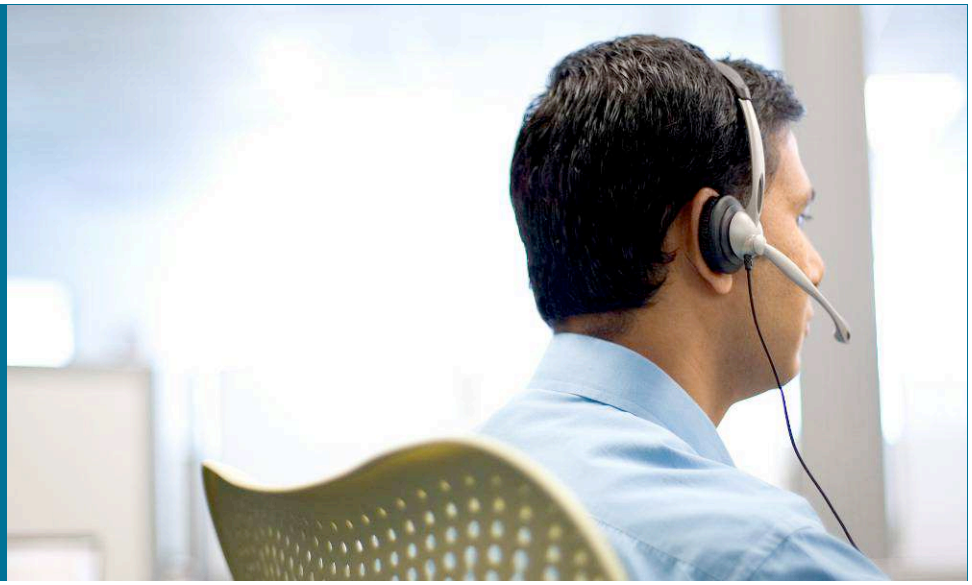


The QFP Packet Processing Chip Set

Will Eatherton (Speaker), Don Steiss (Speaker),

James Markevitch

Cisco Systems,

Cisco Development Organization

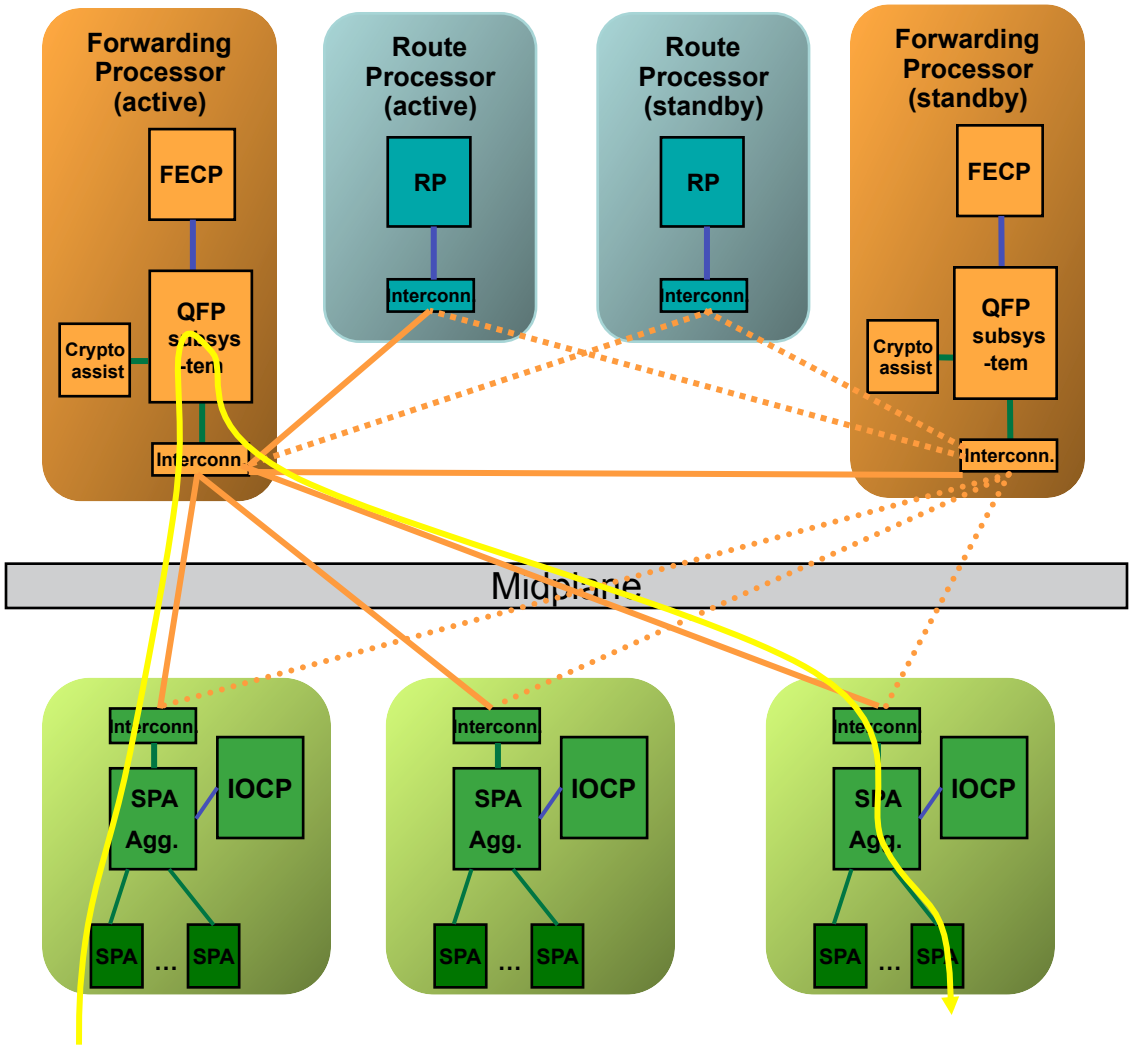
 © 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

Agenda

- Overview of Cisco's ASR-1000 Quantum Flow Processor
- Architecture and Implementation Tradeoffs
- Software Development and Debug Environments
- Silicon Details and Design Methodology
- Results
- Summary

© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

ASR 1000 Building Blocks

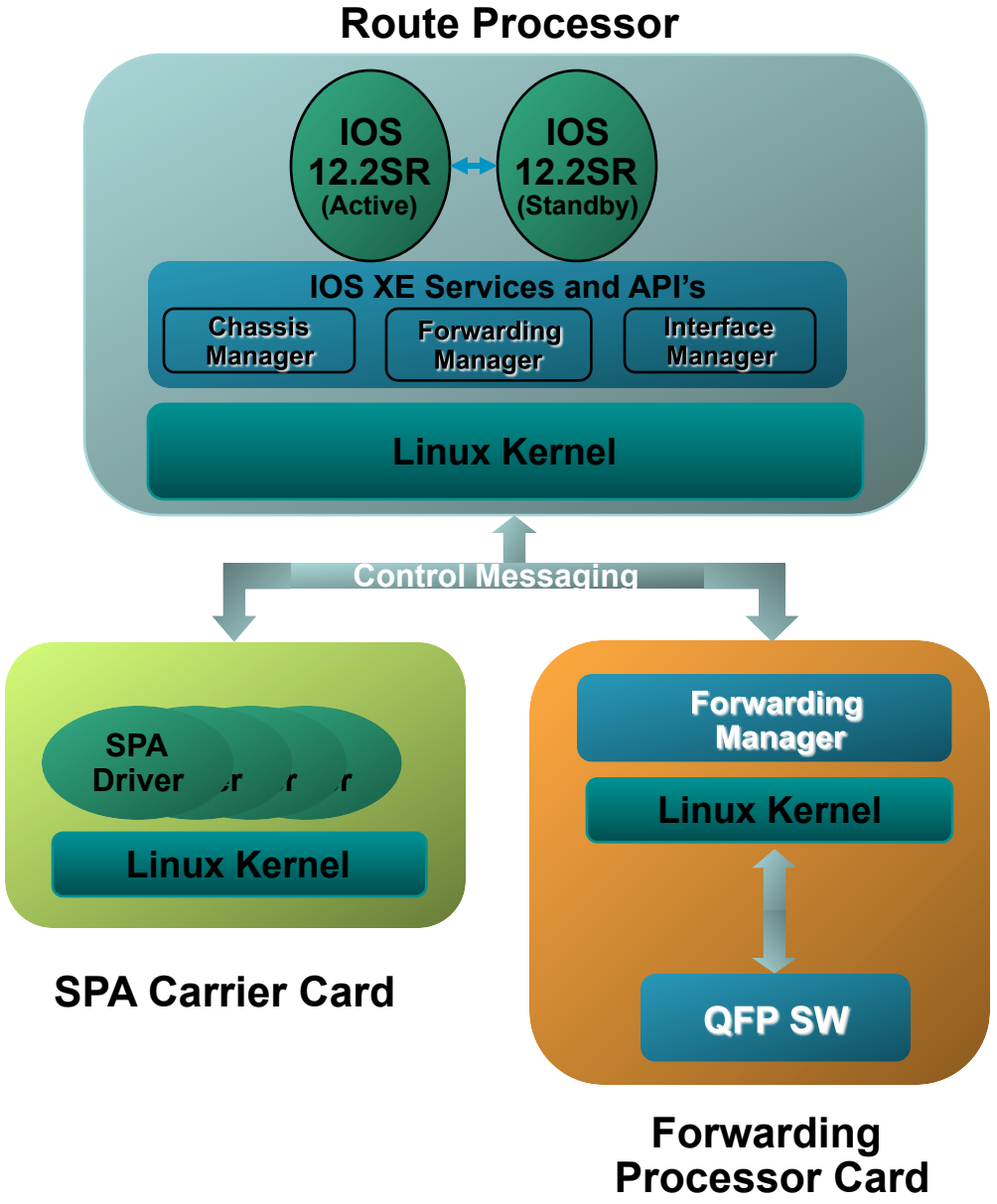


- 3 Chassis types
  - 6 / 4 / 2 RU
- RP (Route Processor)
  - Handles control plane traffic
  - Manages system
- ESP/FP (Forwarding Processor)
  - Handles forwarding plane traffic
- Shared Port Adapter Carrier Card
  - Houses the SPAs
- SPAs
  - Provide interface connectivity
- Centralized Forwarding Architecture
  - All traffic flows through the ESP/FP

© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

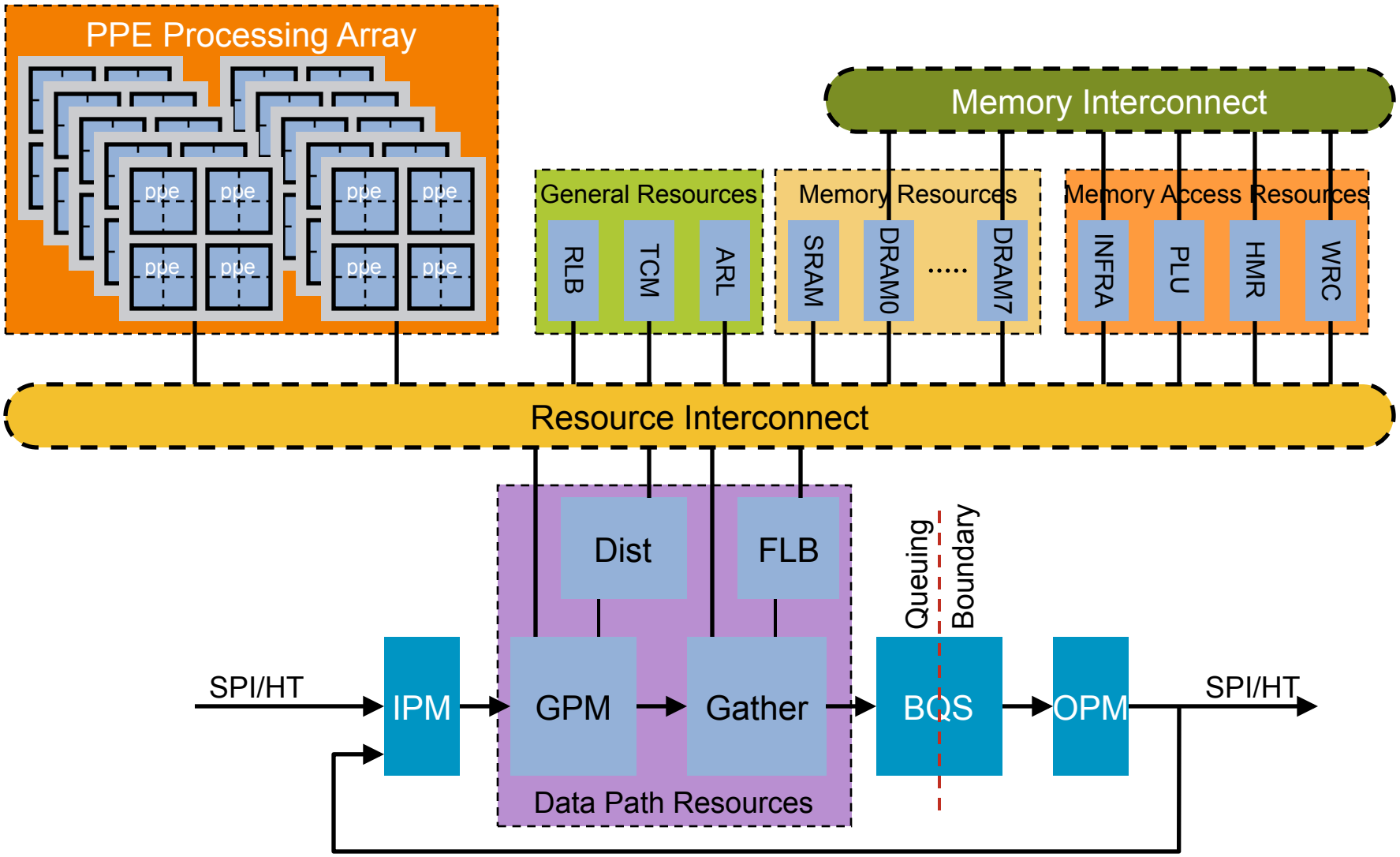
ASR 1000 Software Architecture – IOS XE

- IOS XE is:
  - IOSd
  - plus
  - IOS XE Services/API's
  - plus
  - QFP Datapath Software
- IOS runs as its own Linux process for control plane (Routing, SNMP, CLI etc). Linux kernel with multiple processes running in protected memory
- QFP Datapath Software
  - Fully multiprocessor code base covering range of features across Security, Voice, Deep packet inspection
  - Code base is ANSI-C, written to run based on packet loop and with function libraries for OS like services (Memory Mgmt, Timers, ..)



© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

QFP10 Architecture



© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

Architecture and Implementation Tradeoffs

- Programming Environment

  Assembly code vs. HLL: ANSI C with typical C runtime system for portability and productivity. Assembly coded Hardware Abstraction Layer (HAL).

- Instruction Set Architecture and Implementation

  ISA make vs. buy: off-the-shelf ISA to accelerate production software development.

  Implementation make vs. buy: custom microarchitecture and circuits to improve power, performance and area.

- Latency Hiding

  Implicit vs. explicit mechanisms: Both; threading is implicit, non-blocking messages and data prefetch are explicit.

© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

Architecture Tradeoffs – Continued

- Memory Subsystem

  Ordering: intentionally weak. Loads and stores have no guaranteed MP ordering. Barriers, indivisible operations and serialization are provided. Atomic ordering facilities are provided by special resources.

  D$ parameters: enough capacity to cover parts of the working set with high locality, allocation policies to avoid cache pollution.

  I$ parameters: a large as possible and large second level cache bandwidth to reduce thread performance interference.

- Processor and Resource Communication

  Address mapping and/or message passing: message passing hardware infrastructure with an address mapped layer visible in the programming model.

© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

Architecture Tradeoffs – Continued

- Accelerators: what goes in hardware ?

  Message passing coprocessor: performance (latency hiding), code size, encapsulation.

  Scheduling: performance, encapsulation.

  Crypto: performance (parallelism), stable algorithms.

  Lookup/Classification/Deep Packet Inspection algorithms: performance (parallelism), stable algorithms.

© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

Traffic Manager

- 128K queues and the ability to set Max/Min/Excess BW
- Two Priority Queues can be enabled for each QoS policy applied.
- The number and makeup of layers inside the QoS policy are flexible. The possible hierarchies are not tied to any existing hierarchies used in networks today.
- The traffic manager can schedule multiple levels of hierarchy in one pass through the chip.

  Queuing operations can be split across multiple passes through the chipset (for example, ingress shaping or queuing followed by egress queuing).

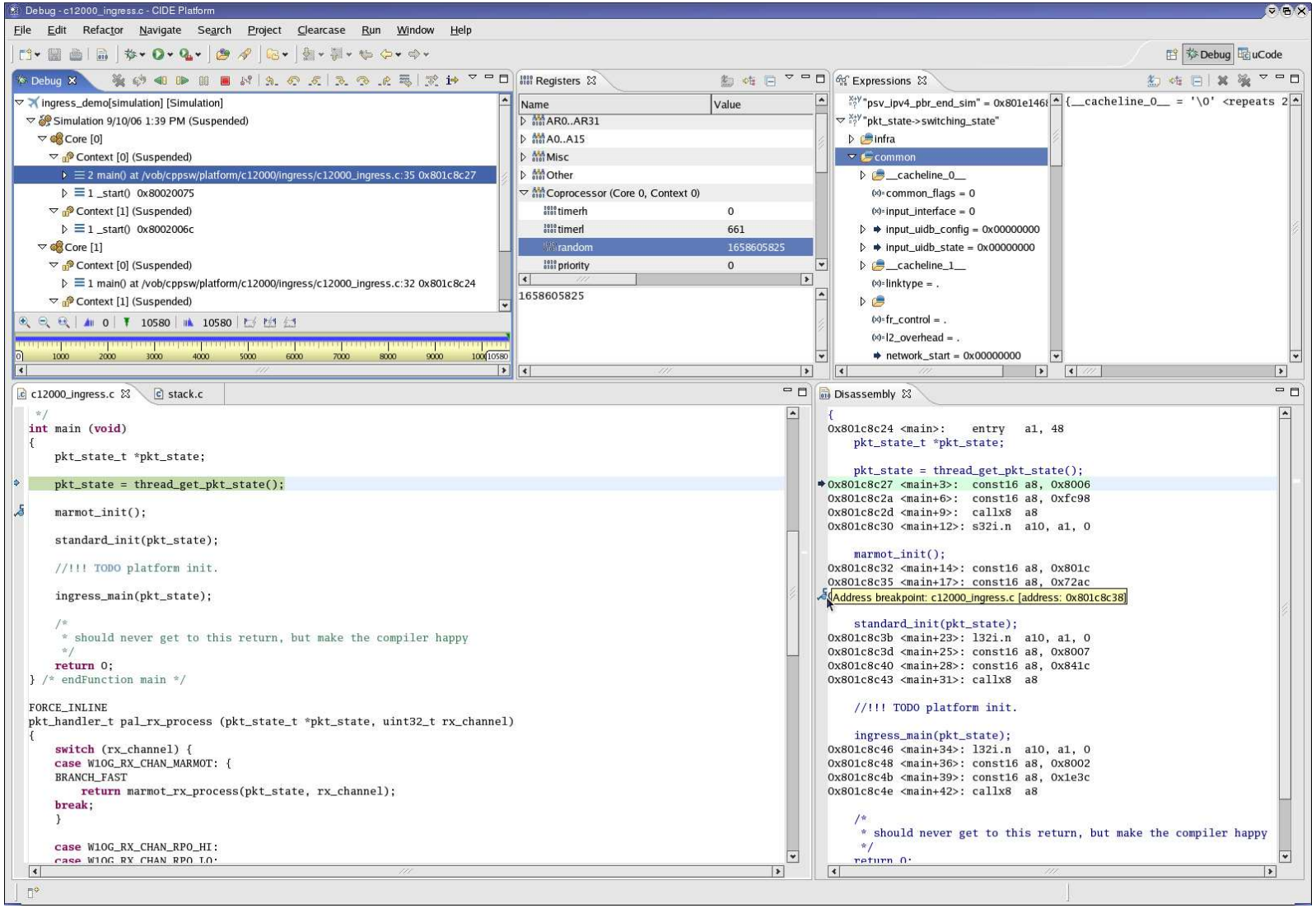
© 2006 Cisco Systems, Inc. All rights reserved.

Software Code Development & Debug Environments

- Simulation, code debug and hardware debug software components integrated in the Eclipse platform.
  - Multiple speed vs. accuracy processor simulation options
  - Compiler, assembler, linker, code debugger
  - Chipset state inspection (registers, interrupt sources)
- Enabled production code development and performance tradeoffs in parallel with hardware development.
- The same environment is available for hardware analysis and debug running on the router's control plane processors.

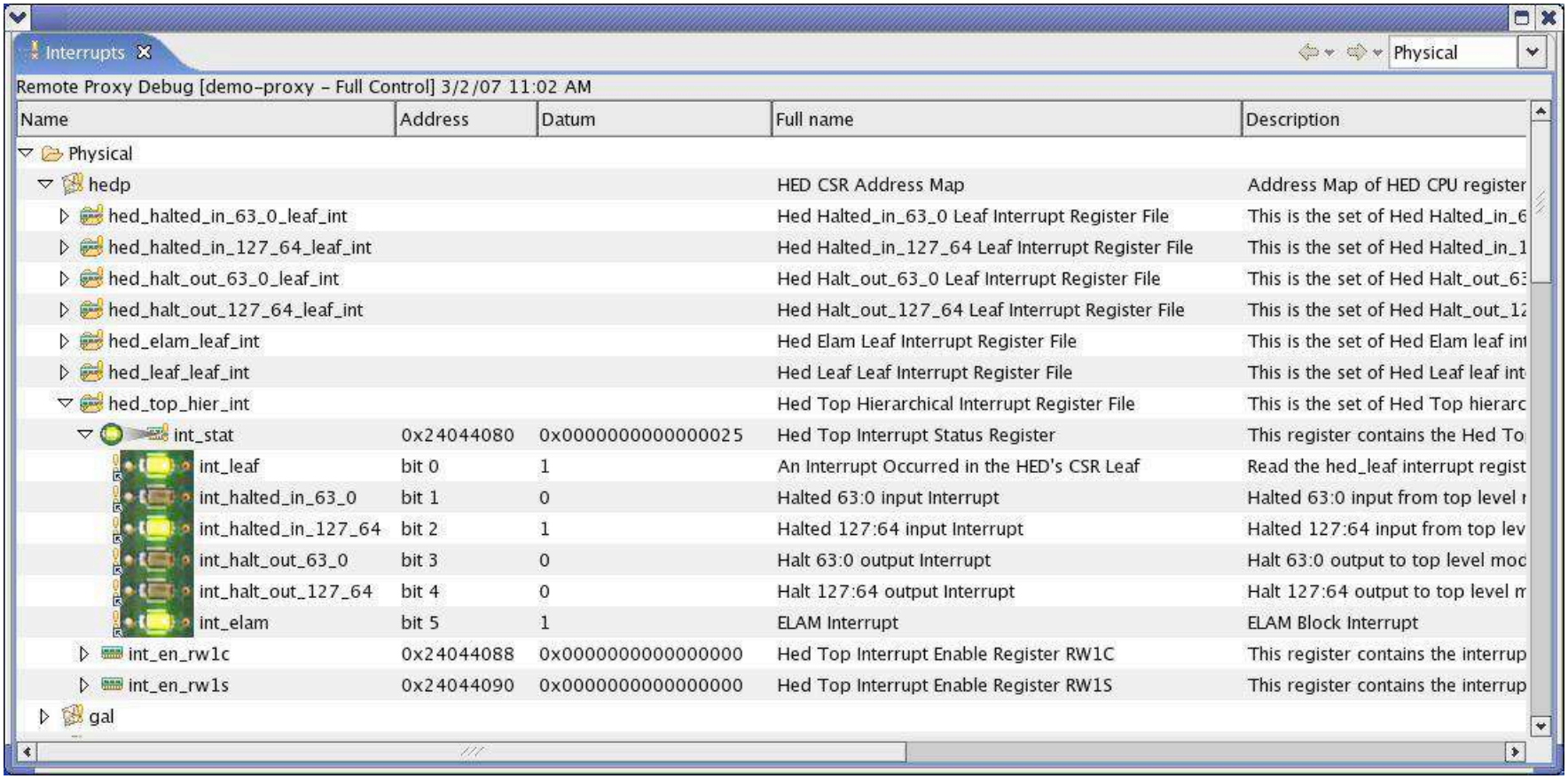
© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

QFP Code Debug Environment



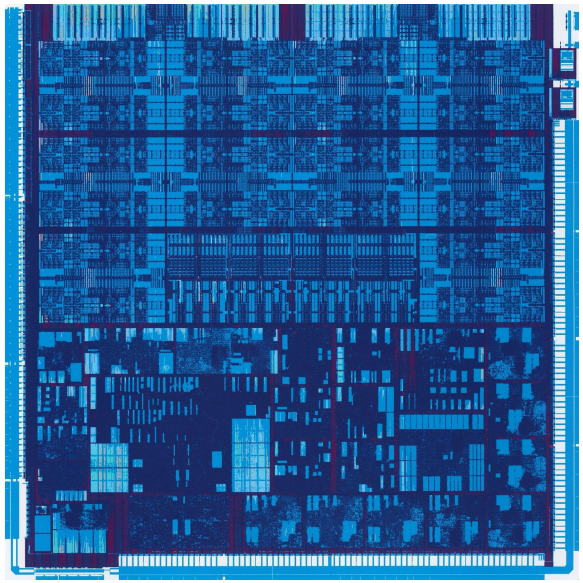
© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

QFP Hardware Debug Environment Example



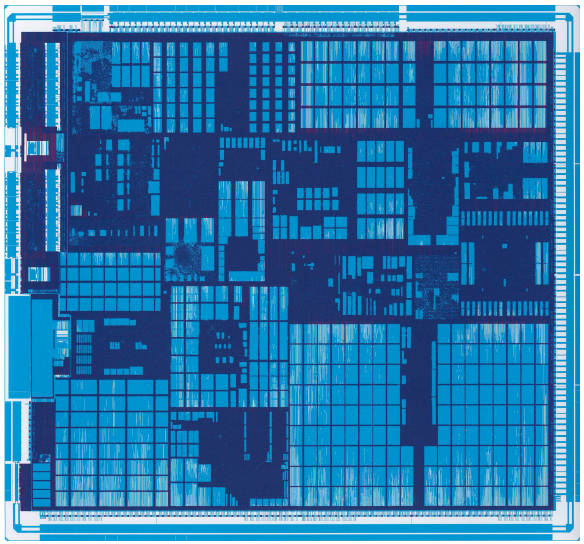
© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

QFP10 Chipset



Multi-Core Packet Processor

- 1.2 GHz/400 MHz
- 40 custom multi-threaded CPUs
- TI 90nm, 8-layers metal
- 19.54 x 19.54 (382 mm^2)
- 307 million transistors
- 20 Mb SRAM
- 1019 I/O, including 800 MHz DDR

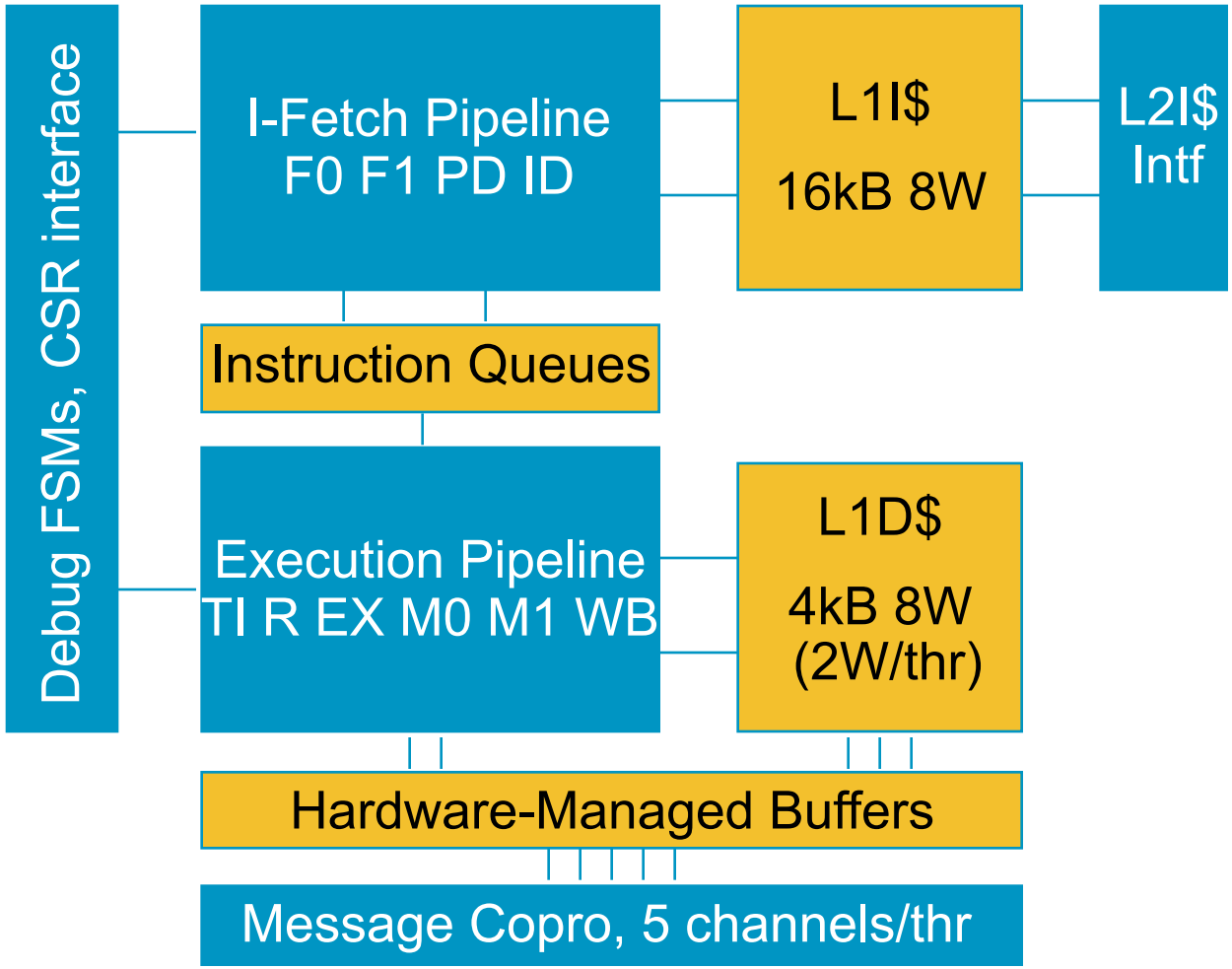


Traffic Manager and Interface Chip

- 400 MHz
- Buffering, 200K queues, hardware HQF scheduling
- TI 90nm, 8-layers metal
- 19.0 x 17.48 (332 mm^2)
- 522 million transistors
- 70 Mb SRAM
- 1318 I/O, including 800 MHz DDR

© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

Packet Processing Engine (PPE) Structure



© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

Hardware Design Methodology

- Customer Owned Tooling (COT) with GDSII handoff to the foundry partner.
  - Cell library designed and characterized by Cisco
  - SRAM compiler and I/O IP from Texas Instruments.
  - Extensive crosschecking at Cisco and Texas Instruments.
- Synthesis/P&R outside of processor array
  - Multiple commercial synthesis and physical design tools.
  - In-house tools for RTL code generation and documentation.

© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

Hardware Design Methodology - Processor

- Cell-based to leverage ASIC tools.
- Static CMOS with selective use of domino circuits
- Schematic design entry with physical specifications in instance names for fast, deterministic placement.
- Polygon-level artwork in critical modules.
- Autorouted signals with many pre-routes above the cell level
- Multiple functional reference models and equivalence checking

© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

Processor Design Verification

- Strong block level verification methodology
- Heavy use of code generators and constraint solving
- Processor verification statistics:

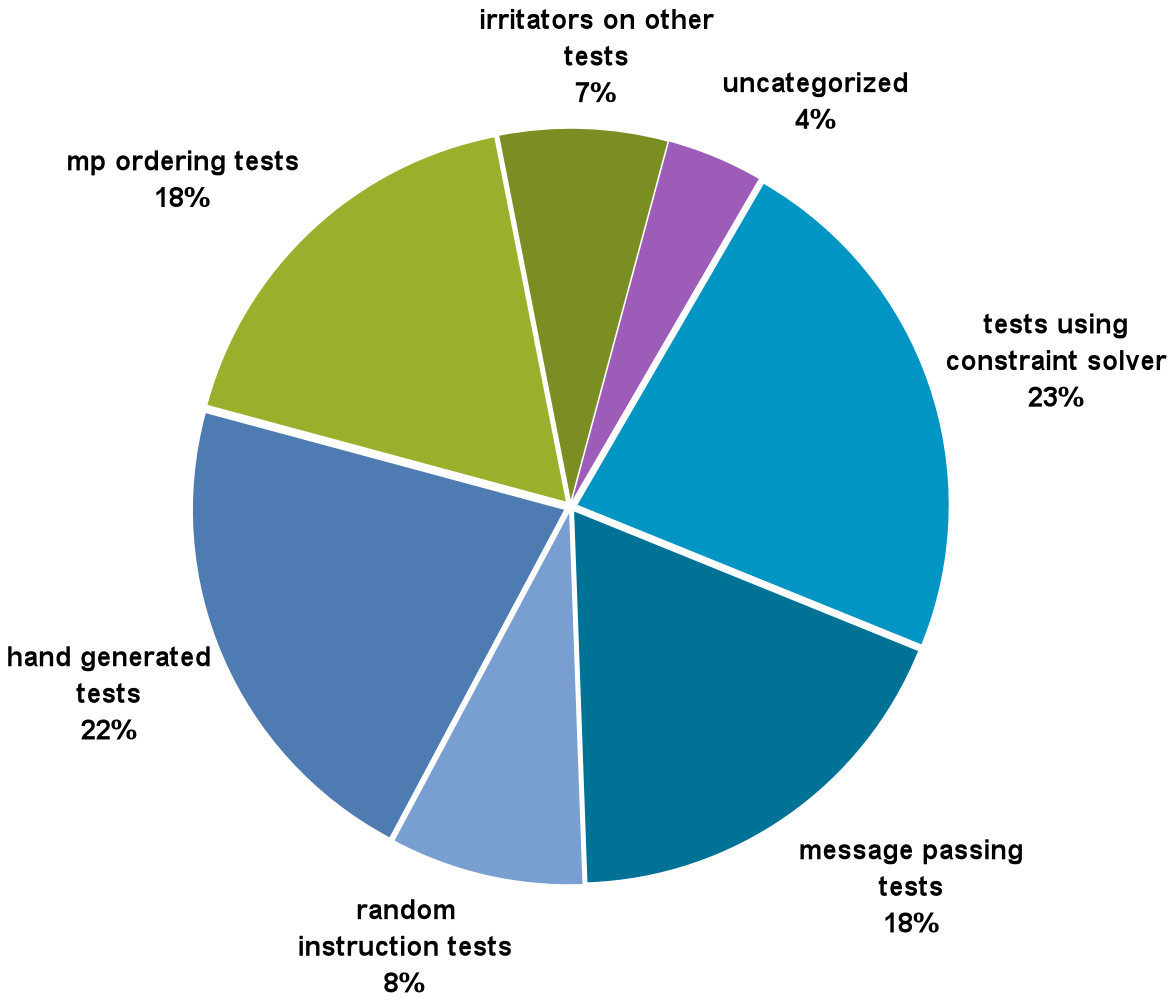
  1,903,282 RTL simulation runs

  95,185,665,935 total clock cycles

  17,421 total test failures

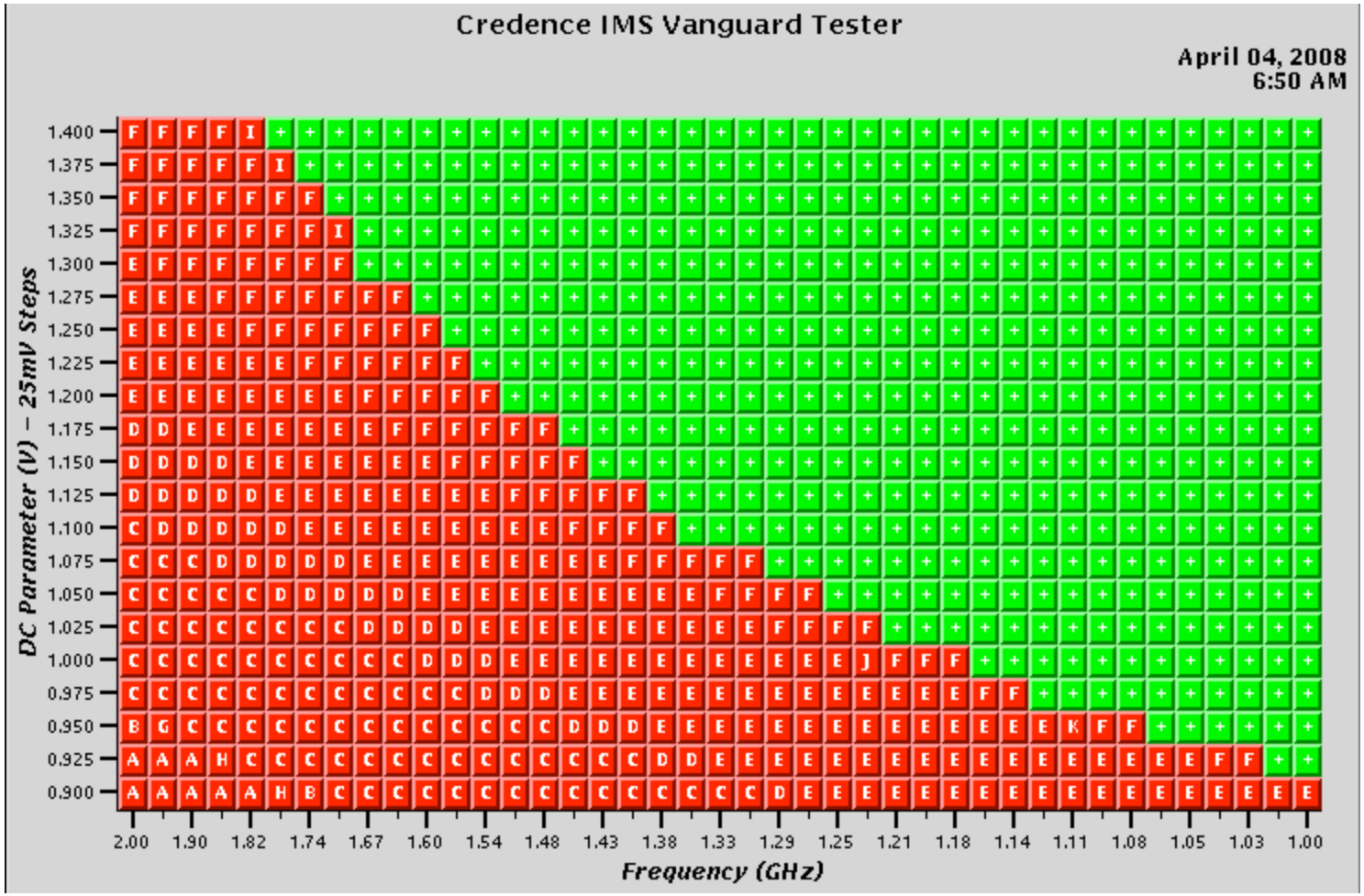
© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

Processor Design Verification – Continued
Failures by Test Source



© 2006 Cisco Systems, Inc. All rights reserved.

Results – Shmoo at Room Temp



© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

Results - Functional

- On rev 1.0 silicon:

  Silicon delivered for system integration, January 2007

  First packets through production software, February 2007

  Customer testing, August 2007

- Product Launch March 4, 2008
- All “mission mode” registers and instructions in the ISA are used in production software.

© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

Summary

- Router Architecture today involves partitioning of software across multiple CPU complexes, Multiple Cores per CPU, Multiple threads per Core
- Processor architecture in networking is still evolving
- Many architectural and implementation tradeoffs are a result of software engineering complexity that rivals hardware complexity and often 10x more staffing.

© 2006 Cisco Systems, Inc. All rights reserved. Cisco Public

© 2006 Cisco Systems, Inc. All rights reserved.