USC
Low Power
CAD

Low Power Design Methodologies and Techniques: An Overview

Massoud Pedram
Department of EE-Systems
University of Southern California
Los Angeles CA 90064
pedram@ceng.usc.edu

Massoud
Pedram

USC
Low Power
CAD

Outline

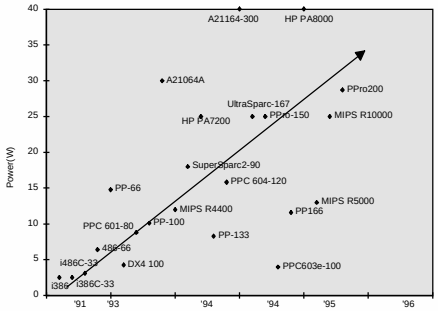
- Introduction
- Analysis/Estimation Techniques
- Synthesis/Optimization Techniques
- Summary

Massoud
Pedram

USC
Low Power CAD

What Is the Power Problem?

- **Power: Biggest challenge facing the industry**
  - ⇨ **Power goes up 2X per generation for high end CPU**
- **CAD/analysis tools are excellent at enabling tradeoffs**
  - ⇨ **But what if performance cannot be traded-off?**
  - ⇨ **Need innovative CAD solutions to reduce Power**



[Source: Microprocessor Report]

Massoud Pedram

USC
Low Power CAD

Process Technology Trend

	1997	1999	2001	2003	2006	2009	2012
Min. Feature Size (μm)	0.25	0.18	0.15	0.13	0.1	0.07	0.05
Threshold Voltage (V)	0.5	0.45	0.4	0.35	0.3	0.25	0.2
T_{ox}(nm)	5	4	3	3	2	1.5	1
$C_{interconnect}$ (aF/μm)	50	36	29	21	20	16	14
Norminal I_{on}(μA/μm)(NMOS/PMOS)	600/280	600/280	600/280	600/280	600/280	600/280	600/280
Max I_{off}(μA/μm) (For minimum L device)	1	1	3	3	3	10	10
Supply Voltage (V)	2.5	1.8	1.8	1.5	1.2	0.9	0.6
On-chip across-chip clock freq. (MHz)	350	526	727	928	1108	1468	1827
Off-chip peripheral bus freq. (MHz)	75/175	100.263	100/362	125/464	125/554	150/734	150/913
Usable transistors (millions / cm^2)	8	14	16	24	40	64	100
Chip size (cm^2)	3	3.4	3.85	4.3	5.2	6.2	7.5
Chip pad count	256-800	300-976	352-1193	413-1458	524-1968	666-2656	846-3578

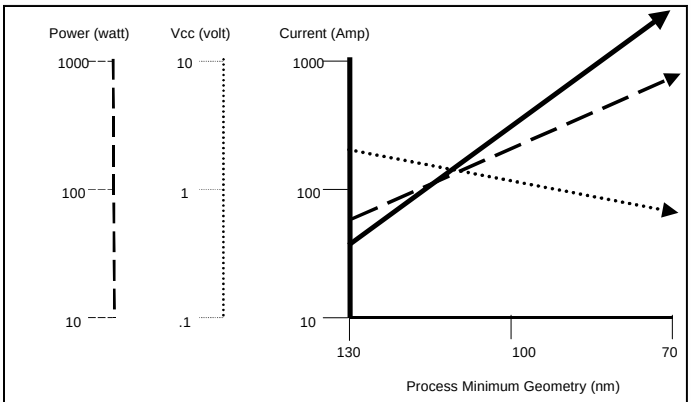
Cost-Perf. Designs - notebooks, desktop personal computers, telecom

Massoud Pedram

Data extracted from NTRS'97

USC
Low Power CAD

Design Technology Trend



Based on Information provided by Shakhar Borkar, Intel

Massoud Pedram

Data extracted from EDAR'99

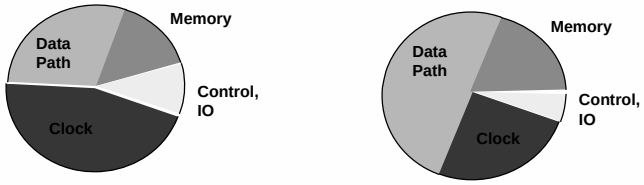
USC
Low Power CAD

Example Applications

- **Portable Electronics (PC, PDA, Wireless)**
- **IC Cost (Packaging and Cooling)**
- **Reliability (Electromigration, Latch-up)**
- **Signal Integrity (Switching Noise, DC Voltage Drop)**
- **Thermal Design**

Where Does the Power Go?

Varies from one design to next



Massoud Pedram

USC
Low Power CAD

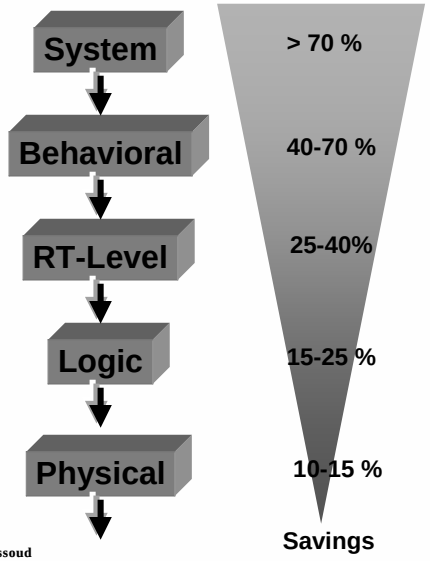
What Has Worked?

- Voltage and process scaling (3x/Generation)
- Design methodologies
  - ⇨ Power management through HW/SW, trade area for lower power
- Architecture Design
- Power down techniques
  - ⇨ Clock gating, dynamic power management
- Dynamic voltage scaling based on workload
- Power conscious RT/ logic synthesis
- Better cell library design and resizing methods
  - ⇨ Cap. reduction, threshold control, transistor layout

Massoud Pedram

USC
Low Power CAD

Opportunities for Power Savings



HW/SW Co-design
Custom ISA
Algorithm Design
Communication Synthesis

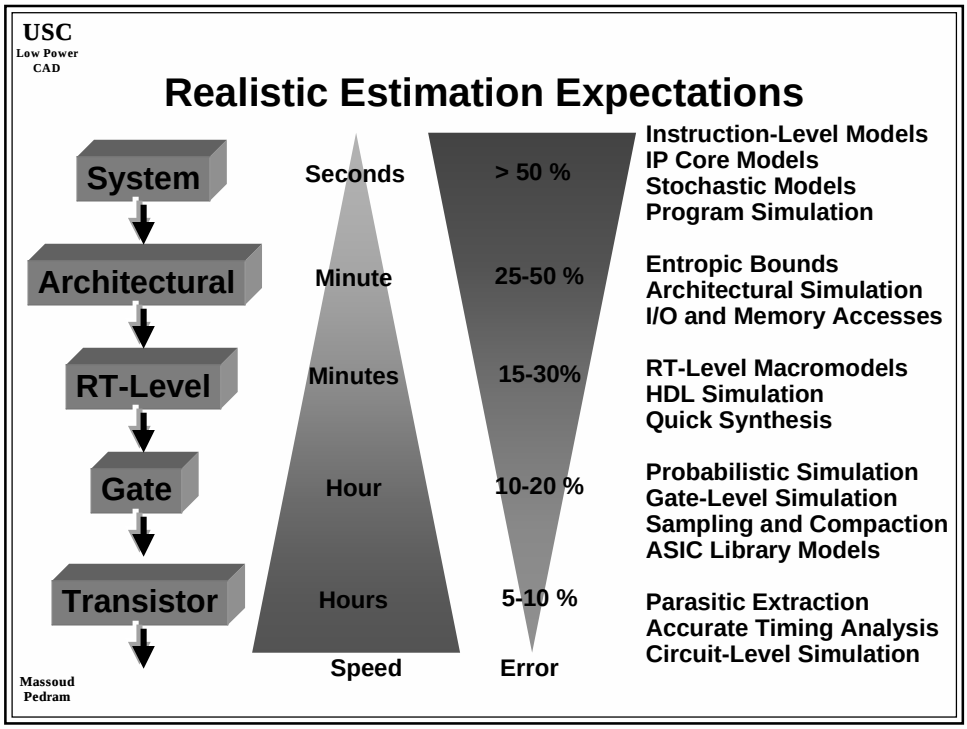
Scheduling, Binding
Pipelining
Behavioral Transformations

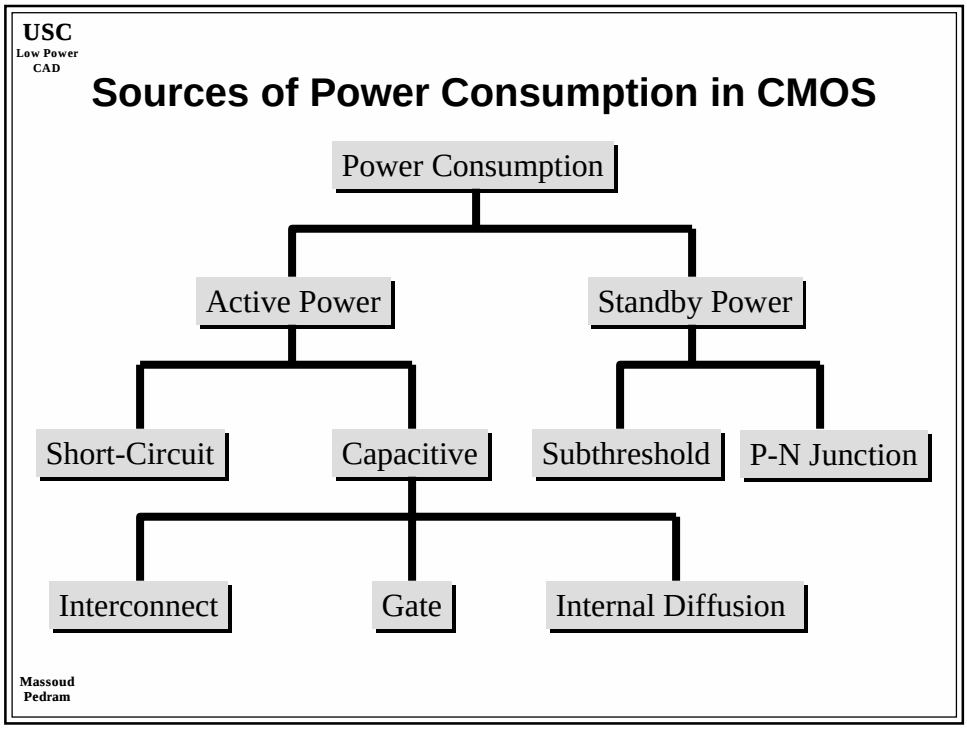
Clock Gating, Precomputation
Operand Isolation
State Assignment, Retiming

Logic Restructuring
Technology Mapping, Rewiring
Pin Ordering & Phase Assignment

Fanout Optimization, Buffering
Transistor Sizing, Placement
Partitioning, Clock Tree Design
Glitch Elimination

Massoud Pedram

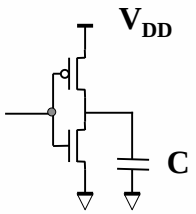

USC
Low Power CAD
Realistic Estimation Expectations
System
Architectural
RT-Level
Gate
Transistor
Seconds
Minute
Minutes
Hour
Hours
Speed
> 50 %
25-50 %
15-30%
10-20 %
5-10 %
Error
Instruction-Level Models
IP Core Models
Stochastic Models
Program Simulation
Entropic Bounds
Architectural Simulation
I/O and Memory Accesses
RT-Level Macromodels
HDL Simulation
Quick Synthesis
Probabilistic Simulation
Gate-Level Simulation
Sampling and Compaction
ASIC Library Models
Parasitic Extraction
Accurate Timing Analysis
Circuit-Level Simulation
Massoud Pedram


USC
Low Power CAD
Sources of Power Consumption in CMOS
Power Consumption
Active Power
Standby Power
Short-Circuit
Capacitive
Subthreshold
P-N Junction
Interconnect
Gate
Internal Diffusion
Massoud Pedram

USC
Low Power CAD

Power Dissipation Equations

CMOS circuits only dissipate power when node voltages are changing



short circuit charge

$$P = 0.5CV_{DD}^{\ 2} fN + Q_{SC}V_{DD} fN + I_{leak}V_{DD}$$

leakage current

***f*: frequency of clocking**
***N*: number of times gate switches in a clock cycle**

Massoud Pedram

USC
Low Power CAD

Power Estimation Techniques

- **Static (non-simulative) - useful for synthesis and architectural exploration**
  - ⇨ **Probability-based**
  - ⇨ **Entropy-based**
- **Dynamic (simulative) - useful for final power evaluation and validation**
  - ⇨ **Direct (flat and hierarchical)**
  - ⇨ **Sampling-based**
  - ⇨ **Compaction-based**
- **Hybrid (high-level simulation + low-level analytical model evaluation)**
  - ⇨ **Power macromodels for datapath, control, memory**
  - ⇨ **Instruction-level models for microprocessors, DSPs**

Massoud Pedram

USC
Low Power CAD

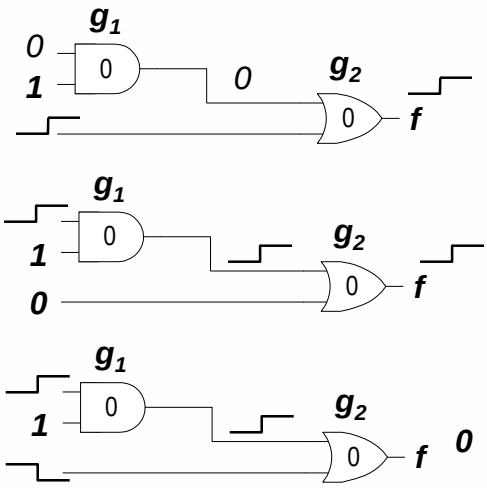
Probabilistic Power Estimation

- Capture primary input and register output statistics
  - ⇨ E.g., signal probability, switching activity, spatio-temporal correlation
- Obtain load capacitances for all nodes
  - ⇨ Estimate at the RT and logic level
  - ⇨ Extract after layout
- Propagate input statistics throughout the circuit
  - ⇨ Account for reconvergent fanout,
  - ⇨ Use appropriate delay model
  - ⇨ Iterate in the case of finite state machines
- Calculate node power and hence total power

Massoud Pedram

USC
Low Power CAD

Zero Gate Delays

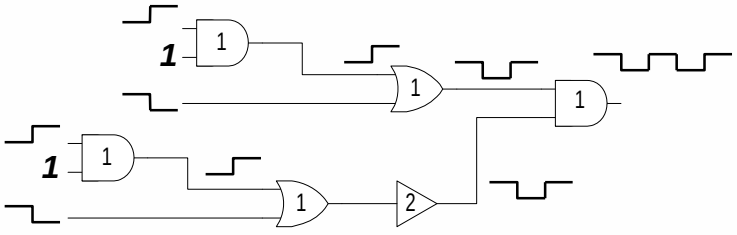


Massoud Pedram

USC
Low Power CAD

General Gate Delays

A gate may switch many times for a vector pair

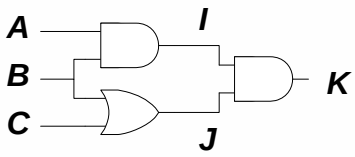


Massoud Pedram

USC
Low Power CAD

Static Probabilities

Define p_g as static probability of g = 1



For AND gate: $p_I = p_A \bullet p_B$

For OR gate: $p_J = 1 - (1 - p_B) \bullet (1 - p_C)$

Above equations assume that *A*, *B* and *C* are uncorrelated

Massoud Pedram

USC
Low Power CAD

Computing Static Probabilities

$$f = a\,b + b\,c$$

Write *f* as a disjoint sum-of-products expression where each product-term has a null intersection with any other

$$f = a\,b + \bar{a}\,b\,c$$

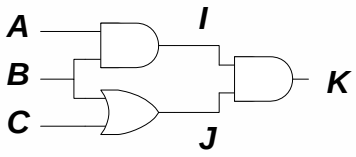
Then,

$$p_f = p_a \bullet p_b + (1 - p_a) \bullet p_b \bullet p_c$$

Massoud Pedram

USC
Low Power CAD

Spatial Correlations



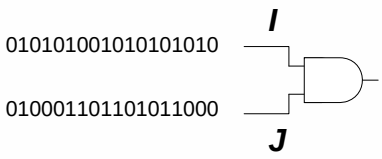
$p_K \neq p_I \bullet p_J$ **since *I* and *J* are correlated**
- both depend on *B*

Need to compute static probabilities taking into account the spatial correlations

Massoud Pedram

USC
Low Power CAD

Temporal Correlations



$p_K = p_I p_J$

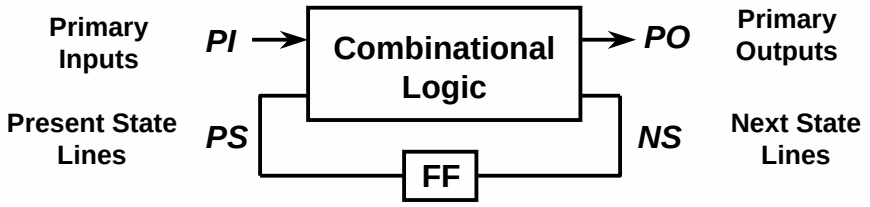
For temporally uncorrelated data,
$sw_K = 2 \bullet p_I p_J \bullet (1- p_I p_J)$
For temporally correlated data, this is not true
E.g., every 1 on input I is immediately followed by a 0

Need to compute switching probabilities taking into account the temporal correlations

Massoud Pedram

USC
Low Power CAD

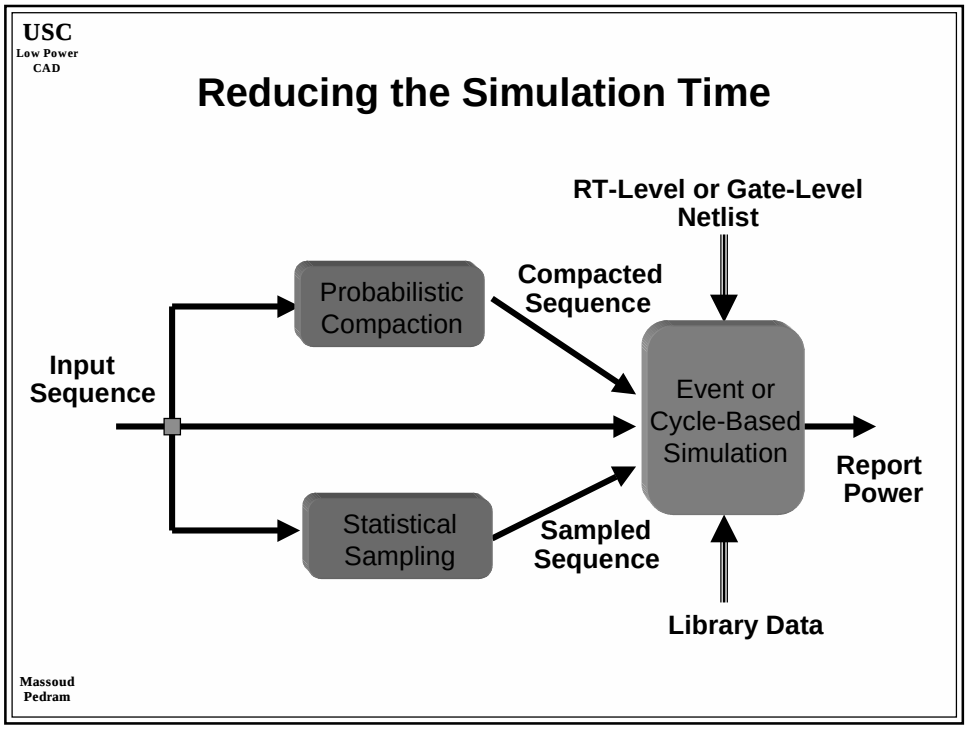
Sequential Circuits

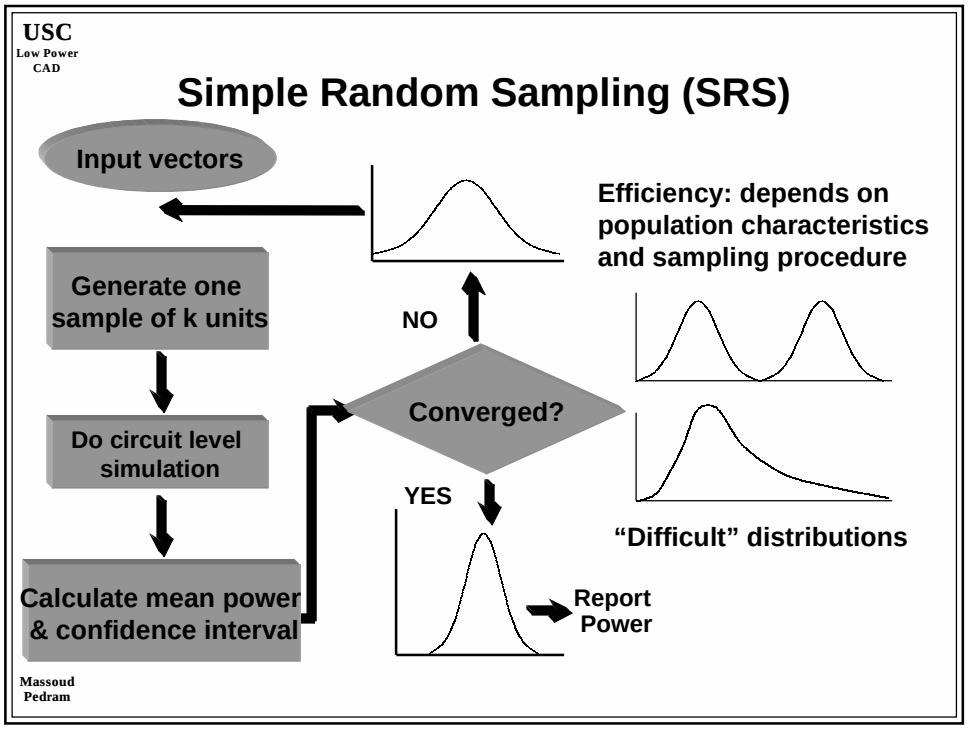


Two issues for sequential circuits:
- **Probability of machine being in a particular state**
- **Correlation in time: Given a primary input and present state, next state is uniquely determined by the next state combinational logic**

Solutions:
- **Solve C-K Equations for calculating steady state probabilities**
- **Unroll the state machine**

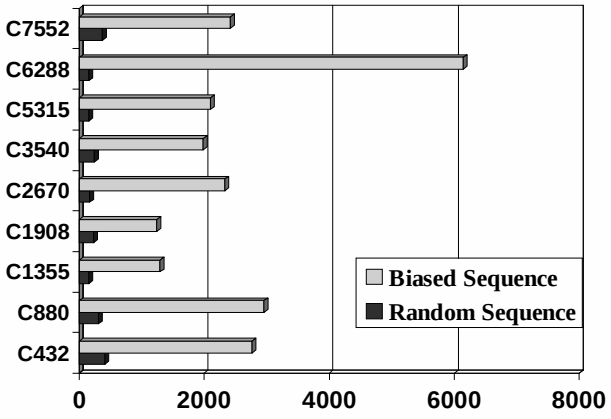
Massoud Pedram


USC
Low Power CAD
Reducing the Simulation Time
RT-Level or Gate-Level Netlist
Probabilistic Compaction
Compacted Sequence
Input Sequence
Event or Cycle-Based Simulation
Report Power
Statistical Sampling
Sampled Sequence
Library Data
Massoud Pedram


USC
Low Power CAD
Simple Random Sampling (SRS)
Input vectors
Efficiency: depends on population characteristics and sampling procedure
Generate one sample of k units
NO
Converged?
Do circuit level simulation
YES
"Difficult" distributions
Calculate mean power & confidence interval
Report Power
Massoud Pedram

USC
Low Power CAD

SRS Results

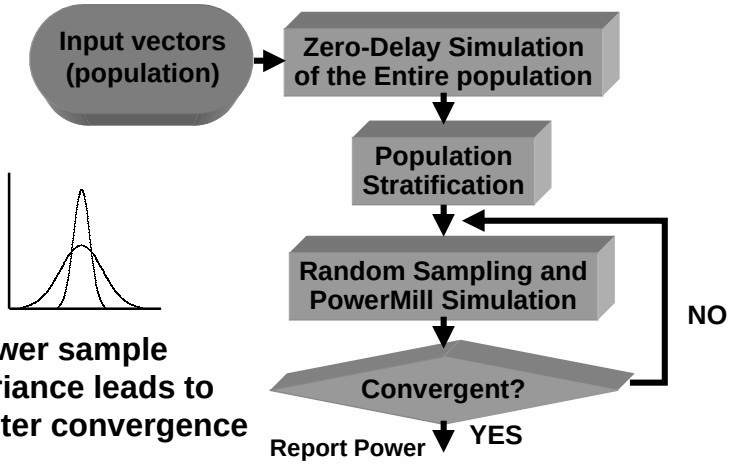


Massoud Pedram

USC
Low Power CAD

Stratified Random Sampling (StRS)

Estimate the transistor-level power, assuming that the zero-delay power estimate is readily available



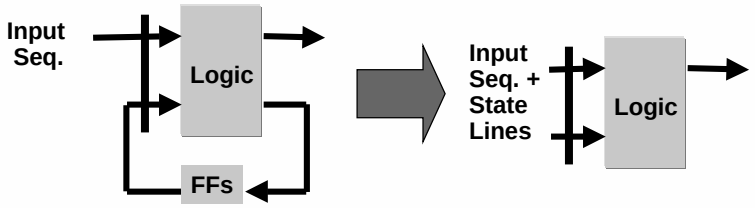
Lower sample variance leads to faster convergence

Massoud Pedram

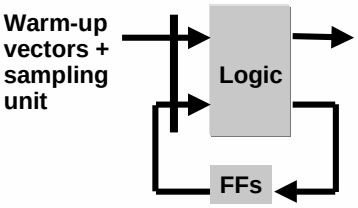
USC
Low Power CAD

Sampling on FSMs

Method 1: Functional simulation + sampling on comb. circuit



Method 2: Do machine warm-up for every sampling unit



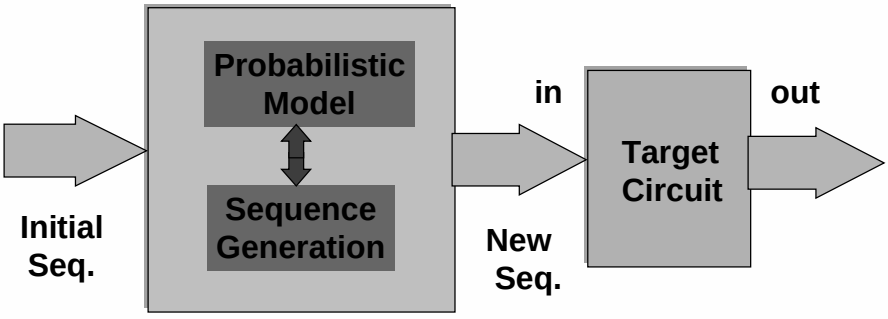
Warm-up sequence length can be quite large

Massoud Pedram

USC
Low Power CAD

Probabilistic Compaction (PC)

Sequence compaction : Generate a new, but shorter, input sequence which exhibits nearly the same spatio-temporal correlations as the initial sequence



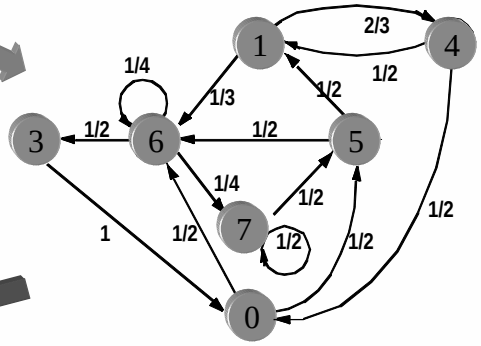
Massoud Pedram

USC
Low Power CAD

PC by Example

Initial sequence
01101010011111010
11100000110111000
10000110101110101
Avg. bit activity
23/16

Compacted sequence
001010
011000
010001
Avg. bit activity
8/5



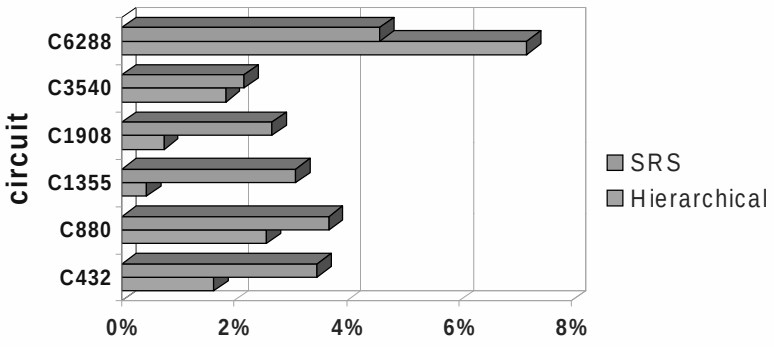
Stochastic State Machine Model

Massoud Pedram

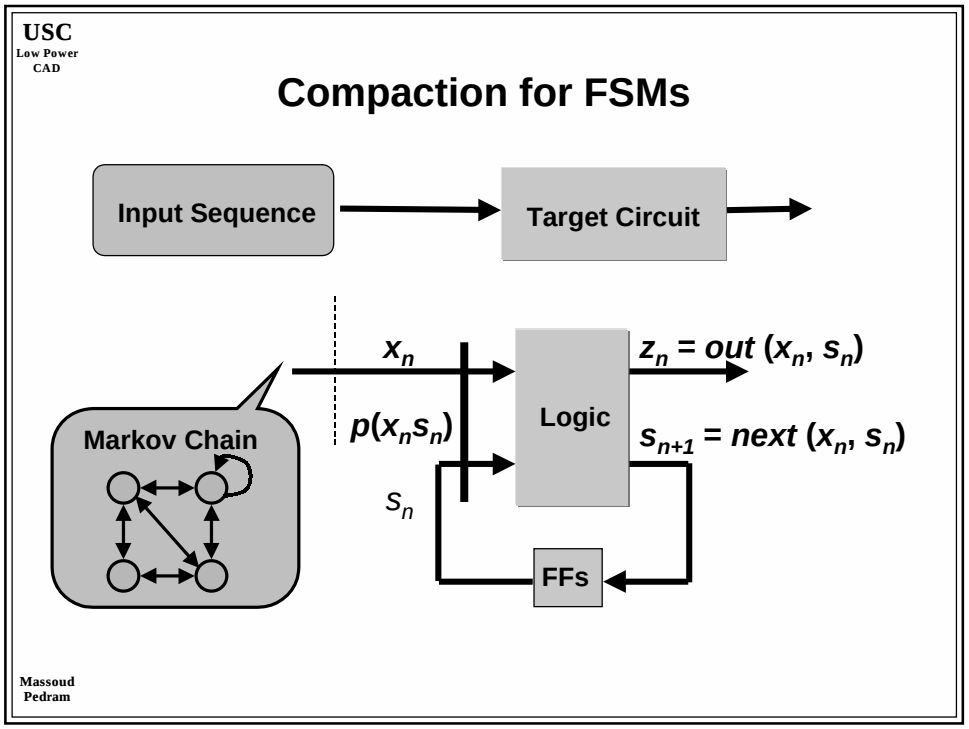
USC
Low Power CAD

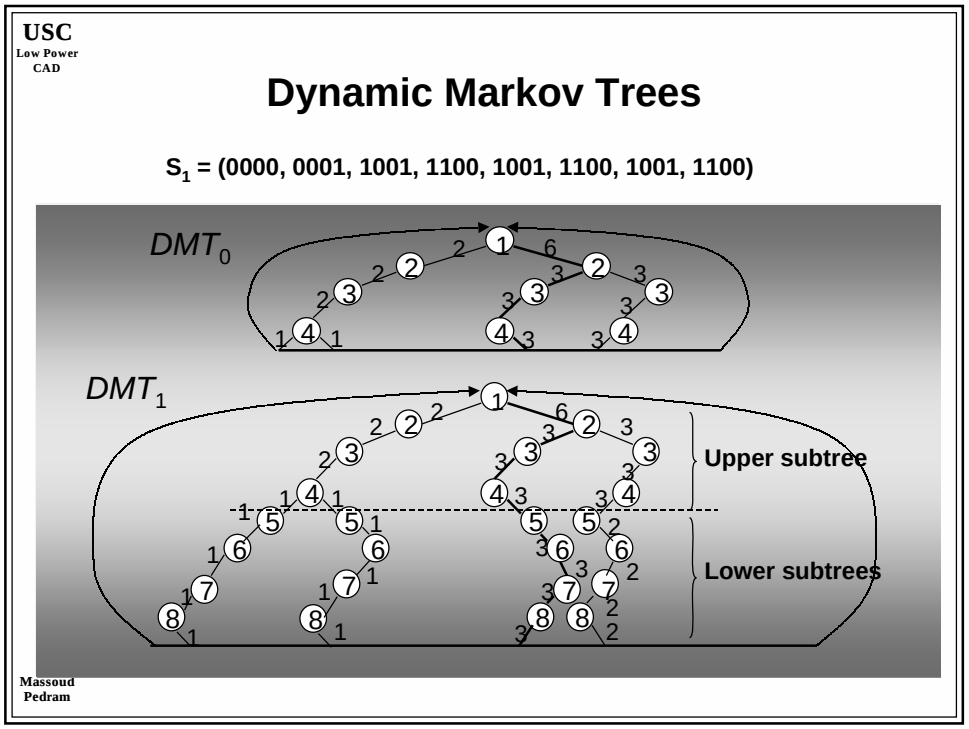
Comparison with SRS

L = 4,000
Compaction Ratio 10



Massoud Pedram

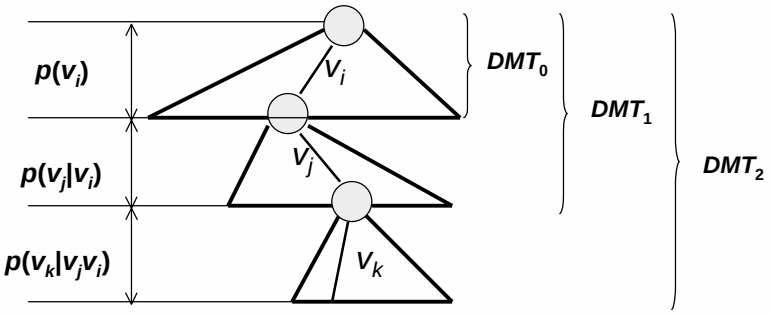

USC
Low Power CAD
Compaction for FSMs
Input Sequence
Target Circuit
x_n
$p(x_n s_n)$
s_n
Markov Chain
Logic
$z_n = out\ (x_n, s_n)$
$s_{n+1} = next\ (x_n, s_n)$
FFs
Massoud Pedram


USC
Low Power CAD
Dynamic Markov Trees
S_1 = (0000, 0001, 1001, 1100, 1001, 1100, 1001, 1100)
DMT_0
DMT_1
Upper subtree
Lower subtrees
Massoud Pedram

USC
Low Power CAD

Higher Order DMTs

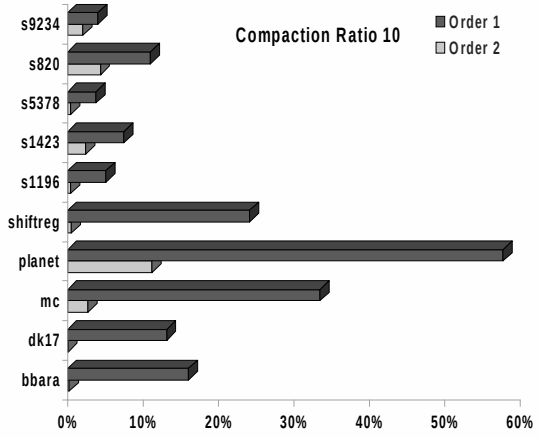
A lag-k Markov chain which correctly models the input sequence, also models the joint *k*-step conditional probabilities of the primary inputs and state lines



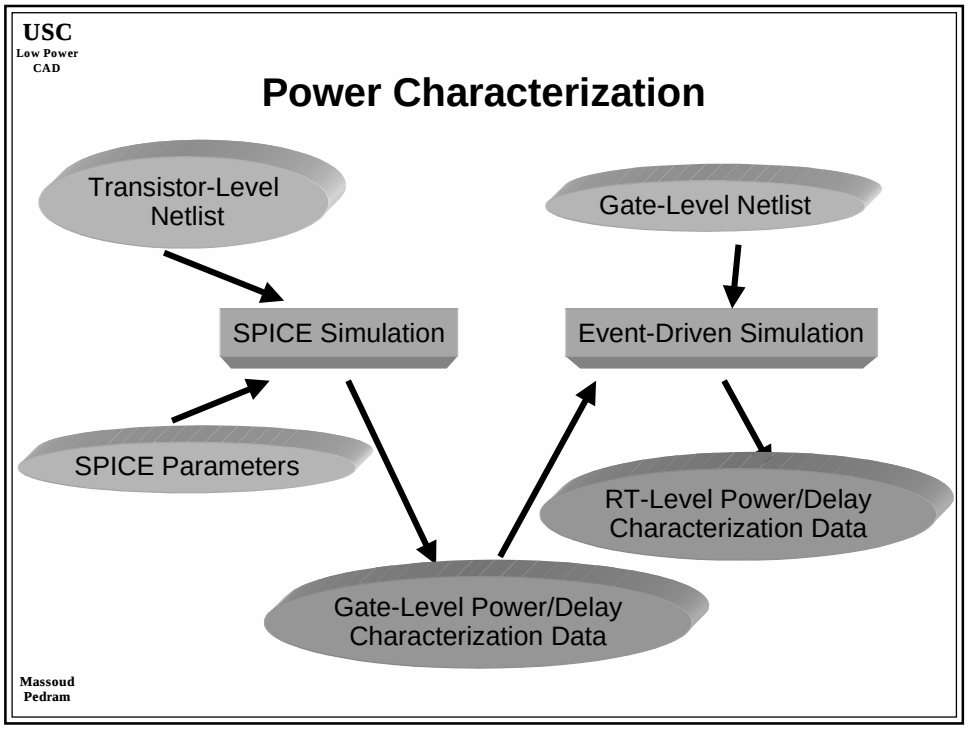
Massoud Pedram

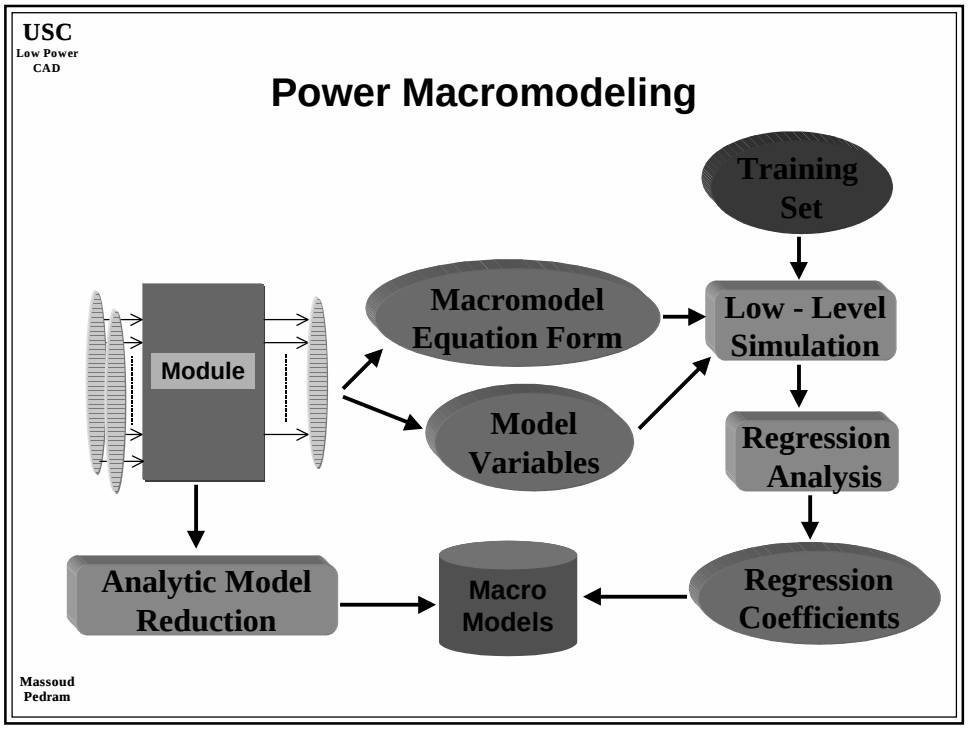
USC
Low Power CAD

High Order DMT Results



Massoud Pedram


USC
Low Power CAD
Power Characterization
Transistor-Level Netlist
Gate-Level Netlist
SPICE Simulation
Event-Driven Simulation
SPICE Parameters
RT-Level Power/Delay Characterization Data
Gate-Level Power/Delay Characterization Data
Massoud Pedram

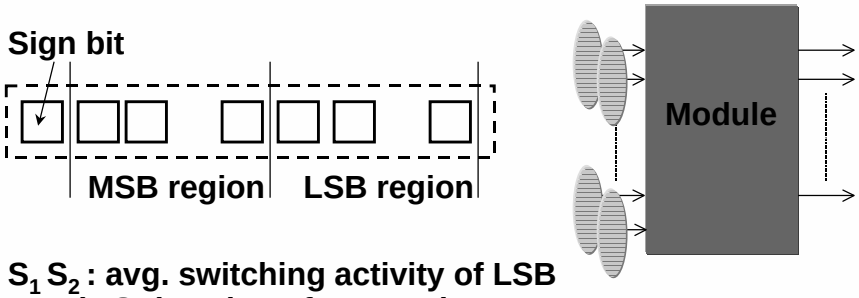

USC
Low Power CAD
Power Macromodeling
Training Set
Macromodel Equation Form
Low - Level Simulation
Module
Model Variables
Regression Analysis
Analytic Model Reduction
Macro Models
Regression Coefficients
Massoud Pedram

USC
Low Power CAD

Dual Bit Type Model

Consider a data path block:

$$Pwr = C_0 + C_1 \cdot S_1 + C_2 \cdot S_2 + C_3 \cdot S_3 + C_4 \cdot S_4$$



S_1 S_2 : avg. switching activity of LSB (MSB) region of operand 1

S_3 S_4 : avg. switching activity of LSB (MSB) region of operand 2

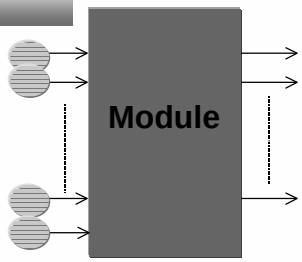
Massoud Pedram

USC
Low Power CAD

Bitwise Data Model

Consider a random logic block:

$$Pwr = C_0 + \sum_{inputs} C_i \cdot S_i$$



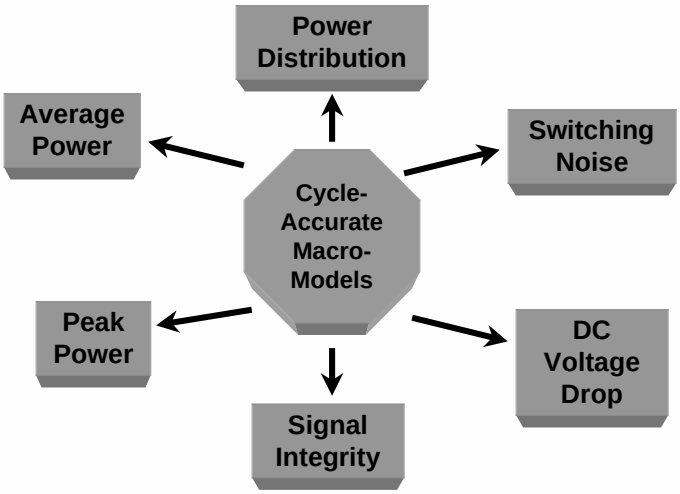
S_i : avg. switching activity of input signal i

More parameters lead to a higher degree of accuracy, but increase the computational overhead

Massoud Pedram

USC
Low Power CAD

Cycle-Accurate Macro-Models



Massoud Pedram

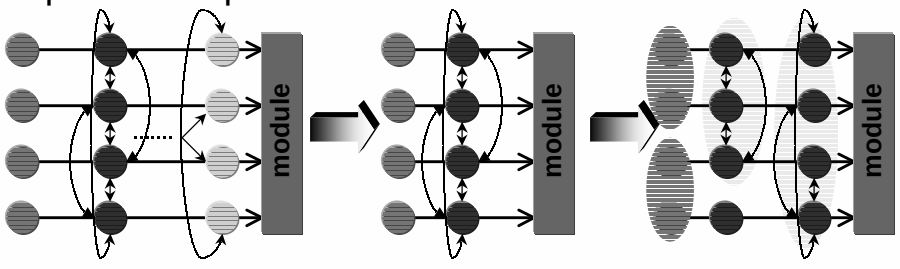
USC
Low Power CAD

Macro-Model Construction

Starting point: an exact power consumption function

$$Pwr = f(V_1, V_2) =$$ Order 1 terms + Order 2 terms + ... + Order n terms

Equation simplification



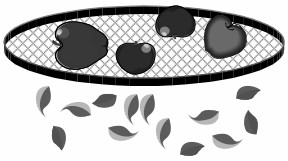
Exact function | After order reduction | After input grouping

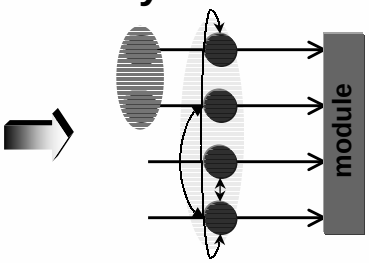
Massoud Pedram

USC
Low Power CAD

Regression Analysis

Variable reduction: find the "most significant" variables, by using a statistical test





After variable reduction

Population stratification: makes the macro-model accuracy less sensitive to different input conditions

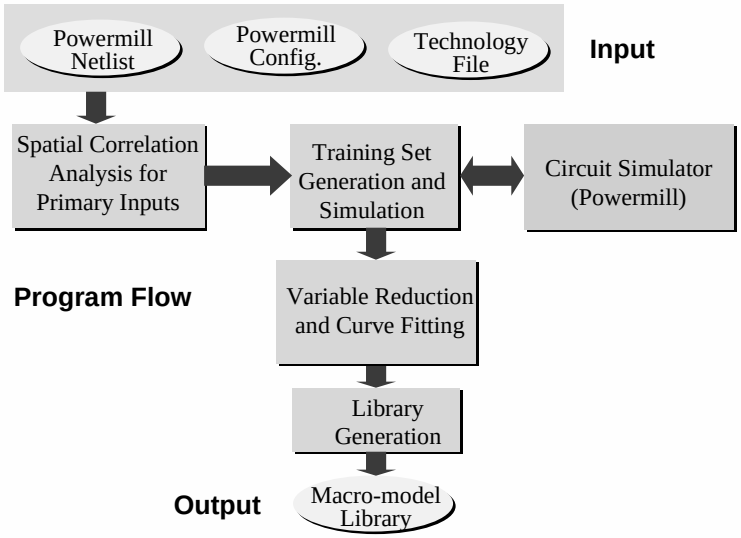
Training set design: reduces the macro-model bias

A general linear regression model: $Pwr = C_0 + \sum_{i=1}^{k} C_i \cdot X_i$

Massoud Pedram

USC
Low Power CAD

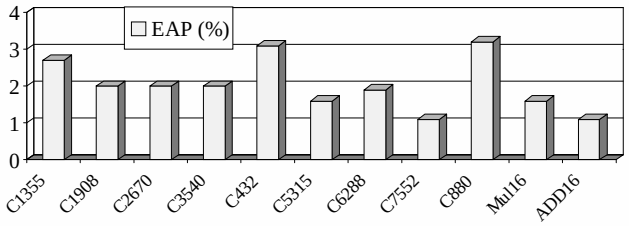
Integrated Macro-Modeling Flow



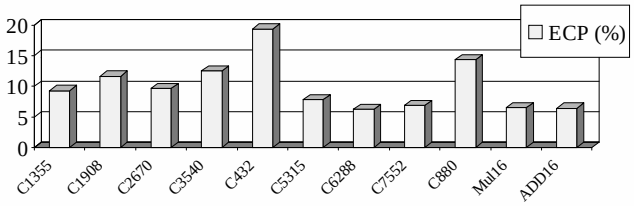
Massoud Pedram

USC
Low Power CAD

Macro-Modeling Results



Average EAP = 2.03%

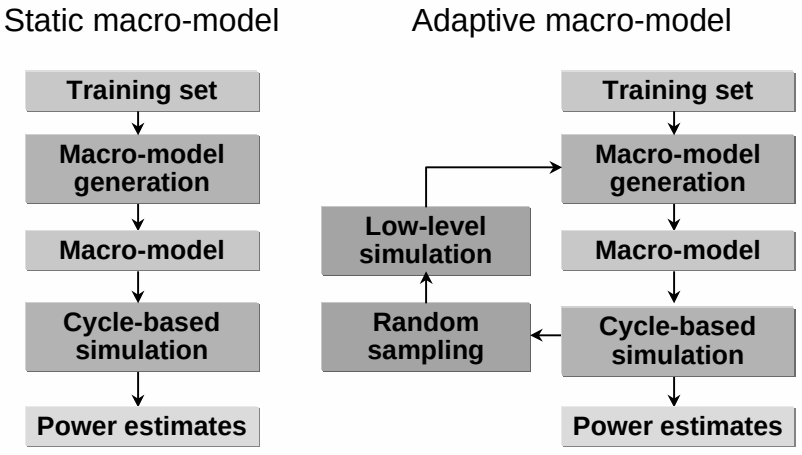


Average ECP = 10.04%

Massoud Pedram

USC
Low Power CAD

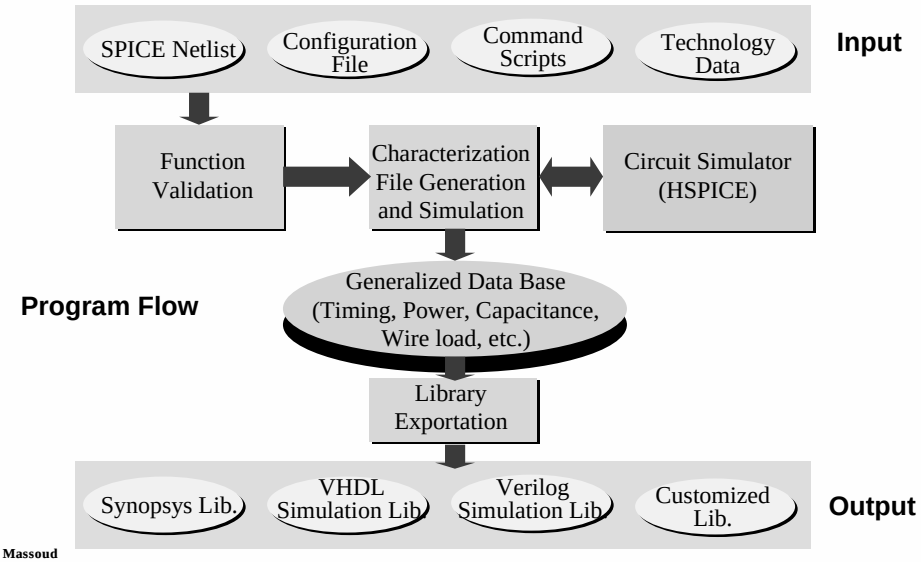
Adaptive Macro-Modeling



Massoud Pedram

USC
Low Power CAD

Standard Cell Characterization



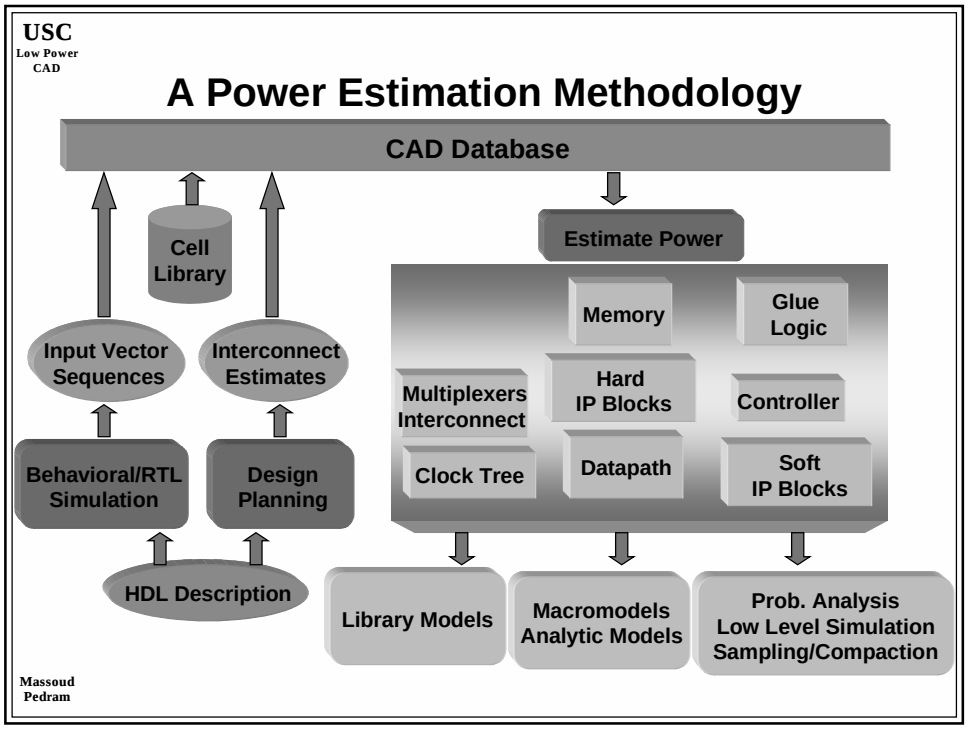
Massoud Pedram

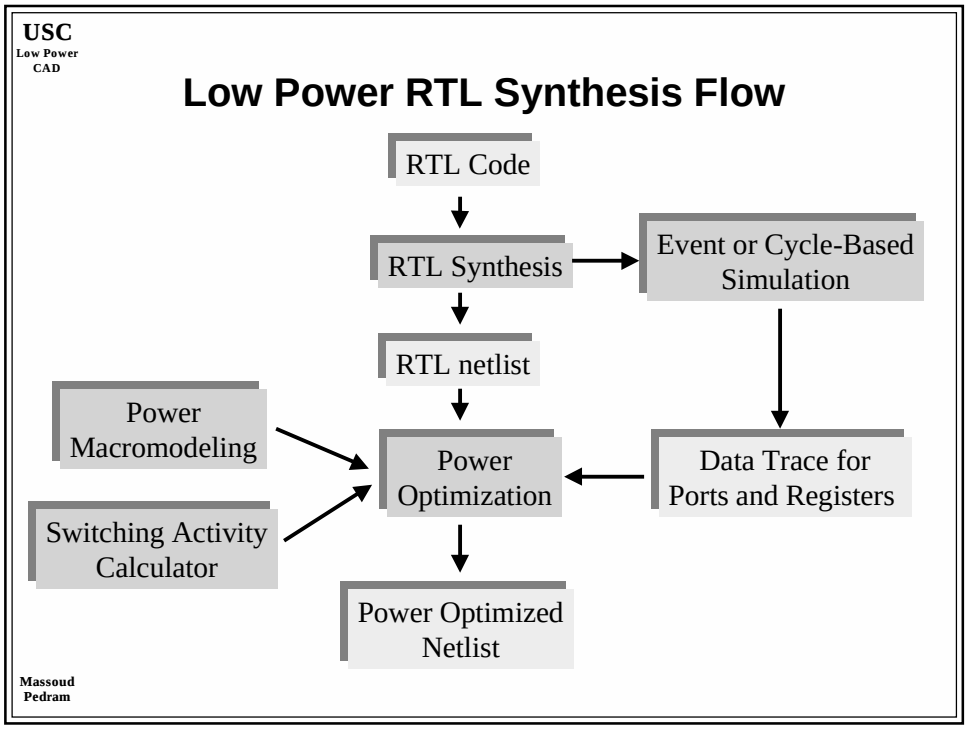
USC
Low Power CAD

Design Tool Characteristics

- **Accurate model**
  - **Include both output load (*C*) and input slope (*S*)**
  - **Nonlinear equation $f=k_0+k_1*C+k_2*C^2+k_3*S+k_4*C*S$**
  - **Variable domain stratification, different sets of coefficients for different strata**
  - **Default sensitization method is pin-to-pin, can be extended to pattern-dependent by customization**
- **Fast model**
  - **Small equation form**
  - **Can be customized to full table-lookup library**
- **Easy library maintenance**
  - **Manage the libraries through Graphic User Interface**
  - **Generalized data base enables data reuse**

Massoud Pedram


USC
Low Power CAD
A Power Estimation Methodology
CAD Database
Cell Library
Estimate Power
Memory
Glue Logic
Input Vector Sequences
Interconnect Estimates
Multiplexers Interconnect
Hard IP Blocks
Controller
Clock Tree
Datapath
Soft IP Blocks
Behavioral/RTL Simulation
Design Planning
HDL Description
Library Models
Macromodels Analytic Models
Prob. Analysis Low Level Simulation Sampling/Compaction
Massoud Pedram


USC
Low Power CAD
Low Power RTL Synthesis Flow
RTL Code
RTL Synthesis
Event or Cycle-Based Simulation
RTL netlist
Power Macromodeling
Power Optimization
Data Trace for Ports and Registers
Switching Activity Calculator
Power Optimized Netlist
Massoud Pedram

USC
Low Power CAD

RTL Optimization for Low Power

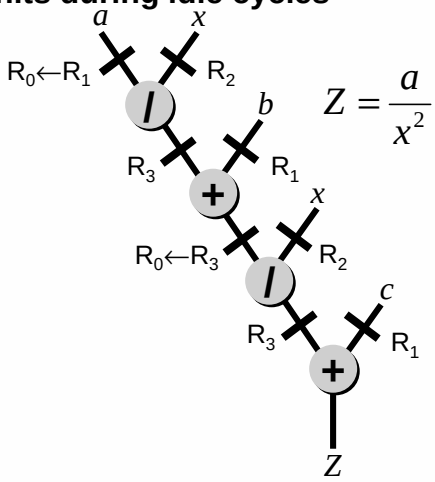
- Resource Assignment
  - ⇨Module and/or register allocation and binding
- Clock Gating
- State Encoding
  - ⇨Two-level and multi-level logic realizations
- Multiple Supply Voltage Scheduling
  - ⇨Pipelined and non-pipelined designs

Massoud Pedram

USC
Low Power CAD

Register Assignment for Low Power

Avoid value change at the input of functional units during idle cycles

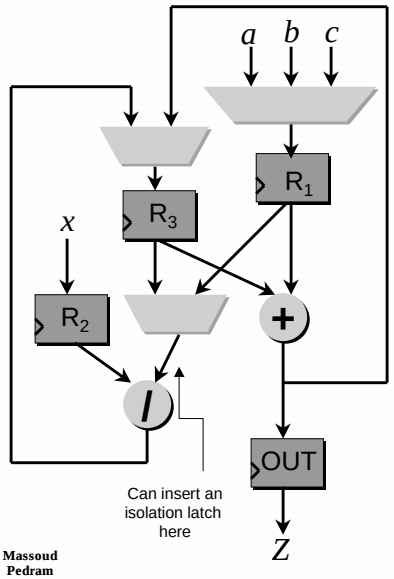


$$Z = \frac{a}{x^2} + \frac{b}{x} + c = \frac{1}{x}\left(\frac{a}{x} + b\right) + c$$

Massoud Pedram

USC
Low Power CAD

An RTL Structure

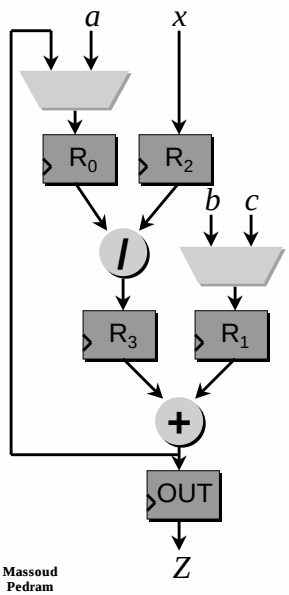


Cycle	R_1	R_2	R_3	/	+
1	a	x		$< a, x >$	
2	b	x	a/x	$< R_1 \text{ or } R_3, x >$	$< b, \frac{a}{x} >$
3	b	x	$a/x+b$	$< \frac{a}{x}+b, x >$	
4	c	x	$\frac{a}{x^2}+\frac{b}{x}$	$< R_1 \text{ or } R_3, x >$	$< \frac{a}{x^2}+\frac{b}{x}, c >$
5					

Massoud Pedram

USC
Low Power CAD

A Low Power RTL Structure



Cycle	R_0	R_1	R_2	R_3	/	+
1	a	b	x		$< a, x >$	
2	a	b	x	a/x	$< a, x >$	$< b, \frac{a}{x} >$
3	$a/x+b$	c	x	$a/x+b$	$< \frac{a}{x}+b, x >$	
4	$a/x+b$	c	x	$\frac{a}{x^2}+\frac{b}{x}$	$< \frac{a}{x}+b, x >$	$< \frac{a}{x^2}+\frac{b}{x}, c >$
5						

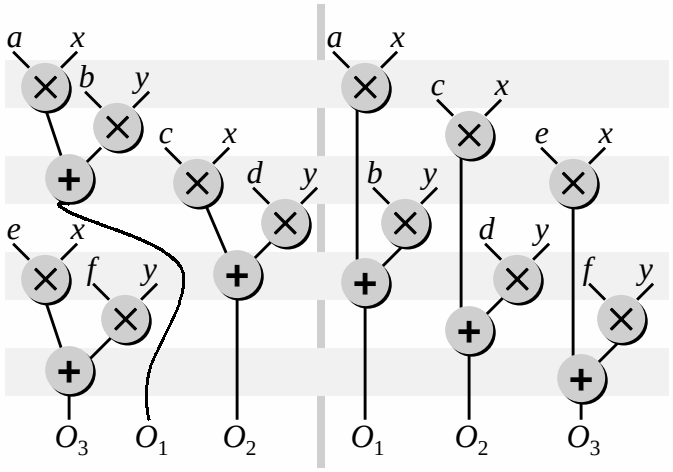
Massoud Pedram

USC
Low Power CAD

Module Assignment for Low Power

Do module binding to minimize the input toggling of modules

$$O_1 = a \cdot x + b \cdot y$$
$$O_2 = c \cdot x + d \cdot y$$
$$O_3 = e \cdot x + f \cdot y$$

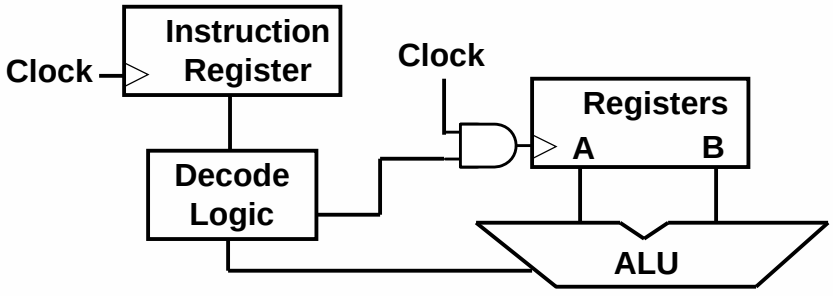


Massoud Pedram

USC
Low Power CAD

Gated Clocks

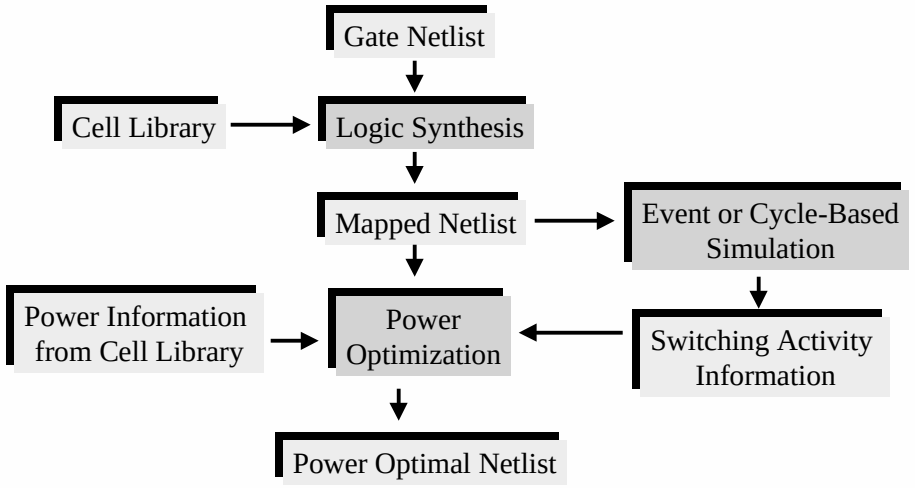
Disable clocking of input registers if the current instruction does not access it



Massoud Pedram

USC
Low Power CAD

Low Power Logic Synthesis Flow



Massoud Pedram

USC
Low Power CAD

Logic Optimization for Low Power

- **Combinational logic optimization**
  - **Technology Mapping**
  - **Gate Resizing**
  - **Pin Assignment**
- **Sequential logic optimization**
  - **Precomputation**
  - **State Re-encoding**
  - **Power management for control units**

Massoud Pedram

USC
Low Power CAD

Technology Mapping

Outputs of nodes of combinational circuit mapped to library gates under the cost function of load capacitance multiplied by the switching activity

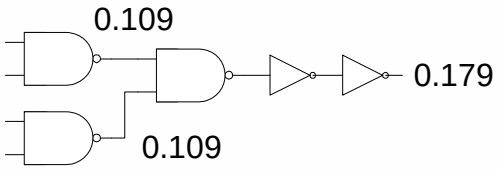
To minimize power dissipation, high switching activity points are either hidden within gates or driven by smaller gates

Minimum power realization under zero delay model can be obtained using dynamic programming

Massoud Pedram

USC
Low Power CAD

Circuit and Gate Library



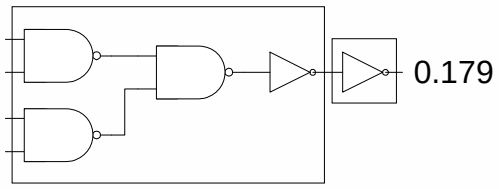
Gate	Area	Intrinsic Capacitances	Input Load Capacitances
INV	928	0.1029	0.0514
NAND2	1392	0.1421	0.0747
AOI22	2320	0.3410	0.1033

Massoud Pedram

USC
Low Power CAD

Minimum Area Mapping

AOI22



Area = 2320 + 928 = 3248

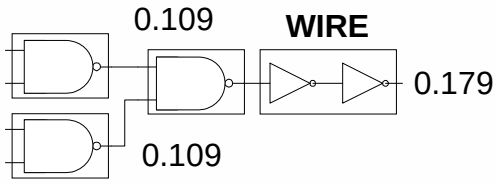
Power = 0.179 (0.3410 + 0.0514) + 0.179 (0.1029)
= 0.0887

Note: Ignore capacitances at circuit inputs

Massoud Pedram

USC
Low Power CAD

Minimum Power Mapping



Area = 1392 * 3 = 4176

Power = 0.109 (0.1421 + 0.0747) * 2 + 0.179 (0.1421)
= 0.0726

Massoud Pedram

USC
Low Power CAD

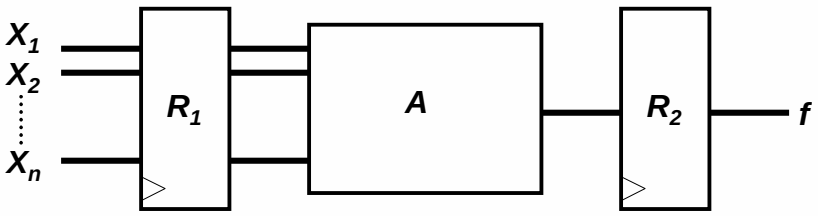
Sequential Precomputation

- Selectively precompute the outputs of the logic circuit one clock cycle before they are required and use the precomputed values to reduce switching activity in the next clock cycle

Massoud Pedram

USC
Low Power CAD

Original Circuit



Predictor functions:

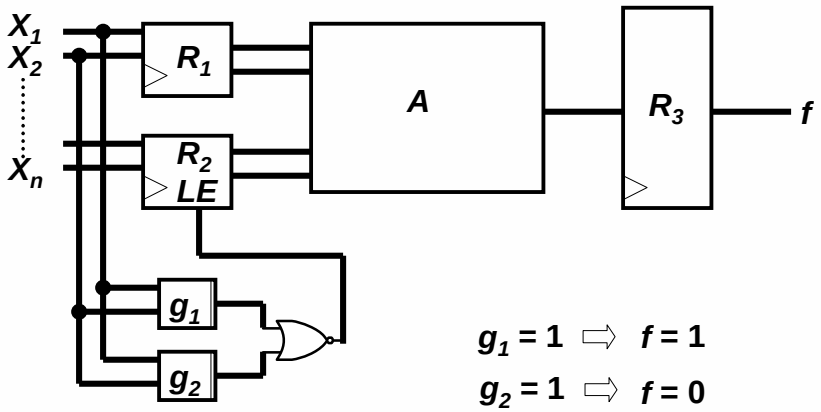
$g_1 = 1 \Rightarrow f = 1$

$g_2 = 1 \Rightarrow f = 0$

Massoud Pedram

USC
Low Power CAD

Subset Input Disabling Architecture



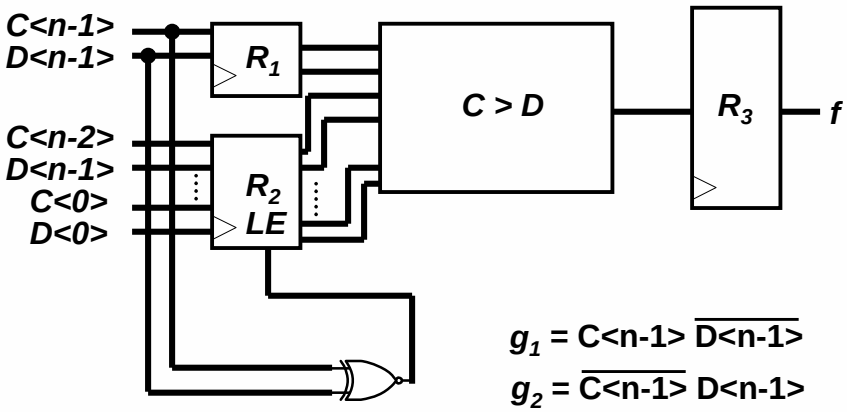
$g_1 = 1 \Rightarrow f = 1$

$g_2 = 1 \Rightarrow f = 0$

Massoud Pedram

USC
Low Power CAD

A Comparator Example



$g_1 = C\langle n-1\rangle \overline{D\langle n-1\rangle}$

$g_2 = \overline{C\langle n-1\rangle} D\langle n-1\rangle$

Massoud Pedram

USC
Low Power CAD

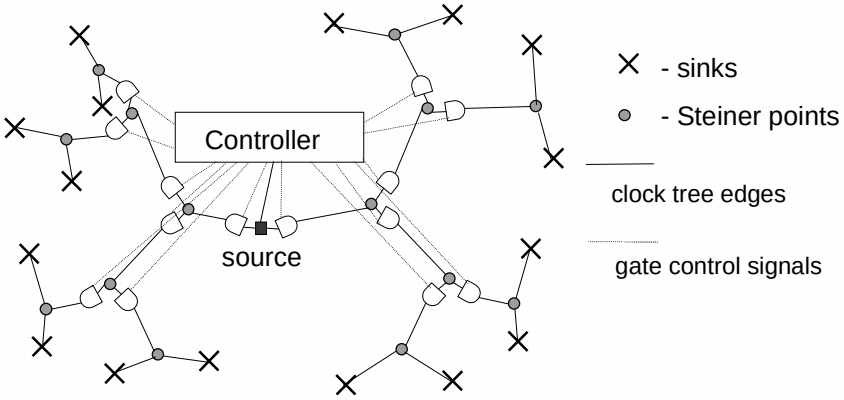
Layout Optimization for Low Power and Noise

- Floorplanning and/or Placement
- Gate Resizing
- Gated Clock Tree Design
- Driver Design
- Chip-Package Interface Design
- Power Bus Planning and Sizing

Massoud Pedram

USC
Low Power CAD

Gated Clock Tree Design

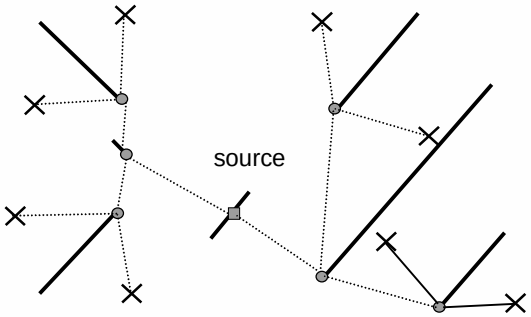


- Clock tree : binary tree
- Controller tree : star tree

Massoud Pedram

USC
Low Power CAD

Clock Tree Layout

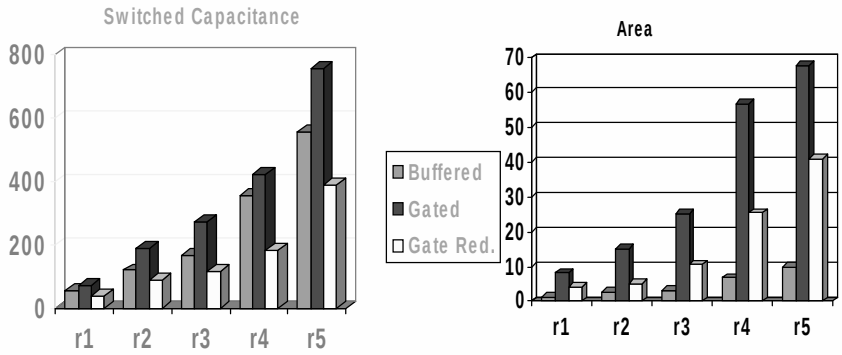


- ⇨ **Use the Deferred Merge Embedding algorithm of to layout the clock tree**
- ⇨ **The bottom-up merging process is guided by the minimum switched capacitance heuristic**
- ⇨ **Top-down phase determines the exact locations of the Steiner nodes in their respective merging sectors**

Massoud Pedram

USC
Low Power CAD

Buffered clock tree versus Gated clock tree



- ⇨ ***The average module activity was set to 40%***

Massoud Pedram

USC
Low Power CAD

Conclusions

- **Analysis tools :**
  - ⇨ **Simulation based techniques provide the best accuracy**
  - ⇨ **Transistor level power estimation is mature; a satisfactory gate and RT level simulation engine is needed**
  - ⇨ **Probability-based techniques are good for relative accuracy analysis and to guide the synthesis engine**
- **Optimization Tools:**
  - ⇨ **Gate level optimization / re-mapping provides 20-40% power reduction**
  - ⇨ **Behavioral and RTL power synthesis provides 40-60% power reduction**
  - ⇨ **System-level optimization provides 2-5 X power reduction**
- **Power analysis, tradeoffs and optimizations for design levels higher than RTL are not mature**

Massoud Pedram