Monolithic 3D DRAM Technology

Deepak C. Sekar, Brian Cronquist, Israel Beinglass, Paul Lim, Zvi Or-Bach

15th June 2011

Outline

- Status of the DRAM industry today
- Monolithic 3D DRAM
- Implications and risks of the technology
- Summary

Outline

- Status of the DRAM industry today

DRAM makers fall in the “endangered species” category

1996
24 key DRAM players

Samsung	Hyundai
Micron	Siemens
NEC	Hitachi
Mitsubishi	Toshiba
Fujitsu	LG Semicon
TI-Acer	Vanguard
Powerchip	ProMOS
Winbond	Oki
IBM	TI
Motorola	Matsushita
Seiko Epson	Nippon Steel
UMC	Mosel Vitelic

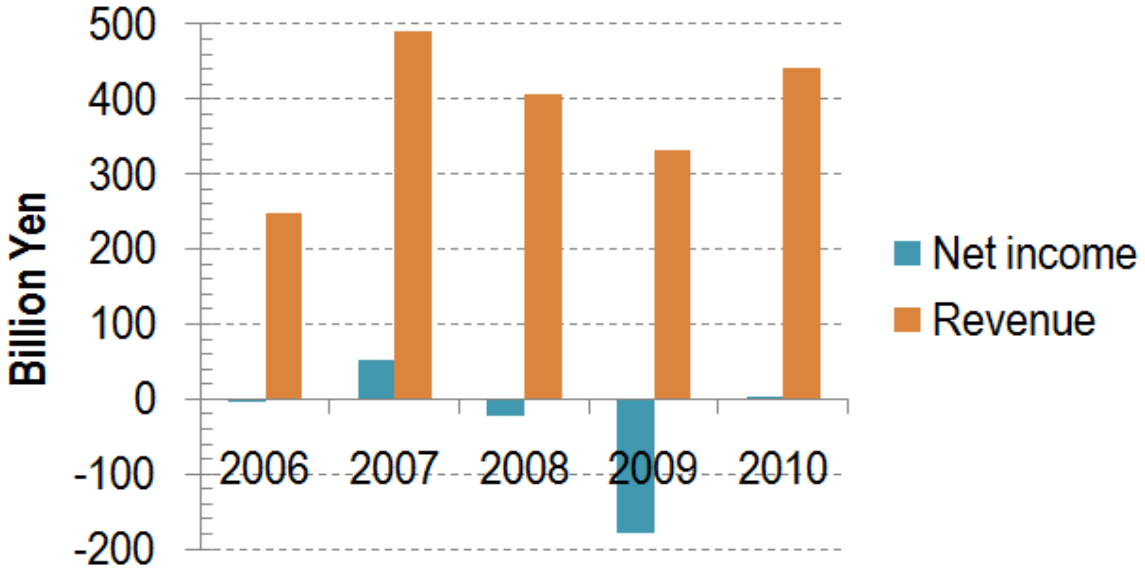
2010
9 key DRAM players

- Samsung
- Hynix
- Micron
- Elpida
- Nanya
- Inotera
- Powerchip
- ProMOS
- Winbond

Why? What are the challenges? We’ll see in the next few slides

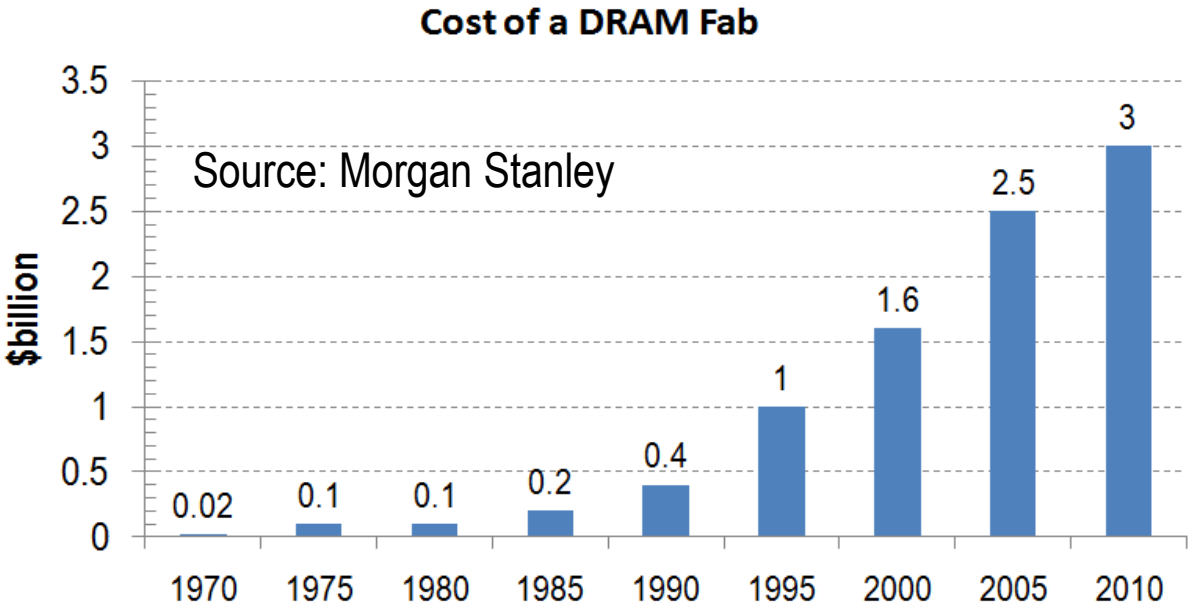
Reason 1: Profitability

Financials of a top-tier DRAM vendor (Elpida) vs. fiscal year



- DRAM has not been a profitable business in the near past
- Balance sheets of most companies' DRAM businesses similar to above ☹

Reason 2: Large fab cost for scaling-down



Today,

Litho Tool Cost = $42M

Etch, CVD, Implant, RTA tools each cost <$5M

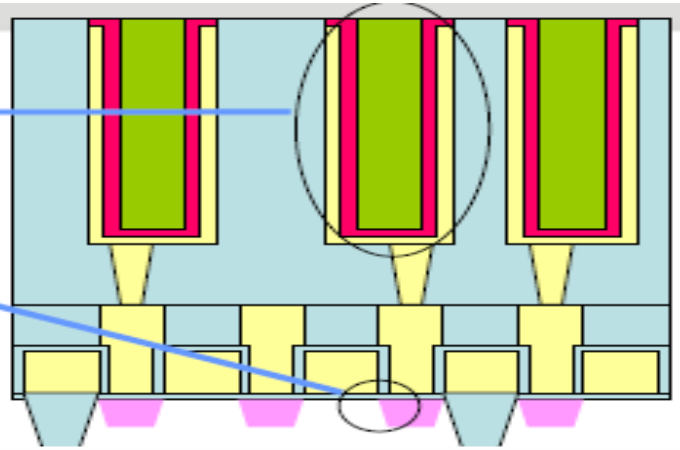
- **Scaling-down → lower cost per bit → but huge litho and fab investment**
- **Hard for unprofitable companies to fund scaled-down fabs. But if they don't fund new "scaled-down" fabs and others do → bad cost per bit → profitability even worse ☹**

Reason 3:
Scaling-down the stacked capacitor challenging

Capacitance
Keep ~25fF

Memory Cell Transistor
Keep low leakage current

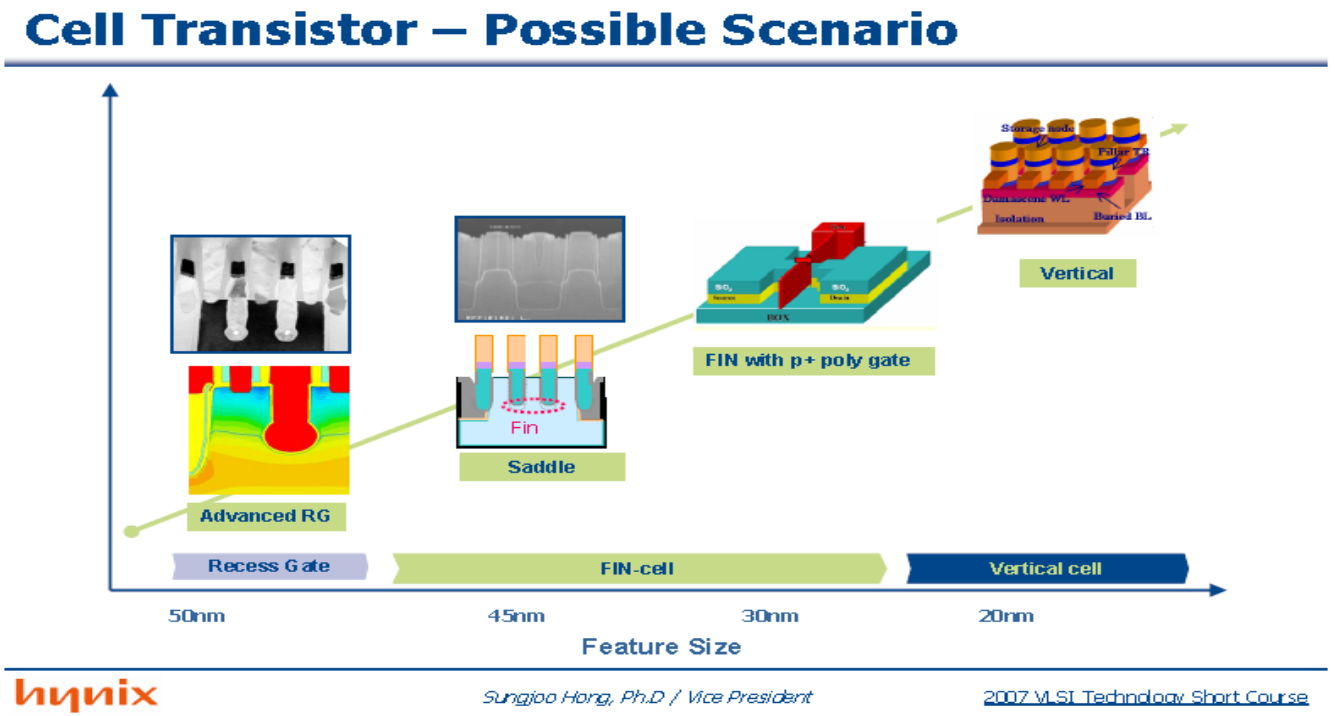
Al_2O_3 (90nm) → HfO_2 (80nm)
→ ZrO_2 (60nm) → ?



Source: ITRS 2010	45nm	32nm	22nm	15nm	10nm
Dielectric constant	40	50	60	65	70
Aspect ratio	47:1	56:1	99:1	147:1	193:1
EOT	0.8nm	0.6nm	0.5nm	0.3nm	0.2nm

Requires >150:1 aspect ratios and exotic new high-k dielectrics!

Reason 4:
The cell transistor needs major updates on scaling-down



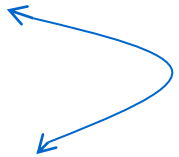
A major new transistor every generation or two!

100nm Planar → 80nm RCAT → 60nm S-RCAT → 35nm Finfet (?) → 20nm Vertical (?)

To recap,
Things don't look good for DRAM vendors because

(1) Low profitability

(2) Cost of scaled-down fabs



Related

(3) Scaling-down stacked capacitor

(4) Cell transistor scaling-down

Common theme → **Scaling-down**

Is there an alternative way to reduce DRAM bit cost other than scaling-down?

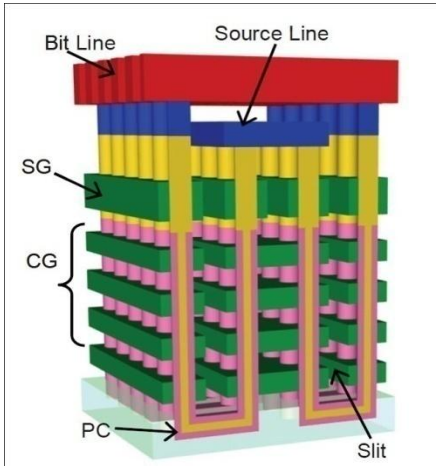
Focus of this presentation

Outline

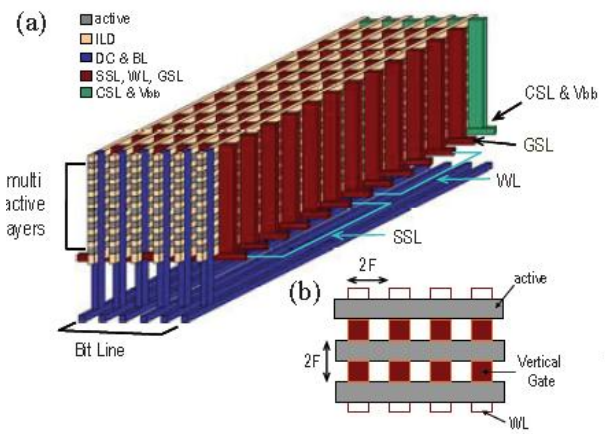
- **Monolithic 3D DRAM**

Key technology direction for NAND flash: Monolithic 3D with shared litho steps for memory layers

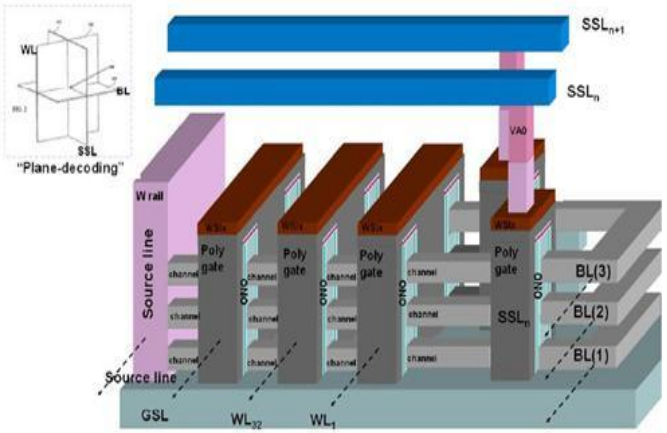
Toshiba BiCS
Poly Si



Samsung VG-NAND
Poly Si



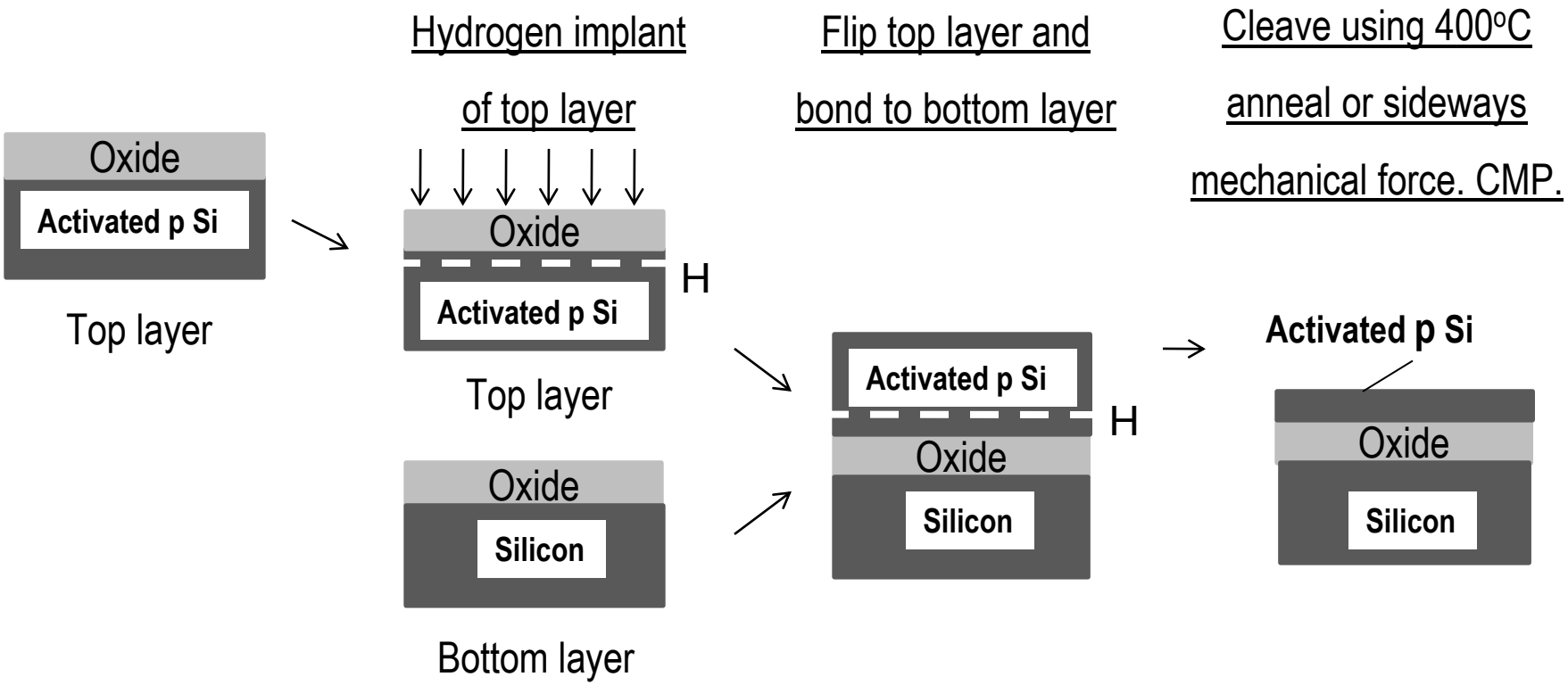
Macronix junction-free NAND
Poly Si



To be viable for DRAM, we require

- **Single-crystal silicon at low thermal budget → Charge leakage low**
- **Novel monolithic 3D DRAM architecture with shared litho steps**

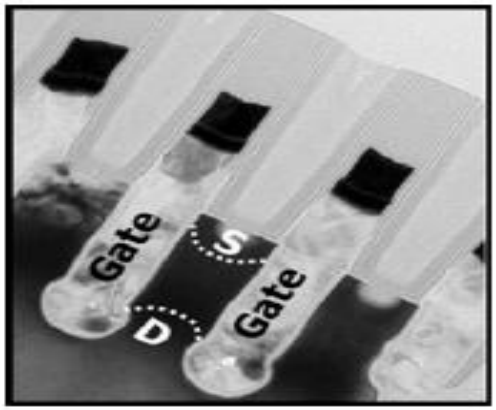
Single crystal Si at low thermal budget



- Obtained using the ion-cut process. It's use for SOI shown above.
- Ion-cut used for high-volume manufacturing SOI wafers for 10+ years.

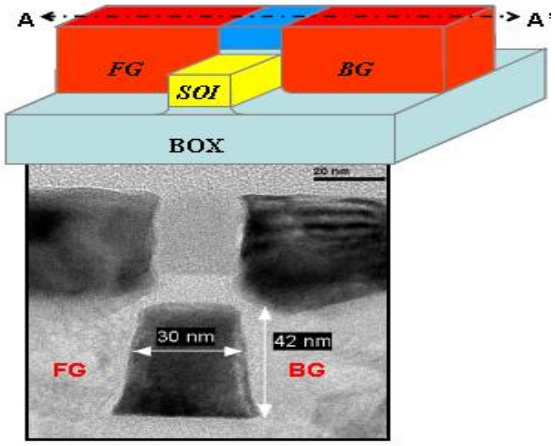
Double-gated floating body memory cell well-studied in Silicon (for 2D-DRAM)

Hynix + Innovative Silicon
VLSI 2010



- 0.5V, 55nm channel length
- 900ms retention
- Bipolar mode

Intel
IEDM 2006

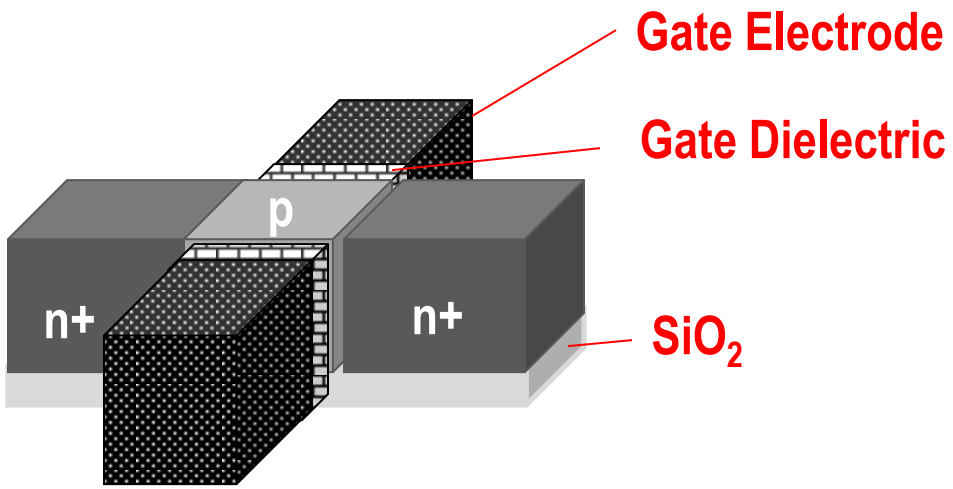


- 2V, 85nm channel length
- 10ms retention
- MOSFET mode

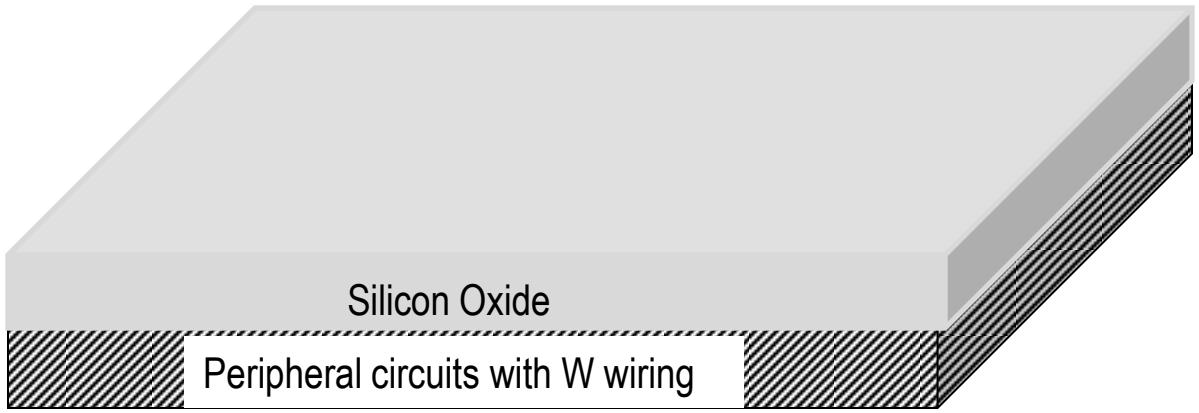
Our novel DRAM architecture

Innovatively combines these well-studied technologies

- Monolithic 3D with litho steps shared among multiple memory layers
- Stacked Single crystal Si with ion-cut
- Double gate floating body RAM cell (below) with charge stored in body

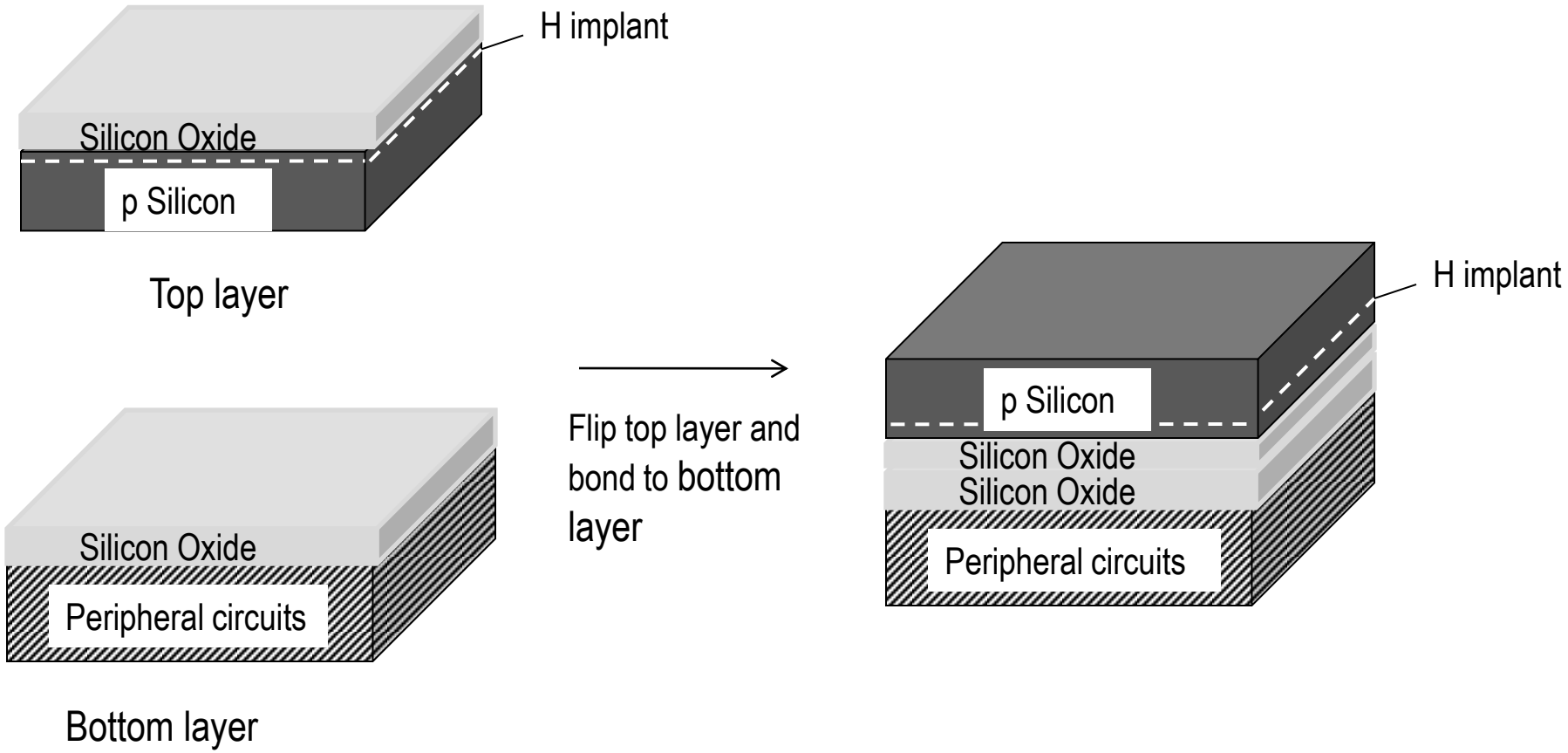


Process Flow: Step 1
Fabricate peripheral circuits followed by silicon oxide layer

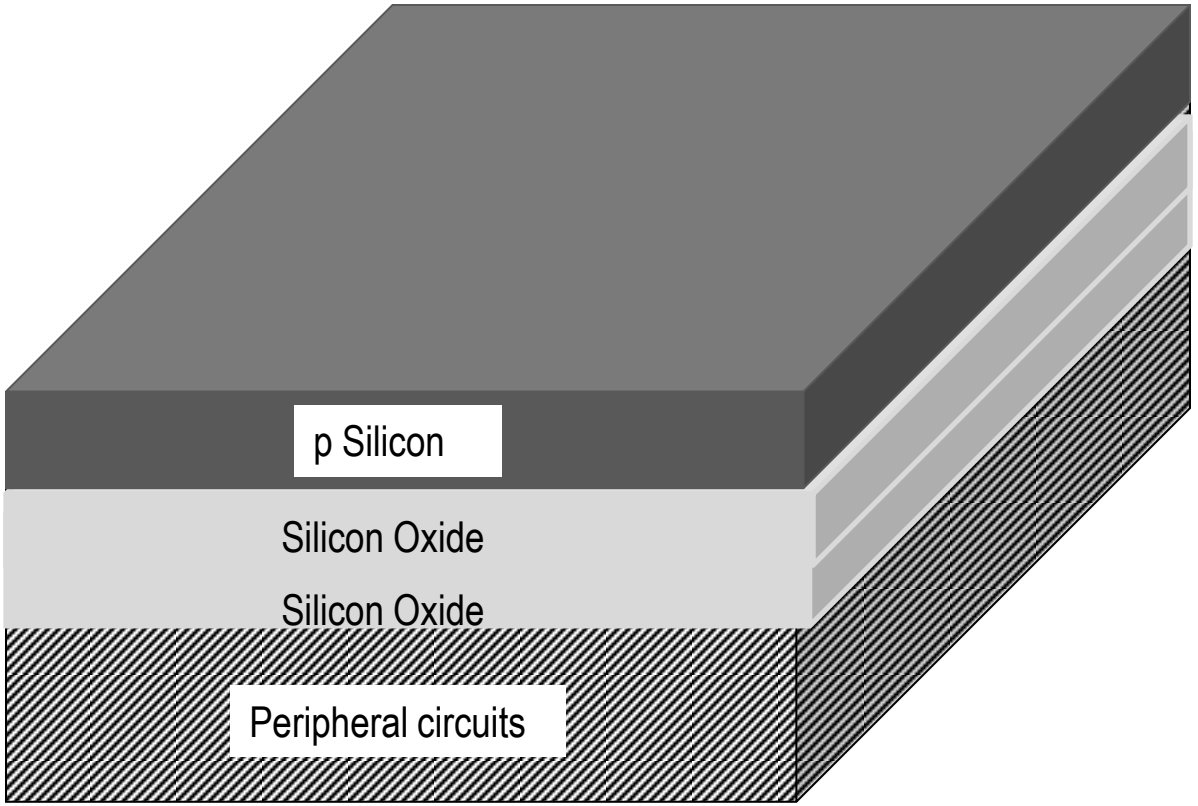


MonolithIC3D™

Process Flow: Step 2
Transfer p Si layer atop peripheral circuit layer

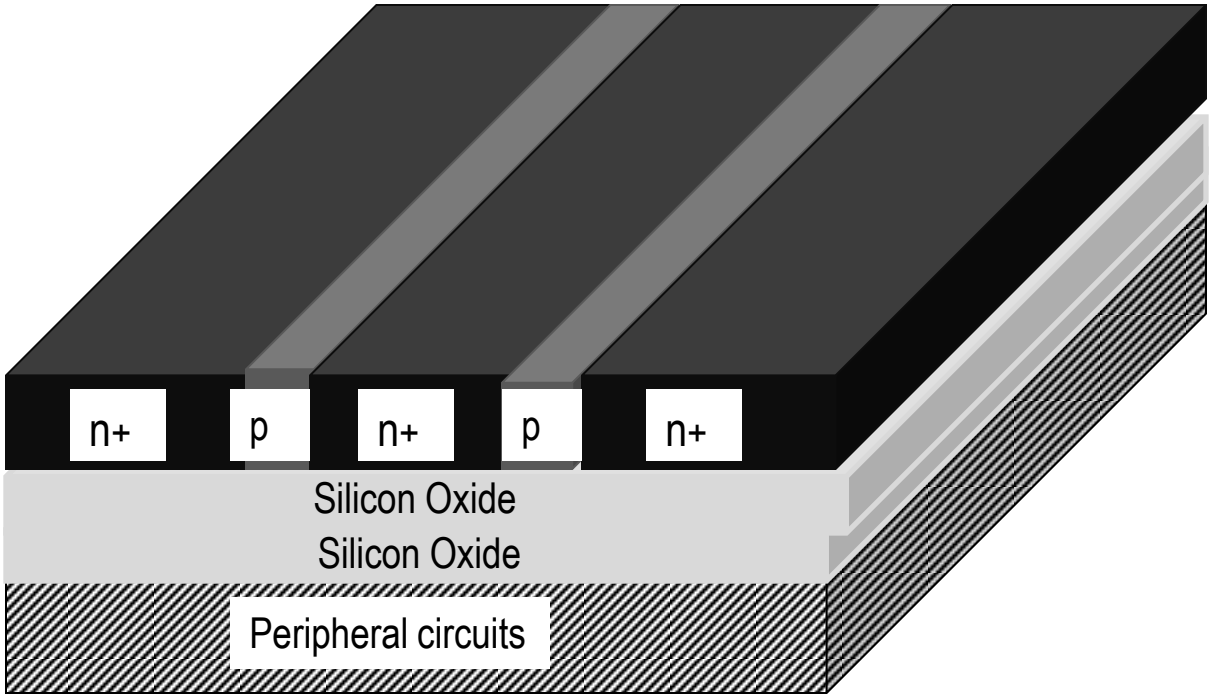


Process Flow: Step 3
Cleave along H plane, then CMP

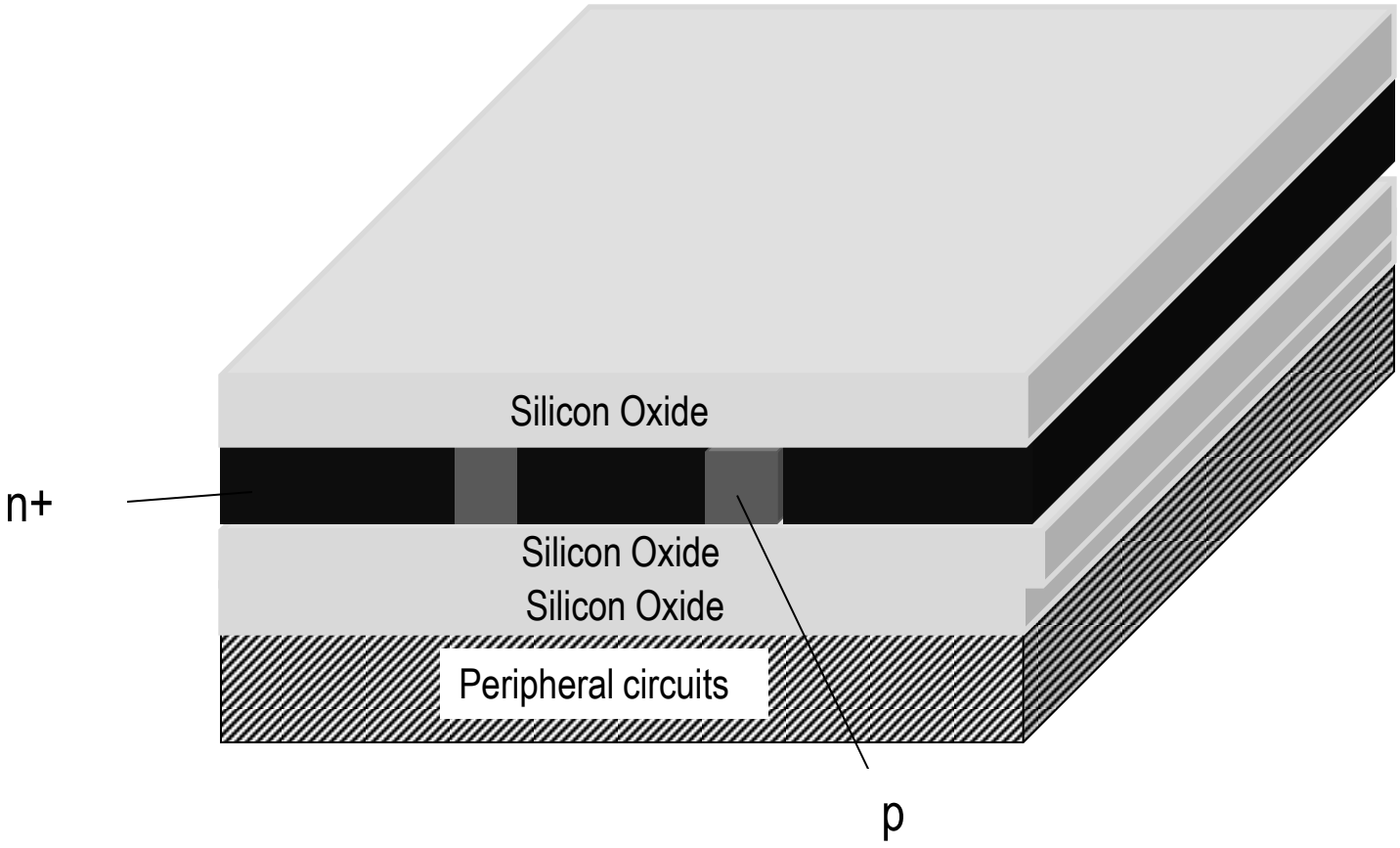


MonolithIC3D™

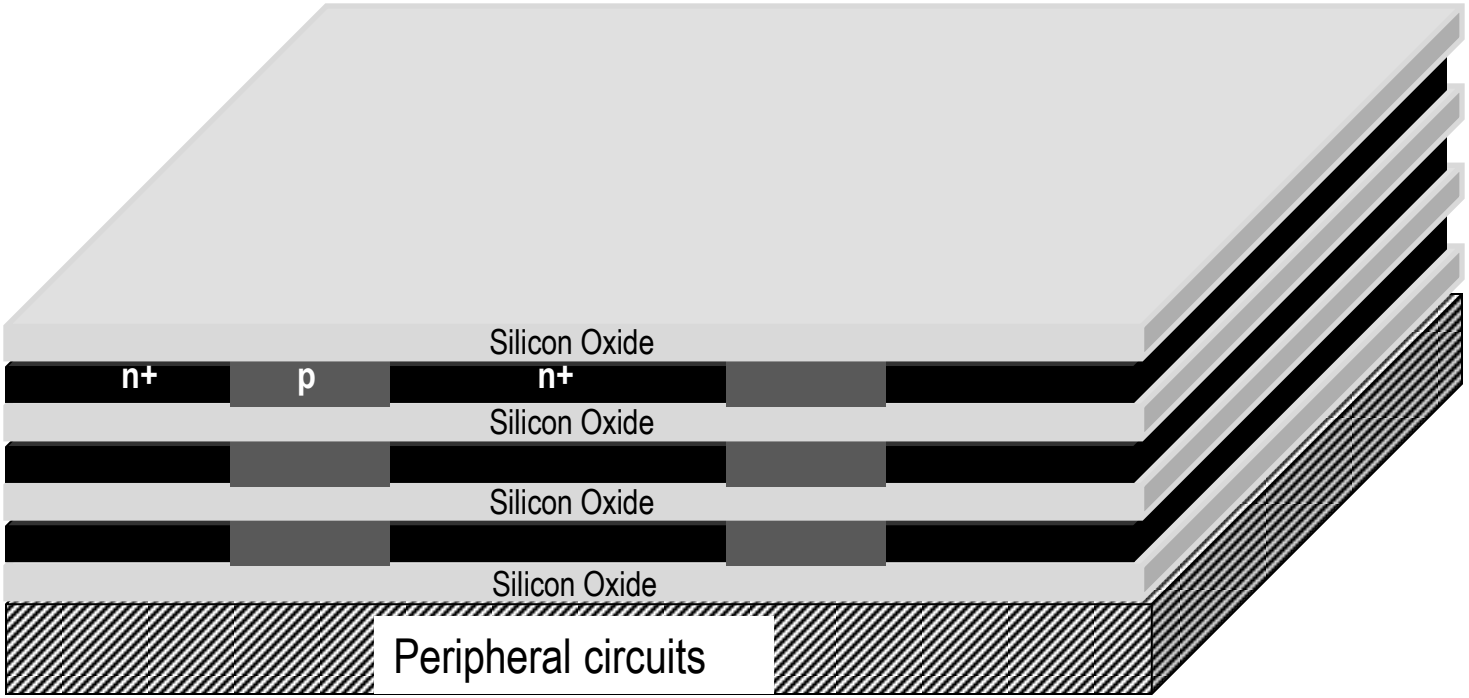
Process Flow: Step 4
Using a litho step, form n+ regions using implant



Process Flow: Step 5
Deposit oxide layer

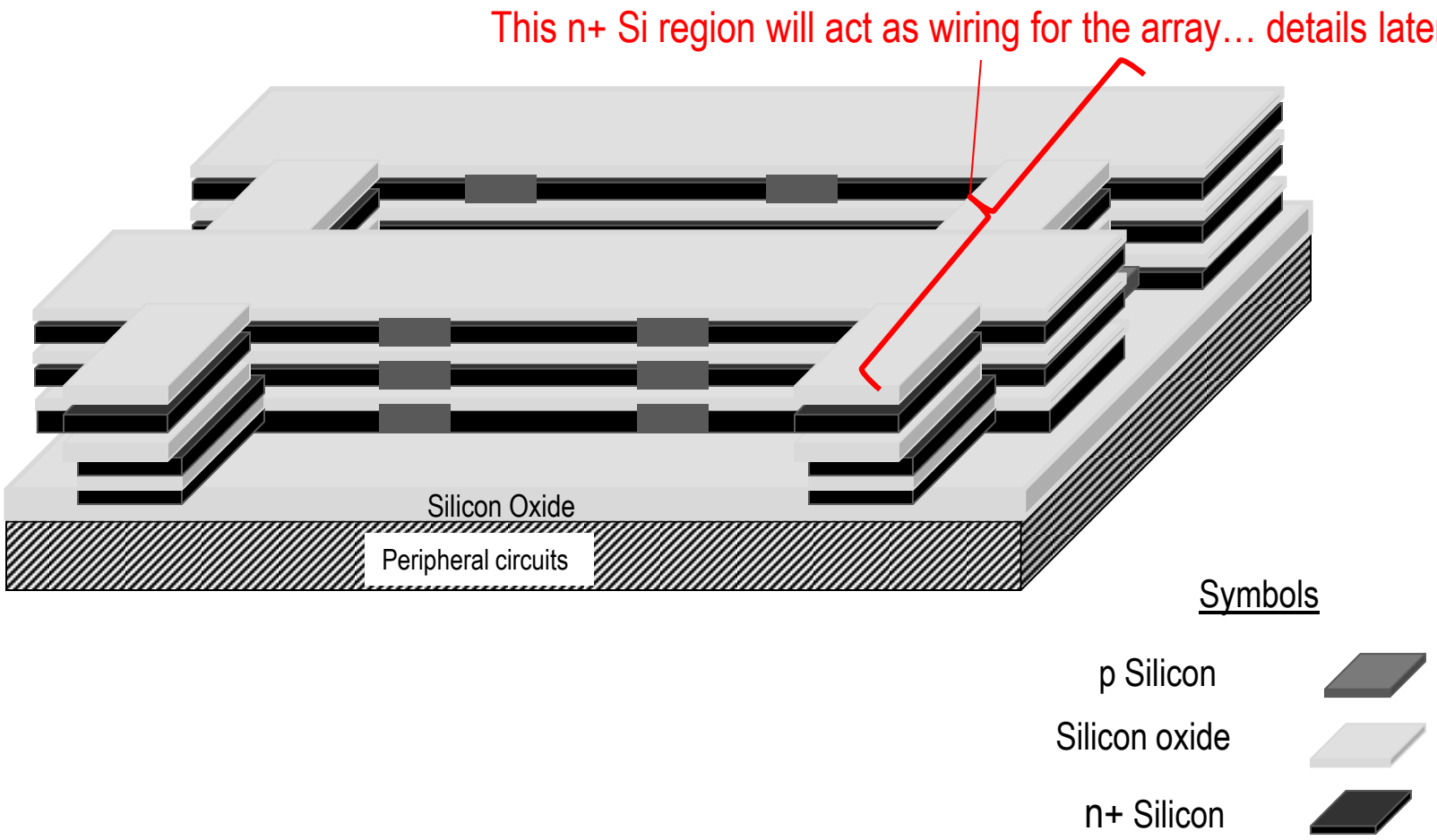


Process Flow: Step 6
Using methods similar to Steps 2-5, form multiple Si/SiO_2 layers, RTA



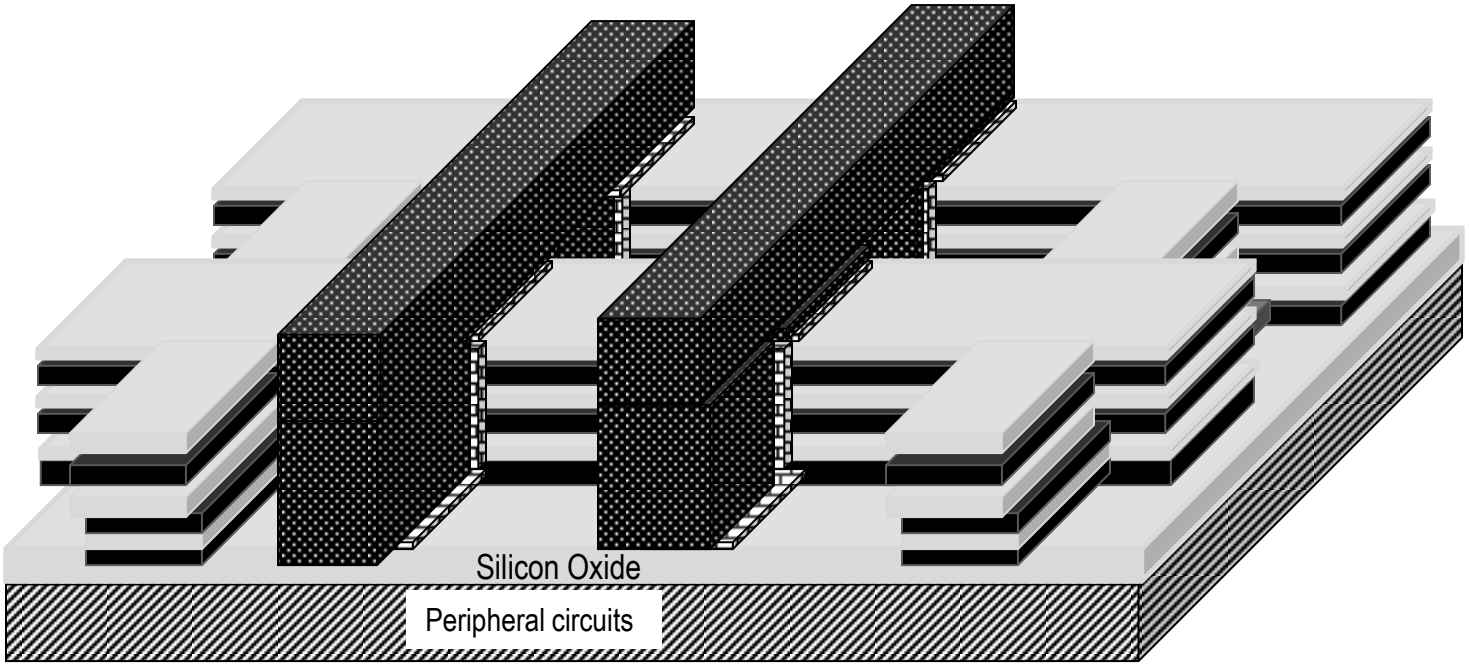
MonolithIC3D™

Process Flow: Step 7
Use lithography and etch to define Silicon regions



MonolithIC3D™

Process Flow: Step 8
Deposit gate dielectric, gate electrode materials, CMP, litho and etch



Symbols

n+ Silicon

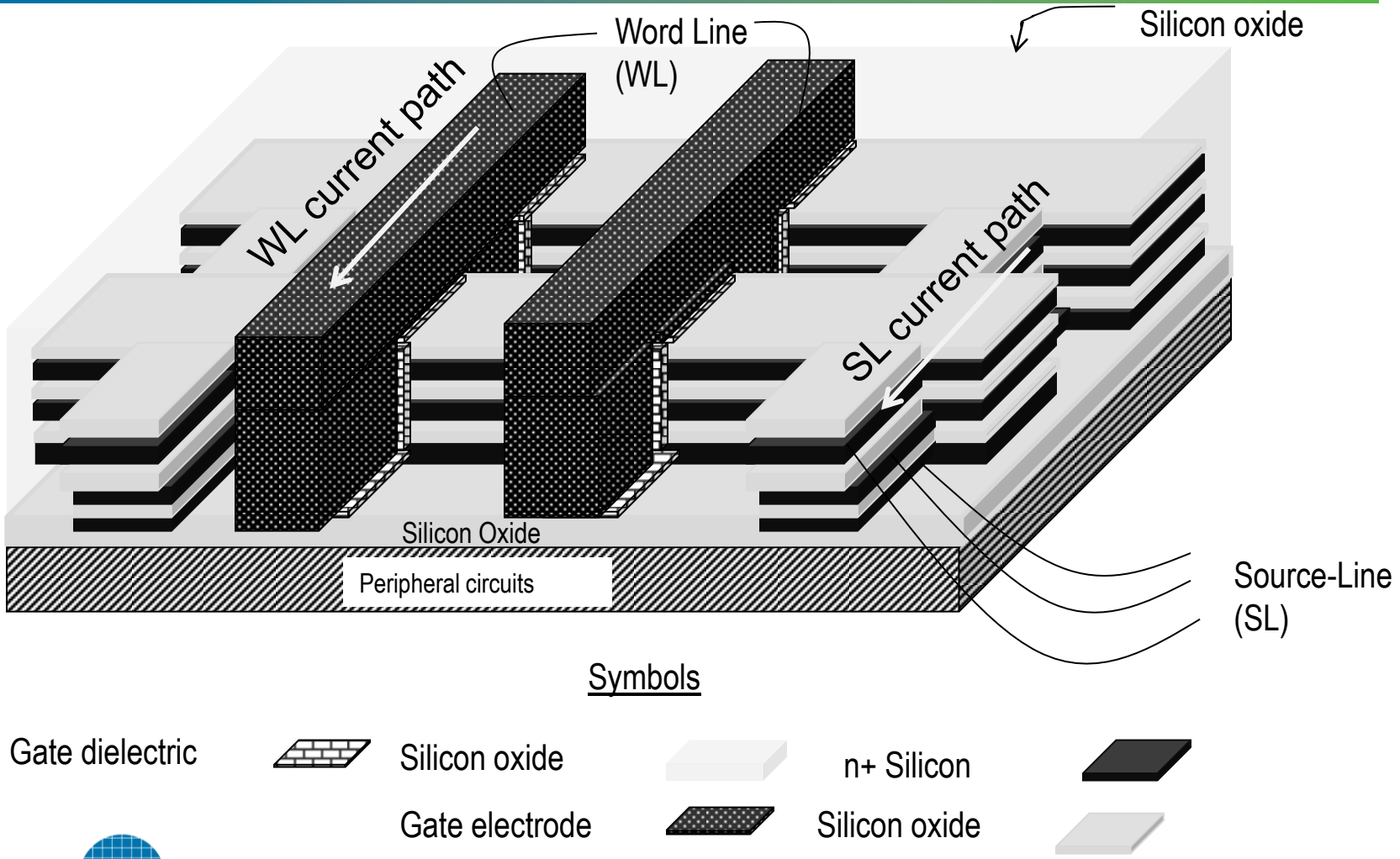
Silicon oxide

Gate electrode

Gate dielectric

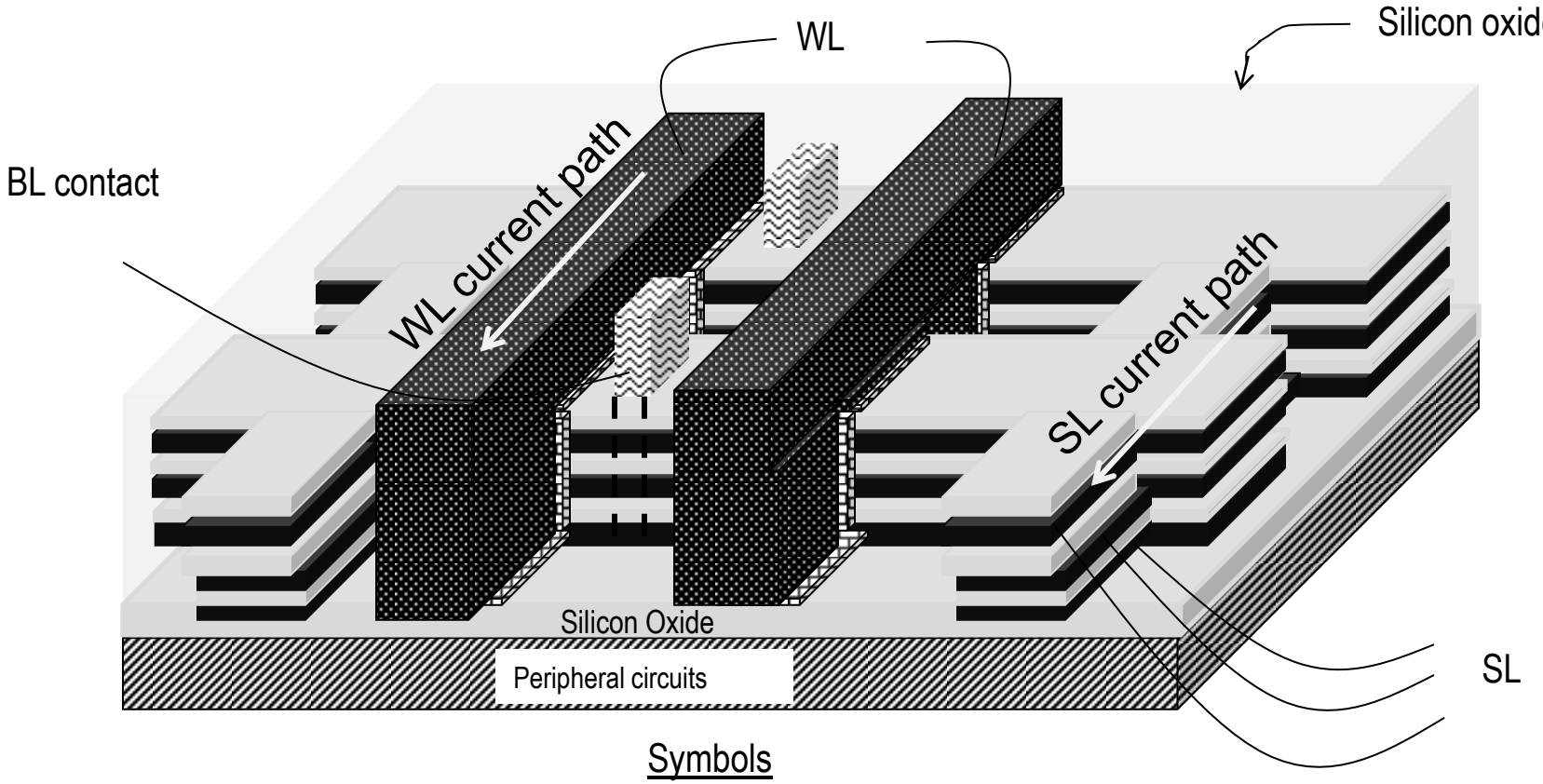
MonolithIC3D

Process Flow: Step 9
Deposit oxide, CMP. Oxide shown transparent for clarity.



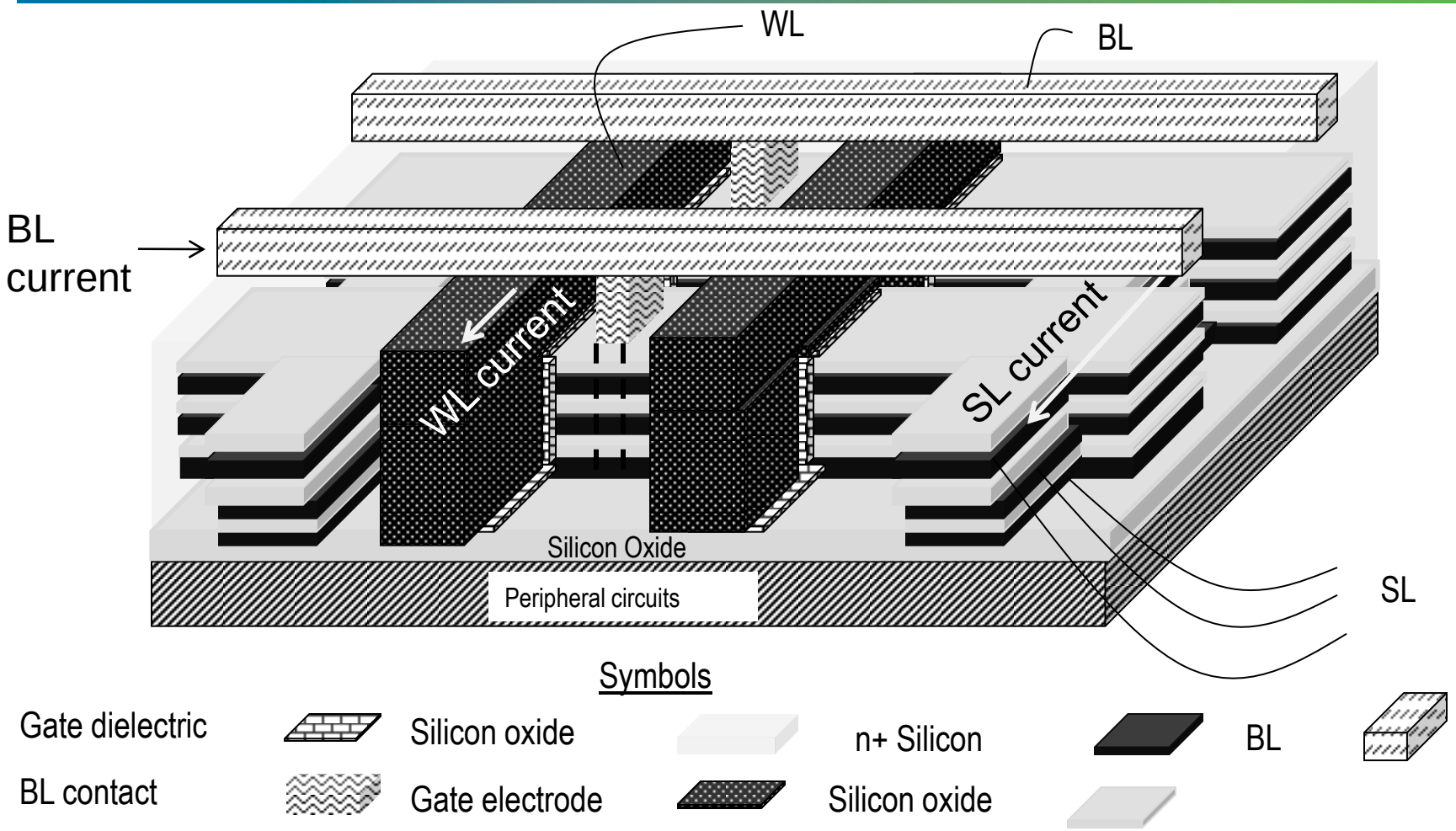
MonolithIC3D

Process Flow: Step 10
Make Bit Line (BL) contacts that are shared among various layers.



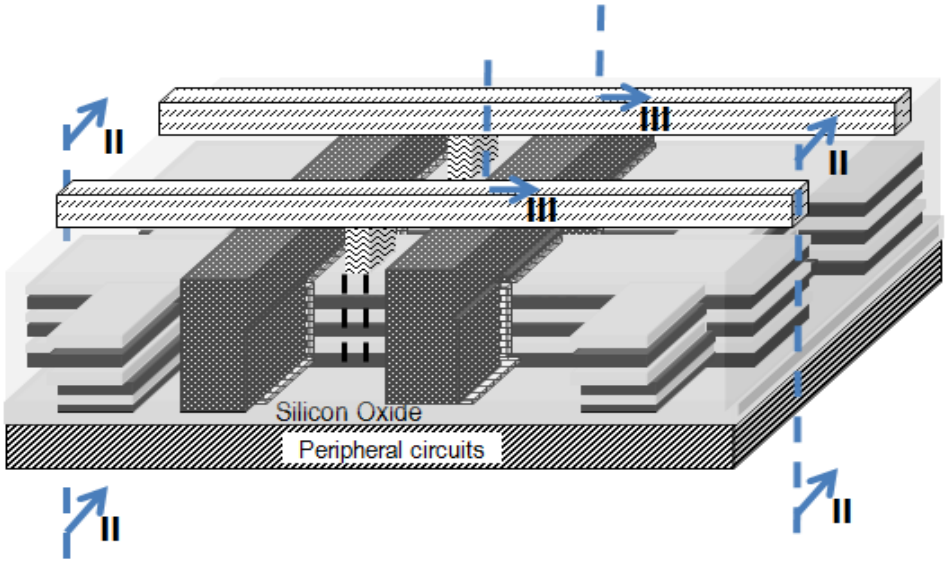
MonolithIC3D

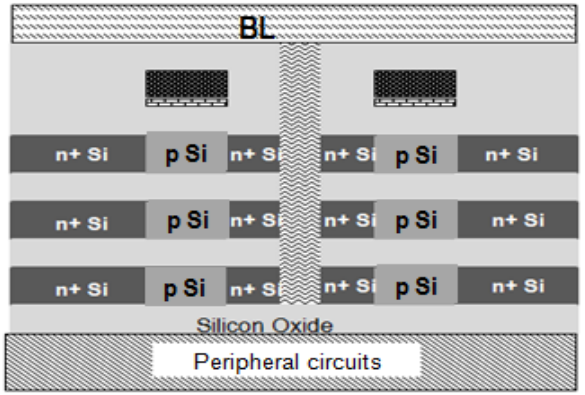
Process Flow: Step 11
Construct BLs, then contacts to BLs, WLs and SLs at edges of memory array using methods in [Tanaka, et al., VLSI 2007]



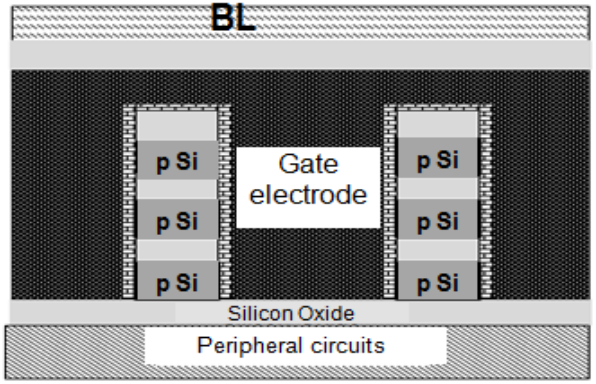
MonolithIC3D

Some cross-sectional views for clarity. Each floating-body cell has unique combination of BL, WL, SL





View along II plane



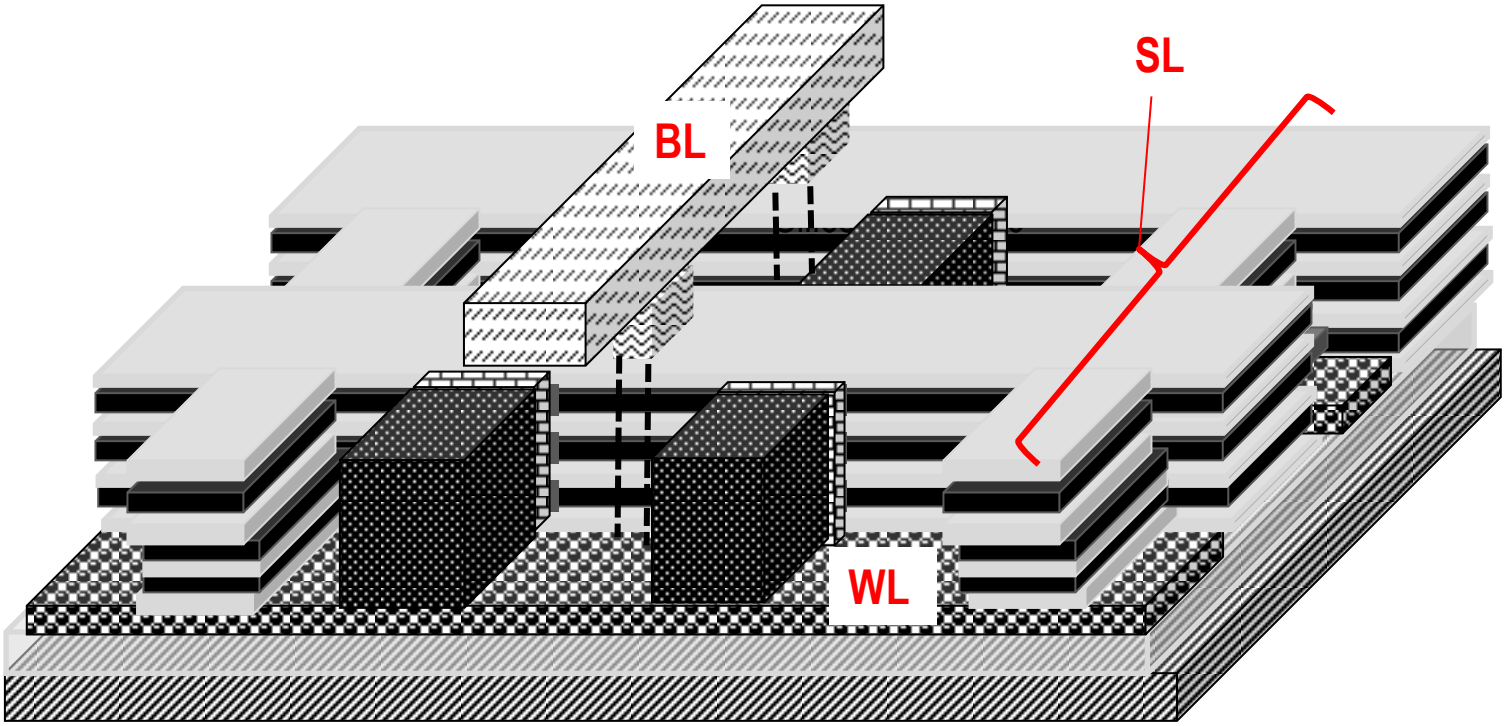
View along III plane

Gate dielectric
BL contact
n+ Silicon
Silicon oxide

Silicon oxide
Gate electrode
BL

MonolithIC3D

A different implementation:
With independent double gates



p Silicon
Silicon oxide
n+ Silicon
Periphery
WL wiring
Gate dielectric
Gate electrode
BL contact
BL

MonolithIC3D

Outline

- Implications and risks of the technology

Density estimation

	Conventional stacked capacitor DRAM	Monolithic 3D DRAM with 4 memory layers
Cell size	$6F^2$	Since non self-aligned, $7.2F^2$
Density	x	3.3x
Number of litho steps	26 (with 3 stacked cap. masks)	~26 (3 extra masks for memory layers, but no stacked cap. masks)

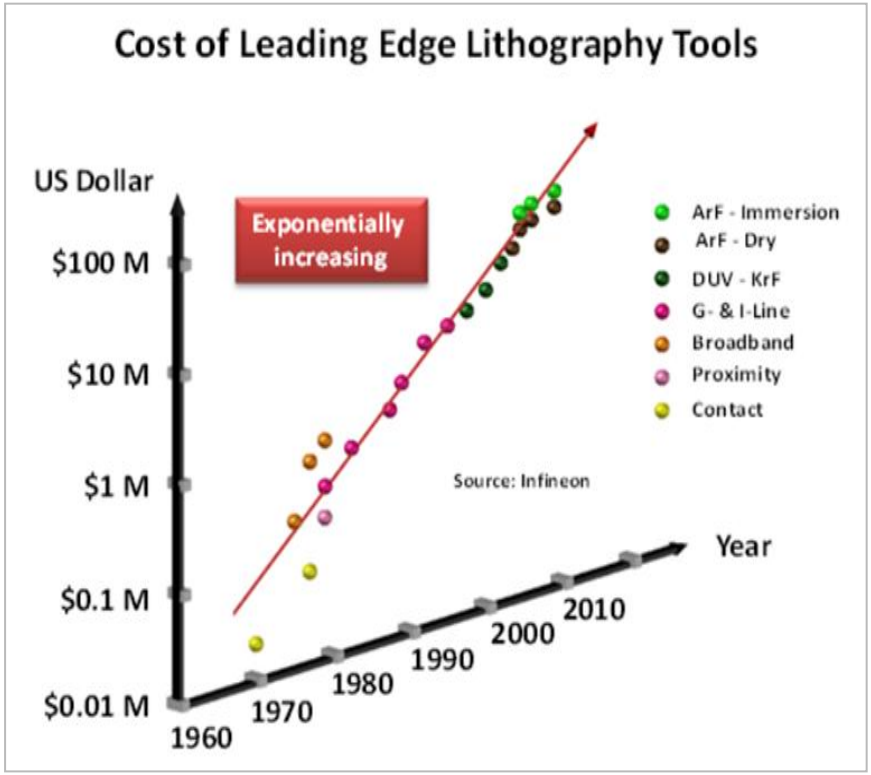
3.3x improvement in density vs. standard DRAM, but similar number of critical litho steps!!!

Negligible prior work in Monolithic 3D DRAM with shared litho steps, poly Si 3D doesn't work for DRAM (unlike NAND flash) due to leakage

Monolithic 3D

Scalability

- Multiple generations of cost per bit improvement possible
  (eg) 22nm 2D →
  22nm 3D 2 layers →
  22nm 3D 4 layers → ...
- Use same 22nm litho tools for 6+ years above. Tool value goes down 50% every 2 years → Cheap ☺
- Avoids cost + risk of next-gen litho



Reduces or avoids some difficulties with scaling-down

EUV delays and risk

(EETimes 2002)
"EUV to be in production in 2007"

(EETimes 2003)
"EUV to be leading candidate for the 32nm node in 2009"

(EETimes 2004)
"EUV to be pushed out to 2013"

(EETimes 2010)
"EUV late for 10nm node milestone in 2015"

Capacitor manufacturing

	45 nm	32 nm	22 nm	15 nm	10 nm
ε	40	50	60	65	70
AR	47	56	99	147	193

Continuous transistor updates

Planar → RCAT → S-RCAT → Finfet → Vertical devices

Risks

- **Floating-body RAM**

  Retention, reliability, smaller-size devices, etc

- **Cost of ion-cut**

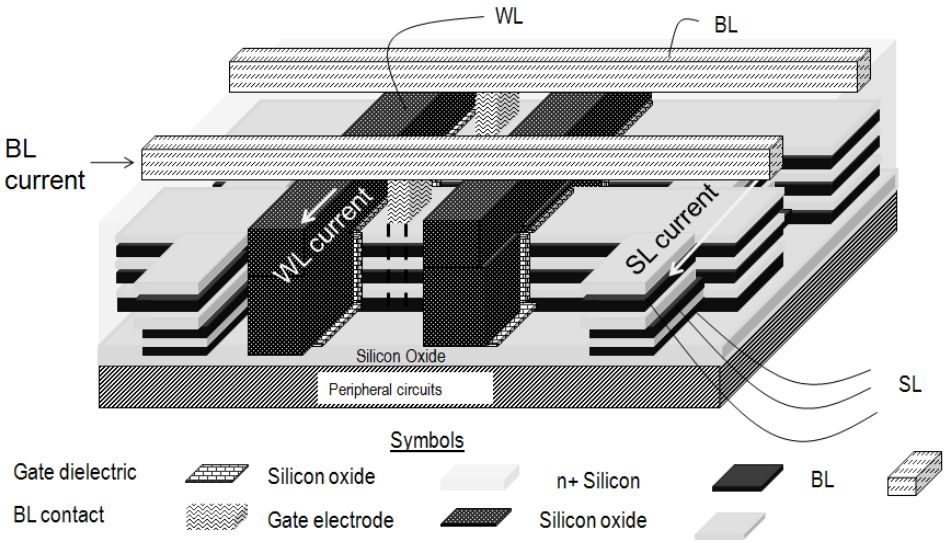
  Supposed to be <$50-75 per layer since one implant, bond, cleave, CMP step.

  But might require optimization to reach this value.

Outline

- **Summary**

Summary of Monolithic 3D DRAM Technology



Monolithic 3D with shared litho steps

Single crystal Si

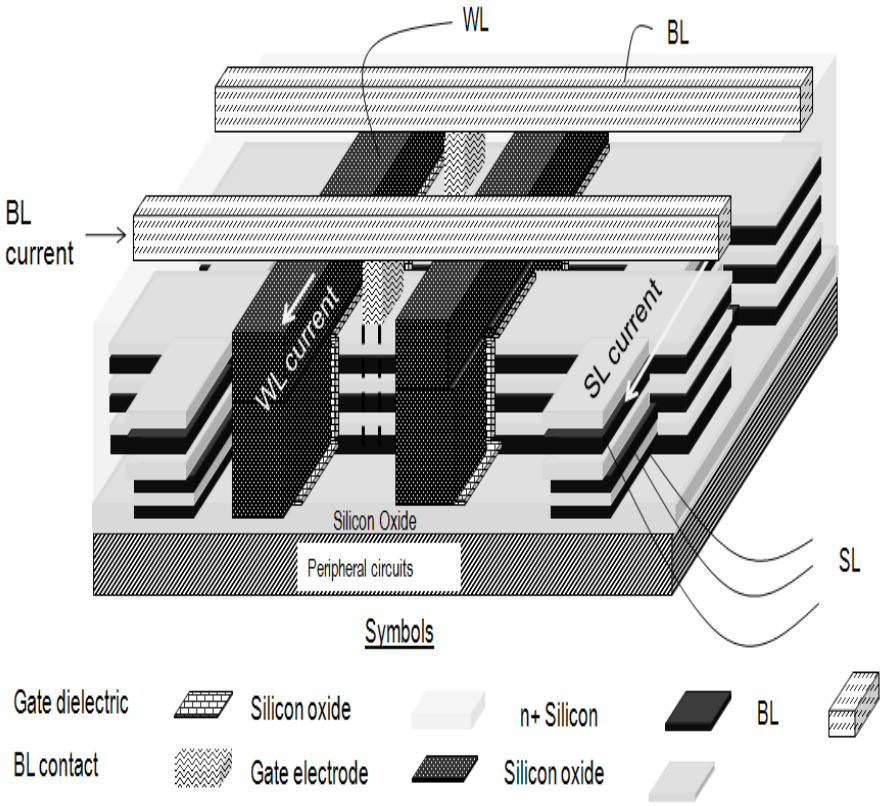
Floating body RAM

Under development...

- 3.3x density of conventional DRAM, but similar number of litho steps
- Scalable (eg) 22nm 2D → 22nm 3D 2 layers → 22nm 3D 4 layers → ...
- Cheap depreciated tools, less litho cost + risk, avoids many cap. & transistor upgrades
- Risks = Floating body RAM, ion-cut cost

Backup slides

A note on overlay

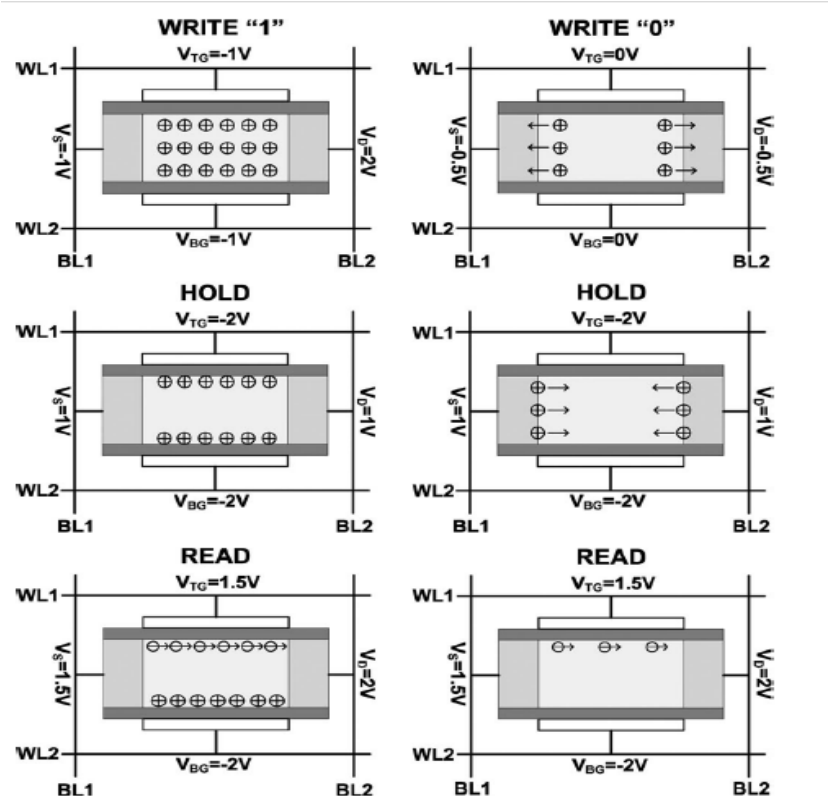


- Implant n+ in p Si regions layer-by-layer, then form gate
  → non self-aligned process
- ITRS <20% overlay requirement
  ASML 1950i = 3.5nm overlay for 38nm printing. <10% overlay.
- So, gate length = 1.2F. Penalty of 0.2F for non-self-aligned process

MonolithIC3D™

Bias schemes for floating body RAM

Bipolar Mode [S. Alam, et al, TED 2010]



MOS Mode [Intel, IEDM 2006]

	V_D	V_{FG}	V_{BG}	V_S
Program	2	1.5	-1.5	0
Erase	-1	0.7	-1.5	0
Read	0.2	0.7	-1.5	0
Hold	0	-1	-1.5	0

Contact processing with shared litho steps

- Similar to Toshiba BiCS scheme [VLSI 2007]

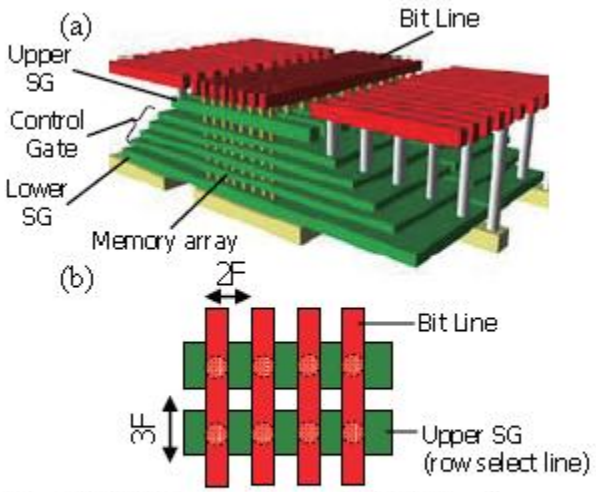


Fig. 3 (a) Birds-eye view of BiCS flash memory, (b) Top down view of BiCS flash memory array.