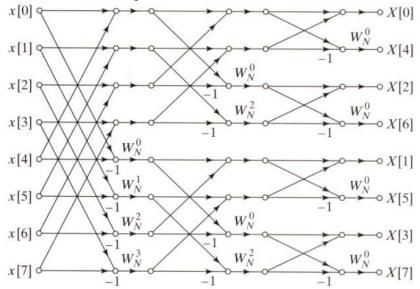
Lab C Report – DFT

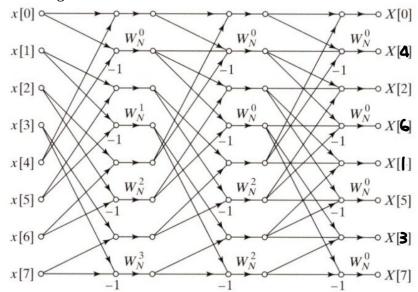
Introduction to the Algorithm and Overall System

The problem is to perform digital Fourier transformation (DFT) to a 1024-size vector, so we implement it by fast Fourier transformation (FFT).

We use the basic Radix-2 FFT. Since the usual FFT has different connecting structure in each stage, which is difficult to perform ARRAY_PARTITION to accelerate the process.^[1]



We adopt the connecting structure shown in the following figure, which has the same structure in each stage:^[1]



(Note: in Wikipedia the result has no bit reversal, but in our experiment and derivation the result must be bit reversal.)

By this algorithm we can first split the array into two blocks, and in each block we may use cyclic ARRAY_PARTITION to accelerate the process. This allow us to choose partition factor with more freedom.

At the final stage we need to do bit reversal process, and we combine it with output stage. The value of cosine, sine and bit reversal are constant and is stored in look-up table to accelerate the process.

Details for Code and Pragma

The main program is a ten-stage FFT, and we use axis as input and output interfaces. To meet the requirement memory bandwidth, the array in each state input (output) is split into two arrays (as block partition) and each of them uses cyclic ARRAY_PARTITION pragma.

```
real_sample[1024], DTYPE imag_sample[1024],DTYPE real_op[1024],DTYPE
 //Write your code here
    a HLS INTERFACE mode=axis port=real_sample
     HLS INTERFACE mode=axis port=imag_sample
     HLS INTERFACE mode=axis port=real_op
    HLS INTERFACE mode=axis port=imag_op
DTYPE R1u[SIZE/2], R2u[SIZE/2], I1u[SIZE/2], I2u[SIZE/2];
DTYPE R1d[SIZE/2], R2d[SIZE/2], I1d[SIZE/2], I2d[SIZE/2];
agma HLS ARRAY_PARTITION variable=R1u,R2u,I1u,I2u,R1d,R2d,I1d,I2d type=cyclic factor=8
 transfer_data_in(real_sample, imag_sample, R1u, R1d, I1u, I1d);
 fft_stage(R1u, R1d, I1u, I1d, 1, R2u, R2d, I2u, I2d);
fft_stage(R2u, R2d, I2u, I2d, 2, R1u, R1d, I1u, I1d);
 fft_stage(R1u, R1d, I1u, I1d, 3, R2u, R2d, I2u, I2d);
 fft_stage(R2u, R2d, I2u, I2d, 4, R1u, R1d, I1u, I1d);
 fft_stage(R1u, R1d, I1u, I1d, 5, R2u, R2d, I2u, I2d);
 fft_stage(R2u, R2d, I2u, I2d, 6, R1u, R1d, I1u, I1d);
 fft_stage(R1u, R1d, I1u, I1d, 7, R2u, R2d, I2u, I2d);
 fft_stage(R2u, R2d, I2u, I2d, 8, R1u, R1d, I1u, I1d);
fft_stage(R1u, R1d, I1u, I1d, 9, R2u, R2d, I2u, I2d);
 fft_stage(R2u, R2d, I2u, I2d, 10, R1u, R1d, I1u, I1d);
bit_reversal(R1u, R1d, I1u, I1d, R2u, R2d, I2u, I2d);
transfer_data_out(R2u, R2d, I2u, I2d, real_op, imag_op);
 bit_reversal_out(R1u, R1d, I1u, I1d, real_op, imag_op);
```

The code also shows that bit reversal and output process are combined into one stage.

This structure of FFT is not suitable for DATAFLOW, since all stages run sequentially and in each stage the memory read and write pattern is not sequential.

```
void fft_stage(DTYPE Ruin[SIZE/2], DTYPE Rdin[SIZE/2], DTYPE Iuin[SIZE/2], DTYPE Idin[SIZE/2],
    int stage, DTYPE Ruout[SIZE/2], DTYPE Rdout[SIZE/2], DTYPE Iuout[SIZE/2], DTYPE Idout[SIZE/2]){
    int SIZE4 = SIZE >> 2;
    int angle = 0;
    for(int i=0; i<SIZE/2; i++){
    #pragma HLS PIPELINE II=1

#pragma HLS UNROLL factor=16
    if(i%2==0){
        Ruout[i] = Ruin[i>>1] + Rdin[i>>1];
        Iuout[i] = Iuin[i>>1] + Idin[i>>1];
    }
    else{
```

Computation in FFT is accelerated by UNROLL pragma.

```
const DTYPE W_real[]={1.0, 0.99998116, 0.9999247, 0.9998306, 0.9996988, 0.9995294,
const DTYPE W_imag[]={-0.0, -0.0061358847, -0.012271538, -0.01840673, -0.024541229,
const int BITR[]={0, 512, 256, 768, 128, 640, 384, 896, 64, 576, 320, 832, 192, 704
480, 992, 16, 528, 272, 784, 144, 656, 400, 912, 80, 592, 336, 848, 208, 720, 464,
1008, 8, 520, 264, 776, 136, 648, 392, 904, 72, 584, 328, 840, 200, 712, 456, 968,
24, 536, 280, 792, 152, 664, 408, 920, 88, 600, 344, 856, 216, 728, 472, 984, 56, 5
```

Cosine, sine and bit reversal are computed and made as look-up tables to accelerate the process. We can make these look-up tables since the size of problem is fixed.

Timing, Performance, Utilization and Trade-off

The trade-off in this program is time vs. resource. If we use larger unroll factor and array partition factor we can reduce the computation time, but it will consume more resource.

The tightest resource is DSP. In FPGA xc7z020clg400-1, the best timing performance we can obtain is using 8 as array partition factor and 16 as unroll factor:

```
        Modules & Loops
        Issue Type
        Violation Type
        Distance
        Slack
        Latency(cycles)
        Latency(cycles)
        Latency(ns)
        Iteration Latency
        Interval
        Trip Count
        Pipeline
        BRAM(%)
        DSP(%)
        FF(%)
        LUT(%)
        URAM(%)

        • o dft
        pipeline_VITIS_LOOP_36_1
        -
        3099
        3.099£4
        -
        3100
        -
        no
        56
        94
        23
        85
        0

        • o dft_Pipeline_VITIS_LOOP_36_1
        -
        514
        -
        no
        0
        0
        -
        -
        0
        0

        • oft_Pipeline_VITIS_LOOP_41_2
        -
        514
        5.140E3
        -
        514
        -
        no
        0
        0
        -
        -
        0
        0
        -
        -
        0
        0
        -
        -
        0
        0
        -
        0
        0
        -
        0
        0
        -
        0
        0
        -
        0
        0
        -
        0
        0
        -
        0
        0
        -
        0
        0
        -
        0
        0
        -
        0</t
```

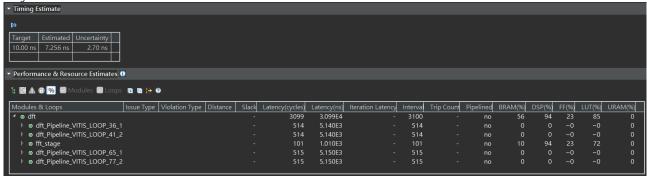
There is another trade-off between clock frequency and cycle count, which will be shown in the next section.

Score

We calculate the score under different constraint and synthesis parameters:

1. Limited to resource used in FPGA, target clock time=10ns:

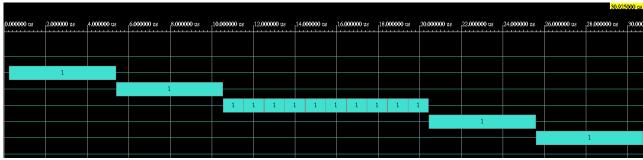
Csynthesis



Cosimulation

Modules & Loops	Avg II	Max II	Min II	Avg Latency	Max Latency	Min Latency
				3059	3059	3059
ø dft_Pipeline_VITIS_LOOP_36_1				512	512	512
ø dft_Pipeline_VITIS_LOOP_41_2				512	512	512
▶ @ fft_stage	99	99	99			
ø dft_Pipeline_VITIS_LOOP_65_1						
▶ o dft_Pipeline_VITIS_LOOP_77_2						

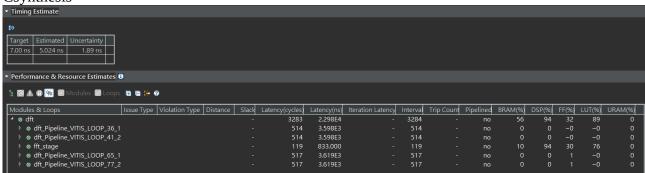
Waveform



Score: $(1 \div (7.256 \times 10^{-9})) \div (309250 \div 10) = 4456.49$

2. Limited to resource used in FPGA, target clock time=7ns:

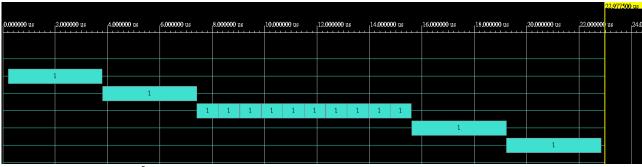
Csynthesis



Cosimulation

Modules & Loops	Avg II	Max II	Min II	Avg Latency	Max Latency	Min Latency
4 ⊚ dft				3245	3245	3245
				512	512	512
▶ @ dft_Pipeline_VITIS_LOOP_41_2				512	512	512
P ⊚ fft_stage	117	117	117	115	115	115
ø dft_Pipeline_VITIS_LOOP_65_1				515	515	515
o dft_Pipeline_VITIS_LOOP_77_2				515	515	515

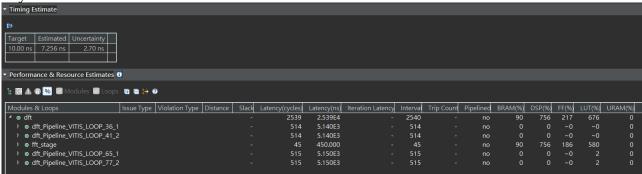
Waveform



Score: (1÷(5.024×10⁻⁹))÷(229775÷7)=6063.81

3. Unlimited to resource used in FPGA, target clock time=10ns, unroll factor=128:

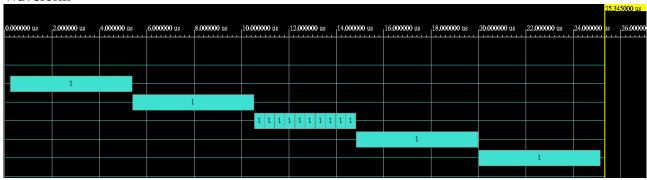
Csynthesis



Cosimulation

Modules & Loops	Avg II	Max II	Min II	Avg Latency	Max Latency	Min Latency
				2501	2501	2501
▶ ø dft_Pipeline_VITIS_LOOP_36_1				512	512	512
				512	512	512
▶ ⊚ fft_stage		43	43		41	41
				513	513	513
▶ @ dft_Pipeline_VITIS_LOOP_77_2				513	513	513

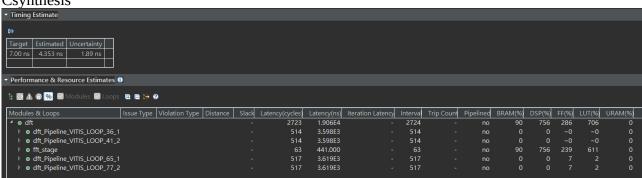
Waveform



Score: (1÷(7.256×10⁻⁹))÷(253450÷10)=5437.64

4. Unlimited to resource used in FPGA, target clock time=7ns, unroll factor=128:

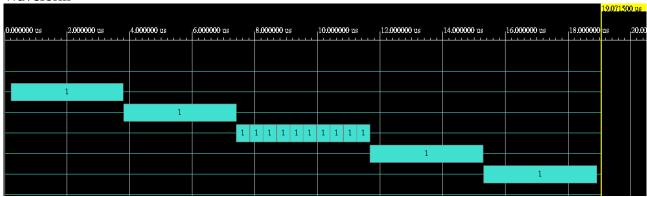
Csynthesis



Cosimulation

Modules & Loops	Avg II	Max II	Min II	Avg Latency	Max Latency	Min Latency
4 ⊚ dft				2687	2687	2687
Ø dft_Pipeline_VITIS_LOOP_36_1				512	512	512
				512	512	512
▶ ⊚ fft_stage				59	59	59
Ø dft_Pipeline_VITIS_LOOP_65_1				515	515	515
				515	515	515

Waveform



Score: (1÷(4.353×10⁻⁹))÷(190715÷7)=8431.88

Final Result (Score)

Resource	Lim	nited	Unlimited		
Target Clock	10ns	7ns	10ns	7ns	
Score	4456.49	6063.81	5437.64	8431.88	

Reference

[1] https://zh.wikipedia.org/zh-tw/%E5%BA%93%E5%88%A9%EF%BC%8D%E5%9B%BE %E5%9F%BA%E5%BF%AB%E9%80%9F%E5%82%85%E9%87%8C%E5%8F%B6%E5%8F%98%E6%8D%A2%E7%AE%97%E6%B3%95#%E5%96%AE%E4%B8%80%E5%9F%BA %E5%BA%95