Computation structures

Pierre Wolper

Email: Pierre.Wolper@ulg.ac.be
URL: http://www.montefiore.ulg.ac.be/~pw/

http://www.montefiore.ulg.ac.be/~pw/cours/struct.html

References

Course objectives

   (a) To study in detail how a simple programmable machine can be built.
   (b) To introduce processor performance enhancement techniques.
   (c) To overview the concepts used in order to make the task of programming a computer more manageable: interrupts and supervisor (privileged) mode, virtual memory, ...

2. Operating systems and Concurrent programming
   (a) To introduce programming with processes and the organization of an operating system kernel.
   (b) To develop the concepts pertaining to programming with processes and interprocess communication.
   (c) To study classical parallel programming problems.
   (d) To give a brief introduction to questions concerning parallel machines and how they are programmed.

Exercises and programming assignments

• Exercises on processor design and microcode.
• Exercises and a programming assignment in assembly language on low level programming and operating system kernels.
• Exercises and a programming assignment on using parallel processes and interprocess communication (using C as programming language).
Digital electronic circuits: review

We will use basic facts about digital circuits.

- In digital circuits, an electrical signal is used to represent a Boolean (binary) piece of information (0, 1), for example [0, 1] volts represents 0 and [2, 3] volts represents 1.

- Boolean circuits can perform operations on the represented Boolean values.

- We will use three types of circuits: combinatorial circuits, memory elements and synchronous sequential circuits.

Digital circuits

- Each $f_i$ is a Boolean function ($\land, \lor, \neg$) of $x_1, \ldots, x_n$. If the inputs ($x_i$) change, there is a certain delay before the outputs ($f_i$) stabilize at the correct value.

- When LD is active (value 1), the values of $x_1, \ldots, x_n$ are stored at the address given by $a_1, \ldots, a_k$ at the next (0 → 1) clock transition. The values of the outputs $y_1, \ldots, y_n$ are equal to the values stored at the address $a_1, \ldots, a_k$.

- A register is a one location memory (no address)
Digital circuits
Synchronous sequential circuits

- At each clock transition, a Boolean function of the inputs and of the previous register contents is loaded in the registers.
- The outputs are a Boolean function of the inputs and of the register contents.

2. A large capacity memory is included in the “data” part (the machine’s main memory).
3. The behavior of the control part is influenced by the data part:
   - control influenced - by the results of operations on the data;
   - control determined by a program stored in the data path memory.

From synchronous circuits to processors

A processor is a synchronous sequential circuit with a few peculiarities.

1. The circuit is divided into a “data” and a “control” part. The “data” part performs operations on memorized data; the “control” determines which operations are performed.

A simple data path

In the data path shown below, one finds memory elements and a computation element (a combinatorial circuit) interconnected by a bus.
A bus is a collection of shared interconnection lines.

- Each element with an output connected to the bus can be active (0 or 1) or inactive (disconnected – this is called 3 state logic).

- The bus presented is a simple bus with a centralized control managed by the control unit.

The signals needed for the data path to operate will be generated by the control unit.

To design the control unit one must first define the machine that is to be built.

A machine architecture:
The \( \beta \) machine

- A 32 bit architecture: the registers contain 32 bits and the memory addresses are 32 bits long, which makes it possible to address \( 2^{32} \) bytes of memory.

- The machine has 32 (0 à 31) 32-bit registers and a 32-bit program counter whose value is always a multiple of 4. Register 31 always contains the value 0.

- All instructions are 32-bit long.

\[ \beta \] Machine: instruction formats

All operations of the \( \beta \) Machine (except those accessing the main memory) are performed on registers.

There are two possible instruction formats:

- The format without a literal (\textit{valeur constante})

  \[
  \begin{array}{cccccccccc}
  31 & 26 & 25 & 21 & 20 & 16 & 15 & 11 & 10 & 0 \\
  \hline
  \text{Opcode} & \text{Re} & \text{Ra} & \text{Rb} & \text{unused} \\
  \end{array}
  \]

- The format with a literal

  \[
  \begin{array}{cccccc}
  31 & 26 & 25 & 21 & 20 & 16 & 15 & 0 \\
  \hline
  \text{Opcode} & \text{Re} & \text{Ra} & \text{value (2's complement)} \\
  \end{array}
  \]
### Machine: arithmetic instructions

<table>
<thead>
<tr>
<th>Opcode</th>
<th>name</th>
<th>Opcode</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0:20</td>
<td>ADD</td>
<td>0:30</td>
<td>ADDC</td>
</tr>
<tr>
<td>0:21</td>
<td>SUB</td>
<td>0:31</td>
<td>SUBC</td>
</tr>
<tr>
<td>0:22</td>
<td>MUL</td>
<td>0:32</td>
<td>MULC</td>
</tr>
<tr>
<td>0:23</td>
<td>DIV</td>
<td>0:33</td>
<td>DIVC</td>
</tr>
</tbody>
</table>

**ADD(Ra,Rb,Rc):**

\[ \text{PC} \leftarrow \text{PC} + 4 \]
\[ \text{Reg[Rc]} \leftarrow \text{Reg[Ra]} + \text{Reg[Rb]} \]

**ADDC(Ra,literal,Rc):**

\[ \text{PC} \leftarrow \text{PC} + 4 \]
\[ \text{Reg[Rc]} \leftarrow \text{Reg[Ra]} + \text{SEXT(literal)} \]

SEXT(literal) represents the value contained in the instruction extended from 16 to 32 bits, the sign being preserved.

The definitions of SUB, MUL and DIV are similar.

### Machine: comparison instructions

<table>
<thead>
<tr>
<th>Opcode</th>
<th>name</th>
<th>Opcode</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0:24</td>
<td>COMPEQ</td>
<td>0:34</td>
<td>COMPEQC</td>
</tr>
<tr>
<td>0:25</td>
<td>COMPLT</td>
<td>0:35</td>
<td>COMPLTC</td>
</tr>
<tr>
<td>0:26</td>
<td>CMPLE</td>
<td>0:36</td>
<td>CMPLEC</td>
</tr>
</tbody>
</table>

**COMPEQ(Ra,Rb,Rc):**

\[ \text{PC} \leftarrow \text{PC} + 4 \]
\[ \text{if } \text{Reg[Ra]} = \text{Reg[Rb]} \text{ then } \text{Reg[Rc]} \leftarrow 1 \text{ else } \text{Reg[Rc]} \leftarrow 0 \]

**COMPEQC(Ra,literal,Rc):**

\[ \text{PC} \leftarrow \text{PC} + 4 \]
\[ \text{if } \text{Reg[Ra]} = \text{SEXT(literal)} \text{ then } \text{Reg[Rc]} \leftarrow 1 \text{ else } \text{Reg[Rc]} \leftarrow 0 \]

CMPLT and CMPLTC compare according to the \(<\) relation, CMPLE and CMPLEC according to \(\leq\).

### Machine: logical instructions

<table>
<thead>
<tr>
<th>Opcode</th>
<th>name</th>
<th>Opcode</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0:28</td>
<td>AND</td>
<td>0:38</td>
<td>ANDC</td>
</tr>
<tr>
<td>0:29</td>
<td>OR</td>
<td>0:39</td>
<td>ORC</td>
</tr>
<tr>
<td>0:2A</td>
<td>XOR</td>
<td>0:3A</td>
<td>XORC</td>
</tr>
</tbody>
</table>

**AND(Ra,Rb,Rc):**

\[ \text{PC} \leftarrow \text{PC} + 4 \]
\[ \text{Reg[Rc]} \leftarrow \text{Reg[Ra]} \& \text{Reg[Rb]} \]

**ANDC(Ra,literal,Rc):**

\[ \text{PC} \leftarrow \text{PC} + 4 \]
\[ \text{Reg[Rc]} \leftarrow \text{Reg[Ra]} \& \text{SEXT(literal)} \]

OR and ORC compute “or”, XOR and XORC exclusive “or”.

### Machine: shift instructions

<table>
<thead>
<tr>
<th>Opcode</th>
<th>name</th>
<th>Opcode</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0:2C</td>
<td>SHL</td>
<td>0:3C</td>
<td>SHLC</td>
</tr>
<tr>
<td>0:2D</td>
<td>SHR</td>
<td>0:3D</td>
<td>SHRC</td>
</tr>
<tr>
<td>0:2E</td>
<td>SRA</td>
<td>0:3E</td>
<td>SRAC</td>
</tr>
</tbody>
</table>

**SHL(Ra,Rb,Rc):**

\[ \text{PC} \leftarrow \text{PC} + 4 \]
\[ \text{Reg[Rc]} \leftarrow \text{Reg[Ra]} \ll \text{Reg[Rb]}_{4:0} \]

**SHLC(Ra,literal,Rc):**

\[ \text{PC} \leftarrow \text{PC} + 4 \]
\[ \text{Reg[Rc]} \leftarrow \text{Reg[Ra]} \ll \text{literal}_{4:0} \]

The content of Ra is shifted by a number of positions given by the 5 least significant bits of Rb and the result is placed in Rc. 0 bits are introduced in the freed positions.

SHR and SHRC shift to the right (0 bits are introduced in the freed positions).

SRA and SRAC also shift to the right, but the sign bit (Reg[Ra]_{31}) is introduced in the freed positions.
### Machine: memory access instructions

<table>
<thead>
<tr>
<th>Opcode</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x18</td>
<td>LD</td>
</tr>
<tr>
<td>0x19</td>
<td>ST</td>
</tr>
</tbody>
</table>

LD(Ra, literal, Rc) :
- \(PC \leftarrow PC + 4\)
- \(EA \leftarrow \text{Reg}[Ra] + \text{SEXT}(\text{literal})\)
- \(\text{Reg}[Rc] \leftarrow \text{Mem}[EA]\)

ST(Rc, literal, Ra) :
- \(PC \leftarrow PC + 4\)
- \(EA \leftarrow \text{Reg}[Ra] + \text{SEXT}(\text{literal})\)
- \(\text{Mem}[EA] \leftarrow \text{Reg}[Rc]\)

EA is called the “effective address”.

### Machine: memory access instructions II

<table>
<thead>
<tr>
<th>Opcode</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x1F</td>
<td>LDR</td>
</tr>
</tbody>
</table>

LDR(label, Rc) :
- \(PC \leftarrow PC + 4\)
- \(EA \leftarrow PC + 4 \times \text{SEXT}(\text{literal})\)
- \(\text{Reg}[Rc] \leftarrow \text{Mem}[EA]\)

In LDR, the instruction given to the assembler contains a label. (label). In the assembled instruction, the literal is computed from the label as follows:

\[
literal = ((\text{OFFSET(}label) - \text{OFFSET(current inst.)}) ÷ 4) - 1
\]

### Machine: branch instructions

<table>
<thead>
<tr>
<th>Opcode</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x1B</td>
<td>JMP</td>
</tr>
</tbody>
</table>

JMP(Ra, Rc) :
- \(PC \leftarrow PC + 4\)
- \(EA \leftarrow \text{Reg}[Ra] \& 0xFFFFFFFFFC\)
- \(\text{Reg}[Rc] \leftarrow PC\)
- \(PC \leftarrow EA\)

The 2 least significant bits of Ra are set to 0 to force the address to be a word address. The address of the instruction following the JMP is saved in Rc.

### Machine: branch instructions II

<table>
<thead>
<tr>
<th>Opcode</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x1D</td>
<td>BEQ/BEF</td>
</tr>
<tr>
<td>0x1E</td>
<td>BNE/BT</td>
</tr>
</tbody>
</table>

BEQ(Ra, label, Rc) :
- \(PC \leftarrow PC + 4\)
- \(EA \leftarrow PC + 4 \times \text{SEXT}(\text{literal})\)
- \(\text{TMP} \leftarrow \text{Reg}[Ra]\)
- \(\text{Reg}[Rc] \leftarrow PC\)
- \(\text{if} \quad \text{TMP} = 0 \quad \text{then} \quad PC \leftarrow EA\)

In BEQ, the instruction given to the assembler contains a label. In the assembled instruction, the literal is computed as follows:

\[
literal = ((\text{OFFSET(}label) - \text{OFFSET(current inst.)}) ÷ 4) - 1
\]

BNE is similar to BEQ except that the test is “different from 0”

TMP is used because Ra and Rc could be the same register.
An implementation of the $\beta$ Machine:
The machine ULg01

We start from the data path described previously.

- The static RAM will be used to implement the registers.
- The dynamic RAM will be used to implement the main memory.
- One still has to introduce a program counter (PC) and a control unit.

A detailed view of the PC of ULg01

Only 20 bits are used because the machine being built will not have more than 4 megabytes (1 megaword) of memory.

A detailed view of the SRAM of ULg01

The circuit is designed to produce the value 0 when register 31 is read.
The DRAM module must contain refresh circuitry.

The clock speed is chosen so that the DRAM can operate at the same speed as the rest of the machine. In an optimized implementation, the DRAM operates at a lower speed.

The ALU flags are used as address bits into the ROM. This allows the implementation of conditional instructions. The flags are:

- \( E \) whose value is 1 is the output of the ALU is equal to 0xFFFFF
- \( C \) the complemented carry bit,
- \( \times \) the sign bit (bit 31).

The left-hand part of the control unit is used to separate the different parts of an instruction and connect them to the appropriate lines of the bus.

All that needs to be done to finish the implementation of the \( \beta \) Machine, is to write the microcode that has to be placed in the ROM.

The instruction **OR(Ra, Rb, Rc)**

<table>
<thead>
<tr>
<th>Opcode</th>
<th>Phase</th>
<th>Flags</th>
<th>LATCH Flags</th>
<th>ALU Flags</th>
<th>DR</th>
<th>SEL</th>
<th>PC+</th>
</tr>
</thead>
<tbody>
<tr>
<td>101001</td>
<td>0000</td>
<td>*</td>
<td>1</td>
<td>0000000</td>
<td>01</td>
<td>001</td>
<td>0</td>
</tr>
<tr>
<td>101001</td>
<td>0001</td>
<td>*</td>
<td>1</td>
<td>0000000</td>
<td>01</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>101001</td>
<td>0010</td>
<td>*</td>
<td>1</td>
<td>0000000</td>
<td>01</td>
<td>010</td>
<td>0</td>
</tr>
<tr>
<td>101001</td>
<td>0011</td>
<td>*</td>
<td>1</td>
<td>0000000</td>
<td>01</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>101001</td>
<td>0100</td>
<td>*</td>
<td>1</td>
<td>0000000</td>
<td>01</td>
<td>000</td>
<td>0</td>
</tr>
<tr>
<td>101001</td>
<td>0101</td>
<td>*</td>
<td>1</td>
<td>111001</td>
<td>101</td>
<td>011</td>
<td>0</td>
</tr>
<tr>
<td>101001</td>
<td>0110</td>
<td>*</td>
<td>1</td>
<td>0000000</td>
<td>100</td>
<td>110</td>
<td>1</td>
</tr>
<tr>
<td>101001</td>
<td>0111</td>
<td>*</td>
<td>1</td>
<td>0000000</td>
<td>000</td>
<td>101</td>
<td>0</td>
</tr>
</tbody>
</table>

The control unit is microprogrammed and allows up to 16 phases by instruction. The microprogram is stored in a ROM.
### The instruction **JMP(Ra, Rc)**

<table>
<thead>
<tr>
<th>Opcode</th>
<th>Phase</th>
<th>Flags</th>
<th>Latch</th>
<th>ALU</th>
<th>LD</th>
<th>DR</th>
<th>PC+</th>
</tr>
</thead>
<tbody>
<tr>
<td>011011</td>
<td>0000</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>011</td>
<td>001</td>
<td>0</td>
</tr>
<tr>
<td>011011</td>
<td>0001</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>001</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>011011</td>
<td>0010</td>
<td>*</td>
<td>0</td>
<td>111110</td>
<td>001</td>
<td>011</td>
<td>0</td>
</tr>
<tr>
<td>011011</td>
<td>0011</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>011</td>
<td>000</td>
<td>0</td>
</tr>
<tr>
<td>011011</td>
<td>0100</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>101</td>
<td>110</td>
<td>0</td>
</tr>
<tr>
<td>011011</td>
<td>0101</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>100</td>
<td>110</td>
<td>1</td>
</tr>
<tr>
<td>011011</td>
<td>0110</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>000</td>
<td>101</td>
<td>0</td>
</tr>
</tbody>
</table>

**Reminder:** The literal that is found in the instruction interpreted by the microcode is computed (by the assembler) from the label found in the symbolically written instruction as follows

\[
literal = \left( (\text{OFFSET}(\text{label}) - \text{OFFSET}(\text{current inst.})) \div 4 \right) - 1
\]

### The instruction **BEQ(Ra, label, Rc)**

<table>
<thead>
<tr>
<th>Opcode</th>
<th>Phase</th>
<th>Flags</th>
<th>Latch</th>
<th>ALU</th>
<th>LD</th>
<th>DR</th>
<th>PC+</th>
</tr>
</thead>
<tbody>
<tr>
<td>011101</td>
<td>0000</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>011</td>
<td>001</td>
<td>0</td>
</tr>
<tr>
<td>011101</td>
<td>0001</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>001</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>011101</td>
<td>0010</td>
<td>*</td>
<td>0</td>
<td>111110</td>
<td>001</td>
<td>011</td>
<td>0</td>
</tr>
<tr>
<td>011101</td>
<td>0011</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>111</td>
<td>000</td>
<td>0</td>
</tr>
<tr>
<td>011101</td>
<td>0100</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>101</td>
<td>110</td>
<td>0</td>
</tr>
<tr>
<td>011101</td>
<td>0101</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>100</td>
<td>110</td>
<td>1</td>
</tr>
<tr>
<td>011101</td>
<td>0110</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>000</td>
<td>101</td>
<td>0</td>
</tr>
</tbody>
</table>

### L’instruction **LD(Ra, literal, Rc)**

<table>
<thead>
<tr>
<th>Opcode</th>
<th>Phase</th>
<th>Flags</th>
<th>Latch</th>
<th>ALU</th>
<th>LD</th>
<th>DR</th>
<th>PC+</th>
</tr>
</thead>
<tbody>
<tr>
<td>011000</td>
<td>0000</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>011</td>
<td>001</td>
<td>0</td>
</tr>
<tr>
<td>011000</td>
<td>0001</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>001</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>011000</td>
<td>0010</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>010</td>
<td>010</td>
<td>0</td>
</tr>
<tr>
<td>011000</td>
<td>0011</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>011</td>
<td>000</td>
<td>0</td>
</tr>
<tr>
<td>011000</td>
<td>0100</td>
<td>*</td>
<td>1</td>
<td>100110</td>
<td>100</td>
<td>011</td>
<td>0</td>
</tr>
<tr>
<td>011000</td>
<td>0101</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>101</td>
<td>101</td>
<td>0</td>
</tr>
<tr>
<td>011000</td>
<td>0110</td>
<td>*</td>
<td>1</td>
<td>000000</td>
<td>000</td>
<td>101</td>
<td>0</td>
</tr>
</tbody>
</table>

**Reminder:**

The literal that is found in the instruction interpreted by the microcode is computed (by the assembler) from the label found in the symbolically written instruction as follows

\[
literal = \left( (\text{OFFSET}(\text{label}) - \text{OFFSET}(\text{current inst.})) \div 4 \right) - 1
\]