



# The Alpha 21364 and 21464 Microprocessors: Continuing the Performance Lead Beyond Y2K

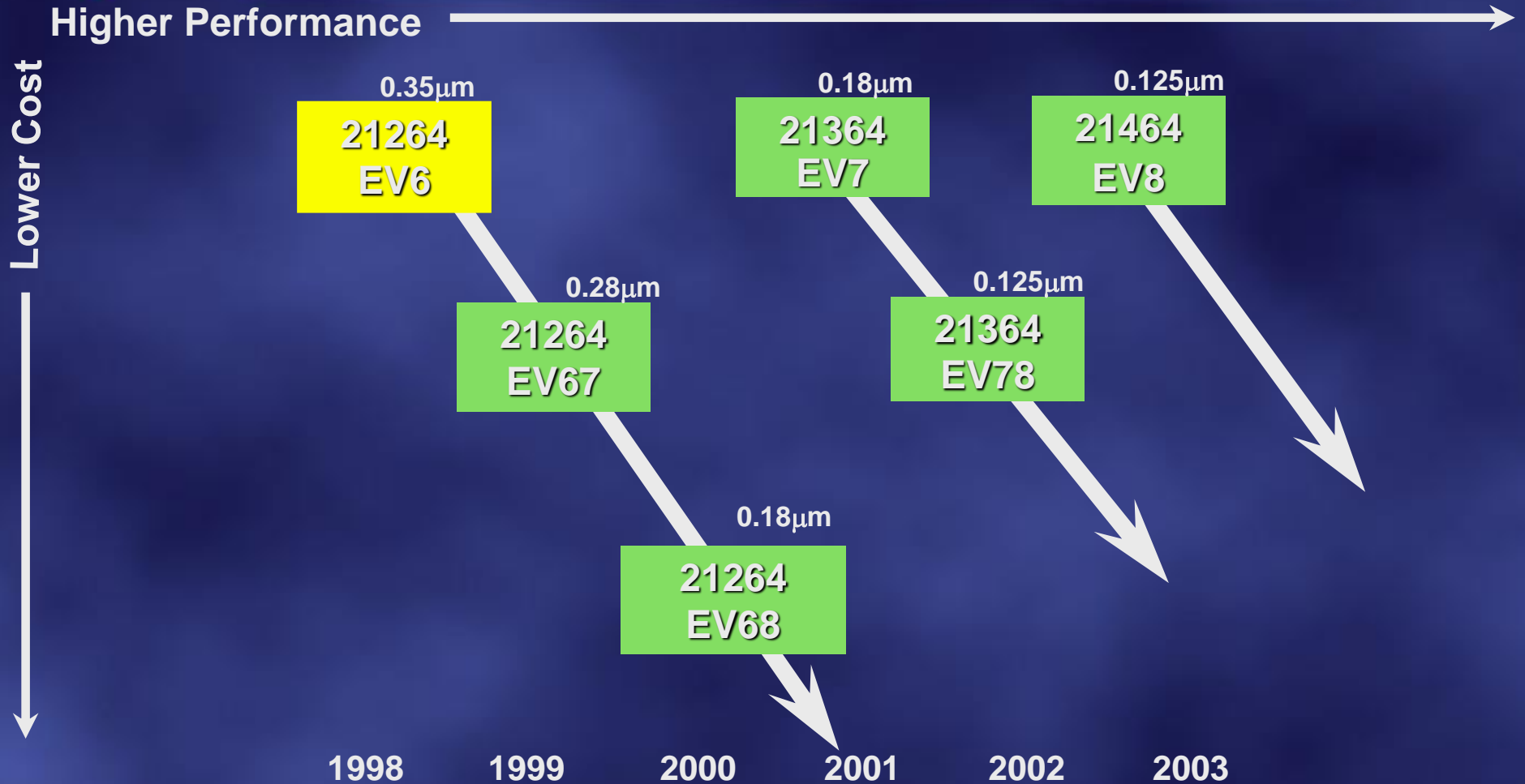
*Shubu Mukherjee, Ph.D.*

Principal Hardware Engineer  
VSSAD Labs, Alpha Development Group  
Compaq Computer Corporation  
Shrewsbury, Massachusetts

Slides: 1998 Microprocessor Forum (Peter Bannon) and 1999 Microprocessor Forum (Joel Emer)



# Alpha Microprocessor Roadmap





# Alpha 21264 Microprocessor

## ◆ Architectural Features

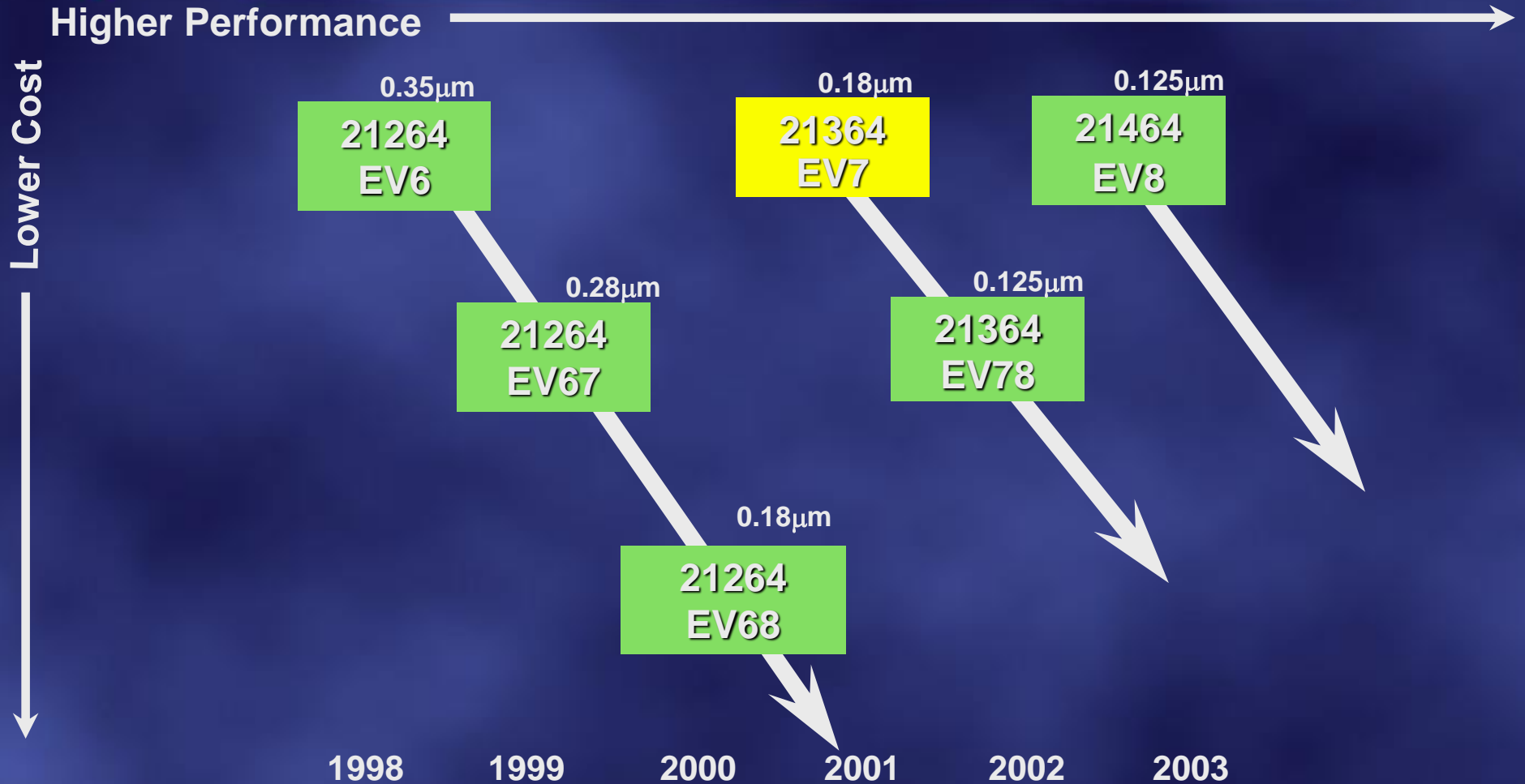
- First “Out-of-Order” Alpha
- Four-wide superscalar
- ...

## ◆ Performance

- World’s Fastest Microprocessor ([www.spec.org](http://www.spec.org), 11/17/99)
- 39 SPECINT95, 68 SPECFP95 @ 700 Mhz
  - Intel Pentium III @ 733 Mhz delivers 36 SPECINT95, 30 SPECFP95



# Alpha Microprocessor Roadmap





# Alpha 21364 Goals

---

- ◆ Leadership single stream performance
  - Higher operating frequency
  - Integrated memory interface
- ◆ Leadership multiprocessor performance
  - Integrated system / multiprocessor interface



# Alpha 21364 Features

---

## ◆ System-on-a-Chip

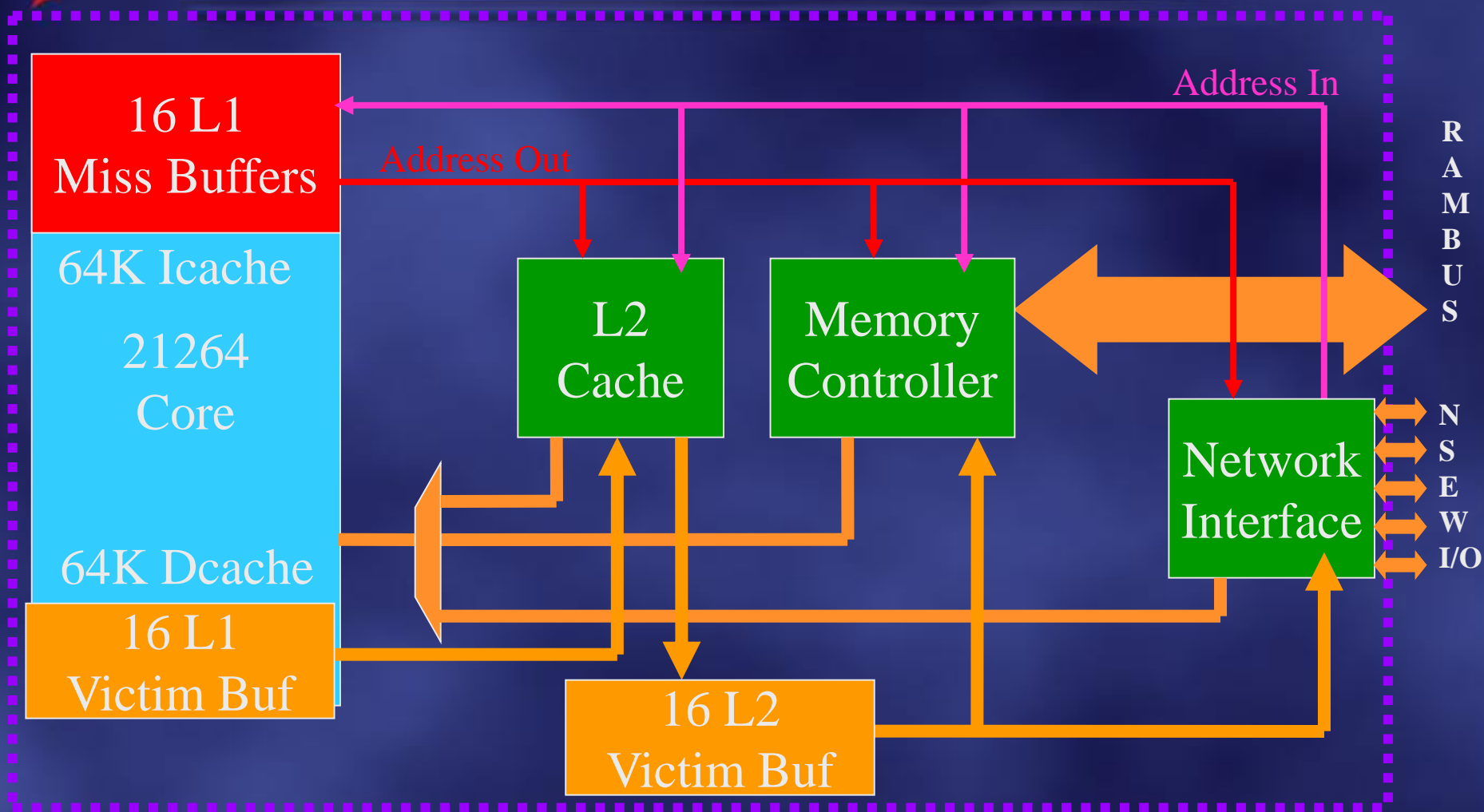
- Alpha 21264 core with enhancements
- Integrated L2 Cache
- Integrated memory controller
- Integrated network interface

## ◆ Fault-Tolerance

- Support for lock-step operation to enable high-availability systems.

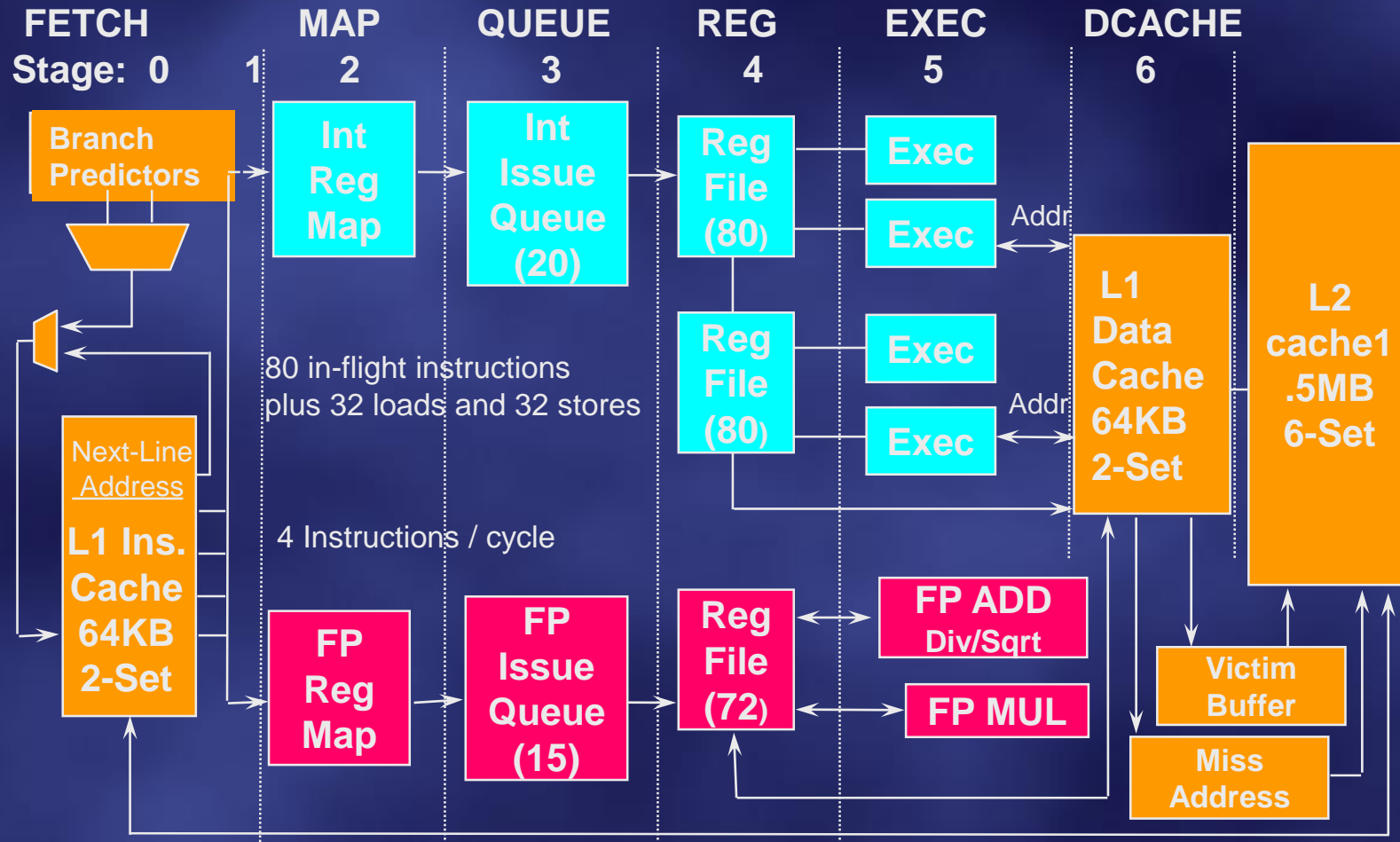


# 21364 Chip Block Diagram





# 21364 Core







# Integrated L2 Cache

---

- ◆ 1.5 MB
- ◆ 6-way set associative
- ◆ 16 GB/s total read/write bandwidth
- ◆ 16 Victim buffers for L1 -> L2
- ◆ 16 Victim buffers for L2 -> Memory
- ◆ ECC SECDED code
- ◆ 12ns load to use latency



# Integrated Memory Controller

---

- ◆ Direct RAMbus
  - High data capacity per pin
  - 800 MHz operation
  - 30ns CAS latency pin to pin
- ◆ 6 GB/sec read or write bandwidth
- ◆ 100s of open pages
- ◆ Directory based cache coherence
- ◆ ECC SECDED



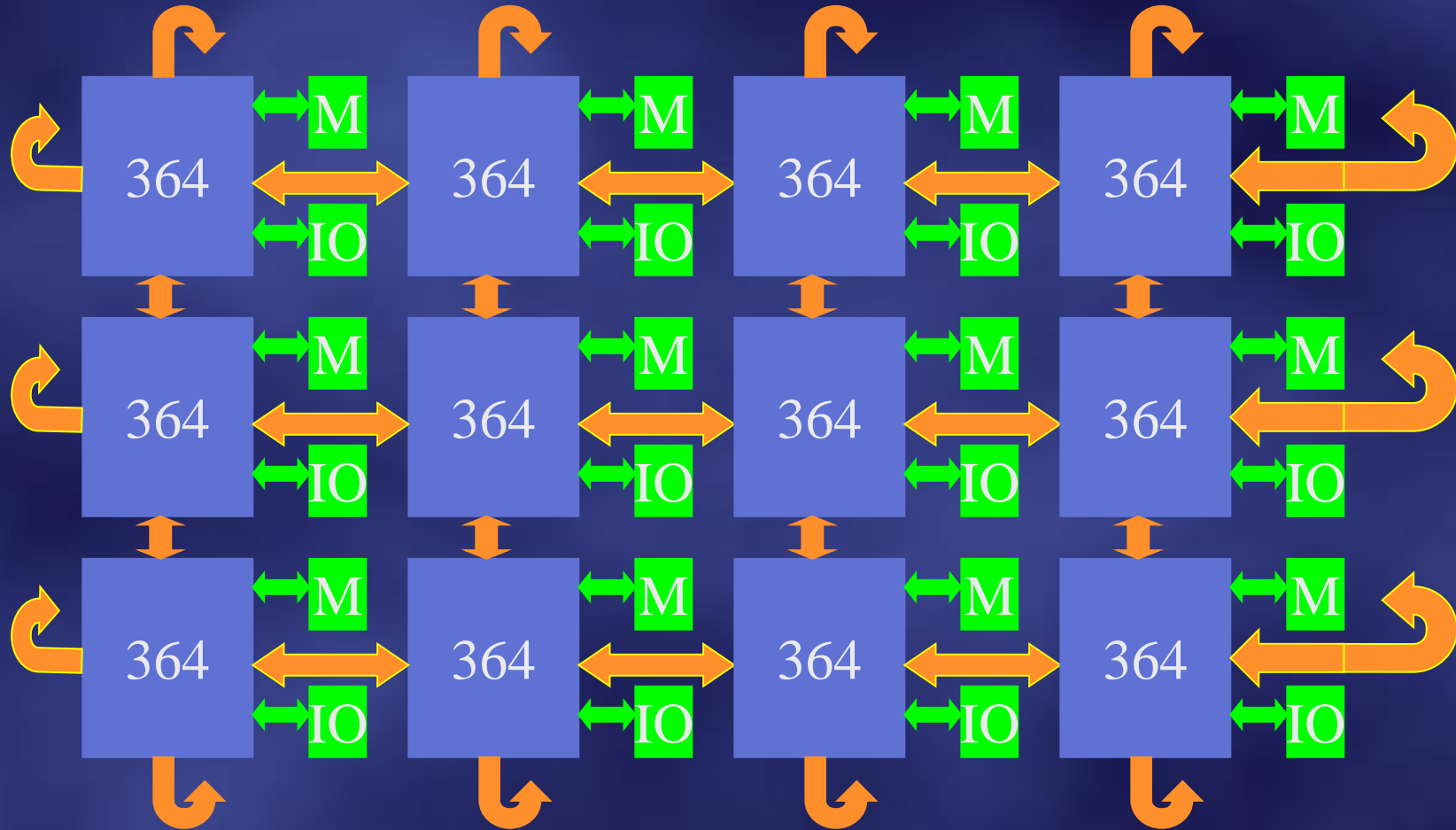
# Integrated Network Interface

---

- ◆ Direct processor-to-processor interconnect
- ◆ 10 GB/second per processor
- ◆ 15ns processor-to-processor latency
- ◆ Out-of-order network with adaptive routing
- ◆ Asynchronous clocking between processors
- ◆ 3 GB/second I/O interface per processor



# 21364 System Block Diagram





# Alpha 21364 Technology

---

- ◆ 0.18  $\mu\text{m}$  CMOS
- ◆ 1000+ MHz
- ◆ 100 Watts @ 1.5 volts
- ◆ 3.5  $\text{cm}^2$
- ◆ 6 Layer Metal
- ◆ 100 million transistors
  - 8 million logic
  - 92 million RAM



# Alpha 21364 Status

---

- ◆ 70 SPECint95 (estimated)
- ◆ 120 SPECfp95 (estimated)
- ◆ RTL model running
- ◆ Tapeout: Summer 2000



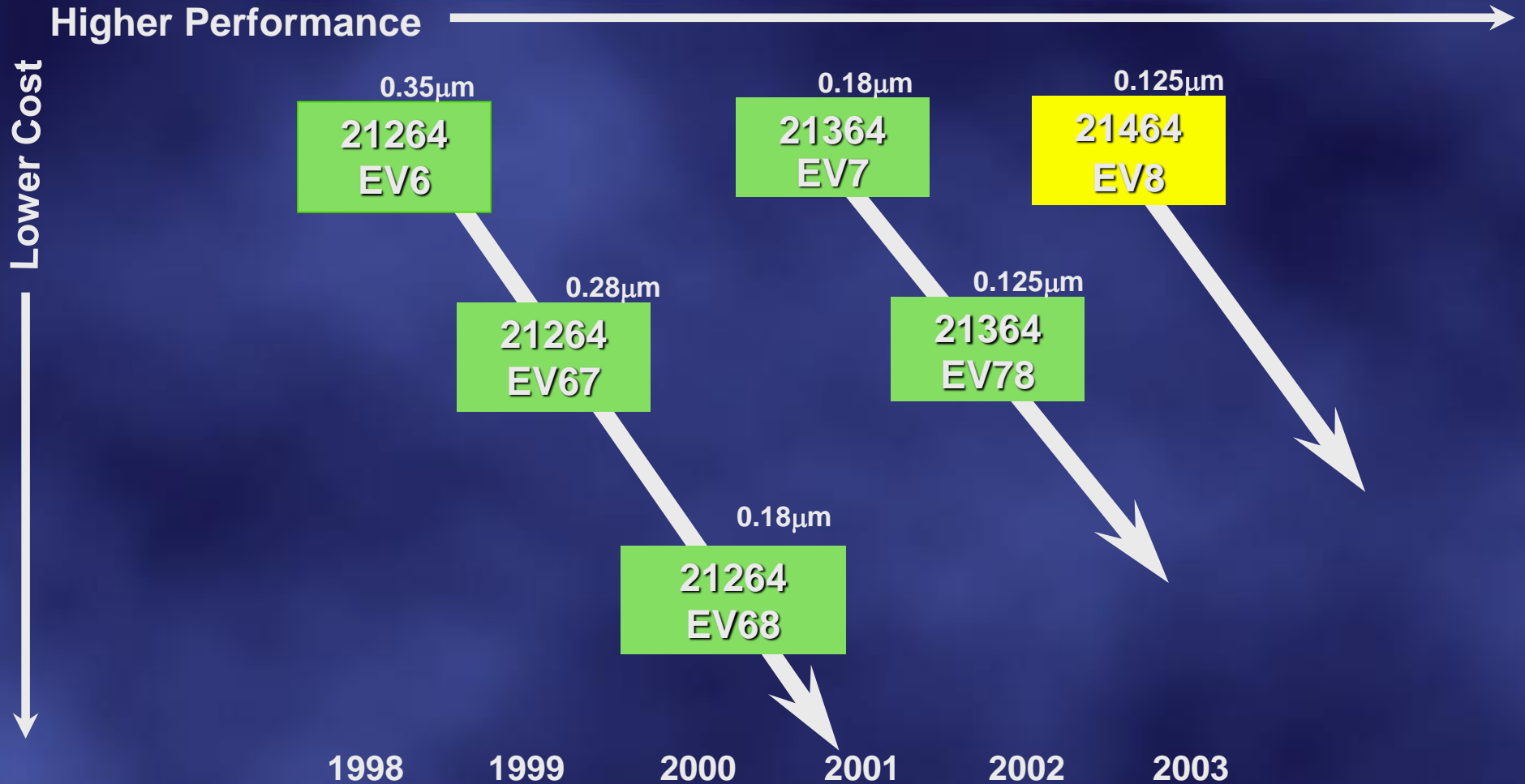
# 21364 Summary: System on a Chip

---

- ◆ Integrated L2 cache and memory controller
  - outstanding single processor performance
- ◆ Integrated network interface
  - high performance multi-processor systems
  - scales to large number of processors



# Alpha Microprocessor Overview







# Alpha 21464 Goals

---

- ◆ Leadership single stream performance
  - Higher operating frequency / better technology
  - New microarchitecture
  - Integrated memory interface (like 21364)
- ◆ Leadership multiprocessor performance
  - Simultaneous Multithreading (with minimal change/cost)
  - Integrated system / multiprocessor interface (like 21364)



# Alpha 21464 Technology Overview

- ◆ Leading edge process technology – 1.2-2.0GHz
  - 0.125 $\mu$ m CMOS
  - SOI-compatible
  - Cu interconnect
  - low-k dielectrics
- ◆ Chip characteristics
  - ~1.2V Vdd
  - ~250 Million transistors



# Alpha 21464 Architecture Overview

---

- ◆ Enhanced out-of-order execution
- ◆ 8-wide superscalar
- ◆ Large on-chip L2 cache
- ◆ Direct RAMBUS interface
- ◆ On-chip router for system interconnect
- ◆ Glueless, directory-based, ccNUMA
  - for up to 512-way multiprocessing
- ◆ 4-way simultaneous multithreading (SMT)



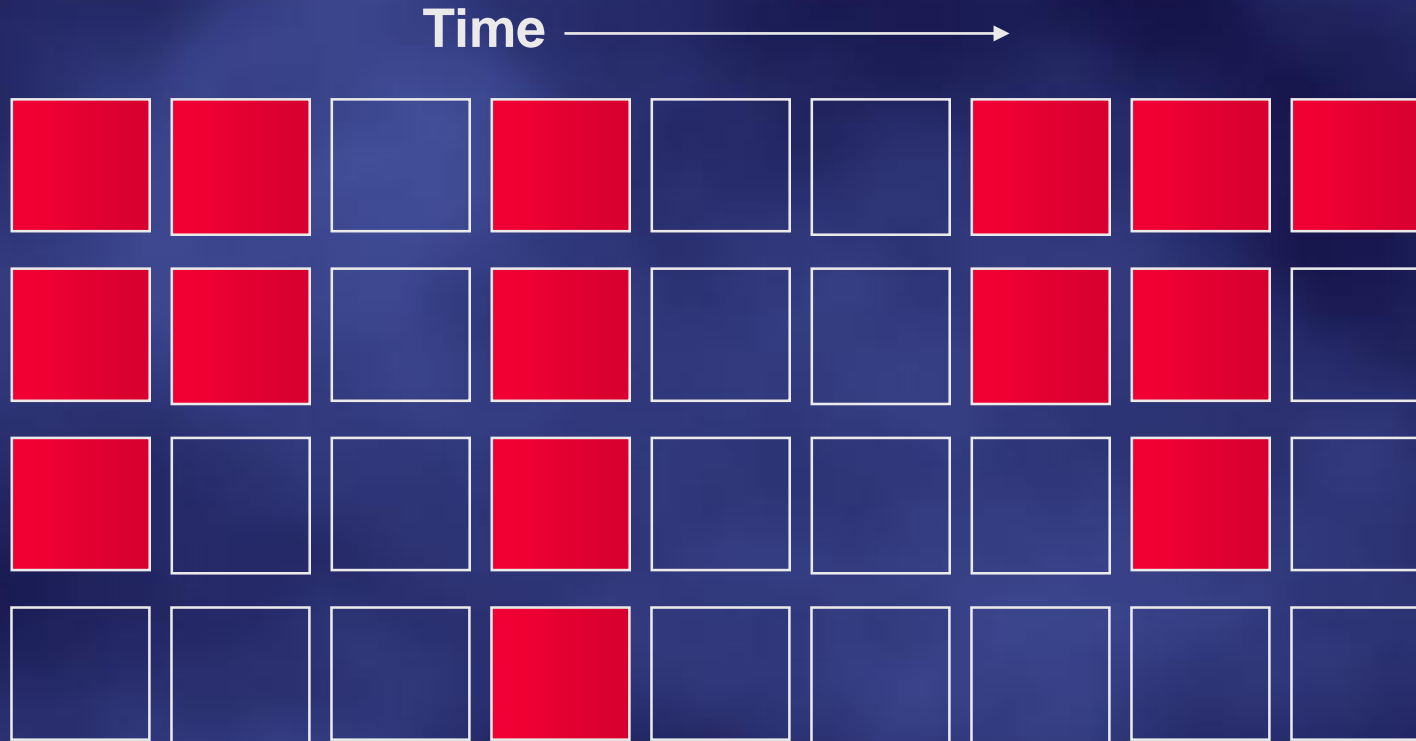
# Instruction Issue



Reduced function unit utilization due to dependencies



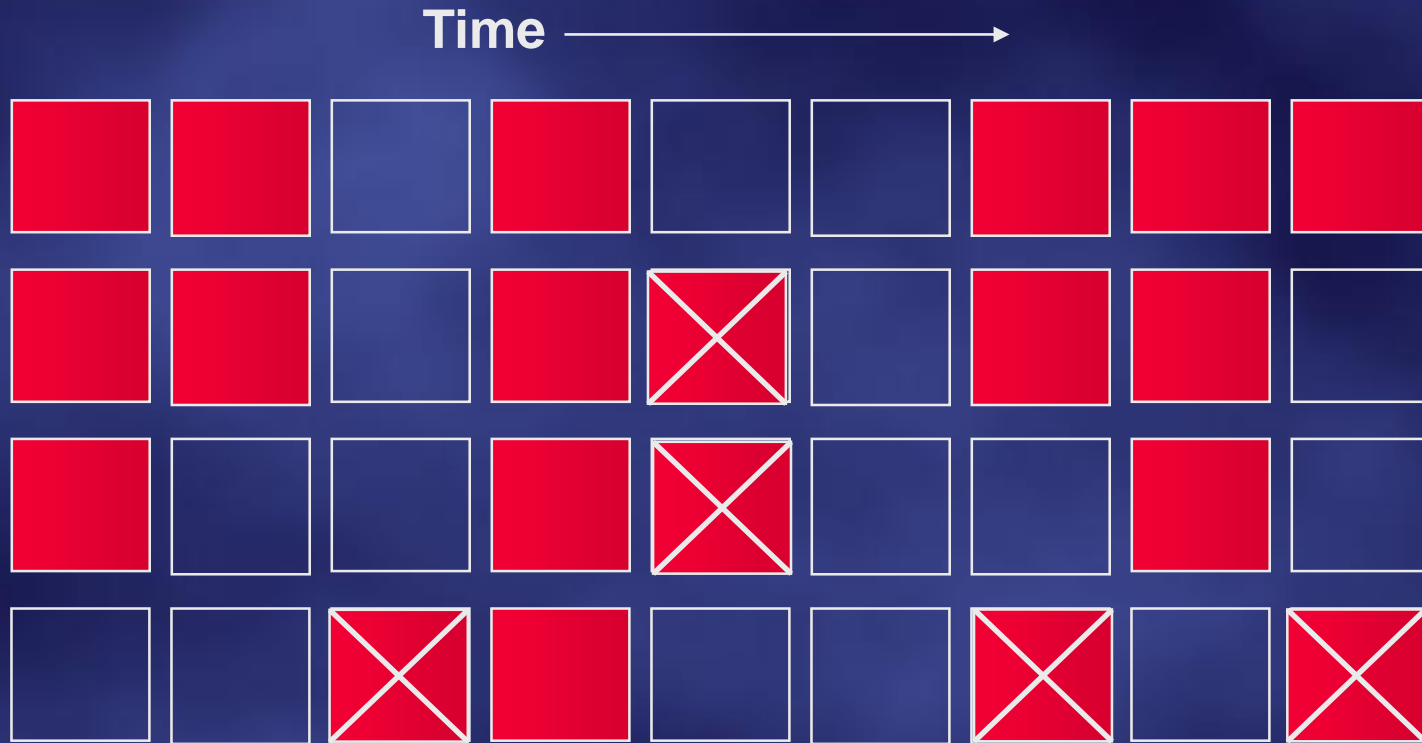
# Superscalar Issue



Superscalar leads to more performance, but lower utilization



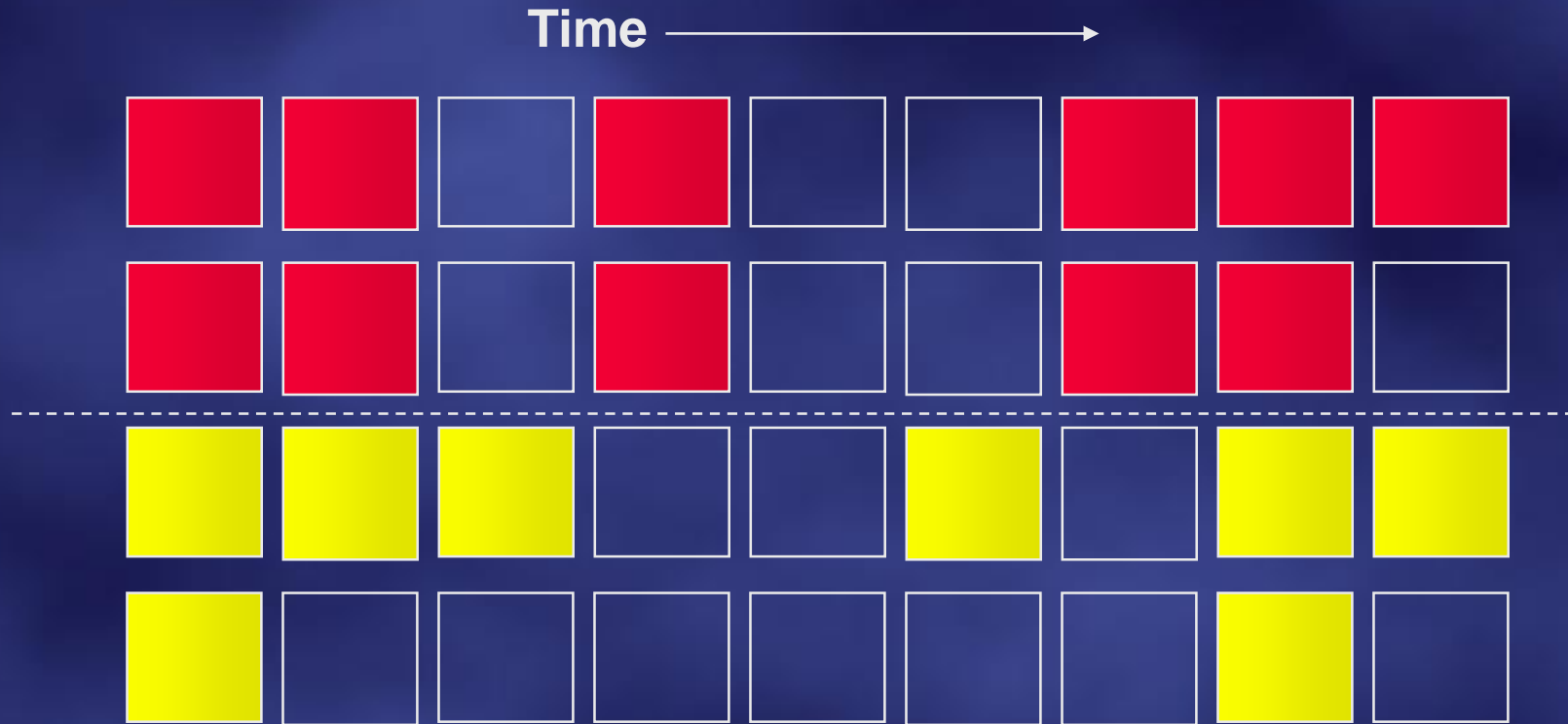
# Predicated Issue



Adds to function unit utilization, but results are thrown away



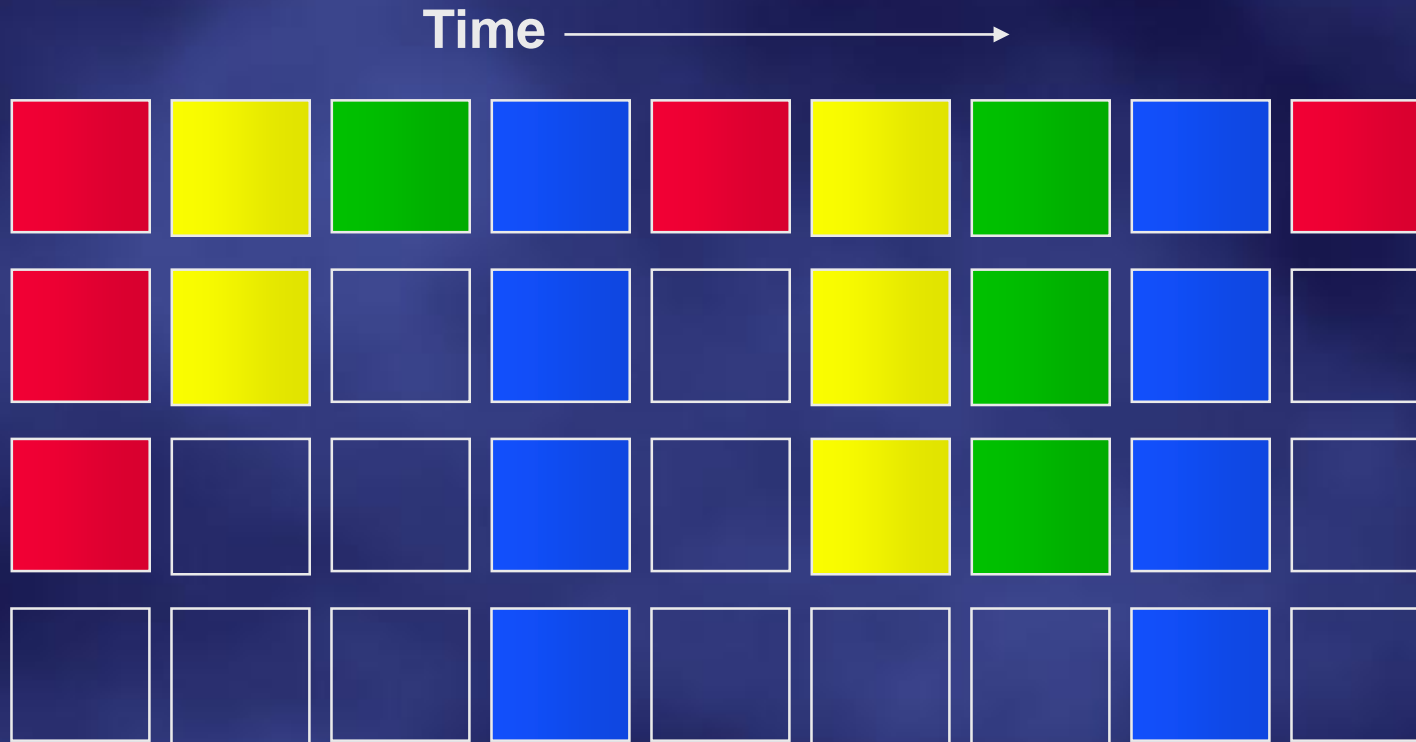
# Chip Multiprocessor



Limited utilization when only running one thread



# Fine Grained Multithreading

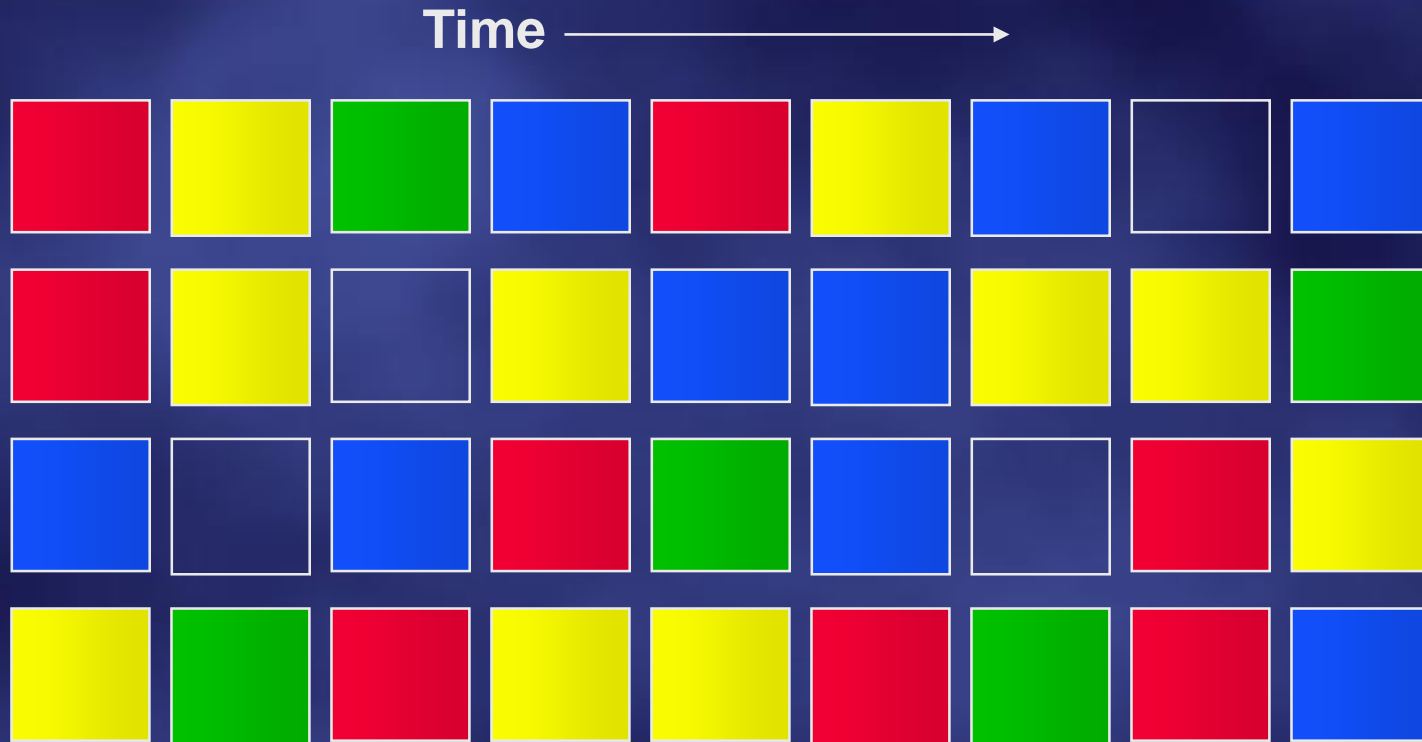


Intra-thread dependencies still limit performance





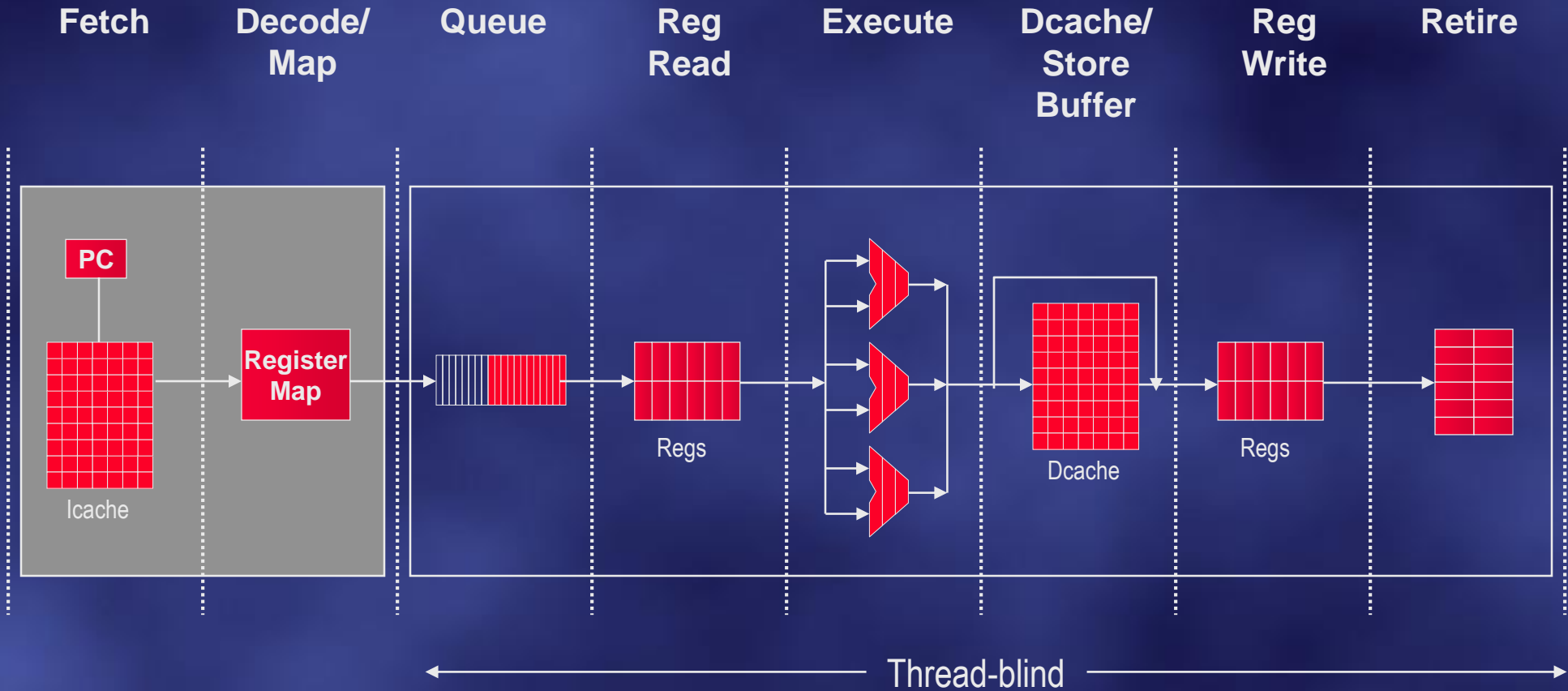
# Simultaneous Multithreading



Maximum utilization of function units by independent operations

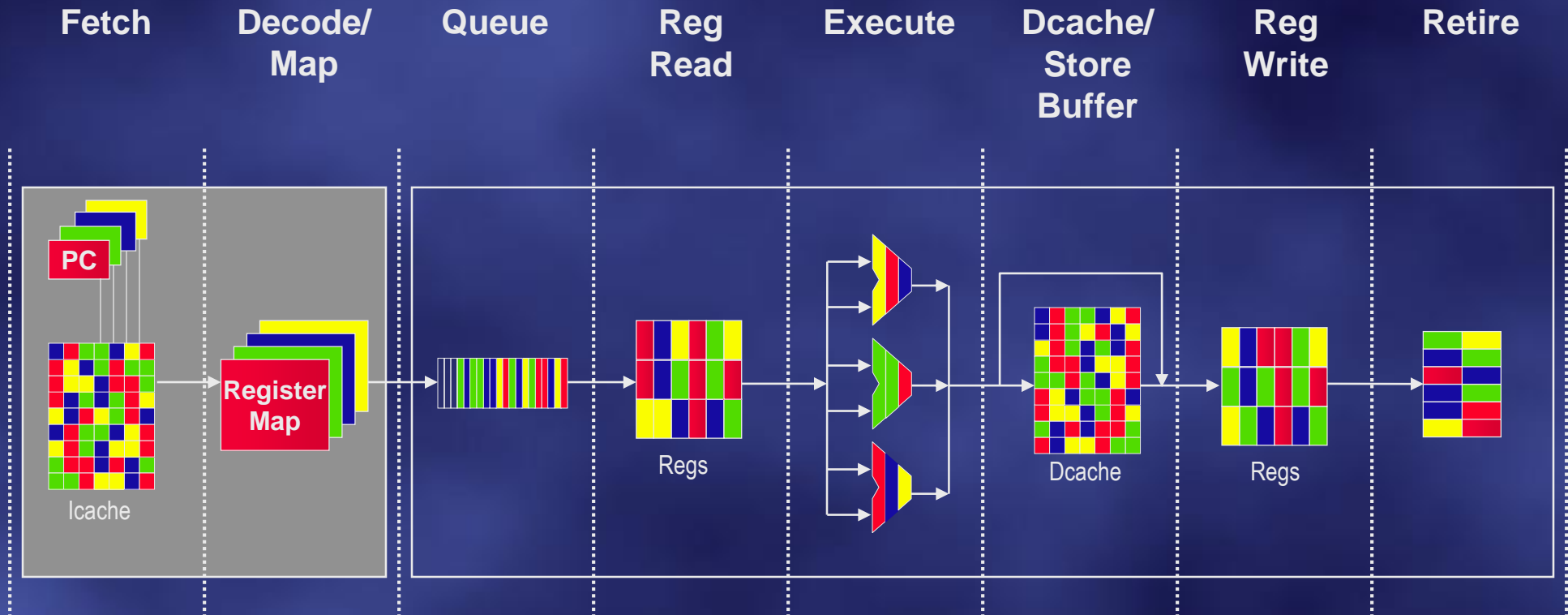


# Basic Out-of-order Pipeline





# SMT Pipeline





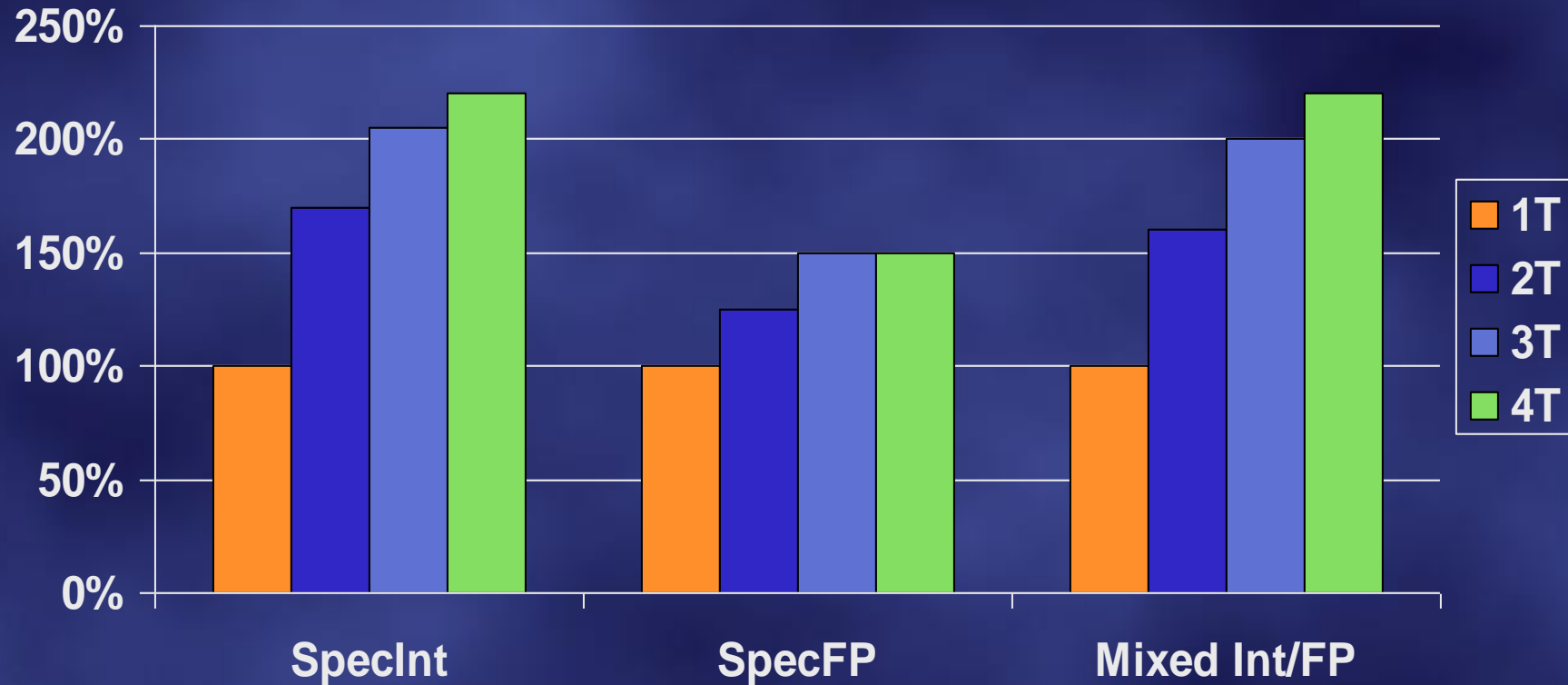
# Changes for SMT

---

- ◆ Basic pipeline – unchanged
  
- ◆ Replicated resources
  - Program counters
  - Register maps
  
- ◆ Shared resources
  - Register file (size increased)
  - Instruction queue
  - First and second level caches
  - Translation buffers
  - Branch predictor



# Multiprogrammed workload





# Decomposed SPEC95 Applications





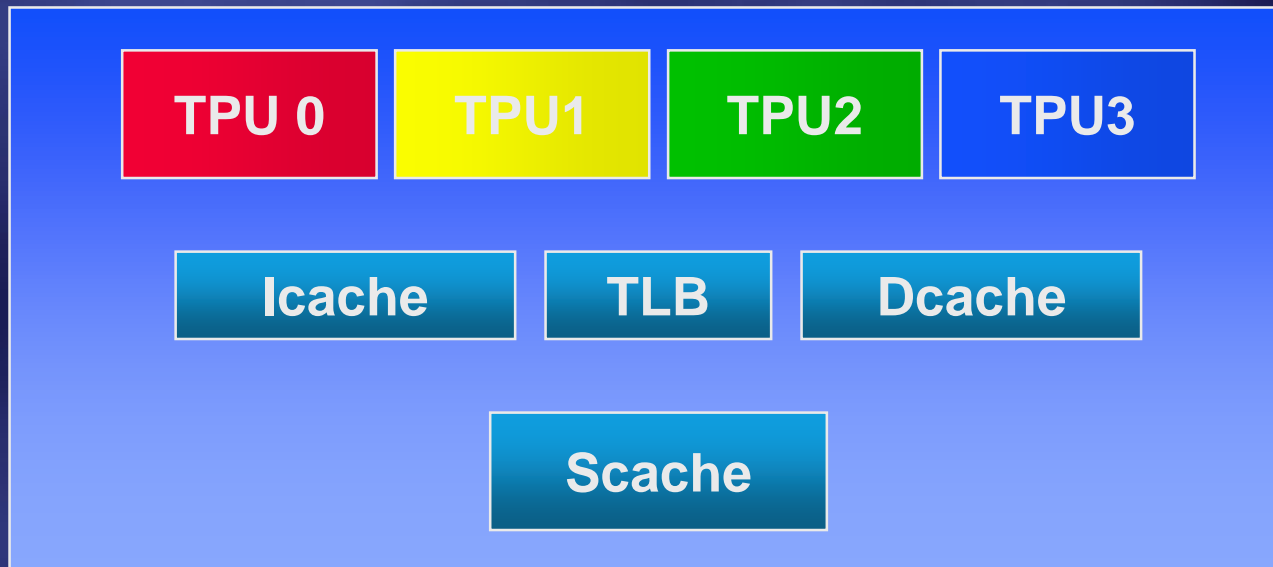
# Multithreaded Applications





# Architectural Abstraction

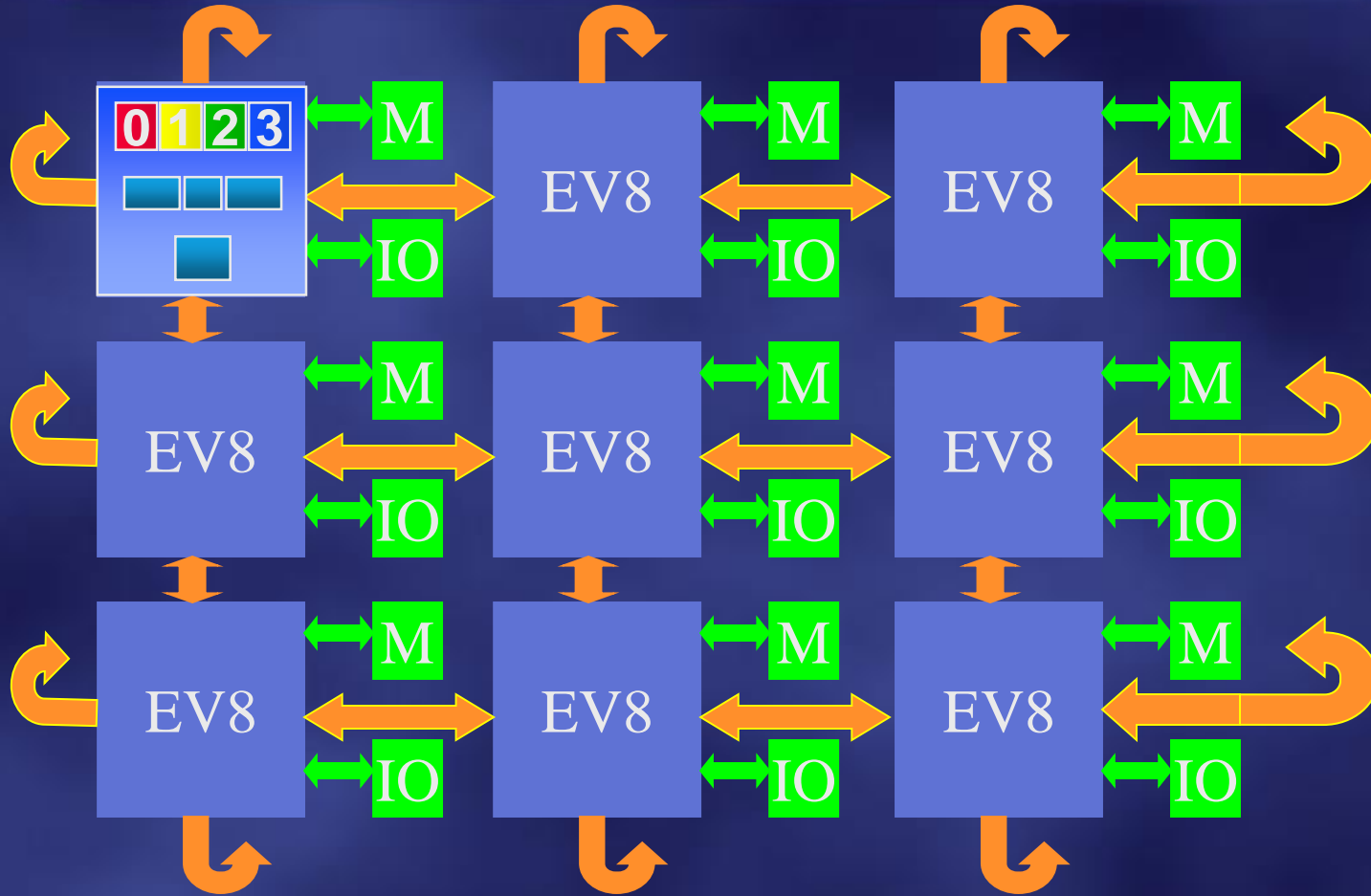
- ◆ 1 Processor with 4 Thread Processing Units (TPUs)
- ◆ Shared hardware resources







# 21464 System Block Diagram





# Alpha 21464 Summary

---

- ◆ Leadership single stream performance
  - Higher operating frequency / better technology
  - New microarchitecture
  - Integrated memory interface (like 21364)
- ◆ Leadership multiprocessor performance
  - Simultaneous Multithreading (with minimal changes/cost)
  - Integrated system / multiprocessor interface (like 21364)



# Maintain Performance Lead Beyond Y2K

---

## ◆ Alpha 21364

- Reuses 21264 microprocessor core
- System on a chip

## ◆ Alpha 21464

- New microarchitecture
- System on a chip
- Simultaneous Multithreading



# My Current Research: Beyond 21464?

---

- ◆ **The Truth Project** (w/ Joel Emer)
  - Examines different microarchitectural issues
- ◆ **The Multinet Project** (w/ Rick Kessler)
  - Tightly-coupled multiprocessor networks
- ◆ **The Reliant Project** (w/ Steve Reinhardt)
  - Self-Checking Microprocessors using SMT, ISCA submission
- ◆ **Asim** (w/ VSSAD Labs)
  - Performance Model for Alphas beyond 21464