Realization of Multiple-Output Functions by Sequential Look-Up Table Cascade

Hui QIN†, Tsutomu SASAO‡,††, Munehiro MATSUURA†, Shinobu NAGAYAMA†, Kazuyuki NAKAMURA‡, and Yukihiro IGUCHI†††

† Department of Computer Science and Electronics, Kyushu Institute of Technology
680-1, Kawazu, Iizuka, Fukuoka, 820-8502 Japan
‡† Center for Microelectronics Systems, Kyushu Institute of Technology
680-1, Kawazu, Iizuka, Fukuoka, 820-8502 Japan
††† Department of Computer Science, Meiji University
1-1-1, Higasimita, Tamaku, Kawasaki, Kanagawa, 214-8571 Japan

E-mail: † {qinhui,nagayama}@aries02.cse.kyutech.ac.jp, ‡ {sasao,matsuura}@cse.kyutech.ac.jp,
†† nakamura@cms.kyutech.ac.jp, ††† iguchi@cs.meiji.ac.jp

Abstract A look-up table (LUT) cascade is a new type of a programmable logic device (PLD), which provides an alternative way to realize multiple-output functions. Two types of LUT cascades exist: A combinational LUT cascade, and a sequential LUT cascade. Comparison with a combinational LUT cascade, the sequential LUT cascade is more flexible. A prototype of a sequential LUT cascade has been custom-designed with 0.35μm CMOS technology. Simulation results show that the sequential LUT cascade is 77 to 229 times faster than software programs on SH-1 with the same clock frequency, and 35 to 88 times faster than software programs on PentiumIII with the same clock frequency, but is a little bit slower than the commercial FPGAs.

Key words LUT cascade, Multiple-output function, Reconfigurable logic, Programmable logic device.
1. Introduction

Programmable logic devices (PLDs) are widely used for prototyping and final products to reduce turn-around time and financial risk. In this paper, we consider a realization of multiple-output logic functions by using PLDs. Various methods exist to realize multiple-output logic functions. Among them, RAMs and programmable logic arrays (PLAs) directly implement logic functions. However, when the number of input variables \( n \) is large, the necessary amount of the hardware becomes too large. Thus, field programmable gate arrays (FPGAs) are often used. However, FPGAs require both physical and logic design. Also, without the complete physical design, the prediction of the performance of FPGAs is hard because the area and delay for the interconnections are often much larger than that for logic cells.

A look-up table (LUT) \([1]\) cascade is a new type of a PLD. Since it has a memory-like structure and the interconnections are simple, prediction of the circuit performance is easy. Two types of LUT cascades exist: A combinational LUT cascade, and a sequential LUT cascade.

A combinational LUT cascade shown in Fig. 1 consists of several memories (cells) connected in series to realize a function. Although the combinational LUT cascade is simple and fast, the logical capability is low. Once the number of inputs, outputs of cell, and the number of cells are fixed, the number of realizable functions is limited. Thus, an efficient use of memory is hard.

A sequential LUT cascade shown in Fig. 2 simulates a combinational LUT cascade sequentially using just one memory for logic. Since the number of inputs, the number of outputs of the cell and the number of cells can be changed by using programmable connection network, the sequential LUT cascade can implement wider range of functions. Also, memory-packing \([1]\) technology can be used to reduce total amount of memory.

In this paper, we show a prototype of a sequential LUT cascade custom-designed with 0.35\(\mu m\) CMOS technology, and compare its performance with microprocessors and FPGAs. The rest of the paper is organized as follows: Section 2 explains the architecture and features of the sequential LUT cascade. Section 3 presents the circuit design of the prototype. The experimental results are shown in Section 4, and Section 5 concludes the paper.

2. Sequential LUT Cascade

In this section, we explain the architecture and features of a sequential LUT cascade.

The sequential LUT cascade shown in Fig. 2 consists of five parts: The Input Reg. stores the values of primary inputs; the Output Reg. stores the values of primary outputs; the Memory for Logic stores the LUT data for cells, where all of the LUT data are stored in one memory. The Control Part consists of the Control block that generates control signals and the Memory for Interconnections that stores the information of the interconnections among cells; the Programmable Connection Network implements the interconnections among cells.

A sequential LUT cascade simulates the combinational LUT cascade sequentially. Although it is slower than the combinational one, its logical ability is much higher. In the sequential LUT cascade, the number of inputs, outputs of the cell, and the number of cells can be changed by using programmable connection network.

Compared with the combinational LUT cascade, the sequential LUT cascade has the following features:
- Requires only one memory for LUT data.
- Can implement wide range of functions.
- Can use memory-packing \([1]\) to reduce total amount of memory.
- 30 to 100 times faster than a microprocessor.

[Example 1] Consider a 4-bit adder.

\[
\begin{align*}
& x_4 \ x_3 \ x_2 \ x_1 \\
+ & y_4 \ y_3 \ y_2 \ y_1 \\
= & S_4 \ S_3 \ S_2 \ S_1
\end{align*}
\]

The adder can be implemented by a combinational LUT cascade with four independent cells as shown in Fig. 3. Note that the number of cells is four.

However, if we use a sequential LUT cascade, we need just one memory for LUT data. Figs. 4 (a)-(d) show the operation of the 4-bit adder by the sequential LUT cascade. Suppose that the memory for logic with five address lines \((A_4, A_3, A_2, A_1, A_0)\), and eight outputs \((D_7, D_6, D_5, D_4, D_3, D_2, D_1, D_0)\).

The combinational cascade consists of four cells. In this case, the sequential cascade requires four steps to compute the outputs of the
adder. We partition the memory into four pages, and assume that the two most significant bits of the address denote the page. In the first step, the top two most significant bits are set to (0,0), as shown in Fig. 4(a). This corresponds to the page 1 in the sequential cascade, and the first cell in the combinational cascade. We have to read the memory to obtain \( C_1 \) and \( S_1 \). Since the first cell has only two inputs, the least significant bit of the address can be either 0 or 1. This operation is symbolically denoted by:

\[
(A_4, A_3, A_2, A_1, A_0) \leftarrow (0, 0, x_1, y_1, \times),
\]

where \( \times \) denotes either 0 or 1.

After reading the values of \( (C_1, S_1) \), \( C_1 \) is transferred to the least significant bits of the address for the next lookup. At the same time, \( S_1 \) is set to the least significant bit of the output register. This wiring is done by the programmable connection network in Fig. 2. The information for the connection for this step is stored in the memory for interconnections. This operation is symbolically denoted by:

Read \( (D_1, D_0) \), and let \( (C_1, S_1) \leftarrow (D_1, D_0) \).

Let OUT_REG[0] \( \leftarrow S_1 \).

In the second step, the 2nd page is used to obtain \( (C_2, S_2) \) as shown in Fig. 4(b). Symbolically, the operation is denoted by:

\[
(A_4, A_3, A_2, A_1, A_0) \leftarrow (0, 1, x_2, y_2, C_1).
\]

Read \( (D_1, D_0) \), and let \( (C_2, S_2) \leftarrow (D_1, D_0) \).

Let OUT_REG[1] \( \leftarrow S_2 \).

In the third step, the 3rd page is used to obtain \( (C_3, S_3) \) as shown in Fig. 4(c). Symbolically, the operation is denoted by:

\[
(A_4, A_3, A_2, A_1, A_0) \leftarrow (1, 0, x_3, y_3, C_2).
\]

Read \( (D_1, D_0) \), and let \( (C_3, S_3) \leftarrow (D_1, D_0) \).

Let OUT_REG[2] \( \leftarrow S_3 \).

In the last step, the 4th page is used to obtain \( (C_{out}, S_4) \) as shown in Fig. 4(d). Symbolically, the operation is denoted by:

\[
(A_4, A_3, A_2, A_1, A_0) \leftarrow (1, 1, x_4, y_4, C_3).
\]

Read \( (D_1, D_0) \), and let \( (C_{out}, S_4) \leftarrow (D_1, D_0) \).

Let OUT_REG[4:3] \( \leftarrow (C_{out}, S_4) \).

Note that it simulates the combinational LUT cascade in Fig. 3. In this way, the 4-bit adder is evaluated by accessing the memory four times. (End of Example)

### 3. Circuit Design

This section presents a transistor-level design using 0.35μm 3.3V CMOS technology, and evaluates the delay of the prototype using circuit simulator SPICE.

#### 3.1 Circuit Design of the Prototype

The specifications of the prototype sequential LUT cascade are as follows:

- The number of external inputs is at most 32.
- The number of outputs is at most 24.
- The number of cells of LUT cascades, \( s \) is at most 8.
- The maximum number of inputs for each cell, \( k \) is 13. Each cell may have different number of inputs.
- The maximum number of outputs for each cell is 8.
- It can use memory-packing to reduce total amount of memory.

Fig. 5 shows the architecture of the prototype. It consists of the following components:

**Input Reg.:** 32-bit, stores the values of primary inputs.

**Output Reg.:** 24-bit, stores the values of primary outputs.

**MAR:** 13-bit memory address register, stores the values of intermediate address.

**64K-bit SRAM:** Stores the LUT data for cells. 13 inputs and 8 outputs.

**Control Part:** Includes two blocks. The RAM stores the information of the interconnections among cells, and the control block consists of one counter and several logic gates that generate the control signals for the LUT cascade.

**Shifter Network:** Consists of several barrel shifters and a 32-bit register that stores the values of intermediate outputs of cells. It implements the programmable interconnections among cells.

The prototype has two important parts: The 64K-bit SRAM and the shifter network. Since the SRAM should operate as a data path in the cascade mode, we use an asynchronous SRAM. Because design methods of SRAMs are described in literature [2], we just show the circuit design of the barrel shifter. The delay of an \( n \)-bit shifter is proportional to \( \log n \), so combined with the fast transmission gates, shift can be fast [2]. Furthermore, we reduced chip area by using a single \( n \)-channel pass transistor instead of a full transmission gate.

Fig. 6 shows the detail design of a 4-by-4 barrel shifter.

#### 3.2 Operation Modes in Cascades

The sequential LUT cascade has two modes: The **configuration mode** and the **cascade mode**. In the configuration mode, all of the address lines of two RAMs are directly connected to the inputs through two multiplexers. 8-bit data input lines which connected to \( \text{DIN} \) in Fig. 5 are used to store LUT data for cells into the 64K-bit SRAM, while another 28-bit data input lines are used to store the information of interconnection into the RAM in the control part. In the cascade mode all the address lines of 64K-bit SRAM are connected to the MAR, while the address lines of the RAM in the control part are connected to the control block, and both RAMs are in
To use the LUT cascade, first, the chip is set to the configuration mode: We need to store LUT data into the 64K-bit SRAM and store the interconnection information into the memory for interconnection. Then, the chip is set to the cascade mode. Finally, we can get the desired result after the evaluation time.

### 3.3 Operation Speed of the Prototype

To obtain the operation speed of the sequential cascade, we partition the sequential LUT cascade into three parts as shown in Fig. 7. The **In Part** consists of input pad (I Pad) and input register, the **Out Part** consists of output pad (O Pad) and output register, and the others are represented as the **Sequential Part** dotted line. Fig. 8 shows the delay time in the signal path, where the values are obtained by SPICE simulation for a 0.35µm, 3.3V CMOS process.

In Fig. 8, the delay for **In Part** is denoted by Delay1, and the delay for **Out Part** is denoted by Delay4. The delay for the sequential part depends on the number of clocks. In the prototype, during one clock cycle we can access the 64K-bit SRAM once. Note that before the memory access, we need to setup the MAR with the values of cell address. The time to setup the MAR for the first cell is denoted by Delay2, which is shorter than the time for the rest of the cells because its path is C to D. The time to setup the MAR for the other cells is equal to one clock cycle. Thus, when the combinational cascade consists of $s$ cells, the time to setup the MAR for cells is equal to $(\text{Delay2} + (s-1) \times \text{CLK})$, where CLK is 4.5 ns. However, we also need one clock cycle to transfer the outputs of the last cell to the point E. And then after Delay3, the primary outputs are out of the sequential part. Therefore, the delay time for the sequential part is equal to $(\text{Delay2} + s \times \text{CLK} + \text{Delay3})$. And the total delay of the sequential LUT cascade is obtained by the following:

\[
\text{Delay} = \text{Delay1} + \text{Delay2} + s \times \text{CLK} + \text{Delay3} + \text{Delay4}
\]

\[
= 0.63 + 2.8 + 4.5s + 0.8 + 1.64
\]

\[
= 4.5s + 5.9 \text{ (ns)}
\]

[ Example 2 ] Consider the 4-bit adder in Example 2.1, where $s$ is
4. The evaluation time is $4.5 \times 4 + 5.9$ (ns).

Fig. 9 shows the distribution of the evaluation time for the 4-bit adder. First, after Delay1, the values of $(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$ are stored into the Input Reg. Second, before accessing the 64K-bit SRAM, we need to spend Delay2 to setup the MAR with the values of $(0, 0, 0, x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$ for Cell1, where $x$ denotes either 0 or 1. Then, to obtain $C_2$ and $S_2$, we need to spend one clock cycle to setup the MAR with the values of $(0, 0, 1, x_2, y_2, C_1, x, x, x, x, x, x, x, x)$ for Cell2, meanwhile $S_1$ is transferred to the point of E. During the next clock, the values of $(0, 1, 0, x_3, y_3, C_2, x, x, x, x, x, x, x, x, x)$ have to be sent to the MAR for Cell3 to obtain $C_3$ and $S_3$, meanwhile $S_1$ is stored in the 32-bit register, and $C_2$ is transferred to the point of P. Similarly, to obtain $C_{out}$ and $S_4$, the values of $(0, 1, 1, x_4, y_4, C_3, x, x, x, x, x, x, x, x, x)$ have to be sent to the MAR for Cell4, meanwhile $S_2$ is stored in the 32-bit register, and $S_3$ is transferred to the point of E. During the next clock, $S_3$ is stored in the 32-bit register, then $C_{out}$ and $S_4$ are transferred to the point of E. During Delay3, both $C_{out}$ and $S_4$ are stored in the 32-bit register, and then the values of $(S_1, S_2, S_3, C_{out}, S_4)$ are transferred to the output register. Finally, we can get the evaluated results after Delay4.

(End of Example)

4. Experimental Results

We evaluated the performance of the prototype, and compared it with microprocessors and FPGAs. To make the argument simple, we selected functions from MCNC combinational benchmark set [3].

Table 1 compares the performance of the prototype with two microprocessors and commercial FPGAs. In this table, “Name”, “In”, and “Out” denote the function name, the number of inputs, and the number of outputs, respectively. The column “#s” denotes the number of cells in the LUT cascade. To obtain the number of cells, we used the newly developed logic synthesis tool [4], where the number of inputs for each cell $k$ is set to 10. The column “Time” denotes the evaluation time for the prototype estimated by the equation 1 in Section 3.3, in nano second.

4.1 Comparison with Microprocessors

At least two approaches exist to implement software for a combinational benchmark functions. The first approach is to simulate the multi-level combinational circuit by some logic simulator. The second approach is to represent the function by a binary decision diagram (BDD), and then traverse the BDD by a special program [5], [6]. In this experiment, we used the second approach, since it was faster than the first approach for the benchmark functions.

We represented benchmark functions using binary decision diagrams (BDDs) for characteristic functions (CFs). The numbers of nodes in BDDs for CFs were reduced by using sifting algorithm [7]. Then, we generated a table that represents the BDD, and we used a special program to traverse the BDD data.

To compare the performance of the prototype with the speed for the software programs, we used RISC microprocessor SH-1 (SH7020) 20MHz [8] and PentiumIII 1GHz with 256KB cache [9]. SH-1 is an embedded MPU used in DVDs, navigation systems, digital cameras, etc. On the other hand, PentiumIII is a high-performance CPU for desktop PCs. For SH-1, we compiled the software program using SH C/C++ compiler [10] with the optimization option for speed, and obtained the CPU time using SH simulator [11]. For PentiumIII, we compiled the software program using GNU C-compiler gcc with -O2 option, and obtained the CPU time by executing it on Linux with 4GB memory.

In Table 1, the column “SH-1” denotes the average CPU time per test vector for the software program on SH-1, in micro second. When the number of inputs for benchmark function is smaller than 17, we obtained the average CPU time for the benchmark function using the exhaustive test (i.e., $2^n$ test vectors, where $n$ is the number of inputs). When the number of inputs for benchmark function is larger than or equal to 17, we obtained the average CPU time for the benchmark function using 1,000,000 random test vectors. Similarly, the column “PenIII” denotes the average CPU time per test vector for the software program on PentiumIII, in micro second. PentiumIII was too fast to obtain the average CPU time for the benchmark function using 1,000,000 random test vectors. Thus, for all benchmark functions, we used 1,000,000 random test vectors. In the last three columns, the column “SH-1” denotes the relative speed of LUT cascades to the speed of SH-1, where the speed of SH-1 with 222MHz is set to 1. Similarly, the column “PenIII” denotes the relative speed of LUT cascades to the speed of PentiumIII, where the speed of PentiumIII with 222MHz is set to 1. That is, they are calculated by

$$\text{LUT/SH-1} = \frac{\text{SH-1 time} \times 20\text{MHz}/222\text{MHz}}{	ext{LUT cascade time}}$$

$$\text{LUT/PenIII} = \frac{\text{PenIII time} \times 1\text{GHz}/222\text{MHz}}{	ext{LUT cascade time}}$$

Note that 222MHz is the clock frequency for the LUT cascades.

Table 1 shows that the sequential LUT cascades are 77 to 229 times faster than SH-1, and 35 to 88 times faster than PentiumIII when the clock frequencies for LUT cascade and MPUs are equal.

4.2 Comparison with FPGAs

The same MCNC benchmark functions were implemented by FPGAs. Firstly, each benchmark function was optimized by SIS tool [12] with script.algebraic and then it was converted to Verilog HDL source. Secondly, by Synplify Pro (version: 7.3.3)[13] it was optimized for Altera EP20K30EFC144-1 FPGA device (1.8V, 0.18µm)[14]. Finally, Quartus (version 2000.09)[14] was used to map it into the FPGA.

In Table 1, the column “LEs” denotes the number of logic elements actually used in the FPGA. The column “Delay” denotes the delay time for each benchmark function obtained by Quartus, in
Table 1 Comparison of sequential LUT cascades with MPUs and FPGA.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>misex2</td>
<td>25</td>
<td>18</td>
<td>5</td>
<td>28.4</td>
<td>4.19</td>
<td>0.36</td>
<td>35</td>
<td>13.2</td>
<td>133</td>
<td>57</td>
</tr>
<tr>
<td>misex3</td>
<td>14</td>
<td>14</td>
<td>4</td>
<td>23.9</td>
<td>38.6</td>
<td>0.30</td>
<td>263</td>
<td>23.4</td>
<td>146</td>
<td>57</td>
</tr>
<tr>
<td>mlp6</td>
<td>12</td>
<td>12</td>
<td>7</td>
<td>37.4</td>
<td>42.0</td>
<td>0.34</td>
<td>1074</td>
<td>29.0</td>
<td>101</td>
<td>41</td>
</tr>
<tr>
<td>in4</td>
<td>32</td>
<td>20</td>
<td>8</td>
<td>41.9</td>
<td>45.6</td>
<td>0.39</td>
<td>109</td>
<td>17.3</td>
<td>98</td>
<td>42</td>
</tr>
<tr>
<td>chkn</td>
<td>29</td>
<td>7</td>
<td>5</td>
<td>28.4</td>
<td>24.3</td>
<td>0.22</td>
<td>178</td>
<td>24.8</td>
<td>77</td>
<td>35</td>
</tr>
<tr>
<td>b3</td>
<td>32</td>
<td>20</td>
<td>7</td>
<td>37.4</td>
<td>45.3</td>
<td>0.37</td>
<td>103</td>
<td>18.4</td>
<td>109</td>
<td>45</td>
</tr>
<tr>
<td>b2</td>
<td>16</td>
<td>17</td>
<td>6</td>
<td>32.9</td>
<td>43.0</td>
<td>0.33</td>
<td>216</td>
<td>25.9</td>
<td>118</td>
<td>45</td>
</tr>
<tr>
<td>amd</td>
<td>14</td>
<td>24</td>
<td>7</td>
<td>37.4</td>
<td>54.6</td>
<td>0.43</td>
<td>107</td>
<td>16.5</td>
<td>132</td>
<td>52</td>
</tr>
<tr>
<td>apex4</td>
<td>9</td>
<td>19</td>
<td>4</td>
<td>23.9</td>
<td>49.0</td>
<td>0.39</td>
<td>820</td>
<td>23.5</td>
<td>185</td>
<td>74</td>
</tr>
<tr>
<td>bc0</td>
<td>26</td>
<td>11</td>
<td>5</td>
<td>28.4</td>
<td>31.9</td>
<td>0.25</td>
<td>417</td>
<td>24.0</td>
<td>101</td>
<td>40</td>
</tr>
<tr>
<td>intb</td>
<td>15</td>
<td>7</td>
<td>3</td>
<td>19.4</td>
<td>32.3</td>
<td>0.25</td>
<td>393</td>
<td>36.2</td>
<td>150</td>
<td>58</td>
</tr>
<tr>
<td>prom2</td>
<td>9</td>
<td>21</td>
<td>3</td>
<td>19.4</td>
<td>49.4</td>
<td>0.38</td>
<td>764</td>
<td>25.1</td>
<td>229</td>
<td>88</td>
</tr>
<tr>
<td>tial</td>
<td>14</td>
<td>8</td>
<td>3</td>
<td>19.4</td>
<td>34.0</td>
<td>0.27</td>
<td>357</td>
<td>20.9</td>
<td>158</td>
<td>63</td>
</tr>
<tr>
<td>p1</td>
<td>8</td>
<td>18</td>
<td>6</td>
<td>32.9</td>
<td>46.2</td>
<td>0.38</td>
<td>91</td>
<td>14.0</td>
<td>127</td>
<td>52</td>
</tr>
<tr>
<td>m4</td>
<td>8</td>
<td>16</td>
<td>3</td>
<td>19.4</td>
<td>39.8</td>
<td>0.30</td>
<td>204</td>
<td>19.9</td>
<td>185</td>
<td>70</td>
</tr>
<tr>
<td>s9dn</td>
<td>27</td>
<td>7</td>
<td>5</td>
<td>28.4</td>
<td>28.2</td>
<td>0.26</td>
<td>31</td>
<td>15.9</td>
<td>89</td>
<td>41</td>
</tr>
<tr>
<td>gary</td>
<td>15</td>
<td>11</td>
<td>4</td>
<td>23.9</td>
<td>29.8</td>
<td>0.24</td>
<td>181</td>
<td>20.4</td>
<td>112</td>
<td>45</td>
</tr>
<tr>
<td>exam</td>
<td>10</td>
<td>10</td>
<td>2</td>
<td>14.9</td>
<td>34.3</td>
<td>0.26</td>
<td>148</td>
<td>18.8</td>
<td>207</td>
<td>79</td>
</tr>
<tr>
<td>t2</td>
<td>17</td>
<td>16</td>
<td>5</td>
<td>28.4</td>
<td>42.3</td>
<td>0.33</td>
<td>67</td>
<td>14.4</td>
<td>134</td>
<td>52</td>
</tr>
</tbody>
</table>

Relative speed of LUT = \( \frac{\text{Delay time of FPGA}}{\text{LUT cascade time}} \).

Table 1 shows that the prototype is a little bit slower than the FPGAs for many functions. Note that FPGA uses 0.18 µm CMOS technology, while the LUT cascade uses 0.35 µm CMOS technology. In spite of the advantages of process technology in FPGAs, the sequential LUT cascade gives a competitive performance to the FPGAs.

5. Conclusions

In this paper, we have shown a realization of multiple-output functions by a sequential LUT cascade. A prototype of a sequential LUT cascade has been custom-designed by using 0.35 µm CMOS technology. Our experiment results show that the sequential LUT cascades is 77 to 229 times faster than software programs on SH-1 with the same clock frequency as the LUT cascade, and 35 to 88 times faster than software programs on PentiumIII with the same clock frequency, but is a little bit slower than the commercial FPGAs.

Acknowledgments

This research is partly supported by JSPS, the Grant in Aid for Scientific Research, and MEXT, the Kitakyushu area innovative cluster project.

References


