Kapitel 10
Betriebssystemunterstützung

Virtuelle Speicherverwaltung
Schutzmechanismen
Unterbrechungsbehandlung

10.3 Schutzmechanismen

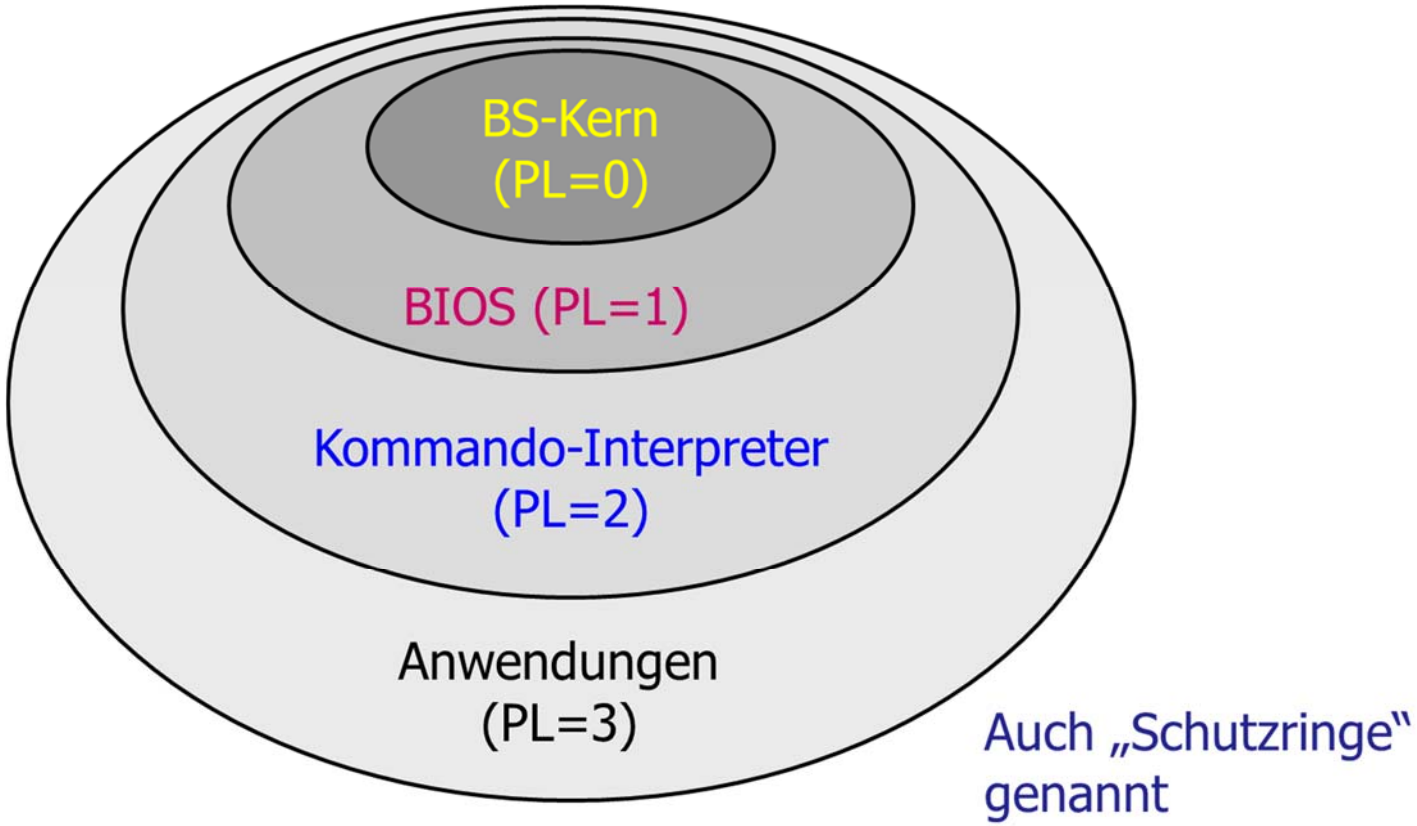
- Moderne Mikroprozessoren bieten Schutzmechanismen an, um während der Laufzeit von Programmen unerlaubte Speicherzugriffe zu verhindern
- Dies geschieht im wesentlichen durch:
  - Trennung der Systemsoftware, z.B. des Betriebssystems, insbesondere des Ein-/Ausgabe-Subsystem (BIOS, Basic I/O-System), von den Anwendungsprozessen
  - Trennung der Anwendungsprozessen voneinander; ist dies nicht gewährleistet, könnte ein fehlerhaftes Anwenderprogramm andere, fehlerfreie Programme beeinflussen (Schutzebenen und Zugriffsrechte)

10.3 Schutzmechanismen

- **Beispiel Intel IA32:**
  - **Schutzebenen (Priviledge Levels, PL):**
    - Wichtigstes Mittel zur Realisierung von Schutzmechanismen
  - Zweischutzebenen (bei Seitenverwaltung der x86-Prozessoren):
    - Betriebs-Systemmodus (supervisor mode)
    - Benutzermodus (user mode)
    - Ein Auftrag im Benutzermodus darf keine Daten des höher privilegierten Betriebssystemmodus benutzen
  - Vierstufige Hierarchie bei Segmentverwaltung:
    - Privileg-Ebene PL = 0 entspricht der vertrauenswürdigsten Ebene (most trust level)
    - Privileg-Ebene PL = 3 entspricht der am wenigsten vertrauenswürdigen Ebene (least trust level)

10.3 Schutzmechanismen

- **Beispiel Intel IA32:**
  - **Schutzebenen (Priviledge Levels, PL):**
    - **Schutzmaßnahmen der Segmentverwaltung**
      - Zugriffsschutz durch Zuweisen von Privilegstufen (Priviledge Levels)
      - Hierarchisches Schutzmodell mit vier Privilegstufen

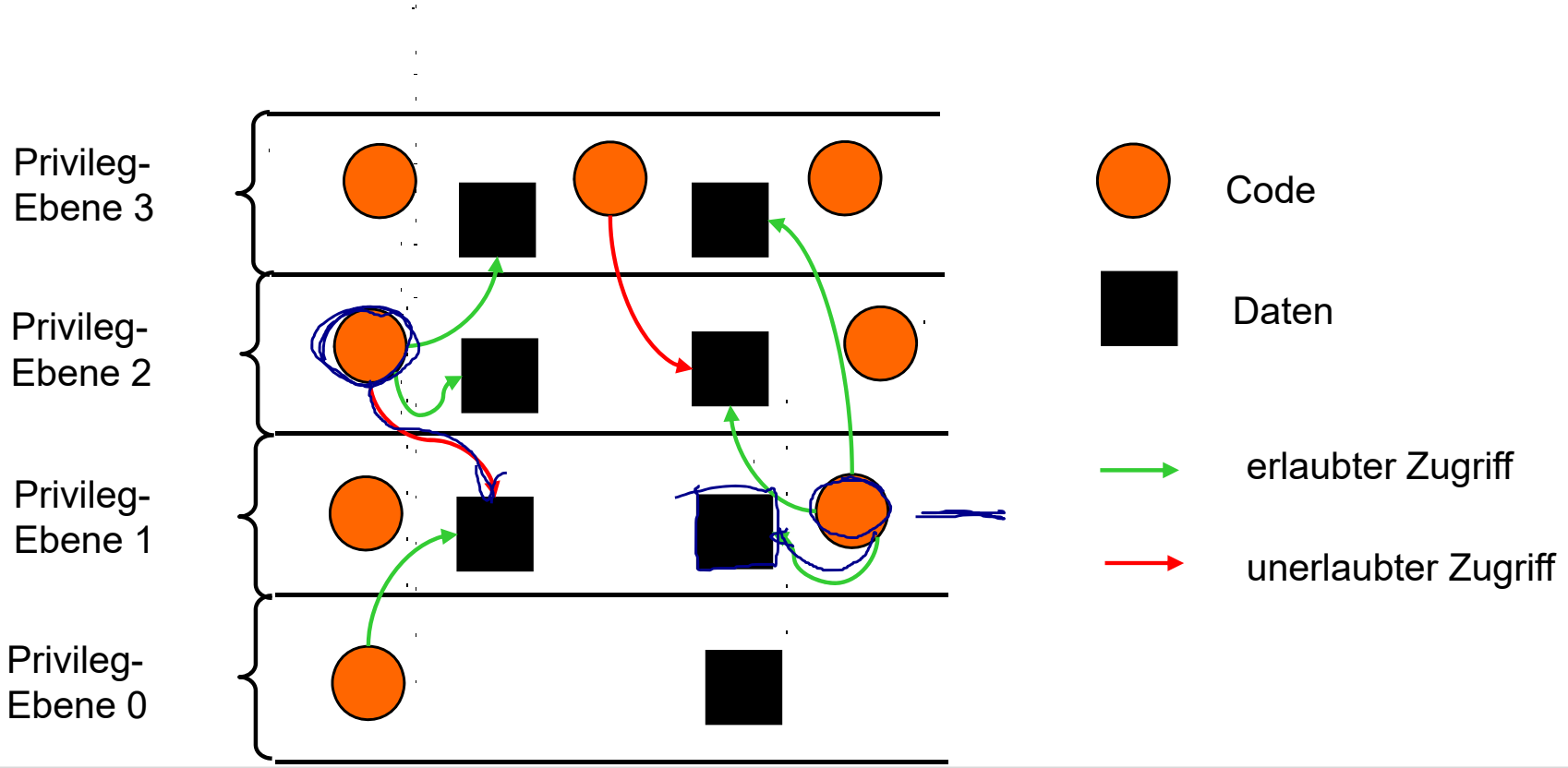


10.3 Schutzmechanismen

- **Beispiel Intel IA32:**
  - **Schutzebenen (Priviledge Levels, PL):**
    - **Schutzmaßnahmen der Segmentverwaltung**
      - Hierarchisches Schutzmodell mit vier Privilegstufen
        - Bei der Realisierung von Schutzmechanismen muss zwischen einem Zugriff auf Daten und einem Zugriff auf Programmcode unterschieden werden.
        - Regeln für den Zugriffsschutz für eine vierstufige Hierarchie
      - Vertrauenswürdigkeit
        - Ein Prozess darf nur auf Daten zugreifen, die höchstens genauso vertrauenswürdig sind wie er selbst.
        - Ein Prozess darf nur Code benutzen, der mindestens genauso vertrauenswürdig ist, wie er selbst.

10.3 Schutzmechanismen

- **Beispiel Intel IA32:**
  - **Schutzebenen (Priviledge Levels, PL):**
    - **Schutzmaßnahmen der Segmentverwaltung**
      - Vertrauenswürdigkeit: Ein Prozess darf nur auf Daten zugreifen, die höchstens genauso vertrauenswürdig sind wie er selbst.

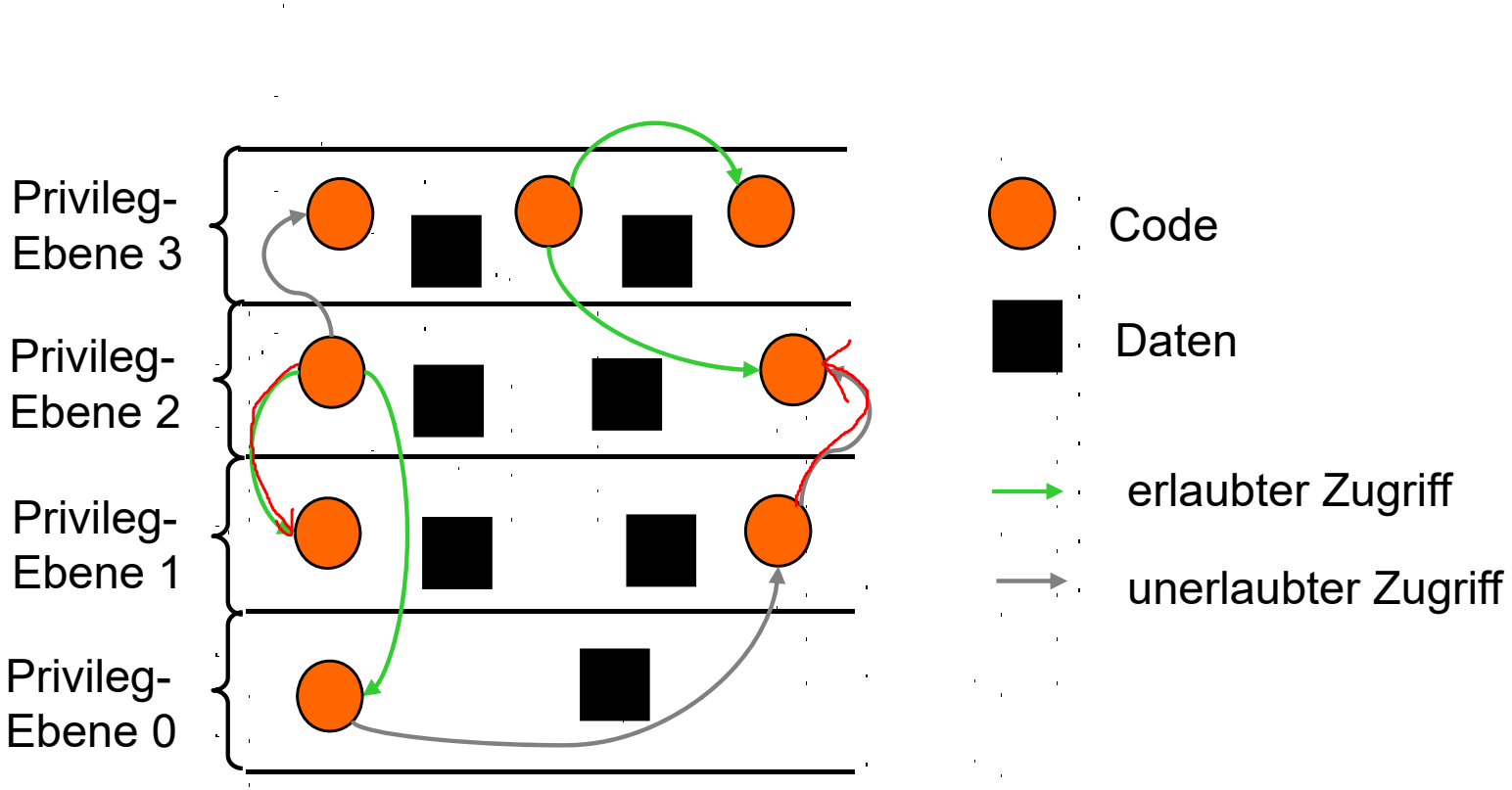


10.3 Schutzmechanismen

- **Beispiel Intel IA32:**
  - **Schutzebenen (Priviledge Levels, PL):**
    - **Schutzmaßnahmen der Segmentverwaltung**
      - Vertrauenswürdigkeit: Ein Prozess darf nur Code benutzen, der mindestens genauso vertrauenswürdig ist, wie er selbst.
        - Regel stellt sicher, dass der aufgerufene Code wenigstens den gleichen Sicherheitsansprüchen genügt, wie der aufrufende.
        - Sie verhindert, dass in einer aufgerufenen Prozedur mit niedrigem Privileg vor dem Rücksprung in die höhere Privilegebene die Rücksprungadresse auf dem Stack manipuliert und dadurch geschützter Code auf der höheren Ebene angesprungen wird.

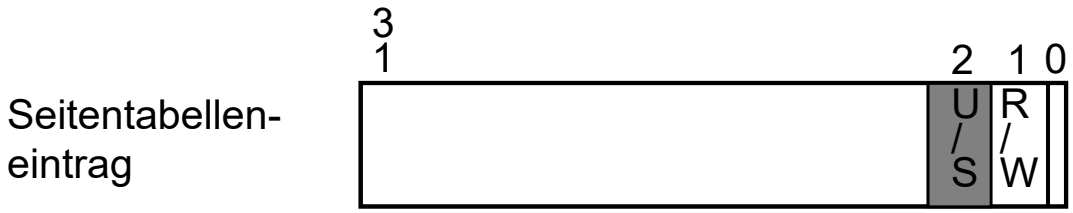
10.3 Schutzmechanismen

- **Beispiel Intel IA32:**
  - **Schutzebenen (Priviledge Levels, PL):**
    - **Schutzmaßnahmen der Segmentverwaltung**
      - Vertrauenswürdigkeit: Ein Prozess darf nur Code benutzen, der mindestens genauso vertrauenswürdig ist, wie er selbst.



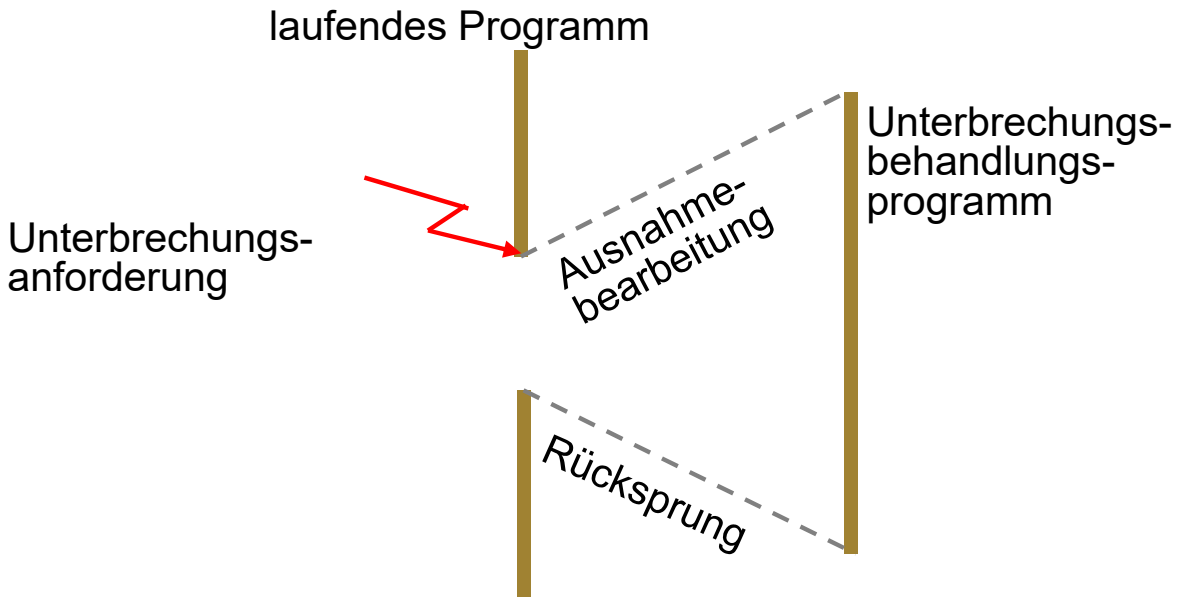
10.3 Schutzmechanismen

- **Beispiel Intel IA32:**
  - **Schutzebenen (Priviledge Levels, PL):**
    - **Schutzmaßnahmen der Seitenverwaltung**
      - Systemmodus (Supervisor-Level)
        - Für Betriebssystem, Systemsoftware und geschützte Systemdaten (z.B. Seitentabellen)
      - Benutzermodus (User-Level)
        - Für Anwendungen (Code, Daten des Benutzers)



10.4 Unterbrechungsbehandlung

- Bei der Abarbeitung eines Programms können Situationen eintreten, die eine Unterbrechung dieser Abarbeitung erforderlich machen, um auf die eingetretene Situation entsprechend reagieren zu können.
- Unterbrechungsbehandlung (Interrupt, Exception Processing) Reaktion eines Prozessors auf eine Unterbrechungsanforderung (Ausnahmebearbeitung)



10.4 Unterbrechungsbehandlung

- **Unterbrechungsbehandlung (Interrupt, Exception Processing):**
  - Reaktion eines Prozessors auf eine Unterbrechungsanforderung (Ausnahmebearbeitung)
  - Eine Ausnahmesituation erfordert eine vorübergehende Unterbrechung oder gar den Abbruch des laufenden Programms
  - Die Ausnahmebehandlung erfolgt durch eine Ausnahmeroutine (Interrupt Service Routine), deren Aktivierung durch eine Hardware-Komponente (Unterbrechungs-System, Interrupt System) im Steuerwerk unterstützt wird.
  - Die Ausnahmeroutine hat Ähnlichkeit mit dem Aufbau eines Unterprogramms

10.4 Unterbrechungsbehandlung

- **Unterbrechungsbehandlungsroutine vs. Unterprogramm**
  - **Aktivierung:**
    - call subroutine bei Unterprogramm
    - Hardware-Aktivierung durch externes Signal bei Ausnahmeroutine
  - **Beendigung:**
    - ret - Befehl bei Unterprogrammen (return from subroutine)
    - reti - Befehl bei Ausnahmebehandlung (return from interrupt)
  - **Einsprungadresse**
    - ins Unterprogramm: direkt im Programm
    - bei Ausnahmebehandlung über Interrupttabelle

10.4 Unterbrechungsbehandlung

- **Unterbrechungsbehandlungsroutine vs. Unterprogramm**
  - Status:
    - Unterprogrammaufruf sichert meist nur den PC auf den Stack
    - Ausnahmebehandlungs-Aufruf meist auch das PSW
  - Aufruf:
    - Unterprogrammaufrufe werden immer durchgeführt
    - die meisten Ausnahmebehandlungen werden nur aktiviert, falls das Interrupt-Enable-Bit im Steuerregister (PSW) gesetzt ist

10.4 Unterbrechungsbehandlung

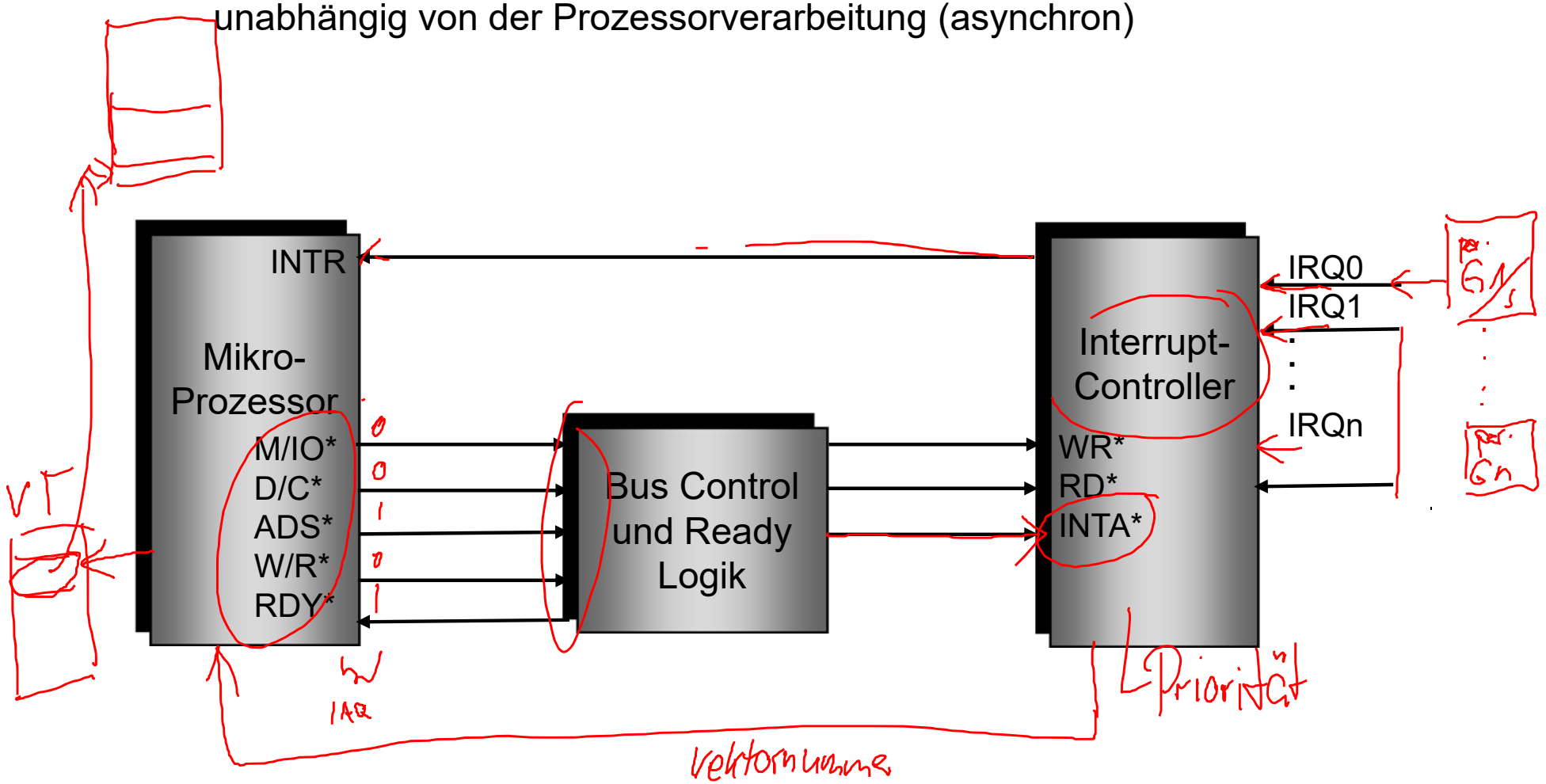
- **Arten von Unterbrechungen**
  - **Exceptions (Alarme)**
    - Programmunterbrechungen, die synchron zur Prozessorverarbeitung ausgelöst werden
    - **Faults**
      - Programmunterbrechungen, die vor der Ausführung der eine Ausnahmeverarbeitung verursachenden Instruktion gemeldet und bedient werden
      - Beispiel: Von der Seitenverwaltung wird gemeldet, wenn auf ein Segment oder eine Seite zugegriffen wird, die nicht im Hauptspeicher liegt
      - Es wird der Prozessorzustand gerettet, der eine nochmalige Ausführung der die Ausnahmeverarbeitung verursachende Instruktion erlaubt

10.4 Unterbrechungsbehandlung

- **Arten von Unterbrechungen**
  - **Interrupts**
    - Haben immer eine prozessorexterne Ursache und erfolgen in der Regel unabhängig von der Prozessorverarbeitung (asynchron)
    - **Nichtmaskierbare Interrupts**
      - Werden über einen speziellen Interrupteingang (NMI) gemeldet
      - Nichtmaskierbaren Interrupts wird immer stattgegeben
    - **Maskierbare Interrupts**
      - Werden über den Interrupteingang (INTR) gemeldet
      - Maskierbaren Interrupts wird nur stattgegeben, wenn das Interrupt-Enable Flag (IF Flag) im Steuerregister gesetzt ist:
        - Privilegierte Befehle zum Setzen und Zurücksetzen des IF-Flags
          - CLI (Clear Interrupt Enable Flag)
          - STI (Set Interrupt Enable Flag)

10.4 Unterbrechungsbehandlung

- **Arten von Unterbrechungen**
  - **Interrupts**
    - Haben immer eine prozessorexterne Ursache und erfolgen in der Regel unabhängig von der Prozessorverarbeitung (asynchron)



10.4 Unterbrechungsbehandlung

- **Arten von Unterbrechungen**
  - **Unterbrechungsbedingungen (Beispiel Intel IA-32)**

Vektornummer	Unterbrechungs-bedingung	Quelle	Art	
0	Division durch 0	DIV, IDIV	Fault	Prozessor entdeckt bei Ausführung eines Divisionsbefehls einen Divisor mit dem Wert 0
1	Debug Exception	jeder Befehl	Traps/ Faults	
2	NMI	extern	NMI	
...	...	...	...	...
10	Invalid TSS	JMP,CALL, IRET, INT (intern)	Fault	

10.4 Unterbrechungsbehandlung

- **Arten von Unterbrechungen**
  - **Unterbrechungsbedingungen (Beispiel Intel IA-32)**

Vektornummer	Unterbrechungs-bedingung	Quelle	Art	
11	Segment not Present	Segment-Registerbef.,	Fault	Present-Bit eines Deskriptors ist zurückgesetzt
12	Stack Fault	Zugriff auf Stack	Fault	
13	General Protection	Speicherzugriff (int.)	Fault	Verletzung von Zugriffsrechten
14	Page Fault	Speicherzugriff (int.)	Fault	Seite nicht im Speicher
16	FP Error	ESC, FP-Bef.	Fault	Fehler bei der Ausführung eins Gleitkommabefehls
0-255	Software Interupts	INTn	Trap	
32-255	Maskierbare Interrupts	Extern	Inter.	

10.4 Unterbrechungsbehandlung

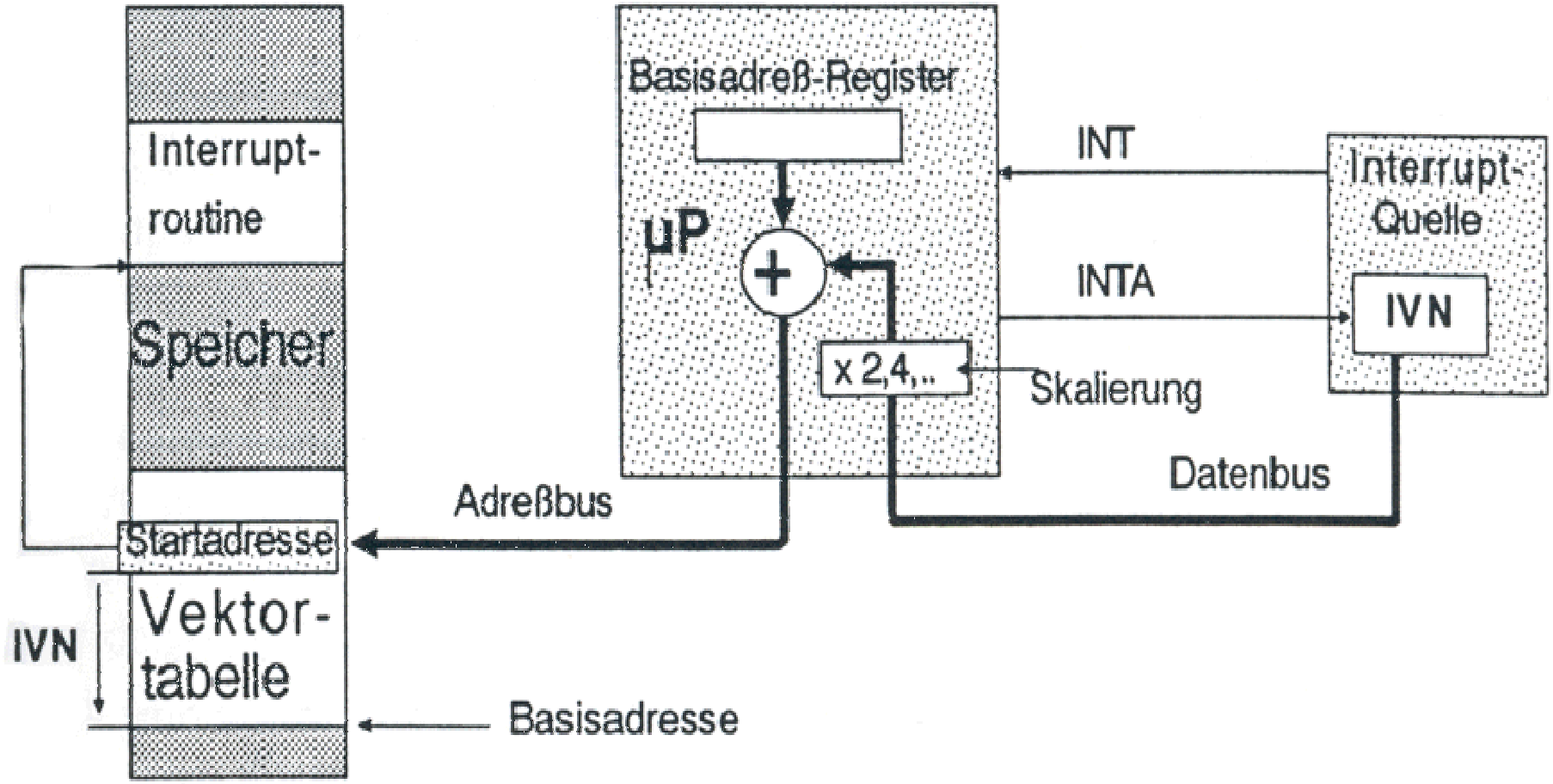
- **Arten von Unterbrechungen**
  - **Vektortabelle**
    - Tabelle an spezieller Speicheradresse (oft in einer der untersten Speicherseiten)
    - Enthält die Startadressen der Behandlungsroutinen
    - Interrupt-Quelle liefert bei Interrupt eine Interrupt-Vektor-Nummer (IVN), welche den Eintrag in der Interruptvektor-Tabelle charakterisiert
    - Basis-Adressregister enthält die Basisadresse der Tabelle

10.4 Unterbrechungsbehandlung

- **Arten von Unterbrechungen**
  - **Priorität von Interrupts und Exceptions**
    - Werden am INTR- und am NMI-Eingang gleichzeitig Unterbrechungsanforderungen angezeigt, dann hat die am NMI-Eingang angezeigte Vorrang
    - Bei der Ausführung eines Befehls können mehrere Exceptions bzw. Fehlerbedingungen angezeigt werden
    - Festgelegte Reihenfolge bei der Überprüfung von Fehlerbedingungen

10.4 Unterbrechungsbehandlung

- **Berechnung der Startadresse der Unterbrechungsbehandlungsroutine**

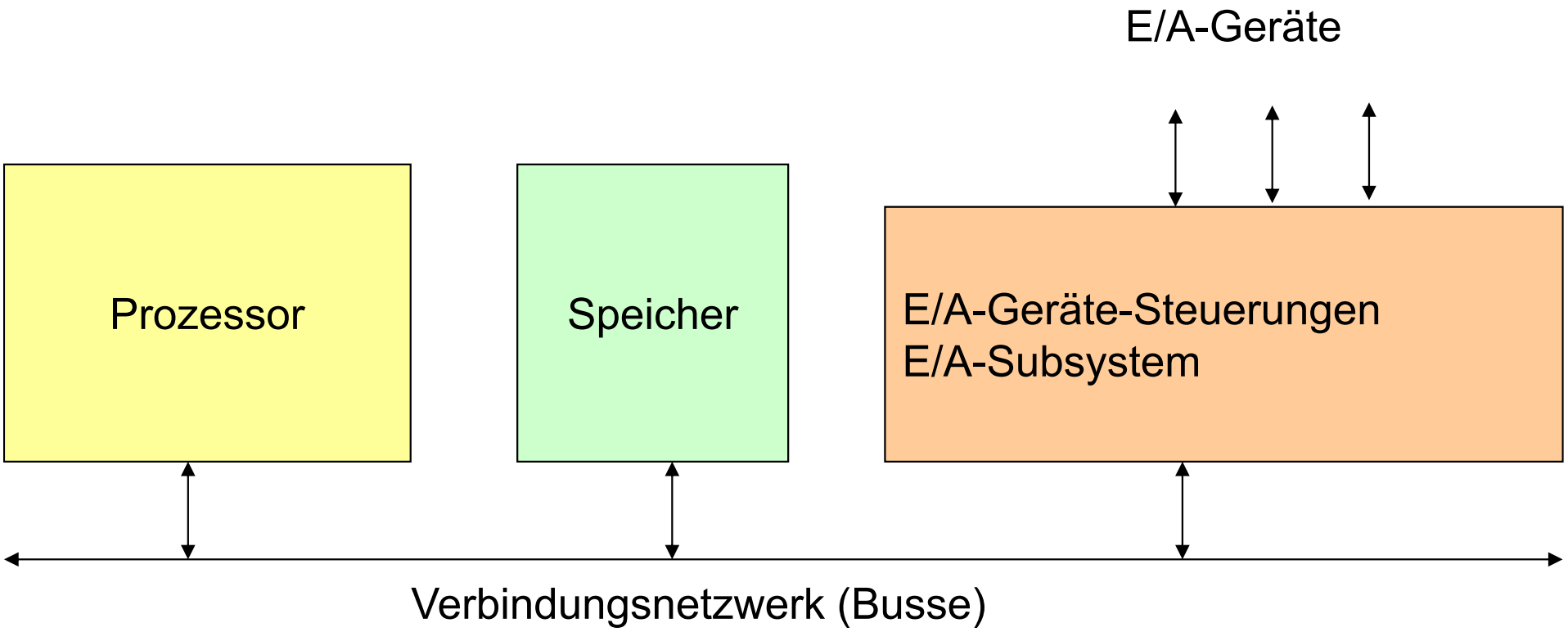


Kapitel 11
Parallele Rechnerstrukturen

Parallele Prozessorarchitekturen
Parallelrechner

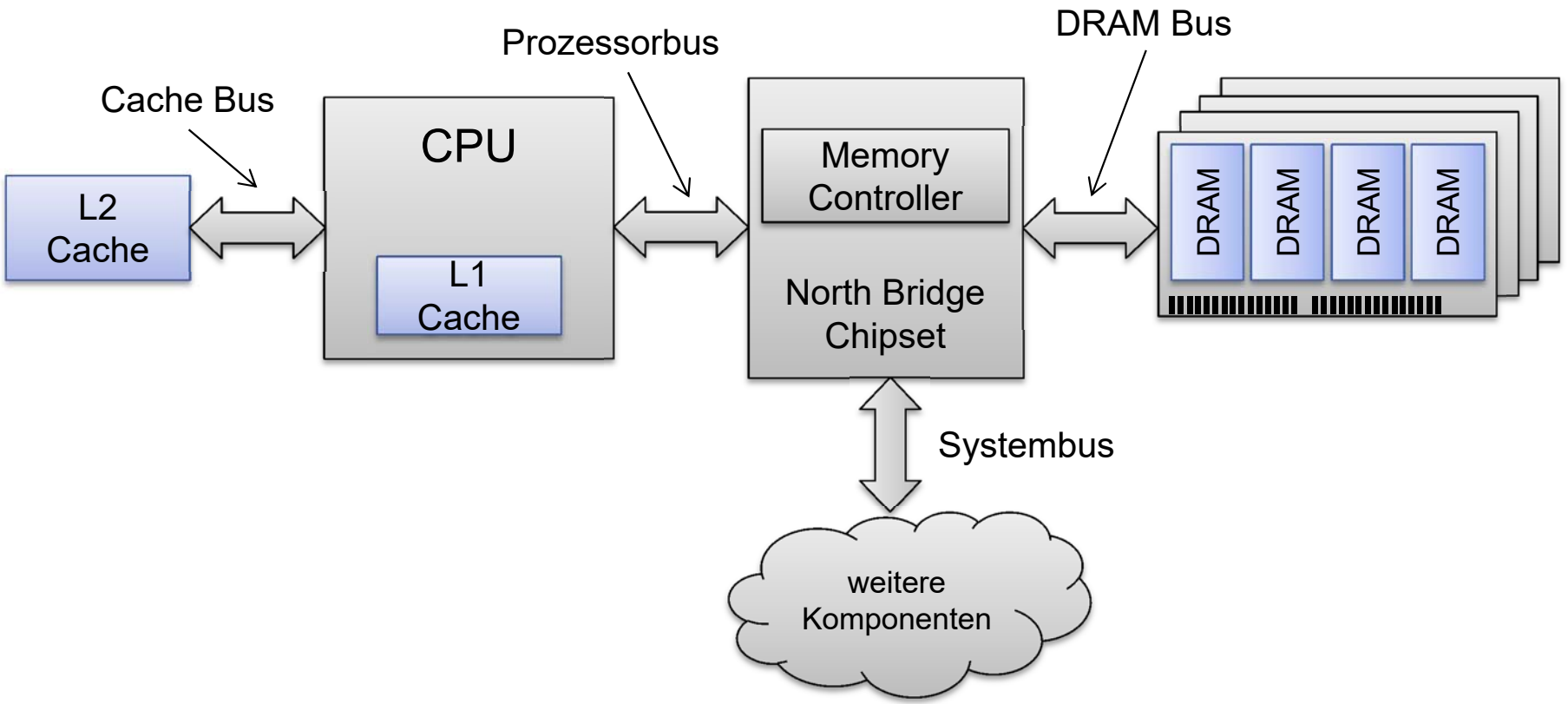
11.1 Motivation

- **Einfaches Rechnermodell**



11.1 Motivation

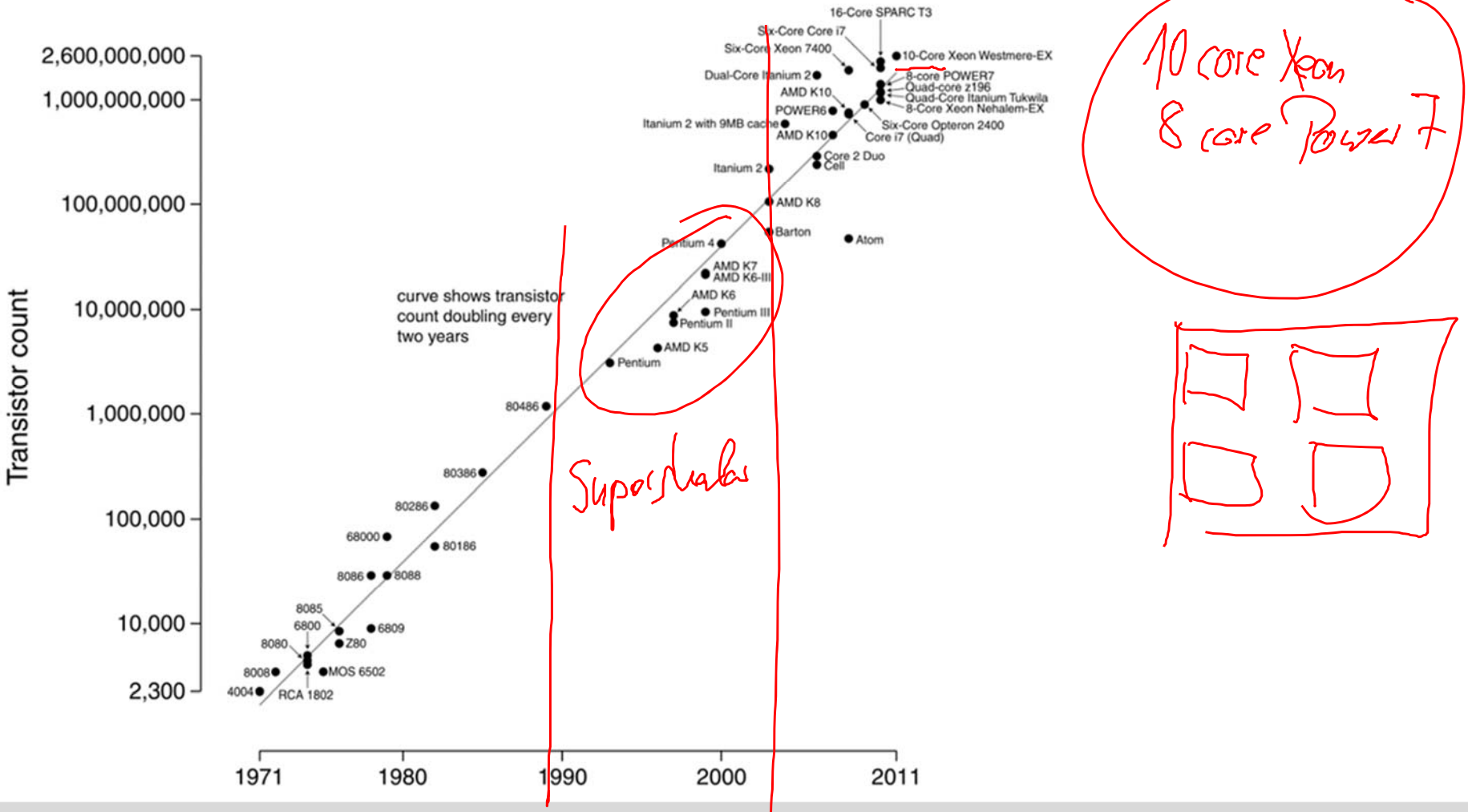
- **Beispiel: Aufbau klassischer PC**



11.1 Motivation

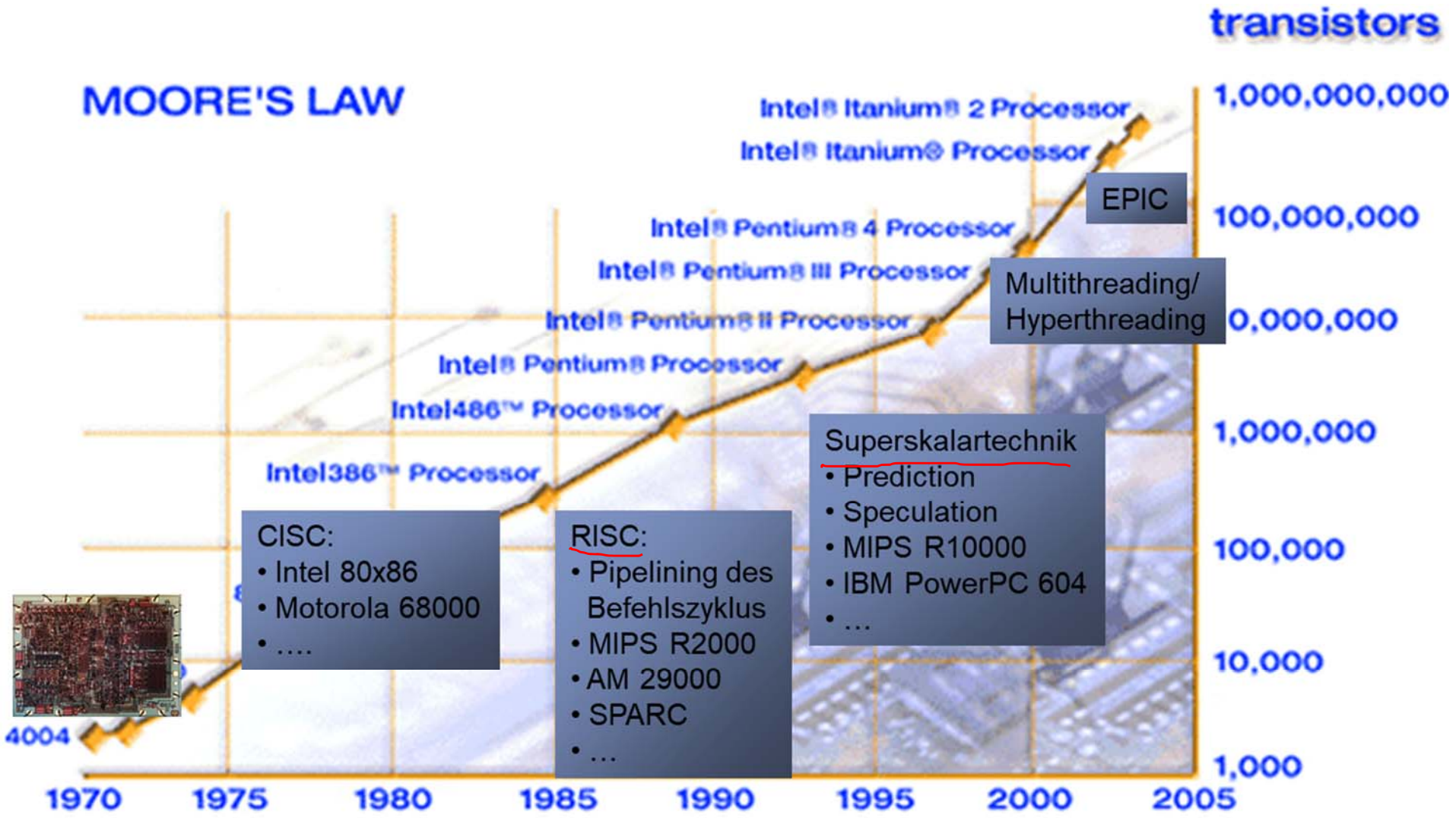
Entwicklung der Rechnertechnologie

Moore‘s Law am Beispiel von Mikroprozessoren



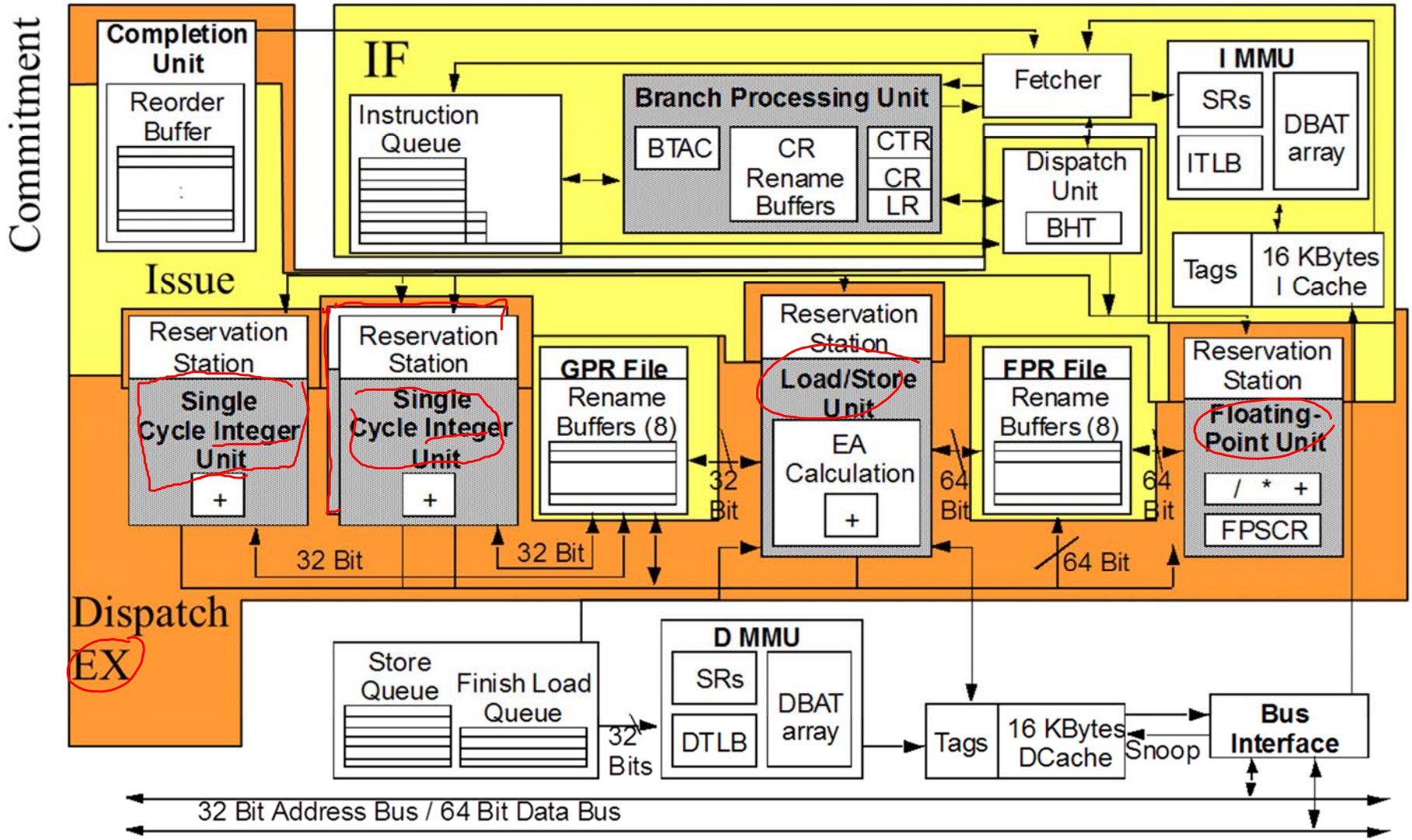
11.1 Motivation

- **Entwicklung der Mikroprozessoren**



11.1 Motivation

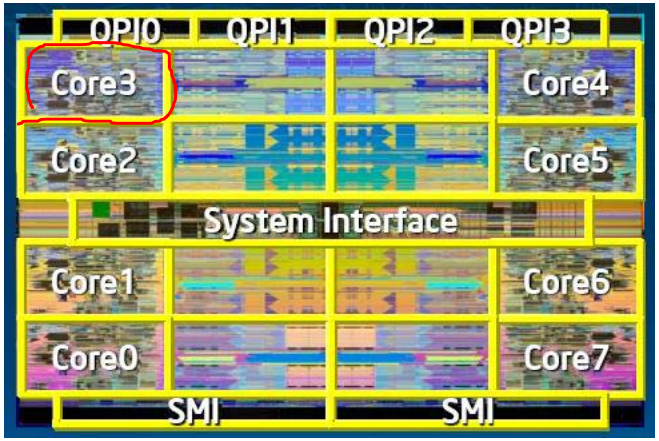
- **Superskalartechnik**



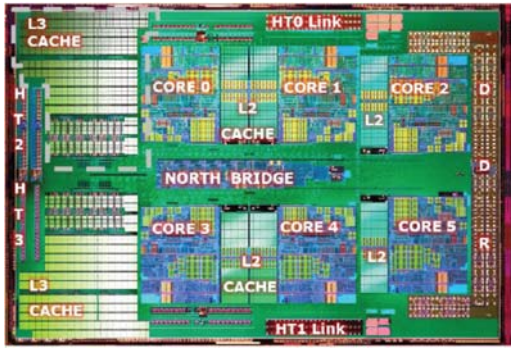
11.1 Motivation

- **Multicore, Manycore: homogene Strukturen**

Intel Nehalem – 8 cores

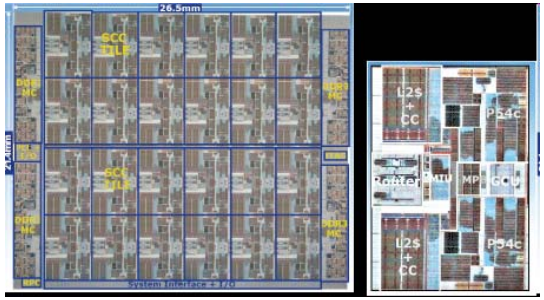


AMD „Magny Cours“ – 6 cores



Quelle: Pat Conway, **AMD**: Blade Computing with the AMD Opeteron™ Processor ("Magny Cours"), HotChips-21, Standford, 2009
http://www.hotchips.org/archives/hc21/

Intel SCC – 48 cores



http://download.intel.com/pressroom/images/rockcreek/scc-h-rack.jpg

- Integration mehrerer identischer Prozessorkerne auf einem Chip
- Thread and Task-level Parallelism
- Gleiche Befehlssatzarchitektur

11.1 Motivation

Power Systems

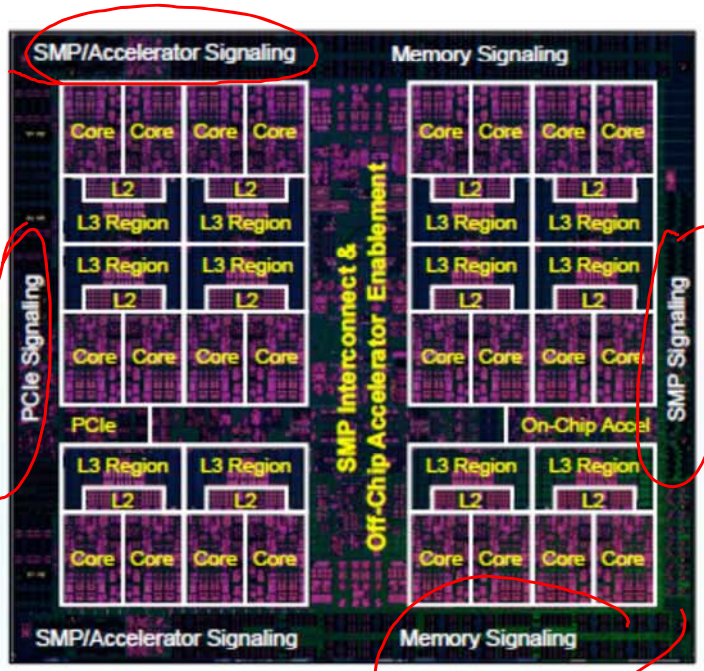
POWER9 Processor – Common Features

IBM

New Core Microarchitecture
- Stronger thread performance
- Efficient agile pipeline
- POWER ISA v3.0

Enhanced Cache Hierarchy
- 120MB NUCA L3 architecture
- 12 x 20-way associative regions
- Advanced replacement policies
- Fed by 7 TB/s on-chip bandwidth

Cloud + Virtualization Innovation
- Quality of service assists
- New interrupt architecture
- Workload optimized frequency
- Hardware enforced trusted execution



14nm finFET Semiconductor Process
- Improved device performance and reduced energy
- 17 layer metal stack and eDRAM
- 8.0 billion transistors

Leadership Hardware Acceleration Platform
- Enhanced on-chip acceleration
- Nvidia NVLink 2.0: High bandwidth, advanced new features
- CAPI 2.0: Coherent accelerator and storage attach (PCIe G4)
- New CAPI: Improved latency and bandwidth, open interface

State of the Art I/O Subsystem
- PCIe Gen4 – 48 lanes

High Bandwidth Signaling Technology
- 16 Gb/s interface
  - Local SMP
- 25 Gb/s interface – 25G Link
  - Accelerator, remote SMP

4

Quelle:http://www.hotchips.org/wp-content/uploads/hc_archives/hc28/HC28.23-Tuesday-Epub/HC28.23.90-High-Perform-Epub/HC28.23.921-.POWER9-Thompto-IBM-final.pdf

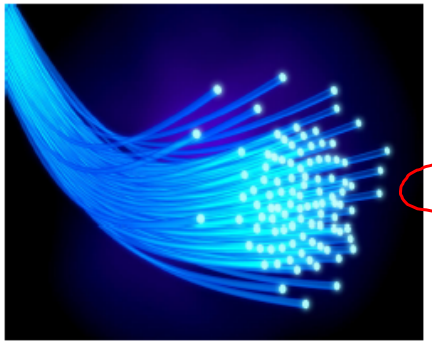
11.1 Motivation

IBM: POWER9

Power Systems

IBM

Modular Constructs → High-speed 25 Gb/s Signaling



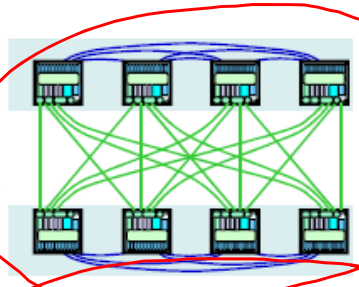
Utilize Best-of-Breed
25 Gb/s Optical-Style
Signaling Technology

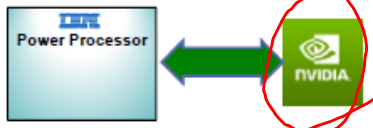
Multi-Drawer SMP Interconnect

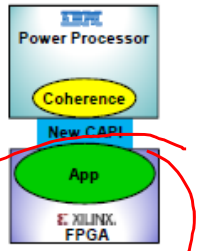
NVLINK 2 GPU Accelerator Attach

Open CAPI Accelerator Attach

Flexible & Modular
Packaging
Infrastructure







11.1 Motivation

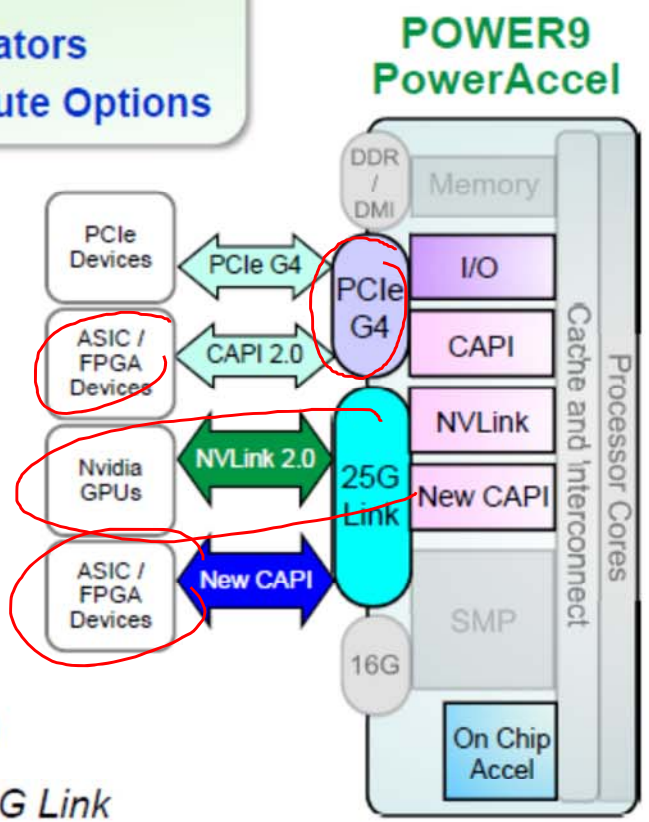
Power Systems

POWER9 – Premier Acceleration Platform

IBM

- **Extreme Processor / Accelerator Bandwidth and Reduced Latency**
- **Coherent Memory and Virtual Addressing Capability for all Accelerators**
- **OpenPOWER Community Enablement – Robust Accelerated Compute Options**

- **State of the Art I/O and Acceleration Attachment Signaling**
  - **PCIe Gen 4** x 48 lanes – 192 GB/s duplex bandwidth
  - **25G Link** x 48 lanes – 300 GB/s duplex bandwidth
- **Robust Accelerated Compute Options with OPEN standards**
  - **On-Chip Acceleration** – Gzip x1, 842 Compression x2, AES/SHA x2
  - **CAPI 2.0** – 4x bandwidth of POWER8 using *PCIe Gen 4*
  - **NVLink 2.0** – Next generation of GPU/CPU bandwidth and integration
  - **New CAPI** – High bandwidth, low latency and open interface using *25G Link*

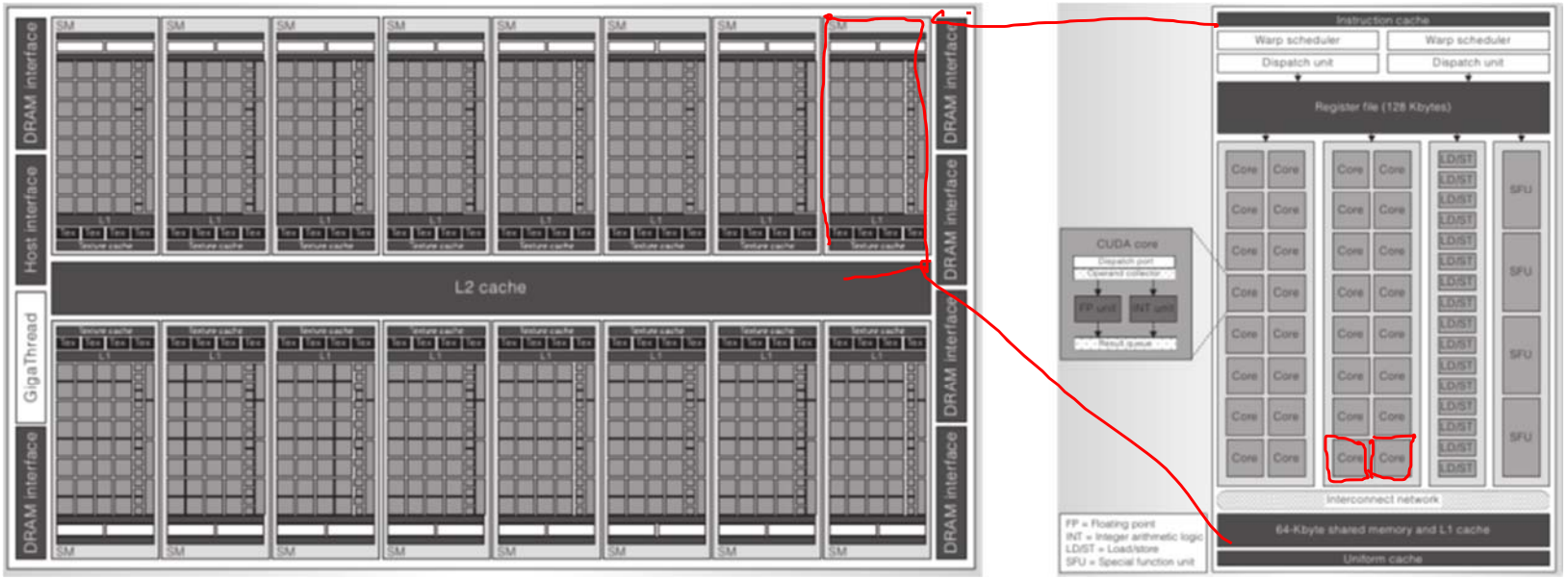


14

Quelle:http://www.hotchips.org/wp-content/uploads/hc_archives/hc28/HC28.23-Tuesday-Epub/HC28.23.90-High-Perform-Epub/HC28.23.921-.POWER9-Thompto-IBM-final.pdf

11.1 Motivation

- Entwicklung der Mikroprozessoren
- Beschleuniger, GPGPUs: Beispiel NVIDIA



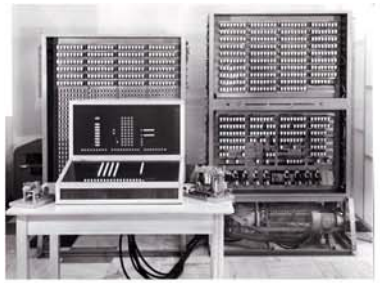
Quelle: N.J. Nickolls, J.W:Dally: The GPU computing Era. IEEE Micro, 31(2), 2010

- Datenparallelismus
- ALU Replikation, e.g. FP Beschleuniger
- Host/Master ←→ Accelerator/Slave

11.1 Motivation

- **Entwicklung im Bereich der Höchstleistungsrechner**

$$\text{MFlops} = \frac{\text{Anzahl der ausgeführten Gleitkommainstruktionen}}{10^6 \times \text{Ausführungszeit}}$$



Zuse Z1 1 Flops (10^0)
1941

Cray 2

GigaFlops (10^9)
1985

Intel ASCI Red

TeraFlops (10^{12})
1997

Roadrunner

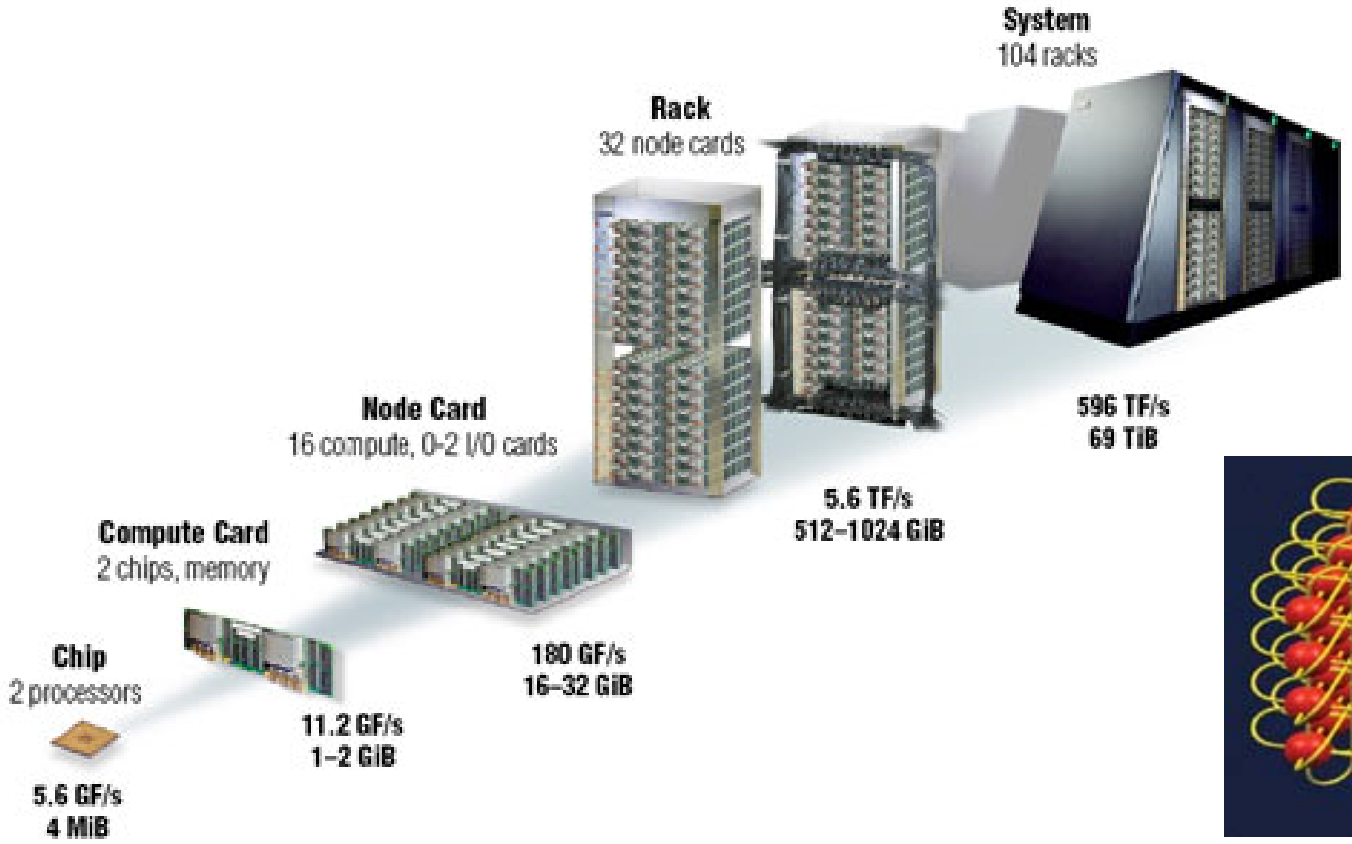
PetaFlops(10^{15})
2008

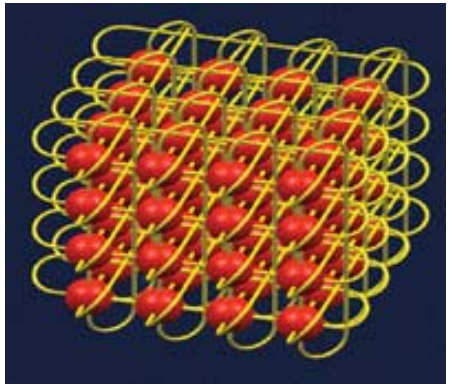
?

ExaFlops(10^{18})

11.1 Motivation

Multiprozessorsysteme:





BlueGene/L: Interconnection Network

Quelle: https://asc.llnl.gov/computing_resources/bluegenel/photogallery.html