What is the Problem to be Solved?

• Just configuring the protocol?
• Participating in the Internet
  – and/or running Virtual Private Networks
A Life Cycle

- Initial address space
- Customers’ Space
- Manage Space
- Route
- Renumber
- Justify
- Allocated Space

- Release
"BGP Transmits Policies"

Wrong!
Policies are in Routers

- **Advertising Policies**
  - Outbound to other AS
  - BGP advertisement sources
  - Outbound route filters
  - Route must be in internal routing table

- **Acceptance Policies**
  - Inbound AS filters
  - Inbound route filters
BGP Stack

- BGP
- TCP
- IP
- medium
Policy vs. Protocol Flow

All equivalent from a policy standpoint!
AS Path

1 → 2 → 4 → 5
Network Layer Reachability Information (NLRI)

- Prefix
- Attributes of prefix
  - AS_PATH to reach it
  - Next hop
  - Community
- Address Aggregates
  » Aggregate only
  » Aggregate and all more specific
  » Aggregate and some more specific
  » Holes

- • • •
Stop!
What are you going to Advertise?
Advertising Affects

- The way the world sees you/sends to you
- Binary
  - Routes to which you provide routing
- Quantitative Preferences
  - Multi-Exit Discriminators to your Neighbors
  - AS Path Manipulation to all
Advertise

- Routes Assigned/Allocated to You
- Routes Assigned/Allocated to Customers
- Routes for which you provide Transit
Routes Eligible to Advertise

- Are reachable by your IGP or static routes
- Unless they are black holes
  - Which conceptually are reachable
Do Not Advertise

- Spoofed source addresses
- Your internal addresses
- RFC1918 space
- Known rogues?
  - RBL?
Stop!
What are you going to Accept?
Do Not Accept

- RFC1918 source or destination
- Unexpected sources not assigned/allocated to peers
- Your internal addresses from peers
Accept

- It depends
- Only those routes you will do something about
- Otherwise default
Definitions and Issues
Transit vs. Peer

- Two meanings
  - General BGP Peering
  - Economic
Peering vs. Transit Provider

• **Peer relationship**
  – Mutual benefit customers reach one another
  – No monetary exchange
  – Each advertises customer routes

• **Transit Provider relationship**
  – Customer pays for service
  – Full routes available to customer
Closest Exit Routing

- Paths are not optimized end-to-end
- Paths are optimized for each AS
Asymmetrical Routing

- No guarantee that traffic leaving your AS at one point
- Will return at the same point
- Remember
  - Each AS in both directions makes decisions on its information
Operational Relationships 1
Addresses and Delegation

Address authority -> Reverse DNS

Address delegation -> Prefixes

Prefixes -> DNS

DNS -> Hosts
Autonomous System

- Basis of exterior routing
- AS originate routes for some prefixes they want to be visible
- AS advertise routes to one another
  - Advertisement may not contain all addresses
  - Not all advertisements need to be accepted
Current AS Definition

RFC 1930

- Connected group of IP CIDR blocks
- Run by one or more network operators
- Single routing policy
  - announced to the general Internet
  - announced with BGP-4
AS Number

- 16 bit number
- Numbers assigned by registries
- Private ASNs
  - 64512 through 65535
When is an AS Not Needed?

- Enterprise connects to Internet only via ISP
- Enterprise uses provider AS if it needs one at all
When is an AS Possibly Needed?

- Enterprise connects to Internet via multiple ISP
- AS may be needed for load-sharing or keepalive
When is an AS Needed?

Supplier AS 4
Research Partner AS 5
Internal Researchers
Commercial ISP AS 2
Academic network AS 3
Routes/Aggregates

Address Authority

AS 1

Organization 1
Prefix Block A

Organization 1
Prefix Block B

Organization 3
Prefix Block C

Prefix Block A.1

Prefix Block A.2

Internal Only

Aggregation
Relationships to Adjacent AS

AS 1

AS 2
Announces A, C

AS 3
Announces A

AS 4
Announces A
(B in future)

Prefix Block A
Prefix Block B
Prefix Block C
Joining the Club

- Get/join an AS
- Register AS in appropriate registries
- Peer AS to other cooperating AS
Establishing an AS (1)
Obtain routable address space

- Apply to registry
  - RIPE, APNIC, ARIN
  - If immediate need for /19 or /20*
- Obtain addresses from upstream ISP
  - If /19 or /20 cannot be justified
- Registry needs
  - Network design
  - Justification for address space
Establishing an AS (2)

AS Number Request

- In request to AS number registry
  - Administrative and technical contacts
  - Autonomous system name
  - Router description
  - Deployment schedule
  - Networks (by name) connected by the router(s)
  - Internet addresses of the routers
Establishing an AS (3)
Registering in Routing Registry

- Minimum requirements
  - Maintainer object
  - AS object
  - Route object (s)
Establishing an AS (4)
Operational deployment

- Build configuration
  - Policy implementation
  - Ingress/egress filtering
- Establish security procedures
- Start BGP connections
Registry Issues

- Not all providers use registries
- If multihomed to any that does use registry...it’s needed
- **Procedure** is useful as a checklist that everything is defined
Registry-Based Tools

- Web-available Servers
- Routing Arbiter Tool Set
- Other tools
Routing Registry Objects

- Basic
  - AS
  - Route
  - Maintainer

- Additional
  - Inter-AS Network
  - Community
  - Router
Refinements

Route
- Route set
- Route macro

AS
- AS set
- AS macro

Maintainer
- Person
- Role
Operational Relationships 3:
Registries, Domains, etc.
Turning it On
BGP Configuration Overview

- Plans and policies first!
- Define system of BGP speakers
- Specific BGP speaker configuration
  - Identifier
  - BGP process
  - Neighbors
  - NLRI to advertise
  - Filters and other policy mechanisms

*Cisco commands used as examples*
Consider

- Register in registry
- Use rtconfig, etc. to generate configuration
Router ID and loopback interface

interface loopback 0
ip address 192.168.0.1 255.255.255.0
**router bgp command**

`router bgp AS-number`

- The parameter is a true AS number
- Subcommands will be needed -- minimally,  
  - network  
  - neighbor
network subcommand for BGP

```
network ip-address [mask]
```

- Establishes eligibility of address ranges to be advertised
- But does not actually cause anything to be advertised
  - Different than Cisco IGP network statements
Basic `neighbor` subcommand for eBGP

```
neighbor ip-address
    remote-as as-number
```

- Identifies peer in external AS
- For eBGP, `as-number` is different from our own
- Will not advertise anything unless network statements are defined
Bilateral Exchange
Refining the Configuration

Single and Multiple Links
to a Single Provider
The BGP Tunnel

\texttt{ebgp-multihop} needed when neighbor is not on same subnet
Load Balancing 1: IP Level to Single Provider
Basic IBGP
Basic neighbor subcommand (iBGP)

```plaintext
neighbor ip-address remote-as as-number
```

- Identifies peer in our AS
- For iBGP, `as-number` is our own
- Will not advertise anything unless `network` statements are defined
- Not needed in single-router configurations
Another Non-BGP Alternative
OSPF Routing Domain

ISP 1

D1-A0 ASBR1

D1-A0 ASBR2

Default Route (0.0.0.0/0)
Metric Type 1
Equal Metrics

Static routes
Multiple OSPF Defaults

ISP 1 POP

Default Route (0.0.0.0/0)
Metric Type 2
Higher Metric to ISP 2 (Backup)

ISP 2 POP

Static routes

D1-A0 ASBR1
D1-A0 ASBR2

Default Route (0.0.0.0/0)
Metric Type 2
Higher Metric to ISP 2 (Backup)
Blackhole Route

- Establish static route to your block(s)
  \texttt{ip route 1.2.3.4 255.255.240.0 null0}
- Redistribute/import into BGP
- Suppress more-specific prefix advertising
Effects of Blackholing

- No route flapping outside your AS
  - If your internal routes go up or down
- Incoming traffic for specific routes that are down
  - Doesn’t match any internal route
  - Automatically discarded without concerning anyone else
BGP Path Selection
Next Hop Access

Advertised route via R1

Advertised route via R2
General commands

- Pattern matching
- Action/set
Scope: MED vs. Local Preference vs. Weight
Administrative Weight

Advertised route via R1

Advertised route via R2

Rules in this router set R1 weight to 100, R2 weight to 500
Weight example for load sharing

Default local preference 200

Primary ISP

Default local preference 500
All routes \(^{\text{AS\_Backup}}\) +
local preference 100

Backup ISP
Tiebreaker for Equal Weight:
Local Preference

Advertised route via R1, local preference 100

Advertised route via R2, local preference 500
Local Preference example for load sharing

Default local preference 200

Default local preference 500
All routes ^ AS_Backup +
local preference 100
Mutual Backup & Private Peering

AS65534 --> AS65533
AS65530 --> AS65535

AS65533 --> AS65535
Prefer locally originated routes

Advertised route via R1

Locally defined via R2
Shortest AS Path
AS Path Prepending

- Applies to routes you advertise
- Makes them less attractive to others
- Increases AS_PATH length
  - your AS put in the path twice
Limitations of Prepending
External Paths Preferred

Route Learned from iBGP

Route Learned from eBGP

R1

R2
Lowest MED

MED=100

MED=500

Remote AS

R2

R1
Closest Neighbor

IGP metric to R1=500

IGP metric to R1=100
Lowest BGP router ID

R1 1.1.1.1

R2 2.2.2.2
Administrative Distance

- Preference factors for installing routes in main routing table
  - Cisco calls it administrative distance
- Will affect which sources of routing are installed
- Since we advertise only internally reachable routes
  - This affects those advertised
Permit/Deny Filtering

- Routes can be blocked or permitted to enter BGP
- Criteria include:
  - Advertising address
  - Route address being advertised
  - AS_PATH to reach route
    » Can use UNIX regular expressions
An ISP Topology

POP1
Dial /25
8 x /28 /24

POP2
Dial /25
8 x /28 /28

POP3
Dial /25
8 x /28 /28

POP4
Dial /25
8 x /28 /24

Core 1
Core 2
An ISP Address Plan

/19*

128.0.0.0/20

<table>
<thead>
<tr>
<th>/23</th>
<th>/23</th>
<th>/23 (4x25)</th>
<th>/23</th>
</tr>
</thead>
<tbody>
<tr>
<td>POP Dialups Block A</td>
<td>Internal Block B</td>
<td>Customers Block C</td>
<td>Customers Block D</td>
</tr>
<tr>
<td>/25</td>
<td>/25</td>
<td>/25</td>
<td>/25</td>
</tr>
<tr>
<td>/25</td>
<td>/25</td>
<td>/30</td>
<td>/25</td>
</tr>
<tr>
<td>/30</td>
<td>/25</td>
<td>/25</td>
<td>/24</td>
</tr>
<tr>
<td>/28</td>
<td>/24</td>
<td>/24</td>
<td>/24</td>
</tr>
</tbody>
</table>
Aggregation is better than Aggravation

- Blackhole routes for your blocks
  - Avoid more-specifics
  - Use NO-EXPORT when controlling load to upstream
- Encourage customers to aggregate
  - Proxy aggregation hard to administer
- Understand which blocks you can advertise
  - And do ingress/egress filtering
### Origination vs. Advertising

#### AS 65000

<table>
<thead>
<tr>
<th>128.0.0.0/20</th>
</tr>
</thead>
<tbody>
<tr>
<td>/23 POP Dialups</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

- **AS65000**
  - 128.0.0.0/19

- **AS64444**
  - 192.0.0.0/16
  - an AS65000 Customer

---

<table>
<thead>
<tr>
<th>192.0.0.0/16</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS64444</td>
</tr>
</tbody>
</table>
Aggregating your Own Traffic

**AS65000**

128.0.0.0/19

 Suppress more specific routes unless required by multihoming
Advertising with NO-EXPORT

AS63333  
64.0.0.0/12

Assigns  
64.0.0.0/22

AS62222

Advertises  
64.0.0.0/22  NO-EXPORT

Assigns  
64.0.4.0/22

AS61111

Advertises  
64.0.4.0/22  NO-EXPORT

96.1.0.0/16

AS61000  
96.1.0.0/16