



# Skalierbare Rechensysteme für Echtzeitanwendungen

Stefan Aust

Institut für Informatik

TU Clausthal

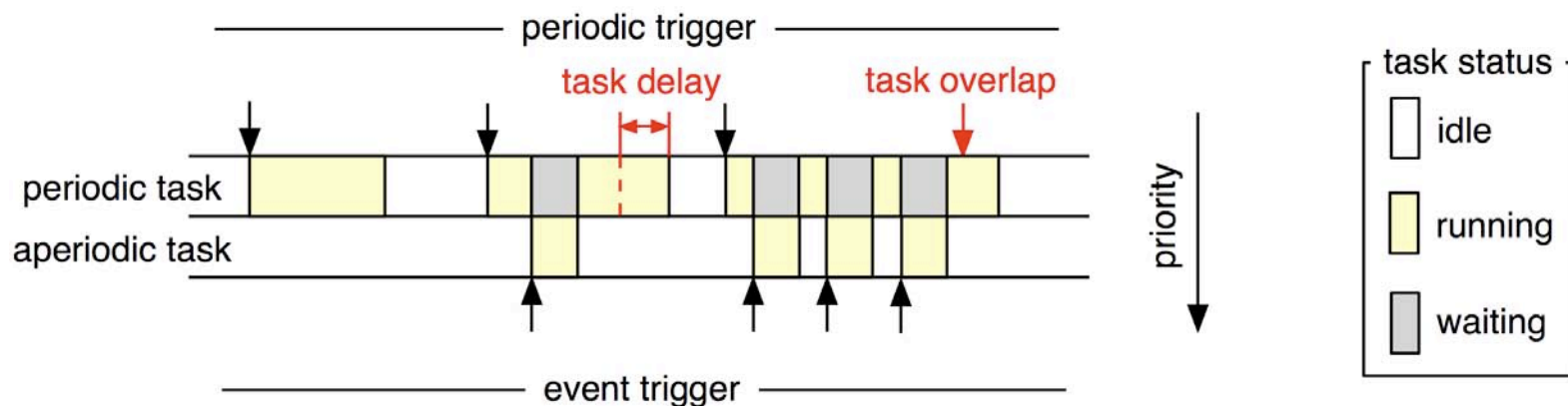
Echtzeit 2011:

„Herausforderungen durch Echtzeitbetrieb“

3. und 4. November 2011 in Boppard am Rhein

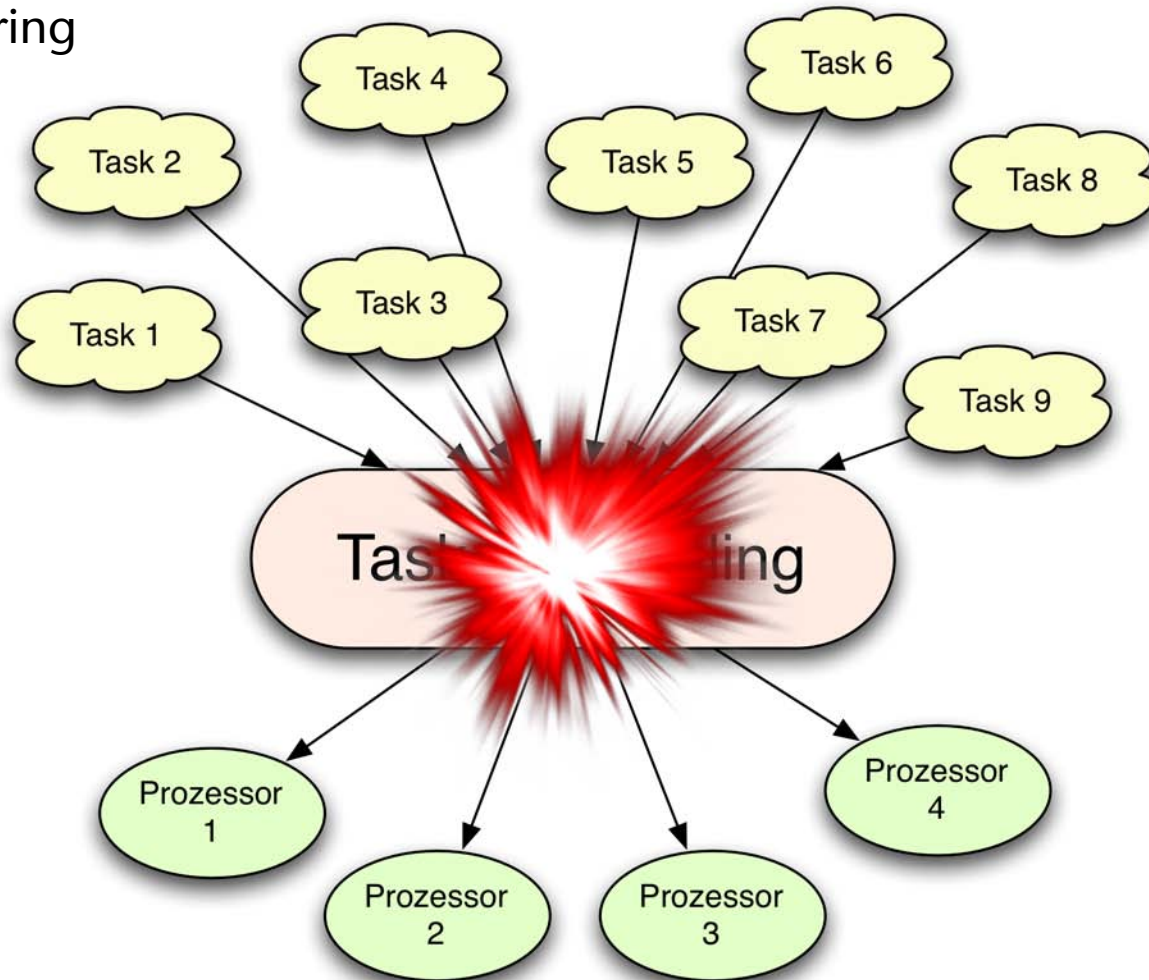
# Multitasking

- Task Delay
    - Die Ausführung einer Task wird durch Tasks höherer Priorität unterbrochen, so dass sich die Zeit der Ausführung signifikant verlängert
  - Task Overlap
    - Die Zeit der Ausführung einer Task verlängert sich über den nächsten Triggerzeitpunkt hinaus
- ☞ Multitasking führt leicht zu einem nicht deterministischen Verhalten



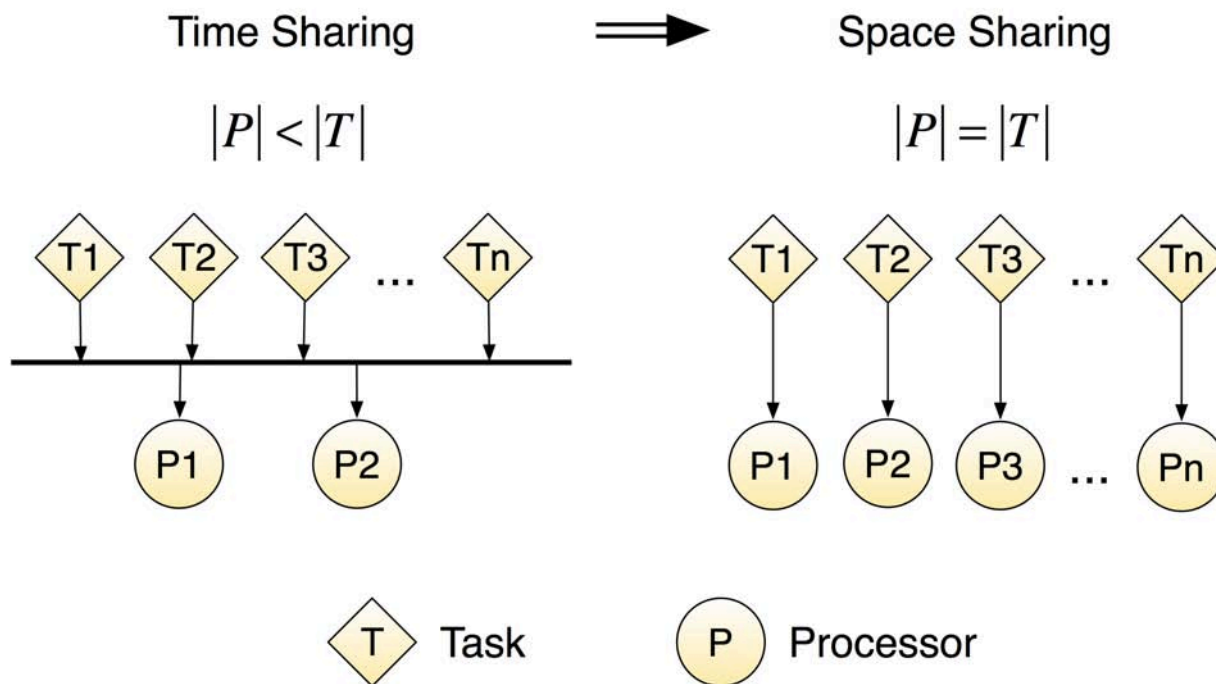
# Multitasking

- Time Sharing



## Von Multi-Core zu Many-Core

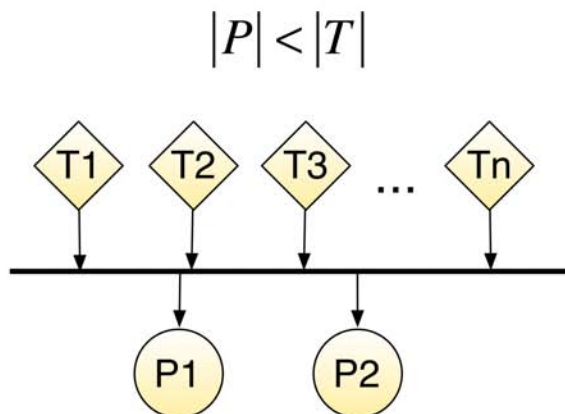
- “To pull a bigger wagon, it is easier to add more oxen than to grow a gigantic ox.” – Gropp, Lusk, Skjellum: *Using MPI*.



## Von Multi-Core zu Many-Core

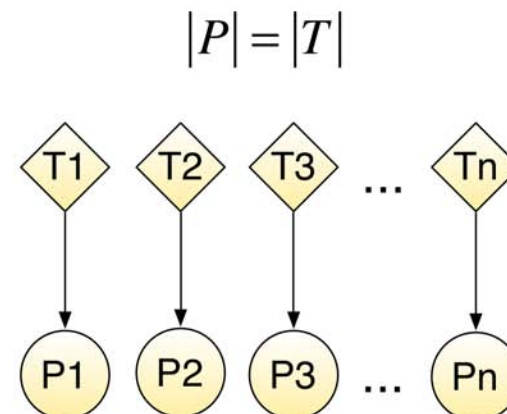
### Time Sharing

- Die Menge an ablaufbereiten Tasks ist deutlich größer als die Menge der zur Verfügung stehenden Prozessoren
- Zeitliche Aufteilung der Prozessor-Rechenleistung an die einzelnen Tasks (Multitasking)



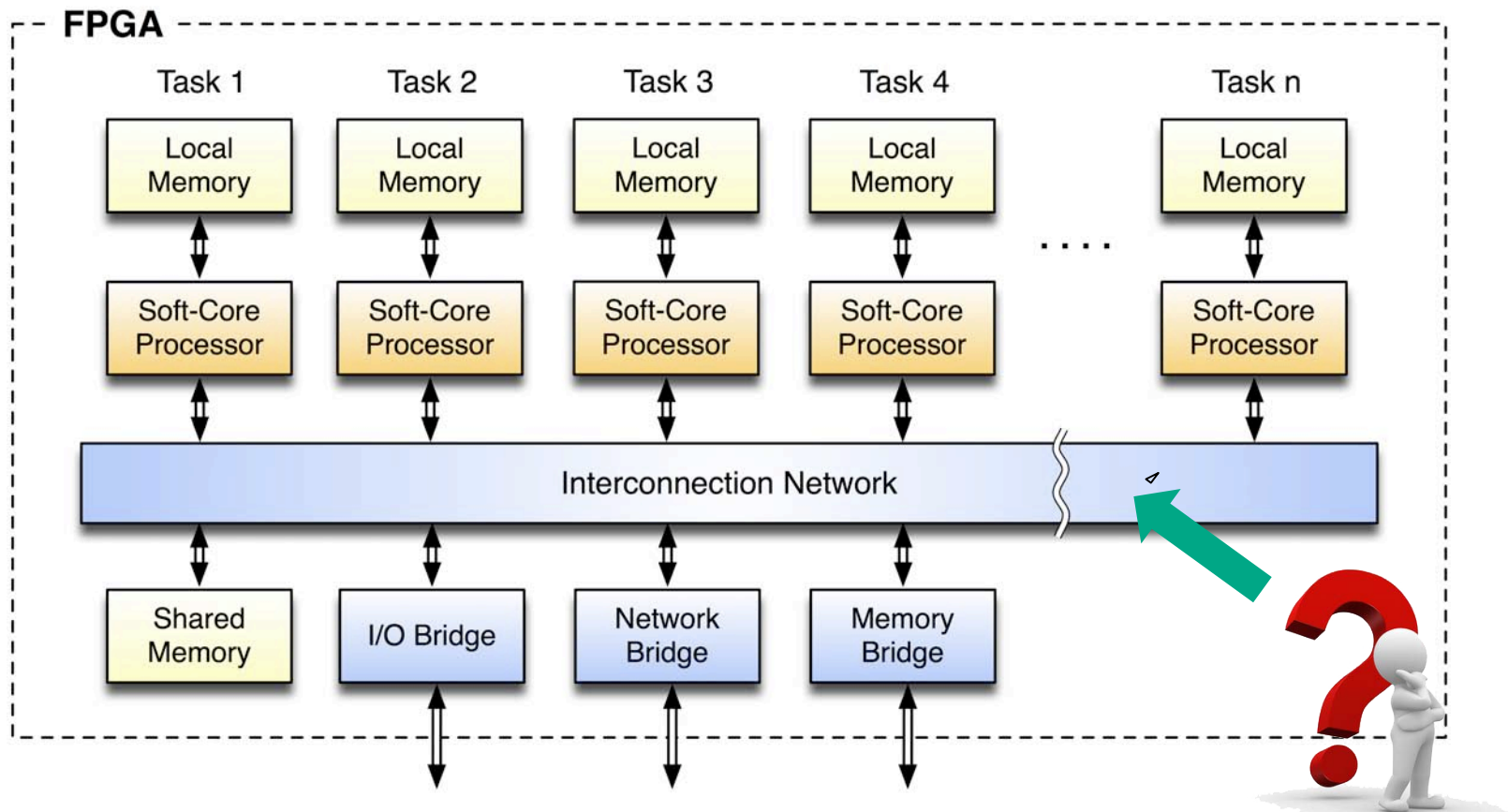
### Space Sharing

- Die Menge an Prozessoren entspricht zahlenmäßig der Menge an geplanten Tasks
- Räumliche Aufteilung der Chipfläche auf die einzelnen Tasks (**Software-First-Design**)



## Space Sharing

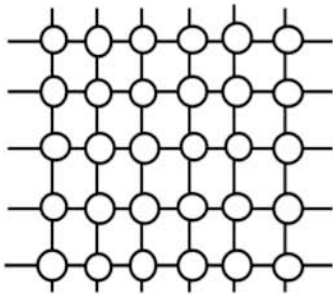
- Architektur des Multi-Prozessor System-on-Chip (MPSoC)



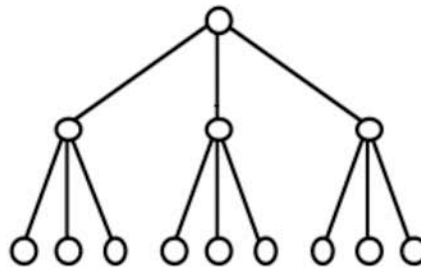
## Interprozessor-Kommunikation

- Statische Topologien für Netzwerke On-Chip (NoC)
  - Jeder Kommunikationsknoten ist zusätzlich auch **Routerknoten**
  - Mehrdimensionale Netzwerkverbindungen nötig
  - Nicht-deterministisches Verhalten aufgrund von:
    - Hotspots: zusätzliche Latenzen bei erhöhtem Datenaufkommen
    - Multihops: zeitvariante Übertragungen durch unterschiedliche Pfadlängen

a)



b)



c)

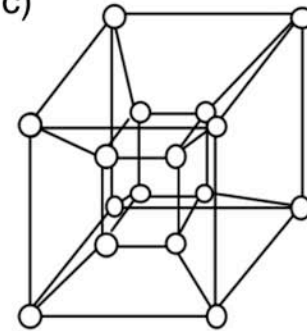


Bild: Statische Netzwerktopologien, a) 2D-Gitter, b) Baum, c) 4D-Hypercube

## Interprozessor-Kommunikation

- Dynamische Topologien für Netzwerke On-Chip (NoC)
  - Mehrstufige Verbindungsnetzwerke:
    - Bestehend aus in Stufen angeordneten Kreuzschaltern
    - Schalterstufen sind verbunden durch Permutationen
    - Jeder Eingang kann mit jedem Ausgang verbunden werden
    - Paketvermittelndes oder leitungsvermittelndes Netz

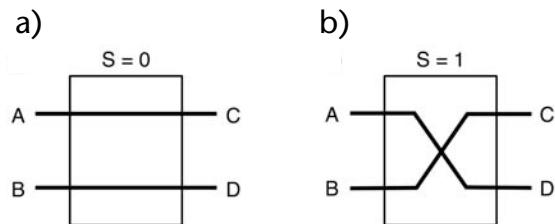


Bild: Kreuzschalter, a) gerade, b) gekreuzt

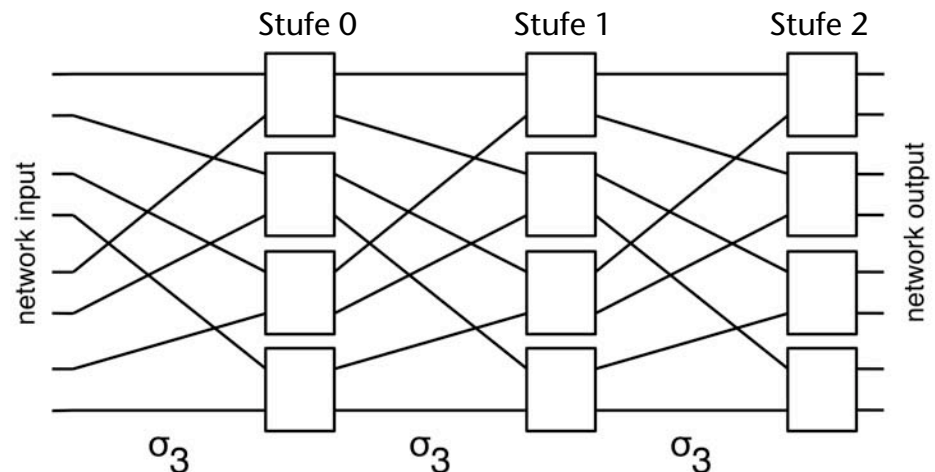


Bild: Omega-Netz der Größe  $8 \times 8$



# Interprozessor-Kommunikation

- Nicht-blockierungsfreie Netze

- Minimum an Hardware um jeden Eingang mit jedem Ausgang zu verbinden
- Äußerst einfache Routingalgorithmen durch Pfadeindeutigkeit
- Nicht alle Verbindungen zur gleichen Zeit möglich

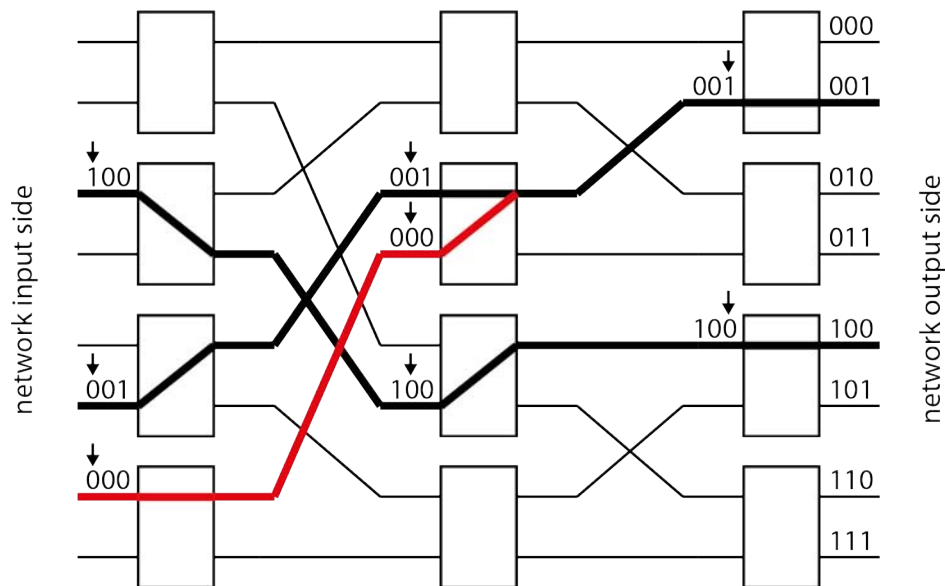


Bild: Routing im Baseline-Netz der Größe 8x8

## Interprozessor-Kommunikation

### ■ Blockierungsfreie Netze

- Erweiterung auf  $2\log_2 N - 1$  Stufen mit  $N/2$   $2 \times 2$  Kreuzschalter
- Jeder Eingang kann zu jeder Zeit mit jedem freien Ausgang verbunden werden
- Maximale Bandbreite skaliert mit  $O(N)$

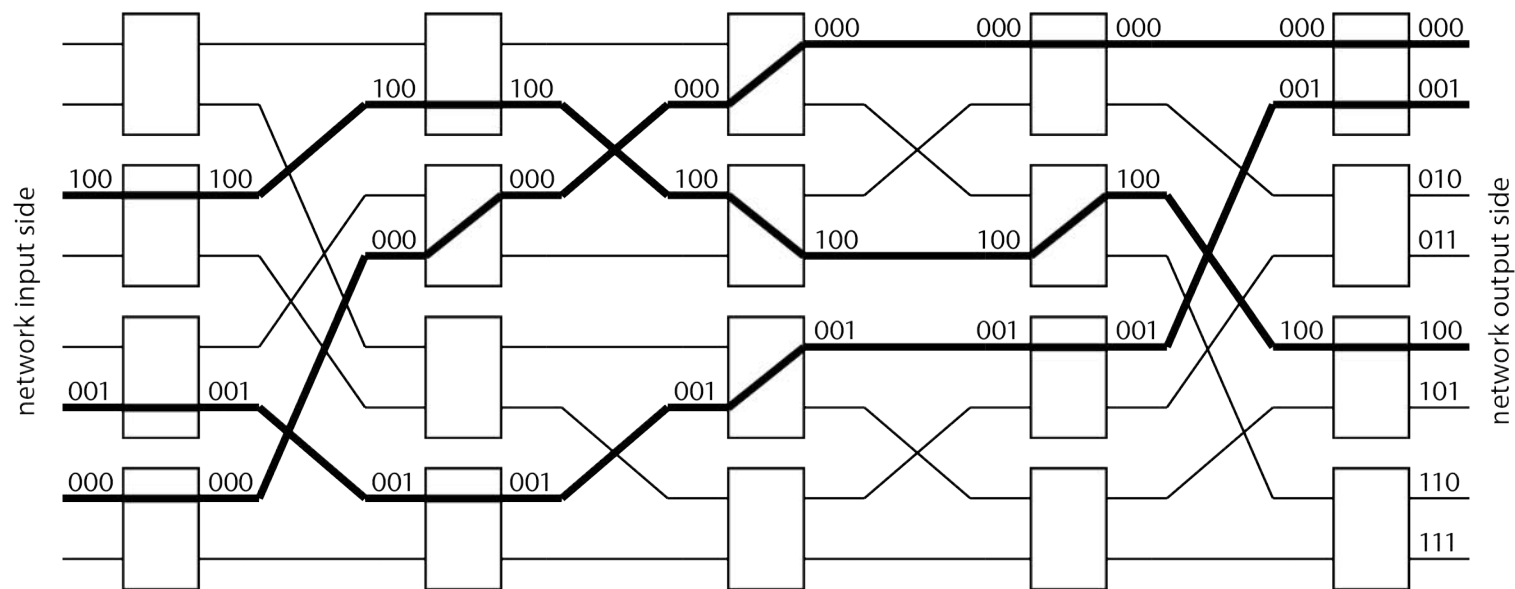


Bild: Routing im Beneš-Netz der Größe 8x8

## Interprozessor-Kommunikation

### ■ Blockierungsfreie Netze

- Komplexes Routing aufgrund von  $N/2$  alternativen Pfaden
  - bisher bekanntes Verfahren: Schleifen-Routing mit  $N*8$  langen Routingsequenzen pro Verbindung
    - ➔ 4096 Einzelschritte (bzw. Takte) bei 512 Prozessoren erforderlich

➔ Skalierbarkeit erfordert parallele Routing-Algorithmen

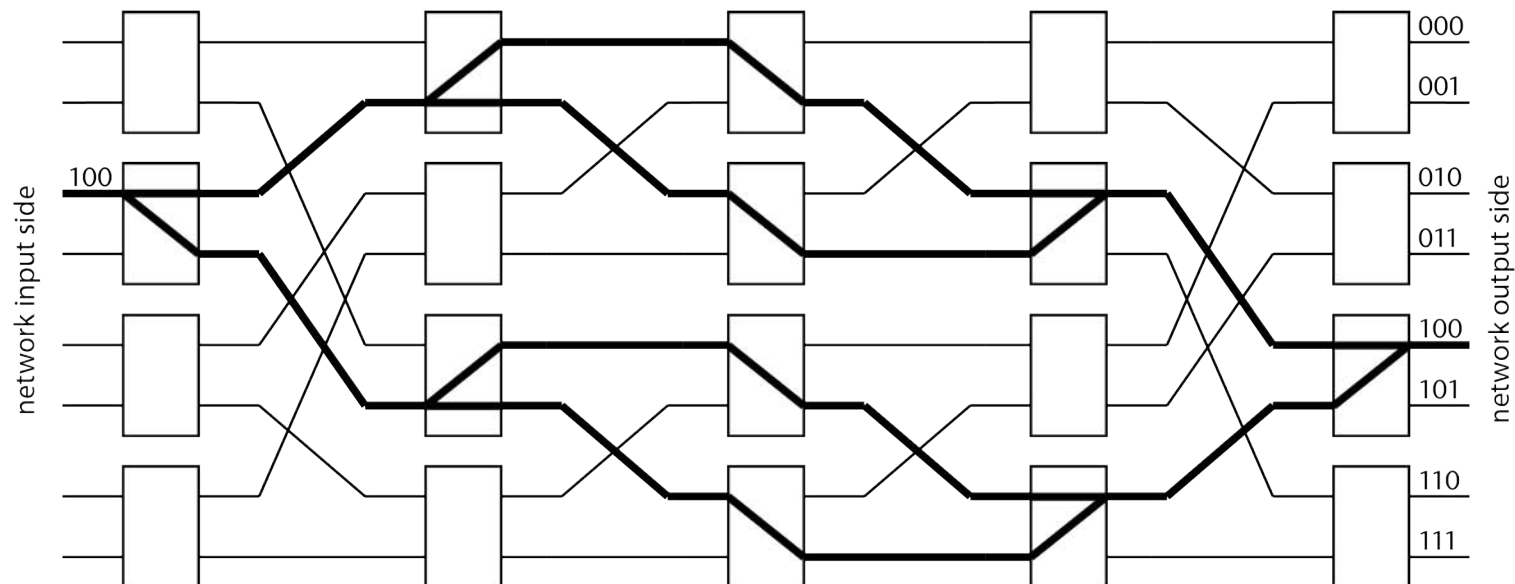


Bild: Alternative Pfade im Beneš-Netz der Größe 8x8

## Interprozessor-Kommunikation

- Blockierungsfreie Netze

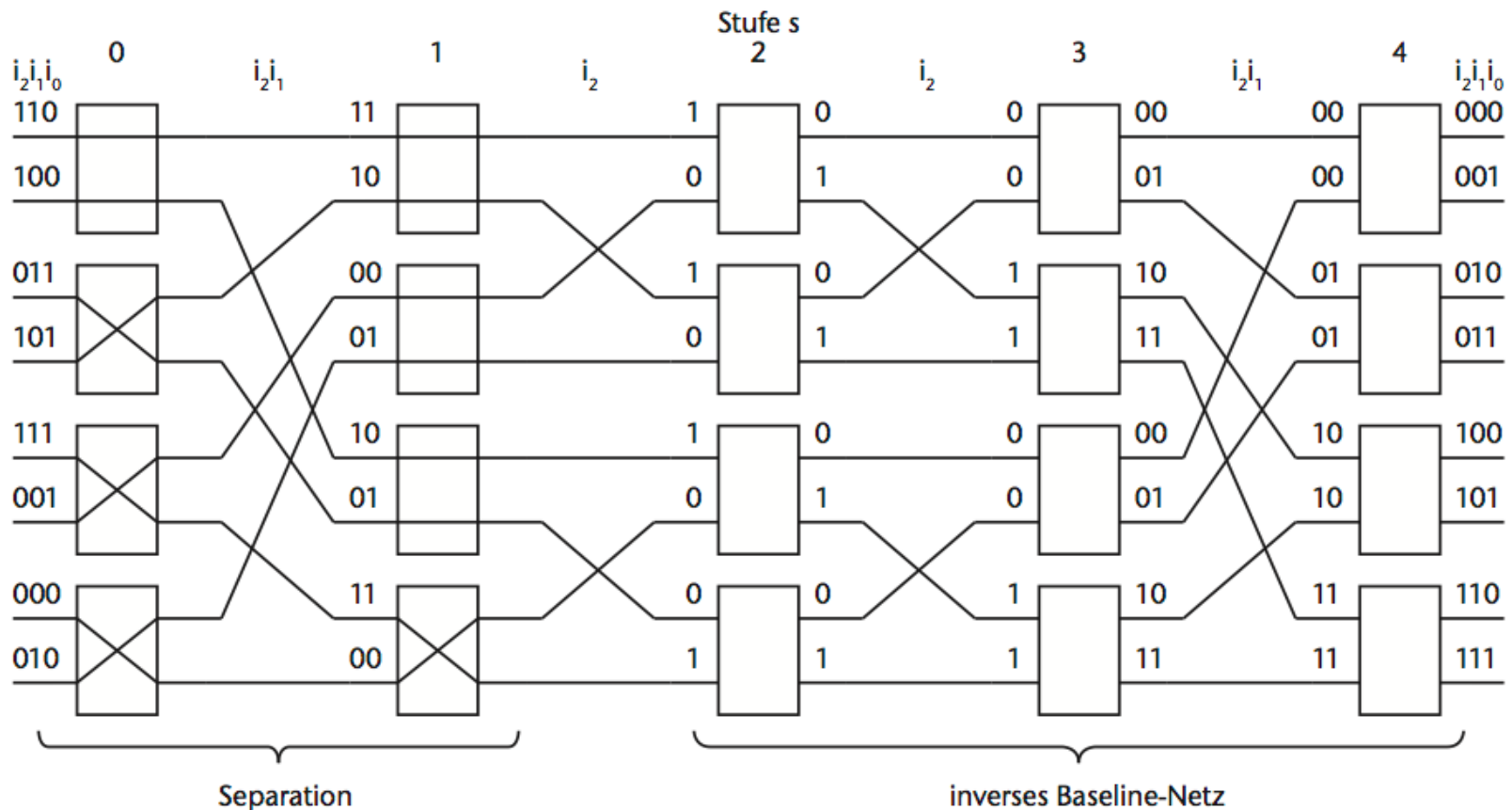


Bild: Echtzeitfähiges Routing im Beneš-Netz

## Interprozessor-Kommunikation

- Blockierungsfreie Netze

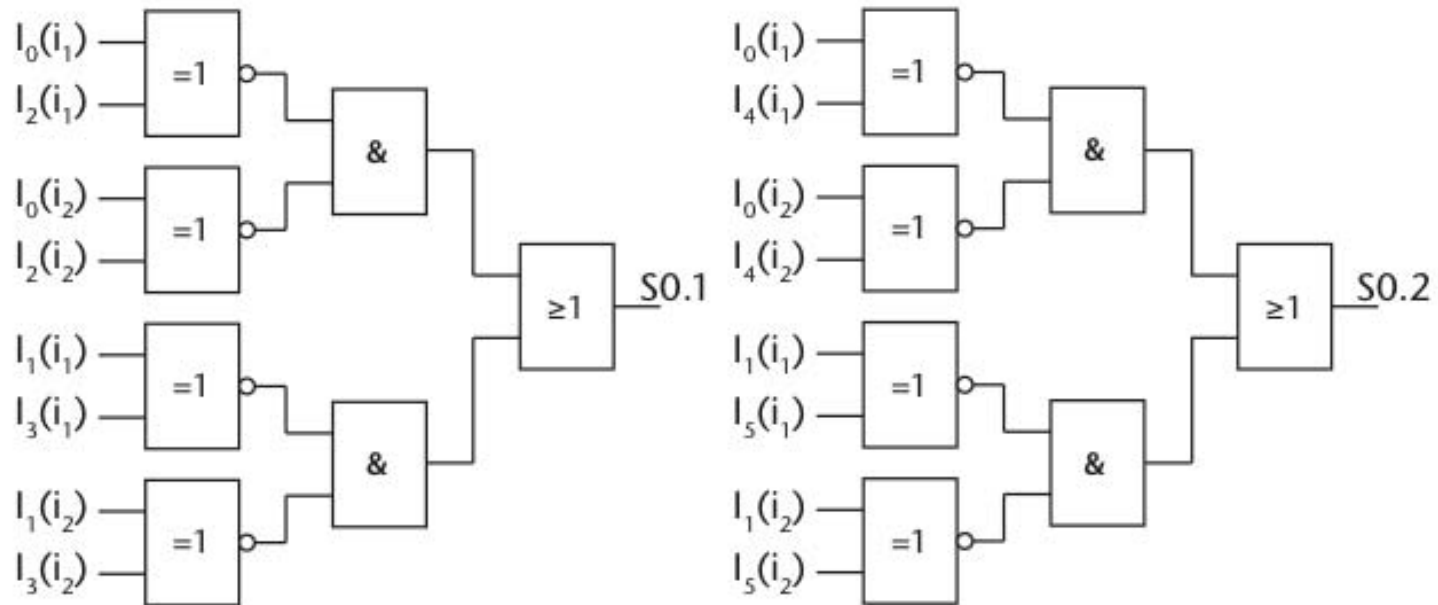
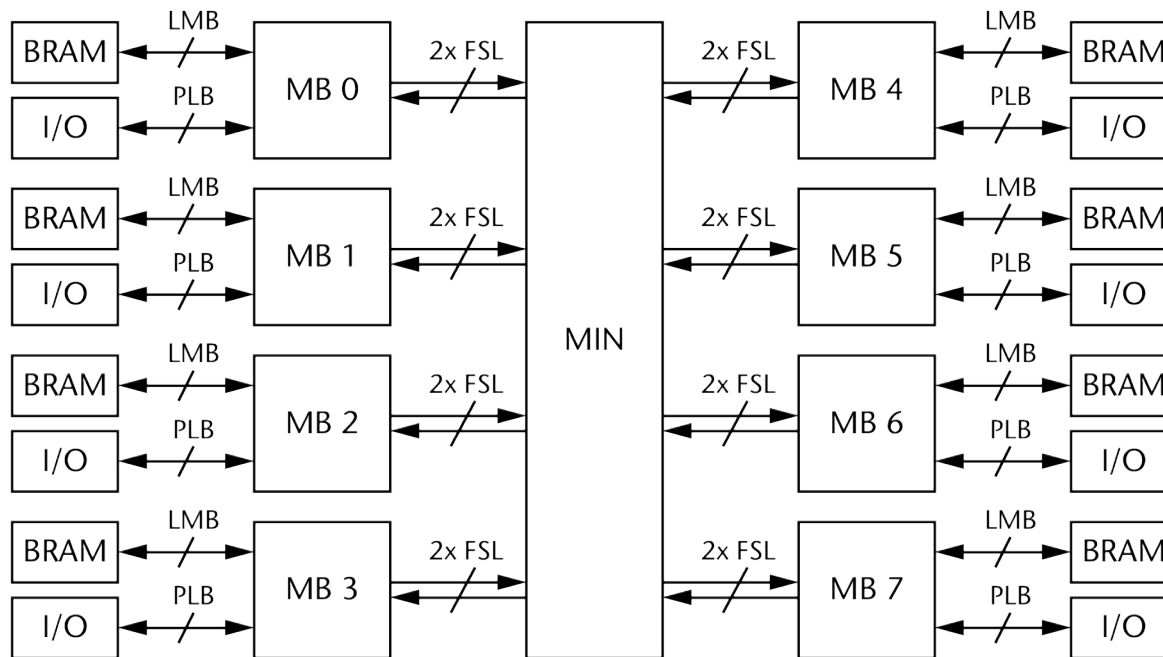


Bild: Schaltnetz zur Bestimmung der Schalterstellungen der ersten Stufe

# Implementierung im Xilinx FPGA

## Multi-Prozessor System-on-Chip (MPSoC)



BRAM: Local Memory (Block-RAM)

FSL: Fast Simplex Link

I/O: Input / Output (digital, RS232, etc.)

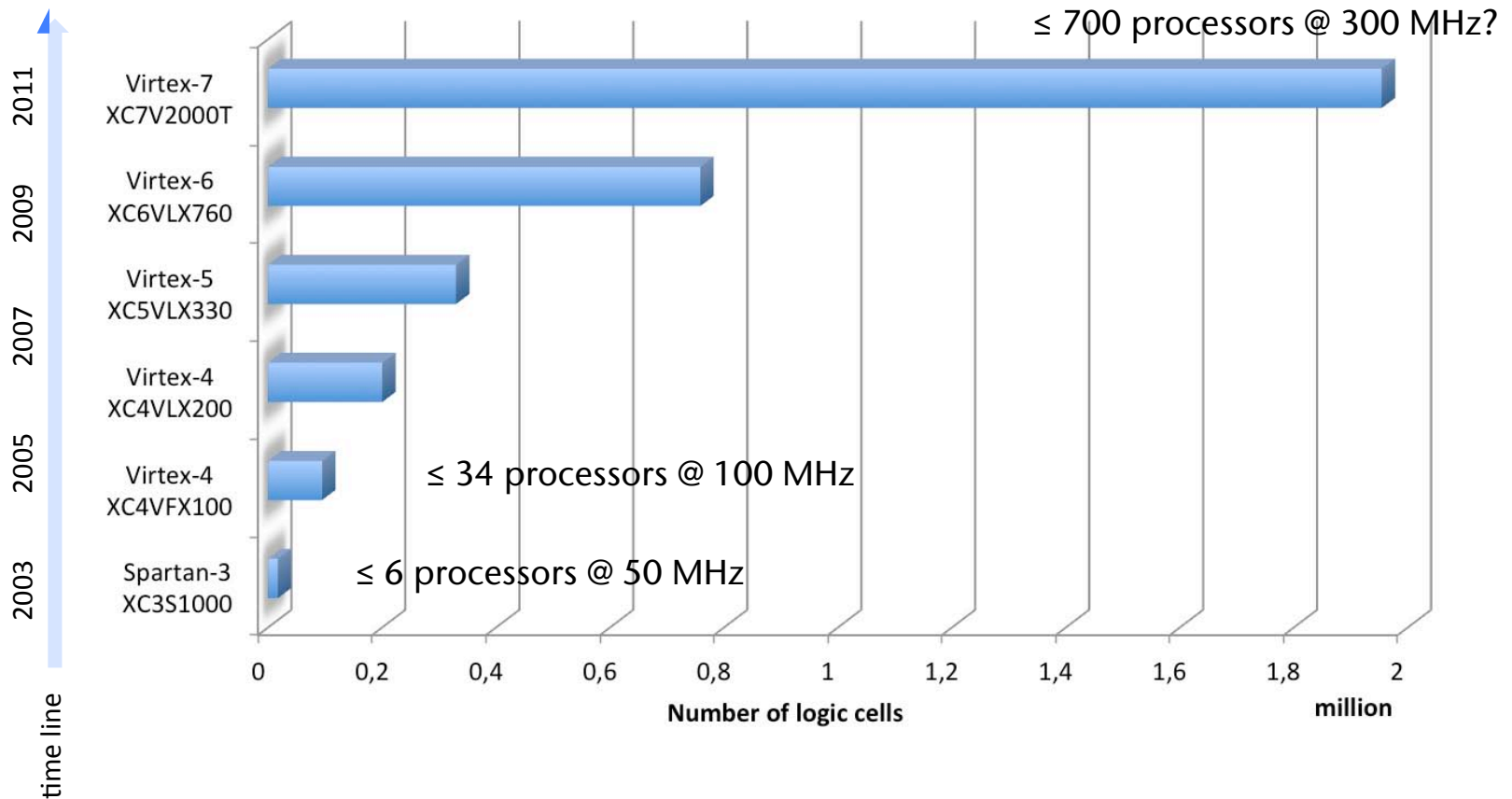
LMB: Local Memory Bus

MB: Softprozessor (MicroBlaze)

MIN: Multistage Interconnection Network

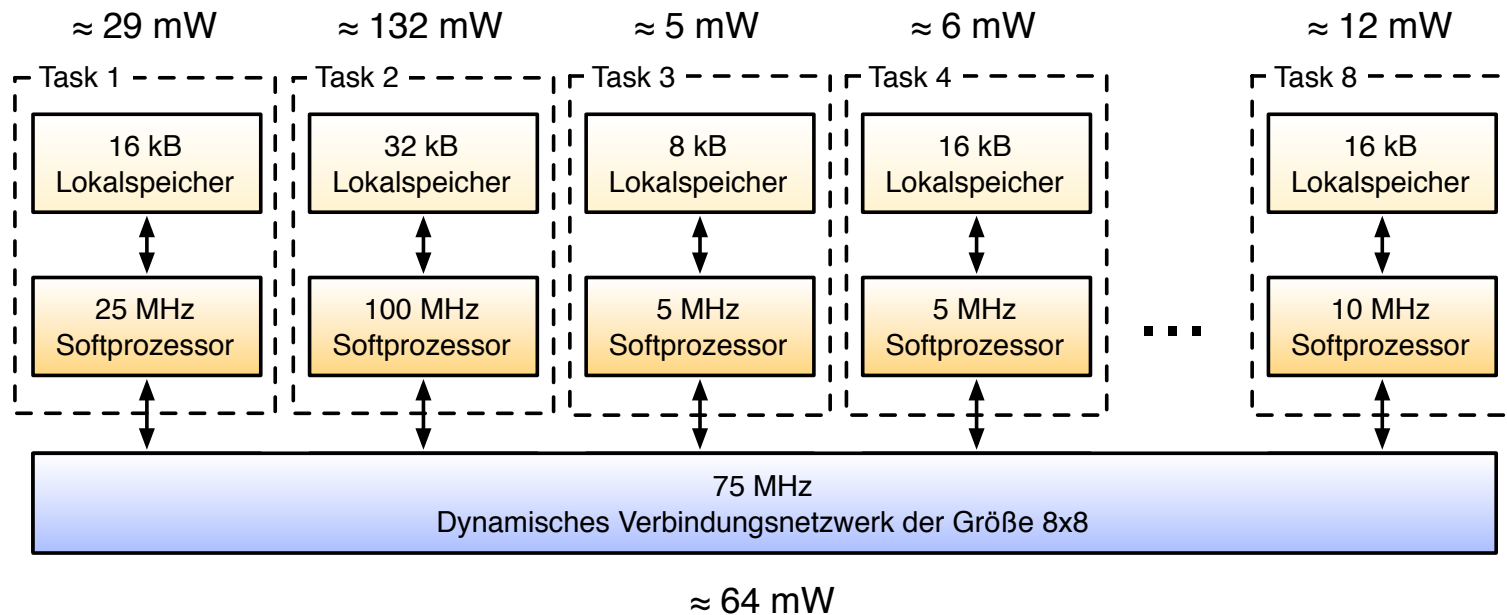
PLB: Processor Local Bus

## Beyond Moore's Law



## Optimierung des Leistungsbedarfs

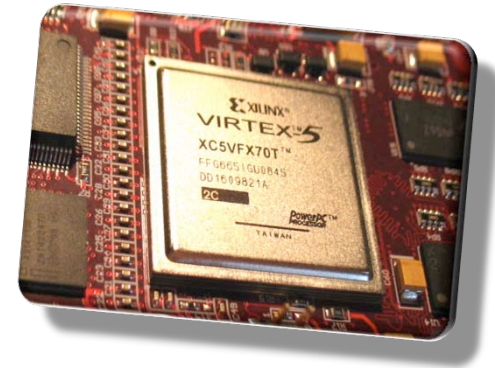
- Dynamische Verlustleistung in einem Multiprozessorsystem





## Zusammenfassung

- Space Sharing
  - Statische Abbildung von Tasks auf je einen Prozessor
  - Many-Core Architektur
  - Umsetzung in FPGA-Hardware
  - Software-First-Design
- Interprozessorkommunikation
  - Blockierungsfreies mehrstufiges Netz
  - Direkte Kommunikation zwischen Sender und Empfänger
  - Verzögerungsfreies Routing



## Vielen Dank für Ihre Aufmerksamkeit!

- Kontakt

- Stefan Aust  
TU Clausthal  
Institut für Informatik  
Arnold-Sommerfeldstr. 1  
38678 Clausthal-Zellerfeld

Tel.: +49 53 23 72-71 72

E-Mail: [stefan.aust@tu-clausthal.de](mailto:stefan.aust@tu-clausthal.de)