

DIE PC-ARCHITEKTUR

Computer sind strikt in verschiedene Bereiche eingeteilt. Das hat zwar historische Gründe, aber eben nicht ausschließlich. Von Alexander Köpf

Gaming-Rechner beziehungsweise Personal Computer im Allgemeinen sind im Grunde schon seit jeher gleich aufgebaut. Das schlagende Herz ist die zentrale Recheneinheit – der Prozessor. Drum herum sind verschiedene andere Komponenten aufgereiht:

Arbeitsspeicher, Grafikkarte, Soundkarte, sämtliche Eingabegeräte und vieles mehr. Doch warum ist das eigentlich der Fall? Wieso sind die verschiedenen Komponenten so klar voneinander getrennt?

Die Geschichte beginnt in Deutschland

Um die Frage von oben zu beantworten, müssen wir weit in die Geschichte des Computers zurückblicken. Konkret ins Jahr 1937. Der deutsche Bauingenieur und Erfinder Konrad Ernst Otto Zuse (1910 bis 1995) stellte in diesem Jahr den mechanischen Rechner »Z1« fertig. Die Z1 gilt als erster Vorläufer des modernen Computers, da sie über ein frei programmierbares Rechenwerk basierend auf binären Zahlen (0 und 1) und zusätzlich bereits über ein Eingabe-/Ausgabewerk und ein Speicherwerk verfügte. Zum Programmieren wurden Lochkarten verwendet. Da es sich um ein mechanisches Rechenwerk handelte, verhalten sich die einzelnen Elemente jedoch immer wieder, weshalb die Z1 als unzuverlässig gilt.

1941 stellte Zuse zusammen mit Helmut Schreyer dann die »Z3« fertig. Sie war der erste funktionsfähige Digitalrechner und gilt als auch als erster Universalrechner der Welt. Das bedeutet, sie wurde nicht für einen speziellen Zweck gebaut, sondern konnte viele unterschiedliche Berechnungen ausführen.

Von-Neumann-Architektur

Im Jahr 1945 beschrieb der österreichisch-ungarische und später US-amerikanische Mathematiker John von Neumann (1903 bis 1957) eine konzeptionelle Rechenmaschine, die in Form des EDVAC Ende der 1940er-Jahre umgesetzt werden sollte. Dieses Konzept bildet die formale Grundlage, auf

der dann auch praktisch alle modernen PCs basieren – die Von-Neumann-Architektur.

Andere Rechner aus dieser Zeit waren fest an Programme gebunden. Das heißt, sie waren entweder ausschließlich so verdrahtet, dass sie nur eine bestimmte Aufgabe erfüllen konnten, oder wurden mithilfe von Lochkarten gesteuert, die ihrerseits nur eine bestimmte Aufgabe ausführen konnten. Übrigens: Lochkartensysteme wurden schon ab Mitte des 18. Jahrhunderts verwendet, etwa um Webstühle zu steuern.

Konvergente Entwicklung

Ob und gegebenenfalls inwieweit von Neumann sich durch die Arbeiten Zuses inspirieren ließ, ist unklar. Es gilt aber als unwahrscheinlich. Die Entwicklung wird historisch daher als konvergent erachtet. Sie erfolgte demnach unabhängig voneinander.

Das ist gerade deshalb so spannend, weil es darauf hindeutet, dass es im Grunde nicht viele andere Möglichkeiten gibt, eine Rechenmaschine zu konstruieren. Die strikte Aufteilung in Rechenwerk, Steuereinheit, Ein- und Ausgabe, Speicher und Datenbus ist daher eine Notwendigkeit und keine zufällige Anordnung. Dennoch gibt es zumindest in Teilbereichen andere Ansätze. Diese finden sich in der sogenannten Flynnschen Klassifikation:

- **SISD (Single Instruction, Single Data):** Traditionell werden darunter Einkernprozessoren verstanden, die ihre Daten sequenziell abarbeiten, also eine Berechnung zu je einem Datensatz. Aber auch heutige Computer mit Mehrkernprozessoren zählen zu dieser Kategorie, da jeder Kern respektive Thread Daten immer noch nacheinander abarbeitet.
- **SIMD (Single Instruction, Multiple Data):** Hierbei handelt es sich um sogenannte Vektorprozessoren, die eine Rechenoperation gleichzeitig auf vielen Daten ausführen. Die Berechnung von Spielen in Grafikprozessoren beispielsweise basieren im Wesentlichen auf Matrizenoperationen, die wiederum von Vektorprozessoren realisiert werden können. GPUs sind diesen in ihrer Funktionsweise daher sehr ähnlich.
- **MISD (Multiple Instruction, Single Data):** Ob es diese Architektur in der Realität überhaupt gibt, ist in Fachkreisen umstritten. Wenn überhaupt, findet sie sich jedoch in leistungsfähigen Supercomputern.
- **MIMD (Multiple Instruction, Multiple Data):** Prozessoren auf Basis von MIMD führen mehrere Rechenoperationen gleichzeitig auf verschiedenen Daten aus. Jeder Prozessor hat dabei Zugriff auf die Daten aller anderen Prozessoren. Auch dieses Konzept wird in Großrechnern realisiert.



Im deutschen Technikmuseum in Berlin steht ein Nachbau der Z1.
Foto: Mike Peel (www.mikepeel.net), CC BY-SA 4.0,
<https://commons.wikimedia.org/w/index.php?curid=58211423>

In der Praxis sehen wir Mischformen

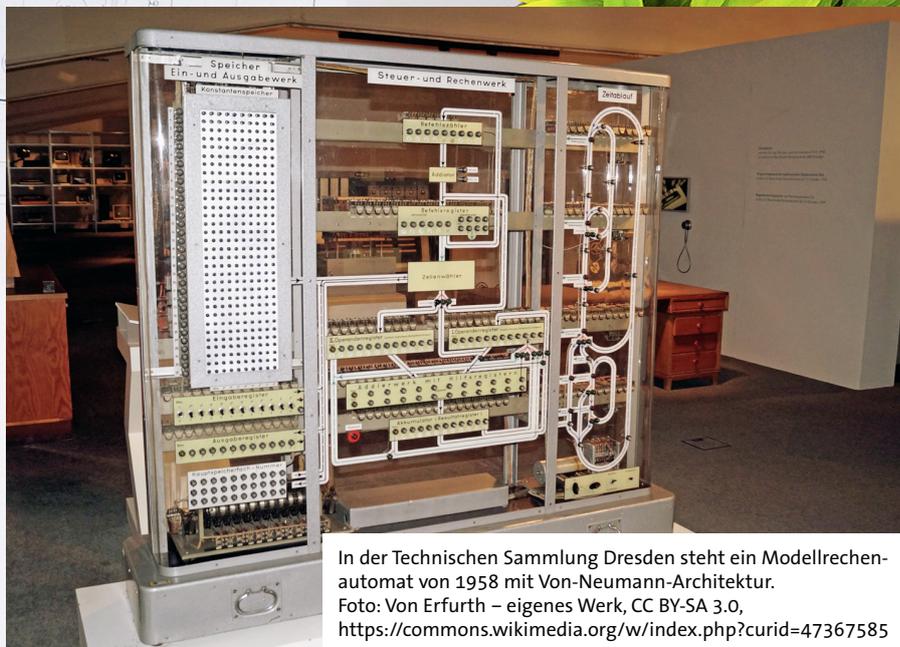
Ganz so trennscharf sind die einzelnen Architekturkonzepte in der Praxis dann aber doch nicht. So fließen Aspekte von SIMD auch in moderne Gaming-Prozessoren ein, die an sich unter die SISD-Architektur fallen – dazu im nächsten Absatz mehr. Prozessoren auf Basis der Von-Neumann-Architektur zählen ebenfalls zu den SISD-Rechnern.

Von-Neumann-Architektur hört man heutzutage jedoch eher sehr selten. Vielmehr ist die Rede von »x86«. Die x86-Architektur fußt wiederum auf von Neumanns Konzept und wurde als »8086« 1978 von Intel entwickelt. Die »80« fiel später einfach weg und wurde durch ein »x« ersetzt. x86 beinhaltet dabei eine ganze Reihe Befehlsätze, die auch SIMD-Elemente beinhalten. Das sind zum Beispiel 3DNow, SSE und AVX.

Harvard-Architektur

Neben der Von-Neumann-Architektur gibt es auch noch die sogenannte Harvard-Architektur. Die beiden Ansätze sind sich grundlegend recht ähnlich, unterscheiden sich jedoch in einem wichtigen Punkt: Von Neumanns Ansatz differenziert nicht zwischen Befehlen und Daten. Beides wird im selben Speicher hinterlegt. Die Hardware-Architektur beispielsweise in Form des »Mark I« aus dem Jahr 1944 verwendet unterschiedliche Speicher für Befehle und Daten. Der Vorteil des Konzepts besteht darin, dass diese gleichzeitig und nicht in zwei aufeinander folgenden Buszyklen geladen werden müssen.

Moderne Prozessoren basieren zwar grundlegend auf den Konzepten von Neumanns, nutzen jedoch auch Elemente der Harvard-Architektur. So finden sich innerhalb von Prozessoren verschiedene Cache-Speicher, die unabhängig vom Datenbus an die Prozessorkerne angebunden sind.



In der Technischen Sammlung Dresden steht ein Modellrechenautomat von 1958 mit Von-Neumann-Architektur. Foto: Von Erfurth – eigenes Werk, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=47367585>

Das heißt, auf Programme respektive Befehle und Daten können sie getrennt voneinander zugreifen, auch wenn sie sich letztlich im selben Speicher (Arbeitsspeicher) befinden.

Ein Computer ist praktisch ein großer Prozessor

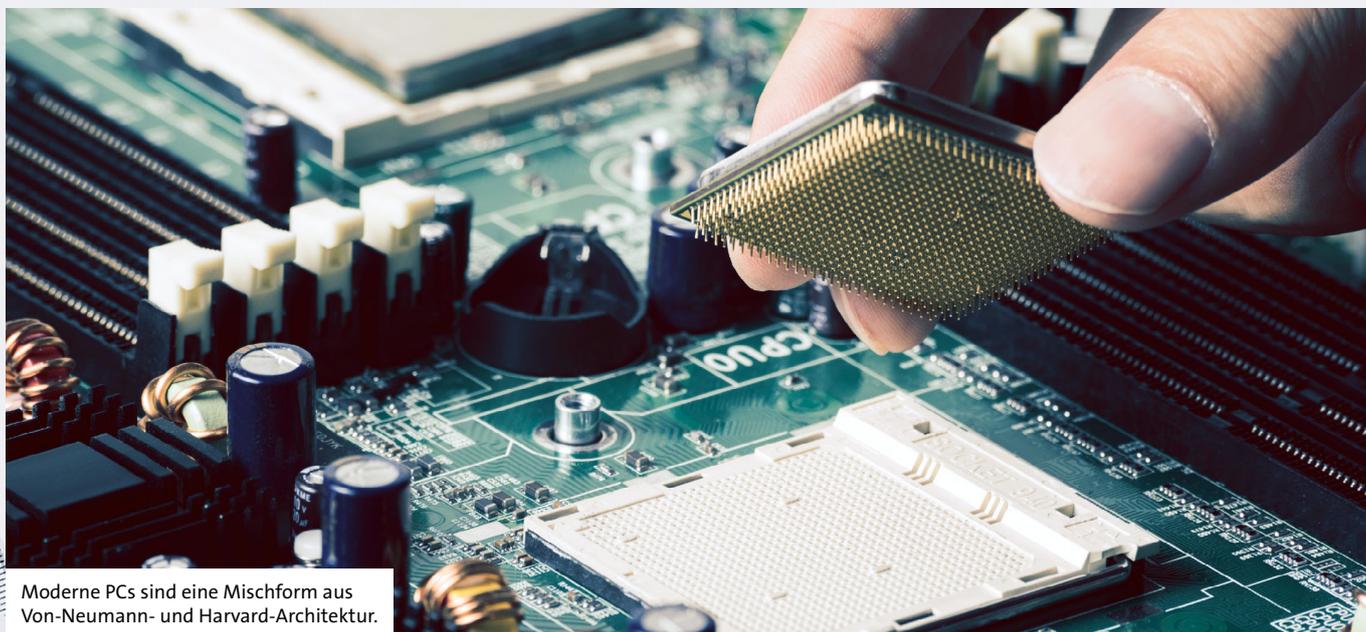
Die heutige Einteilung in Prozessor, Grafikkarte, Arbeitsspeicher und so weiter ist also eine direkte Folge sehr früher Entwicklungen. Im Grunde könnt ihr euch das so vorstellen: Die ersten Computer waren riesig, wenn man so will, waren sie ein einziger großer Prozessor. Daher werden die Begriffe Prozessor, Recheneinheit oder Rechner oft auch synonym verwendet.

Im Laufe der Zeit sind die einzelnen Komponenten massiv geschrumpft. Manche Teile wurden von dem, was wir heute unter Prozessor verstehen, jedoch auf die Platine (das Mainboard) ausgelagert. Der Arbeitsspeicher etwa oder auch der sogenannte PCH (Plattform Controller Hub), der verschiede

Peripheriegeräte steuert. Im Prozessor selbst befindet sich aber noch ein Ein- und Ausgabewerk in Form eines IO-Chips.

Ein IO-Chip wiederum ist ein Co-Prozessor, der die CPU bei ihren Berechnungen unterstützt oder ergänzt. Eine Grafikkarte beziehungsweise ein Grafikchip ist im Grunde nichts anderes, abgesehen davon, dass dedizierte Grafikkarten einen eigenständigen Speicher in Form von Grafikspeicher nutzen.

Abschließende Frage an euch: Habt ihr das alles schon gewusst oder konnten wir euch noch etwas Neues erzählen? ★



Moderne PCs sind eine Mischform aus Von-Neumann- und Harvard-Architektur.