

Internet Routing Protocols Lecture 03 Inter-domain Routing

Advanced Systems Topics

Lent Term, 2009

**Timothy G. Griffin
Computer Lab
Cambridge UK**

Autonomous Routing Domains

A collection of physical networks glued together using IP, that have a unified administrative routing policy.

- **Campus networks**
- **Corporate networks**
- **ISP Internal networks**
- **...**

Autonomous Systems (ASes)

An autonomous system is an autonomous routing domain that has been assigned an Autonomous System Number (ASN).

... the administration of an AS appears to other ASes to have a single coherent interior routing plan and presents a consistent picture of what networks are reachable through it.

RFC 1930: Guidelines for creation, selection, and registration of an Autonomous System

AS Numbers (ASNs)

ASNs are 16 bit values (soon to be 32 bits)

64512 through 65535 are “private”

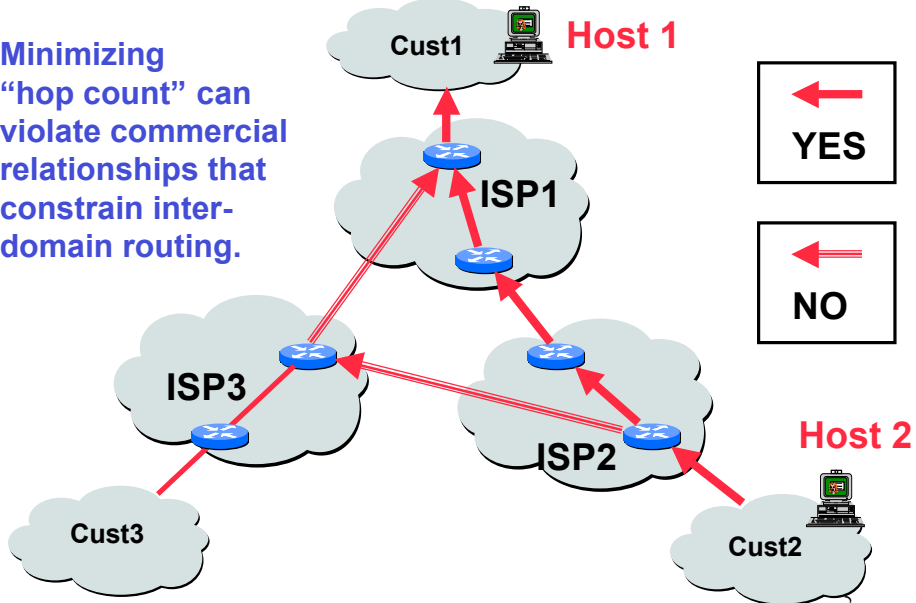
Currently nearly 30,000 in use.

- **JANET: 786**
- **MIT: 3**
- **Harvard: 11**
- **UC San Diego: 7377**
- **AT&T: 7018, 6341, 5074, ...**
- **UUNET: 701, 702, 284, 12199, ...**
- **Sprint: 1239, 1240, 6211, 6242, ...**
- **...**

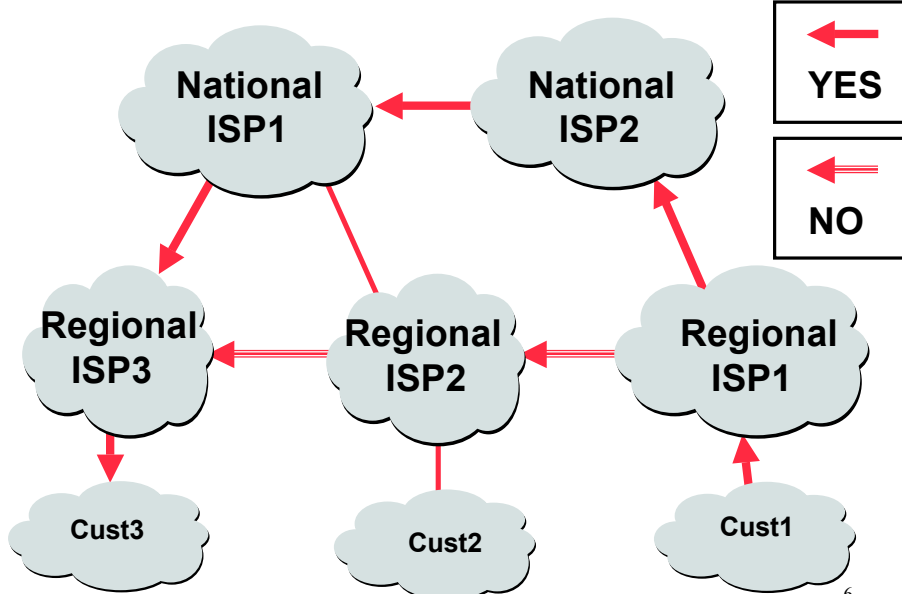
ASNs represent units of routing policy

Policy-Based vs. Distance-Based Routing?

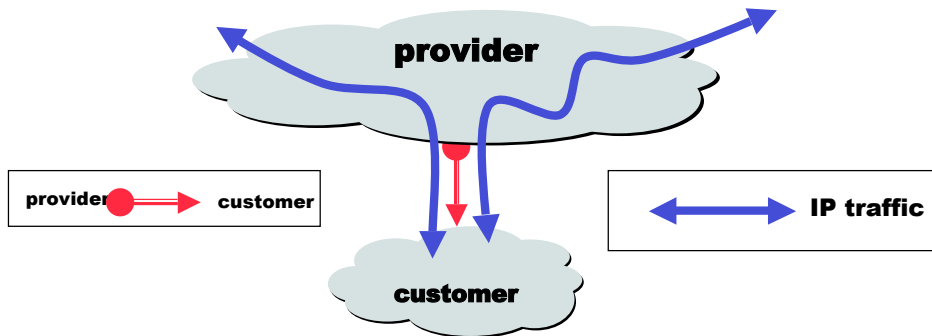
Minimizing "hop count" can violate commercial relationships that constrain inter-domain routing.



Why not minimize "AS hop count"?

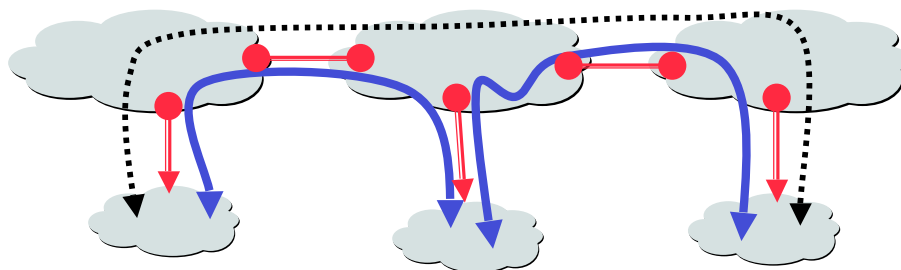


Customers and Providers



Customer pays provider for access to the Internet

The “Peering” Relationship

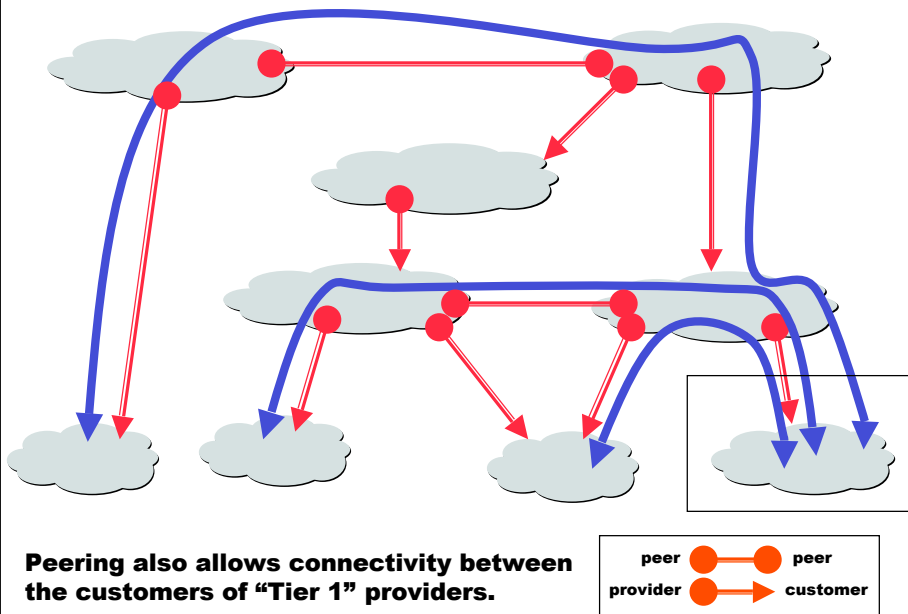


Peers provide transit between their respective customers

Peers do not provide transit between peers

Peers (often) do not exchange \$\$\$

Peering Provides Shortcuts



Peering Wars

Peer

- Reduces upstream transit costs
- Can increase end-to-end performance
- May be the only way to connect your customers to some part of the Internet ("Tier 1")

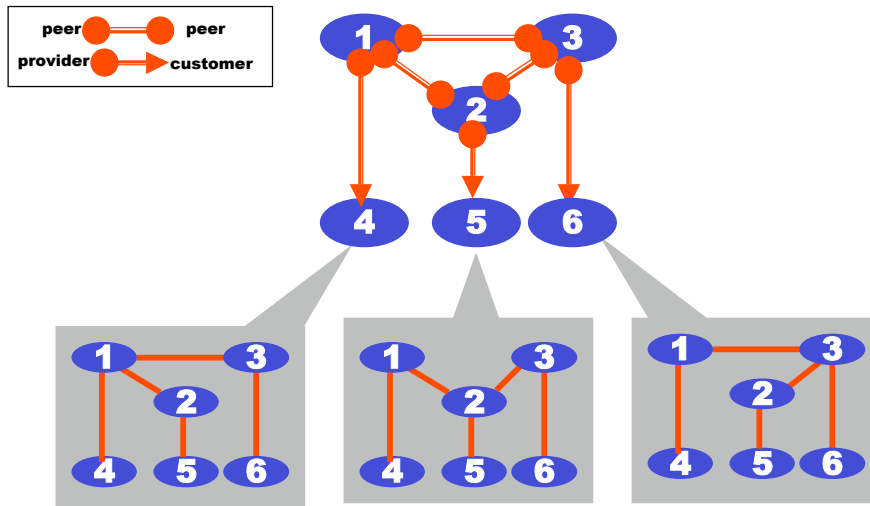
Don't Peer

- You would rather have customers
- Peers are usually your competition
- Peering relationships may require periodic renegotiation

Peering struggles are by far the most contentious issues in the ISP world!

Peering agreements are often confidential.

AS Graphs Depend on Point of View



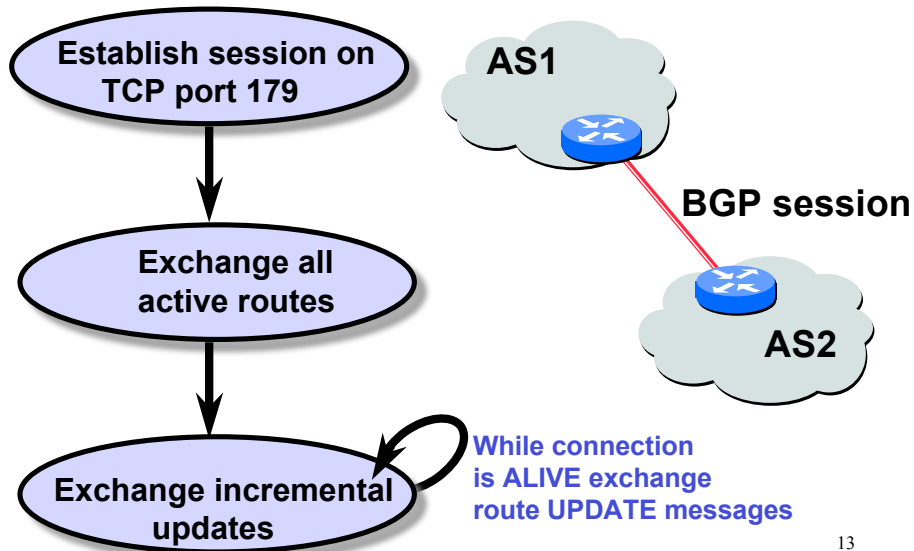
BGP-4

- **BGP = Border Gateway Protocol**
- Is a **Policy-Based** routing protocol
- Is the **de facto EGP** of today's global Internet
- Relatively simple protocol, but configuration is complex and the entire world can see, and be impacted by, your mistakes.

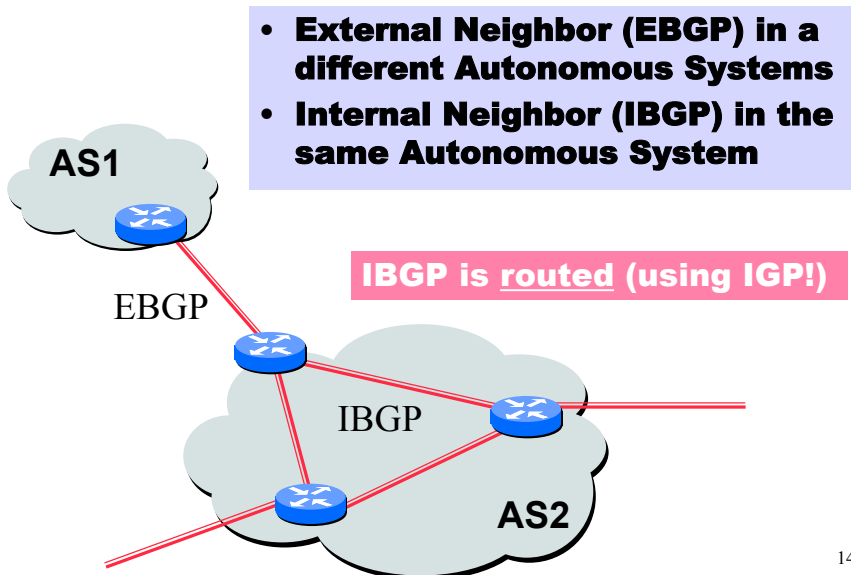
- **1989 : BGP-1 [RFC 1105]**
 - Replacement for EGP (1984, RFC 904)
- **1990 : BGP-2 [RFC 1163]**
- **1991 : BGP-3 [RFC 1267]**
- **1995 : BGP-4 [RFC 1771]**
 - Support for Classless Interdomain Routing (CIDR)
- **2006 : BGP-4 [RFC 4271]**

12

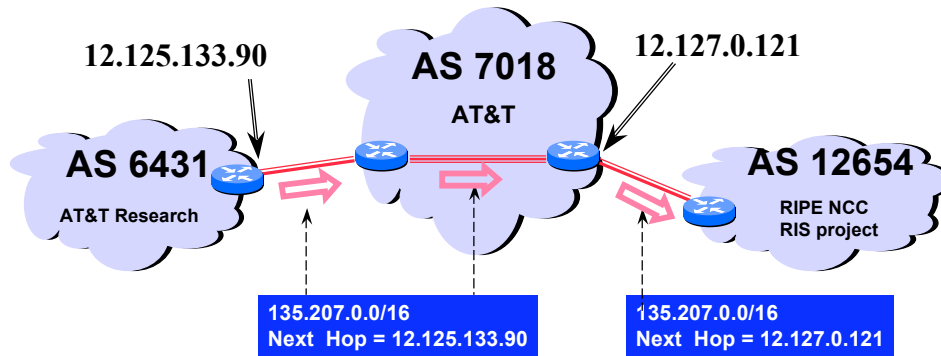
BGP Operations (Simplified)



Two Types of BGP Sessions



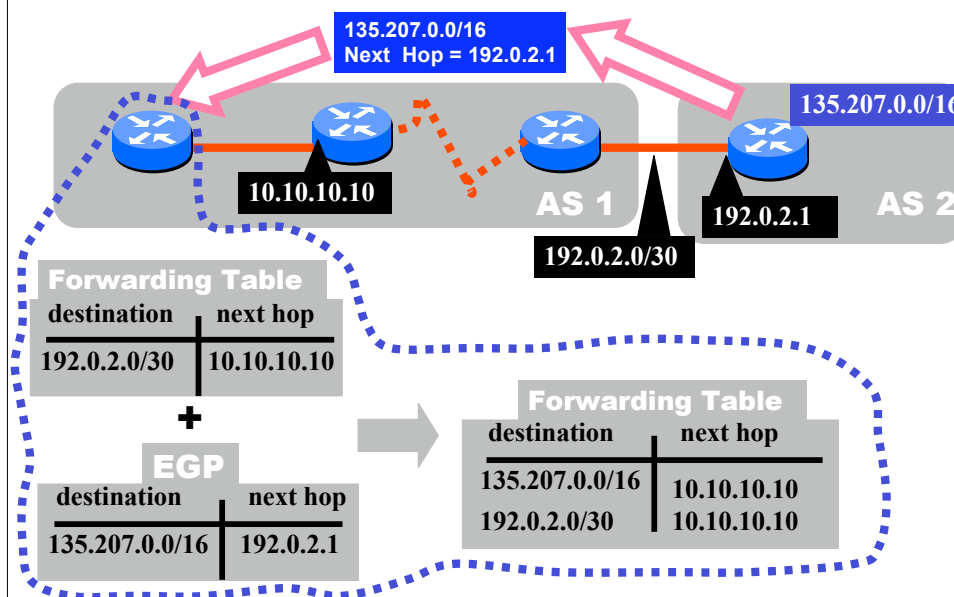
BGP Next Hop Attribute



Every time a route announcement crosses an AS boundary, the Next Hop attribute is changed to the IP address of the border router that announced the route.

15

Join EGP with IGP For Connectivity



Four Types of BGP Messages

- **Open** : Establish a peering session.
- **Keep Alive** : Handshake at regular intervals.
- **Notification** : Shuts down a peering session.
- **Update** : Announcing new routes or withdrawing previously announced routes.

announcement
 =
prefix + attributes values

17

BGP Attributes

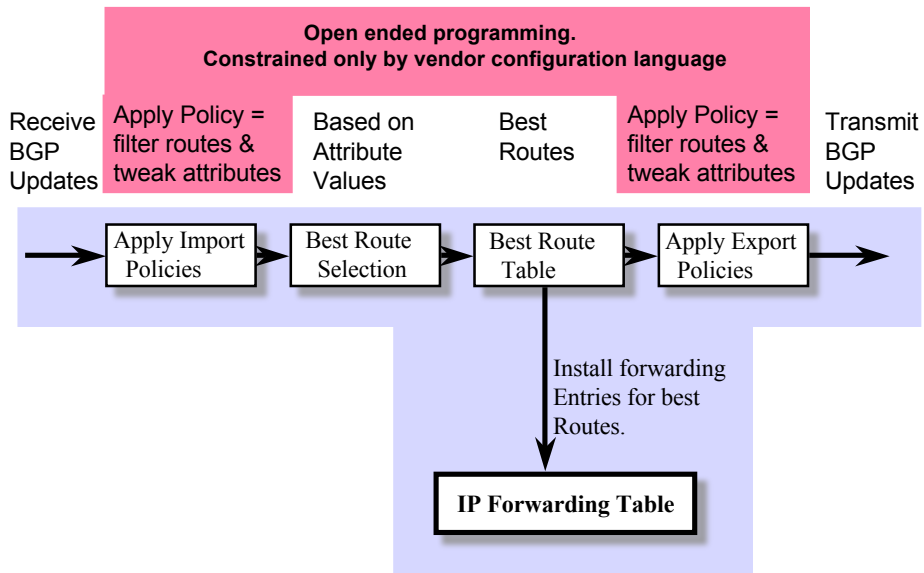
Value	Code	Reference
1	ORIGIN	[RFC1771]
2	AS_PATH	[RFC1771]
3	NEXT_HOP	[RFC1771]
4	MULTI_EXIT_DISC	[RFC1771]
5	LOCAL_PREF	[RFC1771]
6	ATOMIC_AGGREGATE	[RFC1771]
7	AGGREGATOR	[RFC1771]
8	COMMUNITY	[RFC1997]
9	ORIGINATOR_ID	[RFC2796]
10	CLUSTER_LIST	[RFC2796]
11	DPA	[Chen]
12	ADVERTISER	[RFC1863]
13	RCID_PATH / CLUSTER_ID	[RFC1863]
14	MP_REACH_NLRI	[RFC2283]
15	MP_UNREACH_NLRI	[RFC2283]
16	EXTENDED COMMUNITIES	[Rosen]
...		
255	reserved for development	

Most important attributes

From IANA: <http://www.iana.org/assignments/bgp-parameters>

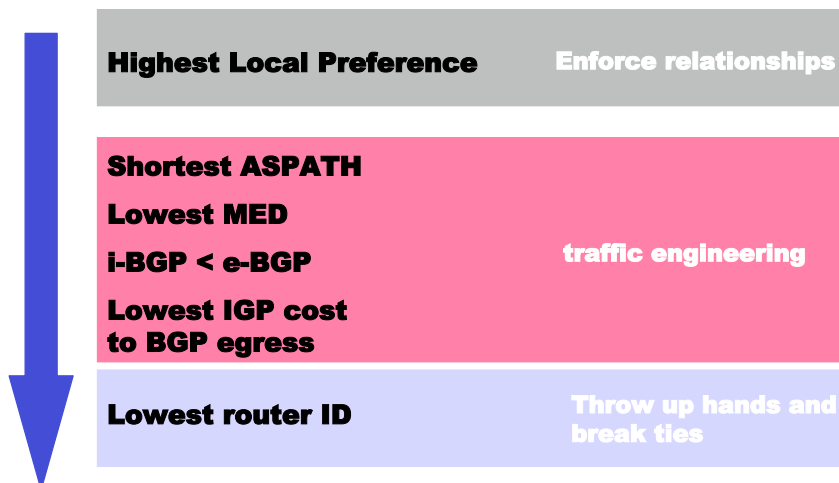
Not all attributes need to be present in every announcement

BGP Route Processing

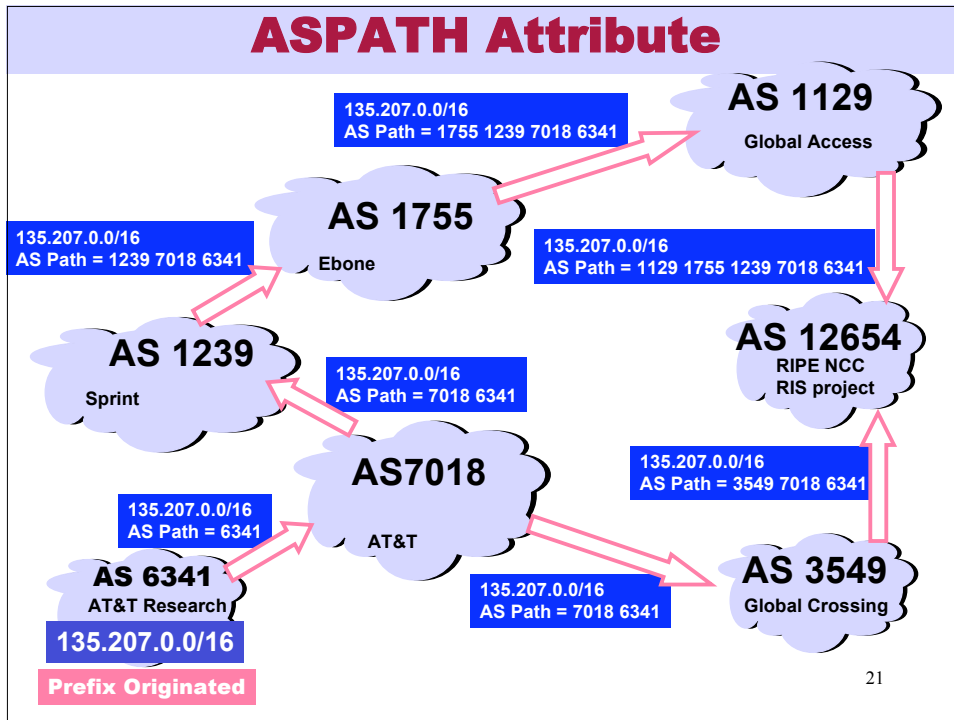


19

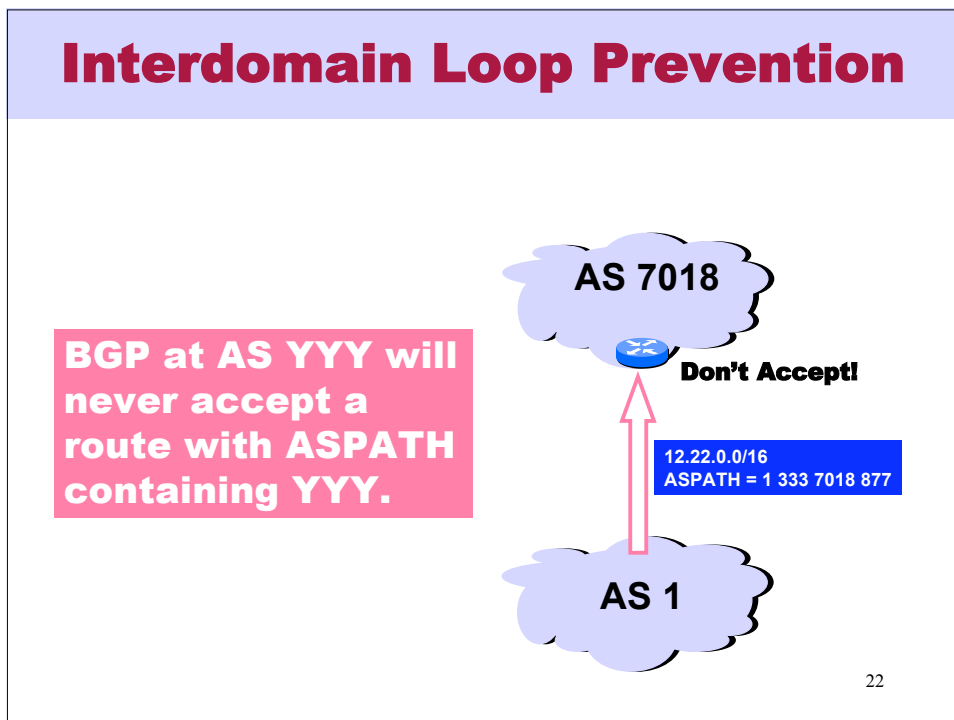
Route Selection Summary



ASPATH Attribute

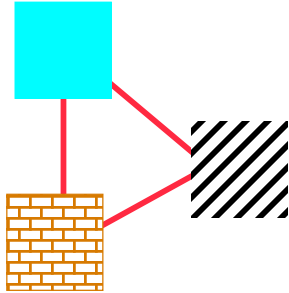


Interdomain Loop Prevention

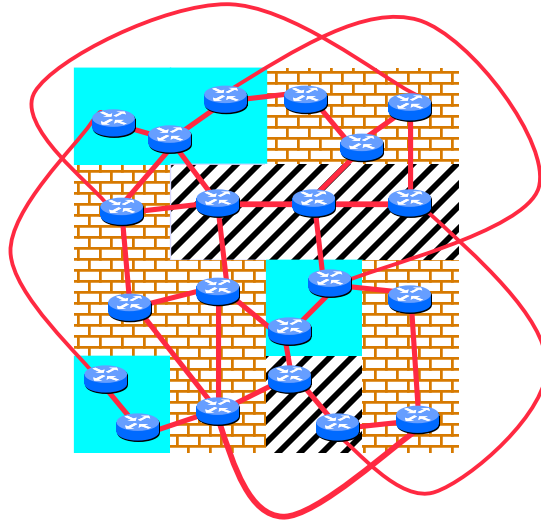


AS Graphs Do Not Show "Topology"!

BGP was designed to throw away information!



The AS graph may look like this.

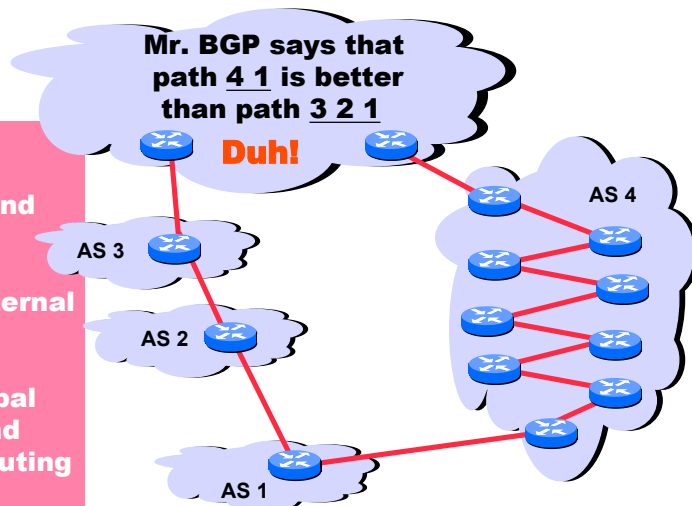


Reality may be closer to this...

Shorter Doesn't Always Mean Shorter

In fairness: could you do this "right" and still scale?

Exporting internal state would dramatically increase global instability and amount of routing state

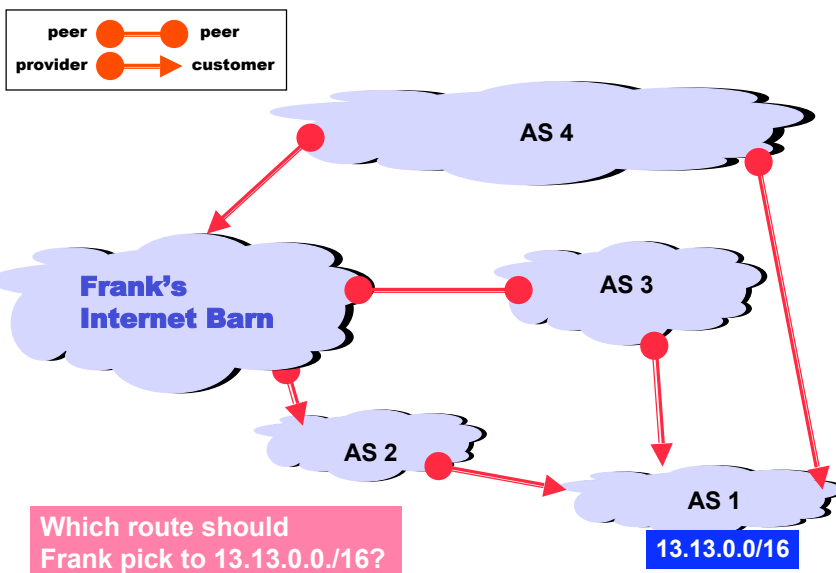


Implementing Customer/Provider and Peer/Peer relationships

Two parts:

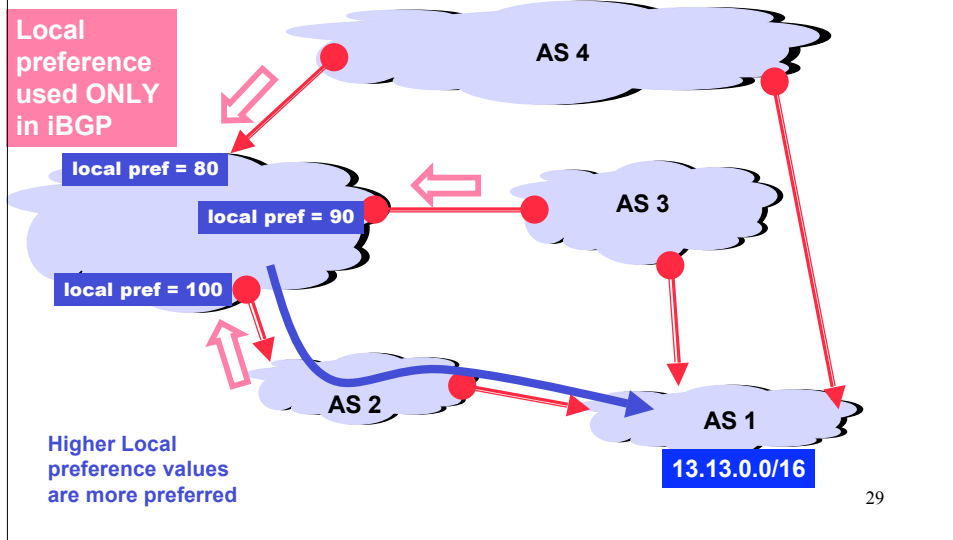
- Enforce transit relationships
 - Export all (best) routes to customers
 - Send only own and customer routes to all others
- Enforce order of route preference
 - provider < peer < customer

So Many Choices



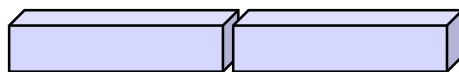
28

LOCAL PREFERENCE



How Can Routes be Classified? BGP Communities!

A community value is 32 bits



By convention, first 16 bits is ASN indicating who is giving it an interpretation

community number

Used for signaling within and between ASes

Very powerful BECAUSE it has no (predefined) meaning

Community Attribute = a list of community values. (So one route can belong to multiple communities)

Reserved communities

no_export = 0xFFFFF01: don't export out of AS
no_advertise 0xFFFFF02: don't pass to BGP neighbors

RFC 1997 (August 1996)