SUMMARY A 32-bit arithmetic logic unit (ALU) is designed for a rapid single flux quantum (RSFQ) bit-parallel processor. In the ALU, clocked gates are partially replaced by clockless gates. This reduces the number of D flip flops (DFFs) required for path balancing. The number of clocked gates, including DFFs, is reduced by approximately 40%, and size of the clock distribution network is reduced. The number of pipeline stages becomes modest. The layout design of the ALU and simulation results show the effectiveness of using clockless gates in wide datapath circuits.

key words: SFQ digital circuit, ALU, wide datapath circuit, bit-parallel processor, clockless gate

1. Introduction

Rapid single flux quantum (RSFQ) circuits [1] and their energy-efficient derivatives [2]-[4] are expected to be used for realizing high-performance and energy-efficient computing systems, which cannot be achieved using CMOS technologies. Substantial progress has been made in the field of superconductive electronics in the past decades. The current SFQ manufacturing technology is capable of accommodating over 800,000 Josephson junctions (JJs) on a die [5]. It is expected that a 32-bit/64-bit SFQ microprocessor will be developed in the near future [6].

Several SFQ microprocessors were studied [7]-[15]. Until now, most successfully demonstrated SFQ microprocessors are up to 8-bit bit-serial microprocessors [9], [10], [14], [15]. A 32-bit bit-serial microprocessor would require at least 320 ps for processing one word even if it is operated at 100 GHz, and it is not superior to existing high-end CMOS microprocessors. Therefore, it is desired to develop 32-bit/64-bit bit-parallel microprocessors. To develop such a microprocessor, it is necessary to establish a method for designing a wide datapath circuit.

In SFQ digital circuits, SFQ appears as a voltage pulse (SFQ pulse), and it is used as the carrier of information. Logic values of 1 and 0 is typically represented by the presence and absence of a data pulse, respectively, between two consecutive clock pulses. Ordinary logic gates are with a clock input and are referred to as clocked gates. A clocked gate has a latching function, which stores data until a clock pulse arrives. All clocked gates must be supplied with clock pulses. The zero-skew clocking scheme or concurrent-flow clocking scheme are generally used in these gates [16]-[19].

In these clocking schemes, the number of clocked-gate stages must be the same in all paths from the input to the output of a circuit. A large number of clocked buffers, i.e., D flip flops (DFFs), must be inserted for this path balancing. In addition, clock pulses must also be supplied to these DFFs. The clocking schemes lead to a deep-pipelined circuit, which is referred to as a gate-level-pipelined circuit, where the number of pipeline stages (pipeline depth) is the same as the number of clocked-gate stages.

Although implementing a wide datapath circuit as a gate-level-pipelined circuit results in an extremely high throughput, it requires a large number of DFFs for path balancing and a large clock distribution network. One way to solve this problem is to use clockless gates. A well-known example of a clockless gate is an asynchronous AND gate [20], [21] which can be derived from a simpler Muller C-element (coincidence junction) [1]. A clockless AND gate and a clockless NIMPLY (not imply) gate based on a non-destructive read out (NDRO) were proposed [22], [23]. A clockless dynamic AND gate has been proposed in recent years [24]. Dorojevets et al. suggested the use of asynchronous AND gates and a hybrid wave pipeline for designing a wide datapath circuit [12] and proposed 8-bit arithmetic logic units (ALUs) [25], [26] and a 16-bit adder [27].

In this paper, we present the design of a 32-bit ALU using clockless gates. The ALU consists of 10 gate stages, each of which is composed of clockless gates. These stages do not require DFFs for path balancing. The number of clocked gates, including DFFs, is reduced by approximately 40%. There are 6 pipeline stages. The partial replacement of clocked gates by clockless gates reduces the number of DFFs required for path balancing, and size of the clock distribution network. It also makes the number of pipeline stages modest. We create a layout design the ALU with the clockless gates based on an NDRO [22] and clocked gates in the CONNECT cell library [28] for the AIST ADP2 fabrication process [29]. The ALU consists of 16,727 JJs. Its size is 2.475 mm × 7.080 mm = 17.5230 mm². The estimated clock cycle is 92.7 ps, and the operating frequency is 10.8 GHz. The latency is 556.2 ps (6 cycles). The design has been verified using the static timing analysis and behavior abstraction tools in [30].

The rest of this paper is organized as follows: In Section 2, we explain the clockless gates based on the NDRO. In Section 3, we present the design of the 32-bit ALU with clockless gates. Section 4 concludes the paper.

†The author is with the Graduate School of Informatics, Kyoto University
2. Clockless Gates

In the proposed ALU, we use the clockless AND gate and clockless NIMPLY gate based on an NDRO [22], [23].

An NDRO is widely used in RSFQ circuits. It has three inputs, \(c\), \(s\), and \(r\), and one output, \(d\). There are two states, \(s0\) and \(s1\). The state is set to \(s1\) by pulse arrival at \(s\) and reset to \(s0\) by pulse arrival at \(r\). At \(s1\), a pulse is produced at \(d\) by pulse arrival at \(c\). At \(s0\), no pulse is produced at \(d\) by pulse arrival at \(c\). The interval between the arrival times at \(c\), \(s\), and \(r\), must be sufficiently long.

The clockless AND gate consists of an NDRO and two delay elements, as shown in Fig. 1. It has two inputs, \(i1\) and \(i2\), and an output, \(o\). \(i1\) is connected to \(c\) through delay element \(delay1\). \(i2\) is directly connected to \(s\) and \(r\) through delay element \(delay2\). \(delay2\) delays a pulse more than \(delay1\). A pulse is produced at \(o\) only when pulses arrive at both \(i1\) and \(i2\). When pulses arrive at both \(i1\) and \(i2\), their arrival times must be close to each other.

A clockless AND cell was realized using the AIST 10-kA/cm\(^2\) Advanced Process (ADP2) [29]. It consists of 25 JJs. Its size is 60 \(\mu\)m × 60 \(\mu\)m. The delay from pulse arrival at \(i1\) to pulse output is 20.9 ps. Note that the timing of the pulse output depends only on that of the pulse arrival at \(i1\). The delay varies according to the bias current, similar to clocked gates. The delay shown above is the nominal delay at 100 \% bias current. When pulses arrive at both \(i1\) and \(i2\), they must arrive within approximately 14 ps. If the first pulse arrives at \(i1\), the second pulse must arrive at \(i2\) within 14.3 ps. If the first pulse arrives at \(i2\), the second pulse must arrive at \(i1\) within 14.0 ps.

The clockless NIMPLY gate consists of an NDRO and two delay elements, as shown in Fig. 2. It calculates \(\overline{i1 \land i2}\). The difference between the clockless NIMPLY and AND gates is the position of \(delay2\). In the clockless NIMPLY gate, \(delay2\) is inserted in the path from \(i2\) to \(s\) instead of that from \(i2\) to \(r\). A pulse is produced at \(o\) only when a pulse arrives at \(i1\) and no pulse arrives at \(i2\). When pulses arrive at both \(i1\) and \(i2\), their arrival times must be close to each other.

A clockless NIMPLY cell was also realized. Its JJ count, size, and delay are the same as those of the AND gate. When pulses arrive at both \(i1\) and \(i2\), they must arrive within approximately 10 ps. If the first pulse arrives at \(i1\), the second pulse must arrive at \(i2\) within 10.0 ps. If the first pulse arrives at \(i2\), the second pulse must arrive at \(i1\) within 9.8 ps.

We implement a clockless XOR compound gate using two NIMPLY gates and a CB, as shown in Fig. 3. Its JJ count is 57 (= 25 × 2 + 7). We directly connect the NIMPLY gates and CB.

### Table 1: ALU operations

<table>
<thead>
<tr>
<th>Operation</th>
<th>Op-XOR</th>
<th>Op-AND</th>
<th>Op-ADD</th>
<th>Inv-X</th>
<th>Inv-Y</th>
<th>C-in</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADD</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SUB</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AND</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>OR</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>XOR</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>XNOR</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>EQ</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

The arithmetic operations are more complex and hardware consuming than the logic operations. Therefore, an adder is used as the main component of the ALU. High-performance adders typically use prefix trees, which generate carries in \(\log_2 n\) stages, where \(n\) is the number of bits of the datapath. For example, the Kogge-Stone adder [31] and Sklansky adder [32] were used in 4-bit/8-bit SFQ adders. However, the former requires numerous long wires, e.g., \(n\) wires spanning \(n/2\) bit positions, and the latter requires numerous (maximum \(n\)) fanouts.

A parallel-prefix adder structure with a moderate number of long wires and fanouts is required for wide datapath SFQ adders. Similar to the 16-bit SFQ adder in [27], we develop an adder structure based on the sparse-tree adder [33]. In the sparse-tree adder, every four carries are calculated using a carry-merge tree (a prefix tree), and the sum is calculated using 4-bit conditional sum-generators in which the carry-merge logic is serially connected. In [27], a carry-skip adder was used instead of the conditional sum-generators. In
our adder, every four carries are calculated using the prefix tree, and they are used to calculate the other carries in parallel similar to the Sklansky adder. An adder structure based on the sparse-tree adder is suitable for wider (e.g., 64-bit) SFQ ALUs. The adder structure must be tuned according to the data width.

A structural diagram of the ALU is shown in Fig. 4. The ALU consists of nine types of blocks. The details of the blocks are shown in Fig. 5. Blocks \( pq^* \), \( PG^* \), and \( C^* \) consist of clockless gates. Block \( pq^* \) produces a 'propagation' signal, i.e., XOR of the inputs, and a 'generation' signal, i.e., AND of the inputs. As mentioned in the previous section, we can implement a clockless XOR compound gate using two clockless NIMPLY gates and a CB. Block \( SEL \) produces the result for a specified logic operation. Block \( PG^* \) calculates the prefix operation, and block \( C^* \) calculates a carry. Block \( S \) calculates a sum. Note that \( c \) is 0 for logic operations.

The ALU consists of 10 gate stages. The second, fourth, sixth, and ninth stages are composed of clockless gates. There are 6 pipeline stages. We target the operation of the ALU at 10 GHz. On this basis, we determine the gate stage that should be clockless so that the delay at each pipeline stage is balanced considering fanouts and wire lengths.

In the second stage, in each \( pq^* \), a clockless XOR compound gate is used instead of a clocked XOR gate. In the fourth, sixth, and ninth gate stages, \( PG^* \) and \( C^* \) are used instead of \( PG \) and \( C \), respectively. In other words, clockless AND gates are used instead of clocked AND gates. These stages no longer require DFFs for path balancing; the number of DFFs is reduced by approximately 200. Overall, the number of clocked gates, including DFFs, is reduced by more than 300 (approximately 40%) to 454.

We use the zero-skew clocking scheme for the clocked gates. Even though the concurrent-flow clocking scheme provides a higher operating frequency, clock skew accumulates. This leads to a large time difference in the clock cycle between the input and output of the circuit. This makes it difficult to control the clock cycle of the entire system. Note that the use of clockless gates significantly reduces the number of clocked gates and the DFFs to be supplied with clock pulses.

We create a layout design of the ALU using the clockless gates based on the NDRO and clocked gates in the CONNECT cell library [28] for the AIST ADP2 fabrication process [29]. The design is shown in Fig. 6. The number of JJs is 16,727. The size of the ALU is 2,475 mm \( \times \) 7,080 mm = 17,523 mm\(^2\). The estimated clock cycle is 92.7 ps, and the operating frequency is 10.8 GHz. The latency is 556.2 ps (6 cycles). We have verified the design using the static timing analysis and behavior abstraction tools in [30].

Compared to the design with only clocked gates, 194 DFFs are removed, 79 clocked AND gates are replaced with clockless AND gates, and 32 clocked XOR gates are replaced with clockless XOR compound gates in the new design. Furthermore, 305 (= 194+79+32) splitters for clock distribution are removed. The number of JJs of a DFF with two PTL re-
Fig. 5 Details of the blocks

Fig. 6 Layout design of the ALU
receivers (one for data input and the other for clock input) and a PTL driver is 14 (= 6 + 3 × 2 + 2). The number of JJs of a clocked AND gate with three receivers and a driver is 25 (= 14 + 3 × 3 + 2), and that of a clocked XOR gate with three receivers and a driver is 22 (= 11 + 3 × 3 + 2). On the other hand, the number of JJs of a clockless AND gate with two receivers and a driver is 33 (= 25 + 3 × 2 + 2), and that of a clockless XOR compound gate with four receivers and a driver is 71 (= 57 + 3 × 4 + 2). For a clockless XOR compound gate, two splitters with a driver are required for splitting the data inputs into two NIMPLY gates. The number of JJs of a splitter with a driver is 5 (= 3 + 2). Therefore, the number of JJs in the designed ALU is 1,710 (= 14×194+(25–33)×79+(22–(71+5×2))×32+5×305) less than that of the design with completely clocked gates.

We may reduce the JJ count further using a clockless dynamic AND gate [24] instead of the clockless AND gate based on the NDRO and by developing a clockless XOR gate. From the point of view of the JJ count, it may be better to not replace clocked XOR gates with clockless XOR compound gates.

We can change the number of pipeline stages by changing the number of stages of the clockless gates. For example, we can reduce the number of pipeline stages to 4 by utilizing clockless gates in the second, third, fifth, sixth, eighth, and ninth gate stages. The gate stage in which clockless gates will be used should be selected such that the delay at each pipeline stage is balanced considering fanouts and wire lengths. Furthermore, as the number of pipeline stages decreases, the level of clockless gates in each pipeline stage may increase, and the cycle time may increase. As the clockless gates have the limitation that the pulse arrival times at the inputs must be close to each other, the maximum level of clockless gates in each pipeline stage may be limited by the timing jitter on paths.

4. Conclusion

We have shown the feasibility of realizing a wide datapath circuit using clockless gates by designing a 32-bit ALU. We partially replace clocked gates with clockless gates. This reduces the number of DFFs required for path balancing and size of the clock distribution network, and makes the number of pipeline stages modest.

One of the merits of using clockless gates in the development of bit-parallel processors is the flexibility in determining the number of pipeline stages of component circuits. The results of this study show that the use of clockless gates and the zero-skew clocking scheme for clocked gates is an effective method for designing wide datapath component circuits for 32-bit/64-bit microprocessors.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number 18H05211 and also supported through the activities of VDEC, The University of Tokyo, in collaboration with Cadence Design Systems.

References


