



How hard could it be?

Understanding network traffic
at the picosecond level

Firmware Development Kit Expansion Pack

- Low(er) latency PCS/MAC



Firmware Development Kit Expansion Pack



- Low(er) latency PCS/MAC
- PCIe DMA Engine

Firmware Development Kit Expansion Pack



- Low(er) latency PCS/MAC
- PCIe DMA Engine
- HW Accelerated TCP Engine (ATE)

Firmware Development Kit Expansion Pack



- Low(er) latency PCS/MAC
- PCIe DMA Engine
- HW Accelerated TCP Engine (ATE)
- Software, drivers, tooling

Firmware Development Kit Expansion Pack



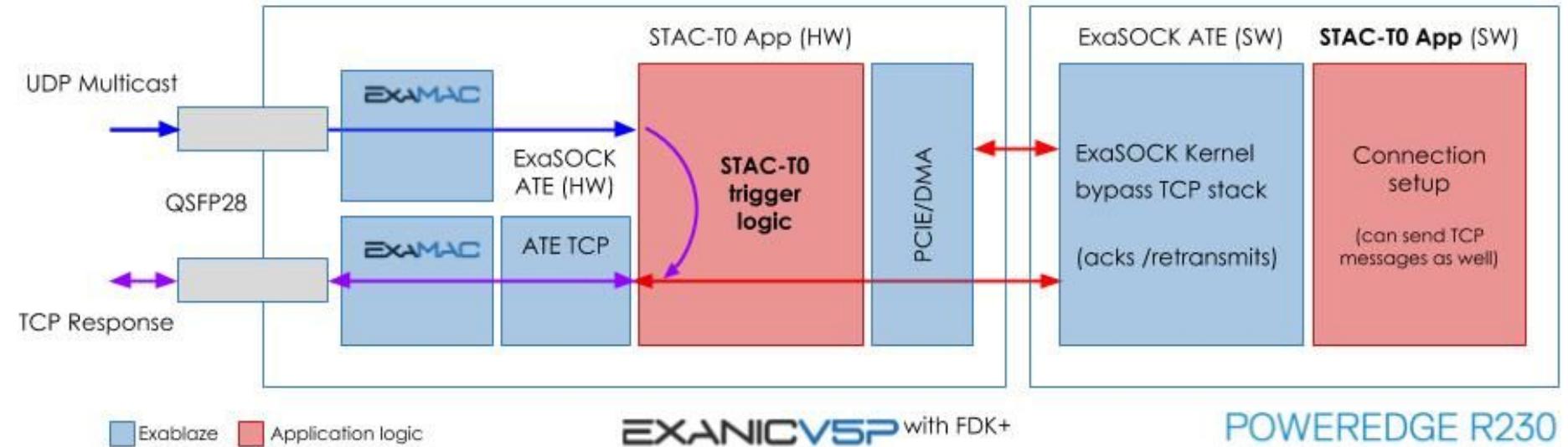
- Low(er) latency PCS/MAC
- PCIe DMA Engine
- HW Accelerated TCP Engine (ATE)
- Software, drivers, tooling
- Extensive examples.

Firmware Development Kit Expansion Pack



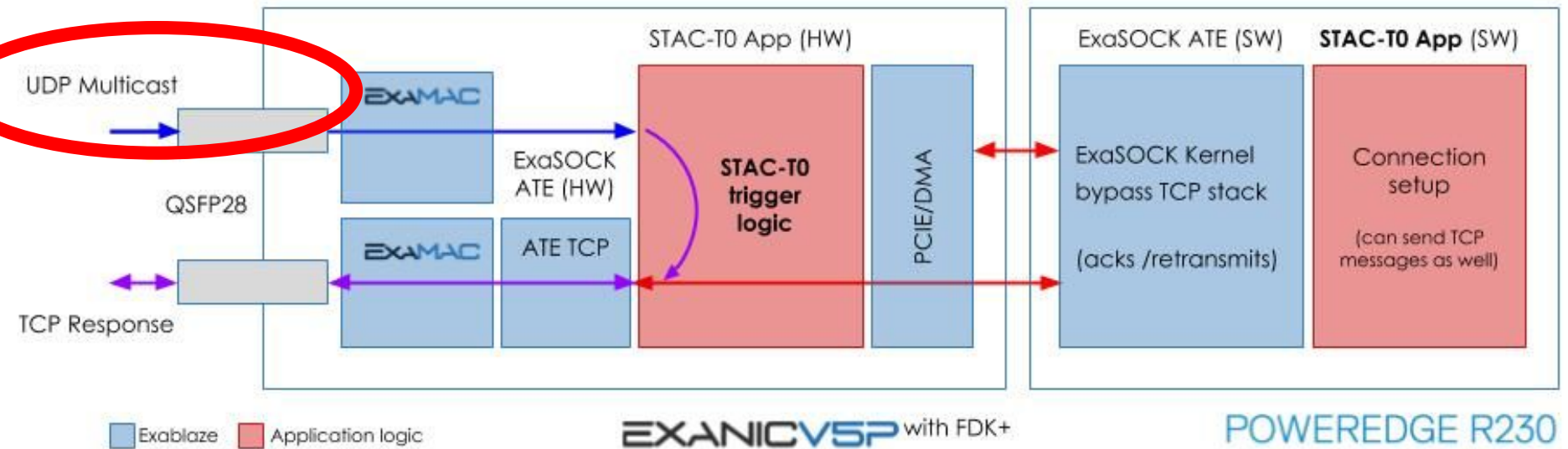
- **Low(er) latency PCS/MAC**
- PCIe DMA Engine
- **HW Accelerated TCP Engine (ATE)**
- Software, drivers, tooling
- Extensive examples.

STAC-T0 (tick-to-trade)



STAC-T0 (tick-to-trade)

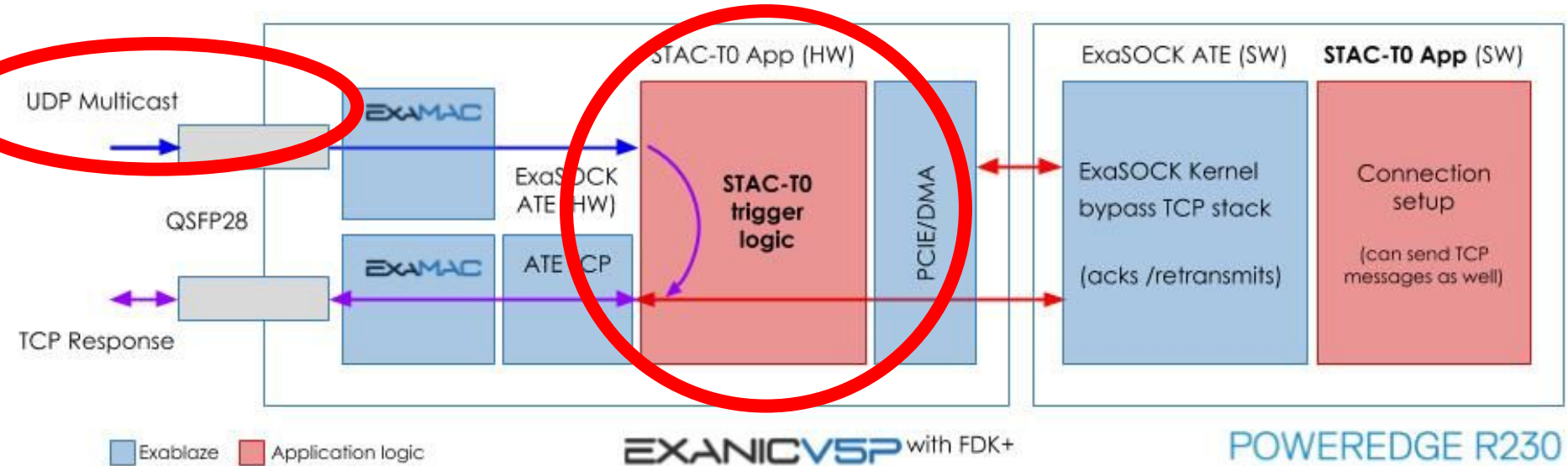
UDP Multicast in



STAC-T0 (tick-to-trade)

UDP Multicast in

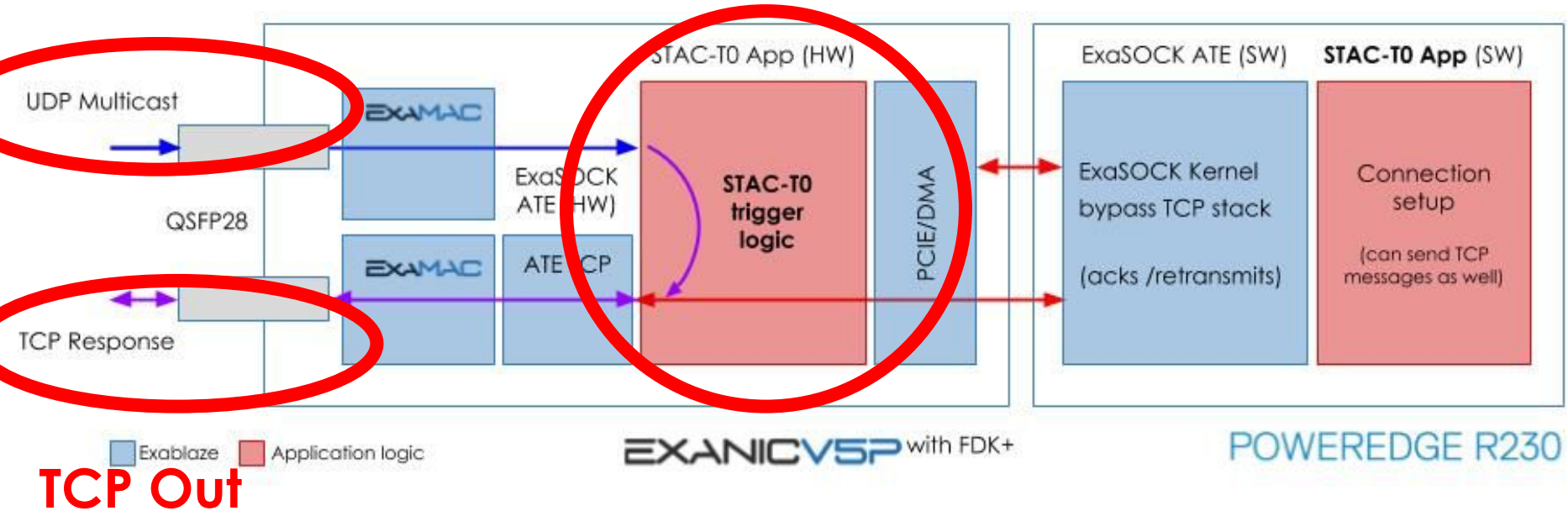
Trigger logic



STAC-T0 (tick-to-trade)

UDP Multicast in

Trigger logic



STAC-T0 (tick-to-trade)

31ns*

Min. actionable latency

* Subject to final validation

STAC-T0.β1.*.ACTIONABLE.MIN)

STAC-T0 (tick-to-trade)

31ns*

Min. actionable latency

* Subject to final validation

STAC-T0.β1.*.ACTIONABLE.MIN)

STAC-T0 (tick-to-trade)

31ns*

Min. actionable latency

* Subject to final validation

STAC-T0.β1.*.ACTIONABLE.MIN)

Possible reasons

Possible reasons

1. STAC can't measure things

Possible reasons

1. ~~STAC can't measure things~~
2. It's harder than it looks

Possible reasons

1. ~~STAC can't measure things~~

2. **It's harder than it looks**

Enter the Picosecond

Problem:

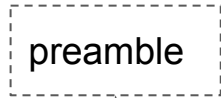
When did a field in my
packet arrive?

In an ideal world...

An Ethernet frame:

In an ideal world...

An Ethernet frame:



7B of
0101010

In an ideal world...

An Ethernet frame:



Start of frame delimiter
(1B)

In an ideal world...

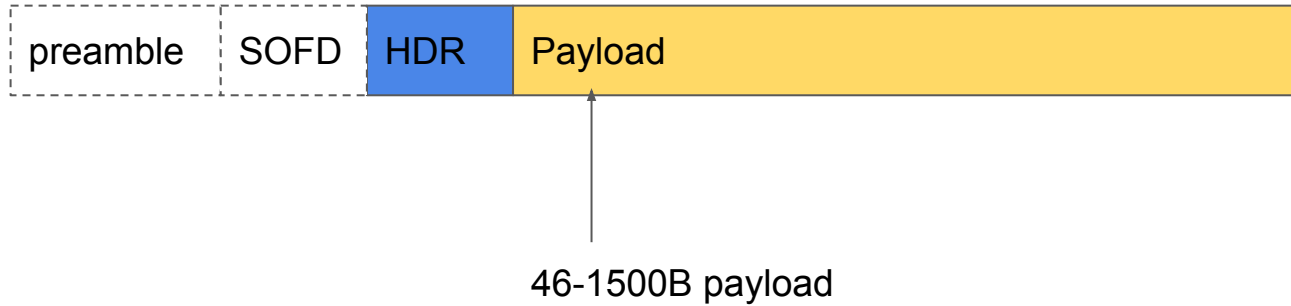
An Ethernet frame:



14B - SRC/DST MAC
address, ether type

In an ideal world...

An Ethernet frame:



In an ideal world...

An Ethernet frame:



4B (32b) frame check sequence

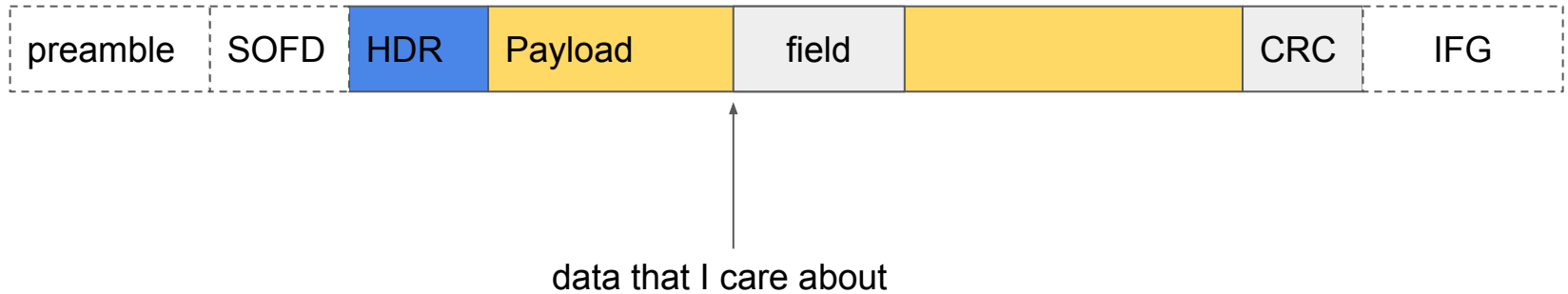
In an ideal world...

An Ethernet frame:



In an ideal world...

An Ethernet frame:



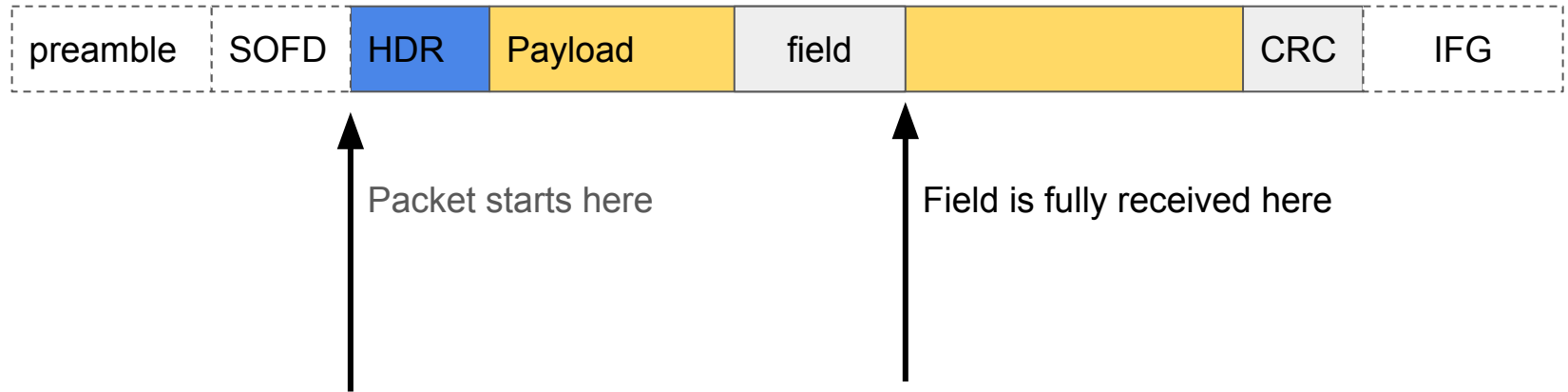
In an ideal world...

An Ethernet frame:



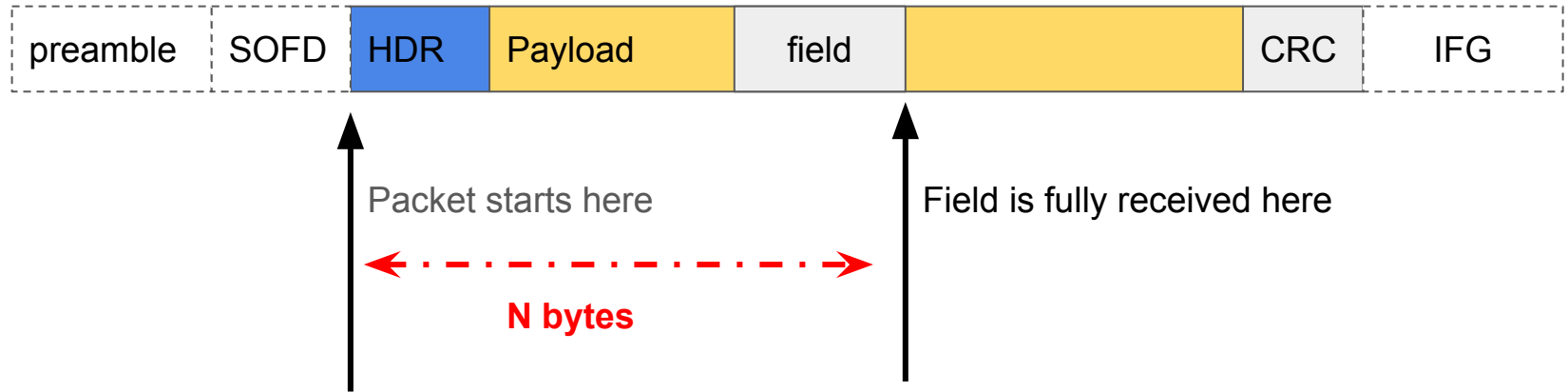
In an ideal world...

An Ethernet frame:




In an ideal world...

An Ethernet frame:



Ideal calculation

Bytes offset into the packet


$$\text{Delay} = N$$

Ideal calculation

$$\text{Delay} = N \times 8$$

Convert to bits



Ideal calculation

Line rate (G-bits per second)



$$\text{Delay} = N \times 8 \times 1 / 10 \text{ Gb/s}$$

Ideal calculation

Convert to picoseconds (10^{-12})

$$\text{Delay} = N \times 8 \times 1 / 10 \text{ Gb/s} \times 1/\text{ps}$$


Ideal calculation

$$\text{Delay} = N \times 8 \times 100$$

Cancels out



Ideal calculation

simplifies


$$\text{Delay} = N \times 800$$

Ideal calculation - Example

42B

Ideal calculation - Example

Delay = 42B

Ideal calculation - Example

$$\text{Delay} = 42 \times 800 =$$

Ideal calculation - Example

$$\text{Delay} = 42 \times 800 = 33,600\text{ps}$$

Finished?

Ideal calculation

Line rate (G-bits per second)

$$\text{Delay} = N \times 8 \times \frac{1}{10} \text{ Gb/s}$$


Ideal calculation

Line rate (G-bits per second)

$$\text{Delay} = N \times 8 \times \frac{1}{10 \text{ Gb/s}}$$


Does 10G Ethernet really run at 10 Gb/s?

Ideal calculation

Line rate (G-bits per second)

$$\text{Delay} = N \times 8 \times \frac{1}{10 \text{ Gb/s}}$$


Does 10G Ethernet really run at 10 Gb/s? **No!**

- 10GbE *average* rate is 10Gb/s

- 10GbE *average* rate is 10Gb/s
- Uses 64b/66b encoding

- 10GbE *average* rate is 10Gb/s
- Uses 64b/66b encoding
- $66/64 \times 10 = 10.3125\text{Gb/s}$ ***bit rate***

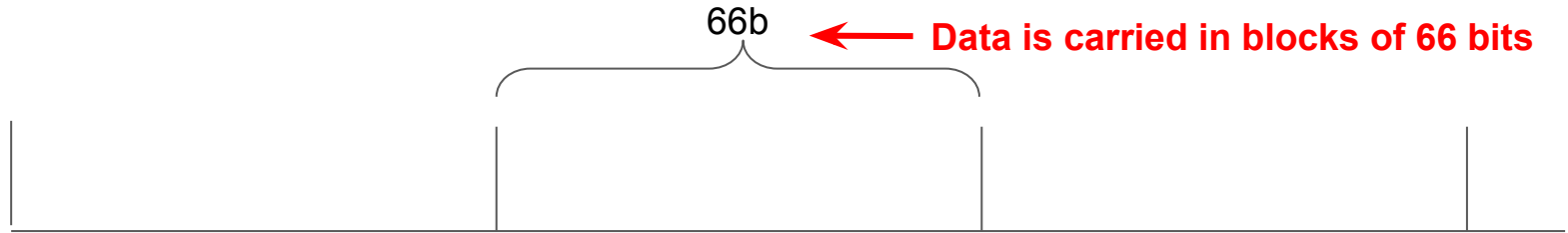
- 10GbE *average* rate is 10Gb/s
- **Uses 64b/66b encoding**
- $66/64 \times 10 = 10.3125\text{Gb/s}$ bit rate

64b/66b encoding in 1 minute or less...

64b/66b encoding in 1 minute* or less...

**not a STAC measurement*

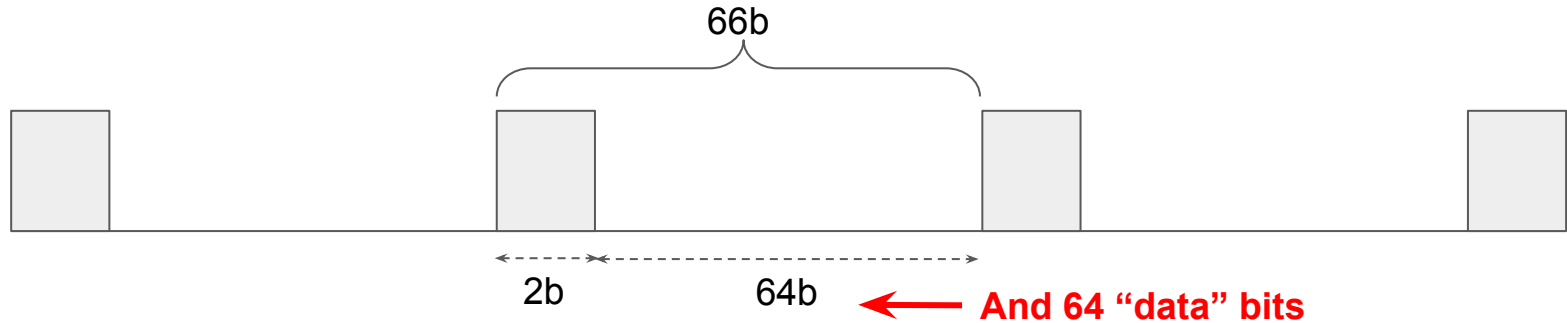
64b/66b encoding in 1 minute or less...



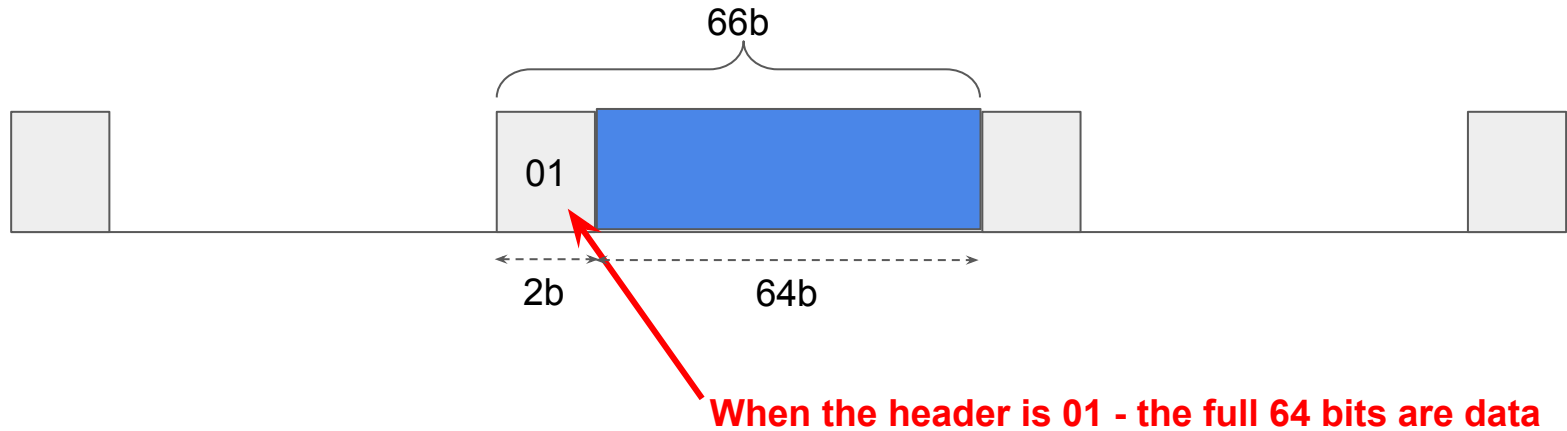
64b/66b encoding in 1 minute or less...



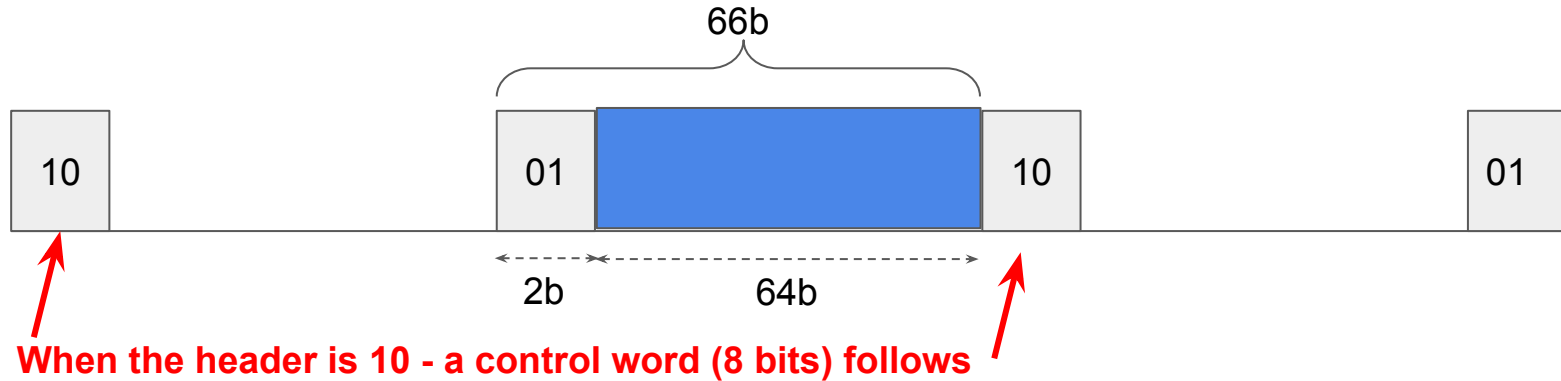
64b/66b encoding in 1 minute or less...



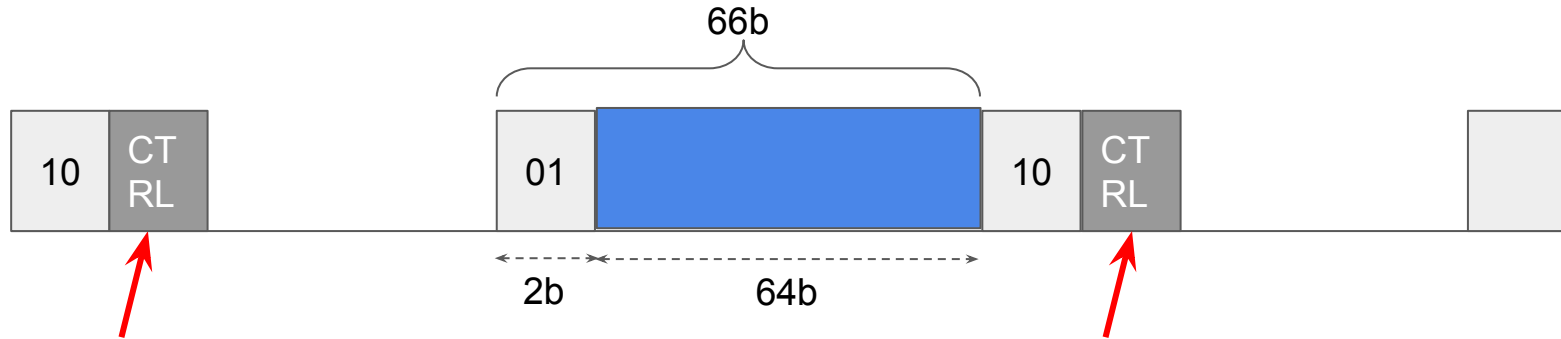
64b/66b encoding in 1 minute or less...



64b/66b encoding in 1 minute or less...

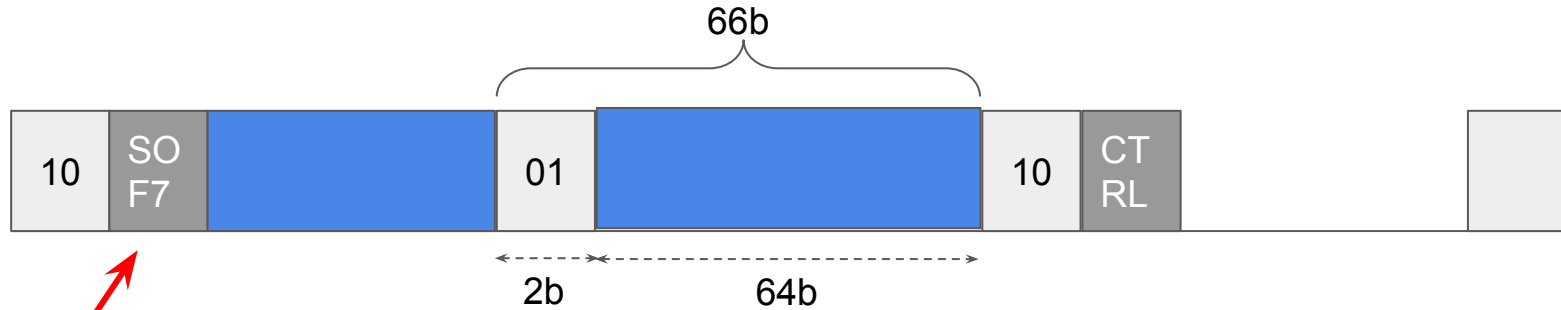


64b/66b encoding in 1 minute or less...



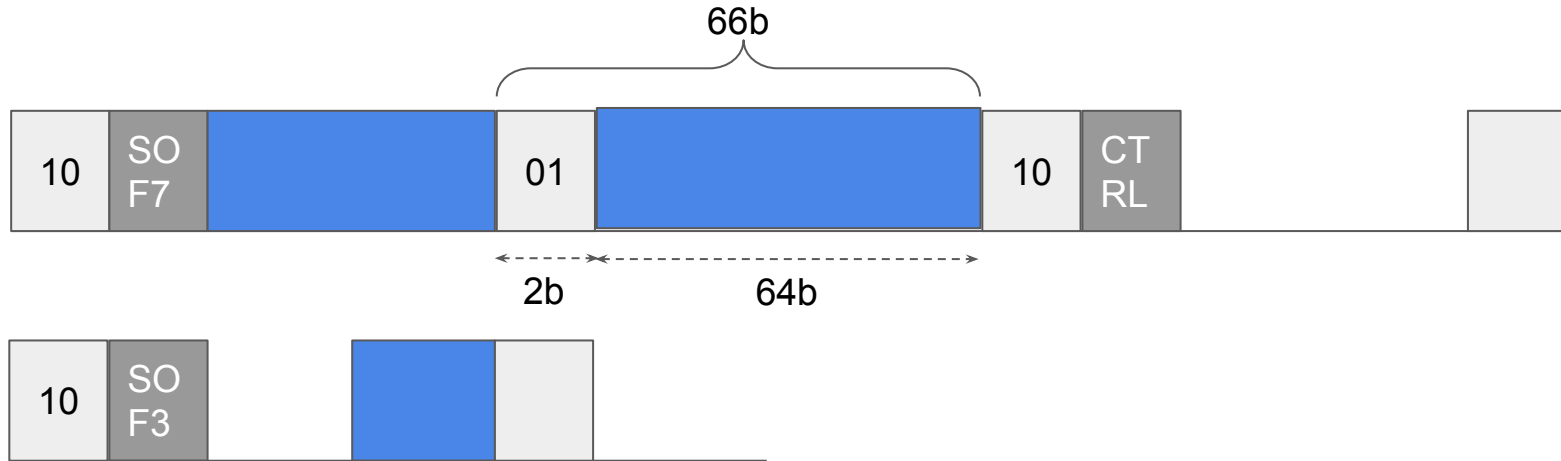
The control word can have a number of values (256), but the most important ones for this discussion are ...

64b/66b encoding in 1 minute or less...



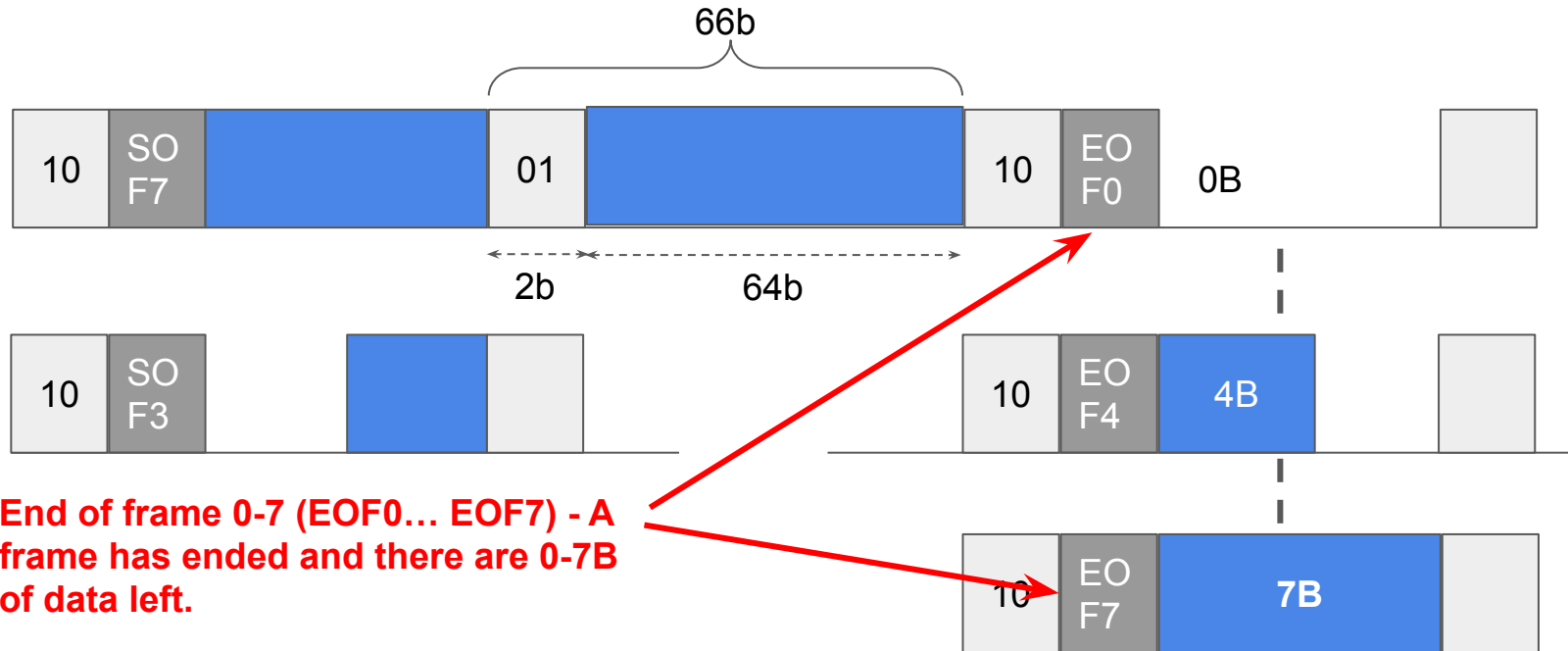
Start of frame 7 (SOF7) - A frame has started there are 7B (56b) of data

64b/66b encoding in 1 minute or less...



And... start of frame 3 (SOF3) - A frame has started there are 4B of idle, then 3B of data (24bits)

64b/66b encoding in 1 minute or less...



The making of an Ethernet frame



The making of an Ethernet frame



The making of an Ethernet frame



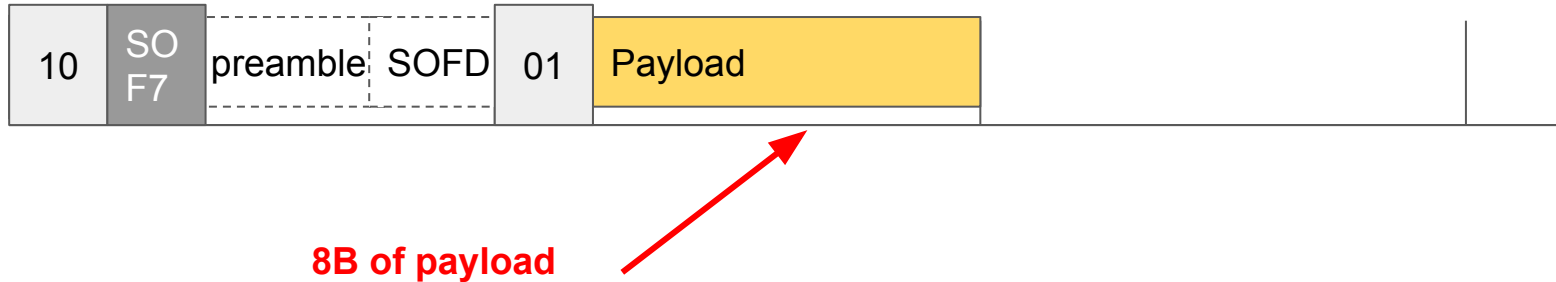
6B of preamble

YES! 6B!

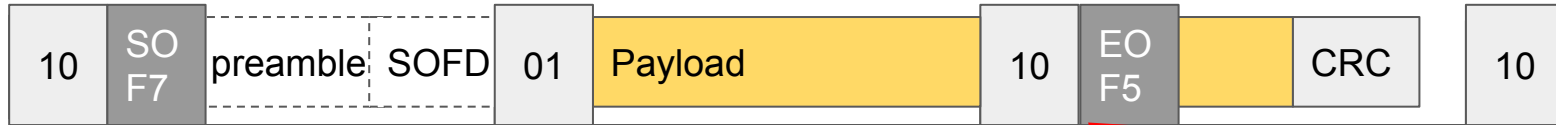
The making of an Ethernet frame



The making of an Ethernet frame

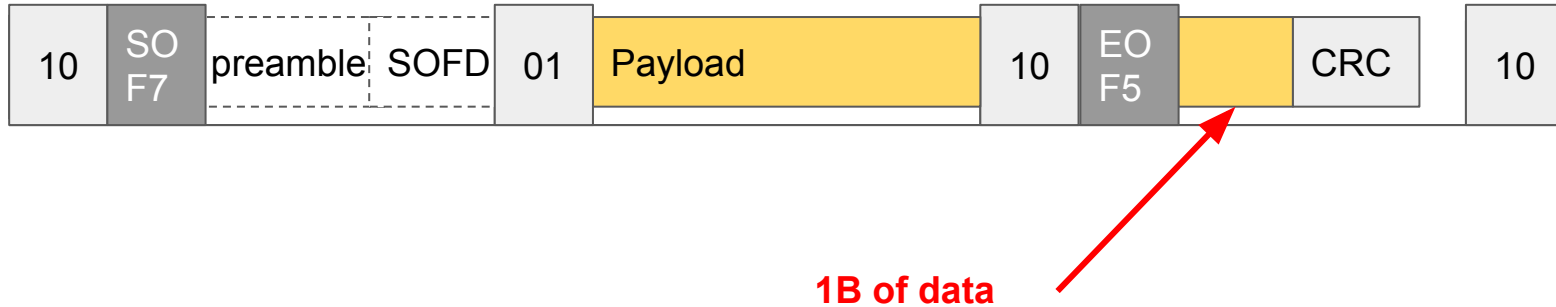


The making of an Ethernet frame

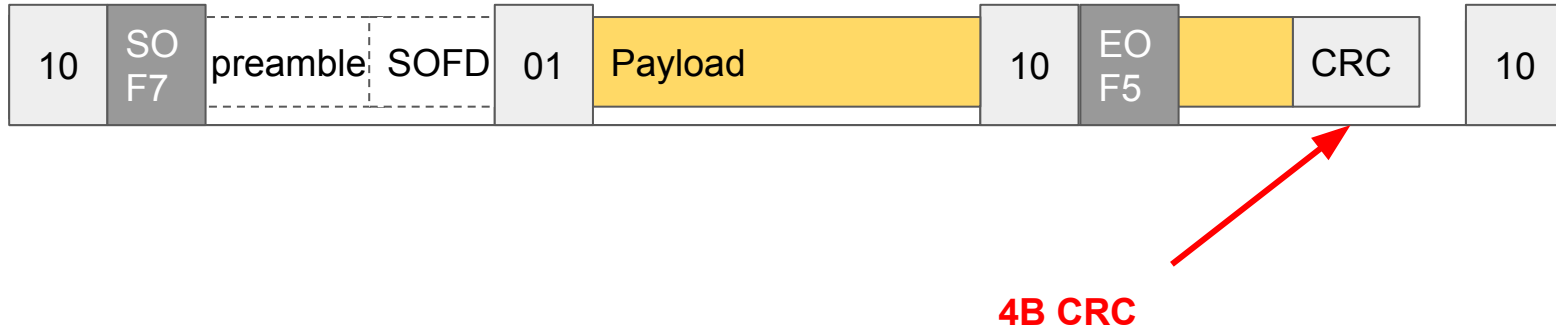


Control word

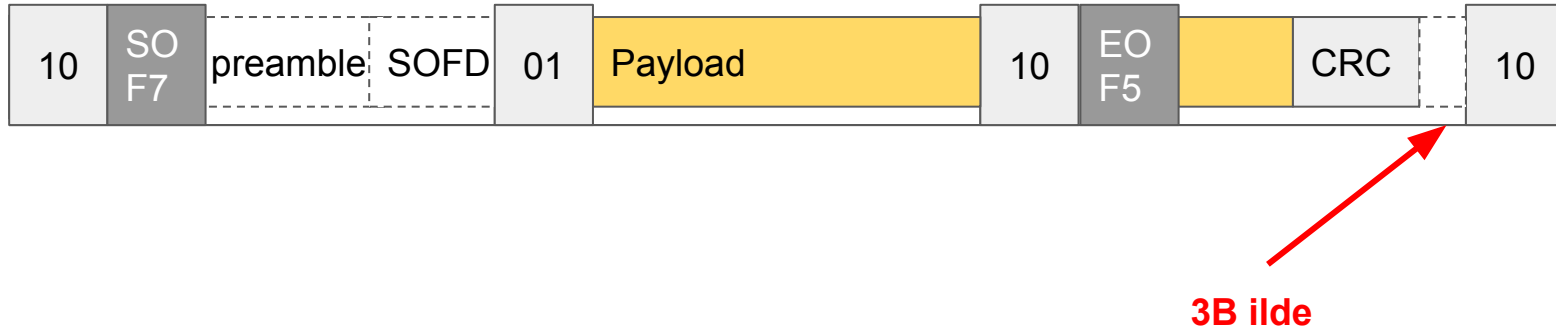
The making of an Ethernet frame



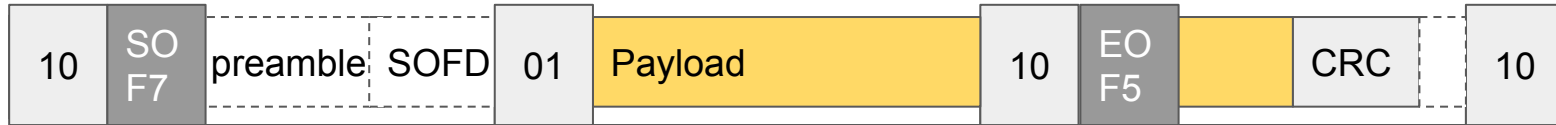
The making of an Ethernet frame



The making of an Ethernet frame



The making of an Ethernet frame



Possible reasons

1. ~~STAC can't measure things~~

2. It's harder than it looks

Where's the hard part??

Considering the encoding layer raises questions....

Considering the encoding layer raises questions....

- When does the frame start? And when it is timestamped?

Considering the encoding layer raises questions....

- When does the frame start? And when it it timestamped?
- When does the frame end?

Considering the encoding layer raises questions....

- When does the frame start? And when it it timestamped?
- When does the frame end?
- How long is the frame? (in bits and in picoseconds)

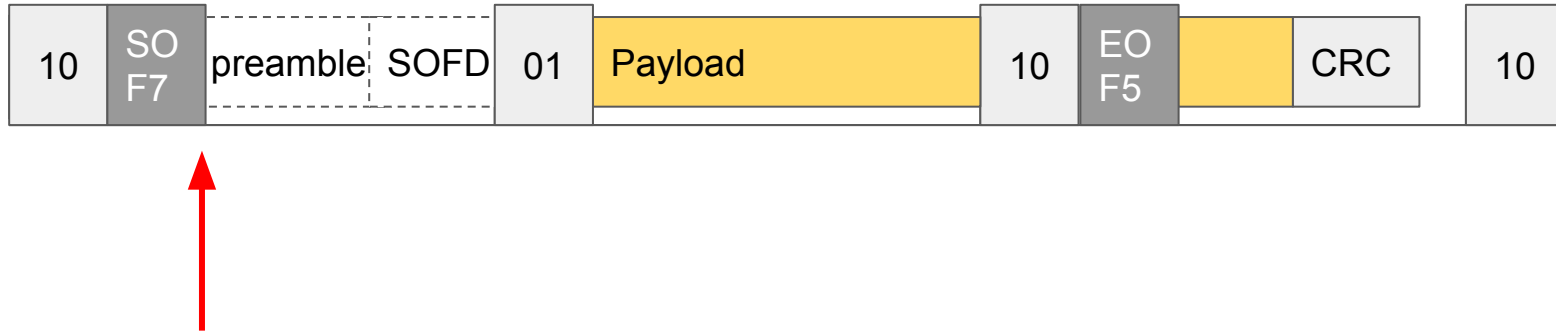
Considering the encoding layer raises questions....

- When does the frame start? And when it it timestamped?
- When does the frame end?
- How long is the frame? (in bits and in picoseconds)
- How far (ps) into the frame is an arbitrary offset?

Encoding layer questions....

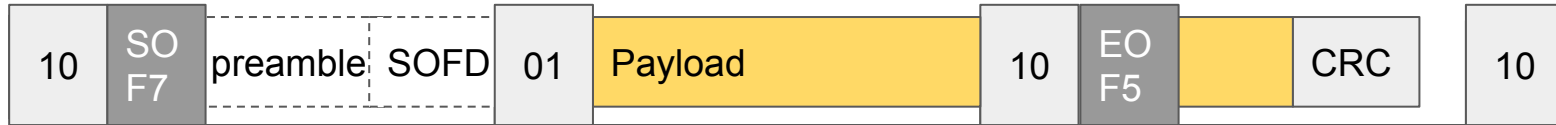
- **When does a frame start? (and is timestamped?)**

When does a frame start?



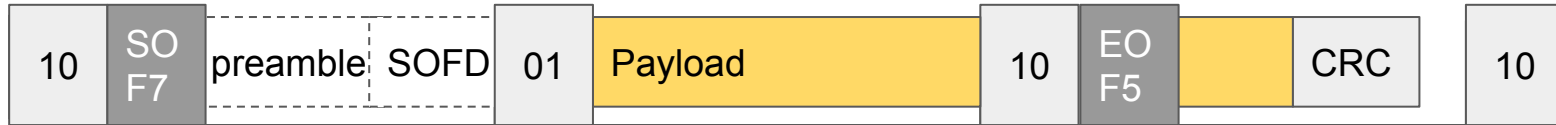
When the 64/66 SOF7 signal is found?

When does a frame start?



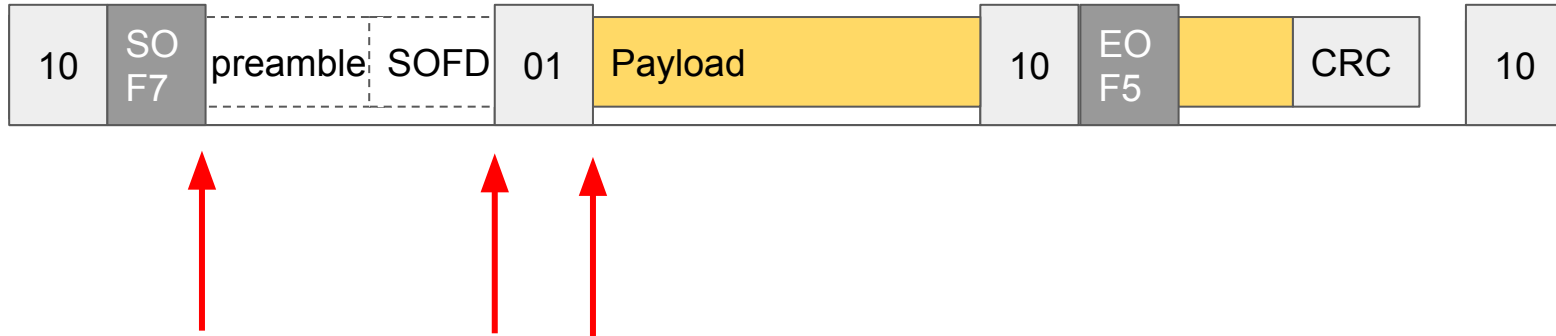
Or when the SOFD finishes?

When does a frame start?



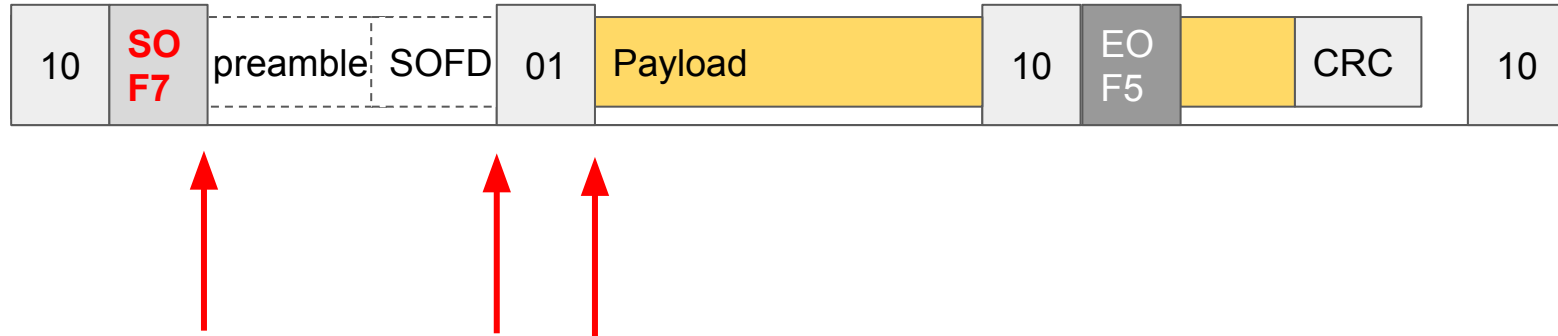
Or when the payload starts?

When does a frame start?



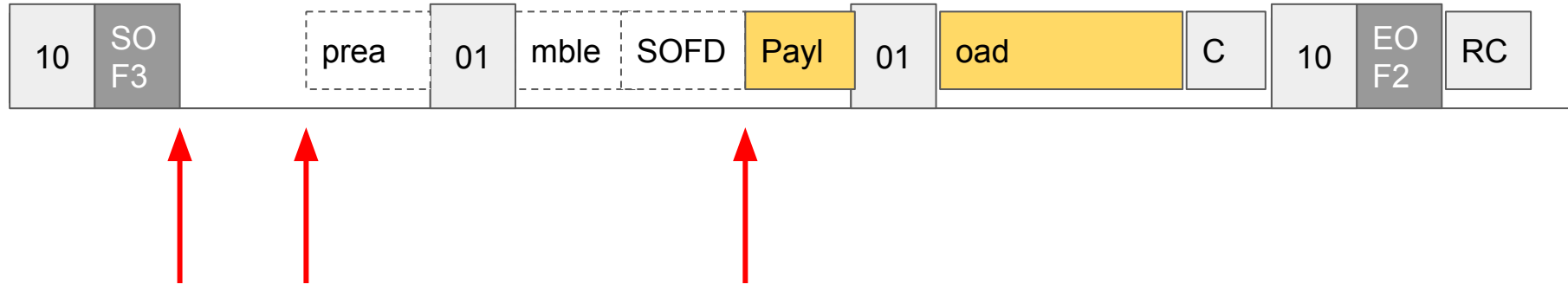
Or when the payload starts?

When does a frame start?



Or when the payload starts?

When does a frame start? And when is it timestamped

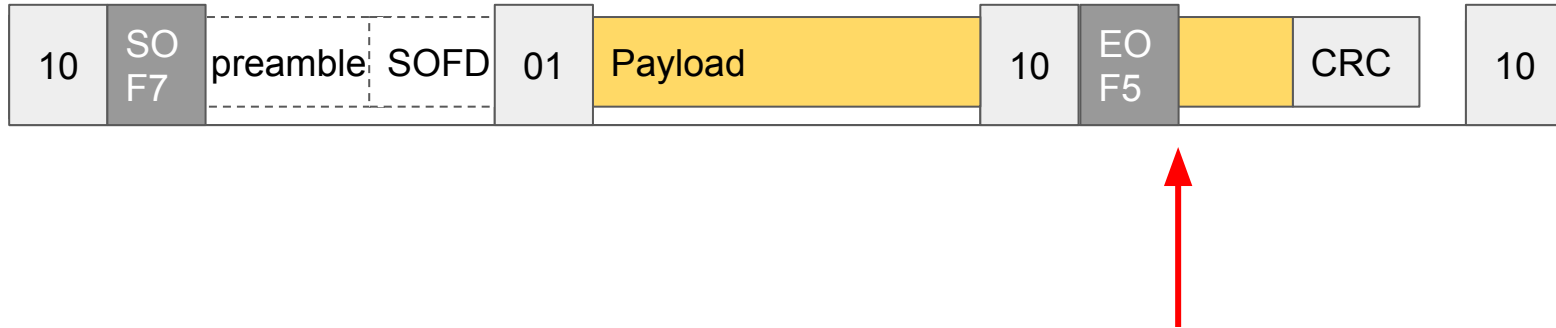


What about the SOF3 case?

Encoding layer questions....

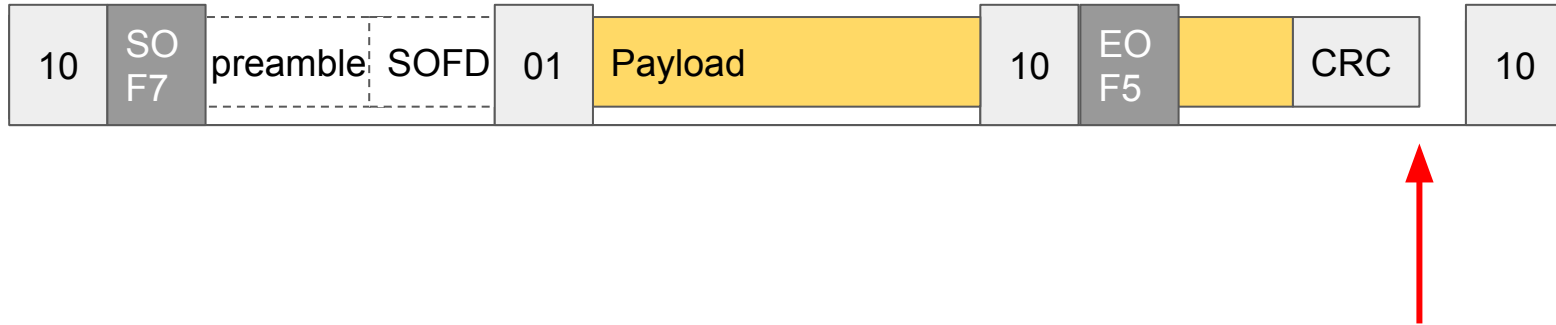
- When does the frame start? And when is it timestamped?
- **When does the frame end?**

When does a frame end?



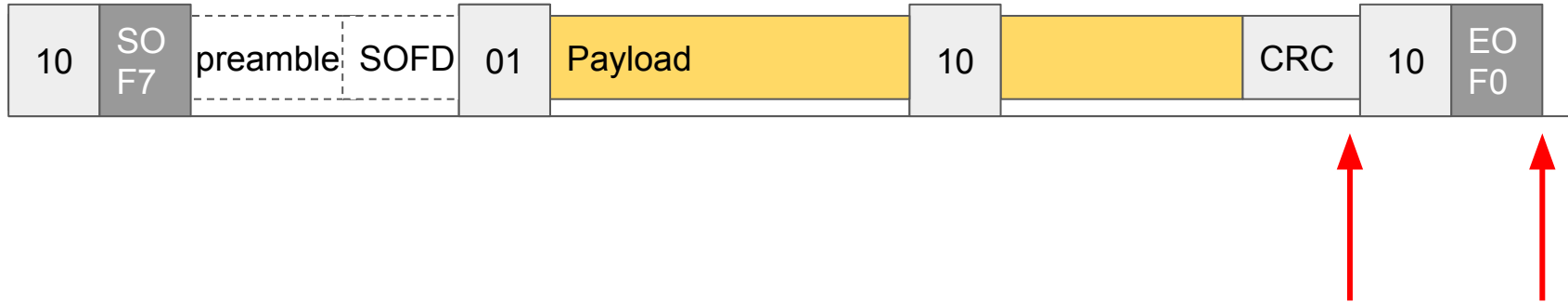
When the 64/66 EOF1-7 signal is found?

When does a frame end?



Or when the frame CRC arrives?

When does a frame end?

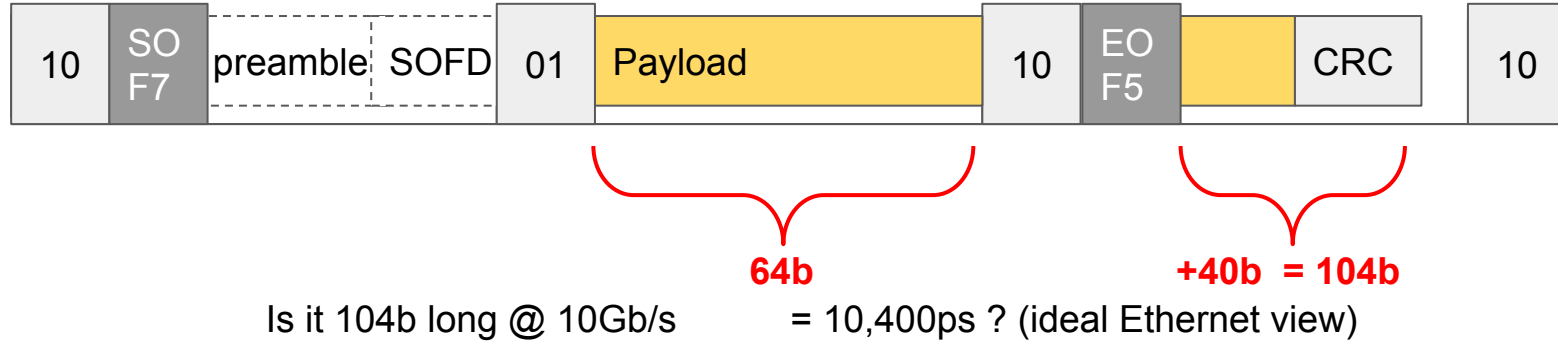


What about the EOF0 case?

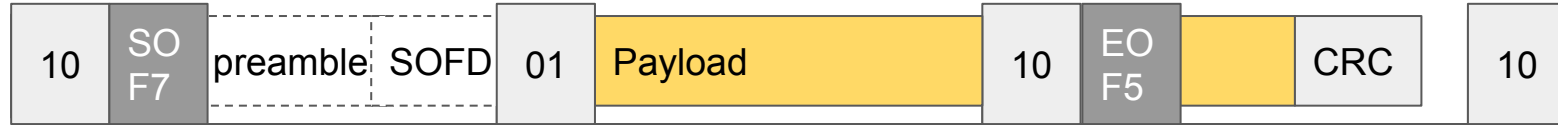
Encoding layer questions....

- When does the frame start? And when is it timestamped?
- When does the frame end?
- How long is a frame?

How long is the frame? (In bits / picoseconds)



How long is the frame? (In bits / picoseconds)



56b

Is it 104b long @ 10Gb/s

Is it 160b long @ 10Gb/s

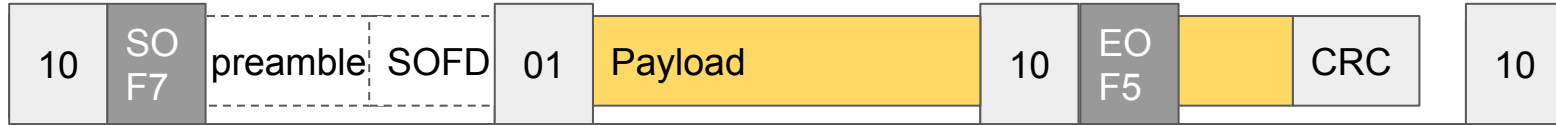
+64b

= 10,400ps ? (ideal Ethernet view)

= 16,000ps ? (ideal Ethernet view + preamble/SOFD)

+40b = 160b

How long is the frame? (In bits / picoseconds)



64b

Is it 104b long @ 10Gb/s

Is it 160b long @ 10Gb/s

Is it 168b long @ 10Gb/s

+64b

= 10,400ps ? (ideal Ethernet view)

= 16,000ps ? (ideal Ethernet view + preamble/SOFD)

= 16,800ps ? (ideal Ethernet view + XGMII preamble/SOFD)

+40b = 160b

How long is the frame? (In bits / picoseconds)

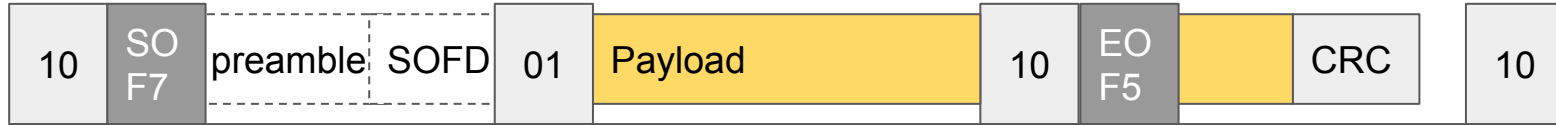


64

+2 +8 +40 = 114b

- Is it 104b long @ 10Gb/s = 10,400ps ? (ideal Ethernet view)
- Is it 160b long @ 10Gb/s = 16,000ps ? (ideal Ethernet view + preamble/SOFD)
- Is it 168b long @ 10Gb/s = 16,800ps ? (ideal Ethernet view + XGMII preamble/SOFD)
- Is it 114b long @ 10.3125Gb/s = 11,054ps ? (PCS[SOF7], ideal Ethernet view)

How long is the frame? (In bits / picoseconds)



- 56 +2 +64**
- Is it 104b long @ 10Gb/s = 10,400ps ? (ideal Ethernet view)
- Is it 160b long @ 10Gb/s = 16,000ps ? (ideal Ethernet view + preamble/SOFD)
- Is it 168b long @ 10Gb/s = 16,800ps ? (ideal Ethernet view + XGMII preamble/SOFD)
- Is it 114b long @ 10.3125Gb/s = 11,054ps ? (PCS[SOF7], Ethernet view)
- Is it 172b long @ 10.3125Gb/s = 16,679ps ? (PCS[SOF7], Ethernet view + preamble/SOFD)
- +2 +8 +40 = 172b**

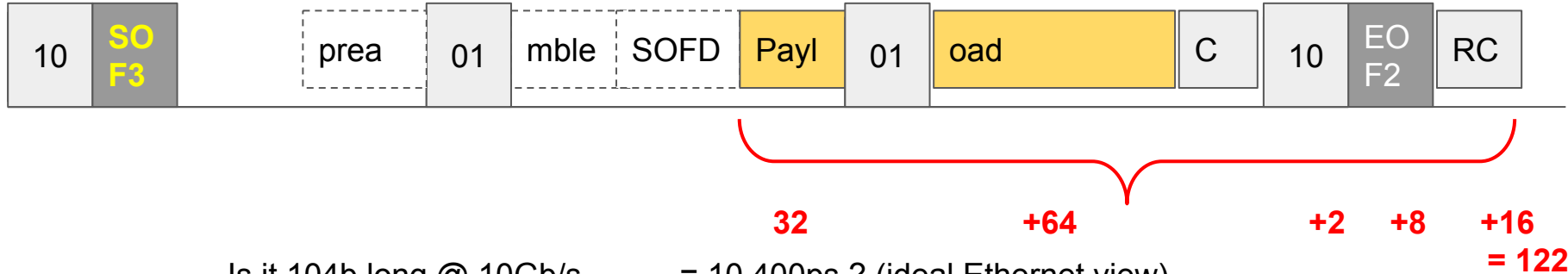
How long is the frame? (In bits / picoseconds)



- Is it 104b long @ 10Gb/s = 10,400ps ? (ideal Ethernet view)
- Is it 160b long @ 10Gb/s = 16,000ps ? (ideal Ethernet view + preamble/SOFD)
- Is it 168b long @ 10Gb/s = 16,800ps ? (ideal Ethernet view + XGMII preamble/SOFD)
- Is it 114b long @ 10.3125Gb/s = 11,054ps ? (PCS[SOF7], Ethernet view)
- Is it 172b long @ 10.3125Gb/s = 16,679ps ? (PCS[SOF7], Ethernet view + preamble/SOFD)

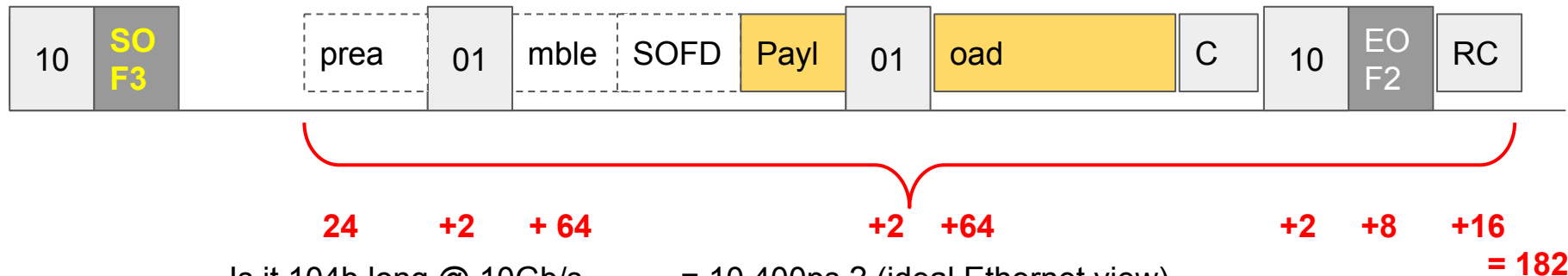
What about the SOF3 case?

How long is the frame? (In bits / picoseconds)



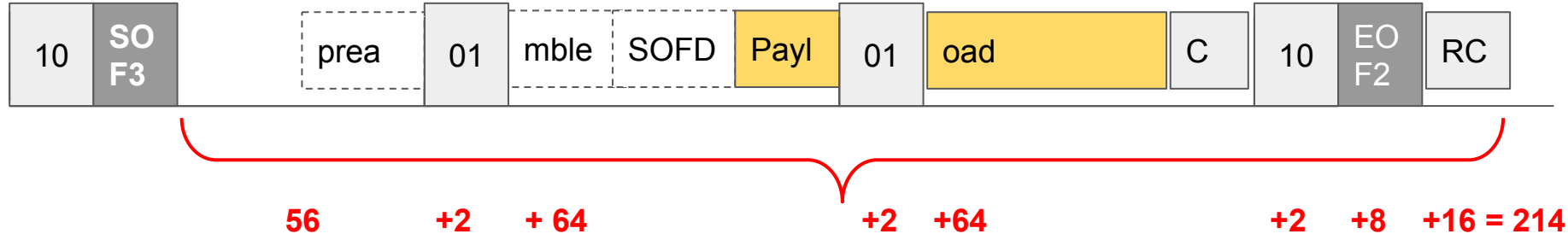
- Is it 104b long @ 10Gb/s = 10,400ps ? (ideal Ethernet view)
- Is it 160b long @ 10Gb/s = 16,000ps ? (ideal Ethernet view + preamble/SOFD)
- Is it 168b long @ 10Gb/s = 16,800ps ? (ideal Ethernet view + XGMII preamble/SOFD)
- Is it 114b long @ 10.3125Gb/s = 11,054ps ? (PCS[SOF7], Ethernet view)
- Is it 172b long @ 10.3125Gb/s = 16,679ps ? (PCS[SOF7], Ethernet view + preamble/SOFD)
- Is it 122b long @ 10.3125Gb/s = 11,830ps ? (PCS[SOF3], Ethernet view)

How long is the frame? (In bits / picoseconds)



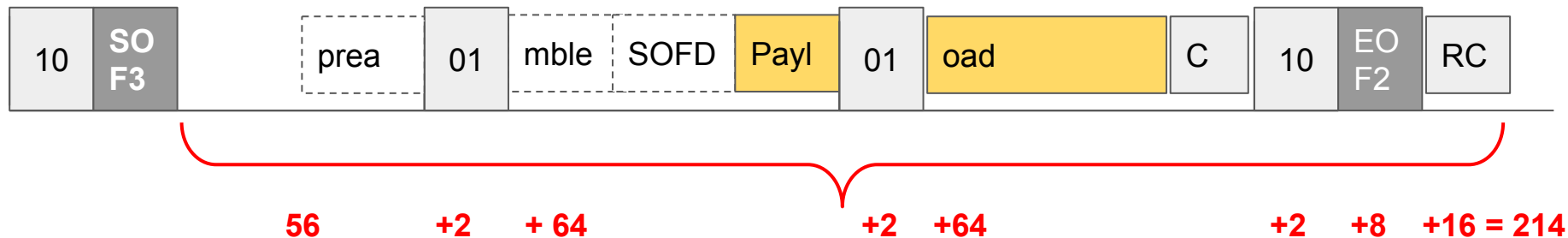
- Is it 104b long @ 10Gb/s = 10,400ps ? (ideal Ethernet view)
- Is it 160b long @ 10Gb/s = 16,000ps ? (ideal Ethernet view + preamble/SOFD)
- Is it 168b long @ 10Gb/s = 16,800ps ? (ideal Ethernet view + XGMII preamble/SOFD)
- Is it 114b long @ 10.3125Gb/s = 11,054ps ? (PCS[SOF7], Ethernet view)
- Is it 172b long @ 10.3125Gb/s = 16,679ps ? (PCS[SOF7], Ethernet view + preamble/SOFD)
- Is it 122b long @ 10.3125Gb/s = 11,830ps ? (PCS[SOF3], Ethernet view)
- Is it 182b long @ 10.3125Gb/s = 17,648ps ? (PCS[SOF3], Ethernet view + preamble/SOFD)

How long is the frame? (In bytes / picoseconds)



- Is it 104b long @ 10Gb/s = 10,400ps ? (ideal Ethernet view)
- Is it 160b long @ 10Gb/s = 16,000ps ? (ideal Ethernet view + preamble/SOFD)
- Is it 168b long @ 10Gb/s = 16,800ps ? (ideal Ethernet view + XGMII preamble/SOFD)
- Is it 114b long @ 10.3125Gb/s = 11,054ps ? (PCS[SOF7], Ethernet view)
- Is it 172b long @ 10.3125Gb/s = 16,679ps ? (PCS[SOF7], Ethernet view + preamble/SOFD)
- Is it 122b long @ 10.3125Gb/s = 11,830ps ? (PCS[SOF3], Ethernet view)
- Is it 182b long @ 10.3125Gb/s = 17,648ps ? (PCS[SOF3], Ethernet view + preamble/SOFD)
- Is it 214b long @ 10.3125Gb/s = 20,752ps ? (PCS[SOF3/SOF7] Ethernet view)

How long is the frame? (In bytes / picoseconds)



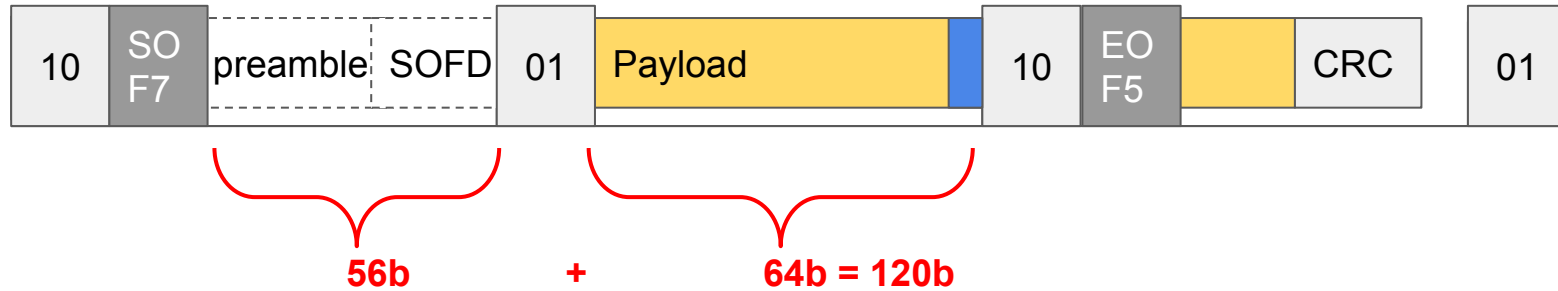
- Is it 104b long @ 10Gb/s = 10,400ps ? (ideal Ethernet view)
- Is it 160b long @ 10Gb/s = 16,000ps ? (ideal Ethernet view + preamble/SOFD)
- Is it 168b long @ 10Gb/s = 16,800ps ? (ideal Ethernet view + XGMII preamble/SOFD)
- Is it 114b long @ 10.3125Gb/s = 11,054ps ? (PCS[SOF7], Ethernet view)
- Is it 172b long @ 10.3125Gb/s = 16,679ps ? (PCS[SOF7], Ethernet view + preamble/SOFD)
- Is it 122b long @ 10.3125Gb/s = 11,830ps ? (PCS[SOF3], Ethernet view)
- Is it 182b long @ 10.3125Gb/s = 17,648ps ? (PCS[SOF3], Ethernet view + preamble/SOFD)
- Is it 214b long @ 10.3125Gb/s = 20,752ps ? (PCS[SOF3/SOF7] Ethernet view)

... and many more...

Encoding layer questions....

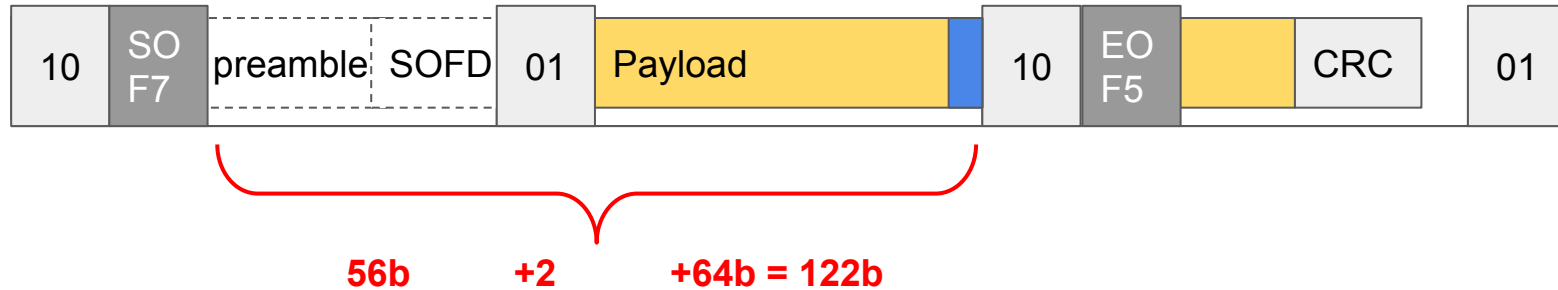
- When does the frame start? And when it it timestamped?
- When does the frame end?
- How long is the frame? (in bits and in picoseconds)
- **How far (ps) into the frame is an arbitrary offset?**

How far (ps) into the frame is an offset?



Is it 120b @ 10.000Gb/s = 12,000ps ?

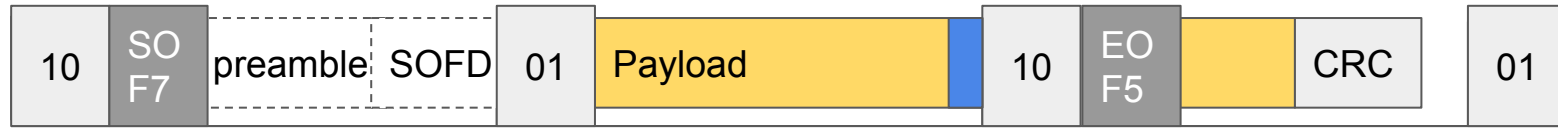
How far (ps) into the frame is an offset?



Is it 120b @ 10.000Gb/s = 12,000ps ?

Is it 122b @ 10.3125Gb/s = 11,830ps ?

How far (ps) into the frame is an offset?



56b

+2

+64b = 122b

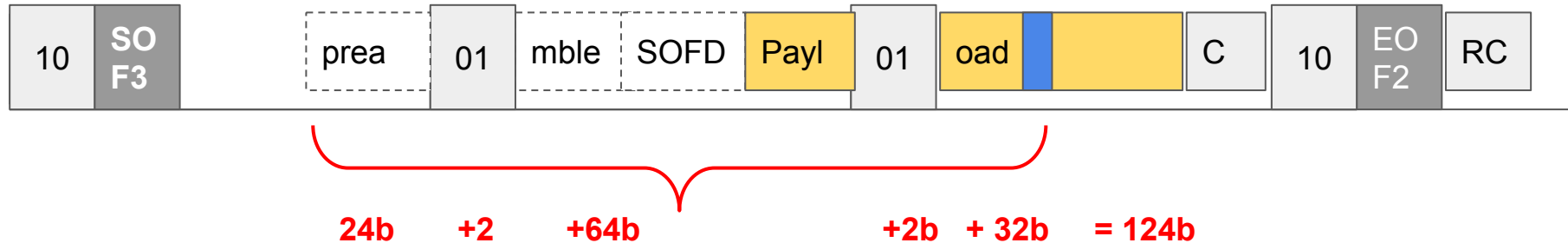
Is it 120b @ 10.000Gb/s = 12,000ps ?

Is it 122b @ 10.3125Gb/s = 11,830ps ?



Note: this is smaller than above!!!

How far (ps) into the frame is an offset?

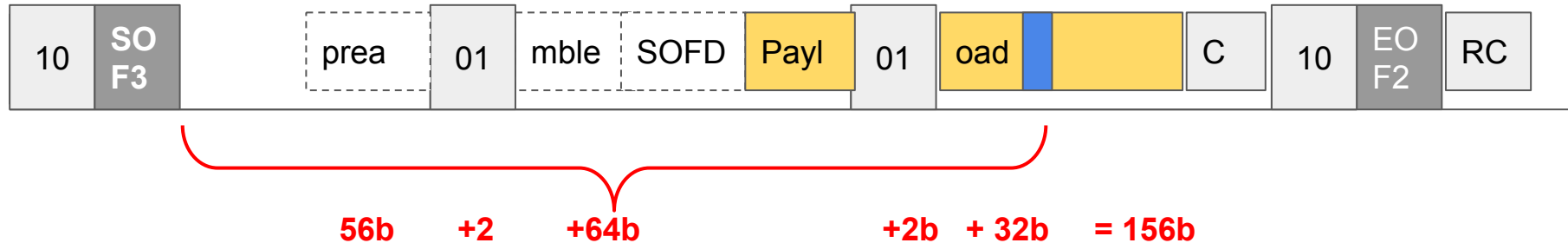


Is it 120b @ 10.000Gb/s = 12,000ps ?

Is it 122b @ 10.3125Gb/s = 11,830ps ?

Is it 124 @ 10.3125Gb/s = 12,024ps ?

How far (ps) into the frame is an offset?



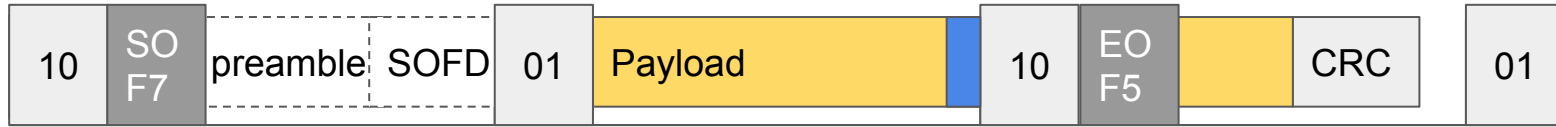
Is it 120b @ 10.000Gb/s = 12,000ps ?

Is it 122b @ 10.3125Gb/s = 11,830ps ?

Is it 124 @ 10.3125Gb/s = 12,024ps ?

Is it 156 @ 10.3125Gb/s = 15,127ps ?

How far (ps) into the frame is an offset?



64b = 120b

Is it 120b @ 10.000Gb/s = 12,000ps ?

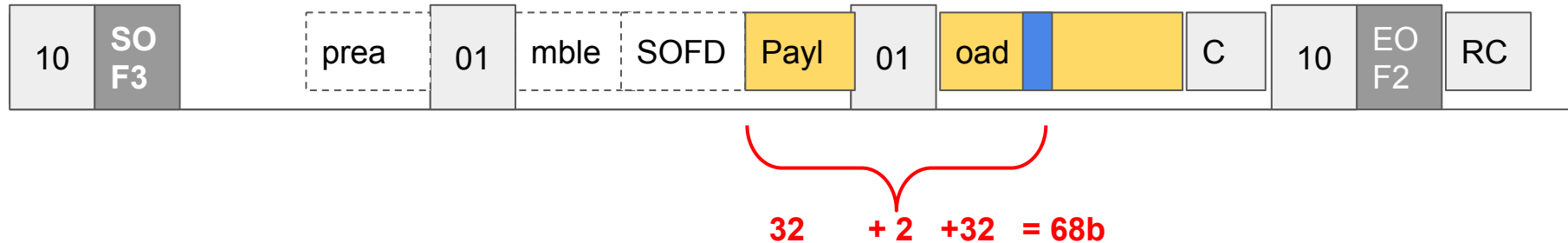
Is it 122b @ 10.3125Gb/s = 11,830ps ?

Is it 124 @ 10.3125Gb/s = 12,024ps ?

Is it 156 @ 10.3125Gb/s = 15,127ps ?

Is it 64b @ 10.3125Gb/s = 6,206ps ?

How far (ps) into the frame is an offset?



Is it 120b @ 10.000Gb/s = 12,000ps ?

Is it 122b @ 10.3125Gb/s = 11,830ps ?

Is it 124 @ 10.3125Gb/s = 12,024ps ?

Is it 156 @ 10.3125Gb/s = 15,127ps ?

Is it 64b @ 10.3125Gb/s = 6,206ps ?

Is it 68b @ 10.3125Gb/s = 6,594ps ?

Implications for uncertainty

1. Ethernet protocol has an **average** rate of 10Gb/s at layer 2, but PCS effects are visible at individual packet sizes.

Implications for uncertainty

1. Ethernet protocol has an **average** rate of 10Gb/s at layer 2, but PCS effects are visible at individual packet sizes.
2. At picosecond scales, PCS encodings must be taken into account and 10.3125Ghz must be used.

Implications for uncertainty

1. Ethernet protocol has an **average** rate of 10Gb/s at layer 2, but PCS effects are visible at individual packet sizes.
2. At picosecond scales, PCS encodings must be taken into account and 10.3125Ghz must be used.
3. Timestamps at PCS SOF3/7 and Ethernet layer SOFD have different absolute offsets*. Since both SOF3 and SOF7 may appear, these need to be accounted for.

Our recommendations....

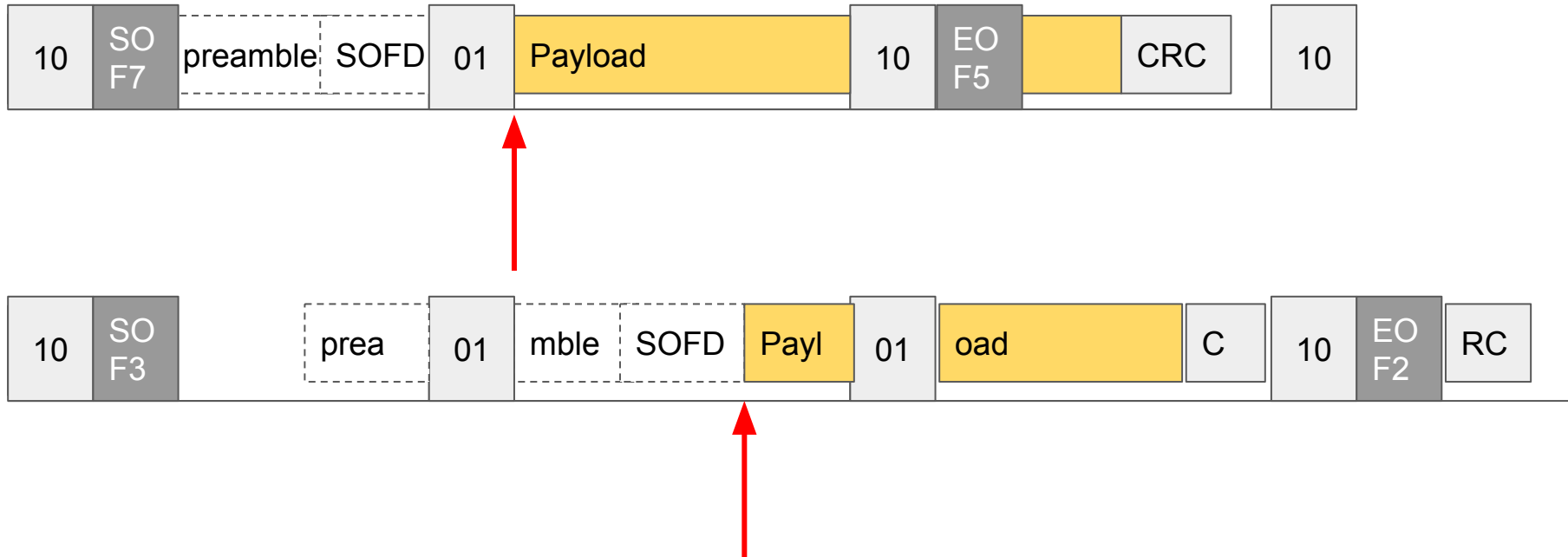
Our recommendations....

- When does the frame start?

Our recommendations....

- When does the frame start? ***At the start of the payload (DST MAC)***

When does a frame start? *At the start of the payload (DST MAC)*



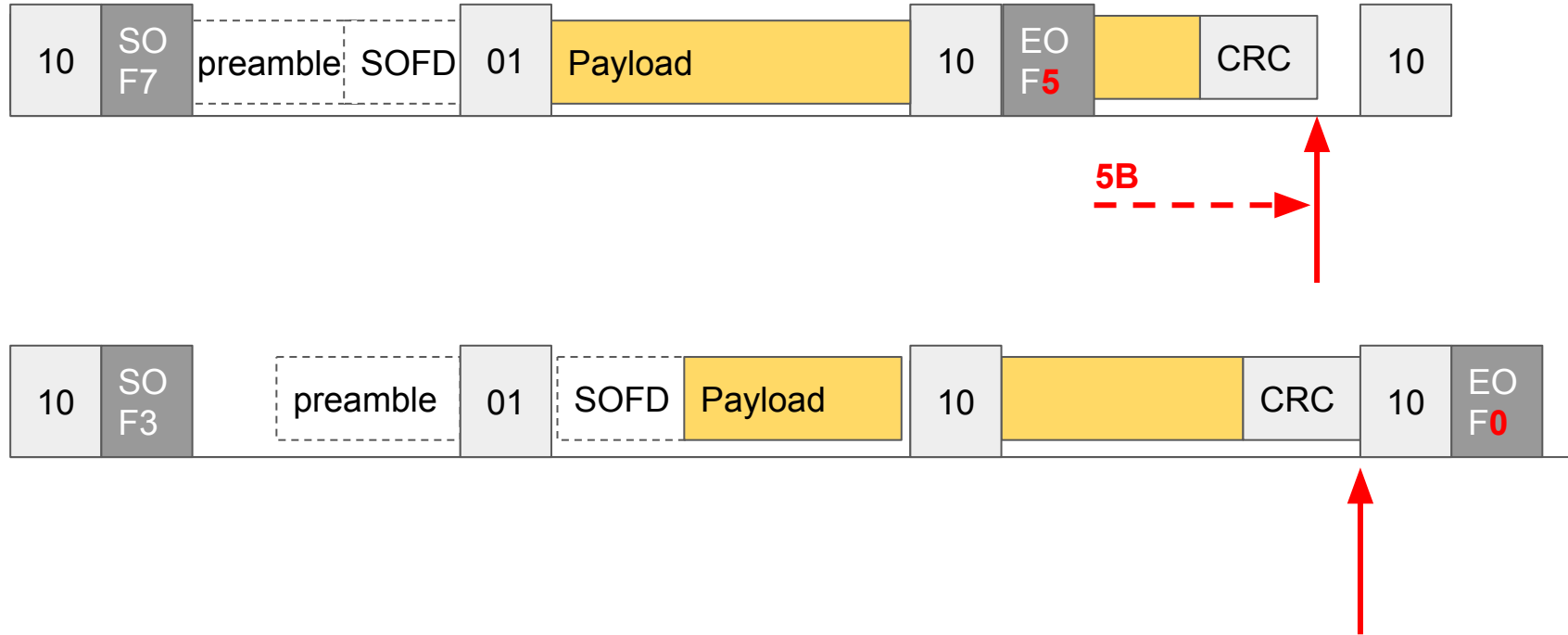
Our recommendations....

- When does the frame start? *At the start of the payload*
- When does the frame end?

Our recommendations....

- When does the frame start? *At the start of the payload*
- When does the frame end? ***At the end of the CRC***

When does a frame end? *At the end of the CRC*



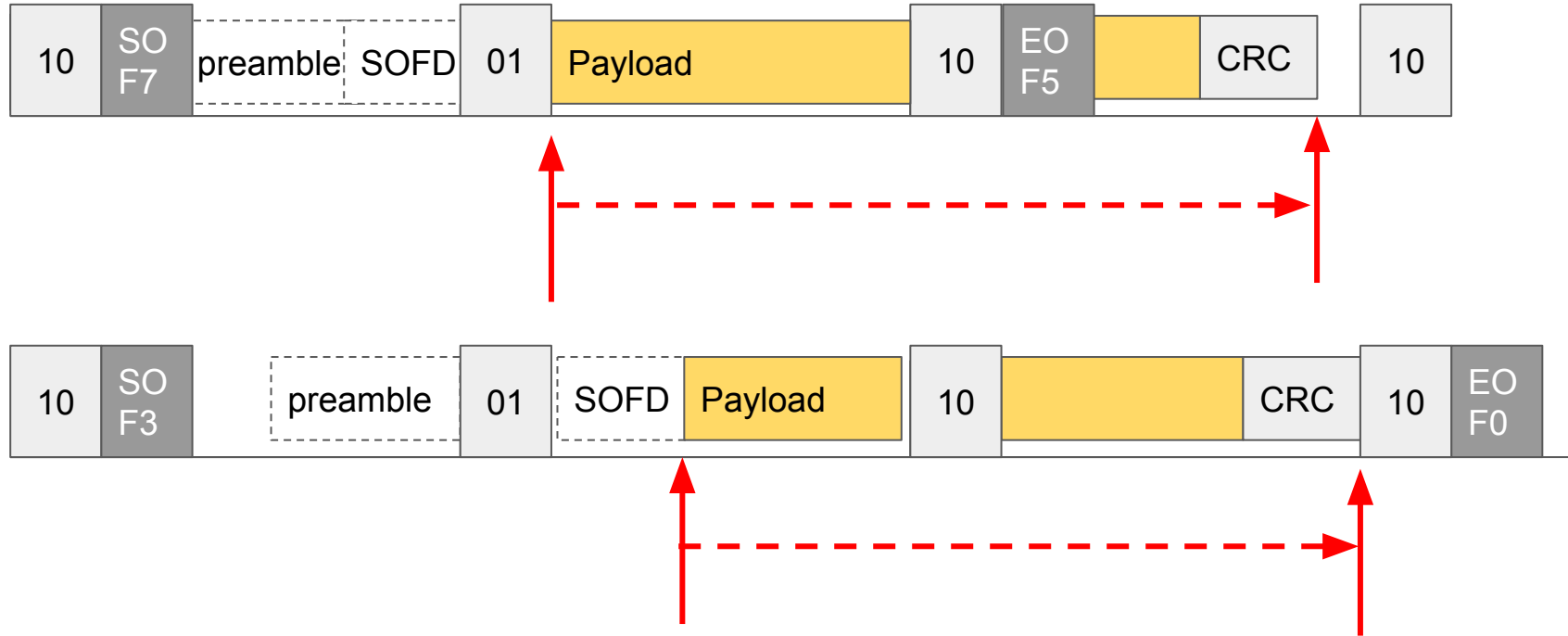
Our recommendations....

- When does the frame start? *At the start of the payload*
- When does the frame end? *At the end of the CRC*
- How long is the frame?

Our recommendations....

- When does the frame start? *At the start of the payload*
- When does the frame end? *At the end of the CRC*
- How long is the frame? **(CRC - payload) @ 10.3125G**

How long is the frame? *(CRC - payload) @ 10.3125Gbs*



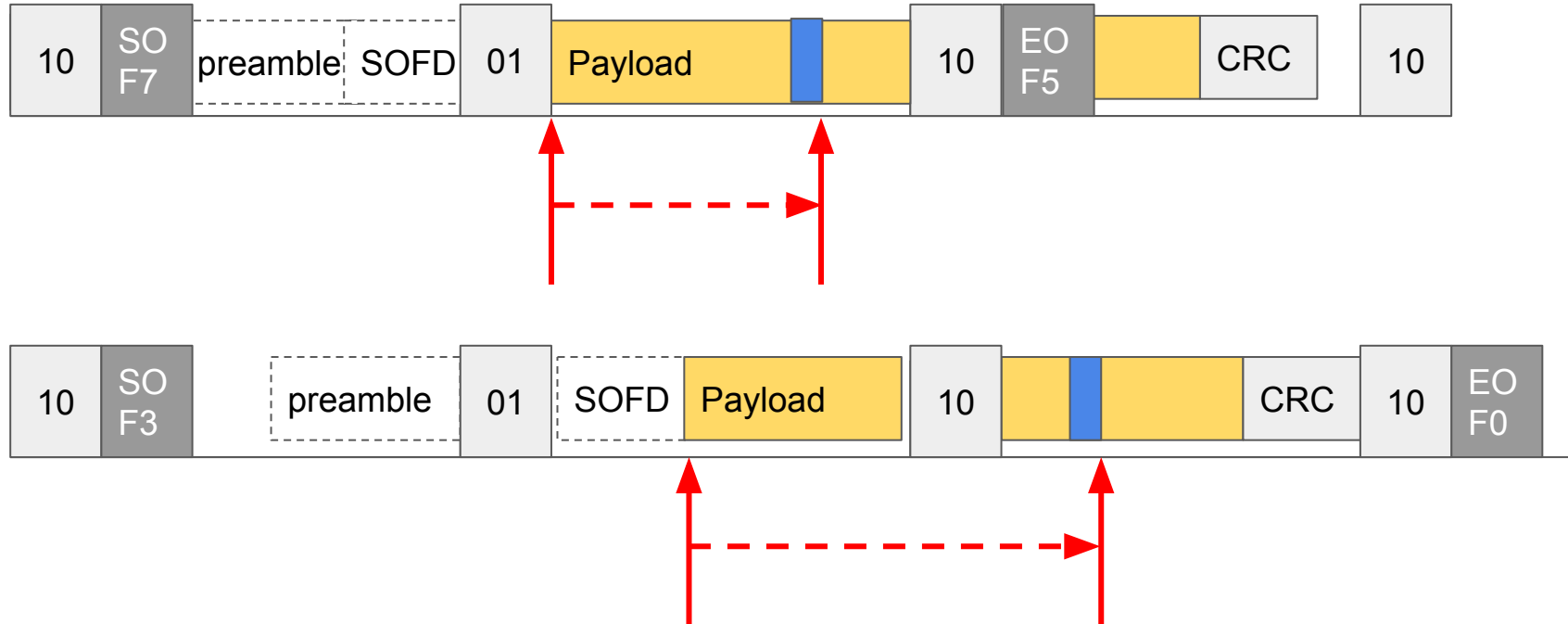
Our recommendations....

- When does the frame start? At the start of the payload
- When does the frame end? At the end of the CRC
- How long is the frame? (CRC - payload) @ 10.3125G
- How far (ps) is an offset?

Our recommendations....

- When does the frame start? At the start of the payload
- When does the frame end? At the end of the CRC
- How long is the frame? (CRC - payload) @ 10.3125G
- How far (ps) is an offset? **(bit offset - payload) @ 10.3125G**

How far is an offset?? *(bit offset - payload) @ 10.3125G*



Our recommendations!

- When does the frame start? **At the start of the payload**
- When does the frame end? **At the end of the CRC**
- How long is the frame? **(CRC - payload) @ 10.3125G**
- How far (ps) is an offset? **(bit offset - payload) @ 10.3125G**

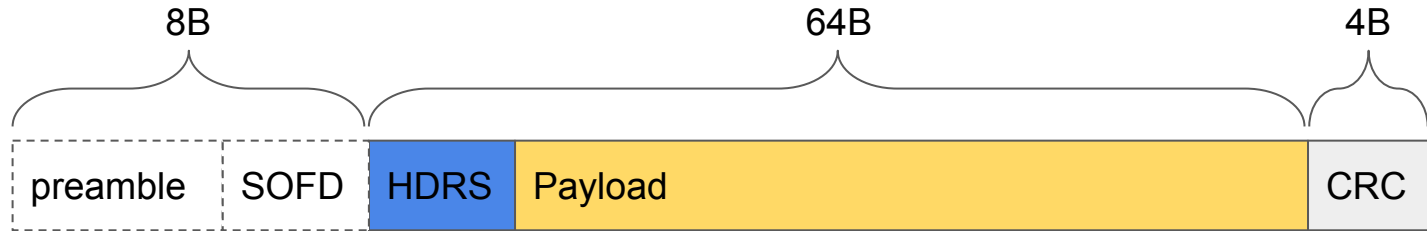
Worked example

1. Message is **64**B long, excluding FCS (4B)

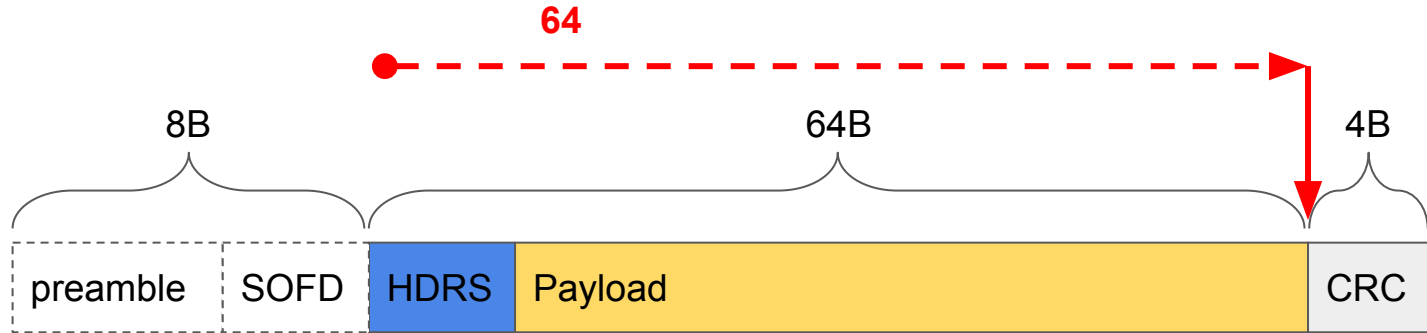
Worked example

1. Message is 64B long, excluding FCS (4B)
- 2. How long is the frame? (in ps)**

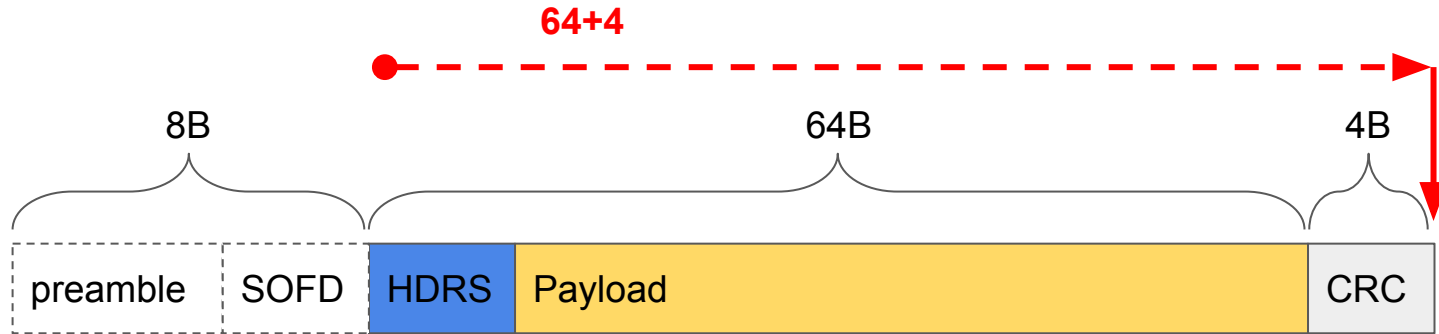
Ideal view of 64B Frame



Ideal view of 64B Frame



Ideal view of 507B Fame



Ideal calculation

$$\text{Delay} = N \times 800$$

Ideal calculation

$$\text{Delay} = 68 \times 800$$

Ideal calculation

$$\text{Delay} = 68 \times 800 = 54,400\text{ps}$$

Ideal calculation

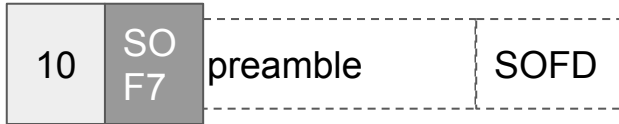
$$\text{Delay} = 68 \times 800 = 54,400\text{ps}$$

Taking into account PCS layer effects...

PCS **SOF7** view of a 68B fame:

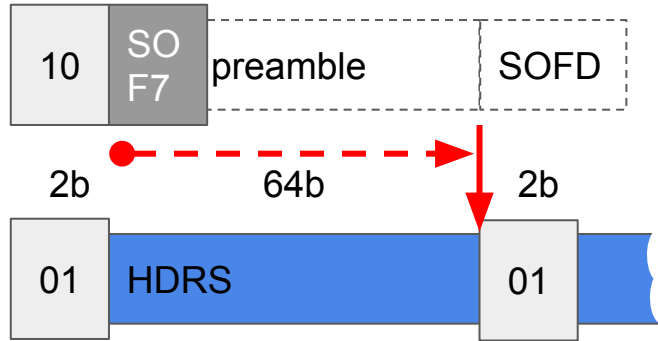
Taking into account PCS layer effects...

PCS **SOF7** view of a 68B fame:



Taking into account PCS layer effects...

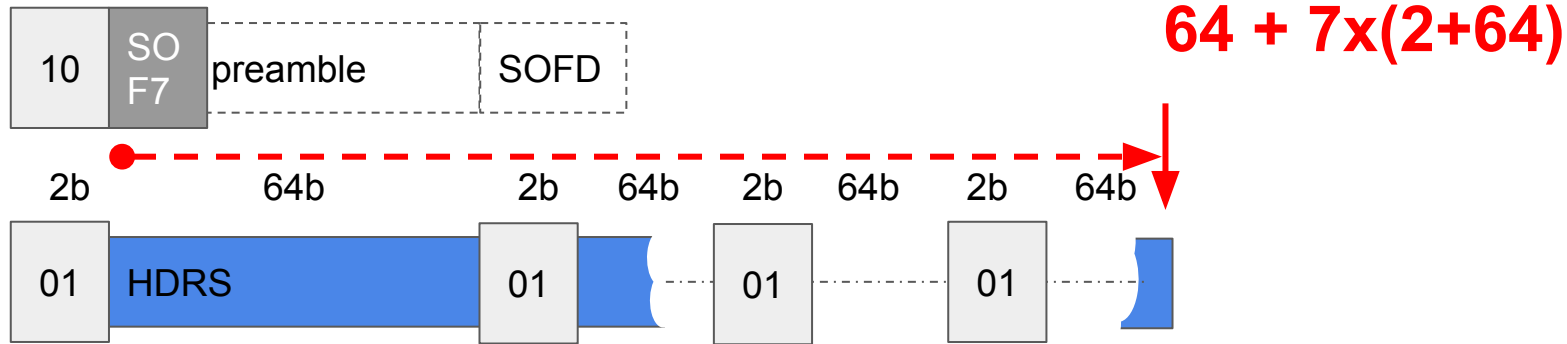
PCS **SOF7** view of a 68B frame:



64b

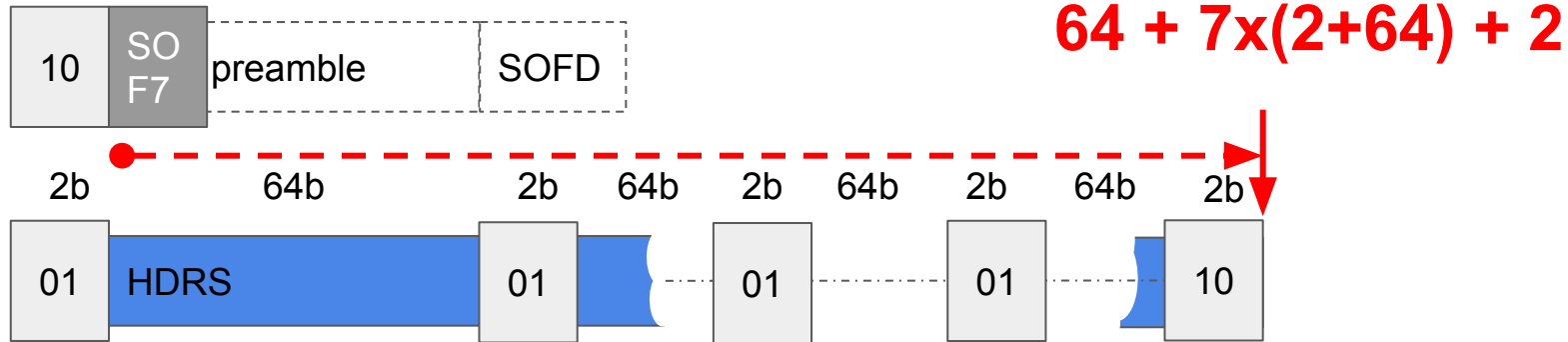
Taking into account PCS layer effects...

PCS **SOF7** view of a 68B frame:



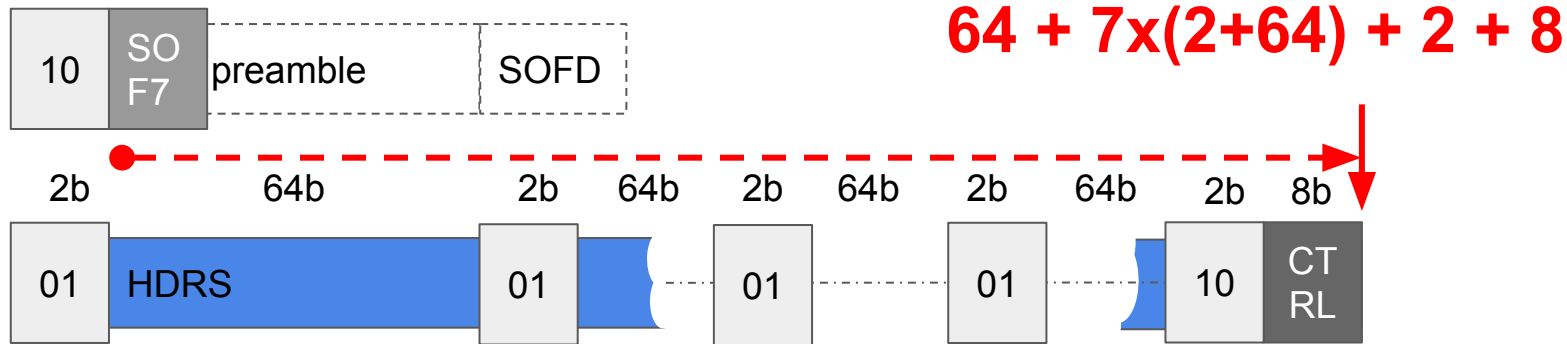
Taking into account PCS layer effects...

PCS **SOF7** view of a 68B frame:



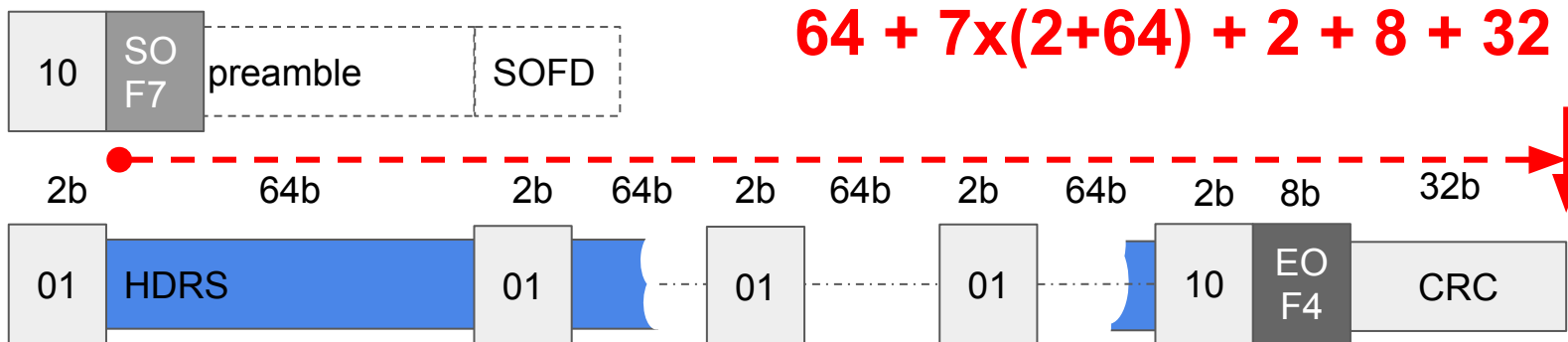
Taking into account PCS layer effects...

PCS **SOF7** view of a 68B frame:



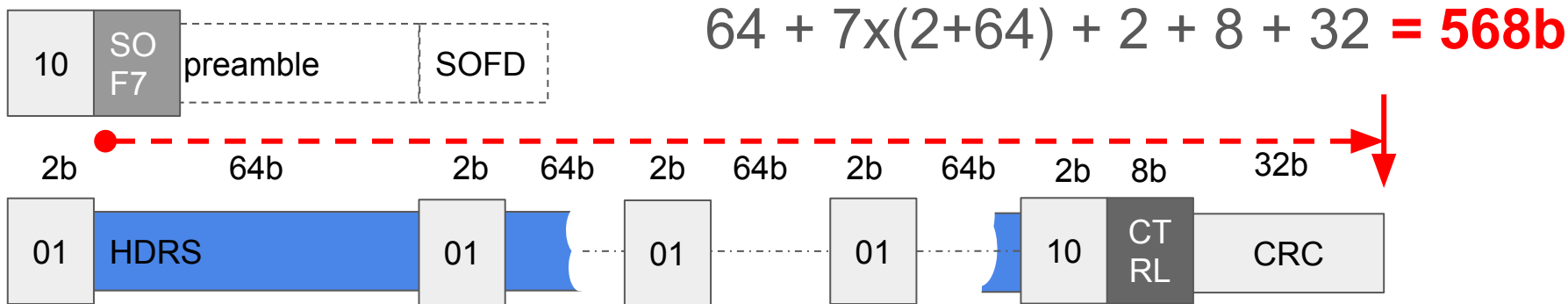
Taking into account PCS layer effects...

PCS **SOF7** view of a 68B fame:




Taking into account PCS layer effects...

PCS **SOF7** view of a 68B frame:



PCS Layer Calculation

bits in the packet


$$\text{Delay} = n$$

PCS Layer Calculation

picoseconds



$$\text{Delay} = n \times 1000$$

PCS Layer Calculation

bit-rate



$$\text{Delay} = n \times 1000 / 10.3125$$

PCS Layer Calculation

simplifies



$$\text{Delay} = n \times 96.97$$

PCS Layer Calculation

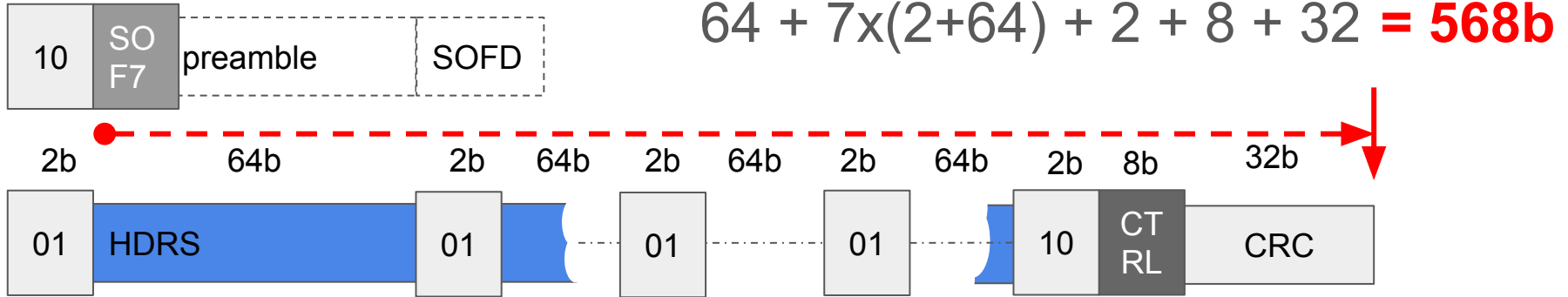
$$\text{Delay} = 568 \times 96.97$$

PCS Layer Calculation

$$\text{Delay} = 568 \times 96.97 = \mathbf{55,079ps}$$

Taking into account PCS layer effects...

PCS **SOF7** view of a 68B frame:

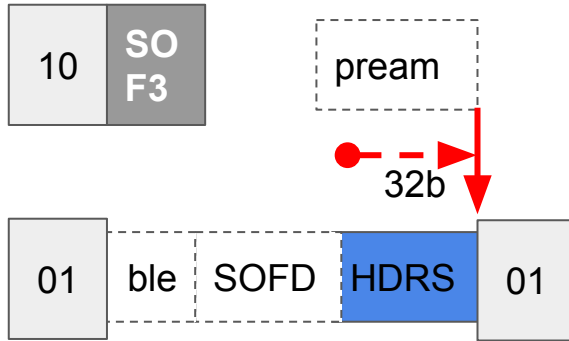


Taking into account PCS layer effects...

PCS **SOF3** view of a 68B fame:

Taking into account PCS layer effects...

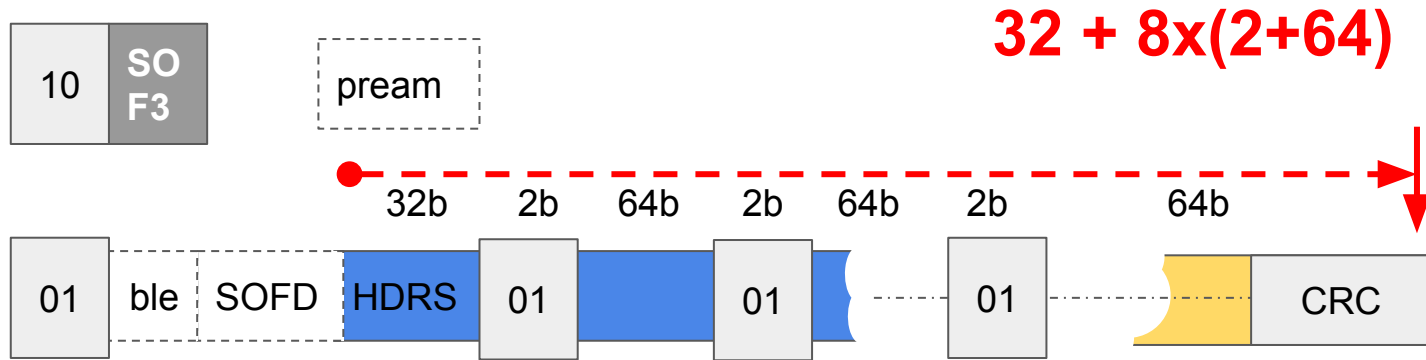
PCS **SOF3** view of a 68B frame:



32

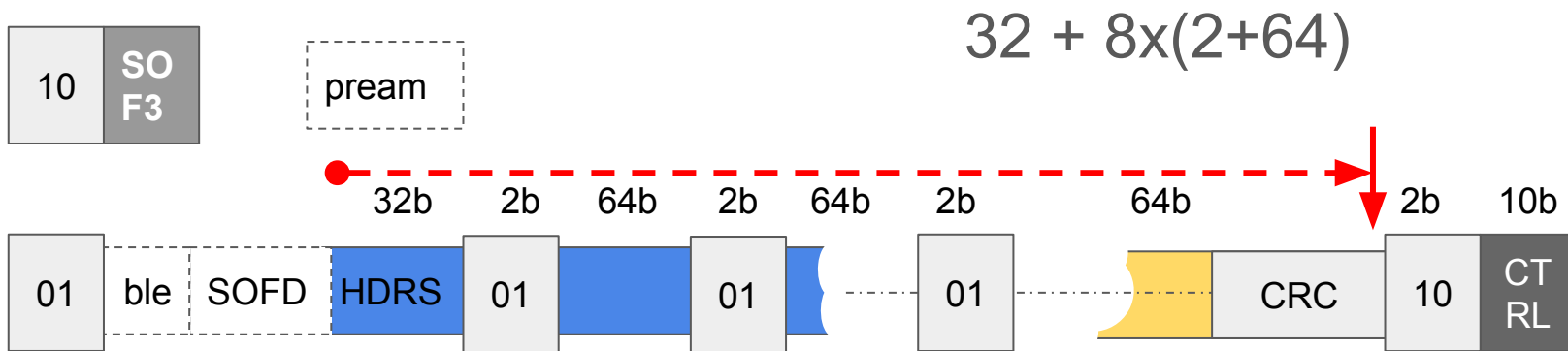
Taking into account PCS layer effects...

PCS **SOF3** view of a 68B frame:



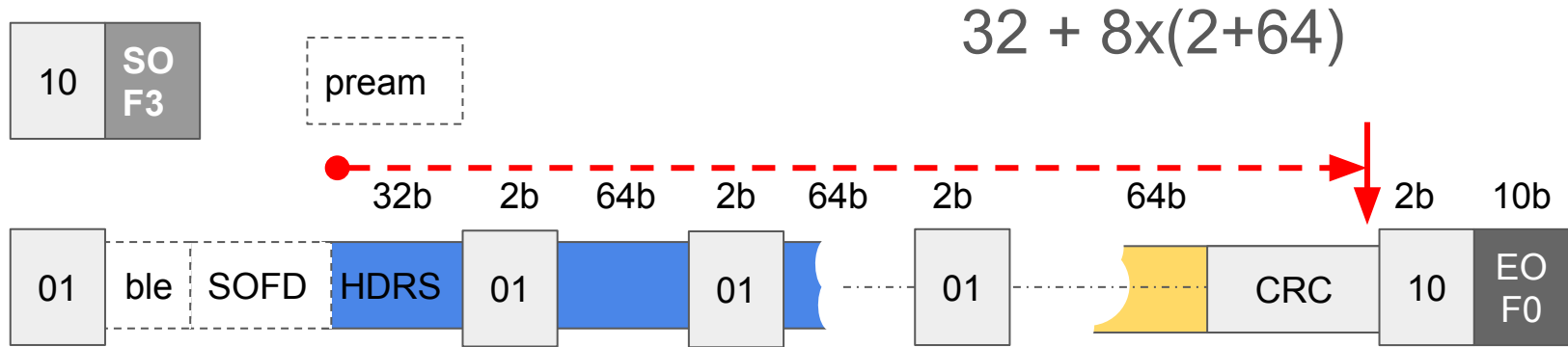
Taking into account PCS layer effects...

PCS **SOF3** view of a 68B fame:



Taking into account PCS layer effects...

PCS **SOF3** view of a 68B frame:

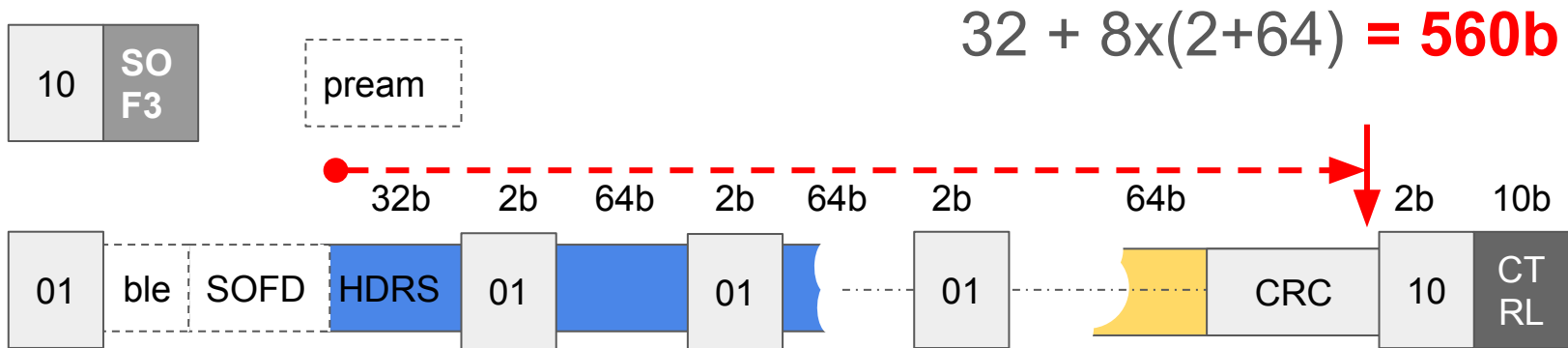


Our recommendations!

- When does the frame start? At the start of the payload
- When does the frame end? At the end of the CRC
- **How long is the frame? (CRC - payload) @ 10.3125G**
- How far (ps) is an offset? (bit offset - payload) @ 10.3125G

Taking into account PCS layer effects...

PCS **SOF3** view of a 68B fame:



PCS Layer Calculation

$$\text{Delay} = 560 \times 96.97 = \mathbf{54,303ps}$$

Results Summary

Ideal frame length: 55,400ps

Results Summary

Ideal frame length: 55,400ps

SOF7 fame length: 55,079ps

Results Summary

Ideal frame length: 54,400ps

SOF7 fame length: 55,079ps

SOF3 frame length: 54,303ps

Results Summary

Ideal frame length: 54,400ps

SOF7 fame length: 55,079ps

SOF3 frame length: 54,303ps

**SOF7 is 679ps
longer than ideal**

Results Summary

Ideal frame length: 55,400ps

SOF7 fame length: 55,079ps

SOF3 frame length: 54,303ps

**SOF3 is 97ps
shorter than ideal**

Results Summary

Ideal frame length: 54,400ps

SOF7 fame length: 55,079ps

SOF3 frame length: 54,303ps



**SOF3 is 776ps
shorter than SOF7**

Conclusions

1. It's harder than it looks to do measurements at the picosecond scale.

Conclusions

1. It's harder than it looks to do measurements at the picosecond scale.
2. Vendors need to specify where/when timestamps are taken to facilitate index offset/frame length calculations

Conclusions

1. It's harder than it looks to do measurements at the picosecond scale.
2. Vendors need to specify where/when timestamps are taken to facilitate index offset/frame length calculations
3. When taking into account PCS layer effects in the STAC-T0 benchmark one frame is 776ps longer/later than expected.

Questions?

(or tick the box)

