Computer Network Design

Network Layer IV

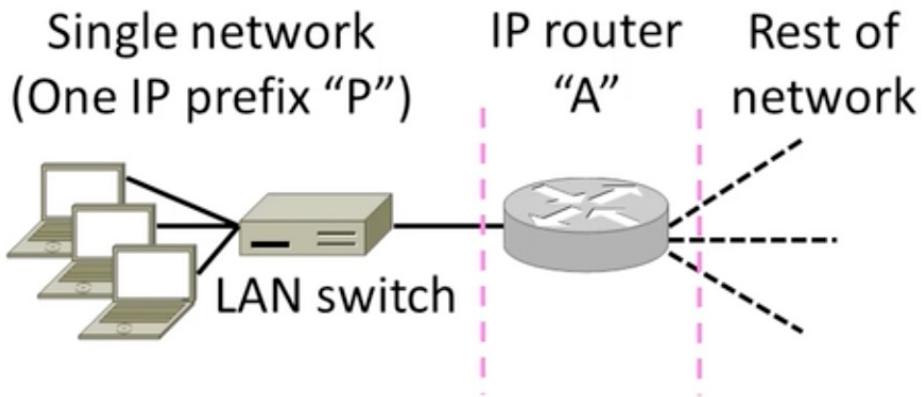
Yalda Edalat – Spring 23

Recap

- In the Internet:
  - Hosts on same network have IP addresses in the same IP prefix
  - Hosts just send off-network traffic to the nearest router to handle
  - Routers discover the routes to use
  - Routers use longest prefix matching to send packets to the right next hop

Host/Router Combination

- Hosts attach to routers as IP prefixes
  - Router need table to reach all hosts

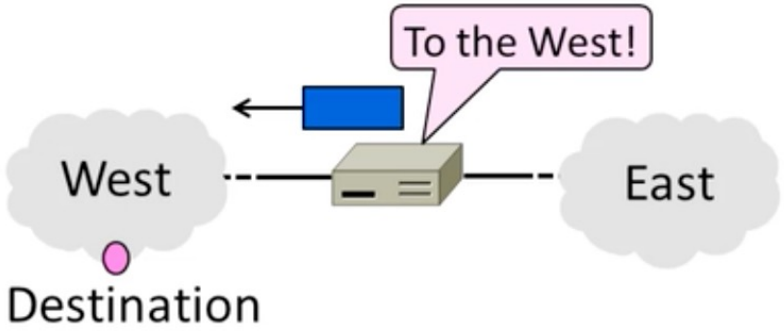


Network Topology for Routing (2)

- Routing now works as before
  - Routers advertise IP prefixes for hosts
  - Router addresses are “/32” prefixes
  - Lets all routers find a path to hosts
  - Hosts find by sending to their router

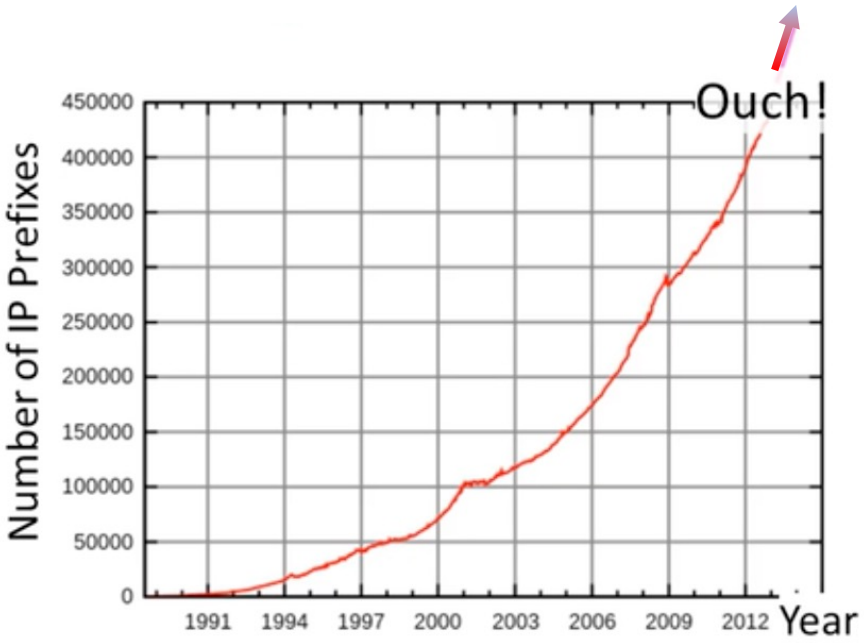
How to Scale Routing?

- How to scale routing with hierarchy in the form of regions
  - Route to regions, not individual nodes



Internet/Internet Routing Growth

- At least a billion Internet hosts and growing …
- Internet growth translates into routing table growth (even using prefixes)



Source: By Mro (Own work), CC-BY-SA-3.0 , via Wikimedia Commons

Impact of Routing Growth

1. Forwarding table grow
   - Larger router memories, may increase lookup time
2. Routing messages grow
   - Need to keeps all nodes informed of larger topology
3. Routing computation grows
   - Shortest path calculations grow faster than size of the network

Techniques to Scale Routing

1. IP prefixes
   - Route to blocks of hosts

2. Network hierarchy
   - Route to network regions

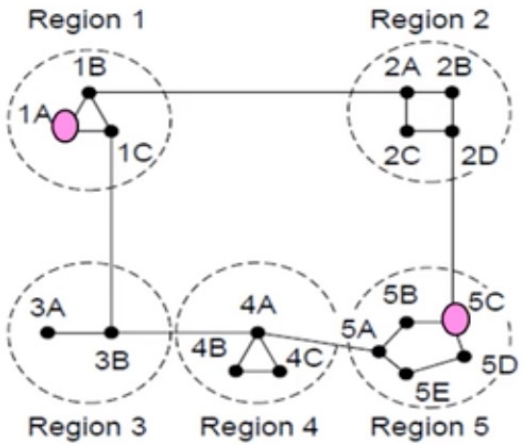
3. IP prefix aggregation
   - Combine, and split, prefixes

Hierarchical Routing

- Introduce a larger routing unit
  - IP prefix (hosts)
  - Region, e.g., ISP network

- Route first to the region, then to the IP prefix within the region
  - Hide details within a region from outside of the region

Hierarchical Routing (2)



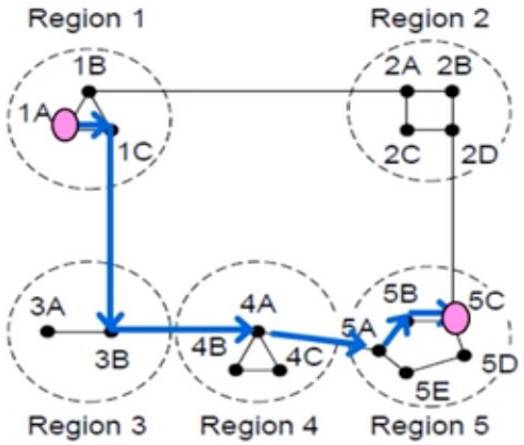
Full table for 1A

Dest.	Line	Hops
1A	–	–
1B	1B	1
1C	1C	1
2A	1B	2
2B	1B	3
2C	1B	3
2D	1B	4
3A	1C	3
3B	1C	2
4A	1C	3
4B	1C	4
4C	1C	4
5A	1C	4
5B	1C	5
5C	1B	5
5D	1C	6
5E	1C	5

Hierarchical table for 1A

Dest.	Line	Hops
1A	–	–
1B	1B	1
1C	1C	1
2	1B	2
3	1C	2
4	1C	3
5	1C	4

Hierarchical Routing (3)



Full table for 1A

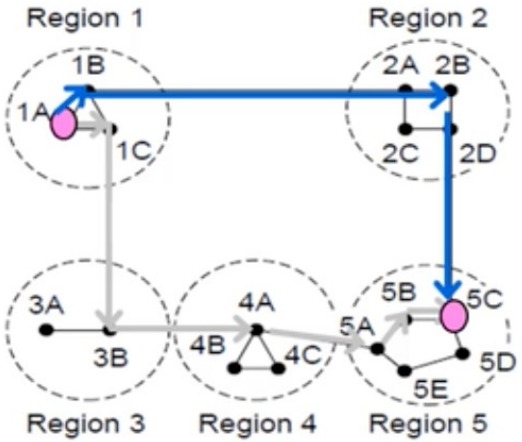
Dest.	Line	Hops
1A	–	–
1B	1B	1
1C	1C	1
2A	1B	2
2B	1B	3
2C	1B	3
2D	1B	4
3A	1C	3
3B	1C	2
4A	1C	3
4B	1C	4
4C	1C	4
5A	1C	4
5B	1C	5
5C	1B	5
5D	1C	6
5E	1C	5

Hierarchical table for 1A

Dest.	Line	Hops
1A	–	–
1B	1B	1
1C	1C	1
2	1B	2
3	1C	2
4	1C	3
5	1C	4

Hierarchical Routing (4)

Penalty is longer paths



Full table for 1A

Dest.	Line	Hops
1A	–	–
1B	1B	1
1C	1C	1
2A	1B	2
2B	1B	3
2C	1B	3
2D	1B	4
3A	1C	3
3B	1C	2
4A	1C	3
4B	1C	4
4C	1C	4
5A	1C	4
5B	1C	5
5C	1B	5
5D	1C	6
5E	1C	5

Hierarchical table for 1A

Dest.	Line	Hops
1A	–	–
1B	1B	1
1C	1C	1
2	1B	2
3	1C	2
4	1C	3
5	1C	4

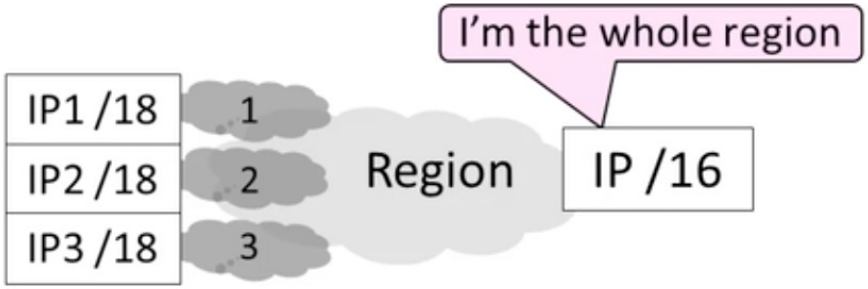
1C is best route to region 5, except for destination 5C

Observations

- Outside a region, nodes have one route to all hosts within the region
  - This gives saving in table size, messages and computation
- However each node may have a different route to an outside region
  - Routing decisions are still made by individual nodes; there is no single decision made by a region

IP Prefix Size

- How to help scale routing by adjusting the size of IP prefixes
  - Split (subnets) and join (aggregation)



Recall

- IP addresses are allocated in blocks called IP prefixes, e.g., 18.31.0.0/16
  - Hosts on one network in same prefix

- A “N” prefix has the first N bits fixed and contains 2^{32-N} addresses
  - E.g., “/24”
  - E.g., “/16”

Key Flexibility

- Routers keep track of prefix lengths
  - Use it for longest prefix matching

Routers can change prefix lengths without affecting hosts

- More specific IP prefix
  - Longer prefix, fewer IP addresses
  - Shorter prefix, more IP addresses

Prefixes and Hierarchy

- IP prefixes already help to scale routing, but we can go further
  - Can use a less specific prefix to name a region made up of several prefixes

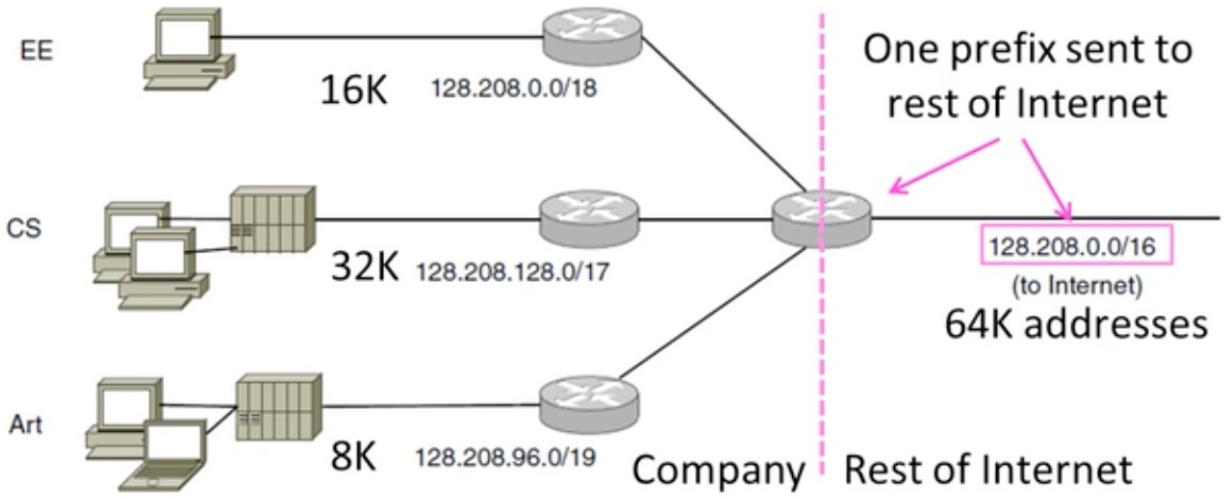
Subnets and Aggregation

- Two use cases for adjusting the size of IP prefixes; both reduce routing table

1. Subnets
   - Internally split one less specific prefix into multiple more specific prefixes
2. Aggregation
   - Externally join multiple more specific prefixes into one large prefix

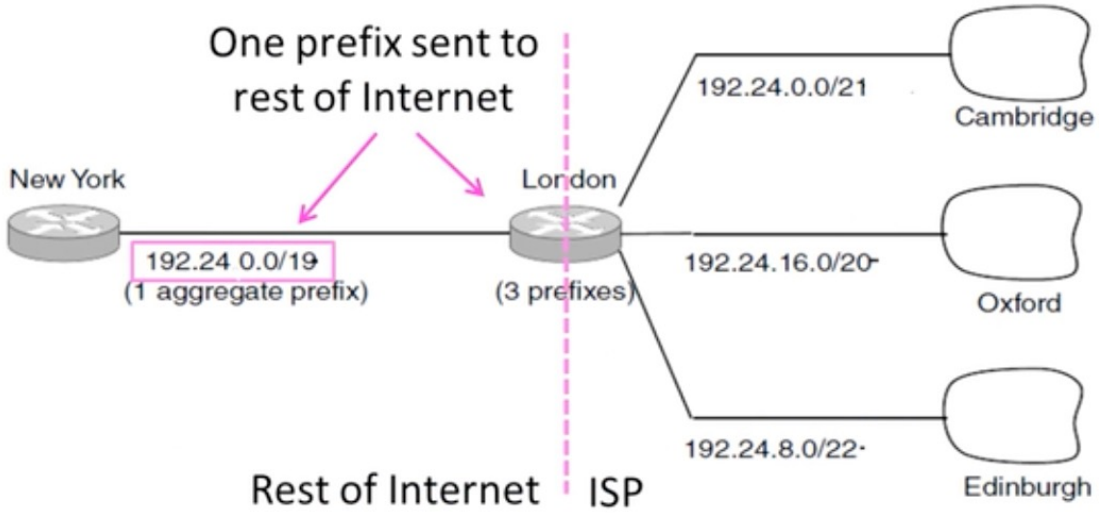
Subnets

- Internally split up one IP prefix



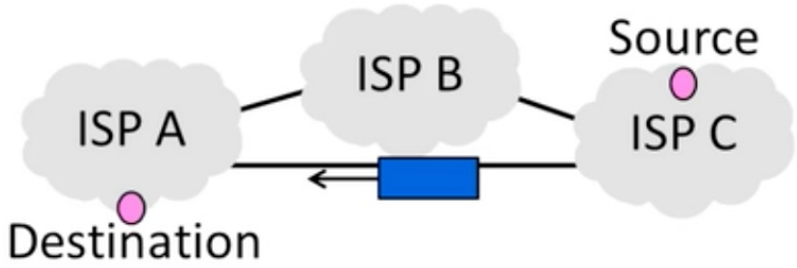
Aggregation

- Externally join multiple separate IP prefixes



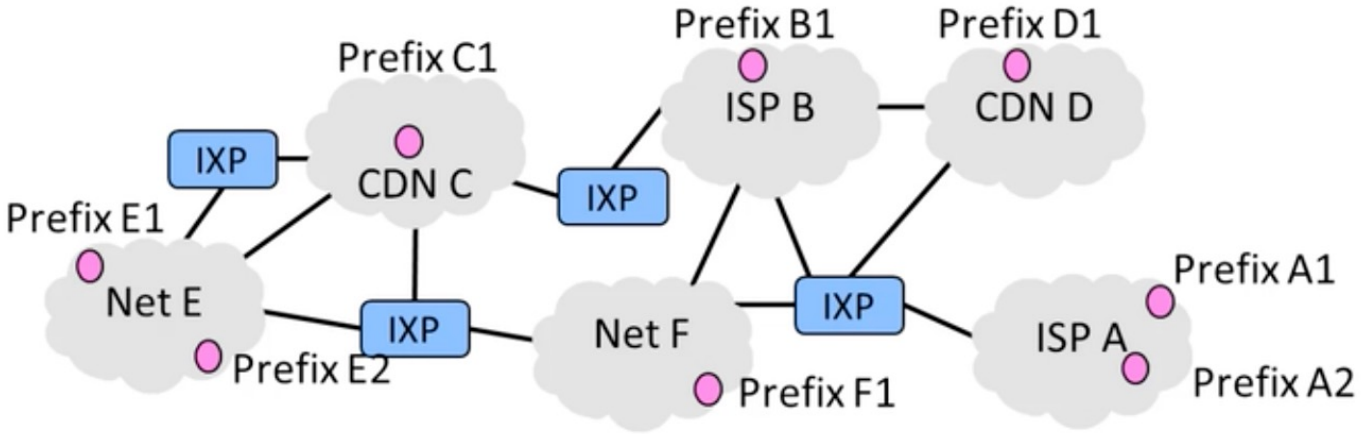
Exterior Gateway Protocol

- How to route with multiple parties, each with their own routing policies
  - BGP computes Internet-wide routes



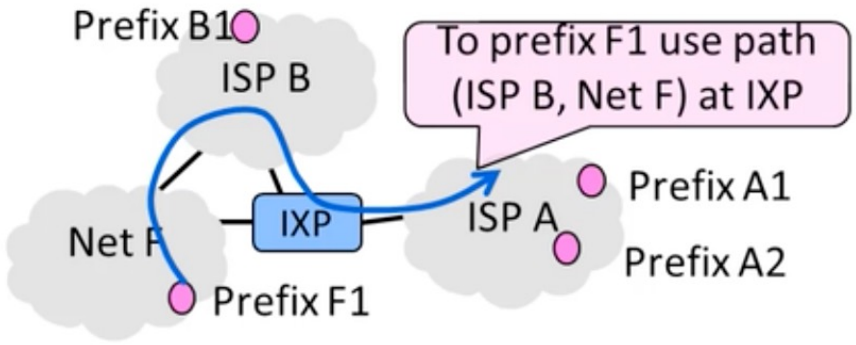
Recall

- Internet is made up of independently run networks
- Each network has its own route preferences (policies)



BGP (Border Gateway Protocol)

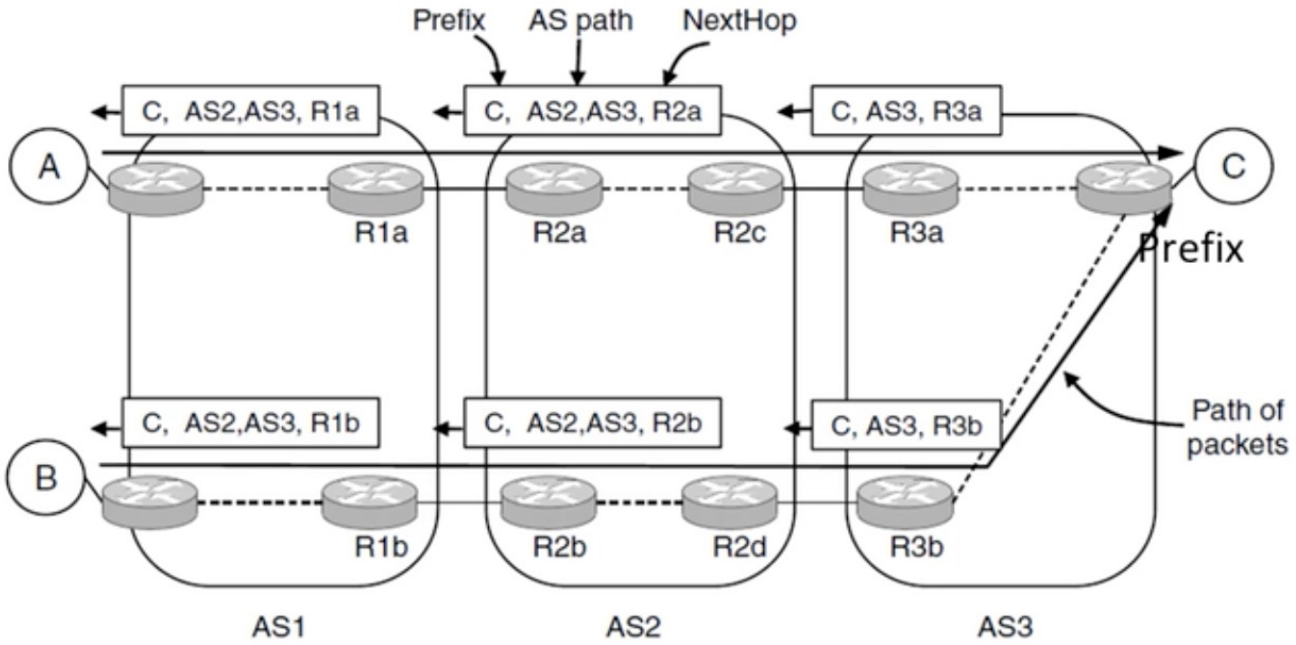
- BGP is the protocol that computes interdomain routes in the Internet
  - Path vector, a kind of distance vector



BGP (2)

- Different parties like ISPs are called AS (Autonomous Systems)
- Border routers of ASes announce BGP routes to each other
- Route announcements contain an IP prefix, path vector, next hop
  - Path vector is list of ASes on the way to the prefix; list is to find loops
- Route announcements move in the opposite direction to traffic

BGP (3)


Prefix
AS path
NextHop
C, AS2,AS3, R1a
C, AS2,AS3, R2a
C, AS3, R3a
A
C
R1a
R2a
R2c
R3a
Prefix
C, AS2,AS3, R1b
C, AS2,AS3, R2b
C, AS3, R3b
Path of packets
B
R1b
R2b
R2d
R3b
AS1
AS2
AS3

BGP(4)

- Policy is implementation in two ways:
  1. Border routers of ISP announce paths only to other parties who may use those paths
     - Filter out paths others can't use
  2. Border routers of ISP select the best path of the ones they hear in any, non-shortest way

To do

- No quiz next week
- Review session next week