Low Power Techniques

Keshava Murali
shavakmm@gmail.com
http://asic-soc.blogspot.com

Increasing Challenges of Power

- Increasing device densities
- Increasing clock frequencies
- Lowering supply voltage
- Lowering transistor threshold voltage

High power consumption➔higher temperature➔heat sinks, ceramic packaging (expensive)

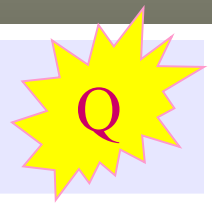
Power Management

- Manage power in all modes of operation
- Dynamic power during device operation, Static power during standby
- Maintain device performance while minimizing power consumption
- Performance available when required
- Power minimized while providing required performance

Require power awareness in every stage of design cycle

Power architecture	Power aware design	Power aware implementation	Power aware verification
-define spec -refine power architecture -determine which technique -identify power interdependencies among blocks	-capture RTL based on power requirement	-Libraries with power models -Special cells -power aware logic synthesis -power aware physical synthesis -Achieve best power, timing and QoR	-Voltage becomes functional -coverage metrics for low power methods -verification for different power modes

Is it possible to have single specification of power intent???



Power Has Broken the Rules of Scaling

Source: ITRS 2005	90nm	65nm	45nm
Device Length (nm) ⇩	1x	0.7x	0.5x
Frequency (GHz) ⇧	1x	1.43x	2x
Integration Capacity (BT) ⇧	1x	2x	4x
Voltage (V) ⇘	**1x**	**0.85x**	**0.75x**
V_{TH} (V) ⇘	**1x**	**0.85x**	**0.75x**
I_{OFF} (nA/um) ⇧⇧	**1x**	**~3x**	**~9x**
Dynamic Power (W) ⇘	**1x**	**>0.7x**	**>0.5x**
Dynamic Power Density (W/cm²) ⇧	**1x**	**> 1.43x**	**> 2x**
Leakage Power Density (W/cm²) ⇧	**1x**	**~2.5x**	**~6.5x**
Power Density (W/cm²) ⇧	**1x**	**~2x**	**~4x**

Cadence Design Systems Inc. estimates that 90-nm standard transistors are about 40 times leakier than the standard-voltage 130-nm transistors

Types of Power Consumption

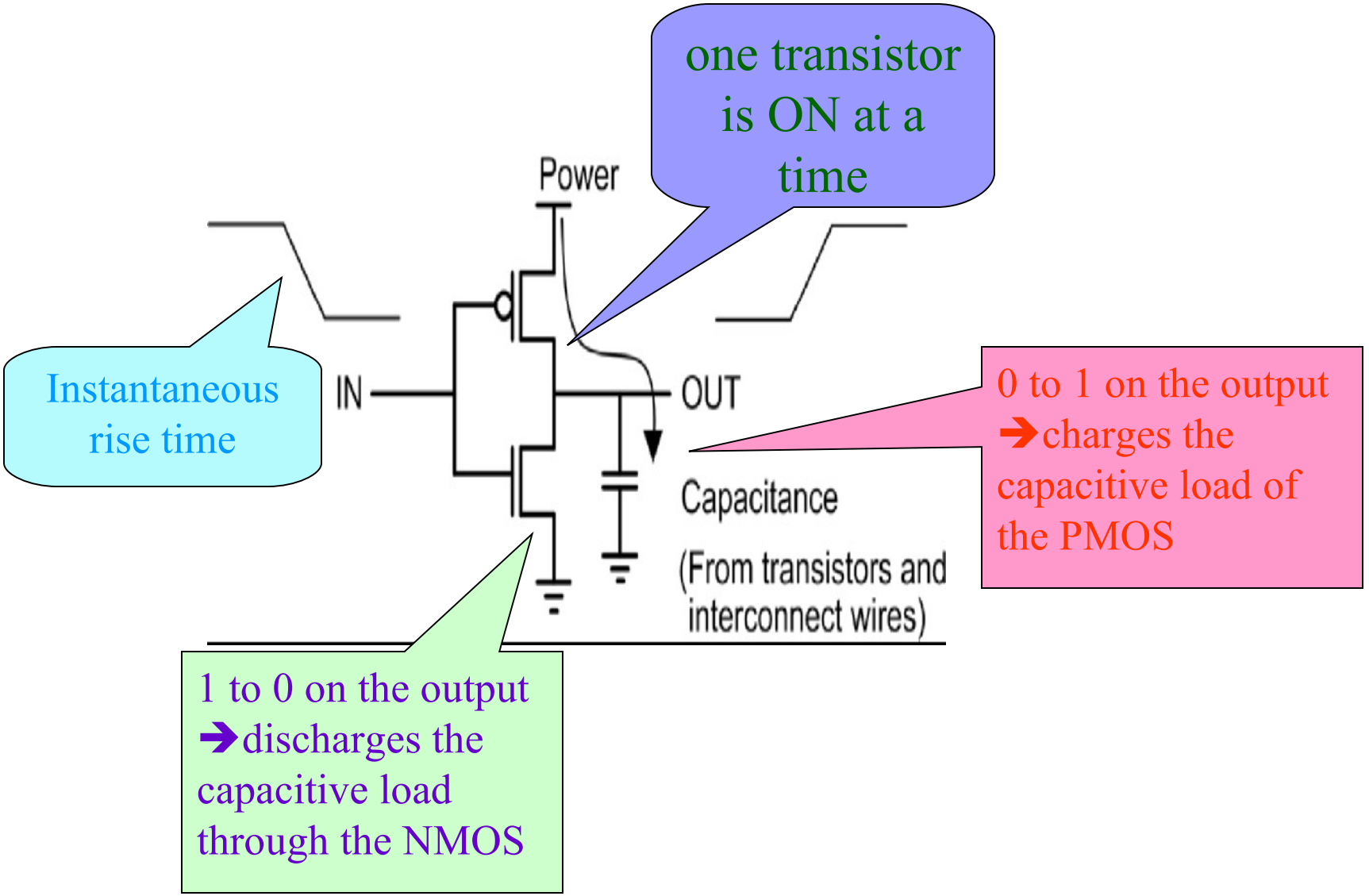
Dynamic power

- During the switching of transistors
- Depends on the clock frequency and switching activity
- Consists of switching power and internal power.

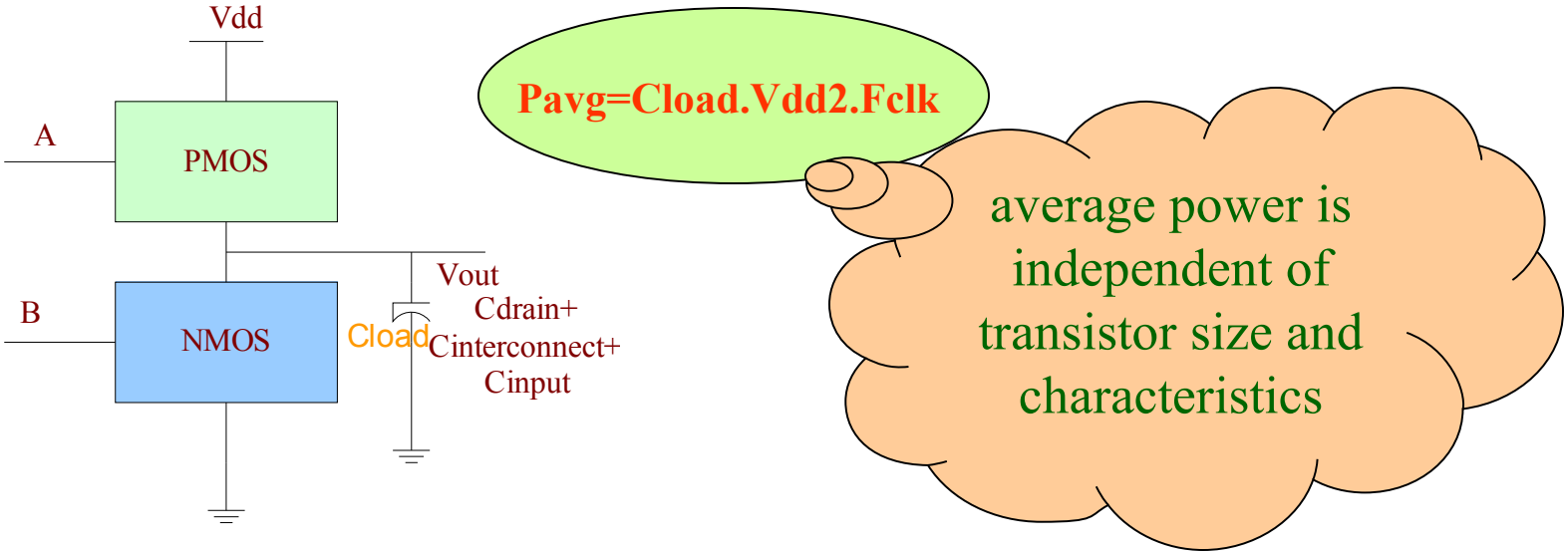
Static Power

- Transistor leakage current that flows whenever power is applied to the device
- Independent of the clock frequency or switching activity.

Dynamic Power



Dynamic Power Contd....



Cload depends on:

- **Output node capacitance of the logic gate:** due to the drain diffusion region.
- **Total interconnects capacitance:** has higher effect as technology node shrinks.
- **Input node capacitance of the driven gate:** due to the gate oxide capacitance.

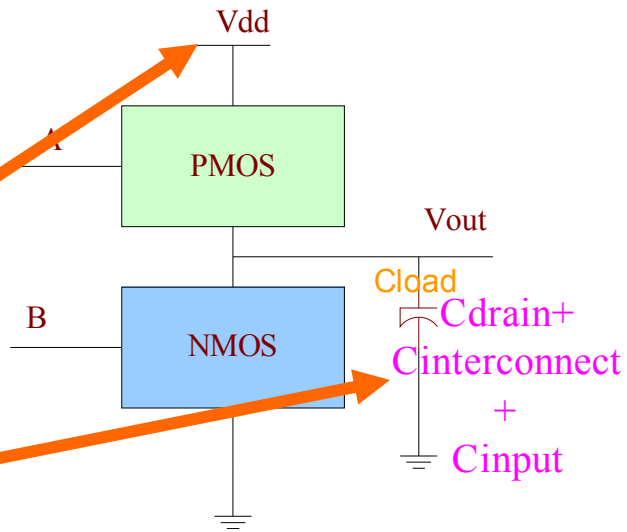
Internal power

- Power consumed by the cell when an input changes, but output does not change
- Internal node voltage swing can be only Vi which can be smaller than the full voltage swing of Vdd leading to the partial voltage swing.

How to reduce dynamic power?

Pavg α Cload.Vdd2.Fclk

Reduce Vdd
Reduce Cload
Reduce Fclk



Short Circuit Power

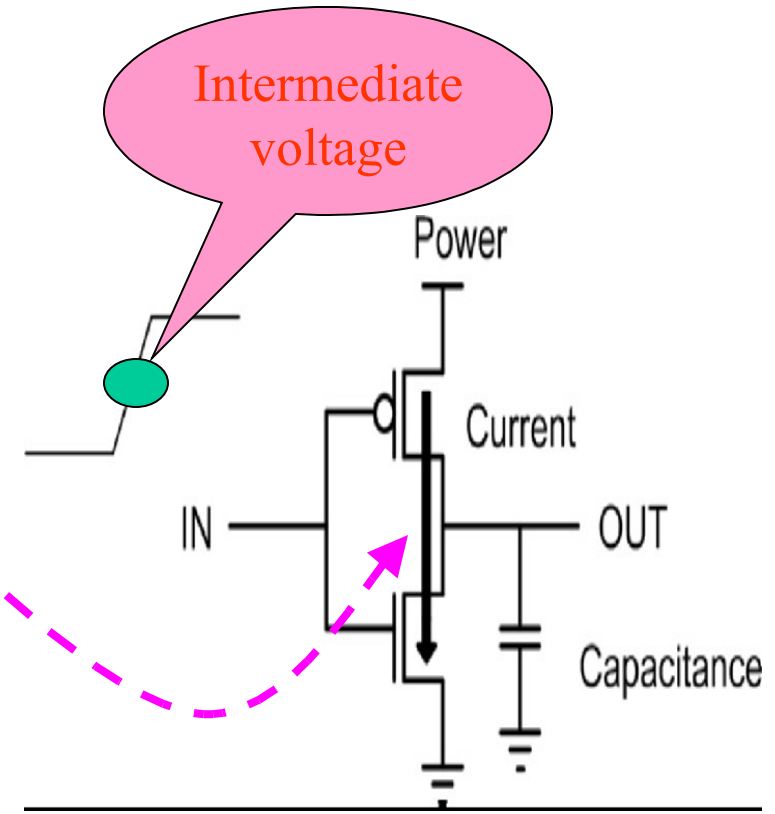
Finite rise and fall time

VTn < Vin < Vdd - |VTp|

➔Both PMOS and NMOS are conducting for a short duration of time

➔short between supply power and ground

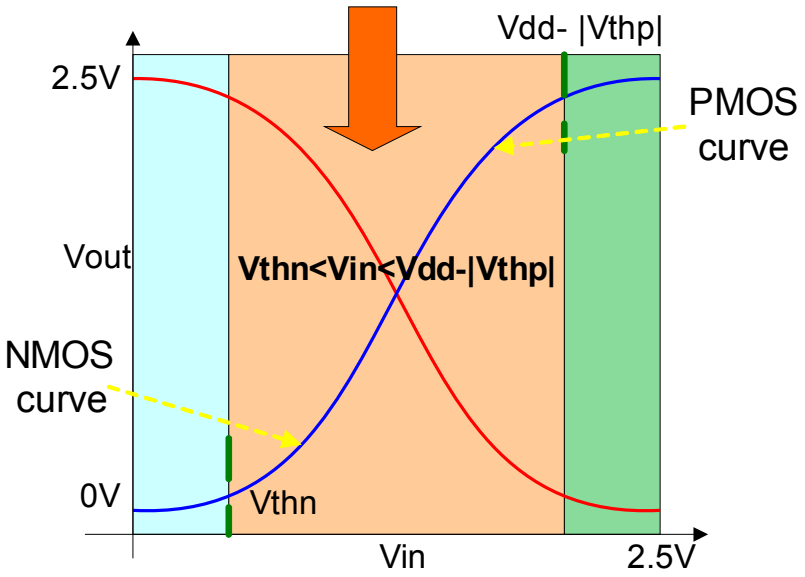
➢Lower threshold voltages and slower transitions result in more internal power consumption.



Short Circuit Power-Analysis

Condition	PMOS	NMOS
Vin < Vth	ON (sat)	OFF (cutoff)
Vin = Vth	Linear (towards cutoff)	Linear (towards sat)
Vin > Vth	OFF (cutoff)	ON (sat)

More rise/fall time➔more short circuit
Lower threshold voltage➔more short circuit



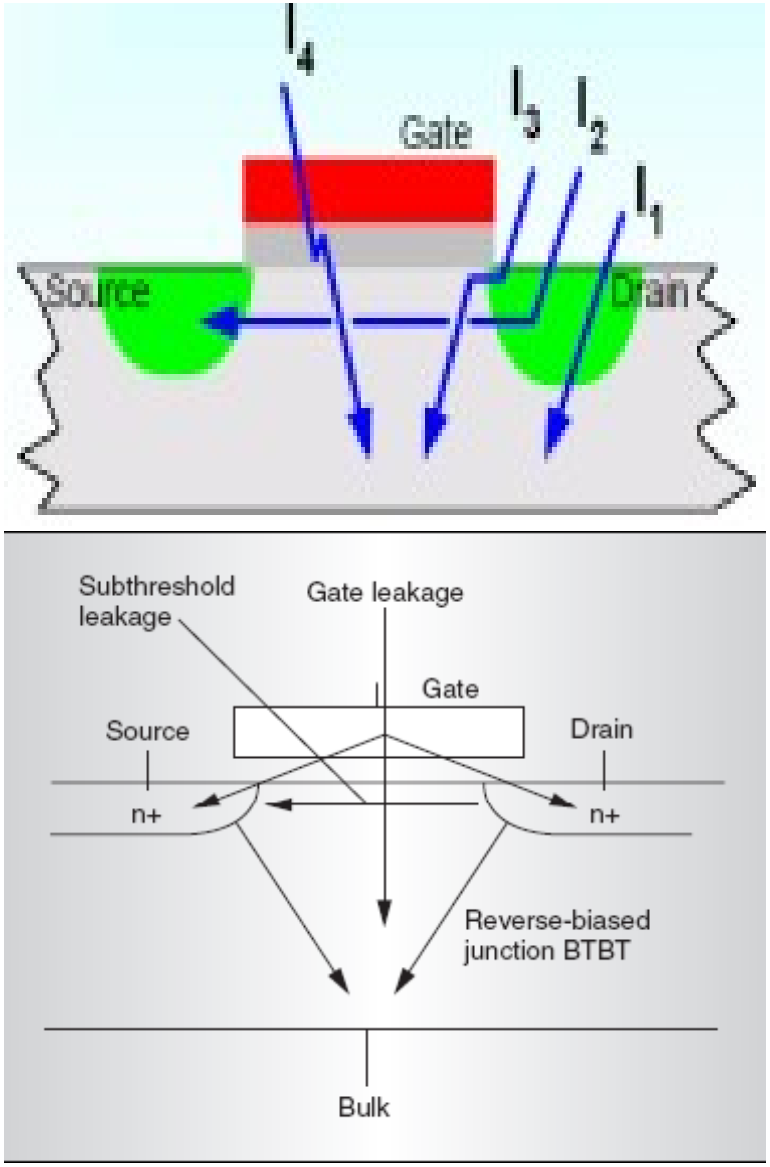
$$Pavg(short\ circuit) = 1/12.k.\tau.Fclk.(Vdd-2Vt)3$$

To get equal rise/fall ➔ balance transistor sizing

If Vdd<Vthn+|Vthp| can we eliminate short circuit current?????

Leakage Power

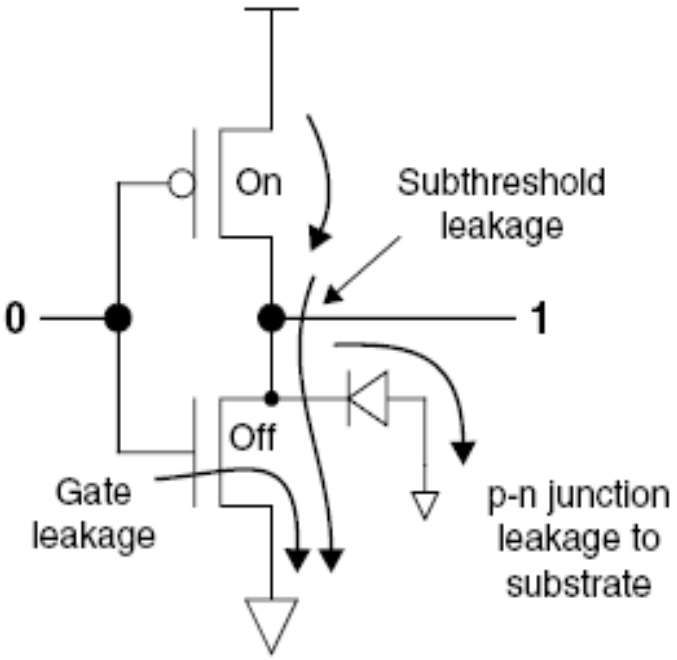
- Diode reverse bias current or Reverse-biased, drain- and source-substrate junction band-to-band-tunneling (BTBT) –I1
- Sub threshold current – I2
- Gate induced drain leakage – I3
- Gate oxide tunneling – I4

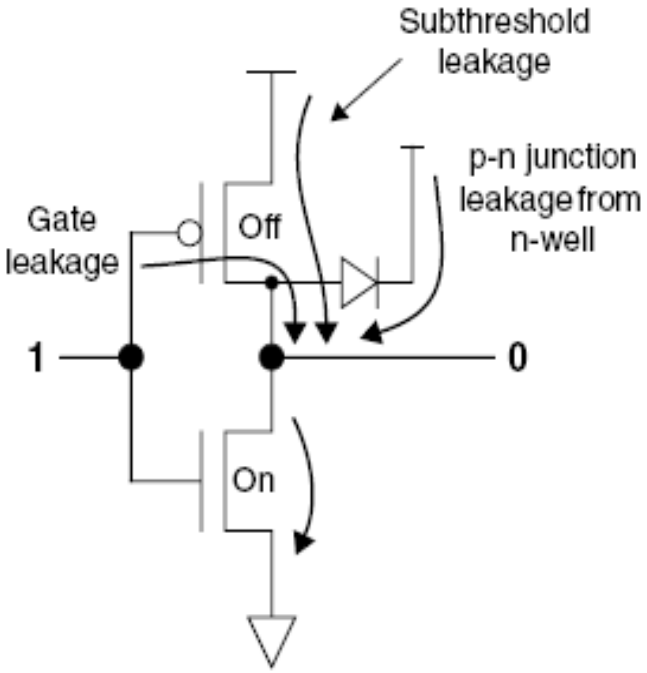


Leakage Power Contd....

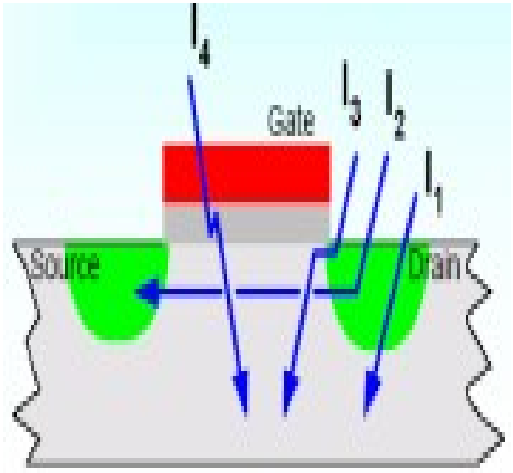
-does not depend on input transition, load capacitance

-remains constant





Reverse Biased Diode Current (Junction Leakage)-I1



Parasitic diodes formed between the diffusion region of the transistor and substrate

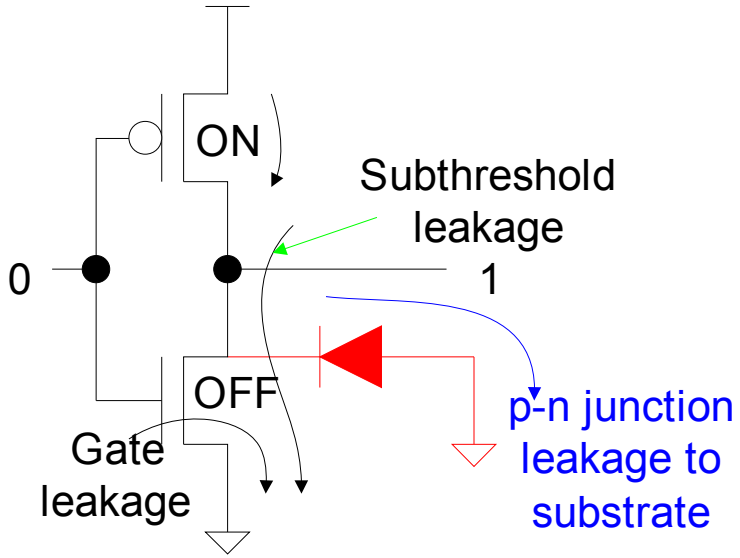
Ireverse=A.Js.(e(q.Vbias/kT)-1)
where,
Vbias --> reverse bias voltage across the junction
Js --> reverse saturartion current density
A --> junction area

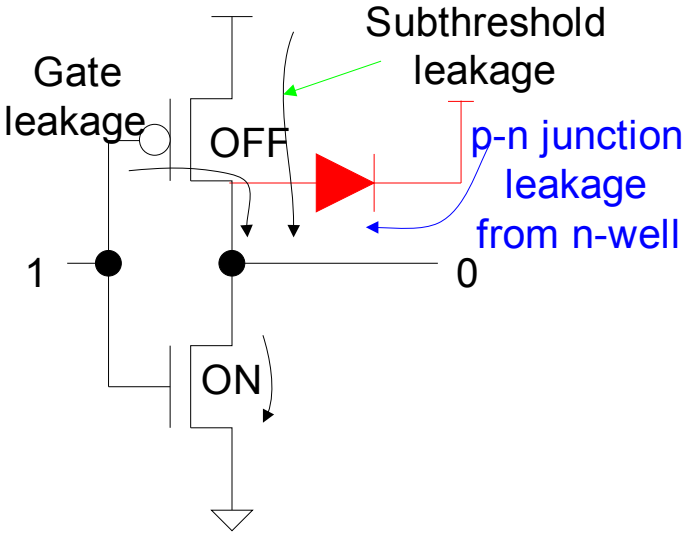
How to reduce?

Decrease in junction area ➔ depends material

Can we adjust* Vbias *to control junction leakage? Q

Reverse Biased Diode Current (Junction Leakage)-I1 Contd…

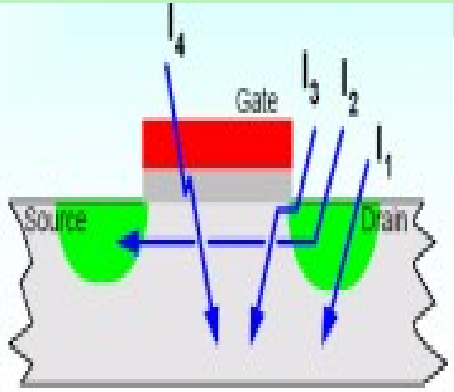


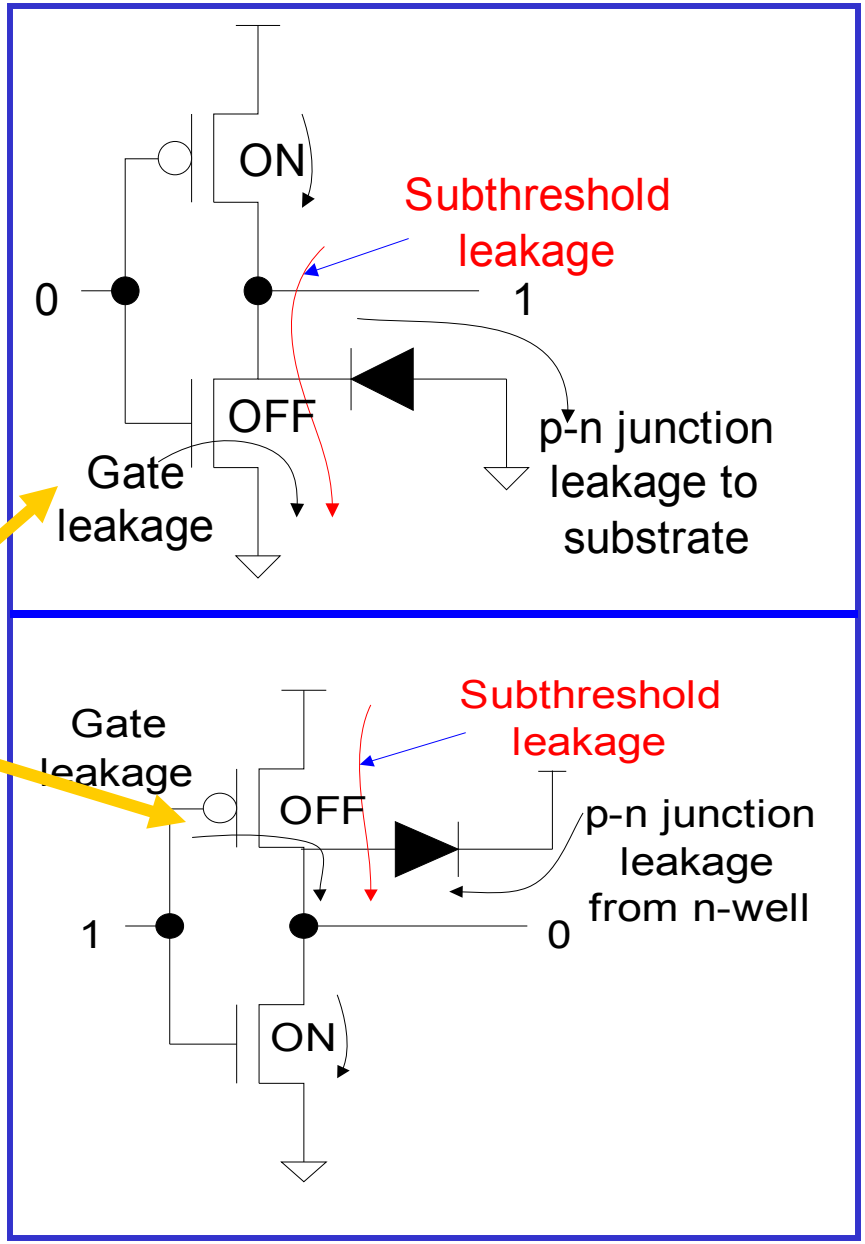


Sub threshold Current – I2 (Isub)

- Always flows from source to drain
- **Vgs <~ Vth** → carrier diffusion causes sub threshold leakage

- Vgs<=0 → accumulation mode
- 0<Vgs<<Vth → depletion mode
- **Vgs~Vth → weak inversion**
- Vgs>Vth → Inversion





How to reduce sub threshold leakage?

$$I_{sub} = I_0 \cdot e^{\frac{q}{nk_BT}(V_{gs} - V_{th0} + \gamma V_{bs} + \eta V_{ds})} \left(1 - e^{\frac{-qV_{ds}}{k_BT}}\right) \quad (1)$$

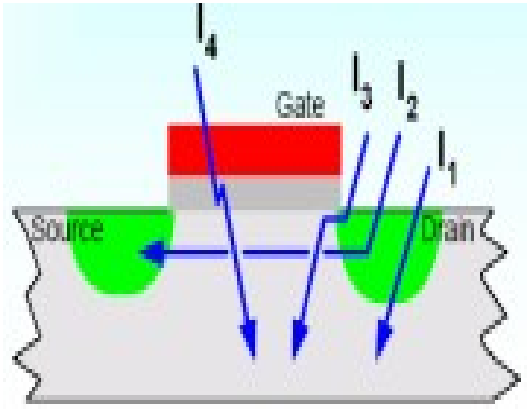
Isub exponentially scales with Vth ➔vary Vth

- Higher Vth results in lower leakage, longer delay
- Optimize the design with the balance application of low Vth (LVT) and high Vth devices (HVT).
- Older technologies - more threshold variation
- Newer technologies produce around 30 mV threshold variation

Does this equation valid below 90nm????

Gate Induced Drain Leakage (GIDL) - I3

- Caused by high field effect in the drain junction of MOS transistors



When Vgs <= 0V; Vd = Vdd

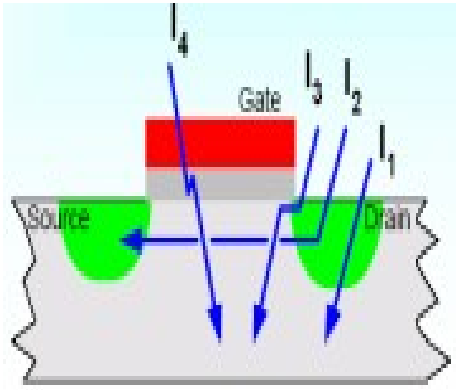
➔avalanche multiplication and band-to-band tunneling

➔Minority carriers underneath the gate are swept to the substrate

GIDL increases with:

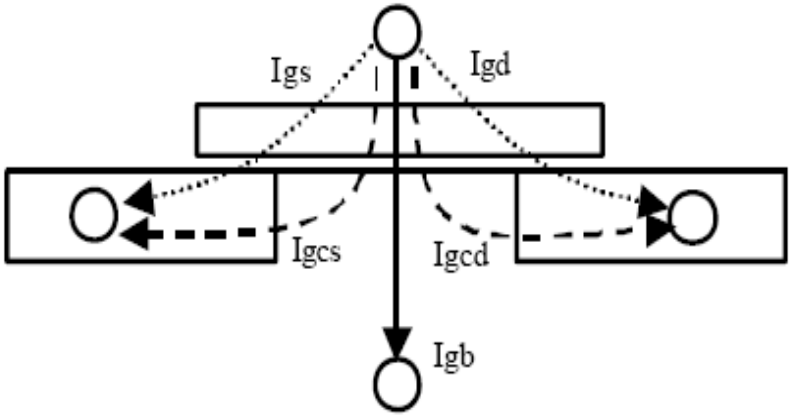
- Higher supply voltage
- thinner oxide
- increase in Vdb and Vdg.

Gate Oxide Tunnelling - I4



- Due to high electric field across a thin gate oxide
- Fowler-Nordheim tunneling: conduction band of the oxide layer

- Direct tunneling through the silicon oxide layer if it is less than 3–4 nm thick



How to reduce gate leakage?

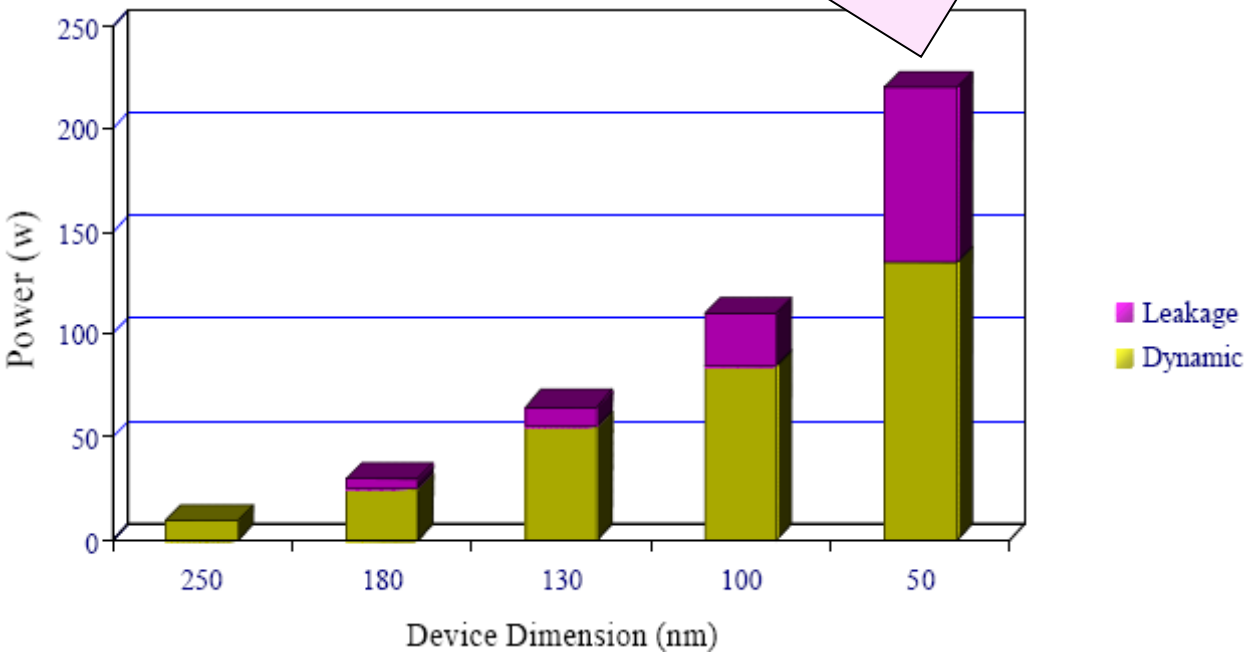
Improve fab chemistry

Reached fundamental limit of gate oxide thickness????

Leakage Power Trends

Scaling: Boon or Curse???
Should be done for Voltage and Threshold voltage to gain the performance

- leakage current increases exponentially.
- Leakage power is catching up with Dynamic Power.

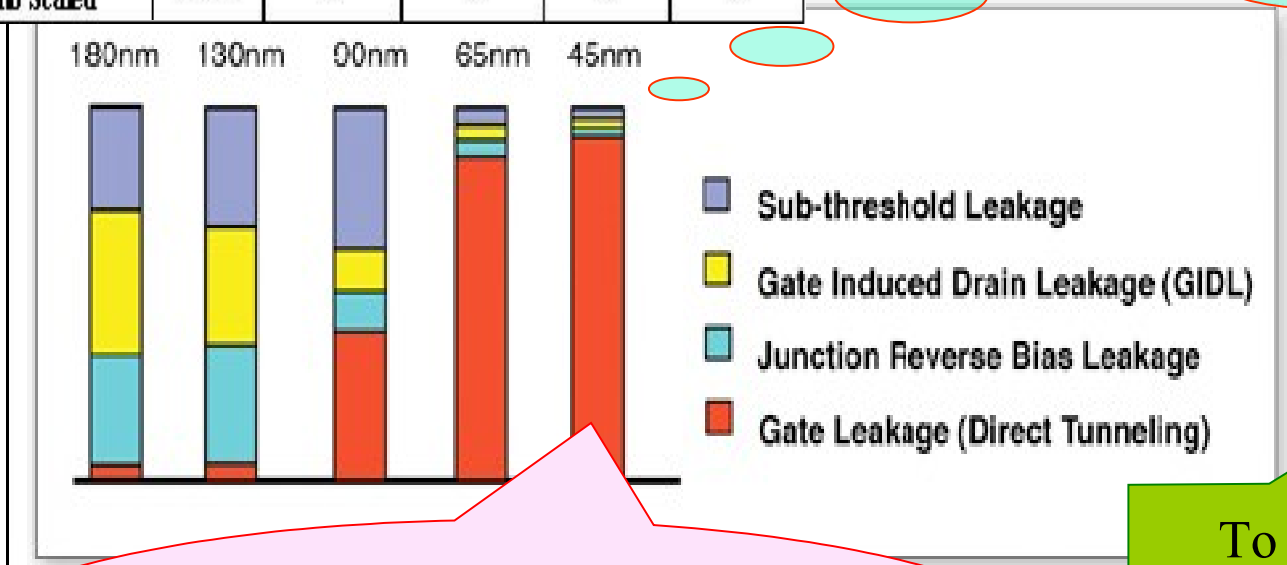


- At 90 nm and below, leakage power management is essential.
- Thinner gate oxides have led to an increase in gate leakage current.

Technology shrinking vs Leakage components

	L (μm)	Tox(μm)	Isub	Igate	Ijunc
Long Channel	>1	>3	x	x	x
Short Channel	>0.18	>3	Y	x	x
Very Short Channel	>0.090	>2	Y	Y	x
Nano Scaled	<0.090	<2	Y	Y	Y

Other leakages are low due to improvements in the fabrication process and material.



45 nm and below==>increased electric field==>increased gate leakage

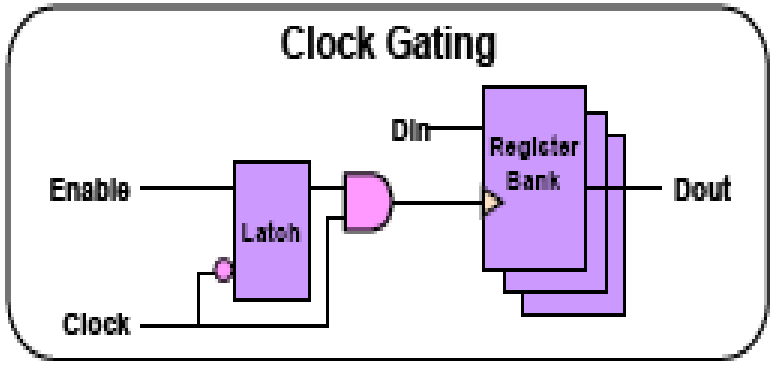
To counteract this voltage is scaled down to around 1V

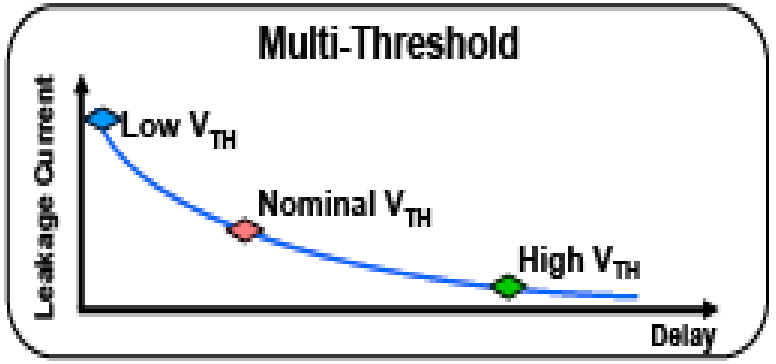
Low Power Design Techniques

Dynamic Power	Leakage Power	Design	Architectural	Process Technology
Clock gating	Multi Vt	Multi Vt	Pipelining	Multi Vt
Variable frequency	Power gating	Clock gating	Asynchronous	PD SOI
Variable power supply	Back (substrate) bias	Power gating		FD SOI
Multi Vdd	Use new devices- FinFet, SOI	Multi Vdd		FinFet
Voltage islands		DVFS		Body Bias
DVFS				Multi oxide devices
				Minimize capacitance by custom design

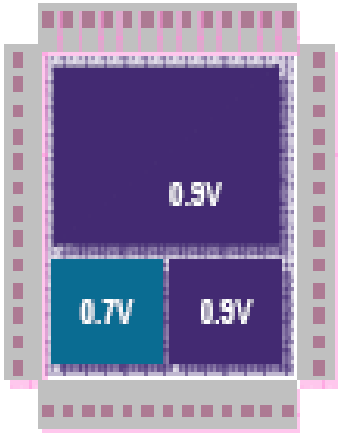
Low Power Design Techniques

Basic techniques

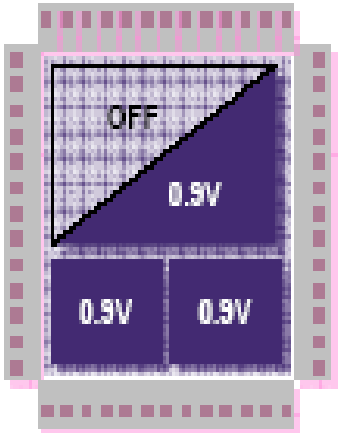




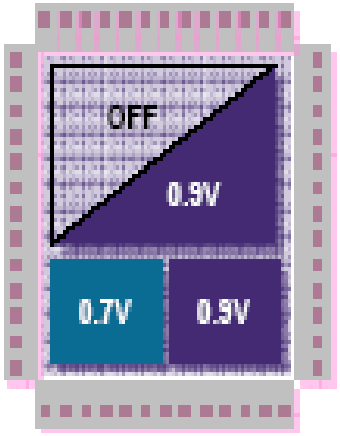
Advanced techniques



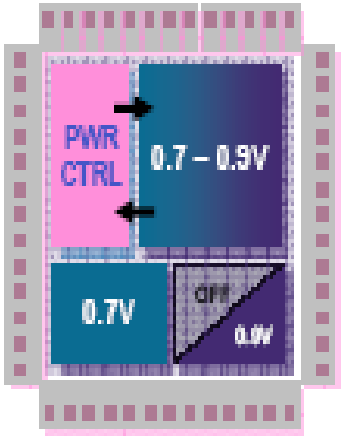
Multi-Voltage (MV)



MTCMOS power gating (shut down)

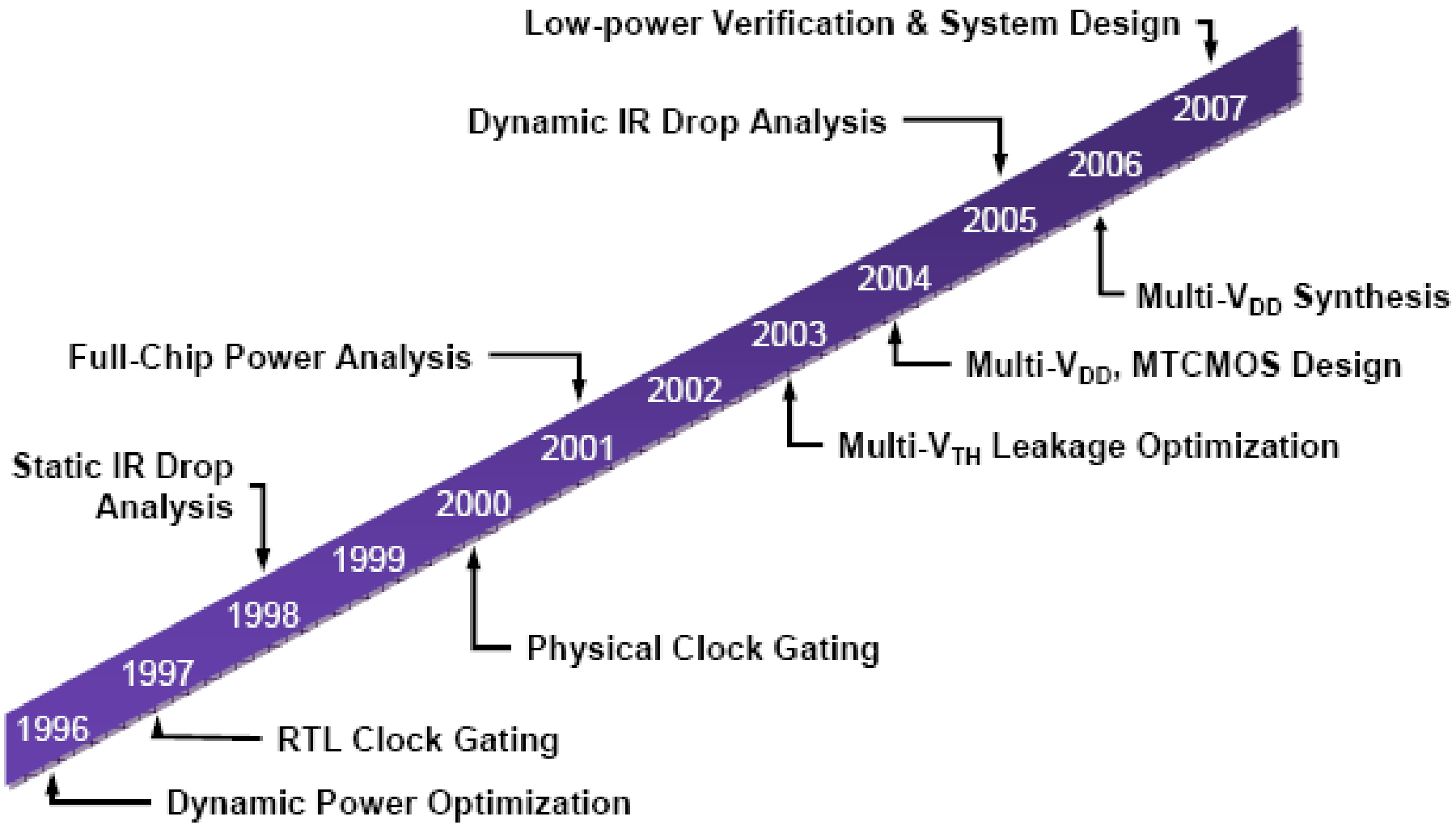


MV with power gating



Dynamic Voltage Frequency Scaling (DVFS)

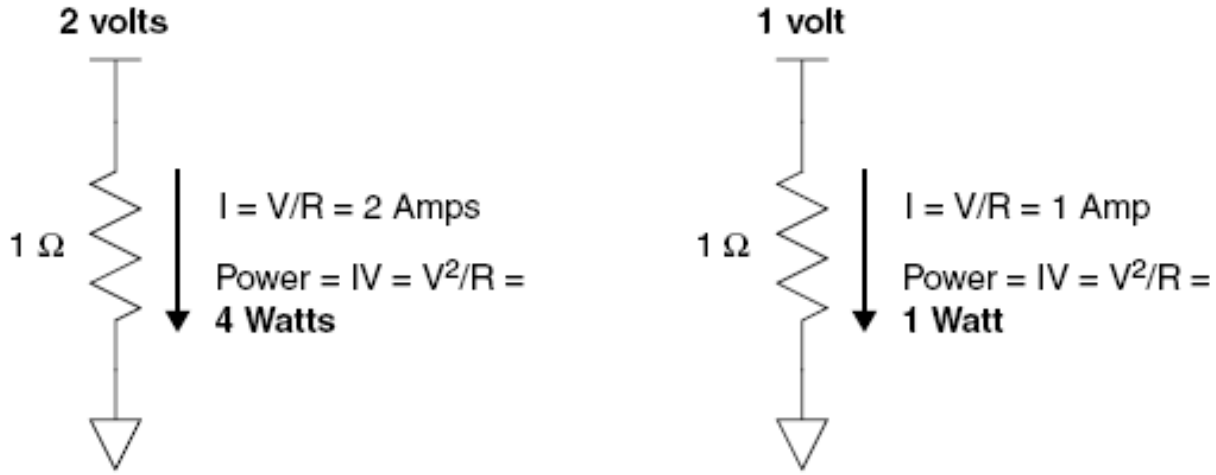
Evolution of low power techniques



Source: SNUG 2007

Supply Voltage Reduction - Voltage Scaling

- Scale both Vdd and Vth to maintain performance
- Quadratic reduction in supply voltage==>cubic reduction of power
- This equation deviates when Vdd reaches sub threshold voltage level i.e. Vdd ~ Vth
- Dynamic power reduction decreases... sub threshold leakage increases
- ==>**puts limit on scaling** !!!! Don't expect any more rigorous scaling !!!!!!!



Clock Gating

- Clock tree consume more than 50 % of dynamic power
- Turn off the clock when it is not needed
- Gate the clocks of flops which have common enable signal

The components of this power are:

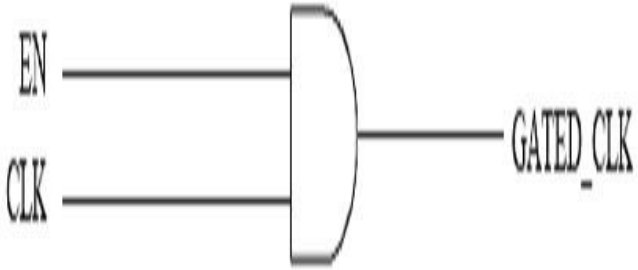
- Power consumed by combinatorial logic whose values are changing on each clock edge
- Power consumed by flip-flops and
- The power consumed by the clock buffer tree in the design.

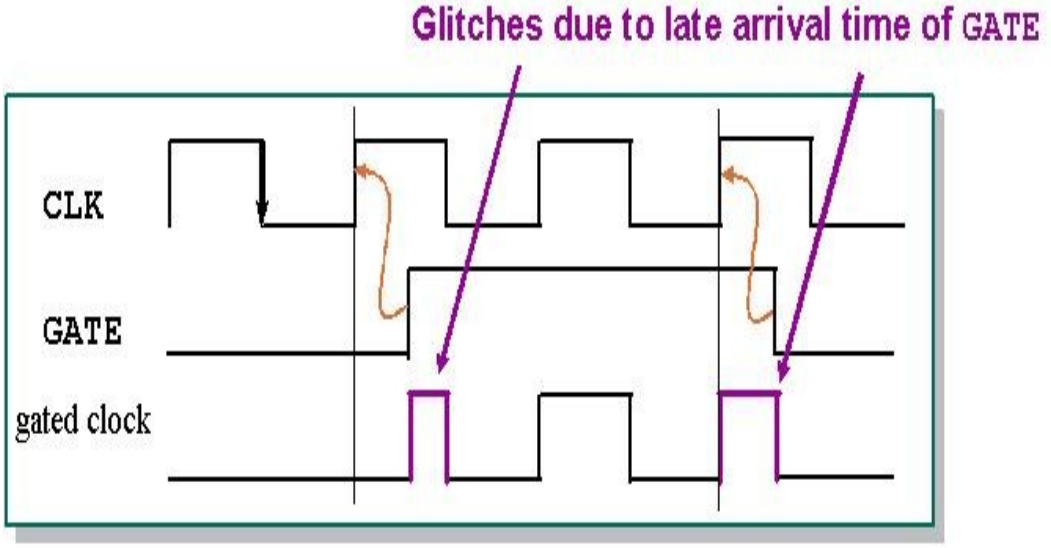
There are two types of clock gating styles available. They are:

1) Latch-based clock gating
2) Latch-free clock gating.

Latch free clock gating

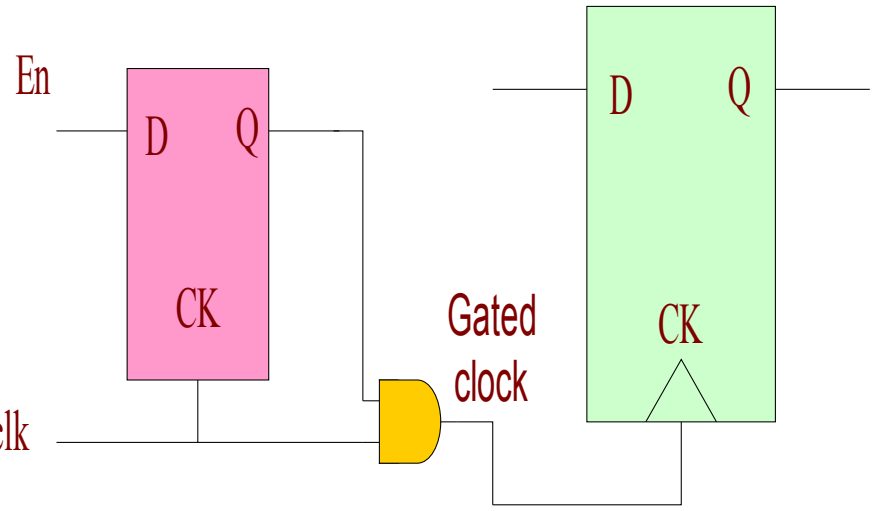
- Uses a simple AND or OR gate
- Glitches are inevitable
- Less used





Latch based clock gating

- Adds a level-sensitive latch
- Holds the enable signal from the active edge of the clock until the inactive edge of the clock
- Less glitch
- ICG-Integrated Clock Gating cells
- Easy adoption by EDA tools
- No change/modification required



Multi Threshold (MVT) technique

- Use both LVT and HVT cells.
- LVT gates on critical path while HVT gates off the critical path.
- Footprint and area of low Vt and high Vt cells are same as that of nominal Vt cells.
- Multi Vt optimization is placement non disturbing.
- This enables swapping of cells.
- Increase fabrication complexity

Different multi vt flows

One (single) pass flow

Compile with a set of libraries

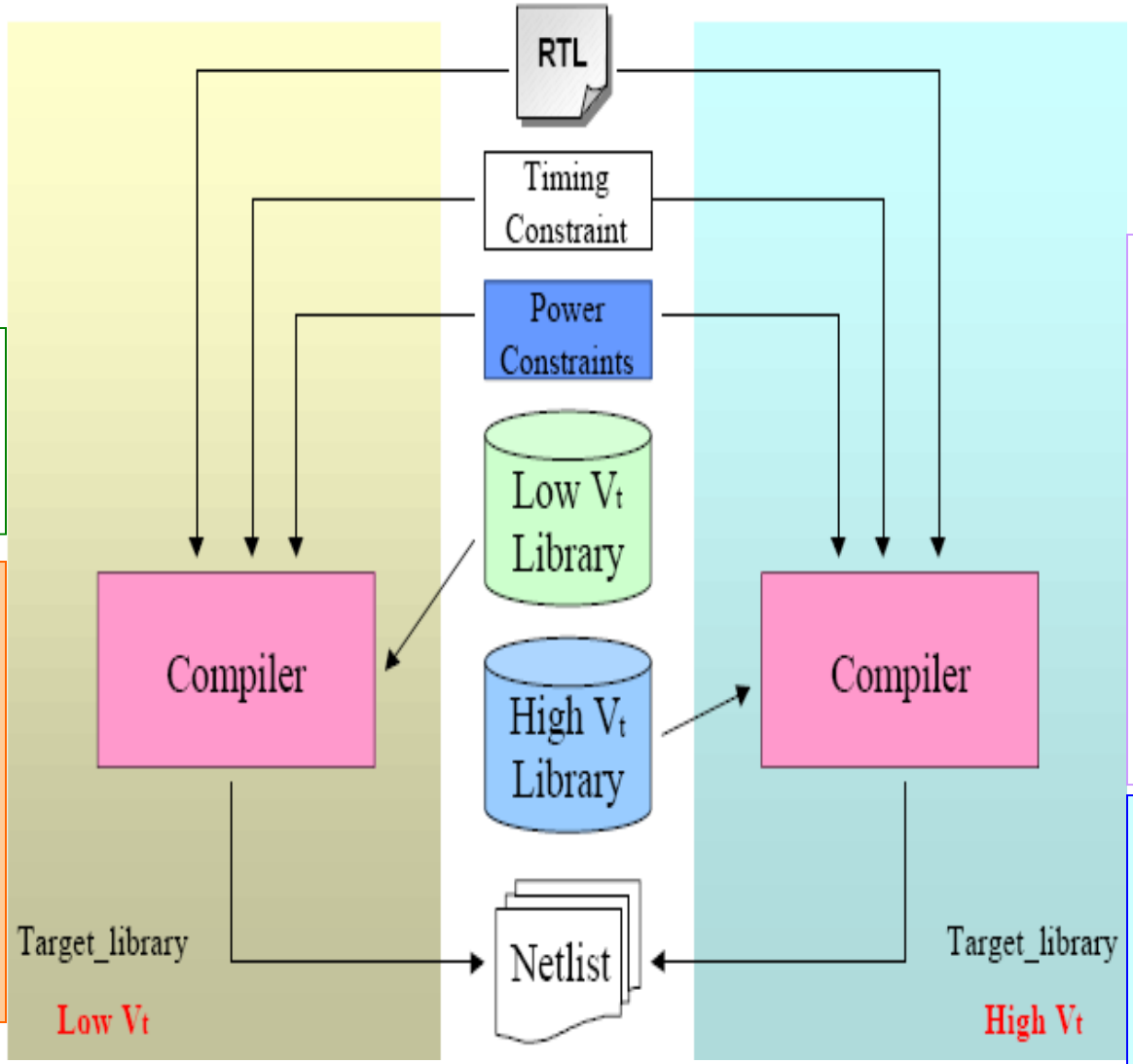
-Rvt-Lvt

-Rvt-Hvt

-Multi Vt

-Lvt

-Hvt



Two pass flow

-Compile with a set of libraries

-Incremental compilation with another set of libraries

-Lvt to Mvt

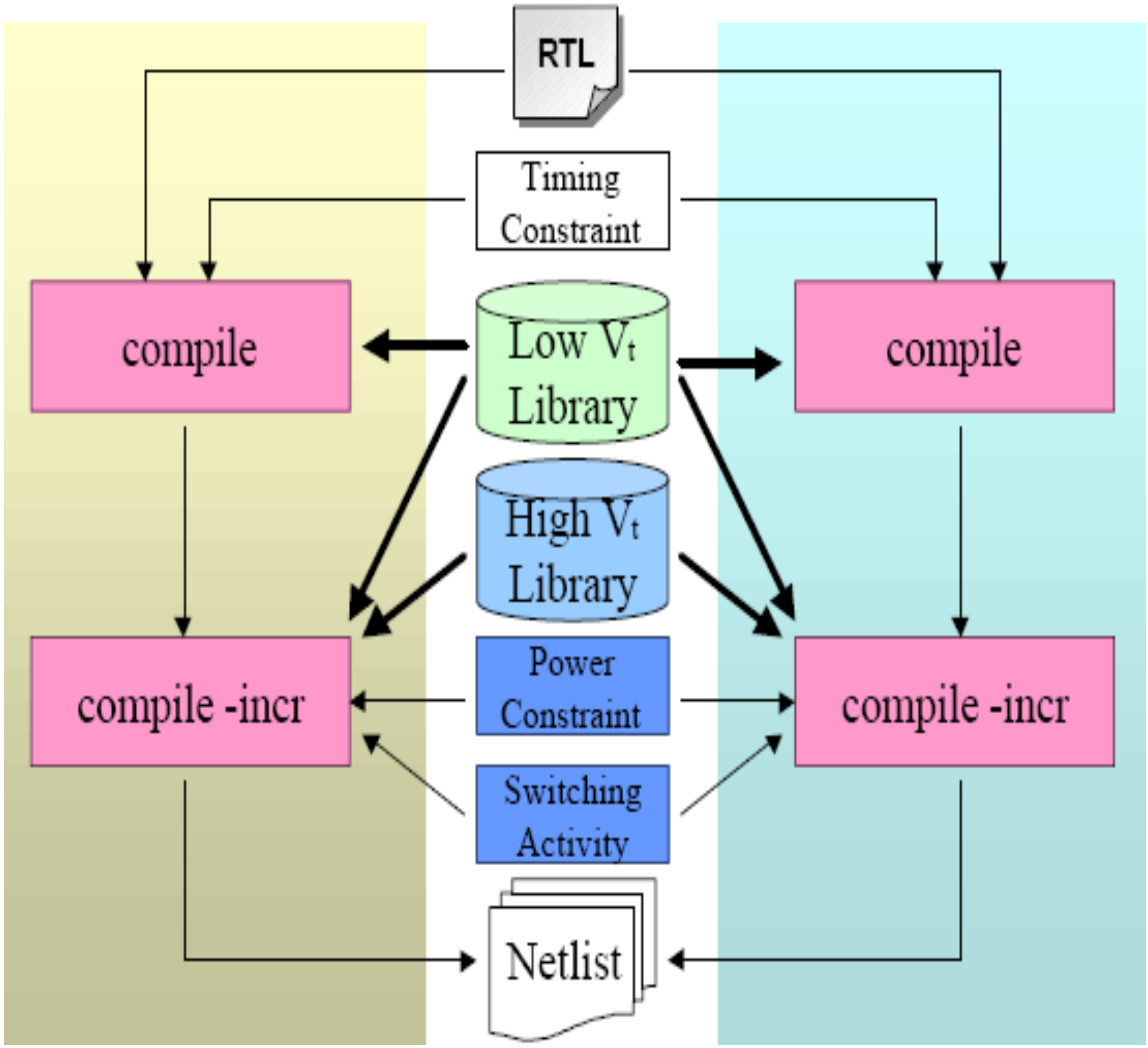
-Hvt to Mvt

-Rvt to Multi Vt

Low Vt to Multi-Vt

-Least cell count

-Good for tight timing constraint



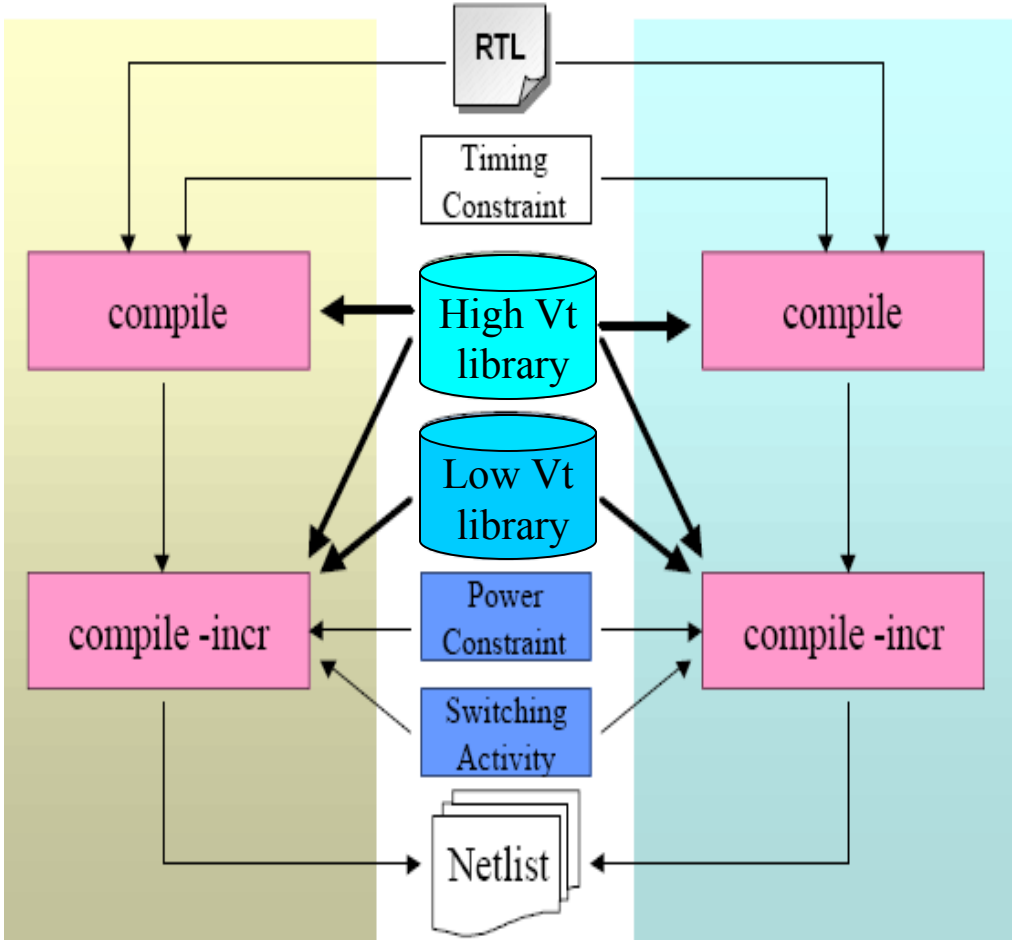
-Highest leakage power

-Less opportunity for leakage optimization

High Vt to Multi-Vt

-Least leakage power

-Good for leakage critical design

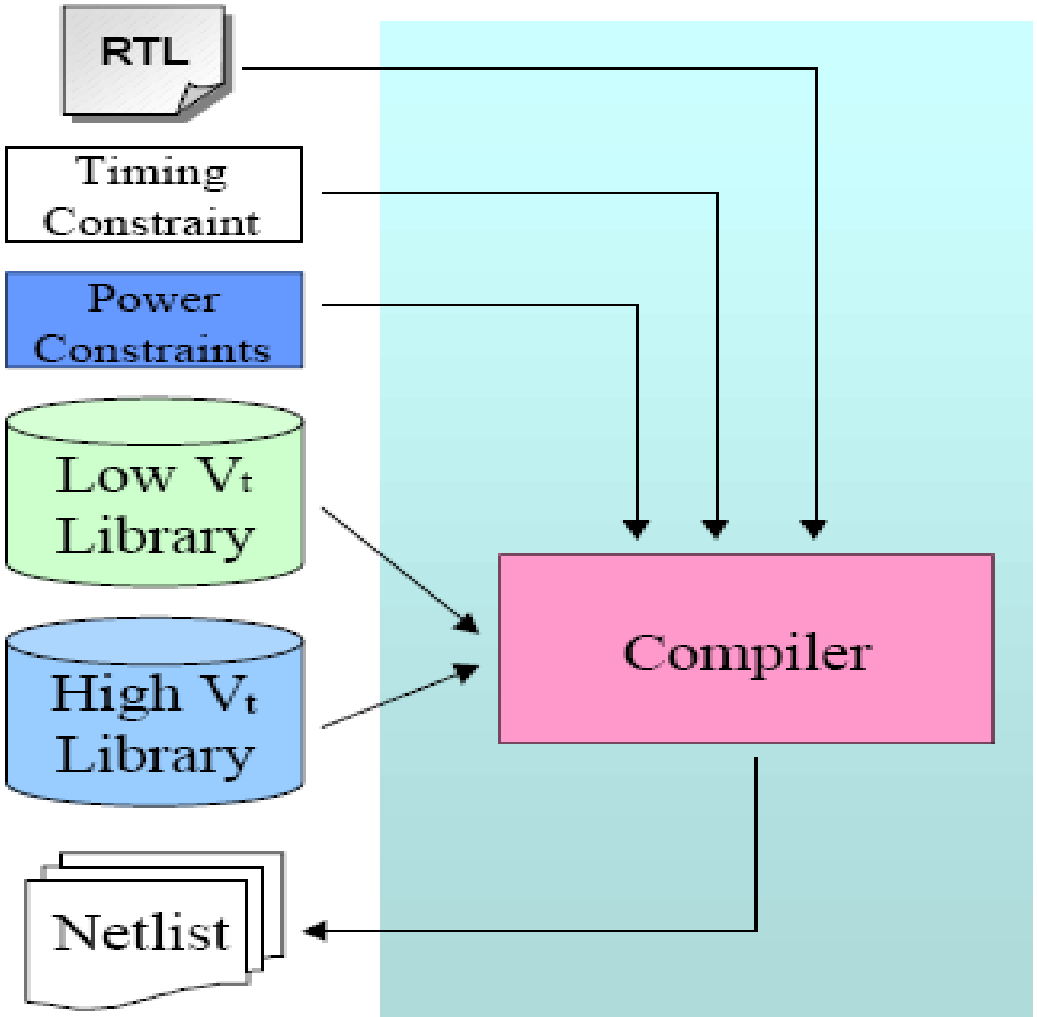


-Higher cell count

With different timing constraints it works as well balanced flow

Compile With Multi-Vt Libraries-Multi Vt One Pass Flow

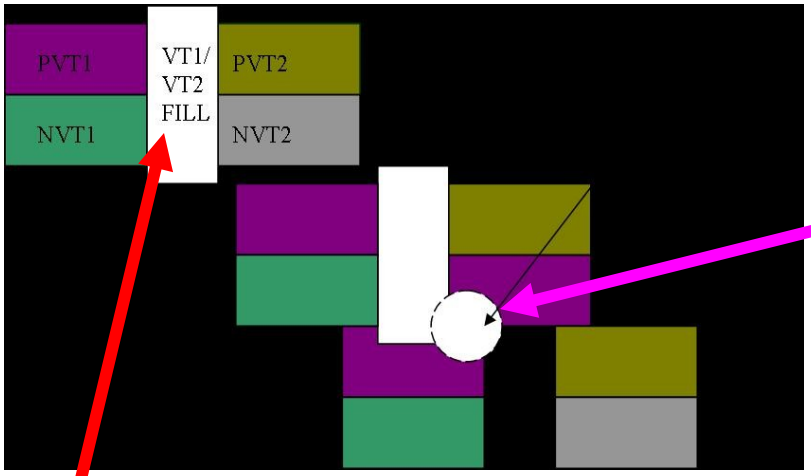
Overall good result



Can be used for most of the designs

Multi Vt Spacing requirement

- Low Vt have different well implantation
- Could overlap to adjacent High Vt cell
- Provide buffer space around edges



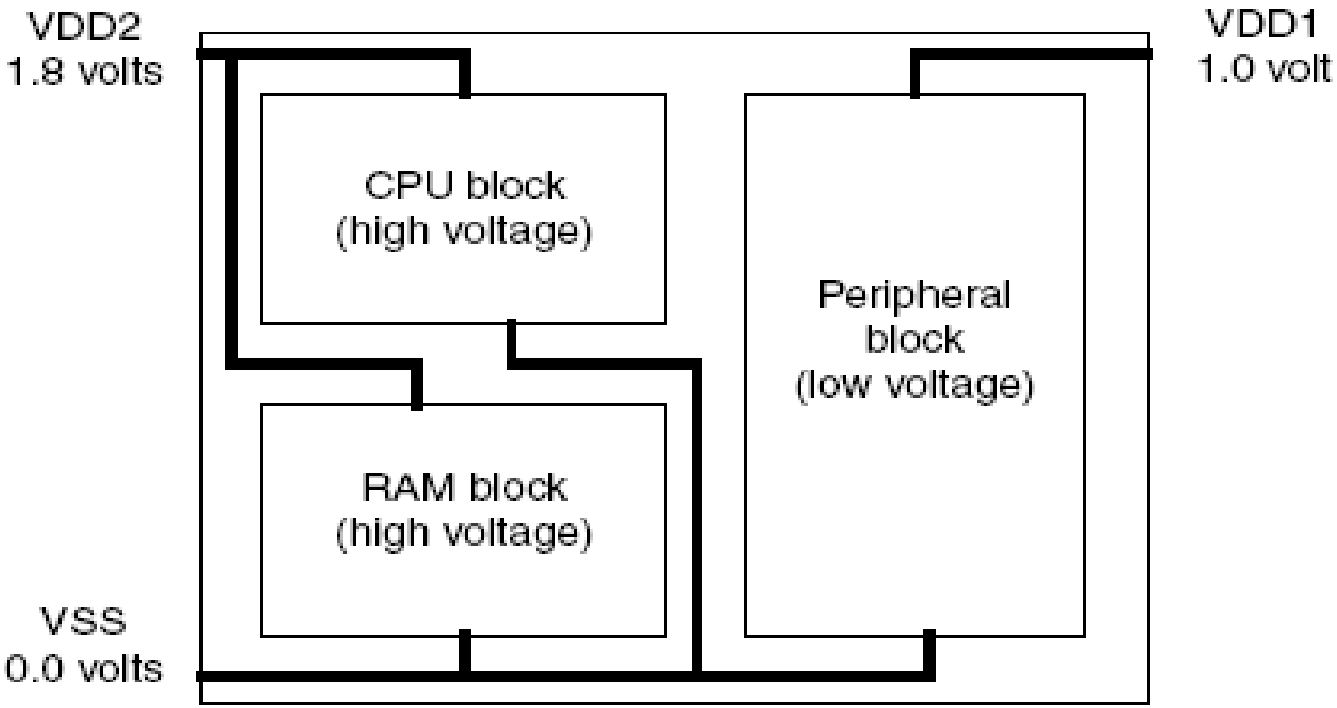
Placing opposite Vt filler cell can create gap in implant regions
➔ violation of DRC

Between Hvt Cells-place only Hvt filler cell
Between Lvt Cells-place only Lvt filler cell

IC Compiler handles the issue automatically

Multi Vdd (Voltage)

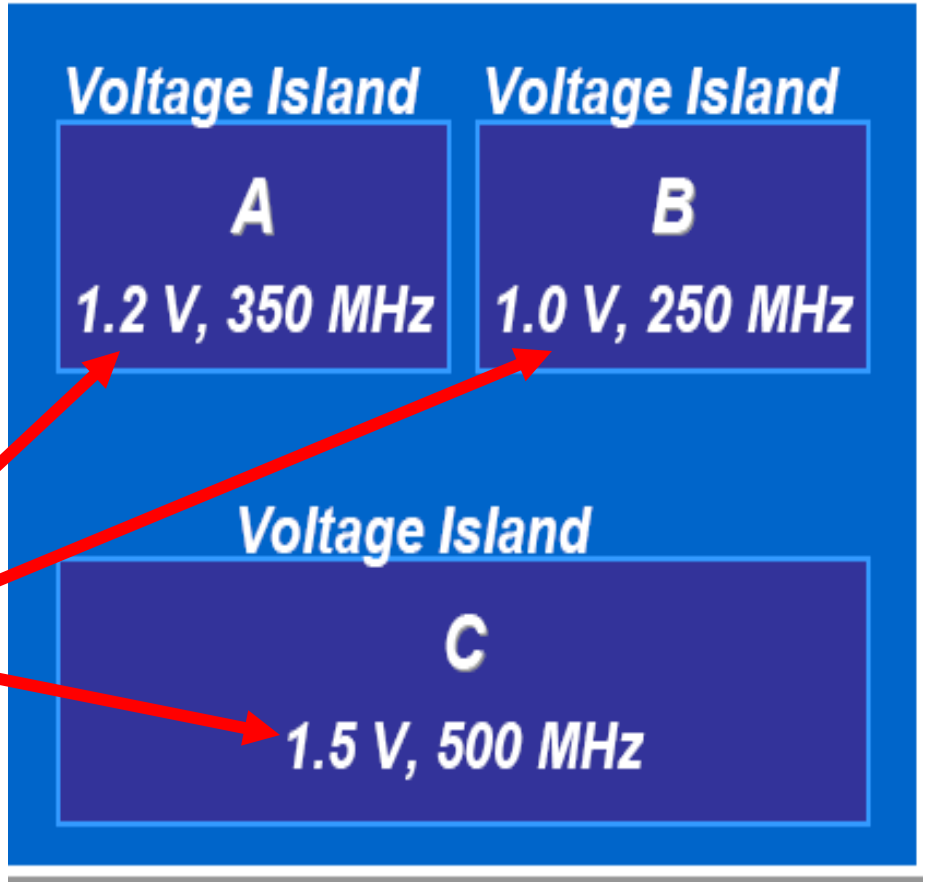
- Dynamic power is directly proportional to power supply.
- Reduce voltage when performance demand is less.
- Provide different voltage to different blocks.



Static Voltage Scaling (SVS)

Multiple Supply
Multi-Voltage (MV) Islands
- Voltage areas with fixed, single voltages

Different but fixed voltage is applied to different blocks or subsystems of the SoC design.

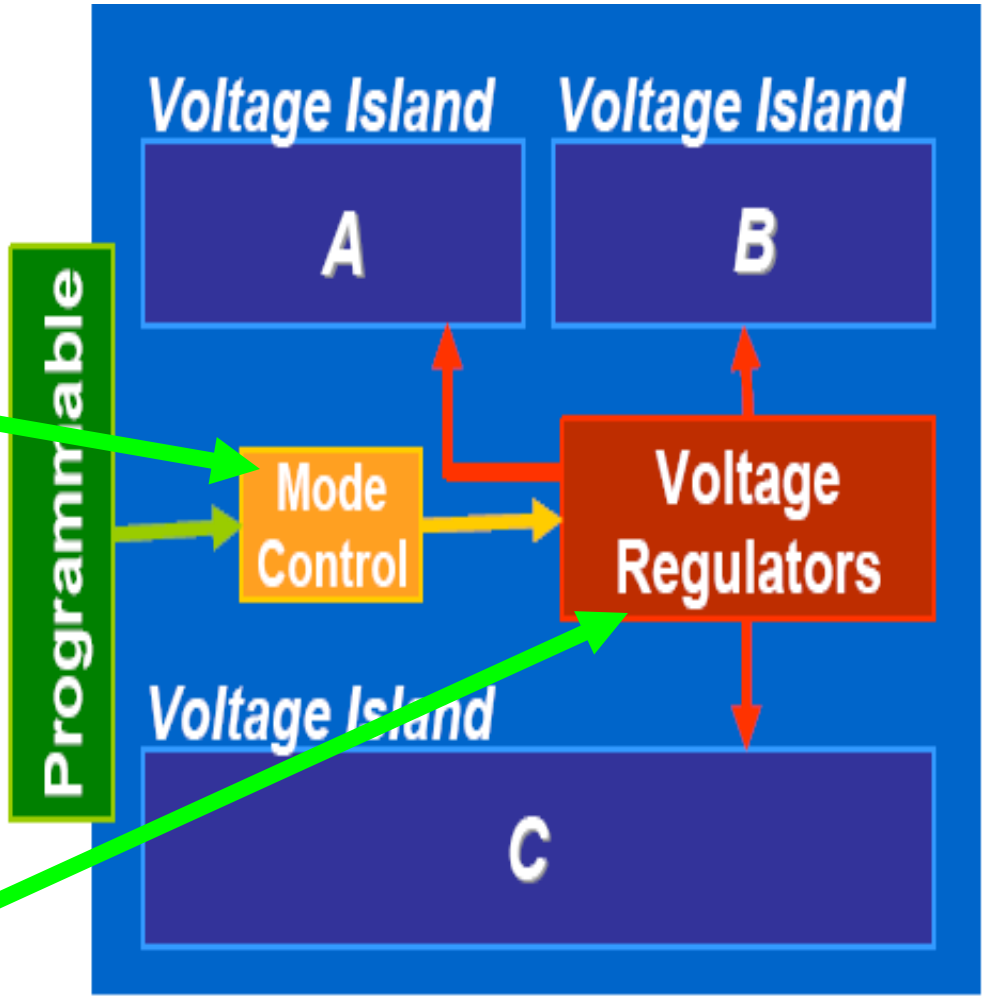


Dynamic Voltage and Frequency Scaling (DVFS)

- **Voltage areas with multiple, but fixed voltages**
- **Software controlled**

Voltage as well as frequency is dynamically varied as per the different working modes of the design

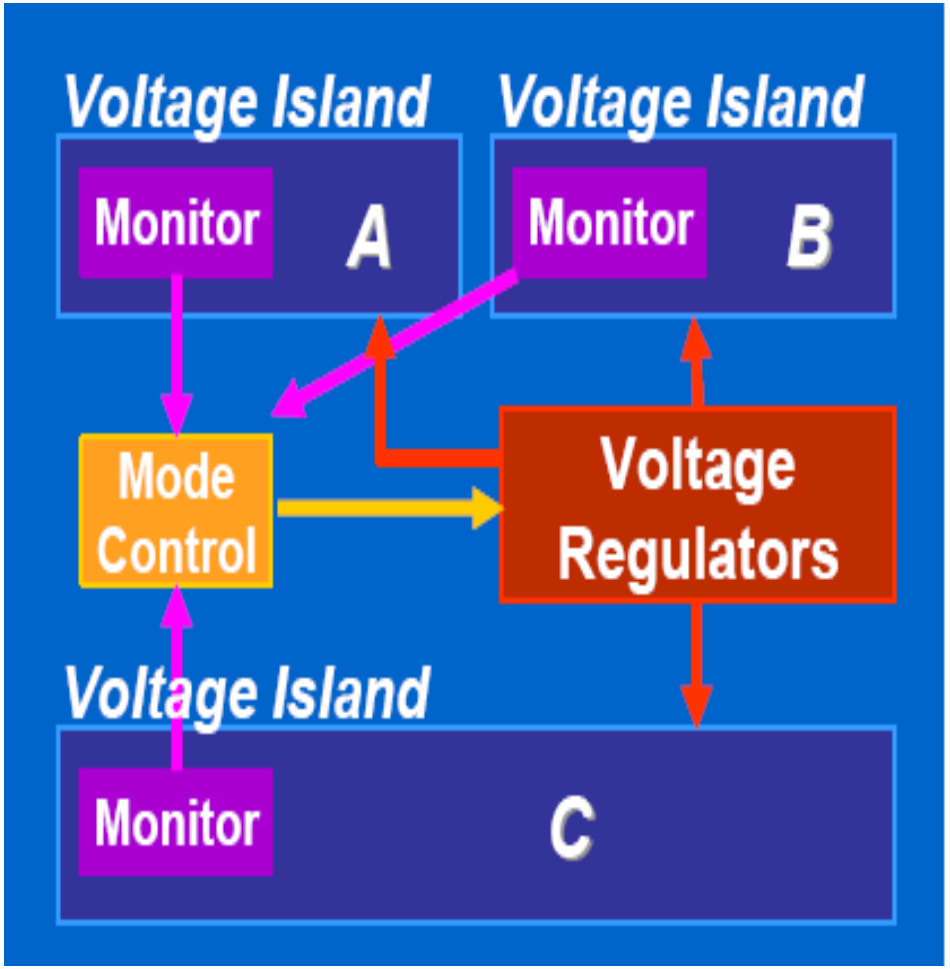
When high speed of operation is required voltage is increased to attain higher speed of operation with the penalty of increased power consumption



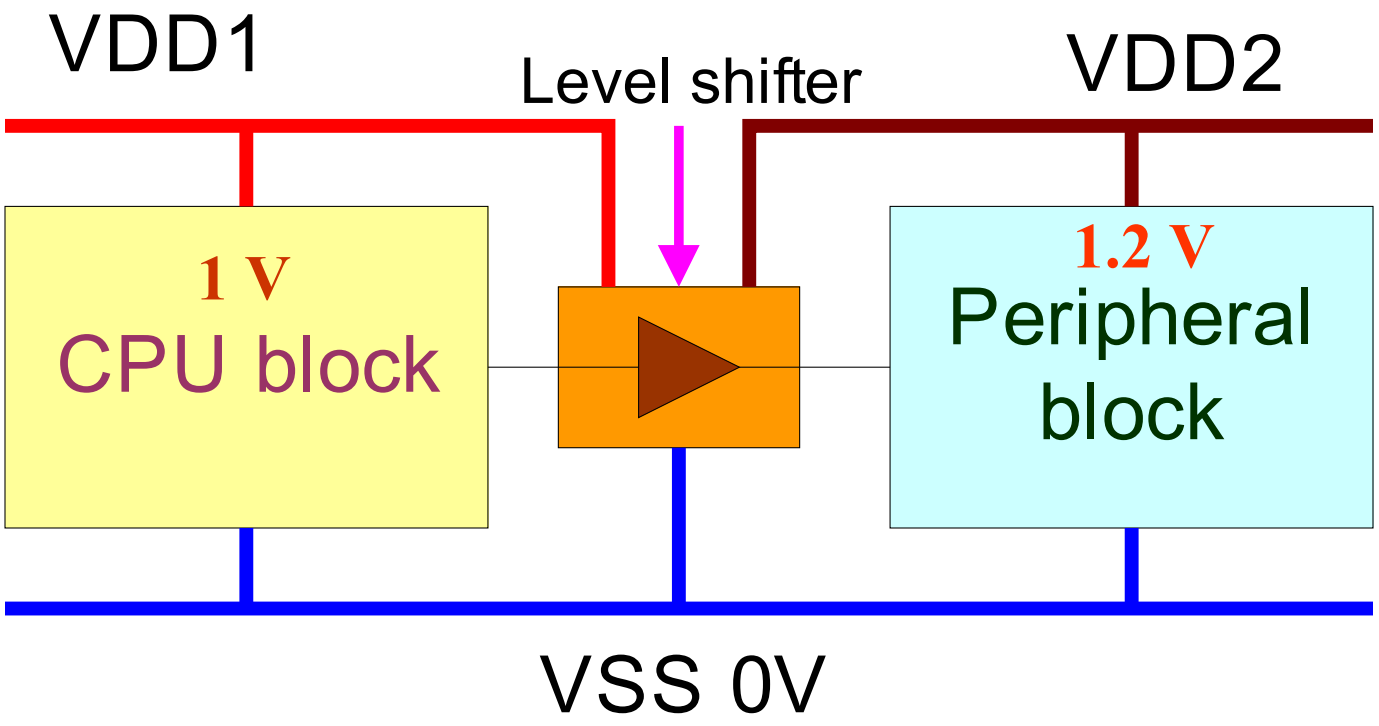
Adaptive voltage Scaling (AVS)

- Voltage areas with variable VDD
- Software controlled

- Voltage is controlled using a control loop.
- An extension of DVFS.

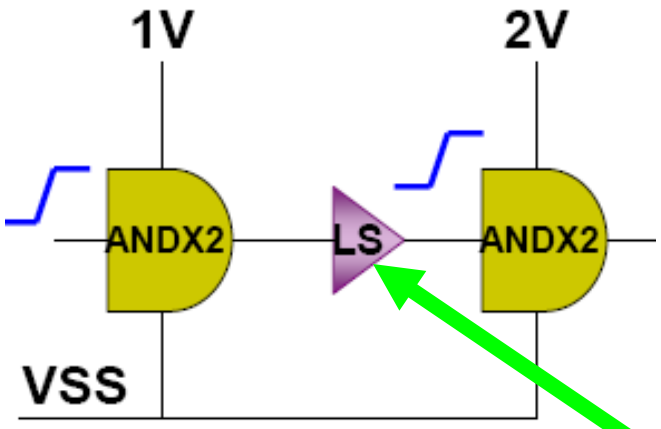


Multi Voltage Design Challenges: Level Shifters



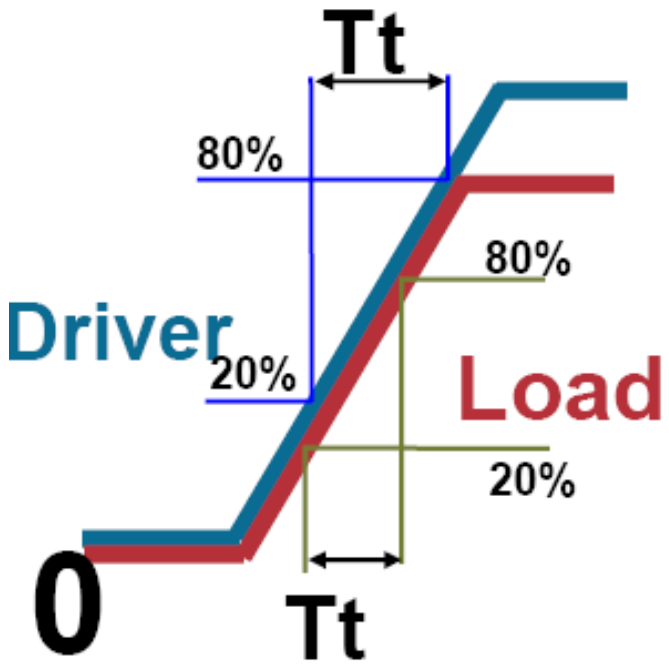
- **Multi power domain interface-** **voltage swing should match.**
- Propagation delay should be less.

Library description of level shifter



- Type of conversion performed
- High-to-low
- Low-to-high
- Or both

- Identities of the respective power pins that must be connected to each power supply.



- Supported voltage levels

Floor planning and Power Planning

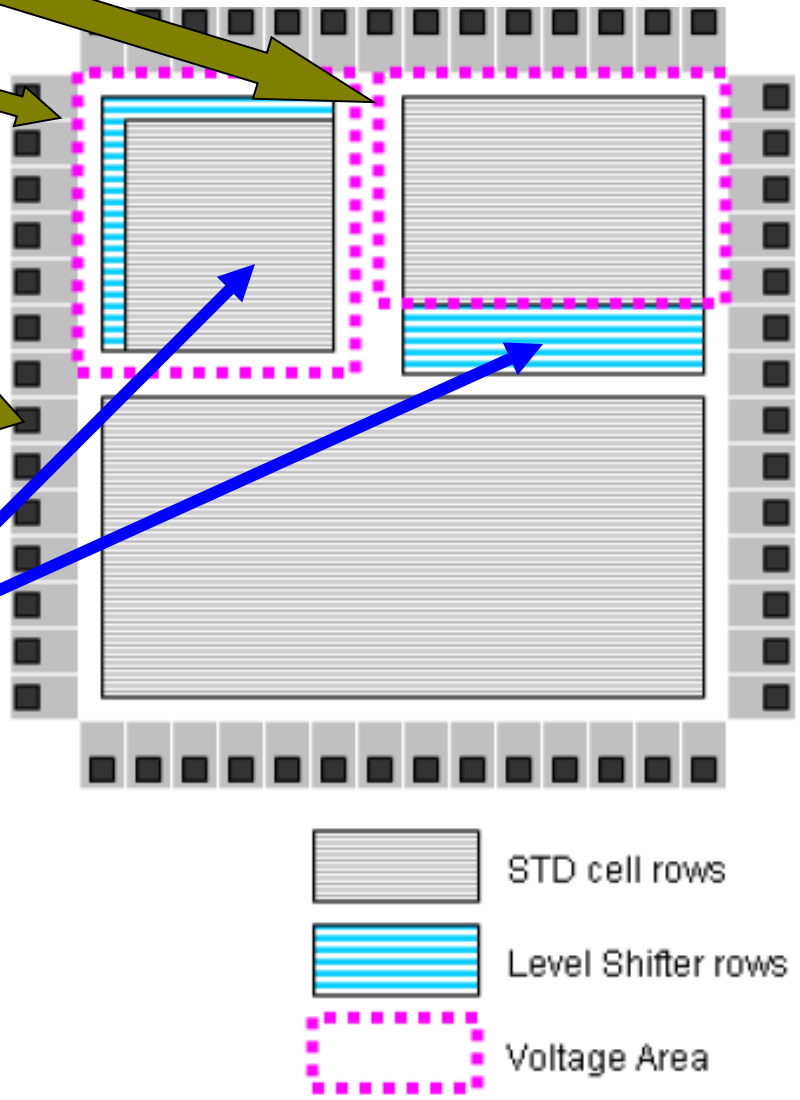
Every power domain requires independent local power supply and grid structure

- May have a separate power pad.
- In flip-chip designs power pad can be taken out near from the power domain.

Separate rows for standard cells and special cells

Burning issue

Local on chip voltage regulation or external separate supply???



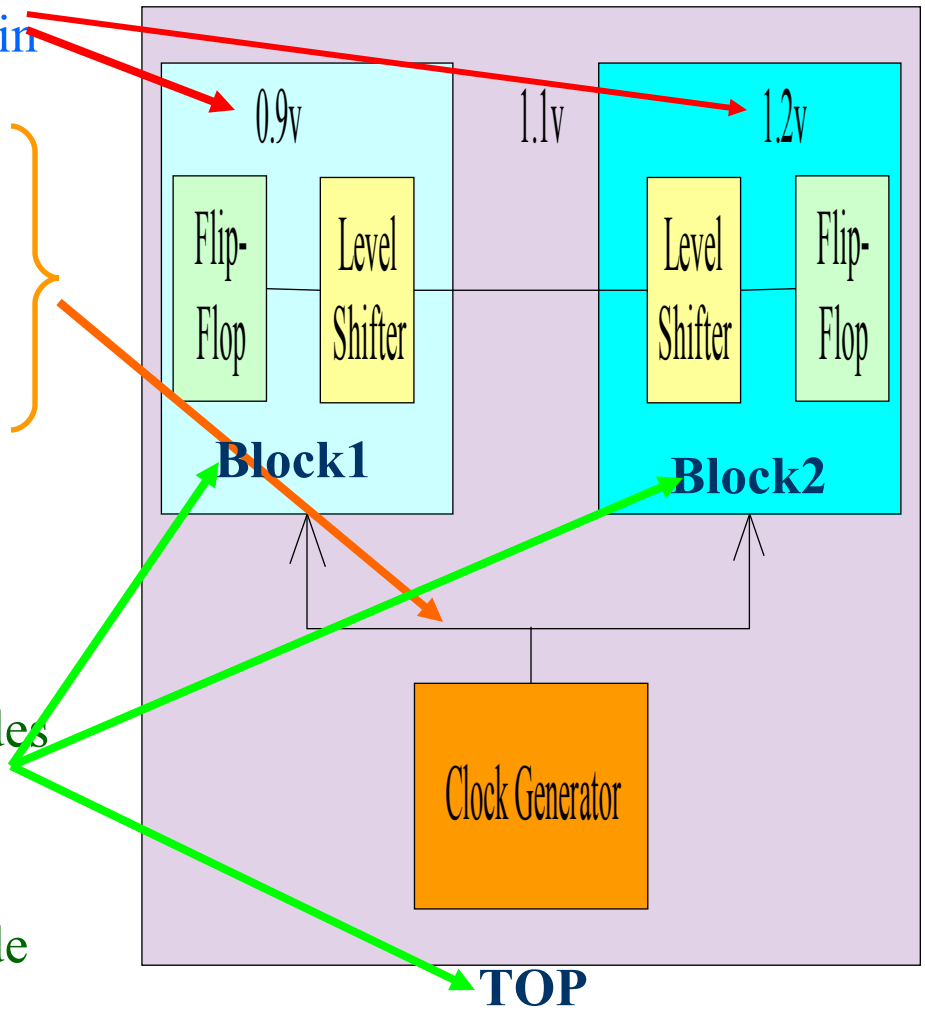
Multi Voltage Designs: Timing Issues

Clock

- Libraries should be characterized for different voltage levels that are used in the design
- Clock Tree Synthesis (CTS) tools should be aware of different power domains
- Clock tree is routed through level shifters to reach different power domains.

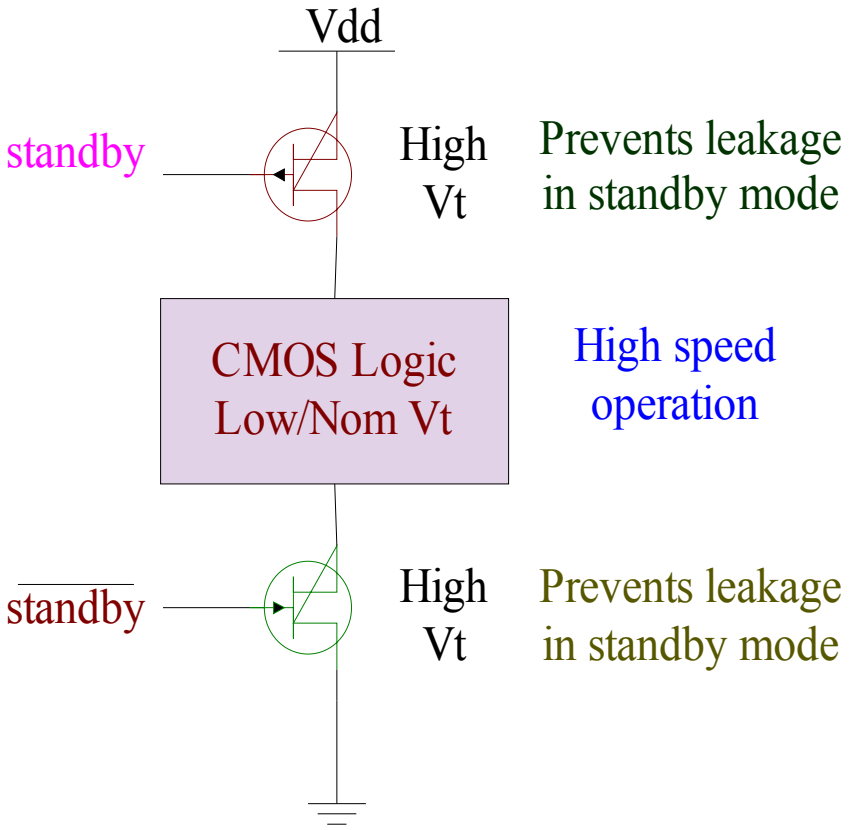
Static Timing Analysis (STA)

- For each supply voltage level or operating point constraints should specified.
- There can be different operating modes for different voltages.
- Constraints need not be same for all modes and voltages.
- The performance target for each mode can vary.

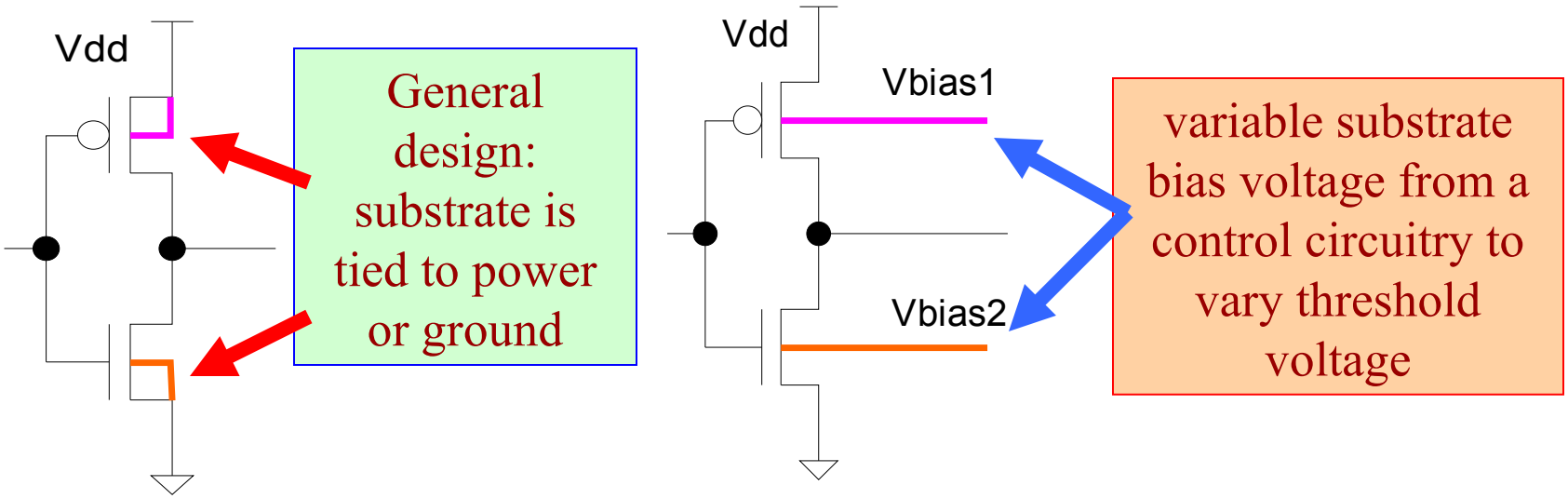


Multiple Threshold CMOS (MTCMOS) Circuits

- Use Hvt and Lvt cells
- Mainly used ion CMOS full custom design
- Extensively used in Power gating
- Called as “sleep transistor”



Variable Threshold CMOS (VTCMOS)-Substrate biasing



Pros

- Considerable power reduction
- Negligible area overhead

Cons

- Requires either twin well or triple well technology to achieve different substrate bias voltage levels at different parts of the IC

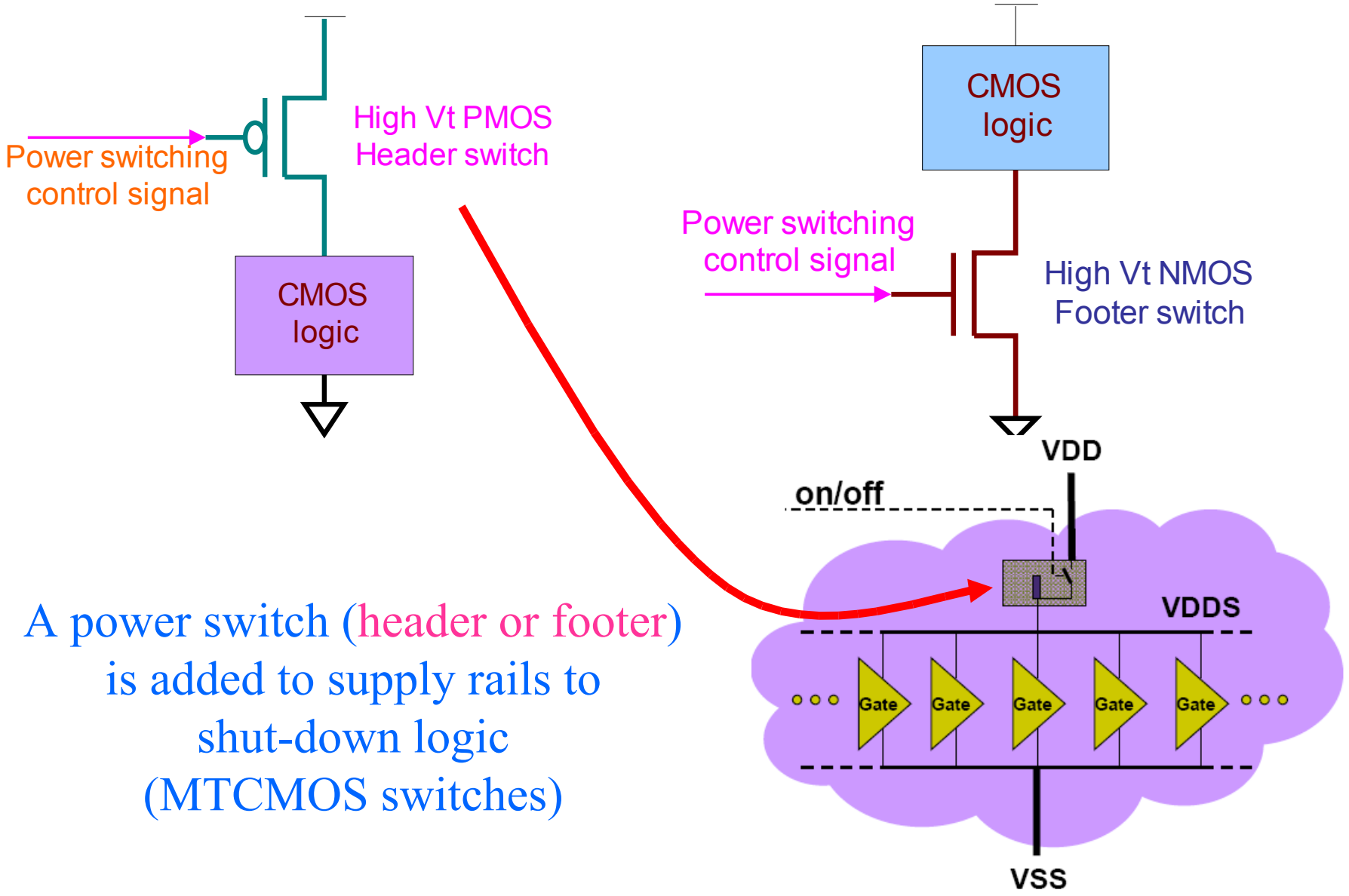
Power-gating

- Circuit blocks that are not in use are temporarily turned off
- Affects design architecture more compared to the clock gating
- It increases time delays as power gated modes have to be safely entered and exited

How to shut down????

- Either by software or hardware
- Driver software can schedule the power down operations
- Hardware timers can be utilized
- A dedicated power management controller is the other option.
- Switch off the block by using external power supply for long term
- Use CMOS switches for smaller duration switch off
- Header switch (PMOS) or footer switch (NMOS)

Header –Footer Switches



A power switch (header or footer) is added to supply rails to shut-down logic (MTCMOS switches)

Power-gating parameters

Power gate size

- Should handle the switching (rush) current
- Big enough not to have IR drop
- Footer gates are smaller for the same amount of current (NMOS has twice mobility of PMOS)

Gate control slew rate

- Larger the slew larger the time taken to switch off or switch on

Simultaneous switching capacitance

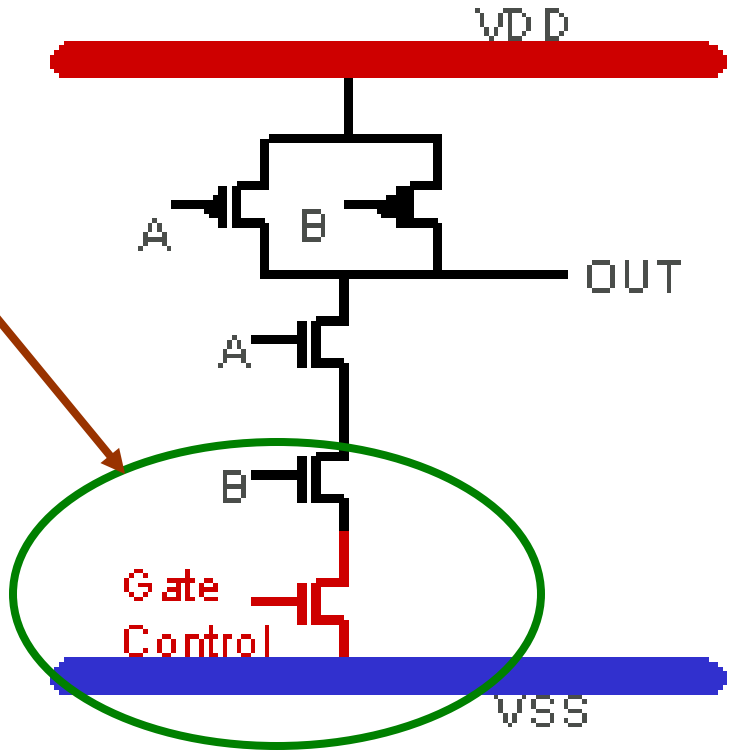
- Refers to the amount of circuit that can be switched simultaneously without affecting the power network integrity
- Rush current may damage the circuitry
- Switch the block step by step

Power gate leakage

- Should have less leakage
- Use High Vt transistors ==> slower switching

Fine-grain power gating

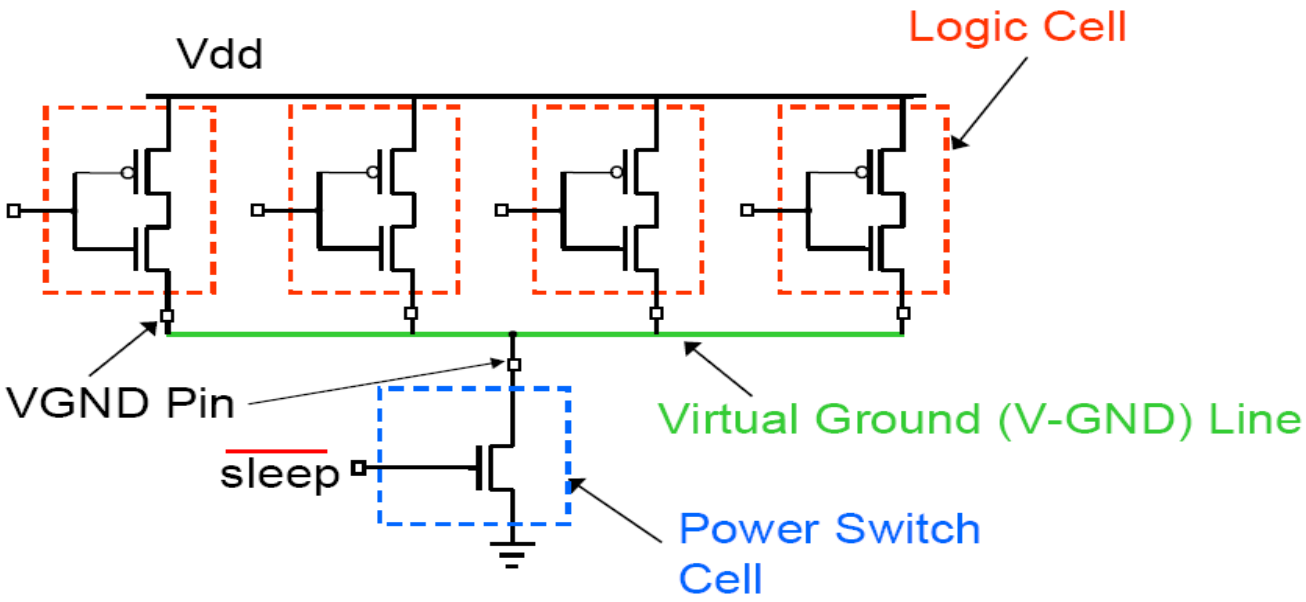
- Add a sleep transistor to every cell
- Switching transistor as a part of the standard cell logic
- ~10X leakage reduction



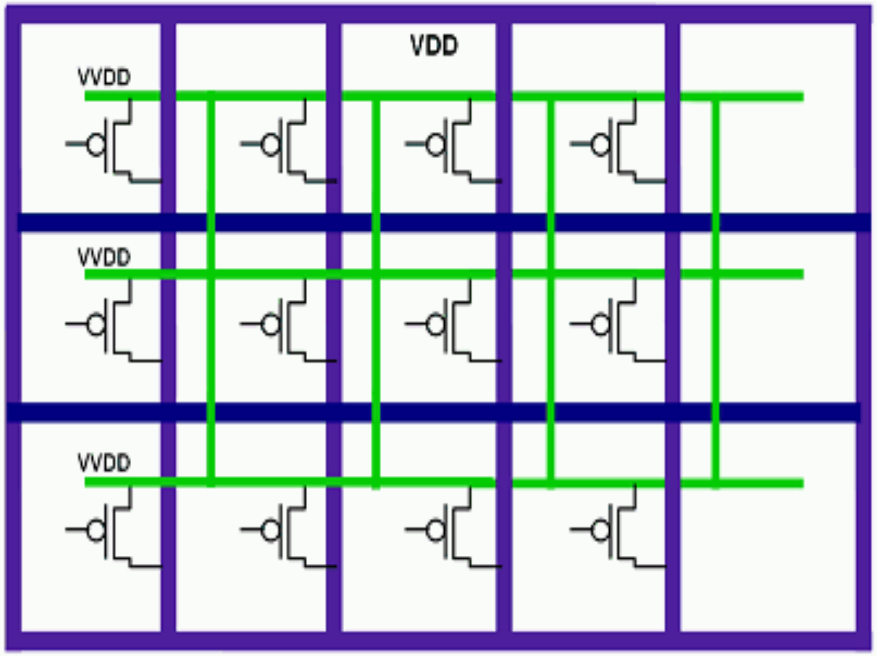
- Large area penalty
- Creates timing issues

Coarse-grain power gating

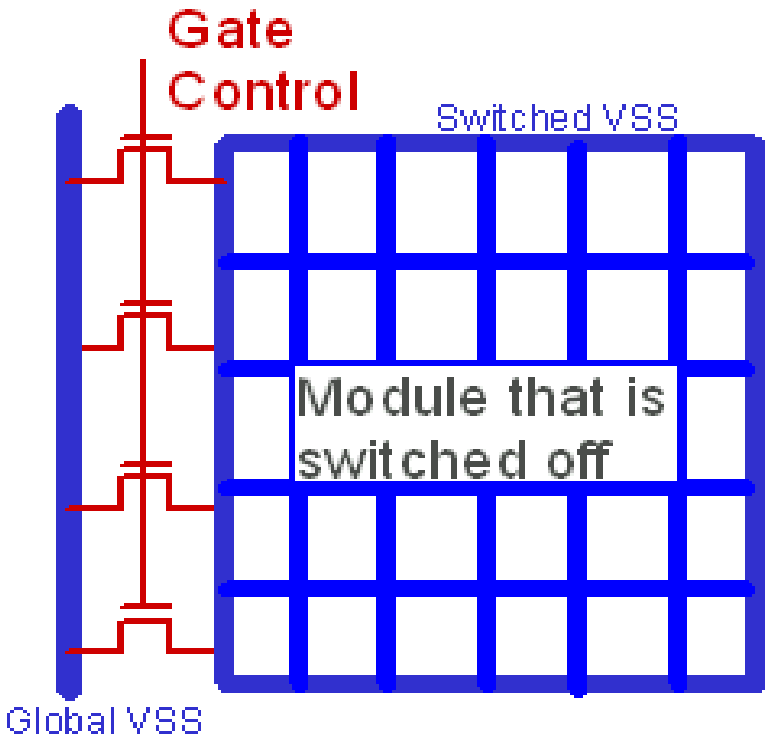
- Grid style sleep transistors
- Power-gating transistor is a part of the power distribution network
- Less sensitive to PVT variation
- Introduces less IR-drop variation
- Imposes a smaller area overhead
- Switching capacitance is a major issue.....switch on blocks one by one, use counters, daisy chain logic
- The global power is the higher layers of metal (Metal 5 and 6 in a 6 metal layer process), while the switched power is in the lower layers (Metal 1 and 2).



Coarse-grain power gating-Column or Ring based



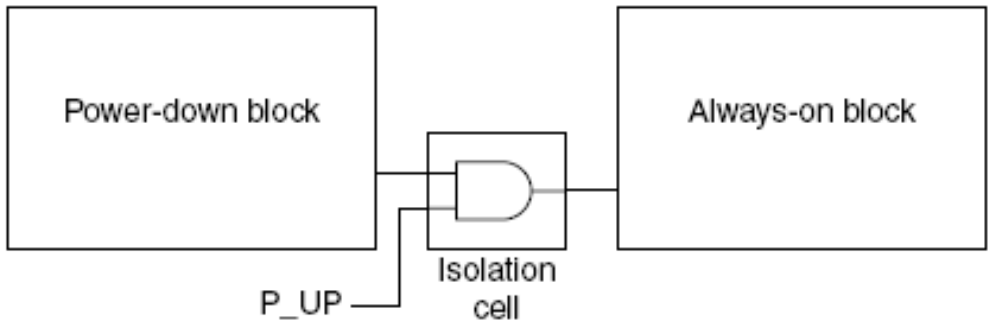
- Column-based methodology
- Gates are inserted within the module



- Ring-based methodology
- Power gates are placed around the perimeter of the module

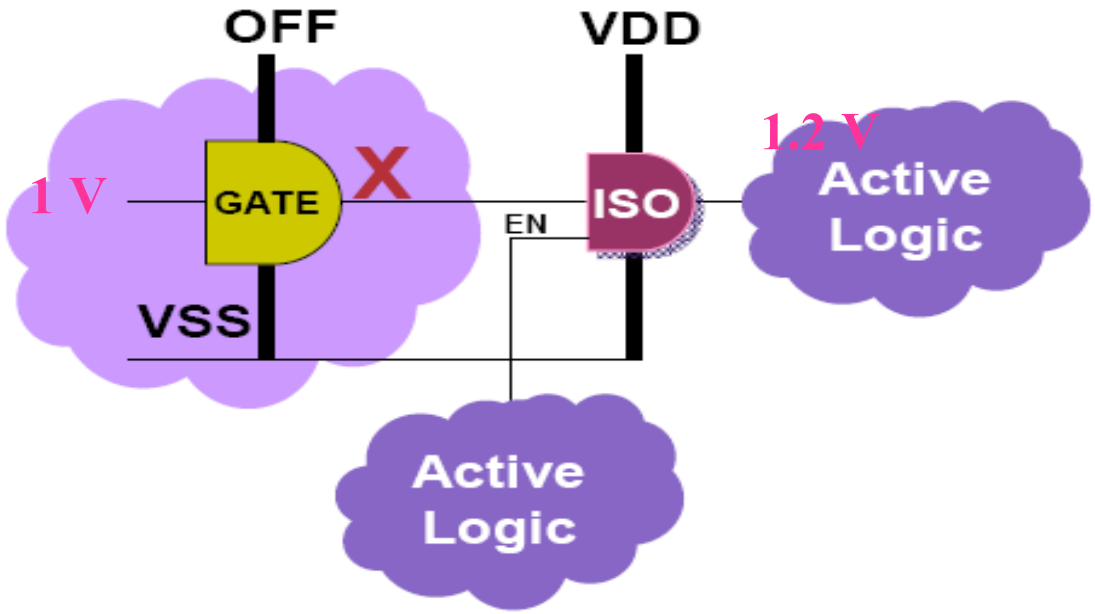
Isolation Cells

- Isolate power gated block from the normally on block.
- Isolation cells are specially designed for low short circuit current when input is at threshold voltage level.
- Isolation cell provides a known, constant logic value to an always-on block when the power-down block has no power.
- Can hold a logic 1 or 0, or can hold the signal value latched at the time of the power-down event.
- Isolation cells must themselves have power during block power down periods.



Enable level shifter

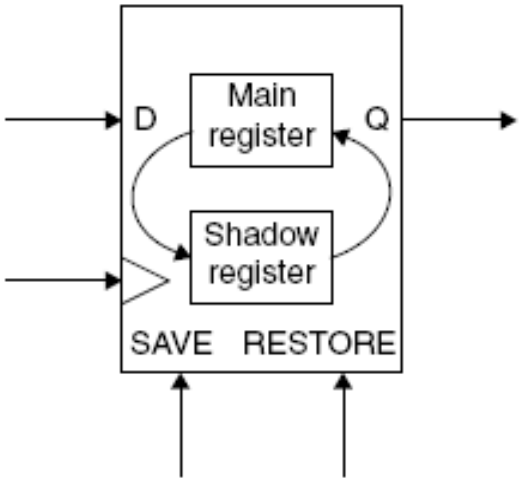
- Acts as level shifter and Isolation cell

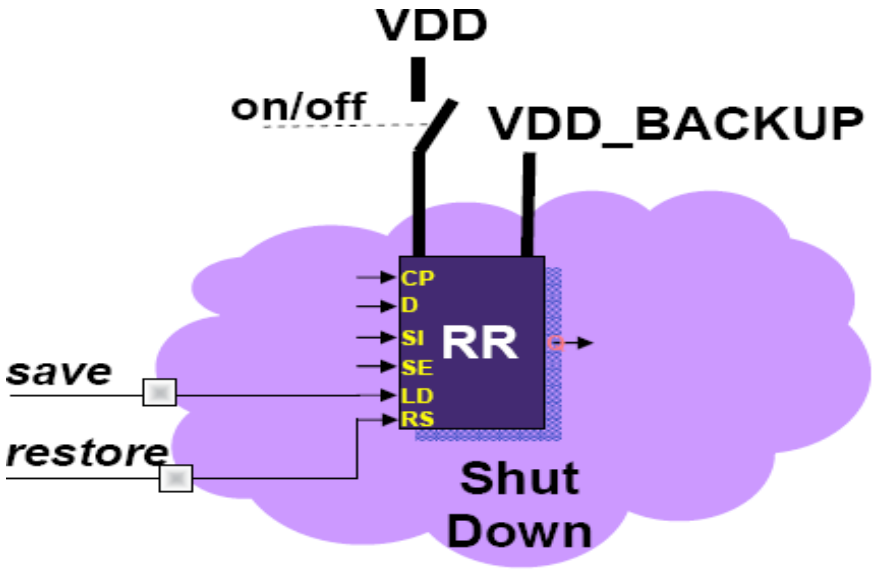


- The power switching can be combined with multi-voltage operation.
- The interface cells between different blocks must perform both level shifting and isolation functions.

Retention Registers

- Special low leakage flip-flops used to hold the data of main register of the power gated block.
- Always powered up.
- Power gating controller controls the retention mechanism





- SAVE signal saves the register data into the shadow register prior to power-down and the RESTORE
- Signal restores the data after power-up.

Always on logic

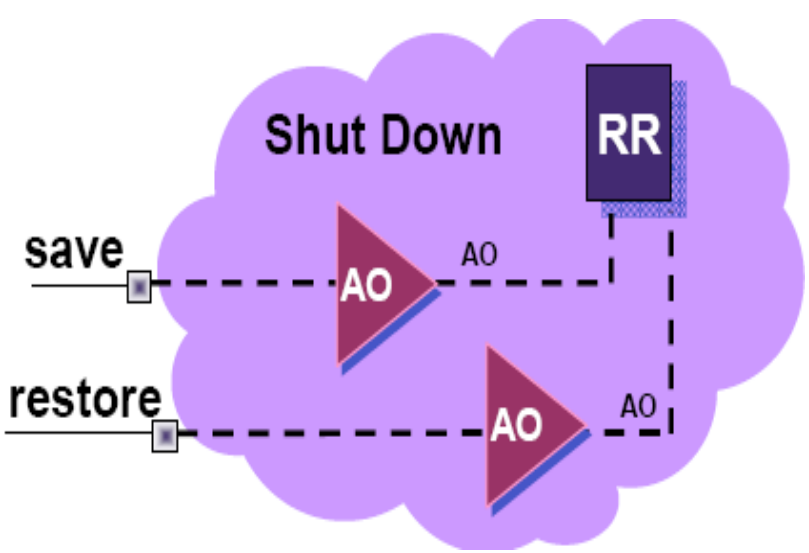
Some logic needs to stay active during shut-down

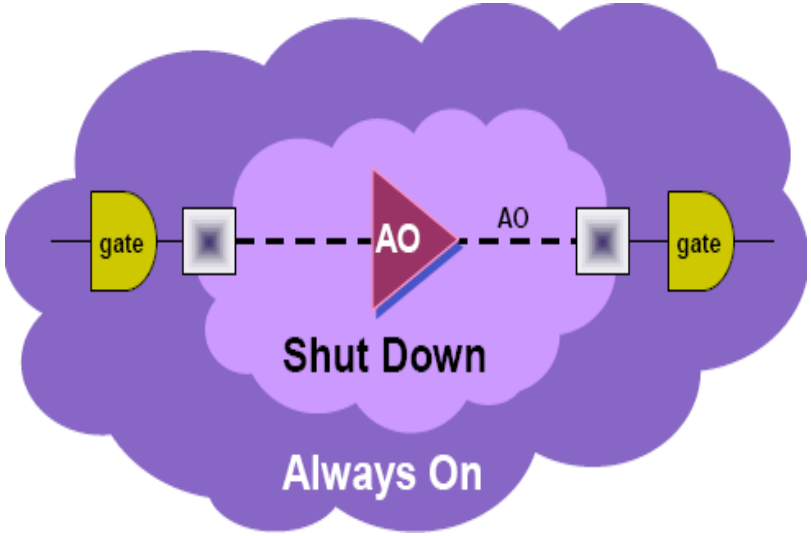
Internal enable pins (ISO/ELS)

I Power switches

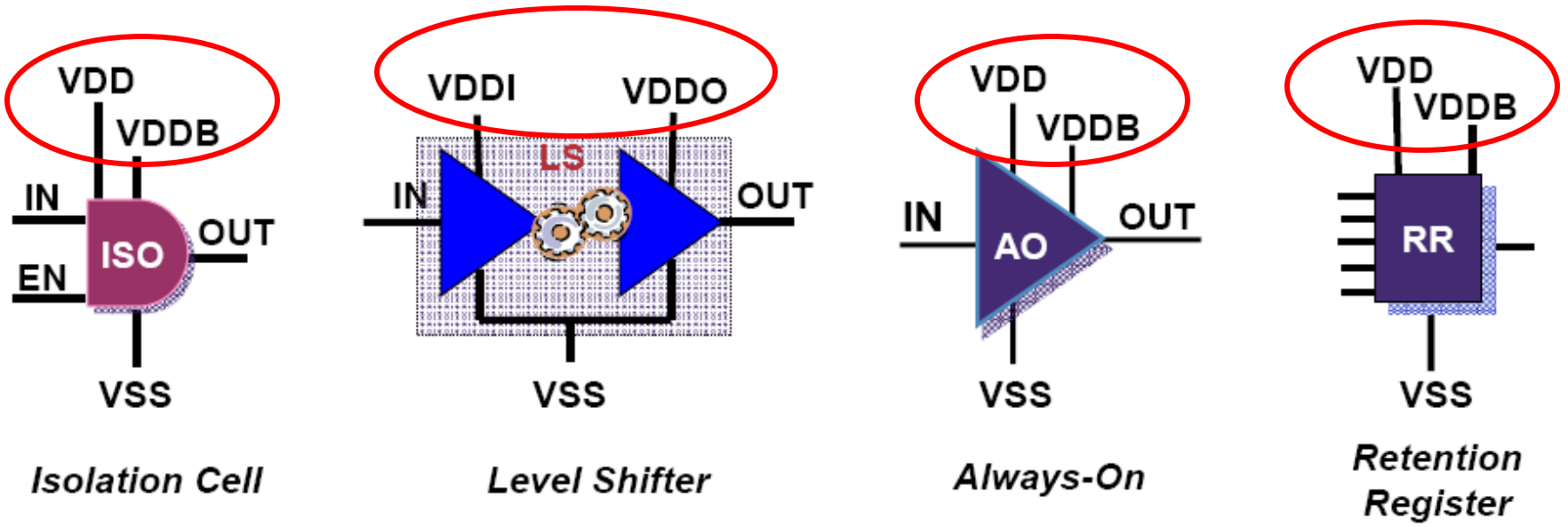
P Retention registers

R User-specific cells





Low-Power Infrastructure

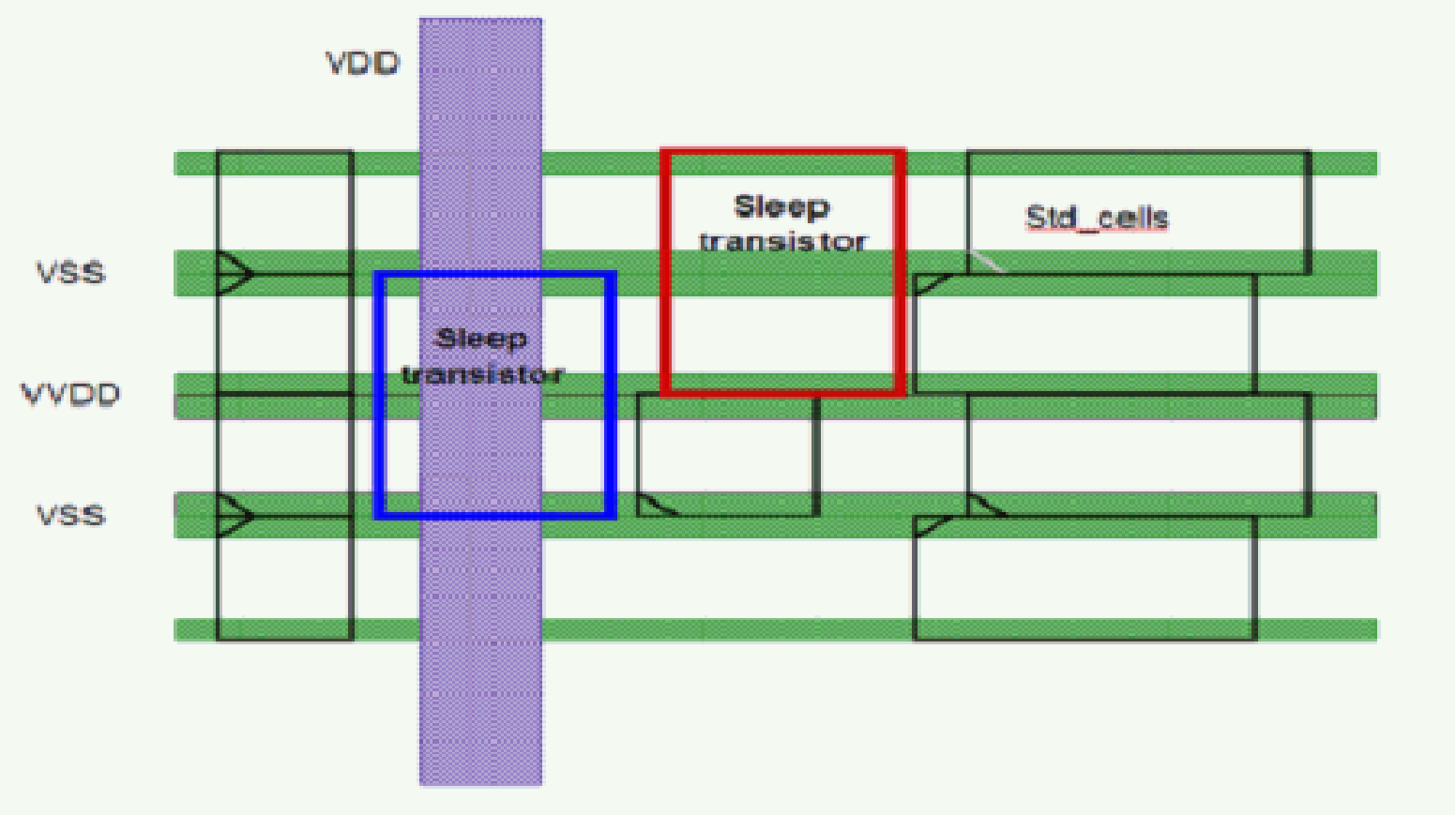


Low-power design requires new cells with multiple power pins

Additional modeling information in ".lib" is required to automatically handle these cells

Layout Constraints

- Occupy two rows of standard cell placement
- The sleep transistors need to be placed as close as possible to the metal straps to minimize IR drops



Library syntax of special cells

```
cell (Retention_DFF) {
retention_cell:"ret_dff";
  area : 1.0;
  ...
 pg_pin(VDDG) {
   voltage_name : VDDG;
   pg_type : backup_power;
 }
 ...
 pin(RETN) {
   direction : input;
   capacitance : 1.0;
   nextstate_type : data ;
   related_power_pin :VDDG;
   related_ground_pin:VSSG;
   retention pin
     (save_restore, "1" );
  }

  pin(Q) {
power_down_function:"!VDD+VSS";
related_power_pin   : VDD;
related_ground_pin : VSS;
direction : output;
```

```
cell (Enable_Level_Shifter) {
  is_level_shifter : true;
  level_shifter_type : LH ;
  input_voltage_range(0.7,1.4);
  output_voltage_range(0.7,1.4);

pg_pin(P1) {
  voltage_name : VDD1;
  pg_type : primary_power;
  std_cell_main_rail : true;
}
pg_pin(P2) {
  voltage_name : VDD2;
  pg_type : primary_power;
}
...
pin(A) {
  direction : input;
  related_power_pin : P1;
  related_ground_pin : G1;
  level_shifter_data_pin:true;
}
pin(EN) {
  direction : input;
  related_power_pin : P1;
  related_ground_pin : G1;
  level_shifter_enable_pin:true;
} ...
```

```
cell ( Simple_Switch ) {
  ...
  switch_cell_type:
          coarse_grain;

  pg_pin ( VDDG ) {
    pg_type :primary_power;
    direction : input;
    voltage_name : VDD;
  }
 ...
   pg_pin(VDD) {
    voltage_name : VDD;
    pg_type:internal_power;
    direction : inout;
    switch_function : "SLEEP";
    pg_function : "VDDG";
  }
 ...
 pin ( SLEEP ) {
    switch_pin : true;
    capacitance: 0.034;
  }
 ...
```

Input vector control (IVC)

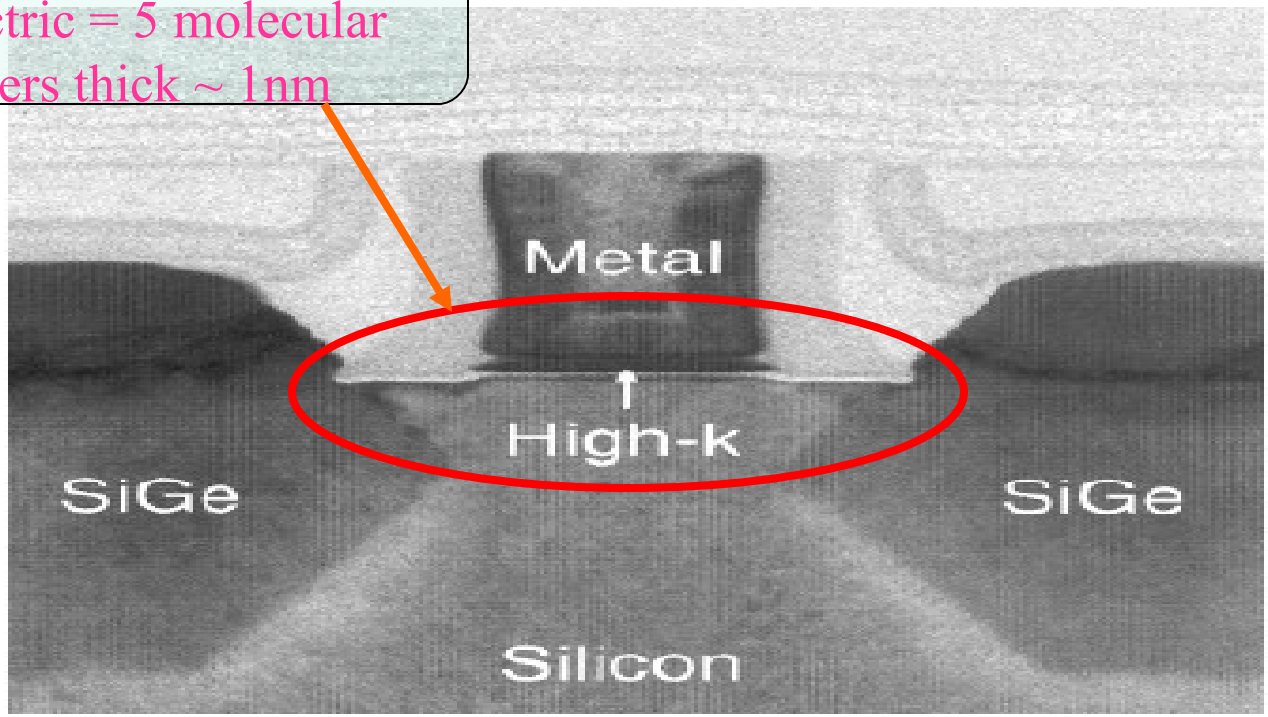
- **Sub threshold leakage and gate leakage are input vector dependent.**

- Force primary inputs, latches and flip-flops into certain logic values when they are not in active state

	NAND (W_p=0.3375um,W_n=0.225um)		
Input	I_{sub} (A)	I_{gate} (A)	I_t (A)
000	2.62E-11	1.41E-10	1.67E-10
001	2.56E-09	2.65E-09	5.21E-09
010	9.19E-11	1.01E-08	1.02E-08
011	5.16E-09	5.16E-09	1.03E-08
100	4.31E-11	5.41E-11	9.73E-11
101	2.63E-09	2.56E-09	5.19E-09
110	1.18E-10	1.17E-10	2.35E-10
111	7.72E-09	7.54E-09	1.53E-08

Improvement in Process technology

For 90nm and 65nm dielectric = 5 molecular layers thick ~ 1nm

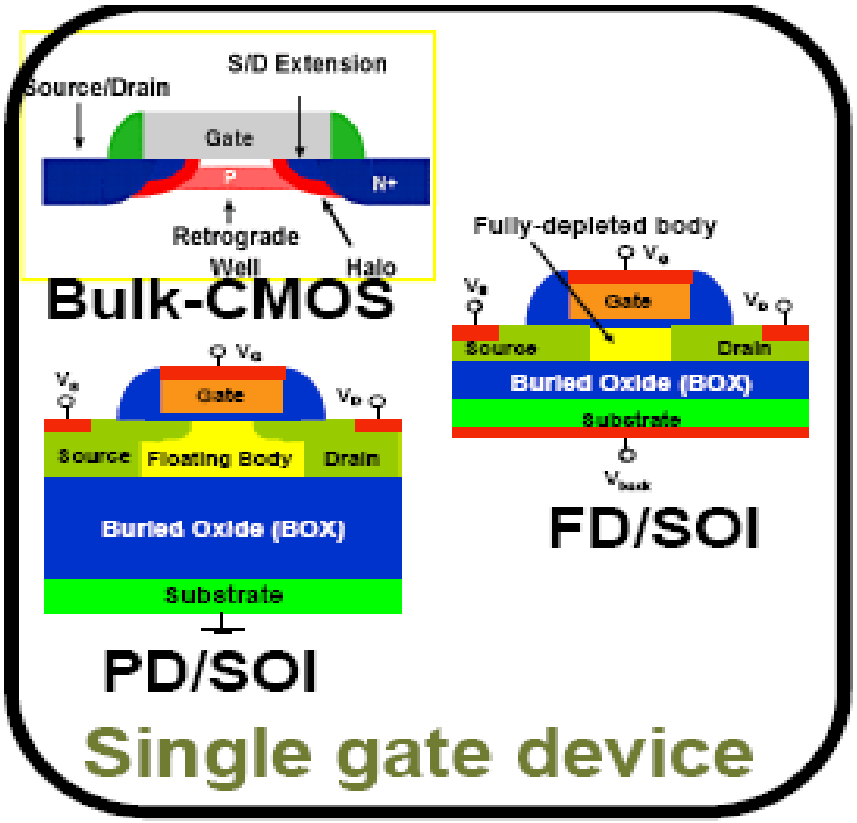


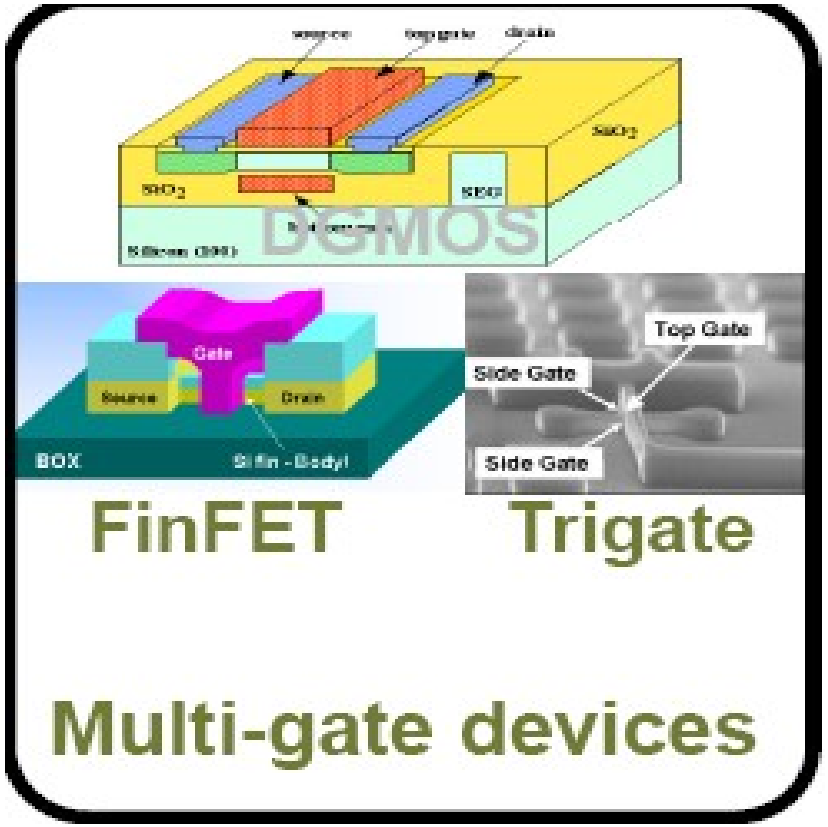
25x reduction in gate leakage

5x reduction in sub threshold leakage

Improvement in Process technology (Contd...)

- New devices: SOI, FinFET





2003 2009

Tradeoffs

Power-reduction Technique	Power Benefit	Timing Penalty	Area Penalty	Methodology Impact			
				Architecture	Design	Verification	Implementation
Multi-Vt Optimization	Medium	Little	Little	Low	Low	None	Low
Clock Gating	Medium	Little	Little	Low	Low	None	Low
Multi-supply Voltage	Large	Some	Little	High	Medium	Low	Medium
Power Shut-off	HUGE	Some	Some	High	High	High	High
Dynamic and Adaptive Voltage Frequency Scaling	Large	Some	Some	High	High	High	High
Substrate Biasing	Large	Some	Some	Medium	None	None	High

Tradeoffs Contd...

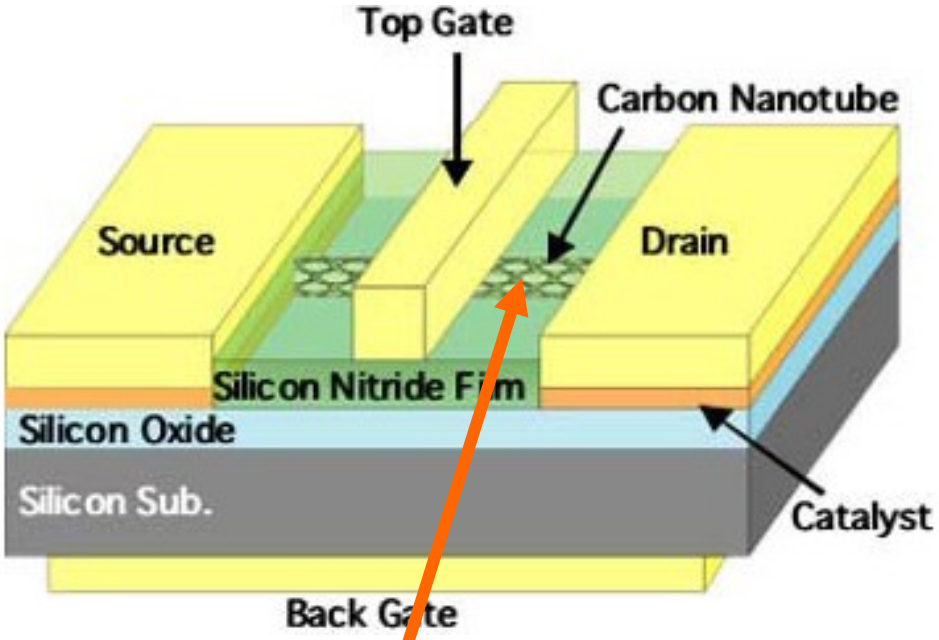
	Leakage power	Dynamic power	Timing	Area penalty	Methodology impact	Methodology change
Low-power optimization	10%	10%	0%	10%	None	None
Multi-V_t	6X	0%	0%	0%	Low	Multi-V_t library needed
Clock gating	0%	20%	0%	<2%	Low	Clock-gating cells needed and extra overhead in STA
Multisupply voltage	2X	40% to 50%	0%	<10%	Medium	Microarchitecture and methodology needs to be domain-aware; Need voltage regulators and level shifters; verification and analysis challenge
Power shutoff	10X to 50X	0%	4% to 8%	5% to 15%	Medium-high	Insertion of switch cells; retention flops; wake-up and shutdown time analysis
Dynamic voltage frequency scaling	2X to 3X	40% to 70%	0%	<10%	High	Multimode optimization and analysis flow needed: Clock synchronization
Substrate biasing	10X	–	10%	<10%	High	Maintain well separation; multiple power rail distribution; timing analysis

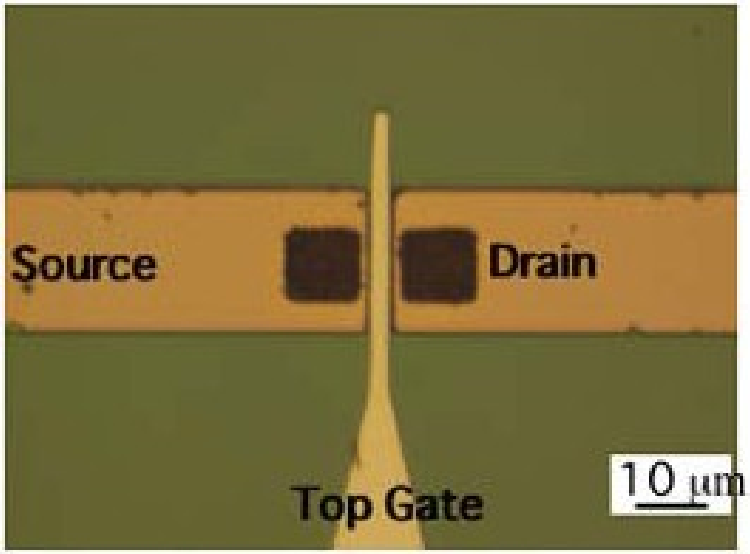
Source: Cadence Design Systems

Future low power strategy????

- Asynchronous Design - Solution to Dynamic Power?
  - Clock is a third to half the total dynamic power
  - Let's get rid of the clock
  - Micro pipeline: A Simple Asynchronous Design Methodology
- Is Hi-k sufficient for 22nm and 16nm?
  - Whether this type of transistor structure (hi-k, metal gate) will continue to scale to the next two generations—22 nm and 16 nm—is a question for the future.
- Is there a simple, coherent power strategy that unifies the best of DVFS, power gating, asynchronous ?
- How do we represent and verify very complex power intent such as asynchronous ? Can we separate function from implementation?

Carbon Nano tubes ???





Channel is a coil of carbon hexagons

Mobility up to 70x silicon

Spintronics ???

- Information is stored (written) into spins as a particular spin orientation (up or down)
- The spins, being attached to mobile electrons, carry the information along a wire
- The information is read at a terminal.

References

- http://asic-soc.blogspot.com
- www.cadence.com
- www.synopsys.com
- SNUG 2007 and 2008 presentations on low power