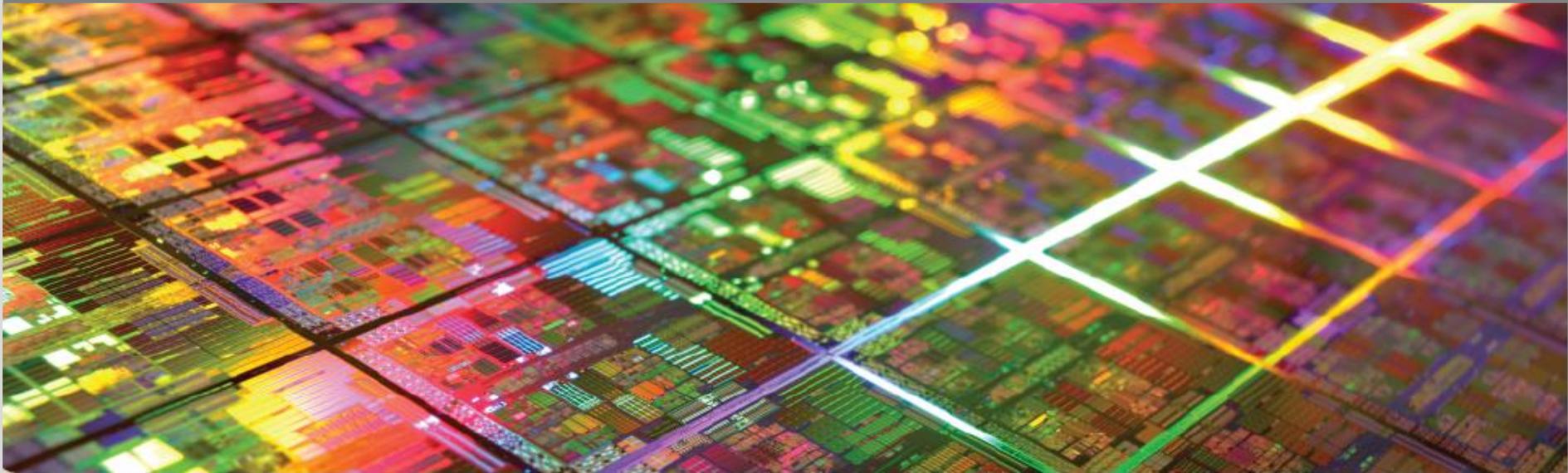


Rechnerstrukturen

Vorlesung im Sommersemester 2017

Prof. Dr. Wolfgang Karl

Institut für Technische Informatik (ITEC), Lehrstuhl für Rechnerarchitektur und Parallelverarbeitung



Persönliches

- Prof. Dr. Wolfgang Karl
 - Professur für Entwurf von Systemen in Hardware/Organisation Innovativer Rechnerarchitekturen,
 - Lehrstuhl für Rechnerarchitektur und Parallelverarbeitung
 - Institut für Technische Informatik der KIT-Fakultät für Informatik, Karlsruher Institut für Technologie
 - Büro: Raum 314.1, Technologiefabrik, Haid-und-Neu-Str. 7
 - Tel.: 0721 608 43771
 - Email: karl@kit.edu
 - Sprechstunde:
 - Dienstag, 16:00 – 17:00 Uhr und nach Vereinbarung

Lehrstuhl

■ **Leitung:**

- Prof. Dr. Wolfgang Karl

■ **Sekretariat (Technologiefabrik, 2. Stock):**

- Frau Amjad
- Frau Wehrmann

■ **Mitarbeiter (Technologiefabrik, 2. Stock):**

- Thomas Becker, M.Sc.
- Dipl.-Inform. Michael Bromberger
- Dipl.-Math. Markus Hoffmann

Organisatorisches

■ Termine:

KW	Vorlesung			Übung	
	Di 09:45-11:15 HSaF	Do 08:00-09:30 Gaede		Di 09:45-11:15 HSaF	Do 14:00-15:30 Gaede
KW17	25.04.2017	27.04.2017			
KW18	02.05.2017				04.05.2017
KW19	09.05.2017	11.05.2017			
KW20	16.05.2017				18.05.2017
KW21	23.05.2017	Feiertag			Feiertag
KW22	30.05.2017	01.06.2017			
KW23		08.06.2017		06.06.2017	
KW24	13.06.2017	Feiertag			Feiertag
KW25		22.06.2017		20.06.2017	
KW26	27.06.2017	29.06.2017			
KW27		06.07.2017		04.07.2017	
KW28	11.07.2017	13.07.2017			
KW29	18.07.2017			20.07.2017	
KW30	25.07.2017				27.07.2017

Organisatorisches

■ **Übungen:**

- Vertiefung des in der Vorlesung behandelten Stoffs an Beispielen und Aufgaben
- Übungsleiter: Thomas Becker
 - Thomas.becker@kit.edu
 - Raum: Technologiefabrik, 2. Stock, 315.1

■ **Klausurtermin:**

- Voraussichtlich 14. August 2017, 11:00 Uhr
- Stoff: Vorlesung und Übung

■ **Informationen zur Vorlesung**

- <http://capp.itec.kit.edu/teaching>

Aufbau der Vorlesung

1. Grundlagen

- Einführung
- Allgemeine Grundlagen des Entwurfs von Rechenanlagen
- Formen des Parallelismus und Klassifizierungen von Rechnerarchitekturen
- Bewertung von Rechensystemen
- Zuverlässigkeit, Verfügbarkeit und Fehlertoleranz

2. Prozessortechniken

- Von RISC zu Superskalar
- Instruction-level parallelism (ILP):
 - Superskalartechniken
 - VLIW, EPIC
- Multithreading

Aufbau der Vorlesung

3. Multiprozessoren

- Allgemeine Grundlagen, Verbindungsnetze, Leistungsfähigkeit
- Speichergekoppelte Multiprozessoren
 - SMP und DSM
 - Speicherkonsistenz und Cache-Kohärenz
- Nachrichtengekoppelte Multiprozessoren
- Chip-Multiprozessoren, Multi-core, Many-core
- Höchstleistungsrechner
- Cluster-Computing, Warehouse-scale Computers

4. Weitere Rechnerstrukturen

- Vektorrechner und Feldrechnerprinzip
- SIMD-Verarbeitung in Mikroprozessoren

5. Ausblick: Aktuelle Forschungsthemen

Hinweis

■ Vom Lehrstuhl angebotene Lehrveranstaltungen:

■ Vorlesungen:

- Rechnerorganisation
- Digitaltechnik und Entwurfsverfahren
- Rechnerstrukturen
- Mikroprozessoren I,
- Mikroprozessoren II
- Heterogene parallele Rechnerstrukturen

} Im Turnus mit Prof. Asfour,
Prof. Hanebeck, Prof. Henkel, Prof. Tahoori

} Regelmäßig im Sommersemester

} Regelmäßig im Wintersemester

■ Praktika:

- Basispraktikum Technische Informatik: Hardware-naher Systementwurf
- Projektorientiertes Software-Praktikum (Parallele Numerik)

■ Seminare

- Ausgewählte Kapitel der Rechnerarchitektur

Literatur

- Dubois, M.; Annavaram, Stenström, P.: Parallel Computer Organization and Design. Cambridge University Press, 2012
- M.Hennessy, J.L., Patterson, D.A.: Computer Architecture: A Quantitative Approach. Morgan Kaufmann, 5.Auflage 2012.
- U. Bringschulte, T. Ungerer: Microcontroller und Mikroprozessoren, Springer, Heidelberg, 2. Auflage 2007
- Theo Ungerer: Parallelrechner und parallele Programmierung, Spektrum-Verlag 1997

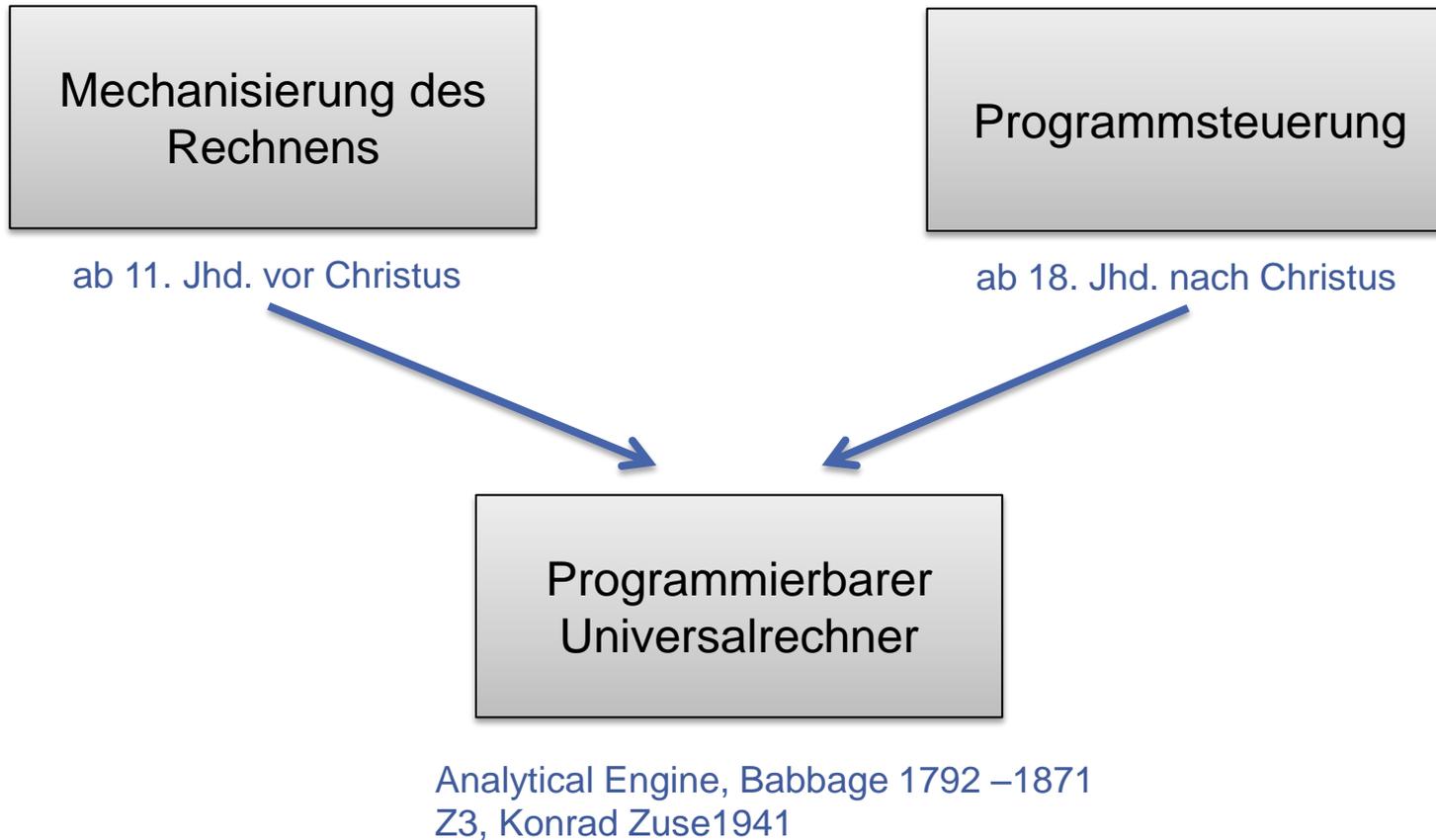
Vorlesung Rechnerstrukturen

Kapitel 1: Grundlagen

■ 1.1 Einführung

Zur Geschichte

Genealogie des programmierbaren Universalrechners



Zur Geschichte

Genealogie des programmierbaren Universalrechners

Mechanisierung des Rechnens

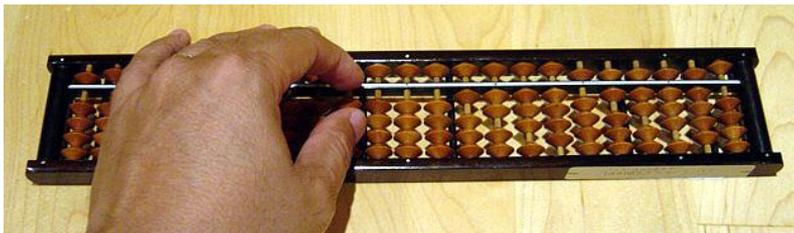
ab 11. Jhd. vor Christus



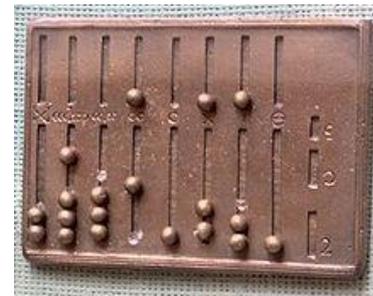
Chinesischer Suan Pan



Russische Stschoty



Japanischer Soroban



Rekonstruktion eines römischen Abakus

Quelle: Wikipedia

Zur Geschichte

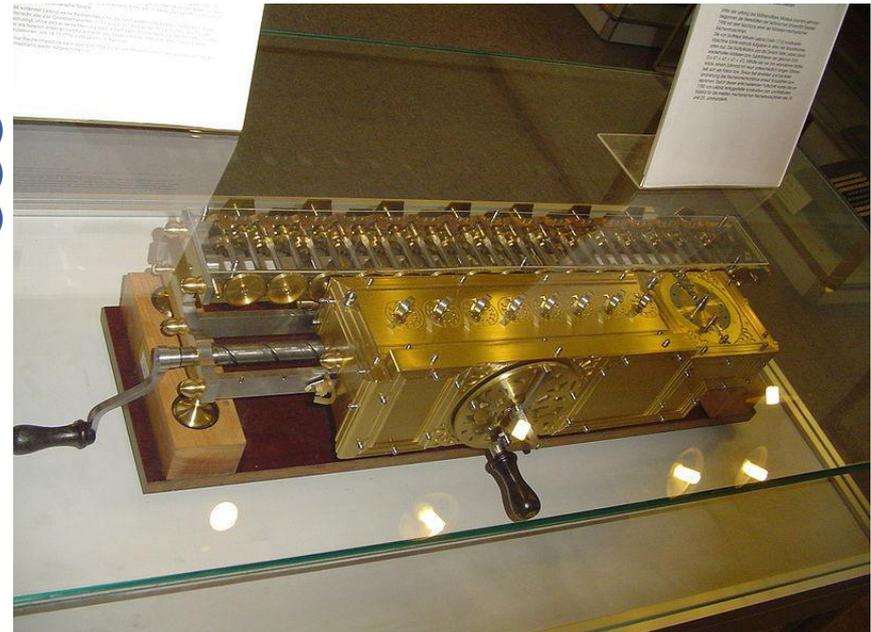
Genealogie des programmierbaren Universalrechners

Mechanisierung des Rechnens

ab 17. Jhd. n. Chr.: Vierspezies Rechenmaschine

Schickard(1592 -1635)
Pascal (1629 -1662)
Leibniz (1646 -1716)

Leibniz: Idee einer auf dem dualen
Zahlensystem basierenden
„Machina Arithmeticae Dyadicae“



Quelle: Wikipedia

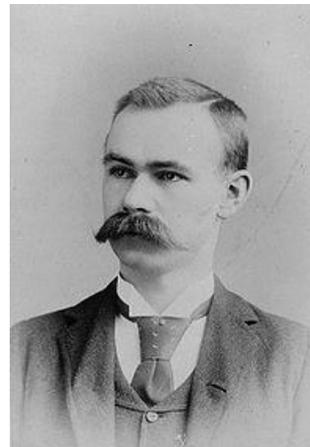
Zur Geschichte

Genealogie des programmierbaren Universalrechners

Programmsteuerung



Ab 18. Jhd. Lochkarten-
gesteuerter Webstuhl
Jacquard (1752 –1834)



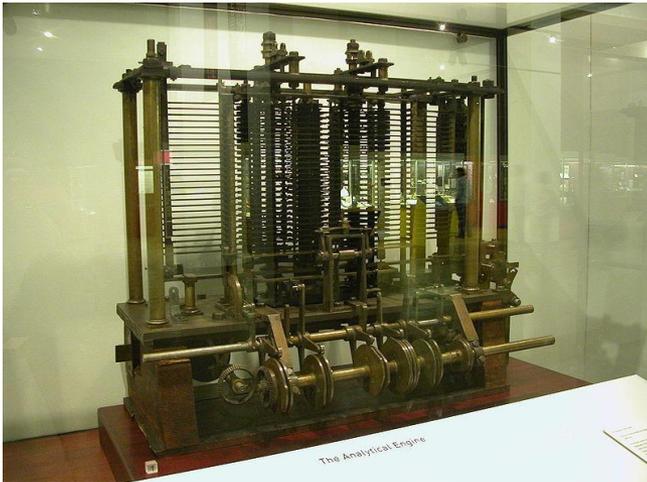
Hollerith (1860 –1929)
1889: Patent auf Tabelliermaschine:
Lochkartenstanzer, -leser, -sortierer



Quelle: Wikipedia

Zur Geschichte

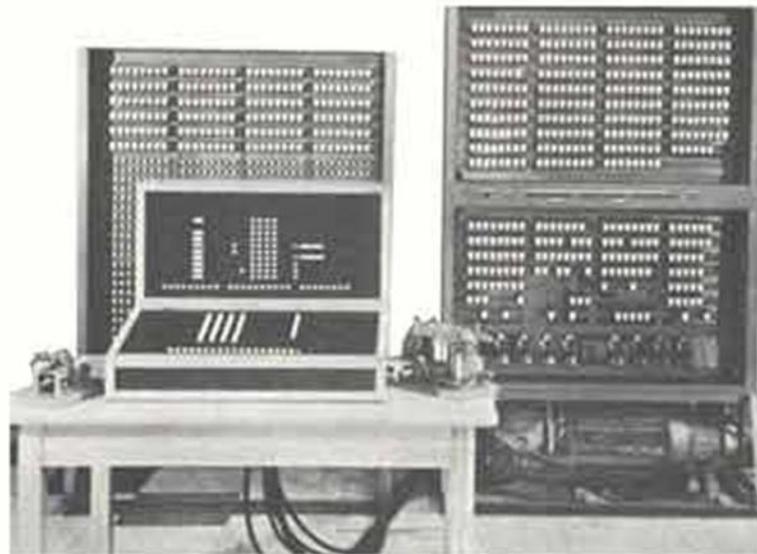
Genealogie des programmierbaren Universalrechners



Quelle: Wikipedia

Versuchsmodell der Analytical Engine
Babbage (1791 -1871)

Programmierbarer
Universalrechner



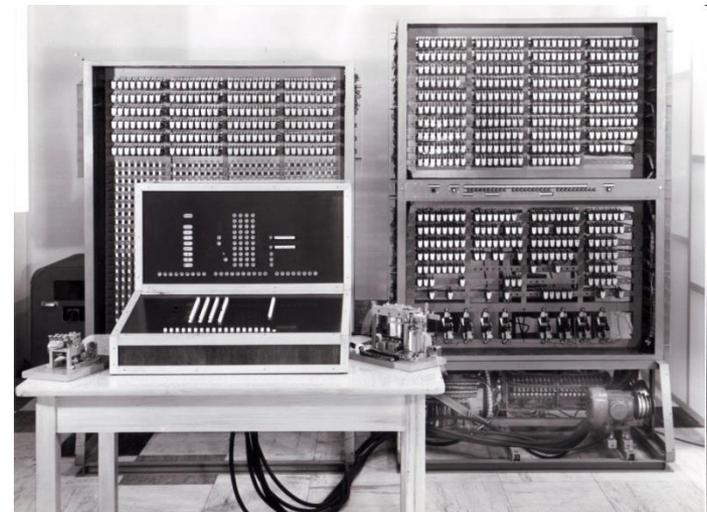
Z3: Konrad Zuse (1910 – 1995): programmierbarer
Rechner (voll funktionsfähig)

Quelle: Horst Zuse

Zur Geschichte

Zuse Z3 (1941):

- „der erste funktionsfähige, frei programmierbare, auf dem binären Zahlensystem (Gleitkommazahlen) und der binären Schaltungstechnik basierende Rechner der Welt.“
- Speicherkapazität:
 - 64 Worte zu je 22 Bit
- 4 Grundrechenarten, Quadratwurzel
 - 1 Addition benötigte 3 Takte
 - 1 Multiplikation: ~ 3s
- Taktfrequenz: 5 Hz
- Rechenleistung: ~ 1 Gleitkommaoperation pro Sekunde

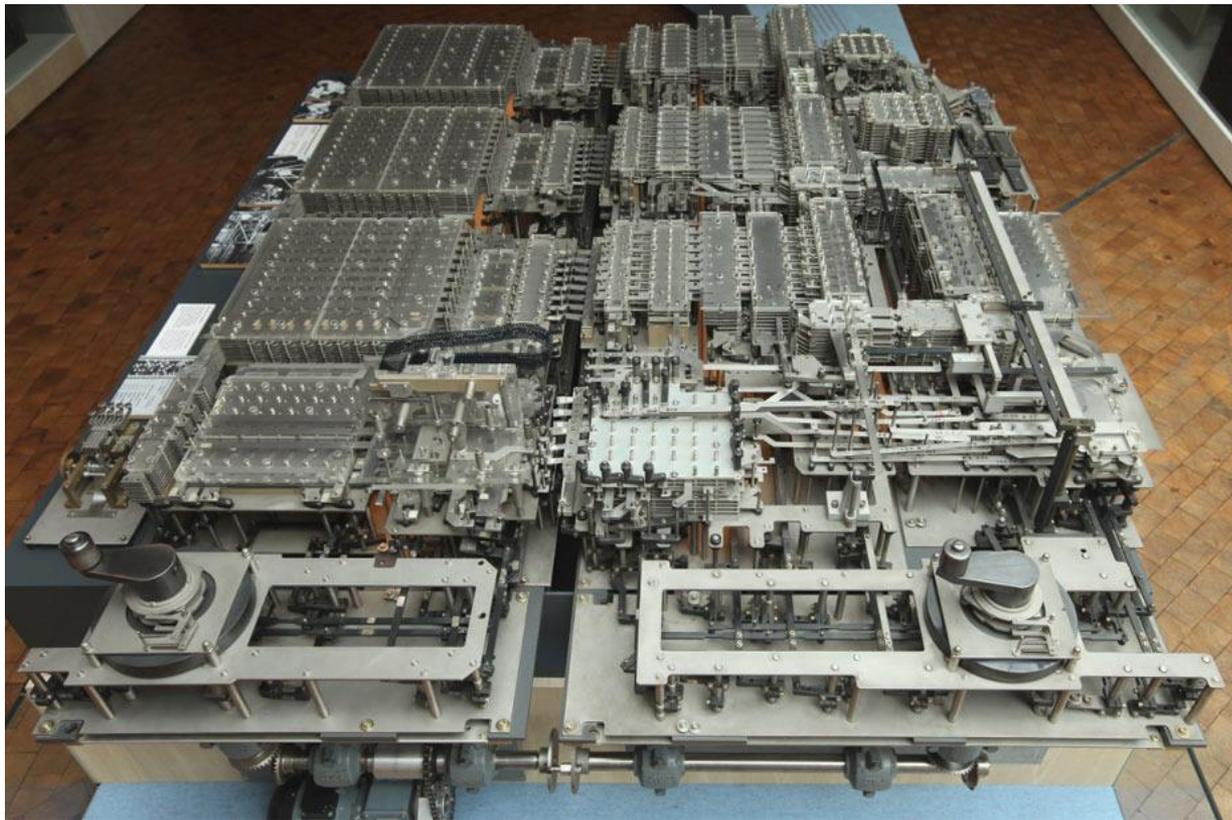


Quelle: H.Zuse, www.zuse.de

Zur Geschichte

Zuse Z1 (1936):

- Die Z1 im Deutschen Technikmuseum Berlin

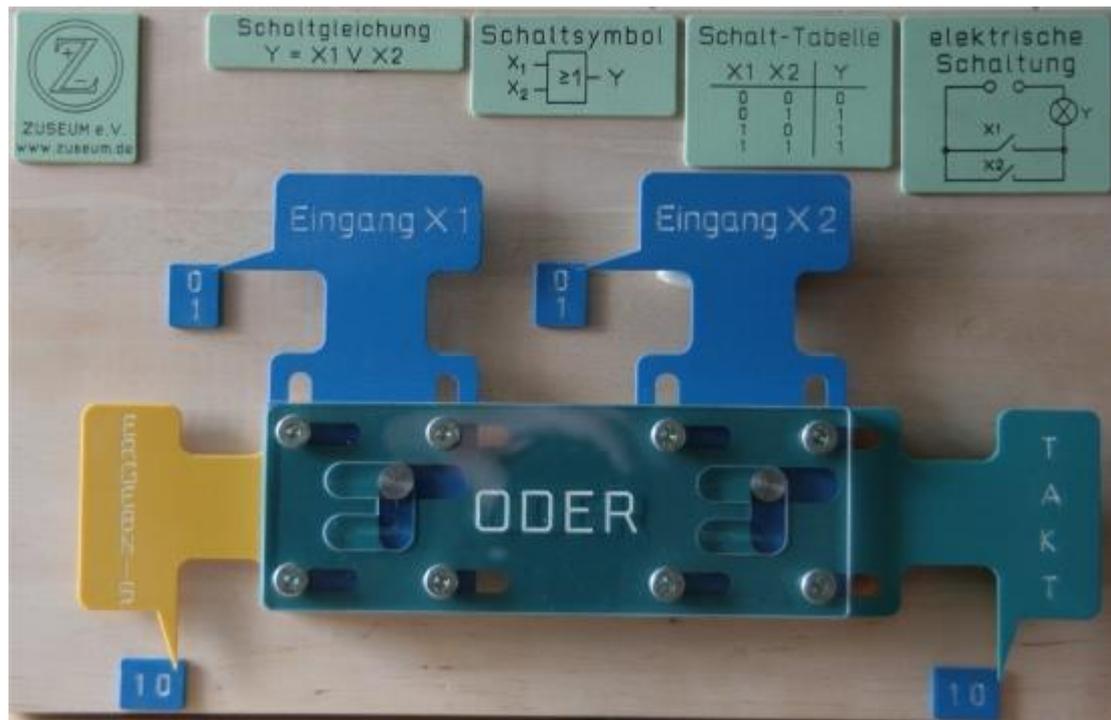


Quelle: Freie Universität Berlin

Zur Geschichte

Zuse Z1 (1936):

- Mechanischer Rechner
 - Logische ODER-Funktion



Quelle: H.Zuse, www.zuse.de

Einführung

Rechnerarchitektur (Disziplin)

- Allgemeine Strukturlehre mit deren Hilfsmittel
- Ingenieurwissenschaftliche Disziplin, die bestehende und zukünftige Rechenanlagen beschreibt, vergleicht, **beurteilt**, verbessert und **entwirft**.
- Betrachtet den Aufbau und die Eigenschaften des Ganzen (Rechenanlage), seiner Teile (Komponenten) und seiner Verbindungen (Globalstruktur, Infrastruktur)

Begriffsklärung

Der Begriff Rechnerarchitektur (Systemsicht)

■ Definition nach Amdahl, Blaauw, Brooks (1967)

„Computer architecture is defined as the attributes and behavior of a computer as seen by a machine-language programmer. This definition includes the instruction set, instruction formats, operation codes, addressing modes, and all registers and memory locations that may be directly manipulated by a machine language programmer.

Implementation is defined as the actual hardware structure, logic design, and data path organization of a particular embodiment of the architecture.“

- Beschreibung der Attribute und des funktionalen Verhaltens eines Systems, wie es von einem Anwender, der in Maschinensprache programmiert, gesehen wird.
- Spezifiziert die konzeptionelle Struktur und das funktionale Verhalten und betrifft nicht Details der Hardware und der technischen Ausführung des Rechners
- Die Definition behandelt nur das äußere Erscheinungsbild des Rechners und klammert die internen Vorgänge ausdrücklich aus.

Begriffsklärung

Rechnerarchitektur (Systemsicht)

- Heutige Sichtweise (Hennessy/Patterson, 2003):
 - Befehlssatzarchitektur (Instruction Set Architecture)
 - Beschreibung der Attribute und des funktionellen Verhaltens eines Rechners
 - Sichtweise des Maschinenprogrammierers
 - Schnittstelle zwischen Hardware und Software
 - Spezifikation der Befehlssatzarchitektur
 - Ausführungsmodell
 - Datenformate, Datentypen
 - Adressierungsarten
 - Befehlsformat und Befehlssatz
 - Logischer Adressraum
 - Unterbrechungssystem
 - ...

Begriffsklärung

Rechnerarchitektur (Systemsicht)

- Heutige Sichtweise (Hennessy/Patterson, 2003):
 - Organisation
 - Höhere Aspekte des Rechnerentwurfs:
 - Entwurf der internen CPU
 - Art und Anzahl der internen Ausführungseinheiten
 - Art und Stufenzahl der Befehlspipeline
 - Grad und Verwendung der Superskalartechnik, VLIW, EPIC, Multithreading
 - Speicher- und Cachesystem
 - Busstruktur
 - ...
 - Hardware
 - Betrifft die speziellen Hardware-Eigenschaften des Rechners, einschließlich dem Logik-Entwurf bis hin zur Verpackungstechnik

Begriffsklärung

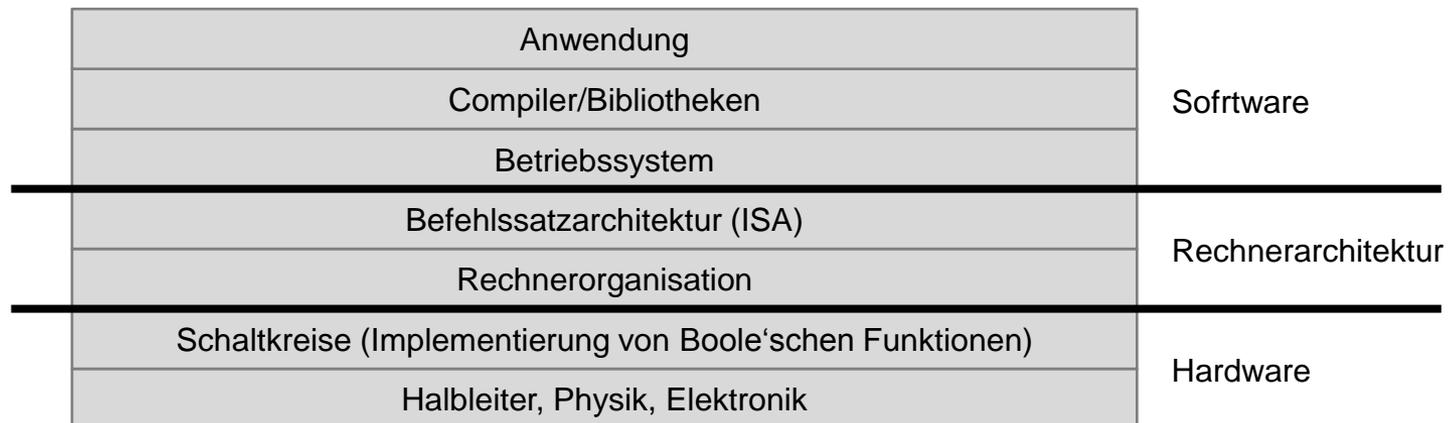
Der Begriff Rechnerarchitektur (Systemsicht)

■ Heutige Sichtweise (Hennessy/Patterson, 2011):

■ Rechnerarchitektur umfasst

- Befehlssatzarchitektur (Instruction Set Architecture)
 - Organisation
 - Hardware
- } Implementierung

■ Heutige Sichtweise (Dubois, Annavaram, Stenström, 2012)



Vorlesung Rechnerstrukturen

- Der Begriff „**Rechnerstrukturen**“:
 - Hardwarestruktur eines Rechners auf einer konzeptionellen Darstellungsebene.
 - Anwendersicht und Operationsprinzip eines Rechners.

Vorlesung Rechnerstrukturen

Kapitel 1: Grundlagen

- 1.1 Einführung, Begriffsklärung
- 1.2 Entwurf von Rechenanlagen - Entwurfsfragen

Einführung

Architektur



Rechnerarchitektur (Disziplin)

Entwurf einer Rechenanlage

- Ingenieurmäßige Aufgabe der Kompromissfindung zwischen
 - Zielsetzungen
 - Einsatzgebiet, Anwendungsbereich, Leistung, Verfügbarkeit ...
 - Randbedingungen
 - Technologie, Größe, Geld, Energieverbrauch, Umwelt,...
 - Gestaltungsgrundsätzen
 - Modularität, Sparsamkeit, Fehlertoleranz ...
 - Anforderungen
 - Kompatibilität, Betriebssystemanforderungen, Standards

Entwurfsfragen

Zielsetzungen

■ Einsatzgebiete, Rechnermärkte

■ Personal Mobile Devices (PMC)



- Drahtlose Geräte mit Multimedia Schnittstelle
- Beispiele: Mobiltelephone (Smart phones), Tablets
- Niedrige Kosten, niedriger Energieverbrauch
- Hohe Leistung für Multimedia-Anwendungen

■ Desktop Computing



- PCs bis Workstations
- Günstiges Preis-/ Leistungsverhältnis
- Ausgewogene Rechenleistung für ein breites Spektrum von Anwendungen, einschließlich interaktiver Anwendungen (Graphik, Video, Audio) oder WEB-Anwendungen

Entwurfsfragen

Zielsetzungen

- Einsatzgebiete, Rechnermärkte
 - Server
 - Rechen- und datenintensive Anwendungen, transaktionsorientierte Anwendungen
 - Hohe Anforderungen an die Verfügbarkeit und Zuverlässigkeit, Energieeffizienz
 - Skalierbarkeit
 - Große Datei-Systeme, Ein-/Ausgabesysteme
 - Abgeschlossene Räume, kostenintensiv

Quelle: LRZ: SuperMUC

Entwurfsfragen

Zielsetzungen

■ Einsatzgebiete, Rechnermärkte

■ Server

■ Höchstleistungsrechner

- Hohe Rechenleistung bezüglich Gleitkommaverarbeitung
- Große, kommunikationsintensive Anwendungen
- Batch-Programme



Quelle: LRZ: SuperMUC

■ Server im kommerziellen Bereich, Warehouse-Scale Computers

- Mainframes, Clusters
- Durchsatzorientierte Anwendungen
- Software as a Service (SaaS)

Entwurfsfragen

Zielsetzungen

■ Einsatzgebiete

■ Eingebettete Systeme (Embedded Systems)

- Mikroprozessorsysteme, eingebettet in Geräten, daher nicht unbedingt sichtbar

- Beispiele: Automobil, Unterhaltungselektronik, Telekommunikation, Haushaltsgeräte, ...

- Rechensysteme sind auf spezielle Aufgabe zugeschnitten

- Hohe Leistungsfähigkeit für spezielle Anwendung

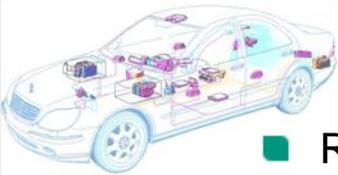
- Spezialprozessoren, Prozessorkerne mit anwendungsspezifischen Komponenten

- Breites Preis-/Leistungsspektrum

- Von einfachen 8-, 16-Bit Microcontrollern bis hin zu komplexen Spezialprozessoren

- Echtzeitanforderungen

- Abwägen der Anforderungen an die Rechenleistung, Speicherbedarf, Kosten, Energieverbrauch



Entwurfsfragen

Zielsetzungen

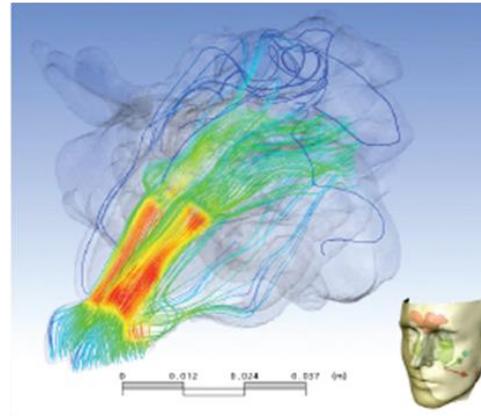
■ Anwendungsbereiche:

■ Technisch-wissenschaftlichen Bereich:

- Hohe Anforderungen an die Rechenleistung, insbesondere Gleitkommaverarbeitung

■ Beispiele:

- Rechnergestützte Simulation
- Strömungsmechanik
- Modellierung der globalen klimatischen Veränderungen
- Struktur von Materialien
- ...
- Medizintechnik



Rechenzeit auf einer HP XC 4000:
(15 TFLOPS): ~4 Tage

Rechenzeit auf einem Rechner
im PFLOPS-Bereich: <40 min

Entwurfsfragen

Zielsetzungen

- Anwendungsbereiche:
 - Kommerzieller Bereich
 - Datenbankanwendungen
 - WEB, Suchmaschinen
 - Optimierung von Geschäftsprozessen, Unterstützung von Geschäftsentscheidungen (Risikoanalyse)
 - Soziale-Netzwerke, Video-Sharing, On-line-Shopping
 - Eingebettete Systeme
 - Verarbeitung digitaler Medien
 - Automatisierungstechnik
 - Automobil
 - Telekommunikation
 - Medizintechnik
 - ...

Entwurfsfragen

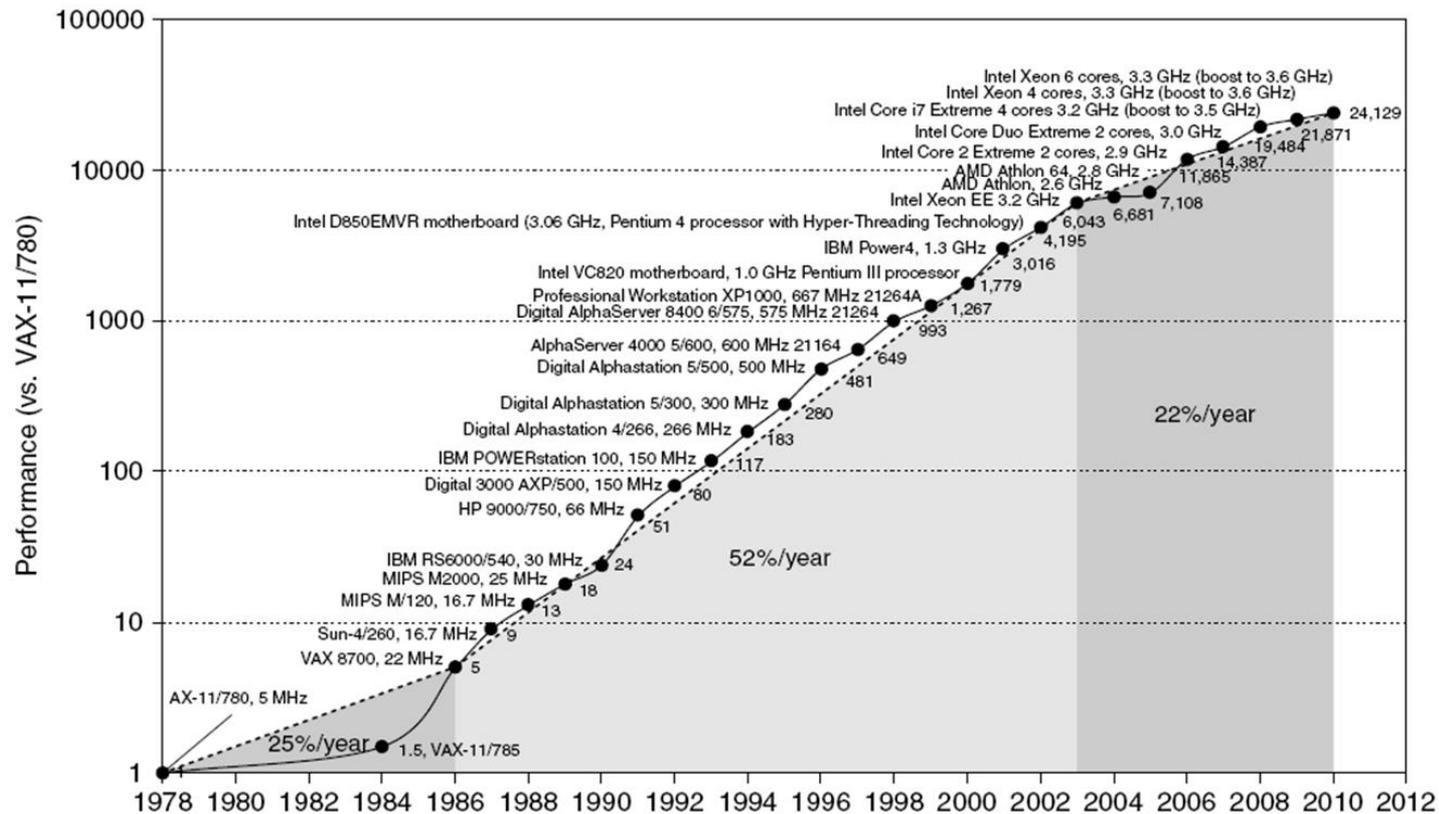
Zielsetzungen

- Rechenleistung
 - Bandbreite, Durchsatz
 - Ausgeführte Arbeit in einem gegebenen Zeitintervall
 - Latenz, Antwortzeit
 - Zeit zwischen dem Start und dem Ende eines Ereignisses
- Bewertung der Leitungsfähigkeit

Entwurfsfragen

Zielsetzungen

■ Rechenleistung



Quelle: J. Hennessy, D. Patterson: Computer Architecture – A Quatative Approach, Morgan Kaufmann Publishers, 5th Ed., 2012

Copyright © 2012, Elsevier Inc. All rights reserved.

Entwurfsfragen

Zielsetzungen

- Zuverlässigkeit
 - Bei Ausfällen von Komponenten muss ein betriebsfähiger Kern bereit sein
 - Techniken der Fehlertoleranz
 - Verwendung redundanter Komponenten
 - Wichtig für sicherheitskritische Anwendungen
 - Wichtig im kommerziellen Bereich
 - Bewertung mittels stochastischer Verfahren
 - Fehlerwahrscheinlichkeit
 - Überlebenswahrscheinlichkeit
 - Mittlere Lebensdauer
 - Ausfallrate
- Verfügbarkeit
 - Wahrscheinlichkeit, ein System zu einem beliebigen Zeitpunkt fehlerfrei anzutreffen

Entwurfsfragen

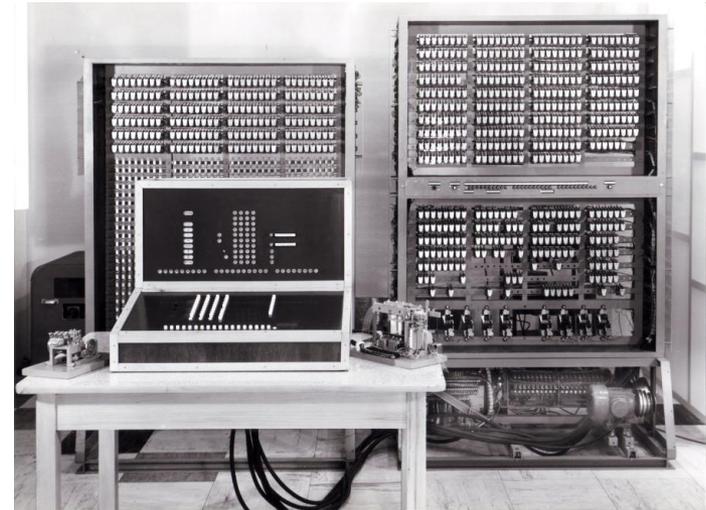
Zielsetzungen

- Energieverbrauch, Leistungsaufnahme
 - Mobile Geräte
 - verfügbare Energiemenge durch Batterien und Akkumulatoren begrenzt
 - möglichst lange mit vorhandener Energie auskommen
 - möglichst wenig Energie soll in Wärme umgesetzt werden, um eine Überhitzung zu vermeiden
 - Green IT
 - Rechnerhersteller bieten „green HW“ an:
 - niedriger Energieverbrauch
 - ökologische Produktion
 - einfaches Recycling

Trends in der Rechnerarchitektur

Entwicklungen im Bereich der Höchstleistungsrechner

- Zuse Z3 (1941):
 - „der erste funktionsfähige, frei programmierbare, auf dem binären Zahlensystem (Gleitkommazahlen) und der binären Schaltungstechnik basierende Rechner der Welt.“
 - Speicherkapazität: 64 Worte zu je 22 Bit
 - 4 Grundrechenarten, Quadratwurzel
 - 1 Addition benötigte 3 Takte
 - 1 Multiplikation: ~ 3s
 - Taktfrequenz: 5 Hz



Quelle: H.Zuse, www.zuse.de

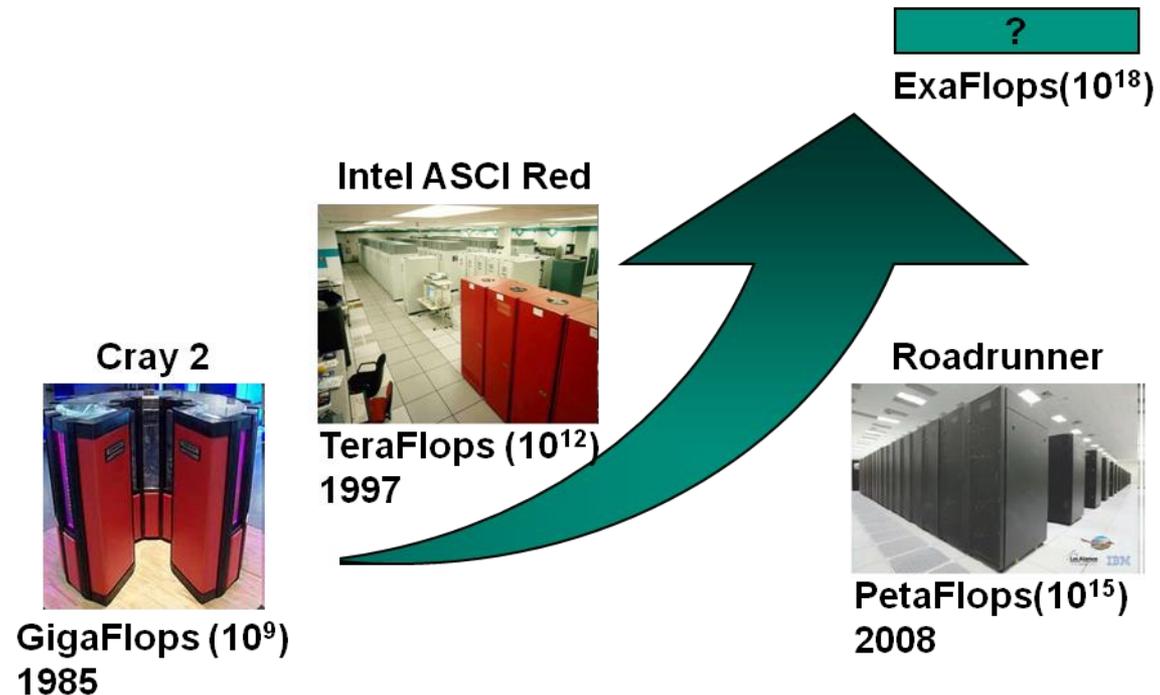
- Rechenleistung: ~ 1 Gleitkommaoperation pro Sekunde

Trends in der Rechnerarchitektur

Entwicklungen im Bereich der Höchstleistungsrechner

- Maßzahl für die Operationsleistung (Gleitkomma-Verarbeitung)

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



Trends in der Rechnerarchitektur

Entwicklungen im Bereich der Höchstleistungsrechner

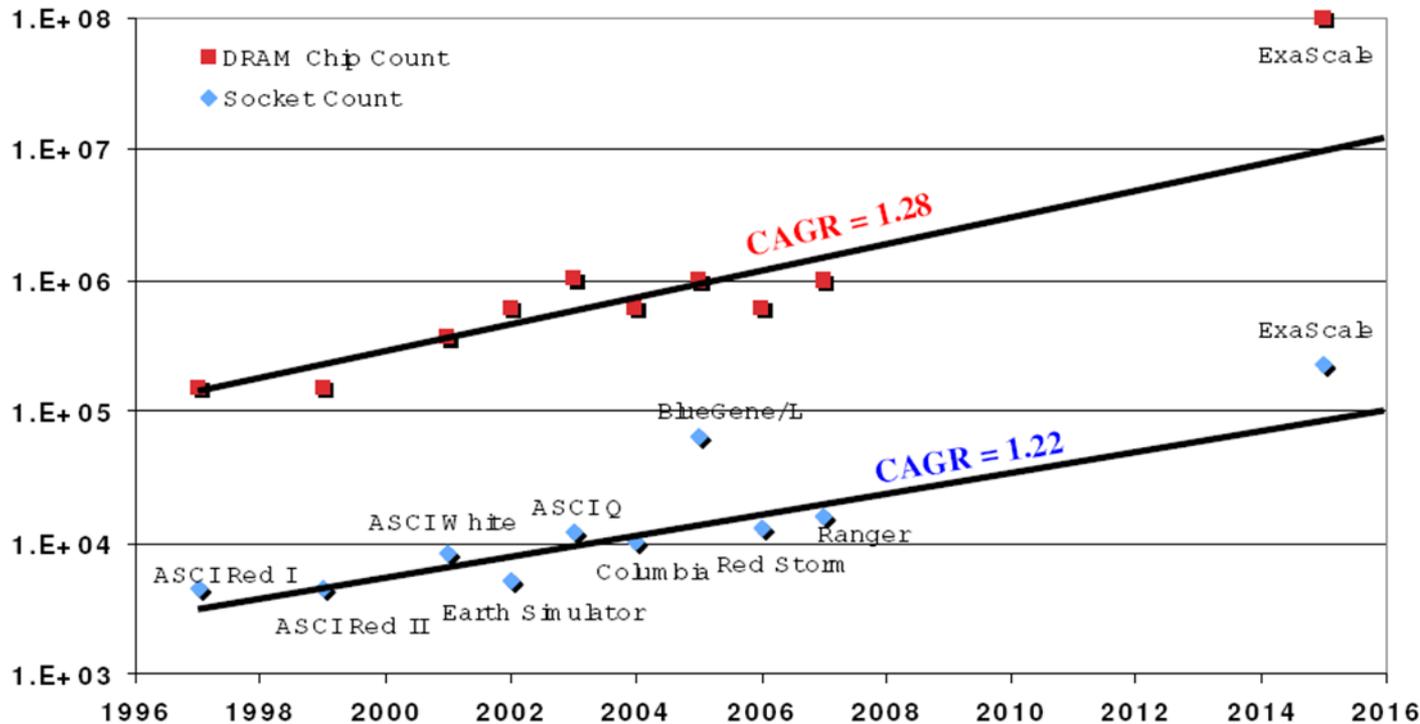
- Herausforderungen bezüglich der Verlustleistung
 - Stand 2010:
 - Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P, NUDT (Nr. 1, TOP 500):
 - 17.8 MW
 - Projektion auf Exascale mit heutiger Technologie:
 - 4 GW !!!!!
 - Verlustleistung für Systeme im ExaFlops-Bereich kann höchstens im Bereich 20 – 40 MW liegen



Trends in der Rechnerarchitektur

Entwicklungen im Bereich der Höchstleistungsrechner

- Herausforderungen bezüglich der Zuverlässigkeit, Verfügbarkeit
 - Entwicklung der Anzahl der Komponenten



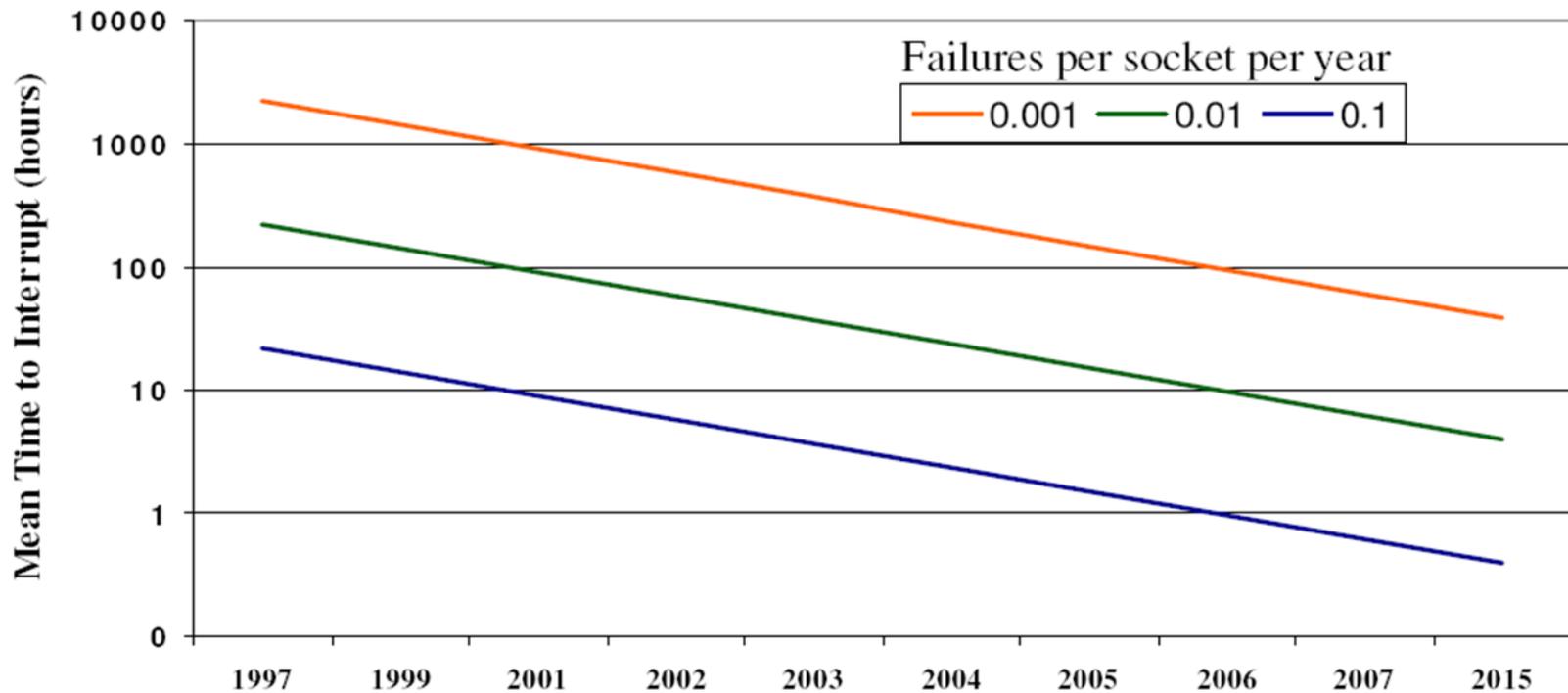
CAGR: Component Annual Growth Rate

* Quelle: ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems

Trends in der Rechnerarchitektur

Entwicklungen im Bereich der Höchstleistungsrechner

- Herausforderungen bezüglich der Zuverlässigkeit, Verfügbarkeit
 - Entwicklung der Fehlerraten als Funktion der Anzahl der Komponenten



* Quelle: ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems

Trends in der Rechnerarchitektur

Entwicklungen im Bereich der Höchstleistungsrechner

- Weltweite Forschungsaktivitäten bezüglich ExaScale-Rechner
 - Herausforderungen
 - Verlustleistung, Energie

 - Hauptspeicher (DRAM), permanenter Speicher
 - Kapazität
 - Zugriffsgeschwindigkeit
 - Mithalten mit der Rechengeschwindigkeit

 - Zuverlässigkeit und Verfügbarkeit

 - Parallelität und Lokalität

Entwurfsfragen

Randbedingungen

■ Technologische Entwicklung



Konrad Zuse: Z1
Mechanische
„Schaltungstechnik“

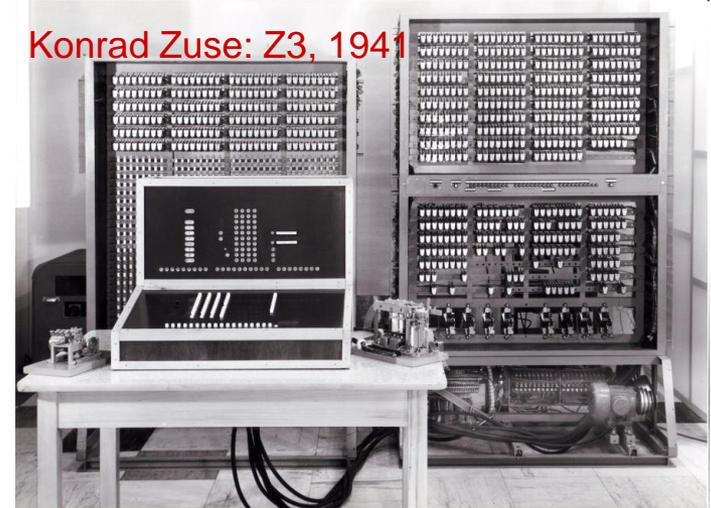
Relais: ab ~1940
Elektromechanisch
Schaltzeit: ms-Bereich



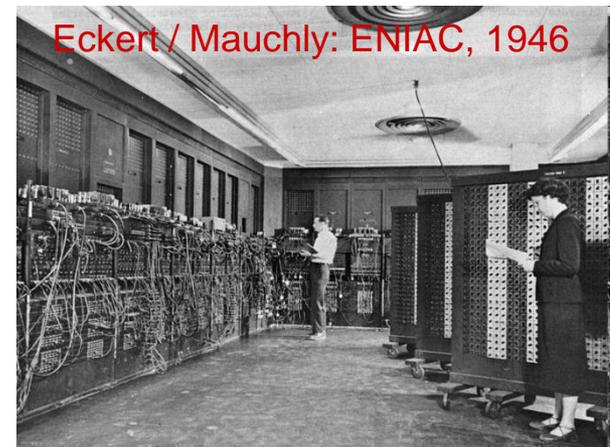
Elektronenröhre: ab ~1945
Elektronisch
Schaltzeit: ms-Bereich



Konrad Zuse: Z3, 1941



Eckert / Mauchly: ENIAC, 1946



Entwurfsfragen

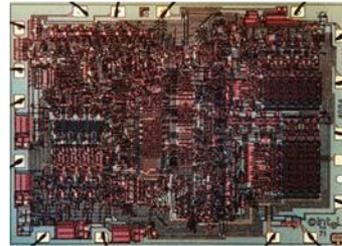
Randbedingungen

■ Technologische Entwicklung

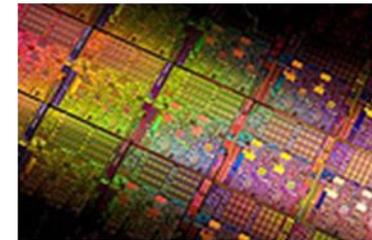
Transistoren: ab ~1947
Halbleitertechnologie
Schaltzeit: ns-Bereich



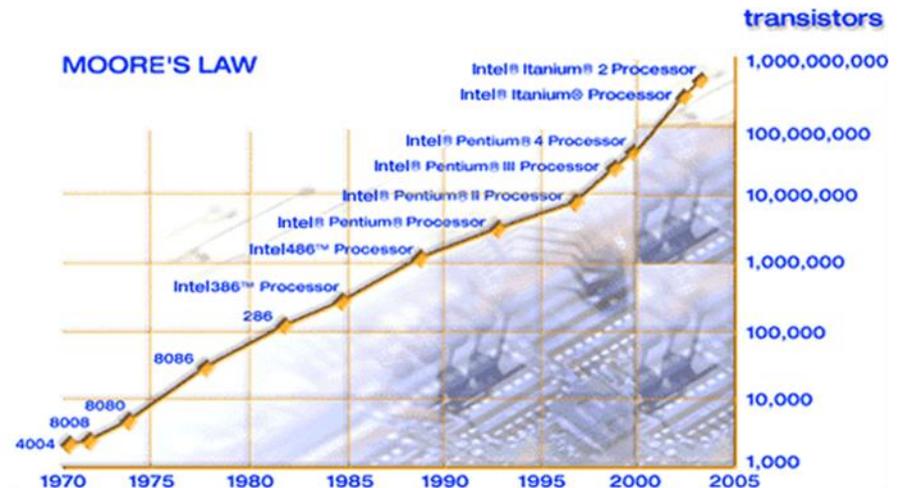
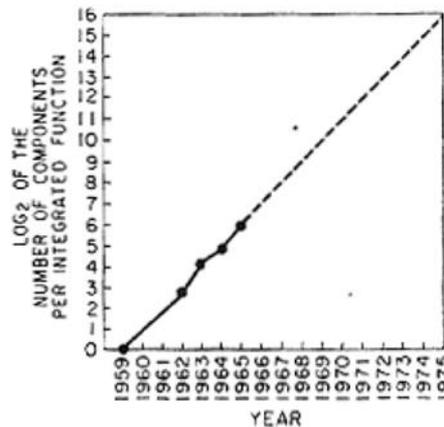
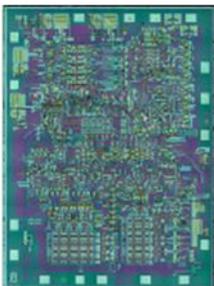
Intel 4004, 1971
~ 2300 Transistoren
Strukturbreite: 10µm



Intel „Nehalem“, 2010
~ 2 Mrd. Transistoren
Strukturbreite: 32nm



Planartechnik: ab ~1958
Integrierte Schaltkreise
Schaltzeit: ms-Bereich



Entwurfsfragen

Randbedingungen

- Technologische Entwicklung
 - Mikrominiaturisierung setzt sich fort (ITRS Roadmap)
 - Verkleinerung der Strukturbreiten
 - Stand 2016: 14nm (Intel-Skylake-Mikroarchitektur)
 - Verbesserung um den Faktor 0,88 pro Jahr in den letzten Jahren
 - Anzahl der Transistoren verdoppelt sich alle 18 Monate
 - Mehrere Milliarden Transistoren auf einem Chip
 - Erhöhung der Integrationsdichte
 - Verbesserung um den Faktor 1,28 pro Jahr
 - Vorhersagen durch die Semiconductor Industry Association (SIA), (<http://www.sia-online.org>)
 - International Technology Roadmap for Semiconductors 2003 Edition (<http://public.itrs.net>)

Entwurfsfragen

Randbedingungen

■ Technologische Entwicklung

■ Mikrominiaturisierung setzt sich fort (ITRS Roadmap)

■ Verkleinerung der Strukturbreiten führt zu erhöhten Problemen der Zuverlässigkeit auf Schaltkreis- / Chip-Ebene

- Erhöhung der elektrischen Felder führt zu negativen elektrischen Effekten

- Erhöhung der Temperatur

- Zeitliche und räumliche Erhöhung der Variabilität führt zu erhöhter Fehleranfälligkeit

 - Herausforderung für Prozess, Schaltungs- und Systementwickler

- Erhöhung der Anfälligkeit bezüglich transienter Fehler

Entwurfsfragen

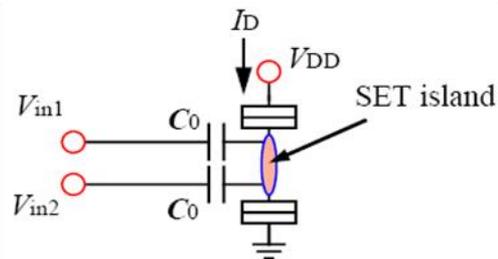
Randbedingungen

■ Technologische Entwicklung

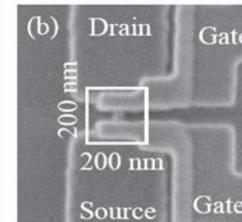
■ Mikrominiaturisierung setzt sich fort (ITRS Roadmap)

- Erforschung zukünftiger Fertigungstechnologien auf der Grundlage von Kohlenstoff, Nanotechnologie

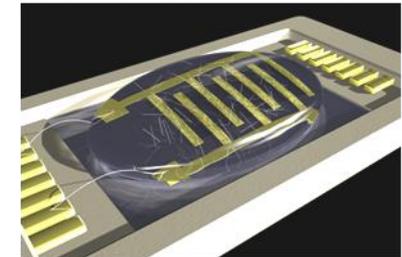
Single Molecule Diode



Single Electron Transistor (SET)



Carbon Nano Tube



Quelle: CFN, KIT

Entwurfsfragen

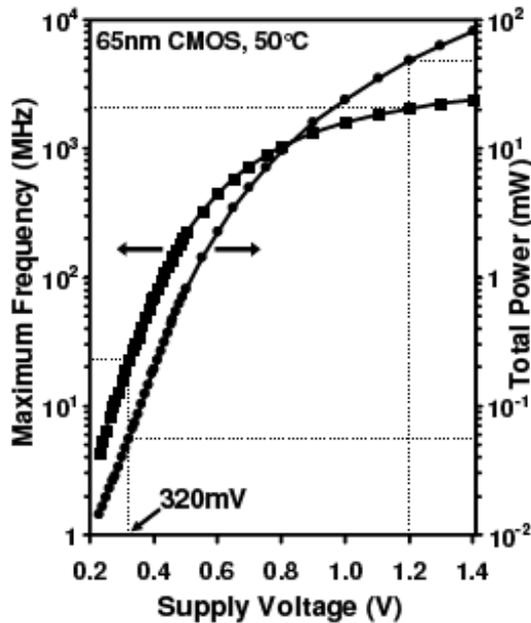
Randbedingungen

■ Elektrische Leistung und Energie

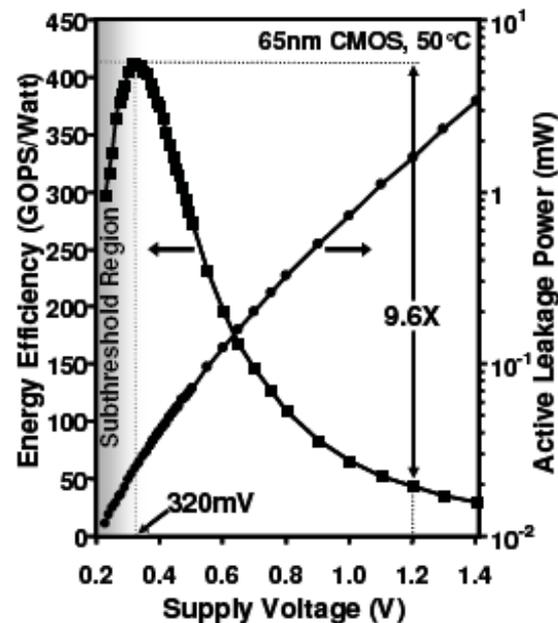
■ Leistungsaufnahme bei CMOS-Schaltungen

■ $P_{\text{total}} = P_{\text{switching}} + P_{\text{shortcircuit}} + P_{\text{static}} + P_{\text{leakage}}$

■ $P_{\text{switching}} = C_{\text{eff}} * V_{\text{dd}}^2 * f$



(a)



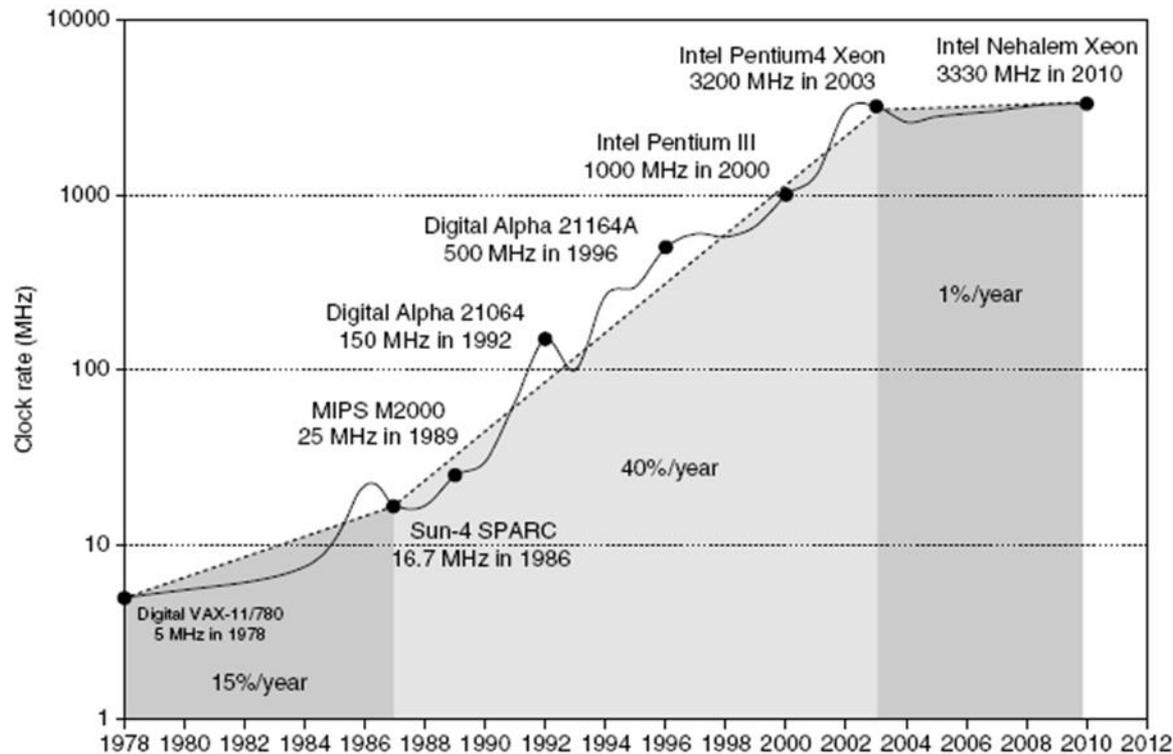
(b)

Quelle: ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems

Entwurfsfragen

Randbedingungen

- Elektrische Leistung und Energie
 - Entwicklung der Taktfrequenz



Quelle: J. Hennessy, D. Patterson: Computer Architecture –
 A Quatative Approach, Morgan Kaufmann Publishers, 5th Ed., 2012

Copyright © 2012, Elsevier Inc. All rights reserved.

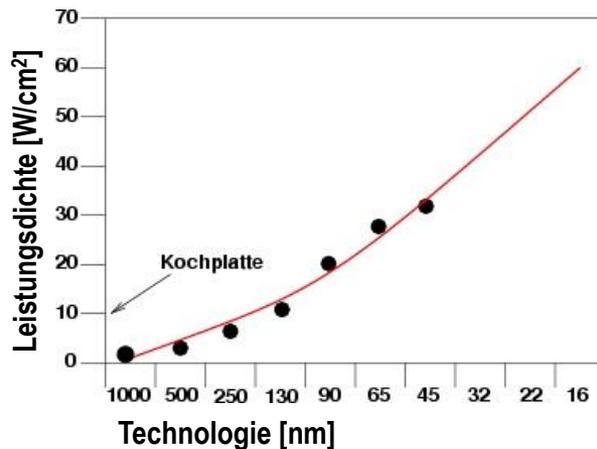
Entwurfsfragen

Randbedingungen

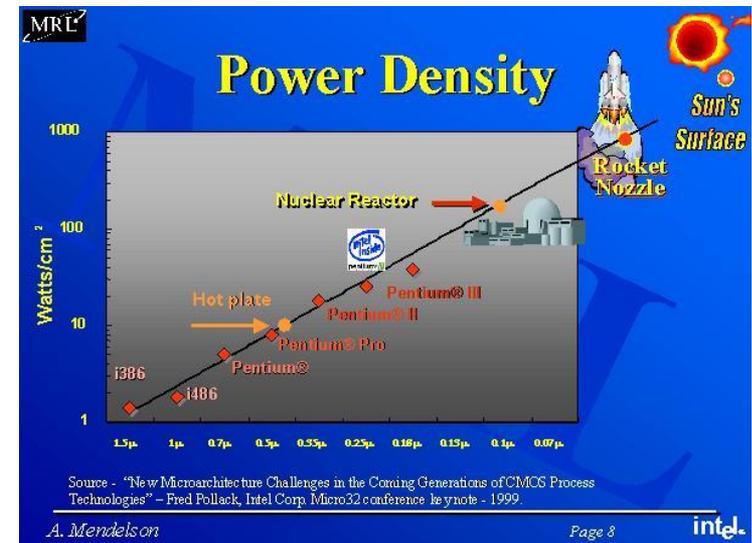
■ Elektrische Leistung und Energie

■ Leistungsdichte

■ Verlustleistung pro Fläche (Watt/cm²)



Problem: Umwandlung in Wärme



Beispiel: 3.3 GHz Intel Core i7: elektr. Leistung 130 W

- Abführung der Wärme von einem 1.5 x 1.5 cm großen Chip
- Grenze, was mit Luft gekühlt werden kann.

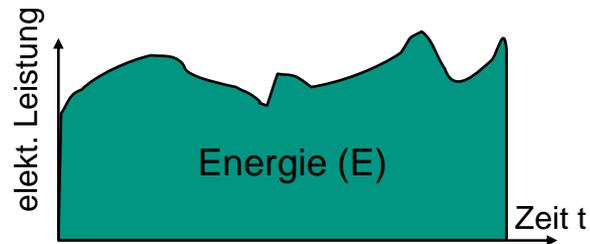
Entwurfsfragen

Randbedingungen

■ Elektrische Leistung und Energie

■ Zusammenhang zwischen Energie E, Leistung P und Zeit t:

■ $P = E / t$ bzw. $E = P * t$



Entwurfsfragen

Randbedingungen

- Technologische Entwicklungen
 - Entwicklung der DRAM-Technologie (Henn./Patt. ,12)
 - Integrationsdichte
 - Verbesserung um etwa 40% - 60% pro Jahr
 - DRAM Kapazität:
 - Steigerung: 25-40%/Jahr (mit fallender Tendenz)
 - Flash Kapazität:
 - Steigerung 50-60%/year
 - 15-20X billiger/Bit als DRAM
 - Magnetspeicher-Technologie:
 - Steigerung der Kapazität um 40%/Jahr
 - 15-25X billiger/Bit als Flash
 - 300-500X billiger/Bit als DRAM

Entwurfsfragen

Randbedingungen

- Kosten eines Rechners / Herstellungskosten
 - Lernkurve
 - Herstellungskosten sinken im Lauf der Zeit
 - Maß: Ausbeute, d.h. der Anteil der hergestellten Komponenten, die die Testphase erfolgreich bestehen

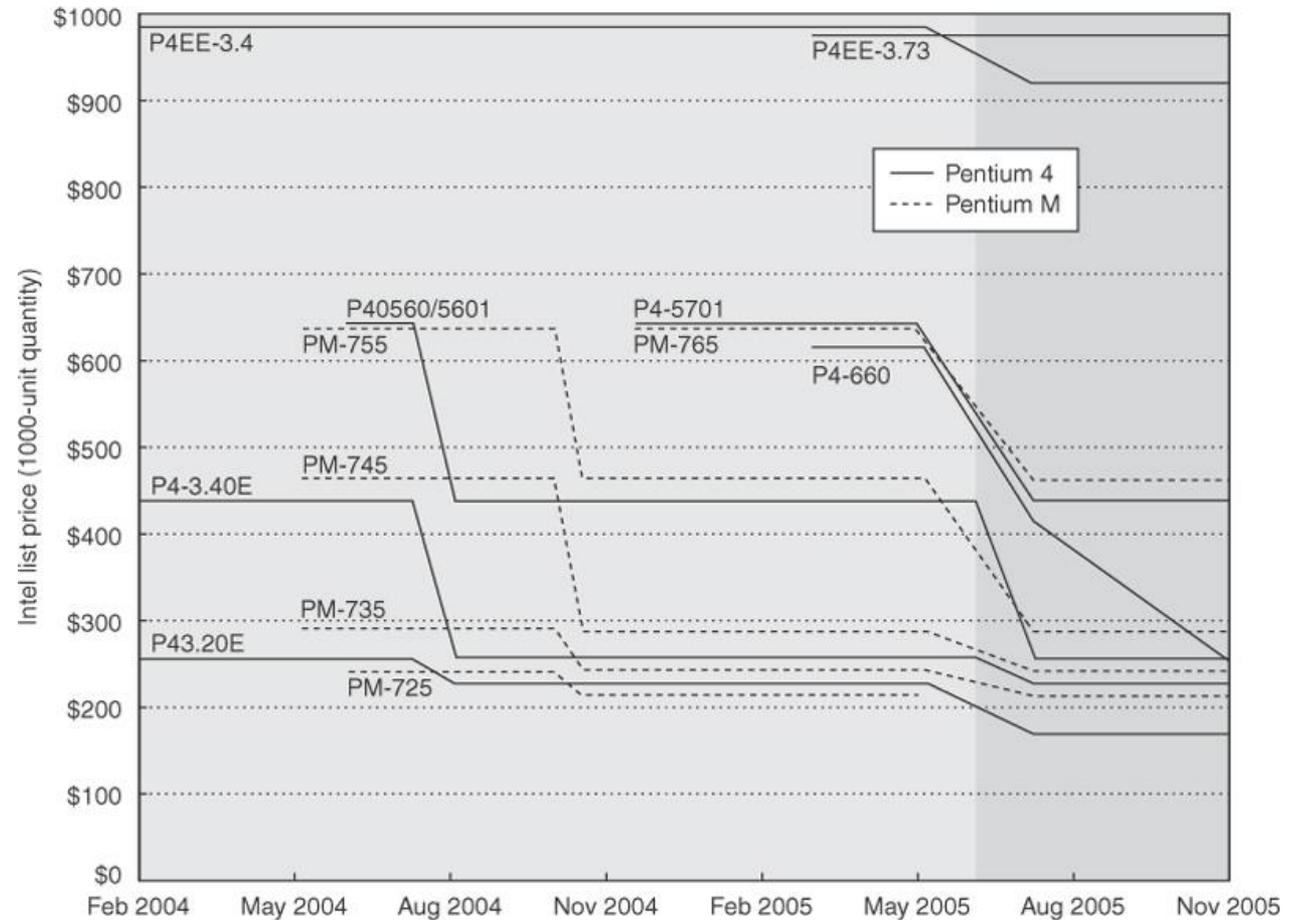
- Preis
 - DRAMs
 - Preis orientiert sich weitgehend an Herstellungskosten
 - Mikroprozessoren
 - Preis abhängig von Herstellungsmenge, Volumen, Wettbewerb

Entwurfsfragen

Randbedingungen

■ Preisentwicklung

Fallstudie: Preisentwicklung bei Pentium 4, Pentium M

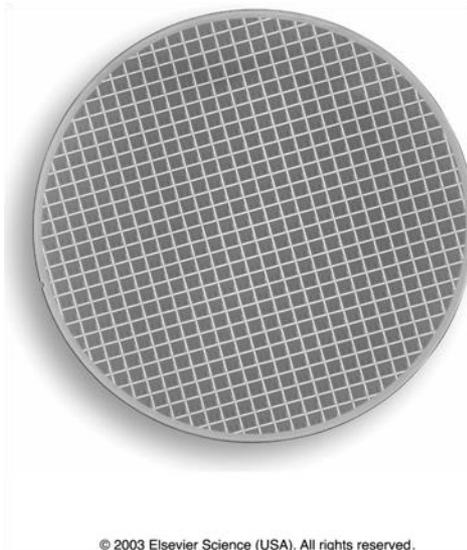


© 2007 Elsevier, Inc. All rights reserved.

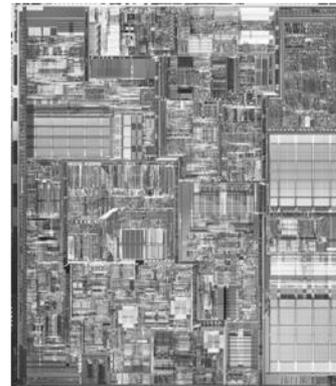
Entwurfsfragen

Randbedingungen

- Herstellungskosten eines integrierten Schaltkreises:
 - Kosten des Dies + Kosten für das Testen des Dies + Kosten für das Packaging und den endgültigen Test in Bezug auf die endgültige Testausbeute



Quelle: Hennessy j., Patterson, D.:
Computer Architecture A Quantative Approach.
Morgan Kaufmann Publ., 3. Auflage, 2003



Quelle: Hennessy j., Patterson, D.:
Computer Architecture A Quantative Approach.
Morgan Kaufmann Publ., 3. Auflage, 2003

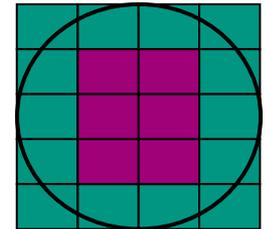


Entwurfsfragen

Randbedingungen

■ Kosten eines integrierten Schaltkreises

$$\text{Kosten des Dies} = \frac{\text{Kosten des Wafers}}{\text{Dies pro Wafer} \times \text{Ausbeute}}$$



$$\text{Anzahl der Dies} = \frac{\pi \times \left(\frac{1}{2} \times \text{Durchmesser des Wafers}\right)^2}{\text{Fläche des Dies}} - \frac{\pi \times \text{Durchmesser des Wafers}}{\sqrt{2} \times \text{Fläche des Dies}}$$

■ Ausbeute (Die Yield)

$$\text{Ausbeute} = \text{Wafer Ausbeute} \times 1 / (1 + \text{Defekte pro Flächeneinheit} \times \text{Die Fläche})^N$$

Entwurfsfragen

Randbedingungen

- Kosten eines integrierten Schaltkreises / Ausbeute (Die Yield)
- Empirisches Modell durch Beobachtung der Ausbeute
 - Annahme:
 - Die Defekte sind zufällig verteilt über den Wafer
 - Die Ausbeute ist umgekehrt proportional zur Komplexität des Herstellungsprozesses
 - Waferausbeute (Wafer yield):
 - berücksichtigt, dass ein Wafer vollständig defekt ist und nicht getestet zu werden braucht
 - Defekte pro Flächeneinheit (defects per unit area):
 - Maß für die zufällig auftretenden Defekte bei der Herstellung
 - In 2010: ~ 0.016-0.057 Defekte pro Quadratzentimeter
 - Hängt von der Reife des Herstellungsprozesses ab
 - α : Maß für die Komplexität des Herstellungsprozesses
 - In 2010: Näherungswert für $N = 11.5-15.5$ (40 nm)

Entwurfsfragen

Randbedingungen

- Kosten eines integrierten Schaltkreises / Ausbeute (Die Yield)
 - Beispiel: Wie ist die Ausbeute für Dies mit einer Seitenlänge von $1,5\text{ cm}$ und $1,0\text{ cm}$ unter der Annahme der Defektdichte von $0,4$ pro cm^2 und $\alpha = 4$

$$\text{Die yield} = \left(1 + \frac{0,4 \times 2,25}{4,0}\right)^{-4} = 0,44$$

$$\text{Die yield} = \left(1 + \frac{0,4 \times 1,00}{4,0}\right)^{-4} = 0,68$$

Entwurfsfragen

Randbedingungen

■ Kosten eines integrierten Schaltkreises

■ Fazit:

- Der Herstellungsprozess diktiert die Kosten für den Wafer, die Wafer Ausbeute und die Defekte pro Flächeneinheit
- Die Kosten pro Chip wachsen ungefähr mit der Quadratwurzel der Chipfläche. Der Entwickler hat einen Einfluss auf die Chipfläche und daher auf die Kosten, je nachdem welche Funktionen auf dem Chip integriert werden und durch die Anzahl der I/O Pins