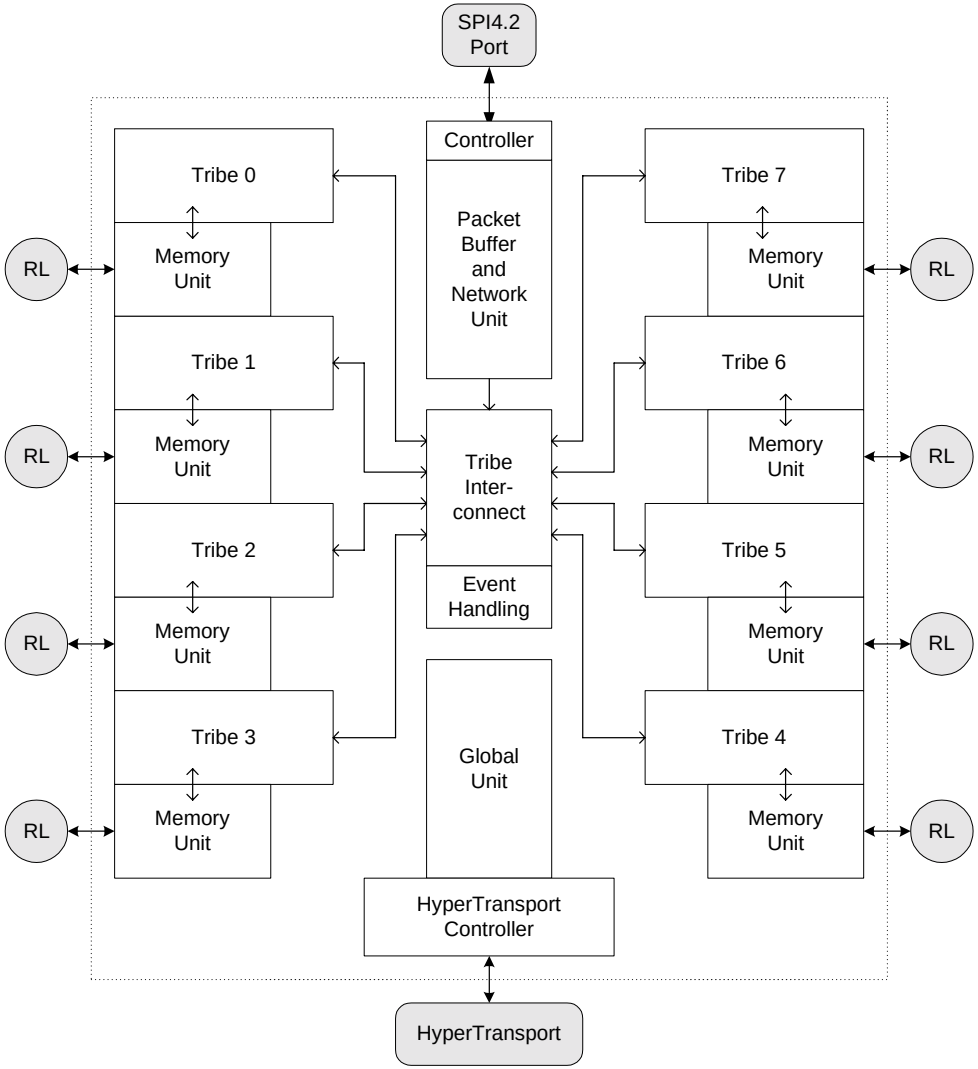
A Massively Multithreaded Packet Processor

Steve Melvin
Mario Nemirovsky
Enric Musoll
Jeff Huynh
Rodolfo Milito
Hector Urdaneta
Koroush Saraf

Outline

- "Porthos" Chip
- Motivation
  - Application Requirements
  - External Memory Accesses Rate
  - Configuration of Multiple Memory Ports
- Porthos Microarchitecture
  - Tribe
  - Network Unit
- Project Status
- Summary

“Porthos” Block Diagram



Application Requirements

- *Stateful* Applications
- Memory Access Characteristics
  - Packet Rate (1 – 10MPPS)
  - Memory Accesses per Packet (100s – 1000s)
  - Little Locality (low spatial, low temporal)
  - Large Footprint (100s of MB)
- Packet Rate vs. DRAM Cycle Time:
  - 100 – 200 Packets Required to Cover Latency
- But Packet Dependencies Do Exist

Application Requirements - Summary

- Flexibility and Programmability is Critical
  - General Purpose ISA with Instruction Cache
  - Not Microengine with Control Store
- Multithreaded Design
  - Can’t Use Standard Processor Cores
  - SMT Adapted to Large Numbers of Threads
  - Size and Power Constrained
  - Efficient Handling of Packet Dependencies Critical
- Attention to Memory Configuration

REAPS

- *Random External Accesses Per Second*
- More Critical than Memory Bandwidth
- Measurements
  - Overall REAPS Capability vs. Application Requirements
  - MIPS/REAPS Ratio
- High REAPS Suggest
  - Multiple DRAM Ports
  - Narrow Buses
  - Short Bursts
  - Memory Technology That Supports Fast Address Rates

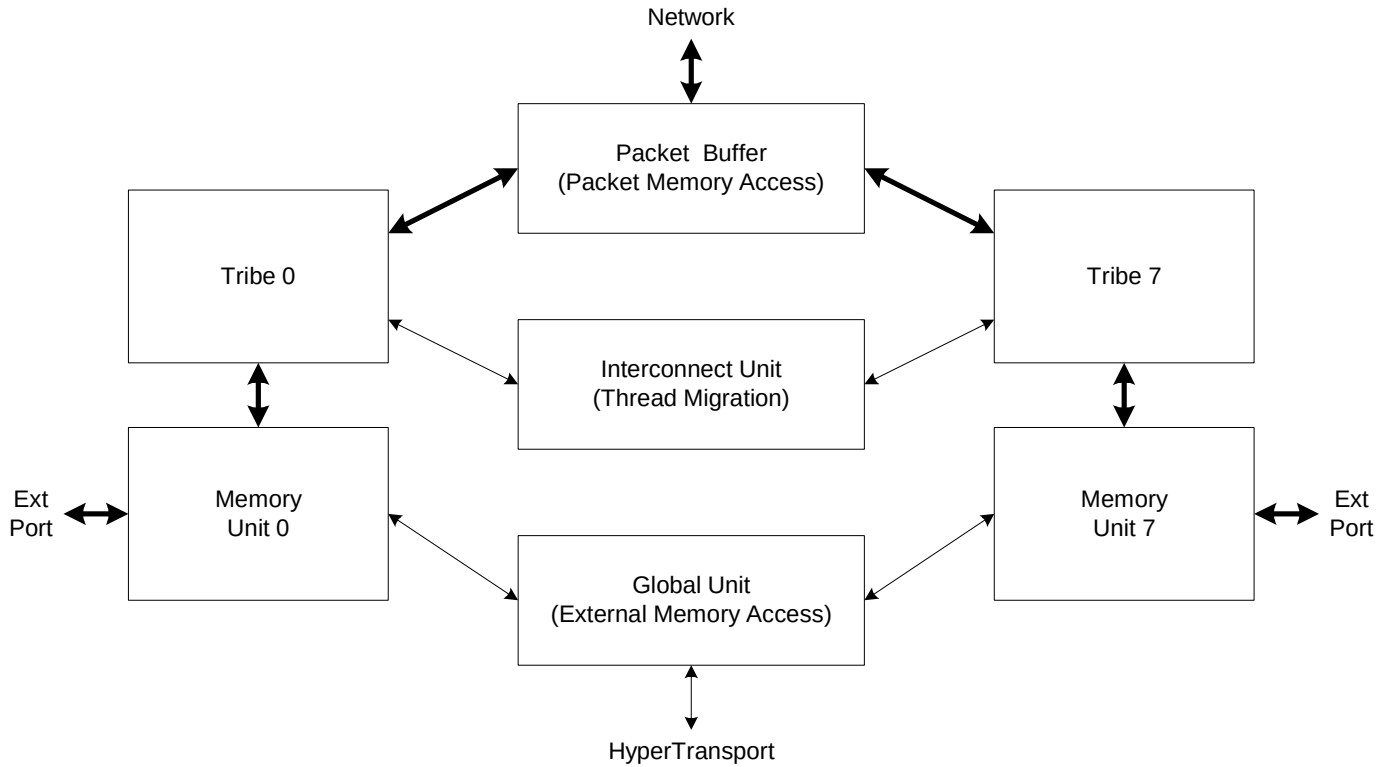
MIPS/REAPS Analysis Summary

Application Requirements at OC48, 500b packets			
App.	*Inst/packet*	*GIPS*	*GREAPS*
CEF	70	0.09	0.01
Classification	400	0.50	0.05
Re-assembly	400	0.50	0.08
Monitoring	2500	3.13	0.38
IDS	3500	4.38	0.51
Processor Capabilities			
Intel, 2Ghz		3.0	0.08
SB1250		2.4	0.16
Porthos		9.6	2.40

Processing/Memory Configurations

- Independent
  - Processing and Memory Uncoupled
  - Internal Switch
- Coupled
  - Processing Dedicated to Memory Port
- Loosely Coupled
  - Preferential Memory Port
  - Flat Memory Space and Switch for Non-local Accesses

Loosely Coupled Configuration
(Two of Eight Porthos Tribes Shown)



Tribe Microarchitecture

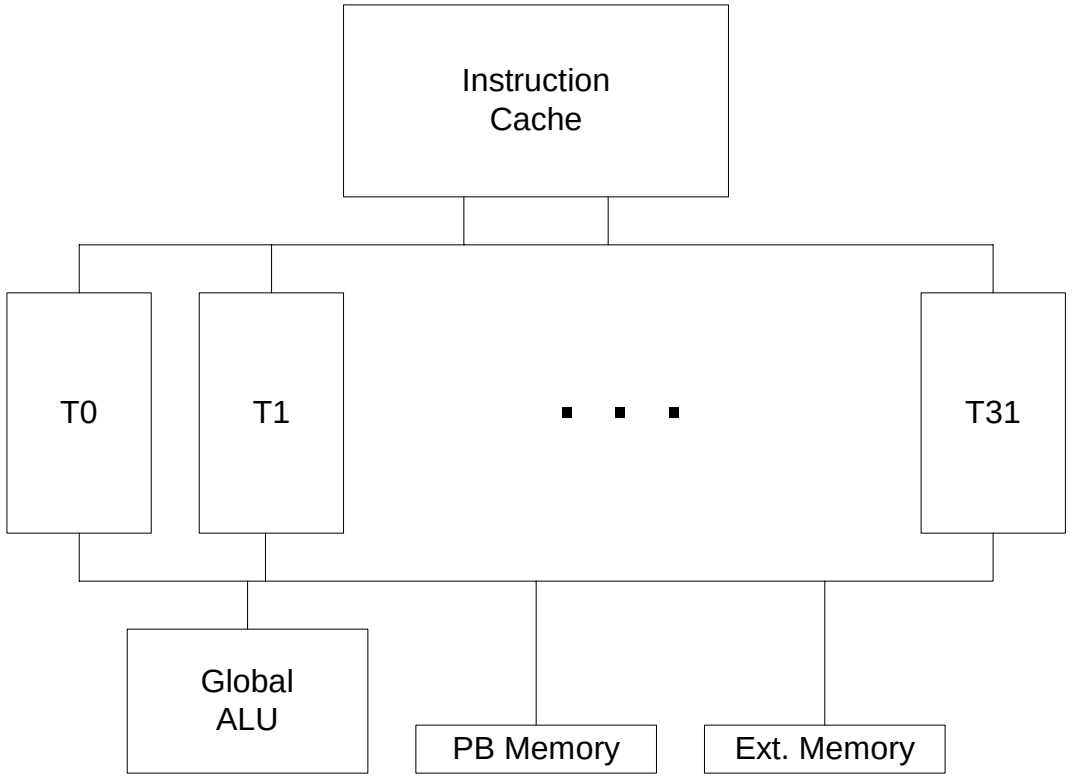
- 32 Threads Execute Simultaneously
  - Three Decoupled Pipelines (SMT / SMP / SMT)
  - General Purpose ISA (MIPS-like)
  - Each Thread Has a Separate Register File
- Optimized for Silicon Area and Tribe IPC, Not Thread IPC
  - No Branch Prediction
  - No Out of Order Execution
  - But Non-Dependent Instructions Can be Concurrent
- Two Memory Ports (Packet Buffer and External Memory)
- Balance of Local ALUs (32) and Global ALU

Tribe's Three Decoupled Pipelines

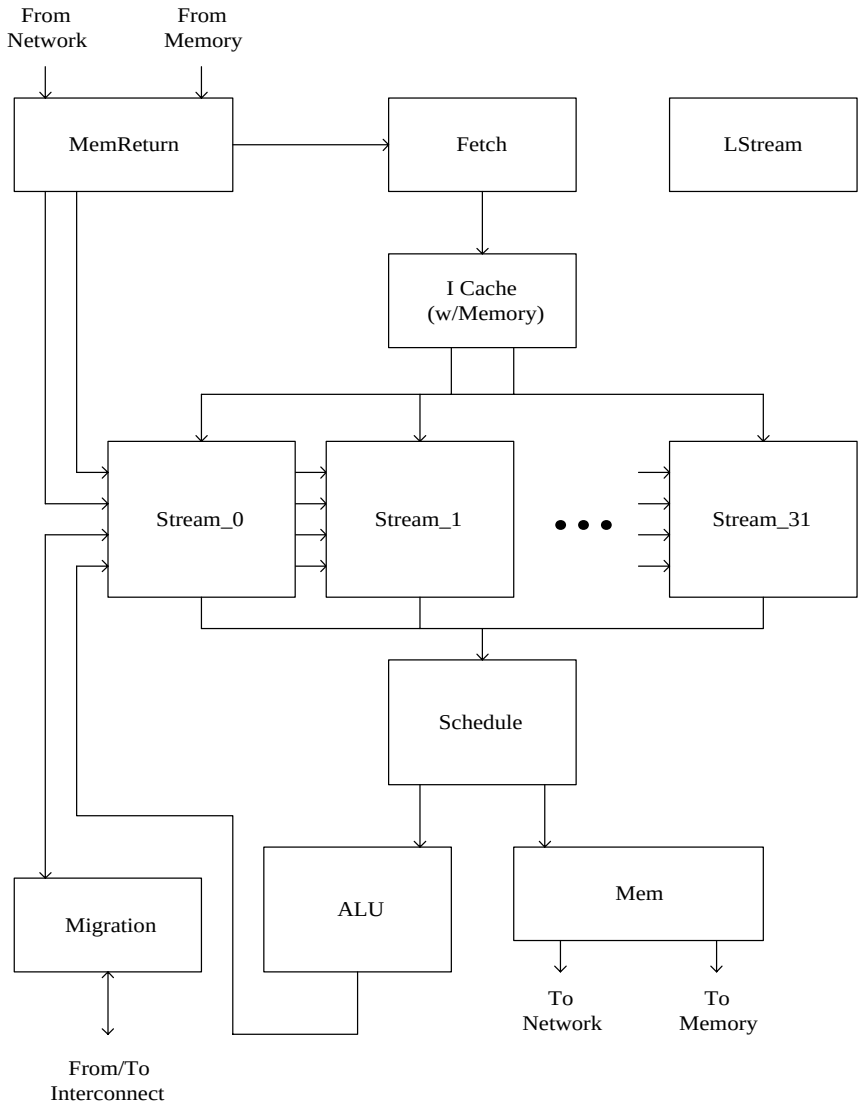
SMT
8 IPC Peak
Fetch 4 inst. each
from 2 Threads

SMP
32 IPC Peak
Independent
execution

SMT
3 IPC Peak
3 Threads



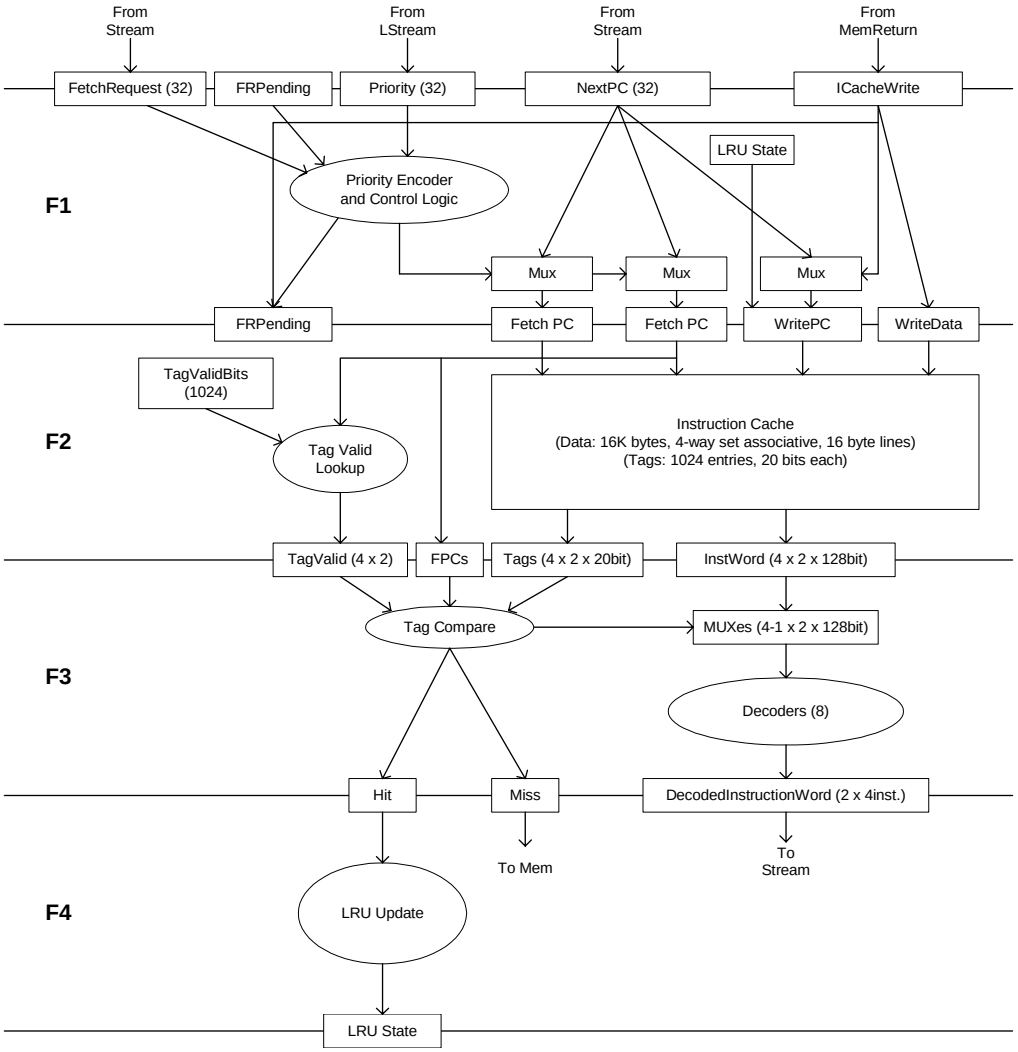
Tribe Block Diagram



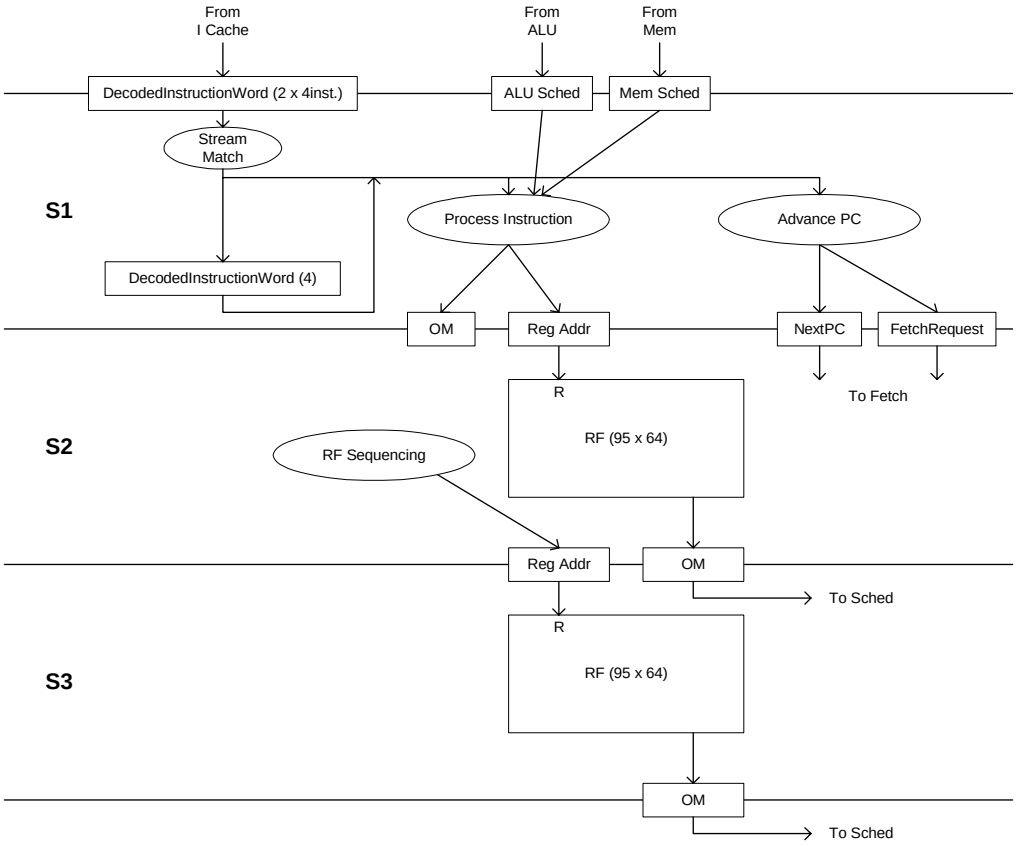
Tribe Per-Thread Resources

- Tribe Register Files
  - 1 Port per Thread (32 Ports Total) – Small Area
  - ½ Port per Thread Alternative Implementation
- Stream ALUs
  - Instruction Distributions are Very Skewed
  - ADDIU, Branch and some Logical
  - 60% of Dynamic Frequency
  - No Shift, Multiply/Divide, 64-bit Arithmetic

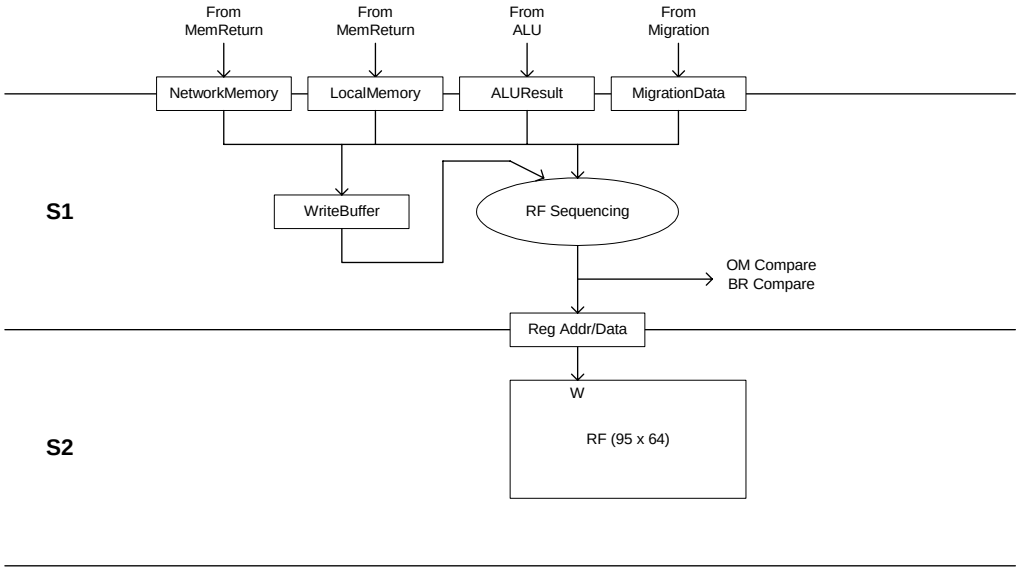
Fetch Pipeline



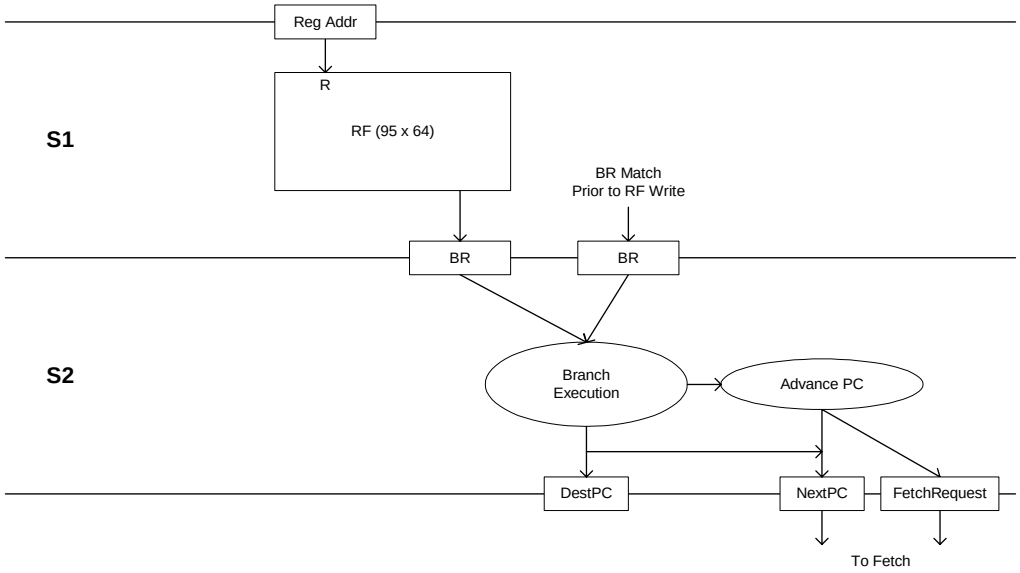
Stream Pipeline – Operand Read



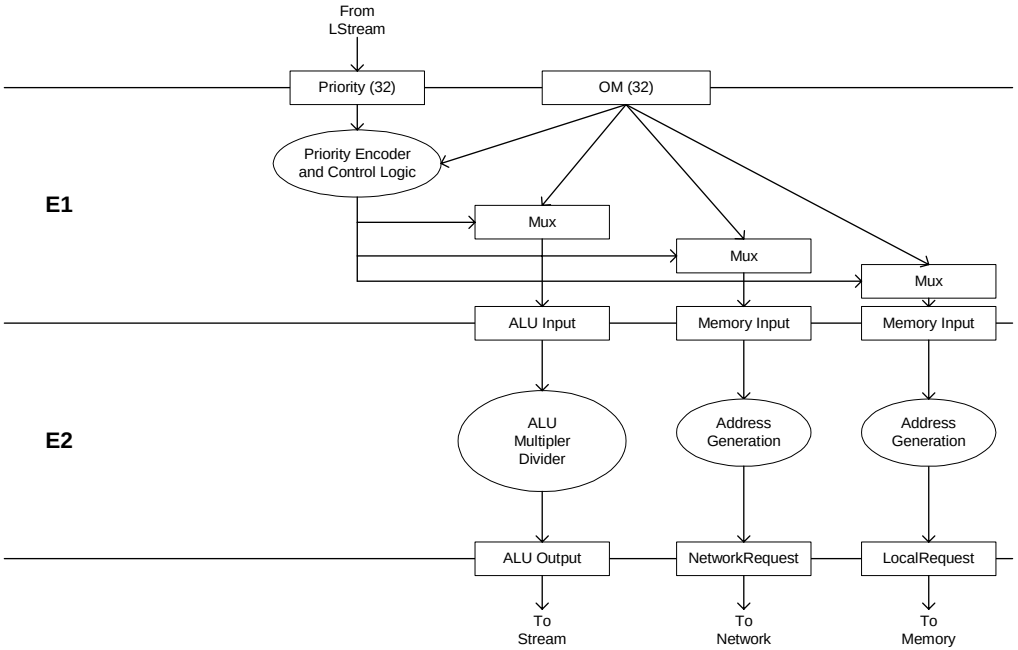
Stream Pipeline – Register Write



Stream Pipeline – Branch Resolution



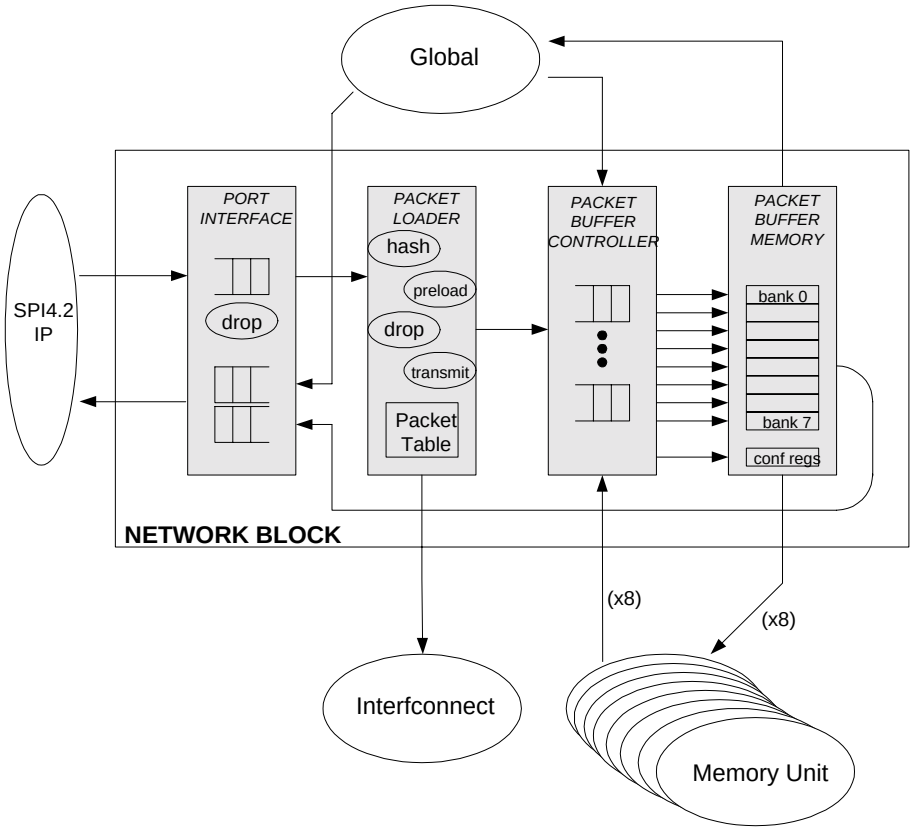
Execute Pipeline



Thread Scheduling Summary

- Used in Two Places Independently
  - Fetch: Two Threads to Fetch
  - Execute: ALU, Packet Buffer and External Memory
- Primary Policy: Seniority Scheduling
  - Not Fair
  - Favors Oldest Packet
- Software Override
  - Background and Interrupt Threads
  - Non-starved Priority Levels

Network Block Diagram



Project Status

- Implementation Details - Projected
  - 0.15 um, 300Mhz
  - 120 mm^2, 12W
- Status
  - RTL >90% Completed on All Major Blocks
  - Some Synthesis and Timing Analysis
  - Extensive Performance Analysis, Software, System Design
- Performance
  - 9.6 GIPS Sustained (4 IPC per Tribe)
  - 80 MIPS/mm^2 , Higher Possible
    - Scalable to Higher DRAM and Clock Frequencies
    - Tribe and Memory Unit Size Projected at 6.5 mm^2

Summary

- Fundamental Application Requirements Dictate
  - 100s of Threads
  - Multiple DRAM Ports, Short Bursts
  - Loosely Coupled Processing/Memory Architecture
- Massive Multithreading
  - Hardware Support for 100s of Threads
  - Small Number of Tightly Integrated SMT Engines
  - General Purpose ISA
  - Single Threaded Performance is Sacrificed