Need for / Usage of Performance and Diagnostic Metrics Destination Options Header

Nalini Elkins: Inside Products, Inc.

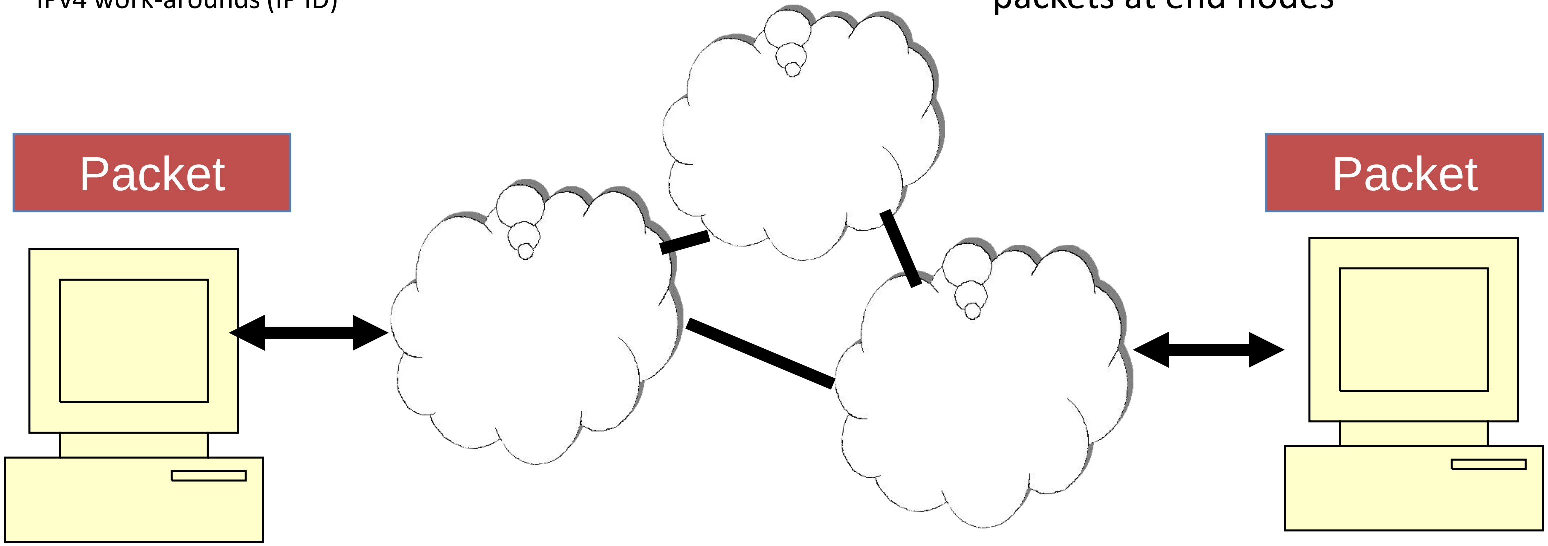
Mike Ackermann : BCBS Michigan

(Team : Keven Haining : US Bank, Sigfrido Perdomo: DTCC, William Jouris: Inside Products, David Boyes : Sine Nomine)

Introduction

- Network traffic needs to be monitored
  - Diagnostics
  - Performance
  - Repair
  - Failures
- Traditionally done by:
  - instrumentation at hosts, router
  - IPv4 work-arounds (IP ID)

- Not all operators own all parts of network
  - Enterprises, business partners, software defined networks, Infrastructure as a Service
- Only visible portion of traffic is packets at end nodes



Not Zero-Impact

- **Measurement at middle hardware and hosts not a zero-impact solution**
- Data capture, measurement impacts performance
- Footprint of agents is substantial
- Diagnostics require packet (headers)
- More metrics needed

IPv4 Work-Arounds

- **No unified place for performance / diagnostic metrics**
- IPv4 IP ID field used as de facto packet sequence number
- Doesn't work for some platforms
- Not available in IPv6 (moved to fragment header)
- Timestamps for response time not available

Metrics Needed

- Packet sequence number
  - Speeds diagnostics
  - Many use cases given in Internet Draft, last IETF
  - IPv4 IP ID
- End-to-end response time WITHOUT agents
  - Service Level Agreements
  - First Mover Advantage
  - Separate metrics needed for quick triage:
    - Inbound network time
    - Server time
    - Outbound network time

Why Packet Sequence Number?

- Why current metrics not good enough?
  - TCP sequence number
    - distinguish between retransmit / duplicate packets
    - packets dropped and retransmitted (sender may have sent 4 times, we only received once)
    - not applicable to non-TCP traffic
- UDP
  - No current metric
  - Would have to change all apps
- Hashing technique (packet sequence number)
  - Known problem with packet duplication
  - Packets dropped and retransmitted (sender may have sent 4 times, we only received once)
  - Time / overhead delay

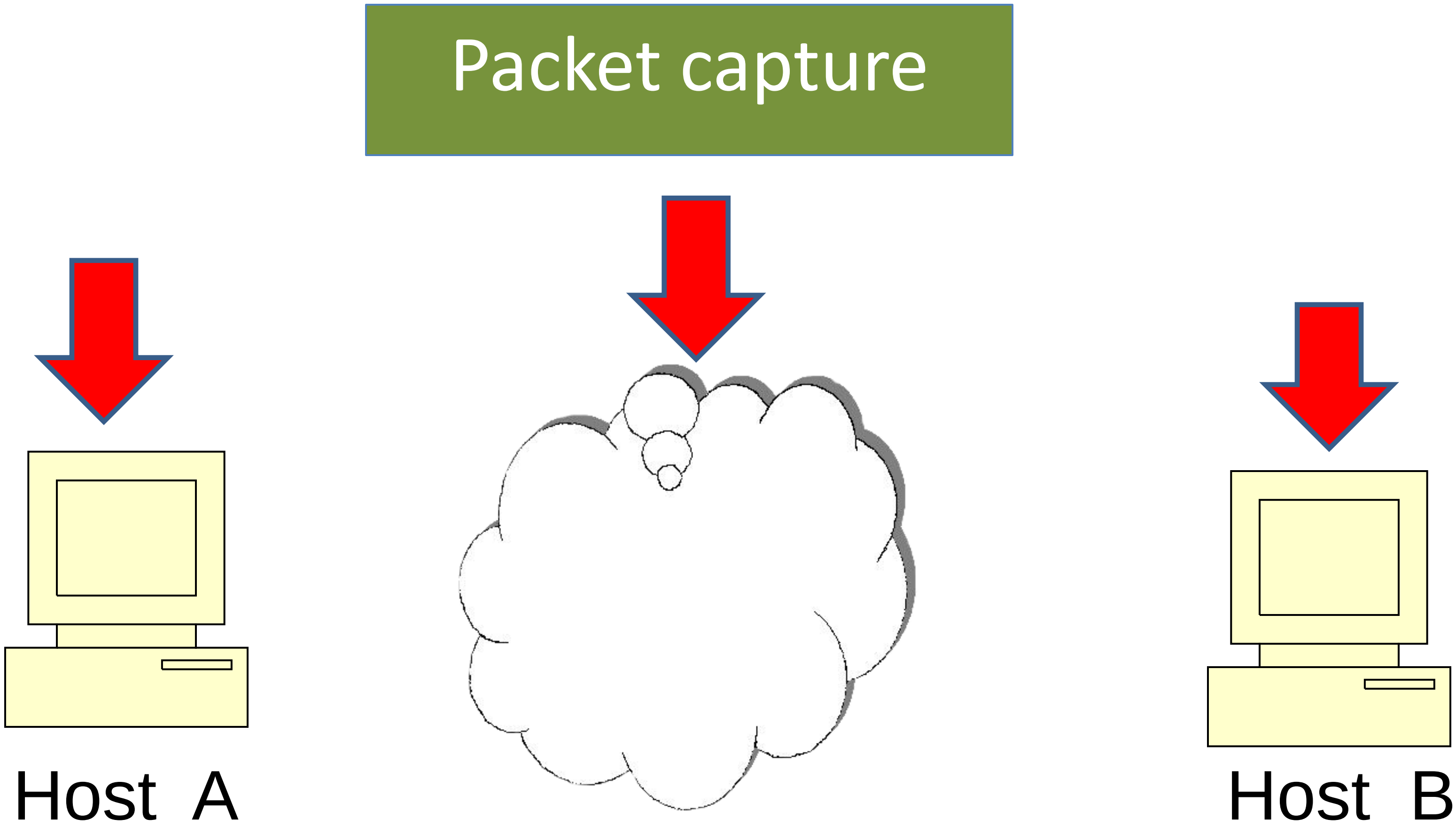
Requirement

- In basic IP transport
- Unmolested by middle systems

Solution

- **Implementation** of existing extension header : Destination Options Header (DOH)
- Performance and Diagnostic Metrics (PDM) DOH

Response Time Measurements
Packet Capture



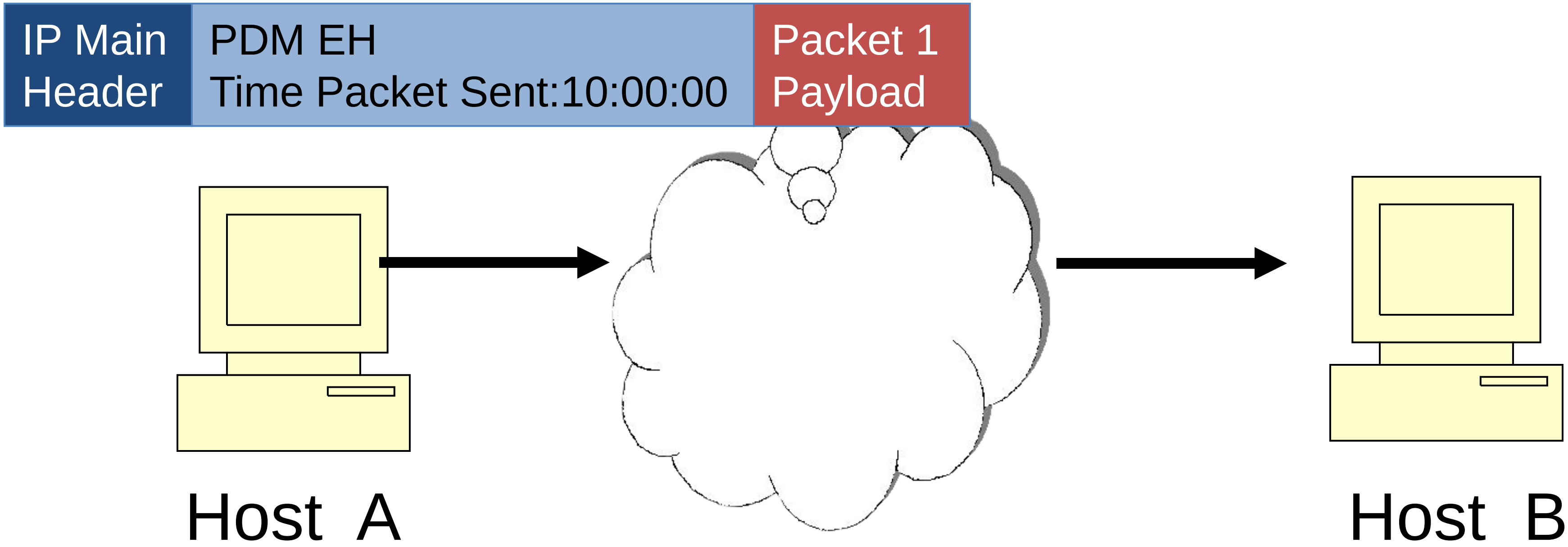
Packet capture

Host A

Host B

Response Time Measurement
Step 1

- Packet 1 sent from source host A
- Time-stamped leaving Host A
- Timestamp is in PDM extension header



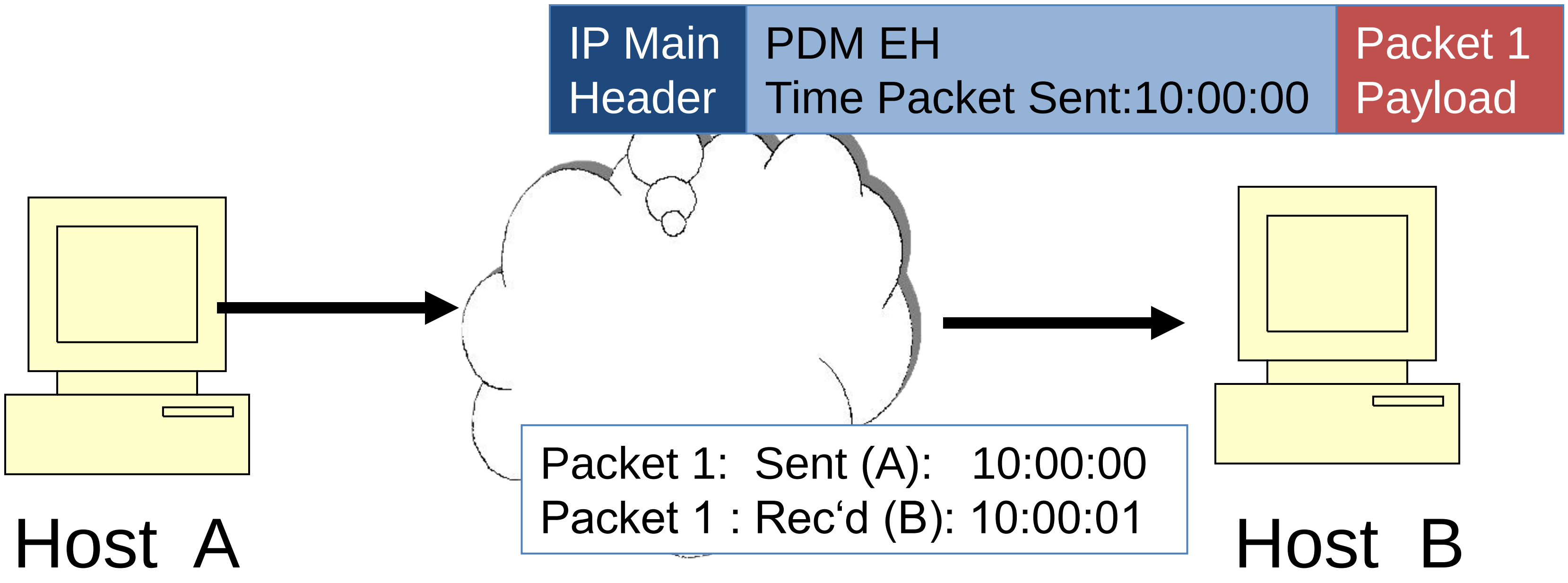
IP Main Header

PDM EH
Time Packet Sent:10:00:00

Packet 1 Payload

Response Time Measurement
Step 2

- Packet 1 received at Host B
- Time-stamped leaving Host A
- Inbound network time = Packet 1 rec’d (B) – Packet 1 sent (A)



IP Main Header

PDM EH
Time Packet Sent:10:00:00

Packet 1 Payload

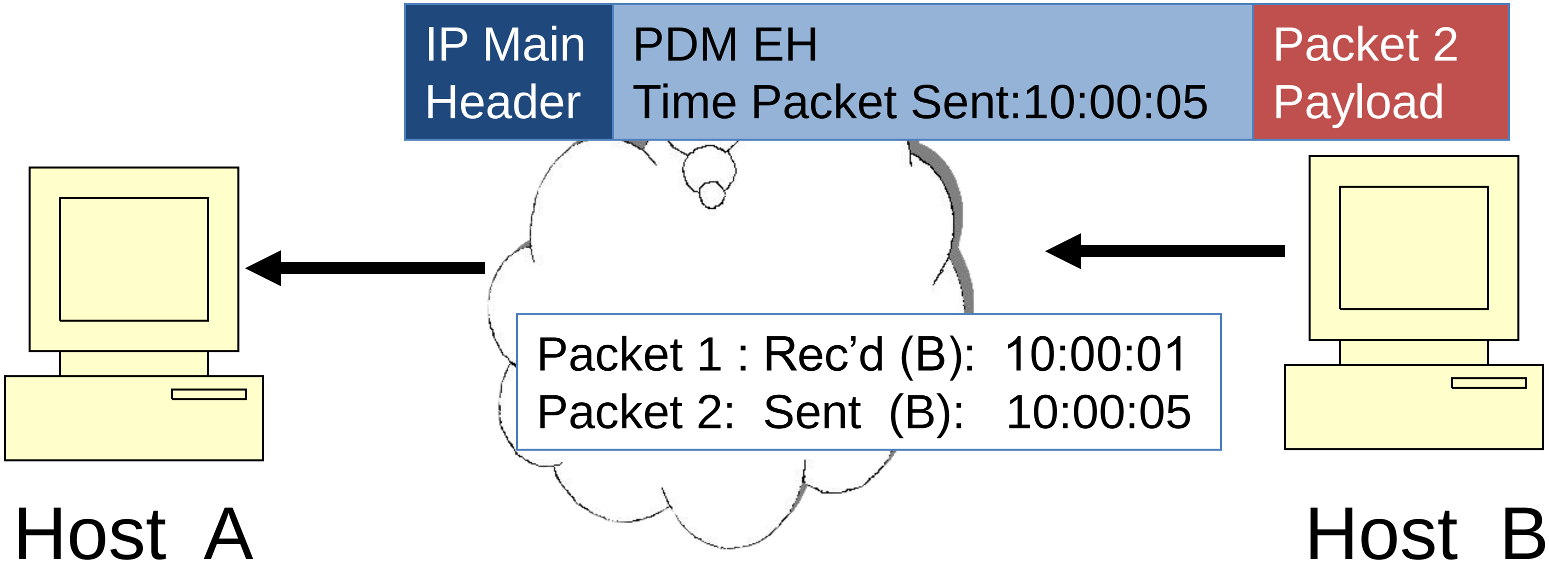
Packet 1: Sent (A):   10:00:00
Packet 1 : Rec‘d (B): 10:00:01

Host A

Host B

Response Time Measurement
Step 3

- Packet 2 sent from Host B (response to Packet 1)
- Time-stamped leaving Host B
- Processing Time = Packet 2 sent (B) - Packet 1 rec'd (B)



IP Main Header

PDM EH
Time Packet Sent:10:00:05

Packet 2 Payload

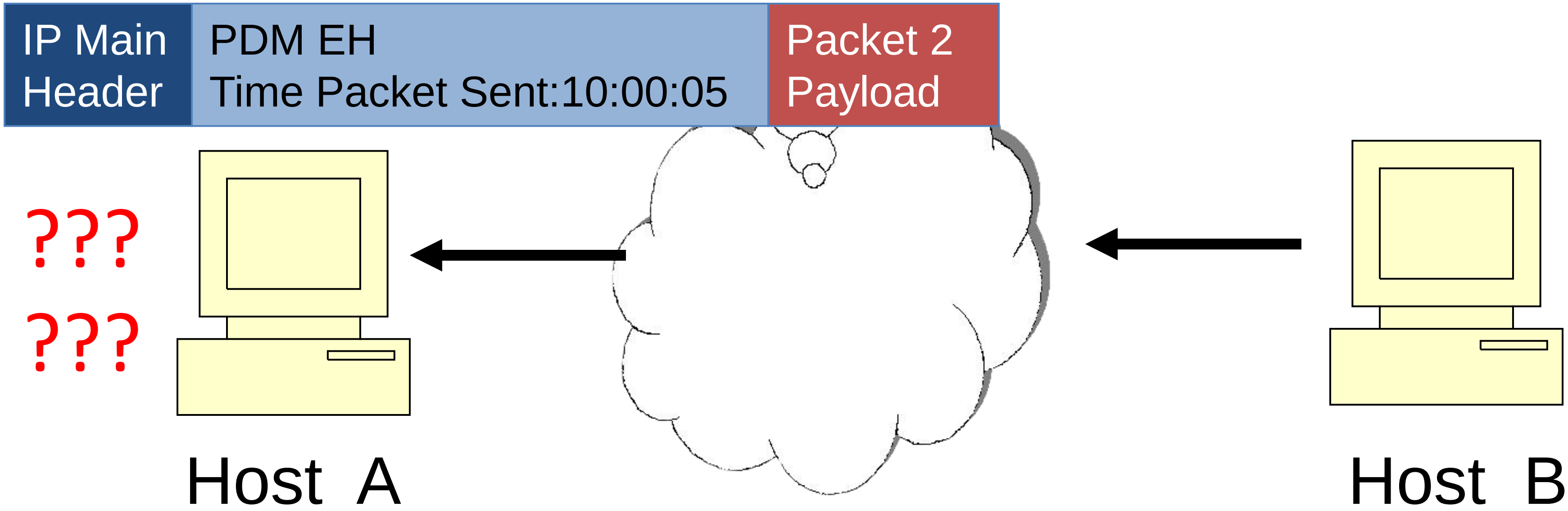
Packet 1 : Rec'd (B): 10:00:01
Packet 2: Sent (B): 10:00:05

Host A

Host B

When Did it Get to Host A?

- When did Packet 2 to arrive at Host A?
- Return route may not be the same, may be congestion, packet might never arrive.



IP Main Header

PDM EH
Time Packet Sent:10:00:05

Packet 2 Payload

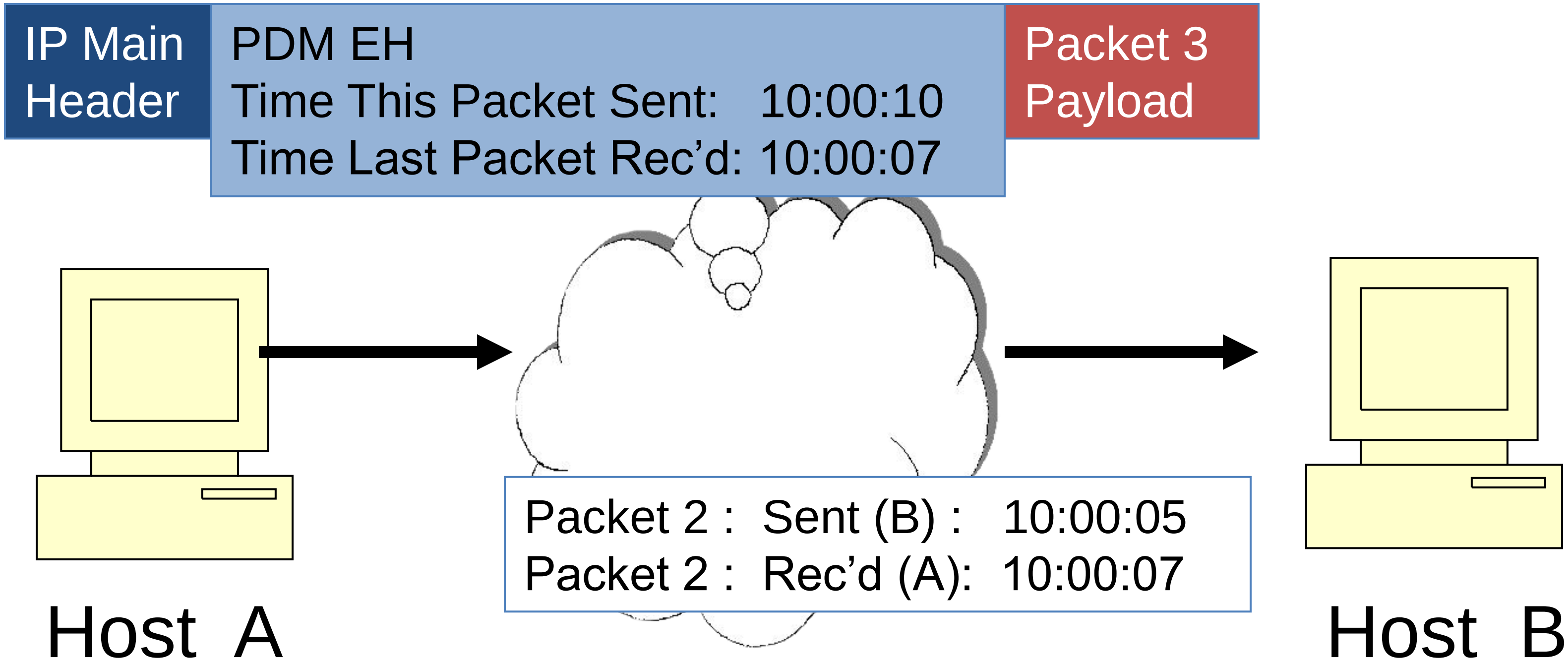
???
???

Host A

Host B

What is Needed?

- With each packet, add “Time Last Packet Received” in PDM EH
- When Packet 3 sent, has when Packet 2 got to Host A
- Outbound Network time = Last rec’d (A) – Time sent (B)
- Processing Time (A) = Packet 3 sent (A) - Last rec’d (A)



IP Main Header

PDM EH
Time This Packet Sent: 10:00:10
Time Last Packet Rec’d: 10:00:07

Packet 3 Payload

Packet 2 : Sent (B) : 10:00:05
Packet 2 : Rec’d (A): 10:00:07

Appendix

PDM Destination Options EH

Size (bits)	Field Name	Description
8	Next Header	Points to next header or payload
8	Reserved	Set to 0.
8	Option Type	To be assigned by IANA
8	Option Length	Length
16	Packet Sequence Number	Initialized at 0 and monotonically incremented for protocol packet on the connection. 16-bit unsigned integer. This field will obviously wrap quickly. It is intended for human use.
64	Timestamp (This packet)	A 64-bit unsigned integer field containing a timestamp. This is the time this packet was sent. NTP format timestamp
64	Timestamp (Last Packet)	A 64-bit unsigned integer field containing a timestamp. This is the time the last packet was received. NTP format timestamp
64	Application Specific	To be used by end-nodes to convey information