

Kapitel 1

Aufbau und Funktion eines Personal Computers

In diesem Kapitel werden wir in den Aufbau und die Funktion eines Personal Computers (PCs) einführen. Der Kern eines PCs besteht aus der Hauptplatine, auch *Motherboard* genannt, auf dem der Prozessor, Speicher, Ein-/Ausgabebausteine und –schnittstellen untergebracht sind. Wir beschäftigen uns ausführlich mit der Hauptplatine und den wichtigsten Komponenten, die auf ihr zu finden sind: das sind der *Mikroprozessor* als „Gehirn“ eines Mikrorechners und der *Chipsatz*, der die Steuerung der übrigen Komponenten – insbesondere des Hauptspeichers und der Peripheriemodule – vornimmt und sie mit dem Prozessor verbindet. Dabei beschränken wir uns auf Platinen und Bausteine, die für den Einsatz in den sog. *Desktop*-PCs vorgesehen sind, also PCs, die stationär im Bereich eines Schreibtisches Verwendung finden. Auf Exemplare, die ihren Einsatz in mobilen Systemen finden, können wir hier aus Platzgründen nicht eingehen. Sie werden erst im Kapitel 7 behandelt. Zum Abschluß des Kapitels 1 werden wir uns mit dem Hauptspeicher auseinandersetzen und dazu die gebräuchlichen Speichermodule und ihre wichtigsten Parametern beschreiben.

Doch zunächst beginnen wir mit einer kurzen Darstellung aktueller Computersysteme, mit den verschiedenen Arten von Computern und Entwicklungstrends.

1.1 Einführung

1.1.1 Aktuelle Computersysteme

In diesem Kapitel soll ein kurzer Überblick über aktuelle Computersysteme gegeben werden. Zunächst stellen wir die verschiedenen Arten von Computern vor. Dann betrachten wir am Beispiel von *Desktop*-Systemen deren internen Aufbau, der vor allem durch den Chipsatz geprägt wird. Danach werden die

aktuellen Desktop-Prozessoren der beiden führenden Hersteller AMD und Intel vorgestellt und miteinander verglichen. Im Weiteren beschreiben wir die Funktionsprinzipien der aktuellen Speichermodule sowie Ein- und Ausgabeschnittstellen. Schließlich gehen wir auch auf die Bedeutung von Graphikadaptern ein und geben einen Ausblick auf die künftige Entwicklung.

Die Entwicklung neuer Prozessorarchitekturen und Computersysteme ist rasant. Die Chiphersteller vermelden fast täglich neue technologische und architektonische Verbesserungen ihrer Produkte. Daher fällt es natürlich auch schwer, einen aktuellen Schnappschuss der Entwicklung wiederzugeben – zumal dieser dann nach kurzer Zeit wieder veraltet ist. Trotzdem wollen wir im Folgenden versuchen, den aktuellen Stand zu erfassen.

1.1.2 Arten von Computern

Obwohl es uns meist nicht bewusst ist, sind wir heutzutage von einer Vielzahl verschiedenster Computersysteme umgeben. Die meisten Computer, die wir täglich nutzen, sind nämlich in Gebrauchsgegenständen eingebaut und führen dort Spezialaufgaben aus. So bietet uns beispielsweise ein modernes Mobiltelefon („Handy“) die Möglichkeit, Telefonnummern zu verwalten, elektronische Textnachrichten (SMS) zu versenden, Musik abzuspielen oder sogar Bilder oder Filme aufzunehmen. Ähnliche Spezialcomputer findet man in Geräten der Unterhaltungselektronik (z.B. CD-, DVD-, Video-Recordern, Satelliten-TV-Empfängern), Haushaltstechnik (z.B. Wasch- und Spülmaschinen, Trockner, Mikrowelle), Kommunikationstechnik (z.B. Telefon- und Fax-Geräte) und auch immer mehr in der KFZ-Technik (z.B. intelligentes Motormanagement, Antiblockier- und Stabilisierungssysteme). Diese Spezialcomputer oder so genannten eingebetteten Systeme (*Embedded Systems*) werden als Bestandteile größerer Systeme kaum als Computer wahrgenommen. Sie müssen jedoch ein weites Leistungsspektrum abdecken und insbesondere bei Audio- und Videoanwendungen bei minimalem Energiebedarf Leistungen erbringen, die bei Universalrechnern nur so genannte Supercomputer erreichen.

Solche Systeme basieren meist auf Prozessoren, die für bestimmte Aufgaben optimiert wurden (z.B. Mikrocontroller, Signal- oder Netzwerkprozessoren). Aufgrund der immensen Fortschritte der Mikroelektronik ist es sogar möglich, Prozessorkerne zusammen mit zusätzlich benötigten digitalen Schaltelementen auf einem einzigen Chip zu realisieren (*System on a Chip* – SoC).

Neben diesen eingebetteten Systemen gibt es auch die so genannten *Universalcomputer*. Gemeinsames Kennzeichen dieser Computersysteme ist, dass sie ein breites Spektrum von Funktionen bereitstellen, die durch dynamisches Laden entsprechender Programme implementiert werden. Neben Standardprogrammen für Büroanwendungen (z.B. Schreib- und Kalkulationsprogram-

me) gibt es für jede nur erdenkliche Anwendung geeignete Software, die den Universalcomputer in ein anwendungsspezifisches Werkzeug verwandelt (z.B. Entwurfs- und Konstruktionsprogramme, Reiseplaner, Simulatoren usw.).

Derartige Universalcomputer unterscheiden sich hinsichtlich der Größe und Leistungsfähigkeit. Die kleinsten und leistungsschwächsten Universalcomputer sind kompakte und leichte Taschencomputer (*Handheld Computer*), die auch als PDAs (*Personal Digital Assistant*) bekannt sind. Sie verfügen über einen nichtflüchtigen Speicher, der auch im stromlosen Zustand die gespeicherten Informationen behält. PDAs können mit einem Stift über einen kleinen berührungsempfindlichen Bildschirm (*Touch Screen*) bedient werden und sind sogar in der Lage, handschriftliche Eingaben zu verarbeiten.

Mobile Systeme, populär auch in *Notebooks*, *Laptops* oder *Netbooks* unterschieden, sind ebenfalls portable Computer. Sie haben im Vergleich zu PDAs größere Bildschirme, eine richtige Tastatur und ein Sensorfeld, das als Zeigeinstrument (Maus-Ersatz) dient. Sie verfügen auch über deutlich größere Speicherkapazitäten (sowohl bzgl. Haupt- als auch Festplattenspeicher) und werden immer häufiger als Alternative zu ortsfesten *Desktop-Computern* verwendet, da sie diesen insbesondere bei Büro- und Kommunikationsanwendungen ebenbürtig sind. Um eine möglichst lange vom Stromnetz unabhängige Betriebsdauer zu erreichen, werden in Notebooks stromsparende Prozessoren eingesetzt.

Desktop-Computer oder *PCs* (*Personal Computer*) sind Notebooks vor allem bzgl. der Rechen- und Graphikleistung überlegen. Neben den typischen Büroanwendungen werden sie zum rechnergestützten Entwurf (*Computer Aided Design* – CAD), für Simulationen oder auch für Computerspiele eingesetzt. Die dazu verwendeten Prozessoren und Graphikadapter produzieren hohe Wärmeleistungen (jeweils in der Größenordnung von ca. 100 Watt), die durch große Kühlkörper und Lüfter abgeführt werden müssen.

Weitere ortsfeste Computersysteme sind die so genannten *Server*. Im Gegensatz zu den Desktops sind sie nicht einem einzelnen Benutzer zugeordnet. Da sie Dienstleistungen für viele über ein Netzwerk angekoppelte Desktops oder Notebooks liefern, verfügen sie über eine sehr hohe Rechenleistung (*Compute Server*), große fehlertolerierende und schnell zugreifbare Festplattensysteme¹ (*File Server*, *Video-Stream Server*), einen oder mehrere Hochleistungsdrucker (*Print Server*) oder mehrere schnelle Netzwerkverbindungen (*Firewall*, *Gateway*).

Server-Systeme werden in der Regel nicht als Arbeitsplatzrechner genutzt, d.h. sie verfügen weder über leistungsfähige Graphikadapter noch über Peripheriegeräte zur direkten Nutzung (Monitor, Tastatur oder Maus).

Um sehr rechenintensive Anwendungen zu beschleunigen, kann man mehrere Compute Server zu einem so genannten *Cluster Computer* zusammenschalten. Im einfachsten Fall, werden die einzelnen Server-Systeme über einen *Switch* mit Fast- oder Gigabit-Ethernet zusammengeschaltet. Über diese

¹ Meist so genannte RAID (*Redundant Array of Independent Disks*).

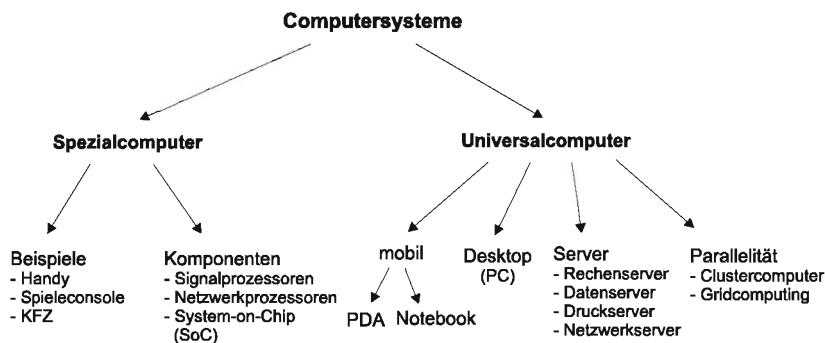


Abb. 1.1 Übersicht über die verschiedenen Arten von Computersystemen.

Verbindungen können dann die einzelnen Compute Server Daten untereinander austauschen und durch gleichzeitige (parallele) Ausführung von Teilaufgaben die Gesamtaufgabe in kürzerer Zeit lösen. Die maximal erreichbare Beschleunigung hängt dabei vom Grad der Parallelisierbarkeit, der sog. *Körnigkeit* (*Granularity*), der Programme ab. Cluster-Computer sind vor allem für grobkörnige (*coarse grained*) Parallelität geeignet. Hier sind die Teilaufgaben zwar sehr rechenintensiv, die einzelnen Programmteile müssen jedoch nur geringe Datenmengen untereinander austauschen.

Je feinkörniger (*fine grained*) ein paralleles Programm ist, desto höher sind die Anforderungen an das dem Cluster zu Grunde liegende Netzwerk. Um eine hohe Beschleunigung der feinkörnigen Programme zu erreichen, muss die Netzwerkverbindung sowohl eine hohe Datenrate bereitstellen als auch möglichst geringe Latenzzeiten aufweisen.

Betrachtet man die über das Internet verbundenen Computersysteme, so erkennt man, dass diese ähnlich wie bei einem Cluster Computer organisiert sind. Auch hier kann jeder Netzwerkknoten mit jedem beliebigen anderen Knoten Daten (oder Programme) austauschen. Aufgrund der komplexen Vermittlungsstrategien des Internetprotokolls (IP) muss man allerdings mit höheren Latenzzeiten und geringeren Datenraten rechnen, d.h. man ist auf grobkörnige Parallelität beschränkt. Trotzdem hält dieser „weltweite“ Cluster-Computer extrem hohe Rechenleistungen bereit, da die angeschlossenen Desktop-Systeme im Mittel nur zu ca. 10% ausgelastet sind. Um diese immense brachliegende Rechenleistung verfügbar zu machen, entstanden in den letzten Jahren zahlreiche Forschungsprojekte zum so genannten *Grid Computing*.

Der Name *Grid* wird in Analogie zum *Power Grid* verwendet, bei dem es um eine möglichst effektive Nutzung der in Kraftwerken erzeugten elektrischen Energie geht. Die Kernidee des Grid Computings besteht darin, auf jedem Grid-Knoten einen permanenten Zusatzprozess laufen zu lassen, über den dann die Leerlaufzeiten des betreffenden Desktop-Computers für das

Grid nutzbar gemacht werden können. Diese Software wird als *Grid Middleware* bezeichnet. Die am weitesten verbreitete Grid Middleware ist das Globus Toolkit. Neben der Grid Middleware wird auch ein so genannter *Grid Broker* benötigt, der für jeden eingehenden Benutzerauftrag (*Job*) geeignete Computerkapazitäten (*Resources*) sucht und der nach der Bearbeitung die Ergebnisse an den Benutzer weiterleitet. In Analogie zum *World Wide Web* (WWW) spricht man beim Grid-Computing auch von einem *World Wide Grid* (WWG). Es bleibt abzuwarten, ob sich dieser Ansatz genauso revolutionär entwickelt wie das WWW.

Nach dem Überblick im Abschnitt 1.1 über die verschiedenen Arten moderner Computersysteme werden wir im weiteren Verlauf des Kapitels den Aufbau von Desktop-Systemen genauer betrachten und anschließend die Architektur der aktuellen Desktop-Prozessoren von AMD und Intel vorstellen. Als verbindenden Komponenten kommt den Chipsätzen eine besondere Bedeutung zu.

1.1.3 Entwicklungstrends

In diesem Unterabschnitt sollen kurz aktuelle Entwicklungstrends skizziert werden, die sich bzgl. der Technologie und Architektur von Computersystemen abzeichnen. Zu den technologischen Trends zählen die weitere Verkleinerung der Strukturen, die Silicon-on-Isolator- und die Kupfertechnologie. Zu den architektonischen Trends gehören Dual- bzw. Multi-Core-Prozessoren, höhere Speicherbandbreiten durch *Prefetching* und die Unterstützung der Sicherheit und Zuverlässigkeit.

Verkleinerung der Strukturen

Moderne Prozessoren werden in CMOS-Technologie realisiert. Die *Strukturgröße* gibt an, wie klein man die geometrischen Strukturen zur Realisierung der Transistoren auf dem Chip² ätzen kann. Immer kleinere Strukturgrößen werden aus folgenden beiden Gründen angestrebt:

- Bei der Herstellung werden gleich mehrere Chips auf einer Halbleiterschleibe, einem so genannten *Wafer*, geätzt. Aus verfahrenstechnischen Gründen hängt die Ausbeute, d.h. der Prozentsatz funktionsfähiger Chips, (und damit der Gewinn) von der Chipfläche ab. Daher darf die Fläche der einzelnen Chips auf dem Wafer nicht zu groß werden. Um dies zu erreichen, muss man bei steigender Zahl der Transistoren (Komplexität) die Strukturgröße verringern.
- Die maximal mögliche Taktfrequenz hängt sowohl von der Geschwindigkeit der Funktionsschaltnetze (z.B. des Rechenwerks) als auch von den Signallaufzeiten auf den Verbindungsleitungen zwischen den Registern und

² Im Englischen *Die* („Plättchen“) genannt.

diesen Funktionsschaltnetzen ab. Demnach kann die Taktfrequenz erhöht werden, wenn die Strukturgröße verkleinert wird.

Leider hat die Verkleinerung der Strukturgröße auch eine Schattenseite: Durch die höhere Transistordichte pro Flächeneinheit steigt auch die spezifische Wärmeleistung (gemessen in Watt pro Quadratzentimeter). Da Halbleiterbausteine bei zu hohen Temperaturen ($> 100^\circ$ Celsius) zerstört werden, muss für ausreichende Wärmeableitung bzw. Kühlung gesorgt werden. Die spezifische Wärmeleistung aktueller Prozessoren liegt in einem Bereich von 80 bis 100 Watt/cm². Zum Vergleich liefert eine 2 kW-Herdplatte (bei 400° Oberflächentemperatur) weniger als 1 Watt/cm².

Um die Wärmeleistung zu reduzieren, entwickelt man insbesondere für Prozessoren in mobilen Computersystemen (Notebooks) intelligente Power-Management-Systeme. So kann beispielsweise sowohl die Taktfrequenz als auch die Betriebsspannung per Software geregelt werden. Da die Leistung quadratisch von der Spannung und linear von der Frequenz abhängt, ergibt sich dadurch eine kubische Leistungsanpassung. Sowohl AMD als auch Intel bieten mittlerweile Prozessoren in 65- bzw. 45-nm-Technologie an.

Silicon-on-Isolator (SOI)

Durch eine vergrabene Oxid-Schicht gelingt es, die Transistoren auf dem Chip vollständig voneinander zu isolieren. Damit kann – bei unveränderter Architektur – die Prozessorleistung um bis zu 30% gesteigert werden. Bei gleicher Taktrate kann die Leistungsaufnahme um bis zu 70% gesenkt werden. Diese Technologie wird bereits beim Opteron von AMD eingesetzt.

Kupfertechnologie

Hier wird zur Herstellung von leitenden Verbindungen zwischen den Transistoren Kupfer anstatt Aluminium verwendet. Der Widerstand der Leiterbahnen sinkt dadurch um 40%, die Signallaufzeiten werden verkürzt und die Taktrate kann um 35% erhöht werden. Auch diese technologische Neuerung wird bei PC-Prozessoren seit einigen Jahren schon angewandt.

Dual-Core-Prozessoren und Multicore-Prozessoren

Nachdem sich beim Intel Pentium 4 die mehrfädige Programmausführung durch *Hyper-Threading* etabliert hatte, wurden sowohl von AMD als auch von Intel schon im Jahr 2005 neue Prozessoren mit 2 bis 4 Kernen auf einem Chip auf den Markt gebracht. Während Hyper-Threading den Prozessen nur zwei *logische* Prozessoren bereitstellt, hat man bei den Multi-Core-Prozessoren echte Hardware-Parallelität (*Multiprocessing*), d.h., das Betriebssystem muss nicht immer zwischen den beiden Programmfäden (*Threads*) umschalten. Wegen der hohen Wärmeentwicklung bei dicht nebeneinander liegenden Prozessorkernen muss jedoch die Taktfrequenz reduziert werden. Außerdem kommt es bei dieser Architektur auch verstärkt zu Zugriffskonflikten bei den gemeinsamen schnellen Pufferspeichern, den sog. *Caches*, und dem Hauptspeicher.

Erhöhung der Speicherbandbreite

In den kommenden Jahren wird sich durch immer ausgefeiltere Speicherarchitekturen die erreichbare Bandbreite erhöhen. Durch *Prefetching* beim Speicherzugriff lassen sich trotz gleicher Halbleitertechnologie im Mittel enorme Steigerungsraten erreichen. So haben sich DDR2-Speicher bereits durchgesetzt und von DDR3 werden weitere Steigerungen erwartet. Die Erhöhung der Speicherbandbreite wirkt sich unmittelbar auf die Leistung eines Computersystems aus, da bislang noch große Geschwindigkeitsunterschiede zwischen Register- und Speicherzugriff bestehen. Es ist daher sehr wirksam, den Speicher-*Flaschenhals* zu beseitigen. Eine weitere Maßnahme in die gleiche Richtung ist die Vergrößerung der Speicherkapazität der schnellen Zwischenspeicher, der sog. Caches. Da hiermit die Trefferrate vergrößert wird, trägt auch sie dazu bei, die Speicherbandbreite des Gesamtsystems zu erhöhen.

Sicherheit und Zuverlässigkeit

Computerviren und andere unerwünschte Programme wie Viren, *Spyware* oder Trojaner richten immer größere Schäden an. Es ist daher sehr wichtig, Prozessorarchitekturen zu entwickeln, die derartige Angriffe frühzeitig erkennen und entsprechend reagieren. Daher implementieren die Hersteller Schutzmechanismen für Prozessoren, die ein unerlaubtes Ausführen von Programmen verhindern. Da Computerviren meist durch Pufferüberläufe eingeschleust werden, blockiert man durch entsprechende Hardware das Schreiben nach einen solchen Überlauf.

Neben der Sicherheit soll auch die Zuverlässigkeit von Computersystemen durch verbesserte Architekturen erhöht werden. Da beim Ausfall eines Computers sehr hohe Kosten entstehen können, wäre es wünschenswert, fehlerhafte Systemteile frühzeitig zu erkennen und trotzdem fehlerfrei weiterzuarbeiten. So könnte man beispielsweise drei Prozessorkerne parallel betreiben und deren Ergebnisse ständig miteinander vergleichen. Liefern zwei dieser Prozessorkerne übereinstimmende Ergebnisse, während die Ergebnisse des dritten Prozessorkerns davon abweichen, so sind diese Ergebnisse wahrscheinlich fehlerhaft. Trotz dieses Ausfalls kann aber das Gesamtsystem zuverlässig weiterarbeiten. In sicherheitskritischen Bereichen ist eine derartige *Fehlertoleranz* oft wichtiger als hohe Rechenleistung. Daher wird es in Zukunft gewiss auch Prozessoren geben, die auf eine hohe Zuverlässigkeit optimiert sind.

1.2 Komponenten eines Personal Computers

Ein PC (*Personal Computer*) ist ein Rechner, der – wie der Name es nahe legt – nur von einer Person (oder wenigen Personen) genutzt wird. Im Gegensatz zu zentralen (Groß-)Rechnern steht somit beim PC dem Benutzer die gesamte Rechenleistung exklusiv zur Verfügung. Da ein PC am Arbeitsplatz

des Benutzers steht, wird er auch als *Desktop PC* bezeichnet³. Ein *Desktop PC* verfügt über alle Hard- und Software-Komponenten, die zur interaktiven Arbeit mit lokalen Anwendungen – wie z.B. Textverarbeitung, Tabellenkalkulation, Programmierung usw. – nötig sind. Darüber hinaus verfügen alle modernen PCs über Netzwerkschnittstellen, um mit anderen Rechnern Daten auszutauschen oder auf zentrale Dateiserver bzw. auf das Internet zuzugreifen. Leistungsfähige PCs, die über eine solche Netzwerkschnittstelle und besonders schnelle Graphiksysteme verfügen, werden auch als *Workstations* bezeichnet. Ein *Server* ist ein Rechner, der Betriebsmittel, wie Dateisysteme, Drucker, Internetverbindungen usw., für andere Rechner bereitstellt. Ein *Notebook PC* oder *Laptop PC* fasst alle Komponenten eines Personal Computers auf kleinstem Raum zusammen. Da diese mit Akkumulatoren (Akkus) betrieben werden können, sind sie portabel. Eine weitere Miniaturisierung findet man bei *Handheld PCs* oder *Palmtop PCs*.⁴

Während bei mobilen Systemen alle Komponenten in einem Gerät integriert sind, können bei stationären PCs die einzelnen Komponenten – wie die Systemeinheit mit ihren Schnittstellen sowie die Ein- und Ausgabegeräte – unterschieden werden. Die wichtigsten Eingabegeräte sind Tastatur und Maus, die wichtigsten Ausgabegeräte Monitor und Drucker.

Die Systemeinheit besteht aus einem Gehäuse mit Netzteil und den Schnittstellen zur Peripherie. Die Abbildung 1.2 gibt einen ersten, groben Überblick über die grundlegenden Komponenten innerhalb der Systemeinheit eines PCs:

- Hauptplatine mit integrierten Schnittstellen,
- Graphikkarte(n),
- Schnittstellenkarten für PCI-Express oder andere Bussysteme,
- Festplattenlaufwerk(e) (*Hard-Disk Drive* – HDD),
- Diskettenlaufwerk(e), die jedoch kaum noch zu finden sind,
- CD-ROM- bzw. DVD-Laufwerke,
- Monitor, Tastatur und Maus.

Auf der Hauptplatine (*Motherboard*, *Mainboard*) befinden sich der Prozessor, der Hauptspeicher und Steckplätze für Bussysteme mit unterschiedlicher Geschwindigkeit. Über speziell auf die Prozessoren abgestimmte Chipsätze werden die genormten Peripheriebus-Signale aus den Prozessorbus-Signalen abgeleitet. An diese Busse können über Steckkontakte Schnittstellenkarten angeschlossen werden, die dann standardisierte Schnittstellen für Massenspeicher oder Peripheriegeräte bereitstellen. Auf den meisten aktuellen Hauptplatinen sind auch bereits einfache parallele und serielle Standardschnittstellen sowie meist auch eine ganze Anzahl von USB-Schnittstellen (*Universal Serial Bus*) integriert. Ebenso findet man meist auch eine (S)ATA-Schnittstelle in paralleler oder serieller Form (*(Serial) Advanced Technology Attachment*)

³ *Desktop*: Schreibtisch-Oberfläche

⁴ *Notebook*: Notizbuch, *Lap*: Schoß, *Palm*: Handfläche

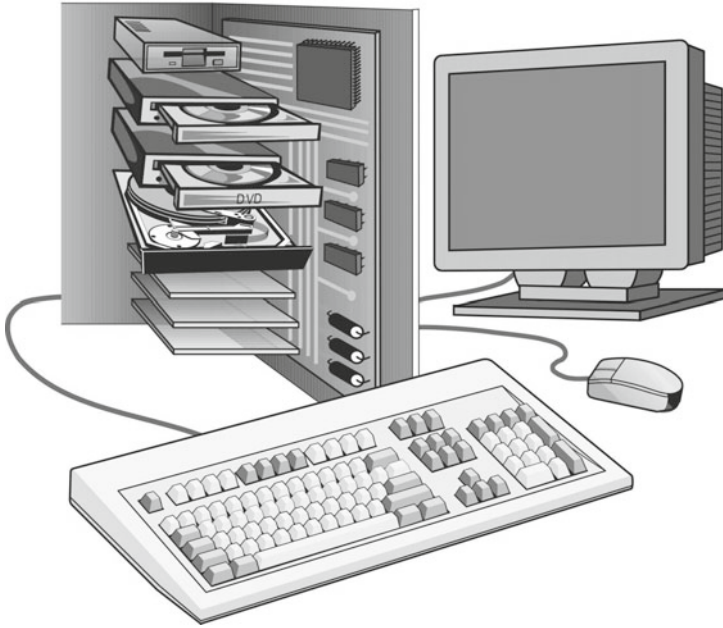


Abb. 1.2 Der Aufbau eines PCs.

für magnetomotorische und optische Massenspeicher mit einem integriertem Laufwerkscontroller (*Integrated Device Electronic* – IDE). Neben Controllern für Laufwerke ist bei heutigen Hauptplatinen meist auch schon eine Ethernet-Netzwerkschnittstelle zur Integration des PCs in ein lokales Netzwerk vorhanden.

Zunächst lassen wir bei unserer Betrachtung die externen Geräte eines PCs außer Betracht und behandeln ausschließlich den im PC „eingebetteten“ Mikrorechner, der im Wesentlichen als Steckkartensystem realisiert ist. Abbildung 1.3 zeigt schematisch die wesentlichen Komponenten, die sich auf der Hauptplatine befinden (s. auch Abbildung 1.4):

- der Mikroprozessor, auch CPU (*Central Processing Unit*) genannt, der in einem speziellen Sockel eingesteckt wird,
- Module des Hauptspeichers, für die eine unterschiedliche Anzahl von Steckplätzen vorhanden sind,
- ein Steckplatz für eine Graphikkarte,
- den Peripherie- oder Erweiterungsbussen mit Steckplätzen zur Aufnahme verschiedener Steckkarten (Erweiterungskarten, *Add-On Cards*),
- der so genannte *Chipsatz*, eine Sammlung von mehreren hochintegrierten Bausteinen, die insbesondere die Verbindung der eben genannten Komponenten und die Unterstützung der Kommunikation zwischen ihnen zur

Aufgabe haben; daneben enthalten sie aber noch eine ganze Reihe von Steuermodulen zum Anschluss interner und externer Geräte;

- verschiedene Steuer- und Schnittstellenbausteine unterschiedlicher Komplexität, die direkt an den Bausteinen des Chipsatzes oder am Peripheriebus angeschlossen sind und z.T. über Steckverbinder mit externen Komponenten gekoppelt werden.

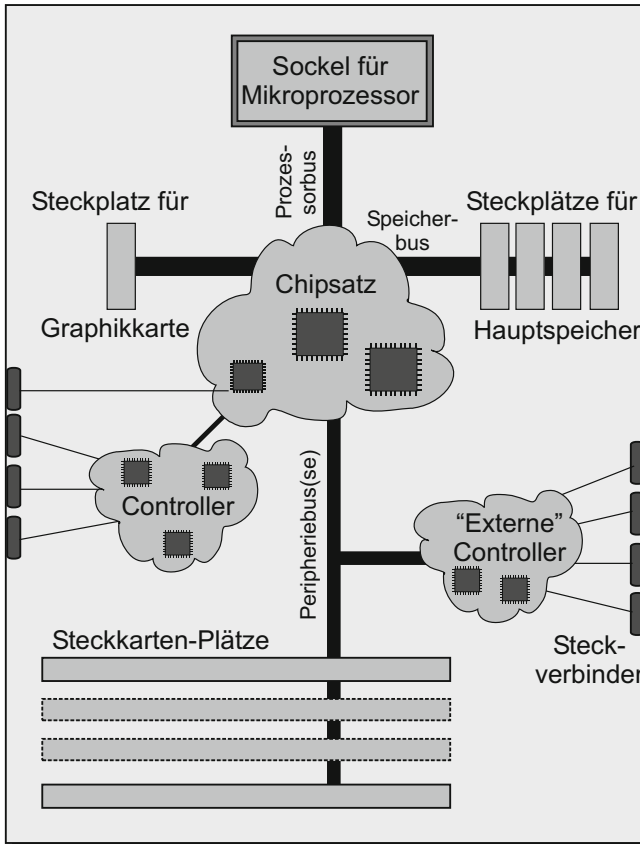


Abb. 1.3 Der prinzipielle Aufbau einer Hauptplatine.

Die erwähnten Sockel und Steckplätze sind in der Regel exakt auf spezielle Bausteine, also Prozessoren, Speichermodule und Graphikkarten, abgestimmt. Hochleistungs-PCs enthalten darüber hinaus auch Sockel und Steckplätze für mehr als einen Prozessor bzw. zwei Graphikkarten. Die in den Chipsätzen integrierten Steuer- und Schnittstellenmodule, aber auch die oben erwähnten „externen“ Steuer- und Schnittstellenbausteine werden typischerweise *Controller* genannt. Zu den Erweiterungskarten gehören insbesondere Gra-

phikkarten, Audiotkarten (*Soundcards*) und Netzwerkkarten (*LAN ards*). Der Fortschritt der Integrationstechnik ermöglicht es aber, immer mehr von diesen Erweiterungen direkt auf der Hauptplatine unterbringen, z.B. den Controller für den Netzwerkanschluss und die Audioverarbeitung. Andererseits kann der PC durch Erweiterungskarten wieder mit „Altlast“-Schnittstellen (*Legacy*) ausgerüstet werden, die in den letzten Jahren nicht mehr zum Standardumfang eines PCs gehören und daher nicht mehr vom Chipsatz zur Verfügung gestellt werden, wie zum Beispiel die früher weit verbreiteten parallelen und seriellen Schnittstellen⁵. Ein Steckverbinder-Modul, das auf der Hauptplatine zur Gehäuserückwand zeigt, liefert die Anschlüsse für eine ganze Reihe von Standard-Ein-/Ausgabegeräten, wie z.B. Tastatur, Maus, Netzwerk, Lautsprecher usw. (Auf dieses Modul werden wir weiter unten eingehen.) Über eine Reihe von Steckern können die oben erwähnten Massenspeicher angeschlossen werden, wobei Flachbandkabel mit mehr oder weniger Kupferadern verwendet werden.

In diesem Kapitel werden wir uns ausschließlich mit der Hauptplatine eines PCs und ihren Komponenten beschäftigen. Nicht behandelt werden der mechanische Aufbau eines PCs, die verschiedenen Ausprägungen von gebräuchlichen Hauptplatinen, Sockel für Prozessor und Chipsätze, Fragen der Kühlung und Lüftung, Steckverbinder, Netzteile und Spannungsversorgung, Überwachung von wichtigen Parametern (Temperatur, Spannung usw.)

Die im PC-Bereich eingesetzten Mikroprozessoren werden als *x86-kompatible Prozessoren* bezeichnet, da sie sich auf den ersten 16-Bit-Prozessor von Intel, den 8086, aus dem Jahr 1979 zurückführen lassen.⁶ Über die Leistungsfähigkeit der x86-Prozessoren entscheidet nicht zuletzt die komplexe virtuelle Speicherverwaltung und ihre Unterstützung durch die Speicherverwaltungseinheit (*Memory Management Unit* – MMU). Selbst eine oberflächliche Beschreibung der virtuellen Speicherverwaltung und ihrer vielfältigen Funktionen würde den Rahmen dieses Kapitels sprengen. Wir werden dieses Thema daher erst in Kapitel 2 behandeln.

⁵ Diese waren früher unter den Bezeichnungen Centronics- und V.24-Schnittstellen bekannt.

⁶ Über diese Prozessoren werden Sie noch in diesem Kapitel eine Reihe von Details erfahren und die Hauptkomponenten kennen lernen.

1.3 Hauptplatine und ihre Komponenten

1.3.1 Hauptplatine

In Abbildung 1.4 ist exemplarisch eine moderne Hauptplatine für den Intel-Prozessor Core 2 gezeigt. Die in der Abbildung dargestellten Komponenten sind Gegenstand dieses und z.T. auch der folgenden Kapitel.

Um Ihnen eine grobe Vorstellung von der Größe einer Hauptplatine zu geben, seien hier nur die Maße einer typischen Platine für den *Desktop*-Bereich gegeben (vgl. Abbildung 1.4). Ihre „genormte“ Größe wird als *ATX-Formfaktor* bezeichnet und belegt die folgende Rechteckfläche: $9,6 \times 12 \text{ Zoll}^2 = 24,4 \times 30,5 \text{ cm}^2$. (Sie ist damit nur wenig größer als eine DIN-A4-Seite mit $21,0 \times 29,7 \text{ cm}^2$.)

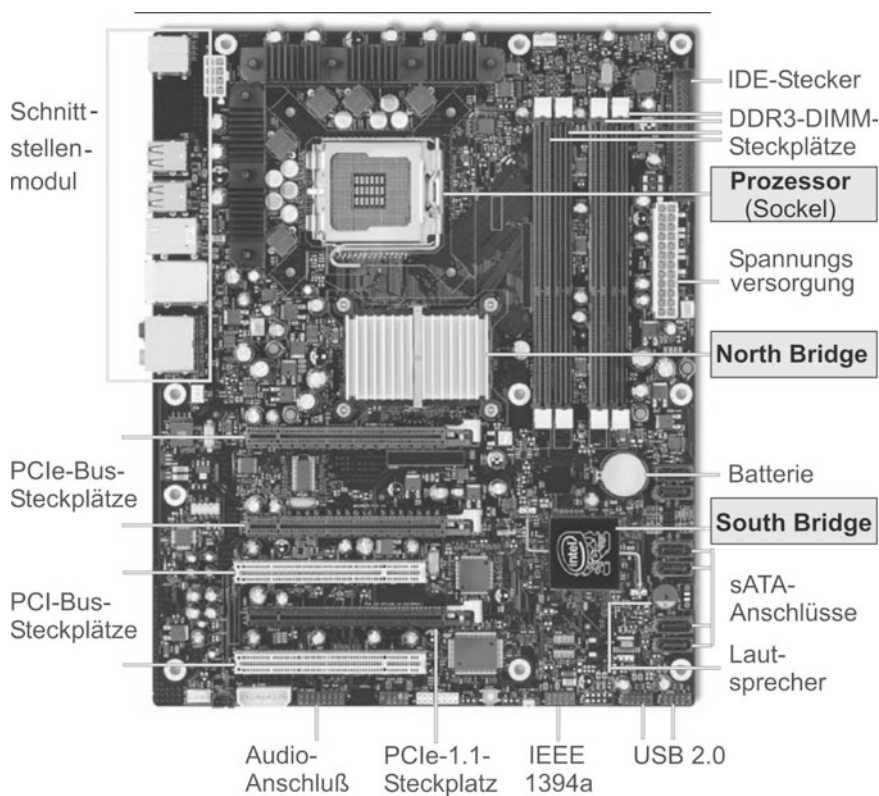


Abb. 1.4 Photo einer Hauptplatine (*Intel Desktop Board DX48BT2*).

Die Hauptplatine eines PCs (ebenso wie seine Einsteckkarten) enthielten noch vor etwas mehr als zwei Jahrzehnten Dutzende von integrierten Bausteinen (*Integrated Circuits* – ICs). Die rasante Entwicklung der Höchstintegrationstechnik (*Very Large Scale Integration* – VLSI) ermöglichte es seitdem, immer mehr Komponenten auf dem Prozessorchip selbst bzw. in sehr wenigen „Hilfsbausteinen“ unterzubringen. Die wichtigsten dieser Bausteine, die zum Aufbau eines PCs benötigt werden und auf der Hauptplatine Platz finden, werden zusammenfassend als *Chipsatz* (*Chipset*) bezeichnet. Der erste dieser Chipsätze wurde im Jahre 1988 von der Firma *Chips and Technologies* auf den Markt gebracht. Seither bieten die Prozessorhersteller (vor allem AMD und Intel) sowie auf die Entwicklung von Chipsätzen spezialisierte Firmen kurz nach dem Erscheinen eines neuen Prozessors auch die dazu passenden Chipsätze an.

Ein Chipsatz besteht meist aus ein bis drei Chips, die benötigt werden, um den Prozessor mit dem Speichersystem und Ein-/Ausgabebussen zu koppeln. Da diese drei Haupteinheiten mit unterschiedlichen Geschwindigkeiten arbeiten, benötigt man *Brückenbausteine* (*Bridges*), welche die vorhandenen Geschwindigkeitsunterschiede ausgleichen und für einen optimalen Datenaustausch zwischen den Komponenten sorgen. Die Brückenbausteine müssen auf die Zeitsignale (*Timing*) des Prozessors abgestimmt werden. Die im PC eingesetzten Bussysteme und Schnittstellen unterscheiden sich sehr stark in der Anzahl der Daten- und Adresssignale, der Taktfrequenzen und der verwendeten Spannungspegel sowie der zugrunde liegenden Busprotokolle. Aufgabe der Brückenbausteine ist insbesondere die elektrisch/physikalische Anpassung der verschiedenen Bussignale sowie die Berücksichtigung der unterschiedlichen Übertragungsleistungen.

Der Chipsatz hat also großen Einfluss auf die Leistungsfähigkeit eines Computersystems und muss daher optimal auf den eingesetzten Prozessor und die verwendeten Speicher-/Bustypen zugeschnitten sein. Die Bausteine des Chipsatzes übernehmen im eigentlichen Sinne die Steuerung des Systems.⁷ Obwohl ein Chipsatz für eine bestimmte Prozessorfamilie entwickelt wird, kann er meist auch eine Vielzahl kompatibler Prozessoren unterstützen.

Chipsätze verfügen über z.T. sehr unterschiedliche Komponenten, Anschlüsse, Register und Funktionen. Es ist Aufgabe des sog. *BIOS* (*Basic Input/Output System*), diese Unterschiede vor dem Betriebssystem zu „verstecken“, das meist für viele Prozesstypen und Familien einsetzbar sein muss. Das BIOS ermöglicht dem Betriebssystem und den darunter laufenden Anwendungsprogrammen den Zugriff auf die Hardware-Komponenten. Dazu muss es genaue Kenntnisse über den Aufbau der Hauptplatine, die Anzahl der Steckplätze sowie die verwendeten Bausteine und Komponenten besitzen. Als Beispiele seien hier nur der Zugriff auf die Plattenlaufwerke und die Ab-

⁷ Auf die speziellen Komponenten zur Steuerung und Überwachung der Versorgungsspannung, der Verlustleistung, der eingesetzten Lüfter und bestimmter physikalischer Größen und Einheiten (System-/Power-/Takt-„Management“) können wir aus Platzgründen leider nicht eingehen.

frage der Tastatur genannt. Abbildung 1.5 zeigt die Lage des BIOS zwischen dem Betriebssystem und der Hardware, im Wesentlichen repräsentiert durch den Prozessor und den Chipsatz. Als wesentliche Aufgaben des BIOS seien hier aufgeführt:

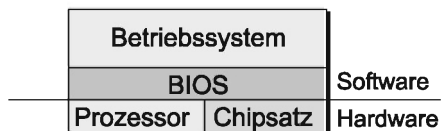


Abb. 1.5 Das BIOS zwischen Betriebssystem und Hardware-Bausteinen.

- Ausführung eines PC-Selbsttests, der automatisch nach dem Einschalten des Systems aufgerufen wird (*Power-On Selftest* – POST). Dadurch werden z.B. die Größe des Hauptspeichers ermittelt und ein Test seiner Speicherzellen durchgeführt.
- Konfiguration der auf der Hauptplatine oder in den Steckplätzen eingesetzten Komponenten, insbesondere der am PCI- bzw. PCIe-Bus angeschlossenen Geräte.
- Ausführung eines *Setup*-Programms, durch das der Anwender BIOS-Daten ändern und Einfluss auf den Systembetrieb nehmen kann. Als Beispiel sei die manuelle Zuweisung von Interrupt-Kanälen zu bestimmten Komponenten genannt.

Abbildung 1.6 zeigt das Blockschaltbild einer Platine und die Lage und Verbindungen der Bausteine des Chipsatzes⁸. Häufig sind die wesentlichen Funktionen eines Chipsatzes (heute noch) auf zwei Brücken-Bausteine aufgeteilt, die wegen ihrer Lage auf der senkrecht stehenden Platine anschaulich als *North Bridge* und *South Bridge* bezeichnet werden.

North Bridge

Sie verbindet den Prozessor mit allen Komponenten, die einen möglichst schnellen Datentransfer benötigen. Das sind insbesondere der Hauptspeicher und die Graphikeinheit. Da die North Bridge sich insbesondere um die Steuerung der Zugriffe auf den angeschlossenen Hauptspeicher kümmern muss, wird sie auch als Speicher-Controller-Hub (*Memory Controller Hub* – MCH) bezeichnet. Einige MCHs enthalten außerdem einen eigenen Graphikcontroller. Sie werden dementsprechend als *Graphics Memory Controller Hub* (GMCH) bezeichnet.

⁸ Die in der Abbildung verwendeten Bezeichnungen werden im weiteren Verlauf des Abschnitts erklärt.

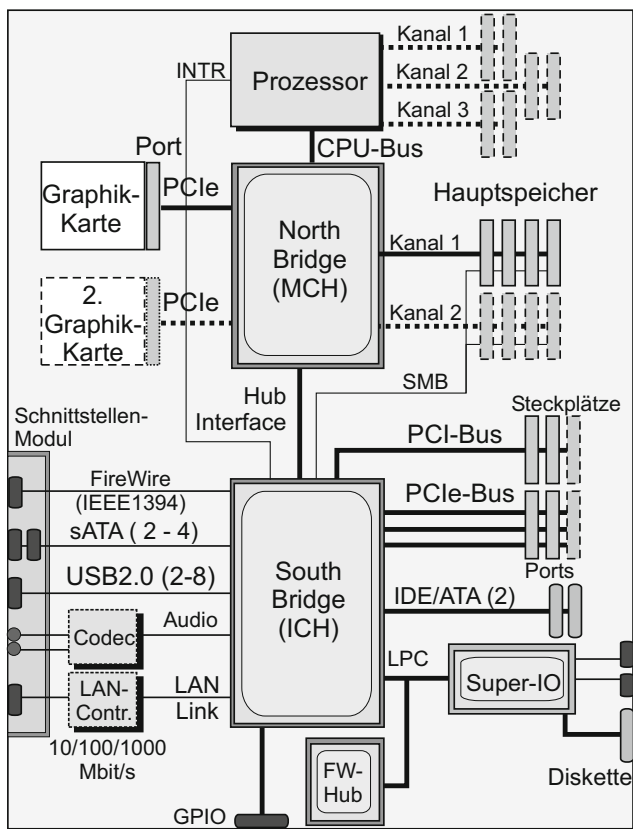


Abb. 1.6 Blockschaltbild einer Hauptplatine mit Hub-Architektur.

South Bridge

Sie verbindet als zweiter Baustein den Prozessor oder Hauptspeicher mit einer Reihe von integrierten oder extern hinzugefügten Controllern, die insbesondere zur Steuerung von Massenspeicher- oder Ein-/Ausgabegeräten dienen. Daher rührt ihre Bezeichnung als *Ein-/Ausgabecontroller-Hub* (*I/O Controller Hub* – ICH). Sie enthält dazu insbesondere eine Vielzahl von USB-2.0-Schnittstellen, über die heutzutage sehr viele externe Geräte an den PC angeschlossen werden können. Zusätzlich sichert sie die Kompatibilität zu älteren Systemkomponenten. Dazu enthält sie oft noch die vom „legendären“ IBM-AT-kompatiblen PC eingeführten Komponenten, wie z.B. die batteriegepufferte Echtzeituhr (*Real Time Clock* – RTC).

Super-I/O-Baustein

Häufig sind einige dieser einfachen Komponenten aber auch in einem besonderen Baustein integriert, dem *Super-I/O-Baustein*, der über die unten

beschriebene LPC-Schnittstelle an die *South Bridge* angeschlossen wird. Die weitere Entwicklung der PC-Technik wird die Super-I/O-Bausteine voraussichtlich überflüssig machen, da einerseits immer mehr Komponenten zusätzlich in die *South Bridge* integriert werden können, andererseits aber auch immer mehr Geräte über den USB an den PC angeschlossen und so einige Schnittstellen überflüssig werden. Neben den bereits erwähnten Komponenten stellt der Super-I/O-Baustein oft noch alle oder einige der folgenden Schnittstellen und Komponenten zur Verfügung, auf die wir jedoch im Rahmen dieses Buches nicht näher eingehen können:

- eine Infrarot-Schnittstelle als weit verbreitete „drahtlose“ serielle Schnittstelle (*Infrared Data Association – IrDA*),
- einen Tastatur- und Maus-Controller,
- einen *Floppy-Disk Controller* (FDC),
- einen so genannten MIDI-Port (*Musical Instrument Digital Interface*) zur digitalen Steuerung von Musikinstrumenten,
- einen Hardwaremonitor zur Überwachung des Rechnerzustands über Temperaturfühler und Spannungsmessern sowie zur Steuerung und Regelung der eingesetzten Lüfter,
- eine Komponente zur Steuerung und Regelung der Versorgungsspannung und der Leistungsaufnahme (*Power Management*).

Firmware Hub

Das oben beschriebene BIOS befindet sich in einem Festwertspeicher, dem BIOS-ROM, auf der Hauptplatine, der eine Speicherkapazität von 4 bis 8 Mbit hat und im sog. *Firmware Hub* (FWH) untergebracht ist. Wegen der geringen Zugriffsgeschwindigkeiten dieser Speichertypen wird der Inhalt des BIOS-Speichers nach dem Einschalten des PCs aus dem Festwertspeicher ausgelesen und in einem besonderen Bereich des Hauptspeichers abgelegt, von dem aus er schneller bearbeitet werden kann (*Shadowing*). Üblicherweise enthält der FWH noch eine Hardwarekomponente zur Erzeugung von Pseudo-Zufallszahlen (*Random Number Generator – RNG*), die z.B. für Verschlüsselungsalgorithmen benutzt werden können. Der FWH wird gewöhnlich im Multiplexbetrieb gemeinsam mit dem Super-I/O-Baustein über die unten beschriebene LPC-Schnittstelle angesprochen.

Die beiden Brückenbausteine – North und South Bridge – sind heute durch eine schnelle dedizierte Verbindung miteinander gekoppelt, die als *Hub Interface* bezeichnet wird. Sie ist sehr viel leistungsfähiger als bei älteren Chipsätzen, bei denen die North Bridge und die South Bridge sowie alle Ein-/Ausgabeschnittstellen über den relativ langsamen PCI-Bus, der eine maximale Übertragungsrate von 133 MB/s erreichte, angebunden wurden. Bei Intel wird diese Verbindung als DMI (*Direct Media Interface*) bezeichnet und überträgt maximal 2 GB/s⁹. AMD (*Advanced Micro Devices*) setzte dafür

⁹ GB/s: Gigabyte/Sekunde, also 10⁹ Byte/Sekunde.

früher den eigenentwickelten *HyperTransport*-Bus ein, der maximal bis zu 52 GB/s übertragen soll. Bei neueren Chipsätzen verwendet AMD jedoch einen 4×4 Leitungen breiten *PCI-Express-Bus* (PCIe-x4), der (nach der Spezifikation 2.0) maximal 4 GB/s übertragen kann (s. Unterabschnitt 1.3.2.). Wie wir aus der Abbildung 1.4 sehen können, muss der MCH mit einem eigenen Kühlkörper ausgestattet werden, da er mit sehr hohe Taktraten arbeitet. Dagegen kommt der ICH meist ohne Kühlkörper aus.

Zusammenfassend spricht man bei neueren Chipsätze meist auch von einer *Hub-Architektur* (*Hub Architecture*). Dabei ist abzusehen, dass durch die Fortschritte der Halbleiter-Integrationstechnologie konsequenterweise bereits in naher Zukunft die beiden Brückenbausteine, North Bridge und South Bridge, zu einem einzigen *Hub*-Baustein zusammengefasst werden oder die North Bridge vollständig auf dem Prozessor-Chip Platz finden wird. So hat bereits die Firma Intel einen Ein-Chip-Chipsatz (Typkennung P55) auf den Markt gebracht, der die letztgenannte Variante realisiert.

1.3.2 North Bridge

Wegen ihrer weit reichenden Funktionen wird die North Bridge auch als *System Controller Hub* bezeichnet. Sie muss den Datenfluss zwischen den verschiedenen Komponenten steuern und setzt dazu Schreib- und Leseanforderungen des Prozessors, des Graphikcontrollers oder einer Komponente an der South Bridge in Speicher-Buszyklen um. Dazu muss die North Bridge die Protokollanpassungen zwischen dem Prozessorbus, dem Speicherbus und der Graphikcontroller-Schnittstelle (s.u.) vornehmen. Zum Geschwindigkeitsausgleich speichert sie alle transportierten Daten zwischen. Dies geschieht in Registersätzen, die als Warteschlangen organisiert sind (*First in, First out* – FIFO) und für jede Übertragungsrichtung (*Read/Write* – RD/WR) und Quelle/Ziel-Kombination getrennt vorhanden sind. Die Puffer besitzen unterschiedliche Größen und erlauben mit ihren dediziert ausgelegten Verbindungen den simultanen Transport von Daten.

Die Verbindung des Prozessorbausteins mit der North Bridge wird über den bereits genannten Prozessorbus vorgenommen, der auch als CPU-Bus oder – anschaulich – als *Front Side Bus* (FSB) bezeichnet wird. In den Anfangszeiten der Chipsätze wurde dazu der Systembus des Prozessors, bestehend aus Adress-, Daten- und Steuerbus, verwendet. Zur Erhöhung der Leistungsfähigkeit ersetzte die Firma AMD den Systembus aber schon seit Jahren durch ihre HyperTransport-Verbindung, von der gleich drei Schnittstellen auf den Prozessoren zur Verfügung gestellt werden (vgl. Abschnitt 1.4.1).

HyperTransport ist eine Technologie, die für den Einsatz an verschiedensten Stellen einer Hauptplatine vorgesehen ist. Dabei handelt es sich um eine Punkt-zu-Punkt-Verbindung, die über zwei getrennte unidirektionale Pfade eine Vielzahl von Halbleiterbausteinen – also Prozessoren, Brücken, Speiche-

reinheiten und Peripheriekomponenten – im Vollduplex-Betrieb miteinander koppelt. Die Verbindungspfade sind als parallele Busse ausgelegt. Ihre Breite kann in Abhängigkeit von den angeschlossenen Komponenten 2, 4, 8, 16 oder 32 Bit betragen, wobei die Verbindungen der beiden Übertragungsrichtungen unterschiedliche Breiten haben können. In HyperTransport-Systemen sind Taktfrequenzen bis zu 800 MHz möglich, wobei eine Zweiflanken-Übertragung (*Double Data Rate* – DDR, *Doubled Pumped*) eingesetzt wird. Die Übertragungsrate pro Richtung liegt damit bei 400 bis 1.600 MegaTransfers pro Sekunde (MT/s). Sie entspricht einer Datenrate von 100 bis 6.400 MB/s für jede der beiden Richtungen einer Verbindung, also einer maximalen Gesamtrate von 12,8 GB/s. Die Übertragung der Informationen geschieht in Form von kurzen Paketen mit einer maximalen Länge von 64 Bytes. Dabei werden Anforderungs-, Antwort- und Rundspruchpakete unterschieden, die jeweils Befehle, Adressen und Daten enthalten können. Die Übertragung der Signale auf den Busleitungen geschieht differentiell mit niedrigen 1,2V-Spannungsepegeln (*Low-Voltage Differential Signalling* – LVDS).

Die Firma Intel hingegen hielt sehr lange an ihrem speziellen Systembus, dem *Front Side Bus* (FSB) des Pentium 4, fest. Erst in der neuesten Prozessorgeneration, dem Intel Core i7, hat auch Intel den Übergang zu einer speziellen Punkt-zu-Punkt-Schnittstelle zur North Bridge vollzogen. Der *Quick-Path Interconnect* (QPI) besteht aus zwei einzelnen Verbindungen (*Links*) für jede Richtung des Datentransfers. Jeder Link überträgt 20 unidirektionale Signale in differenzieller Form, d.h. jeweils auf einem Leitungspaar in unterschiedlichen logischen Pegeln. Dazu kommt für jede Richtung noch ein Leitungspaar für die Übermittlung eines Taktsignals, sodass der QPI insgesamt aus 84 Leitungen besteht. Von den 20 Datenleitungen werden 16 für die parallele Übertragung von Zwei-Byte-Daten benutzt, die restlichen vier werden zur Fehlerkorrektur und zur Kontrolle der Übertragungen eingesetzt. Der QPI arbeitet mit (zunächst) 3,2 GHz, wobei nach dem Zweiflankenverfahren (*Double Data Rate* – DDR) beide Flanken des Taktes zur Übertragung verwendet werden. Damit sollen nun Übertragungsraten bis zu 25,6 GB/s ermöglicht werden, davon je 12,8 GB/s für jede Übertragungsrichtung.

Wesentlicher Bestandteil konventioneller North Bridges ist der *Speicher-Controller* (*Memory Controller*), der über ein oder zwei unabhängige Schnittstellen („Kanäle“) den Zugriff auf die Speichermodule des Hauptspeichers erlaubt. Über jede Schnittstelle können Daten mit einer Übertragungsrate von max. 12,8 GB/s geschrieben oder gelesen werden. Der Controller erzeugt – oft für unterschiedliche Typen von Speicherbausteinen – alle benötigten Steuersignale und übernimmt ggf. das Auffrischen der dynamischen Speicherbausteine (*Refresh*). Diese speichern die Information in winzigen Kondensatoren und würden ohne diese Maßnahme im Sekundenbereich ihre Information verlieren. Bei Verwendung eines ECC-geschützten Speichers (*Error Correcting Code*) übernimmt der Speicher-Controller die Überprüfung und Reaktion auf Speicherfehler durch spezielle zusätzliche ECC-Bits. Der Speicher-Controller enthält Puffer zur Bearbeitung mehrerer Schreib-/Lesezugriffe und sorgt für

deren geordnete Abarbeitung. Dadurch werden insbesondere Ladezugriffe (*Line Fill*) auf den schnellen Zwischenspeicher, dem *Cache*, unterstützt. Auf den Aufbau der Speichermodule, die zur Realisierung des Hauptspeichers eingesetzt werden, gehen wir in einem eigenen Unterabschnitt 1.5 ausführlich ein.

Bereits mit dem Athlon 64 begann die Firma AMD damit, den Speicher-Controller aus der North Bridge in den Prozessorchip selbst zu verlagern und so der CPU über dedizierte Speicherschnittstellen („Kanäle“) mit jeweils 64 Bit Breite einen schnelleren konfliktfreien Zugriff auf den Speicher zu ermöglichen. Diesen Schritt hat die Firma Intel mit ihrer Prozessorfamilie Intel Core i7 nun nachgeholt und ermöglicht den Speicherzugriff sogar über drei unabhängige Kanäle. Damit wird eine maximale Übertragungsrate von zusammen bis zu 25,6 GB/s erreicht. In Abbildung 1.6 ist die Integration des Speicher-Controllers in den Prozessorbaustein durch die gestrichelt gezeichneten Hauptspeicher-Steckplätze angedeutet. Bei einigen Intel-Prozessoren der i7-Familie ist bereits die gesamte North Bridge auf dem Prozessor-Chip untergebracht. Dies ist in Abbildung 1.6 durch die gestrichelt gezeichnete Umrahmung des Prozessors und der North Bridge skizziert.

Die zweite wichtige Komponente, die über die North Bridge angesteuert wird, ist die *Graphikeinheit*, die aus einem oder zwei Graphikcontrollern besteht¹⁰. Der Anschluss der Graphikcontroller an die North Bridge geschieht heutzutage meist über den *PCI-Express*, auch als PCIe bezeichnet, der als skalierbare Variante des älteren PCI-Busses zuerst in den Intel-Chipsätzen eingeführt wurde. Die PCIe-Verbindung wird nicht mehr als paralleler Bus ausgeführt, sondern besteht aus einer oder mehreren bidirektionalen seriellen Verbindungen, die als *Lanes* bezeichnet werden. Jede *Lane* stellt eine Nutz-Transferrate¹¹ von 500 MB/s je Richtung bereit (PCIe-2.0-Spezifikation). Um auch dem Leistungsbedarf künftiger Graphikanwendungen gerecht zu werden, wurden für den Anschluss von Graphikkarten von Anfang an gleich 16 *Lanes* verwendet, die Übertragungen mit 16-facher Geschwindigkeit erlauben. Die Schnittstelle, der Graphikport, wird dementsprechend mit PCIe-x16 bezeichnet und ermöglicht eine maximale Übertragungsrate von ca. 8 GB/s in jeder Richtung, also zusammen 16 GB/s.

Bei Chipsätzen für PCs im Niedrig-Kosten-Bereich oder für den Einsatz in tragbaren Geräten (*Laptops*, *Notebooks* usw.) ist der Graphikcontroller häufig in der North Bridge selbst integriert, die dann – wie bereits gesagt – als *Graphics Memory Controller Hub* (GMCH) bezeichnet wird.

¹⁰ Der Einsatz von zwei getrennten, aber verzahnt zusammen arbeitenden Graphikcontrollern wird insbesondere von Computerspielern geschätzt.

¹¹ d.h. ohne Anrechnung von eingefügten Steuerbits

1.3.3 South Bridge

Die Hauptaufgabe der *South Bridge* ist es, die Kommunikationsverbindungen und -vorgänge zwischen den unterschiedlichen Ein-/Ausgabeeinheiten zu regeln. Die Anforderungen an die Übertragungsgeschwindigkeiten dieser Verbindungen sind dabei – gemessen an den über die North Bridge laufenden Übertragungen – relativ gering. Als weitere Aufgabe stellt die South Bridge eine Reihe von internen Controllern zur Verfügung, die seit den Anfangstagen des PCs dazugehören, wie z.B. einen Zeitgeber/Zähler (*Timer/Counter*), eine batteriegepufferte Echtzeituhr (*Real Time Clock* – RTC) mit dem sog. CMOS-RAM zur Speicherung wichtiger Systemdaten sowie einen Tastatur-, DMA- (*Direct Memory Access*) und Interrupt-Controller (*Programmable Interrupt Controller* – PIC). Häufig bietet sie auch allgemein verwendbare digitale Ein-/Ausgabeleitungen (*General Purpose I/O* – GPIO), die der Anwender unter Programmkontrolle für beliebige Steuer- und Abfrageaufgaben einsetzen kann.

Für die Kommunikation mit den Ein-/Ausgabegeräten stellt die South Bridge eine ganze Reihe von Standard-Ein-/Ausgabeschnittstellen zur Verfügung, die wir nun kurz behandeln wollen.

USB-Schnittstelle

Der momentan noch übliche USB 2.0 ist eine Schnittstelle zwischen Computer und Peripheriegeräten (z.B. Drucker, Scanner usw.), die von einem Firmenkonsortium definiert wurde (u.a. Compaq, IBM, DEC, Intel, Microsoft) und als Ziel die einfache Erweiterbarkeit des PCs um unterschiedliche Peripheriegeräte ohne das früher übliche Kabelgewirr hatte. Eine ausführliche Beschreibung des USB finden Sie z.B. in [3]. Der Anschluss der Geräte geschieht beim USB über eine kostengünstige Schnittstelle mit einheitlichen billigen Steckern; die Übertragung läuft seriell über ein abgeschirmtes 4-Draht-Kabel, über das auch die Spannungsversorgung für Geräte mit niedrigem Leistungsbedarf ($< 500\text{ mA}$) geführt wird. An den so genannten „Wurzelknoten“ (*Host Controller, Root Hub*) können bis zu 127 Geräte angeschlossen werden, die in Form eines Baumes angeordnet sind: Die „Blätter“ werden von den Endgeräten, die Verzweigungen durch die *Hubs* (engl. für Nabe, Mittelpunkt, Kern...) gebildet. Das sind spezielle externe Geräte oder integrierte Einheiten mit einer Schnittstelle in Richtung des Wurzelknotens (*Upstream Port*) und bis zu acht Schnittstellen in Richtung der Endgeräte (*Downstream Ports*). Ein Hauptvorteil von USB liegt darin, dass neue Geräte bei laufendem Rechner und ohne Installation von Gerätetreibern (*Hot Plug and Play*) hinzugefügt werden können. Dazu überwacht ein *Hub* alle an ihn angeschlossenen Geräte und ihre Versorgungsspannung. Er informiert den *Host Controller* über alle Änderungen, z.B. über das Entfernen oder

Hinzufügen eines neuen Gerätes. Er kann jeden *Downstream Port* individuell aktivieren, deaktivieren oder zurücksetzen. Zwischen dem Wurzelknoten und einem Endgerät dürfen maximal sieben Kabelsegmente mit einer Gesamtlänge von 35 m, also maximal sechs *Hubs*, liegen.

Die USB-Schnittstelle überträgt Daten in drei Geschwindigkeiten, die in gemischter Form eingesetzt werden können: 1,5 Mbit/s (*Low Speed*), 12 Mbit/s (*Full Speed*) und 480 Mbit/s (*High Speed*), was im letzten Fall einer Rohdatenrate von maximal 60 MB/s entspricht. Die gesamte Übertragung im USB wird vom *Host Controller* gesteuert. Dazu fragt er in einer festgelegten Reihenfolge, aber mit unterschiedlicher Wiederholrate, alle angeschlossenen Geräte nach Übertragungswünschen ab. Dieses so genannte *Polling* geschieht in Zeitrahmen von 1 ms, die für die *High-Speed*-Übertragung noch einmal in acht 125- μ s-Rahmen unterteilt werden. Die Daten werden in Form von Datenblöcken, sog. Paketen, versendet. In jedem Rahmen gibt es reservierte Zeitschlitzte für unterschiedliche Datentypen und Übertragungen:

- Kontinuierlich und in Echtzeit anfallende Daten, die keine Unterbrechung und Verzögerung erlauben, werden in Form der *isochronen Übertragung*¹² ausgetauscht. Es findet keinerlei Fehlerüberwachung oder erneute Übertragung eines fehlerhaften Datenpaket statt. Beispiele für solchermaßen übertragene Daten sind Sprach-, Audio- und Video-Daten.
- Als sog. *Interruptdaten* werden nicht periodisch, spontan auftretende Datenmengen gesendet, die z.B. von der Tastatur oder der Maus stammen. Hier findet eine Fehlererkennung mit eventueller Wiederholung des Pakets statt.
- Zu den so genannten *Steuerdaten* (*Control Data*) zählen alle Daten zur Identifikation, Konfiguration und Überwachung der Geräte und der *Hubs*.
- Massendaten, d.h. größere Datenmengen ohne Echtzeitanforderungen, werden mit der *Bulk-Übertragung* transportiert. Hier findet eine Fehlererkennung mit eventuellen Wiederholungsversuchen statt. Typische Geräte für diese Übertragungsart sind Drucker, Scanner und Modems.

Im USB wird der Erhalt jedes Pakets vom Empfänger quittiert. Dazu hat der drei Möglichkeiten: Er kann den fehlerfreien Erhalt durch ein ACK-Quittungspaket bestätigen (*Acknowledge*) bzw. den fehlerhaften Fall durch das Auslassen dieser Quittung anzeigen. Durch ein NAK-Paket (*Non Acknowledge*) kann er mitteilen, dass momentan kein Senden oder Empfangen eines Pakets möglich ist. Durch ein STALL-Paket wird angezeigt, dass das angesprochene Gerät außer Betrieb oder ein Eingriff des *Host Controllers* nötig ist.

Viele PCs sind mittlerweile bereits mit der leistungsfähigeren **USB-3.0-Schnittstelle** (SuperSpeed) ausgestattet. Diese bietet eine höhere maximale Geschwindigkeit von 5 Mbit/s (also 625 MByte/s) im Vollduplex-Betrieb. Dazu findet eine Abkehr vom reinen Polling durch den Host-Controller statt.

¹² Isochron bedeutet sinngemäß: zum gleichen Zeitpunkt im Zeitrahmen auftretend.

Die mögliche Stromentnahme über den USB 3.0 wurde von bisher 100 mA auf 150 mA erhöht. Nach Anmeldung beim Controller können aber auch bis zu 900 mA (statt bisher 500 mA) über die USB-Leitungen entnommen werden. Für die realisierte Duplexfähigkeit mussten die USB-3.0-Stecker um weitere Kontakte ergänzt werden, darunter auch neue Masseanschlüsse. Jedoch wurde beim Entwurf der Stecker auf Abwärtskompatibilität geachtet, d.h. alte USB-2.0-Stecker passen in neue USB-3.0-Buchsen, die dazu tiefer aufgebaut sind.

FireWire-Schnittstelle (IEEE 1394)

Der IEEE-1394-Bus ist eine standardisierte Weiterentwicklung der FireWire-Schnittstelle der Firma Apple¹³. Typische Anwendungen liegen in den Bereichen Audio, Video und Multimedia, wo man ihn z.B. in Festplatten, CD-ROM- und DVD-Laufwerken und -Brennern findet. So wird er insbesondere als i.Link in CamCordern häufig eingesetzt. Der Standard IEEE 1394a sieht Übertragungsraten bis zu 400 Mbit/s vor, der Standard IEEE1394b solche bis 3200 Mbit/s. Dabei unterstützt er – wie der USB – neben der asynchronen auch die isochrone Übertragung. Bei der asynchronen Übertragung wird wiederum jedes Paket vom Empfänger durch ein Quittungspaket (ACK) beantwortet und im Fehlerfall automatisch wiederholt. Bei der isochronen Übertragung wird auf diese Quittierung verzichtet. Andere Gemeinsamkeiten mit dem USB sind die kostengünstigen Stecker und Verbindungskabel, die einfache Erweiterbarkeit des Bussystems, die (eingeschränkte) Spannungsversorgung über das Anschlusskabel, das Hinzufügen bzw. Entfernen von Geräten während des Betriebs (*Hot Plug and Play*)

Anders als eine USB-Vernetzung hat ein System, das auf dem FireWire basiert, keine Baumstruktur. Hingegen lässt der FireWire fast beliebige Netzstrukturen zu – solange dabei keine Schleifen auftreten. Alle Verbindungen sind Punkt-zu-Punkt-Verbindungen, wobei in den Endgeräten Verzweigungen realisiert werden können. Dabei kann jeder Knoten im FireWire bis zu 27 Anschlüsse (*Ports*) für weitere Knoten besitzen. Zur Kopplung von FireWire-Bussen können aber auch externe Einheiten, sog. *Repeater* oder Brücken (*Bridges*), eingesetzt werden. Insgesamt kann ein FireWire-System maximal 1023 Teilbusse mit jeweils höchstens 63 Knoten umfassen. Auf den Verbindungsleitungen können unterschiedliche Übertragungsgeschwindigkeiten verwendet werden. Die Knoten im FireWire sind Rechner oder Endgeräte, wobei Busse aber auch ohne Rechner arbeiten können. Das heißt, dass – anders als beim USB – eine Datenübertragung auch direkt zwischen Endgeräten stattfinden kann. Das Bussystem ist selbstkonfigurierend, d.h. nach dem Einschalten oder Rücksetzen ermitteln die Knoten selbst, wer von ihnen die

¹³ Vereinfachend bezeichnen wir im Weiteren den Bus meist nur noch als FireWire.

Funktion des Wurzelknotens (*Root Node*) wahrnehmen darf. Die maximale Entfernung zwischen zwei Knoten beträgt beim FireWire mit Kupferkabel-Verbindungen ungefähr 72 m, durch Einsatz einer Glasfaserverbindung können bis zu 1600 m überbrückt werden. Dabei dürfen zwischen zwei Knoten höchstens 16 Kabelsegmente liegen.

PCI-Bus-Schnittstelle

Der PCI-Bus (*Peripheral Component Interconnect*) war lange Zeit der am weitesten verbreitete Busstandard für Ein-/Ausgabekarten. Eine ausführliche Beschreibung findet sich in [3]. Der PCI-Bus wurde ursprünglich von Intel eingeführt. Später fand man aber PCI-Steckplätze bei allen Desktop-Computern, d.h. PCI wurde auch durch Chipsätze anderer Prozessorhersteller unterstützt. Der PCI-Bus überträgt die Daten entweder auf einem 32-Bit-Datenbus mit 33 MHz oder einem 64-Bit-Datenbus, der mit 66 oder 133 MHz (*extended PCI* bzw. PCI-X) getaktet wird. Dabei ist der PCI-X kein Bus im herkömmlichen Sinne mehr, sondern eine Punkt-zu-Punkt-Verbindung zum Anschluss einer einzigen Komponente. Aus den Kenndaten ergibt sich eine Datentransferrate von 532 MB/s bis zu 1,064 GB/s. Der PCI-Bus wird auf modernen Hauptplatinen nur noch aus Kompatibilitätsgründen als „Erblast“ (*Legacy*) implementiert und wird wohl in kurzer Zeit vom Markt verschwinden. Deshalb wollen wir ihn im Rahmen dieses Buches nicht weiter beschreiben.

PCIe-Schnittstelle

Das Grundkonzept des PCIe (auch: PCI-E) wurde bereits im Unterabschnitt 1.3.2 kurz beschrieben. Der PCIe wurde ab 2004 von der Firma Intel eingeführt; inzwischen werden seine Spezifikationen von einer Firmengruppe (*PCI Special Interest Group* – PCI-SIG) betreut, der fast alle namhaften Chiphersteller angehören. Beim PCIe handelt es sich um eine skalierbare, d.h. den Anforderungen des Einsatzbereichs anpassbare Anzahl von seriellen Punkt-zu-Punkt-Verbindungen, den sog. *Lanes* (Bahn, Weg), die die angeschlossenen Bausteine über schnelle Schalteinheiten mit dem Computersystem verbinden. Durch die Bündelung von mehreren parallelen PCIe-*Lanes* können leicht verschiedene, den jeweiligen Erfordernissen angepasste Übertragungsbandbreiten realisiert werden. Die Anzahl der Lanes einer Verbindung wird mit einem kleinem vorangestellten x angegeben. So sind also Systeme mit PCIe-x1, PCIe-x2 bis zu PCIe-x32 möglich. (Realisiert wurden im PC-Bereich bis heute jedoch nur Systeme mit bis zu PCIe-x16.) Durch die Anpassung der Anzahl der Lanes an die Erfordernisse einer Verbindung spart man u.U. sehr viele Verbindungsleitungen, denn insgesamt werden für eine

Lane nur vier Leitungen benötigt, je zwei für eine Richtung.¹⁴ Durch diesen drastischen Wegfall von Leitungen vereinfachen und verbilligen sich auch die Steckverbindungen. Die sind so realisiert, dass eine Erweiterungskarte mit n *Lanes* auch in einen Slot für m *Lanes* eingesetzt werden kann, solange $m > n$ ist.

Anders als beim Vorgänger, dem PCI-Bus, der die vom Prozessor ausgehenden Adressen und Daten direkt übertrug, werden im PCIe die Daten in Form von Paketen übertragen. Das sind Datenblöcke, die durch einen Paketkopf mit Quelladressen und Paketnummer sowie eine Prüfsumme am Ende (*Cyclic Redundancy Check* – CRC) ergänzt werden. Werden durch die Prüfsumme Übertragungsfehler angezeigt, so fordert der PCIe-Controller automatisch eine erneute Aussendung desselben Pakets an. Durch die Paketnummer wird sichergestellt, dass das erneut ausgesendete Paket vom Empfänger in die richtige Reihenfolge gesetzt werden kann, auch wenn bereits Folgepakete eingetroffen sind.

Im Gegensatz zum älteren PCI-Bus gibt es bei PCI-Express keine Einsteckplätze (*Slots*) an einem gemeinsamen Bus, sondern *geschaltete Ports*. Dies bedeutet, dass den einzelnen PCI-Express-Karten stets die volle Bandbreite zur Verfügung steht, da hier keine Zugriffskonflikte wie beim PCI-Bus auftreten können. Die South Bridge ist dabei in der Lage, simultan über mehrere PCIe-Verbindungen Daten und das in beiden Richtungen – vom bzw. zum Gerät – zu übertragen. Dazu verwendet sie einen sog. Kreuzschienenschalter (*Crossbar Switch*), der es erlaubt, der mehrere Quellen von Datenübertragungen mit ihren jeweiligen Zielen gleichzeitig verbinden kann.

Die PCIe-Spezifikation sieht auch vor, dass Einsteckkarten im laufenden Betrieb gewechselt werden können (*Hot Plugging*). Von dieser Fähigkeit können vor allem mobile Systeme wie Notebooks profitieren. Neben der im Unterabschnitt 1.3.2 erwähnten Verbindung von Graphikadapter mittels PCIe-x16 sind im Serverbereich durch den Maximalausbau von 32 *Lanes*, PCIe-x32 genannt, auch bidirektionale Hochgeschwindigkeitsverbindungen mit bis zu 16 GB/s pro Übertragungsrichtung, also insgesamt 32 GB/s möglich.

IDE-Schnittstelle

IDE ist eine standardisierte Schnittstelle zum Anschluss von nichtflüchtigen Speichermedien, wie Festplattenlaufwerken, CD-ROMs und DVD-Laufwerken. Sie ist auch unter dem Namen ATA (*Advanced Technology Attachment*) bekannt. Die Beschränkung auf Plattengrößen von 528 MB wurde durch die Erweiterung zu EIDE (Extended IDE) aufgehoben. Zur Zeit

¹⁴ Auf jedem Leitungspaar werden die Signale zur Erhöhung der Störsicherheit differenziell übertragen, d.h. mit jeweils entgegengesetztem logischen Pegel. Störungen wirken sich in der Regel auf beide Leitungen gleichartig aus und können daher durch die Differenzbildung der Pegel herausgefiltert werden.

beträgt die maximale Datenrate des EIDE-Busses 133 MB/s. Man muss allerdings beachten, dass diese Datenrate nur dann erreicht wird, wenn die Laufwerkselektronik die Daten bereits in ihrem internen (Cache-)Speicher hat. Die permanente Datenrate zwischen Festplatte und Laufwerkselektronik liegt deutlich unter dem o.g. Wert.

Das *IDE Interface* bietet zwei getrennte Schnittstellen („Kanäle“) für den Anschluss von Massenspeichergeräten, also Festplatten- (*Hard Disk Drive* – HDD), CD-ROM-, DVD-Laufwerken (*Digital Versatile Disk*) und natürlich die immer beliebter werdenden Programmiergeräte für CD-ROMs oder DVDs („Brenner“). Die Bezeichnung IDE (*Integrated Drive Electronics*) zeigt an, dass diese Schnittstelle Geräte mit integrierten Controllern verlangt und selbst keinerlei Ansteuerlogik zur Verfügung stellt. Die IDE-Schnittstellen im PC wurden von der ANSI (*American National Standards Institute*) standardisiert. Diese „genormten“ Schnittstellen wurden zuerst im IBM-AT-PC eingesetzt und tragen daher die Bezeichnung ATA (*Advanced Technology Attachment*). Die anschließbaren Geräte werden dementsprechend ATAPI-Geräte (*ATA Packet Interface*) genannt.

Die beiden o.g. Kanäle werden als primärer und sekundärer Kanal (*Primary, Secondary Channel*) bezeichnet und erlauben jeweils den Anschluss von bis zu zwei Geräten. Diese werden *Master* und *Slave* genannt. Sie müssen ihre Funktion durch kleine Steckbrücken auf ihrer Platine und ihre Platzierung in den entsprechenden Steckplatz zugewiesen bekommen. Jedes der maximal 4 IDE-Geräte kann unabhängig voneinander durch die IDE-Schnittstelle aktiviert oder deaktiviert werden. Die Deaktivierung versetzt die Leitungen des „abgeschalteten“ Gerätes in den (hochohmigen) *Tristate*-Modus. In diesem Modus kann das Gerät während des Betriebs des PCs aus dem Rechner entfernt bzw. (wieder) eingesetzt werden (*Hot Swap*).

Die ANSI hat im Laufe der Jahre eine ganze Reihe von Standards der ATA-Schnittstelle erstellt, die sich insbesondere durch die Übertragungsleistung der Schnittstelle unterscheiden. Üblich sind heute die Standards Ultra-ATA/66 und Ultra-ATA/100, die eine maximale (theoretische) Übertragungsrate von 66 bzw. 100 MB/s erreichen. Diese Raten werden jedoch nur im DMA-Modus erreicht, bei dem der Gerätecontroller als *Bus Master* selbst die Datenübertragung durchführt. (Daher stammt auch die häufig verwendete Bezeichnung: UltraDMA – UDMA.) Außerdem werden sie nur bei Lesezugriffen erreicht. Schreibzugriffe ermöglichen nur eine um ca. 10 % geringere Übertragungsrate (88,9 MB/s). Noch langsamer sind die sog. PIO-Zugriffe (*Programmed Input/Output*), bei denen der Prozessor jedes Datum selbst übertragen muss. Hier werden maximal 16 MB/s erreicht.

Die IDE-Schnittstellen können in zwei verschiedenen Modi arbeiten:

- Im sog. *Legacy Mode* müssen den Geräten der IDE-Kanäle bestimmte Eingänge des Interrupt-Controllers fest zugewiesen werden. Wie weiter unten beschrieben, sind dies IRQ14 und IRQ15. Außerdem werden die Steuer- und Statusregister ihrer Controller und festgelegten Adressen im

Ein-/Ausgabe-Adressbereich (*I/O Address Space*) des Prozessors angesprochen.

- Im *Native Mode* werden sie über spezielle Register in ihrem Konfigurationsbereich definiert und benötigen daher keine festgelegten Interrupt-Eingänge und I/O-Adressen.

Serielle ATA-Schnittstelle (SATA, eSATA)

Das *Serial ATA Interface* (*SATA Interface*) wurde aus dem oben beschriebenen IDE/ATA-Standard entwickelt und dient wie dieser dem Datenaustausch mit nichtflüchtigen Speichermedien. Um die Zahl der benötigten Adern zu verringern und damit die Kabelführung zu vereinfachen, wurde ein serielles Übertragungsprotokoll eingeführt. Kompatibilität wird durch SATA/ATA-Adapter erreicht, über die SATA-Geräte auch an der IDE-Schnittstelle eingesetzt werden können (*Standard-ATA Emulation*).

Während die ersten SATA-Spezifikationen eine Datenrate von 150 MB/s (SATA I) bzw. 300 MB/s (SATA II) vorsahen, arbeitet der neue Standard SATA III¹⁵ bereits mit einer Datenrate von 600 MB/s, was einer Brutto-Datenrate von ca. 6 Gbit/s entspricht.¹⁶ Um diese hohen Datenraten sicher zu erreichen, benutzt man die Signalübertragung mit dem so genannten LVDS-Verfahren (*Low Voltage Differential Signalling*), das zur Unterdrückung von Fehlern, die durch elektrische Störungen hervorgerufen werden, die Signale über Leitungspaare mit entgegengesetztem Signalpegel und niedrigen Spannungsdifferenzen überträgt.

Das *SATA Interface* stellt ebenfalls zwei unabhängig arbeitende Schnittstellen (gekennzeichnet durch 0 bzw. 1 in den Signalbezeichnungen), die jede für sich aktiviert und deaktiviert werden kann. Diese Schnittstellen werden durch zwei unabhängige Controller im *Bus Master*-Modus betrieben, d.h. sie können selbständig auf den Hauptspeicher zugreifen. Im Unterschied zum Standard-ATA sind jedoch nur Punkt-zu-Punkt-Verbindungen, also keine Master/Slave-Konfigurationen möglich. Insgesamt können somit nur bis zu zwei SATA-Geräte, d.h. Festplattenlaufwerke (*Hard-Disk Drive* – HDD) oder ATAPI-Geräte, betrieben werden. Im Gegensatz zur parallelen ATA-Schnittstelle ist mit SATA ein Wechsel des Speichermediums im laufenden Betrieb möglich (*Hot Plugging*). Jedoch ist ein unerwartetes Entfernen bzw. Hinzufügen eines SATA-Gerätes nicht zulässig; es muss erst durch Software vorbereitet werden, indem z.B. das Betriebssystem die gewünschte Schnittstelle in den *Power-Down*-Modus schaltet.

Der SATA-Standard sieht nur den Anschluss von Geräten innerhalb des PCs vor. Die eingesetzten Kabel sind deshalb nicht gegen elektromagnetische Störungen abgeschirmt und die Stecker nicht für den Anschluss von externen

¹⁵ offizielle Bezeichnung: SATA 6Gb/s nach der SATA Revision 3.0

¹⁶ Wird jedoch eine SATA-Schnittstelle im PIO-Modus betrieben, so reduziert sich die Übertragungsrate auf maximal 16 Mbit/s.

Geräten ausgelegt. Um auch den externen Anschluss von Geräten über die SATA-Schnittstelle zu ermöglichen, wurden im neuen Standard auch Vorgaben zu externen Steckern und Anschlusskabeln gemacht. Diese Vorgaben definieren nun die externe serielle ATA-Schnittstelle – *External Serial ATA* oder kurz eSATA.

Audio-Schnittstelle

Heutzutage finden sich auf den PC-Hauptplatinen Audio-Schnittstellen, die zwei verschiedenen Standards genügen.

AC'97-Schnittstelle: Die AC'97-Schnittstelle (*AC'97 Link*) ist der ältere Standard.. Sie bietet dem PC-Entwickler die Möglichkeit, sehr kostengünstig Audio- und Modemfunktionen schon auf der Hauptplatine (*Onboard Sound*) zu realisieren und auf den Einsatz einer teuren Audio-Steckkarte (*Sound-card*) zu verzichten. Für diese Funktionen ist lediglich ein so genannter Codec (Codierer/Decodierer, besser: Converter/Deconverter) erforderlich, der in einem einzigen Baustein einen Analog/Digital- sowie einen Digital/Analog-Wandler für mehrere (Stereo-)Kanäle zur Verfügung stellt. Bei dieser einfachen Lösung muss jedoch der Prozessor die „Rechenaufgaben“ zur Erzeugung von Audio-/Modem-Signalen übernehmen, die auf einer „Soundkarte“ von Spezialchips geleistet werden. Die AC'97-Schnittstelle überträgt einzelne Stereo-Signale mit einer Abtastrate von bis zu 96 kHz und einer Datenbreite von 20 Bits. Im Mehrkanal-Betrieb in bis zu sechs getrennten Zeitkanälen werden im Zeit-Multiplexverfahren (*Time Division Multiplex Access* – TDMA) über eine einzige Leitung noch 48 kHz erreicht. So können maximal sechs verschiedene Codecs über die AC'97-Schnittstelle mit Ausgabedaten versorgt werden. Man spricht in diesem Fall von bis zu sechs „Ausgabekanälen“. Bei den Codecs kann es sich um Audio-Codecs (AC), um Modem-Codecs (MC) oder aber um kombinierte Audio/Modem-Codecs (AMC) handeln. Die AC'97-Schnittstelle verfügt häufig über einen gesonderten Steckplatz, der mit AMR (*Audio/Modem Riser Slot*) bezeichnet wird und z.B. eine Steckkarte zum Anschluss des PCs an das Telefonnetz aufnehmen kann.

High Definition Audio: Die Audio-Schnittstelle moderner Hauptplatinen genügt den Anforderungen des *High Definition Audio-Standards* (HD-Audio), der im Jahr 2004 erlassen wurde und auf Entwicklungen der Firma Intel beruht. Längerfristig soll der HD-Audio-Standard die oben beschriebene AC'97-Schnittstelle ablösen. Audio-Controller nach dem HD-Audio-Standard übertragen Stereo-Signale mit einer Abtastrate von 192 kHz und einer Datenbreite von 32 Bits in bis zu acht getrennten Zeitkanälen (im sog. Zeitmultiplex-Verfahren, *Time Division Multiplex Access* – TDMA). Gibt man diese Kanäle auf sieben Lautsprechern und einem speziellen Bass-

Lautsprecher (*Subwoofer*) aus, so erhält man z.B. eine Rundum-Beschallung, die als *7.1 Surround* bezeichnet wird. Der Standard sieht aber auch die simultane Ausgabe von zwei oder mehr unabhängigen Audio-Strömen vor. So kann man sich z.B. mit fünf Lautsprechern und dem Bass-Lautsprecher (*5.1 Surround*) für den Musikgenuss zufrieden geben und die beiden restlichen Kanäle z.B. für eine zweite (Sprach-)Übertragung über einen Kopfhörer benutzen. Auch andere Aufteilungen sind möglich, um so z.B. zwei verschiedene Ausgaben in zwei verschiedenen Räumen zu unterstützen.

Der HDA-Standard erweitert auch die Möglichkeiten zur gleichzeitigen Aufzeichnung von Tönen, Geräuschen, Sprache und Musik durch eine Vielzahl von unabhängigen Mikrophonen (*Array Microphone*). Dadurch wird insbesondere die Spracherkennung und die „verständliche“ Übermittlung von Sprache über das Internet (*Voice over IP*) unterstützt.

Netzwerkschnittstelle (LAN-Link)

PCs sind in der Regel an ein Lokales Netz (*Local Area Network* – LAN) angeschlossen. Dabei handelt es sich meist um das so genannte *Ethernet*, bei dem alle Rechner auf dasselbe physische Verbindungsmedium zugreifen und daher Kollisionen auftreten können. Zur Strukturierung der verwendeten Hardware- und Software-Technologien und Protokolle wird die Netzwerkschnittstelle eines Rechners gewöhnlich in mehrere Schichten eingeteilt („Schichtenmodell“). Die beiden unteren Schichten des Schichtenmodells sind die MAC-Schicht (*Media Access Control Layer*) und die PHY-Schicht (*Physical Layer*). Die MAC-Schicht regelt hauptsächlich den kollisionsfreien Zugriff auf das gemeinsam genutzte Medium. Die unterste PHY-Schicht beschäftigt sich mit den mechanischen und physikalischen Eigenschaften des Verbindungsmediums und um die bitweise Datenübertragung über diese Verbindung.¹⁷

Die Netzwerkschnittstellen-Hardware eines PCs besteht in der Regel aus zwei Controllern, die die erwähnten beiden unteren Schichten des Ethernet realisieren, also einem MAC- und einem PHY-Controller. Je nach Aufbau des Chipsatzes existieren heute drei Lösungsansätze:

- Beide Controller sind extern in getrennten Bausteinen realisiert. Die Ankopplung des MAC-Controllers an die South Bridge geschieht z.B. über den oben beschriebenen PCIe. Die Verbindung zwischen den Controllern geschieht über eine dedizierte schnelle Schnittstelle, die z.T. mehrfach für unterschiedliche Übertragungsgeschwindigkeiten vorhanden sein muss.
- Beide Controller sind zusammen in einem externen Ethernet-Controller, also außerhalb der South Bridge, realisiert. Auch in diesem Fall bietet die South Bridge eine spezielle schnelle Schnittstelle zum Ethernet-Controller oder verwendet dazu den PCIe.

¹⁷ Sie wird daher im Deutschen als Bitübertragungsschicht bezeichnet.

- Der MAC-Controller ist in der South Bridge integriert, der PHY-Controller muss extern als eigenständiger Baustein ergänzt werden. In diesem Fall werden beide Bausteine durch die im ersten Fall beschriebenen dedizierten Schnittstellen verbunden, die nun von der South Bridge zur Verfügung gestellt werden.

Moderne PCs unterstützen verschiedene Ethernet-Standards, die den Netzwerkanschluss über verdrehte Paare von Kupferkabeln (*Twisted Pair*) vornehmen: Das Standard-Ethernet überträgt mit 10 Mbit/s, das Fast-Ethernet mit 100 Mbit/s und das Gigabit-Ethernet mit 1000 Mbit/s (1 Gbit/s). Ethernet überträgt die Daten in Form von Paketen, die – neben einer jeweils 6 Byte langen Empfänger- und Sender-Adresse und einer 4 Byte langen Prüfsumme – bis zu 1500 Byte von Benutzerdaten enthalten können.

SMBus

Der SMBus (*System Management Bus*) verbindet wichtige Komponenten der Hauptplatine miteinander. Eine seiner Aufgaben ist es, Steuer- und Überwachungsinformationen zwischen den Brückenbausteinen und den Speichermodulen zu übertragen. Er unterstützt außerdem Überwachungsfunktionen (*Monitoring*) der folgenden Bauteile und physikalischen Größen: Batterie, Lüfter, Temperaturen, Betriebsspannungen für alle Komponenten, PCI-Takt, unterschiedliche Betriebszustände (*Power-Down*-Modi, Prozessor-Stop-Modus) usw. Die Ergebnisse seiner Überwachungsfunktion kann er durch Meldung an den Prozessor weiterreichen. Der SMBus erlaubt häufig auch einem externen Mikrocontroller den Zugriff auf bestimmte Systemkomponenten.

Der SMBus ist ein langsamer serieller Bus, der aus dem von der Firma Philips vor ca. 20 Jahren entwickelten I²C-Bus zur Verbindung von Integrierten Bausteinen (ICs) hervorging. Er besitzt ein einfaches Übertragungsprotokoll.

LPC-Schnittstelle

Die wichtigsten Peripheriekomponenten, wie Tastatur- und Maus-Controller usw. sind häufig in einem externen Baustein, dem *Super-I/O*-Baustein (vgl. Abbildung 1.6), untergebracht, der an der *South Bridge* über die LPC-Schnittstelle (*Low Pin Count*) mit insgesamt neun Signalen angeschlossen ist. Adressen und Daten werden über vier Leitungen nacheinander und in mehreren Teilen übertragen (Multiplexbetrieb). Von den restlichen Signalen dient jeweils eines zur Übermittlung des Taktes bzw. zur Kennzeichnung einer laufenden Übertragung. An der LPC-Schnittstelle ist sehr oft auch das BIOS-ROM im *Firmware Hub* (FWH) angeschlossen.

Interrupt-Controller

In einem komplexen Mikrorechner-System wie einem PC werden immer wieder Unterbrechungsanforderungen durch die Peripheriekomponenten an den Prozessor gestellt. Diese so genannten *Interrupts* werden dem Prozessor in regelmäßigen oder unregelmäßigen zeitlichen Abständen über spezielle Leitungen übermittelt. Sie zeigen dem Prozessor z.B. an, dass eine Komponente Daten übertragen will oder die Ausführung einer Routine zur Durchführung einer bestimmten Aktion wünscht. Da dazu nicht jeder Komponente eine eigene Leitung zur Verfügung gestellt werden kann, muss ein *Interrupt-Controller* eingesetzt werden, der insbesondere auch die Aufgabe hat, gleichzeitig vorliegende Unterbrechungsanforderungen in eine geeignete Prioritäten-Reihenfolge zu setzen und dem Prozessor über die Anforderungsquelle mit momentan höchster Priorität zu informieren.

Der in der *South Bridge* integrierte Interrupt-Controller unterstützt drei unterschiedliche Möglichkeiten, Unterbrechungsanforderungen der Peripheriekomponenten anzunehmen und zum Prozessor weiterzuleiten.

1. Variante: In der ersten Variante, die als „Altlast“ (*Legacy*) vom IBM AT-PC der 80er Jahre des letzten Jahrhunderts übernommen wurde, besitzt der Interrupt-Controller 15 Eingänge (IRQ0,...,IRQ15 – ohne IRQ2)¹⁸, die den verschiedenen Interrupt-Quellen zugeordnet werden. Dazu enthält die *South Bridge* eine besondere Schaltung, *IRQ-Router* genannt, die die Unterbrechungssignale der internen Brückenkomponenten mit den oben erwähnten Eingängen IRQ0,...,IRQ13 (ohne IRQ2) des Interrupt-Controllers verbindet. Die Eingänge IRQ14 und IRQ15 werden über zwei Anschlüsse der *South Bridge* nach Außen geführt. Sie sind für die Geräte reserviert, die an den IDE-Schnittstellen betrieben werden, und zwar IRQ14 für den primären Kanal (*Primary Channel*, s.u.) und IRQ15 für den sekundären Kanal (*Secondary Channel*). Akzeptierte Unterbrechungsanforderungen reicht der Interrupt-Controller über ein Ausgangssignal (INTR) an den Prozessor weiter.

2. Variante: Bei der zweiten Variante werden Unterbrechungsanforderungen seriell über eine einzelne (bidirektionale) Leitung (*Serial Interrupt Request – SERIRQ*) an den Interrupt-Controller weitergereicht. Dazu werden über SERIRQ Zeitrahmen aus 32 Schlitzen (*Slots*), eingerahmt durch einen Start- und einen Stop-Slot, ausgesandt. Jeder Slot kann einer möglichen Interrupt-Quelle zugewiesen werden. Einige Slots sind für spezielle Anforderungen reserviert, darunter die oben genannten 15 Unterbrechungssignale IRQ0,...,IRQ15 (ohne IRQ2). Die restlichen Slots sind frei belegbar.

Jede Komponente, die eine Unterbrechungsanforderung an den Interrupt-Controller stellen will, markiert dies in dem ihr zugewiesenen Slot

¹⁸ Die Abkürzung IRQ steht für *Interrupt Request*.

des SERIRQ-Zeitrahmens. Der Interrupt-Controller wertet den empfangenen Zeitrahmen aus, entscheidet über die Unterbrechungsanforderung mit der momentan höchsten Priorität und leitet sie über das o.g. Signal INTR an den Prozessor weiter.

3. Variante: In der dritten Variante benutzt der Brückenbaustein eine Erweiterung des oben beschriebenen Interrupt-Controllers, die als APIC (*Advanced Programmable Interrupt Controller*) bezeichnet wird. Dieser Controller unterstützt bis zu 24 verschiedene Interrupt-Quellen, die seinen Eingängen IRQ0,...,IRQ23 zugewiesen werden. Dabei werden die IDE-Geräte – wie oben beschrieben – wiederum mit den externen Eingängen IRQ14 und IRQ15 verbunden. Die IRQ-Eingänge IRQ16,...,IRQ23 werden insbesondere von den internen Komponenten (USB, LAN Link usw.) benutzt.

Der APIC unterscheidet sich von der oben beschriebenen „Legacy“-Variante insbesondere durch die Art, wie die Unterbrechungsanforderungen Komponenten an den Controller herangeführt werden. Dies geschieht nicht durch die Eingangssignale des Interrupt-Controllers, sondern durch „normale“ Schreibzugriffe auf den Hauptspeicher. Dazu wird dem APIC eine bestimmte Speicherzelle zugewiesen, die als *IRQ PIN Assertion Register* (IRQPA) bezeichnet wird. In dieses „Register“ schreibt eine „unterbrechungswillige“ Komponente ihre spezifische Kennung, die so genannte Interrupt-Vektornummer (IVN).

Der APIC liest die IVN aus dem IRQPA-Register, aktiviert den entsprechenden Interrupt-Eingang und löscht sofort das IRQPA-Register für die nächste Anforderung. Er entscheidet dann über die Prioritäten der aktuellen Unterbrechungsanforderungen. Dabei wird die Prioritäten-Reihenfolge nicht durch die Nummern der zugeordneten Controller-Eingänge festgelegt, sondern sie kann frei zugeordnet werden. Die Interrupt-Vektornummer wird einer Komponente bei der Systeminitialisierung zugewiesen und in ein bestimmtes Register ihres Konfigurationsbereichs eingetragen.

Neben der oben beschriebenen „konventionellen“ Methode, die Unterbrechungsanforderung über das Signal INTR an den Prozessor weiterzureichen, kann der APIC ein weiteres Verfahren dazu anwenden – sofern dieses vom Prozessor unterstützt wird. Dieses wird als *FSB Interrupt Delivery* bezeichnet. Hierbei schreibt die *South Bridge* eine „Unterbrechungsnachricht“ (*Interrupt Message*) in bestimmte Speicherzellen des Hauptspeichers. Durch diese Nachricht und ihre Zieladresse wird die Interrupt-Quelle spezifiziert – für den Einsatz in einem Mehrprozessorsystem aber auch der angesprochene Prozessor.

Der Prozessor liest regelmäßig die Speicherzellen und informiert sich dadurch über eventuelle Unterbrechungswünsche. Aus der gefundenen IVN ermittelt er die Startadresse der verlangten Unterbrechungsroutine (*Interrupt Service Routine* – ISR) und führt sie aus.

Die Vorteile des beschriebenen APIC-Verfahrens sind:

- Die Durchführung eines speziellen Buszugriffes zur Ermittlung der IVN (*Interrupt Acknowledge Cycle*) ist nicht nötig.

- Es wird keine zusätzliche Busleitung zum Prozessor verlangt.
- In einem System sind mehrere APICs mit eigenen Interrupt-Vektoren einsetzbar – insbesondere in einem Mehrprozessorsystem.

Prozessor- und Systemsteuerung

Die Prozessor- und Systemsteuerung übernimmt üblicherweise die Funktion der Steuerung und Regulierung des Energieverbrauchs im PC. Diese Komponente wird *Advanced Configuration and Power Interface* (ACPI) genannt. Durch ACPI kann das Betriebssystem – nach Vorgaben im BIOS – vielfältige Aufgaben übernehmen, die werbewirksam als TCO-Funktionen (*Total Cost of Ownership*) bezeichnet werden, da sie helfen sollen, die Gesamtkosten für den PC-Besitzer zu vermindern:

- Steuerung und Überwachung der Systemkomponenten,
- Steuerung der Leistungsaufnahme (*Power Management* – ACPI/ APM) durch verschiedene Stromspar-Systemzustände (*Low-power States*) und der Deaktivierung nicht gebrauchter Komponenten und Schnittstellen (AC'97 bzw. HDA für Audio und Modem, ATA/IDE, SATA, LAN, USB, SMBus),
- Takterzeugung und -überwachung,
- Systemdiagnose und Meldung von Fehlern; dazu gehören z.B. ECC-Fehler, aber auch Warnungen, wenn das PC-Gehäuse geöffnet wird (*Intruder Detect*),
- Behebung von Systemblockaden, zu deren Erkennung die Zeitgeber-Bausteine (*Timer*) verwendet werden.

Bei heutigen Hauptplatinen werden die unterschiedlichen Komponenten der South Bridge in einem besonderen Platinenbereich mit einem Steckermodul verbunden, das durch eine Aussparung im Gehäuse nach Außen zugänglich ist. Die Lage dieses Steckermoduls wurde bereits in Abbildung 1.4 gezeigt. Abbildung 1.7 zeigt für die dort dargestellte Hauptplatine Intel DX48BT2 die „Außenansicht“ des Steckermoduls mit den verschiedenen Anschlüssen der beschriebenen Schnittstellen.

Das gezeigte Steckermodul bietet acht USB-Anschlüsse. Die zugehörige Hauptplatine bietet, wie fast alle modernen Hauptplatinen, darüber hinaus noch weitere, leichter zugängliche USB-Anschlüsse auf der Frontplatte oder einer Seite des PC-Gehäuses. Die Kennung RJ45 in der Abbildung steht für den Anschluss der Netzwerkschnittstelle, zusätzlich gekennzeichnet durch die Skizze eines kleinen Rechnernetzes. Lautsprecher, Mikrophon und *Line In*¹⁹ kennzeichnen die Anschlüsse der Audio-Schnittstelle. Auch diese ist häufig zusätzlich im vorderen Bereich des PC-Gehäuses zugänglich.

¹⁹ *Line In* bezeichnet die Steckbuchse für ein Audio-Eingabegerät, das – anders als das Mikrophon – ohne Vorverstärker auskommt, also z.B. der Ausgang einer Stereoanlage.

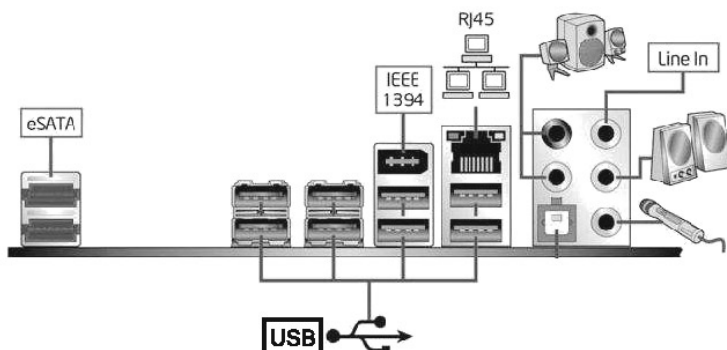


Abb. 1.7 Ein Schnittstellenmodul.

1.4 Prozessoren für Personal Computer

Die im PC-Bereich eingesetzten Mikroprozessoren werden bekanntermaßen als x86-kompatible Prozessoren bezeichnet. Über die Komponenten dieser Prozessoren haben Sie bereits im Band 2 [29] dieser Buchreihe eine Menge von Details erfahren, die wir hier nicht mehr wiederholen werden. Dort wurde auch die allgemeine Architektur dieser Prozessorklasse beschrieben. Gemeinsam ist allen modernen PC-Prozessoren, dass sie sowohl 32-bit-Daten wie auch 64-bit-Daten direkt verarbeiten, also als 32-bit- bzw. 64-bit-Prozessoren verwendet werden können. Dazu implementieren sie eine Architektur Erweiterung, die zunächst von AMD unter dem Begriff AMD64 eingeführt wurde (vgl. Unterabschnitt 1.4.1.1).

Über die Leistungsfähigkeit der x86-Prozessoren entscheidet nicht zuletzt die komplexe virtuelle Speicherverwaltung und ihre Unterstützung durch die Speicherverwaltungseinheit (*Memory Management Unit* – MMU). Selbst eine oberflächliche Beschreibung der virtuellen Speicherverwaltung und ihrer vielfältigen Funktionen würde den Rahmen dieses Kapitels sprengen. Wir werden diesem Thema daher das folgende Kapitel 2 vollständig widmen.

1.4.1 Prozessoren der Firma AMD

1.4.1.1 AMD64

Die modernen Prozessoren der Firma AMD, die in PCs zum Einsatz kommen, besitzen die so genannte AMD64-Architektur, die eine Erweiterung der seit den 80er Jahren des letzten Jahrhunderts verwendeten 32-bit-Architektur der x86-Prozessoren auf 64 bit darstellt. Die Erweiterung besteht hauptsächlich

in der „Verlängerung“ der Register auf 64 bit, aus der Verdopplung der Anzahl der Architekturregister von 8 auf 16 sowie dem Hinzufügen einiger weiterer Register (vgl. Abbildung 1.10 in Abschnitt 1.4.2). Die Verdopplung der Länge der Adressregister auf 64 Bit erweitert den physikalischen Adressraum von 4 GB auf 2^{64} Byte. Die Fähigkeiten der Multimedia-Rechenwerke (3Dnow!) wurden durch Implementierung der SSE-Befehle (*Streaming SIMD Extension*) der Firma Intel verbessert. Zur Realisierung der genannten Erweiterungen musste die benötigte Chipfläche moderat nur um ca. 5 % vergrößert werden.

Die Kompatibilität der AMD64-Architektur zur 32-bit-x86-Architektur wurde dadurch gewahrt, dass die 32-bit-Befehle beibehalten und lediglich durch neue 64-bit-Befehle ergänzt wurden. Dazu wurden einige der „überflüssigen“, d.h. selten gebrauchten, 1-byte-Befehle der x86-Architektur geopfert. Sie dienen in der AMD64-Architektur nun als Präfixe, also vorangestellte Kennungen, für die neuen 64-bit-Befehle.

Die AMD64-Architektur sieht drei verschiedene Betriebsarten vor:

- im 32-bit *Native Mode* („angeborene Betriebsart“) arbeitet ein AMD64-Prozessor wie ein herkömmlicher 32-bit-Prozessor, kann aber vom zusätzlichen SSE-Befehlssatz und einer eventuellen Hardwareerweiterung (z.B. einer schnelleren Speicherschnittstelle und größeren internen Caches) profitieren.
- im 64-bit *Native Mode* werden die neuen 64-bit-Befehle (mit ihren Präfixen) und die Registererweiterungen verwendet. Zur Erreichung der vollen Leistungsfähigkeit muss natürlich auch ein entsprechendes 64-bit-Betriebssystem²⁰ eingesetzt werden.
- im 64-bit-Kompatibilitätsmodus kann solch ein 64-bit-Betriebssystem durch die Verwendung spezieller Systemsoftware auch 32-bit-Software ausführen.

32-bit-Software kann ohne jegliche Veränderung oder Anpassung auch auf der AMD64-Architektur laufen; wegen der o.g. Erweiterungen wird sie jedoch z.T. sogar schneller ausgeführt. Programme im 64-bit-Modus werden durch die benötigten Präfixe im Mittel zwar 15 % länger, führen ihre Aufgaben aber – nach AMD-Angaben – bis zu 20 % schneller aus.

1.4.1.2 Die AMD-Mikroarchitektur K10

Der Prozessorkern der K10-Mikroarchitektur – von AMD selbst als *Family 10h* bezeichnet – unterstützt die eben beschriebene AMD64-Architektur, d.h. es handelt sich um einen 32/64-bit-Prozessor (s. Abbildung 1.8). Er realisiert eine zweistufige Cache-Hierarchie: Der lokale L2-Cache (*Level-2 Cache*) ist 512 kB groß und enthält als „unified“ Cache sowohl Befehle als

²⁰ d.h. ein Betriebssystem, das selbst für die 64-bit-Erweiterung geschrieben wurde,

auch Daten. Er arbeitet mit der vollen Prozessor-Taktgeschwindigkeit als so genannter *Victim Cache*, das heißt, er enthält nur Datenblöcke, die aus dem L1-Cache verdrängt worden sind. Über jeweils 256 bit breite Datenpfade werden Befehle und Daten in 64 kB große, getrennte L1-Caches geschrieben, dem I-Cache (*Instruction Cache*) bzw. D-Cache (*Data Cache*). Beide Caches sind als zweifach assoziative Caches (*2-Way set-assoziative Caches*) mit 64 Byte großen Einträgen (*Cache Lines*) ausgeführt. Der Daten-Cache arbeitet nach dem Rückschreibverfahren (*Write-Back Cache*) und setzt als Verdrängungsstrategie das LRU-Verfahren (*Least-recently used*) ein. Die Adressierung von Befehlen und Daten im L2-Cache wird durch eigene Übersetzungstabellen mit jeweils 512 Einträgen, den *Translation Lookaside Buffers* (ITLB, DTLB) unterstützt. (Auf ihre Realisierung wird in Kapitel 2 eingegangen.) Zwei weitere Adressen-Übersetzungstabellen, L1-ITLB und L1-DTLB, die 48 Einträge umfassen, erleichtern den Zugriff auf Befehle und Daten in den L1-Caches. Die TLBs unterstützen Seitengrößen von 2 KB bis zu 1 GB.

Um Fehlzugriffe auf die L1-Caches möglichst zu vermeiden, verfügt der K10-Kern über je eine eigene „Vorab-Ladeeinheit“ (*Prefetcher*) für Befehle und Daten, die im Bild mit IPF bzw. DPF bezeichnet sind: Wenn nach einem erfolglosen Zugriff auf den L1-Cache (*Cache Miss*) ein Eintrag aus dem L2-Cache angefordert wird, holt der entsprechende Prefetcher nicht nur den angeforderten 64-Byte-Eintrag aus dem L2-Cache (oder sogar dem Arbeitsspeicher), sondern auch die 64 Byte des folgenden Cache-Eintrags. Während des Ladens von neuen Befehlen aus dem Speicher werden automatisch zusätzliche Decodier-Informationen (*Predecode Bits*) erzeugt und mit den geladenen Befehlen im L1-Befehls-Cache abgelegt. Darunter sind z.B. Kennungen für den Beginn und das Ende der geladenen x86-Befehle, die ja als komplexe CISC-Befehle (*Complex Instruction Set Computer*) zwischen einem und 16 byte lang sein können. Diese Informationen helfen der Befehls-Ladeeinheit (*Fetch Unit*), die unterschiedlich langen Instruktionen zu erkennen und zu trennen.

Diese *Ladeeinheit* überträgt die auszuführenden Befehle aus dem 64 kB großen I-Cache (*Instruction Cache*) in den 32 byte großen *Fetch Buffer* in der Decoder-Einheit (*Decode Unit*). Dieser Puffer kann in einem einzigen Takt über einen 256 bit breiten Datenpfad gefüllt werden. Weitere 152 bit dienen der Fehlerüberprüfung durch Paritätsbits oder tragen die oben erwähnten zusätzlichen Vorab-Decodierinformationen, die bereits beim Eintrag der Befehle in den I-Cache gewonnen wurden. Diese Bits werden von der Befehls-Ladeeinheit ausgewertet. Aufgrund dieser Bits entscheidet sie, welchen der vorhandenen Decodern (s.u.) die einzelnen Befehle zugeordnet werden. Wird unter den geladenen Befehlen ein Befehl erkannt, der zu einer Änderung des Programmflusses führt, also z.B. ein Sprung- oder Verzweigungsbefehl, ein Unterprogrammaufruf oder -rücksprung, so wird die Sprungziel-Vorhersage-Einheit (*Branch Prediction Unit*) aktiviert. Diese umfasst einen *Branch Target Buffer* (BTB) mit maximal 2048 Einträgen für die Adressen von Sprungbefehlen und ihren Zielen sowie eine Tabelle mit 16384 Zwei-Bit-Zählern

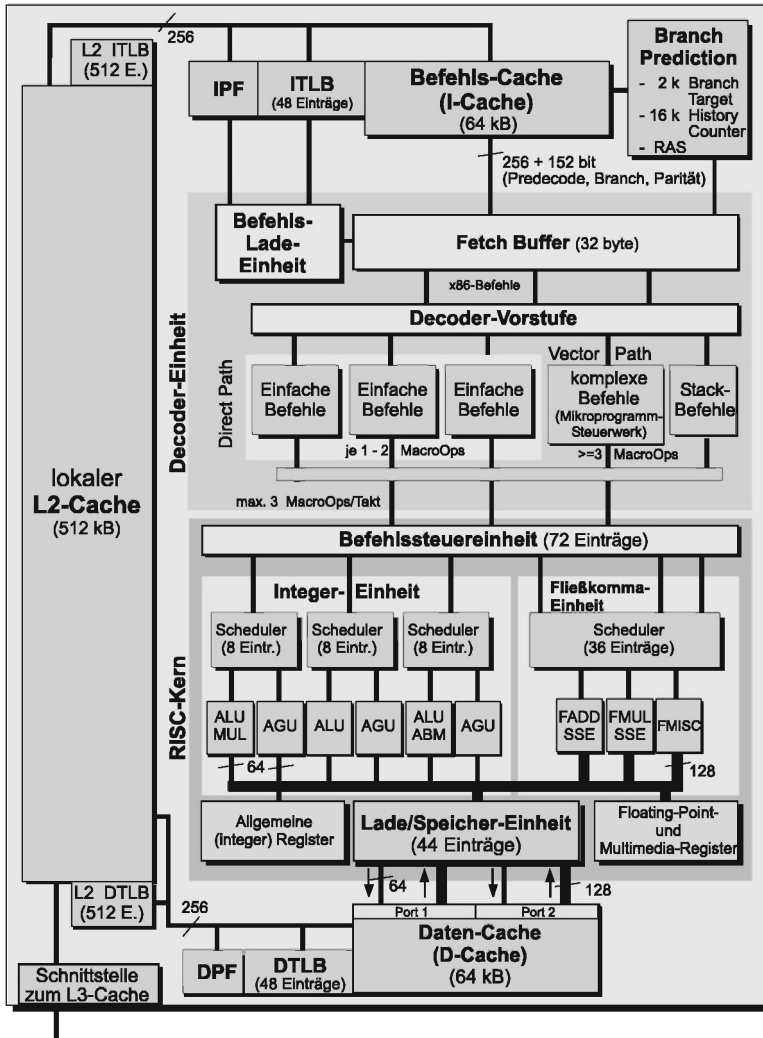


Abb. 1.8 Der Aufbau der AMD K10-Mikroarchitektur

(*Global History Bimodal Counter* – GHBC), die Auskunft über die „Vorgeschichte“ der Befehle speichern. Der Zugriff auf den BTB und die GHBC-Tabelle erfolgt parallel zum Zugriff des Befehlsholers auf den L1-Befehls-Cache. Als Besonderheit ermöglicht die Sprungziel-Vorhersage-Einheit auch die Vorhersage von Verzweigungen und Sprüngen mit indirekter Adressierung, bei denen also die Zieladresse erst zur Laufzeit des Programms, z.B. aus Register- und Speicherinhalten, berechnet wird. Dazu besitzt sie eine Tabelle mit 512 Einträgen. Die Vorhersage von Zieladressen indirekter Sprünge hat zu einer Verringerung falsch vorhergesagter Sprünge von bis zu 12% ge-

führt. Die richtige Vorhersage von Sprüngen führt beim AMD-K10 durch das Holen des Zielbefehls zu einer Verzögerung von nur einem Takt; aus fehlerhaften Vorhersagen resultieren jedoch mindestens zehn Takte, die benötigt werden, um die Pipeline zu leeren und die Ausführung an der richtigen Stelle fortzusetzen. Zur kurzzeitigen Ablage der Rücksprungadressen bei Unterprogrammaufrufen ist ein Hardware-Stack mit 24 Einträgen vorhanden, der so genannte *Return Address Stack* (RAS), auf den erheblich schneller zugegriffen werden kann, als auf einem im Arbeitsspeicher angelegten Stack.

Der K10-Kern ist dreifach superskalar und kann also in jedem Taktzyklus mit der Bearbeitung von bis zu drei x86-Befehlen beginnen, die parallel von der *Decoder-Einheit* aus dem *Fetch Buffer* in einen Vor-Decoder übertragen werden. Dieser stellt zunächst fest, ob es sich bei den Befehlen um einfache oder komplexe Instruktionen bzw. Instruktionen zur Verwaltung des Stacks handelt, und weist diese dann einem geeigneten Decoder zu. Für einfache Befehle stehen auf dem so genannten „direkten Weg“ (*Direct Path*) drei parallel arbeitende Decoder zur Verfügung, die diese Befehle in jeweils eine oder zwei RISC-ähnliche Instruktionen (*Reduced Instruction Set Computer*), so genannte MacroOps (Macro-Operationen), umsetzen. Eine MacroOp kann jeweils eine ganzzahlige oder Gleitkomma-Operation mit einer Schreib/Lese-Operation verschlüsseln. Komplexe Befehle werden von einem Mikroprogramm-Steuerwerk auf dem *Vector Path* verarbeitet, das sie in Folgen von wenigstens drei MacroOps, zum Teil aber sehr viel mehr, umwandelt. Der letzte Decoder ist auf die Entschlüsselung von Stack-Befehle (*Sideband Stack Optimizer*) spezialisiert. Die beschriebenen fünf Decoder können pro Takt maximal drei MacroOps zu den Verarbeitungseinheiten im RISC-Kern weitergeleitet werden.

Die zentrale Instanz im RISC-Kern ist die *Befehlssteuereinheit* (*Instruction Control Unit*), die mit Hilfe eines Anordnungspuffers (*Reorder Buffer*) die Ausführung der in MacroOps codierten Befehle steuert und überwacht. Dieser Puffer kann in 24 Registern jeweils bis zu drei MacroOps aufnehmen, so dass er insgesamt maximal 72 Einträge enthält. Die Teilaufgaben der Befehlssteuereinheit sind: Zuteilung der MacroOps an die Ausführungseinheiten, Interrupt- und Ausnahmeverarbeitung, Auflösung von Registerabhängigkeiten und Setzen der Statusflags. Außerdem bestimmt sie die Ausführungsreihenfolge der Befehle, die in der Regel von der vorgegebenen Programmordnung abweicht (*out-of-order Execution*) sowie deren Beendigung in der vorgesehenen Programmordnung (*in-order Completion*). Für die Verwaltung der ausgeführten Befehle steht der Befehlssteuereinheit ein Satz von 40 Umbenennungsregistern (*Rename Registers*) zur Verfügung, welche zur Vermeidung von Registerabhängigkeiten zunächst die Ergebnisse der noch nicht beendeten Befehle bis zu deren Abschluss (*Completion*) aufnehmen.²¹ Aus dem Anordnungspuffer können simultan bis zu sechs MacroOps an die folgenden Ausführungseinheiten weitergeleitet werden.

²¹ Mit den x86-Architekturregistern und weiteren technischen Maßnahmen ergibt sich insgesamt ein Satz von 96 Arbeitsregistern.

Die *Ausführungseinheiten* (*Execution Units*) sind in zwei Untereinheiten angeordnet: der Integer-Einheit und der Gleitkomma-Einheit. Beide Untereinheiten sind nach dem Fließbandprinzip (*Pipelining*) realisiert, bei dem die Befehle parallel, aber in mehreren überlappenden Phasen zeitlich versetzt bearbeitet werden. Die Integer-Einheit besteht aus drei unabhängigen „Fließbändern“ (*Pipelines*, kurz: *Pipes*), die Gleitkomma-Einheit nur aus einer.

Die *Integer-Einheit* (*Integer Unit*) übernimmt die Verarbeitung von Befehlen mit ganzzahligen Operanden und Ergebnissen, aber auch die Adressberechnung für die Lade-/Speicherbefehle (*Load/Store*). Ihr können von der Befehlssteuereinheit in jedem Takt bis zu drei MacroOps zugewiesen bekommen. Diese werden zunächst von einer Verwaltungseinheit, dem so genannten *Scheduler*, in so genannte MicroOps umgewandelt. Diese MicroOps sind sehr einfache Instruktionen mit einer festen Länge, die eine Ganzzahlen-Operation und die Adressberechnung für eine zusätzliche Lade-/Speicher-Operation umfassen können. Sie werden in einer Reservierungsstation (*Reservation Station*) mit acht Einträgen abgelegt. Dort warten sie, bis die vorgesehene Ausführungseinheit frei ist und alle benötigten Operanden zur Verfügung stehen. Jede Ausführungseinheit besteht aus einer Arithmetisch/Logischen Einheit (*Arithmetic and Logical Unit* – ALU) und einem Adressrechenwerk (*Address Generation Unit* – AGU), die die Ganzzahlen- und Adressberechnung der MicroOps parallel ausführen können. Dabei ist es auch möglich, dass die ALU und die AGU die Teiloperationen aus zwei verschiedenen MacroOps parallel ausführen. Die Integer-Ausführungseinheiten sind unterschiedlich aufgebaut: Die ALU einer Ausführungseinheit besitzt zusätzlich einen Parallel-Multiplizierer (MUL) und muss daher alle Multiplikationsbefehle ausführen, eine zweite besitzt ein spezielles Rechenwerk zur Ausführung von Bitmanipulationen (ABM).

Die *Gleitkomma-Einheit* (*Floating Point Unit* – FPU) verarbeitet sämtliche Gleitkomma-Befehle sowie alle Multimedia-Befehle. Zu den letztgenannten zählen die Befehlssätze für ganzzahlige Operationen MMX (*Multi-Media Extension* von Intel) bzw. 3DNow! (von AMD) sowie die Befehle für Gleitkommazahlen, die so genannten SSE-Instruktionen (*Streaming SIMD Extension*), die bereits in verschiedenen Ergänzungen (Versionen von 1 bis 4a) vorliegen. Die Gleitkomma-Einheit verfügt insgesamt über 120 Register, einschließlich vieler Umbenennungsregister, sowie eine eigene Verwaltungseinheit. Dieser *Scheduler* enthält eine Reservierungsstation mit 36 Einträgen. Jeweils drei Einträge sind dabei so verbunden, dass sie parallel den angeschlossenen Ausführungseinheiten zugewiesen werden können. Diese Ausführungseinheiten bearbeiten verschiedene Operationsklassen: Die Erste ist für die Gleitkomma-Addition (FADD) und die oben erwähnten SSE-Befehle zuständig, die Zweite führt – neben den SSE-Operationen – Gleitkomma-Multiplikationen (FMUL) aus und die Dritte erledigt die restlichen anstehenden Gleitkomma-Operationen (FMISC – *Miscellaneous*). Insgesamt können also zwei SSE-Operationen simultan ausgeführt werden und mit jedem

Prozessortakt können Fließkamma- und Integer-Einheit zusammen maximal neun MicroOps gleichzeitig beenden.

Die letzte Komponente des K10-Kerns ist die *Lade/Speicher-Einheit* (*Load/Store Unit*), die die Zugriffe auf den L1-Daten-Cache verwaltet. Dieser ist als Zweiport-Speicher mit 128 bit breiten Datenpfaden ausgelegt und ermöglicht so den gleichzeitigen Zugriff auf zwei getrennte Speicherbereiche. Mit jedem Takt kann die Lade/Speicher-Einheit entweder zwei parallele 128-Bit-Lese- oder zwei 64-Bit-Schreiboperationen, aber auch eine Kombination aus beiden Zugriffen ausführen. Leseoperationen auf den L1-Daten-Cache können dabei zum Teil außerhalb der Programmreihenfolge (*out-of-Order*) durchgeführt werden. Natürlich kann auch die Befehlssteuereinheit selbst Lade/Speicher-Operationen umordnen und so eine weiteren Erhöhung des Befehlsdurchsatzes erreichen.

1.4.1.3 K10-Mehrkernprozessoren

Die modernen PC- und Server-Prozessoren der Firma AMD enthalten bis zu sechs K10-Kerne auf einem einzigen Chip. Wie in Abbildung 1.9 gezeigt, haben die Kerne Zugriff auf einen gemeinsamen, bis zu 6 MB großen Level-3-Cache (L3-Cache). Dieser Cache ist – wie bereits oben erwähnt – als *Victim Cache* („Opfer-Cache“) organisiert, der – bis auf hier nicht darzustellende Ausnahmen – nur die Einträge aufnimmt, die aus einem der L2-Caches der Kerne verdrängt wurden. Diese Einträge, die 64 Byte lang sind, werden automatisch im L3-Cache gelöscht, wenn sie von einem der Kerne wieder angefordert werden. Die Zugriffe der Kerne auf den L3-Cache wird von einer speziellen Komponente verwaltet, die als *System Request Interface* (SRI) bezeichnet wird. Wollen mehrere Prozessorkerne gleichzeitig den L3-Cache ansprechen, so wird von ihr der Zugriff nach rotierenden Prioritäten (*Round Robin*) vergeben, d.h. der zuletzt zugreifende Kern muss sich in der Warteschlange „ganz hinten“ wieder anstellen.

Ein weiterer zentraler Bestandteil eines AMD-Prozessors ist der integrierte Speicher-Controller, der über zwei getrennte 64 bit breite Speicherkanäle den Anschluss von DDR2- bzw. DDR3-Speichermodulen erlaubt (vgl. Abbildung 1.6 und Abschnitt 1.5). In einer speziellen Betriebsart können aber beide Kanäle auch zu einem einzigen 128 bit breiten Speicherkanal zusammengefasst und dadurch unter einer beliebigen Adresse auf ein 16-Byte-Datum zugegriffen werden. Durch die Integration des Speicher-Controllers auf den Prozessorchip werden der zeitaufwändige Umweg der Speicherzugriffe über den Prozessorbus und die *North Bridge* sowie Konflikte mit gleichzeitig stattfindenden Zugriffen auf Peripheriekomponenten vermieden. Zugriffe der Kerne auf den (externen) Hauptspeicher werden von der SRI-Komponente zum Speicher-Controller durchgeschaltet.

Für die Kommunikation mit den Peripheriekomponenten stehen dem Prozessor bis zu drei unabhängige HyperTransport-Schnittstellen (HT1 – 3, vgl.

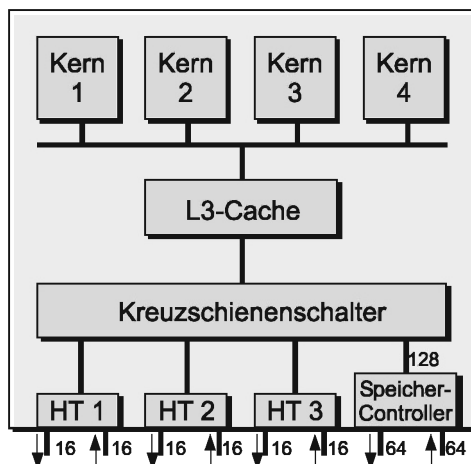


Abb. 1.9 Aufbau eines AMD-Prozessors mit vier K10-Kernen

Abschnitt 1.3.2) zur Verfügung, die über jeweils zwei 16-bit-Pfade den Datenaustausch im Vollduplexverfahren, d.h. gleichzeitige Lese- und Schreibzugriffe ermöglichen. Das SRI erkennt Zugriffe der Kerne auf Peripheriekomponenten und leitet sie an die richtige HT-Schnittstelle weiter. Über diese Schnittstelle(n) können aber auch externe Komponenten auf den Speicher-Controller – und damit auf den Hauptspeicher – zugreifen. Realisiert wird diese Vielfalt von Zugriffsmöglichkeiten durch den Kreuzschienenschalter (*Crossbar Switch*), der simultan mehrere Verbindungen unterhalten kann. So kann z.B. ein Prozessorkern über eine HT-Schnittstelle eine externe Komponente ansprechen und gleichzeitig einer weiteren externen Komponente über eine zweite HT-Schnittstelle der Zugriff auf den Hauptspeicher ermöglicht werden. Zu bemerken ist hier aber, dass realisierte K10-Prozessorfamilien für den PC-Bereich (mit den Bezeichnungen Athlon bzw. Phenom) lediglich eine HT-Schnittstelle besitzen. Drei HT-Schnittstellen sind nur in den so genannten Opteron-Prozessoren für den Server-Bereich zu finden und ermöglichen dort den Einsatz von Hauptplatinen mit mehreren Mehrkern-Prozessoren. Diese Prozessoren enthalten vier oder sechs K10-Kerne.

Die genannten PC-Prozessoren der Phenom-Familie besitzen drei oder vier Kerne. Sie arbeiten mit einer Taktfrequenz von maximal 2,8 GHz und verbrauchen dabei bis zu 140 Watt. Ihr L3-Cache ist auf 2 MB beschränkt. Phenom-Prozessoren werden in 65-nm-Technologie hergestellt, die moderneren Phenom-II-Prozessoren in 45-nm-Technologie. Die Anzahl der Adressleitungen ist auf 48 beschränkt, so dass Speicher mit maximal 256 Terabyte physikalisch adressiert werden können.

Die letzte, in Abbildung 1.9 mit *Power Management* bezeichnete Komponente ist nur der Einfachheit halber als eigenständige Einheit gezeichnet. In Wirklichkeit ist sie über alle Kerne und Komponenten des Prozessors verteilt.

Ihre Aufgabe ist es, denn „Stromverbrauch“ und damit die Leistungsaufnahme des Prozessors in Abhängigkeit von den aktuell ausgeführten Funktionen intelligent zu steuern. Dazu kann sie u.a. die Taktfrequenzen der Prozessorkerne unabhängig voneinander einstellen, in den Kernen momentan nicht benutzte Funktionsblöcke (z.B. bestimmte Rechenwerke) abschalten oder die Versorgungsspannung einzelner Kerne und des Speicher-Controllers zeitweise absenken.

1.4.2 Prozessoren der Firma Intel

Der große Erfolg der AMD64-Architekturserweiterung der Firma AMD für die ursprünglichen x86-kompatiblen 32-bit-Prozessoren zwang die Firma Intel, mit etwas Verspätung ihre Prozessoren ebenfalls auf 64 bit zu erweitern, d.h. insbesondere die Register und die Verarbeitungsmöglichkeit der Integer-Rechenwerke auf 64 bit zu „verbreitern“. Im Wesentlichen übernahm man dazu die Lösungsansätze der Firma AMD. Die verschiedenen Betriebsarten im ursprünglichen 32-bit-Modus werden wir in Kapitel 2 ausführlich beschreiben. Im 64-bit-Betrieb wird bei der Firma Intel die durch die Erweiterung ermöglichte Unterscheidung der Betriebsart als *Intel IA-32e Mode* (kurz für: *Intel Instruction Architecture 32 Extension Mode*) bezeichnet. Sie umfasst

- den Kompatibilitätsmodus (*Compatibility Mode*), in dem ein 64-bit-Betriebssystem die meisten Programme unverändert ausführen kann, die für die 32-bit-Prozessorarchitektur entwickelt wurden,
- den 64-bit-Modus (*64-bit Mode*), in dem ein 64-bit-Betriebssystem Anwendungen ausführen kann, die speziell für die 64-bit-Prozessorarchitektur geschrieben wurden und auf sämtliche 64-bit-Daten- und Adressregister zugreifen können, insbesondere also auch einen 64-bit-Adressraum unterstützen.

Abbildung 1.10 zeigt die so genannten *Architekturregister*, d.h. die Register, die vom Programmierer direkt angesprochen werden können.

Alle betrachteten Prozessoren besitzen den Satz der ursprünglichen universellen Register (*General Purpose Register* – GPR) ihres „Stammvaters“ Intel 8086, die jedoch im Laufe der Entwicklungsgeschichte von ursprünglich 8 bzw. 16 bit Länge auf – wie besprochen – bis zu 64 bit verlängert wurden. Dazu kommen die acht Gleitkommaregister (MMXi/FPRi) der ersten FPU Intel 8087, die von 80 auf 64 bit beschränkt wurden und gleichzeitig für die Verarbeitung der MMX-Daten dienen. Für die Verarbeitung der SSE-Gleitkomma-Operanden wurden 16 128-bit-Register (XMMi) hinzugefügt, die als zweiter Registersatz auch für die MMX-Befehle zur Verfügung stehen.

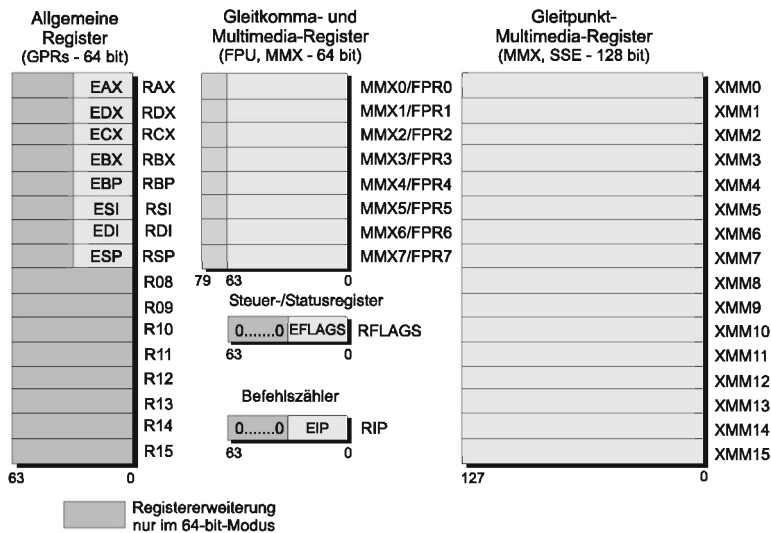


Abb. 1.10 Die Architekturregister der PC-Prozessoren.

1.4.2.1 Der Prozessorkern Intel Core 2

Wie fast alle modernen Prozessoren realisiert die Intel Core 2-Mikroarchitektur eine interne Havard-Architektur (s. Abbildung 1.11):

- Befehle werden im 32 kB großen I-Cache (*Instruction Cache*) abgelegt, der 8-fach assoziativ organisiert ist und durch einen TLB (*Translation Lookaside Buffer*) mit 128 Einträgen zur Unterstützung der Adressumsetzungen ergänzt wird.
- Der D-Cache (*Data Cache*) ist ebenfalls 32 kB groß und 8-fach assoziativ ausgelegt. Er besitzt jedoch zwei unabhängige Zugriffspfade (*Zweiportspeicher*, *dual ported*), sodass simultan auf zwei verschiedene Einträge zugegriffen werden kann.

Aus dem I-Cache werden die Befehle von der Befehlslade-Einheit über einen 128 bit breiten Datenpfad dem *Decoder* zugeführt. Dieser ist zweistufig realisiert: Zunächst werden durch den Vordecoder Beginn und Ende der x86-CISC-Befehle markiert und festgestellt, ob es sich um einfache oder komplexe Befehle handelt. Ausserdem wird ermittelt, ob Verzweigungs- oder Sprungbefehle vorliegen, und in diesem Fall die Sprungvorhersage-Einheit (*Branch Prediction Unit*) aktiviert, die das spekulative Nachladen der Befehle ab einer geeigneten Programmstelle veranlasst. Die vordecodierten Befehle werden in einem 32-Byte-Puffer (*Fetch Buffer*) zwischengelagert. Von dort können dann bis zu sechs Befehle gleichzeitig in die Befehls-Warteschlange übertragen werden, die bis zu 18 Einträge aufnimmt. Zur Beschleunigung der Befehlsverarbeitung können dabei zwei geeignete x86-Befehle zu einem

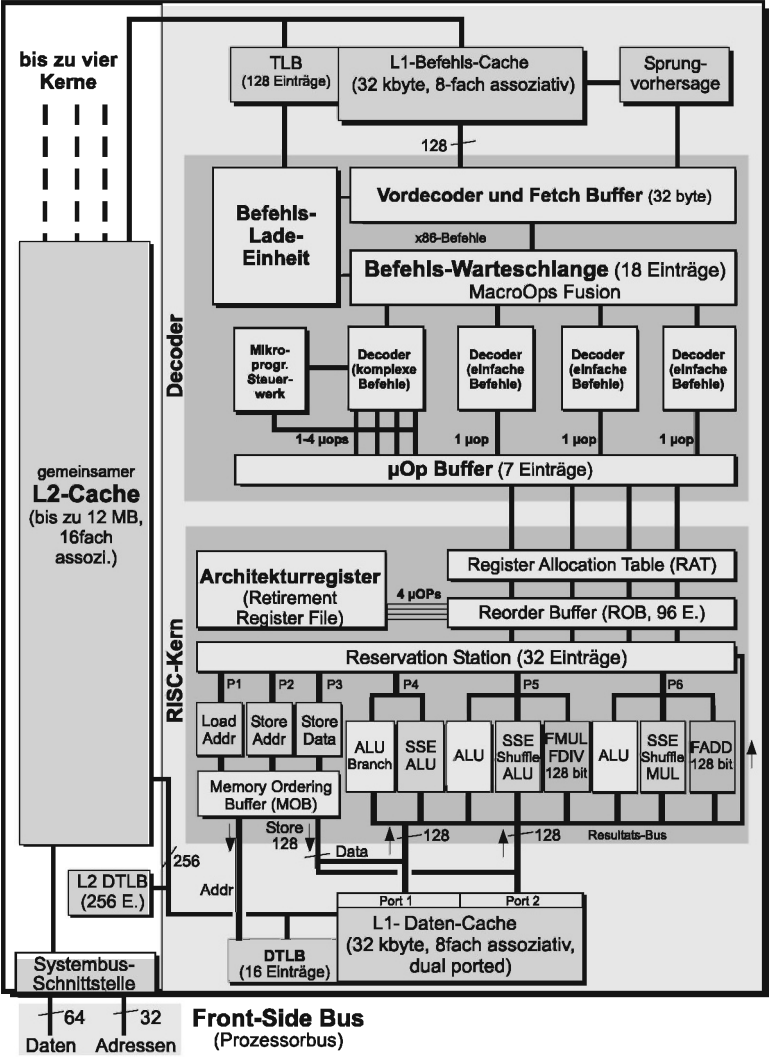


Abb. 1.11 Der Aufbau des Intel Core 2

einzigsten kombiniert und weiterbehandelt werden. Ein wichtiges Beispiel für solch eine Befehlsverkettung ist ein Vergleichsbefehl und ein darauf folgender bedingter Verzweigungsbefehl. Dieses Verfahren wird von Intel als *MacroOps Fusion* bezeichnet. Die eigentliche Entschlüsselung wird durch vier nachfolgende, parallel arbeitende Decoder vorgenommen. Drei von ihnen sind für einfache Befehle ausgelegt, die in jeweils eine so genannte Micro-Operation (μ op) umgesetzt werden. Der vierte Decoder erzeugt für komplexe Befehle entweder direkt 1 – 4 μ ops oder aber – für sehr komplexe Befehle – mit Hil-

fe eines Mikroprogramm-Steuerwerks eine lange Folge von μ ops. Durch die vier unabhängig arbeitenden Decoder ist die Core 2-Architektur also 4-fach superskalar. (Ist unter den codierten x86-Befehlen ein durch MacroOps Fusion zusammengesetzter Befehl, so erhöht sich der Superskalaritätsgrad – für solche Takte – sogar auf 5.) Die durch die Decoder erzeugten μ ops werden parallel in einem Puffer abgelegt, der bis zu sieben Einträge enthalten kann.

Der *RISC-Kern* entnimmt dem μ op-Puffer pro Taktperiode bis zu vier decodierte μ ops. Zunächst weist er ihnen – zur Vermeidung von Zugriffskonflikten auf gleiche Architekturregister – mit Hilfe einer Tabelle (*Register Allocation Table* – RAT) ein freies temporäres Register aus dem Satz so genannter Umbennungsregister (*Rename Register*) als Ausgaberegister zu. Danach trägt er die μ ops in den Anordnungspuffer (*Reorder Buffer* – ROB). Der ROB kann maximal 96 μ ops verwalten und ist für ihre Ausführung in der vorgegebenen Programmreihenfolge verantwortlich. Nach ihrer erfolgreichen Verarbeitung löscht er pro Taktzyklus ebenfalls bis zu vier μ ops und überträgt ihre Ergebnisse mit Hilfe der RAT aus den Umbenennungsregistern in die zugeordneten Architekturregister. Ausführungsbereite μ ops werden im Viererpack in die 32 Einträge umfassende Reservierungsstation (*Reservation Station*) übertragen. Hier warten sie, bis ein geeignetes Operationswerk frei ist und ihre benötigten Operanden zur Verfügung stehen.

Bis zu sechs μ ops können gleichzeitig (über die Pfade P1 – P6) den Operationswerken zugewiesen werden – davon jeweils drei Load/Store-Operationen und drei Rechenoperationen. Die Load/Store-Operationen werden mit Hilfe eines Pufferspeichers, dem *Memory Ordering Buffer* (MOB) in eine möglichst günstige Ausführungsreihenfolge gebracht. Gleichzeitig können bis zu zwei Load/Store-Operationen auf den Zweiport-L1-Datencache zugreifen. Für die Rechenoperationen stehen drei Gruppen von unterschiedlichen Operationswerken zur Verfügung, darunter Integer-Rechenwerke (*Arithmetic Logical Unit* – ALU), Rechenwerke für die Verarbeitung der MMX/SSE-Befehle (*Multimedia Extension*, *Streaming SIMD Extension*) sowie Gleitkommabefehle (FMUL, FDIV, FADD). Eine ALU führt auch die Berechnung der Zieladresse von Verzweigungsbefehlen (*Branch*) aus, zwei weitere erzeugen für den *Shuffle*-Befehl beliebige Permutationen aus den vorgegebenen Datenwörtern. Die von den Rechenwerken verarbeiteten Daten sind bis zu 128 bit breit. Dem entspricht die Breite des Resultats-Busses, der die Ergebnisse in die entsprechenden Register oder den L1-Daten-Cache transportiert.

1.4.2.2 Core 2–Mehrkernprozessoren

Die Core 2-Architektur erlaubt es, bis zu vier Prozessorkerne auf einem einzigen Chip unterzubringen und miteinander zu verbinden. Abbildung 1.11 zeigt auf der linken Seite, dass diese Verbindung über einen gemeinsamen L2-Cache geschieht, der bis zu 12 MB groß und 16-fach assoziativ organisiert ist. Für den Datenzugriff auf diesen Cache steht eine gesonderte Adress-

Übersetzungstabelle L2-DTLB (*Data Translation Lookaside Buffer*) mit 256 Einträgen zur Verfügung. Die bis zu vier Kerne müssen über eine gemeinsame Systembus-Schnittstelle sowohl auf den Hauptspeicher als auch auf sämtliche Peripheriekomponenten zugreifen. Daher stellt diese Schnittstelle sicher einen Engpass in der Core-2-Architektur dar.

Es werden heute Core-2-Prozessoren mit einem, zwei oder vier Kernen angeboten, die den Namenszusatz Solo, Duo oder Quad führen. Die Taktfrequenz beträgt zwischen 1 und 3,33 GHz. Die Prozessoren werden in 65- oder 45-nm-Technologie gefertigt.

1.4.2.3 Der Prozessorbus des Core 2

Beim (semi-)synchrone Busprotokoll stimmt die Buszykluszeit mit der Zykluszeit des Prozessortakts überein. Bei realen Prozessoren der älteren Generationen entsprach ein Buszyklus häufig jedoch mehreren Prozessortaktzyklen. Üblich waren bis zu vier Perioden des Prozessortakts pro Buszyklus. Mit dem Begriff Systemtakt wurde typischerweise der Prozessortakt bezeichnet, von dem der Bustakt (durch Frequenzteilung) abgeleitet wurde. In späteren Jahren gab es dann Prozessoren, bei denen Prozessor- und Bustakt übereinstimmten. Während bei den universellen Hochleistungsprozessoren die Verarbeitungsgeschwindigkeit, vorgegeben durch den Prozessortakt, sehr schnell gesteigert wurde, konnte die Zugriffszeit der Speicherbausteine aus technologischen Gründen nicht entsprechend verringert werden. Ein erster Schritt aus diesem Dilemma bestand darin, die Frequenz des Bustakts nur mäßig zu steigern, den Prozessor intern aber mit einer viel höheren Taktfrequenz zu betreiben, die aus dem Bustakt durch Vervielfachung gewonnen wird. Als Folge davon wird heute typischerweise der Bustakt mit dem Begriff „Systemtakt“ bezeichnet, nicht mehr der (interne) Prozessortakt. Die interne Taktvervielfachung hilft natürlich wenig, den oben beschriebenen Engpaß der langsamen Datenübertragung zwischen dem Prozessor und dem Speicher zu vermeiden. In einem zweiten Schritt wurden daher die folgenden Maßnahmen ergriffen:

- Der externe Zugriff auf einen Speicherbaustein und seine Speicherzellen wurde durch verschiedene Methoden, die hauptsächlich auf den Einsatz von schnellen Registern im Speicherbaustein beruhen, entkoppelt (s. Abschnitt 1.5). Auf diese kann der Prozessor schneller zugreifen als auf den eigentlichen Speicherinhalt.
- Aus dem Bustakt werden durch Vervielfachung (um einen geringen Faktor) schnellere Strobe-Signale gewonnen, die zur Synchronisation der Datenübertragung zwischen dem Prozessor und den Registern der Speicherbausteine dienen. Technisch gesehen, werden die genannten Strobe-Signale von den Flanken des Bustaktes abgeleitet. Bei der Nutzung beider Bustaktflanken gewinnt man (theoretisch) eine Verdopplung der Übertragungsrates. Daher spricht man hier von dem Zweiflanken-Übertragungsverfahren

(*Double Transition* – DT, *Double Data Rate* – DDR) oder einer 2x-Übertragung. Weitere Verdopplungen führen zur 4x- oder 8x-Übertragungsrate.

In Abbildung 1.12 ist der Systembus des Intel Core 2 als Beispiel für einen 200-MHz-Bustakt dargestellt, der als *Front-Side Bus* (FSB) bezeichnet wird. Der Systembus der neueren Core 2-Prozessoren kann mit einer Frequenz von bis zu 400 MHz, also 1600 MT/s, betrieben werden²². (Für diese Taktrate müssen die Zeiten in Abbildung 1.12 noch einmal halbiert werden.) Auf diesem Bus werden die Adressen im 2x-Modus übertragen, gesteuert durch das Signal 2x-Strobe. Seine Flanken liegen jeweils in der Mitte der Impuls- bzw. Pausenzeit des Bustakts.

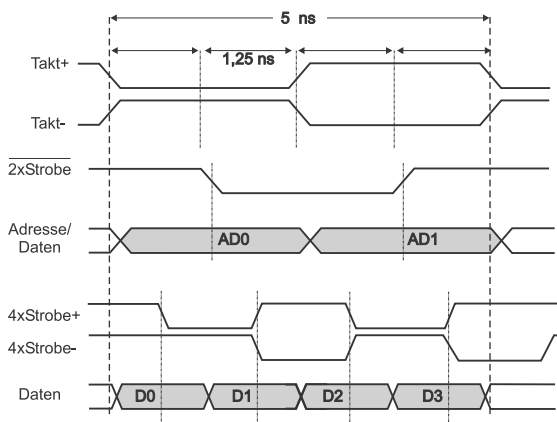


Abb. 1.12 Der Core 2-Bus mit Taktvervielfachung.

Wie bereits gesagt, ist der Core 2-Datenbus 64 bit (8 byte) breit. Dazu kommen noch einmal 8 Fehlererkennungsbits (*Error Correcting Code* – ECC). Die Daten werden im 4x-Modus transferiert, so dass bei einem 200-MHz-Takt bis zu 800 MT/s ausgeführt und damit maximal 6,4 GB/s übertragen werden können. Zur Steuerung werden Strobe-Signale 4xStrobe+, 4xStrobe- eingesetzt, die während eines Buszyklus vier Flanken aufweisen und zu diesen Zeitpunkten das Vorliegen gültiger Daten anzeigen. Auffällig ist die Form dieser Strobe-Signale: Im inaktiven Zustand liegen sie beide auf dem hohen Potential, im aktiven Zustand nehmen sie jeweils entgegengesetzte Pegel ein. Diese Form der Signalerzeugung wird differenzielle Übertragung genannt. Ihre Vorteile liegen hauptsächlich darin, dass sie mit geringeren Spannungspegeln auskommen und unempfindlicher gegen Störungen sind. Denn ein äußeres Störsignal wirkt sich mit großer Wahrscheinlichkeit auf beide Strobe-Signale

²² Die Firma Intel spricht deshalb – vereinfachend, aber marketinggerecht – von einem 1600-MHz-Bus.

gleichartig aus und seine Wirkung kann daher vom Empfänger der Signale durch einfache Differenzbildung herausgefiltert werden. Zum Abschluss sei noch darauf hingewiesen, dass in Abbildung 1.12 auch der Takt zur Unterdrückung von Störsignalen in differentieller Form übertragen wird.

1.4.2.4 Der Prozessorkern Intel Core i7

Die modernste Mikroarchitektur der Firma Intel für PC-Prozessoren trägt die Bezeichnung Core i7, wird aber auch (nach ihrem Codenamen) Nehalem-Architektur genannt. In Abbildung 1.13 sind die Hauptmerkmale dieser Architektur dargestellt. Da sie in vielen Details mit der Core 2-Architektur übereinstimmt, beschränken wir uns im Weiteren darauf, auf die Unterschiede und Ergänzungen hinzuweisen.

- Der größte Unterschied besteht wohl darin, dass jeder i7-Prozessorkern einen eigenen („privaten“) L2-Cache mit einer Kapazität von 256 kB besitzt, der 8-fach assoziativ organisiert ist. Für die Adress-Umsetzung von Datenzugriffen besitzt dieser Cache eine Übersetzungstabelle (*Translation Lookaside Buffer* – L2-TLB) mit 512 Einträgen.
- Der *Prefetch Buffer* ist gegenüber dem Core 2 auf 16 Byte halbiert, der *μOp Buffer* hingegen auf 28 Einträge vervierfacht worden.
- Zur Verringerung der Anzahl der von den Operationswerken auszuführenden *μops* werden in der mit *μops Fusion* indexierten *μops Fusion* bezeichneten Einheit mehrere *μops* zu einer einzigen zusammengefasst²³.
- Der Satz der Architekturregister und die zur Verwaltung benutzte Tabelle (*Register Allocation Table* – RAT) wurden zweifach realisiert. Auf den Grund dafür, gehen wir zum Schluss dieses Unterabschnitts unter der Überschrift *Hyper-Threading* ein.
- Die Anzahl der Einträge im *Reorder Buffer* wurde um ein Drittel auf 128 erhöht, in der Reservierungsstation auf 128 vervierfacht.
- Die Operationswerke zur Berechnung der Adressen in Load/Store-Befehlen wurden zu allgemeinen Adressrechenwerken erweitert (*Address Generation Unit* – AGU).
- Die Systembus-Schnittstelle des Core 2 wurde beim i7 durch eine Schnittstelle zu einem L3-Cache ersetzt, auf den wird im folgenden Unterabschnitt eingegangen werden.

²³ Das Verfahren der *μops Fusion* wurde bereits beim mobilen Pentium M eingeführt. Ob es auch beim Core 2 realisiert wurde, ist leider den ausgewerteten Unterlagen nicht eindeutig zu entnehmen.

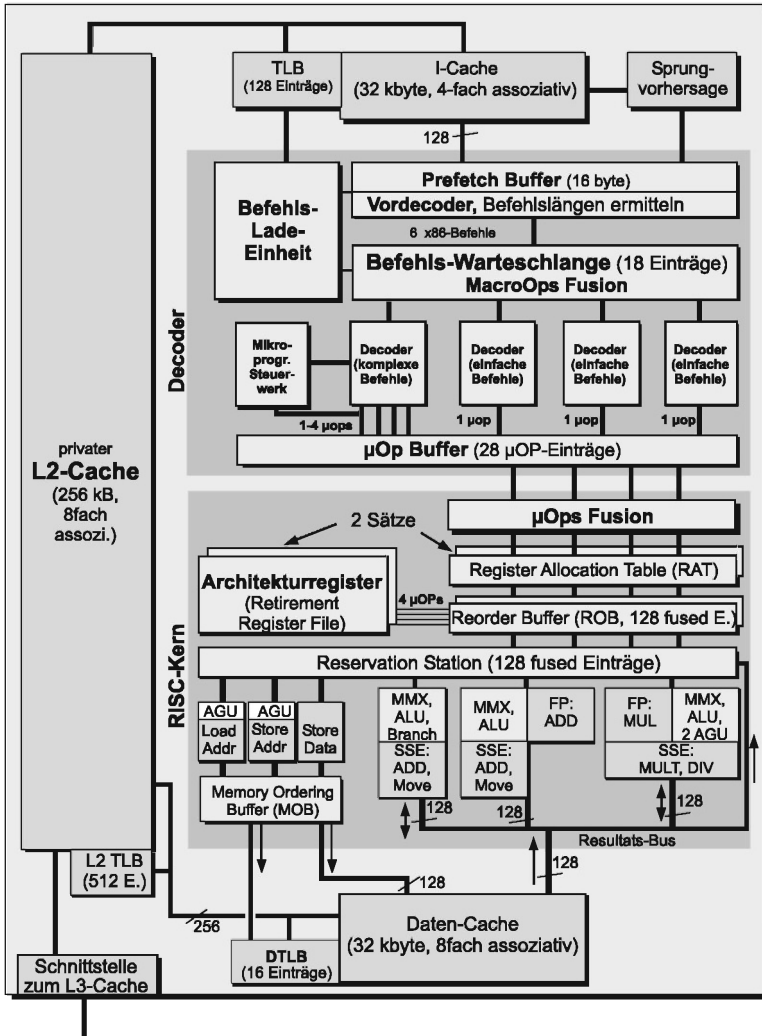


Abb. 1.13 Die Architektur des Intel Core i7

Hyper-Threading

Bereits seit Ende 2002 konnten die Pentium-4-Prozessoren (P4) mit Taktraten über 3 GHz im *Hyper-Threading Mode* arbeiten. Das Verfahren, das dahinter steht, wird in der Rechnerarchitektur üblicherweise als *Simultaneous Multi-Threading* (SMT) bezeichnet und Prozessoren, die diese Fähigkeiten besitzen, werden *simultan mehrfädige Prozessoren* genannt. Mehrfädige Prozessoren sind in der Lage, verschiedene Abschnitte und Befehlsfolgen eines Programms, die unabhängig voneinander sind, bzw. mehrere Programme si-

multan auszuführen. Diese Befehlsfolgen werden *Kontrollfäden* (*Threads*) genannt. So können mehrere logische (virtuelle) Prozessoren auf einem einzigen physischen Prozessor(-kern) arbeiten. Für das Betriebssystem und die Anwendungssoftware liegt damit ein (virtuelles) Mehrprozessorsystem vor – was insbesondere von den modernen *Multitasking*-Betriebssystemen²⁴ für die „gleichzeitige“ Ausführung verschiedener Prozesse genutzt werden kann. Simultan bearbeitete *Threads* können dabei sowohl zu einem Prozess, als auch zu verschiedenen Prozessen gehören. Unter einem *Prozess* verstehen wir dabei ein in Ausführung befindliches oder ausführbares Programm mit seinen Daten (vgl. Abschnitt 2.5).

Beim Intel Core i7 sind es maximal zwei *Threads*, die simultan bearbeitet werden können. Für den zweiten *Thread* stellt der Prozessorkern u.a. jeweils eine weitere Version der folgenden Komponenten zur Verfügung: des (oben erwähnten) Architekturregistersatzes, der in Kapitel 2 beschriebenen Register zur Virtuellen Speicherverwaltung, des Umordnungspuffers (*Reorder Buffer* – ROB) mit Register-Zuordnungstabelle (RAT) sowie des Interruptcontrollers (APIC) und einiger weiterer Steuermodule. Die Gesamtheit der mehrfach ausgeführten Register speichert den so genannten *Architektur-Status* (*Architectural State*) der beiden logischen Prozessoren. Insbesondere muß für jede μ op festgehalten werden, zu welchem der beiden aktiven *Threads* sie gehört. Alle anderen Prozessorkomponenten sind unverändert und werden von beiden *Threads* gemeinsam benutzt. Die Erweiterung des P4 um die Mehrfädigkeit verlangte z.B. lediglich die Erhöhung der Transistorzahl um ca. 1 %, der Chipfläche um ca. 5 %. Die erreichte Leistungssteigerung für typische Programme gab die Firma Intel aber mit bis zu 35 % an.

1.4.2.5 i7-Mehrkern-Prozessoren

Intel wird Prozessoren mit bis zu acht i7-Kernen anbieten. 6-Kern- und 8-Kern-Prozessoren werden im Server-Bereich eingesetzt. Im PC-Bereich werden Prozessoren mit bis zu vier Kernen verwendet. Abbildung 1.14 zeigt beispielhaft einen 8-Kern-Prozessor.

Das Bild zeigt, dass die i7-Kerne über die im letzten Unterabschnitt erwähnte Schnittstelle auf einen gemeinsamen L3-Cache mit einer Kapazität von 8 MB zugreifen können. Der Speicher-Controller wurde aus der North Bridge auf den Prozessorchip verlagert. Das beschleunigt einerseits den Zugriff auf den Speicher, da der zeitraubende Umweg über den Prozessorbus vermieden wird, und verhindert außerdem Zugriffskonflikte, wenn gleichzeitig einer der Kerne auf den Speicher, ein anderer auf eine Peripheriekomponente zugreifen möchte. Der Speicher-Controller unterstützt drei unabhängig arbeitende Speicher-Kanäle (vgl. Abbildung 1.6) mit einer Breite von 64 bit.

²⁴ Auf Multitasking-Betriebssysteme gehen wir ausführlicher in Kapitel 8 ein.

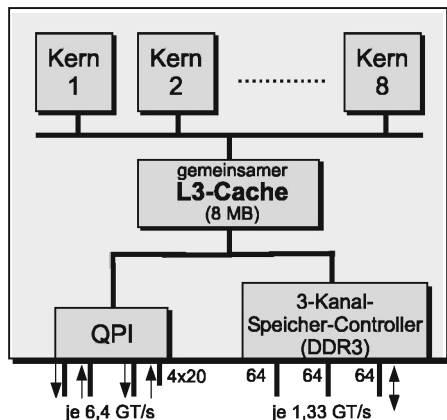


Abb. 1.14 Achtkern-Prozessor mit Intel Core i7

Jeder Kanal erlaubt eine maximale Übertragungsrate von 1,33 GT/s (Giga-Transfers pro Sekunde), also 10,67 GB/s.

Die wichtigste Neuerung der i7-Prozessoren ist sicher die neue Schnittstelle zur Peripherie. Der noch im Core 2 benutzte *Front-Side Bus* (FSB) wurde durch den *QuickPath Interconnect* (QPI) ersetzt, der bereits im Unterabschnitt 1.3.2 beschrieben wurde. Auf vier unidirektionalen Verbindungen, die je 20 Leitungen umfassen, kann er jeweils bis zu 6,4 GT/s ausführen und ist damit erheblich leistungsfähiger als der FSB. Außerdem ermöglicht er eine Vielzahl von Verbindungstopologien zwischen i7-Mehrkern-Prozessoren, auf die im Rahmen dieses Buches jedoch nicht eingegangen werden kann.

Die von Intel angebotenen Core i7-Prozessoren für den PC-Bereich sind in 32- oder 45-nm-Technologie gefertigt und werden mit Taktraten von 2,66 bis 3,33 GHz betrieben. Ihre Verlustleistung liegt bei 130 W bei einer Betriebsspannung von 0,8 bis 1,375 Volt.

1.5 Hauptspeicher

1.5.1 Speichermodule

Der Hauptspeicher eines PCs besteht heutzutage vollständig aus synchronen, dynamischen RAM-Bausteinen (SDRAMs). Dynamische RAMs (*Random Access Memory*) besitzen Speicherzellen, die die Information als Ladungsträger in kleinen Kapazitäten (Kondensatoren) speichern. Durch unvermeidliche Leckströme werden diese Kondensatoren im Laufe der Zeit entladen, sodass der Speicher-Controller die gesamte gespeicherte Information

in regelmäßigen Abständen von wenigen Millisekunden (max. 64 ms) wieder auffrischen (*Refresh*) muss. Auch der Lesezugriff auf eine Speicherzelle zerstört deren Inhalt²⁵, so dass dieser danach erneut eingeschrieben werden muss. Der Zugriff auf den Speicher geschieht mit Hilfe eines Taktes, durch den die auszulesende Speicherzelle in ein Pufferregister geladen bzw. aus dem die eingeschriebene Information in die Speicherzelle gelangt. Das Pufferregister kann vom Controller erheblich schneller angesprochen werden (bis zu 800 MHz) als die Speicherzellen selbst (ca. 20 MHz). Wegen dieser Taktsteuerung spricht man von „synchronen Speichern“.

Seit Anfang 2001 haben sich von den verschiedenen SDRAM-Bausteinen die DDR-SDRAMs, kurz: DDR-RAMs, im großen Umfang durchgesetzt. Ihr großer (namensgebender) Vorteil ist die Eigenschaft, mit jedem Zugriff vorausschauend gleich zwei nebeneinander liegende Speicherwörter in das Pufferregister zu laden (*Prefetch*) und diese dann in einer einzigen Taktperiode durch zwei Datentransfers zu lesen – je einen mit der positiven und der negativen Taktflanke. Beim Schreiben geht man analog vor, d.h. nachdem ein 64 Bit-Wort mit der ersten Taktflanke zwischengespeichert wurde, werden mit der fallenden Taktflanke ein zweites 64-Bit-Wort übertragen und beide Wörter gleichzeitig in den Speicher eingeschrieben. Zusammen mit einer zweikanaligen Ankopplung können pro Taktzyklus vier Datenwörter zwischen Prozessor und Speicher übertragen werden.²⁶

Bei der zweiten Generation der DDR-RAMs, dem DDR2, wurde die Anzahl der vorausschauend geladenen Speicherwörter auf vier erhöht, die dann in zwei Taktperioden mit vier Datentransfers aus dem Pufferregister gelesen werden. Der aktuelle Standard, DDR3, erhöht die Anzahl der simultan gelesenen Speicherwörter weiter auf 8, die in vier Taktperioden übertragen werden. Die beschriebene Entwicklung machte eine Reduzierung der Betriebsspannung von 2,5 V bei DDR-RAMs über 1,8 V bei DDR2-RAMs auf 1,5 V bei den DDR3-RAMs nötig. Momentan werden in PCs hauptsächlich DDR2- und DDR3-Speicherbausteine eingesetzt. Daher beschränken wir unsere weiteren Erläuterungen auf diese Bausteine. Auf ihre genauen Spezifikationen und Eigenschaften gehen wir in einem Unterabschnitt ein.

Durch die beschriebene Integration der DRAM-Speicher-Controller in die *North Bridge* des Chipsatzes bzw. in den Prozessorchip selbst wird einerseits der Aufbau der Speichermodule vereinfacht und kostengünstiger. Andererseits verlangen die großen Stückzahlen und die geforderte Kompatibilität eine gewisse Standardisierung der Module. In Abbildung 1.15 ist eine Platine mit einem DDR2/DDR3-DRAM-Speichermodul skizziert.

Diese Platine ist ca. 133 x 30 mm² groß und wird in einem Steckplatz der Hauptplatine untergebracht. Dazu besitzt sie auf jeder ihrer Platinenseiten an der Unterkante eine Reihe von 120 Steckkontakten, die direkt in die

²⁵ Man spricht deshalb von einem „flüchtigen“ Speicher (*volatile Memory*).

²⁶ Dieses Verfahren wird von Intel als *quad-pumped* bezeichnet.

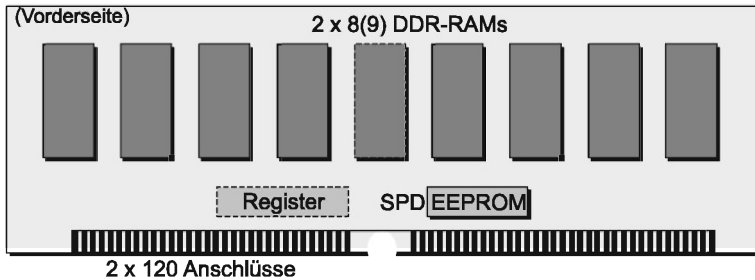


Abb. 1.15 Der Aufbau eines DDR2/DDR3-Speichermoduls

Platinenoberfläche eingeztzt werden („direkte Steckung“).²⁷ Die beidseitigen Steckkontakte geben dem Speichermodule die Bezeichnung DIMM (*Dual Inline Memory Module*). DIMMs werden in einseitig bestückte DIMMs (*single-sided DIMMs*) und zweiseitig bestückte DIMMs (*double-sided DIMMs*) unterschieden, je nachdem, ob sie nur auf einer oder aber auf beiden Platinenseiten DRAM-Bausteine tragen. Leider bedingt der Aufbau eines DIMMs und seine Verbindung mit der Hauptplatine über einen Steckplatz eine Verzögerung der Signale zwischen Speicher und Speicher-Controller in der North Bridge. Daher erreichen DIMMs noch nicht die Übertragungsraten, die man in speziellen Systemen (wie z.B. Spielecomputern oder Graphikkarten) durch direkt auf der Systemplatine eingelötete schnelle Speicherbausteine erhält.

Die DIMMs unterscheiden sich nun wesentlich darin, welche Speicherbausteine auf ihnen implementiert sind. Wie gesagt, kann es sich dabei um DDR2- oder DDR3-RAMs handeln. Auch die Kapazität der Bausteine kann variieren. So können z.B. Speicherbausteine mit einer Datenbreite von $n = 4, 8, 16$ Bit eingesetzt werden. Die Anzahl der Speicherwörter pro Baustein kann bis zu $m = 512 \text{ M}$ ²⁸ betragen. Die Baustein-Organisation wird gewöhnlich zu $m \times n$ angegeben, d.h. der Baustein enthält m Datenwörter zu je n Bits. Für die Kapazität des Bausteins erhält man daraus $m \cdot n$ Bit. Für $n = 16$ und $m = 512 \text{ M}$ erhält man beispielsweise die Organisation $512 \text{ M} \times 16$ und die Kapazität 8 Gbit, also 1 GB. Die Anzahl der Bausteine und ihre Kapazität bestimmen dann die Gesamtorganisation und -kapazität des DIMMs. Ist das in Abbildung 1.15 dargestellte DIMM z.B. mit DDR3-RAM-Bausteinen mit einer Organisation von $256 \text{ M} \times 4$, d.h. einer Kapazität $256 \text{ M} \cdot 4 = 1/8 \text{ GB}$ bestückt, so hat das DIMM – bei einer Bestückung mit 16 RAM-Bausteinen – eine Kapazität von 2 GB. Durch eine Erweiterung um zwei zusätzliche Bausteine erhält man Platz für die Aufnahme von Prüfinformationen. Bei

²⁷ In früheren Jahren wurden auch Steckkarten eingesetzt, bei denen die übereinander liegenden Steckkontakte auf beiden Platinenoberflächen jeweils miteinander verbunden waren, so genannte SIMMs (*Single Inline Memory Modules*).

²⁸ M steht für Mega, also 2^{20} , G im Folgenden für Giga, also 2^{32} .

der Angabe der DIMM-Kapazität wird dieses zusätzliche Speichervolumen jedoch meistens nicht berücksichtigt.

Die DDR-RAM-Bausteine unterstützen außerdem den „Seitenmodus“ (*Page Mode*), bei dem durch einen Lesezugriff auf ein bestimmtes Speicherwort ein ganzer Block von Speicherzellen, eine so genannte Seite, in ein internes Pufferregister übertragen wird. Aus diesem Puffer kann der Speicher-Controller sehr viel schneller lesen als aus dem Speicher selbst. Die effektiven Zugriffszeiten auf weitere Speicherwörter im selben Block werden dadurch wesentlich verkleinert. Dies wirkt sich insbesondere für Übertragungen ganzer Datenblöcke, sog. *Bursts*, leistungssteigernd aus. Typische Seitengrößen von DDR-RAMs liegen bei 1 oder 2 kB. Pro Baustein können dabei mehrere Seiten in entsprechenden Pufferregistern des Bausteins abgelegt werden, auf die dann wahlfrei zugegriffen werden kann. Diese aktuell gespeicherten Seiten werden als „offene Seiten“ bezeichnet. Ein einkanaliger Controller kann bis zu 32 offene Seiten verwalten. Außerdem sind die Speicherzellen von DDR-Bausteinen meist in vier bis acht Speicherbereiche, sog. Bänke, unterteilt. Bei der vorgestellten Organisation des DIMMs überträgt sich diese Bankaufteilung auf das gesamte Speichermodul. Bei der Selektion der Speicherwörter wird das Verfahren der „verschränkte Bankadressierung“ (*Bank Interleaving*) angewandt. Spricht man die verschiedenen Bänke mit geeigneten Adresssignalen an²⁹, so kann erreicht werden, dass konsekutive Zugriffe mit großer Wahrscheinlichkeit auf verschiedene Bänke ausgeführt werden.

Abbildung 1.16 skizziert den möglichen Aufbau eines DIMM-Moduls. Dieses DIMM enthält auf jeder Oberfläche neun DDR3-RAM-Bausteine mit der oben gewählten Organisation von $256 \text{ M} \times 4$. Acht der Bausteine jeder Seite enthalten jeweils 32 Bit der Daten, zusammen also 64-Bit-Daten (8 Byte) – was der Datenbusbreite der typischen PC-Prozessoren entspricht. In den restlichen beiden Bausteinen wird für diese Datenbits eine 8-Bit-Prüfinformation gespeichert, auf jeder Seite des Moduls also 4 Bit. Speichermodule mit dieser ECC-Erweiterung (*Error Correcting Code*) werden hauptsächlich im Server- und Workstation-Bereich eingesetzt, im Heim- und Bürocomputer-Bereich sind sie (heute noch) eher unüblich. Neben den größeren Kosten haben sie noch den Nachteil, dass sie – durch die Prüfsummenberechnung – Schreibzugriffe auf den Speicher verlangsamen. Um die Transferrate zwischen Hauptspeicher und Prozessor zu erhöhen, unterstützen die meisten Chipsätze den parallelen Zugriff auf zwei Module, d.h. es können gleichzeitig 128 Bit übertragen werden.

Wie bereits gesagt, besitzt jede Oberfläche des DDR-DIMMs 120 Steckanschlüsse, das DIMM insgesamt also 240 Anschlüsse. Für die Datenbits werden insgesamt 64 Anschlüsse benötigt. Dazu kommen noch die acht Anschlüsse für die Realisierung der oben erwähnten ECC-Fehlerüberprüfung/Korrektur – unabhängig davon, ob diese implementiert ist oder nicht. Die Adresse ei-

²⁹ Welche Signale das sind, hängt von der inneren Organisation der Speicherbausteine, insbesondere auch von der realisierten Seitengröße, ab und kann hier nicht weiter behandelt werden.

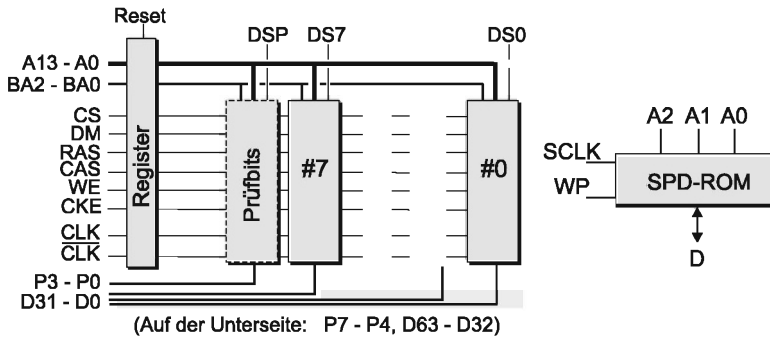


Abb. 1.16 Blockschaltbild einer Oberfläche des Speichermoduls

nes Speicherwortes benötigt insgesamt 28 Bit, was gerade den 256M Wörtern jedes Bausteins entspricht. Diese Adressen werden jedoch in drei Teile aufgeteilt, die z.B. folgendermaßen aussehen können: Drei Bits BA2 – BA0 (*Bank Address*) legen eine der oben beschriebenen Bänke fest. Die folgenden 14 Adressbits, die in den Speicherbausteinen eine Zeile von Speicherzellen („Zeilenadresse“) selektieren, werden zunächst über die Adressleitungen A13 – A0 übertragen. Diese Übertragung wird durch das Steuersignal RAS (*Row Address Strobe*) aktiviert. Erst nach einer gewissen Verzögerung folgen über dieselben Leitungen die restlichen 11 Adresssignale, angezeigt durch das Signal CAS (*Column Address Strobe*). Diese selektieren innerhalb der angesprochenen Speicherzeile ein bestimmtes Speicherwort („Spaltenadresse“). Die zusätzlichen Anschlüsse der DDR-DIMMs werden hauptsächlich für die differentielle Übertragung der Takt- und Steuersignale benötigt.

Das betrachtete DIMM besitzt außerdem noch ein Register. Nach diesem Register wird dieses Speichermodul als *registered DIMM* oder *buffered DIMM* bezeichnet. Dieses Register dient als Treiberbaustein für die Adress- und Steuerleitungen (in Abbildung 1.16: A13 – A0, D63 – D0), die zu allen Speicherbausteinen geführt werden müssen. Es entlastet somit die Ausgangstreiberschaltungen in der *North Bridge*. Nachteilig wirkt sich jedoch aus, dass die Signale beim Durchlauf durch das Register eine Verzögerung von einer Taktperiode erleiden, der Zugriff auf *registered DIMMs* also relativ langsam ist. Es existieren deshalb auch *North Bridges*, deren Speicher-Controller auch mit schnelleren Speichermodulen ohne Register – sog. *unregistered DIMMs* bzw. *unbuffered DIMMs* – arbeiten können. Auf dem in Abbildung 1.15 gezeigten DIMM erkennt man auch das so genannte SPD-ROM (*Serial Presence Detect*), ein kleiner Festwertspeicher zur Aufnahme von Steuerinformationen in einer 128-byte-Speichertabelle. Das SPD-ROM hat eine Kapazität von 128 oder 256 Byte. Nur die Bedeutung der unteren 128 Byte ist fest vorgegeben, die oberen 128 Byte können vom Hersteller frei belegt werden.

1.5.2 Spezifikationen

Wie von allen wichtigen Komponenten des PCs muss auch von den Speichermodulen erwartet werden, dass sie strengen Spezifikationen gehorchen. Nur so ist es möglich, dass Speicherbausteine, die mit ihnen bestückten Speichermodule, Hauptplatinen und Chipsätze – insbesondere ihre integrierten Speicher-Controller – problemlos funktionieren und „zusammenarbeiten“, auch wenn sie von verschiedenen Herstellern stammen. Bei den Speichermodulen werden diese Standards einerseits von der Firma Intel, andererseits von einem Industriegremium (von ca. 120 Firmen) in den USA vorgegeben. Dieses Gremium nennt sich *Joint Electron Devices Engineering Council* und gibt die sog. JEDEC-Spezifikationen heraus.

Die maximale Übertragungsleistung eines DDR-RAM-Bausteins wird typischerweise durch die „nominale Datenrate“ vorgegeben. Diese wird bezogen auf die Frequenz des Speichertaktes $freq$ in MHz, mit dem der Baustein als synchrones dRAM intern angesprochen werden kann³⁰, und die Anzahl π der mit jedem Takt vorausschauend geladenen Speicherwörter, bei DDR2 also $\pi=4$, bei DDR3 $\pi=8$ Wörter. Die Ein-/Ausgabe-Taktfrequenz (E/A-Takt – *I/O Clock*) hingegen, mit der die Übertragung zwischen dem Speichercontroller und den Speicherbausteinen durchgeführt wird, ist bei DDR2-RAMs doppelt, bei DDR3-RAMs viermal so hoch wie die Rate des Speichertaktes. Als Angabe für die nominale Datenrate bekommt man $\pi \cdot freq$, für die wir die Einheit MT/s (Megatransfers/s) benutzen wollen. Die JEDEC-Spezifikation verwendet daher für diese Speicher die Bezeichnung DDR2-($4 \cdot freq$) bzw. DDR3-($8 \cdot freq$). Beispielsweise steht DDR2-800 für einen DDR2-Speicherbaustein, der intern mit 200 MHz getaktet wird. Sein Pufferregister wird mit der doppelten Ein-/Ausgabefrequenz, also 400 MHz, angesprochen. Die maximale Übertragungsrate beträgt 800 MT/s. Denselben Wert erreicht ein DDR3-800-DIMM mit der halben Speicherfrequenz von 100 MHz.

Da vom Hauptspeicher des PCs mit jedem Transfer 8 Byte übertragen werden, erhält man eine Übertragungsrate von $\pi \cdot freq \cdot 8$ MB/s.³¹ In der JEDEC-Spezifikationen für DDR2- und DDR3-Bausteine und damit aufgebaute Speichermodule werden die heute gebräuchliche Speichermodule in der Form PC2-($4 \cdot freq \cdot 8$) bzw. PC3-($8 \cdot freq \cdot 8$) angegeben, wobei der in den Klammern stehende Wert zur Vereinfachung (auf ganzzahlige Vielfache von 100 MHz) gerundet wird. Häufig findet man auch die genaueren Bezeichnungen aus Modul- und Bausteinbeschreibung, also z.B.: PC2-($4 \cdot freq \cdot 8$)/DDR($4 \cdot freq$). In Tabelle 1.1 werden die wichtigsten Daten einiger dieser Module zusammengefasst. Dazu ist jedoch zu bemerken, dass die Module in realen Anwendungen die aufgeführten („theoretischen“) Maximalwerte kaum

³⁰ Davon zu unterscheiden ist jedoch die Zugriffs- bzw. Zykluszeit auf eine einzelne Speicherzelle, die auch heute noch im vielen ns-Bereich (40 ns oder mehr) liegen.

³¹ Sie kann durch den Einsatz eines zweiten unabhängigen Speicherkanals noch verdoppelt werden.

erreichen können. Diese setzen z.B. voraus, dass alle Zugriffe auf offene Speicherseiten (*Page Hit*, s.o.) geschehen. Diese Maximalwerte geben daher eigentlich nur die maximale Übertragungsleistung der Schnittstelle zwischen Speicher-Controller und den Speicherbausteinen an.

Tabelle 1.1 Gebräuchliche Speichermodule.

DIMM- Bezeichnung	Baustein- Bezeichnung	Speichertakt in MHz	E/A-Takt in MHz	nominale Ü.- Rate in MT/s	Ü.-Leistung in GB/s
PC2-3200	DDR2-400	100	200	400	3,2
PC2-4200	DDR2-533	133	267	533	4,3
PC2-5300	DDR2-667	167	333	667	5,3
PC2-6400	DDR2-800	200	400	800	6,4
PC2-8500	DDR2-1066	267	533	1067	8,5
PC3-6400	DDR3-800	100	400	800	6,4
PC3-8500	DDR3-1066	133	533	1067	8,5
PC3-10600	DDR3-1333	167	667	1333	10,7
PC3-12800	DDR3-1600	200	800	1600	12,8

Die JEDEC-Spezifikationen legen alle wichtigen Parameter eines Speichermoduls fest. Dazu gehören insbesondere

- Anzahl, Form und Lage der Anschlusskontakte des Speichermoduls sowie die Größe des Moduls,
- die Kapazität und Organisation der verwendeten Speicherbausteine, also die Anzahl ihrer Datenanschlüsse, die Anzahl der Speicherbänke und deren Aufbau in Speicherwörtern $\times$ Bitlänge,
- die Zeit- und Spannungswerte der Daten-, Adress- und Steuersignale sowie die zulässigen Verzögerungszeiten zwischen den Steuersignalen (*Latency*); das Zeitverhalten allein wird von ca. 30 Parametern bestimmt.

Wie bereits gesagt, wird die Adressierung der Speicherbausteine und die Datenübertragung vom bzw. in den Speicherbaustein mit Hilfe der beiden Steuersignale RAS (*Row Address Strobe*) und CAS (*Column Address Strobe*) durchgeführt. Diese Signale bestimmen die wichtigsten Zeitparameter für den Zugriff auf den Speicher.

RAS-to-CAS Delay: (t_{RCD}) Dies ist die Zeit, die zwischen der Übernahme der Zeilenadresse im Speicherbaustein mindestens vergeht, bevor durch das Anlegen einer Spaltenadresse das gewünschte Speicherwort selektiert werden kann.

CAS Latency: (t_{CL} , kurz: *CL*) Hierdurch wird angegeben, wie lange es dauert, bis das selektierte Speicherwort an den Ausgängen des Bausteins zur Verfügung steht.

RAS Precharge Time: (t_{RP}) Die so genannte Vorladezeit gibt an, welche Zeit der Speicherbaustein für das Vorladen benötigt, bevor eine Speicherseite aktiviert werden kann.

RAS Pulse Width Time: (*Bank Active Time*, t_{RAS}) Das ist die minimale Zeit in Taktzyklen, die nach der Eingabe einer Zeilenadresse mit Hilfe des RAS-Signals vergeben muss, bevor die selektierte Zeile wieder geschlossen werden darf.

Die genannten Zeiten werden bei den heute üblichen synchronen, d.h. getakteten, dynamischen RAMs (SDRAM) in Zyklen des Ein-/Ausgabetaktes $t_{CK} = 1/(2 \cdot \text{freq})$ angegeben. Typische Werte für die CAS-Latenz bei DDR2- und DDR3-RAMs liegen zwischen $t_{CL}=3 \cdot t_{CK}$ und $t_{CL}=11 \cdot t_{CK}$. Um die Schreibweise zu vereinfachen, werden diese Parameter in der Form: CL=2 bis CL=11 dargestellt, d.h. die Einheit t_{CK} wird nicht angegeben.

Bei der Darstellung der beiden anderen Parameter wird lediglich die Einheit t_{CK} weggelassen. Für t_{RCD} und t_{RP} findet man als typische ebenfalls Werte zwischen 3 und 11. Für t_{RAS} sind Werte zwischen 6 und 28 üblich, wobei t_{RAS} meist ein Vielfaches der übrigen Werte ist. Häufig wird der Wert t_{RAS} aber auch nicht angegeben, was wir im Folgenden ebenso halten werden.

Wenn man nun noch berücksichtigt, dass es gepufferte (*registered* – R) und nicht gepufferte (*unbuffered* – U) DIMMs gibt, so ergibt die Zusammenfassung des bisher Gesagten in allgemeiner Form die folgende Bezeichnung für ein DDR-Speichermodul. Die Angabe DDR2(4·freq) bzw. DDR3(8·freq) wird dabei zur Vereinfachung meist weggelassen.

PC2(4·freq)R/U-(t_{CL}, t_{RCD}, t_{RP}) bzw.
PC3(8·freq)R/U-(t_{CL}, t_{RCD}, t_{RP}).

Dazu wollen wir nun lediglich zwei Beispiele angeben:

- PC2-6400R-555 bezeichnet ein mit einem internen Speichertakt von 200 MHz betriebenes, gepuffertes DDR-Speichermodul, das eine CAS-Latenz, eine RAS/CAS-Verzögerungszeit und eine RAS-Vorladezeit von jeweils fünf Ein-/Ausgabetaktperioden, also $t_{CL} = t_{RCD} = t_{RP} = 12,5$ ns, besitzt.
- PC3-8500U-777 kennzeichnet ein mit 133 MHz intern getaktetes, ungepuffertes Modul, das eine CAS-Latenz, eine RAS/CAS-Verzögerungszeit und eine RAS-Vorladezeit von jeweils sieben Ein-/Ausgabetaktperioden, also $t_{CL} = t_{RCD} = t_{RP} = 13,3$ ns, aufweist.

Die wichtigsten Parameter des Speichermoduls werden vom Hersteller im SPD-ROM (*Serial Presence Detect*) einprogrammiert. Das BIOS liest nach dem Einschalten des PCs diese Parameter über den oben beschriebenen SMBus ein und teilt sie dem Speicher-Controller in der Host-Brücke mit. Dieser ist damit in der Lage, das Speichermodul „zeitgerecht“ anzusprechen.

Neben den oben beschriebenen Parametern finden sich im SPD-ROM u.a. Informationen zu den folgenden Moduleigenschaften:

- Typ der eingesetzten Speicherbausteine: SDRAM, DDR-RAM, DDR2-RAM, DDR3-RAM usw.,
- gepuffertes oder nicht gepuffertes Modul (*registered/unbuffered DIMM*),
- Kapazität, Anzahl der Zeilen-/Spalten-Adressleitungen und Speicherbänke sowie der Datenleitungen der Speicherbausteine; je nach Kapazität besitzen heutige RAM-Bausteine 4, 8, 16 oder 32 Datenleitungen;
- Existenz der ECC-Fehlerkorrektur,
- mögliche *Burst*-Längen, d.h. Anzahl der unmittelbar hintereinander ausgeführten Transfers pro Speicherzugriff: 1, 2, 4, 8,
- sowie eine JEDEC-Identifikation für den Modulhersteller und den Herstellungsort, eine Artikel- und Seriennummer sowie das Herstellungsdatum.

Zum Abschluss sei noch erwähnt, dass die in Abbildung 1.4 gezeigte Hauptplatine DX48BT2 der Firma Intel den Einsatz aller in Tabelle 1.1 beschriebenen DDR3-DIMMs erlaubt und so über seine beiden Kanäle eine maximale Übertragungsrate zwischen Hauptspeicher und North Bridge von bis zu 25 GB/s ermöglicht. Da die Übertragungsrate des CPU-Busses jedoch auf maximal 12,8 GB/s beschränkt ist, kann dieser schnelle Speicherzugriff nur zur Hälfte ausgenutzt werden.

Technische Informatik 3

Grundlagen der PC-Technologie

Schiffmann, W.; Bähring, H.; Hönig, U.

2011, XV, 488 S. 188 Abb., 3 Abb. in Farbe., Softcover

ISBN: 978-3-642-16811-6