Network Layer


Network Layer

• Network Layer
  • Routing (Control Plane)
  • Forwarding (Data Plane)
  • Router Architecture Overview

• Data Plane
  • Forwarding
  • Internet Protocol (IP)
  • Generalized Forwarding and SDN

✓ Control Plane
  ✓ Routing (Per-Router Control)
    • Algorithms
  ✓ Protocols
    • Policies
  • Software Defined Networking (Logically Centralized Control)
Routing & Scalability

Our routing study thus far: **idealized**
- All routers identical
- Flat Network
- **Not** true in practice

**Scale:** with billions of destinations
- Can not store all destinations in routing tables
- Routing table exchange would swamp links

**Administrative autonomy**
- Internet: network of networks
- Each network admin may want to control routing in its own network
Internet Approach to Scalable Routing

• Aggregate routers into regions known as autonomous systems (AS) or domains

• Intra-AS routing
  • Routing among hosts and routers in the same AS (network)
  • All routers in AS must run same intra-domain protocol
  • Routers in different AS can run different intra-domain routing protocol
  • Gateway router: At edge of its own AS, has link(s) to router(s) in other ASes

• Inter-AS routing
  • Routing among ASes
  • Gateways perform inter-domain routing (as well as intra-domain routing)
Interconnected ASes

- Forwarding table configured by both intra-AS and inter-AS routing algorithm
  - Intra-AS routing determine entries for destinations within AS
  - Inter-AS & intra-AS determine entries for external destinations
Inter-AS Tasks

• Suppose router in AS1 receives datagram destined outside of AS1
  • Router should forward packet to gateway router
    • Which gateway router?

• AS1 must
  • Learn which destinations are reachable through AS2 and which destinations through AS3
  • Propagate reachability info to all routers in AS1
  • Job of inter-AS routing
Intra-AS Routing

• Also known as **Interior Gateway Protocols (IGP)**

• Most common intra-AS routing protocols
  • **RIP**: Routing Information Protocol
  
  • **OSPF**: Open Shortest Path First
    (IS-IS protocol similar to OSPF)
  
  • **IGRP**: Interior Gateway Routing Protocol
    (Cisco proprietary for decades, until 2016)
OSPF (Open Shortest Path First)

• Open: publicly available

• Uses link-state algorithm
  • Link state packet dissemination
  • Topology map at each node
  • Route computation using Dijkstra

• Router floods OSPF link-state advertisements to all other routers in entire AS
  • Carried in OSPF messages directly over IP (rather than TCP or UDP)
  • Link state: for each attached link

• IS-IS routing protocol: Similar to OSPF
OSPF: Advanced Features

• **Security**: all OSPF messages authenticated (to prevent malicious intrusion)

• **Multiple** same-cost **paths** allowed (only one path in RIP)
• For each link multiple cost metrics for different **TOS**
  • Example: Satellite link cost set low for best effort ToS & high for real-time ToS

• Integrated unicast and **multicast** support
  • Multicast OSPF (MOSPF) uses same topology database as OSPF

• **Hierarchical** OSPF in large domains.
Hierarchical OSPF

- **Two-level hierarchy**: local area, backbone.
  - Link-state advertisements only in area
  - Each node has detailed area topology only know direction (shortest path) to nets in other areas.
- **Area border routers**: Summarize distances to nets in own area, advertise to other Area Border routers.
- **Backbone routers**: Run OSPF routing limited to backbone.
- **Boundary routers**: Connect to other ASes
Network Layer

• Network Layer
  • Routing (Control Plane)
  • Forwarding (Data Plane)
  • Router Architecture Overview

• Data Plane
  • Forwarding
  • Internet Protocol (IP)
  • Generalized Forwarding and SDN

✓ Control Plane
  ✓ Routing (Per-Router Control)
    • Algorithms
    • Protocols
    ✓ Policies
  • Software Defined Networking (Logically Centralized Control)
Internet Inter-AS Routing: BGP

- **BGP (Border Gateway Protocol):** The de facto inter-domain routing protocol
  - Glue that holds the Internet together

- BGP provides each AS a means to
  - **eBGP:** Obtain subnet reachability information from neighboring ASes
  - **iBGP:** Propagate reachability information to all AS-internal routers.
  - Determine **good** routes to other networks based on reachability information and **policy**

- Allows subnet to advertise its existence to rest of Internet: **I am here!**
eBGP & iBGP Connections

gateway routers run both eBGP and iBGP protocols
BGP Basics

- **BGP session**: Two BGP routers (peers) exchange BGP messages over semi-permanent TCP connection:
  - Advertising **paths** to different destination network prefixes (BGP is a **path vector** protocol)

- When AS3 gateway router 3a advertises path AS3, X to AS2 gateway router 2c, AS3 **promises** to AS2 it will forward datagrams towards X
Path Attributes & BGP Routes

• Advertised prefix includes BGP attributes
  • Prefix + attributes = route

• Two important attributes
  • **AS-PATH**: list of ASes through which prefix advertisement has passed
  • **NEXT-HOP**: indicates specific internal-AS router to next-hop AS

• **Policy-based routing**
  • Gateway receiving route advertisement uses import policy to accept/decline path (e.g. never route through AS Y)
  • AS policy also determines whether to advertise path to other neighboring ASes
BGP Path Advertisement

• AS2 router 2c receives path advertisement AS3, X (via eBGP) from AS3 router 3a

• Based on AS2 policy, AS2 router 2c accepts path AS3, X, propagates (via iBGP) to all AS2 routers

• Based on AS2 policy, AS2 router 2a advertises (via eBGP) path AS2, AS3, X to AS1 router 1c
BGP Path Advertisement

Gateway router may learn about multiple paths to destination:

- AS1 gateway router 1c learns path **AS2, AS3, X** from 2a
- AS1 gateway router 1c learns path **AS3, X** from 3a
- Based on policy, AS1 gateway router 1c chooses path **AS3, X, and advertises path within AS1 via iBGP**
BGP Messages

- BGP messages exchanged between peers over TCP connection

- BGP messages
  - **Open**: opens TCP connection to remote BGP peer and authenticates sending BGP peer
  - **Update**: advertises new path (or withdraws old)
  - **Keepalive**: keeps connection alive in absence of **Updates**; also ACKs **Open** request
  - **Notification**: reports errors in previous message; also used to close connection
Q: How does router set forwarding table entry to distant prefix?

Recall: 1a, 1b, 1c learn about destination X via iBGP from 1c: path to X goes through 1c
1d: OSPF intra-domain routing: to get to 1c, forward over outgoing local interface 1
Q: How does router set forwarding table entry to distant prefix

- Recall: 1a, 1b, 1c learn about destination X via iBGP from 1c: Path to X goes through 1c
- 1d: OSPF intra-domain routing: to get to 1c, forward over outgoing local interface 1
- 1a: OSPF intra-domain routing: to get to 1c, forward over outgoing local interface 2
BGP Route Selection

- Router may learn about more than one route to destination AS, selects route based on
  - Local preference value attribute: policy decision
  - Shortest AS-PATH
  - Closest NEXT-HOP router: hot potato routing
  - Additional criteria
Hot Potato Routing

- 2d learns (via iBGP) it can route to X via 2a or 2c
- **Hot potato routing:** Choose local gateway that has least intra-domain cost (e.g., 2d chooses 2a, even though more AS hops to X): don’t worry about inter-domain cost!
BGP: Achieving Policy Via Advertisements

• Suppose an ISP only wants to route traffic to/from its customer networks (Does not want to carry transit traffic between other ISPs)
  • A advertises path Aw to B and to C
  • B chooses not to advertise BAw to C:
    • B gets no revenue for routing CBAw, since none of C,A, w are B’s customers
    • C does not learn about CBAw path
  • C will route CAw (not using B) to get to w
BGP: Achieving Policy Via Advertisements

• Suppose an ISP only wants to route traffic to/from its customer networks (does not want to carry transit traffic between other ISPs)
  • A,B,C are provider networks
  • X,W,Y are customers (of provider networks)
  • X is dual-homed: attached to two networks
  • policy to enforce: X does not want to route from B to C via X
    • .. so X will not advertise to B a route to C
Why Different Intra-, Inter-As Routing?

Policy
- Inter-AS: admin wants control over how its traffic routed, who routes through its net
- Intra-AS: single admin, so no policy decisions needed

Scale
- Hierarchical routing saves table size, reduced update traffic

Performance
- Intra-AS: Can focus on performance
- Inter-AS: Policy may dominate over performance
Acknowledgements

• The following materials have been used in preparation of this slide set:

   7th Edition
   James Kurose, Keith Ross
   Pearson
   2016

   5th Edition
   Larry Peterson, Bruce Davie
   Morgan Kaufmann
   2011