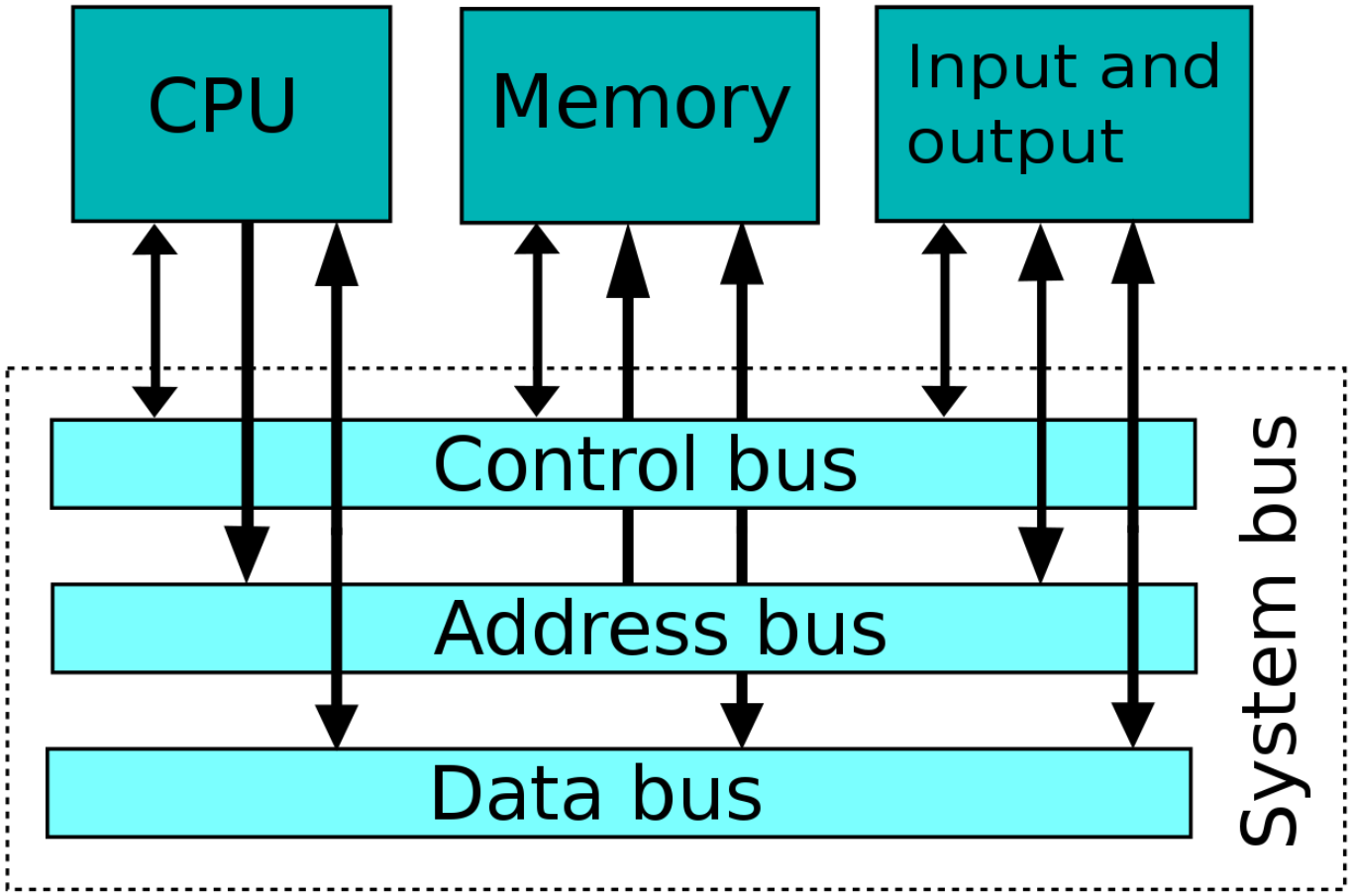
Contemporary Computer Architecture TDSN13

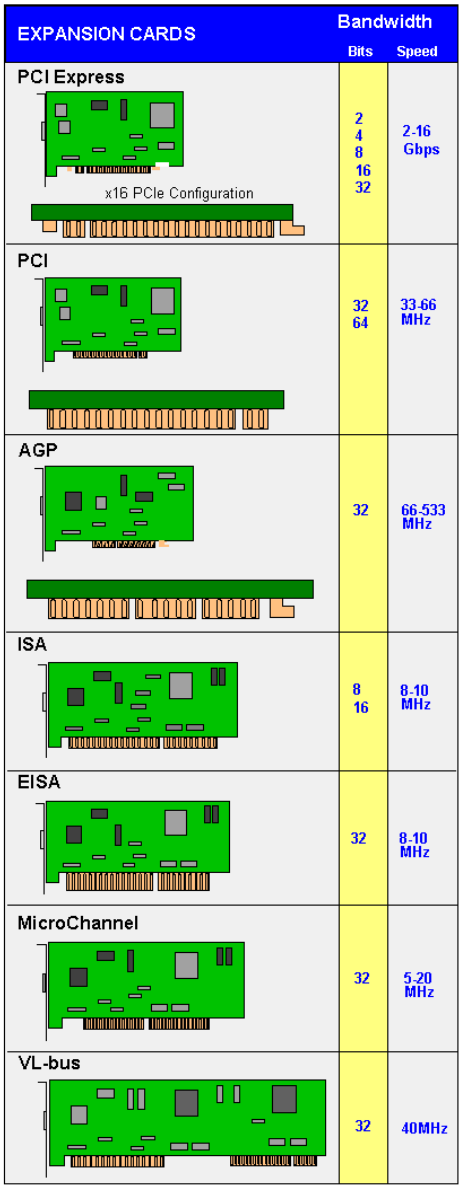
LECTURE 10 – HIGH PERFORMANCE BUSSES

ANDREAS AXELSSON (ANDREAS.AXELSSON@JU.SE)

Data Busses


CPU
Memory
Input and output
Control bus
Address bus
Data bus
System bus

Data busses in a PC



Current:

PCI Express - (Parallel/Serial)

PCI Express (PCIe) is the current bus interface, superseding PCI. All new desktop PCs and Macs have PCIe slots.

USB - (Serial)

Permanently or temporarily attach almost anything (flash drives, hard drives, printers, phones, cameras, etc.).

Previous:

PCI - (Parallel)

PCI was popular in all hardware platforms but was superseded by PCI Express (PCIe). Transition motherboards may have one PCI slot.

FireWire - (Serial)

Mostly used for digital camera connections. Popularized by Apple, adapters were required to use FireWire on new Macs.

AGP - (Parallel)

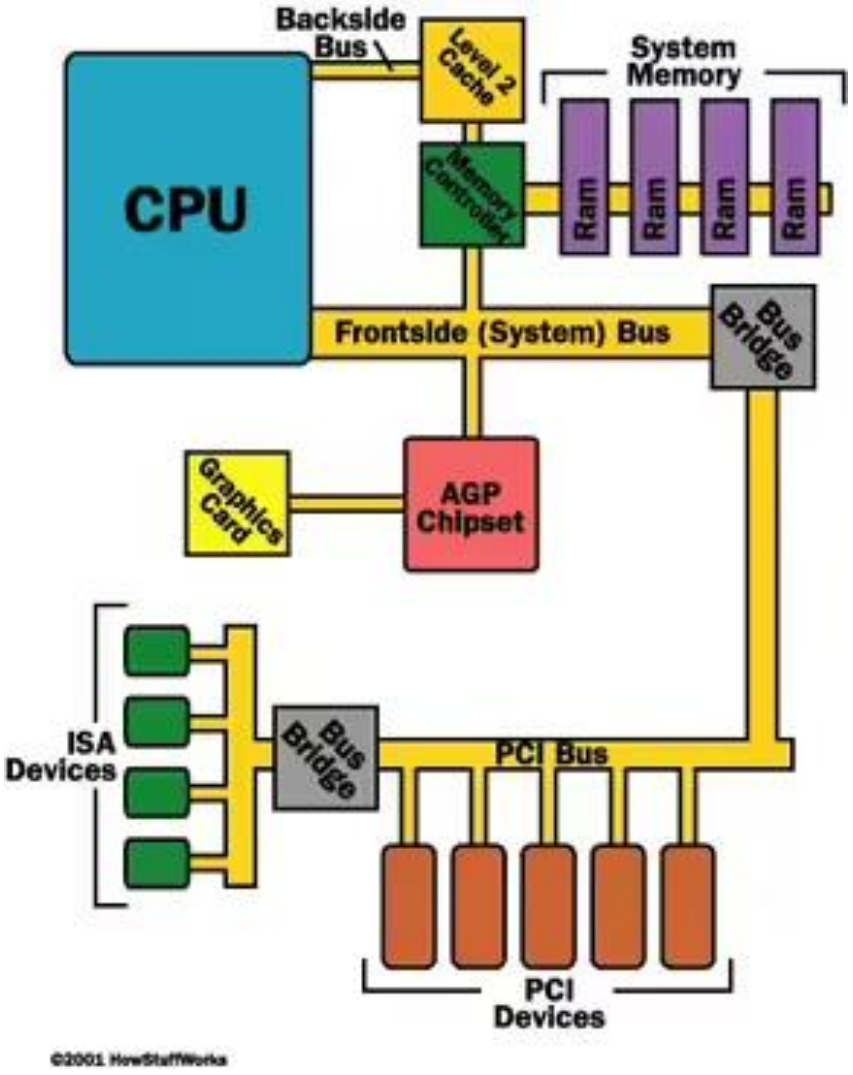
The graphics interface between PCI and PCI Express. AGP was faster than PCI and freed up a PCI slot.

ISA - (Parallel)

Pronounced "**eye**-suh" and debuting on the IBM PC AT, ISA was the evolution of the first PC bus in 1981.

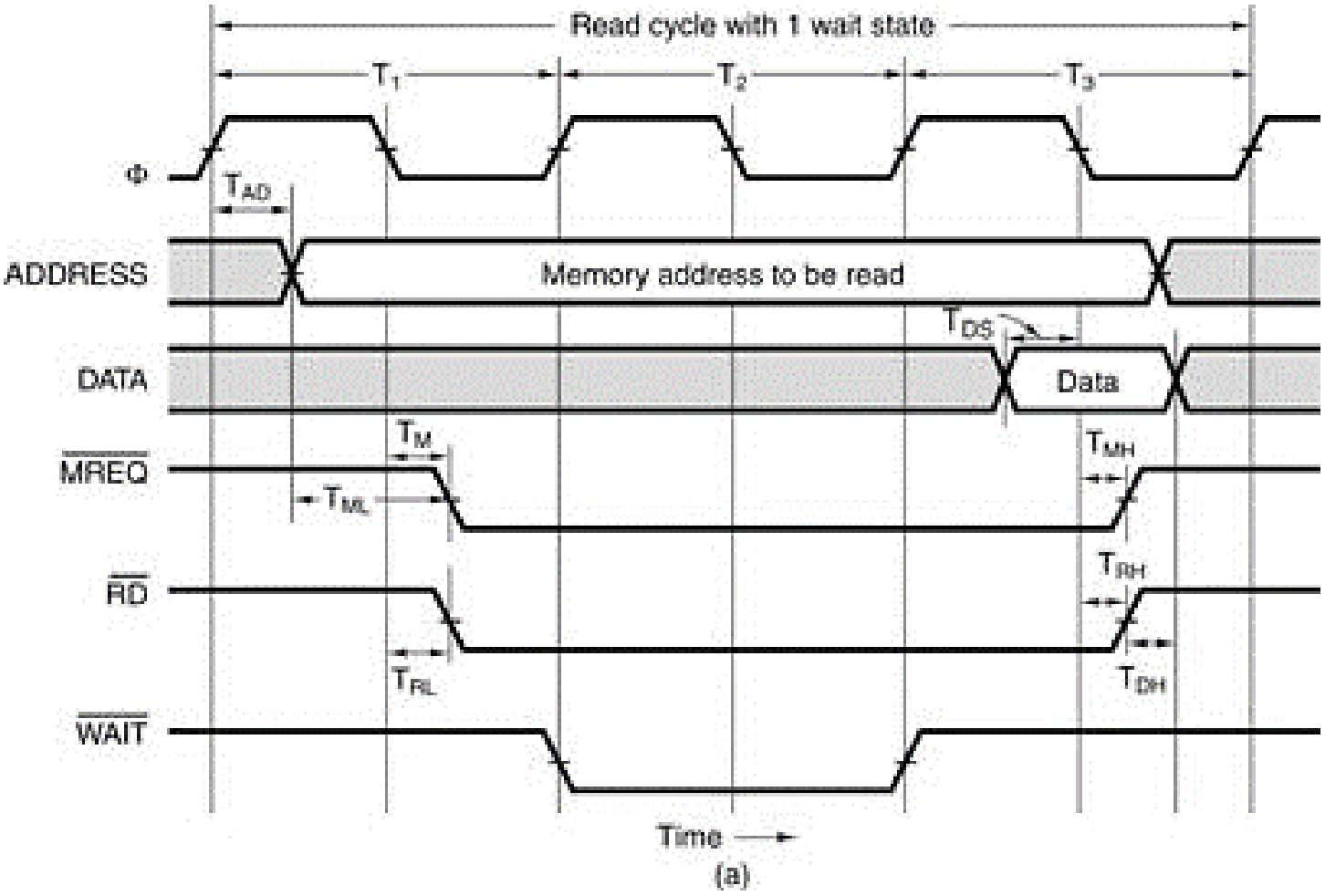
https://www.pcmag.com/encyclopedia/term/pc-data-buses

Legacy busses (ISA / PCI / AGP)

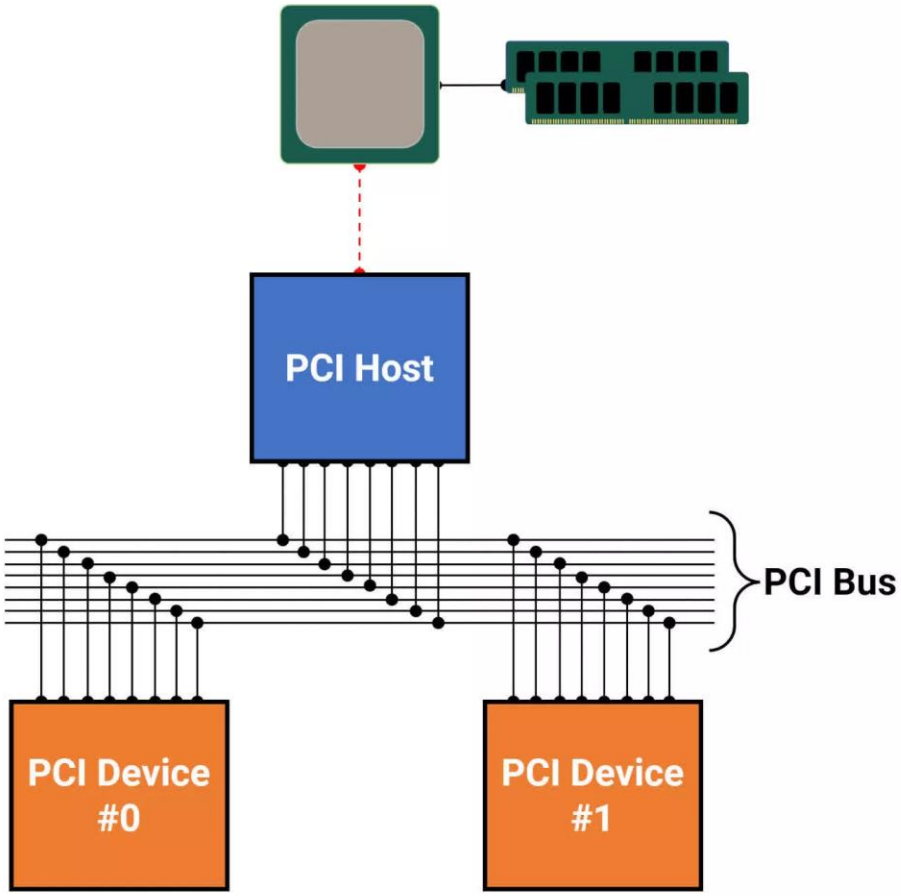


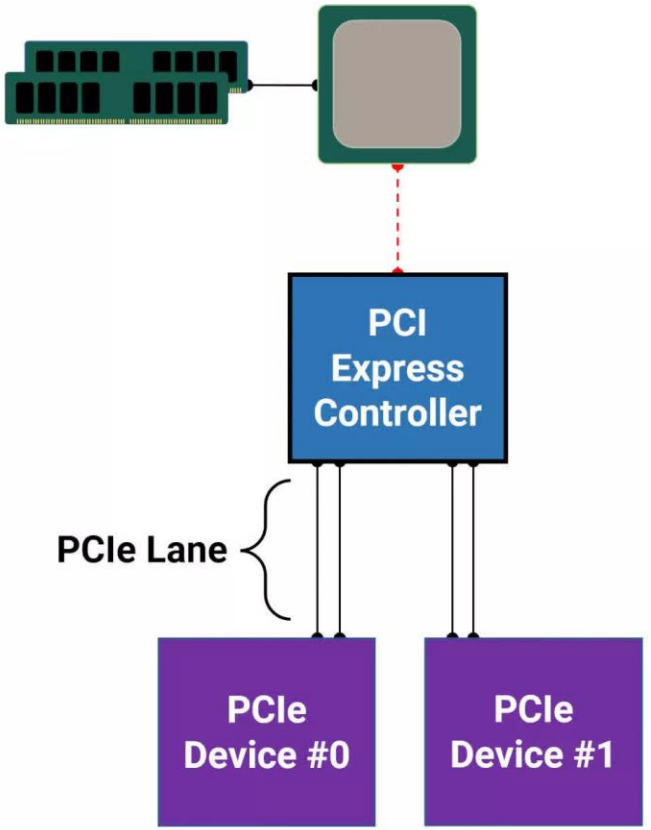
Bus Type	Bus Width	Bus Speed	MB/sec
ISA	16 bits	8 MHz	16 MBps
EISA	32 bits	8 MHz	32 MBps
VL-bus	32 bits	25 MHz	100 MBps
VL-bus	32 bits	33 MHz	132 MBps
PCI	32 bits	33 MHz	132 MBps
PCI	64 bits	33 MHz	264 MBps
PCI	64 bits	66 MHz	512 MBps
PCI	64 bits	133 MHz	1 GBps

Parallel synchronous busses

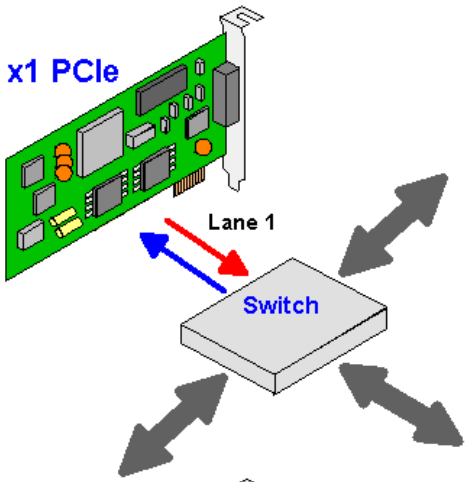


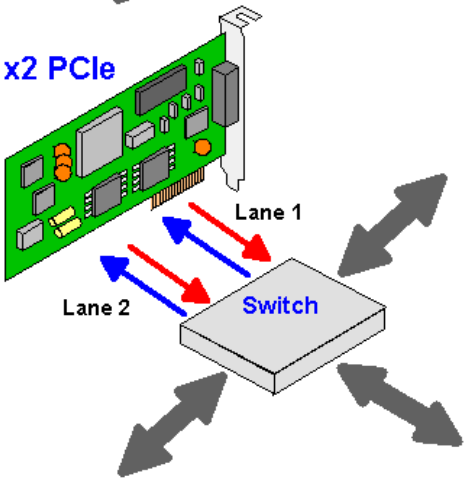
PCI vs PCIe

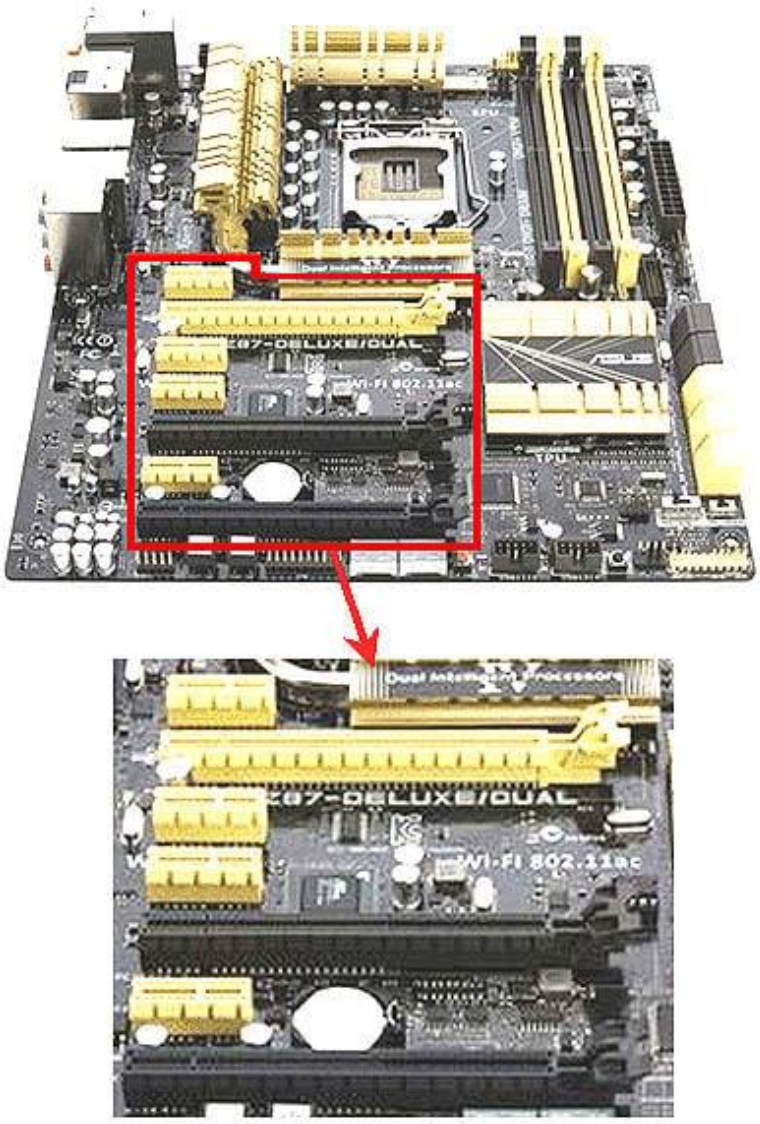

PCI Host
PCI Bus
PCI Device #0
PCI Device #1


PCI Express Controller
PCIe Lane
PCIe Device #0
PCIe Device #1

High speed serial busses (PCIe etc)

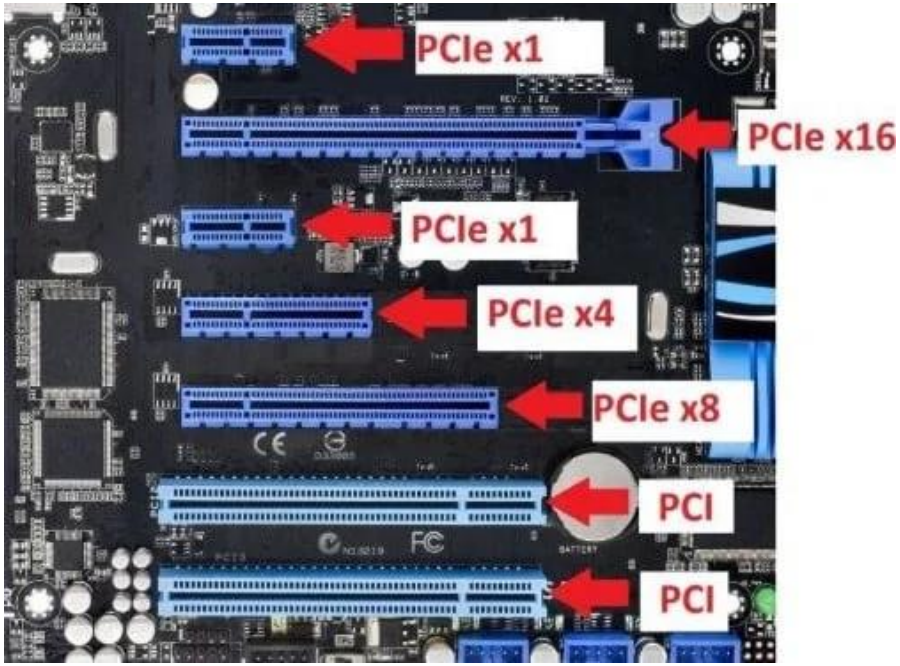






https://www.pcmag.com/encyclopedia/term/pci-express

PCI Express



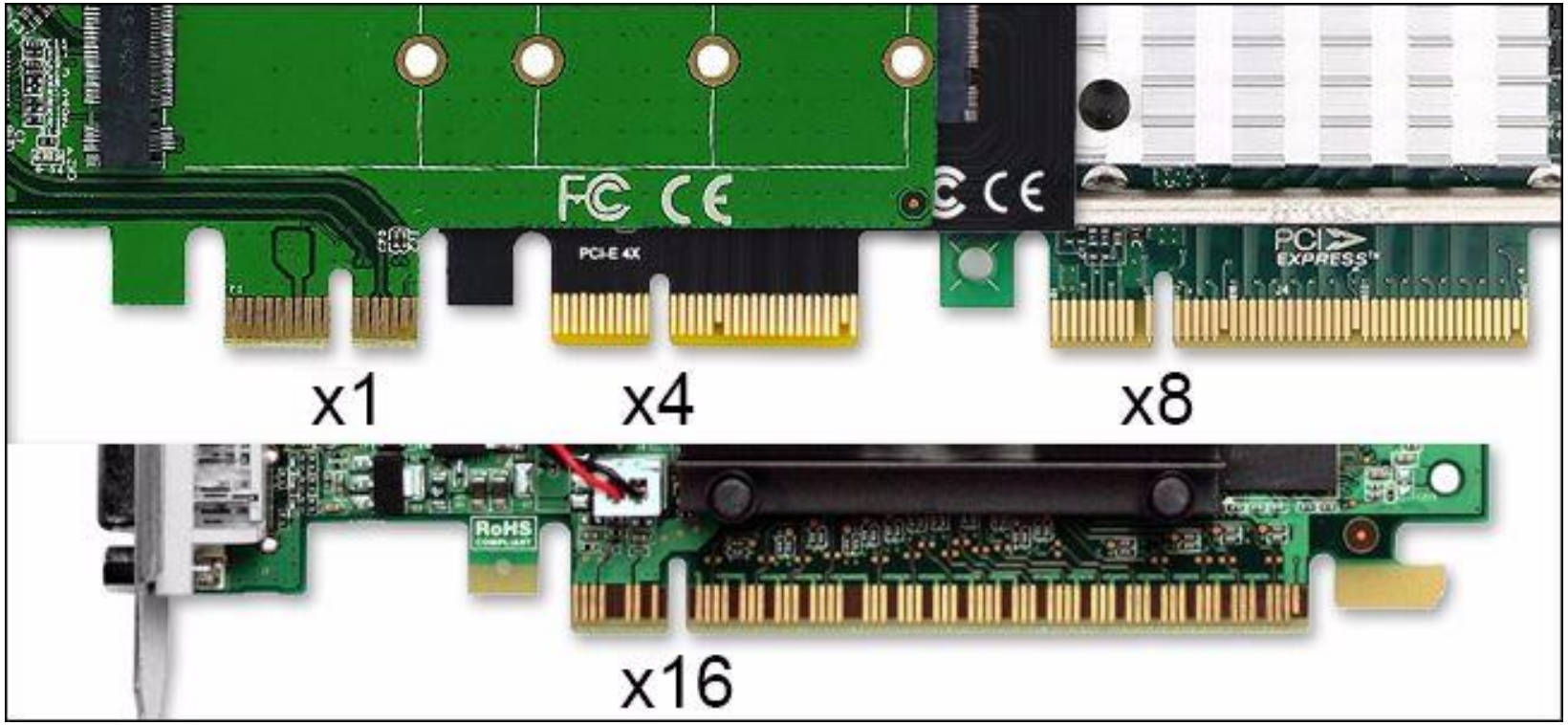
PCI Express connector pinout (x1, x4, x8 and x16 variants)

Pin	Side B	Side A	Description
1	+12 V	PRSNT1#	Must connect to farthest PRSNT2# pin
2	+12 V	+12 V	Main power pins
3	+12 V	+12 V	
4	Ground	Ground	
5	SMCLK	TCK	SMBus and JTAG port pins
6	SMDAT	TDI	
7	Ground	TDO	
8	+3.3 V	TMS	
9	TRST#	+3.3 V	
10	+3.3 V aux	+3.3 V	Aux power & Standby power
11	WAKE#	PERST#	Link reactivation; fundamental reset [23]
Key notch			
12	CLKREQ# [24]	Ground	Clock Request Signal
13	Ground	REFCLK+	Reference clock differential pair
14	HSOp(0)	REFCLK−	Lane 0 transmit data, + and −
15	HSOn(0)	Ground	
16	Ground	HSIp(0)	Lane 0 receive data, + and −
17	PRSNT2#	HSIn(0)	
18	Ground	Ground	
PCI Express x1 cards end at pin 18			
19	HSOp(1)	Reserved	Lane 1 transmit data, + and −
20	HSOn(1)	Ground	
21	Ground	HSIp(1)	Lane 1 receive data, + and −
22	Ground	HSIn(1)	
23	HSOp(2)	Ground	Lane 2 transmit data, + and −
24	HSOn(2)	Ground	
25	Ground	HSIp(2)	Lane 2 receive data, + and −
26	Ground	HSIn(2)	
27	HSOp(3)	Ground	Lane 3 transmit data, + and −
28	HSOn(3)	Ground	
29	Ground	HSIp(3)	Lane 3 receive data, + and −
30	PWRBRK# [25]	HSIn(3)	"Power brake", active-low to reduce device power
31	PRSNT2#	Ground	
32	Ground	Reserved	
PCI Express x4 cards end at pin 32			
33	HSOp(4)	Reserved	Lane 4 transmit data, + and −
34	HSOn(4)	Ground	
35	Ground	HSIp(4)	Lane 4 receive data, + and −
36	Ground	HSIn(4)	
37	HSOp(5)	Ground	Lane 5 transmit data, + and −
38	HSOn(5)	Ground	
39	Ground	HSIp(5)	Lane 5 receive data, + and −
40	Ground	HSIn(5)	
41	HSOp(6)	Ground	Lane 6 transmit data, + and −
42	HSOn(6)	Ground	
43	Ground	HSIp(6)	Lane 6 receive data, + and −
44	Ground	HSIn(6)	
45	HSOp(7)	Ground	Lane 7 transmit data, + and −
46	HSOn(7)	Ground	
47	Ground	HSIp(7)	Lane 7 receive data, + and −
48	PRSNT2#	HSIn(7)	
49	Ground	Ground	
PCI Express x8 cards end at pin 49			

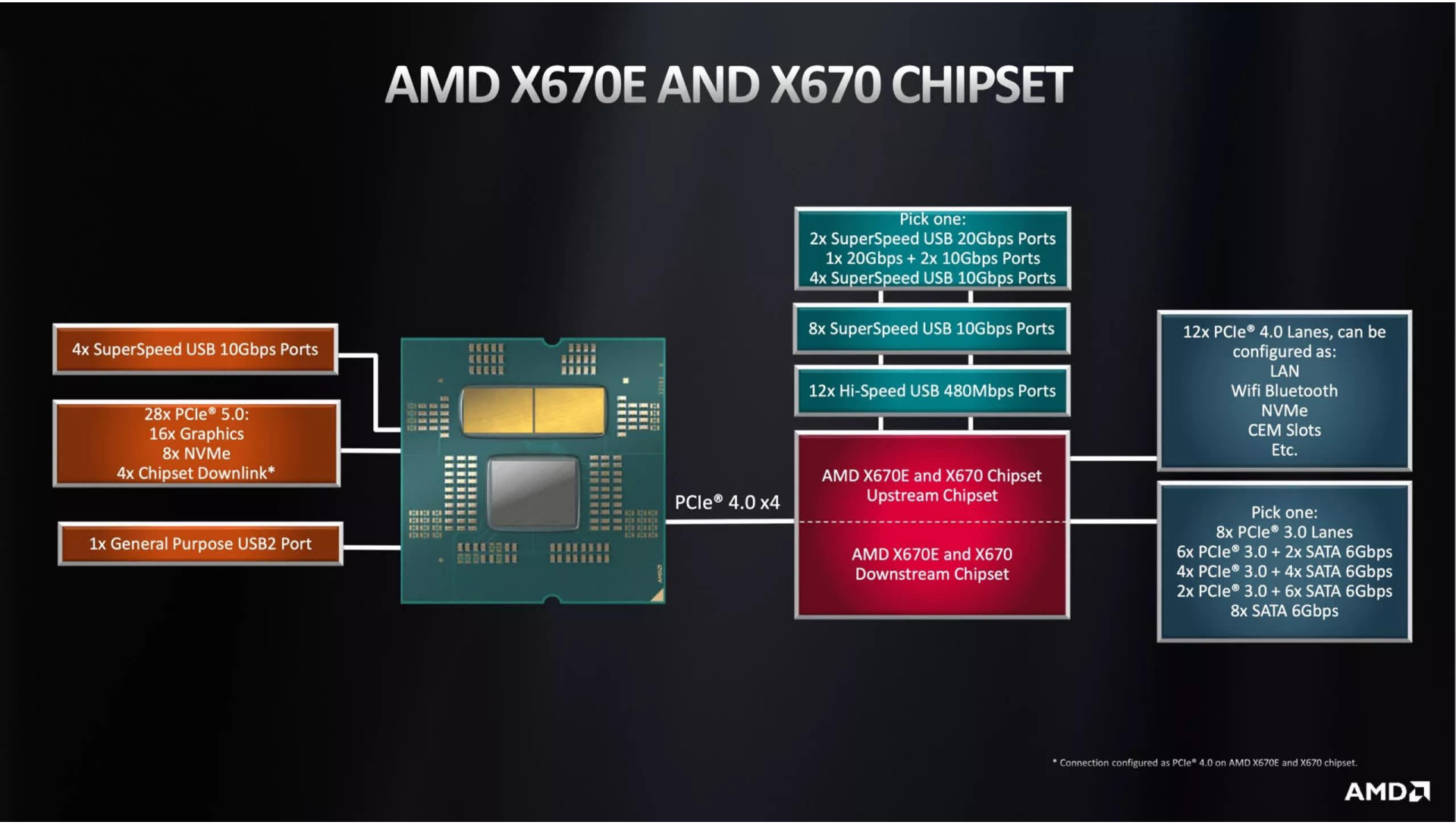
Pin	Side B	Side A	Description
50	HSOp(8)	Reserved	Lane 8 transmit data, + and −
51	HSOn(8)	Ground	
52	Ground	HSIp(8)	Lane 8 receive data, + and −
53	Ground	HSIn(8)	
54	HSOp(9)	Ground	Lane 9 transmit data, + and −
55	HSOn(9)	Ground	
56	Ground	HSIp(9)	Lane 9 receive data, + and −
57	Ground	HSIn(9)	
58	HSOp(10)	Ground	Lane 10 transmit data, + and −
59	HSOn(10)	Ground	
60	Ground	HSIp(10)	Lane 10 receive data, + and −
61	Ground	HSIn(10)	
62	HSOp(11)	Ground	Lane 11 transmit data, + and −
63	HSOn(11)	Ground	
64	Ground	HSIp(11)	Lane 11 receive data, + and −
65	Ground	HSIn(11)	
66	HSOp(12)	Ground	Lane 12 transmit data, + and −
67	HSOn(12)	Ground	
68	Ground	HSIp(12)	Lane 12 receive data, + and −
69	Ground	HSIn(12)	
70	HSOp(13)	Ground	Lane 13 transmit data, + and −
71	HSOn(13)	Ground	
72	Ground	HSIp(13)	Lane 13 receive data, + and −
73	Ground	HSIn(13)	
74	HSOp(14)	Ground	Lane 14 transmit data, + and −
75	HSOn(14)	Ground	
76	Ground	HSIp(14)	Lane 14 receive data, + and −
77	Ground	HSIn(14)	
78	HSOp(15)	Ground	Lane 15 transmit data, + and −
79	HSOn(15)	Ground	
80	Ground	HSIp(15)	Lane 15 receive data, + and −
81	PRSNT2#	HSIn(15)	
82	Reserved	Ground	

Legend	
Ground pin	Zero volt reference
Power pin	Supplies power to the PCIe card
Card-to-host pin	Signal from the card to the motherboard
Host-to-card pin	Signal from the motherboard to the card
Open drain	May be pulled low or sensed by multiple cards
Sense pin	Tied together on card
Reserved	Not presently used, do not connect

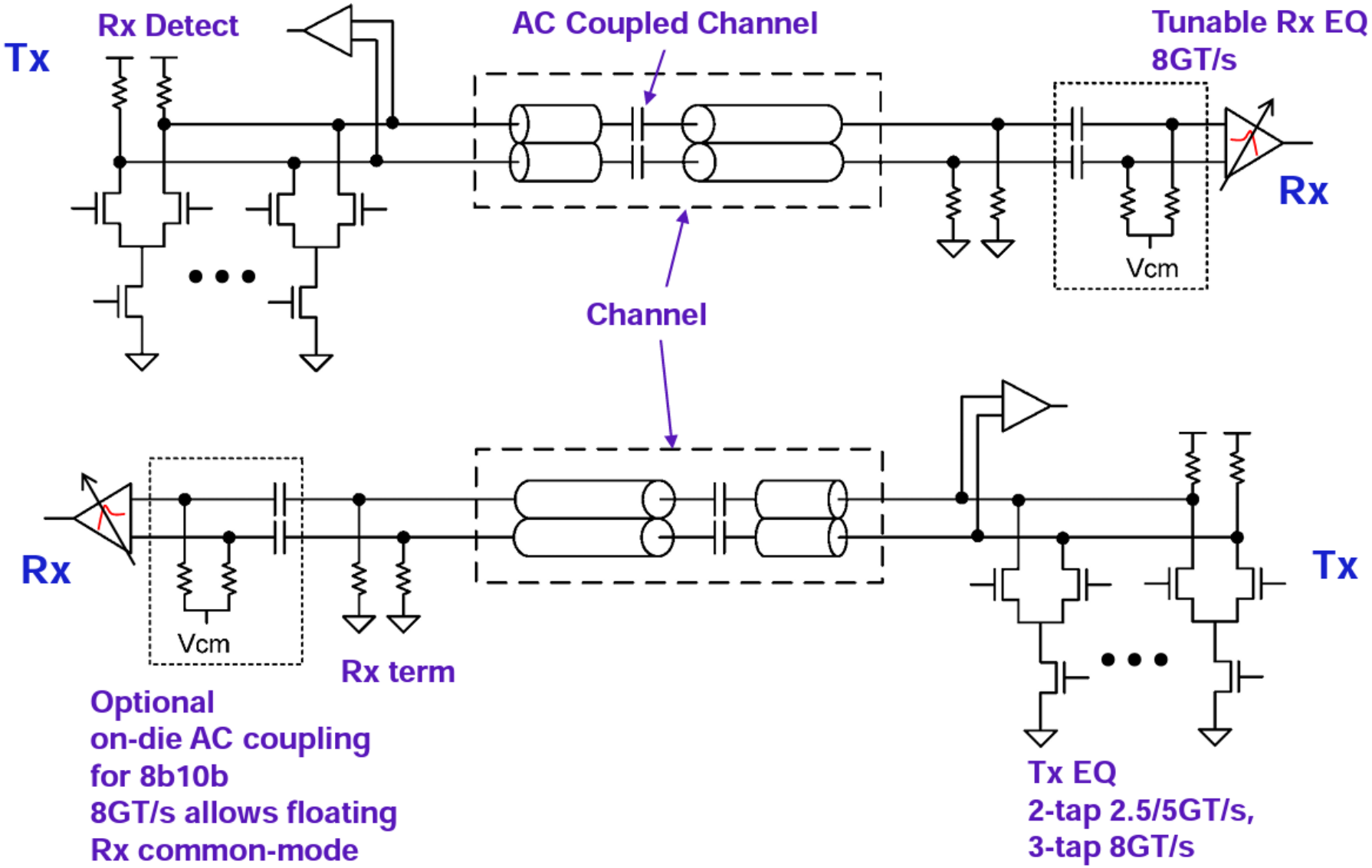
PCI Express



Structure of modern computer

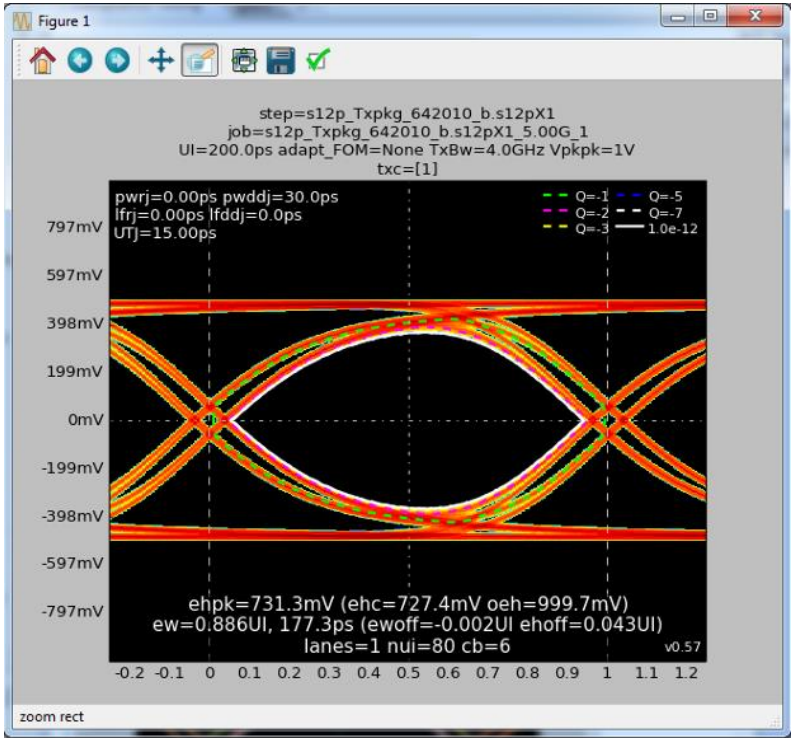


High speed serial busses (PCIe etc)



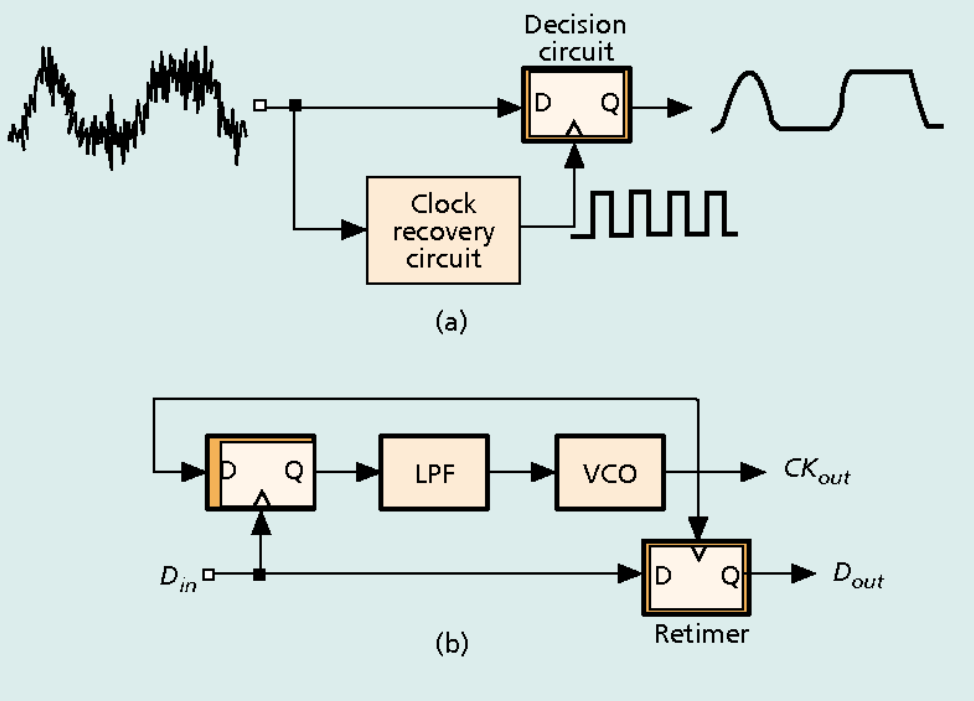
High speed serial busses (PCIe etc)

- Uses differential signaling (LVDS)
  - VP/VN-pair with +-400mV differential voltage
- Simliar tech used for USB, DisplayPort, HDMI
- Embed clock signal into bitstream (uses 8B/10B encoding etc)
  - Also equalizes 1's and 0's (in average same number of 0 and 1 transmitted)



High speed serial busses (PCIe etc)

- Uses differential signaling (LVDS)
  - VP/VN-pair with +-400mV differential voltage
- Simliar tech used for USB, DisplayPort, HDMI
- Embed clock signal into bitstream (uses 8B/10B encoding etc)
  - Also equalizes 1's and 0's (in average same number of 0 and 1 transmitted)

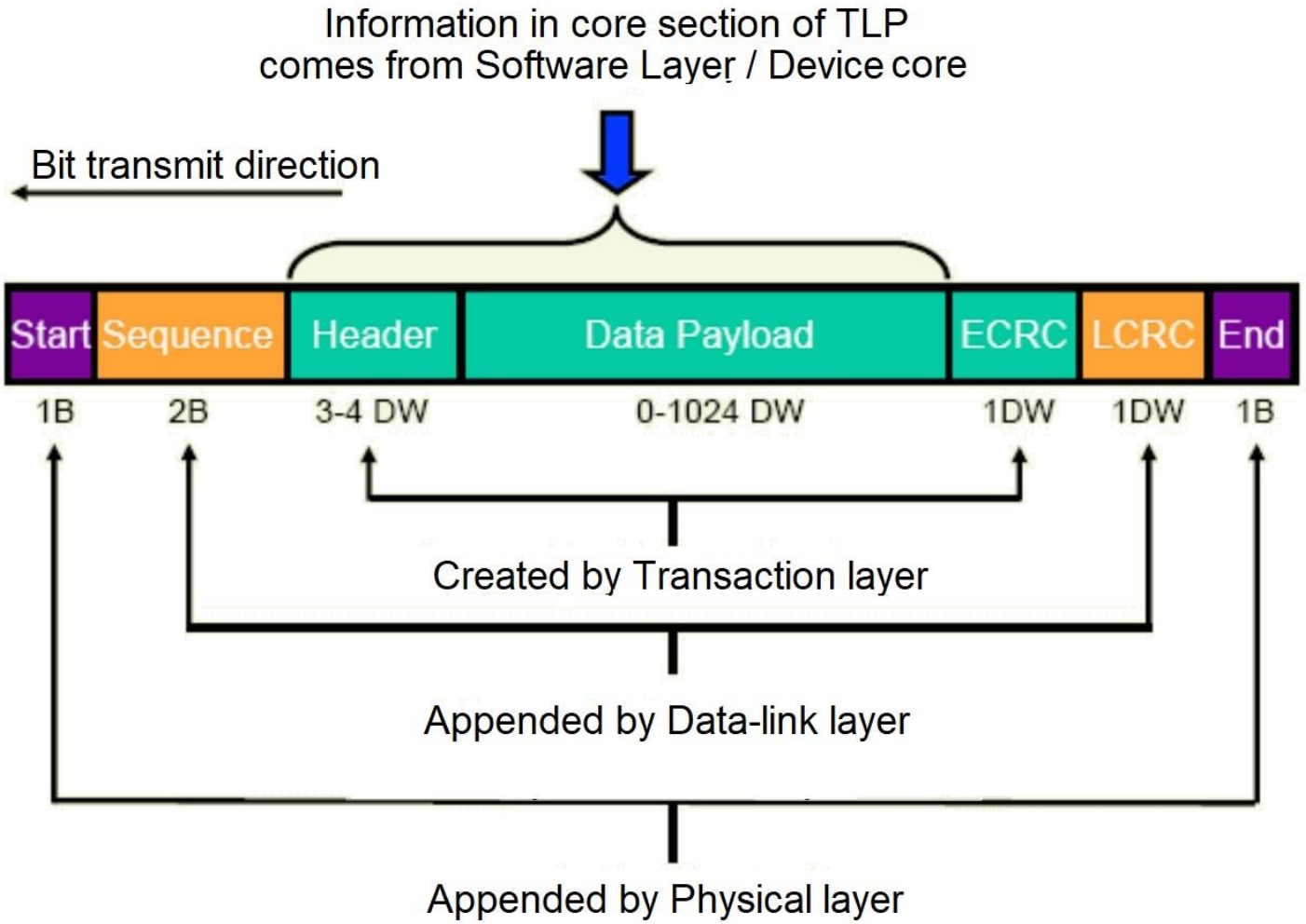


High speed serial busses (PCIe etc)

PCI Express: Unidirectional Bandwidth in x1 and x16 Configurations				
Generation	**Year of Release**	**Data Transfer Rate**	**Bandwidth x1**	**Bandwidth x16**
PCIe 1.0	2003	2.5 GT/s	250 MB/s	4.0 GB/s
PCIe 2.0	2007	5.0 GT/s	500 MB/s	8.0 GB/s
PCIe 3.0	2010	8.0 GT/s	1 GB/s	16 GB/s
PCIe 4.0	2017	16 GT/s	2 GB/s	32 GB/s
PCIe 5.0	2019	32 GT/s	4 GB/s	64 GB/s
PCIe 6.0	2021	64 GT/s	8 GB/s	128 GB/s

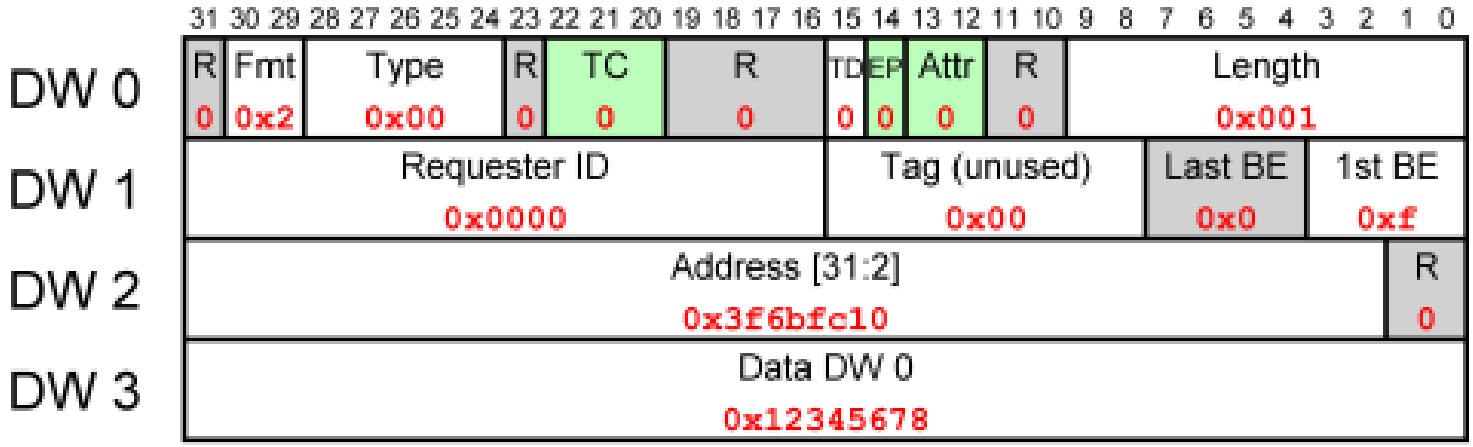
Table: PCI-SIG introduced the first generation of PCI Express in 2003. With each new generation comes a doubling of data transfer rate and total bandwidth per lane configuration, the latter of which is expressed in both unidirectional and bidirectional measurements, depending on the source. To find the total unidirectional bandwidth for each lane configuration, simply multiply the x1 bandwidths listed in the table above by two, four, eight or 16. Multiply the number resulting from that calculation by two to calculate total bidirectional bandwidth. *Source: PCI-SIG*

High speed serial busses (PCIe etc)



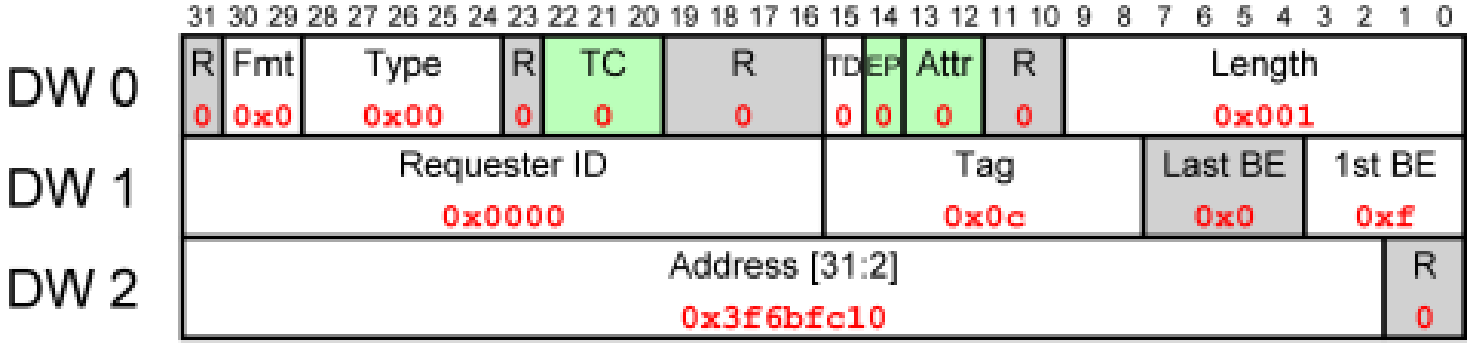
PCIe Protocol – Transaction Layer

Write Packet TLP



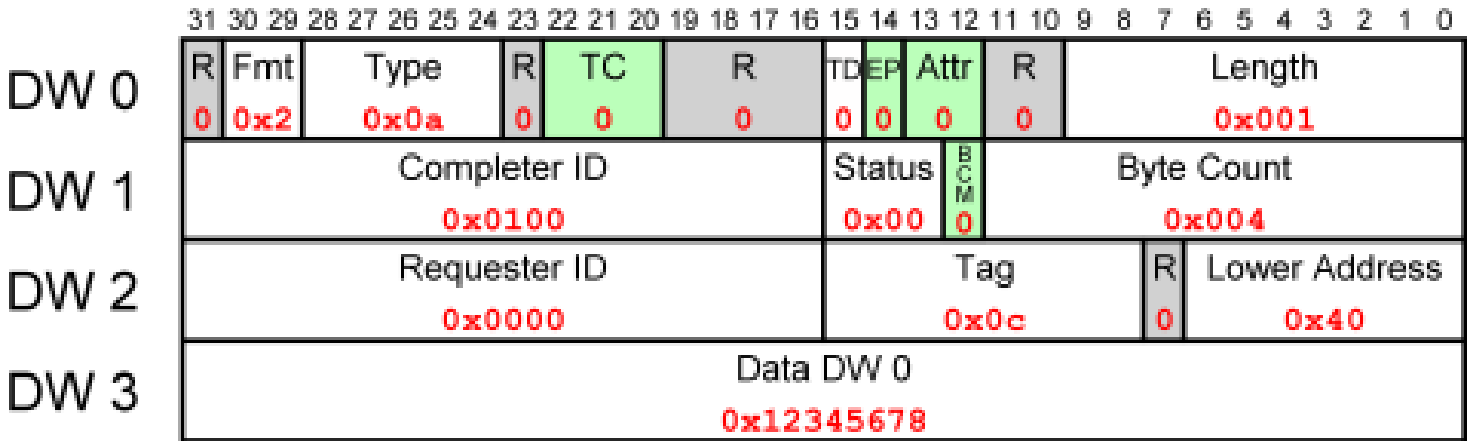
PCIe Protocol – Transaction Layer

Read Request TLP



PCIe Protocol – Transaction Layer

Completion TLP



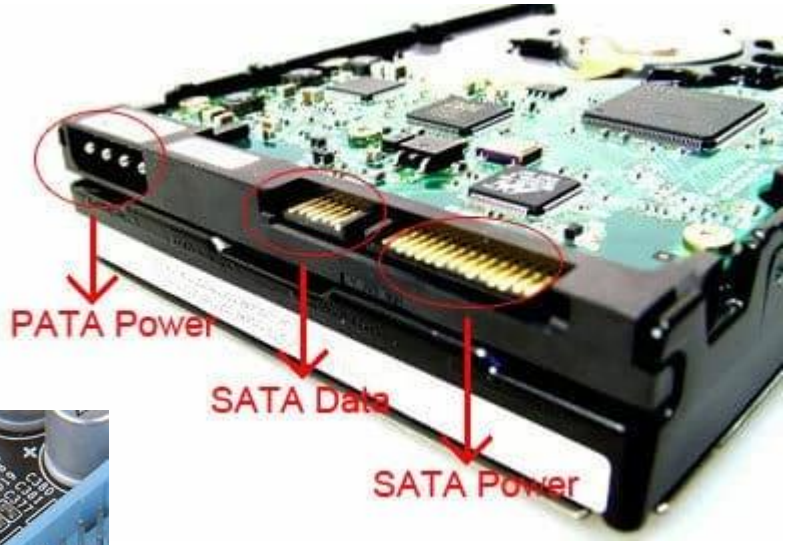
Hard disk interfaces

Hard disk interfaces – SATA

SATA (Serial AT Attachment)

600 MB/s

Very common in computers still

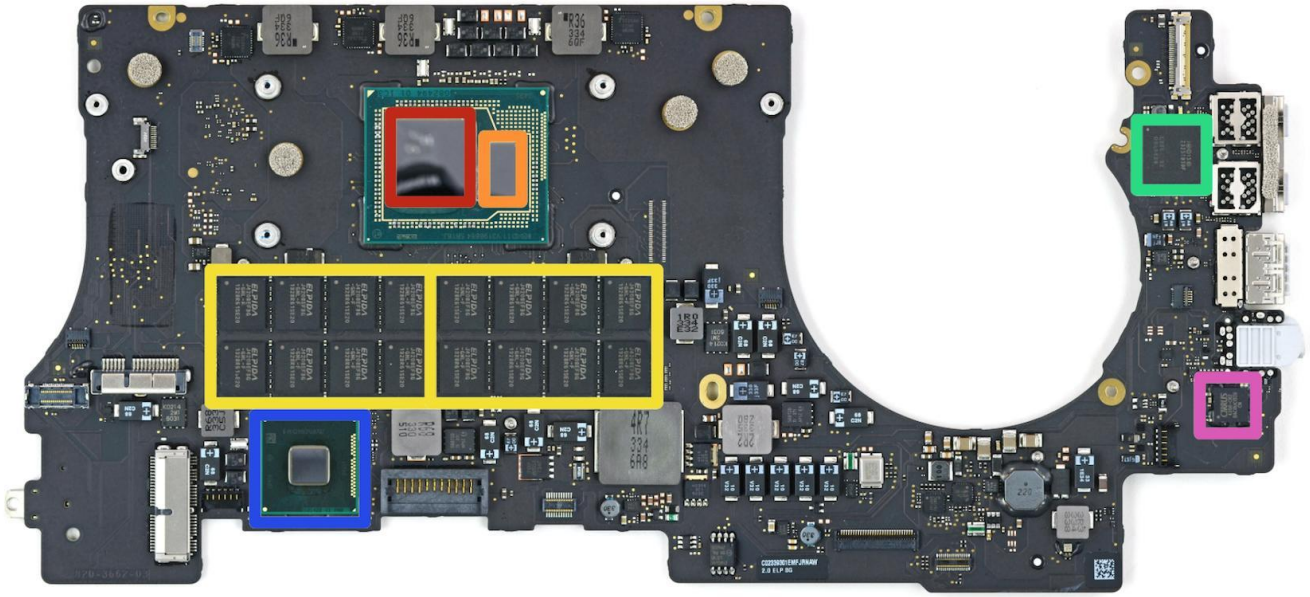


Hard disk interfaces – NVMe

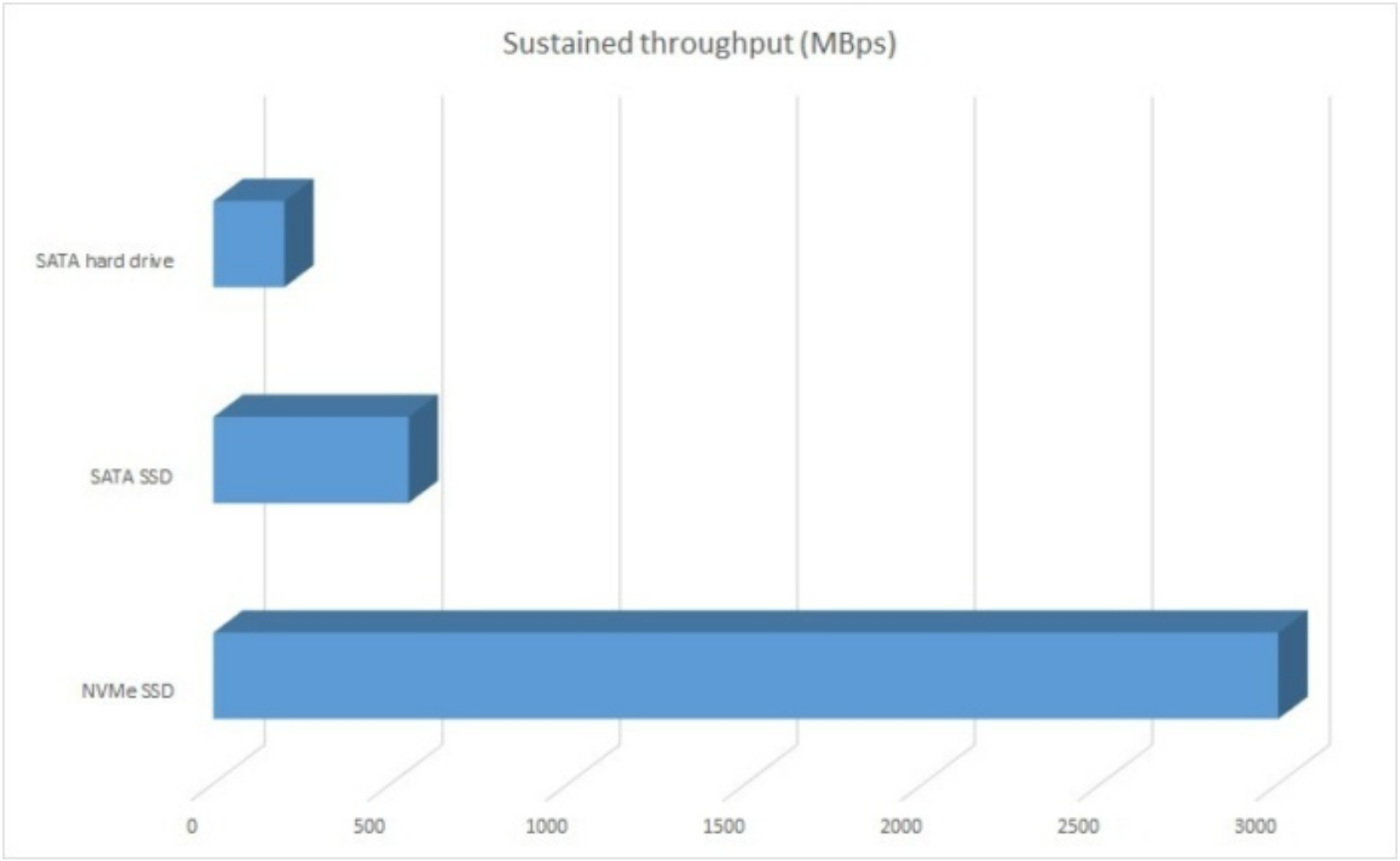
Uses PCIe bus to transfer data

In many modern laptops soldered to motherboard.

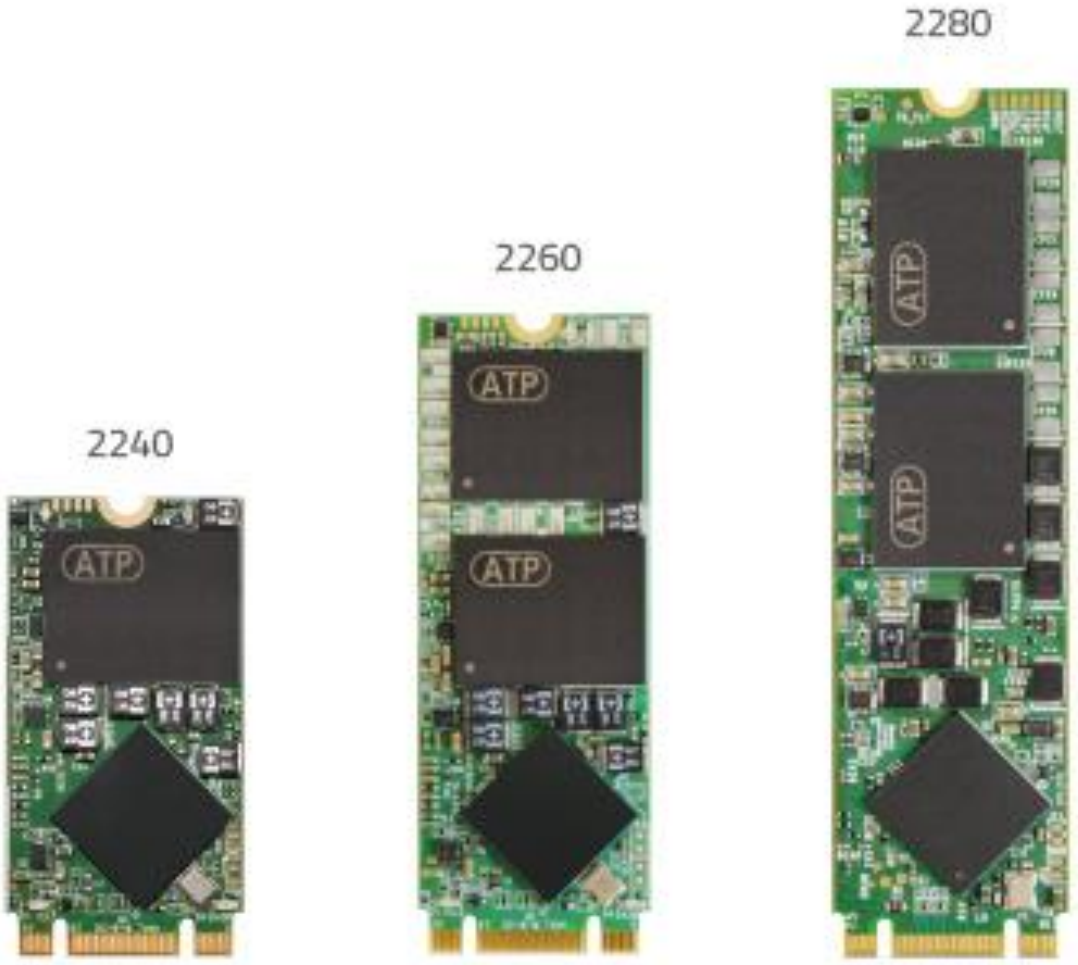
Sometimes as M.2. SSD using PCIe 3.0 or above, using 4 lanes



Hard disk interfaces



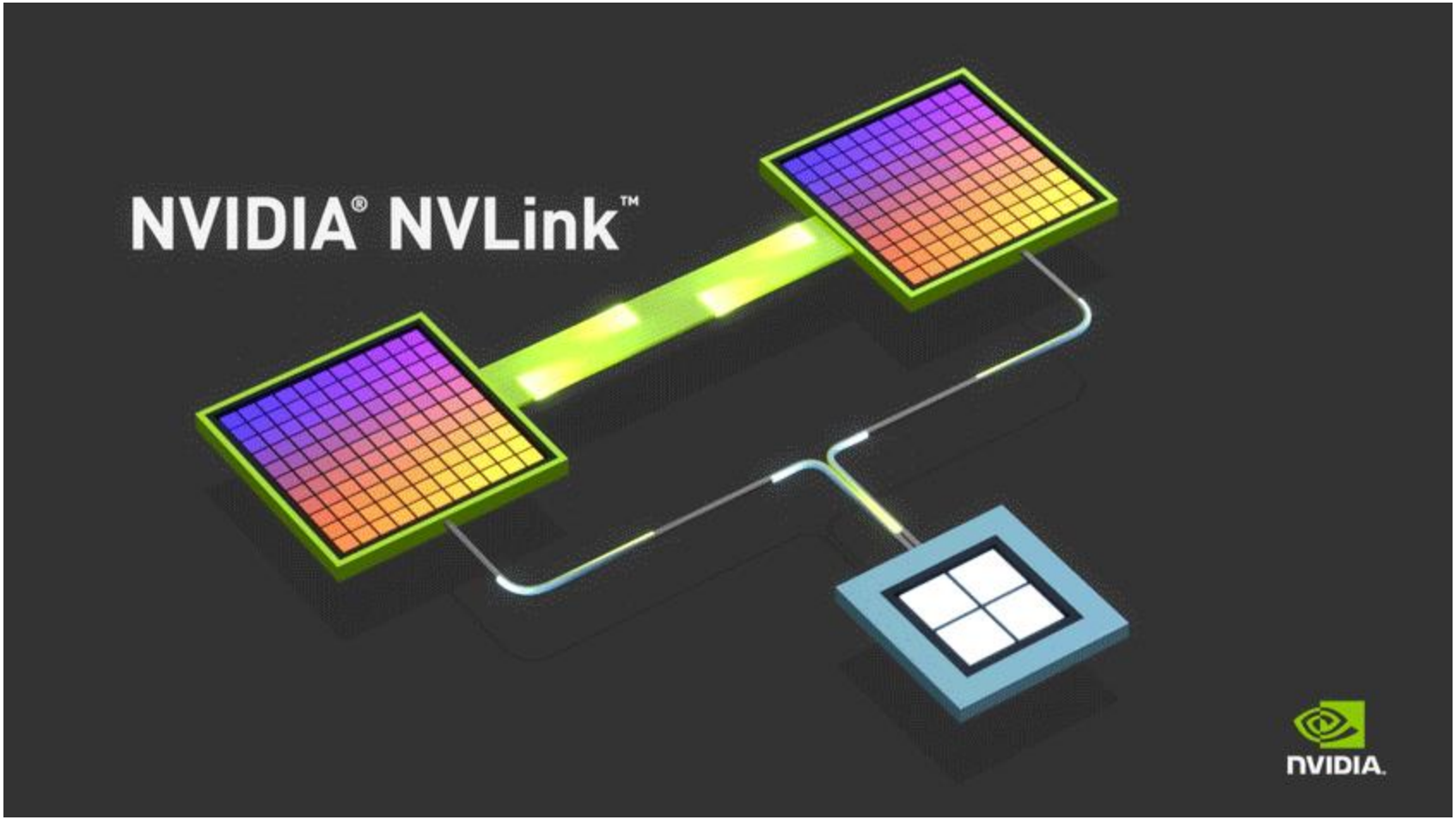
Form factors – M.2 (size)



Form factors – M.2 (keying)

Key ID	Pin Location	Interface
A	8-15	2x PCIe x1 / USB 2.0 / I2C / DP x4
B	12-19	PCIe x2 / SATA /USB 2.0 / USB 3.0 / HSIC / SSIC / Audio / UIM / I2C
C	16-23	Reserved for Future Use
D	20-27	Reserved for Future Use
E	24-31	2x PCIe x1 / USB 2.0 / I2C / SDIO / UART / PCM
F	28-35	Future Memory Interface (FMI)
G	39-46	Not Used for M.2; for Custom/Non-Standard Apps
H	43-50	Reserved for Future Use
J	47-54	Reserved for Future Use
K	51-58	Reserved for Future Use
L	55-62	Reserved for Future Use
M	59-66	PCIe x4 / SATA

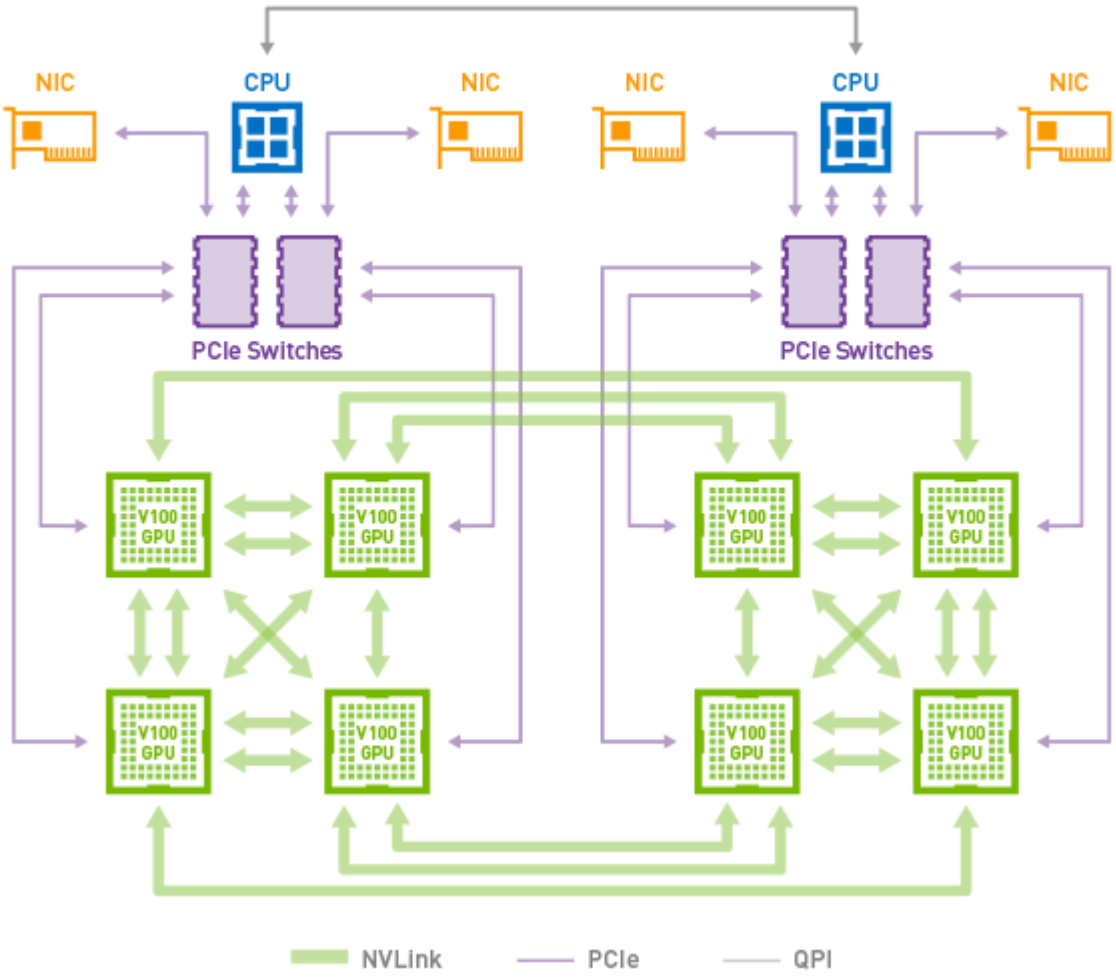
NVLink



NVLink

- NVLink 1.0, 2.0, 3.0 and 4.0
- Consists of links made up of differential signals (lanes)
  - 4 or 8 lanes for a link per direction. 50 GB/s per link bidirectional rate
- NVLink 2.0, 3.0 and 4.0 has 6, 12 and 18 links correspondingly
- Ampere A100 has 12 links, and a rate of 600 GB/s
- Hopper has 18 links, and a rate of 900 Gb/s
- The NVLink can form a mesh network

NVLink


NIC
CPU
NIC
NIC
CPU
NIC
PCIe Switches
PCIe Switches
V100 GPU
V100 GPU
V100 GPU
V100 GPU
V100 GPU
V100 GPU
V100 GPU
V100 GPU
NVLink
PCIe
QPI

Introduction

Questions?

Contact information

Andreas Axelsson

Email: andreas.axelsson@ju.se

Mobile: 0709-467760