Network: Global Internet Routing, Policy Routing Analysis

Qiao Xiang, Congming Gao

https://sngroup.org.cn/courses/cnns-xmuf23/index.shtml

12/12/2023

This deck of slides are heavily based on CPSC 433/533 at Yale University, by courtesy of Dr. Y. Richard Yang.

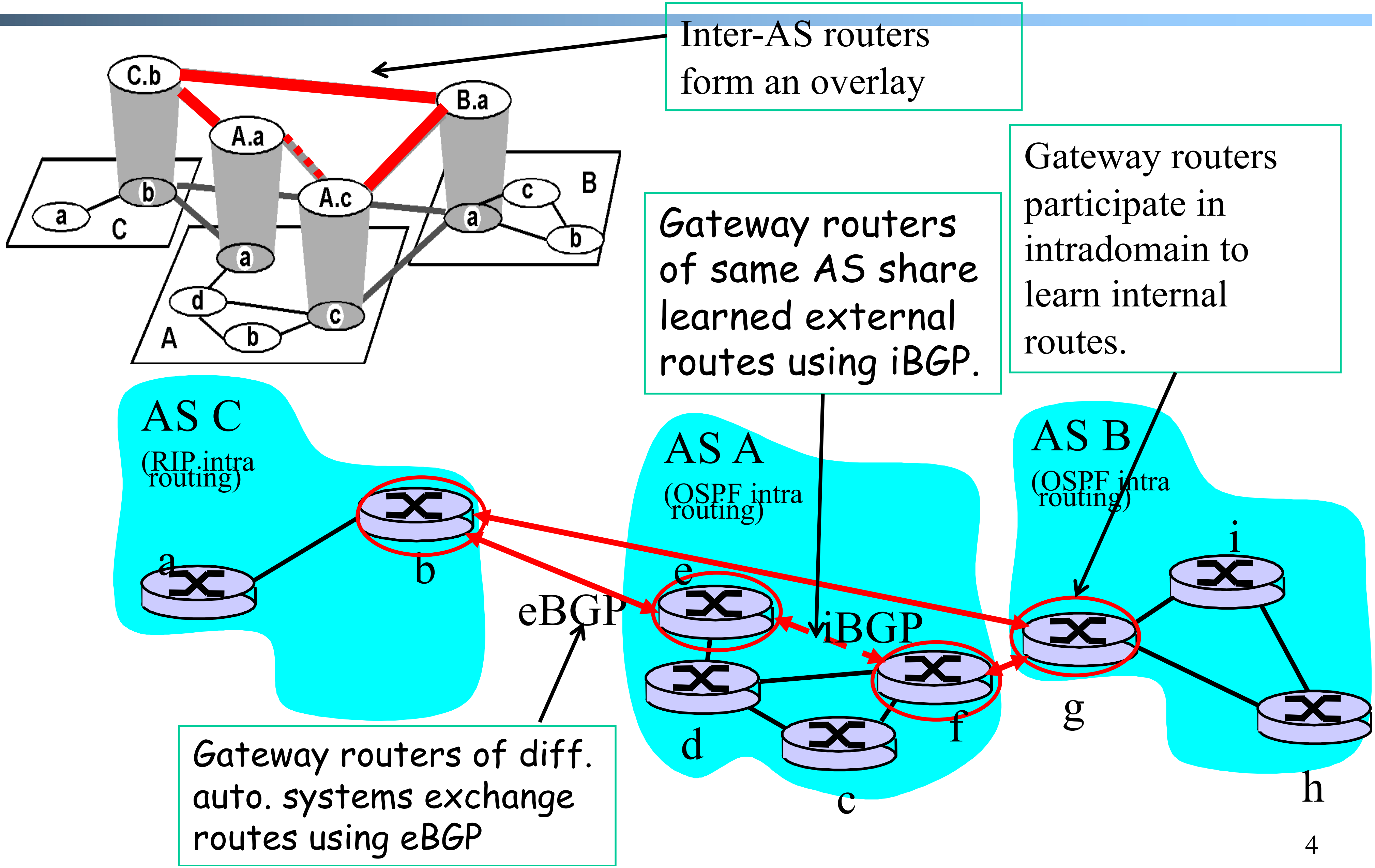
Outline

- Admin and recap
- Network control plane
  - Routing
    - Link weights assignment
    - Routing computation
      - Basic routing computation protocols
      - Global Internet routing
        - Basic architecture
        - BGP (Border Gateway Protocol): The de facto Inter-domain routing standard
          - Basic operations
          - BGP as a policy routing framework (control interdomain routes)
        - Policy/interdomain routing analysis
          - Global preference aggregation and Arrow's Theorem
          - Local preference aggregation

Recap: Internet Routing Architecture

- Interdomain routing uses a path vector protocol based on AS topology
  - improves scalability, privacy, autonomy
- Only a small # of routers (gateways) from each AS in the interdomain level
  - improves scalability
- Autonomous systems have flexibility to choose their own intradomain routing protocols
  - allows autonomy

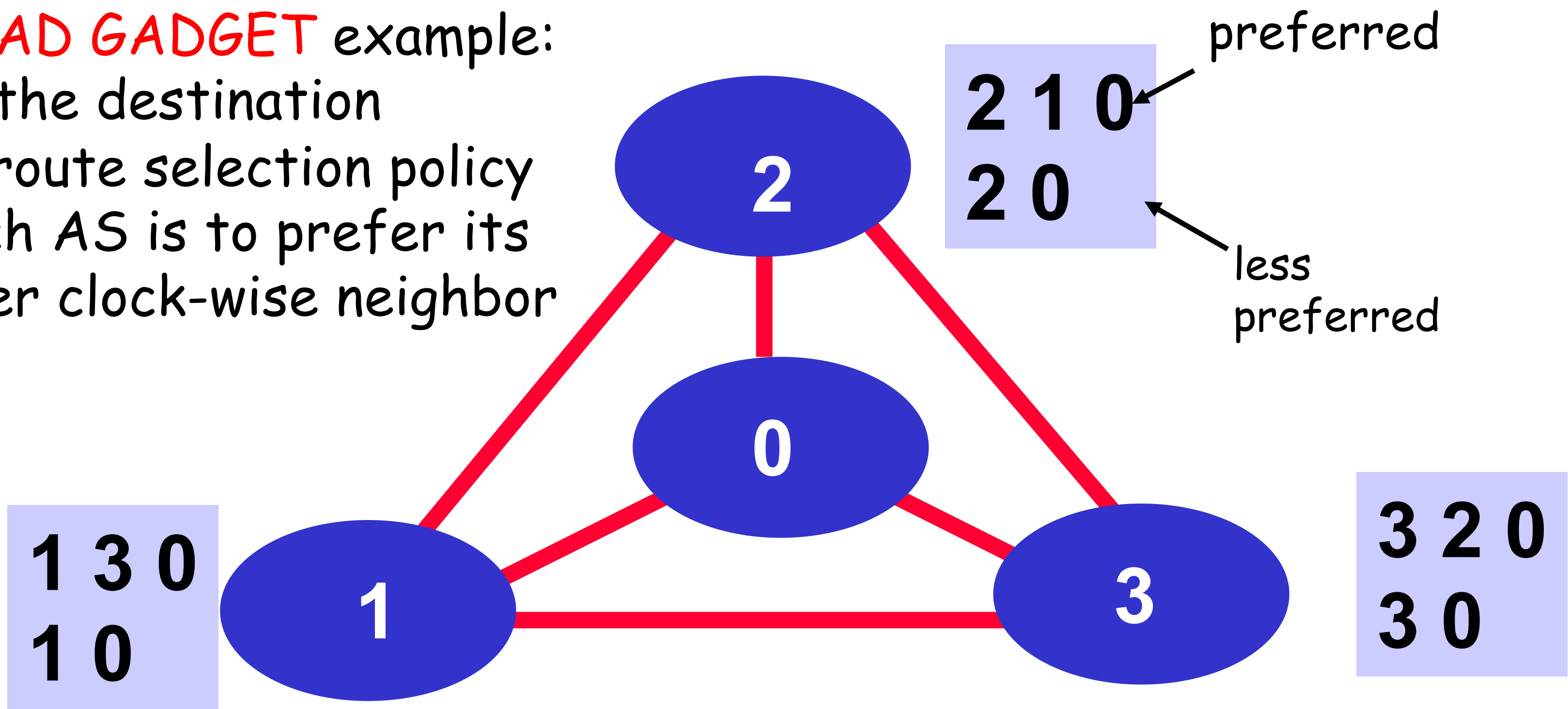
Recap: Routing with Autonomous Systems



Recap: Policy Routing as a Preference Aggregation System

- A policy routing system can be considered as a system to aggregate individual preferences, but aggregation may not be always successful.

The BAD GADGET example:
- 0 is the destination
- the route selection policy of each AS is to prefer its counter clock-wise neighbor



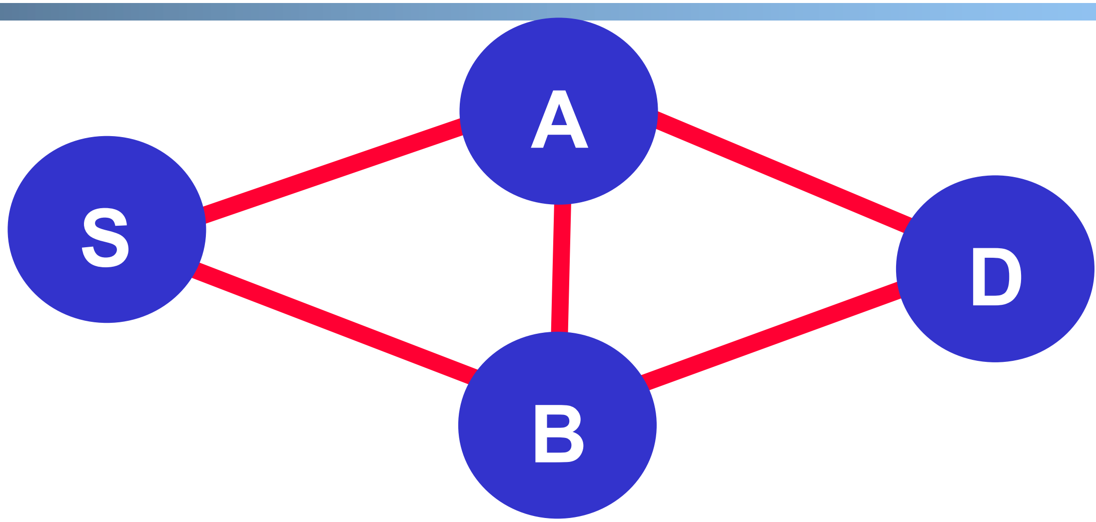
Policy (preferences) aggregation fails: routing instability !

General Framework of Preference Aggregation

- Also called Social Choice
  - Given individual preferences, define a framework to aggregate individual preferences:
    - A set of choices: a, b, c, …
    - A set of voters 1, 2, …
      - Each voter has a preference (ranking) of all choices, e.g.,
        - » voter 1: a > b > c
        - » voter 2: a > c > b
        - » voter 3: a > c > b
    - A well-specified aggregation rule (protocol) computes an aggregation of ranking, e.g.,
      - Society (network): a > b > c

For more details: see Semih Salihoglu (2007). Interdomain Routing as a Social Choice.

Example: Aggregation of Global Preference



- Choices (for S->D route): SAD, SBD, SABD, SBAD
- Voters S, A, B, D
- Each voter has a preference, e.g.,
  - S: SAD > SBD > SABD > SBAD
  - ...

Arrow's Impossibility Theorem

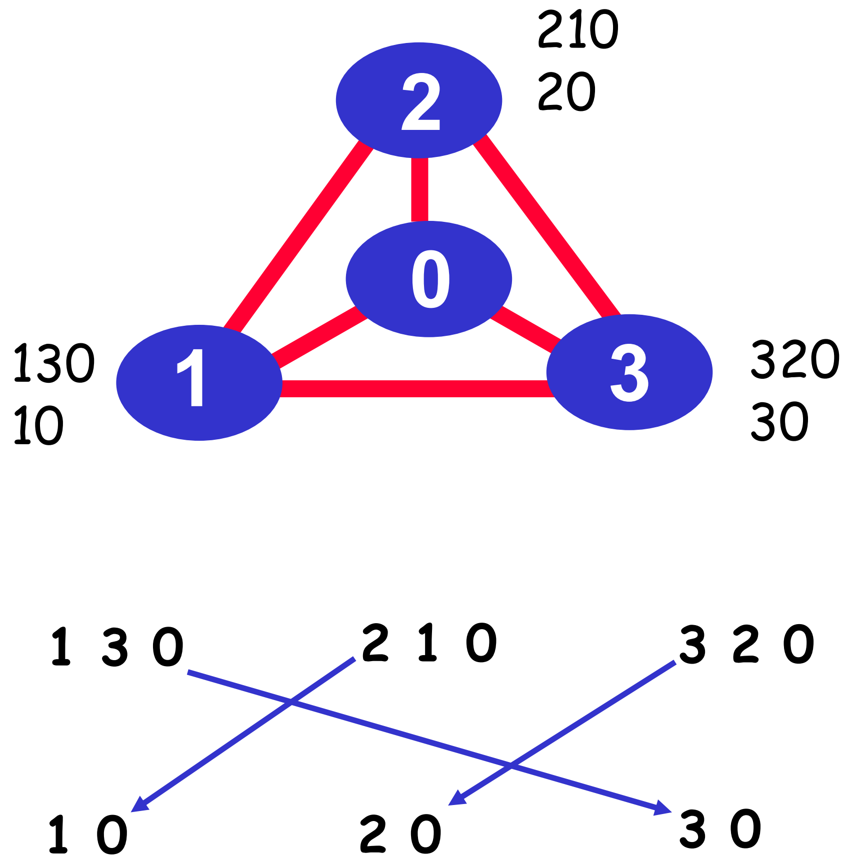
- Axioms:
  - Transitivity
    - if a > b & b > c, then a > c
  - Unanimity:
    - If all participants prefer a over b (a > b) => a > b
  - Independence of irrelevant alternatives (IIA)
    - Social ranking of a and b depends only on the relative ranking of a and b among all participants
- Result:
  - Arrow's Theorem: Any constitution that respects transitivity, unanimity and IIA is a dictatorship.

Outline

- Admin and recap
- Network control plane
  - Routing
    - Link weights assignment
    - Routing computation
      - Basic routing computation protocols
      - Global Internet routing
        - Basic architecture
        - BGP (Border Gateway Protocol): The de facto Inter-domain routing standard
          - Basic operations
          - BGP as a policy routing framework (control interdomain routes)
          - Policy/interdomain routing analysis
            - Global preference aggregation and Arrow's Theorem [Optional]
            - *Local preference aggregation*

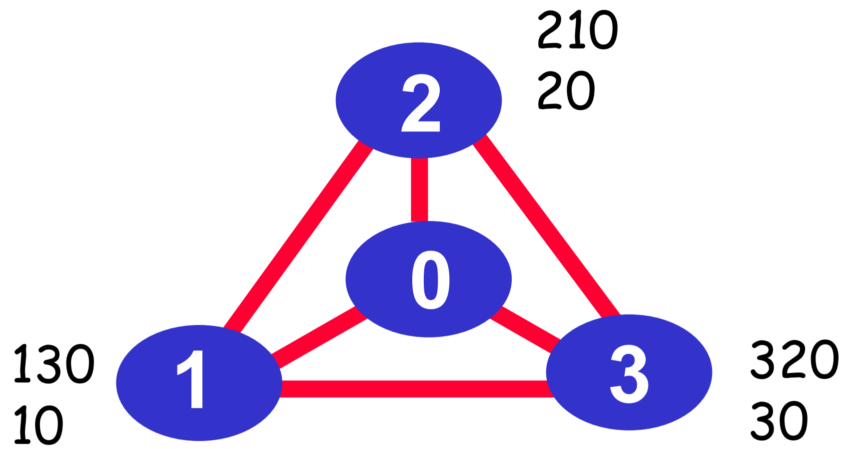
BGP w/ Local Preference

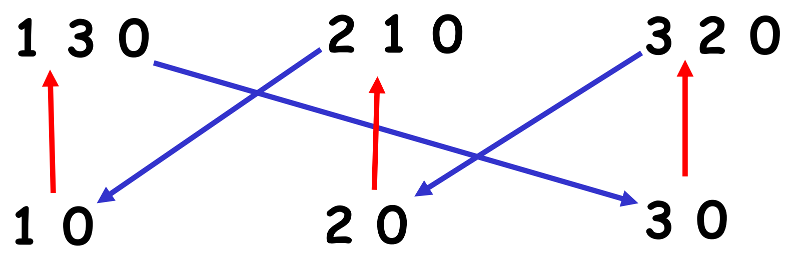
- BGP preferences are typically local (only on paths start from itself)
- Hence the preferences have dependency (priority)
  - The "closer" a node to the destination, the more "powerful" it may be



Complete Dependency: P-Graph

- Complete dependency can be captured by a structure called P-graph
- Nodes in P-graph are feasible paths
- Edges represent priority (low to high)
  - A directed edge from path N_1P_1 to P_1
    - intuition: to let N_1 choose N_1P_1, P_1 must be chosen and exported to N_1
  - A directed edge from a lower ranked path to a higher ranked path
    - intuition: the higher ranked path should be considered first

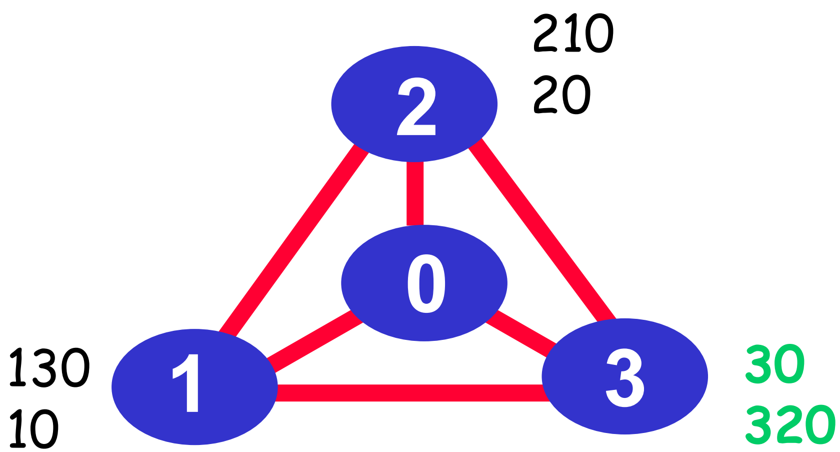


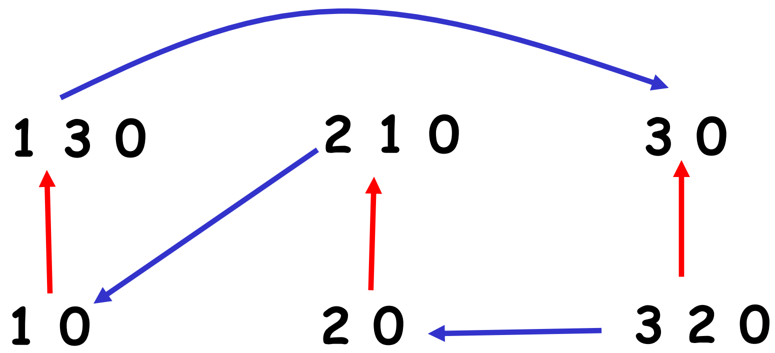


Any observation on the P-graph?

P-Graph and BGP Convergence

- If the P-graph of the networks has no loop, then policy routing converges.
  - intuition: choose the path node from the partial order graph with no out-going edge to non-fixed path nodes, fix the path node, eliminate all no longer feasible; continue
- Example: suppose we swap the order of 30 and 320



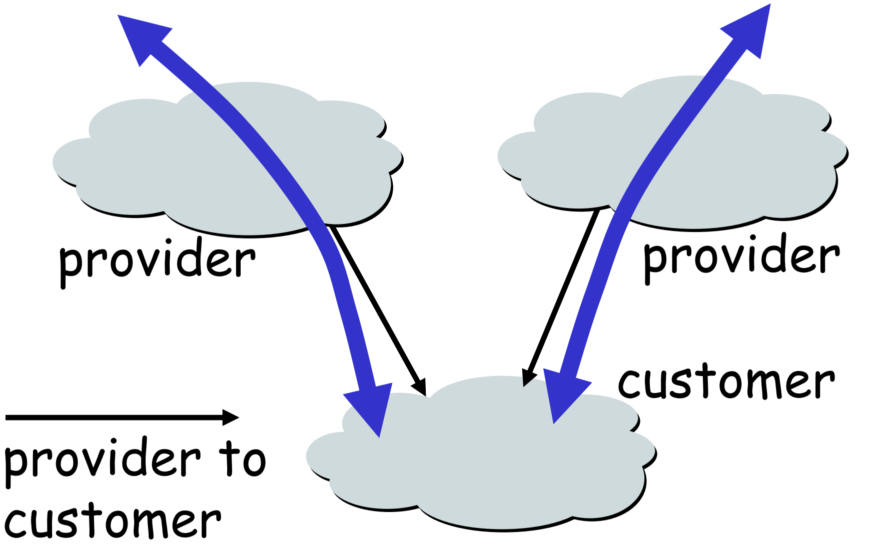


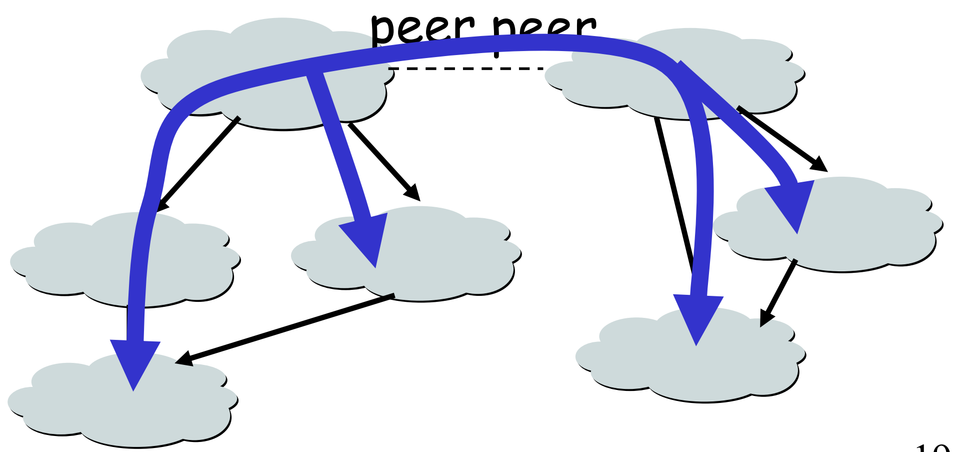
Outline

- Admin and recap
- Network control plane
  - Routing
    - Link weights assignment
    - Routing computation
      - Basic routing computation protocols
      - Global Internet routing
        - Basic architecture
        - BGP (Border Gateway Protocol): The de facto Inter-domain routing standard
          - Basic operations
          - BGP as a policy routing framework (control interdomain routes)
          - Policy/interdomain routing analysis
            - Global preference aggregation and Arrow's Theorem
            - Local preference aggregation
            - ***Economics and interdomain routing patterns***

Internet Economy: Two Types of Business Relationship

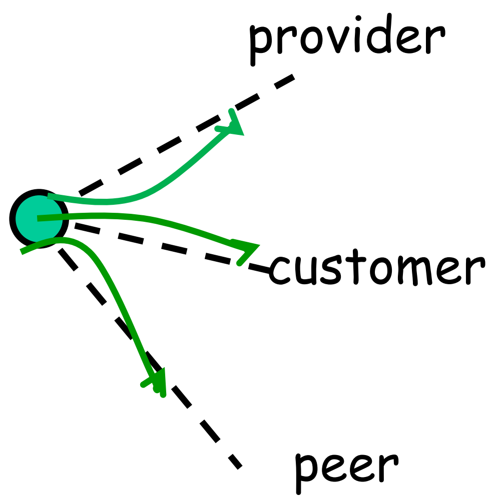
- *Customer provider relationship*
  - a provider is an AS that connects the customer to the rest of the Internet
  - customer pays the provider for the transit service
  - e.g., XMU is a customer of CERNET and China Telecom
- *Peer-to-peer relationship*
  - mutually agree to exchange traffic between their respective **customers** only
  - there is no payment between peers





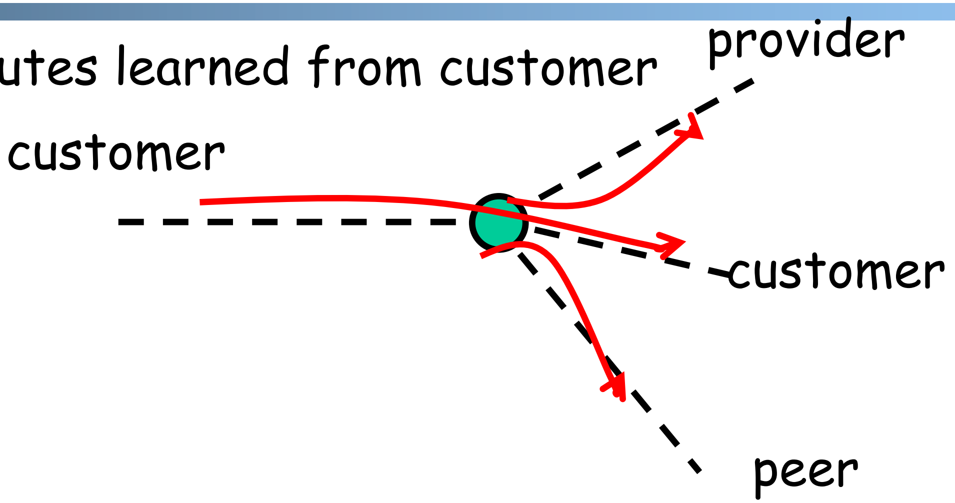
Route Selection Policies and Economics

- Route selection (ranking) policy:
  - the typical route selection policy is to prefer customers over peers/providers to reach a destination, i.e., Customer > pEer/Provider



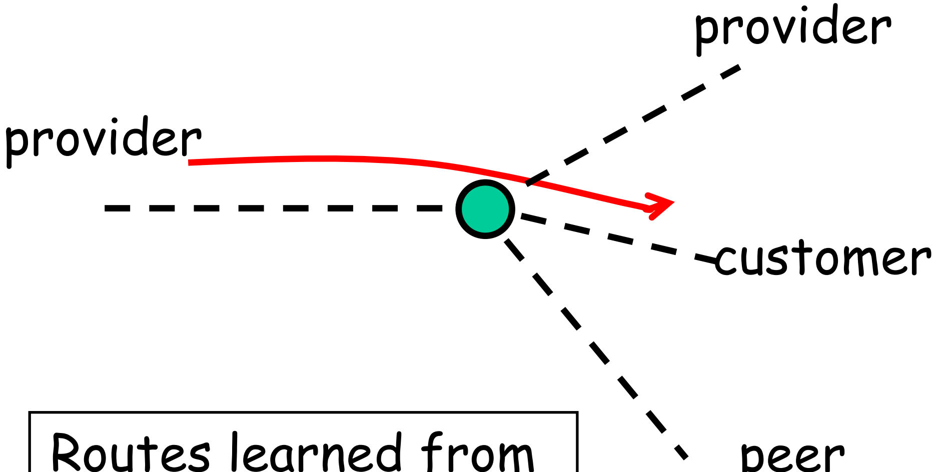
Export Policies and Economics

case 1: routes learned from customer



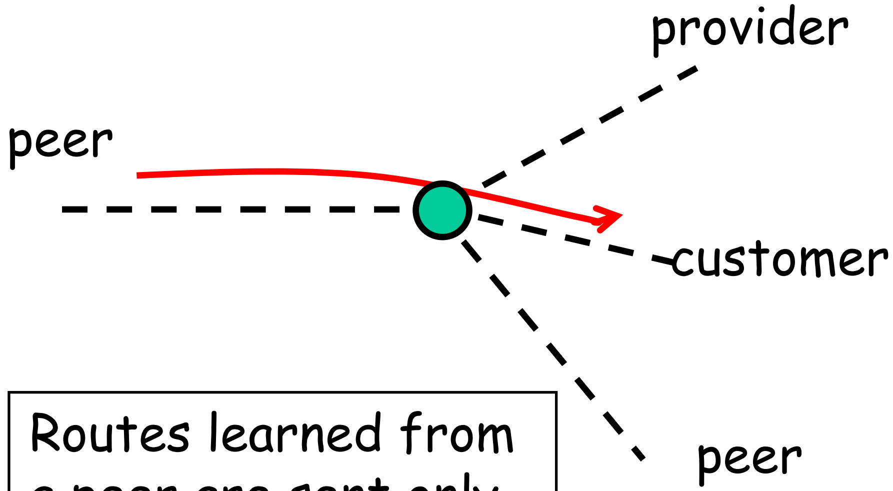
Routes learned from a customer are sent to all other neighbors

case 2: routes learned from provider



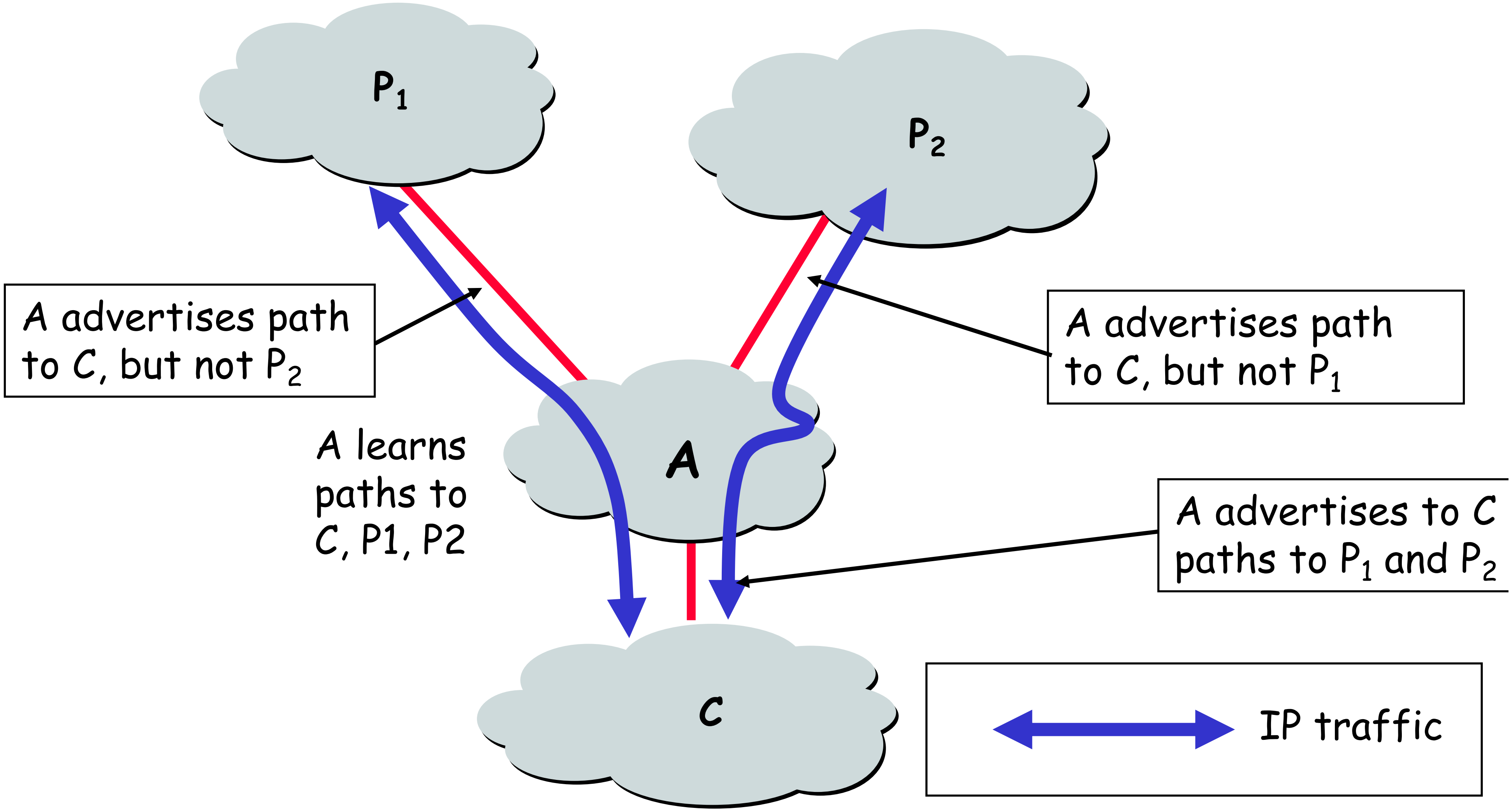
Routes learned from a provider are sent only to customers

case 3: routes learned from peer



Routes learned from a peer are sent only to customers

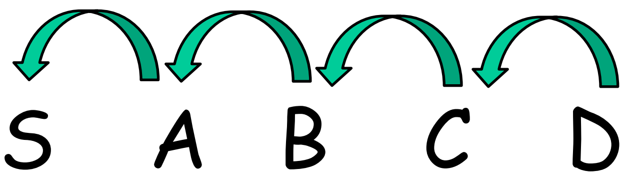
Example: Typical Export -> No-Valley Routing



Suppose P_1 and P_2 are providers of A; A is a provider of C

Typical Export Policies Route Patterns

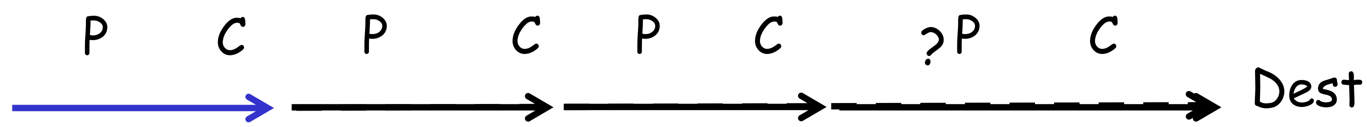
- Assume a BGP path SABCD to destination AS D. Consider the business relationship between each pair:



- Three types of business relationships:
  - PC (provider-customer)
  - CP (customer-provider)
  - PP (peer-peer)

Typical Export Policies Route Patterns

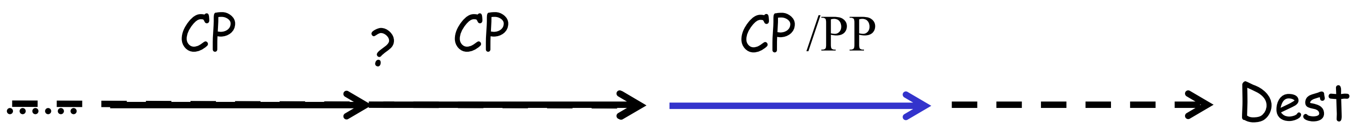
- Invariant 1 of valid BGP routes (with labels representing business relationship)



Reasoning: only route learned from customer is sent to provider; thus after a PC, it is always PC to the destination

Typical Export Policies Route Patterns

- Invariant 2 of valid BGP routes (with labels representing business relationship)



Reasoning: routes learned from peer or provider are sent to only customers; thus all relationship before is CP.

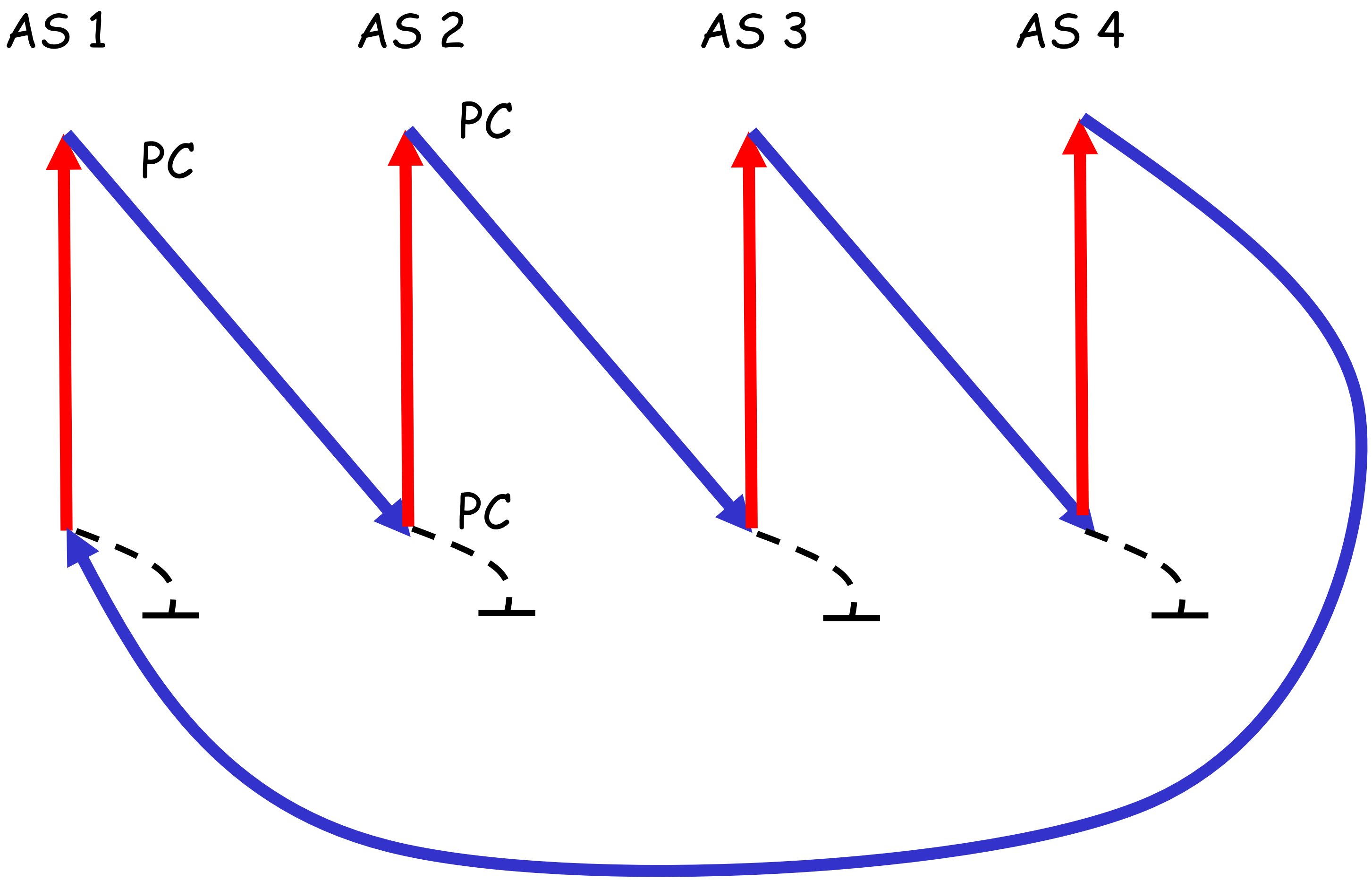
Stability of BGP Policy Routing

- Suppose
  1. there is no loop formed by provider-customer relationship in the Internet
  2. each AS uses typical route selection policy:
     C > E/P
  3. each AS uses the typical export policies

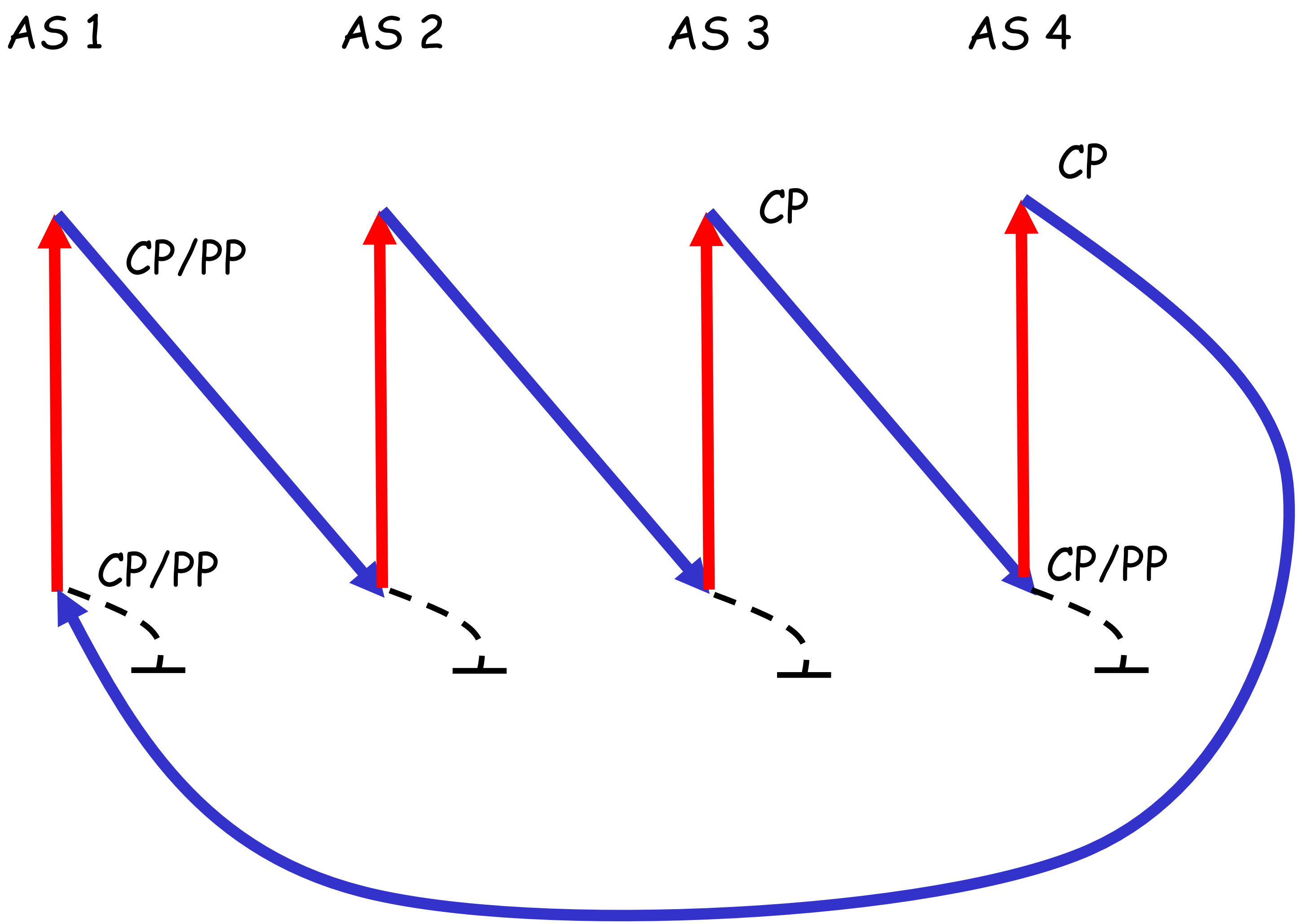
- Then policy routing always converges (i.e., is stable).

Case 1: A Link is PC

Proof by contradiction. Assume a loop in P-graph. Consider a fixed link. in the loop



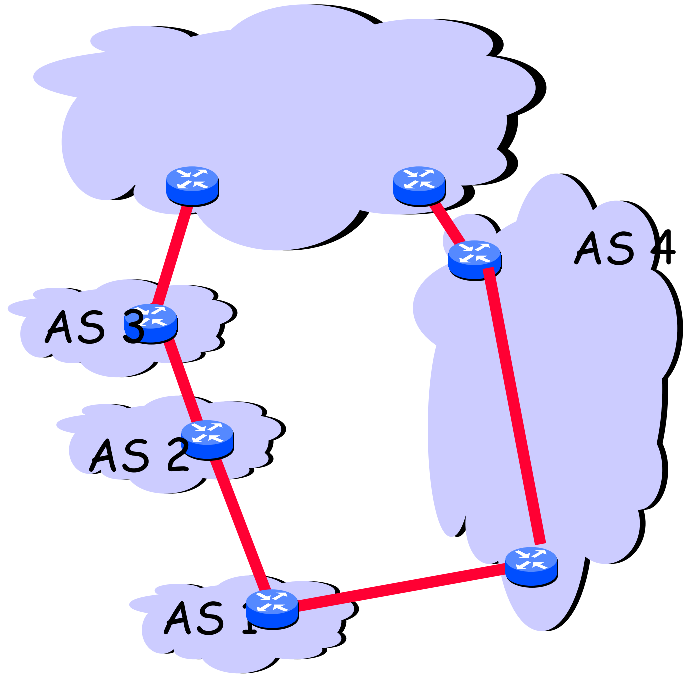
Case 2: Link is CP/PP



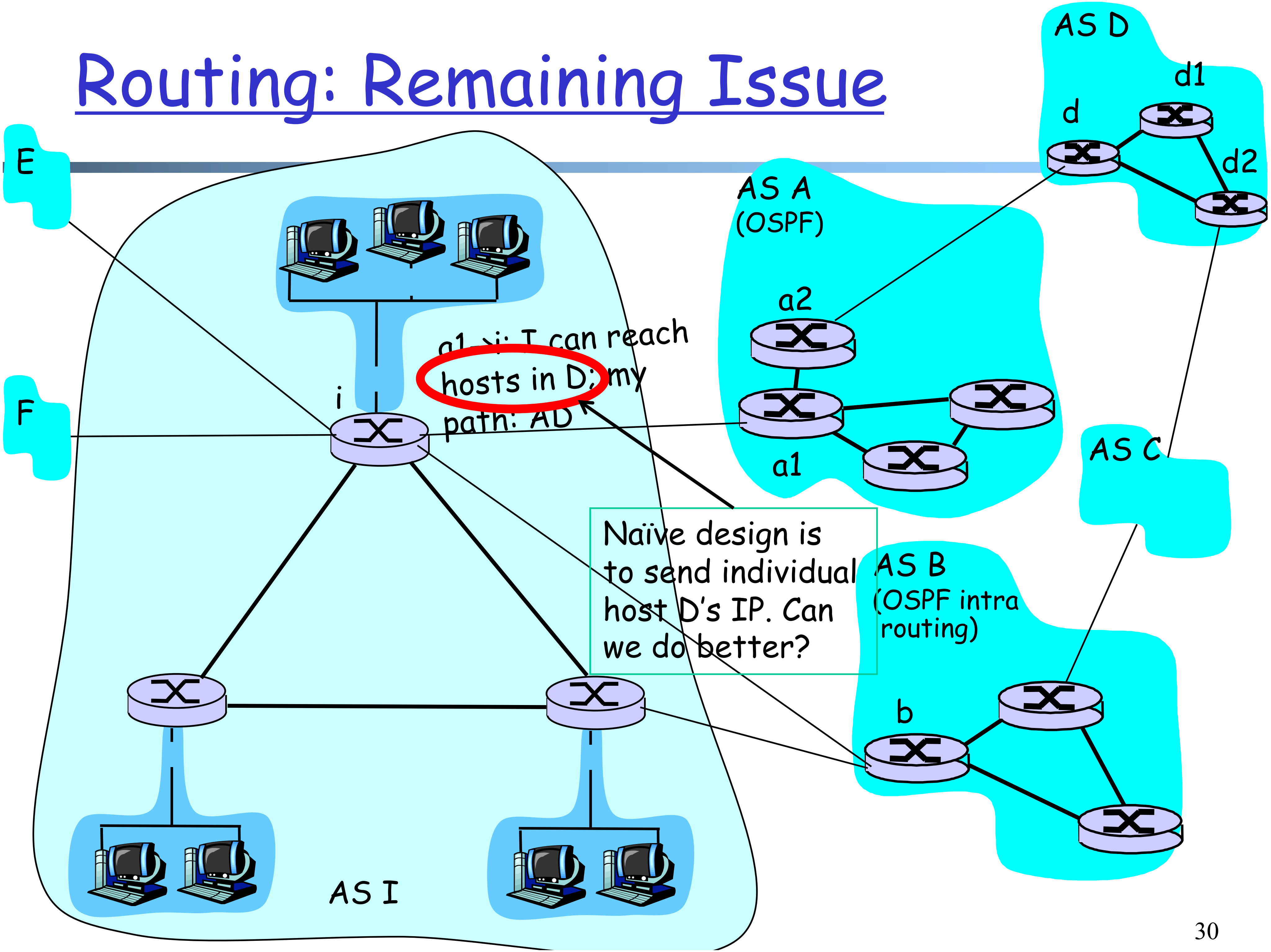
Summary: BGP Policy Routing

- Advantage
  - satisfies current demand

- Issues
  - policy dispute can lead to instability
    - current Internet economy provides a stability framework, but if the framework changes, we may see instability
  - Hierarchical routing can be inefficient



Routing: Remaining Issue



Outline

- Admin and recap
- Network control plane
  - Routing
    - Link weights assignment
    - Routing computation
      - Basic routing computation protocols
      - Global Internet routing
        - Basic architecture
        - BGP (Border Gateway Protocol): The de facto Inter-domain routing standard
          - Basic operations
          - BGP as a policy routing framework (control interdomain routes)
          - Policy/interdomain routing analysis
            - Global preference aggregation and Arrow's Theorem
            - Local preference aggregation
            - Economics and interdomain routing patterns
          - ***IP addresses for Interdomain routing***

IP Addressing Scheme: Requirements

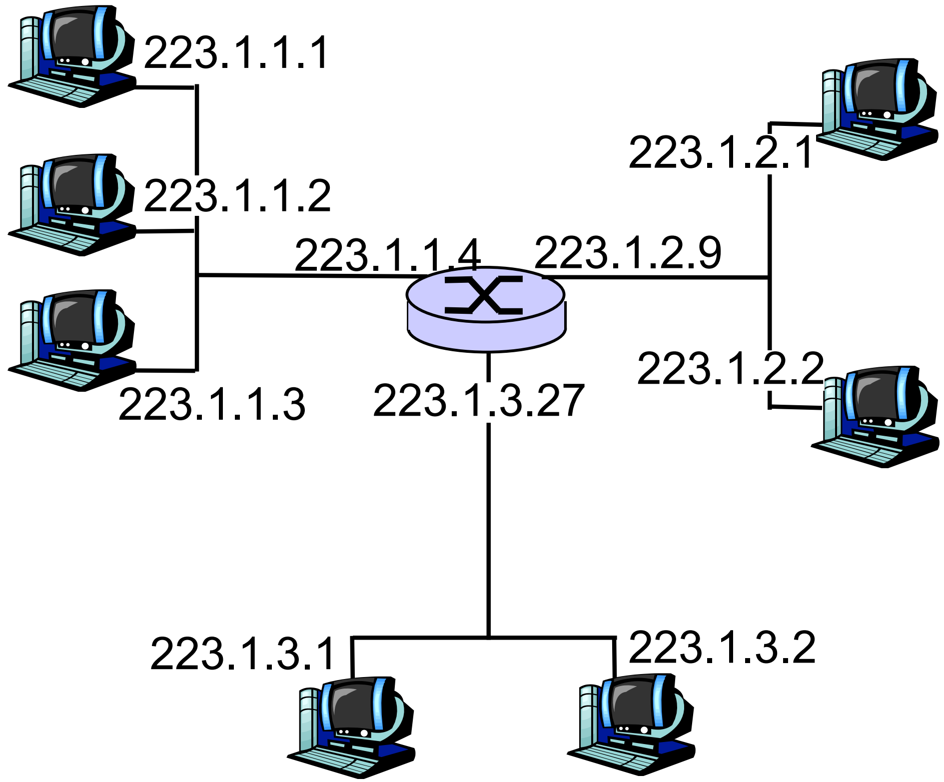
- Uniqueness: We need an address to uniquely identify each destination

- Aggregability : Routing scalability needs flexibility in aggregation of destination addresses
  - we want to aggregate as a large set of destinations as possible in BGP announcements

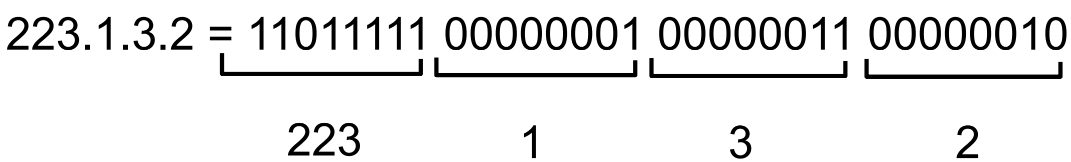
- Current: the unit of routing in the Internet is a classless interdomain routing (CIDR) address

IP Address: Uniqueness

- IPv4 address: A 32-bit unique identifier for an *interface*
- *interface:*
  - routers typically have multiple interfaces
  - host may have multiple interfaces




```
%/sbin/ifconfig -a
```




```
e.g., /etc/sysconfig/network-scripts/ifcfg-enp0s25
%ifup
```

Classless InterDomain Routing (CIDR) Address: Aggregation

- A CIDR address partitions an IP address into two parts
  - A prefix representing the network portion, and the rest (host part)
  - address format: a.b.c.d/x, where x is # bits in network portion of address

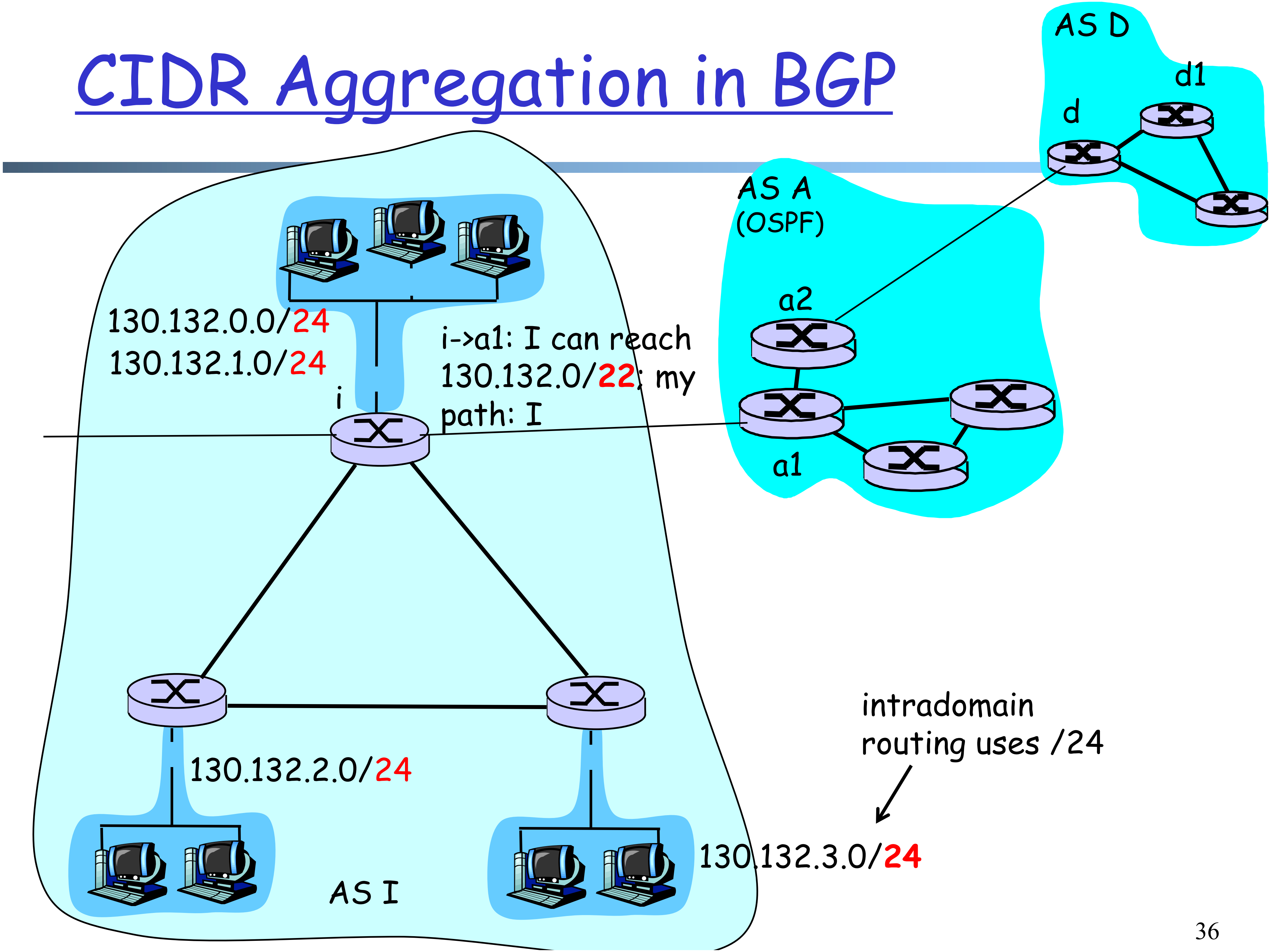
network part | host part

11001000 00010111 00010000 00000000

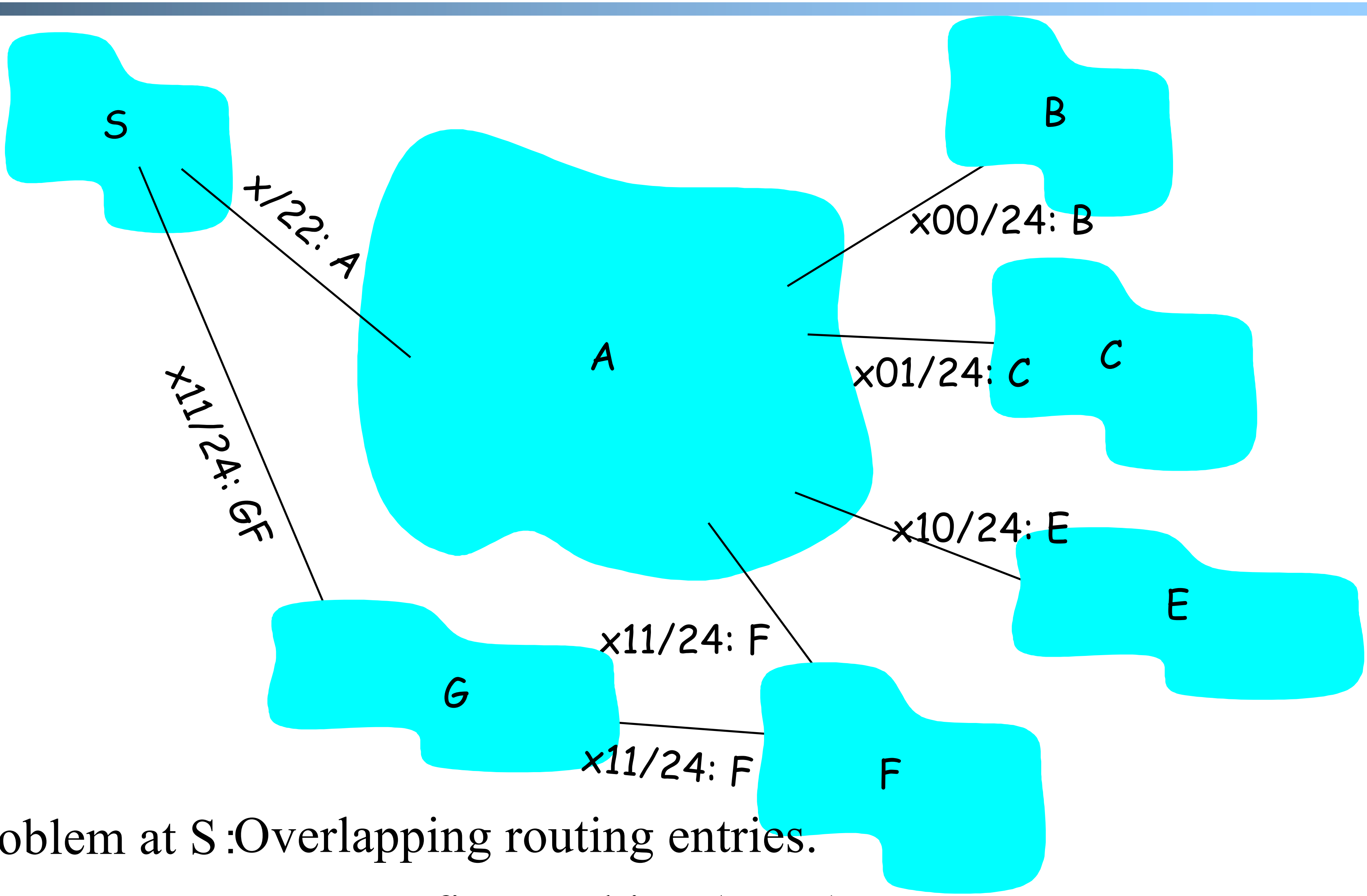
200.23.16.0/23

Some systems use mask (1's to indicate network bits), instead of the /x format

CIDR Aggregation in BGP



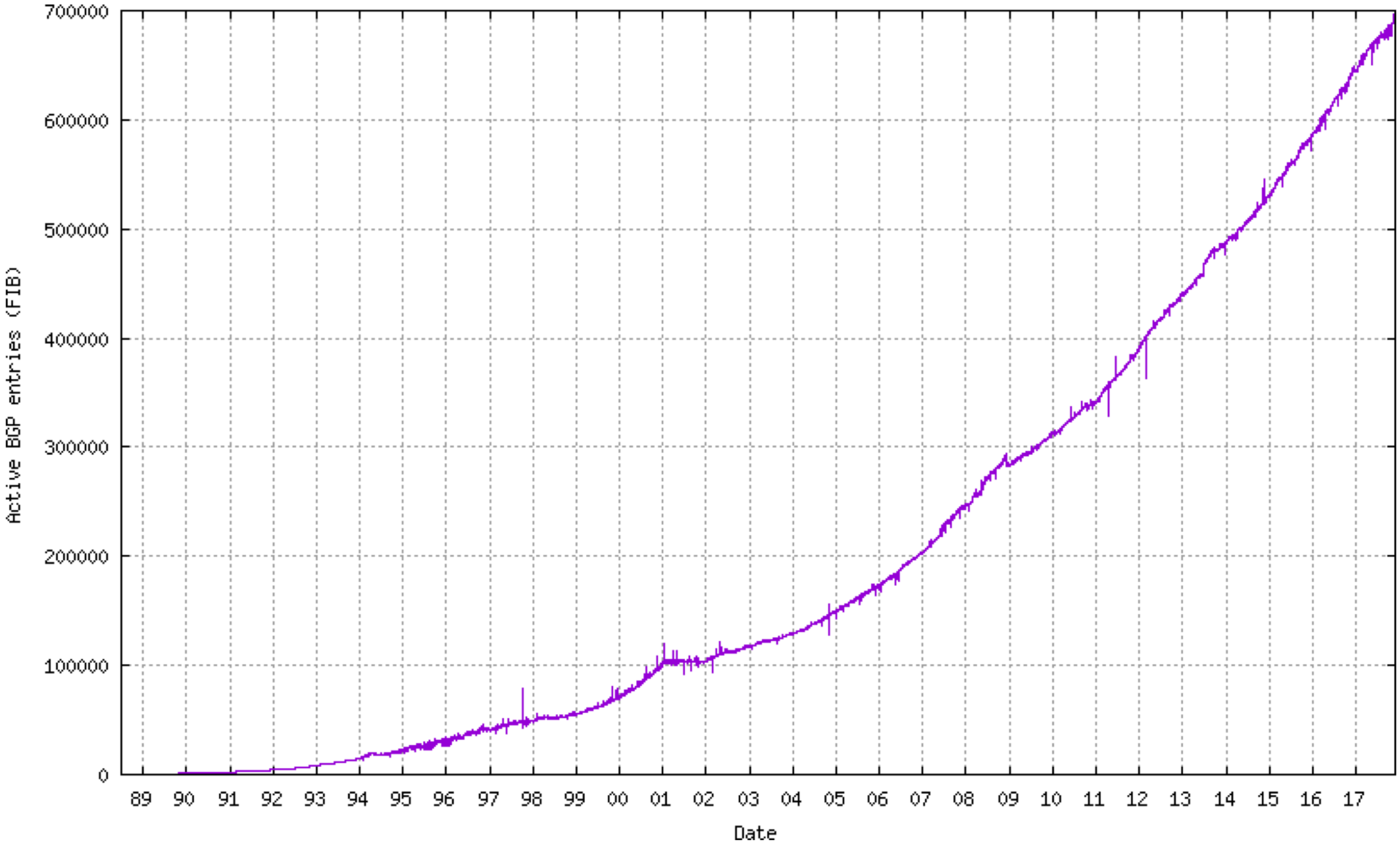
CIDR Aggregation in BGP



Problem at S :Overlapping routing entries.

Solution :   Longest prefix matching (LPM)

Routing Table Size of BGP (number of globally advertised, aggregated entries)



Active BGP Entries (http://bgp.potaroo.net/as1221/bgp-active.html)

Internet Growth (http://www.caida.org/research/topology/as_core_network/historical.xml)