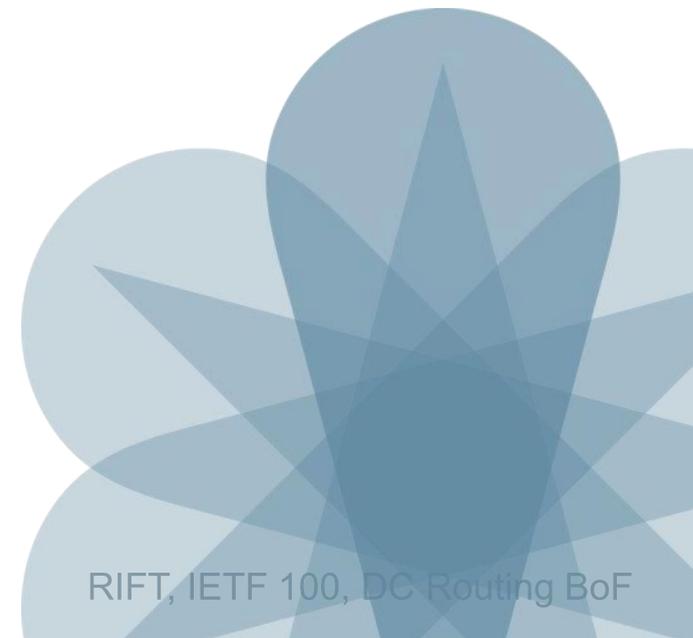


# RIFT: ZERO OPEX ROUTING PROTOCOL FOR IP FABRICS

TONY PRZYGIENDA PRZ@JUNIPER.NET

DRAFT-PRZYGIENDA-RIFT @ IETF



# RIFT VS. DRAFT-DT-RTGWG-DCROUTING-REQUIREMENTS

Problem / Attempted Solution	Vs. draft-dt-rtgwg-dcrouing-requirements
01. As Close to Zero Necessary Configuration as Possible (Contradicts 02)	✓ (next revision)
02. Peer Discovery/Automatic Forming of Trees/Preventing Cabling Violations (Contradicts 01)	+
03. Minimal Amount of Routes/Information on ToRs	+
04. High Degree of ECMP (BGP needs lots knobs, memory, own-AS-path violations) and ideally NEC and LFA	✓
05. Traffic Engineering by Next-Hops, Prefix Modifications	✓
06. See All Links in Topology to Support PCE/SR	✓
07. Carry Opaque Configuration Data (Key-Value) Efficiently	✓
08. Take a Node out of Production Quickly and Without Disruption	✓ (do we need GR?)
09. Automatic Disaggregation on Failures to Prevent Black-Holing and Back-Hauling	+
10. Minimal Blast Radius on Failures (On Failure Smallest Possible Part of the Network "Shakes")	+
11. Fastest Possible Convergence on Failures	✓

# RIFT: ZERO OPEX ROUTING ALGORITHM FOR CLOS UNDERLAY

---

- GENERAL CONCEPT
- AUTOMATIC DISAGGREGATION
- OPTIONAL HORIZONTAL LINKS
- AND MORE BEYOND THAT

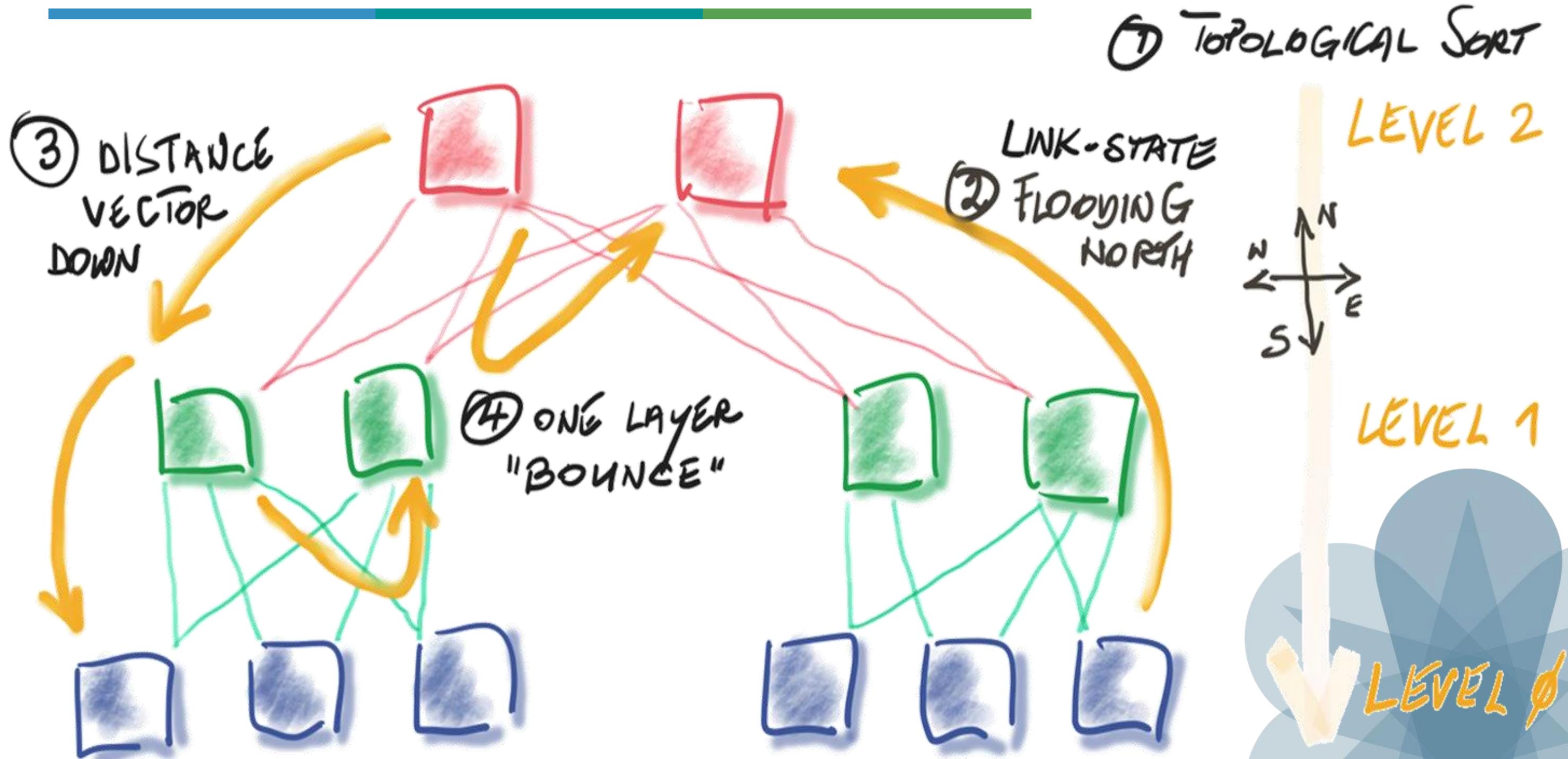
BUT IT'S SO NEW ...

*"Man cannot discover new oceans unless he has the courage to lose sight of the shore." --- Andre Gide*

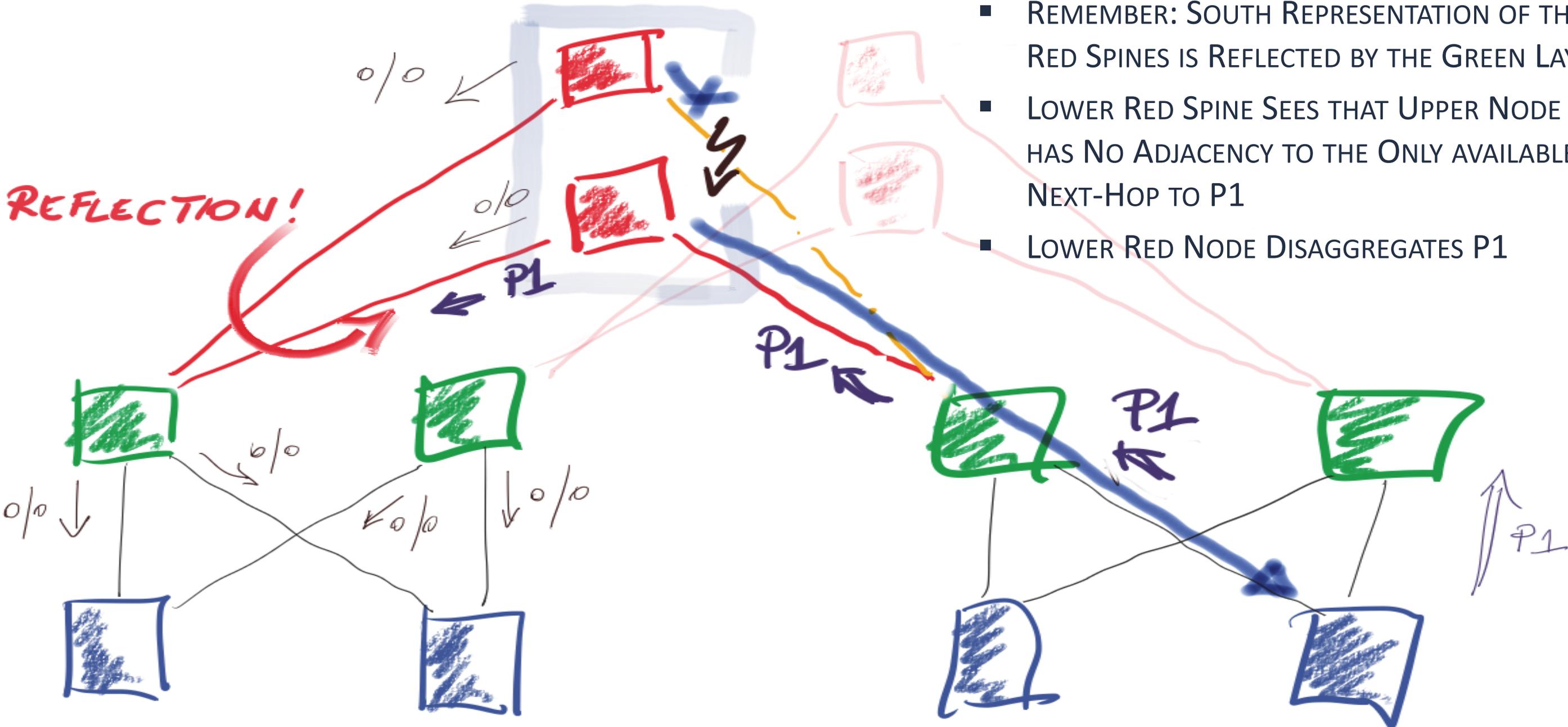
WELL, YOU MUST BE ...

*"The reasonable man adapts himself to the world: the unreasonable one persists in trying to adapt the world to himself. Therefore all progress depends on the unreasonable man." --- Bernard Shaw*

# LINK-STATE UP, DISTANCE VECTOR DOWN & BOUNCE



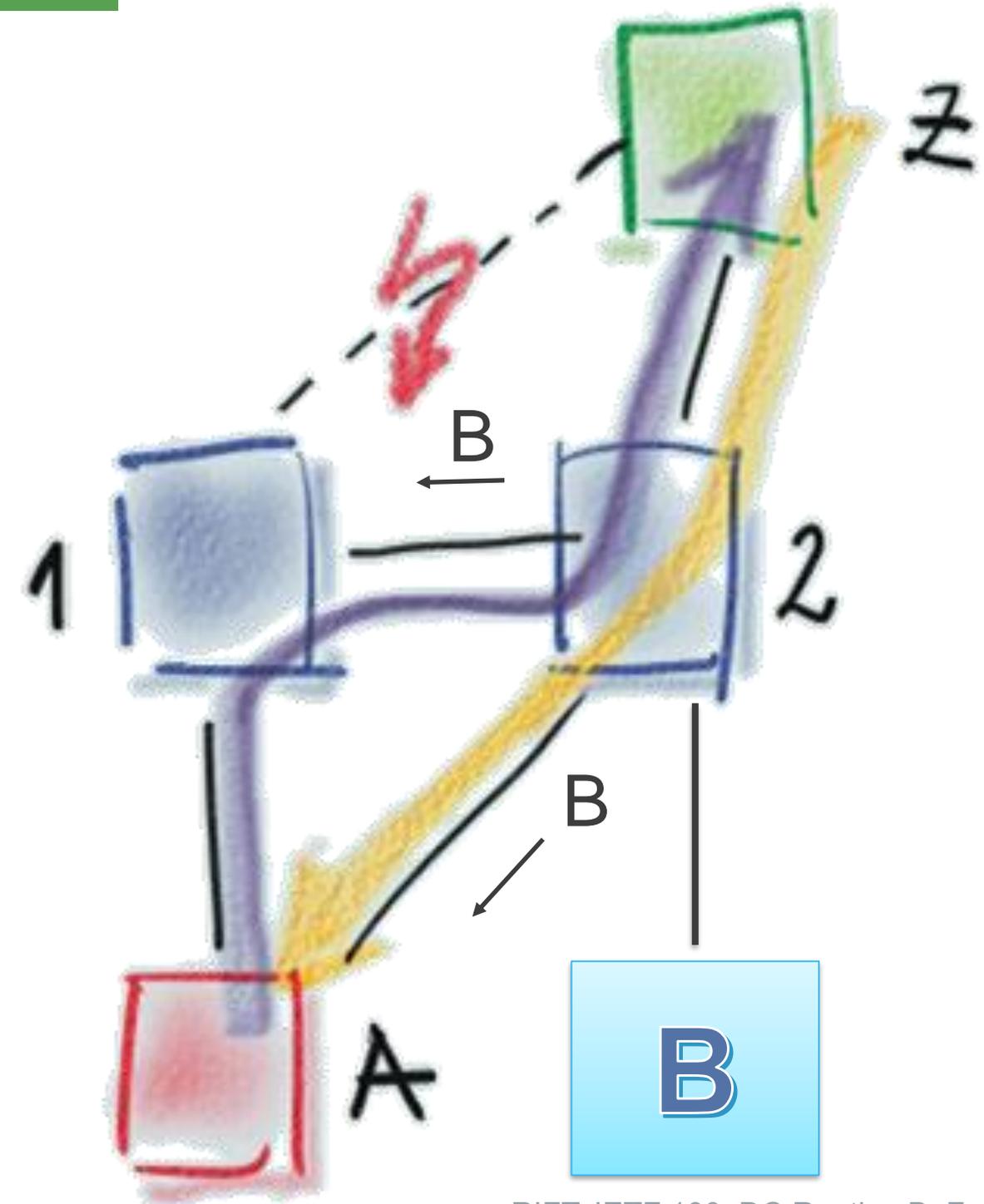
# AUTOMATIC DISAGGREGATION



- REMEMBER: SOUTH REPRESENTATION OF THE RED SPINES IS REFLECTED BY THE GREEN LAYER
- LOWER RED SPINE SEES THAT UPPER NODE HAS NO ADJACENCY TO THE ONLY AVAILABLE NEXT-HOP TO P1
- LOWER RED NODE DISAGGREGATES P1

# OPTIONAL HORIZONTAL LINKS FOR FAILURE PROTECTION

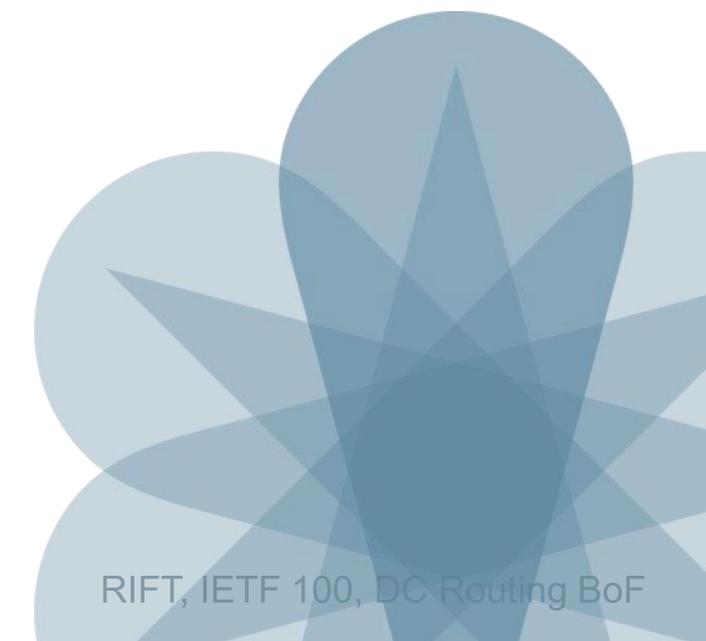
- LEVELS CAN INSTALL OPTIONAL HORIZONTAL LINKS
- LEVEL 0 IS SPECIAL:
  - LEAF-2-LEAF CONNECTION THAT CANNOT BE USED EXCEPT FOR LEAF-2-LEAF TRAFFIC
- LEVEL > 0 USES HORIZONTAL LINKS FOR FAILURE PROTECTION ONLY
  - SINGLE NODE PROTECTION: NODE THAT LOST NORTHBOUND LINKS BUT HAS NEIGHBORS THAT CAN REACH HIGHER LAYERS USES THE HORIZONTAL LINK
  - N:N-1 PROTECTION: FULL MESH IN A LEVEL CAN PROVIDE UP TO N-2 NORTHBOUND PROTECTION
  - HORIZONTAL DISAGGREGATION CAN HEAL COMPLEX FAILURES (NOT DIFFERENT FROM SOUTHBOUND DISAGGREGATION)



# RIFT DOES ON TOP

---

- AUTOMATIC FLOOD REDUCTION
- LEAF-TO-LEAF BI-DIRECTIONAL SHORTCUTS
- POSSIBLE TRAFFIC ENGINEERING VIA “FLOODED DV OVERLAY” WITH POLICIES
- COMPLETELY MODEL BASED PACKET FORMATS
- CHANNEL AGNOSTIC DELIVERY, COULD BE QUICK, TCP, UDP
- PREFIXES TO TOPOLOGY ELEMENT MAPPING BASED ON HASH FUNCTIONS LOCAL TO EACH NODE
  - ONE EXTREME POINT IS PREFIX PER FLOODED ELEMENT = BGP UPDATE
- PURGING (GIVEN COMPLEXITY) IS OMITTED
- POLICY CONTROLLED KEY-VALUE STORE SUPPORT



# SUMMARY OF RIFT ADVANTAGES COMPARED TO DV OR LS MODIFICATIONS

---

## ■ ADVANTAGES OF BOTH LINK-STATE AND DISTANCE VECTOR

- FASTEST POSSIBLE CONVERGENCE
- AUTOMATIC DETECTION OF TOPOLOGY
- MINIMAL ROUTES ON TORs
- HIGH DEGREE OF ECMP
- FAST DE-COMMISSIONING OF NODES
- MAXIMUM PROPAGATION SPEED WITH FLEXIBLE # PREFIXES IN AN UPDATE

## ■ NO DISADVANTAGES OF NEITHER LINK-STATE NOR DISTANCE VECTOR

- REDUCED FLOODING
- AUTOMATIC NEIGHBOR DETECTION

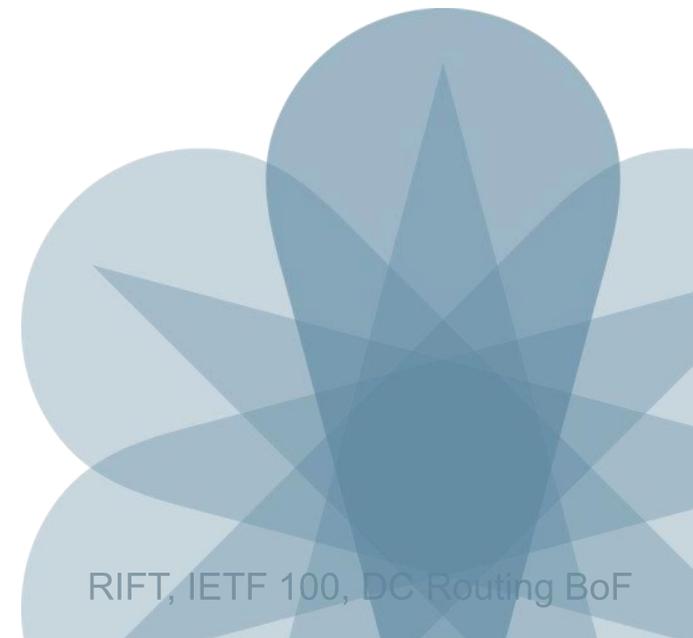
## ■ UNIQUE RIFT ADVANTAGES

- AUTOMATIC DISAGGREGATION ON FAILURES
- KEY-VALUE STORE
- HORIZONTAL LINKS USED FOR PROTECTION
- MINIMAL BLAST RADIUS ON FAILURES
- CAN UTILIZE **ALL** PATHS THROUGH FABRIC WITHOUT LOOPING

# IS THERE MORE THAN BITS OVER POWERPOINT ?

---

- YES ;-)



THANK YOU FOR YOUR ATTENTION

