

PCI Express - Peter Fabian Arnold(s0574033)

Der Verbindungsstandard PCIe(Peripheral Component Interconnect Express) ist eine sehr schnelle interne Schnittstelle für Erweiterungskarten, wie GPUs und SSDs, in Computer Systemen. Die PCIe Spezifikation beschreibt das genutzte Software Protokoll, sowie die elektrischen und mechanischen Eigenschaften der PCIe Steckverbinder Anschlüsse und Erweiterungskarten.

PCIe wurde 2003 als Verbesserung von PCI und AGP eingeführt und seit 2010 sind die alten Ports fast vollständig abgelöst. Auf der logischen Ebene ist der PCIe Standard vollkommen mit PCI kompatibel.

Bei dem veralteten PCI Standard wurde eine parallele Busstruktur genutzt, bei der sich alle Komponenten die verfügbare Bandbreite des Busses teilen mussten. Anders als PCI nutzt PCIe eine moderne Switch-Architektur, bei der die Geräte über serielle Direktverbindungen mit einem Switch verbunden sind. Der Switch verbindet die PCIe Baugruppe direkt mit dem CPU, RAM und anderen Baugruppen, so können alle Geräte die volle Bandbreite der Verbindungen ausnutzen, ohne die Datenrate anderer Geräte zu beeinflussen.

Wie auch beim OSI Standard funktioniert die Kommunikation bei PCIe mit verschiedenen Schichten: Physical Layer, Data Link Layer, Transaction Layer.

Physical Layer:

PCIe nutzt Lanes zum Versenden der Daten. Jede Lane besteht aus 2 einseitig gerichteten Kabelpaaren(4 Kabel pro Lane). Jeweils ein Kabelpaar wird zum Senden/ Empfangen genutzt. PCIe ist also voll Duplex-fähig.

Ein PCIe Slot kann 1 bis 32 Lanes lang sein, dabei kann ein Gerät mit wenigen Lanes auch in größeren Slots genutzt werden, die maximal nutzbaren Lanes werden vom Host ermittelt. So kann ein PCIe x1 Gerät(mit einer Lane) in jedem PCIe Slot genutzt werden.

Um die Anzahl von Lanes auszunutzen, verwendet PCIe Data Striping, dabei wird das nachfolgende Byte jeweils auf der nächsten Lane übertragen, dies ermöglicht eine sehr hohe Datenrate und eine niedrige Latenz.

Datenübertragung:

PCIe 1.0 nutzte zum Kodieren der Daten den 8bit/10bit Code, aus dem sich die Sendefrequenz beim Empfangen rekonstruieren lässt. Da die Effizienz von 8bit/10bit Code nur 80% ist, wird ab PCIe 3.0 128bit/130bit Code mit 98,4% Effizienz genutzt.

Bis PCIe 5.0 wurden Non Return to Zero(NRZ) Signale genutzt, ab PCIe 6.0 PAM-4. Die Abkürzung PAM-4 steht für Puls Amplituden Modulation mit 4 Zuständen, also können mit PAM-4 bei gleicher Frequenz und Bandbreite doppelt so viele Daten übertragen werden(NRZ = 2 Zustände). Bei PCIe 6.0 x16 können bis zu 121 GB/s übertragen werden.

Data Link Layer:

Das Data Link Layer legt die Reihenfolge der zu sendenden TLPs(Transaction Layer Packages) fest und synchronisiert diese beim Empfangen.

Zudem muss das Data Link Layer die korrekte Übertragung der Daten mittels Acknowledgement(ACK) Signalen sicherstellen, dazu gibt es zwei Systeme: Jedem TLP wird ein Tag(Ident Tag) zugewiesen und am Ende jedes TLP wird eine 32 Bit Prüfsumme(Link Cyclic Redundancy Code) angehängt. Wenn das Data Link Layer eine fehlerhafte oder fehlende Übertragung feststellt, sendet es ein NACK(negative acknowledgment) und die Pakete ab dem Fehler müssen neu übertragen werden.

Außerdem kann das Data Link Layer DLLPs(Data Link Layer Packages) versenden. Diese Pakete enthalten die ACK/NACK Signale, Informationen über das Power Management und Flow Control Credits, diese werden für das Transaction Layer übertragen und enthalten Informationen über die Kapazitäten der an der Übertragung beteiligten Geräten.

Transaction Layer:

Das Transaction Layer empfängt und sendet Transaction Layer Packages(TLPs). Die TLPs werden auf einem 32 Bit Bus eingelesen und bestehen aus 32 Bit langen Double Words(DWs). Der Header eines TLP kann 3-4 DWs lang sein und das Payload 0-1023 DWs. Es gibt verschiedene Arten von TLPs wie z.B. write, read request oder completion TLPs.