Computer Organization and Architecture (EET 2211)

Chapter 3
A Top-Level View of Computer Function and Interconnection

PCI Express

- The **peripheral component interconnect (PCI)** is a popular high-bandwidth, processor-independent bus.

- Delivers better system performance for high speed I/O subsystems.

- The bus-based PCI scheme has not been able to keep pace with the data rate demands of attached devices.

- Hence, a new version, known as PCI express (PCIe) has been developed which is intended to replace bus-based schemes such as PCI.

Contd.

- The key requirement for PCIe is high capacity to support the needs of higher data rate I/O devices, such as Gigabit Ethernet.

- Another requirement is the need to support time-dependent data streams for various applications such as video-on-demand and audio redistribution etc. which usually put real-time constraints on servers.

PCI Physical and Logical Architecture

- Fig.1 shows a typical configuration that supports the use of PCIe.
- A **root complex** device, also referred to as a *chipset* or a *host bridge*, connects the processor and memory subsystem to the PCI Express switch fabric comprising one or more PCIe and PCIe switch devices.

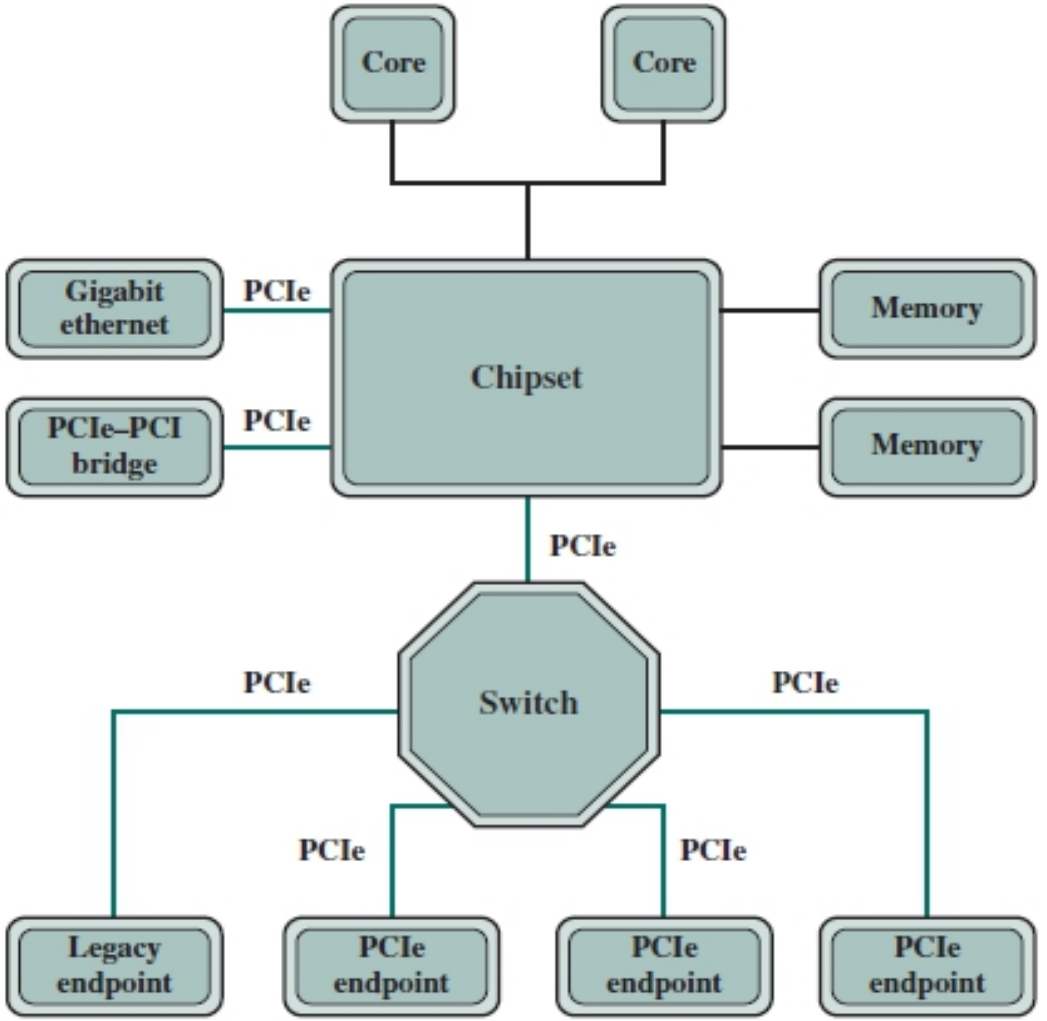
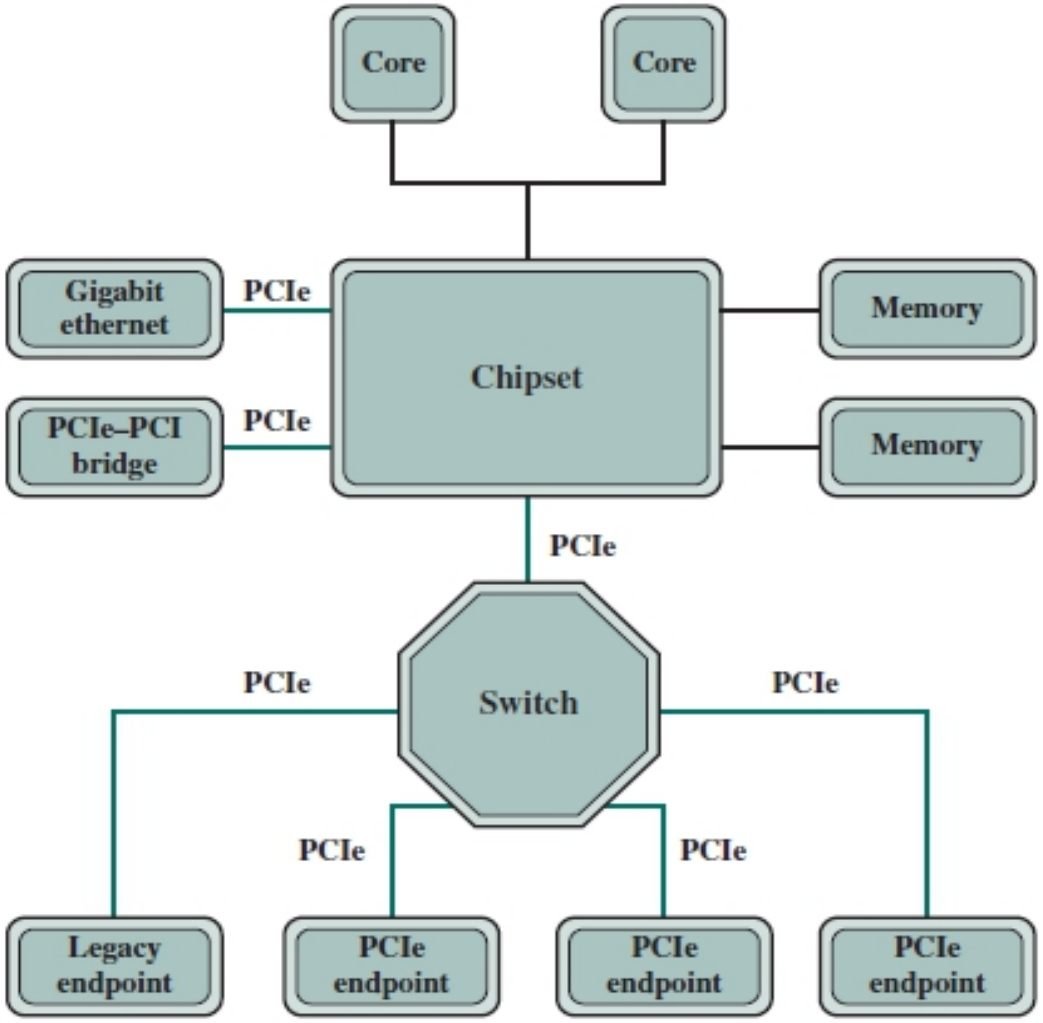


Fig.1. Typical Configuration using PCI

Contd.

- The root complex acts as a buffering device, to deal with difference in data rates between I/O controllers and memory and processor components.
- The root complex also translates between PCIe transaction formats and the processor and memory signal and control requirements.
- The chipset will typically support multiple PCIe ports, some of which attach directly to a PCIe device, and one or more that attach to a switch that manages multiple PCIe streams.

Contd.



- PCIe links from the chipset may attach to the following kinds of devices that implement PCIe:
- **Switch:** The switch manages multiple PCIe streams.
- **PCIe endpoint:** An I/O device or controller that implements PCIe, such as a Gigabit Ethernet switch, a graphics or video controller, disk interface, or a communications controller.
- **Legacy endpoint:** Legacy endpoint category is intended for existing designs that have been migrated to PCI Express, and it allows legacy behaviors such as use of I/O space and locked transactions.
- **PCIe/PCI bridge:** Allows older PCI devices to be connected to PCIe-based systems.

PCIe Protocol Layers

The PCIe protocol architecture includes the following layers:

- **Physical:** Consists of the actual wires carrying the signals, as well as circuitry and logic to support ancillary features required in the transmission and receipt of the 1s and 0s.
- **Data link:** Is responsible for reliable transmission and flow control. Data packets generated and consumed by the DLL are called Data Link Layer Packets (DLLPs).
- **Transaction:** Generates and consumes data packets used to implement load/ store data transfer mechanisms and also manages the flow control of those packets between the two components on a link. Data packets generated and consumed by the TL are called Transaction Layer Packets (TLPs).

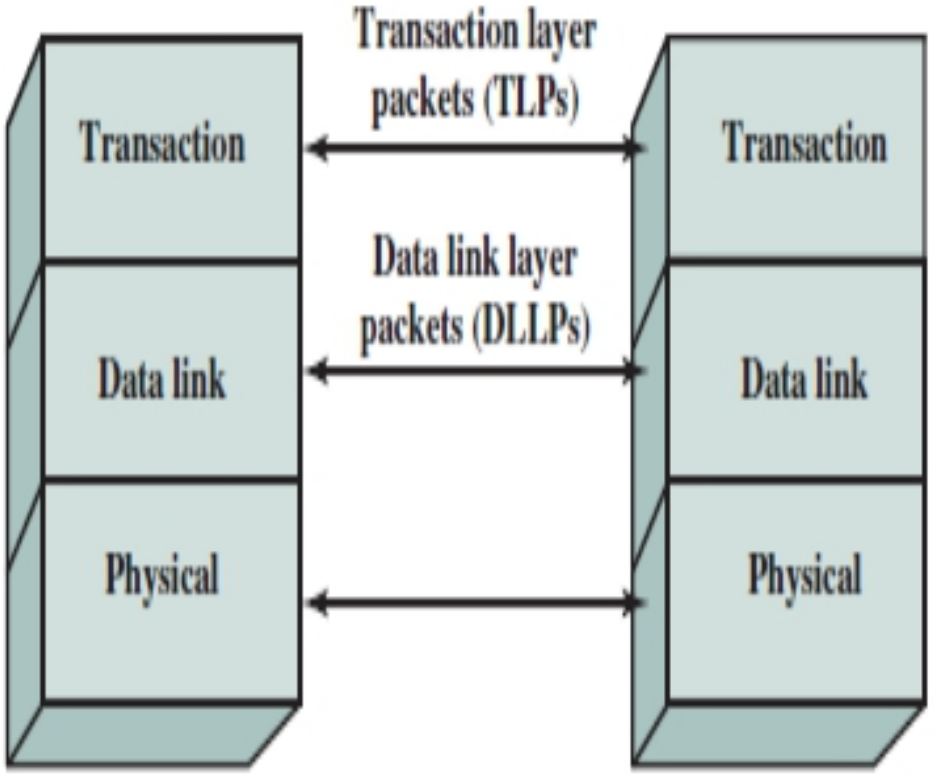


Fig.2. PCIe Protocol Layers

PCIe Physical Layer

- PCIe is a point-to-point architecture.
- Each PCIe port consists of a number of bidirectional lanes where as in QPI, the lane refers to transfer in one direction only.
- Transfer in each direction in a lane is by means of differential signaling over a pair of wires.
- A PCI port can provide 1, 4, 6, 16, or 32 lanes.

PCIe Multilane Distribution

- PCIe uses a multilane distribution technique.
- Fig.3. shows an example for a PCIe port consisting of four lanes.
- Data are distributed to the four lanes 1 byte at a time using a simple round-robin scheme.

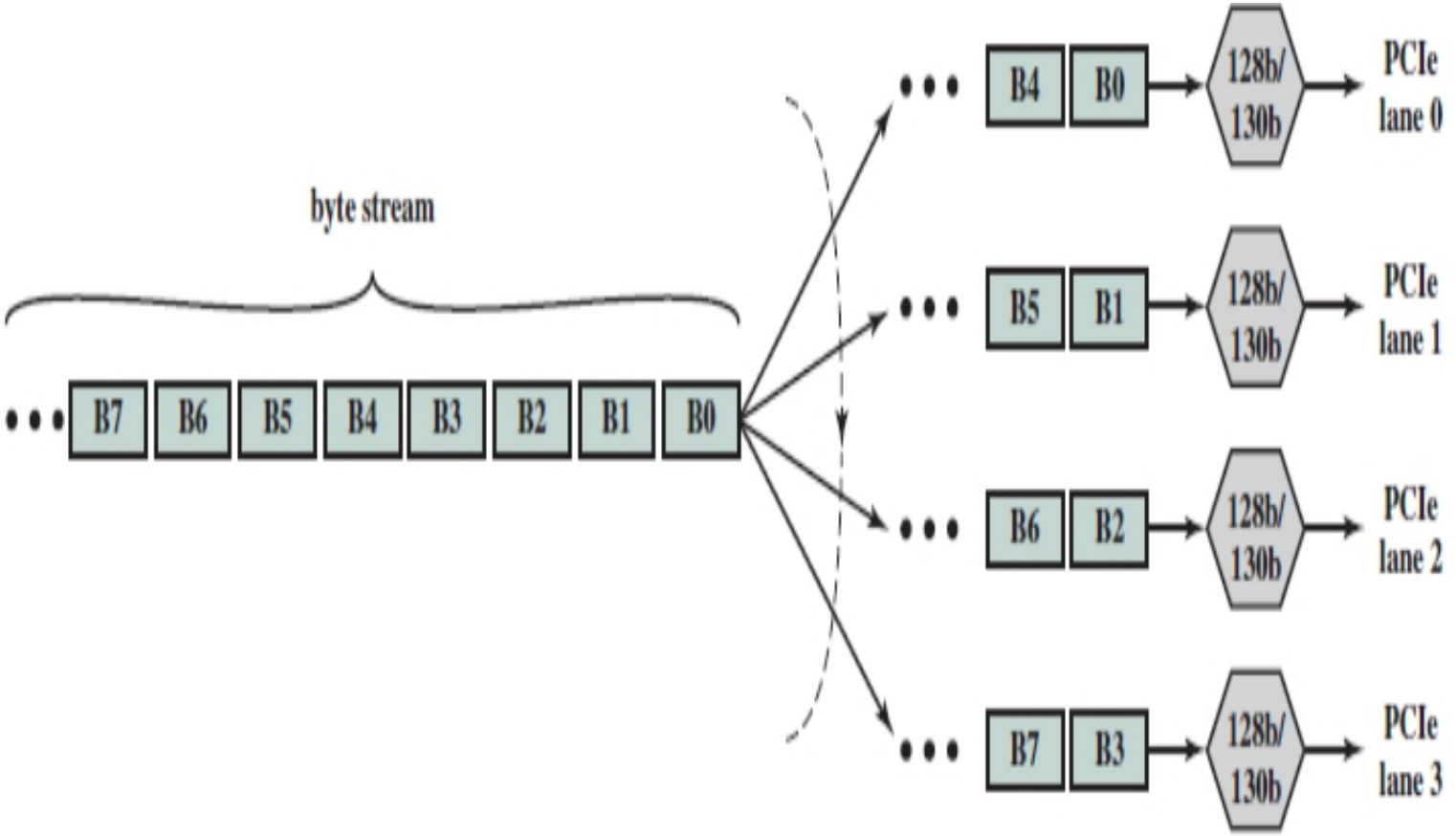


Fig.3. PCIe Multilane Distribution

Contd.

- At each physical lane, data are buffered and processed 16 bytes (128 bits) at a time.
- Each block of 128 bits is encoded into a unique 130-bit code word for transmission; this is referred to as 128b/130b encoding.
- Thus, the effective data rate of an individual lane is reduced by a factor of 128/130.

PCIe Transmit and Receive Block Diagrams

- Fig.4. illustrates the use of scrambling and encoding.
- Data to be transmitted are fed into a scrambler.
- The scrambled output is then fed into a 128b/130b encoder, which buffers 128 bits and then maps the 128-bit block into a 130-bit block.
- This block then passes through a parallel-to-serial converter and transmitted one bit at a time using differential signaling.

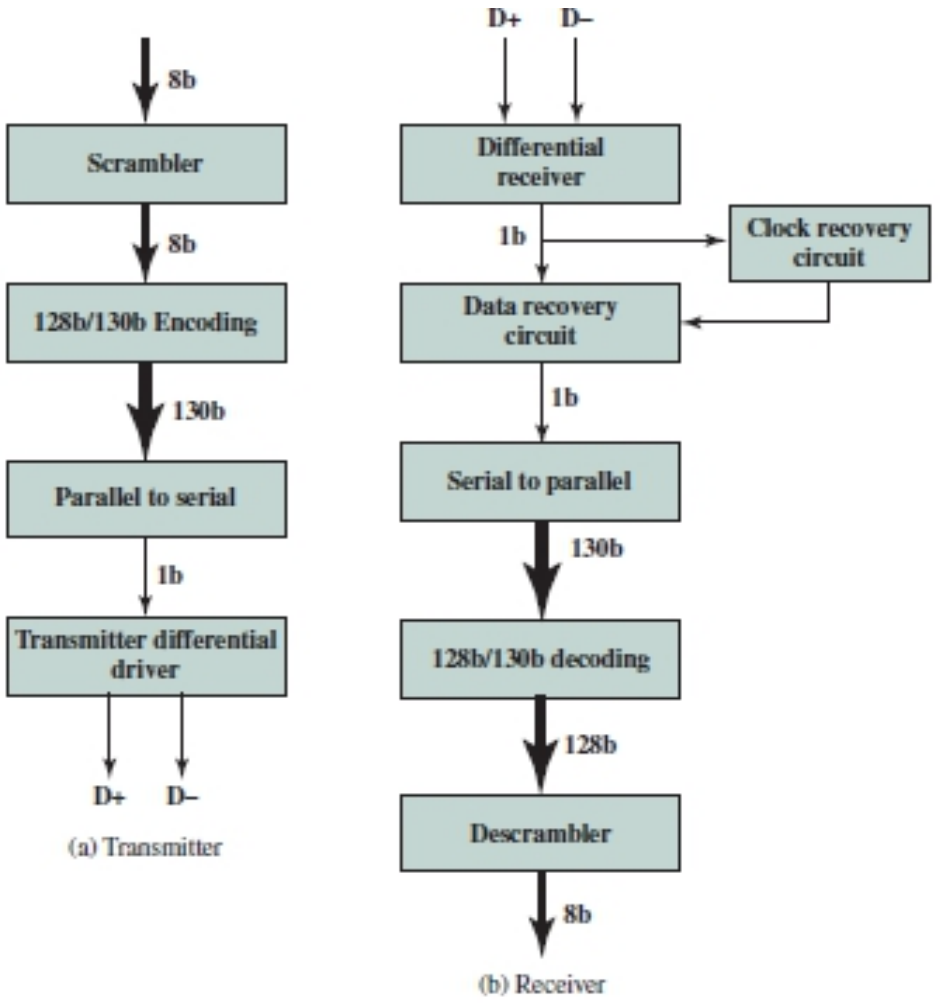


Fig.4. PCIe Transmit and Receive Block Diagrams

Contd.

- At the receiver, a clock is synchronized to the incoming data to recover the bit stream.
- This then passes through a serial-to-parallel converter to produce a stream of 130-bit blocks.
- Each block is passed through a 128b/130b decoder to recover the original scrambled bit pattern, which is then descrambled to produce the original bit stream.
- Using these techniques, a data rate of 16 GB/s can be achieved.
- Each transmission of a block of data over a PCI link begins and ends with an 8-bit framing sequence intended to give the receiver time to synchronize with the incoming physical layer bit stream.

PCIe Transaction Layer

- The transaction layer (TL) receives read and write requests from the software above the TL and creates request packets for transmission to a destination via the link layer.
- Most transactions use a *split transaction* technique, which works in the following manner:
- A request packet is sent out by a source PCIe device, which then waits for a response, called a *completion* packet.
- The completion following a request is initiated by the completer only when it has the data and/or status ready for delivery.
- Each packet has a unique identifier that enables completion packets to be directed to the correct originator.
- With the split transaction technique, the completion is separated in time from the request, in contrast to a typical bus operation in which both sides of a transaction must be available to seize and use the bus.

Address spaces and transaction types

- The TL supports four address spaces:
- **Memory:** The memory space includes system main memory. It also includes PCIe I/O devices. Certain ranges of memory addresses map into I/O devices.

- **I/O:** This address space is used for legacy PCI devices, with reserved memory address ranges used to address legacy I/O devices.

- **Configuration:** This address space enables the TL to read/write configuration registers associated with I/O devices.

- **Message:** This address space is for control signals related to interrupts, error handling, and power management.

PCIe Data Link Layer

- The purpose of the PCIe data link layer is to ensure reliable delivery of packets across the PCIe link.

- The DLL participates in the formation of TLPs and also transmits DLLPs.

Data link layer packets

- Data link layer packets originate at the data link layer of a transmitting device and terminate at the DLL of the device on the other end of the link.

- There are three important groups of DLLPs used in managing a link: flow control packets, power management packets, and TLP ACK and NAK packets.

- Power management packets are used in managing power platform budgeting.

- Flow control packets regulate the rate at which TLPs and DLLPs can be transmitted across a link.

Transaction layer packet processing

- The DLL adds two fields to the core of the TLP created by the TL:
  - a 16-bit sequence number and a
  - 32-bit link-layer CRC (LCRC).

- Whereas the core fields created at the TL are only used at the destination TL, the two fields added by the DLL are processed at each intermediate node on the way from source to destination.

- When a TLP arrives at a device, the DLL strips off the sequence number and LCRC fields and checks the LCRC.

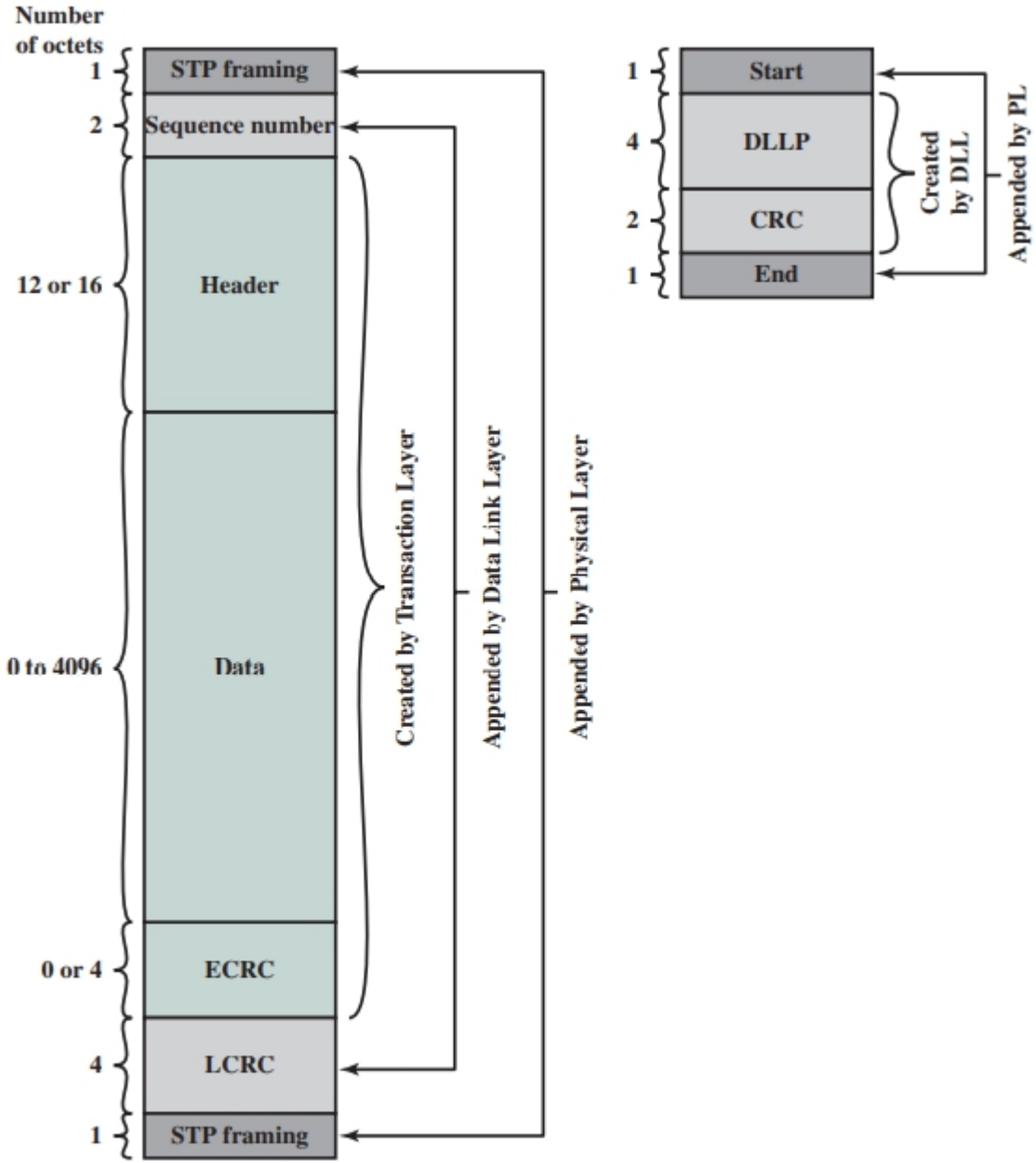


Figure 3.25 PCIe Protocol Data Unit Format

Review Questions

1. What general categories of functions are specified by computer instructions?

2. List and briefly define the possible states that define an instruction execution.

3. List and briefly define two approaches to dealing with multiple interrupts.

4. What types of transfers must a computer's interconnection structure (e.g., bus) support?

5. List and briefly define the QPI protocol layers.

6. List and briefly define the PCIe protocol layers.

Thank You !