

OSPF and IS-IS Evolution

Jukka Honkola
Helsinki University of Technology
Jukka.Honkola@hut.fi

Abstract

Open Shortest Path First (OSPF) and Intermediate System to Intermediate System (IS-IS) are two most widely used link-state routing protocols. Both date from the late 1980s and share a common ancestry. The protocols have evolved considerably over their existence and there is still active development going on. I survey the development of the protocols, taking into account the changes made to OSPF after version 2 and changes related to IP routing in IS-IS. Convergence issues and traffic engineering extensions are also considered.

KEYWORDS: OSPF, IS-IS, IGP routing, IPv4, IPv6

1 Introduction

Open Shortest Path First (OSPF) and Intermediate System to Intermediate System (IS-IS) are both link state routing protocols designed to be used within a single autonomous system (AS). In this paper I will examine the changes made to the protocols and analyze the effect of the changes. Only changes affecting layer 3, the network layer, are considered as far as it is possible to separate layers 2 and 3. In the case of IS-IS, I will restrict myself to examine its use in IP routing.

The terms “router” and “intermediate system” both describe a node which forwards packets towards their destination. I use the term “router” when there is no need to separate the OSPF and IS-IS behavior. Otherwise, intermediate system (IS) refers to IS-IS and router to OSPF. Also, when discussing about the common aspects of the protocols, the terminology used is from OSPF whenever there are different terms for a common concept (e.g. LSA vs. LSP).

This paper consists of a short introduction of the two protocols in section 2. The introduction is a very broad view of the two protocols and does not go into details. A survey of changes made and proposed is presented in section 3. The changes and proposals have been classified into categories. In section 4 I examine which changes have been implemented and are in actual use.

2 OSPF and IS-IS

OSPF [24] and IS-IS [7] are routing protocols using link-state routing. OSPF is designed for routing IP and IS-IS was originally designed for routing ConnectionLess Network Protocol (CLNP) (ISO 8473). Extensions to IS-IS required for IP routing have been defined in [3].

IS-IS was based on a routing protocol for DEC Phase V, which the ISO took as basis for IS-IS. An early draft of IS-IS

was then taken as a basis for OSPF. Specifications for both protocols were published at the end of 1980s. A protocol based on an early draft of IS-IS was deployed in NSFnet starting in 1988. Both protocols have been deployed since the turn of the 1990s.[21]

The protocols closely resemble each other, OSPF having been developed on ideas present in IS-IS[21]. Therefore, I will describe the protocols together, and point out any differences between them.

IS-IS sends the routing information packets directly over the link layer. It does not encapsulate the packets in IP as does OSPF. Thus, in an AS that uses IS-IS there is non-IP network traffic. This might cause problems in some network configurations as mentioned in annex E of [3]. The example topology is one where there is a bridge/router connecting two Ethernet local area networks (LANs) with an IS-IS router on each. The bridge/router routes IP packets and acts as a bridge for all other protocols. The IS-IS routers exchange routing information using IS-IS packets, and therefore do not know about the existence of the bridge/router. When the routers start to forward IP packets to one another, the bridge/router might drop them because they are not addressed to its Ethernet address.

2.1 Basic functions

The routers keep a local copy of the network topology in a link state database. The topology is interpreted as a weighted graph where the nodes represent routers and the edges links. On both protocols, a broadcast LAN is modeled as a pseudonode which has edges to all routers connected to the LAN. Without the pseudonode, a broadcast LAN would have N^2 edges connecting N routers while the pseudonode cuts the number of edges to N . All routers flood information about the state of their neighborhood to other routers in the AS. The information is stored in Link State Advertisements (LSAs) in OSPF and in Link State PDUs (LSPs) in IS-IS. LSAs are encapsulated in Link State Update (LSU) or Database Description packets. Again, the broadcast LANs are an exception. A Designated Router (DR) is chosen for a LAN and acts as a pseudonode, that is, sends LSAs describing links to all routers connected to the LAN. The graph is used to calculate a shortest path tree by using Dijkstra’s shortest path first (SPF) algorithm. Routing tables are then constructed based on the tree.[7][24]

The routers are considered *neighbors* if they are directly connected. Neighboring routers can form an *adjacency*. Adjacencies are formed by sending hello packets to all neighbors. Routers attach information about their capabilities to the hello packets. The hello packets are sent periodically

over all interfaces on which the protocol is activated to determine the state of adjacent routers. Adjacent routers exchange routing information and keep their databases synchronized. Thus, adjacencies generate network traffic. In IS-IS, all neighboring routers even on broadcast LANs become adjacent. The adjacency in IS-IS does not entail as much network traffic as in OSPF.[7][24]

In order for the routing to work, the databases containing the network topology must remain identical. If the network topology changes, the routers noticing the change then flood new LSAs which causes new SPF calculations in all routers in an area. The database synchronization mechanisms are similar. In OSPF the contents of the database are sent in Database Description or LSU packets, containing the LSAs in the database. In IS-IS portions of database are sent in Complete Sequence Numbers PDUs (CSNP) and Partial Sequence Numbers PDUs (PSNP). The packets containing link state information are numbered to guarantee that only the most current information is used.[7][24]

2.2 Hierarchical Routing

Both protocols support the division of an AS into separate areas. There is a special backbone area connecting all normal areas. All inter-area traffic will travel across the backbone. The concept is broadly the same in both protocols, but the terminology differs.

Routers inside an area are called internal routers in OSPF and Level 1 (L1) and Level 1/2 (L1/L2) intermediate systems in IS-IS. Routers belonging to the backbone are called Area Border Routers (ABR) in OSPF and Level 2 (L2) intermediate systems in IS-IS. In OSPF the backbone is composed by all ABRs and possibly routers internal to the backbone. In IS-IS the Level 2 Subdomain contains all Level 2 capable ISs.[7][24]

Furthermore, in OSPF areas with only one ABR can be classified as *stub areas*. Stub areas do not receive information about AS external routes. This eases the memory and processing requirements for the routers inside the stub area.[7][24]

The IS-IS L1 areas are analogous to OSPF stub areas. They do not have information about other areas at all but rely on default routes for inter-area traffic. The routes may be suboptimal if there are more than one L1/L2 routers in a L1 area because the inter-area traffic is always routed to the nearest L2 router in an area.[7]

In OSPF, routers can belong to several different areas, typically to the backbone and one other area. In IS-IS, the routers are always inside just one area but some routers exchange traffic with both L1 and L2 routers. The L1/L2 routers reside in a L1 area. To summarize, in OSPF the area border lies on router and in IS-IS it lies on a link connecting two areas.[7][24]

The internal topology of an area is hidden from outside. Similarly, internal routers are only aware of the network topology in their own area. The backbone will route inter-area traffic to the ABR at the border of the correct area.[7][24]

OSPF backbone area is assigned an identifier of 0. The backbone has a connection to all other areas. The connec-

tion does not have to be a physical one, a virtual connection suffices.[24]

The IS-IS L2 subtopology only routes to the edge of an area. L2 routers have no knowledge about other areas. L1/L2 routers only know the topology of their own area in addition to the L2 topology. This is a different approach than in OSPF. A technique to insert some knowledge about the areas into L2 subtopology is presented in [18].

The IS-IS approach scales better while the OSPF approach gives sometimes shorter routes to destinations in areas. The differences are not very significant anymore.

Indiscriminate use of areas can cause problems especially when the network topology changes. For example, some prefixes can become unreachable even when a route exists to them if a link breaks inside of an area. Even without changes in topology a bad area division can result in sub-optimal routes to destinations inside an area.[39]

3 Changes and extensions

In this section I present changes and extensions made to the protocols. First there are descriptions of a few general issues, namely Multi-Protocol Label Switching, Bidirectional Forwarding Detection and convergence. After those, changes and extensions to the two protocols are described. On many occasions, both protocols have incorporated similar changes.

3.1 MPLS

A new packet forwarding technique, Multi-Protocol Label Switching (MPLS), associates a simple label to each packet. The label is typically determined by the destination address of the packet and/or the route by which it arrived to the network. Forwarding is then performed according to the label eliminating the need to look at the packet headers. All packets following the same path through the network are equivalent as far as the forwarding in the network is considered. The label thus describes the forwarding equivalence class (FEC) of a packet.[34]

There are no rules about the assignment of labels to the packets. Packets can be given different labels depending on their origin and the information is preserved in the label. Therefore the labels add some information about the packet that can not be deduced based on the headers alone. This extra information makes it possible to perform fairly detailed routing of packets. Also, the calculations need only be performed at the edge of the network as the packet enters the network. The path that the labelled packet takes is called a Label Switched Path (LSP, unfortunately the acronym is the same as the acronym for the IS-IS Link State PDU). A LSP is defined to be the sequence of hops that forwards the packet according to the label assigned to it by the router in the beginning of the path. The packet may have a stack of different labels.[34]

The label based forwarding is simpler than using the destination address stored in the packet header. This makes it possible to use simple switches inside the network and only have complex routers at the edges assigning labels.[34]

Both OSPF and IS-IS contain extensions to facilitate MPLS-based traffic engineering (TE). MPLS TE extensions

have also been implemented for both protocols.[15][16].

3.2 Bidirectional Forwarding Detection

Bidirectional Forwarding Detection (BFD) is a purpose built liveness testing protocol. Its aim is to quickly notice if a path between two network entities breaks down. Thus it can be used to test the state of the link between two routers.[13]

The normal mechanisms to test the liveness of a link notice a link failure in a matter of seconds, depending on the interval for sending periodic hello packets if there is no hardware failure detection on the link. For example, in OSPF the minimum interval for hello packets is 1s, and the link is considered down when three hello packets are lost. Also, the hello detection only tests the link from one forwarding engine to another.[13]

BFD is intended for use with any protocol independent of layer. In routers, it is appropriate to implement BFD in forwarding plane to keep it independent from the control functions. This enables the use of BFD even when doing graceful restart.[13]

BFD works by establishing a BFD session. The session is tied to the application using it and the application provides the necessary information for BFD to start a session (e.g. the address of the other party). BFD packets are encapsulated in a packet of the appropriate protocol for transmission. BFD can work either in asynchronous mode or in demand mode. In asynchronous mode BFD sends control packets periodically while in demand mode they are sent only when there is a reason to test the path. BFD also has an echo function, where BFD sends a stream of packets which are looped back to the sender. If several packets in a row are lost, the session is declared down. The use of echo function allows reduction of control traffic in both modes.[13]

BFD is useful in reducing the time needed to notice if a link fails and therefore also the time needed to establish new routes.[13]

3.3 Convergence

Convergence time is the time an IGP takes to fully adapt to changed situation. There are four aspects that influence convergence time: event detection, SPF processing, IGP advertisement (e.g. LSA flooding) and Forwarding Information Base (FIB) update. While the IGP is not converged packets may be lost. Furthermore, with today's high bandwidth networks even short failures result in many lost packets.[28]

Customers of Internet Service Providers (ISPs) typically use the packet loss and convergence time as indicators of the service quality.[28]

Both OSPF and IS-IS incorporate similar features to decrease convergence time. BFD can be used to shorten the time to notice a link failure (event detection). Incremental SPF calculation algorithms have been implemented at least in Cisco's and Juniper's routers (SPF processing).

The flooding of LSAs can be artificially delayed in order to avoid congesting the network with LSAs for example when there are route flaps. Removing the restrictions enables faster flooding of the LSAs reducing convergence time. Also, the LSAs are not necessarily sent forward until

the SPF calculation is performed. This is not such a problem if the incremental SPF algorithm is used, as it is significantly faster especially in larger topologies.[28]

The FIB update is an implementation issue and thus is outside the scope of this paper.

3.4 Network Traffic

Traffic Engineering

The TE extensions are very similar for both protocols. The extensions specify new type-length-value (TLV) fields for distributing extra information about links. There is also defined an extension to TLVs, namely sub-TLVs which are just TLVs inside a TLV. In OSPF, the TLV describing the attributes of a single link is called a Link-TLV, while in IS-IS it is an extended IS reachability TLV. The TLV has sub-TLVs for total bandwidth, reservable bandwidth, unreserved bandwidth, TE metric and administrative class. The bandwidth sub-TLVs describe the real bandwidth of the link. The TE metric is a metric set by the network administration and used for TE calculations. The administrative group can be used for example to divide the links to different classes according to their properties. The administrative classes can be used for example to constrain some packets to travel only through links belonging to some administrative class.[14][38]

In IS-IS the link metric is also increased from 6 bits in the original IS reachability TLV to 24 bits. IS-IS had originally four different classes of metrics, each 6 bits long. The space used by the four metrics (three of which were not used in practice) was combined to get a larger metric.[14][38]

The extra information is used to build a traffic engineering database which is in addition to the link state database. The TE database can then be utilized in various ways.[14][38]

One can, for example, perform constraint based source routing, where the source typically is the router through which the packet entered the network. Constrained SPF calculation is NP-hard or in some cases even unsolvable. Thus, the computational cost of CSPF can be prohibitive.[14][38]

Another possibility is to use the TE database to monitor the network state. The database typically contains information about the bandwidth of the links and unreserved bandwidth, for example.[14][38]

Yet another way of utilizing the extra information is to perform global traffic engineering. In global TE, one computer uses the TE database to compute routes for the whole network.[14][38]

Packet Priority

When network topology gets larger, the amount of control traffic increases as there is more potential for failures and more routers sending control data. If the network is already somewhat congested, events causing a lot of control traffic, such as link failures, add to the congestion. When the congestion gets bad enough, router control packets can be lost. This in turn can cause adjacencies to drop because of lost Hello packets adding to the congestion. Events of this sort can easily lead to a positive feedback loop where the amount of router control traffic significantly adds to the congestion problem.

A proposal to alleviate the problem is presented for OSPF in [9]. Routers should treat critical OSPF control packets, namely Link State Acknowledgement (LSAck) and Hello, at higher priority than the rest of packets. Alternatively, if the described approach can not be used, all control packets received over adjacencies are interpreted as attesting the liveness of an adjacent router. Naturally, all other control packets are still prioritized over normal traffic.[9]

The reasoning for the prioritizing is straightforward. If a LSAck does not arrive in time, the LSA will be resent. If Hello packets are lost or delayed, the link will eventually be declared down by adjacent routers generating a lot of LSA traffic. Furthermore, a received, changed LSA will trigger a SPF calculation, adding to the load of a router, making packet loss even more probable.

LSA Traffic Reduction

In OSPF, LSAs are refreshed every 30 minutes, causing unnecessary traffic in stable network topologies. Use of DoNotAge LSAs has been proposed to cut this traffic.[27] DoNotAge LSAs have been defined in “Extending OSPF to Support Demand Circuits”[23]. Normally LSAs have to be refreshed every 1800s, or they expire after 3600s. DoNotAge LSAs only have to be refreshed after a (configurable) forced flood interval. Flood reduction capable routers only issue DoNotAge LSAs and also set the DoNotAge bit on LSAs that they receive before flooding them.[27] Routers that do not recognize DoNotAge LSAs will prevent the use of DoNotAge LSAs. Also, routers which have originated DoNotAge LSAs should prematurely age them if a router not supporting them is spotted.[23]

This approach reduces unnecessary network traffic when the topology remains stable, but has some tradeoffs. LSAs are not refreshed which might cause LSA corruption in routers. This should be noticed with periodic LSA checksum computation and, according to OSPF specification, should result in restarting the router[24].[23]

Another drawback is the loss of some network management functionality. OSPF MIB [1] specifies `ospfExternalLSASumSum` and `ospfAreaLSASumSum` management variables which verify that the link state databases are identical. The verification is done by taking the sum of the individual LSA checksums in the database. However, when DoNotAge LSAs are used, similar LSAs might have different checksums as they can have different sequence numbers.[23]

Coping with Exceptional Situations

OSPF stub router advertisement describes a technique which can be used to prevent other routers to send transit traffic to a router. The router sets high costs to all its outgoing links. This causes the other routers to look for alternative routes around the burdened router because of the high cost associated in routing through it. However, packets addressed to directly attached networks and to networks reachable only through the router in question will be routed normally. This allows overloaded routers to avoid some of the traffic normally going through them and also creates a possibility to

prepare for maintenance by redirecting all possible transit routes to other routers.[32]

If an IS-IS transit router goes down for short period, packets bound to other AS might be lost. The routes existing prior to the router failure will be established quickly by IS-IS when the router comes back up. However, external routes learned from e.g. BGP will take much longer to become established. Therefore the other routers will send external traffic to the recently failed router which has not yet learned external routes. This leads to the router dropping those packets.[22]

The problem could be avoided. Because other routers have functioning routes which were used when the router failed, those routes should be used until the restarted router learns all necessary information. The proposed solution is to inform other routers by the use of the overload bit in LSPs to avoid routing traffic through the restarting router. The overload bit should be cleared once the information about the external routes is learned or by timer.[22]

IS-IS Mesh Groups

Some organizations operate IS-IS over point to point links (e.g. ATM virtual circuits) where the ISs form a full mesh topology. The flooding of IS-IS packets in full mesh topology results in unnecessary transmissions, as an IS sends a received flooding packet over every interface except the one over which it was received[2].

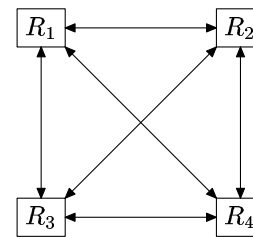


Figure 1: A Full Mesh Topology of Four Routers

Now, if router R_1 floods a LSP, it will be received by R_2 , R_3 and R_4 . R_2 sends the LSP to R_3 and R_4 , even though they already received the LSP from R_1 . According to [2] there will be a total of $N - 2$ extra transmissions in a full-mesh system with N routers. The Mesh Group RFC describes a way to build a separate flooding topology to avoid the unnecessary transmissions.

The links will be assigned two new attributes. The attribute `meshGroupEnabled` can be in three different states: `meshSet`, `meshBlocked` or `meshInactive`. The other attribute, `meshGroup`, is an integer describing the mesh group the link belongs to. The idea is to break the full mesh topology into different groups and to restrict the links used for control traffic between the groups. LSPs will be sent only over links with attribute `meshInactive`. LSPs will also be sent over `meshSet` links if the `meshGroup` attribute is different than on the link from where the LSP was received. This is the case where the router at the other end of the link belongs to different mesh group. LSPs will not be sent over `meshBlocked` links. To ensure database synchronization, CSNPs will be sent over links with attribute `meshSet` or

meshBlocked. CSNPs will be sent over meshInactive links only when initializing. The attributes are set by hand by the network administrator, which creates a possibility for errors which might lead to lost packets.[2]

Point-to-Point over Broadcast LAN

If there are only two routers in a broadcast network, it may be advantageous to treat the network as a point-to-point (p2p) link. There is no need to elect a designated router which reduces network traffic and the amount of information the routers need to store. Also, the configuration of the p2p link is simpler than that of a LAN. Furthermore, the p2p link can be unnumbered, saving the scarce address resources of IPv4. A broadcast network is normally depicted as a vertex (pseudonode) in the link state database while a p2p network is just an edge connecting two routers. Therefore, the p2p over LAN approach can save resources.[36]

Even a broadcast LAN with more than two routers can utilize this approach by dividing the LAN into logical subnets with just two routers in each and treating them as p2p links. With more than two routers, the only resource saved is the address space. However, the approach enables one to assign different costs for links to different neighbors. This is impossible in normal LANs.[36]

In contrast to the useful things, some problems also arise from the approach. The potential for misconfiguration increases as the operation of the router over the LAN can be configured manually. If the link is unnumbered, it is impossible to monitor or ping the interface in question. However, the potential of the approach in diminishing unnecessary control traffic and simpler operation outweighs the drawbacks.[36]

The point-to-point over LAN approach has become necessary with the advent of long-reach Ethernet. A range of up to 10km (over fiber)[8] is planned, which makes it an attractive choice for connecting routers when greater ranges are not needed. When long-reach Ethernet is used to connect two routers, the link is logically a point-to-point link while the technology used is a broadcast one.

3.5 IPv6 routing

Both OSPF and IS-IS have provisions for routing IPv6 [6] [12]. IS-IS protocol is extended to support IPv6 while OSPF requires deployment of a new protocol.

In IS-IS, only minor changes are required. The IPv6 support consists of two new TLVs which contain IPv6 reachability information and IPv6 interface address. Also, a value for IPv6 is added to the “protocols supported” field in LSPs.[12]

In OSPF the changes are a bit more substantial. All addressing information is removed from OSPF packet headers. Therefore, OSPFv3 is in a sense protocol independent as there are no IP addresses in LSA headers. OSPFv3 uses link local addresses for communication except in the case of virtual links where global or site-local addresses have to be used. OSPFv3 also allows multiple instances of the protocol to run over a single interface.[6]

A new LSA, Link LSA, is added to distribute router addresses to neighbors. Furthermore, in OSPF, any addressing information is removed from the OSPF packet headers. This makes OSPF protocol independent. Router IDs are kept at

length of 32 bits, which makes it impossible to assign their IPv6 addresses as router IDs.[6]

The main difference between IS-IS and OSPFv3 is that OSPFv3 is a different protocol from OSPFv2. It thus requires the use of two different protocols in networks with both IPv4 and IPv6 traffic. With IS-IS the routing can be done with a single protocol. The use of a single protocol is advantageous in many ways. Separate instances of routing protocols (e.g. OSPFv2 and v3) consume additional resources because of the need to run two copies of the software in a router. The independent instances can also interact with each other in subtle ways. With one protocol, the interactions are more explicit and the overhead of running two protocols can be avoided. On the other hand, running different protocols for IPv4 and IPv6 can increase the robustness of the network. If there are problems with one version, the other may still work normally. Anyway, when there is IPv6 and IPv4 traffic simultaneously the network topologies will have to be carefully thought out.[19]

There is an extension to IS-IS describing how to have multiple topologies on single network infrastructure. The multi-topology (MT) extensions facilitate the migration to IPv6 by allowing the IPv6 network topology to be different from the IPv4 network while upgrading the infrastructure.[30]

When having multiple topologies in an AS, each adjacency can belong to multiple topologies. An IS will advertise information over the adjacency, such as reachability, according to the topology that the adjacency belongs to. For example, In figure 2, the adjacency between the IS₁ and IS₂ belongs to topology 1 and between IS₁ and IS₃ to topology 2. Now, IS₁ will not advertise reachability of IS₃ to IS₂ because the adjacencies belong to different topologies.[30]

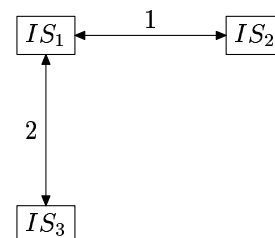


Figure 2: An Example of Multiple Topologies

The topologies are described with three new TLVs, one for advertising the MTs that the IS participates in. The remaining two are for IS reachability information and for IP reachability information. They differ from the standard TLVs only in that they have a MT identifier attached to them.[30]

The topologies all share the same level 1 / level 2 boundaries. An IS can become partitioned in a single MT and remain connected in others. Partition repair is not implemented for multiple topologies.[30]

3.6 Restarting

In modern routers, there are usually separate processors for packet forwarding and control functions. This creates a possibility to keep forwarding packets even when the control plane is being restarted. As the router can not react to

changes in network topology while the control plane is being restarted, there must be some mechanism to inform the restarting router if the network topology does not remain stable. The mechanism is called Graceful Restart or Non-Stop Forwarding.[25][35]

Graceful restart has been specified for both OSPF and IS-IS. The mechanism functions in a similar fashion for both protocols. Graceful restart requires cooperation of all neighboring routers. The restarting router first notifies all neighboring routers about the restart. The helping routers then keep the adjacency up during the restart if the network remains stable. If the helpers receive LSAs that differ from earlier LSAs, meaning that the network topology has changed, they immediately flood LSAs reporting the restarting router as being down. This is done for safety, namely to avoid routing loops or blackholes.[25][35]

The time allowed for graceful restart is limited by timers. If the restart takes too long, the helpers report the restarting router down as in the case of network topology change.[25][35]

Graceful restart techniques depend on the correct functioning of the forwarding engine. Thus, it can be used to avoid rerouting during planned restarts. Also restarts due to the control plane failures can benefit from graceful restart if the control functions can report the situation to helpers before having to restart. If the restart occurs due to problems in the forwarding plane, graceful restart can naturally not be utilized.

3.7 Other Miscellaneous Extensions

OSPF

A new LSA class for OSPF, called Opaque LSA is defined in [5]. Opaque LSA provides a way to include generic information to other routers in LSAs, therefore enabling future expandability. The RFC defines three new types of LSAs, 9, 10 and 11. The different types are flooded to different scopes. The LSA type 9 has link-local flooding, which is not available in plain OSPF. Routers are required to flood the opaque LSAs even if they do not recognize the opaque type. Most extensions to OSPF use opaque LSAs to distribute information.[5]

IS-IS LSPs use a TLV (Type, Length, Value) encoding for additional information. Therefore there is no need for a similar extension for IS-IS as the TLV approach can easily incorporate new types of information. IS-IS routers are also required to flood LSPs with unknown TLVs.

Optional router capabilities for OSPF are normally advertised in Hello packets, but the bits reserved for the purpose have been used up. A proposal to use opaque LSAs for the purpose is presented in [20].

There is also a proposal to include an extra part to the end of OSPF Hello packets. The mechanism is called Link Local Signaling (LLS). LLS also provides a way to include information in Hello packets. The LLS mechanism is backwards compatible, because routers not supporting LLS will never see the extra part.[40]

OSPF NSSA Option [26] defines an expanded stub area. Not-So-Stubby Area (NSSA) can have AS border routers in the area. This can be useful, for example, when an ISP has

a customer connection at the end of a slow link. The router connecting to the customer network can be configured as a NSSA. It will not receive information about other AS external routes (similar to stub area) but it can advertise AS external routes to other routers.[26]

The extension defines a new type of LSA, namely the Type 7 AS external LSA, which can be flooded in a NSSA, but not outside it. The ABRs perform the translation of Type 7 LSAs to type 5 LSAs (the normal type of AS external LSA) at the NSSA border.[26]

IS-IS

A change to IS-IS has been proposed which would allow distribution of Level 1 prefixes to Level 2 ISs and vice versa. This makes the protocol less scalable but can improve routes [18]. This extension makes the IS-IS Level 2 sub-topology behave more like OSPF backbone. The Level 2 ISs now can choose the shortest route to the destination instead of the shortest route to the Level 1/2 IS in the correct area.

A new TLV for IS-IS to include CSNP and PSNP checksum is proposed in [29]. The protocols should verify the integrity of the packets, but fail to do so sometimes.

A new TLV for experimental use for IS-IS protocol is defined in [4]. The experimental TLV may be used by anyone to test extensions to IS-IS. The experimental TLV is assigned a code of 250, and it must contain an organizationally unique identifier (OUI) assigned by the manufacturer of the router. Currently the manufacturers pick the TLV code for experimental features more or less randomly. This can cause interoperability problems if two manufacturers pick the same code for two different features. The experimental TLV also conserves the limited amount of TLV codepoints in IS-IS.[4]

Implementations ignore the experimental TLV if the OUI does not match. Routers are, however, required to forward the experimental TLVs even if they ignore the contents. This allows manufacturers to add experimental features to their routers while allowing for interoperability.[4]

IS-IS LSPs can be fragmented to a maximum of 256 fragments. The fragment size is bounded. Therefore the amount of information that can be distributed in a LSP is also bounded. It is proposed to extend the number of fragments beyond the 256 limit by using virtual routers, which in reality equate to the router, to advertise the required additional information. For backwards compatibility, the router can advertise links with cost of 0 to the virtual routers. The amount of information that needs to be distributed is expected to grow with the increasing use of TE extensions.[11]

IS-IS does not have a provision for dynamically assigning names to ISs. Static assignment is possible, but requires extensive manual configuration and is error-prone. A new TLV is defined for advertising the name of the IS. This makes it possible to assign a name to an IS and having that IS announce its address/name mapping to other ISs.[37]

3.8 Corrections

If there is an ABR in OSPF which is not connected to backbone, packets may be lost. IBM and Cisco have implemented methods [41] to correct the problem. The OSPF specification states however that there can not be such ABRs in page

26, section 3.1 :“The OSPF backbone always contains all area border routers” [24]. The problem arises when an ABR which is not connected to the backbone (not even with virtual link) receives traffic not destined to any of the areas attached to the ABR. The ABR will then drop the traffic as it will not have any routes except for the areas attached to it. The standard solution is to add a virtual link to backbone. However, it can be advantageous to have a router belonging to several areas while not being part of a backbone. For example, a router which is used to connect customers to the service provider’s network might make use of two areas for redundancy.[41]

The proposed methods redefine an ABR to be a router which is connected to several areas, one of which is the backbone. A router is allowed to connect to several areas but it will not be considered an ABR if there is no backbone connection. The multi-area router will behave as an ABR would except that it does not announce itself as being an ABR. This allows for correct routing in all attached areas. The multi-area router can still forward traffic from one area to another if the shortest path goes through it, but it does not actively attract inter-area traffic.[41]

There is a possibility of routing loops in BGP/MPLS VPNs if the VPN uses OSPF. The problem arises from the fact that converting routes from OSPF to BGP, and vice versa, some information may be lost. A solution is described in [33]. If the provider’s router has already sent a route to the customer’s router, it will ignore the same route if the customer’s router sends it back.[33]

3.9 Security

A new value for IS-IS authentication TLV for use with HMAC-MD5 authentication is specified. This allows for cryptographic authentication while the original specification only included clear-text passwords [17].

BFD protocol, if used with network layer protocols, creates a risk of denial of service attacks. A proposed solution is to set the TTL field of the packets to maximum when transmitting and refuse to accept a packet if the TTL field is less than maximum. This mechanism works for single hop communication. For multiple hop communication, some other mechanism should be used[13].

In OSPFv3, IPv6 Authentication Header (AH) and Encapsulating Security Payload (ESP) can be used to authenticate and encrypt OSPF packets[6]. The use of AH and ESP is far from straightforward and a draft exists describing their use [10].

OSPFv2 is vulnerable to router spoofing from outside the attached network due to its use of IP in distributing routing information. IP can carry traffic multiple hops (as it was meant to) and thus the attacker can try to inject false information to routers even when not directly connected to the network under attack.[31]

The IS-IS approach of distributing the routing information directly over link layer protects it fairly well from that kind of attacks as the attacker must have a direct, single hop, connection to the IS.[31]

4 Implementations

In this section I examine which changes have been implemented. I use as sources public documentation of Cisco’s and Juniper’s router software. The information has been compiled from the feature descriptions in the documentation. Thus, it is possible that some changes that have been implemented do not appear in documentation, especially if the feature is not configurable. Therefore the following information may contain omissions and serves only as a rough guide to the usefulness of the proposed extensions to the protocols. It is not meant to be a guide to implemented features in the routing software. The software versions are IOS 12.3(7) (Early Release) for Cisco and JUNOS 6.2 for Juniper.

Feature	Cisco	Juniper
TE Extensions	•	•
Hostnames	•	•
Prefix Distribution	•	•
Mesh Groups	•	•
SNP Checksums		•
IPv6	•	•
Crypto Auth	•	•
BFD		•
Multitopologies	•	•
Graceful Restart	•	•
p2p over LAN	•	•

Table 1: Implemented Features for IS-IS

Feature	Cisco	Juniper
TE Extensions	•	•
IPv6	•	•
NSSA	•	•
BFD		•
Graceful Restart	•	•
LLS	•	
DoNotAge LSA	•	
Stub Router adv.	•	•
Opaque LSA	•	•

Table 2: Implemented Features for OSPF

Nearly all of the changes described in this paper have been implemented by both vendors. A few newer ones only by one. The changes described here and not implemented are: “TLV for Experimental Use” and “Extending the Number of IS-IS LSP Fragments Beyond the 256 Limit”.

The experimental TLV provides guaranteed interoperability with routers of different manufacturers. Apparently there have been few problems yet with TLV codepoints collisions affecting network operations. The extension of the number of TLV fragments might become required in the future as the amount of information distributed in LSPs increases.

BFD seems to be implemented on Juniper’s routers only. Cisco has an “accelerated Hello” mechanism for OSPF, which has the same aims, namely the reduction of the time needed to notice a link failure. The BFD approach is more flexible, BFD being protocol independent.

The point-to-point over LAN feature seems to be implemented only for IS-IS. This comes as a bit of a surprise, as the feature appears to be fairly protocol independent.

5 Conclusions

The protocols have evolved considerably during their existence. During the whole 15 years or so that the protocols have existed, most of the work has been done to OSPF, but lately the development of IS-IS has also been active. Currently there seems to be some amount of cooperation in the development of the two protocols. The cooperation is not surprising, as the protocols are very similar and aimed to the same areas of operation.

At the moment the “hottest” new directions seem to be convergence time reduction and traffic engineering. Most new internet-drafts or RFCs are aimed to one or another. For example, the BFD protocol is aimed to decrease convergence time while the MPLS extensions are aimed to facilitate traffic engineering.

Work is also done in reducing the control traffic overhead. This is likely to become important as TE is likely to increase the amount of control traffic that needs to be sent.

Adding new features to the protocols require new types of messages. This brings forward another class of extensions, namely those which aim to expand the facilities for router-to-router communication. An example of this would be the LLS extension for OSPF. These extensions are by nature very protocol dependent.

References

- [1] F. Baker and R. Coltun. OSPF Version 2 Management Information Base. RFC 1253, 1991.
- [2] R. Balay, D. Katz, and J. Parker. IS-IS Mesh Groups. RFC 2973, 2000.
- [3] R. Callon. Use of OSI IS-IS for Routing in TCP/IP and Dual Environments. RFC 1195, 1990.
- [4] P. Christian. TLV for Experimental Use. Internet Draft, 2004.
- [5] R. Coltun. The OSPF Opaque LSA Option. RFC 2370, 1998.
- [6] R. Coltun, D. Ferguson, and J. Moy. OSPF for IPv6. RFC 2740, 1999.
- [7] E. David Oran. OSI IS-IS Intra-domain Routing Protocol. RFC 1142, 1990.
- [8] H. Frazier and G. Pesavento. Ethernet takes on the first mile. *IT Professional*, 3(4):17–22, 2001.
- [9] E. Gagan Choudhury. Prioritized Treatment of Specific OSPF Packets and Congestion Avoidance. Internet Draft, 2003.
- [10] M. Gupta and N. S. Melam. Authentication/Confidentiality for OSPFv3. Internet Draft, 2003.
- [11] A. Hermelin, S. Previdi, and M. Shand. Extending the Number of IS-IS LSP Fragments Beyond the 256 Limit. Internet Draft, 2004.
- [12] C. Hopps. Routing IPv6 with IS-IS. Internet Draft, 2003.
- [13] D. Katz and D. Ward. Bidirectional Forwarding Detection. Internet Draft, 2003.
- [14] D. Katz, D. M. Yeung, and K. Kompella. Traffic Engineering (TE) Extensions to OSPF Version 2. RFC 3630, 2003.
- [15] K. Kompella and Y. Rechter. IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching. Internet Draft, 2003.
- [16] K. Kompella and Y. Rechter. OSPF Extensions in Support of Generalized Multi-Protocol Label Switching. Internet Draft, 2003.
- [17] T. Li and R. Atkinson. Intermediate System to Intermediate System (IS-IS) Cryptographic Authentication. RFC 3567, 2003.
- [18] T. Li, T. Przygienda, and H. Smith. Domain-wide Prefix Distribution. RFC 2966, 2000.
- [19] M. Lind, V. Ksinant, S. D. Park, A. Baudot, and P. Savola. Scenarios and Analysis for Introducing IPv6 into ISP Networks. Internet Draft, 2004.
- [20] A. Lindem, N. Shen, R. Aggarwal, S. Shaffer, and J. Vasseur. Extensions to OSPF for Advertising Optional Router Capabilities. Internet Draft, 2004.
- [21] V. Manral and M. Bhatia. IS-IS and OSPF Difference Discussions. Internet Draft, 2003.
- [22] D. McPherson. Intermediate System to Intermediate System (IS-IS) Transient Blackhole Avoidance. RFC 3277, 2002.
- [23] J. Moy. Extending ospf to support demand circuits. RFC 1793, 1995.
- [24] J. Moy. OSPF Version 2. RFC 2328, 1998.
- [25] J. Moy, A. Lindem, and P. Pillay-Esnault. Graceful OSPF Restart. RFC 3623, 2003.
- [26] P. Murphy. The OSPF Not-So-Stubby Area (NSSA) Option. RFC 3101, 2003.
- [27] P. Pillay-Esnault. OSPF Refresh and Flooding Reduction in Stable Topologies. Internet Draft, 2003.
- [28] S. Poretsky. Benchmarking Applicability for IGP Route Convergence. Internet Draft, 2003.
- [29] T. Przygienda. Optional Checksums in Intermediate System to Intermediate System (ISIS). RFC 3358, 2002.
- [30] T. Przygienda, N. Shen, and N. Sheth. M-ISIS: Multi Topology (MT) Routing in IS-IS. Internet Draft, 2003.

- [31] K. J. Repage and O. L. Moigne. OSPF Security Vulnerabilities Analysis. Internet Draft, 2003.
- [32] A. Retana, L. Nguyen, R. White, A. Zinin, and D. McPherson. OSPF Stub Router Advertisement. RFC 3137, 2001.
- [33] E. C. Rosen and P. Pillay-Esnault. Using an LSA Options Bit to Prevent Looping in BGP/MPLS IP VPNs. Internet Draft, 2004.
- [34] E. C. Rosen, A. Viswanathan, and R. Callon. Multiprotocol Label Switching Architecture. RFC 3031, 2001.
- [35] M. Shand and L. Ginsberg. Restart signaling for IS-IS. Internet Draft, 2004.
- [36] N. Shen, A. Lindem, J. Yuan, A. Zinin, R. White, and S. Previdi. Point-to-Point Operation Over LAN in Link State Routing Protocols. Internet Draft, 2003.
- [37] N. Shen and H. Smith. Dynamic Hostname Exchange Mechanism for IS-IS. RFC 2763, 2000.
- [38] H. Smit and T. Li. IS-IS extensions for Traffic Engineering. Internet Draft, 2003.
- [39] M. Thorup. OSPF Areas Considered Harmful. Internet Draft, 2003.
- [40] A. Zinin, B. Friedman, L. Nguyen, A. Roy, and D. Young. OSPF Link Local Signaling. Internet Draft, 2004.
- [41] A. Zinin, D. M. Yeung, and A. Lindem. Alternative Implementations of OSPF Area Border Routers. RFC 3509, 2003.