Dimension Reduction of Hyperspectral Data on Reconfigurable Computers

Jacqueline Le Moigne, *NASA/Goddard Code 935*
Pen-Shu Yeh, *NASA GSFC Code 564*
Joanna Joiner, *NASA GSFC, Code 910.3*
Gregory Donohoe, *University of Idaho*
Tarek El-Ghazawi, *George Washington University*
Abhishek Agarwal, *George Washington University*
Esam El-Araby, *George Washington University*
Wei Xia, *Global Science &Technology*

Description and Objectives

Investigate Use of Reconfigurable Computing for On-Board Automatic Processing of Remote Sensing Data

- Choice of Basic Methodologies to be Performed On-Board

- Dimension Reduction and Cloud Detection:
  > reduce communication bandwidth
  > make subsequent computations simpler, faster and more accurate.

- Computation Engine: RDPP (under ESTO/ATIP &AIST), now FPPA ("Field Programmable Processor Array")

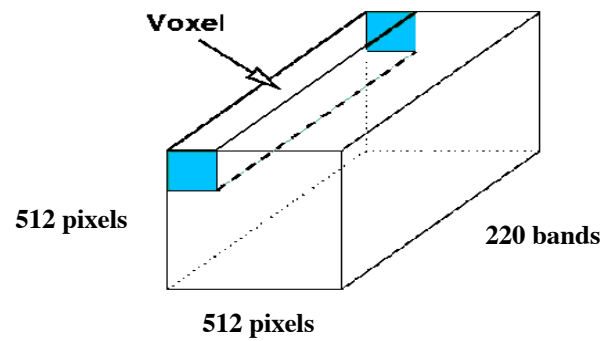
- Preliminary Feasibility Studies on Hybrid Reconfigurable Architecture, the SRC-6E

- Results Available for Designing Future Missions

- Building Capabilities of Future Formation Flying Systems

Wavelet-Based Dimension Reduction

- **Dimension Reduction: Application-Oriented Compression**
  - Large Volumes of Hyperspectral Data
  - High Computational Cost



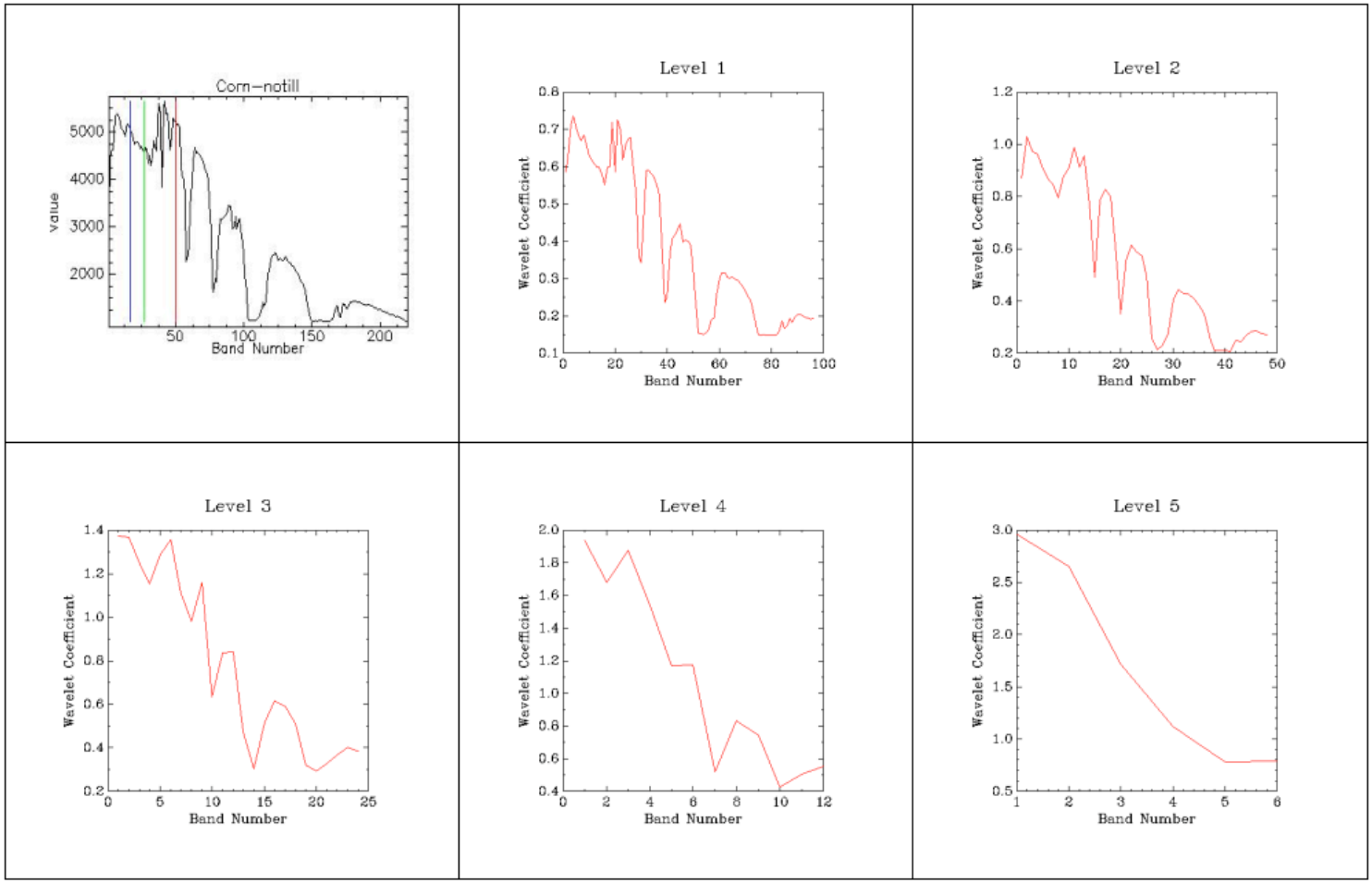
- **Principal Component Analysis (PCA):**
  - Most Common Method Dimension Reduction
  - Does Not Preserve Spectral Signatures
  - PCA Complex and Global computations: difficult for parallel processing and hardware implementations

- **Applications:**
  - Classification of Hyperspectral Data: AVIRIS (220 Bands)
  - Analysis of High Spectral Resolution Data: AIRS (2378 Bands)

- **Wavelet-Based Dimension Reduction:**
  - Preserves Spectral Signatures
  - High-Performance Implementation
  - Simple and Local Operations

Wavelet-Based Dimension Reduction (2)

- Multi-Resolution Wavelet Decomposition of Each Pixel 1-D Spectral Signature



Wavelet-Based Dimension Reduction (3)

(1) Multi-Resolution Wavelet Decomposition of Each Pixel 1-D Spectral Signature

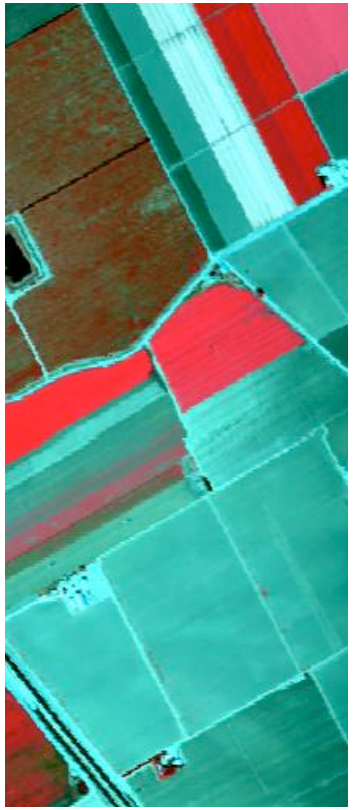
(2) Reconstruction Using only Low-Pass Information

(3) Similarity Measure (e.g., Correlation) between Original Signature and Reconstructed Signature for each Decomposition Level

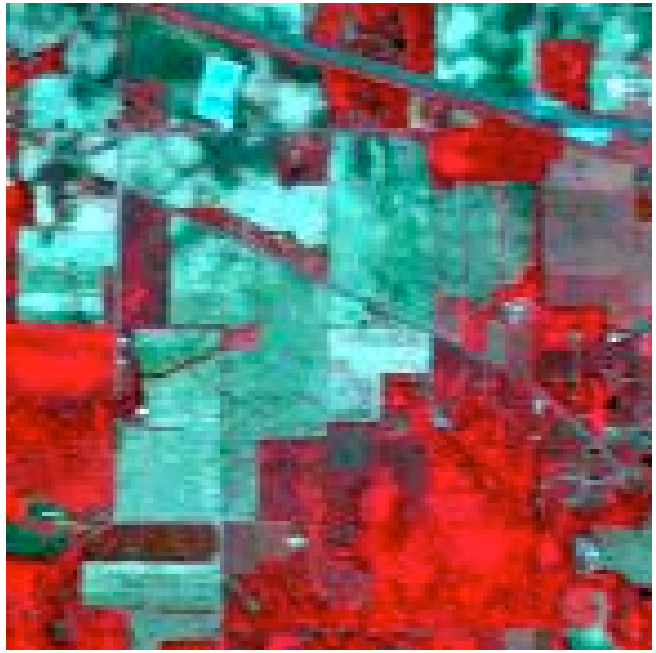
(4) Record Level that Satisfies Quality of "Good" (Percentage Threshold-Defined) Reconstruction ==> Histogram Building

(5) From Histogram, Choose Optimum Level of Decomposition and Build Corresponding Dimension Reduced Image

Wavelet-Based Dimension Reduction (4)



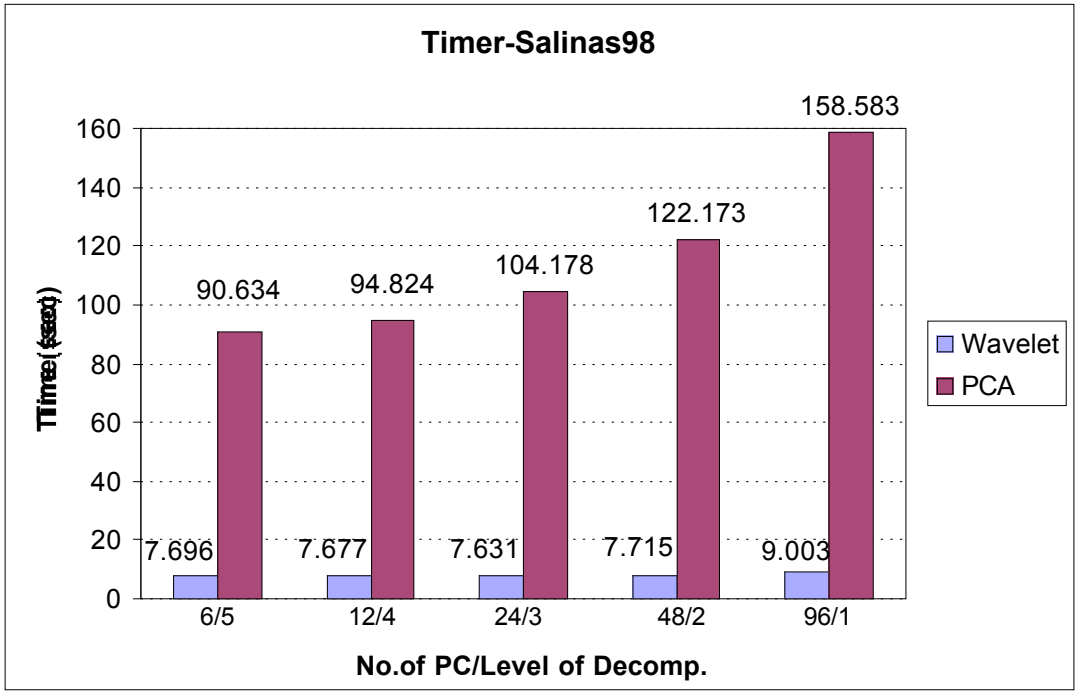
AVIRIS: SALINAS'98



AVIRIS: INDIAN PINES'92

Wavelet-Based Dimension Reduction (5)

Complexity: Wavelet-Based = O(MN) ; PCA = O(MN^2+N^3)



WAVELET

	No.of PC/Level of Wavelet Decomposition				
	6/5	**12/4**	**24/3**	**48/2**	**96/1**
Timer GLOBAL	7.696	7.677	7.631	7.715	9.003
IO_R	0.406	0.412	0.411	0.412	0.41
Comp.	7.253	7.19	7.069	6.692	7.939
IO_W	0.037	0.075	0.151	0.311	0.654

PCA

	No.of PC/Level of Wavelet Decomposition				
	6/5	**12/4**	**24/3**	**48/2**	**96/1**
Timer GLOBAL	90.634	94.824	104.178	122.173	158.583
IO_R	0.423	0.395	0.395	0.394	0.394
Comp.	90.173	94.355	103.633	121.478	157.568
IO_W	0.038	0.074	0.15	0.301	0.621

(Timings on 500 MHz Pentium III)

Wavelet-Based Dimension Reduction (6)

Classification Accuracy - Salinas'98

Classification Method	Reduction Method	Classification Accuracy				
		No. of Component/Level of Decomposition				
		6/5	12/4	24/3	48/2	96/1
Maximum Likelihood	*PCA*	97.2278	97.9587	98.3321	98.2313	40.3501
	Wavelet	97.8968	98.5979	98.6208	98.5681	40.3501
Minimum Distance	*PCA*	93.7844	93.8096	93.821	93.821	93.8187
	Wavelet	93.5736	93.6927	93.8086	93.821	93.821
Parallelepiped	*PCA*	81.7678	82.2054	80.2053	74.638	64.8277
	Wavelet	76.5488	76.4915	76.2326	74.8213	73.6689

Classification Accuracy - IndianPines'92

Classification Method	Reduction Method	Classification Accuracy				
		No. of Component / Level of Decomposition				
		6/5	12/4	24/3	48/2	96/1
Maximum Likelihood	*PCA*	72.6214	73.0496	72.2066	70.186	15.3486
	Wavelet	75.1372	78.429	82.4435	81.5335	15.3486
Minimum Distance	*PCA*	42.3659	42.6602	42.968	43.0483	43.142
	Wavelet	41.1214	41.9912	42.2722	42.3792	42.3926
Parallelepiped	*PCA*	36.9865	37.2541	36.8259	35.8624	31.1521
	Wavelet	32.5037	32.7579	33.7883	33.4404	33.0925

- Implemented on the HIVE (8 Pentium Xeon/Beowulfs-Type System): 6.5 times faster than sequential implementation
- Classification Accuracy Similar or Better Than PCA
- Faster than PCA

AIRS Data Experiments

Newly Acquired Dataset - AIRS Data

- AIRS = high spectral resolution spectrometer
  - **2378 bands in the thermal infrared (3.7-4.61 and 6.2-8.22 and 8.8-15.4 µm)**
  - **4 bands in the visible (0.4 - 1.0 µm).**
- Ranges selected for determination of atmospheric temperature with accuracy of 1°C in layers 1 km thick, and humidity with an accuracy of 20% in layers 2 km thick in the troposphere.
- Coverage:
  - **Global,**
  - **Twice Daily (Daytime and Nighttime)**
- Data produced by AIRS/AMSU-A/HSB system on Earth Observing System (EOS) polar orbiting platform(EOS Aqua).
- Level1B radiance data.
- Resolution:
  - **13.5 km at nadir 41 km x 21.4 km at the scan extremes**
- Volume: **122.1 MB / file (244 granules per day)**

AIRS Data Experiments (2)

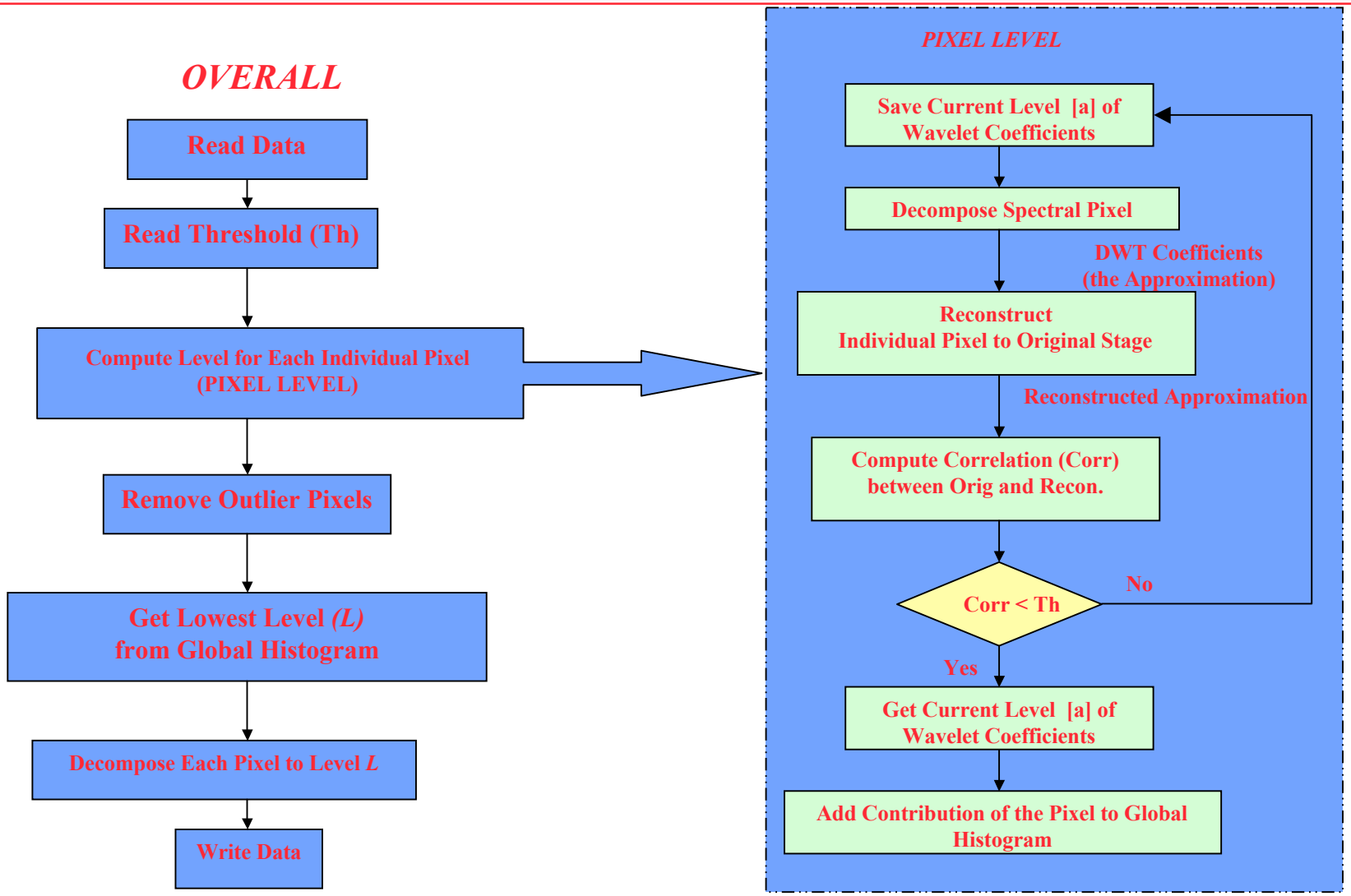
- **Currently, Scientists use 281 channels "appropriately" selected at NOAA/NESDIS**

- **Channels arranged in 3 major intervals (and 17 sub-intervals):**
  - 3.74 - 4.61 micron, channels 1865-2378
  - 6.20 - 8.22 micron, channels 1263-1864
  - 8.80 - 15.4 micron, channels 1-1262

- **Investigation:**
  - Adapt PCA and Wavelet-Based Dimension Reduction for 3 or 17 intervals.
  - Experiments show that if 17 intervals considered and a threshold of 92%, some of the data can be reduced by using 4 levels of decomposition

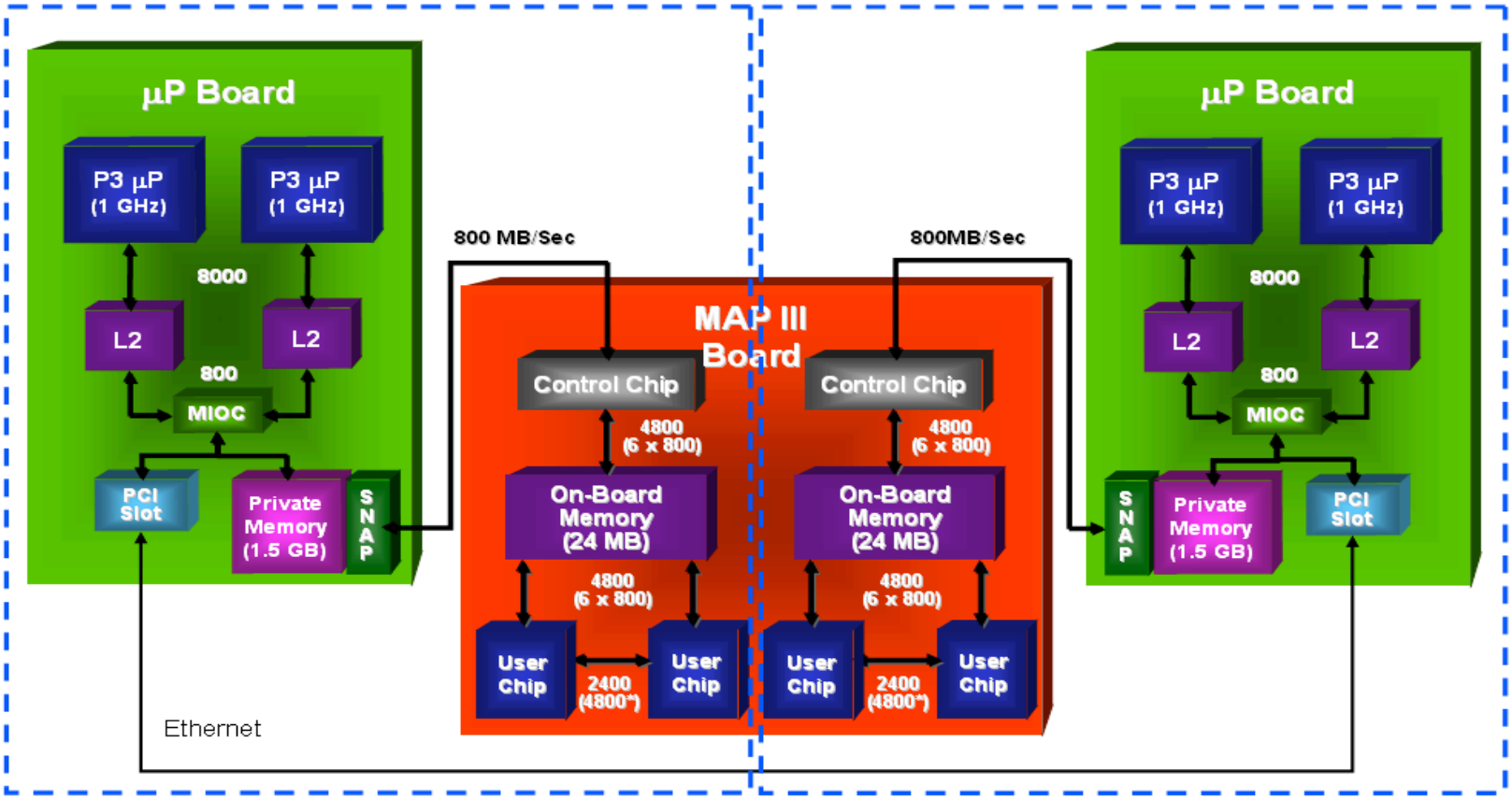
- **Questions/Issues:**
  - Anomalies in some channels ("popping")
  - Consider specific characteristics of selected channels
  - Validation procedure

The Algorithm

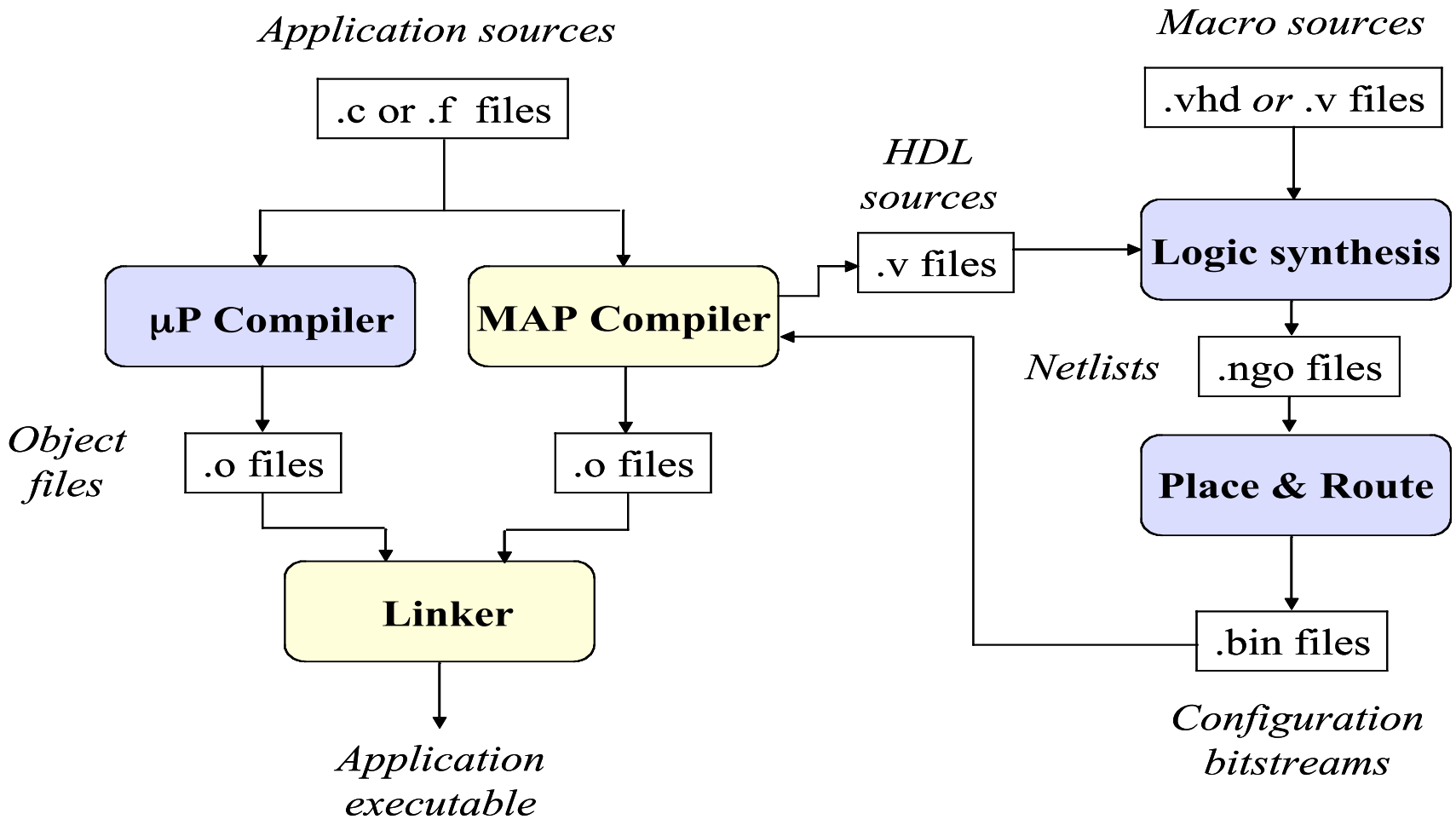


Prototyping Wavelet-Based Dimension Reduction of Hyperspectral Imagery on a Reconfigurable Computer, the SRC-6E

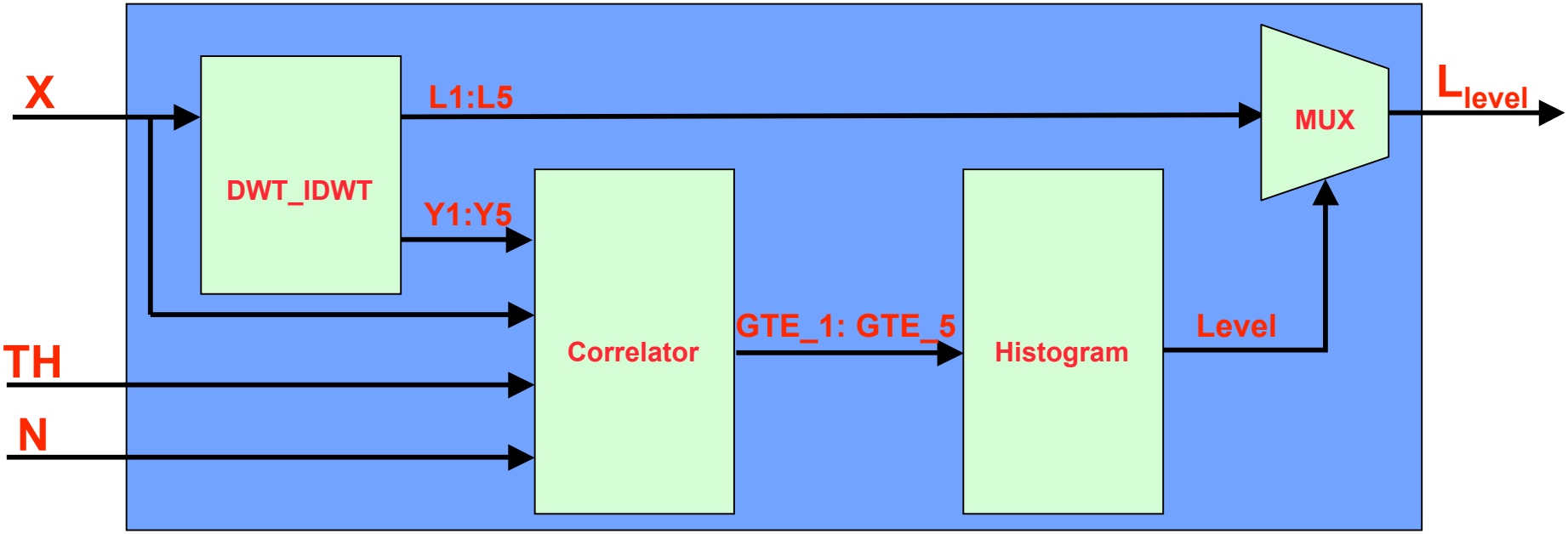
Hardware Architecture of SRC-6E



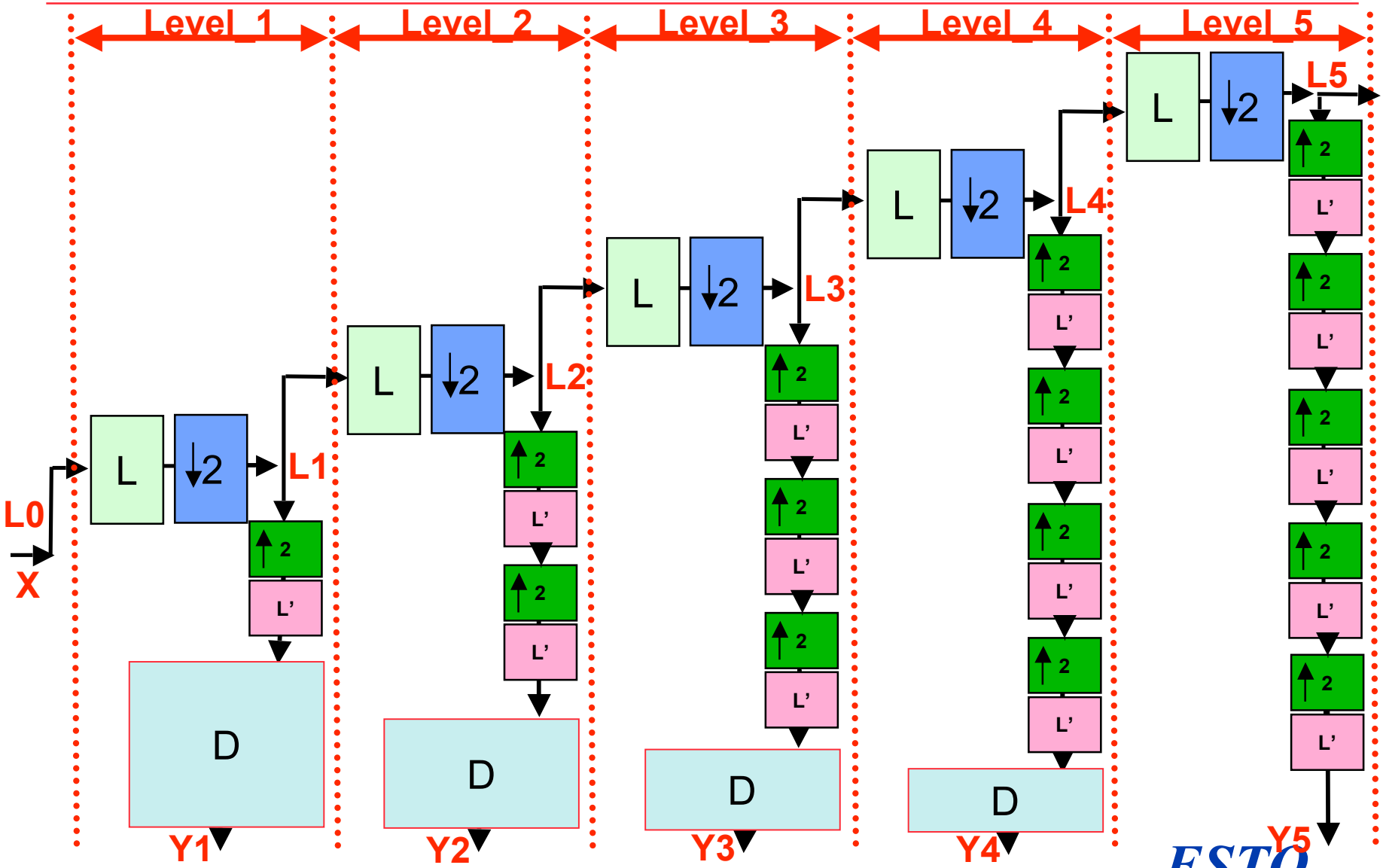
SRC Compilation Process



Top Hierarchy Module

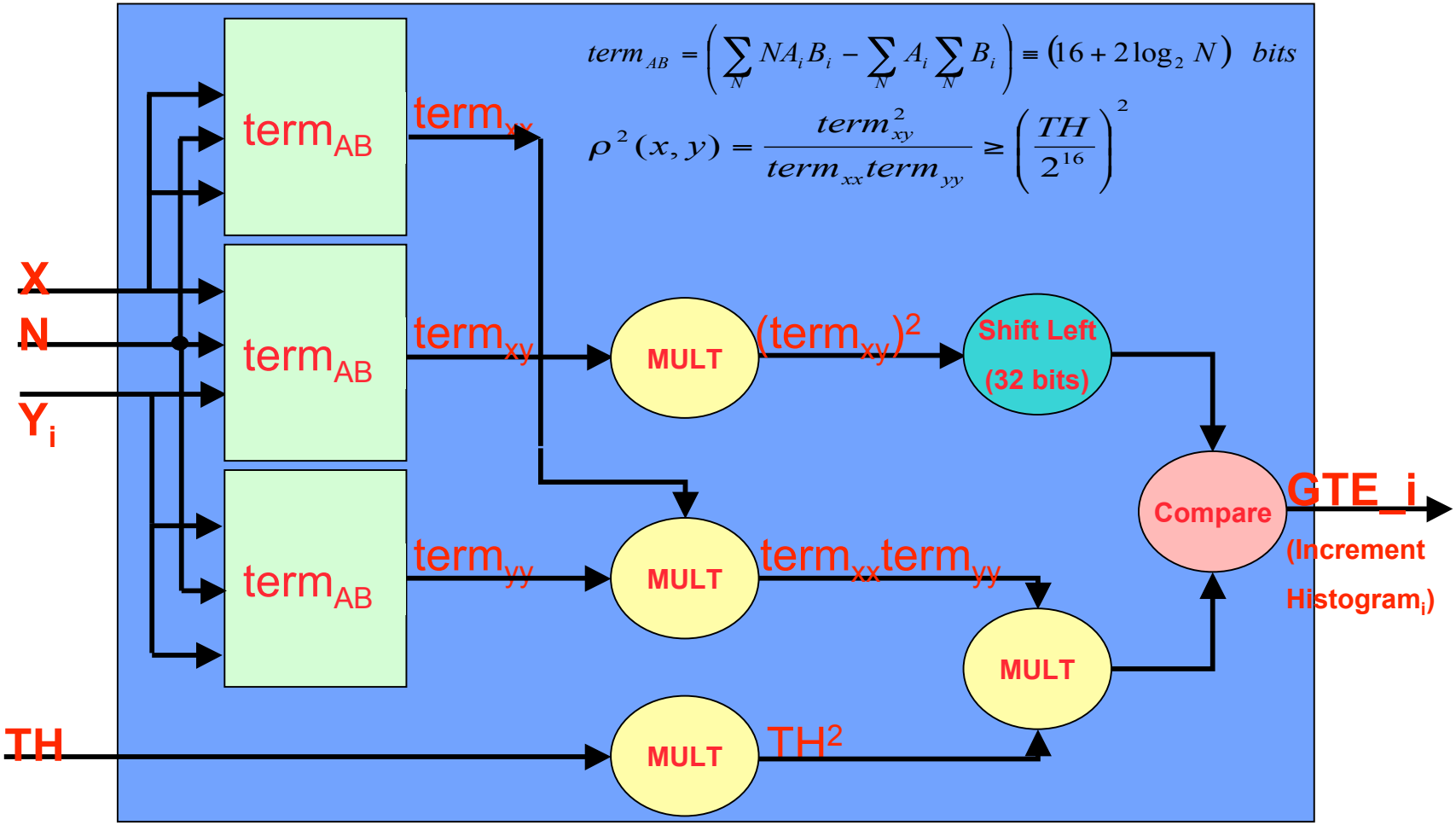


Decomposition and Reconstruction Levels of Dimension Reduction (DWT_IDWT)

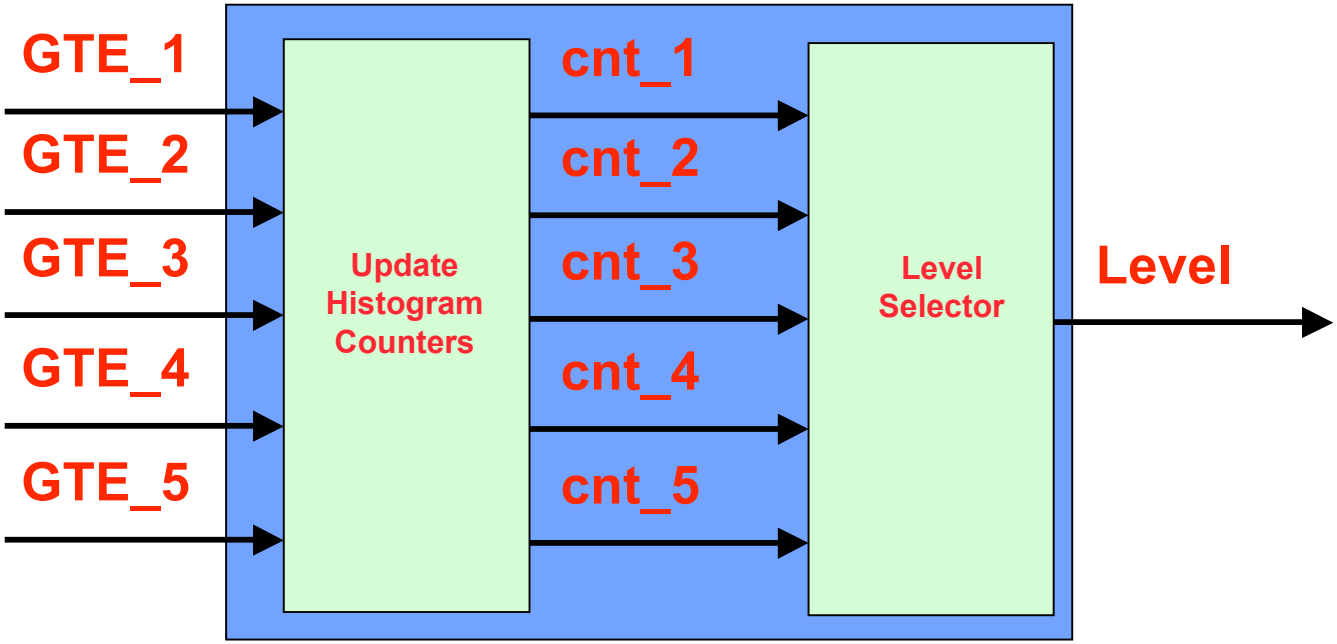


ESTO
Earth Science Technology Office

Correlator Module



Histogram Module

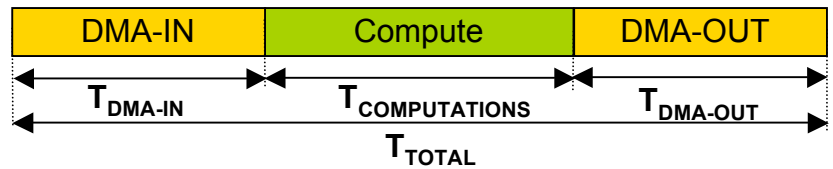


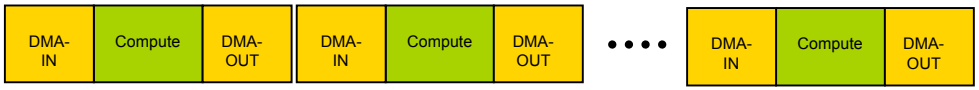
Resource Utilization and Operating Frequency

Slice Utilization (%)		Total 4-Input LUTs Utilization (%)				MULT18X18 (%)	Frequency (MHz)
Flip Flops	Total	Logic	Route-Throu	Shift-Registers	Total		
58	65	43.22	0.61	8.20	52.03	81	100

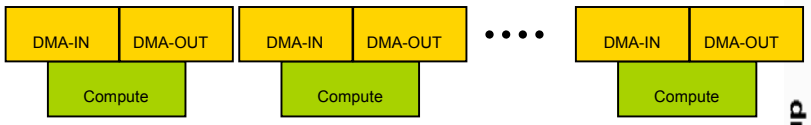
SRC Experiment Setup and Results

- **Salinas'98**
  - 217 X 512 Pixels, 192 Bands = 162.75 MB
  - Number of Streams = 41
  - Stream Size = 2730 voxels ≈ 4 MB

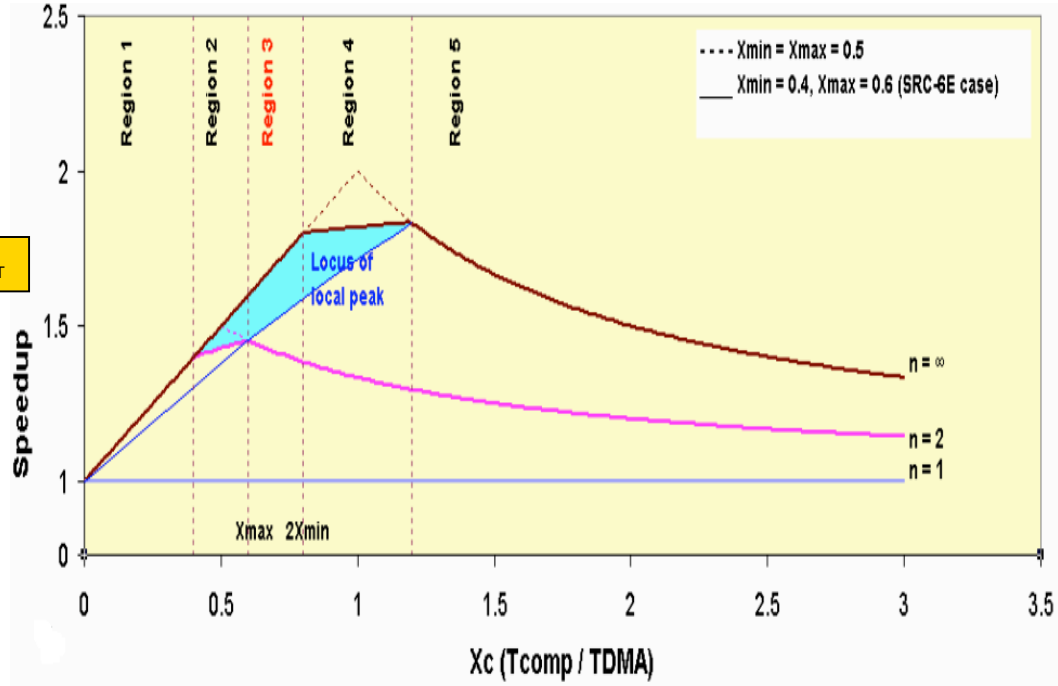




- **Non-Overlapped Streams**
  - T_{DMA-IN} = 13.040 msec
  - T_{COMP} = 0.62428 msec
  - $T_{DMA-OUT}$ = 22.712 msec
  - T_{Total} = 1.49 sec
  - Throughput = 109.23 MB/Sec

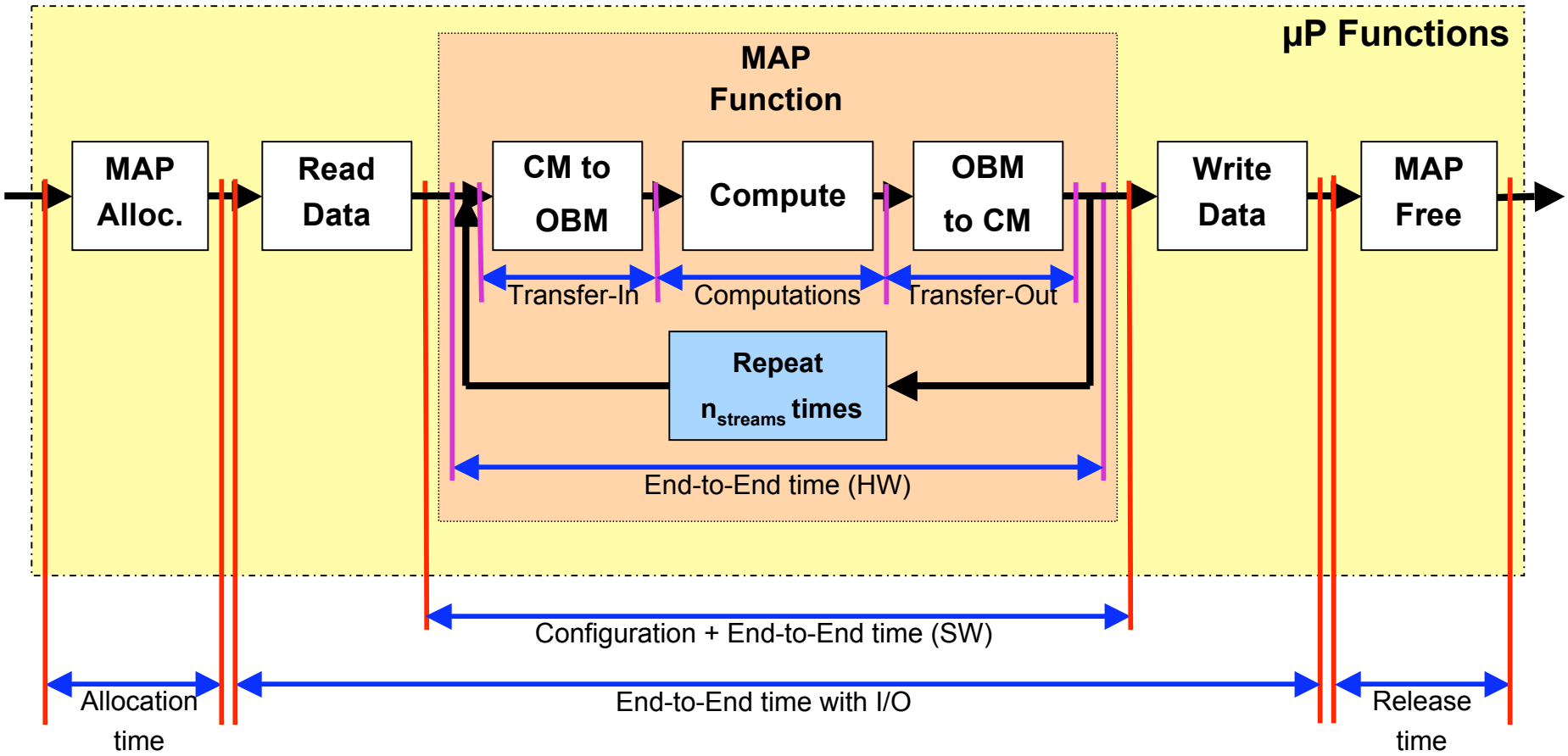


- **Overlapped Streams**
  - T_{DMA} = 35.752 msec
  - T_{COMP} = 0.62428 msec
  - X_c = 0.0175
  - Throughput = 111.14 MB/Sec

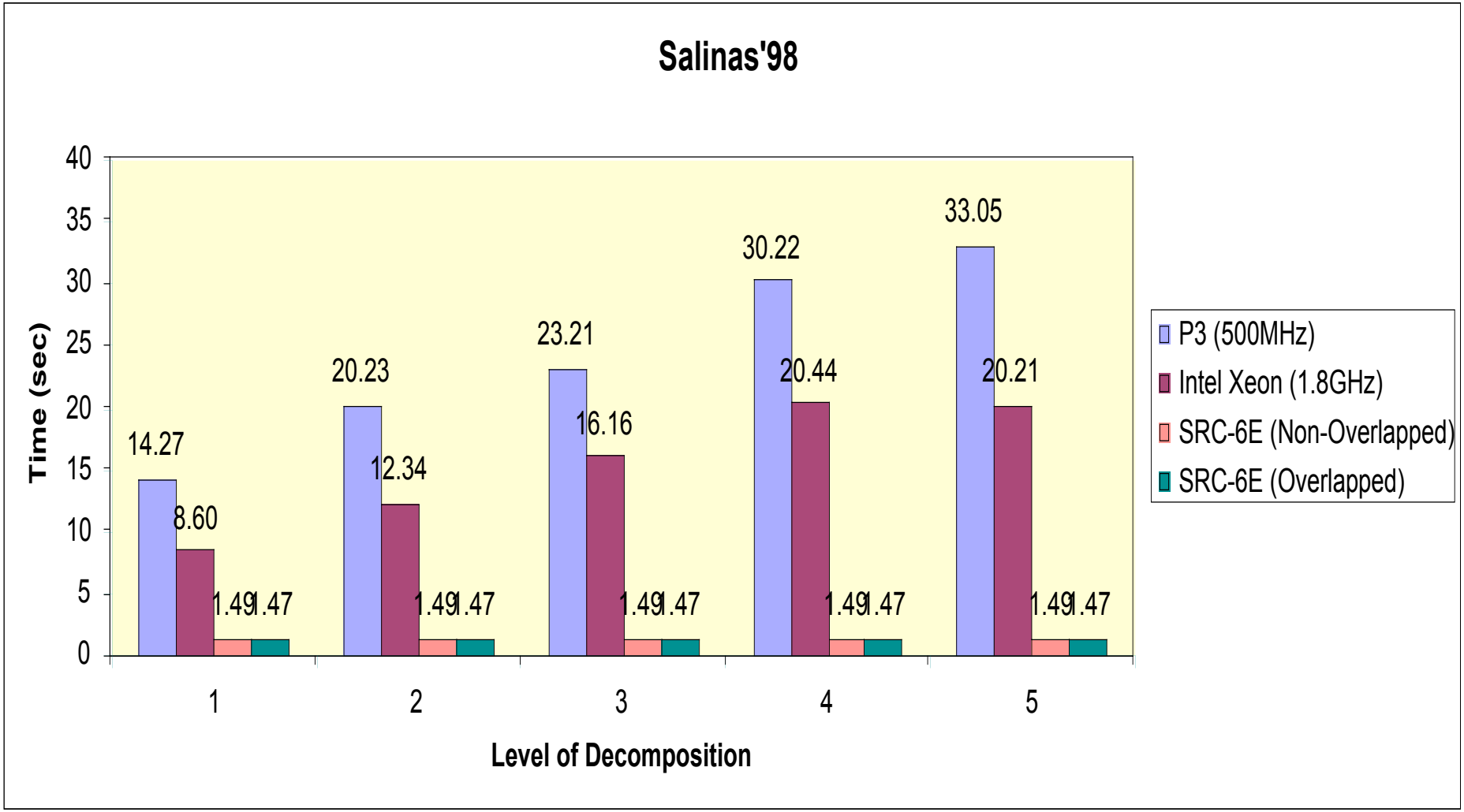


- **$Speedup_{non\text{-}overlapped}$ = (1+ X_c) = 1.0175 (insignificant)**

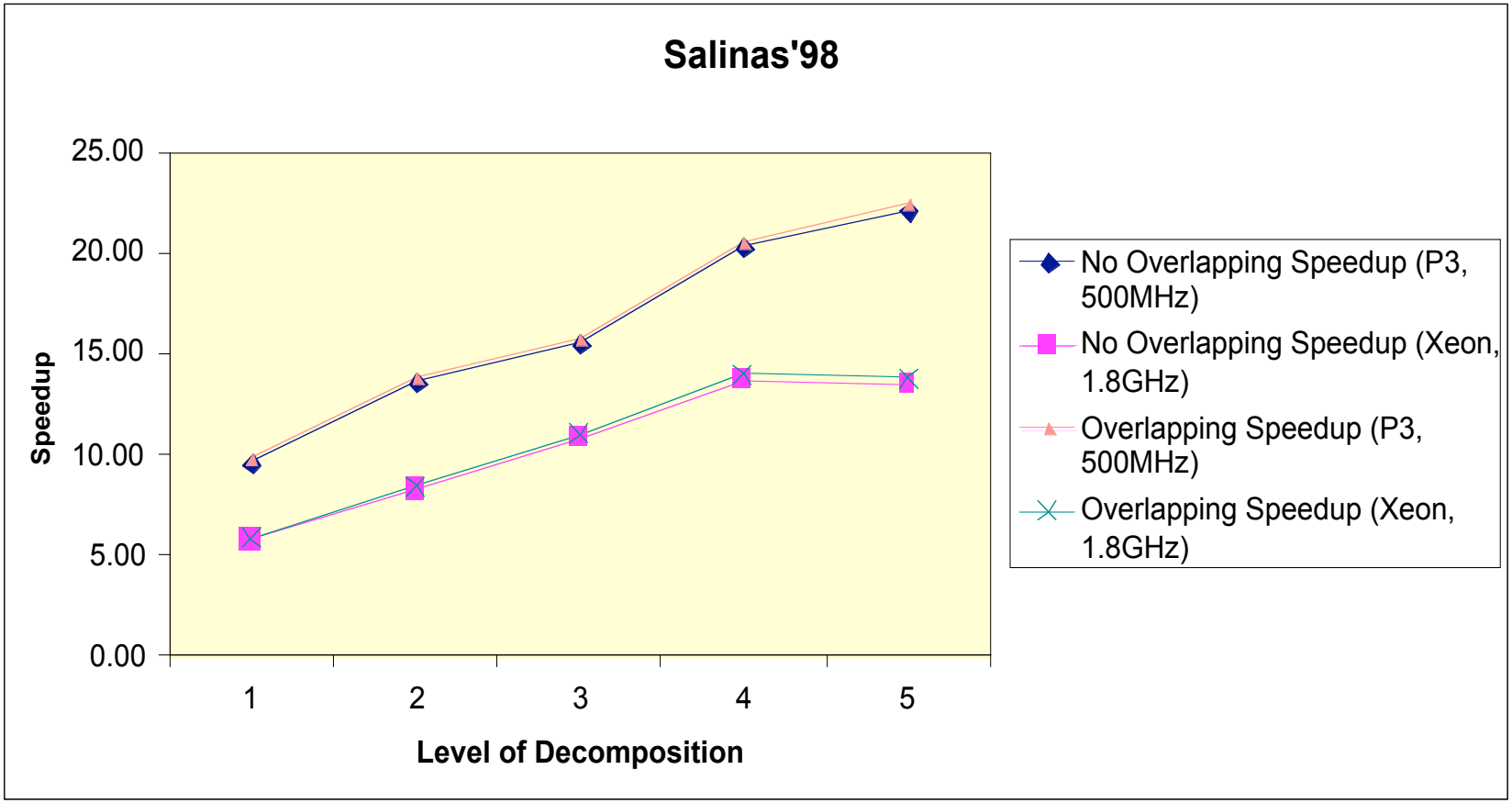
Measurements Scenarios



Execution Time



Speedup Results



Summary Prototyping on the SRC-6E

- We prototyped the automatic wavelet-based dimension reduction algorithm on a reconfigurable architecture
- Both coarse-grain and fine-grain parallelism are exploited
- We observed a 10x speedup using the P3 version of SRC-6E. From our previous experience we expect this speedup to double using the P4 version of SRC machine
- These speedup figures were obtained while I/O is still dominating. The speedup can be increased by improving I/O

Designing a Reconfigurable Processor for On-Board Dimension Reduction

On-Board Wavelet-Based Dimension Reduction

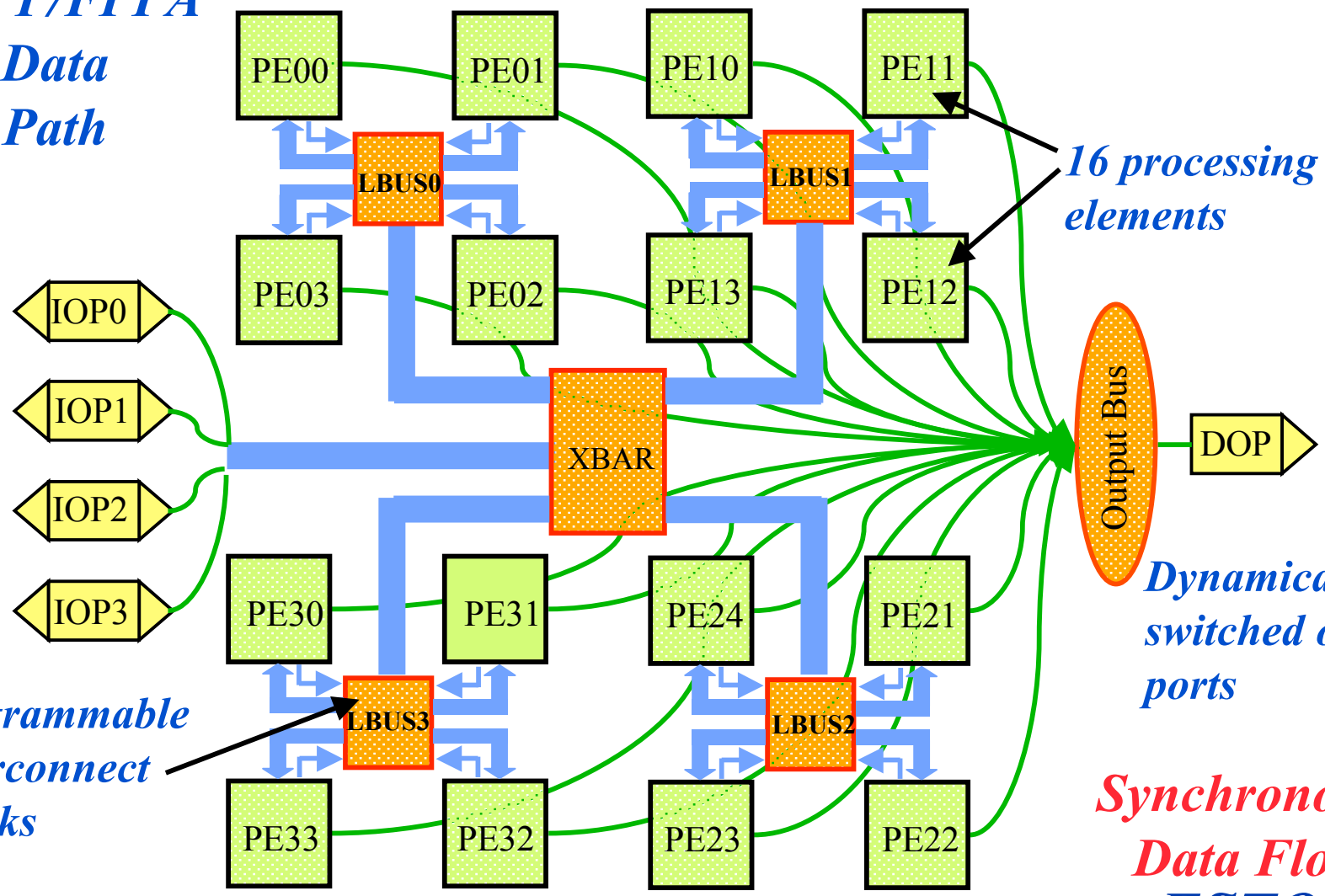
GOAL: *Design a Reconfigurable subsystem to perform a wide range of spacecraft instrument data computations at very high throughput and low power consumption*

- **On-Board Reconfigurable Computing Design Issues**
  - Low Power Consumption
  - Radiation-Hard Capability

- **Previously ESTO funded Studies**
  - Reconfigurable Data Path Processor (RDPP)
  - Demonstrated:
    - o Focal plane array readout correction
    - o Fourier Transform Hyperspectral Imager Data Conversion
    - o Digital Filtering, Convolution
  - Simulink-Based Graphical Programmer Interface
    - o Simplifies RDPP Platform Programming
    - o Supports Multiple RDPP Chips

RDPP and New Field Programmable Processor Array (FPPA) Architectures

RDPP/FPPA Data Path



16 processing elements

Dynamically-switched output ports

Programmable Interconnect Blocks

Synchronous Data Flow

ESTO

Earth Science Technology Office

FPPA Demonstration

- **4-Level Wavelet Decomposition**
  - Daubechies DAUB-4 wavelet basis set
  - Requires one FPPA chip
  - No external chips required
  - One input sample per clock cycle
    - About 50 Megasamples/second
  - Typical dimensionality-reduction operation

- **Wavelet-Based Dimension Reduction**
  - Will be implemented on 16 FPPAs (total of 256 PEs)

Conclusions and Future Work

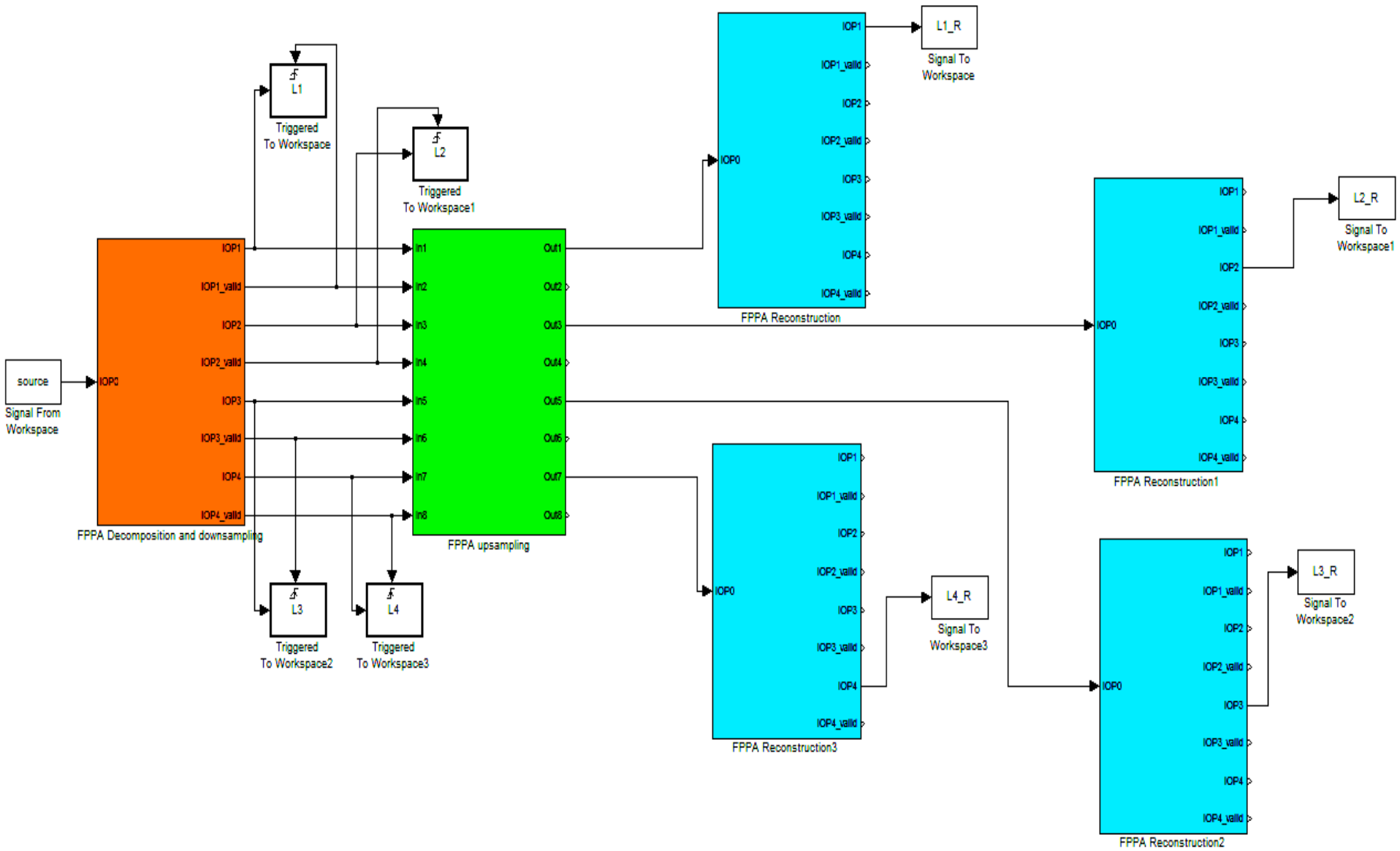
- **Accomplished**
  - **Wavelet-Based Dimension Reduction Demonstrated on AVIRIS 220-Bands Data and Being Generalized on AIRS 2378-Bands Data**
  - **Demonstrated Wavelet Decomposition on FPPA Simulator**
  - **Wavelet-Based Data Reduction Demonstrated on Hybrid SRC-6E Reconfigurable Architecture**
- **Future Work**
  - **Design Necessary FPPA Software Modules and Testbed Architecture**
  - **Complete Testing of AIRS Data Sequentially, on SRC-6E and on FPPA**

APPENDIX

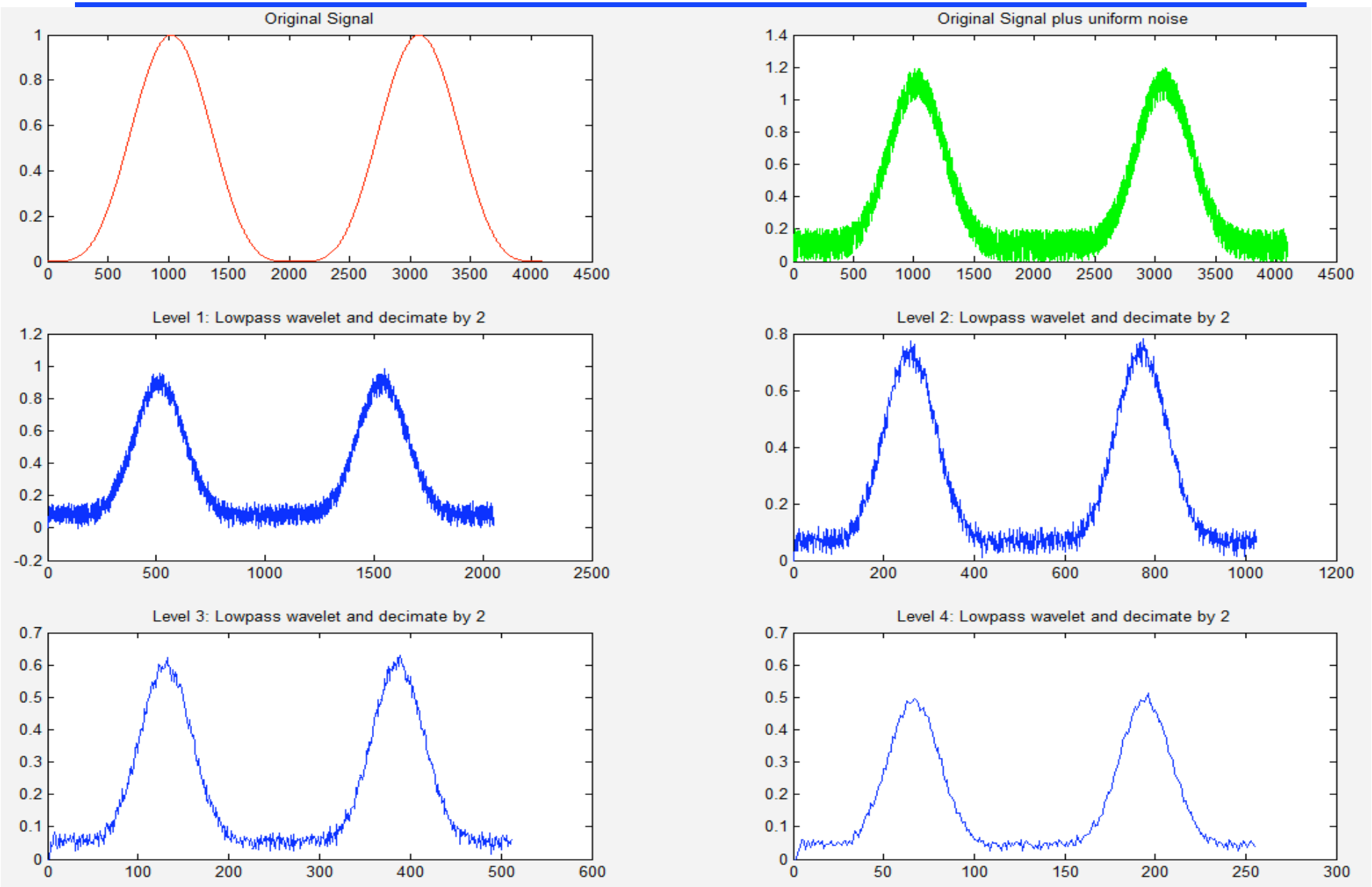
Update Histogram Counters

GTE_1	GTE_2	GTE_3	GTE_4	GTE_5	count_en_1	count_en_2	count_en_3	count_en_4	count_en_5
0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	1	0	0	0	0
1	1	0	0	0	0	1	0	0	0
1	1	1	0	0	0	0	1	0	0
1	1	1	1	0	0	0	0	1	0
1	1	1	1	1	0	0	0	0	1

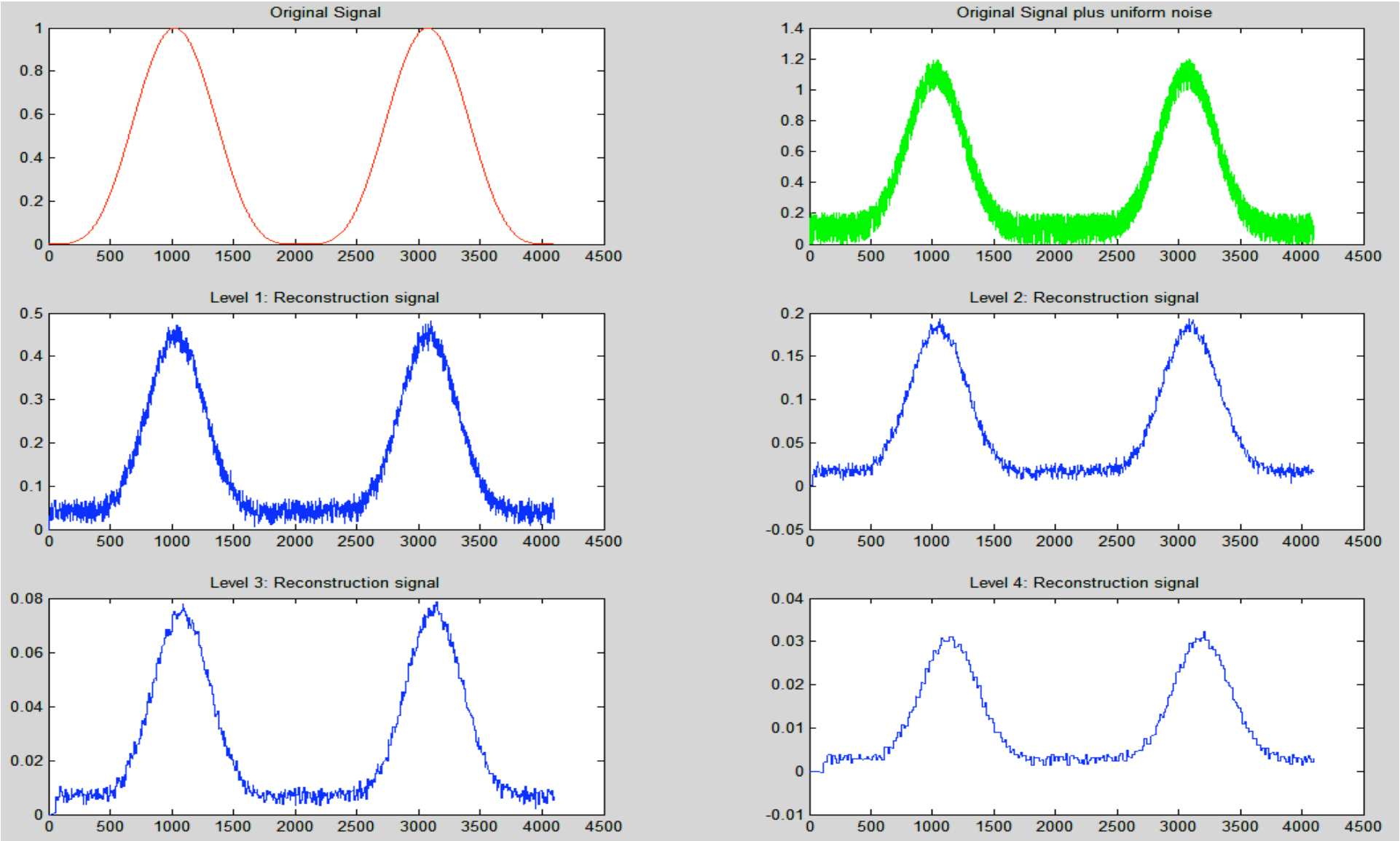
Simulink Model



Signals at Each of the Four Levels



Signals at each of the reconstruction four levels



Current Simulink PEs model capabilities

- **Configuration mode**
  - **8 basic PEs functional block**
  - **1 unconditional PE block**
  - **1 generalize conditional PE (NO RESTRICTION)**
  - **1 exact model of the conditional PE**

- **Run Time model**
  - **User define Boolean fire pattern through a GUI interface exist on every PE model.**

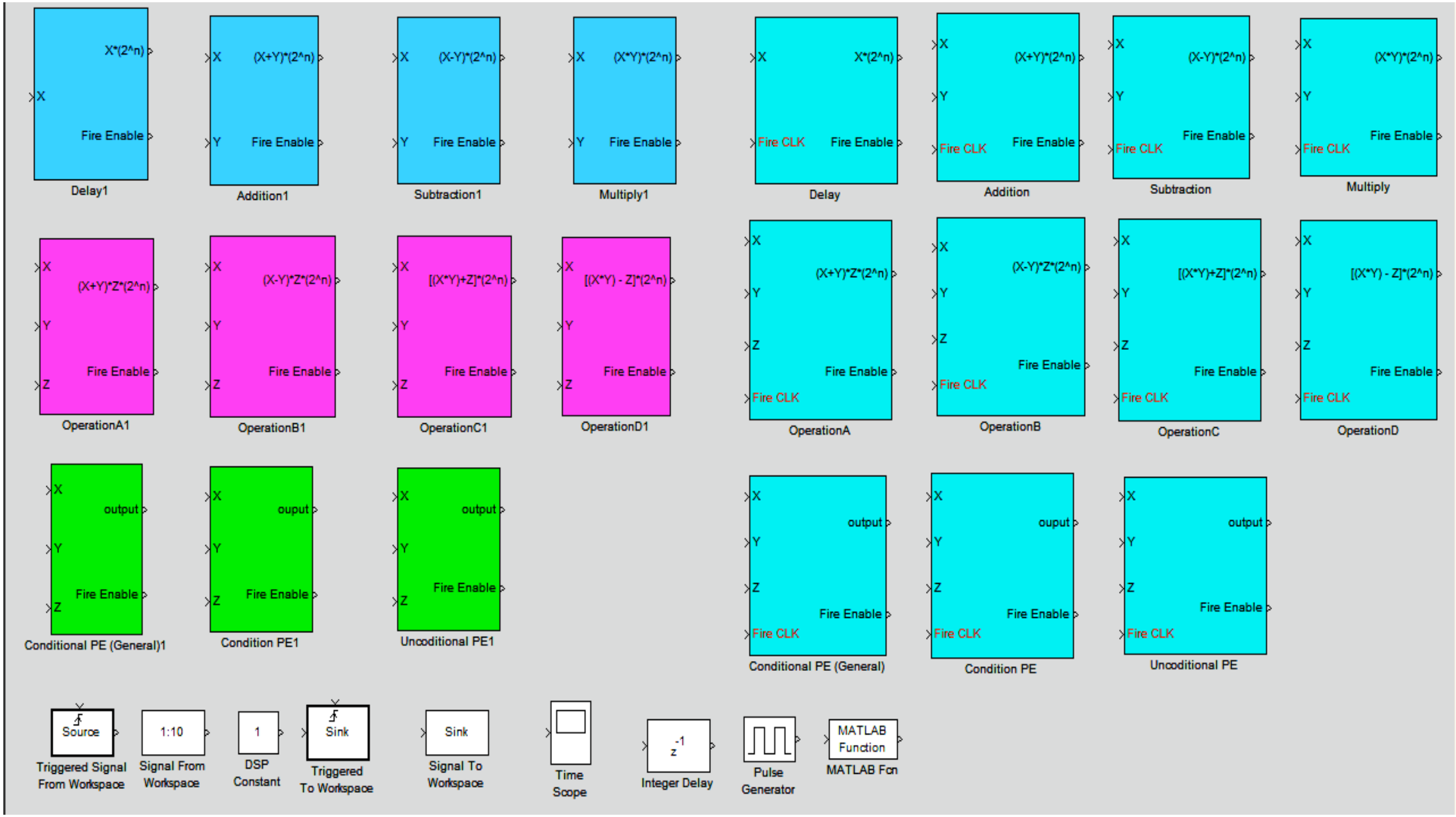



Image of the Current Simulink FPPA PE model library