



Inband Network Telemetry Extensions

Ramesh Sivakolundu
Cisco Systems Inc.



C97-738306-01 © 2017 Cisco and/or its affiliates. All rights reserved.

Inband Network Telemetry (INT)

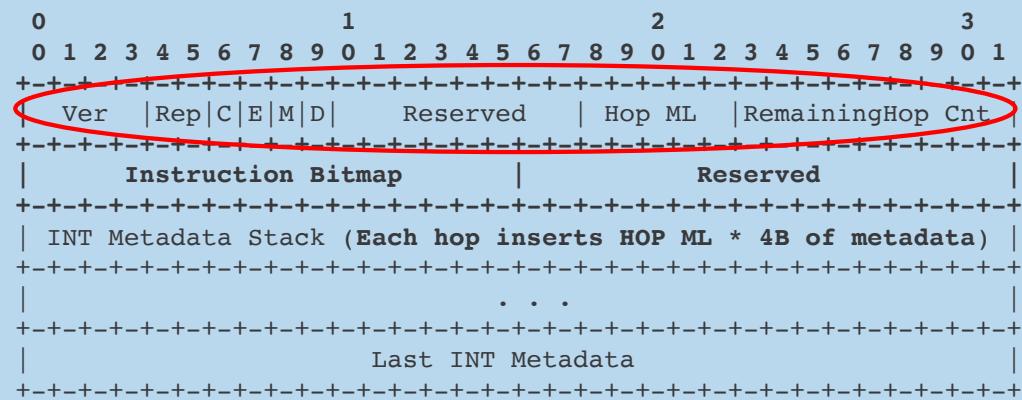
(Version 2, proposed)

INT Metadata Header

- **Ver (4b)**: INT Metadata header version. Should be 1 for this version.
- **Rep (2b)**: Replication Requested. Support is optional. Useful to explore all the valid physical forwarding paths (ECMP, LAG).
- **C (1b)**: Copy. If replication requested for data packets, the INT sink must be able to distinguish original packets from replicas.
- **E (1b)**: Max Hop Count Exceeded. This flag is set if a device cannot prepend its own metadata due to Remaining Hop Count reaching zero.
- **M (1b)**: MTU Exceed. This flag must be set if a device cannot add all of the requested metadata because doing so will cause the packet length to exceed egress link MTU.

• **D (1b): Domain Specific Flag**

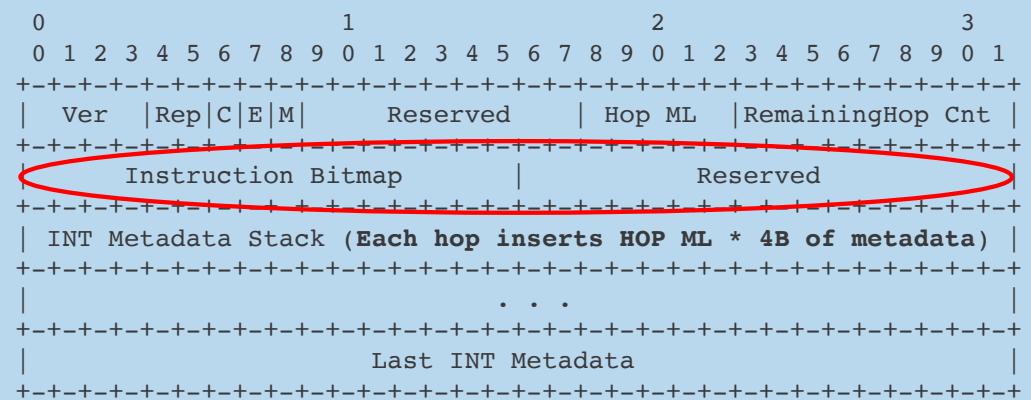
- **R (7b)**: Reserved.
- **Hop ML (5b)**: Per-hop Metadata Length, the length of metadata in 4-Byte words to be inserted at each INT hop.
- **RemainingHopCnt (8b)**: The remaining number of hops that is allowed to add their metadata to the packet.



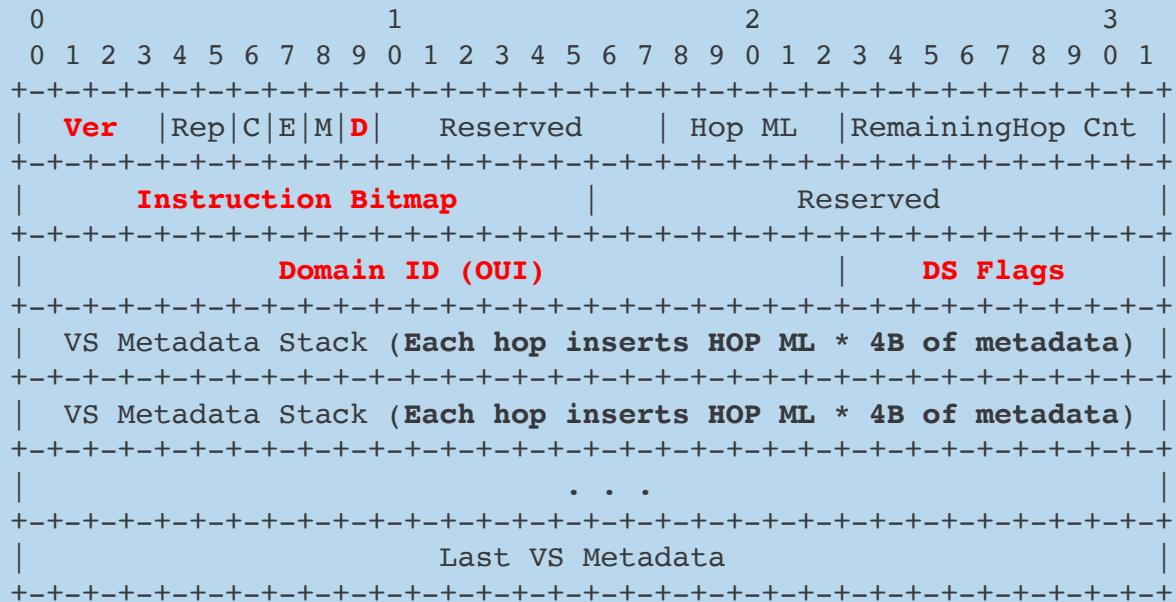
INT Metadata Header – Instruction Bitmap

INT instructions are encoded as a bitmap in the 16 bit INT Instruction field: each bit corresponds to a specific standard metadata.

- bit0 (MSB): Switch ID
- bit1: Level 1 Ingress port ID (16 bits) + Egress Port ID (16 bits)
- bit2: Hop latency
- bit3: Queue ID (8 bits) + Queue occupancy (24 bits)
- bit4: Ingress timestamp
- bit5: Egress timestamp
- bit6: Egress port ID
- bit7: Level 2 Ingress port ID (32 bits) + Egress Port ID (32 bits)
- bit8: Egress port TX utilization
- **Bit 14: Domain Specific Metadata**
- bit15: Checksum Complement
- The remaining bits are reserved.



Metadata stack /w Vendor Specific Data

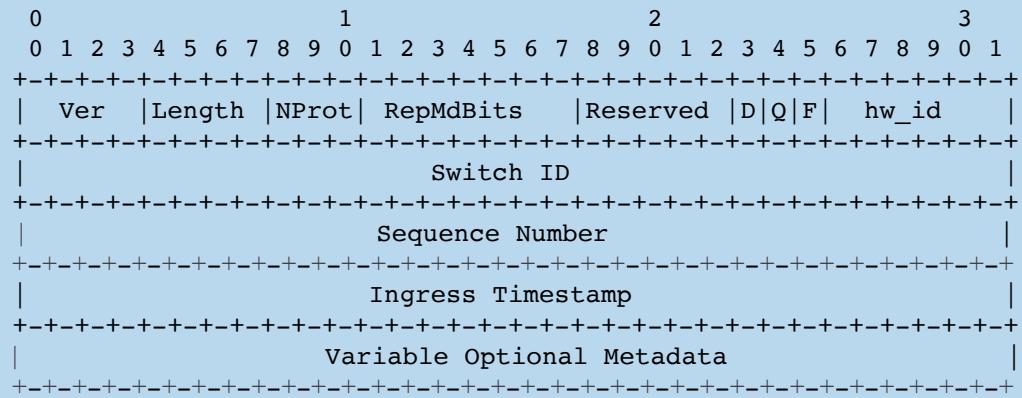


- **Ver (4b):** INT Metadata header version. Should be **2** for this version.
- **D (1b): Domain Specific Flag**
- **R (8b):** Reserved.
- **Hop ML (5b):** Per-hop Metadata Length, the length of metadata in 4-Byte words to be inserted at each INT hop, **including the Domain Specific Metadata**.
- **RemainingHopCnt (8b):** The remaining number of hops that is allowed to add their metadata to the packet.
- **Domain ID (24b): Domain ID, the unique OUI of the vendor**
- **DS Flags (8b): Domain Specific Flags.**

INT - Telemetry Report Export

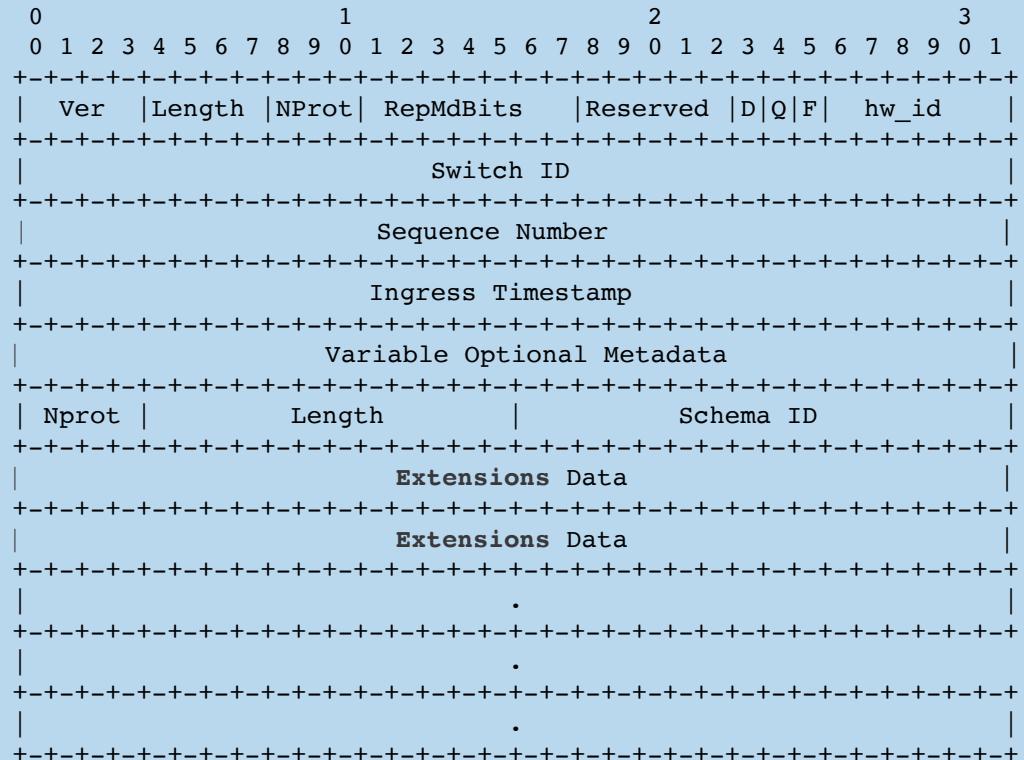
Telemetry Report Header

Telemetry Report Header (16+ Octets)



- **Length:** Indicates length of report header in multiple of 4 octets, including the optional metadata.
- **Nprot: Next Protocol**
 - 0 – Ethernet
 - 1 – IPv4
 - 2 – IPv6
 - 3 – Telemetry Extensions
- **RepMdBits: Report Metadata Bits** (each bit represents 4 octets of optional metadata)
 - 0 – Ingress ID (16 bits) & Egress Port ID (16 bits)
 - 1 – Hop Latency
 - 2 – Queue ID (8 bits) + Queue Occupancy (24 bits)
 - 3 – Egress Timestamp (32 bits)
 - 4 – Queue ID (8 bits) + Drop Reason (8 bits) + Padding (16 bits)
 - 5 – Egress Port TX Utilization
- **D** – Dropped – Indicates at least one packet was dropped.
- **Q** – Congested Queue
- **F** – Tracked Flow Association - indicated report is for a tracked flow.
- **hw_id:** Hardware subsystem within the source that generated this report
- **Switch ID:** The unique ID of a switch (administratively assigned).
- **Sequence Number:** Reflects the sequence of reports from a specific hw_id
- **Ingress Timestamp:** Device local time when packet was first received on the ingress physical or logical port, in nanoseconds

Telemetry Report Header /w Extensions



- **Nprot (3b): Next Protocol**
 - 0 – Ethernet
 - 1 – IPv4
 - 2 – IPv6
 - 3 – Telemetry Extensions
- **Length (13b):** Indicates length of telemetry extensions header in multiple of 4 octets.
- **Schema ID (16b):** 2-octet unsigned integer identifying the schema of the telemetry extensions data. The following ranges of Schema ID values are defined:
 - 0 – 0xFFFF: Available for private or experimental use
 - 0xF000 – 0xFEFF : Reserved
 - 0xFF00 – 0xFFFF : Reserved for specification by P4.org. This range is set aside for future uses such as specification of field units in INT and Telemetry Reports.
 - 0xFFFF : No schema, indicates that the telemetry extensions header has no content.
- **Extensions Data :** Variable length field. The field is interpreted as specified by schema identified by the Schema ID.

INT – IPv6 Extension Header

INT in IPv6

IPv6 HbyH header Option:

