Border Gateway Protocol (BGP) Basics

Guest Talk
Ola Nordström (nalo@cc.gatech.edu)

Overview

- Distance Vector vs. Link State
  - Path Vector
- BGP Protocol
  - Conceptual Model
  - Message Types
  - Attributes
  - I-BGP and E-BGP
- BGP Operations
  - Route Selection
  - Aggregation
  - Community Attributes
  - Peering Relations
  - Routing Policies
  - Business Impacts

BGP Protocol

- BGP-4 RFC 1771 (1995)
  - RFC 1997 BGP Communities Attribute
  - RFC 1998 Use of Community Attribute
  - RFC 2283 Multiprotocol Extensions for BGP-4
  - ... several others
- Inter Domain Routing Standard
- CIDR Support

Distance Vector Problems

- RIP
  - Relies on Hop Count
    - Earlier Version Had Max Hop Count
  - Large Table Exchanges
    - Link and CPU overhead
  - Active Route Reappearance
    - Increased Convergence
  - Slow Convergence In General

- EIGRP
  - Good Convergence Times

Link State Problems

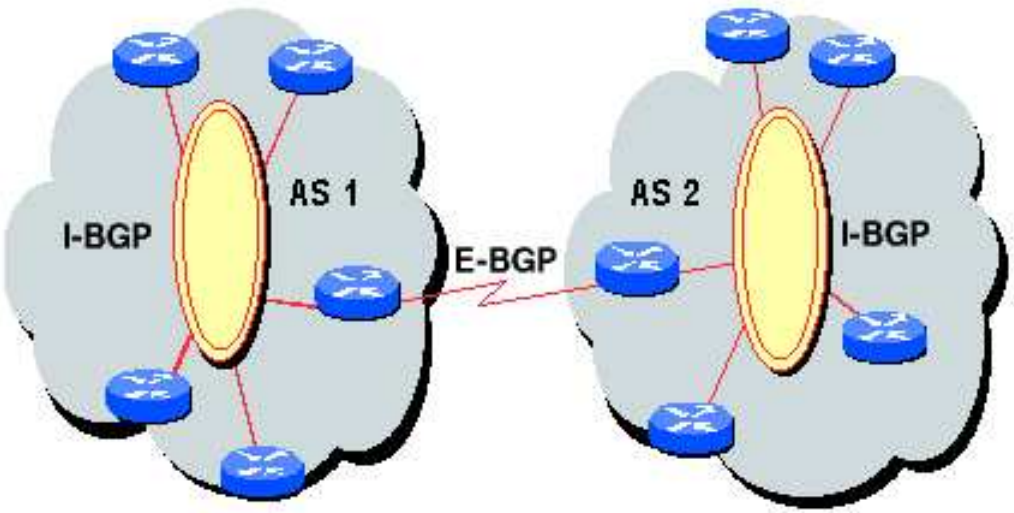
- OSPF
  - No Hop Count
  - Bandwidth Representation
  - Good Convergence via Link State Updates
  - Hierarchy and CIDR support

- Problem
  - Large Networks (10000’s)
  - Route oscillation and Link-State re-transmission too intensive

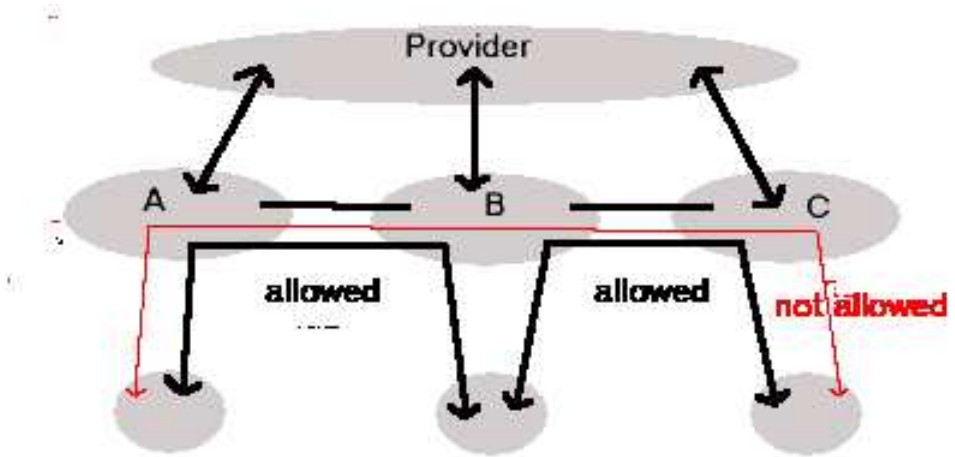
BGP Protocol

- BGP-4
  - Deployment Began in 1993
  - CIDR Support
- Path Vector
  - Sequence of AS Numbers Identify Network Path
- Routers with BGP
  - Speakers, Peers, Neighbours
- Explicit Notification of Errors
- Streamlined Message Exchanges
  - Full exchange on startup
  - Incremental Changes for topology changes
  - Reduced Processing Overhead

Simple Example



Peering Relationships



- Peers A, B, C provide Transit between their customers
  - Why?
- Do not provide Transit between other ASs
  - A to C, C to A traffic must go through provider

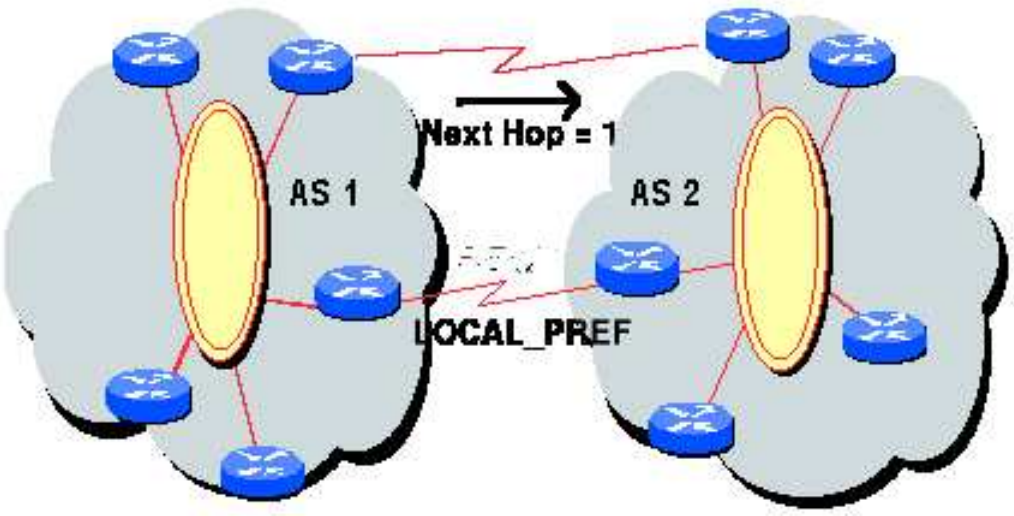
BGP Protocol Specifics

- Run Over TCP
  - Port 179
- 4 Message Types
  - Open
    - Conn. Establishment
  - Update
    - Announce or withdraw paths
    - NLRI
    - Path Attributes
    - Unfeasible Routes
  - Notification
    - Indicate Errors
    - Close session
  - Keepalive
    - Determine Rechability

BGP Update Messages

- NLRI
  - Carries CIDR Information
    - <len,prefix>
- Withdrawn Routes
  - Ex: <18, 199.43.290.0> = 192.213.128.0/18
- Path Attributes
  - 8 bits, 1-16 Used, 16> Reserved for development
  - ORIGIN
  - AS_PATH
  - NEXT_HOP
  - LOCAL_PREF
  - AGGREGATOR
  - COMMUNITY

Local Pref and AS Path

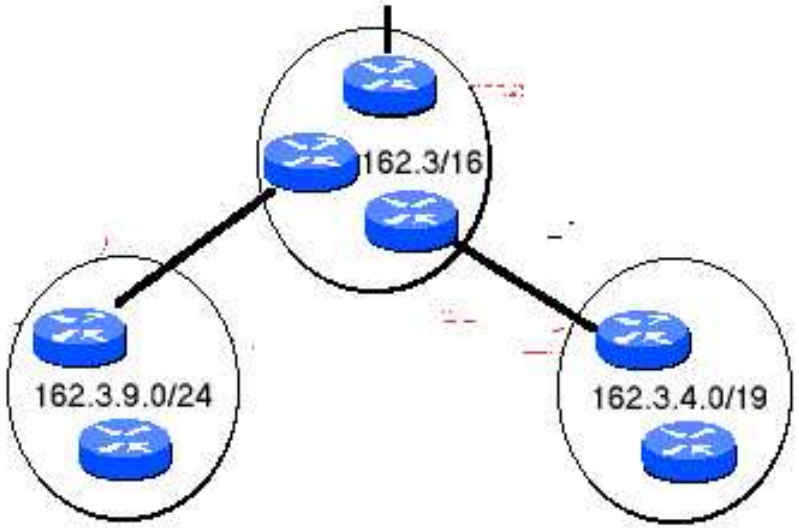


- LOCAL_PREF determines which link is used between AS's
  - Local to AS
- AS1's speaker announces a route to AS2, "next hop" is set
  - next hop = IP addr of announcing border router
  - not changed when announced to I-BGP neighbours

BGP Attributes

- Describe characteristics of a prefix
- Transitive or non-transitive
- Used to Select Routes
- BGP Speaker Picks at Most One Route
  - Maintains secondary routes
- Path Attributes
  - AS sequences
  - Loop detection
  - Allows for routing polices

Aggregation

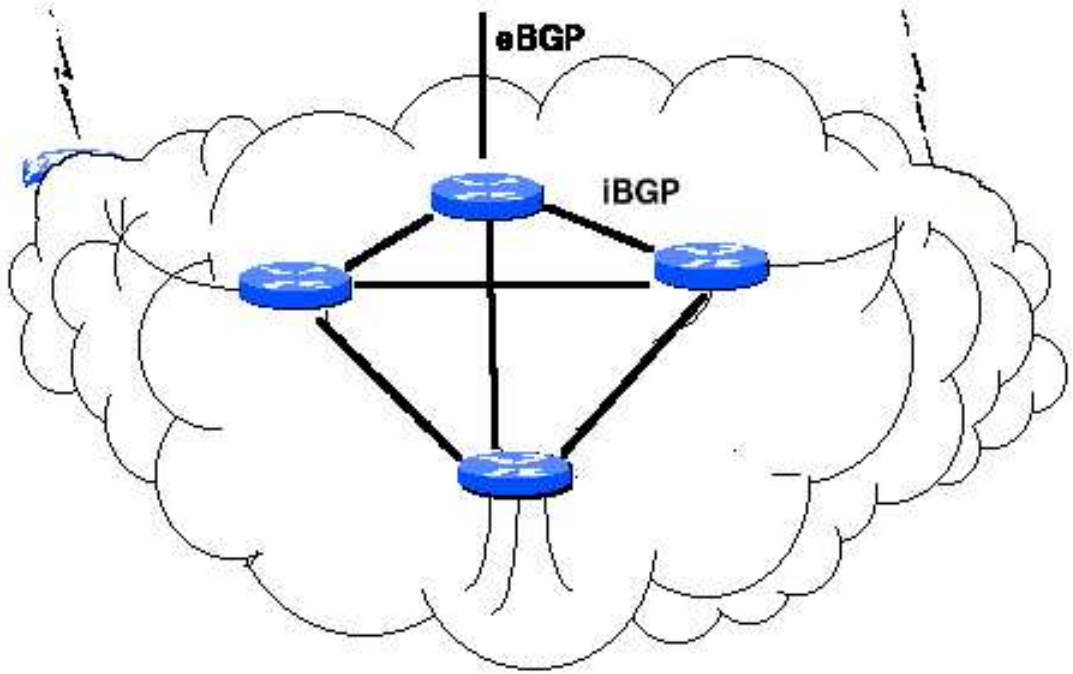


- Part of CIDR
- Allows provider to aggregate customer addresses into blocks
  - eg. 162.3/16
- Requires providers to filter customer announcements
- Reduces routing tables sizes
  - Core routers have a Max Prefix-Limit

I-BGP and E-BGP

- External BGP
  - Inter AS
    - Usually when people say BGP, they mean E-BGP
- Internal BGP
  - Intra AS
  - Same Protocol as E-BGP

I-BGP



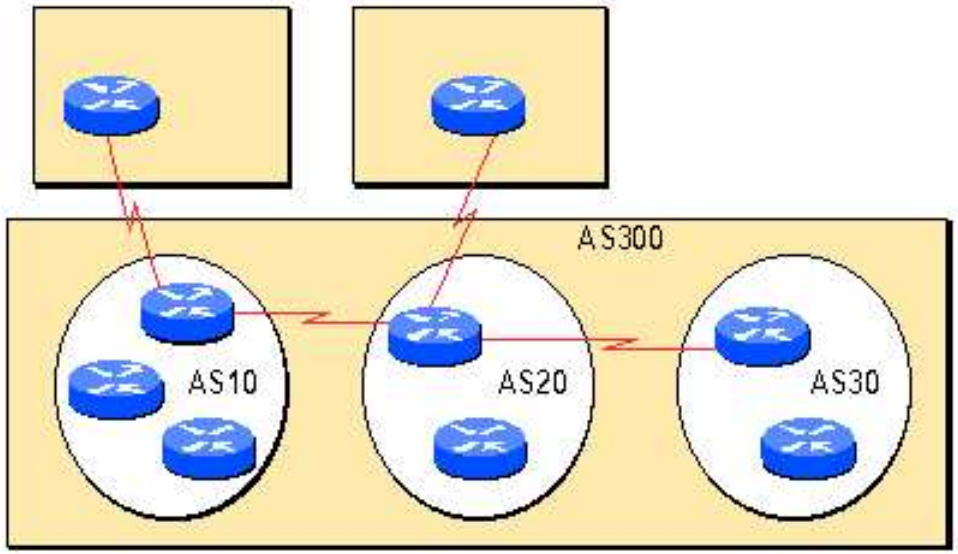
(C) Paul Ferguson

- I-BGP does not re-advertise routing updates.
  - Loop Detection done from AS-PATH attribute
  - AS-PATH only appended to routes crossing AS Boundaries
- Leads to Full Mesh I-BGP Topologies
  - n-1 peering sessions for each I-BGP speaker
  - n(n-1)/2 total session

Reducing the I-BGP Mesh

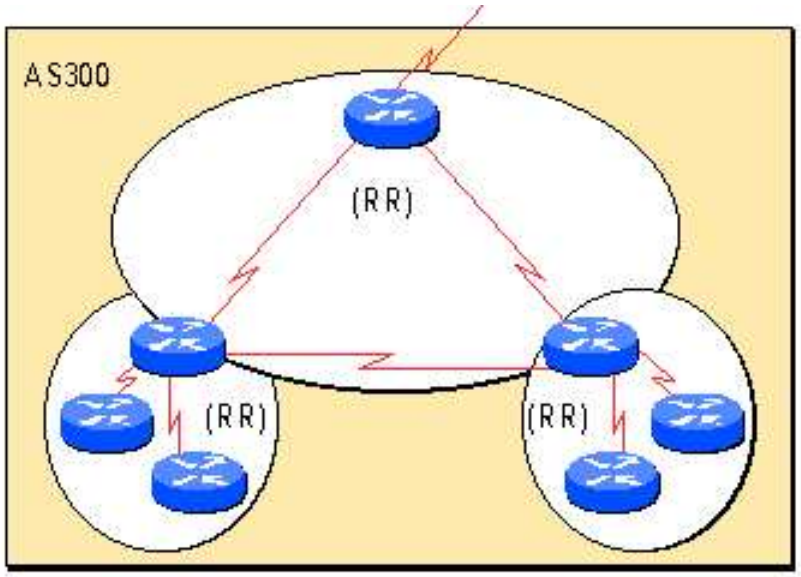
- I-BGP does not scale well
  - 9 routers = 36 I-BGP sessions
- Extensions to the rescue!
  - Confederations
  - Route Reflectors

Confederations



- Divide AS into sub-AS's, still advertise a single AS to peers

Route Reflectors

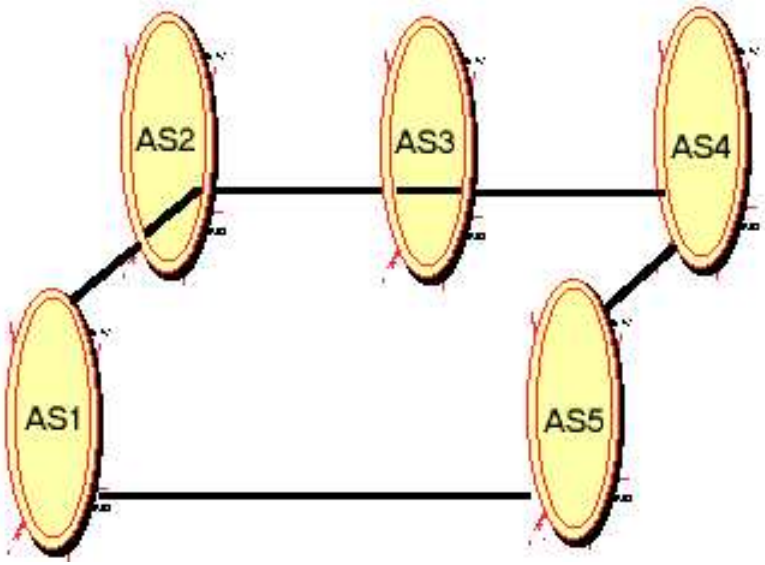


- I-BGP re-advertisement restrictions are relaxed
- Route Reflector "reflects" routes to reflector clients
- Can be multiple Route Reflectors
  - Must be fully meshed
  - Reflector clients may not peer outside their reflector domain

Route Selection

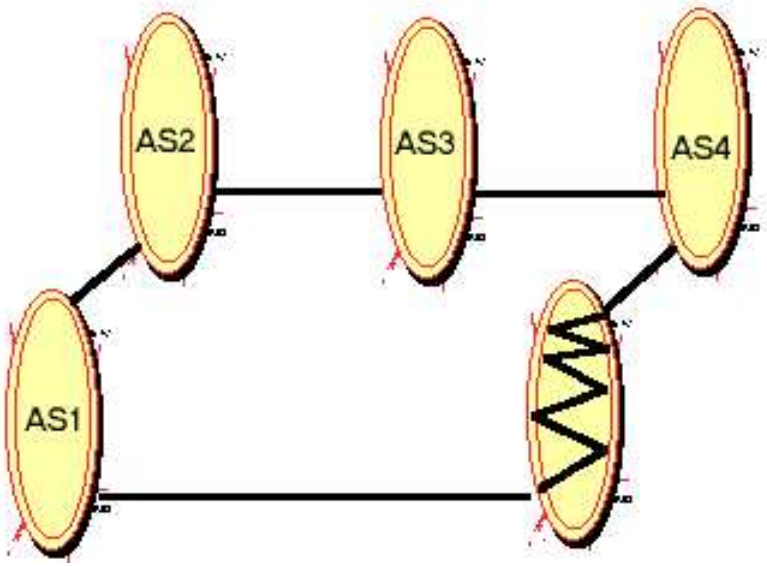
- Governed by Management Policies
  - Built for Business
- Routes always matched to most specific prefix
- Route Selection
  - 1 Highest Local Preference
  - 2 Shortest AS Path
  - 3 Lowest MED
  - 4 Lowest IGP cost to BGP egress
  - 5 Lowest Router ID

2 - Shortest AS Path



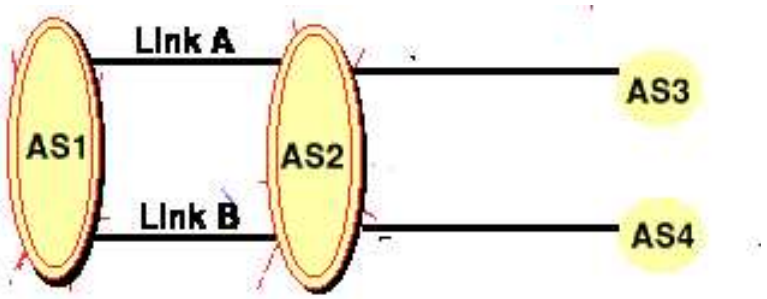
- Path AS1 - AS4
  - Obviously AS1, AS5, AS4

2.5 - Shortest AS Path (not always shortest)



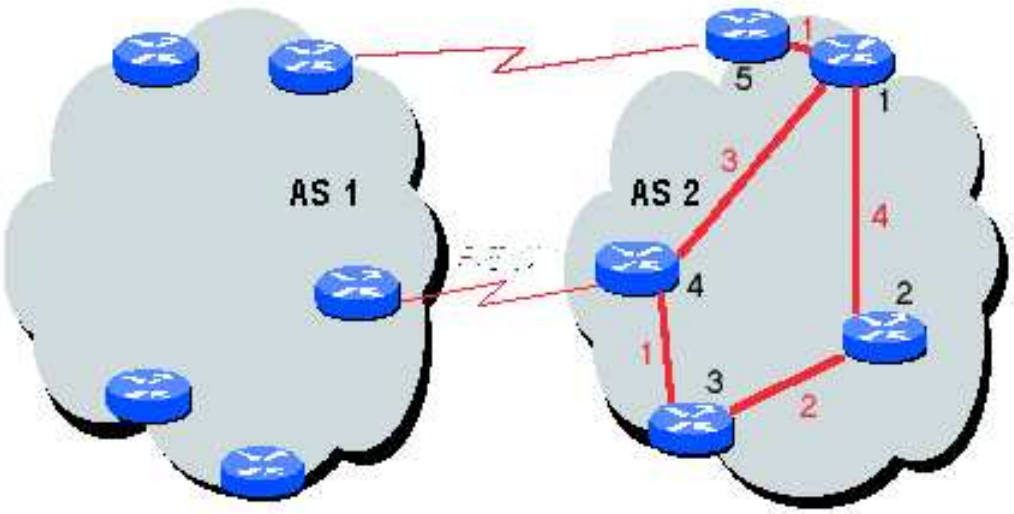
- Not really shortest path, but this is OK
- Do not want ASs' exporting their internal state
  - Increase routing instability

3 - Multi Exit Discriminator (MED)



- Metric Information Expressing Degree of Preference
- Non-transitive
- Set by One AS used by Another
- Ex:
  - AS2 sets MED for AS3's prefix to 10 for Link A
  - AS2 sets MED for AS3's prefix to 50 for Link B
  - AS1 tries to go through Link A to go to AS3
  - If Link A goes down, Link B still available
- Typically only used in Provider-Subscriber Relationships

4 - Lowest IGP Cost



- Local System Selects Route with minimum cost to NEXT HOP
- Ex:
  - AS2: I-BGP router 2 - AS1
    - picks lowest cost path 2, 3, 4

5 - Lowest Router ID

- If all routes learned via I-BGP
  - I-BGP neighbour with lowest BGP identifier used
- If exactly one route learned via E-BGP
  - Select that route
- Multiple Routes learned via E-BGP
  - Select Route learned from E-BGP neighbour with lowest ID

Other BGP Topics

(Covered Next Class)

- Route Flapping & Flap Dampening
- Multi-Homing
  - Effects on Aggregation
- BGP Communities
- Egress and Ingress Filtering
  - BGP Policies

References

- BGP4 Inter-Domain Routing in the Internet.
  - John W. Stewart III, Addison-Wesley 2001.
- Internet routing Architectures. Sam Halabi
  - Cisco Press 2000.
- A BGP Tutorial. T. Griffin (AT&T Labs)
  - 10 IEEE Int. Conf. on Network Protocols, Paris, France 2002.
- RFC 1771 - BGP-4
- RFC 1772 Application of the border Gateway Protocol in the Internet
- Introduction to the Border Gateway Protocol (BGP)
  - NANOG Meeting Feb. 1997. Paul ferguson

Many Examples and Ideas Were Taken from the Griffin and Ferguson presentations.