# Broadcom BCM5600 StrataSwitch

## A Highly Integrated Ethernet Switch On A Chip

**Andrew Essen and James Mannos**

**Broadcom Corporation**

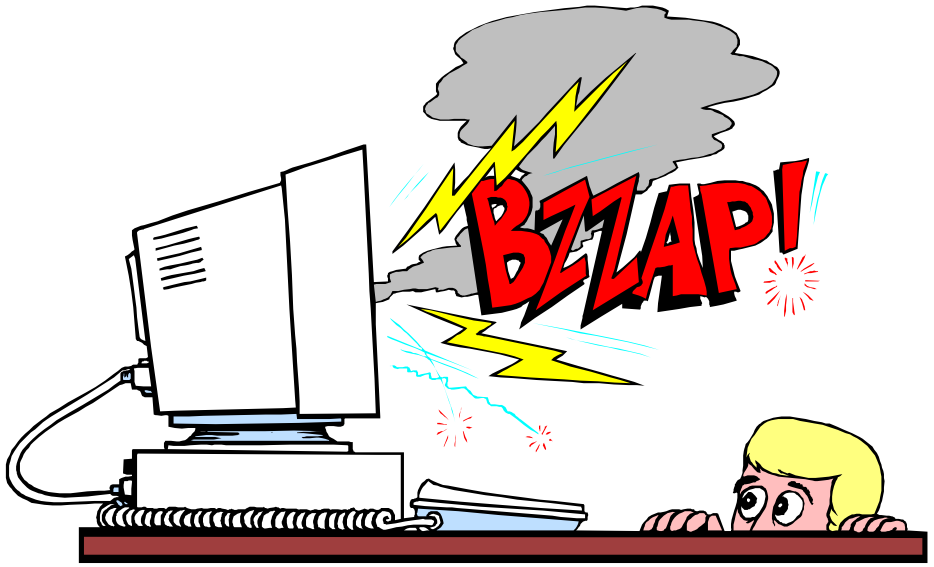
# Outline

- Introduction
- Networking Basics
- Description of BCM5600
- Design Process
- Vital Statistics
- Conclusion

# Background

- **A single modern processor can now saturate a LAN which used to be easily shared by multiple users.**

- **Multimedia applications, especially those with real-time requirements, place a huge strain on bandwidth**




BROADCOM®

# New Technology to the Rescue

- **New Network Architectures - replace shared media and hubs with dedicated media and switches**
- **Provision for Increased Bandwidth - replace 10Mb/s with 100Mb/s (FE) and 1000Mb/s (Gig)**
- **New Protocols - L4 to L7 filtering, Class of Service (802.1p), Virtual LANS (802.1Q)**
- **Advancing Semiconductor Technology - DSM (< .25μ), System-on-chip architectures, and large amounts of integrated memory.**

BROADCOM®

# Project Goals

| Features | Benefits |
|---|---|
| Switch On Chip | - Lowest Cost/Port<br>- Faster System Development<br>- Modules can be easily "bolted" on |
| 24 10/100 & 2 Gig Ports Non-blocking | - High Capacity<br>- High Performance |
| Non-blocking L2/L3 Switch | - Line speed switching and routing |
| Line speed L4-L7 filtering. | - Fast Filter Processor & Flexible Rules Engine<br>- Content Aware traffic classification based on any combination of designated fields. |
| Advanced Features | - 4-levels of COS (Class of Service).<br>- Support of Virtual LANs, Trunking, Flow Control, Mirroring, Tagging & Spanning Tree<br>- Elimination of Head of Line Blocking<br>- Broadcast/Multicast |
| On board SRAM caches | - Allows use of low-cost, external SDRAM |
| PCI interface | - Flexible connectivity and expandability. |

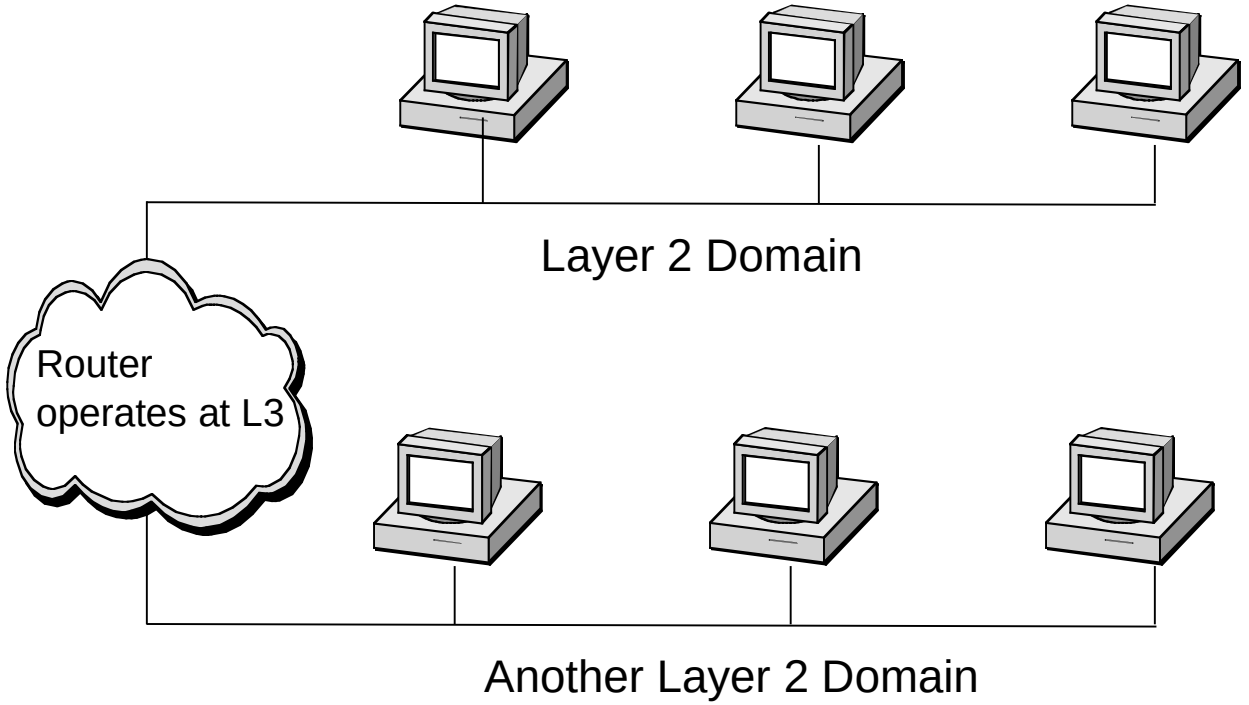
# What Makes a SOC so Tough?

- Network traffic is chaotic and asynchronous
- Can have large differential flows of data
- Behavior often appears non-deterministic because of packet drop
- Differing packet sizes & formats along with a huge amount of state - very difficult to verify
- 60M transistors, 1 MB SRAM, 4M+ gates.

BROADCOM®

# Traditional Network

**(Before Switched LAN's)**




Each L2 broadcast domain is connected by a router

# Layer 2 Switching

- **Appears to end stations as though everybody is still on a single shared media**
- **Lookup destination MAC address and forward packet to correct port**
- **Numerous packets can be “on the wire” simultaneously**
  - Up to 52 with the BCM5600
    - 26 full duplex ports
  - Versus 1 with a traditional shared Ethernet

# Layer 3 Switching

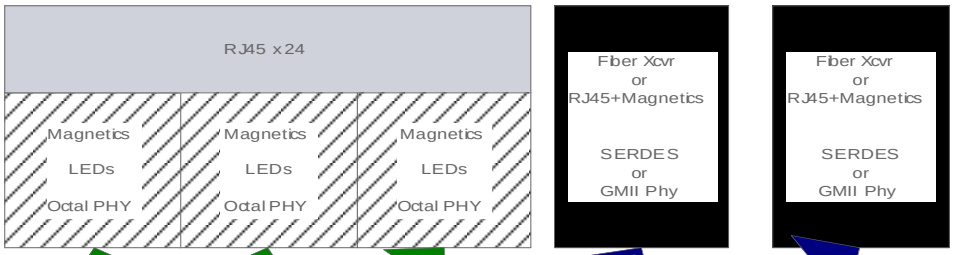
- **End stations must be kept aware of routing configuration**
- **L3 Switching is performed when the L2 destination is the router itself**
  - Lookup IP destination in routing table
  - Update L2 addresses
  - Decrement TTL
  - Forward packet to correct port

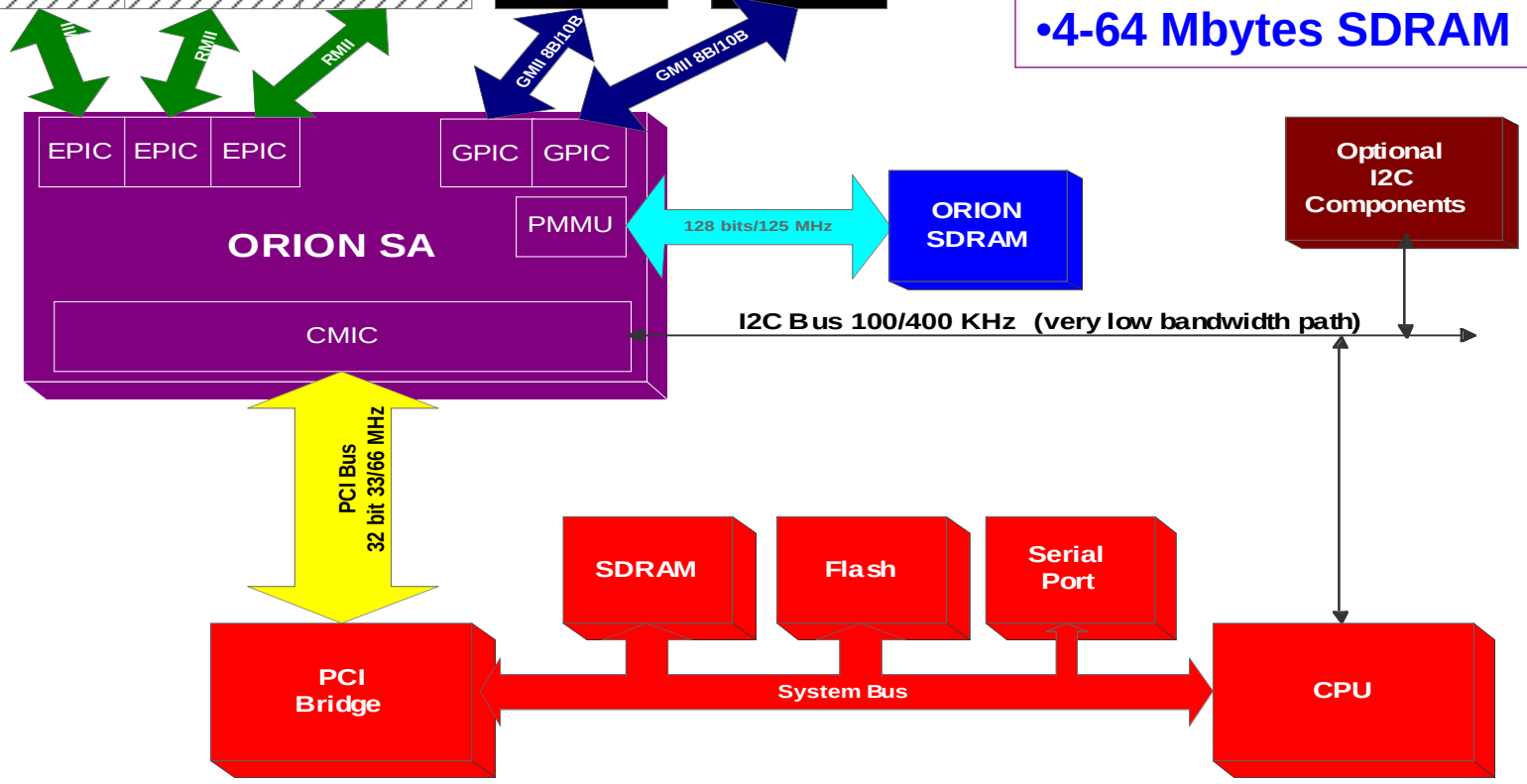
# Layer 4-7 Processing

- **Pattern matches result in actions**
  - Kind of like a simple awk script
  - Actions include drop, change priority, change destination, send to CPU
- **Applications include:**
  - Detect traffic type based on TCP port and reprioritize
  - Support VoIP and multimedia applications such as real-time video
  - Simple firewall by examining IP addresses

BROADCOM®

# StrataSwitch 24+2 Switch



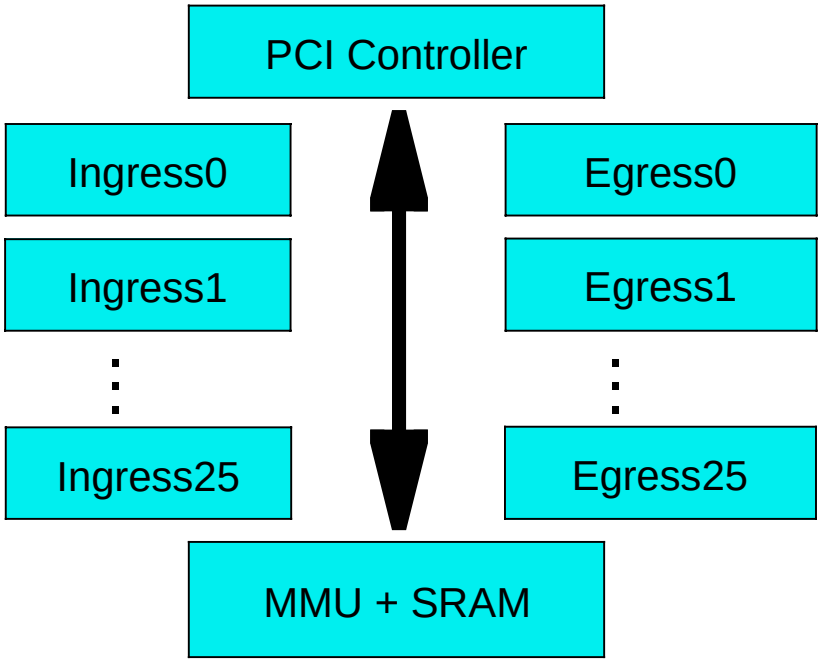




***Additional Silicon Required***

- 3-5218
- 2-5400
- Integrated Microcontroller
- 4-64 Mbytes SDRAM

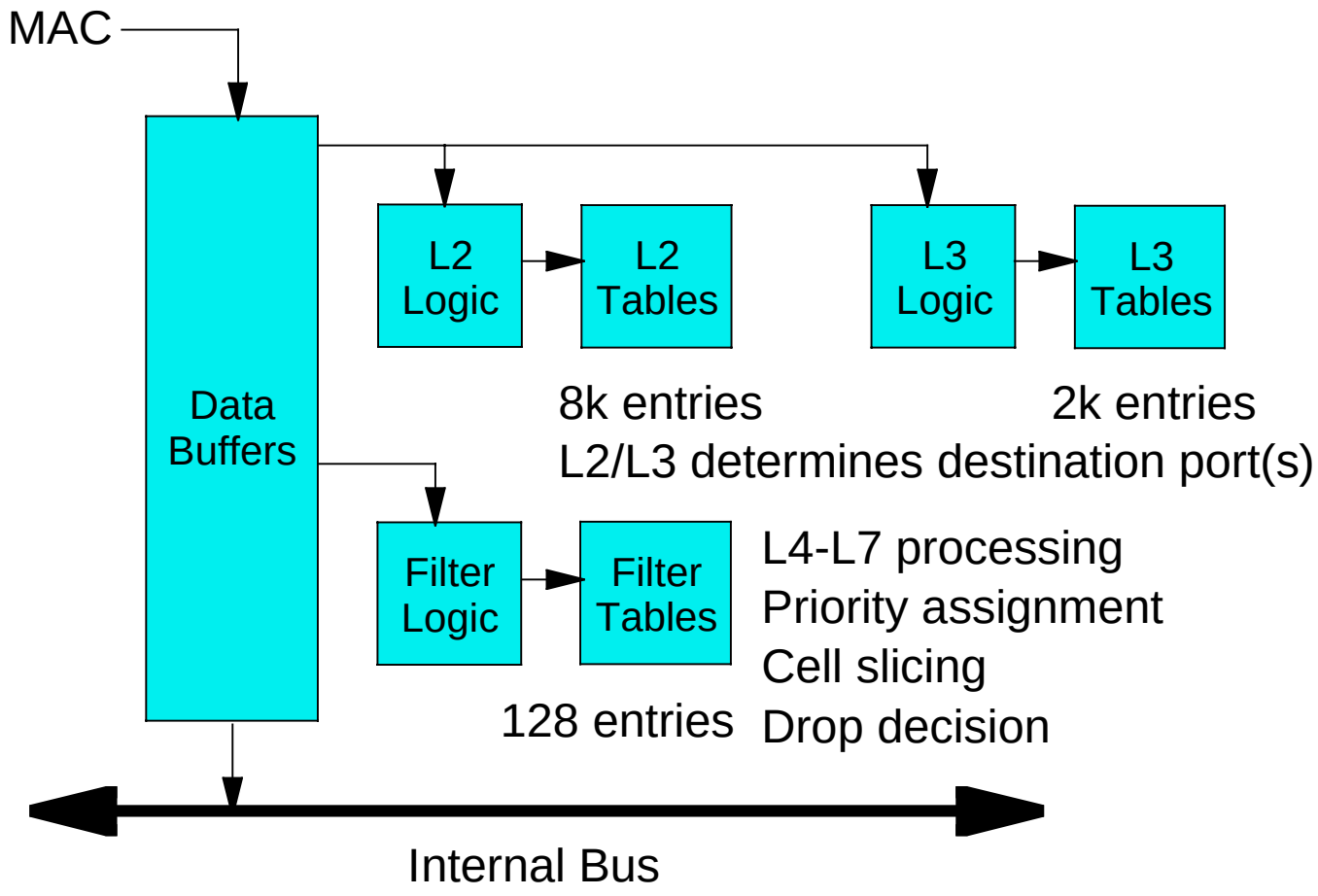
BROADCOM®

# Simplified Block Diagram

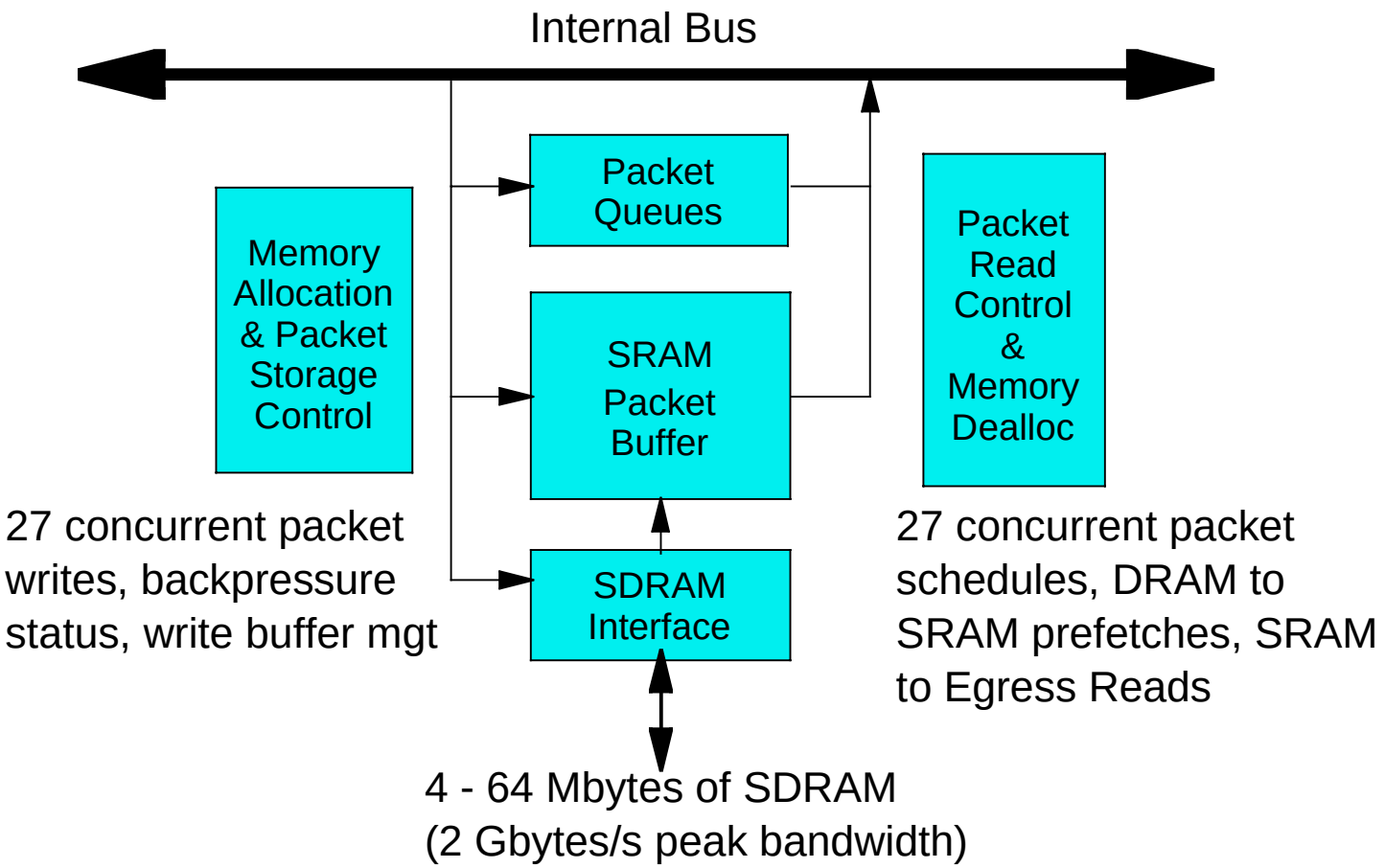



All packets travel as cells along a very wide central bus

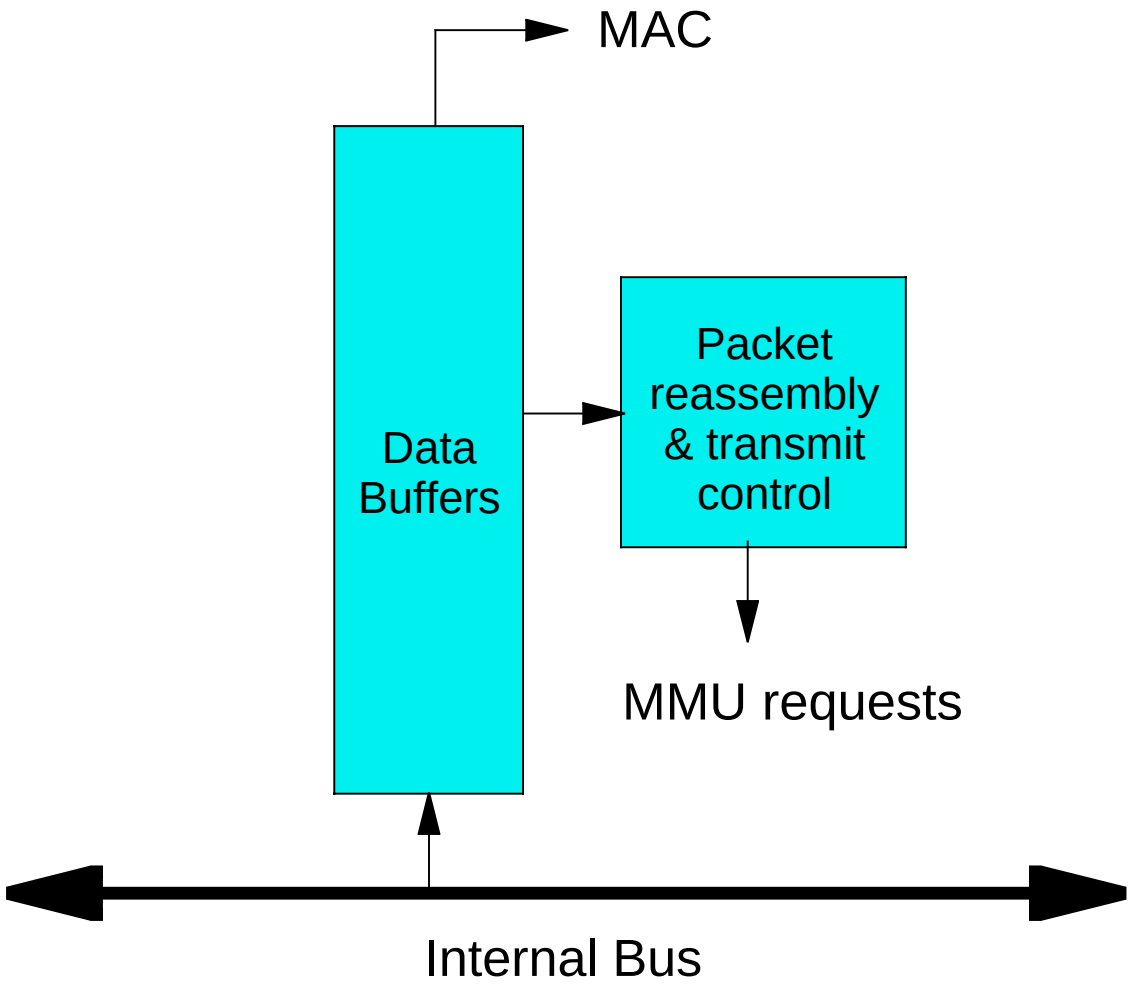
# Ingress Details



# MMU Details


Internal Bus
Packet Queues
Memory Allocation & Packet Storage Control
SRAM Packet Buffer
Packet Read Control & Memory Dealloc
27 concurrent packet writes, backpressure status, write buffer mgt
SDRAM Interface
27 concurrent packet schedules, DRAM to SRAM prefetches, SRAM to Egress Reads
4 - 64 Mbytes of SDRAM
(2 Gbytes/s peak bandwidth)

# Egress Details



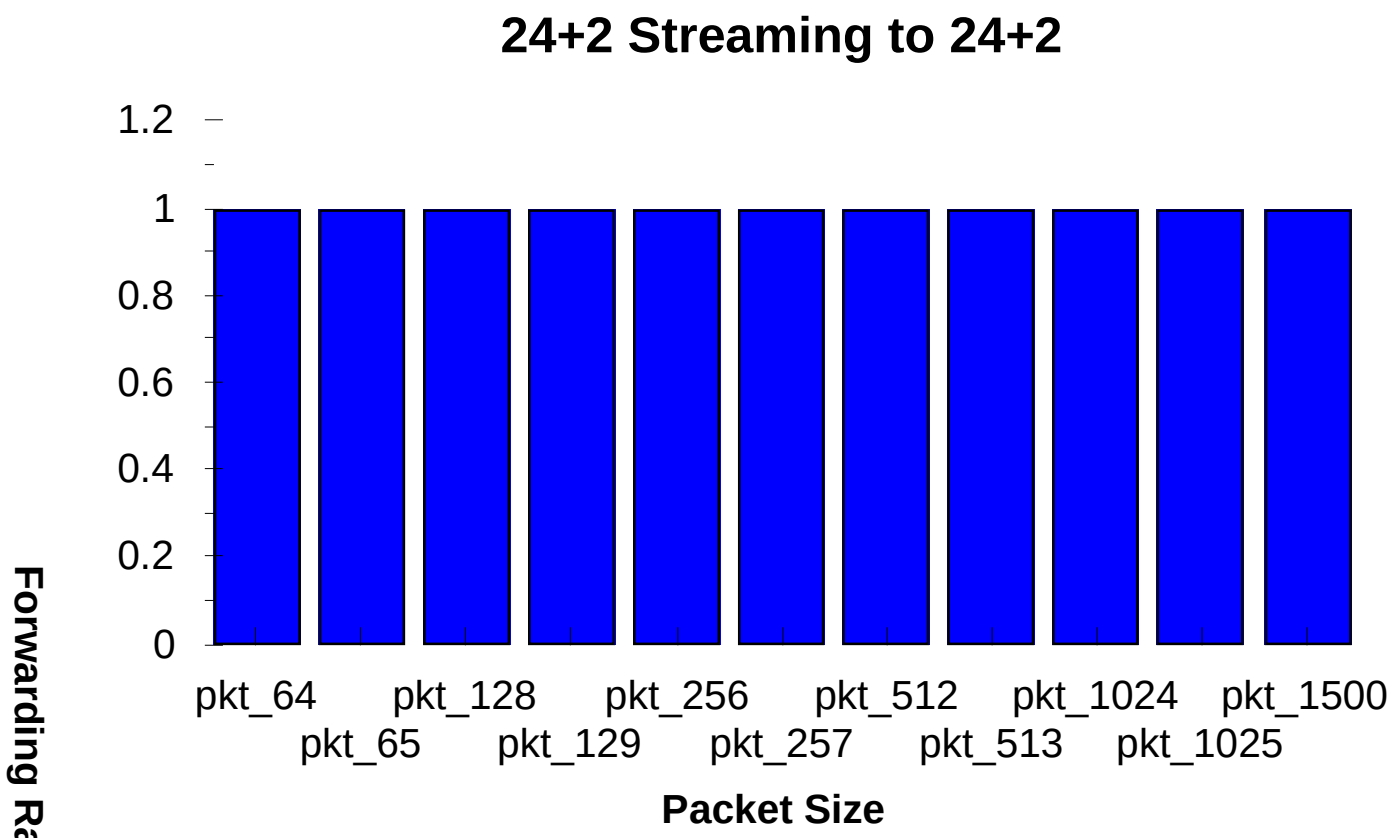
# PCI Details

- **32 bit/ 66 MHz operation**
- **Also behaves as port 27**
  - Contains mini- ingress & egress
  - Allows easy processing of unusual packets
    - E.g., routing non- IP packets

# Simulation

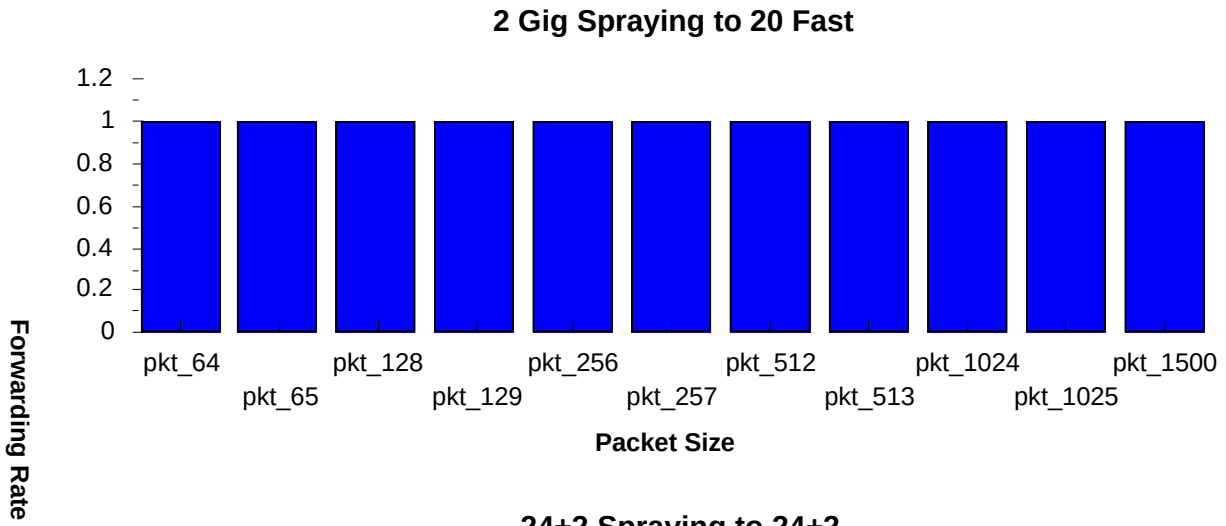
- **“Cycle accurate” C++ model**
- **Validate architecture**
  - First implementation of initial concepts
  - Try out various algorithms
  - Find bottlenecks
- **Optimize queue and buffer sizes**
  - With regards to finite real estate

# Simulation Results

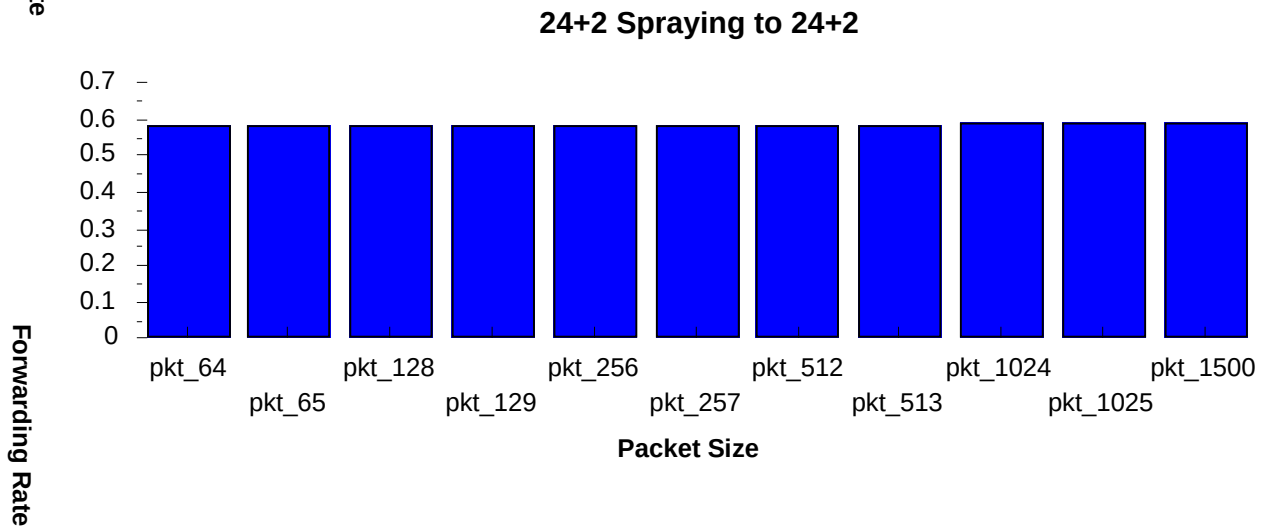



Wire speed operation with various packet sizes

# More Results

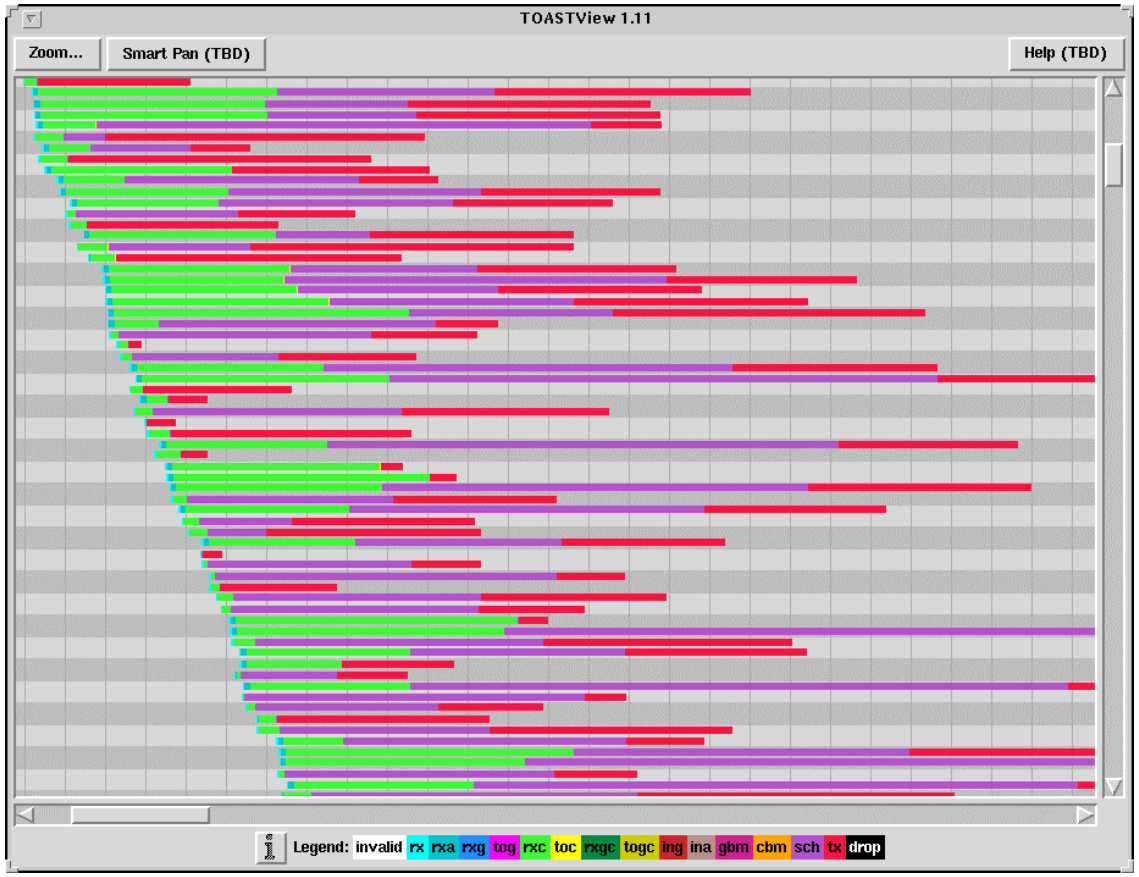



**Wire speed**




**Fast ports are overloaded**

# Simulated Packet Flow

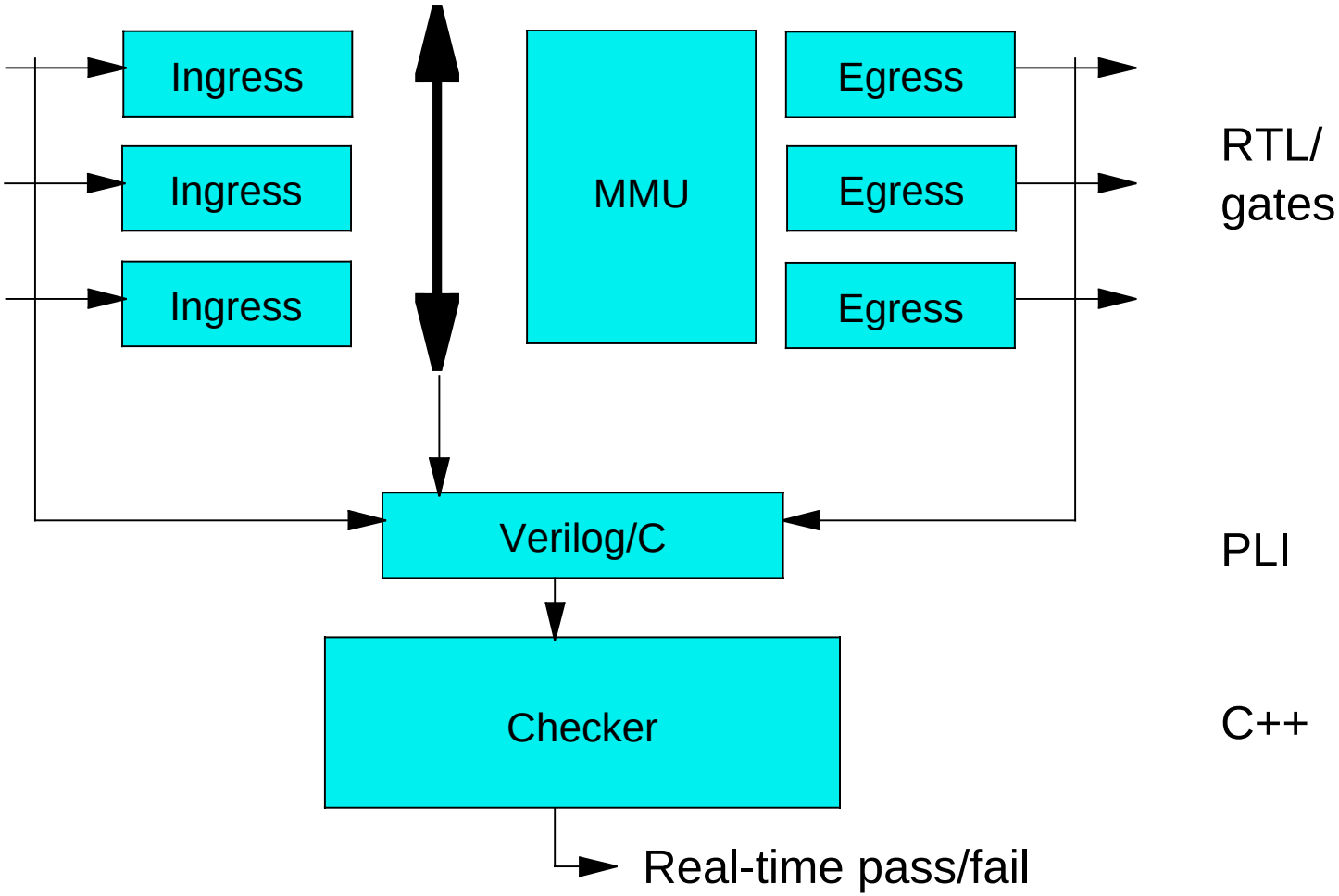



Timeline of individual packets in simulation

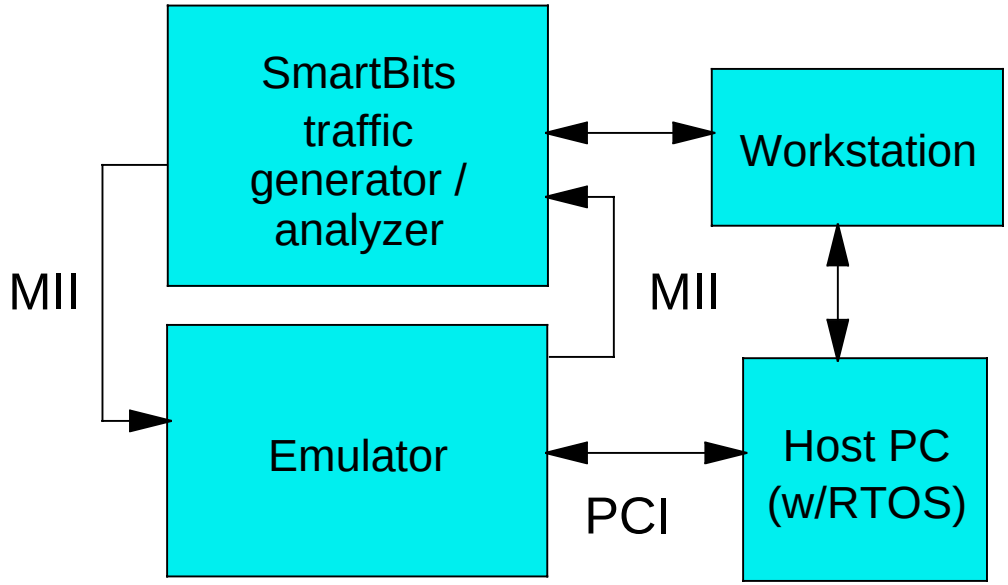
# Verification

- **Verification is complicated by…**
  - Multithreaded, asynchronous nature
  - Long operations
    - 160k cycles for 1500B packet at 10 mbits
  - Long time to steady state
    - Filling 64 Mbytes takes a while at 100 Hz
  - Dropped or redirected packets may be correct
- **Multiple solutions**
  - Extensive unit testing
  - Full chip testing with automated checking
  - Emulation

# Checker


Ingress
Ingress
Ingress
MMU
Egress
Egress
Egress
RTL/
gates
Verilog/C
PLI
Checker
C++
Real-time pass/fail

# Emulation




- **50 KHz is a lot better than 50 Hz!**
  - Allows software development
  - Millions of packets per day
  - Test environment is similar to that for silicon debug

# Emulation Setup

SmartBits unit, Host PC, and speedbridges

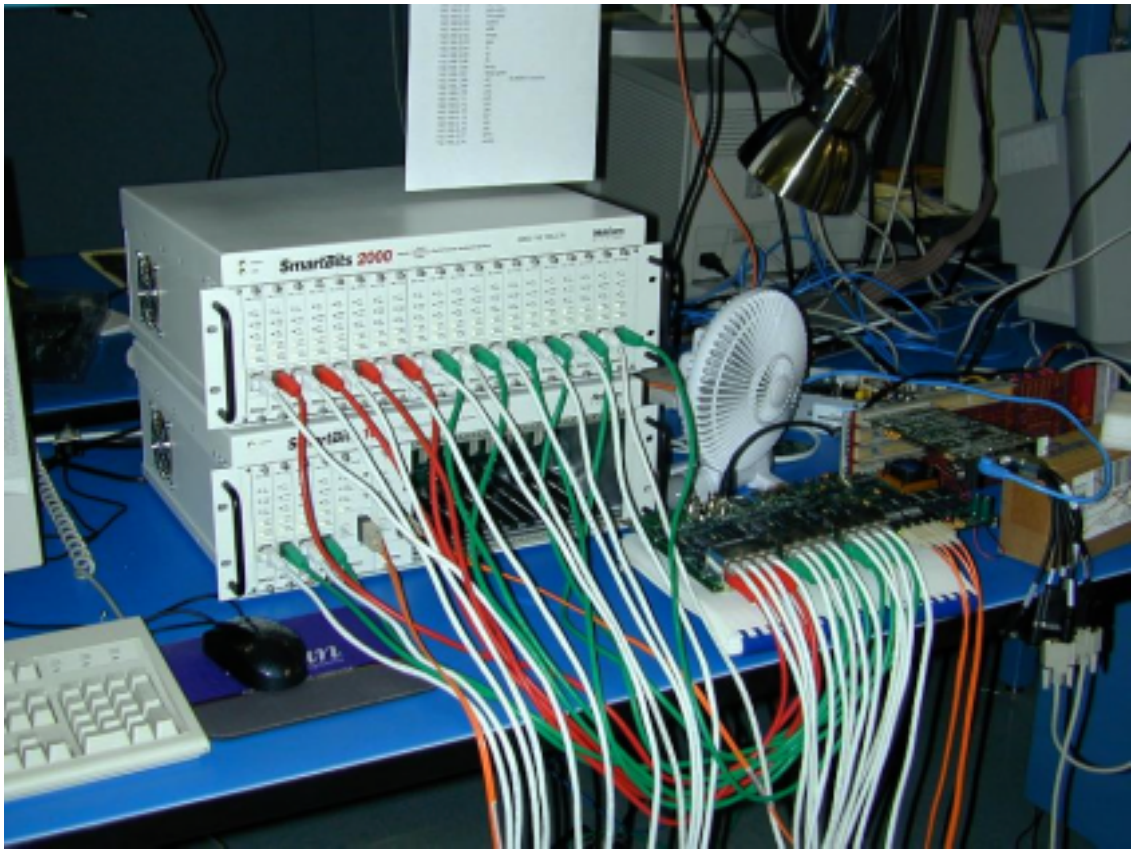
BROADCOM®

# The Emulator

One of three cabinets

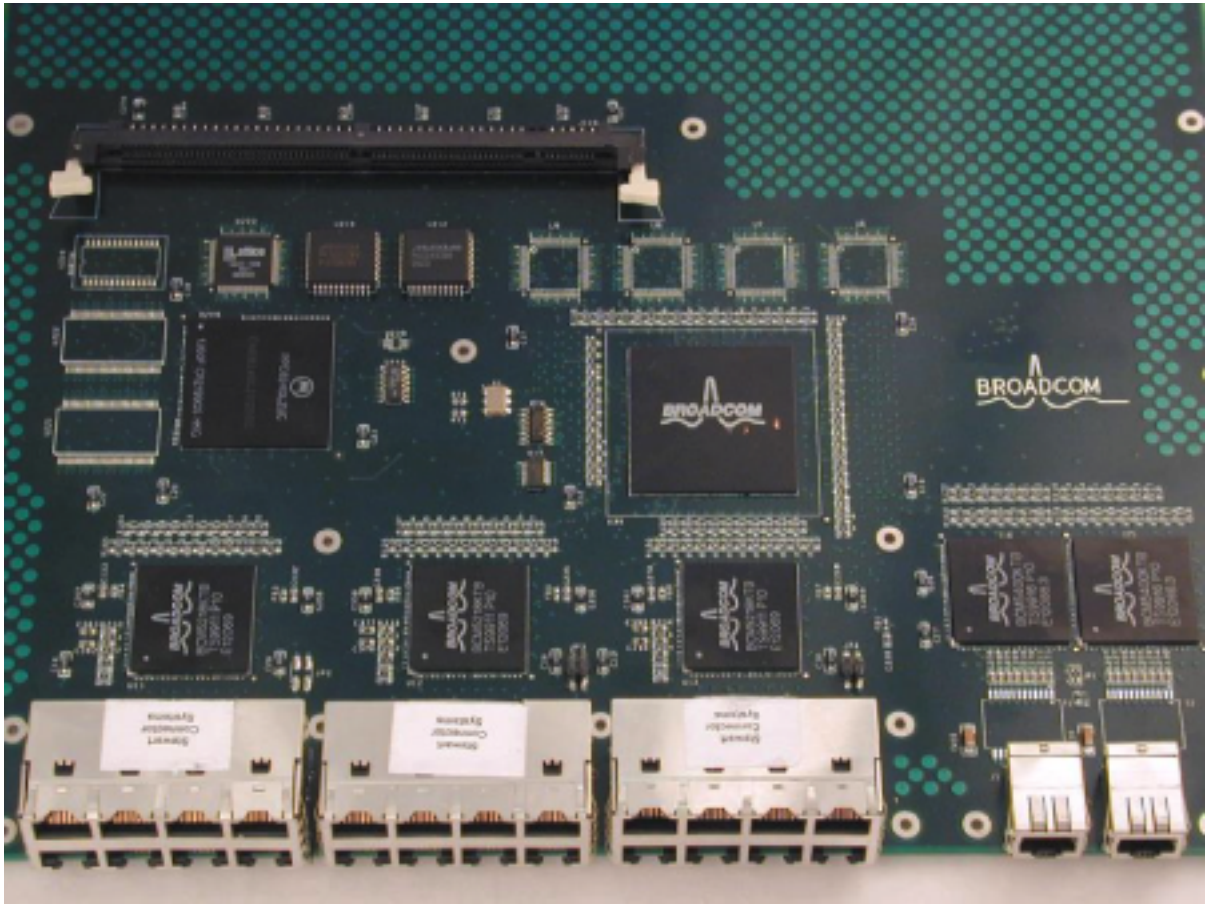
BROADCOM®

# System Bringup

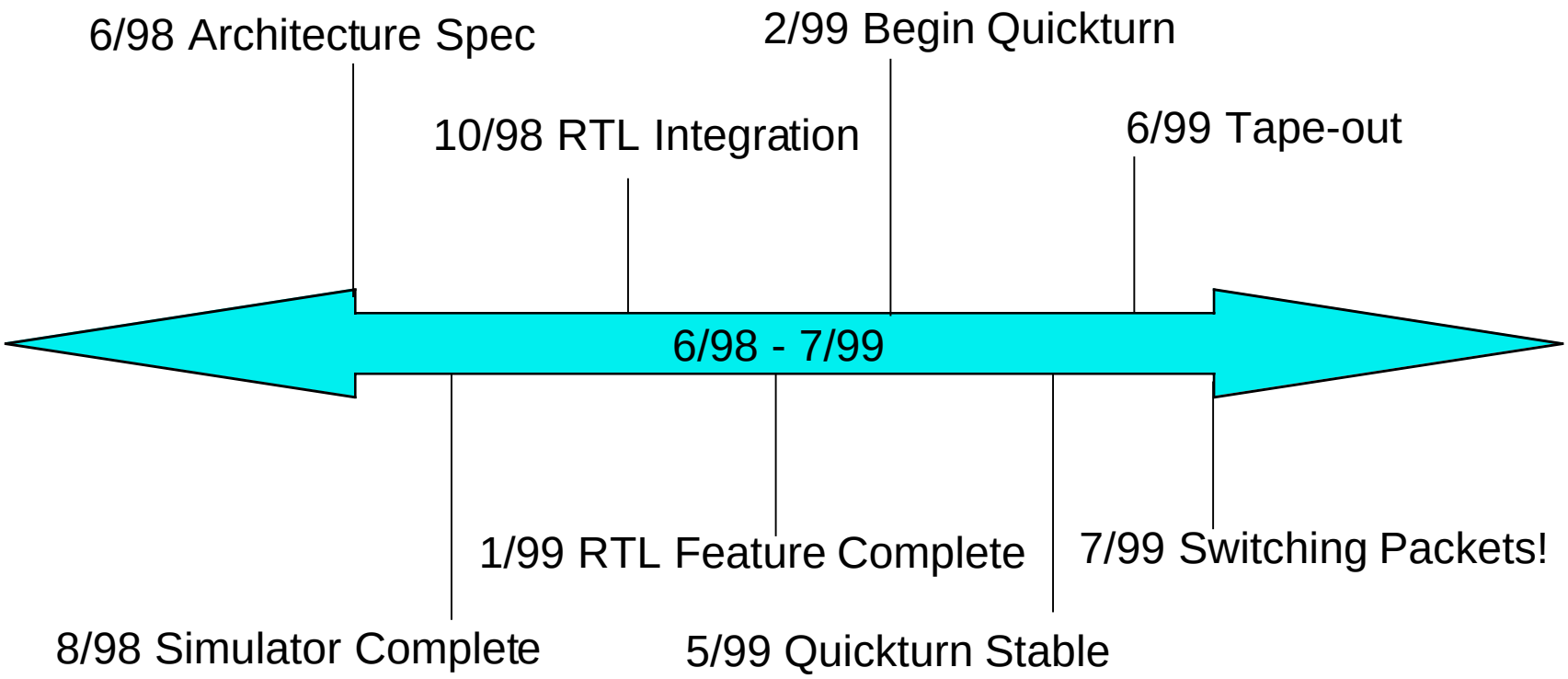



SmartBits and board

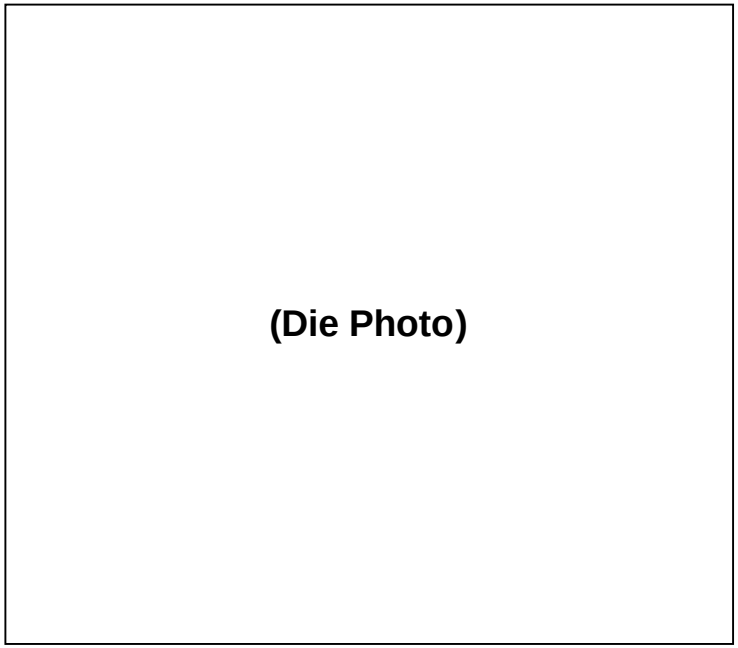
# Reference Board



# Design Timeline



# Vital Statistics




- 24+2 ports
- 60 million transistors
- 1MB of embedded SRAM
- 133 MHz
- 0.25u 5 metal CMOS
- 2.5V core, 3.3V I/O
- 600 ball TBGA package

BROADCOM®

# Summary

- First integration of 24FE + 2GE, line-speed, L2-L7 switch on a ***single*** chip
- Enables convergence of voice, video, and data to the desktop.
- First pass silicon in less than 12 months, read PCI ID in minutes, switching packets the next day!
- Good flow from solid spec through emulation was crucial to the success.

BROADCOM®