INTRODUCTION TO

HIGH-PERFORMANCE DIGITAL PROCESSING

INTELLECTUAL PROPERTIES

of

Rice Electronics

Filename: Processing IP Rev M

OVERVIEW

Rice Electronics has created Intellectual Property (IP) for advanced digital processing. These are broadly defined and discussed below as,

- Architecture IP (or "Architectures")
- Modelling IP (or "Models)
- Trigonometric Multiplication IP (or "TMs)

The company's IP extends to unique "Multi-Mode" Architectures, enabling more than 90% size reduction for processors in such areas as;

- Neural Networks

- Wireless Communications Infrastructure (e.g. 5G Networks)

- Pattern/Image Recognition

The Architectures serve as paradigms for high-performance hardware structures, which may be scaled to suit system requirements.

The Architectures incorporate;

- Proprietary building blocks such as the Trigonometric Multipliers (TM) discussed herein

- Proprietary multi-tier structures employing the TMs

The Rice Architectures enable high-performance/low-complexity hardware implementation(s). These can achieve order-of-magnitude improvements (in terms of complexity/cost/performance) over existing hardware entities, such as;

- FPGA or ASIC Form of Specialized/Dedicated Processors

- General Purpose Programmable Processors

- Digital Signal Processors or Graphics Processing Units (DSPs or GPUs)

ARCHITECTURE IP

Operational Modes

The Rice Architecture is capable of "multi-mode" operation. Within a given Mode, various complex Functions may be executed (e.g. such as large FFTs or multi-dimensional convolutions).

The Architecture encompasses three primary Modes. These span the three traditional realms of processing for a broad range of sensors, systems and infrastructure. Accordingly, the "Architecture IP" is a framework for major components of systems, or systems-on-chip (SOC).

The Modes and Functions of the Architecture may be invoked via simple parametric controls (i.e., requiring no explicit programming). The three primary Modes, and examples of their functionality, are given below;

Frequency Domain Mode – Functions include computation of Fourier transforms up to 4096 points for applications such as;

- Signal Analysis/Detection

- Advanced Modulation Techniques (e.g. OFDM)

- Frequency Domain Filtering

Spatial Domain Mode – Functions include high-speed 1-D or 2-D convolutions (e.g. up to 32 X 32) for "intelligent" systems including;

- Neural Networks

- Image Recognition

Time Domain Mode – Functions for high-bandwidth signal processing in;

- Waveform Generation

- Up-conversion/Down-conversion

- Signal Modulation

MODELLING IP

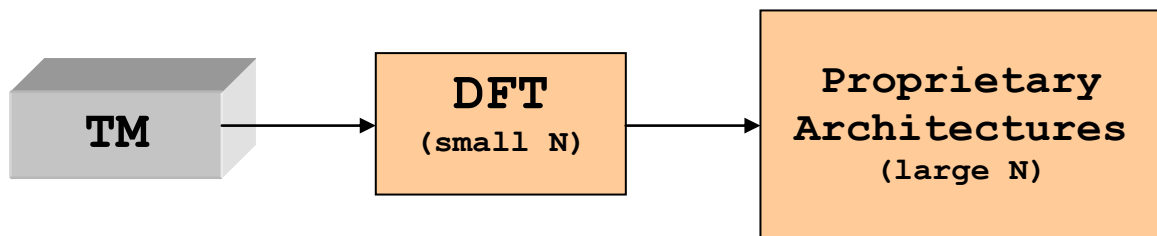# CONTACT COMPANY FOR FURTHER INFORMATION REGARDING MODELLING IP

CONTACT COMPANY FOR FURTHER INFORMATION REGARDING MODELLING IP

ARCHITECTURE IP STRUCTURAL HIERARCHY

The Architecture IP greatly simplifies the construction of high-performance hardware processors.

The Rice IP base employs a Hierarchical Architecture approach to minimize physical structure. With this approach, proprietary "Trigonometric Multipliers" (TMs) replace conventional multipliers (using 100s of gates, instead of 1000s).

In the Hierarchical approach, TMs enable highly efficient "small" DFTs, which in turn create optimal larger Architectures. That is, the TMs serve as "building blocks" for the larger scale "Architecture IP", as exemplified in Figure 1 below.

HIERARCHICAL ARCHITECUTURE APPROACH

FIGURE 1

The hierarchical Architectures greatly reduce logic complexity, data movement and memory organization. Also, computational complexity is far less than traditional implementations of challenging processing functions, with most multiplication operations being eliminated. As a point of comparison, for large Fourier transform functions;

Rice Architecture IP $\approx$ **2N($\sqrt{N}$)** <u>addition</u> operations for N-point real-value FFT

Traditional approach $\approx$ **(N/2) log$_2$N** <u>complex multiplies</u> for N-point complex FFT

The Architecture hierarchy is also captured within the body of the Modelling IP. Both structural and computational results of the approach are depicted therein.

TRIGONOMETRIC MULTIPLIERS (TM) IP

The Proprietary TM replaces traditional multiplier building blocks. The TM (requiring hundreds of gates versus thousands for conventional multipliers) drastically reduces logic complexity of the Architectures.

The TM is specialized for arithmetic calculations associated with many transform operations. The basic TM calculation as illustrated below;

Function of the Trigonometric Multiplier (TM)

- *The TM performs the basic computation stated below:*

$$A \cdot (e^{-2\pi i k/N}) = A \cdot \cos(2\pi k/N) - i[A \cdot \sin(2\pi k/N)]$$

- *Where N and k are integers,*

- *N is a power of 2,*

- *A is a fixed point operand,*

- *Cosine and Sine values are fixed point numbers generated internally to the TM*

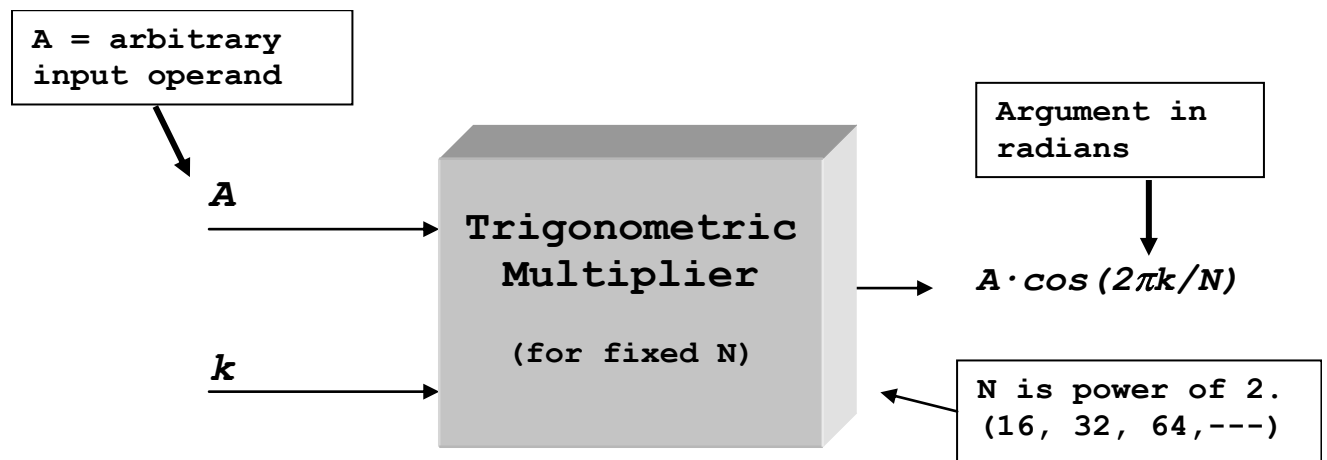**TRIGONOMETRIC MULTIPLIER CALCULATION**

FIGURE 2

The TMs may be implemented in hardware using only addition operators (i.e., precluding need for any conventional multiplication circuitry).

Although not discussed herein, the TMs are naturally applicable to the calculation of DCTs and stand-alone small DFTs, enabling significant optimization of such functions. However, this discussion focuses on Architectures for challenging large-vector operations, wherein the TM is a building block.

The Rice IP includes variations of the TM which may be instantiated in the Architecture to meet system requirements at minimal complexity. The Modelling IP can be used to support such hardware optimization.

The TM can generate trigonometric values internally, thus mitigating the need for external storage of such values in ROM or other memory.
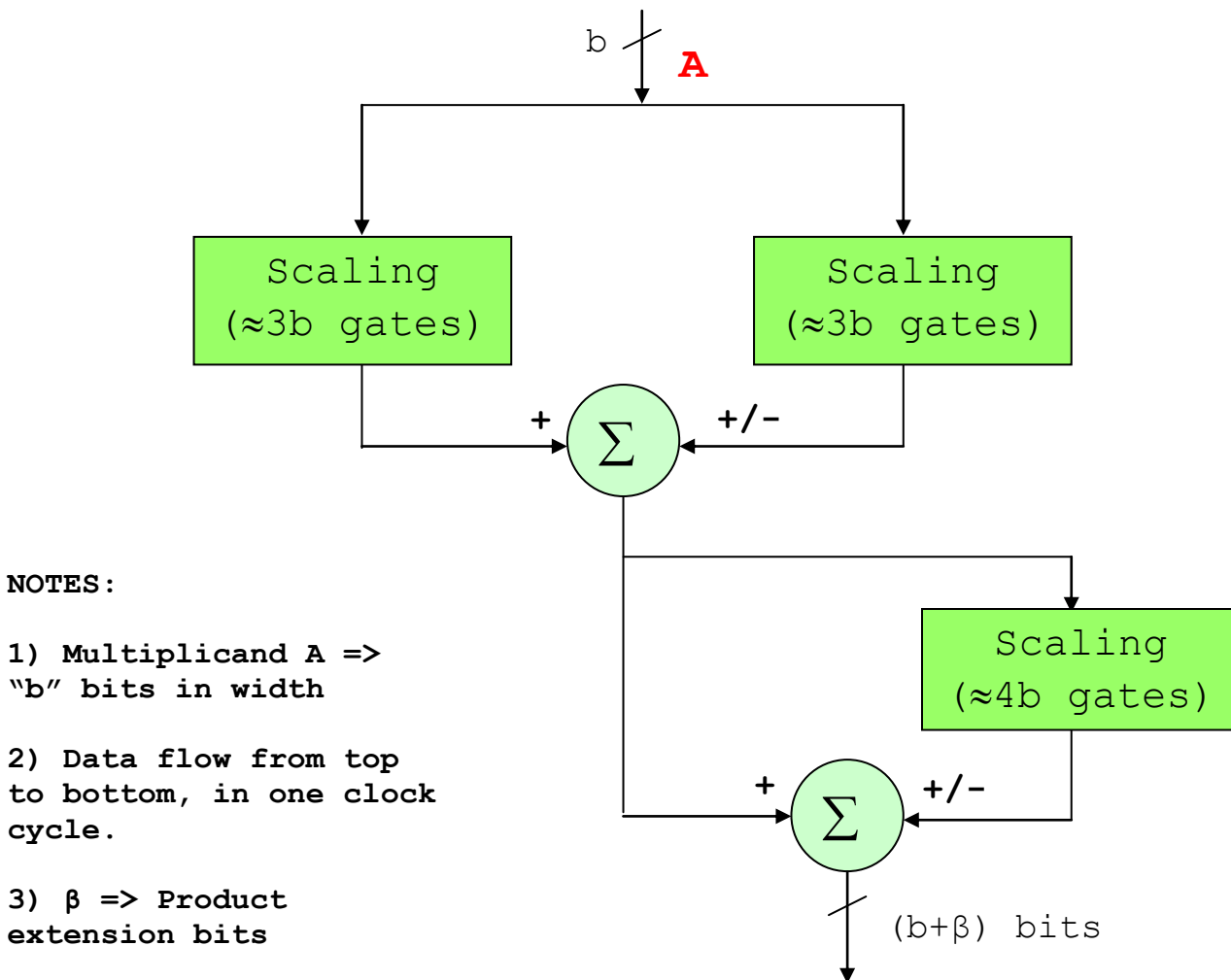
The TM resembles a "modular" computational building block. It presents a straightforward interface as shown in Figure 3 below.

```
 ┌─────────────────┐
 │ A = arbitrary   │
 │ input operand   │
 └─────────────────┘
          │
          ▼
     A                                   ┌─────────────┐
  ─────────►  ┌──────────────────┐       │ Argument in │
              │  Trigonometric   │       │ radians     │
              │   Multiplier     │       └─────────────┘
              │                  │              │
              │  (for fixed N)   │              ▼
  ─────────►  └──────────────────┘   ───►  A·cos(2πk/N)
     k                                  ┌─────────────────┐
                                        │ N is power of 2.│
                                        │ (16, 32, 64,---)│
                                        └─────────────────┘
```

TRIGONOMETRIC MULTIPLIER INTERFACES

FIGURE 3

Its simplistic structure lends the TM to high clock rates. Clock periods are typically limited by the propagation delay through one or two adders, and two multiplexers; as exemplified in Figure 4.

b   **A**

Scaling
(≈3b gates)

Scaling
(≈3b gates)

+   Σ   +/-

NOTES:

**1) Multiplicand A =>
"b" bits in width**

**2) Data flow from top
to bottom, in one clock
cycle.**

**3) β => Product
extension bits**

Scaling
(≈4b gates)

+   Σ   +/-

(b+β) bits

TRIGONOMETRIC MULTIPLIER DELAY PATH - EXAMPLE

FIGURE 4

FREQUENCY DOMAIN MODE - EXAMPLE (FFT)

Structurally, the TMs;

- Are characterized by simple interface and operation

- Consist of basic library elements (e.g. adders, multiplexers, registers)

- Can trade complexity for numeric precision (fixed point, 8 to 24 bits)

- Are synchronous in operation (requiring one or two clock cycles to produce most products within the Architecture structure)

The TM may be used to build efficient, small transforms (e.g., an N-point DFT performed with N^2 to $N^2/2$ addition operations, and no multiplies). Within the Architecture, such DFTs can implement semi-independent "processing Sections". These Sections can then be applied to construction of the Multi-Mode Architecture(s). Greater (or fewer) Sections can be employed in parallel to scale the Architecture complexity, and related performance (execution times).

The Rice Architecture has three operational Modes. Within a given Mode, various Functions may be executed. An example is the Frequency Domain Mode, wherein a large FFT Function may be executed. Such a function is critical for communications devices and infrastructure. In 5G radio systems, FFT sizes might approach 4K (4096) points to support advanced signal modulation schemes.

Figures 5 and 6 below relate to implementation of the Architecture with emphasis on a 4K point FFT transform. The Figures reflect lower and higher complexity implementations, respectively. As seen, the Architecture scales such that execution time (column 3) is inversely proportional to complexity (column 2). The Architecture can be ultimately scaled such that execution times approach 4096 clock cycles (or $\approx$ 8μsec execution time @ 500MHz clock speed).

The use of modular Sections resolves certain issues inherent in high-performance transform design, as related to memory structure and data movement. These are alleviated by the Sections' use of small, self-sufficient local memories (referenced in column 2 of the Figures). Also of note, is that the size of "Global" memory (column 4) remains constant as the Architecture is scaled.

Data Resolution	Approximate Complexity	Clock Cycles	"Global" Memory: Size / Accesses
16 bit input data f(n) 20 bit output data F(j)	2,000 gates +1024 bytes "local" RAM +256 bytes "local" ROM	≈ 2^17 (131,072)	4096 words RAM / 2^17 accesses
12 bit input data f(n) 16 bit output data F(j)	1,500 gates +1024 bytes "local" RAM +256 bytes "local" ROM	≈ 2^17 (131,072)	4096 words RAM / 2^17 accesses

Low-Complexity 4096 Point Real-Valued Transform

FIGURE 5

Data Resolution	Approximate Complexity	Clock Cycles	"Global" Memory: Size / Accesses
16 bit input data f(n) 20 bit output data F(j)	8,000 gates +4096 bytes "local" RAM +1024 bytes "local" ROM	≈ 2^15 (32,768)	4096 words RAM / 2^15 accesses
12 bit input data f(n) 16 bit output data F(j)	6,000 gates +4096 bytes "local" RAM +1024 bytes "local" ROM	≈ 2^15 (32,768)	4096 words RAM / 2^15 accesses

High-Complexity 4096 Point Real-Valued Transform

FIGURE 6

The execution times seen in Figures 5 and 6 can be further reduced by modifying the Section structure itself. This accelerates processing at the expense of greater complexity. For example, increased complexity of ≈ 50% essentially halves execution time. <u>This is achieved by increased complexity of individual Sections, as opposed to a greater number of Sections in parallel.</u>

Still, the FFT execution time (column 3) scales with the number of parallel Sections employed in the Architecture. Execution times remain inversely proportional to overall complexity (column 2). The Architecture can be scaled such that execution times decrease to nearly 2048 clock cycles (or ≈ 4μsec execution time @ 500MHz clock speed). This scaling relationship is expressed as follows;

T_E α (1/S)

Where,

T_E = Execution Time

S = Number of Sections Employed

Advantages of shorter execution time are greater throughput (or bandwidth), and decreased latency of data vectors "streaming" through the transform. Decreased latency can be a critical system-level consideration, especially in control systems.

Figures 7 and 8 below reflect the impact of Section modifications upon the transforms described by Figures 5 and 6, respectively. Differences in the modified (accelerated) transforms include;

- Reduction in clock cycles by factor of 2 (column 3 of Figures 7 and 8)

- Increased complexity of ≈ 50% (column 2 of Figures 7 and 8)

- Reduction by ≈ 2 bits of "best data resolution" (column 1 of the Figures)

- "Global" Memory size unchanged, but requiring faster access via "split memory" organization (allowing simultaneous dual access to 2 Memory halves). "Local RAM" memory requiring either faster access (cycle) times, or increased complexity.

Data Resolution	Approximate Complexity	Clock Cycles	"Global" Memory: Size / Accesses
14 bit input data f(n) 18 bit output data F(j)	3,000 gates +1024 bytes "local" RAM +256 bytes "local" ROM	≈ 2^16 (65,536)	4096 words RAM / 2^17 accesses
12 bit input data f(n) 16 bit output data F(j)	2,300 gates +1024 bytes "local" RAM +256 bytes "local" ROM	≈ 2^16 (65,536)	4096 words RAM / 2^17 accesses

2X Speed, Low-Complexity 4K Real-Valued Transform

FIGURE 7

Data Resolution	Approximate Complexity	Clock Cycles	"Global" Memory: Size / Accesses
14 bit input data f(n) 18 bit output data F(j)	12,000 gates +4096 bytes "local" RAM +1024 bytes "local" ROM	≈ 2^14 (16,384)	4096 words RAM / 2^15 accesses
12 bit input data f(n) 16 bit output data F(j)	9,000 gates +4096 bytes "local" RAM +1024 bytes "local" ROM	≈ 2^14 (16,384)	4096 words RAM / 2^15 accesses

2X Speed, High-Complexity 4K Real-Valued Transform

FIGURE 8

SUMMARY

Rice Electronics has created Intellectual Property (IP) for specialized digital processing Architectures. The Architectures employ novel building blocks in a hierarchical approach; facilitating rapid hardware development and unique complexity/ performance trade-offs. The Company's IP eliminates most conventional multiplication from the high-performance Architecture. This allows hardware implementations based primarily upon simple addition operations. The IP paradigms are independent of hardware target technologies.

The Rice Architectures may operate in multiple Modes, executing complex functions such as transforms, convolutions and correlations.

The Company's IP addresses size, cost and power constraints of many high-performance systems. Examples may include wide-bandwidth digital radio and real-time pattern recognition systems. For such systems, the IP enables specialized hardware circuits comparable in performance to traditional solutions, at a small fraction of the complexity. The IP described herein provides alternate system solutions to DSP processors, GPUs and numerous customized logic entities.

As discussed herein, the Company's digital processing IP base includes;

- Architecture IP

- Trigonometric Multiplier (TM) IP

- Modelling IP

NOTES:

This document contains preliminary information.

Some Intellectual Properties referenced in this document may have patents pending.

Contact:

Greg Rice

ricetronics@gmail.com

ABBREVIATIONS

ASIC	Application Specific Integrated Circuit
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
DSP	Digital Signal Processing (or Processor)
FFT	Fast Fourier Transform
5G	Fifth Generation
GPU	Graphics Processing Unit
IP	Intellectual Property
FPGA	Field Programmable Gate Array
OFDM	Orthogonal Frequency Domain Multiplexing
SoC	System-on-Chip

Filename: Processing IP Rev M