Advanced Computer Networks

Border Gateway Protocol

Jianping Pan
Summer 2007

Feedback on S1

- Reading
- Presentation
- Project

Review: routing

- Internet addressing
- Internet routing
  - distance vector routing
  - link state routing
  - hierarchical routing

Internet structure

- Tiered service provider networks
  - tier-1 service provider
    - not a customer of other service providers
    - global coverage
  - tier-2 service provider
    - a customer of tier-1service providers
    - a service provider of lower-tier service providers
    - continent/country/region coverage
  - tier-3 service provider
    - a service provider of end customers
    - local coverage
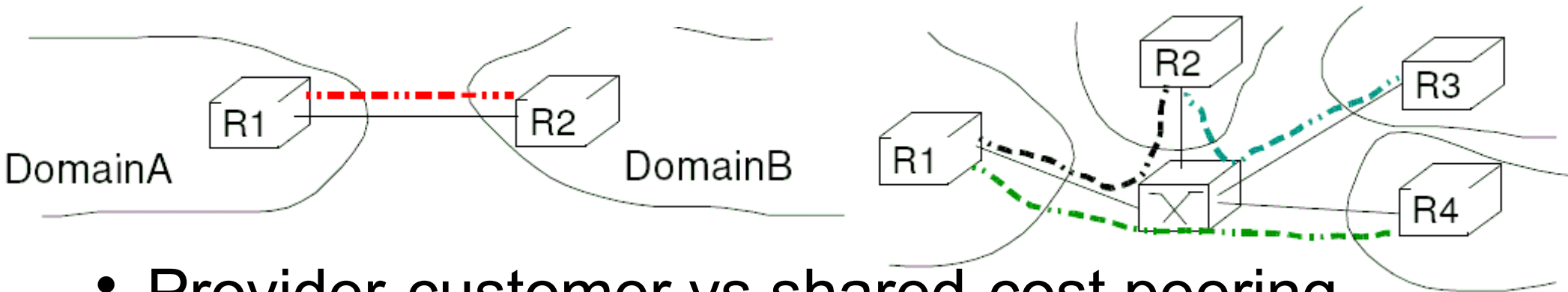
Inter-domain routing

- Domain
  - normally, Autonomous System (AS)
- Transit domain
  - normally service providers
- Stub domain
  - normally regular customers
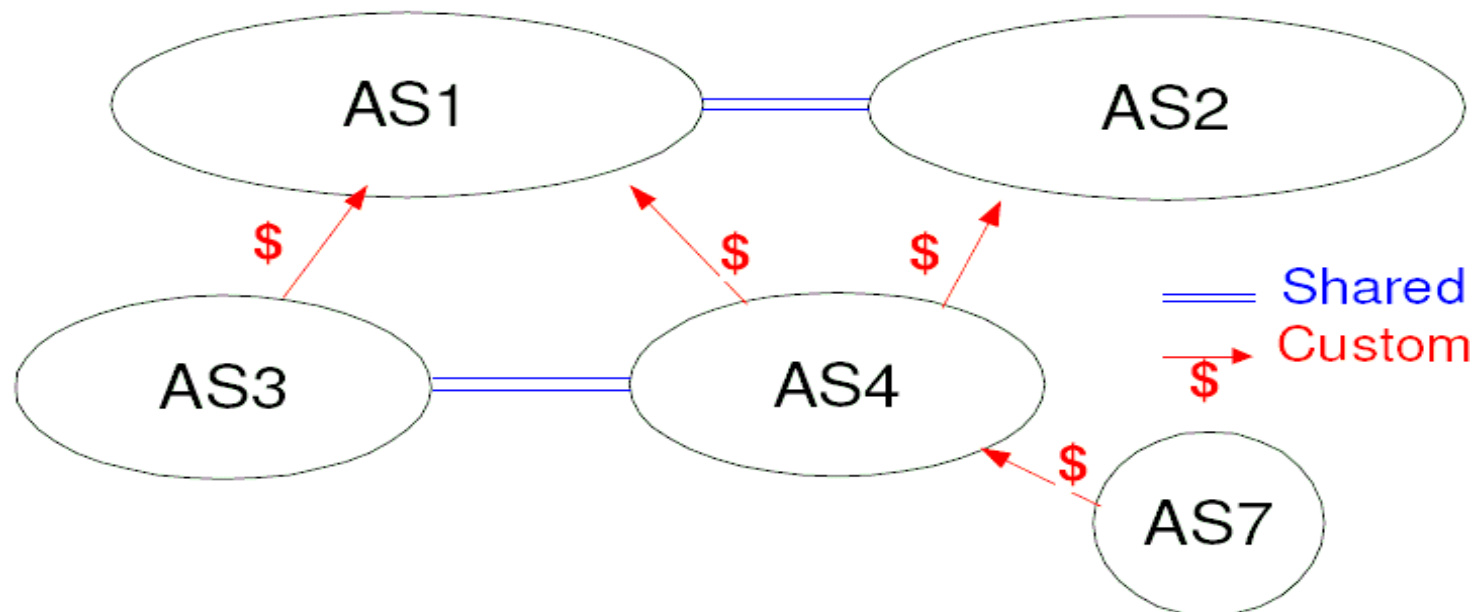 - single-homed
  - multi-homed

6/27/07

http://www.info.ucl.ac.be/~obo/BGP/
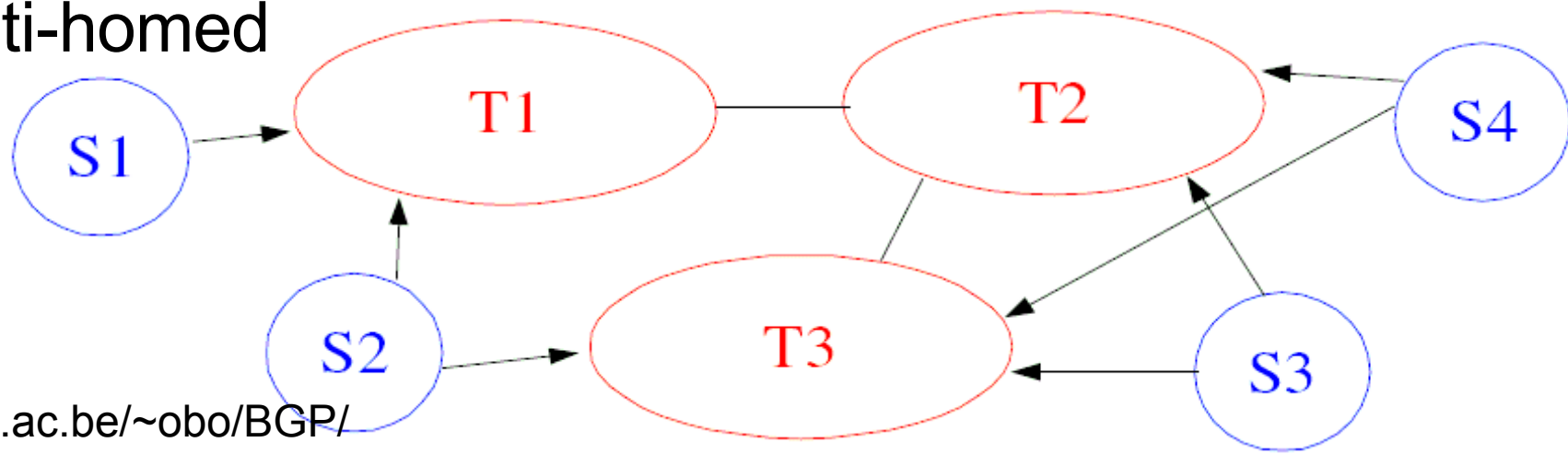
Inter-domain peering

- Private vs public peering

- Provider-customer vs shared-cost peering

Routing policies

- Inter-domain routing is mostly policy-based
  - connectivity
  - peering
  - contract, etc.
- Routing policy specification language
  - import filter
    - from AS# accept list-of-AS
  - export filter
    - to AS# announce list-of-AS

Border gateway protocol
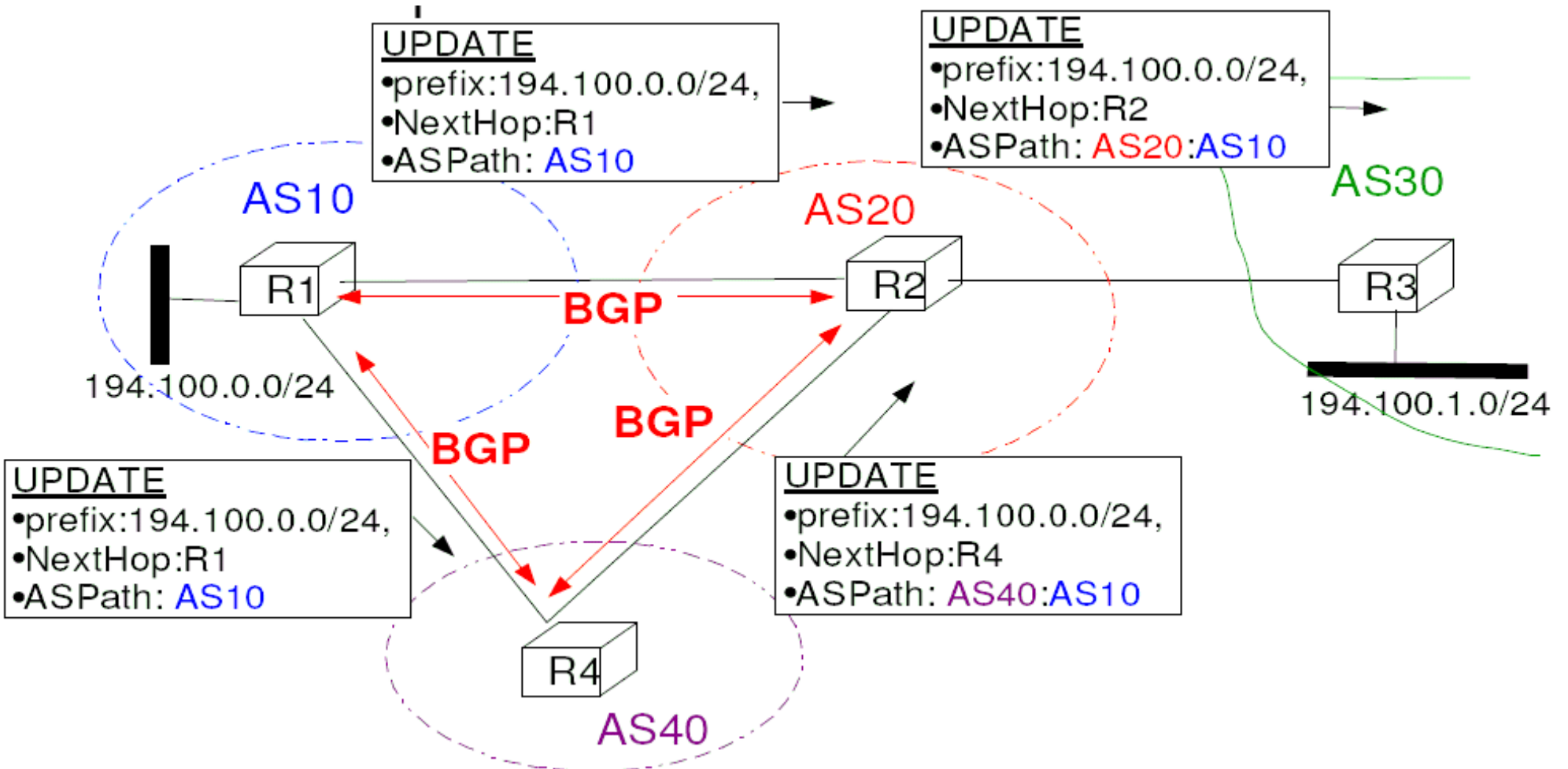
- BGP, currently version 4 [RFC1771]
    - transported by TCP: port 179
    - path vector routing
        - prefix: AS path
        - path vector for loop detection
        - only single path advertised
    - incremental update
        - initially advertise known routes to all prefixes
        - then only update routes that change
        - intended to be stable during operation

BGP messages

- OPEN
  - establish BGP session
- UPDATE
  - route announcement or withdraw
- NOTIFICATION
  - error notification
- KEEPALIVE
  - ensure at least one message every 30 seconds
- ROUTE_REFRESH
  - used to support graceful restart

BGP in action

- Example

UPDATE
- prefix:194.100.0.0/24,
- NextHop:R1
- ASPath: AS10

UPDATE
- prefix:194.100.0.0/24,
- NextHop:R2
- ASPath: AS20:AS10

AS30

AS10

AS20

R1

BGP

R2

R3

194.100.0.0/24

BGP

BGP

194.100.1.0/24

UPDATE
- prefix:194.100.0.0/24,
- NextHop:R1
- ASPath: AS10

UPDATE
- prefix:194.100.0.0/24,
- NextHop:R4
- ASPath: AS40:AS10

R4

AS40

Q: how to obtain "prefix"?
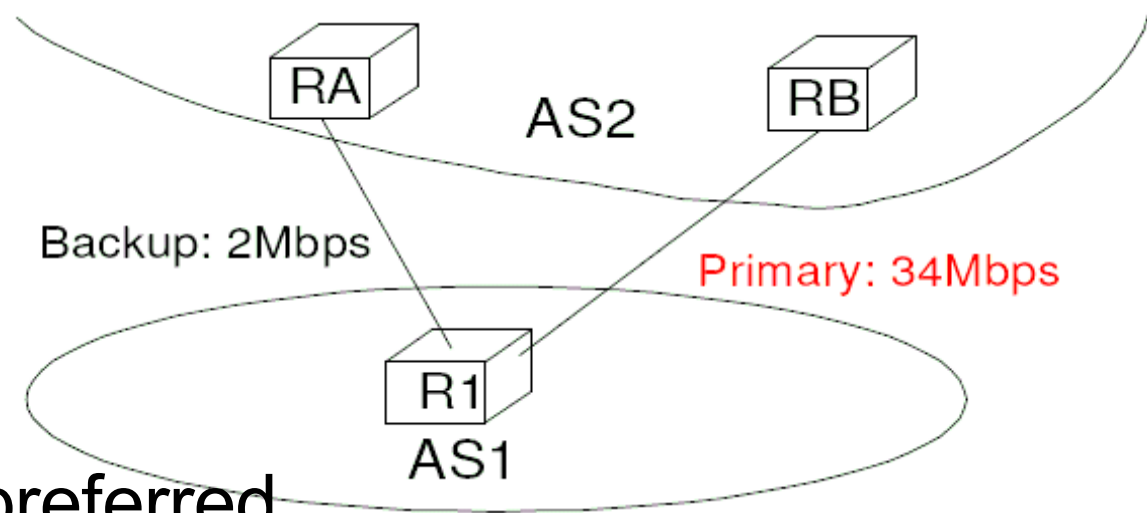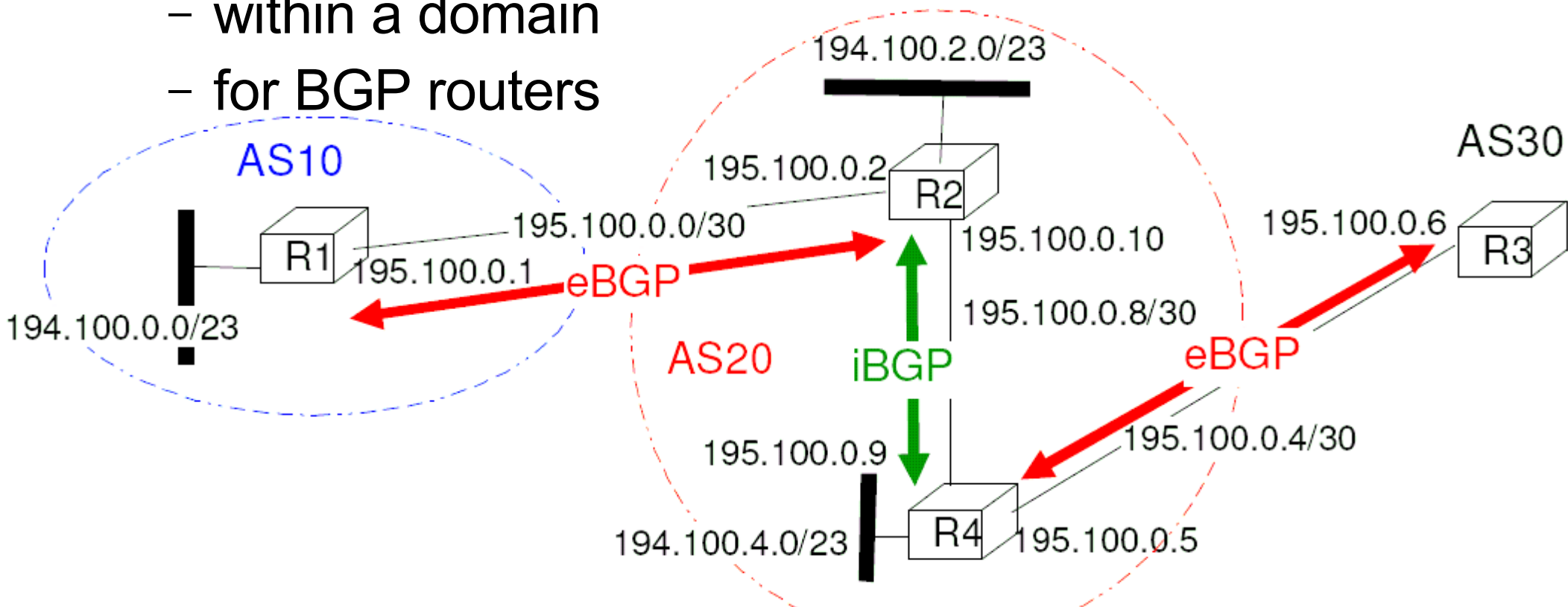
Path selection

- Most specific route
  - longest prefix match
- Local preference
  - highest preference preferred
- AS path length
  - shortest path preferred
- Multi-exit discriminator (MED)
  - lowest MED preferred
- There are other selection criteria

eBGP and iBGP

- eBGP

- iBGP
  - within a domain
  - for BGP routers

194.100.2.0/23

AS10

AS30

195.100.0.2

R2

195.100.0.6

R3

R1

195.100.0.0/30

195.100.0.10

195.100.0.1

eBGP

195.100.0.8/30

194.100.0.0/23

AS20

iBGP

eBGP

195.100.0.9

195.100.0.4/30

194.100.4.0/23

R4

195.100.0.5

Student presentation

- Dale Lyons: "talky" BGP
    - [LMJ97] C. Labovitz, G. R. Malan, and F. Jahanian, "Internet Routing Instability". In Proceedings of ACM SIGCOMM'97, September 1997.

Further discussion

- MRAI
  - minimal routing advertisement interval
    - e.g. 30 seconds
- BGP dampening
  - vendor-dependent
  - e.g., Cisco
    - withdraw penalty: 1000
    - cutoff threshold: 2000
    - reuse threshold: 750
    - half time: 15 minutes

This lecture

- BGP
  - path vector routing
    - prefix: AS-path
  - policy-based routing
    - filter, local preference, MED
  - design goal
    - scalable and stable
  - reality check

- Explore further
  - http://www.cl.cam.ac.uk/~tgg22/interdomain/
  - http://www.routeviews.org/

Next lectures

- July 4
  - [GR00] Lixin Gao and Jennifer Rexford, "Stable Internet Routing Without Global Coordination". In Proceedings of the 2000 ACM SIGMETRICS international conference on Measurement and modeling of computer systems. 2000.

- Course project checkpoint