OSPF

This lecture gives an overview of Open Shortest Path First (OSPF), a commonly used intra-domain routing protocol. After this lecture you should be able to: select between RIP, EIGRP, and OSPF based on router capabilities, determine if two routers are adjacent based on Hello packet contents, determine if a received LSA would be flooded or not based on its age and sequence number, manually calculate the shortest paths and SPF-based routing table for a simple network, explain meaning of OSPF area ID 0.0.0.0, find typical router ID value(s) based on its IP address(es).

Introduction

The purpose of an IP network is to route packets. Routing tables tell each router or host how to forward packets. This lecture discusses how these routing tables are set up and maintained.

Simple routers that contain only a default route can be configured using DHCP. The routing tables for simple networks can be configured manually. However, for more complex networks the routing tables should be updated automatically to cope with the addition/deletion of routers and links between them.

In this lecture we describe OSPF, the most common Internal Gateway Protocol (IGP), a routing protocol used to configure the routers within an Autonomous System (AS).

Comparison of IGP Protocols

Routers within a domain almost always use the same routing protocol. Different router models support different protocols and this will dictate the routing protocol chosen for a particular network.

RIP is an older protocol whose path cost calculations are based one the number of hops.

EIGRP is a proprietary protocol limited to Cisco routers and so can only be used in networks that consist exclusively of Cisco equipment.

OSPF is an open protocol and is supported by all modern routers. Its performance and features are generally considered equal to EIGRP.

Exercise 1: Which IGP would you use if your network included routers that only supported RIP? If you had a mixture of (modern) Cisco and Juniper Networks routers? If your company had a "Cisco only" policy?

OSPF

Published as RFC 2328 (OSPF v2), OSPF differs from earlier routing protocols in using what is called "link state" routing as compared to "distance-vector routing."

Distance-vector routing means routers exchange cost of *routes* to all other routers. Link-state routing means routers only exchange the cost of *links* to their immediately neighbouring routers. Link state routing algorithms are believed to converge faster and be more reliable.

OSPF includes some features that make it relatively complex:

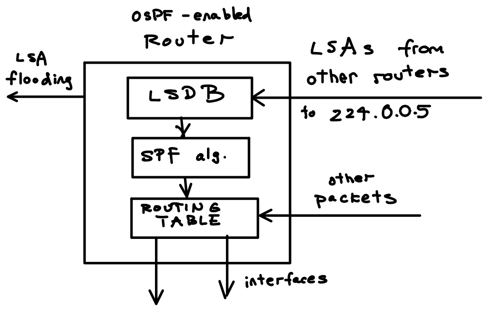
- the AS can be segmented into "areas" to limit the size of routing databases,
- use of "Designated Routers" to aggregate routing information when OSPF is used on broadcast-capable networks such as Ethernet,
- routing can be a function of the IP header's TOS (Type of Service) field (although this is not widely used), and
- routes can be propagated into and out of an area.

We cannot cover the operation of OSPF in detail. The purpose of this lecture is to introduce the most important concepts and terminology.

The operation of OSPF can be divided into the following steps:

- each router must be configured with some basic information (interface IP addresses and netmasks, link costs, area ID(s), password(s), ...)
- each router monitors its interfaces to discover what other routers it can communicate with

- each router distributes its link state information to other routers, possibly indirectly using a Designated Router
- each router collects link state information from other routers to keep its link state database up to date
- each router computes shortest (lowest cost) paths to every other router in the area and uses this to derive a routing table



Hello Packets

Link state is derived by having each router transmit a "Hello" packet every 10 seconds on every interface that is configured for OSPF. Like other OSPF packets, it is an IP packet with protocol 89. The destination IP address is set to 224.0.0.5 which is a multicast IP address that is not assigned to any host or network.

Exercise 2: Does OSPF use UDP or TCP?

Each Hello packet includes the addresses of routers in the same area from which Hello packets have been received.

A router receiving a Hello packet that includes its own IP address establishes an "adjacency" to the source of the packet (meaning that these two routers can communicate).

If a Hello packet is not seen for 40 seconds the link is determined to be down and this fact is propagated to other routers.

Exercise 3: A router with IP address 1.1.1.1 receives a Hello packet from 3.3.3.3 with the IP addresses 1.1.1.1 and 2.2.2.2 in the Neighbors field. What routers can you be certain are adjacent?

Link State Advertisements

The adjacency information derived from Hello packets is used to generate Link State Advertisement (LSA) packets that describe the links between routers.

The information contained in an LSA includes:

- the "advertising router" which generated the LSA
- an age value (seconds since the LSA was generated) to determine when an LSA is too old
- a sequence number that is used to detect duplicate LSAs
- an error-detecting checksum
- IDs (IP addresses) and metrics for individual links

There are several different types of LSAs because they can come from routers or DRs (see below) and can advertise routes between two routers, within an area, between areas or routes outside the AS.

Flooding

The LSAs are propagated through the area by flooding. Individual LSAs are aggregated and transmitted inside link state update (LSU) packets.

When a router receives an LSA it sends it to every connected router except the one it was received from if the sequence number is larger than the one already stored for that router and the age value is less than the maximum age.

Exercise 4: Why would you not flood an LSA whose sequence number was the same as one already stored?

Exercise 5: The standard maximum LSA age is one hour. What range of values would you expect to find in the LS Age field of an LSA packet?

Designated Routers

With broadcast links such as Ethernet LANs every router connected to a LAN is potentially "adjacent" to every other router on the LAN. This would result in unnecessarily large amounts of traffic when exchanging LSAs.

Routers on a LAN can "elect" one router to act as the Designated Router (DR). The DR collects LSAs and maintains the link state database for all of the routers on the LAN.

The election is done by looking at the priority of each router. The router with highest priority becomes the DR and the one with second-highest priority is the backup DR (BDR).

The multicast address 224.0.0.6 is used to send LSAs to the DR; it is the "address" of the DR.

Exercise 6: Would you configure the fastest or the slowest routers with the highest priority?

Database Transfers

LSAs are only sent when a link comes up, down and every 30 minutes thereafter. It would therefore take at least 30 minutes for a router to acquire a complete link state database. To avoid this delay, OSPF allows a router to request a copy of the link state database from another router. This is done by sending Link State Request and Link State Acknowledge packets.

Routing Table Computation

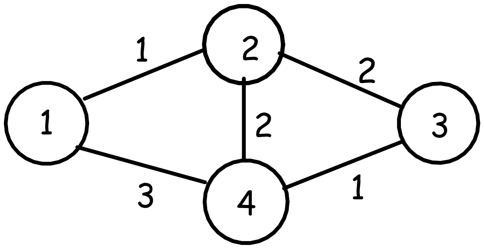
The computation of the routing table involves computing the shortest (lowest total cost) path to each destination.

We can model a network as a graph of nodes (routers) and edges (links). Each edge has a cost. The routing problem is to find the least-cost path through the graph. This is a graph theory problem and similar optimization problems exist in other areas, such as transportation or scheduling.

The algorithm typically used, Djikstra's Shortest-Path Algorithm, is fairly involved and will not be covered here.

Although the shortest-path algorithm finds a complete path to a given destination, only the next link on that path is required when forwarding a packet.

Exercise 7: Find the routing table for host 1 for the network with the link costs shown in the diagram below.



OSPF Areas

We may want to segment an AS into separate 'areas' for routing efficiency.

For example, there may be geographically separate networks connected by slow or expensive links (e.g. branch offices). Since the traffic for each office has to go through a specific router, there is no need to extend the routing algorithm outside that network.

Each area is assigned a 32-bit ID. The backbone area connects all other areas and must have area ID of 0.0.0.0.

Area Border Routers (ABRs) connect these areas to the backbone network.

Router ID

Routers are assigned a 32-bit ID that should be unique within an AS. Typically, this is the smallest (or sometimes the largest) IP address of the router.

Exercise 8: Why would a router have multiple IP addresses?

Aging

Link status database entries are re-flooded every 30 minutes. If an entry in the link status database is more than an hour old it is flushed.

Path Costs

Link costs are not defined by the OSPF standard. The rules used to compute link costs are typically based on link speed. A common rule is to compute the cost as inversely proportional to the data rate. For example $10^8/datarate$ would give a cost of 1 for a data rate of 100 Mb/s and a cost of 10 at a data rate of 10 Mb/s. As the maximum data rate increases the constant also needs to increase.

Metrics can also include factors such as link delay, packet loss rate and a links maximum packet length.

Exercise 9: Which of these would go into the numerator and which into the denominator of a cost function?

External Routes

Most networks include one or more routers that provide links to the Internet. These Autonomous System Border Routers (ASBRs) need to run BGP (the standard Exterior Gateway Protocol) on the inter-domain interface and OSPF on the intra-domain interface in order to provide a gateway to the internet.

These default routes can be forwarded by a specific type of LSA and will then be included in routing tables.

Authentication

To prevent false routing information from being accidentally or intentionally injected into the network, OSPF packets can include authentication. The authentication can consist of password (not secure) or digest authentication.

Digest authentication computes a one-way "hash" of the packet contents and the password using the MD5 algorithm. This way the password is not sent but the receiver can check that the sender knew the password and that the message contents were not altered.

Exercise 10: How does the receiver authenticate the message? Why should the hash function be one-way?