

Prozeßmodule für produktionstaugliche Submikrometer-MOS-Technologien

H. Noll

Austria Mikro Systeme International AG

1. Einleitung

Die Herstellung integrierter Schaltkreise in MOS- bzw. CMOS-Technologien hat sich weltweit zu einem der bedeutendsten Wirtschaftszweige entwickelt. Der Gesamtmarkt beträgt 1995 etwa 100 Mrd. US \$ und hat ein Steigerungspotential von 20–30% jährlich. Das nächste Jahrzehnt wird vom Informationszeitalter geprägt und das Verlangen nach höherer Integrationsdichte, um komplexe elektronische Systeme auf den Chip zu bringen, nimmt ständig zu. Gigabit Speicherchips werden um das Jahr 2000 produktionsmäßig zur Verfügung stehen. Das Hauptaugenmerk wird auf den kostenmäßigen Entwicklungsaufwand der Herstellungstechnologien und die Vereinheitlichung gerichtet. Dies resultiert aus der Forderung nach schnelleren und dichter gepackten Schaltkreisen. Um diese Ziele zu erreichen, werden weltweit Anstrengungen unternommen, um die Prozeßtechnologien bis in den sogenannten „deep submicron“ Bereich voranzutreiben. In Abb. 1 ist eine „CMOS Technology Roadmap“ abgebildet, wie sie sich aus heutiger Sicht darstellt.

CMOS Technology Roadmap

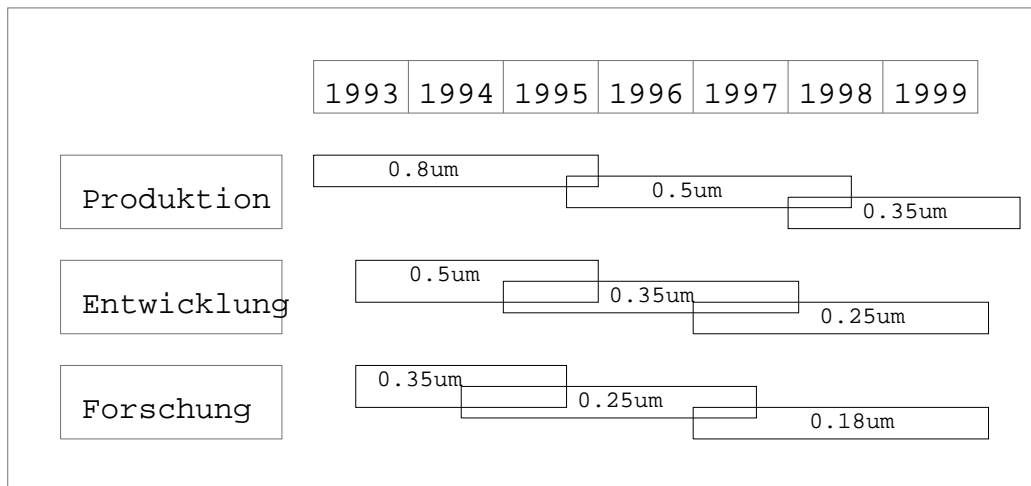


Abb. 1: CMOS Technology Roadmap

Ausgehend von konventionellen Prozessen mit Strukturgrößen im Mikrometerbereich bedeutet das Vordringen in Submikrometer-Prozesse einen enormen Technologiesprung, was den Einsatz neuer Materialien für elektrisch aktive Dünnschichten und Dielektrika, Maschinen und Prozeßmodule betrifft. Bei MOS-Transistoren treten — physikalisch bedingt — Kurzkanaleffekte in den Vordergrund, Dünnschichten müssen

mit höchster Reinheit und Qualität im Nanometerbereich abgeschieden werden und Maschinen arbeiten teilweise an den physikalischen Grenzen.

Im folgenden werden die wichtigsten Prozeßmodule, die heute und in naher Zukunft für produktionstaugliche und hochzuverlässige Submikrometer-Schaltkreise eingesetzt werden, vorgestellt:

2. Lithographie

Die Belichtungs und Fotolacktechnologie ist eine der wichtigsten und dynamischsten Bereiche im IC-Herstellungsprozeß. Sie stellt ca. 50% des Gesamtaufwandes dar. Fotolithographische Belichtungstechniken bis zu 0.18 μm Prozessen prägen das Bild. Elektronen-, Röntgen- und Ionenstrahlolithographie haben die Hürde der absoluten Notwendigkeit und der produktionstechnischen Umsetzung noch nicht überwunden.

Ausgehend von Wafer-Steppern mit einer Belichtungswellenlänge von 435 nm (g-line) über 365 nm (i-line), „deep UV“ bei 248 nm und Excimer-Laser-Belichtungsmethoden (KrF, ArF) wird vor allem an den äußerst kritischen Fotolacktechniken gearbeitet. Mehrlagen-Fotolack und Oberflächenbelichtungs-Techniken mit kombinierter naß-chemischer Entwicklung und Trockenätzung sind in der Lage, rein optische bedingte Limitierungen bezüglich Auflösungsvermögen und Fokustiefe weitgehend zu entschärfen. Hierzu zählen Prozesse wie CARL (Chemically Amplified Resist Lines), DESIRE (Diffusion Enhanced Silylation of Resist) und *Reverse Image* Verfahren. Auch neue Maskentechniken wie „Phase Shift Masks“ tragen wesentlich zur enormen Entwicklung auf diesem Gebiet bei.

3. Transistorstrukturen

Der Grund für die Skalierung von Transistorstrukturen in den Submikrometerbereich ist nicht nur die Erreichung höherer Packungsdichten, sondern auch die Erzielung höherer Schaltgeschwindigkeiten. Größere Transistorströme werden benötigt, um die Geschwindigkeit für das Laden und Entladen von parasitären Kapazitäten zu erhöhen. Dazu werden kürzere Kanalbreiten und hohe elektrische Gateoxid-Felder benötigt. Eine der wichtigsten Bedingungen bei der Wahl der geeigneten Transistor-Architektur ist die Minimierung der Leckströme im Sperrzustand (ca. 1 pA pro μm Transistorbreite) [2]. Erhöhte Leckströme treten vor allem durch Erniedrigung der Einsatzspannung (Drain Induced Barrier Lowering), welche durch die kurze Kanallänge bedingt ist, durch den sogenannten „bulk punchthrough“ und durch „gate induced drain leakage“ (GIDL) auf. Aufgrund von hohen elektrischen Feldstärken (Hot Electron Effekte) kommt es zu einer Degenerierung des Gate-Oxids und damit zu einer drastischen Lebensdauerverkürzung der Bauelemente. „Punchthrough“ und Erniedrigung der Einsatzspannung werden als typische Kurzkanaleffekte bezeichnet. Die beiden anderen Effekte sind elektrische Feldstärkeneffekte.

Um all diesen oben genannten, limitierenden Faktoren entgegenzutreten, werden verschiedene Transistor-Architekturen in Abhängigkeit von der benötigten Prozeßtechnologie eingesetzt. Dabei wird dem Drain/Source-Engineering besonderes Augenmerk gewidmet. „Lightly Doped Drain“ (LDD) Strukturen zur Verminderung der „Hot Electron“ Effekte [4, 5] werden angefangen von 1.2 μm bis hinunter zu 0.5 und 0.35 μm -Prozeßtechnologien erfolgreich eingesetzt. Diese Hot-Electron-Effekte werden in hohem Maße von der Versorgungsspannung (Vdd) beeinflusst. Deshalb ist auch unter

anderem eine Limitierung der Versorgungsspannung auf 3.3 Volt für 0.5 μm und 0.35 μm Prozesse, sowie 2.5 Volt für 0.25 μm und 1.5 Volt für 0.18 μm Prozesse, erforderlich.

Für 0.25 μm und 0.18 μm Technologien sind weitere vielversprechende Alternativen wie LATID (Large Angle Tilt Implanted Drain), FOND (Fully Overlapped Nitride-Etch Defined Devices) und GOLD Devices im Entwicklungsstadium.

Besondere Beachtung bei Submikrometer Device-Strukturen findet das Gate-Dielektrikum. Mit der notwendigen Verringerung der Transistor-Strukturen geht auch eine Verringerung der Dicke des Gateoxids einher, um die geforderten elektrischen Parameter wie Treiberfähigkeit und Schaltgeschwindigkeit zu erreichen. Während man sich bei Technologien bis 0.35 μm mit der Aufbringung von hochqualitativen thermischen Oxiden begnügt, werden für „deep submicron“ Devices sogenannte „Nitrided Gate Oxides“ eingesetzt. Die Vorteile liegen in den verbesserten elektrischen Eigenschaften und im niedrigerem thermischen Budget beim Aufwachsen des Dielektrikums (NO Atmosphäre bei 850°C). Die Dicke beträgt für 0.25 μm Prozesse ca. 6 nm. „Nitrided Gate Oxides“ bieten weiters eine bessere Barriere gegenüber der Bor Diffusion bei p-dotierten Poly-Gates. Weiters werden höhere QBD (Charge to Breakdown) Werte erreicht. Für die Herstellung werden „Rapid Thermal Processing“ Techniken verwendet.

4. Transistor-Isolationstechniken

Für die Isolation der einzelnen Bauelemente unterscheidet man zwei Kategorien: Erstens, die Isolation von Bauelementen, die sich in einer gemeinsamen Wanne befinden und parasitäre Leckströme zwischen benachbarten Bauelementen verhindern soll (z. B. n^+ zu n^+ Leckströme in einer p-Wanne). Zweitens, eine Isolation zwischen zwei benachbarten Wannern, die eine mögliche Wechselwirkung von NMOS- und PMOS-Transistoren verhindert.

Konventionelle LOCOS (Local Oxidation) Techniken werden seit mehr als zwanzig Jahren für MOS-Technologien benutzt. Die Limitierung dieses relativ einfach zu handhabenden Prozeßschrittes liegt in der Bildung des sogenannten „bird's beaks“, der geometrische Einschränkungen für eine Skalierung der aktiven Transistorgebiete mit sich bringt. Dieser „bird's beak“ ragt ca. 0.7 μm pro Seite in das aktive Gebiet und beeinflusst damit zu einem wesentlichen Teil die Packungsdichte der Bauelemente. Aus diesem Grund konzentrieren sich eine Reihe von Forschungs- und Entwicklungsaktivitäten auf die Reduktion des „bird's beak encroachments“. Im folgenden werden einige auf LOCOS basierende Isolationstechniken für Submikrometer-Technologiegenerationen dargestellt:

4.1 SWAMI (Side Wall Masked Isolation)

Der Prozeß beruht auf der Ätzung von konischen Silizium-Gräben (Trenches), gefolgt von der Strukturierung einer Nitrid/Oxid-LOCOS-Schicht. Die laterale Diffusion von Sauerstoff wird durch das Aufbringen einer dünnen Nitridschicht an den Seitenwänden des Siliziums weitgehend unterbunden. Dabei wird eine Reduktion des „bird's beaks“ auf 0.25–0.3 μm erzielt. Diese Technik ist bis zu 0.7 μm Prozesse einsetzbar.

4.2 PBL (Poly Buffered LOCOS) und FPBL (Framed Poly Buffered LOCOS)

Dieser Isolationsmodul wird industrieweit für die Herstellung von Schaltkreisen bis zu $0.5\ \mu\text{m}$ Prozeßtechnologien eingesetzt. Er entsteht durch die Einfügung einer Polysiliziumschicht zwischen der Nitridmaske und der ersten Oxidschicht (Pad Oxid) Dieser „Poly Buffer“ erlaubt die Einführung dünner Pad-Oxide für die Reduktion der lateralen Sauerstoffdiffusion. Die darauffolgende Entfernung der Polyschicht mittels Plasmaätzung kann aber sehr leicht das darunterliegende Silizium schädigen. Dadurch ergeben sich Limitierungen bezüglich der minimalen Oxiddicke, sodaß der „bird's beak“ nicht kleiner als $0.2\ \mu\text{m}$ gemacht werden kann.

Mit FPBL, einer ähnlichen Technik wie PBL, können „bird's beaks“ unter $0.15\ \mu\text{m}$ erreicht werden [3].

4.3 PELOX (Poly Encapsulated LOCOS)

Die Ausgangsschichten beim PELOX Prozeß bestehen aus einem Sandwich von $1400\ \text{Å}$ Nitrid und $500\ \text{Å}$ Oxid. Ein Naßätzschritt formt einen Graben im Oxid. Ein abermaliger Oxidationsschritt überzieht die geöffneten Siliziumflächen und eine darauffolgende Polysiliziumschicht füllt die Gräben. Nach der Feldoxidation (ca. $7000\ \text{Å}$) bildet sich ein „bird's beak“ der kleiner als $0.1\ \mu\text{m}$ ist. Diese enorme Verbesserung macht den Einsatz bis zu $0.35\ \mu\text{m}$ Technologien möglich.

4.4 Shallow Trenches

Diese Technologie gehört heute noch zu den schwierigsten. Nach einem Oxidationsschritt werden „Shallow Trenches“ von ca. $400\ \text{nm}$ Breite geätzt. Diese Strukturen werden danach mit einer etwa $100\ \text{nm}$ dicken Polysiliziumschicht ausgekleidet. Durch das Aufbringen von CVD-SiO₂ (z.B. BPSG oder TEOS) werden die „Trenches“ aufgefüllt. Man nennt diese Technologie auch BOX (buried oxide) Isolation. Die dadurch entstehende Dünnschichtlagen und deren Topographie müssen nun mittels geeigneter Methoden lokal entfernt und planarisiert werden. Plasma- bzw. naßchemische Rückätztechniken haben sich nicht besonders bewährt. Die derzeit einzige Möglichkeit wird in der CMP Technik (Chemical-Mechanical-Polishing) gesehen. Die Entwicklung eines geeigneten BOX-CMP Prozesses befindet sich noch im Anfangsstadium [3]. Man ist jedoch zuversichtlich, daß diese Art von Isolationstechnik für die Produktion von $0.25\ \mu\text{m}$ und $0.18\ \mu\text{m}$ Schaltkreisen eingesetzt werden wird.

Neben der Shallow-Trench-Isolation gibt es noch eine Reihe von weiteren Techniken, wie „Moderate Depth Trench and Refill Isolation“, „Deep Narrow Trench“ usw.

5. Kontakt und Verbindungs- (Interconnect-) Technologie

Die Kontakt- und Interconnect-Technologie wird als sogenanntes „back end“ bei der Herstellung von integrierten Schaltkreisen bezeichnet. Sie trägt in höchstem Maße zu den elektrischen Eigenschaften der Schaltkreise und zur Maximierung der erreichbaren Packungsdichte bei.

Um die Kontaktwiderstände und das unerwünschte „Junction Spiking“ (Diffusion von Aluminium ins Silizium und umgekehrt) wesentlich zu reduzieren, werden Mehrschicht-Kontaktstrukturen mit der Formation von Siliziden und Diffusionsbarrieren eingeführt. Für die Herstellung dieser komplexen Strukturen können verschiedene

Materialien wie Pt, W oder Ti eingesetzt werden. Im allgemeinen haben Submikrometer Kontaktstrukturen folgendes Aussehen:

Metallsilizid-Diffusionsbarriere (TiN, TiW usw.) – Metallsilizid – Aluminium [1,6]

Mit der Verkleinerung der Kontaktstrukturen zu Submikron-Geometrien müssen auch — bedingt durch die Limitierung der heutigen Metall-Sputtertechniken — neue Materialien und Verfahren für das Auffüllen der Kontaktlöcher eingesetzt werden. Verfahren wie „Planarized Aluminium“ oder „Tungsten CVD“ haben neue Maßstäbe bezüglich Zuverlässigkeit (Elektromigrationsverhalten, Kantenbedeckung und niederohmiges Widerstandsverhalten) gesetzt. Ähnliche Mehrschichtstrukturen und Techniken werden auch für die Verbindung von Mehrlagen-Metallisierungsebenen eingesetzt (Via-Kontakte).

Einen besonderen Stellenwert in der Interconnect-Technologie nehmen die Isolationschichten zwischen den einzelnen metallischen Verbindungsebenen ein. Mit der notwendig gewordenen vertikalen Skalierung dieser Dünnschichten werden auch erhöhte Anforderungen bezüglich zuverlässiger Isolation der elektrisch aktiven Leiterbahnen und der Unterdrückung von unerwünschten parasitären Kapazitäten gestellt. An Stelle von bis heute üblichen Phosphor-dotierten Plasmaoxiden treten BPSG (Bor-Phosphor-Silikatglas), TEOS (Tetraethylorthosilikat) und Ozon-TEOS-Schichten in den Vordergrund. Diese Materialien belasten aufgrund der Abscheideeigenschaften bei relativ niedrigen Temperaturen weitaus weniger das thermische Gesamtbudget des Prozesses und sind auch in der Lage Submikron-Leiterbahnabstände relativ zuverlässig aufzufüllen. Die dennoch auftretenden Unebenheiten (feinste Spalten und Löcher) können mit geeigneten Planarisierungstechniken eliminiert werden. Dafür stehen Technologien wie SOG („Spin On Glass“) mit anschließendem Rückätzen und CMP („Chemical Mechanical Polishing“) zur Verfügung.

6. Zusammenfassung

Der Einstieg in Submikrometer-Prozeßtechnologien bedeutet eine vollkommene Neuorientierung bezüglich der Einführung und Entwicklung neuer Materialien, Maschinen und Prozeßmodule. Ausgehend von der lateralen und vertikalen Skalierung der Transistorstrukturen in den tiefen Submikron-Bereich bedarf es — auch aufgrund physikalischer Limitierung — einer Erniedrigung der IC- Versorgungsspannungen. Konträr dazu kommt jedoch die Forderung zur Erhöhung der Transistortreiberfähigkeit. Um diesen Anforderungen gerecht zu werden, werden neue Transistor-Architekturen, komplexe fotolithographische Belichtungstechniken, neue Bauelement-Isolationstechniken, hochzuverlässige Dünnschichtmaterialien und Metallisierungsschemata eingeführt. Die in diesem Abstrakt grob dargestellten Prozeßmodule erheben in keiner Weise Anspruch auf Vollständigkeit. Die produktionstauglichen Prozeßmodule für 0.5 μm -Technologien sind heute — wie vorgestellt — größtenteils definiert. Prozeßvarianten für 0.25 μm bzw. 0.18 μm -Technologien sind teilweise noch im vollen Entwicklungsstadium. Die erfolgversprechendsten Ansätze dazu wurden diskutiert.

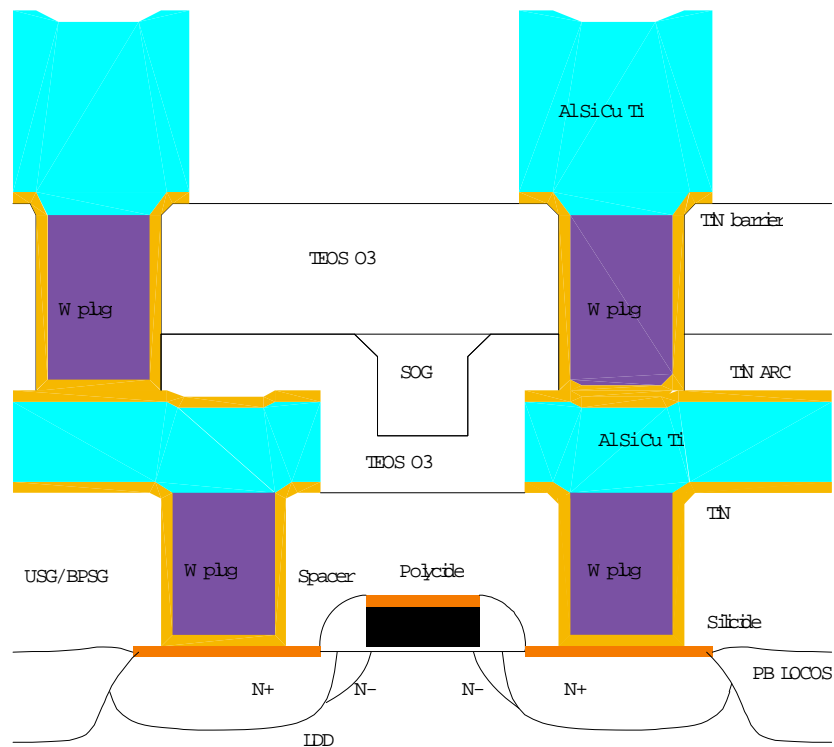


Abb. 2: Typischer Aufbau eines „state of the art“ 0.5 μm NMOS Transistors

Literaturverzeichnis

- [1] S. Wolf: „Silicon Processing for the VLSI Era“, Vol 1: „Process Technologie“, Vol 2: „Process Integration“, Lattice Press, 1990
- [2] S. M. Sze: „VLSI Technologie“, Mc Graw Hill International, 1983
- [3] IMEC — Persönliche Mitteilung
- [4] A. Bauer, H. Noll, H. Pimingsdorfer: „Hot Carrier Supression for an Optimized 10V CMOS Process“, SISDEP 93 Intl. Conference of Semiconductor Devices and Processes, Vienna, 1993
- [5] A. Bauer, H. Noll, H. Pimingsdorfer: „Optimizing Reliability of CMOS Processes Using Process and Device Simulation“, IASTED, Intl. Conf. on Modelling and Simulation, Pittsburgh, 1993
- [6] O. Stelmaszyk: „Untersuchungen zum Kontaktsystem Si/TiSi₂/TiW/Al für VLSI Schaltkreise“, Diplomarbeit TH Darmstadt, 1987