MIT 6.829: Computer Networks	Fall 2017

Lecture 15: Routing and BGP

Scribe: William Moses and Aaron Zalewski

One of most important problems in networking: routing, how you get your packets from one side to the other.

Two topics: forwarding and routing.

Routing has two types (intradomain and interdomain). Not going to discuss much about intradomain routing (assuming prior knowledge). Today going to be discussing interdomain routing and the subproblems of policies and BGP.

The internet is really a network of networks. It is organized in some hierarchical structure. Each enterprise has its own internal network. There is still a structure of the ISPs so that MIT can send to Berkeley for example.

The big ones of these are called Autonomous Systems (AS). Each one is used to route at an interdomain level. Inside each one there is autonomy of how it does its routing. Nobody will impose what to run internally.

To be able to exchange traffic, it has to connect itself to the rest of the internet. Usually where MIT and customer networks go to an ISP, which have their own networks, which connect to bigger ISPs (say global ISPs like AT&T).

There is a routing protocol (intradomain routing) that connect these autonomous systems togehter.

1 Intradomain routing

Time for a review.

1.1 Forwarding vs routing

There are routers/switches. These receive packets and must be able to send them to their destinations. These routers connect to other routers and eventually connect to their destination. Must be able to forward packets (by looking at IP in destination field), process through its internal lookup table, and determine which link to send the packet.

This forwarding process must happen **very** quickly.

How does the router generate these forwarding tables?

Lookup tables are generated via the routing protocol.

Forwarding: act of forwarding data packet.

Routing: process of generating forwarding tables that allows this to happen.

At the router, where the packet came from is irrelevant. Forwarding at the router is only determined by the destination of the packet.

Not going to talk about most of the intradomain routing protocols available.

Objective: Allow each node in the network to find the shortest path to all other reachable nodes.

- Network can be viewed as a graph
- Links have costs
- "Shortest path" means the path with minimum link cost

Nodes form a **shortest path tree** for each destination.

Requirements for routing protocol:

- Correctness (route must lead to correct destination)
- Completeness - protocol should find a route if the destination is reachable
- Convergence (if network graph does not change, the routes must eventually converge)
- Loop-freedom (after convergence, the routes have no loops) If a packet is looped once, the router won't be able to tell if its already been seen (simply adding congestion).

Overhead not as big of a problem here (bigger when discussing interdomain routing).

1.2 Link State Protocol

Each node tells everyone inside the network about its neighbors. This forms a bigger network. Flooding protocol (sending from everyone except those who got the packet).

- Based on Dijkstra's shortest path algorithm
- Each node tells everyone about its neighbors and link costs
- Each node obtains the network graph
- Each node locally computes paths form itself to everyone

1.3 Distance Vector

- Based off of Bellman-Ford shortest path alg.
- A node tells only its neighbors its shortest path cost to every node in the network
- Each node updates its shortest path of its neighbors
- No one has the full graph

2 Interdomain routing

All intradomain requirements still hold, but now we have a few more requirements

2.1 Scalability

- Small routing tables: Can't have an entry per machine, causes large lookup delay
- Small message overhead and fast convergence: Link going up or down shouldn't cause routing messages to spread to the whole internet

Need less information with increasing distance to destination. Typically implemented with hierarchy. Both routing and addressing on the internet is hierarchical.

Across ASs, we run a different routing protocol (called BGP).

2.2 Hierarchical Addressing

Analogy: Mail - an address has varying layers of hierarchy (ie. country, state, street address). Exception, IPs use prefix hierarchy as opposed to mail talking about suffix.

- Each IP address has 4 bytes (i.e. 18.0.1.2)
- The IP address space is divided into line segments (contiguous chunk of addresses)
- Each segment is described by a prefix.
- A prefix is of the form x/y where x is the prefix of all addresses in the and y is the length of the segment in bits
- For example 128.9/16 is all the segments in the range 128.9.0.0 to 128.9.255.255
- When routing, will send packet to entry in its forwarding table w/ longest prefix match

2.3 Policy

Shortest path is not only metric; ISPs want to maximize revenues.

Transient links are in consumer-provider relationship. Peering link, two ISPs that are equally big that themselves try to avoid paying own provider by peering locally.

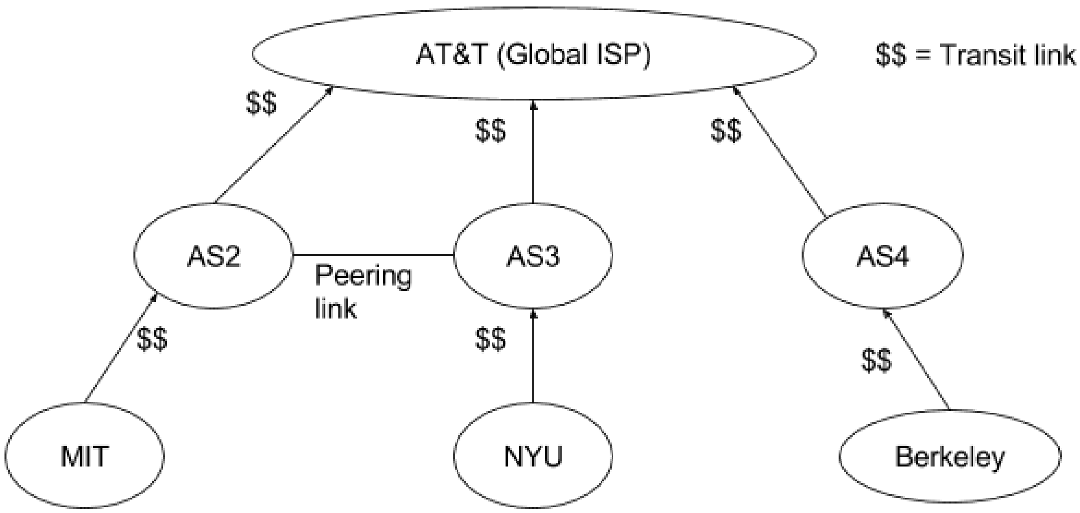
Transit: A larger AS will take and route traffic for a smaller AS, in return for money

Peering: Two ASes send their traffic to each other, avoid having to pay their transit provider

An ISP does not take a packet unless it makes (or saves) it money.

An ISP will prefer to route traffic through its customer (even if it's not the shortest path), since then it will get paid

Figure 1: Transit vs Peering



3 BGP

No security, because verifying signatures would take too long (and potentially lead to DDOS type attacks).

"So, the success of this system is the effectiveness to which we can enforce antitrust laws on ISPs"

1. Advertise whole path (needed to prevent loops, since the AS is not necessarily choosing the shortest path. An AS will reject the route if the path has its own AS number)

2. Incremental updates. Send updates only when best/current route changes. Has two types of update messages: announcement ("`P: {AS-20, AS-6}`") and withdrawals ("withdraw P").

Forwarding table will contain the actual best link, routing table will contain all paths. If a link goes down, the AS uses the next best route.

3.1 Enforcing Policies (and making money) using BGP

Route Exports: Controls incoming traffic

- Advertise customers to everyone and everyone to customers
- Advertise peer to customers but not customers to peer

Route Imports

- Prefer Customer > Peer > Provider

BGP

1. Update from neighbor AS
2. Remove loops; rank imported routes and update routing table
3. Pick best route for forwarding
4. export best route if desirable
5. send update to neighbors