RIPE

Bit Indexed Explicit Replication

BIER

Stateless Multi-point Replication

RIPE

Greg Shepherd, May 2015

Background – IPMulticast History

- Steven Deering, 1985, Stanford University
- RFC988, 1986 (Obsoleted by RFC1112, 1989)
- Multicast is part of the IP protocol stack
- Intended as an Internet-wide end-to-end service

Background – IPMulticast Uses

- Any applications with multiple receivers
  - One-to-many or many-to-many
- Live video distribution
- Collaborative groupware
- Periodic data delivery—“push” technology
  - Stock quotes, sports scores, magazines, newspapers, adverts
- Reducing network/resource overhead
  - More than multiple point-to-point flows
- Distributed interactive simulation (DIS)
  - War games
  - Virtual reality

Background – IPMulticast Challenges

- Explicit Tree Building Protocol
  - Tree state per flow
  - RPF tree building can have multicast taking different paths than unicast
  - Convergence times negatively impacted by tree state
    - No way to aggregate state without sacrificing optimal delivery
      - Choose between state explosion or data flooding
- Data-driven events
- Specialized skill set to troubleshoot and maintain

Background – Today

- The benefits of multi-point services are well understood
- The challenges of the current solutions often result in a failed cost/ benefit analysis
- Only those networks with an overwhelming business need have successful multicast deployments
- Much of the community have come to think of multicast as a failed technology
- Can we do better?

The BIER Epiphany

- Consider MY topology rather than a global topology
- Only encode the end-receivers in the packet header
  - Not the intermediate nodes
- Assign end-receivers a Bit Position from a Bit String
  - The smallest identifier possible
  - Advertise in the IGP
- Encode the Bit String in the packet header
  - Using some sort of encapsulation
- Create a Bit Forwarding Table on all BIER nodes to allow multicast packet forwarding using the Bit String in the packet
  - Derived from the RIB, SPF based
- We call it, Bit Indexed Explicit Replication (BIER)

IETF

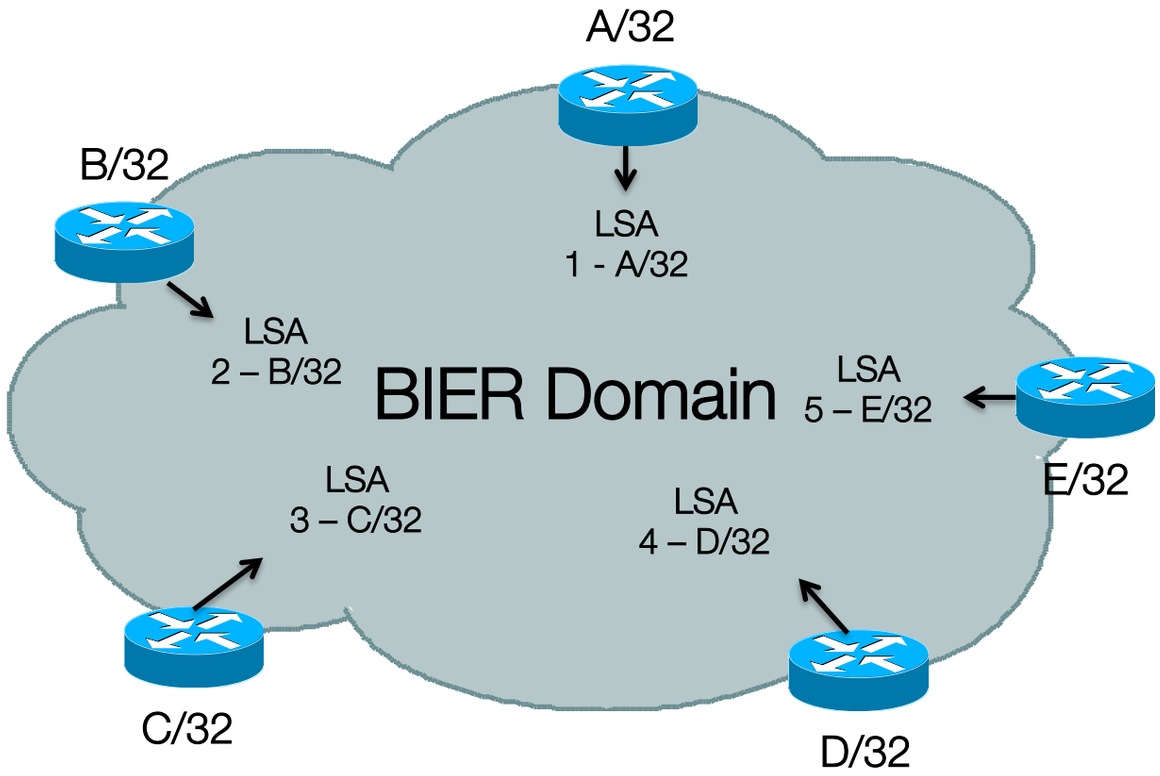
- The BIER idea was presented in a BOF at the IETF in Hawaii.
  - November 2014.
- BIER WG 1st meeting at IETF 92, March 2015
- Vendors collaborating
  - Cisco
  - Ericsson
  - Alcatel-Lucent
  - Juniper
  - Huawei
- Received very good traction and support

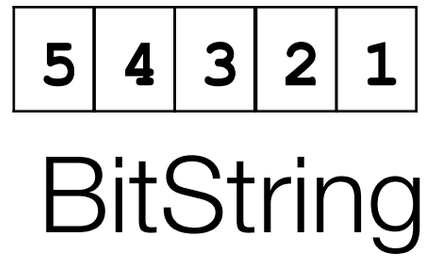
IETF drafts

- draft-ietf-bier-problem-statement
- draft-ietf-bier-architecture
- draft-ietf-bier-encapsulation-mpls
- draft-ietf-bier-use-cases
- draft-ietf-bier-mvpn
- draft-ietf-bier-ospf-extensions
- draft-ietf-bier-isis-ranges

BIER Solution Overview

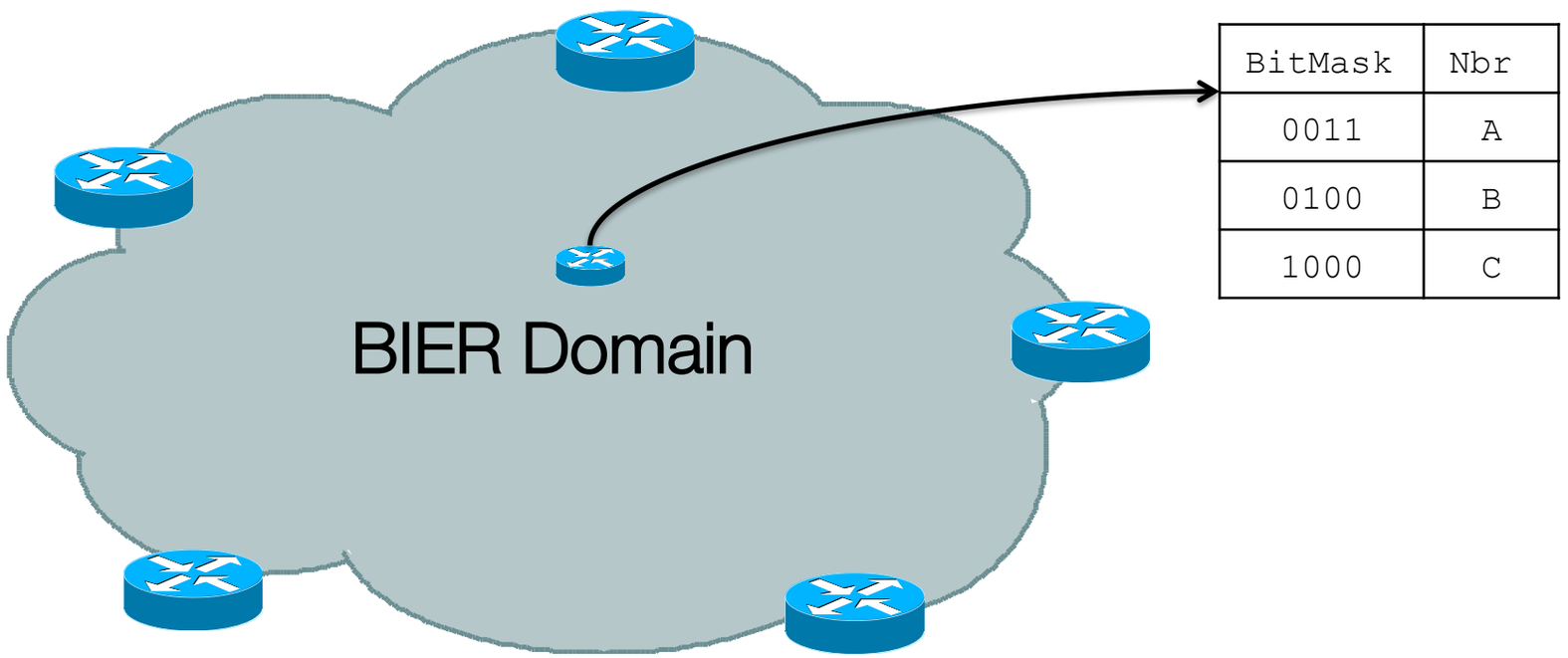
Basic Idea BIER





1. Assign a unique Bit Position from a BitString to each BFER in the BIER domain.
2. Each BFER floods their Bit Position to BFR-prefix mapping using the IGP (OSPF, ISIS)

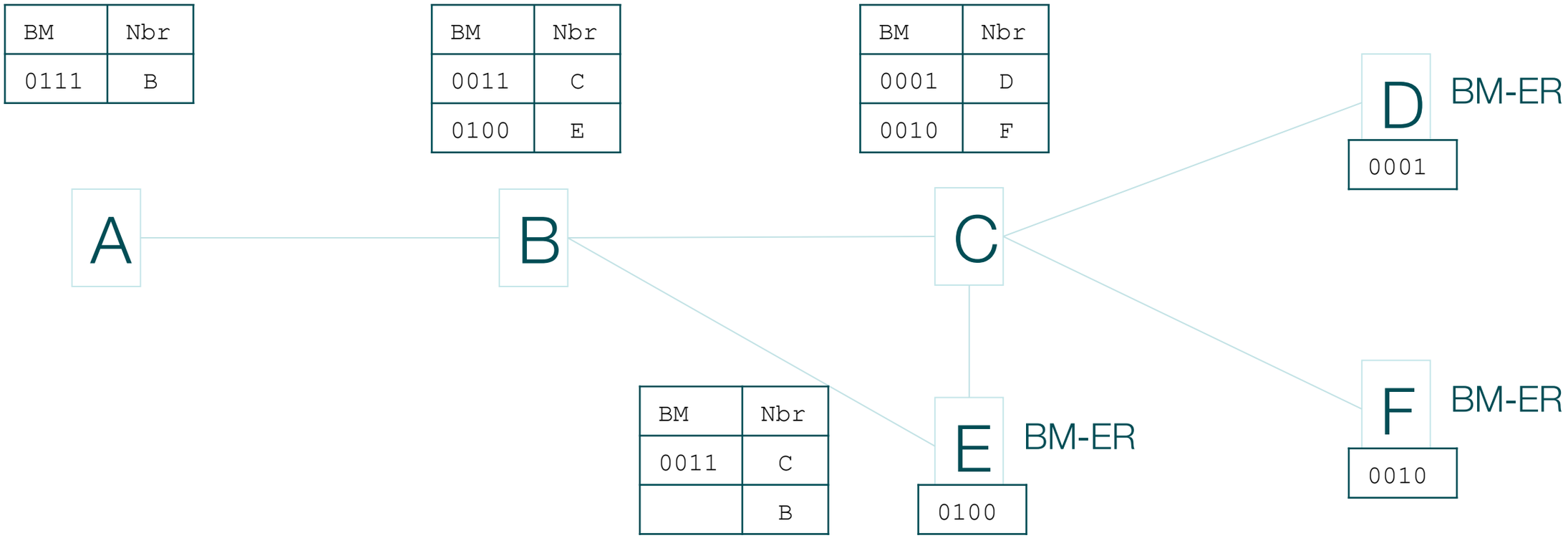
Basic Idea BIER



BitMask	Nbr
0011	A
0100	B
1000	C

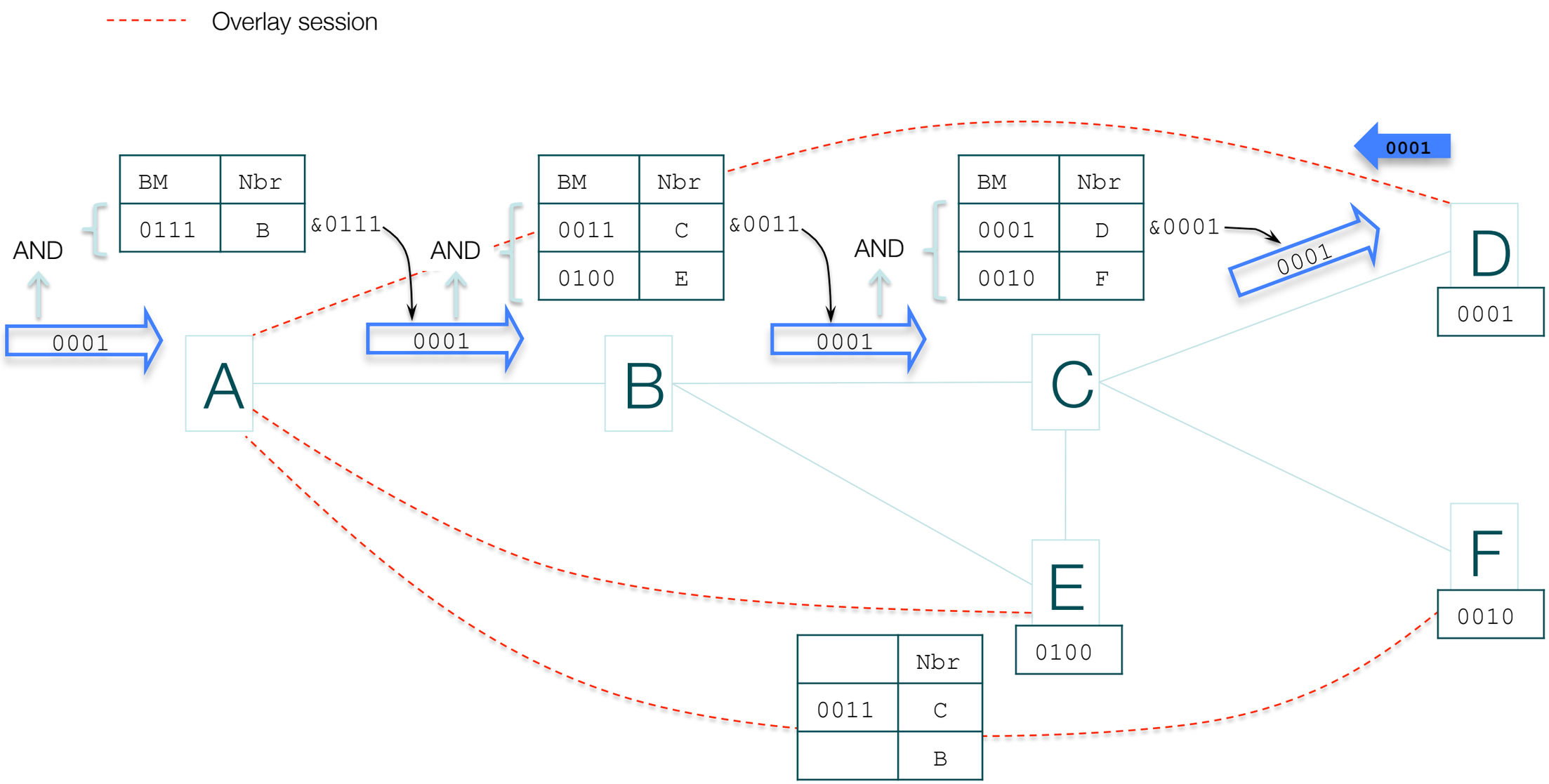
1. Assign a unique Bit Position from a mask to each edge router in the BIER domain.
2. Each edge router floods their bit-position-to-ID mapping with a new LSA – OSPF or ISIS
3. All BFR's use unicast RIB to calculate a best path for each BFR-prefix
4. Bit Positions are OR'd together to form a Bit Mask per BFR-nbr
5. Packets are forwarded and replicated hop-by-hop using the Bit Forwarding Table..

Bit Index Forwarding Table

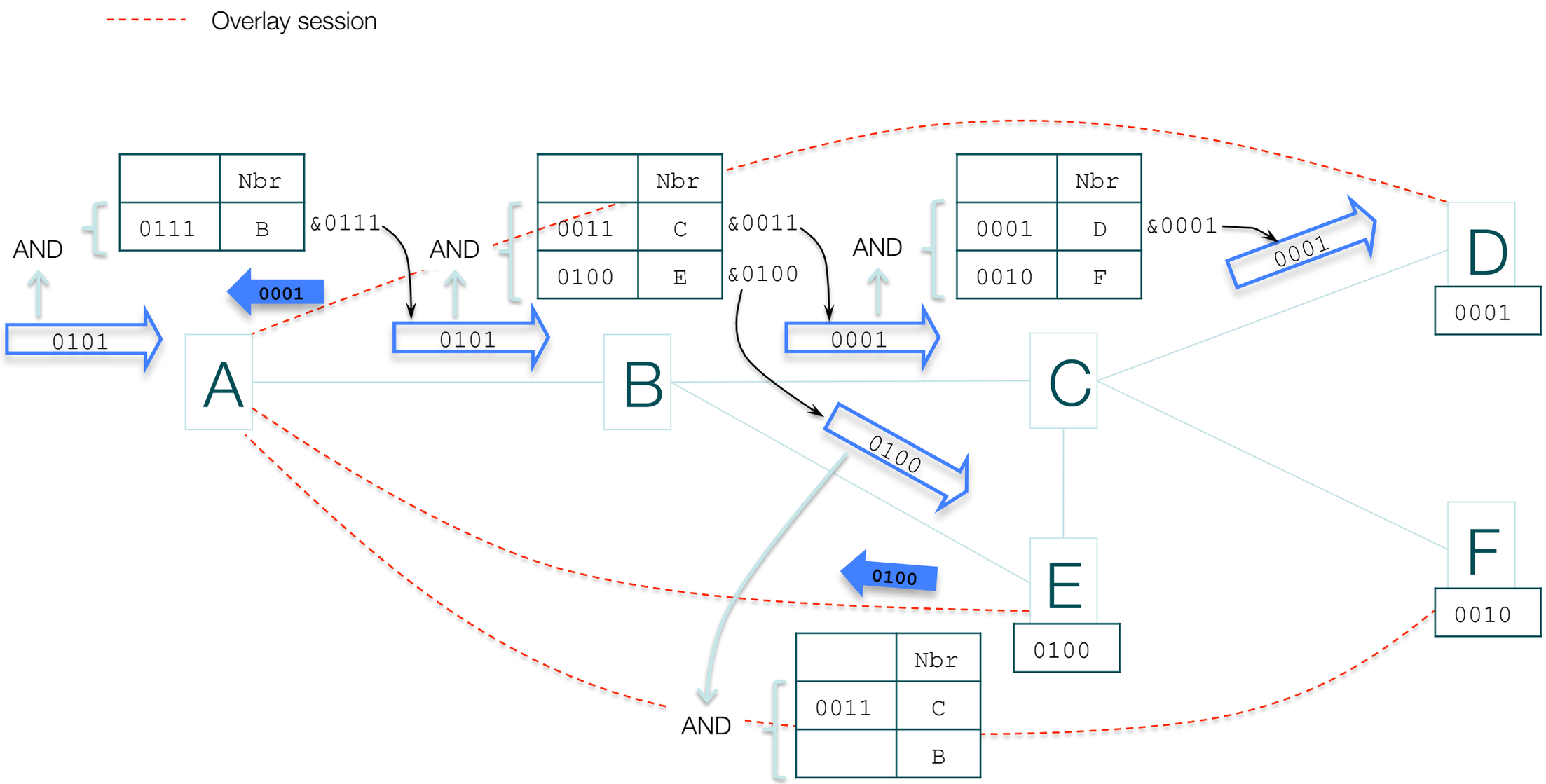


- D, F and E advertise their Bit positions in the IGP (flooded).
- A, B and C know the mapping between the Bit and RID,
- Based on shortest path route to RID, the Bit Mask Forwarding Table is created.

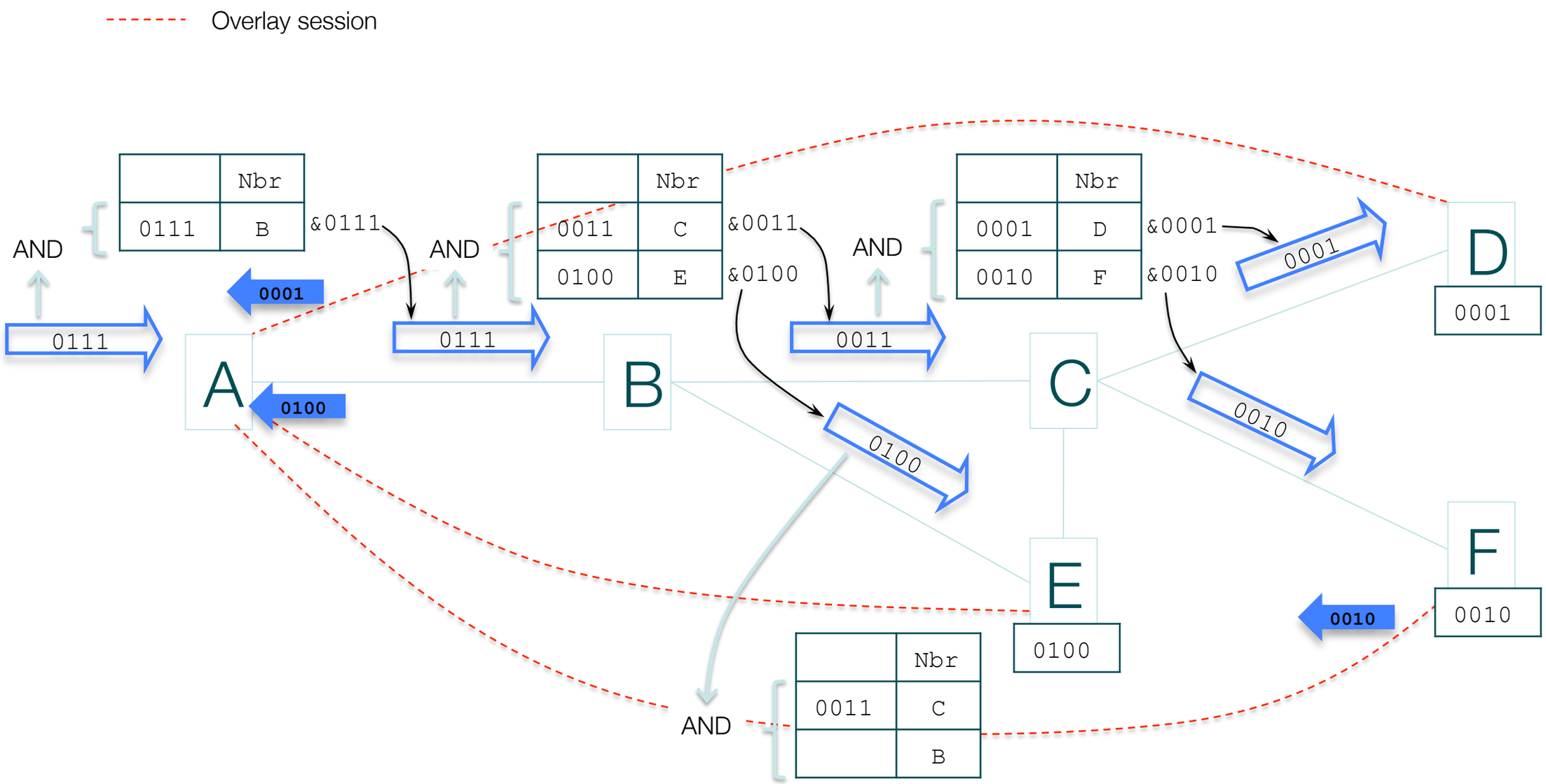
Forwarding Packets



Forwarding Packets



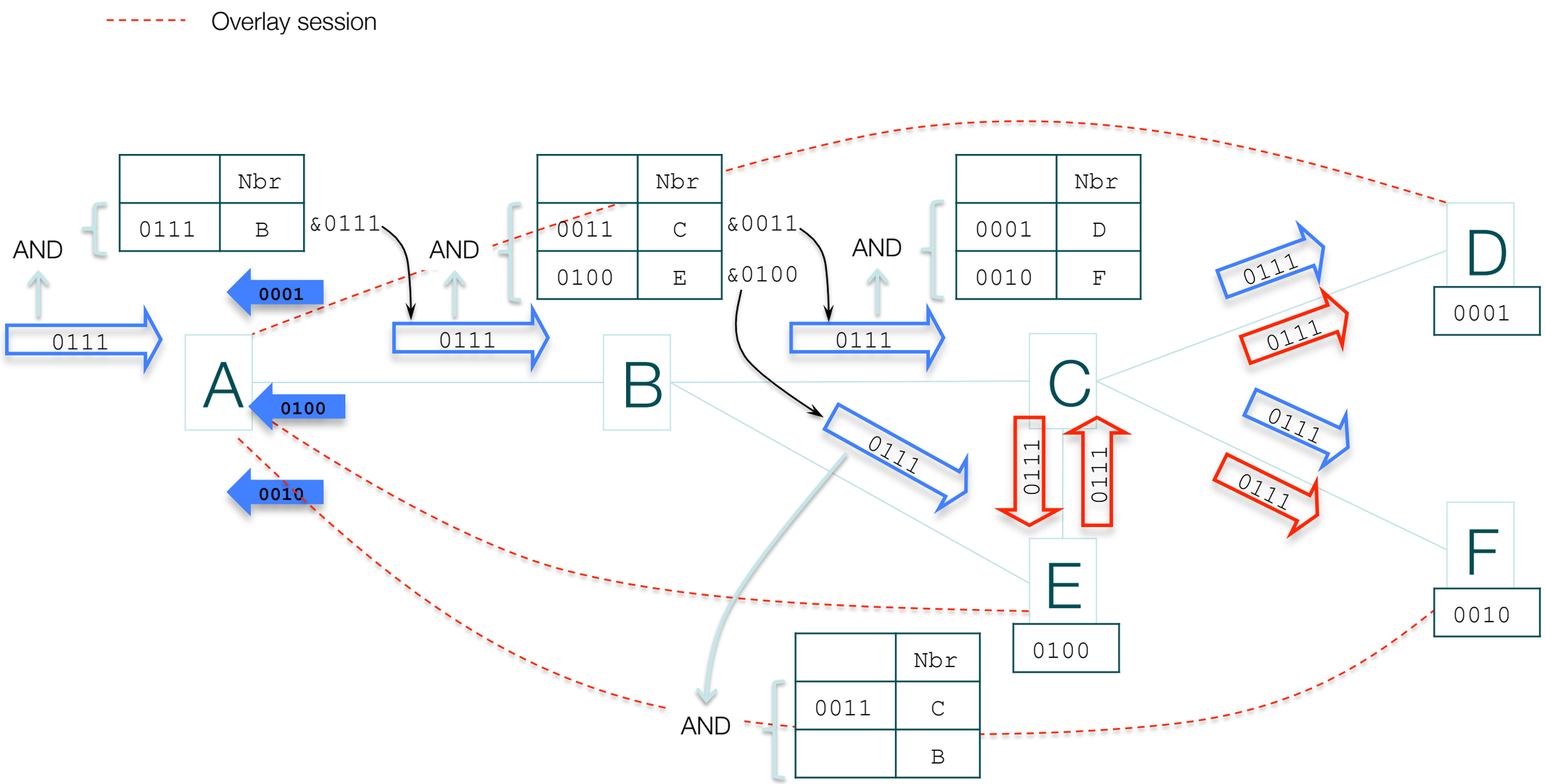
Forwarding Packets



Forwarding Packets

- As you can see from the previous slides, the result from the bitwise AND (&) between the Bit Mask in the packet and the Forwarding table is copied in the packet for each neighbor.
- This is the key mechanism to prevent duplication.
- Look at the next slide to see what happens if the bits are not reset
- If the previous bits would not have been reset, E would forward the packet to C and vice versa.

Forwarding Packets

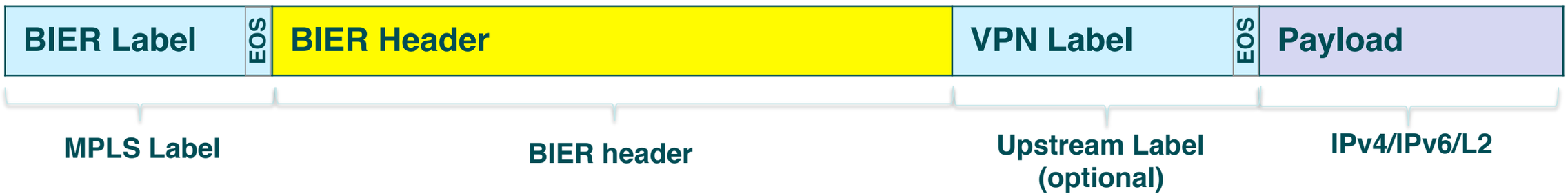


How many Bits and Where?

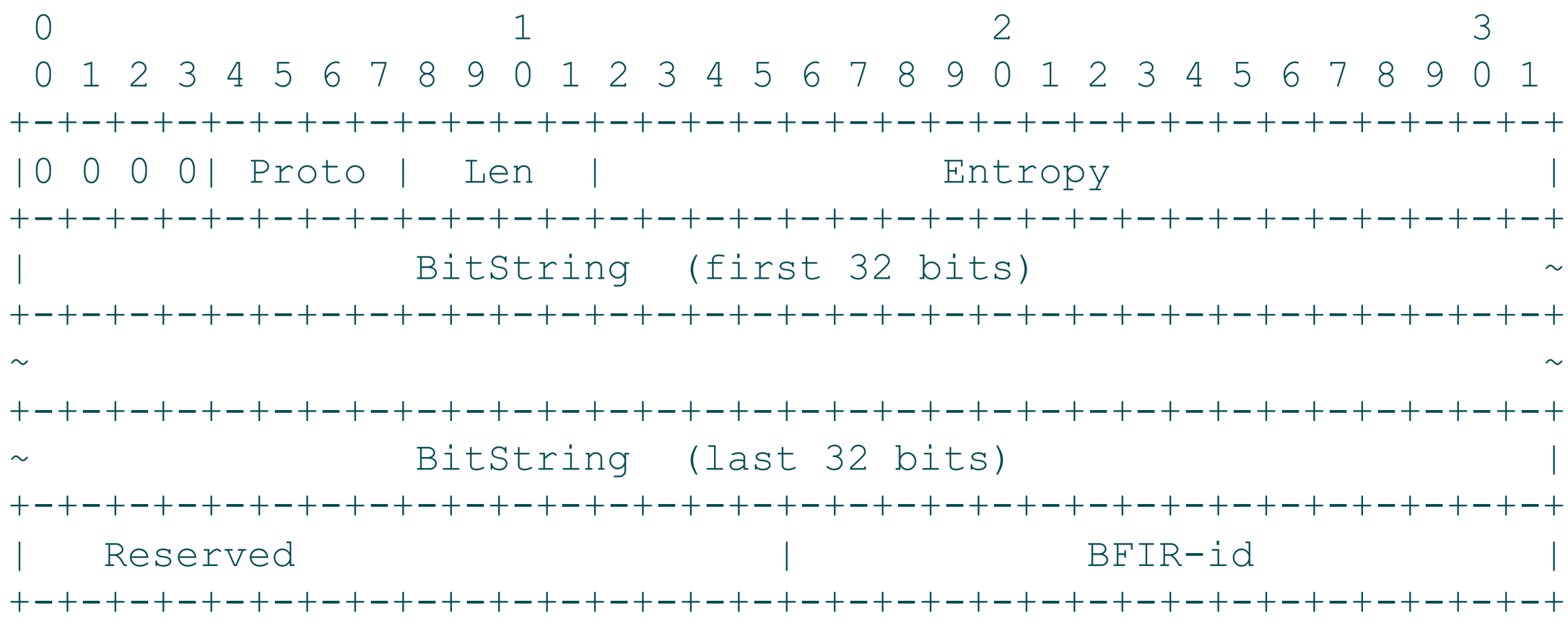
- The number of multicast egress routers that can be addressed is depending on the number of Bits that can be included in the BitString
- The BitString length is depending on the encapsulation type and router platform.
- We identified 5 different encoding options, most attractive below;
  1. MPLS, below the bottom label and before IP header.
  2. IPv6, extensions header.

MPLS encapsulation

- Multiple vendors have confirmed 256 bits is workable on today's programmable platforms
- WG is using 256 bits as a starting point



BIER Header



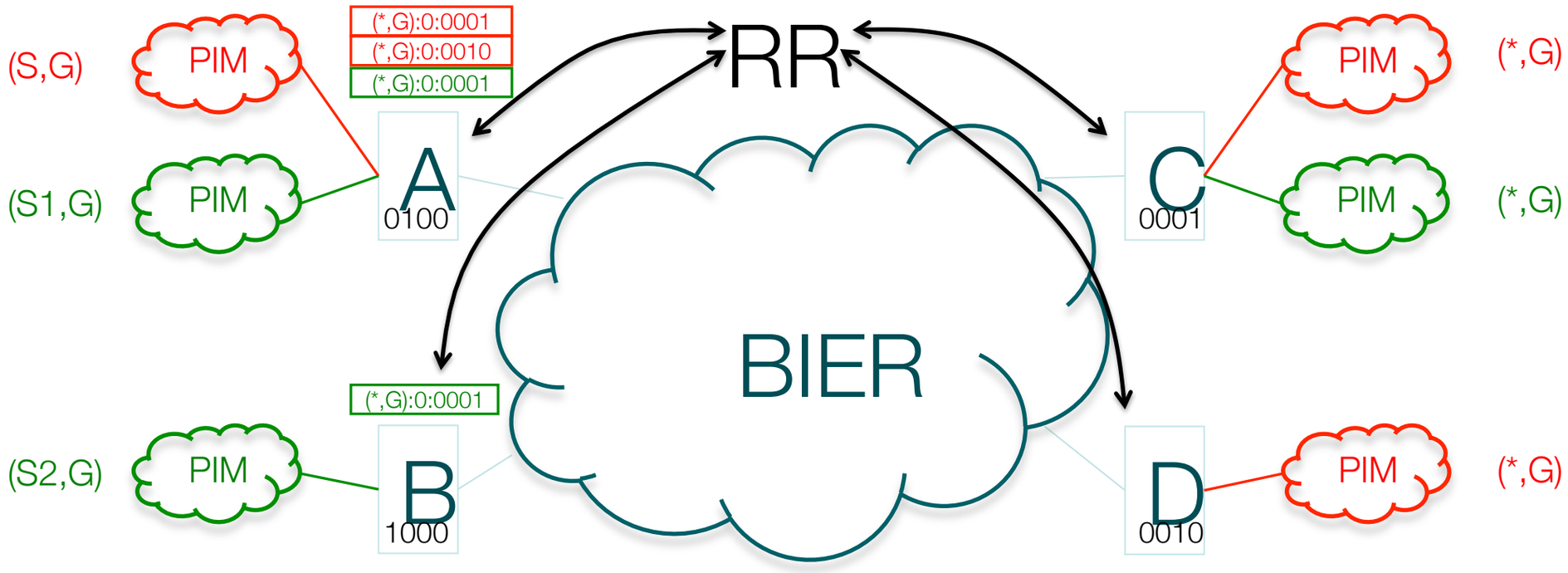
- Documented in draft-ietf-bier-mpls-encapsulation

MVPN over BIER

MVPN over BIER

- BIER replaces PIM, mLDP, RSVP-TE or IR in the core
- BIER represents a full mesh (P2MP) connectivity between all the PE's in the network
- There is no need to explicitly signal any MDT's (or PMSI's)
- Current MVPN solutions have many profiles
  - This is partly due to the tradeoff between 'State' and 'Flooding'
  - Different C-multicast signaling options
- MVPN over BIER, there is one profile
  - BGP for C-multicast signaling
- No need for Data-MDTs

MVPN over BIER



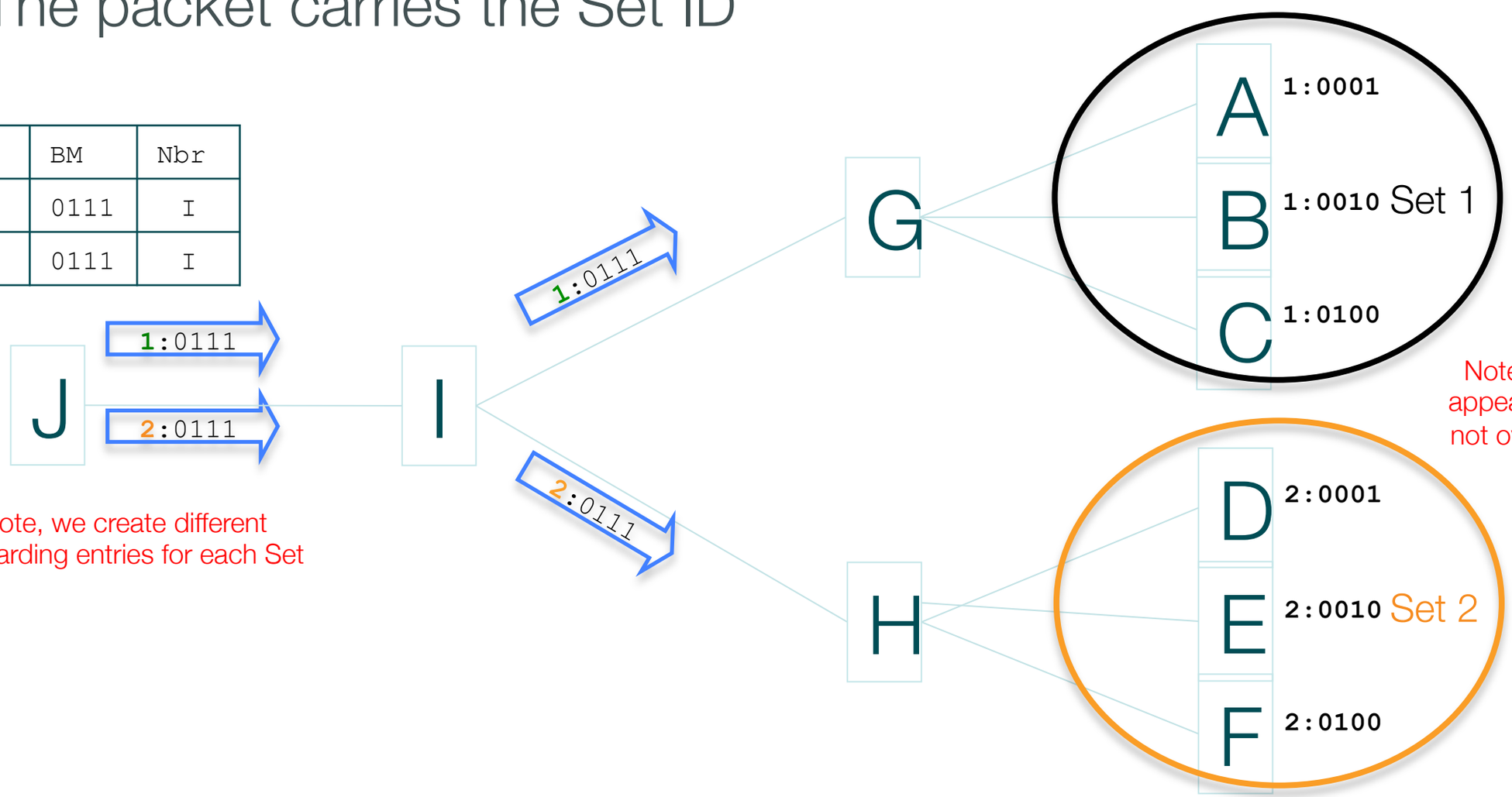
- The BGP control plane defined for MVPN can be re-used.
- PIM (S,G)/(*,G) can be translated into BGP updates.
- Requirement, we depend on Leaf AD routes for explicit tracking!
- Big difference, there is no Tree per VPN…!!!
- The BIER packets needs to carry Source ID and upstream VPN context label

Sets and Areas

BIER Sets

- To increase the scale we group the egress routers in Sets
- Each Bit Position is unique in the context of a give Set
- The packet carries the Set ID

Set	BM	Nbr
1	0111	I
2	0111	I



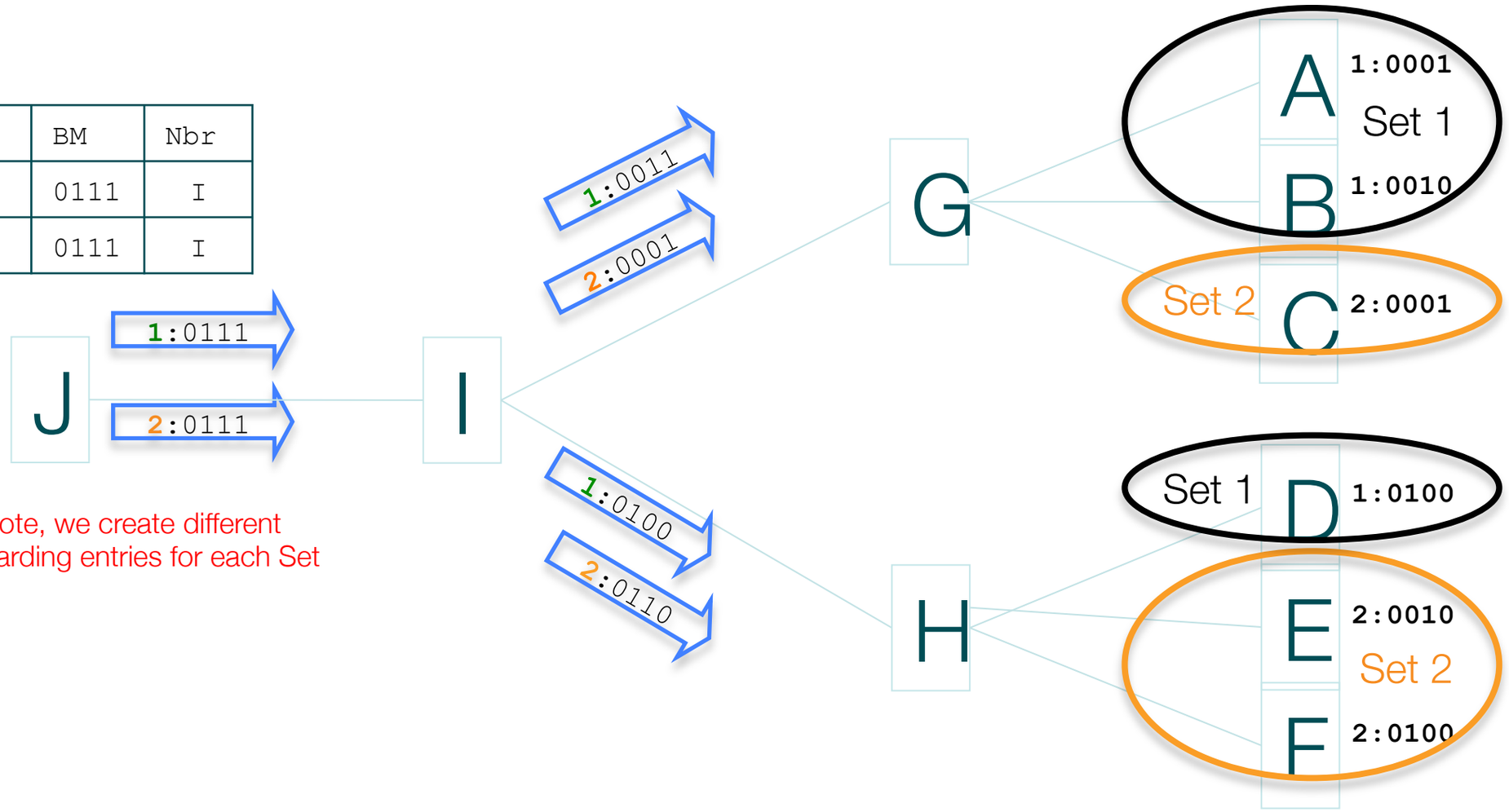
Note, we create different forwarding entries for each Set

Note, Bit Positions 1,2,3 appear in both Sets, yet do not overlap due to different Set assigments

BIER Sets

- There is no topological restriction which set an egress belongs to
- But it may be more efficient if it follows the topology

Set	BM	Nbr
1	0111	I
2	0111	I



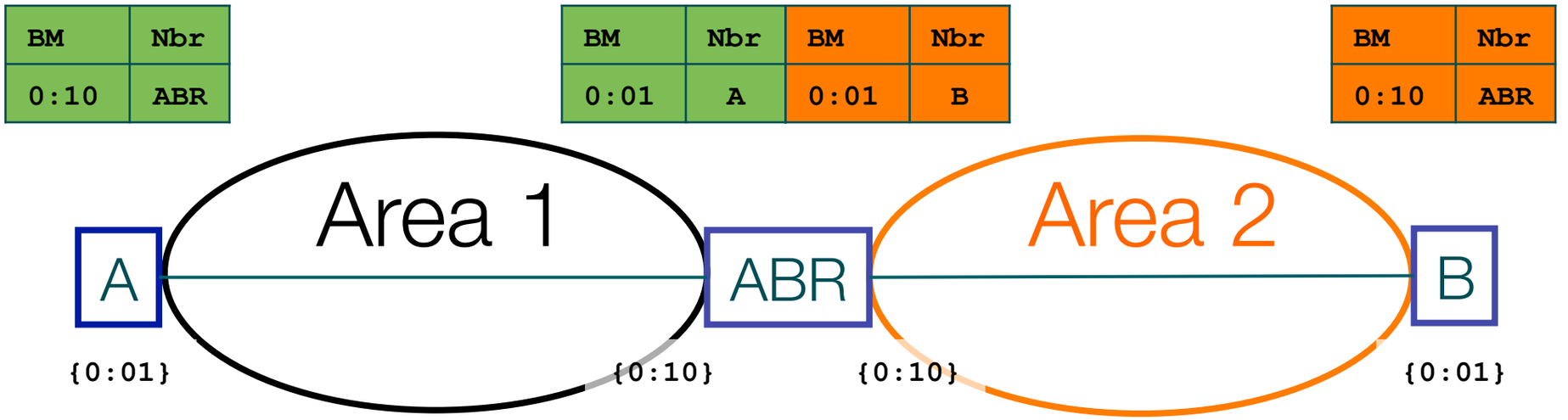
Note, we create different forwarding entries for each Set

BIER Sets

- If a multicast flow has multiple receivers in different Sets, the packet needs to be replicated multiple times by the ingress router, once for each set
- Is that a problem? We don't think so…
- The Set identifier is part of the packet.
- Can be implemented as MPLS label.

BIER Area

- A bit Mask only needs to be unique in its own area.
- ABR's translate Bit Masks between area's.
- Requires a IP lookup and state on the ABRs.
- This is very similar for 'Segmented Inter-AS MVPN'.



Conclusion

Advantages

- Packets forwarded via BIER follow the unicast path towards the receiver, inheriting unicast features like FRR and LFA.
- There is no per multicast flow state in the network.
- Multicast convergence is as fast as unicast, there is no multicast state to re-converge, signal, etc.
- Nice plugin for SDN, its only the ingress and egress that need to exchange Sender and Receiver information.
- The core network provides a many-2-many connectively between all BIER routers by default following the IGP.
- No Multicast control protocol in the network.

Disadvantages

- The Bit String length has an upper bound and may not cover all deployment scenarios.
- Using sets to increase the number of egress routers may require the ingress to replicate the packet multiple times.
- Using area's requires the ABR to have state.
- Existing low-end platforms are less flexible to adopt BIER.
- ASIC/Merchant spin required for low-end platforms

Questions?


NL
1st BoF
BIER
BIT INDEXED EXPLICIT REPLICATION
bier@ietf.org - http://trac.tools.ietf.org/bof/trac/
IETF Nov 2015
Honolulu
Hawaii