Introduction to routing protocols

Primary router responsibility:

- determining the best path
- forwarding packets

Router components:

1. CPU
2. RAM
    - OS
    - running config
    - IP routing table
    - ARP cache
    - packet buffer
3. ROM
    - bootstrap instructions
    - diagnostic soft
    - scaled-down OS
4. Flash memory (OS storage)
5. NVRAM (start-up config)

Bootup process:

1. Power-On Shelf Test (hardware testing)
2. Bootstrap loading (program to locate OS and load it to RAM)
3. OS loading
4. Config file loading (broadcast request if it doesn't exist)
    - interface addresses
    - routing info
    - passwords
    - etc

Interface - physical connector on router. Samples:

- LAN interface (RJ-45 UTP cable; ethernet encapsulation)
- WAN (serial; PPP, frame relay, high-level data link control)

Cases for establishing static routing:

- Network consists of few routers.
- Network is connected to internet through single ISP.

- Hub-and-spoke topology (central hub).

Dynamic routing protocols perform:

- network discovery
- update and maintain routing tables

Routing table principles:

- Router makes decision alone.
- Different routers do not always have the same information in routing table.
- Info about the path doesn't provide info about return path.

Routes are added only after interfaces are configured.

IP packet format:

- version
- IP header length
- destination & source address
- TTL
- protocol (upper-layer)
- etc

Ethernet frame fields:

- preamble
- source & destination address
- start-of-frame
- type/length
- etc

Hop-count - number of routers that packet must travel.

Bandwidth - data capacity of link.

Equal cost load balancing - using multiple interfaces with same metric to split data.

Unequal cost load balancing is used in EIGRP.

By default 4 equal paths are allowed.

ICMP unreachable message is sent is there is no appropriate record in routing table and no default route.

Network address is determined by adding IP address and subnet mask.

Cables can be *straight-through* or *crossover*.

Recursive lookup - multiple lookups in routing table before forwarding (e.g. to check connectivity and then determine next-hop).

If interface is down, all static routes to it are removed.

In static routing when exit interface is ethernet network, both exit interface and next-hop IP should be configured.

Route summarizing is used for merging different networks.

Cisco Discovery Protocol (CDP)

Getting information about directly connected devices. Operates on Layer 2.
Info messages are periodically sent to neighbors:

- type of connected devices with interfaces
- interfaces to make the connections
- model numbers

Information provided about each neighbor:

- device id
- address list (for network layer)
- port identifier
- capabilities list
- platform

Dynamic routing protocols

Compare to static routing:

- independent of network size
- advanced knowledge required
- automatically adapts to topology
- suitable for all topologies
- less secure
- uses more resources
- route depends on current topology

Routing protocol purposes:

- remote network discovery
- maintaining up-to-date routing information
- choosing best path
- ability to find new best path is current is unavailable

Routing domain - collection of routers under a common administration.
Interior gateway protocols are used within routing domain or individual networks.
Exterior gateway protocols are under control of different organisations.
Types of IGPs:

- Distance vector routing - routes are advertised as vectors of distance (e.g. hop-count metric) and direction (next-hop router or exit interface). Complete routing tables are sent to all neighbors periodically. Bellman-Ford or Ford-Fulkerson algorithm. Used when:
  - simple networks
  - not enough administrative knowledge
  - specific networks (e.g. hub-and-spokes)
  - worst-case convergence time are not a concern
- Link-state routing - creating complete view of the topology by gathering information from other routers. After network has converged updates are sent when there is topology change. Used when:
  - hierarchical network design
  - good administrative knowledge
  - fast convergence is crucial

Classful routing protocols don't send subnet mask info. Examples: IGRP, RIPv1.

Classless protocols include subnet mask in updates. Support discontiguous networks.

Convergence -state when all routing tables are consistent.

Metric is used to evaluate difference between available paths. Types:

- hop count - number of routers
- bandwidth
- load
- delay
- reliability - probability of failure
- cost - metric determined by OS or administrator

Administrative distance defines the preference of a routing source. The lower value is for more preferred route source.

Cold start - situation when routers know nothing about the connected devices.

Loopback interface - software-only interface that emulates physical interface. It is used by routing protocols, can be assigned an IP address.

Nullo interface simulates exit interface. In is always up and discards traffic. Shouldn't be created or configured.

Distance vector routing

Distance is number of hops to destination, direction represents exit interface.

RIP:

- Hop-count metric (16 equal infinity).
- Updates are broadcasted every 30 seconds.

IGRP:

- Bandwidth, delay, load and reliability as composed metric.

- Broadcasting updates every 90 seconds by default.

EIGRP:

- Unequal cost load balancing.
- Diffusing Update Algorithm to calculate the shortest path.
- No periodic updates.

Entire routing table is broadcasted to 255.255.255.255 regularly.
Routing protocols algorithm do the following:

- send and receive routing information
- calculate the best path
- detect and reach topology changes

Distance vector protocols features:

- simple implementation
- low resource requirements
- slow convergence
- limited scalability
- routing loops

Initial exchange - updates including only information about directly connected networks.

Speed of achieving convergence consists of:

- speed of propagating a change to neighbors
- speed of calculating best path using collected information

EIGRP uses bounded updates:

- non-periodic
- partial updates are sent only when topology is changed
- only routers that need the information are updated

Routing loop - condition in which packet is continuously transmitting without reaching the destination. May be a result of:

- Incorrectly configured static routes.
- Incorrectly configured route redistribution.
- Routing tables were not updated due to slow convergence.
- Incorrectly installed discard routes.

Count to infinity problem - situation when updates increase metric to infinity that is no longer reachable. Solution - setting maximum value of hops.

Split horizon rule - not advertising a network through the interface from which the update came. In other words, router send records that don't contain current exit interface in routing table.

Route poisoning is used to mark the route as unreachable and send updates to other routers.

Split horizon with poison reverse - principle to ensure that router with unreachable network is not susceptible to incorrect updates about that network.

Time to live (TTL) - 8-bit counter in IP header that limits number of hops the packet can traverse.

RIP version 1

Features:

- distance vector protocol
- the only metric is hop count
- routes with hop count greater than 15 are unreachable
- broadcasting every 30 seconds
- classful
- no VLSM & CIDR support
- administrative distance 120
- by default updates are sent to all RIP-configured interfaces (even if there is no RIP-device) (can be solved using passive interface)
- can have two exit interfaces with the same network address even if networks are discontiguous
- if no version type is specified, receives v1 and v2 updates

Timers:

- invalid (setting metric to 16 if update from current route hasn't been received for 180 seconds)
- flush (removing route after 240 seconds silence)
- holddown (unreachable route stays in this position for 180 seconds to let routers learn about the failure)

Holddown timer usage:

1. Router receives triggered update indicating that some network changed is no longer accessible.
2. Router marks the network as possibly down and starts the holddown timer.
3. If update with better metric is received, holddown timer is removed.
4. Other updates are discarded.
5. Packets are still forwarded.

Triggered updates (immediately sent):

- Interface changes state.
- Changing state to/from unreachable.
- Route is installed in routing table.

Message format:

1. data link frame header (MACs)
2. IP header packet (IPs, protocol field-17 for UDP)
3. UDP segment header (src/dst port = 520)
4. RIP message

- command (request/responce)
- version
- address family id (2 for IP, 0 for requesting full table)
- IP routes (25 maximum) + metric

Information exchange:

1. Each RIP-configured interface sends request for full routing table.
2. RIP-enabled neighbors send response.
3. Routes are added or updated if needed.
4. After updating triggered updates are sent to RIP-enabled interfaces.

When there is a packet collision (in hubs), RIP _JITTER contains randomly chosen time to wait before next update.

Boundary router has interfaces in more than one classful network.

Update rules:

- If the update and the received interface belong to the same major network, the subnet mask of the interface is applied to tie network in routing table.
- If networks differ, the classful mask of the network is applied.

VLSM and CIDR

Not supported by RIPv1 and IGRP.

Designed for more effective use of IP addresses.

Network classes:

- A (1-126 in first octet; subnet mask 255.0.0.0)
- B (128-191; 255.255.0.0)
- C (192-223; 255.255.255.0)

Significant bits refer to network, insignificant to host.

Variable length subnet mask (VLSM) can be thought as sub-subnetting.

- Conserves address space.
- Ability to specify a different subnet mask for the same network number and different subnets.

E.g. 10.0.0.0/8 can be divided in 256 networks 10.x.0.0/16.

Classless inter-domain routing (CIDR) - form of route summarizing.

- Reduces the number of entries in routing updates and routing tables.
- Reduces bandwidth utilization.

Algorithm (calculating a route summary):

1. List networks in binary format.
2. Count the number of left-most matching bits.
3. Copy the matching bit and add the rest zero bits.

Supernet - group of major network addresses summarized as single network.

RIP version 2

Features:

- classless
- next-hop addresses are included in updates
- multicasting instead of broadcasting (takes up less bandwidth and less processing of non-RIP devices)
- authentication option available (accepting packets from devices with the same password)
- holddown and other timers, triggered updates, split horizon, poisoning
- auto-summarizing by default as in RIPv1 (can be disabled)

RIPv2 message format:

  - command (request/responce)
  - version
  - address family id (2 for IP, 0 for requesting full table)
  - route tag
  - subnet mask
  - next hop address (used to identify a better hext-hop address than router address) (0.0.0.0 refers to the best next-hop address)
  - IP routes (25 maximum) + metric

The routing table

Source network types:

- directly connected
- static routes
- dynamic protocols

Level 1 route is a route with subnet mask equal or less than the classful mask of network address. Such routes can function as:

- default route - static route with the address 0.0.0.0/0
- supernet route (mask less than classful)
- network route (mask is equal to classful)

Ultimate route includes:

- either a next-hop IP address
- and/or an exit interface

Parent route doesn't contain any next-hop IP address or exit interface. It is automatically created when a route with a greater mask than the classful mask is entered into the routing table.

Level 2 route - route that is a subnet of a classful network address.

Parent route can be subnetted or *variably subnetted* (VLSM).

Routing behaviour (not the same as protocols) (affects lookup process only):

- classful
- classless

Route lookup process:

1. Examining 1 level routes.
   (a) If the best match is ultimate route then forward.
   (b) Otherwise (parent route) continue.
2. Examining child routes.
   (a) If there is a match then forward.
   (b) Otherwise continue.
3. If classful behavior then drop. If classless behavior then continue searching supernets , including the default route, with less match.
4. Forward if there is a match.
5. Drop.

EIGRP

Features:

- distance vector
- authentication allowed
- auto-summarizing by default

Message format:

1. Data link frame header
2. IP packet header
3. EIGRP packet header
4. Type/length/values (TLV) types

EIGRP packet header fields:

- opcode
  - update
  - query and reply
  - reply

    - hello
- autonomous system (AS) number - number to track multiple EIGRP instances

TLV:

- parameters
    - weights for composite metric (only bandwidth and delay by default, equally weighted)
    - hold timer (time for neighbors to wait before considering the advertising router down)
- IP internal (used for advertising routes within an AS)
    - delay (sum of delays from src to dst in units of 10 microseconds)
    - bandwidth (lowest configured bandwidth of any interface along the route)
    - subnet mask (prefix length)
    - destination address (24bits + additional 32bits if needed)
- IP external (used for importing external routes into routing process)

EIGRP is capable of several different routing protocols IP, IPX, AppleTalk using protocol-dependent modules (PDM).

Reliable Transport Protocol (RTP) - protocol for exchanging information packets. Can send packets either multicast or unicast.

Packet types:

- hello (discover neighbors and form adjacency) (multicast unreliable delivery)
- update (propagate routing information is sent only when necessary and to routers that need it) (multicast or unicast dependent on quantity)
- acknowledgement (when reliable delivery is used; contain a nonzero ack number) (always unicast)
- query and reply (used by DUAL; always reliable delivery) (queries are multicast, responses are unicast)

Hello's are sent every 5 seconds (60 on slow connections).

Holdtime is three times the hello interval.

DUAL finite state machine (FSM) track all routes, uses efficient loop-free least cost path.

Wildcard mask - inverse subnet mask.

Null0 is included as a child route if:

- There is at least one EIGPR-learned subnet.
- Auto-summarizing is enabled.

Null0 is always selected if there is no match regardless the classless behavior.

Metric consists of:

- bandwidth (link characteristic)
- delay (set by administrator)
- reliability (probability of fail)

- load (amount of traffic utilizing a link)

$EIGRP metric = \left(\frac{10.000.000}{\text{bandwidth kbps}} + \frac{\text{sum of delays}}{10}\right) * 256$

Diffusing update algorithm (DUAL) (uses topology table and neighbor table to build the routing table):

- Successor - neighboring router that is used for packet forwarding and is the least-cost to the destination network).
- Feasible distance (FD) - the lowest calculated metric to reach the destination.
- Feasible successor (FS) - neighbor who has a loop-free backup path to the same network as successor by satisfying the FC.
- Reported distance (RD) - total metric along a path to destination (neighbor's FD).
- Feasibility condition (FC) is met when a neighbor's RD is less than the local router FD.

Topology table holds information about the successor, FD and any FS with their RD.

Passive state means that DUAL is not performing computations.

Link-state protocols

Dijkstra's algorithm (shortest path first).

Routing process:

1. learning about directly connected networks (detecting that interface is up)
2. saying hello to neighbors on directly connected networks
3. building link-state packet (LSP) containing the state of directly connected links
4. flooding LSP to all neighbors, who then store it in database
5. using database to construct a complete map of the topology and compute the best path to each destination network

LSP includes:

- interface IP and subnet mask
- type of network (ethernet, point-to-point)
- link cost
- any neighbor routers on that link
- sequence number and aging info (help to manage flooding and to keep the database up-to-time)

Unlike distance vector protocols, SPF is calculated after the flooding is complete.

LSPs are sent when:

- initial startup
- topology change

Overall advantages:

- building a topological map
- fast convergence
- event-driven updates
- hierarchical design (concept of areas)

IS-IS is mainly used by ISPs and carriers.

OSPF

Features:

- link-state routing protocol
- AD 110
- authentication is allowed

Message format:

1. Data link frame header
2. IP packet header
3. OSPF packet header
4. OSPF packet type-specific data

Packet types:

- hello
- database description (DBD) (check for database synchronization between routers)
- link state request (LSR) (request for specific database records)
- link state update (LSU) (reply to LSR)
- link state acknowledgment (LSAck) (confirm receipt of the LSU)

Usage of hello packets:

- discover neighbors and establish adjacencies
- advertise parameters on which routers must agree to become neighbors
- elect designated router (DR) and backup DR (BDR) on multiaccess networks

OSPF packet fields:

- type (hello, DD, LSR, LSU, LSAck)
- router ID
- area ID
- network mask (associated with sending interface)

- hello interval
- router priority
- DR ID if any
- BDR ID if any
- list of neighbors (IDs)

Hello interval - time before hellos (10s on multiaccess and point-to-point, 30 on non-broadcast multiaccess segments e.g. frame relay).

Dead interval - time to wait until neighbor is considered down if there is no hellos. 4 times the hello interval.

To establish connection network type, dead and hello intervals must be the same.

DR is responsible for updating all other routers when a change occurs in the multiaccess network.

LSA contains route information for destination networks.

Each router has link-state database containing LSAs from neighbors.

OSPF area - group of routers that share link-state info.

Multiarea is used for isolating unstable networks and for storing smaller databases.

How to choose router ID:

1. configured IP address
2. if not, choose highest IP of any loopbacks
3. if not, highest active IP of any physical interface

Router ID is chosen with first 'network' command and changes only if ospf process is killed.

Flapping link - network that cycles between up and down.

SPF shedule delay - 5 seconds delay after receiving an LSU before running SPF. Used to minimize flapping link problem.

Hold time - 10 seconds delay between rerunning SPF.

Retransmit interval is time before repeated advertisement is sent if there was no ack).

$OSPF Cisco metric $= \frac{10^8}{\text{bps}}$$

The cost of OSPF route is accumulated value from one router to the destination.

Multiaccess network - network with more than 2 devices on the same shared media (e.g. ethernet LAN).

OSPF network types:

- point-to-point
- broadcast multiaccess
- nonbroadcast multiaccess (NBMA)
- point-to-multipoint
- virtual links

Challenges in multiaccess networks:

- creation of multiple adjacencies $Total adjacencies $= \frac{n(n-1)}{2}$$

- extensive flooding of LSAs (the problem is to introduce each device to other devices)

DR and BDR are chosen to solve flooding problem (only in multiaccess).

1. DR is a router with highest interface priority.
2. BDR is a router second highest interface priority.
3. If priorities are equal ther compare by router IDs.
4. Other router are marked as DROther.

Neighbor states:

- full
- 2way (two drothers form adjacency)

DR remains RD until:

- DR fails
- OSPF process on it fails
- multiaccess interface on it fails

Routers with priority 0 will never be DR or BDR.

*Autonomous system boundary route*r (ASBR) is located between OSPF routing domain and non-OSPF network.

OSPF area types:

- Backbone area forms the kernel. It should distribute routing information between non-backbone areas.
- Standard area is created by default. It receives link updates, summary routes and external routes.
- Stub area don't receive information about external routes but receives it from other areas.
- Totally stubbed area don'r receive information from other areas. Cannot contain ASBR.

LSA types:

1. router link state update
2. network link state
3. summary network LSA
4. ASBR summary
5. AS external
6. + some other ...