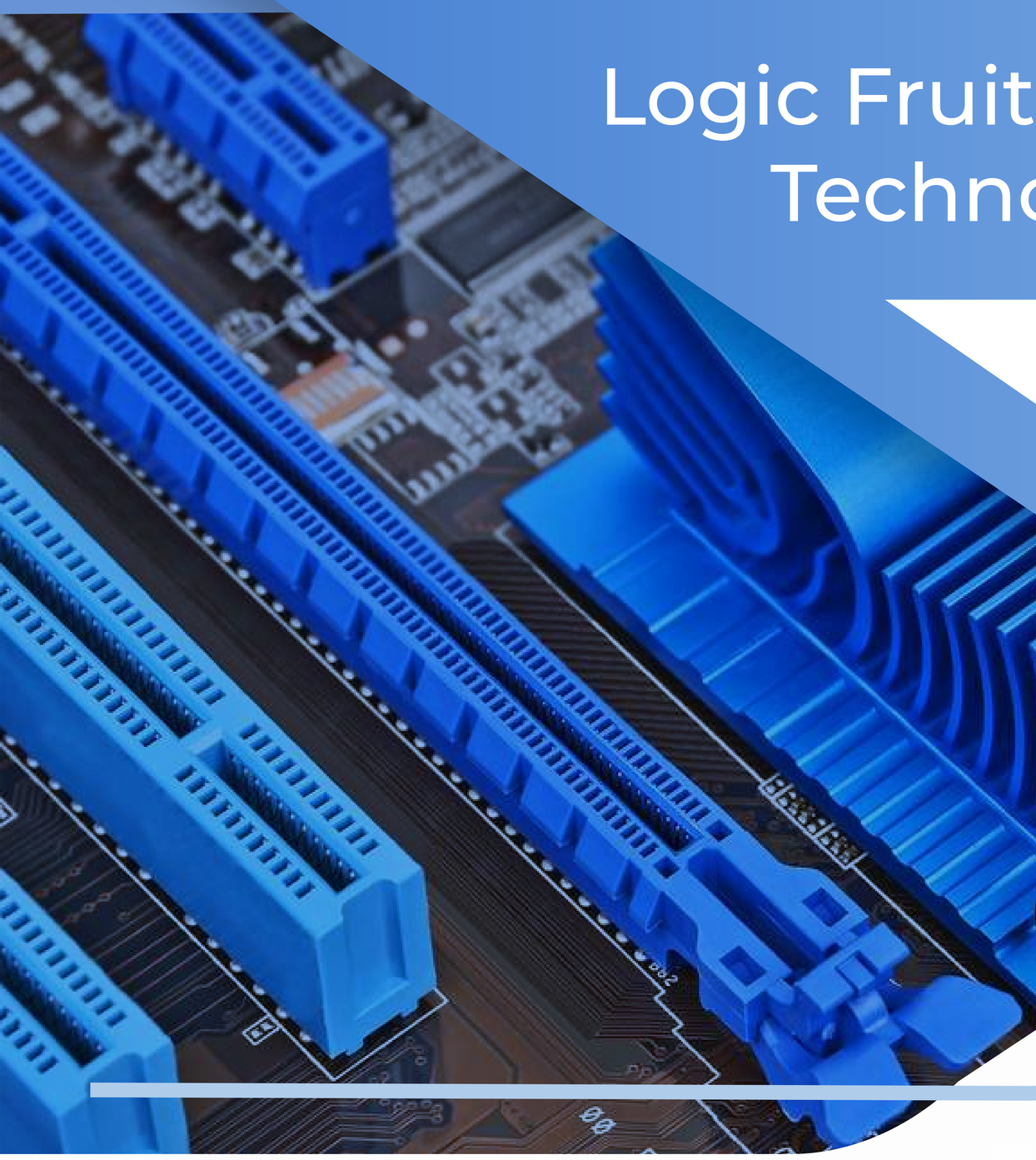
logic fruit Technologies

Logic Fruit Technologies

WHITEPAPER



QUICK INSIGHT INTO

PCI EXPRESS® BASE SPECIFICATION REVISION 6.0 VERSION 0.9

PHYSICAL LAYER LOGICAL SUB- BLOCKS

This white paper presents an overview of the lowest hierarchical layer of PCI Express® Base Specification Revision 6.0 Version 0.9 and outlines the modifications and advancements in the Gen6 Physical Layer Logical Sub-Block.

By:
Sahil Singh & Gonuguntla Deepthi

Contents

Quick Insight into PCI Express® Base Specification Revision 6.0 Version 0.9 Physical Layer Logical Sub-Blocks

Introduction:

The Physical Layer is present at the bottom of the PCIe hierarchy, which includes all logic functions and circuitry for data transfer operations through the physical interface. It isolates the Transaction and Data Link Layers from the signaling technology used for Link data interchange. Information, in the form of TLP and DLLP bytes received from the Transaction layer and Data Link Layer, needs to be transmitted and received across the PCIe Express link in an appropriate serialized format at a compatible frequency supported by the devices connected to each other across a link.

PCIe Gen6 Physical layer can transceive data at the maximum rate of 64 GT/s per lane, which is equipped with the backward compatibility of 32 GT/s, 16GT/s, 8GT/s, 5GT/s and 2.5GT/s. This layer is responsible for the generation of Ordered Sets on the transmit side which will further be used for Link Initialization and Training, Clock Tolerance compensation and indication for entry into and exit from low power state on the link.

The Physical Layer is divided into the logical and electrical sub-blocks. This paper describes the logical sub-block. The logical sub-block directs control and management functions of the Physical Layer.

The Evolving PCIe Interconnect:

As the dependency of computing shifts more towards Artificial Intelligence, Machine Learning, and Cloud Computing, the necessity for large bandwidth has surged. The need for more speed than ever has set an unprecedented demand for faster processing and high speed transfer interconnect, and has brought new challenges along with it. PCI Express® Base Specification Revision 6.0 Version 0.9 incorporates some major changes in PCIe technology as defined below:

Transfer Speed: Data transfer rate doubled to 64 GT/s from 32 GT/s in Gen5.

Signaling: PAM4 signaling is deployed for Gen6 instead of NRZ. Nyquist frequency remains unchanged as 32 GT/s for 64 GT/s data rate, thus helping with the channel loss, but PAM4 encoding is expected to deteriorate the bit error rate (BER).

Forward Error Correction: This mechanism is used to deal with high error rate caused due to PAM-4. Since FEC works on the fixed sized code words, a new data transfer mode, FLIT (Flow Control Unit), is introduced in PCIe generations to enable forward error correction PCIe Gen6 can support only flit mode transfer, whereas, Gen1/2/3/4/5 allows both flit and non-flit mode transfer.

Encoding: Data transfer in Gen5 supports 128b/130b encoding, which is now replaced by 1b/1b encoding in Gen6. Two-bit sync header, present in 128b/130b encoding, has been removed for 1b/1b encoding. Since, Flits are of constant size, the need for PHY layer framing token (4 bytes) for every TLP and DLLP gets totally eliminated.

Ordered Sets: The definition of ordered sets have been changed in Gen6 and all ordered sets continue to be in multiples of 128 except SKPOS, whose nominal length has been increased to 40 bytes. This time FTS ordered set has not been considered as a part of Gen6 Ordered Sets list, meanwhile a new training sequence, TS0, is added to the same list. TS0 ordered sets are used during Recovery.Equalization state at 64 GT/s. Since, Flit mode only supports control SKP ordered sets, therefore, Gen6 doesn't have standard SKP OS.

Gen 6 Ordered Set Blocks:

The new definition adopted for Gen6 Ordered sets is shown below:

Electrical Idle Exit Ordered Set (EIEOS): Transmitting EIEOS periodically, ensures that the Electrical Idle Exit circuitry of the receiver can detect an exit from electrical idle.

EIEOS	Symbol Number	Value
	0 - 7	00h
	8 - 15	FFh

Electrical Idle Ordered Set (EIOS): At 64.0 GT /s, before a Transmitter enters Electrical Idle, it must always send one Electrical Idle Ordered Set (EIOS), unless otherwise specified.

EIOS	Symbol Number	Value
	0, 2, 4, 6, 8, 10, 12, 14	0Fh
	1, 3, 5, 7, 9, 11, 13, 15	F0h

Skip Ordered Set (SKP OS): SKP OS are mainly inserted to compensate for the discrepancies in the transmitter and receiver clocks. Only Control SKP Ordered Sets are transmitted in 1b/1b Encoding. A transmitted SKP Ordered Set is 40 symbols (40B), and a received SKP Ordered Set can be 24, 32, 40, 48 or 56 symbols (24B, 32B, 40B, 48B or 56B). The transmitted SKP Ordered Set consists of 24 bytes of SKPs (F00F_F00Fh), followed by 8 bytes of SKP_END (FFF0_00F0h) and the rest of the bytes are PHY Payload (8B).

SKPOS	**Standard:** Doesn't exist	**Control:** SKP Symbol - F00F_F00Fh SKP END Symbol - FFF0_00F0h

Start of Data Stream Ordered Set (SDS): In 1b/1b encoding, the transmitter must always send two back to back SDS when the conditions to send an SDS is met.

SDS	Symbol Number	Value
	0, 4, 8, 12	B1h
	1-3, 5-7, 9-11, 13-15	C6h

TS0 Ordered Set: TS0 Ordered Set of Gen6 has the first 8 symbols of the Ordered Set replicated in the second 8 symbols of the Ordered Set. This Ordered Set is used in phase 0/1, phase2 stages of Recovery. Equalization. Each Unit Interval(UI) of a TS0 ordered set is either a 00b or 10b. Due to half scrambling, Bits {6,4,2,0} are identical to bits {7, 5, 3, 1}

Symbol Number	Description
0, 8	**TS0 identifier** - 33h (Unscrambled)
1, 9	**Equalization Byte 0 - Half Scrambled** ✓ **Bits 3,1 - Equalization Control (EC)** (00: Phase 0 01: Phase 1 10: Phase 2 11: Phase 3) ✓ **Bit 5 - Reset EIEOS Interval Count** 0b: Do not reset 1b: Reset ✓ **Bit 7 - Use Preset** 0b: Use Coefficients 1b: Use Preset
2, 10	**Equalization Byte 1 -** Half Scrambled ✓ **Bits 7, 5 ,3, 1** **Phases 0 & 1** - FS[3:0] **Phase 2** - First Pre-Cursor Coefficient [3:0] ($\|C_{-1}\|$) **Others** - Reserved
3, 11	**Equalization Byte 2** - Half Scrambled ✓ **Bits 3, 1** **Phases 0 & 1** - FS[5:4] **Phase 2** - Post-Cursor Coefficient [1:0] ($\|C_{+1}\|$) **Others** - Reserved ✓ **Bits 7, 5** **Phases 0 & 1** - LF[1:0] **Phase 2** - Post-Cursor Coefficient [3:2] ($\|C_{+1}\|$) **Others** - Reserved

4, 12	**Equalization Byte 3** - Half Scrambled **Bit 1** **Phases 0 & 1** - LF[2] **Phase 2** - Post-Cursor Coefficient [4] ($\|C_{+1}\|$) **Bits 7, 5, 3** **Phases 0 & 1** - LF[5:3] **Phase 2** - Second Post-Cursor Coefficient [2:0] ($\|C_{-2}\|$) **Others** - Reserved
5, 13	**Equalization Byte 4 -** Half Scrambled **Bits 7, 5, 3, 1** **Phases 0 & 1** - Transmitter Preset [3:0] **Phase 2** - If Use Preset is 1b, then Transmitter Preset [3:0], Else Cursor [3:0] ($\|C_{0}\|$) **Others** - Reserved
6, 14	**Equalization Byte 5** - Half Scrambled **Bits 3, 1** Phases 2 - Cursor [5:4] ($\|C_{0}\|$) Others - Reserved **Bit 5** Retimer Equalization Extend **Bit 7** Reserved
7, 15	If DC Balance adjustment needed at the start of the TS0: 00h or FFh – Unscrambled Else: **Bits 1,3,5,7** Byte level even parity - Half Scrambled Symbol 7 = Symbol 0 ^ Symbol 1 ^ … Symbol 6 Symbol 15 = Symbol 8 ^ Symbol 9 ^ … Symbol 14

TS1/TS2 Ordered Set: TS1/TS2 Ordered Sets of Gen6 have the first eight symbols of the Ordered Set replicated in the second portion of eight symbols of the Ordered Set. TS1/TS2 ordered sets are used for bit alignment, symbol/block alignment and for exchanging physical layer parameters during Link Training.

Symbol Number	Description
0, 8	TS1/TS2 Identifier - Unscrambled • 1Bh for TS1 • 39h for TS2

1, 9	• **Link Number** in Configuration, Hot Reset, or Recovery. RcvrCfg state - Scrambled • PAD is encoded as F7h • As a Receiver in Recovery.Idle, this Byte is only used to check for PAD (F7h) • **Equalization Byte 0** in Recovery and Loopback for TS1 Ordered Set – Scrambled **Bits 1:0** - Equalization Control (EC) **Bit 2** - Reset EIEOS Interval Count **Bits 6:3** - Transmitter Preset in Recovery. **Bit 7** - Use Preset/Equalization Redo • **Equalization Byte 0** in Recovery for TS2 Ordered Set – Scrambled **Bits 2:0** - Reserved **Bits 5:3** – Equalization Request Data Rate 000b 8.0 GT/s 001b 16.0 Gt/s 010b 32.0 GT/s 011b 64.0 GT/s Others Reserved **Bit 6** - Quiesce Guarantee **Bit 7** - Request Equalization • Reserved in other states - Scrambled
2, 10	• **Lane Number** in Configuration or Hot Reset state - Scrambled • PAD is encoded as F7h • As a Receiver in Recovery.Idle, this Byte is only used to check for Pad (F7h) • **Equalization Byte 1** in Recovery – Scrambled **Bits 5:0** - Cursor |C0| for the current data rate. **Bit 6** - Transmitter Precoding on **Bit 7** - Retimer Equalization Extend bit • Reserved in other states - Scrambled
3, 11	• **Equalization Byte 2** in Recovery for TS1 Ordered Sets, Reserved for TS2 Ordered Sets - Scrambled **Bits 3:0** - $\|C_{-1}\|$[3:0] for the current data rate **Bits 6:4** - $\|C_{-2}\|$[2:0] for the current data rate. **Bit 7** - Reject Coefficient Values bit • Reserved in other states - Scrambled
4, 12	• **Equalization Byte 3** in Recovery for TS1 Ordered Sets, Reserved for TS2 Ordered Sets - Scrambled **Bits 4:0** - $\|(C_{+1})\|$ [4:0]for the current data rate . **Bits 7:5** – Reserved • Reserved in other states - Scrambled
5, 13	**Data Rate Identifier** - Scrambled **Bit 0** - Reserved **Bit 5:1** - Data Rates Supported 00001b - Only 2.5 GT /s supported 00011b - 2.5 and 5.0 GT /s supported 00111b - 2.5, 5.0, and 8.0 GT /s supported

	01111b - 2.5, 5.0, 8.0, and 16.0 GT /s supported 11111b - 2.5, 5.0, 8.0, 16.0, and 32.0 GT/s supported 10111b - 2.5, 5.0, 8.0, 16.0, 32.0, and 64.0 GT/s supported **Bit 6** - Autonomous Change/ Selectable De-emphasis **Bit 7** - speed_change
6, 14	**Training Control** Used in Recovery.RcvrCfg to Disable, Hot Reset, or Loopback. Reserved for TS2 Ordered Sets in Configuration - Scrambled **Bits 3:0** 0000b - Deassert 0001b - Assert Hot Reset 0010b - Assert Disable 0100b - Assert Loopback 0101b - Assert Loopback 0110b - Assert Loopback 1000b - Assert Compliance Receive Bit 1100b - Assert Loopback and Compliance Receive bits Others : Reserved **Bit 7:4** - Reserved
7, 15	• If DC Balance needs adjustment at the start of the TS1 or TS2: DC Balance Symbol - Unscrambled • **Else:** Byte level even parity over Symbols 0-6 (or 8-14) - Scrambled Symbol 7 = Symbol 0 ^ Symbol 1 ^ ... Symbol 6 Symbol 15 = Symbol 8 ^ Symbol 9 ^ ... Symbol 14

Flit Data Stream mode:

The Flit has a fixed size. Each Flit is of 256 bytes length. The allocation of bytes in Flit are as follows:

236 bytes for TLP (Bytes from 0 to 235)

6 bytes for DLLP (Bytes from 236 to 241)

8 bytes for CRC (Bytes from 242 to 249)

6 bytes for ECC (Bytes from 250 to 255)

A flit is interleaved across the Lanes in the Link in a Byte aligned fashion. Each Byte corresponds to a Symbol in the 64.0 GT/s and above Data Rates. The 8 bytes CRC protects the TLP and DLLP bytes (ECC bytes are not included), and The 6 bytes ECC protects the entire flit including the CRC.

Forward Error correction is 3-way interleaved ECC. Each interleave has 2 ECC bytes. This interleaving is added so that burst error of up to 16 bits in any lane does not impact more than one byte in each interleaved ECC code.

Flit layout in a x2 link is shown in *Table1.*

	Lane 0	Lane 1
TLP bytes	0	1
	2	3
	….	…
	232	233
	234	235
DLP bytes	DLP0	DLP1
	DLP2	DLP3
	DLP4	DLP5
CRC bytes	CRC0	CRC1
	…	…
	CRC6	CRC7
ECC bytes	ECC0A	ECC0B
	ECC0C	ECC1A
	ECC1B	ECC1C

***Table1**. Flit layout in a x2 link*

Steps involved in Tx and Rx PHY Layer:

The physical layer consists of two main section: a Transmit section which prepares outgoing information, travelled all the way through Transaction layer and Data Link layer, for transmission over the external channel and to reach at the other end of the same channel; and a Receive section that identifies and translates the received information in the required structure before passing it to the Data Link layer. An abstract representation of the processing of the data symbols in the Flit mode at the 64 GT/s rate on the

Transmitter and the Receiver side is demonstrated in *Figure 1* and *Figure 2* respectively.

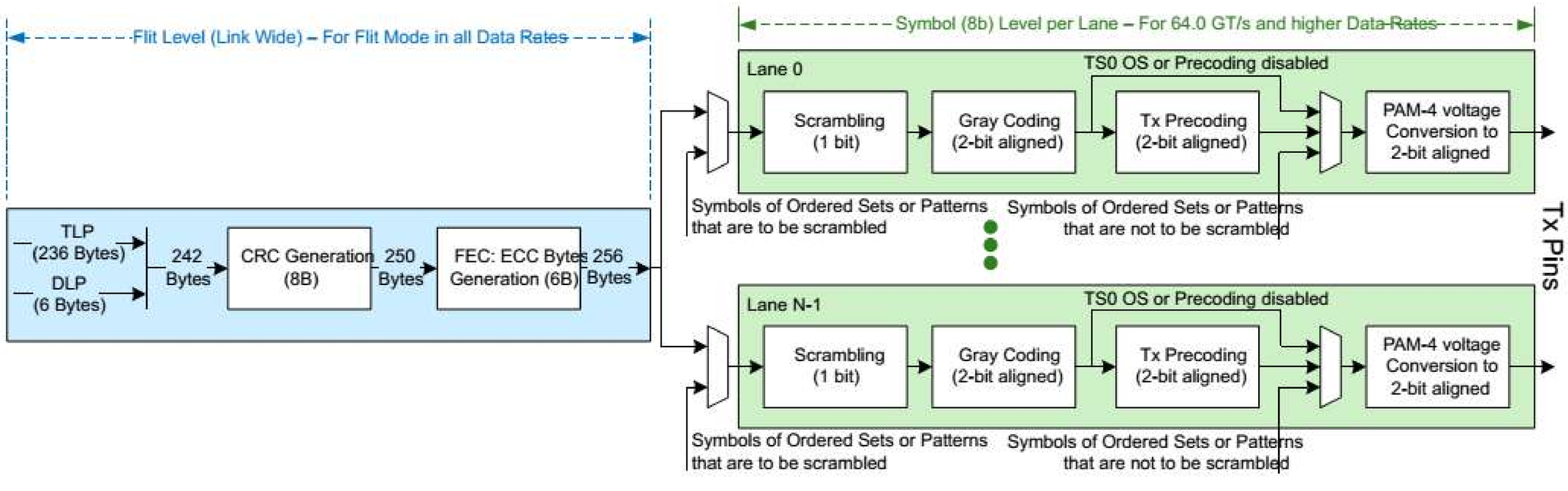


Figure 1: Flit mode processing with 1b/1b encoding on the transmit side

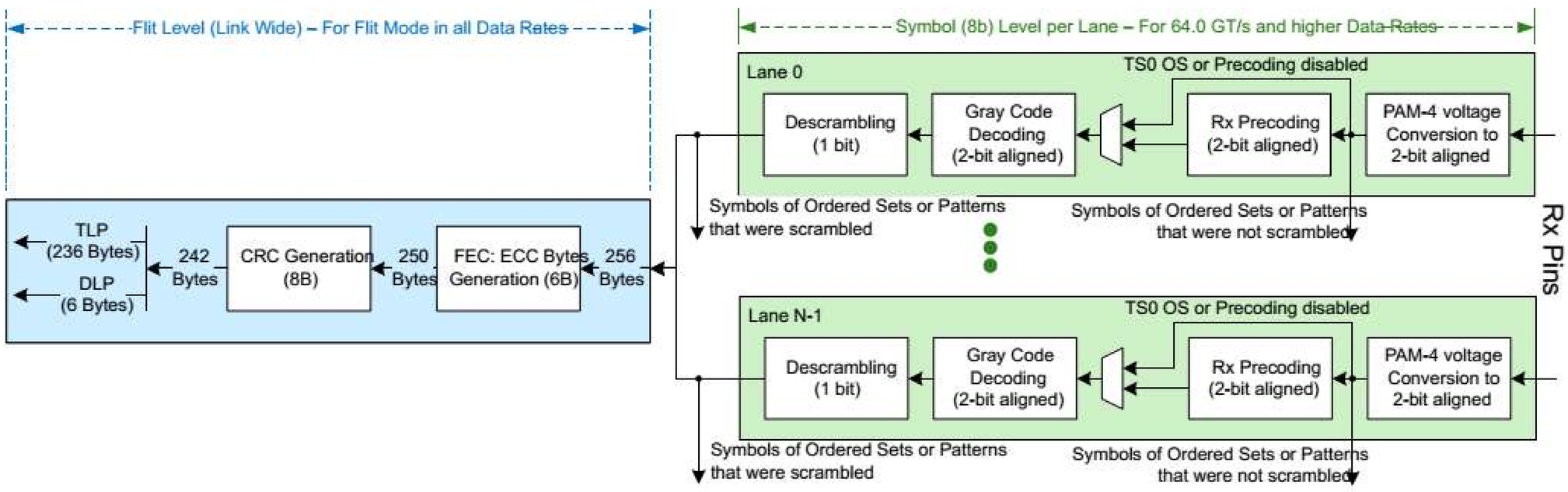


Figure 2: Flit mode processing with 1b/1b encoding on the receive side

The data entering the Physical layer has to pass through following stages before getting out from the transmitter or after entering into the receiver:

- ☞ The fixed sized data in the form of TLP and DLLP bytes is used to generate the 8 bytes of CRC at the transmitter, which will further protect the TLP and DLLP bytes. On the Receive side the CRC bytes are again generated and compared against the received CRC bytes (after the FEC decode) by the CRC checker and any mismatch represents an uncorrectable error.

- ☞ Now the algorithmic evaluation of 3-way interleaved ECC is done using the combined TLP, DLLP and CRC bytes. It ensures the protection of the entire flit including itself. Each interleaved ECC pair will further be used at the receiver by the FEC decoder to correct a single byte error.

- ☞ The appended CRC and ECC bytes completes the Flit which will be directed towards the transmitter's scrambler along with the Ordered Sets to be scrambled, whereas, on receive side Flit data and Ordered Sets will be separated from the descrambled output from the receiver, and the Flit data moves towards the FEC decoder.

- ☞ At the transmitter, the data will now be scrambled to prevent repetitive patterns in the data stream. This involves XORing the data stream with a pattern generated by a Linear The data entering the Physical layer has to pass through following stages before getting out from the transmitter or after entering into the receiver:

- Only scrambled bits, at the transmitter, on a 2-bit aligned boundary are gray coded. Gray Coding is only enabled in the PAM4 mode (at 64.0 GT/s and above Data Rates). The 2-bit aligned input 00b, 01b, 10b, 11b is converted to 00b, 01b, 11b, and 10b respectively. The inverse gray coding at the receive side is completely identical, except the input will be coming from the Inverse Precoder.

- The output from the gray coder will be treated as the input to the Precoder at the transmitter. Precoding also works on 2-bit aligned boundary, and only the scrambled bits are Precoded. At the receiver, Inverse Precoding is performed to get back the gray coded data. Precoding is bypassed for all Symbols of TS0 Ordered Sets.

The truth table for the Precoding function for the transmit and receive side is shown in *Table 2* and *Table 3*. In *Table 2*, P_n and T_n are the input and output of the Precoder and T_{n-1} is the output from the Precoder in the previous UI, and in *Table 3*, R'_n and P'_n are the input and output of the inverse Precoder and R'_{n-1} is the input of the inverse Precoder in the previous UI.

T_n				
P_n	T_{n-1}			
	00	01	11	10
00	00	11	01	10
01	01	00	10	11
11	11	10	00	01
10	10	01	11	00

Table 2. Truth Table for the Precoding

P'_n				
R'_n	R'_{n-1}			
	00	01	11	10
00	00	01	11	10
01	01	10	00	11
11	11	00	10	01
10	10	11	01	00

Table 3. Truth Table for Inverse Precoding

- On the transmit side, unscrambled symbols of Ordered Sets are MUXed with the Precoder output and TS0 Ordered Sets. This MUXed output is sent over the link after the PAM4 signaling, whereas on the receive side, PAM4 voltage is converted into a 2-bit aligned quantity.

Enhanced Bit Error Rate (BER) in Gen6:

Introduction of PAM4 signaling, where two bits are encoded using four levels in the same unit interval (UI), has worsened the bit error rate significantly as compared to lower data rates (2.5GT/s, 5.0GT/s, 8.0GT/s, 16.0GT/s and 32GT/s) BER of 10-12. The four levels map to the voltage value of -400mV, -133mV, +133mV, +400mV. Because of four voltage levels and reduced amplitude for each voltage level, 64 GT/s PAM4 signaling is sensitive to noise and burst errors. As a result, the probability of the errors in the Flit data and Ordered Sets received by the receiver has increased. The bit error without accounting for any burst error is called First Bit Error Rate (FBER) and it is expected to be around 10-6. This necessitates the inclusion of techniques to minimize the bit errors in the Flit data, and to consider some relaxation during the detection of errored Ordered Sets.

Methods incorporated for minimizing High BER in Gen6

Below mentioned approaches are used to reduce the higher bit error rate:

- **Below mentioned approaches are used to reduce the higher bit error rate:** To minimize the high First Bit Error rate in the data stream and improve the bandwidth efficiency a light weight FEC, to keep low latency, and strong CRC, for high reliability, are used. It eliminates the need for re-transmission of the flits to some extent. FEC works on fixed size Flits, where a Flit consists of a fixed number of bytes of TLPs, DLLPs, CRC and ECC. Here, 6 bytes long ECC plays the crucial role of error correction of the entire flit at the receiver's end.

- **Gray coding:** In PAM4 signaling the most likely error scenario is +/-1 on the voltage level over one Unit Interval. Table 4 demonstrates resulting voltage under this error scenario of ±1 on voltage level and resulting bit error with Gray coded encoding. Gray coding results in at most a single bit flip within that Unit interval. Thus Gray coding reduces errors on certain channels.

Voltage Level (Gray coded Encoding)	Resulting voltage level	
	Error of +1	Error of -1
0 (00b)	1 (01b)	0 (00b) (no error)
1 (01b)	2 (11b)	0 (00b)
2 (11b)	3 (10b)	1 (01b)
3 (10b)	3 (00b) (no error)	2 (11b)

Table 4. Effect of +/-1 voltage level error on the wire for various PAM4 voltage levels

- **Precoding:** Precoding in 64 GT/s works similar to precoding at 32GT/s, except here it works on 2-bit aligned boundary with PAM-4 encoding. Precoding along with gray coding is effective on certain channels in reducing the impact of errors. In cases where error voltage magnitude +/-1 and DFE errors occur in consecutive UIs, Precoding ensures that only two bit errors appear. Precoding may not be effective for cases where error voltage magnitude is >+1 or <-1 and for cases where errors are introduced in non-contiguous UI's due to DFE.

Relaxation in Ordered Set Detection

At 64 GT/s, Ordered Sets exhibit a greater number of errors as compared to the corresponding Ordered Sets of the low data rates. Thus, relaxation is required while detecting the presence of Ordered Set in the continuous stream of data. The following relaxations are included in each OS:

i)EIEOS:

When processing Ordered Sets outside Data Stream:

If any 5 or more aligned and consecutive bytes in each of the first and last 8 bytes of a Block match the corresponding bytes of an EIEOS and either Symbol 0 or Symbol 8 of the Block matches the corresponding byte of an EIEOS, the Ordered Set must be considered as an EIEOS. A Receiver is permitted to ignore up to one bit mismatch for this comparison.

While in Locked phase, a Receiver must switch the block boundary by > 1UI on an EIEOS Ordered Set or should shift the Block boundary by ≤ 1UI only if it receives an EIEOS followed by a valid Ordered Set.

While in Aligned Phase searching for an EIEOS, the 8-byte boundary can begin in any aligned UI and the bits must be looked at prior to performing gray coding, precoding, or descrambling in the Receiver side.

When processing Ordered Sets during or end of a Data Stream:

If any 5 or more aligned bytes of the first aligned 8 byte chunk matches the corresponding bytes of EIEOS, the Ordered Set must be considered as an EIEOS.

ii)EIOS:

When processing Ordered Sets outside Data Stream:

If any 5 or more aligned bytes of the first 8 bytes of a Block match the corresponding bytes of an EIOS and either Symbol 0 or Symbol 8 of the Block matches the corresponding byte of an EIOS, the Ordered Set must be considered as an EIOS. When an EIOS is received, the Lane is ready to enter Electrical Idle irrespective of what is received in the last 7 Bytes of the Block.

When processing Ordered Sets during or end of a Data Stream:

If any 5 or more aligned bytes of the first aligned 8 byte chunk matches the corresponding bytes of EIOS, the Ordered Set must be considered as an EIOS.

iii)SKPOS:

When processing Ordered Sets outside Data Stream:

A SKP OS is considered to begin if any 5 or more aligned bytes of the first 8 bytes of a Block match a SKP, and either Symbol 0 is a SKP or Symbol 8 of the Block is either a SKP or a SKP_END. It looks at each subsequent aligned 8 byte chunks, if any 5 or more aligned bytes match SKP_END or the current chunk is the fifth 8 byte chunk after the start of the SKP OS (i.e., current set is bytes 40 through 47 from the start of the SKP OS), the SKP OS will terminate after the next aligned 8 byte chunk.

When processing Ordered Sets during or end of a Data Stream:

If any 5 or more aligned bytes of the first aligned 8 byte chunk match the corresponding bytes of SKP OS, then the beginning of the SKP OS must be considered.

Then it looks at each subsequent aligned 8 byte chunks and applies the following rule:

If at least each of any 5 aligned bytes matches SKP_END or the current chunk is the fifth 8 byte chunk after the start of the SKP OS (i.e., current set is bytes 40 through 47), the SKP OS will terminate after the next aligned 8 byte chunk. But it is necessary that the SKP OS must terminate with the receipt of 5 aligned bytes of SKP_END.

iv)TS0/TS1/TS2:

In 1b/1b encoding, a received TS0, TS1, or TS2 Ordered Set is considered valid when either half is valid. Conditions for a half being valid are stated as:

- The first half is valid if Symbol 0 is a valid TS0/TS1/TS2 encoding, Symbols 0 to 7 passes its parity check after decoding Gray code and descrambling, and Symbol 7 is not equal to a DC balance pattern prior to performing gray code and descrambling.
- The second half is valid if Symbol 8 is a valid TS0/TS1/TS2 encoding, Symbols 8 to 15 passes its parity check after decoding Gray code and descrambling, and symbol 15 is not equal to a DC balance pattern prior to performing gray code and descrambling.
- Both halves are valid if Symbols 0-7 are identical to Symbols 8-15.

The type of the Ordered Set (TS0, TS1, or TS2) is determined by the first symbol of the valid half. The even bits in the half-scrambled Symbols must be ignored after descrambling in the Receiver. (Note: because the gray code was done by forcing the even bit to be identical to the odd bit in the Transmitter, the even bit position may not be identical to the odd bit position on the Receiver after descrambling)

Flit Mode Negotiation:

Flit Mode is supported at all data rates. Flit Mode negotiation occurs during the initial Link Training (at 2.5GT/s), if both ports support. When Link is Up then link never transitions to Flit mode from Non-Flit Mode and vice versa. Operation at 64.0 GT/s or higher is permitted only if Flit mode negotiation is succeeded.

Flit Mode is negotiated in Configuration state. Flit Mode is enabled, if the following conditions are satisfied:

- When Link is down.
- The Flit Mode Supported bit is transmitted in the 'Data Rate Identifier' field (Symbol 4, Bit 0) in the TS1 and TS2 Ordered Sets in Polling and Configuration states since entering the Polling State.
- The received eight consecutive TS2 Ordered Sets on all Lanes had the set Flit Mode Enabled in the Data Rate Identifier field (Symbol 4, Bit 0) to 1b of the currently configured Link that caused the transition from Polling. Configuration to Configuration state.

Flit Mode for Lower Data Rates:

PCIe Generations support two data stream modes (Flit and Non-Flit modes) and three types of encoding (8b/10b, 128b/130b, 1b/1b). A Data Stream in the non-Flit Mode is defined as a contiguous collection of TLPs, DLLPs (or DLPs), and Logical Idle/IDL Token, starting at the end of an Ordered Set and ending with another Ordered Set or a Link Electrical Idle. A Data Stream in Flit Mode is defined as a set of Flits, starting at the end of the first SKP Ordered Set after an SDS Ordered Set, and ending with the last Flit prior to an Ordered Set other than SKP Ordered Set that causes the Link to exit out of L0 state or if the Link enters Electrical Idle.

Even though the Error Correction is required only with the high FBER (First Bit Error Rate) with the PAM4 signaling at 64.0 GT/s and higher Data Rates, it will be deployed for the lower Data Rates for consistency. Flit mode negotiation happens as described in the previous section. Non-Flit mode remains unchanged from earlier PCIe Specifications whereas Flit mode needs some changes.

The following modifications are required for lower data rates to operate in Flit mode:

☞ At lower data rates, the fixed-sized flits will be sent, similar to the 64 GT/s, along with the SKP OS, based on the type of encoding, as shown in *Table 5.*

Current Data Rate	Encoding	Data Stream	Ordered Set
2.5GT/s, 5.0GT/s	8b/10b	Flit Mode	Same Ordered Sets are used in Flit Mode as well as Non-Flit Mode.
8.0GT/s, 16.0GT/s, 32.0GT/s	128b/130b	Flit Mode	Alternates between Standard SKP OS and Control SKP OS, when SKP OS needs to be sent. Rest of the Ordered Sets are identical for non-Flit Mode and Flit Mode in the corresponding Data Rate.

***Table 5**. Ordered Sets in Flit mode*

☞ **SKP Ordered Set Interval During Data Processing:**

The SKP Ordered Set insertion interval for data rates of 2.5 GT/s and 5.0 GT/s in Flit mode remains unchanged (between 1180 and 1538 Symbols), but occurs at the Flit boundary. For the data rates above 8.0GT/s the SKP OS insertion interval is tabulated in *Table 6* in terms of Flits.

Encoding and Clocking Modes	Link Width				
	x16	x8	x4	x2	x1
Common Clock/SRNS Mode with 128b/130b encoding	374	187	93	46	23
SRIS Mode with 128b/130b encoding	37	18	9	4	2

***Table 6**. SKP OS insertion in Data stream in terms of Flits*

☞ **Flit Mode Data Path on Transmit side and Receiver side:**

At lower data rates (2.5GT/s, 5.0GT/s, 8.0GT/s, 16.0GT/s, 32GT/s) when Flit Mode is enabled, data is transmitted in the form of Flits.

An abstract representation of the Flit Mode and Non-Flit Mode Data processing with 8b/10b encoding for 2.5 GT/s and 5.0 GT/s as well as with 128b/130b encoding for 8.0 GT/s, 16.0 GT/s, and 32.0 GT/s for the Transmit and Receive side is shown in Figure 3 and Figure 4 respectively.

On the Transmit side, CRC and ECC bytes are generated for Flit from the received TLPs and DLLPs from the Data link layer. These Flits are passed to the data path which is identical to Non-Flit Mode. Flits are scrambled and in case of 32.0GT/s precoding is performed, if enabled and corresponding encoding is done as per data rate and transmitted with NRZ signaling.

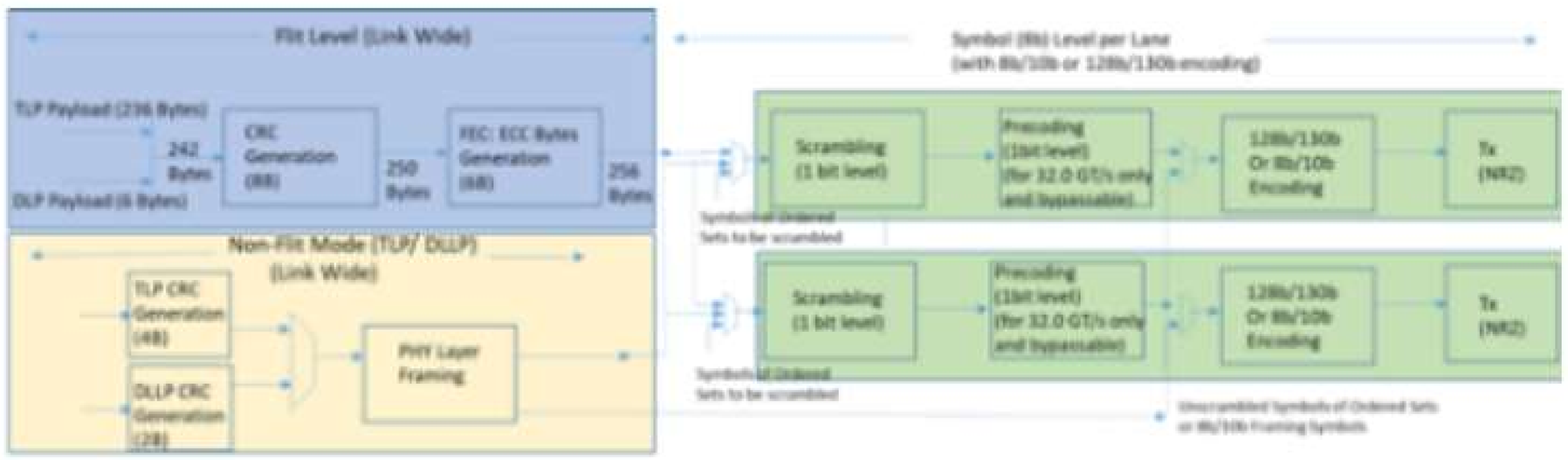


***Figure 3**. Flit Mode and Non-Flit Mode processing with 8b/10b and 128b/130b encoding on the Transmit side*

On the Receive side, data path till descrambler is identical to Non-Flit Mode. After that Flit passes through FEC Decoder and CRC checker. Based on CRC check, a decision is made whether the received flit is accepted or retried. The valid Flit's TLP and DLLP are passed to the Data Link Layer.

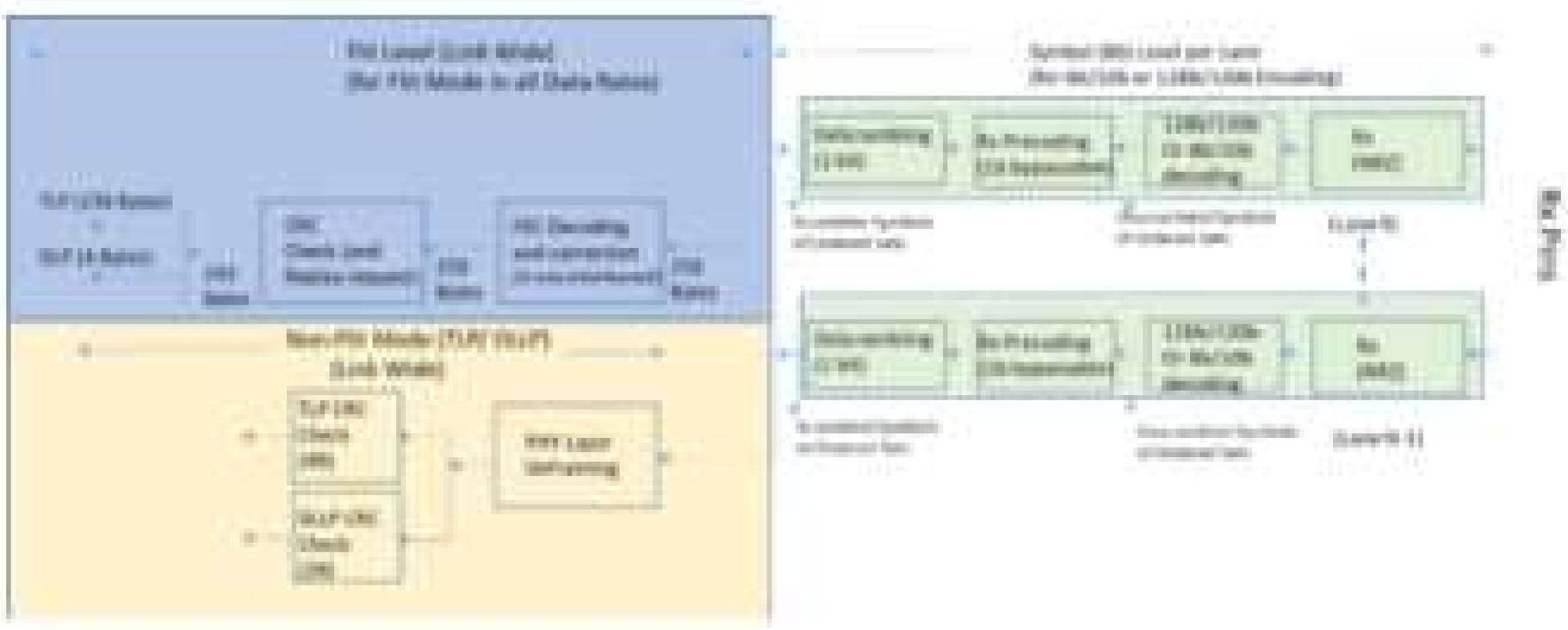


***Figure 4**. Flit Mode and Non-Flit Mode processing with 8b/10b and 128b/130b encoding on the Receive side*

References:

PCI Express® Base Specification Revision 6.0 Version 0.9.

About Us:

At Logic Fruit, we specialize in Architecting, Implementing and Validation high-quality real-time high throughput FPGA/SoC embedded solutions, and Developing Proof-of-concept (PoC) designs/prototypes real-time data generation, acquisition and analysis.

Our engineers have expertise with many high speed protocols and interfaces, including 1G, /10G/40G/100G Ethernet, PCIe(Gen1-Gen6), USB3.0/4.0, CPRI/ORAN, DisplayPort, ARINC818 etc.

The team also has deep expertise in Signal processing for wireless and Imaging based solution development, software-defined radio (SDR), as well as encryption, protocol compliance, signal generation, data analysis, IoT technology, and multiple image processing techniques.

LFT and PCIe:

We have worked on RTL IP and sub-system design for Latest generation of PCIe and other high speed serial protocols, and development of device drives. Verification is done using the latest methodologies like UVM and RTL-SW co-simulation. We also perform FPGA prototyping, validation, and testing with real DUTs. Development of reference hardware is done with different kinds of PCIe devices and links.

logic fruit Technologies

Thank You!

Does anyone have any questions?

Gurugram (Headquarter)

806, 8th Floor
BPTP Park Centra Sector–30,
NH–8 Gurgaon – 122001
Haryana (India)

info@logic-fruit.com

+91-0124 4643950

Bengaluru (R&D House)

Sy. No 118, 3rd Floor,
Gayathri Lakefront,
Outer Ring Road, Hebbal,
Bangalore - 560 024

sales@logic-fruit.com

+91 80-69019700/01

United States (Sales Office)

Logic Fruit Technologies
INC 691 S Milpitas Blvd
Ste 217 (Room 9) Milpitas
CA 95035

info@logic-fruit.com

+1-408 338 9743