



JUNIPER
NETWORKS

BETTER IP MULTICAST ARCHITECTURE BREWING: BIER

KRZYSZTOF GRZEGORZ SZARKOWICZ

JUNIPER NETWORKS

DISCLAIMERS AND EXPECTATIONS

- THIS IS A INTRODUCTORY MATERIAL BASED ON CURRENT OPEN STANDARDS
- NONE OF THOSE THINGS CONSTITUTE COMMITMENTS TO PRODUCT SPECIFICATIONS, OFFERINGS OR RELEASE DATES BY JUNIPER AT THIS POINT IN TIME

IETF

- CURRENTLY BEING RE-CHARTERED AS "STANDARDS TRACK"
- ARCHITECTURE AND ENCAPSULATION
 - RFC 8279: MULTICAST USING BIT INDEX EXPLICIT REPLICATION (BIER)
 - RFC 8296: ENCAPSULATION FOR BIER IN MPLS AND NON-MPLS NETWORKS
- IGP DRAFTS IN SHEPHERD WRITE-DOWNS, LC PASSED, CONSENSUS BUILT
- MVPN/EVPN OVERLAY DRAFTS MATURE AND LC'ED
- SEVERAL SILICON VENDORS SUPPORT TODAY OR NEXT GENERATION PRODUCT MAP
- SEVERAL VENDORS IMPLEMENTED CONTROL PLANE
- HEALTHY INTEREST FROM SPs AND A ENTERPRISE VERTICAL

AGENDA

- OVERVIEW
- BIER: UNDERLAY CONCEPTS
- BIER: OVERLAY INTEGRATION
- BIER: ADVANCED ASPECTS

OVERVIEW

- CURRENT MULTICAST IN THE CORE
- NUCLEUS OF BIER IDEA
- FIRST AND SECOND OBVIOUS AND WRONG SOLUTIONS

CURRENT MULTICAST IN THE CORE

- STATE IN THE CORE ROUTERS SCALES AT BEST WITH #VPNS
 - HARD TO OFFER LARGE-SCALE COMPETITIVE MVPN SERVICE
- SLOW AND “NERVOUS” PIM CONVERGENCE COMPARED TO UNICAST
 - PIM TUNNELS BASED ON DIFFERENT TRIGGERS
 - ONLY TRADE-OFF BETWEEN MORE STATE (S-PMSI) AND EXCESSIVE REPLICATION (I-PMSI) POSSIBLE
- MULTIPLE P2MP TUNNEL SIGNALING TECHNOLOGIES WITH DIFFERENT STATE AND PROTECTION PROPERTIES
 - MLDP LSPs
 - P2MP RSVP
- ASM AND SSM ASSYMETRY
- ROOT NODE PLACEMENT PROBLEMS FOR PIM AND MP2MP LDP
- CONCEPTUALLY INTERNET SCALES BECAUSE “EDGE STATE SHOULD NOT SPILL INTO THE CORE” AND CURRENT MULTICAST VIOLATES THAT AXIOM

NUCLEUS OF BIER IDEAS

- MULTICAST IN THE CORE SHOULD
 - CONVERGE AT SPEED AND STABILITY OF UNICAST
 - PROVIDE "STATE COMPRESSION" LIKE A "LONGEST PREFIX LOOKUP"
 - ALLOW FOR MP2MP S-PMSI AND I-PMSI AT THE SAME "STATE COST"
 - NOT NEED UPSTREAM LABELS
 - NOT NEED RPF
 - NOT FORM LOOPS
 - MORE "MULTICAST SERVICE" AT THE EDGE NOT CAUSE MORE STATE IN THE CORE
 - BE EASY TO DEBUG: BUILT-IN OAM
 - ALLOW FOR "SERVICE SEPARATION" IF SO DESIRED BASED ON E.G.
 - SERVICE: MVPN, EVPN, "PLAIN" SELECTIVE MP2MP DISTRIBUTION
 - CUSTOMER OR FLOW
- MAYBE UNICAST SHOULD BE EVEN A SPECIAL CASE OF MULTICAST

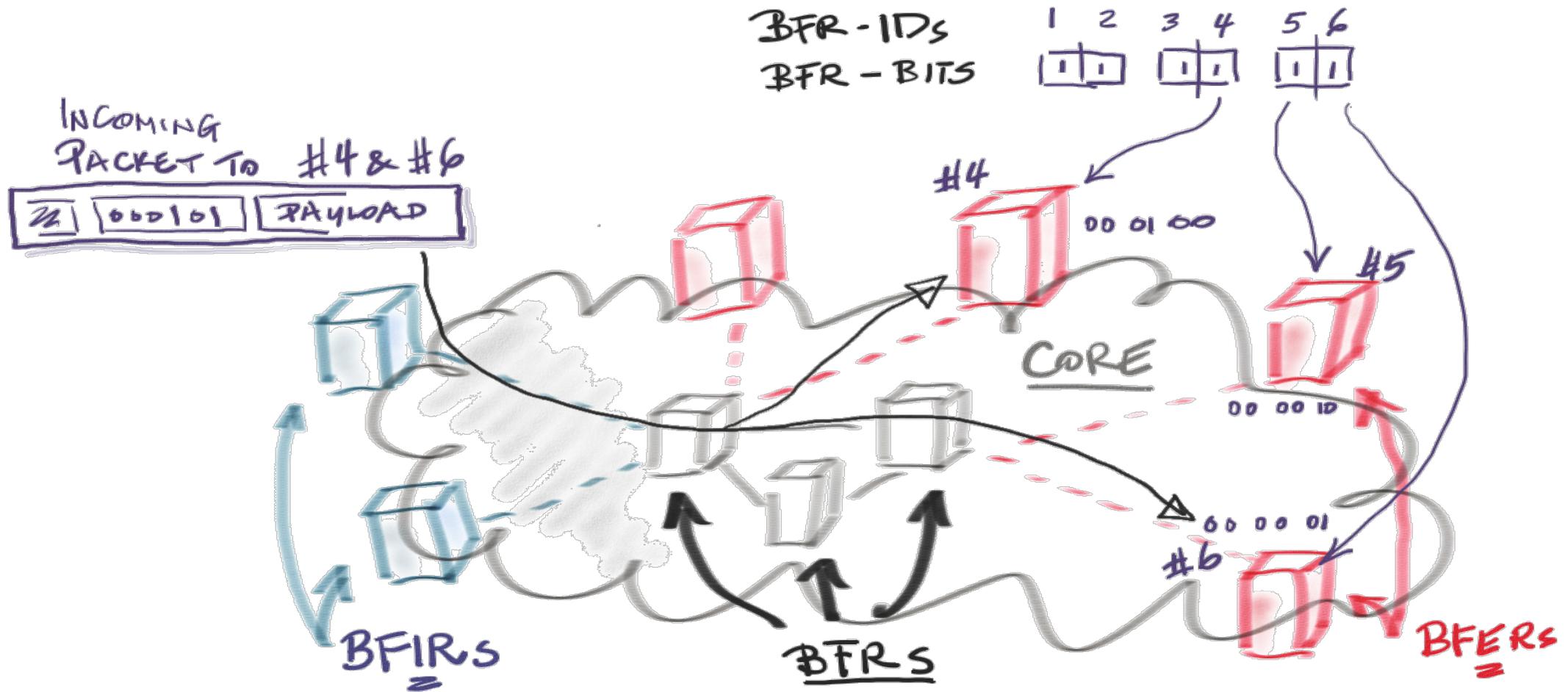
IT'S EASY: PUT STATE ON THE PACKET

- FIRST IDEA: PUT SET OR RECEIVERS ON EACH PACKET
 - INCREDIBLY POOR EFFICIENCY
 - INCREDIBLY POOR SCALE (128 BYTES HEADER IS ONLY ~ 30 RECEIVERS)
 - REMOVING RECEIVERS "IN FLIGHT" IS BAD FOR SPEED OR OTHERWISE LOOPING IS VERY HARD TO CONTROL
- SECOND IDEA: PUT STACK OF RECEIVER LABELS ON EACH PACKET
 - SAME INCREDIBLY POOR SCALE (4 BYTES IS 4 BYTES)
- STILL NO OAM
- EXTRA RAT-HOLE:
 - BUILD HIERARCHICAL MULTICAST TREES (ATM DID IT)
- WELL: THE **SMALLEST WAY** TO INDICATE ABSENCE OR PRESENCE IS ONE BIT

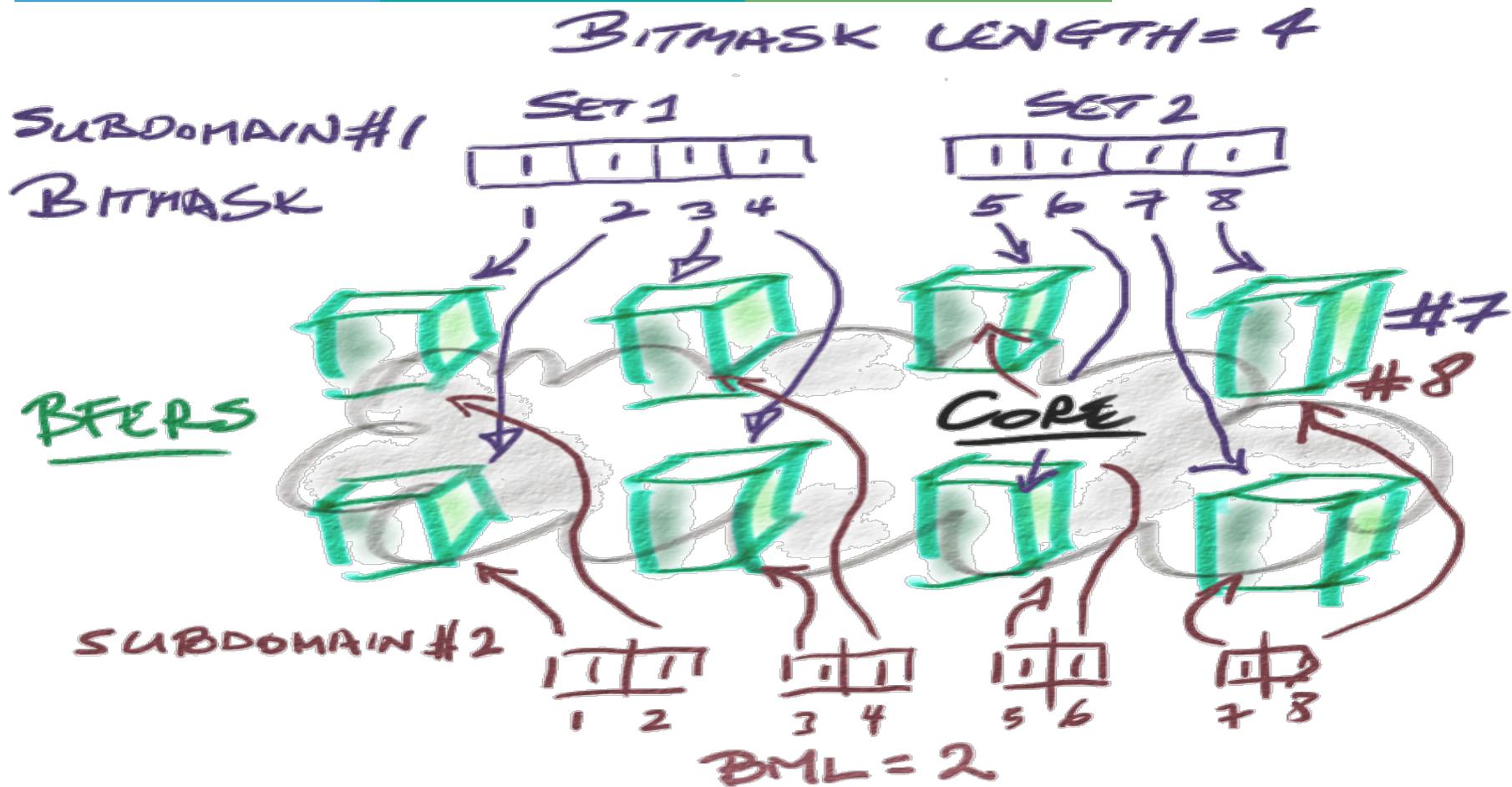
BIT INDEX EXPLICIT REPLICATION (BIER) UNDERLAY

- BIER FORWARDING ROUTERS (BFR)
 - BIER FORWARDING INGRESS ROUTERS (BFIR)
 - BIER FORWARDING EGRESS ROUTERS (BFER)
- BIER BITMASK, BITMASK LENGTH (BML) AND SETS
- BIER SUBDOMAINS
- BIT INDEX FORWARDING TABLES (BIFT)

BFRs, BFIRs, BFERs & BIER BITMASKS

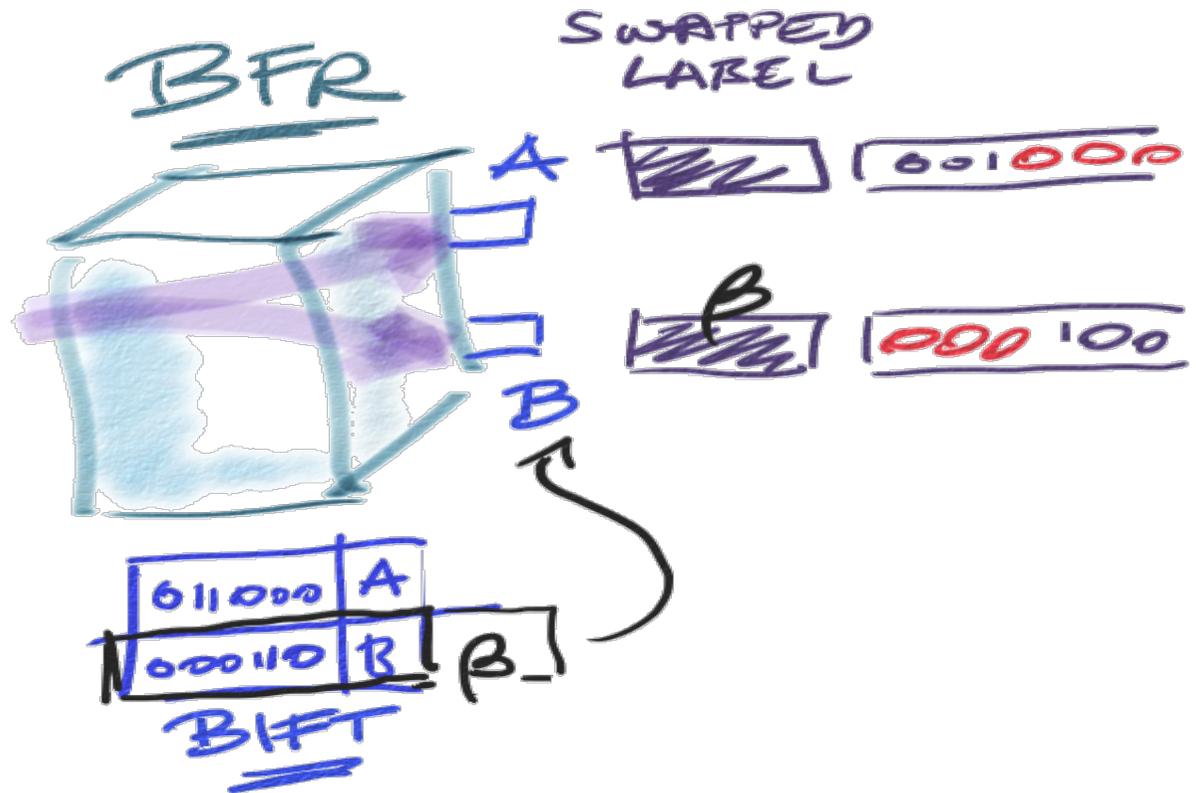
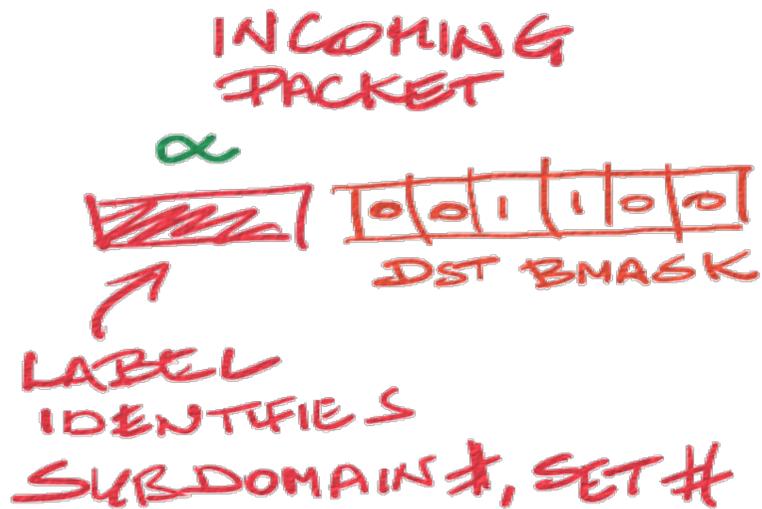


SUBDOMAINS, BITMASK LENGTH AND SETS



BIER FORWARDING & BIFT

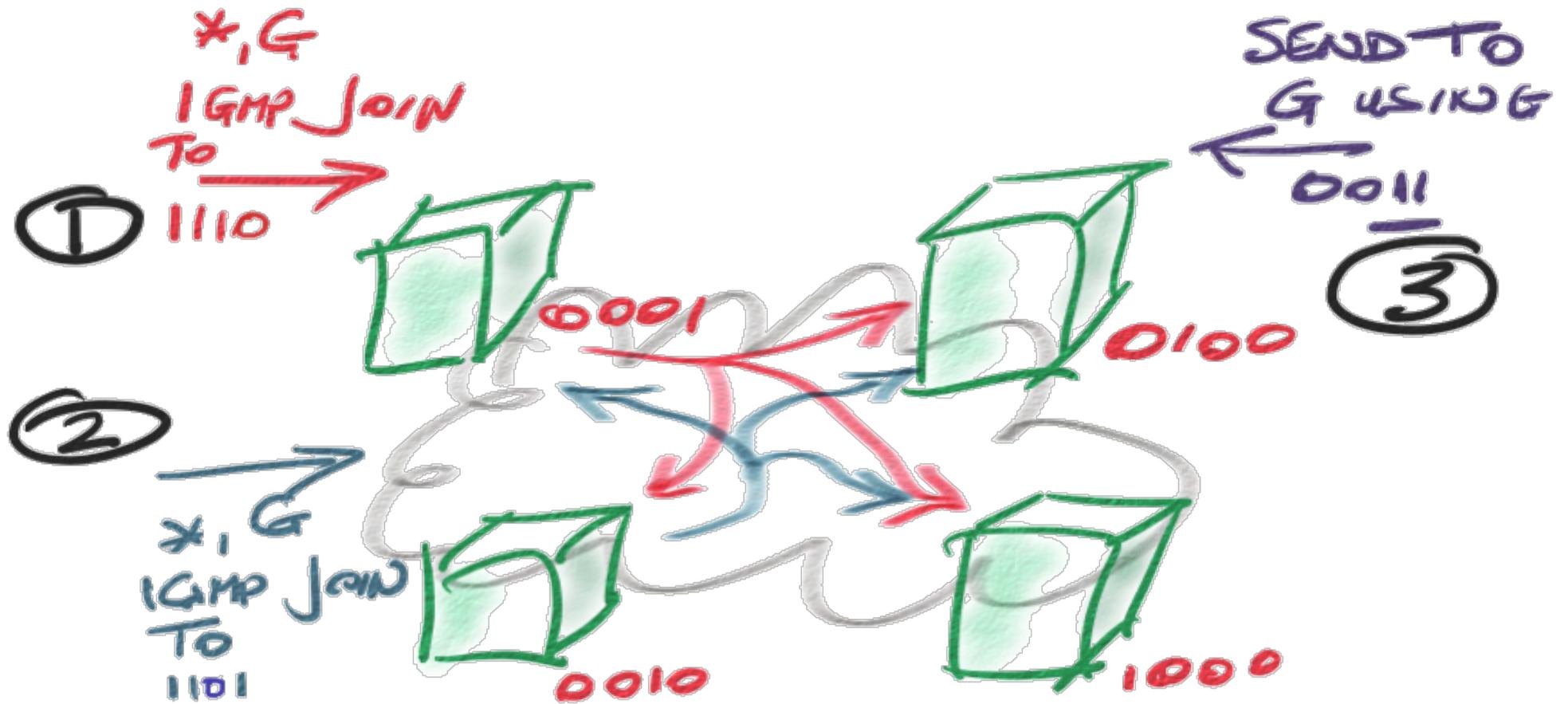
IGP ADVERTISES LABEL α = SUBDOMAIN# SET#



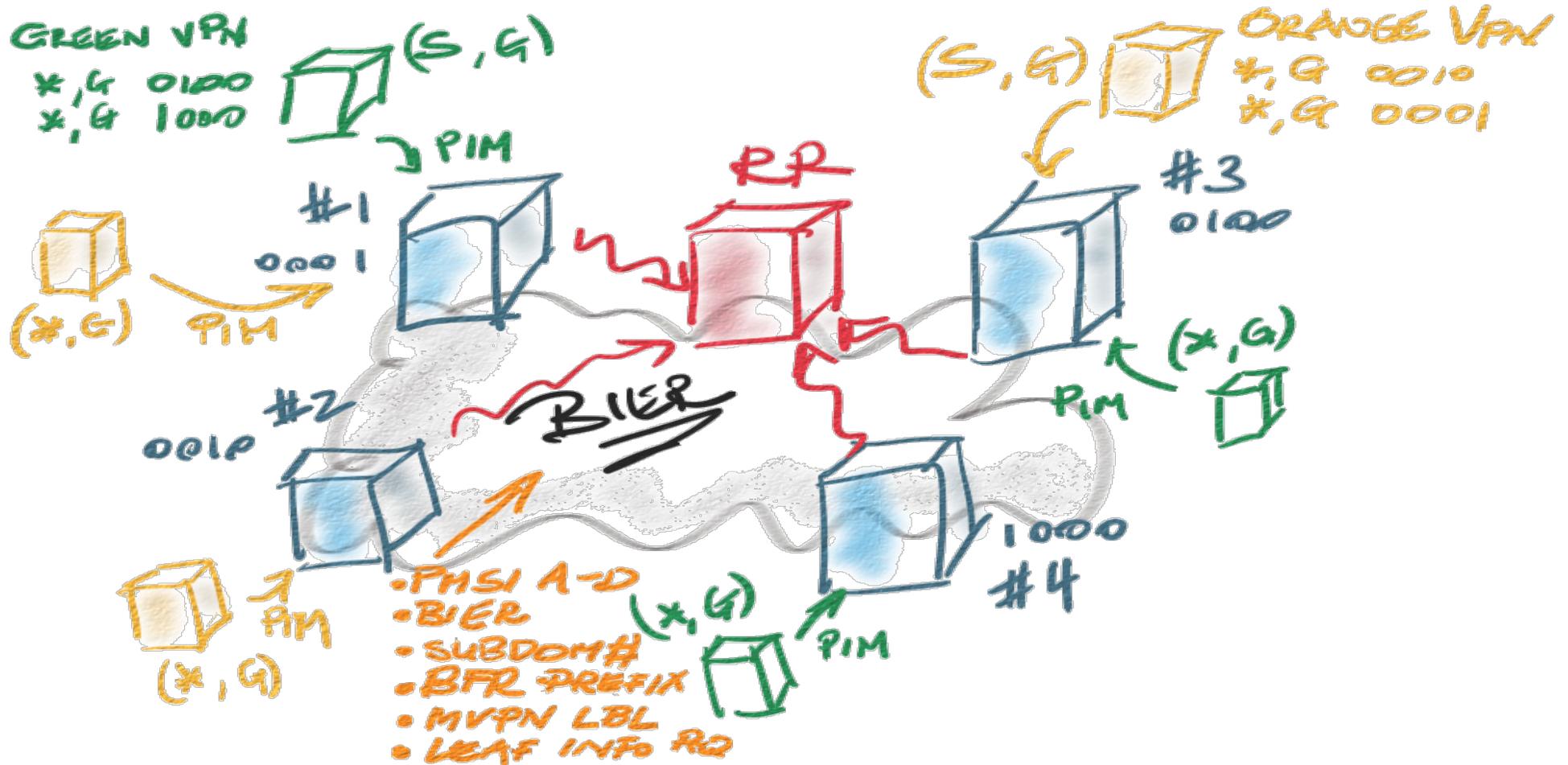
BIER OVERLAY

- IGMP
- MVPN/EVPN
- BIER ADVANTAGES

BIER: IGMP OVERLAY



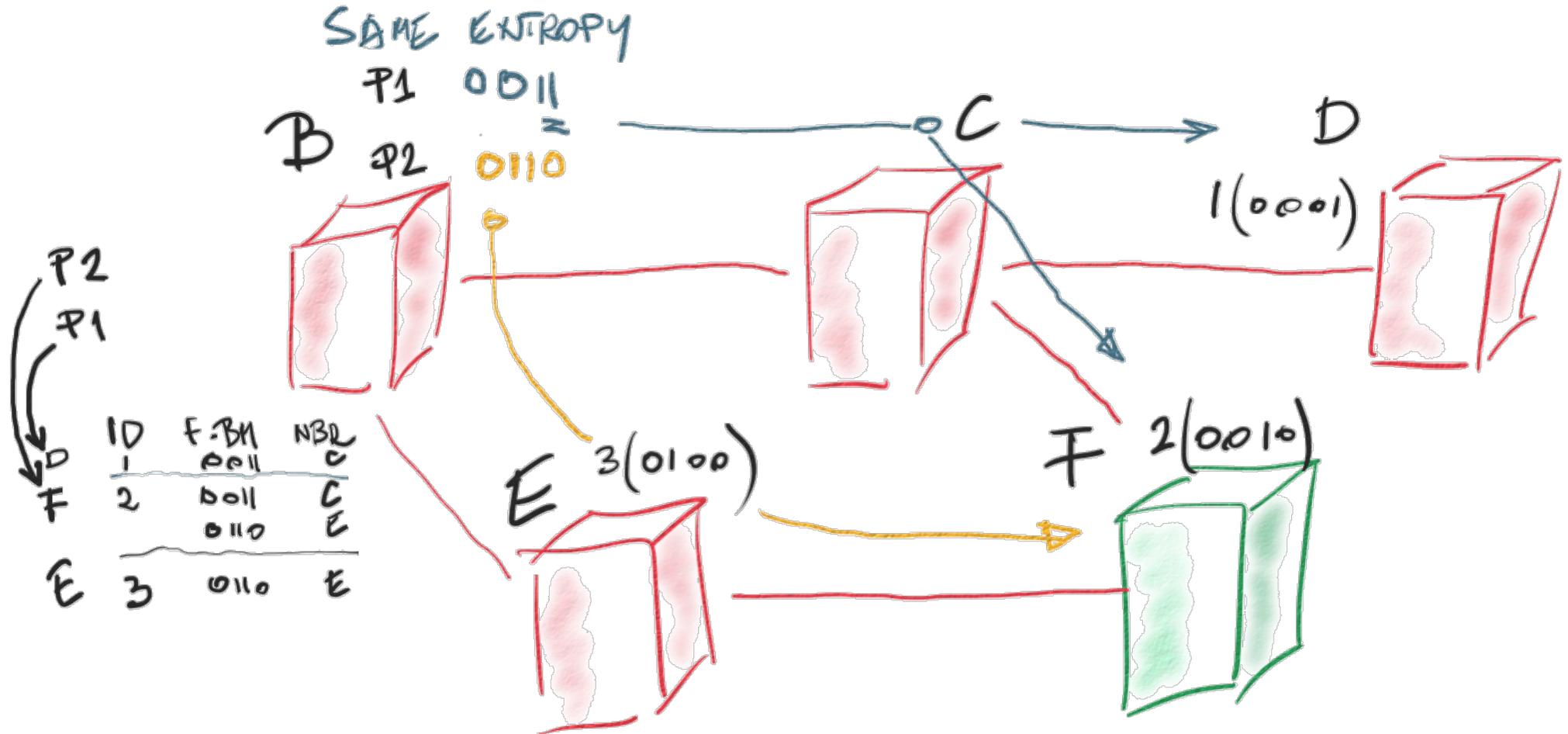
BIER: MVPN OVERLAY



BIER ADVANTAGES

- UNICAST SPEED & PROTECTION
- NO PIM
- NO RSVP, MLDP
- UNICAST ECMP
- NO PER VPN OR (S,G) STATE IN THE CORE
- OAM

ECMP: ADVANCED CONSIDERATIONS



JUNOS IMPLEMENTATION STATUS

- CURRENT STATUS: INTEGRATED AND TESTED
- PoC IS PRODUCT QUALITY DESIGN (I.E. NOT A PROTOTYPE BUT PRE-PRODUCTION CODE WITH ARCHITECTURE INCLUDING ALL BIER CAPABILITIES)
- SCOPE OF PoC SUPPORT
 - MX & VMX (FULL TRIO MICROCODE)
 - 256 BML ONLY (ARTIFICIAL RESTRICTION TO SIMPLIFY TESTING)
 - MPLS ENCAPSULATION ONLY (PREPARED FOR ETHERNET ENCAPS)
 - 4 SETS ONLY (ARTIFICIAL RESTRICTION TO SIMPLIFY TESTING)
 - MULTIPLE SUBDOMAINS
 - NON-DETERMINISTIC ECMP ONLY
 - NO NSR/RESTART
 - SPRING CANNOT BE ENABLED AT SAME TIME AS BIER (ARTIFICIAL RESTRICTION TO SIMPLIFY TESTING)
- ONLY FIRST BACKUP PROTECTION NEXT-HOP IS USED
- REMOTE LFA BACKUP PROTECTION UNTESTED
- OSPF
 - v2 ONLY
 - OSPF DRAFT VERSION -08 ONLY
 - AREA 0.0.0.0 ONLY
 - NO TUNNELING OVER BIER INCAPABLE ROUTERS, I.E. ALL ROUTERS MUST BE BIER CAPABLE (UNICAST AND BIER FULLY CONGRUENT)
 - MULTIPLE TOPOLOGIES
- MVPN OVERLAY
 - SELECTIVE AND INCLUSIVE TUNNELS

An aerial night view of a city, likely Chicago, with a network of glowing blue lines and nodes overlaid on the scene. The lines form a complex web of connections across the cityscape, with bright blue nodes at the intersections. The city lights are visible in the background, and the sky is a mix of blue and orange from the sunset or sunrise. A dark semi-transparent rectangle is positioned on the left side of the image, containing the word "THANKS" in white capital letters.

THANKS