Low Power Address Bus Encoding Techniques

Massoud Pedram
Dept. of EE-Systems
University of Southern California
Los Angeles CA

CAD Seminar, UCLA CS Dept.
May 21, 2001

Outline

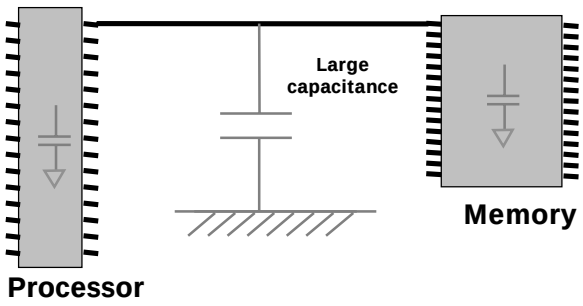
- Introduction
- Low Power Encoding techniques
  - SRAM Address Bus
  - DRAM Address Bus
- Summary

Introduction

- Average power reduction of a battery-powered embedded system while meeting key minimum performance and quality-of-service criteria is an important design driver
- Much of the power savings is at the system-level and can be achieved thru dynamic power management (DPM) and power-aware architecture organization
- Our research focuses on these two approaches to system-level power reduction and is expected to deliver 2-4 X power savings

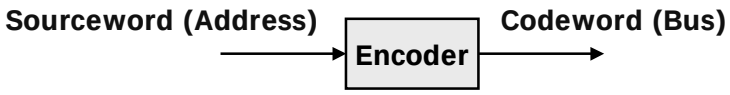
Why Focus on the Memory Bus

The processor-memory bus is highly capacitive. Significant power is consumed to drive these busses



Memory Bus Power Minimization

- Physical Layer
  - Conductor types
  - Signaling
  - Voltage levels
- Data Link Layer
  - Encoding
  - Packetization



Encoding Techniques

- Functional Classification
  - Algebraic
  - Permutation
  - Probabilistic
- Physical Classification
  - Redundant
  - Irredundant
- Behavioral Classification
  - Static
  - Dynamic

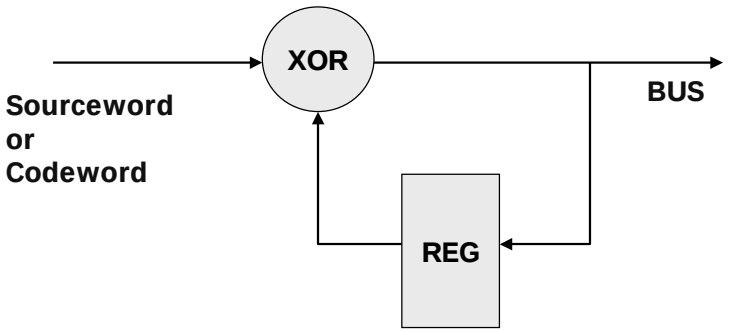
Background

- Limited Weight Codes
- Transition signaling
- Codebook-based codes

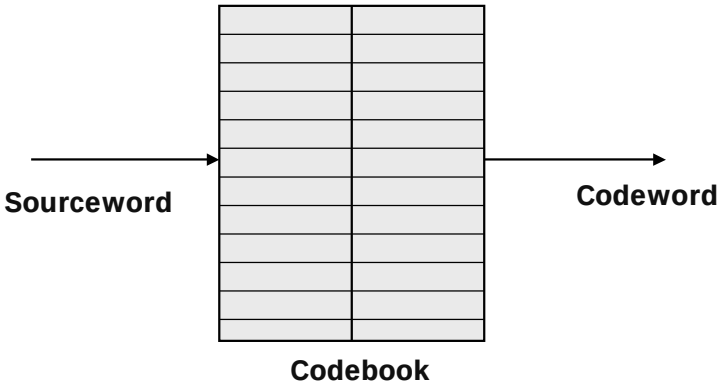
Limited Weight Codes

- One-Limited Weight Code:
0000100, 1000000...0
- Two-Limited Weight Code:
01010, 11000, 10...0001
- N-Limited Weight Code:
Exactly N **1**s, the rest **0**s

Transition Signaling



Codebook



Inspiration

- Find a code with minimum number of ones in the codewords
- Use redundant bits if needed
- Exploit address locality

Minimum Hamming Distance

- 2^N symbols, equal probability
- For a Hamming distance of one, at least 2^N-1 bits are needed
- Easily implemented when using **transition signaling** and **Limited Weight Codes**

000…00	**symbol #1**
000…01	**symbol #2**
000…10	…
…	…
010…00	
100…00	**symbol #2^N**

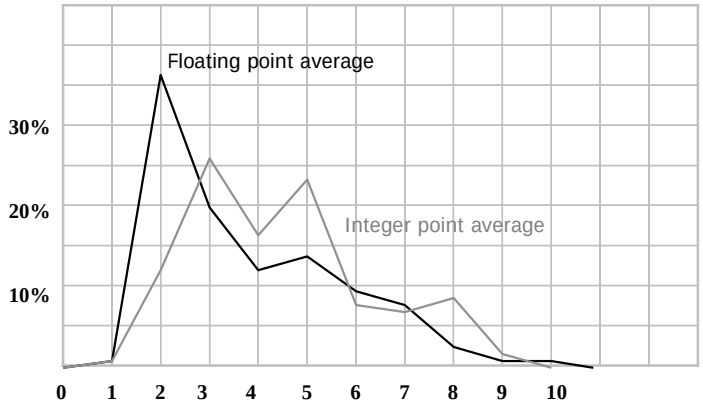
LWC + Transition Signaling

Symbol	Codeword	Bus (0000000)
#2	0000001 → XOR	0000001
#4	0000100	0000101
#5	0001000	0001101
#4	0000100	0001001
#1	0000000	0001001
#6	0010000	0011001
#8	1000000	1011001

Disadvantages

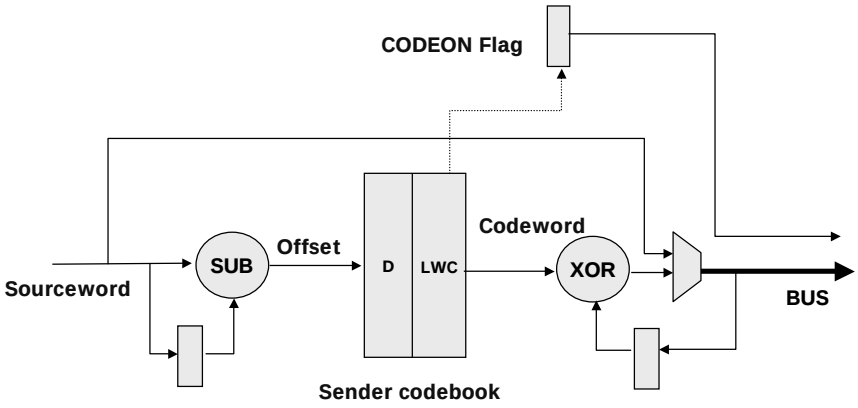
- Number of redundant bits too large
- Address locality not exploited

Common Branch Displacements

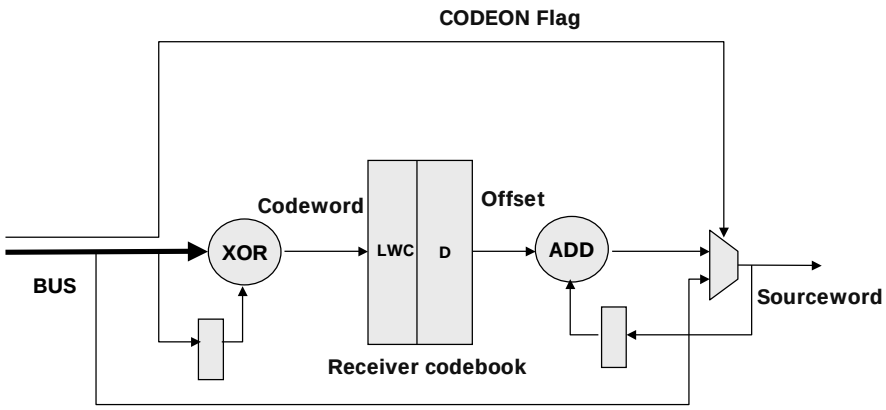


Computer Architecture, A Quantitative Approach
Hennessy and Patterson

ALBORZ Encoder



ALBORZ Decoder



Apollo

ALBORZ Codebook

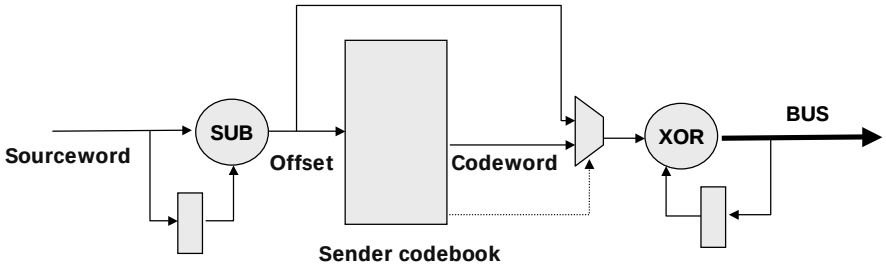
- Fixed codebook
  - Codebook is a read-only memory
  - Mapping between offsets and LWCs is fixed
- Adaptive Codebook
  - Codebook is a read/write memory
  - Offsets are dynamically replaced in the codebook
  - Same update policy is used in the encoder and the decoder

Apollo

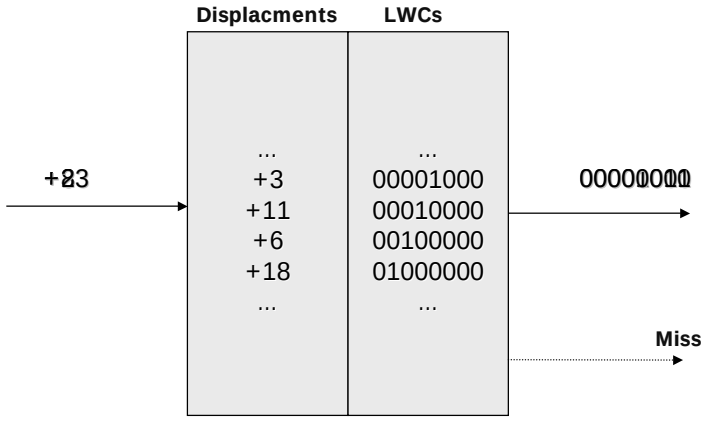
ALBORZ Code

Address	Displacement	Codebook Output	BUS
39	?		0 00111001
40	+1	00000000	1 00111001
48	+8	00010000	1 00101001
49	+1	00000000	1 00101001
50	+1	00000000	1 00101001
23	-27	01000100	1 01101101
24	+1	00000000	1 01101101
196	+172		0 11000100
197	+1	00000000	1 11000100

Irredundant ALBORZ



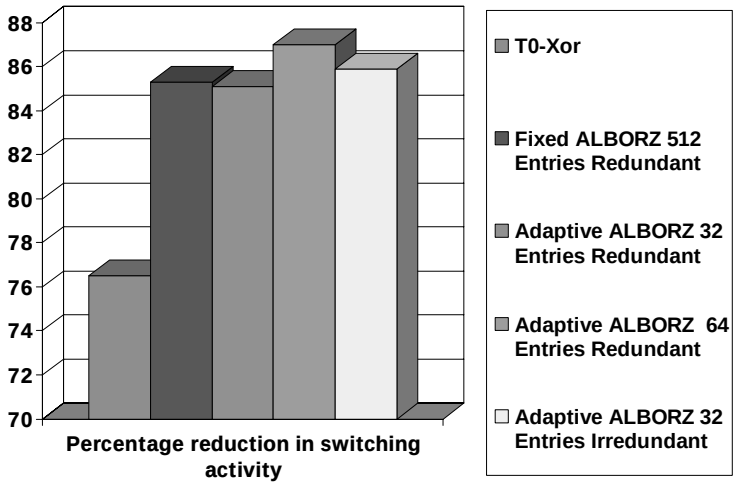
Irredundant Codebook



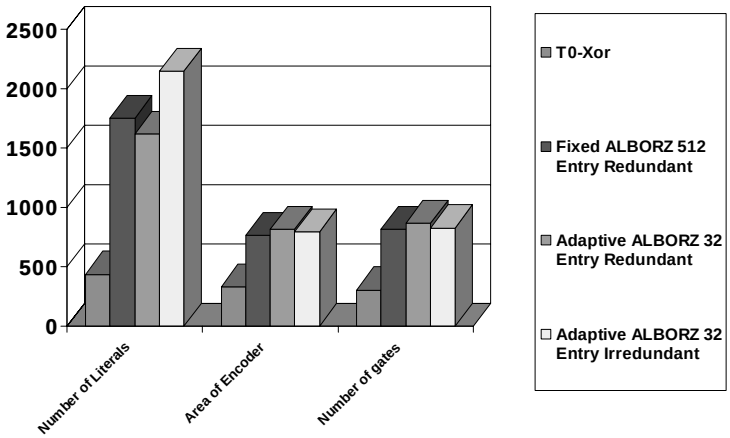
Experimental Results

Benchmark	Base Case Pwr	Fixed ALBORZ 512 Entry Redundant Pwr	Adaptive ALBORZ Redun.		Adaptive ALBORZ 32 Entry Irredun. Pwr
			32 Entry	64 Entry	
Art	34.3	2.765	2.466	2.110	2.382
		92.0%	92.9%	93.9%	93.1%
Gzip	34.7	4.005	3.103	2.651	3.483
		88.5%	91.1%	92.4%	90.0%
Vortex	37.3	6.048	6.244	5.554	4.918
		85.9%	83.3%	85.2%	86.9%
Equak	35.4	6.383	7.673	6.472	6.576
		82.1%	78.4%	87.8%	81.5%
Gcc	35.4	6.373	7.651	6.881	7.490
		82.1%	78.5%	80.6%	78.9%
Parser	37.0	4.177	4.273	3.614	4.560
		88.8%	88.5%	90.3%	87.3%
Vpr	35.9	7.021	6.085	5.502	5.928
		80.5%	83.1%	84.7%	83.5%
% power reduction		85.3%	85.1%	87.0%	85.9%

Summary Results

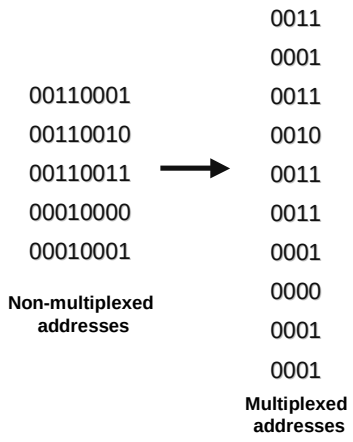


Hardware Implementation



DRAM Address Bus

- The DRAM address bus is multiplexed between the row and column addresses
  - Conventional
  - Page Mode
  - Burst Mode



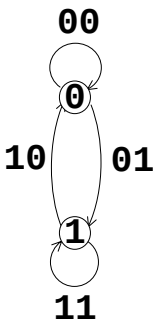
Binary versus Pyramid Code

Binary Code

SRAM	DRAM
00	0 0
01	0 1
10	1 0
11	1 1
SA=6	SA=4

Pyramid Code

SRAM	DRAM
00	0 0
01	0 1
11	1 1
10	1 0
SA=4	SA=2

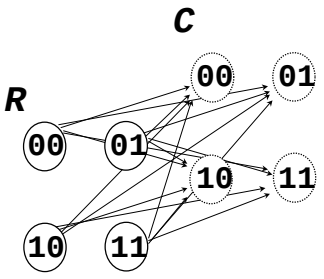


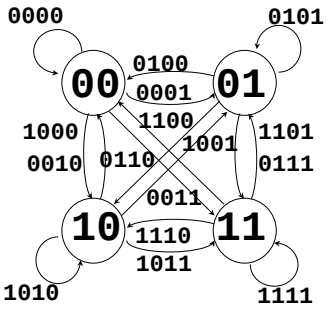
First-Order Comparison

Binary	Gray	Bus-Inv	T0	Gray DRAM	Pyramid
0000	0000	0000-0	0000-0	00 00	00 00
0001	0001	0001-0	0000-1	00 01	00 01
0010	0011	0010-0	0000-1	00 10	01 01
0011	0010	0011-0	0000-1	00 11	01 00
0100	0110	1011-1	0000-1	01 00	00 10
0101	0111	1010-1	0000-1	01 01	10 01
0110	0101	0110-0	0000-1	01 10	01 10
0111	0100	0111-0	0000-1	01 11	10 10
1000	1100	0111-1	0000-1	10 00	10 00
1001	1101	0110-1	0000-1	10 01	00 11
1010	1111	1010-0	0000-1	10 10	11 01
1011	1110	1011-0	0000-1	10 11	01 11
1100	1010	0011-1	0000-1	11 00	11 10
1101	1011	0010-1	0000-1	11 01	10 11
1110	1001	1110-0	0000-1	11 10	11 11
1111	1000	1111-0	0000-1	11 11	11 00
=20	=16	=28	=2	=32	=16

Apollo

Merged Row/Column Graph

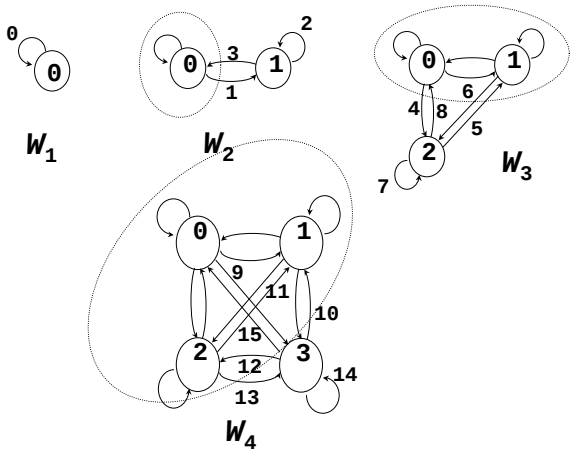




- *forward-edge* (*u*,*v*) representing address <*uv*>; weight equal to *H*(*u*,*v*); *internal switching activity* of address <*uv*>
- Consider two consecutive addresses <u_1v_1> and <u_2v_2>; back-edge (v_1, u_2) ; weight equal to *H*(v_1, u_2); *external switching activity* on the bus
- **Back-edges are not shown in the graphs**

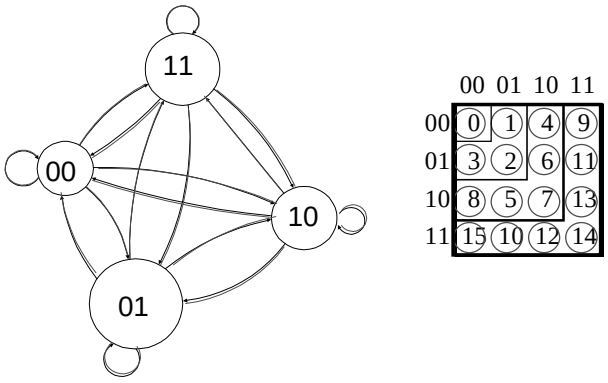
Apollo

Eulerian Cycle Traversal Order



Theorem: A Eulerian cycle of the merged RC graph yields a power-optimal multiplexed code for sequential addressing of the corresponding address space

Pyramid Code



Pyramid code virtually eliminates all the external switching activity, if the access pattern exhibits a pure sequential pattern. As a result, Pyramid code applied to a conventional DRAM bus can cut the switching activity in half

Pyramid Encoding Function

```
1:      Edge (k,j,dir) {
2:              if (dir==1)
3:                      return <k,j>;
4:              else
5:                      if (k==j)
6:                              return <k,0>;
7:                      else
8:                              return <j,k>;
9:      }
10:
11:     Pyramid_Encoder (x) {
12:     p = ⌊√x⌋;
13:     q = x − p^2;
14:             return Edge(p,q/2+q%2, q%2);
15:     }
```

Theorem : The Pyramid encoding function generates a power-optimal multiplexed code for a conventional mode DRAM address bus

An Example 4-Bit Encoder

Take 2^4 as an example, the original 4-bit address $b_3b_2b_1b_0$ will be encoded into Pyramid address $a_3a_2a_1a_0$. The Boolean functions describing the encoded bits are given below

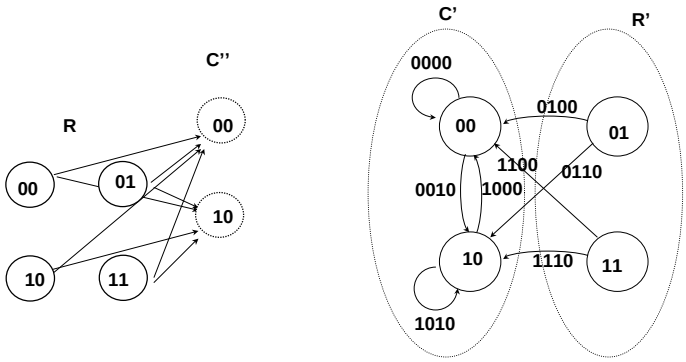
$$a_3 = b_2b_0 + b_3\overline{b_0}$$

$$a_2 = b_3b_1 + b_1\overline{b_0} + \overline{b_3}\,\overline{b_2}b_0 + b_3b_2\overline{b_0}$$

$$a_1 = b_2\overline{b_0} + \overline{b_3}b_2b_1 + b_3b_2\overline{b_1} + b_3\overline{b_2}b_0$$

$$a_0 = \overline{b_3}b_2 + b_3\overline{b_2}b_1 + b_3b_1\overline{b_0} + \overline{b_2}b_1\overline{b_0}$$

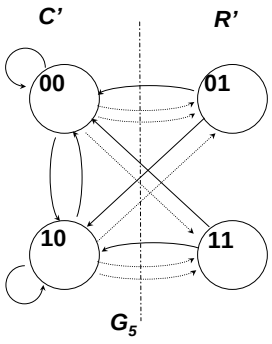
Merged RC Graph for an Aligned Access w/ L=2



Notice again that the forward-edges that represent the internal switching activities are shown while the back-edges that represent the external switching activities are not shown

Apollo

Modified Merged RC Graph for Burst DRAM



Theorem: A Eulerian cycle of the modified merged RC graph yields a power-optimal multiplexed code for sequential burst-mode addressing of the corresponding address space

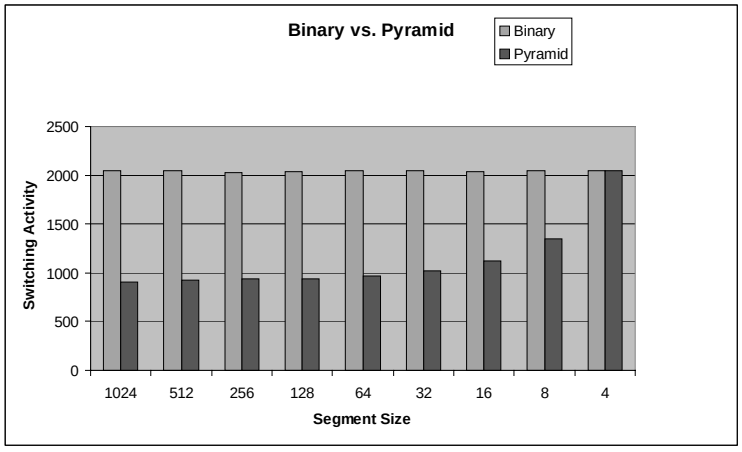
Apollo

Example Burst Pyramid Code

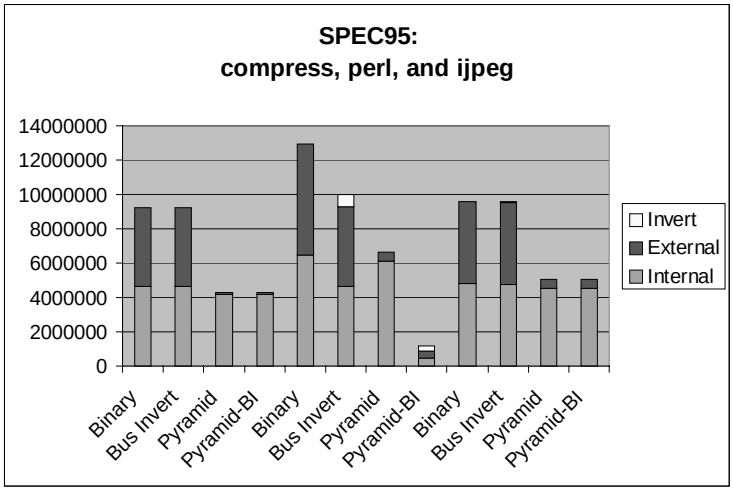
- It is then easy to construct the *Burst Pyramid code*. For the example in the previous figure, we obtain the following code*:* {0000, 0100, 0110, 1100, 0010, 1010, 1110, 1000}
- The four underlined addresses are added to the original Pyramid code and cause external switching activity represented by the back-edges (00,01), (00,01), (10,11), and (10,11)
- The encoding function is:

$$a_0 = b_0$$
$$a_1 = b_3\overline{b_2} + b_2\overline{b_1}$$
$$a_2 = \overline{b_3}b_2 + b_2\overline{b_1}$$
$$a_3 = b_3b_1 + b_3b_2 + b_2b_1$$

Experimental Results I

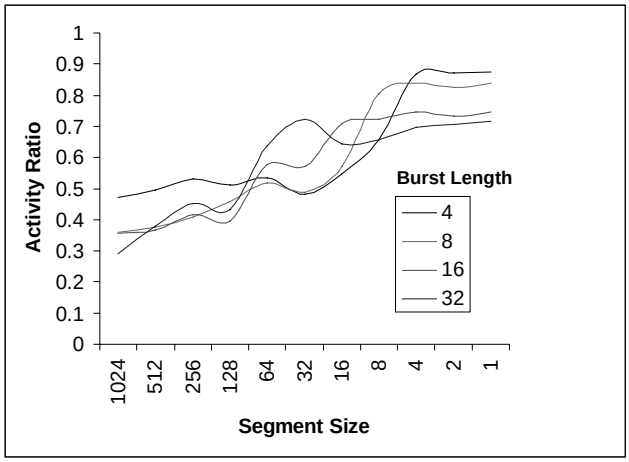


Experimental Results II



Apollo

Experimental Results III



Apollo

Summary

- Power-aware encoding techniques are needed to reduce power dissipation in highly capacitive address busses.
- Alborz encoding cuts the switching activity on SRAM address bus by as much as 83%
- Burst Pyramid encoding cuts in half the switching activity for sequential access pattern on a DRAM bus
- Future work will focus on developing irredundant, adaptive codes for both address and data busses.