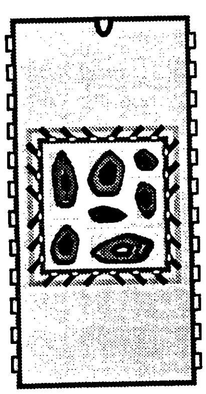
Cold Chip Design Techniques

Anantha Chandrakasan - GRAD STUDENT
Samuel Sheng
Robert Brodersen

Dept. of EECS,
U.C. Berkeley

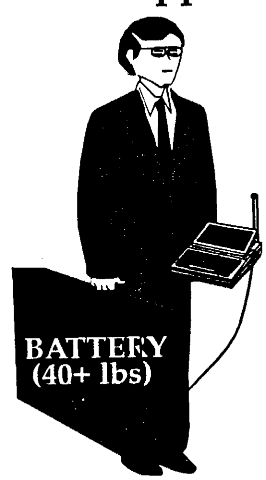
Why "Cold" Chips?

Heat Dissipation in "Hot" Chips



- DEC Alpha: 30W @ 3.3V, 200Mhz
- SUN Viking: 8W @ 5V, 50Mhz

Portable Applications



- Multimedia Terminals
- Laptop Computers
- Digital Cellular Telephony

General Purpose vs. Application Specific Computing

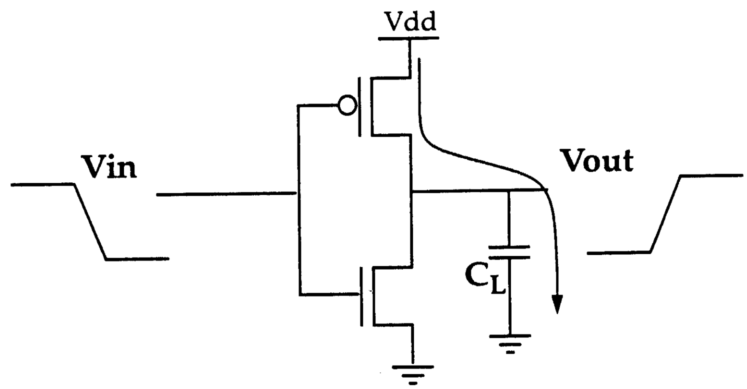
- General Purpose Computing (e.g. Laptop computers)
  - Desire maximum achievable throughput.
  - Power down techniques.
  - Reduction of supply voltage from 5V to 3.3V.
- Application Specific Computing (e.g. Video Compression)
  - Power down techniques not as appropriate.
  - No advantage in making the computation faster once throughput requirements are met.

Many new and exciting applications are application specific in nature.

Ways to Reduce Power

- Technology.
- Logic Styles.
- Architectures.
- Algorithms.

Where Does Power Go?

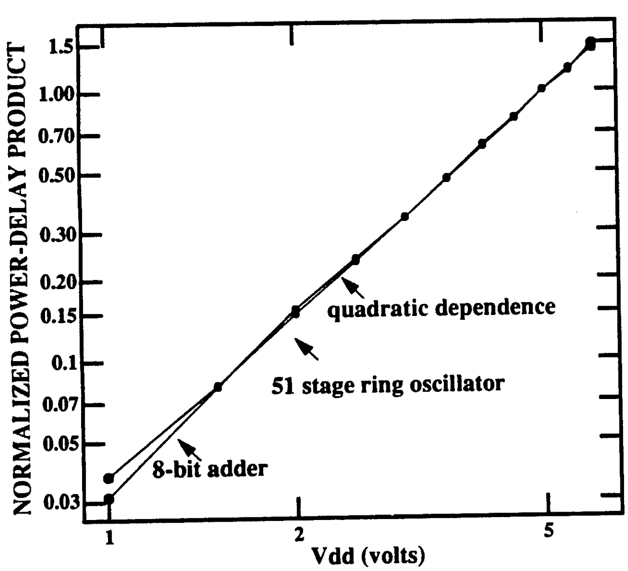


$$\text{Energy/transition} = C_L * V_{dd}^2$$

$$\text{Power} = \text{Energy/transition} * f = C_L * V_{dd}^2 * f$$

- Need to reduce C_L, V_{dd}, and f to reduce power.

Normalized Energy



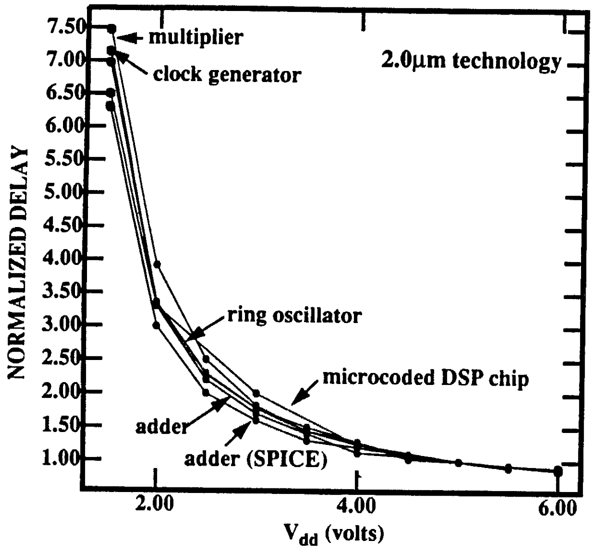
$$E_t = C_L * V_{dd}^2$$

$$\frac{E_{(Vdd=2)}}{E_{(Vdd=5)}} = \frac{(C_L) * (2)^2}{(C_L) * (5)^2}$$

$$E_{(Vdd=2)} \approx 0.16\ E_{(Vdd=5)}$$

- Strong function of voltage (V^2 dependence).
- Relatively independent of logic function and style.

Normalized Delay



$$T_d = \frac{C_L * V_{dd}}{I}$$

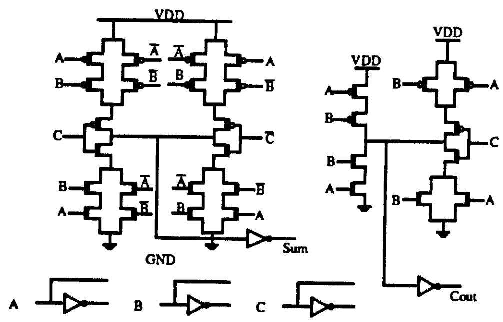
$$I \sim (V_{dd} - V_t)^2$$

$$\frac{T_{d(Vdd=2)}}{T_{d(Vdd=5)}} = \frac{(2) * (5 - 0.7)^2}{(5) * (2 - 0.7)^2}$$

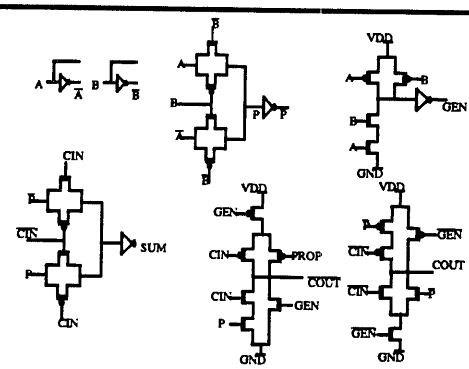
$$\approx 4$$

- Lowering V_{dd} reduces energy but increases delays.
- Relatively independent of logic function and style.

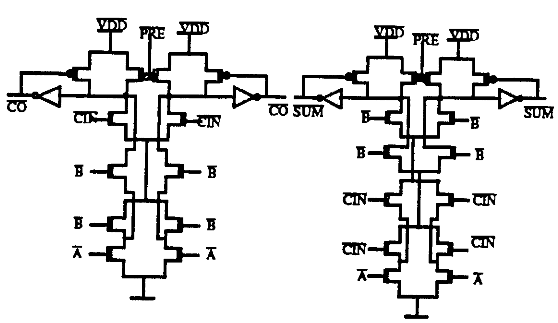
How About Logic Styles?



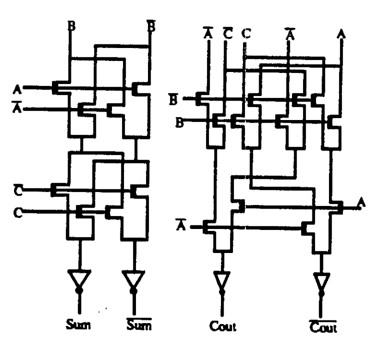
CONVENTIONAL CMOS Adder



OPTIMIZED static Adder

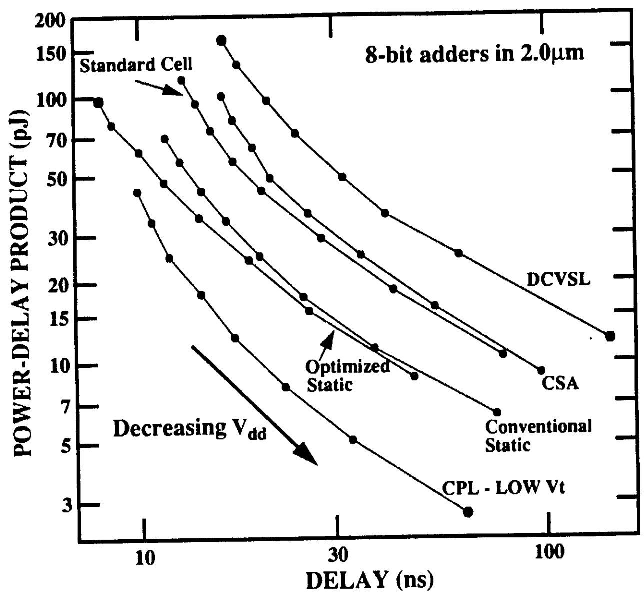


DCVSL Adder



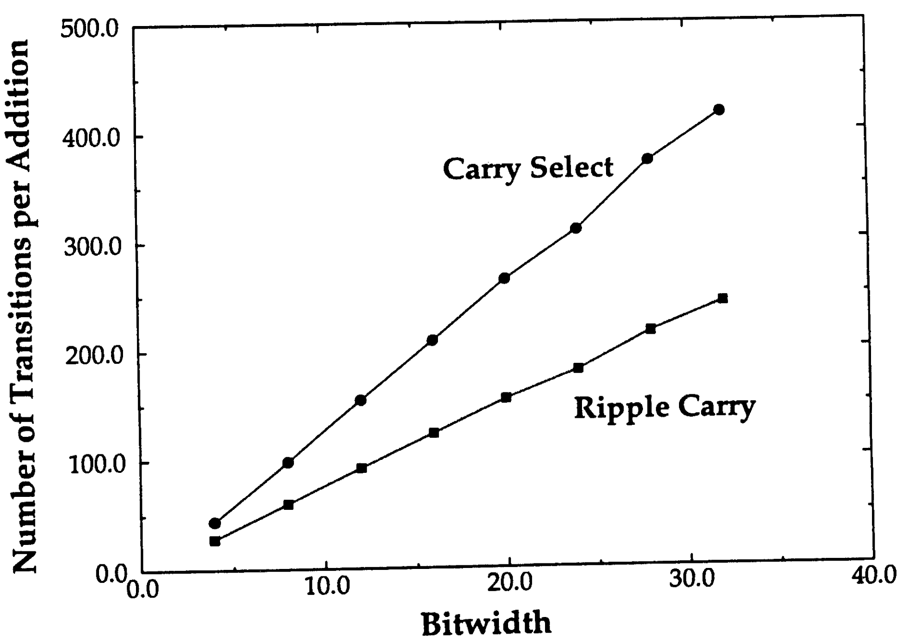
CPL Adder

What Logic Style to Use?

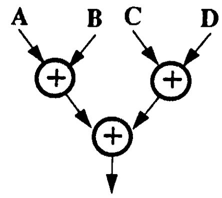


- Power-delay product improves as voltage decreases.
- The "best" logic style minimizes power-delay for a given delay constraint.

Optimizing Logic Structures for Reduced Activity

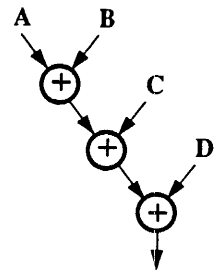


Mutiple Signal Transitions



(A + B) + (C + D)

Tree vs. Chain



(A + B) + C + D

Inputs	Normalized # of Transitions Tree	Normalized # of Transitions Chain
4	1	1.45
8	1	2.5

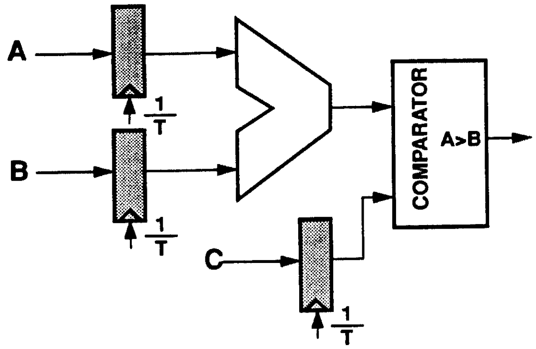
- Can be reduced by reducing the logic depth and balancing signal paths.

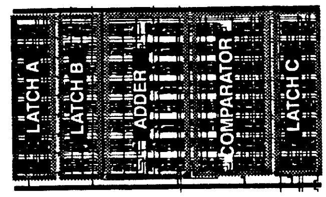
Architecture

- Desire to operate at lowest possible speeds (using low supply voltages).
- Use Architecture optimization to compensate for slower operation.

Approach: Trade-off AREA for lower POWER

Architecture Trade-offs - Reference Datapath

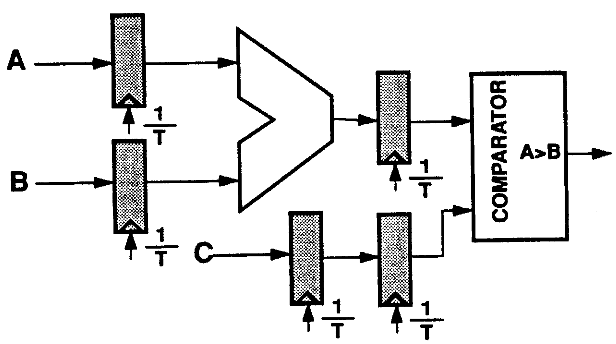


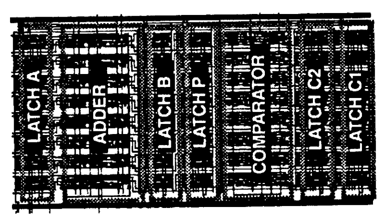


Area = 636 x 833 μ^2

- Critical path delay ⇒ $T_{adder} + T_{comparator}$ (= 25ns)
  ⇒ f_{ref} = 40Mhz
- Total capacitance being switched = C_{ref}
- $V_{dd} = V_{ref}$ = 5V
- Power for reference datapath = $P_{ref} = C_{ref} V_{ref}^2 f_{ref}$

Pipelined Datapath

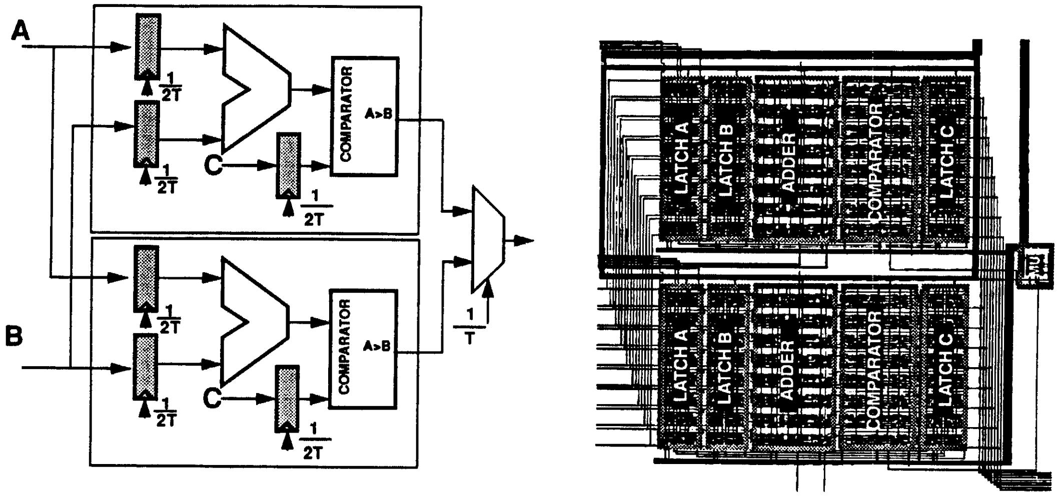




Area = 640 x 1081 μ^2

- Critical path delay is less ⇒ max [T_{adder} , $T_{comparator}$]
- Keeping clock rate constant: $f_{pipe} = f_{ref}$
  Voltage can be dropped ⇒ $V_{pipe} = V_{ref}$ / 1.7
- Capacitance slightly higher: $C_{pipe} = 1.15C_{ref}$
- $P_{pipe} = (1.15C_{ref})\,(V_{ref}/1.7)^2\, f_{ref} \approx 0.39\, P_{ref}$

Parallel Datapath



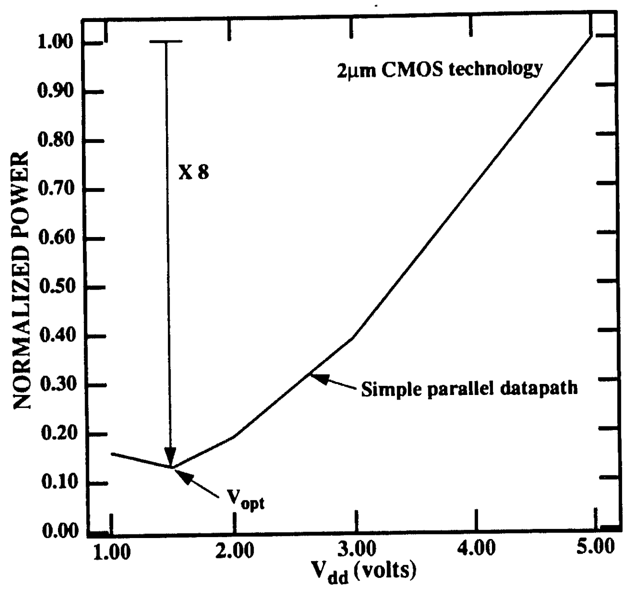
Area = 1476 x 1219 μ^2

- The clock rate can be reduced by half with the same throughput ⇒ $f_{par} = f_{ref} / 2$
- $V_{par} = V_{ref} / 1.7$, $C_{par} = 2.15C_{ref}$
- $P_{par} = (2.15C_{ref})\,(V_{ref}/1.7)^2\,(f_{ref}/2) \approx 0.36\,P_{ref}$

Architecture Summary for a Simple Datapath

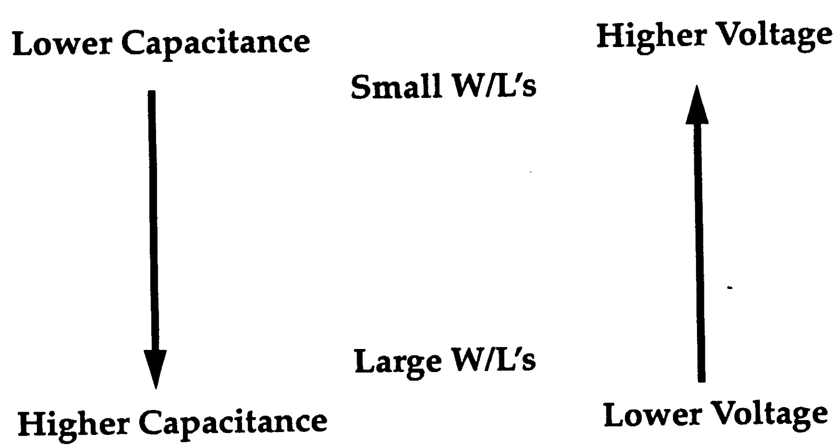
Architecture type	Voltage	Area	Power
Simple datapath (no pipelining or parallelism)	5V	1	1
Pipelined datapath	2.9V	1.3	0.39
Parallel datapath	2.9V	3.4	0.36
Pipeline-Parallel	2.0V	3.7	0.2

How Low a Voltage Can We Use?



- Too much parallelism adds overhead circuitry. Overhead circuitry can dominate with high levels of parallelism.
- An architecture driven optimum operating voltage can be found.

Transistor Sizing Strategy for Minimizing Power



- Larger sized devices are useful only when interconnect dominated.
- Minimum sized devices are usually optimal for low-power.

Conclusions

- POWER is the most critical consideration in portable systems.
- AREA (transistors) can be traded for POWER.

SLOWER IS BETTER - The best design will used the slowest circuits to meet the computation requirements.