draft-ietf-pim-bier-frr-07

Review / Discussion

BIER-WG, IETF122 Bangkok

Toerless Eckert, Futurewei USA (tte@cs.fau.de) (as RTG-DIR reviewer)

Why this ...

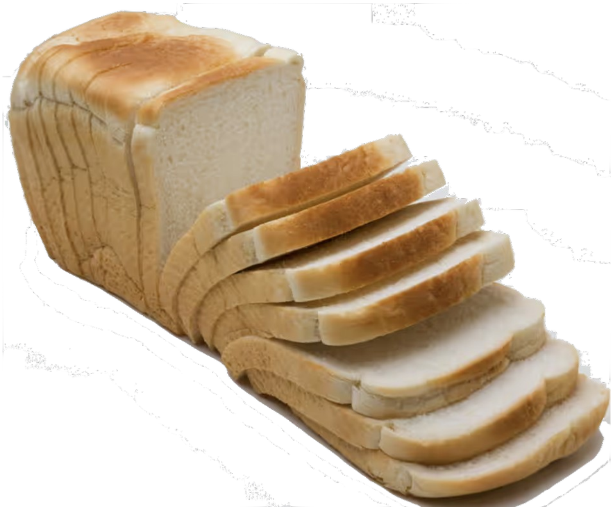
could be the key selling feature for BIER

BIER

Many Service Provider customers over decades adament about „sub 50msec protection for multicast services

- Not even RSVP-TE/P2MP made it work well for node protection
- Design attempts / patents for stateful / PIM – ETOOCOMPLEX
- Only „ECMP“ (dual-plane) style FRR (MoFRR) works reasonably well with stateful multicast (PIM, mLDP – But that‘s poor-mans-live-live (more traffic)
- draft-ietf-pim-mofrr-tilfa-09 – Does not work for cases where basic MoFRR does not work.
- FRR for BIER-TE.. Tried to define... Too complex (IMHO)

BIER+FRR



So, when BIER-FRR was worked on, i did not check

sigh

So now here at tailend of doc (drafted by RTG dir review

FRR for multicast can work best exactly with only BIER

Not even possible with proposed „encoded trees instead of bitstring“

So could be unique selling point for BIER in SPs that like FRR

But significant text shortcomings

Issues (1)

- Not doing a good sales job (see prior slide)
  - Stateful multicast node protection way too difficult:
    - Need to have backup p2p tunnels to every next-next-hop to compensate for failures of next-hop
    - Many duplicate copies over same link (next-hop with fan-out of 10 -> may need 10 FRR tunnel copies over same link)
    - And need to expand signaling protocol (pim/mLDP next-next-hop signaling)
    - Would need even backup pre-established multicast FRR trees to optimize
    - Terrible!
  - BIER makes it simple because you just need to „somehow tunnel enough around the the failure“
    - And once the packet exits that tunnel, BIER replication will take over again. Where ever you exited!
      - Exit point of tunnel may only cause more or less duplicate packets on a transit link, but never duplicate BFER reception!
    - You only need to determine the right set of bits (BFER) for the FRR-tunneled packet
    - There are even some new IGP algorithms to invent/explore to determine the best possible tunnels (some good heuristics, some NP complete optimization issues – depending how well you want to minimize FRR traffic!)
    - Many new degrees of freedom to optimize

  - (Almost) no new interop issue. Could just implement totally locally on single routers!
    - Except: Need to know which routers are BFR able & willing to receive unicast encapsulated BIER packets on which address/SID
    - Note: This candidate IGP extension dependency is also not noted
    - No signal at all needed if you can assume that any of the pre-existing known SID/Addresses of all known BFER can receive unicast FRR tunneled packets.

Issues ()

- Not mentioning MPLS.
  - Is MPLS/SR-MPLS not a desirable option for networks where BIER is targeted ?
    - Not tried to toroughly vet if/what differences in explanations might be missing (hopefully not much)
    - Suggestion: should describe BIER-FRR is applicable equally to IP and MPLS unicast forwarding
    - Forward looking: Mainly consider networks supporting SR – SR-MPLS / SRv6 (TI-MoFRR backup tunnels)
    - Can still consider RLFA with IP / MPLS tunnels – but „classic“ MPLS (LDP signaling) goes away anyhow.

- Confusing use/classifications of „tunnel-based“ vs. „FRR-based“ Contradicting between section 2 and 3.
  - Suggestion: Consider „Tunnel based“ to be ONLY for special case of pre-existing unicast link-protection:
    - No BIFT work required – packets solely need to be encapsulated with protected unicast adjacency to primary BFR-NBR
    - BIFT changes only required if FBM in failure case need to differ from FBM in failure case
    - And existing FRR protected unicast adjacency code can be used. No need to do ANY FRR work in BIER

Issues (3)

Q: Are there actual implementations of „partial deployment BIER“ ?

- With automatic remote BIER-NBR calculation in IGP ?
- That is a new IGP calculation i have not seen required for unicast.
  - And it‘s not an „exact science“ (aka: there is some good leeway in selecting „best remote next-hop-BFR“ from IGP)
- So i doubt that IGMP implementations in commercial routers where eager to implement it for BIER
- Its a cool feature of BIER though, so it would be great to have

Reason: Doc partially discusses „remote BFR-NBR“

- Which would only be of interest if we did have support in BIER implementations for this
- If it is rare/likely not supported, we could remove this complexity
- Or move all considerations avbout this to a separate section – and expand on complexities there.

- It would require some form of directed fast BFD ... So not sure how important for FRR this is....

Issues (4)

Needs simpler reference example to explain the core aspects of operation

Not a single complex topology that can represent everthing that will ever need to be explain – but is much harder to understand

Example 1: How can the order of processing BIFT entries impact the number number of packet copies to send when FRR happens ?

BFR-id	F-BM	BFR-NBR	BF-BM	BBFR-NBR	BFA	BEA	
1	Abm	A	...	...	...	0	First packet copy
2	Bbm	B	Bbm|Abm	A	...	1	Second packet copy

BFR-id	F-BM	BFR-NBR	BF-BM	BBFR-NBR	BFA	BEA	
2	Bbm	B	Bbm|Abm	A	...	1	Only packet copy
1	Abm	A	...	...	...	0	

Issues (5)

Example 2: Why do i ever need to send two or more FRR packet copies
... If i could simply put all necessary bits into one copy

Answer: Node failure can partition Q-space

```
                                    ------- BFER7
                                   /
           ----BFR3 --- BFR4----- BFR1 ---- BFER1
          /                      /
BFIR---BFRa ------ PLR --L1-- Pbfr
          \                      \
           ----BFR5 --- BFR6----- BFR2 ---- BFER2
                                   \
                                    ------- BFER8
```

Packet for BFER1,2,7,8 arrives at PLR, Pbfr failed. PLR needs to have previously created to FRR tunnels – one to BFR3 to reach BFER1,7 (first Q-space), one to BFR5 to reach BFER2,8 (second Q-space)

Issues (5)

Why the heck do i need to know what P and Q space are ?

Because that is the standard FRR terminology, See RFC7490

- Should use this existing Unicast FRR terminology to correctly explain what tunnels to build – unicast FRR experts understand it as easily as possible.
- Not sure how much of terminology rehash is needed (Read the fine RFC), but explaining BIER-FRR may benefit from even expanding it:
- Bitset-Q-Space: set of routers from whom all BFER in Bitset are reachable.

Draft only describes one node-protection FRR strategy

- Determine for each BFER reachable through BFR-NBR the BFR-Nbr-Next-hop
- Create backup tunnel to each BFR-Nbr-Next-hop with FBM being „all“ BFR reachable through them
- Problems: Not the only strategy, likely not the best, and even within this strategy, order of BFR-Nbr-Next-hop matters.

Issues (6)

No mention of BIER-ECMP and how to deal with it in BIER-FRR

Two proposed BIFT implementations are IMHO just examples

- Should be explained
- Would be useful to discuss if NPU would do better with even different approaches
- Eg.: If BIER is actually implemented as sequential replication, then simpler FRR should be possible (putting FRR into adjacency code – but returning completed FBM which may be different in case of FRR)

More explanations of IGP algos to calculatedesired tunnels may be required (as done by more normative unicast FRR RFCs).

Conclusions ?

Mike McBride made good suggestions on list already:

Pick up for this to-be-RFC whatever you feel is doable

- Document is just informational so expections should not be high
- Of course, i don‘t know what more concerns IETF/IESG will have.

If/when there is interest for a better version, spin out a new RFC with maybe experimental stat

- Feasibility of BIFT extensions of course at the core of adoption considerations!
- But see first slide: could be a big win for BIER adoption if we make/sell it best!