MPLS NETWORK ACTION (MNA) HEADER ENCODINGS (DRAFT-JAGS-MPLS-MNA-HDR-03)

Jaganbabu Rajamanickam (jrajaman@cisco.com)
Rakesh Gandhi (rgandhi@cisco.com)
Royi Zigler (royi.zigler@broadcom.com)
Haoyu Song (haoyu.song@futurewei.com)
Kireeti Kompella (kireeti.ietf@gmail.com)

CONTRIBUTORS

- Jisu Bhattacharya (jisu@cisco.com)
- Bruno Decraene (bruno.decraene@orange.com)
- Weiqiang Cheng (chengweiqiang@chinamobile.com)
- Xiao Min (xiao.min2@zte.com.cn)
- Luay Jalil (luay.jalil@verizon.com)
- Jie Dong (jie.dong@huawei.com)
- Tianran Zhou (zhoutianran@huawei.com)
- Bin Wen (Bin_Wen@cable.comcast.com)
- Sami Boutros (sboutros@ciena.com)
- Tony Li (tony.li@tony.li)
- John Drake (jdrake@juniper.net)

ABBREVIATIONS

Abbreviations	Meaning
AD	Ancillary Data
HBH	Hop-By-Hop Scope
I2E	Ingress-To-Egress Scope
IHS	Ingress-To-Egress, Hop-By-Hop, Select Scope
ISD	In-Stack Data
LSE	32-Bit Label stack Entry
MNA	MPLS Network Action
NA	Network Action
NAI	Network Action Indicator
NAI-Opcode	Network Action Indicator Opcode
NAL	Length of Network Action in number of LSEs
NASL	Length of Network Action Sub-Stack in number of LSEs
NASS	Network Action Sub-Stack
NASI	Network Action Sub-Stack Indicator
O-Bit	Ordering Bit, indicates that the NASS requires ordering and MUST process the NAIs in the order it is encoded else it MUST be dropped
P	Post-Stack Network Action Presence
PSD	Post-Stack Data
UOH	Unknown Opcode Handling

AGENDA

- Scope
- High Level Solution
  - Network Action Sub-Stack Header
  - Network Action Encoding
- Reserved Network Action Opcodes
- Network Action Ordering
- Backwards Compatibility
- Advantages
- Next Steps

SCOPE

- Solution for MNA Encoding Formats Carried in Label Stack
- Address MNA Requirements [draft-ietf-mpls-mna-requirements]
- Align with MNA Framework [draft-ietf-mpls-mna-fwk]

HIGH-LEVEL SOLUTION

MPLS Network Action (MNA) header contains two parts:

1. **Network Action Sub-Stack Header**
   a. Label to indicate the presence of Network Action Sub-Stack (NASI)
   b. NASS encoding parameters indicating the structure of NASS
      - Includes Scope, Ordering, Length, etc. parameters of NASS

2. **Network Action Encoding**
   a. Network Action is encoded in **TLV** format
      - **T**ype – Network Action Indicator Opcode
      - **L**ength – Network Action Length (NAL)
      - **V**alue – Ancillary Data (optional)

NETWORK ACTION SUB-STACK HEADER

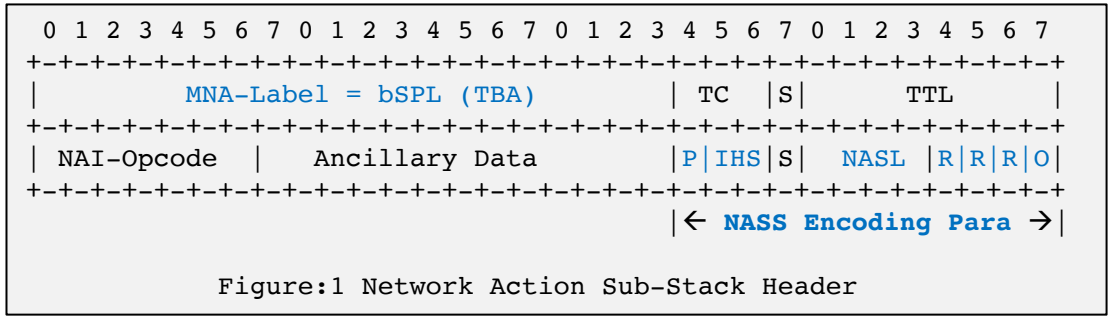


Figure:1 Network Action Sub-Stack Header

Network Action Sub-Stack Header Contains **MNA Label** and **NASS Encoding Parameters**.

MNA-Label: A new bSPL value (value TBA) is assigned to indicate the presence of the MPLS Network Action Sub-Stack (NASS).

NASS Parameters: The TTL and TC fields in the second LSE are used to encode NASS encoding parameters. These parameters are common for all NAIs encoded in this NASS. The NASS encoding parameter contains

P (**20th Bit**): Indicates the presence of Post-Stack NAs

IHS (**21-22 Bits**): Two-Bit value indicates the scope of the NASS (I2E (0) or HBH (1) or Select (2))

NASL (**24-27 Bits**): Network Action Sub-Stack Length. Four-Bit value indicates the Length of Current NASS

R (**28-30 Bits**): Reserved Bits for future use

O (**31st Bit**): Ordering Bit, this indicates that the NAIs encoded in this NASS MUST be processed in the order it is encoded

NETWORK ACTION ENCODING

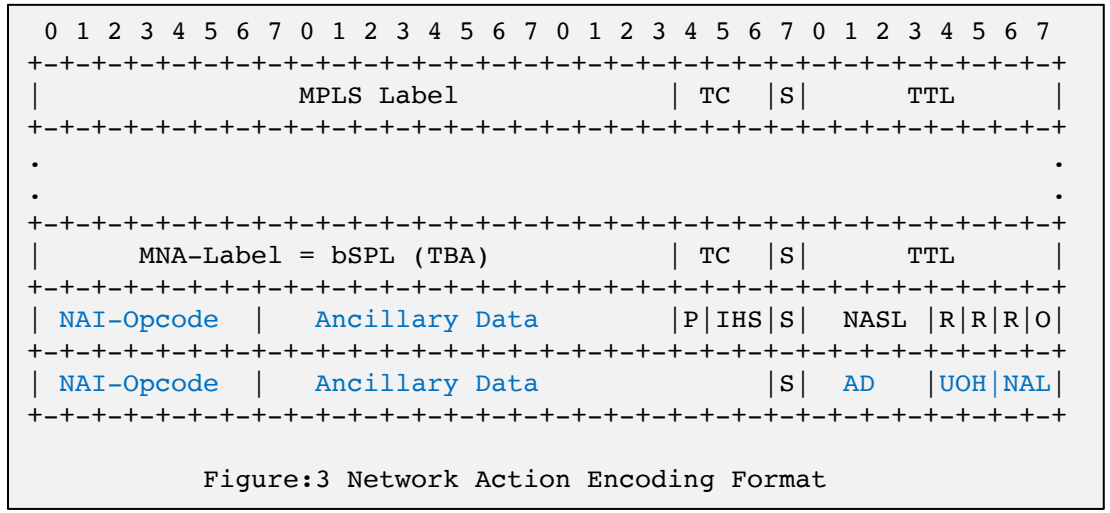


Figure:3 Network Action Encoding Format

Network Actions are encoded in the **TLV** Format.

NAI-Opcode (Type) : 7-bit value indicates the NAI Opcode value. This In-Stack NAI-Opcode will use an IANA registry to allocate the opcodes

NAL (Length) (Network Action Length): 2-bit value to indicate additional LSEs used to encode Ancillary Data

Ancillary-Data/AD (Value): 20 bits Ancillary Data carried to process Network Action corresponding to the NAI-Opcode

UOH (Unknown Opcode Handling): 2-bit value defines the action a node should take when it does not understand this specific opcode. (0 – Skip and process Next NA, 1 – Drop, 2 – Stop processing MNA and forward)

NETWORK ACTION SUB-STACK SCOPE

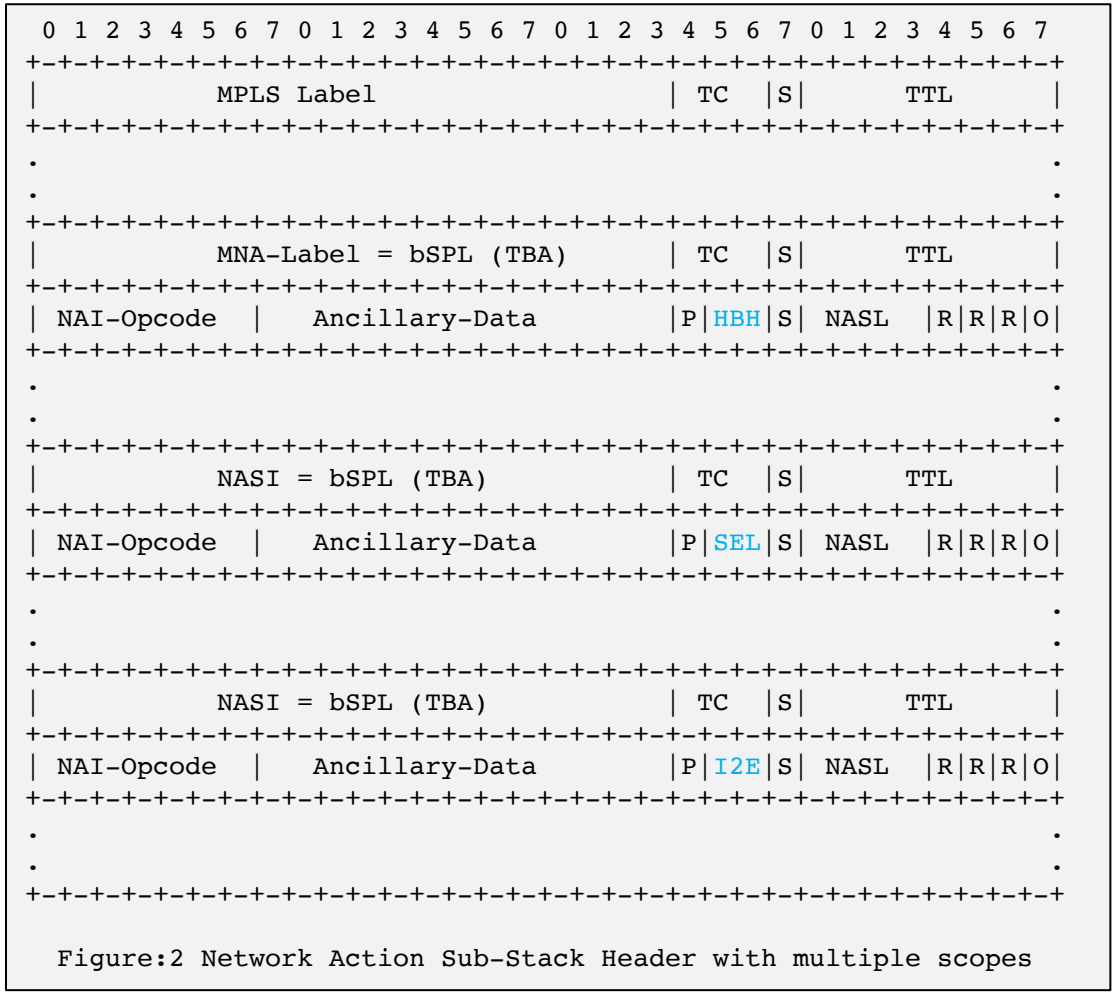


Figure:2 Network Action Sub-Stack Header with multiple scopes

NASS has 2-bit scope for:

- **I2E Scope (Ingress-To-Egress, value 00b):** This NASS is processed only on the Egress Node.
- **HBH Scope (Hop-By-Hop, value 01b):** This NASS is processed on all the nodes on the path of the packet.
- **Select Scope (value 10b)**: This NASS processed only on the select nodes.

Separating NASS based on the scope makes it easier for the mid-point nodes to only process HBH/Select scoped NASS.

The MPLS stack may carry up to three different scoped NASS.

”P” Bit is set with respect to the scope of the Post-Stack Data encoding, (i.e.) If only HBH scoped Post-Stack NAs are encoded then the “P” bit will be set only in the NASS with the HBH scope.

RESERVED NETWORK ACTION OPCODES – (1/3)

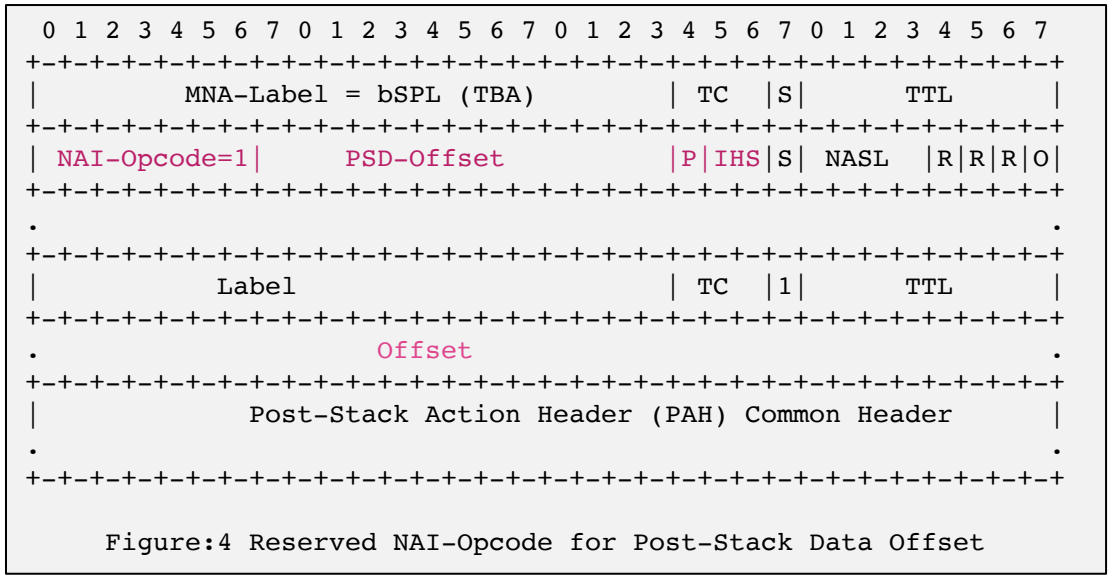


Figure:4 Reserved NAI-Opcode for Post-Stack Data Offset

Reserved NAI-Opcode Value:1 (Post-Stack Data Offset) – Opcode reserved to indicate the Post-Stack Action Header (PAH) start after the BOS in bytes. It can be placed at the 2nd LSE or after the 2nd LSE.

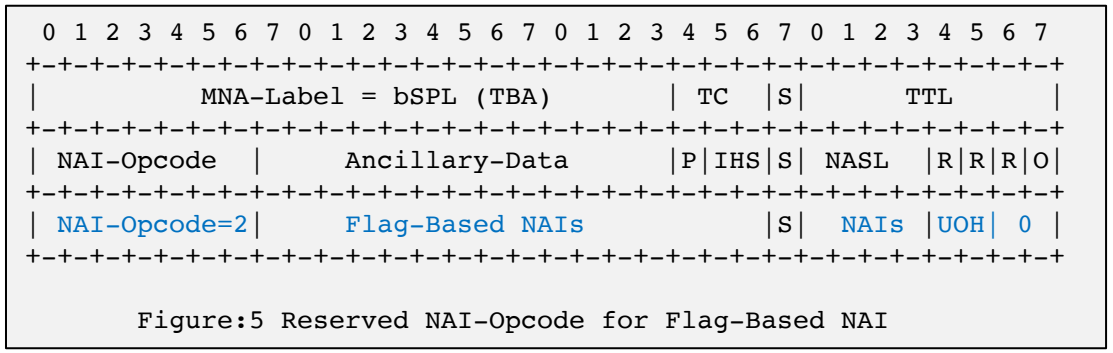


Figure:5 Reserved NAI-Opcode for Flag-Based NAI

Reserved NAI-Opcode Value:2 (Flag-Based NAIs) – Opcode reserved to carry Flag-Based NAIs. This Opcode does not require any Ancillary Data to process the Network Action. The Flag Bits offset value are read from Left to Right.

RESERVED NETWORK ACTION OPCODES – (2/3)

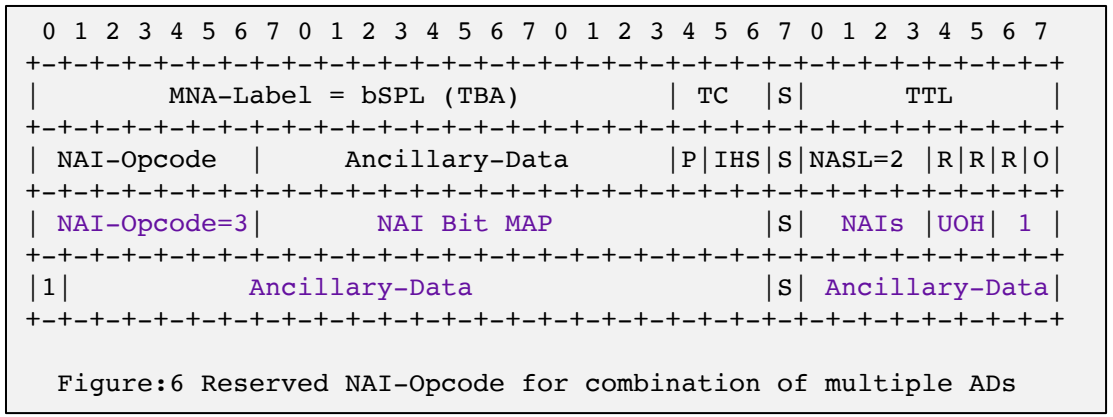


Figure:6 Reserved NAI-Opcode for combination of multiple ADs

Reserved NAI-Opcode Value:3 (Combination of Multiple NAIs) – Opcode reserved to carry combination of Multiple NAIs that May require AD. In this case the, 3rd LSEs AD carries set of NAI Bit Maps and the next LSEs carries 30-bit of AD corresponding to NAI Bit Map encoded.

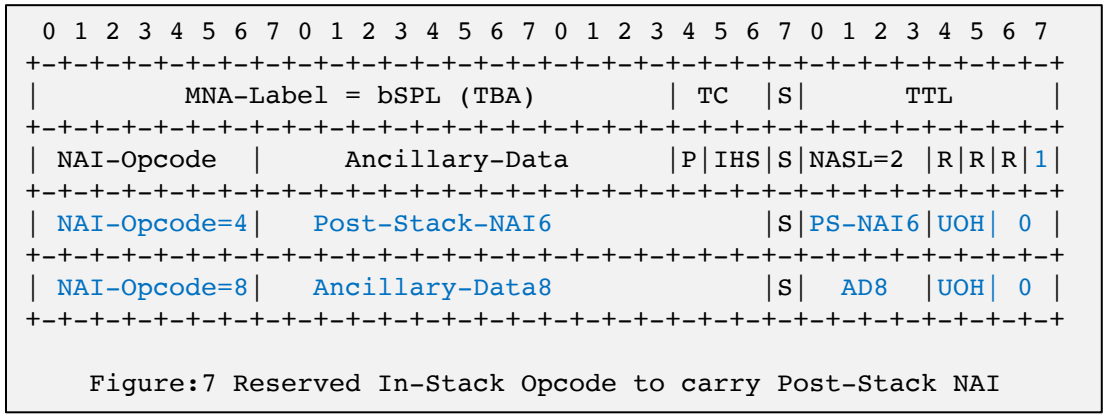


Figure:7 Reserved In-Stack Opcode to carry Post-Stack NAI

Reserved NAI-Opcode Value:4 (In-Stack and Post-Stack NA ordering) – Opcode is used for PSD NAI that needs to processed in specific order

- This opcode can be used for specifying ordering between In-Stack and Post-Stack NAIs.
- In this case, Post-Stack NAI Opcode “6” MUST be processed before In-Stack NAI Opcode “8”
- O-Bit (Ordering Bit) is set, if the nodes that are not capable of processing the NASS in order that are encoded then the packet MUST be dropped.

RESERVED NETWORK ACTION OPCODES – (3/3)

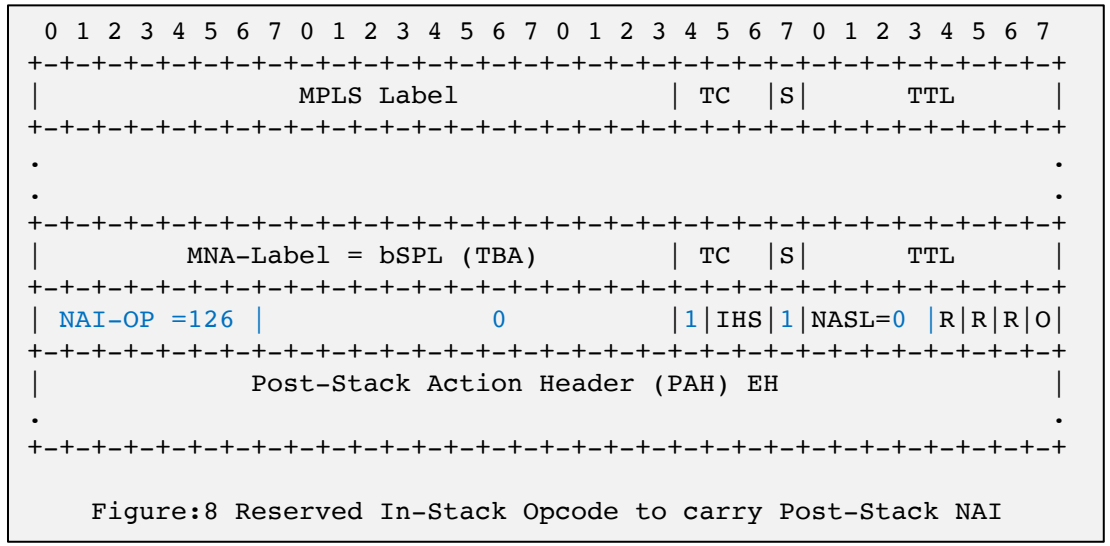


Figure:8 Reserved In-Stack Opcode to carry Post-Stack NAI

Reserved NAI-Opcode Value:126 (In-Stack No-Operation) – Opcode reserved to fill-in unused 20 bits of the 2nd LSE, if the field is not used.

One of the example is that when MNA carries only the Post-Stack Data, then the No-Operation In-Stack NAI opcode “126” will be used to fill-in.

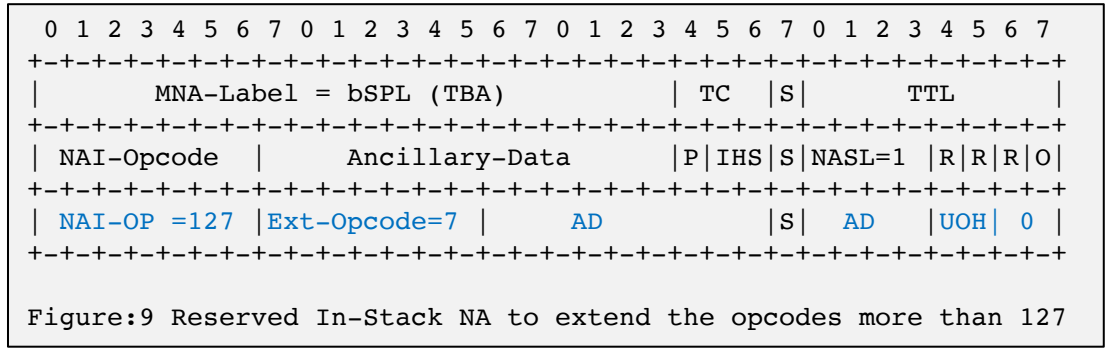


Figure:9 Reserved In-Stack NA to extend the opcodes more than 127

Reserved NAI-Opcode Value:127 (Opcode Extension) – Opcode reserved to extend the NAI-Opcode value beyond value 127.

NETWORK ACTION ORDERING – (1/2)

Network Action Ordering: In some cases, the node that encapsulates the MNA expects the other nodes to process the Network Actions in a certain order. The below example provides a framework to inform the order of the NA processing.
O-Bit (Ordering Bit) – This bit MUST be set to indicate the nodes to follow the order at which NAs are encoded. In general, processing the NAs in order is a complex process (Especially ordering between In-Stack and Post-Stack NAs) . Based on the Nodes capability it could decide to process or drop the packet.

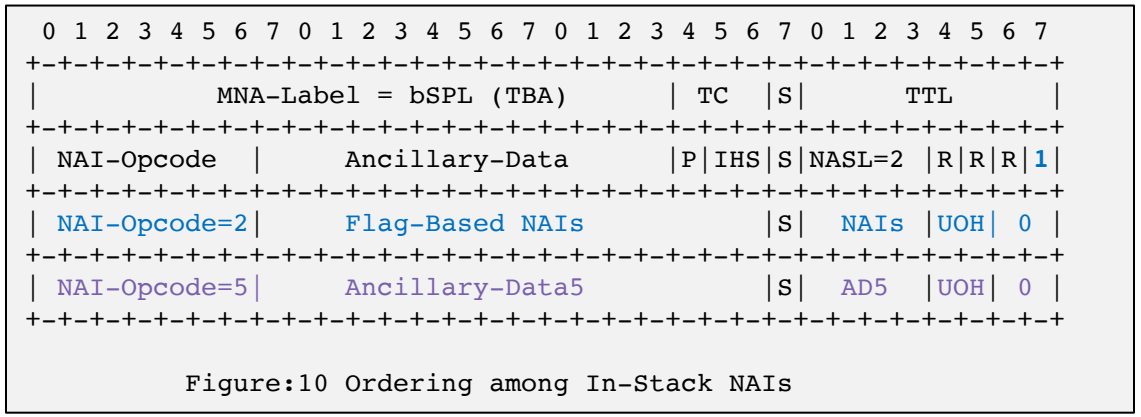


Figure:10 Ordering among In-Stack NAIs

Ordering Among In-Stack NAIs: The order of processing the NA should follow the order of NAs encoded.
Example NAI-opcode 5 is executed after NAI-opcode 2.

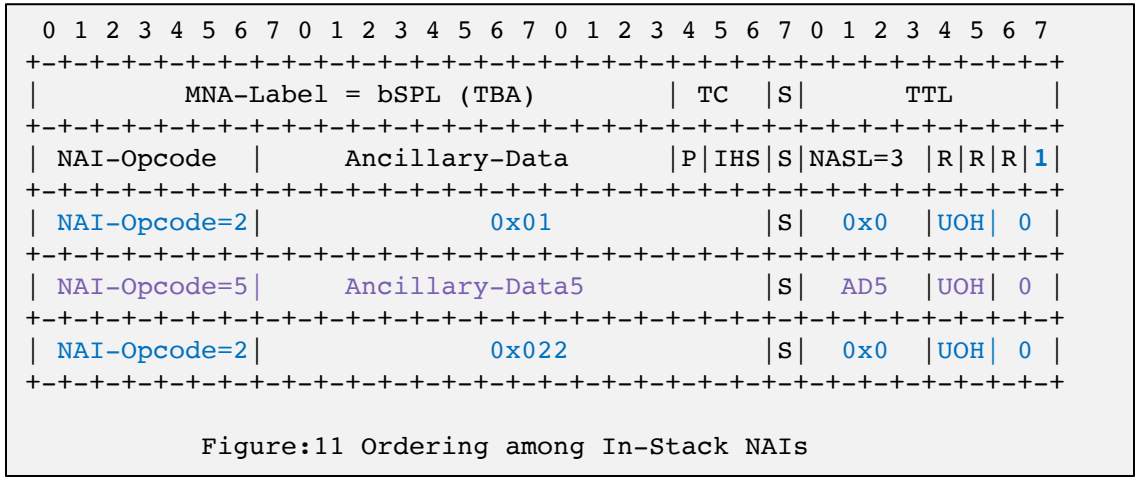


Figure:11 Ordering among In-Stack NAIs

Ordering Among In-Stack NAI with AD and without AD: In this case some of the Flag-Based NAI are required to be processed before example NAI-Opcode “5” and some are required to be processed after example NAI-Opcode “5”.

NETWORK ACTION ORDERING – (2/2)

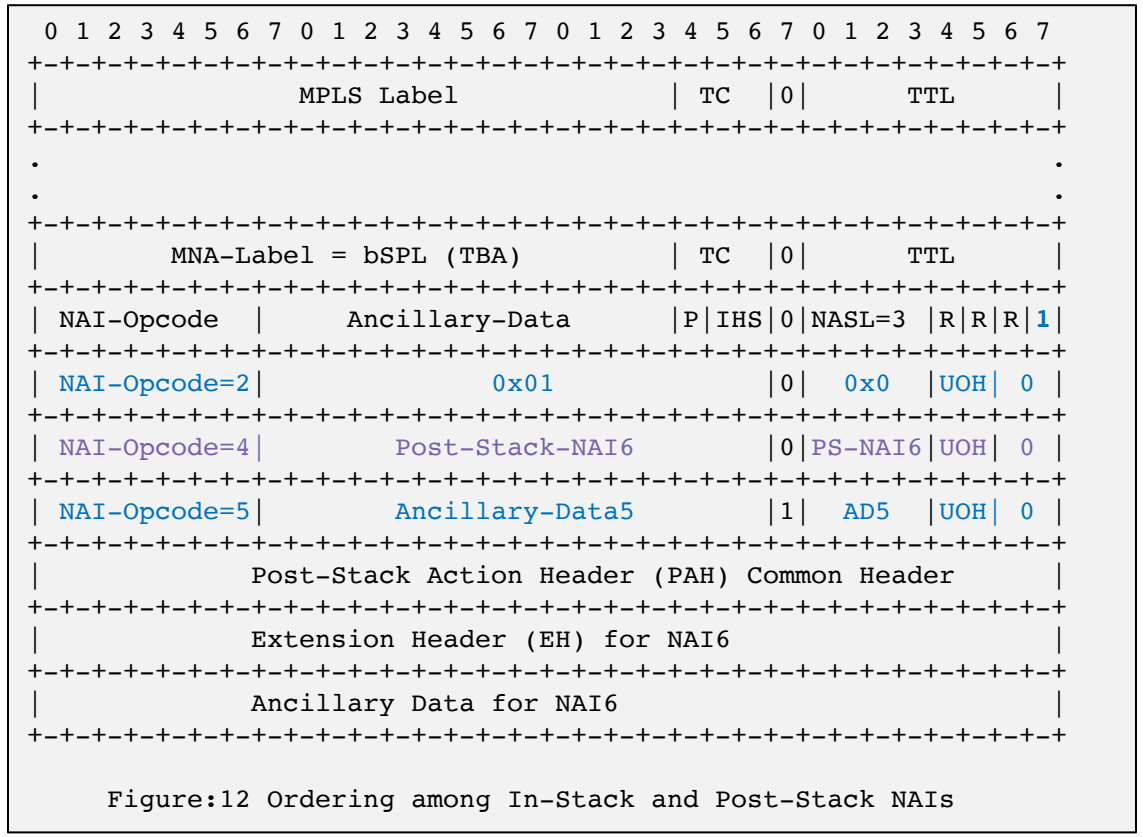


Figure:12 Ordering among In-Stack and Post-Stack NAIs

Ordering Among In-Stack and Post-Stack NAIs : A reserved opcode “4” is used to indicate the order of Post-Stack NAI processing with respect to the In-Stack NAI. In this example, Post-Stack NAI “6” is processed before In-Stack Opcode “5”.

BACKWARDS COMPATIBILITY

1. Node that needs to process MNA MUST signal the MNA capability
    a. If the MNA Label is exposed at the top on the received packet on the non-supporting node, the packet will be dropped
2. Penultimate node (or segment routing endpoint node) with TTL propagation behavior does not corrupt the NASS encoding parameters as they are added in the 2nd LSE
3. Penultimate node that does not recognize the exposed MNA bSPL after forwarding label lookup does not drop packets
4. ECMP behavior is not adversely affected by the MNA encoding
    a. Hashing for ECMP may use label fields in the LSEs in label stack. Label fields in the LSEs in NASS do not change for a flow
5. Label fields in LSEs in the NASS do not alias with an existing reserved bSPL
6. Can co-exist with existing G-ACh
    a. Opcode to specify the byte-offset for Post-Stack Network Action after the BOS

ADVANTAGES

1. MNA encoding follows the MNA framework and requirements - both In-Stack and Post-Stack Data
2. MNA encoding has **flexibility** to encode Network Actions in a desired order
3. MNA encoding is **extensible** by defining new Network Action opcodes and Flag-Based LSEs
    a. Easy to extend for user-defined Network Actions
4. MNA encoding is **hardware parser friendly**
    a. Length field of the MNA sub-stack allows to easily skip sub-stack
    b. Network Action Opcode, Ancillary Data and Length in a single LSE are easy to process in hardware
        - Network Action Length allows to carry Ancillary Data of different lengths (including > 20-bits)
        - Network Action Length allows node to easily skip only the not-supported Network Actions
5. **MSD efficient** encoding when enabling Network Action as only needed NAI opcodes are added
6. Backwards compatible in an existing network
7. MNA encoding is **ECMP friendly**
    a. Labels in the label stack does not change for a flow
    b. Encoding allows to change AD in the TTL field for a flow if there is such a need

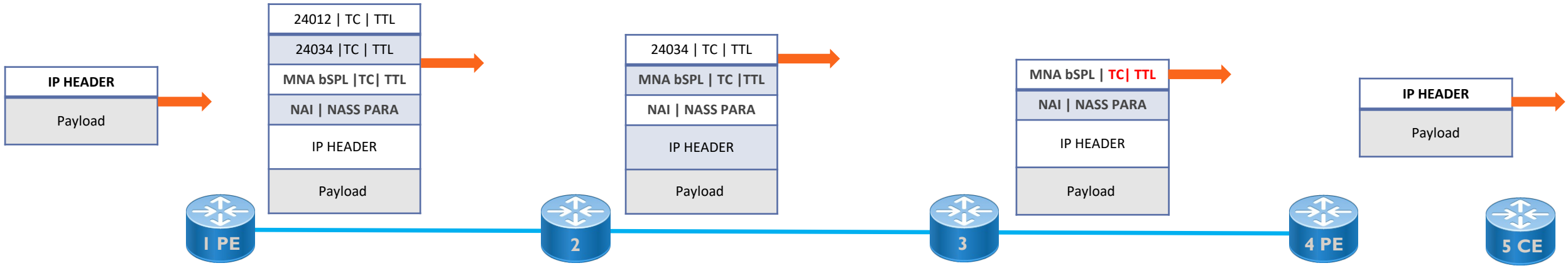
NEXT STEPS

- Welcome review comments and feedbacks
- Address pending review comments, some cleanup
- Clarify more on the Select Scope and Ordering

THANK YOU!

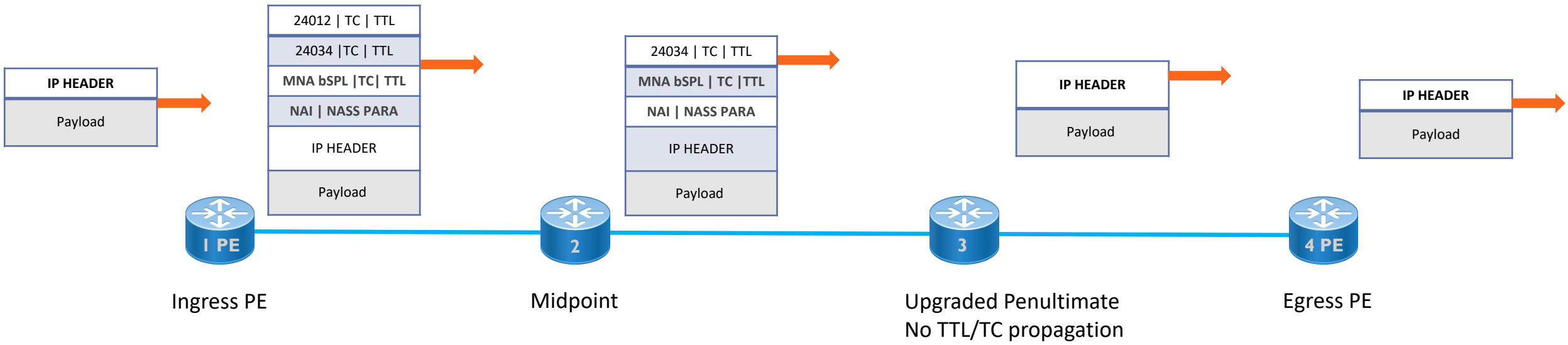
APPENDIX

PACKET WITH MNA – LEGACY PENULTIMATE



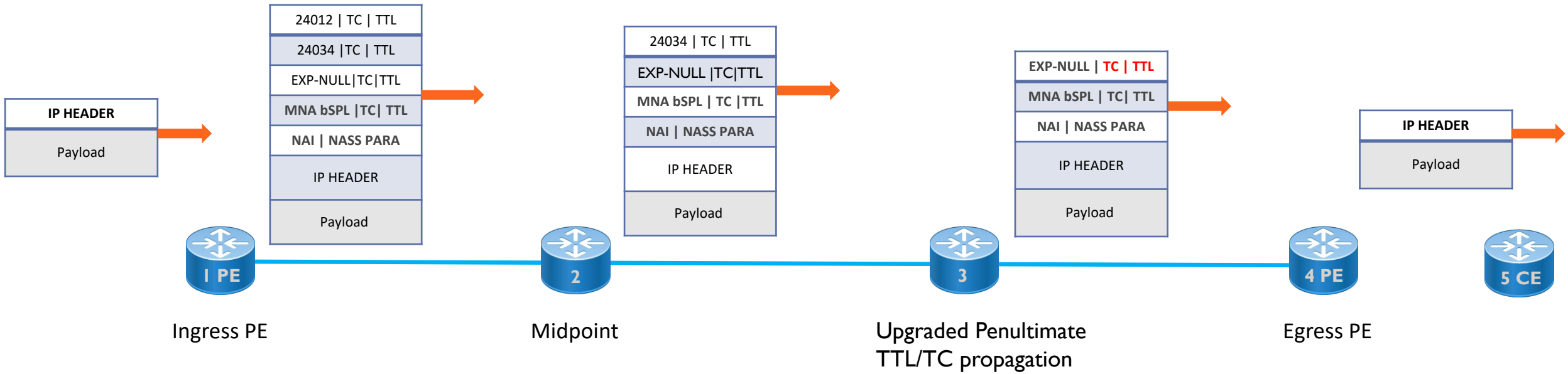
1. Egress / Segment Endpoint node:
   - NASS needs to be received at the egress node to perform the Network Action.
   - Egress node needs to remove NASS in case it is at the top of the received packet or it is exposed after popping the label and not send it towards CE.
2. Penultimate:
   a. Penultimate legacy node that does not recognize the MNA bSPL will not pop the NASS after popping the forwarding label.
   b. Penultimate node recognizes the MNA bSPL after popping the forwarding label and removes the NASS.
      - NASS is not reached to egress PE
      - EXP-NULL label can be added to avoid PHP popping the NASS. This is not an issue when VPN label is used.
      - Capability can be used for the PHP node to know if EXP-NULL label needs to be added.

PACKET WITH MNA – UPGRADED PENULTIMATE



- Penultimate node recognizes the MNA bSPL after popping the forwarding label and removes the NASS.
- NASS is not reached to Egress PE

PACKET WITH MNA – UPGRADED PENULTIMATE - EXP-NULL/VPN LABEL



- EXP-NULL label can be added to avoid PHP popping the NASS. This is not an issue when VPN label is used.
- Capability can be used for the PHP node to know if EXP-NULL label needs to be added.

NETWORK ACTION ENCODING EXAMPLES- 1/3

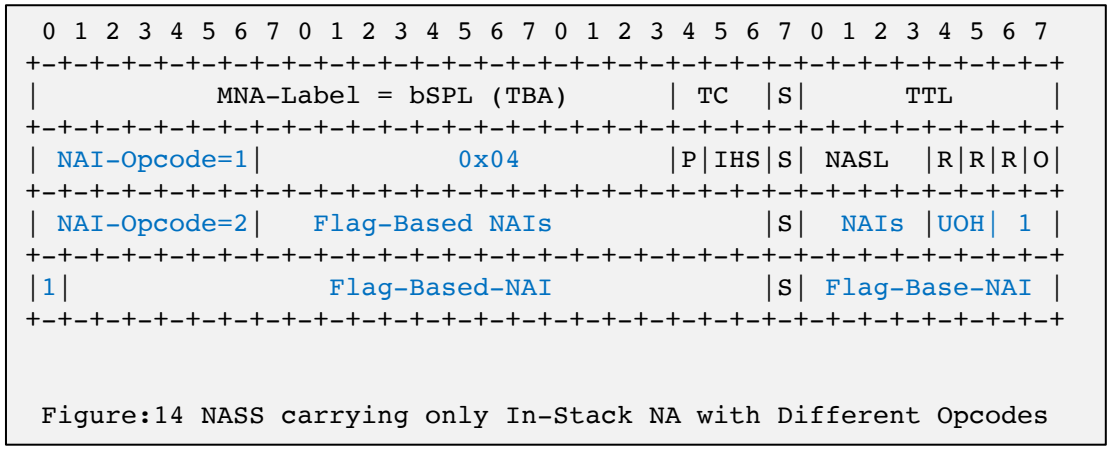


Figure:14 NASS carrying only In-Stack NA with Different Opcodes

NASS Carrying In-Stack NA with 12 bits AD and NA with more than 20-bits of AD:

NETWORK ACTION ENCODING EXAMPLES - 2/3

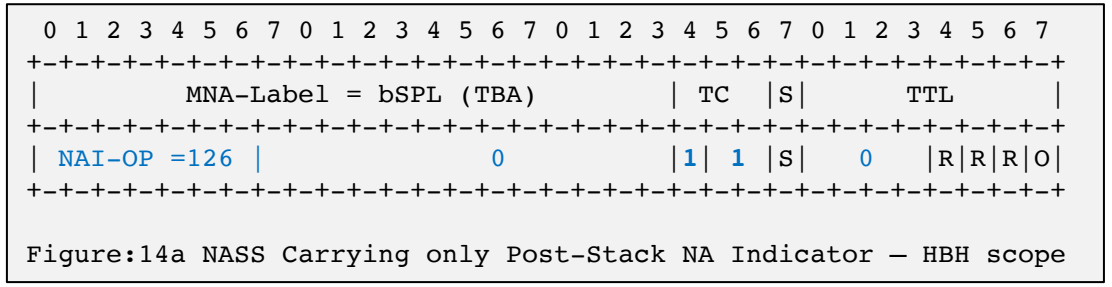


Figure:14a NASS Carrying only Post-Stack NA Indicator – HBH scope

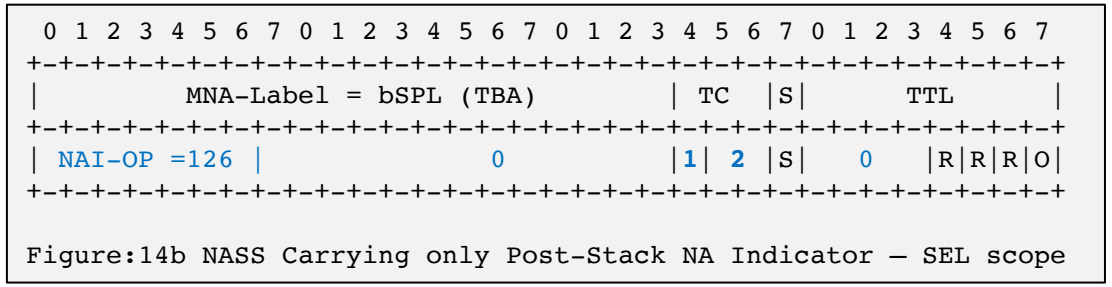


Figure:14b NASS Carrying only Post-Stack NA Indicator – SEL scope

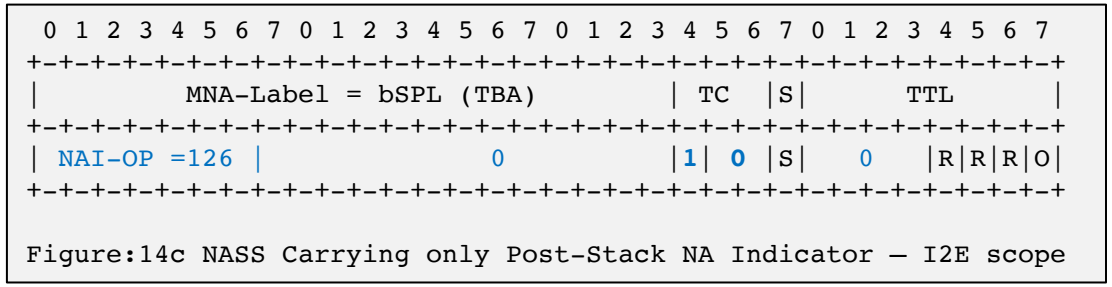


Figure:14c NASS Carrying only Post-Stack NA Indicator – I2E scope

NASS Carrying Only Post-Stack NA Indicator:

- P-Bit is set and the IHS scope value is set

NETWORK ACTION ENCODING EXAMPLES- 3/3

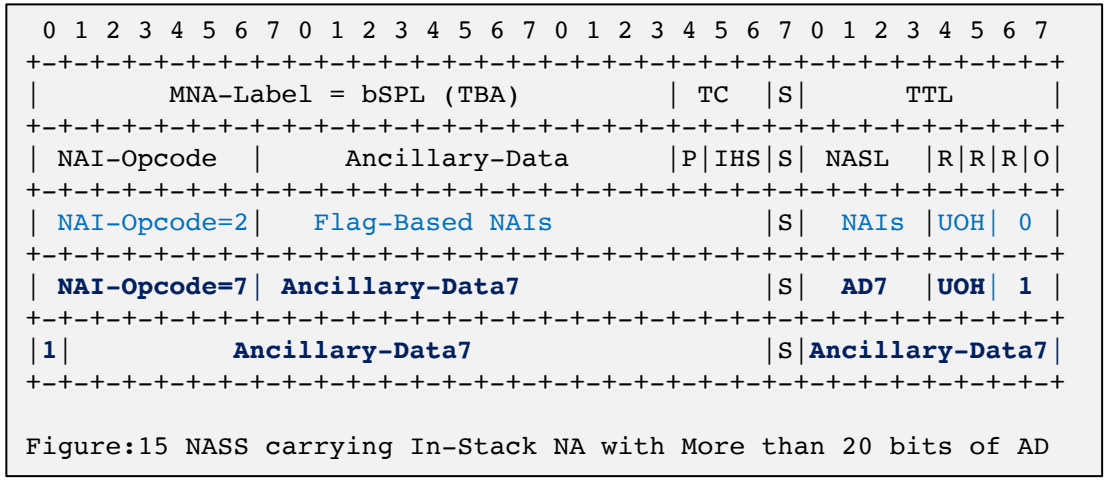


Figure:15 NASS carrying In-Stack NA with More than 20 bits of AD

NASS Carrying In-Stack NA with More than 20 Bits of AD:

- In this example, In-Stack NA with opcode "7" carried more than 20 bits of AD.

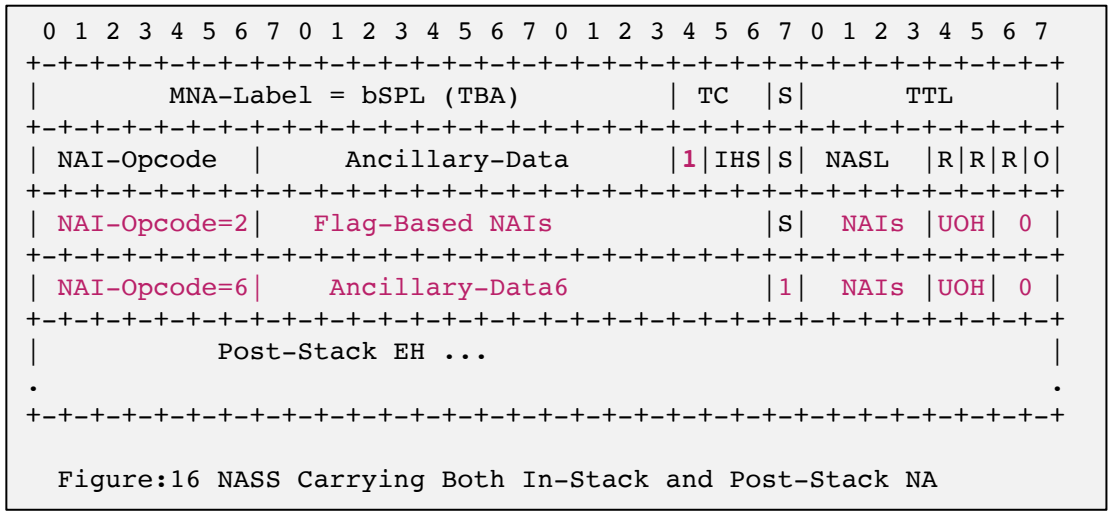


Figure:16 NASS Carrying Both In-Stack and Post-Stack NA

NASS Carrying both In-Stack and Post-Stack Network Actions:
P-Bit is set to indicate the presence of Post-Stack Data and the In-Stack Opcode "2" and "6" are encoded.

THANK YOU!