BIER
Bit Indexed Explicit Replication

Greg Shepherd– Distinguished Engineer
shep@cisco.com
Agenda

• Background
• Intro to BIER
• IETF Status
• Conclusion
Introduction

A little History…
Background – IP Multicast History

- Steven Deering, 1985, Stanford University
- RFC988, 1986 (Obsoleted by RFC1112, 1989)
- Multicast is part of the IP protocol stack
- Intended as an Internet-wide end-to-end service
- Primary focus – to create an L2-overlay on top of IP

“Your routed network broke my L2 applications!”
What is the problem with Multicast

• Each Tree has its own unique receiver population.
• To efficiently forward/replicate there is a Tree per flow!
• This means State is created in the network.
• State convergence times can cripple a Multicast Network
Background – IP Multicast Challenges

- Explicit Tree Building Protocol
  - End-to-end tree state per flow
  - RPF tree building can cause multicast traffic to take different paths than unicast traffic
  - Convergence times negatively impacted by tree state
  - No way to aggregate state without sacrificing optimal delivery
    - Choose between state explosion or data flooding

- Data-driven events

- Specialized skill set to troubleshoot and maintain
  - High operational costs
Multicast Routing State

• State is created in the network using a Multicast routing protocol like PIM, mLDP, RSVP-TE, Tree-SID.
• State means resources consumed, memory/CPU.
• Convergence is impacted by the amount of State.
• In order to manage the network, the protocol needs to be understood by the network operator.
• Different levels of complexity based on protocol choice.
Why?
IP Multicast isn’t IP

- IP routing protocols calculate SPF on a topology
- Unicast RIB only holds topology state
- IP Group address is an abstraction, not a destination
  - Identifies a “statefull overlay”
  - Tree built using unicast topology, but forwarding on tree state.
- IP Multicast was ’wedged’ into IP because that’s all we had.
- MPLS opened the door to a new forwarding model
- Opening our eyes to a need for a dedicated multi-point forwarding model took a bit more time..
Today

• The value of multi-point (multicast) services are well understood
• The challenges of the current solutions (PIM) often result in a failed cost/benefit analysis
• Only those networks with an overwhelming business need have successful multicast deployments
• Customers often say they hate multicast, when in reality they hate PIM
• Customers value network replication but not often at the cost of deploying and maintaining PIM in their network
• Can we do better?
What is BIER?
What is BIER?

• BIER is a new forwarding paradigm to forward and replication multicast packets through the network.

• Packets are forwarded using a special Header that is embedded into the packet.

• Routers build a special forwarding table to forward/replicate using the BIER header.

• BIER forwarding is State Less!
How does BIER work

• We give the Egress routers an identifier.
• The Ingress router includes the “identifiers in the packet.
How does BIER work

- Packet is forwarded hop-by-hop using the “identifier”
- Each “identifier” is forwarded along the unicast (SPF) path.
How does BIER work

• The smaller the identifier, the more we can fit into a single packet, how small can an identifier be?

• A single Bit!!!

• With BIER the Egress Identifier is a Bit Position.

• We include a BitString of 256 bits into the packet.

• Manipulations of a BitString is much easier compared to including a list of numbers.
How does BIER work

- Each Egress BIER router has a unique Bit Position.
- Routers maintain a forwarding table of Bitmask's
Bit Index Forwarding Table (BIFT) by IGP

- The IGP carries an extension to map the Bit Position to the Routers Loopback address.
- Using this mapping information the BIFT is created.
- This Table follows the IGP SPF.
Suppose A leans about D’s interest, in the blue multicast flow. (via BGP, SDN, STATIC, etc…)

**Forwarding Packets**

```
<table>
<thead>
<tr>
<th>BM</th>
<th>Nbr</th>
</tr>
</thead>
<tbody>
<tr>
<td>0111</td>
<td>B</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BM</th>
<th>Nbr</th>
</tr>
</thead>
<tbody>
<tr>
<td>0011</td>
<td>C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BM</th>
<th>Nbr</th>
</tr>
</thead>
<tbody>
<tr>
<td>0100</td>
<td>E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BM</th>
<th>Nbr</th>
</tr>
</thead>
<tbody>
<tr>
<td>0001</td>
<td>D</td>
</tr>
<tr>
<td>0010</td>
<td>F</td>
</tr>
<tr>
<td>0100</td>
<td>E</td>
</tr>
</tbody>
</table>
```

BFR-ID 1
BS:0001

BFR-ID 2
BS:0010

BFR-ID 3
BS:0100

D

F
Suppose A learns about D and E’s interest, in the blue multicast flow. (via BGP, SDN, STATIC, etc…)

Forwarding Packets
Suppose A learns about D, E and F's interest, in the blue multicast flow. (via BGP, SDN, STATIC, etc…)

Forwarding Packets
Forwarding Packets

- As you can see from the previous slides, the result from the bitwise AND (&) between the Bit Mask in the packet and the Forwarding table is copied in the packet for each neighbor.
- This is the key mechanism to prevent duplication.
- Look at the next slide to see what happens if the bits are not reset.
- If the previous bits would not have been reset, E would forward the packet to C and vice versa.
Forwarding Packets (wrong behavior)

A

BM | Nbr
---|---
0111 | B

AND

0111

A

BM | Nbr
---|---
0111 | C
0100 | E

AND

0111

B

BM | Nbr
---|---
0001 | D
0010 | F
0100 | E

AND

0111

BFR-ID 1
BS:0001

D

AND

0111

E

BM | Nbr
---|---
0011 | C
0111 | B

AND

0111

F

BFR-ID 2
BS:0010
How many Bits and Where?

• The number of multicast egress routers that can be addressed is depending on the number of Bits that can be included in the BitString.

• The BitString length is dependent on the encapsulation type and router platform.

• IETF BIER WG has agreed to 256bits as minimum required support.

• Encap
  1. MPLS, below the bottom label and before IP header.
  2. Native w/ BIER Ethertype
MPLS encapsulation

• The Top Label is allocated by BIER from the downstream platform label space.

• The BIER Header follows directly below the BIER label.

• There is a single BIER label on top, unless the packet is re-encapsulated into a unicast MPLS tunnel.

• The VPN label is allocated from the upstream context label space (optional).
BIER Header

Sets and Areas
BIER Sets

- To increase the scale we group the egress routers in Sets.

<table>
<thead>
<tr>
<th>Set</th>
<th>BM</th>
<th>Nbr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0111</td>
<td>I</td>
</tr>
<tr>
<td>2</td>
<td>0111</td>
<td>I</td>
</tr>
</tbody>
</table>

Note, Bit Positions 1, 2, 3 appear in both Sets, and do not overlap due to Sets.

Note, we create different forwarding entries for each Set.
BIER Sets

<table>
<thead>
<tr>
<th>Set</th>
<th>BM</th>
<th>Nbr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0111</td>
<td>I</td>
</tr>
<tr>
<td>2</td>
<td>0111</td>
<td>I</td>
</tr>
</tbody>
</table>

Note, we create different forwarding entries for each Set

- There is no topological restriction which set an egress belongs to
BIER Sets

• If a multicast flow has multiple receivers in different Sets, the packet needs to be replicated multiple times by the ingress router, for each set once.

• Is that a problem? We don’t think so…

• The Set identifier is part of the packet.

• Can be implemented as MPLS label.
BIER Areas

• The ABR removes the BIER header from Area 0, and imposes a new BIER header for Area 1 and 2.

• The new BIER header can be determined by a Group/Label lookup.
  • Look for the inner IPv4/6 packet group address, do a lookup in the MFIB
  • Requires flow state on the ABR.

• Similar to Segmented Inter-AS MVPN
MVPN over BIER
MVPN over BIER

- BIER replaces PIM, mLDP, RSVP-TE or IR in the core.
- BIER represents a full mesh (P2MP) connectivity between all the PE’s in the network.
- There is no need to explicitly signal any MDT’s (or PMSI’s).
- With MVPN there are many profiles,
  - This is partly due to the tradeoff between ‘State’ and ‘Flooding’.
  - Different C-multicast signaling options.
- MVPN over BIER, there is one profile.
  - BGP for C-multicast signaling.
- No need for Data-MDTs.
MVPN over BIER

- The BGP control plane defined for MVPN can be re-used.
- Big difference, there is no Tree per VPN…!!!
- The BIER packets needs to carry Source ID and upstream VPN context label
IETF

- The BIER idea was presented in a BOF at the IETF in Hawaii.
  - November 2014.
- A new BIER Working Group has been formed (bier@ietf.org)
- BIER architecture became RFC 8279 (November 2017)
- BIER work re-chartered as Standards Track (March 2018)
- Vendors collaborating (co-authoring) with us;
BIER Conclusion
Stateless

• There is no Multicast receiver or flow state in the core network (only edge).
  • Imposition of the BIER Header may be done by application, removes state from ingress.

• There is no tree state in the network.

• There is no tree building protocol or logic in the network.

• There is only topology state for the BFER’s, derived from unicast routing.
Scale

• Since there is no flow and tree state, converges as fast as unicast.
• Compared to Ingress Replication, saves 256x (minimum)
Simplicity

• No Reverse Path Forwarding (RPF)
• No Rendezvous Points
• No shared tree / source tree switchover
• No receiver driven tree building
• No flow state
• BIER is like unicast
More information and references

• bier@cisco.com
• https://dcloud-cms.cisco.com/demo/cisco-bit-indexed-explicit-replication-v2
• https://datatracker.ietf.org/wg/bier/charter/