



Verteidigung Studienarbeit

Entwurf und Implementierung eines statischen Backbones für die Kommunikation mit dynamischen Nutzerpartitionen auf einem Multi-FPGA-Board

Albert Schulz

Dresden, 20.10.2016

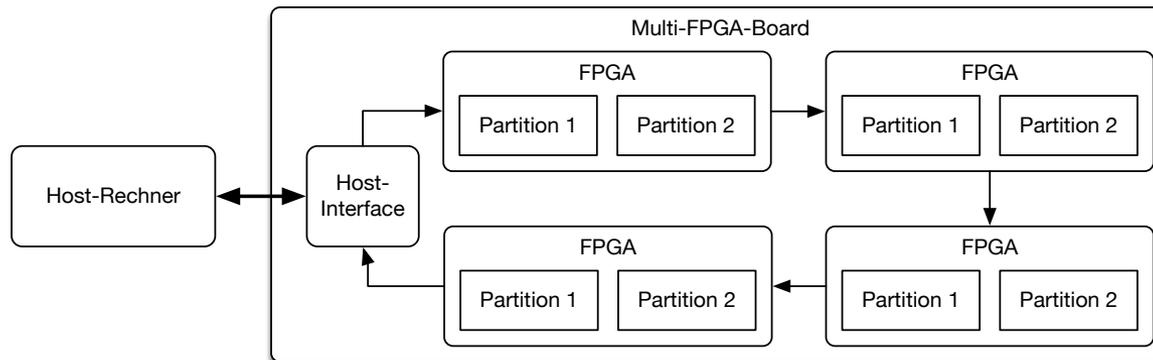


Gliederung

1. Aufgabenstellung
2. Zielarchitektur
3. Grundlage des Entwurfs
4. Entwurf und Implementierung
5. Auswertung
6. Zusammenfassung

1. Aufgabenstellung

- große FPGAs → Optimale Auslastung durch mehrere Nutzerdesigns pro FPGA
- mehrere große FPGAs auf Multi-FPGA-Boards
- Kommunikation zwischen Host und Nutzerdesigns nötig



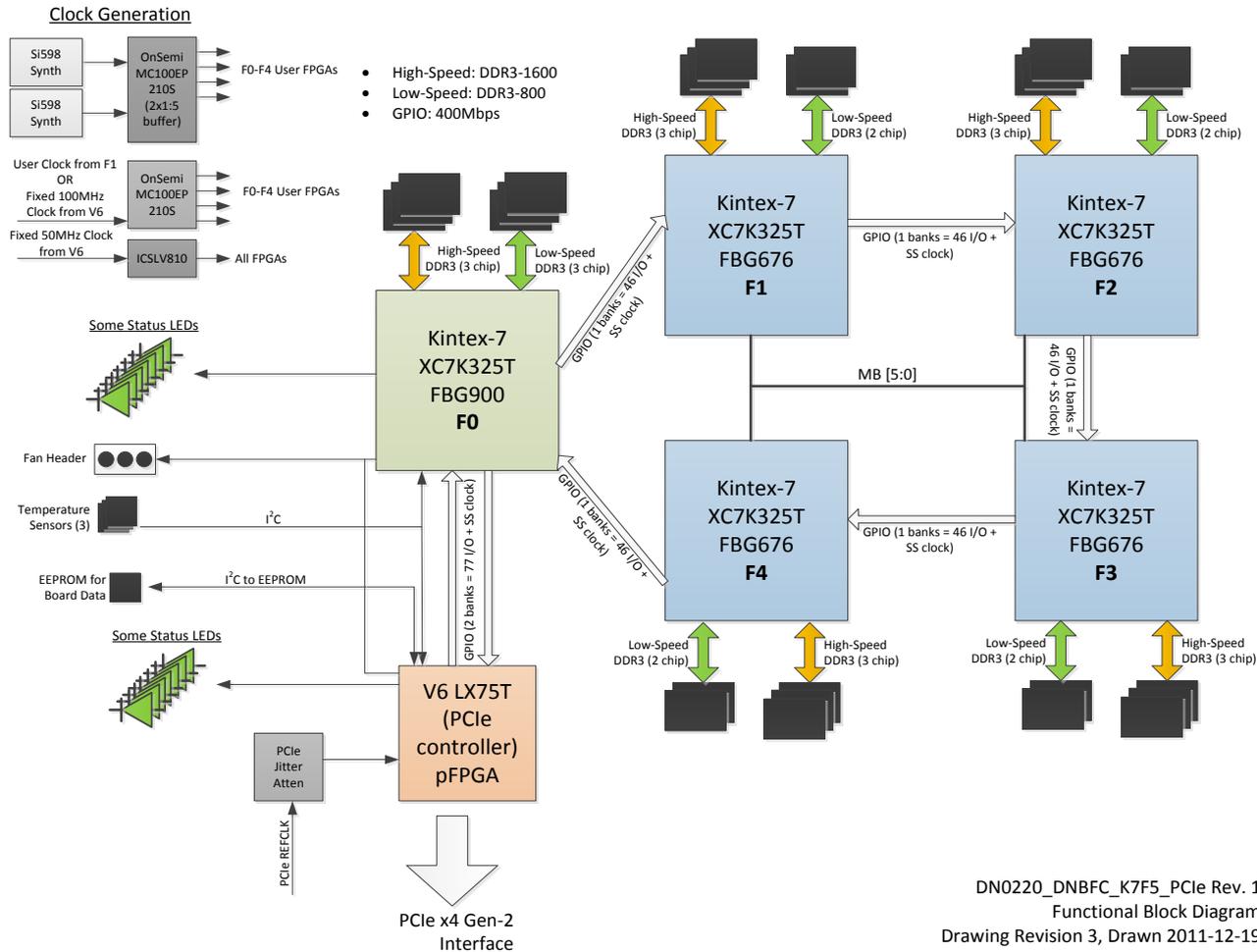
1. Aufgabenstellung

- Wesentliche Punkte der Aufgabenstellung:
 - Datenübertragung zwischen Host und Partitionen über gemeinsamen Bus
 - Entwurf eines paketbasierten Protokolls
 - Möglichkeit zur latenzarme Übermittlung von Interrupt- und Statusinformationen
 - generisch mehrere Nutzerpartitionen adressierbar

2. Zielarchitektur

- DNK7 F5 PCIe-Board mit 5 Kintex-7 FPGAs
 - 4x gFPGA: „User FPGAs“
 - 1x dFPGA: „Dataflow Manager“
- Ring-Verbindung der FPGAs über 46 GPIO-Pins
- PCIe-Interface zu Host-Rechner (mittels zusätzlichen pFPGA)
- FPGAs durch Chip-ID (3 Bit) identifizierbar

2. Zielarchitektur

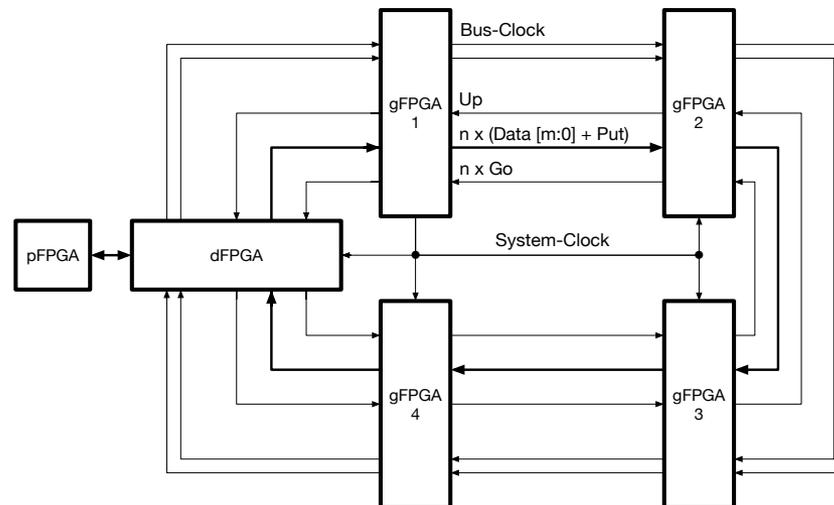


DN0220_DNBFC_K7F5_PClE Rev. 1
Functional Block Diagram
Drawing Revision 3, Drawn 2011-12-19

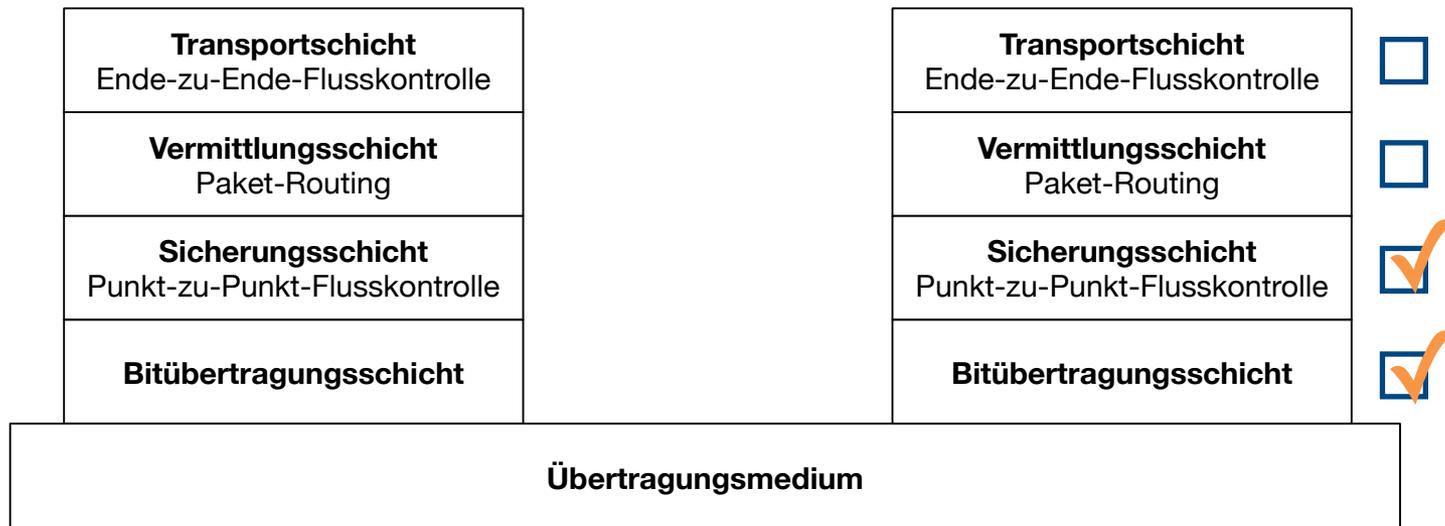
[1]

3. Grundlage des Entwurfs

- systolische Datenübertragung über Ringbus
 - Flusskontrolle zwischen benachbarten Chips über Go-Signal
 - Keine Sicherung der Datenintegrität (Annahme eines störungsfreien Kanals)
- IP-Core für Host-Kommunikation per DMA

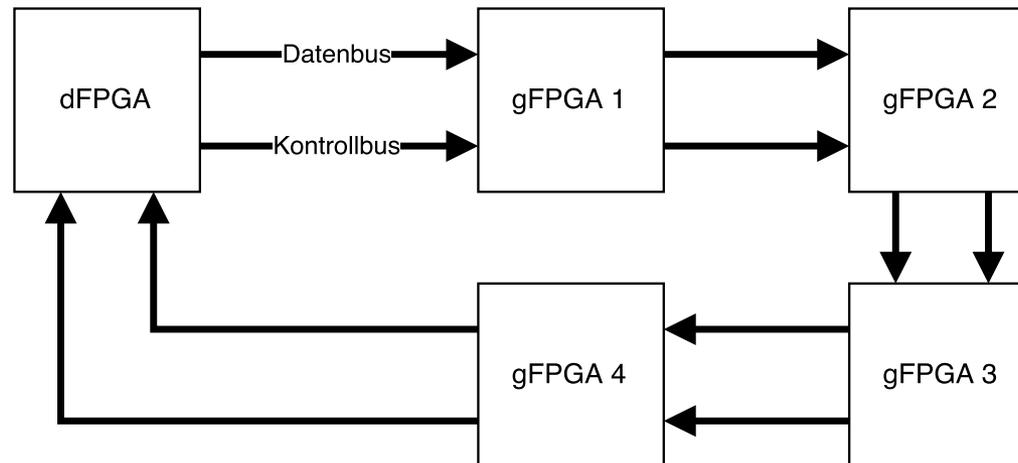


3. Grundlage des Entwurfs



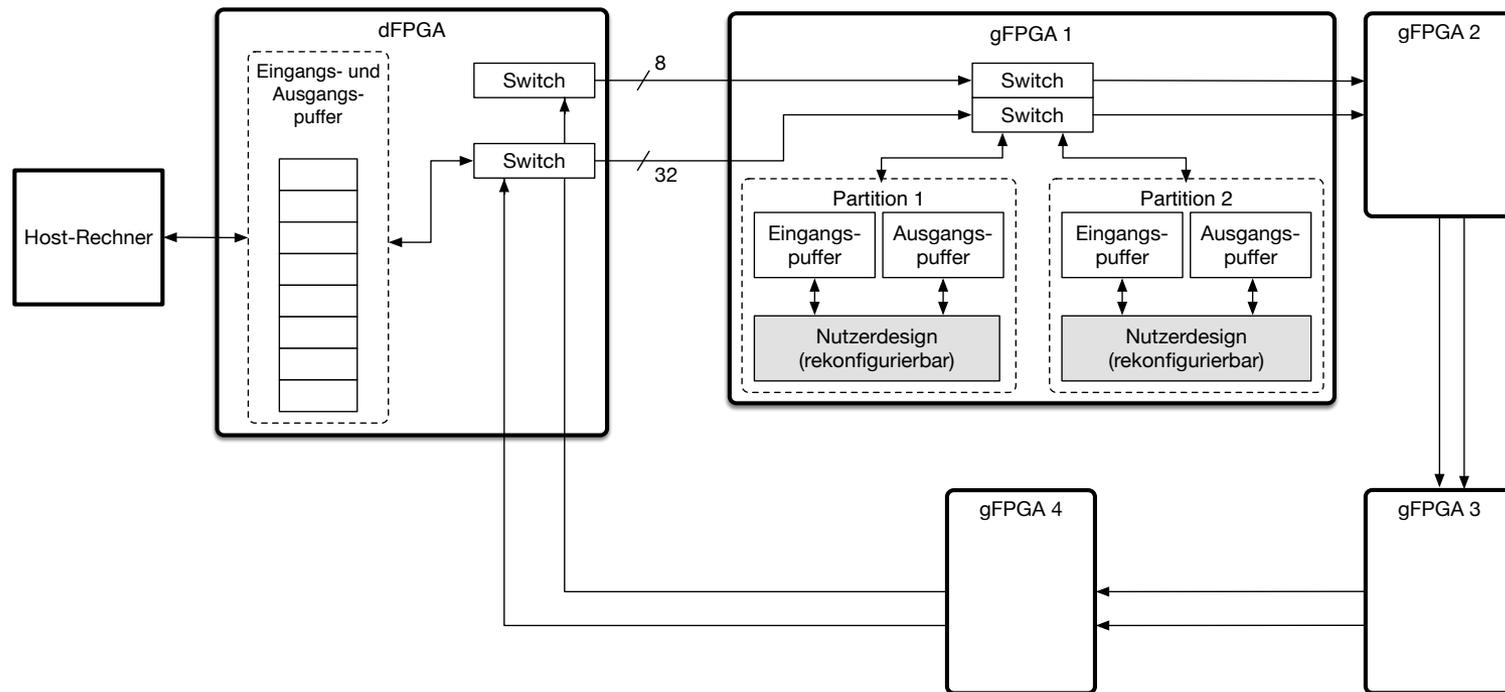
4. Entwurf und Implementierung - Struktur

- 2 getrennte Busse:
 - 32-Bit Datenbus (hohe Datenrate)
 - 8-Bit Kontrollbus (geringe Verzögerung)



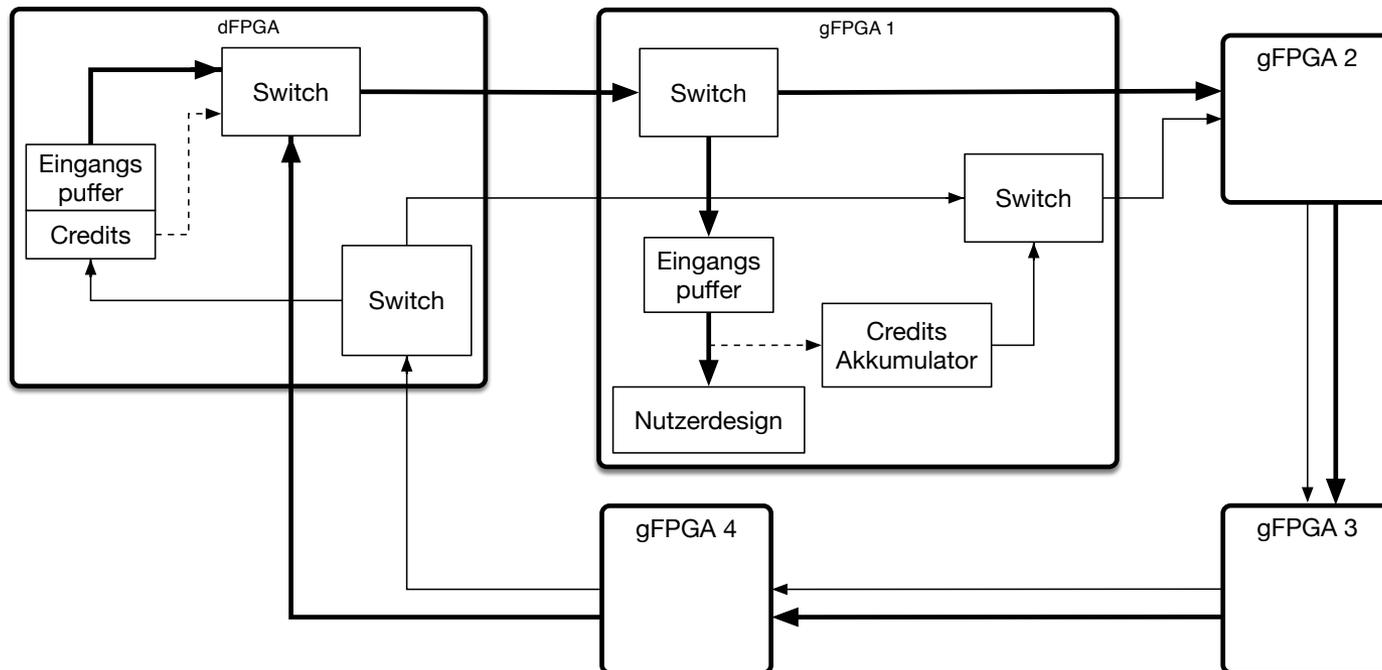
4. Entwurf und Implementierung - Struktur

- Systemstruktur auf Netzwerkschicht



4. Entwurf und Implementierung - Flusskontrolle

- „Credit-based-Flow-Control“ zwischen dFPGA und Nutzerdesigns



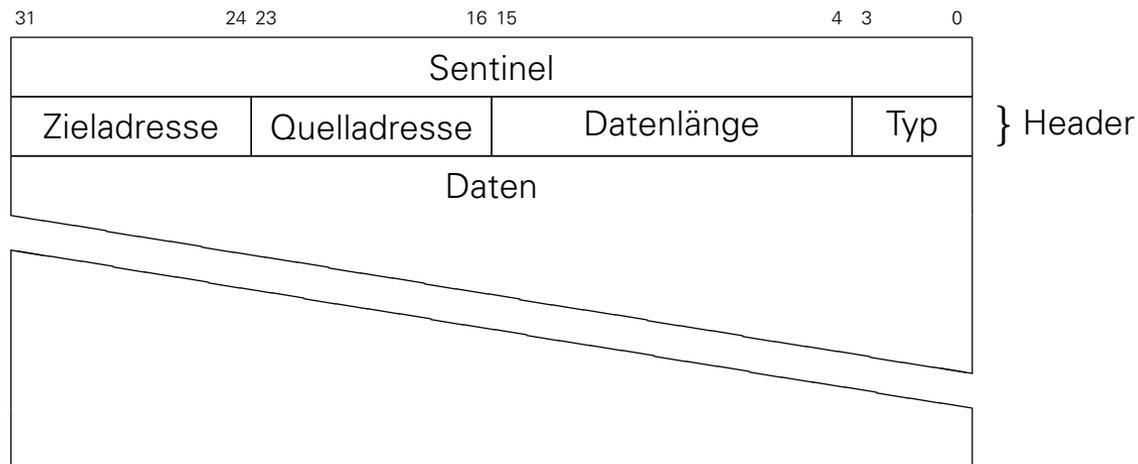
4. Entwurf und Implementierung - Flusskontrolle

- Trigger für Senden von Credit-Paket:
 - nach Lesen von $N/2$ Worten aus Empfangspuffer (N =Puffergröße)
 - max. Creditanzahl von 255 erreicht
 - geringere, gleichmäßige Auslastung des Busses

4. Entwurf und Implementierung - Adressierung

- Identifikation der Nutzerdesigns über 8-bit Adresse:
 - 3-Bit Chip-ID
 - 5-Bit ID für Partition auf dem Chip
- Adressierung
 - Downstream: nur Zieladresse
 - Upstream: Ziel- & Quelladresse

4. Entwurf und Implementierung - Paketformat

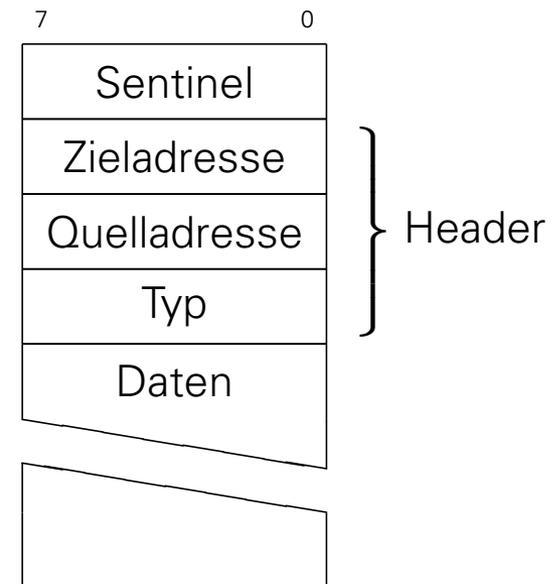


- Paketformat für Datenbus:
 - Sentinel zur Synchronisation
 - Typen: Nutzdaten (RAM, Konfiguration)
 - Datenlänge zwischen 0 und 4.095 Worte

4. Entwurf und Implementierung - Paketformat

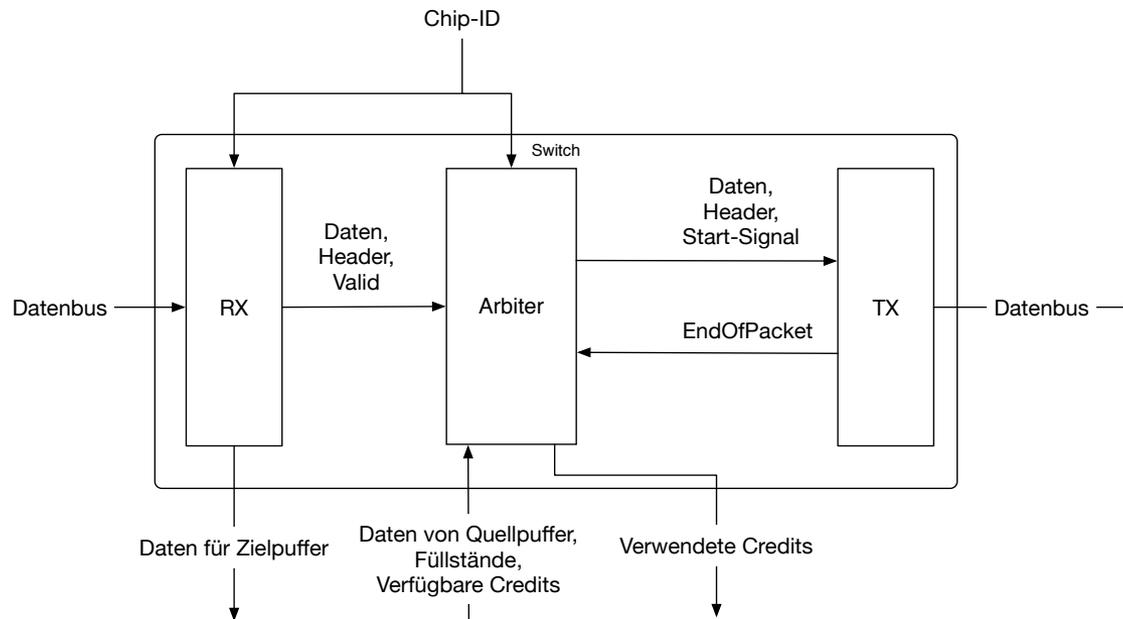
- **Paketformat für Kontrollbus:**
 - Typen: Credits, Status, Reset
 - Daten abhängig vom Typ

Pakettyp	Datenformat	Funktion
0x00	unsigned byte	Credits
0x01	-	Reset
0x02	4-Byte-Wort (Big Endian)	Status



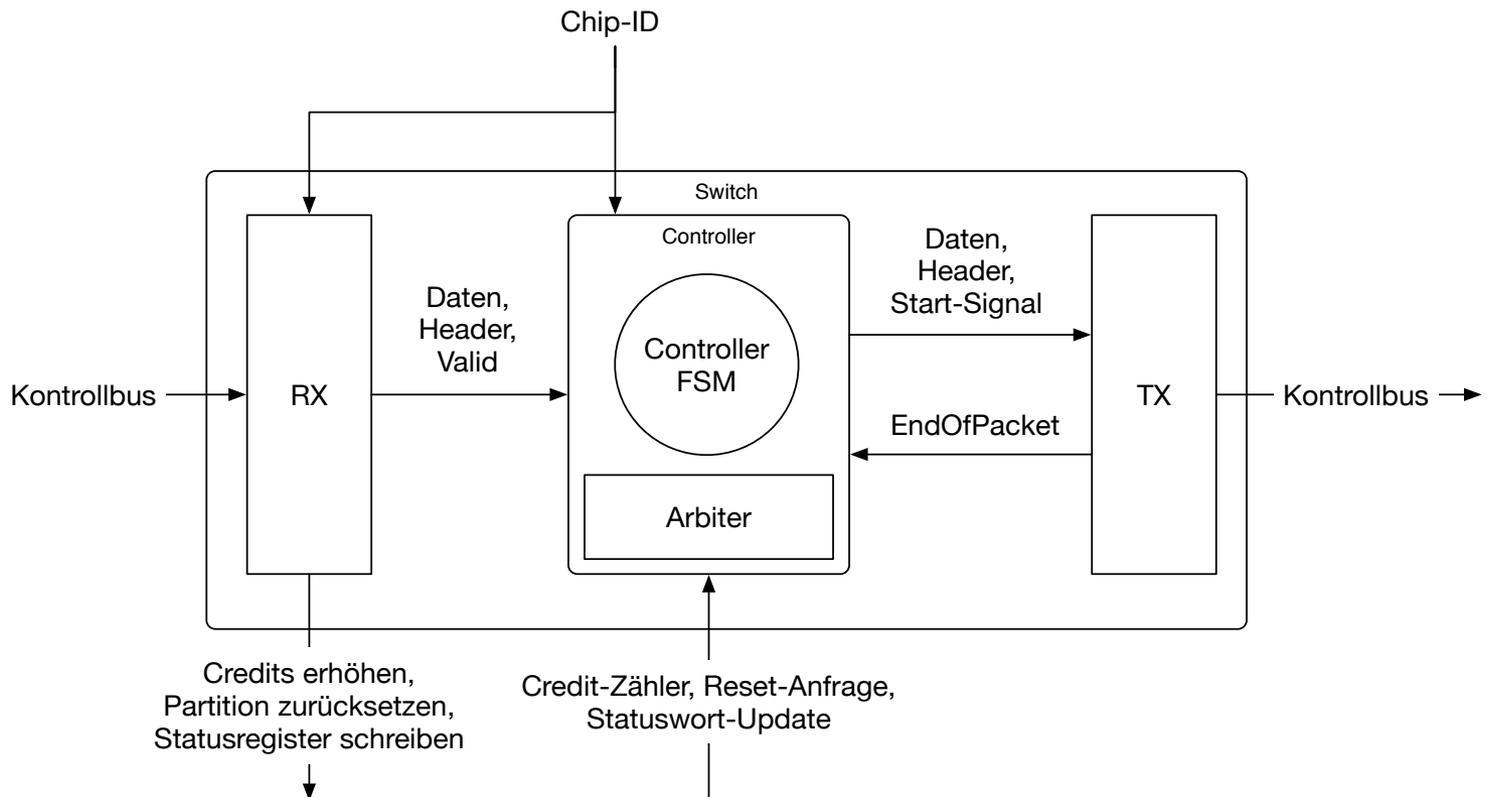
4. Entwurf und Implementierung - Komponenten

- Datenbus Switch:
 - Verbindung zwischen Partitionen und Sicherungsschicht



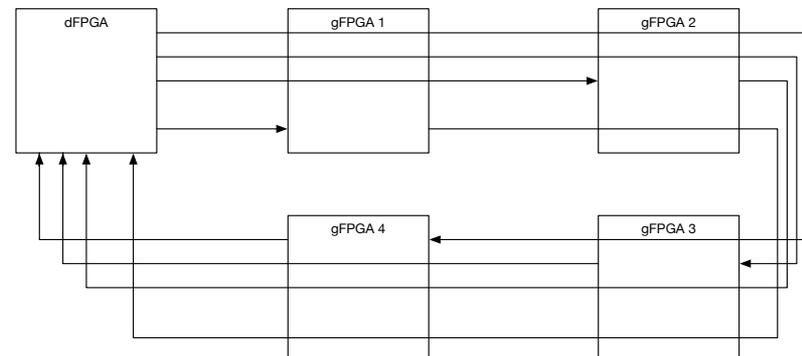
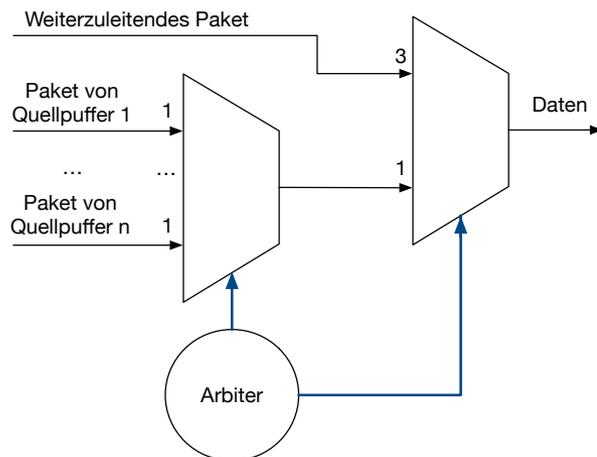
4. Entwurf und Implementierung - Komponenten

- **Kontrollbus Switch**



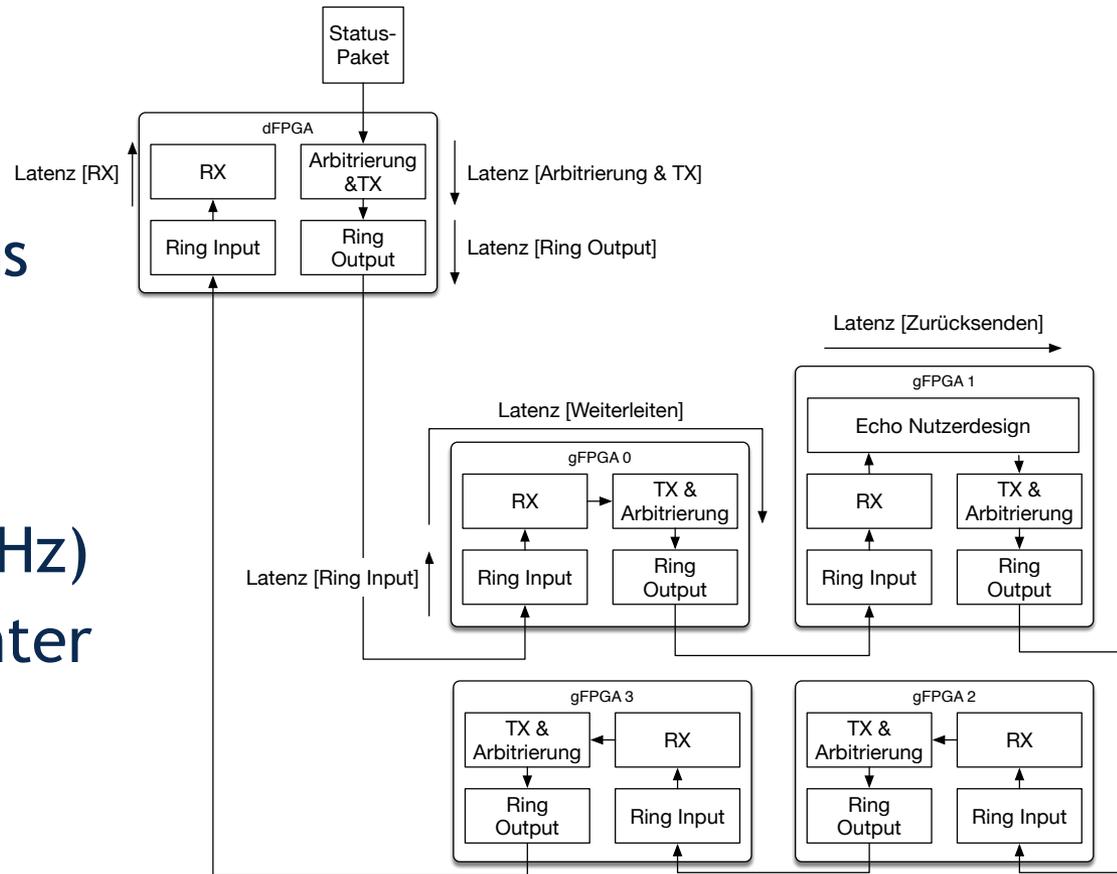
4. Entwurf und Implementierung - Komponenten

- Bus-Arbiter
 - 3:1 Gewichtung der Weiterleitung für faire Buszuteilung
 - Arbitrierung auf Paketbasis
 - gleiche Datenrate pro Kanal bei gleicher Paketgröße



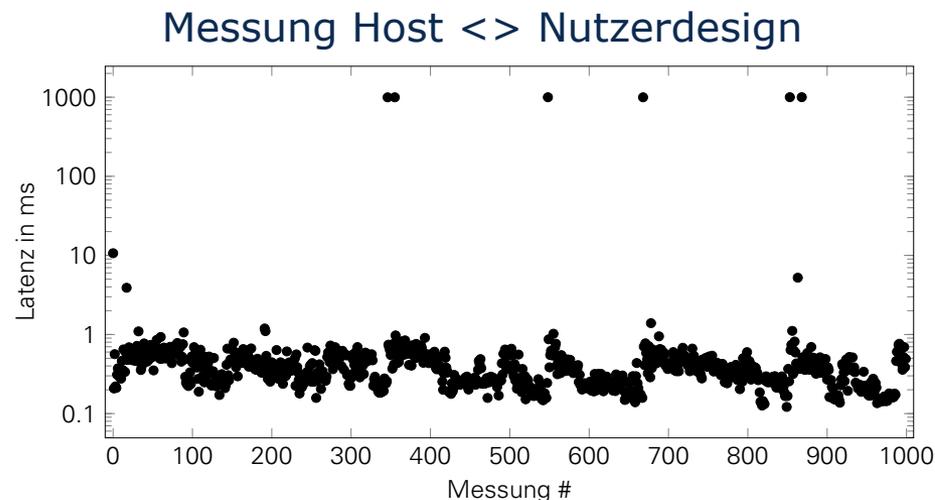
5. Auswertung - Verzögerungen

- Verzögerung Statuswort:
 - On-Board Status senden & empfangen: ~92 Taktzyklen (~612ns, 150 MHz)
 - unverändert unter Vollast



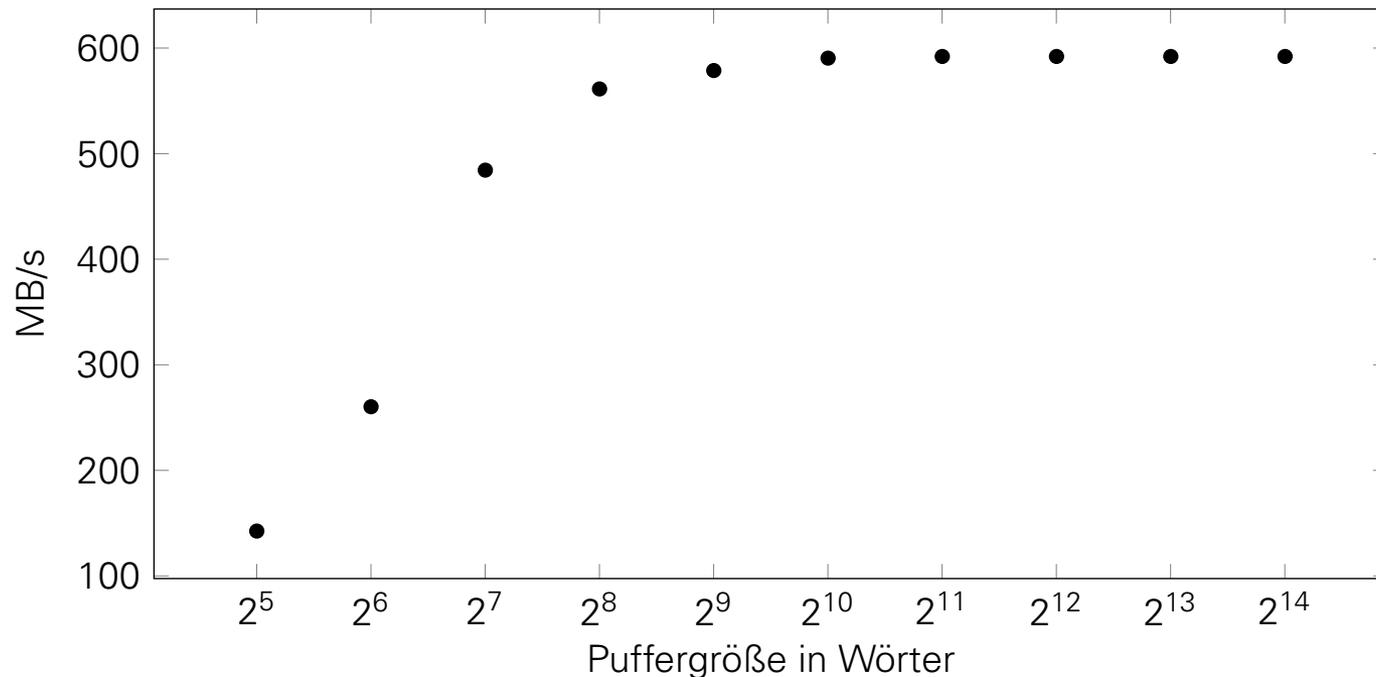
5. Auswertung - Verzögerungen

- Verzögerung Nutzdaten (hin- & rücksenden):
 - Host <> Nutzerdesign: $\sim 410\mu\text{s}$
 - dFPGA <> Nutzerdesign: ~ 80 Taktzyklen ($\sim 500\text{ns}$ bei 150 MHz)



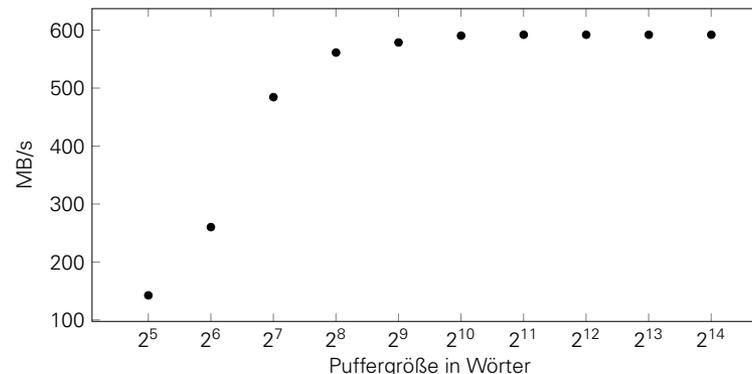
5. Auswertung - Datenrate (dFPGA - 1 Nutzerdesign)

- Max. Datenrate bei 150 MHz: 600 MB/s
- Datenrate für 1 Down- & Upstream-Kanal (On-Board):



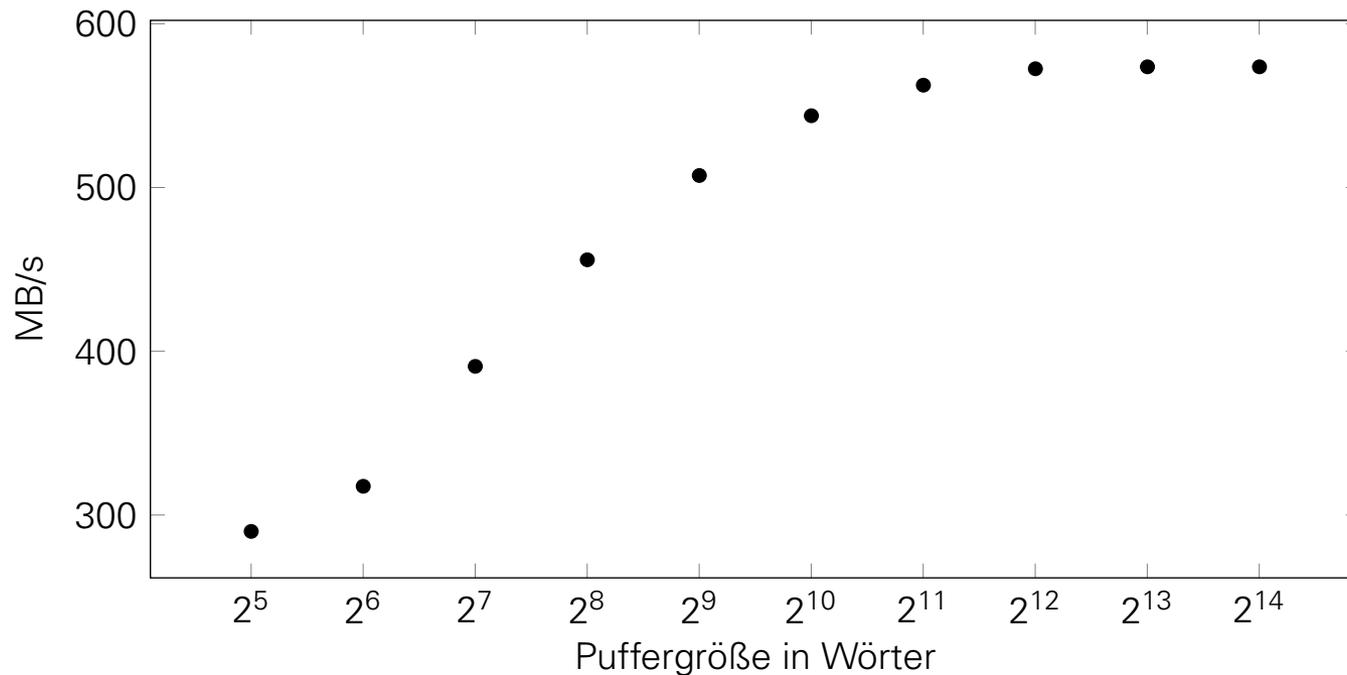
5. Auswertung - Datenrate (dFPGA - 1 Nutzerdesign)

- limitiert durch Flusskontrolle, Protokoll- und Arbitrierungsoverhead
- Flusskontrolle begrenzt bis Puffergröße: 128 Worte
 - 561,29 MB/s (93,5%)
- Maximum:
 - 592,05 MB/s (98,6%) ab 2.048 Worte



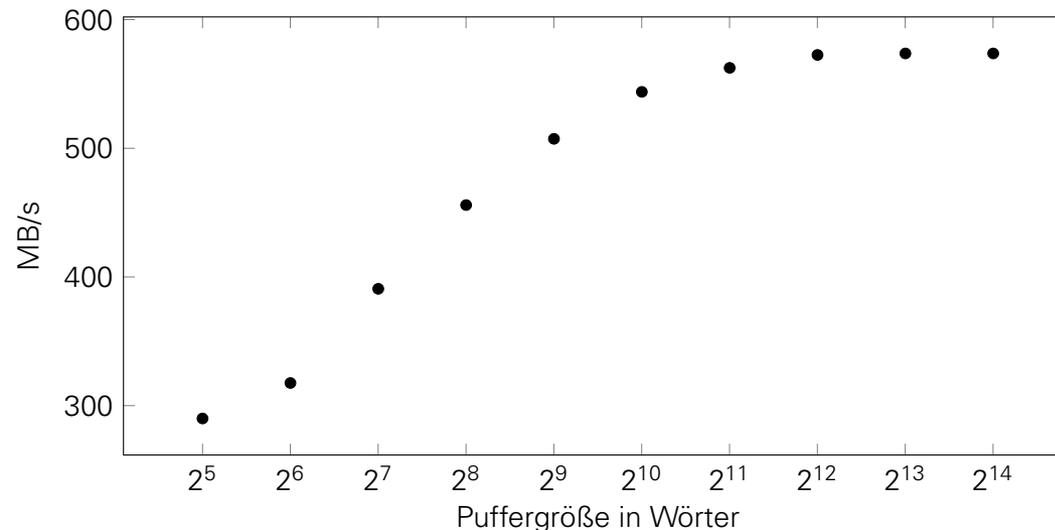
5. Auswertung - Datenrate (dFPGA - 8 Nutzerdesigns)

- Datenrate für 8x Down- & Upstream-Kanäle (On-Board)



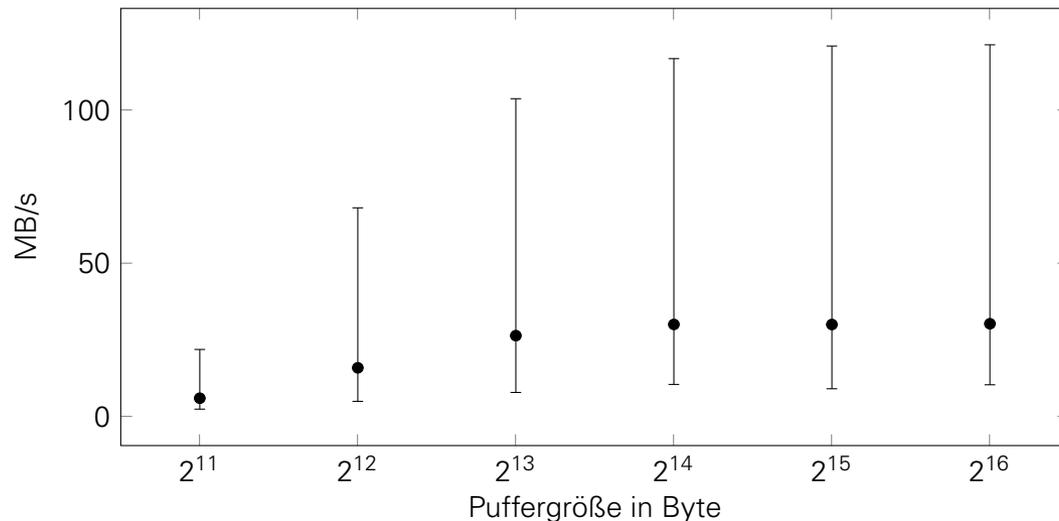
5. Auswertung - Datenrate (dFPGA - 8 Nutzerdesigns)

- Geringere Gesamtdatenrate aufgrund kleiner Pakete
- Maximum: 573,69 MB/s (95,6%) ab Puffergröße von 8.192 Worten
- gleiche, faire Datenrate pro Kanal
- Puffergröße von 1.024 Worten als Kompromiss zwischen Datenrate und Ressourcenverbrauch



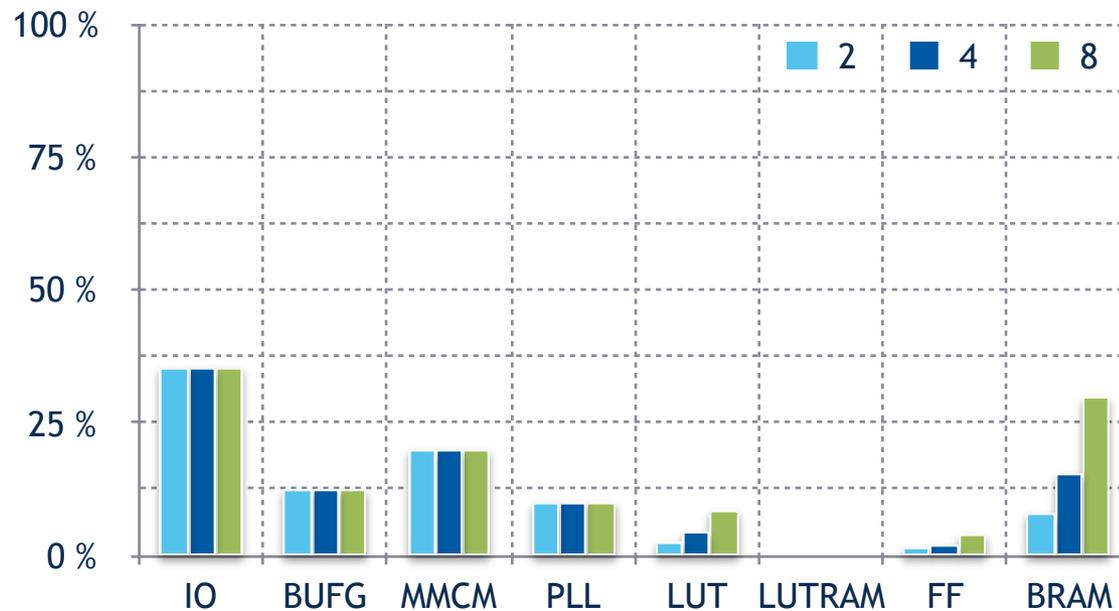
5. Auswertung - Datenrate (Host - 8 Nutzerdesigns)

- Datenrate für 8 Down- & Upstream-Kanäle zwischen Host & Nutzerdesigns
 - Anstieg \emptyset Datenrate bis 16 KiB Puffer, danach kein weiterer Einfluss messbar
 - \emptyset 30 MB/s, max. 120,9 MB/s, min. 9 MB/s



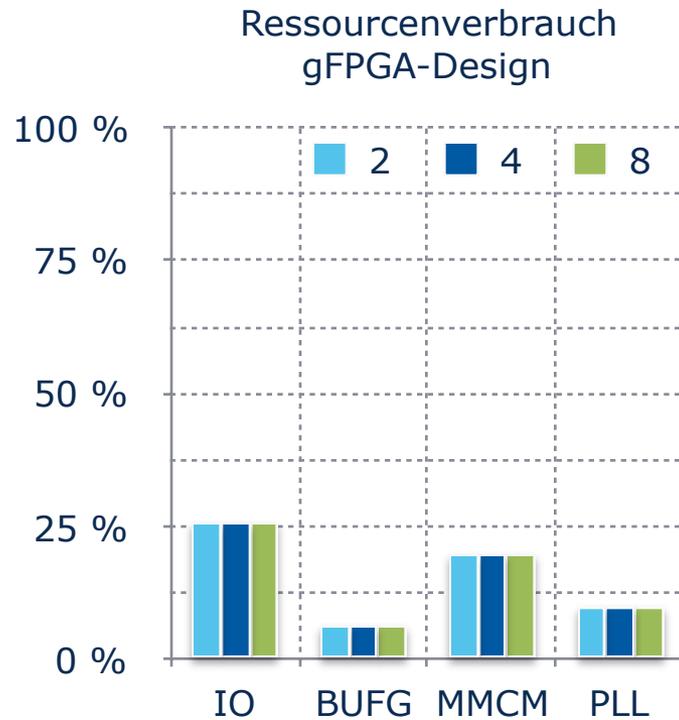
5. Auswertung - Ressourcenverbrauch

Ressourcenverbrauch dFPGA Design



Verfügbare Ressourcen	
IO	400/500
BUFG	32
MMCM	10
PLL	10
LUT	203.800
LUTRAM	64.000
FF	407.600
BRAM	445

5. Auswertung - Ressourcenverbrauch



Ressource	Verfügbar	gFPGA		
		2	4	8
LUT	203.800	778	1.273	1.969
		0,38%	0,62%	0,97%
LUTRAM	64.000	32	32	32
		0,05%	0,05%	0,05%
Flip-Flops	407.600	751	1.073	1.529
		0,18%	0,26%	0,38%
Block-RAM	445	4	8	16
		0,9%	1,8 %	3,6%

5. Auswertung - Ressourcenverbrauch

- ausreichend Logik- und Speicherressourcen für 2-8 Nutzerdesigns oder Erweiterung des Backbones
- 8 MMCM für Custom-Takte zuteilbar

6. Zusammenfassung

- funktionsfähige & praktisch getestete VHDL-Implementierung des Backbones
 - geringe Verzögerung auf dem Kontrollbus
- max. Taktfrequenz: 150 bzw. 171,4 MHz
- Datenrate:
 - Flaschenhals: Host-dFPGA-Kanal
 - On-Board Optimierungen: Pipelining zum Verdecken von Arbitrierungsoverhead, Mindest-Paketgröße, Taktfrequenz erhöhen
- ausreichend Platz auf FPGAs für mehrere Designs oder Erweiterung des Backbones, z.B.
 - Zugriff auf Speicher
 - ICAP Zugriff für dynamische Rekonfiguration

Quellen

- [1] http://www.dinigroup.com/product/data/DNK7_F5PCIe/files/Hardware_Manual_DNK7_F5_PCIe_REV4.pdf

Anhang

Anhang - Abwägung getrennter Daten- & Kontrollbus

- Vorteil: 25% höhere max. Datenrate (40-Bit Bus)
- Nachteil: höhere Design-Komplexität
 - Multiplexen von Daten- und Kontrollpaketen
 - höhere Priorität für Kontrollpakete, um niedrige Latenz zu gewährleisten

Anhang - Ressourcenverbrauch

Ressource	Verfügbar	dFPGA			gFPGA		
		2	4	8	2	4	8
LUT	203.800	5.174 2,54%	9.024 4,43%	17.325 8,5%	778 0,38%	1.273 0,62%	1.969 0,97%
LUTRAM	64.000	80 0,13%	80 0,13%	80 0,13%	32 0,05%	32 0,05%	32 0,05%
Flip-Flops	407.600	5.490 1,35%	9.203 2,26%	16.584 4,01%	751 0,18%	1.073 0,26%	1.529 0,38%
Block-RAM	445	35 7,88%	69 15,51%	133 29,89%	4 0,9%	8 1,8%	16 3,6%

Tabelle 4.6: Verbrauch ausgewählter Ressourcen für 2, 4 und 8 Partitionen pro gFPGA