Architectural Power Optimization by Bus Splitting

Cheng-Ta Hsieh
Verplex System

Massoud Pedram
Dept. of EE-systems
University of Southern California

Outline

- Prior Work
- Shared Bus v.s. Point-to-Point Connection
- Power Model and Examples
- Algorithm
- Summary

Prior Work

- Low Voltage Swing [Nakagome, ‘93]
- Bus Encoding
  - Bus-Invert Code for Data Bus [Stan, ‘95]
    - Uses INV line to dynamically signal the receiver that the transmitted data is inverted
  - Gray Code for Address Bus [Su,’94]
    - Arranges the program in gray code order
  - T0 Code for Address Bus [Benini,’97]
    - Uses redundant line INC to auto-increment address for consecutive accesses

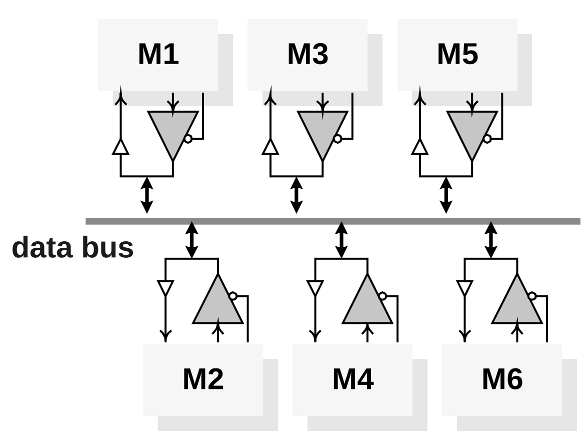
Shared-Bus Architecture

- Advantage
  - Simple topology
  - Lower area cost
  - Extensibility
- Disadvantage
  - Larger load per data-bus line
  - Longer delay for data transfer
  - Larger power consumption
  - Lower bandwidth

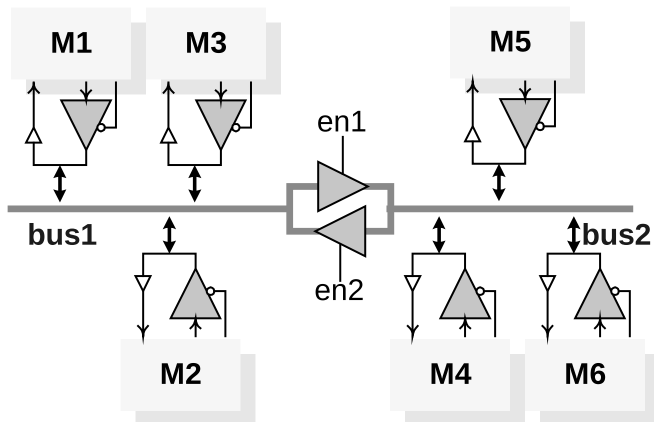
Point-to-Point Connection

- Pros
  - Smaller load per bit
  - Speed and power
- Cons
  - High area cost
  - Routability

Monolithic-Bus Architecture



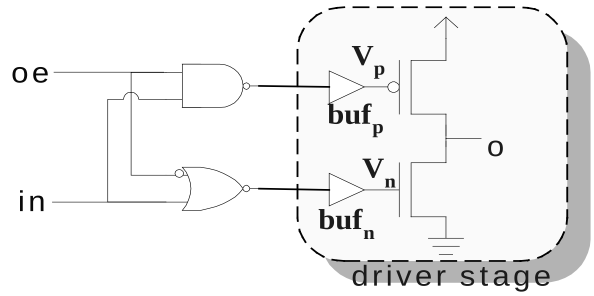
Split-Bus Architecture



Benefits of Split Bus Architecture

- Smaller parasitic load
- Larger timing slack
- Smaller driver size
- Lower power consumption
- Lower noise problems

Probabilistic Power Model of Driver



$$sw(V_p) = 2\,prob(in)\,prob(oe)[1 - prob(in)\,prob(oe)]$$

$$sw(V_n) = 2\,prob(in=0)\,prob(oe)[1 - prob(in=0)\,prob(oe)]$$

$$E_{driver} = 0.5(sw(V_p)C_{eff,bufp} + sw(V_n)C_{eff,bufn})V_{dd}^2$$

Accurate Power Model of Split-Bus

- $V_{BUS1,i}$: logic value on bus1 at clock cycle *i*
- at clock cycle *i*, $M_{SRC(i)}$ *sends data to* $M_{DST(i)}$

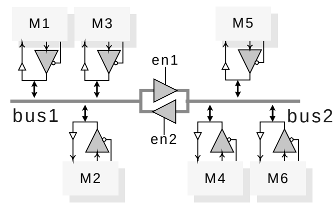
$V_{BUS1,i}$	$= V_{BUS1,i-1}$	*if* $M_{SRC(i)} \notin BUS1$ *and* $M_{DST(i)} \notin BUS1$
	$= V_i$	*otherwise*
$V_{BUS2,i}$	$= V_{BUS2,i-1}$	*if* $M_{SRC(i)} \notin BUS2$ *and* $M_{DST(i)} \notin BUS2$
	$= V_i$	*otherwise*

$$E = E_{BUS1} + E_{BUS2} + E_{driver}$$

$$= 0.5C_{BUS1}\sum_{i=1}^{p}(V_{BUS1,i-1} \oplus V_{BUS1,i})\frac{V_{dd}^2}{p} + 0.5C_{BUS2}\sum_{i=1}^{p}(V_{BUS2,i-1} \oplus V_{BUS2,i})\frac{V_{dd}^2}{p}$$

$$+\sum_{i=1}^{n} E_{driver,Mi} + E_{driver,BUF1} + E_{driver,BUF2}$$

Probabilistic Power Model of Bus



- xfer(BUS1,BUS2)
  - probability of bus1 transferring data to bus2
- prob(BUS1)
  - the probability for the bus1 having a logic value '1' in a clock cycle

$$sw(BUS1) = 2\,prob(BUS1)[1 - prob(BUS1)]xfer(BUS1)$$

$$sw(BUS2) = 2\,prob(BUS2)[1 - prob(BUS2)]xfer(BUS2)$$

$$E = 0.5(C_{BUS1}sw(BUS1) + C_{BUS2}sw(BUS2))V_{dd}^2$$

$$+\sum_{i=1}^{n} E_{driver,Mi} + E_{driver,BUF1} + E_{driver,BUF2}$$

Simple Power Model for Examples

- Monolithic Bus

$$E1 = 0.5 \cdot sw \cdot C_{BUS} V_{dd}^2$$

- Split Bus (assume all data has same *sw*)

$$E2 = 0.5V_{dd}^2[sw \cdot C_{BUS1} \overbrace{\sum_{i\in BUS1}\sum_{j\in BUS1, i\neq j} xfer(M_i, M_j)}^{\text{bus1}} + C_{BUS2} \overbrace{\sum_{i\in BUS2}\sum_{j\in BUS2, i\neq j} xfer(M_i, M_j)}^{\text{bus2}}$$

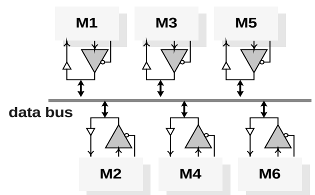
$$+(C_{BUS1}+C_{BUS2}) \underbrace{\sum_{i\in BUS1}\sum_{j\in BUS2} xfer(M_i, M_j)}_{\text{bus1->bus2}} + (C_{BUS1}+C_{BUS2}) \underbrace{\sum_{i\in BUS2}\sum_{j\in BUS1} xfer(M_i, M_j)}_{\text{bus2->bus1}}]$$

- Voltage and Capacitive Load

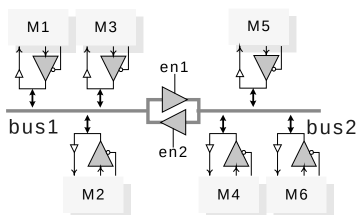
$$V_{dd} = \frac{C_{BUS1}}{|BUS1|} = \frac{C_{BUS2}}{|BUS2|} = \frac{C_{BUS}}{n} = 1$$

Example 1

- 6 modules, xfer(M_i,M_j)=1/6(6-1)=1/30



$$E1 = 0.5 \cdot (6 \cdot 0.5) = 1.5$$



$$E2 = 0.5(\overbrace{\frac{1}{5}\times 3\times 0.5}^{\text{bus1}} + \overbrace{\frac{1}{5}\times 3\times 0.5}^{\text{bus2}} + \underbrace{\frac{3}{5}\times 6\times 0.5}_{\text{inter-bus}}) = 1.2$$

Example 2

- Homogeneous transfer probabilities
  - 2*k* modules
  - |BUS1|=*k*-*a*, |BUS2|=*k*+*a*
  - xfer(M_i,M_j)=1/2*k*(2*k*-1)
- Results

$$E1 = 0.5 \cdot (2k \cdot 0.5) = 0.5k$$

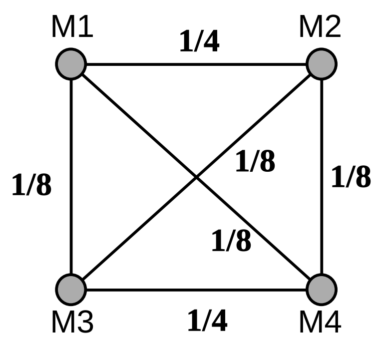
$$E2 = 0.25\frac{3k^3 - k^2 + a^2(k-1)}{2k^2 - k}$$

$$\frac{E1 - E2}{E1} = 0.5\frac{k^3 - k^2 - a^2(k-1)}{2k^3 - k^2}$$

k=2,*a*=0, power saving =16%

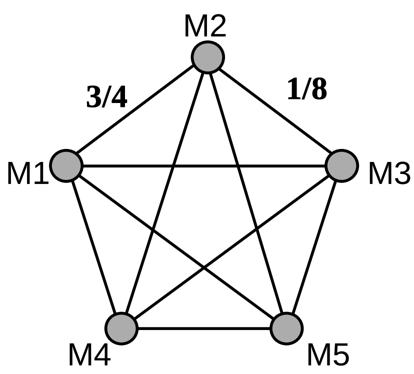
k→∞, *a*=0, power saving = 25%

Example 3



Architecture	Energy
BUS={M1,M2,M3,M4}	1
BUS1={M1,M2} BUS2={M3,M4}	0.75
BUS1={M1,M3} BUS2={M2,M4}	0.875
BUS1={M1,M4} BUS2={M2,M3}	0.875

Example 4



Architecture	Energy
BUS={M1,M2,M3,M4,M5}	1.25
BUS1={M1,M2} BUS2={M3,M4,M5}	0.66
BUS1={M1,M2,M3} BUS2={M4,M5}	0.79
BUS1={M2,M3} BUS2={M1,M4,M5}	1.13

un-labeled probabilities = 1 / 64

Bus-Splitting with Fix Module Order

- Calculate the switching activities of the data on bus1 and bus2 for each possible buffer positions at segment *i*, *i=2…n-2*.
- Calculate energy consumption *E(i)* for buffer position at segment *i*, *i=2…n-2.*
- Find the minimum *E(i)*
- *Complexity: O(num_of_clock_cycles x n)*

Bus-Splitting with Arbitrary Module Order

- NP-hard Problem
  - The problem is equivalent to 'minimum cut into two bounded sets' by converting $xfer(m_i,m_j)$ to weights in the graph
- Solutions Space
  - Number of feasible splitting: $2^{n-1}-1$
- Probabilistic power model can be used to speed up the searching
  - In practice, *n* <= 30 can be solved efficiently

Exact Algorithm

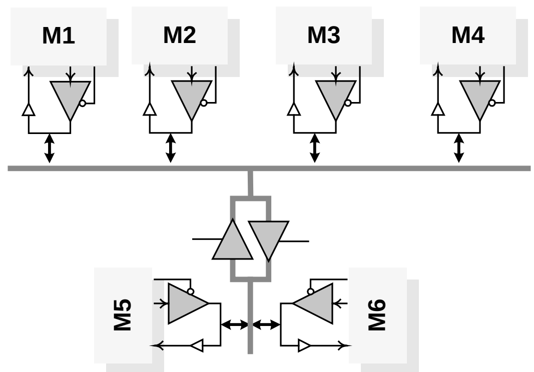
- For each possible order, find optimal buffer position with fix-module-order algorithm
- If all the data transition is uncorrelated, the previous algorithm can be sped up to $O(n^2)$ by only calculating the energy difference in adjacent buffer positions
- If n is too large, heuristic algorithm can be applied to cluster modules first

Heuristic Algorithm

- Clustering
  - Maximize the intra-transfer probability in each cluster
  - Avoid big cluster size (example 4)
  - Recursive max-weight matching is used.
- After Clustering, the effective number of modules is reduced. Then the exact algorithm is applied.

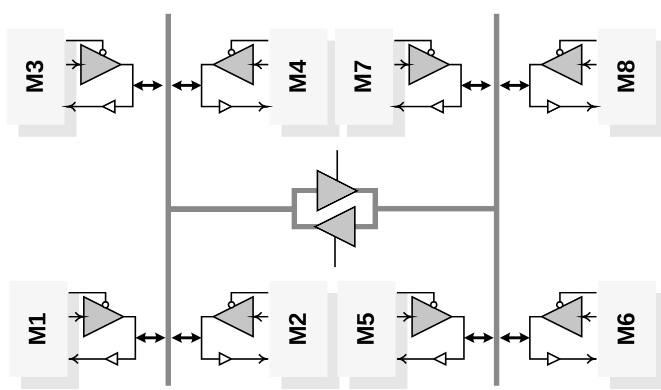
Topology Variation

- T-Shaped Bus

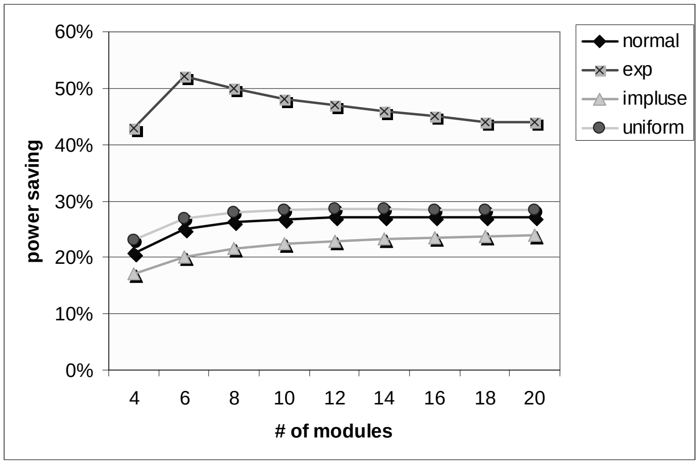


Topology Variation (cont'd)

- H-Shaped Bus



Experimental Results



- Each point shows the average power saving of the split bus over the monolithic bus for 500 randomly generated test cases under different distributions.

Summary

- Split bus can improve timing and power dissipation for on-chip data exchange
- Split bus can save up to 50% power
- T-shaped and H-shaped can further improve the bus performance
- Multi-way splitting can be used if a large number of modules are connected to a shared bus