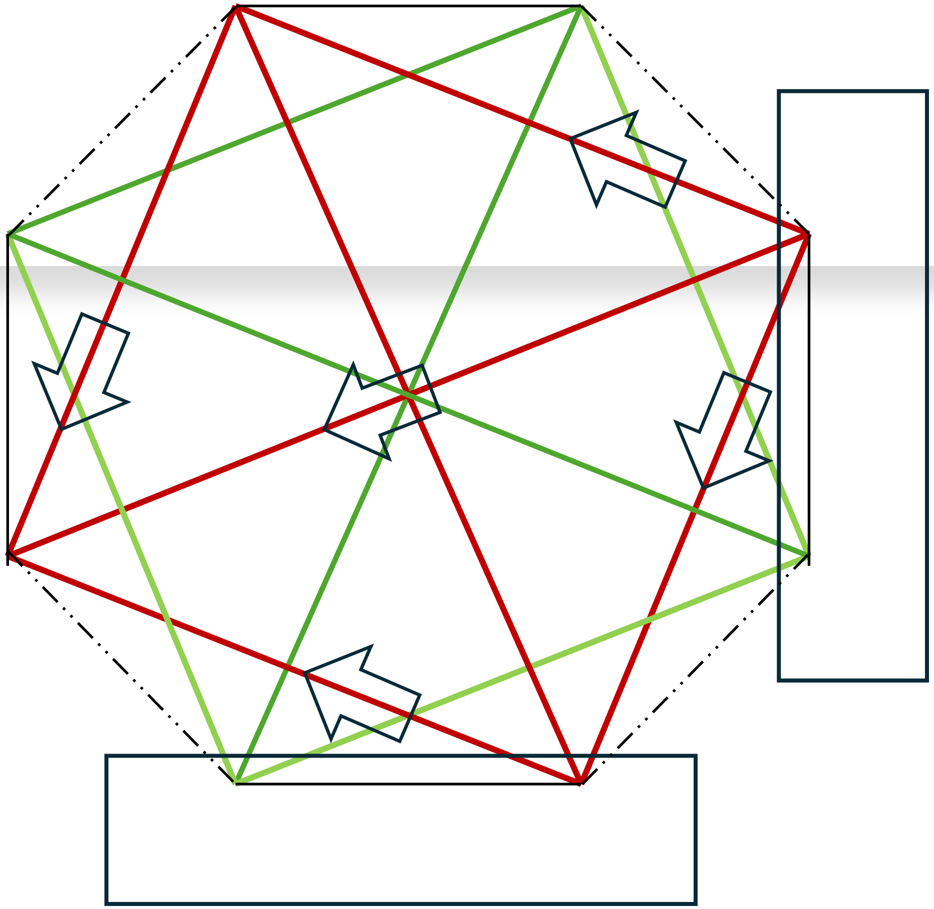
Further considerations for RIFT dragonfly(+) routing

IETF 123

Tony P, HPE of former Juniper

Original Proposal

- Allows for at most one additional hop alternate routing
- Implies that for each dragon fly node there exist either a direct path or one hop alternate to reach any other node
- Any standard silicone will work since the split horizon is built based on packet arrival direction
  - arriving into the dragonfly or
  - being already in the dragonfly

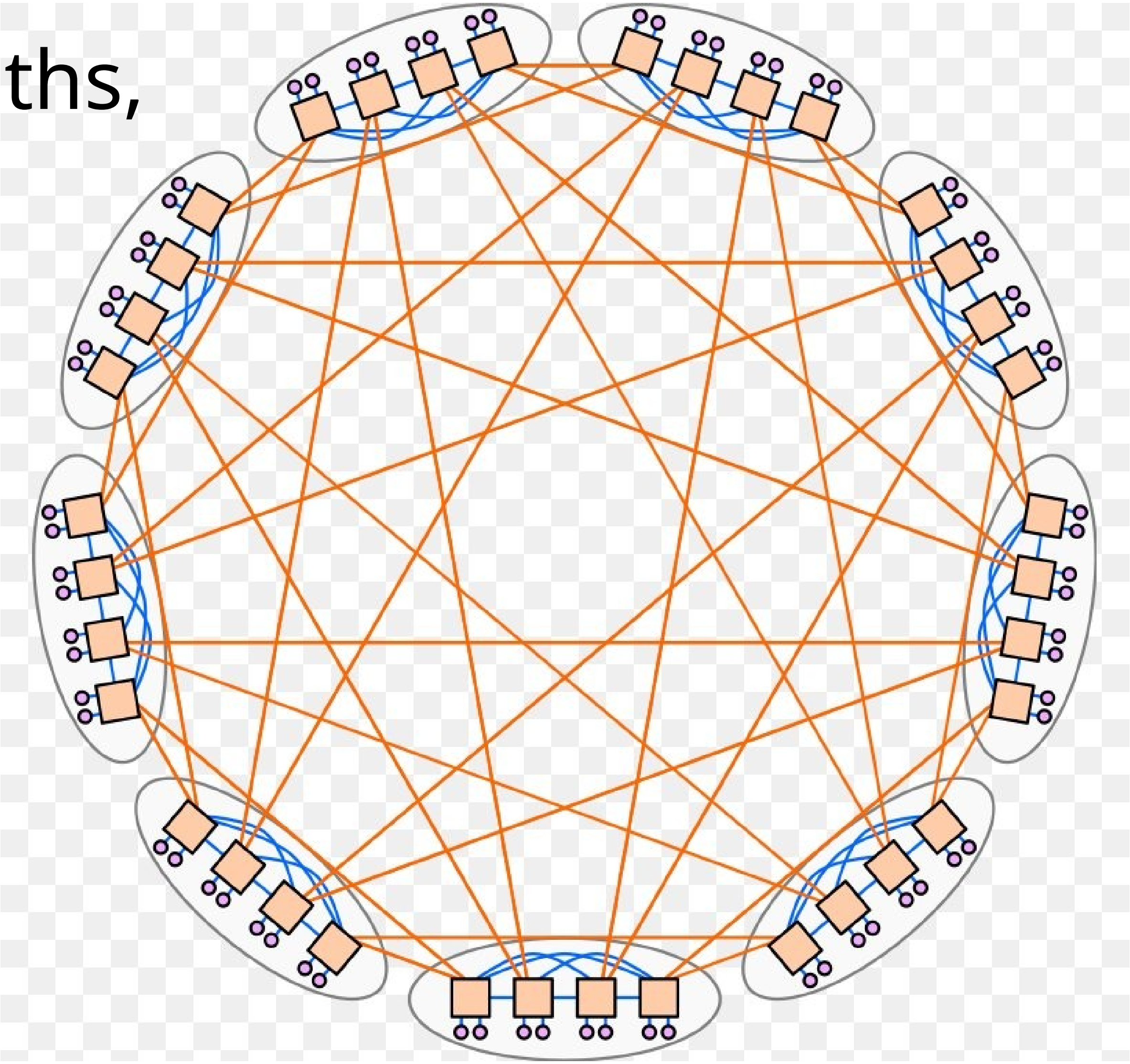


Fun Intermezzo

- Ask AI for “picture of dragonfly+ topology with 4 planes?”
- Get this ;-)
- Looks like our overlords are still dumb as proverbial rocks when it comes to anything non-trivial

Massive Failures, Longer Paths, Funny Variants

- Dragonfly can have tons planes, funny variants, we deal with CLOS variant connected by DF+ in RIFT only
- Link failures may make 2-hop max. alternate infeasible
- Even then “blue is south”, “orange is north” albeit this would be 4 ToFs poorly connected



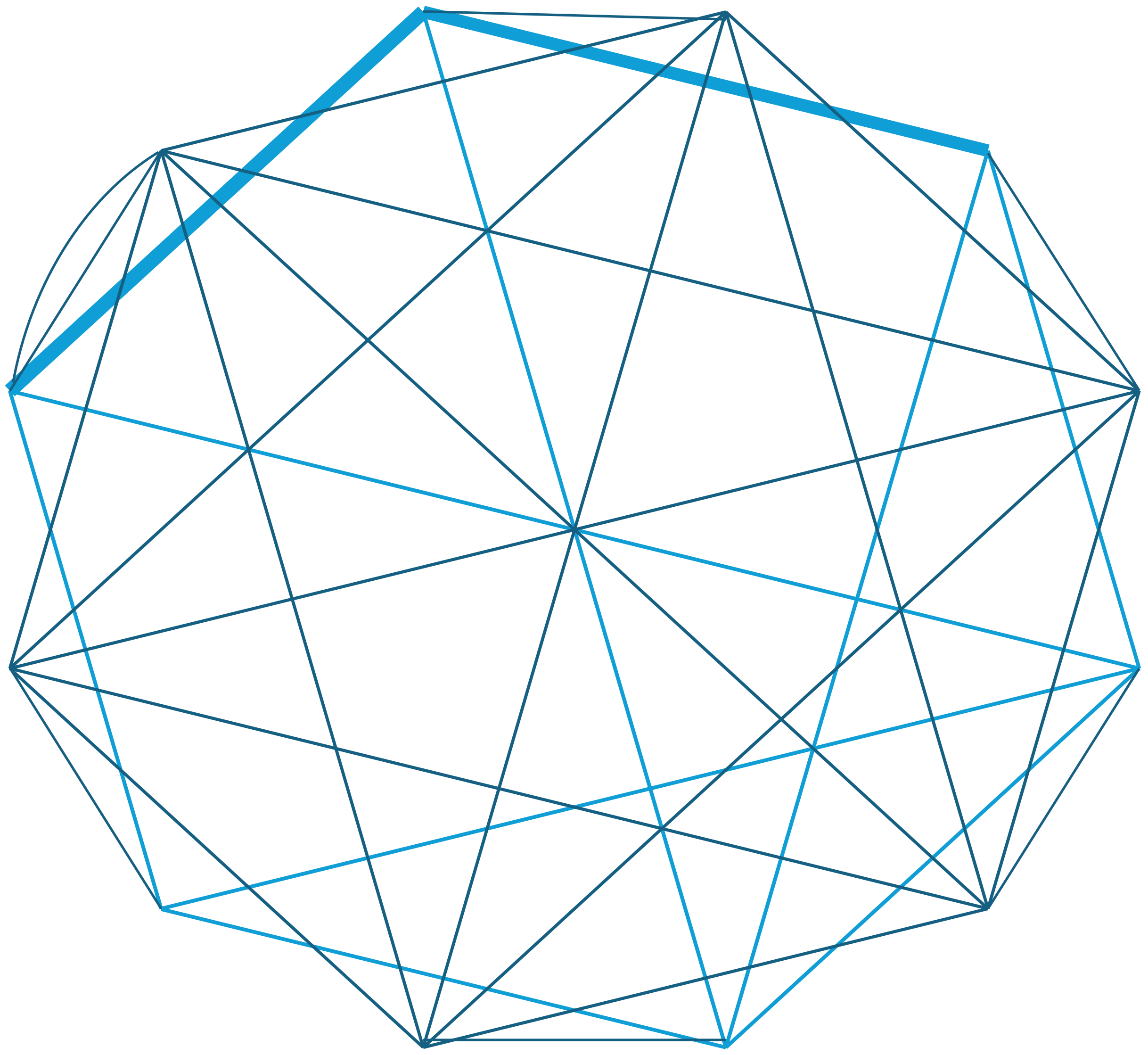
Kim, John, ACM SigArch, 08

Any Number of Alternates

- Originally 2 FIBs allowing 1 alternate nexthop
  - Packet from south interfaces either forwarded south or hit northbound
  - Packet from north either matches prefix visible in the fabric and is “sunk south” or FIB MUST have a direct next-hop match to destination that can ”sink south”
- New proposal differentiates FIB _per_ northbound interface
  - Compute SPF (or whatever agreed on computation is used that delivers a shortest spanning) from the view of the sender (we are on p2p links)
  - Use only the results where the path becomes strictly shorter (loop prevention in case of spanning rather than shortest spanning tree)
  - Discard any routes going through the originator (i.e. dragonfly leaf originating the source prefix). Multicast is interesting ;-)

Example

- Light blue plane
  - Links failed or funny dragonfly variant
- A to F loop-free paths
  - A-F, A-D-F, A-D-E-F, A-B-E-F, A-B-C-D-F, A-B-C-E-F
- FIB for C on interface B-C for A->F
  - DST in F -> D or E



Properties

- No source specific routing necessary
- Source can use northbound loop free path of any length
- Few calculations necessary
- Rather than one FIB for all northbound a FIB per northbound interface