Power Gating for Ultra-low Leakage: Physics, Design, and Analysis

Ken Choi
Jerry Frenkil
Kimiyoshi Usami
Mark McDermott

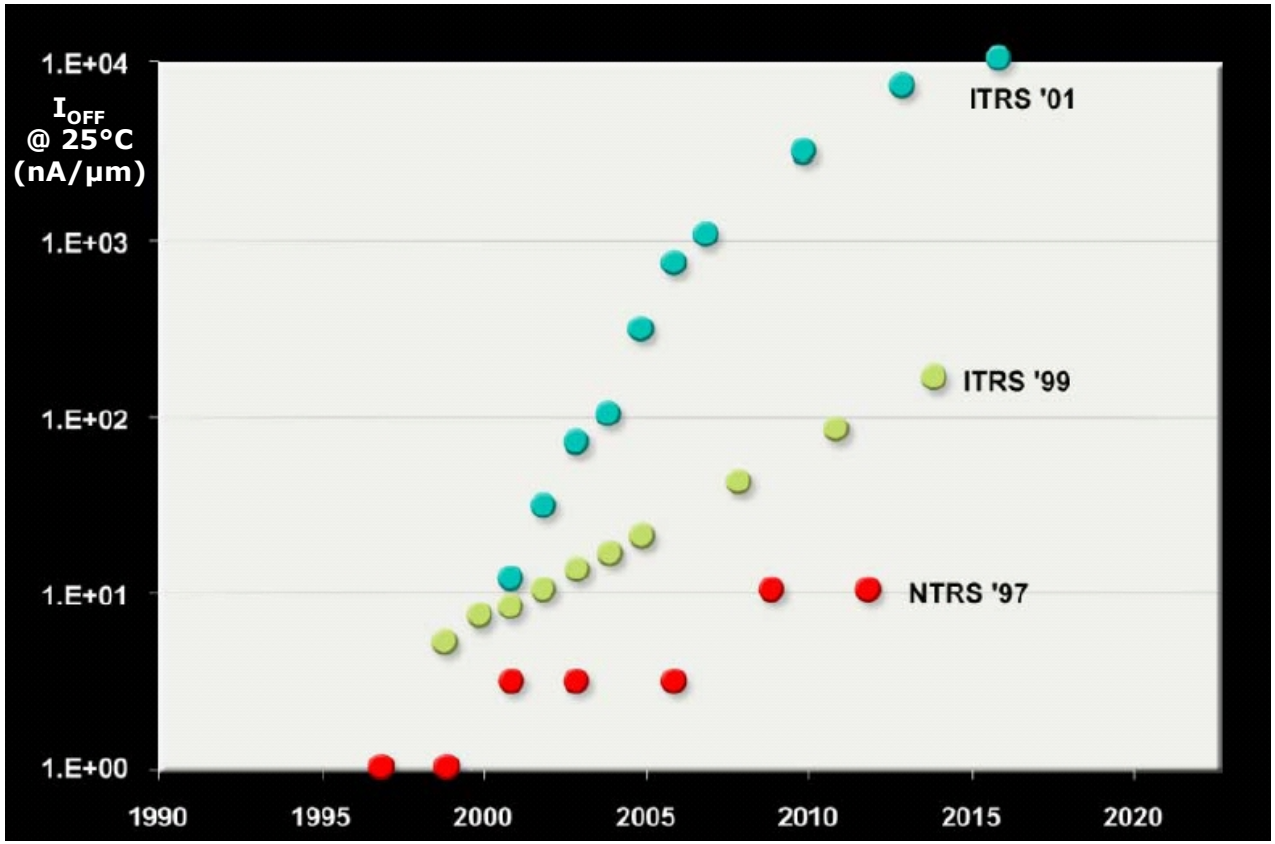
Spring 2017

The Leakage Problem: Market Demands

- Increased functionality with longer battery life….

- … but keep the battery small and lightweight

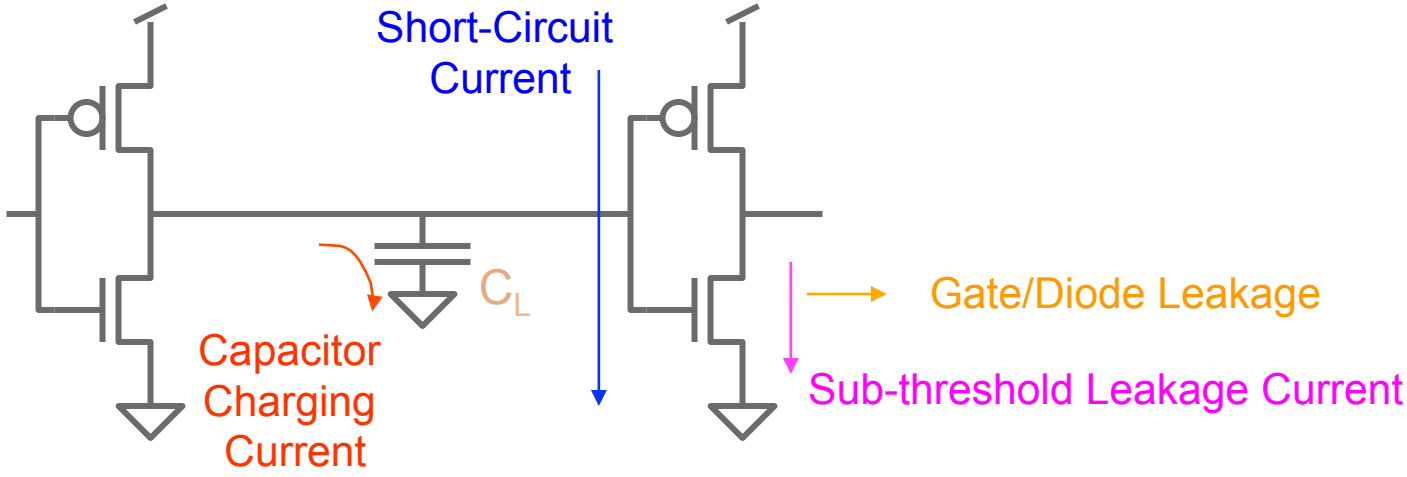
The Leakage Problem: Advanced Technology



"Leakage will become a major industry crisis, threatening the survival of CMOS itself"

ITRS2005 Executive Summary

Power Review



1. Capacitor or switching power (~50% of total power @ 90 nm)
   - Energy consumed is ½ CV^2 per transition
2. Short-circuit or internal power (10-15% of dynamic power)
   - When both p and n transistors turn on during signal transition
3. Sub-threshold leakage power (dominates when inactive)
   - Transistors do not turn off completely
4. Gate and diode leakage power (not negligible)
   - Gate oxide tunneling and parasitic source and drain diodes leak to substrate

Leakage Types

- **Standby Leakage**
  - Leakage from blocks in logical standby state
  - Reduction techniques
    - Coarse-grained MTCMOS power gating
    - Body bias control and power supply collapse

- **Active Leakage**
  - Leakage from operationally active blocks
    - Leakage from off-state instances within the active block
    - Gate leakage from active transistors
  - Reduction techniques
    - Fine-grained MTCMOS power gating
    - Multiple threshold voltage cell swapping
    - Long channel devices
    - Input vector control
    - Off-off stacking

Leakage Control Techniques

- **Process oriented techniques**
  - High Vt processes
  - Thick gate oxide processes
  - High-K gate oxide processes

- **Design oriented techniques**
  - Transistor sizing
  - Transistor stacking
  - Input vector control
  - Dual/Multi-Vt cell swapping
  - Body Biasing (VTCMOS)
  - Dynamic voltage scaling (DVS)
  - MTCMOS power gating

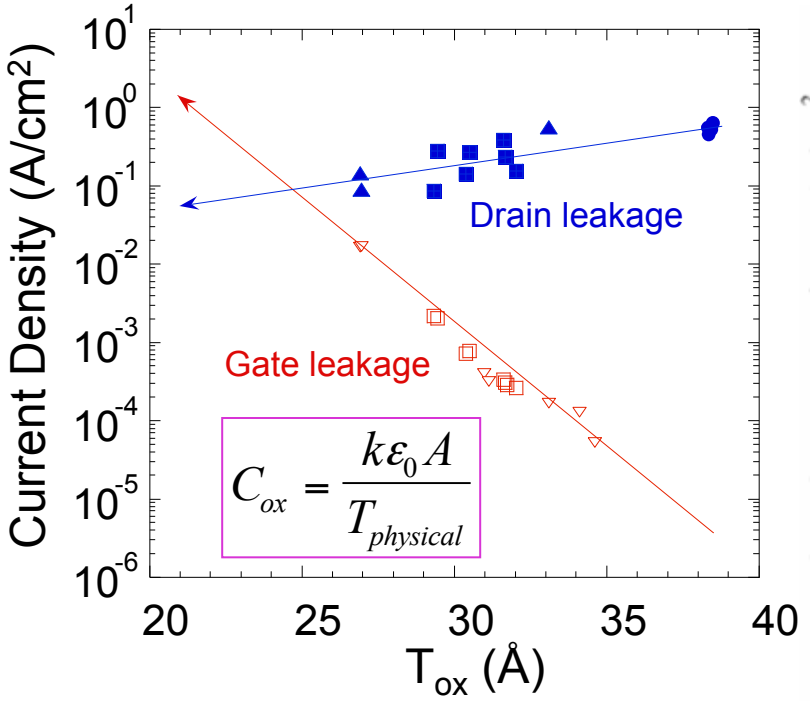
Low-Leakage Process Techniques

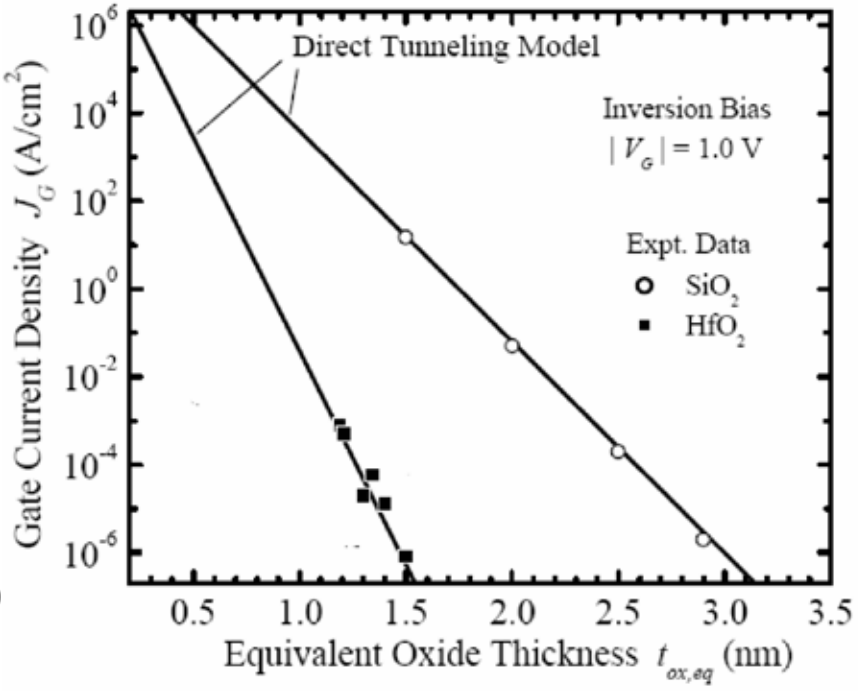
- **High Vt Processes**
  - Foundry companies often provide low-power technology libraries which have high-Vt standard cells for deep sub-micrometer technologies (below 90nm)
  - Subthreshold leakage is reduced at the expense of reduced drive currents

- **Thick gate oxide processes**
  - Achieve the same Cox with a thicker Tox by using high-k dielectric process

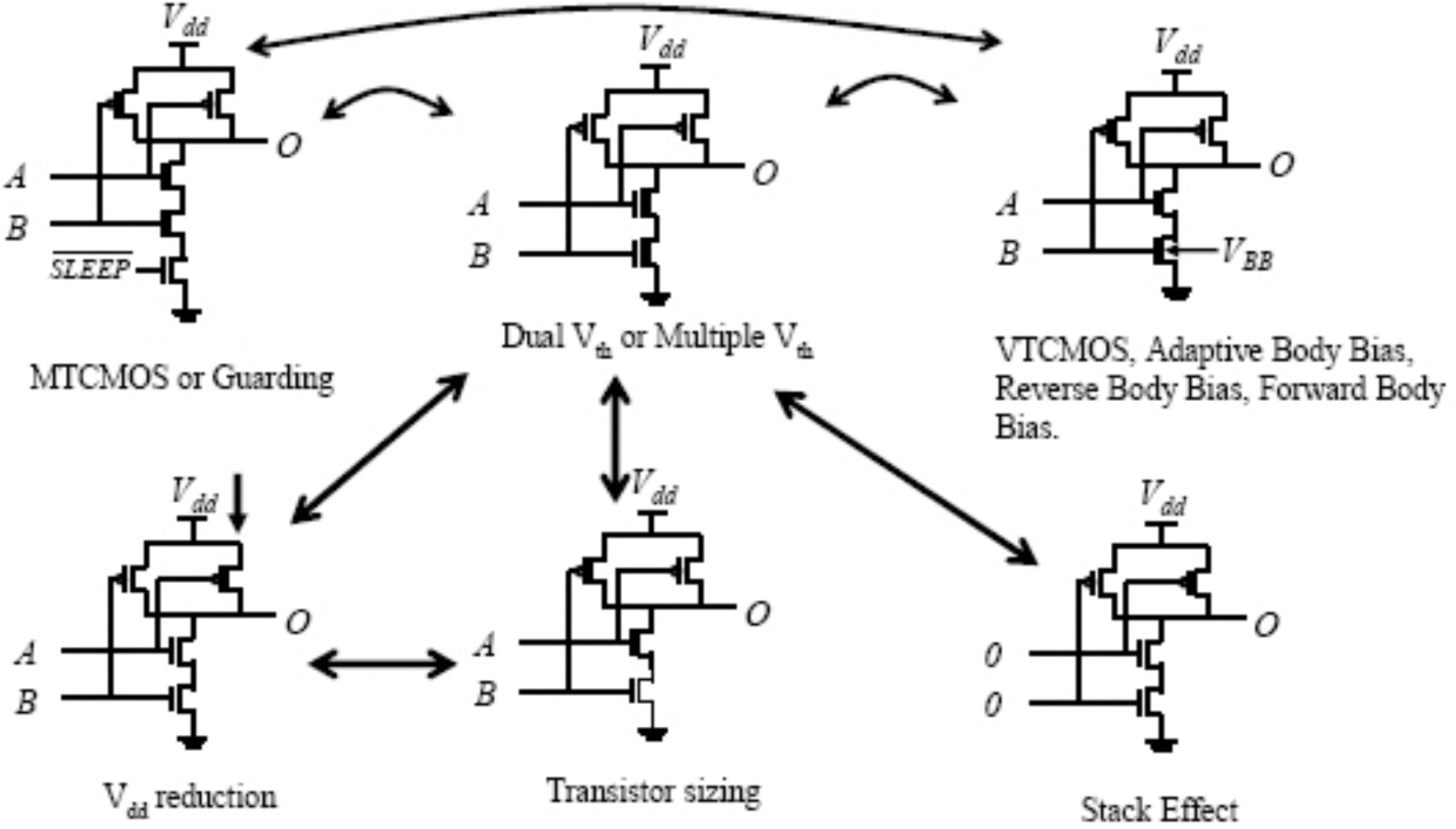
Low-Leakage Process Techniques

- **Gate leakage reduction with High-K gate dielectric**
  - As gate oxide thicknesses were thinned for 45nm nodes and below, the industry has moved to using high-k dielectrics
    - High-k gate dielectrics enable the process to achieve the same Cox with a thicker Tox

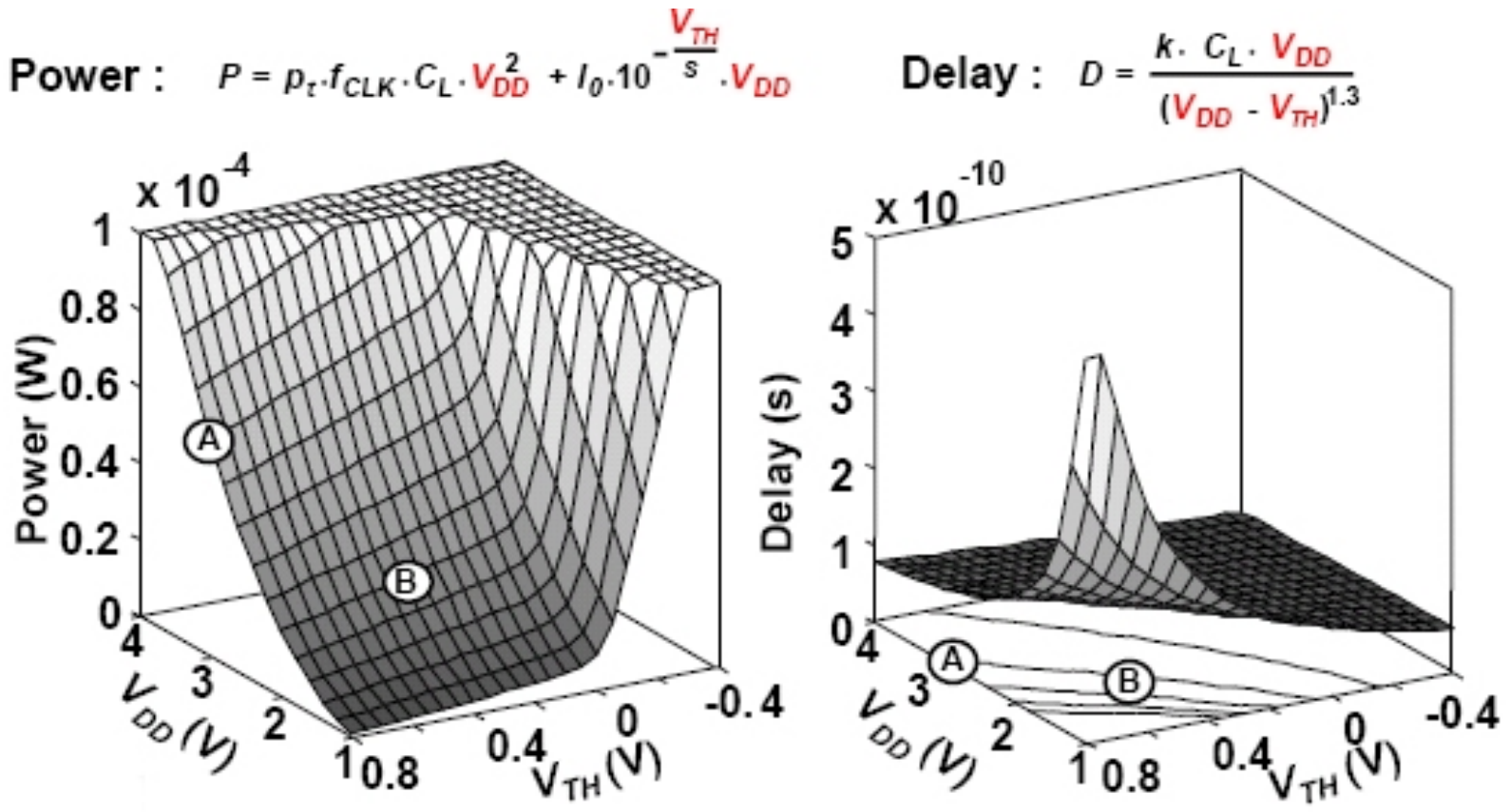




Low-Leakage Design Techniques Overview



Relationship between Power and Delay

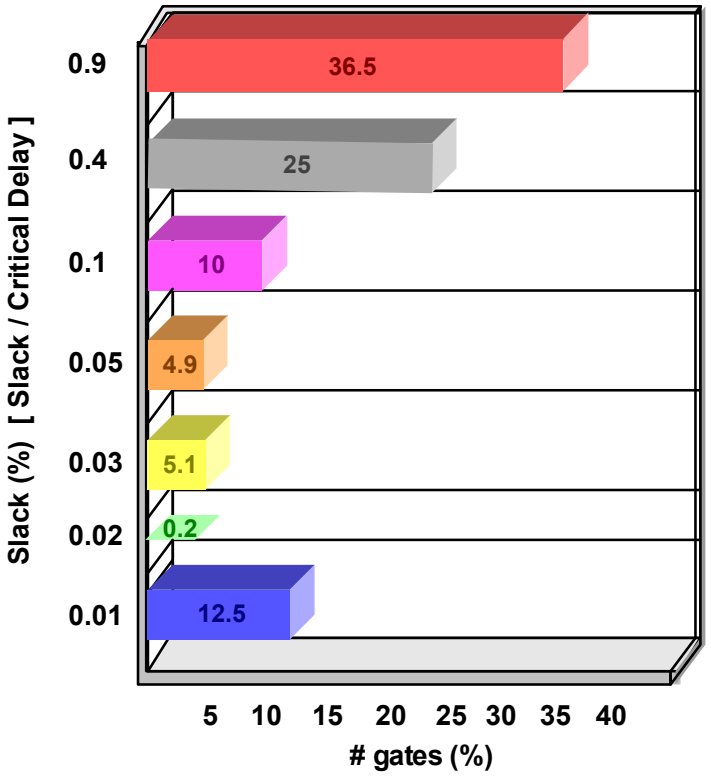


[Kuroda]

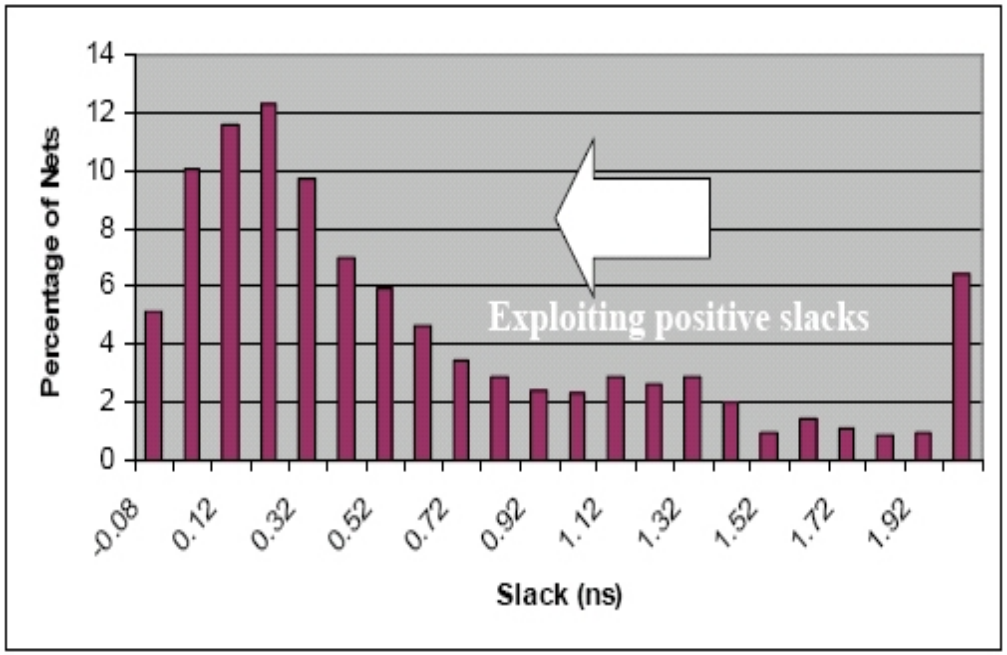
- **Power is reduced while delay is unchanged if both Vdd and Vth are lowered such as from A to B**

Rationale of Low Power Approaches

- More than 60% of gates have their slack larger than 25% of the critical path delay (ISCAS/ MCNC / ARM9 functional blocks)
- Exploiting the positive slack



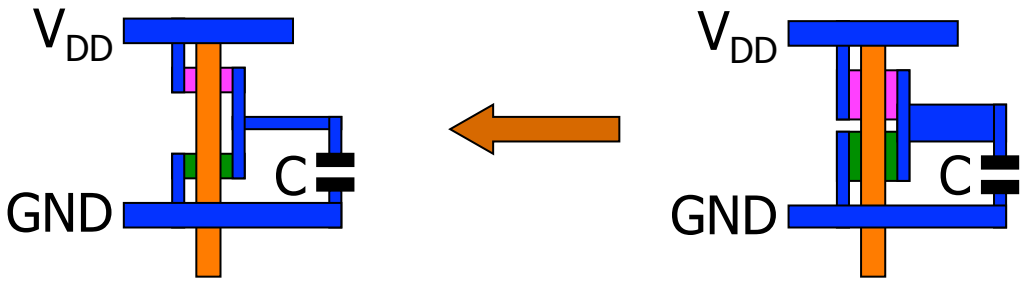
[ISLPED 03, K. Choi]



Trends of Low Leakage Techniques

- **Temporal granularity (transition time)**
  - Static techniques (infinity) to Dynamic techniques (ns)
- **Spatial granularity**
  - Coarse grained (full-chip level) to Fine grained (block or gate level)
- **Variable granularity**
  - One variable to dual / multiple variables

1. Transistor Sizing



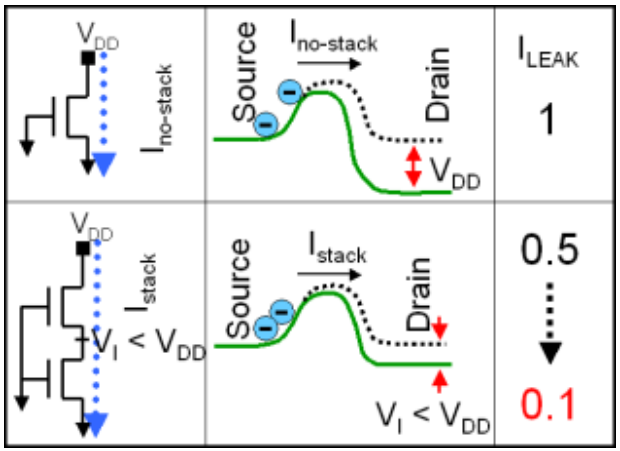
- Shorter width means lower leakage, more delay, and lower dynamic power
- Issues
  - Given delay constraints, finding the optimal size for minimum power
- Pros
  - Leakage reduction in both active and sleep modes
  - Fine-grained optimization is possible
  - Simultaneously optimize width with Vdd and Vth
- Cons
  - Design automation complexity is high
  - Limited amount of leakage reduction
  - Greater variability in drive strength

2. Transistor Stacking

$$I_{leakage} \propto 10^{(V_{GS}-V_{TH})/S}$$

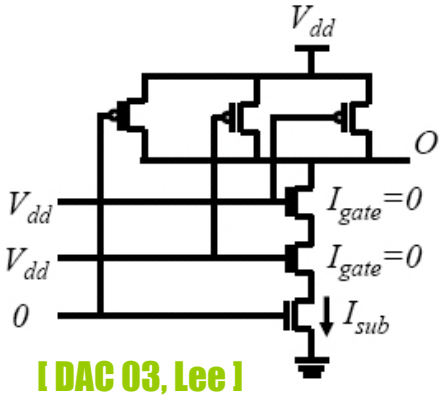
$$V_{TH} = V_{TH0} - \gamma V_{BS} - \lambda V_{DS}$$

Body effect DIBL



- **Main factors**
  - Negative Vgs, lowered signal rail (Vdd-Vs), lower DIBL (lower Vdd-Vs), and larger Body Effect (negative Vbs)
- **Issues**
  - Given logic topology, maximizing stacking
- **Pros**
  - Design complexity is low
  - No impact to technology scaling
  - Area and dynamic power overhead is generally low
- **Cons**
  - New cell library is needed

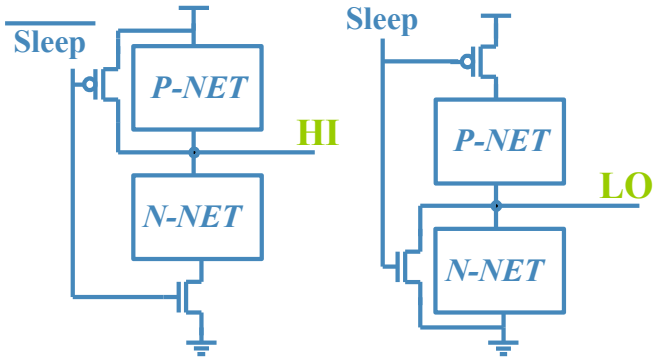
3. Input Vector Control (IVC)



[DAC 03, Lee]

State	I_{sub} (nA)	I_{gate} (nA)	I_{total} (nA)
000	0.382	0.000	0.382
001	0.709	6.339	7.048
010	0.709	1.275	1.984
011	5.626	12.677	18.303
100	0.676	0.000	0.676
101	3.804	6.339	10.143
110	3.804	0.000	3.804
111	28.273	19.015	47.288

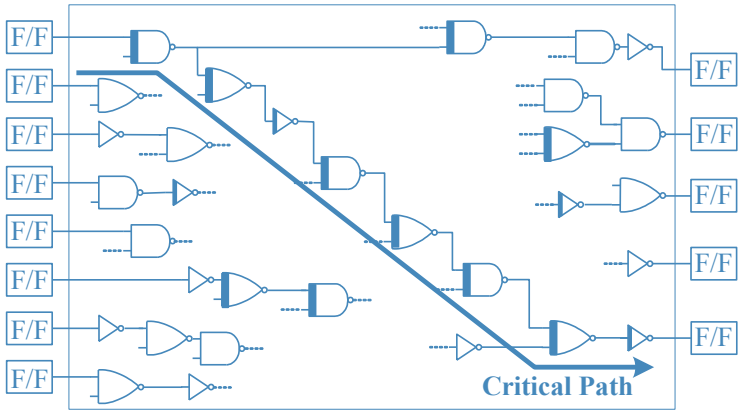
5x reduction

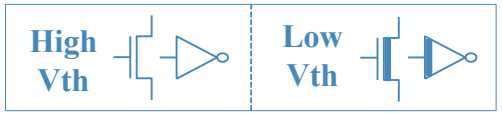


[VLSI Sym. 05, K. Choi]

- Different input vectors generate different leakage currents, so there must be optimal primary input vectors which lead to minimum leakage power in standby mode
- Issues
  - Given logic topology, finding optimal input vector for minimum power
- Pros
  - Easy to implement
  - Overhead is low in terms of delay, area, and dynamic power
  - No impact to technology scaling
- Cons
  - Relatively less effective on leakage reduction

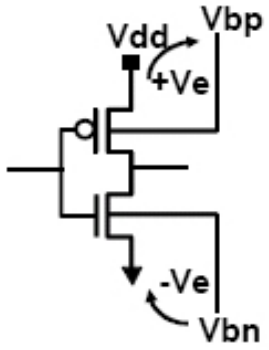
4. Dual/Multi-Vth Cell Swapping

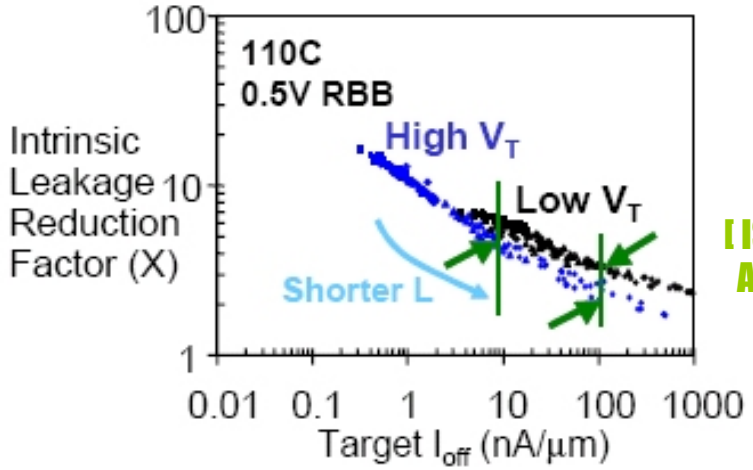




- Low-Vth cells on critical paths, High-Vth cells on non-critical paths
- Issues
  - Given delay constraints, finding the optimal Vths for minimum power without compromising the delay
- Pros
  - No area overhead
  - Leakage reduction in both active and sleep modes
- Cons
  - Critical paths are still leaky
  - Limited amount of leakage reduction
  - Weak technology scaling

5. Body Biasing (VTCMOS)

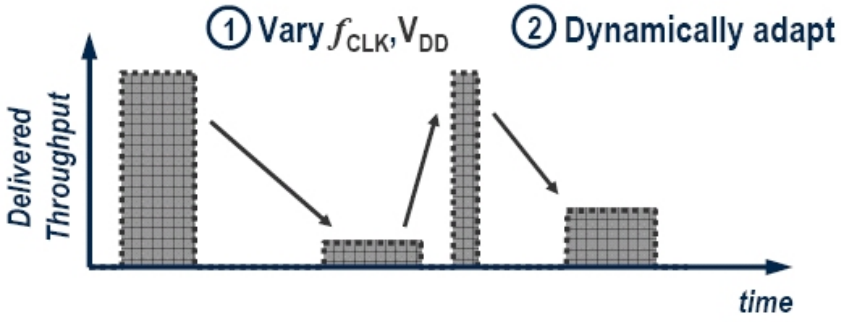


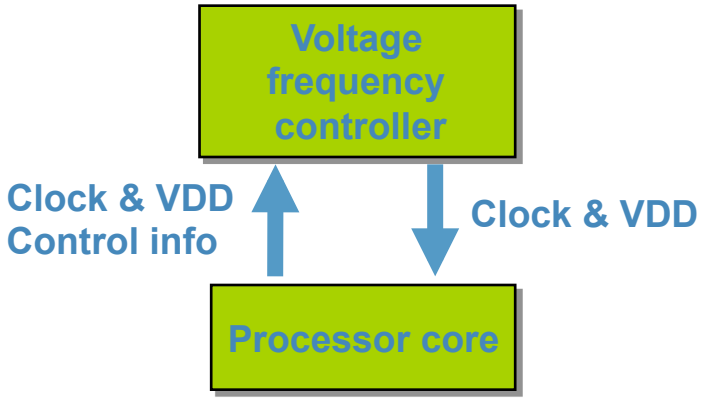


[ISLPED 01, A. Keshavarzi]

- **Reverse Body Biasing (RBB)**
  - Active Mode: No Bias (= Low Vth) and Sleep Mode: RBB (= High Vth)
- **Forward Body Biasing (FBB)**
  - Active Mode: FBB (= Low Vth) and Sleep Mode: No Bias (= High Vth)
- **Issues**
  - Given delay constraints, finding the optimal biasing for minimum power
- **Pros**
  - Useful as a post-silicon tuning method for yield enhancement
  - Dynamic approach is suitable for use with DVS
- **Cons**
  - Less effective at shorter channel length and lower Vth
  - Process complexity is high (Triple well is needed)

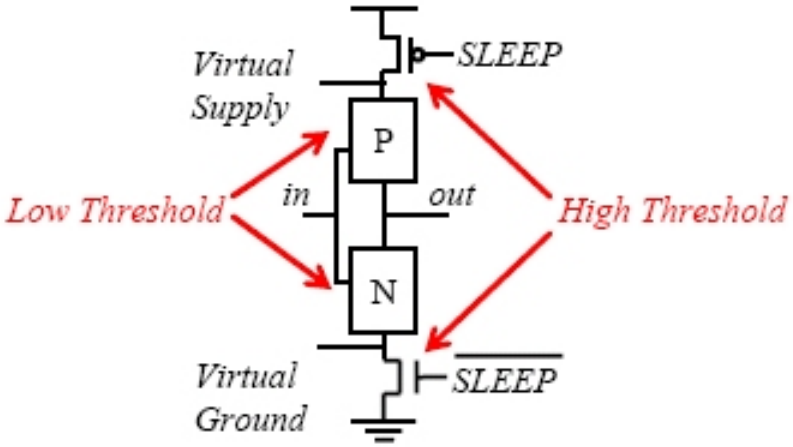
6. Dynamic Voltage Scaling (DVS)





- Dynamically scale energy/operation with throughput and exploit data dependent computation times to vary the supply voltage
- Issues
  - Given delay constraints, finding the optimal Vdd and frequency for minimum power
- Pros
  - Very effective to reduce total energy (dynamic and leakage power)
- Cons
  - Difficult to implement
  - Additional Control circuits are needed for monitoring and scaling

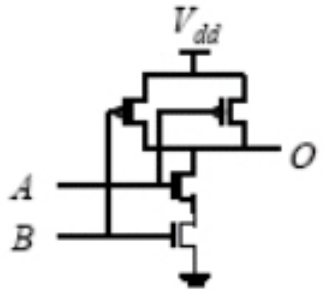
7. MTCMOS Power Gating



- High Vth transistors gate leakage power during sleep mode for low Vth circuits
- Issues
  - Given delay constraints, finding the optimal clustering and the proper switch size for minimum power
- Pros
  - Most powerful leakage control scheme (10x-100x reduction)
  - Dynamic fine-grained approach can be used to reduce active leakage
- Cons
  - Implementation complexity is high

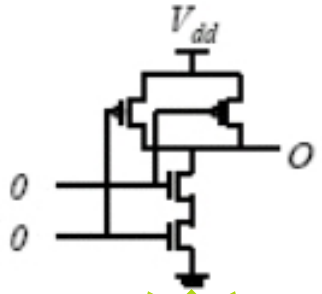
Leakage Reduction Comparison

1. Transistor Sizing



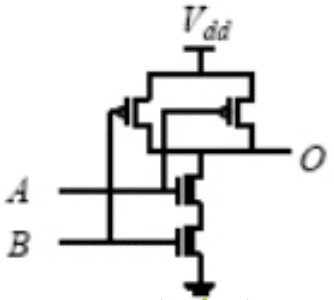
1.2X – 2X

2. Transistor Stacking / 3. IVC



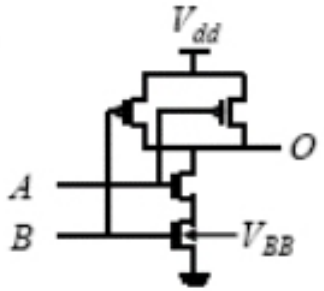
2X – 5X

4. Dual/Multi-Vth Cell Swapping



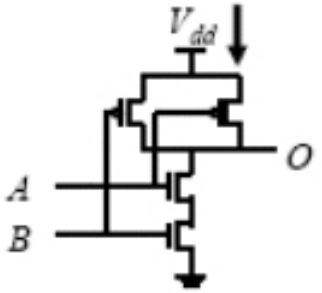
1.5X – 3X

5. Body Biasing (VTCMOS)



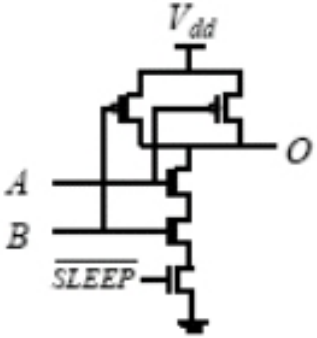
5X – 10X

6. DVS (Dynamic Voltage Scaling)



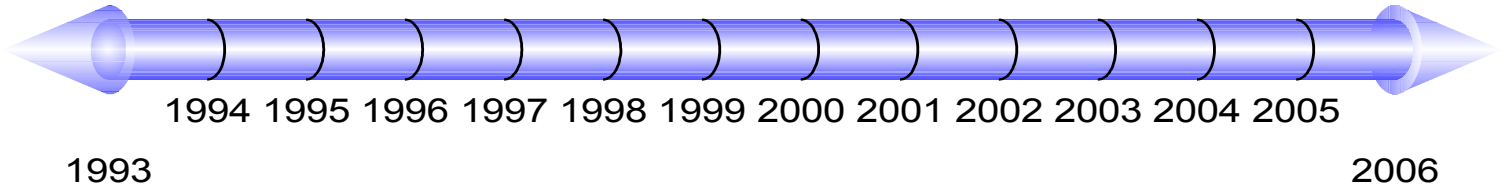
3X – 5X

7. MTCMOS Power Gating



10X – 100X

Historical Footprint for Power-Gating



Switched Source-Impedance MOS ➡ MTCMOS ➡ Power Gating

[1993, Itoh, Switched Source-Impedance MOS]

[1995, Mutoh, MTCMOS]

[1998, Kao, Mutual Excusive]

[1999, Mutoh, ACM]

[2000, Sakurai, BGMOS and Super-Cutoff CMOS]

[2002, Usami, Selective MTCMOS]

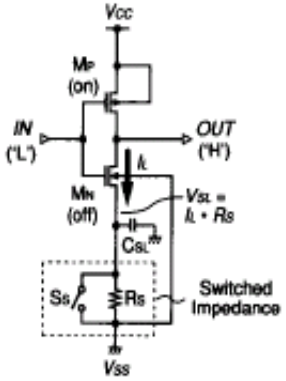
[2002, Anis, Gate-Clustering]

[2003, S.Kim, Mode Transition Method]

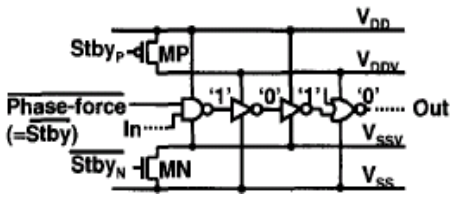
[2003, Sakurai, Zigzag SCCMOS]

[2005, K. Choi/Sakurai, Optimal Zigzag]

[2007, Frenkil/Venkatraman, CoolPower]



[Switched Source-Impedance MOS]



[Zigzag SCCMOS]

Power Gating Variants

- Power Gating current reduction
- Coarse-grained Power Gating
- Fine-grained Power Gating
- Selective Power Gating
- Super-cutoff Power Gating
- Zig-Zag Power Gating
- Advanced Power Gating schemes

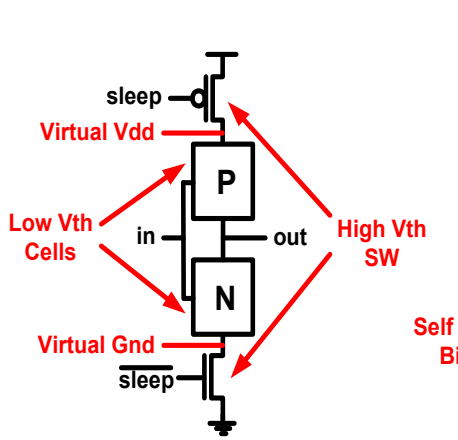
Power Gating Current Reduction

- **Weak Inversion Current Reduction**

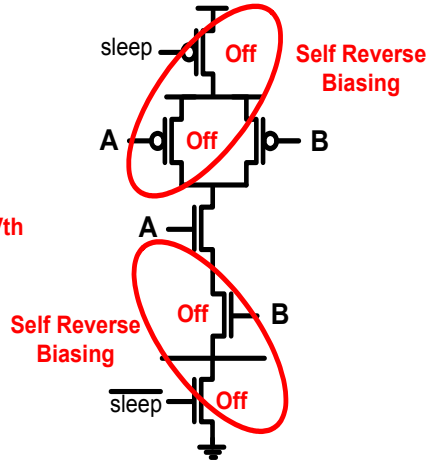
$$\begin{cases} I_{LEAK} \propto 10^{(V_{GS}-V_{TH})/S} \\ V_{GS} - V_{TH} = V_{GS} - \left(V_{TH0} - \gamma V_{BS} - \delta V_{GS2} - \lambda V_{DS}\right) \end{cases}$$

Subthreshold swing

Negative V_{GS} | High V_{TH} | Body bias | Double gate | DIBL



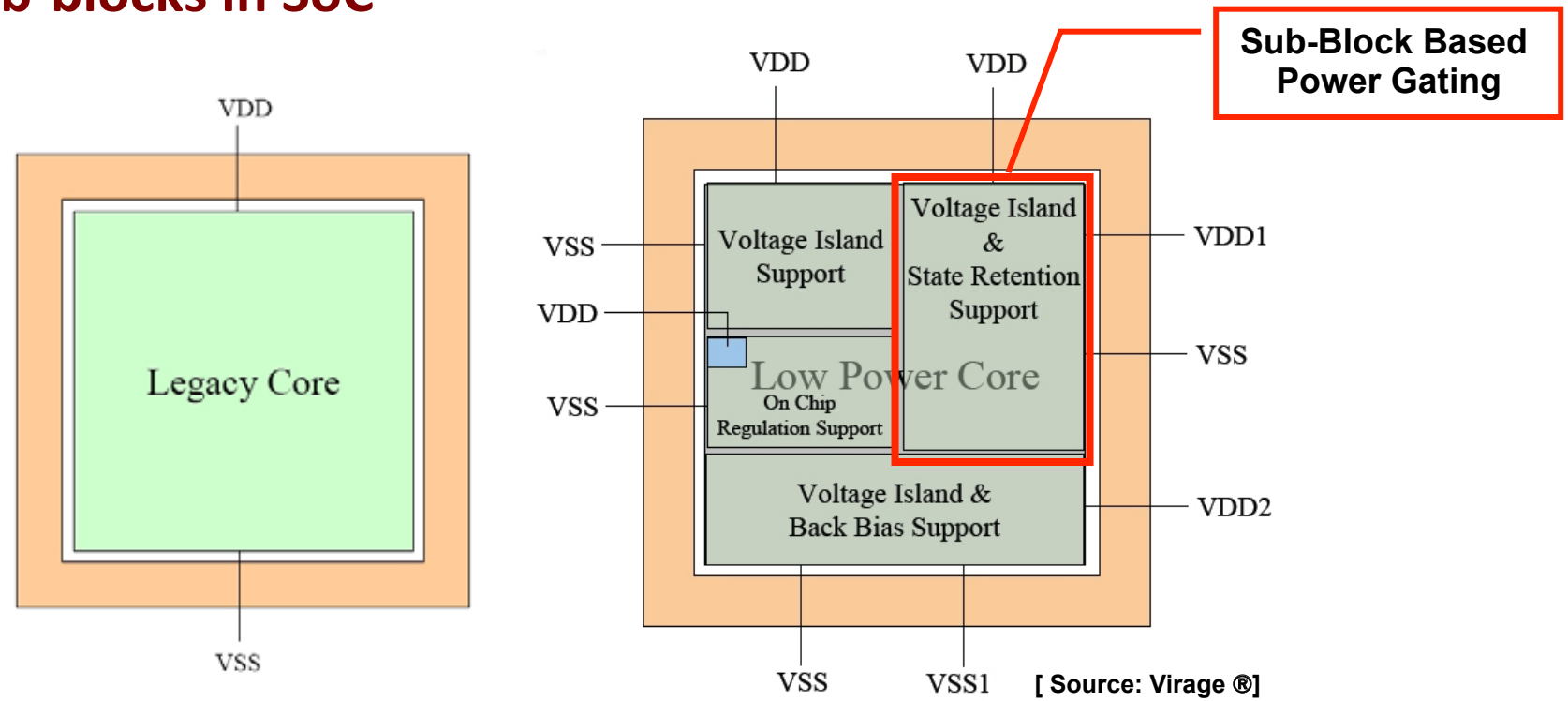
High-Vth Switch



Self Reverse Biasing (Off-Off Stacking)

Coarse Grained Power Gating

- Power Gating switch cell is a part of the power distribution network
- Power management block controls the power turn on and off of sub-blocks in SoC



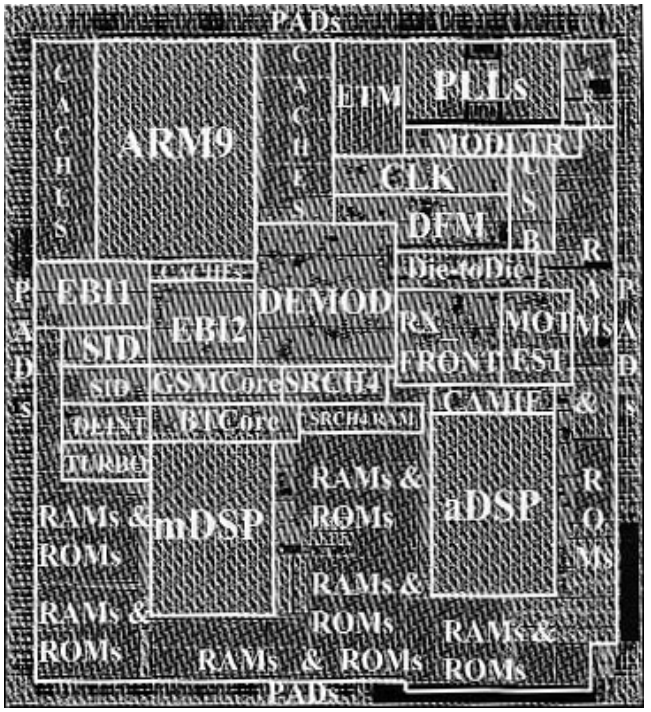
Traditional Chip Design

Modern Low-Power Chip Design

Fine-Grained Power Gating

- **Switch-in-Cell**
  - Switching transistor is encapsulated as a part of the standard cell

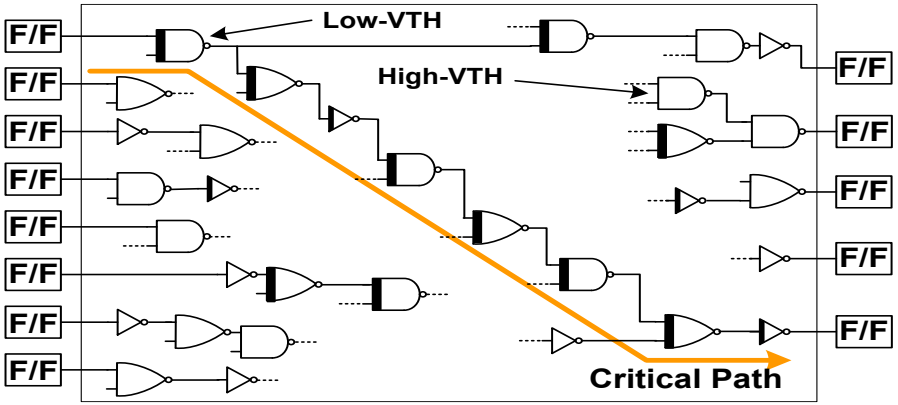
Diagram	NAND2*(non-footswitch)*	fs_NAND2*(footswitch)*
SYMBOL	a, b, nz, HVT	a, b, nz, LVT, slpb, HVT
SCHEMATICS	a, b, nz	a, b, nz, LVT, slpb, HVT
LAYOUT	V_{DD}, a, b, nz, nwell, V_{SS}	1.25X, V_{DD}, slpb, a, b, nz, nwell, LVT Layer, V_{SS}



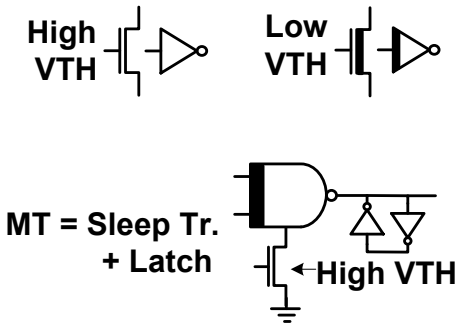
[Source: ISSCC. 05, G. Uvieghara]

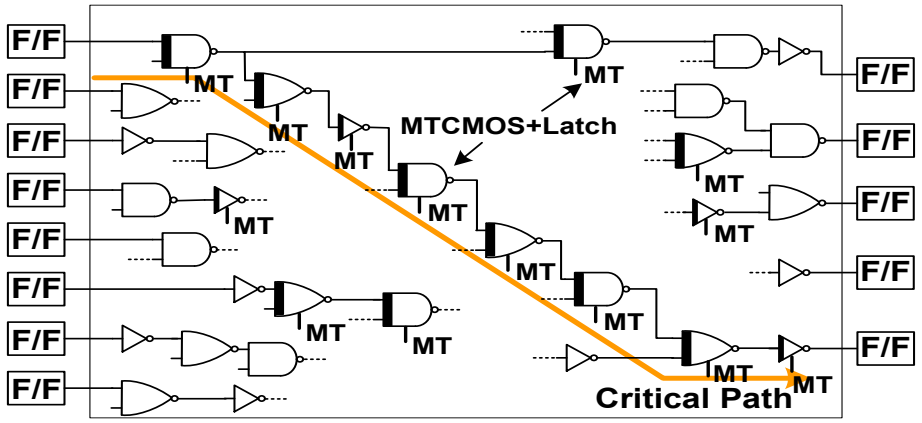
Selective Power Gating

- Selective MTCMOS
  - Low-Vt cells on critical path and high-Vt cells on non-critical paths
  - Apply Power Gating only for the low-Vt cells



(a) Dual VTH Scheme

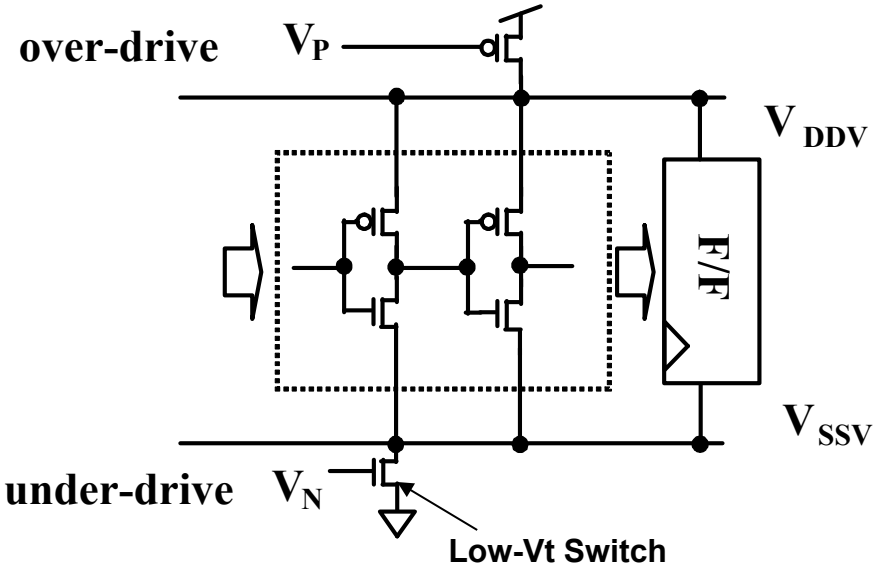


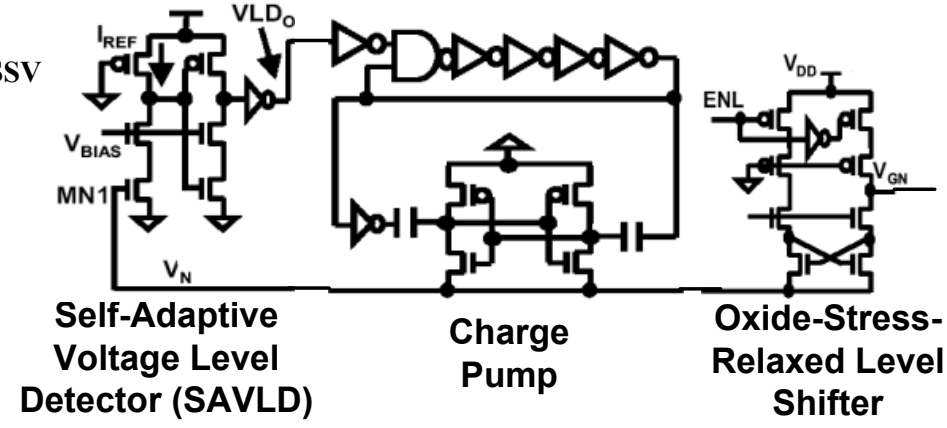


(b) Selective MTCMOS Scheme

Super Cutoff Power Gating

- Use gate voltage of the PG switch transistor to reduce leakage instead of using high-vt switch transistor
  - Overcome technology scaling issue with high-Vt switch

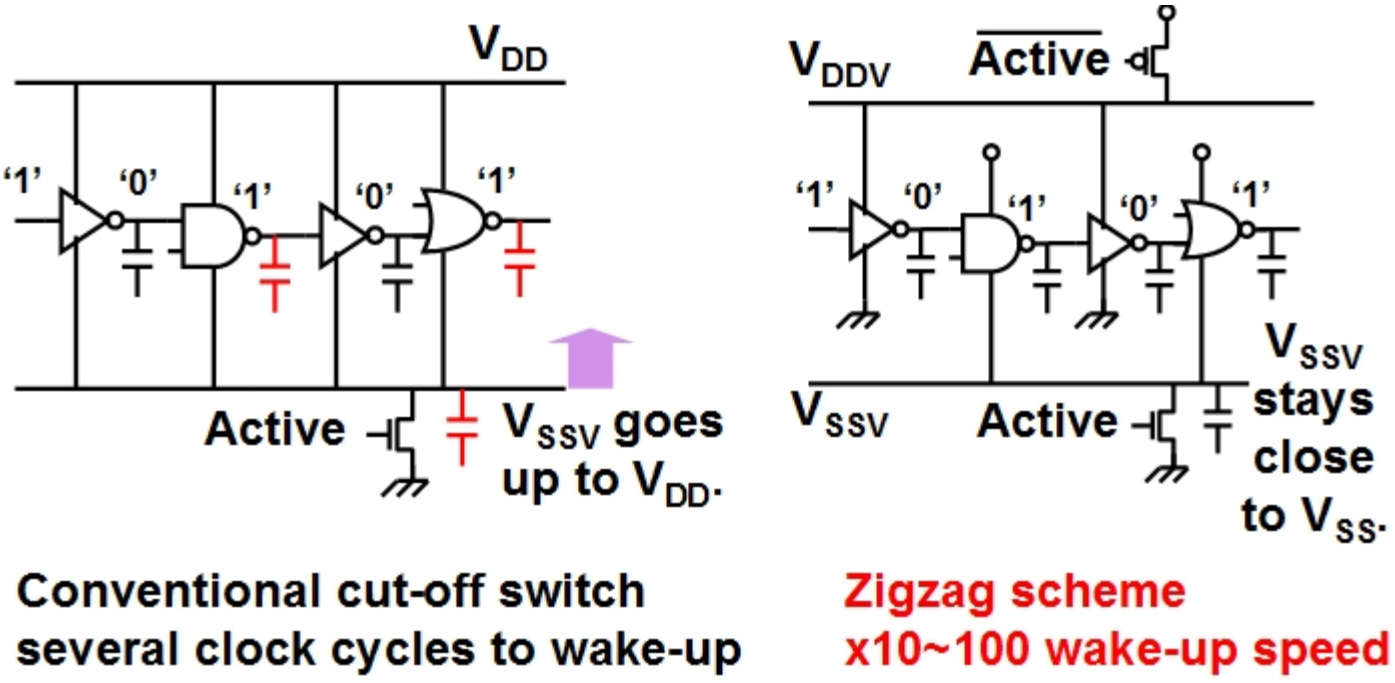




[V_N (negative voltage) generator for NMOS cut-off switch with self-adaptive voltage level detector, charge pump, and oxide-stress-relaxed level shifter, T. Sakurai, 2002]

Zig-Zag Power Gating

- Use header and footer switches in a ZigZag, alternating on successive instances
  - Overcome long wakeup time for Power Gating scheme

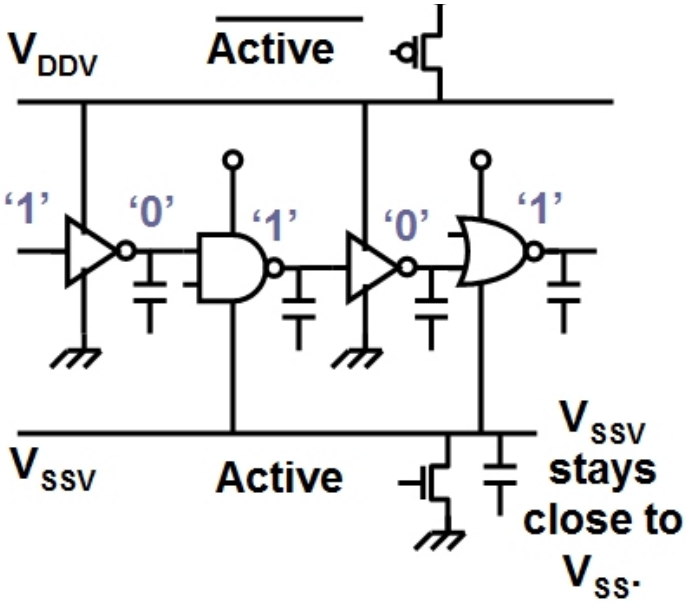


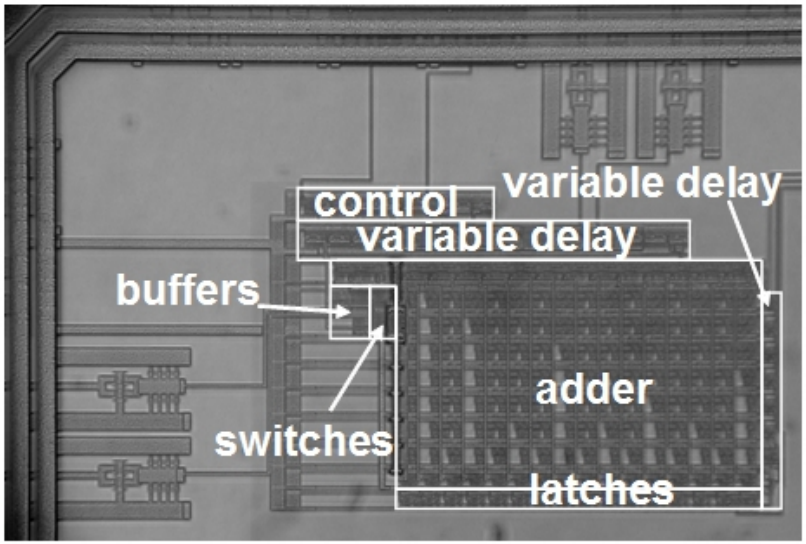
Conventional cut-off switch
several clock cycles to wake-up

Zigzag scheme
x10~100 wake-up speed

Zigzag Super Cut-off CMOS

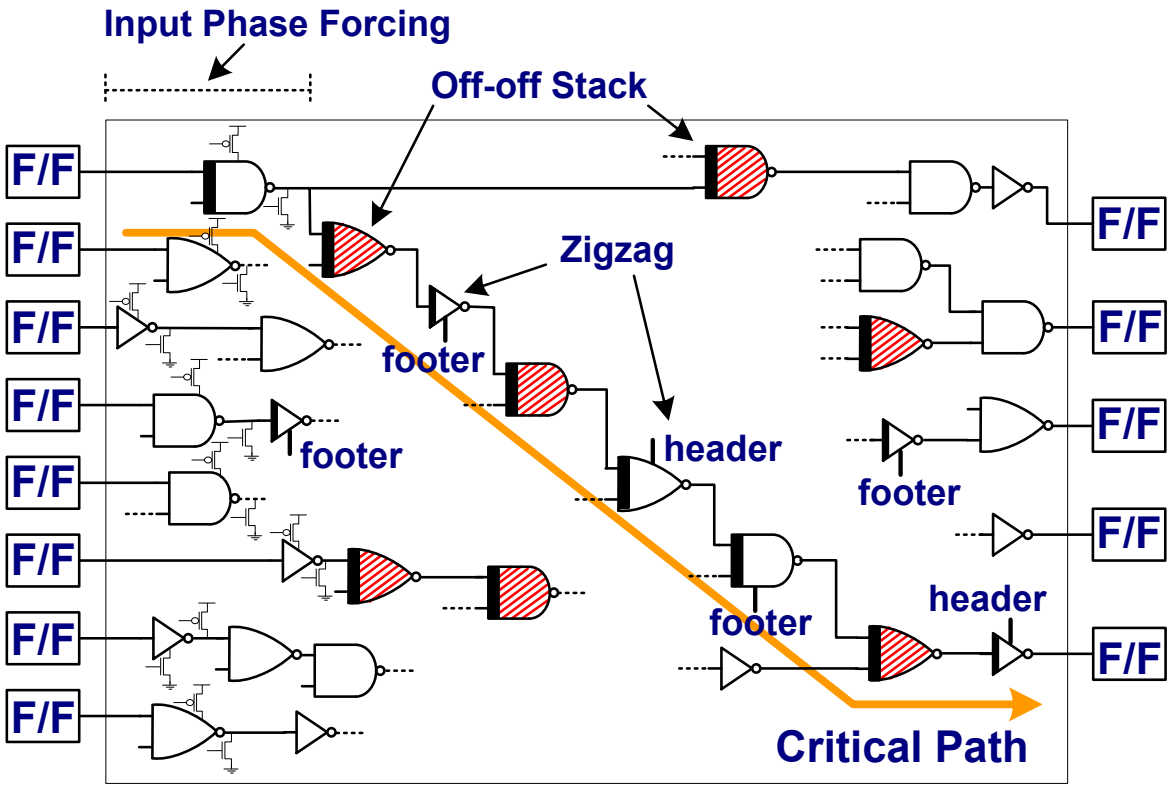
- Combine Zigzag Power Gating and Super Cut-Off Power Gating
  - Overcome long wakeup time and technology scaling issue for general Power Gating scheme



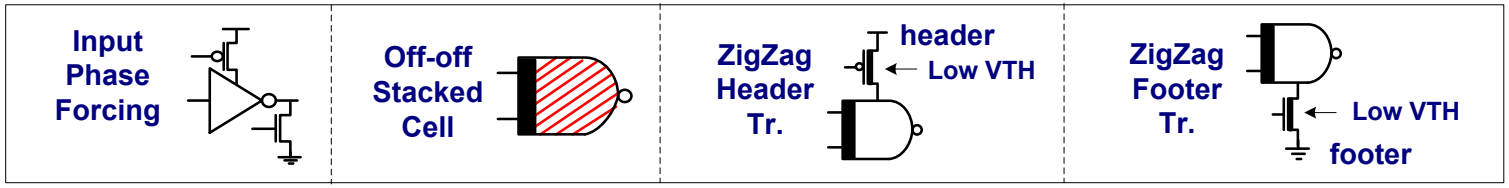


Node	Wakeup time (1)	Delay of the adder (2)	(1):(2)
.18µm	0.3ns	1.9ns	16%
J. Tschanz et al., JSSC, pp. 1838-1843, Nov. 2003.			200%

Optimized Zigzag Super Cut-off CMOS



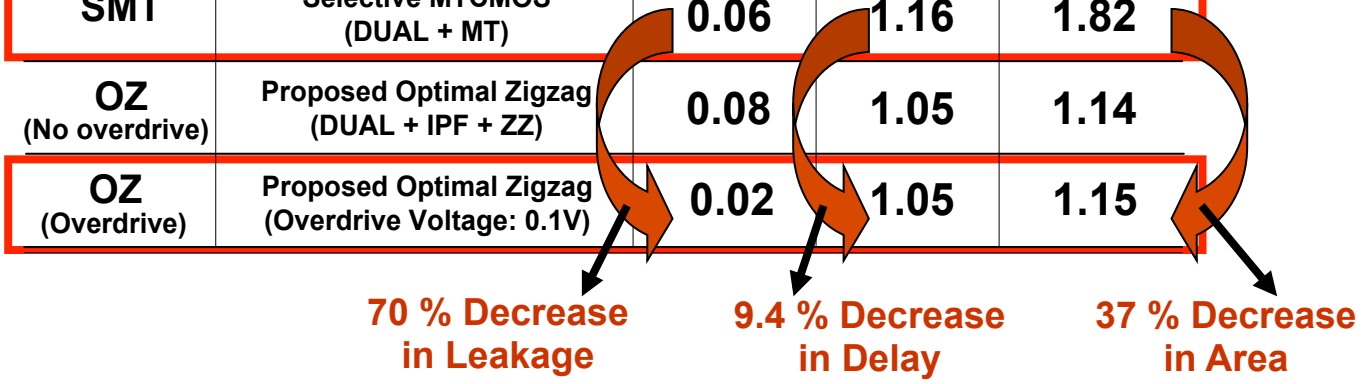
- Main benefits
  - Optimized IPF (36% of gates on critical path are off-off stacked on avg)
  - No keeper circuit required (zig-zag)
  - Worst-case guarantee (power-gating switch is not shared)
  - Fast wake-up
  - Small ground bounce



Results and Summary

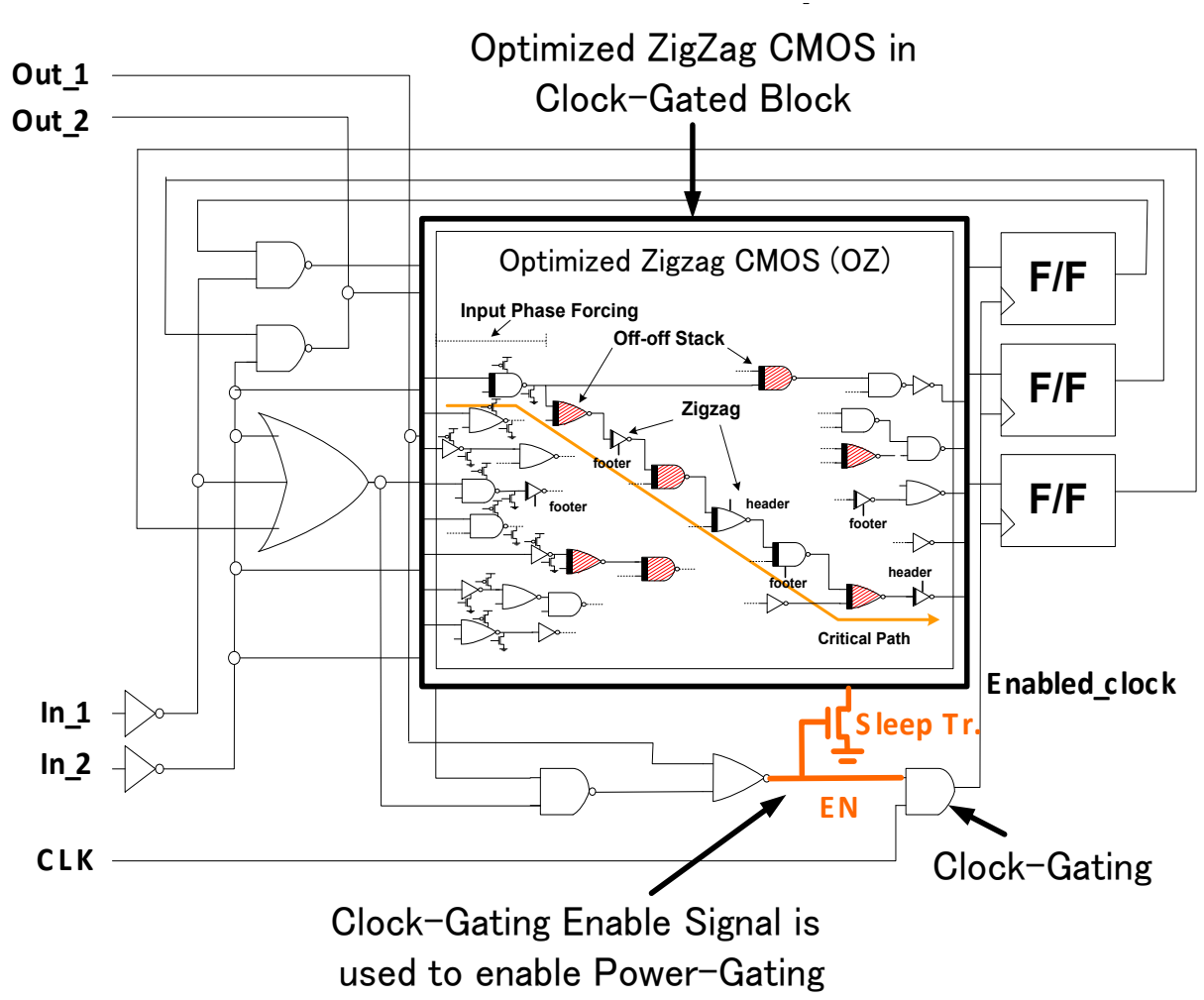
- 10 benchmark: c432, c499, c880, c1355, c1908, c2670, c3540, c5315, c6288, c7552
- Tech.: 65nm (Vdd : 0.9V, diff-VTH: 0.1V, sleep tr. size: 2W, overdrive : 0.1V)
- % of Low VTH cells / total cells: 23.4 % on average (min.: 9.19% / max.: 48.01%)
- % of IPF on critical path : 36.4 % on average (min.: 28.01% / max.: 47.05%)
- % of IPF over Low Vth cells : 20.1 % on average (min.: 11.9% / max.: 30.12%)

Power Gating Scheme	Description	Normalized by ORG		
		Power(leak)	Delay	Area
ORG	All Low-VTH (Original Circuit)	1.0	1.0	1.0
HVT	All High-VTH (Diff-Vt: 0.1 volt)	0.06	1.51	1.0
DUAL	Dual VTH (% of Low VTH: 23.4%)	0.32	1.00	1.0
MT	Gate-Level MTCMOS (2W-sleep tr. sizing)	0.07	1.10	2.07
IPF	Input Phase Forcing Only (Off-off Stacking only)	0.37	1.02	1.03
SMT	Selective MTCMOS (DUAL + MT)	0.06	1.16	1.82
OZ (No overdrive)	Proposed Optimal Zigzag (DUAL + IPF + ZZ)	0.08	1.05	1.14
OZ (Overdrive)	Proposed Optimal Zigzag (Overdrive Voltage: 0.1V)	0.02	1.05	1.15

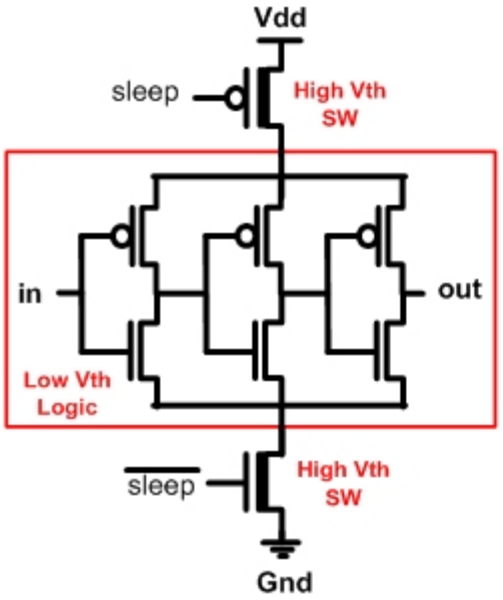


Advanced Fine-Grained Power Gating

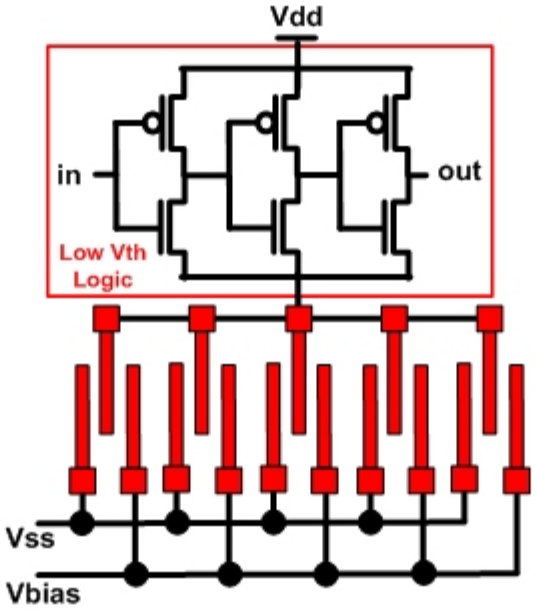
- Use Clock Gating enable signal to control Power Gating circuits
  - Requires fast wakeup and small speed loss



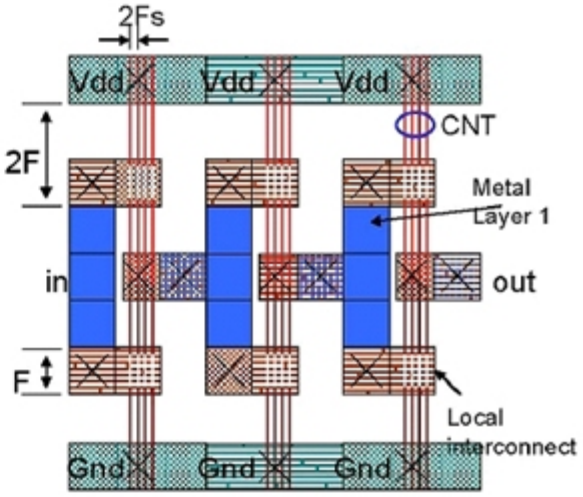
Advanced Power Gating Scheme with Nanotubes



(a) Low-Power Design for CMOS (MTCMOS Power Gating, 3 Stage Inverter, 10x leakage power reduction in 65 nm)



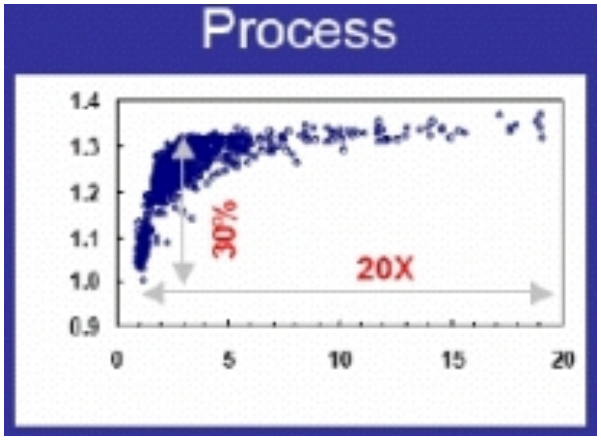
(b) Hybrid PowerGating Design with CMOS and CNT (Carbon Nano Tubes)

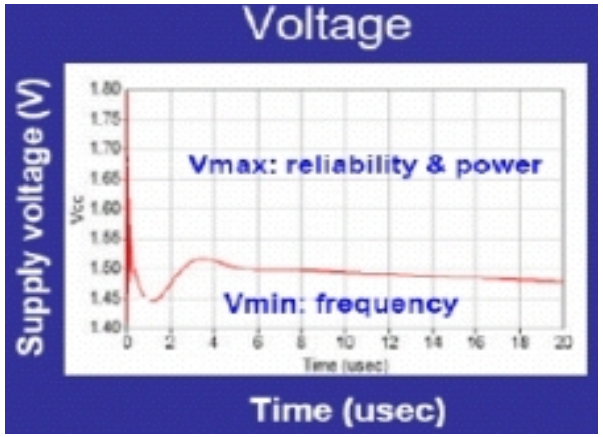


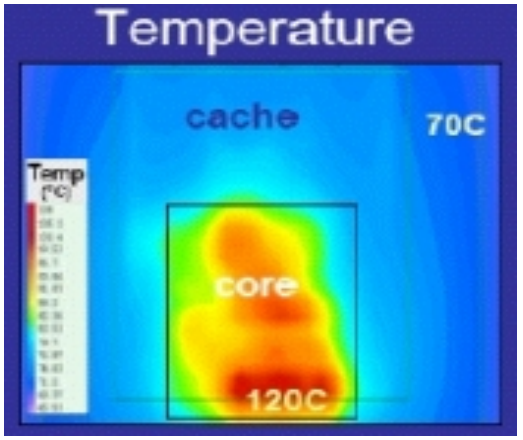
(c) 3 Stage CNT FET Inverter (100x leakage reduction and 10x fast speed over 65 nm CMOS design)

One Last Thing . . .

- PVT (Process, Voltage, and Temperature) variations should be very carefully considered with the leakage reduction techniques







Quick Summary: Leakage Physics and Control

- Leakage has become a very tough problem
- Leakage worsens with each new process generation
- Designers must solve the leakage problem
  - Process engineers do not have realistic solutions on the horizon
- Multiple design solutions exist for reducing leakage
- MTCMOS power gating will become the dominant solution for ultra-low Leakage.

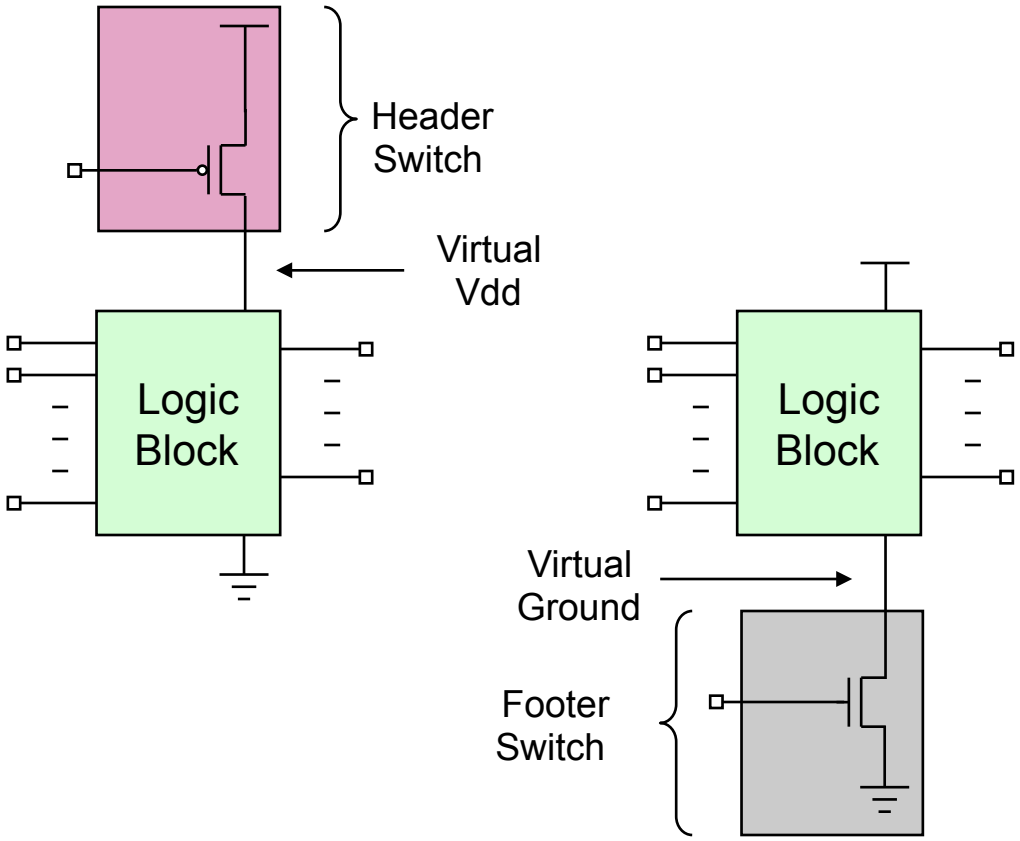
Power Gating Design Issues

- Logic Design
- Physical Design
- Design Verification
- Tools, Flows, and Methodologies

Power Gating Issues – Logic Design

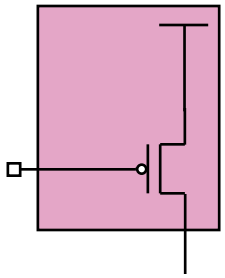
- **High-side or Low-side switching**
  - Use of header switches to gate the power supply or footer switches to gate ground
- **Domain partitioning**
  - Partitioning the design into sections with always-on power and those with gated power
- **Signal interfacing**
  - Conditioning the outputs of power-gated blocks such that they can not float
- **State retention**
  - Retaining some or all of the internal state when power is removed

High-side or Low-side Switching

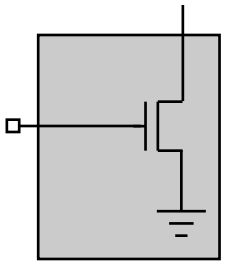


High-side or Low-side Switching

- **Header: PMOS device used to gate power supply**
  - Advantages
    - Less susceptible to gate leakage than footers
    - Can be used with multiple supply voltages
  - Disadvantages
    - ~2X larger than equivalent resistance footer



- **Footer: NMOS device used to gate the ground**
  - Advantages
    - ~1/2 the size of equivalent resistance header
  - Disadvantages
    - Does not reduce gate leakage



Refresh with contemporary CMOS technology.

1. the 2X area is no longer applicable; PFET IDSATs are now comparable to NFETs.
2. The virtual node will get closer to GND with a PFET pwr gate vs. an NFET pwr gate?
3. Switching performance is better? No pull-down de-biasing with solid GND grid.
4. PFET switches allows the use of different VDD values to different blocks in the same SoC design.

Domain Partitioning

- **Which modules must be always powered on?**
  - Assign these modules to the always-on domain
- **Which blocks can be power-gated?**
  - Assign these to the power-gated domain
    - Switch enables must be available
  - How much leakage will power gating save? Is it worth the effort?
    - How much does the module leak when powered?
    - How long is the module idle time?
- **How many separate modules are worth power-gating?**
  - Only a few large domains: Coarse-grained power gating
  - A large number of very small domains: Fine-grained power-gating

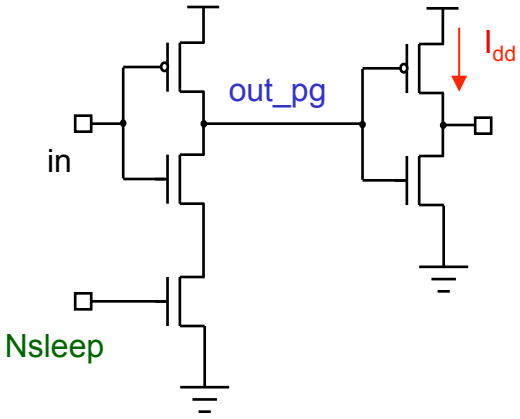
Domain Partitioning: Coarse / Fine Grained

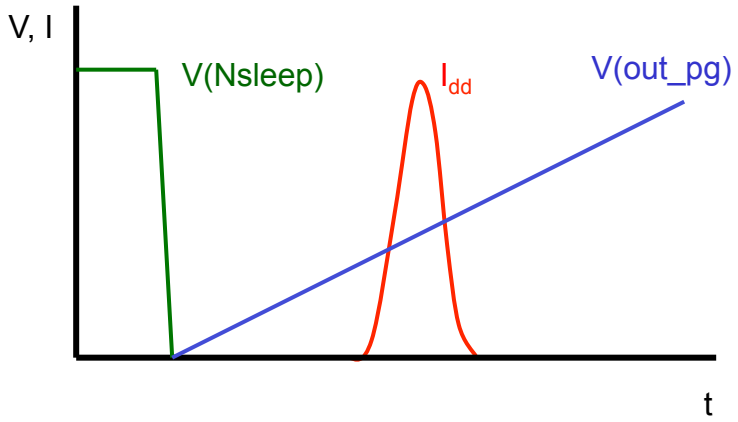
- **Coarse Grained**
  - Typically used to reduce *standby* leakage power
- **Fine Grained**
  - Useful for reducing *active* leakage
    - Small logic modules can be powered off, even while rest of module is active
  - Finding / creating enables can be difficult
  - Physical design may be difficult
  - Number of isolation cells needed can be substantial

Signal Interfacing

- **Power-gated logic outputs become high-impedance drivers when switches are opened….**

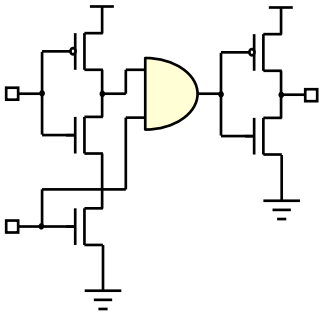
- **… but high-impedance drivers must not drive powered logic**
  - **High impedance nets can float to intermediate voltages resulting in large through currents in receiving logic**



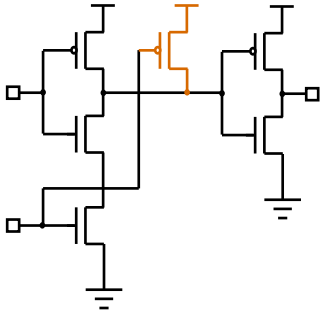


Signal Interfacing – Different Types

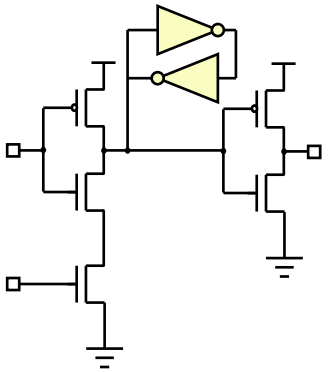
- **Buffers: AND, OR gates**
  - AND used with footers; pulls signal low
  - OR used with headers; pulls signal high



- **Pull-up, Pull-down**
  - Pull-up used with footers; pulls net high
  - Pull-down used with headers: pulls net low

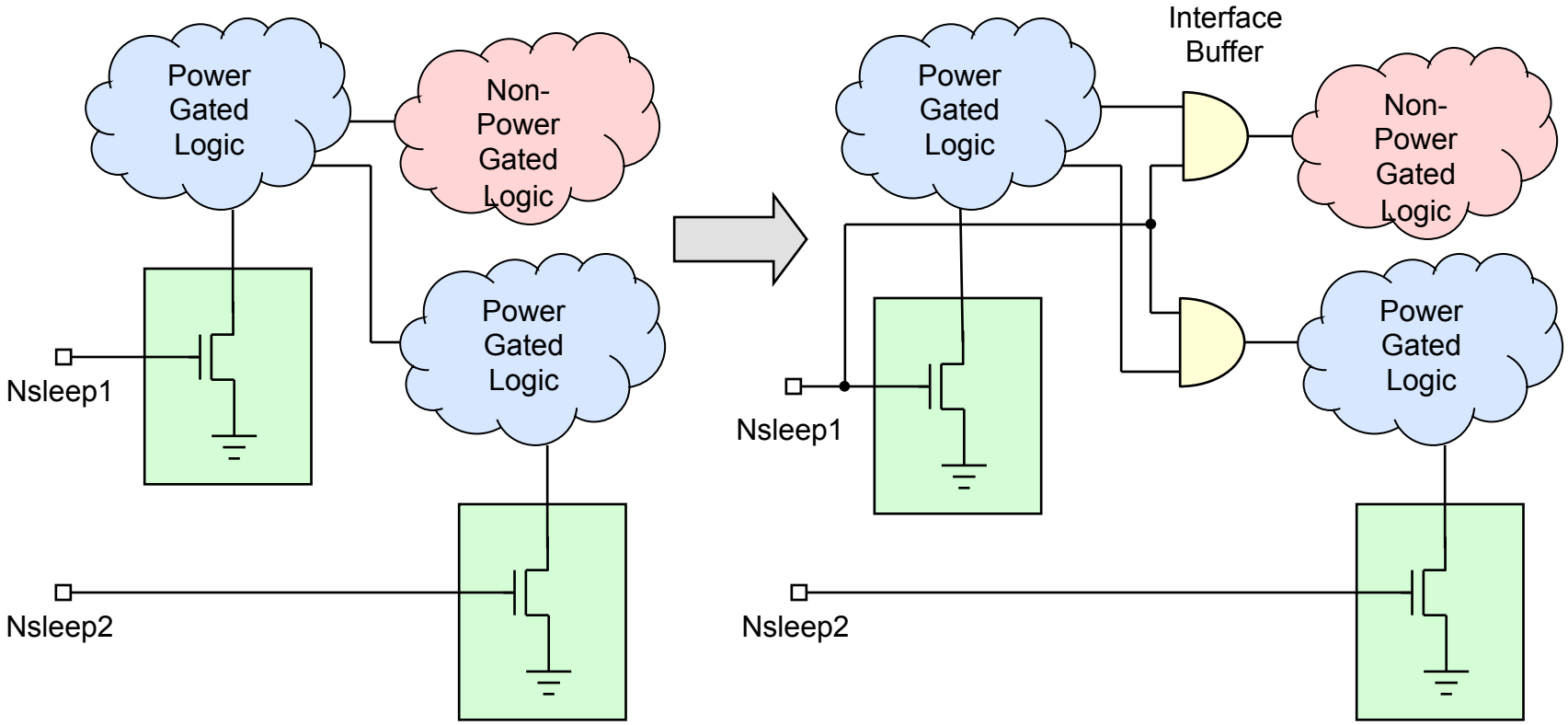


- **Half-latch**
  - Maintains most-recent state of output



Signal Interfacing for Multiple Domains

- **All outputs of power-gated blocks must be isolated**



State Retention

- Remember critical state so that when power is turned back on the logic resumes operation from the point at which power was removed

- Key question:
  - How much / which state needs to be retained?

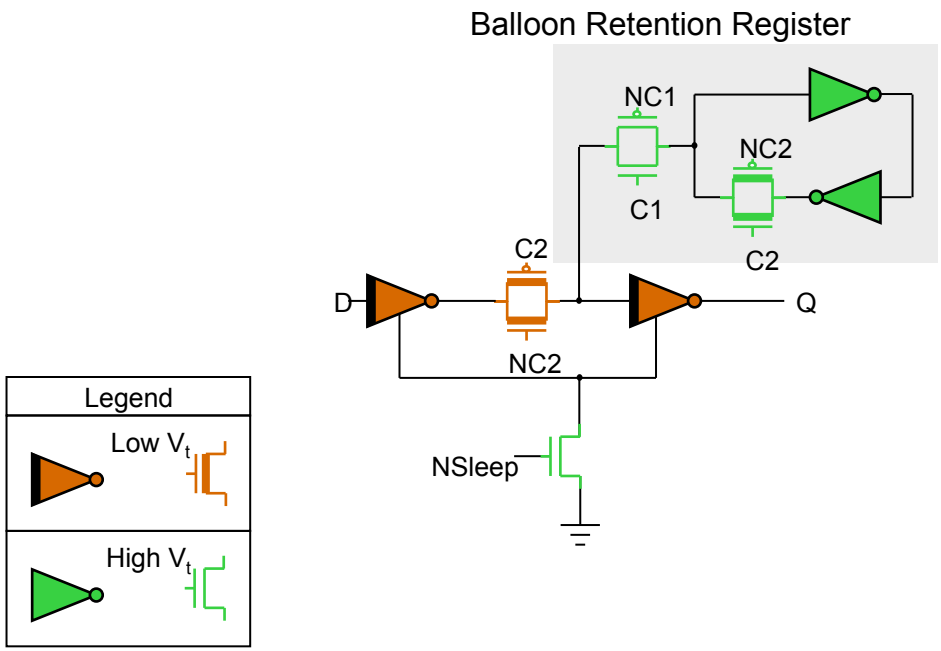
- Answer:
  - It depends on the application
    - Generally, data does not need to be saved
    - But control registers must be saved

State Retention Methods

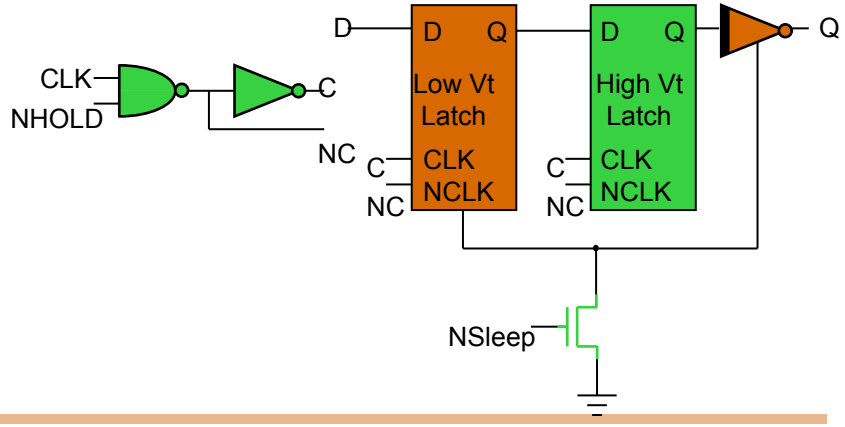
- **Scan-out/scan-in**
  - State in power-gated domain is scanned out into an always-on domain prior to opening switches
  - State is scanned back in after switches are re-closed
  - Power manager design can be complex
  - Lengthy power-down and power-up time
- **State retention registers (SRFF)**
  - Requires both switched and always-on power
  - Fast power-down and power-up
- **State retention voltage**
  - Reduce supply to ~ zero, but not so low as to lose state
  - Medium speed power-down and power-up

State Retention Register Design

- **Balloon Retention Register**
  - State is saving in an always-on shadow register
  - Main register uses low-Vt devices
  - Balloon uses high-Vt devices



- **Always-on Slave**
  - Modified master-slave flip-flop
  - Slave is always powered
  - Master latch, clock buffers, and output buffers are power-gated



State Retention Tradeoffs

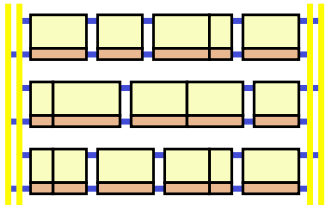
Method / Parameter	Area	Power-down time	Recovery Time	Leakage Reduction	Power manager design
Scan-in / Scan-out	Additional logic needed for scanning state in/out	Slowest – must scan state out	Slowest – must scan state back in	Some leakage from additional logic	Most complex
SRFF	Biggest registers, but little additional logic	Fastest	Fastest	Registers still leak, amount depends on register design	Least complex
Reduced voltage	Least	Almost as fast as SRFF	Moderate – depends upon voltage regulator	Depends upon voltage for virtual rail	Depends upon number of voltage levels

Power Gating Issues – Physical Design

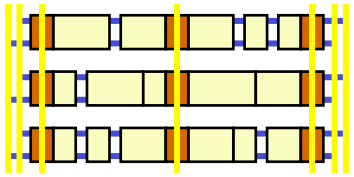
- Switch placement
- Switch sizing
- Rush current & Wakeup time control
- Library requirements

Switch Placement Options

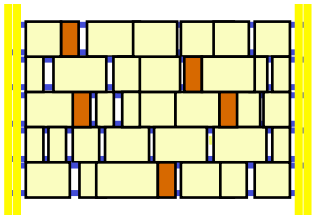
- **Switch-in-cell**
  - Each standard cell has a switch transistor embedded in the cell itself -> no standalone or separate switch cells
    - Each instance is power-gated
- **Grid of switches**
  - Switches are arrayed across the power-gated block
- **Filler switches**
  - Switches placed in available slots, otherwise occupied by filler cells
- **Ring of switches**
  - Ring separates *internal virtual* rails from *external real* rails



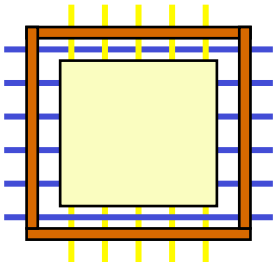
Switch-in-cell



Grid of switches

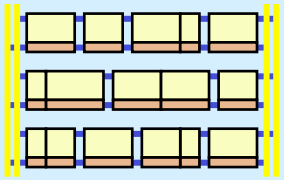
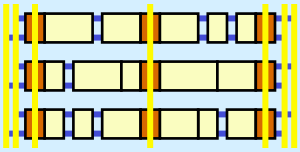
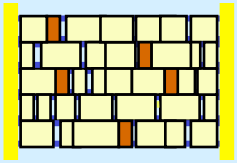
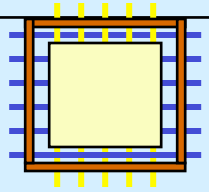


Filler Switches



Ring of switches

Switch Placement Tradeoffs

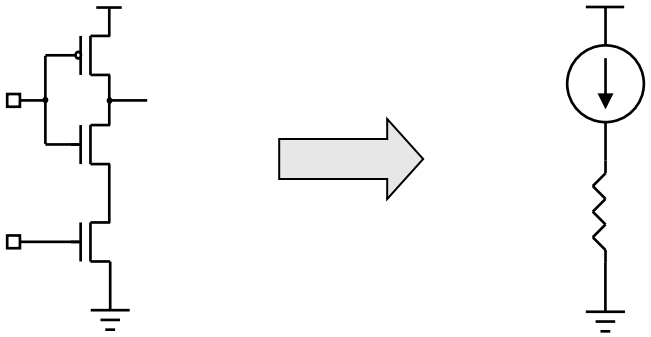
Placement Type	Merits	Drawbacks
Switch-in-cell 	Easy P&R; No floorplanning required	Very large area overhead
Grid of switches 	Area efficient	Must route a third rail
Filler switches 	Easy P&R; No floorplanning required; Area efficient	Specialized library layouts required IR droop issues due to irregular switched power grid
Ring of switches 	Useful for hard layout IP	Significant area overhead

Switch Sizing

- **Switches must be sized to meet electrical specs**
  - Voltage drop spec
  - Timing impact

- **Conceptual model – switch as a resistor in the power delivery network**

- **Critical issue: current estimation**



Switch Sizing Methods

Technique	Current Calculation Method	Merits	Drawbacks
Average Current Method (ACM)	Average (DC) currents assumed	Simple – avoids dynamic current analysis	Unrealistic; not worst case
Switch in Cell	Discharge currents simulated during cell characterization	Easy to characterize	Large area overhead; significant library development
Mutually Exclusive Switching	Discharge currents obtained from input vector simulation	Provides upper bound on switch sizing	Potential oversizing
Min-Max Window	Trapezoidal discharge waveform over mix-max switching window	Clustering minimizes max simultaneous switching	Potential oversizing; not worst case
Event Driven (VEDA)	Composite discharge waveform composed of currents from all switching events	All possible events are considered; sizing covers worst case conditions	Potential oversizing

Switch Sizing Tradeoffs

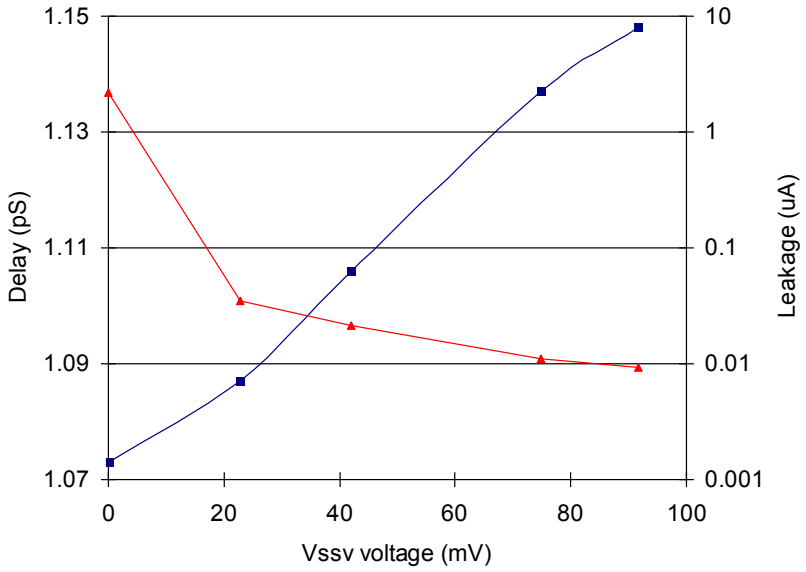
- **Fundamental tradeoff: area, performance, leakage**
  - Voltage drop across switch impacts performance and leakage

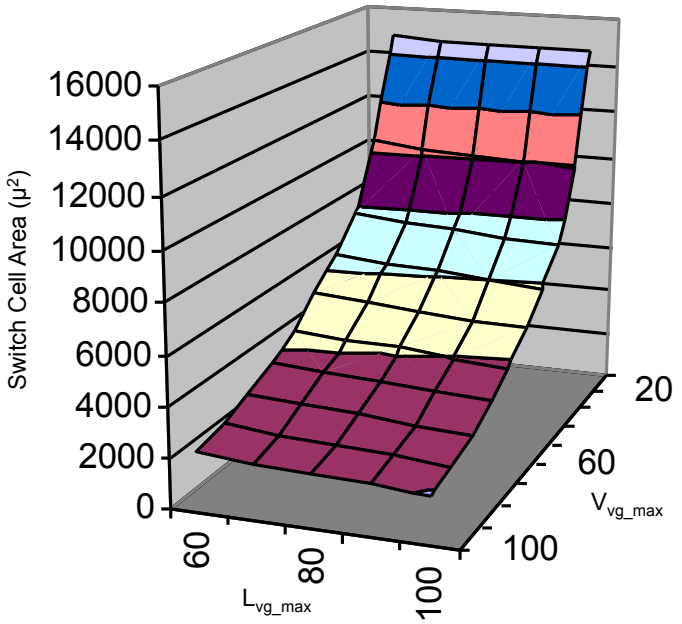
Switch Size	Voltage Drop	Performance Impact	Area	Off-state Leakage
Bigger switches	Smaller	Smaller	Larger	Larger
Smaller switches	Larger	Larger	Smaller	Smaller

- **Implications**
  - For least leakage – choose smallest switches without timing violations
  - For smallest area – choose smallest switches without timing violations
  - For fastest chip – choose biggest switches that area will allow

Switch Sizing Tradeoff Summary

- Larger switches
  -> More area, more leakage, higher performance
- Smaller switches
  -> Less area, less leakage, lower performance





Rush Current and Wake-up Time

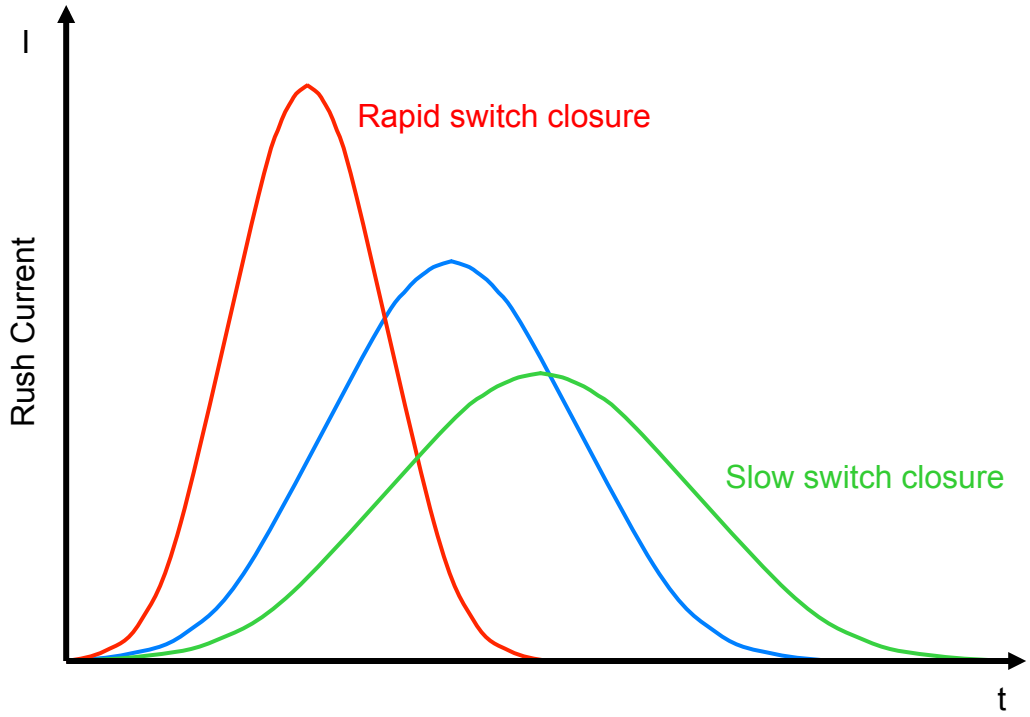
- **Rush Current: the current that flows through the switches when the switches transition from open to closed**

- **Wakeup Time: the time required for the virtual rail to settle down to an operational voltage after the switches have been closed**

- **Control Motivation**
  - Minimize peak current flow when switches are enabled
    - No EM problems
  - Minimize wake-up time
    - Switched logic must be operational within specified recovery time
  - Minimize voltage drop in always-on logic

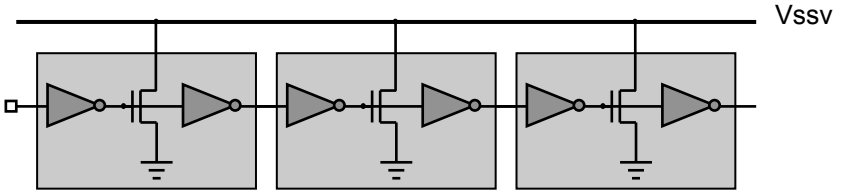
Rush Current & Wakeup Time Tradeoff

- Closing switches rapidly => large rush current
- Closing switches slowly => lengthy wakeup time

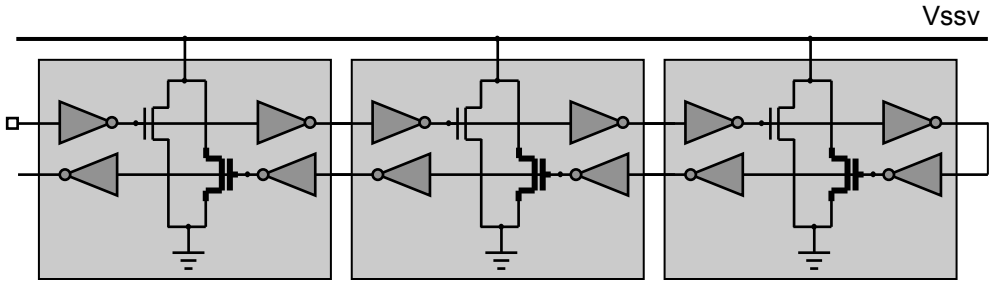


Rush Current Control Techniques

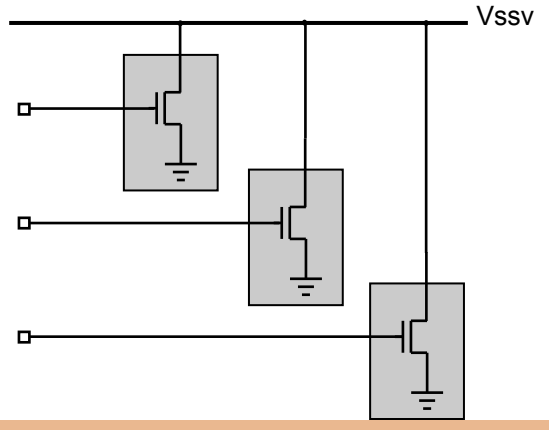
- **Single pass cascade**
  - Sleep enable traverses switch network in one direction



- **Two pass cascade**
  - Sleep enable traverses switch network in two directions – first with small switches and second with large switches



- **Staggered turn-on**
  - Multiple sleep enables turn on with staggered delays according to sleep controller outputs



Power Gating Library Requirements

- **Power Gating compatible libraries must contain**
  - Switch cells
    - Different sized switches desirable
    - Switches with and without built-in delay buffers
  - State retention registers
    - May / may not be necessary depending upon retention strategy
  - Isolation cells

- **Special Case: Switch-in-cell implementation**
  - A power-gated version of each logic cell is needed
  - No switch cells are needed
    - Since a switch is embedded in each logic cell
  - Isolation cells may / may not be needed
    - Depends upon whether isolation is built-in to power-gated logic cells

Power Gating Issues - Design Verification

- Logical verification
- Voltage drop verification
- Rush current and wakeup time verification

Logical Design Verification

- **Are the switches connected and controlled correctly?**
  - Are always-on blocks in fact always powered?
  - Do the power-gated blocks turn-off when and as expected?
  - Do the power-gated blocks turn-on when and as expected?
    - Each power state transition should be simulated / verified
- **Are all signal outputs of power-gated blocks isolated?**
- **Can state be stored and restored?**
  - Are retention controls controllable and observable for test?
  - Is there a minimum power-down time before power gating controller can attempt to power up?

Voltage Drop Verification

- **Verify voltage drop during active mode operation**
  - Check dynamic voltage drop
    - Virtual rail dynamic voltages must be within voltage drop limit
    - Real rail dynamic voltages must be within voltage drop limit
  - Check current densities
    - Virtual rail current densities must be below EM limit
    - Real rail current densities must be below EM limit
    - Switch currents must be below EM limit

- **Verify voltage drop during mode transitions**
  - Check rush currents and wake-up times

Rush Current & Wakeup Time Verification

- **Verify rush currents for powering-up operation**
  - Analyze transient current flow when switches close
  - Must be performed concurrently with dynamic voltage drop analysis
    - Voltage drop affects current flow, and current flow affects voltage drop
  - Check currents
    - Peak currents must be below current limit
  - Check voltages
    - Max voltage drop on real rail must be below voltage drop limit

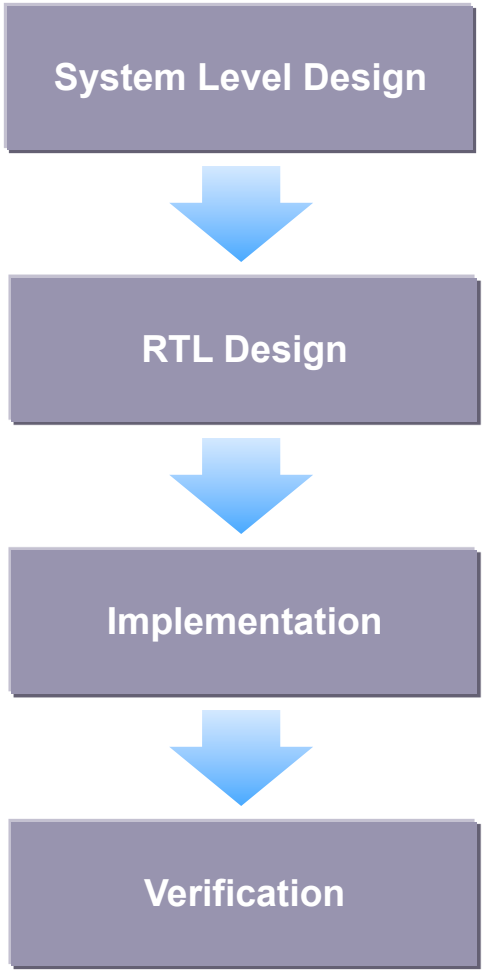
- **Verify wakeup times**
  - Check time required for power-gated logic to see a steady voltage
  - Check time required for always-on logic to see a steady voltage

Power Gating Design Flow and Design Tools

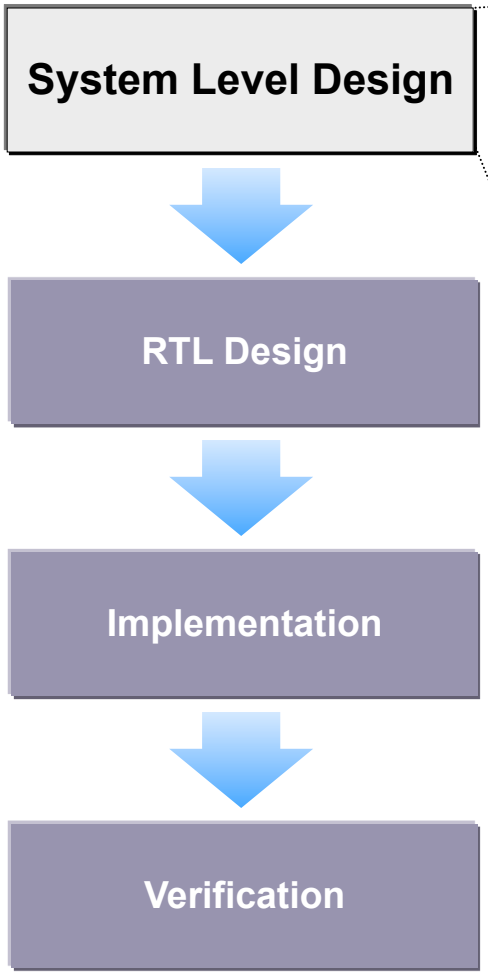
- Power Gating design flow
- Power Gating commands for power intent
- Power Gating tool requirements

Generalized Low Power Design Flow

- **System Level Design**
  - Goal: Explore architectures, algorithms, and protocols
- **RTL design**
  - Goal: Generate RTL to match system level design
- **Implementation**
  - Goal: Synthesize RTL to gates, floorplan, place & route
- **Verification**
  - Goal: Verify correct functional and electrical operation

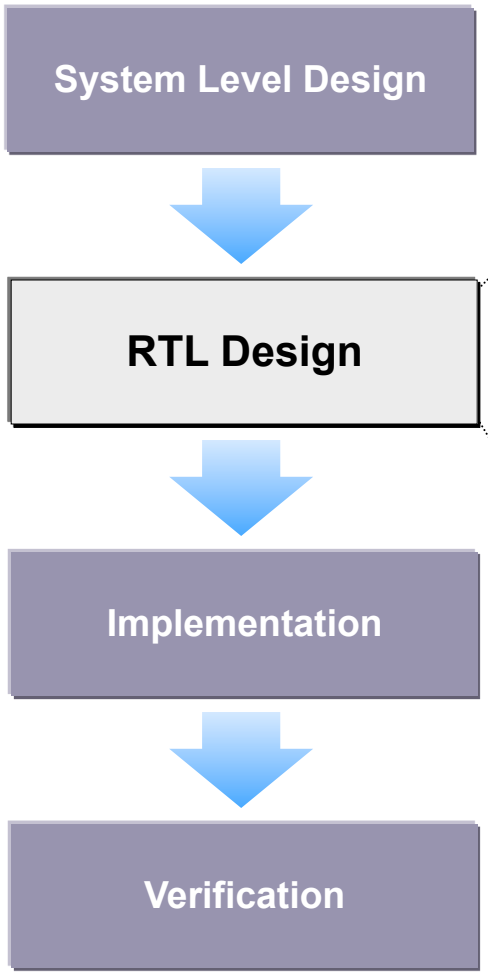


Power Gating Aware System Level Design



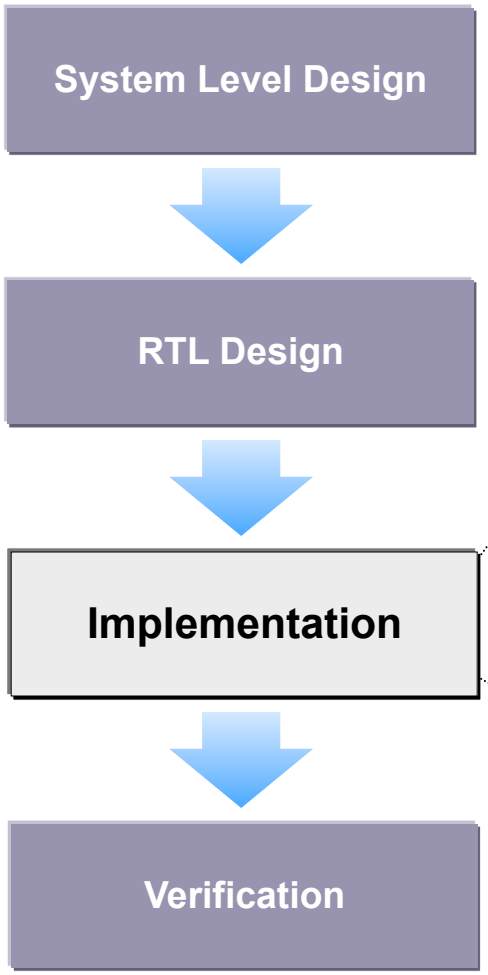
- Define power states and domains
  - Number of different domains
- Define power-down and power-up protocols
- Define interface and isolation protocols
- Define retention strategy
  - Save and restore methods
  - Full or partial state retention
- Decide upon high-side or low-side switching

Power Gating Aware RTL Design



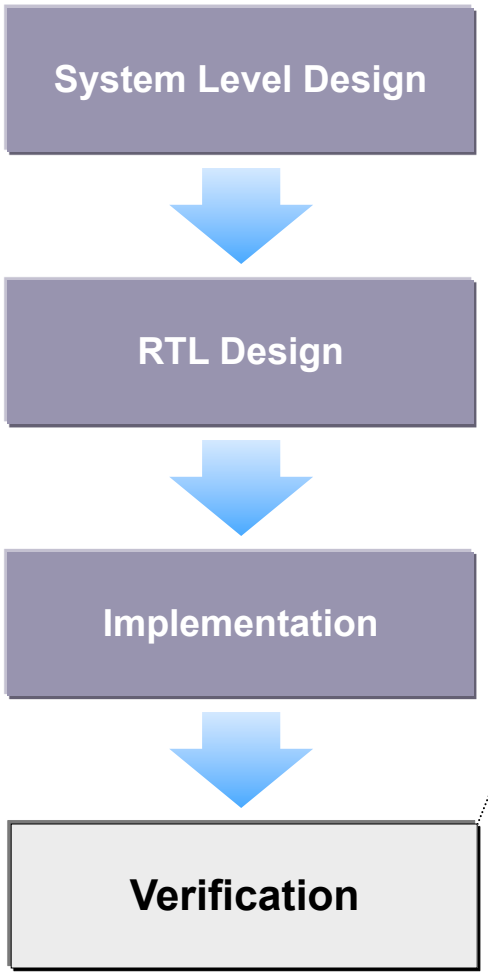
- Specify which registers must maintain state
- Design power gating controller
  - Generate sleep enable signals
- Design rush current control scheme
- Simulate RTL to confirm power gating behavior
  - Always on, power-up, & power-down
- Analyze power to determine if design will meet power specs
- Create power gating design constraints in CPF / UPF

Power Gating Aware Implementation



- Synthesize RTL to gates
  - Specify which registers must maintain state
- Floorplan, P&R design
- Insert, size, & place switches
- Buffer and distribute sleep enables
- Tune rush current control circuit

Power Gating Aware Verification



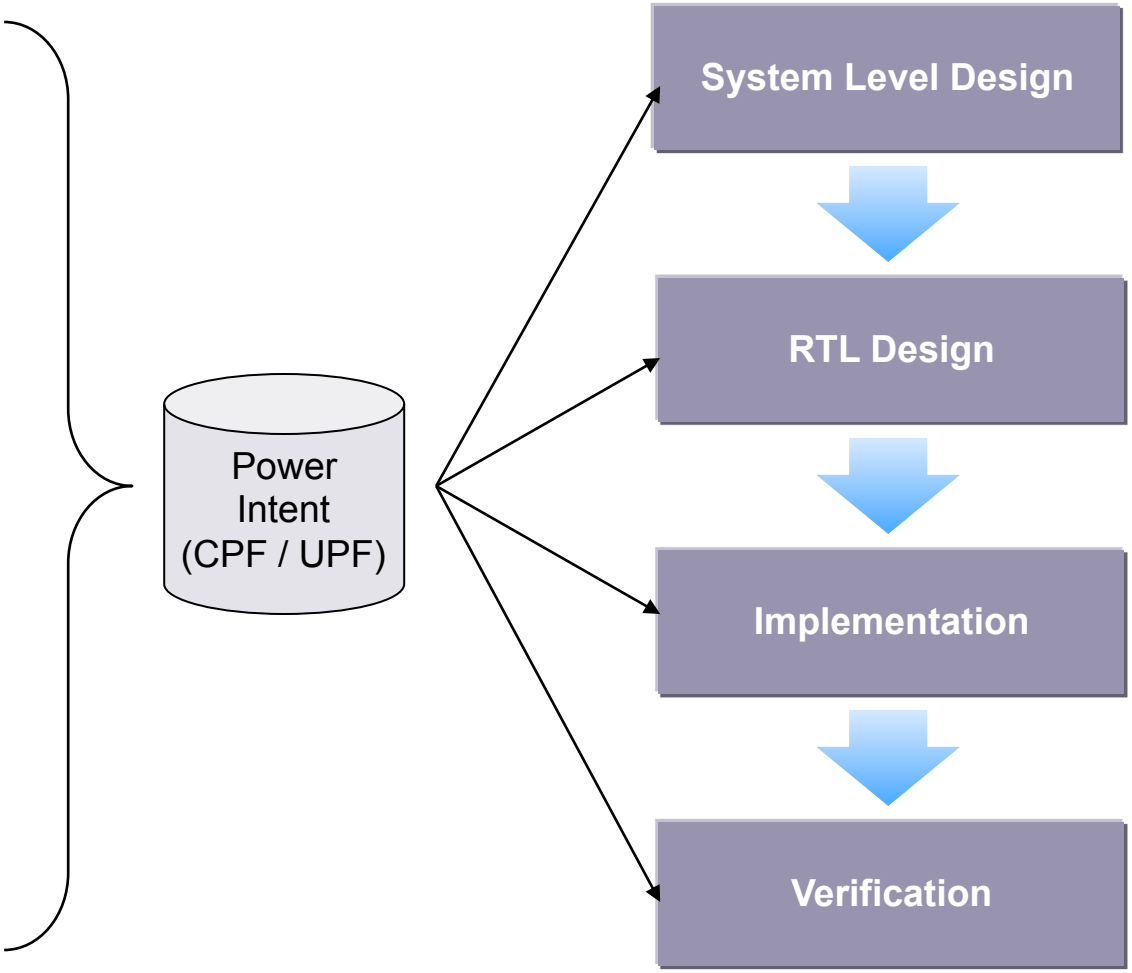
- Power specs
  - Leakage power
  - Dynamic power
- Voltage Drop
  - Real rails & Virtual rails
- EM
  - Real rails & Virtual rails
- Rush Currents
- Wakeup time
- Timing effects

Power Gating Design Constraints

- Specify power intent to multiple tools with a single constraint language
  - Example: Power domains, logical and physical
- Two different (but very similar) TCL-based languages
  - CPF: Common Power Format
  - UPF: Unified Power Format
- Some CPF commands
  - set_ * commands: general commands
  - define_*_cell commands: library cell descriptions
  - create_*_rule commands: design intent
  - update_*_rule commands: implementation directives

Power Intent

- **Power domains**
  - Logical partitions
  - Physical partitions
- **Power logic**
  - Switches
  - Isolation cells
  - State retention logic
- **Power modes**
  - Mode definitions
  - Transition definitions



CPF Example Snippet

```
# define power domains
create_power_domain -name PDcore -default
create_power_domain -name PDau -instances alu_inst/aui \
        -shutoff_condition pcu_inst/pau[2]

# define condition which triggers state retention
create_state_retention_rule -name sr_rule -domain PDrf \
        -restore_edge {!pcu_inst/[rf[1]}

# define which signals will be isolated, under which condition, and how
create_isolation_rule -name iso-rule1 -from PDau \
        -isolation_condition {!pcu_inst/pau[0]} \
        -isolation_output low

# define voltages
create_nominal_condition -name high -voltage 1.2
create_nominal_condition -name low  -voltage 0.0

# define power modes
create_power_mode PM1 -default -domain_conditions {PDcore@high PDau@high}
create_power_mode PM2 -domain_conditions {PDcore@high PDau@low }
```

Power Gating SoC Tool Requirements

- **Key SoC tools must be Power Gating aware**
  - RTL functional simulator
    - Needed to simulate power gating power-up / power-down behavior
    - Must be CPF or UPF compliant to understand power intent
  - RTL power analyzer
    - Needed to analyze module level power consumption and to predict power savings from Power Gating and other low power techniques
  - Synthesizer
    - Needed to insert isolation cells and SRFFs where appropriate
  - Switch sizer
    - Power Gating specific tool needed to correctly size switches
  - Placer and router
    - Needed placing and connecting switches according to desired topology
  - Voltage drop analyzer
    - Needed to analyze voltage drop across switches and verify rush current and wake-up time behavior.

- **Power Gating awareness can be controlled by CPF / UPF**

Quick Summary: Power Gating Design Issues

- **Power Gating presents unique issues to designers**
  - Logic design, physical design and verification
- **Unfamiliar trade-offs must be confronted efficiently**
  - Headers vs footers
  - Switch size vs area vs leakage reduction vs delay
  - Switch placement vs area vs ease of layout
  - Power controller complexity vs power verification
- **New tools and methodologies are required**
  - Existing tools are evolving to become power aware
  - But new methodologies accompany the new tools
  - Usage of CPF / UPF will become requirements for efficient flows

BACKUP

References (1 of 3)

1. International Technology Roadmap for Semiconductor (ITRS), 2006 Edition, Executive Summary
2. Kaushik Roy, et. al., "Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits," *Proceedings of the IEEE*, pp 305-327, Feb. 2003.
3. L.Wei, Z.Chen, M.Johnson, and K.Roy, "Design and optimization of low voltage high performance dual threshold CMOS circuits," *DAC'98, pp.489-494, Jun.1998.*
4. L. Wei, Z. Chen, K. Roy, Y. Ye, and V. De, "Mixed Vth CMOS circuit design methodology for low power applications," *DAC, 1999*
5. J.P.Harter and F.Najm, "A gate-level leakage power reduction method for ultra-low-power CMOS circuits," *CICC'1997, pp.475-478, Oct. 1997.*
6. A. Abdollahi, F. Fallah, and M. Pedram, "Leakage current reduction in CMOS VLSI circuits by input vector control," *IEEE Transaction on VLSI, Feb. 2004.*
7. S.Mutoh, T.Douseki, Y.Matsuya, T.Aoki, S.Shigematsu, and J.Yamada, "1-V power supply high-speed digital circuit technology with multi threshold-voltage CMOS," *IEEE J. Solid-State Circuits, vol. 30, pp. 847-854, Aug. 1995.*
8. K.Usami, N.Kawabe, and M.Koizumi, "Automated selective multi-threshold design for ultra-low standby applications," *ISLPED'02*, pp.202-206, Aug. 2002.
9. K.Min, H.Kawaguchi, and T.Sakurai, "Zigzag super cut-off CMOS (ZSCCMOS) block activation with self-adaptive voltage level controller: An alternative to clock-gating scheme in leakage dominant era," *ISSCC'2003*, pp.400-401, Feb. 2003.
10. K-w Choi, Y. Xu, and T. Sakurai, "Optimal zigzag (OZ): an effective yet feasible power-gating scheme achieving two orders of magnitude lower standby leakage," *Proceeding of 2005 Symposium on VLSI Circuits*, pp. 312-315, June 14-18, 2005.
11. J. Frenkil and S. Venkatraman, "Power Gating Design Automation", *Closing the Power Gap Between ASIC and Custom*, Springer, 2007, pp. 251-280.

References (2 of 3)

12. J. Tschanz, J. Kao, S. Narendra, R. Nair, D. Antoniadis, A. Chandrakasan, and V. De, "Adaptive body bias for reducing impacts of die-to-die and within-die parameter variations on microprocessor frequency and leakage", *ISSCC*, February 2002, Vol.1, p. 422-478.
13. J. Hu, Y. Shin, N. Dhanwada, and R. Marculescu, "Architecting voltage islands in Core-based System-on-a-Chip Designs," *ISLPED*, 2004.
14. M. Horiguchi, T. Sakata, and K. Itoh, "Switched-source-impedance CMOS circuit for low standby subthreshold current gigascale LSI's, " *Proceeding of 2005 Symposium on VLSI Circuits*, pp. 47-48, May 19-21, 1993.
15. J. Kao, S. Narendra, and A. Chandrakasan, "MTCMOS hierarchical sizing based on mutual exclusive discharge patterns," *Proc. of ACM/IEEE Design Automation Conference*, pp.495-500, 1998.
16. S. Mutoh, S. Shigematsu, G. Yoshinori, and S. Konaka, "Design method of MTCMOS power switch for low-voltage high-speed LSIs," *Proc. of ACM/IEEE Design Automation Conference. (ASP-DAC)*, pp. 113-116, 1999.
17. M. Anis, A. Shawki, and M. Elmasry, "Design and optimization of MTCMOS circuits," *IEEE Trans on CAD, vol. 22, pp. 1324-1342, Oct. 2003.*
18. G. Uvieghara, et. al., "A highly-integrated 3G CDMA2000 1X cellular baseband chip with GSM/AMPS/ GPS/Bluethooth/Multimedia capabilities and ZIF RF support," *ISSCC, 2004.*
19. K-w Choi and J. Frenkil, "VEDA: Vectorless event-driven approach for optimal switch sizing of Power-Gating circuits to reduce two orders of magnitude of leakage power," *SAME conference in Nice, France,* Oct., 2006.
20. R.S. Chakraborty, et. al., **"**Hybridization of CMOS With CNT-Based Nano-Electromechanical Switch for low leakage and robust circuit design", *IEEE transactions on Circuits and Systems,* Vol. 54, Nov. 2007.
21. S. Bhunia, M. Tabib-Azar, and D. Saab, "Ultralow-Power reconfigurable computing with complementary nano-electromechanical carbon nanotubes switches," *ASP-DAC* 2007.

References (3 of 3)

22. B. Calhoun, et. al., "Power Gating and Dynamic Voltage Scaling", *Leakage in Nanometer CMOS Technologies,* Springer, pp. 41-76, 2005.
23. S. Kosonocky, et. al., "Enhanced Multi-Threshold (MTCMOS) Circuits Using Variable Well Bias", *ISLPED*, pp. 165-169 2001.
24. Q. Wang, "Common Power Format Tutorial", *www.si2.org*, 2007.
25. P. Royannez, et. al., "90nm Low Leakage SoC Design Techniques for Wireless Applications", *ISSCC,* 2005.
26. T. Hattori, et. al., " Hiearchical Power Distribution and Power Management Scheme for a Single Chip Mobile Processor", *Proc. of ACM/IEEE Design Automation Conference*, pp.292-295, 2006.
27. T. Lueftner, et. al., "A 90nm CMOS Low-PowerGSM/EDGE Multimedia-Enhanced Baseband Processor With 380-MHz ARM926 Core and Mixed-Signal Extensions", *ISSCC, 2006*
28. Y. Kanno, et. al., "Hierarchical Power Distribution with 20 Power Domains in 90-nm Low-Power Multi-CPU Processor", *ISSCC*, 2006.
29. Y. Kanno, et. al., "Hierarchical Power Distribution with Power Tree in Dozens of Power Domains for 90-nm Low-Power Multi-CPU SoCs", *IEEE Journal of Solid State Circuits*, vol. 42, pp. 74-83, January 2007.
30. Z. Hu, et. al., "Microarchitectural Techniques for Power Gating Execution Units", *ISLPED*, 2004.
31. S. Rele, et. al., "Optimizing Static Power Dissipation by Functional Units in Superscalar Processors", *International Conference on Compiler Construction*, pp 261-275, April 2002.
32. K. Usami, et. al., "Dynamic Sleep Control for Finite-State-Machines to Reduce Active Leakage Power", *IEICE Transactions on Fundamentals of Electronics, Communications, and Computer Systems*, Dec. 2004.
33. K. Usami, et. al, "An Approach for Fine-grained Run-Time Power Gating using Locally Extracted Sleep Signals", *IEEE International Conference on Computer Design*, Oct. 2006.

Power 101

Recall that power is defined as:

$$Power\ (watts) = Vdd * I_{ckt}$$

And that power is also defined as:

$$Power\ (watts)\ = \frac{Energy\ (watt{-}sec)}{T_{program}\ (sec)}$$

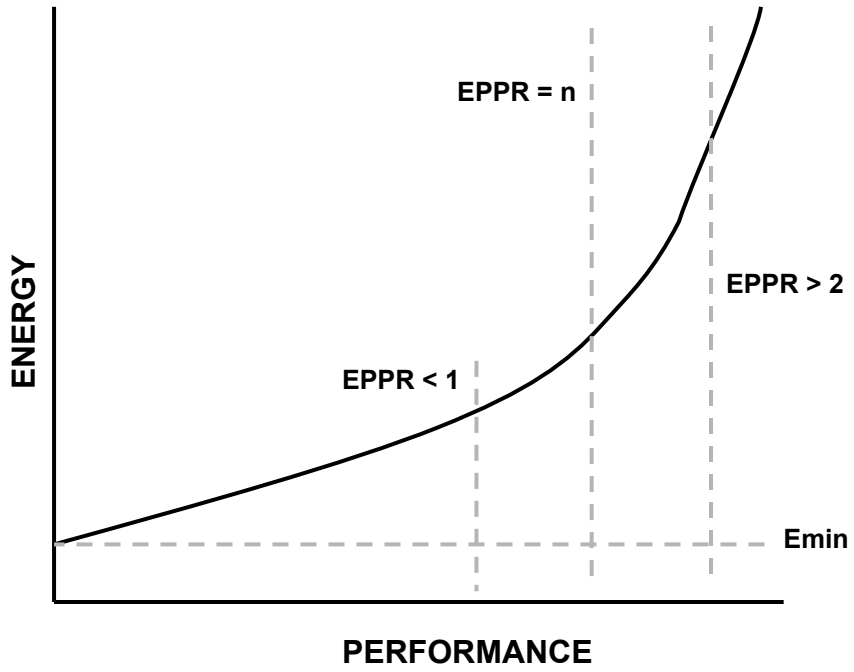
Where $T_{program}$ is the program (workload) execution time and is defined as the number of instructions required to execute the program multiplied by the clock period per instruction:

$$T_{program} = Num_{Inst} * T_{clk}$$

Energy per instruction is another measurement of the efficiency of computational element and is typically measured in Joules/Instruction. A variation of this measurement is Watts/IPS or the more widely used variant: MIPS/Watt. The derivation of this equality is shown below:

$$\frac{Joules}{Instruction} = \frac{\frac{Joules}{Second}}{\frac{Instructions}{Second}} = \frac{Watts}{IPS}$$

EPPR: Energy Performance Percentage Ratio



- **EPPR < 1** where an increase in energy results in a proportional increase in performance
- **1<= EPPR <= 2** region where a moderate increase in energy produces reasonable performance increase
- **EPPR > 2** where an increase in energy provides a marginal increase in performance while asymptotically approaching a finite performance limit.

Energy Optimized: EPPR < 1

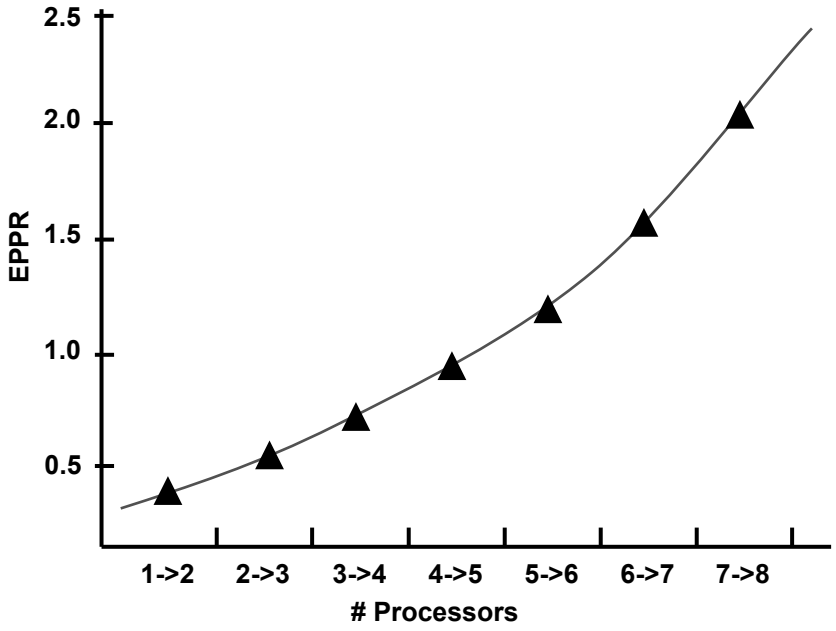
- Run at low VDD
- Mostly small devices
- Shallow pipeline

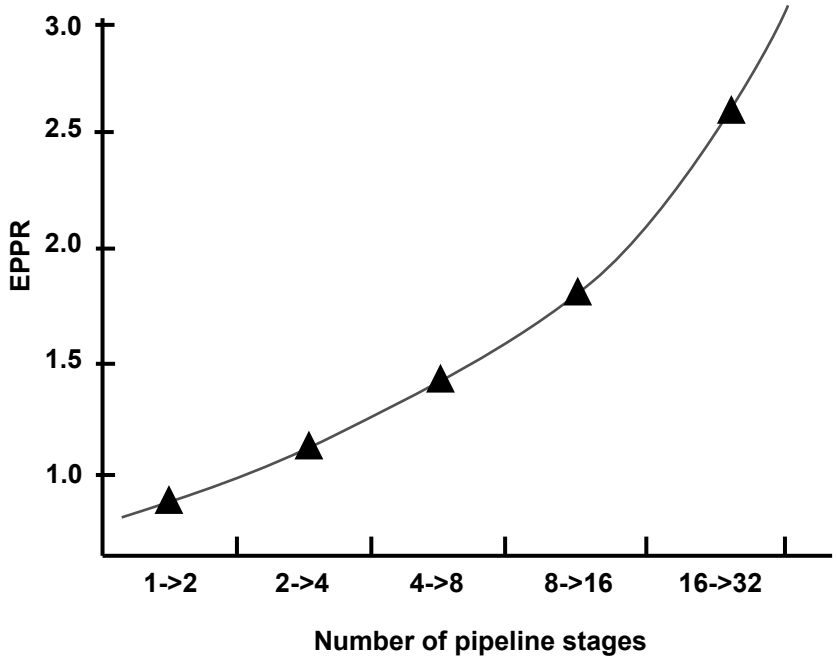
Energy – Performance: 1<= EPPR <= 2

- Run at nominal VDD
- Moderate pipeline depth, moderate instruction level parallelism

Performance optimized: EPPR > 2

- Run at maximum VDD
- Little concurrency in application
- Deep pipelines


2.5
2.0
1.5
1.0
0.5
EPPR
1->2
2->3
3->4
4->5
5->6
6->7
7->8
Processors


3.0
2.5
2.0
1.5
1.0
EPPR
1->2
2->4
4->8
8->16
16->32
Number of pipeline stages

Exam #2 Problem

This problem looks at the tradeoffs between active power and leakage power: A 65nm microprocessor runs at 3.3GHz with 54W of active power and 29W of leakage power

Assumptions:
1) Frequency scales linearly (proportionally) with supply voltage (VDD)
2) Leakage scales linearly with VDD and linearly with junction temperature.
3) Junction temperature scales linearly with power

By scaling **frequency and supply voltage** together what is the highest frequency, supply voltage and junction temperature that will reduce total power to 16W?

NOTE: This problem does not have a closed form solution, you will need to solve it accordingly.

Solution

New_VDD = Old_VDD ✖ VDD_scale

Pactive ∝ VDD^2*Freq ∝ VDD_scale^3

Tj_scale ∝ Pactive ∝ VDD_scale^3

∴ Pleak ∝ VDD_scale ✖ Tj_scale ∝ VDD_scale^4

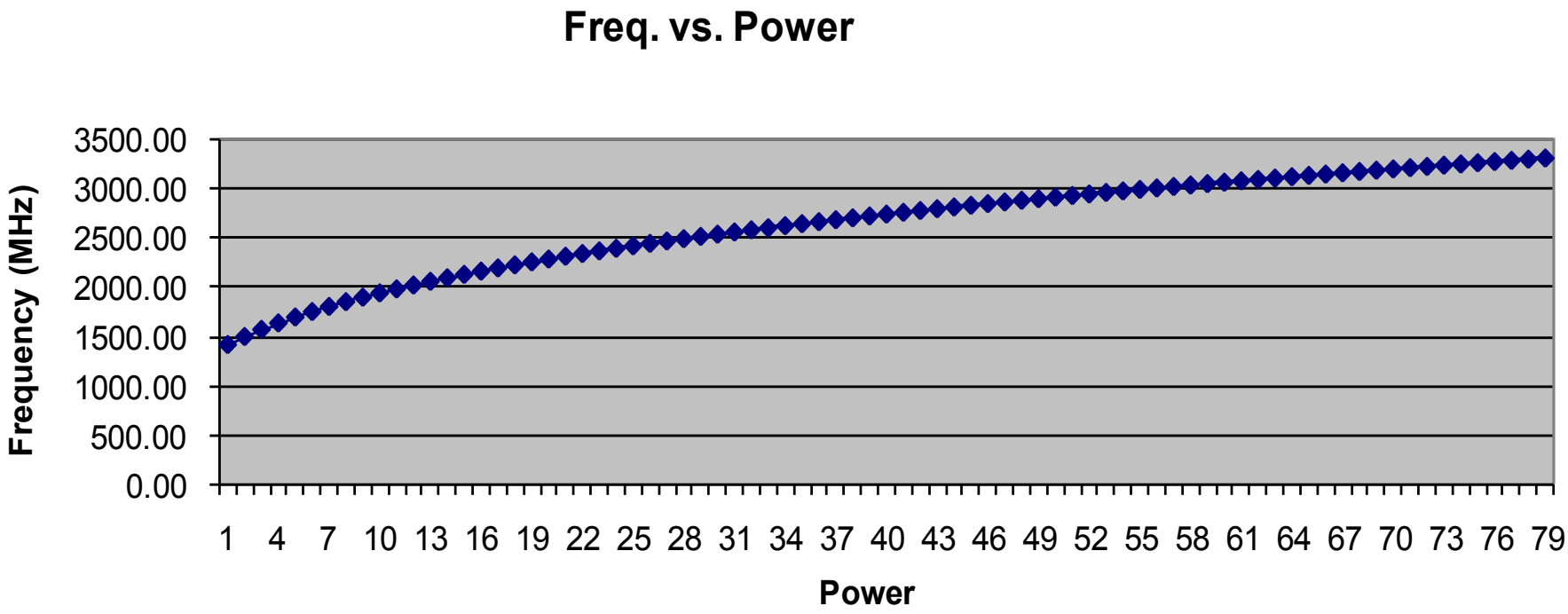
Ptotal = Pactive ✖ VDD_scale^3 + Pleak ✖ VDD_scale^4

VDD_scale = [Ptotal /(Pactive + Pleak ✖ VDD_scale)]^(1/3)

(*no closed form solution for VDD_Scale but we can iterate to solve)*

VDD_scale = [16W/(54W+29W ✖ VDD_scale)]^(1/3) = .606

∴ NewFreq = OldFreq ✖ VDD_scale = 3.3GHz ✖ 0.606 = 2.002GHz


Freq. vs. Power
Frequency (MHz)
3500.00
3000.00
2500.00
2000.00
1500.00
1000.00
500.00
0.00
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52 55 58 61 64 67 70 73 76 79
Power