

Chip-Design zwischen Teilchenphysik und UMTS-Handys

Die fortschreitende Miniaturisierung und die zunehmende Komplexität in der Halbleiter-Elektronik gehen einher mit immer aufwändigeren Entwurfswerkzeugen für Chips.

Ullrich Pfeiffer und Thomas Röwer

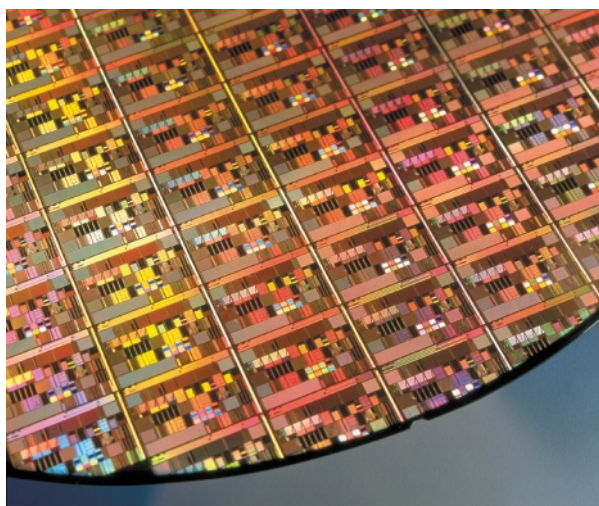
Die Kunst des Chip-Designs besteht in dem intelligenten Vernetzen elektronischer Komponenten auf einem Mikrochip. Ein Fortschritt ist in zweierlei Hinsicht denkbar: Kluge Köpfe entwickeln neue kreative Schaltungen, oder die Chip-Komponenten selbst werden durch neue Prozesstechnologien kleiner und besser. Da die physikalischen Grenzen der Miniaturisierung ohne Zweifel in naher Zukunft erreicht werden, wird weltweit fieberhaft nach neuen Konzepten für mikroelektronische Systeme gesucht.

Die Erfindung des Transistors im Dezember 1947 markiert ohne Zweifel den Grundstein der kontaktlosen Elektronik [1]. Bis zur Erfindung des Chips im Juli 1958 verging allerdings rund ein weiteres Jahrzehnt [2]. Die Idee, Transistoren, Widerstände und Kapazitäten auf einem gemeinsamen, monolithischen Substrat zu verschalten, hat seitdem unseren Alltag revolutioniert. Der Erfolg beruht darauf, dass sich die Chip-Technologie besonders zur Miniaturisierung eignet und dadurch eine ständig steigende Funktionalität bei gleichzeitig reduzierten Kosten ermöglicht. Infolge der rasenden Entwicklung erhöht sich die Komplexität integrierter Schaltungen, gemessen in Transistoren pro Chip, dramatisch. Bereits 1965 beobachtete Intel-Mitbegründer Gordon Moore, dass sich die Transistordichte seit der Erfindung der Chips jedes Jahr verdoppelt hatte [3]. Moore sagte voraus, dass sich dies auch in absehbarer Zukunft nicht ändern würde: Bis heute verdoppelt sich die Kapazität von Speicherchips (DRAMs) etwa alle 18 Monate, was der derzeitigen Definition des Mooreschen Gesetzes entspricht. Früher oder später muss es allerdings eine Abweichung vom exponentiellen Wachstum geben. Daher bieten die Zeitskala und die physikalischen Ursachen Anlass für Spekulationen.

Design integrierter Schaltungen

Chips bestehen aus aktiven und passiven Strukturen, die durch gezieltes Dotieren, Oxidieren und Ätzen einer dünnen Silizium-Scheibe (Wafer) geformt werden. Darüber verlaufen Verbindungsleitungen in von einander isolierten Metalllagen. Beim derzeitigen Stand der Silizium-Prozesstechnologie lassen sich Strukturen mit lateralen Abmessungen von 130 nm in Serie fertigen. Die dazu notwendigen Schritte sind im Infokasten „Prozesstechnologie“ beschrieben [4].

In den Anfängen des Chip-Designs wurden die Mas-



Der Umstieg von Aluminium auf Kupfer als Leiter auf Chips ist ein Beispiel für Fortschritt durch neue Technologien. Daneben kommt aber auch der Platzierung der einzelnen Komponenten auf dem Chip eine entscheidende Bedeutung zu. (Quelle: IBM)

ken verschiedener Prozessschritte allein durch graphische Eingabe erstellt (Abb. 1a). Effizienter wurde das Design schließlich durch wieder verwendbare Elemente wie Transistoren oder Widerstände (Abb. 1b). In den 70er Jahren setzte sich der Metall-Oxid-Halbleiter-Transistor (MOS-Transistor) als dominierendes Element hoch integrierter Schaltungen durch. In dieser Zeit wurde auch der Prozessor, und damit die Basis für zukünftige Software-Produkte, entwickelt. Schaltungen wurden damals hauptsächlich auf Gatterniveau entworfen (Abb. 1c). Innerhalb eines Gatters sind Transistoren bezüglich ihrer Ein- und Ausgänge zu AND-, OR-Funktionen oder Speichern verschaltet. In den 80er Jahren wurde dieses Prinzip um so genannte Standardzellenbibliotheken erweitert. Das Layout einer Standardzelle zeichnet sich durch eine gewisse Regelmäßigkeit, wie zum Beispiel gleiche Höhe oder gleiche Anordnung der Spannungsversorgung, aus. Standardzellen lassen sich dadurch in gleichhohen Reihen anordnen, automatisch platzieren und verdrahten.

Zur automatischen Synthese digitaler Chips werden heutzutage überwiegend „Programmiersprachen“ (Hardware Description Language, HDL) benutzt (Abb. 1d). Das heißt, es muss nur noch die Funktionalität beschrieben werden – die eigentliche Schaltung erzeugt ein „Compiler“. In Verbindung mit Standardzellenbibliotheken vereinfacht dies das Design enorm. Das Maskenlayout der Schaltung lässt sich nun automatisch aus dem Layout einzelner Elemente zusammensetzen. Die Kehrseite dieser Automatisierung ist allerdings eine schlechtere Ausnutzung der Prozesstechnologie.

Dr. Ullrich Pfeiffer,
Dr. Thomas Röwer,
IBM Watson Research Center,
Yorktown Heights, USA.
E-mail: ullrich@us.ibm.com,
roewer@us.ibm.com

Grenzen des Chip-Designs

Doch wo liegen die Grenzen der Miniaturisierung? Die Längenskala, auf der Transistor-Abmessungen verglichen werden, ist die physikalische Länge L_g der Transistor-Steuerelektrode (Gate). Je kürzer das Gate, desto schneller durchlaufen Elektronen den Transistor und desto höher ist seine maximale Schaltfrequenz. Idealerweise sollte sich ein MOS-Transistor wie ein Schalter verhalten können, dessen Zustand sich je nach Spannung am Gate ändert. Im eingeschalteten Zustand fließt ein möglichst hoher Strom zwischen Source- und Drain-Gebiet, während im ausgeschalteten Zustand, das heißt unterhalb der Transistor-Schwellenspannung (V_i), lediglich ein geringer Leckstrom übrig bleibt. Reduziert man die Transistorabmessungen, wird es jedoch schwierig, dieses ideale Schaltverhalten aufrecht zu erhalten, da der Transistor-Leckstrom zunimmt und das effektive Ausschalten des Transistors somit verhindert wird (Kurzkanal-Effekt).

Der bisherige Schlüssel zur Miniaturisierung heißt *Constant-Field Scaling*. Dabei werden alle Abmessungen der Transistoren um den gleichen Prozentsatz reduziert. Das Auftreten hoher Feldstärken versucht man zu vermeiden, indem gleichzeitig die Betriebsspannung reduziert wird. Leider gibt es Größen, die nicht mit dem gleichen Faktor skalieren. Tatsache ist, dass die sog. Temperaturspannung kT/q und die Bandlücke von Silizium E_g unabhängig von der Größe der Struktur sind. Dies führt einerseits dazu, dass die Schwellenspannung nicht in dem Maße skalierbar ist wie andere Parameter, und andererseits zum Nichtskalieren von Potentialbarrieren, der Verarmungszonen-Breite und somit zum Auftreten von Kurzkanal-Effekten. Abb. 2 zeigt die Betriebsspannung (V_{DD}), die Transistor-Schwellenspannung (V_i) sowie die Gateoxid-Dicke (t_{ox}) beim Skalieren der Transistor-Abmessungen.

Alle zwei Jahre sagen internationale Experten-Teams im Rahmen der International Technology Roadmap for Semiconductors (ITRS) die Chip-Entwicklung der nächsten 15 Jahre voraus (Abb. 3, [5]). Vergleicht man die Prognosen aus Abb. 3 mit den neuesten Rekord-

meldungen der Forschungslaboratorien, dann scheinen die Vorhersagen bereits jetzt überholt zu sein. Der gegenwärtige Rekord liegt bei $L_g = 15$ nm und wurde im Dezember 2001 auf dem International Electron Devices Meeting in Washington bekannt gegeben [6]. Eine zuverlässige Serienfertigung in diesem Bereich wird aber, entsprechend der ITRS-Prognose, noch einige Jahre auf sich warten lassen. Bis dahin müssen noch signifikante Entwurfs- und Prozessbarrieren überwunden werden. Grenzen existieren auch in der Lithografie, dem Herstellen dünner Gate-Oxide-Schichten und dem Kontrollieren von Dotierstoff-Fluktuationen.

Ebenso bereitet die Signalverzögerung auf den Leitungen schon lange Schwierigkeiten. Neben fundamentalen physikalischen Effekten spielen aber auch eher praktische Überlegungen eine wichtige Rolle. Schon heute erreichen Chips allein aufgrund ihrer Komplexität die Grenzen der Softwarewerkzeuge, die für den Entwurf verwendet werden. Eine der Schlüsselfragen hinsichtlich der Serienfertigung scheint zu sein, ob man im Bereich unter 100 nm überhaupt in der Lage bleibt, die prozentualen Abweichungen der Transistorparameter in akzeptablen Grenzen zu halten. Zumal man versucht, Kanallängen von einigen nm und Gate-Oxid-Dicken von wenigen Atomlagen zu erreichen.

Entwurfs- und Simulationsmethoden

Der enorme Fortschritt der Prozesstechnologie ermöglicht immer größere Chips mit aufwändigeren Schaltungen. Dadurch werden die Entwurfswerkzeuge wichtiger und gleichzeitig komplizierter, da sie mehr physikalische Gegebenheiten berücksichtigen müssen. Seitdem die Grenze von $0,25 \mu\text{m}$ unterschritten wurde, ist zum Beispiel die Signalverzögerung innerhalb der Gatter geringer als auf den Verbindungsleitungen. Dies hat dazu geführt, dass die Platzierung digitaler Bauelemente schon bei der Synthese der Logik berücksichtigt werden muss. Außerdem ist es immer schwieriger, das „Übersprechen“ zwischen den einzelnen Leitern einzudämmen. Moderne digitale Chips haben zum Beispiel ca. 1,5 km Verdrahtung auf einer Fläche von 8×8

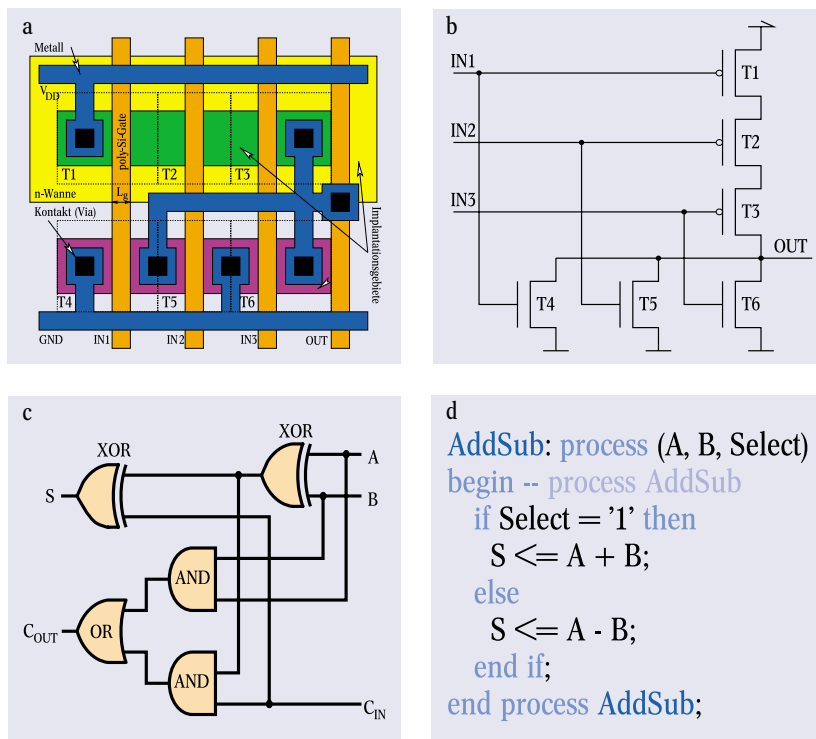


Abb. 1:

Hauptentwicklungsstufen des Schaltungsentwurfs

► **a) Layoutniveau:** Das Chip-Layout besteht aus Polygonen, die mit den Masken der Prozesstechnologie identifiziert werden. Unter Einhaltung der Prozess-Regeln wird hierbei die Geometrie und Verschaltung der Bauelemente definiert. Dargestellt ist ein NAND-Gatter mit drei Eingängen.

► **b) Transistorniveau:** Die Elemente im Schaltplan werden einer Bibliothek mit vordefiniertem Layout entnommen. Nach der Definition des Schaltplans wird im Layout lediglich die Platzierung und Verdrahtung der Elemente vorgenommen. Dargestellt ist der zu Abb. 1a) gehörende Schaltplan eines NAND-Gatters.

► **c) Gatterniveau:** Im Schaltplan (hier ist ein 1-Bit Volladdierer dargestellt) werden nicht mehr einzelne Transistoren verbunden, stattdessen wird auf logische Gatter wie „AND“ oder „XOR“ zurückgegriffen. Das Layout der Gatter ist in einer Bibliothek vordefiniert und bietet standardisierte Schnittstellen zu Nachbarzellen.

► **d) Hardwarebeschreibungssprache (HDL):** Die Funktion der Schaltung wird in Sprachen wie VHDL oder Verilog beschrieben. „Compiler“ erzeugen daraus die Schaltung, die auf Standardzellen mit vordefiniertem Layout zurückgreifen. Das Beispiel zeigt, wie wenig Programm-Code zum addieren und subtrahieren benötigt wird.

mm². Der Fortschritt der Prozesstechnologie ist daher nicht alleine verantwortlich für die rasante Chip-Entwicklung, vielmehr benötigt man Software-Werkzeuge, die es dem Designer erlauben, die neu gewonnene Komplexität zu bewältigen. Diese Werkzeuge sind aufwändige Softwarepakete, die sich je nach Art des Designs (digital, analog oder mixed-signal) unterscheiden. Sie ermöglichen es, das Verhalten der Schaltung im Zeit- oder Frequenzbereich vorauszuberechnen, oder digitale Logik zu synthetisieren, zu platzieren und zu verdrahten. Abbildung 4 zeigt einen typischen „Screen-Shot“ der verwendeten analogen Design-Software [7].

Digitales Chip-Design

Anwendungsspezifische Schaltungen (Application Specific Integrated Curcuits, ASICs) werden heute hauptsächlich durch die automatische Synthese von HDL-Code und anschließendem automatischen Platzieren und Verdrahten der Bibliothekselemente entworfen. Je nach Anwendung wird größeres Gewicht auf Geschwindigkeit, Größe oder Leistungsaufnahme gelegt. Bei modernen Prozessoren wird hingegen nur zum Teil auf eine automatische Synthese vertraut. Hier ist Rechengeschwindigkeit das alles überragende Entwurfsziel. Darum werden die kritischen Teile der Schaltung „von Hand“ erzeugt, um die technologischen Möglichkeiten optimal zu nutzen. Für sehr große ASICs werden ganze Schaltungsteile wie zum Beispiel eingebettete Prozessoren oder Schnittstellen als vordefinierte Blöcke implementiert. Die Verifikation der Ergebnisse bleibt aber schließlich dem Designer überlassen. Je komplexer das Design, umso fehleranfälliger wird die Verifikation der Ergebnisse. Komplexe Prozessor-Designs benötigen etwa spezielle Experten-Teams, die sich ausschließlich mit der Automatisierung und der Verifikation beschäftigen. Wird ein Designfehler erst während der Herstellung des Chips gefunden, ist ein weiterer zeitaufwändiger und teurer Design-Zyklus erforderlich.

Analoges Chip-Design

Der Entwurf analoger Chips unterscheidet sich grundlegend von dem digitaler, da sich die Automatisierung schwierig gestaltet. Erste Schaltungskonzepte werden mit Papier und Bleistift berechnet, bevor man sich mit zeitaufwändigen Simulatoren an die Optimierung der Schaltung heranwagt. Dabei hängt der Designer von der Qualität der Transistormodelle ab, die in Form einer analogen Design-Bibliothek vorliegen. Um das Verhalten möglichst genau vorhersagen zu können, werden Transistorbreiten verwendet, die bis zu fünfmal höher liegen als es die Grenze der verwendeten Siliziumtechnologie vorschreibt. Während die digitale Chip-Entwicklung von modularen und wieder verwertbaren Blöcken profitiert, beinhaltet die analoge Chip-Entwicklung eine Vielzahl verschiedener Techniken und Schaltungen für ein und die selbe Aufgabe. Die geeignete Wahl der Schaltung fällt deshalb nicht leicht und ist oft nur durch Optimieren und Abwägen der Schaltungsparameter möglich. Das Rauschen, die Linearität und der Stromverbrauch der Schaltung lassen sich oft nicht in gleicher Weise verbessern.

Konfigurierbare Chips

Besteht der Wunsch, die Funktionalität eines Chips grundlegend zu ändern, stößt man bei ASICs auf Pro-

bleme, da ihre Funktionalität bereits bei der Synthese festgelegt werden muss. Daher bieten eingebettete Prozessoren, bei denen neue Programm-Routinen auch nach der Produktion noch geladen werden können, eine sinnvolle Alternative. Eine Zwischenstellung dazu nehmen sogenannte Field Programmable Gate Arrays (FPGAs) ein. Dabei sind große Teile der Schaltung bereits vordefiniert. Die kleinsten Bausteine eines FPGAs werden „Cells“ genannt und haben vergleichbare Funktionen, wie die Gatter eines ASICs. Die Verdrahtung ist ebenfalls vorhanden, allerdings müssen hier noch „Schalter“ gesetzt werden, um die benötigten Verbindungen auszuwählen. Diese Schalter können irreversibel oder reversibel sein. Reversible Schalter bestehen aus einem Transistor, der durch ein Speicherelement gesteuert wird. Bei irreversiblen Schaltern werden mit einem Programmierstrom spezielle Leiterbahnsegmente wie eine Sicherung durchgebrannt (*fuse*) oder offene Verbindungen dauerhaft verschmolzen (*anti-fuse*). Je nach Typ des FPGAs stehen zusätzliche Makros, d.h. in ASIC-Form entworfene Blöcke wie Speicher, Prozessoren oder Schnittstellen, zur Verfügung. Moderne FPGAs¹⁾ bieten zum Beispiel die Möglichkeit, vier Power-PC-Prozessoren mit 300 MHz, sechzehn 3,125-GHz-Schnittstellen, 3,8-MBit-RAM und über 50 000 Cells zu implementieren.

Im Vergleich zu ASICs sind FPGAs sehr viel fle-

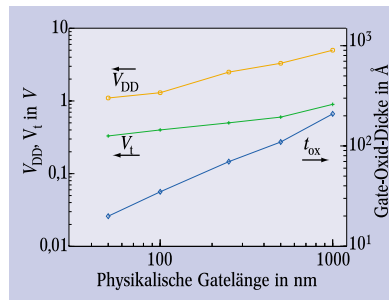


Abb. 2: Betriebsspannung V_{DD} , Schwellenspannung V_I und Gateoxid-Dicke t_{ox} eines MOSFET bei Skalierung der Transistor-Abmessungen. Da die „Temperaturspannung“ kT/q und die Silizium-Bandlücke E_g skaleninvariant sind, lassen sich einige Transistorparameter nicht im gleichen Maße skalieren.

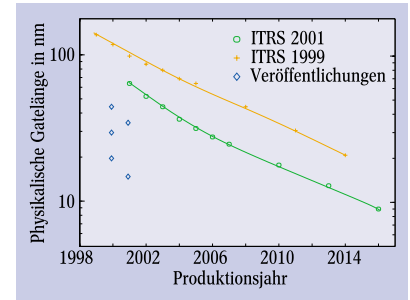


Abb. 3: Vergleich von Prognosen mit aktuellen Publikationen der physikalischen Gatelänge. Bereits jetzt scheinen die Vorhersagen der Experten-Teams überholt zu sein. Bis zu einer zuverlässigen Serienfertigung wird man aber, den Prognosen entsprechend, noch einige Jahre warten müssen.

ditionen auszuwählen. Diese Schalter können irreversibel oder reversibel sein. Reversible Schalter bestehen aus einem Transistor, der durch ein Speicherelement gesteuert wird. Bei irreversiblen Schaltern werden mit einem Programmierstrom spezielle Leiterbahnsegmente wie eine Sicherung durchgebrannt (*fuse*) oder offene Verbindungen dauerhaft verschmolzen (*anti-fuse*). Je nach Typ des FPGAs stehen zusätzliche Makros, d.h. in ASIC-Form entworfene Blöcke wie Speicher, Prozessoren oder Schnittstellen, zur Verfügung. Moderne FPGAs¹⁾ bieten zum Beispiel die Möglichkeit, vier Power-PC-Prozessoren mit 300 MHz, sechzehn 3,125-GHz-Schnittstellen, 3,8-MBit-RAM und über 50 000 Cells zu implementieren.

1) hier speziell der XILINX Virtex-II Pro

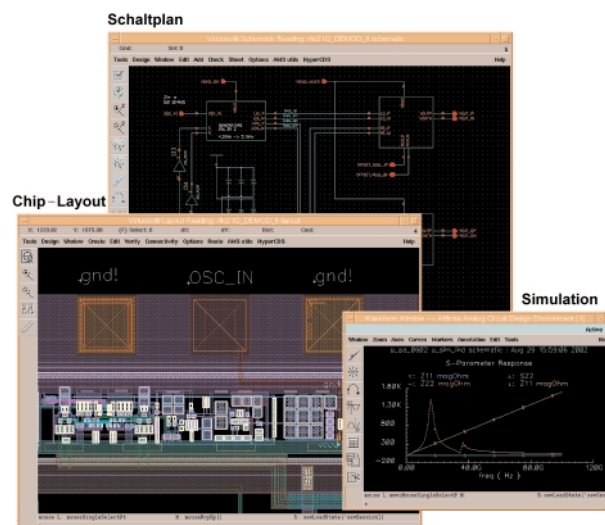


Abb. 4: Typischer „Screen-Shot“ beim Design eines analogen Chips. Im Schaltplan werden Komponenten miteinander logisch vernetzt. Das Layout des Chips besteht aus einer Überlagerung einzelner Polygone, die sich in den „Masken“ der Prozesstechnologie wiederfinden. Simulationsergebnisse werden in Signaldiagrammen im Zeit- oder Frequenzbereich dargestellt.

xibler, da sie häufig umkonfiguriert werden können. Durch die höhere Spezialisierung sind ASICs allerdings schneller, können mehr Funktionalität integrieren und verbrauchen in gleicher Konfiguration weniger Leistung. Ein großer Vorteil von FPGAs ist der geringere Entwicklungsaufwand. Da alle Hardwarestrukturen schon bestehen, muss der Baustein nur noch „programmiert“ werden. Dabei wird die Funktionalität wie bei ASICs mit einer HDL beschrieben, danach synthetisiert, für das jeweilige FPGA optimiert und in den FPGA geladen. Beim Laden wird die „Cell“-Funktionalität festgelegt und die Verdrahtung durch Beschreiben des Speicherelements oder Durchbrennen der Sicherungen konfiguriert. Das aufwändige Platzieren und Verdrahten von Standardzellen entfällt.

Beispiele

Im Folgenden zeigen wir anhand aktueller Projekte aus Forschung und Industrie, wie die diskutierten Methoden des Chip-Designs eingesetzt werden.

Chipentwicklung in der Teilchenphysik

Um den Ursprung und den Aufbau der Materie zu erforschen und insbesondere das Higgs-Boson und vielleicht sogar supersymmetrische Teilchen nachzuweisen, wird am europäischen Teilchenlabor CERN in Genf derzeit der Large Hadron Collider (LHC) gebaut. Der Detektor ATLAS, eins von vier Experimenten am LHC, wird so hoch wie ein fünfstöckiges Haus sein. In

seinem Inneren sollen Protonen mit einer Energie von bis zu 7 TeV aufeinander geschossen werden [8]. Dabei kommt es zu Reaktionen, wie sie Sekundenbruchteile nach dem „Big Bang“ auftraten. Im Zentrum des Detektors finden pro Sekunde rund 40 Millionen Zusammenstöße statt; bei jeder Kollision entstehen durchschnittlich 1600 neue Teilchen, die beim Durchkreuzen der Detektoren eine Datenflut von Terabyte pro Sekunde hinterlassen. Innerhalb einer Stunde würde man damit die Festplatten der größten Internet-Suchmaschine Google zum Überlaufen bringen. Ein äußerst schneller und effizienter Filter (Trigger) für physikalisch relevante Daten ist daher Voraussetzung für das ATLAS-Experiment.

Am Kirchhoff-Institut für Physik der Universität Heidelberg wird die erste Triggerstufe (Level-1 Trigger) des ATLAS-Experiments entwickelt [9]. Die Echtzeit-Elektronik hat dabei die Aufgabe, ca. 7200 analoge Detektorsignale innerhalb von 400 ns zu digitalisieren, zu filtern, zwischen zu speichern und für die Datenauslese bereit zu stellen. Parallelität, Geschwindigkeit und Kompaktheit sind beim Entwurf der benötigten Prozessoren von entscheidender Bedeutung. Abbildung 6a zeigt den digitalen Signalprozessor (Pre-Prozessor ASIC). Der Chip wurde mit Verilog HDL beschrieben und besitzt 24 Speicherblöcke mit einer Größe von insgesamt 66 kbit. Er wurde in ein 0,6- μm -Prozess hergestellt und hat eine Fläche von 60 mm². Abbildung 6b zeigt das Layout eines Chip-Gehäuses (Multi-Chip Modul), das insgesamt sieben Chips mit einer Leistungsaufnahme von mehr als 7 W in sich trägt und in der Lage ist, vier analoge Trigger-Signale auf einer Fläche von nur 14 cm² zu verarbeiten.

Mobilfunk-Systeme der dritten Generation (3G)

3G ist innerhalb Europas auch unter dem Kürzel UMTS (oder Universal Mobile Telecommunications System) bekannt und wird als die Zukunft der mobilen Kommunikation betrachtet. Nach den Spezifikationen des 3rd Generation Partnership Project (3 GPP) [10] sollen Mobilfunkdaten innerhalb definierter Frequenzbänder nach dem *Frequency Division Duplex*-Verfahren (FDD) übertragen werden – von der Basisstation zum Mobilfunkgerät innerhalb 2110–2170 MHz und in der umgekehrten Richtung innerhalb von 1920–1980 MHz. Dabei hat jeder Kanal eine Bandbreite von 5 MHz, der jeweils von verschiedenen Mobilfunkanbietern genutzt wird. Innerhalb eines Kanals teilen sich alle Nutzer eines (*Wideband Code Division Multiple Access*, WCDMA) denselben Frequenzbereich zur selben Zeit. Die dadurch auftretenden Überlagerungen werden reduziert, indem jeder Benutzer eine eindeutige Bitfolge (orthogonalen Code) erhält, mit der sein digitaler Datenstrom multipliziert wird.²⁾

Das eigene Signal wird somit je nach der verwendeten Codelänge verstärkt. Mit anderen Worten, ein WCDMA-Signal kann immer noch empfangen werden, selbst wenn es 10-mal schwächer ist (–20 dB) als benachbarte Signale oder das thermische Rauschen. Der Preis, den man dafür zahlt, ist eine höhere Übertragungsbandbreite, die besondere Anforderungen an das analoge Empfangsteil eines Mobilfunkgerätes stellt. Für die Übertragung der Daten wird schließlich ein digitales Modulations-Verfahren verwendet (*Quadrature Phase-Shift Keying*, QPSK).

Chips, die in mobilen 3G-Geräten eingesetzt werden, müssen aufgrund des hohen Wettbewerbs am

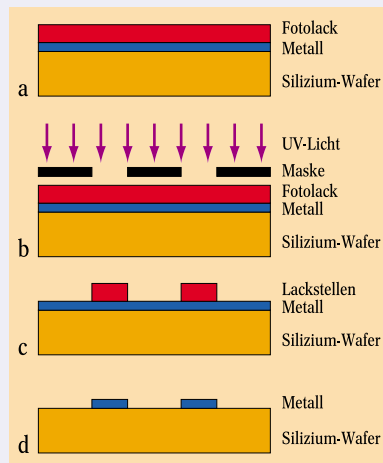
2) Dies ist ein Grundprinzip aus der Signaltheorie und kann leicht verifiziert werden, indem man sich einen zu übertragenden digitalen Datenstrom aus $[-1,1]$ Werten vorstellt, der mit einem sich wiederholenden Code aus $[-1,1]$ Werten höherer Frequenz multipliziert wurde. Der Empfänger kennt den Code, multipliziert den empfangenen Datenstrom ebenfalls, und erhält somit das ursprünglich übertragene Signal zurück. Signale von Nachbarn, die mit einem anderen Code multipliziert wurden, werden schwächer empfangen, wenn man sie über das Zeitintervall eines Datenbits integriert.

Prozesstechnologie

Im Laufe der Chip-Produktion werden die unten beschriebenen Prozessschritte in leicht abgeänderter Form mehrfach wiederholt, um Bauelemente und Verdrahtung zu erstellen.

► a) Zunächst wird auf den Wafer (gelb) das gewünschte Material, zum Beispiel Oxid oder Metall (blau), aufgebracht. Dies geschieht durch Oxidation des Siliziums oder durch Depositionen des Materials im Hochvakuum. Dann wird ein Photolack aufgebracht, indem der Lack auf den rotierenden Wafer getropft wird. Durch die Rotationsgeschwindigkeit lässt sich dabei die Schichtdicke einstellen.

► b) Nun wird der Photolack durch eine Maske belichtet. Dies ist der entscheidende Schritt für die weitere Verkleinerung der Strukturen in der Chipfertigung. In der Vergangenheit wurden Quecksilberdampflampen ($\lambda = 436 \text{ nm}$ und später 365 nm) und Excimer-Laser ($\lambda = 248 \text{ nm}$ (KrF) und 193 nm (ArF)) benutzt. Für zukünftige Technologien wird über noch kürzere Wellenlängen wie zum Beispiel Röntgenstrahlung nachgedacht. Da die Spaltbreite in den Masken und die Wellenlänge die gleiche Größenordnung haben, treten Interferenzen auf. Um eine Verbreiterung des Belichtungsmusters zu verhindern, werden so genannte Phase-Shift-Masken benutzt. Durch eine Phasendrehung um 180 Grad in benachbarten Spalten wird negative Interferenz im Dunkelbereich erreicht und somit eine größere Flankensteilheit in der Beleuchtungsintensität.



► c) Durch die Belichtung werden zusätzliche Vernetzungen im Polymerlack induziert und somit die Löslichkeit im Entwickler reduziert (Negativ-Lack). Nach der Entwicklung werden somit die unbelichteten Bereiche entfernt. Die nun noch vorhandenen Lackstellen dienen als Maskierung für den folgenden Prozessschritt.

► d) Im gezeigten Beispiel wird die vorher aufgebrauchte Schicht durch reaktives Ionen-Ätzen entfernt. Dabei hat der Ätzabtrag eine chemische und eine physikalische Komponente. Das Verfahren ermöglicht maßhaltiges Übertragen der vordefinierten Struktur. Es könnte aber auch zum Beispiel eine Ionenimplantation folgen, um das Silizium lokal zu dotieren.

Markt vor allem günstig in der Herstellung sein. Sie dürfen möglichst wenig elektrische Leistung und Platz verbrauchen. Daher werden für analoge Empfänger von 3G-Geräten so genannte *single-chip*-Entwicklungen angestrebt. Diese Chips sollen aus Kompatibilitätsgründen in der Lage sein, sich an möglichst viele verschiedene Mobilfunkstandards anzupassen. Aufgrund des geringen Platzes muss aber möglichst auf externe Bauelemente wie hochwertige Filter (z.B. *Surface Acoustic Wave Filter*) verzichtet werden. Daher sind vor allem so genannte *direct-conversion*-Empfänger interessant, die Signale direkt, d. h. ohne Zwischenfrequenz und Bandpass-Filter, in das Basisband modulieren [11]. Hohe Anforderungen werden an die Linearität der analogen Schaltungen gestellt, um die Intermodulationsprodukte³⁾ der Nachbarkanäle im Empfangskanal zu verringern. Zum Einsatz kommen dabei bipolare Chip-Technologien. Transistoren aus Silizium-Germanium zum Beispiel haben ein geringes Rauschen und einen niedrigen Stromverbrauch und zeichnen sich besonders bei extrem hohen Frequenzen aus [12].

Abbildung 7 zeigt einen *direct-conversion*-Empfänger, der bei IBM in einem Silizium-Germanium (SiGe) BiCMOS-0,25- μm -Prozess hergestellt wurde [13]. Der entwickelte Chip besteht aus einem rauscharmen Eingangsverstärker, einem Mischer, einem Synthesizer, einem variablen Verstärker sowie einem programmierbaren Basisbandfilter.

Ethernet mit 10 GBit/s

Die Telekommunikation ist eines der Gebiete, die am meisten von der rasanten Entwicklung in der Mikroelektronik profitiert haben. Während die Übertragungsraten immer weiter anwachsen, werden die Geräte immer kleiner bei geringerem Stromverbrauch. Momentan werden die ersten Geräte, die 10 GBit/s nach dem Ethernet-Standard übertragen können, entwickelt [14].

Hier soll der Entwurf des Chips, der die Schnittstelle zwischen Glasfaser und dem TCP/IP-Router darstellt, beschrieben werden. Der Entwicklungs-Schwerpunkt liegt auf hoher Integration und niedrigem Leistungsverbrauch. Analoge und digitale Schaltungsteile werden auf einem Chip zusammengefasst, um Synergien auszunutzen und Preisvorteile realisieren zu können. Dies hat außerdem zur Folge, dass auch die analogen Schaltungsteile in CMOS-Technologie implementiert werden. Eine CMOS-Implementierung bietet erhebliche Leistungs- und Preis-Vorteile gegenüber bipolaren Transistoren, erschwert aber den Schaltungsentwurf erheblich.

Der Chip ist in drei große Einheiten unterteilt. Im ersten Teil werden die seriellen Eingangsdaten in einen 64 bit breiten Datenstrom umgewandelt und dabei um das 64fache verlangsamt. Die gesamte Datenverarbeitung findet im zweiten Teil statt. Hier wird das Ethernet-Protokoll ausgewertet, Prüfsummen werden berechnet und die Daten decodiert. Schließlich wird der Datenstrom wieder auf 4 bit verengt, um Leitungen auf der Platine zu sparen. Jeder dieser drei Teile besteht zudem aus einem Empfangs- und einem Sendeteil. Die Senderseite erhält 4-bit-Daten, erzeugt das Ethernet-Protokoll und sendet diese in serieller Form auf die Glasfaser. Diese drei Teile (10 Gbit/s seriell-parallel Wandler, digitale Protokollauswertung und analoge Platinenschnittstelle) werden von separaten Teams ent-

wickelt und dann zum Gesamtchip zusammengesetzt. Dies ist eine nicht zu unterschätzende Aufgabe, da sichergestellt werden muss, dass der Digitalteil die „analogen“ Signale richtig interpretiert und umgekehrt.

Diese Anwendung zeigt, dass auch heute noch die gesamte Breite der Schaltungsentwurfs-Methoden genutzt wird. Die sehr hohe Frequenz der Ein- und Ausgangsdaten verlangt die sorgsame manuelle Methode, während der langsame aber komplexe Digitalteil durch HDL-Codierung erstellt wird.

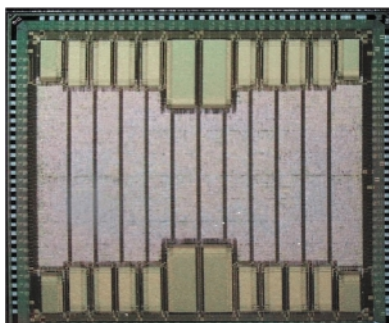
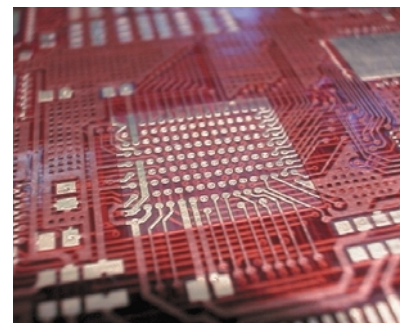


Abb. 6: Der Signalprozessor für die erste Triggerstufe der Kalorimeter am ATLAS-Experiment sorgt für die Vorverarbeitung und die Datenauslese von Triggersignalen in Echtzeit (links). Er wurde vollständig mit der Hardware-Beschreibungssprache Verilog entworfen, simuliert und syntheti-



siert. Chips arbeiten im Level-1-Trigger möglichst effizient in kompakten Gehäusen, sog. Multi-Chip-Modulen (MCMs). Das rechte Foto zeigt das MCM-Substrat, auf dem Chips mit 100 μm breiten Leitungen verbunden werden.

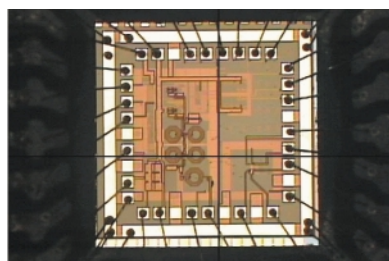


Abb. 7: Empfänger für ein UMTS-Handy. Der rein analoge Chip wurde am IBM Watson Research Center entwickelt und besitzt eine Fläche von $1,7 \times 1,7 \text{ mm}^2$.

Ausblick

Die ultimative Skalierungsgrenze ist die Längenskala der Atome. Nähert man sich diesem Bereich, ist mit wohl verstandenen quantenmechanischen Effekten zu rechnen, welche die Anwendbarkeit der zur Zeit benutzten Gesetze für elektrische Stromkreise einschränken. Die Quantenmechanik wird aber für solche Systeme auch ganz neue Methoden der Informationsverarbeitung eröffnen, bei denen die klassische Boolesche Logik durch Quantenlogik ersetzt wird. Auf diesem Sektor fehlt es zur Zeit noch an theoretischen Grundlagen, jedoch gibt es bereits eine Reihe von Forschungsaktivitäten auf diesem Gebiet.

In der Zwischenzeit versucht man möglichst lange durch alternative Transistorstrukturen und neue Materialien an der Siliziumtechnologie festzuhalten. Derzeit sind MOSFET-Transistoren, die auf *Silicon-on-insulator* (SOI), SiGe sowie verspanntem Silizium (*strained silicon*) basieren, und sog. *double-gate*-Transistoren⁴⁾ am Erfolg versprechendsten. Abbildung 8 zeigt eine Elektronen-Mikroskop-Aufnahme von *double-gate*-Transistoren [15]. Der Transistorkanal wird dabei von drei Seiten eingeschnürt. Der Strom fließt hauptsächlich entlang der beiden Seiten, daher die Bezeichnung *double-gate*. Laut Vorhersagen sollen damit Frequenzen von mehreren Terahertz möglich sein.

3) Nichtlinearitäten der Übertragungsfunktion eines Schaltkreises in der Form $y(t) = \alpha_1 x(t) + \alpha_2 x(t)^2 + \alpha_3 x(t)^3$ führen zu Intermodulation, d. h. Termen mit $\cos(n\omega_1 \pm m\omega_2)$, wenn neben dem ursprünglichem Signal ω_1 ein Nachbarsignal ω_2 vorhanden ist und $x(t) = A_1 \cdot \cos(\omega_1 t) + A_2 \cdot \cos(\omega_2 t)$ gilt.

4) Eine äquivalente Bezeichnung von Intel lautet *tri-gate*.

Ein weiteres Transistor-Konzept, das wahrscheinlich in der 90 nm oder 65 nm Prozess-Generation zum Einsatz kommen wird, beruht auf verspanntem Silizium (Abb. 9). Dabei wird eine dünne Siliziumschicht auf ein SiGe-Substrat aufgewachsen. Da die Gitterkonstante von SiGe größer ist als die von Si, treten Verspannungen in der Si-Deckschicht auf. Diese Verspannungen erhöhen die Elektronen- und Löcherbeweglichkeit und ermöglichen somit bis zu 35 % schnellere Transistoren [16].

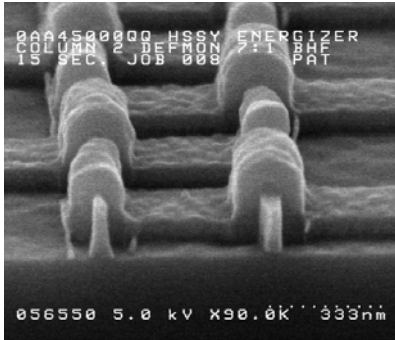


Abb. 8: Elektronen-Mikroskop-Aufnahme eines *double-gate*-Transistors von IBM. Das Gate umklammert einen Silizium-Steg, der Source und Drain darstellt.

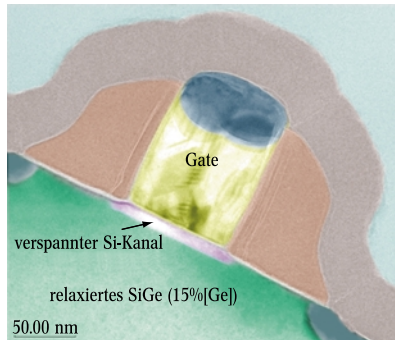


Abb. 9: Elektronen-Mikroskop-Aufnahme eines *strained-silicon*-Transistors von IBM. Die verspannte Silizium-Schicht, im Bild violett dargestellt, bildet den leitenden Transistor-Kanal.

Diese Forschungsergebnisse lassen erwarten, dass das Mooresche Gesetz auch in den nächsten Jahrzehnten weiter Bestand haben wird. Zeit genug bleibt also, um mit ganz neuen informationsverarbeitenden Methoden aufzuwarten.

Literatur

[1] *J. Bardeen, W. Shockley und W. Brattain*, U.S. Patent 2,524,035 (1950) www.eepatents.com, www.uspto.gov.
 [2] *J. S. Kilby*, U.S. Patent 3,138,743 (1964).
 [3] *G. E. Moore*, in *Electronics* 38, April 1965. www.intel.com/research/silicon/moorespaper.pdf.
 [4] *S. M. Sze*. *Semiconductor Devices: Physics and Technology*. John Wiley & Sons, 2. Auflage (2001).
 [5] Semiconductor Industry Association SIA. *International Technology Roadmap for Semiconductors*

2001 Edition. SIA, 2001. <http://public.itrs.net>.
 [6] *B. Yu et al.*, IEEE International Electron Devices Meeting, December 2001;
 [7] Cadence Design Systems; www.cadence.com.
 [8] <http://atlas.web.cern.ch/Atlas>.
 [9] *P. Hanke et al.*, Fifth Workshop on Electronics for LHC Experiments, September 1999. <http://web.kip.uni-heidelberg.de/atlas>.
 [10] The 3rd Generation Partnership Project (3GPP). www.3gpp.org.
 [11] *B. Razavi*, RF Microelectronics. Prentice-Hall, (1998).
 [12] *D. Harame et al.*, IEEE International Caracas Conference on DEVICES, CIRCUITS and SYSTEMS, 2002.
 [13] *S. Reynolds et al.*, IEEE Bipolar/BiCMOS Circuits and Technology Meeting, September 2002.
 [14] IEEE 802.3ae 10Gb/s Task Force. IEEE Draft P802.3ae. <http://grouper.ieee.org/groups/802/3/ae/index.html>.
 [15] *J. Kedzierski et al.*, IEEE International Electron Devices Meeting, December 2001.
 [16] *K. Rim*, IEEE International Solid State Circuits Conference, 2001.

Die Autoren

Seine Begeisterung für physikalische Experimente entwickelte **Ullrich Pfeiffer** schon früh beim Bau von Lautsprecher-Chassis zusammen mit seinem Vater. Er studierte Physik an der Universität Heidelberg und entwickelte während seiner Doktorarbeit Echtzeit-Elektronik für die Vorverarbeitung von Kalorimeter-Signalen in der ersten Trigger-Stufe des ATLAS-Experiments am CERN. Sein derzeitiges Forschungsgebiet sind Chips für die drahtlose Kommunikation unter Einsatz schnellster Silizium-Germanium Chip-Technologien. **Thomas Röwer** studierte Elektrotechnik mit Schwerpunkt Halbleitertechnologie an der RWTH Aachen. Während des Studiums absolvierte er ein Praktikum bei MEMC Electronic Materials in St.Louis, MO, USA, und arbeitete in der Waferproduktion. Seine Doktorarbeit über wiederverwendbare Blöcke im ASIC-Entwurf hat Thomas Röwer an der ETH Zürich geschrieben. Derzeit beschäftigt er sich mit dem Schaltungsentwurf für Kommunikationsanwendungen.

