Distribution of Source Address Validation State in BGP (SAV-D)

draft-haas-savnet-bgp-sav-distribution-00

Jeffrey Haas, Juniper Networks

<jhaas@juniper.net>
Goals

• This proposal deals with the efficient distribution of SAV Tables to forwarding elements in the network for SAV enforcement using BGP as the transport protocol.

• The BGP speakers that participate in originating the SAV-D state are “SAV-D Controllers”.
Non-Goals

• This draft does not deal with how the controller calculates the SAV Tables for the network.
Operational Model

• Interfaces on forwarding elements are provisioned to be part of logical SAV Tables.
• This provisioning is in the form of “membership”.
• Four classes of membership are currently defined:
  • Interface-specific: The SAV Rules are for a specific interface on this forwarding element.
  • Interface-set: The SAV Rules are for a group of interfaces on a specific forwarding element based on a group number.
  • Peer-AS: The SAV Rules are for interfaces that are peering with a specific BGP AS Number.
  • Origin-AS: This interface is interested in SAV Rules associated with a specific Origin AS.
A simplified Network-Wide SAV Table

<table>
<thead>
<tr>
<th>Router</th>
<th>Interface</th>
<th>Prefixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>a</td>
<td>{P1}, {P3}, {P4}</td>
</tr>
<tr>
<td>R1</td>
<td>b</td>
<td>{P2}, {P5}</td>
</tr>
<tr>
<td>R2</td>
<td>c</td>
<td>{P2}, {P5}</td>
</tr>
</tbody>
</table>
A simplified Network-Wide SAV Table

<table>
<thead>
<tr>
<th>Router</th>
<th>Interface</th>
<th>Prefixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>a</td>
<td>{P1}, {P3}, {P4}</td>
</tr>
<tr>
<td>R1</td>
<td>b</td>
<td>{P2}, {P5}</td>
</tr>
<tr>
<td>R2</td>
<td>c</td>
<td>{P2}, {P5}</td>
</tr>
</tbody>
</table>

**Interface-specific** membership could be used to populate each of the interfaces on the above routers.

For example, interface a will get SAV Rules permitting prefixes {P1}, {P3}, {P4}
A simplified Network-Wide SAV Table

<table>
<thead>
<tr>
<th>Router</th>
<th>Interface</th>
<th>Prefixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>a</td>
<td>{P1}, {P3}, {P4}</td>
</tr>
<tr>
<td>R1</td>
<td>b</td>
<td>{P2}, {P5}</td>
</tr>
<tr>
<td>R2</td>
<td>c</td>
<td>{P2}, {P5}</td>
</tr>
</tbody>
</table>

**Interface-set** membership could be used to populate interfaces that share common properties.

For example, interfaces b and c can receive the same prefixes, {P2}, {P5}.
Peer-AS membership could be used to populate interfaces share a common peer-AS.

For example, interfaces b and c both peer with AS 200 and will have \{P2\}, \{P5\}.
A simplified Network-Wide SAV Table

<table>
<thead>
<tr>
<th>Router</th>
<th>Interface</th>
<th>Prefixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>a</td>
<td>{P1}, {P3}, {P4}</td>
</tr>
<tr>
<td>R1</td>
<td>b</td>
<td>{P2}, {P5}</td>
</tr>
<tr>
<td>R2</td>
<td>c</td>
<td>{P2}, {P5}</td>
</tr>
</tbody>
</table>

Origin-AS membership can be used to cover a BGP prefix cone.

For example, a has membership for routes originated from AS 100, AS 300, AS 400.
SAV-D NLRI

• A new AFI/SAFI for distributing SAV enforcement state.
• The NLRI carries sources and provides the key for memberships.
• NLRI is a new Route Distinguisher Type (8 octets) + IPv4 (AFI 1) or IPv6 (AFI 2) source prefix.
• Since the same source prefix may belong to a very large number of groups, a controller needs the ability to originate the same prefix multiple times.
  • Route Distinguisher is 4 octets BGP Identifier of originating router as a Global Administrator, and 2 octets of Local Administrator field.
Membership via Extended Communities

• Device Scoped membership is signaled using a Node Target Community and:
  • SAV-D Interface Member - carries an ifIndex
  • SAV-D Device-Specific Group Member - carries a 6 octet Local Admin Field.
• Non-scoped membership:
  • SAV-D Neighbor-AS Member - carries the neighbor AS.
  • SAV-D Origin-AS Member - carries the origin AS.
Efficient distribution using RT-Constrain (RFC 4684)

• Interface membership can be signaled using RT-Constrain as a “subscription” to SAV Rules of a given membership type.

• Device-scoped membership only subscribes for Node Target specific SAV-D routes, not the specific local types.
Carrying Traffic Handling Actions

• Traffic handling actions in other SAV drafts are modeled after BGP Flowspec (RFC 8955, et al.)

• Flowspec currently signals these actions via additional Extended Communities. Redirect, rate limit, etc.
  • Combinations of these are already problematic in Flowspec, so this is mostly example usage. Current work for BGP Flowspec could be leveraged for SAV-D.
Why not just do this in Flowspec?

• The ”membership” is the novel piece of work for this proposal. Flowspec might benefit from some of these same behaviors - but maybe not all.

• SAV Rules could be implemented by the forwarding element’s firewall, but maybe VRF mode or other enforcement is used.

• SAV Rules are without explicit order, they’re longest match. (Are they? Compare vs. uRPF...)

• There’s a legitimate discussion to be had as to where SAV enforcement takes place in the forwarding pipeline. With firewall? Before firewall? After firewall and before dst forwarding?
A simplified Network-Wide SAV Table

<table>
<thead>
<tr>
<th>Router</th>
<th>Interface</th>
<th>Prefixes</th>
<th>Subscribes To</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>a</td>
<td>{P1}, {P3}, {P4}</td>
<td>ifIndex A</td>
</tr>
<tr>
<td>R1</td>
<td>b</td>
<td>{P2}, {P5}</td>
<td>Peer-AS 200</td>
</tr>
<tr>
<td>R2</td>
<td>c</td>
<td>{P2}, {P5}</td>
<td>Origin-AS 500</td>
</tr>
</tbody>
</table>

Router Interface Prefixes Subscribes To

R1 a \{P1\}, \{P3\}, \{P4\} ifIndex A
R1 b \{P2\}, \{P5\} Peer-AS 200
R2 c \{P2\}, \{P5\} Origin-AS 500

AS 100 \{P1\}
AS 200 \{P2\}
AS 300 \{P3\}
AS 400 \{P4\}
AS 500 \{P5\}

R1
R2
Controller
Questions?