



**TECHNISCHE
UNIVERSITÄT
DRESDEN**

Institut für Technische Informatik, Professur für VLSI-Entwurfssysteme, Diagnostik und Architektur

PCI Express

Dirk Wischeropp

Dresden, 07.06.2011



**DRESDEN
concept**
Exzellenz aus
Wissenschaft
und Kultur

1 Einleitung

2 Architektur

3 Layering

4 Zusammenfassung



1 Einleitung

2 Architektur

3 Layering

4 Zusammenfassung

- **PCIe (peripheral component interconnect express)**
- **Nachfolger von PCI und AGP**
 - **Kernunterschied: seriell ↔ parallel**
- **Spezifiziert durch PCI-SIG**
 - **Über 900 Unternehmen (AMD, Intel, Nvidia, ...)**
 - **1992 gegründet**
 - **Spezifikation: <http://www.pcisig.com>**

- **Kommunikation zwischen Komponenten unterschiedlicher Hersteller**
- **Weiter Bereich (Desktop, Server, Embedded, ...)**
- **PCI SW Kompatibilität → keine Neuen Treiber**
- **Performance: Bandbreite, Latency, Overhead, Skalierbarkeit**
- **Wenig Kosten verursachen**

- **2003: PCIe 1.0a**
- **2005: PCIe 1.1 → Orientierung an dieser Spez.**
- **2007: PCIe 2.0**
- **2010: PCIe 3.0**

- **Wesentlicher Unterschied ist die Übertragungsbandbreite**
- **Konstante Weiterentwicklung**



1 Einleitung

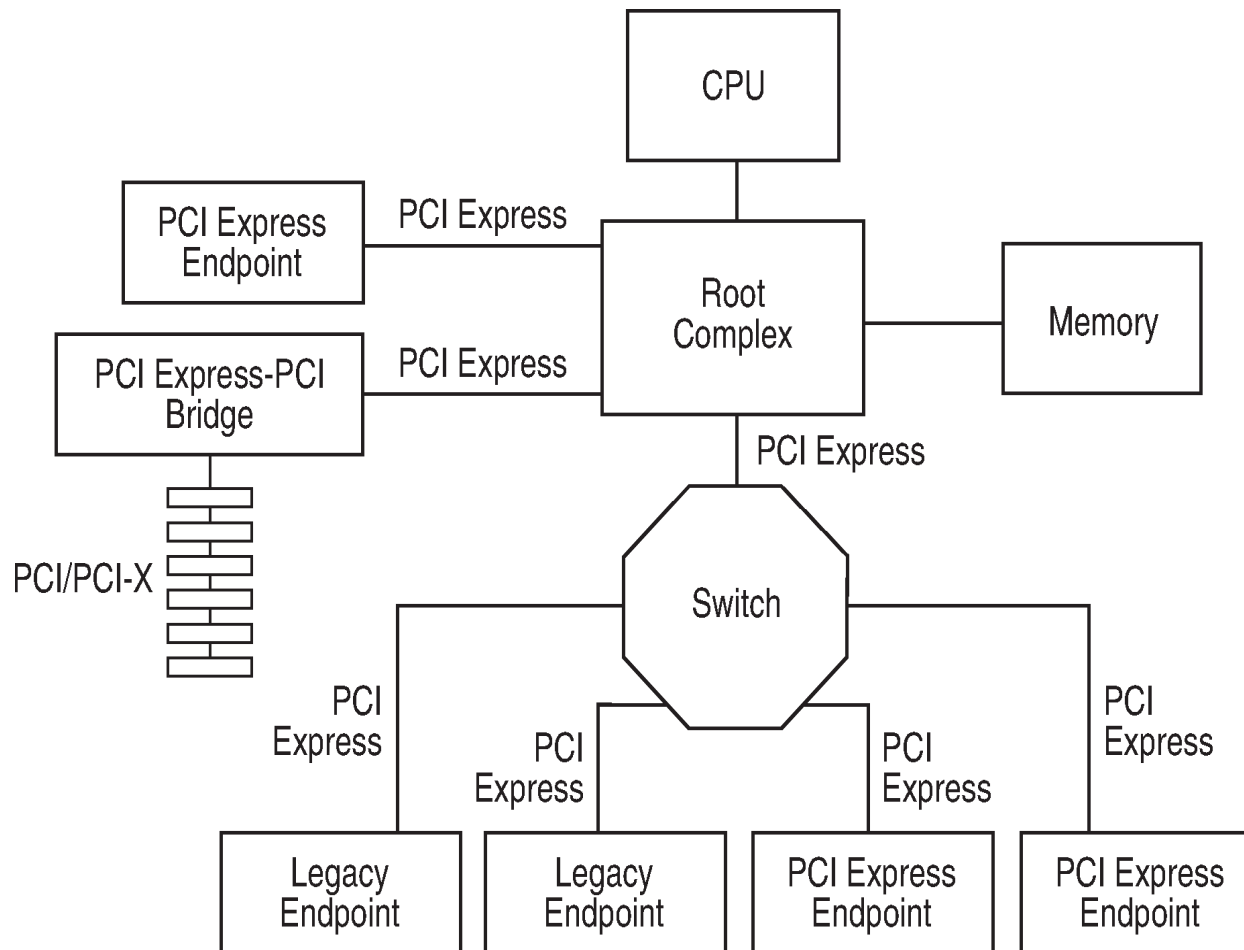
2 Architektur

3 Layering

4 Zusammenfassung

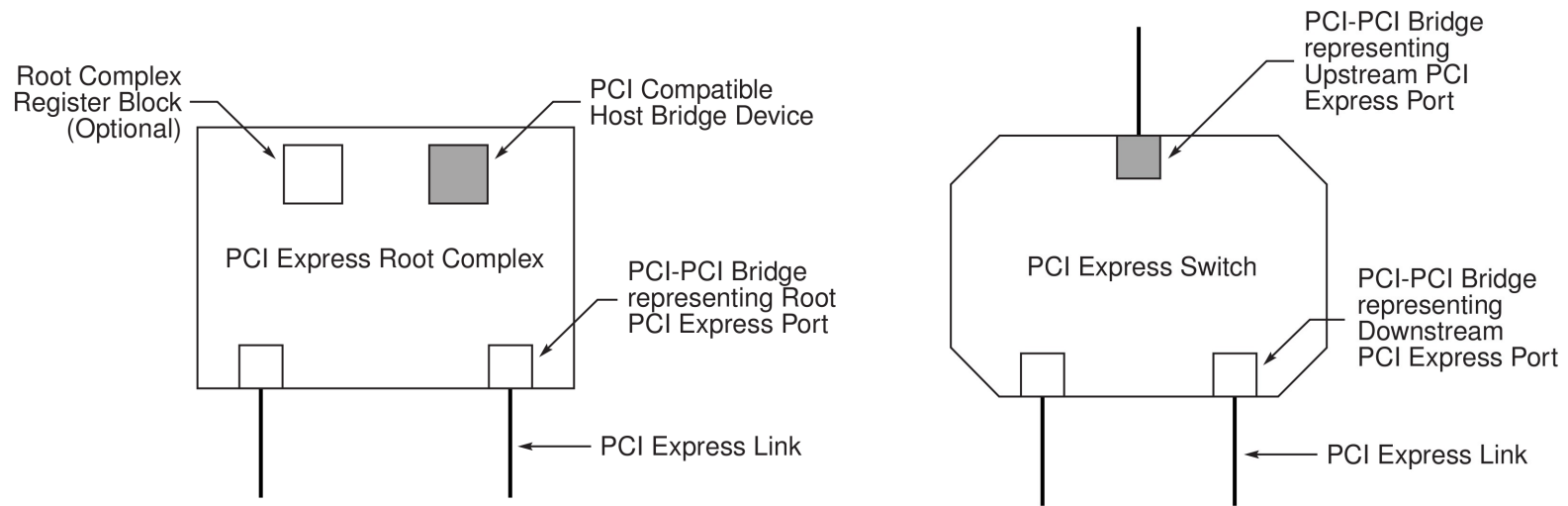
- **Serielle Point-to-Point Verbindung**
 - **Kein BUS-Sharing**
 - **Kein Clock-Skewing (Leitungsparameter)**
 - **höhere Taktrate**
- **Full duplex**
 - **gleichzeitig Lesen und Schreiben**
- **Taktrate unabhängig von anderer Peripherie**
 - **Komponente kann System nicht ausbremsen**
- **Softwarekompatibilität zu PCI 3.0**

Topologie



Beispiel Topologie. Quelle: Abb. 1-2 aus [1]

Root Complex und Switch



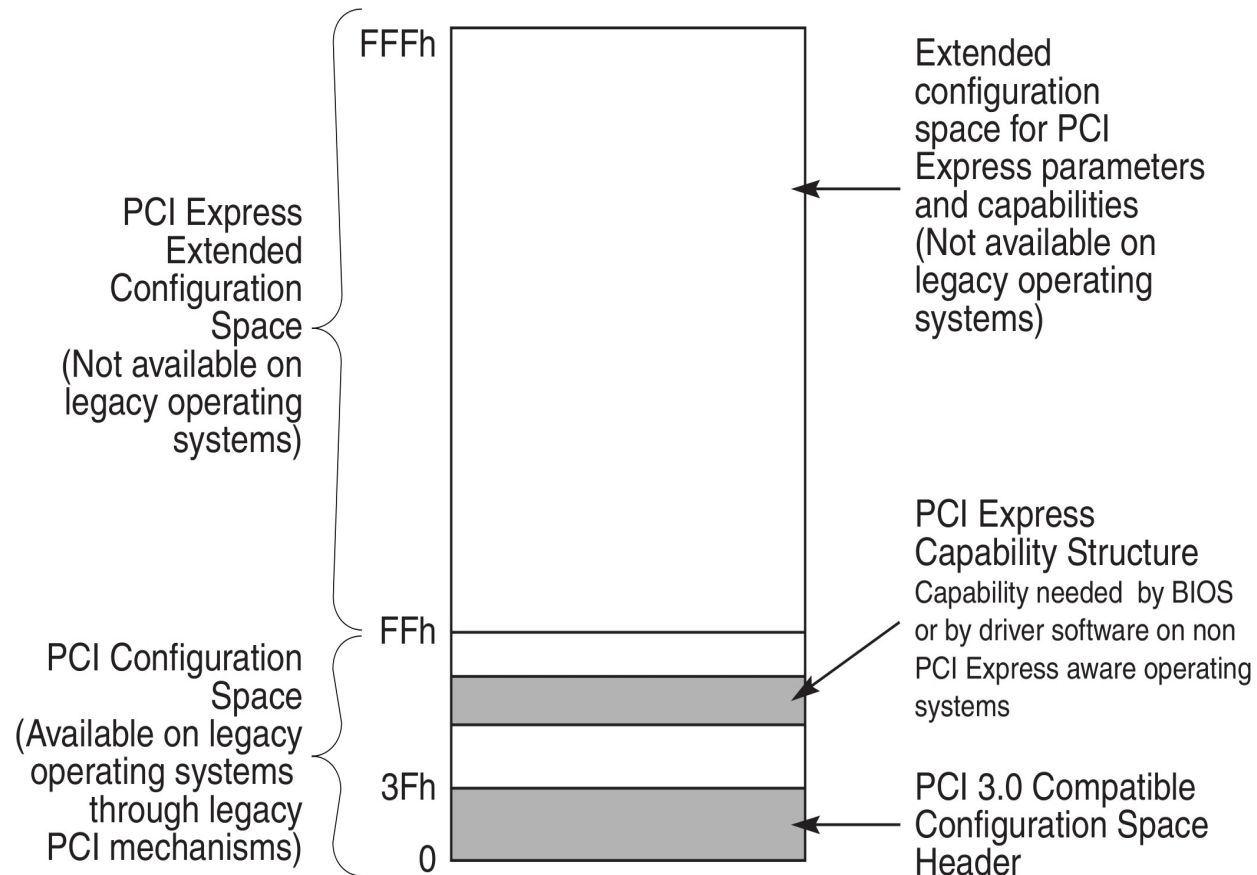
Quelle: Abb. 7-1 und 7-2 aus [1]

- **PCIe Link wird als PCI Bus dargestellt**
- **Nur ein Device pro Bus**
- **Aber: bis zu 8 Funktionen pro Device**

```
$ lspci -vt
```

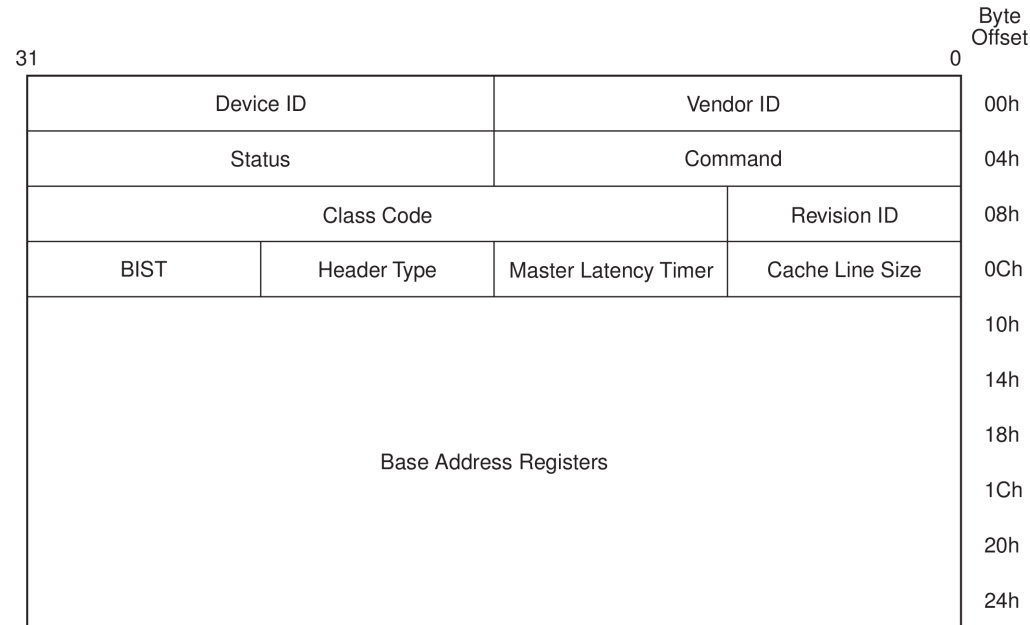
```
-[0000:00]--+-00.0 Advanced Micro Devices [AMD] RS780 Host Bridge
  +-02.0-[01]---+-00.0 ATI Technologies Inc Barts PRO [ATI Radeon HD 6800 Series]
  |
  | \-00.1 ATI Technologies Inc Barts HDMI Audio [Radeon HD 6800 Series]
  +-0a.0-[02]----00.0 Ltd. RTL8111/8168B PCI Express Gigabit Ethernet controller
  +-11.0 ATI Technologies Inc SB7x0/SB8x0/SB9x0 SATA Controller [IDE mode]
  +-12.0 ATI Technologies Inc SB7x0/SB8x0/SB9x0 USB OHCI0 Controller
  +-12.1 ATI Technologies Inc SB7x0 USB OHCI1 Controller
  +-12.2 ATI Technologies Inc SB7x0/SB8x0/SB9x0 USB EHCI Controller
  +-13.0 ATI Technologies Inc SB7x0/SB8x0/SB9x0 USB OHCI0 Controller
  +-13.1 ATI Technologies Inc SB7x0 USB OHCI1 Controller
  +-13.2 ATI Technologies Inc SB7x0/SB8x0/SB9x0 USB EHCI Controller
  +-14.0 ATI Technologies Inc SBx00 SMBus Controller
  +-14.1 ATI Technologies Inc SB7x0/SB8x0/SB9x0 IDE Controller
  +-14.2 ATI Technologies Inc SBx00 Azalia (Intel HDA)
  +-14.3 ATI Technologies Inc SB7x0/SB8x0/SB9x0 LPC host controller
  +-14.4-[03]---+-06.0 Brooktree Corporation Bt878 Video Capture
  |
  | +-06.1 Brooktree Corporation Bt878 Audio Capture
  |
  | \-0e.0 Texas Instruments TSB43AB23 IEEE-1394a-2000 Controller (PHY/Link)
  +-14.5 ATI Technologies Inc SB7x0/SB8x0/SB9x0 USB OHCI2 Controller
  +-18.0 Advanced Micro Devices [AMD] Family 10h Processor HyperTransport Configuration
  +-18.1 Advanced Micro Devices [AMD] Family 10h Processor Address Map
  +-18.2 Advanced Micro Devices [AMD] Family 10h Processor DRAM Controller
  +-18.3 Advanced Micro Devices [AMD] Family 10h Processor Miscellaneous Control
  \-18.4 Advanced Micro Devices [AMD] Family 10h Processor Link Control
```

Konfigurationsraum



Quelle: Abb. 7-3 aus [1]

Konfigurationsraum PCI Header „00“



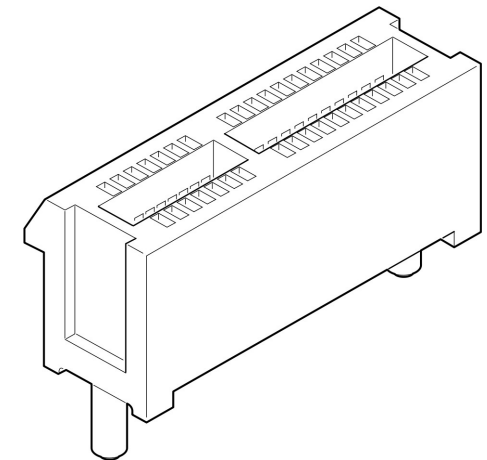
Quelle: Abb. 7-5 aus [1]

- **Identifikation des Device → passenden Treiber laden**
- **System setzt BARs → Adressierung im Speicher- bzw. I/O-Bereich**

- **PCIe Capability Structure:**
 - **Zeigt an, dass es sich um ein PCIe Device handelt**
 - **PCIe spezifische Kontroll- und Status-Register (Power Management)**
 - **Angepasste SW auf altem OS kann Funktionen nutzen**
- **Extended Capabilities:**
 - **Error Reporting, Device Seriennummer, Hersteller spezifische Fähigkeiten**
 - **OS Treiber müssen erweitert werden**

- **Kommunikation findet über Links statt**
- **Logische Verbindung zwischen 2 Komponenten**
- **Auf physikalischer Ebene bestehen Links aus Lanes (x1, x2, x4, x8, x16, x32)**
- **Jede Lane überträgt Daten full-duplex mit 2,5Gbit/s**
- **Datenraten bis zu 80Gbit/s möglich**
- **Verbindung von Peripherie über Steckkarten**

Pin #	Side B		Side A	
	Name	Description	Name	Description
1	+12V	12 V power	PRSNT1#	Hot plug presence detect
2	+12V	12 V power	+12V	12 V power
3	RSVD	Reserved	+12V	12 V power
4	GND	Ground	GND	Ground
5	SMCLK	SMBus (System Management Bus) clock	JTAG2	TCK (Test Clock), clock input for JTAG interface
6	SMDAT	SMBus (System Management Bus) data	JTAG3	TDI (Test Data Input)
7	GND	Ground	JTAG4	TDO (Test Data Output)
8	+3.3V	3.3 V power	JTAG5	TMS (Test Mode Select)
9	JTAG1	TRST# (Test Reset) resets the JTAG interface	+3.3V	3.3 V power
10	3.3Vaux	3.3 V auxiliary power	+3.3V	3.3 V power
11	WAKE#	Signal for link reactivation	PWRGD	Power good



Quelle: Abb. 1-1 aus [2]

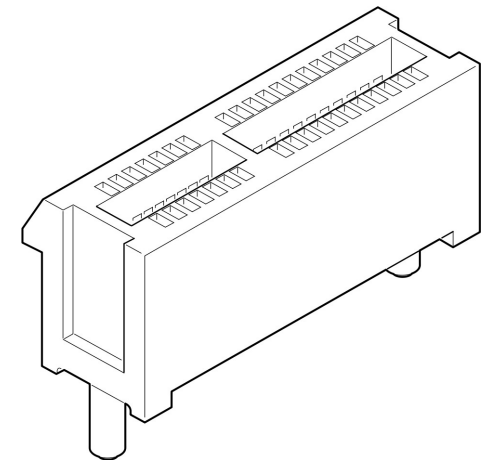
Quelle: Tab. 5-1 aus [2]

Steckkartenverbindungen (2)

Pin #	Side B		Side A	
	Name	Description	Name	Description
Mechanical Key				
12	RSVD	Reserved	GND	Ground
13	GND	Ground	REFCLK+	Reference clock (differential pair)
14	HSOp(0)	Transmitter differential pair, Lane 0	REFCLK-	
15	HSOn(0)		GND	Ground
16	GND	Ground	HSIp(0)	Receiver differential pair, Lane 0
17	PRSENT2#	Hot plug presence detect	HSIn(0)	
18	GND	Ground	GND	Ground

End of the x1 Connector

Quelle: Tab. 5-1 aus [2]



Quelle: Abb. 1-1 aus [2]



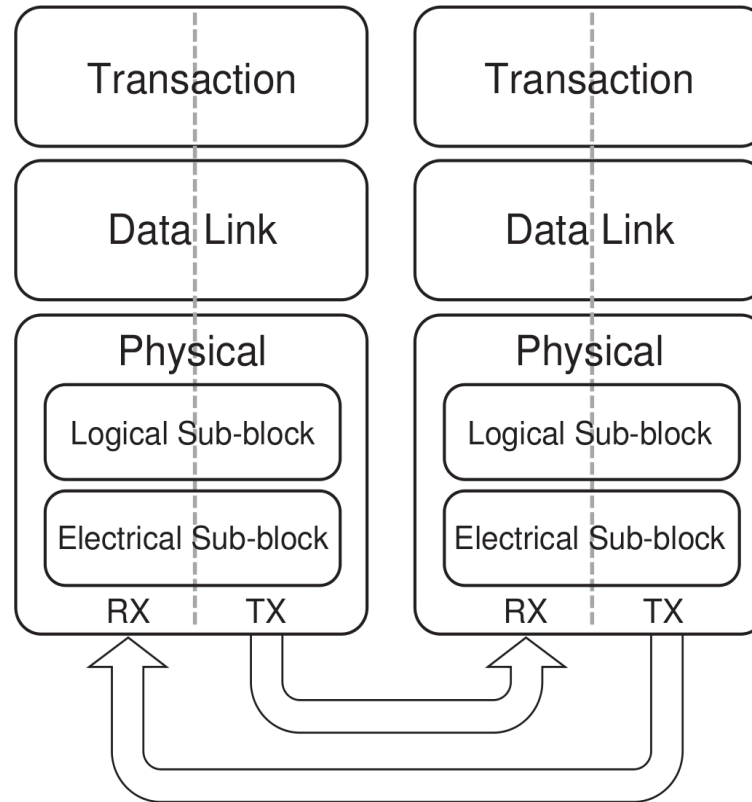
1 Einleitung

2 Architektur

3 Layering

4 Zusammenfassung

Layers



Quelle: Abb. 1-4 aus [1]

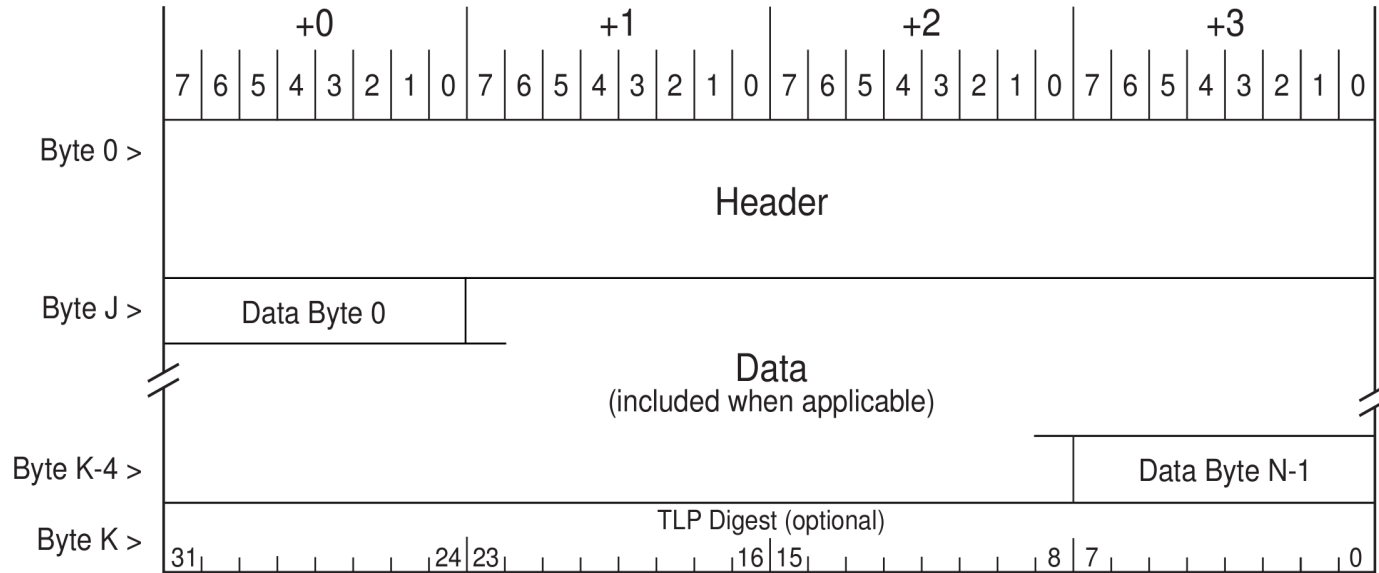
- **Datenaustausch auf Lanes**
- **2 Sub-Layer**
- **Electrical Sub-Layer:**
 - **Bit-Stream ↔ elektrisches Signal**
 - **Differentielle Signalpaar pro Lane und Richtung**
 - **Taktrate für PCIe 1.x : 2,5Gbit/s**

- **Logical Sub-Layer:**
 - **Data Link Stream ↔ Bit-Stream**
 - **8b10b encoding und decoding:**
 - **Gleichspannungsausgleich (max. 5 mal '1')**
 - **Taktrückgewinnung**
 - **Aber: 20% Overhead (geändert in PCIe 3.0)**
 - **Framing (STP, SDP, END)**
 - **Übertragung Symbolweise auf Lanes aufgeteilt**

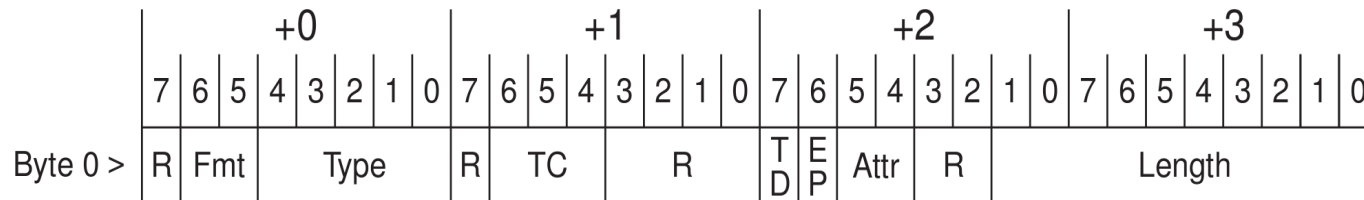
- **Sichert die Integrität zwischen 2 Link Partnern**
- **32bit Prüfsumme am Ende (LCRC)**
- **TLP Sequenz-Nummer um Reihenfolge zu erhalten**
- **Spezielle DLLPs:**
 - **TLP kam korrekt an → ACK DLLP**
 - **TLP kam nicht korrekt an → NACK DLLP**
 - **Powermanagement**

- **Ende-zu-Ende Kommunikation**
- **4 Arten von TLP:**
 - **Memory Read/Write**
 - **I/O READ/Write**
 - **Configuration Read/Write**
 - **Message (Signalisieren von Ereignissen)**
- **Datenaustausch erfolgt immer als Requester und Completer**
- **Routing über Memory bzw. I/O Adresse oder über ID (Bus-, Device-, Function-Nr.)**

TLP



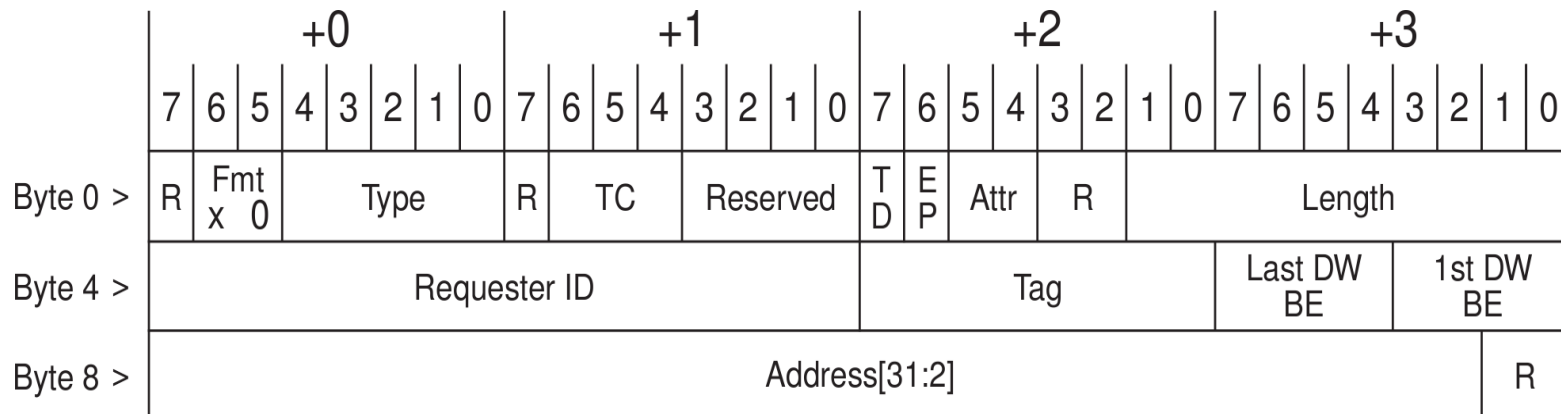
Quelle: Abb. 2-3 aus [1]



Quelle: Abb. 2-4 aus [1]

TLP Beispiel

- **32bit Memory Read Request:**
 - Fmt[0] = '0' → 3 DW Header
 - Fmt[1] = '0' → keine Daten (Read)
 - Type = „00h“ → Memory Read/Write
 - Length = „004h“ → 4 DW Payload bei Completion
 - Requester ID = Bus-, Dev-, Func-Nummer
 - Tag = einzigartig für TLP → benötigt Completion TLP



Quelle: Abb. 2-14 aus [1]

1 Einleitung

2 Architektur

3 Layering

4 Zusammenfassung

- **Serieller Bus-Standard**
- **Software-seitig kompatibel zu PCI**
- **Kommunikation über Links**
- **Paket-basierte Datenübertragung**
- **3 Übertragungsschichten**
- **Übertragung durch Requester und Completer**

**[1] PCI-SIG, „PCI Express® Base Specification
Revision 1.1“**

**[2] PCI-SIG, „PCI Express™ Card
Electromechanical Specification Revision 1.0“**

[3] <http://en.wikipedia.org> und externe Verweise