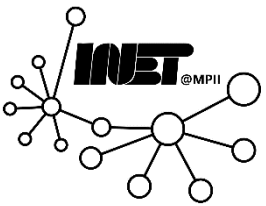
Routing: BGP

Prof. Anja Feldmann, Ph.D.



Inter-AS routing: BGP

The *de facto* standard: ***Border Gateway Protocol (BGP)***

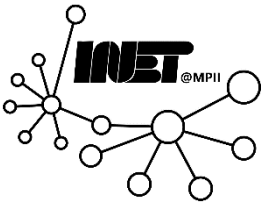
BGP provides each ***AS*** a means to:

- Obtain subnet reachability information from neighboring ASs
- Propagate reachability information to all routers in the AS
- Determine *"good"* routes to subnets based on reachability information and routing policy.

Allows a subnet to advertise its existence to rest of the Internet: *"I am here"*

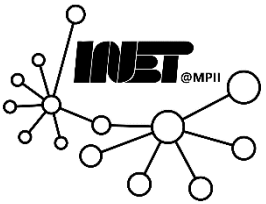
Issues:

- Which routing algorithm?
- How are routes advertised?
- How to implement routing policies?



BGP-4

- BGP = Border Gateway Protocol
- Is an exterior routing protocol (EGP)
- Is a Policy-Based routing protocol
- Is the de facto EGP of today's global Internet
- Has a reputation for being complex
- Supports hierarchical routing
- Is a distance vector protocol

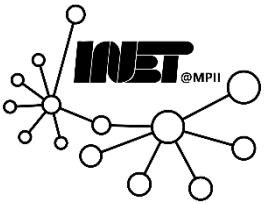


BGP history

- 1989: BGP-1 [RFC 1105]
  - Replacement for EGP (1984, RFC 904)
- 1990: BGP-2 [RFC 1163]
- 1991: BGP-3 [RFC 1267]
- 1995: BGP-4 [RFC 1771] (only 57 pages!)
  - Support for CIDR

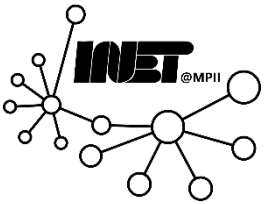
Changes primarily driven by scalability issues

Development dominated by Cisco



Routing tasks: BGP

- Neighbor?
  - Discovery
  - Maintenance
- Database?
  - Granularity
  - Maintenance – updates
  - Synchronization
- Routing table?
  - Metric
  - Calculation
  - Update



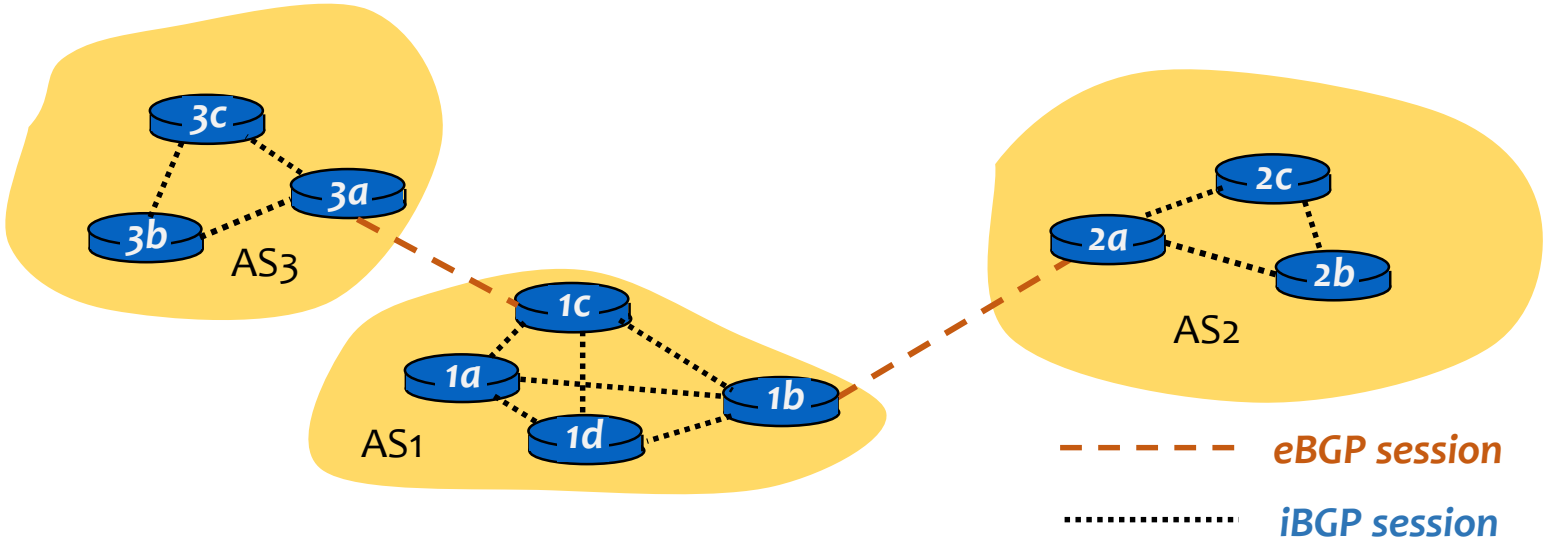
BGP Basics

Pairs of routers (***BGP peers***) exchange routing info over *semi-permanent TCP* connections:

BGP sessions

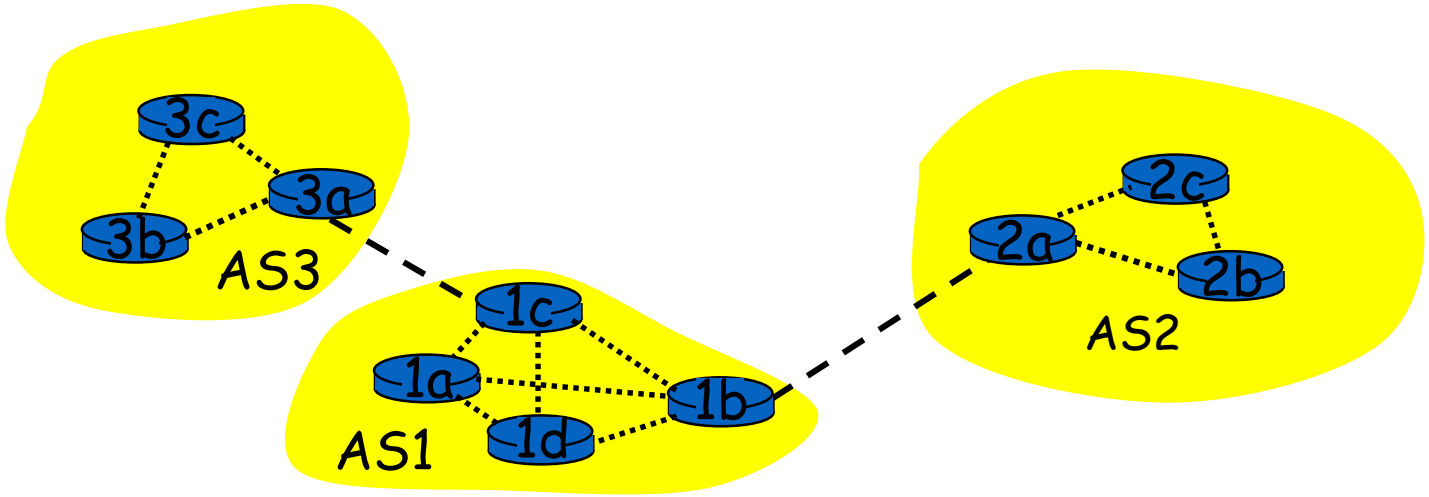
- *Note that BGP sessions do not correspond to physical links.*

- When ***AS2*** advertises a prefix to ***AS1***, ***AS2*** is promising it will forward any datagrams destined to that prefix towards the prefix.
  - ***AS2*** can aggregate prefixes in its advertisement



Distributing reachability info

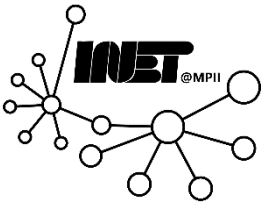
- With eBGP session between 3a and 1c, AS3 sends prefix reachability Info to AS1
- 1c then use iBGP to distribute this new prefix reach. Info to all routers in AS1
- 1b then re-advertise the new reach. Info to AS2 over 1b-to-2a eBGP session
- When routers learn about a new prefix, they create entries for the prefix their forwarding table



Routing policy

Reflects goals of network provider

- *Which routes to accept from other ASes?*
- *How to manipulate the accepted routes?*
- *How to propagate routes through network?*
- *How to manipulate routes before they leave the AS?*
- *Which routes to send to another AS?*

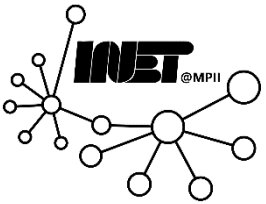


Policies with BGP

BGP provides capabilities for enforcing various policies

- Policies are ***not*** part of BGP!
- Policies are ***used to configure*** BGP

- BGP ***enforces*** policies by …
  - Choosing paths from multiple alternatives
  - Controlling advertisements to other ASes



Routing policy: Examples

- ***Honor business relationships***

  (e.g., customers get full-table; peers only customer prefixes)

  (e.g., prefer customer routes over peer routes over upstream routes)

- Allow customers a choice of route

  (e.g., on customer request do not export prefix to AS x, etc.)

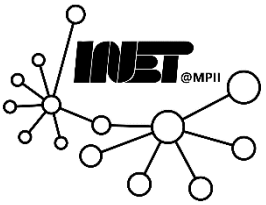
- Enable customer traffic engineering

  (e.g., prepend x times to all peers or to specified AS)

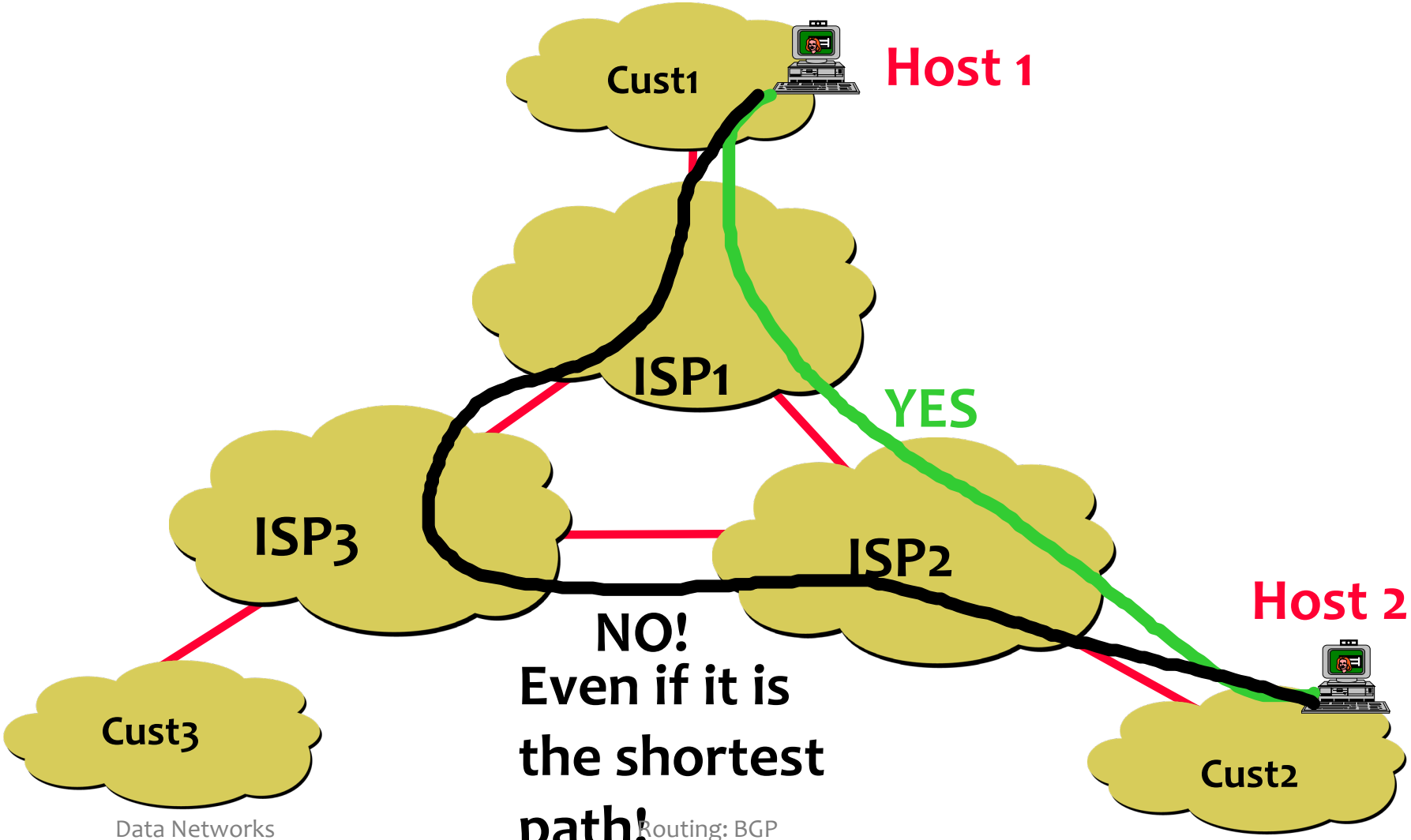
- Enable DDoS defense for customers

  (e.g., blackholing by rewriting the next hop)

- …

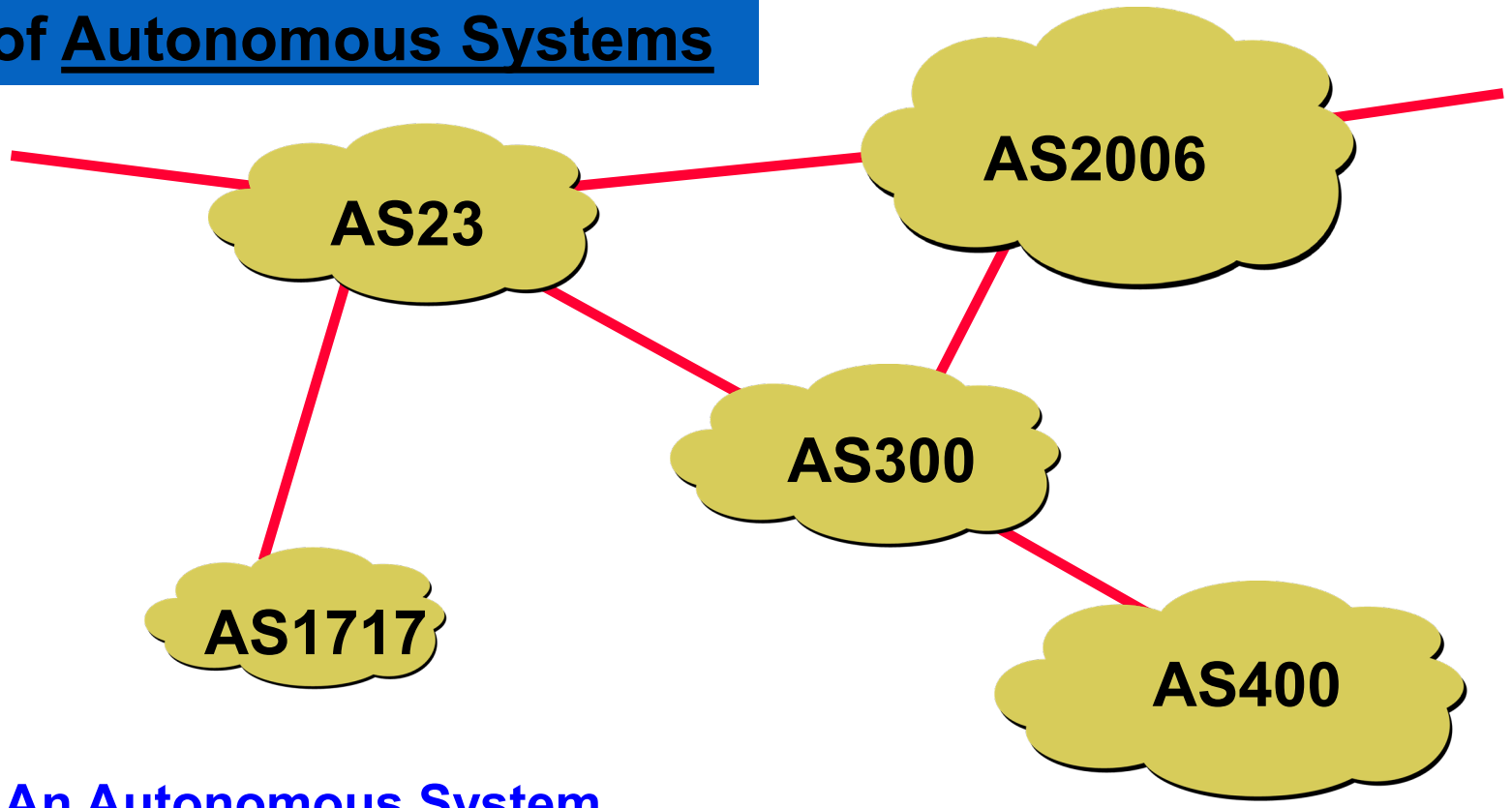


Why policy should win over distance metrics



Current Internet architecture

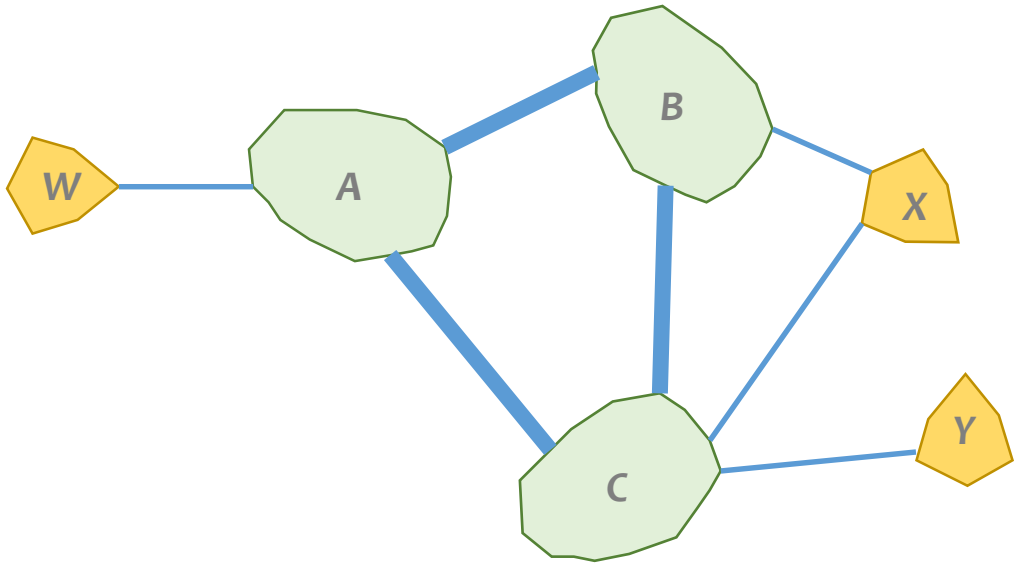
Arbitrary Internetwork of Autonomous Systems

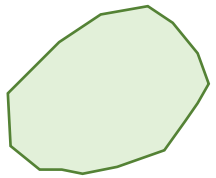


An Autonomous System is a unified administrative domain with a consistent routing policy

A few years ago about 7000 AS numbers are assigned, about 4200 in use

BGP routing policy



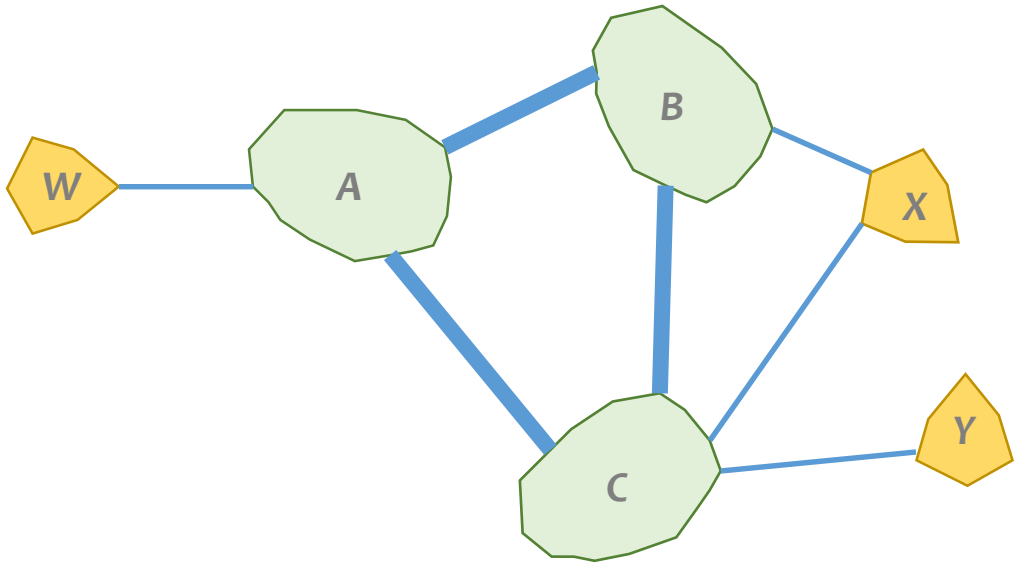


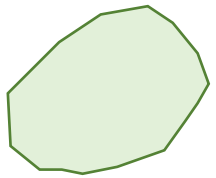
Provider network

Customer network

- ***A***, ***B***, and ***C*** are ***provider*** networks
- ***X***, ***W***, ***Y*** are ***customer*** (of *provider* networks)
- ***X*** is ***dual-homed*** *(i.e., attached to two networks)*
- ***X*** does not want to route from ***B*** via ***X*** to ***C***
  - … so ***X*** *will not* advertise to ***B*** a route to **C**

BGP routing policy



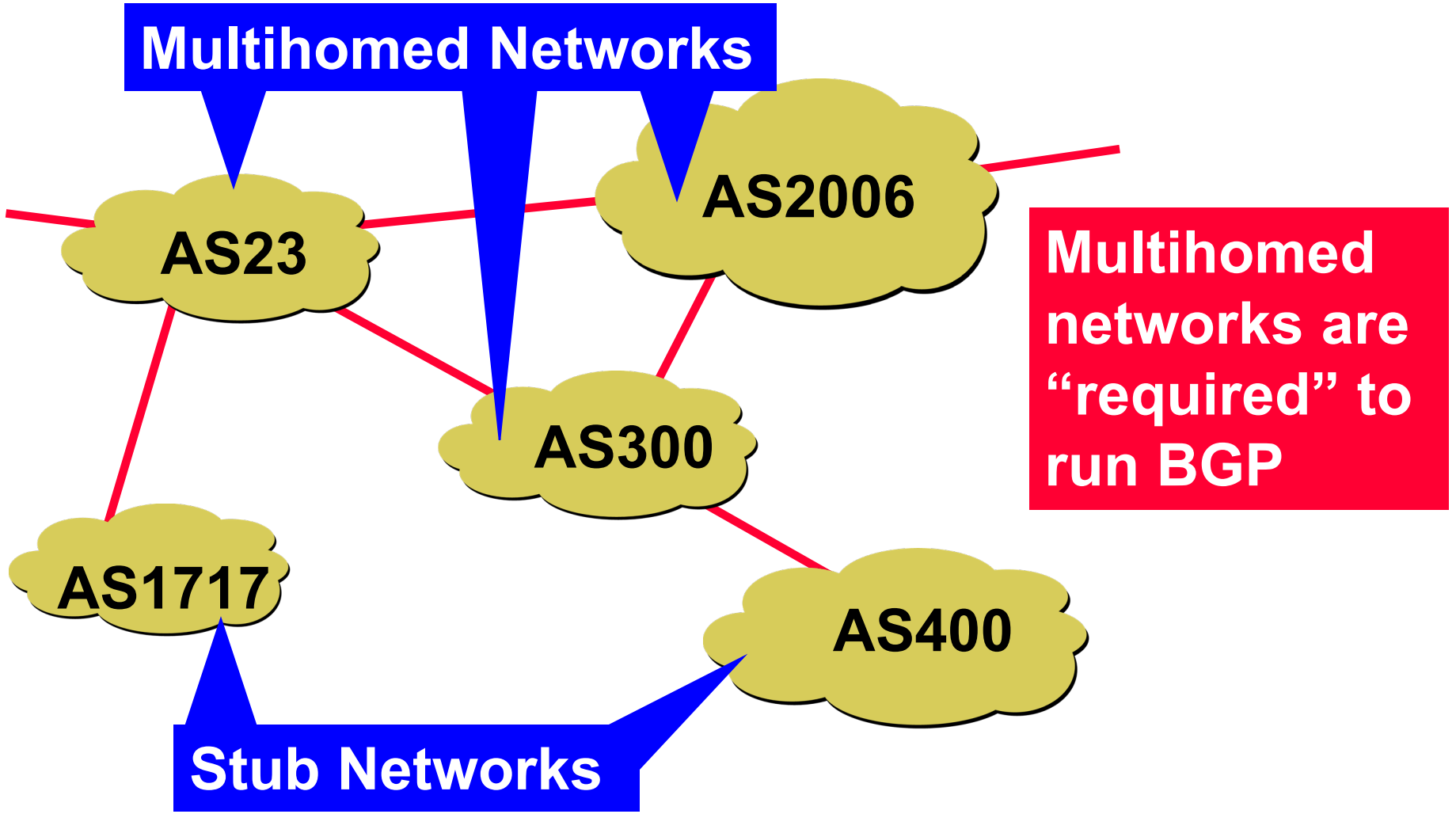


Provider network

Customer network

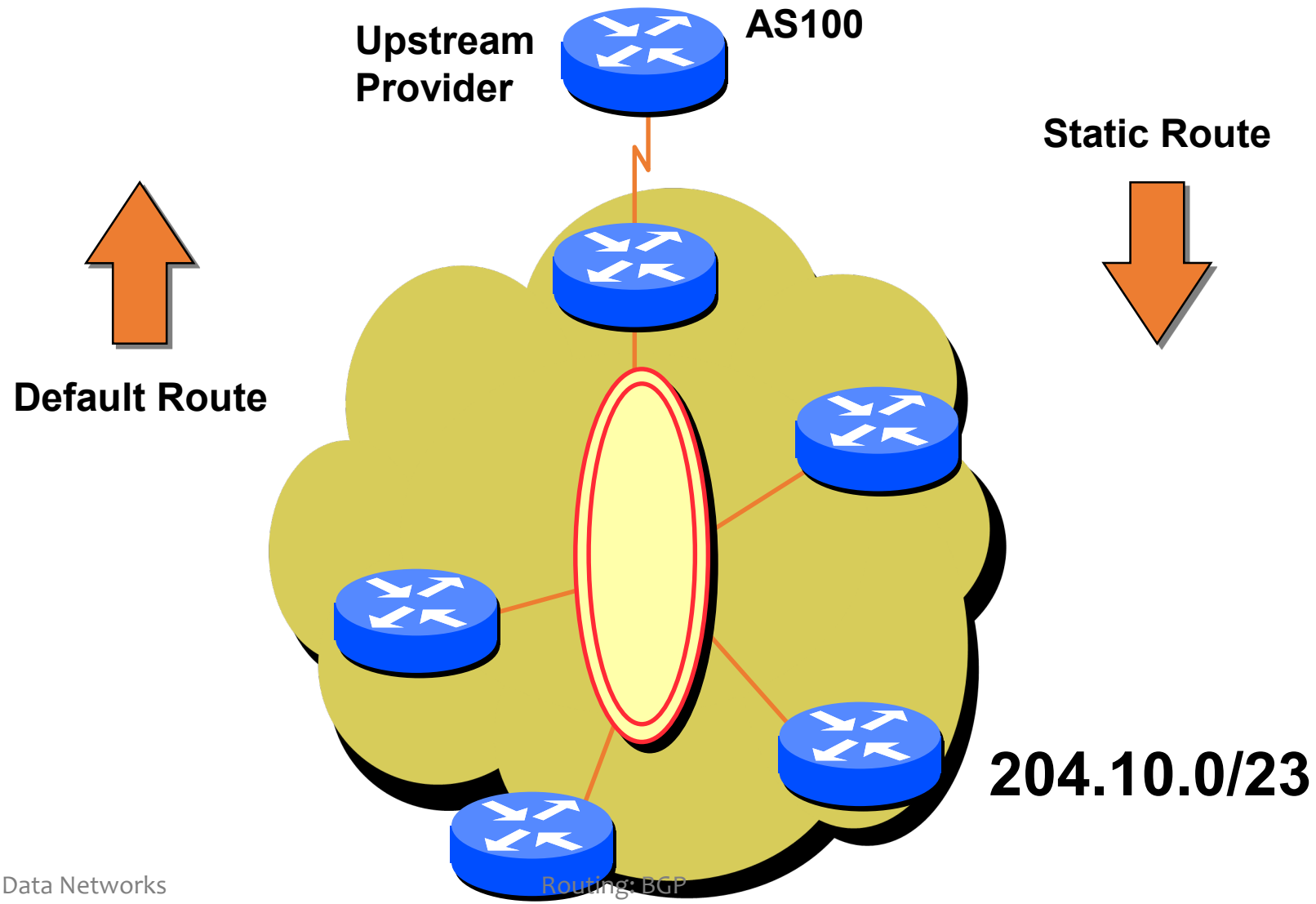
- ***A*** advertises to ***B*** the path ***A-W***; ***B*** advertises to ***X*** the path ***B-A-W***
- *Should B advertise to C the path BAW?*
  - *No way!* ***B*** gets no *"revenue"* for routing ***C-B-A-W*** since neither ***W*** nor ***C*** are ***B***'s *customers*
  - ***B*** wants to *force* ***C*** to route to ***W*** via ***A***
  - ***B*** wants to route *only* to and from its *customers!*

Stub vs. Multihomed networks

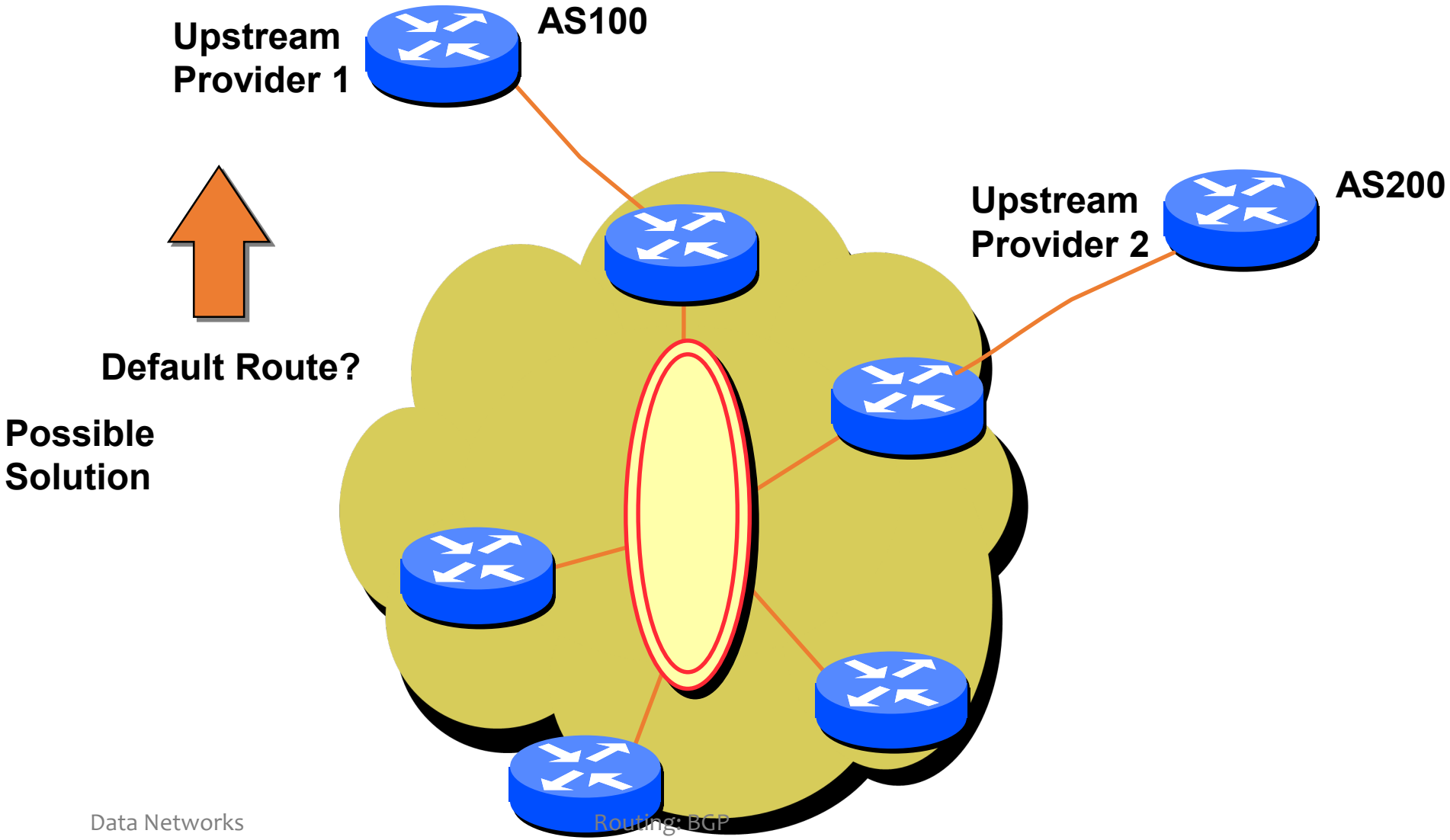


Multihomed networks are "required" to run BGP

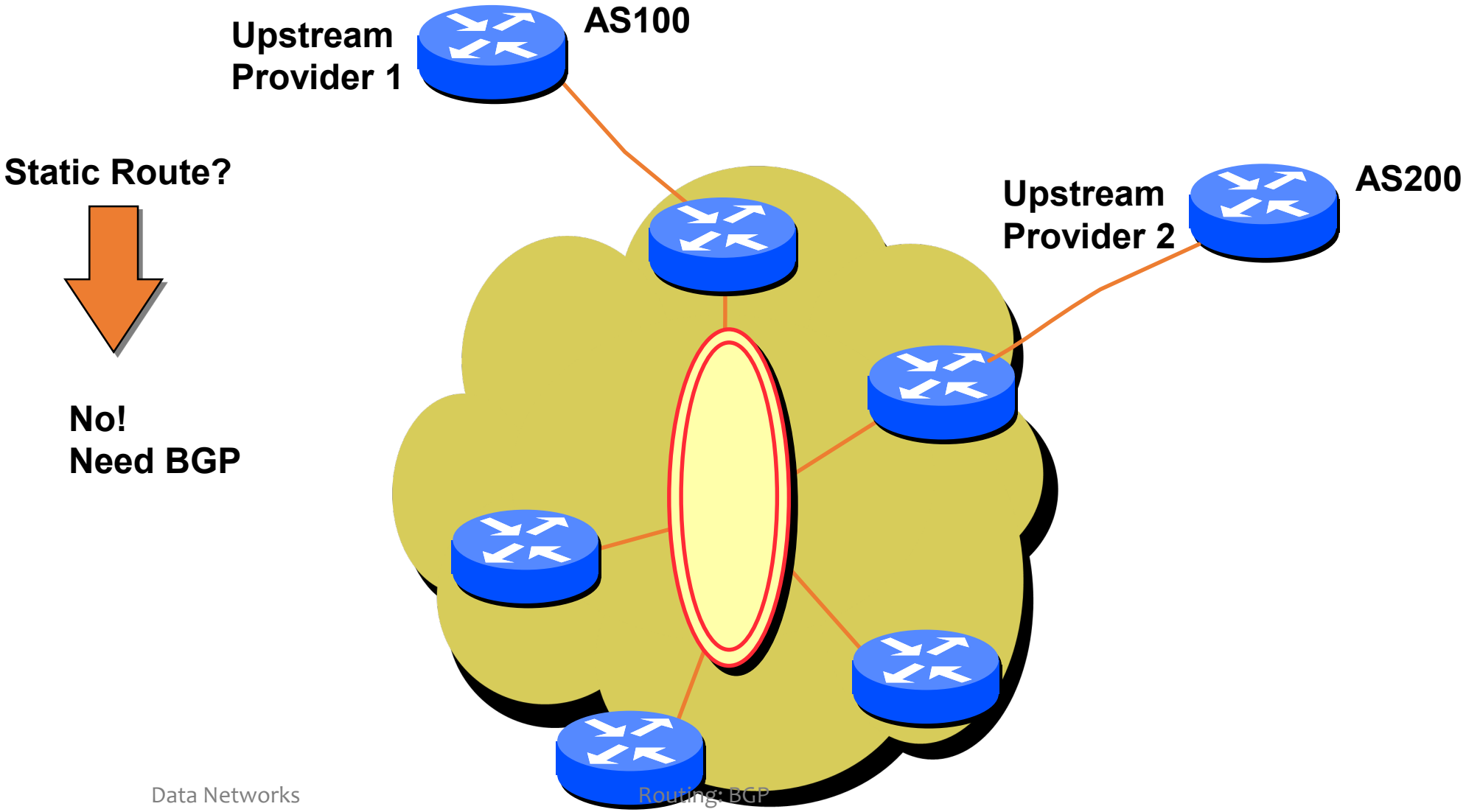
Routing at Stub ASs



Routing at Stub ASs – Multiple providers

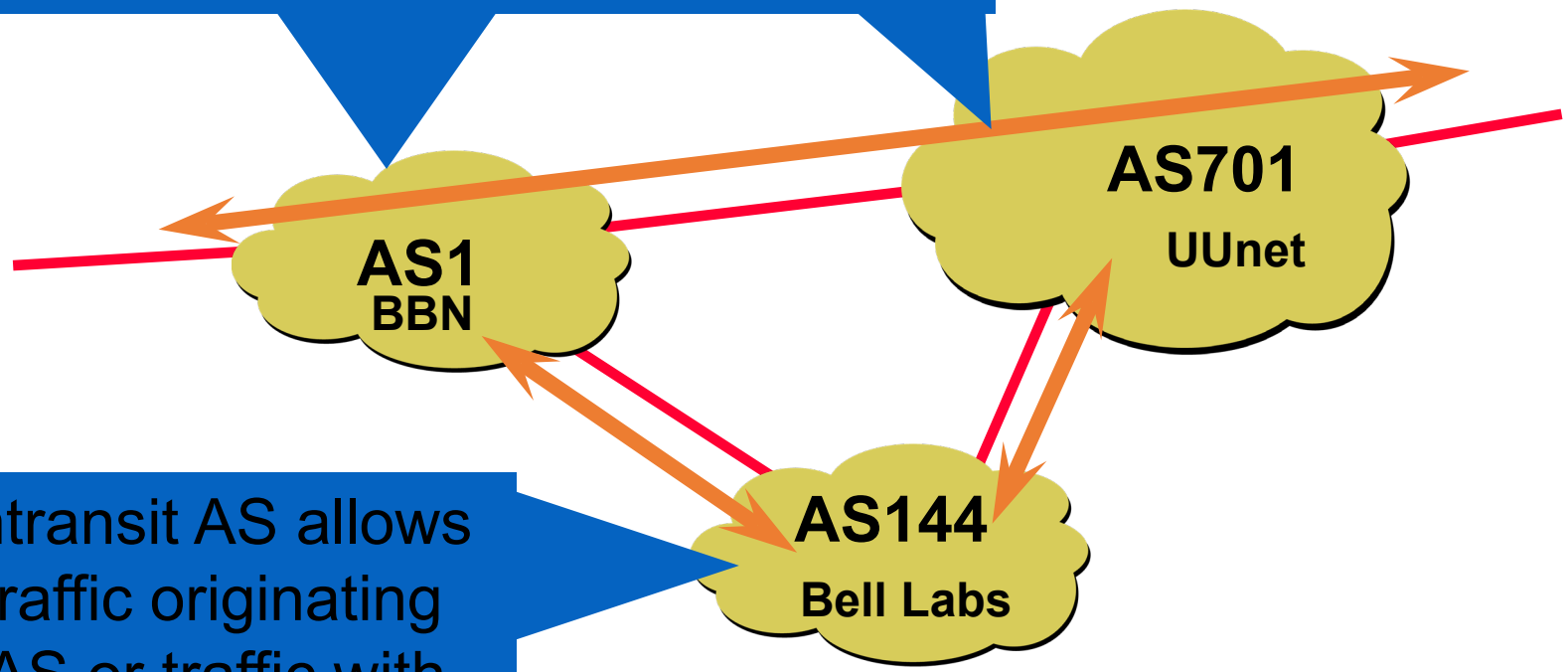


Routing at Stub ASs – Multiple providers



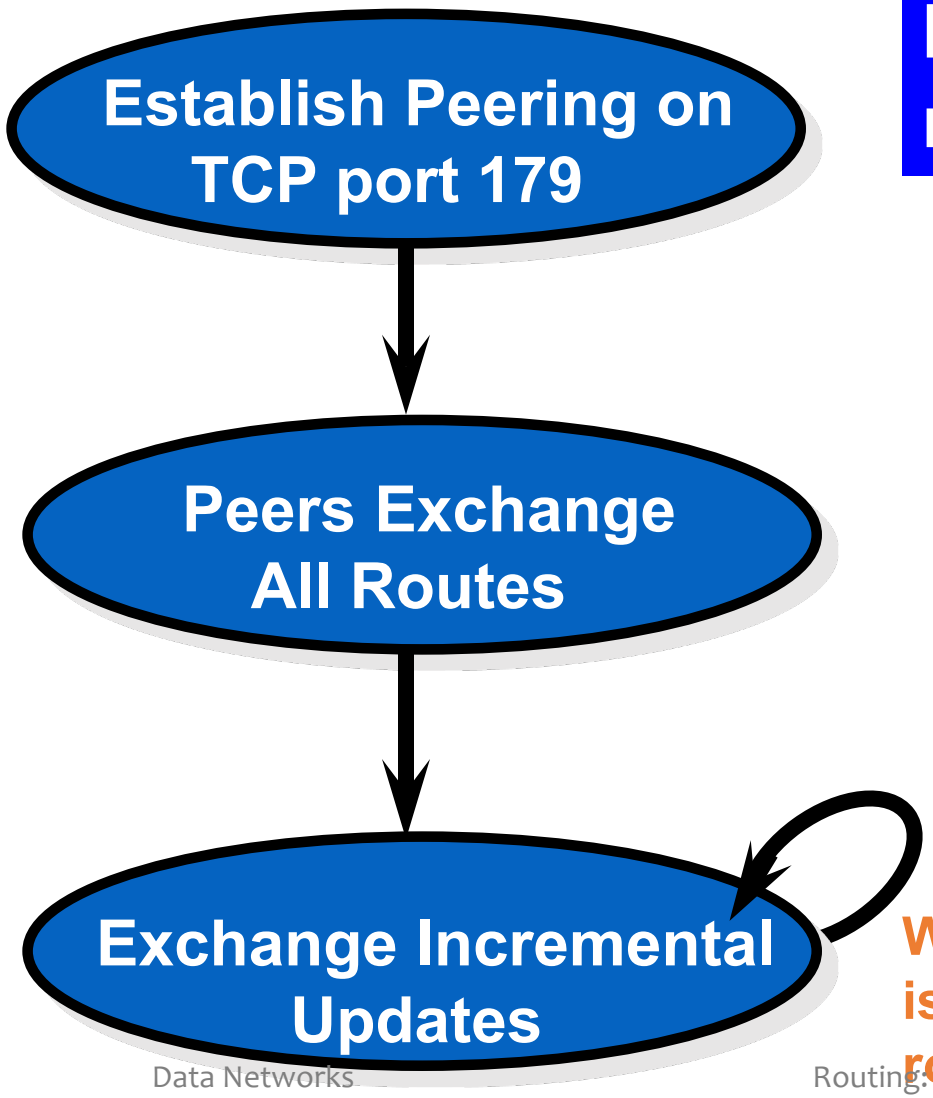
Policy: Transit vs. Nontransit

A transit AS allows traffic with neither source nor destination within AS to flow across the network

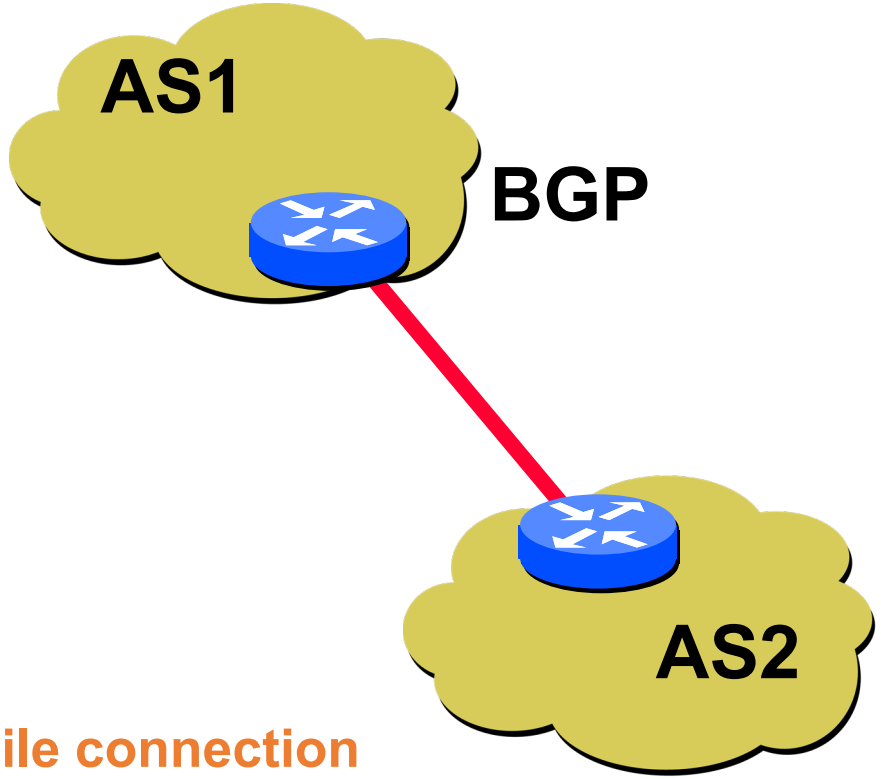


A nontransit AS allows only traffic originating from AS or traffic with destination within AS

BGP operations simplified



BGP Route = network prefix + attributes



While connection is ALIVE exchange route UPDATE messages

BGP messages

Peers exchange ***BGP messages*** using *TCP*

- ***OPEN***:
  - Opens TCP connection to peer
  - Authenticates sender
- ***UPDATE***:
  - Advertises new routes (or withdraws old)
- ***KEEPALIVE***:
  - Keeps conn alive in absence of UPDATES, ACKs OPEN request
- ***NOTIFICATION***:
  - Reports errors in previous message; closes a connection

Process:

- *Initialization*: Open ⟹ Updates for all routes
- *Ongoing*: Updates for changed routes