designs by putting ourselves in the *designers'* shoes, rather than those of the *users* (i.e., network operators). In conducting the comparison, we do not retell the well-known technical differences, but use IS-IS and OSPF as two landmarks in our exploration of the design space. Hence, we compare the two protocols by keeping a question in mind: *How should one design a new link-state routing protocol from scratch?*

In the rest of this paper, we first give an overview of link-state routing protocols and identify ten design questions. We then compare the design choices based on those questions. We hope that the outcome of our study would be generally interesting to people in network protocol designs, especially in routing protocol design.

2 LINK-STATE ROUTING OVERVIEW

This section briefly describes the basic functionality of a link-state routing protocol, and presents ten design questions that must be answered in achieving the functionality. Since IS-IS and OSPF use somewhat different terminologies, instead of translating between the two, we keep the description simple by using the terminology from OSPF and TCP/IP, together with the definition of a few general terms.

Generally speaking, all link-state routing protocols use Dijkstra's shortest path computation algorithm to compute the best path(s) for each of all the destinations. Since the Dijkstra algorithm computes over a given topology map made of nodes and links, the

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
AINTEC'17, November 20–22, 2017, Bangkok, Thailand.
© 2017 Association for Computing Machinery.
ACM ISBN 978-1-4503-5551-3/17/11…$15.00
https://doi.org/10.475/123_4

core function of a link-state routing protocol is to disseminate the network topology to all routers in the same network. Once a router obtains the topology map, it then independently runs Dijkstra's algorithm on the topology to find the best path(s) to each of all the destinations. Therefore, the main task of the protocol is to describe, discover, and disseminate the topological information of the network.

Link-state routing protocols work in relatively stable wire-connected networks consisting of routers, hosts, and links. For simplicity, this paper only considers two general types of communication media: point-to-point links and broadcast networks, such as Ethernet. The topological connectivity information, called *link-state* (LS) data, has two categories: (a) *adjacency*, i.e., how routers are connected; and (b) *reachability*, i.e., which destinations can be reached directly via a given router. The costs of links between neighboring routers, and between routers and their directly-reachable destinations are represented by *metrics*. To represent LS data, the following questions must be answered:

(1) How to uniquely identify a router?
(2) How to organize and identify LS data?
(3) How to format LS data in packets?

Discovering adjacency information is the job of link-state routing protocols, while reachability information is obtained by some other means, such as static configurations or routing redistributions. Each router contributes its own piece of LS data to compose the global knowledge on the topology, namely Link-State DataBase (LSDB). A link-state routing protocol is responsible for ensuring that all the routers in the same network maintain the same topology map for their own best route calculations. To achieve this goal, the design needs to answer the following questions:

(4) Which types of messages are in protocols?
(5) How to establish adjacencies between neighbors?
(6) How to synchronize data between neighboring routers over a point-to-point link?

In a broadcast network with n adjacent routers, there can be up to $O(n^2)$ adjacent router pairs, which may lead to inefficiency in adjacency establishment/detection, SPF calculation, and LSDB synchronization. The main concern in efficiency is about the number of messages. For example, because every update message is supposed to traverse every link once, $O(n^2)$ adjacencies mean $O(n^2)$ transmissions. Therefore, the design questions are as follows.

(7) How to establish adjacencies in a broadcast network?
(8) How to synchronize data in a broadcast network?

The final step for the routing protocol is to find a path to a given destination. As long as a router's own LSDB is up to date, it can utilize SPF calculations to independently calculate the best path to the destination, and then inject the results into its forwarding table, which maps destinations to neighboring routers via the address pair of local and neighbor's interfaces. Before the SPF calculation, we need to answer a question:

(9) How to abstract a graph from an LSDB?

To make a routing system scale to large networks, a straightforward method is to split the large network into smaller networks, and then organize them into a hierarchy. The question here is as follows:

(10) How to design a hierarchical routing scheme?

In addition, there are other requirements in routing protocol design, such as stability against frequent changes, tolerance to faults, and data authenticity. This paper does not cover them, but focuses on the major differences between IS-IS and OSPF.

3 DESIGN COMPARISON

In each subsection below, we address one question in Section 2. We first compare the two protocols, and then explore the design space by analyzing the differences and proposing some alternatives. As the details may vary in implementations, our comparison is based only on the specifications of Integrated IS-IS [1, 3] and OSPFv2 [10]. Some differences between OSPFv2 and OSPFv3 [4] are also discussed.

3.1 Router Identifier

Each router should be uniquely identified by a Router ID (RID). The key question is where the uniqueness comes from. IS-IS directly uses the address of the router (OSI NSAP address) assigned by the operator as its RID. In practice, the RID may be one of the router's MAC addresses or IP addresses, if available.

In OSPF, a router's RID is the highest/lowest IP address on its active logical (loopback) interfaces and physical interfaces. In OSPFv2, to identify a neighbor, a router uses the neighbor's RID for a point-to-point link, but uses the neighbor's interface IP address for a broadcast network. This inconsistency is addressed in OSPFv3 where the RID is used in both cases. Binding the RID with an IP address might bring inconvenience when the IP address is changed.

The uniqueness of RID can be obtained in either a *direct* or an *indirect* way. In the direct approach, all RIDs are directly assigned by a center or manually configured by an operator. In the indirect approach, each router independently derives its own RID from some of its own properties which may be obtained in some centralized ways. The indirect approach such as using MAC addresses can enable plug-and-play, as every MAC address assigned to only one manufacturer of networking devices. RIDs may carry some semantics besides uniqueness.

3.2 Link-State Data Organization

LS data need to be organized and identified to enable orderly dissemination and update. Let an *LS Unit* (LSU) be a minimal piece of LS data which can be individually sent and updated. LSU is called LS PDU (LSP) in IS-IS or LS Advertisements (LSA) in OSPF.

In IS-IS, all LS data advertised by the same router is simply put in a single LSU identified by the advertising router's RID. Therefore, it is very likely that an LSU may be bigger than the MTU of the underlying MAC layer. IS-IS's solution is to split an LSU into a set of fragments, each of which is identified by the RID and a number.

In OSPF, an LSU is identified by three values: *LS-Type*, *LS-ID* and the advertising router's RID. There are eleven LS-Types including five basic types, each of which has it own specific rule to determine the value of LS-ID. Multiple LSUs can also be included in a single packet. Although the granularity is finer, an LSU can also be big, such as LS-Type 1 (Router-LSA), which is handled by IP fragmentation.

IS-IS's one-dimensional structure is simple, but so coarse that any small change will trigger an update of whole LSU. OSPF's organization is three-dimensional and finer than that of IS-IS, which makes LS update efficient, but requires more overhead to represent the current state of LSDB. An alternative is a two-dimensional structure, where an LSU is identified by "RID+X". This "X" could be a type ID which identifies a type of data. The granularity of "X" should be determined according to engineering tradeoffs, such as the one between the total number of LSUs and the size of LSUs.

3.3 Link-State Data Packet Format

An LSU consists of a header part and a content part in a packet. Both protocols' headers include the advertising router's RID, *Sequence Number*, *Age* and *Checksum*.

The Sequence Number is used to determine whether an LSU is older than another LSU from the same router. Both protocols use an incremental sequence number, which means the LSU with a larger sequence number is newer. Once the Sequence Number reaches a maximum value, this LSU should be completely purged by setting the Age of LSU zero/maximum and re-flooding it (see next paragraph) through the whole network before the next update.

Both protocols are soft-state, which means LS data will be invalid unless refreshed. To implement the soft state of LSU, as time goes by, the value of Age will decrease downward from a positive number to zero in IS-IS (so the field is actually called Remaining Lifetime), or will increase upward from zero to a maximum in OSPF. The count-down method in IS-IS needs two values: zero and age, while OSPF needs one more: maximum. The Age field also serves other purposes, e.g., removing looped LSUs (similar to TTLs), and quickly purging an old LSU before its age expires by setting the age to the maximum in OSPF or zero in IS-IS.

In IS-IS, the content of LS data is organized in a Type-Length-Value (TLV) format. Adjacency and reachability information is put in two different types of TLVs, *IS-Neighbor* and *IP-Internal-Reachability*, respectively. TLVs can be nested. A router will forward an LSU without any change even if some types of TLVs are not understandable to the router.

In OSPF, each LS-Type has a predefined field format. The LS-ID field in the header is actually also a part of content. For example, the LS-ID of Type 1 (Router-LSA) is the advertising router's RID; the LS-ID of Type 2 (Network-LSA) is the the designated router's interface IP address. For a router, a single LSU of Type 1 (Router-LSA) contains both its adjacency information on point-to-point neighbors and its reachability information on the directly connected stub IP subnets. The extensibility for new features is provided by *Opaque* LSAs, which will be propagated by routers even if the routers do not understand the inside content.

In general, the header part should be minimized for efficiency. An LSU header should include the RID, the Sequence Number and the Age at least, while the function of Checksum may be replaced by other methods for data integrity. For the content part, the TLV format provides extensibility, compatibility and flexibility, while in OSPF the static field format is more hardware-friendly, and the extensibility is provided by the *Opaque* mechanism.

3.4 Message Types

Both IS-IS and OSPF use various types of messages to implement five functions. OSPF assigns each function to a single message type, while IS-IS adopts four types for those five functions.

In OSPF, there are five types of messages: 1) the Hello message for the adjacency establishment; 2) the Update message containing LSUs; 3) the Request message requesting for LSUs from other routers; 4) the Acknowledgement (ACK) message acknowledging the receipt of LSUs; 5) the Database Description (DBDS) message describing a router's LSDB. Each Request, ACK or DBDS message, describes a set of LSUs by listing the headers of LSUs. A DBDS message describes all LSUs in the LSDB and is only used for the initial synchronization.

In IS-IS, there are four types of messages: the Hello message, the Update message, and the other two types are Complete Sequence Number PDU (CSNP) and Partial Sequence Number PDU (PSNP), both of which describe a set of LSUs by listing its ID, Sequence Number, Age (Remaining Lifetime) and Checksum for each LSU. The CSNP covers all LSUs in the LSDB, while the PSNP covers a subset. The CSNP functions as the DBDS message and an implicit ACK message in a broadcast network. The PSNP functions as the Request message in a broadcast network and an explicit ACK message in a point-to-point link.

In checking whether those types of messages are indispensable, the obvious observation is that Hello and Update messages are necessary. For the rest three types, the design heavily depends on the synchronization method. Theoretically, Request and ACK messages (also PSNPs) can be replaced by periodically sending DBDS (CSNP) messages and pushing Update messages. The details will be discussed in Section 3.6 and 3.8.

3.5 Adjacency in Point-to-Point Link

An adjacency between two neighboring routers means the bidirectional forwarding capability (BFC) between them. Both protocols adopt 3-way handshaking to discover neighbors and check the BFC. The 3-way handshaking process is done if both routers saw its own RID appearing in the other's Hello message. For example, given two routers R1 and R2, a handshaking process will be as follows: 1) R1 sends out a Hello message without any neighbor's RID inside; 2) R2 replies with a Hello message with R1' RID; and 3) R1 sends out a Hello message with R2's RID.

The distinction between two protocols is on the timing of advertising adjacency. In IS-IS, two routers advertise their adjacency to the rest of the network once the 3-way handshaking is completed, while in OSPF, after handshaking, the adjacency will not be advertised until two routers's LSDBs are synchronized.

IS-IS separates the adjacency establishment from the synchronization. After advertising the adjacency, two routers will immediately synchronize their LSDBs. This separation simplifies the process of adjacency establishment. However, there will be a time period during which two routers with different LSDBs can exchange traffic, which may lead to loops or black holes.

In OSPF, the adjacency establishment includes the initial synchronization of LSDBs. In implementation, a router runs an individual state machine for each neighbor, which traces the whole process of 3-way handshaking, the designated router election and the LSDB

Table 1: Message Exchanges in P2P Synchronization. The rows in the sender's column are organized according to functions, and the rows in the receiver's column are the corresponding responses.

Sender		Receiver	
OSPF	**IS-IS**	**OSPF**	**IS-IS**
DBDS	CSNP	Request	Update
Request	PSNP	Update	
Update		ACK	PSNP, Update

synchronization. This method prioritizes the convergence of LSDBs over the simplicity of implementation.

To take advantage of both approaches, one may use an *advertising-after-synchronizing* method, in which the adjacency establishment is separated from the synchronization as in IS-IS, but the advertisement of adjacency is postponed until the initial synchronization is completed as in OSPF.

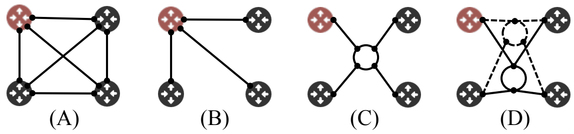
3.6 Sync over Point-to-Point Link

For two routers connected by a point-to-point (p2p) link, the two protocols adopt similar methods for synchronization, except that IS-IS performs the initial synchronization by pushing data, and may have periodic resynchronization.

In OSPF, two routers conduct the initial synchronization during their adjacency establishment by starting with the exchange of DBDS messages. Then each router will request newer or unknown LSUs and reply to the other's requests. After that, upon detecting any change in its own LSDB, a router will directly send updates to its neighbor and wait for acknowledgments. The unacknowledged updates will be retransmitted after a time interval. As long as all updates have been delivered successfully, two LSDBs are synchronized.

In IS-IS, once an adjacency is established, both routers must send its own CSNP, the equivalent of DBDS, to the other router. By comparing the received CSNP against its own LSDB, a router can know which LSU is stale or missing in the other's LSDB, and then pushes those LSUs to the other. So it is not necessary for a router to actively request what it is missing. After the initial synchronization, updates will be directly flooded. Upon receiving updates, a receiver router replies with a PSNP as the acknowledgment to the sender, except in the case where the receiver's LSU is newer than that sent by the sender. In that case, instead of a PSNP, the newer LSU will be sent back to the sender by the receiver. Optionally, a router may periodically send CSNPs to restart synchronization as an enhancement of reliability.

Table 1 summarizes the message exchanges between neighboring routers in both protocols. There are two options for transmitting differences in the initial synchronization: push in IS-IS and pull in OSPF. Dynamic changes can be propagated by triggered reliable updates. One inspiration from IS-IS is that the synchronization can also be done by periodically sending CSNPs as implicit requests. It will be practical if the cost of representing the state of LSDB is reduced.



	Virtual Adjacency	Synchronization Relationship	Graph for SPF Calculation
IS-IS	C	A/B	D
OSPF	B	B	C

Figure 1: Structures in Broadcast Network. Red router – the designated router; solid circle – pseudo-node; dash circle – IP subnet.

3.7 Adjacency in Broadcast Network

To improve the efficiency in a broadcast network, both IS-IS and OSPF try to make the network sparse with a *Designated Router* (DR), but in slightly different ways. During adjacency establishment, a DR is elected among the routers according to the preconfigured priority of each router.

The election is preemptive in IS-IS, which means a new router with higher priority can take the status of DR from the old DR. In OSPF, the DR status is sticky so that the DR will not be changed unless it crashes. So the DR election is deterministic in IS-IS and non-deterministic in OSPF. In OSPF, a backup DR is also elected to enable the graceful switchover when the DR crashes. In IS-IS, any router could be a backup DR, as the exclusive information maintained by the DR in IS-IS is much less than that in OSPF, which may be the reason why the election in IS-IS can be preemptive.

To sparsify adjacencies, both protocols construct a virtual star graph. In IS-IS, the center of star is a pseudo-node representing the local network, as shown in Figure 1(C). Speaking on behalf of this pseudo-node, the DR advertises the list of routers in the network. The RID of pseudo-node is assigned by the DR by concatenating the DR's RID and a number. In OSPF, the center of star is the DR adjacent to all other routers, as shown in Figure 1(B). In both protocols, all routers know the existence of each other, but only advertise their adjacency to one node, i.e., the pseudo-node in IS-IS or the DR in OSPF.

The choice of DR election method depends on whether a DR is heavy-duty and maintains much information. If so, then the DR status should be sticky; otherwise, the deterministic property is preferred. To sparsify adjacencies, IS-IS's virtual graph is intuitive, as it is consistent with the physical topology. And later we will see that OSPF also adopts this graph as the one for the SPF calculation.

The synchronization in a broadcast network and the graph for SPF calculation will be discussed in Section 3.8 and 3.9, respectively. The comparisons are summarized in Figure 1.

3.8 Sync in Broadcast Network

To sparsify synchronization relationships, both protocols utilize DRs again. In OSPF, the synchronization relationship is the same as the adjacencies. All routers synchronize their LSDBs only with the DR's in the same manner as in a p2p link, shown in Figure 1(B). Any update from a non-DR router will be first sent to the DR, and then be forwarded to the others by the DR.

In IS-IS, the synchronization relationship is more distributed like a hybrid of Figure 1(A) and Figure 1(B). Instead of one-to-one synchronization, routers broadcast updates directly to each other, as in Figure 1(A). However, there is no explicit acknowledgment from receivers—only the DR periodically broadcasts CSNPs as implicit acknowledgments, as in Figure 1(B). Upon receipt of a CSNP from the DR, if there is any difference, the receiver prepares a PSNP request or some LSU updates, and then broadcasts them after a random time interval, during which, if receiving the same PSNP or updates from others, the router will suppress its own copy. Only the DR will reply to a PSNP request.

The DR in IS-IS is like a 'spokesman' speaking for the pseudo-node and sending CSNPs on behalf of non-DR routers, whereas the DR in OSPF is like a 'leader' leading the adjacency establishment and the LSDB synchronization among non-DR routers. Flooding updates in OSPF needs one more hop than that in IS-IS. However, OSPF does not need periodic CSNP, and uses the same synchronization method in both types of media. A question is whether we can design a synchronization method that can directly flood updates in both types of media with less cost than periodic CSNPs.

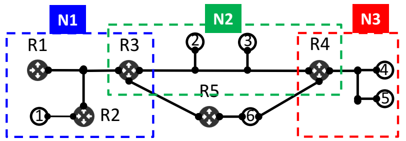
3.9 Graph for SPF Calculation

For SPF calculation, both protocols abstract a weighted graph from the network topology by mapping routers and destinations to vertices. The calculation results should be the same from both protocols. The main distinction is on whether the adjacency and the reachability are separately represented in the graph.

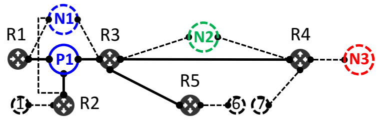
For Integrated IS-IS and OSPF, the reachability is about the sets of interface addresses, i.e., IP subnets. In the example shown in Figure 2(A), each media is an IP subnet. For simplicity, only the three IP subnets with more than two interfaces, i.e., N1, N2 and N3, are explicitly labeled, and other subnets can be identified by their two ends. Only N1 is configured as a broadcast network for routers. If two interfaces on the same host are attached to two different IP subnets, the host will be mapped to two distinct destinations, e.g., Host 6 in Figure 2(A) is mapped to Host 6 and 7 in Figure 2(B) and Figure 2(C).

IS-IS considers adjacency and reachability separately. The topology graph consists of two parts: the core part, representing the adjacency, is composed of routers; and the margin part, representing the reachability, is composed of destinations. On a path, only routers are intermediate vertices, whereas IP subnets are end vertices, which means changing IP address will not trigger an SPF recalculation. Figure 2(b) shows IS-IS's graph, where the adjacencies between routers are represented by solid lines, while the reachability is represented by dash lines.

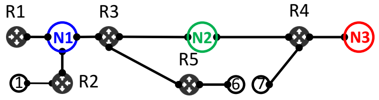
OSPF considers adjacency and reachability as a whole. Basically there is one-to-one mapping between links in the topology and edges in the graph. Two routers connected via an IP subnet in a broadcast network are not connected directly, but indirectly via a vertex denoting that transit subnet. Figure 2(c) shows OSPF's graph, where R3 and R4 are connected via N2. So an IP subnet may be an intermedia vertex on a path, which means changing the IP address of subnet may trigger an SPF recalculation.



(A) Physical Topology



(B) Graph in IS-IS



(C) Graph in OSPF

Figure 2: A networking system, where routers (black disc), non-router entities (small circle), and IP subnets (big circle) are connected via media (solid/dash line) on interfaces (black dot). 3 IP subnets, N1, N2 and N3. N1 is a broadcast network for routers. P1 is a pseudo-node of a broadcast network in IS-IS.

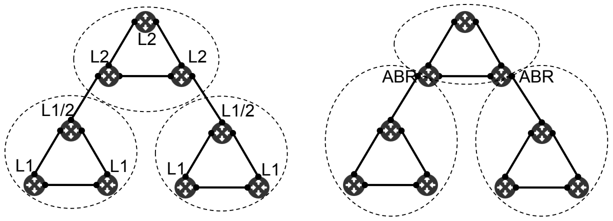
To sparsify the graph for SPF calculation in a broadcast network, both protocols abstract a star graph with a pseudo-node representing the local network as the center of the star. N1 in Figure 2 is an example of broadcast network. In IS-IS, as shown in Figure 1(D), all routers have both the adjacency to the pseudo-node advertised by the DR and the reachability to the subnet advertised by all routers. In OSPF, as shown in Figure 1(C), the graph is derived from the membership information of network (Network-LSA) advertised by the DR. The actual details in OSPF are more complex than the above description, please see Section 2.1 in [10].

We suggest that adjacency should be separated from reachability in the graph. A benefit is that a change on the reachability will not necessarily trigger an SPF recalculation. Moreover, those two types of information should be separately considered in every protocol module.

3.10 Hierarchical Routing

For scalability, both protocols support two-level hierarchical routing by splitting the network into a single backbone area and multiple non-backbone areas. To avoid loops, usually two non-backbone areas can only communicate via the backbone area. Figure 3 illustrates the concept of hierarchical routing in both protocols.

In IS-IS, as shown in Figure 3(A), each router normally belongs to a single area, and the boundaries between areas cross links. There are two types of roles for routers: L2 routers at the higher level compose the backbone, while L1 routers at the lower level can only communicate with other L1 routers within the same area. All L2 routers must be interconnected via L2 routers, even from different areas. A router configured to play both roles at the same time is an L1/2 router, which is responsible for distributing the reachability information and forwarding traffic between two levels. Therefore, each non-backbone area must have at least one L1/2 router.



(A) IS-IS Areas (B) OSPF Areas

Figure 3: Hierarchical routing among areas.

In OSPF, as shown Figure 3(B), an area is a group of IP subnets along with attached routers, and the boundaries between areas cross some routers, called Area Border Routers (ABRs), which are similar to L1/2 routers in IS-IS. The backbone area at the higher level is a designated area (with ID 0.0.0.0), and all other areas at the lower level must be connected directly to the backbone area via ABRs.

Although the two protocols draw boundaries either across links or on routers, both protocols need each border router (L1/2 router in IS-IS, or ABR in OSPF) to maintain two LSDBs of two neighbor areas at the same time. In terms of the knowledge of LS data, boundaries cross routers in both protocols. An alternative design is to draw boundaries across links by making each router maintain only one LSDB for its own area. In this design, between areas, adjacency information is not shared, whereas reachability information is shared after summarized.

4 SUMMARY AND DISCUSSIONS

This section summarizes the lessons learned from IS-IS and OSPF, and possible choices in future design. Each following item corresponds to one problem in Section 2.

- **RID:** The uniqueness of RID can be obtained either in a direct way in a centralized manner or in an indirect way based on some routers' own properties; the semantics of RID may be utilized for purposes other than identification.
- **LSU structure:** The LS data can be organized in different dimensions; two-dimensional organization is an option to make the tradeoff between the total number of LSUs and the size of LSUs.
- **LSU format:** The packet of LSU should include RID, Sequence Number and count-down Age; the content can be encoded with TLV format or static field format with opaque options. The adjacency and reachability information should be clearly separated into different types of data.
- **Adjacency establishment:** Adjacency establishment consists of 3-way handshaking and an optional initial synchronization. An alternative choice is to postpone the adjacency advertisement until the initial synchronization is completed.
- **Message types:** Hello and Update messages are necessary; the design of other message types heavily depends on the specific synchronization method.
- **Sync in p2p link:** The synchronization of LSDBs can be done by an initial synchronization followed by reliable triggered updates; an alternative is the periodic synchronization if the cost of representing the current state of LSDB is cheap.
- **Adjacency in broadcast network:** The adjacencies in a broadcast network can be sparsified into a virtual star topology. The choice of DR election method, either sticky or preemptive, depends on the role of DR.
- **Sync in broadcast network:** The DR can play different roles in the synchronization; a future design would ideally unify the synchronization method for both types of media, directly flood updates, and have low cost.
- **Graph:** Adjacency information should be separated from reachability information in the graph for SPF calculation; a broadcast network is sparsified with a pseudo-node representing the network.
- **Hierarchical routing:** The network is separated into two tiers with a single backbone area at tier 1. Boundaries between areas are across routers in terms of the knowledge of LS data; an alternative is to draw boundaries across links by making each router maintain only the information for its own area.

An observation beyond the design choices in the above technical parts is a complement to the reason for diverse protocol designs: *design style*. Roughly speaking, it seems that the design choices of IS-IS largely reflects more considerations on logical abstraction, whereas those in OSPF reflect more pragmatic consideration. We found that the design choices on various issues in the same protocol mostly follow similar style, which we believe is not by chance, but reflects the coherence of the designers' view through the whole design process.

5 RELATED WORK

Given routing is a fundamental topic in the networking research area, it is somewhat surprising that the research community has paid little attention to the dual existence of IS-IS and OSPF, two routing protocols that share the same origins and the same functional goals, but exhibit so many differences in their design choices. To the best of our knowledge, up to now, Perlman's comparison [11], one of the IS-IS designers, represents the only academic paper on the topic.

Perlman pointed out that "OSPF has favored optimizing routing, whereas IS-IS has favored minimizing storage and computing in level-1 routers", an observation with which we concur. At the same time, we believe that our comparative examination of the two protocols covers much broader issues in the design space. We also largely learned the differences between IS-IS and OSPF protocols from the existing works represented by Bhatia et al.'s informative Internet draft [2] and Doyle's book [5]. However, the focus of our effort is not on the comparison itself, but to gain insight that can help guide new link-state routing protocol designs.

6 CONCLUSIONS

IS-IS and OSPF protocols fulfill the same network routing functions through different designs. This paper discussed their differences on ten design questions centered around how to name routers, encoding routing information, synchronizing information among routers over different types of networks, abstracting the topology to compute shortest paths, and scaling the routing protocol. Their different approaches lead to different tradeoffs in message transmission, computing, storage, protocol complexity, and network stability.

We believe that routing protocol designs, and network protocol designs in general, are still in the stage of being a design art, rather than following well-established principles and guidelines. Thus learning from successful, as well as unsuccessful, protocol designs offers an effective way to evolve the field towards maturity. We hope that this paper makes a contribution toward that ultimate goal. Moreover, we believe that the future Internet architecture should be designed to make it easier to solve the above common protocol design issues, e.g., naming and efficient information dissemination, and that designers of future routing protocols should take advantage of the built-in features of such a network architecture.

7 ACKNOWLEDGMENTS

This work is partially supported by the National Key Research and Development Program of China (2016YFB0801303), and by the US National Science Foundation under the awards CNS-1345318 and CNS-1344495.

REFERENCES

[1] 1990. Information Technology, Telecommunications and Information Exchange Between Systems, Intermediate System-to-Intermediate System Routing Information Exchange Protocol for Use in Conjunction With IS0 8473. *ISO 70589* (1990).

[2] Manav Bhatia, Vishwas Manral, and Yasuhiro Ohara. 2006. IS-IS and OSPF Difference Discussions, draft-bhatia-manral-diff-isis-ospf-01. *IETF Internet Draft* (January 2006).

[3] R. Callon. 1990. Use of OSI IS-IS for Routing in TCP/IP and Dual Environments. *RFC 1195* (1990).

[4] R. Coltun, D. Ferguson, and J. Moy. 1999. OSPF for IPv6. *RFC 2740* (December 1999).

[5] Jeff Doyle. 2005. *OSPF and IS-IS: Choosing an IGP for Large-Scale Networks*. Addison-Wesley Professional.

[6] Hannes Gredler and Walter Goralski. 2005. *The Complete IS-IS Routing Protocol*. Springer.

[7] Phil Gross. 1992. Choosing a Common IGP for the IP Internet (The IESG's Recommendation to the IAB). *RFC 1371* (October 1992).

[8] J.M. McQuillan, I. Richer, and E. Rosen. 1980. The New Routing Algorithm for the ARPANET. *IEEE Transactions on Communications* 28, 5 (1980), 711–719.

[9] John T. Moy. 1998. *OSPF: Anatomy of an Internet Routing Protocol*. Addison-Wesley Professional.

[10] John T. Moy. 1998. OSPF Version 2. *RFC 2328* (April 1998).

[11] Radia Perlman. 1991. A Comparison between Two Routing Protocols: OSPF and IS-IS. *IEEE Network* 5, 5 (September 1991), 18–24.

[12] Lixia Zhang, Alexander Afanasyev, Jeffrey Burke, Van Jacobson, kc claffy, Patrick Crowley, Christos Papadopoulos, Lan Wang, and Beichuan Zhang. 2014. Named Data Networking. *SIGCOMM Computer Communication Review* 44, 3 (July 2014), 66–73.