

Capítulo 16. Tecnologia Back End

Prof. Ioshiaki Doi
FEEC-UNICAMP
22/11/02

- Índice

- 16.1. Introdução
- 16.2. Contato
- 16.3. Interconexão e Vias
- 16.4. Dielétrico entre Metais

16.1. Introdução

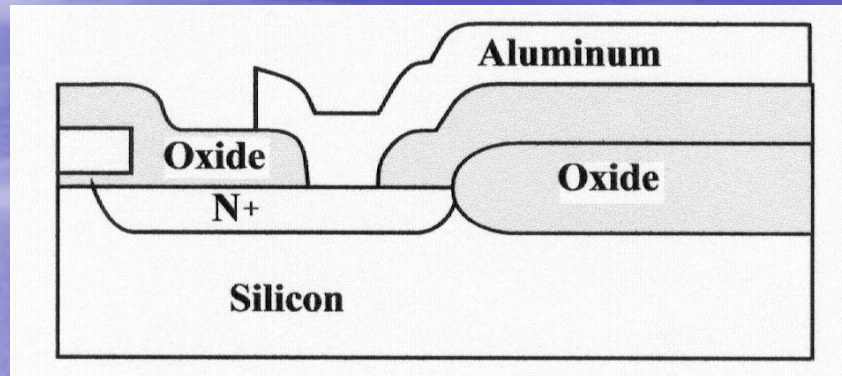
- Fabricação de um chip de CI : envolve centenas de etapas de processos.

- **Processos Front End**

- **Isolação;**
- **Formação de poços;**
- **Formação do transistor, etc.**

- **Processos Back End**

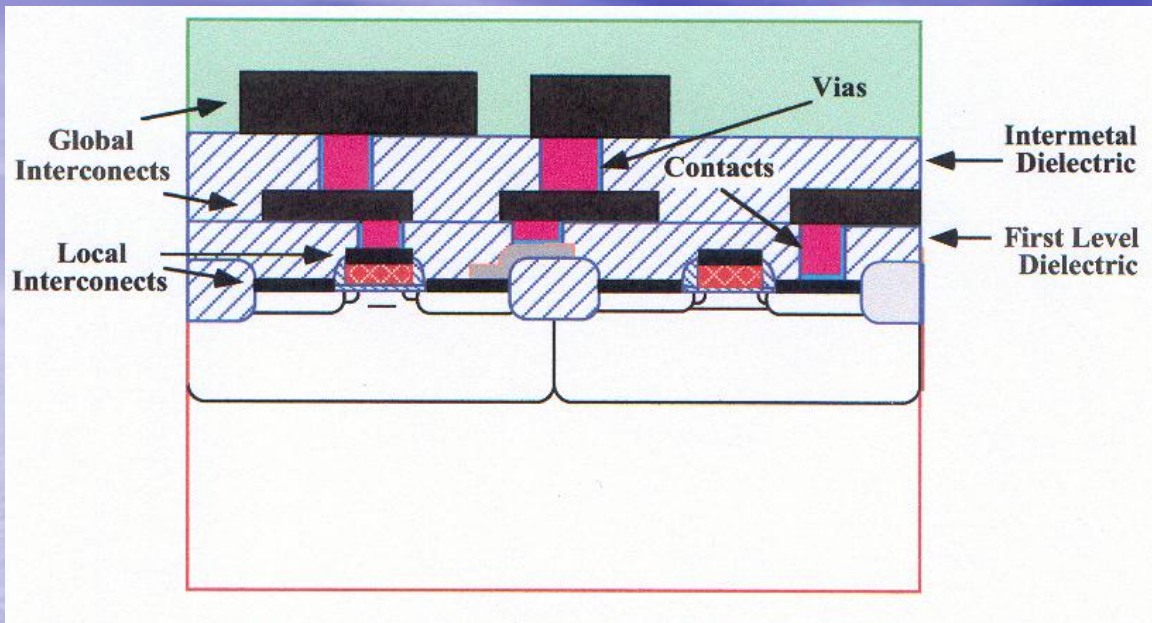
- **Interconexão;**
- **Contatos e vias;**
- **Dielétricos intermetálicos;**
- **Passivação, etc.**



- Estrutura dos primeiros CI's :

bastante simples para padrões atuais.

- **Conexão** : linha de Al conecta regiões ativas – transistores e resistores – em Si;
- **Passivação** : SiO_2 passiva a superfície do Si e isola a interconexão de Al com outras regiões e o substrato;
- **Contato** : Al em contato com o Si.



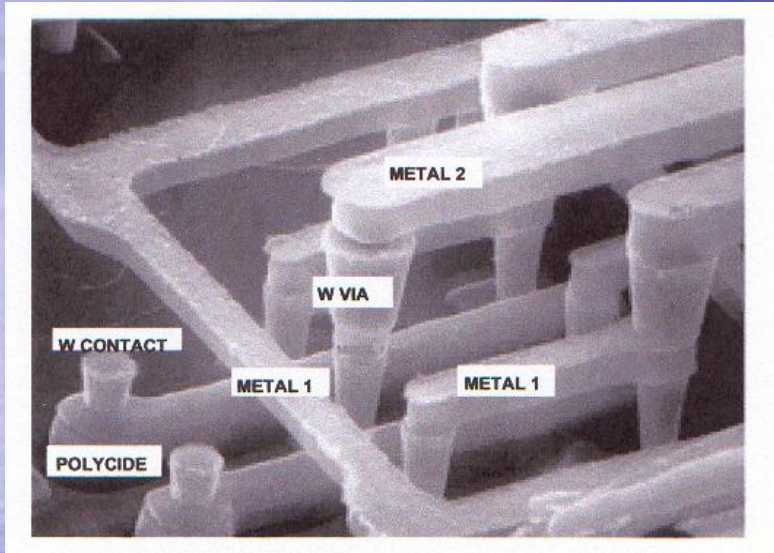
Esquemático de estrutura back end: interconexão, contatos e vias separados por camada dielétrica.

- Mais níveis de interconexão metálica \Rightarrow aumenta funcionalidade e velocidade do circuito.
- **Interconexão local** (nível de interconexão mais baixo – em tecnologia MOS, conexão de Gates, Sources e Drains) : Si-poli, Silicetos, TiN, W.
 - tolera resistividade mais alta que a interconexão global por não ser muito longa. Mas deve suportar alta temperatura de processo, pois são depositados antes da interconexão global.

- **Interconexão global : normalmente Al.**

- Todos níveis de interconexão acima da interconexão local.
- Distâncias longas : conecta diferentes dispositivos e diferentes partes do chip.
- Metais de baixa resistência.

- **Contato ohmico** : conecta a interconexão com os dispositivos.
- **Dielétrico** : camada dielétrica de alta resistência, normalmente SiO_2 , isola as regiões ativas da interconexão local. Usado também para isolamento (IMD) da interconexão local da global.
- **Vias** : conexão de dois níveis de interconexão global.



- SEM de um circuito da IBM, mostrando as interconexões com as camadas dielétricas removidas.

- Processo back end está tornando mais importante no desempenho global do chip.
 - Fração maior da estrutura total e processamento.
 - Está começando a dominar a velocidade total do circuito.

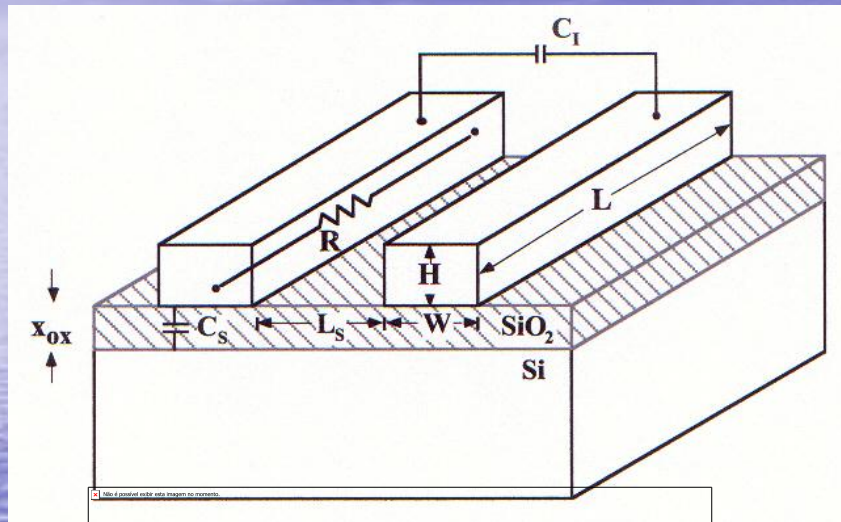
Table 16.1 Future projections for back end technology taken from the SIA NTRS

Year of 1st DRAM Shipment	1997	1999	2003	2006	2009	2012
Minimum Feature Size, F_{\min} (nm)	250	180	130	100	70	50
DRAM Bits/Chip	256M	1G	4G	16G	64G	256G
DRAM Chip Size (mm ²)	280	400	560	790	1120	1580
MPU Chip Size (mm ²)	300	360	430	520	620	750
Wiring Levels - Logic	6	6-7	7	7-8	8-9	9
Min metal CD (nm)	250	180	130	100	70	50
Min contact/via CD nm	280/ 360	200/ 260	140/ 180	110/ 140	80/100	60/70
Metal Aspect Ratio	1.8	1.8	2.1	2.4	2.7	3.0
Contact aspect ratio (DRAM)	5.5	6.3	7.5	9	10.5	12
Via aspect ratio (logic)	2.2	2.2	2.5	2.7	2.9	3.2
Metal resistivity (μ -cm)	3.3	2.2	2.2	2.2	<1.8	<1.8
Interlevel metal dielectric constant	3.0-4.1	2.5-3.0	1.5-2.0	1.5-2.0	<1.5	<1.5

- \uparrow Nível de metais.
- \uparrow razão de aspecto de interconexão, contato e vias.
- \downarrow resistividade dos metais. \Rightarrow não pode ser atingido por Al.
- \downarrow constante dielétrica para dielétricos intermetal. Não pode ser obtido por SiO₂.

• **Solução : interconexão de Cu e dielétricos de "low κ ".**

- A limitação da velocidade de interconexão pode ser estimada de forma bastante simples.



Time delay:

$$\tau_L = 0.89RC$$

Onde: $R = \rho \frac{L}{WH}$ (resistência da linha).

$C = C_s + C_1$ (Capacitância total associada a linha: capacitância da linha – substrato e capacitância de acoplamento entre linhas adjacentes).

- Time delay (rise time) devido a interconexão global :

$$\tau_L = 0.89RC = 0.89K_I K_{OX} \varepsilon_0 \rho L^2 \left(\frac{1}{Hx_{ox}} + \frac{1}{WL_S} \right)$$

Onde : K_{ox} , constante dielétrica do óxido;

x_{ox} , espessura do óxido;

ρ , resistividade da linha de interconexão;

$K_I \approx 2$, fator que considera efeitos de outras interconexões em sistemas de interconexão multicamadas.

- **O que ocorre com as dimensões das estruturas com a evolução da tecnologia?**

- L_s e $W \rightarrow \approx F_{\min}$ e segue o escalamento;
- X_{ox} e H também segue F_{\min} mantendo a relação $H/W =$ constante;

- **Estas espessuras no entanto não tem decrescido tão rapidamente, especialmente para interconexão global.**

- **Assumindo que x_{ox} , H , L_s e W sejam iguais a F_{\min} temos :**

$$\tau_L = 3.56 K_{ox} \epsilon_0 \rho \frac{L^2}{(F_{\min})^2}$$

- Interconexão local : $L \downarrow$ com \downarrow do F_{\min} .

∴ Time delay RC para interconexão local mantém aproximadamente constante

- Interconexão global : o comprimento aumenta, devido ao aumento da área do chip para cada nova geração de tecnologia.

- O comprimento médio de interconexão global mais longa num circuito pode ser aproximado por :

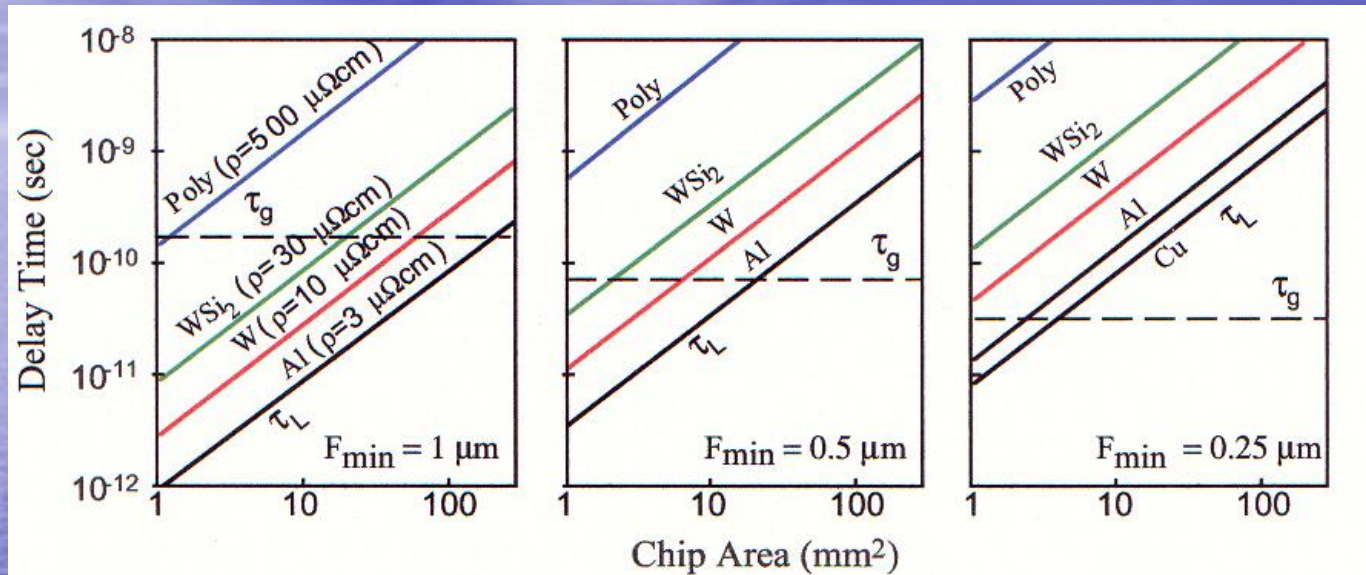
$$L_{\max} = \frac{\sqrt{A}}{2}$$

$$\tau_L = 0.89 K_{ox} \epsilon_0 \rho \frac{A}{(F_{\min})^2}$$

Time delay associado com a interconexão global.

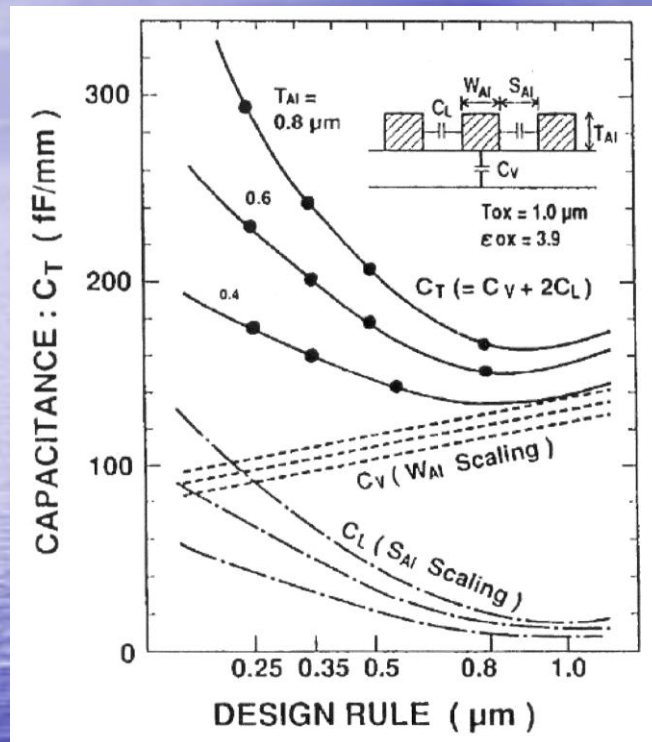
A : área do chip

- Time delay da interconexão e gate versus área do chip.



- Evolução da tecnologia $\Rightarrow F_{\min} \downarrow$ e $A \uparrow$. Ambos aumentam o time delay da interconexão global.
- Mesmo supondo que A permanece constante, time delay \uparrow devido a $\downarrow F_{\min}$. Pois, com a diminuição da área seccional da interconexão, $R \uparrow$, mas o C não \downarrow proporcionalmente, por causa da diminuição da espessura do dielétrico de isolamento das linhas. Portanto, $\Rightarrow RC \uparrow$.

• Capacitância versus design rule



• Gate time delay, τ_g .

- τ_g depende somente de F_{min} e não de $A \Rightarrow$ decresce com \downarrow do F_{min} .
- Para F_{min} grandes, geralmente o τ_g é grande em relação ao delay para interconexão global de Al.

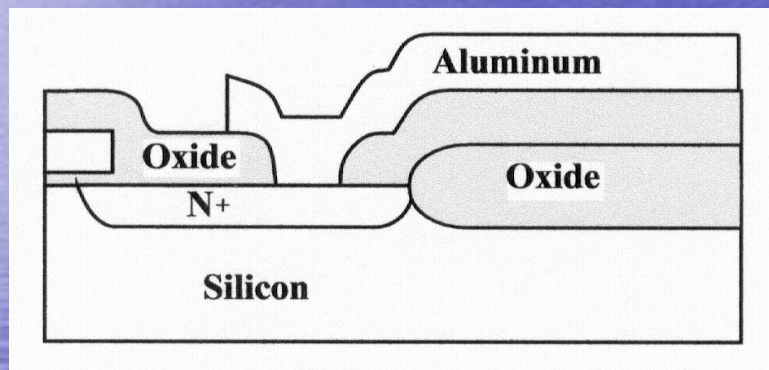
- Mas, para a tecnologia de dimensões mínimas menores e áreas do chip grandes, o delay da interconexão global torna-se bastante grande comparado ao delay do gate e pode ter um grande impacto na performance do circuito.

- Na atual tecnologia CMOS com linhas de Al o delay devido a interconexão pode ser cerca de 30-40% do delay de todo o circuito;
- Tecnologia da geração anterior, 15-20%.

- **Nas tecnologias de gerações próximas, o delay da interconexão pode ser maior do que 50% do delay total do circuito. Isto mostra a importância das estruturas back end.**

16.2. Contato

• Contato Metal - Silício



- Primeiras tecnologias : bastante simples.

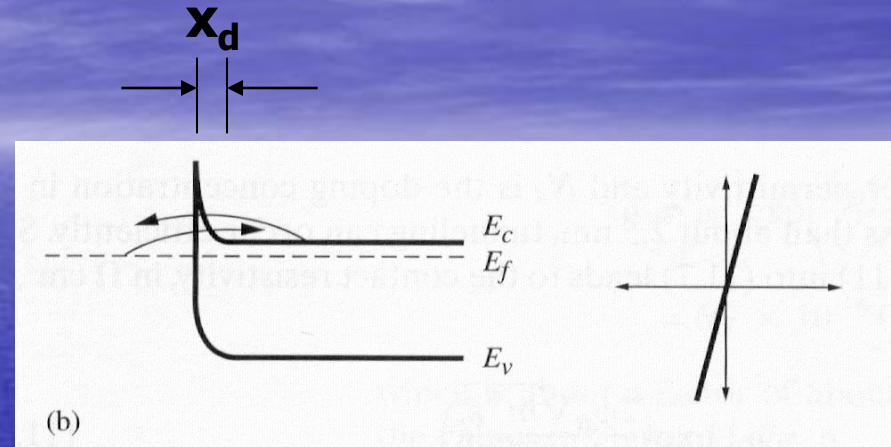
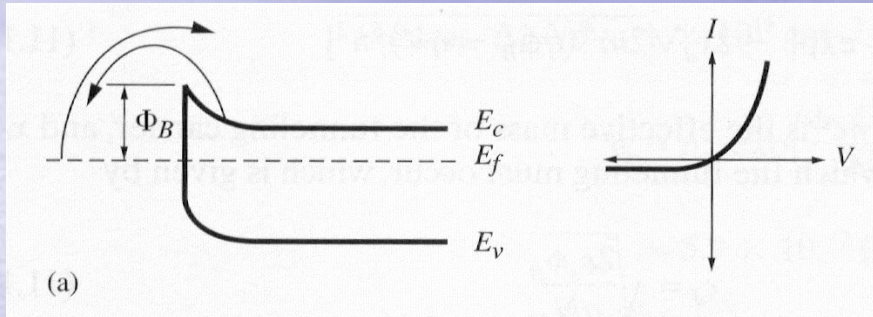
• Contato Al/Si.

- **Provê conexões de baixa resistência entre as interconexões de metal e as regiões ativas.**

- Com a evolução da tecnologia o contato torna-se importante, pois a resistência de contato é inversamente proporcional a área de contato. A área de contato torna-se também menor com a \downarrow do F_{\min} .

• Tipos de Contatos

Contato Schottky



- Devido a diferença na função de trabalho entre o metal e o semiconductor, existe uma barreira ϕ_B , na interface.

Contato ohmico

- Existe uma vasta variedade de metais de diferentes funções de trabalho. Assim deveria ser possível a escolha de um metal que produzisse menor ϕ_B de contato das regiões do dispositivo e o silício.

- Na prática esta escolha não é possível de ser feita de forma generalizada devido a outros requisitos como a resistividade e estados de superfície na junção metal/semicondutor.

- Produz contatos de alto ϕ_B para regiões N (contato Schottky) e contatos de ϕ_B relativamente baixos (contato ôhmico) para silício do tipo p.

- Nota : ϕ_B (tipo n) + ϕ_B (tipo p) = E_g do semicondutor

- No contato Schottky, a energia térmica da emissão termoiônica possibilita os portadores a superar a barreira, permitindo assim o fluxo de alguma corrente sobre a barreira.
- Contudo para maioria dos metais a barreira é bastante alta para Si do tipo N e não há fluxo suficiente de corrente sob as condições normais de operação.

- Outra forma de superar a barreira Schottky é por tunelamento.
- Se o Si for altamente dopado, a largura da barreira (região de depleção, x_d), torna-se suficientemente estreita que muito dos portadores podem atravessar a barreira por tunelamento. Os portadores podem fluir virtualmente em ambas as direções, através do contato. Isto ocorre para dopagens $> 6 \times 10^{19} \text{ cm}^{-3}$.

- Resistividade de contato específico :

$$\rho_c \equiv \left(\frac{\partial J}{\partial V} \right)^{-1}_{\mathbf{V} \approx \mathbf{0}}$$

$$\rho_c = \rho_{co} \exp \left(\frac{2\phi_B \sqrt{m^* \epsilon_s}}{h \sqrt{N_D}} \right)$$

Para contato ohmico

ϕ_B – barreira de Schottky;

N – concentração da dopagem no semicondutor;

ρ_{co} – constante dependente do metal e semicondutor;

ρ_c – resistividade de contato em ohm-cm².

m^* – massa efetiva do portador;

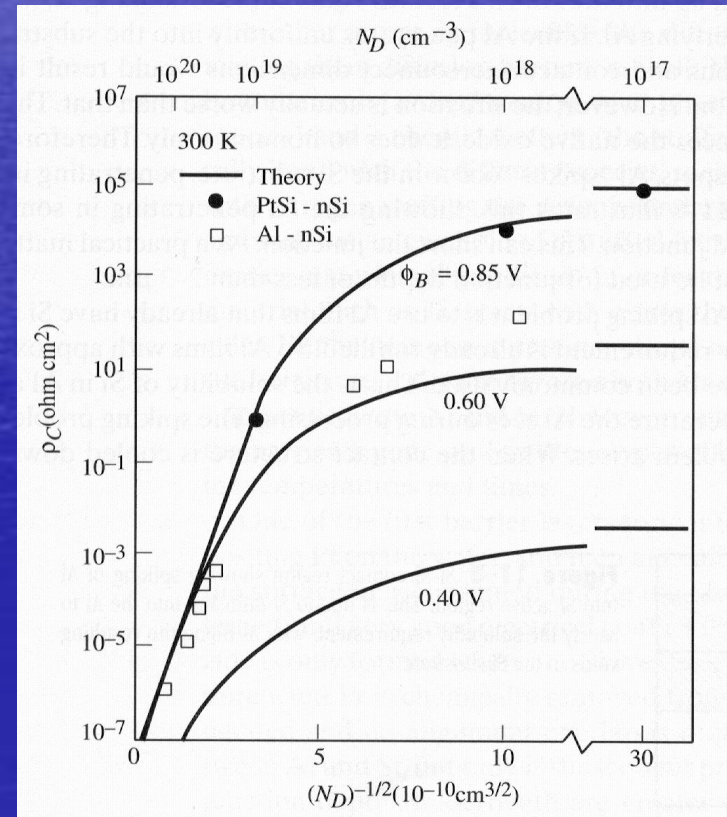
h – constante de Planck

ϵ_s – permissividade do Si.

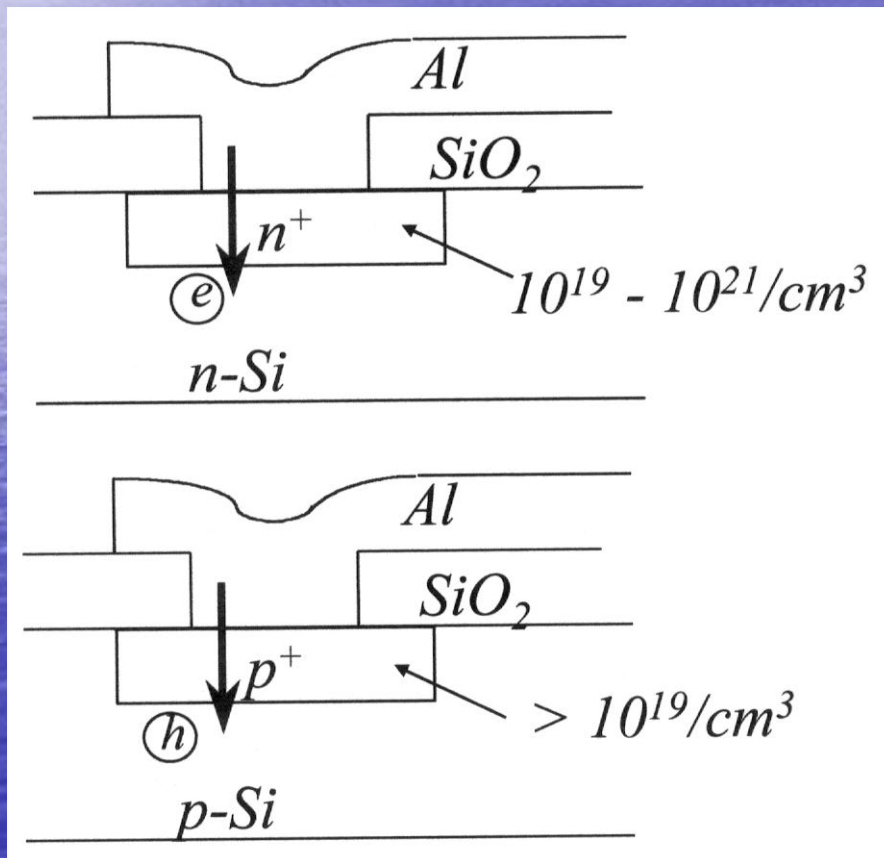
- A resistência de contato decresce com aumento de $(N_D)^{1/2}$, pois $x_d \downarrow$ e torna mais eficiente o tunelamento.
- A corrente de tunelamento através da barreira de Schottky depende da largura da barreira, portanto do N_D .

Resistência de contato versus N_D para tres valores de ϕ_B metal/semicondutor.

- Baixa dopagem : emissão termoiônica controla o processo que é praticamente independente da dopagem.
- Dopagem alta : ρ_c dominado por tunelamento que depende fortemente do nível de dopagem.



- Solução para abaixar a resistência de contato:
 - 1). Usar Si altamente dopado como semiconductor de contato.
 - 2). Escolher metal com baixa altura de barreira Schottky.



- Na prática :
 - $N_D, N_A > 10^{20} \text{ cm}^{-3}$

- **Resistência de contato R_c**

- Para densidade de corrente uniforme sobre a área de contato :

$$R_c = \rho_c / (\text{área de contato})$$

- ρ_c de contato metal – Si : $\sim 10^{-5}$ a 10^{-6} ohm-cm
- ρ_c de contato metal – metal : $< 10^{-7}$ ohm-cm.

- **Requisito para contato metal - Si**

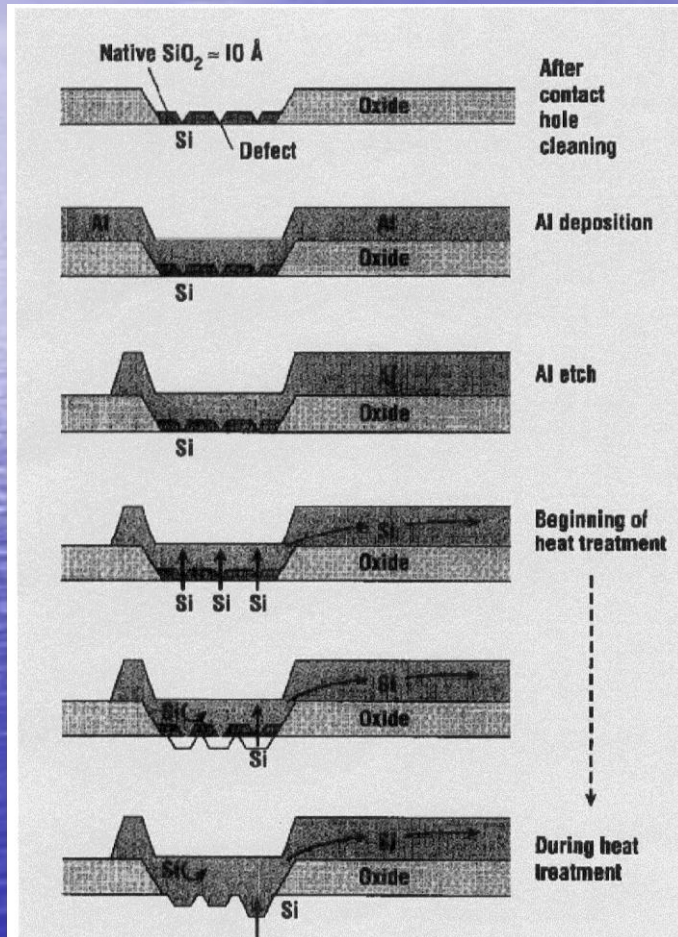
- **Baixa resistência de contato;**
- **Boa estabilidade térmica.**

1. Baixa resistência de contato implica em superfície do Si altamente dopado e bom contato entre o metal e Si.
2. Bom contato requer interface livre de contaminação e resíduos.
3. Boa estabilidade térmica significa não degradação da estrutura de contato durante processos térmicos subsequentes e nem afetar outras áreas como a camada ativa abaixo dos contatos.

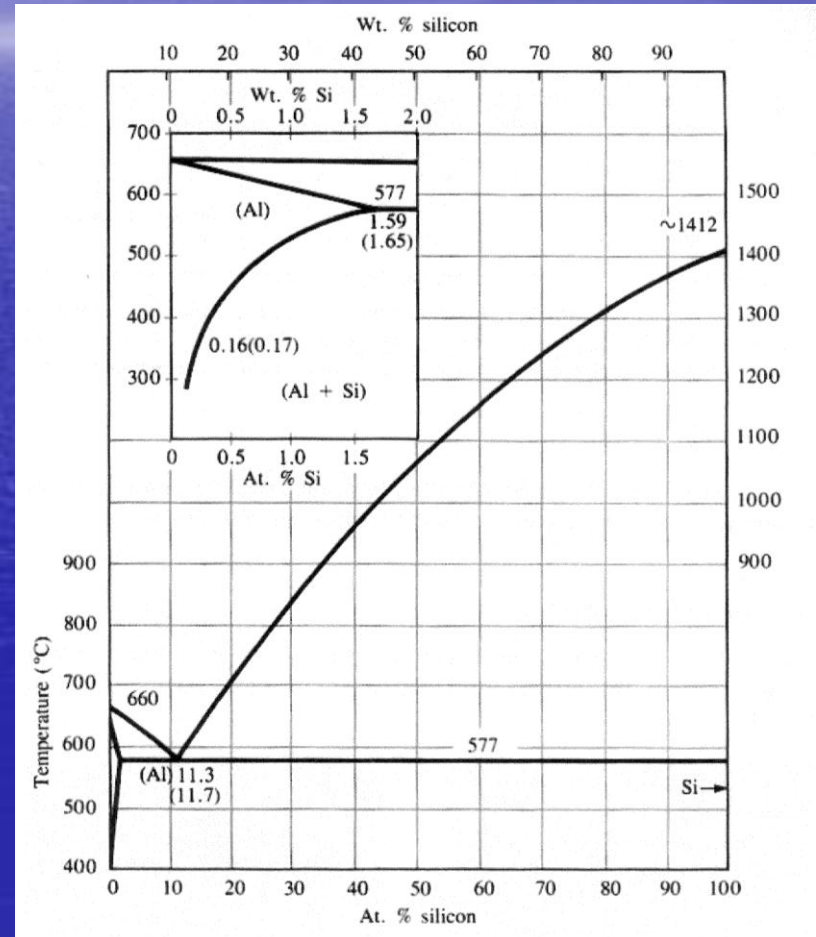
• Problema de “spiking” de Al

- O Al faz um bom contato com o Si pela redução do óxido nativo da superfície, que também remove outras impurezas da superfície.
- Para assegurar um bom contato físico após a metalização do Al é feito o recozimento térmico (sintering) a 450°C.
- Contudo, o Si é solúvel em Al. Aproximadamente 0.5 at% a 450°C e cerca de 1.0 at% a 500°C.
- Como a difusão do Si em Al policristalino é bastante grande, uma quantidade significativa de Si é removida para o Al, deixando um buraco no Si, que é rapidamente preenchido pelo Al. Como a penetração do Al não ocorre de forma uniforme, ocorre “spikes” de Al no substrato de Si ⇒ pode provocar curto!

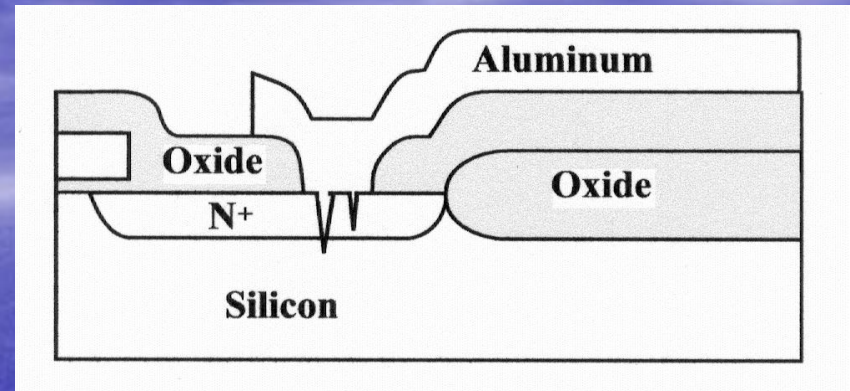
“Spiking” do Al em Si



Solubilidade do Si em Al



- O Al pode penetrar em determinados locais mais profundamente do que a junção difundida, colocando em curto a junção.



• Problema do "spiking" de Al : Soluções.

1. Adicionar 1 – 2 at% de Si em Al.
2. Usar camada de barreira de difusão. Esta é a melhor solução. Ti ou $TiSi_2$ para bom contato e aderência, e TiN para barreira.

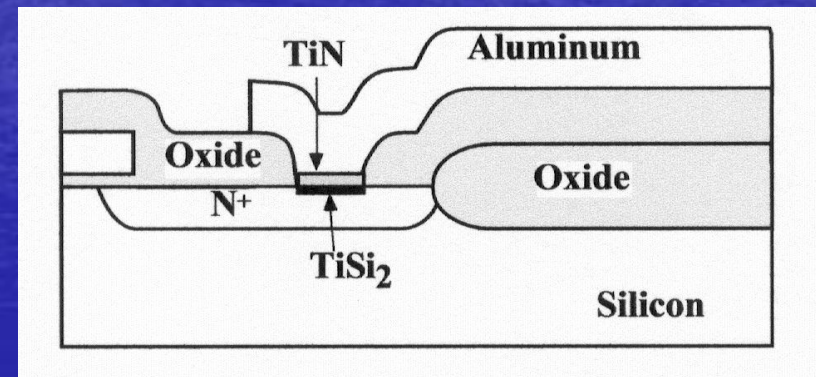
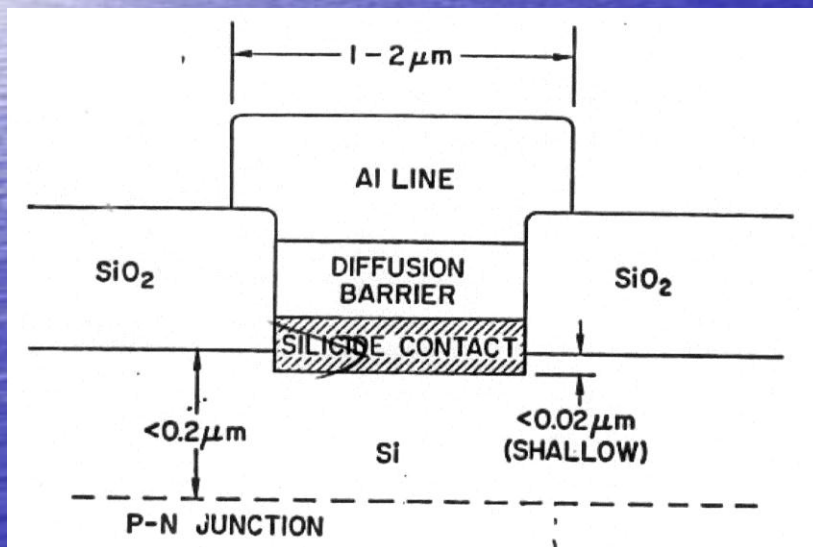
- 1 at% de Si em Al é geralmente utilizado.
- Solubilidade do Si em Al a 500°C, temperatura mais alta no processamento de Al.

- Reduz o problema de "spiking" de Al.
- Mas, quando a estrutura de contato é resfriada abaixo de 500°C pode precipitar o Si.

- Normalmente a precipitação ocorre na interface e pode causar reação de nucleação deixando nódulos de Si.
- há bastante Al dissolvido nestes nódulos o que faz com que eles sejam do tipo p (Al é dopante tipo p em Si), causando aumento da resistência de contato para a região N⁺.

• Uso da barreira de difusão

- Desempenho do circuito torna-se limitado pela resistência de contato, com a redução da área de contato.
- Barreiras de difusão colocadas entre o Al e o Si minimiza a interação entre eles.



**Camada de contato/barreira
TiSi₂/TiN entre o Si e Al**

- Requisitos para a barreira :
 - Barreira para interdifusão química entre Si e Al na temperatura de processamento até 450 - 500°C.
 - Estabilidade térmica.
 - Baixo stress (coef. Expansão térmica próximo do Si, $2.6 \times 10^{-6} \text{C}^{-1}$).
 - Boa aderência com Al, Si e SiO₂. \Rightarrow há reação interfacial entre a barreira e Si e Al.
 - Boa condutividade elétrica e baixa resistência de contato a ambos Si e Al.

- Tipos de barreiras :

1). **Barreira passiva.** Deve ser quimicamente inerte a Si e Al e ser uma boa barreira de difusão. Exemplo : TiN. Alguns materiais são quimicamente inertes mas a difusão entre os contornos dos grãos é significativa.

2). **Barreira "stuffed".** A difusão pode ser reduzida fazendo-se "stuffing" dos contornos de grãos com outras espécies (N ou O, por exemplo). Ti-W stuffed em N₂ é um exemplo deste tipo de barreira (obtido por sputtering de Ti-W em ambiente de N₂).

3). **Barreira sacrificial.** Material de barreira é sacrificado para prevenir reação entre o Al e Si. Ti é um exemplo. O Ti reage com Si e Al formando siliceto e aluminide. Uma vez consumido o material da barreira, a barreira desaparece. Esse tipo de barreira somente é efetivo para limitada temperatura de processamento.

- A tabela abaixo relaciona as camadas de barreira mais comuns usadas para contatos. Permite comparar as suas temperaturas de processamento e modos de falhas.

Table 16.2 Barrier Layer Properties

Structure	Failure Temperature (°C)	Failure Mechanism (reaction products)
Al/PtSi/Si	350-400	Compound formation (Al_2Pt , Si)
Al/TiSi ₂ /Si	400	Diffusion ($\text{Al}_5\text{Ti}_7\text{Si}_2$, Si at 550°C)
Al/NiSi/Si	400	Compound formation (Al_3Ni , Si)
Al/CoSi ₂ /Si	400	Compound formation (Al_9Co_2 , Si)
Al/Ti/PtSi/Si	450	Compound formation (Al_3Ti)
Al/Ti ₃₀ W ₇₀ /PtSi/Si	500	Diffusion (Al_2Pt , Al_{12}W at 500°C)
Al/TiN/TiSi ₂ /Si	550	Compound formation (AlN, Al_3Ti)