

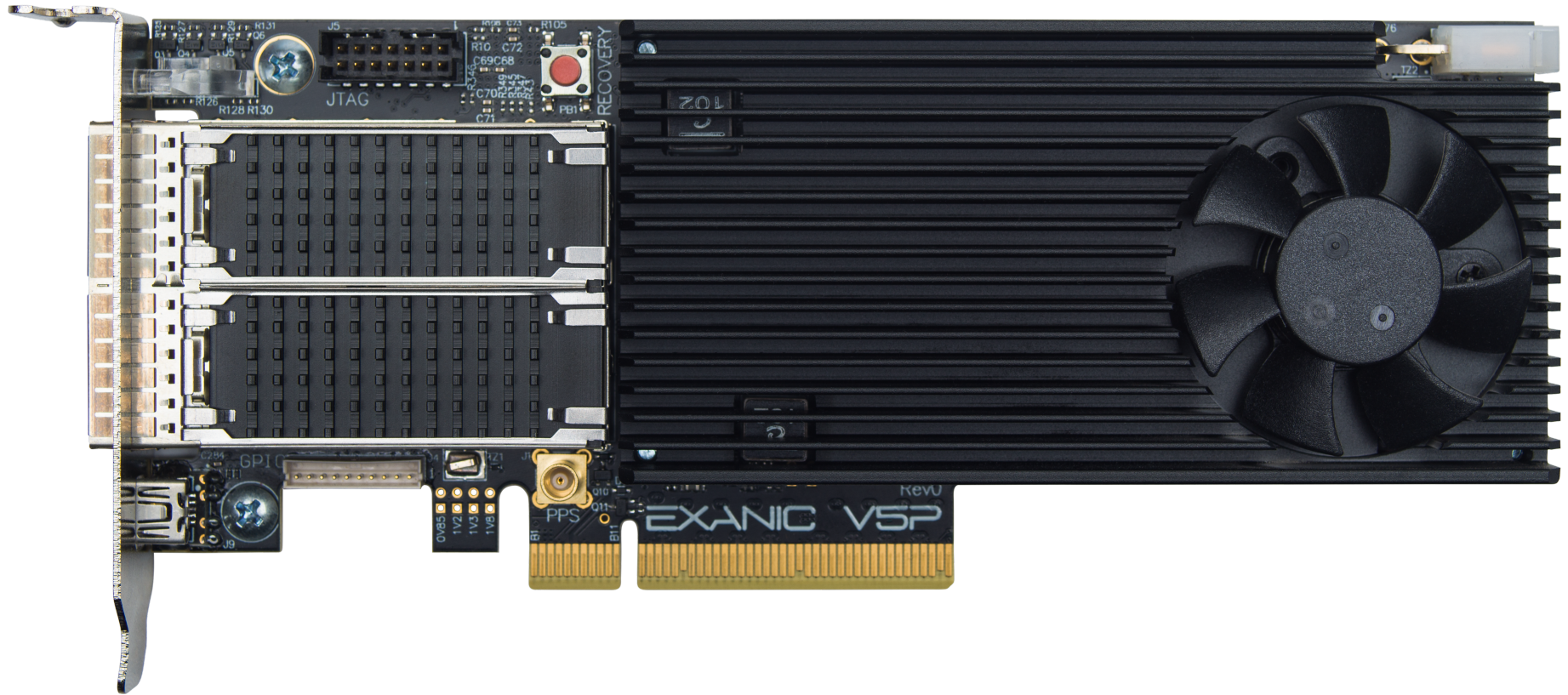


# Plain Old PTP Better than you think?

Dr Matthew Grosvenor  
STAC Summit NYC, Fall 2018

Disclaimer:  
**Not STAC measurements**

# Trading in nanoseconds



“PTP isn't good enough for a nanosecond world ”

- someone in this room (probably)

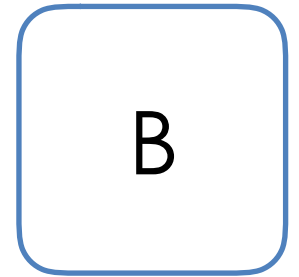
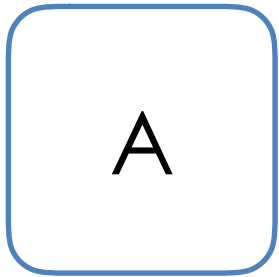


# Part I

## Time Sync theory in 5 mins or less

# 1 Pulse Per Second (1PPS)

# 1 Pulse Per Second (1PPS)



# 1 Pulse Per Second (1PPS)

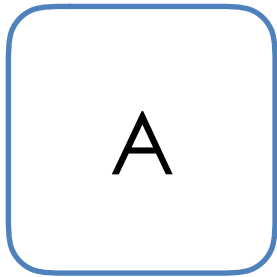
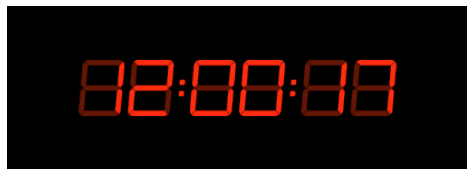




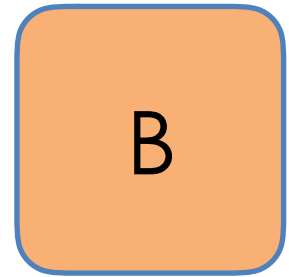
# 1 Pulse Per Second (1PPS)



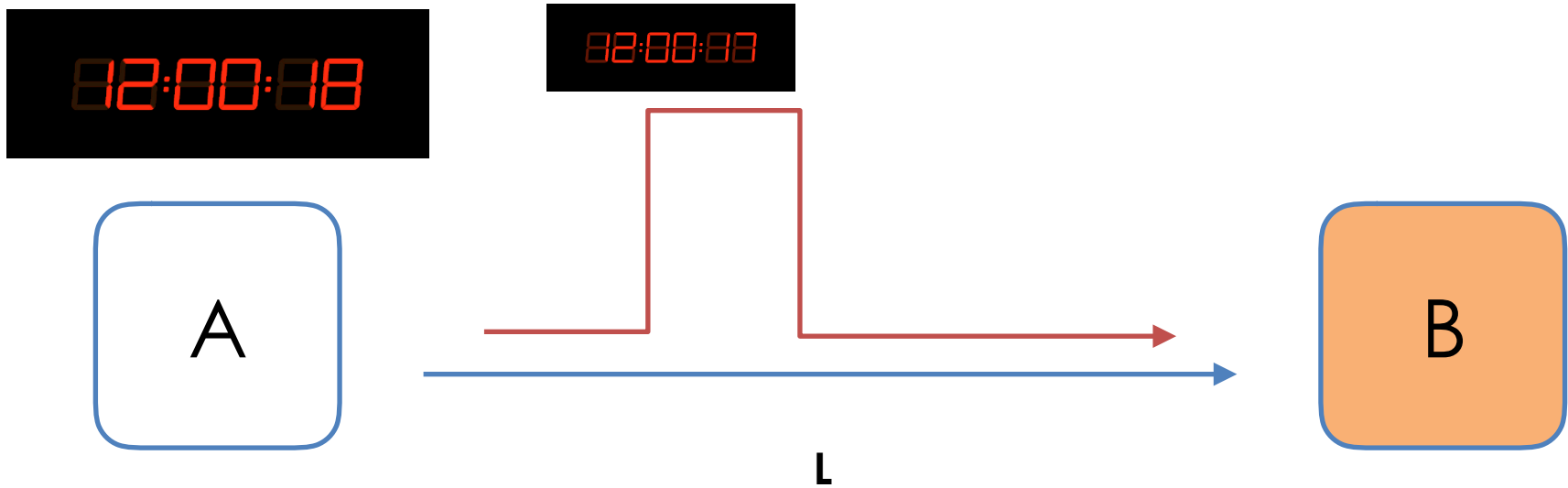
# 1 Pulse Per Second (1PPS)



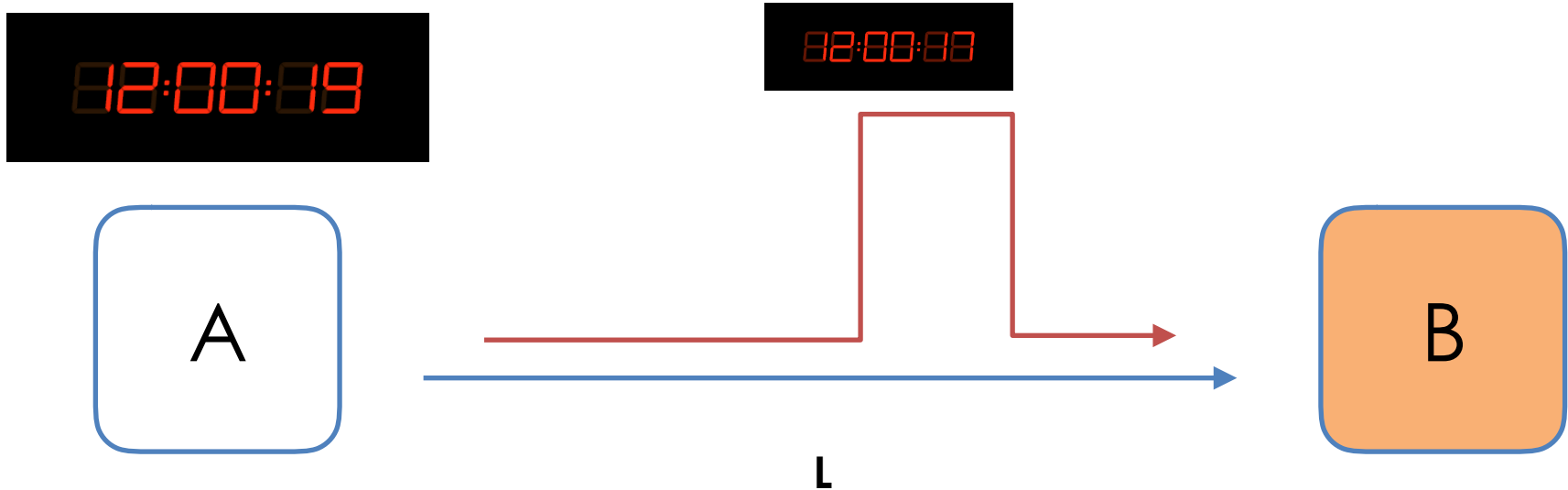
L



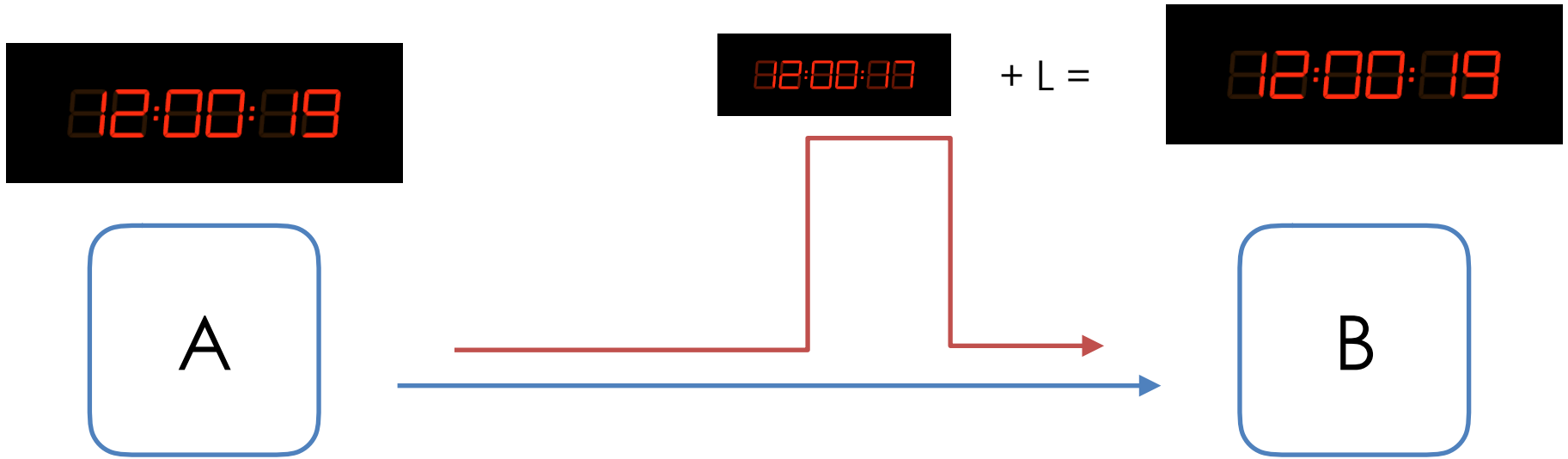
# 1 Pulse Per Second (1PPS)



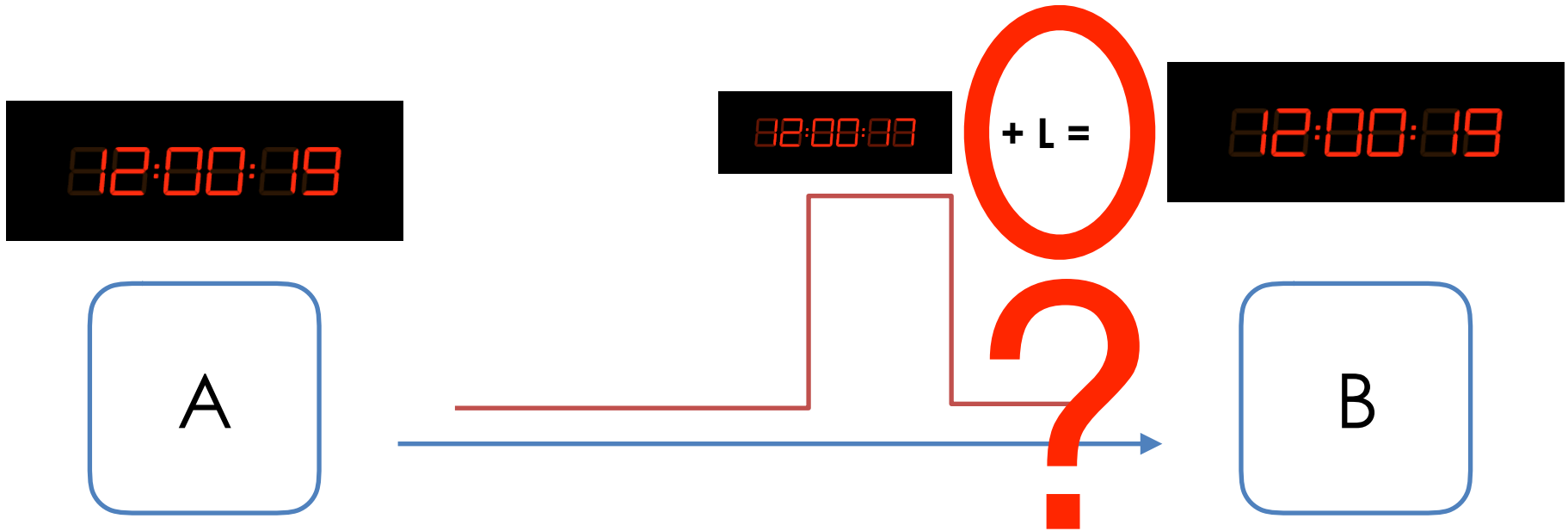
# 1 Pulse Per Second (1PPS)



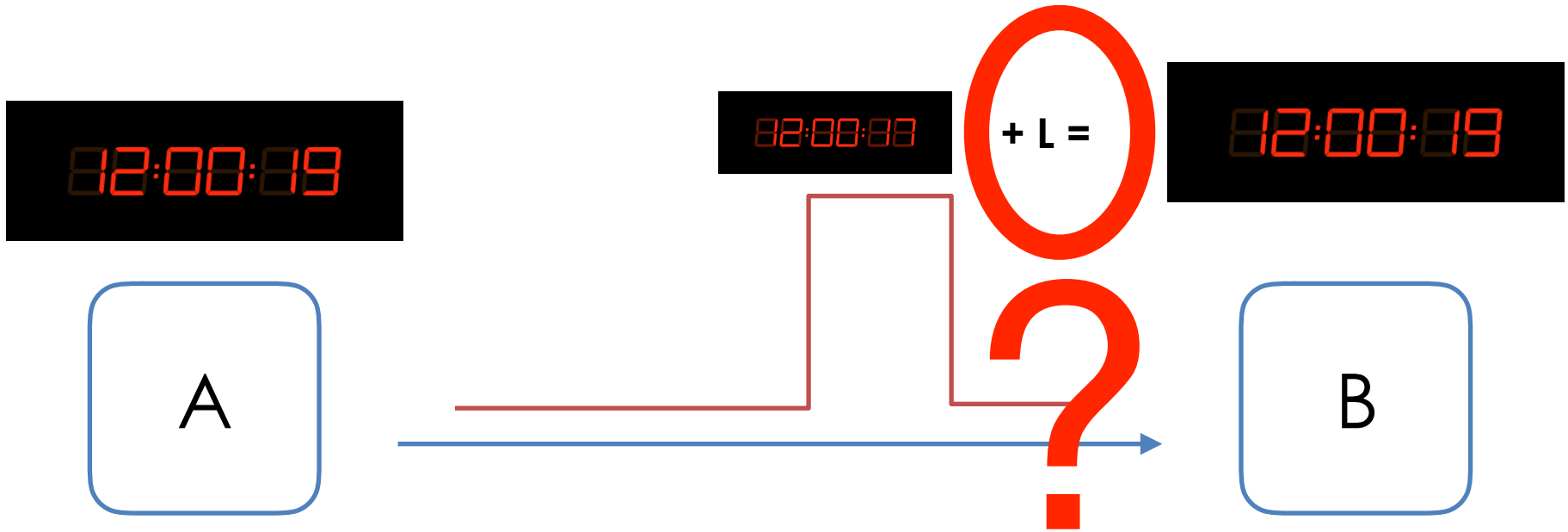
# 1 Pulse Per Second (1PPS)



# Problem 1 - what is $L_t$ ?

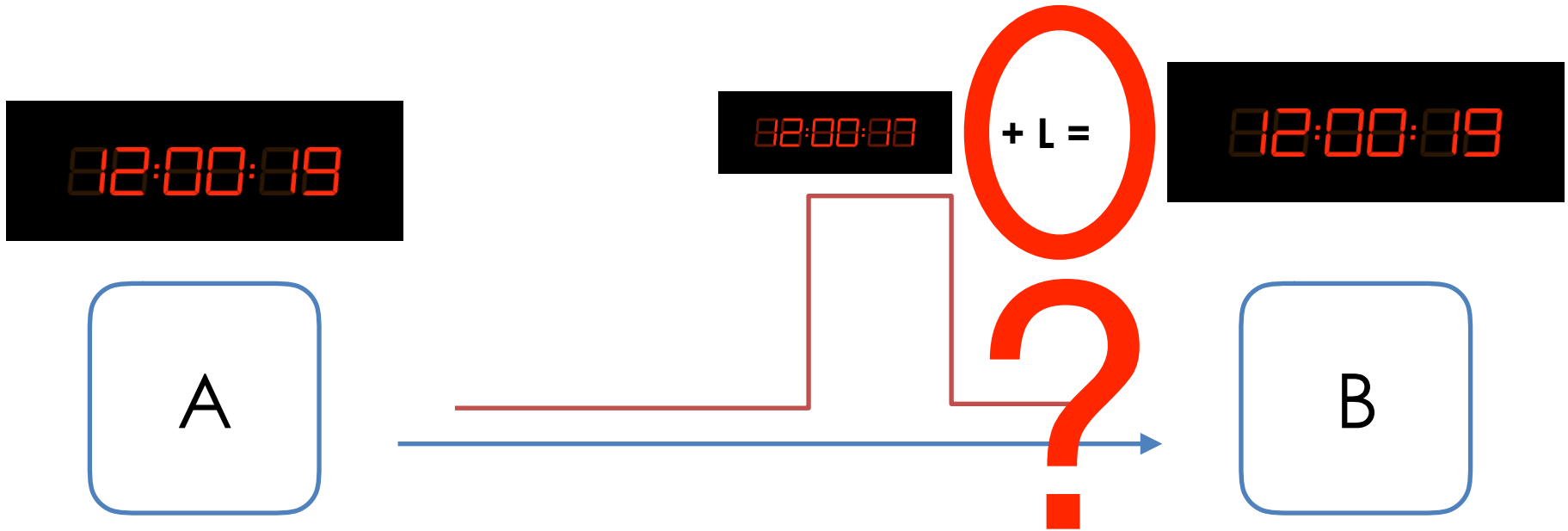


# Problem 1 - what is $L_t$ ?



Solution: assume  $V_{\text{prop}} = 5\text{ns/m}$

# Problem 1 - what is $L_t$ ?

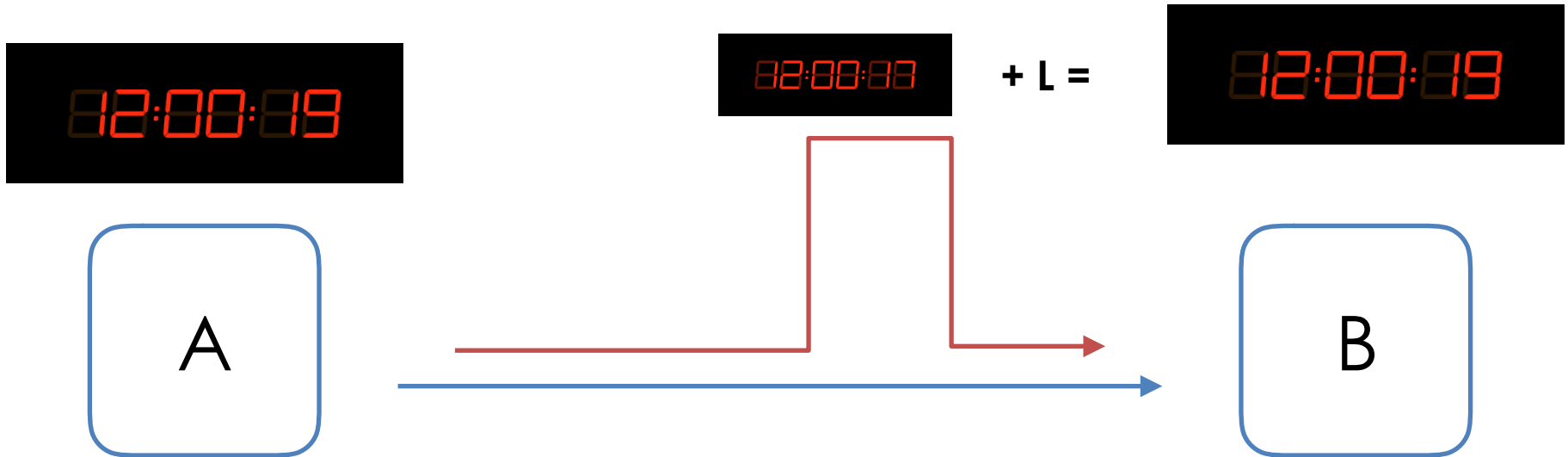


Solution: assume  $V_{\text{prop}} = 5\text{ns/m}$

$$L_t = V_{\text{prop}} \times L$$



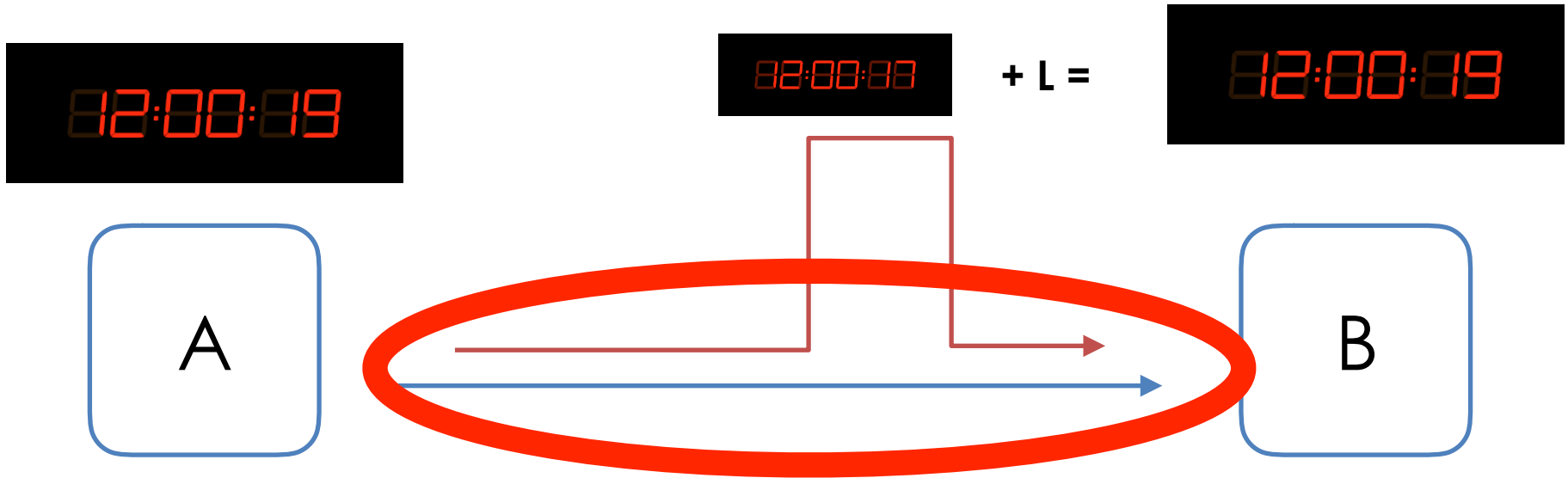
# Problem 1 - what is $L_t$ ?



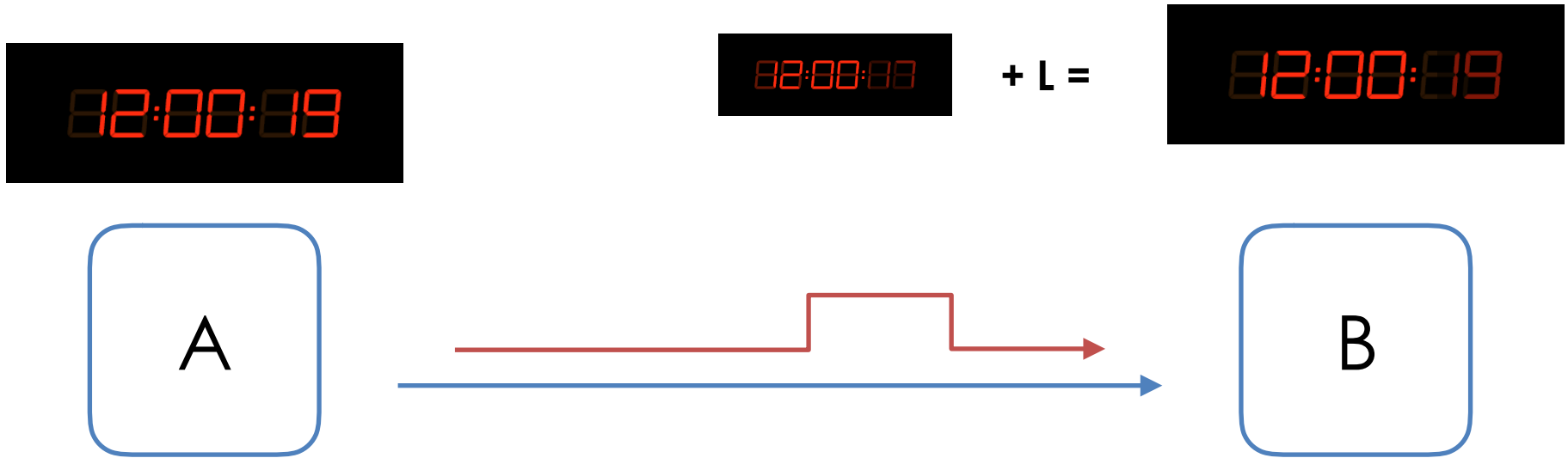
Solution: assume  $V_{\text{prop}} = 5\text{ns/m}$

$$L_t = V_{\text{prop}} \times L$$

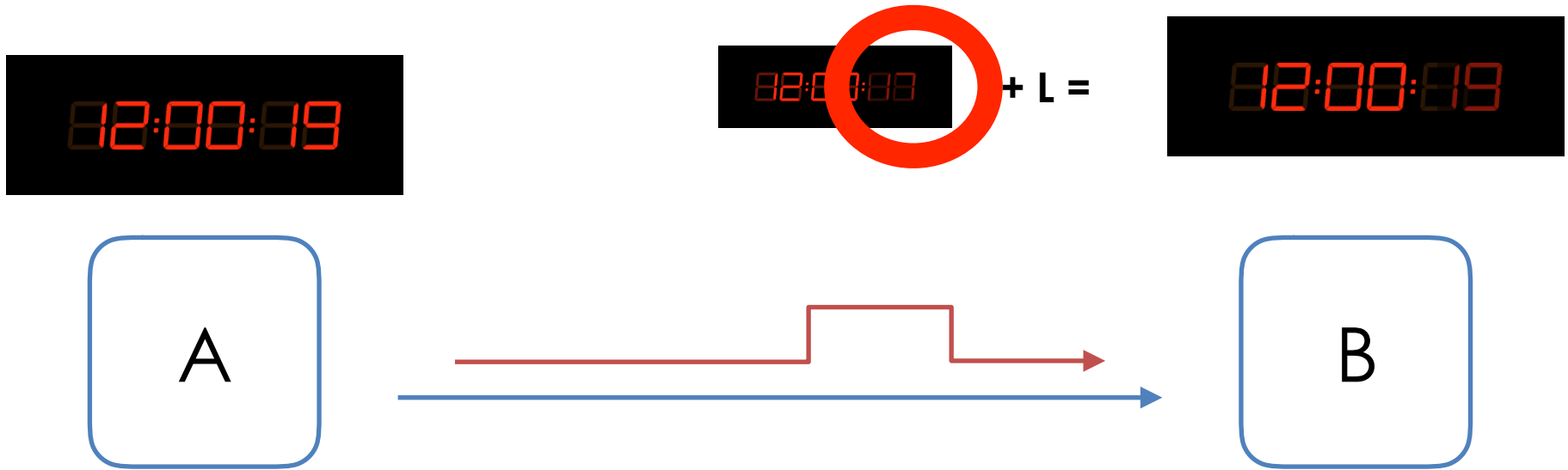
# Problem 2 - what if D is long?



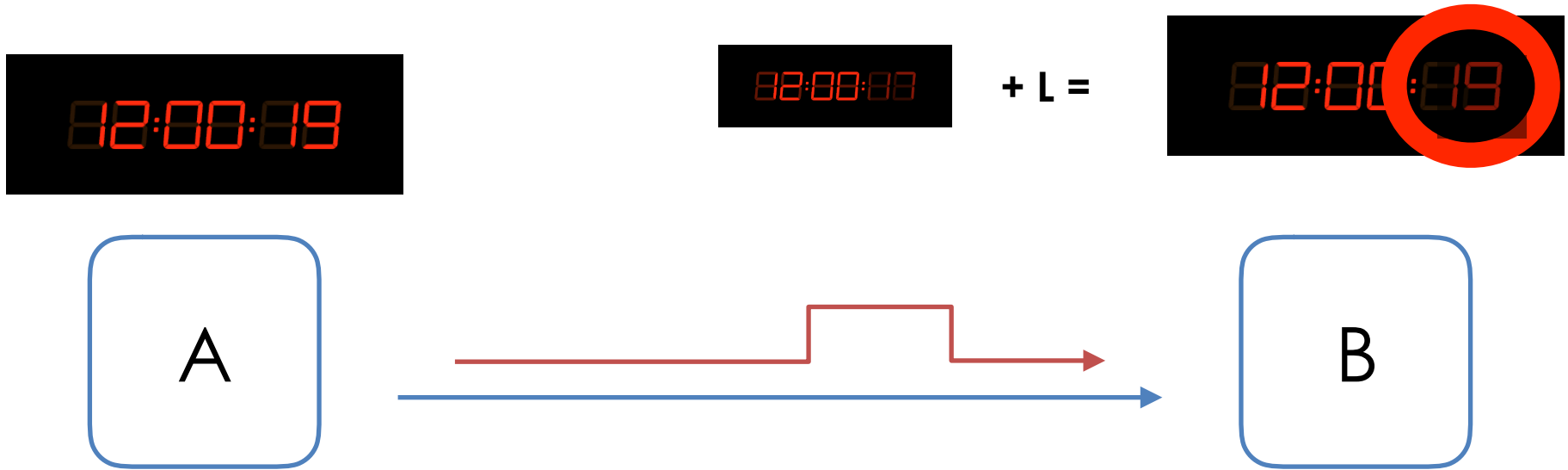
# Problem 2 - what if D is long?



# Problem 2 - what if D is long?



# Problem 2 - what if D is long?

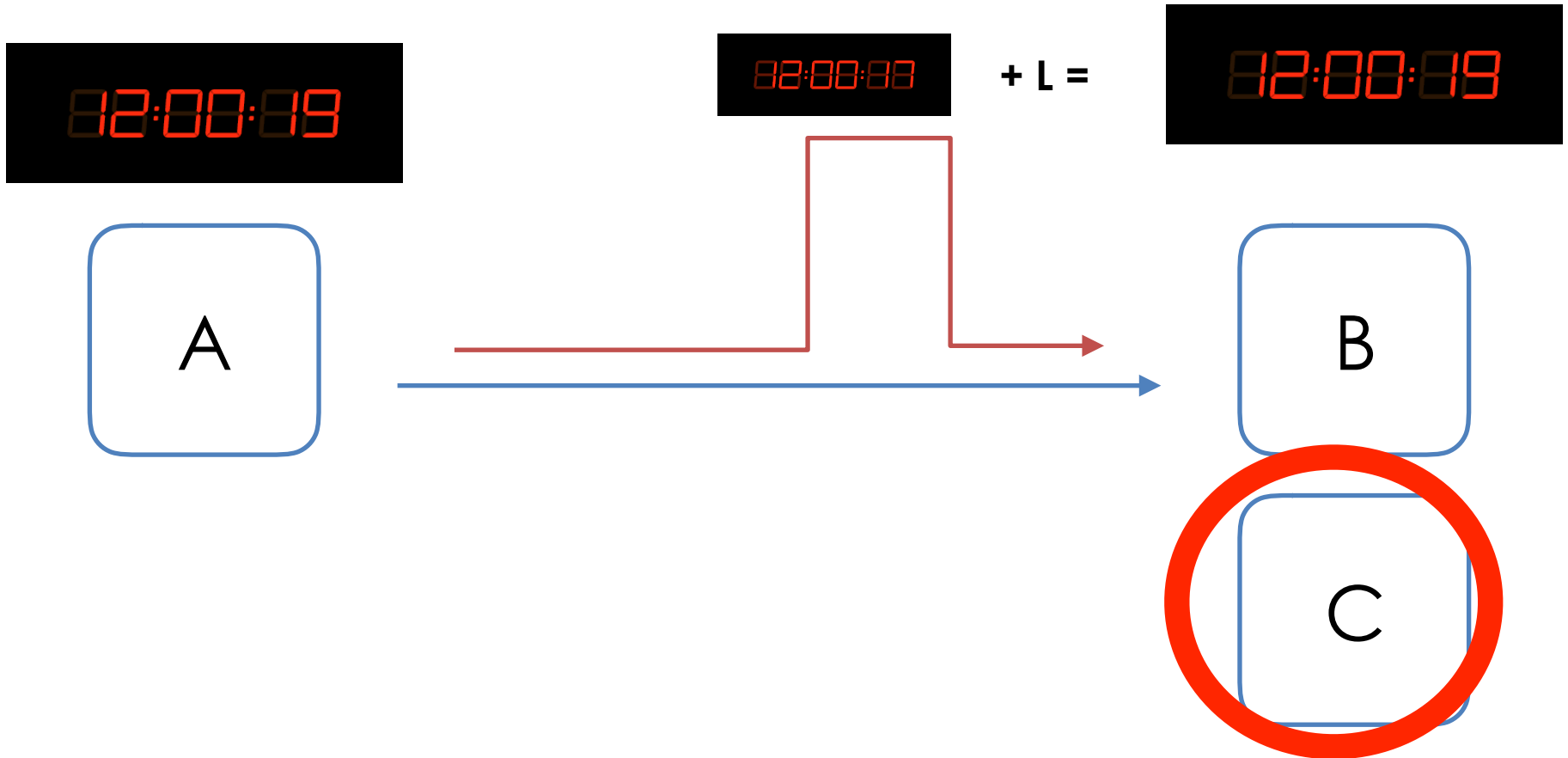


# Problem 2 - what if D is long?

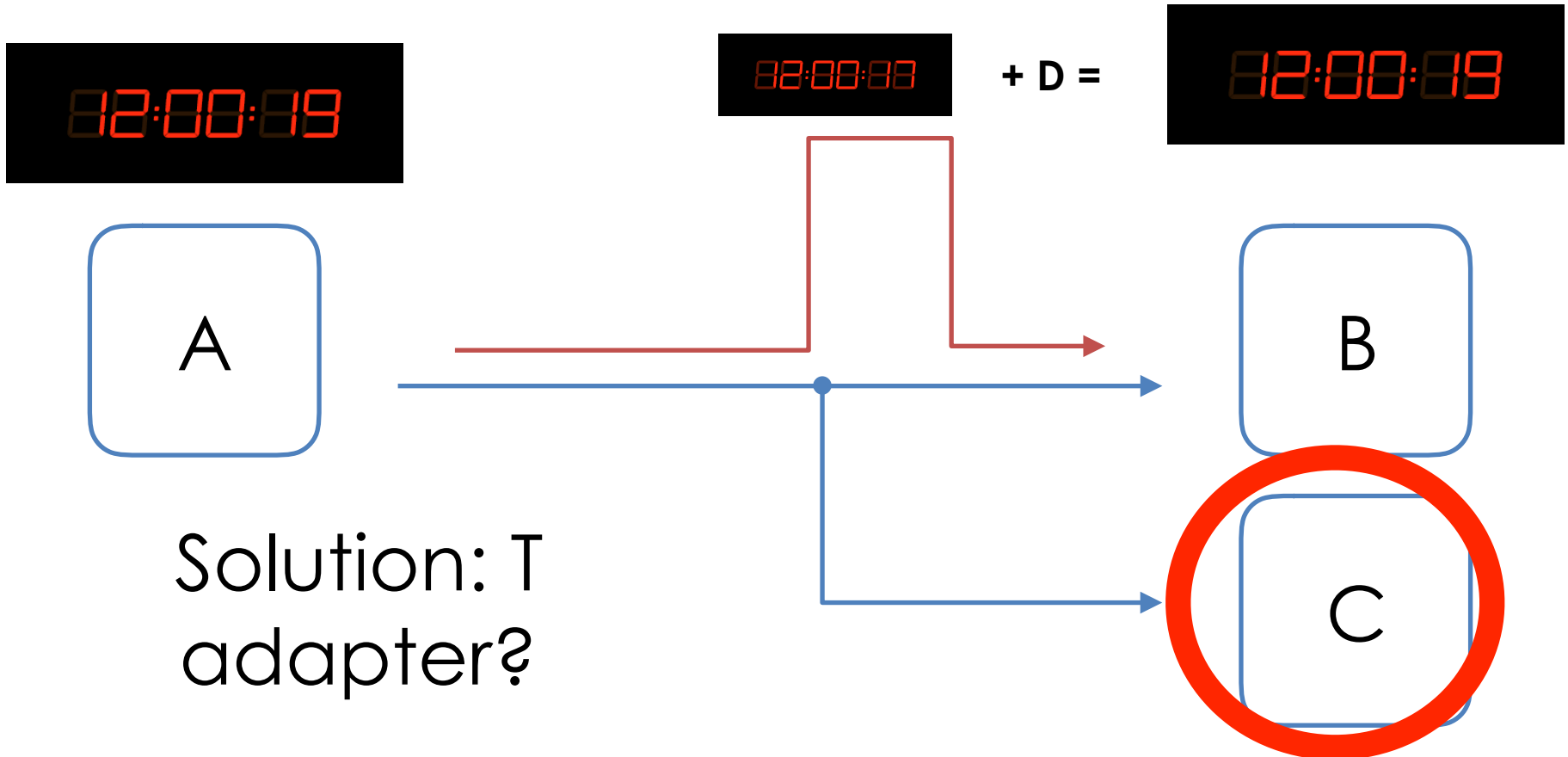


Solution: Amplifier ?

# Problem 3 - multi-host?



# Problem 3 - multi-host?





# 1 PPS Conclusions:

# 1 PPS Conclusions:

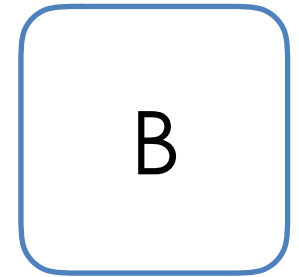
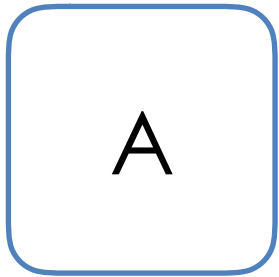
- Doesn't scale well

# 1 PPS Conclusions:

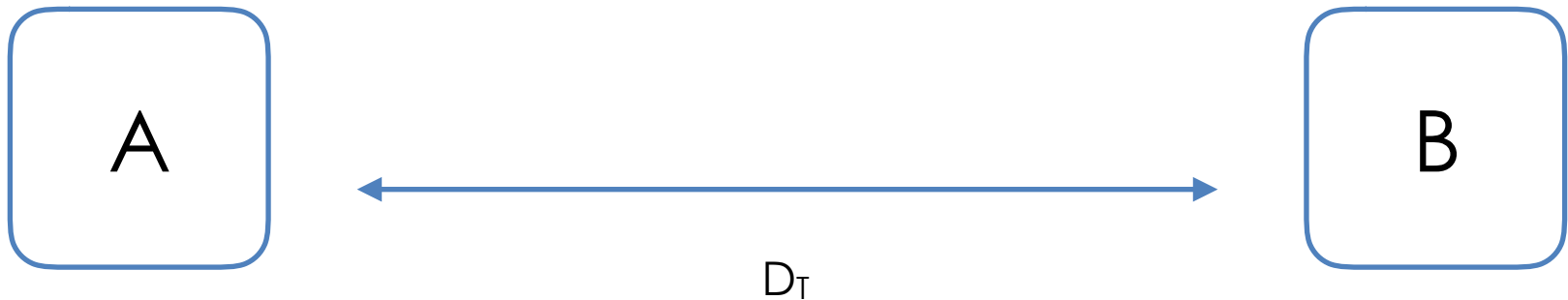
- Doesn't scale well
- Unsuitable distrib. time sync

# What about PTP?

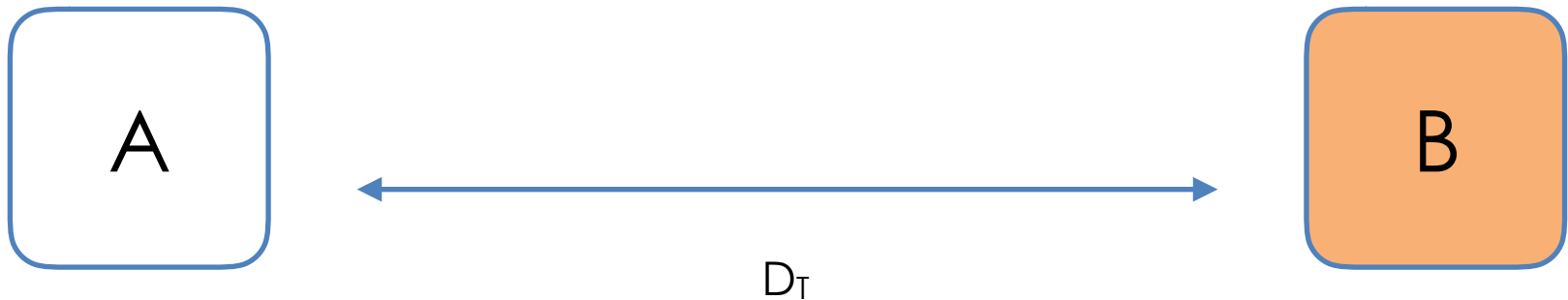
# Fundamentals of PTP



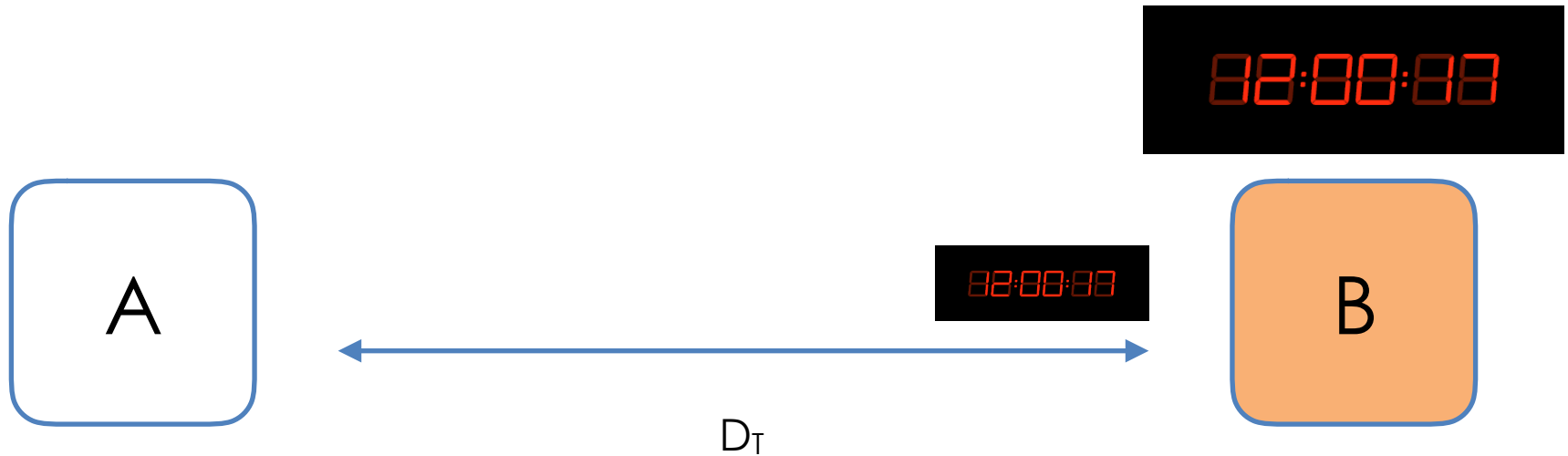
# Fundamentals of PTP



# Fundamentals of PTP

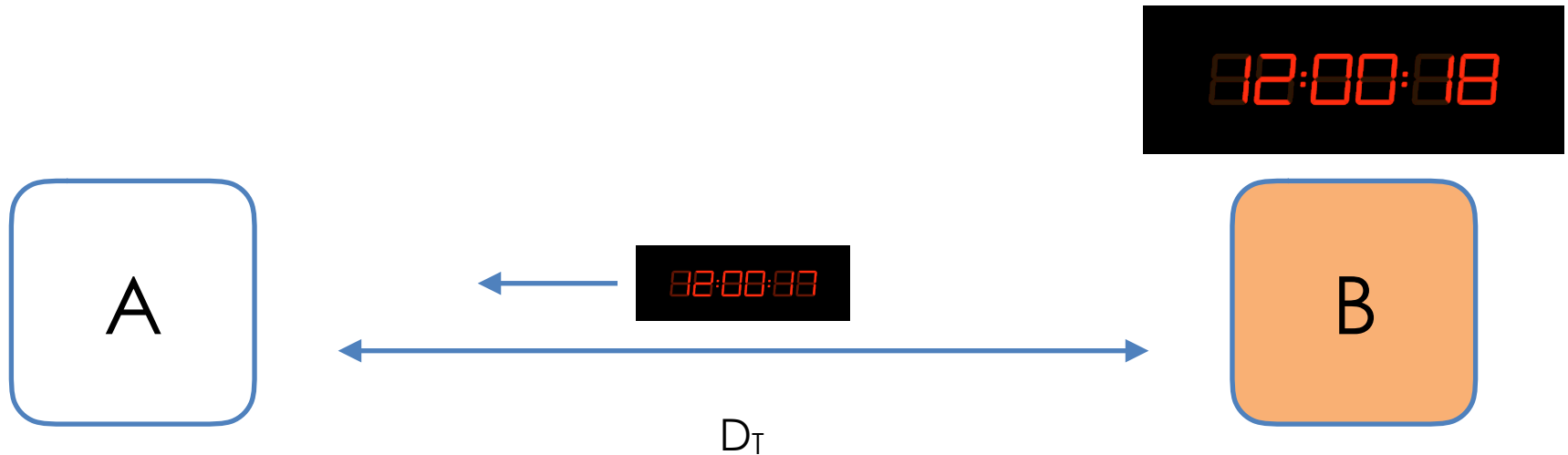


# Fundamentals of PTP

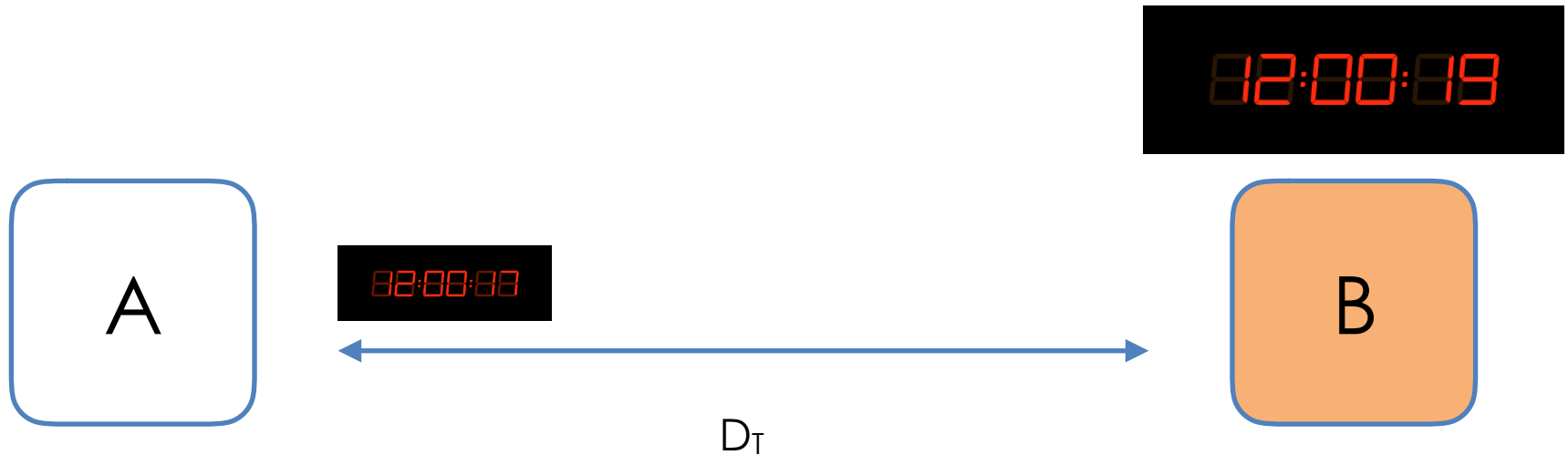




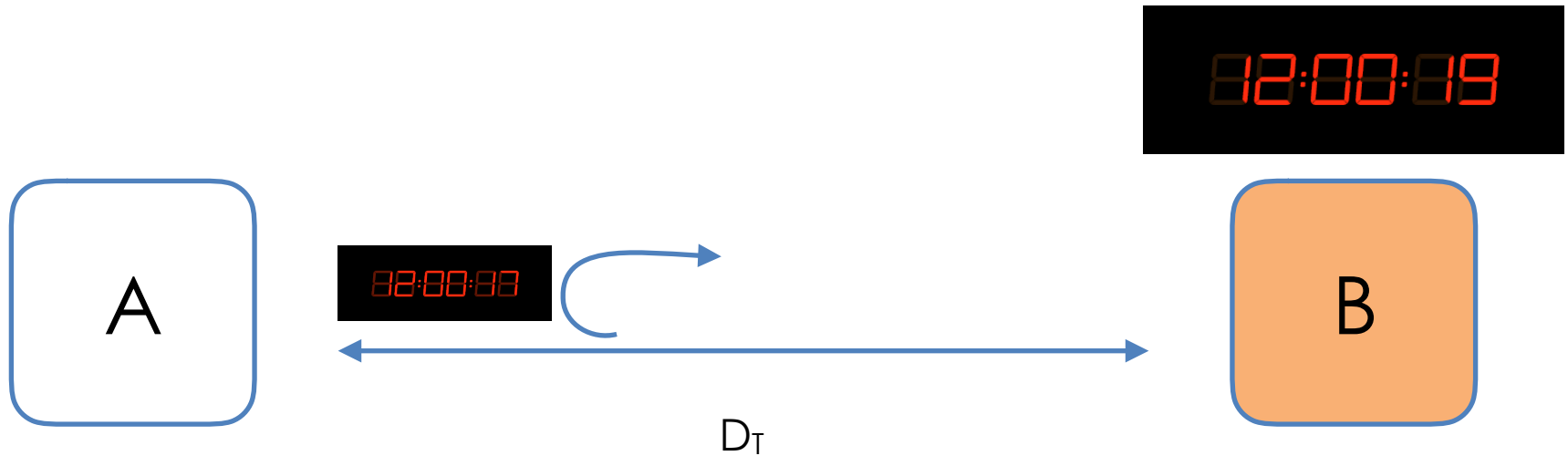
# Fundamentals of PTP



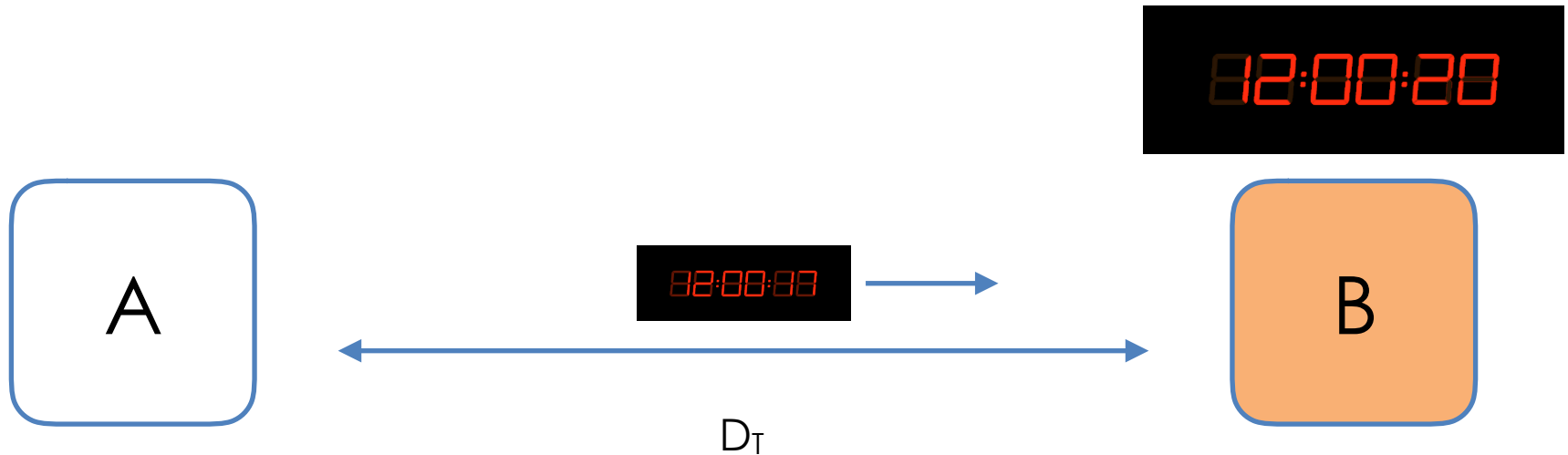
# Fundamentals of PTP



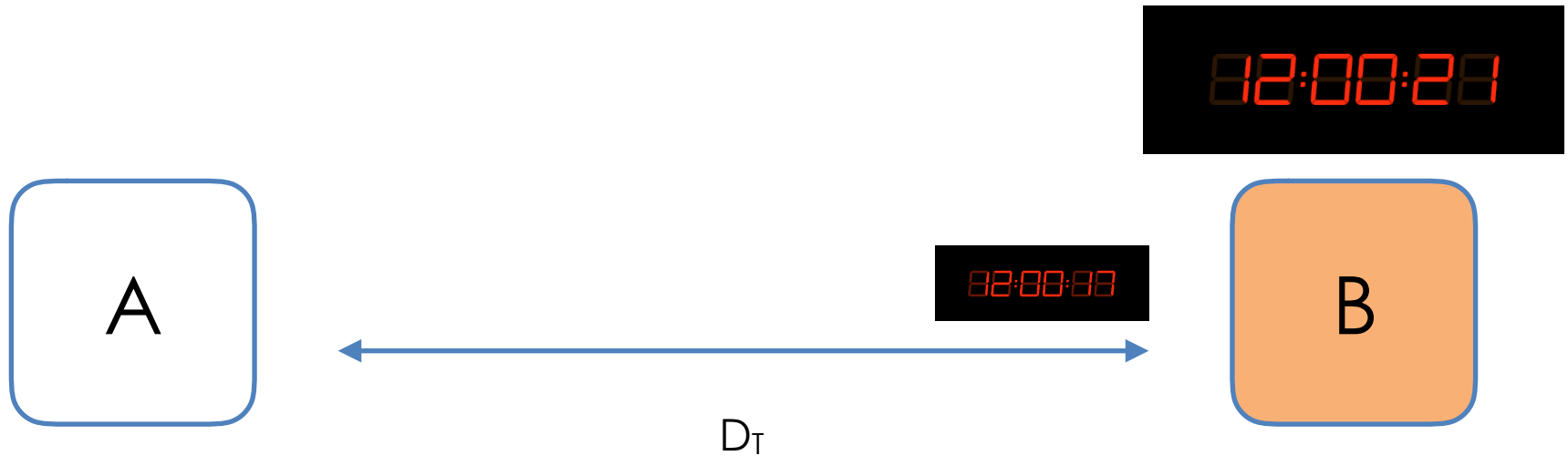
# Fundamentals of PTP



# Fundamentals of PTP

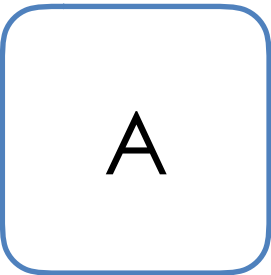


# Fundamentals of PTP

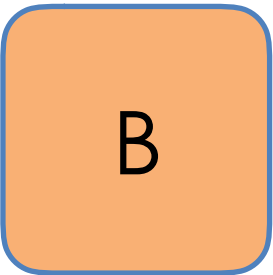


# Fundamentals of PTP

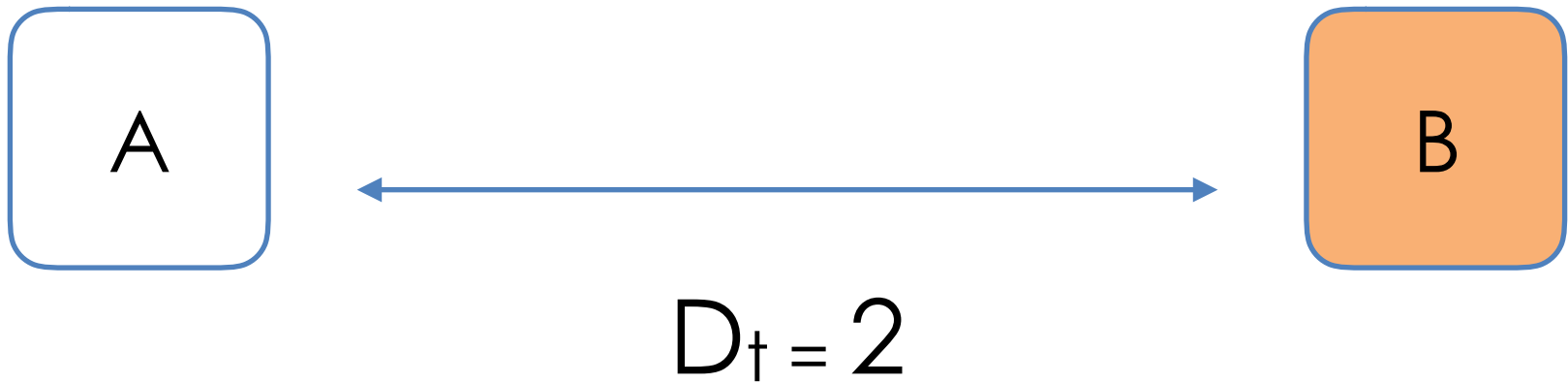

$$12:00:21 - 00:00:00 = 2x D_t$$



$D_t$



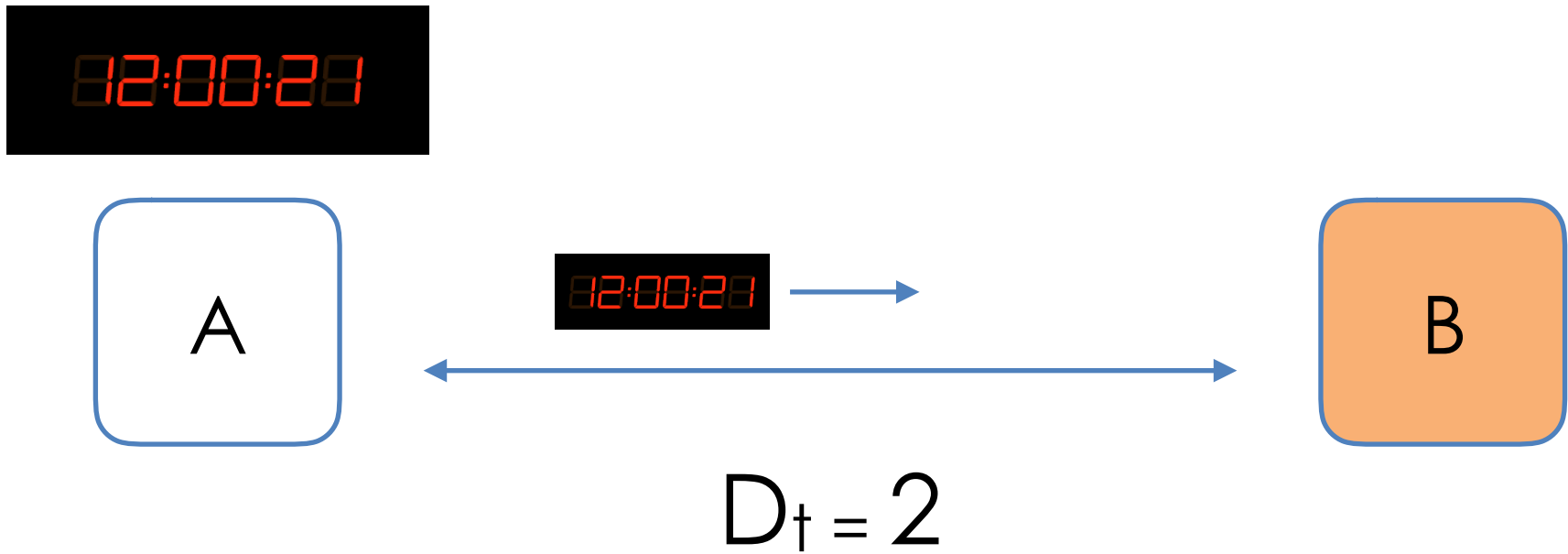
# Fundamentals of PTP



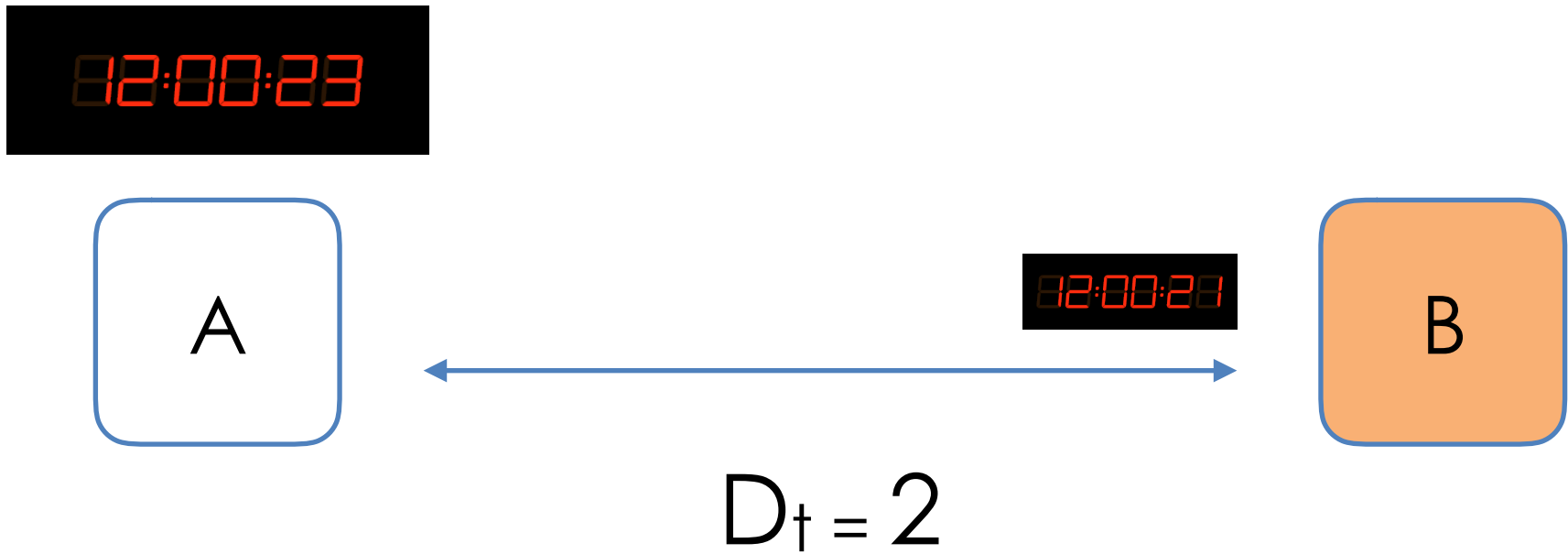
# PTP solves PPS estimation problem



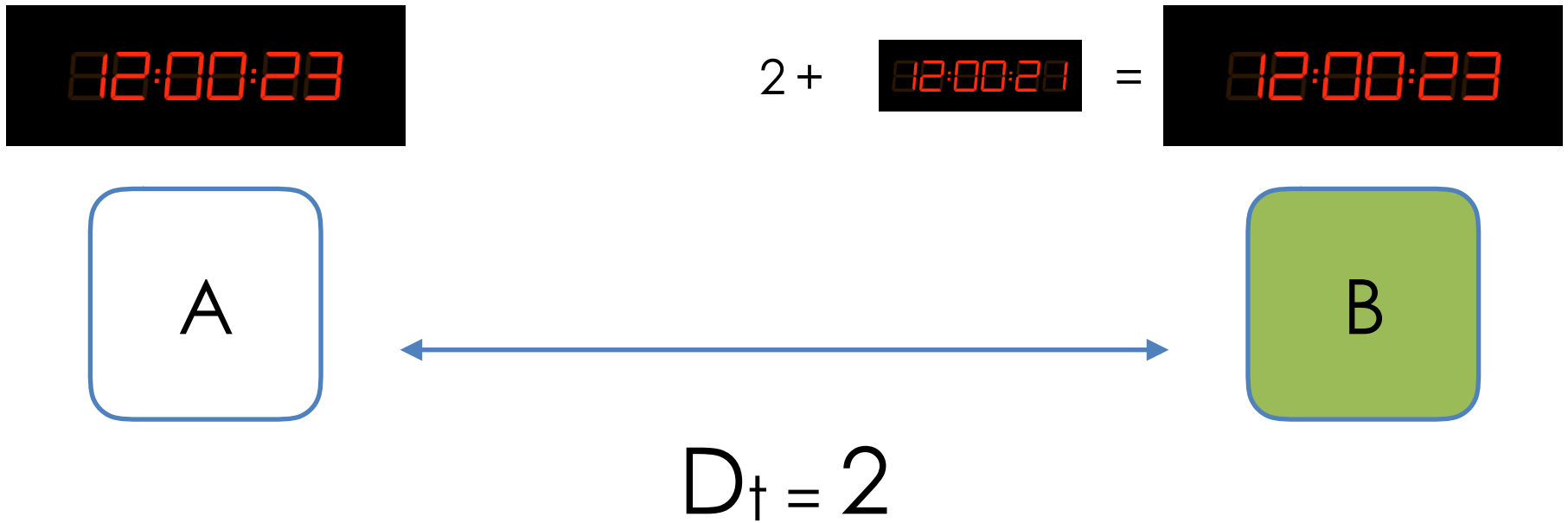
# Fundamentals of PTP



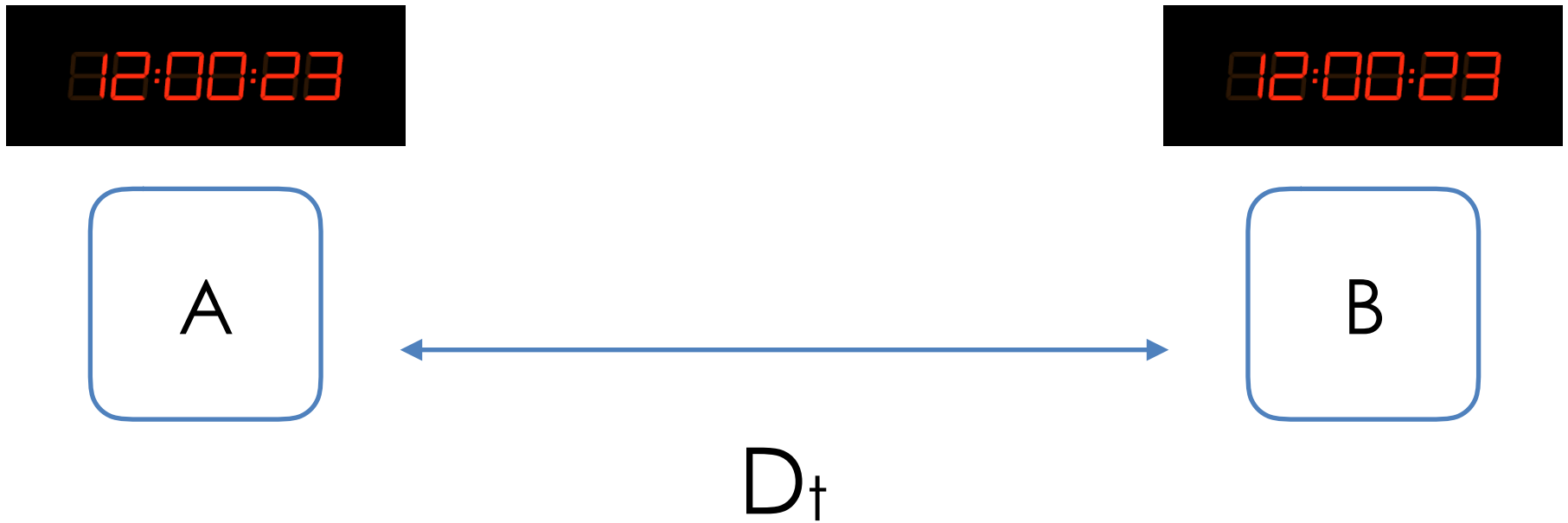
# Fundamentals of PTP



# Fundamentals of PTP



# What if D is long?

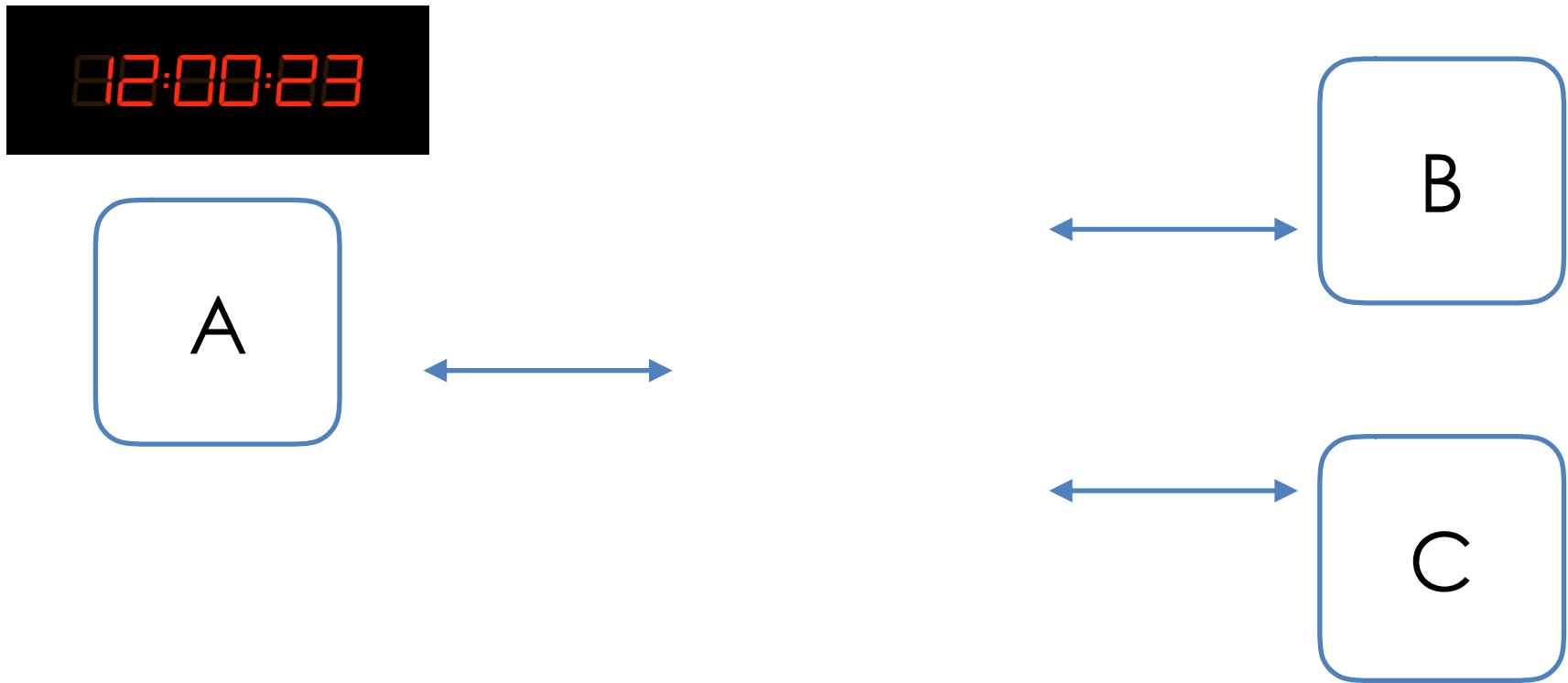


# What if D is long?

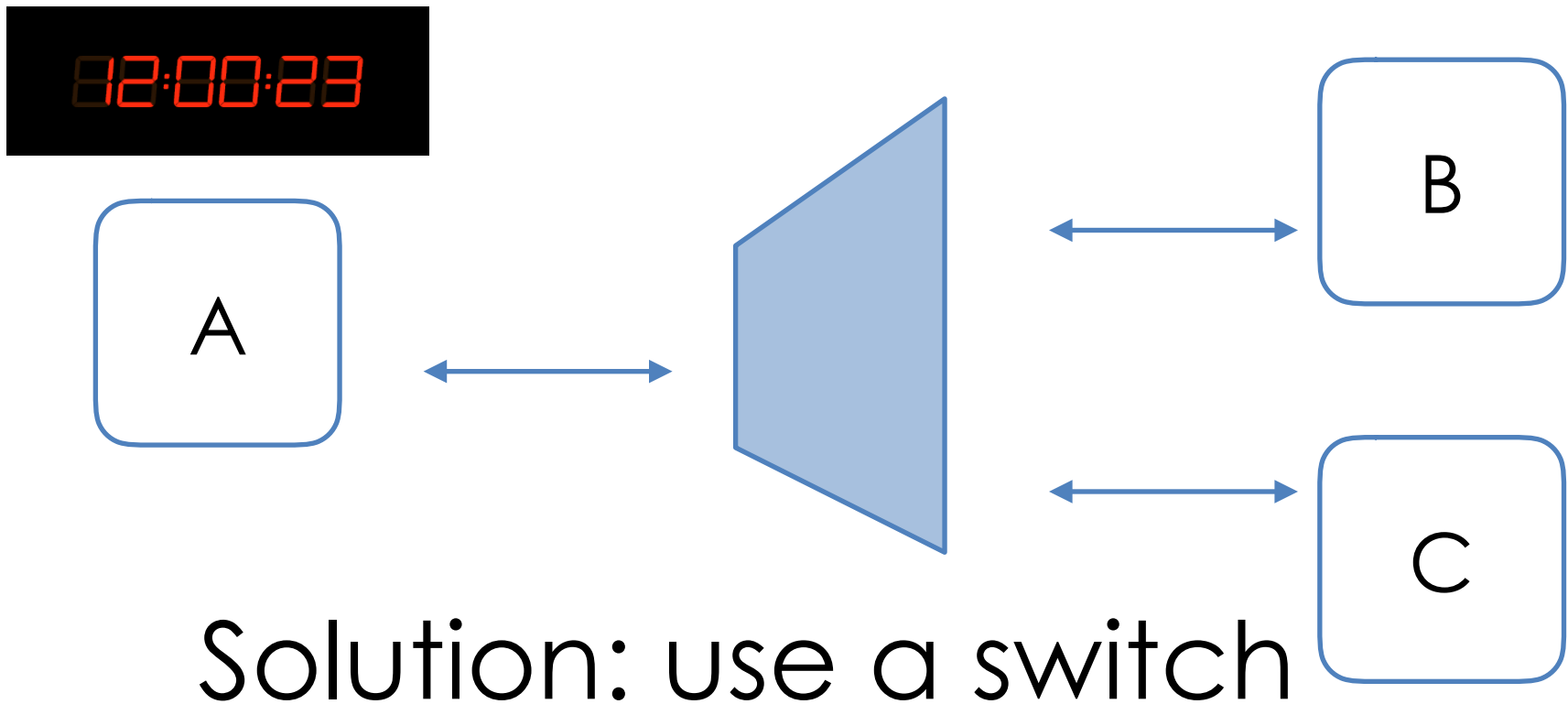


## Solution: LR Transceivers

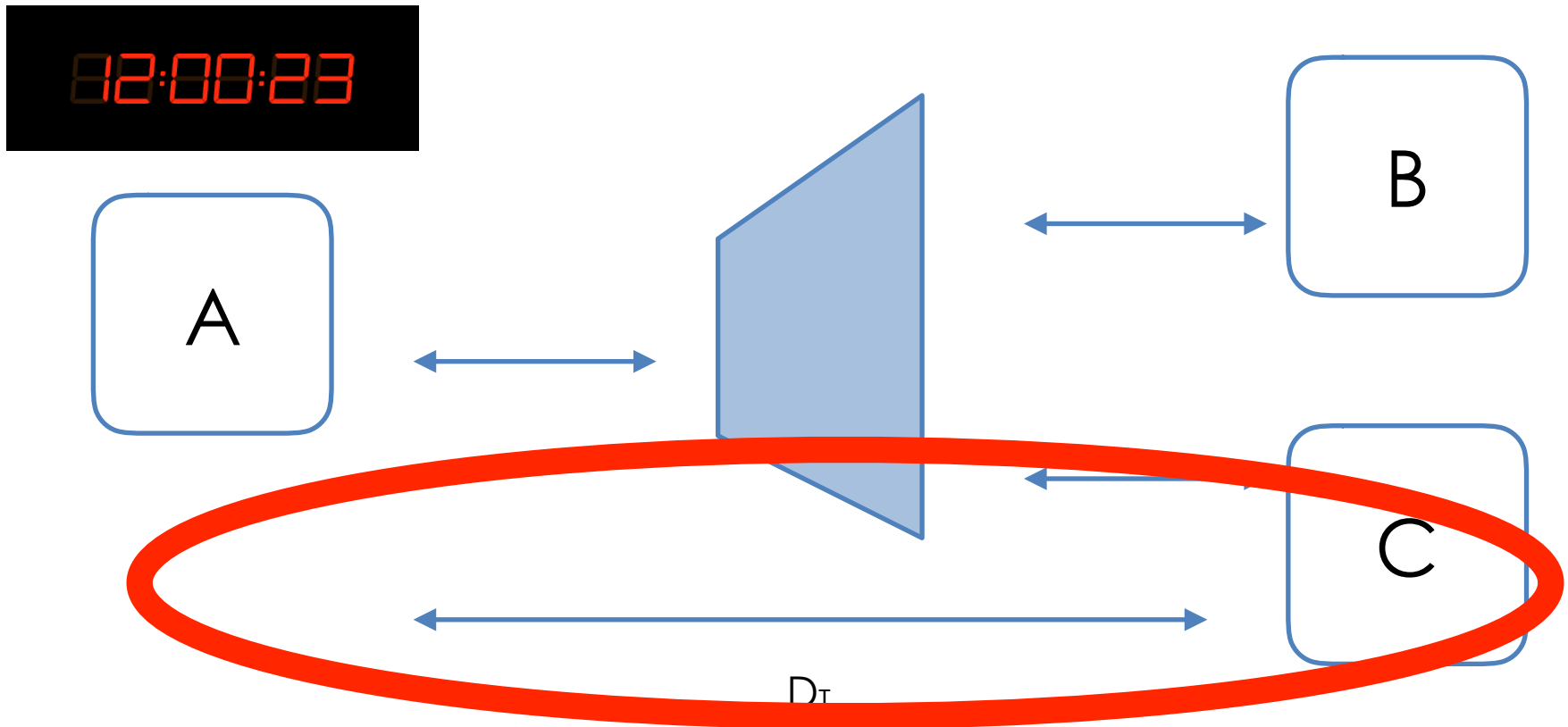
# PTP - Multi-host?



# PTP Problem - multi-host?

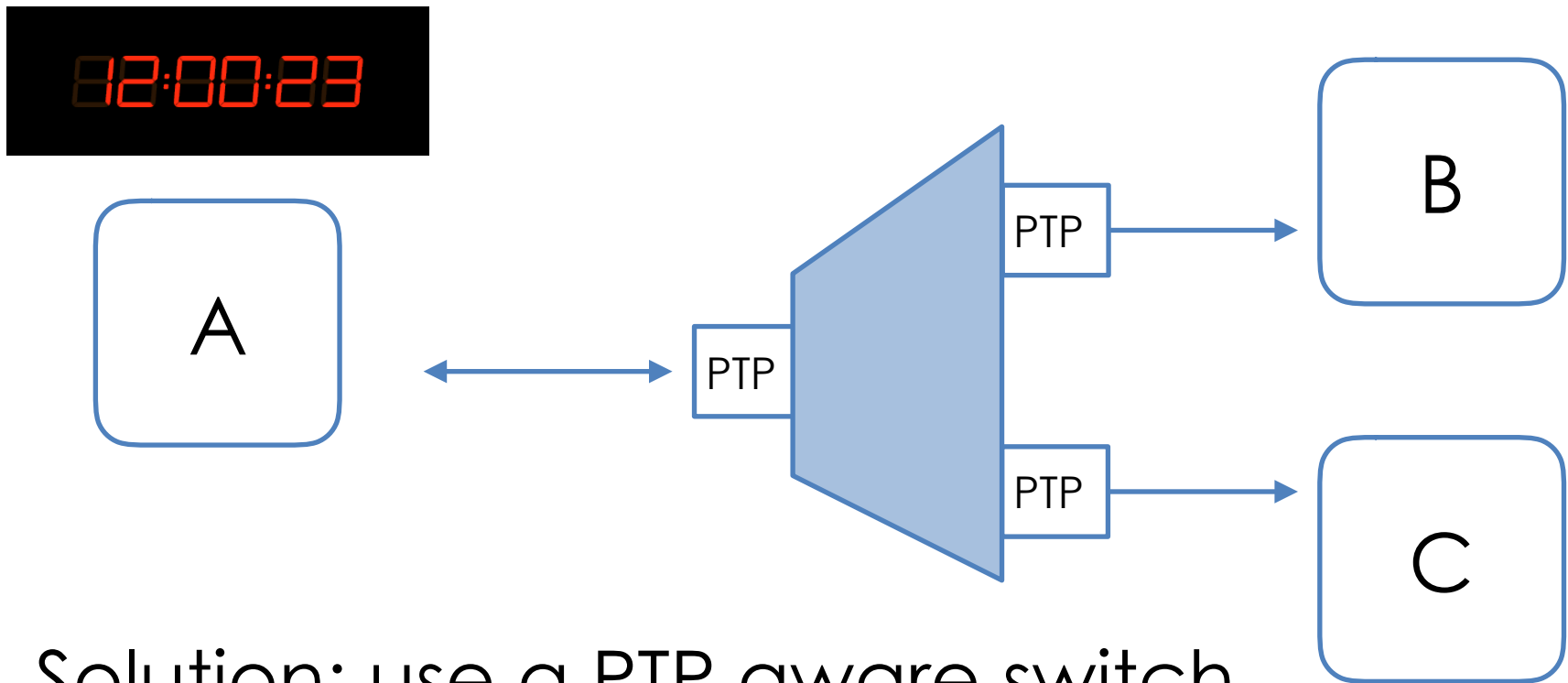


# PTP Problem - switch queuing?





# PTP Problem - switch queuing?



Solution: use a PTP aware switch

# PTP on paper:

# PTP on paper:

- Better quality sync ( $D_t$  est.)

# PTP on paper:

- Better quality sync ( $D_t$  est.)
- Scales longer distance

# PTP on paper:

- Better quality sync ( $D_t$  est.)
- Scales longer distance
- Scales to more hosts



# Part II

## PTP

### in

# practice

# PTP in Practice

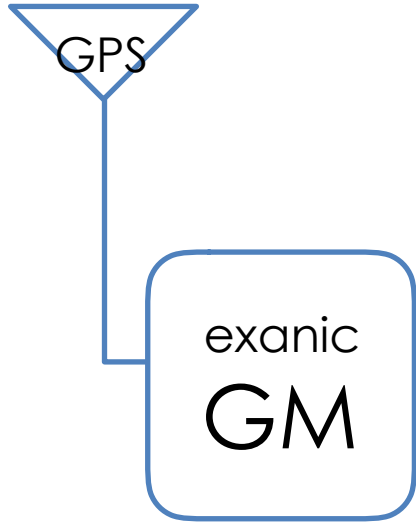


# PTP in Practice

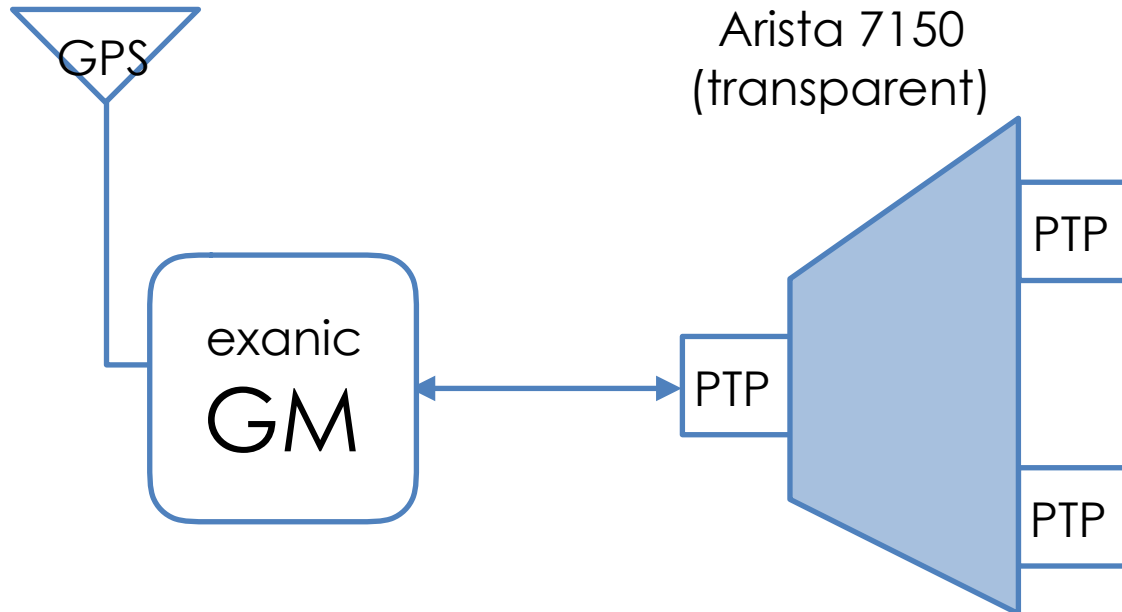




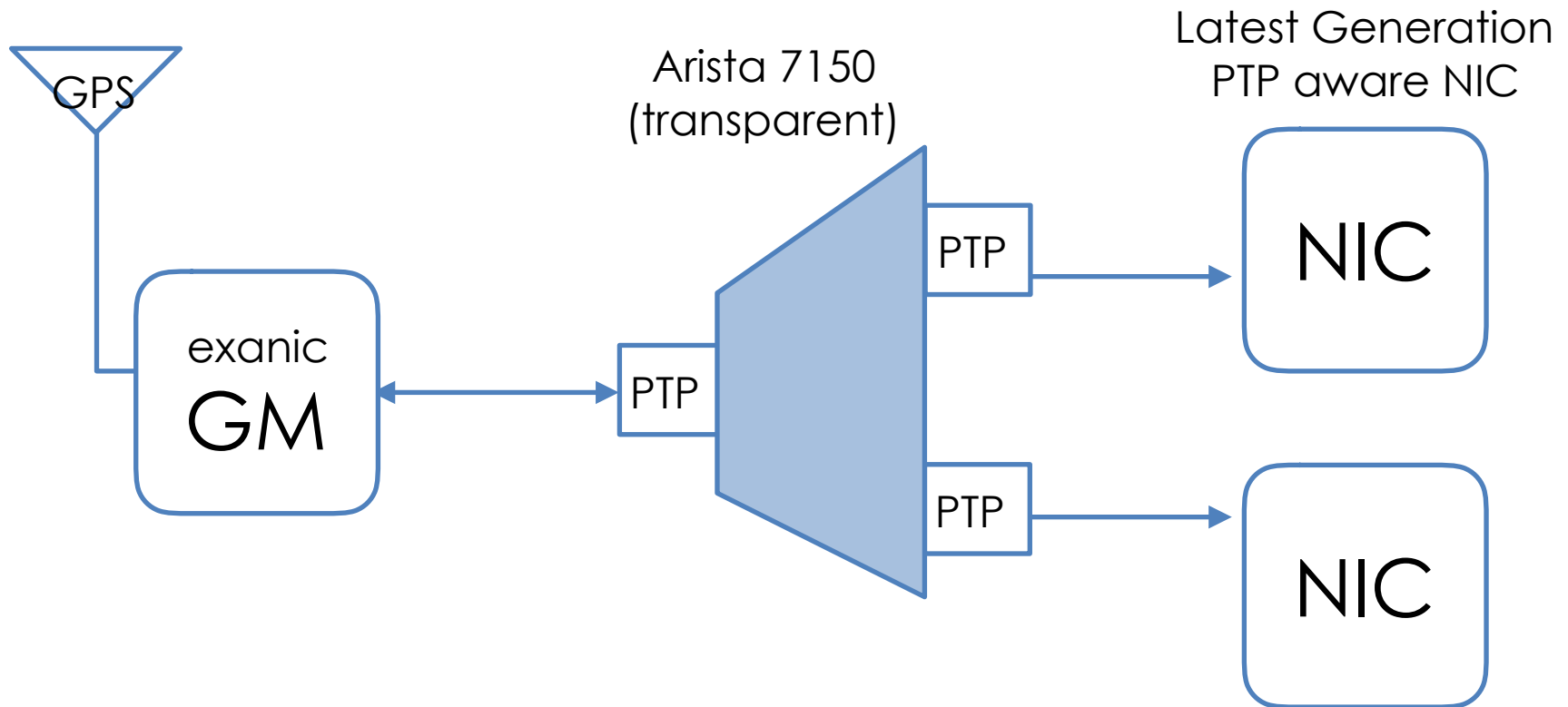
# PTP in Practice



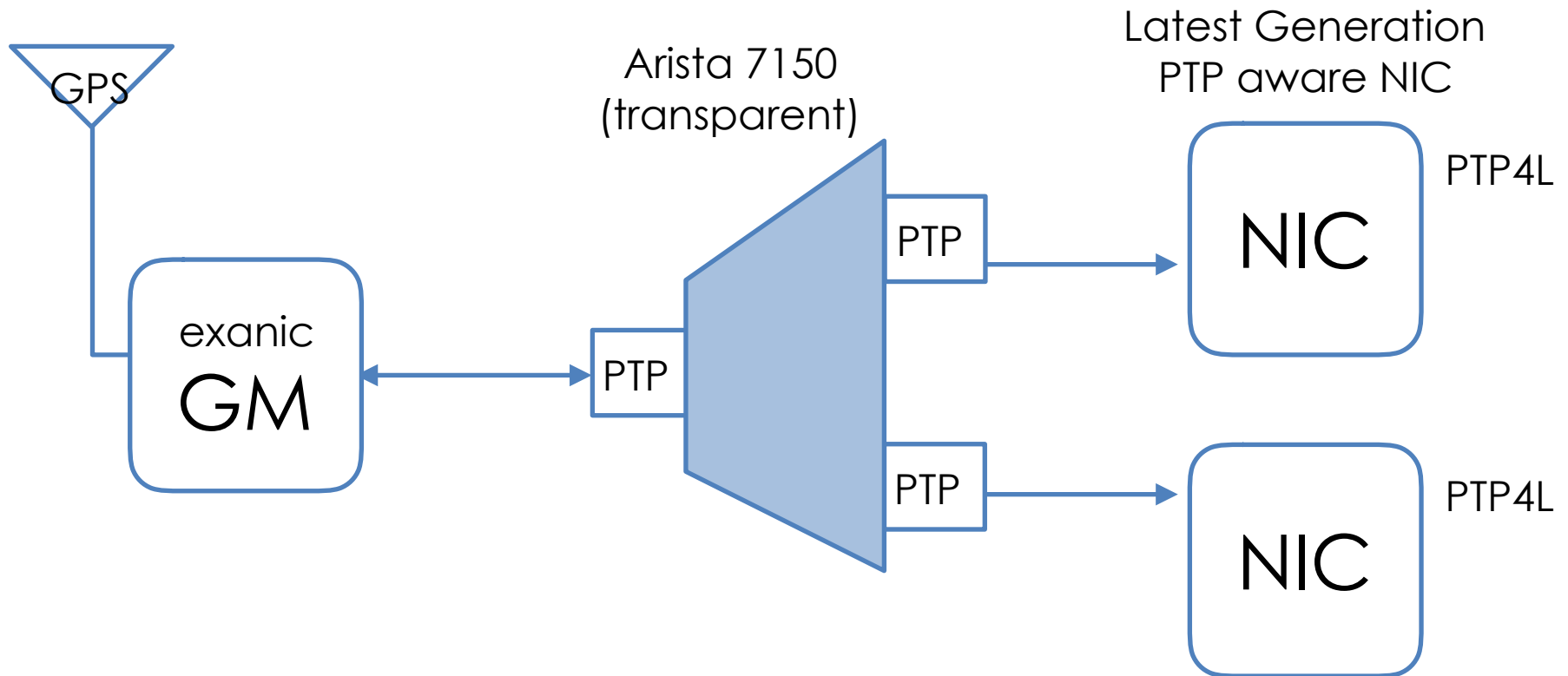
# PTP in Practice



# PTP in Practice



# PTP in Practice



# How do you measure PTP Sync Quality?

# PTP4L Logs

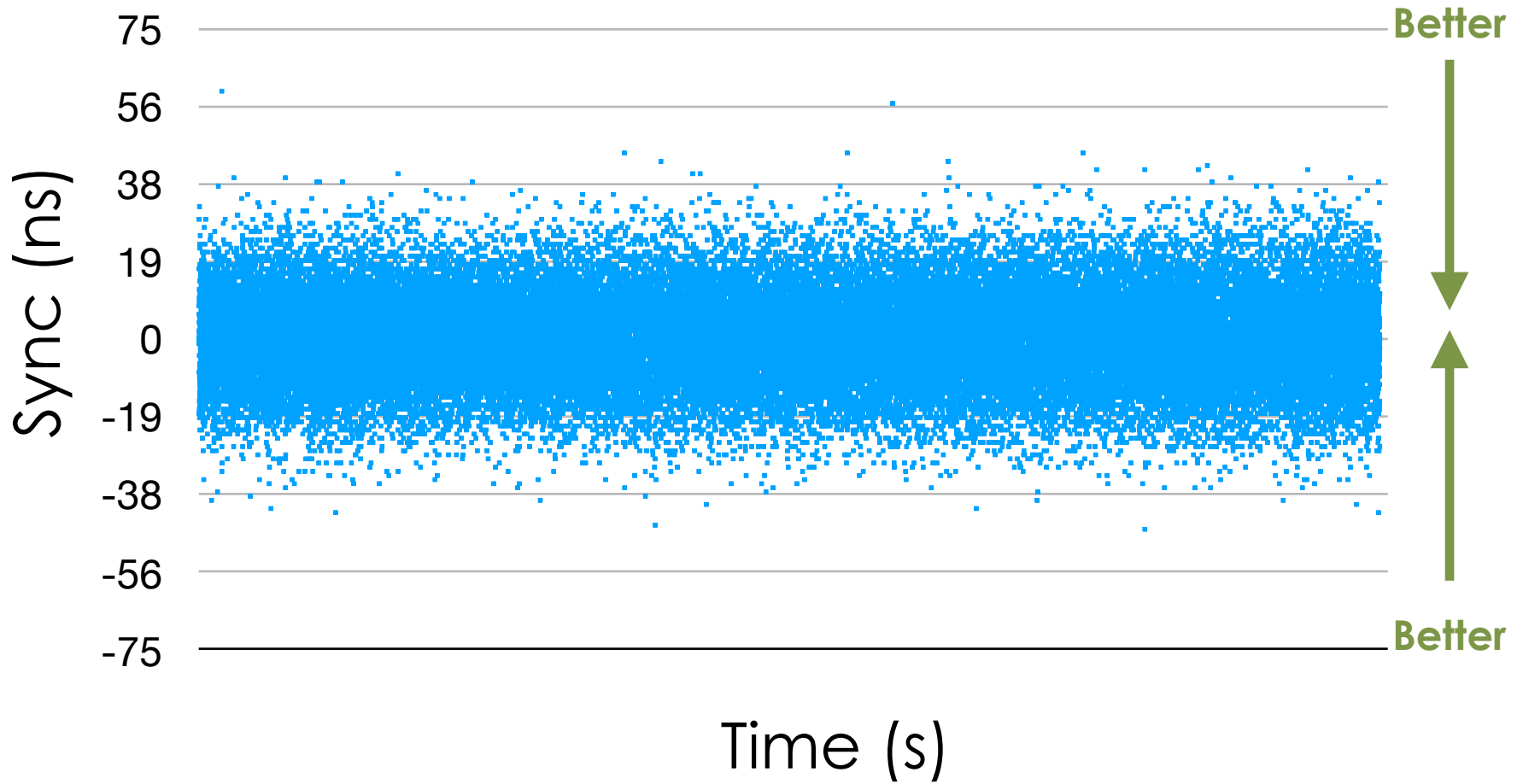
```
ptp4l[336539.407]: selected /dev/ptp2 as PTP clock
ptp4l[336539.408]: port 1: INITIALIZING to LISTENING on INIT_COMPLETE
ptp4l[336539.408]: port 0: INITIALIZING to LISTENING on INIT_COMPLETE
ptp4l[336539.703]: port 1: new foreign master 643f5f.ffffe.01347e-1
ptp4l[336539.780]: port 1: peer detected, switch to P2P
ptp4l[336541.703]: selected best master clock 643f5f.ffffe.01347e
ptp4l[336541.703]: port 1: LISTENING to UNCALIBRATED on RS_SLAVE
ptp4l[336542.703]: master offset      -76 s0 freq      -246 path delay      63
ptp4l[336543.703]: master offset      -75 s2 freq      -245 path delay      57
ptp4l[336543.704]: port 1: UNCALIBRATED to SLAVE on MASTER_CLOCK_SELECTED
ptp4l[336544.703]: master offset      -76 s2 freq      -321 path delay      57
ptp4l[336545.703]: master offset      -18 s2 freq      -286 path delay      57
ptp4l[336546.703]: master offset         6 s2 freq      -267 path delay      57
ptp4l[336547.703]: master offset        17 s2 freq      -254 path delay      57
ptp4l[336548.703]: master offset        29 s2 freq      -237 path delay      57
ptp4l[336549.703]: master offset        17 s2 freq      -241 path delay      57
ptp4l[336550.703]: master offset         -2 s2 freq      -255 path delay      57
ptp4l[336551.703]: master offset         2 s2 freq      -251 path delay      57
ptp4l[336552.703]: master offset       -16 s2 freq      -269 path delay      57
```

# PTP4L Logs

```
ptp4l[336539.407]: selected /dev/ptp2 as PTP clock
ptp4l[336539.408]: port 1: INITIALIZING to LISTENING on INIT_COMPLETE
ptp4l[336539.408]: port 0: INITIALIZING to LISTENING on INIT_COMPLETE
ptp4l[336539.703]: port 1: new foreign master 643f5f.ffffe.01347e-1
ptp4l[336539.780]: port 1: peer detected, switch to P2P
ptp4l[336541.703]: selected best master clock 643f5f.ffffe.01347e
ptp4l[336541.703]: port 1: LISTENING to UNCALIBRATED on RS_SLAVE
ptp4l[336542.703]: master offset      -76 s0 freq      -246 path delay      63
ptp4l[336543.703]: master offset      -75 s2 freq      -245 path delay      57
ptp4l[336543.704]: port 1: UNCALIBRATED to SLAVE on MASTER_CLOCK_SELECTED
ptp4l[336544.703]: master offset      -76 s2 freq      -321 path delay      57
ptp4l[336545.703]: master offset      -18 s2 freq      -286 path delay      57
ptp4l[336546.703]: master offset         6 s2 freq      -267 path delay      57
ptp4l[336547.703]: master offset        17 s2 freq      -254 path delay      57
ptp4l[336548.703]: master offset        29 s2 freq      -237 path delay      57
ptp4l[336549.703]: master offset        17 s2 freq      -241 path delay      57
ptp4l[336550.703]: master offset         -2 s2 freq      -255 path delay      57
ptp4l[336551.703]: master offset         2 s2 freq      -251 path delay      57
ptp4l[336552.703]: master offset       -16 s2 freq      -269 path delay      57
```

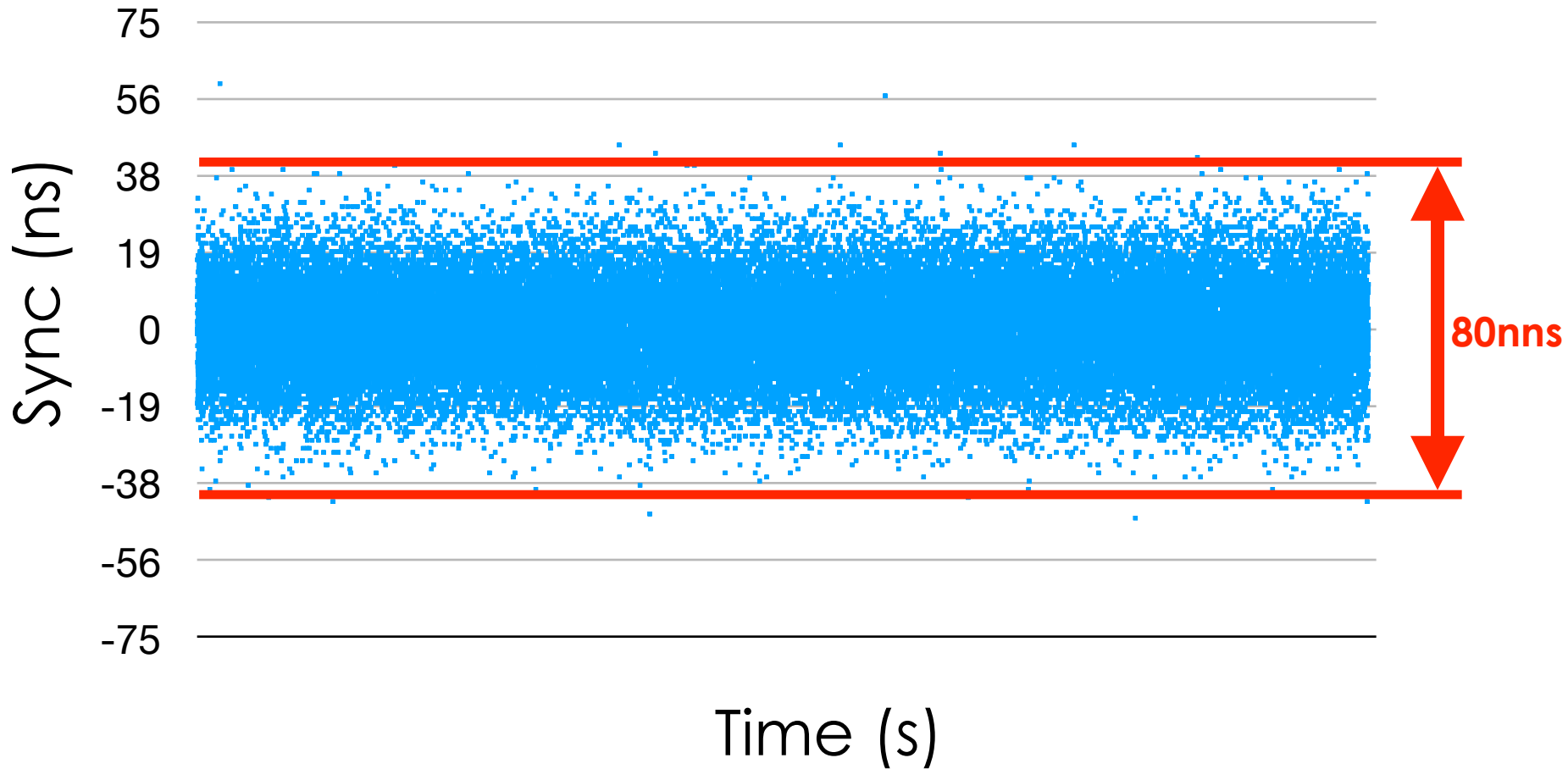
**Perceived Sync**

# PTP4L Perceived Sync

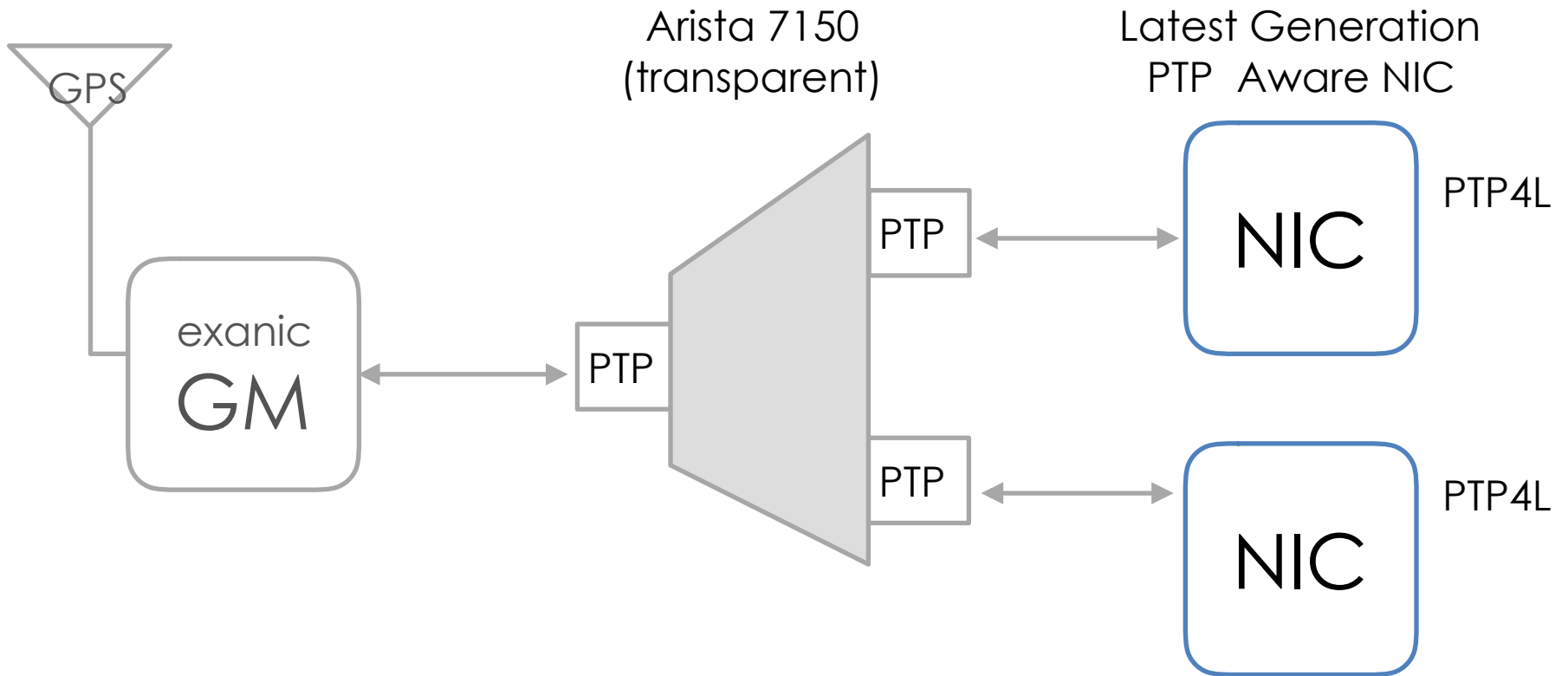




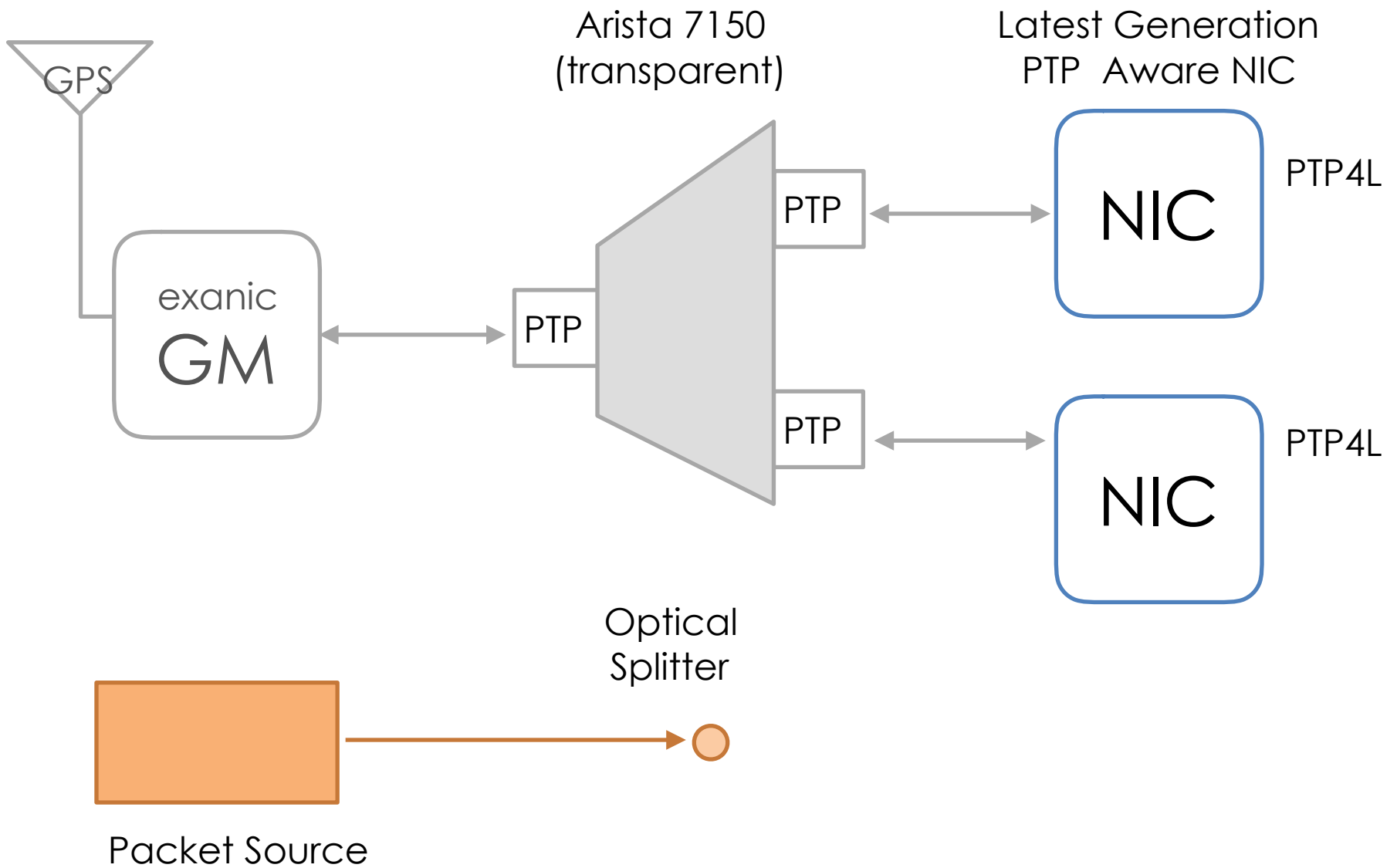
# PTP4L Perceived Sync

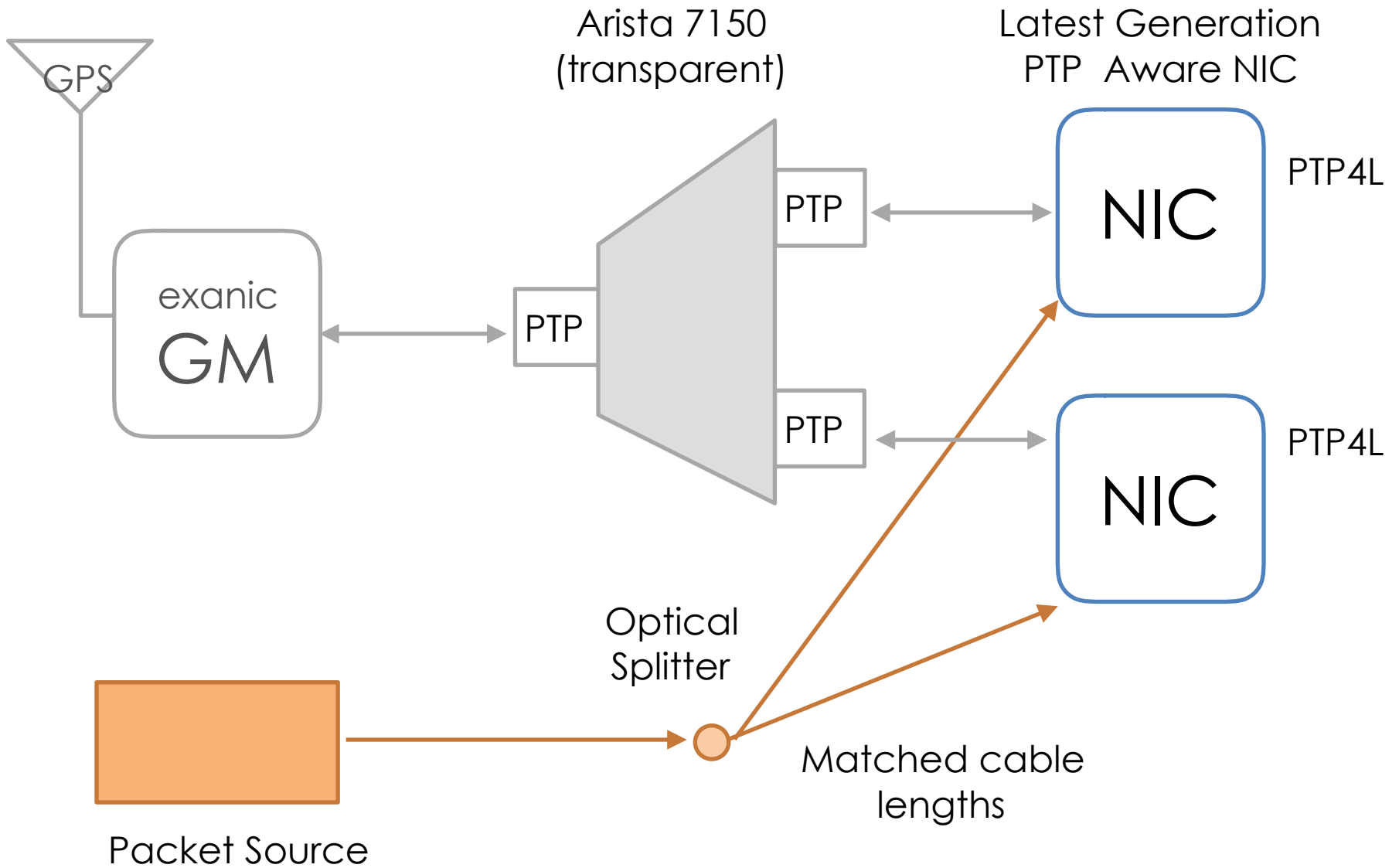


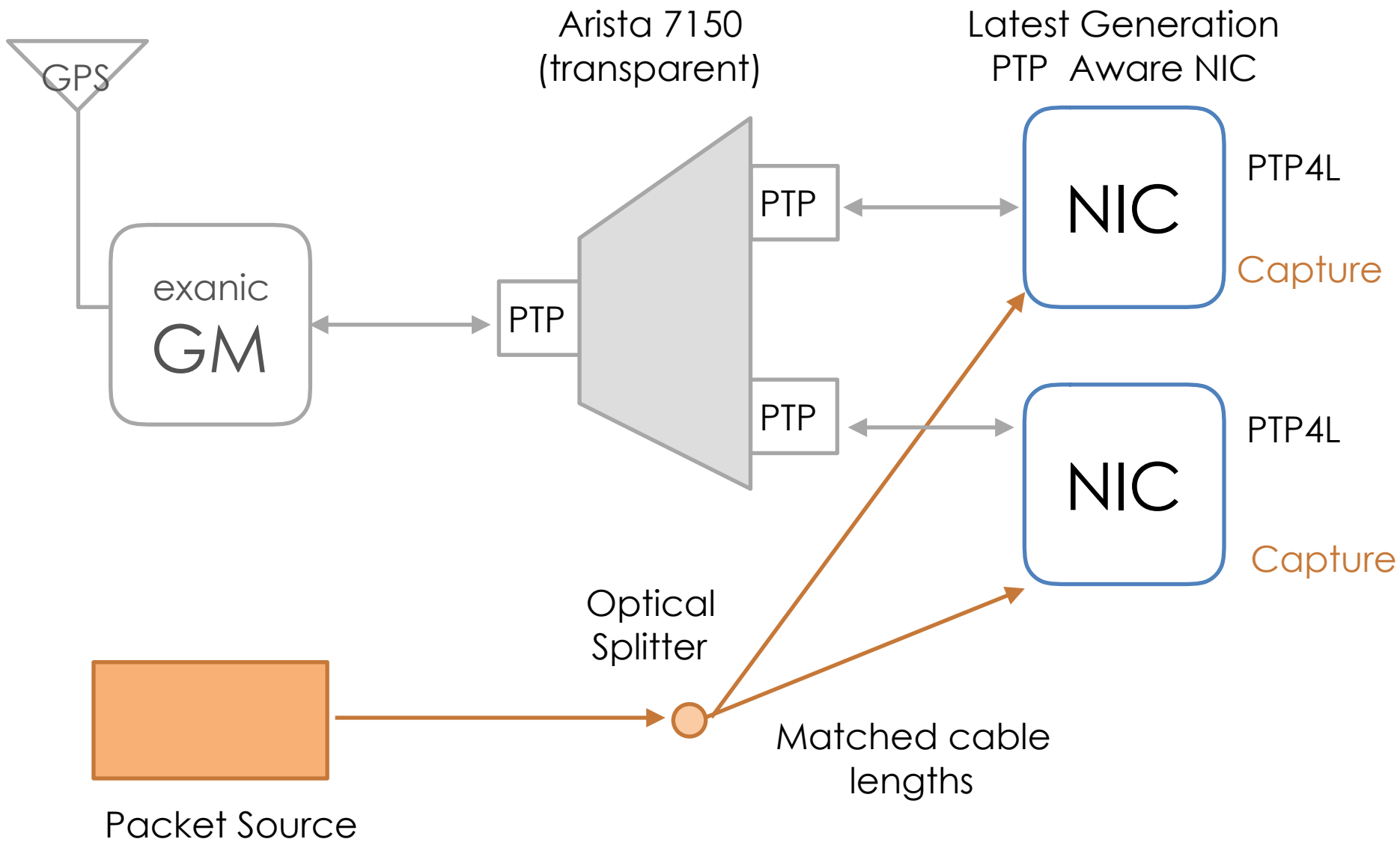
# Another way to measure PTP Sync Quality?



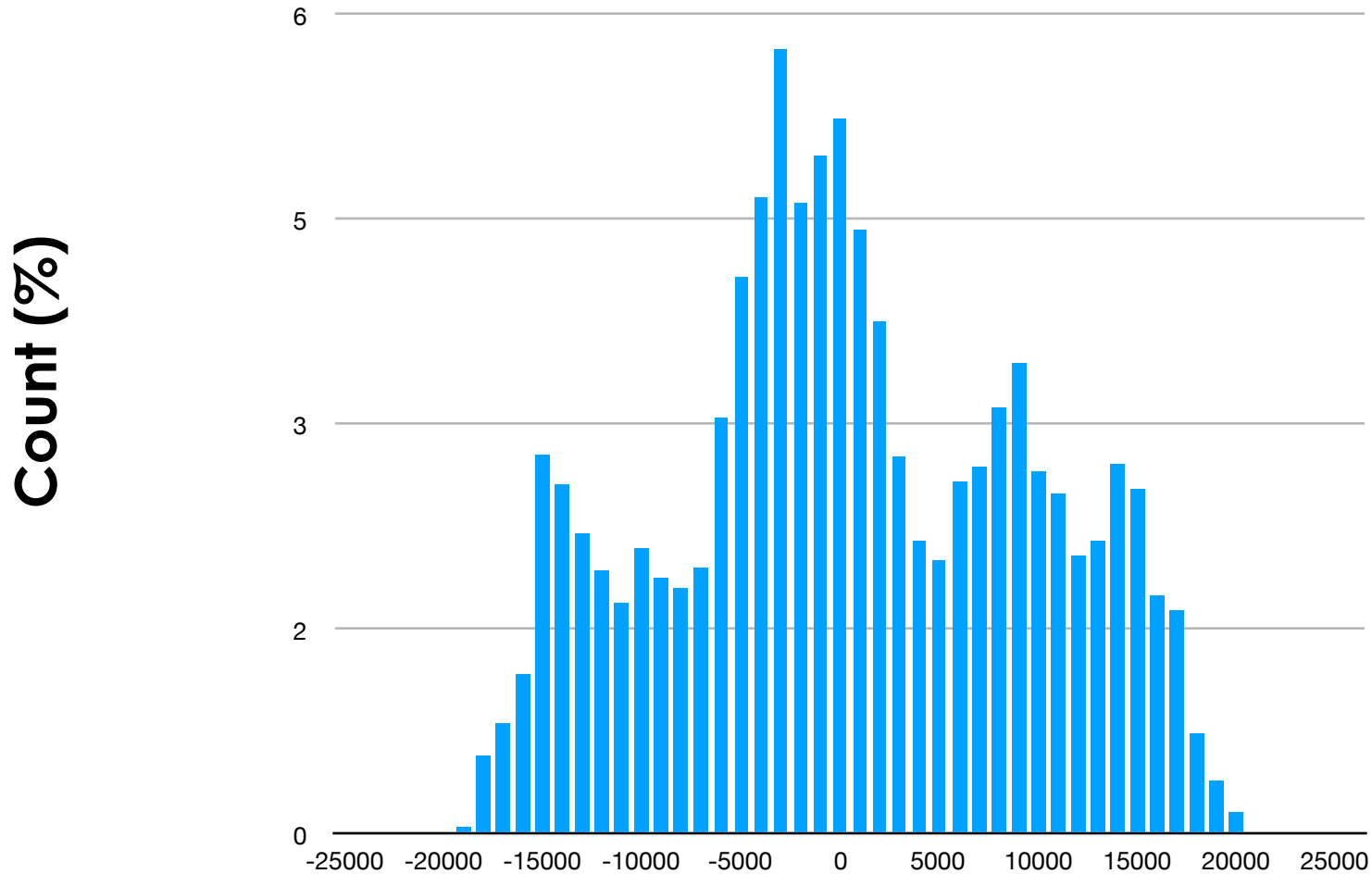
Packet Source



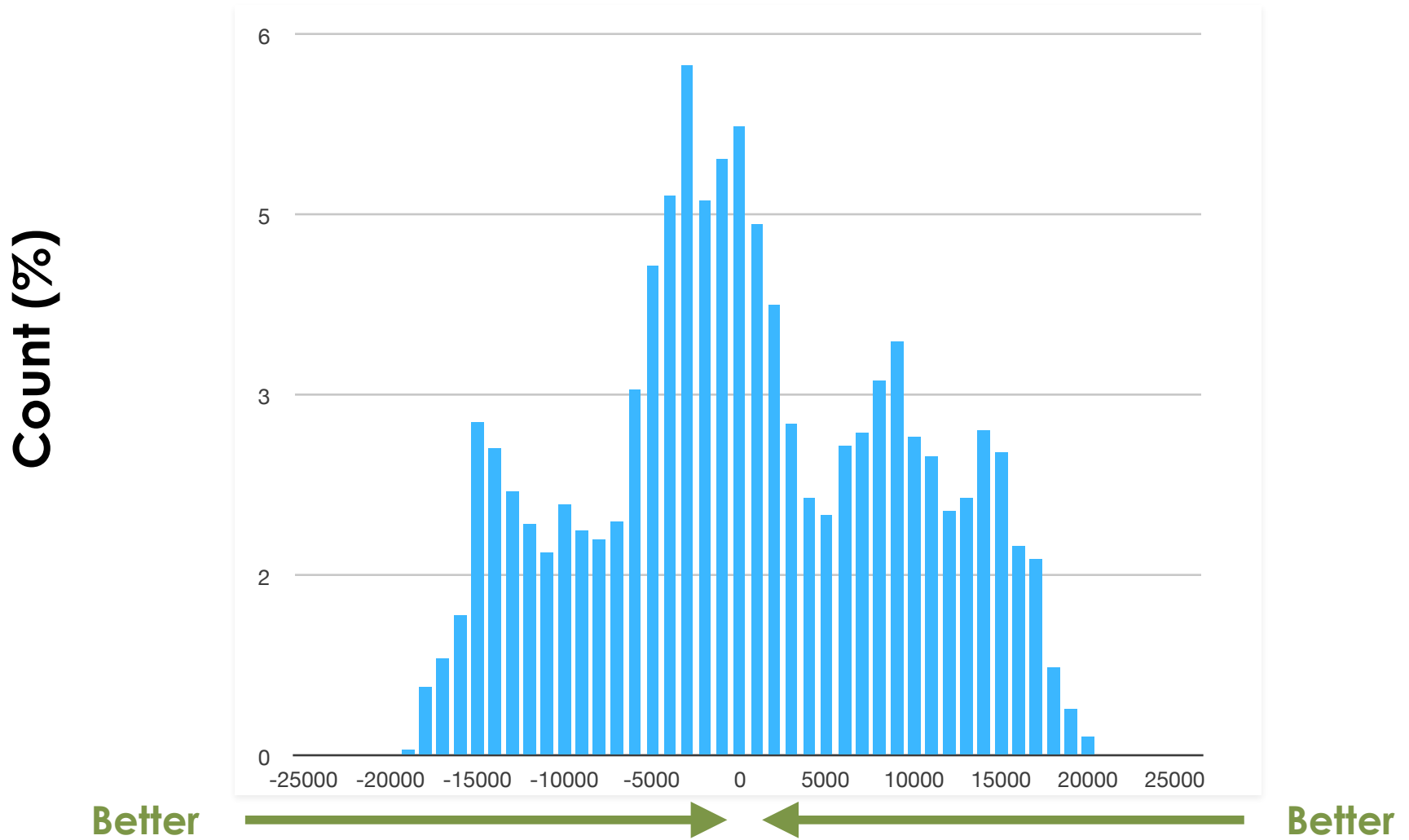




# PTP4L Measured Sync



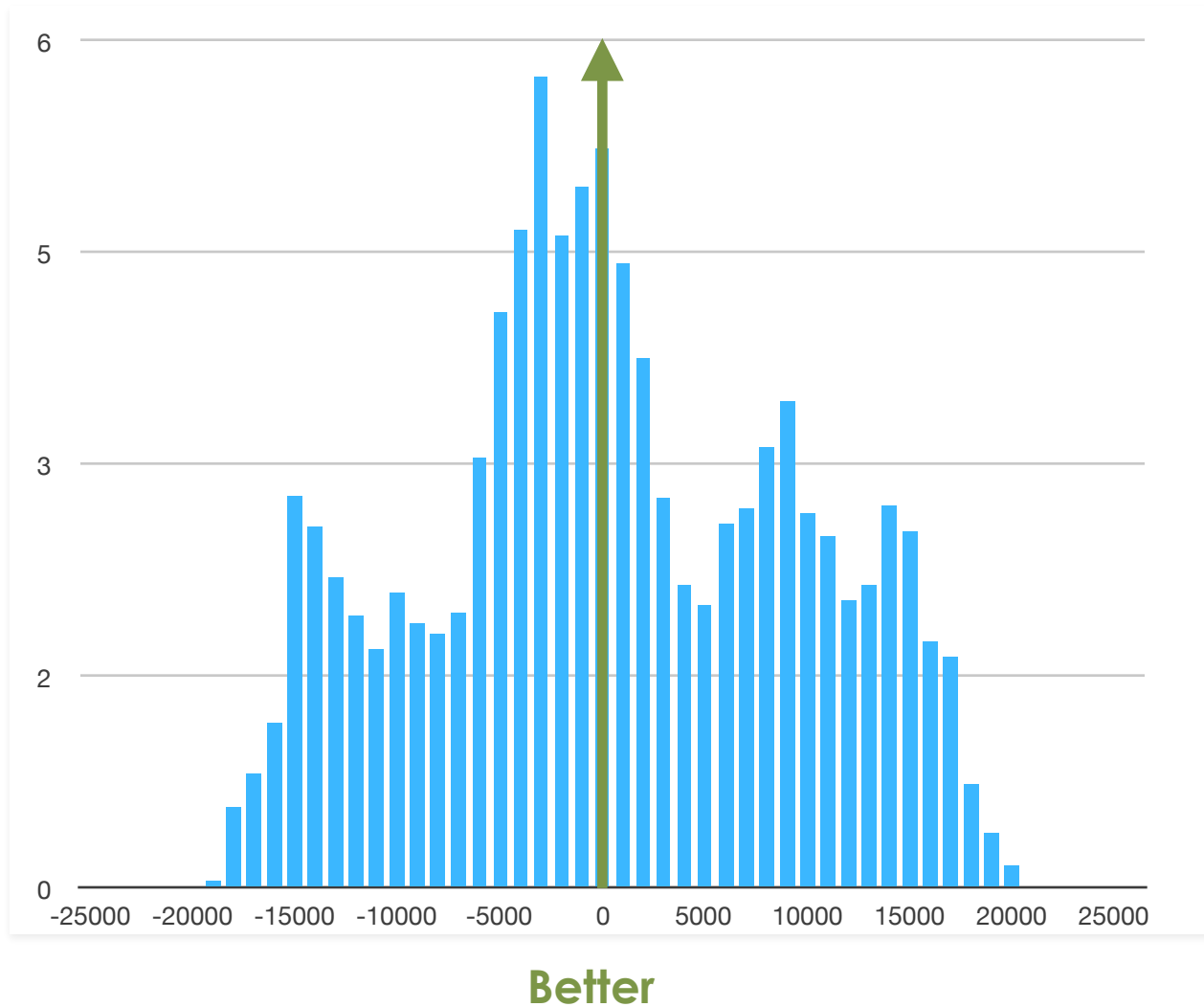
# PTP4L Measured Sync



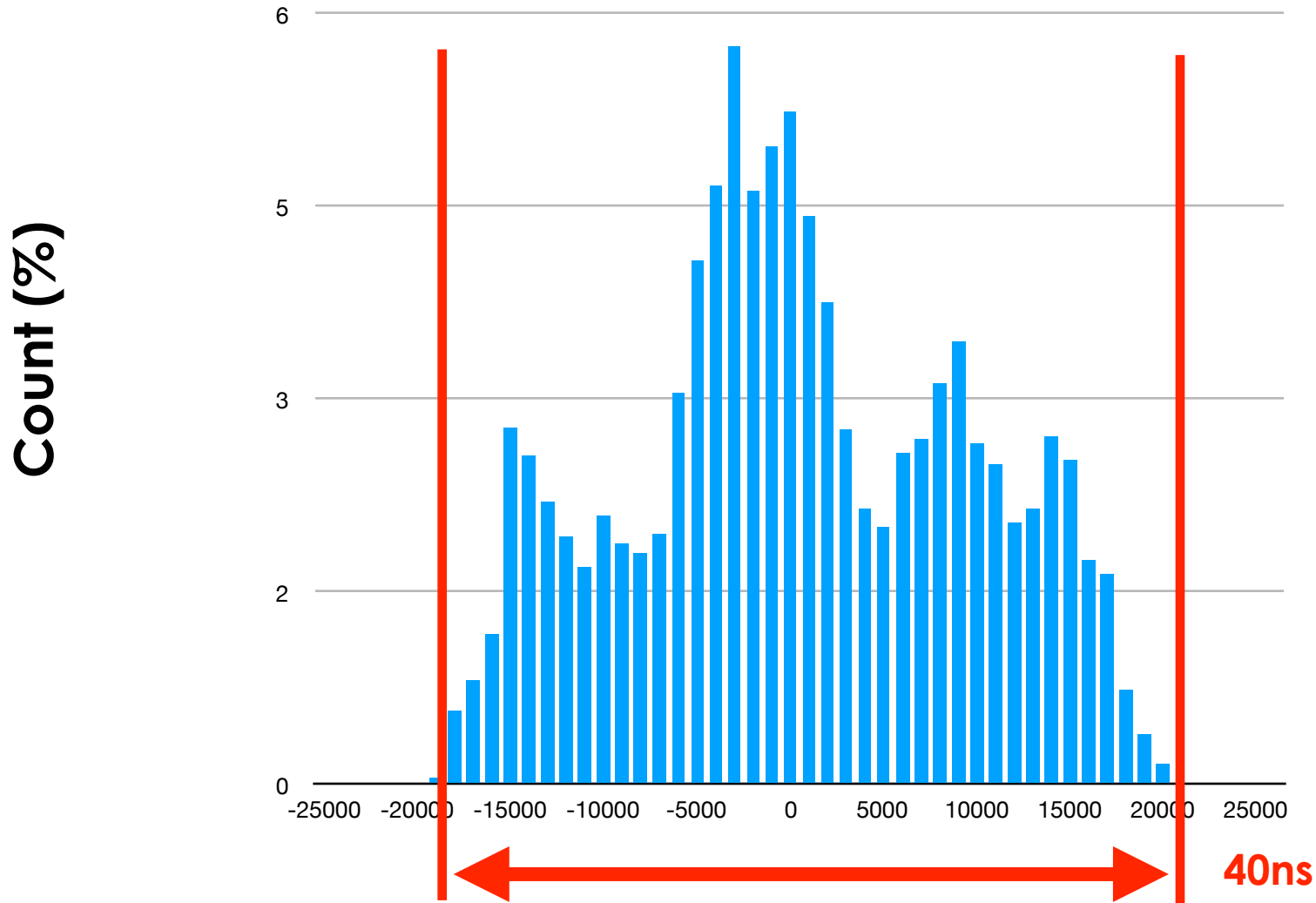


# PTP4L Measured Sync

Count (%)



# PTP4L Measured Sync



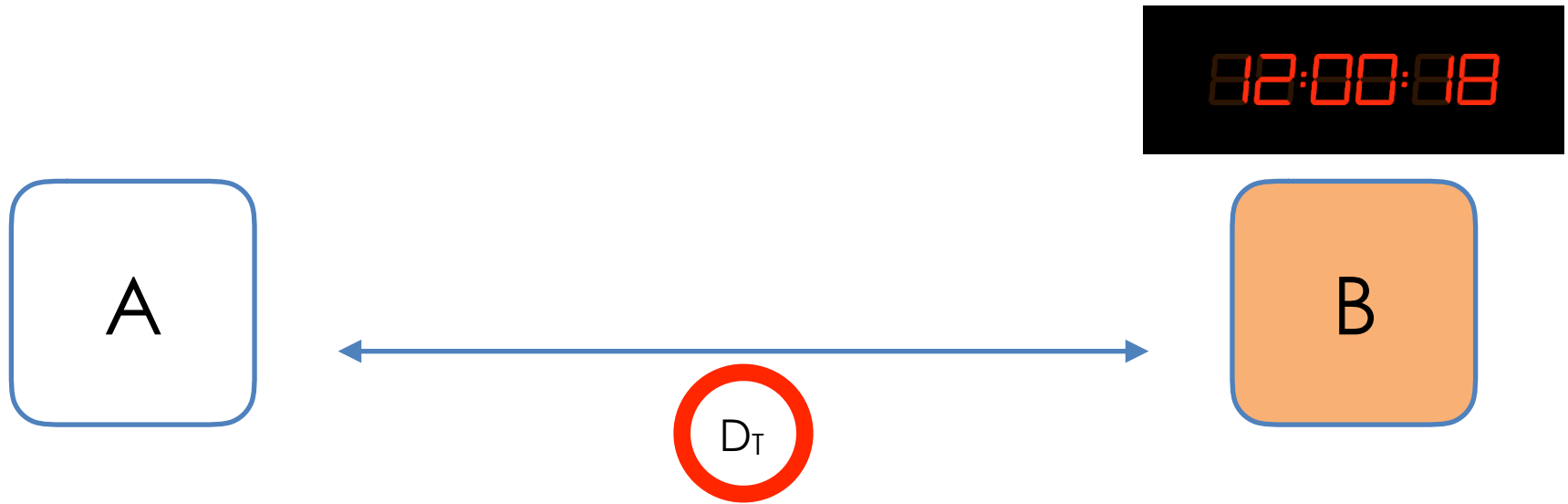
# PTP4L Sync Summary

Configuration	Perceived Sync	Measured Sync
Standard	+/- 40ns	+/- 20ns

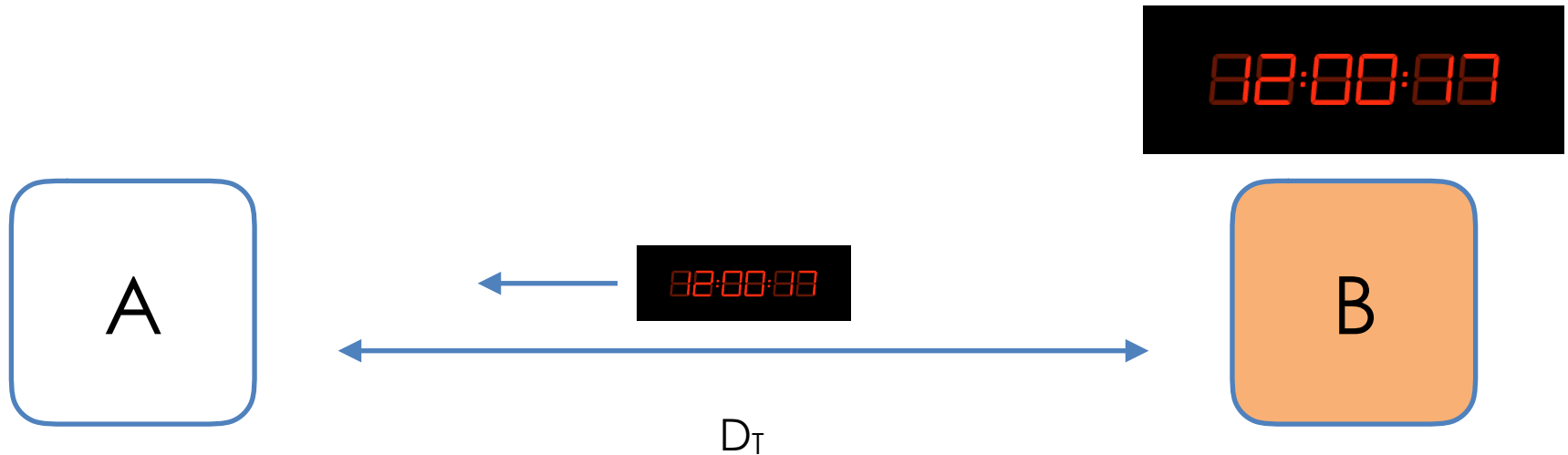
“PTP isn't good enough for a nanosecond world ”

- someone in this room (probably)

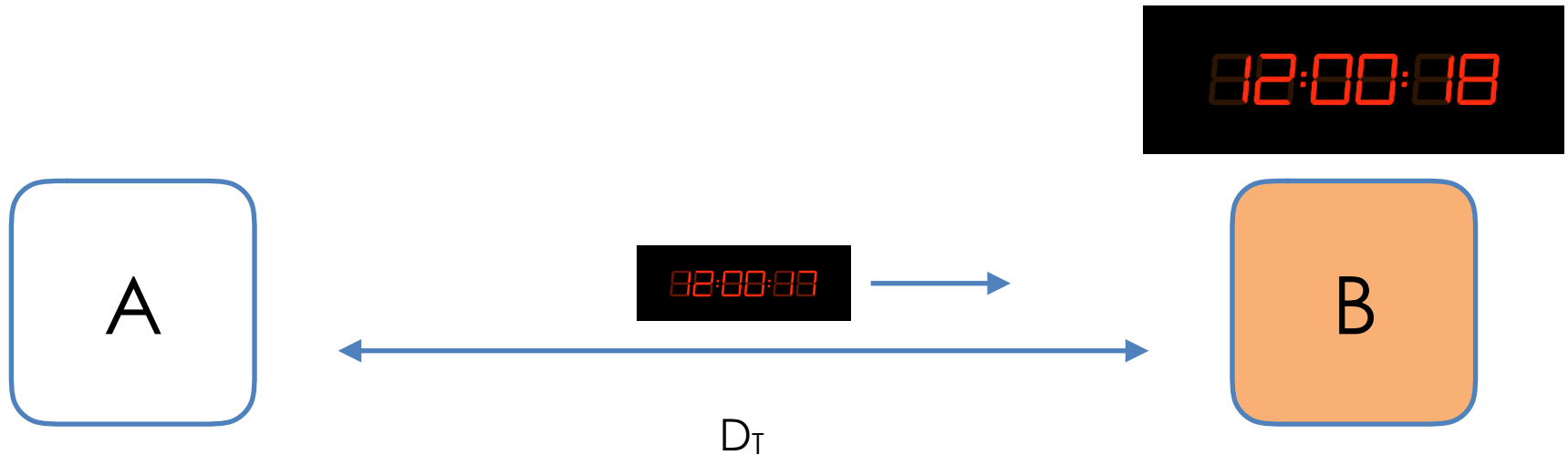
# Fundamentals of PTP



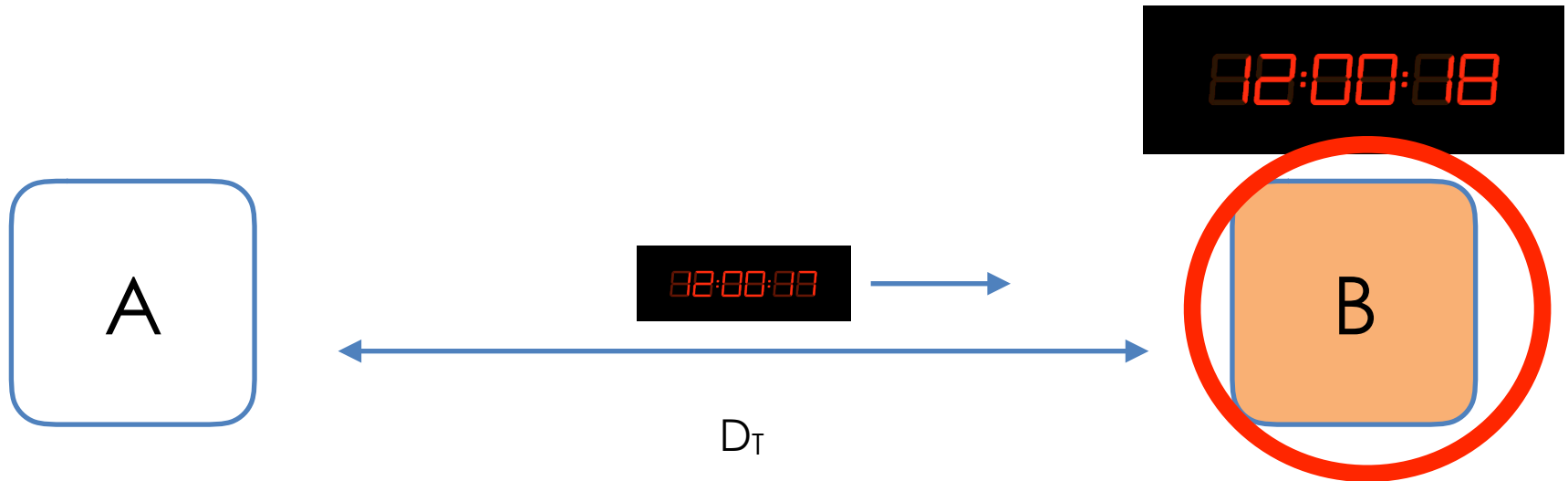
# Fundamentals of PTP



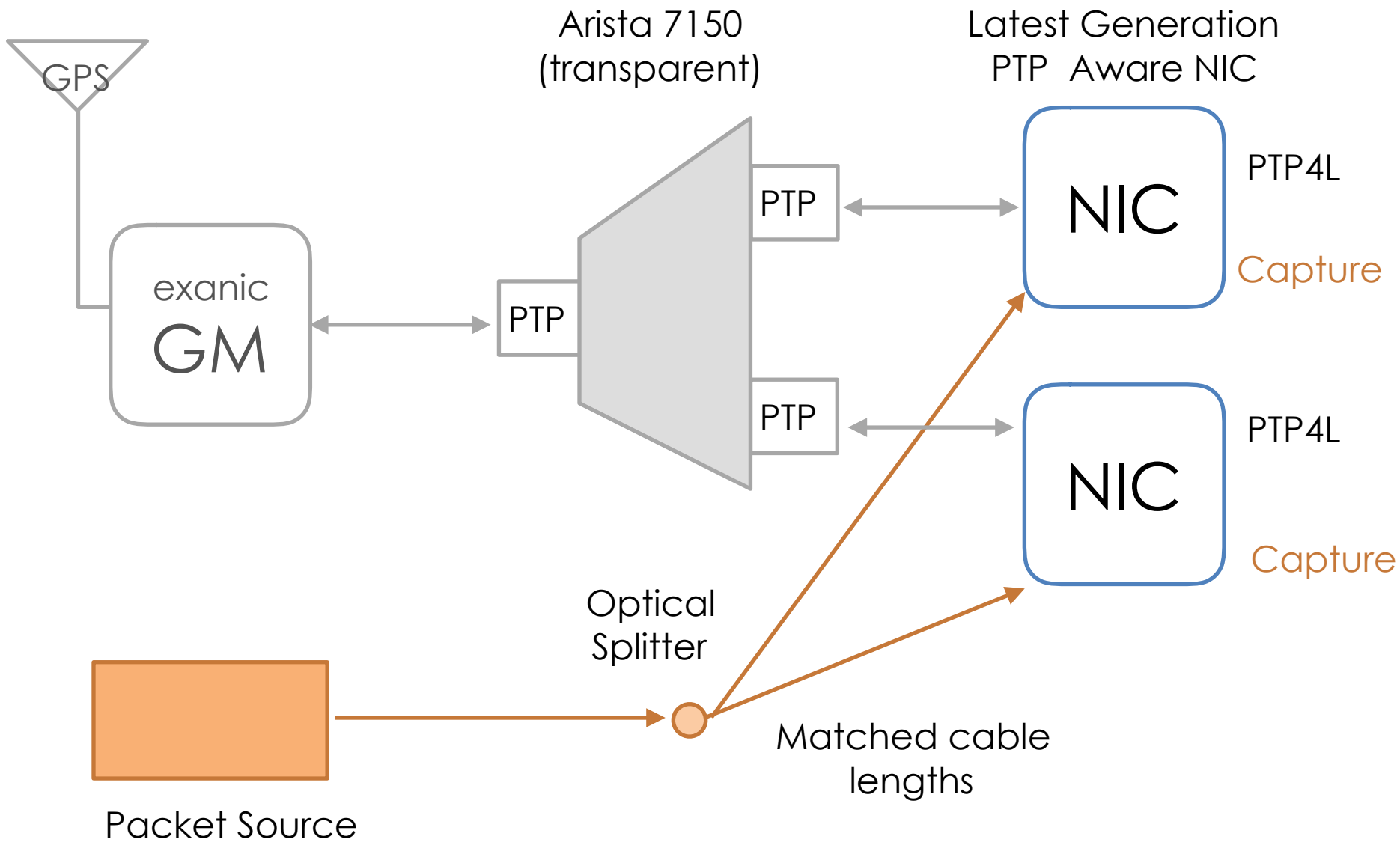
# Fundamentals of PTP

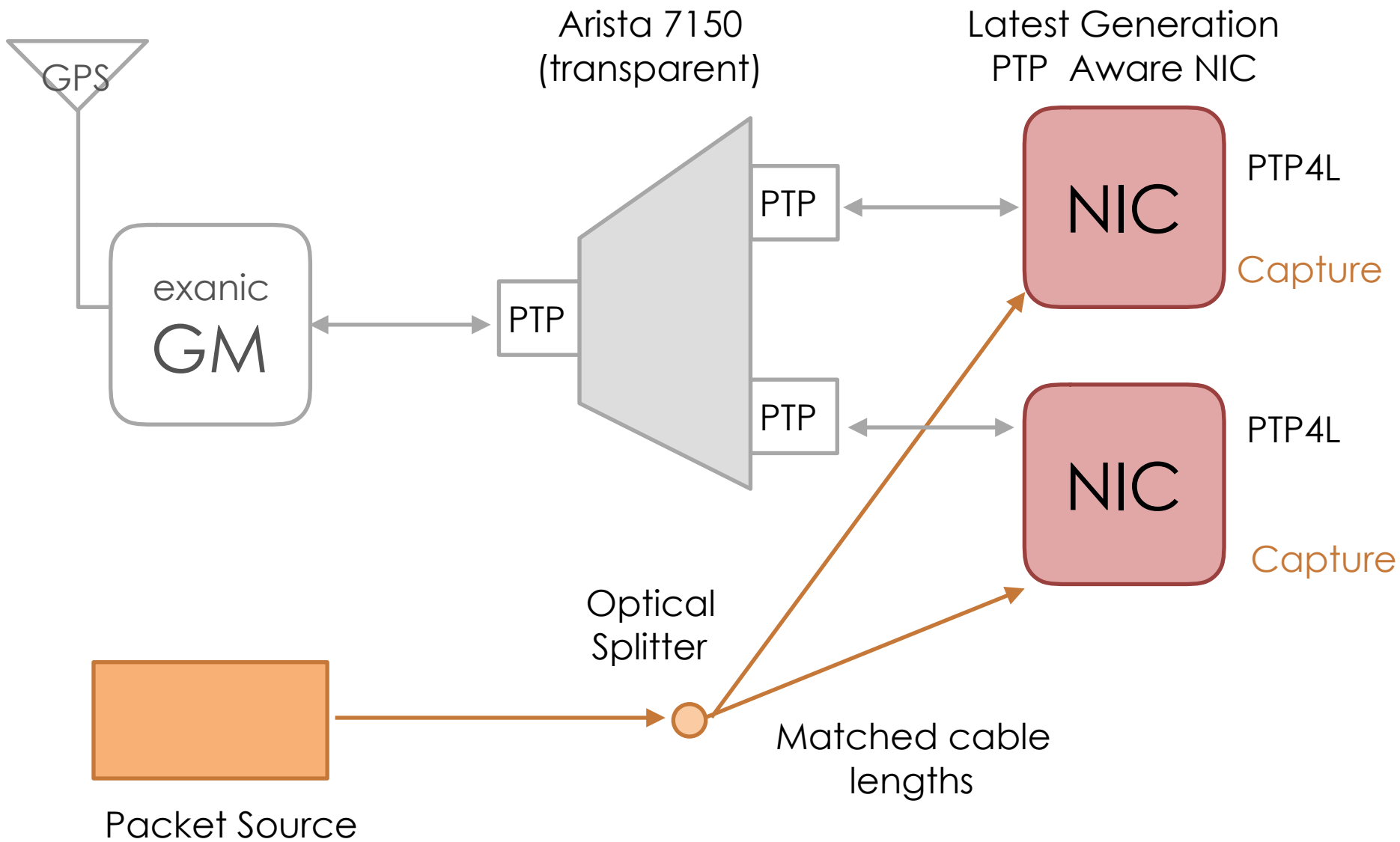


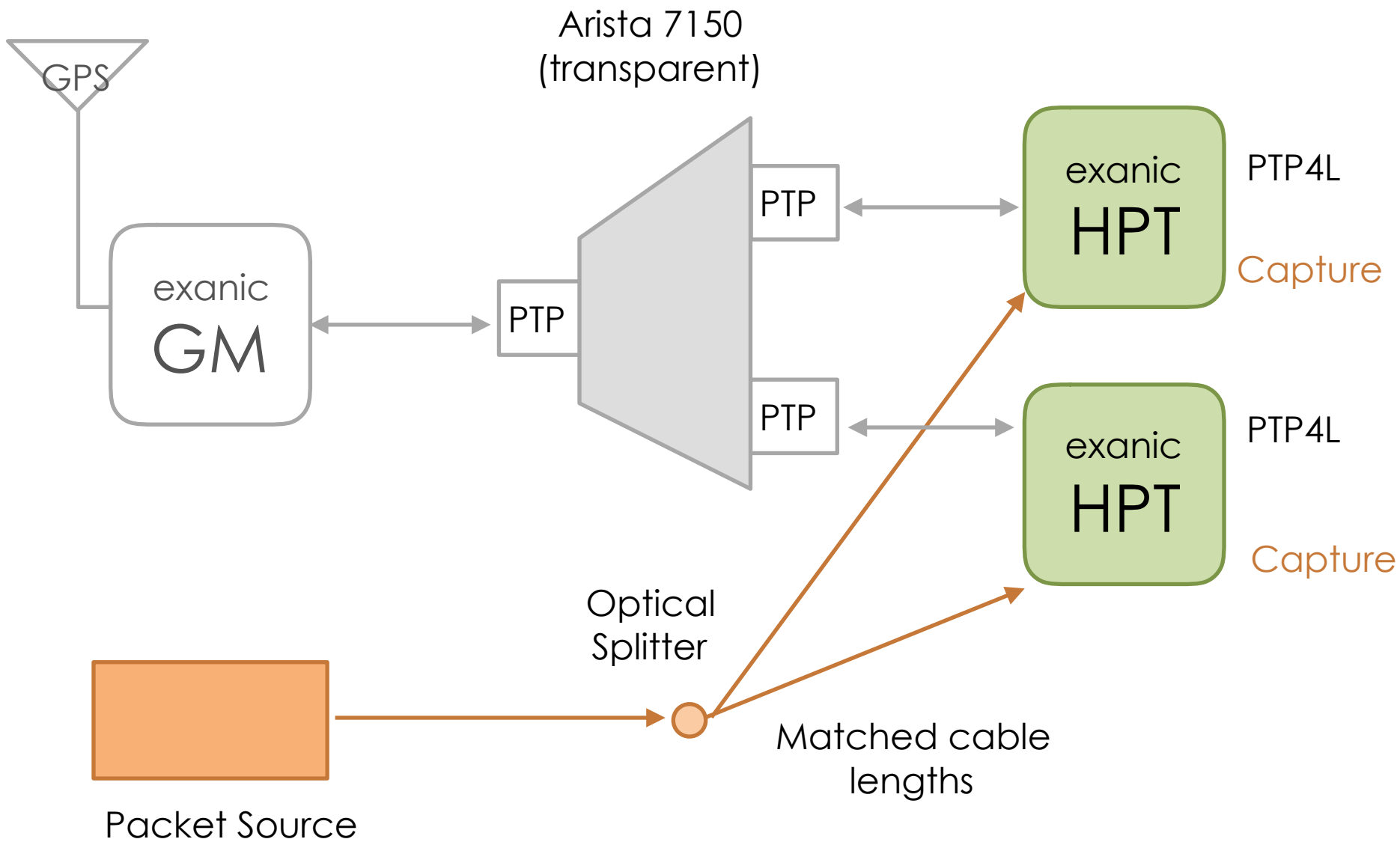
# Fundamentals of PTP



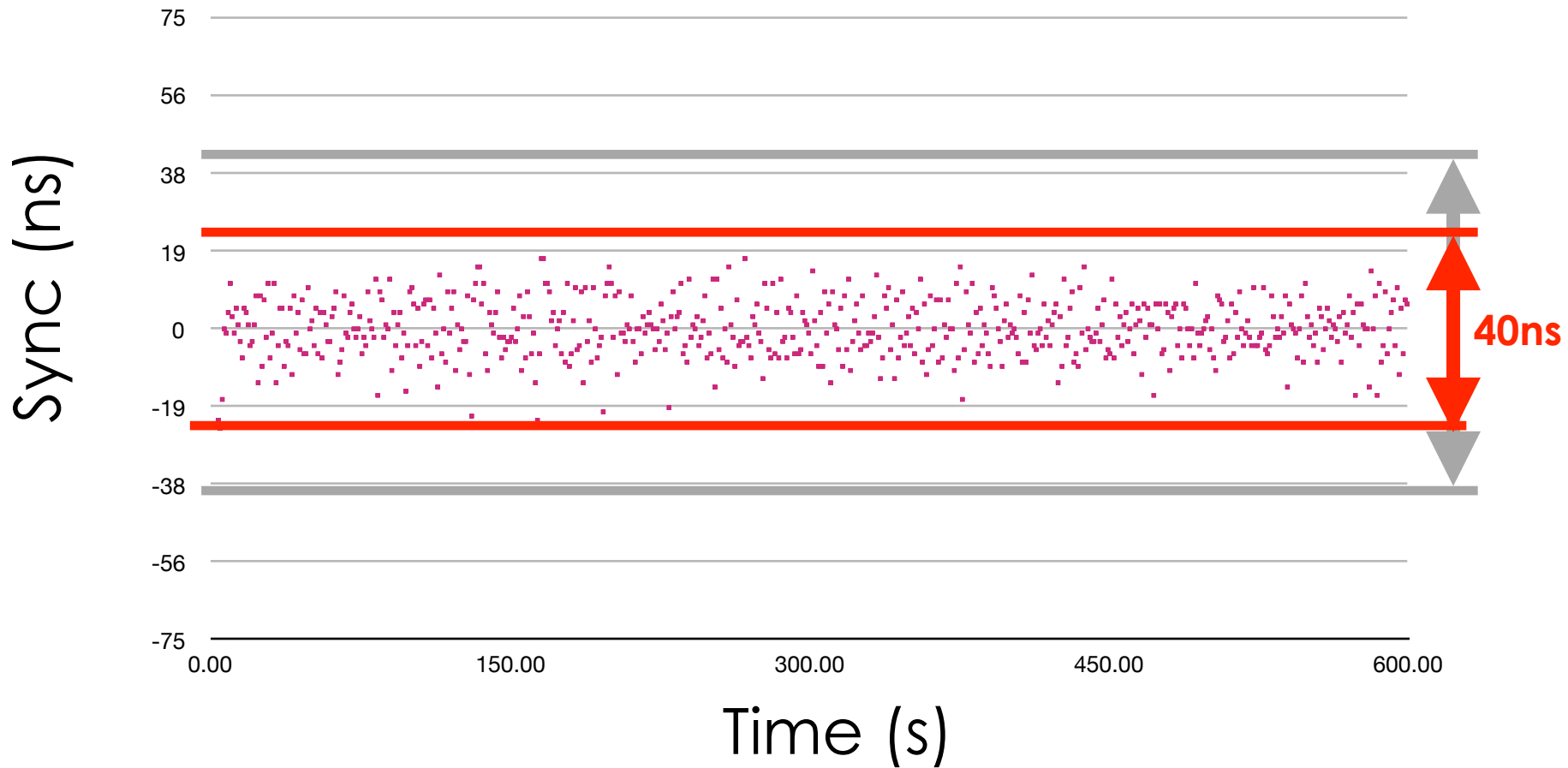




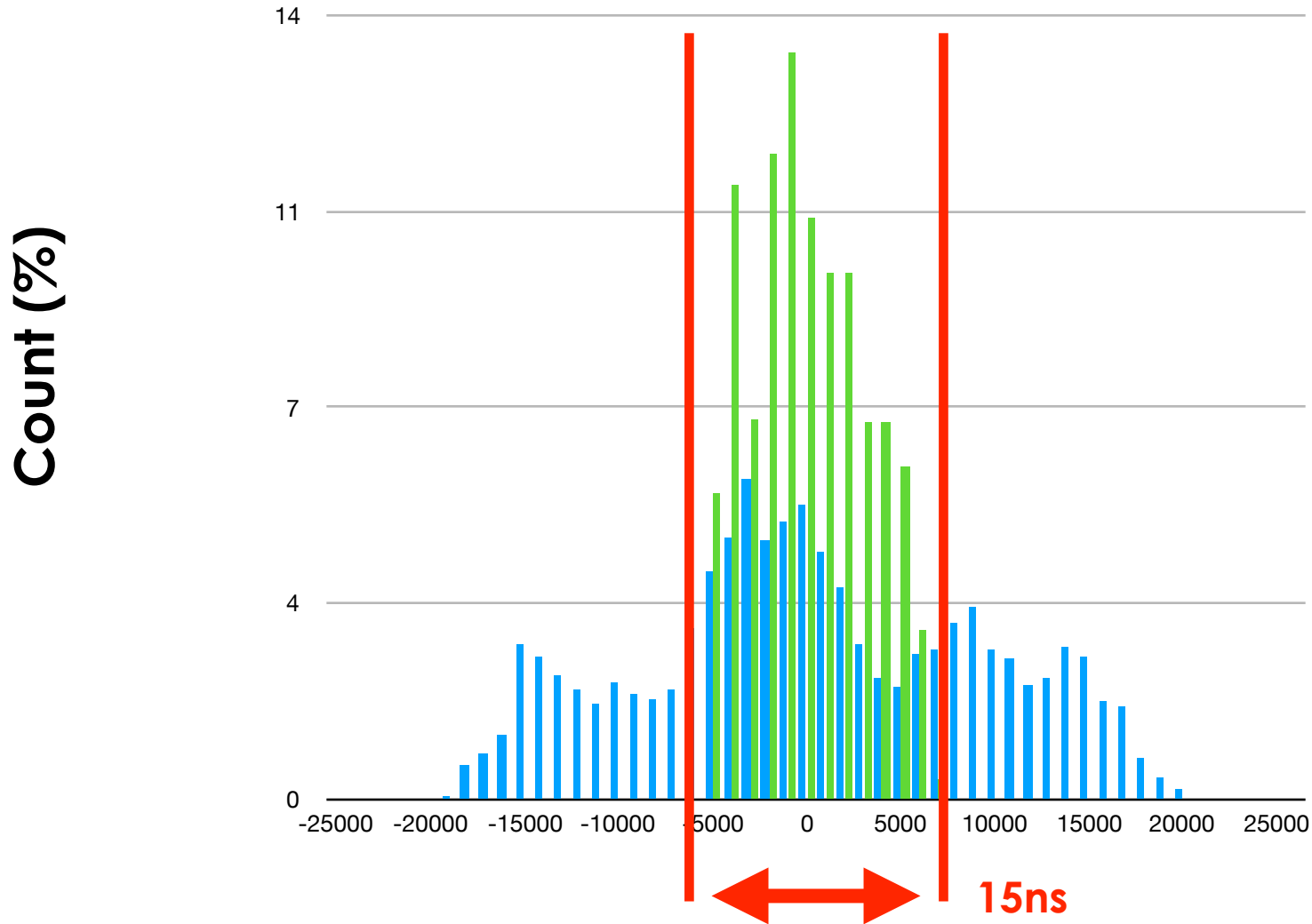




# PTP4L Perceived Sync



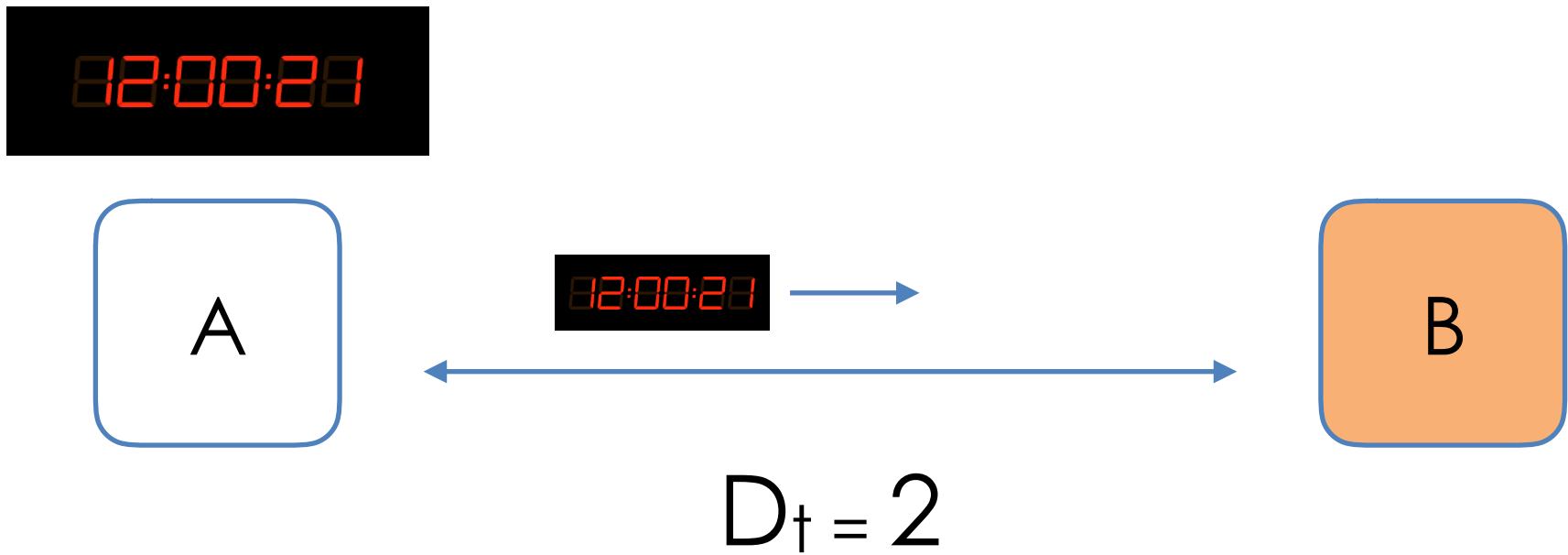
# PTP4L Measured Sync

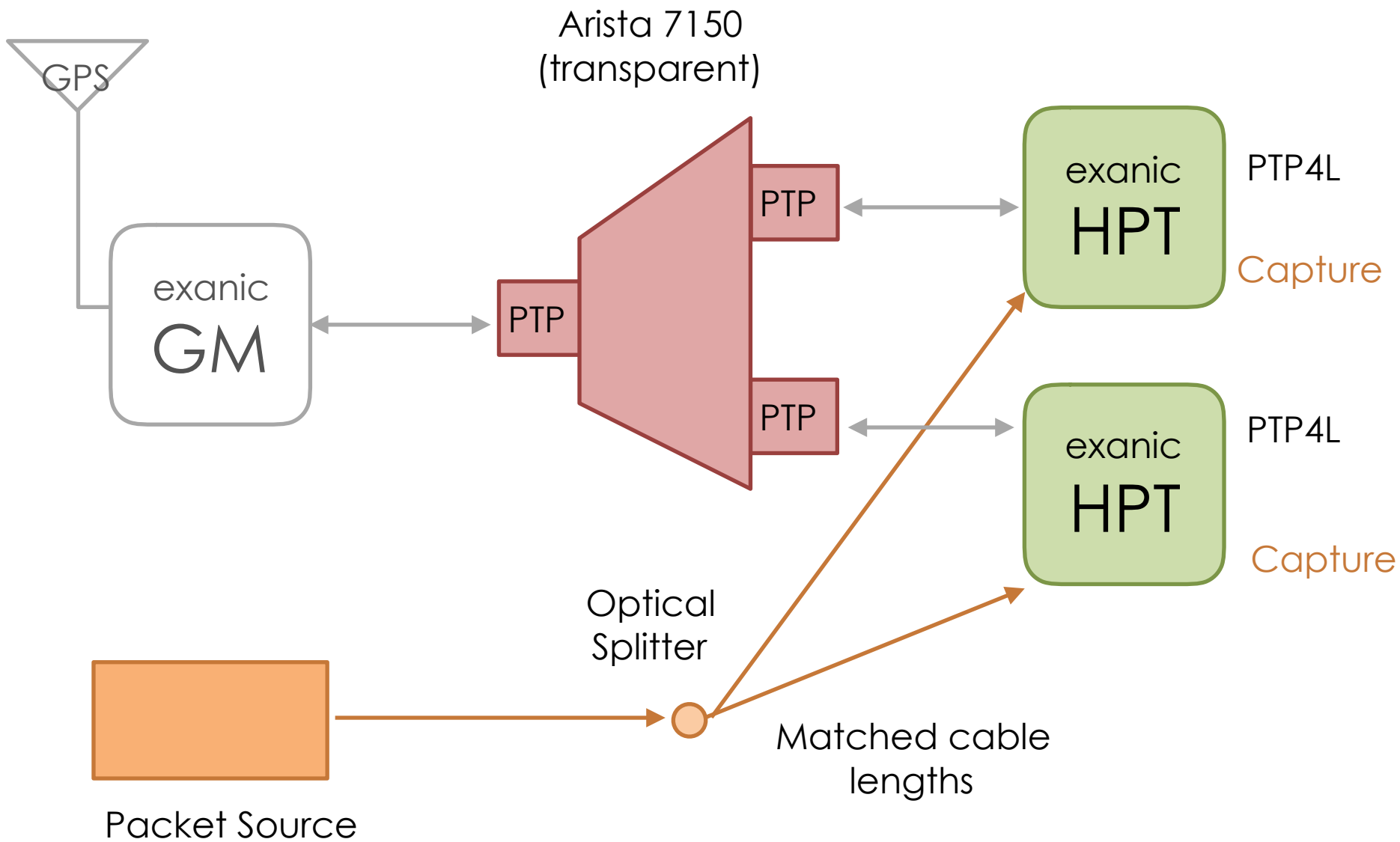


# PTP4L Sync Summary

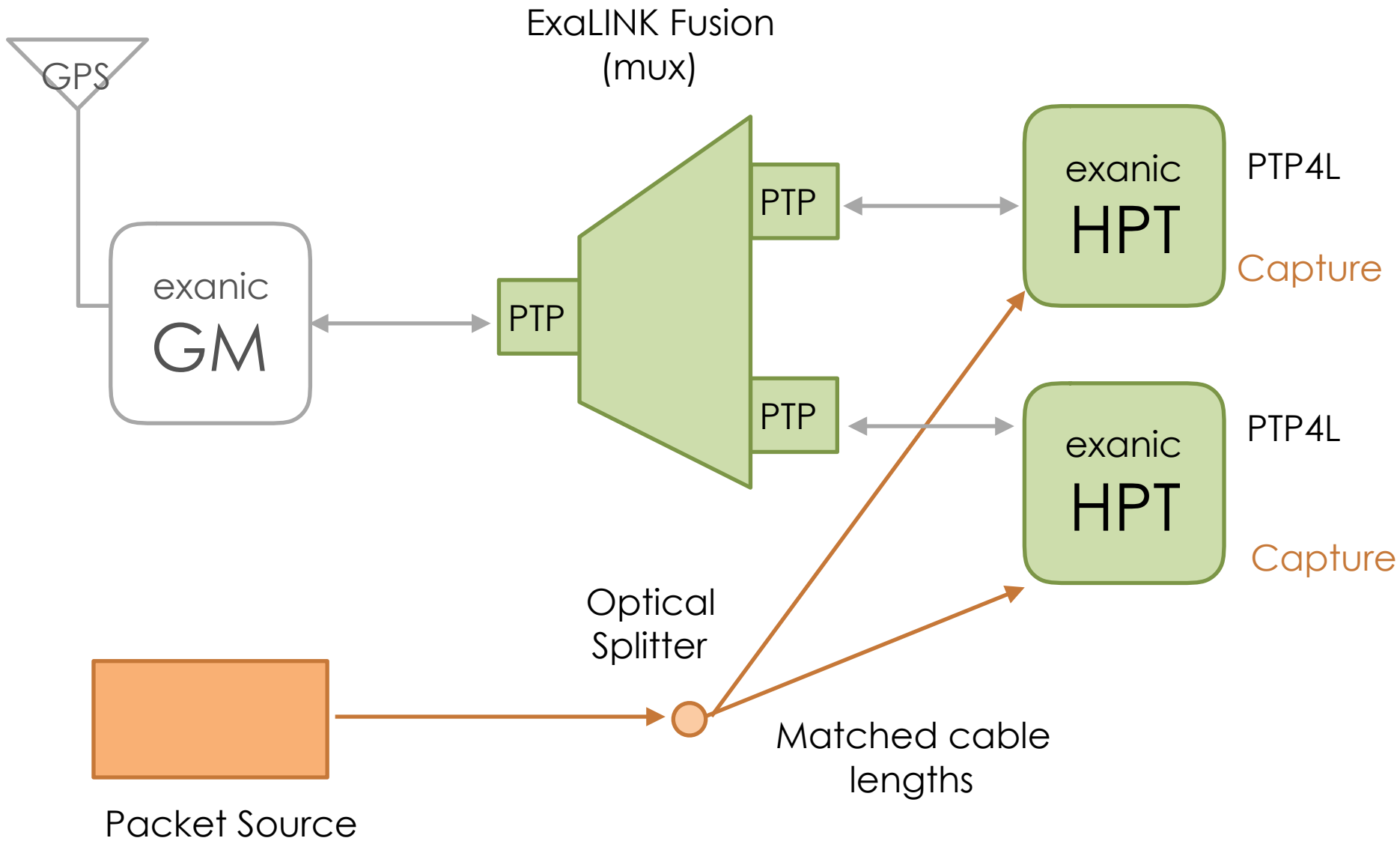
Configuration	Perceived Sync	Measured Sync
Standard	+/- 40ns	+/- 20ns
ExaNIC HPT	+/- 20ns	+/- 7.5ns

# Fundamentals of PTP

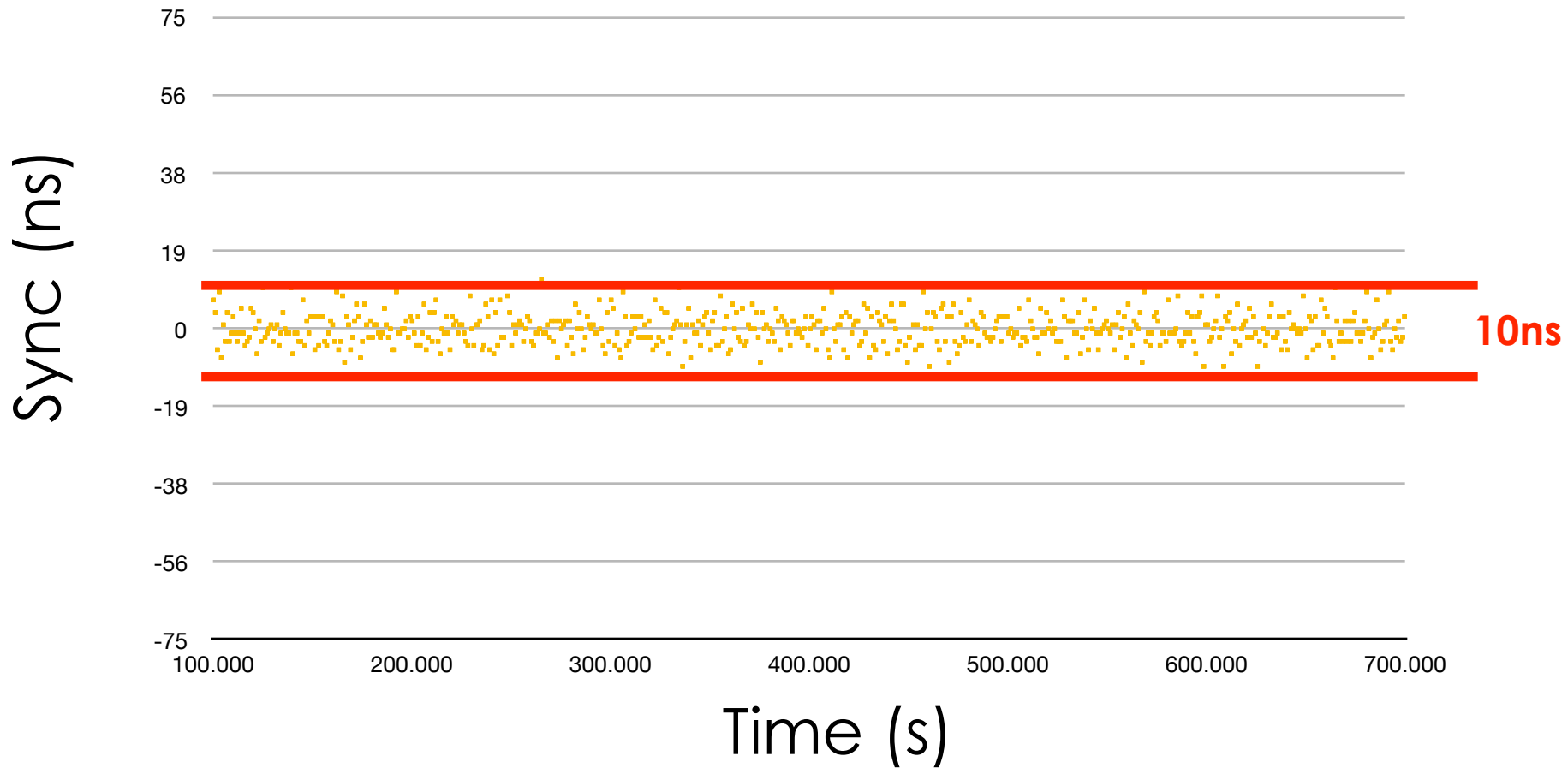




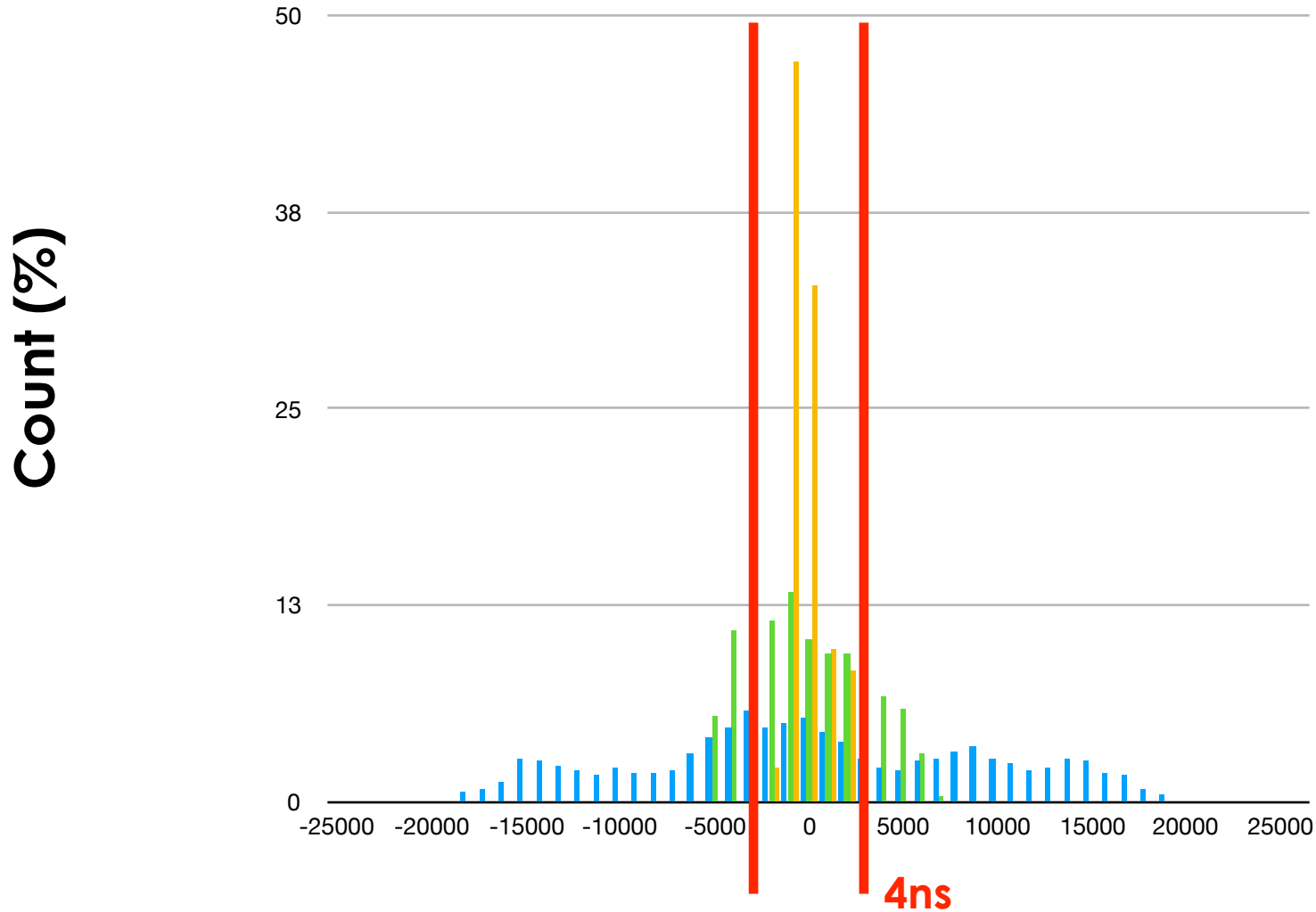




# PTP4L Perceived Sync

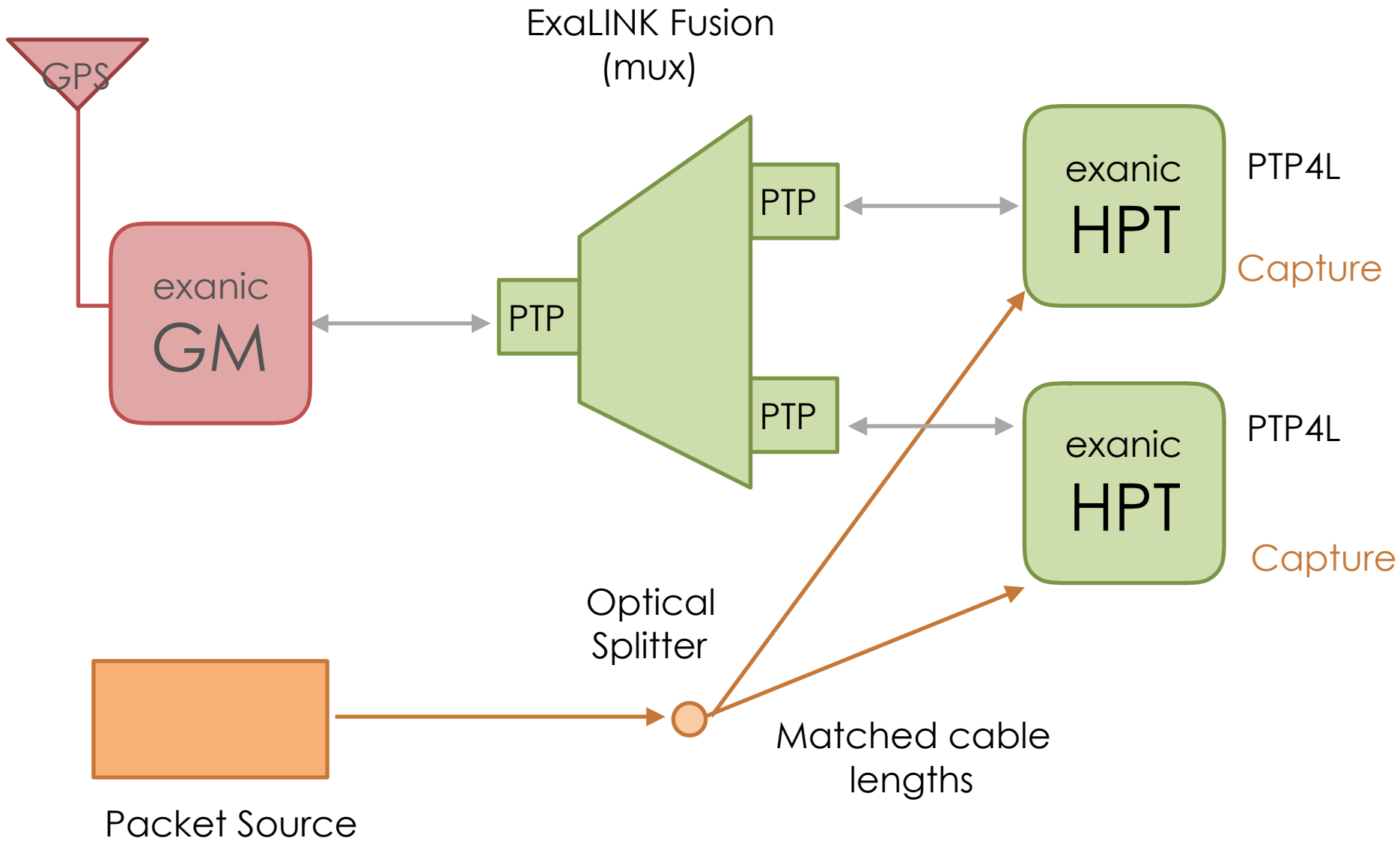


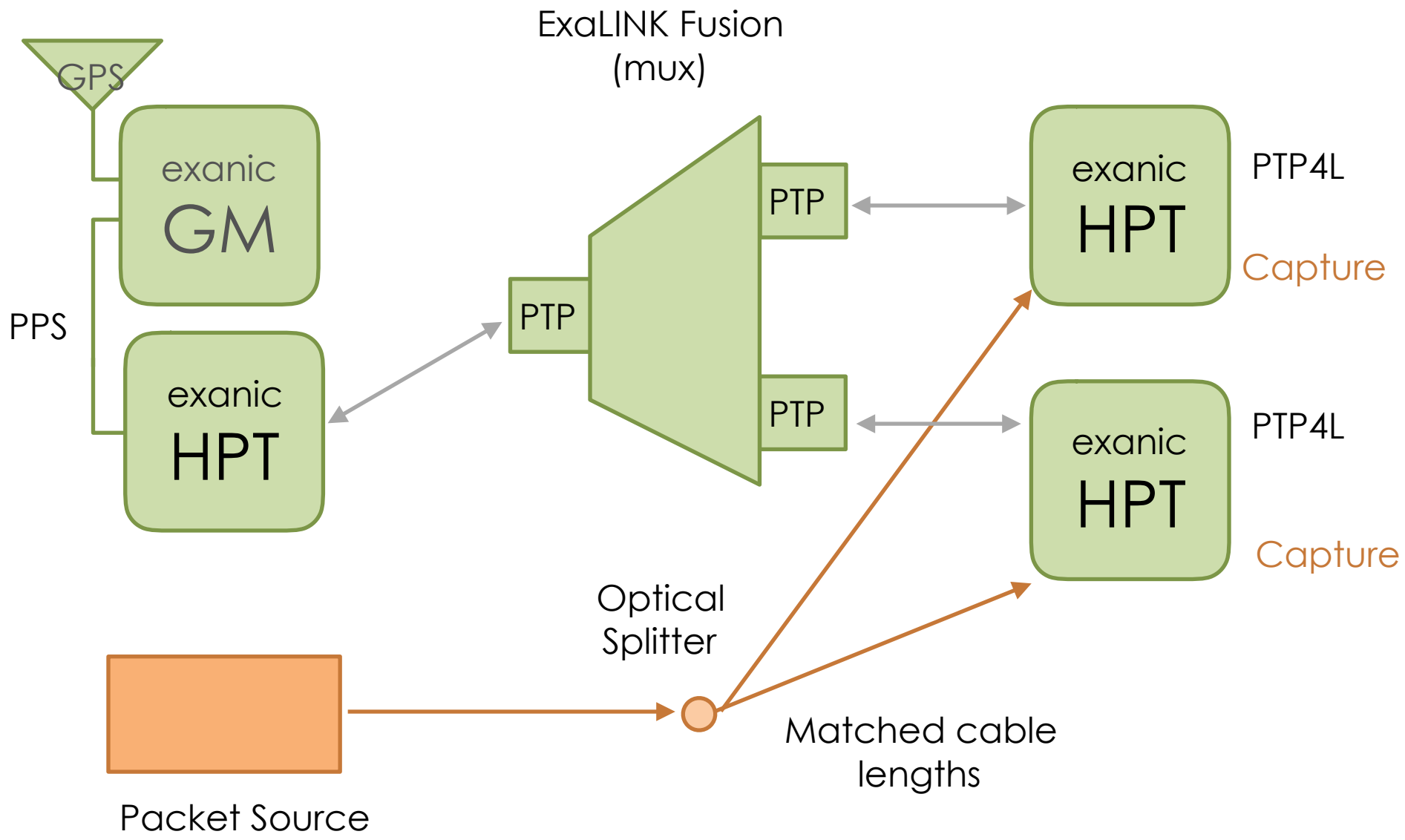
# PTP4L Measured Sync



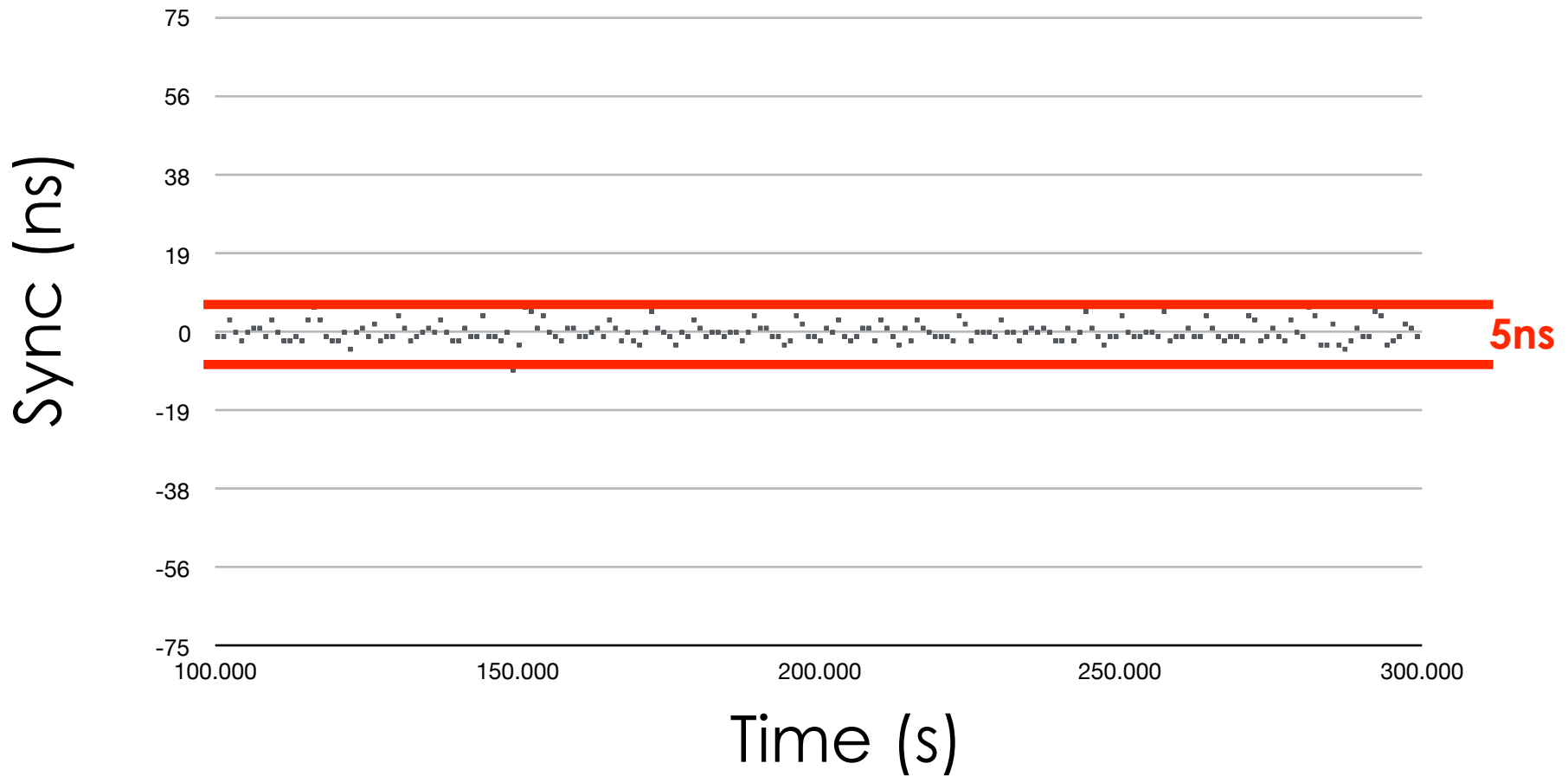
# PTP4L Sync Summary

Configuration	Perceived Sync	Measured Sync
Standard	+/- 40ns	+/- 20ns
ExaNIC HPT	+/- 20ns	+/- 7.5ns
ExaLINK Fusion	+/- 5ns	+/- 2ns

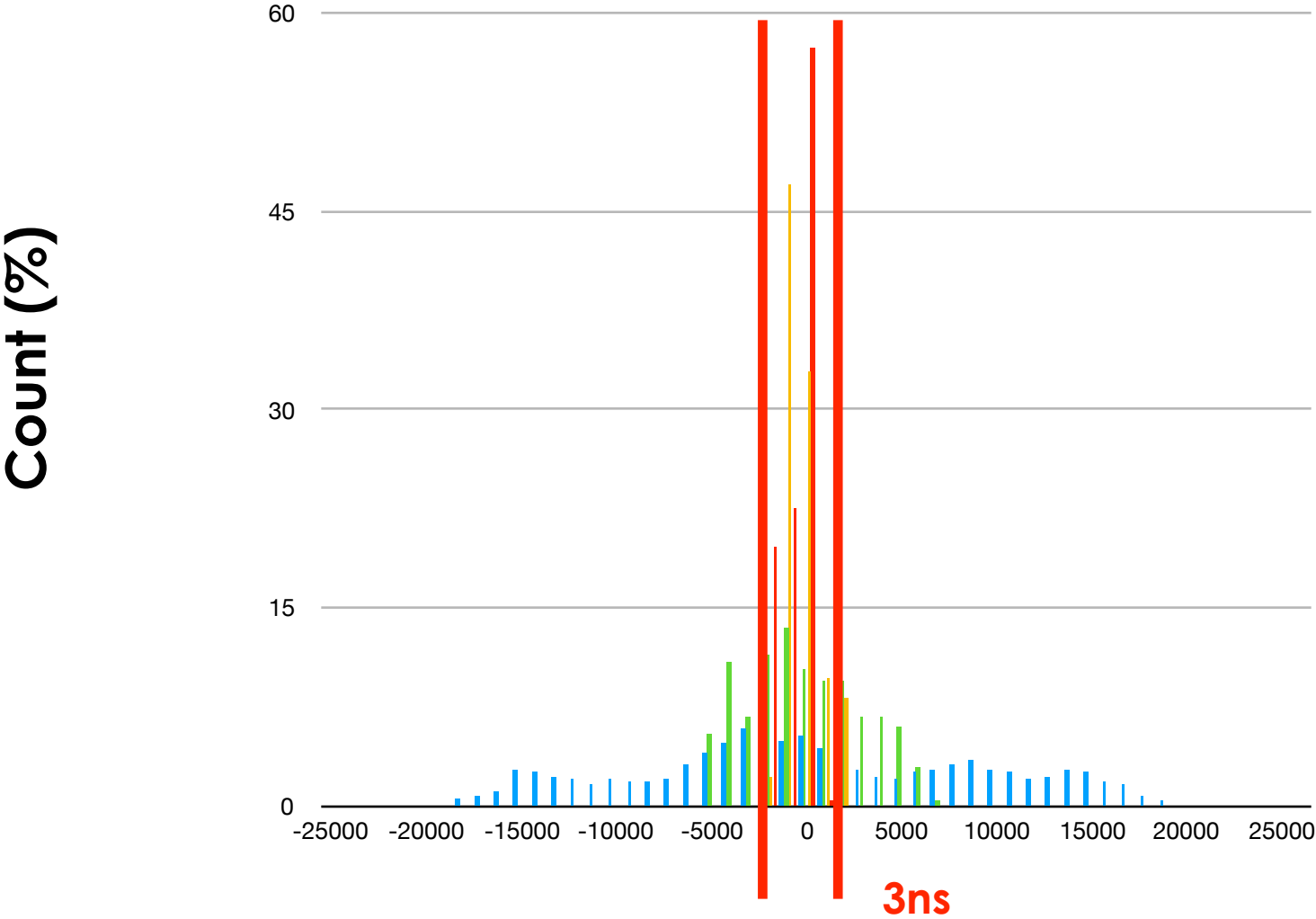




# PTP4L Perceived Sync



# PTP4L Measured Sync



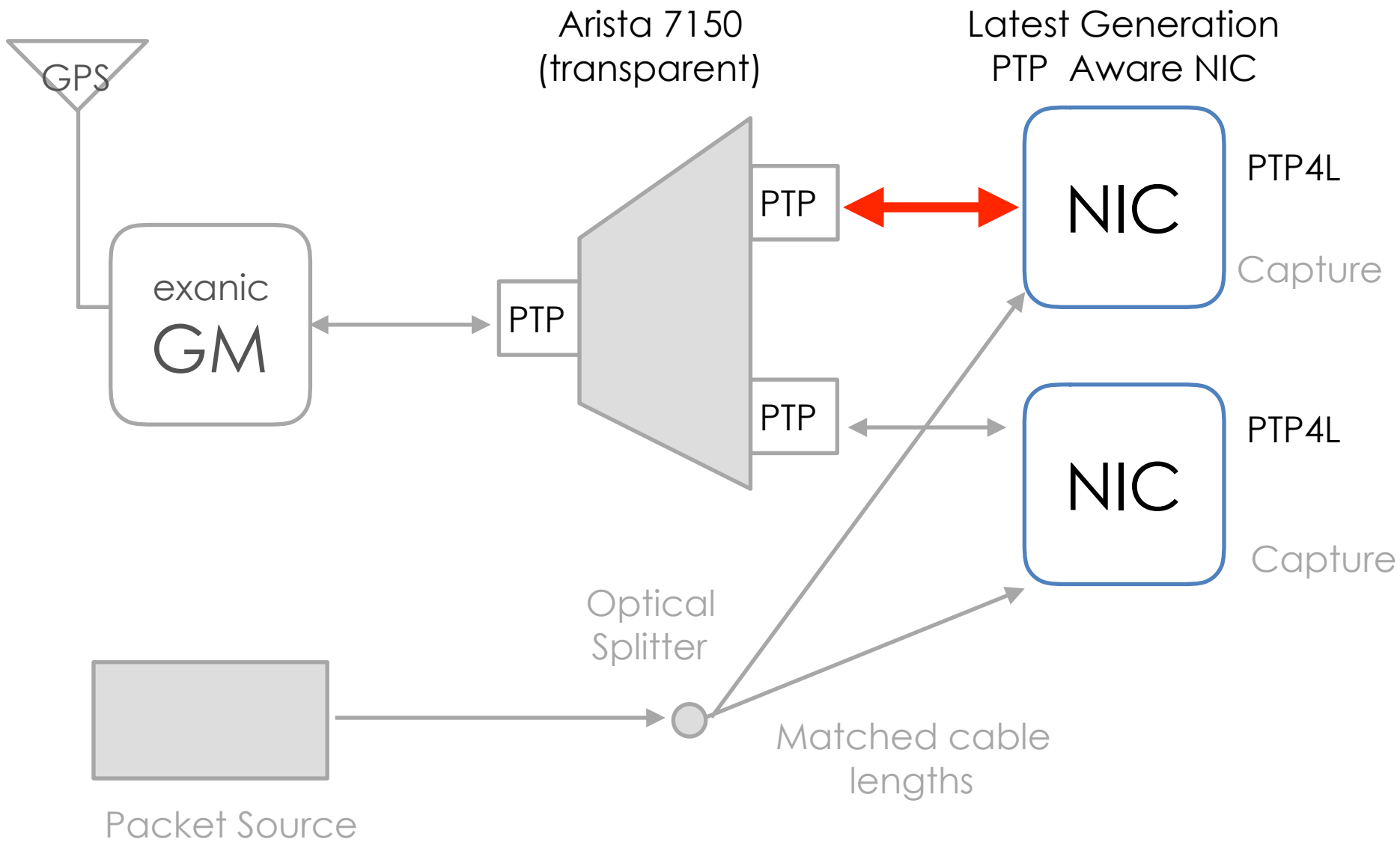


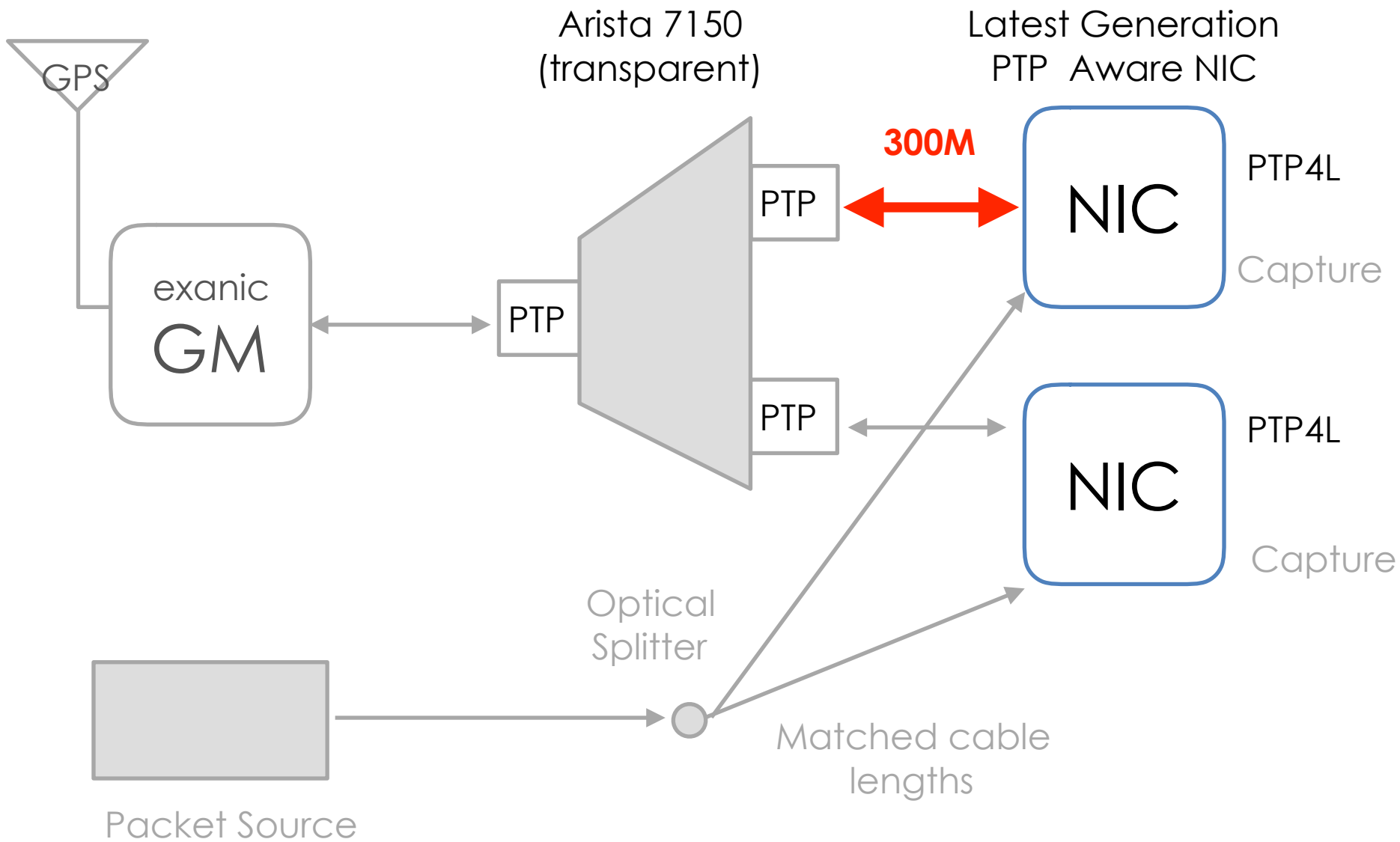
# PTP4L Sync Summary

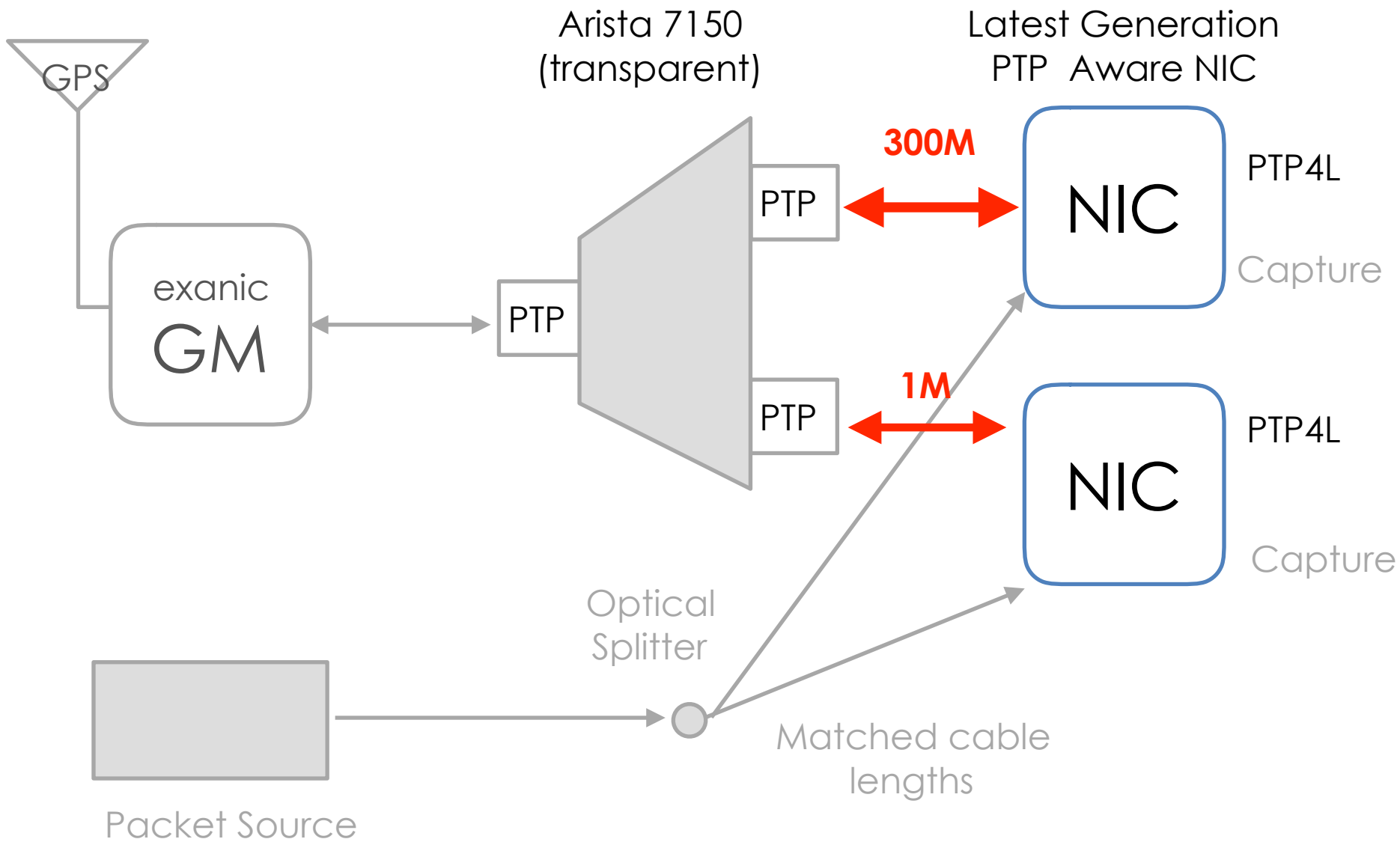
Configuration	Perceived Sync	Measured Sync
Standard	<b>+/- 40ns</b>	<b>+/- 20ns</b>
ExaNIC HPT	<b>+/- 20ns</b>	<b>+/- 7.5ns</b>
ExaLINK Fusion	<b>+/- 5ns</b>	<b>+/- 2ns</b>
GM / HPT	<b>+/- 2.5</b>	<b>+/- 1.5</b>

# PTP4L Sync Summary

Configuration	Perceived Sync	Measured Sync
Standard	+/- 40ns	+/- 20ns
ExaNIC HPT	+/- 20ns	+/- 7.5ns
ExaLINK Fusion	+/- 5ns	+/- 2ns
GM / HPT	+/- 2.5	+/- 1.5







# PTP4L Sync Summary

Configuration	Perceived Sync	Measured Sync
Standard	+/- 40ns	+/- 20ns
ExaNIC HPT	+/- 10ns	+/- 5ns
ExaLINK Fusion	+/- 5ns	+/- 2ns
GM / HPT	+/- 2.5	+/- 1.5

# Plain Old PTP is capable of

Plain Old PTP is capable of  
 $\pm 1.5\text{ns}$  (measured) time sync:



Plain Old PTP is capable of  $\pm 1.5\text{ns}$  (measured) time sync:

- Standard Ethernet network

Plain Old PTP is capable of  $\pm 1.5\text{ns}$  (measured) time sync:

- Standard Ethernet network
- Open source software

Plain Old PTP is capable of  $\pm 1.5\text{ns}$  (measured) time sync:

- Standard Ethernet network
- Open source software
- Familiar network hardware

Plain Old PTP is capable of  $\pm 1.5\text{ns}$  (measured) time sync:

- Standard Ethernet network
- Open source software
- Familiar network hardware
- Scale distance and hosts

“PTP **is** good enough for a  
nanosecond world ”

- people in this room (hopefully?)



Questions?



Tick the box for more info...