

Zeitschrift: Bulletin des Schweizerischen Elektrotechnischen Vereins, des Verbandes Schweizerischer Elektrizitätsunternehmen = Bulletin de l'Association suisse des électriciens, de l'Association des entreprises électriques suisses

Herausgeber: Schweizerischer Elektrotechnischer Verein ; Verband Schweizerischer Elektrizitätsunternehmen

Band: 78 (1987)

Heft: 11

Artikel: VLSI-Technologie

Autor: Solo de Zaldivar, J.

DOI: <https://doi.org/10.5169/seals-903867>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 01.04.2025

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

VLSI-Technologie

J. Solo de Zaldivar

Dieser Beitrag gibt eine Übersicht über den Stand und die Zukunftsaussichten der Siliziumgate-MOS-Technologie. Die grundlegenden Techniken, ihre Vor- und Nachteile bei bestimmten Anwendungen sowie wirtschaftliche Aspekte werden erörtert. Abschliessend werden die wichtigsten technologischen Entwicklungen präsentiert, die alle dem gleichen Ziel dienen, der Erhöhung der Packungsdichte.

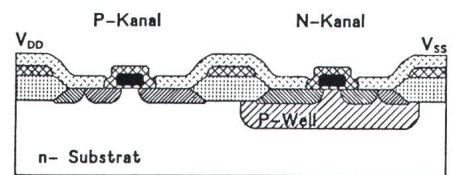
Cet article donne une vue d'ensemble sur l'état actuel et futur de la technologie MOS. Les techniques de base, leurs avantages et leurs inconvénients dans certaines applications ainsi que le côté économique y sont mentionnés. Finalement les principaux développements technologiques sont décrits, ayant tous le même but: augmenter la densité des éléments sur la puce.

1. MOS-Technologien

Die bedeutendsten Varianten der MOS-Technologie sind einerseits die NMOS-Technologie, welche nur n-Kanal-Transistoren verwendet, und die CMOS-Technologie (Complementary MOS), die erlaubt, auf derselben Schaltung p- und n-Kanal-MOS-Transistoren zu integrieren. Die Figur 1 zeigt das Schaltschema und das Layout eines CMOS-Inverters (Grundelement aller logischen CMOS-Schaltungen) in p-Wannentechnik, die Figur 2 den zugehörigen Querschnitt auf dem Si-Chip. Wie anschliessend gezeigt wird, bietet die CMOS-Technologie gegenüber der NMOS-Technologie eine Reihe von Vorteilen, welche sie heute schon zur führenden Technologie für VLSI-Schaltungen¹ macht.

Bis Anfang der achtziger Jahre hat die CMOS-Technologie eine sekundä-

¹ VLSI (Very Large Scale Integration) ist die Bezeichnung für IC-Technologien, welche die Integration von über 100 000 Transistoren auf einem einzelnen Chip ermöglichen.



Figur 2 Inverterquerschnitt

A-A' vgl. Fig. 1

- | | |
|-----------------------------|----------------|
| ■ Polysiliziumgate | ☑ Aluminium |
| ▨ n ⁺ -Diffusion | ▨ LOCOS |
| ▨ p ⁺ -Diffusion | ▨ Siliziumoxid |
| ▨ p ⁻ -Diffusion | |

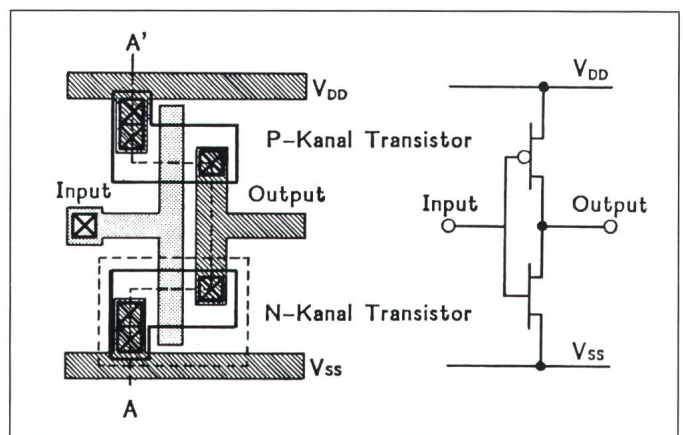
re Rolle gespielt. Sie wurde hauptsächlich für die Herstellung von Uhrenschaltungen und Taschenrechnern, bei welchen der Stromverbrauch auf ein Minimum reduziert werden muss, verwendet. Komplexere Schaltungen dagegen, wie Mikroprozessoren und Speicher, wurden bis dahin in NMOS-Technologie angefertigt.

Die Kraft, welche die IC-Technologie während des letzten Jahrzehnts vorangetrieben hat, ist ohne Zweifel der Zwang, die Komplexität der Schal-

Figur 1 Inverter: Layout und Schaltschema

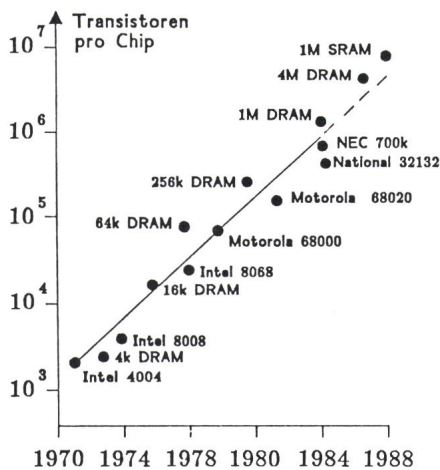
- | |
|---------------------|
| ☒ Kontaktloch |
| □ Diffusionsgebiete |
| ▨ Polysilizium |
| ▨ Metall |

A-A' Schnitt durch Inverter. Man beachte, dass dieser nicht stetig, sondern entlang der gestrichelten Linie verläuft.



Adresse des Autors

J. Solo de Zaldivar, Dipl. El.-Ing., Faselec AG, Räfelstrasse 29, 8045 Zürich.



Figur 3 Entwicklung der Integrationsdichte

tungen stetig zu erhöhen (Fig. 3) [1], was nur durch ständige Vergrößerung der Packungsdichte (Anzahl Transistoren pro Flächeneinheit) d. h. in erster Linie durch Verkleinerung von Transistoren und Leiterbahnen sowie von deren Abständen zu erreichen ist. Dies wiederum führt infolge der höheren Transistor-Steilheit und der Verringerung der Drain-p-n-Übergangskapazität zu einer Erhöhung der Schaltgeschwindigkeit. Da gleichzeitig die parasitären Kapazitäten der Leiterbahnen herabgesetzt werden, ergibt sich insgesamt auch eine Verminderung des *dynamischen* Stromverbrauchs. Trotzdem führt die hohe Packungsdichte der VLSI-Technologie zu einer erhöhten Wärmedichte auf der Chip-Oberfläche. Da CMOS-Schaltungen wegen ihrer Transistorkomplementarität einen vernachlässigbaren statischen Stromverbrauch aufweisen, ist die CMOS-Technologie für VLSI-Anwendungen besonders geeignet. NMOS-Schaltungen dagegen führen hohe statische Ströme, die unter Umständen eine Erhöhung der Betriebstemperatur zur Folge haben können. Ihre Zuverlässigkeit kann dadurch ernsthaft beeinträchtigt werden.

Andere Zuverlässigkeitsaspekte, die für die CMOS-Technologie sprechen, sind bessere Immunität gegen α -Teilchen (SER, Soft-Error-Rate), weniger gravierende Hot-electron-Effekte, vernachlässigbare Elektromigration.

a) Immunität gegen α -Teilchen: Dieser Punkt ist von besonderer Bedeutung für die Zuverlässigkeit grosser Speicher. Im Fall eines CMOS-DRAM-Speichers z. B. können die Speicherzellen in einer p-Wanne

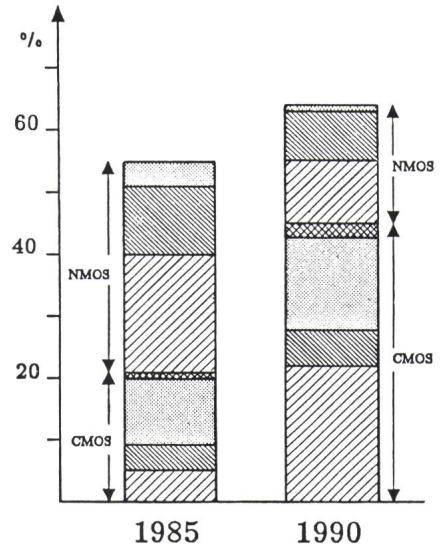
(p-well) plaziert werden. Durch den p-n-Übergang, der zwischen der p-Wanne und dem n-Substrat entsteht, werden Minoritätsträger, die von α -Teilchen im Substrat generiert werden, «abgesaugt» und bleiben so ohne Wirkung auf den Inhalt der einzelnen Speicherzellen (Fig. 2).

b) Hot-electron-Effekte: Dieser zweite Vorteil der CMOS-Technologie betrifft die Erzeugung von heissen Elektronen, d. h. von Elektronen hoher kinetischer Energie, die durch das transversale elektrische Feld am Rande des Drain-p-n-Überganges, unmittelbar unter der Gate-Elektrode eines n-Kanal-Transistors, erzeugt werden.² Da der CMOS-Inverter (Fig. 1) nur im dynamischen Betrieb Strom führt, entstehen in dessen n-Kanal-Transistor heisse Elektronen nur während der kurzen Transitionszeit des Eingangssignals. Beim p-Kanal-Transistor spielt der Hot-Ladungsträger-Effekt praktisch keine Rolle (Löcher als Ladungsträger). Im Gegensatz zum CMOS-Inverter fließen durch einen NMOS-Inverter relativ hohe statische Ströme. Somit werden heisse Elektronen kontinuierlich generiert, wodurch die Lebensdauer der einzelnen Transistoren in hohem Masse beeinträchtigt wird.

c) Elektromigration: Darunter versteht man den Transport von Metallionen in den Leiterbahnen (gewöhnlich aus Aluminium) unter der Wirkung von hohen Stromdichten. Da diese im Falle eines CMOS-IC klein gehalten werden können, kann auch die Zerstörung der Leiterbahnen durch Elektromigration entscheidend vermindert werden.

Die Nachteile der CMOS-Technologie sind einerseits kleinere Packungsdichte und andererseits das Latch-up-Phänomen, bei dem durch äussere Störeinflüsse Halbleiterübergänge verboteenerweise durchschalten, wodurch hohe Ströme fließen, die die Schaltung zerstören können. Diese Nachteile [2] lassen sich aber mit geeigneten Massnahmen stark verringern, so dass die oben erwähnten Vorteile sie zur bestgeeigneten Technologie für VLSI-Anwendungen machen. Die Figur 4 zeigt den Marktanteil der CMOS-Technologie für 1985 sowie den extrapolierten Wert für 1990.

² Dieser wichtige Effekt kann hier aus Platzgründen nicht näher erklärt werden.



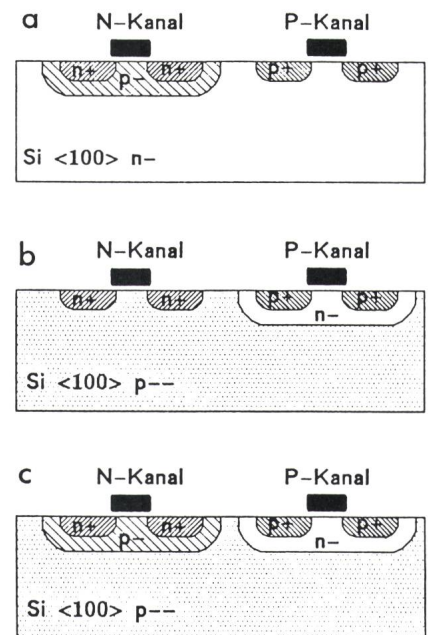
Figur 4 MOS-Marktanteil weltweit

Memory MPU Logic Linear

2. CMOS-Prozesse

2.1 n-, p- und Doppelwannentechnik

In den siebziger Jahren erscheinen erstmals CMOS-Prozesse mit einer p-Wanne (p-well) (Fig. 5a), die in ein



Figur 5 CMOS-Technologien

- p⁻ B-dotiertes Si, 10¹⁴ cm⁻³
- p⁻ B-dotiertes Si, 10¹⁶ cm⁻³
- p⁺ B-dotiertes Si, 10¹⁹ cm⁻³
- n⁻ P-dotiertes Si, 10¹⁵ cm⁻³
- n⁺ As-dotiertes Si, 10²⁰ cm⁻³
- a p-well-Technologie
- b n-well-Technologie
- c Twin-Tub-Technologie

n-Substrat eindiffundiert ist. In diese Wanne sind n-Kanal-Transistoren eingebracht, die so von den p-Kanal-Transistoren, die im Substrat gebildet sind, elektrisch isoliert sind. Später, in den achtziger Jahren, sind n-Wannen-Prozesse (Fig. 5b) entwickelt worden, die sehr schnell einen breiten Anwendungsbereich gefunden haben. Eine dritte Möglichkeit, CMOS-Strukturen zu bauen, besteht darin, beide, n- und p-Kanal-Transistoren, in unterschiedliche Wannen zu plazieren. Diese Prozesse mit p- und n-Wannen im selben Siliziumkristall sind als «Twin-Tub»- oder Doppelwannen-Prozesse bekannt (Fig. 5c).

Welche von diesen drei Möglichkeiten die geeignete ist, hängt von der jeweiligen Schaltungsanwendung ab. Aspekte wie langfristige Zuverlässigkeit und Schaltungsdesign müssen berücksichtigt werden. Im folgenden wird kurz auf die wichtigsten Unterschiede dieser drei Lösungen in bezug auf Schaltungsanwendungen und technologische Standpunkte eingegangen.

Unter dem Einfluss desselben elektrischen Feldes bewegen sich Elektronen zwei- und dreimal so schnell wie Löcher. Diese physikalische Eigenschaft [3] ist der Hauptgrund für den Unterschied in der Steilheit von p- und n-Kanal-Transistoren. Demzufolge liefert, unter gleichen Drain- und Gate-Spannungen, ein n-Kanal-Transistor 2- bis 3mal so viel Strom wie ein gleich dimensionierter p-Kanal-Transistor. Zusätzliche Dotierungseinflüsse, die hier nicht näher behandelt werden, bewirken, dass diese Eigenschaft sich bei Invertieren in p-Wannentechnologie weniger auswirkt als bei solchen in n-Wannentechnologie. Die ersteren zeigen deshalb symmetrischere Schaltcharakteristiken, eine Eigenschaft, die vor allem für statische logische Anwendungen vorteilhaft ist.

n-well-Prozesse werden da angewendet, wo die CMOS-Komplementarität nicht vollständig ausgenutzt wird. Im Fall von dynamischen RAM (DRAM) z. B. werden gewöhnlich die Cell Arrays in ein p-Substrat eingebaut. Die Peripherie der Speicher wird ausschliesslich in CMOS entworfen. Stark gefördert wurde die n-Kanal-Technologie auch durch die Tatsache, dass viele heutige CMOS-ICs als Redesign oder Weiterentwicklungen von NMOS-Schaltungen entstanden sind. Deshalb wiegen bei dieser Technologie n-Kanal-Transistoren (Enhancement- und Depletion-Typ) vor. Nur auf der

Peripherie halb der Schaltung wird, um Strom zu sparen, die CMOS-Komplementarität genützt. Die n-well-Technologie gewährt also eine gewisse Kontinuität der in vielen Halbleiterfirmen etablierten NMOS-Designkultur. Zusätzlich sind der Vollständigkeit halber als typische Applikationen der n-well-CMOS-Prozesse die EPROM-ICs zu erwähnen. Bei diesen bestehen die einzelnen Speicherzellen aus n-Kanaltransistoren, deren Gate nach aussen hin isoliert ist (n-Kanal-Floating-Gate-Transistoren). Mit Hilfe des bereits früher erwähnten Hot-Electron-Effektes werden bei der Programmierung Elektronen aus der Kanalzone durch die Gate-Oxid-Isolation hindurch in das Gate hineingeschossen, von wo sie nur mit Hilfe von UV-Strahlung abfliessen können. Die n-Transistoren werden bei relativ hohen Drainspannungen während der Programmierungszeit in Sättigung gebracht, so dass sehr grosse Substratströme fliessen. Die p-Kanal-Transistoren ermöglichen (zusammen mit den n-Kanal-Transistoren) verlustarme CMOS-Interfacefunktionen.

Ein p-well-Prozess kommt bei dieser Anwendung nicht in Frage, da der erwähnte Substratstrom die Schaltung leicht in Latch-up bringen könnte.

Die dritte erwähnte Möglichkeit ist die Doppelwannen-Technologie. Diese wird, je nachdem, ob man ein n- oder p-Substrat als Ausgangsmaterial verwendet, Ähnlichkeiten mit einer p- oder n-well-Technologie aufweisen.

Ein entscheidender Vorteil des Twin-Tub-Prozesses ergibt sich aus der Tatsache, dass beide Wannendotierungsprofile unabhängig voneinander optimiert werden können. Die n- und p-Oberflächenkonzentration kann, um Punch-through³ zu vermeiden, genügend heraufgesetzt werden. Dieser Vorteil geht auf Kosten des Body Factors⁴ und der Junction-Kapazitäten, die entsprechend erhöht werden. Bei der Twin-Tub-Technologie hat die Wahl des Substrattyps keinen Einfluss auf den Prozess. Diese Flexibilität kann aus fabrikationsstrategischen Überlegungen für den IC-Hersteller von grosser Bedeutung sein.

³ erhöhte Leckströme

⁴ Substrateffektkonstante; sie ist eine Funktion von Gate-Kapazität und Oberflächenkonzentration.

Als zusätzliches Argument für die Twin-Tub-Technologie spricht die Tatsache, dass sich damit besser bipolare Transistoren (für Ausgangstufen) realisieren lassen [3].

2.2 Hochenergetische Ionenimplantation

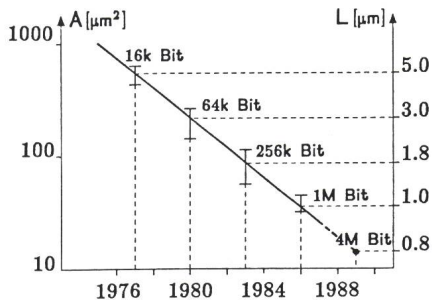
Um sehr hohe Packungsdichten (Submikron-Prozesse) zu erreichen, werden neuerlich die Wannen nicht mehr durch Implantation⁵ mit anschliessender Diffusion gebildet, sondern einzig durch Implantation bei sehr hohen Energien (300 keV für Bor und bis 1 MeV für Phosphor). Die endgültige Eindringtiefe der Wannen ist auf diese Weise nur durch die Implantationsenergie bestimmt. Dadurch wird die in der konventionellen Implantation-Diffusionsmethode entstandene laterale Diffusion eliminiert. Der minimale Abstand zwischen den n+ - und p+ -Gebieten (Drains der n- und p-Kanal-Transistoren) kann deshalb bei Verwendung dieser Technik (als Retrograde Wells bekannt) bis auf 3 µm reduziert werden. Zudem werden die Source-Drain-Junctionkapazitäten und der Body Factor (Substrateffektkonstante) herabgesetzt⁶. In den Feldoxidgebieten (ausserhalb der aktiven Transistorfläche) mit etwa 600 nm Oxiddicke liegt dagegen bei der hochenergetischen Ionenimplantation die maximale Dotierungskonzentration der n- und p-Wanne gerade am Übergang von SiO₂ zu Si, so dass die parasitären Transistoren ohne zusätzliche Implantationen (channel stoppers) eine genügend hohe Schwellspannung (10...15 V) erreichen.

3. Miniaturisierung

In Figur 3 ist der Verlauf des Integrationsgrades über die letzten 15 Jahre aufgezeichnet. Hohe Integrationsdichten sind in den meisten Fällen nur mit entsprechender Verkleinerung der Abmessungen der IC-Bauelemente erreichbar. In Figur 6 wird der Entwicklungsverlauf der minimalen Linienbreite (design rules) sowie der entsprechenden Zellfläche eines DRAM-

⁵ Dotierungsmethode, bei der die Ionen mit hoher Energie ins Silizium eingeschossen werden.

⁶ Weil die max. Konzentration der Wanne unterhalb der Si-Oberfläche liegt.



Figur 6 DRAM-Zellengröße

L Linienbreite
A Zellengröße

Speichers dargestellt. Die Tabelle I enthält die wichtigsten geometrischen Parameter der verschiedenen RAM-Generierungen. DRAM sind Standardprodukte, die in grossen Mengen produziert werden. Sie sind aus diesem Grund für IC-Hersteller das ideale Vehikel, um rasch Fabrikationserfahrung über neue Prozesse und Fabrikationstechniken zu gewinnen. Wegen der Marktdominanz der japanischen Firmen im DRAM-Sektor sahen sich aber in den letzten Jahren viele USA-Firmen gezwungen, auf andere Technologievehikel wie EPROM und SRAM auszuweichen. Trotzdem bleiben DRAM die idealen Vehikel, um die Prozessentwicklung voranzutreiben; sie werden auch in Zukunft als Mass für den aktuellen Technologiestand dienen. Der Mbit-DRAM-Chip ist schon heute kommerziell verfügbar. Spätestens um 1990 wird der 4-Millionen-Transistor-Chip hergestellt. Gleichzeitig bemühen sich europäische Firmen, einen Superchip mit über 6 Millionen Transistoren in Submikron-CMOS-Technologie, den 1-Mbit-SRAM-Chip, zu bauen.

4. Wirtschaftliche Aspekte der IC-Herstellung

Die Herstellung von ICs wird allgemein in die Hauptphasen Wafer-Fabrikation (Wafer Processing), Wafer-Vormessen (Wafer Pretesting), Chip-

Montage (Assembly) und Endmessen (Final Testing) eingeteilt.

Während der Fabrikation der Siliziumscheiben wird durch Waferbrüche, Prozess- und Handhabungsfehler die Endzahl der Si-Scheiben um einen bestimmten Prozentsatz verringert. Im Durchschnitt liegt dieser Verlust für nicht automatisierte Diffusionslinien bei 5 bis 15%.

Der zweite Ausbeutefaktor wird durch den Anteil der Chips einer Scheibe, die den elektrischen Test bestehen, bestimmt. Diese Operation, Vormessen genannt, ist die bedeutendste und entscheidendste. Sie widerspiegelt am deutlichsten die Fabrikationsqualität, die Prozesszuverlässigkeit und die Ausbildung des Personals. Die Vormessausbeute ist stark von der Prozesskomplexität, der Packungsdichte und der Chipfläche abhängig. In der Regel variiert sie zwischen 10% für extrem grosse Chips und 95 % für einfache digitale ICs. Ein oft verwendetes Ausbeutemodell ist die Poisson-Verteilung

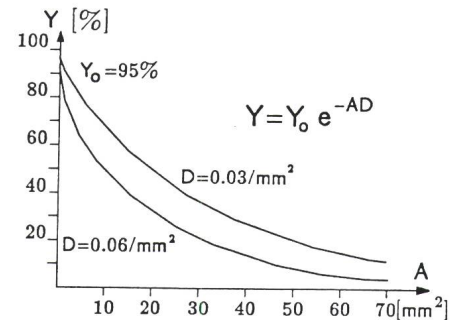
$$Y = Y_0 e^{-DA} \quad (1)$$

mit

Y_0 Parameterausbeute (%)
 D Defektdichte (mm^{-2})
 A Chipfläche (mm^2)

Y_0 sollte für Prozesse, welche die Produktionseinführungsphase absolviert haben, praktisch 100% erreichen. Die Defektdichte ist der entscheidende Parameter, der über die gesamte Qualität einer Diffusionslinie Auskunft gibt. Die meisten IC-Produzenten geben verständlicherweise keine Informationen über Vormessausbeutezahlen. «State of the Art» im Bereich von 2- bis 3- μm -Designregeln ist eine Defektdichte von 3 bis 6 Defekten pro cm^2 je nach Prozesskomplexität und Packungsdichte.

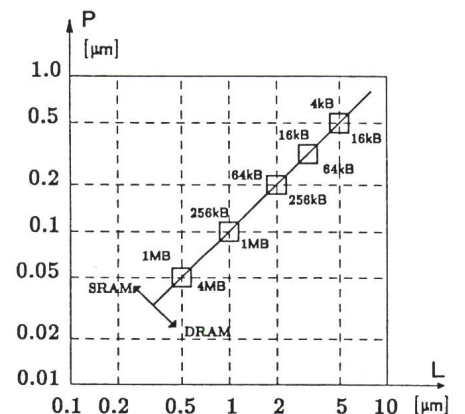
In Figur 7 ist die Vormessausbeute nach dem oben erwähnten Modell in Funktion der Chipfläche für 3 und 6 Defekte pro cm^2 dargestellt. Das starke



Figur 7 Vormessausbeute in Funktion der Chipfläche

Y Vormessausbeute (Yield)
A Chipfläche

Absinken der Ausbeute für grosse Chips ist leicht zu erkennen. Als Faustregel kann angenommen werden, dass die effektive Defektdichte für eine bestimmte Diffusionslinie ungefähr proportional zur Packungsdichte wächst. Gleich grosse Schaltungen, die mit unterschiedlichen Packungsdichten in der gleichen Diffusionslinie produziert werden, erreichen aus diesem Grund verschiedene Ausbeuten.



Figur 8 Zusammenhang zwischen Integrationsgrad (Linienbreite) und kritischer Partikelgröße

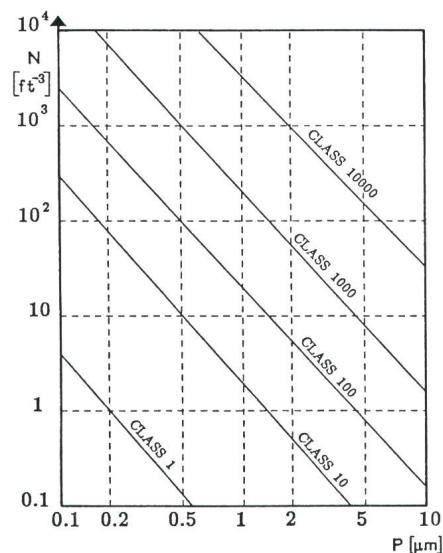
Die Figur 8 zeigt die ungefähre Korrelation zwischen dem Integrationsgrad, gemessen an minimaler Linienbreite und Grösse der Partikel, die im Reinraum noch toleriert werden können. Die Partikel, welche die Siliziumscheibenoberfläche kontaminieren, stammen aus der Luft, von Chemikalien oder von den Einrichtungen.

Die Kontamination aus der Luft geschieht hauptsächlich während des Handlings (manuell oder mit Robotern). Eine weitgehende Vermeidung dieser Kontamination gelingt nur durch den Einsatz einer fortschrittli-

DEVICES	DRAM bit	64k	256k	1M	4M	16M
	SRAM bit	16k	64k	256k	1M	4M
Gate length [µm]		3	2	1.2	0.8	0.5
Gate oxide thickness [nm]		70	45	25	15	10
Capacitor oxide thickness [nm]		40	25	15	--	--
n+ junction depth [µm]		0.4	0.2	0.2	0.2	0.15
p+ junction depth [µm]		0.8	0.35	0.35	0.2	0.15
Well depth [µm]		9	5	3	2	1.5

Tabelle I Speicher-Prozessparameter

DRAM Dynamic Random Access Memory
SRAM Statische RAM



Figur 9 Reinraumdefinition
 N Anzahl Partikel pro Kubikfuss
 P Partikelgrösse

chen Automatisierung in extrem sauberen Räumen (Cleanrooms). Das Diagramm der Figur 9 zeigt den Zusammenhang zwischen Partikelgrösse und Partikeldichte für verschiedene Sauberkeitsgrade (Klassen) der Luft. So bedeutet z.B. Reinraum der Klasse 10, dass es in einem Kubikfuss dieser Luft nicht mehr als 10 Partikel, die grösser als 0,5 µm sind, geben darf.

In den neunziger Jahren zu bauende IC-Linien für VLSI werden Räume der Klasse 1 benötigen, was nur durch den Einsatz von Robotern erreicht werden kann. Der Aufwand an Kapital und Entwicklung für solch hochgezüchtete Produktionsstätten ist derart hoch, dass er sich aus wirtschaftlichen Gründen kaum rechtfertigen lässt. Als Schlüsseltechnologie ist die Mikroelektronik jedoch aus strategischen Gründen für die meisten Elektronikkonzerne zur absoluten Notwendigkeit geworden.

5. Technologieanforderungen

Die sukzessive Verkleinerung der Transistorgeometrie und der Leiterbahnenbreite verlangt zunehmend raffiniertere Produktionstechniken, insbesondere was die Photolithographie, das Plasmaätzen, die Ionenimplantation und die Metallisierung anbetrifft. Der beschränkte Rahmen dieses Artikels gestattet nur, eine kurze Zusammenfassung der erwähnten Technologie zu geben. Bei der Photolithographie sind zu erwähnen:

Tabelle II
Belichtungssysteme

Belichtungssystem	Linienbreite	Richtgenauigkeit
Contact aligners	6.0 - 4.0 µm	1.5 µm
Projection aligners	4.0 - 2.5 µm	1.0 µm
Projection steppers	2.5 - 1.5 µm	0.6 µm
Reduction steppers	1.5 - 0.5 µm	0.3 µm

- Kontaktbelichtung-Maskaligners (Maske in Kontakt mit Si-Scheibe)
- Projektionsbelichtung-Maskaligners (Maske nicht in Kontakt mit Si-Scheibe)
- Wafersteppers (sukzessive Teilbereichsbelichtung)
- Röntgenstrahl-Lithographie
- Elektronenstrahl-Lithographie.

Eine Zusammenfassung der Eigenschaften und Anwendungsbereiche der ersten drei Verfahren ist in Tabelle II gegeben. Die Elektronenstrahl- und Röntgenstrahlolithographie sind im Vergleich zu den optischen Verfahren extrem aufwendig und immer noch unwirtschaftlich. Die ständige Verbesserung der optischen Systeme (Wafersteppers) schiebt den Zeitpunkt ihrer Anwendung immer weiter in die Zukunft hinaus. Die untere Grenze des Auflösungsvermögens für optische Systeme liegt schätzungsweise bei 0,5 µm. Unterhalb dieser Grenze werden neue Entwicklungen, unter anderem die Röntgen- und Elektronenstrahlolithographie, zur Anwendung kommen.

Die physikalische Grenze der Miniaturisierung liegt ungefähr bei 0,25 µm [4]. Solch kleine Dimensionen erfordern völlig neue Fabrikationstechniken in komplett menschenleeren Produktionslinien. Es ist schwierig, den Zeitpunkt für den Produktionsstart dieser sog. ULSI-Technologien (Ultra-Large-Scale-Integration-T.) vorauszusagen. Obwohl die grossen IC-Hersteller sowie die renommierten Forschungslaboratorien sich schon heute mit der Entwicklung von 0,25-µm-Prozessen beschäftigen, so ist doch deren Einführung in die Produktion kaum vor 10 Jahren zu erwarten.

Was das Plasmaätzen anbetrifft, so wird sich die Entwicklung in den nächsten Jahren auf vollautomatische Single-Wafer-Systeme konzentrieren, die folgende Bedingungen erfüllen müssen:

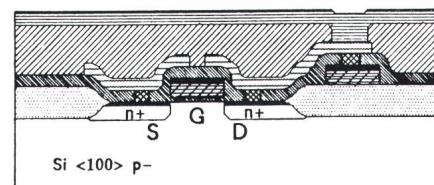
- sehr gute Ätzgleichmässigkeit,
- hohe Ätzselektivität (hohe Ätzrate des zu ätzenden Materials gegenüber der Ätzrate des unterliegenden Substrats),

- Anisotropische Ätzeigenschaften (linienbreite Übertragung von Photolack auf das zu ätzende Material im Verhältnis 1:1),
- grosse Photolackerosionsfestigkeit,
- keine Beeinträchtigung der Schaltungsbaulemente durch im Plasma generierte Ionen,
- grosse Zuverlässigkeit des Systems,
- niedrige Partikelkontamination.

6. Neue Entwicklungen in der VLSI-Technologie

6.1 Metallisierung

Um die enorme Anzahl von Transistoren auf einem VLSI-Chip untereinander zu verbinden, braucht man Metallisationsebenen, in denen ein dichtes Netz von Leiterbahnen Millionen von Diffusionsgebieten und Polysiliziumbahnen (Transistorgates und Interkonnexionen) verbinden. Um eine Verringerung der Siliziumfläche, mehr Designflexibilität und höhere Schaltgeschwindigkeit zu erreichen, werden komplexe integrierte Schaltungen mit doppelter oder sogar dreifacher Metallebene versehen. Den Querschnitt durch eine Doppelmetallstruktur zeigt Figur 10. Um den Fabrikationsprozess



Figur 10 Doppelmetallsystem

- ☒ Wolfram-Kontaktfüllung
- ▨ Polysilizium
- TiSi₂
- ▤ Metall 1 (Al)
- ▥ Metall 2 (Al)
- ▧ Planarisiertes SiO₂
- ▩ LOCOS
- SiO₂

zu vereinfachen, wird gewöhnlich die zweite Metallebene nur Verbindungen mit der ersten Metallebene aufweisen. Die Isolationsschicht zwischen den beiden Ebenen ist meistens aus Siliziumoxid oder Siliziumoxidnitrid. Das am meisten verwendete Metall in der IC-Fabrikation ist ohne Zweifel Aluminium, das vorwiegend als Al/Si- oder als Al/Cu-Legierung eingesetzt wird. Silizium in Aluminium hemmt die gegenseitige Diffusion von Alu und Si und verhindert so die Zerstörung der p-n-Übergänge. Das Kupfer wird dazugegeben, um Elektromigration in den Aluminiumbahnen zu vermindern. Mit der Anwendung von Metallsiliziden auf Diffusions- und Polysiliziumgebieten ist neulich die Verwendung von Wolfram anstelle von Aluminium für die erste Metallebene in den Vordergrund bei der Entwicklung neuer Metallisierungssysteme gerückt.

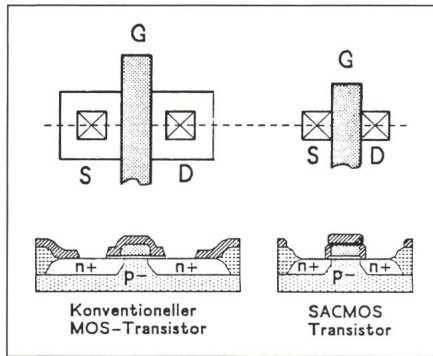
6.2 Metallsilizide

Metallsilizide sind Metall-Silizium-Verbindungen, die wegen ihrer hohen Leitfähigkeit, ihrer thermischen Stabilität und ihrer Prozesskompatibilität im Herstellungsverfahren integrierter Schaltungen auf grosses Interesse stossen. Gerade die Hauptprobleme, die bei der Reduktion der IC-Dimensionen auftauchen, wie zu hoher Widerstand der Polysiliziumleiterbahnen und der Diffusionsgebiete, zu hoher Metall-Silizium-Kontaktwiderstand sowie ungenügende Langzeitstabilität, lassen sich bei der Anwendung von Metallsiliziden besser bewältigen.

Auf Tabelle III sind die am meisten verwendeten Metallsilizide in der

Material	Spezifischer Widerstand [$\mu\Omega\text{cm}$]
Polysilizium	500 - 1000
MoSi ₂	80 - 100
WSi ₂	50 - 70
NiSi ₂	50 - 60
TaSi ₂	35 - 55
Ti	40
TiN	30 - 40
TiSi ₂	12 - 25
CoSi ₂	10 - 20
W	5
Mo	5
Al	2.5

Tabelle III Spezifischer Widerstand von Metallsiliziden und Kontaktierungsmaterialien



Figur 11 Vergleich zwischen konventioneller und SACMOS-Technologie

- Kontaktloch
- Diffusionsgebiet
- Gateelektrode

CMOS-IC-Technologie zusammengestellt. Wegen seiner hohen Leitfähigkeit und des relativ einfachen Bildungs-Verfahrens wird dem TiSi₂ in der zukünftigen VLSI-Silizidtechnologie eine wichtige Rolle zukommen [5]. Selbstjustierende Silizidprozesse verwenden z.B. Ti-Schichten, die durch Aufdampfen oder durch Ionenzerstäubung auf den IC deponiert werden. Während der anschliessenden thermischen Behandlungen wird das Ti auf den offenen Siliziumflächen (Polysiliziumleiterbahnen und Diffusionsgebieten) eine niederohmige TiSi₂-Schicht bilden. Auf den nicht aktiven Flächen verbliebenes Ti wird selektiv entfernt.

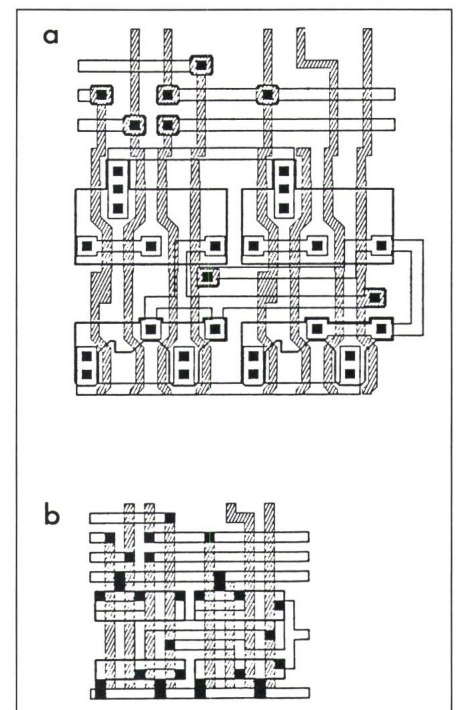
Gute Kontaktierung von extrem un-tiefen p-n-Übergängen (Shallow Junctions) (0,1...0,2 μm), die als Source und Drain von VLSI-MOS-Transistoren dienen, sind mit der traditionellen Al/Si-Metallisierung praktisch unmöglich. Trotz Befügung von Si im Al werden Shallow Junctions, betrieben bei hohen Stromdichten, in ihrer Funktion stark beeinträchtigt, wenn sie mit Al/Si kontaktiert werden.

Shallow Junctions lassen sich zuverlässig nur mit Siliziden kontaktieren [6]. Ein interessantes Silizid-Metallisierungssystem, bestehend aus TiSi₂/TiN/Al, zeigt extrem niedrige Kontaktwiderstände auf p+ - und n+ -Silizium und ist thermisch bis 550 °C stabil [7]. Wegen der während der Montage (Auflöten und Verpacken) auftretenden Temperaturen ist die Anforderung an die Temperaturbeständigkeit des Metallisierungssystems eines IC von grosser Bedeutung.

6.3 Selbstjustierende Kontakte (SACMOS-Technologie)

Andere Entwicklungen der Siliziumgate-CMOS-Technologie zielen ebenfalls in Richtung einer weiteren Erhöhung der Packungsdichte und Schaltgeschwindigkeit sowie einer Verminderung des Stromverbrauchs. Zusätzlich zur Miniaturisierung wollen neue Ideen das Layout eines MOS-Transistors grundsätzlich verändern [8]. In Figur 11 sind als Vergleich die Layouts und Querschnitte der konventionellen MOS-Transistoren und des SACMOS-Transistors (Self-Aligned Contacts) illustriert. Beim SACMOS-Transistor sind die Abstände zwischen den Kontaktlöchern und dem Diffusionsrand sowie den Kontaktlöchern und dem Gaterand eliminiert. Statt dessen wird die Gateelektrode durch Siliziumoxid und Nitrid von Source- und Drainkontakten isoliert. Auf diese Weise erreicht man für eine bestimmte minimale Gatelänge die absolute minimale Transistorfläche.

Anhand eines Layoutbeispiels wird in Figur 12 der Einfluss auf die Pak-



Figur 12 Packungsdichte

- a Konventioneller CMOS-Layout
- b SACMOS-Layout

- Diffusion
- Metall
- Kontakt
- Polysilizium

kungsdichte bei Anwendung der SAC-Struktur ersichtlich [9]: Die benötigte Si-Fläche für gleiche Minimalgeometrie wird um die Hälfte reduziert.

Ein 2- μm -SACMOS-Prozess ermöglicht die Herstellung von Schaltungen mit einer Packungsdichte von 1600 Gates/ mm^2 . Dieser Wert ist mit konventionellen CMOS-Prozessen erst bei 1- μm -Designregeln zu erreichen. Self-Aligned-Contacts-Prozesse im Bereich von 1 μm Gatelänge würden Integrationsdichten ermöglichen, die konventionellen Submikrontechnologien entsprechen.

7. Schlussfolgerungen

Die Miniaturisierung der Strukturen eines IC ist aus wirtschaftlichen und

technischen Gründen die Hauptmotivation in der Entwicklung der VLSI-Technologie. Seit Anfang der siebziger Jahre halbieren sich die Masse der Strukturen durchschnittlich alle vier Jahre, und infolgedessen steigt die Integrationsdichte alle sieben Jahre um das Zehnfache. Diese Gesetzmässigkeit wird in den nächsten Jahren weiter bestehen. Sie wird die Entwicklung neuer Produktionsmethoden sowie die Erfindung neuer technologischer Lösungen fördern. Diese werden vor allem im Bereich der Metallisierung, der Photolithographie, des Plasmaätzens, der Kontaktierungstechnik und der Devicestruktur intensiv erforscht. Da das Scalingprinzip im Submikronbereich auf klare physikalische Grenzen stösst, wird die Entwicklung der VLSI-Technologie, besonders in bezug auf die Anwendung neuer Devicestrukturen, andere Wege einschlagen müssen.

Literatur

- [1] *W.J. McClean*: ICE status 1986. A report on the integrated circuits industry. Scottsdale/Arizona, Integrated Circuit Engineering Corporation (ICE), 1986.
- [2] *J.Y. Chen*: CMOS—the emerging VLSI technology. IEEE Circuits and Devices Magazine 2(1986)2, p. 16...31.
- [3] *G. Zimmer*: CMOS-Technologie. 2. Auflage. München, Verlag R. Oldenbourg, 1987.
- [4] *R.B. Fair* a.o.: Modeling physical limitations of junction scaling for CMOS. IEEE Trans. ED 31(1984)9, p. 1180...1185.
- [5] *H. Kaneko*: Novel submicrometer MOS devices by self-aligned nitridation of silicide. IEEE Trns. ED 33(1986)11, p. 1702...1709.
- [6] *H. Melchior*: Stabile Kontaktstrukturen mit Diffusionsbarrieren für integrierte Schaltungen. Mitteilungen AGEN—(1986)44, S. 13...23.
- [7] *M. Wittmer* and *H. Melchior*: Applications of TiN thin films in silicon device technology. Thin Solid Films 93(1982)—, p. 397...405.
- [8] *J. Solo de Zaldivar*: CMOS-Prozess mit selbstjustierenden Kontakten. NTG-Fachberichte 77(1981), S. 22...25.
- [9] *R.E. Lüscher* and *J. Solo de Zaldivar*: A high density CMOS process. IEEE International Solid-State Circuits Conference 1985, p. 260...261 + 361.