Multicast/BIER As A Service

IETF 109, virtually Bangkok

Zhaohui (Jeffrey) Zhang, Juniper

Eric Rosen, Individaul

Daniel O. Awduche, Verizon

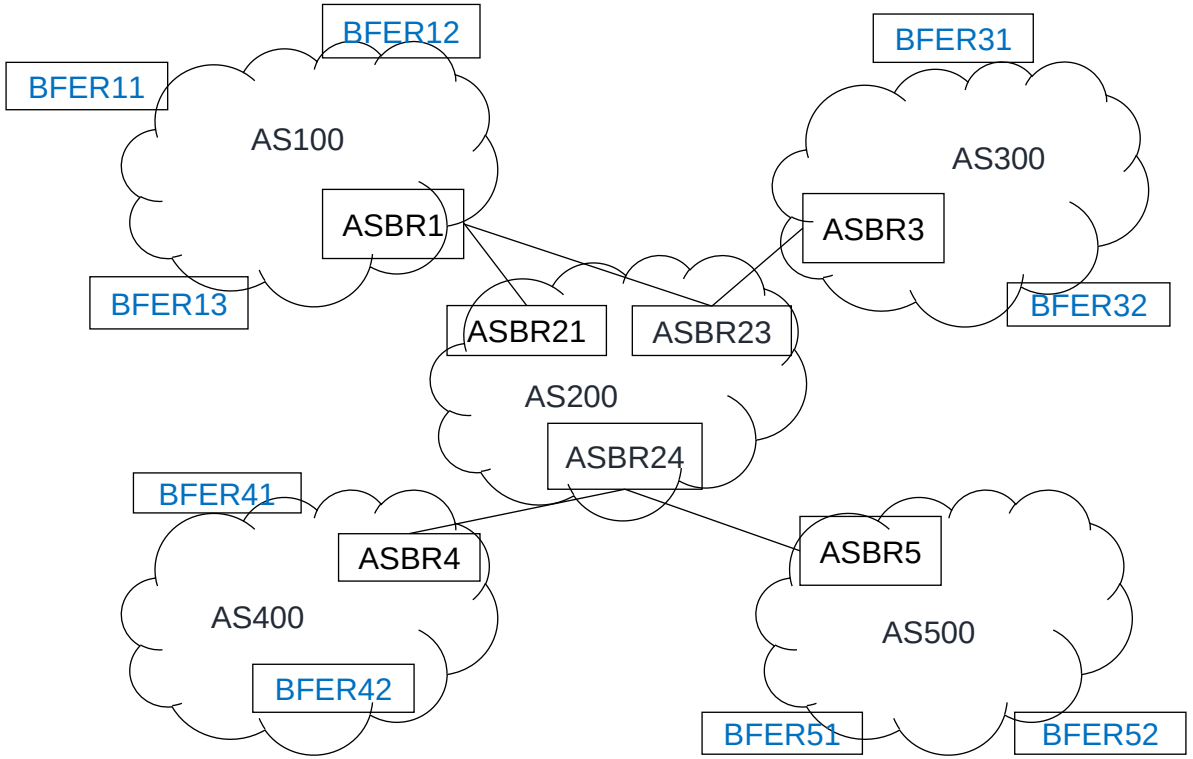
Liang Geng, China Mobile

Summary

- Presented in IETF103, physically in Bangkok
- Stable and ready for WG adoption
  - Seeking comments and adoption
- Two slides for quick review
  - All slides from IETF103 included at the end for references

Single Operator Example

- BFERs all over the places
  - Starting w/o BFRs
    - Essentially Ingress Replication
  - Gradually add BFRs at strategic points
    - E.g. Turn on BIER on ASBR23 & ASBR24
    - Use Tunnel Encapsulation Attribute (TEA) to steer traffic along BIER enabled paths
  - Enable BIER in AS/area where native BIER forwarding is feasible
  - Any kind of tunnel will work
- BGP based BIER signaling
  - draft-ietf-bier-idr-extensions
  - Plus TEA extensions specified in this draft
- IGP signaling in AS/area where most routers are BFRs
  - With BGP-IGP BIER prefix redistribution
- Multi-AS
  - Initially no segmentation
    - Fewer than 256 BFERs or multiple sets used
  - Segmentation can be incrementally added



Multi-operator Example: BIER as a Service

- AS200 is a 3rd party provider providing BIER transport to its BIER customers
- No BFER in the provider, hence no customer (s,g) state
  - What if segmentation is needed?
    - xPMSI state maintained on segmentation points
      - Inclusive or some (s/*,g) Selective PMSI
    - Optionally, a customer equipment (physical or virtual) can be tethered as segmentation point
- Each customer can have their own BIER domain transported through the provider, even with conflicting subdomain-id and BFR-id
  - Use per-customer RD with BIER Prefix
    - BIFT is now per <RD, subdomain-id, bsl, set>; or a sub-domain is now per <RD, subdomain-id>
    - What if we need to redistribute BIER info between BGP and IGP?
      - In IGP signaling, use a BIER Domain sub-TLV to encode the RD and BIER Info
- BFRs in AS200 need to scale on number of BIFTs and BIFT entries
  - E.g., 256 BIFTs, each with 256 entries (64k routes in total)

Slides from IETF103

Current Multicast Use Cases

- Enterprise Applications
  - FSI Financial Data Distribution
- Service Provider
  - Live TV/video distribution inside a provider itself
  - Customer Multicast/BUM for VPN/EVPN
  - Internet Multicast is minimum
    - Mbone is mostly in Internet2, w/o much real usage

Multicast As A Service?

- Only in the form of MVPN or EVPN BUM
- Other multicast transport by SPs virtually non-existent
  - E.g., can an SP provide multicast transport for a non-VPN 3rd party?
    - E.g. for a content provider who does not have its own all-reach network?
  - Lack of confidence/interest on service provider side
    - Complexity and scalability concerns – signaling and per-flow state
    - Profitability concerns
      - Lots of multicast flows are low volume
      - For high volume (e.g. video) traffic, how to bill?
  - Lack of interest on customer side
    - Lack of provider support
    - Content providers resorted to p2p/p2sp (peer to peer or peer to server/peer)
  - Chicken & Egg problem

BIER Enables MaaS

- BIER removes per-flow state
  - Significantly simplifies multicast control plane
  - Significantly improves scalabilities
- BIER can help break the chicken & egg vicious circle
  - It can encourage service providers to provide multicast transport services
    - In addition to using BIER for its own MVPN/EVPN services
  - It can encourage content providers to use multicast for delivery
- Potential use cases for MaaS
  - CDN (large scale high definition live broadcast or content pushing)
  - Any large scale high rate data distribution

Current Common BIER Use Cases

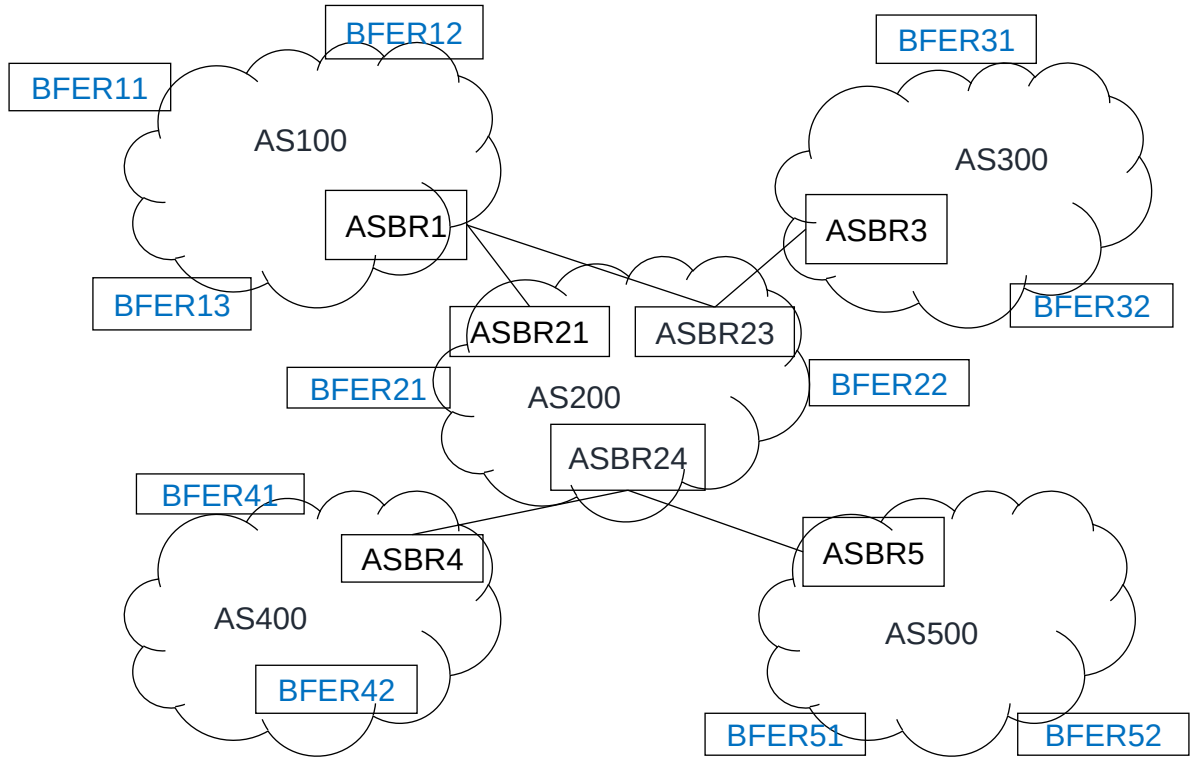
- Current use cases have entire BIER sub-domain (BFERs and BFRs) under the same operator
  - BIER as provider/underlay tunnels for MVPN/EVPN-BUM
    - End-to-end multicast flow in overlay
  - BIER sub-domain as part of end-to-end multicast tree
    - E.g. PIM signaling as “BIER Multicast Flow Overlay”
    - Similar to “mLDP Inband Signaling”
    - An end-to-end multicast tree could have multiple unrelated BIER sub-domains
- Most likely IGP is the BIER signaling protocol

BIER Enabled MaaS

- BFERs/BFRs may be under separate operators
  - BFERs owned by a customer
  - Service providers don’t have to worry about per-flow state at all
    - BFRs do need to know how to route to customer BFERs
- An operator may provide BIER based transport for many customers
  - Independently for each customer
- Mainly BGP signaling
  - OTT tunneling very common
  - IGP signaling may be used in an area/AS where most devices support BIER

A Simple Example

- Single Operator (e.g. a content provider's own all-reach network)
- BFERs all over the places
  - Starting w/o BFRs
    - Essentially Ingress Replication
  - Gradually add BFRs at strategic points
    - E.g. Turn on BIER on ASBR23 & ASBR24
- BGP based BIER signaling
  - draft-ietf-bier-idr-extensions
- Multi-AS but (initially) no segmentation
  - Either have fewer than 256 BFERs or multiple sets are used



A Couple of Details

- In the previous slide, BFER11's shortest path to BFER21 is through ASBR21, which does not support BIER, while BFER23 does
  - For AS100 to send BIER traffic to ASBR23:
    - Only ASBR23 should re-advertise BFER21's BIER info
      - Incongruent unicast/multicast path
- Preventing tunneling to BFERs directly
  - Tunnel Encap Attribute: attached by an BFER, updated by each BFR that changes BGP Next Hop, and used as the BIER neighbor to replicate traffic to
    - BFER42 uses its own BIER prefix as tunnel destination address
    - ASBR24 changes it to its own BIER prefix; ASBR23 changes again
    - For BFER11 to reach BFER42, it tunnels to ASBR23, who then tunnels to ASBR24, who then tunnels to BFER42

Turn on BIER inside an AS/Area

- In the previous slide, BIER traffic are tunneled between a few strategically placed BFRs
  - BFER11 tunnels (Ingress Replicates) to BFER12/BFER13/ASBR1
- If enough routers in AS100 supports BIER, AS100 can run BIER internally
  - The entire network (across ASes) is still a single sub-domain
  - With mixed IGP and BGP signaling for BIER
    - https://tools.ietf.org/html/draft-zwzw-bier-prefix-redistribute used to redistribute BIER info between IGP and BGP
      - BFER11/BFER12/BFER13's BIER Prefixes and BFR-IDs are re-advertised into BGP by ASBR1
      - Other BFERs' BIER prefixes and BFR-IDs are re-advertised into IGP by ASBR1
  - This does require redistribute BFER prefixes into IGP

Segmentation

- If the number of BFERs is very large, segmentation can be used
  - E.g. each AS/area is an independent BIER sub-domain
  - A segmentation point maintains xPMSI (or PIM) state, decapsulate BIER header in the upstream sub-domain and forward to a downstream sub-domain (label switch or PIM based forwarding) with a new BIER header
  - Use Route Targets or policy to restrict BIER info to each sub-domain
- This is reasonable for this single operator case
- If a deployment started with fewer PEs w/o segmentation, segmentation can be introduced incrementally
  - Add a BFR as or convert an existing BFR to a segmentation point
  - Make sure it does not re-advertise BIER information between two sub-domains
  - Make sure BFRs/BFERs in a sub-domain only exchange BIER information among themselves (including the segmentation points)

Multi-Operator Case

- What if AS200 in the earlier simpler example does not belong to the content provider that owns the BFERs?
- With BGP based signaling, it still works
- AS200 is now providing MaaS
  - BIER as a Service (BaaS) to be more accurate

BIER as a Service

- Provided by AS200
- BIER level; no BFER, hence no customer (s,g) state
  - What if segmentation is needed?
    - xPMSI state maintained on segmentation points
      - Inclusive or some (s/*,g) Selective PMSI
    - Optionally, a customer equipment (physical or virtual) can be tethered as segmentation point
- What if different customers have conflicts in subdomain-id and BFR-id?
  - Use per-customer RD with BIER Prefix
    - BIFT is now per <RD, subdomain-id, bsl, set>; or a sub-domain is now per <RD, subdomain-id>
    - What if we need to redistribute BIER info between BGP and IGP?
      - In IGP signaling, use a BIER Domain sub-TLV to encode the RD and BIER Info
- A BFR needs to scale on number of BIFTs
  - E.g., 256 BIFTs, each with 256 entries (64k routes in total)

MaaS Control & Billing

- A provider can have policies to control:
  - Whether/how it re-advertises certain BIER prefixes, e.g. to certain peers only
  - Whether it advertises its own BIER prefixes (with a certain RD)
    - i.e. whether it becomes a BFR for a particular customer
      - This controls the number of BIFTs that it instantiates
- A provider can count traffic and bill accordingly:
  - At an entry point: incoming BIER packets for each BIER label that it advertises
  - At an exit point: outgoing packets for each BIER label that it imposes

Summary

- Scalable MaaS enabled by BIER
- “BIER Transport Service” to be more accurate
  - Leave BFER (and customer specific state) to customers
  - Existing MVPN/EVPN with BIER can provide traditional multicast service
- Incrementally expandable
- With policy control and billing