The Systolic Ring: A Scalable Dynamically Reconfigurable Core for Embedded Systems

Pascal Benoit, Gilles Sassatelli, Michel Robert, Lionel Torres, Gaston Cambon, Thierry Gil

Application example - DCT 2D 8*8

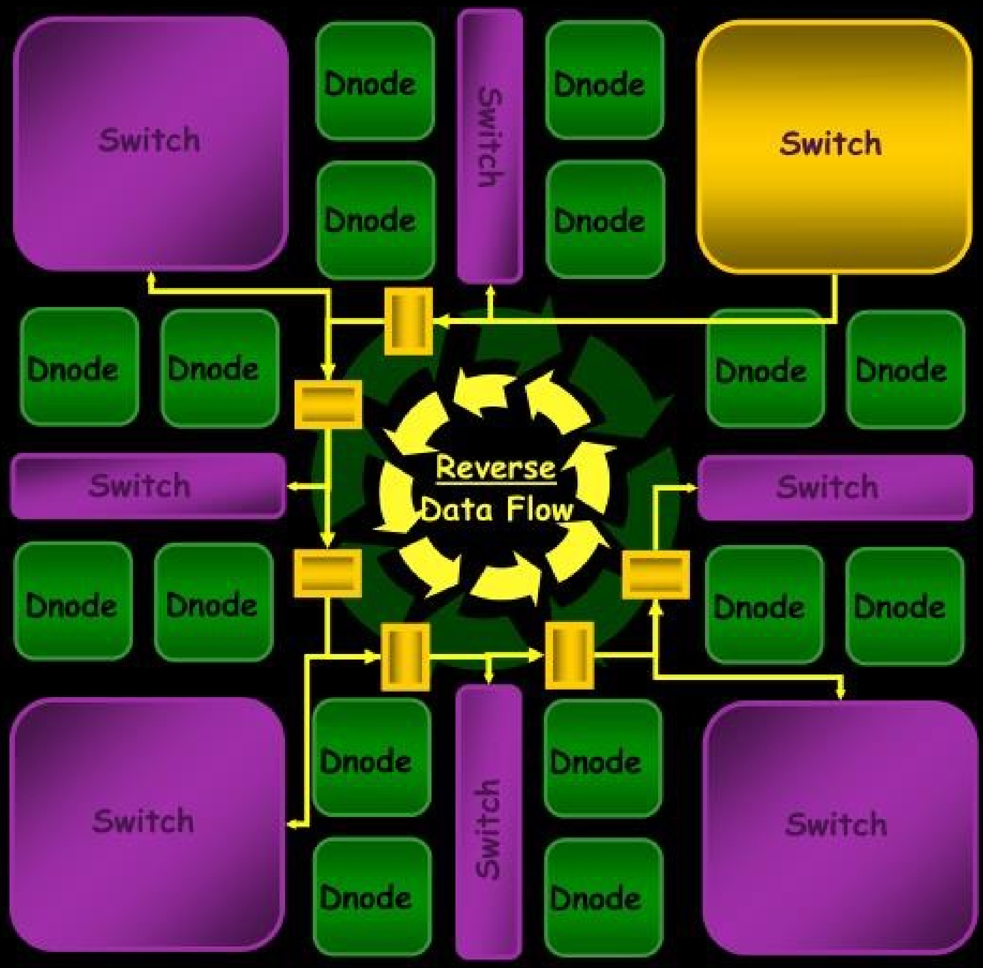
64*64 image example - Comparisons

	Pentium IV Intel	DCT Core Xilinx	TMS320C62 TI	RING-16	RING-64
Cycles #	21248	**4171**	10240	**5120**	**1280**
f (MHz)	1200	80*	300	200	200
Proc. Time (µs)	**17.7**	52.1	**34.1**	**12.8**	**6.4**
Comment	SSE2	*Device dependant	Matrix	Even-Odd decomposition	
Type	Super scalar	Fine Grain Reconfigurable	VLIW	Coarse Grain Reconfigurable	

Only Processing time !!

The Systolic Ring - Operative Layer Topology

Data Flows



Forward Data Flow

Unidirectional data transit between

successive layers (circular pipeline

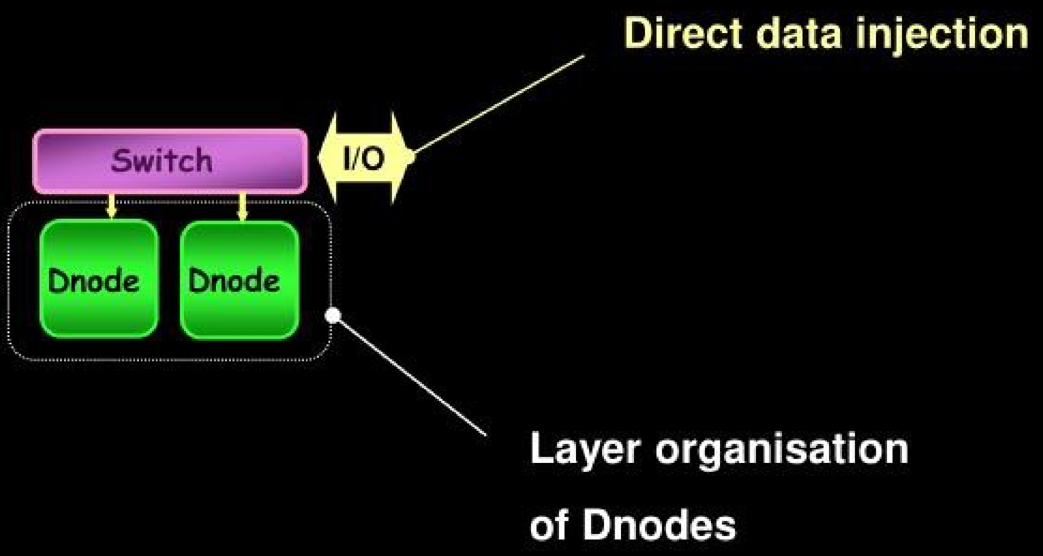
Reverse Data Flow

Feedback pipeline network

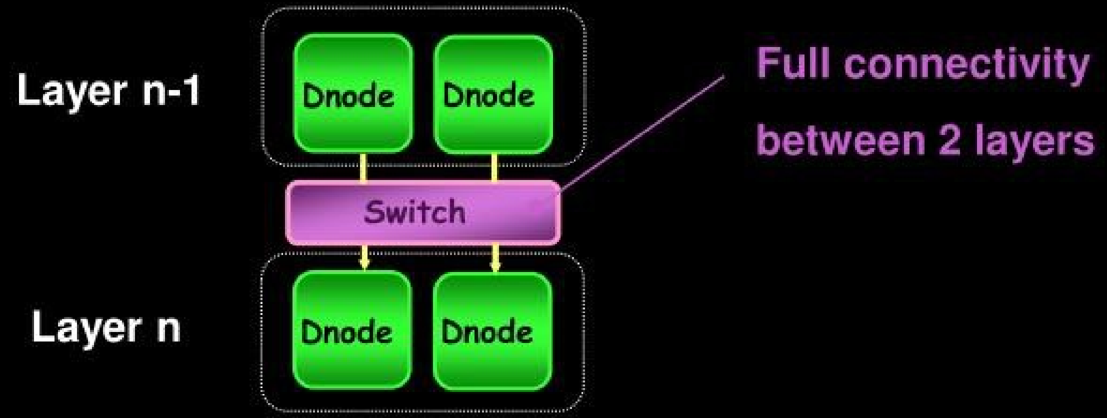
for recursive algorithms

The Systolic Ring - Dnode Clusters

Macro Node

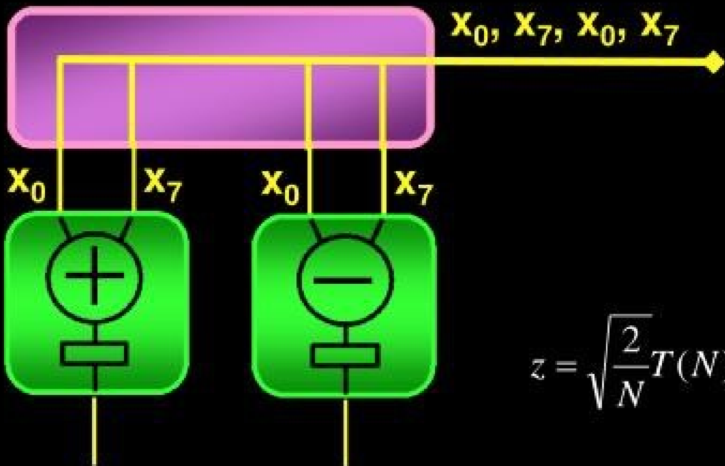


The Systolic Ring - Switch components



Thank You

Application example - 8*8 2D DCT



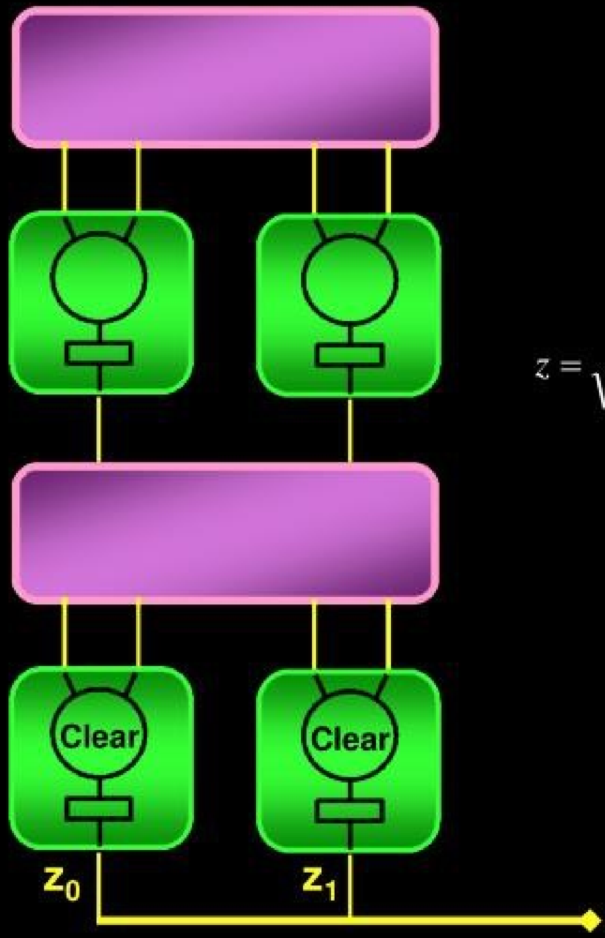
Cycle 0

$$z = \sqrt{\frac{2}{N}}T(N)x \quad \begin{cases} \begin{bmatrix} z_0 \\ z_2 \\ z_4 \\ z_6 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} \\ \beta & \delta & -\delta & -\beta \\ \alpha & -\alpha & -\alpha & \alpha \\ \delta & -\beta & \beta & -\delta \end{bmatrix} \begin{bmatrix} x_0 + x_7 \\ x_1 + x_6 \\ x_2 + x_5 \\ x_3 + x_4 \end{bmatrix} \\ \begin{bmatrix} z_1 \\ z_3 \\ z_5 \\ z_7 \end{bmatrix} = \begin{bmatrix} \lambda & \gamma & \mu & \nu \\ \gamma & -\nu & -\lambda & -\mu \\ \mu & -\lambda & \nu & \gamma \\ \nu & -\mu & \gamma & -\lambda \end{bmatrix} \begin{bmatrix} x_0 - x_7 \\ x_1 - x_6 \\ x_2 - x_5 \\ x_3 - x_4 \end{bmatrix} \end{cases}$$

$$\begin{cases} \alpha = 1/2 \cos(\pi/4) \\ \beta = 1/2 \cos(\pi/8) \\ \delta = 1/2 \sin(\pi/8) \end{cases}$$

$$\begin{cases} \lambda = 1/2 \cos(\pi/16) \\ \gamma = 1/2 \cos(3\pi/16) \\ \mu = 1/2 \sin(3\pi/16) \\ \nu = 1/2 \sin(\pi/16) \end{cases}$$

Application example - 8*8 2D DCT



Cycle 5

$$z = \sqrt{\frac{2}{N}}T(N)x$$

$$\begin{bmatrix} z_0 \\ z_2 \\ z_4 \\ z_6 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} \\ \beta & \delta & -\delta & -\beta \\ \alpha & -\alpha & -\alpha & \alpha \\ \delta & -\beta & \beta & -\delta \end{bmatrix} \begin{bmatrix} x_0 + x_7 \\ x_1 + x_6 \\ x_2 + x_5 \\ x_3 + x_4 \end{bmatrix}$$

α = 1/2 cos (π/4)
β = 1/2 cos (π/8)
δ = 1/2 sin (π/8)

$$\begin{bmatrix} z_1 \\ z_3 \\ z_5 \\ z_7 \end{bmatrix} = \begin{bmatrix} \lambda & \gamma & \mu & \nu \\ \gamma & -\nu & -\lambda & -\mu \\ \mu & -\lambda & \nu & \gamma \\ \nu & -\mu & \gamma & -\lambda \end{bmatrix} \begin{bmatrix} x_0 - x_7 \\ x_1 - x_6 \\ x_2 - x_5 \\ x_3 - x_4 \end{bmatrix}$$

λ = 1/2 cos (π/16)
γ = 1/2 cos (3π/16)
μ = 1/2 sin (3π/16)
ν = 1/2 sin (π/16)

2 transformed samples
computed each *5 clock cycles*

⬇

8*8 2D transformed samples each <u>320 clock cycles</u>

Conclusion

- **DESIGN**
  - ➔ Reconfigurable IP Core for SoC
  - ➔ Assembling Software
  - ➔ RING-8 prototype
- **FEATURES**
  - ➔ Customisable IP Core
  - ➔ Good performance / area trade-off : *Ring-8@200MHz (0.18µ)*

    3.3 mm^2

    1600 MIPS

    Results for DCT, Wavelet Transform, Motion Estimation

Application example - 8*8 2D DCT

Even-Odd frequency decomposition

$$z = \sqrt{\frac{2}{N}}\,T(N)\,x$$

$$\begin{bmatrix} z_0 \\ z_2 \\ z_4 \\ z_6 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} \\ \beta & \delta & -\delta & -\beta \\ \alpha & -\alpha & -\alpha & \alpha \\ \delta & -\beta & \beta & -\delta \end{bmatrix} \begin{bmatrix} x_0 + x_7 \\ x_1 + x_6 \\ x_2 + x_5 \\ x_3 + x_4 \end{bmatrix}$$

α = 1/2 cos (π/4)
β = 1/2 cos (π/8)
δ = 1/2 sin (π/8)

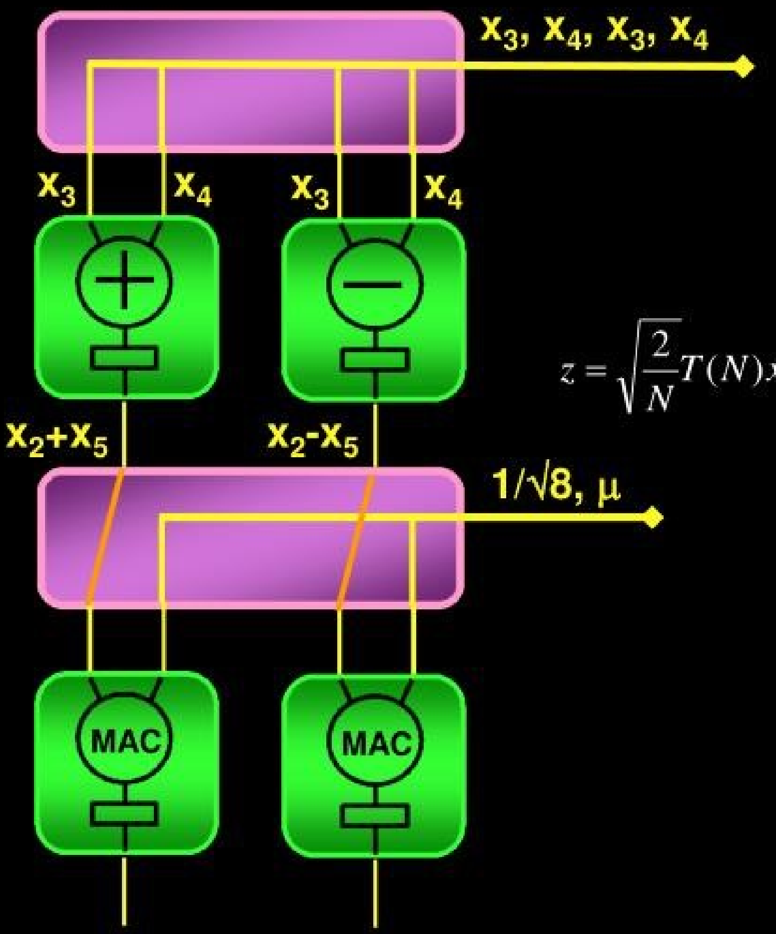
$$\begin{bmatrix} z_1 \\ z_3 \\ z_5 \\ z_7 \end{bmatrix} = \begin{bmatrix} \lambda & \gamma & \mu & \nu \\ \gamma & -\nu & -\lambda & -\mu \\ \mu & -\lambda & \nu & \gamma \\ \nu & -\mu & \gamma & -\lambda \end{bmatrix} \begin{bmatrix} x_0 - x_7 \\ x_1 - x_6 \\ x_2 - x_5 \\ x_3 - x_4 \end{bmatrix}$$

λ = 1/2 cos (π/16)
γ = 1/2 cos (3π/16)
μ = 1/2 sin (3π/16)
ν = 1/2 sin (π/16)

Independent Matrix Products (//)

Application example - 8*8 2D DCT

Cycle 3



$$z = \sqrt{\frac{2}{N}}T(N)x$$

$$\begin{bmatrix} z_0 \\ z_2 \\ z_4 \\ z_6 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} \\ \beta & \delta & -\delta & -\beta \\ \alpha & -\alpha & -\alpha & \alpha \\ \delta & -\beta & \beta & -\delta \end{bmatrix} \begin{bmatrix} x_0 + x_7 \\ x_1 + x_6 \\ x_2 + x_5 \\ x_3 + x_4 \end{bmatrix}$$

$$\begin{bmatrix} z_1 \\ z_3 \\ z_5 \\ z_7 \end{bmatrix} = \begin{bmatrix} \lambda & \gamma & \mu & \nu \\ \gamma & -\nu & -\lambda & -\mu \\ \mu & -\lambda & \nu & \gamma \\ \nu & -\mu & \gamma & -\lambda \end{bmatrix} \begin{bmatrix} x_0 - x_7 \\ x_1 - x_6 \\ x_2 - x_5 \\ x_3 - x_4 \end{bmatrix}$$

$\alpha = 1/2 \cos(\pi/4)$

$\beta = 1/2 \cos(\pi/8)$

$\delta = 1/2 \sin(\pi/8)$

$\lambda = 1/2 \cos(\pi/16)$

$\gamma = 1/2 \cos(3\pi/16)$

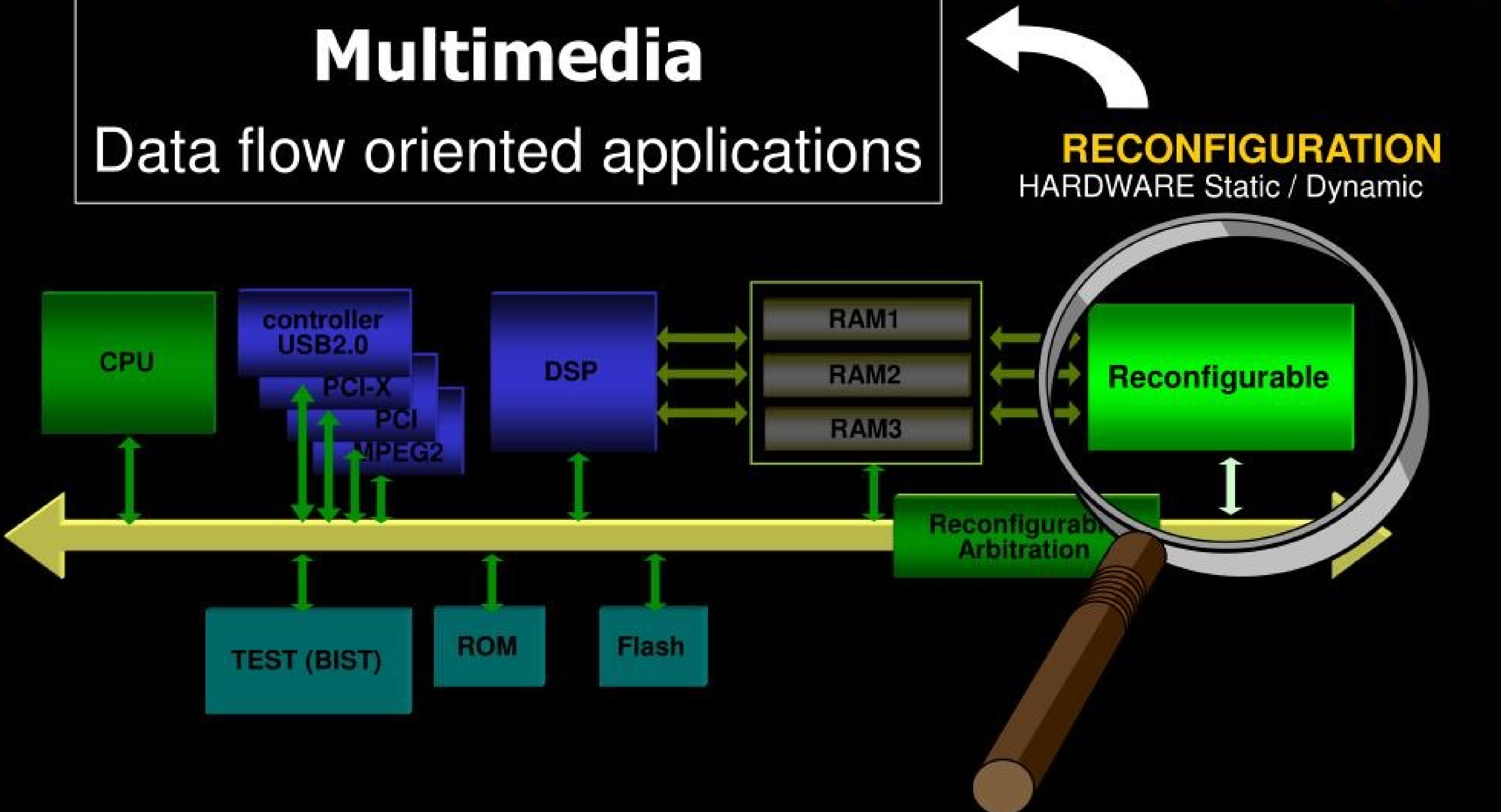
$\mu = 1/2 \sin(3\pi/16)$

$\nu = 1/2 \sin(\pi/16)$

Introduction

Multimedia

Data flow oriented applications



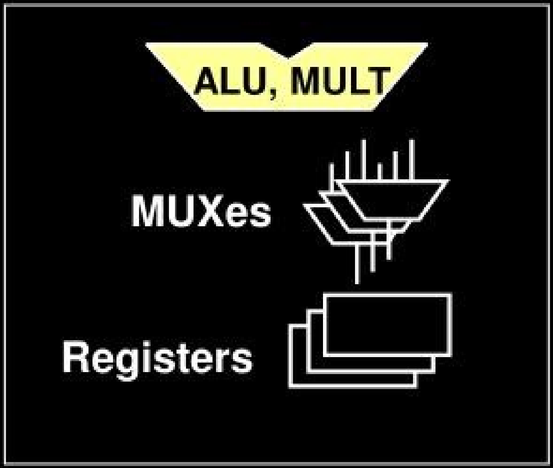
Introduction

What kind of base block is suitable for Multimedia?

- **Coarse Grain:**
  - → Granularity: WORD
    adapted to DSP, data flow oriented processing

SYSTOLIC RING

Coarse Grain
Dynamically Reconfigurable
Architecture

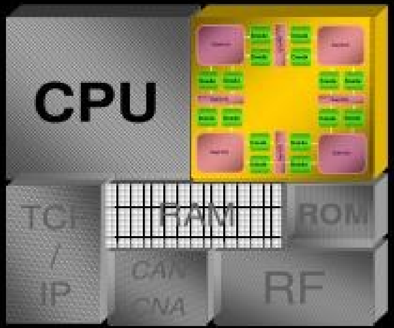


- → Low reconfiguration over-cost
- → High level of performances

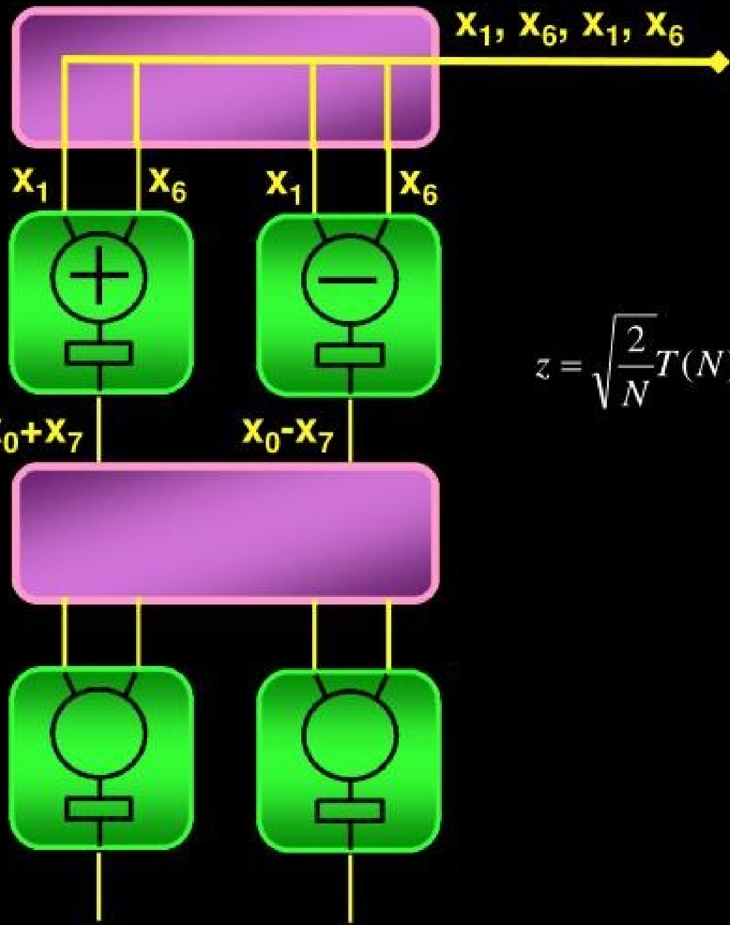
Outline

The Systolic Ring : A Scalable Dynamically Reconfigurable Core for Embedded Systems

- **The Systolic Ring**
  - Building Block
  - Operative Layer Topology
  - System Overview
  - Features
- **Application Example**
  - 8*8 2D DCT
  - Structural Mapping
  - Performance Comparisons
- **Conclusion**



Application example - 8*8 2D DCT



$$z = \sqrt{\frac{2}{N}} T(N) x$$

$$\begin{bmatrix} z_0 \\ z_2 \\ z_4 \\ z_6 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} \\ \beta & \delta & -\delta & -\beta \\ \alpha & -\alpha & -\alpha & \alpha \\ \delta & -\beta & \beta & -\delta \end{bmatrix} \begin{bmatrix} x_0 + x_7 \\ x_1 + x_6 \\ x_2 + x_5 \\ x_3 + x_4 \end{bmatrix}$$

$\alpha = 1/2 \cos(\pi/4)$
$\beta = 1/2 \cos(\pi/8)$
$\delta = 1/2 \sin(\pi/8)$

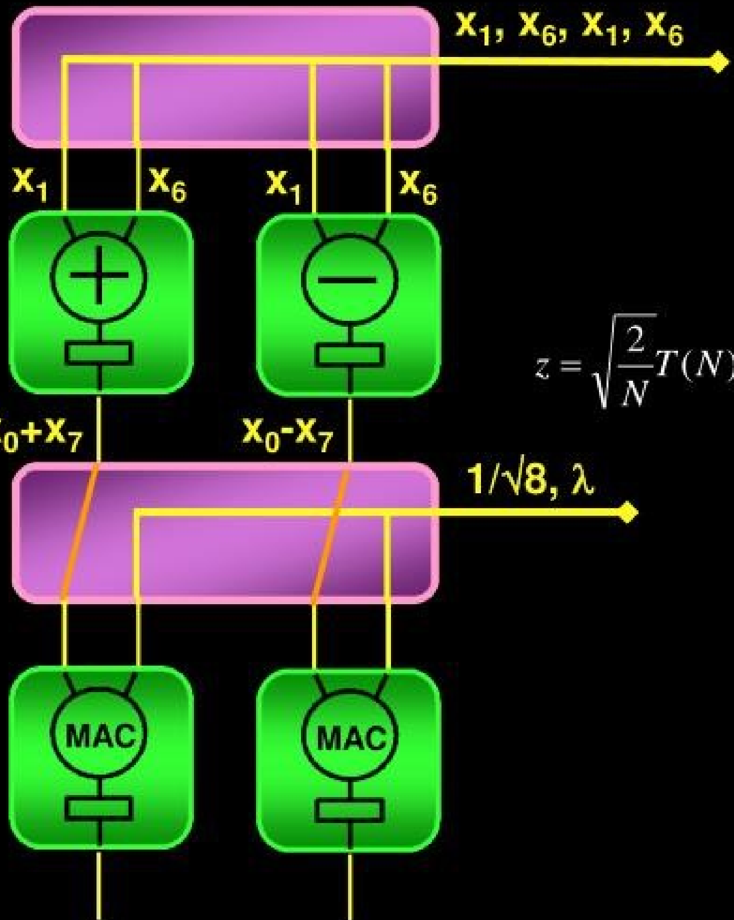
$$\begin{bmatrix} z_1 \\ z_3 \\ z_5 \\ z_7 \end{bmatrix} = \begin{bmatrix} \lambda & \gamma & \mu & \nu \\ \gamma & -\nu & -\lambda & -\mu \\ \mu & -\lambda & \nu & \gamma \\ \nu & -\mu & \gamma & -\lambda \end{bmatrix} \begin{bmatrix} x_0 - x_7 \\ x_1 - x_6 \\ x_2 - x_5 \\ x_3 - x_4 \end{bmatrix}$$

$\lambda = 1/2 \cos(\pi/16)$
$\gamma = 1/2 \cos(3\pi/16)$
$\mu = 1/2 \sin(3\pi/16)$
$\nu = 1/2 \sin(\pi/16)$

Raw-Column product -> MAC operations

Application example - 8*8 2D DCT

Cycle 1



$$z = \sqrt{\frac{2}{N}}T(N)x$$

$$\begin{bmatrix} z_0 \\ z_2 \\ z_4 \\ z_6 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} \\ \beta & \delta & -\delta & -\beta \\ \alpha & -\alpha & -\alpha & \alpha \\ \delta & -\beta & \beta & -\delta \end{bmatrix} \begin{bmatrix} x_0 + x_7 \\ x_1 + x_6 \\ x_2 + x_5 \\ x_3 + x_4 \end{bmatrix}$$

α = 1/2 cos (π/4)

β = 1/2 cos (π/8)

δ = 1/2 sin (π/8)

$$\begin{bmatrix} z_1 \\ z_3 \\ z_5 \\ z_7 \end{bmatrix} = \begin{bmatrix} \lambda & \gamma & \mu & \nu \\ \gamma & -\nu & -\lambda & -\mu \\ \mu & -\lambda & \nu & \gamma \\ \nu & -\mu & \gamma & -\lambda \end{bmatrix} \begin{bmatrix} x_0 - x_7 \\ x_1 - x_6 \\ x_2 - x_5 \\ x_3 - x_4 \end{bmatrix}$$

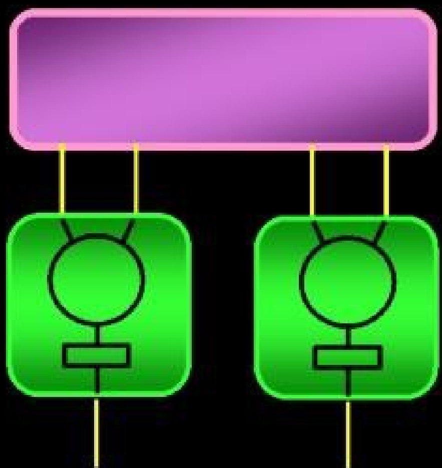
λ = 1/2 cos (π/16)

γ = 1/2 cos (3π/16)

μ = 1/2 sin (3π/16)

ν = 1/2 sin (π/16)

Application example - 8*8 2D DCT



Addition & Subtraction of image samples

$$z = \sqrt{\frac{2}{N}}T(N)x$$

$$\begin{bmatrix} z_0 \\ z_2 \\ z_4 \\ z_6 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} \\ \beta & \delta & -\delta & -\beta \\ \alpha & -\alpha & -\alpha & \alpha \\ \delta & -\beta & \beta & -\delta \end{bmatrix} \begin{bmatrix} x_0 + x_7 \\ x_1 + x_6 \\ x_2 + x_5 \\ x_3 + x_4 \end{bmatrix}$$

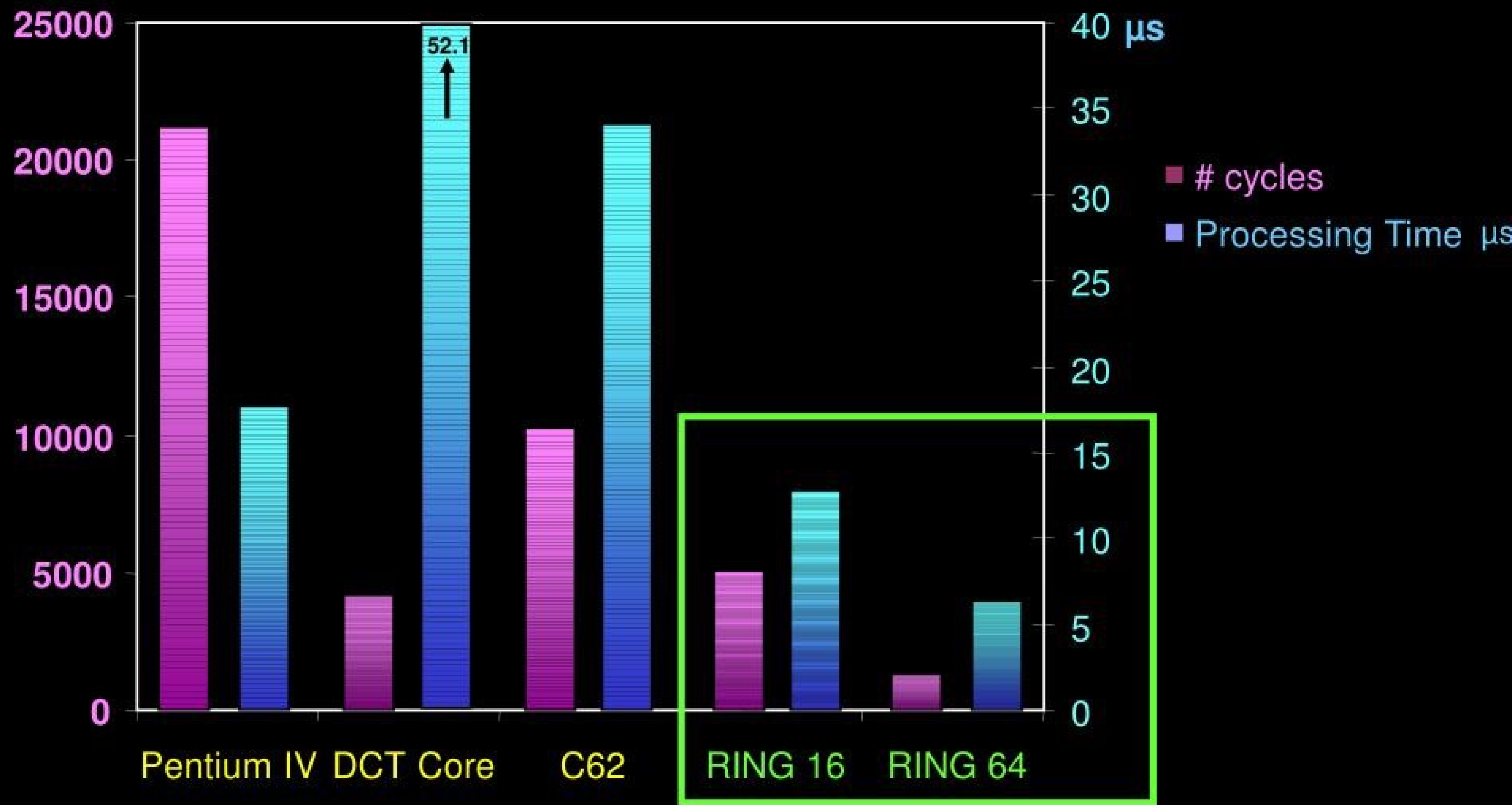
$$\begin{bmatrix} z_1 \\ z_3 \\ z_5 \\ z_7 \end{bmatrix} = \begin{bmatrix} \lambda & \gamma & \mu & \nu \\ \gamma & -\nu & -\lambda & -\mu \\ \mu & -\lambda & \nu & \gamma \\ \nu & -\mu & \gamma & -\lambda \end{bmatrix} \begin{bmatrix} x_0 - x_7 \\ x_1 - x_6 \\ x_2 - x_5 \\ x_3 - x_4 \end{bmatrix}$$

α = 1/2 cos (π/4)
β = 1/2 cos (π/8)
δ = 1/2 sin (π/8)

λ = 1/2 cos (π/16)
γ = 1/2 cos (3π/16)
μ = 1/2 sin (3π/16)
ν = 1/2 sin (π/16)

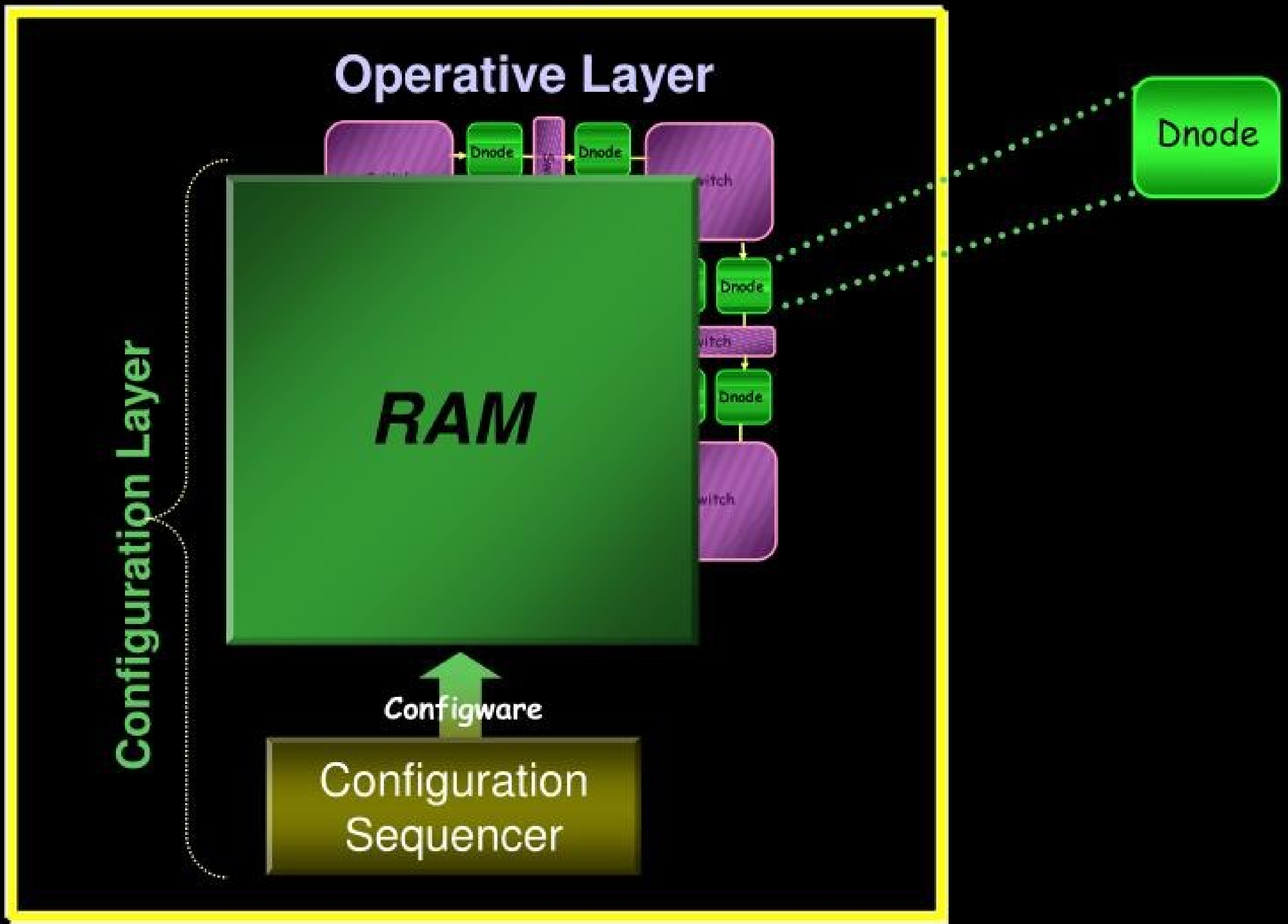
Application example - DCT 2D 8*8

64*64 image example - Comparisons



The Systolic Ring - System Overview

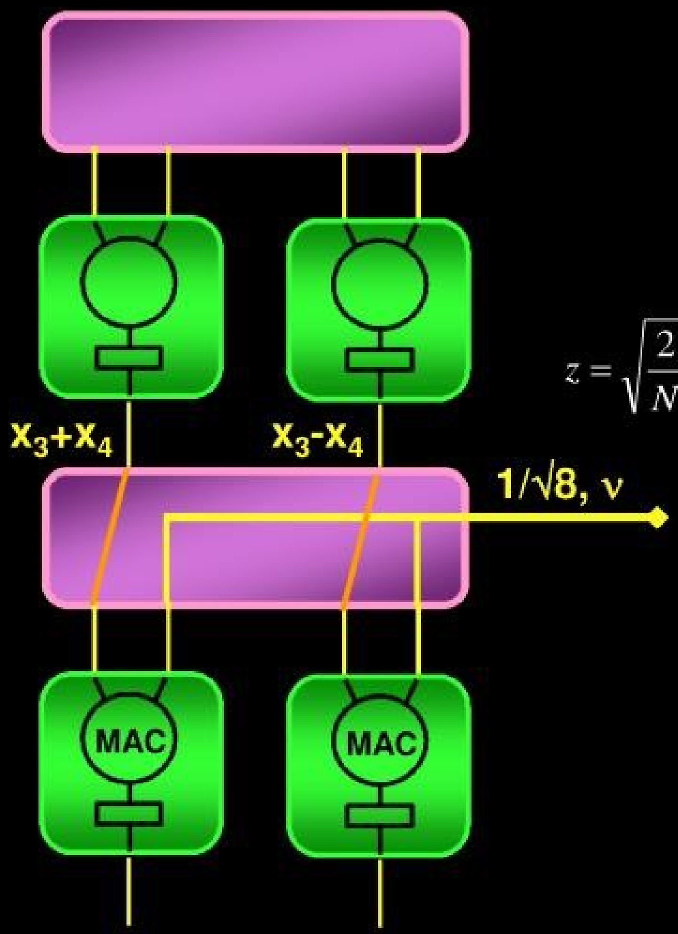
Two-layers based reconfigurable architecture



Coarse Grain Dynamically
Reconfigurable Architecture

Application example - 8*8 2D DCT

Cycle 4



$$z = \sqrt{\frac{2}{N}}T(N)x$$

$$\begin{bmatrix} z_0 \\ z_2 \\ z_4 \\ z_6 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} \\ \beta & \delta & -\delta & -\beta \\ \alpha & -\alpha & -\alpha & \alpha \\ \delta & -\beta & \beta & -\delta \end{bmatrix} \begin{bmatrix} x_0 + x_7 \\ x_1 + x_6 \\ x_2 + x_5 \\ x_3 + x_4 \end{bmatrix}$$

$\alpha = 1/2 \cos(\pi/4)$

$\beta = 1/2 \cos(\pi/8)$

$\delta = 1/2 \sin(\pi/8)$

$$\begin{bmatrix} z_1 \\ z_3 \\ z_5 \\ z_7 \end{bmatrix} = \begin{bmatrix} \lambda & \gamma & \mu & \nu \\ \gamma & -\nu & -\lambda & -\mu \\ \mu & -\lambda & \nu & \gamma \\ \nu & -\mu & \gamma & -\lambda \end{bmatrix} \begin{bmatrix} x_0 - x_7 \\ x_1 - x_6 \\ x_2 - x_5 \\ x_3 - x_4 \end{bmatrix}$$

$\lambda = 1/2 \cos(\pi/16)$

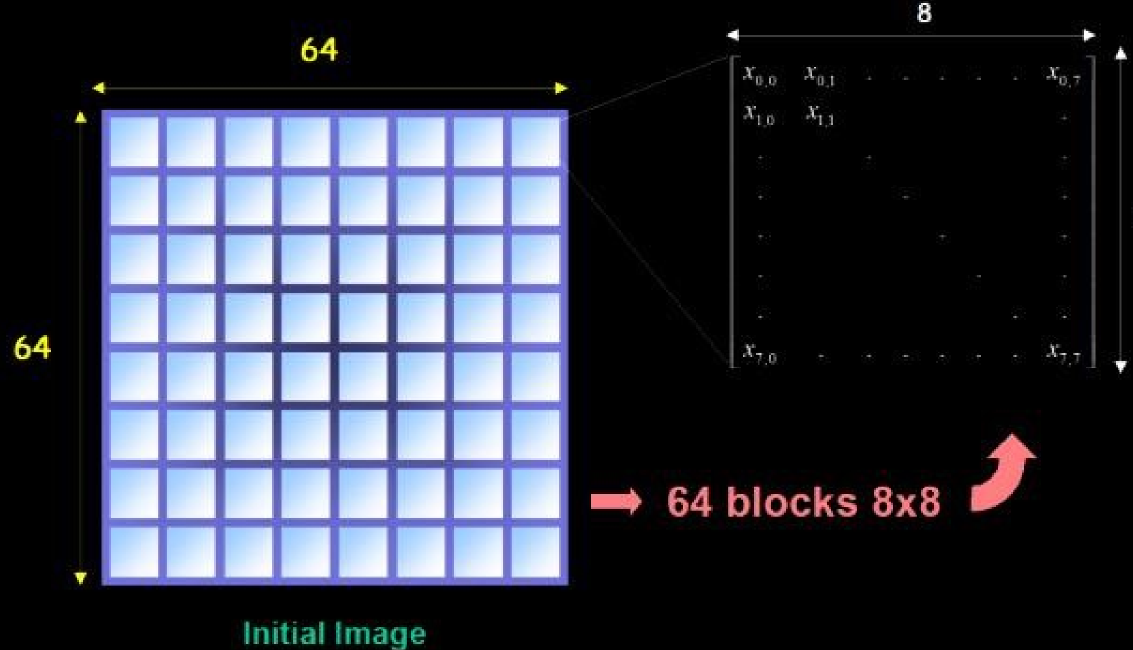
$\gamma = 1/2 \cos(3\pi/16)$

$\mu = 1/2 \sin(3\pi/16)$

$\nu = 1/2 \sin(\pi/16)$

Application example - DCT 2D 8*8

64*64 image example



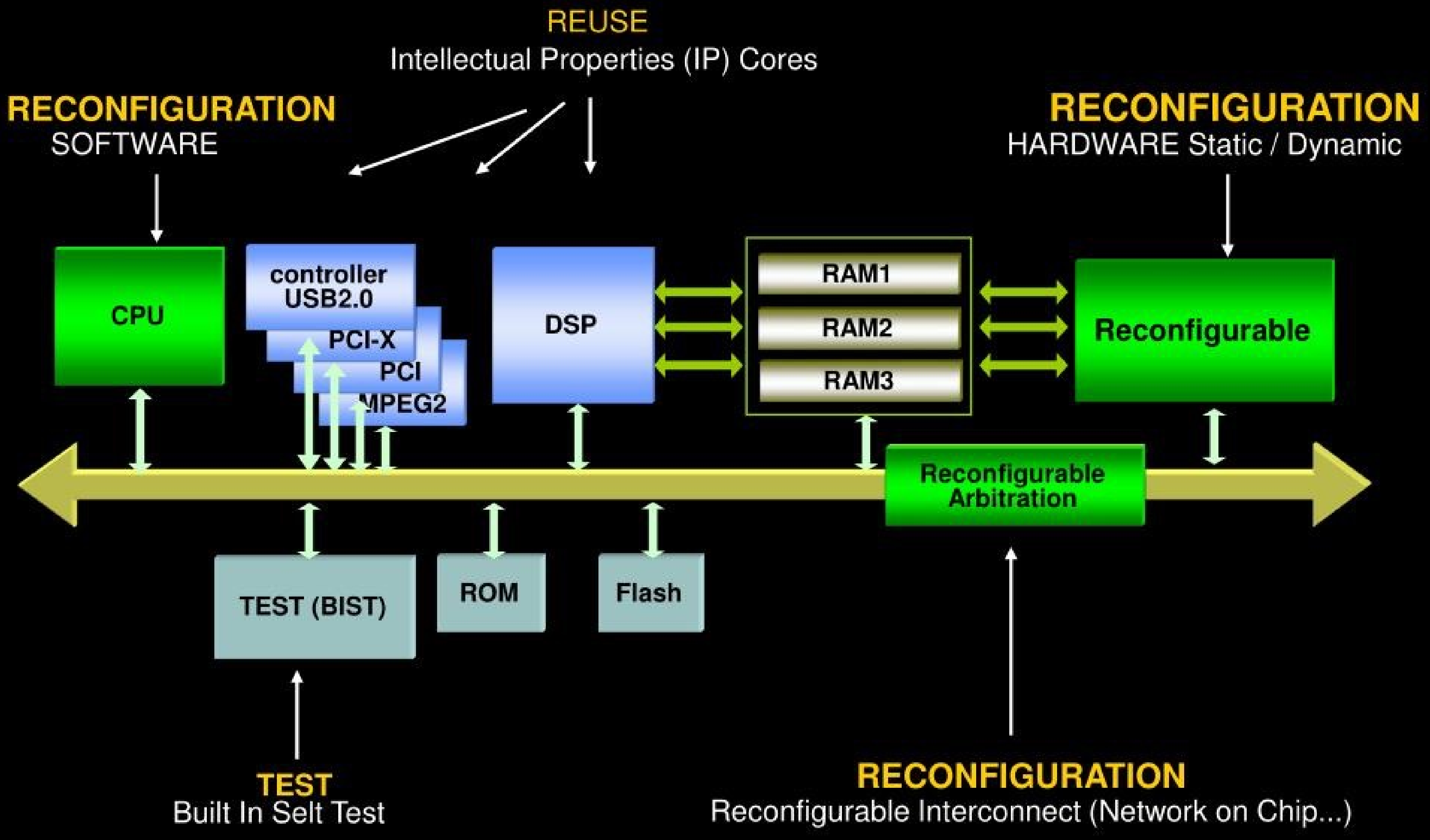
→ 64 blocks 8x8

→ **DCT 2D 8*8 performed by 4 Dnodes**

$$\#\,\text{computation clock cycles} = 320\,.\,(\#\text{ of 8-points 2D blocks}).\,\frac{4}{N}$$

RING-N implementation (N Dnodes)

Introduction - SoC architectures



System on a Chip

Sophia Antipolis
9-10 Oct 2002

The Systolic Ring : A Scalable Dynamically Reconfigurable Core for Embedded Systems

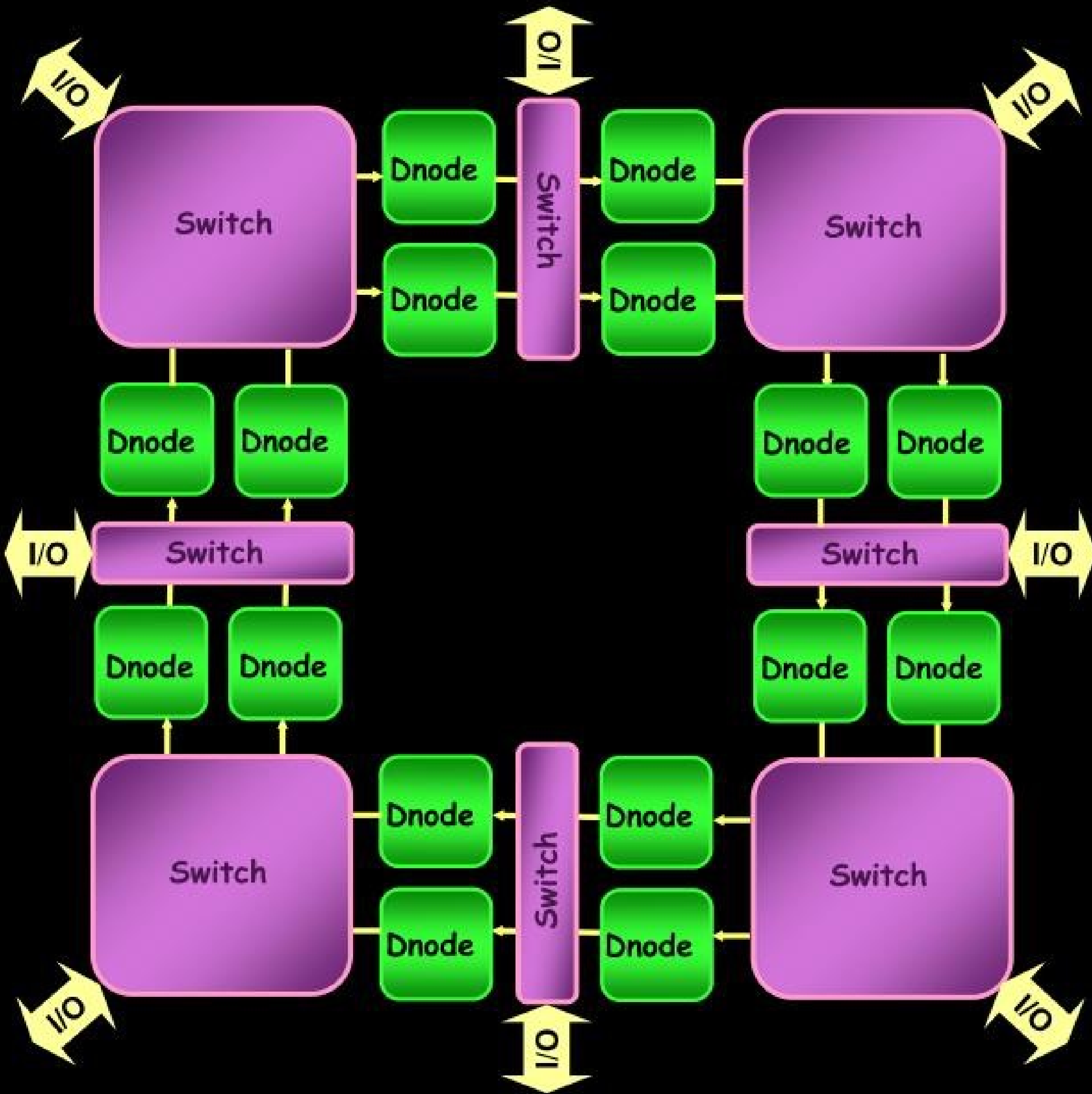
Pascal BENOIT,

G. SASSATELLI, M. ROBERT, L. TORRES, G. CAMBON, T. GIL

LIRM

UNIVERSITE MONTPELLIER II
SCIENCES ET TECHNIQUES DU LANGUEDOC

The Systolic Ring - Operative Layer Topology

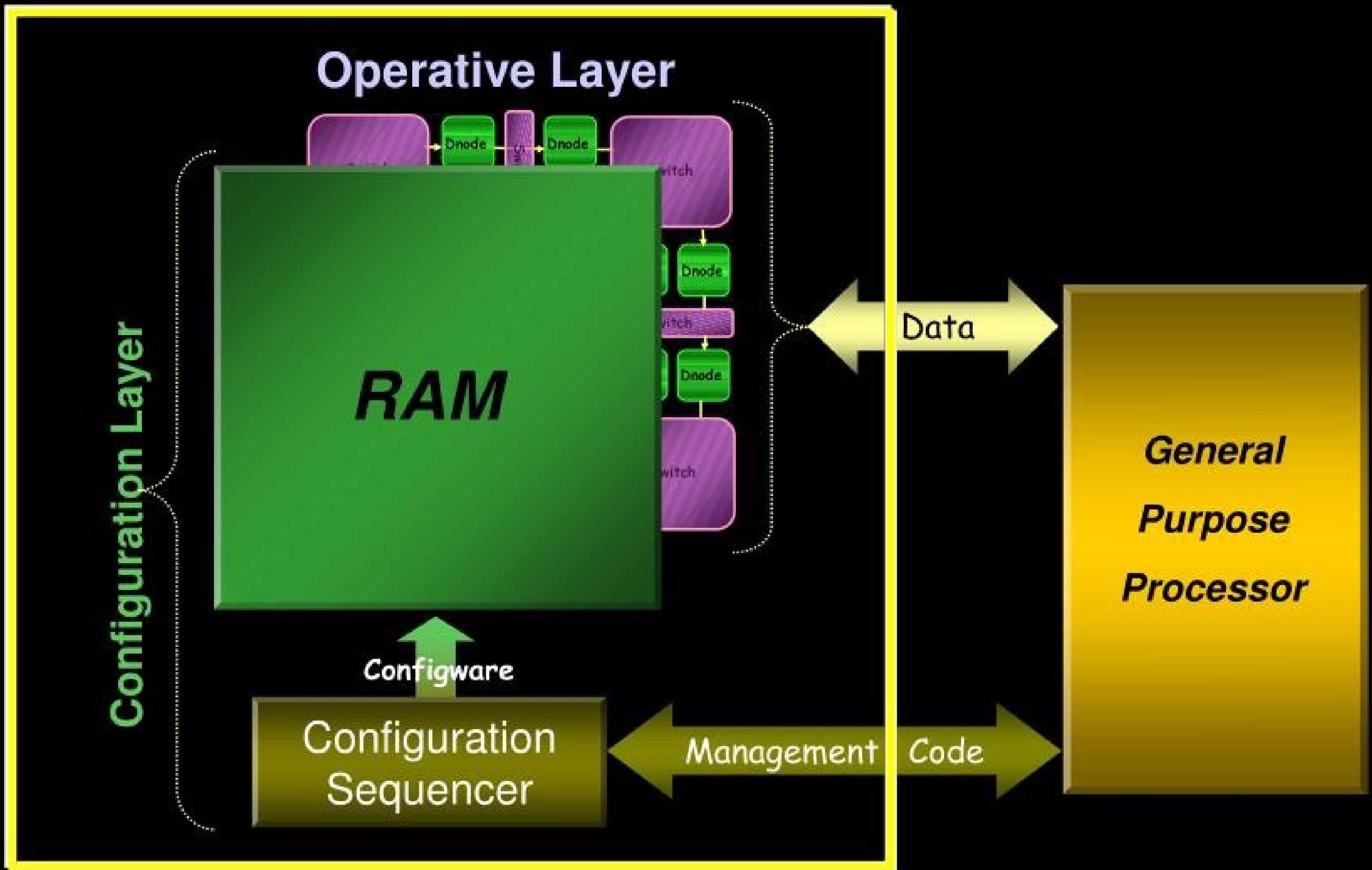


Ring Structure

Customisable...

The Systolic Ring - System Overview

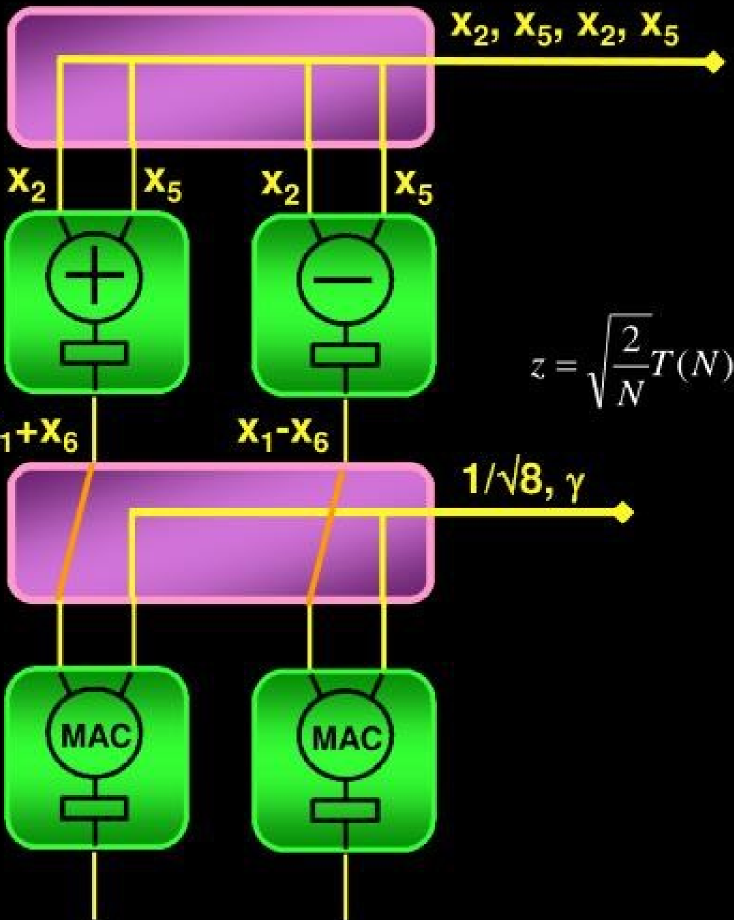
Not a stand-alone solution



Coprocessor for data flow oriented applications

Application example - 8*8 2D DCT

Cycle 2



$$z = \sqrt{\frac{2}{N}}T(N)x$$

$$\begin{bmatrix} z_0 \\ z_2 \\ z_4 \\ z_6 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} & 1/\sqrt{8} \\ \beta & \delta & -\delta & -\beta \\ \alpha & -\alpha & -\alpha & \alpha \\ \delta & -\beta & \beta & -\delta \end{bmatrix} \begin{bmatrix} x_0 + x_7 \\ x_1 + x_6 \\ x_2 + x_5 \\ x_3 + x_4 \end{bmatrix}$$

α = 1/2 cos (π/4)
β = 1/2 cos (π/8)
δ = 1/2 sin (π/8)

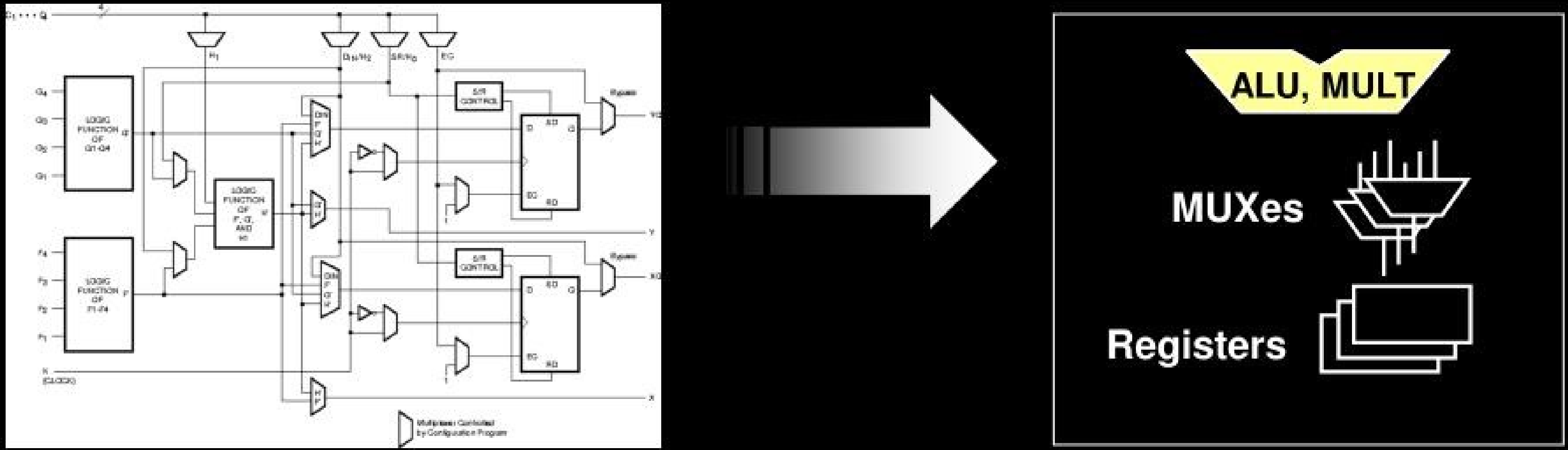
$$\begin{bmatrix} z_1 \\ z_3 \\ z_5 \\ z_7 \end{bmatrix} = \begin{bmatrix} \lambda & \gamma & \mu & \nu \\ \gamma & -\nu & -\lambda & -\mu \\ \mu & -\lambda & \nu & \gamma \\ \nu & -\mu & \gamma & -\lambda \end{bmatrix} \begin{bmatrix} x_0 - x_7 \\ x_1 - x_6 \\ x_2 - x_5 \\ x_3 - x_4 \end{bmatrix}$$

λ = 1/2 cos (π/16)
γ = 1/2 cos (3π/16)
μ = 1/2 sin (3π/16)
ν = 1/2 sin (π/16)

Introduction

What kind of base block is suitable for Multimedia?

- **Fine grain:**
  - ➔ Granularity: BIT
    adapted to Prototyping, Encryption
- **Coarse Grain:**
  - ➔ Granularity: WORD
    adapted to DSP, data flow oriented processing

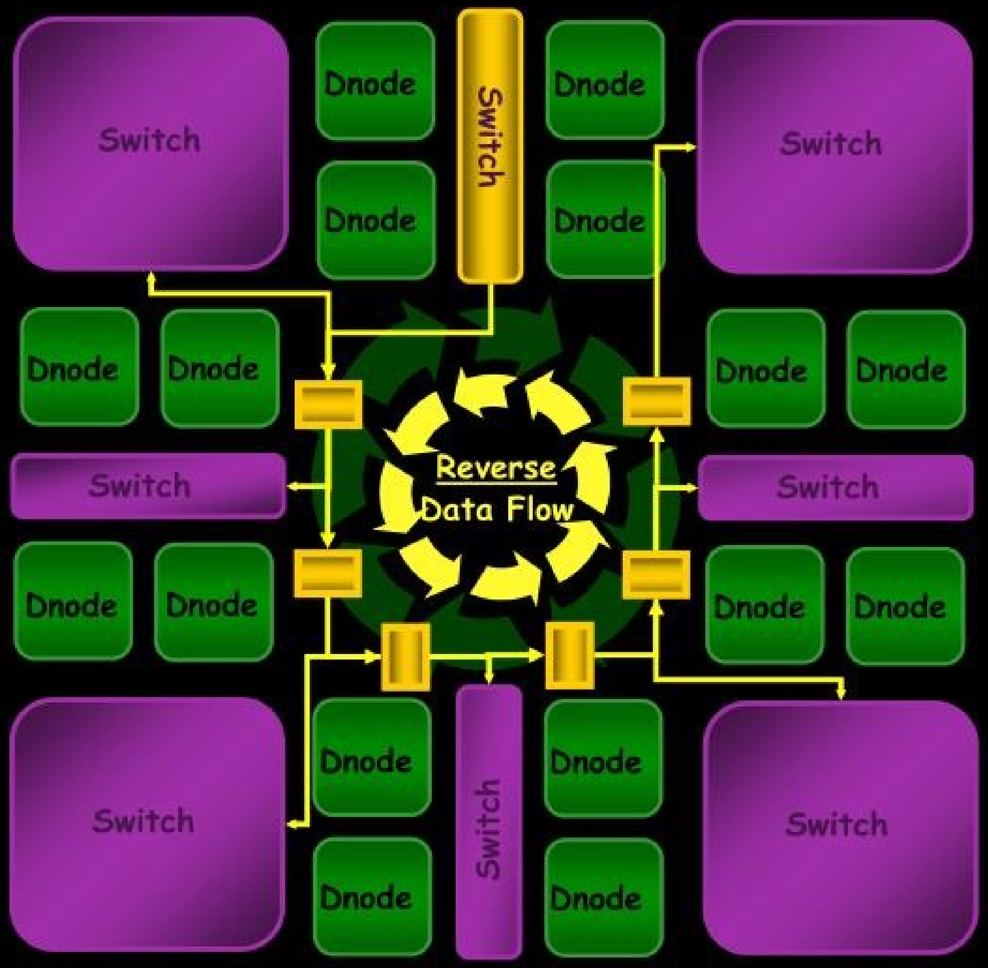


Fine grain:
- ➔ High reconfiguration over-cost
- ➔ Low Functional frequencies

Coarse Grain:
- ➔ Low reconfiguration over-cost
- ➔ High level of performances

The Systolic Ring - Operative Layer Topology

Data Flows



Forward Data Flow

Unidirectional data transit between

successive layers (circular pipeline

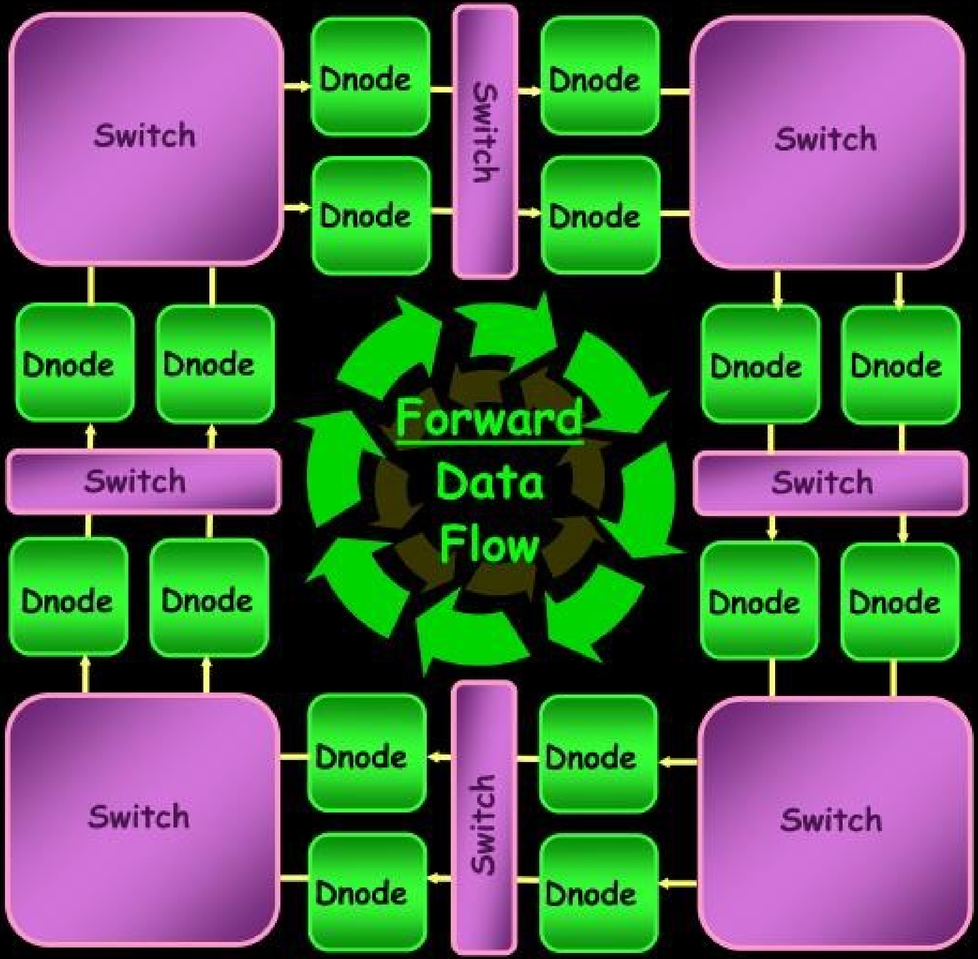
Reverse Data Flow

Feedback pipeline network

for recursive algorithms

The Systolic Ring - Operative Layer Topology

Data Flows



Forward Data Flow

Unidirectional data transit between

successive layers (circular pipeline)

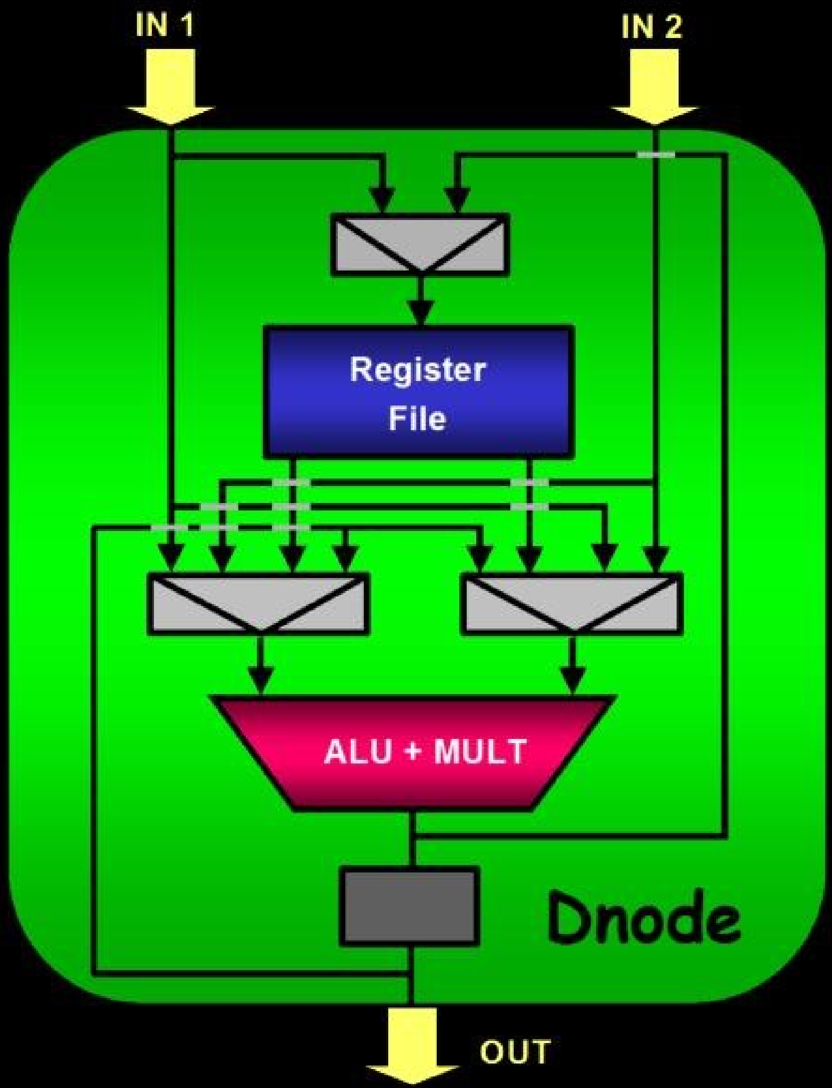
Reverse Data Flow

Feedback pipeline network

for recursive algorithms

The Systolic Ring - Building block

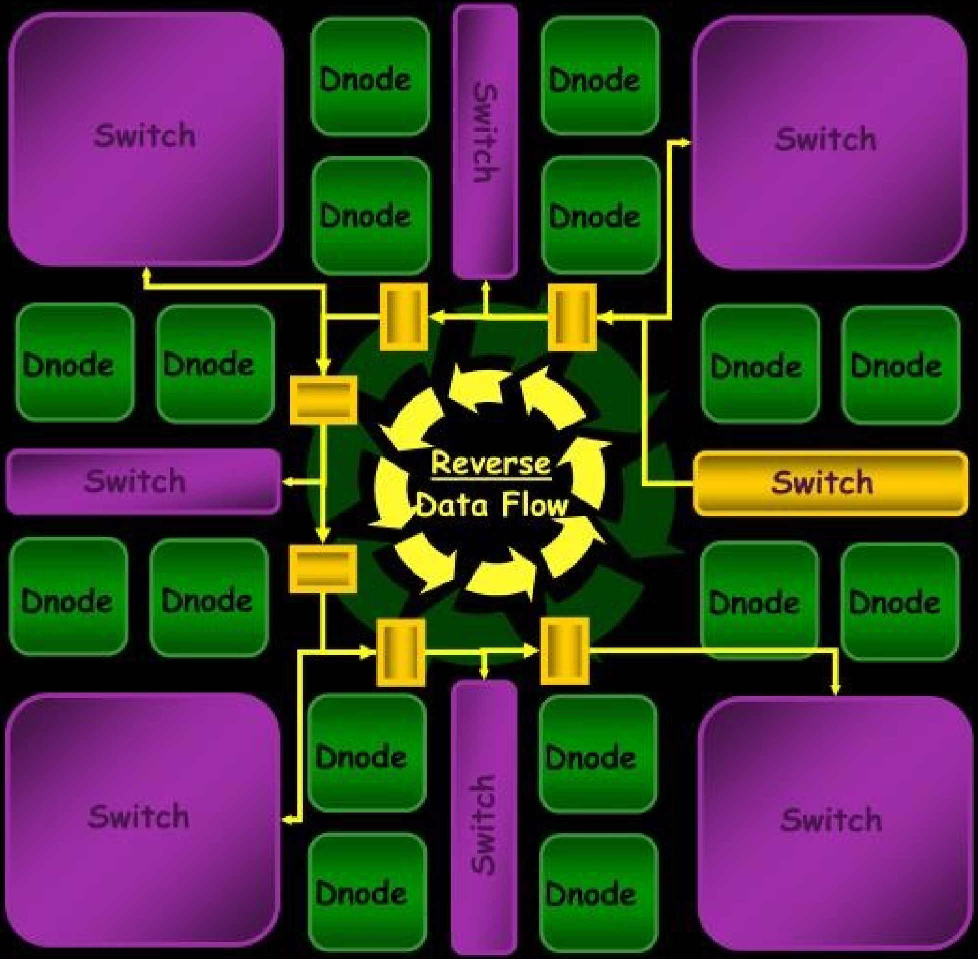
DNODE (Data Node)



- **Data processing oriented block**
  - ➔ ALU + Multiplier ⇑ MAC)
- **Programmable component**
  - ➔ Local Sequencer
    - Dynamic and autonomous configuration management
  - ➔ *one instruction per cycle*

The Systolic Ring - Operative Layer Topology

Data Flows



Forward Data Flow

Unidirectional data transit between

successive layers (circular pipeline

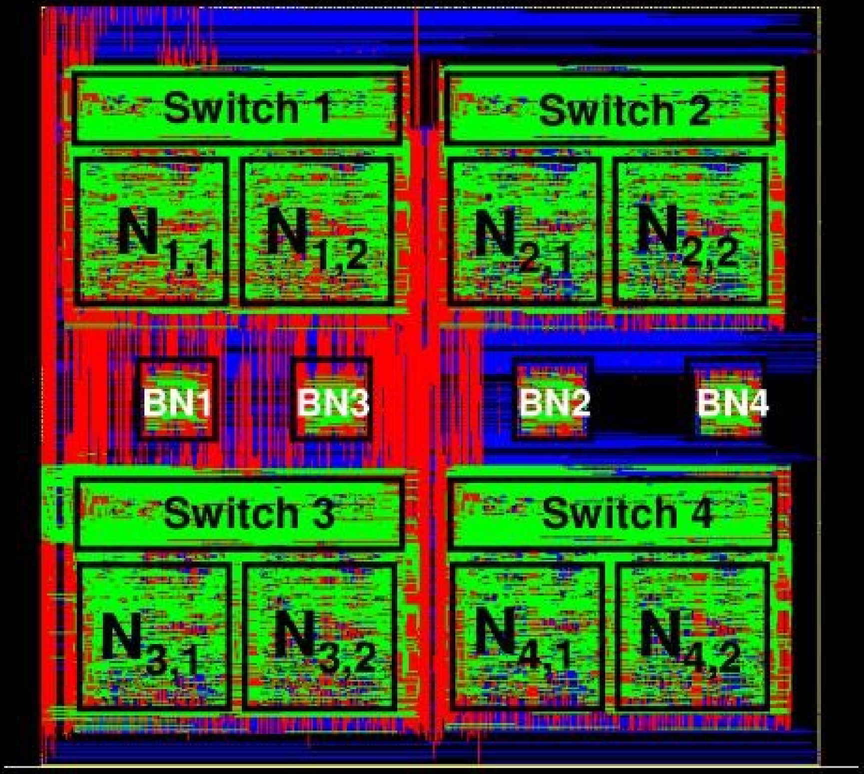
Reverse Data Flow

Feedback pipeline network

for recursive algorithms

The Systolic Ring

Systolic Ring Features



Operative Layer Layout

- **RING-8 (8 Dnodes)**
  - ➔ 0.18µ technology
  - 3.3 mm²
  - 200 MHz
  - 1600 MIPS
  - 1600 MMACs / s

Process geometry dropping ⇨ increase Dnode #