Module 4
(Lecture – 7)

(Network Layer: Router architecture; Internet Protocol (IP) - Forwarding and Addressing in the Internet; Routing algorithms - Link-state routing, Distance vector routing, Hierarchical routing; Routing in the Internet - RIP, OSPF, BGP; Broadcast & multicast routing; ICMP; Next Generation IP - IPv6)

Dr. Nirnay Ghosh

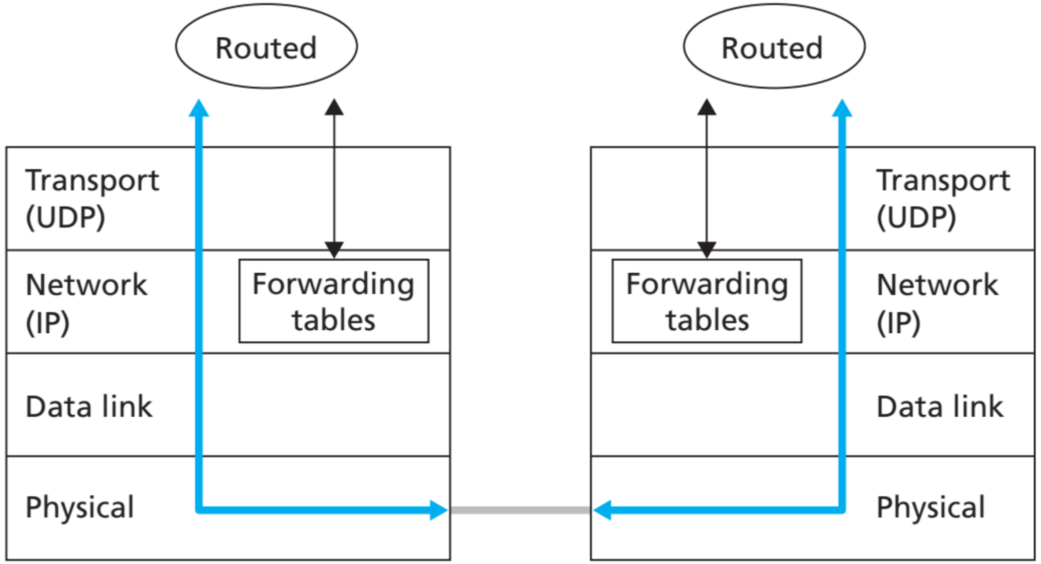
Assistant Professor

Department of Computer Science & Technology

IIEST, Shibpur

Intra-AS Routing in the Internet: RIP

- RIP – included in the Berkeley Software Distribution (BSD) version of the UNIX supporting TCP/IP
- Typically implemented as an application layer process – *"routed"* (route-dee)
  - Runs in neighboring routers
  - Uses UDP Port 520
  - Capable of manipulating the routing tables within the UNIX kernel
  - Sends and receives messages over a standard socket
  - RIP request/response: implemented as UDP segment
  - IP datagram carries the UDP segment



Implementation of RIP as the *routed* (route dee) daemon

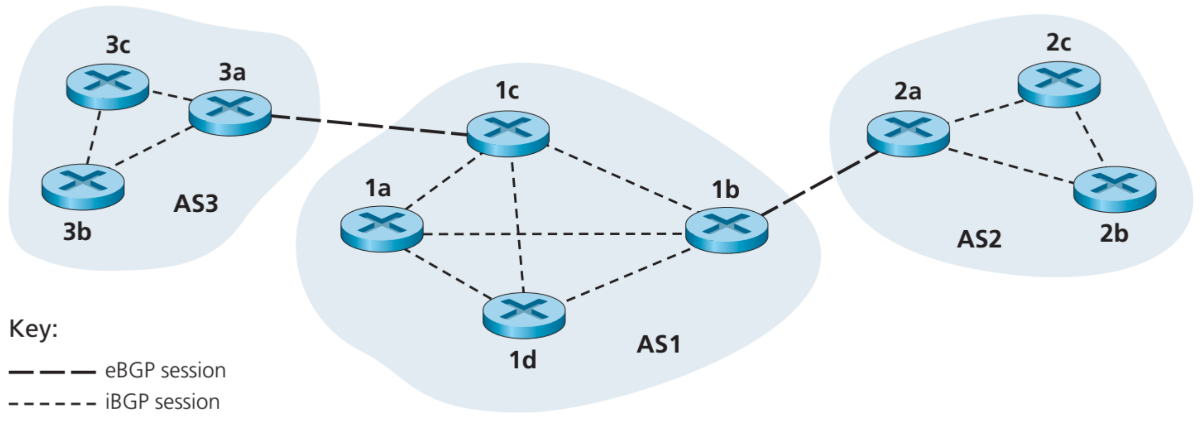
Intra-AS Routing in the Internet: OSPF Protocol

- OSPF – Open Shortest Path First
  - "Open" – routing protocol specification is publicly available
- Improves RIP – introduces a number of advanced features
- Uses Link-State (LS) routing: flooding/broadcasting (OSPF advertisement) of link-state information to *all* routers periodically (interval of 30 secs)
- Each router locally runs Dijkstra's shortest path algorithm
  - Link cost metric: to be determined by the network administrator
  - Generate a shortest-path tree to all subnets – complete topological map of the entire AS
- OSPF advertisement (contained in OSPF message) – carried directly by IP
- No intervention of transport layer protocols
  - Built-in functionalities – reliable message transfer, link-state broadcast

- Advanced features of OSPF
  - Security
    - Routers authenticate exchanged link-state information - preconfigured shared secret keys
    - Sending router: generates MD5 hash of the OSPF packet – appends its secret key - includes the resulting hash value in the packet
    - Receiving router: uses its preconfigured secret key to compute the MD5 hash of the packet – compares with the included hash - verifies packet's authenticity
  - Multiple same-cost paths: allows usage of multiple paths with same cost to the destination
  - Integrated support for unicast & multicast routing
  - Support for hierarchy within a single routing domain:
    - configures any OSPF autonomous system hierarchically into areas
    - Each area runs its own OSPF link-state routing algorithm – broadcasting link-state to *all* routers within the area
    - Area border router: responsible for routing packets outside the area
    - Backbone area: exactly one OSPF area in the AS configured to be the backbone – routes traffic between other areas in the AS (inter-area routing within AS)

Inter-AS Routing in the Internet: BGP

- BGP – Border Gateway Protocol Version 4 (BGP/BGP4) : de facto standard for inter-AS routing protocol in today's Internet
- BGP allows each subnet to advertise its existence to the rest of the Internet
- Services provided by BGP to each AS:
  - Obtain subnet (CIDR-ized prefixes) reachability information from neighboring AS
  - Propagate reachability information to all routers internal to the AS
  - Determine "good" routes to subnets based on reachability information and AS policy
- BGP peers: pair of routers exchanging routing information (both within an AS and in two different ASs)
  - Creates a TCP connection at port 179
  - Used to send all BGP messages in sessions – two types – external BGP (eBGP) and internal BGP (iBGP)

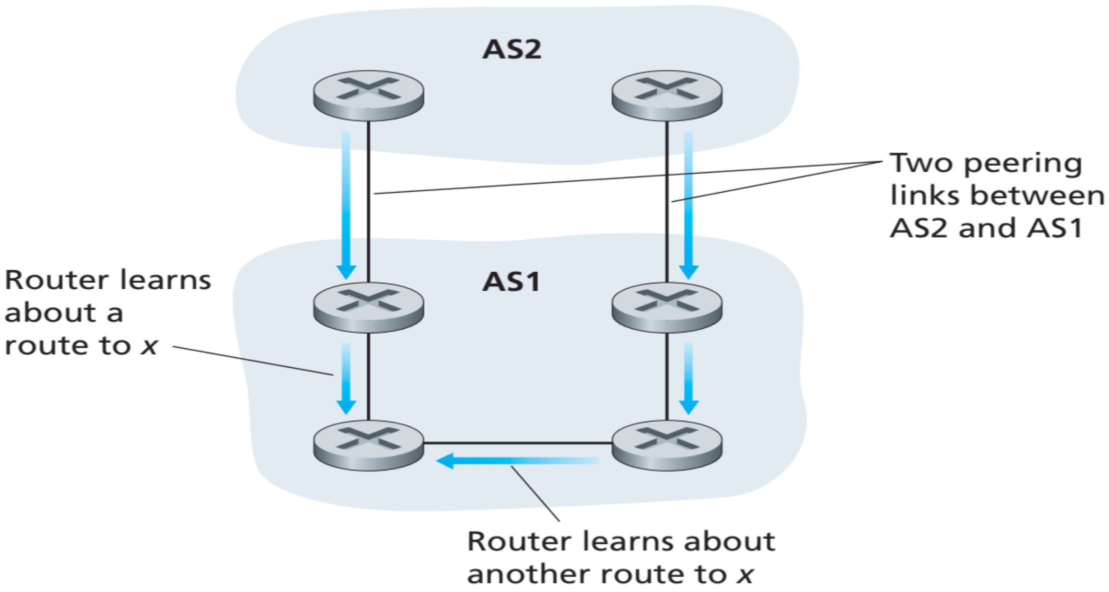


eBGP and iBGP Sessions

- Gateway router(s) (in any AS) uses eBGP/iBGP sessions to gather and propagate subnet reachability information
  - eBGP session: used to receive eBGP-learned prefixes from other ASs
  - iBGP session: used to distribute and re-distribute the prefixes to other routers in the AS
- Other routers create a new entry for this prefix in their forwarding tables

Inter-AS Routing in the Internet: BGP

- Path attributes & BGP Routes
  - Autonomous System Number (ASN): globally unique number to identify an AS
  - BGP Attributes: included while a router advertises a prefix across a BGP session (also termed as *route*)
    - AS-PATH: contains a list of ASs through which an prefix advertisement has passed – a router rejects an advertisement if its ASN is contained in the path list
    - NEXT-HOP: contains the router interface that initiates the AS-PATH – used by routers to properly configure their forwarding tables
    - Preference metric: assigned by the routers to give weights to different routes
- Two routes can have the same AS-PATH to a prefix but different NEXT-HOP values (i.e., different peering link)
  - Router can determine the cost of each peering link
  - Apply *hot-potato routing* to determine the appropriate interface



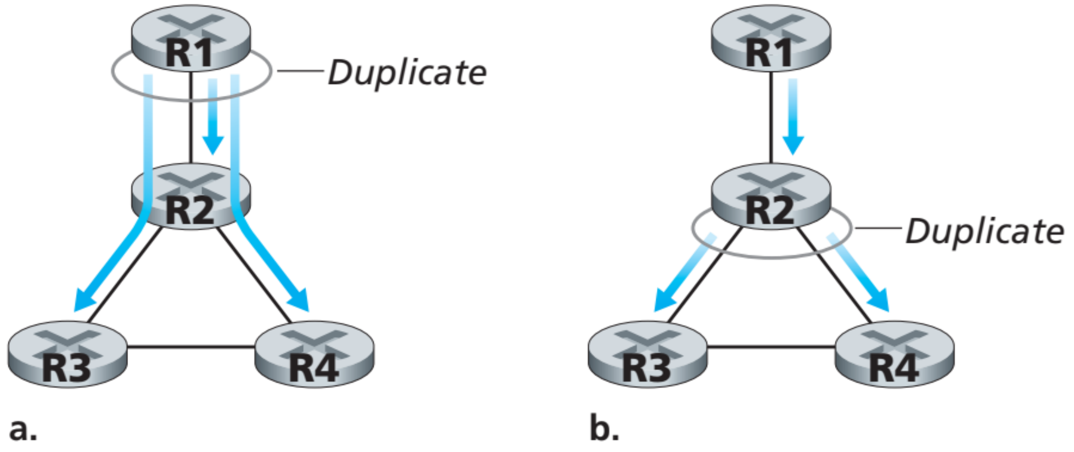
Same AS-PATH but different NEXT-HOP Values in Advertisements

- BGP Route Selection Policy
  - The routes with the highest local preference values are selected
  - From the remaining routes (all with the same local preference value) the route with shortest AS-PATH is selected
  - From the remaining routes (all with the same local preference + AS-PATH length) the router with the closest NEXT-HOP value is selected
    - Router for which the cost of the least-cost path (determined by the intra-AS algorithm) is the smallest
  - If more than one route still remains, the router uses BGP identifier to select the route

Broadcast Routing

- Broadcast routing service – delivery of a packet from a single source node to all other nodes in the network
- Naïve solution: *N-way-unicast approach*
  - Given: *N* destination nodes
  - Source node makes *N* copies of the packet
  - Addresses each copy to different destination
  - Routes using unicast routing
- Drawbacks:
  - Inefficiency: Multiple copies of the same packet occupy the link
  - Additional overhead – obtain addresses of recipients
  - Achieve broadcast by relying on unicast infrastructure

Duplicate creation/transmission

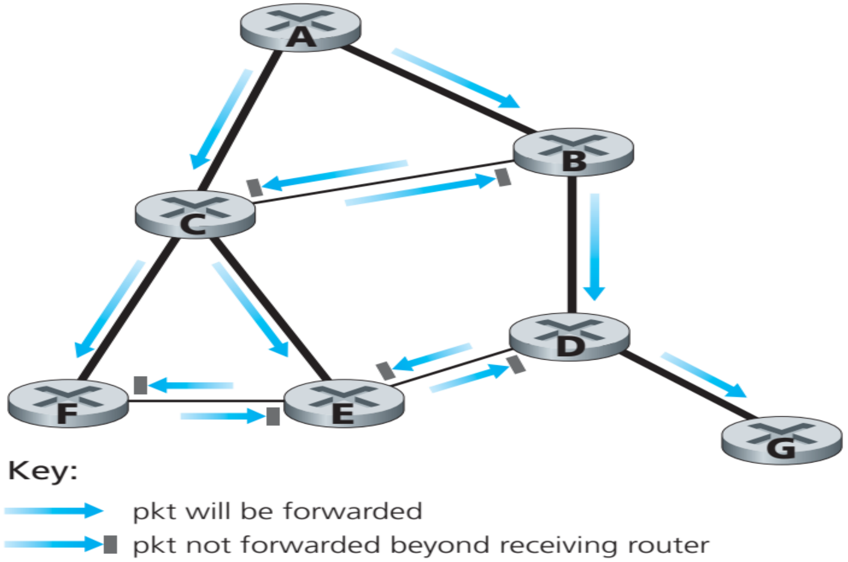


Source Duplication Vs. In-Network Duplication

- Uncontrolled Flooding
  - One of the flooding approaches
  - Source node sends a copy of the packet to all its neighbors
  - Each neighbor duplicates the packet and forwards it to all of its neighbors.
  - Drawback:
    - If a graph cycle is present: one of more copies of each broadcast packet will cycle indefinitely
    - Broadcast storm: endless multiplication of broadcast packets

Broadcast Routing: Controlled Flooding

- Sequence-number-controlled Flooding
  - Source node puts its address as well as a broadcast sequence number into the packet
  - Forwards the packet to all of its neighbors
  - Each node maintains a list of the source address and the sequence of the broadcast packet it has received, duplicated, and forwarded
  - On receiving a broadcast packet, the node checks its list
  - If the packet is in its list, it is dropped
  - Else, it is duplicated and forwarded to all of its neighbors (except the one from which the packet has arrived)
- Gnutella (P2P application) uses a time-to-live (TTL) field to limit the number of hops over which a flooded query will be forwarded
  - TTL field is decremented at each node – limited scope flooding



Reverse Path Forwarding

- Reverse Path Forwarding (RPF)
  - Router checks if the received broadcast packet has arrived on the link that is on its own least-cost (shortest) unicast path back to the source
  - If so, the packet is transmitted on all of its outgoing link – else it is simply discarded
  - Underlying assumption: router will either receive or has already received copies of packets on the link that is on its least-cost (shortest) path back to the source
  - RPF does not use unicast routing - does not require the router to know the complete least-cost (shortest) path back to the sender
  - RPF needs to know only the next neighbor on its unicast shortest path to the sender