Highly Efficient Design of Pipelined Parallel FFT Using Folding Transformation

Serin Sera Paul¹, Simy M Baby²

M.Tech Student, Applied Electronics, Ilahia College of Engineering & Technology, Mulavoor, Kerala, India¹
Assistant Professor, Dept. of ECE, Ilahia College of Engineering & Technology, Mulavoor, Kerala, India²

ABSTRACT: This paper presents a new parallel pipelined architecture to compute Discrete Fourier Transform (DFT) using FFT architecture. This particular architecture uses folding transformation technique as well as register minimization technique for the design of FFT architecture. Novel architectures for the computation of complex and real valued fast Fourier transform are derived. Pipelining is used to reduce the power consumption. Parallel processing and pipelining exploits concurrency. Parallel processing also aids to the reduction of power consumption by reducing the supply voltage. The power consumption is reduced very effectively using the parallel architecture. This paper includes the comparative study of the speed of operation of FFT architectures using different multipliers.

KEYWORDS: Fast Fourier Transform (FFT), folding, pipelining, parallel processing, register minimization, Vedic-multiplier, Array multiplier, Baugh Wooley multiplier

I. INTRODUCTION

Fast Fourier Transform (FFT) is a commonly used technique for the computation of Discrete Fourier Transform (DFT). DFT computations are required in the fields like filtering, spectral analysis etc. to calculate the frequency spectrum or to identify a system’s frequency response from its impulse response and vice versa. FFT is used in digital video broadcasting and OFDM systems. Much research has been carried out to design pipelined architectures for computation of FFT. The basic one is Radix-2 FFT. Based on the radix-2 FFT approach many algorithms have been developed which includes radix-4 [4], split-radix [3], radix-2²[5] etc. A popularly known algorithm is cooley-Tukey radix-2 FFT [2]. Radix-2 Multipath Delay Commutator (R2MDC) [6] is a classical approach for pipelined implementation of FFT architecture. Radix-2 Single-path Delay Feedback (R2SDF) [7] is another approach with reduced memory obtained by a standard usage of storage buffer in R2MDC. Most of the algorithms require hardware complexity and there is no complete hardware utilisation. The basic aspects like high throughput and low power consumption are required to speed and power requirements while keeping the hardware overhead to a minimum. This paper presents a technique to design the architecture from FFT flow graph. Folding transformation [8], [10] and register minimization [8], [9], [11] are the two important steps included in this FFT algorithm.

Folding Transformation is a technique in which the number of butterflies in the same column is mapped into one butterfly unit. If we consider an FFT of size N, then 2-parallel architecture can be obtained if we consider the folding factor to be N/2 or 4-parallel architecture if considering a folding factor of N/4.
By selecting the appropriate folding sets we can derive the FFT architectures. The folding sets are designed in a way to reduce the number of storage elements and also the latency. The prior FFT architectures had no systematic way of approach. This architecture simplifies the design of FFT and is a systematic approach towards the design of FFT with arbitrary level of parallelism. These are derived either in Decimation-In-Time (DIT) or Decimation-In-Frequency (DIF) flow graphs. FFT architectures can be derived for different radices. Parallel pipelined architectures for the computation of Real valued signals (RFFT) based on radix-2^2 and radix-2^3 and different architectures for the computation of complex valued signals (CFFT) are carried out earlier. This paper is organised into VII sections where section II represents the folding transformation and register minimization based FFT architecture, section III and IV explains about the proposed architecture for CFFT and RFFT respectively. In section VI a comparative study is conducted using different multipliers. Finally, section VII certain conclusions are drawn from the comparative study.

II. FFT ARCHITECTURE USING FOLDING

Folding Transformation as well as Register minimization techniques are used to obtain several FFT architectures. The whole process is explained with the help of 8-point radix-2 DIF FFT which can be extended to different radices. The flow graph of 8-point radix-2 DIF FFT is illustrated in Fig. 1. The twiddle factor indicates a multiplication by W_N^k in between the stages. The Data Flow Graph of Fig.1 is shown in Fig. 2 where each node represents a computation.

DFG is subjected to folding transformation in order to derive a pipelined architecture. For this we require a folding set, which is an ordered set of operations executed by the same functional unit. Every folding set contains number of entries which are called the folding factors. A folding set may include null operations also. Consider two nodes represented as U and V which are connected by an edge e with w(e) delays. The I-th iteration of these nodes be scheduled at K_l+u and K_l+v where l is the number of entries and u and v are the folding orders. The folding equation is represented as

\[ D_F(U \rightarrow V) = Kw(e) - Pu + v - u \]  

where \( P_U \) is the number of pipeline stages. For the DFG in Fig. 2 consider the folding set shown below.

\[ A = \{\phi, \phi, \phi, \phi, A0, A1, A2, A3\} \]
\[ B = \{B2, B3, \phi, \phi, \phi, B0, B1\} \]
\[ C = \{C1, C2, C3, \phi, \phi, \phi, C0\}. \]

We assume that the butterfly operations do not have any pipeline stages. Prior to deriving the folded architecture the folded equations in (1) are to be written for all the edges as shown in (2). \( D_F(A0 \rightarrow B0) = 2 \) means there is an edge with weight 2 from node A to Bin the folded DFG. After obtaining the folding equations we have to determine whether the folding sets are feasible or not.

Fig. 3 Pipelined Data Flow graph of a radix-2 8-point DIF FFT
In the equations obtained some negative delays are observed which needs to be removed. To make sure that the folded architecture has non-negative number of delay the DFG can be pipelined as shown in Fig. 3. For the pipelined DFG the folding equations are given by

\[
\begin{align*}
D_F(A_0 \rightarrow B_0) &= 2 & D_F(B_0 \rightarrow C_0) &= 1 \\
D_F(A_0 \rightarrow B_2) &= -4 & D_F(B_0 \rightarrow C_1) &= -6 \\
D_F(A_1 \rightarrow B_1) &= 2 & D_F(B_1 \rightarrow C_0) &= 0 \\
D_F(A_1 \rightarrow B_1) &= -4 & D_F(B_1 \rightarrow C_1) &= -7 \\
D_F(A_2 \rightarrow B_0) &= 0 & D_F(B_2 \rightarrow C_2) &= 1 \\
D_F(A_2 \rightarrow B_2) &= -6 & D_F(B_2 \rightarrow C_3) &= 2 \\
D_F(A_3 \rightarrow B_1) &= 0 & D_F(B_3 \rightarrow C_2) &= 0 \\
D_F(A_3 \rightarrow B_3) &= -6 & D_F(B_3 \rightarrow C_3) &= 1.
\end{align*}
\] (2)

In the equations obtained some negative delays are observed which needs to be removed. To make sure that the folded architecture has non-negative number of delay the DFG can be pipelined as shown in Fig. 3. For the pipelined DFG the folding equations are given by

\[
\begin{align*}
D_F(A_0 \rightarrow B_0) &= 2 & D_F(B_0 \rightarrow C_0) &= 1 \\
D_F(A_0 \rightarrow B_2) &= 4 & D_F(B_0 \rightarrow C_1) &= 2 \\
D_F(A_1 \rightarrow B_1) &= 2 & D_F(B_1 \rightarrow C_0) &= 0 \\
D_F(A_1 \rightarrow B_1) &= 4 & D_F(B_1 \rightarrow C_1) &= 1 \\
D_F(A_2 \rightarrow B_0) &= 0 & D_F(B_2 \rightarrow C_2) &= 1 \\
D_F(A_2 \rightarrow B_2) &= 2 & D_F(B_2 \rightarrow C_3) &= 2 \\
D_F(A_3 \rightarrow B_1) &= 0 & D_F(B_3 \rightarrow C_2) &= 0 \\
D_F(A_3 \rightarrow B_3) &= 2 & D_F(B_3 \rightarrow C_3) &= 1
\end{align*}
\] (3)

From the above equations we can see 24 registers are required for implementing the folded architecture. As a next step a technique called Lifetime analysis [8], [9], [11] is employed to design the architecture with the minimum number of delays. A lifetime chart is obtained as shown in Fig. 4 for one stage of the 8-point DFG. From the lifetime chart we can analyse that we require only 4 registers to implement the design while considering the outputs of nodes A0, A1, A2 and A3 in the DFG instead of the 16 registers which was used in the straight forward implementation. Next step is Register allocation as shown in Fig. 5. From the folding equations and the table in Fig. 5 the architecture in Fig. 6 can be derived. We can see from the folding sets that half of the time null operations are being executed and therefore the hardware utilization is only 50%.

![Fig. 4 Lifetime chart for variables y0, y1, ..., y](image1)

![Fig. 5 Register allocation table for data shown in fig. 4.](image2)
III. ARCHITECTURE WITH COMPLEX INPUTS (CFFT)

This section presents parallel architecture for complex valued signals based on radix-2 and radix-2^3 algorithms. The approach presented in the previous section can be used to derive all these architectures.

A. 2-parallel radix-2 FFT Architecture

Fig. 7 shows the DFG of radix-2 DIF FFT for N=16 where all the nodes represent radix-2 butterfly operations. Consider the folding sets

\[ A = \{A_0, A_2, A_4, A_6, A_1, A_3, A_5, A_7\} \]
\[ B = \{B_5, B_7, B_9, B_2, B_4, B_6, B_1, B_3\}, \]
\[ C = \{C_3, C_5, C_7, C_0, C_2, C_4, C_6, C_1\} \]
\[ D = \{D_2, D_4, D_6, D_1, D_3, D_5, D_7, D_0\}. \]

We can observe that here the folding sets does not contain any null operations. Thus we can derive the folded architecture using the steps used in the previous section. In this architecture two input samples are processed at the same time. The hardware utilization is 100%. The architecture is shown in Fig. 8. In a similar way the 2-parallel architecture for radix-2 DIT FFT can also be derived using the folding set below. Fig. 9 represents the pipelined DFG and fig. 10 shows the 2-parallel architecture.

\[ A = \{A_0, A_2, A_1, A_3, A_4, A_6, A_5, A_7\} \]
\[ B = \{B_5, B_7, B_0, B_2, B_1, B_3, B_4, B_6\}, \]
\[ C = \{C_6, C_5, C_7, C_0, C_2, C_1, C_3, C_4\} \]
\[ D = \{D_2, D_1, D_3, D_4, D_6, D_5, D_7, D_0\}. \]
B. 4-Parallel Radix-2 FFT Architecture

Consider the folding set shown below using which a 4-parallel architecture can be derived.

\[ A = \{A0, A1, A2, A3\} \quad A' = \{A'0, A'1, A'2, A'3\} \]
\[ B = \{B1, B3, B0, B2\} \quad B' = \{B'1, B'3, B'0, B'2\} \]
\[ C = \{C2, C1, C3, C0\} \quad C' = \{C'2, C'1, C'3, C'0\} \]
\[ D = \{D3, D0, D2, D1\} \quad D' = \{D'3, D'0, D'2, D'1\}. \]

Using the algorithm used in the previous section we can obtain the pipelined DFG as in Fig. 11 and the 4-parallel architecture as shown in Fig. 12.
IV. ARCHITECTURES WITH REAL INPUTS (RFFT)

The input sequence $x[n]$ for RFFT is considered to be real. If $x[n]$ is real then output $X[k]$ is symmetric. i.e.; $X[N-k]=X^*[k]$.

Using this property $(N/2)$ - 1 outputs can be removed which are redundant. A new approach in identifying these redundant samples is proposed in [12]. The shaded regions of the Fig. 13 can be removed as they are all redundant samples identified using the approach in [12] and only $N/2 + 1$ outputs of the FFT are required.

A. 2-Parallel Radix-2 Architecture
The DFG of this architecture is same as Fig. 7 and the folding set is as follows.

$A = \{A0, A2, A4, A6, A1, A3, A5, A7\}$
$B = \{B5, B7, B0, B2, B4, B6, B1, B3\}$,
$C = \{C3, C5, \phi, C0, C2, C4, \phi, C1\}$
$D = \{D2, D4, \phi, D1, \phi, D5, \phi, D0\}$.

The architecture is similar to that shown in Fig. 8 except that first two stages will contain a real data path. The hardware complexity is same as that of the CFFT.

B. 2-parallel Radix-$2^2$ Architecture
Two different scheduling approaches are used to derive two different architectures. It is mainly done by changing the folding order of the butterfly nodes.

1) Scheduling Method 1: The parallel-pipelined architecture is shown in Fig. 14 obtained from Fig. 13. The folding set used is [1], [8]. The scheduling for the architecture is shown in Fig. 15.

$A = \{A0, A2, A4, A6, A1, A3, A5, A7\}$
$B = \{B5, B7, B0, B2, B4, B6, B1, B3\}$,
$C = \{C3, C5, \phi, C0, C2, C4, \phi, C1\}$
$D = \{D2, D4, \phi, D1, \phi, D5, \phi, D0\}$.

2) Scheduling Method 2: This method reduces the number of delay elements and slightly modifies the architecture [1], [8]. The folding set is as follows.
Multipliers play a very important role in FFT architectures as it consumes most of the time. The speed of FFT greatly depends on the multiplier employed in the design. Array Multipliers are used for various architectures due to its regular structure where the multiplier circuit is based on add and shift algorithm. Baugh Wooley Multipliers are generally known for its high speed and low power consumption. And Vedic Multipliers are a particular type of multipliers which are high speed complex multipliers which is an ancient methodology of Indian Mathematics. The speed of FFT architectures varies with the use of different multipliers.

VI. COMPARISON AND ANALYSIS

Researchers had generated parallel pipelined FFT architecture that achieves complete hardware utilization with reduced power consumption [1] when compared to the serial architectures. Here different multipliers including Vedic multiplier, Array Multiplier and Baugh Wooley Multiplier are being employed in separate designs and compared for Time of operation and comparison tables are generated for different parallel FFT architectures. The analysis is done using Modelsim 6.5e.

<table>
<thead>
<tr>
<th>S.no</th>
<th>Multiplier</th>
<th>Time of operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Vedic Multiplier</td>
<td>8500ns</td>
</tr>
<tr>
<td>2</td>
<td>Array Multiplier</td>
<td>6600ns</td>
</tr>
<tr>
<td>3</td>
<td>Baugh Wooley Multiplier</td>
<td>8600ns</td>
</tr>
</tbody>
</table>

Table 1. Time of operation in an 8-point FFT architecture
Table 2. Time of operation in a 16-point FFT architecture

<table>
<thead>
<tr>
<th>Sl. no.</th>
<th>Multiplier</th>
<th>Time of operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Vedic Multiplier</td>
<td>19200ns</td>
</tr>
<tr>
<td>2</td>
<td>Array Multiplier</td>
<td>19500ns</td>
</tr>
<tr>
<td>3</td>
<td>Baugh Wooley Multiplier</td>
<td>21200ns</td>
</tr>
</tbody>
</table>

Table 2. Time of operation in a 16-point FFT architecture

Fig. 16 2-parallel architecture of a radix-2^{2^16} 16-point DIF RFFT

Table 3. Time of operation in a 16-point RFFT architecture

<table>
<thead>
<tr>
<th>Sl. no.</th>
<th>Multiplier</th>
<th>Time of operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Vedic Multiplier</td>
<td>121000ns</td>
</tr>
<tr>
<td>2</td>
<td>Array Multiplier</td>
<td>123000ns</td>
</tr>
<tr>
<td>3</td>
<td>Baugh Wooley Multiplier</td>
<td>123000ns</td>
</tr>
</tbody>
</table>

Table 3. Time of operation in a 16-point RFFT architecture

From the tables we can understand that Vedic multiplier is very much efficient in terms of speed of operation than the other two multipliers.

VII. CONCLUSION

This paper presents a pipelined parallel FFT architecture which has a lesser power consumption compared to serial FFT architectures. It also has the advantage of complete hardware utilization. For a high speed Pipelined parallel FFT architecture a Vedic multiplier can be employed in the particular design. Thus an efficient design can be obtained in terms of power and speed.

REFERENCES

[1] Pipelined Parallel FFT Architectures via Folding Transformation, Manohar Ayinala, Student Member, IEEE, Michael Brown, and Keshab K. Parhi, Fellow, IEEE transactions on very large scale integration (vlsi) systems, vol. 20, no. 6, June 2012.
International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Special Issue 5, December 2014