❖ROUTING and RIP

Router: a device that acts as a junction between two networks to transfer data packets among them.

Routing: the process of forwarding packets hop-by-hop through routers to reach their destination

Routing protocol: A set of messages, rules, and algorithms used by routers for the overall purpose of learning routes. This process includes the exchange and analysis of routing information. Each router chooses the best route to each subnet (path selection) and finally places those best routes in its IP routing table. Examples include RIP, EIGRP, OSPF, and BGP.

■ **Routed protocol and routable protocol:** Both terms refer to a protocol that defines a packet structure and logical addressing, allowing routers to forward or route the packets.

Routers forward packets defined by routed Interior and Exterior Routing Protocols IP routing protocols fall into one of two major categories:

Interior gateway protocols (IGP) or exterior gateway protocols (EGP).

The definitions of each are as follows:

■ **IGP:** A routing protocol that was designed and intended for use inside a single autonomous system (AS)

■ **EGP:** A routing protocol that was designed and intended for use between different autonomous systems

- A static routing table's entries are updated manually by an administrator; a dynamic routing table's entries are updated automatically by a routing protocol.
- A metric is the cost assigned for passage of a packet through a network.
- An autonomous system (AS) is a group of networks and routers under the authority of a single administration.
- RIP is based on distance vector routing, in which each router shares, at regular intervals, its knowledge about the entire AS with its neighbors.
- Two shortcomings associated with the RIP protocol are slow convergence and instability. Procedures to remedy RIP instability include triggered update, split horizons, and poison reverse.
- RIP uses the Bellman-Ford Distance Vector algorithm to determine the best “path” to a particular destination.
- RIPv1 (RFC 1058) is classful and RIPv2 (RFC 2543) is classless.
- RIPv1 routers will receive both Version 1 and 2 updates.
- RIP sends out periodic routing updates (every 30 seconds)
- RIP uses a form of distance as its metric (in this case, hop count)
- RIP supports IP and IPX routing.

- RIP utilizes UDP port 520.
- RIP routes have an administrative distance of 120.
- RIP has a maximum hop count of 15 hops.
- A metric of 16 hops in RIP is considered a poison route or infinity metric.
- RIP uses a round-robin system of load-balancing between equal metric routes, which can lead to pinhole congestion.
- RIPv1 does not support Variable Length Subnet Masks (VLSMs).
- RIPv1 sends updates as broadcasts to address 255.255.255.255.
- RIPv1 (RFC 1058) is classful, and thus does not include the subnet mask with its routing table updates.
- Routing updates are sent via multicast, using address 224.0.0.9
- RIPv2 can interoperate with RIPv1.
- Both routers will continue to increment the metric for the network until they reach a hop count of 16, which is unreachable. This behavior is known as counting to infinity.
- There are several loop avoidance mechanisms:
- **Split-Horizon –** Prevents a routing update from being sent out the interface it was received on. In our above example, this would prevent Router A from sending an update for the 172.18.0.0 network back to Router B, as it originally learned the route from Router B. Split-horizon is enabled by default on Cisco Routers.
- **Route-Poisoning –** Works in conjunction with split-horizon, by triggering an automatic update for the failed network, without waiting for the update timer to expire. This update is sent out all interfaces with an infinity metric for that network.
- **Hold-Down Timers –** Prevents RIP from accepting any new updates for routes in a hold-down state, until the hold-down timer expires. If Router A sends an update to Router B with a higher metric than what is currently in Router B's routing table, that route will be placed in a hold-down state. (Router A's metric for the 172.18.0.0 network is 1; while Router B's metric is 0).
- RIPv2 (RFC 2543) is classless, and thus does include the subnet mask with its routing table updates.
- RIP has four basic timers:
  - ❖ Update Timer (default 30 seconds)
  - ❖ Invalid Timer (default 180 seconds)
  - ❖ Hold-down Timer (default 180 seconds)
  - ❖ Flush Timer (default 240 seconds)

❖OSPF

- The routing algorithm running within an autonomous system is called an intra autonomous system routing protocol.
- OSPF is a link-state protocol that uses flooding of link-state information and a Dijkstra's least-cost path algorithm.
- OSPF divides an AS into areas, defined as collections of networks, hosts, and routers.
- OSPF is based on link state routing, in which each router sends the state of its neighborhood to every other router in the area. A packet is sent only if there is a change in the neighborhood.
- OSPF routing tables are calculated by using Dijkstra's algorithm.
- OSPF sends updates (LSAs) when there is a change to one of its links, and will only send the change in the update. LSAs are additionally refreshed every 30 minutes.
- OSPF employs a hierarchical network design using Areas.
- OSPF will form neighbor relationships with adjacent routers in the same Area.
- OSPF supports only IP routing.
- OSPF traffic is multicast either to address 224.0.0.5 (all OSPF routers) or 224.0.0.6 (all Designated Routers).
- OSPF routes have an administrative distance is 110.
- OSPF uses cost as its metric, which is computed based on the bandwidth of the link.
- OSPF also has a Dead Interval, which indicates how long a router will wait without hearing any hellos before announcing a neighbor as "down."
- Instead of advertising the distance to connected networks, OSPF advertises the status of directly connected links using Link-State Advertisements (LSAs).
- Default for the Dead Interval is 40 seconds for broadcast and point-to-point interfaces, and 120 seconds for non-broadcast and point-to-multipoint interfaces.
- Hello packets are sent out OSPF-enabled interfaces every 10 seconds for broadcast and point-to-point interfaces, and 30 seconds for non-broadcast and point-to-multipoint interfaces.
- The Hello packets also serve as keepalives to allow routers to quickly discover if a neighbor is down.
- BDR stands for Backup Designated Router.
- At 2-way neighbor state of OSPF are Designated and Backup Designated Routers elected.
- In ExStart state of OSPF are master/slave relationships formed between routers.
- DBDs stands for Database Descriptors.
- Area 0 is required for OSPF to function, and is considered the "Backbone" area.
- OSPF routers can belong to multiple areas, and will thus contain separate Topology databases for each area. These routers are known as Area Border Routers (ABRs).

- The ABRs and ASBRs of Standard areas do not automatically generate (or inject) default routes into the area.
- The ABRs of Stub and Totally Stubby areas automatically generate (and inject) a default route (0.0.0.0/0) into the area.
- The sequence number and checksum for each entry is provided by Topology Table.
- OSPF is a link-state protocol that uses flooding of link-state information and a Dijkstra's least-cost path algorithm.
- Only data in a datagram is fragmented.
- OSPF is a standardized Link-State routing protocol, designed to scale efficiently to support larger networks.
- OSPF will form neighbor relationships with adjacent routers in the same Area.
- OSPF is a classless protocol, and thus supports VLSMs.
- The router with the highest priority becomes the DR; second highest becomes the BDR.
- The area 1 stub command must be configured on all routers in the Stub area.
- Totally Stubby areas will not accept Type 3 LSAs to other areas.
- At the network layer, a global identification system that uniquely identifies every host and router is necessary for delivery of a packet from host to host.
- OSPF will elect a Designated Router (DR) for each multi access networks, accessed via multicast address 224.0.0.6.
- **The OSPF process builds and maintains three separate tables: A neighbor table, topology table, routing table.**

  The OSPF process builds and maintains three separate tables:

  - **A neighbor table –** contains a list of all neighboring routers.
  - **A topology table –** contains a list of all possible routes to all known networks within an area.
  - **A routing table –** contains the best route for each known network.
- OSPF routers will only become neighbors if the following parameters within a Hello packet are identical on each router:
  - Area ID
  - Area Type (stub, NSSA, etc.)
  - Prefix
  - Subnet Mask
  - Hello Interval
  - Dead Interval
  - Network Type (broadcast, point-to-point, etc.)
  - Authentication

- Designated and Backup Designated Routers are elected at 2-WAY stage.

- In ExStart state Master/slave relationships are formed between routers to determine who will begin the exchange.

- A router can become an ASBR in one of two ways:
  - By connecting to a separate Autonomous System, such as the Internet
  - By redistributing another routing protocol into the OSPF process.
- the four separate OSPF router types are as follows:
  - **Internal Routers -** all router interfaces belong to only one Area.
  - **Area Border Routers (ABRs) -** contains interfaces in at least two separate areas.
  - **Backbone Routers -** contain at least one interface in Area 0.
  - **Autonomous System Border Routers (ASBRs) -** contain a connection to a separate Autonomous System.
- OSPF forms neighbor relationships, called adjacencies, with other routers in the same Area by exchanging Hello packets to multicast address 224.0.0.5. Only after an adjacency is formed can routers share routing information.
- Each OSPF router is identified by a unique Router ID.
- **The Router ID can be determined in one of three ways:**
  - The Router ID can be manually specified.
  - If not manually specified, the highest IP address configured on any Loopback interface on the router will become the Router ID.
  - If no loopback interface exists, the highest IP address configured on any Physical interface will become the Router ID.
- By default, Hello packets are sent out OSPF-enabled interfaces every 10 seconds for broadcast and point-to-point interfaces, and 30 seconds for non-broadcast and point-to-multipoint interfaces.
- OSPF also has a Dead Interval, which indicates how long a router will wait without hearing any hellos before announcing a neighbor as “down.” Default for the Dead Interval is 40 seconds for broadcast and point-to-point interfaces, and 120 seconds for non-broadcast and point-to-multipoint interfaces.

- four types of OSPF areas:
  - Standard areas
  - Stub areas
  - Totally Stubby areas
  - Not So Stubby areas (NSSA)

- The Neighbor Table provides the following information about each neighbor:
  - The Router ID of the remote neighbor.
  - The OSPF priority of the remote neighbor (used for DR/BDR elections).

- The current neighbor state.
- The dead interval timer.
- The connecting IP address of the remote neighbor.
- The local interface connecting to the remote neighbor.

- The Topology Table provides the following information:
  - The actual link (or route).
  - The advertising Router ID.
  - The link-state age timer.
  - The sequence number and checksum for each entry.

- **LSA types exist:**
  - Router LSA (Type 1) – Contains a list of all links local to the router, and the status and "cost" of those links. Type 1 LSAs are generated by all routers in OSPF, and are flooded to all other routers within the local area.
  - Network LSA (Type 2) – Generated by all Designated Routers in OSPF, and contains a list of all routers attached to the Designated Router.
  - Network Summary LSA (Type 3) – Generated by all ABRs in OSPF, and contains a list of all destination networks within an area. Type 3 LSAs are sent between areas to allow inter-area communication to occur.
  - ASBR Summary LSA (Type 4) – Generated by ABRs in OSPF, and contains a route to any ASBRs in the OSPF system. Type 4 LSAs are sent from an ABR into its local area, so that Internal routers know how to exit the Autonomous System.
  - External LSA (Type 5) – Generated by ASBRs in OSPF, and contain routes to destination networks outside the local Autonomous System. Type 5 LSAs can also take the form of a default route to all networks outside the local AS. Type 5 LSAs are flooded to all areas in the OSPF system.

- **OSPF Neighbor States:**
  - Down – indicates that no Hellos have been heard from the neighboring router.
  - Init – indicates a Hello packet has been heard from the neighbor, but two way communication has not yet been initialized.
  - 2-Way – indicates that bidirectional communication has been established. Recall that Hello packets contain a neighbor field. Thus, communication is considered 2-Way once a router sees its own Router ID in its neighbor's Hello Packet. Designated and Backup Designated Routers are elected at this stage.
  - ExStart – indicates that the routers are preparing to share link state information. Master/slave relationships are formed between routers to determine who will begin the exchange.

- ❖ Exchange – indicates that the routers are exchanging Database Descriptors (DBDs). DBDs contain a description of the router's Topology Database. A router will examine a neighbor's DBD to determine if it has information to share.
- ❖ Loading – indicates the routers are finally exchanging Link State Advertisements, containing information about all links connected to each router. Essentially, routers are sharing their topology tables with each other.
- ❖ Full – indicates that the routers are fully synchronized. The topology table of all routers in the area should now be identical. Depending on the "role" of the neighbor, the state may appear as: • Full/DR – indicating that the neighbor is a Designated Router (DR)
    - Full/BDR – indicating that the neighbor is a Backup Designated Router (BDR)
    - Full/DROther – indicating that the neighbor is neither the DR or BDR

- **OSPF Network Types:**
    - ❖ Broadcast Multi-Access – indicates a topology where broadcast occurs.
        - Examples include Ethernet, Token Ring, and ATM.
        - OSPF will elect DRs and BDRs.
        - Traffic to DRs and BDRs is multicast to 224.0.0.6. Traffic from DRs and BDRs to other routers is multicast to 224.0.0.5.
        - Neighbors do not need to be manually specified.

    - ❖ Point-to-Point – indicates a topology where two routers are directly connected.
        - An example would be a point-to-point T1.
        - OSPF will not elect DRs and BDRs.
        - All OSPF traffic is multicast to 224.0.0.5.
        - Neighbors do not need to be manually specified.

    - ❖ Point-to-Multipoint – indicates a topology where one interface can connect to multiple destinations. Each connection between a source and destination is treated as a point-to-point link.
        - An example would be Point-to-Multipoint Frame Relay.
        - OSPF will not elect DRs and BDRs.
        - All OSPF traffic is multicast to 224.0.0.5.
        - Neighbors do not need to be manually specified.

    - ❖ Non-broadcast Multi-access Network (NBMA) – indicates a topology where one interface can connect to multiple destinations; however, broadcasts cannot be sent across a NBMA network.
        - An example would be Frame Relay.

- OSPF will elect DRs and BDRs.
- OSPF neighbors must be manually defined, thus All OSPF traffic is unicast instead of multicast.

❖BGP

- BGP is considered a “Path Vector” routing protocol.
- BGP utilizes TCP for reliable transfer of its packets, on port 179.
- BGP is an interautonomous system routing protocol used to update routing tables.
- BGP is based on a routing protocol called path vector routing. In this protocol, the ASs through which a packet must pass are explicitly listed.
- A routing protocol that was designed and intended for use inside a single autonomous system (AS) is known as Interior gateway protocols (IGP).
- A routing protocol that was designed and intended for use between different autonomous systems exterior gateway protocols (EGP).
- The Administrative Distance for routes learned outside the Autonomous System (eBGP routes) is 20, while the AD for iBGP and locally-originated routes is 200.
- KEEPALIVE messages are sent periodically (every 60 seconds by default) to ensure that the remote peer is still available.
- UPDATE messages are used to exchange routes between peers.
- Standard attributes supported by all BGP implementations, and are optionally included BGP updates is known as Well-known Discretionary.
- By default, a route originated on the local router will be assigned a weight of 32768.
- The Local Preference attribute is applied to inbound external routes, dictating the best outbound path.
- The default Local preference value is 100.
- The MED (MultiExit Discriminator) attribute is applied to outbound routes, dictating the best inbound path into the AS (assuming multiple paths exist).
- The BGP attribute that identifies the list of traversed AS’s to reach a particular destination is called the AS_PATH.
- The Next Hop attribute in BGP specifies the IP address of the next hop to reach a particular destination.
- The BGP attribute that identifies the originator of the route is called origin.
- Local Preference is a well-known discretionary attribute that provides a preference for determining the best path for outbound traffic
- The Multi-Exit-Discriminator (MED) attribute provides a preference to eBGP peers to a specific inbound router.
- The OPEN message is sent between BGP peers to initiate the session.
- The initial BGP state in the Finite-State Machine (FSM) is Idle.
- After a TCP connection is established and an OPEN message is sent, BGP moves to the OpenSent state.

- There are two types of BGP neighbor relationships:
  - iBGP Peers – BGP neighbors within the same autonomous system.
  - eBGP Peers – BGP neighbors connecting separate autonomous systems.

- BGP **subcategories** of attributes:
  - Well-known Mandatory - Standard attributes supported by all BGP implementations, and always included in every BGP update.
  - Well-known Discretionary - Standard attributes supported by all BGP implementations, and are optionally included BGP updates.
  - Optional Transitive - Optional attribute that may not be supported by all implementations of BGP. Transitive indicates that a non compliant BGP router will forward the unsupported attribute unchanged, when sending updates to peers.
  - Optional Non-Transitive - Optional attribute that may not be supported by all implementations of BGP. Non-Transitive indicates that a non-compliant BGP router will strip out the unsupported attribute, when sending updates to peers.

- BGP Attributes :
  - AS-Path (well-known mandatory) – Identifies the list (or path) of traversed AS's to reach a particular destination.
  - Next-Hop (well-known mandatory) – Identifies the next hop IP address to reach a particular destination.
  - Origin (well-known mandatory) – Identifies the originator of the route. • Local Preference (well-known, discretionary) – Provides a preference to determine the best path for outbound traffic.
  - Atomic Aggregate (well-known discretionary) – Identifies routes that have been summarized, or aggregated.
  - Aggregator (optional transitive) – Identifies the BGP router that performed an address aggregation.
  - Community (optional transitive) – Tags routes that share common characteristics into communities.
  - Multi-Exit-Discriminator (MED) (optional non-transitive) – Provides a preference to eBGP peers to a specific inbound router.
  - Weight (Cisco Proprietary) – Similar to Local Preference, provides a local weight to determine the best path for outbound traffic.

As a BGP peer session is forming, it will pass through several states. This process is known as the BGP Finite-State Machine (FSM):

- Idle – the initial BGP state
- Connect - BGP waits for a TCP connection with the remote peer. If successful, an OPEN message is sent. If unsuccessful, the session is placed in an Active state.
- Active – BGP attempts to initiate a TCP connection with the remote peer. If successful, an OPEN message is sent. If unsuccessful, BGP will wait for a ConnectRetry timer to expire, and place the session back in a Connect State.
- OpenSent – BGP has both established the TCP connection and sent an OPEN Message, and is awaiting a reply OPEN Message. Once it receives a reply OPEN Message, the BGP peer will send a KEEPALIVE message.
- OpenConfirm – BGP listens for a reply KEEPALIVE message.
- Established – the BGP peer session is fully established. UPDATE messages containing routing information will now be sent.

- There are three ways to originate a prefix (in other words, advertise a network) into BGP:
  - By using network statements
  - By using aggregate-address statements
  - By redistributing an IGP into BGP

BGP Path Selection process.

- **Weight –** Prefer the path with the HIGHEST weight (Cisco-proprietary parameter)
- **Local Preference –** Prefer the path with the HIGHEST local preference.
- **Locally Originated –** Prefer local-originated path (prefix) over one's learnt from a neighbor.
- **AS Path** – Prefer the path with the SHORTEST AS Path.
- **Origin Type –** Prefer the path with LOWEST origin type. IGP is lower than EGP. EGP is lower than INCOMPLETE.
- **MED –** Prefer the path with the LOWEST MED (Multi-Exit Discriminator).
- **eBGP over iBGP –** Prefer eBGP learned routes over iBGP learned routes.
- **IGP Metric –** Prefer the path with the LOWEST IGP next-hop.
- **Multipath –** Determines if multiple paths are required.
- **External Paths –** When both paths are external prefer OLDEST path.
- **Router ID –** Prefer path that comes from the BGP router with the LOWEST ID.
- **Neighbor Address –** Prefer path that comes from the LOWEST neighbor address.

❖IPV4-IPV6

- An IPv4 address is 32 bits long and uniquely and universally defines a host or router on the Internet.
- In classful addressing, the portion of the IP address that identifies the network is called the netid.
- In classful addressing, the portion of the IP address that identifies the host or router on the network is called the hostid.
- An IP address defines a device's connection to a network.
- There are five classes in IPv4 addresses. Classes A, B, and C differ in the number of hosts allowed per network. Class D is for multicasting and Class E is reserved.
- The class of an address is easily determined by examination of the first byte.
- Addresses in classes A, B, or C are mostly used for unicast communication.
- Addresses in class D are used for multicast communication.
- Subnetting divides one large network into several smaller ones, adding an intermediate level of hierarchy in IP addressing.
- Supernetting combines several networks into one large one.
- In classless addressing, we can divide the address space into variable-length blocks.
- There are three restrictions in classless addressing:
- The number of addresses needs to be a power of 2.
- The mask needs to be included in the address to define the block.
- The starting address must be divisible by the number of addresses in the block.
- The mask in classless addressing is expressed as the prefix length *(In)* in CIDR notation.
- To find the first address in a block, we set the rightmost 32 - *n* bits to O.
- To find the last address in the block, we set the rightmost 32 - *n* bits to O.
  Subnetting increases the value of *n.*
- The global authority for address allocation is ICANN. ICANN normally grants large blocks of addresses to ISPs, which in tum grant small subblocks to individual customers.
- IPv4 is an unreliable connectionless protocol responsible for source-to-destination delivery.
- Packets in the IPv4 layer are called datagrams. A datagram consists of a header (20 to 60 bytes) and data. The maximum length of a datagram is 65,535 bytes.
- The MTU is the maximum number of bytes that a data link protocol can encapsulate. MTUs vary from protocol to protocol.
- Fragmentation is the division of a datagram into smaller units to accommodate the MTU of a data link protocol.
- A packet has arrived with an M bit value of 0 is the last fragment.
- The IPv4 datagram header consists of a fixed, 20-byte section and a variable options section with a maximum of 40 bytes.

- The options section of the IPv4 header is used for network testing and debugging.
- The six IPv4 options each have a specific function. They are as follows: filler between options for alignment purposes, padding, recording the route the datagram takes, selection of a mandatory route by the sender, selection of certain routers that must be visited, and recording of processing times at routers.
- IPv6, the latest version of the Internet Protocol, has a 128-bit address space, a revised header format, new options, an allowance for extension, support for resource allocation, and increased security measures.
- An IPv6 datagram is composed of a base header and a payload.
- Extension headers add functionality to the IPv6 datagram.
- Three strategies used to handle the transition from version 4 to version 6 are **dual stack, tunneling, and header translation.**
- IPv6 addresses use hexadecimal colon notation with abbreviation methods available
- There are three types of addresses in IPv6: unicast, anycast, and multicast.
- In an IPv6 address, the variable type prefix field defines the address type or purpose.
- The delivery of a packet is called direct if the deliverer (host or router) and the destination are on the same network; the delivery of a packet is called indirect if the deliverer (host or router) and the destination are on different networks.
- In the next-hop method, instead of a complete list of the stops the packet must make, only the address of the next hop is listed in the routing table; in the network specific method, all hosts on a network share one entry in the routing table.
- In the host-specific method, the full IP address of a host is given in the routing table.
- In the default method, a router is assigned to receive all packets with no match in the routing table.
- The routing table for classless addressing needs at least four columns.
- Address aggregation simplifies the forwarding process in classless addressing.
- Longest mask matching is required in classless addressing.
- Classless addressing requires hierarchical and geographical routing to prevent immense routing tables.