CHAPTER 25

Open Shortest Path First

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*, a large, international network connecting research institutions, government agencies, universities, and private businesses. Like the *Interior Gateway Routing Protocol* (IGRP), OSPF was created because the *Routing Information Protocol* (RIP) was, in the mid-1980s, increasingly unable to serve large, heterogeneous internetworks. (For more information about IGRP and RIP, see Chapter 24, "Interior Gateway Routing Protocol and Enhanced IGRP," and Chapter 23, "Routing Information Protocol," respectively.)

OSPF was derived from several research efforts, including the following:

- Bolt, Beranek, and 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. Perlman's research on fault-tolerant broadcasting of routing information (1988)
- BBN's work on area routing (1986)
- An early version of OSI's Intermediate System-to-Intermediate System (IS-IS) routing protocol

For more information about IS-IS, see Chapter 28, "OSI Routing."

As indicated by its acronym, OSPF has two primary characteristics. The first is that it is open, in 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 it is based on the SPF algorithm, which is sometimes referred to as the *Dijkstra algorithm*, named for the person credited with its creation.

Technology Basics

OSPF is a *link state* routing protocol. As such, it 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, but only to their neighbors.

Routing Hierarchy

Unlike RIP, OSPF can operate within a hierarchy. The largest entity within the hierarchy is the *autonomous system* (AS). An AS is a collection of networks under a common administration, sharing 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*. An area is a group 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.

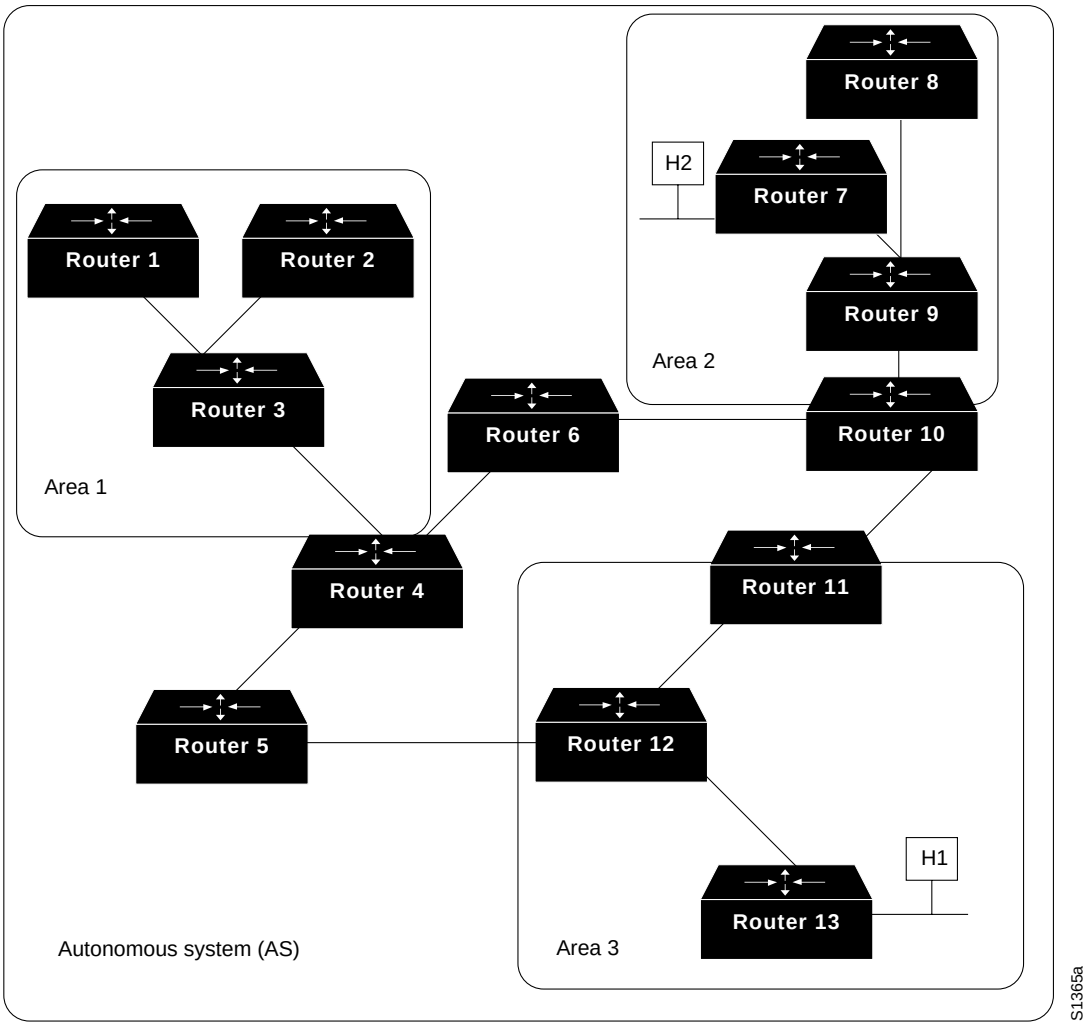
The term *domain* is sometimes used to describe a portion of the network in which all routers have identical topological databases. Domain is frequently used interchangeably with AS.

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 25-1 shows an example of an internetwork with several areas.

Figure 25-1 Hierarchical OSPF Internetwork



In this figure, Routers 4, 5, 6, 10, 11, and 12 make up the backbone. If host H1 in area 3 wishes 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 forwards the packet along the backbone to area border router 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 26, “Exterior Gateway Protocol” and Chapter 27, “Border Gateway Protocol,” respectively.

SPF Algorithm

The 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.

Once a router is assured that its interfaces are functioning, it uses the OSPF *Hello protocol* to acquire *neighbors*. Neighbors 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 keepalives 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. The designated router is responsible, among other things, for generating LSAs for the entire multiaccess network. Designated routers allow a reduction in network traffic and in the size of the topological database.

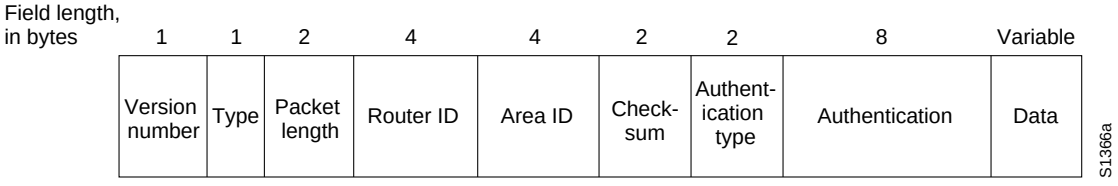
When the link state databases of two neighboring routers are synchronized, the routers are said to be *adjacent*. On multiaccess networks, the designated router determines which routers should become adjacent. Topological databases are synchronized between pairs of adjacent routers. Adjacencies control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

Each router periodically sends an LSA. LSAs are also sent when a router's state changes. LSAs include information on a router's adjacencies. By comparing established adjacencies to link states, failed routers can be quickly detected 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 shown in Figure 25-2.

Figure 25-2 OSPF Header Format



The fields of the OSPF header are as follows:

- *Version number*—Identifies the particular OSPF implementation being used.
- *Type*—Specifies one of five OSPF packet types:
  - *Hello*—Sent at regular intervals to establish and maintain neighbor relationships.
  - *Database description*—Describes the contents of the topological database, and are exchanged when an adjacency is being initialized.
  - *Link state request*—Requests pieces of a neighbor's topological database. They are exchanged after a router has discovered (through examination of database description packets) that parts of its topological database are out of date.

- *Link state update*—Responses to link state request packets. They are also used for the regular dispersal of LSAs. Several LSAs may be included within a single packet.
- *Link state acknowledgment*—Acknowledges link state update packets. Link state update packets must be explicitly acknowledged to ensure that link state flooding throughout an area is a reliable process.

  Each LSA in a link state update packet contains a type field. There are four LSA types:

- *Router links advertisements* (RLAs)—Describe the collected states of the router's links to a specific area. A router sends an RLA for each area to which it belongs. RLAs are flooded throughout the entire area, and no further.
- *Network links advertisements* (NLAs)—Sent by the designated routers. They describe all the routers that are attached to a multiaccess network, and are flooded throughout the area containing the multiaccess network.
- *Summary links advertisements* (SLAs)—Summarize routes to destinations outside an area, but within the AS. They are generated by area border routers, and are flooded throughout the area. Only intra-area routes are advertised into the backbone. Both intra-area and interarea routes are advertised into the other areas.
- *AS external links advertisements*—Describe a route to a destination that is external to the AS. AS external links advertisements are originated by AS boundary routers. This type of advertisement is the only type that is forwarded everywhere in the AS; all others are forwarded only within specific areas.

- *Packet length*—Specifies in bytes the packet's length, including the OSPF header.
- *Router ID*—Identifies the packet's source.
- *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 potential damage suffered in transit.
- *Authentication type*—Contains an authentication type. "Simple password" is an example of an authentication type. All OSPF protocol exchanges are authenticated. The authentication type is configurable on a per-area basis.
- *Authentication*—Contains authentication information and is 64 bits in length.

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. For example, an application 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). For example, if 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.