



C. Musilli Cassino 30/6/2006

# SDRAM

## Synchronous Dynamic Random Access Memory

- **Accesso Casuale**
  - Tempo necessario per accedere ad un qualunque dato e' indipendente dalla posizione che il dato stesso assume all'interno della memoria
- **Volatile**
  - Togliendo l'alimentazione la memoria perde tutte le informazioni in essa contenute
- **Dinamica**
  - La cella di memoria e' costituita da un 'contenitore' che ha bisogno di essere periodicamente 'rinfrescato'
- **Sincrona**
  - Il 'clock' della memoria e' sincronizzato a quello della CPU del sistema

# SDRAM

## 'Ingredienti' di base

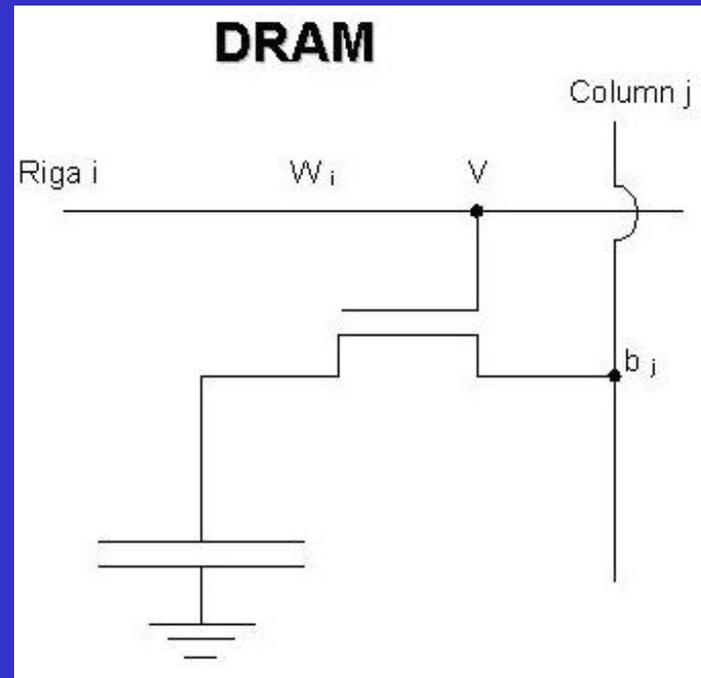
- Gli elementi fondamentali di una memoria DRAM:

Cella di memoria

1. Transistor nmos
2. Capacitore

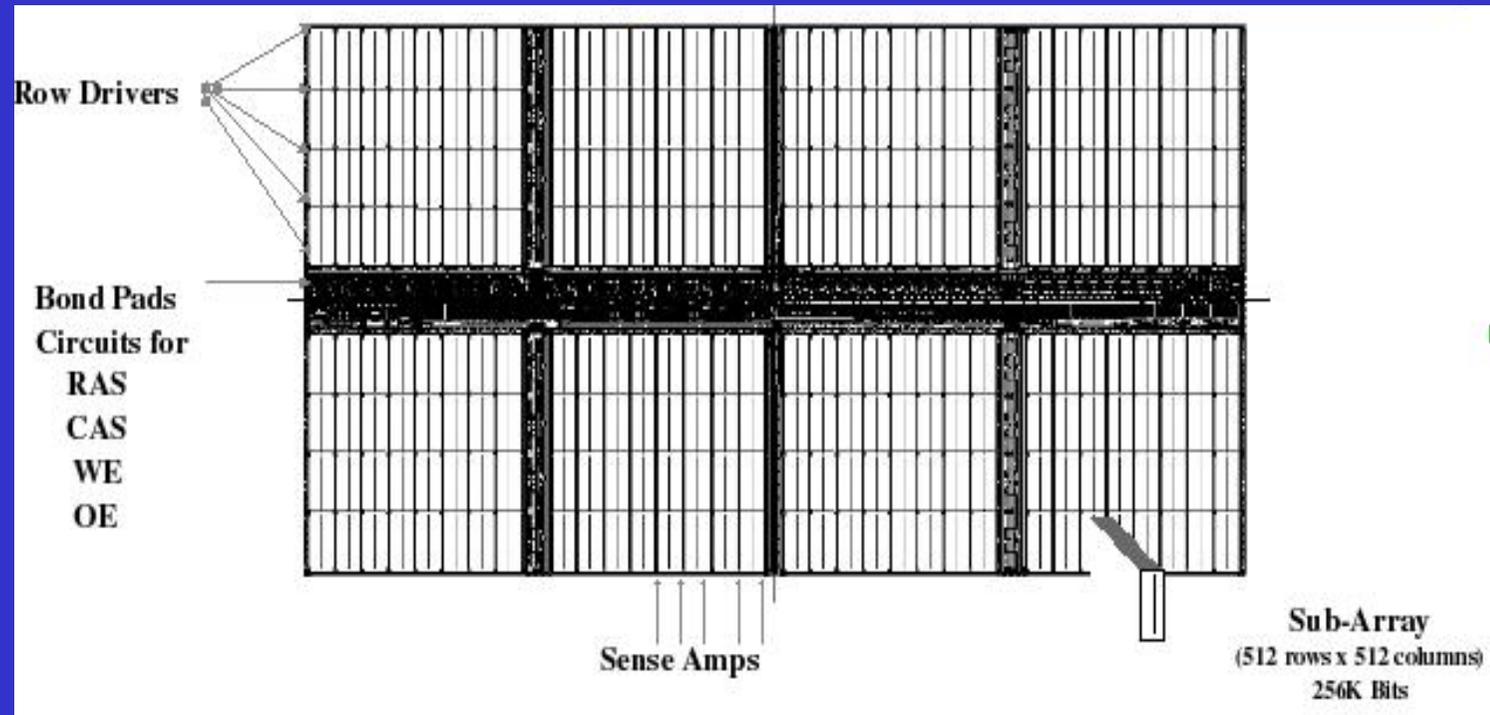
Periferia

1. Transistor nmos/pmos
2. Resistori
3. Capacitori



# SDRAM

## Come e' fatta una 64 MB SDRAM



# SDRAM

- Nel processo di costruzione della memoria dobbiamo:
  - realizzare in diverse zone del wafer un grandissimo numero di transistor, resistori e capacitori (il chip visto prima)
  - replicare tale struttura su tutta la superficie del wafer
  - massimizzare il numero di chip/wafer
  - progettare ed operare in modo da rendere minime le differenze strutturali (geometrie, caratteristiche chimico-fisiche dei materiali)

# Transistor

- Ci concentreremo sulla parte del processo che è deputata alla realizzazione dei transistor (e dei resistori e dei condensatori di periferia)
- Panoramica dei processi che portano alla realizzazione di una memoria DRAM
- Caratterizzazione dei MOS attraverso le misure parametriche

# Cenni sul flusso di processo

- Commento sui processi fotolitografici e di impiantazione che concorrono alla differenziazione delle zone del wafer usate per le diverse funzioni ed alla costruzione dei singoli dispositivi
- Esaminiamo la parte del flusso in cui si realizzano i transistor di periferia e di array
- Brevi cenni sulla costruzione del condensatore che ha il compito di immagazzinare la informazione
- Il processo da cui siamo partiti e` quello relativo ad una memoria DRAM

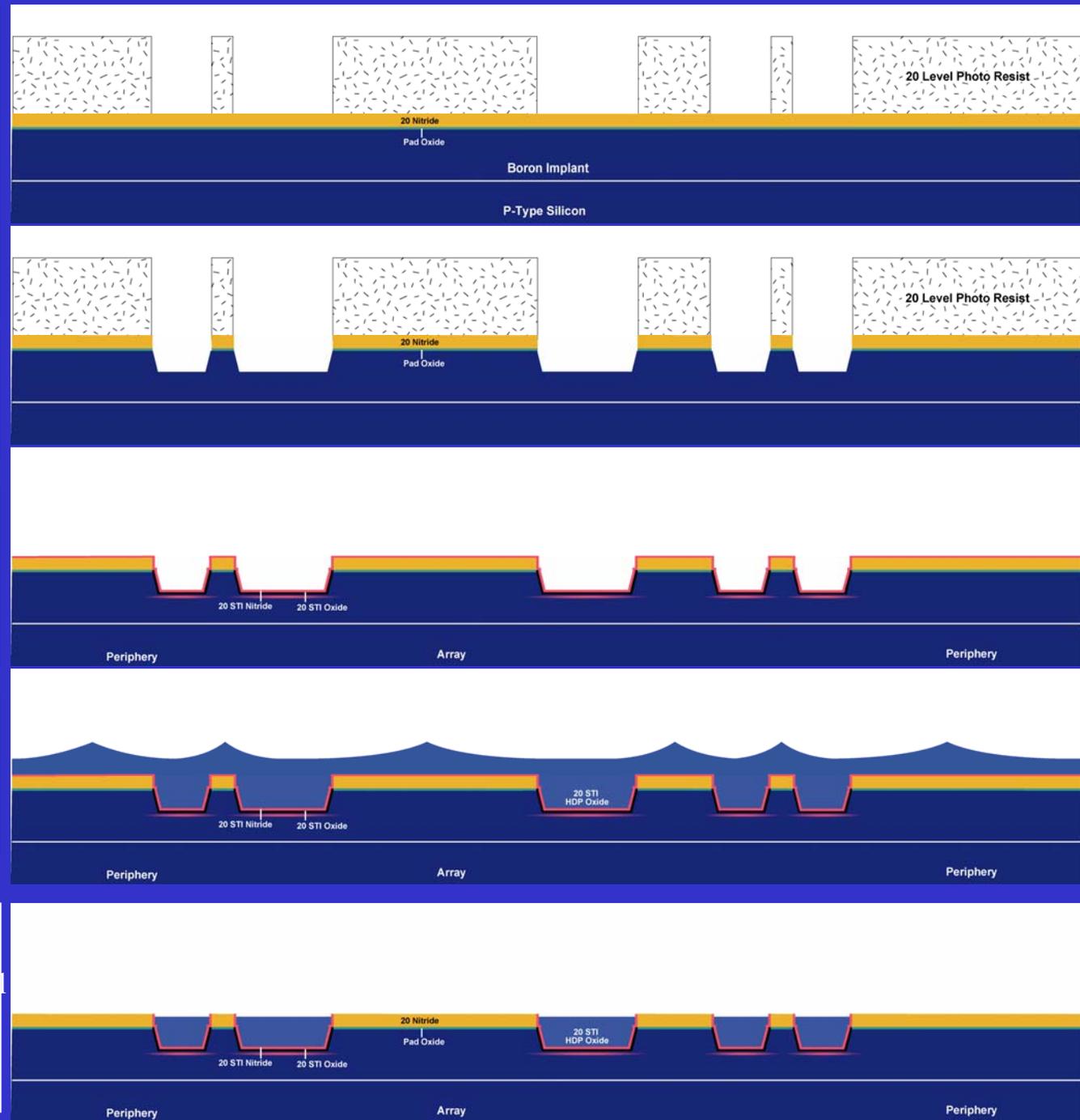
Wafer di Silicio p-type.  
Screen-oxide per impiantazione.  
Impiantazione di tipo p.  
Definizione delle aree attive:  
deposizione di photoresist, sua  
esposizione tramite maschera, e  
sviluppo

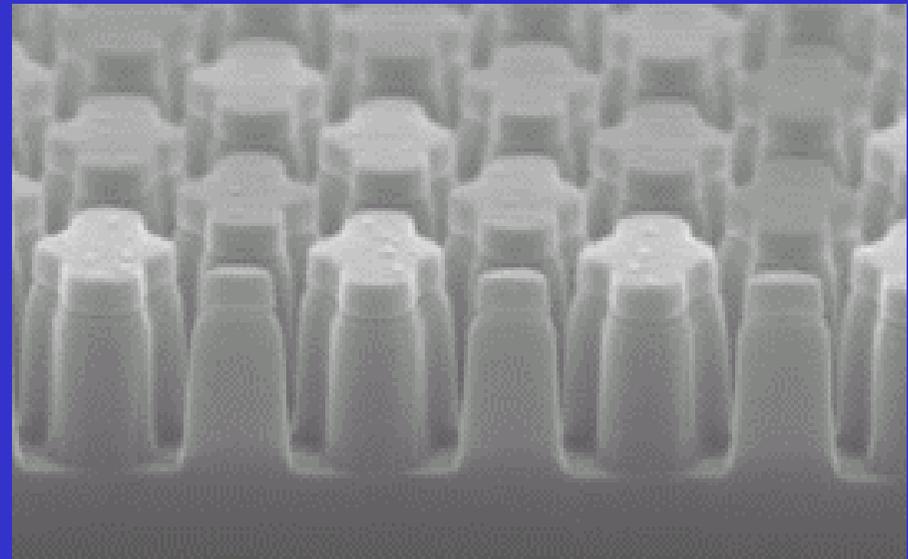
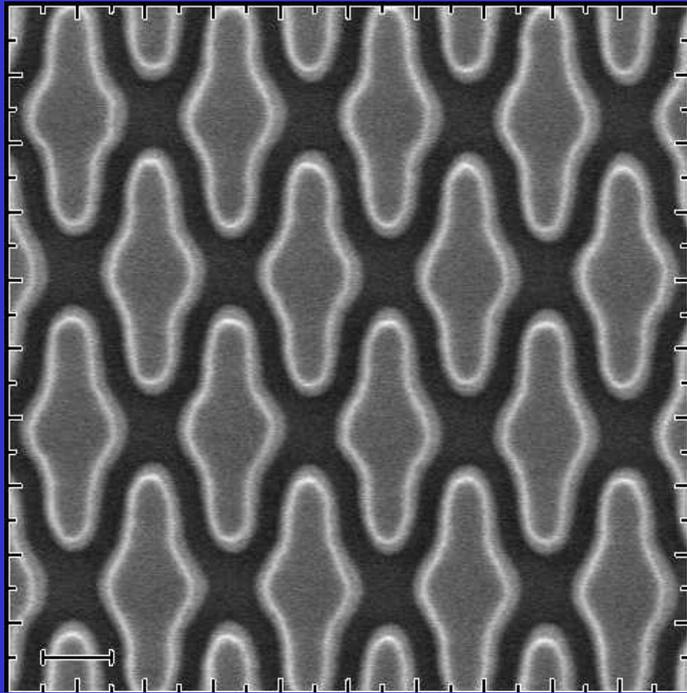
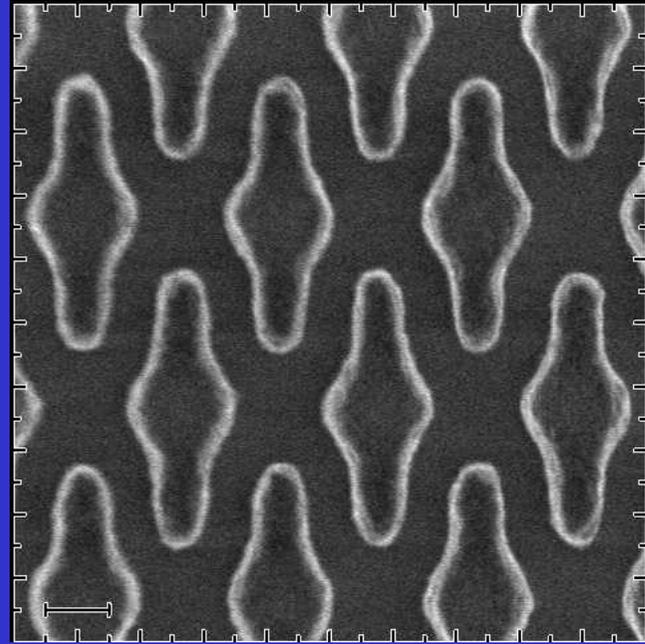
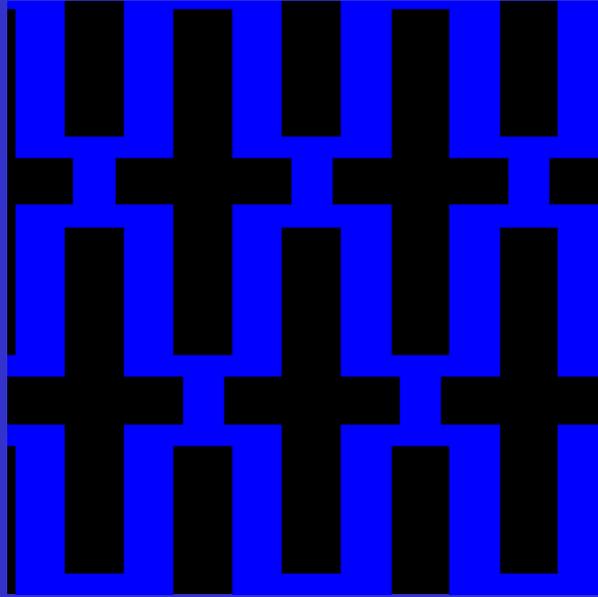
Scavo (etch) del silicio.  
Saranno i 'trench' di isolamento

Preparazione dei trench allo  
scopo di ospitare il materiale  
che fungerà da isolante  
Accrescimento di un sottile  
strato di ossido di silicio

Deposizione di ossido.  
Deve riempire i trench

Dopo la deposizione si procede  
con un processo di rimozione del  
materiale in eccesso e di  
planarizzazione (CMP)

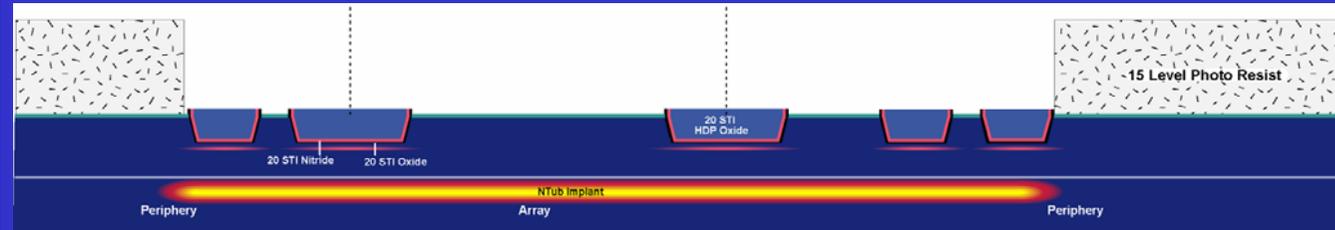




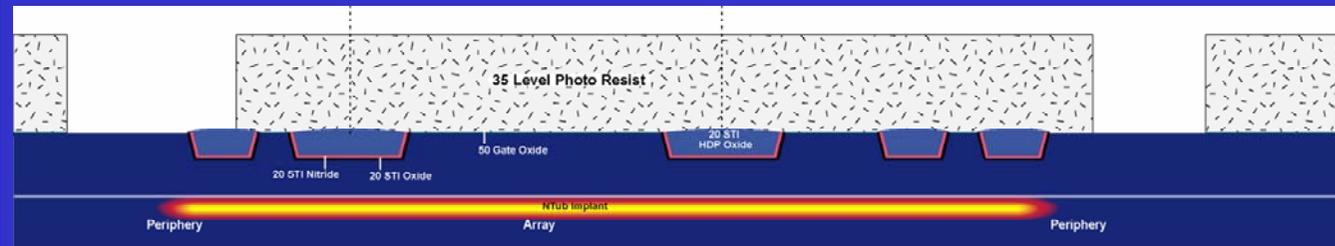
Rimozione del Nitruro.  
Le tre zone sono uguali dal punto di vista del livello di drogaggio



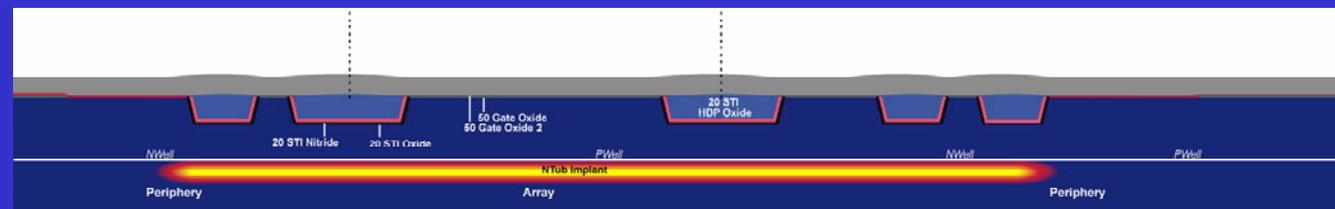
Maschera per effettuare una impiantazione di tipo n in una zona specifica del wafer.  
Drogaggio diverso per diverse funzioni



Accrescimento di un strato di ossido sacrificiale.  
Rimozione e accrescimento dell'ossido di gate.  
In alcuni dispositivi può essere necessario accrescere ossidi di gate di diverso spessore per diverse funzioni



Deposizione del gate dei transistors: film di polisilicio (drogato o non drogato).  
Normalmente si tratta di polisilicio amorfo drogato N

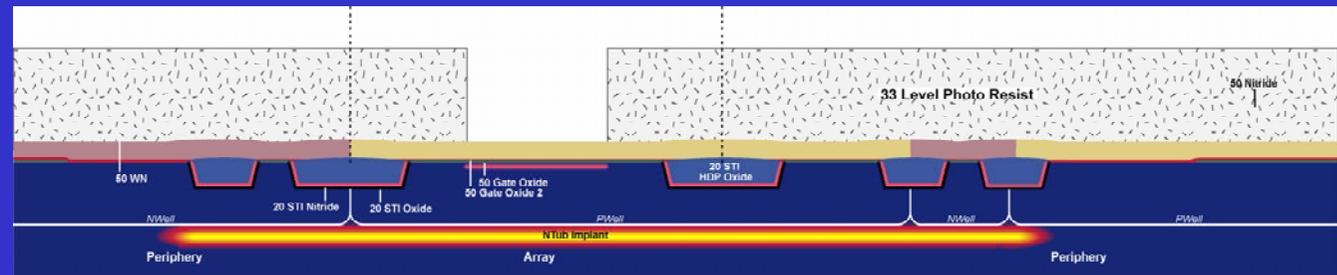
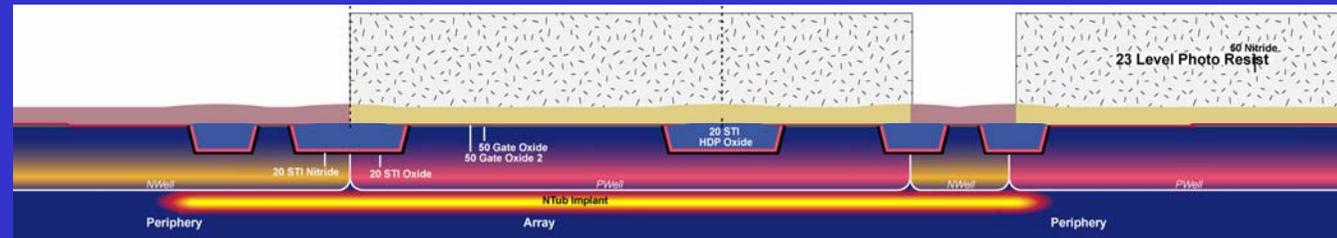
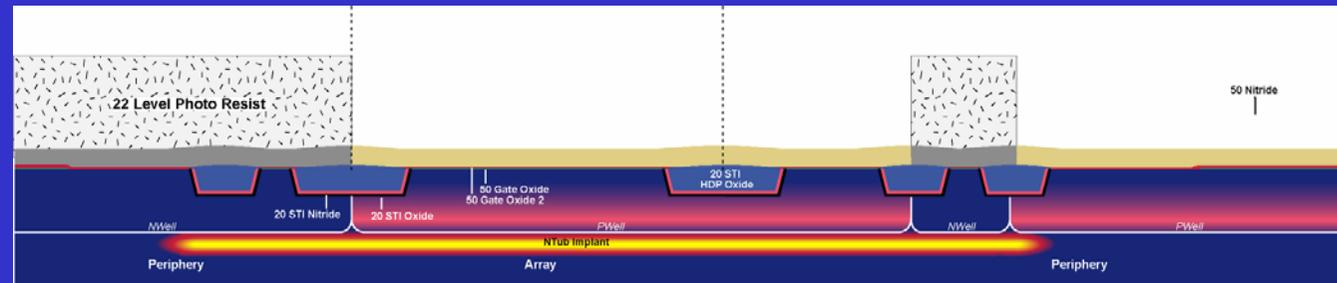


Maschera che definisce le caratteristiche dei dispositivi a canale N.

Impiantazioni per definire  $V_t$  dei transistor e drogaggio dei gate di polisilicio

Ulteriore maschera che definisce le caratteristiche dei dispositivi a canale P. Oltre alle impiantazioni per  $V_t$  e drogaggio del polisilicio, in questo caso ci sono impiantazioni che invertono il drogaggio (da tipo p a tipo n)

Maschera che permette la ulteriore definizione delle caratteristiche dei MOS di array



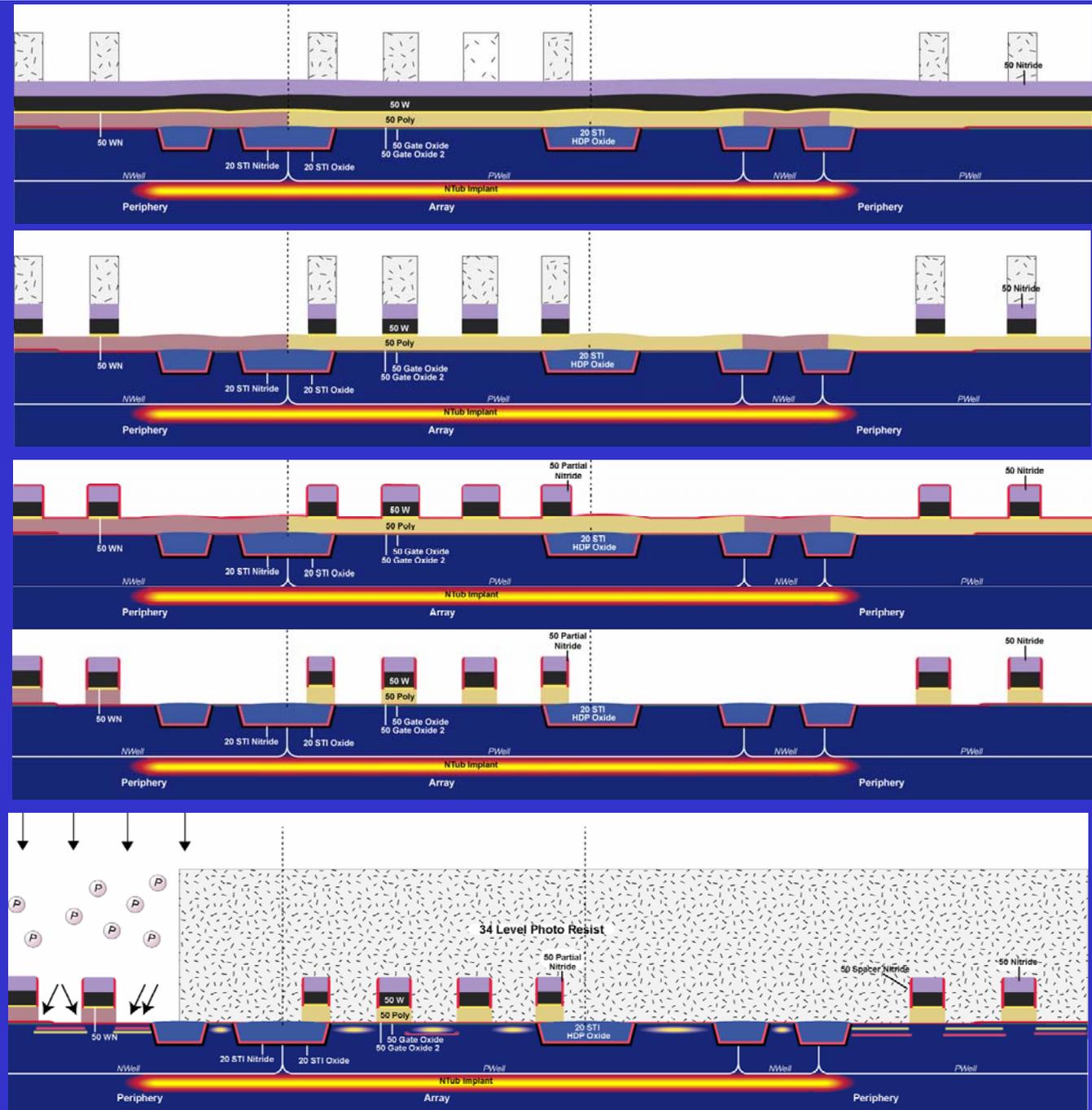
Maschera per la definizione dei gate, dopo la deposizione dei materiali che concorrono alla formazione del gate (oltre al polisilicio)

Si procede alla rimozione dei film non 'protetti' dal photoresist (etch)

Il processo di etch a questo livello è particolarmente delicato poiché si definisce la dimensione dei transistor.

La rimozione procede con vari passi fino al raggiungimento della superficie del silicio

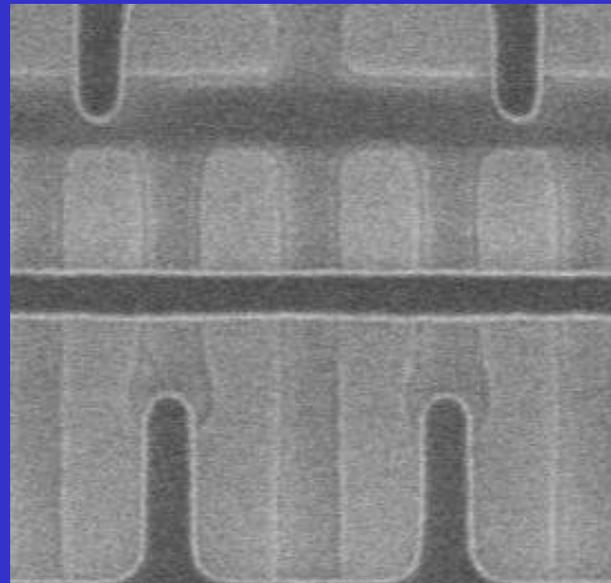
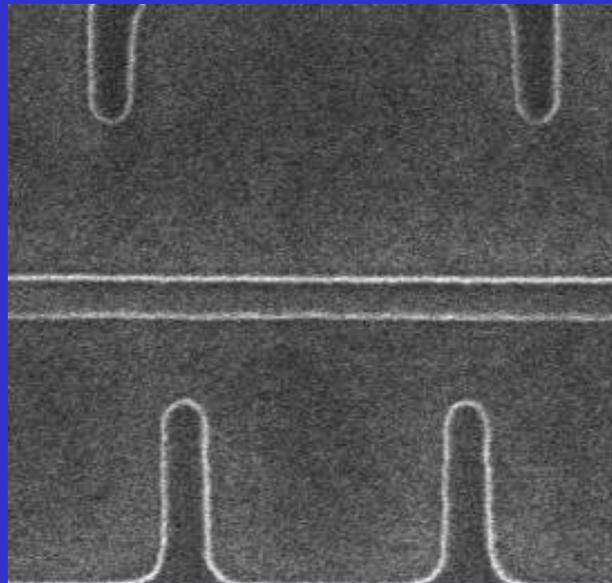
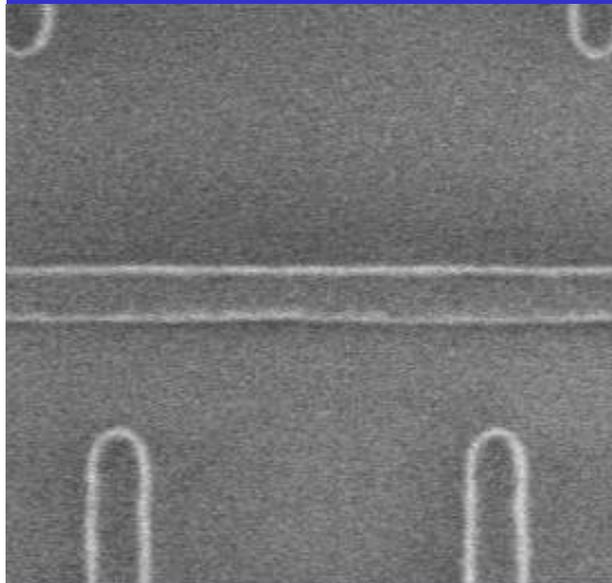
Impiantazioni atte a ottimizzare il funzionamento dei MOS (LDD)



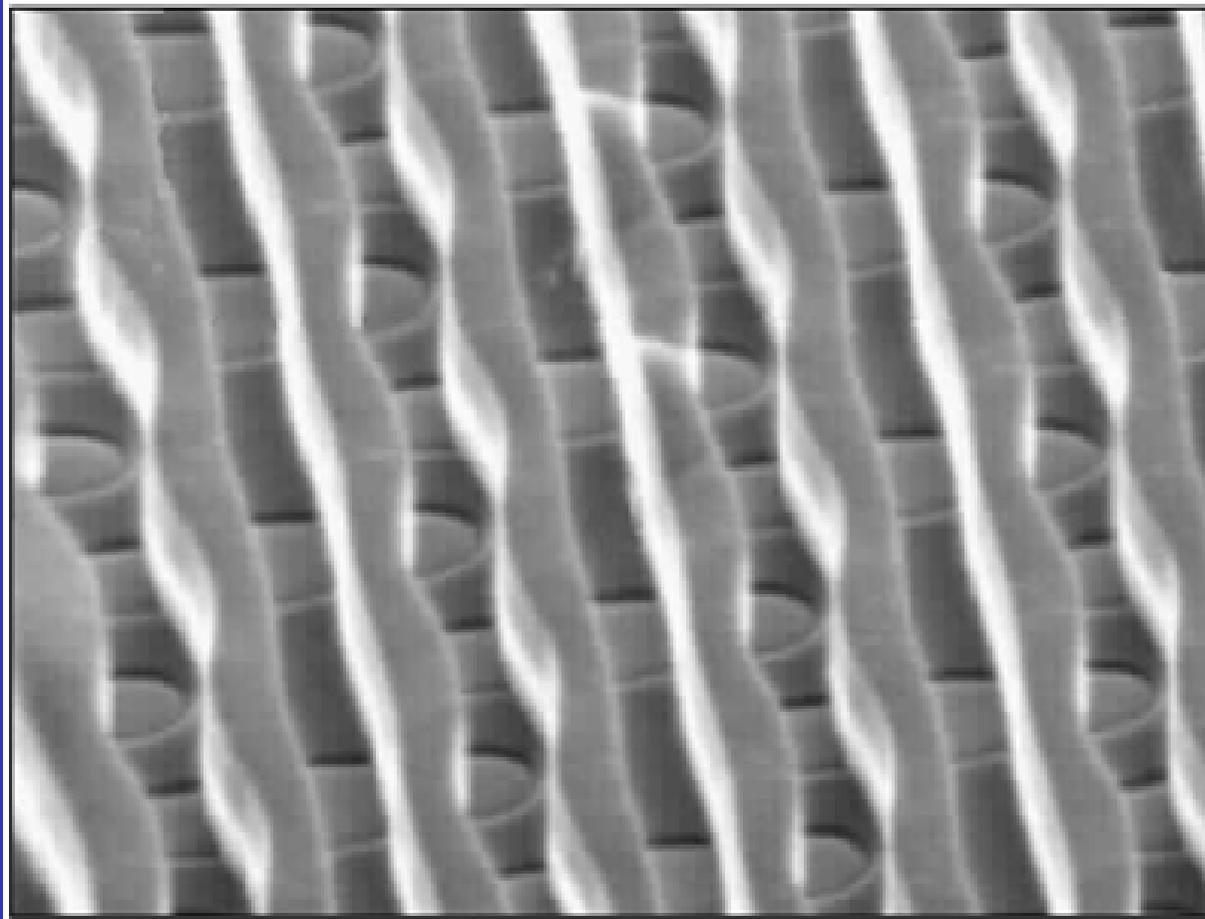
ADI

Partial ACI

final ACI



## Le Wordline sull'Area Attiva

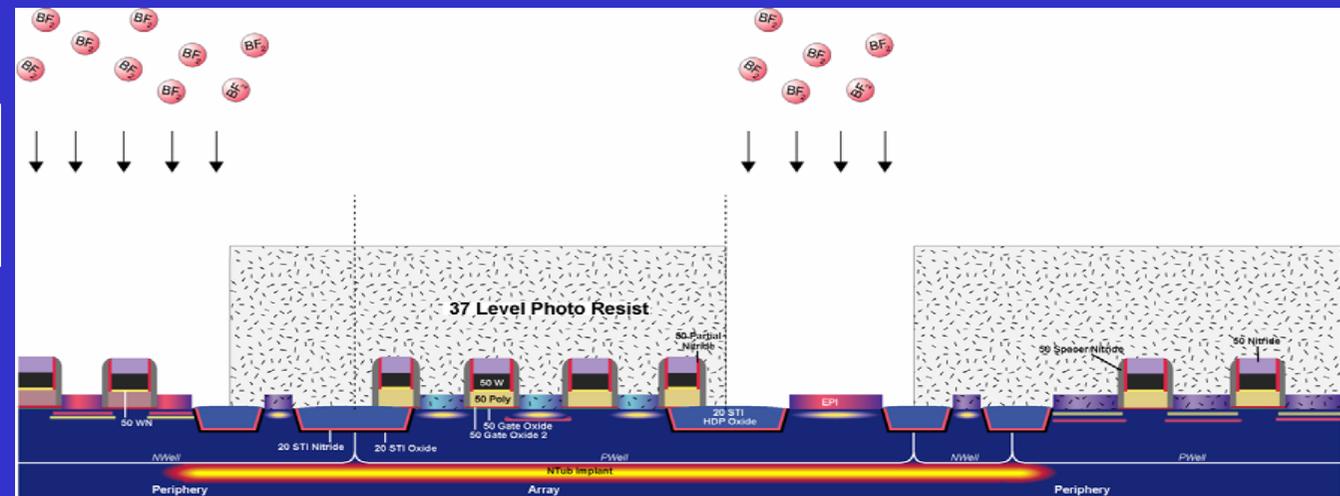
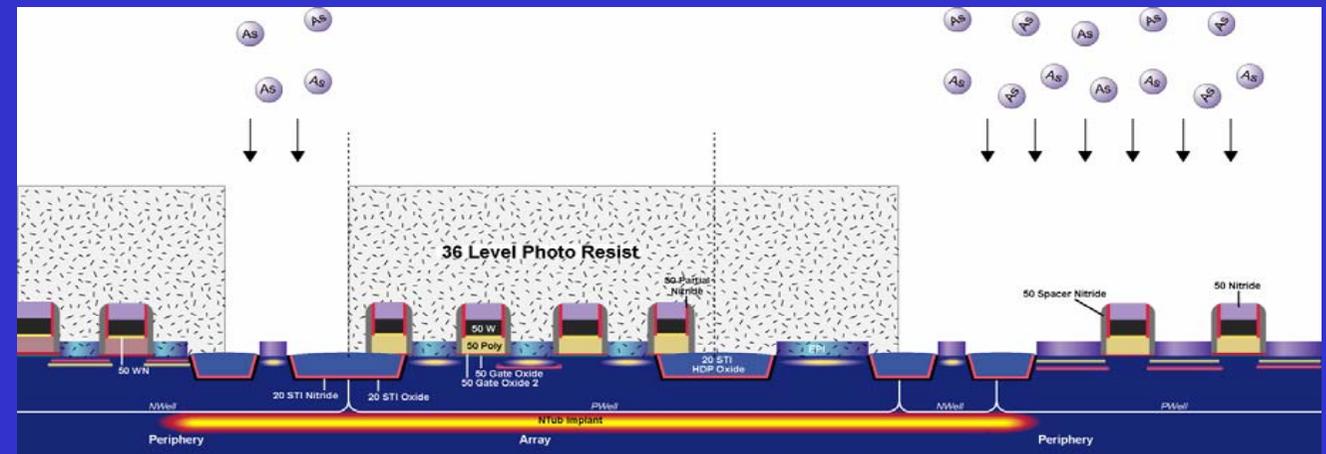
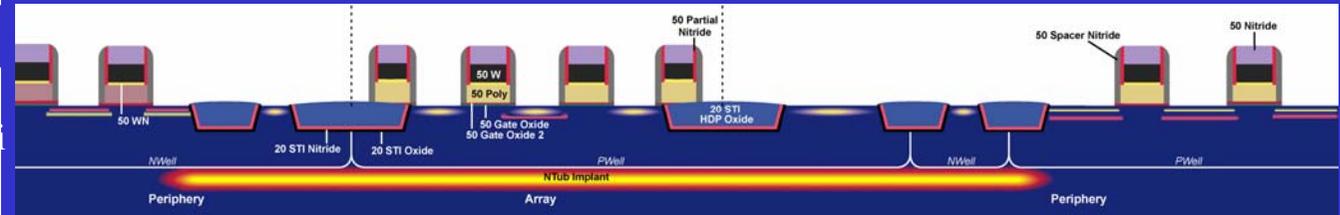
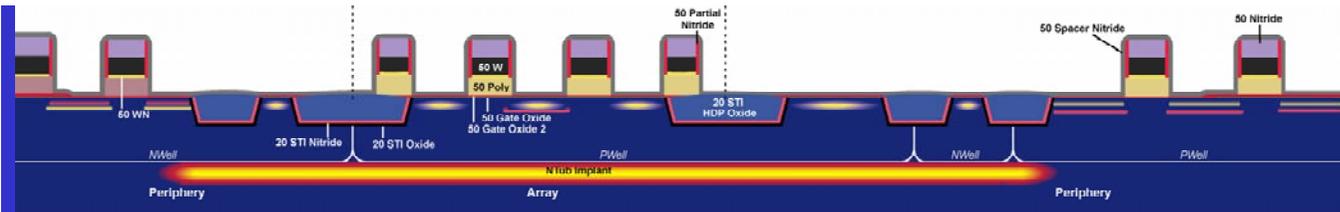


Deposizione dello 'spacer'. E' nitruro di silicio  
Ottimizza le prestazioni dei MOS

Impiantazione su tutto il wafer (blanket) per ottimizzare i profili di drogaggio dei 'source' e 'drain' dei transistor

Con due maschere successive si procede alle impiantazioni che definiscono il livello di Drogaggio dei source e drain dei MOS a canale N ...

.. e dei MOS a canale P



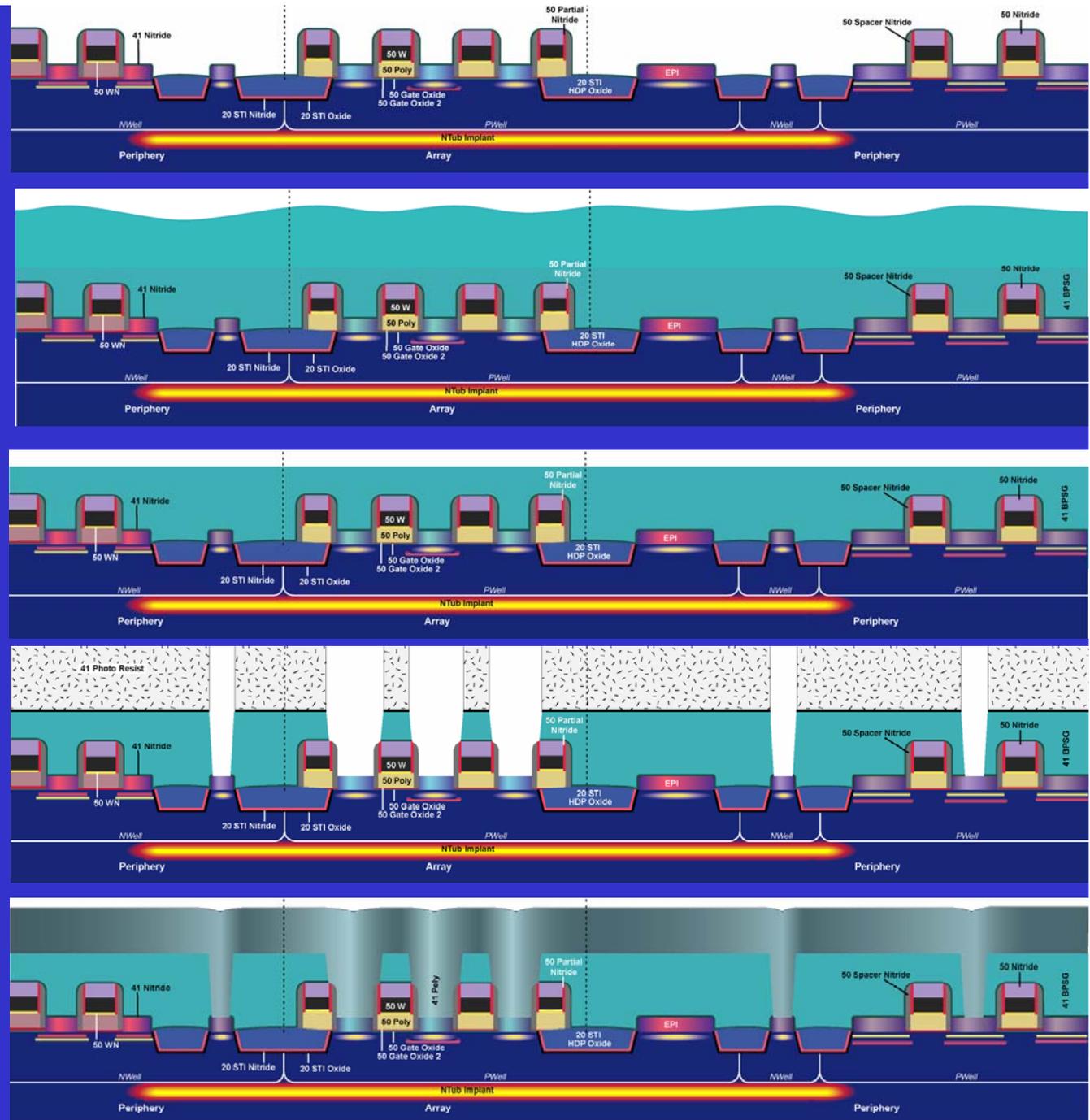
I transistor sono definiti in tutte le zone del wafer.  
Il processo continua per realizzare le altre parti del chip

Si depositano i primi strati di materiale isolante che serviranno come 'supporto' per le unità di stoccaggio della carica e come parte integrante del sistema di contatti

Deposizioni di materiale isolante  
E processi di planarizzazione si susseguono

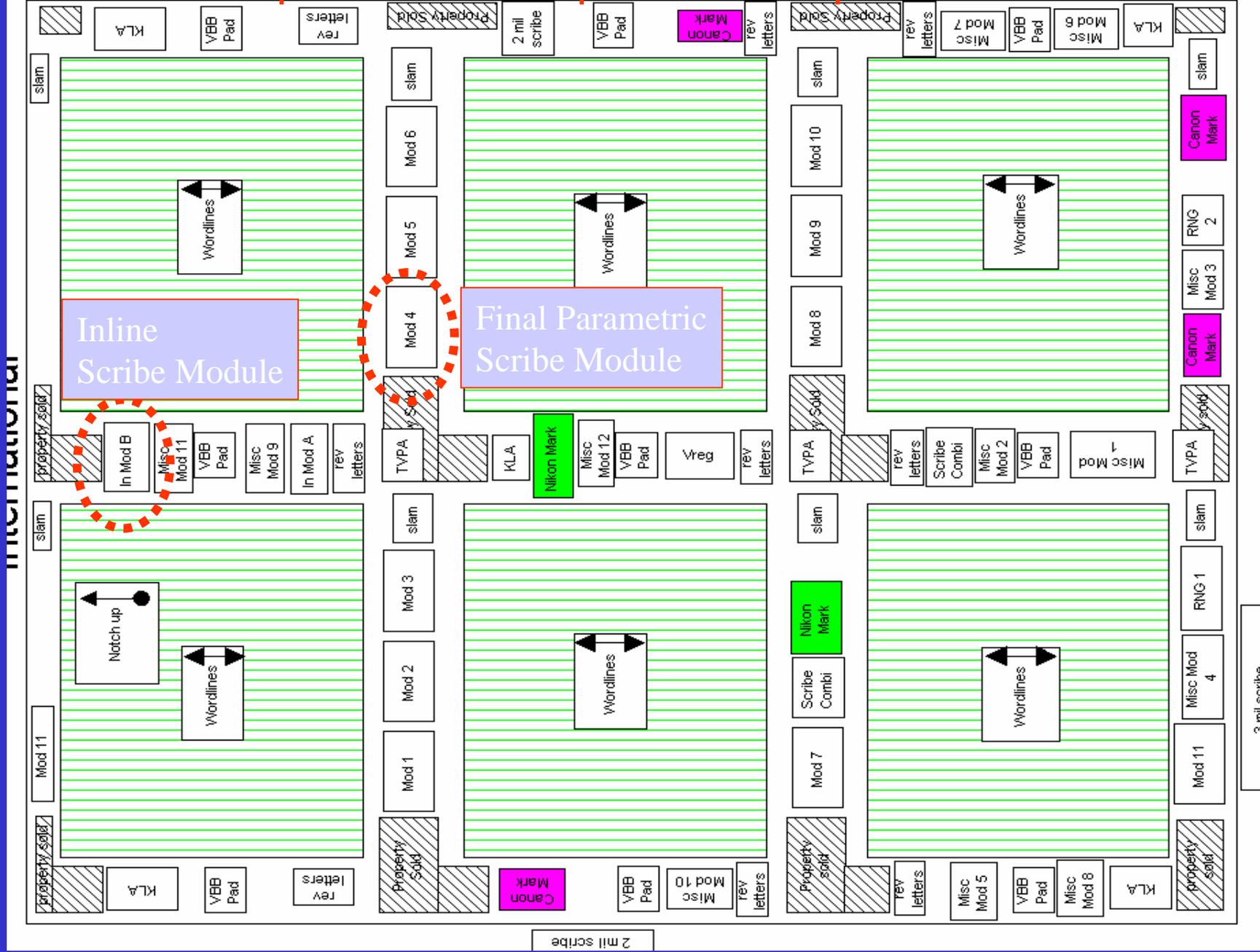
Fino all'apertura dei contatti tramite processi di etch per raggiungere tutti gli elementi dei transistor

I 'buchi' vengono riempiti con materiale conduttivo, che verrà connesso ai livelli di metallizzazione superiore



# **Test Parametrici sui Transistor**

# Esempio di 'shot map'- Strutture parametriche



# Test Parametrico

Varia profondamente in funzione del livello di maturazione raggiunto dal dispositivo in produzione

## Scopi:

- Misurare la 'distanza' tra valori di design e parametri di processo
- Supportare il processo di maturazione/stabilità del dispositivo in produzione
- Contribuire alla soluzione di problemi di resa legati alla funzionalità dei dispositivi
- Costruzione di carte di controllo con lo scopo di segnalare deviazioni dai target predefiniti

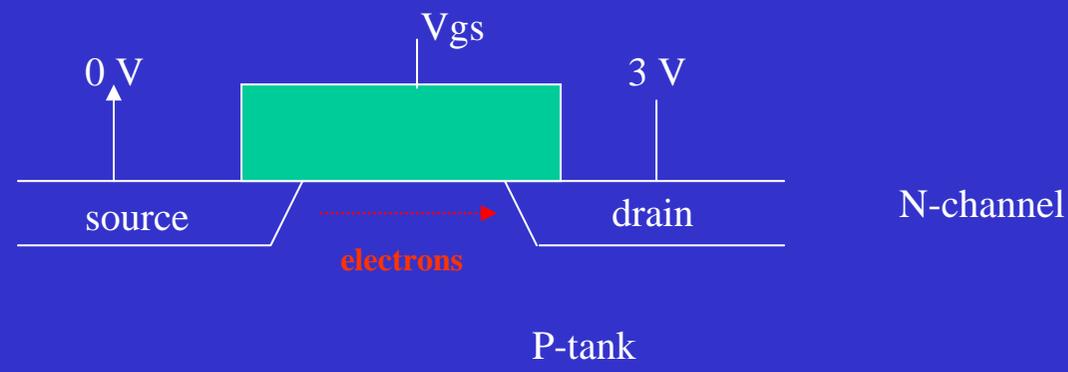
# Transistors

Il transistor può essere considerato un interruttore. Le caratterizzazioni che si effettuano nei test parametrici servono ad assicurare che il transistor:

- Si accenda quando è necessario
- Permetta il passaggio della corrente necessaria al funzionamento del dispositivo
- Impedisca il passaggio della corrente, nello stato OFF

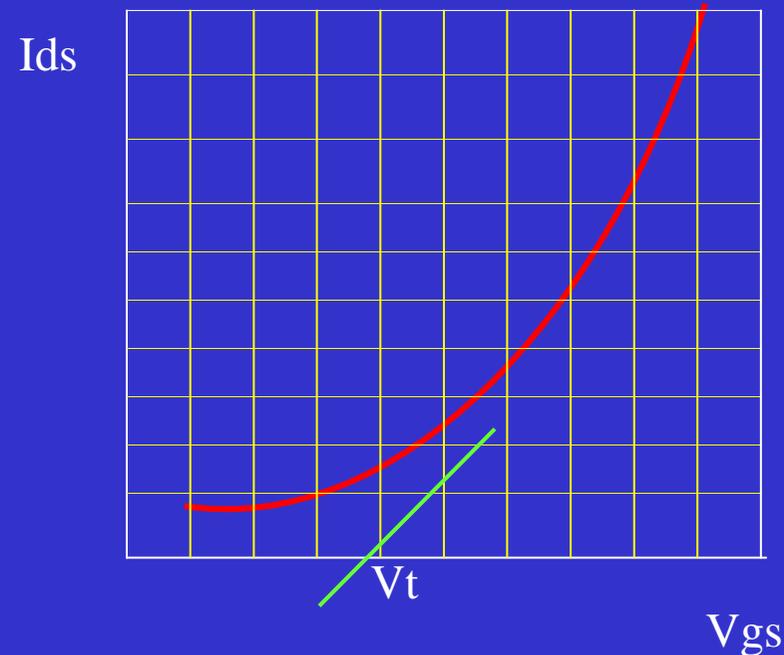
Il cuore del transistor è il gate, dove avviene il processo di ON/OFF. Abbiamo poi due regioni fortemente drogate (source e drain) ai due lati del gate. Il dispositivo è simmetrico. Come identificare source e drain?

La risposta è che ognuna delle due parti può essere source o drain. È il suo utilizzo nel circuito che provoca la differenziazione: il lato che fornisce i portatori è il source



Semplicemente applicando una differenza di potenziale tra source e drain non si ottiene il passaggio di corrente. Come mostrato nel precedente diagramma dobbiamo applicare una tensione al gate ( $V_{gs}$ ). Con un valore appropriato, detto soglia (threshold voltage /  $V_t$ ), il transistor va nello stato ON. Con valori inferiori il transistor rimane nello stato OFF.

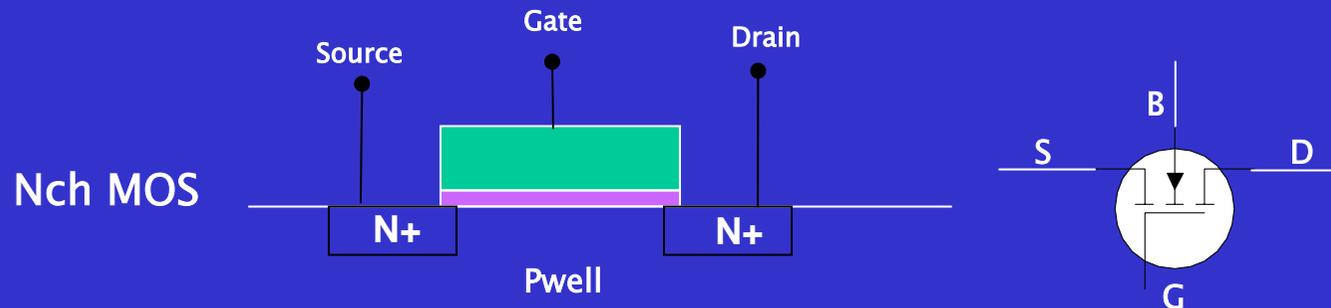
In altre parole, con una differenza di potenziale applicata tra source e drain, non ci sarà passaggio di corrente fino a che non si applichi al gate la tensione di soglia  $V_t$



# Transistor

I test che vengono tipicamente effettuati per determinare le caratteristiche dei transistor utilizzati nei nostri dispositivi sono:

- Tensione di Soglia ( $V_t$ ) e guadagno ( $K_I$ )
- Body effect (delta  $V_t$ )
- Tensione di Sub-threshold (Subvt)
- Corrente di saturazione ( $I_{ds}$ )
- Tensione di Breakdown Drain-Source ( $B_{vdss}$ )



## Tensione di di soglia (threshold voltage $V_t$ )

$$V_{t_0} = \varphi_{ms} - 2 \cdot \varphi_f - \frac{Q_{tot}}{C_{ox}} - \frac{Q_{BO}}{C_{ox}}$$

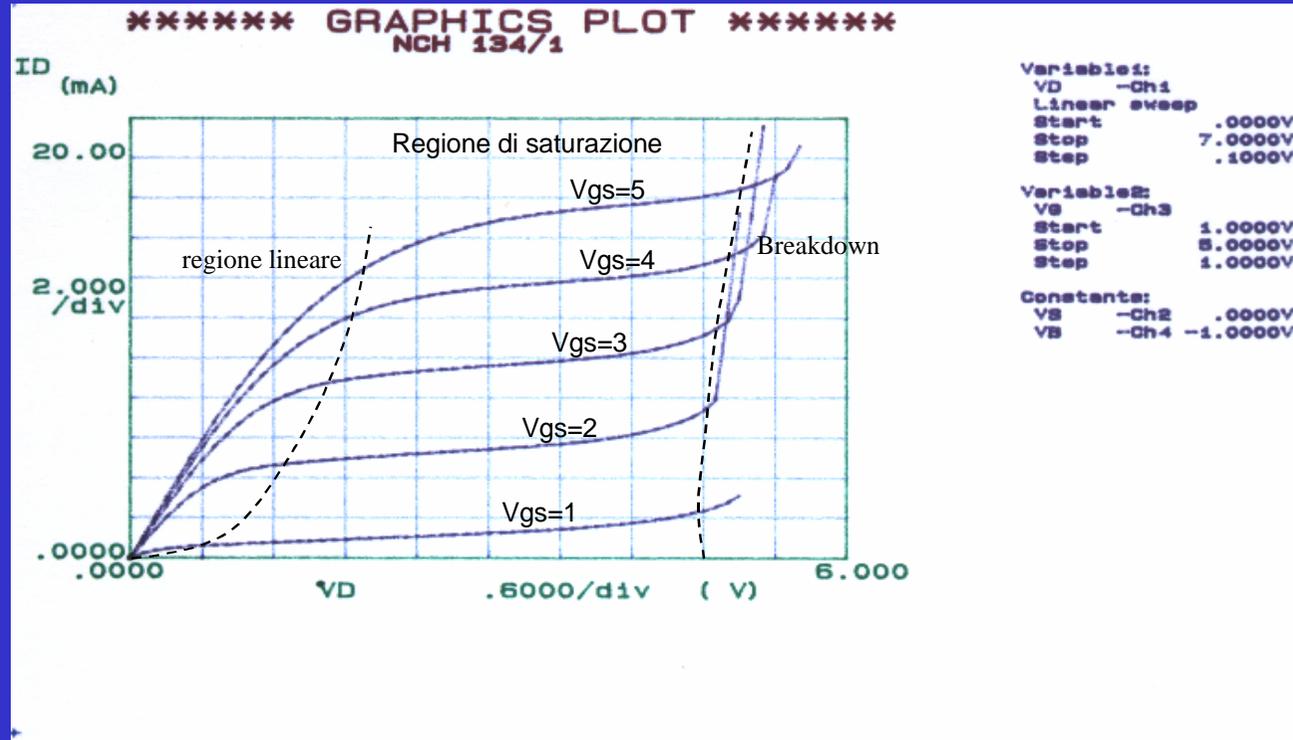
- $Q_{tot}$  è la carica netta per unità di area all'interfaccia ossido/silicio
- $Q_{bo}$  è la carica per unità di area nella regione di svuotamento controllata dal gate
- $\varphi_{ms}$  è la differenza della funzione lavoro tra il gate e il substrato
- $\varphi_f$  è il potenziale di Fermi del substrato
- $C_{ox}$  è la capacità per unità di area dell'ossido di gate

$$\varphi_f = \frac{K \cdot T}{q} \cdot \ln\left(\frac{N_A}{n_i}\right); \quad Q_{bo} = -\sqrt{2 \cdot q \cdot N_A \cdot \varepsilon_{Si} \cdot |-2 \cdot \varphi_f|}; \quad C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}$$

## Controllo della $V_t$

- La differenza di funzione lavoro dipende dal materiale di gate
- Il potenziale di Fermi dipende dal livello di drogaggio del substrato
- $Q_{tot}$  può essere controllata (ridotta) con opportuni metodi di ossidazione
- $Q_{bo}$  può essere modulata dal drogaggio del substrato
- $C_{ox}$  dipende dallo spessore e dalla costante dielettrica del materiale isolante

# Caratteristiche del transistor

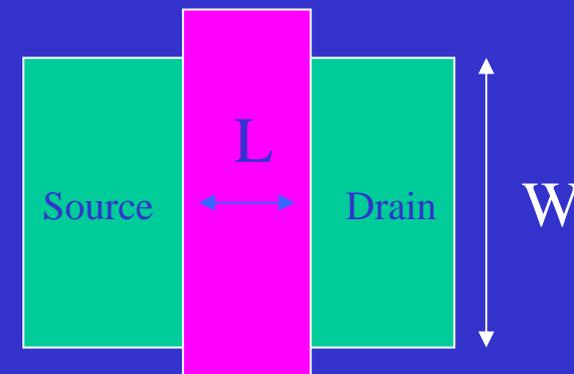


regione lineare :  $I_{sd} \approx K \cdot (V_{gs} - V_t) \cdot V_{sd}$

regione di saturazione :  $I_{sd} \approx \frac{K}{2} \cdot (V_{gs} - V_t)^2$

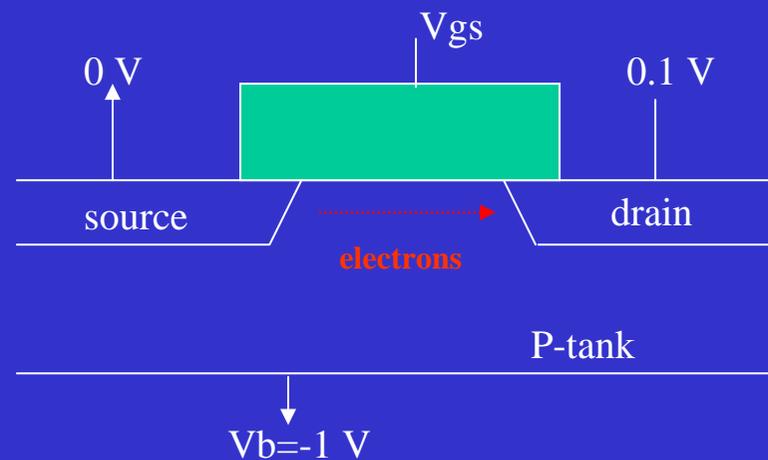
dove :

$$k = \mu \cdot C_{ox} \cdot \frac{W}{L} = \mu \cdot \frac{\epsilon_0 \cdot \epsilon_{ox}}{t_{ox}} \cdot \frac{W}{L}$$

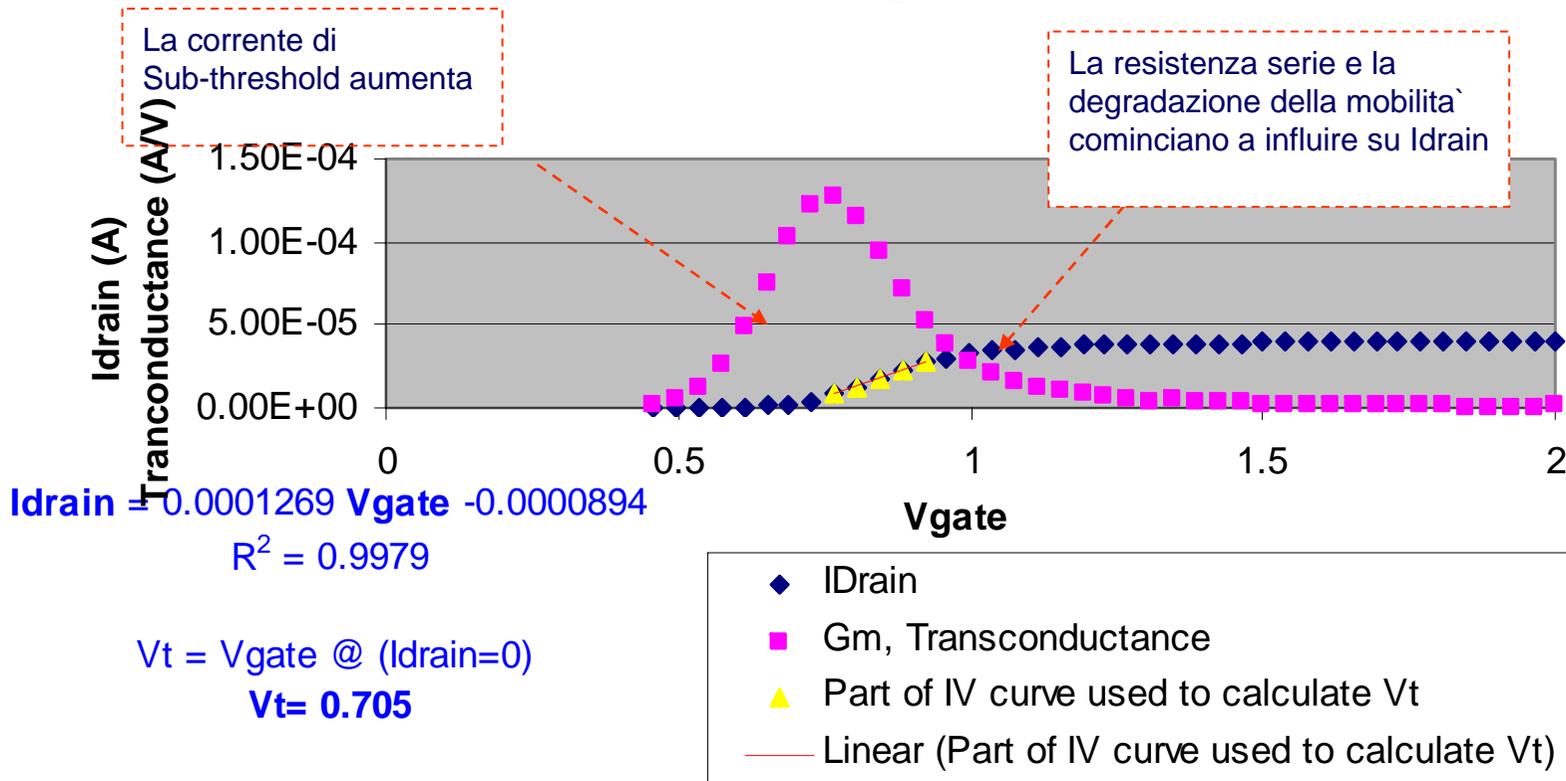


## Procedure di Test

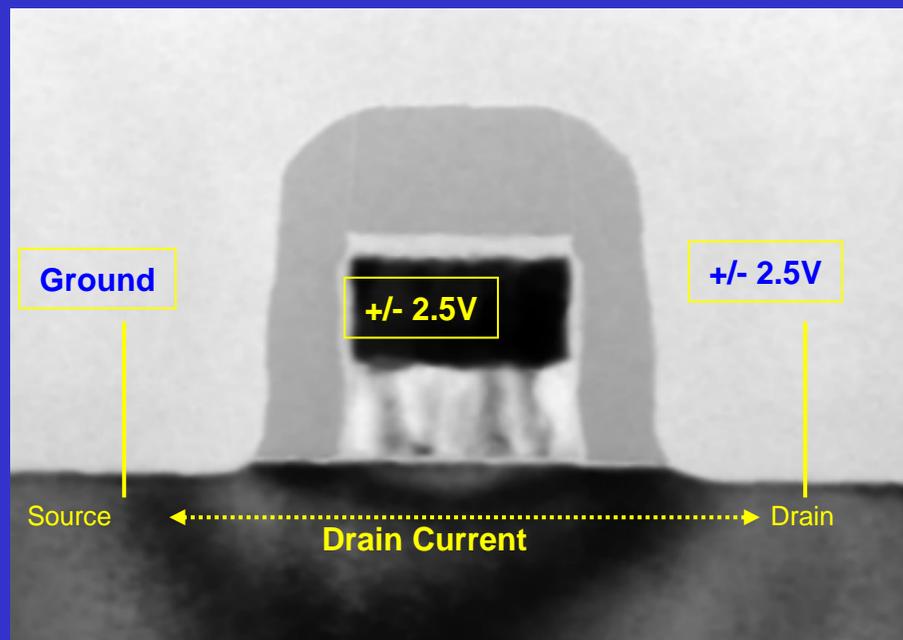
**$V_t$ :**  $V_b = -1\text{ V}$  (dipende dal dispositivo),  $V_s = 0\text{ V}$ ,  $V_d = 0.1\text{ V}$  (negativo per PMOS) e il gate è portato da 0 to  $\pm 3\text{ V}$ . Con una regressione lineare effettuata intorno al punto di massima tranconduttanza ( $G_{\max} = \max(dI_{sd}/dV_{gs})$ ). L'intercetta di tale retta con l'asse delle ascisse ( $V_{gs}$ ) corrisponde al valore di  $V_t$



## N Channel Transistor Vgate-Idrain curve

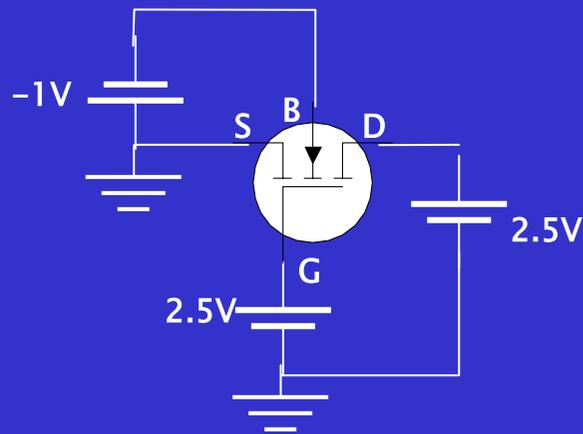


**Ids:** La corrente di drain è misurata con  $V_b = -1V$ ,  $V_s = 0V$ ,  
 $V_d = V_{gs} = \pm 2.5$ . Il valore ottenuto è diviso per  $W$  (larghezza del transistor).  $I_{ds}$  è normalmente espresso in  $A/\mu m$



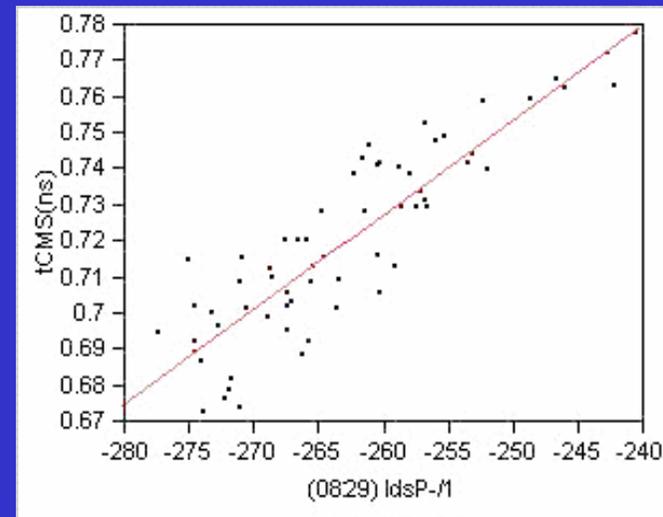
# Corrente di Saturazione ( $I_{ds}$ )

Misura la massima corrente di drain quando il transistor e' in condizioni di saturazione ( $V_{ds} > V_{gs} - V_t$ ). Il risultato viene normalizzato alla larghezza del transistor.



$$I_{ds} = \text{Drain Current} / \text{Width} \left[ \mu\text{A} / \mu\text{m} \right]$$

• Bivariate Fit of tCMS(ns) By (0829) IdsP-/1



**K<sub>lin</sub>**: e` il massimo valore di trasconduttanza normalizzato alla tensione di drain e alla larghezza del transistor. E` il guadagno lineare del dispositivo (unita`: uA/(um\*V^2))

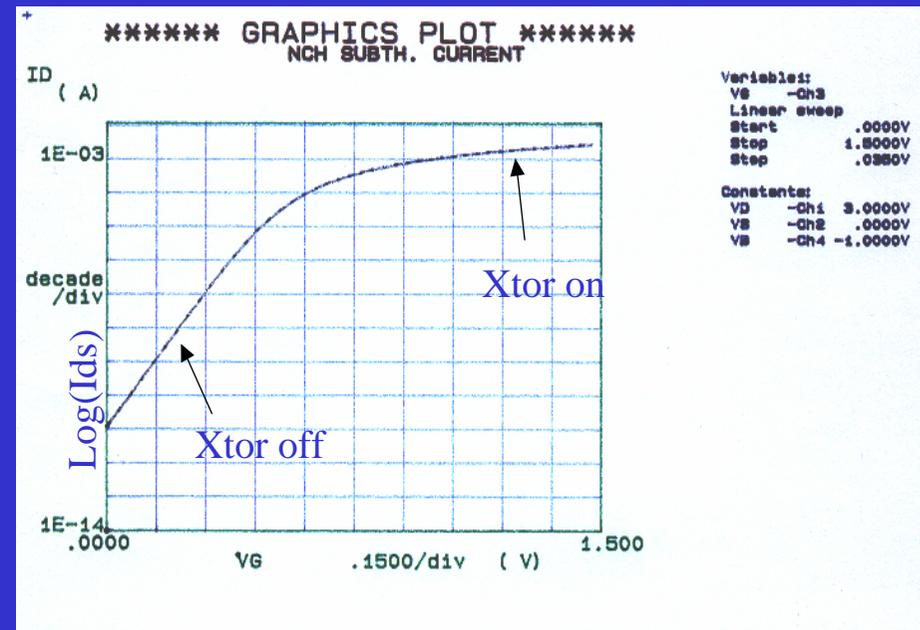
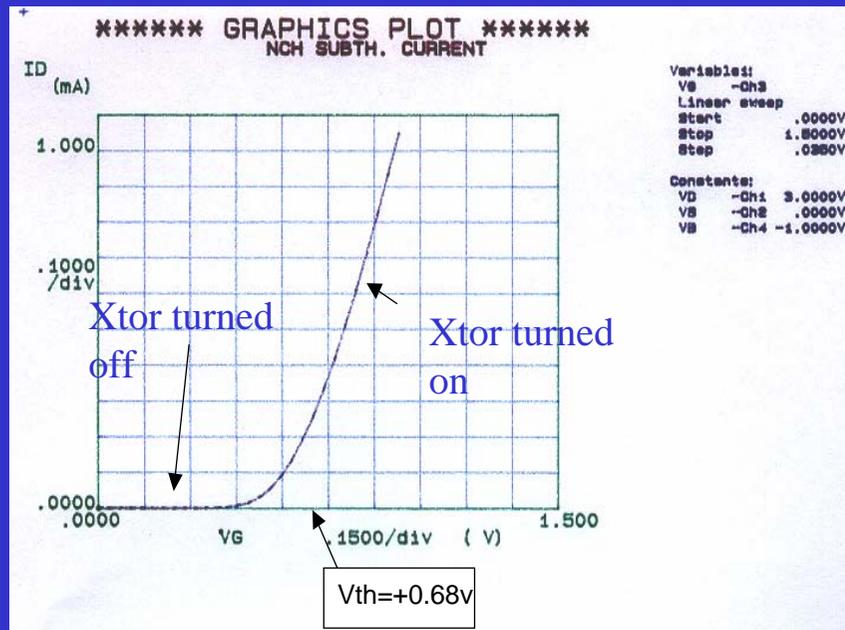
Nella regione lineare:

$$I_{ds} \approx k \cdot (V_{gs} - V_t) \cdot V_{ds}$$

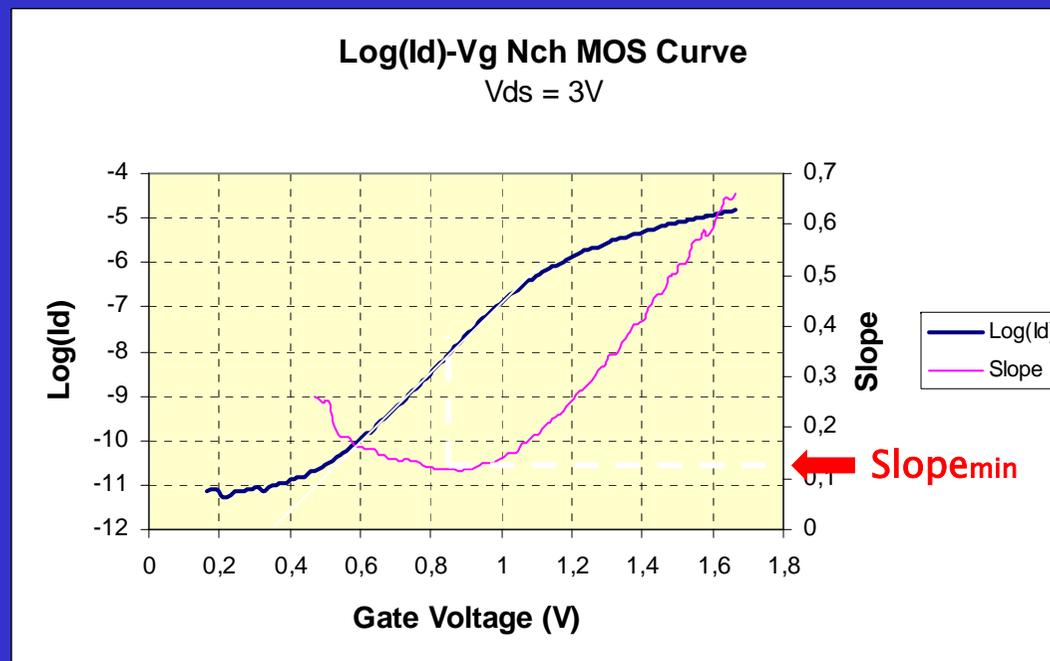
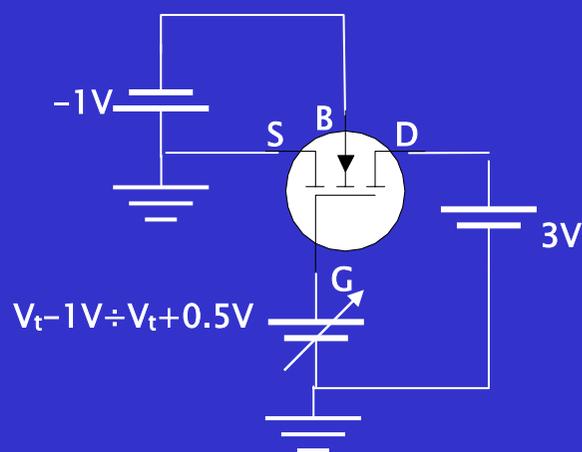
quindi:

$$\frac{dI_{ds}}{dV_{gs}} \approx k \cdot V_{ds} = G_{\max} \Rightarrow K_{lin} = \frac{k}{W} = \frac{G_{\max}}{V_{ds} \cdot W}$$

**SubVt:** La sub-threshold  $V_t$  ( $V_t$  sotto-soglia) è la tensione al di sotto della  $V_t$ , alla quale si ha un'apprezzabile corrente di perdita. La tensione di gate è incrementata, la corrente  $I_{ds}$  è misurata. La sub-threshold  $V_t$  è il valore di  $V_{gs}$  per il quale si misura una corrente pari a  $1\text{pA} \cdot W$



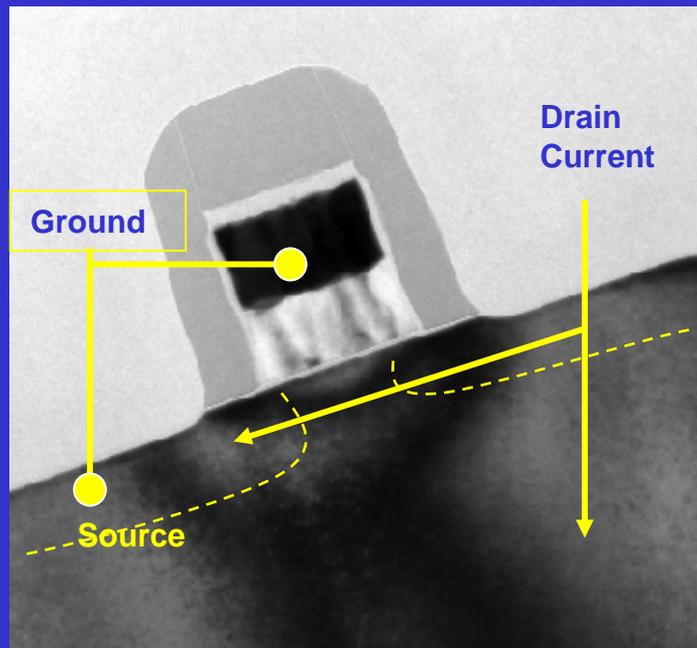
La tensione di *Sub-threshold* (*Subvt*) viene definita come la tensione del gate alla quale la corrente di drain e' pari a  $1\text{pA} \cdot \text{Width}$ .



Dal momento che  $\text{Log}(I_d) - V_g$  e' lineare nella regione di subthreshold, e' possibile calcolare la *Subvt* mediante l'equazione:

$$SubVT = V_g \Big|_{I_d = 1\text{pA} \cdot \text{Width}} = Slope_{\min} \cdot \text{Log}(1\text{pA} \cdot \text{Width}) + Intercept_{\min}$$

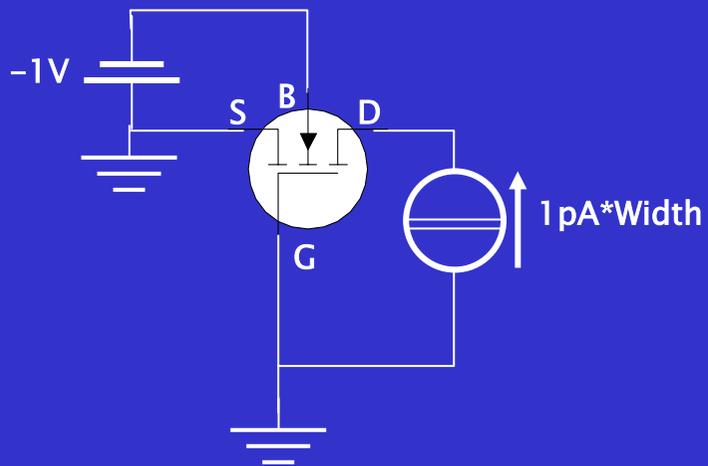
**Bvdss:** Questo parametro misura il minore tra il punch-through source-drain e la rottura della giunzione di drain. La tensione di drain è portata da 0 a +/- 25V, con  $V_{gs}=0$ . Il valore di  $V_d$  per il quale si ha  $I_{drain}=1\mu A \cdot W$  è preso come tensione di break-down



Due componenti possono contribuire a BVDSS: punch-through and junction leakage. Il punch-through avviene quando le regioni di source e drain vengono a contatto

# Tensione di Breakdown Drain-Source

Misura la tensione tra drain e source ottenuta forzando una corrente sul Drain di  $1\text{ nA} \cdot \text{Width}$  quando il transistor e' spento.



$$Bvdss = V_d \Big|_{I_{ds} = 1\text{ nA} \cdot \text{Width}} \text{ [V]}$$

## Body effect

In transistor a canale lungo, uniformemente drogato, la  $V_t$  dipende dalla tensione di substrato  $V_b$ . Per un NMOS, applicando una tensione negativa al substrato, il valore della  $V_t$  aumenta in accordo con la seguente formula

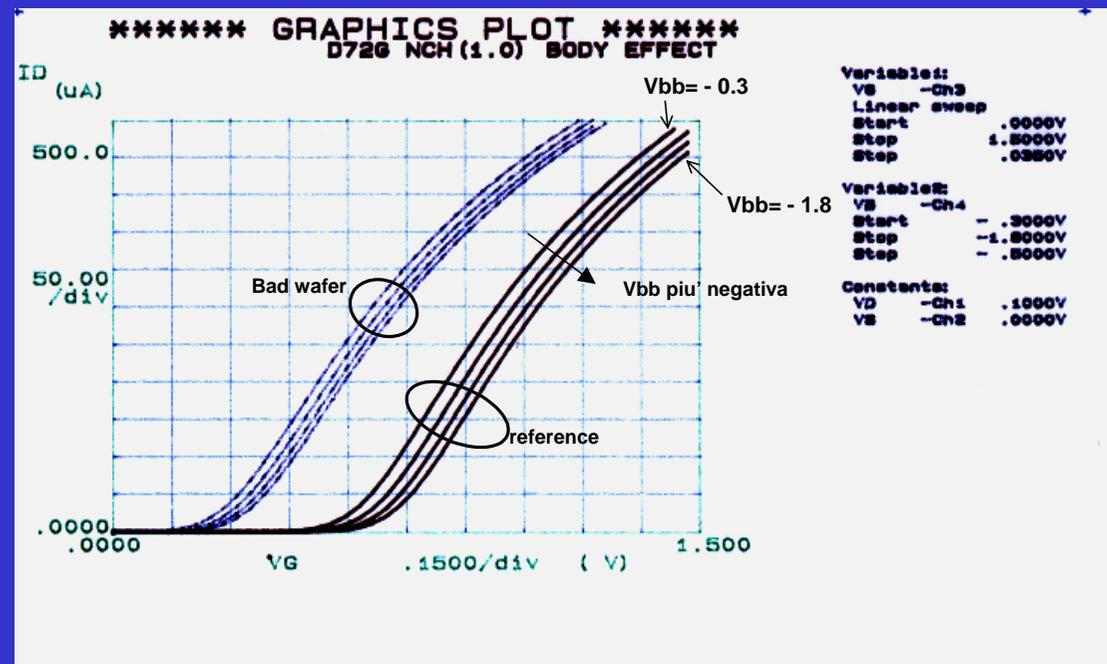
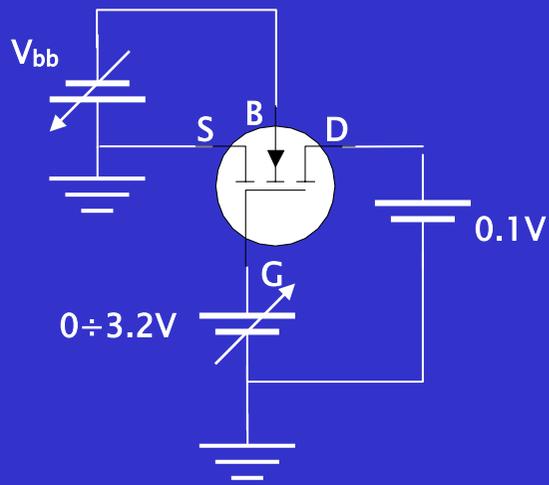
$$\Delta V_t \approx \frac{\sqrt{2 \cdot q \cdot \epsilon_{Si} \cdot N_b}}{C_{ox}} \cdot \sqrt{|\Delta V_b|} = \gamma \cdot \sqrt{|\Delta V_b|}$$

dove  $N_b$  è la concentrazione di specie droganti nel bulk,  $C_{ox}$  è la capacità per unità di area,  $V_b$  è la tensione applicata al bulk.

Poiché il fattore gamma dipende dalla concentrazione nel bulk, la differenza di  $V_t$  misurata può essere usata per identificare deviazione del livello di drogaggio del substrato (dovuto a problemi di processo / contaminazioni).

Il *body effect* e' il fenomeno che descrive la variazione della  $V_t$  quando differenti tensioni di polarizzazione sono applicate al bulk del transistor.

Per valutare il *body effect*, la  $V_t$  viene misurata a diversi valori di  $V_{bb}$ . Vengono quindi riportati I valori assoluti delle differenze delle  $V_t$



DVT: Delta  $V_t$  (N Channel)

$$DVT1 = | V_t @ V_{bs} = -1 - V_t @ V_{bs} = -0.5 |$$

$$DVT2 = | V_t @ V_{bs} = -1 - V_t @ V_{bs} = -1.25 |$$

## Impatto delle dimensioni (L e W) sul funzionamento

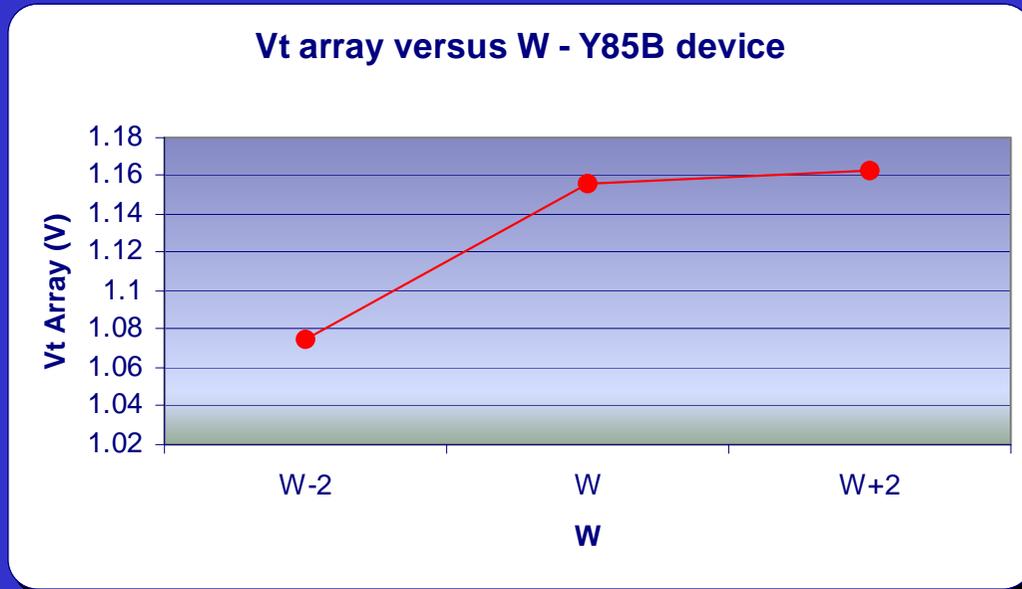
Nel dispositivo vengono utilizzati transistor con differenti dimensioni (lunghezza L di canale e larghezza W). Nelle strutture parametriche vengono riprodotte le tipologie più rappresentative

La misura di transistor con diverse dimensioni può aiutare a capire fallimenti funzionali e separare problematiche di natura fotolitografica da quelle legate ad impiantazioni/deposizioni

### Esempi di transistor

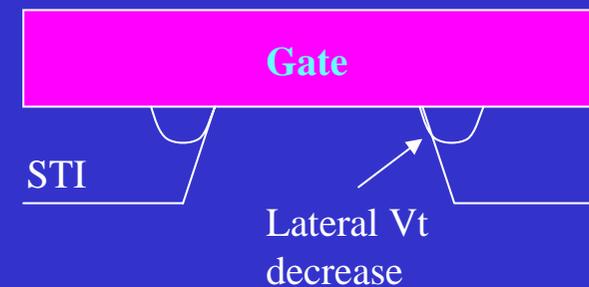
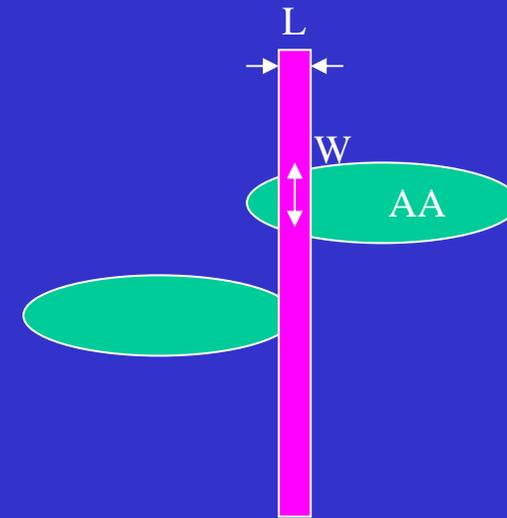
- Short channel --->  $W = 134$ ,  $L = 1$  um as drawn
- Long channel --->  $W = 134$ ,  $L = 64$  um as drawn

## Narrow width effect

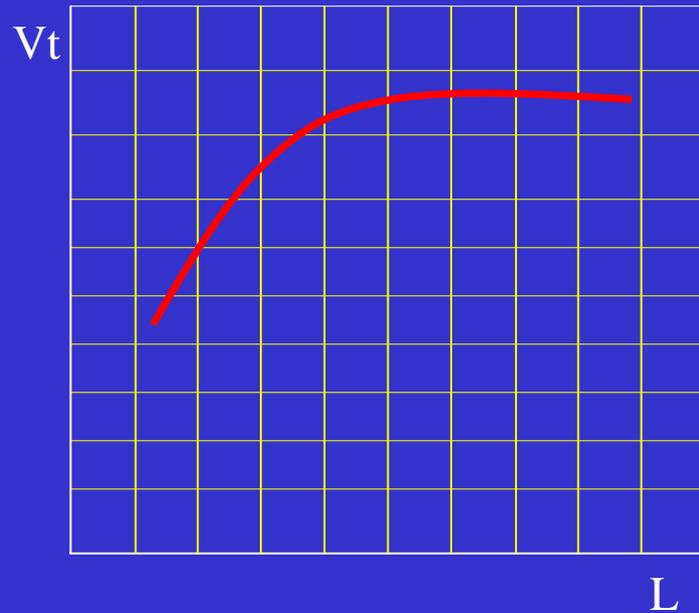


L'undercut dell'STI da luogo a due transistor parassiti con minore  $V_t$  rispetto al nominale.

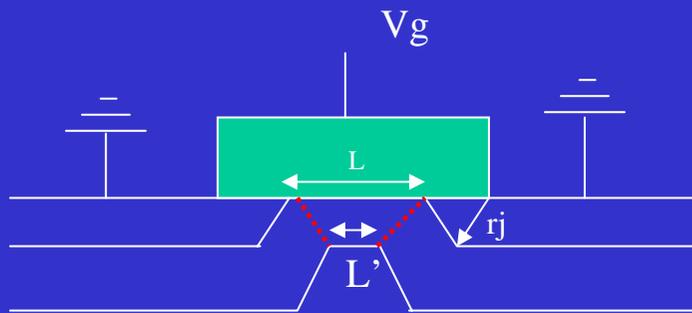
Quando la larghezza del transistor ( $W$ ) diminuisce il contributo dei due transistor parassiti diventa più importante provocando la riduzione della  $V_t$  del nominale



## Short channel effect



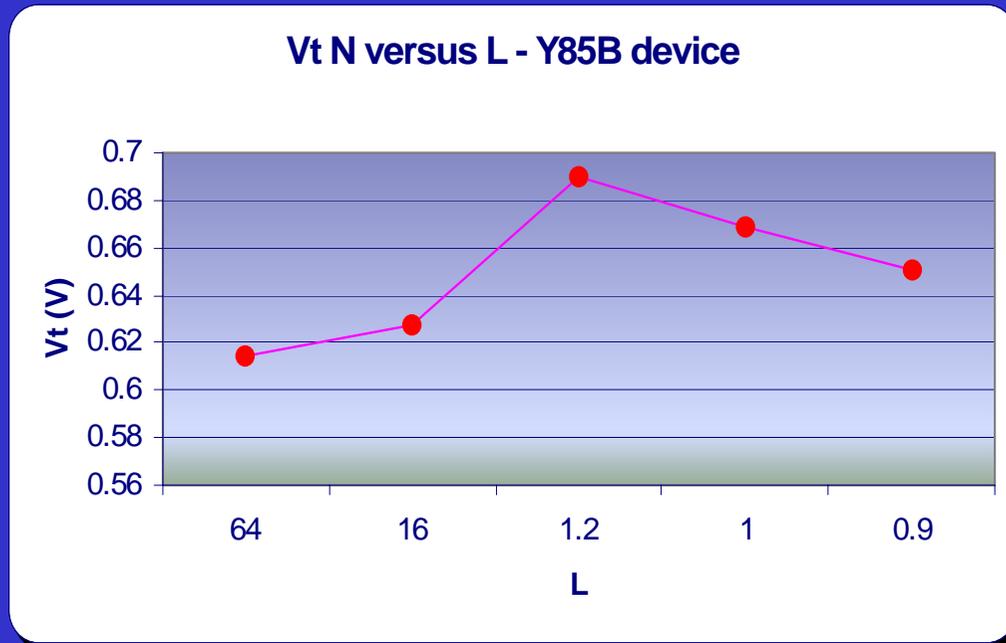
Quando la lunghezza del canale ( $L$ ) diminuisce la carica  $Q_{bo}$  nella regione di svuotamento controllata dal gate diminuisce e cio` provoca la riduzione del valore di  $V_t$



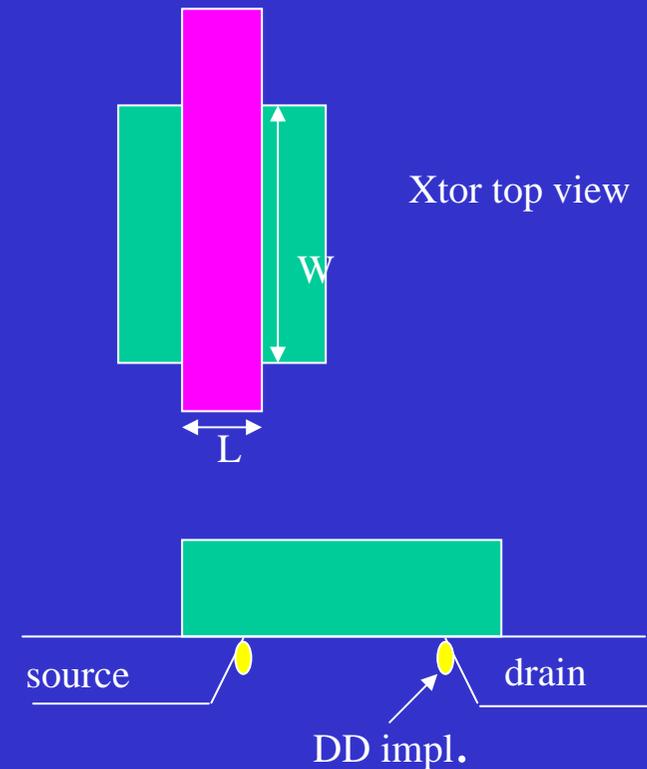
$$\Delta V_t = - \frac{q \cdot N_A \cdot L'}{C_{ox}} \left\{ \frac{r_j}{L} \left( \sqrt{1 + \frac{2 \cdot L'}{r_j}} - 1 \right) \right\}$$

Differenza tra la  $V_t$  di due transistor con canali  $L$  ed  $L'$  (a parita` di tutti gli altri parametri)

## Reverse short channel effect

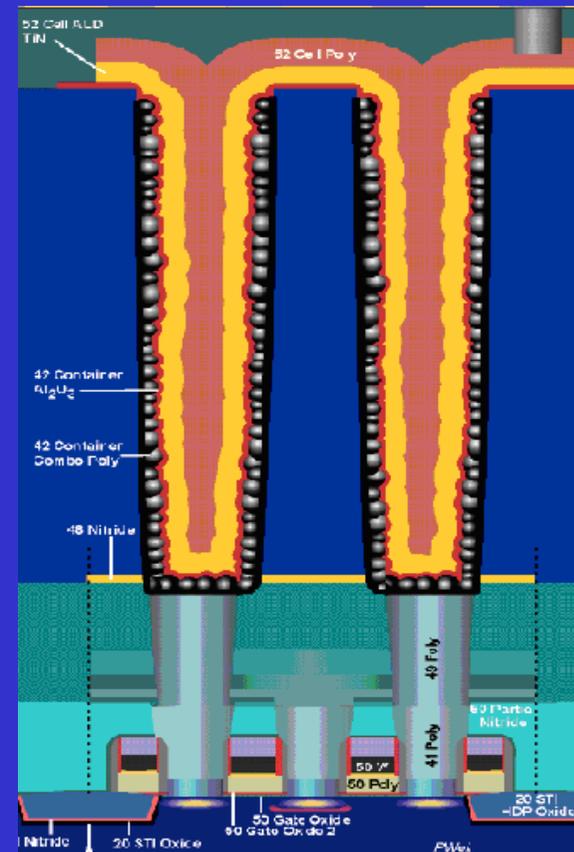
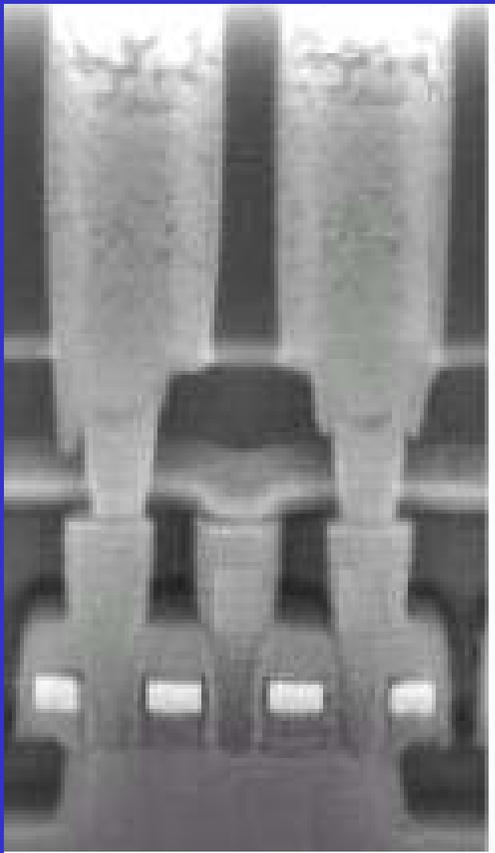


Le impiantazione LDD e responsabile del 'roll-up' della  $V_t$  intorno ad  $L=1.2\mu\text{m}$ . Poi per valori inferiori di  $L$ , l'effetto 'dimensionale' domina e la  $V_t$  diminuisce ('roll-off') lo 'short channel effect' predomina



# Miscellanea

# Cell Capacitor



# Cross section

