Network: Global Internet Routing, Policy Routing Analysis

Qiao Xiang

https://qiaoxiang.me/courses/cnns-xmuf22/index.shtml

11/29/2022

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

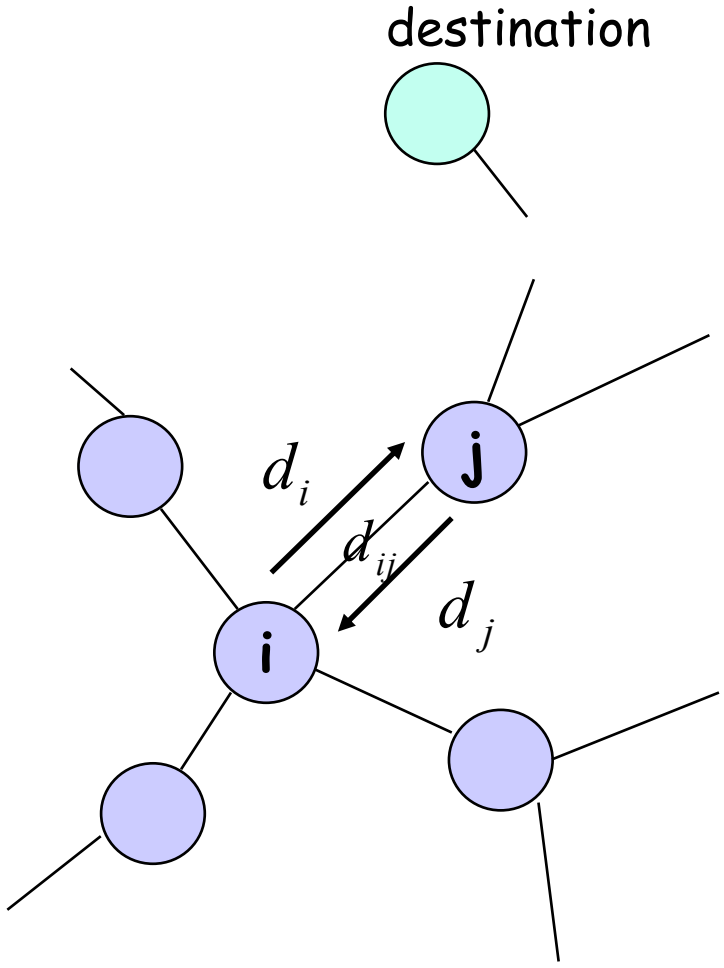
Admin

- Lab 4 due on Dec. 8
  - Time to say goodbye to procrastination ☺
  - When submitting your code, make sure it at least can compile
- Lab 5 to be posted this week
- Class project teams and topics finalized
  - You have until Jan. 5 to finish your project
    - Code, report, slides and pre-recorded presentation (10-15 minutes)
- Guest lecture on Dec. 1

Recap: Routing Computation using Distance Vector/Bellman-Ford Routing

- Distributed computation:
  At node i, computes

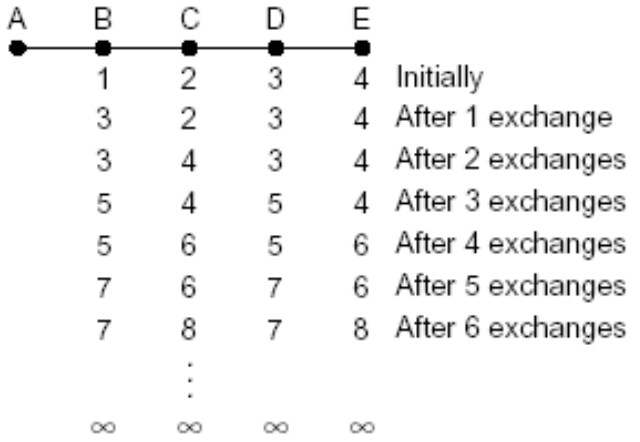
$$d_i = \min_{j \in N(i)} (d_{ij} + d_j)$$



- One way to understand BFA is to consider it as a dynamic programming alg, propagating from dest to other nodes

Recap: Fixing DV/BFA

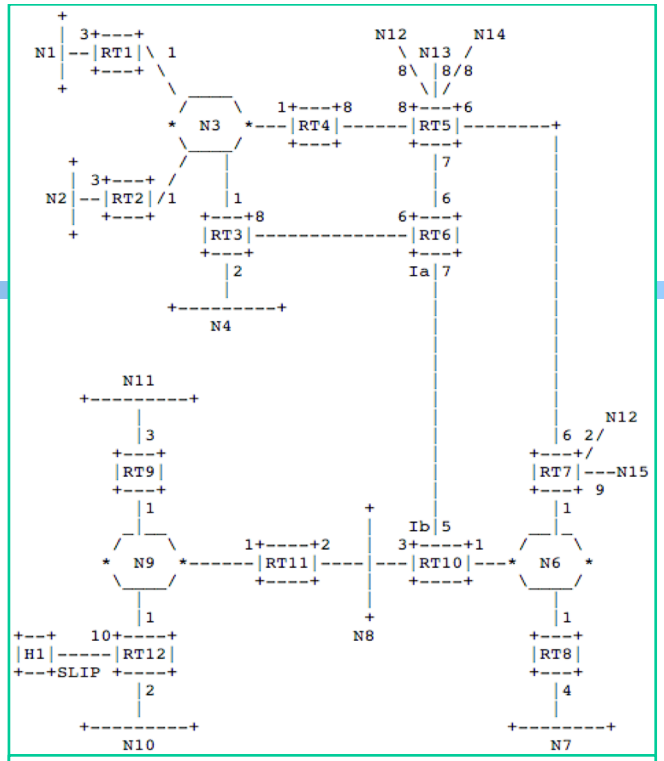
- Property of BFA
  - Bad news may propagate slowly due to loops



- Techniques
  - Reverse poison
    - Avoid two-node loops
  - DSDV
    - Using destination seq to partition into epochs
    - A good example of analysis using global invariants
  - Diffusive Update Alg (DUAL)
    - Utilize backup routes

Recap: Link State Routing

- Basic idea: instead of distributed computing of routes, only distributed state distribution (synchronization)
- Link state distribution can still have much complexity, e.g.,
  - out of order delivery
  - partition and reconnect
  - scalability



FROM

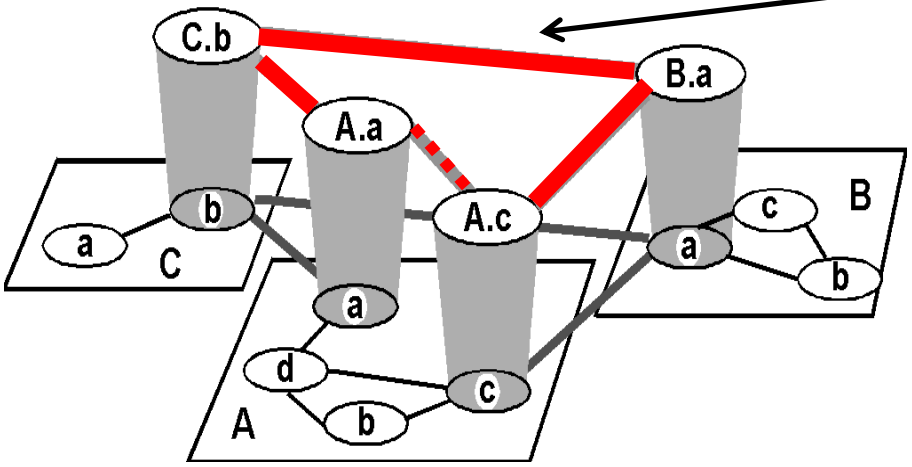
	RT 1	RT 2	RT 3	RT 4	RT 5	RT 6	RT 7	RT 8	RT 9	RT 10	RT 11	RT 12	N3	N6	N8	N9
RT1													0			
RT2													0			
RT3						6							0			
RT4				8									0			
RT5			8			6	6									
RT6		8		7						5						
RT7				6										0		
* RT8														0		
* RT9																0
T RT10						7								0	0	
O RT11															0	0
* RT12																0
* N1	3															
N2		3														
N3	1	1	1	1												
N4			2													
N6							1	1		1						
N7								4								
N8										3	2					
N9									1		1	1				
N10												2				
N11									3							
N12				8		2										
N13				8												
N14				8												
N15							9									
H1												10				

Figure 3: The resulting directed graph

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

Routing with Autonomous Systems



Inter-AS routers form an overlay

Gateway routers of same AS share learned external routes using iBGP.

Gateway routers participate in intradomain to learn internal routes.

AS C (RIP intra routing)

AS A (OSPF intra routing)

AS B (OSPF intra routing)

a

b

e

eBGP

iBGP

d

c

f

g

i

h

Gateway routers of diff. auto. systems exchange routes using eBGP

Summary: Internet Routing Architecture

- Autonomous systems have flexibility to choose their own intradomain routing protocols
  - allows autonomy
- Only a small # of routers (gateways) from each AS in the interdomain level
  - improves scalability
- Interdomain routing using AS topology instead of detailed topology
  - improves scalability/privacy

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*

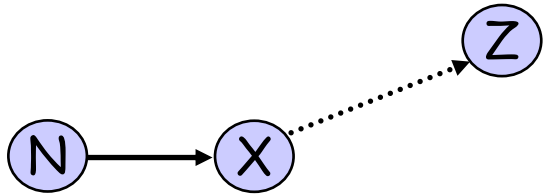
BGP Basic Operations

- BGP is **a Path Vector** protocol
  - similar to Distance Vector protocol
  - a border gateway sends to a neighbor *entire path* (i.e., a sequence of ASNs) to a destination, e.g.,
    - gateway X sends to neighbor N its path to dest. Z:

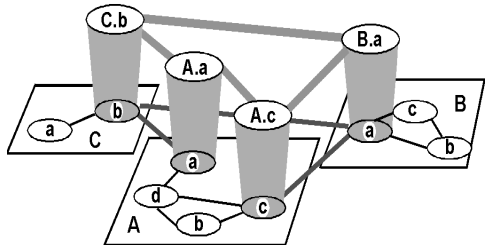
$$\text{path } (X,Z) = X,Y1,Y2,Y3,\ldots,Z$$

  - if N selects path(X, Z) advertised by X, then:

$$\text{path } (N,Z) = N, \text{ path } (X,Z)$$



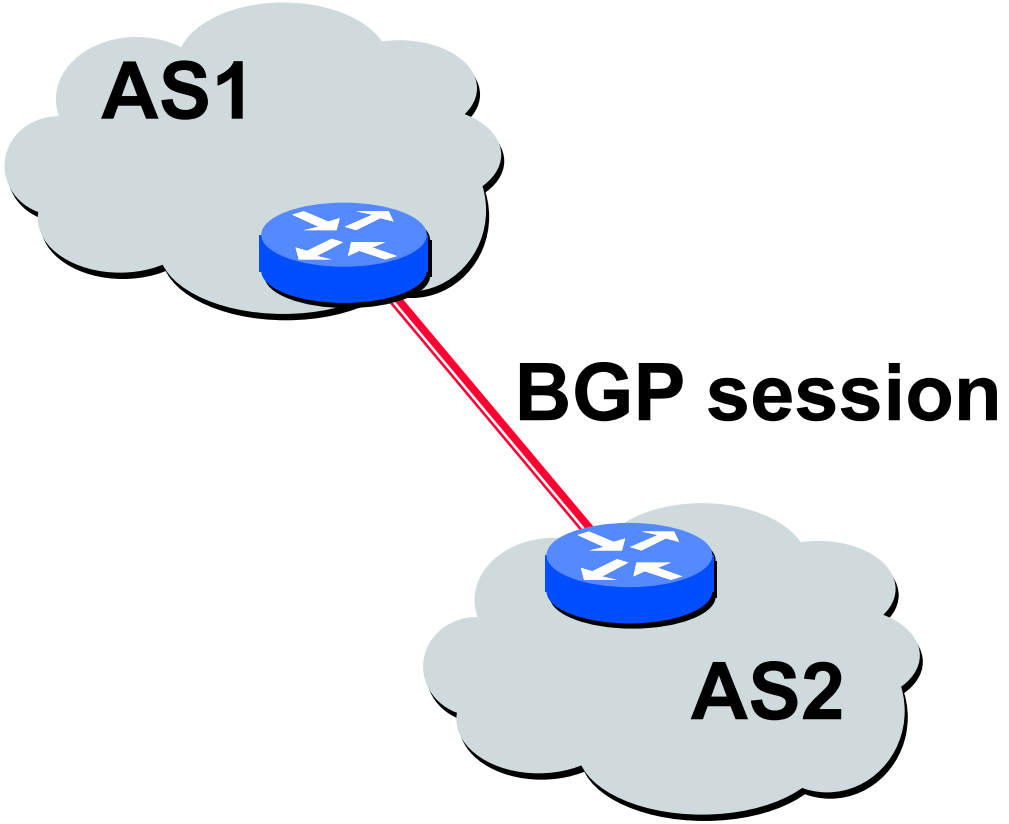
BGP Basic Operations



Establish session on TCP port 179

Exchange all active routes

Exchange incremental updates



while (connection is ALIVE)
exchange UPDATE message
select best available route
if route changes, export to neigh.

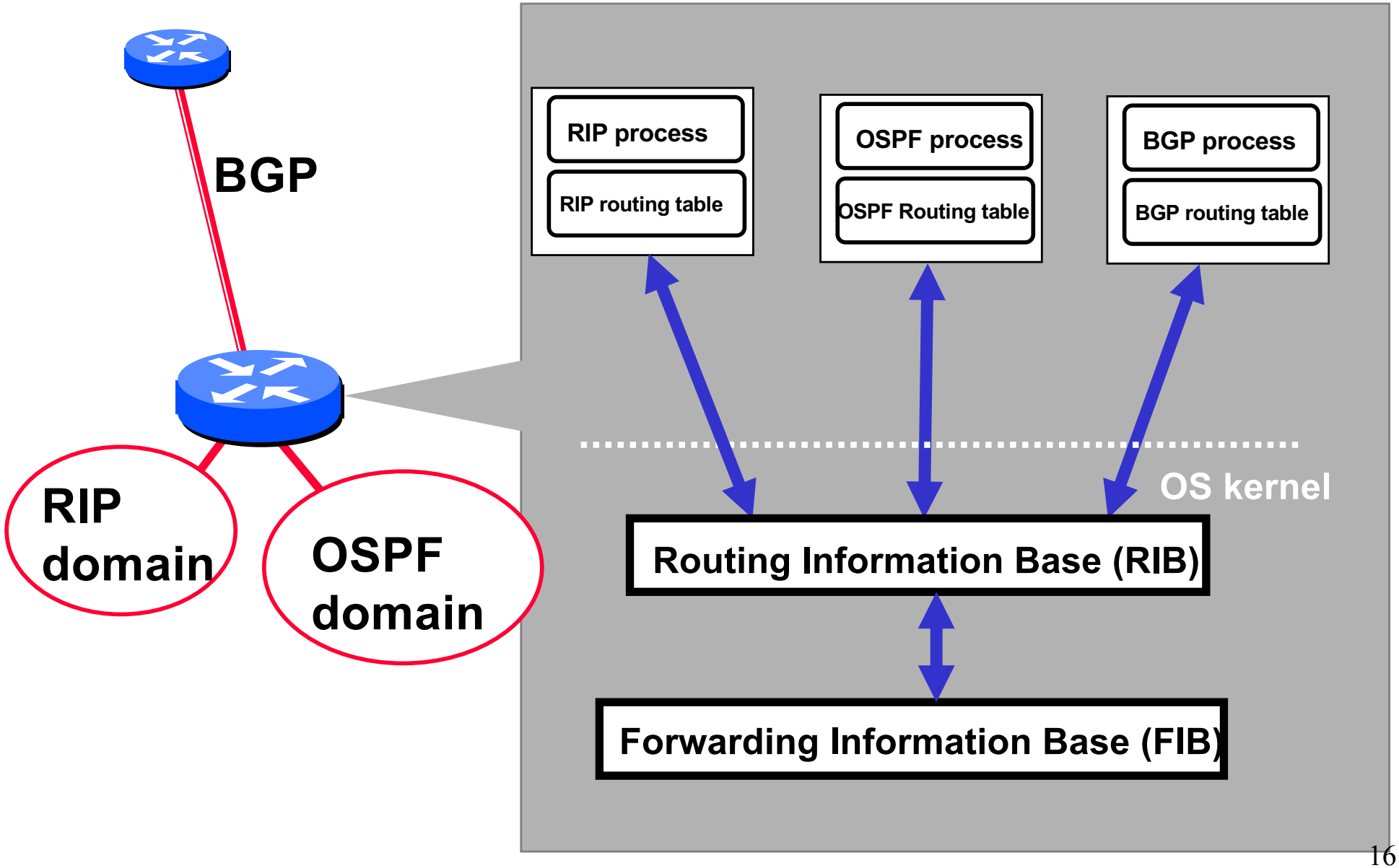
BGP Messages

- Four types of messages
  - OPEN: opens TCP connection to peer and authenticates sender
  - UPDATE: advertises new path (or withdraws old)
  - KEEPALIVE keeps connection alive in absence of UPDATES; also ACKs OPEN request
  - NOTIFICATION: reports errors in previous msg; also used to close connection

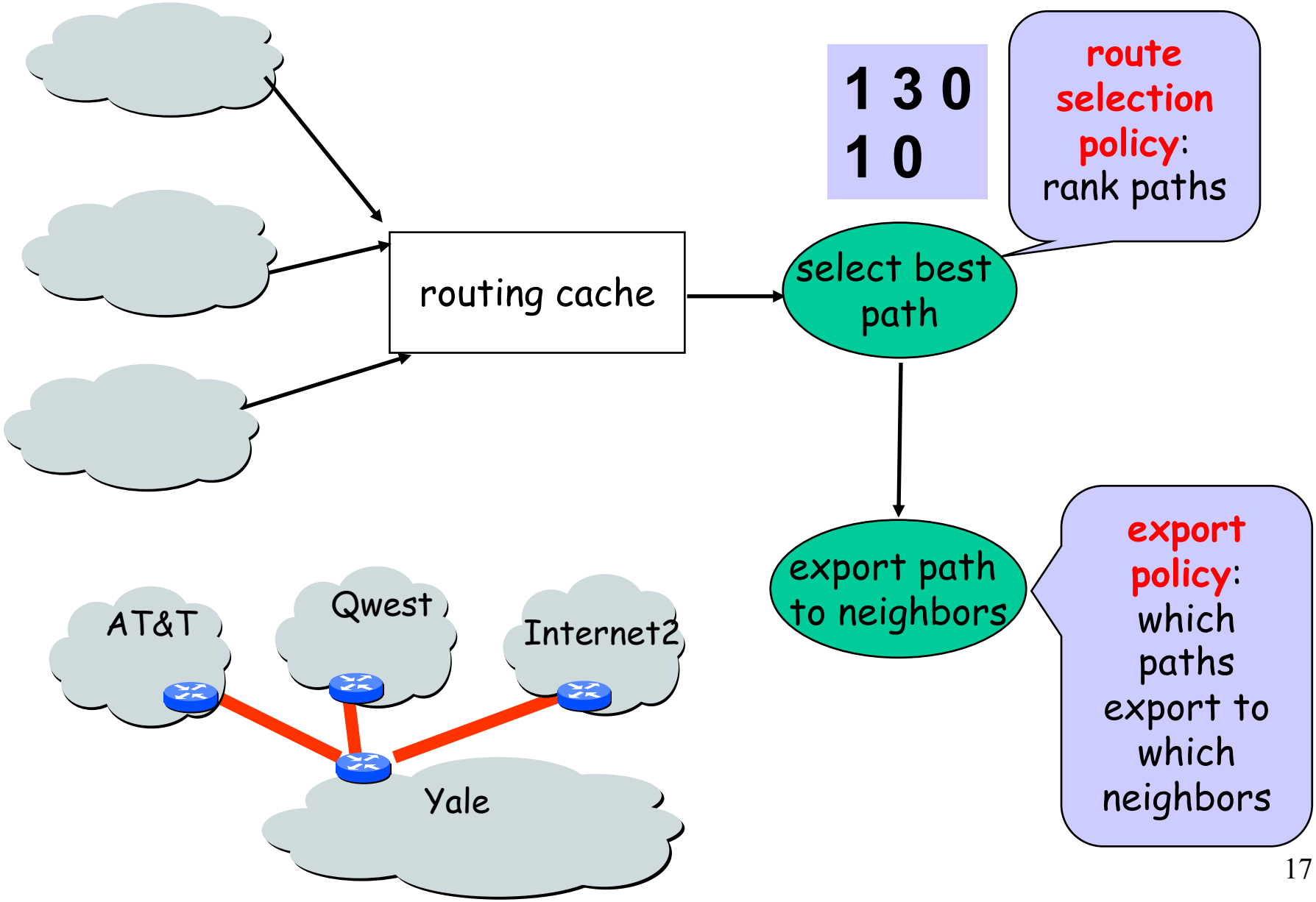
Outline

- 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)*

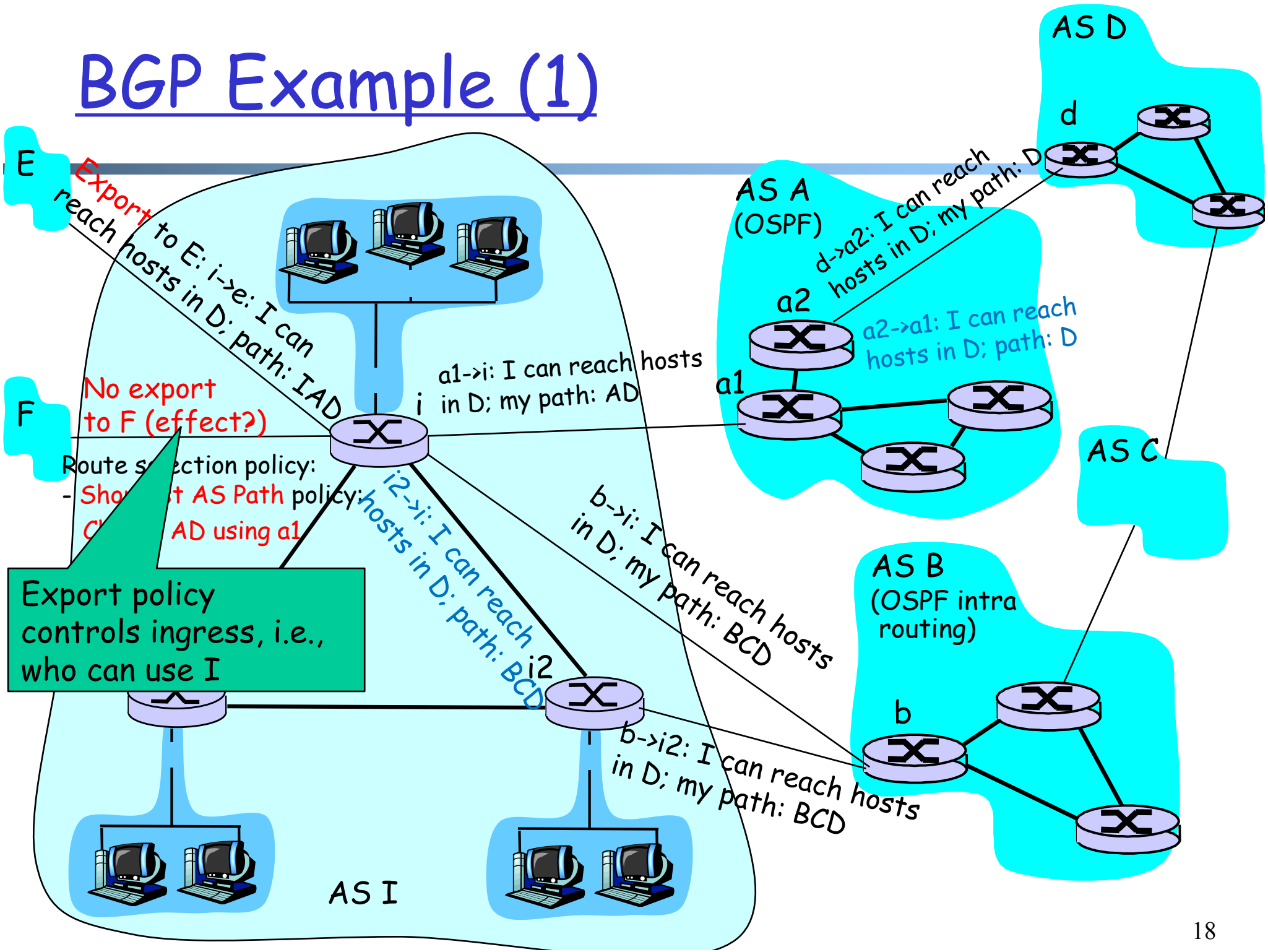
BGP Policy Routing Framework: Label Route Information Sources



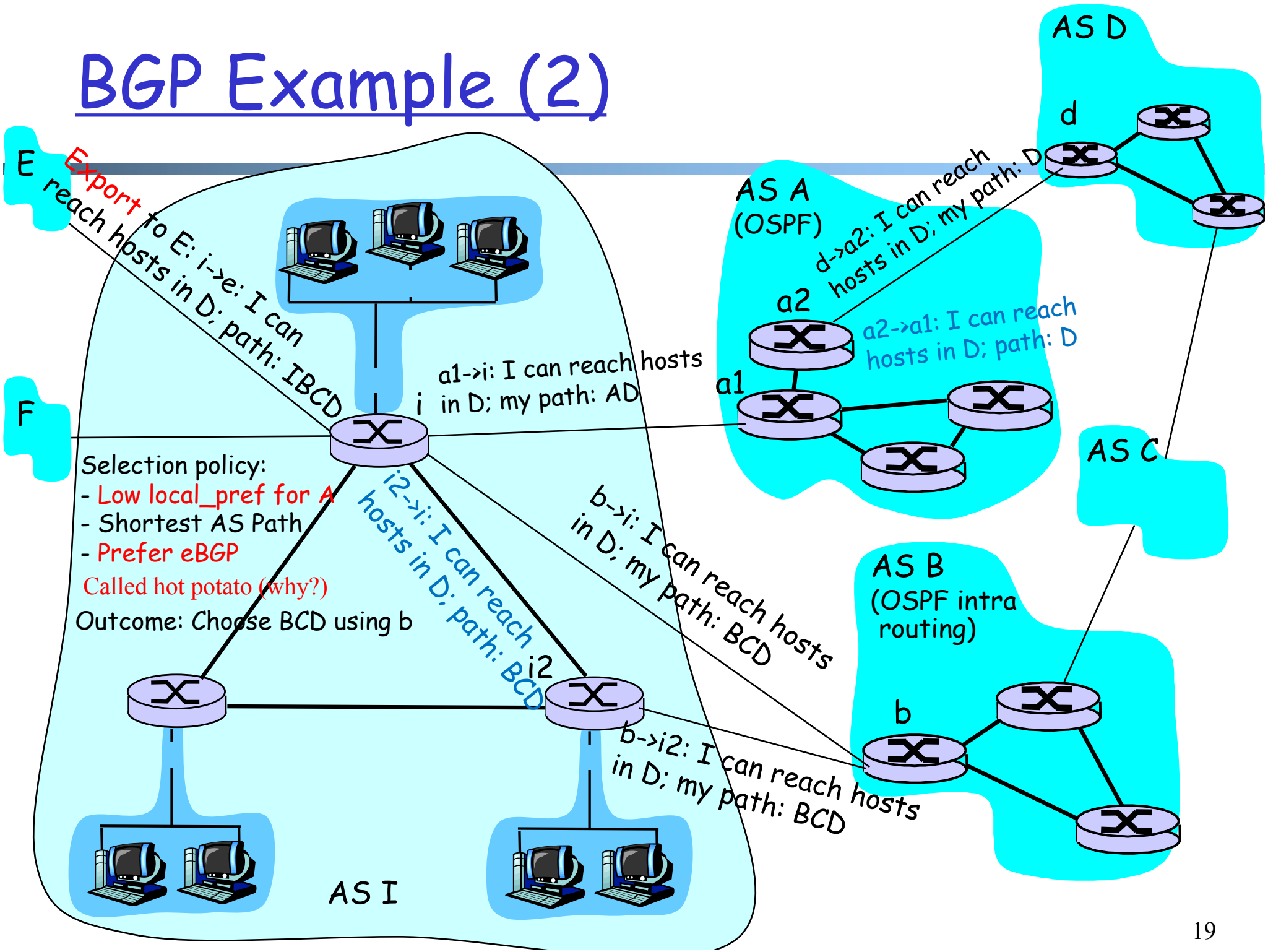
BGP Policy Routing Framework: Decision Components



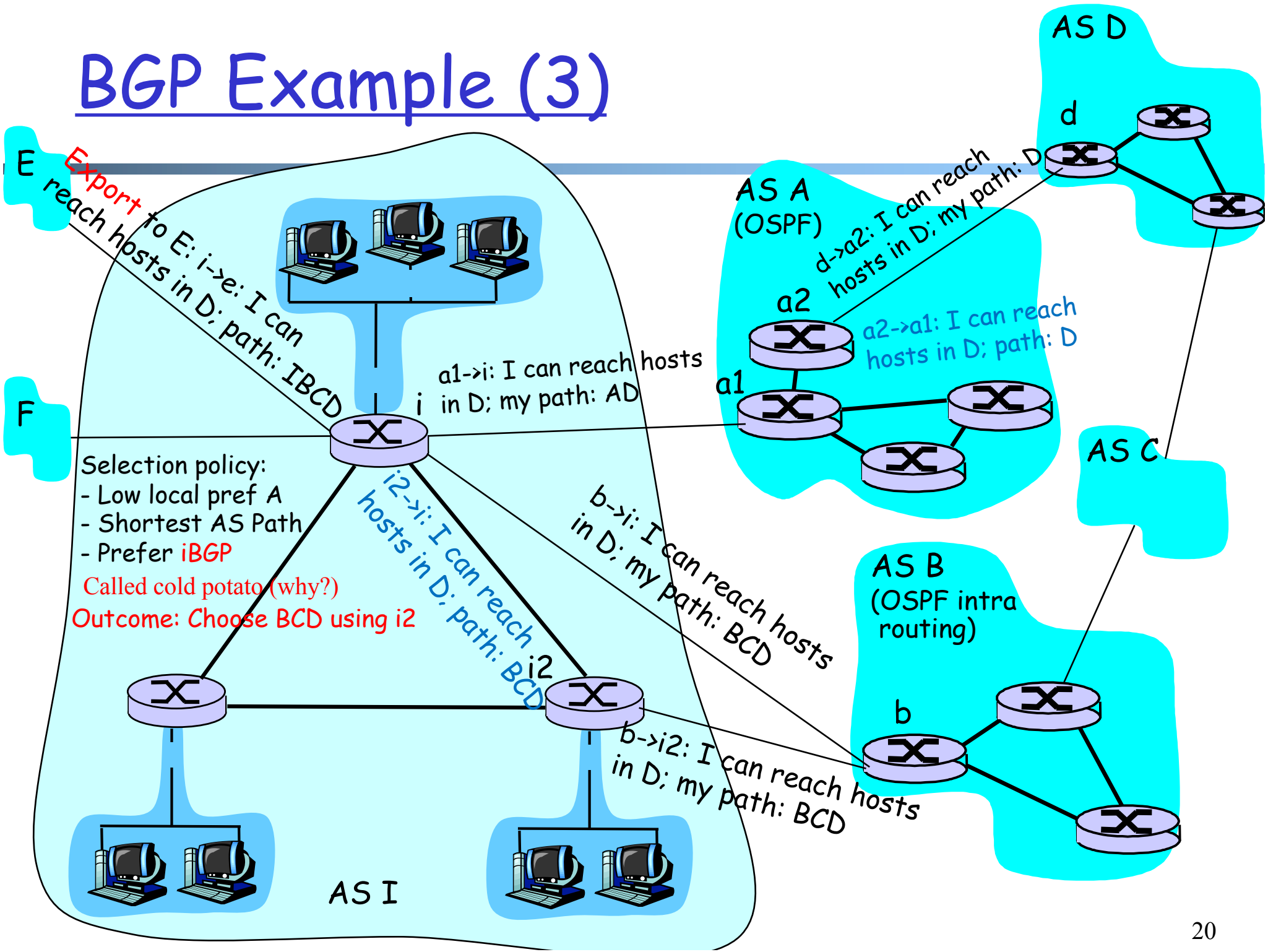
BGP Example (1)



BGP Example (2)



BGP Example (3)



Observing BGP Paths

- Using one of the looking glass servers:
  http://www.bgp4.as/looking-glasses

  https://www.gin.ntt.net/looking-glass/

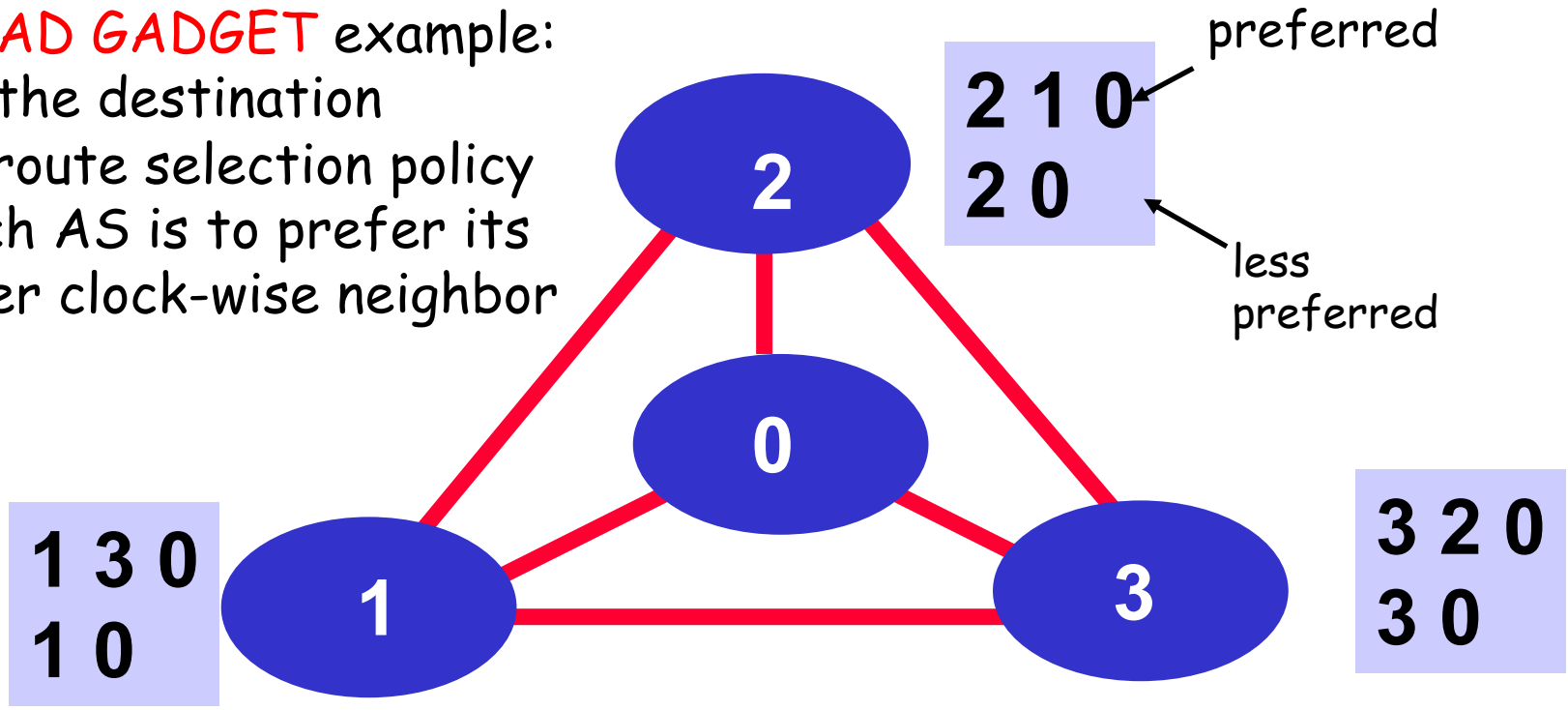
Outline

- 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*

Motivation: Policy Routing Stability

- A policy routing system can be considered as a system to aggregate local 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

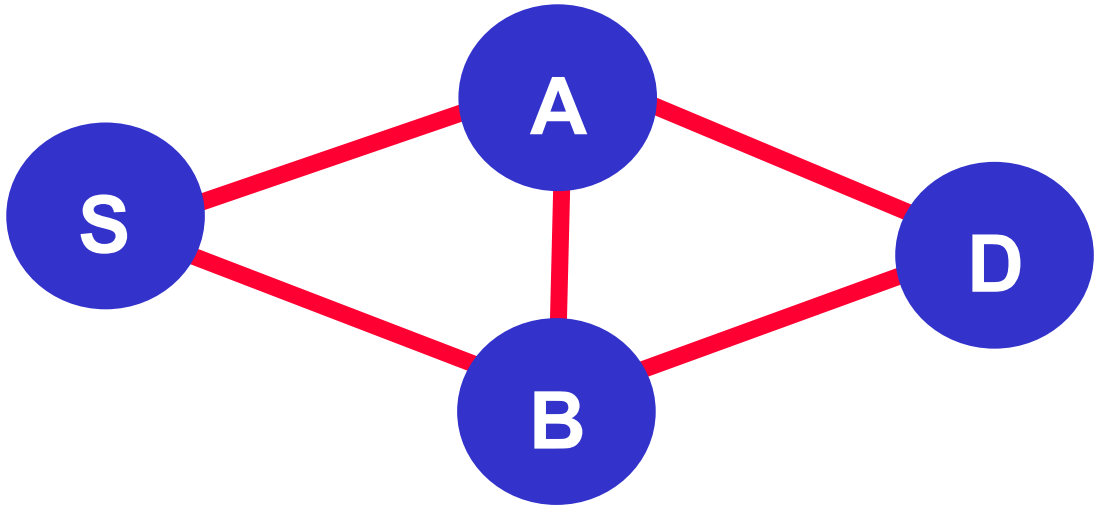


Policy (preferences) aggregation fails: routing instability !

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

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 Aggregation Framework

- 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.