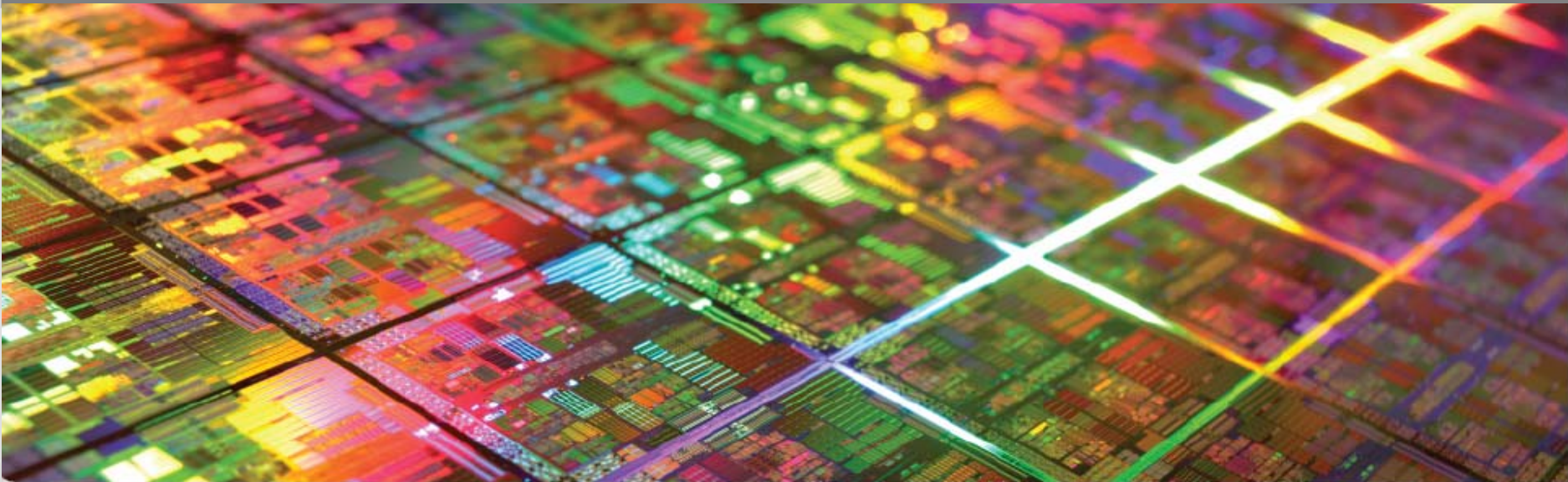


# Rechnerstrukturen

Vorlesung im Sommersemester 2010

Prof. Dr. Wolfgang Karl

Fakultät für Informatik – Lehrstuhl für Rechnerarchitektur und Parallelverarbeitung

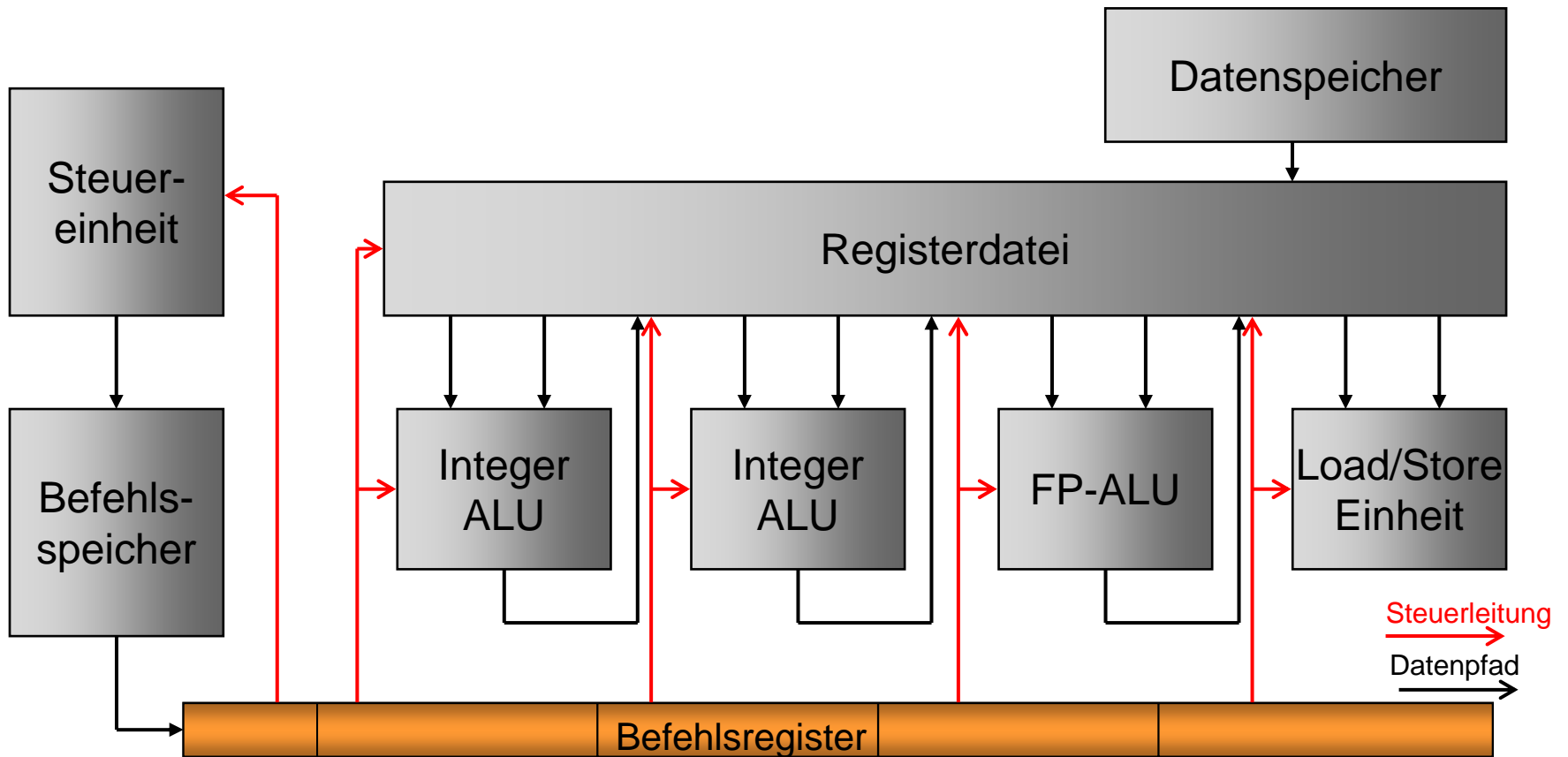


# Vorlesung Rechnerstrukturen

- **Kapitel 2: Parallelismus auf Maschinenbefehlsebene**
- 2.1 Nebenläufigkeit
- 2.1.1 VLIW-Architektur

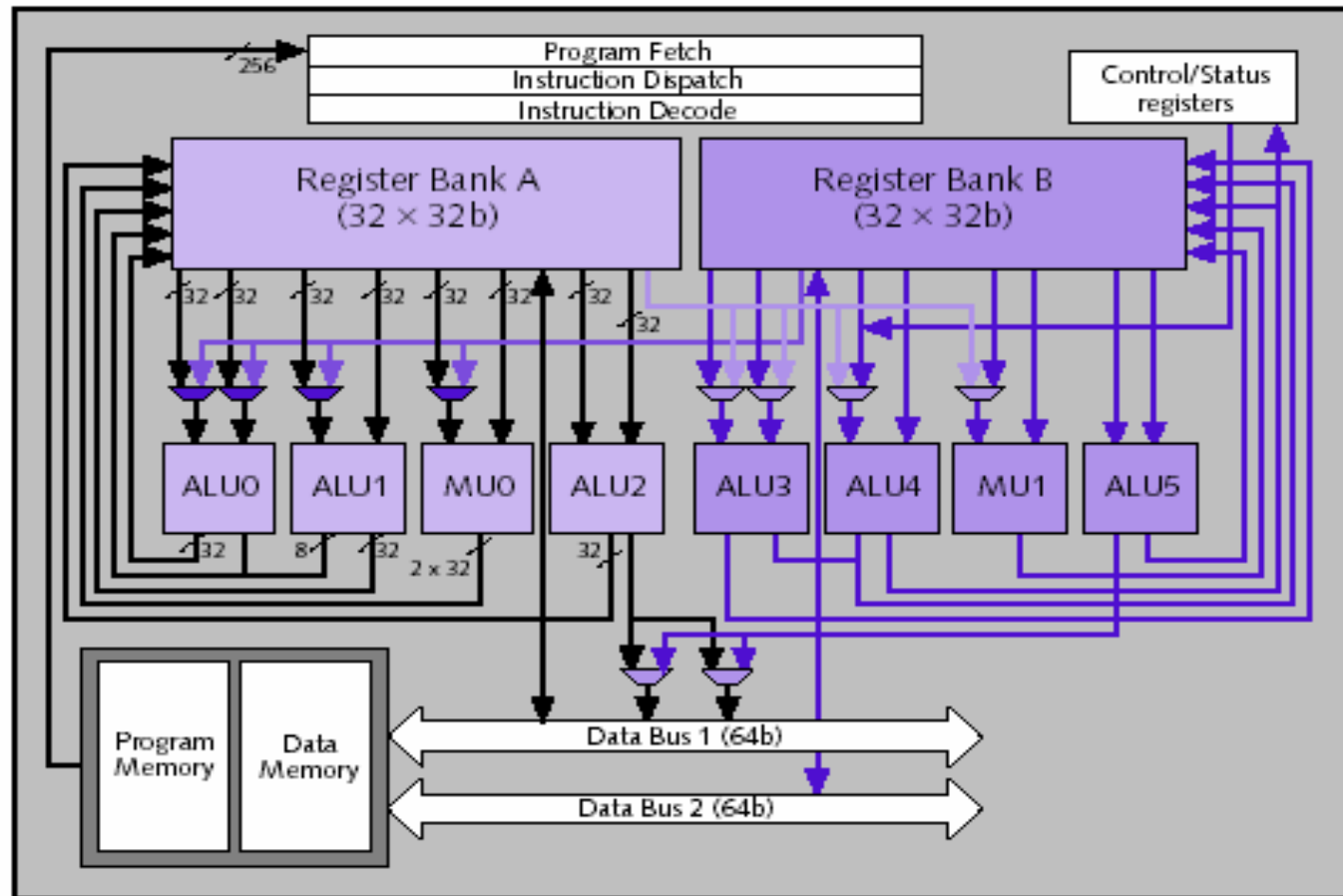
# VLIW-Architektur

- Grundprinzip VLIW
- Prinzipieller Aufbau



# VLIW-Architektur

## ■ Fallstudie TI TMS320C6400



# VLIW-Architektur

## ■ Fallstudie TI TMS320C6400

### ■ Architekturmerkmale

#### ■ 2 x 4 Funktionseinheiten (A und B Seite)

- Jede Seite enthält 16 Register
- Programmierer sieht 32 Register A0 – A15, B0 – B15
- Ausgewählte Register für Boole'sche Ergebnisse von bedingten Befehlen
- 9 32 Bit Lese- und 6 32 Bit Schreibports
- Jeweils ein Crossover-Pfad: Beschränkter wechselseitiger Zugriff

#### ■ Jede Seite enthält

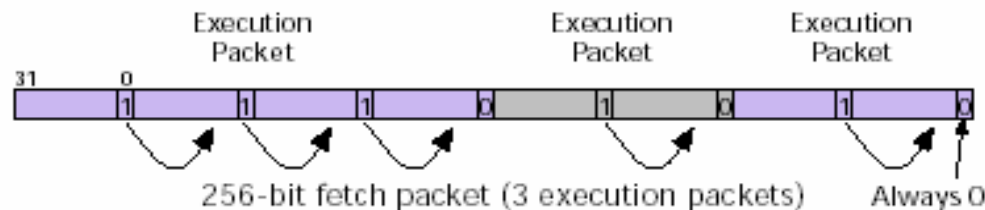
- Eine 40 Bit Integer-ALU (L Unit)
  - Arithmetische und logische Operationen, Vergleiche, Normalisierung, Bit-Count,
- 16 Bit Multiplizierer
  - 16 x 16 → 32 Bit Multiplikation
- 40 Bit Schiebereinheit
- 32 Bit Addierer
  - Adressgenerierung

# VLIW-Architektur

## ■ Fallstudie TI TMS320C6400

### ■ Operationsprinzip

- Holen von 8 32 Bit Befehlen über 256 Bit Befehlsbus (Fetch Packet)
  - Geholte Befehle müssen nicht unbedingt gleichzeitig ausgeführt werden
  - Ein Befehl in einem Fetch Packet ist nicht auf eine Ausführungseinheit beschränkt, jeder Befehl enthält Kodierung, mit der spezifiziert wird, auf welche Einheit der Befehl ausgeführt werden soll
  - Befehle sind nicht positionsabhängig
  - Programmierer / Compiler bestimmt Bindung
- Bis zu 8 Befehle können gleichzeitig ausgeführt werden: Gruppierung von gleichzeitig ausführbaren Befehlen (Execution Packets)
  - Wird im niedrigstwertigen Bit eines Feldes gekennzeichnet: alle nachfolgenden Befehle werden gleichzeitig ausgeführt



# VLIW-Prinzip

## ■ Literatur:

- Hennessy/Patterson: Computer Architecture A Quantative Approach. Kap. 4.1-4.4, 4.8

# Vorlesung Rechnerstrukturen

- **Kapitel 2: Parallelismus auf Maschinenbefehlsebene**
- 2.1 Nebenläufigkeit
- 2.1.3 Multithreading



# Mehrfädigkeit (Multithreading)

- **Grundsätzliche Aufgabe beim Prozessorentwurf:**
  - Reduzierung der Untätigkeits- oder Latenzzeiten
    - Entstehen bei Speicherzugriffen, insbesondere bei Cache-Fehlzugriffen
  - Parallele Ausführung mehrerer Kontrollfäden

# Mehrfädigkeit (Multithreading)

- **Mehrfädige Prozessortechnik**
- Gegeben mehrere ausführbereite Kontrollfäden, Threads
- Ziel: Parallele Ausführung mehrerer Kontrollfäden
- Prinzip
  - Mehrere Kontrollfäden sind geladen
  - Kontext muss für jeden Thread gesichert werden können
  - Mehrere getrennte Registersätze auf Prozessorchip
  - Mehrere Befehlszähler
  - Getrennte Seitentabellen
  - Threadwechsel, wenn gewartet werden muss

# Mehrfädigkeit (Multithreading)

- **Mehrfädige Prozessortechnik**
- Cycle-by-cycle Interleaving (feingranulares Multithreading)
  - Eine Anzahl von Kontrollfäden ist geladen.
  - Der Prozessor wählt in jedem Takt einen der ausführungsbereiten Kontrollfäden aus.
  - Der nächste Befehle in der Befehlsreihenfolge des ausgewählten Kontrollfadens wird zur Ausführung ausgewählt.
- Beispiele
  - Multiprozessorsysteme HEP, Tera
- **Nachteil:**
  - Die Verarbeitung eines Threads kann erheblich verlangsamt werden, wenn er ohne Wartezeiten ausgeführt werden kann

# Mehrfädigkeit (Multithreading)

## ■ Mehrfädige Prozessortechnik

### ■ Block Interleaving

- Befehle eines Kontrollfadens werden so lange ausgeführt, bis eine Instruktion mit einer langen Latenzzeit ausgeführt wird. Dann wird zu einem anderen ausführbaren Kontrollfaden gewechselt.

### ■ Vorteil:

- Die Bearbeitung eines Threads wird nicht verlangsamt, da beim Warten ausführungsbereiter Thread gestartet wird

### ■ Nachteil:

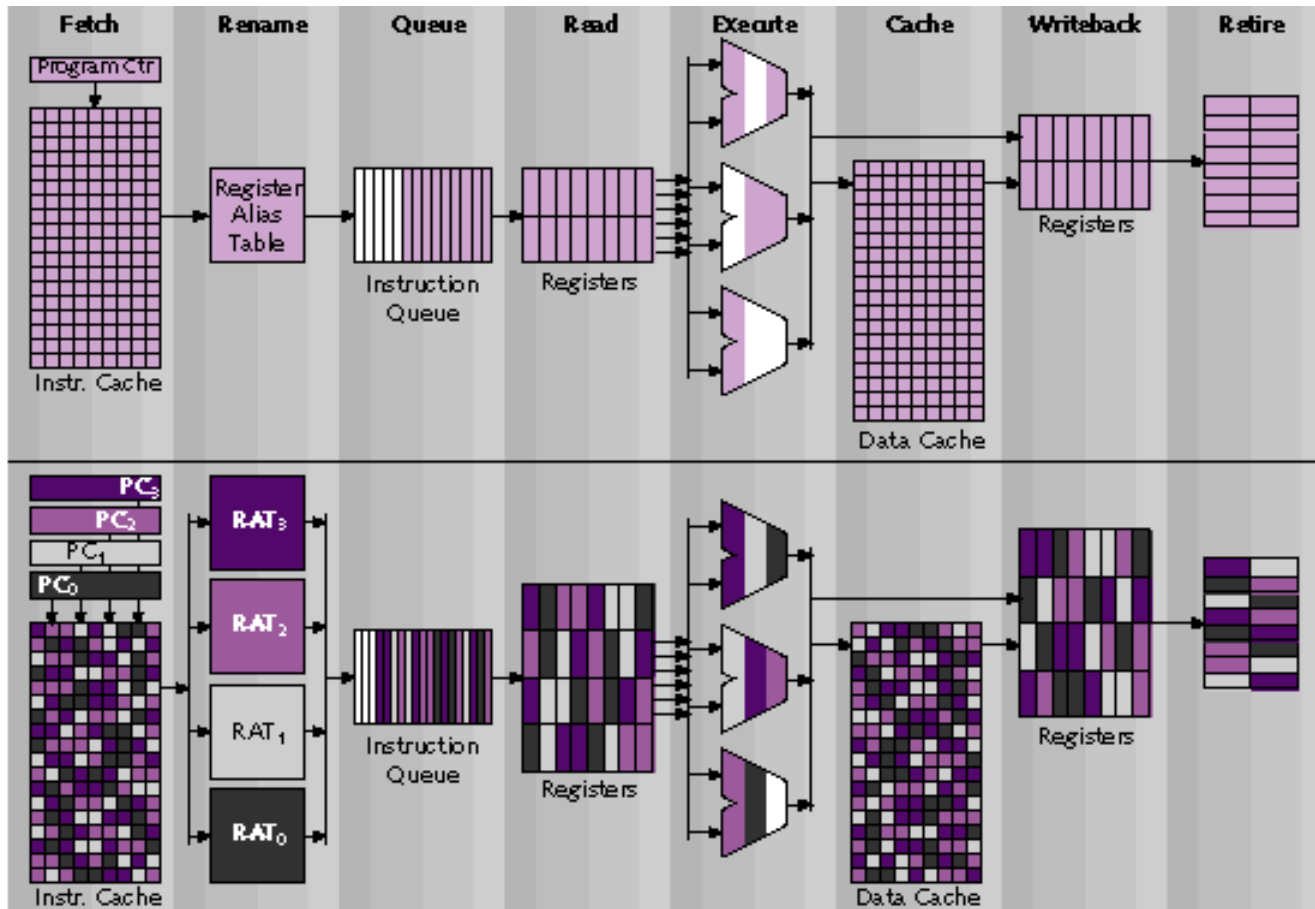
- Bei Thread-Wechsel Leeren und Neustarten der Pipeline,
- Nur bei langen Wartezeiten sinnvoll

# Mehrfädigkeit (Multithreading)

- **Mehrfädige Prozessortechnik**
- Simultaneous Multithreading
  - Mehrfach superskalärer Prozessor
  - Die Ausführungseinheiten werden über eine Zuordnungseinheit aus mehreren Befehlsuffern versorgt.
  - Jeder Befehlsuffler stellt einen anderen Befehlsstrom dar.
  - Jedem Befehlsstrom ist eigener Registersatz zugeordnet.

# Mehrfädigkeit (Multithreading)

- Mehrfädige Prozessortechnik
- Simultaneous Multithreading



# Mehrfädigkeit (Multithreading)

- **Mehrfädige Prozessortechnik**
- Simultaneous Multithreading: Diskussion
  - Abwägen zwischen Geschwindigkeit eines Threads und dem Durchsatz vieler Threads
  - Ein bevorzugter Thread
  - Allerdings kann dies auf Kosten des Durchsatzes gehen, da Befehle anderer Threads möglicherweise nicht bereit stehen
  - Mischen vieler Threads:
  - Geht möglicherweise zu Lasten der Leistung der einzelnen Threads

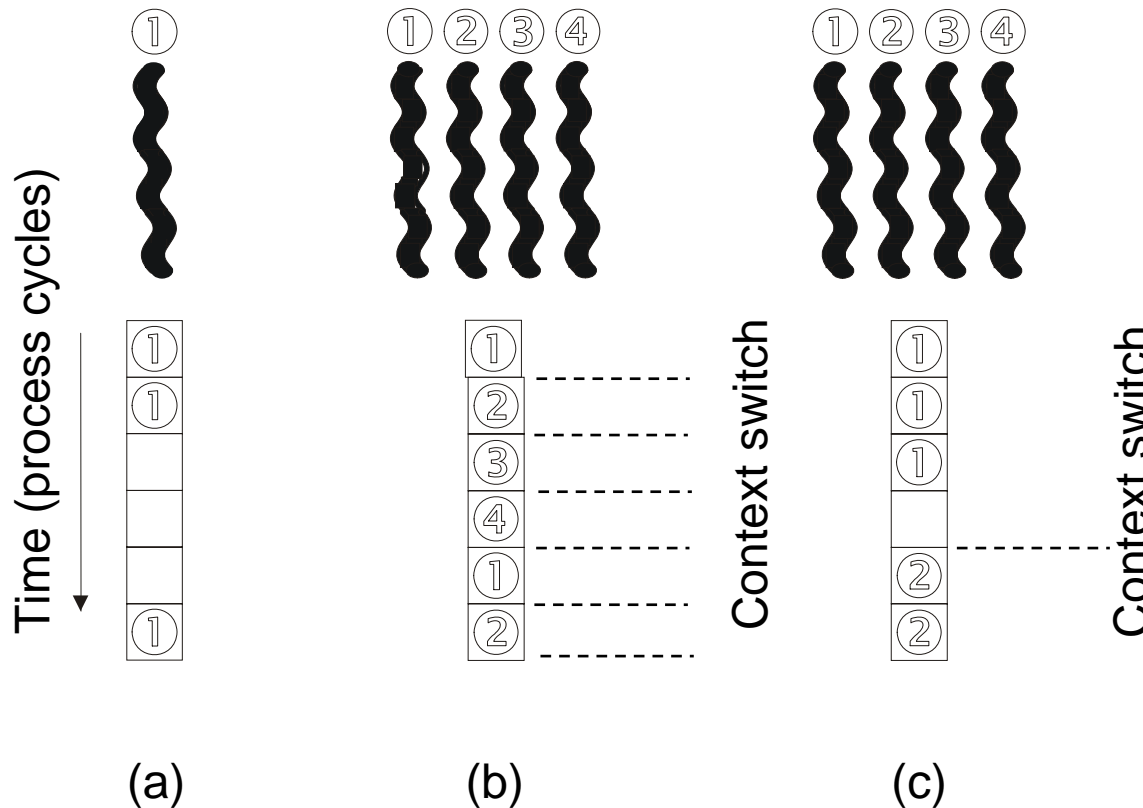
# Mehrfädigkeit (Multithreading)

- **Mehrfädige Prozessortechnik**
- Simultaneous Multithreading: Beispiele
  - Compaq Alpha 21464 (EV8), ursprünglich angekündigt für 2002/2003, Entwicklung aber eingestellt! Entwicklergruppe jetzt bei Intel
  - Intel P4: Hyperthreading
  - Sun Ultra SPARC IV: Chip Multithreading



# Zusammenfassung

## ■ Vergleich von Prozessortechniken



(a): single-threaded scalar

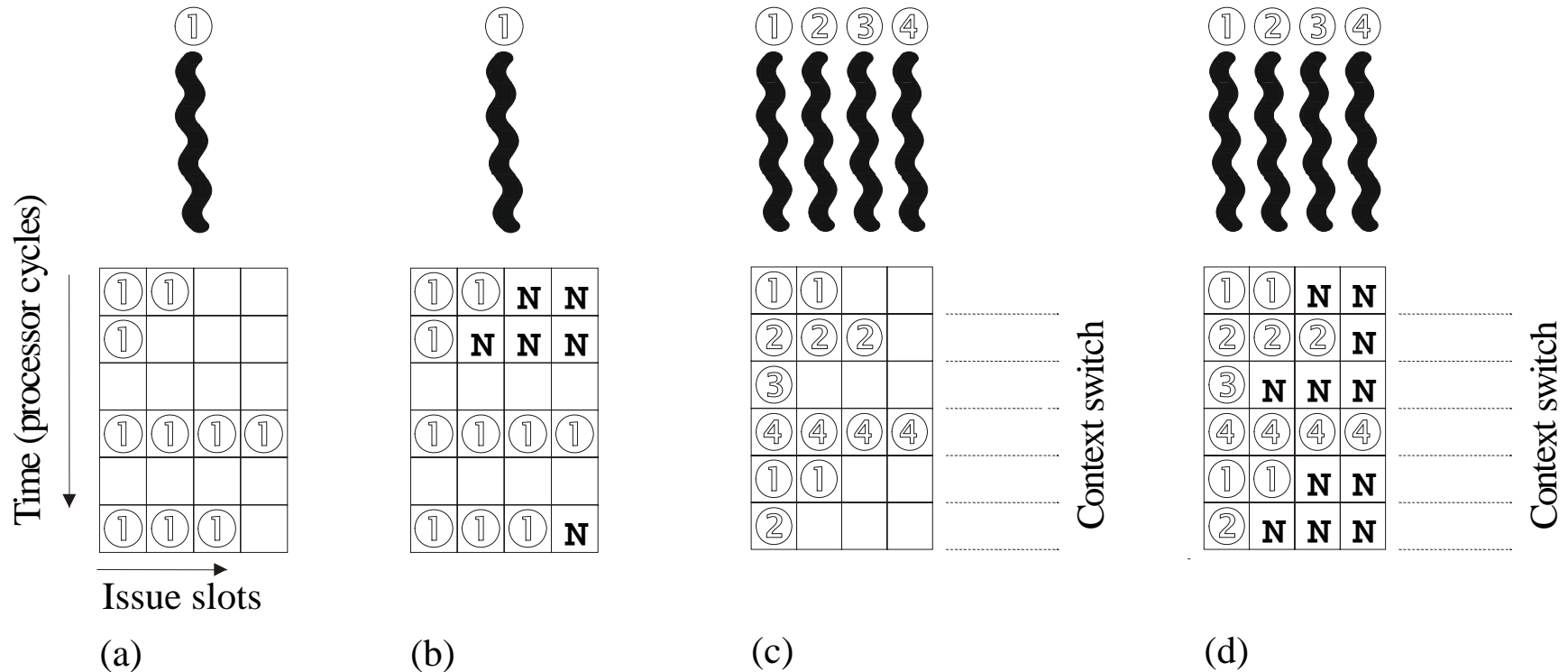
(b) cycle-by-cycle interleaving multithreaded scalar

(c) block interleaving multithreaded scalar

(siehe Brinkschulte, Ungerer: Mikrocontroller und Mikroprozessoren: Kap. 10.4.3)

# Zusammenfassung

## ■ Vergleich von Prozessortechniken



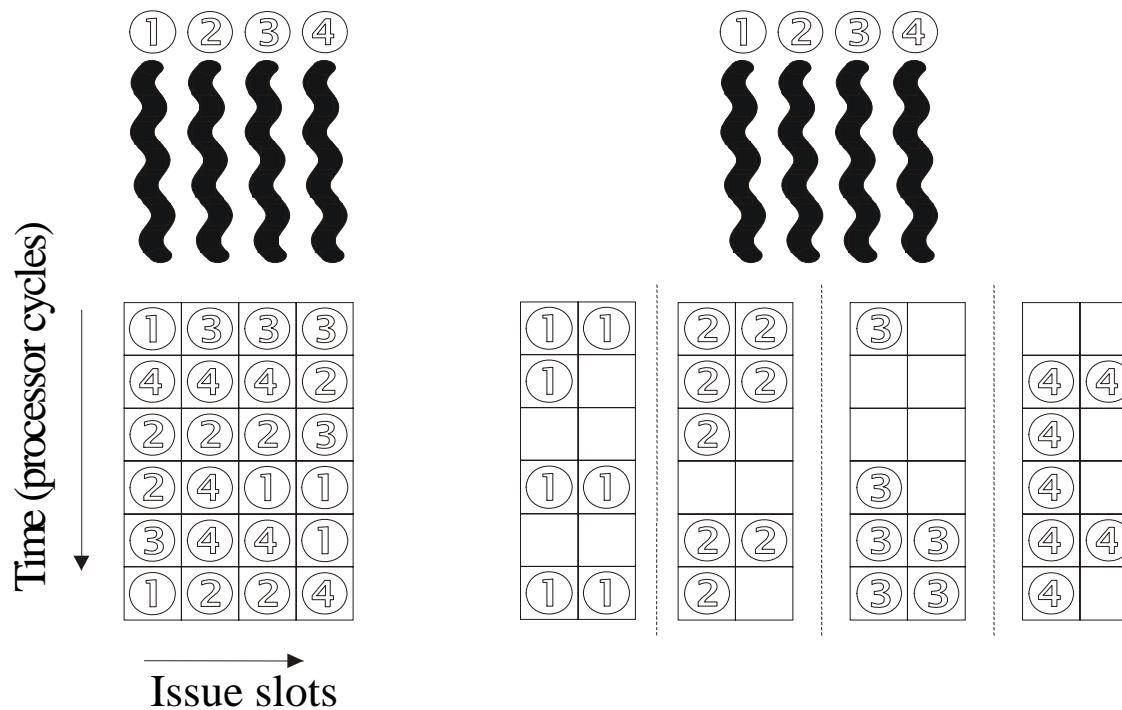
(a) Superskalare Technik  
(b) VLIW-Technik

(c) Cycle-by-cycle Interleaving  
(d) Cycle-by-cycle VLIW

(siehe Brinkschulte, Ungerer: Mikrocontroller und Mikroprozessoren: Kap. 10.4.3)

# Zusammenfassung

## ■ Vergleich von Prozessortechniken



(a)

(b)

(a) Simultaneous Multithreading

(b) Single Chip Multiprocessing

(siehe Brinkschulte, Ungerer: Mikrocontroller und Mikroprozessoren: Kap. 10.4.3)

# Multithreading

## ■ Literatur:

- Brinkschulte, U.; Ungerer, T.: Microcontroller und Mikroprozessoren. Springer, Heidelberg, 2002: Kap.: 10.4.3

# Vorlesung Rechnerstrukturen

- **Kapitel 3: Multiprozessoren – Parallelismus auf Prozess-/Blockebene**
- 3.1 Motivation

# Multiprozessorsysteme

- **Überblick**
- Allgemeine Grundlagen, parallele Programmierung, Verbindungsstrukturen, Leistungsfähigkeit
- Speichergekoppelte Multiprozessoren: SMP und DSM, Cache-Kohärenz und Speicherkonsistenz, Rechnerbeispiele
- Nachrichtengekoppelte Multiprozessoren, Beispielrechner

# Multiprozessorsysteme

## ■ Einordnung:

- Klassifikation nach Flynn: MIMD-Rechner

## ■ Warum Multiprozessorsysteme?

- Hohe Anforderungen von Anwendungen an die Rechenleistung
- Technisch-wissenschaftlicher Bereich
  - Rechnergestützte Simulation
  - Strömungsmechanik
  - Modellierung der globalen klimatischen Veränderungen
  - Evolution von Galaxien
  - Struktur von Materialien
  - .....
- Kommerzieller Bereich
  - Server, Datenbank-Anwendungen, WEB

# Multiprozessorsysteme

- Motivation

- Höchstleistungsrechner:

- TOP500-Liste

- Führt die schnellsten Rechner der Welt auf
    - Erscheint immer im Juni und im November eines Jahres
    - <http://www.top500.org>
    - Beispiel: TOP500 Liste (November 2009)



PRESENTED BY

UNIVERSITY OF MANNHEIM

<http://www.uni-mannheim.de/english/>

ICL-UT INNOVATIVE COMPUTING LABORATORY

<http://icls.uk.ac.uk/>



<http://www.fsl.gov/>

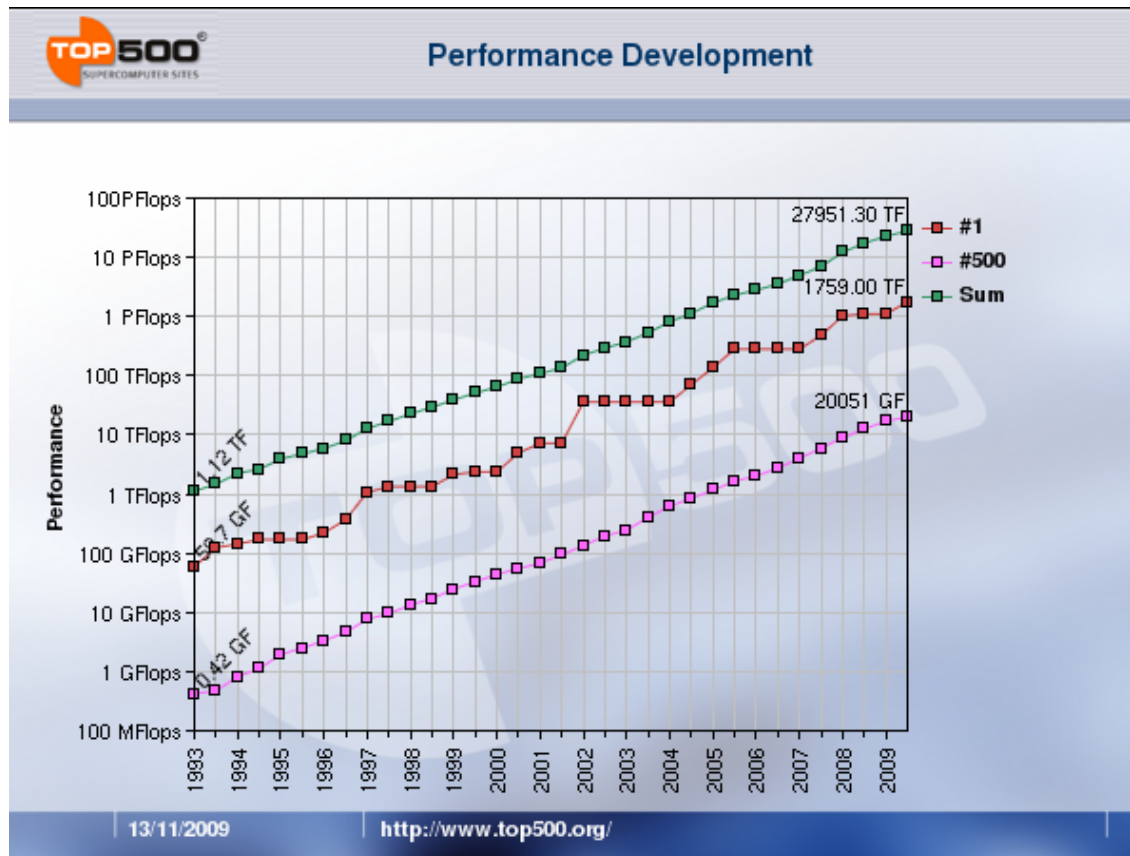
FIND OUT MORE AT <http://www.top500.org/>

	MANUFACTURER/COMPUTER	LOCATION	COUNTRY	CORES	$R_{max}$
1	Jaguar, Cray XT5 6-core 2.6 GHz	DDE / OS / ORNL	USA	224162	1.76
2	Roadrunner, IBM BladeCenter QS22/LS21 Cluster, PowerXCell 3.2 GHz / Opteron 1.8 GHz, Voltaire Iband	DDE / NNSA / LANL	USA	122400	1.04
3	Kraken, Cray XT5 6-core 2.6 GHz	NSF / U of Tennessee	USA	98928	.832
4	Jugene, IBM Blue Gene/P Solution	Forschungszentrum Juelich	Germany	294912	.826
5	Tianhe-1 Cluster, Xeon E5540/E5450 + AMD ATI 4870	National SuperComputer Center in Tianjin/NUDT	China	71680	.563



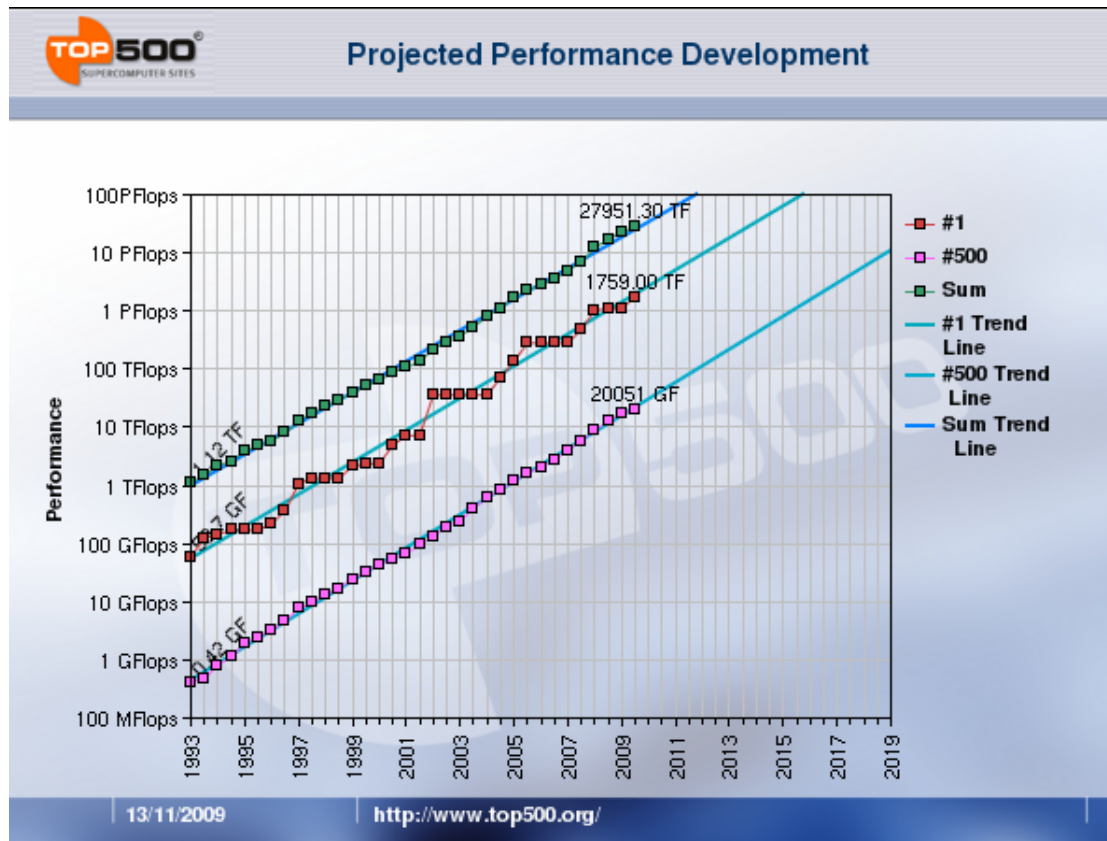
# Multiprozessorsysteme

- Motivation
- Höchstleistungsrechner:
  - TOP500-Liste



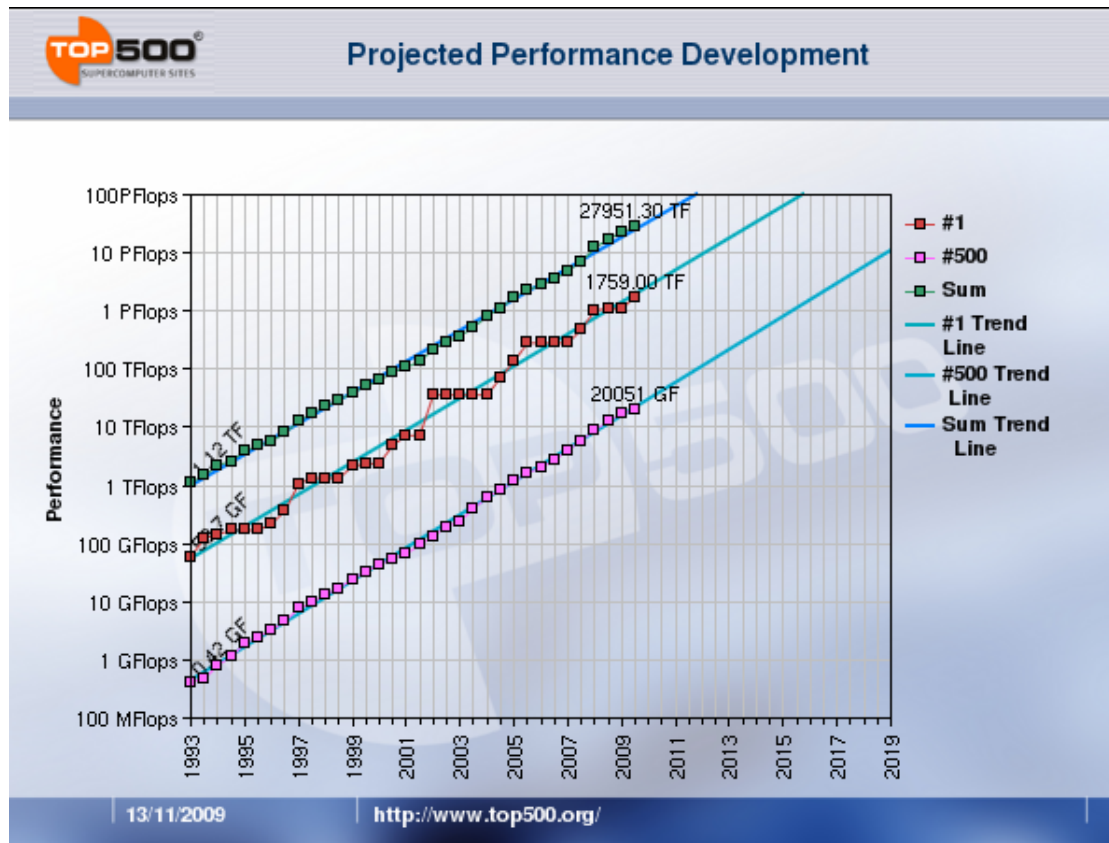
# Multiprozessorsysteme

- Motivation
- Höchstleistungsrechner:
  - TOP500-Liste



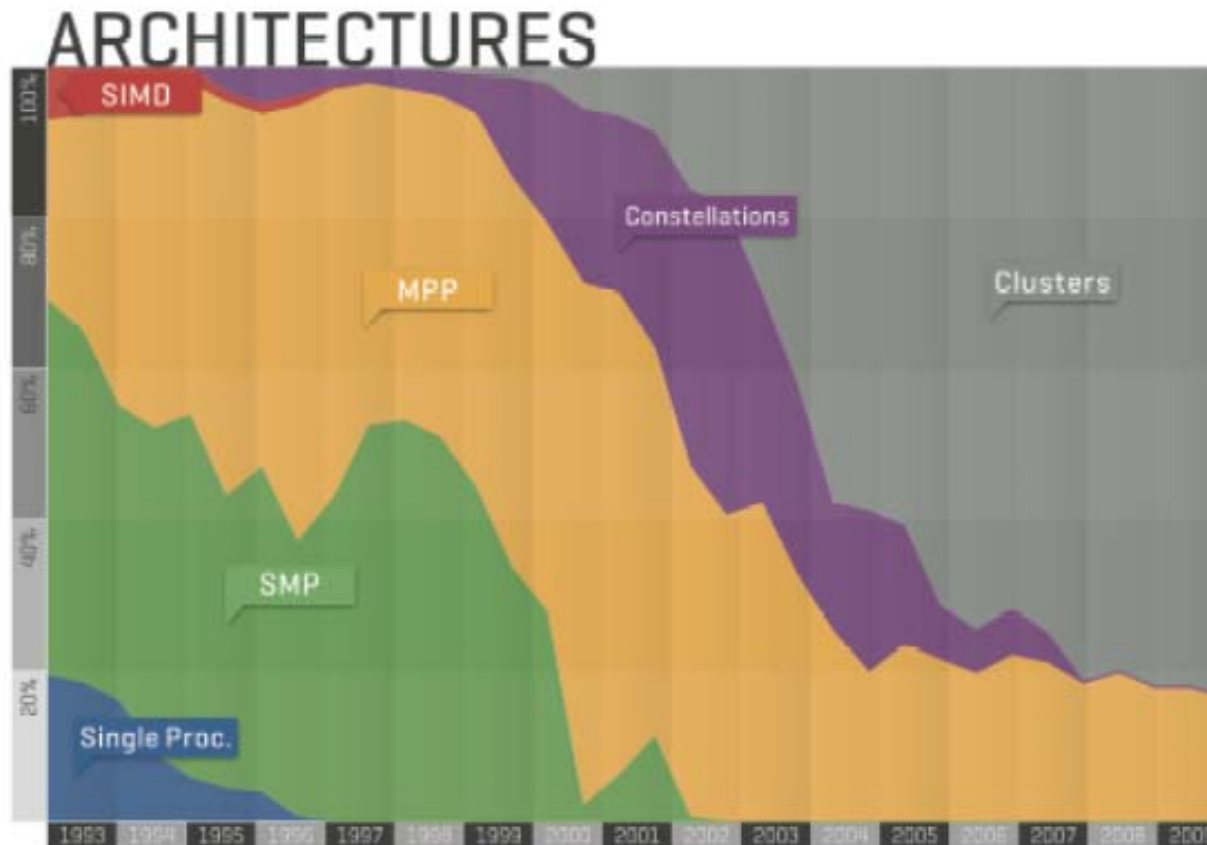
# Multiprozessorsysteme

- Motivation
- Höchstleistungsrechner:
  - TOP500-Liste



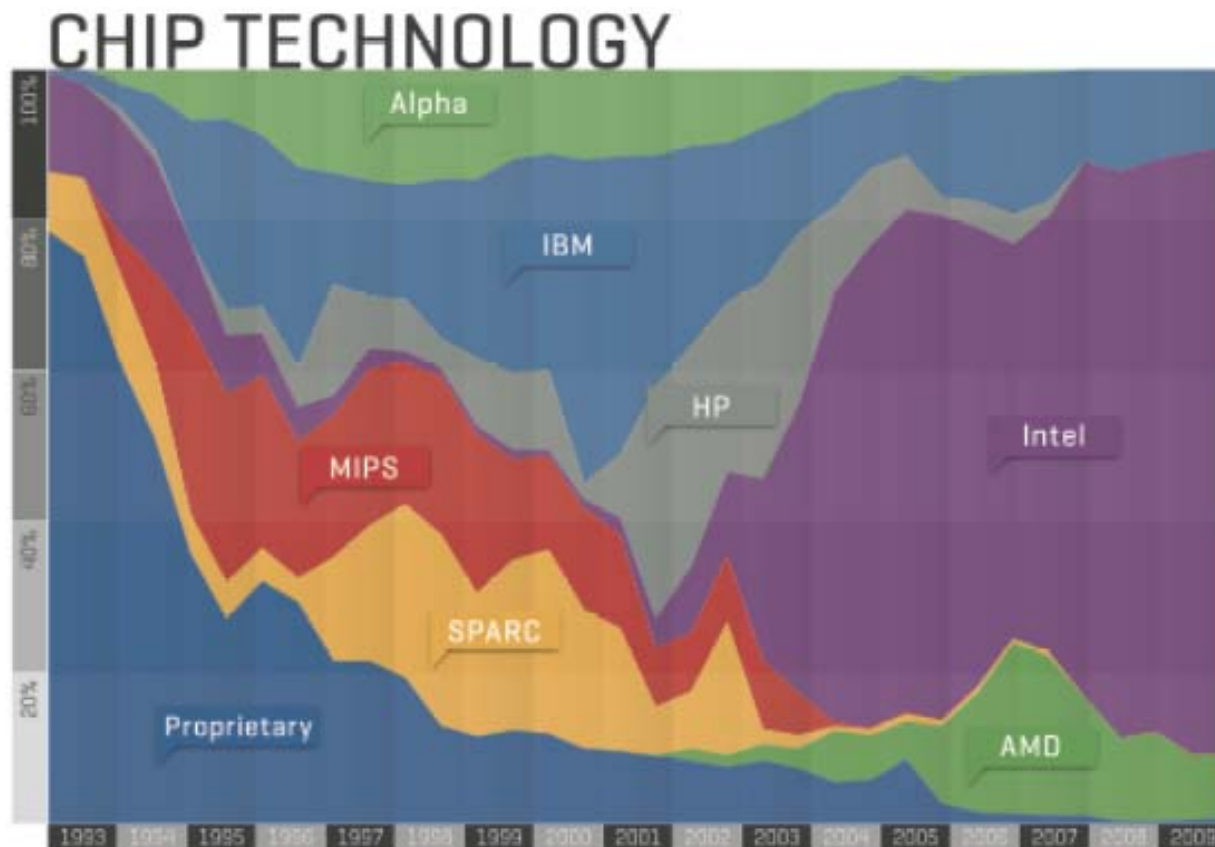
# Multiprozessorsysteme

- Motivation
- Höchstleistungsrechner:
  - TOP500-Liste



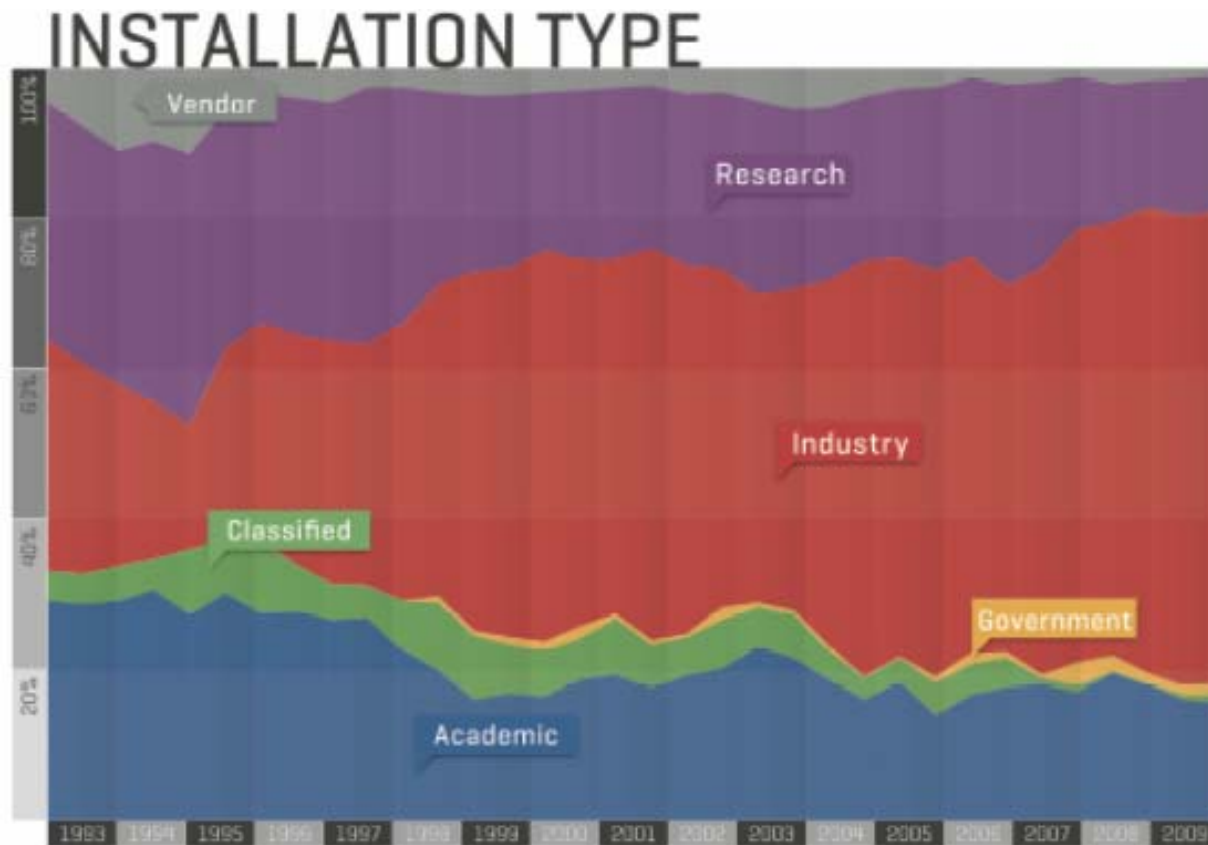
# Multiprozessorsysteme

- Motivation
- Höchstleistungsrechner:
  - TOP500-Liste



# Multiprozessorsysteme

- Motivation
- Höchstleistungsrechner:
  - TOP500-Liste



# Multiprozessorsysteme

- **Motivation**
- Top500 Nr. 1 (November 2009): Cray XT5-HE Opteron Six Core 2.6 GHz

	XT5	XT4	Total
Cabinets	200	84	284
Compute Blades	4,672	1,958	6,630
Quad-core Opteron Processors	37,376	7,832	45,208
Cores	149,504	31,328	180,832
Peak TeraFLOPS	1,375	263	1,639
Nodes	18,688	7,832	26,520
Memory (TB)	300	62	362
Number of disks	13,440	2,774	16,214
Disk Capacity (TB)	10,000	750	10,750
I/O Bandwidth (GB/s)	240	44	284

**Table 1 - Jaguar System Configuration**

Bland A.S., Kendall R.A., Kothe D.B., Rogers J.H., Shipman G.M. *Jaguar: The World's Most Powerful Computer*, <http://www.nccs.gov/computing-resources/jaguar/documentation/2009-cug-meeting/>

# Multiprozessorsysteme

## ■ Motivation

- Top500 Nr. 1 (November 2009): Cray XT5-HE Opteron Six Core 2.6 GHz
  - Massiv-paralleler Multiprozessor mit verteiltem Speicher
  - Knoten: symmetrischer Multiprozessor
    - „Each node has two AMD Opteron model 2356 2.3 GHz quad-core “Barcelona” processors
    - connected to each other through a pair of HyperTransport5 connections.
    - Each of the Opteron processors has 2 MB of level 3 cache shared among the four cores
    - and a DDR2 memory controller connected to a pair of 4 GB DDR2-800 memory modules.
    - The HyperTransport connections between the processors provide a cache coherent shared memory node with eight cores, 16 GB of memory, and 25.6 GB per second of memory bandwidth.
    - The node has a theoretical peak processing performance of 73.6 billion floating point operations per second (GF).

Bland A.S., Kendall R.A., Kothe D.B., Rogers J.H., Shipman G.M. *Jaguar: The World's Most Powerful Computer*, <http://www.nccs.gov/computing-resources/jaguar/documentation/2009-cug-meeting/>



# Multiprozessorsysteme

## ■ Motivation

■ Top500 Nr. 1 (November 2009): Cray XT5-HE Opteron Six Core 2.6 GHz

■ Knoten: symmetrischer Multiprozessor

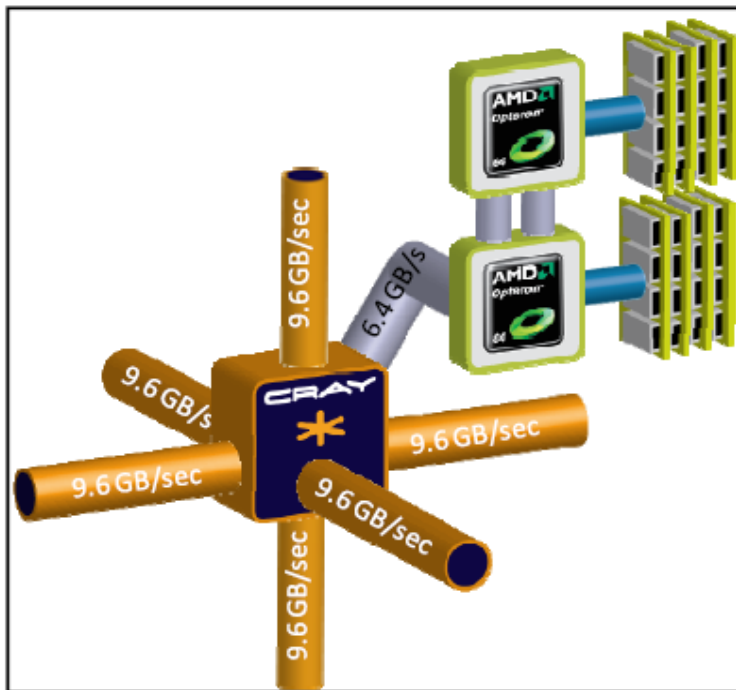


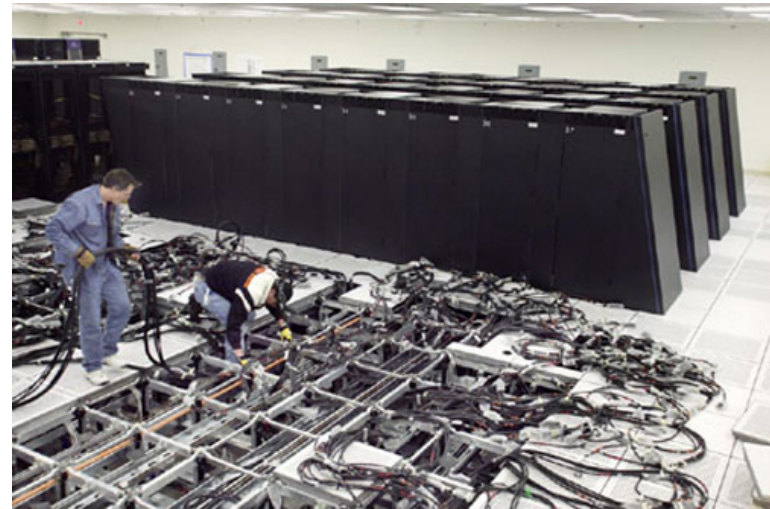
Figure 1 - XT5 Node Configuration

Bland A.S., Kendall R.A., Kothe D.B., Rogers J.H., Shipman G.M. *Jaguar: The World's Most Powerful Computer*, <http://www.nccs.gov/computing-resources/jaguar/documentation/2009-cug-meeting/>

# Multiprozessorsysteme

## ■ Motivation

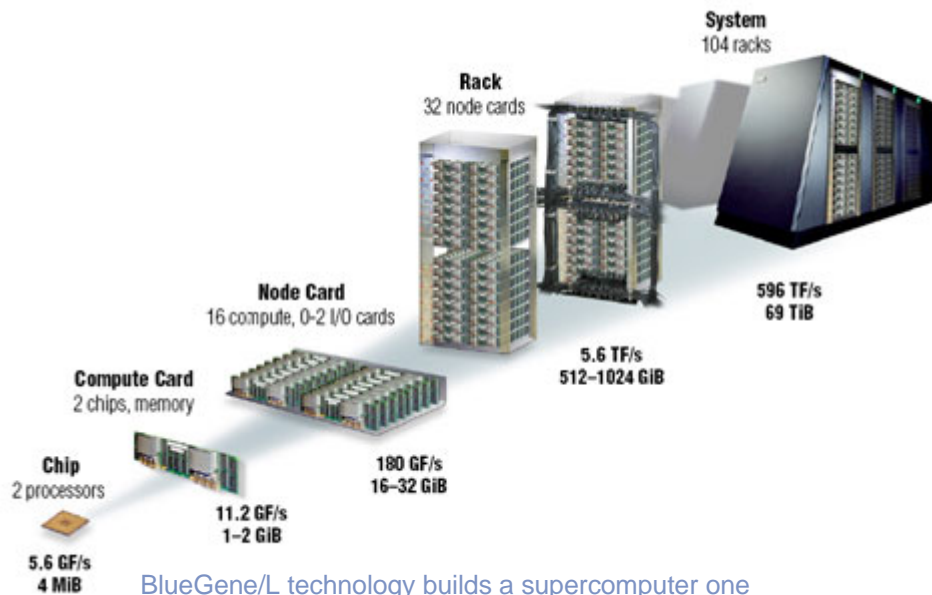
- Top500 Nr. 7 (November 2009): BlueGene/L, Modell: eServer Blue Gene Solution
  - Standort: DOE/NNSA/LLNL
  - Anzahl Prozessoren (PowerPC 440 700 MHz, 2.8 GFlops) : 212992
  - Speicher: 73728 GB
  - Leistung: 478200 GFLOPS (Linpack)
  - Installation: 2007



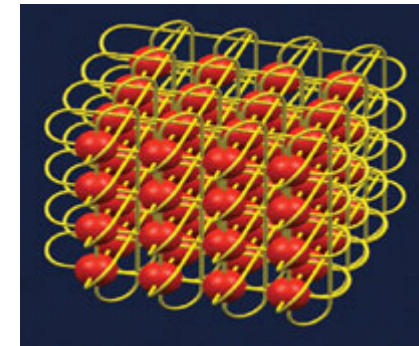
Quelle: [https://asc.llnl.gov/computing\\_resources/bluegenel/photogallery.html](https://asc.llnl.gov/computing_resources/bluegenel/photogallery.html)

# Multiprozessorsysteme

- Motivation
- Top500 Nr. 7 (November 2009): BlueGene/L, Modell: eServer Blue Gene Solution



BlueGene/L technology builds a supercomputer one dual-processor chip at a time. Chips are aggregated into compute cards, which are then assembled into node cards. Each rack holds 2 node cards, and the full machine now comprises 104 racks.

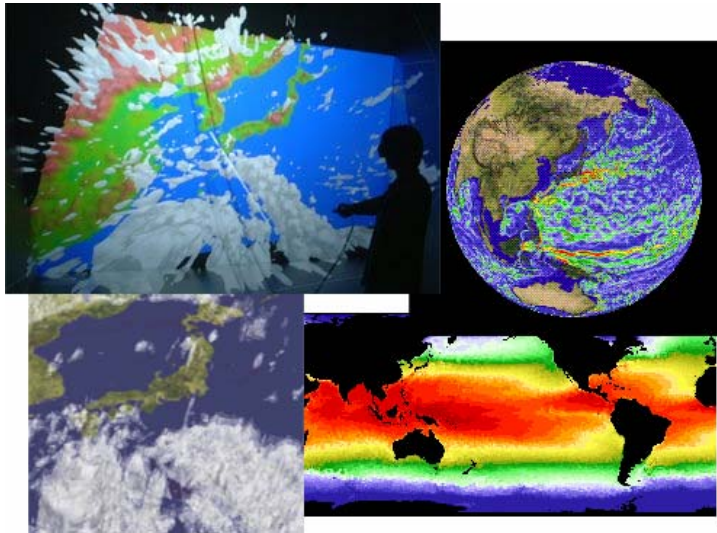


BlueGene/L uses a three-dimensional (3D) torus network in which the nodes (red balls) are connected to their six nearest-neighbor nodes in a 3D mesh. In the torus configuration, the ends of the mesh loop back, thereby eliminating the problem of programming for a mesh with edges. Without these loops, the end nodes would not have six near neighbors.

[https://asc.llnl.gov/computing\\_resources/bluegenel/configuration.html](https://asc.llnl.gov/computing_resources/bluegenel/configuration.html)

# Multiprozessorsysteme

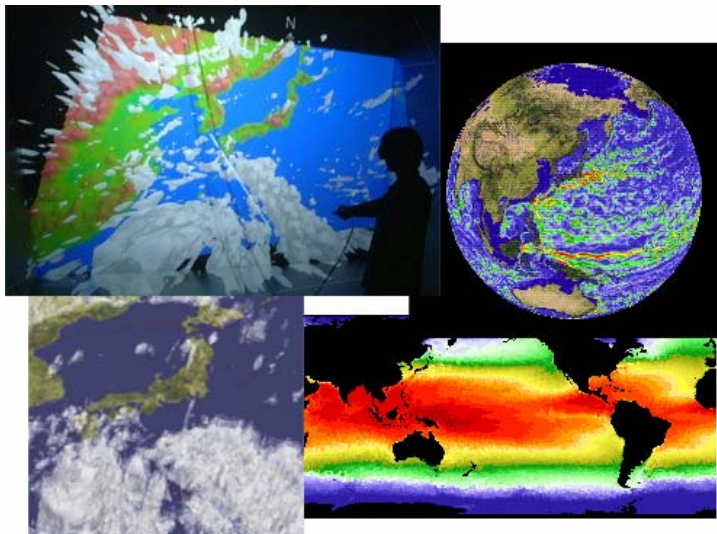
- **Motivation**
- Top500 Nr. 17 (November 2002): Earth Simulator
  - Anzahl Prozessoren: 5120
  - Leistung: 35,86 TFLOPS (Linpack),
  - Anwendung: Klimaforschung



Quelle: The Earth Simulator Center;  
<http://www.es.jamstec.go.jp/esc/research/Perception/index.en.html>

# Multiprozessorsysteme

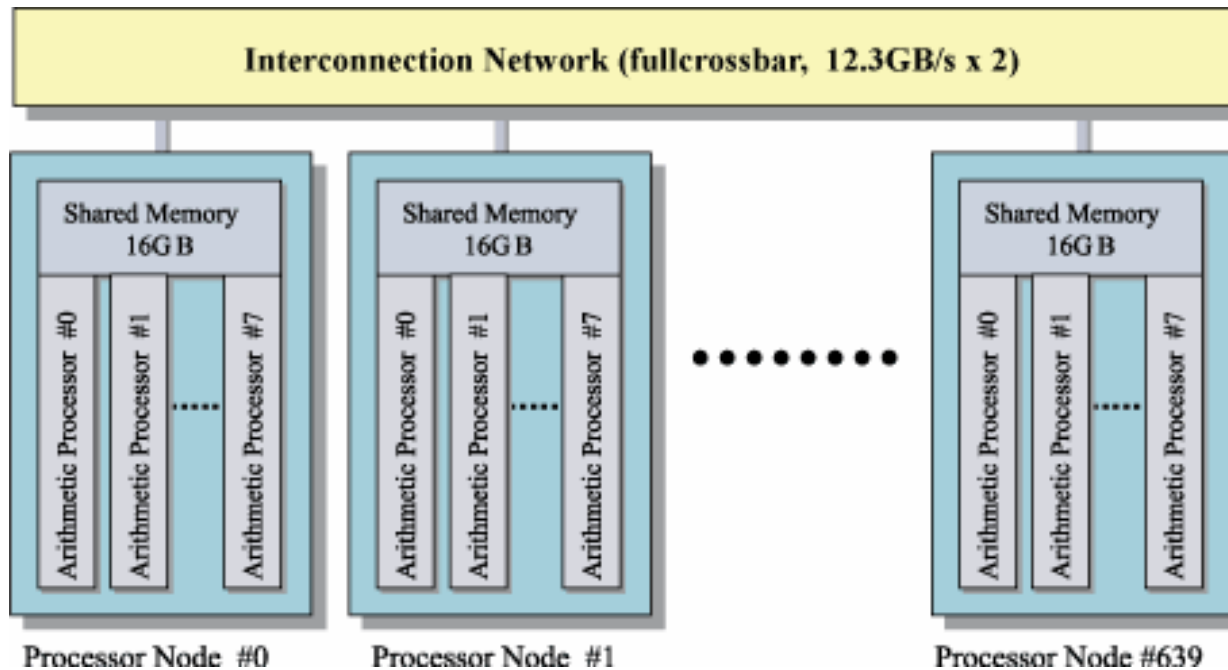
- **Motivation**
- Top500 Nr. 17 (November 2002): Earth Simulator
  - Anzahl Prozessoren: 5120
  - Leistung: 35,86 TFLOPS (Linpack),
  - Anwendung: Klimaforschung



Quelle: The Earth Simulator Center;  
<http://www.es.jamstec.go.jp/esc/research/Perception/index.en.html>

# Multiprozessorsysteme

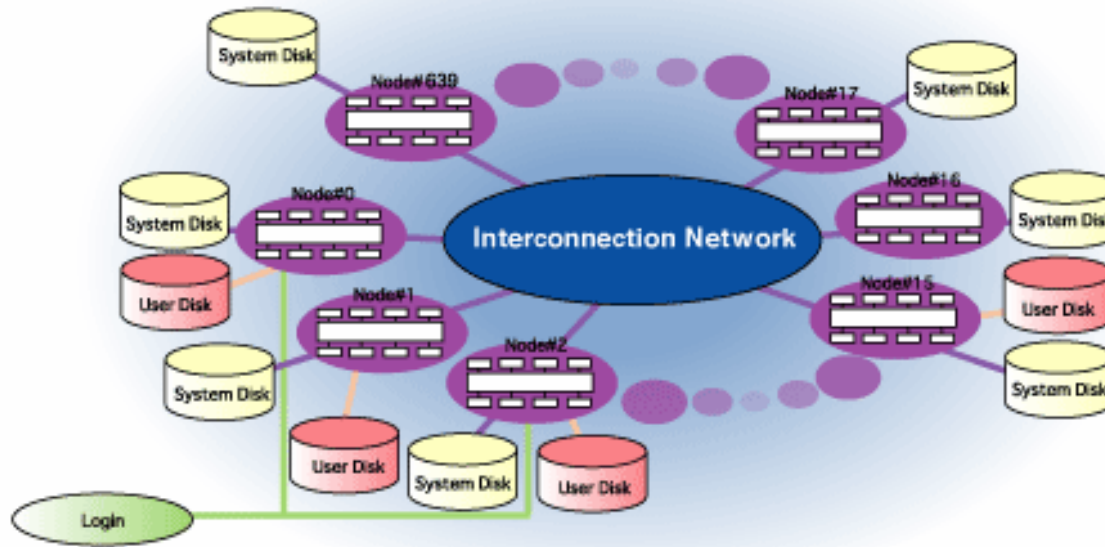
- Motivation
- Earth Simulator
  - Konfiguration



Quelle: The Earth Simulator Center;  
<http://www.es.jamstec.go.jp/esc/research/Perception/index.en.html>

# Multiprozessorsysteme

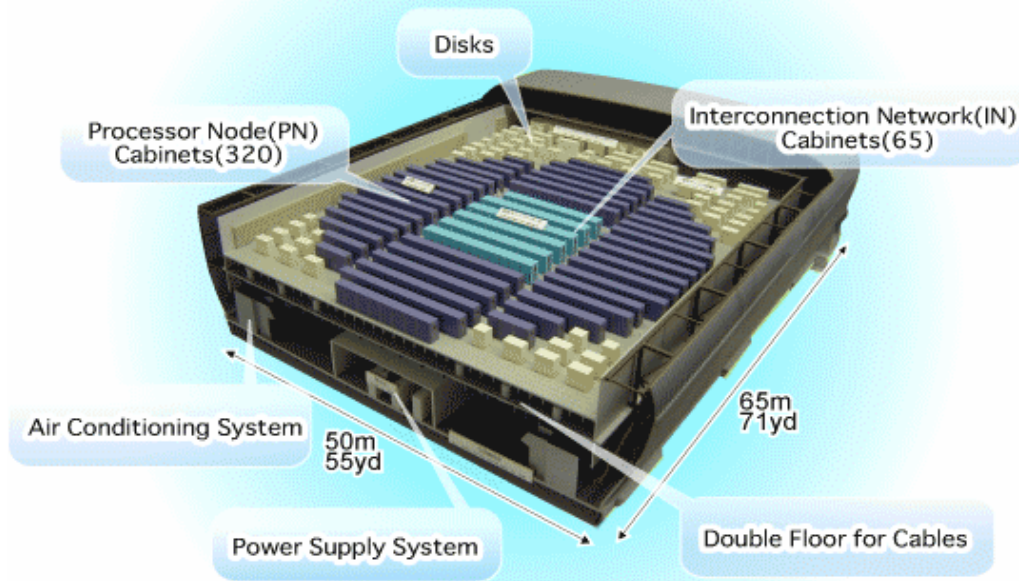
- Motivation
- Earth Simulator
  - Konfiguration



Quelle: The Earth Simulator Center;  
<http://www.es.jamstec.go.jp/esc/research/Perception/index.en.html>

# Multiprozessorsysteme

- Motivation
- Earth Simulator
  - Konfiguration

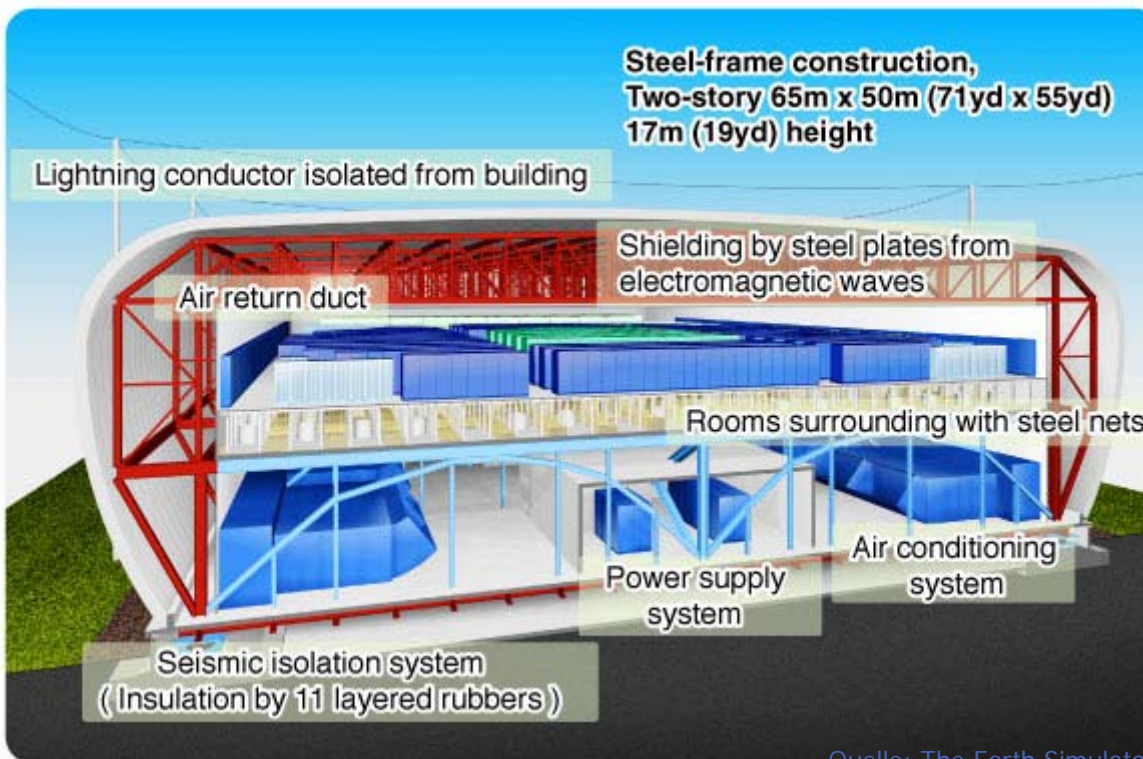


Quelle: The Earth Simulator Center;  
<http://www.es.jamstec.go.jp/esc/research/Perception/index.en.html>



# Multiprozessorsysteme

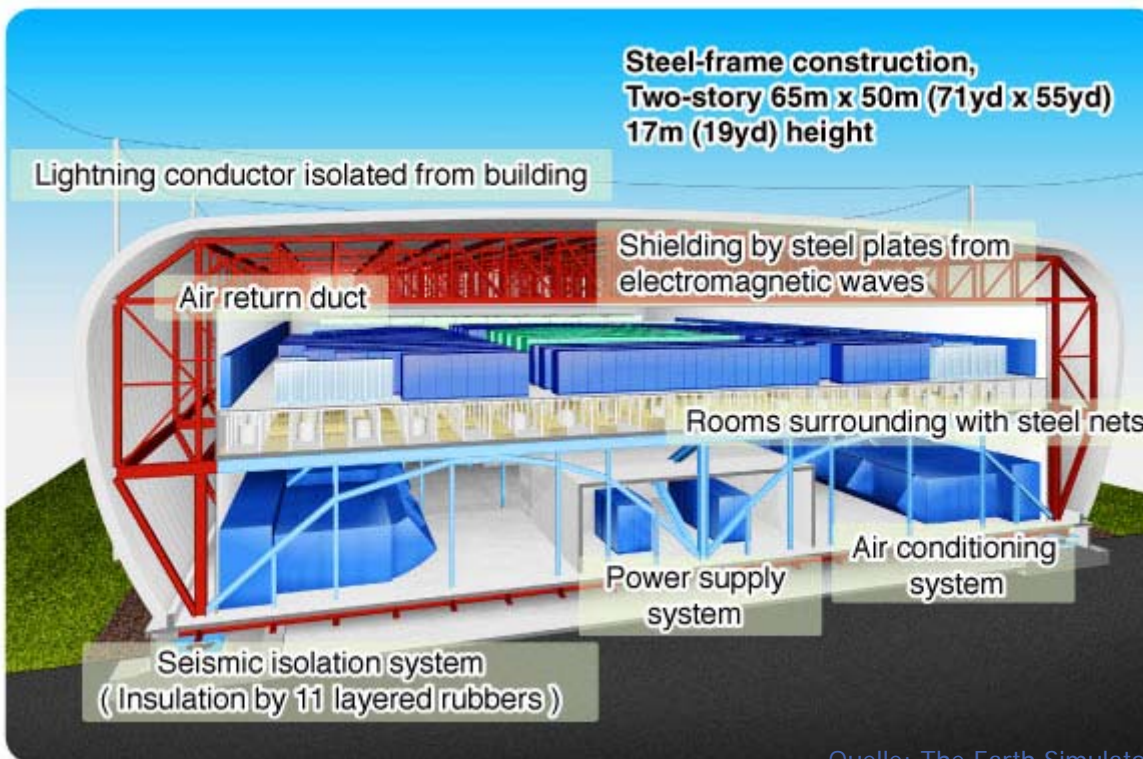
- Motivation
- Earth Simulator
  - Gebäude



Quelle: The Earth Simulator Center;  
<http://www.es.jamstec.go.jp/esc/research/Perception/index.en.html>

# Multiprozessorsysteme

- Motivation
- Earth Simulator
  - Gebäude



Quelle: The Earth Simulator Center;  
<http://www.es.jamstec.go.jp/esc/research/Perception/index.en.html>

# Multiprozessorsysteme

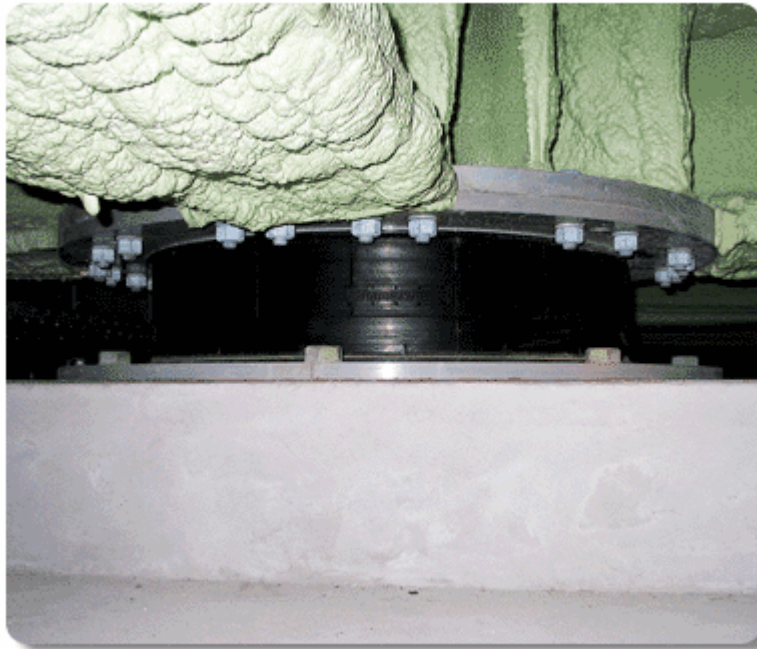
- Motivation
- Earth Simulator
  - Rechnerraum



Quelle: The Earth Simulator Center;  
<http://www.es.jamstec.go.jp/esc/research/Perception/index.en.html>

# Multiprozessorsysteme

- Motivation
- Earth Simulator
  - Rechnerraum



Quelle: The Earth Simulator Center;  
<http://www.es.jamstec.go.jp/esc/research/Perception/index.en.html>

# Multiprozessorsysteme

- **Motivation**
- Barcelona Supercomputer Center
- MareNostrum: BladeCenter JS21 Cluster, PPC 970 (10240) 2.3 GHz, Myrinet



By courtesy of Barcelona  
Supercomputing Center - [www.bsc.es](http://www.bsc.es)



# Vorlesung Rechnerstrukturen

- **Kapitel 3: Multiprozessoren – Parallelismus auf Prozess-/Blockebene**
- 3.2 Allgemeine Grundlagen

# Allgemeine Grundlagen

## ■ Parallele Architekturen

### ■ Definition Parallelrechner:

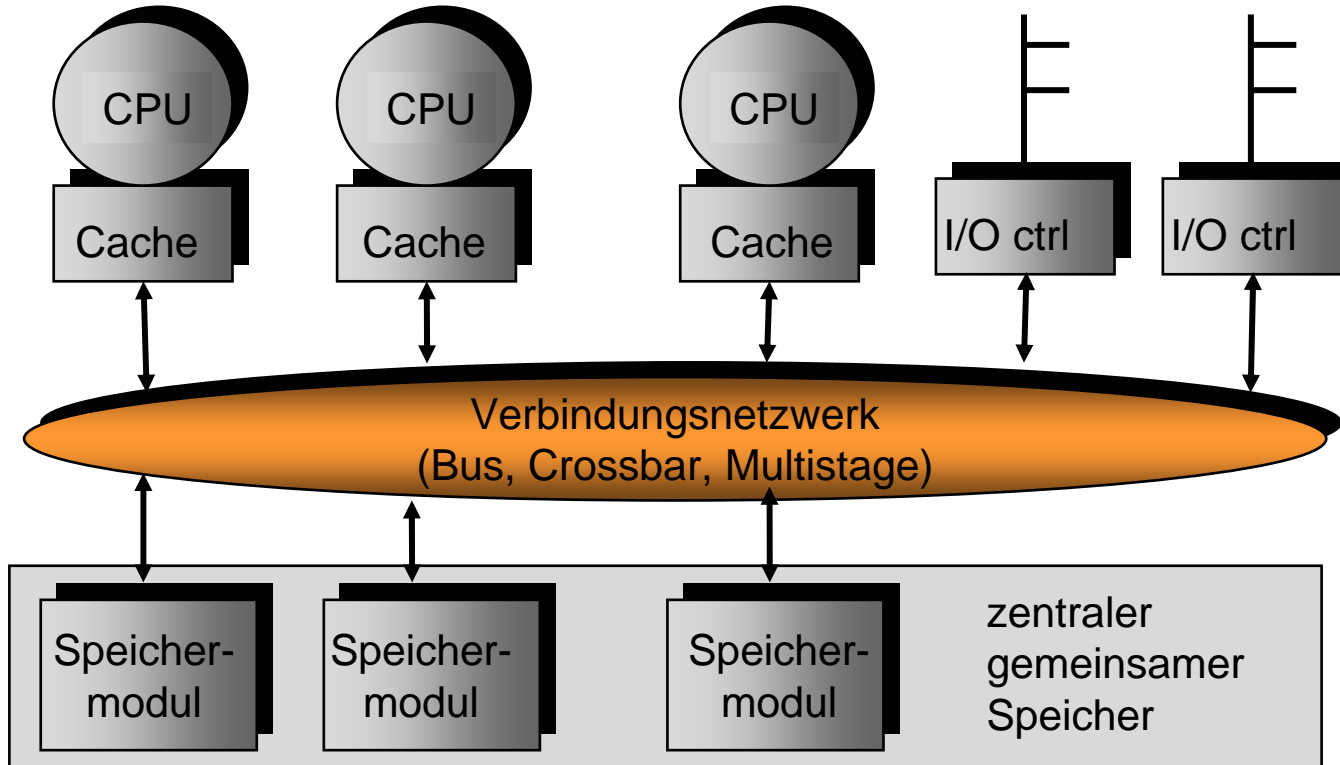
- „A collection of processing elements that communicate and cooperate to solve large problems“ (Almase and Gottlieb, 1989)

- Betrachtung einer parallelen Architektur als eine Erweiterung des Konzepts einer konventionellen Rechnerarchitektur um eine Kommunikationsarchitektur

# Parallele Architekturen

## ■ Multiprozessor mit gemeinsamem Speicher

- UMA: Uniform Memory Access
- Beispiel: symmetrischer Multiprozessor (SMP)
  - Gleichberechtigter Zugriff der Prozessoren auf die Betriebsmittel

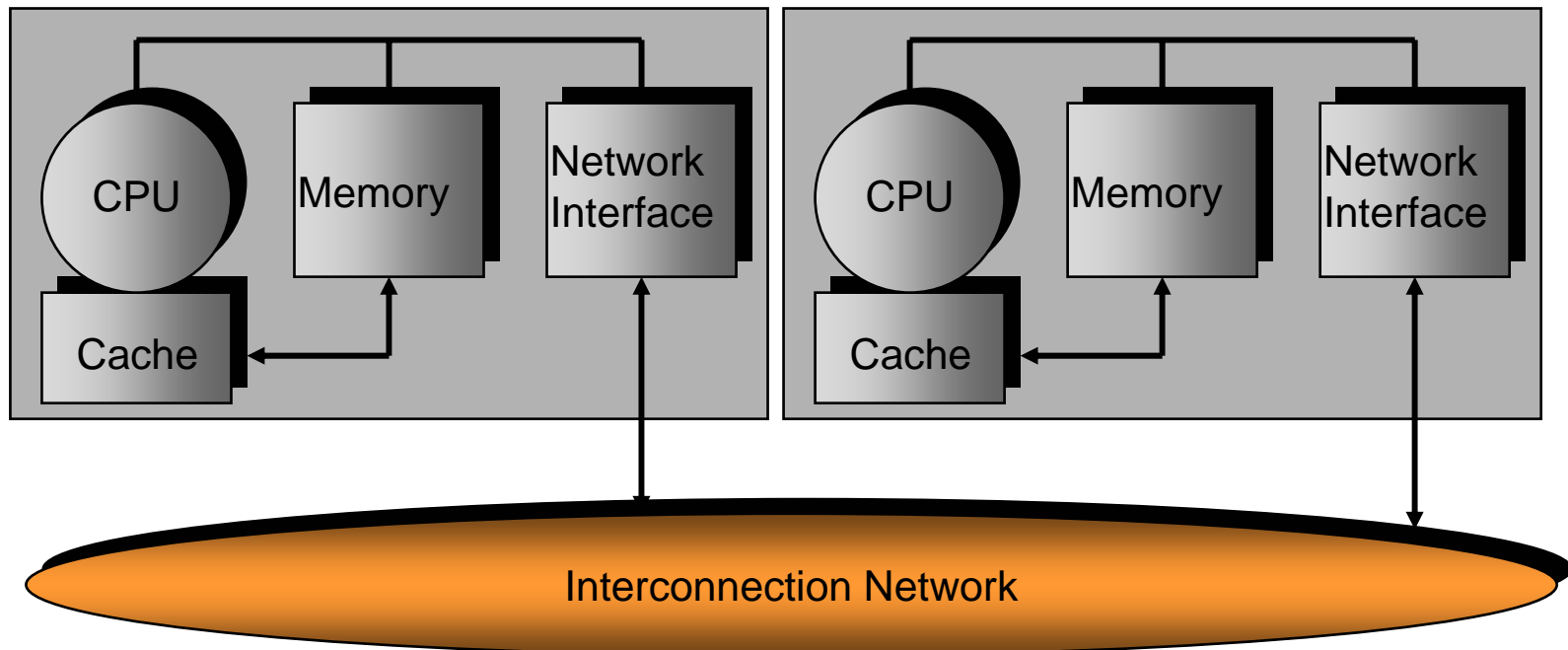




# Parallele Architekturen

## ■ Multiprozessor mit verteiltem Speicher

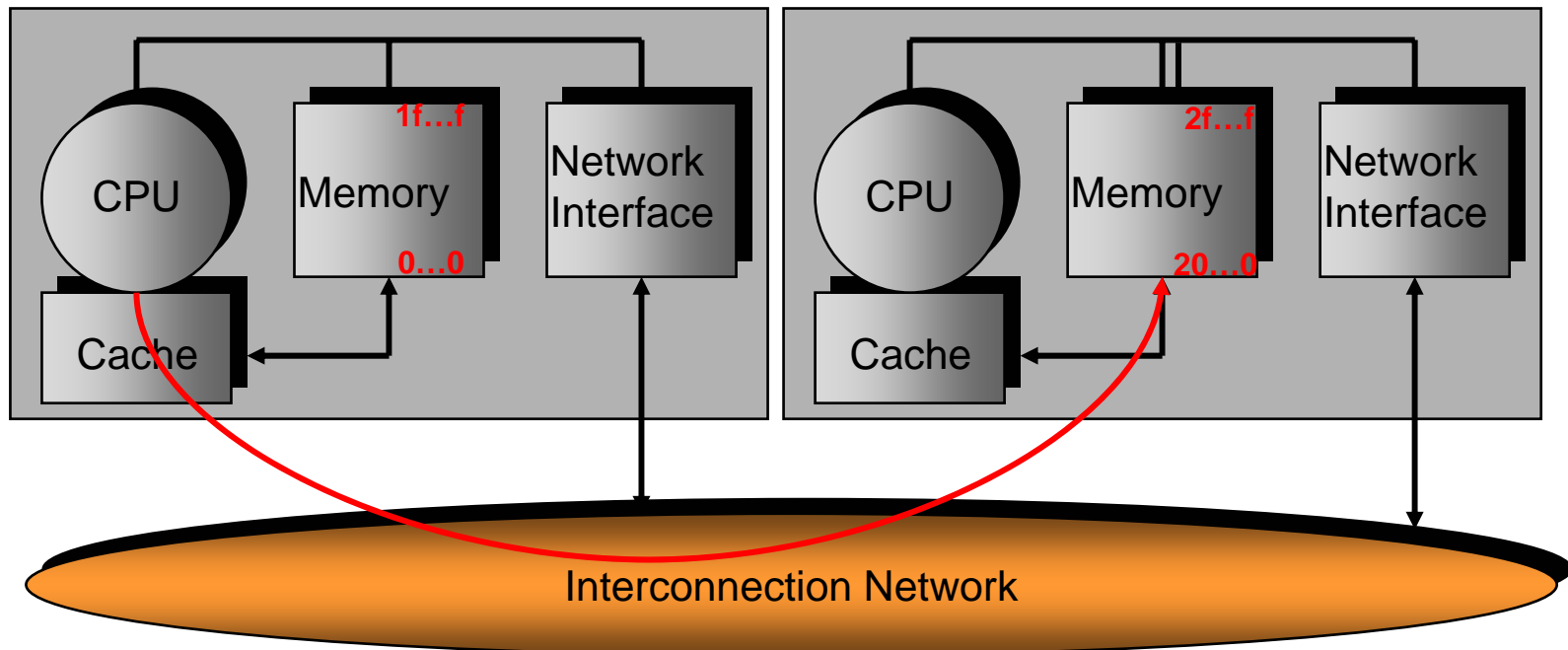
- NORMA No Remote Memory Access
- Beispiel: Cluster



# Parallele Architekturen

## ■ Multiprozessor mit verteiltem gemeinsamen Speicher

- NUMA: Non-Uniform Memory Access
- Globaler Adressraum: Zugriff auf entfernten Speicher



# Parallele Architekturen

## ■ Rechnerarchitektur

### ■ Abstraktion

- Benutzer-/System-Schnittstelle
- Hardware-/Software-Schnittstelle

### ■ Architektur

- Spezifiziert die Menge der Operationen an den Schnittstellen und die Datentypen, auf denen diese operieren

### ■ Organisation

- Realisierung der Abstraktionen

# Parallele Architekturen

## ■ Kommunikationsarchitektur

### ■ Abstraktion

- Benutzer-/System-Schnittstelle
- Hardware-/Software-Schnittstelle

### ■ Architektur

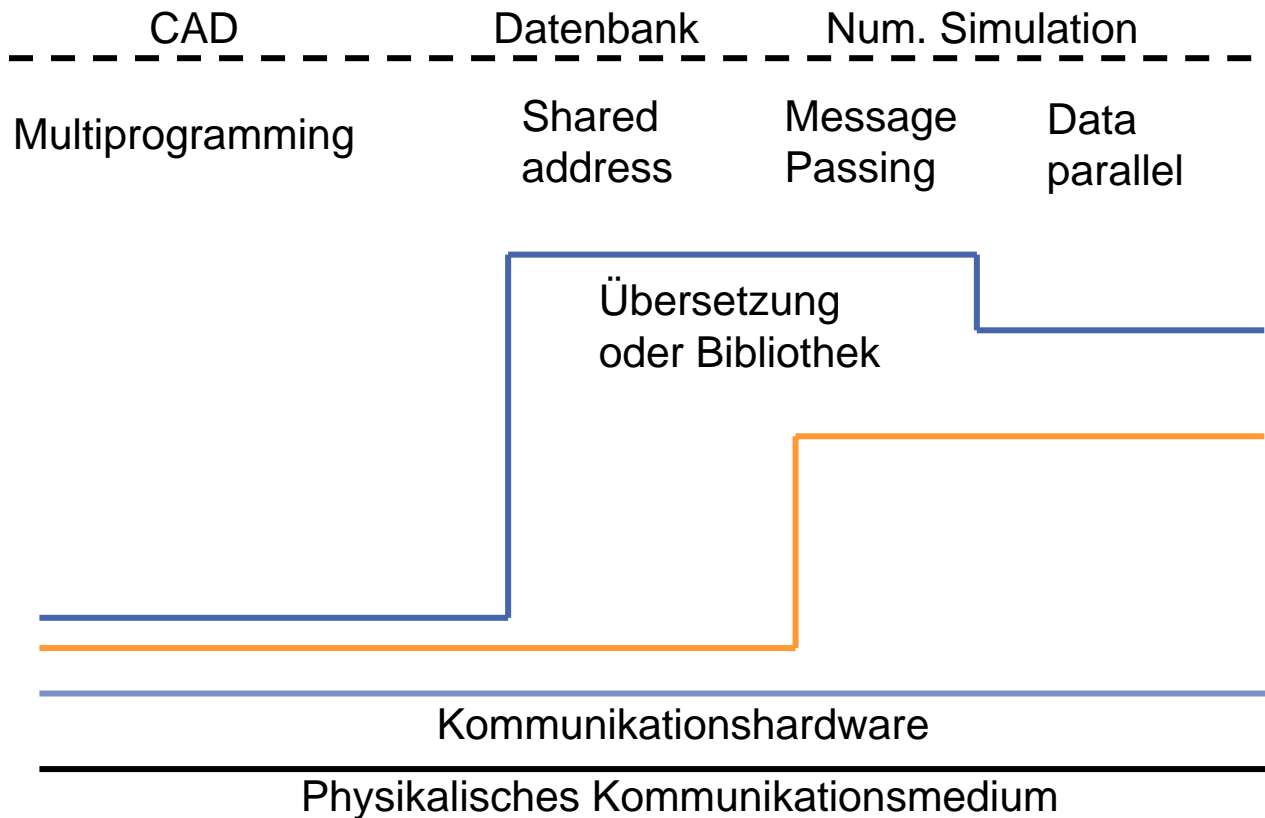
- Spezifiziert die Kommunikations- und Synchronisationsoperationen

### ■ Organisation

- Realisierung dieser Operationen

# Parallele Architekturen

## ■ Abstraktion



**Parallele Anwendung**

**Programmiermodell**

**Kommunikations-  
abstraktion**

**Benutzer/System-  
Schnittstelle**

**Hardware/Software-  
Schnittstelle**