

Statusvortrag Diplomarbeit

Implementierung einer adaptiven Speicherarchitektur für einen heterogenen SHAP-Mehrkernprozessor mit Hardwarebeschleuniger

Marco Zulkowski
Marco.Zulkowski@mailbox.tu-dresden.de

Dresden, 26.03.2013



Gliederung

- 1. Motivation**
- 2. HMP-Systeme und deren Bussysteme**
- 3. Heterogener SHAP-Mehrkernprozessor (Konzept)**
- 4. Zusammenfassung**

1. Motivation

Amdahl's Law (1967)

- Beschreibung eines Geschwindigkeitszuwachses (speedup)
 - bei Ausführung eines Programms mit n Prozessoren
 - mit parallelisierbarem Zeitanteil f des Programms

$$Speedup_{\text{parallel}}(f, n) = \frac{1}{(1-f) + \frac{f}{n}}$$

- Damalige Schlussfolgerung:
Typische Werte für $(1-f)$ sind groß genug um Einzelprozessoren zu favorisieren

1. Motivation

*Everyone knows Amdahl's Law, but quickly forgets it.
-Dr. Thomas Puzak, IBM, 2007*

- Sequenzielle Programmteile stellen nach wie vor Flaschenhals dar
 - ➔ Beschleunigen der sequenziellen Verarbeitungsleistung bei Multi-Core Prozessoren

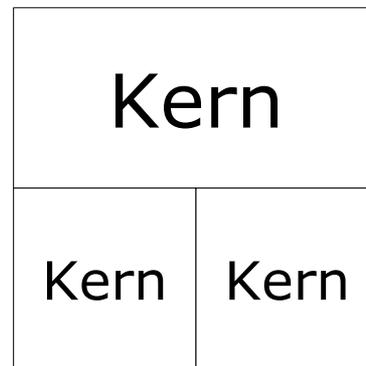
1. Motivation

Mehrkern-Varianten



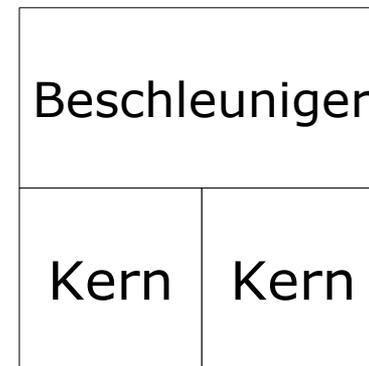
SMP

Symmetrischer
Mehrkernprozessor



AMP

Asymmetrischer
Mehrkernprozessor



HMP

Heterogener
Mehrkernprozessor

[1]

2. HMP-Systeme und deren Bussysteme

Beispiele heterogener Mehrkernprozessoren & -prozessorsysteme

- Highperformance Computing / Desktopsysteme
 - CPU + dedizierte Grafikkarte
 - General Purpose Computation on GPU (GPGPU)
 - z.B. mittels OpenCL, Cuda ...
 - AMD APU (Accelerated Processing Unit)
 - bis zu 4 Prozessorkerne und 384 Shadereinheiten
 - NVIDIA Volta (geplant für ende 2015)
 - ARM-Kern in GPU integriert
 - Cell-Prozessor

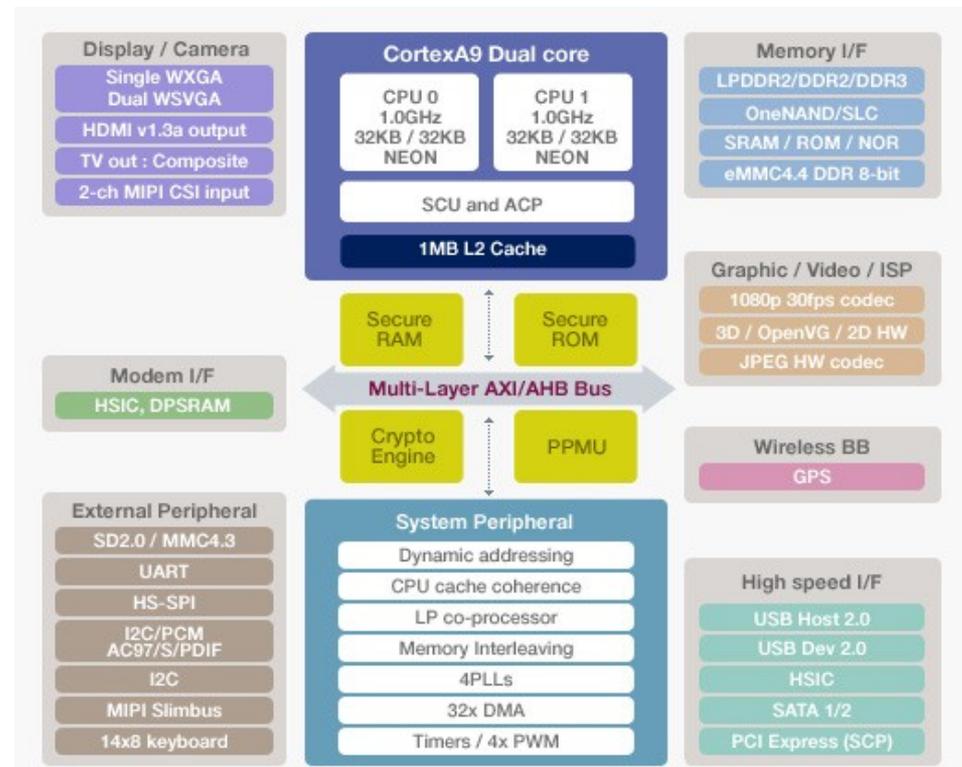
2. HMP-Systeme und deren Bussysteme

Beispiele heterogener Mehrkernprozessoren & -prozessorsysteme

Samsung Exynos 4

- 64 Bit Multilayer
AMBA AHB und AXI

(z.B. Samsung Galaxy S2)



[3]

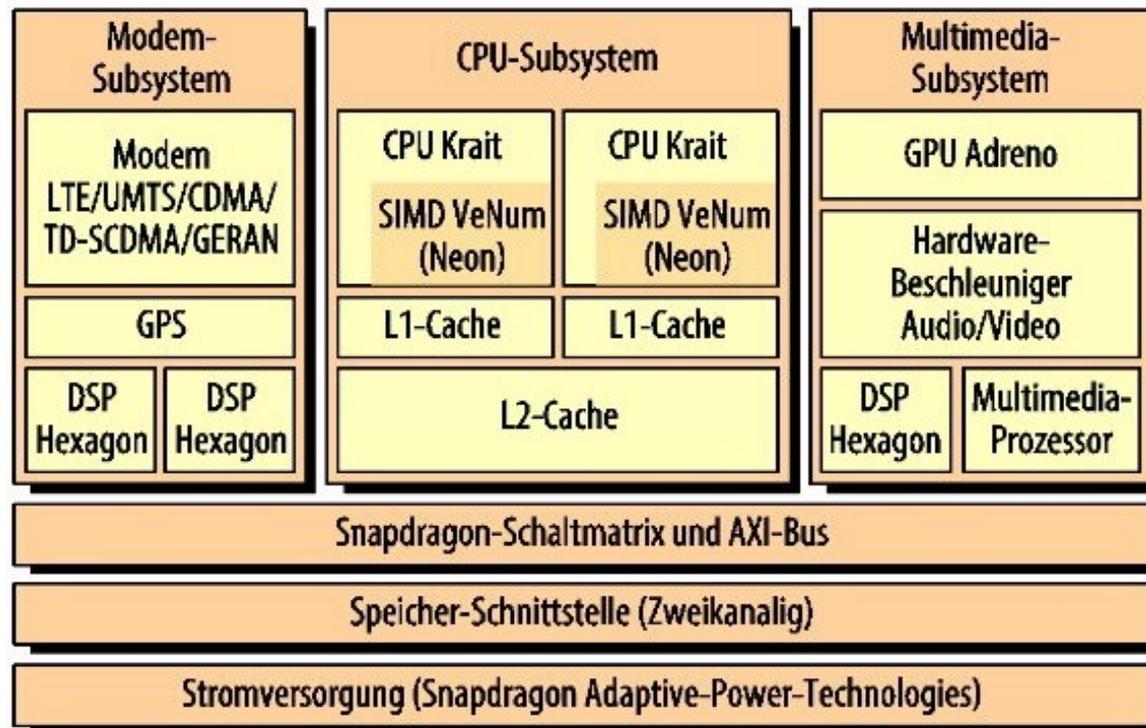
2. HMP-Systeme und deren Bussysteme

Beispiele heterogener Mehrkernprozessoren & -prozessorsysteme

Snapdragon S4

- 64 Bit AMBA AXI

(z.B. HTC One)



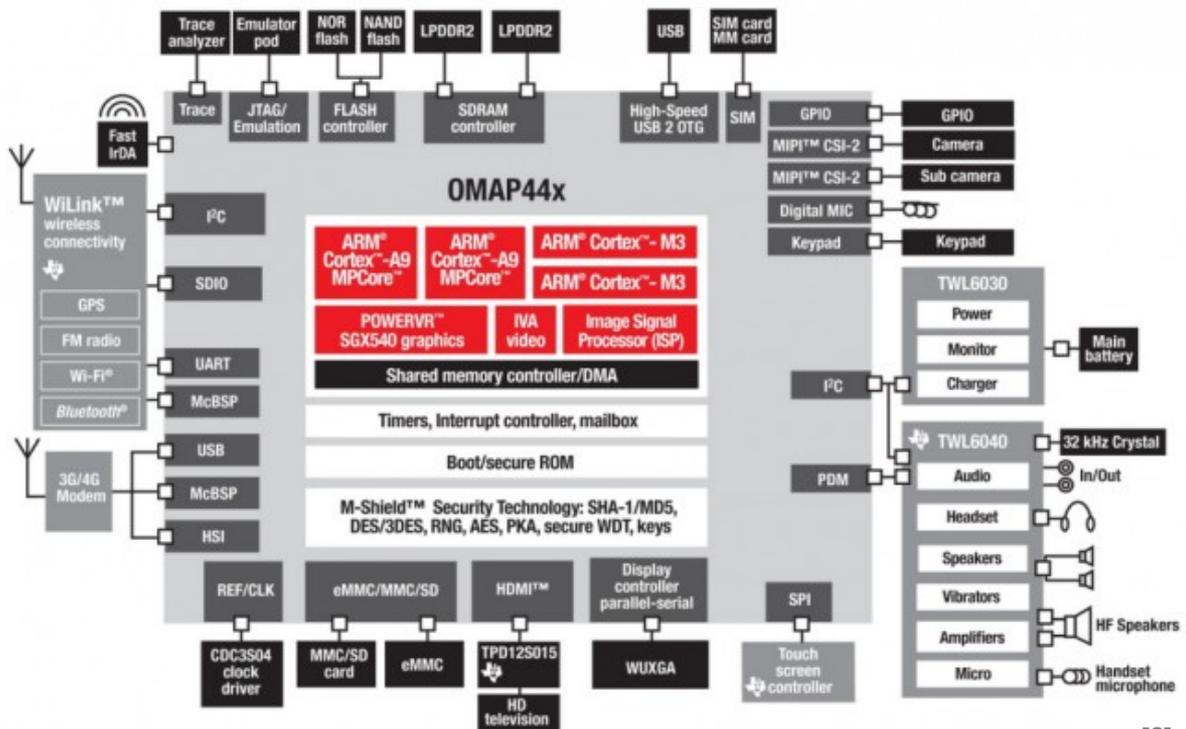
[4]

2. HMP-Systeme und deren Bussysteme

Beispiele heterogener Mehrkernprozessoren & -prozessorsysteme

TI OMAP 4470

- Network-On-Chip
(Arteris)
(z.B. Kindle Fire HD)



[6]

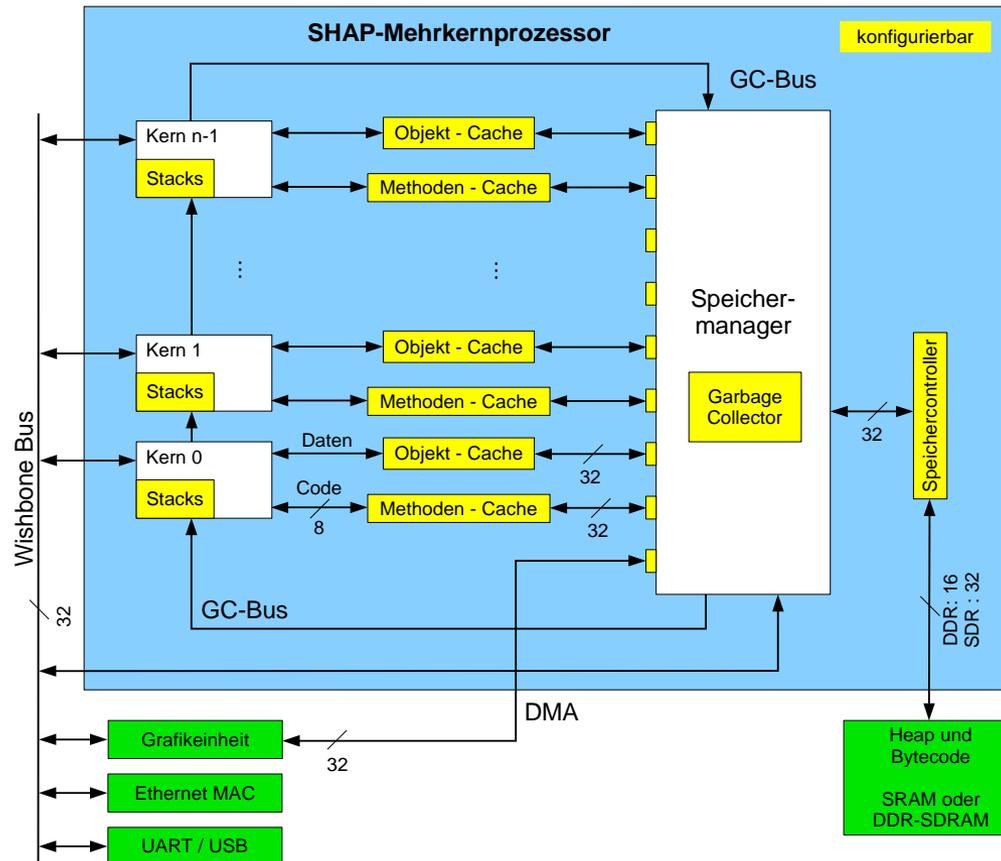
2. HMP-Systeme und deren Bussysteme

Zusammenfassung Bussysteme

	Exynos 4	Snapdragon S4	OMAP 4470
Topologie	Shared-Bus und Crossbar (Multi-Layer)	Crossbar	Network-on-Chip
Bus-Protokoll	AMBA AHB + AXI	AMBA AXI	Proprietär
Datenbreite	64	64	128
Arbitrierung	Keine Angaben	Für jedes Slave-Interface extra	Für jedes Slave-Interface extra
Arbitrierungsalgorithmus	Keine Angaben	Keine Angaben	Statische Prioritäten, AgedPriority, RoundRobin, ...
Kontroll- / Datenpfade	getrennt	getrennt	getrennt

3. Heterogener SHAP-Mehrkernprozessor

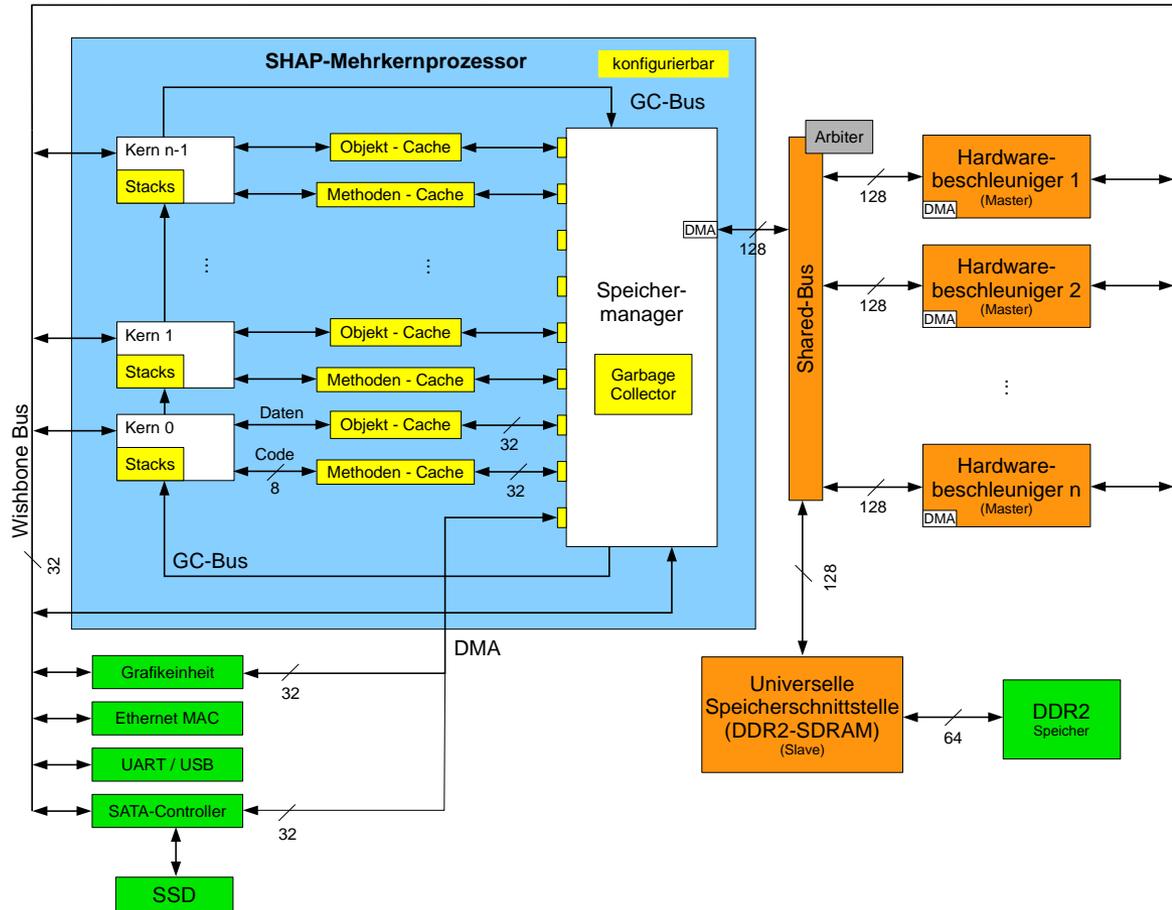
Altes System



[5]

3. Heterogener SHAP-Mehrkernprozessor

Neues System



3. Heterogener SHAP-Mehrkernprozessor

Architekturkonzept:

- Kontrollfluss über vorhandenen Wishbone-Bus
- Ersatz des integrierten Speichercontrollers
 - ➔ Einführung von Shared-Bus
- Shared-Bus:
 - Anbindung des Speichermanagers
 - Anbindung aller Hardwarebeschleuniger
 - Anbindung von DDR2-Controller
- Arbitrierung des Buszugriffes

3. Heterogener SHAP-Mehrkernprozessor

Schnittstelle

- Wiederverwendung der universellen Speicherschnittstelle aus Belegarbeit [15]
- Burstorientierte-Zugriffe (Speichercontroller)
- Adressierung durch Startadresse und Blockanzahl
- Zuordnung von Request zu Master (Tag)
- Allgemein:
 - Vereinfachte Konfiguration
 - Einfaches Protokoll

3. Heterogener SHAP-Mehrkernprozessor

Schnittstelle

	AMBA AXI 4	AMBA AHB 3	Wishbone B4	Universelle Speicherschnittstelle
Topologie	Punkt-zu-Punkt, Crossbar	Shared-Bus	Shared-Bus, Punkt-zu-Punkt, Crossbar	Shared-Bus
Datenbreite (Bit)	8 – 1024	8 - 1024	8, 16, 32, 64	beliebig
Adressierung	Bis 64 Bit	32 Bit	Bis 64 Bit	beliebig
Datenmaskierung	Byte	Nein	Byte	Nein
Tags	Ja	Nein	Ja	Ja
Komplexität	Sehr hoch	hoch	hoch	mittel

[11, 12, 13, 15]

3. Heterogener SHAP-Mehrkernprozessor

Arbitrierung

- Klassisch:
 - Feste Prioritäten
 - Zeitmultiplex (TDMA)
 - Round Robin
- Lottery Scheduling
- Credit-Based Scheduling

3. Heterogener SHAP-Mehrkernprozessor

Lottery Scheduling – Motivation

- Testresultate zeigen:
 - Niedrige Latenzen für hoch-priorisierte Master
 - Bei gleichzeitiger Bandbreitengarantie für andere Master
 - Aber keine harte Echtzeit
- Stochastisch fair
 - Erwartete Zuweisung ist proportional zu Losverhältnissen
- Keine Starvation
 - Jeder Master mit Los „gewinnt“ früher oder später

3. Heterogener SHAP-Mehrkernprozessor

Lottery Scheduling – Motivation

- 4-Port ATM-Switch:
 - Port 1 – Port 3 mit Bandbreitenverhältnis (1:1:4)
 - Port 4 mit niedriger Latenz

Algorithmus	Port 4 Latenz (Takte/Wort)	Port 4 Bandbreite (%)	Port 3 Bandbreite (%)	Port 2 Bandbreite (%)	Port 1 Bandbreite (%)
Static priority	1,39	9,69	45,72	44,58	0,01
TDMA	9,84	10,09	47,29	21,31	21,31
Lottery	1,4	9,67	59,03	17,00	14,30

[8]

3. Heterogener SHAP-Mehrkernprozessor

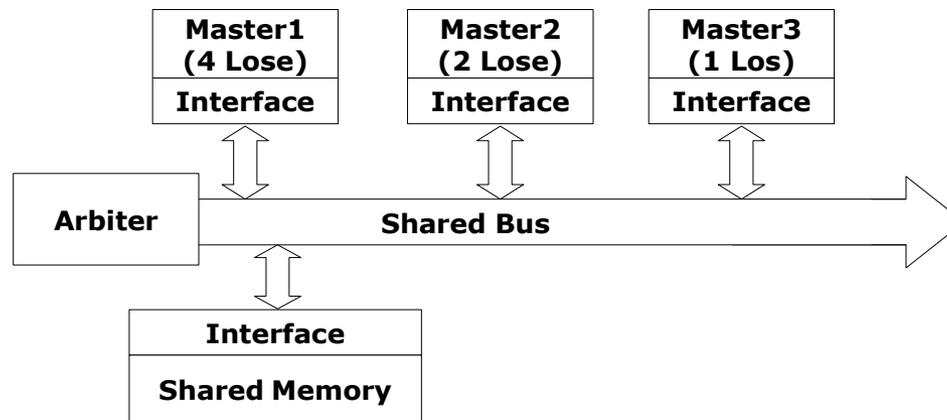
Lottery Scheduling [2, 7, 8]

- 1994 am MIT entwickelt (Prozess- / Threadscheduling)
- Zufallsbasierter Arbitrierungsalgorithmus
- 2001 implementierung des LOTTERYBUS für SoC-Busse
- Diverse Weiterentwicklungen

3. Heterogener SHAP-Mehrkernprozessor

Lottery Scheduling - Prinzip

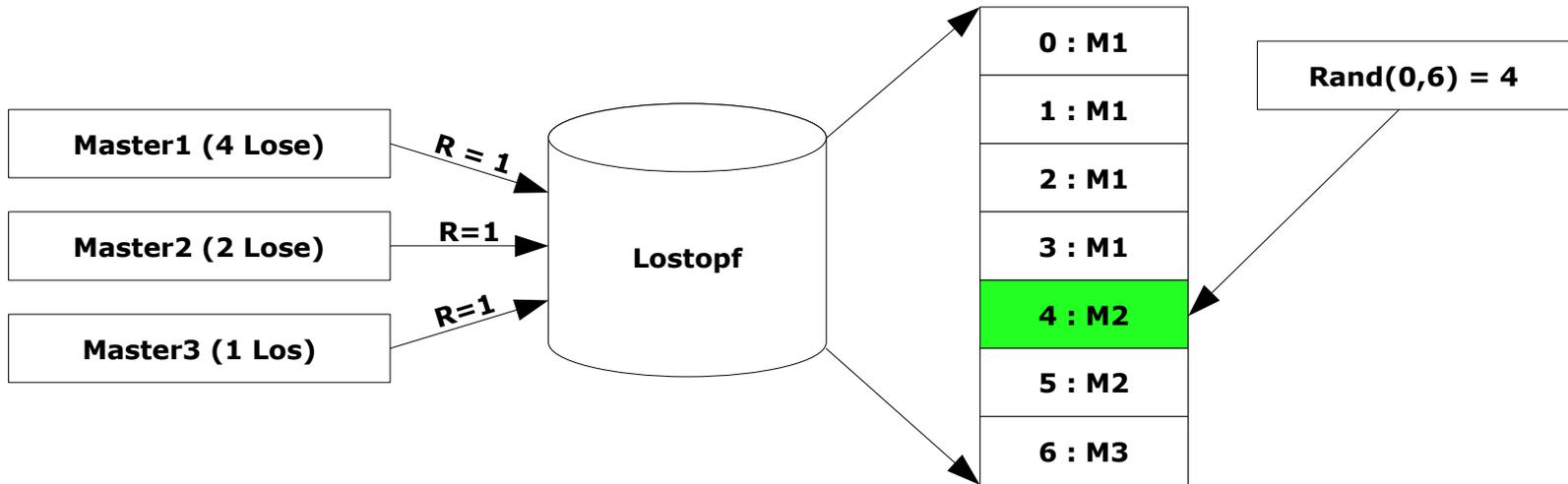
1. Master bekommen Lose (Tickets) zugeordnet
2. Jeder Master mit Request-Wunsch gibt seine Lose in einen Lostopf
3. Arbiter zieht per Zufall ein Los aus dem Lostopf
4. Master dem das gezogene Los gehört bekommt Buszugriff



[8]

3. Heterogener SHAP-Mehrkernprozessor

Lottery Scheduling - Prinzip



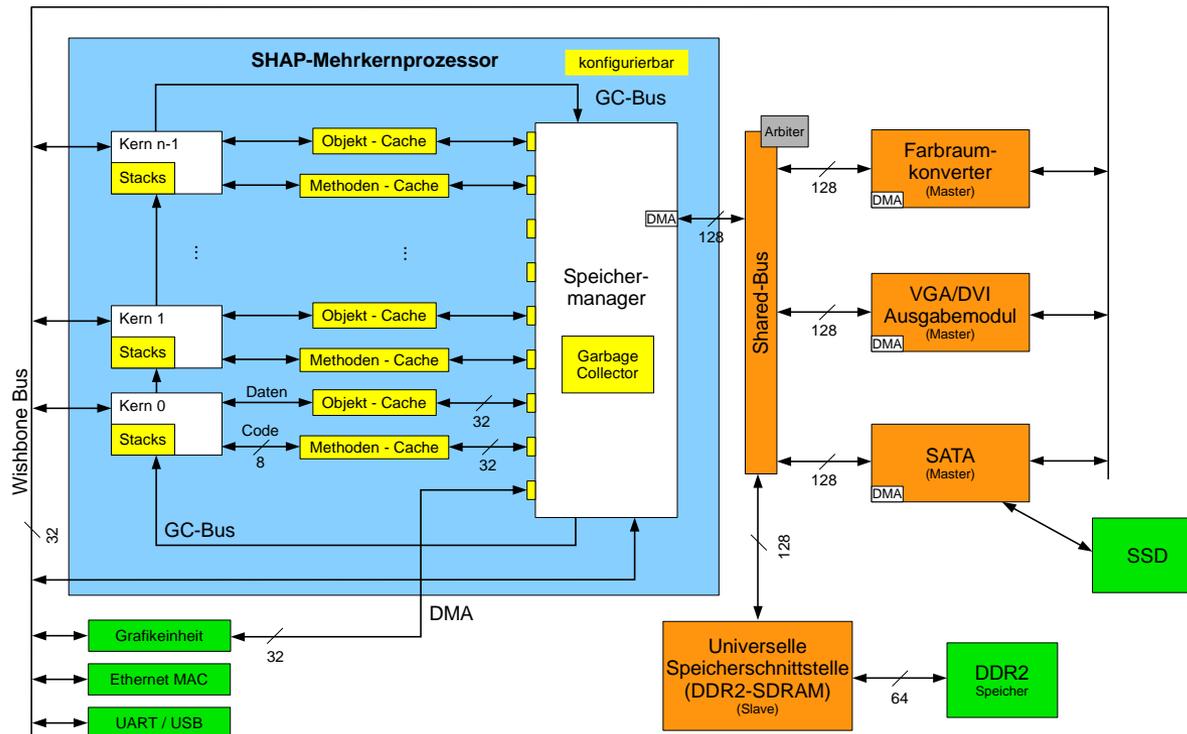
[8]

R...Request

3. Heterogener SHAP-Mehrkernprozessor

Beispielsystem:

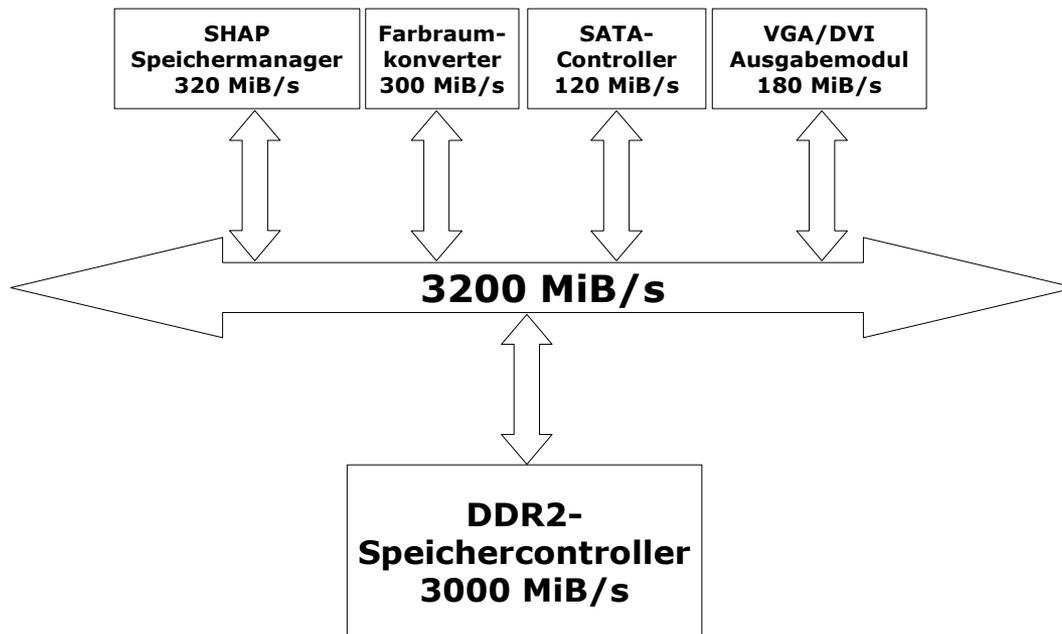
Farbraumkonvertierung von YCbCr nach RGB mittels Hardwarebeschleuniger



3. Heterogener SHAP-Mehrkernprozessor

Beispielsystem:

Farbraumkonvertierung von YCbCr nach RGB mittels Hardwarebeschleuniger
Bandbreitenbedarf: FullHD ~ 1GiB/s



4. Zusammenfassung

- Kurzer Überblick zu heterogenen Mehrkernprozessorsystemen
- Vorstellung des Konzepts
 - Vergleich neues und altes System
 - Arbitrierung (Lotterybus)
 - Beispielsystem

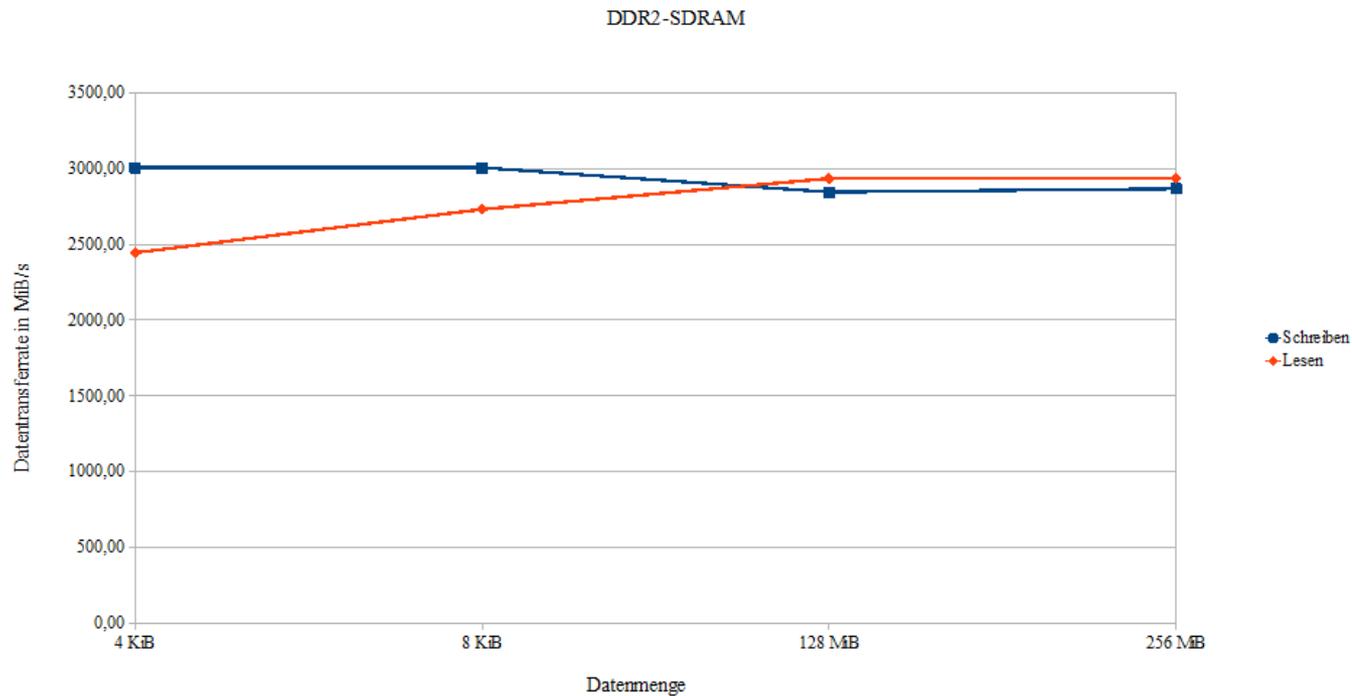
Ausgewählte Quellen

- [1] V. Gupta, R. Nathuji and K. Schwan. An Analysis of Power Reduction in Datacenter using Heterogeneous Chip Multiprocessors. 2011
- [2] M. D. Hill and Michael R. Marty. Amdahl's Law in the Multicore Era. 2008
- [3] <http://www.samsung.com/global/business/semiconductor/file/product/Exynos4210-0-0.jpg>
- [4] <http://www.elektroniknet.de/typo3temp/pics/31183a6fab.jpg>
- [5] M. Zabel. Effiziente Mehrkernarchitektur für eingebettete Java-Bytecode-Prozessoren. Dissertation TU Dresden. 2011
- [6] <http://www.extremetech.com/wp-content/uploads/2011/12/Omap4470-640x416.png>
- [7] K. Lahiri, A. Raghunathan and G. Lakshminarayana. LOTTERYBUS: A New High-Performance Communication Architecture for System-On-Chip Designs. 2001
- [8] K. Lahiri, A. Raghunathan and G. Lakshminarayana. The LOTTERYBUS On-Chip-Communication Architecture. 2006
- [9] C. A. Waldspurger and W. E. Weihl. Lottery Scheduling: Flexible Proportional-Share Resource Management. 1994
- [10] <http://www.heise.de/newsticker/meldung/GTC-2013-ARM-Kerne-auf-Desktop-GPUs-erst-ab-Volta-1826893.html>. 21.03.2013 17:28
- [11] Wishbone B4
- [12] AMBA AXI Protocol Version 2.0 Specification & AMBA Specification Revision 2.0
- [13] On-Chip Communication Architectures: System on Chip Interconnect von Sudeep Pasricha, Nikil Dutt
- [14] <https://www.uplinq.com/2011/sites/default/files/slides/Developing-Next-Gen-3D-Games-Snapdragon-8660-Devices.pdf>
- [15] M. Zulkowski. Universelle Speicherschnittstelle für große externe Speicher. Großer Beleg TU Dresden. 2012
- [16] Clive „Max“ Maxfield. FPGAs: World Class Designs. Newnes, 2010. - ISBN 978-1-85617-621-7
- [17] IBM. Cell Broadband Engine Architecture. Version 1.02. 11. Oktober 2007

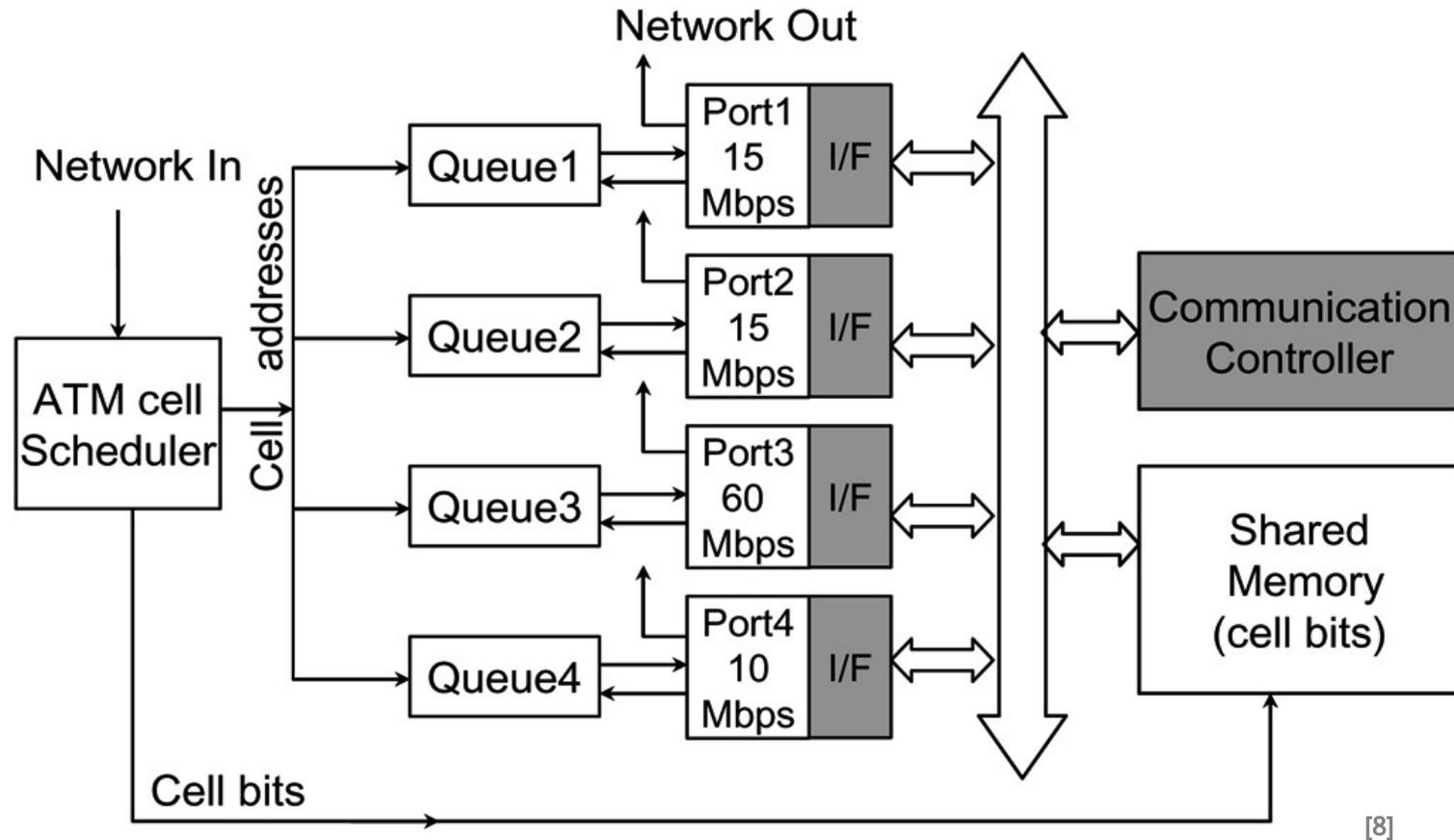
**Vielen Dank
für ihre Aufmerksamkeit!**

Messwerte DDR2-Controller

- Messung mit Blockgröße von 128Byte

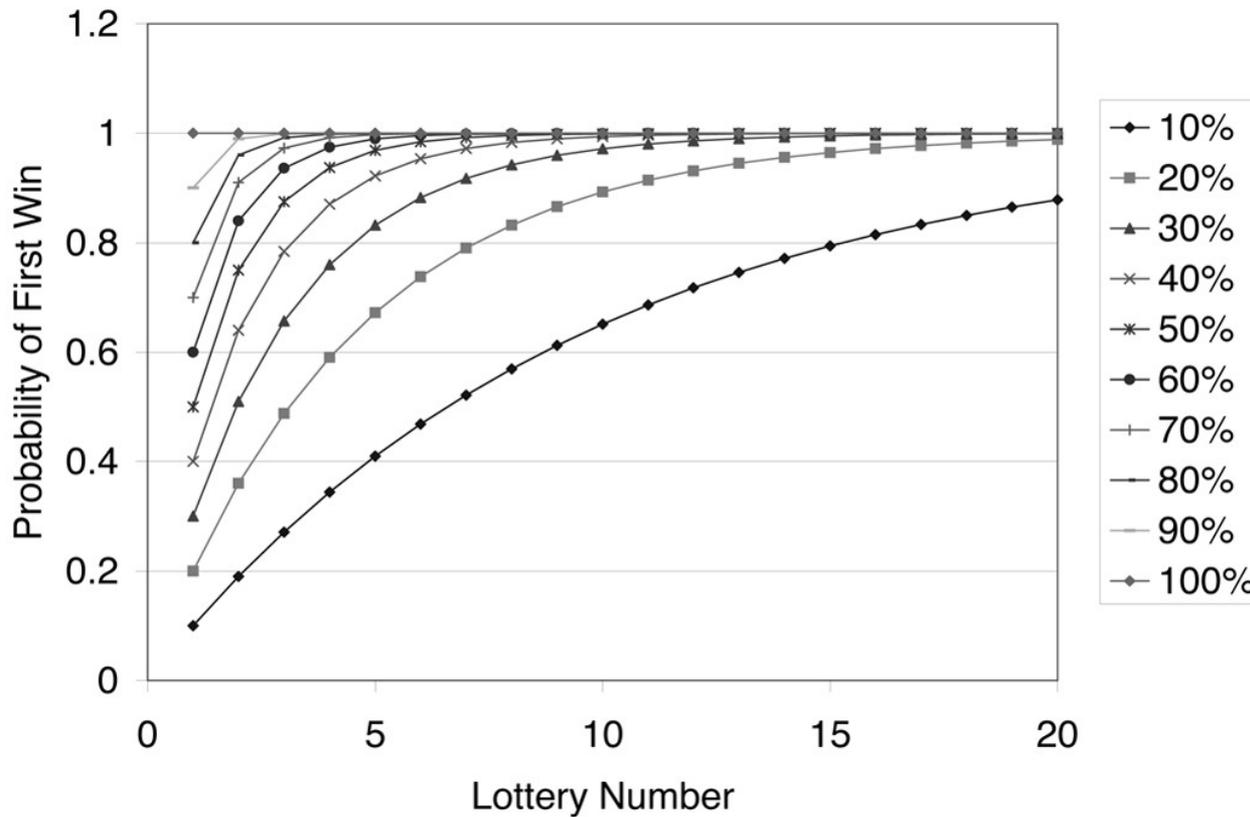


4-Port ATM-Switch



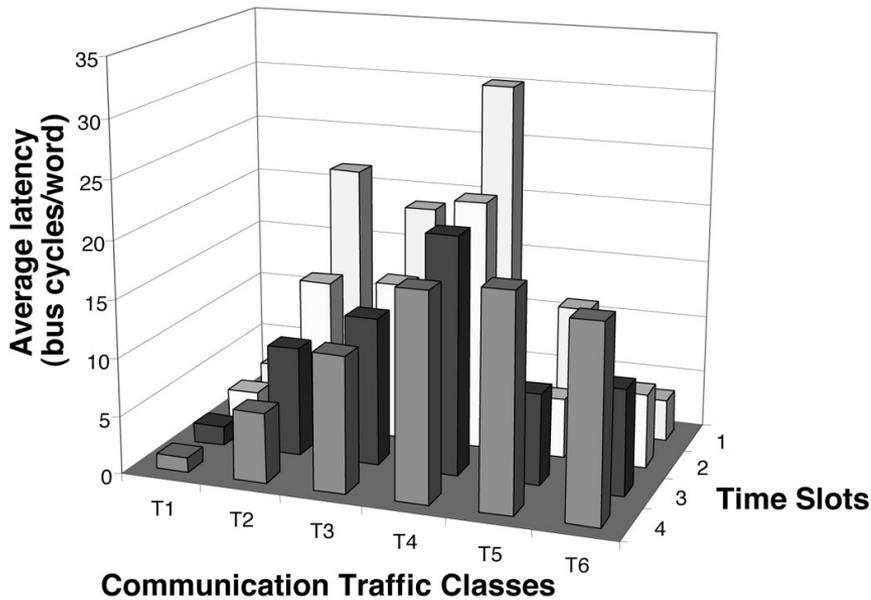
Grant-Wahrscheinlichkeit

$$p(t_i, T, k) = \frac{t_i}{T} * \sum_{n=0}^{k-1} \left(1 - \frac{t_i}{T}\right)^n$$

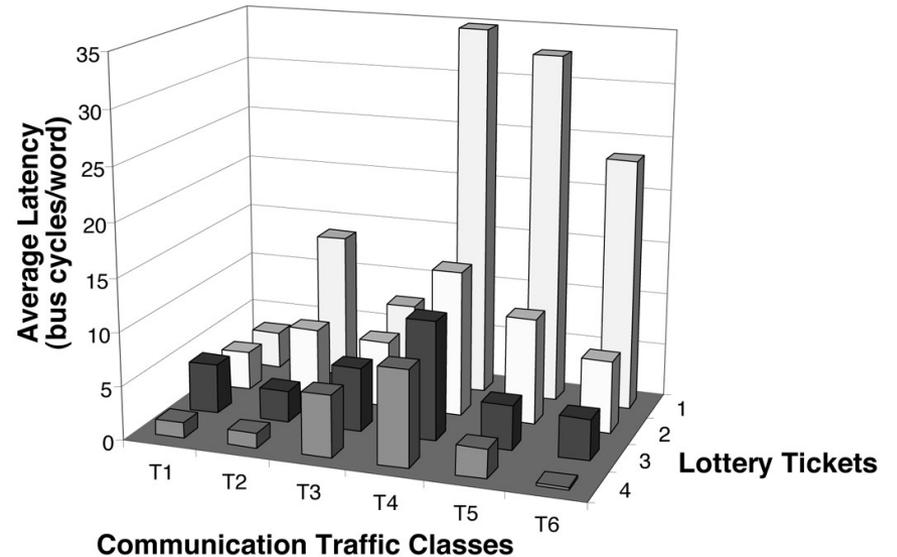


[8]

TDMA vs. Lottery



(a) TDMA-based communication architecture



(b) LOTTERYBUS communication architecture

[8]

Hardwarebeschleuniger

Allgemein:



Beispiel- implementierung:

