Robust and Fault Tolerant Clock Synchronization

Nikolaus Kerö, Oregano Systems **Aneeq Mahmood, ZISS**
Thomas Kernen, Cisco **Felix Ring, ZISS**
Tobias Müller, Oregano Systems **Thomas Bigler, ZISS**

Rationale

- **Common notion of time is crucial for almost all application**
- **Ethernet has become the only viable communication medium**
  - Performance is highly customizable
  - Bandwidth, fault tolerance, latency
- **One single communication medium both for and time**
  - R.I.P. Legacy time transfer systems
- **Ethernet is asynchronous**
  - Time transfer has to be packet based
- **PTP – Precision Time Protocol**

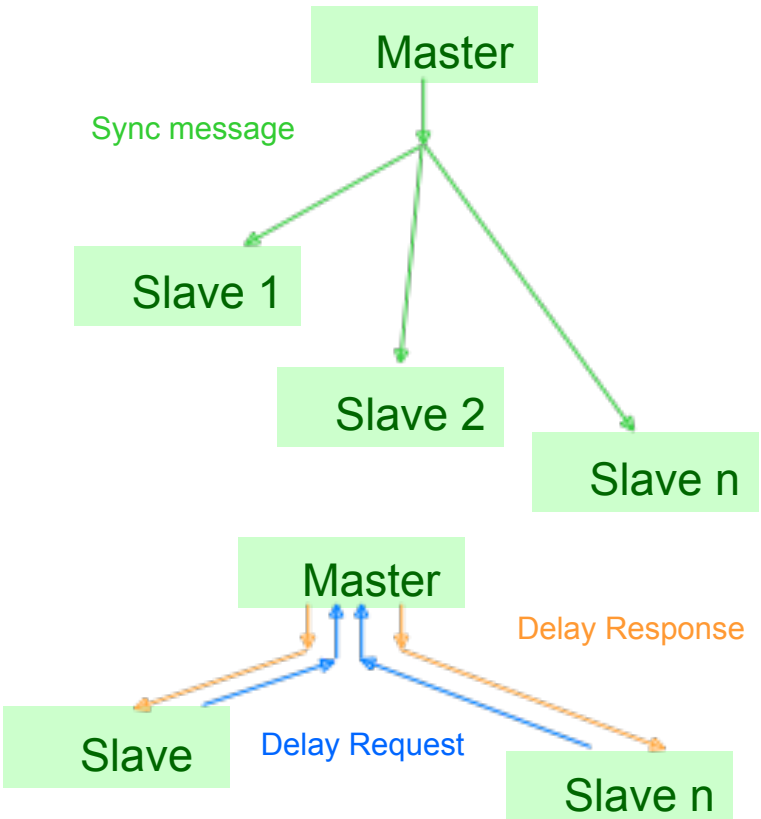
Rationale

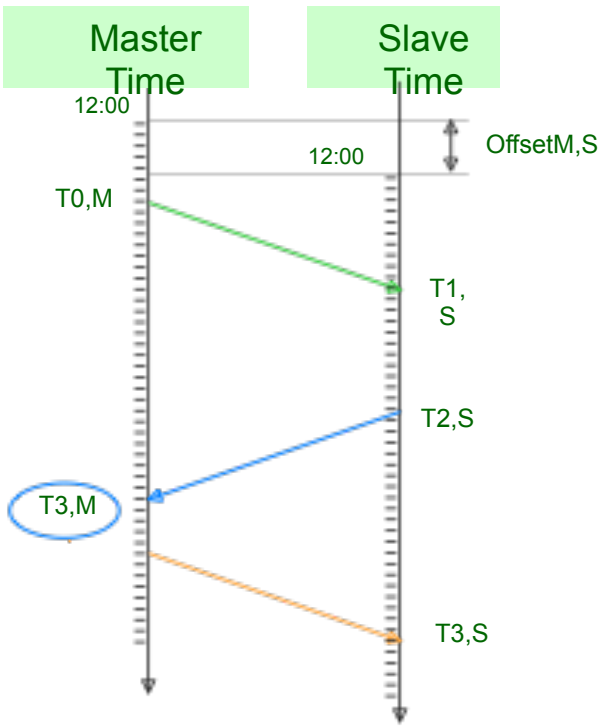
- **PTP – Precision Time Protocol**
  - IEEE1588 Standard
  - Generic protocol for highly accurate time transfer over Ethernet
  - Customizable via PTP-Profiles

- **Adopted as the only means of clock synchronization by various industries**
  - Telecom
  - Power Utility Providers
  - Finance Service Providers
  - Broadcasting

- **Accurate time transfer has become mission critical**

Basic Principles of PTP





$$T_{1,s} - T_0$$

$$Offset =$$

Accuracy in PTP

- **Node dependent**
  - Time stamping resolution
  - Oscillator quality
- **Network related**
  - Packet Delay Variations
  - Performance of PTP aware network devices
  - Different paths upstream and downstream
  - Highly asymmetric network loading
- **Configuration dependent**
  - Message rates

Requirements for Mission Critical Applications

- **Permanent Fault Conditions**
  - Failure of Master
  - Missing PTP event messages
  - Partial network failures
- **Transient Fault Conditions**
  - Packet Delay Variations
  - Path changes
  - Event Message loss
- **Fault Awareness vs Fault tolerance**

Fault Tolerance in PTP (I)

- **BMCA – Best Master Clock Algorithm**
- **Based on PTP Announce messages**
  - Every Master Advertises its clock quality
- **Trigger Conditions**
  - Timeout (Absence of Announce Messages)
  - Content … A better (more accurate) Master joins the network
- **Eventually the best Master will always take over**
- **Extremely robust, but …**

Fault Tolerance in PTP (II)

- **BMCA addresses only a subset of fault conditions**
- **Absence of PTP event messages remains undetected**
  - Errors in PTP devices
  - Misconfigured network and end devices
- **Sudden changes in the network status**
  - Path changes
  - Overloading
- **PTP defined ONLY a protocol but NO specification of the servo**
  - Reaction to network load changes is implementation dependent

Extended PTP Features

- **A number of topics are discussed for the next revision of PTP**
  - Redundancy
  - Security
  - Monitoring
  - …
- **In the meantime there are industry proposals and R&D approaches for**
  - Extended Monitoring
  - Active Redundancy
- **Basic Requirement**
  - PTP Standard must not be violated
  - Any extension has to be standard compliant

Why Monitor a Complete PTP Network?

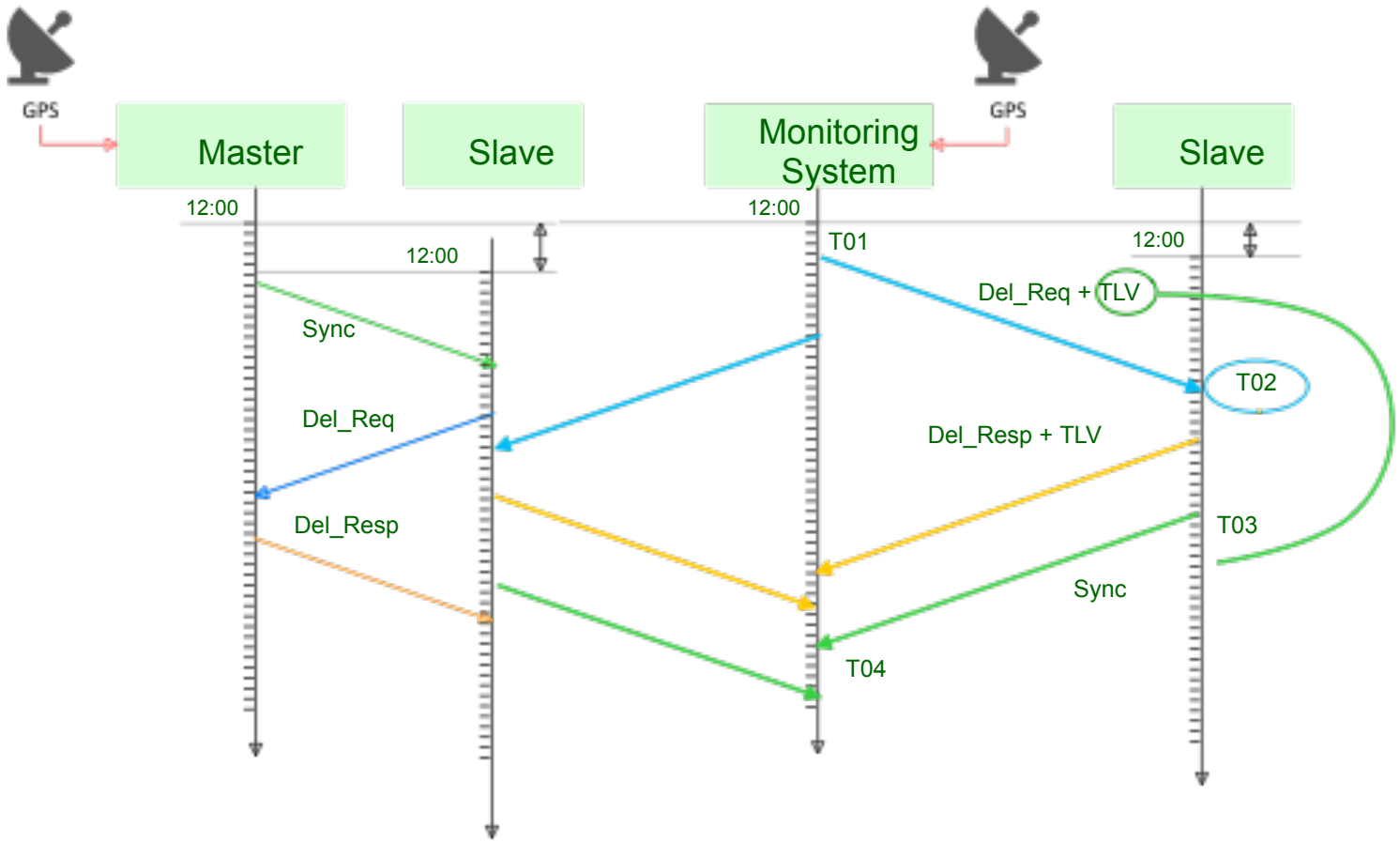
- **Querying Slaves is insufficient**
- **Only current offset as seen by the Slave**
  - Asymmetries are not accounted for

- **Out-of-Band measurement techniques**

- **All auxiliary Masters are hot stand-by devices**
- **Failures of Auxiliary Master have to be detected prior to engaging them**
  - Path to auxiliary Master could be broken
  - Clock quality could have degraded
  - Loss of external time source

- **PTP Network Device need to be monitored as well**

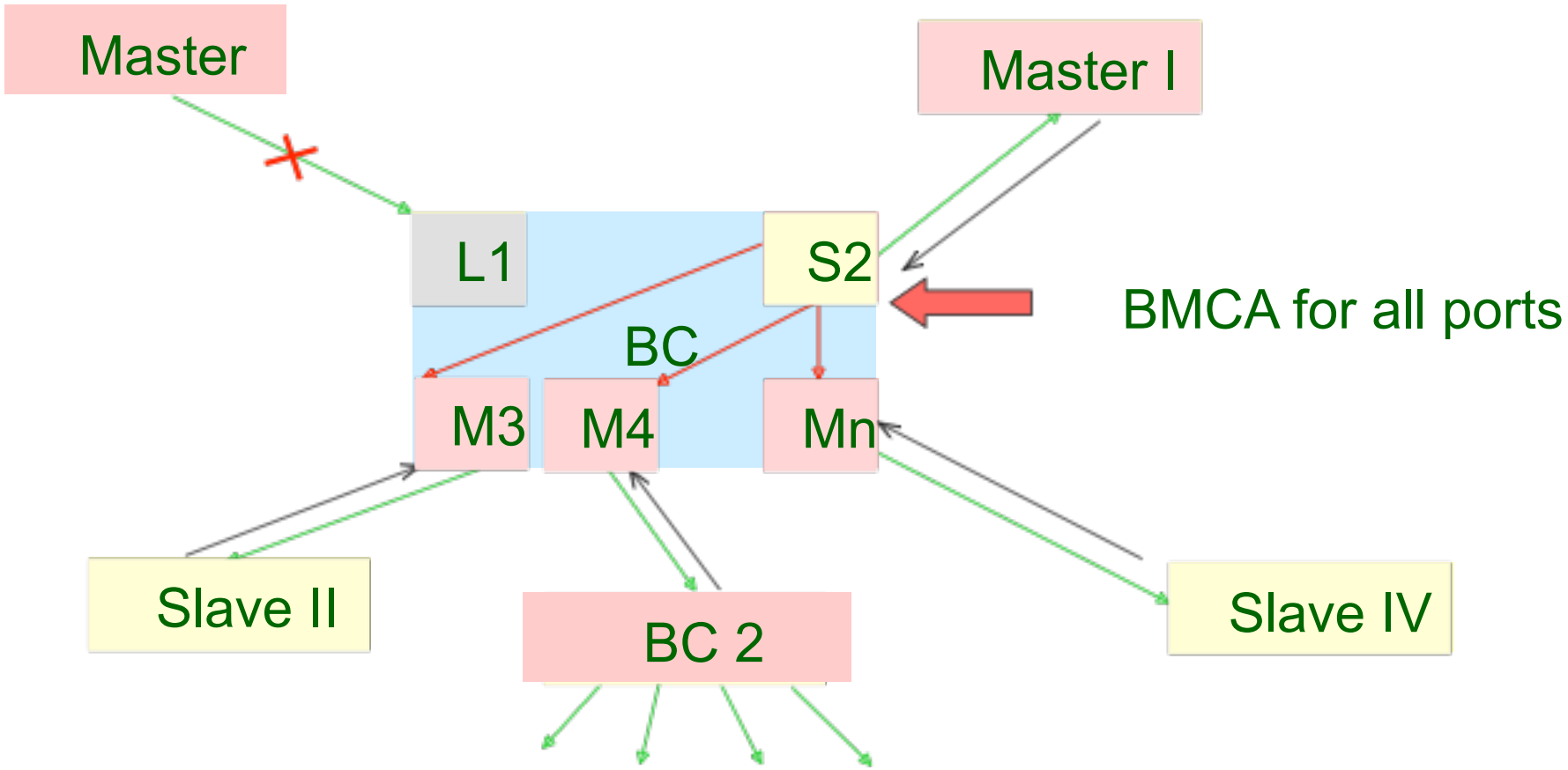
Extended Monitoring



$$T_{02} - T_{}$$

$$Offset =$$

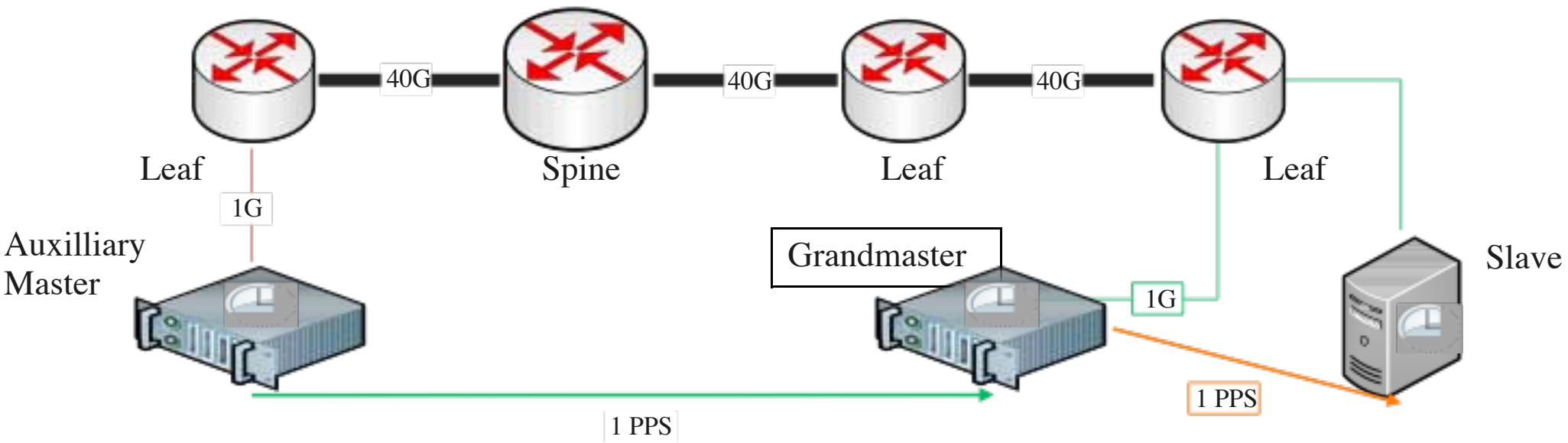
PTP Boundary Clock



Mesurment Setup

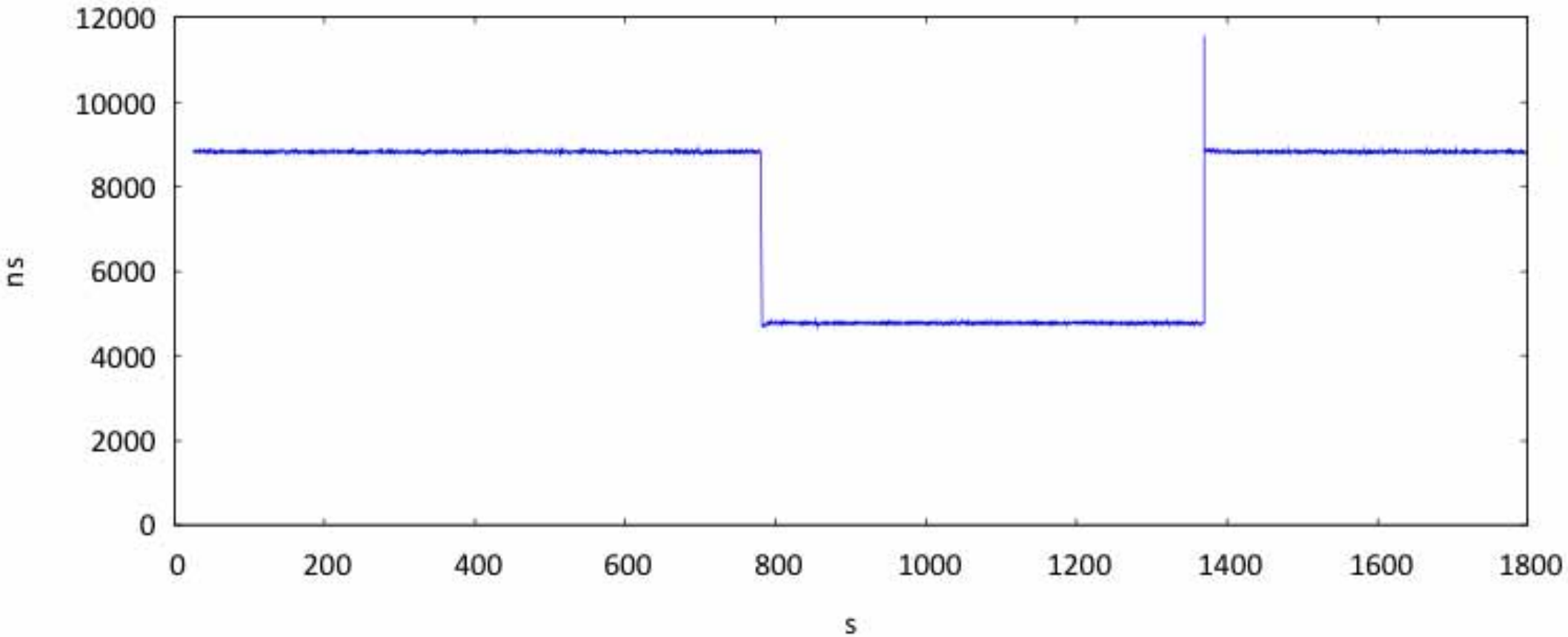
- **Network devices**
  - Standard datacenter products with line rate capabilities
  - Built-in hardware PTP support (Boundary Clocks)
- **2 different PTP Grandmaster vendors**
  - Both Masters synchronized to each other externally
- **1 PTP Slave**
- **PTP message rates**
  - 8/16 Sync and Delay_Request Messages per second
- **Master changes invoked by disabling the respective port**

PTP Unaware Network Topology

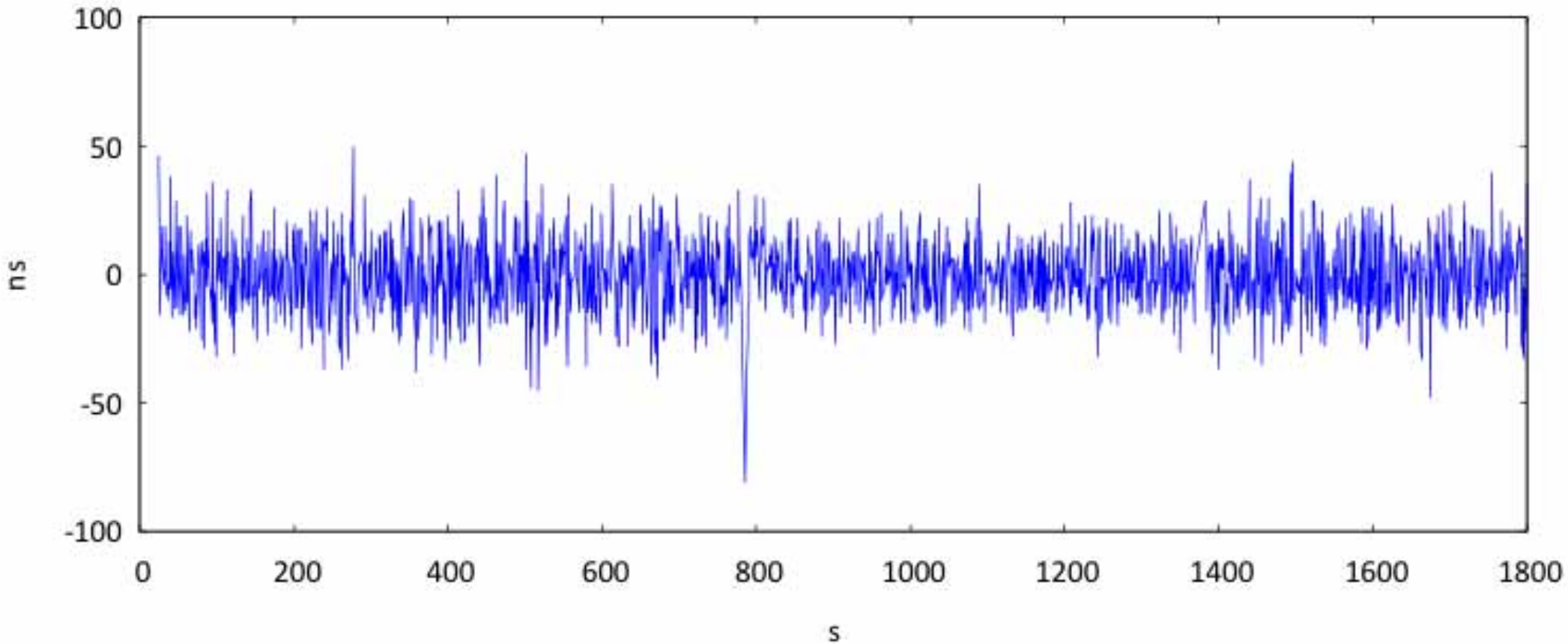


Out-of-Band Measurement

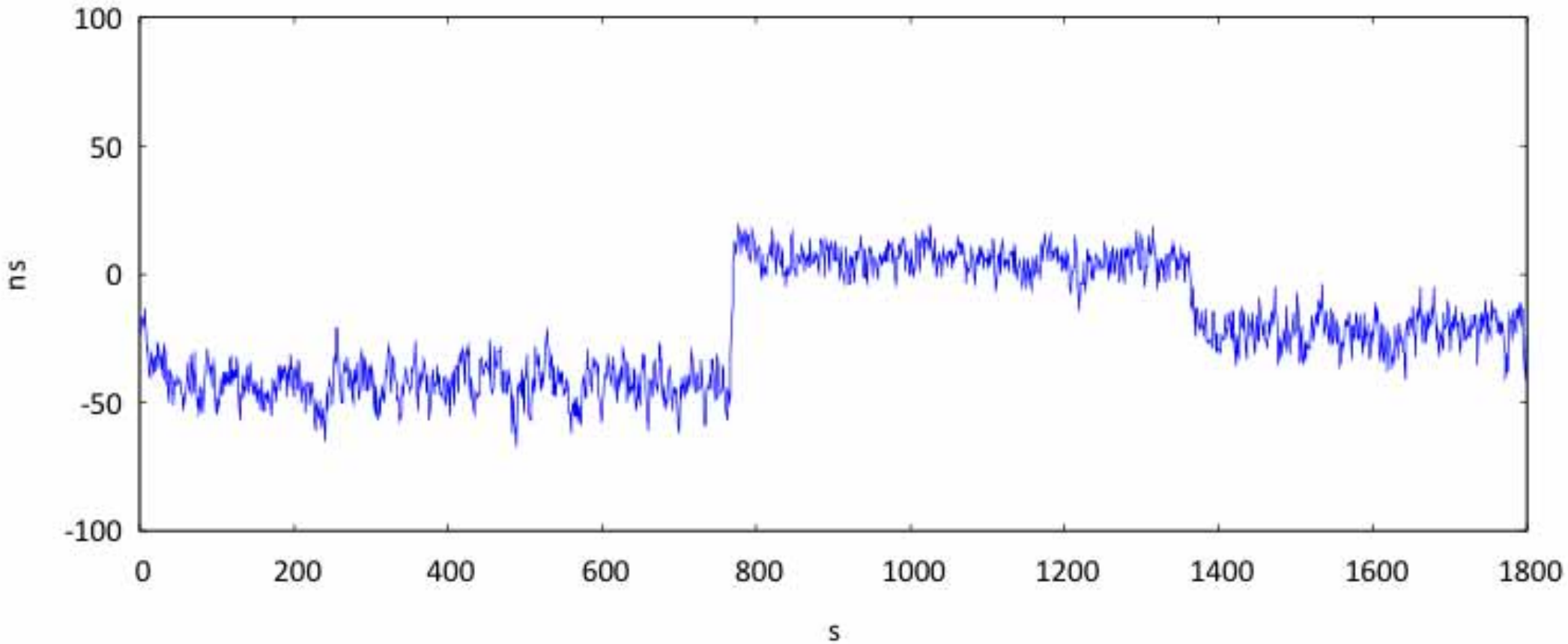
Master to Slave Delay



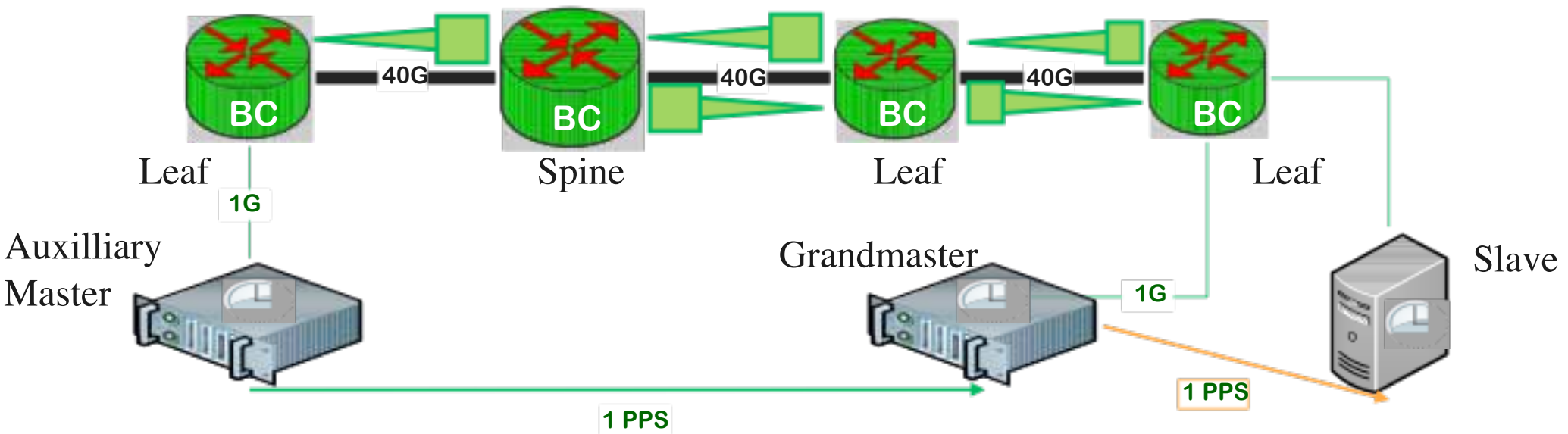
Filtered Offset viewed from the Slave



Offset measured out-of-band

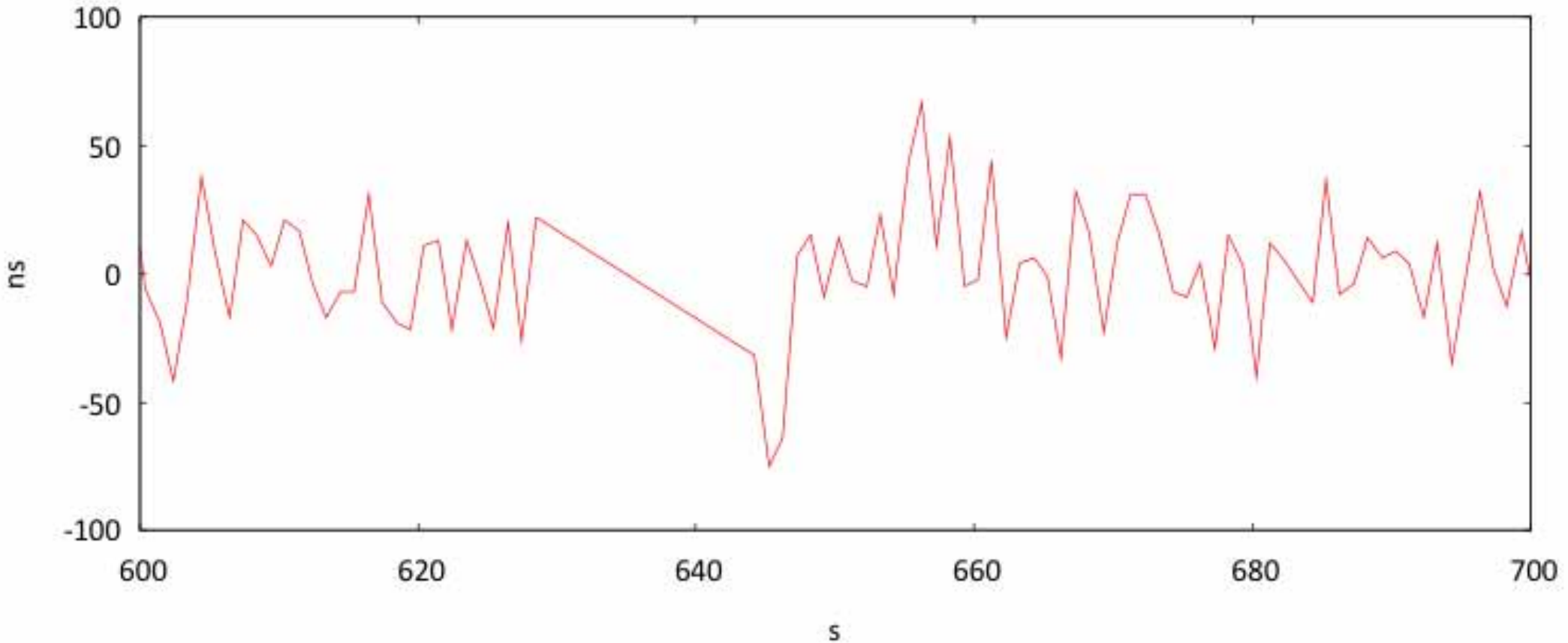


PTP Aware Network Topology

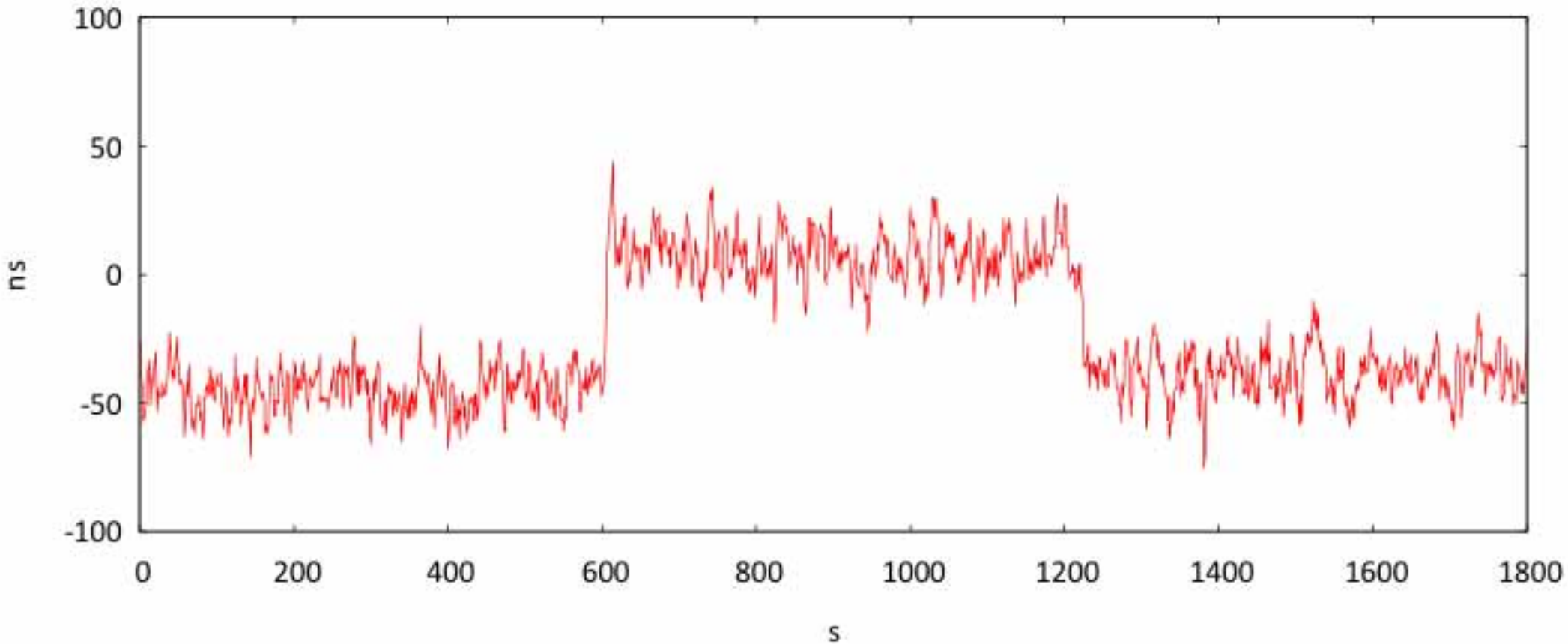


Out-of-Band Measurement

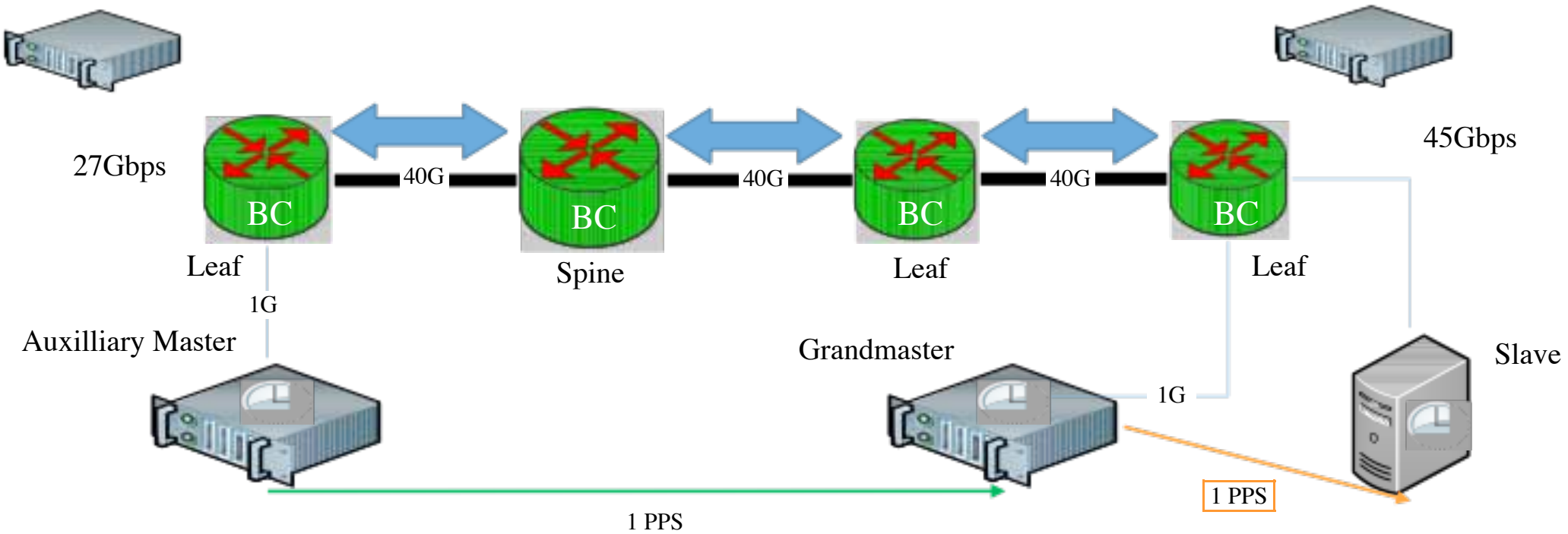
Filtered Offset viewed from the Slave



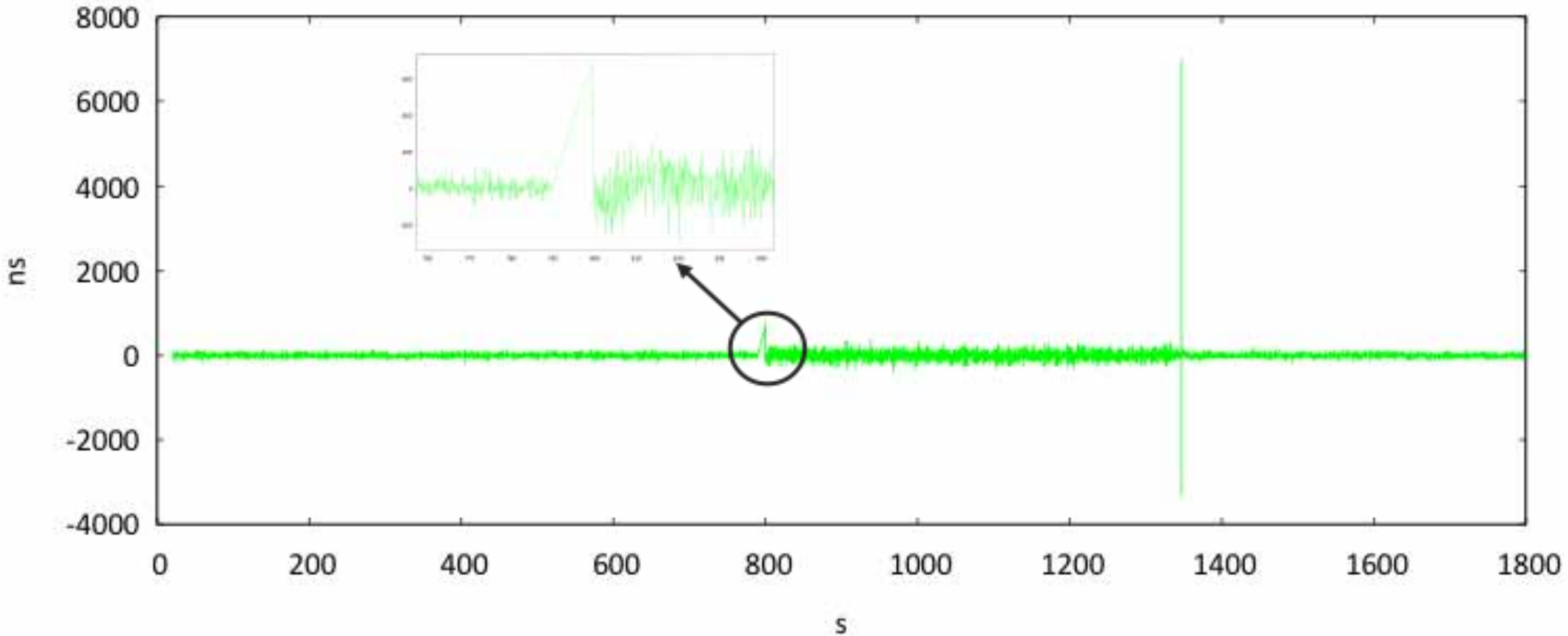
Offset measured out-of-band



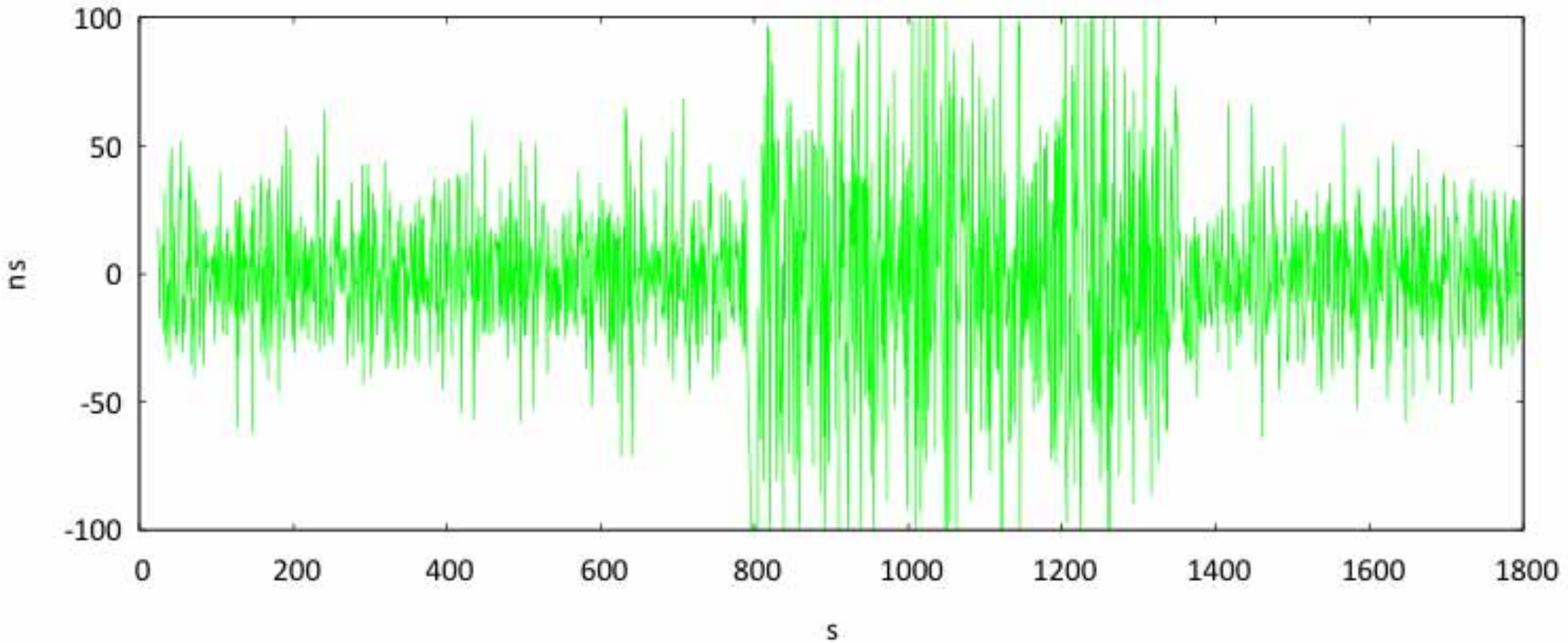
PTP Aware Network Topology with Load



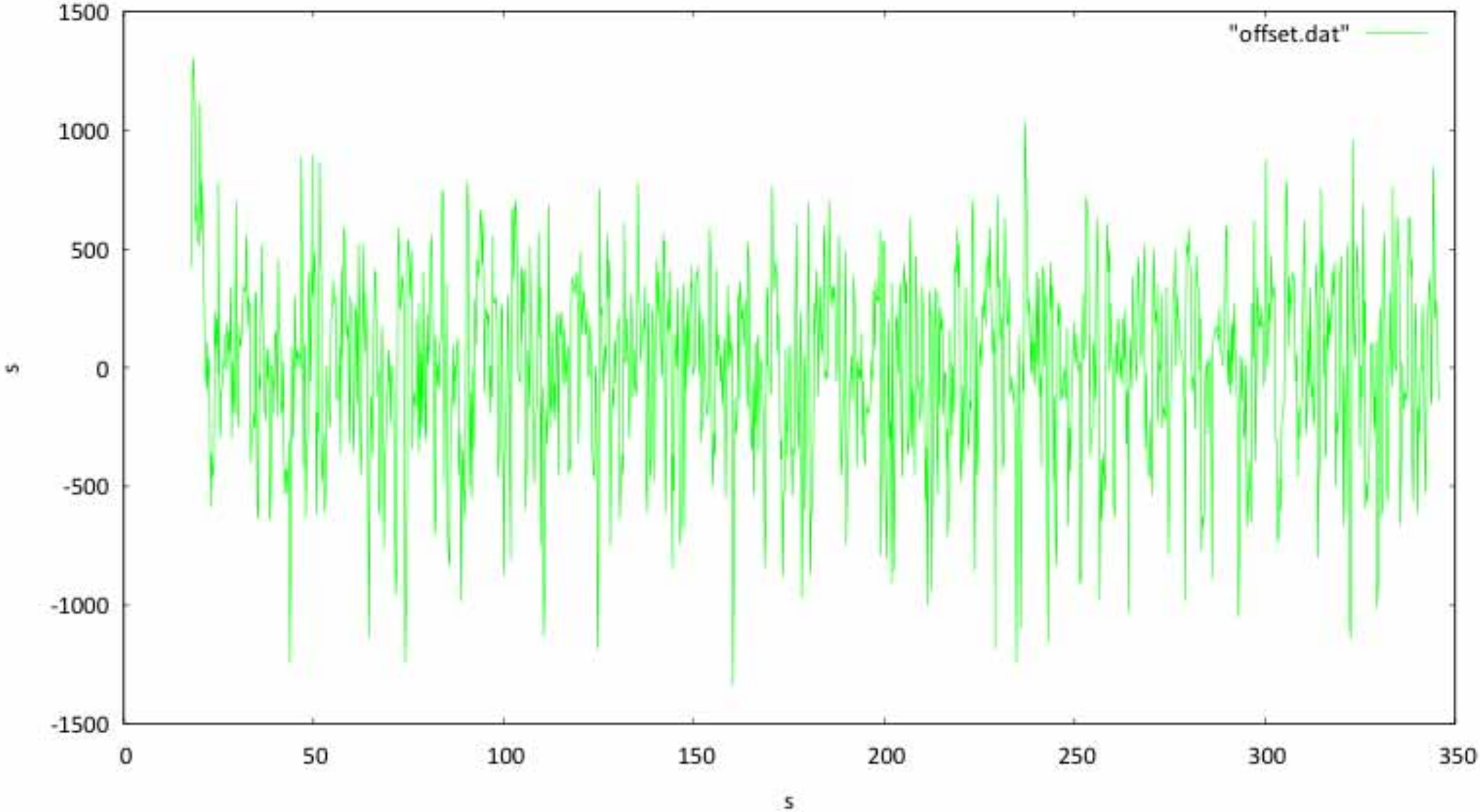
Offset between Master and Slave



Filtered Offset viewed from the Slave

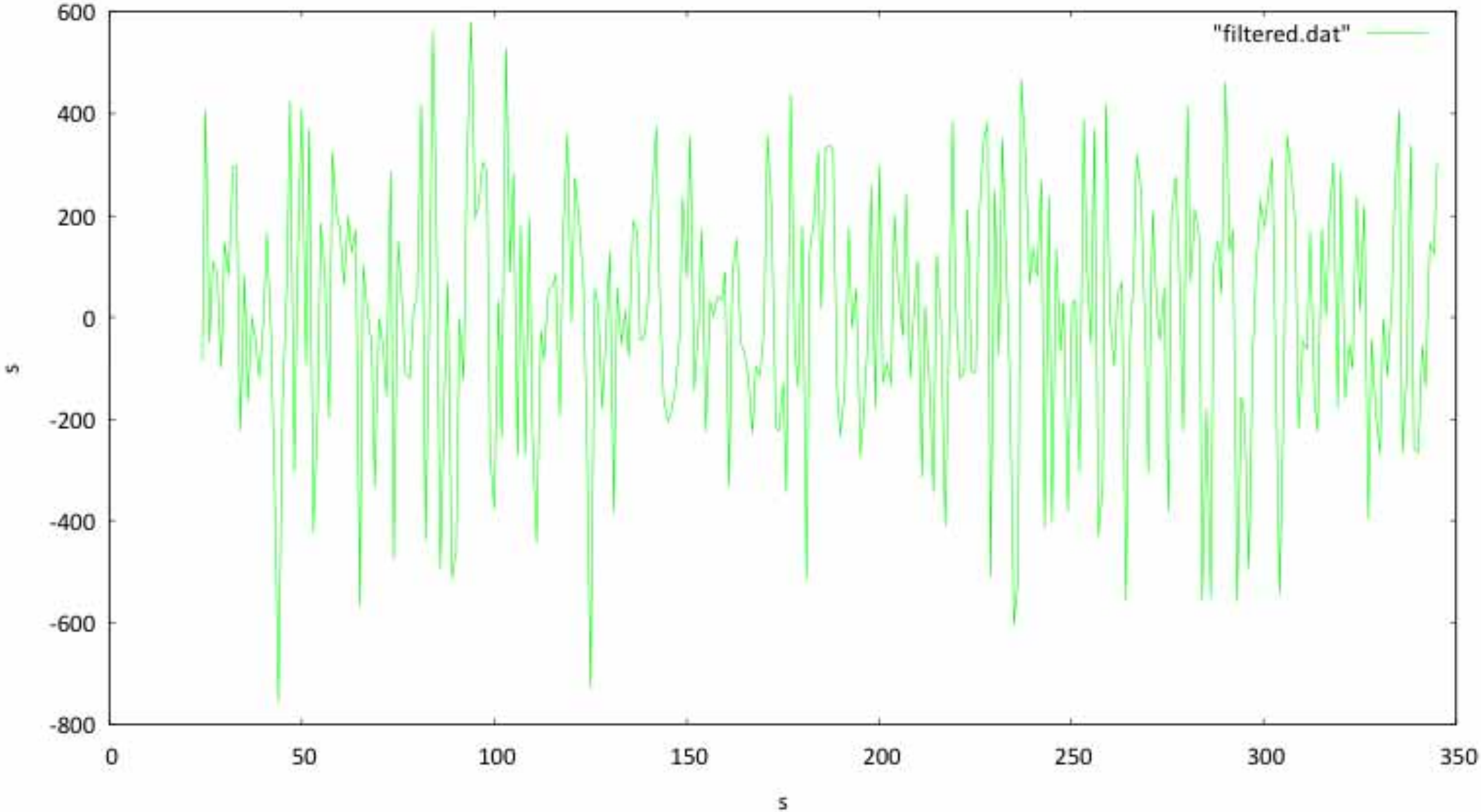


9 Hops with Partial PTP Support (BCs) – 120% load



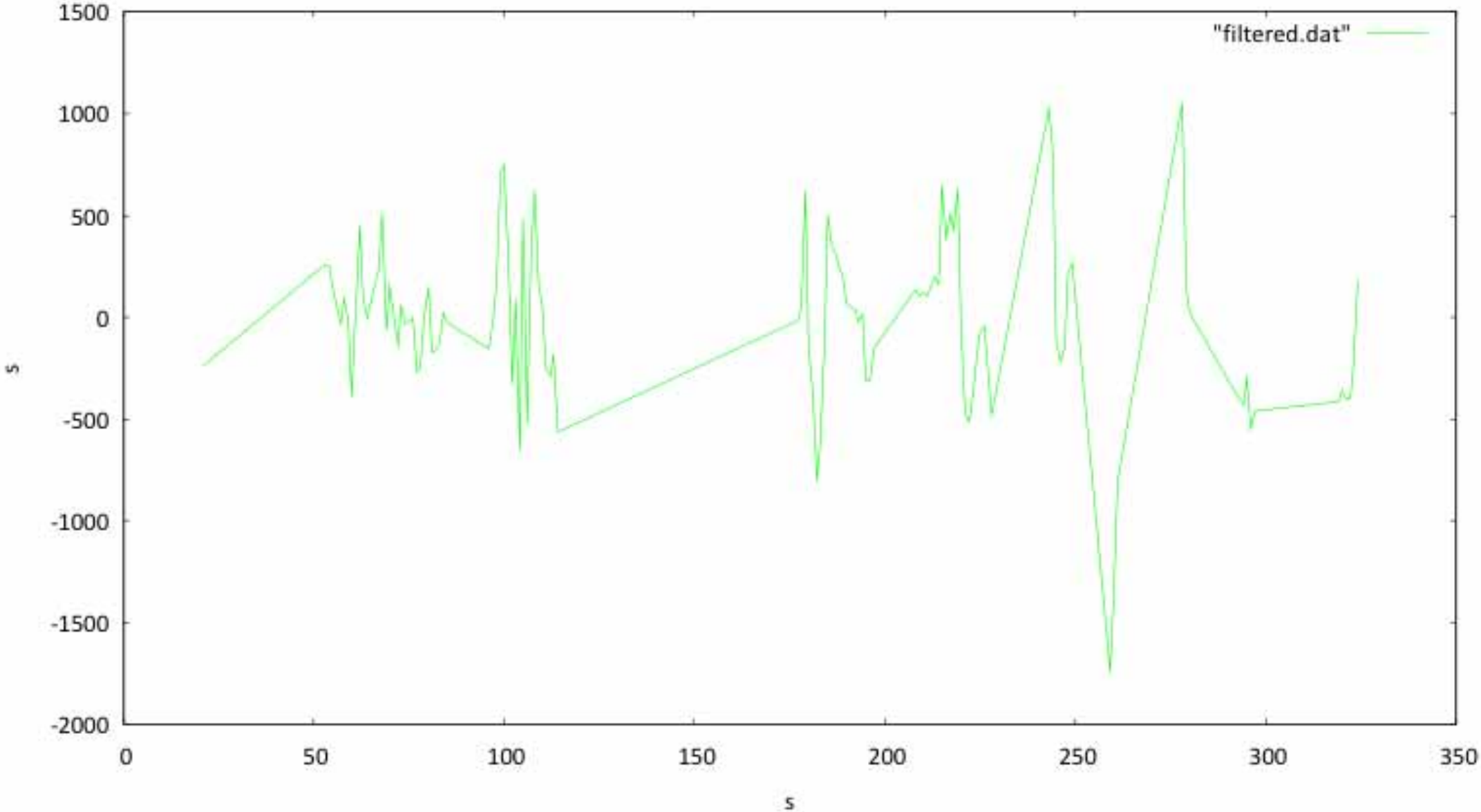
Oregano Systems
The Flavour of Excellence

9 Hops with Partial PTP Support (BCs) – 120% load



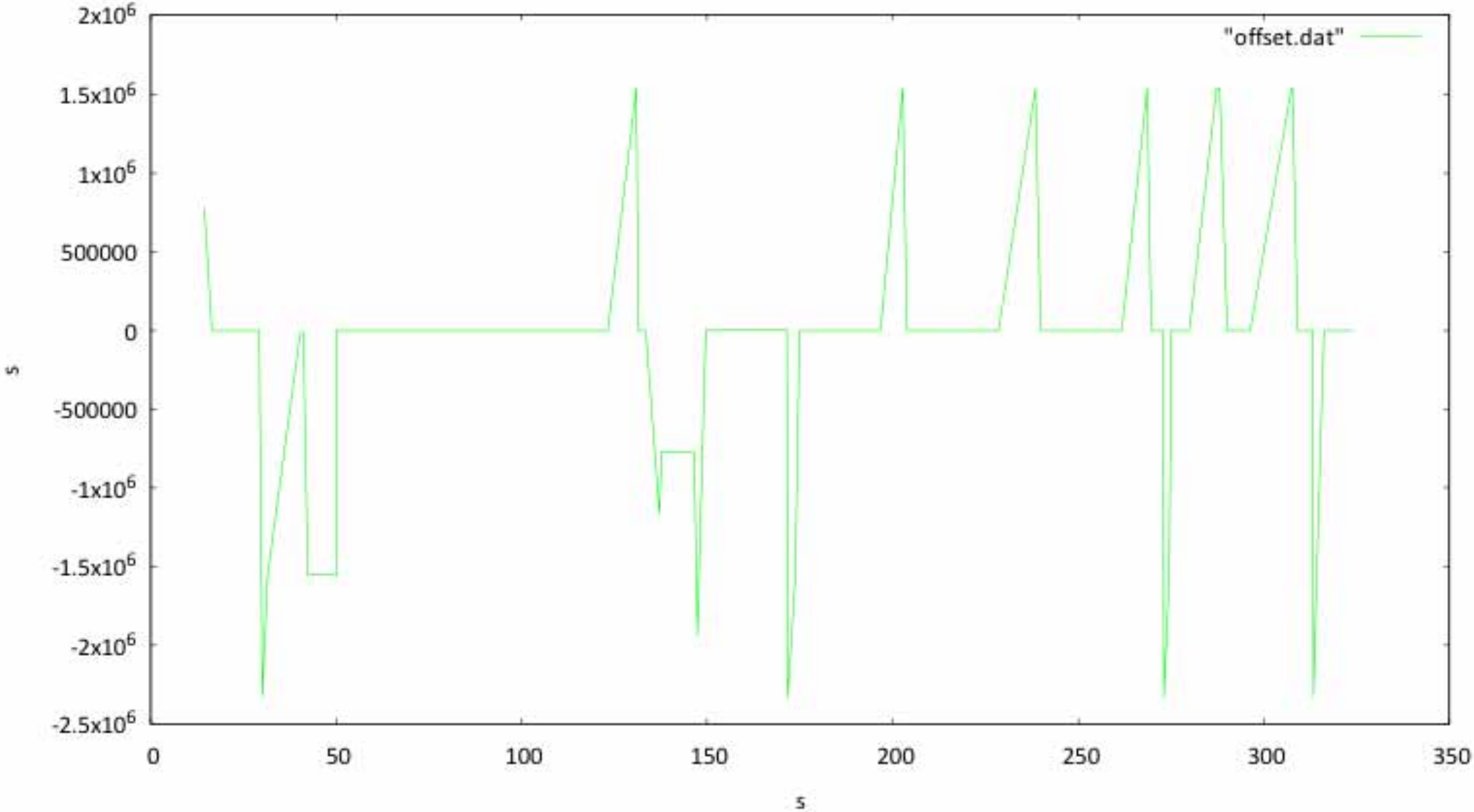
Oregano Systems
The Flavour of Excellence

9 Hops no PTP Support – 120% load



Oregano Systems
The Flavour of Excellence

9 Hops no PTP Support – 120% load



Oregano Systems
The Flavour of Excellence

Use data of more than 1 Master (I)

- **Even extended filtering may prove inadequate to cope with PDVs**
  - Outlier detection
  - Linear input filter
  - Servo optimization
- **How to implemented within PTP?**
  - Use unicast
  - Use different domains
- **Enhancements at the Slaves**
  - Independent protocol engines for every Master
- **One single (physical) clock**

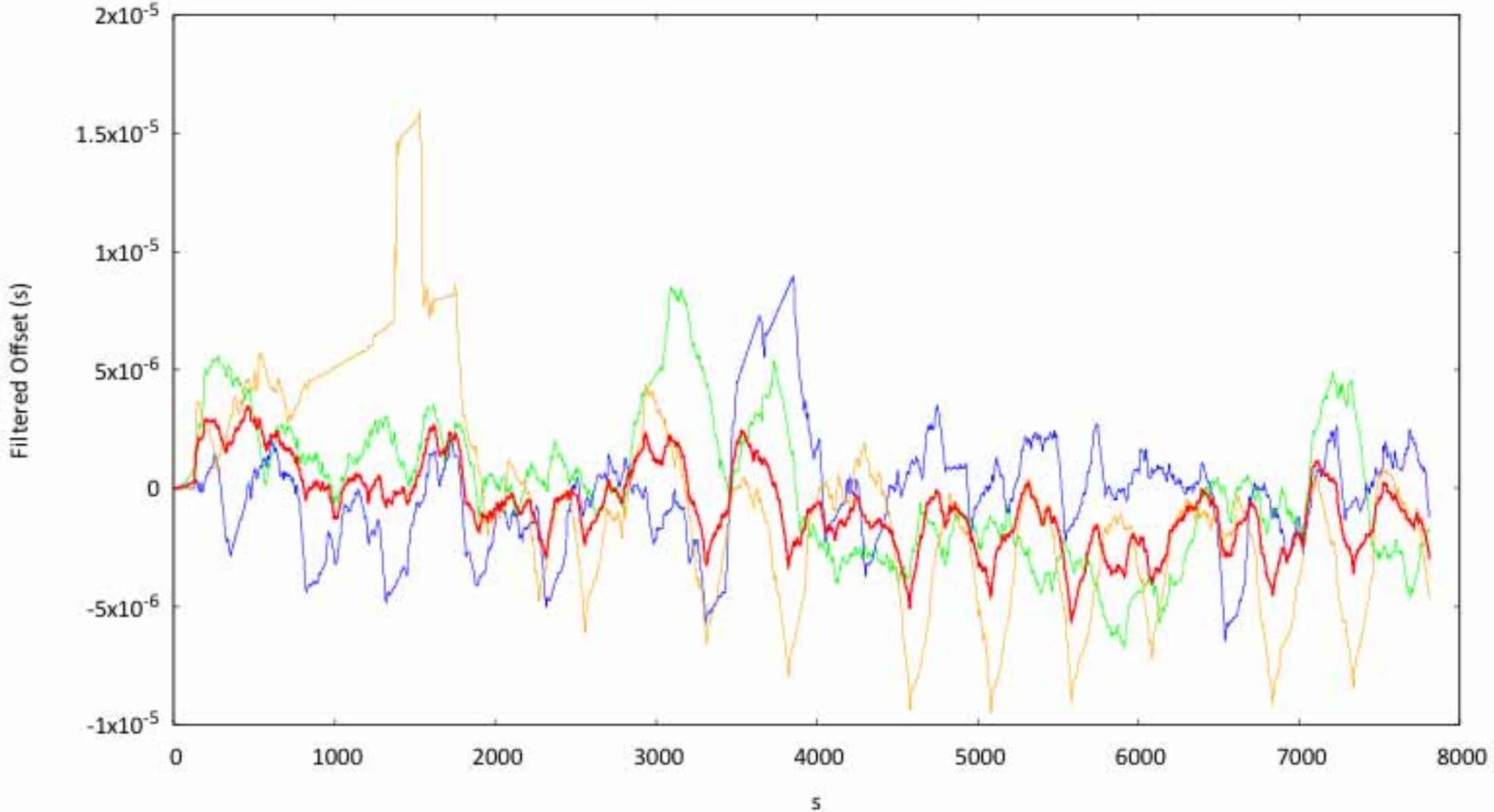
Use data of more than 1 Master (II)

- **Advantages**
  - Cope with deteriorations of the signal path
  - Discard faulty Masters
  - Handle Byzantine faults

- **How to combine the timing information of more than on master**
  - Simple linear combination (mean value) insufficient

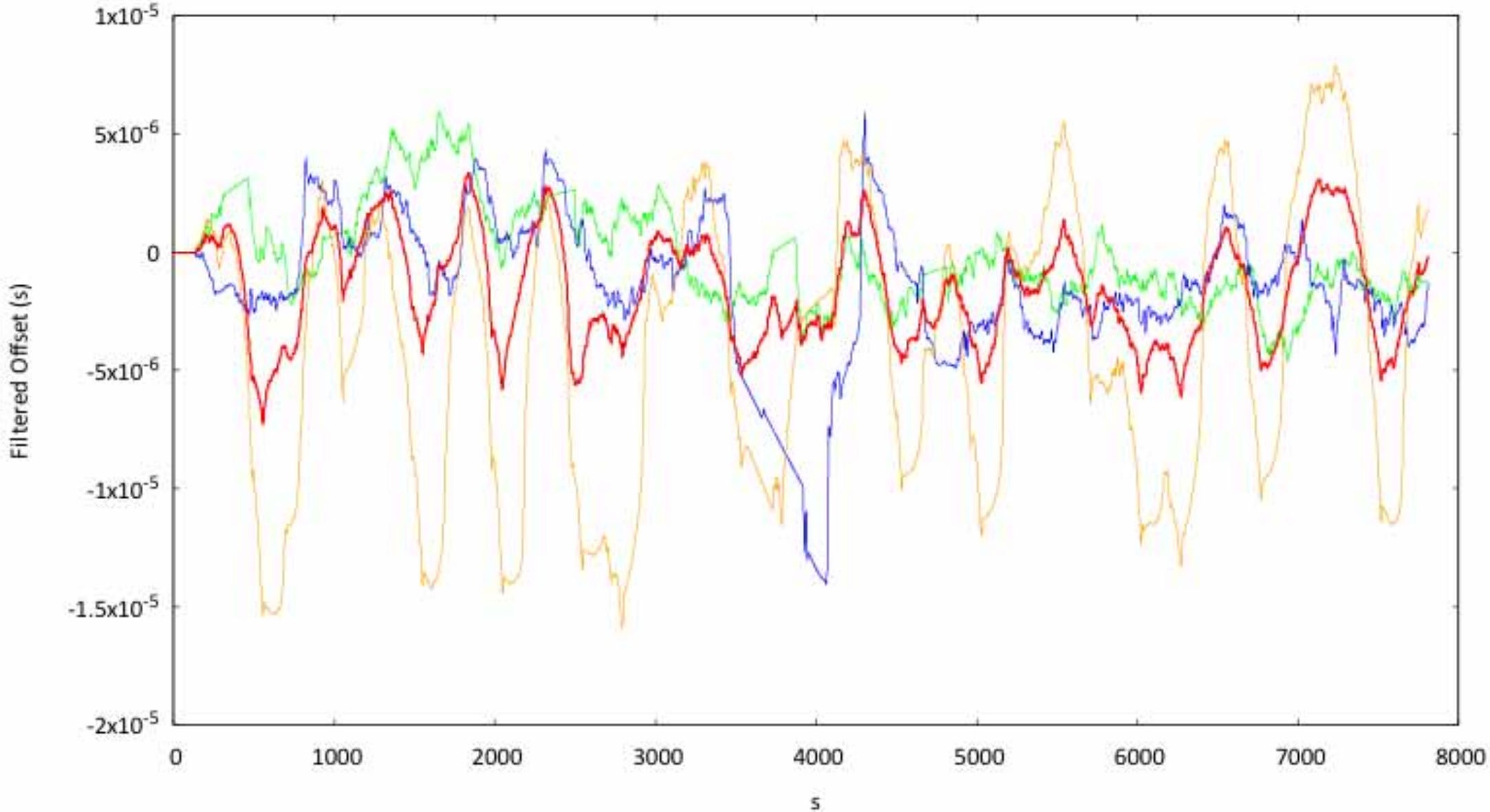
- **Solution**
  - Outlier detection and front end
  - Kalman filter to combine data of multiple Masters
  - PI Servo

Combining Data From 3 Different Masters



Oregano Systems
The Flavour of Excellence

Combining Data From 3 Different Masters



Oregano Systems
The Flavour of Excellence

Conclusions

- **If synchronization is mission critical BMCA may provide insufficient fault tolerance**
  - A number of transient and permanent failures remain undetected
- **Extensive monitoring is crucial**
  - Monitoring systems using out-of-band techniques
- **Combination of more than 1 PTP Masters**
  - Can be accomplished within the IEEE1588 standard
  - Can cope with large PDVs
  - Can cope with faulty Masters and malicious attack
- **Extensions in Firmware and control loop at Slave**

syn1588 ®

Thank You for Your Attention

Strontium-ion optical clock at NPL UK - 100ns deviation per 100 years

Oregano Systems
The Flavour of Excellence

Backup Slides

Oregano Systems
The Flavour of Excellence

Measurement Setup

- **Classic Datacenter Spine-Leaf architecture**
  - Additional hop to verify, if any impact in longer network chain
- **Network devices**
  - Standard datacenter products with line rate capabilities
  - Built-in hardware PTP support (Boundary Clocks)
- **2 different PTP Grandmaster vendors**
  - Both Masters synchronized to each other externally
- **1 PTP Slave**
- **PTP message rates**
  - 4 Announce message per second (log2 -2)
  - Announce timeout = 3
  - 16 Sync and Delay_Request Messages per second (log2 -4)
- **Master changes invoked by disabling the respective port**