Today’s Lecture

• Intro to Border Gateway Protocol
• BGP Prefix Hijacking Attacks
  – Defense Challenges
  – Real-world Example (2008 YouTube “Attack”)
  – Defensive Solution Techniques
• Attacking BGP Sessions
  – Defensive Solution Techniques
What is BGP?

- BGP: Border Gateway Protocol
  - Routing protocol to exchange routing info across networks on the internet
How BGP Works

• AS-level (autonomous system, collection of networks under single administrative org)
• Relationships between ASes (usually private info)
  – Customer/provider (customers pay to providers)
    e.g. USC to Centurylink
  – Peer-to-peer (peers do not pay each other)
    e.g. USC to UCLA
• Each AS announces routes it knows including entire AS path to the destination
  – All routes announced to customers
  – Customer and own routes announced to providers and peers
• Each AS can choose which routes to adopt
  – Preference given to customer routes, then peer, then provider
  – Preference given to short routes

Example – No Relationships

A ➔ 1.2.3.4 ➔ B ➔ 1.2.3.4 BA ➔ C ➔ 1.2.3.4 FEDA ➔ D ➔ 1.2.3.4 DA ➔ E ➔ 1.2.3.4 EDA ➔ F

<table>
<thead>
<tr>
<th>IP</th>
<th>next</th>
<th>hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2.3.4</td>
<td>A</td>
<td>1</td>
</tr>
<tr>
<td>1.2.3.4</td>
<td>B</td>
<td>2</td>
</tr>
<tr>
<td>1.2.3.4</td>
<td>E</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IP</th>
<th>next</th>
<th>hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2.3.4</td>
<td>D</td>
<td>2</td>
</tr>
</tbody>
</table>
Example – With Relationships 2

<table>
<thead>
<tr>
<th>IP</th>
<th>next</th>
<th>hop</th>
<th>ty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2.3.4</td>
<td>A</td>
<td>1</td>
<td>cust</td>
</tr>
</tbody>
</table>

1.2.3.4 A → prov → 1.2.3.4 B → peer → 1.2.3.4 BA → 1.2.3.4 F → prov → 1.2.3.4 E → prov → 1.2.3.4 DA → 1.2.3.4 E

<table>
<thead>
<tr>
<th>IP</th>
<th>next</th>
<th>hop</th>
<th>ty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2.3.4</td>
<td>F</td>
<td>4</td>
<td>cust</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IP</th>
<th>next</th>
<th>hop</th>
<th>ty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2.3.4</td>
<td>E</td>
<td>3</td>
<td>cust</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IP</th>
<th>next</th>
<th>hop</th>
<th>ty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2.3.4</td>
<td>D</td>
<td>2</td>
<td>cust</td>
</tr>
</tbody>
</table>
Example – With Relationships 3

Consider the cash-flow

$\text{Table:}$

<table>
<thead>
<tr>
<th>IP</th>
<th>next</th>
<th>hop</th>
<th>ty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2.3.4</td>
<td>A</td>
<td>1</td>
<td>cust</td>
</tr>
<tr>
<td>1.2.3.4</td>
<td>D</td>
<td>2</td>
<td>peer</td>
</tr>
</tbody>
</table>

$\text{Diagram:}$

- **A** (IP: 1.2.3.4) connected to **B** (IP: 1.2.3.4) with prov.
- **B** connected to **C** (IP: 1.2.3.4) with peer and prov.
- **C** connected to **D** (IP: 1.2.3.4) with prov and **E** (IP: 1.2.3.4) with peer.
- **D** connected to **E** (IP: 1.2.3.4) with peer.
- **E** connected to **F** (IP: 1.2.3.4) with prov.
Quick Note: IP Prefixes

• Set of IPs that are consecutive and differ only in some number of bits on the right
  1.2.3.0
  1.2.3.1
  1.2.3.2
  1.2.3.3 = prefix 1.2.3.0/30
    – First 30 bits are fixed
    – <8 bits>. <8 bits>. <8 bits>. <8 bits>
    – 1.2.3.<0-3>
    – <00000001>. <00000010>. <00000011>. <000000xx>

• 10.0.0.0 - 10.255.255.255 = ???
Quick Note: IP Prefixes

• Set of IPs that are consecutive and differ only in some number of bits on the right
  1.2.3.0
  1.2.3.1
  1.2.3.2
  1.2.3.3 = prefix 1.2.3.0/30
    – First 30 bits are fixed
    – <8 bits> . <8 bits> . <8 bits> . <8 bits>
    – 1.2.3.<0-3>
    – <00000001> . <00000010>. <00000011>. <000000xx>

• 10.0.0.0 - 10.255.255.255 = 10.0.0.0/8
How BGP Works – More Details

• Routers from neighbor ASes exchange periodic updates using TCP sessions
  – If a TCP session dies (e.g., RST) or HELLO messages are absent assume all routes announced by neighbor are not valid anymore
    • Withdraw your announcements of those routes
    • Creates a rippling effect in the Internet
Example – Withdrawal

<table>
<thead>
<tr>
<th>IP</th>
<th>next</th>
<th>hop</th>
<th>ty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2.3.4</td>
<td>A</td>
<td>1</td>
<td>cust</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IP</th>
<th>next</th>
<th>hop</th>
<th>ty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2.3.4</td>
<td>F</td>
<td>4</td>
<td>cust</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IP</th>
<th>next</th>
<th>hop</th>
<th>ty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2.3.4</td>
<td>E</td>
<td>3</td>
<td>cust</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IP</th>
<th>next</th>
<th>hop</th>
<th>ty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2.3.4</td>
<td>D</td>
<td>2</td>
<td>cust</td>
</tr>
</tbody>
</table>
Example – Withdrawal

**Table:**

<table>
<thead>
<tr>
<th>IP</th>
<th>next</th>
<th>hop</th>
<th>ty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2.3.4</td>
<td></td>
<td>1</td>
<td>cust</td>
</tr>
</tbody>
</table>

**Diagram:**

- **A** connected to **B** via **prov**
- **B** connected to **C** via **peer**
- **C** connected to **F** via **prov**
- **F** connected to **E** via **prov**
- **E** connected to **D** via **prov**
- **D** connected to **A** via **prov**

**Notes:**
- Withdrawal at nodes **C** and **F**
BGP Prefix Hijacking

• A hijacking AS announces itself
  – As origin of a prefix it doesn’t own
  – As being close to the origin of a prefix

• Attracts the prefix’s traffic
  – Can drop it (blackholing)
  – Can reroute it to prefix (interception)
Prefix Hijacking

• Originating someone else’s prefix
  – What fraction of the Internet believes it?
• Originating a more-specific prefix
  – Every AS picks the bogus route for that prefix
  – Traffic follows the longest matching prefix
Defense Challenges

• Malicious routes do not propagate to source
  – Source cannot observe the problem easily
• Even if source can observe the problem it is hard to fix
  – No automatic fix – source’s announcements count as much as anyone else’s once they leave the source
  – Must go through human channels
• Interception attacks are very hard to detect
• Source of attacks is not just maliciousness – often it is misconfiguration
2/24/2008, YouTube Outage

• YouTube (AS 36561)
  – Web site www.youtube.com
  – Address block 208.65.152.0/22

• Pakistan Telecom (AS 17557)
  – Receives government order to block access to YouTube
  – Starts announcing 208.65.153.0/24 to provider
  – All packets to YouTube get dropped on the floor

• Mistakes were made
  – AS 17557: announcing to everyone, not just cust
  – AS 3491: not filtering routes announced by AS 17557
    (will come back to this later)

  Lasted 100 minutes for some, 2 hours for others
Timeline (UTC Time)

- **18:47:45**
  - First evidence of hijacked /24 route propagating in Asia
- **18:48:00**
  - Several big trans-Pacific providers carrying the route
- **18:49:30**
  - Bogus route fully propagated
- **20:07:25**
  - YouTube starts advertising the /24 to attract traffic back
- **20:08:30**
  - Many (but not all) providers are using the valid route
Timeline (UTC Time)

- **20:18:43**
  - YouTube starts announcing two more-specific /25 routes
- **20:19:37**
  - Some more providers start using the /25 routes
- **20:50:59**
  - AS 17557 (Pakistan AS) starts prepending (“3491 17557 17557”)
  - Prepending makes routes longer, less desirable
- **20:59:39**
  - AS 3491 (provider) disconnects AS 17557
- **21:00:00**
  - All is well, videos of cats flushing toilets are available
Lessons From the Example

• BGP is incredibly vulnerable
  – Local actions have serious global consequences
  – Propagating misinformation is surprisingly easy

• Fixing the problem required vigilance
  – Monitoring to detect and diagnose the problem
  – Immediate action to (try to) attract the traffic back
  – Longer-term cooperation to block/disable the attack

• Preventing these problems is even harder
  – Require all ASes to perform defensive filtering?
  – Automatically detect and stop bogus route?
  – Require proof of ownership of the address block?
Solution Techniques

• Defensive filtering
  – Know your neighbors

• Anomaly detection
  – Suspect the unexpected

• Checking against registries
  – Establish ground truth for prefix origination
  – May not be up to date

• Signing and verifying
  – Prevent bogus AS PATHs

• Data-plane verification
  – Ensure the path is actually followed
Defensive Filtering

- Filter announcements from customers that are not for customer prefixes
- Filter announcements from customers that have a large AS on the path
- Keep history of prefix origins and prefer bindings that are long-lived
  - Could do the same for adjacencies in AS paths
  - This violates the basic idea of routing – resiliency
  - Doesn’t work on closeness attacks
- Some but not everyone perform defensive filtering
Attacking BGP Sessions

• BGP session runs over TCP
  – TCP connection between neighboring routers
  – BGP messages sent over TCP connection
  – Makes BGP vulnerable to attacks on TCP

• Main kinds of attacks
  – Against confidentiality: eavesdropping
  – Against integrity: tampering
  – Against performance: denial-of-service

• Main defenses
  – Message authentication or encryption
  – Limiting access to physical path between routers
  – Defensive filtering to block unexpected packets
Denial-of-Service Attacks, Part 1

• Overload the link between the routers
  – To cause packet loss and delay
  – ... disrupting the performance of the BGP session

• Relatively easy to do
  – Can send traffic between end hosts
  – As long as the packets traverse the link
  – (which you can figure out from traceroute)

• Easy to defend
  – Give higher priority to BGP packets
  – E.g., by putting packets in separate queue
Denial-of-Service Attacks, Part 2

• Third party sends bogus TCP packets
  – FIN/RST to close the session
  – SYN flooding to overload the router
• Leads to disruptions in BGP
  – Session reset, causing transient routing changes
  – Route-flapping, changing routes back and forth
• Reasons why it may be hard
  – Spoofing TCP packets the right way is hard
    • Difficult to send FIN/RST with the right TCP header
  – Packet filters may block the SYN flooding
    • Filter packets to BGP port from unexpected source
    • ... or destined to router from unexpected source
    • Turn on SYN cookies
Defense: Exploiting the IP TTL Field

- BGP speakers are usually one hop apart
  - To thwart an attacker, can check that the packets carrying the BGP message have not traveled far
- IP Time-to-Live (TTL) field
  - Decremented once per hop
  - Avoids packets staying in network forever
- Generalized TTL Security Mechanism (RFC 3682)
  - Send BGP packets with initial TTL of 255
  - Receiving BGP speaker checks that TTL is 254
  - ... and flags and/or discards all other packets
- Hard for third-party to inject packets remotely
Defense: Resource Public-key Infrastructure (RPKI)

- Resource public-key infrastructure (RPKI)
- Used for origin validation in routes
  - Cannot validate path
- The organization that sells you an IP range also issues you a certificate that you hold this range (no identity information)
  - Binds your address range to your public key
- When you advertise routes you include a ROA (Route Origin Authorization), showing which ASes can advertise this route
  - Signed with your private key
RPKI


• Slides 3-10
Defense: BGPSEC

• Sign everything you announce
  – Origin and AS_PATH

• Use your private key to sign (same key as in RPKI):
  – Prefix
  – AS_PATH
  – Your AS number, neighbor’s AS number pair

• Check everything when you get announcements

• Generate signed announcements only toward neighbors that support BGPSEC
BGPSEC – Open Problems

• Replay is possible
  – Added timers to route announcements
  – Short timers increase overhead, long timers leave you open to attack longer
• Really large signatures
  – 15 x overhead of regular BGP
  – Really problematic at convergence time
  – Disable optimizations such as update packing
• Validating route announcements is expensive computationally
  – Much more than processing BGP updates
• If each router uses a separate public key BGPSEC enables others to learn about internal ISP topology
Recap

• Intro to Border Gateway Protocol
• BGP Prefix Hijacking Attacks
  – Defense Challenges
  – Real-world Example (2008 YouTube “Attack”)
  – Defensive Solution Techniques
• Attacking BGP Sessions
  – Defensive Solution Techniques
• Now:
  – Wrap up CTF 2 Phase 2
• Next Lecture:
  – Article presentations (remote!)