Open Shortest Path First (OSPF)

Background

Open Shortest Path First (OSPF) is a routing protocol developed for *Internet Protocol* (IP) networks by the *interior gateway protocol* (IGP) working group of the Internet Engineering Task Force (IETF). The working group was formed in 1988 to design an IGP based on the *shortest path first* (SPF) algorithm for use in the *Internet*. Similar to the *Interior Gateway Routing Protocol* (IGRP), OSPF was created because in the mid-1980s, the *Routing Information Protocol* (RIP) was increasingly unable to serve large, heterogeneous internetworks. This chapter examines the OSPF routing environment, underlying routing algorithm and general protocol components.

OSPF was derived from several research efforts, including Bolt, Beranek, Newman's (BBN's) SPF algorithm developed in 1978 for the *ARPANET* (a landmark packet-switching network developed in the early 1970s by BBN), Dr. Radia Perlman's research on fault-tolerant broadcasting of routing information (1988), BBN's work on area routing (1986), and an early version of OSI's Intermediate System-to-Intermediate System (IS-IS) routing protocol.

OSPF has two primary characteristics. The first is that the protocol is open, which means that its specification is in the public domain. The OSPF specification is published as *Request For Comments* (RFC) 1247. The second principal characteristic is that OSPF is based on the SPF algorithm, which sometimes is referred to as the *Dijkstra algorithm*, named for the person credited with its creation.

OSPF is a *link-state* routing protocol that calls for the sending of *link-state advertisements* (LSAs) to all other routers within the same hierarchical area. Information on attached interfaces, metrics used, and other variables is included in OSPF LSAs. As OSPF routers accumulate link-state information, they use the SPF algorithm to calculate the shortest path to each node.

As a link-state routing protocol, OSPF contrasts with RIP and IGRP, which are distance-*vector* routing protocols. Routers running the distance-vector algorithm send all or a portion of their routing tables in routing-update messages to their neighbors.

Routing Hierarchy

Unlike RIP, OSPF can operate within a hierarchy. The largest entity within the hierarchy is the *autonomous system* (AS), which is a collection of networks under a common administration that share a common routing strategy. OSPF is an intra-AS (interior gateway) routing protocol, although it is capable of receiving routes from and sending routes to other ASs.

An AS can be divided into a number of *areas*, which are groups of contiguous networks and attached hosts. Routers with multiple interfaces can participate in multiple areas. These routers, which are called *area border routers*, maintain separate topological databases for each area.

A *topological database* is essentially an overall picture of networks in relationship to routers. The topological database contains the collection of LSAs received from all routers in the same area. Because routers within the same area share the same information, they have identical topological databases.

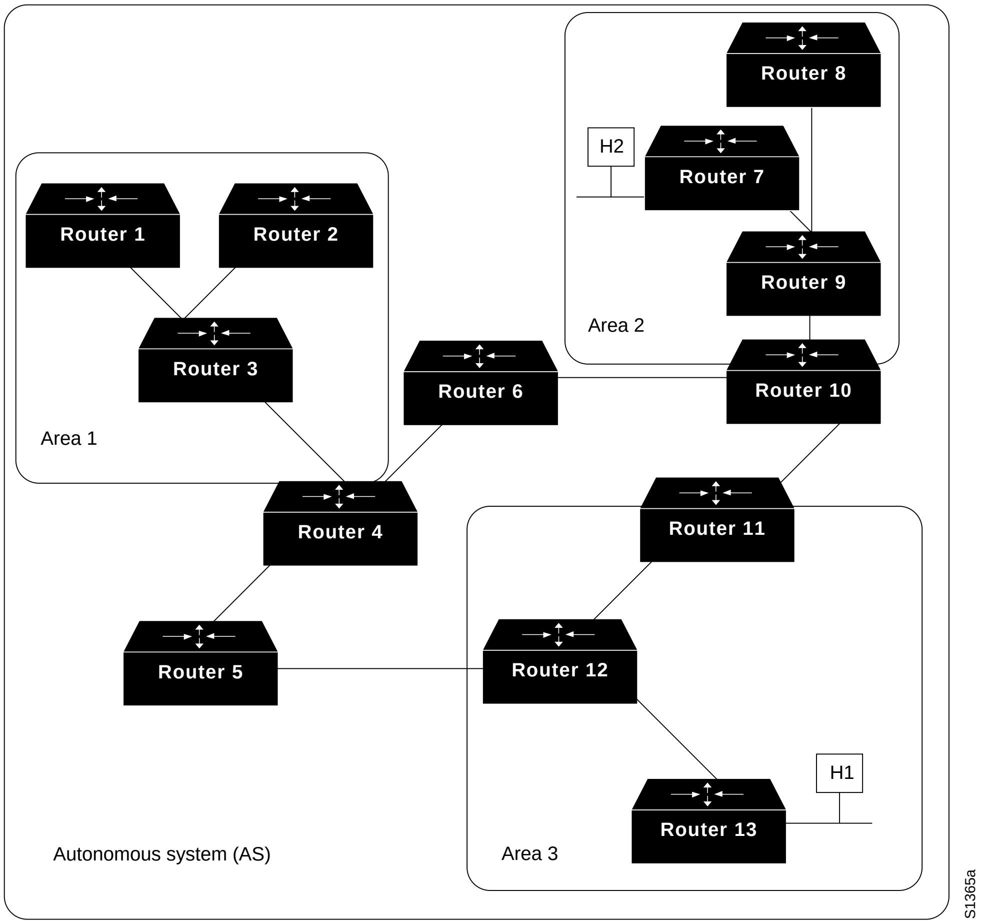
The term *domain* sometimes is used to describe a portion of the network in which all routers have identical topological databases. Domain is frequently used interchangeably with AS.

An area's topology is invisible to entities outside the area. By keeping area topologies separate, OSPF passes less routing traffic than it would if the AS were not partitioned.

Area partitioning creates two different types of OSPF routing, depending on whether the source and destination are in the same or different areas. Intra-area routing occurs when the source and destination are in the same area; interarea routing occurs when they are in different areas.

An OSPF *backbone* is responsible for distributing routing information between areas. It consists of all area border routers, networks not wholly contained in any area, and their attached routers. Figure 42-1 shows an example of an internetwork with several areas.

Figure 42-1 An OSPF AS consists of multiple areas linked by routers.



In the figure, Routers 4, 5, 6, 10, 11, and 12 make up the backbone. If Host H1 in Area 3 wants to send a packet to Host H2 in area 2, the packet is sent to Router 13, which forwards the packet to Router 12, which sends the packet to Router 11. Router 11 then forwards the packet along the backbone to area border Router 10, which sends the packet through two intra-area routers (Router 9 and Router 7) to be forwarded to Host H2.

The backbone itself is an OSPF area, so all backbone routers use the same procedures and algorithms to maintain routing information within the backbone that any area router would. The backbone topology is invisible to all intra-area routers, as are individual area topologies to the backbone.

Areas can be defined in such a way that the backbone is not contiguous. In this case, backbone connectivity must be restored through *virtual links*. Virtual links are configured between any backbone routers that share a link to a nonbackbone area and function as if they were direct links.

AS border routers running OSPF learn about exterior routes through *exterior gateway protocols* (EGPs), such as *Exterior Gateway Protocol* (EGP) or *Border Gateway Protocol* (BGP), or through configuration information. For more information about these protocols, see Chapter 35, “Border Gateway Protocol (BGP).”

SPF Algorithm

The shortest path first (SPF) routing algorithm is the basis for OSPF operations. When an SPF router is powered up, it initializes its routing-protocol data structures and then waits for indications from lower-layer protocols that its interfaces are functional.

After a router is assured that its interfaces are functioning, it uses the OSPF *Hello protocol* to acquire neighbors, which are routers with interfaces to a common network. The router sends hello packets to its neighbors and receives their hello packets. In addition to helping acquire neighbors, hello packets also act as keep-alives to let routers know that other routers are still functional.

On *multiaccess networks* (networks supporting more than two routers), the Hello protocol elects a *designated router* and a backup designated router. Among other things, the designated router is responsible for generating LSAs for the entire multiaccess network. Designated routers allow a reduction in network traffic and in the size of the topological database.

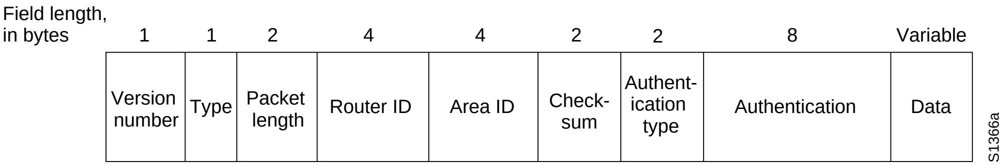
When the link-state databases of two neighboring routers are synchronized, the routers are said to be *adjacent*. On multiaccess networks, the designated router determines which routers should become adjacent. Topological databases are synchronized between pairs of adjacent routers. Adjacencies control the distribution of routing-protocol packets, which are sent and received only on adjacencies.

Each router periodically sends an LSA to provide information on a router’s adjacencies or to inform others when a router’s state changes. By comparing established adjacencies to link states, failed routers can be detected quickly and the network’s topology altered appropriately. From the topological database generated from LSAs, each router calculates a shortest-path tree, with itself as root. The shortest-path tree, in turn, yields a routing table.

Packet Format

All OSPF packets begin with a 24-byte header, as illustrated in Figure 42-2.

Figure 42-2 OSPF packets consist of nine fields.



The following descriptions summarize the header fields illustrated in figure 42-2.

- *Version Number*—Identifies the OSPF version used.
- *Type*—Identifies the OSPF packet type as one of the following:
  - — Hello: Establishes and maintains neighbor relationships.
  - — Database Description: Describes the contents of the topological database. These messages are exchanged when an adjacency is initialized.
  - — Link-state Request: Requests pieces of the topological database from neighbor routers. These messages are exchanged after a router discovers (by examining database-description packets) that parts of its topological database are out of date.
  - — Link-state Update: Responds to a link-state request packet. These messages also are used for the regular dispersal of LSAs. Several LSAs can be included within a single link-state update packet.
  - — Link-state Acknowledgment: Acknowledges link-state update packets.
- *Packet Length*—Specifies the packet length, including the OSPF header, in bytes.
- *Router ID*—Identifies the source of the packet.
- *Area ID*—Identifies the area to which the packet belongs. All OSPF packets are associated with a single area.
- *Checksum*—Checks the entire packet contents for any damage suffered in transit.
- *Authentication Type*—Contains the authentication type. All OSPF protocol exchanges are authenticated. The Authentication Type is configurable on a per-area basis.
- *Authentication*—Contains authentication information.
- Data—Contains encapsulated upper-layer information.

Additional OSPF Features

Additional OSPF features include equal-cost, *multipath routing,* and routing based on upper-layer *type-of-service* (TOS) requests. TOS-based routing supports those upper-layer protocols that can specify particular types of service. An application, for example, might specify that certain data is urgent. If OSPF has high-priority links at its disposal, these can be used to transport the urgent datagram.

OSPF supports one or more metrics. If only one metric is used, it is considered to be arbitrary, and TOS is not supported. If more than one metric is used, TOS is optionally supported through the use of a separate metric (and, therefore, a separate routing table) for each of the eight combinations created by the three IP TOS bits (the *delay*, *throughput*, and *reliability* bits). If, for example, the IP TOS bits specify low delay, low throughput, and high reliability, OSPF calculates routes to all destinations based on this TOS designation.

IP subnet masks are included with each advertised destination, enabling *variable-length subnet masks*. With variable-length subnet masks, an IP network can be broken into many subnets of various sizes. This provides network administrators with extra network-configuration flexibility.