Time Determination for Forensic Analysis of Multipoint Network Traces Taken Across Distributed Hybrid Cloud

Charles Barry
charles@luminouscyber.com

Apr 2022

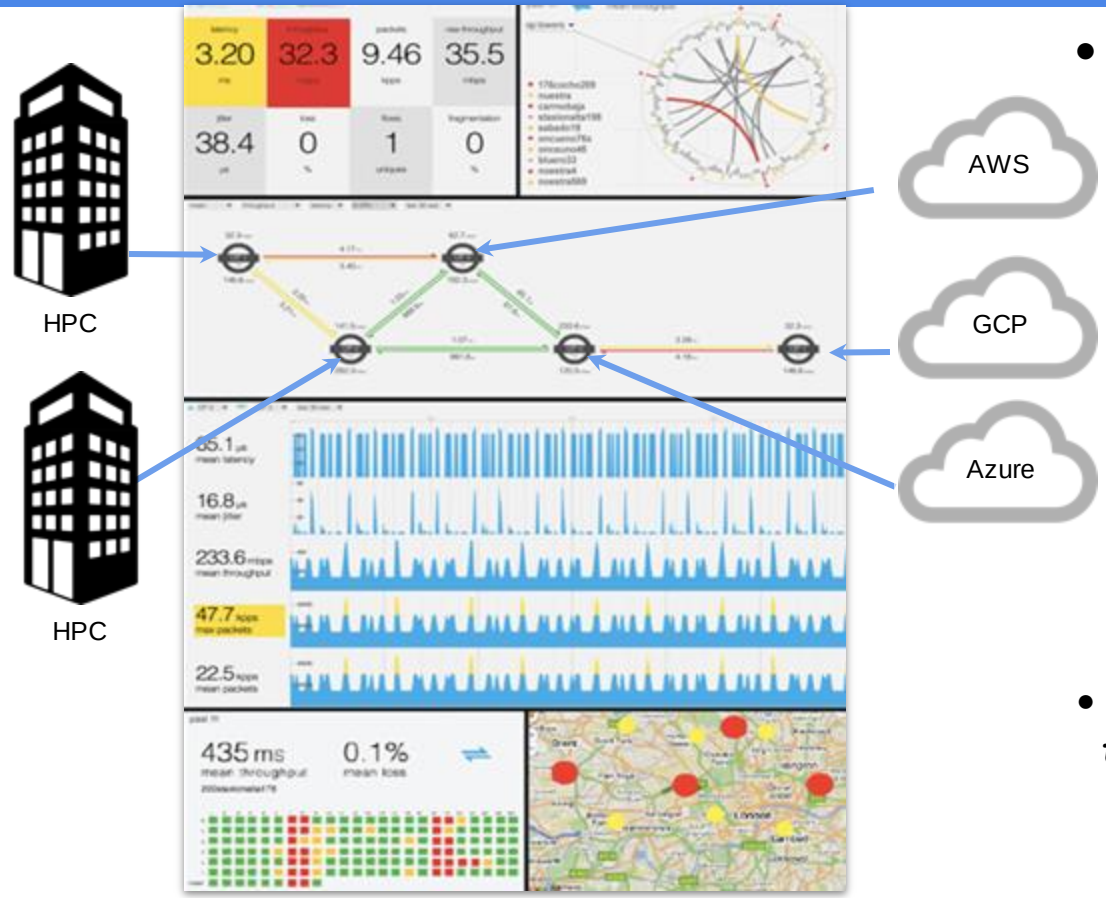
Agenda

- Intro
- Overview
- Problem: Distributed Captures are not Synchronized
- Solution: Time Determination
- Experimental Setup
- Test Results
- Summary

This material is based upon work supported by the U.S. Department of Energy, Office of Science, under Award Number DE-SC-0021595.

Note: See also prior WSTS paper https://wsts.atis.org/wp-content/uploads/sites/9/2019/04/7_02-00_Luminous-Cybernetics_Barry_Time-Determination.pdf

Distributed Hybrid Cloud Analytics



- Analytics Feature Requirements
  - Application Performance Monitoring
  - Network Performance Monitoring
  - Network Security
  - Root Cause Analysis
  - Real-Time Temporal and Spatial Visualization
  - Exploration of Historical data
- All of above need synchronized ***timestamps!***

Problem: Lack of Synchronization in Hybrid Cloud

- Hybrid cloud analytics more important than ever, however:
  - Little conformance among providers {AWS, Azure, GCP, private};
    - *NTP accuracy in the cloud is highly variable, no accuracy guarantees;*
    - *PTP is virtually non-existent, and cloud networks are PTP-Unaware.*
  - NTP/PTP distribute time but provide *no guarantee of timestamp accuracy*
    - Event timestamps subject to linux/OS interrupts, schedulers, TCP stack, etc.

- A new timestamp solution is required for HPC/cloud analytics
  - Accurate, Scalable, Fast lock, Robust

Solution: Time Determination

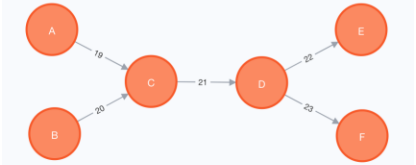
- Packet timestamp analytics to provide causal order of distributed events:
  - Time determination of distributed events without distribution of reference time, phase, or frequency, US10637597B2, US11303374B2
  - ***No** timing packets are exchanged; clients do not have to recover time;*
  - *All packet traffic can be utilized to determine the causal event timestamps*
    - *All packets, interfaces, directions, paths, sizes, classes/priorities*

- Centralized processing with global context ensures causal event order
  - *Superior timestamp performance vs NTP/PTP in every key measure*
    - Time to Lock; accuracy, stability, asymmetry, robustness, scale
    - Timestamp causality

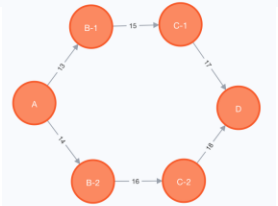
Hybrid Cloud Example Network Topologies

Linear (Multihop)

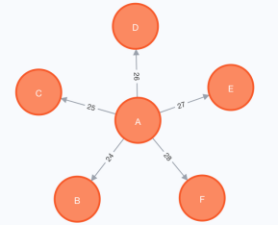
Aggregation/
Bottleneck



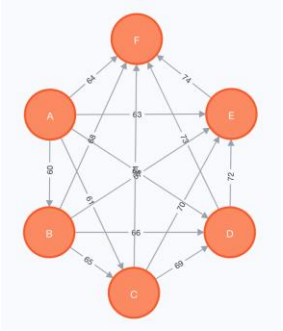
Load Balancer



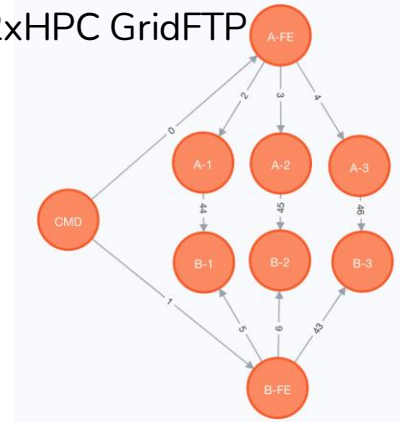
Hub & Spoke



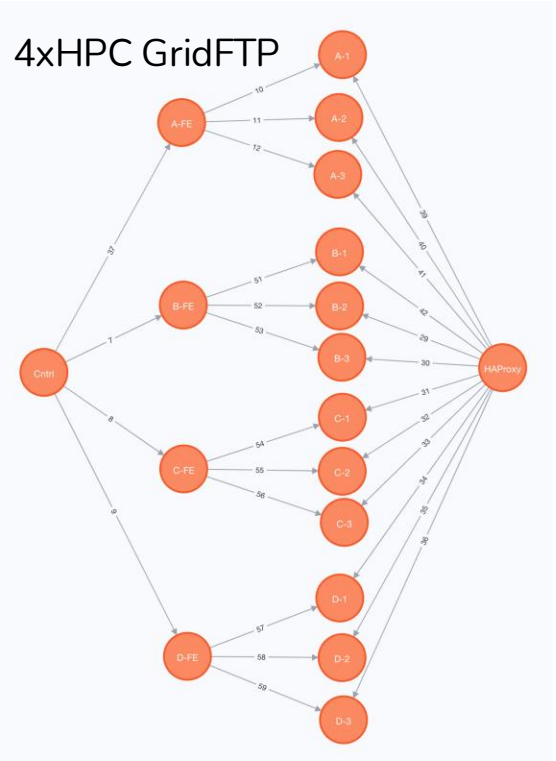
Mesh



2xHPC GridFTP

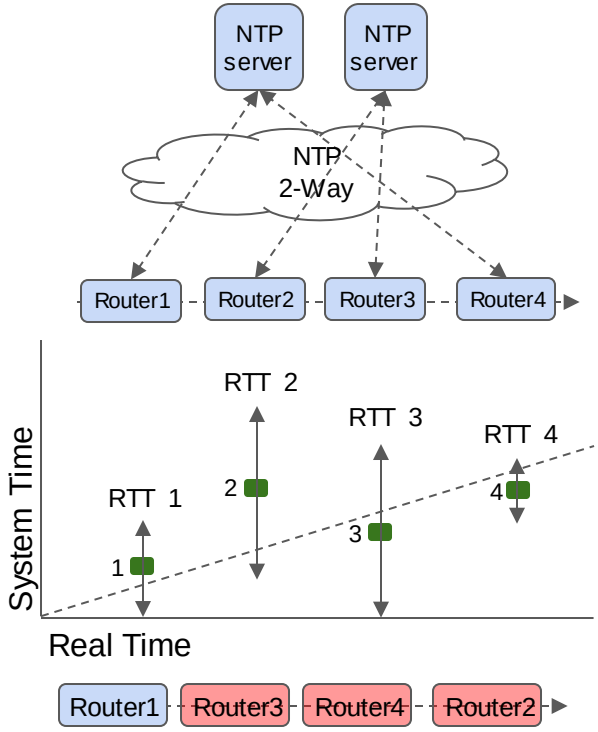


4xHPC GridFTP



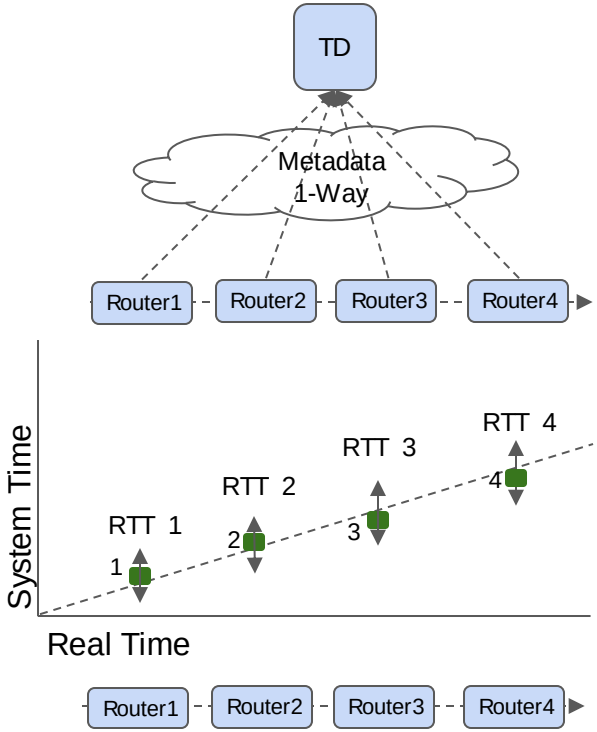
Topologies Created in Neo4J™ Graphical Database

NTP vs TD: non-causal vs causal



NTP event order is incorrect, non-causal

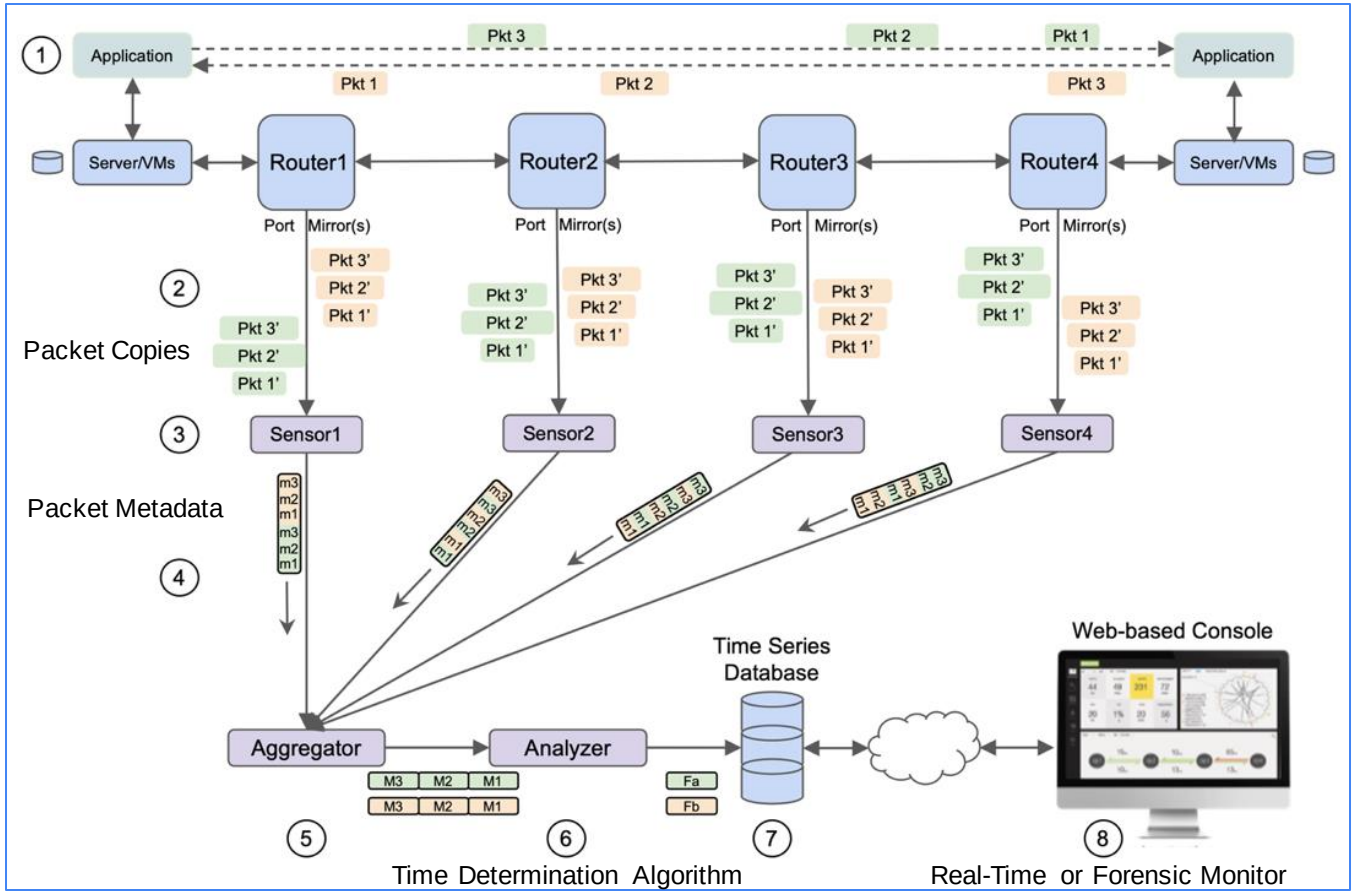
NTP clients may be timed over different paths, over multiple hops and to different servers resulting in per-client asymmetry offset error



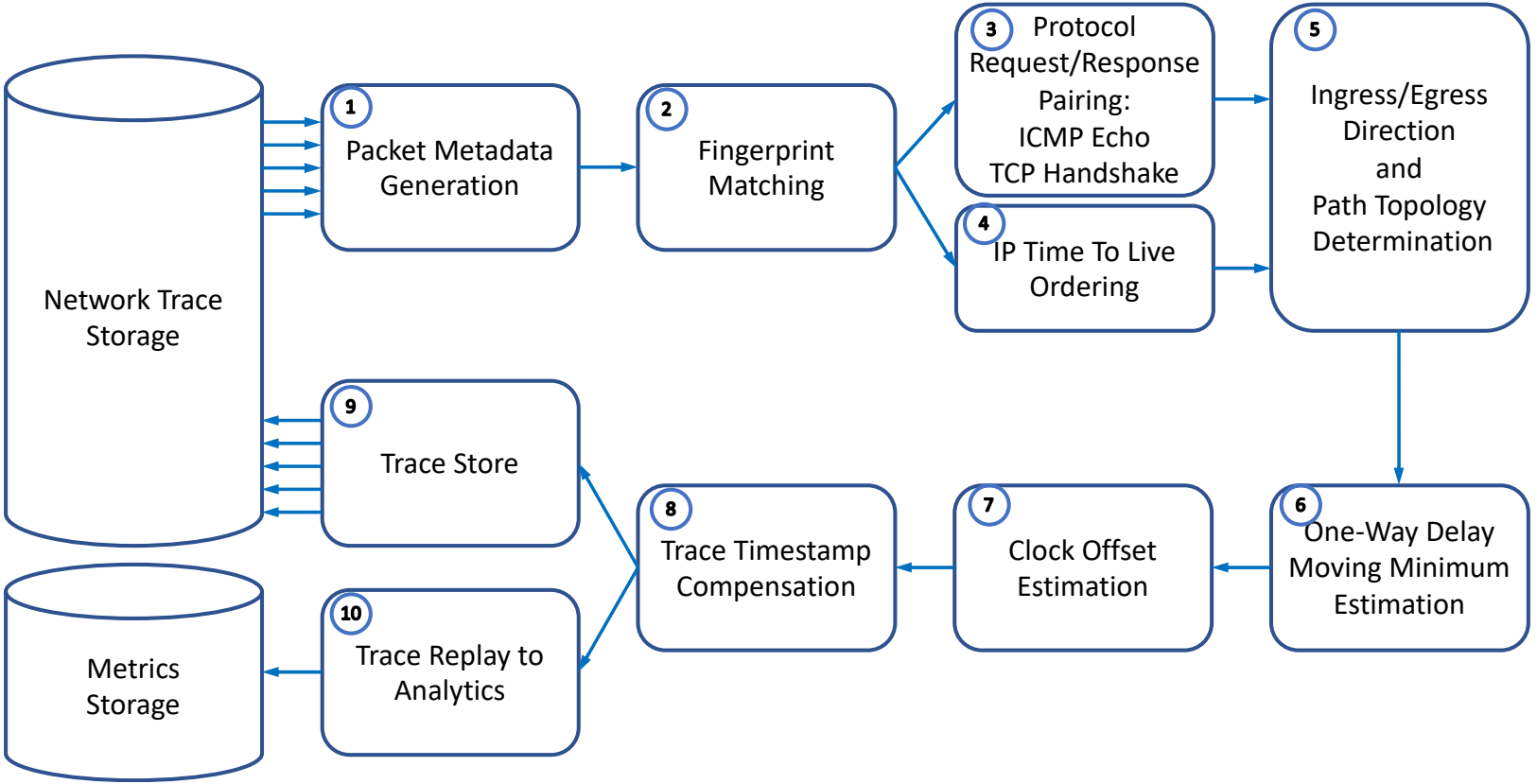
TD event order is correct, causal

TD is synchronized hop by hop (smaller RTT and thus smaller uncertainty) and uses TTL to enforce timestamp causality

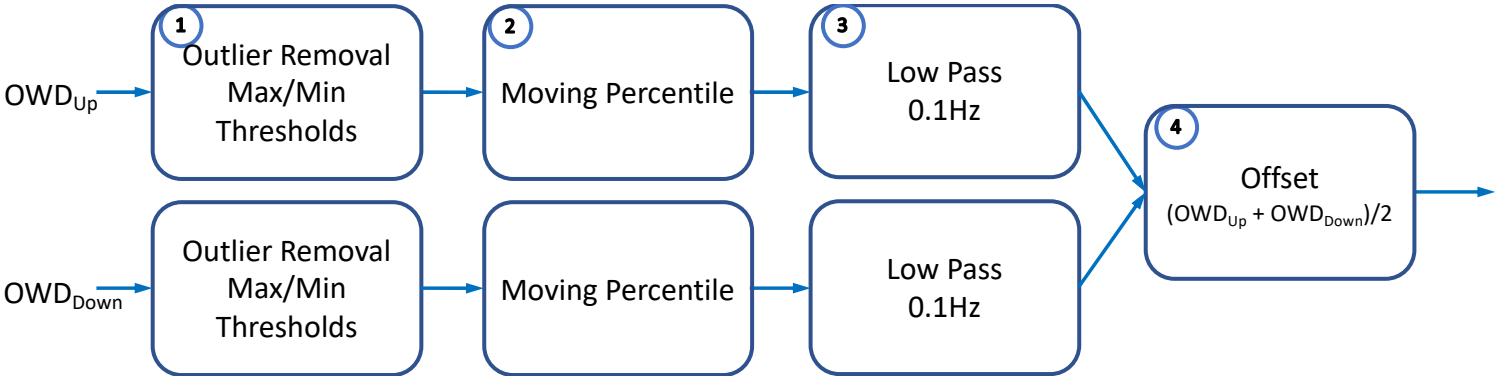
Time Determination: How it works



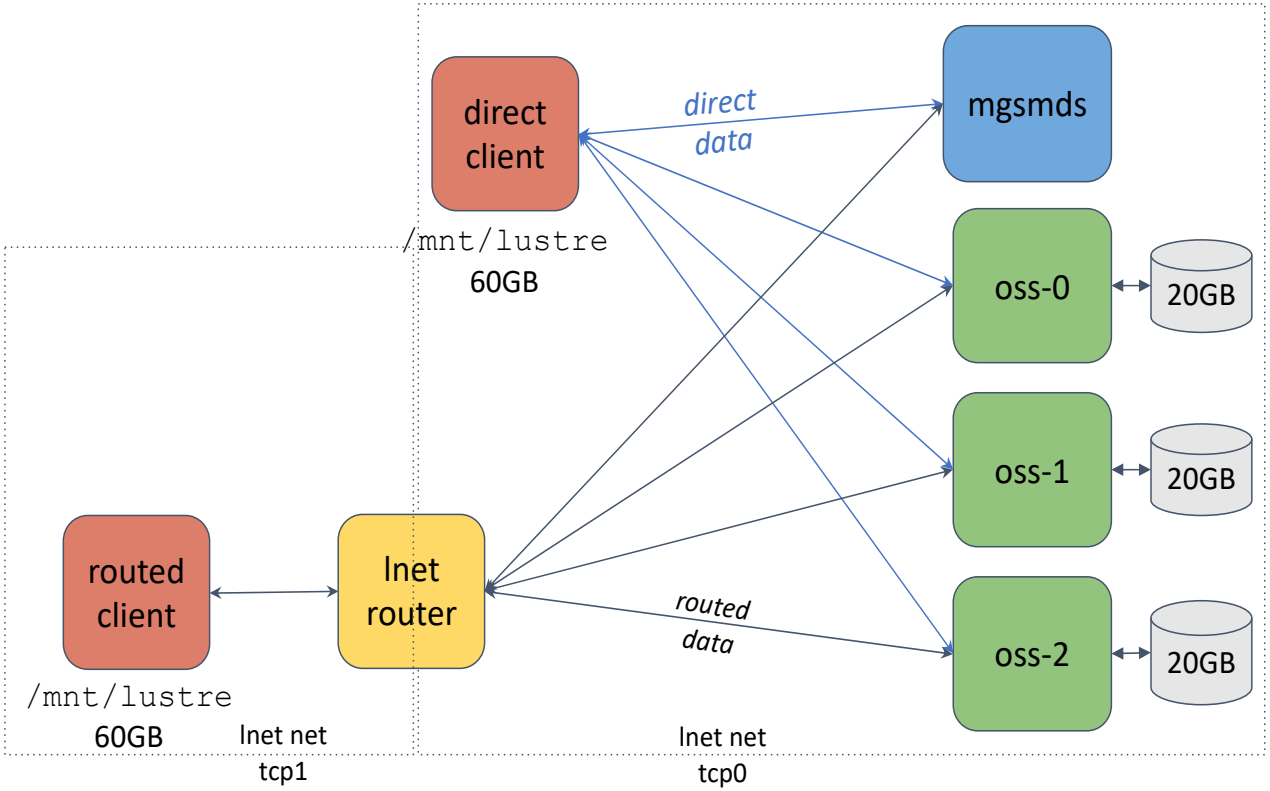
Time Determination from Network Traces



Time Determination Clock Offset estimation

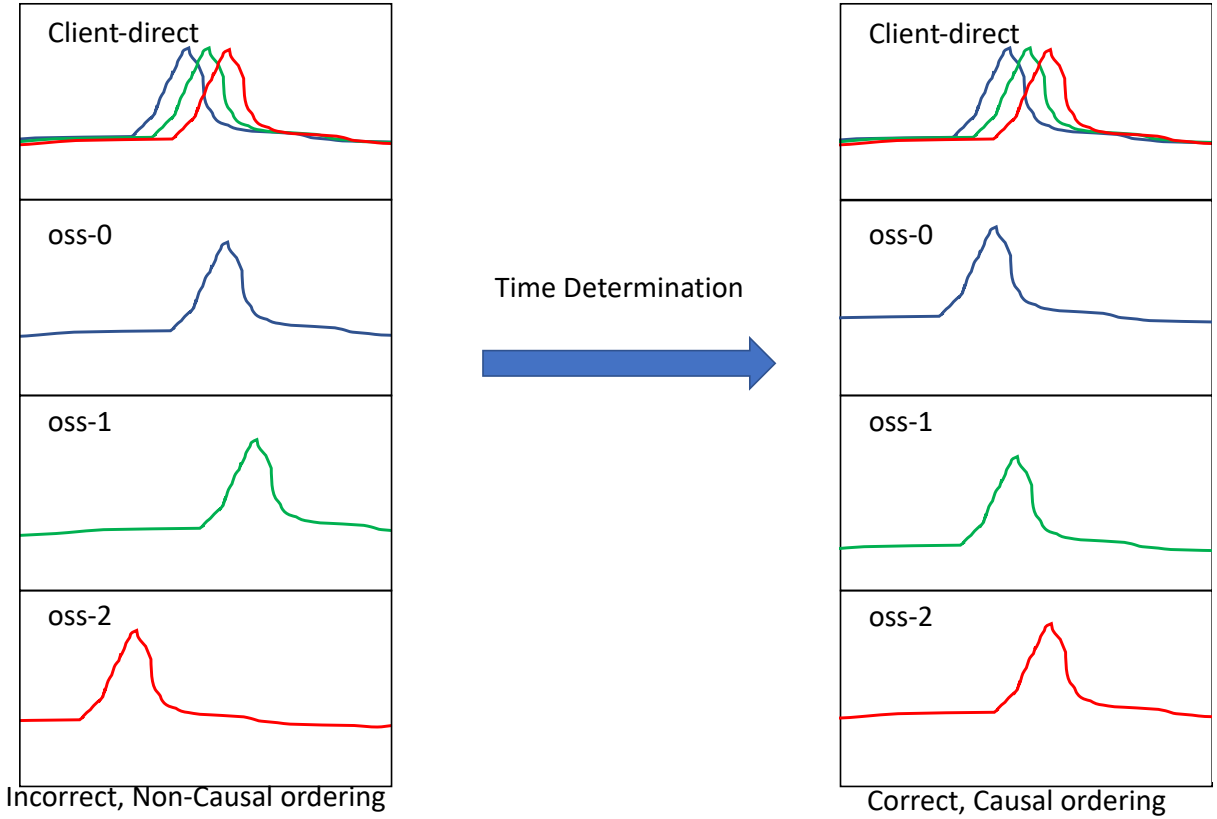


Real-World Application: Lustre™ Database



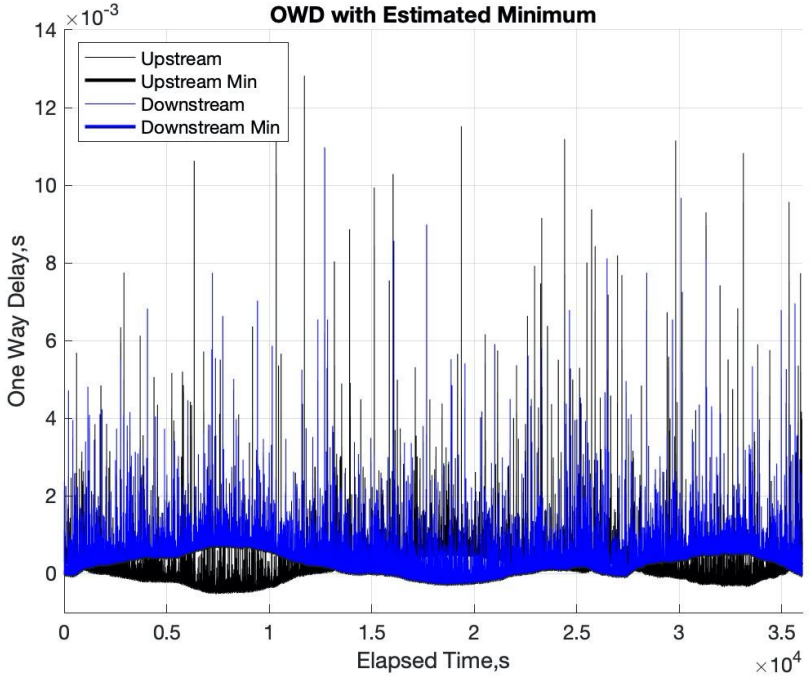
All nodes {routed, direct, router, mgsmds, oss-1,2,3} are instantiated in different virtual machines

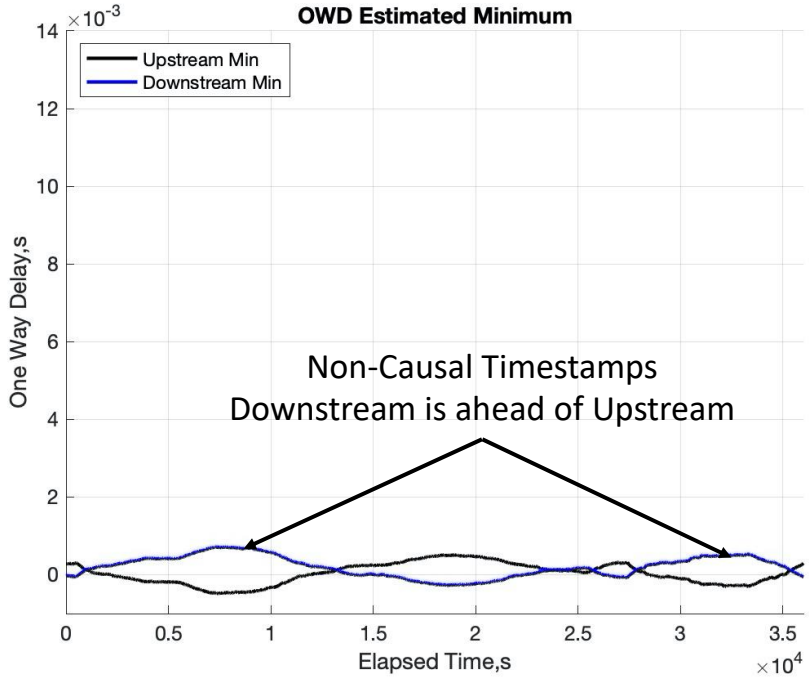
NTP vs TD: non-causal vs causal



Time Determination: One Way Delay (NTP timed)

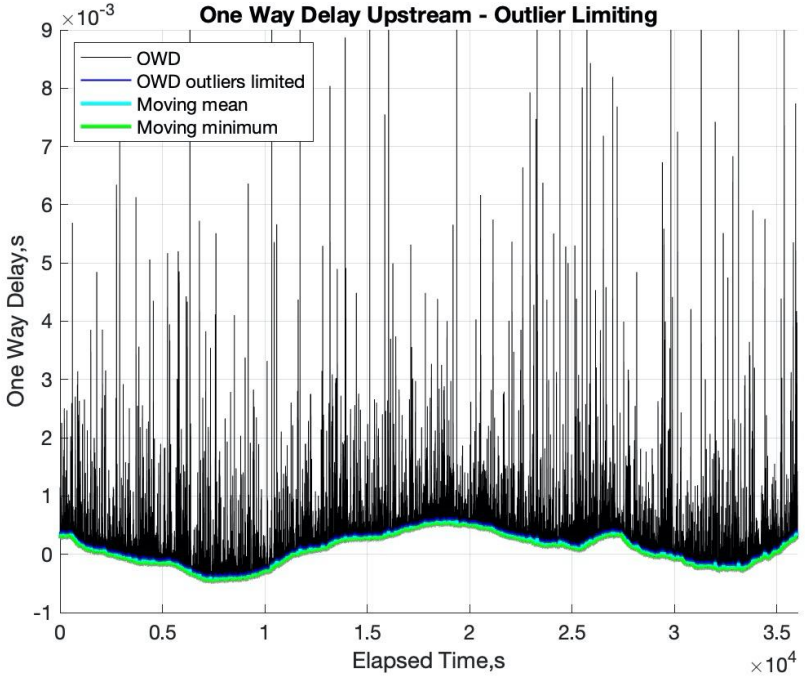
One Way Delay – Raw vs Minimum Estimation

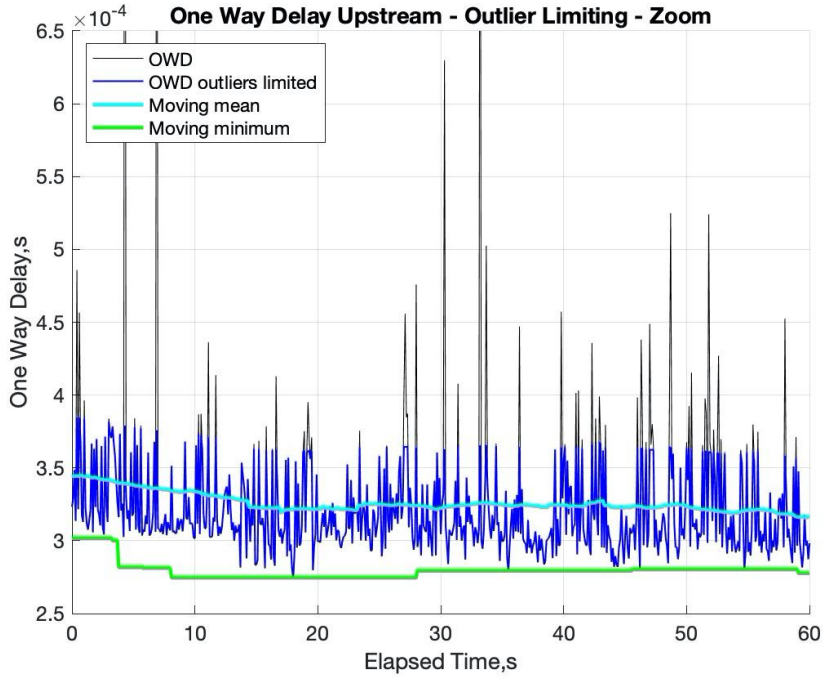




Time Determination: Outlier Removal

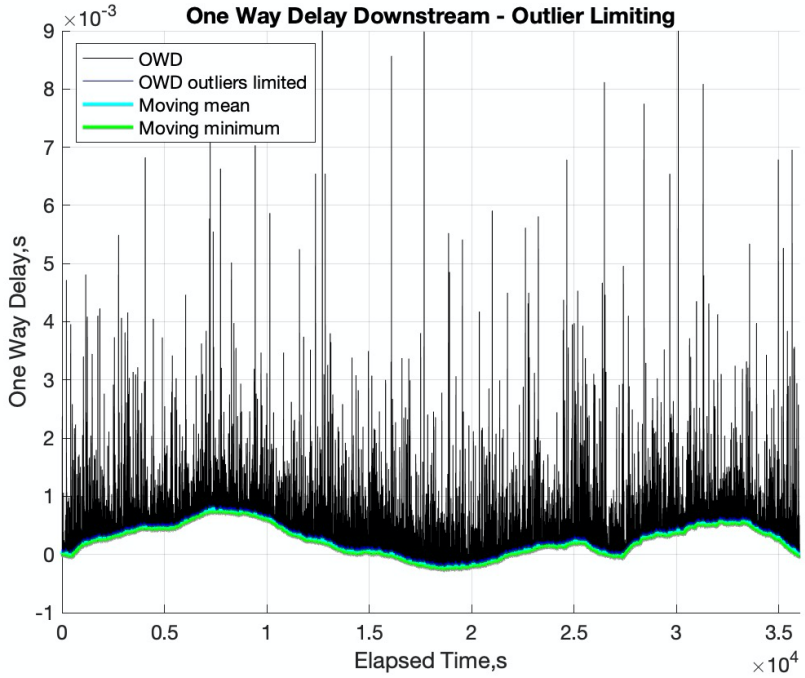
One Way Delay Estimation – Outlier Limiting Upstream

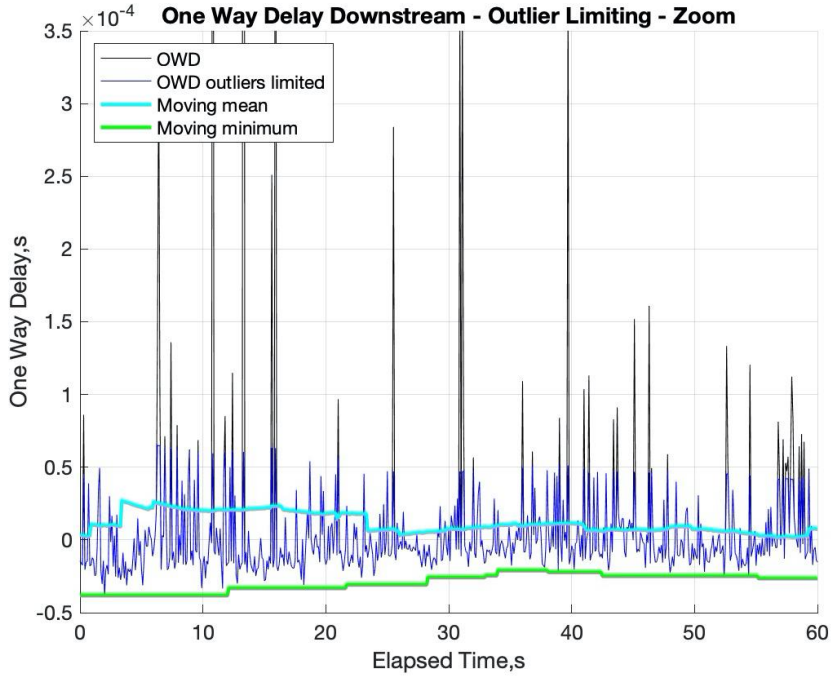




Time Determination: Outlier Removal

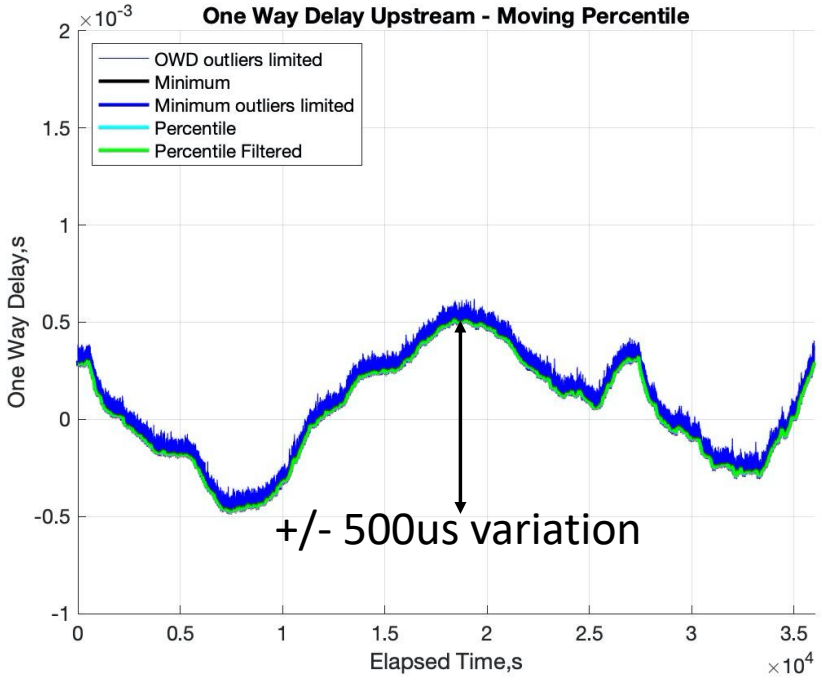
One Way Delay Estimation – Outlier Limiting Downstream

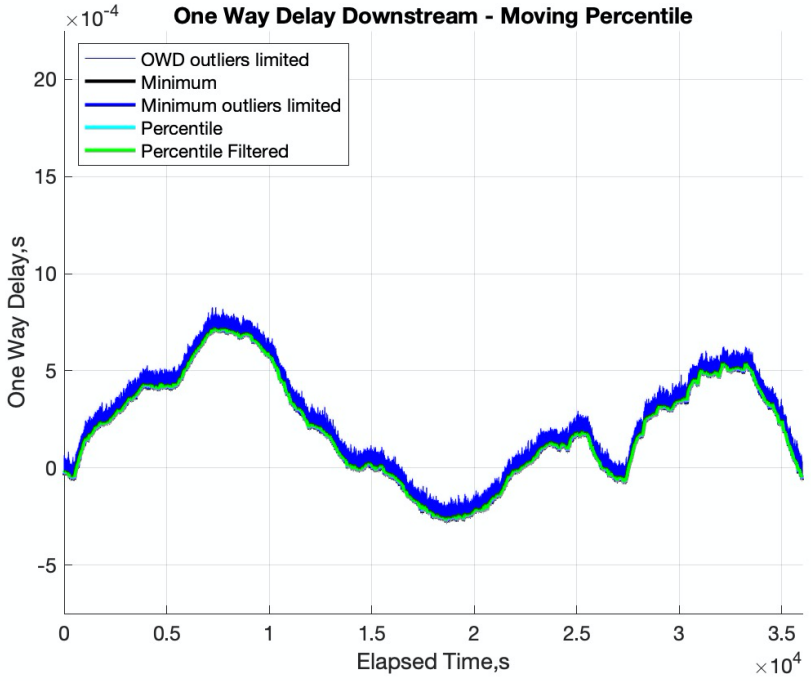




Time Determination: Filtered Moving Percentile

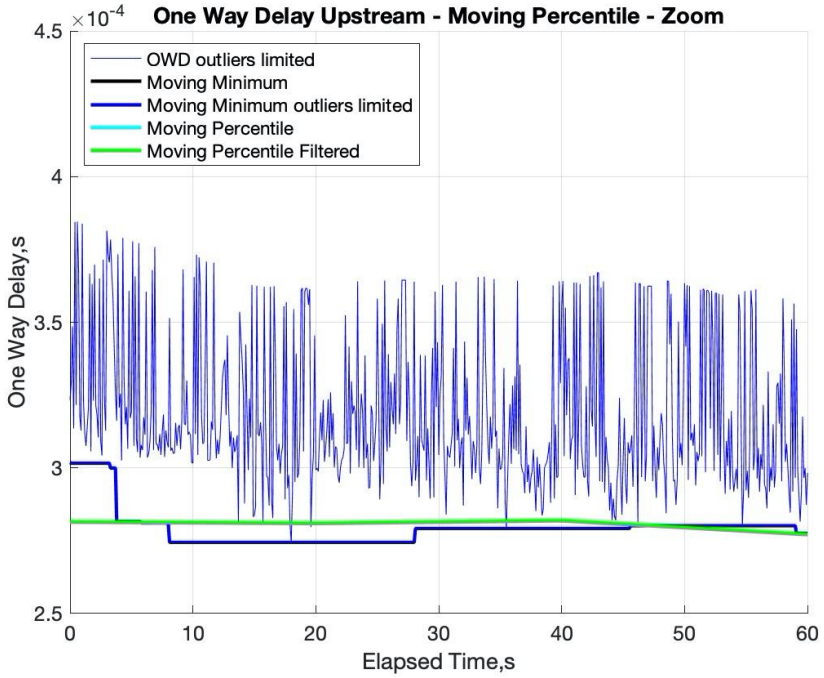
One Way Delay Estimation – Moving Percentile

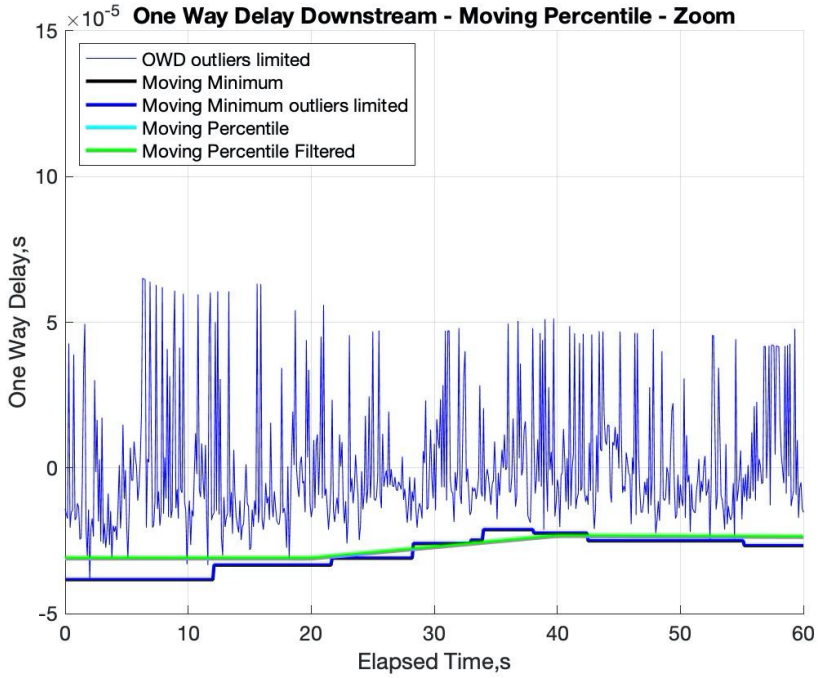




Time Determination: OWD Estimation

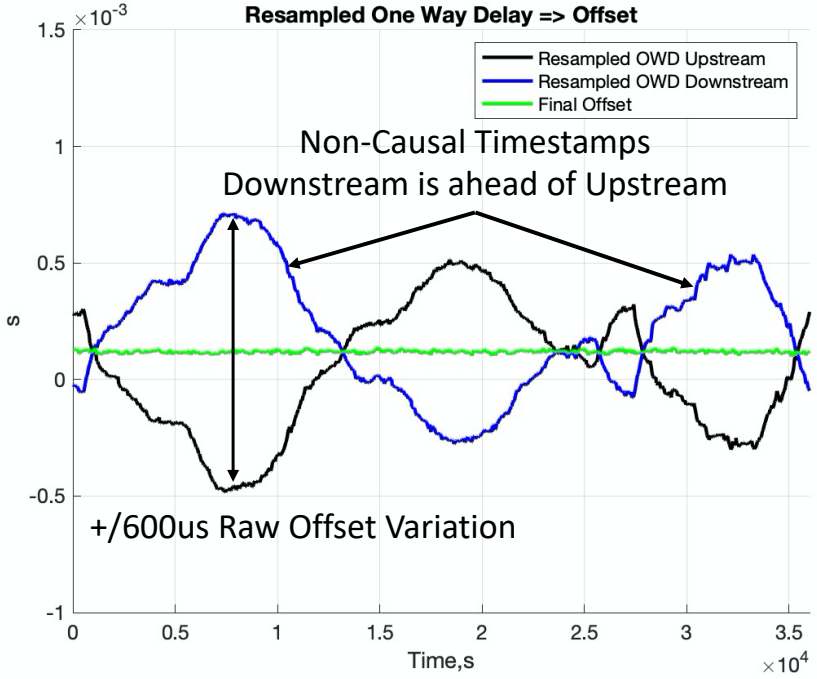
One Way Delay Estimation – Moving Percentile Zoom

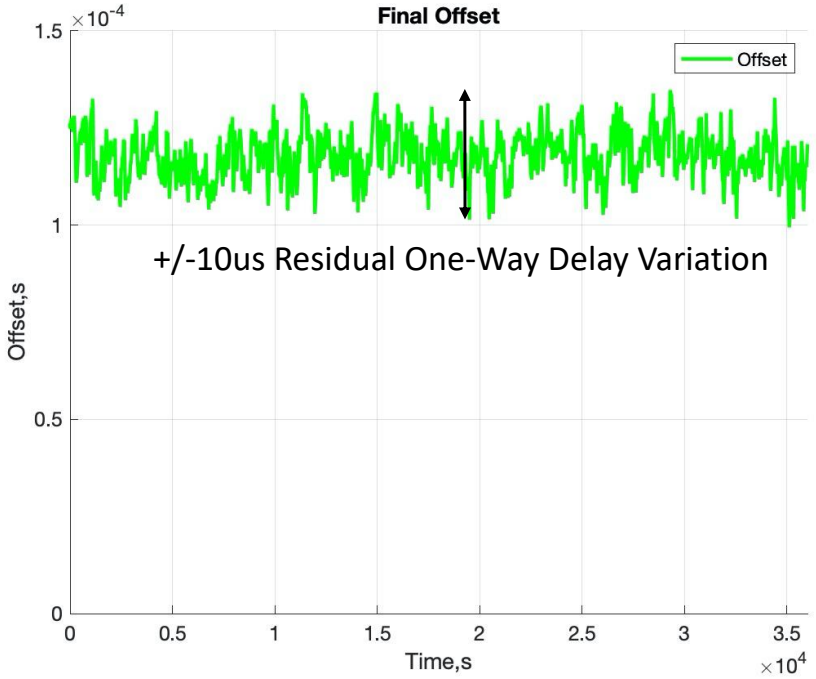




Time Determination: Offset Compensation

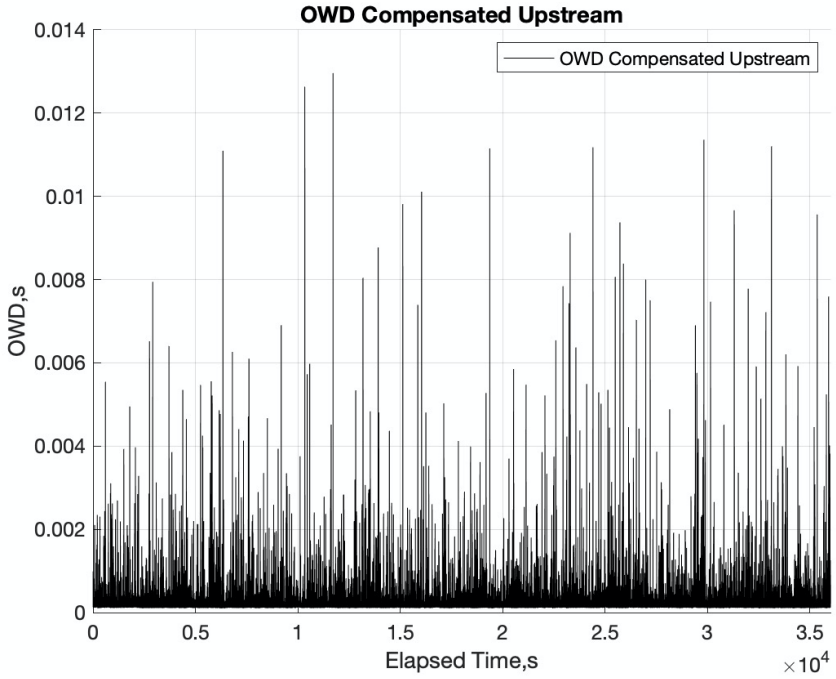
Offset Compensation – Time Aligned Traces

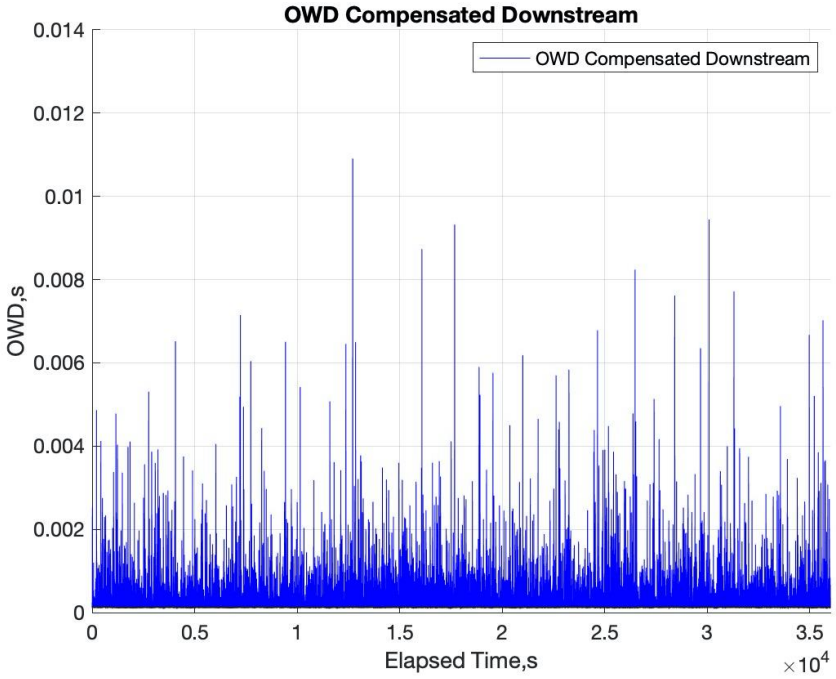




Time Determination: Time Aligned Traces

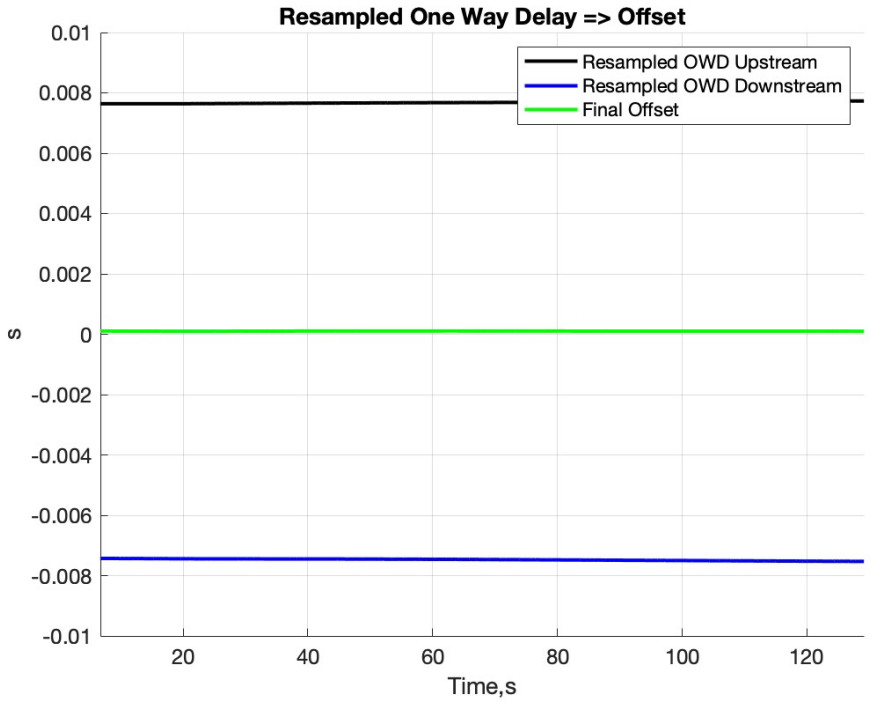
Final Time Aligned Traces

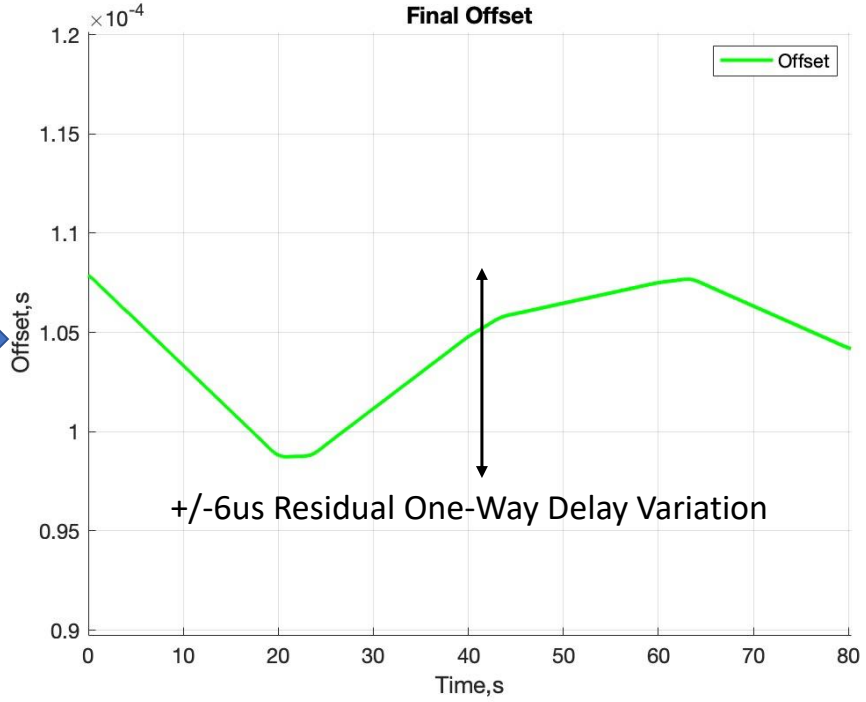




Time Determination: OWD Variation vs Packet/s

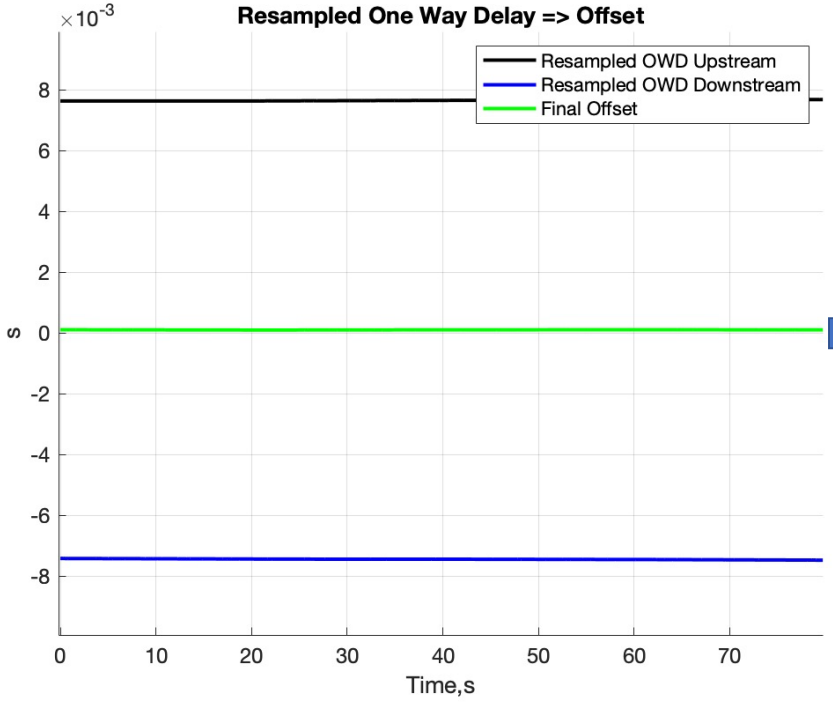
OWD and Time Aligned Offset – 10 packet/s

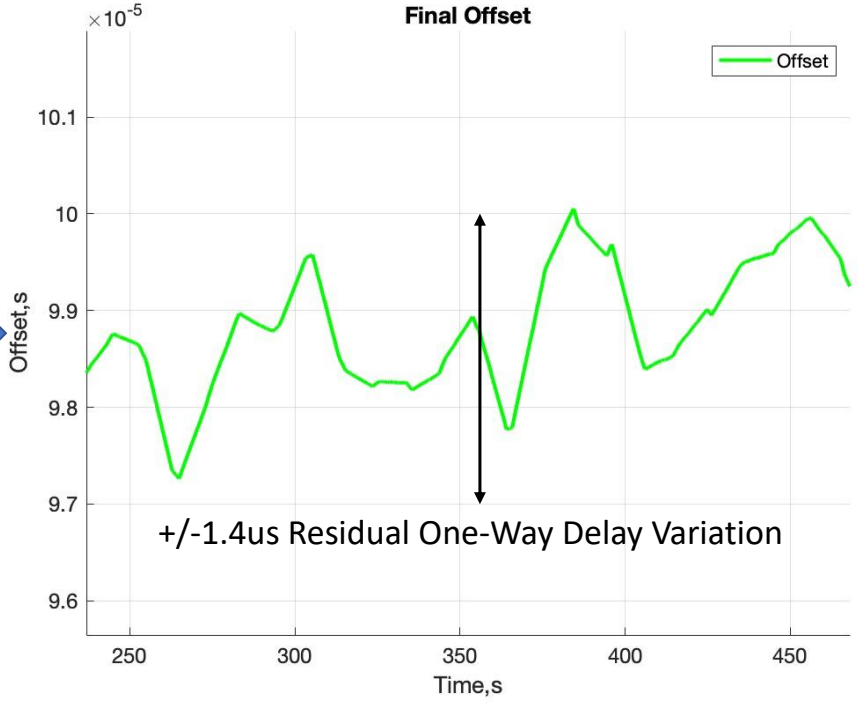




Time Determination: OWD Variation vs Packet/s

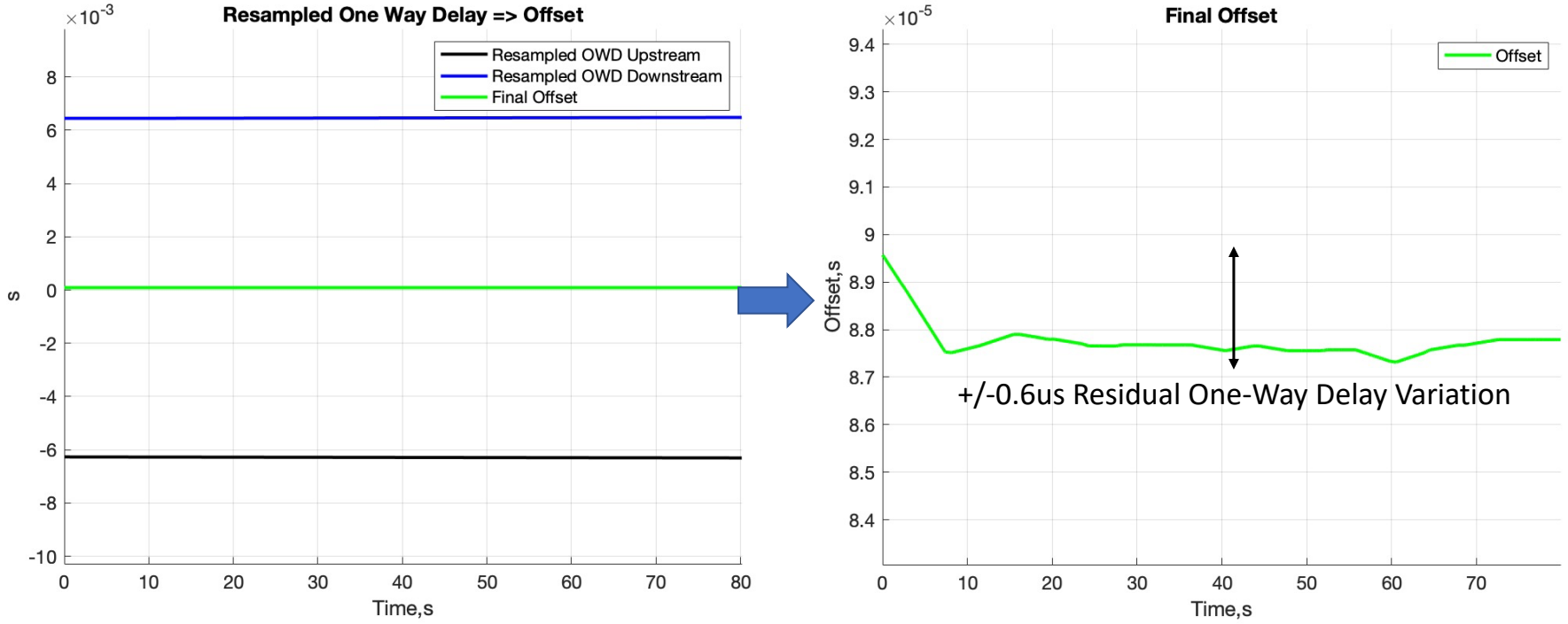
OWD and Time Aligned Offset – 100 packet/s





Time Determination: OWD Variation vs Packet/s

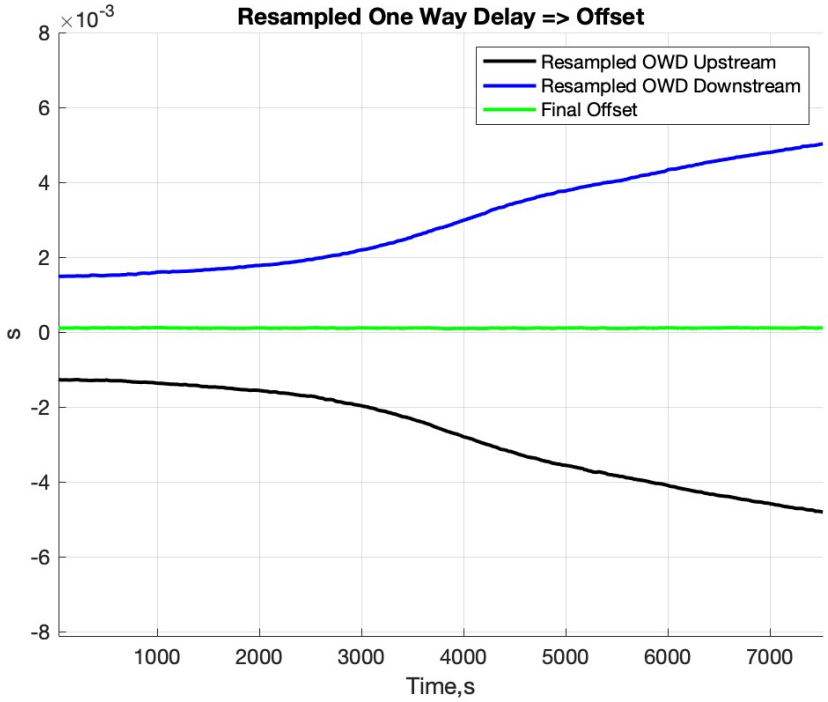
OWD and Time Aligned Offset – 1000 packet/s

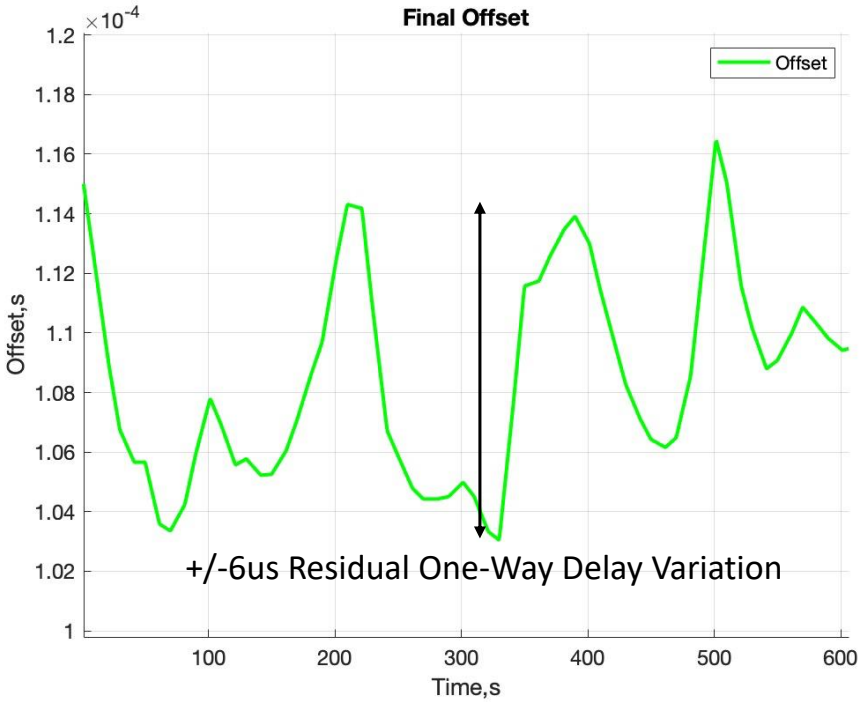


Packet Rate:	10	100	1000
OWD Variation +/- (us)	6-10	1.4	0.6

Time Determination: One Way Delay (Freerun)

Sync to Freerun Examples: mgsmds – oss-2

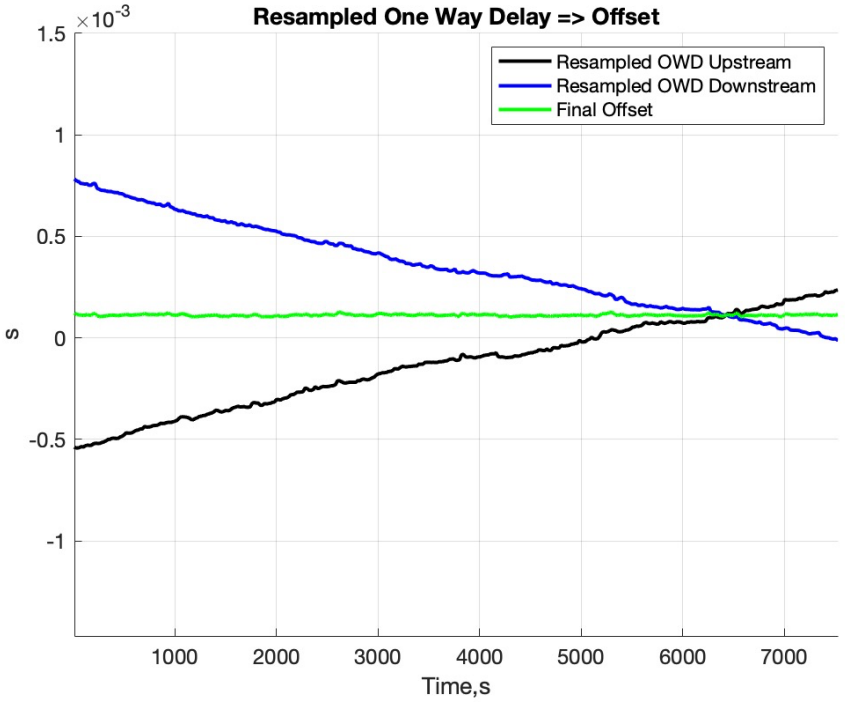


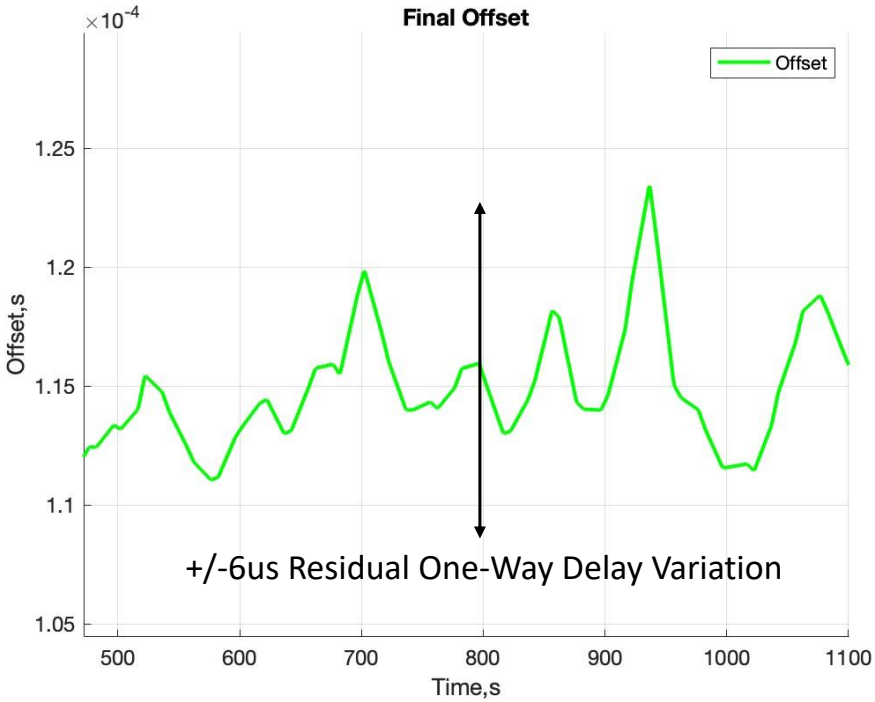


Freerun to TD Time aligned at 10 Packets/s => +/-6μs

Time Determination: One Way Delay (Freerun)

Sync to Freerun Examples: mgsmds – client-proxy





Freerun to TD Time aligned at 10 Packets/s => +/-6μs

Summary: TD Features/Benefits

FEATURE	TIME DISTRIBUTION (NTP,PTP)	TIME DETERMINATION (TD)	BENEFIT of TD
Protocol	Active	Passive	Scalability, Operational ease
Clock Recovery	Distributed Clients independently	Centralized Aggregator full network visibility	Causality is assured for Multi-Point Analytics
Packet Rate	NTP: 1/64s PTP: 128/s	Limited only by Link Rate, to 100M/s	Improved signal to noise in clock recovery
Packet Size	Fixed, typically 96B	All sizes, 64B - MTU	Better performance in multi-hop networks
Packet Class	Typically, Best Effort only	All Classes including Expedited Forwarding	Lower RTT; reduced asymmetry
Network Topology	End-to-end client-server path	Linear, Ring, Partial and Full mesh	Causality is assured for Multi-Point Analytics
Time to Sync Lock	Minutes to Hours	Seconds	Near Real-time analytics
Clock Offset Accuracy	<= End-to-End asymmetry; Causality is not assured	<= Single Hop asymmetry adjusted for causality	Causality is assured for Multi-Point Analytics

This material is based upon work supported by the U.S. Department of Energy, Office of Science, under Award Number DE-SC-0021595.