

Raytheon



Computer Systems, Inc.
MERCURY

MONARCH: **A Morphable Networked micro-ARCHitecture**

John Granacki, USC/Information Sciences Institute

Mike Vahey, Raytheon



MONARCH 

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE 21 MAY 2003		2. REPORT TYPE N/A		3. DATES COVERED -	
4. TITLE AND SUBTITLE MONARCH: A Morphable Networked micro-ARCHitecture				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) USC/Information Sciences Institute and Raytheon				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited					
13. SUPPLEMENTARY NOTES The original document contains color images.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 33	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			



Outline



- ◆ MONARCH Team, Goals & Approach
- ◆ DIVA (Data Intensive Architecture) Leverage: The Chip
- ◆ Raytheon HPPS (High Performance Processor System)
- ◆ MONARCH Architecture & Applications
- ◆ Summary & Conclusions





The Team



Raytheon

USC
INFORMATION
SCIENCES
INSTITUTE

Computer Systems, Inc.
MERCURY

Co-Principal Investigator
Michael Vahey

Principal Investigator
John Granacki

Arlan pool

RESEARCH STAFF

John "Chip" Bodenschatz
Frank Brandon
Reagan Branstetter
Charles Channell
Phil Rosen
Mike Walker

RESEARCH STAFF

Jeff Draper
Pedro Diniz
Jeff LaCoss

RESEARCH STAFF

Vlad Kaufman



MONARCH 



GOALS



- ◆ ***To support multiple classes of military missions*** with a single morphable architecture
- ◆ ***To eliminate processing system redundancies*** through rapid dynamic reconfiguration of front-end filtering and data-reduction processing
- ◆ ***To reduce application development costs*** by allowing the hardware to be mapped to the algorithms both statically and dynamically
- ◆ ***To develop an architecture that can quickly and efficiently adapt to changing situations*** - internal (fault tolerance, sensors configurations) and external (threats change, mission phasing, environment)





Key Ideas



Combines fine, medium and coarse grain processing resources on a single chip

Matches hardware to the algorithms and the control flow mechanisms

Configures memory structures for efficient front-end and back-end processing

Provides flexible gigabyte I/O channels for direct interface to sensors and inter-chip communication

Supports all systems processing requirements with a single MONARCH chip type





Approach



- ◆ Leverage DARPA-sponsored DIVA Project results, Raytheon IRAD-sponsored HPPS and Mercury Stream Co-processing Engine
- ◆ Use DoD missions to drive micro-architecture and morphing concepts and implementation
- ◆ Determine the “sweet spot” for mixing large, small- to-medium and fine-grained elements
- ◆ Through experiments and simulations demonstrate a “single chip” VLSI processing architecture based on DIVA and HPPS



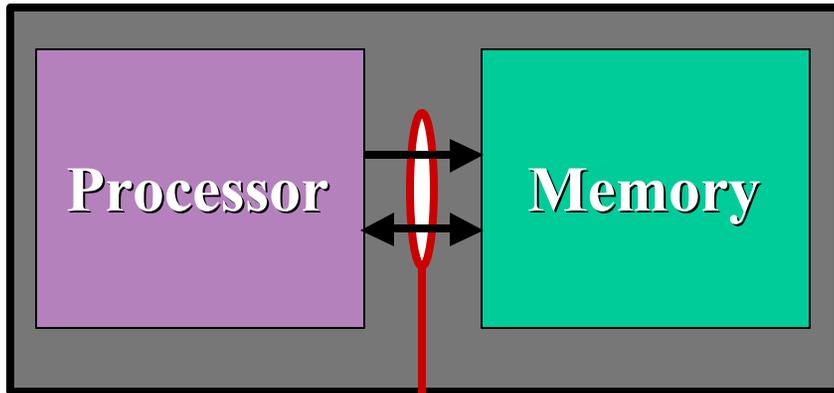


DIVA Leverage: The Chip

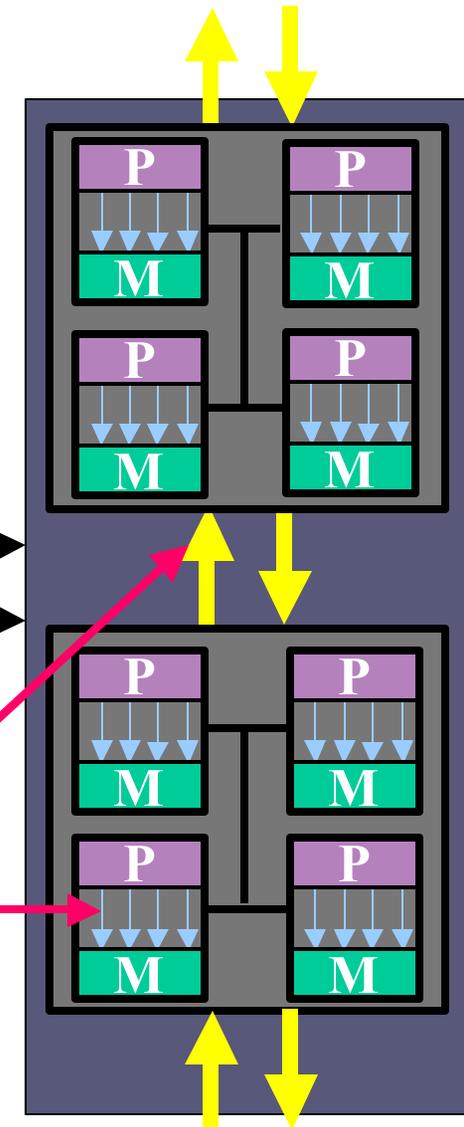




Exploiting The Bandwidth in a System



A
Solution



DIVA Solutions:

- Move concurrent processing on-chip
- More bandwidth and less latency on chip
- Added bandwidth between memories
- Lower latencies throughout system





DIVA Software/Hardware



**Tools &
Applications**

<p><u>Host Runtime Layer</u> Synchronization, Flushing, Thread Management, Host Parcels</p>	<p><u>Application</u></p>
---	---------------------------

**PIM
Applications**

Compiler
Data Placement, Parallelism,
Host-PIM Mapping,
Parcels, Coherence

PIM Backend Compiler
Code generation for scalar
and WideWord

**System
Management**

OS (Linux)
Page Placement,
Paging for Host & PIMs,
Scheduling, PIM Initialization

PIM Runtime Kernel
Parcel Management,
Address Translation Faults,
PIM Context Switches,
Synchronization

**Runtime
Coordination**

**Physical
Hardware**

Host
System

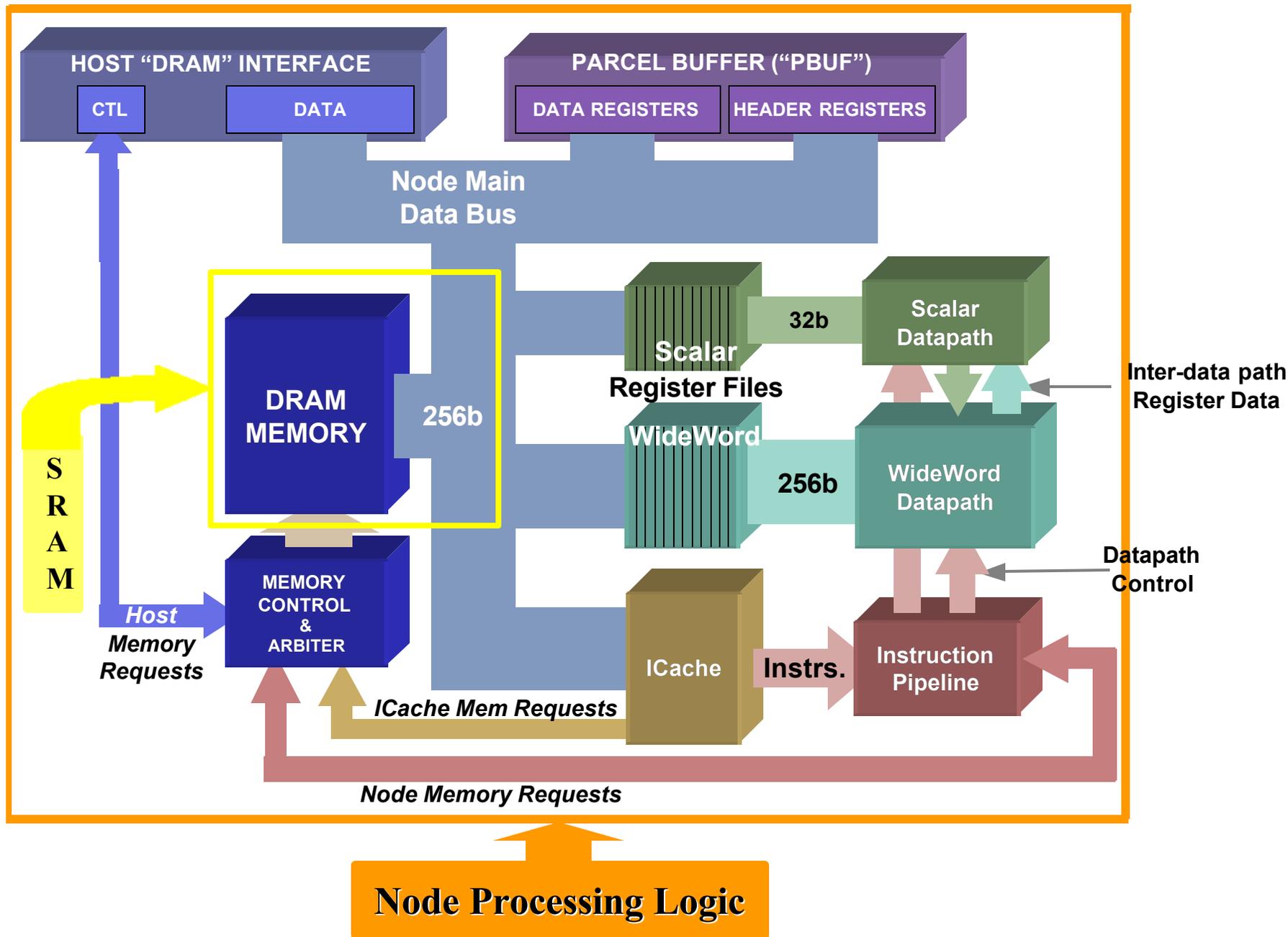
Memory
Controller

PIM VLSI Devices
Processor, Memory Array

**Physical
Hardware**

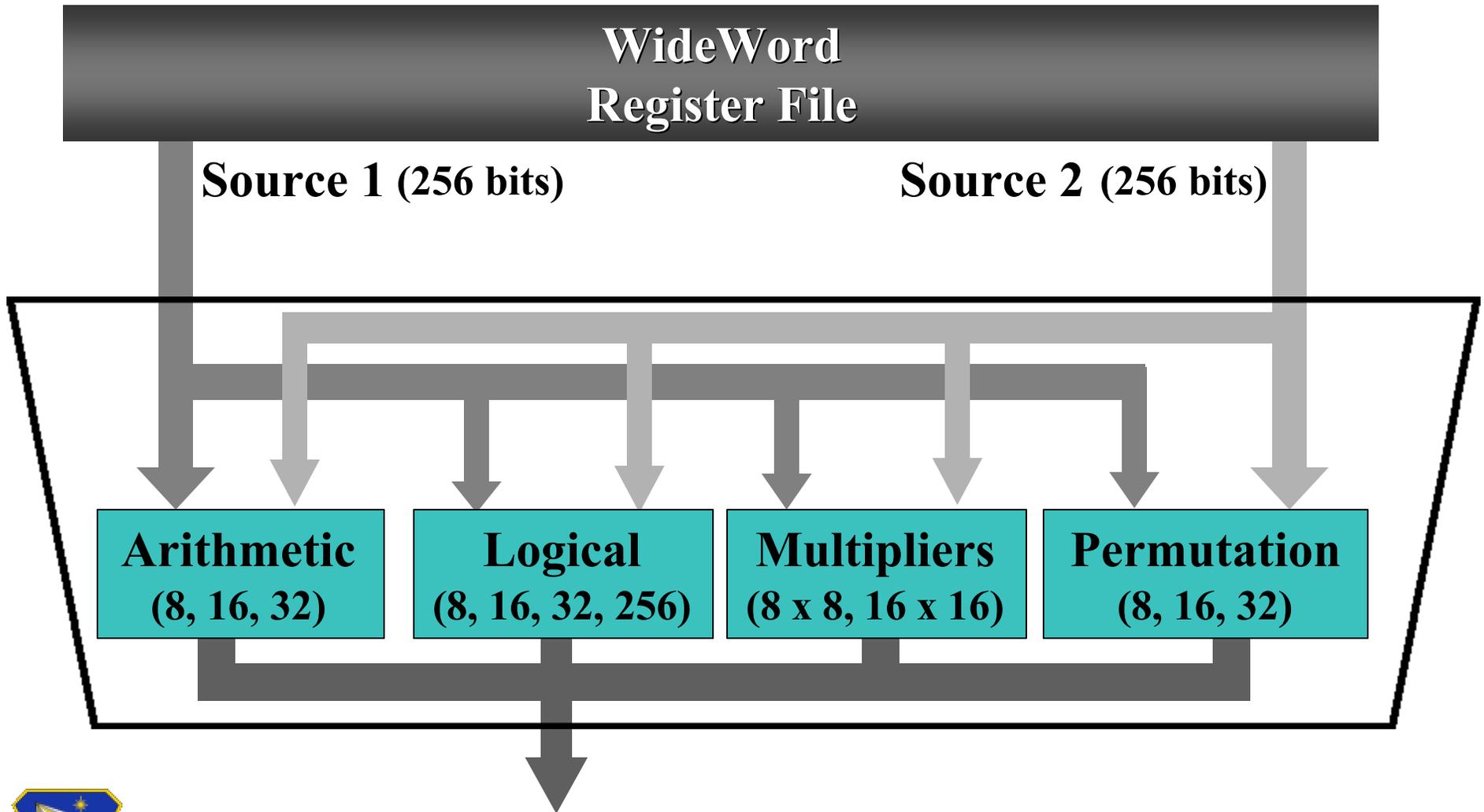


DIVA Node Architecture





WideWord ALU Data Flow

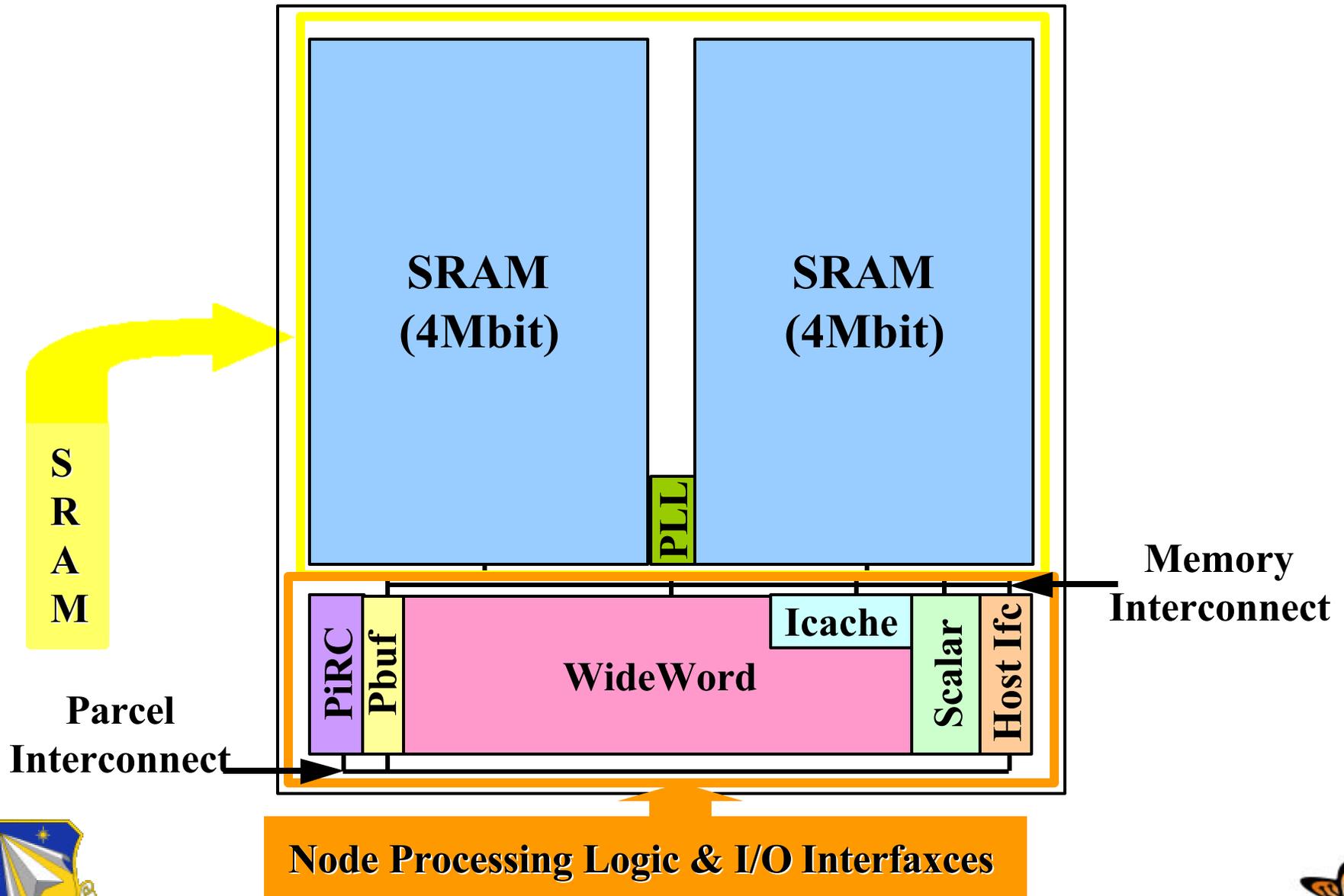


KISS: More compromises in architecture to enable early prototype



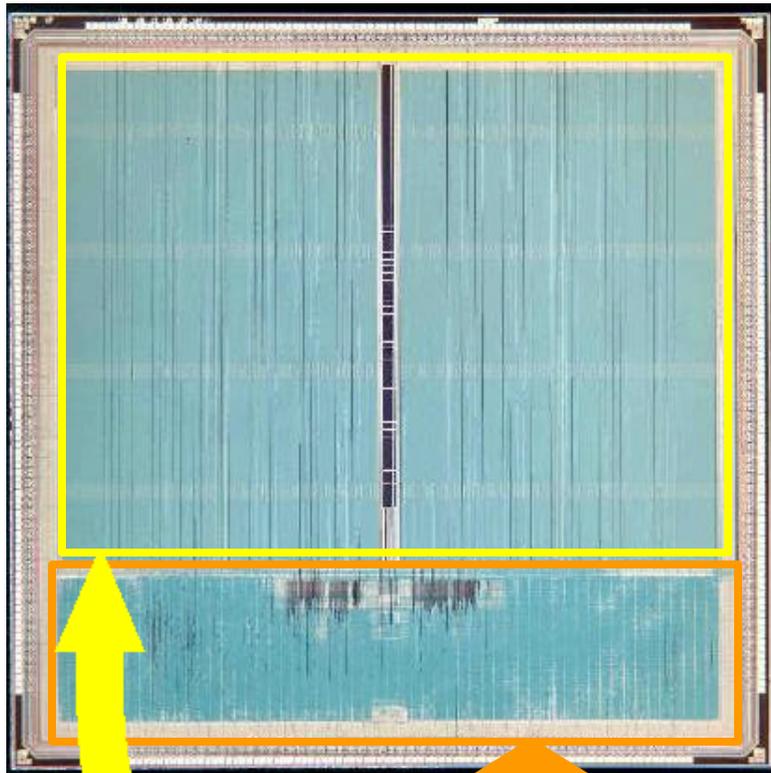


DIVA PIM Chip Floorplan





DIVA PIM Chip



SRAM

Node Processing Logic

- ◆ **Current lab measurements**
 - 640 MOPs (peak, 32-bit ops)
 - 0.8 Watts at 80MHz on cornerturn core loop
- ◆ **Purpose**
 - Demonstrate bandwidth advantages of PIM technology
- ◆ **Key architectural components**
 - High memory bandwidth
 - 256-bit WideWord processing
 - PIM routing component
- ◆ **Chip statistics**
 - 9.8mm X 9.8mm in TSMC 0.18mm
 - ~200K logic cells plus 8Mbit SRAM
 - 352 pins (241 signal pins)
- ◆ **Projected performance for 2nd prototype**
 - 1.6 GOPs
 - 2.5 Watts at 200MHz





HPPS & FPCA ARCHITECTURES

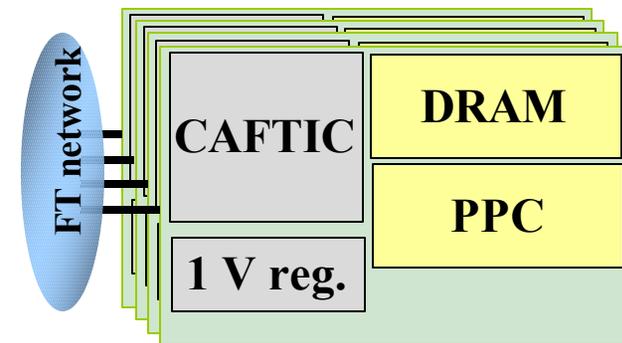
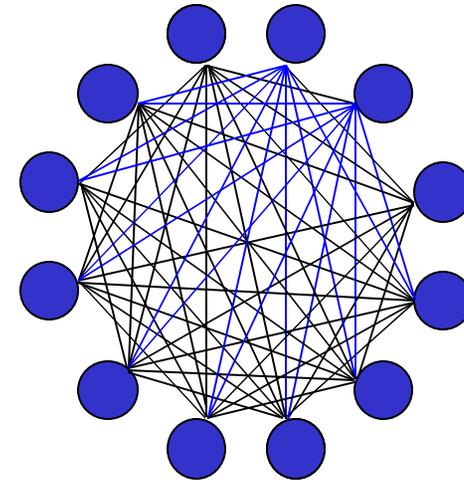




High Performance Processor System

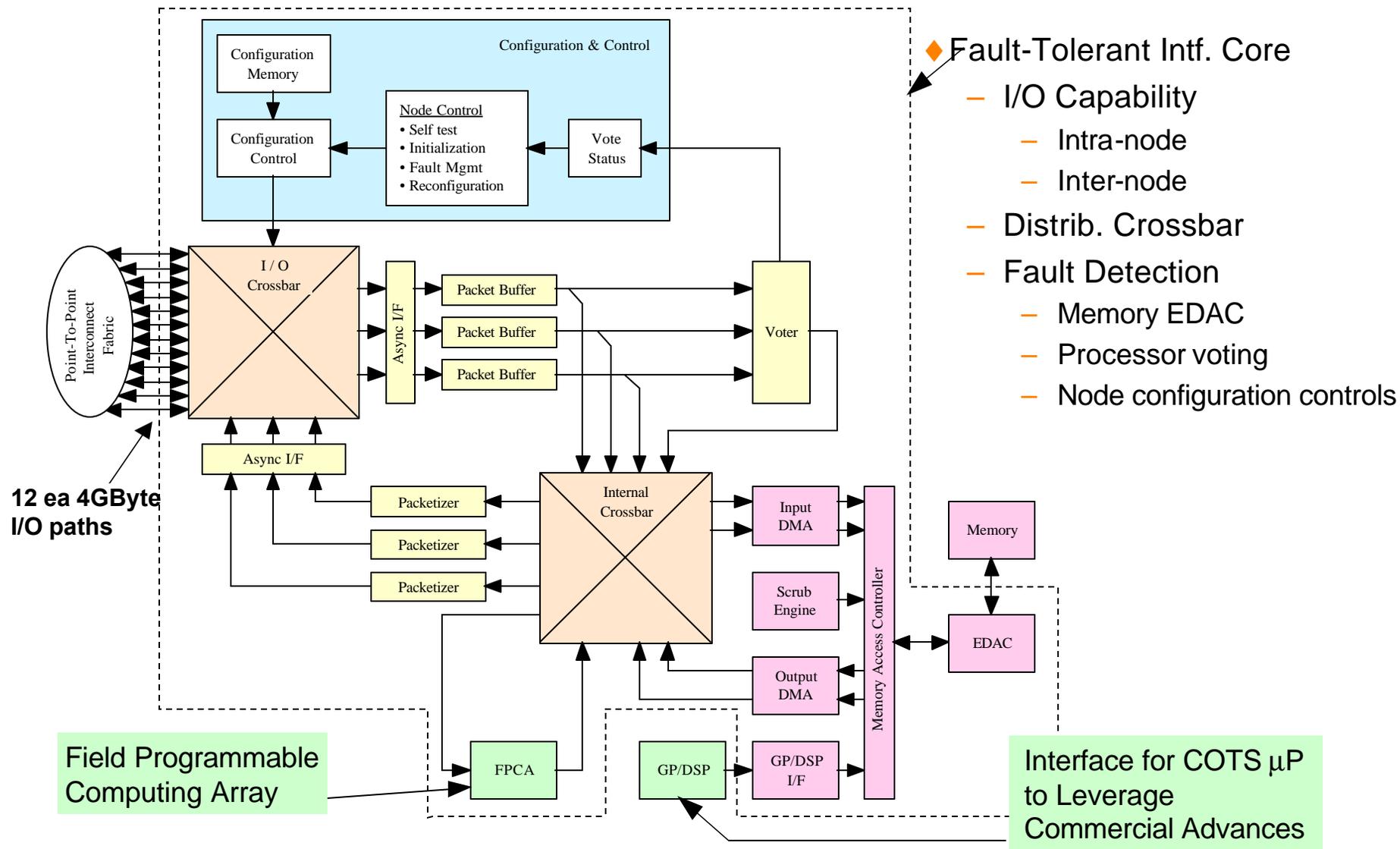


- ◆ Multinode Processor
- ◆ One custom ASIC
- ◆ Innovative voting
- ◆ Inputs for high bandwidth A/D receiver channels or FPCA





HPPS Node Architecture

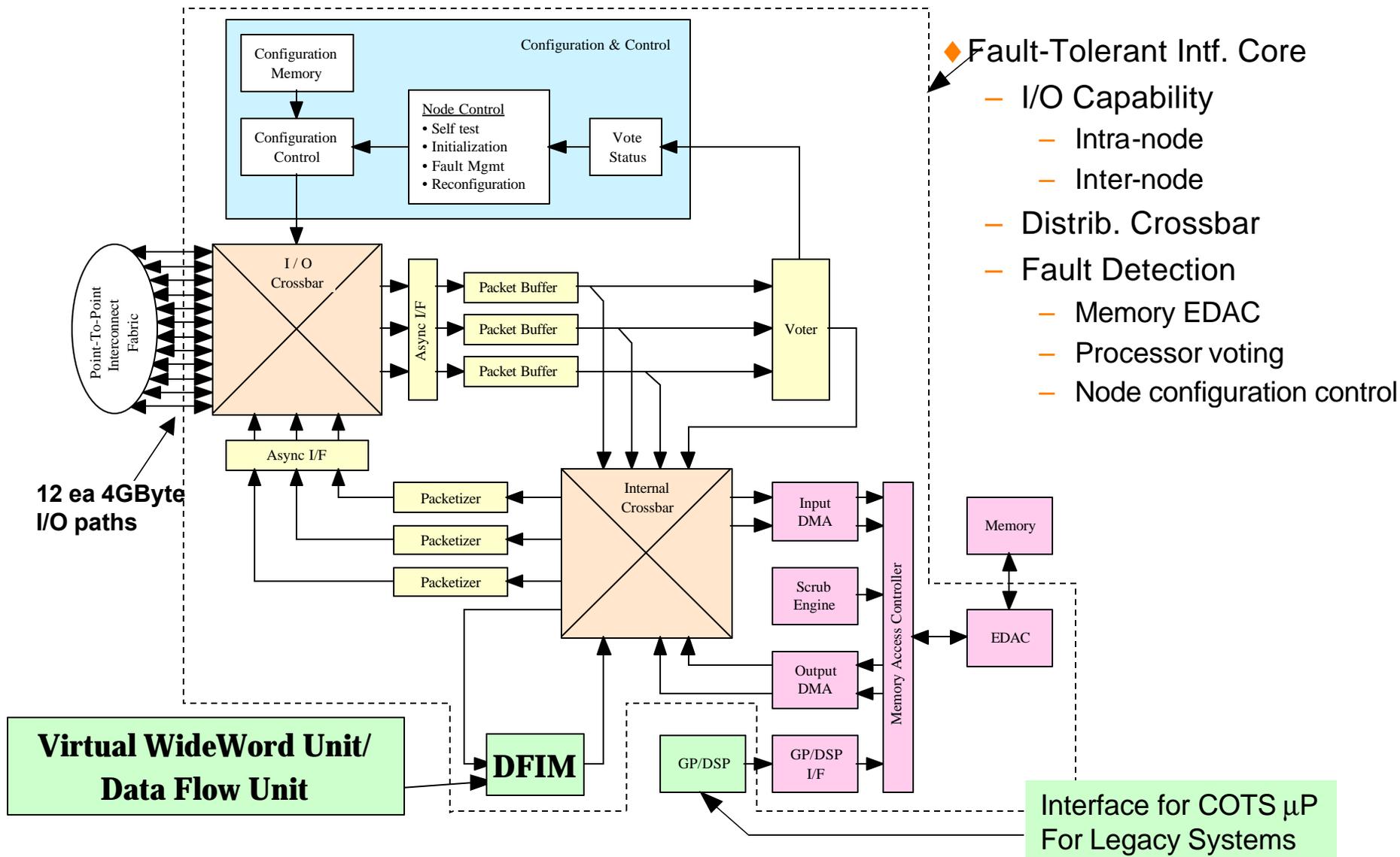




MONARCH: Node Architecture

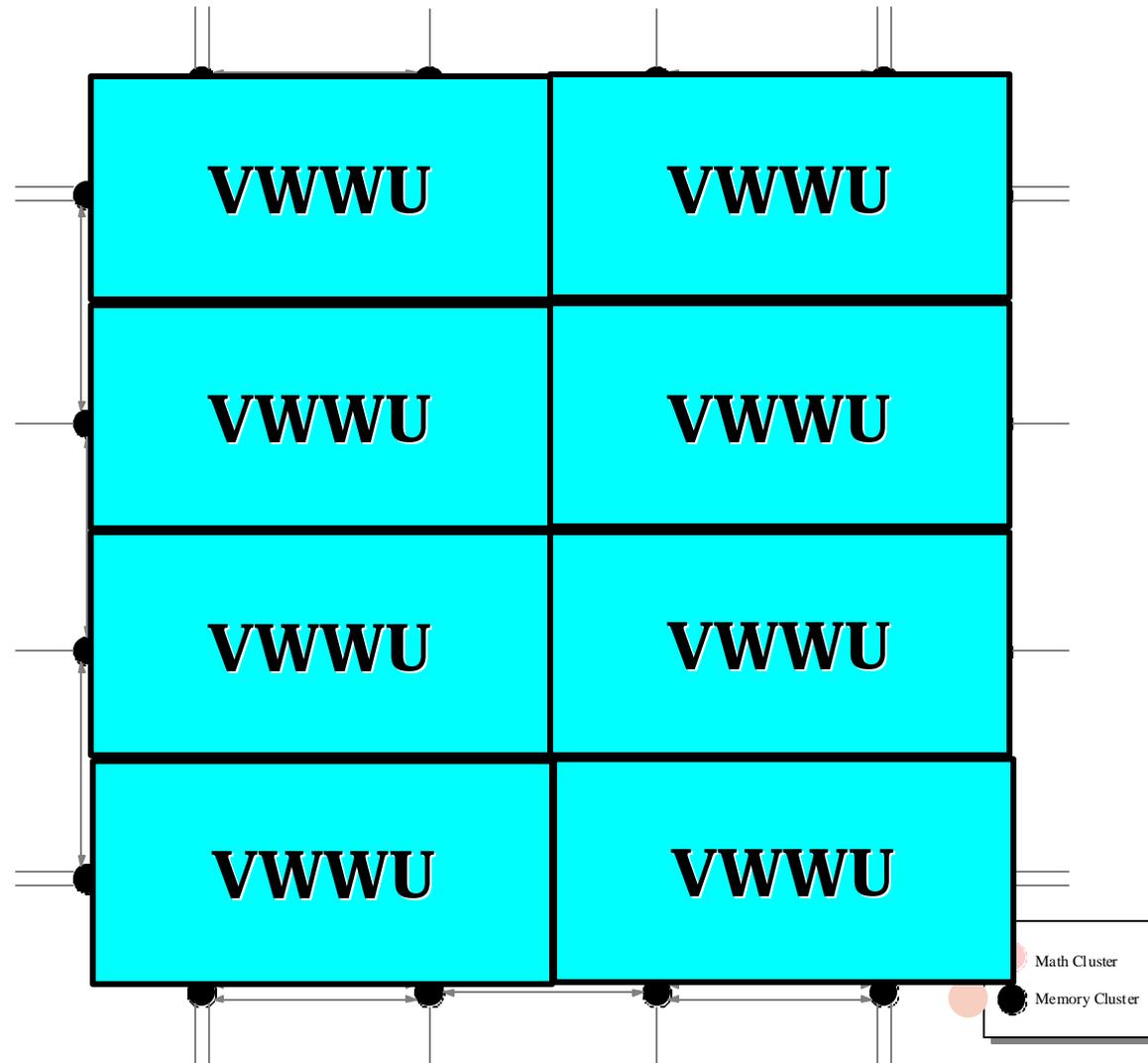


Note: Not to Scale





Virtual WideWord Unit/DFIM





MONARCH ARCHITECTURE

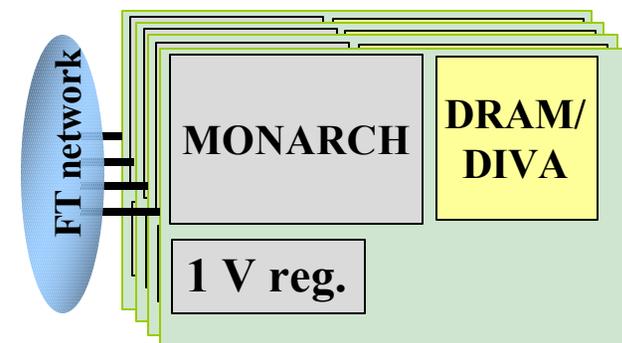
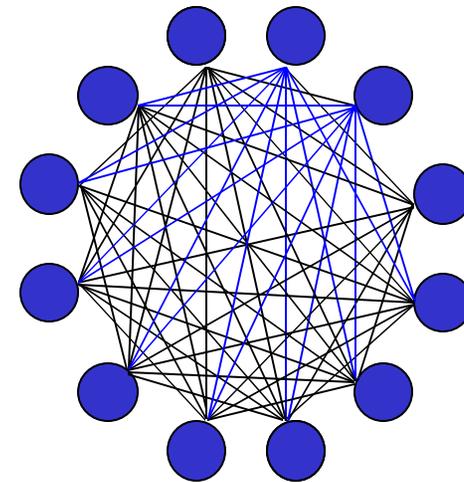




MONARCH Processor System

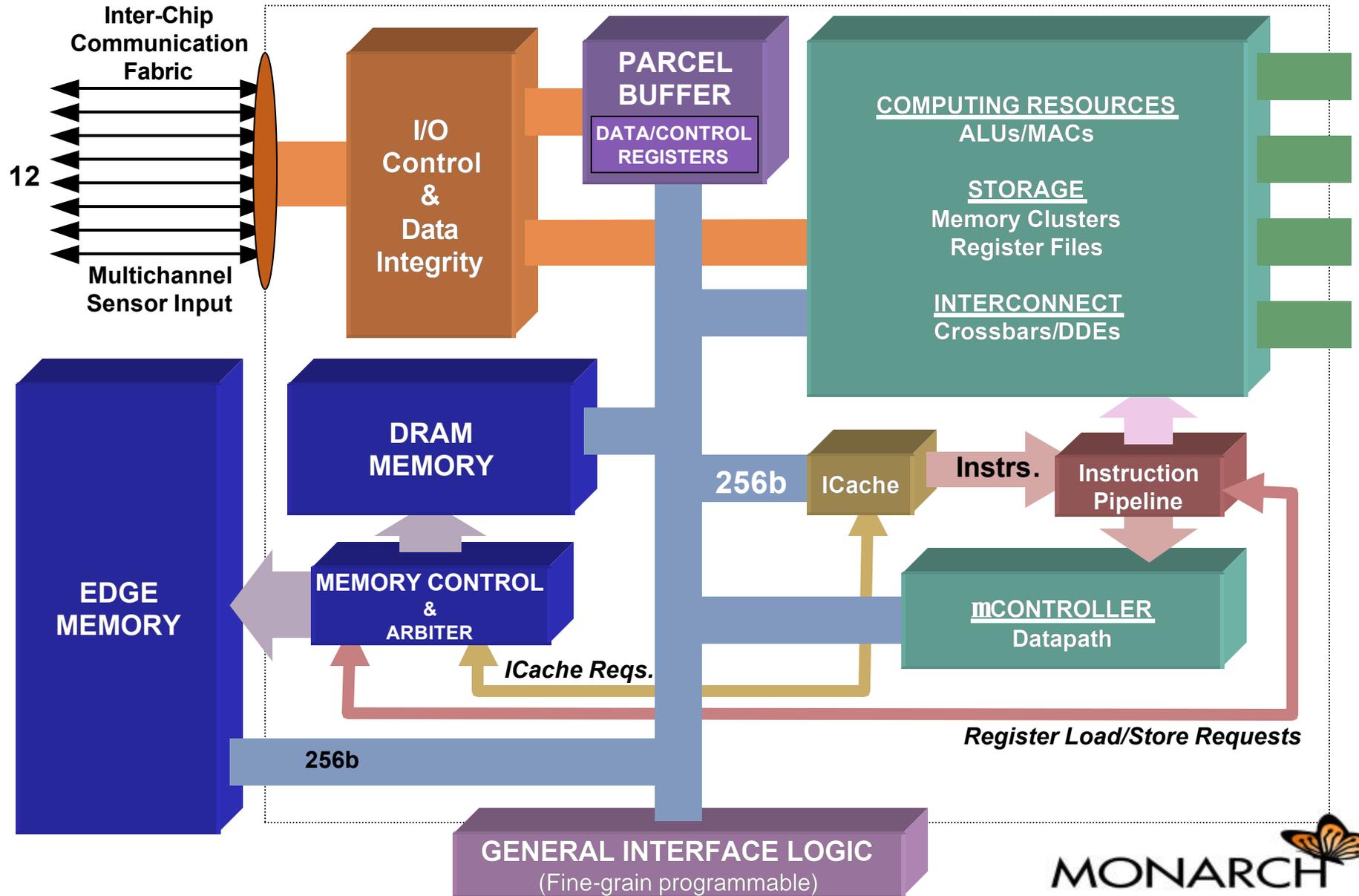


- ◆ Multinode Processor
- ◆ One MONARCH chip
- ◆ Innovative voting
- ◆ Inputs for high bandwidth A/D receiver channels or direct chip-to-chip data transfer



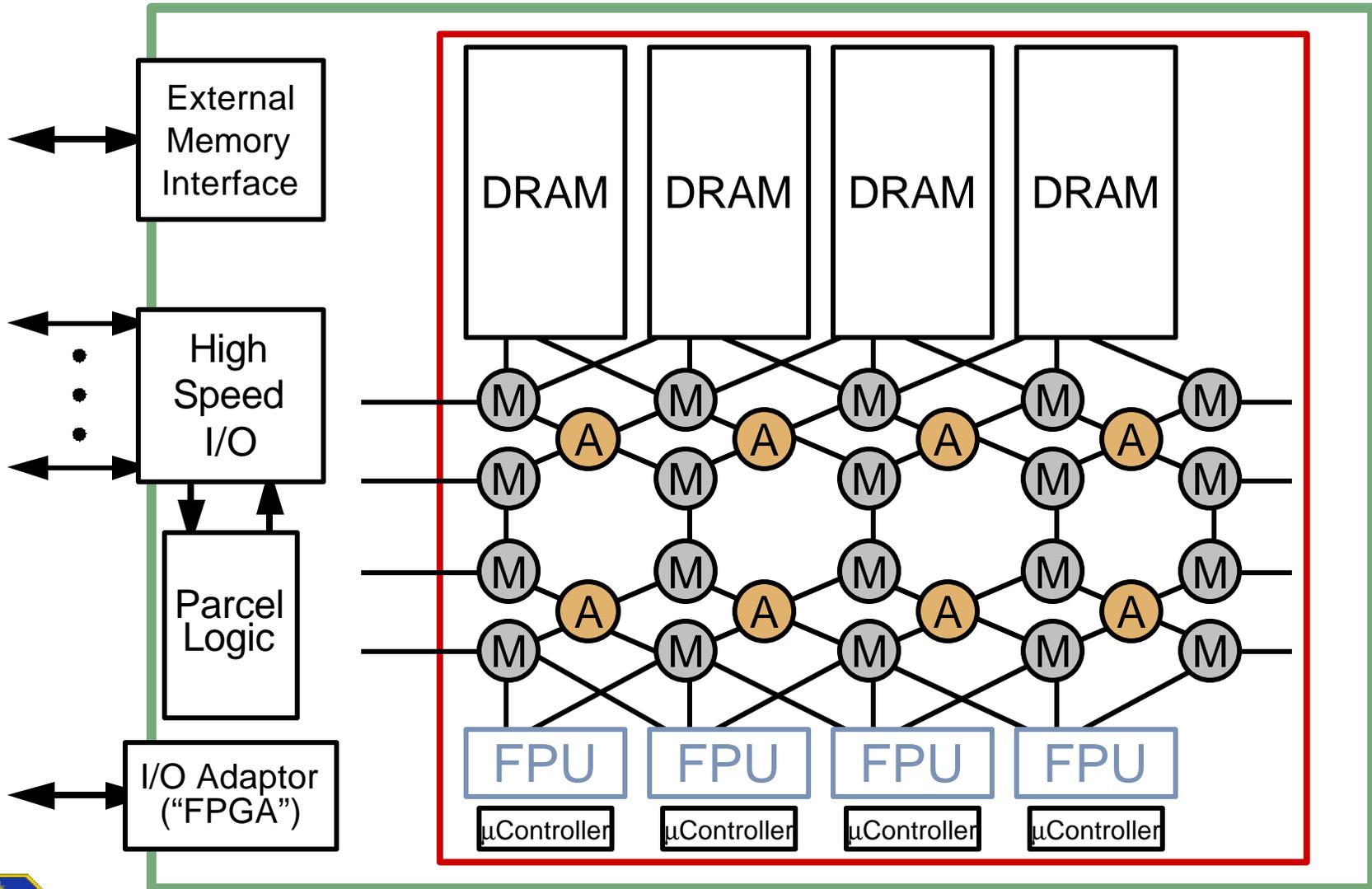


MONARCH Chip Overview



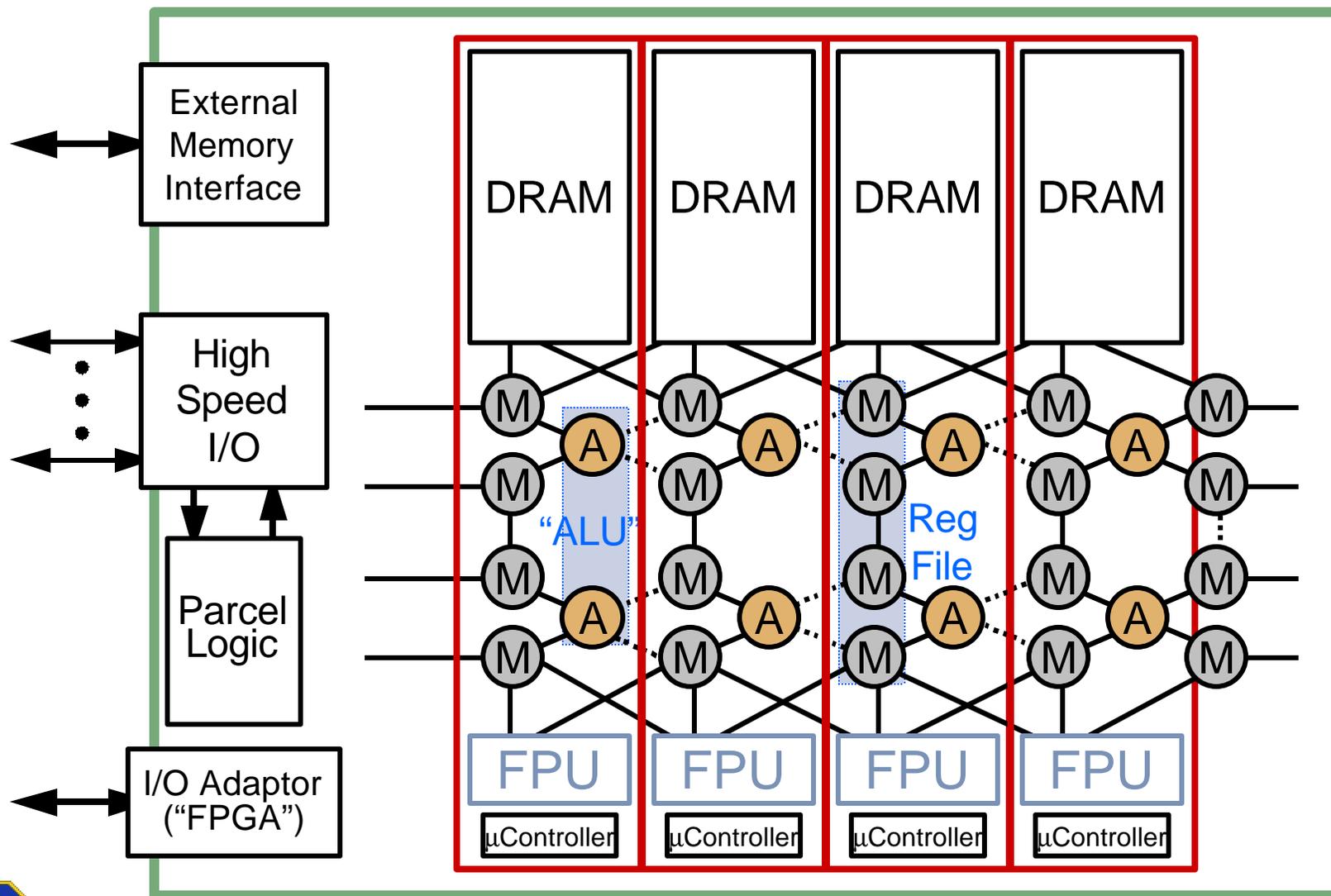


Native "Stream" Mode



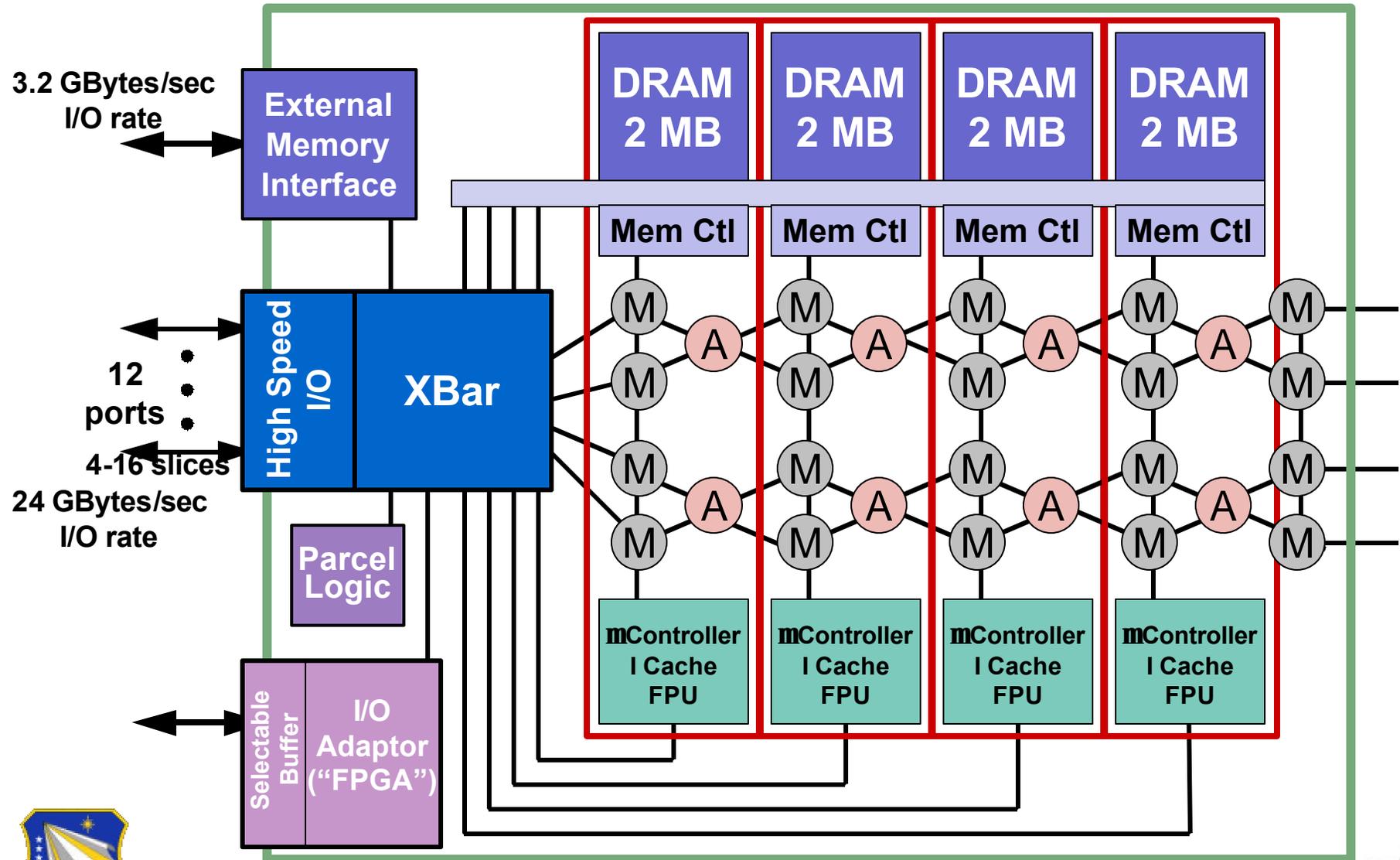


Native Threaded Mode





MONARCH Single Chip Architecture

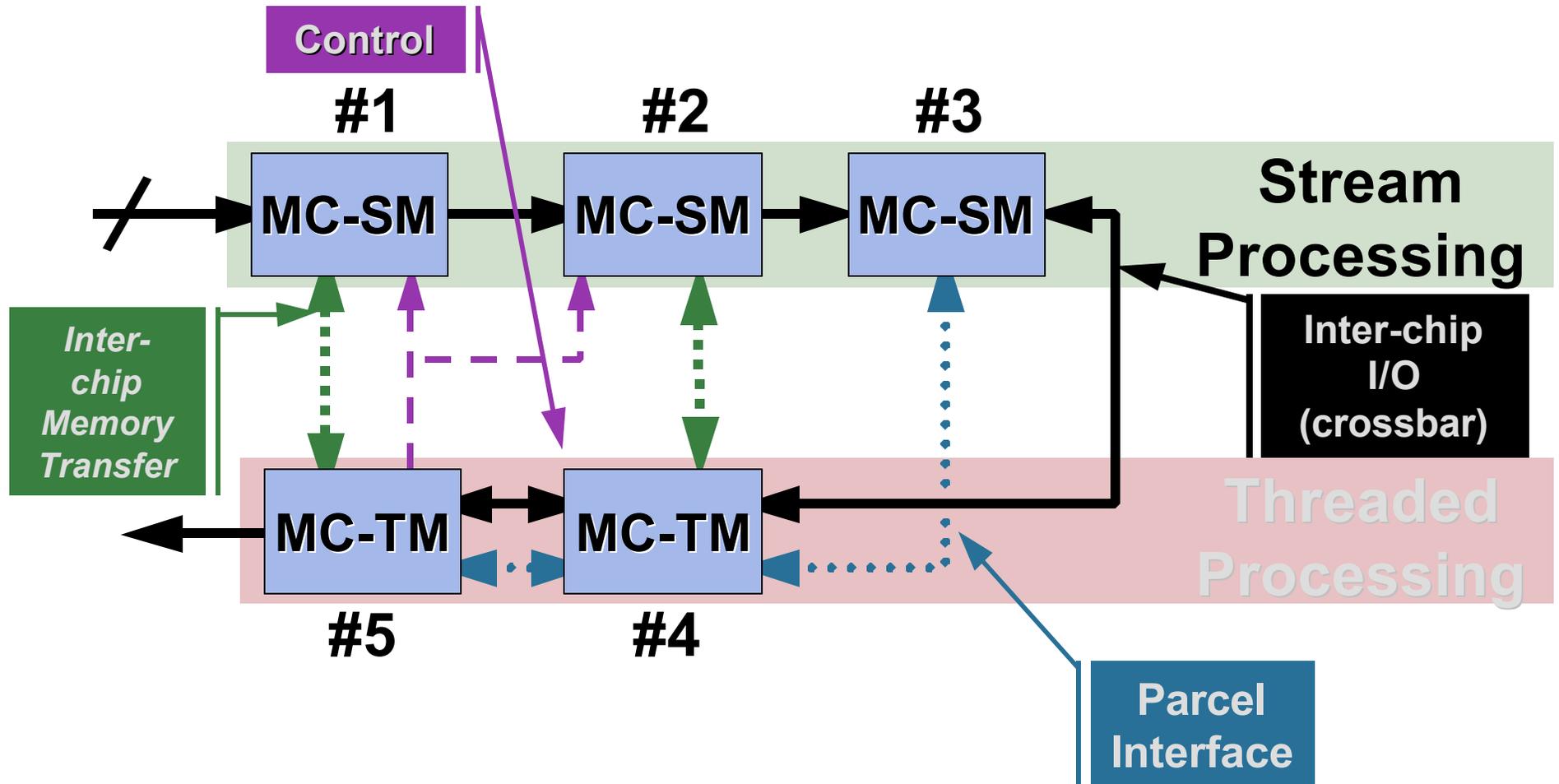


- ◆ 800 MHz Clock
- ◆ 12 GFLOPS
- ◆ 32 MBytes DRAM
- ◆ 256 MALU
- ◆ 512 ops/clock
- ◆ 400 GOPS
- ◆ 320KBytes SRAM
- ◆ 36 Watts





MONARCH Application Processor





MONARCH Architecture Features



- ◆ Dual native mode, high throughput computing
 - Multiple wide word threaded (instruction flow) processors/chip
 - Highly parallel reconfigurable (data flow) processor
- ◆ Large on chip, multiport memories
 - High bandwidth access to memory
 - Extensible with off chip memory
- ◆ High speed, distributed cross bar I/O
 - Integrated with chip processing
 - Scalable I/O bandwidth - multiple topologies
 - Direct connect to high speed I/O devices, e.g., A/D's
- ◆ Rich on chip interconnect
 - Supports on chip topology morphing and fault tolerance
 - Supports multiple computation models (SISD, SIMD, DF, SPMD,...)
- ◆ On chip Morph - Program bus and microcontrollers





Architecture Merger* Features

- Mostly a complementary match and enhancement -



ISSUE	APPROACH	BENEFIT
256 bit wide word processing unit	Each Arithmetic Cluster has 8 32 bit units	1 AC provides same width as WW unit
Instruction Set Mapping	Basic functions same Need to add some insts	Little impact
Large On-chip memory	Similar to Edge Memory Now can have on chip	Performance boost
5 State pipeline, instruction flow decoder	Retain, and mux decoded signals with DF signals	Some hardware growth, but more control modes
Data flow control mode - streaming	Retain - switch mode bit	As above
High speed, multiple channel I/O	Incorporate dist. xbar and use for parcel com	Improved I/O performance
Parcel communications	Retain and map onto other physical protocol	Little impact
On-chip micro controllers	Retain	Performance boost



* Merger of features from DIVA and HPPS processors





Architecture Merger Issues

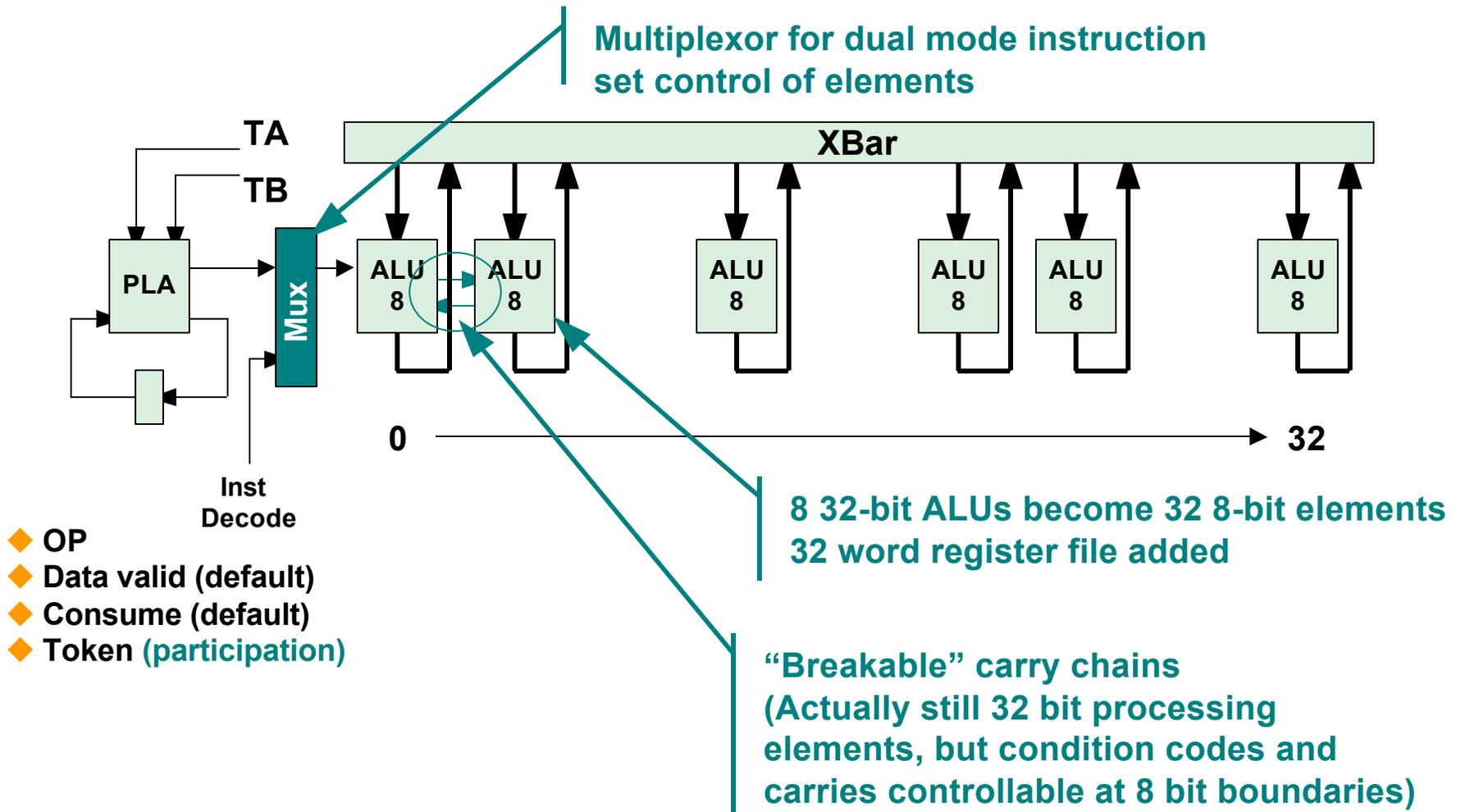


ISSUE	APPROACH	IMPACT
WideWord 8-bit math	Modify array carry-chain logic	Negligible delay
Thread control for array / WideWord	Switch RISC pipeline control into array	TBD
3-port WW register file implementation	Extend array arithmetic clusters	Small area increase
WideWord pipeline length / bypass	TBD / Simulation	Interconnect, Compiler
Minimum I Cache size	Simulation	Area
Data exchange: $W \rightarrow S$ / $S \rightarrow W$	TBD	Area, Interconnect
I/O: Memory map or program?	Memory Map	None
WideWord shifter implementation	TBD (modify array)	Design complexity
Permute implementation	Enhance array x-bars for 8 bit data	Small area increase





FPCA Changes for WideWord





MONARCH Pin Estimate

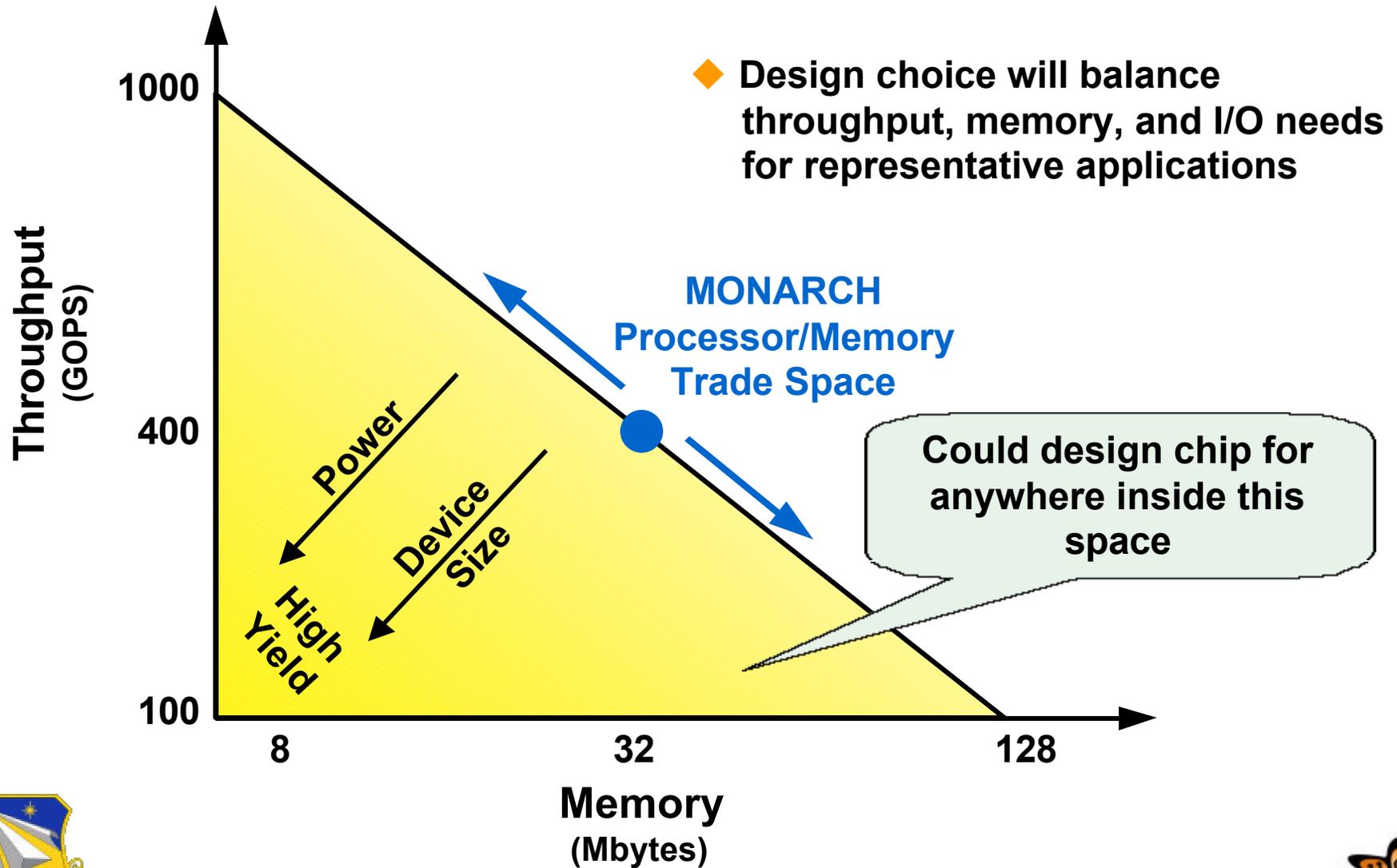


MONARCH I/O Summary					
	number of ports	Wires per port	Total Wires	Type	Clock Rate
High speed ports	12	50	600	LVDS	1 GHz
Inter FPCA Links	4	52	208	LVDS	1-2 GHz
External memory	1	160	160	CMOS	500 MHz
Standard I/O	2	60	120	variable	100+ MHz
Total			1088		





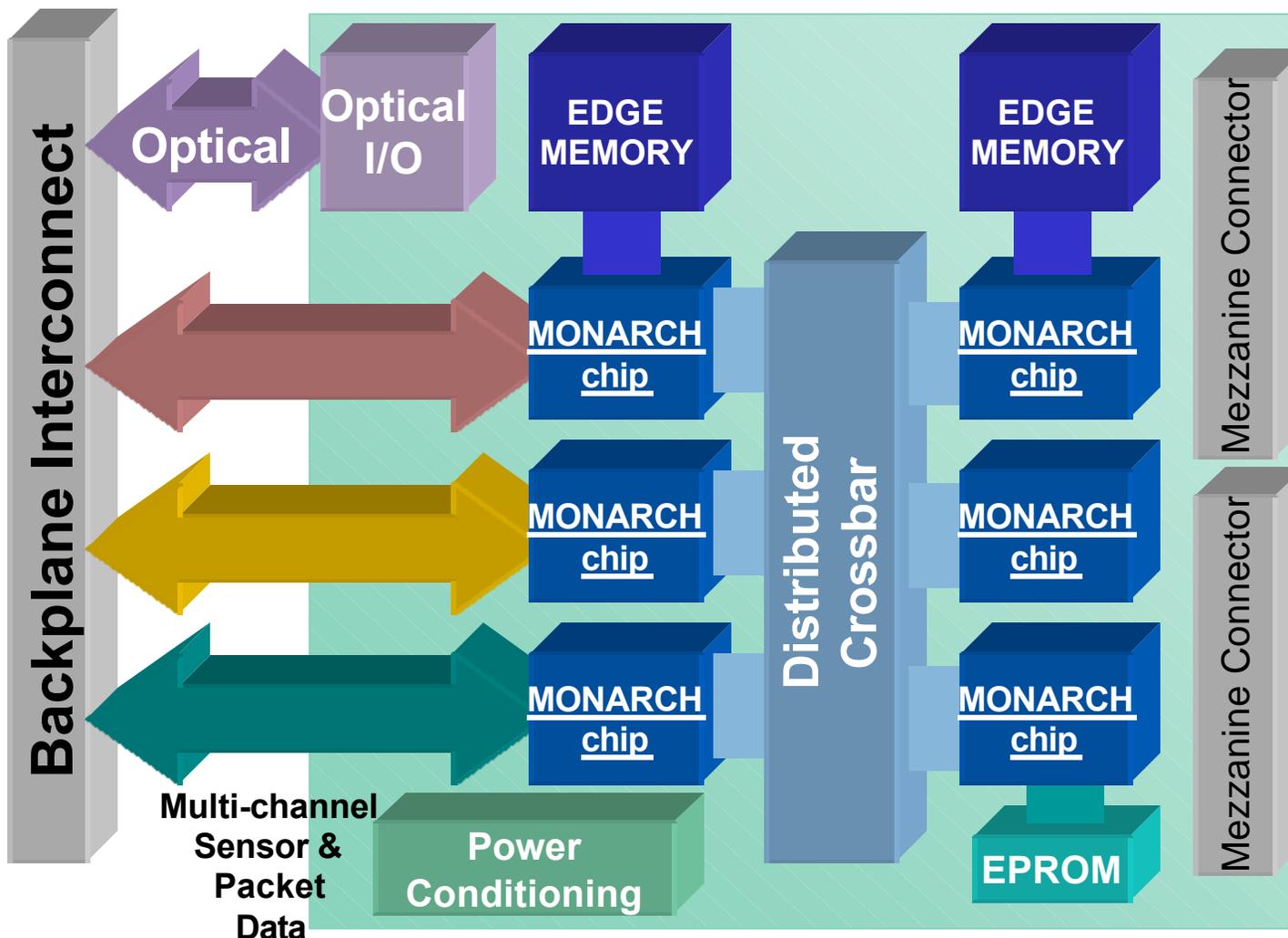
Need to Select Preferred Parameters for 1st MONARCH Chip





MONARCH Processing Card

- 6Ux160 double euro card form factor -



- ◆ 6 MONARCH chips + memory and power conditioning
- ◆ 75 GFLOPS
- ◆ 2.4 TOPS
- ◆ 192 MBytes on-chip DRAM
- ◆ 2 MBytes on-chip SRAM
- ◆ 1 GBytes on-board memory

MONARCH



Summary & Conclusions



- ◆ MONARCH features very attractive for multiple applications
- ◆ Merger of two existing architectures shows good fit
 - “Complementary” but compatible features
 - Rich experience base allows quick design trades
- ◆ “The devil is in the details” --- a lot more work
 - On-chip DRAM organization and access
 - Support for “morphing”
 - Simulation results at application-level
 - Trade offs for FPU capability

