



Extensions to BIER Tree Engineering (BIER-TE) for Large Multicast Domains and 1:1 Protection, v1.2

draft-fluechter-bier-bierte-subset-tunneling-00

Moritz Flüchter <moritz.fluechter@uni-tuebingen.de>,
Michael Menth <michael.menth@uni-tuebingen.de>,
Toerless Eckert , Futurewei USA <tte@cs.fau.de>

<http://kn.inf.uni-tuebingen.de>



- ▶ Formalize / Specification for scaling BIER-TE example from RFC9262, 5.3.6
 - “Ingress unicast TE tunnels” into “egress densely connected” BIER-TE topology subset
 - Where each such egress topology subset fits into one bitstring
 - Based on typical large Wide-Area-Network bandwidth designs
 - Most replication/fan-out happens at egress / towards subscribers
 - Subscriber egress bandwidth often most constrained. Avoid duplicate traffic over links
 - Ingress / e.g.: “Video-head-end” BFIR->Core attachments very fast. Can afford unicast replication

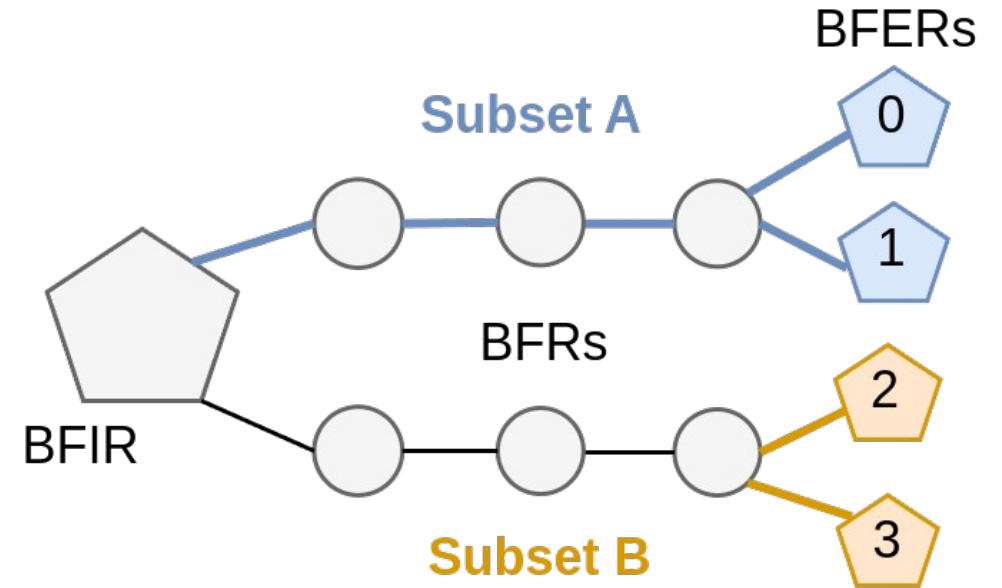
- ▶ Add FRR for resilience
 - Reuse / expect FRR within “egress BIER-TE topology subset
 - Reuse / expect FRR (as desired) in unicast TE tunnel mechanism (SR-TE with TI-LFA etc.)
 - Specify FRR between unicast TE tunnel and egress BIER-TE topology subset
 - <https://arxiv.org/pdf/2409.07082> (early draft)



- ▶ Scalability with BIER subsets
 - Bitstring describes subset of BFERs in domain
 - Set Identifier (SI) in header
 - Simple constraints
 - Subset size must not exceed bitstring length
 - Sum of subsets should cover all BFERs

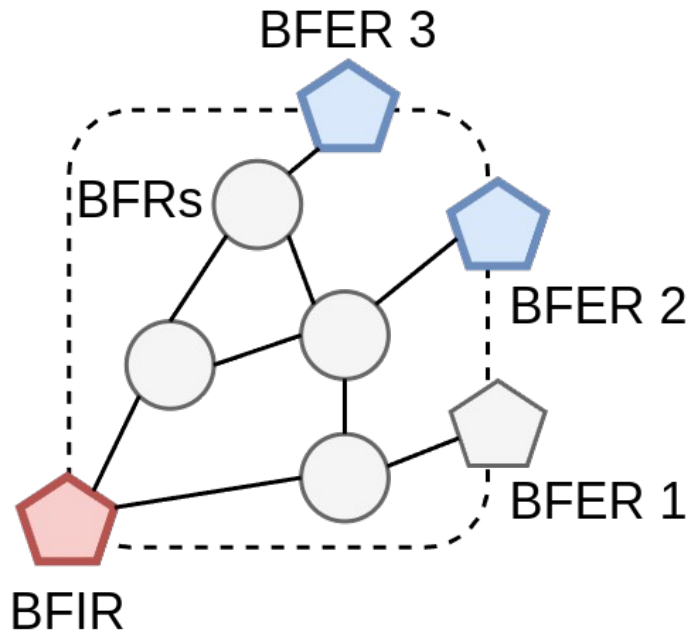
- ▶ BIER-TE subsets
 - Bitstring addresses links and BFERs in same subset
 - BFIR may be unable to reach some BFERs
 - BFER too far away from BFIR
 - Subset not sufficiently large to define path
 - Different approach is needed

- ▶ Contribution
 - Changes to BIER-TE subsets
 - Tunneling mechanism from BFIR to subsets
 - 1:1 protection for MC tree



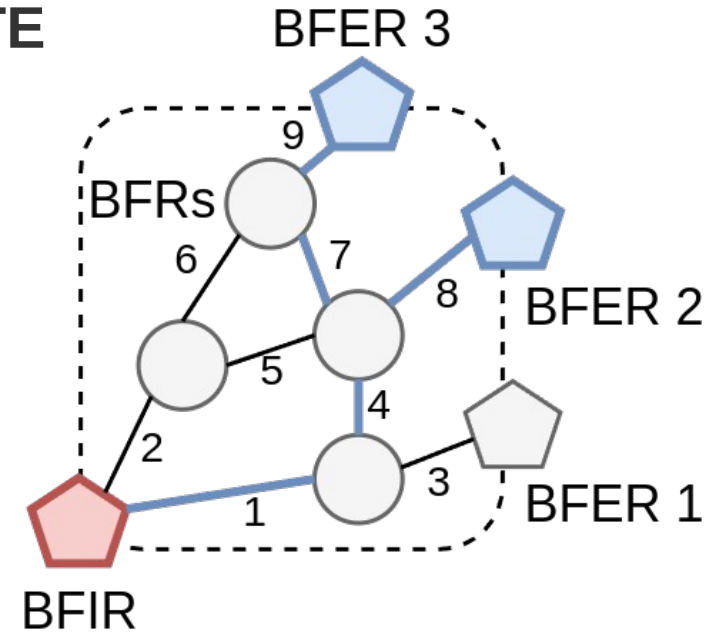


BIER



- ▶ Bitstring identifies destination BFERs
- ▶ Example: "110"
 - Reaches BFERs 2 and 3
- ▶ Multicast tree based on routing underlay
 - BFRs learn next hop for each BFER
 - Actual path depends on routing protocol

BIER-TE



- ▶ Bitstring identifies destination BFERs and links
- ▶ Example: "111001001110"
 - Reaches BFERs 2 and 3
 - Traverses links 1,4,7,8,9
- ▶ Multicast tree explicit in header
 - BFRs forward based on active bits of connected links

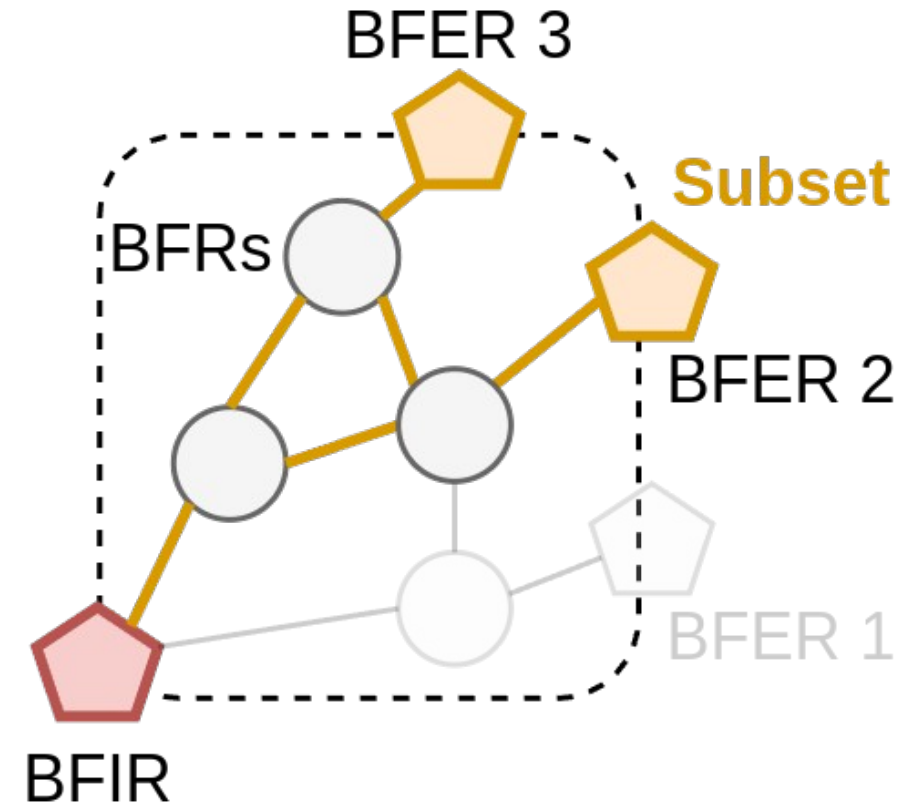


- ▶ Sub-topology of BIER-TE domain
 - Consists of BFERs and links
 - Size of subset must not exceed size of bitstring

- ▶ Reachability requirements
 - BFERs in subset have to be reachable from any link in subset
 - Packet can enter subset at any link and reach BFERs
 - Such is a “one-connected” subset

- ▶ Subset tunneling
 - BFIR may not be connected to a subset
 - Unicast tunnel to a BFR in subset
 - Any BFR that is connected to a subset link
 - Denoted as subset BFIR (S-BFIR)

- ▶ Requirements for 1:1 protection
 - At least two-connected subset required
 - In case of failure the remaining subset is still one-connected
 - Backup S-BFIR(s) for every S-BFIR in subset

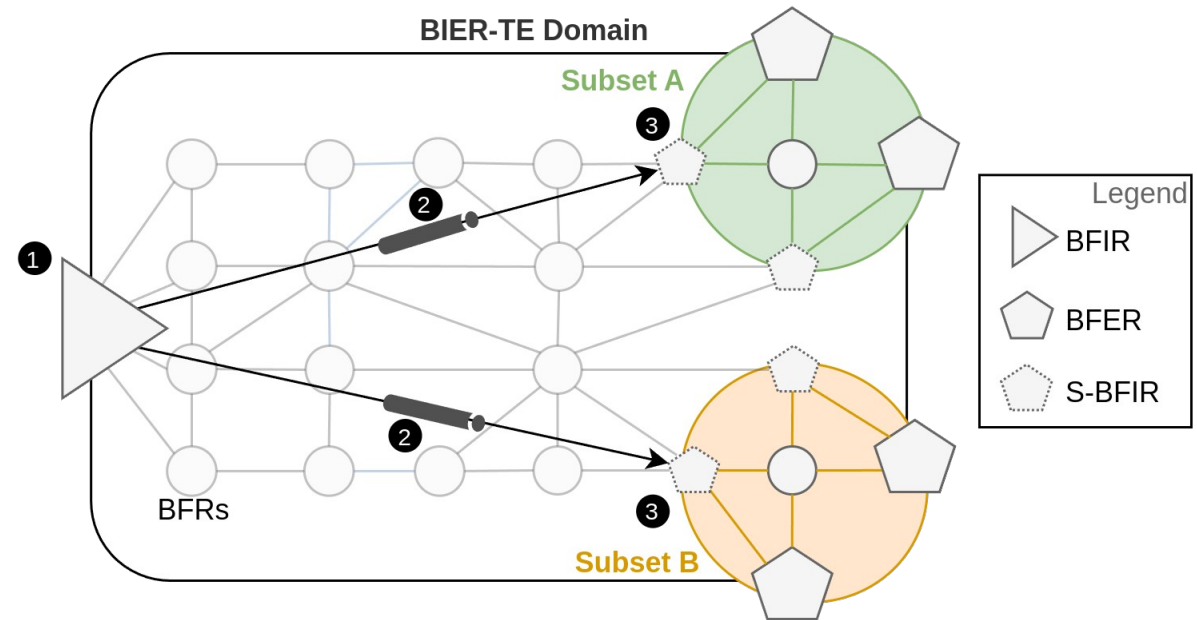




- ▶ Subset tunneling
 - BFIR tunnels packet to border of subset
 - ①. BFIR determines destination subsets
 - ②. Each packet copy is tunneled to an S-BFIR
 - ③. S-BFIR
 - 3. S-BFIR removes tunnel header and forwards packet
 - Standard BIER-TE forwarding inside subset

- ▶ Tunneling protocol
 - Requirements
 - Path steering
 - Egress protection
 - 1:1 FRR protection
 - E.g. MPLS TE, SRv6

- ▶ 1:1 protection cases
 - Tunnel to subset



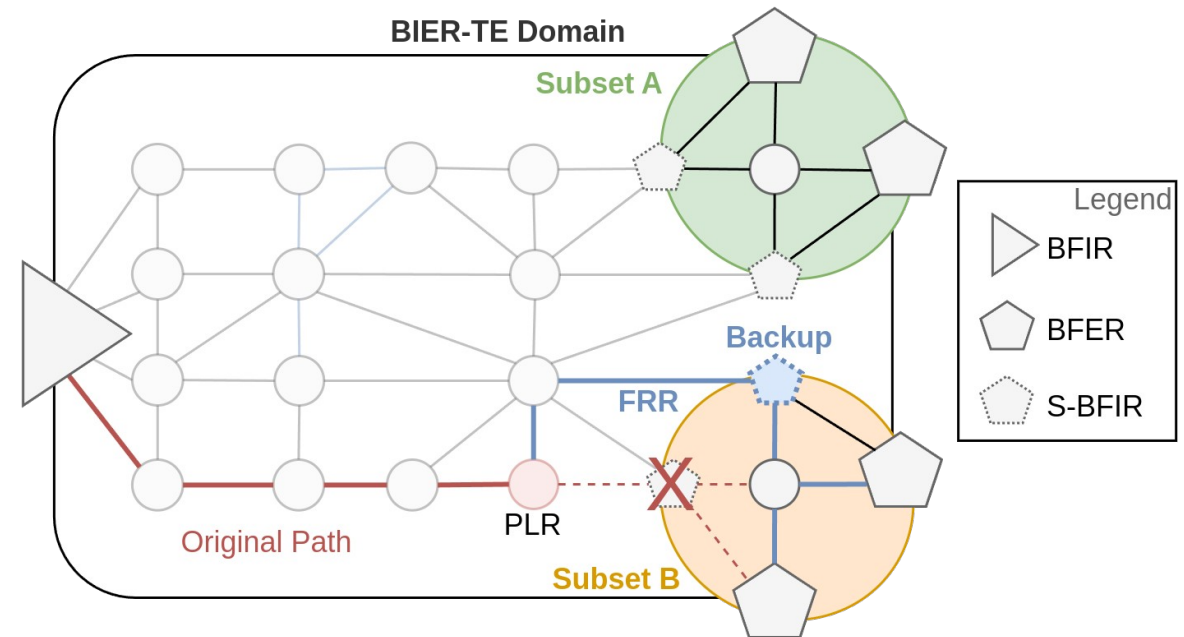
– Over FRR mechanisms e.g. MPLS TE FRR



- ▶ Failure of S-BFIR
 - Packet cannot be delivered to subset
 - Needs to be redirected to another S-BFIR

- ▶ Egress protection mechanism
 - PLR detects failed tunnel egress S-BFIR
 - Reroutes packet to backup S-BFIR

- ▶ Backup S-BFIR handling
 - Recognizes backup role from tunneling header
 - Applies BIER-TE FRR node protection
 - Failed node is primary S-BFIR
 - Reroute packet to NNH of primary S-BFIR
 - Modify BIER-TE header accordingly

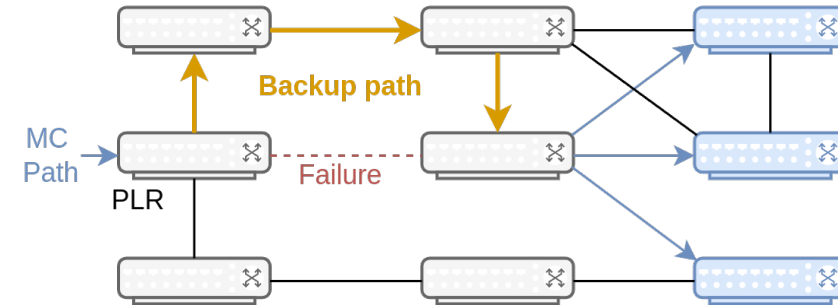




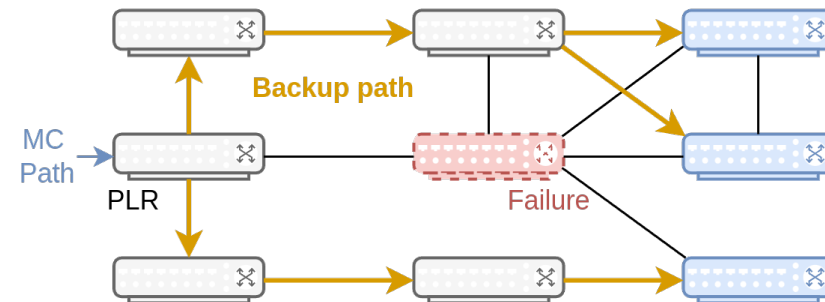
- ▶ 1:1 protection in subset over BIER-TE FRR
 - Required features
 - Link protection
 - Node protection
 - Drafts
 - Eckert et al. [ID.draft-eckert-bier-te-frr]
 - Chen et al. [ID.draft-chen-bier-te-frr]

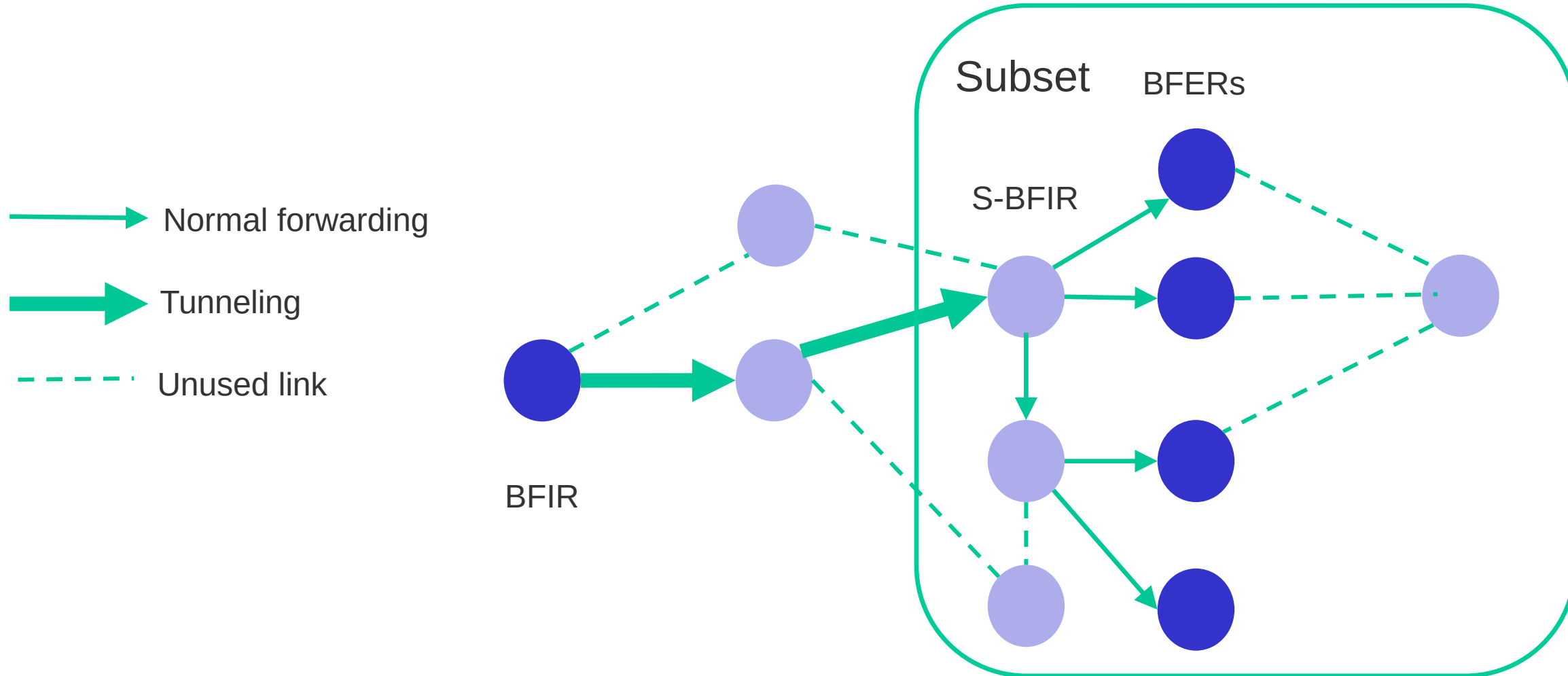
- ▶ Failure handling
 - Point of Local Repair (PLR) discovers failure
 - Changes to packet
 - Unset bits to prevent loops
 - Add backup path over BIER-TE
 - Example challenge: may require replication for FRR adjacency
 - Alternative – pre-set bits for backup
 - Avoids bitstring manipulation on PLR
 - ~~Not sure if/what generic rules can be to do this.~~

Link failure



Node failure

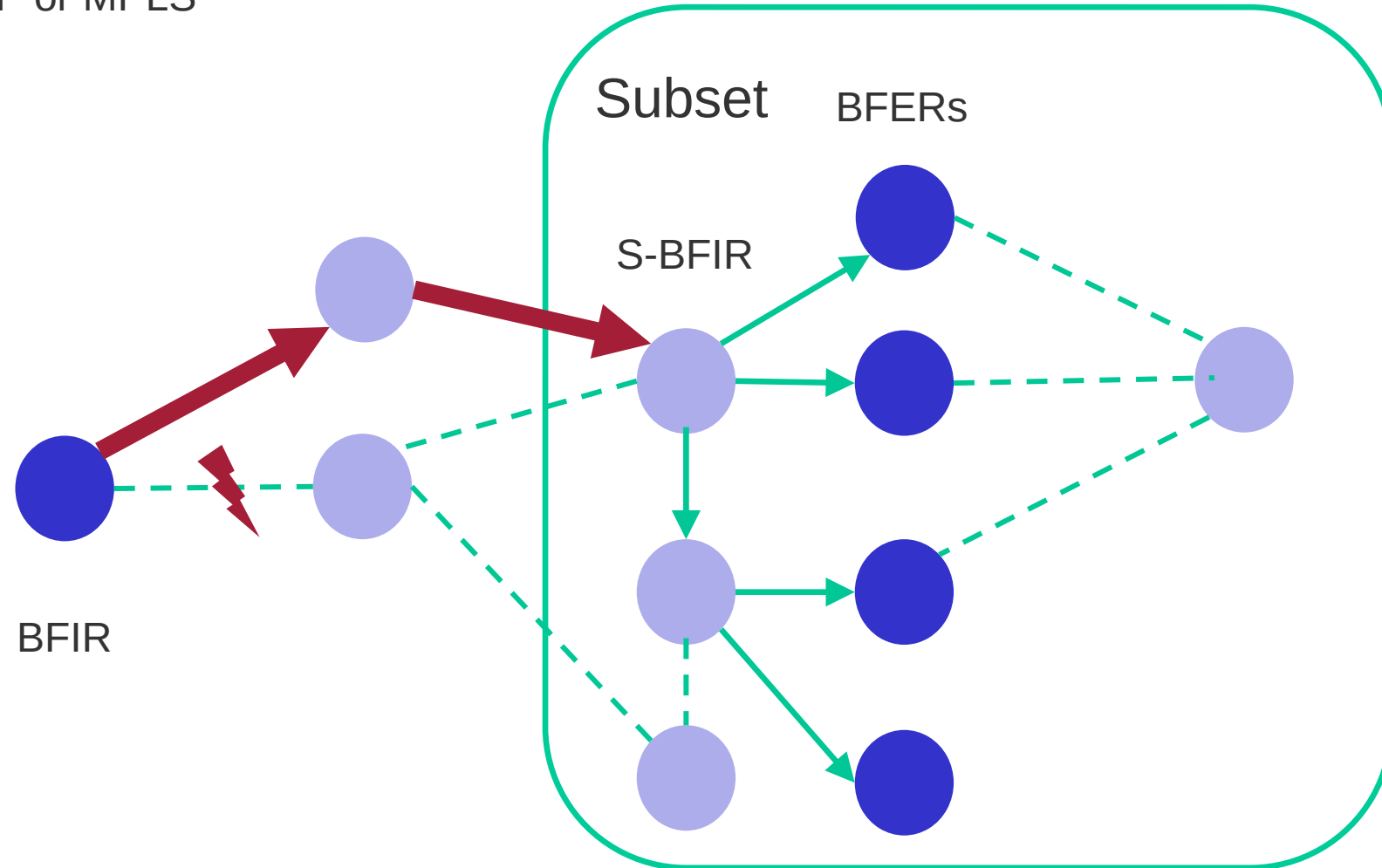
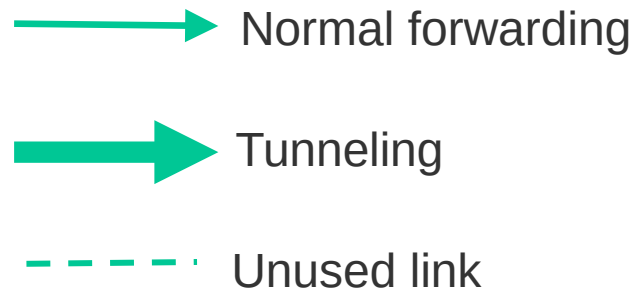






Failure outside the Subset

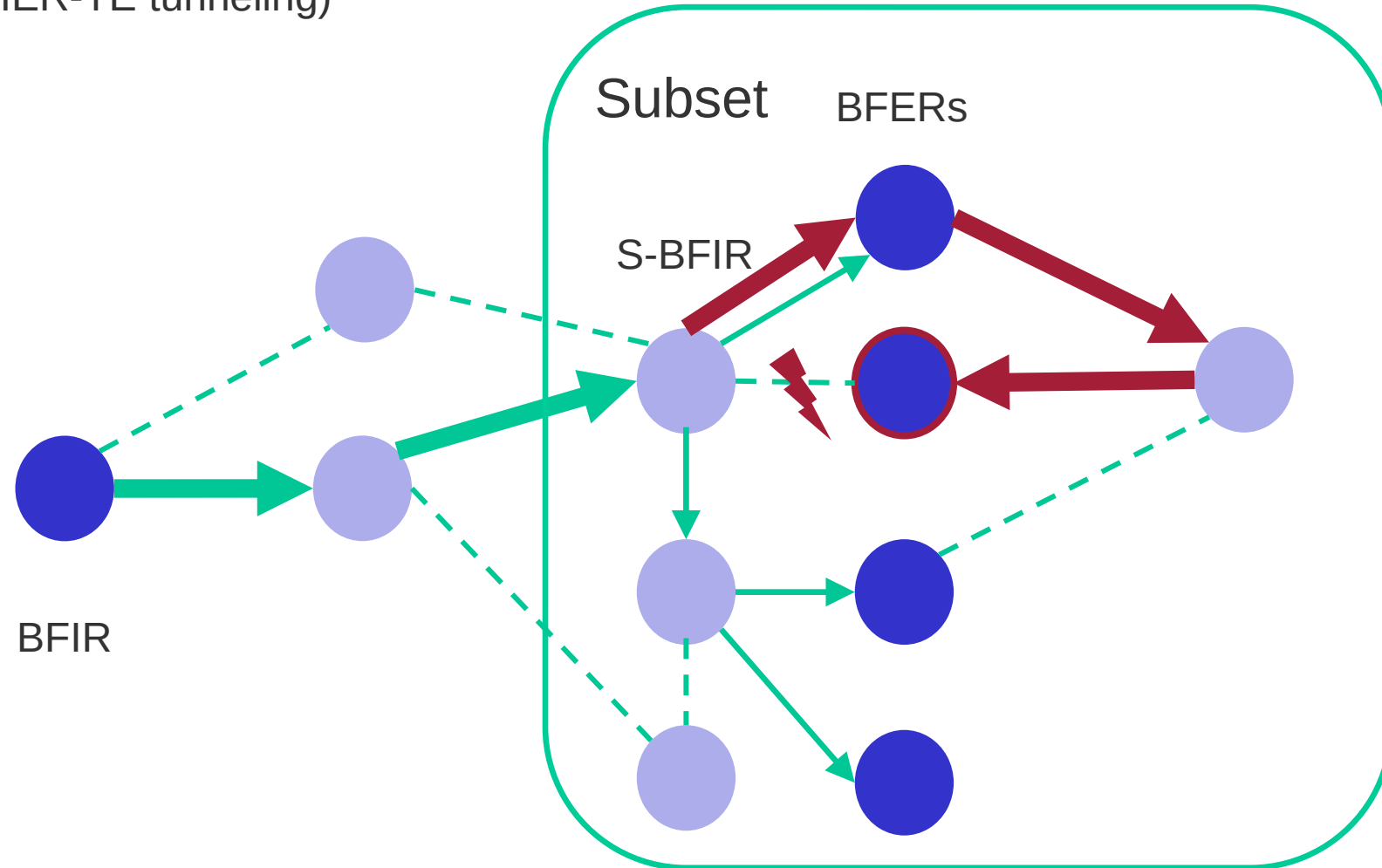
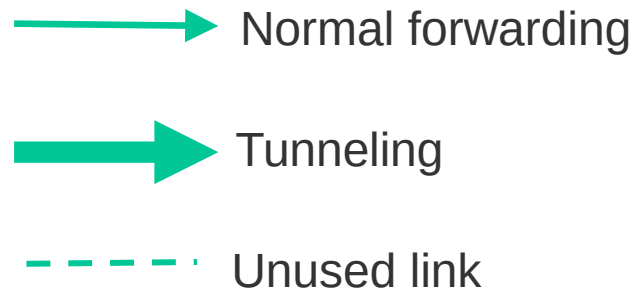
- ▶ Use some FRR method, e.g., for IP or MPLS





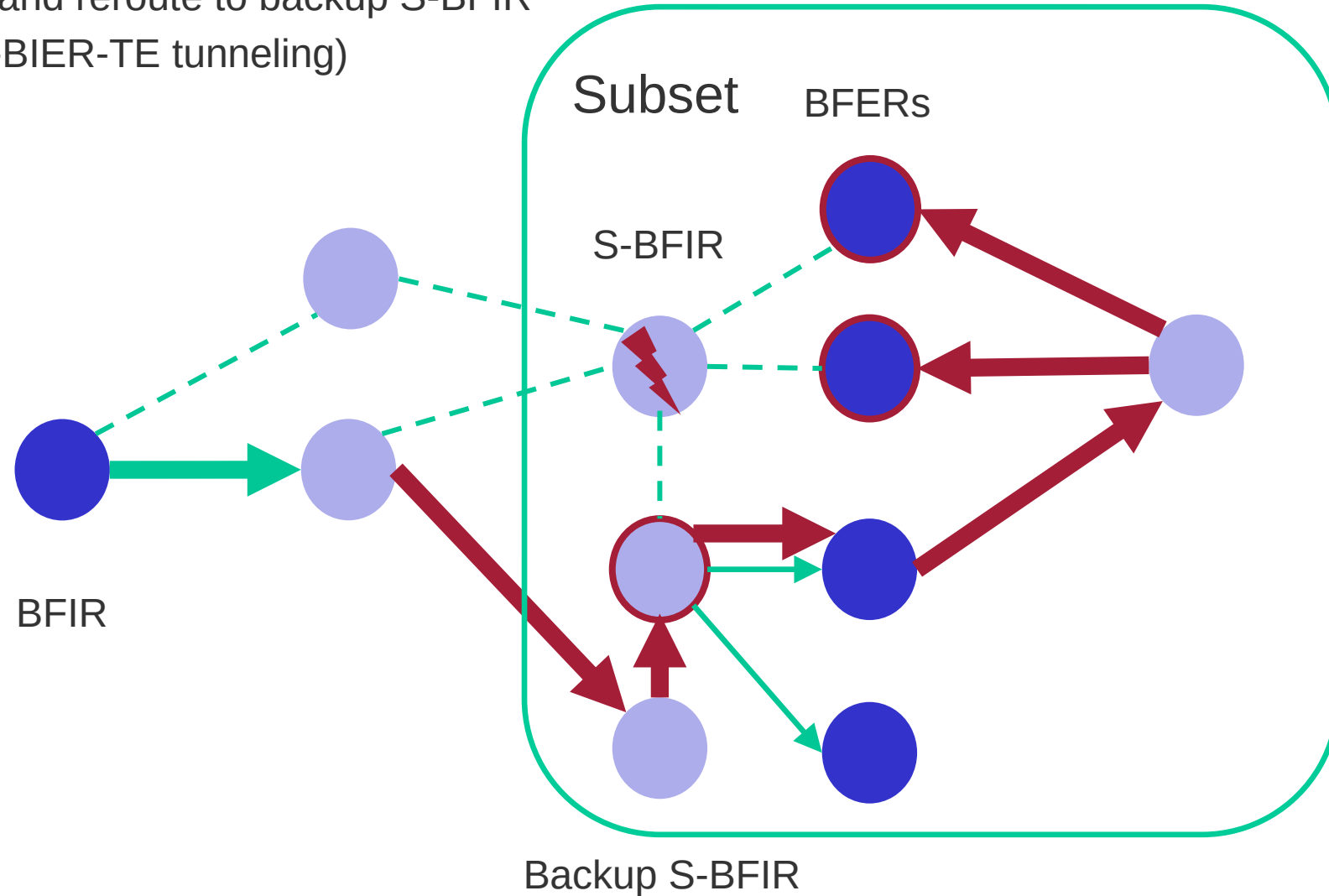
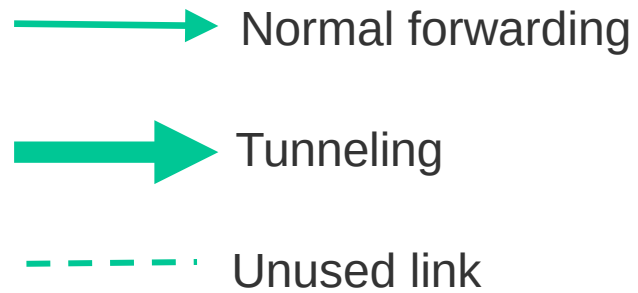
Failure within the Subset

- Use BIER-TE-FRR (BIER-TE-in-BIER-TE tunneling)





- ▶ Use egress protection for S-BFIR and reroute to backup S-BFIR
- ▶ Apply BIER-TE-FRR (BIER-TE-in-BIER-TE tunneling)





- ▶ All main components of concepts validated through P4/Tofino HW implementation
 - BFIR, PLR, S-BFIR, BIER-TE-FRR
 - Unpolished pre-report: <https://arxiv.org/pdf/2409.07082>
 - No recirculation required for unicast-TE parts!
 - BIER-TE, BIER-TE-FRR recirculation required can be minimized “fabric” multicast groups

- ▶ We think this is a core spec to make BIER-TE operationally deployable large-scale

- ▶ Specifies required forwarding features
 - Overlay-lookup (BFIR), SR/FRR lookup based replication into unicast TE-tunnel
 - Bitwise and/or operation on bitstring on PLR in FRR case

- ▶ Solution architecture
 - Combines BIER-TE, SR unicast + replication -> BIER-WG, PIM-WG, MBONED-WG ? *sigh*



Questions?

THANK YOU