High-Level Synthesis for Low Power

Saraju P. Mohanty
smohanty@csee.usf.edu
Dept. of CSE, USF, Tampa, FL

University of South Florida

Major Motivation : Extending battery life for portable applications

Outline of the Talk

1. Sources of power dissipation
2. Dynamic power dissipation details
3. How to reduce dynamic power ?
4. Effect of Frequency on Energy / Power
5. Few low-power research works

University of South Florida

Design Quality Measures

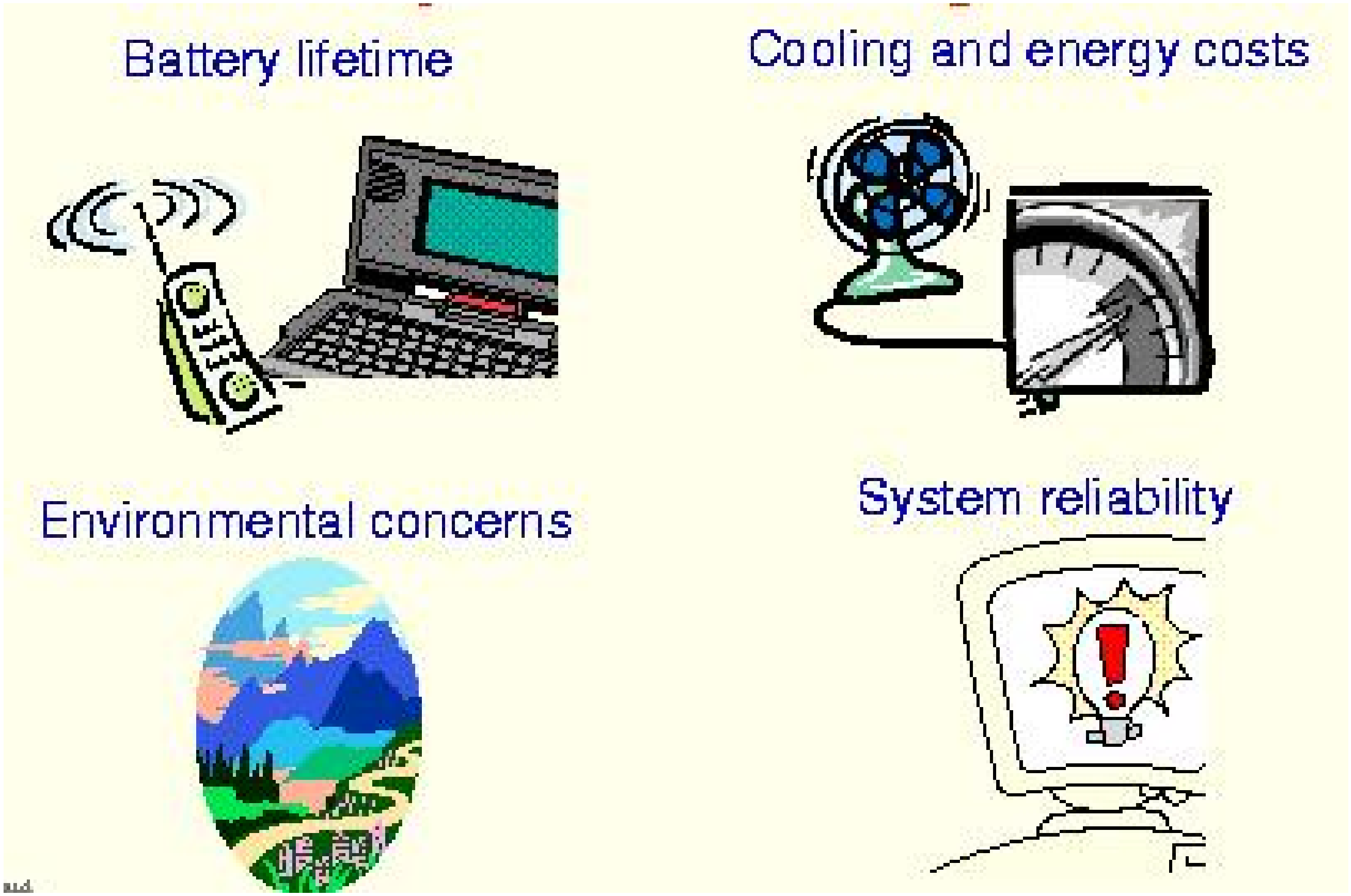
- Area
- Performance
- Power
- Testability(Observability/Controllability)
- Verifiability
- Reliability
- Manufacturability

Quality Measures Vs Objective Functions

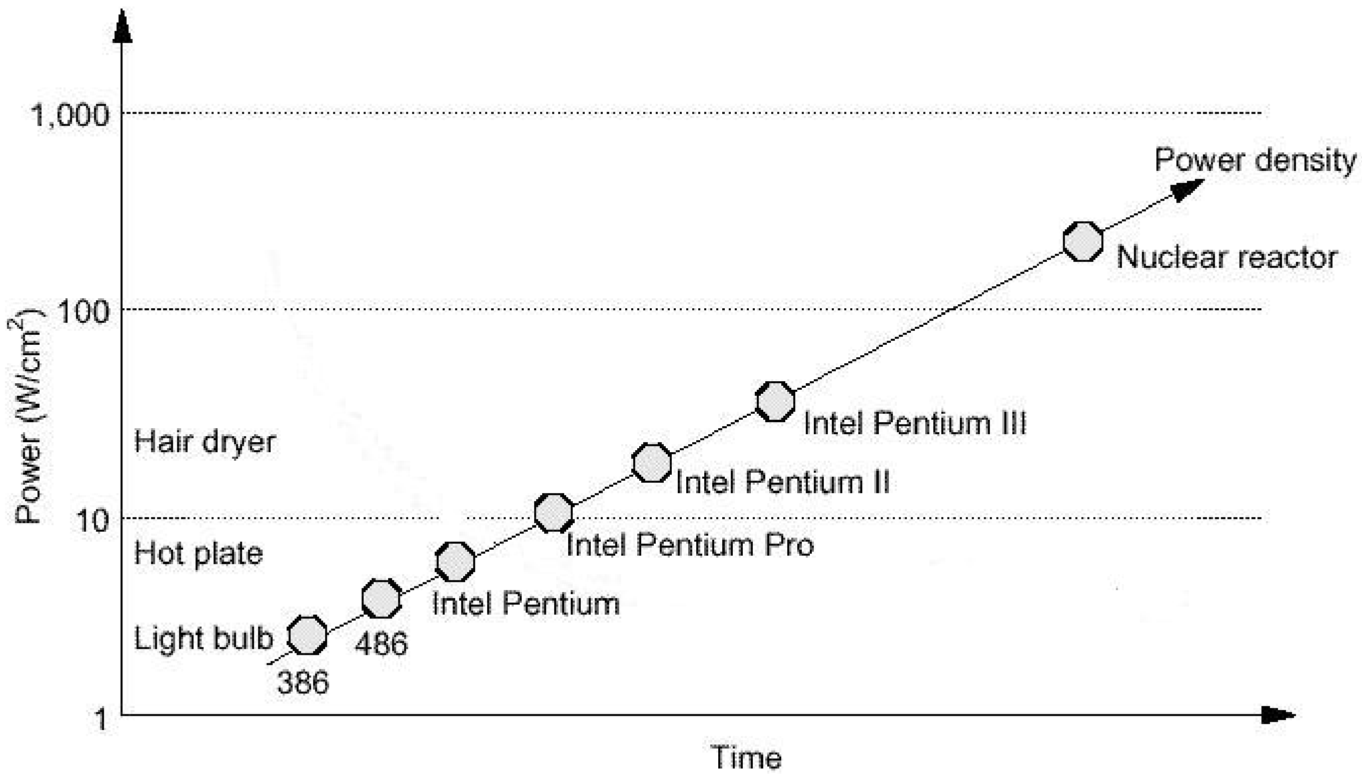
- Quality measures are used to support design decisions
- In order to be acceptable the design must pass the quality measure
- So the designer has to take the quality measures into account in the form of objective functions during the design process

Power is the objective function for low-power design !!

Why Low-Power Synthesis ??



Power Density



Note : chronological time

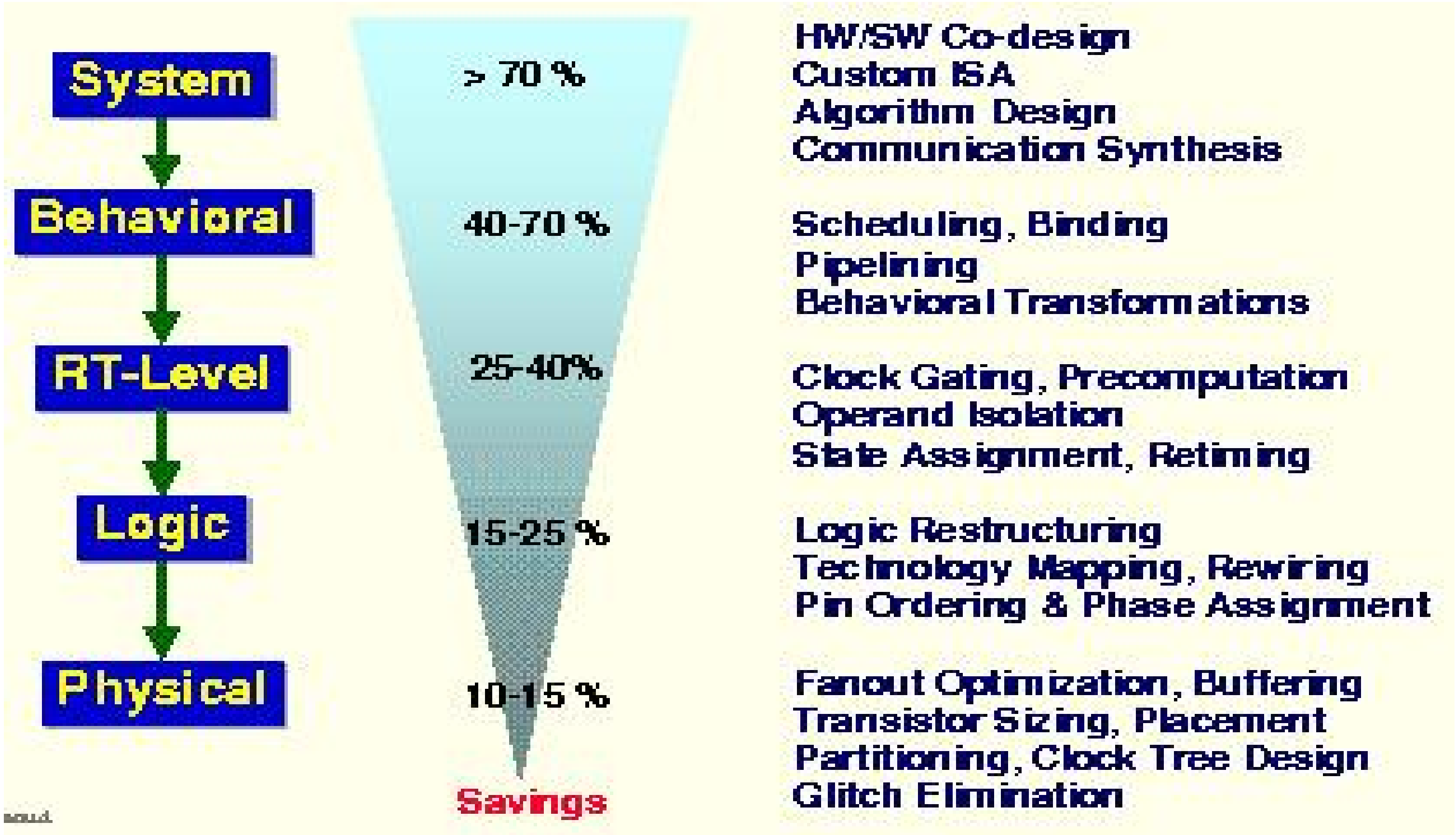
What to reduce for low power ??

- Battery life : Energy / PDP (power-delay-product)
- Battery life and Delay: Action (Energy-delay-product)
- Reliability: Peak Power
- Packaging and Cooling cost, Environment: Average power

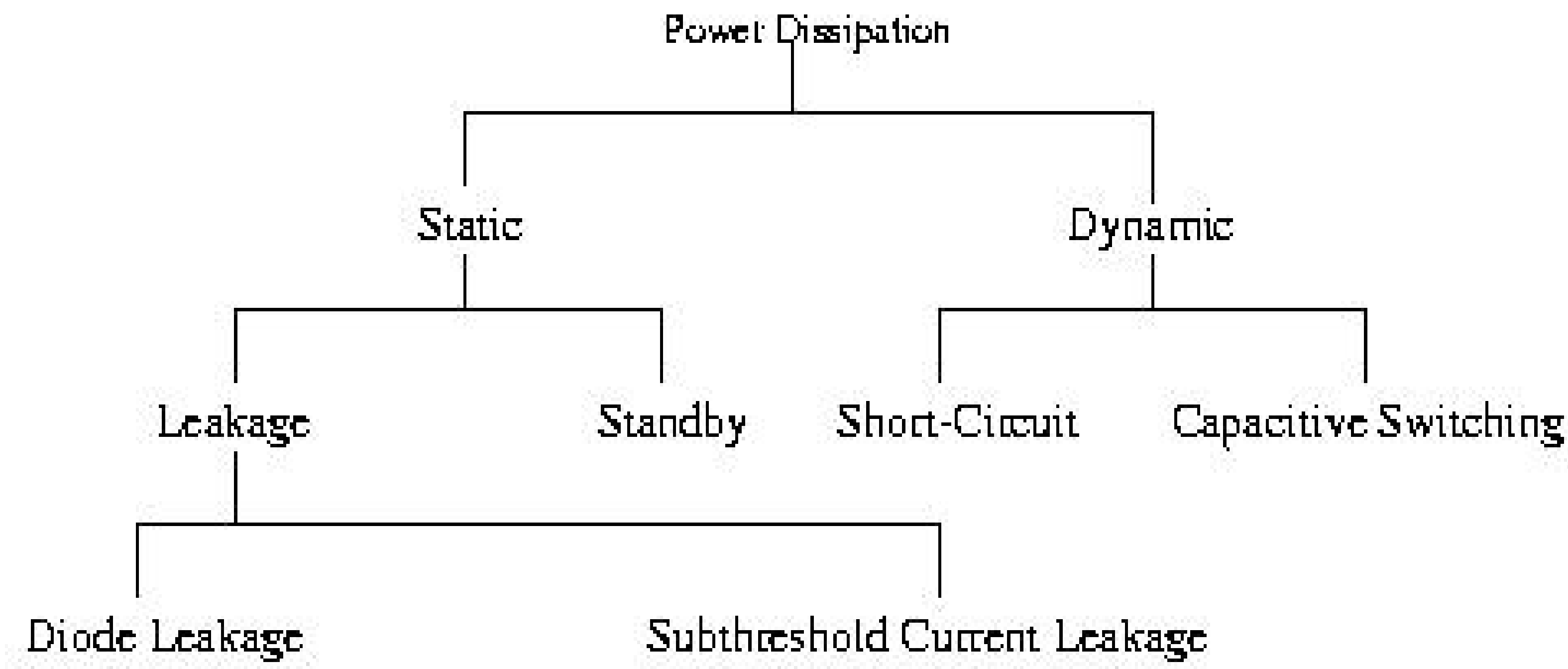
Low-power design: Key Principles

- using the lowest possible supply voltage
- using the smallest geometry, highest frequency devices, but operating them at lowest possible frequency
- using parallelism and pipelining to lower required frequency of operation
- power management by disconnecting the power source when the system is idle
- designing systems to have lowest requirements on subsystem performance for the given user level functionality

Why High-Level Low-Power Synthesis ??



Sources of Power Dissipation (CMOS)



Major component: Dynamic Power

Power Dissipation …….

Leakage Current: This is determined by the fabrication technology and consists of reverse bias current in the parasitic diodes and sub-threshold current that arises from the inversion charges that exists at the gate voltages below the threshold voltage.

Standby Current: It is the DC current drawn continuously from V_{dd} to ground.

Short-Circuit Current: This is the current due to the DC path between the supply and ground during output transition.

Capacitance Current: This flows to charge and discharge capacitance loads during logic changes.

Dynamic Power: Major one

$$P_{dynamic} = \frac{1}{2} C_L V_{dd}^2 N f$$

C_L = load capacitor, V_{dd} = supply voltage,

N = average number of transitions/clock cycle

= E(sw) = 2 $\alpha_{0->1}$ = switching activity

f = clock frequency

Note:

1. N * f is transition density
2. C_L * N (= C_{sw} = C_{eff}) is the effective switching capacitance

Dynamic Power (How to reduce ??)

$$P_{dynamic} = \frac{1}{2} C_L V_{dd}^2 N f$$

- Reduce Supply Voltage (V_{dd})
- Reduce Clock Frequency (f)
- Reduce Switching Activity (N / E(sw))
- Reduce Capacitance (C_L)

Note : we can't reduce all parameters ☹

WHY ???

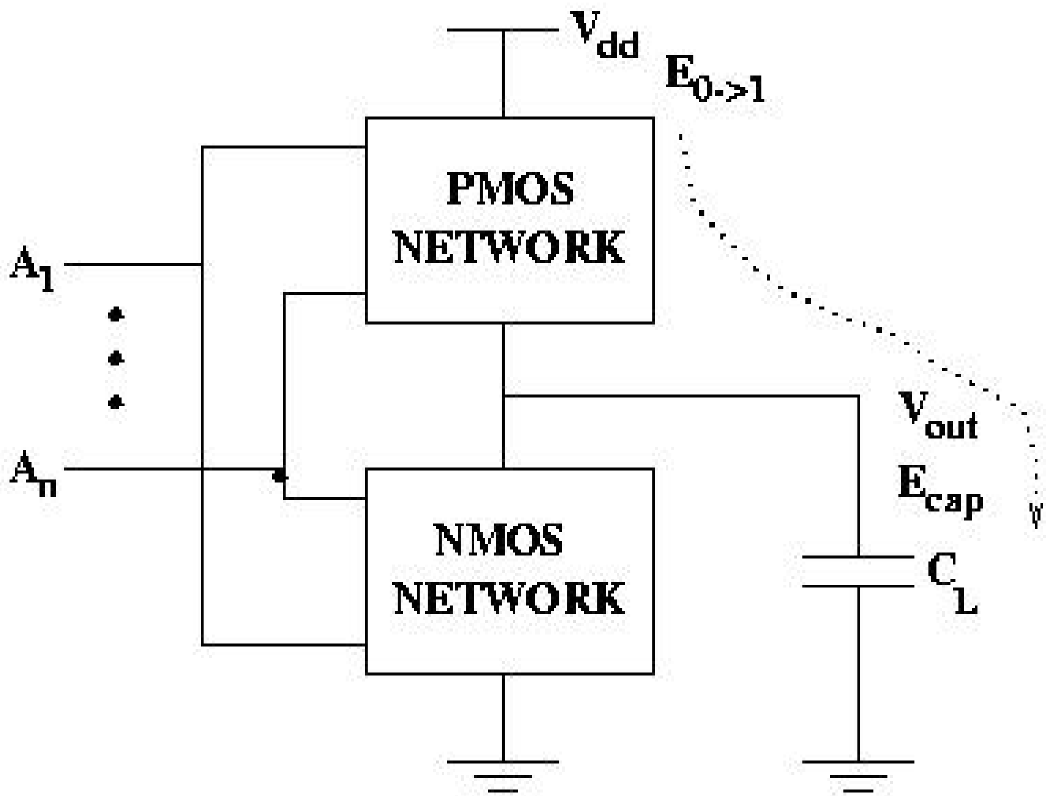
University of South Florida

Reduce V_{dd}, f, N, C_{eff} …..

- Reduce Supply Voltage (V_{dd}) : delay increases ; performance degradation
- Reduce Clock Frequency (f) : only power saving no energy
- Reduce Switching Activity (N / E(sw)) : no switching no power loss !!! Not in fully under designers control. Switching activity depends on the logic function. Temporal/and spatial correlations difficult to handle.
- Reduce Physical Capacitance : done by reducing device size reduces the current drive of the transistor making the circuit slow
- Reduce C_{eff}/C_{sw} : adversely affects the maximum clock frequency

University of South Florida

Power/Energy Dissipation: more details



A general CMOS transistor circuit

Energy Dissipation: Details ……..

$$E_{0\text{->}1} = \int_0^T P(t)\,dt = V_{dd}\int_0^T i_{supply}(t)\,dt = V_{dd}\int_0^{V_{dd}} C_L\,dV_{out} = C_L V_{dd}^2$$

$$E_{out} = \int_0^T P_{out}(t)\,dt = \int_0^T V_{out}\,i_{out}(t)\,dt = \int_0^{V_{dd}} C_L V_{out}\,dV_{out} = \frac{1}{2} C_L V_{dd}^2$$

Note:

1. the difference between the two is the loss
2. Energy doesn’t depend on frequency

Energy to Power

For $\mathbf{N_c}$ clock cycles energy loss :

$$E_{N_c} = C_L V_{dd}^2 n(N_c)$$

$\mathbf{n(N_c)}$: is the number of 0->1 transitions in N_c clock cycles

$$P_{avg} = \lim_{N \to \inf} \left[\frac{E_{N_c}}{N_c} \right] f = \left[\lim_{N \to \inf} \frac{n(N_c)}{N_c} \right] C_L V_{dd}^2 f$$

$$= \alpha_{0->1} C_L V_{dd}^2 f$$

Note: Power depends on frequency

Propagation Delay Vs Frequency

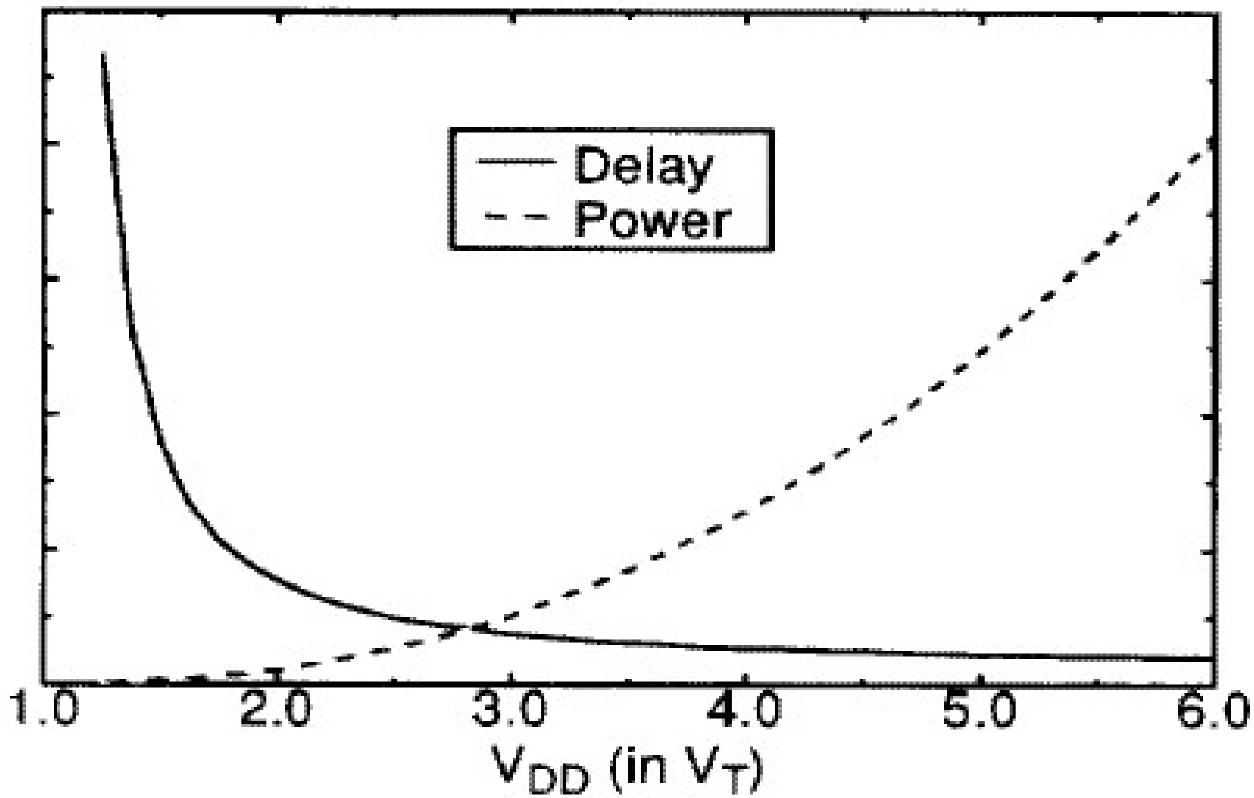
Assuming clock cycle time (T) is proportional to circuit delay (t_d) :

$$T \propto t_d \Rightarrow T = m' \frac{V_{dd}}{(V_{dd} - V_{th})^{\alpha}}, \qquad 1 < \alpha \leq 2$$

Approximately : $T \propto V_{dd}$

For frequency (1/T) : $f_{max} \propto 1/V_{dd}$

Effects of Voltage Change



Energy (Power for constant frequency), Delay Vs Voltage

How much we save ?? Varying V_{dd} / f

Voltage (V_{dd})	Frequency (f)	Power (P_d)	Energy (E_d)
V_{dd}	f_{max}	P_d	E_d
$V_{dd} / 2$	f_{max} *	$P_d / 4$	$E_d / 4$
$V_{dd} / 2$	$f_{max} / 2$	$P_d / 8$	$E_d / 4$
V_{dd}	$f_{max} / 2$	$P_d / 2$	E_d

* Note : f_{max} Vs f

University of South Florida

IEEE Computer magazine: April 2000, page-56

"It is not necessary for a processor to run constantly at maximum frequency to accomplish its work. If we know a computation's deadline, we can adjust the processor's frequency and reduce the supply voltage. For example, a simple MPEG decode runs at a fixed rate determined by every 1/30th of a second. Therefore, we can adjust the processor to run so that it doesn't finish its work ahead of schedule and waste power."

Varying Frequency with Voltage: A Good Idea

T.D. Burd and R.W Brodersen, "Energy efficient CMOS microprocessor design", *Proceedings of the 28th Hawaii International Conference on System Sciences*, 1995, pp. 288 –297.

"Scaling the clock frequency is a third approach which is most beneficial if it coupled with voltage scaling. If the clock frequency is reduced, the delay may be increased (keeping is equal to $1/f_{CLK}$) by reducing the supply voltage and thus saving power. If the voltage is kept constant, then power and throughput reduce linearly with clock frequency."

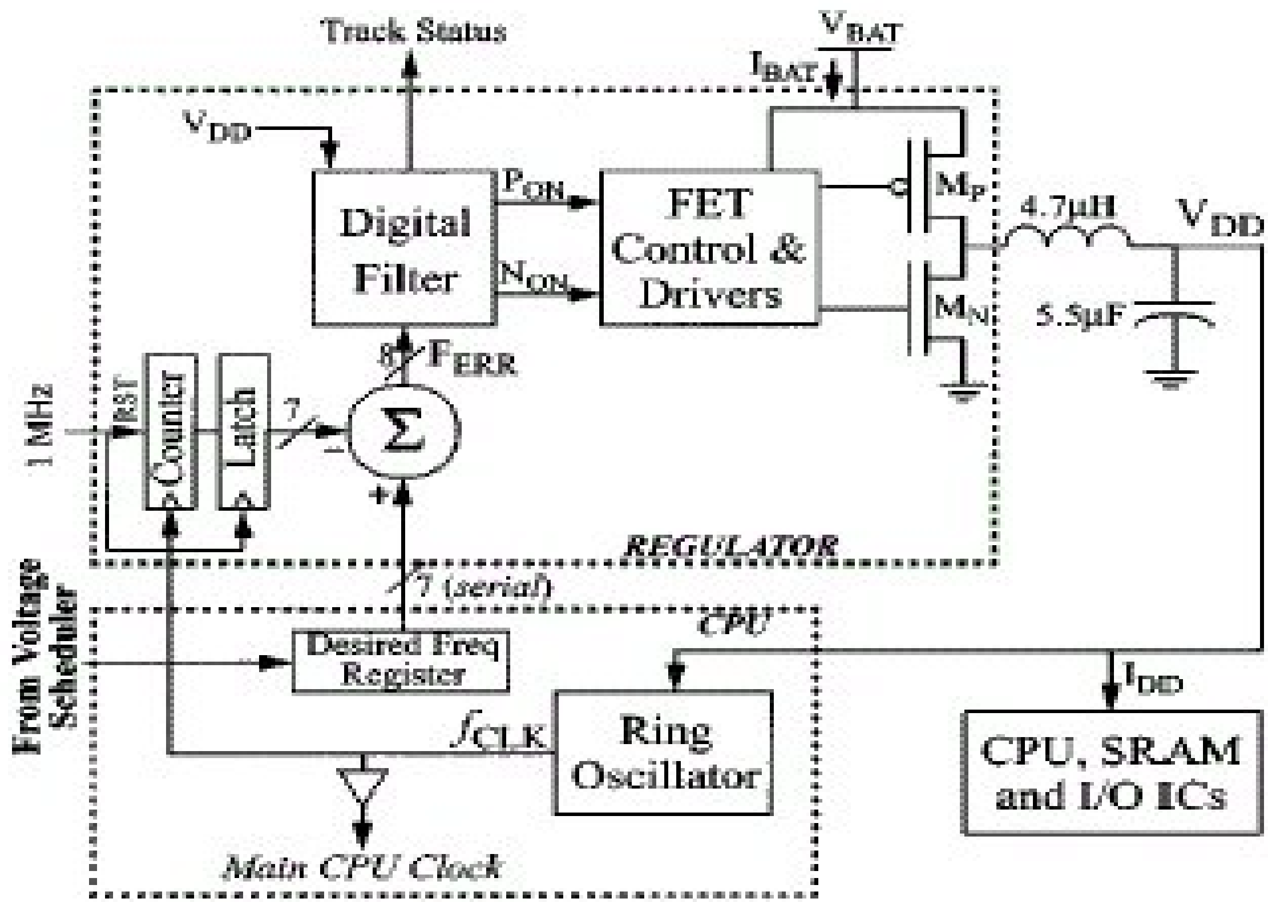
A dynamic voltage scaled microprocessor system

T.D. Burd, T.A.Pering, A.J. Stratakos, R.W.Brodersen, “A Dynamic Voltage Scaled Microprocessor System”, *IEEE Journal of Solid State Circuits*, Volume: 35 Issue: 11, Nov. 2000, Page(s): 1571 –1580.

- Supply Voltage and clock frequency are dynamically varied : system delivers high throughput when required while saves energy(extends battery life) during low speed periods
- The microprocessor operates from 1.2-3.8V and 5-80MHz with 0.54mW/MIPS minimum energy consumption
- Operating System need to have a voltage scheduler that controls f_{CLK} and V_{dd} by writing a desired frequency (MHz) to a coprocessor register

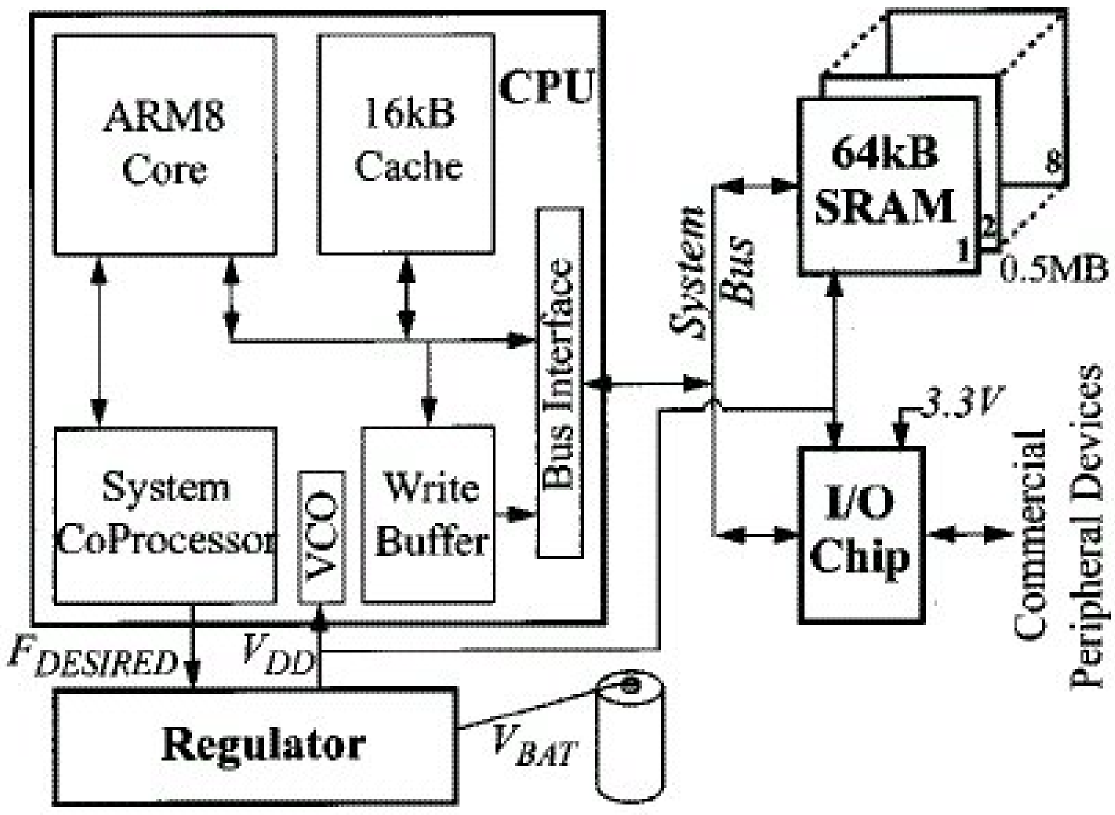
University of South Florida

dynamic voltage scaled system ……



Frequency to Voltage Feedback Loop

dynamic voltage scaled system ……



System Architecture: Four Custom Chips

University of South Florida

dynamic voltage scaled system ……

Conventional Benchmarks like SPEC, MIPS etc. are not useful since they are constructed to measure the peak throughput of the processor. New benchmarks selected which combine computational requirements with realistic latency constraints.

- AUDIO Decryption
- MPEG Decoding
- User Interfaces
- Java Interpreter
- Web Browser
- Graphics Primitive Rendering

Speed-Setting of a Low-Power CPU

K.Govil, E.Chan, H. Wasserman, "Comparing Algorithms for Dynamic Speed-Setting of a Low-Power CPU", *Proc. Of the 1st Annual Intl. Conf. On Mobile Computing and Networking*, *1995*.

A CPU is regarded as a capacitor-based system that satisfies the physical laws:

$$\text{energy/sec} \; \alpha \; \text{voltage}^2 * \text{speed}$$

or $$\text{energy/task} \; \alpha \; \text{voltage}^2$$

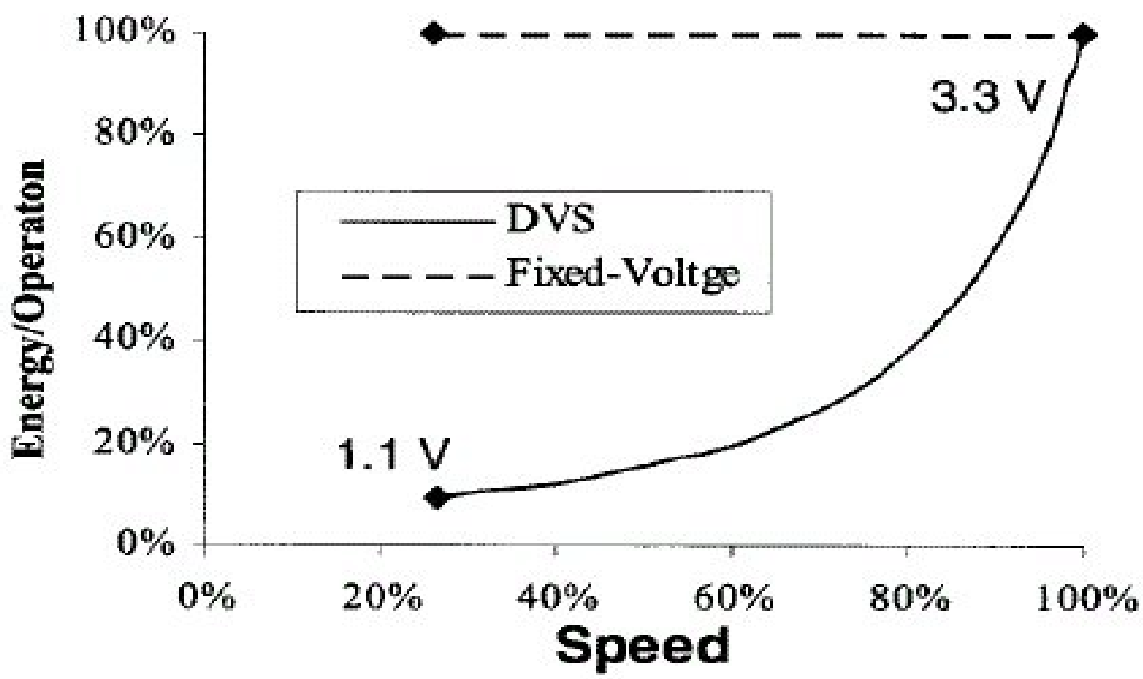
If voltage is reduced directly proportional speed, then

$$\text{energy/task} \; \alpha \; \text{speed}^2$$

Conclusion: CPU capable of dynamic speed setting can save energy is voltage is reduced accordingly

Energy-Speed trade-off

T.A.Pering ,T.D. Burd,, R.W.Brodersen, “Voltage Scheduling in the lpARM Microprocessor System”, *Proceedings of the International Symposium on Low Power Electronics and Design*, 2000, ISLPED’00, pp. 96-10



Energy/Speed Trade-Off

Multiple Operating Voltage Energy Reduction: (MOVER)

M.C.Johnson and K.Roy, "Datapath Scheduling with Multiple Supply Voltages and Level Converters", *Proc. of IEEE Intl. Symposium on Circuits and Systems*, 1997, ISCAS '97, Vol. 3 , Hong Kong, pp. 2152 –2155.

Energy Savings:

two supply voltage : 0-50%
three supply voltage : 0-65%

Area Penalty: 0-170%

Core of MOVER: ILP

MOVER (How it works ??).........

- MOVER find the minimum and maximum bound on the time window in which operation must execute

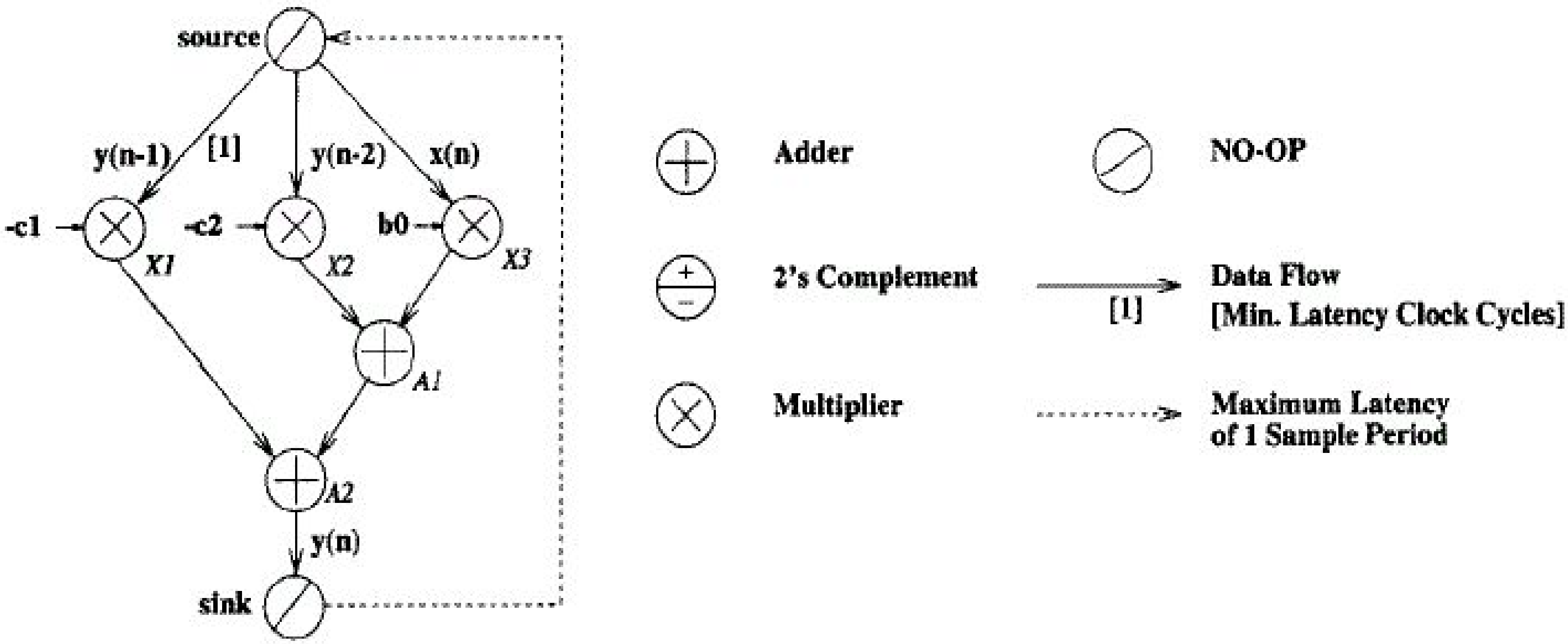
- The DFG is partitioned into two groups (2 supply):
  - Higher voltage operation groups
  - Lower voltage operation groups
  - MOVER 1st fixes the minimum voltage of the lower group and then fixes the minimum voltage for the upper group

- The decision variable $x_{i,l,s} = 1$ indicates that operation *i* begins on clock cycle *l* using supply voltage *s*.

- Objective Functions : Energy and Area

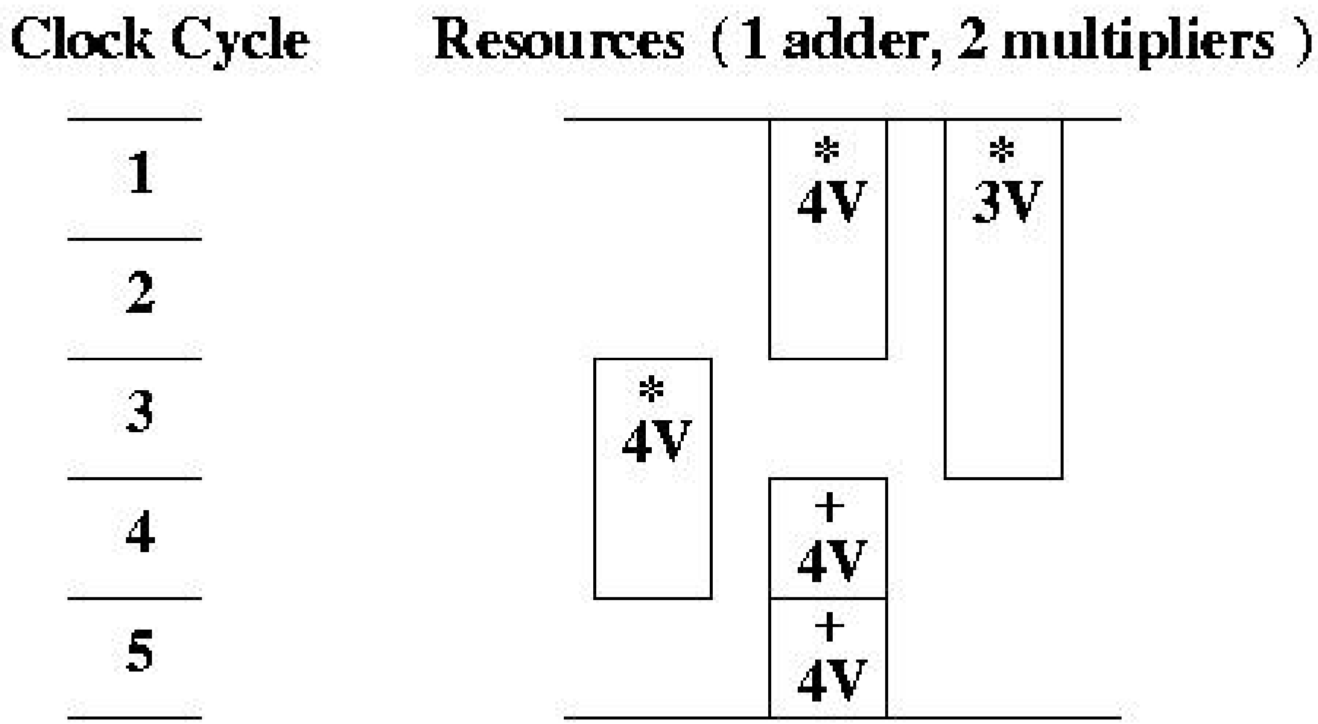
University of South Florida

MOVER ....



Datapath Specification and key notations

MOVER ....



Schedule with variable voltage

Drawbacks of MOVER

- Do not address conditional branches
- Do not consider functional pipelining
- Energy model used is data-intensive
- Exponential worst-case complexity and can't handle large datapaths
- ILP (Integer Linear Programming) is difficult to solve

University of South Florida

Variable Voltage: Effects on Circuit Design

- If Multiple supplies are generated off-chip then additional power and ground pins will be required.
- It may be necessary to partition the chip into separate regions, where all modules in a region operate at the same voltage.
- Some kind of isolation will be required between regions operated at different voltages.
- There may be some limit on the voltage difference that can be tolerated between the regions.
- Protection against latch-up may be needed at the logic interfaces between regions if different voltages.
- New design rules for routing may be needed to deal with signals at one voltage passing through a region at another voltage.

Conclusions

1. Sources of power dissipation
2. Dynamic power dissipation details
3. How to reduce dynamic power ?
4. Effect of Frequency on Energy / Power
5. Few low-power research works