03-01-2024

PCI, PCI-X, PCIe

System Architecture

Silpav @ cdac

x86 → ISA, EISA

Spark → SBUS

⟹ PCI

↓

common protocol

- 32 bit @ 33MHz → for Multimedia storage.
  - plug & play
  - bus mastering
  - architecture independent.

Latest upcomming:

PCIe-6.0 (CXL2)

→ Latency → time taken to respond

→ Bond Width - time taken to send particular data.

PCI

47+2

parallel bus use, Address, data → even parity

Interface control

Error reporting - PERR - parity error - (address & data)
- SERR - system error

** Master only ← [Arbitation : req gnt
System : clk rst

• uses hidden, overlap central arbitation

TRdy → target ready

IRdy → initiator ready

IDselect → to select range of Id address.

↳ happens in configuration space.

→ one inits config RD → read.

+ Req 64#
Ack 64# } for 64 bit

Frame,

→ sustained tristate signals.

LOCK → Read, modify, write (semaphore)

Mechanical form factor also specified, power consumption ⇒ PRSNT [1:0] → power
M66EN → 66 MHz

memory mapped I/O, I/O mapped I/O
↓ ↳ simple 1→1
large can be addressed

1169 → VendorID of CDAC PCIe

- Basic write, basic read → PCI
- Disconnect
  - ~~Retry~~ → with data, → one or more data accepted.
  - without data → TRDY not asserted.
- target abord → last data is discarded. → device select deasserted**

Configuring register → some readonly, write only, general register

PCI EX
64 bit 133 MHz, arbitration ↓, delay transaction replaced by split transactions.
removes wait phase → saves time
→ size of transaction

sequence number → seq n(tag) PCIE, 64 byte - standard block size.
↓
cache line size of CPUs

07-01-2024

PCIe

OSI model - 7 layer

PCIe - 3 layer ⇔ Transaction
Data link → reliable delivery of pack.
Physical

QoS → Quality of service → Same Physical channel → with priority

TC → VC mapping

⇒ with Virtual channel ...

- traffic ~~Virtual~~ class. → 8

usually 4 implemented.

Data link → DLLP → Data link layer packet

Transaction ... ...

Limitations of PCI

- wait states, - stalls bus.
- delay transaction → adds latency
- shared interrupt → delay resour
- video, audio fills bus
- routing all lanes.
- clk to data management.

Bus enumeration numbering bus id, <bus no, device no, function no>

done by root complex

↓

@ each transaction.

↳ TLP header

↓

Transaction layer pa

Physical layer manages the split in frame 1x, 4x, 8x, 16x } combine & send to data link layer.

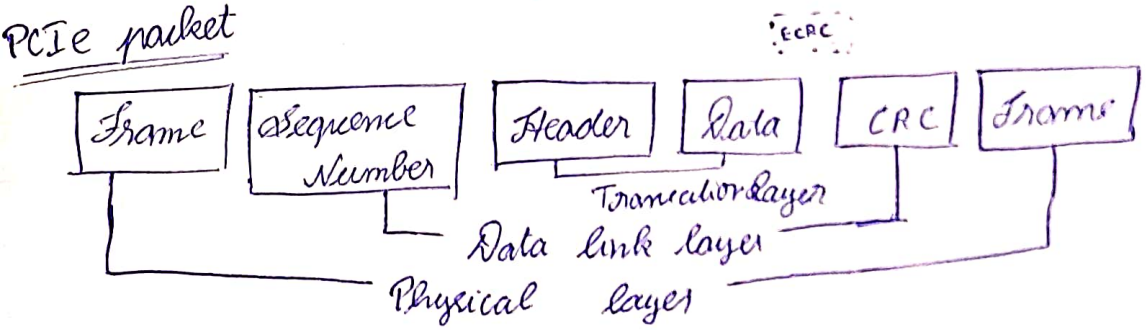
→ lane width, lane interchange, lane speed, TX/RX interchange, skew adjustment } ⇒ configuration

Data link layer buffer size, error detection ↳ ACK / NACK

↓

config. phase.

PCIe packet



RCB → 64byte / 128 byte. → minimum size of data size

MPS → Maximum payload size

4kB (or) 1k DW (Dword) [IEEE specified]

256 Byte (Typically)

Read

MRR → Maximm Read Request.

512 Byte (Typically)

Transfer are d-word → FBE First Byte enable
LBE Last Byte enable. } 4bit. used to denote valid bits.

Interrupt via MSI → Message signal interrupt.

Transation

posted → write, message, msg with data (except config write)

non-posted →
read, ...

Error correction

8B/10B
128b/130b
~~64B/66B~~

1.0V, 2.0V
3.0V, ...
↳ doubled bandwidth

2.0V 3.0V
5GT/s → 8GT/s

NACK → resend from Replay Buffer.

↳ Not done for every packet, depending on time, packet size is done

Express specific config register

PCI header,
PCI device specific New capacity

Express extended config space → optional.

Case study

Mem Read

Mem write.

Design partioning

Xilinx PCIe core.

Completer interface, → Slave decoder

Requestor interface. → Write, DMA Arbiter

Gemini, PCIe interface → Read, DMA, agents.

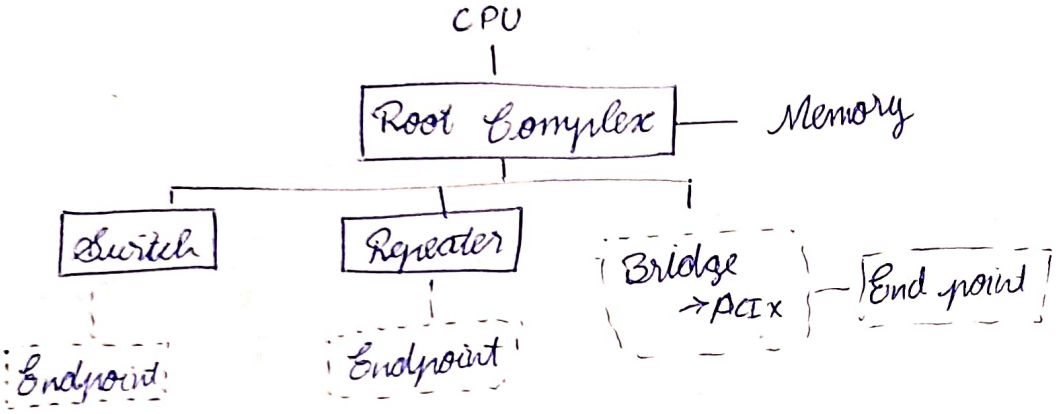
PCIe

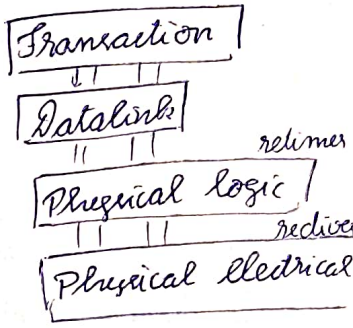
History

2003, 2007, 2010, 2017, 2019, 2021, 2025.. ⟶

1.0	2.0	3.0	4.0	5.0	6.0	7.0
2.5 GT/s	5 GT/s	8 GT/s	16 GT/s	32 GT/s	64 GT/s	

Topology





Root Complex either integrated to CPU (or) external

Repeater - signal conditioner

- retimer
- redriver

Retimer -

Endpoint Legacy, PCIe, Root complex integrated

Signals

PERST# → held low until power rails are stable → initial of transfer

WAKE# } transition ⟺ low power state
CLKREQ# }

} Embedded clocks with each device

Electrical characteristic

85Ω ±20% PCIe connector
100Ω ±20% chip to chip

adj traces 4 time dielectric height.

- AC coupling Cap in TX
  - close to finger rail
  - pads for cap cause impedance mismatch.

RX detect → Polling → Configuration → L0 state (PCIE 1.0V)

@2.5GT/s

Link training data

- varying channel length compensated
- link width determined
- lane polarity, lane speed.

de-skewing lane-lane

↳ recovery state

higher PCIe will beg.

→ link equalization done

PHASE 0

PHASE 1

PHASE 2

PHASE 3 → 10^{-12} error rate

Signal Conditioning

1. Protocol aware retimer - removes jitter,
2. Protocol agnostic linear repeater

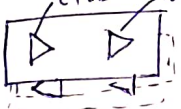
Max loss / signal attenuation

PCIe 5.0 – 35dB

Retimers -

CTLE – Continious time linear equalization

1. Analog only → redriver

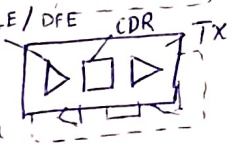


+100ps latency

2. Retimer

CDR - Clock & Data recovery system

DFE - Decision feedback equilizer.



+64ns Latency

LTSSM - Link training & status state machine

- Negotiate link speed
- Fix signal conditioning parameters
- Enter compliant mode for different data rates
- Link recovery
- Enter L0, L1, L2, L3 & L0s powersave

→ FSM Referti video

DLCMSM

Data link control and management statemachine

- error detection & retry, power management