OTTO VON GUERICKE
UNIVERSITÄT
MAGDEBURG

INF

FAKULTÄT FÜR
INFORMATIK

EURO P4
OCTOBER 28 2024

Secure In-Band Network Telemetry for the SCION Internet Architecture on Tofino

M.Sc. Robin Wehner
M.Sc. Lars-Christian Schulz

M.Sc. Tony John
Prof. Dr. David Hausheer

Structure

1. Background: SCION Architecture
2. Motivation
3. The ID-INT Protocol
4. P4 Implementation on Tofino
5. Performance Observations of ID-INT on Tofino
6. Evaluation of ID-INT on Specific Use Case
7. Conclusion & Future Work

1. Background: SCION Architecture

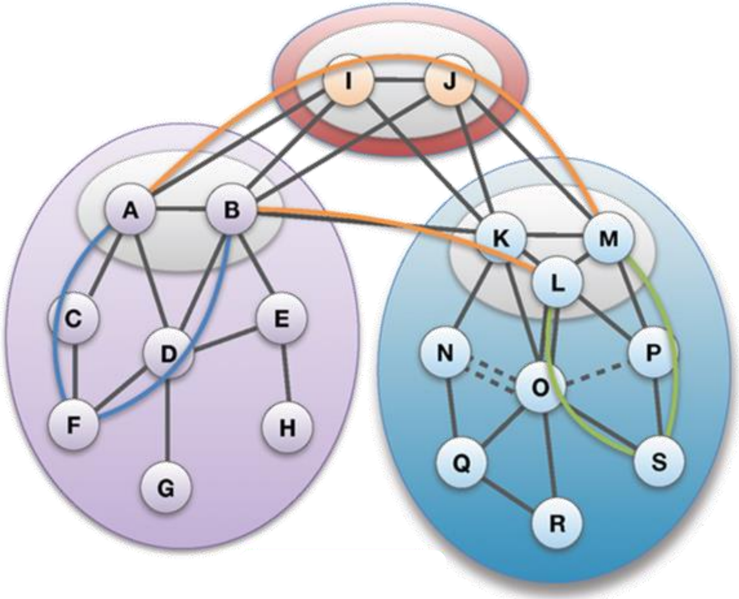
Path-based Network Architecture

Control Plane - Routing

- Constructs and Disseminates Path Segments

Data Plane - Packet forwarding

- Combine Path Segments to Path
- Packets contain Path
- Routers forward packets based on Path
  - Simple routers, stateless operation



SCION™

1. Background: SCION Architecture

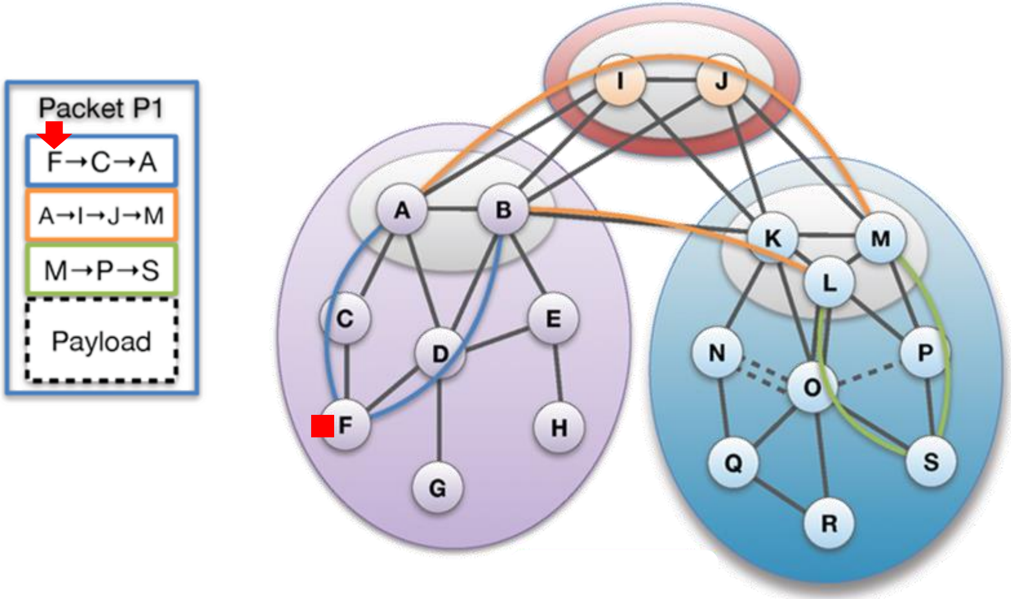
Path-based Network Architecture

Control Plane - Routing

- Constructs and Disseminates Path Segments

Data Plane - Packet forwarding

- Combine Path Segments to Path
- Packets contain Path
- Routers forward packets based on Path
  - Simple routers, stateless operation



Packet P1

F→C→A

A→I→J→M

M→P→S

Payload

SCION

1. Background: SCION Architecture

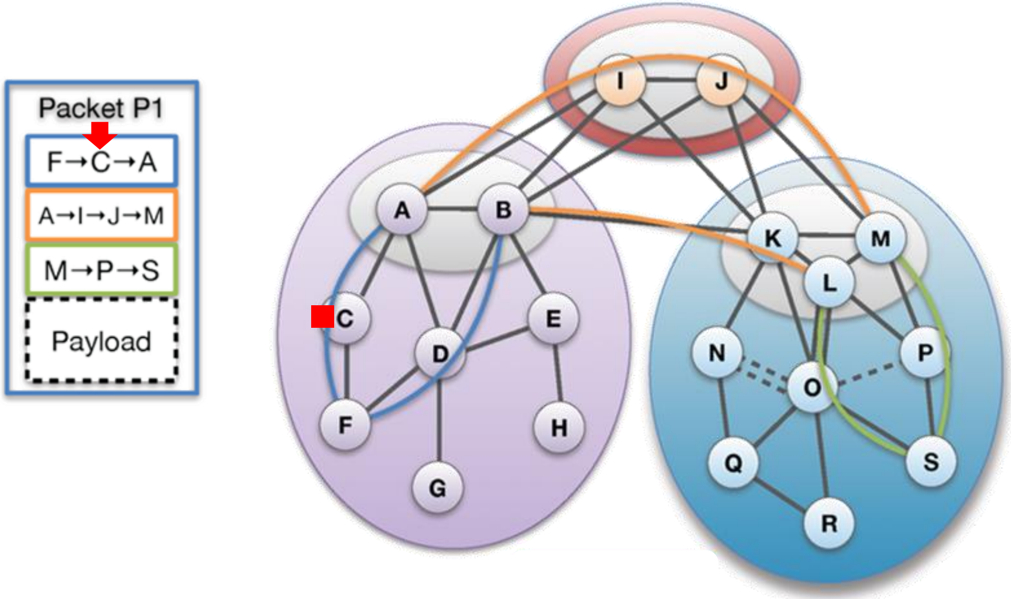
Path-based Network Architecture

Control Plane - Routing

- Constructs and Disseminates Path Segments

Data Plane - Packet forwarding

- Combine Path Segments to Path
- Packets contain Path
- Routers forward packets based on Path
  - Simple routers, stateless operation



Packet P1

F→C→A

A→I→J→M

M→P→S

Payload

SCION

1. Background: SCION Architecture

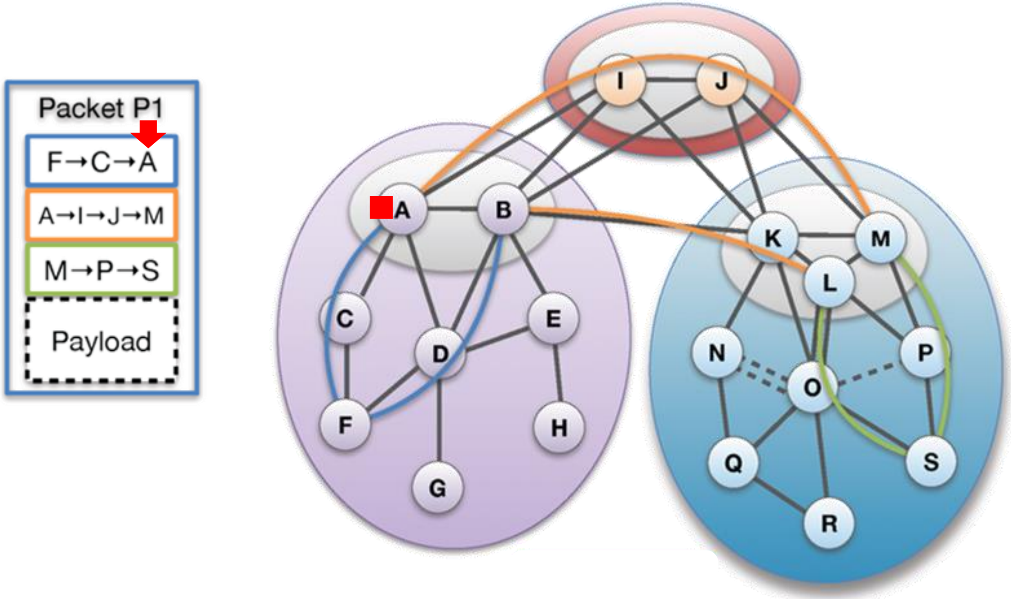
Path-based Network Architecture

Control Plane - Routing

- Constructs and Disseminates Path Segments

Data Plane - Packet forwarding

- Combine Path Segments to Path
- Packets contain Path
- Routers forward packets based on Path
  - Simple routers, stateless operation



SCION

1. Background: SCION Architecture

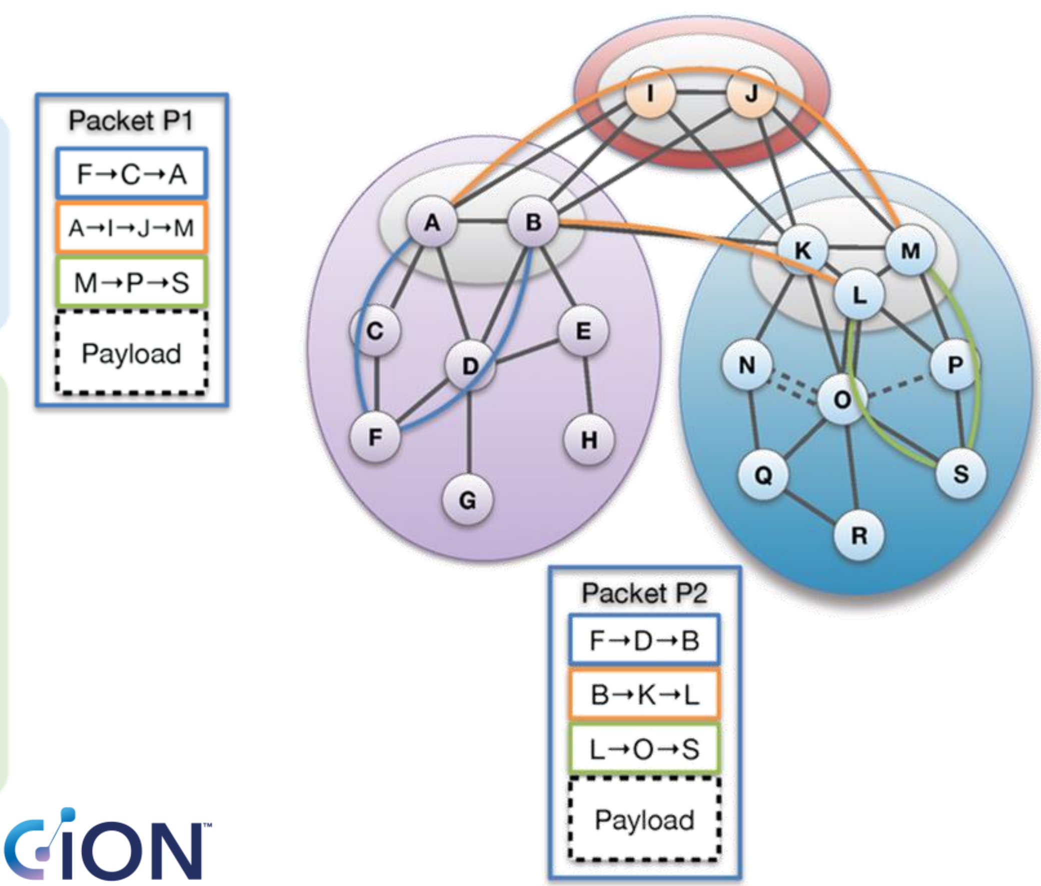
Path-based Network Architecture

Control Plane - Routing

- Constructs and Disseminates Path Segments

Data Plane - Packet forwarding

- Combine Path Segments to Path
- Packets contain Path
- Routers forward packets based on Path
  - Simple routers, stateless operation



SCION™

2. Motivation

- SCION is a path-aware Internet architecture
  - Challenge: How to select an appropriate path?

2. Motivation

- SCION is a path-aware Internet architecture
  - Challenge: How to select an appropriate path?

- General Approach: End-2-end measurements
  - Useful for latency, bandwidth, jitter, etc.
  - Unprecise information on hop-by-hop latency
  - No information on internal router state (e.g. queue length)
  - Insufficient for certain applications that require more detailed network information to optimize path selection (e.g. congestion control)

2. Motivation

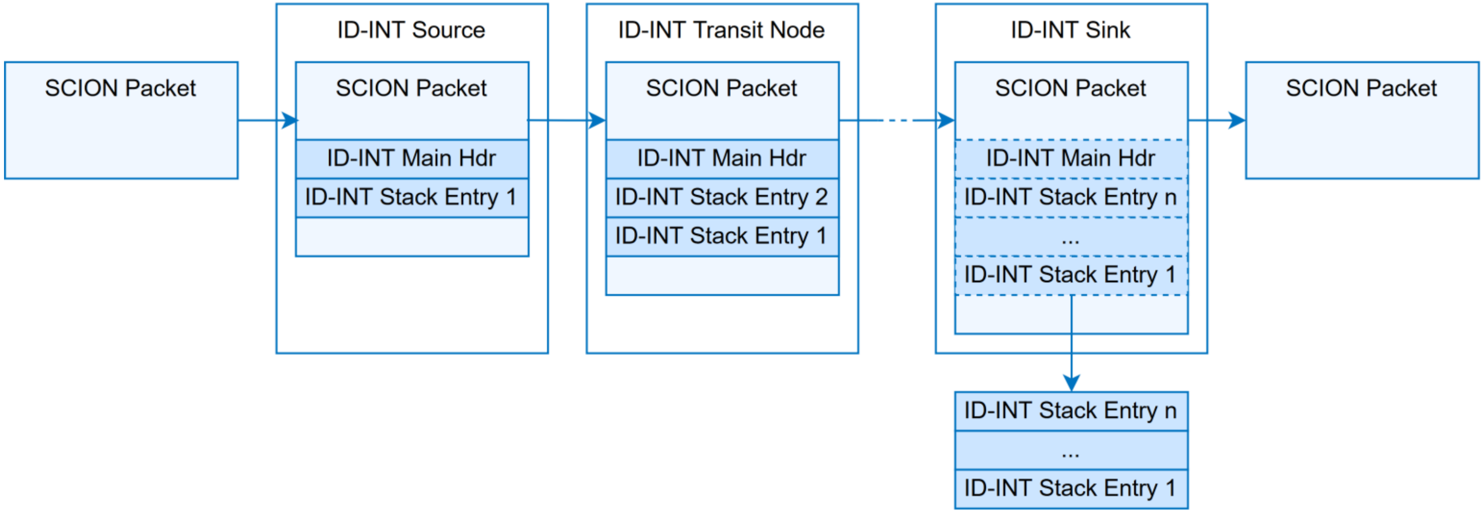
- ❖ SCION is a path-aware Internet architecture
  - ➢ Challenge: How to select an appropriate path?

✓ Our Proposed Approach: Inter-Domain In-band Network Telemetry (ID-INT)

- ➢ Offers fine-granular metadata from inside the network
- ➢ Implemented on Tofino
  - ➢ Hardware offers very precise network metadata (e.g. queue length)
  - ➢ High total throughput in combination with SCION border router

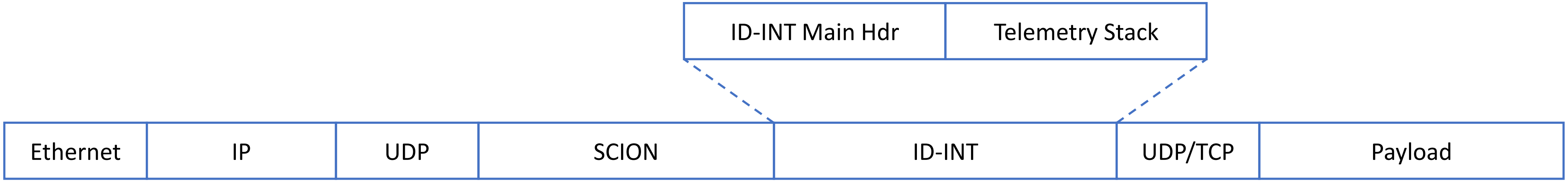
3. ID-INT Design

- Based on P4.org INT standard's INT-MD operational mode
- Extends standard INT to support inter-domain environments by adding verifiable MACs to each stack entry
- Leverages the capabilities of SCION PKI and DRKey



3. ID-INT Header Design

- SCION extension that is inserted after SCION headers

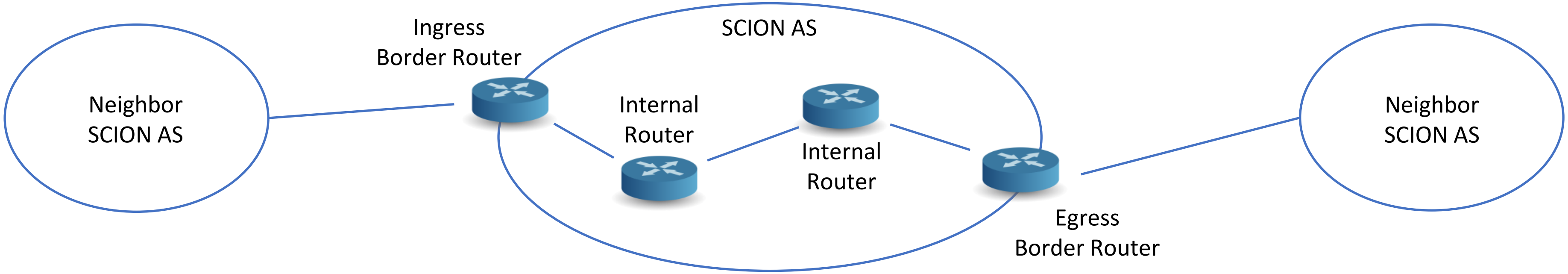


ID-INT Main Hdr | Telemetry Stack

Ethernet | IP | UDP | SCION | ID-INT | UDP/TCP | Payload

3. ID-INT Header Design

- Supports aggregation of metadata, fixed and flexible metadata



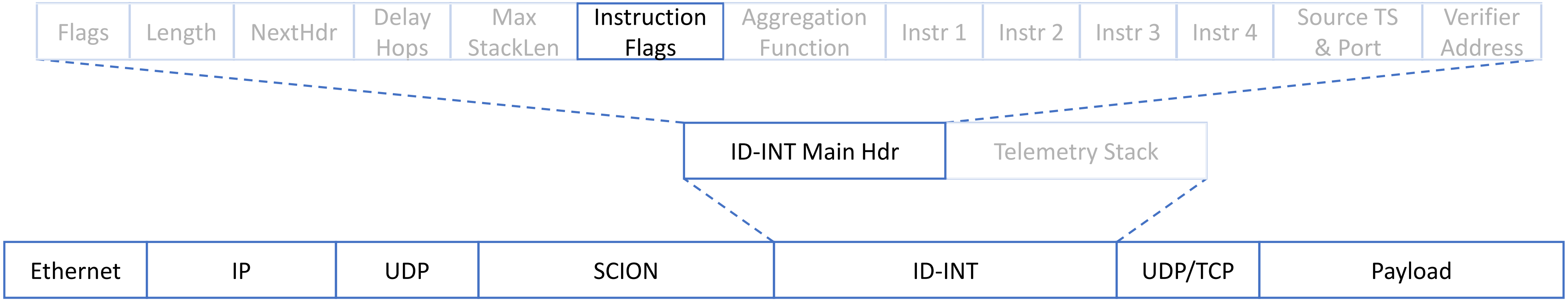
Flags	Length	NextHdr	Delay Hops	Max StackLen	Instruction Flags	Aggregation Function	Instr 1	Instr 2	Instr 3	Instr 4	Source TS & Port	Verifier Address

ID-INT Main Hdr	Telemetry Stack

Ethernet	IP	UDP	SCION	ID-INT	UDP/TCP	Payload

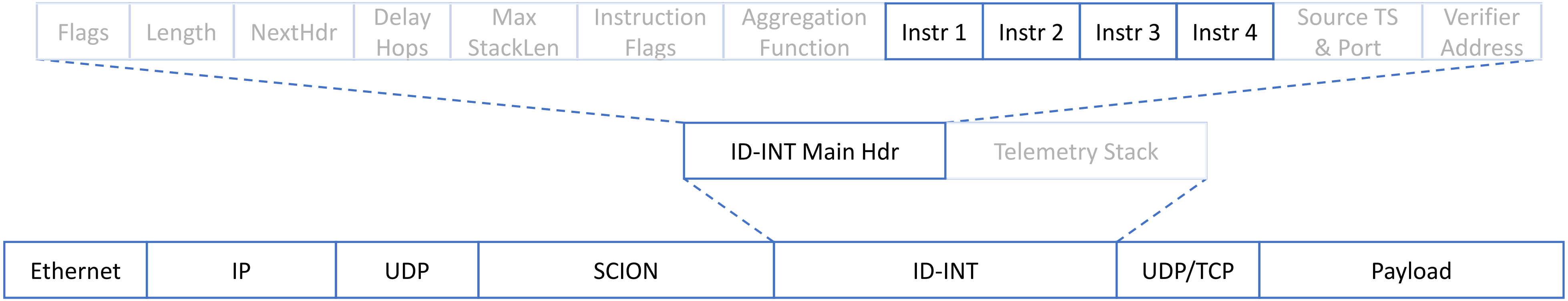
3. ID-INT Header Design

- Supports aggregation of metadata, fixed and flexible metadata
  - **Node Localization:** Node ID, Node Count, Ingress/Egress Interface ID



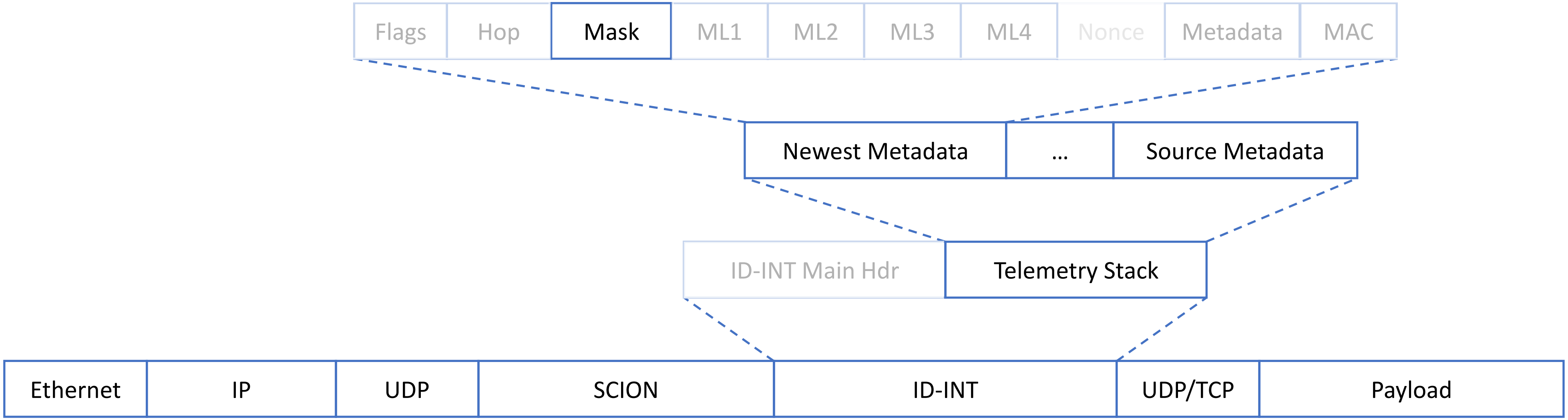
3. ID-INT Header Design

- Supports aggregation of metadata, fixed and flexible metadata
  - **Node Telemetry:** Timestamps, Queue ID, Instantaneous Queue Lengths, Ingress Port, *Device Type, Fan speed, Total Power Draw, Ingress/Egress Interface Speed, Uptime, Ingress/Egress SCION Interface Packet/Drop/Byte Count, Ingress/Egress Total Packet/Drop/Byte Count, …*



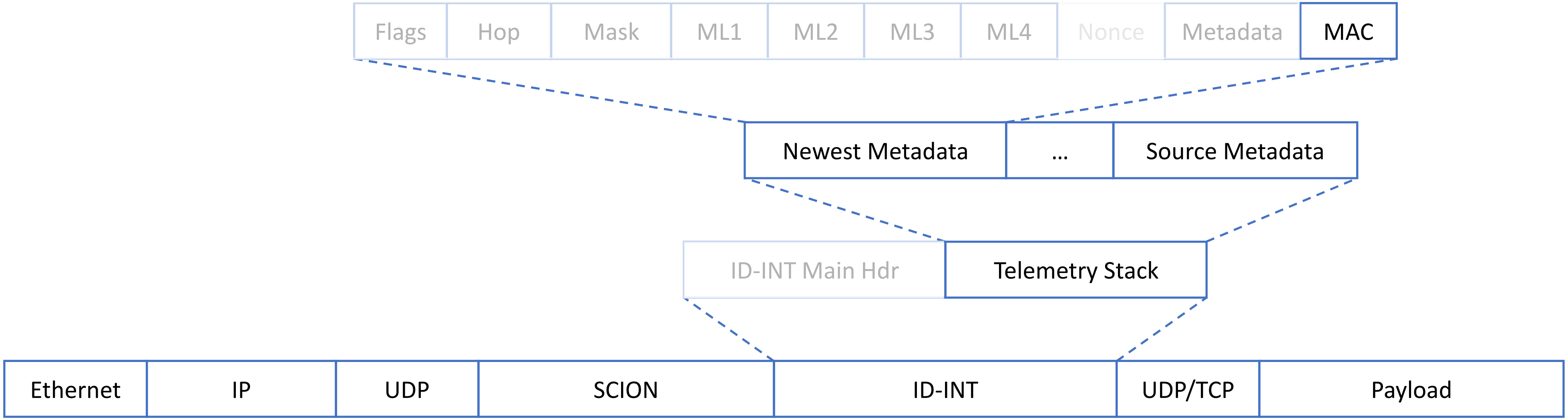
3. ID-INT Header Design

- Unsupported metadata is not added to the stack
- Stack entry size 8-64 Byte
- Telemetry entries are cryptographically secured by AES-CMAC



3. ID-INT Header Design

- Unsupported metadata is not added to the stack
- Stack entry size 8-64 Byte
- Telemetry entries are cryptographically secured by AES-CMAC



3. ID-INT Header Design

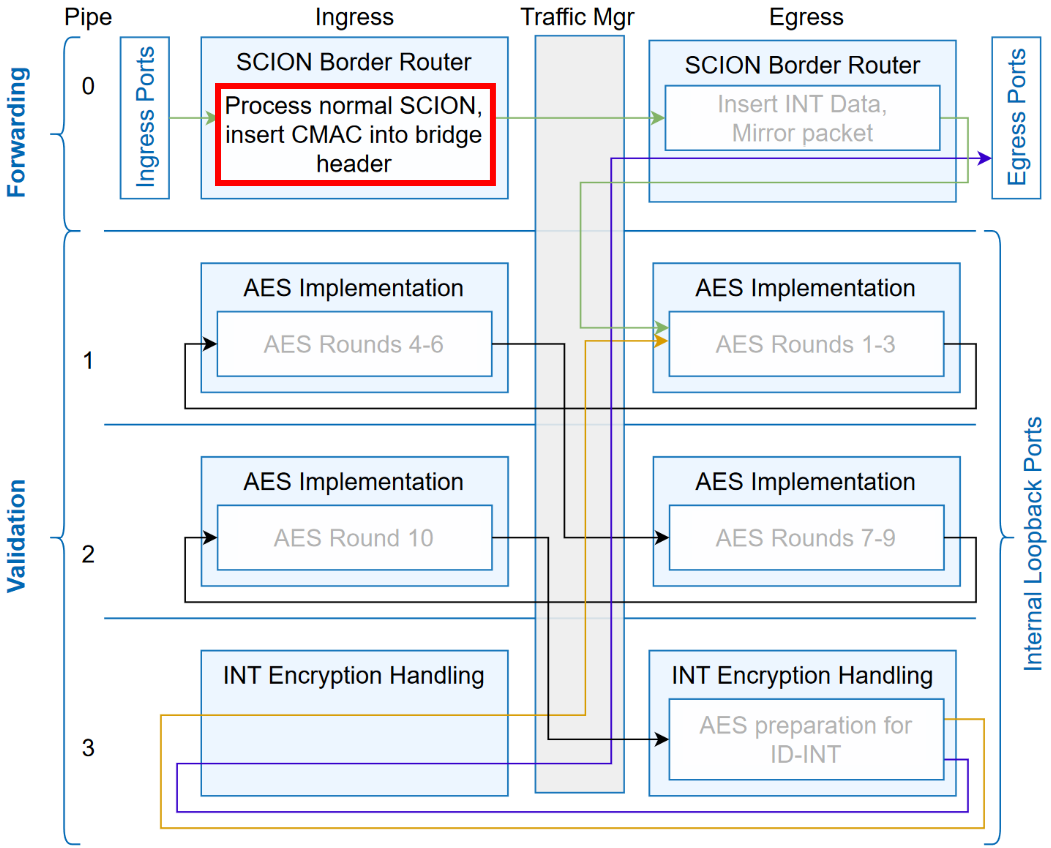
- Unsupported metadata is not added to the stack
- Stack entry size 8-64 Byte
- Telemetry entries are cryptographically secured by AES-CMAC
  - Derived with DRKey:
    - Used to create symmetrical AS-AS level keys in SCION
    - AES as pseudo-random function

4. P4 Implementation on Tofino

- Extends the SCION Border Router implemented in P4 for Tofino 2
- Tofino‘s pipes 0 & 3 manage SCION and ID-INT processing
- Tofino‘s pipes 1 & 2 do AES calculations
  - SCION Hop Field validation
  - ID-INT MAC calculation → Done with an AS-AS level key
  - DRKey derivation (uses AES as pseudo-random function in SCION)

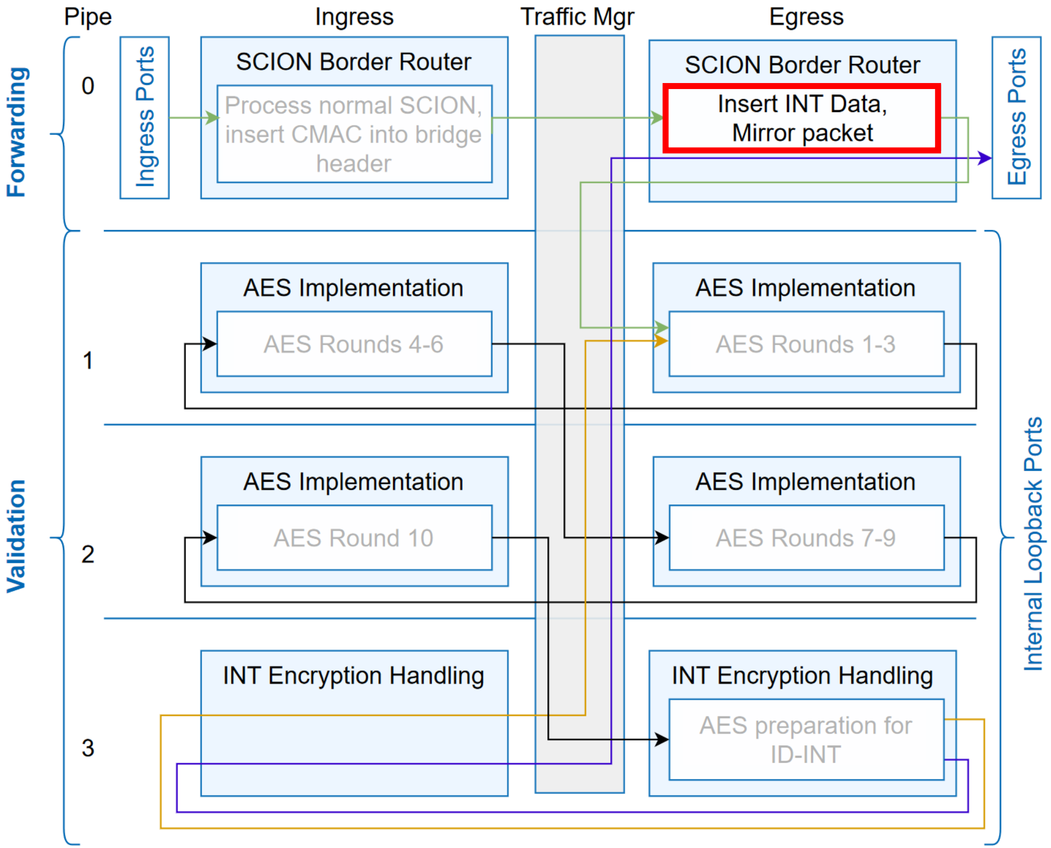
4. P4 Implementation on Tofino

- First process SCION Hop Field
- Process ID-INT
- Validate Hop Field MAC
- Perform DRKey
- Calculate MAC for ID-INT telemetry stack entry
- Forward packet to egress port



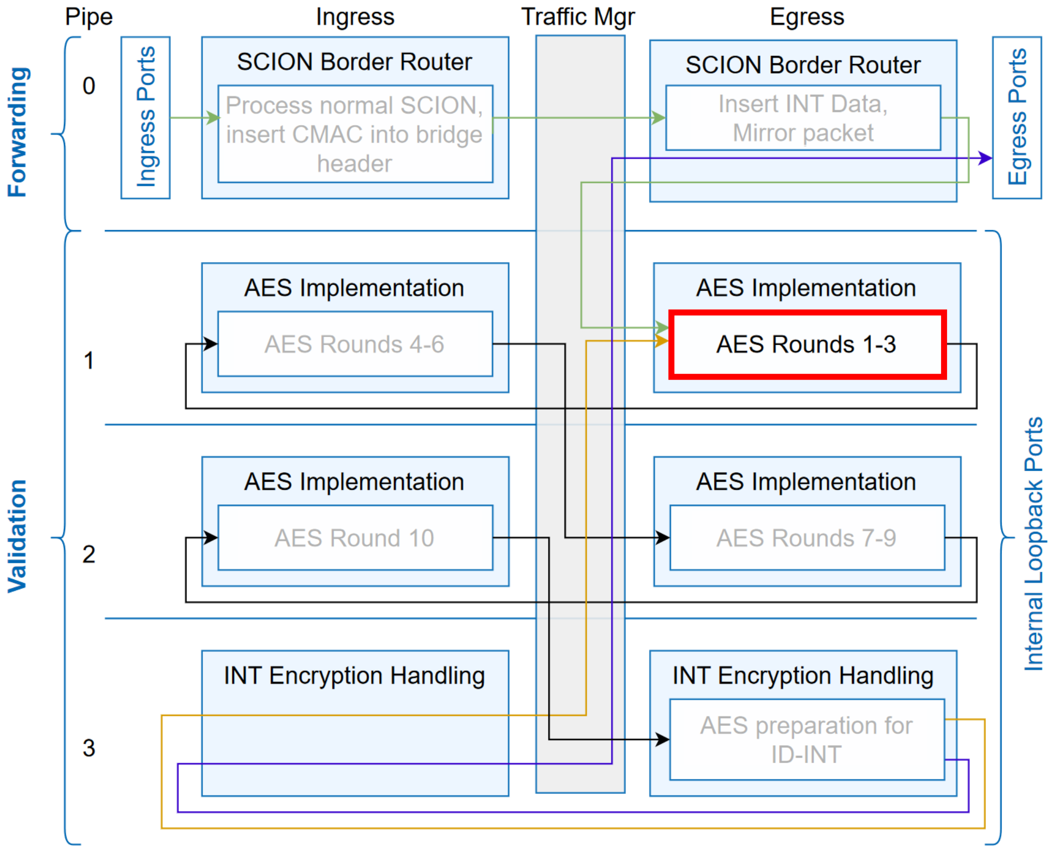
4. P4 Implementation on Tofino

- First process SCION Hop Field
- Process ID-INT
- Validate Hop Field MAC
- Perform DRKey
- Calculate MAC for ID-INT telemetry stack entry
- Forward packet to egress port



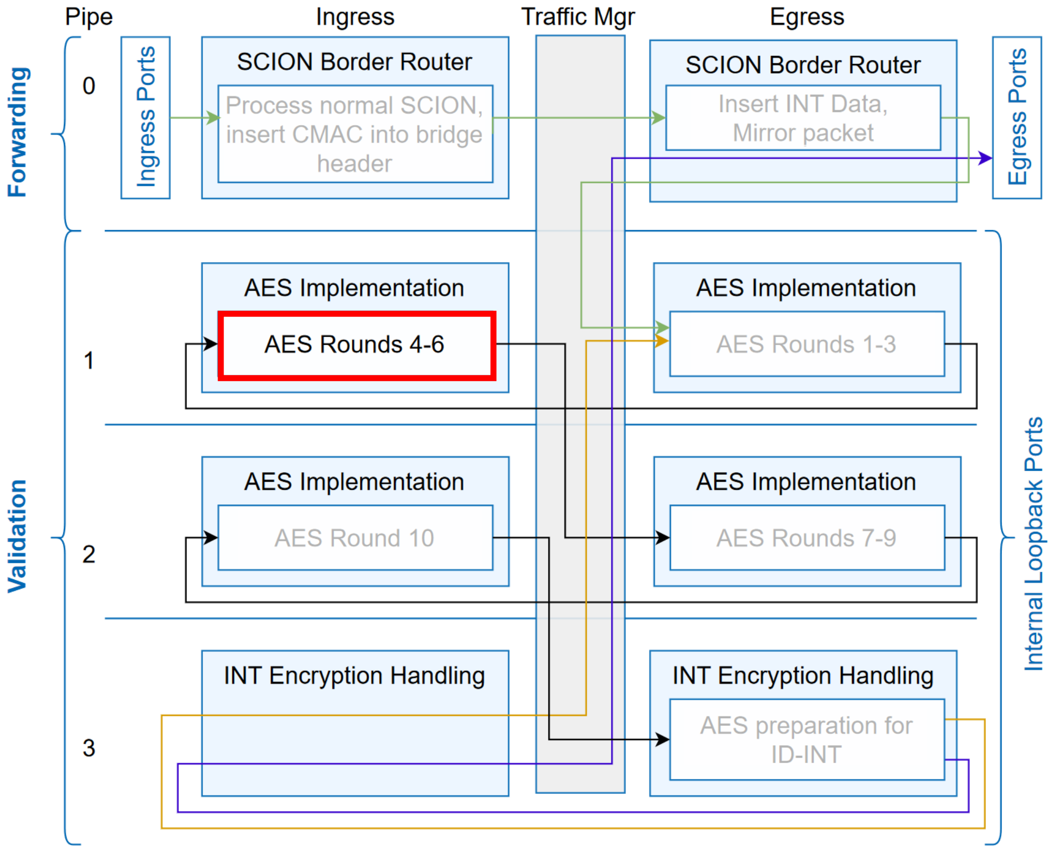
4. P4 Implementation on Tofino

- First process SCION Hop Field
- Process ID-INT
- Validate Hop Field MAC
- Perform DRKey
- Calculate MAC for ID-INT telemetry stack entry
- Forward packet to egress port



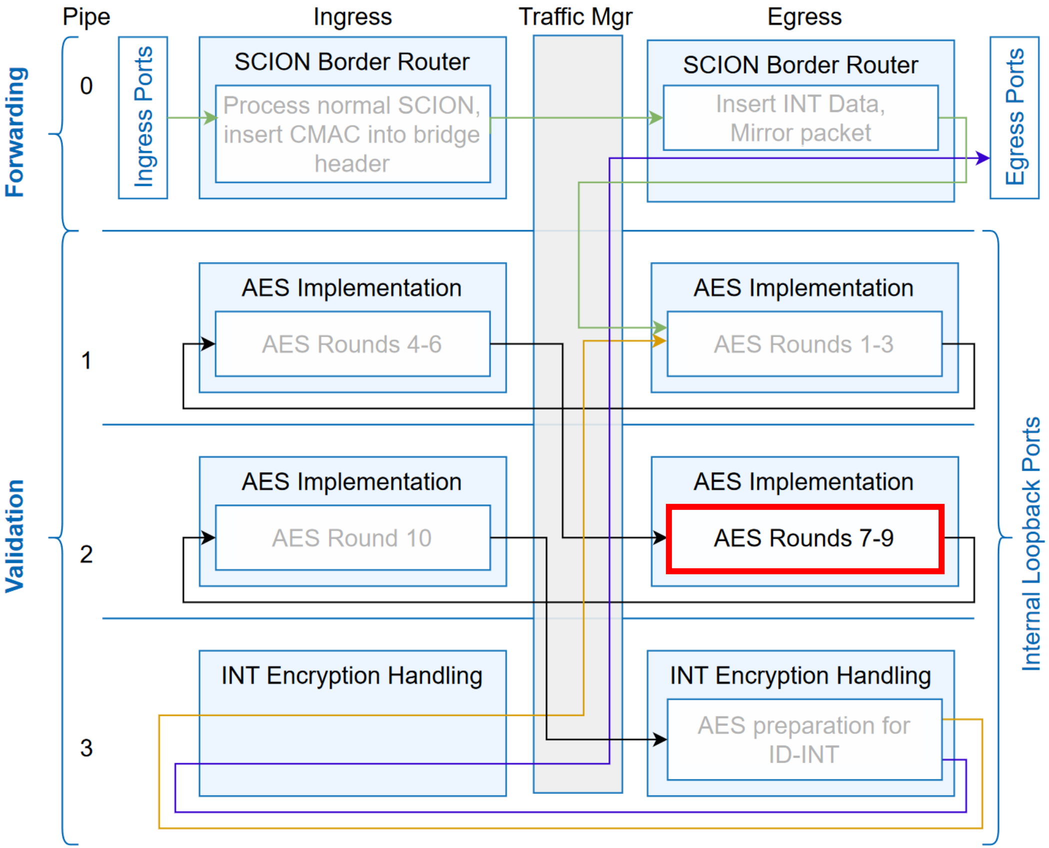
4. P4 Implementation on Tofino

- First process SCION Hop Field
- Process ID-INT
- Validate Hop Field MAC
- Perform DRKey
- Calculate MAC for ID-INT telemetry stack entry
- Forward packet to egress port



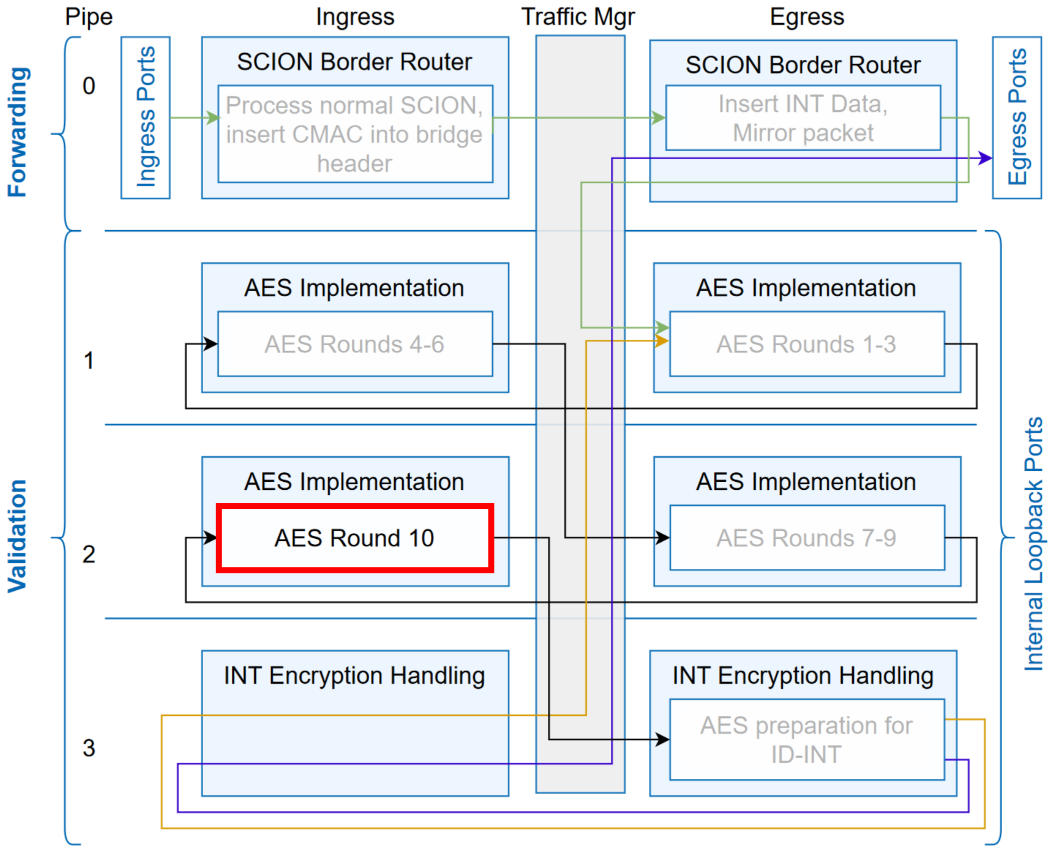
4. P4 Implementation on Tofino

- First process SCION Hop Field
- Process ID-INT
- Validate Hop Field MAC
- Perform DRKey
- Calculate MAC for ID-INT telemetry stack entry
- Forward packet to egress port



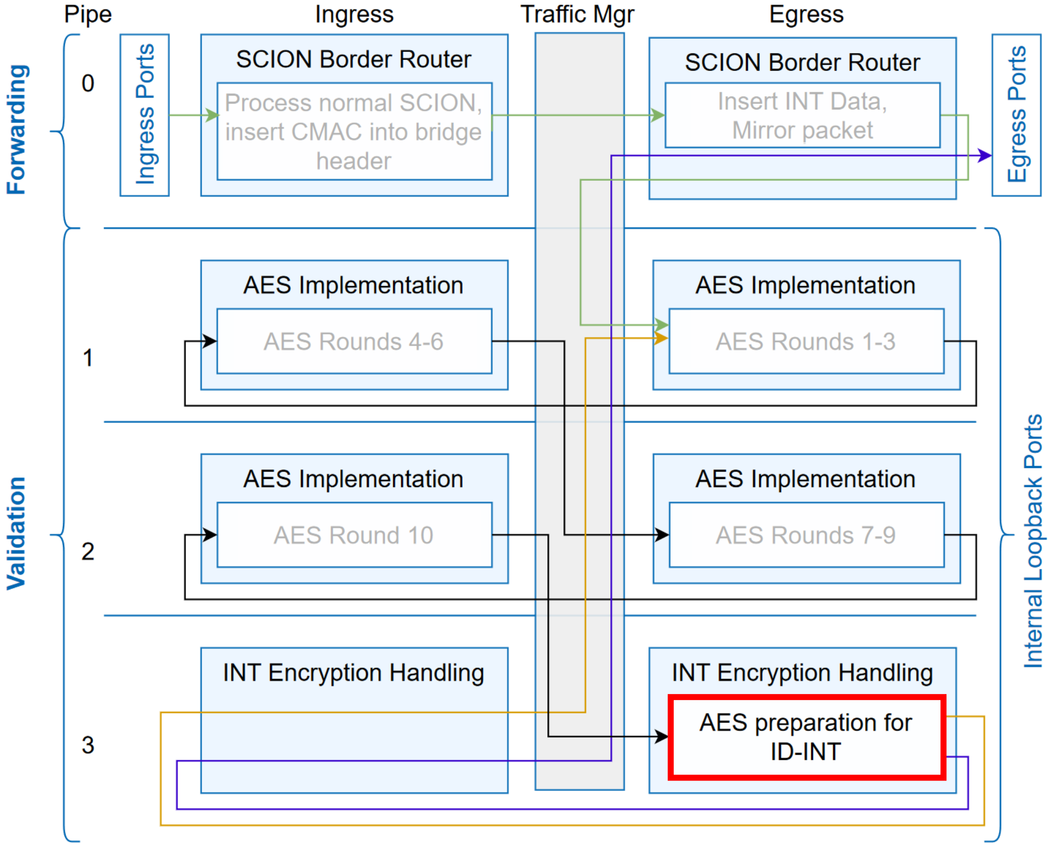
4. P4 Implementation on Tofino

- First process SCION Hop Field
- Process ID-INT
- Validate Hop Field MAC
- Perform DRKey
- Calculate MAC for ID-INT telemetry stack entry
- Forward packet to egress port



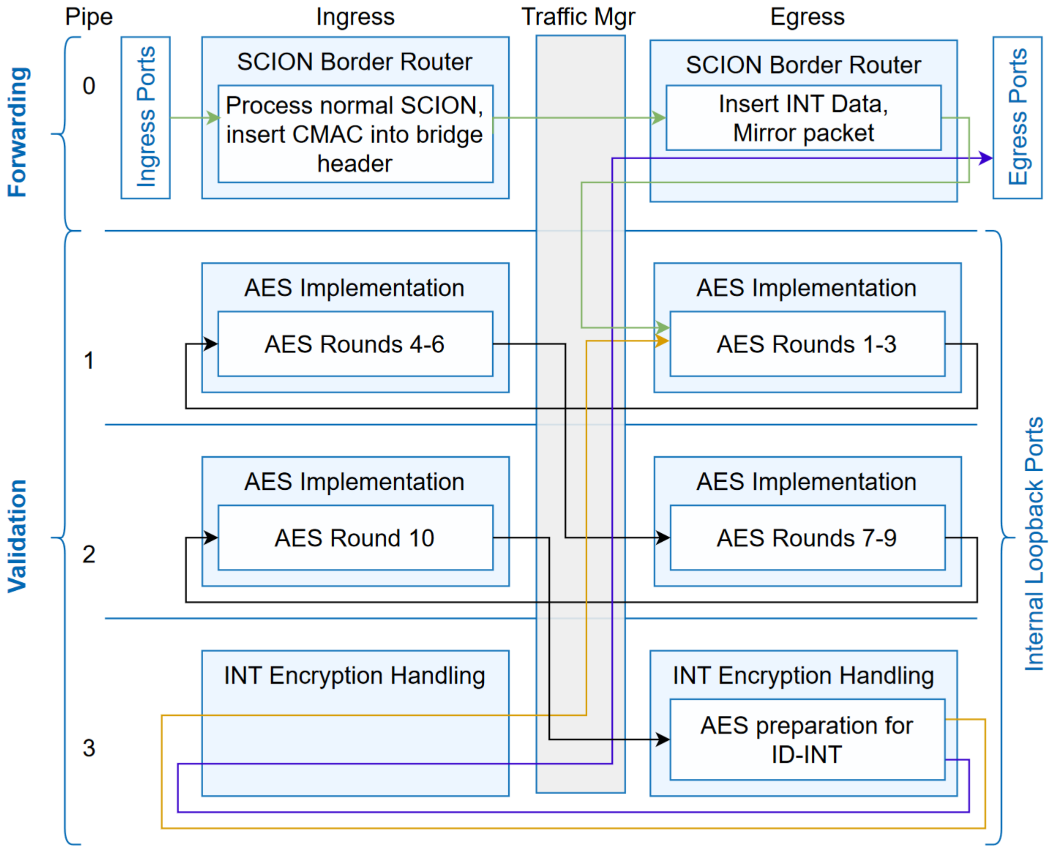
4. P4 Implementation on Tofino

- First process SCION Hop Field
- Process ID-INT
- Validate Hop Field MAC
- Perform DRKey
- Calculate MAC for ID-INT telemetry stack entry
- Forward packet to egress port



4. P4 Implementation on Tofino

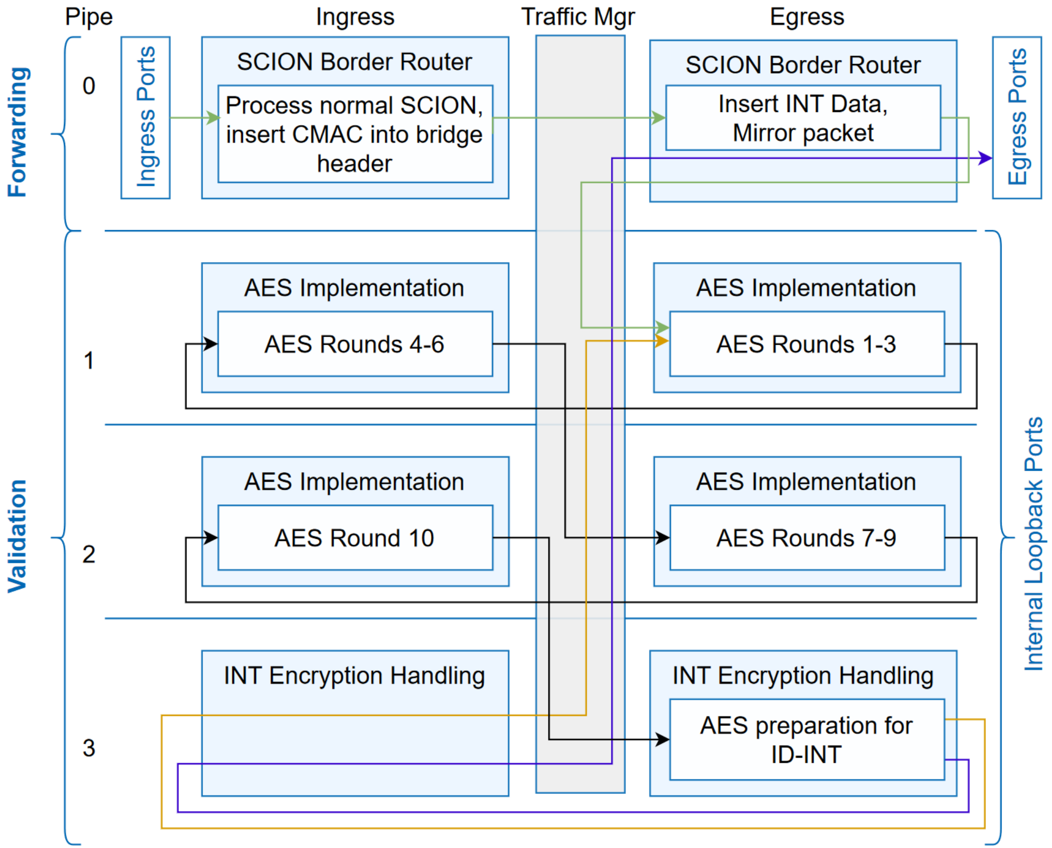
- First process SCION Hop Field
- Process ID-INT
- Validate Hop Field MAC
- Perform DRKey
- Calculate MAC for ID-INT telemetry stack entry
- Forward packet to egress port



5. Performance Observations of our Implementation

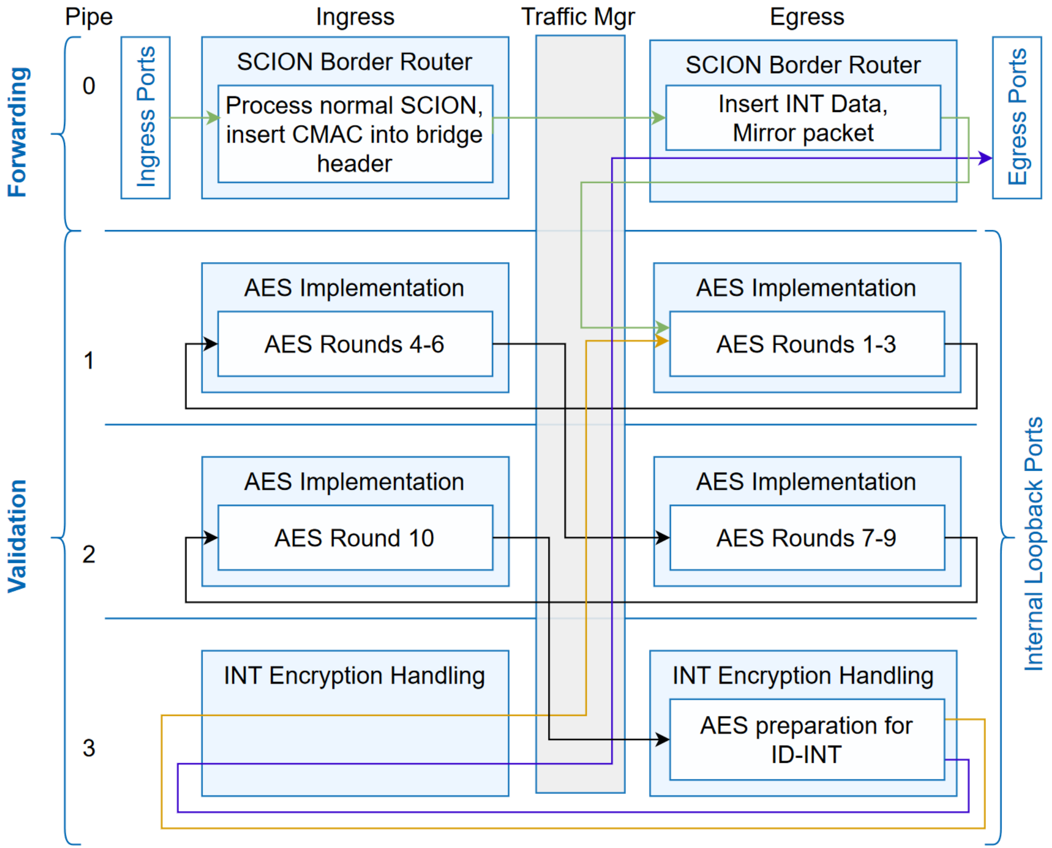
- Best Case: Standard SCION packet without ID-INT
  - 400 Gbps per port

- Worst Case: SCION packet with ID-INT and two hop fields
  - 33.33 Gbps per port



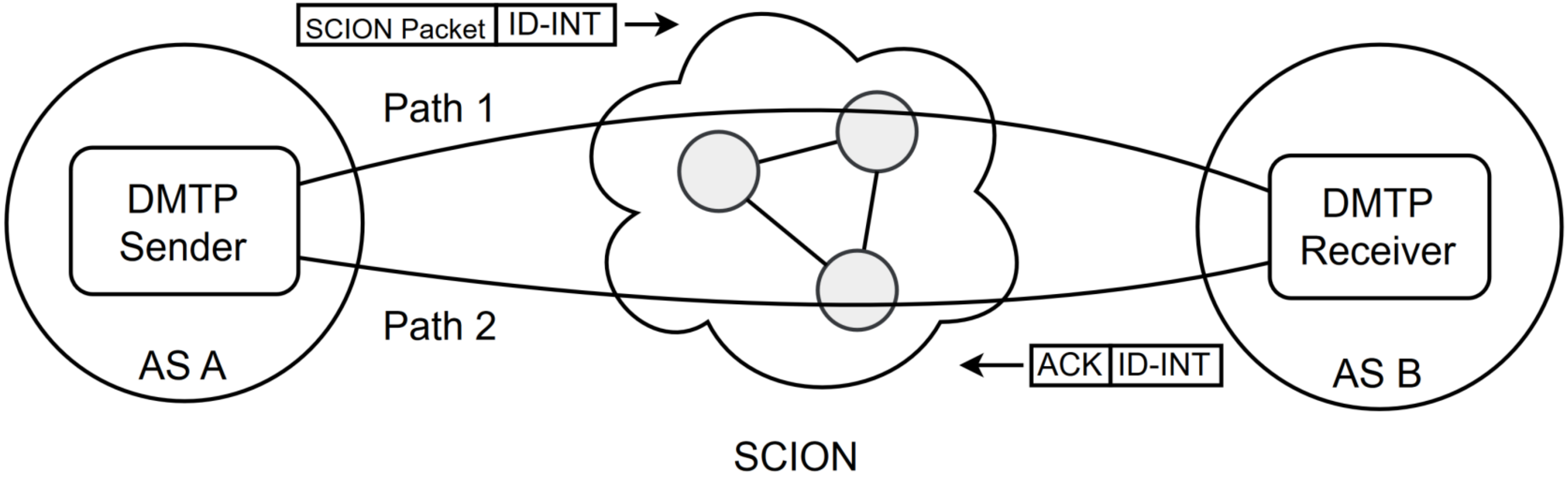
5. Performance Observations of our Implementation

- Two hop fields are a rare case for non-core SCION ASes (path shortcuts)
- Regular Case: SCION packet with one hop field and ID-INT
  - 66.66 - 100 Gbps per port
- ID-INT may not be included in every packet or could be rate limited depending on use case
  - Avg.: Close to 400 Gbps



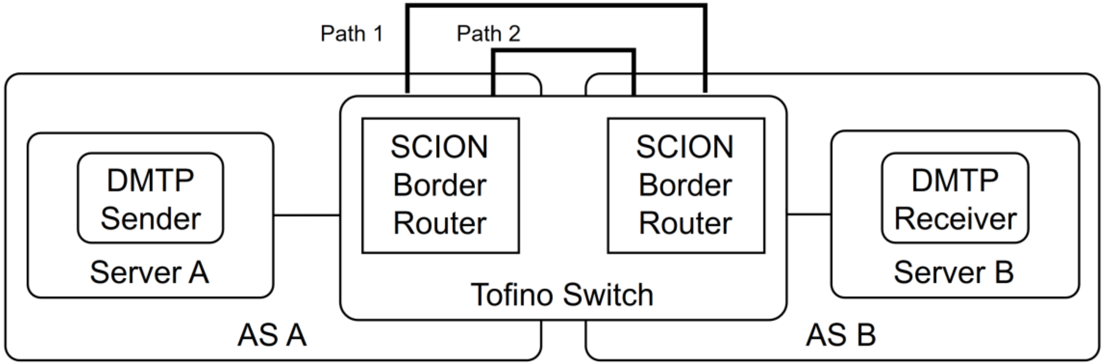
6. Evaluation of ID-INT on DMTP as Specific Use Case

- To prove effectiveness of ID-INT for SCION routing, DMTP is deployed
- Integrated ID-INT into DMTP to select paths based on live statistics



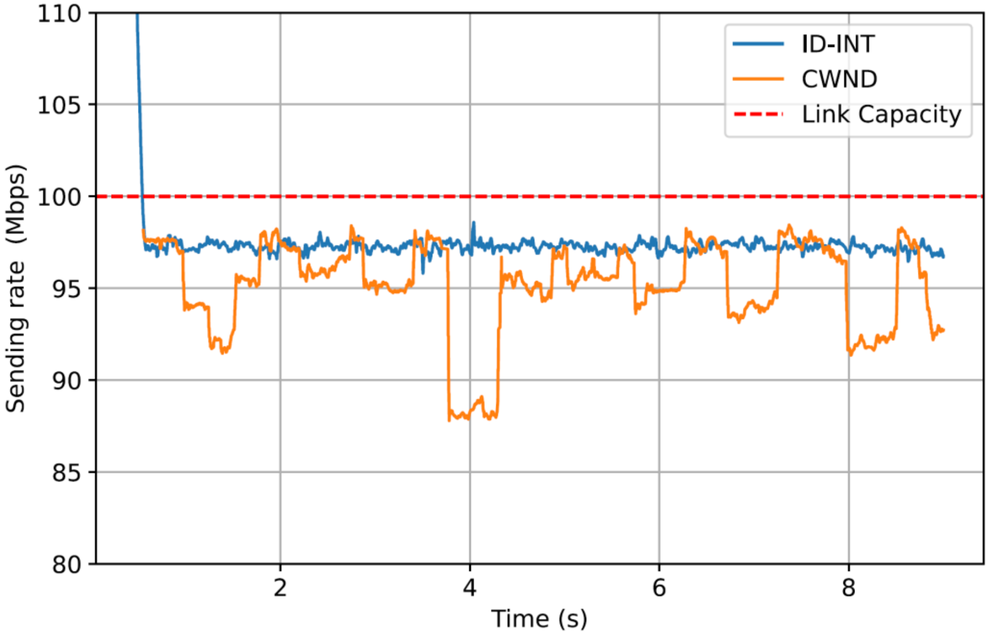
6. Evaluation of ID-INT on DMTP as Specific Use Case

- Evaluation setup used Tofino as 2 routers and a server that ran two SCION end hosts
- Primarily used instantaneous queue length in Tofino as metric
- Bottleneck inside the Tofino introduced to simulate link congestion



6. Evaluation of ID-INT on DMTP as Specific Use Case

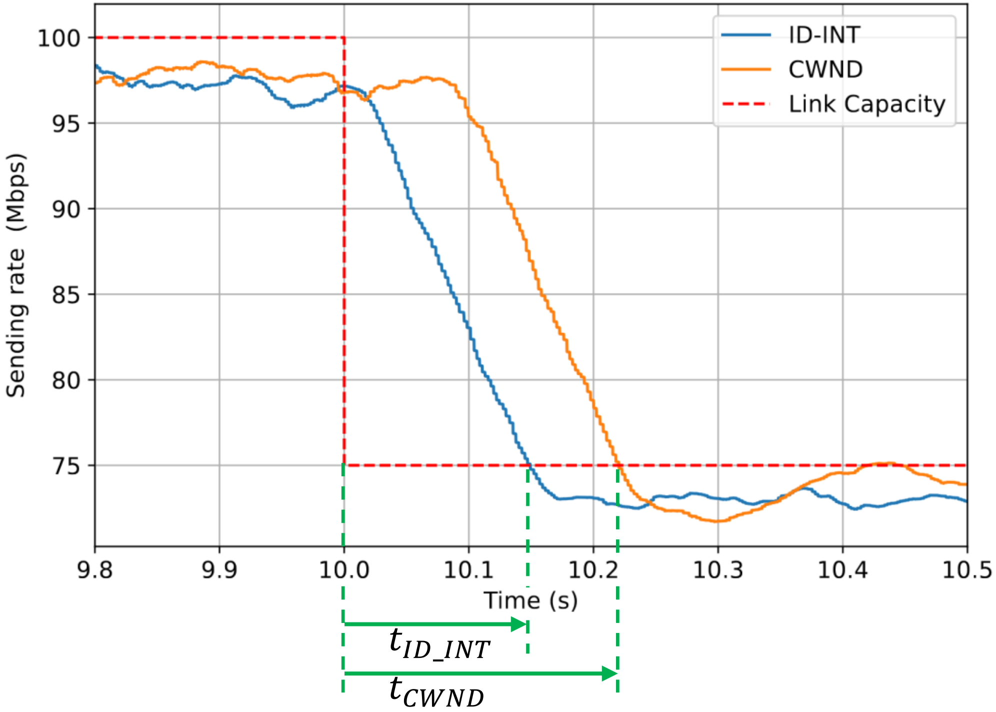
- With constant link capacity of 100 Mbps, the sending rate was more stable using the live metrics of ID-INT



6. Evaluation of ID-INT on DMTP as Specific Use Case

- When dropping the link capacity to 75 Mbps, ID-INT allows a 35 % faster reaction to the changed network conditions

  - In multipath scenario, switching to a different path as a failover would be done faster



7. Conclusion & Future Work

- We presented the first hardware implementation of ID-INT
- We implemented ID-INT in DMTP and proved its effectiveness in assisting path selection in terms of
  - More constant sending rates
  - Faster reaction times and failover under changing network conditions.

7. Conclusion & Future Work

- We presented the first hardware implementation of ID-INT
- We implemented ID-INT in DMTP and proved its effectiveness in assisting path selection in terms of
  - More constant sending rates
  - Faster reaction times and failover under changing network conditions.

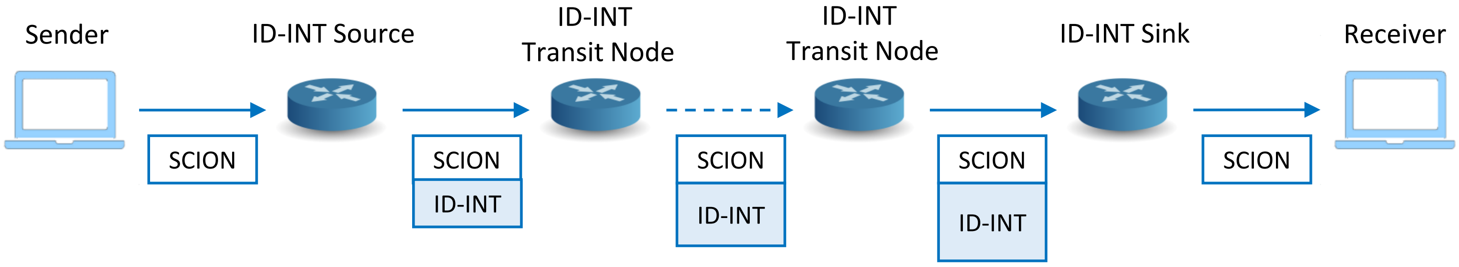
- Improve current implementation
- Implement ID-INT on NetFPGA & XDP and compare performance and available metadata

Thanks for your attention!

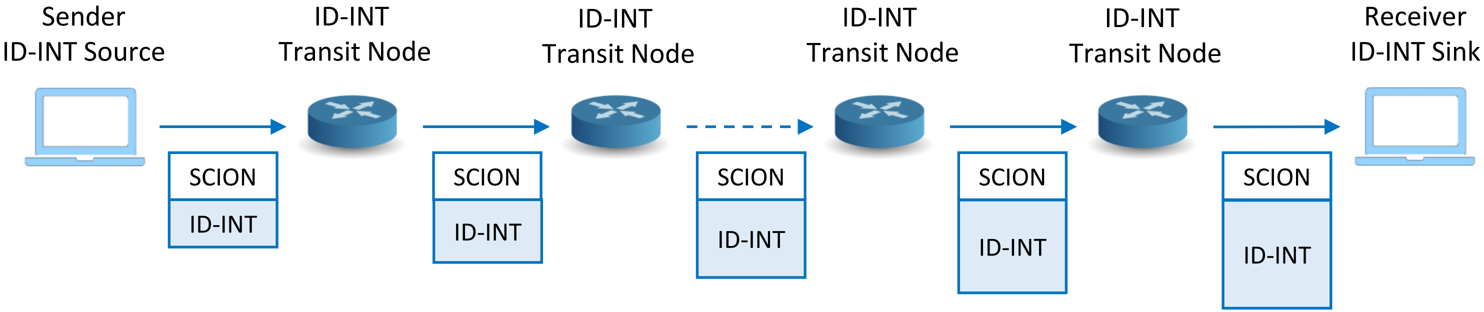
E-Mail: robin.wehner@ovgu.de

3. ID-INT

Infrastructure ID-INT:

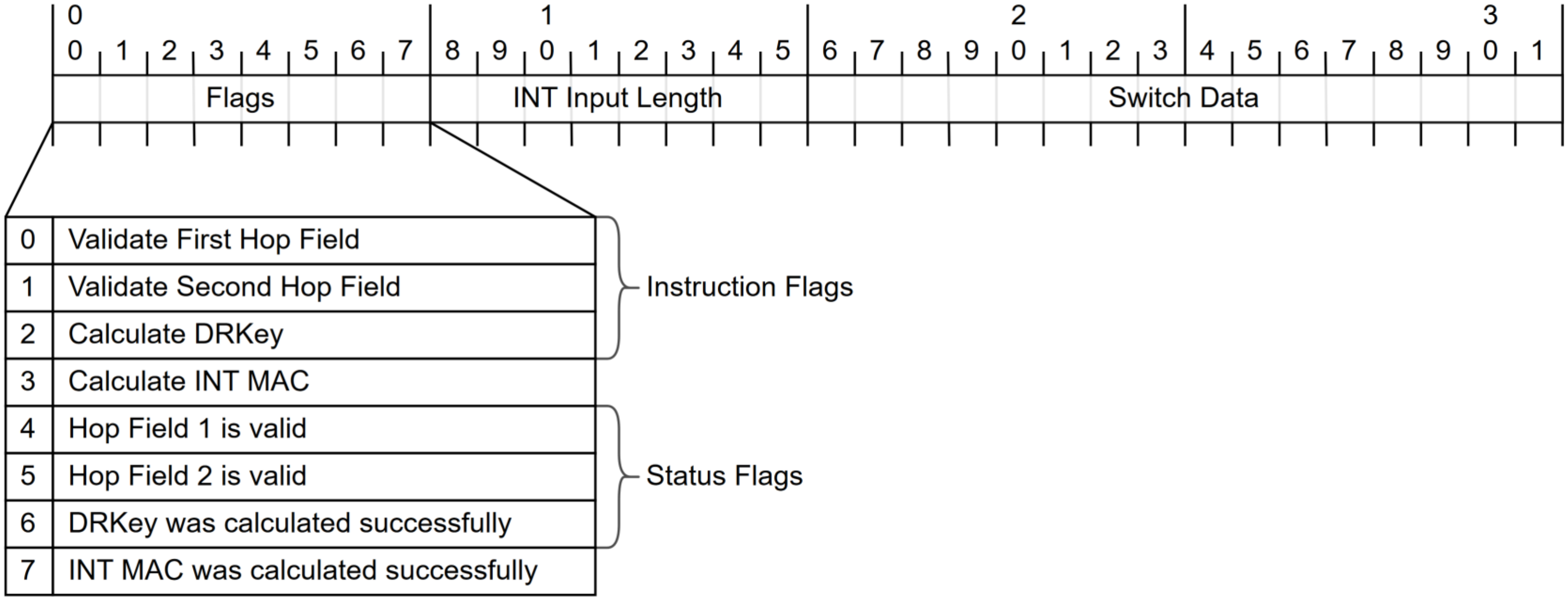


Host ID-INT:



4. Implementation

- Bridge Header to communicate data between pipes:



0	Validate First Hop Field	Instruction Flags
1	Validate Second Hop Field	
2	Calculate DRKey	
3	Calculate INT MAC	
4	Hop Field 1 is valid	Status Flags
5	Hop Field 2 is valid	
6	DRKey was calculated successfully	
7	INT MAC was calculated successfully	

4. Implementation on Tofino

- Currently supported Metadata:
  - Timestamps, Queue ID, Instantaneous Queue Lengths, Ingress Port
- Additional metadata supported by Tofino, but not implemented so far
- Due to Tofino internal functionality it may be impossible to support all combinations of metadata and telemetry fields
  → Allow specific metadata only at specific positions of flexible metadata