Topic & Scope

- **Content:** The course gives ...
  - ... an overview of network processor cards (architectures and use)
  - ... an introduction of how to program Intel IXP network processors
  - ... some ideas of how to use network processors in a multimedia system

Available Resources

- Book:
  Douglas E. Comer: “Network Systems Design using Network Processors”, Pearson Prentice Hall, 2004

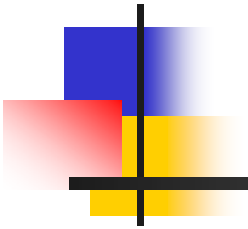
- Other resources will be placed at
  - http://www.ifi.uio.no/~paalh/INF5060
  - Login: *inf5060*
  - Password: *ixp*

- Manuals for IXP1200: .../~paalh/INF5060/IXP1200

- Code: .../~paalh/INF5060/code

Disclaimer

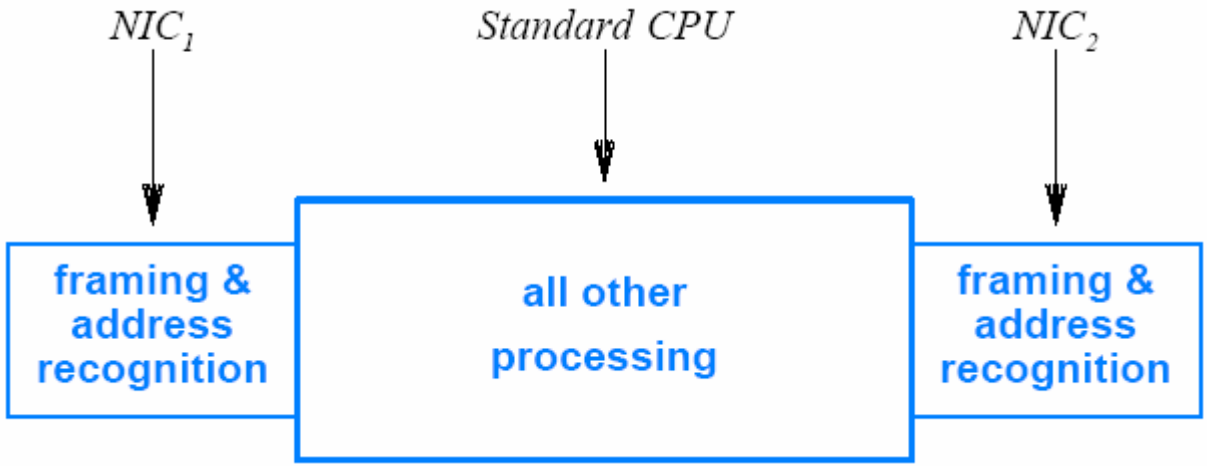
- In the field of network processors, *I am a tyro*
- Definition:
  Tyro ₩*Ty'ro*₩, n.; pl. *Tyros*. A beginner in learning; one who is in the rudiments of any branch of study; a person imperfectly acquainted with a subject; a novice
- Then, by definition,
  in the field of network processors, *we are all tyros*
- In our defense,
  when it comes to network processors, *everyone is a tyro*



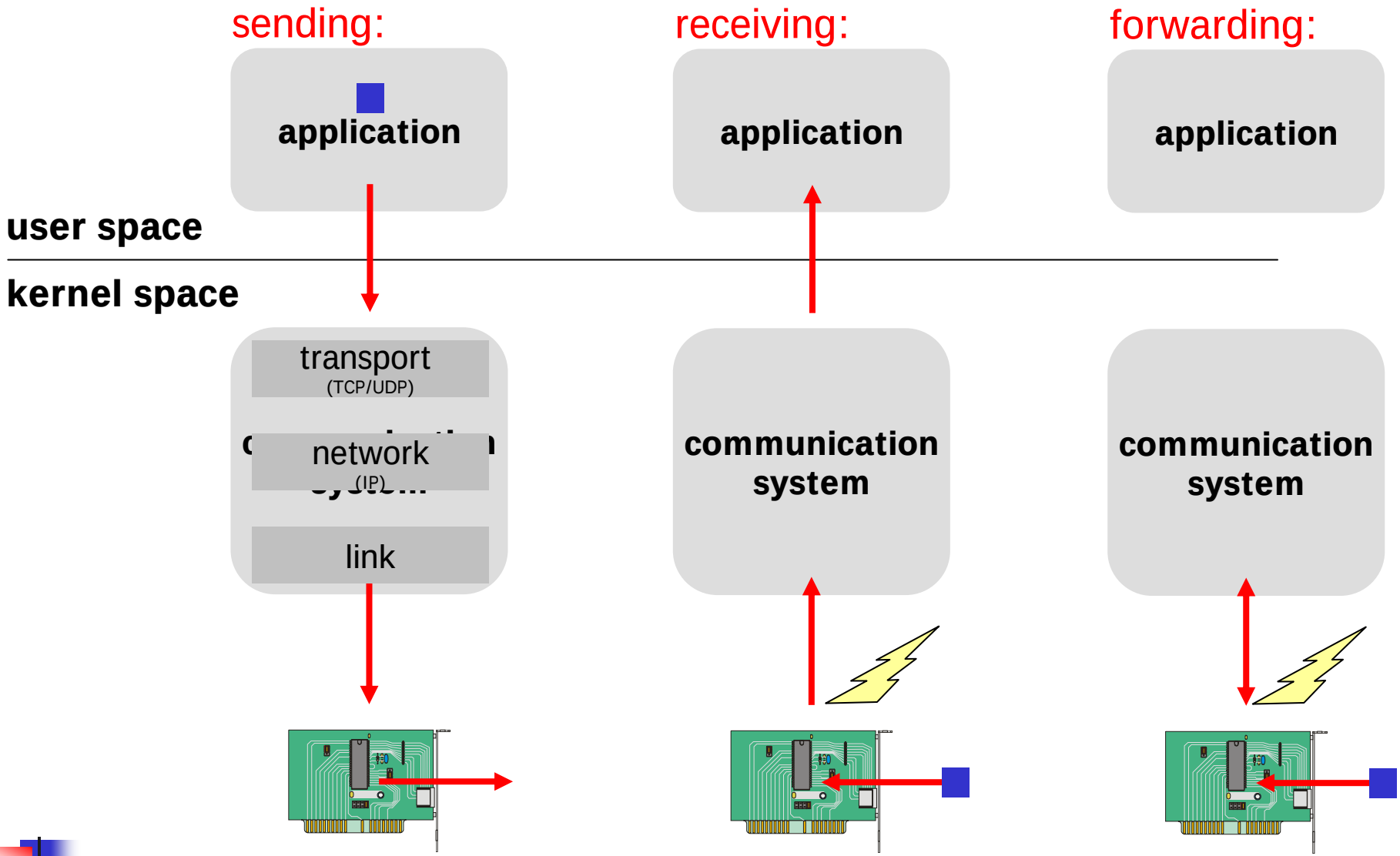
Background and Motivation

Software-Based Network System

- ✓ Uses conventional, shared hardware (e.g., a PC)
- ✓ Software
  - runs the entire system
  - allocates memory
  - controls I/O devices
  - performs all protocol processing
- ✓ First generation network systems:

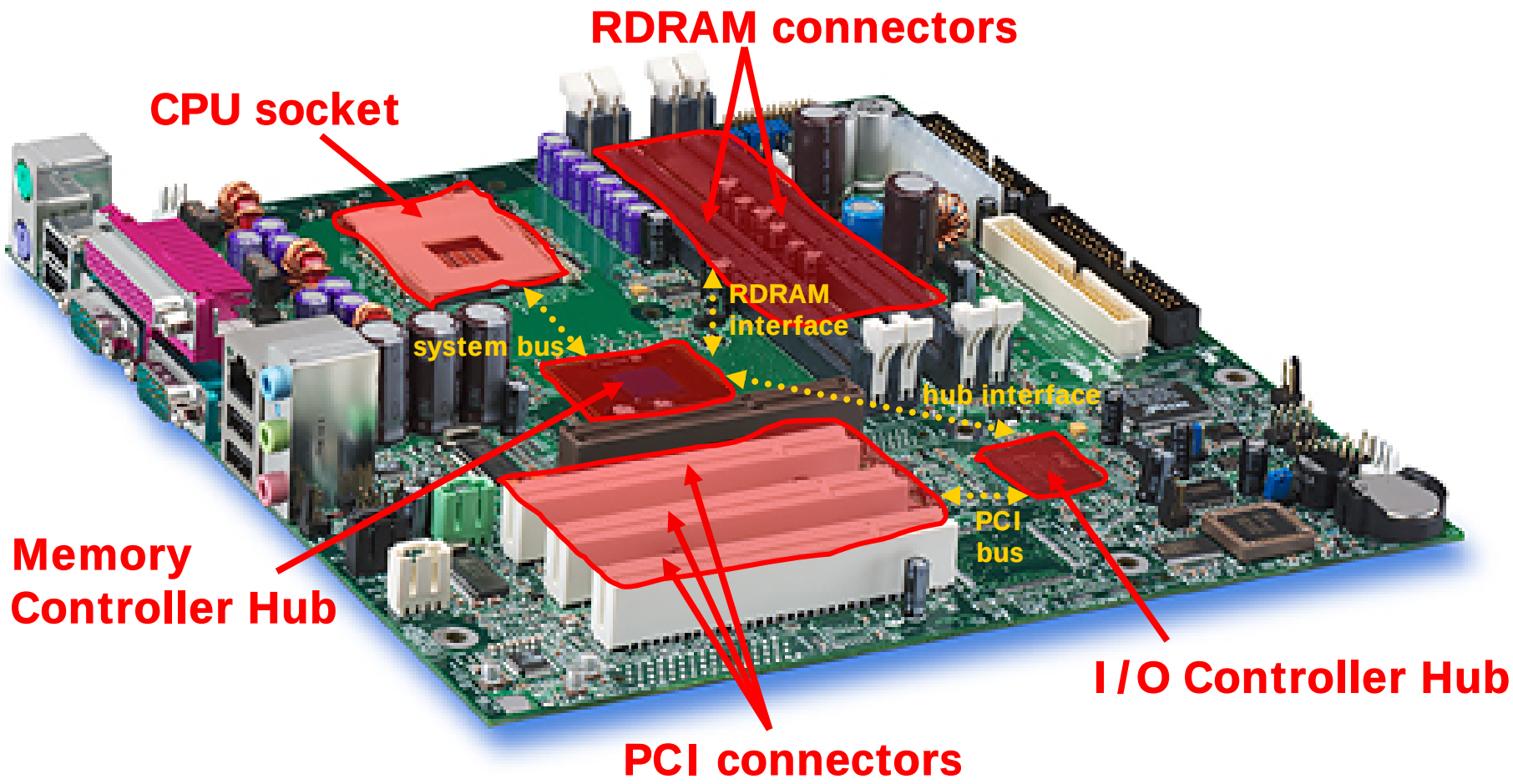


Review of General Data Path on Conventional Computer Hardware Architectures



Review of Conventional Computer Hardware Architectures

Intel D850MD Motherboard - Intel Hub Architecture (850 Chipset) :



Forwarding Example for an Intermediate Node: Intel Hub Architecture

application

user space

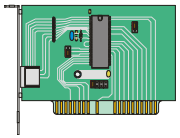
kernel space

communication system

Note:

- one single average MPEG-II DVD stream require ~330-660 packets per second of 1500 Bytes (4-8 Mbps)

- then use smaller packets, add concurrent clients, other applications, ...



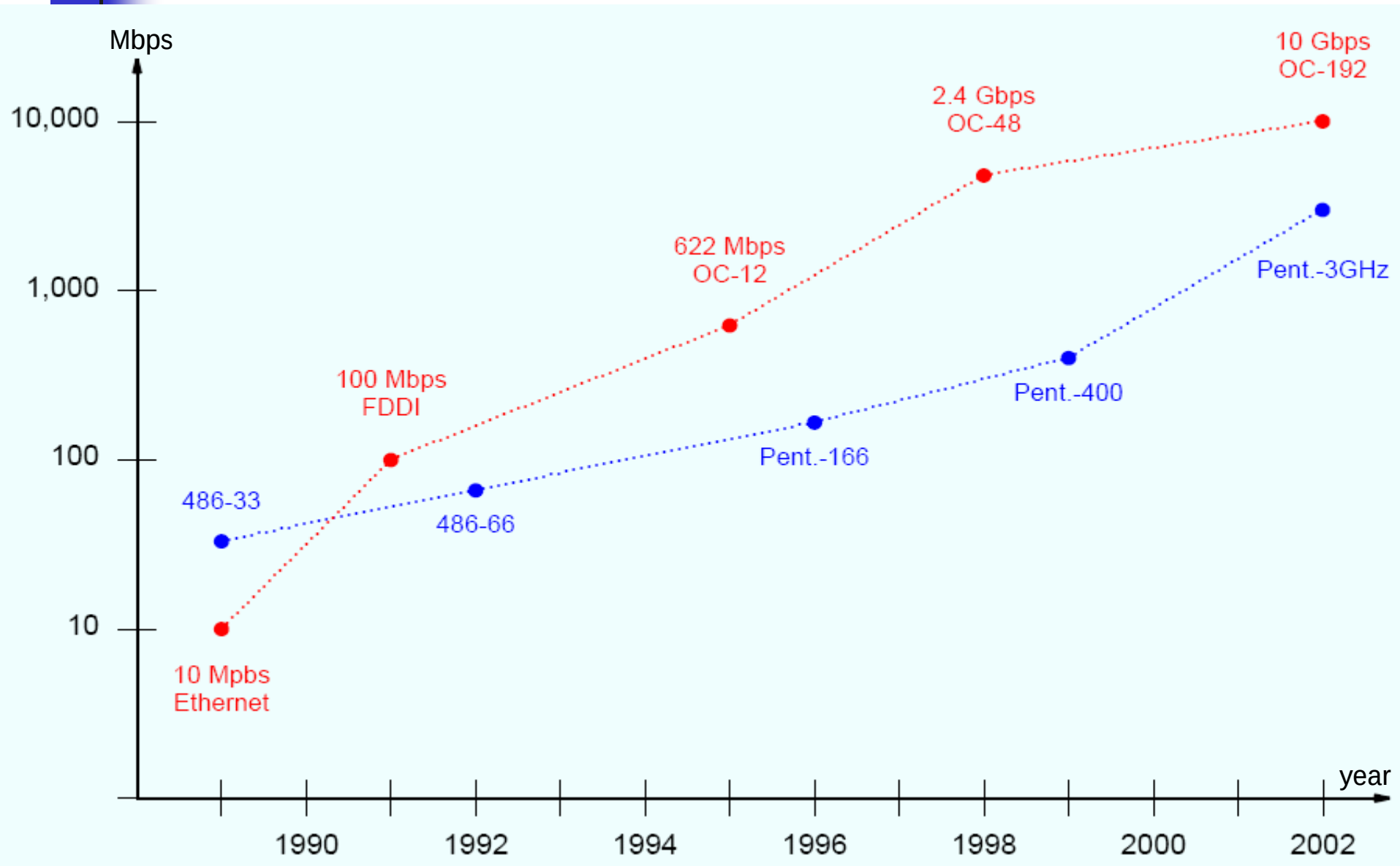
Main Packet Processing Costs

- Copying: used when moving a packet from one memory location to another
  - expensive (proportional to packet size)
  - should be avoided whenever possible (use pointers)
- Checksuming: used to detect errors
  - expensive (proportional to packet size)
  - transport layer: payload + header
  - network layer: header
- Fragmentation/reassembly: needed when packet is larger than smallest MTU
  - generate headers + header checksum
  - receiving many small data fragments

Question:

- Which is growing faster?
  - network bandwidth
  - processing power
- Note: if network bandwidth is growing faster
  - CPU may be the bottleneck
  - need special-purpose hardware
  - conventional hardware will become irrelevant
- Note: if processing power is growing faster
  - no problems with processing
  - network/busses will be bottlenecks

Growth Of Technologies



Packet Rates and Software Processing

- Packet rates (packets per second):

	64 B	1500 B
10BASE-T (10 Mbps)	19.531	833
1000BASE-T (1 Gbps)	1.953.125	83.333
OC-192 (9.95 Gbps)	19.439.453	829.416

- Packet processing (MIPS, assuming 5K instructions per packet):
  - the Comer book uses 10K instructions as an upper bound per packet
  - it varies according to which protocols are used, implementation, data size, etc.
  - more if moved through a fire wall

	64 B	1500 B
10BASE-T (10 Mbps)	97,65	4,17
1000BASE-T (1 Gbps)	9.765,63	416,67
OC-192 (9.95 Gbps)	97.197,27	4.147,08

  - engineering rule: 1GHz general purpose CPU = 1Gbps network data rate

- Note; this is only processing – time must be added to handle interrupts and move data into memory

- Thus, software running on a general-purpose processor is insufficient to handle high-speed networks because the aggregate packet rate exceeds the capabilities of the CPU

The Network System Challenges

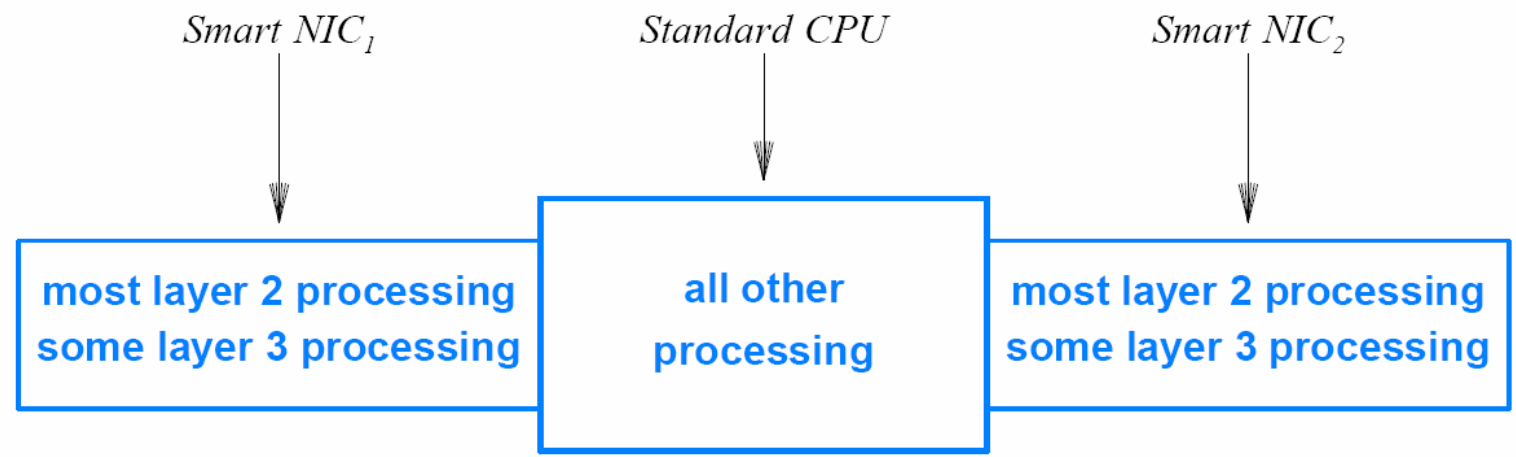
- ✓ Data rates in general keep increasing
- ✓ Network rate > CPU rate > memory, busses and I/O interfaces
- ✓ Protocols and applications keep evolving
- ✓ System design, implementation and testing is time consuming and expensive
- ✓ Systems often contain errors
- ✓ Special-purpose hardware (ASIC) designed for one type of system can usually not be reused
- ✓ Host machine must inspect all incoming packets
- ✓ ...
- ⇨ Challenge: find ways to improve the design and manufacture of complex networking systems

Statement of Hope

- If there is hope, it lies in ...
  - 1990: ... faster CPUs
  - 1995: ... the application specific integrated circuit (ASIC) designers
  - 2002: ... the programmers!
- Programmability
  - we need a programmable device with more capability than a conventional CPU
  - key to low-cost hardware for next generation network systems
  - compared to ASIC designs, it is more flexible, easier and faster to upgrade, and thus, less expensive

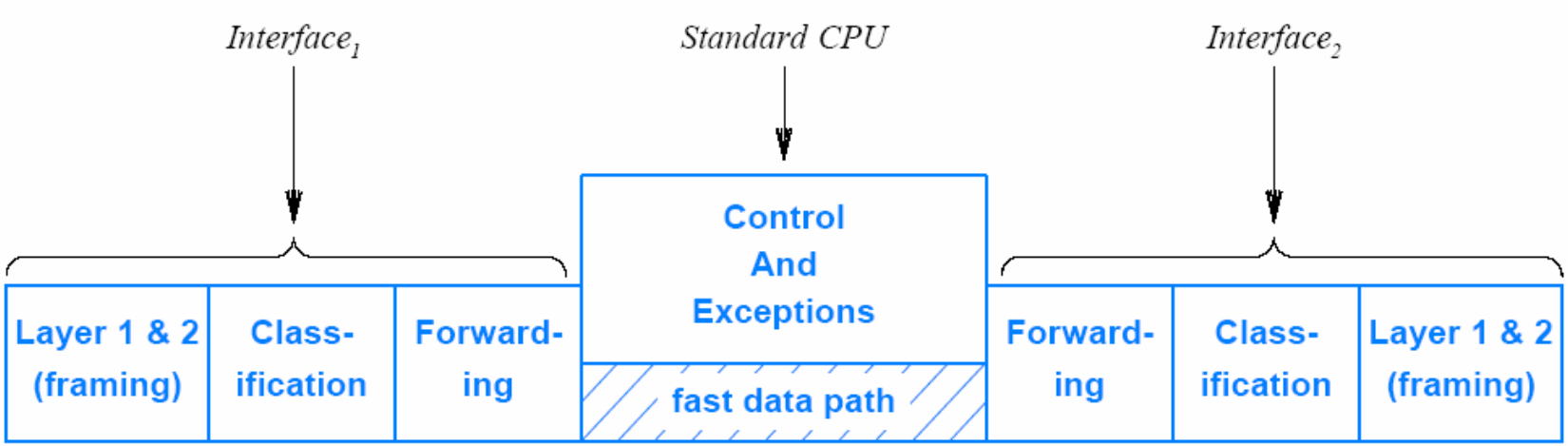
First Generation

- General idea: To optimize computation, move operations that account for the most CPU time from software into hardware
- Onboard...
  - address recognition and filtering
  - onboard buffering
  - DMA
  - buffer and operation chaining
- Add hardware to NIC
  - off-the-shelf chips for layer 2
  - ASICs for layer 3
- Allows each NIC to operate independently
  - effectively a multiprocessor
  - total processing power increased dramatically



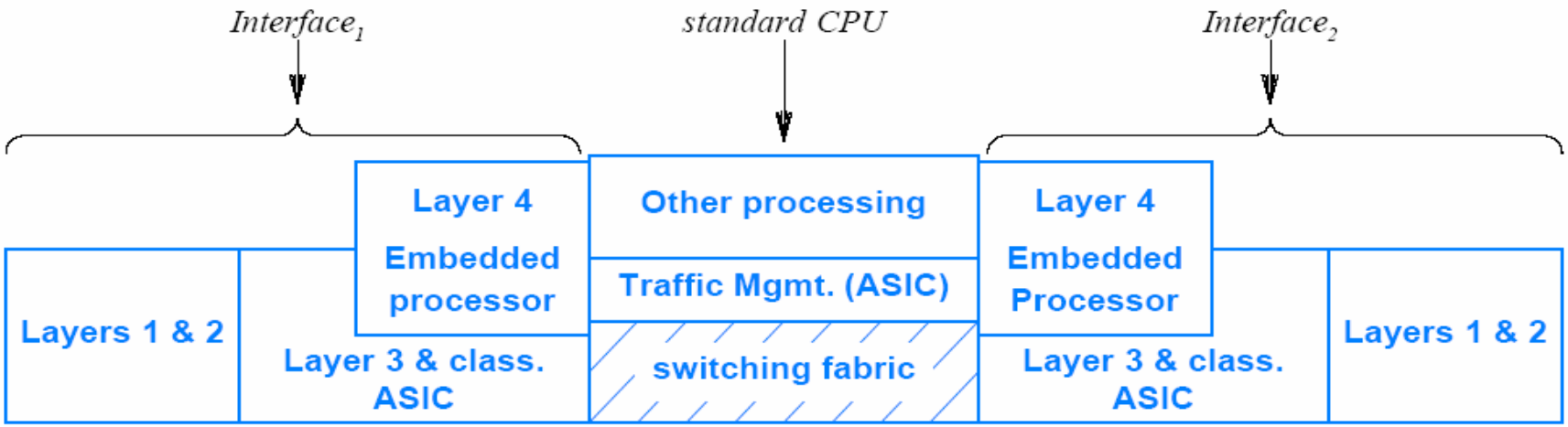
Second Generation (early 1990s)

- ✓ Designed for greater scale
- ✓ Decentralized architecture
  - additional computational power on each NIC
  - NIC implements classification and forwarding
- ✓ High-speed internal interconnection mechanism
  - interconnects NICs
  - provides *fast data path*
- ✓ Multiple network interfaces
- ✓ High-speed hardware interconnects NICs
- ✓ General-purpose processor only handles exceptions
- ✓ Sufficient for medium speed interfaces (100 Mbps)



Third Generation (late 1990s)

- ✓ Functionality partitioned further
- ✓ Additional hardware on each NIC
- ✓ Onboard...
  - ➢ classification
  - ➢ forwarding
  - ➢ traffic policing
  - ➢ monitoring and statistics
  - ➢ ...
- ✓ Almost all packet processing off-loaded from CPU
  - ➢ Special-purpose ASICs handle lower layer functions
  - ➢ Embedded (RISC) processor handles layer 4
  - ➢ CPU only handles low-demand processing



Third Generation (late 1990s)

- Enough, are third generation sufficient??
  - Almost!!
  - But not quite! ;-(
- What's the problem?
  - high cost
  - long time to market
  - difficult to test
  - expensive and time-consuming to change
    - even trivial changes require silicon respin
    - 18-20 month development cycle
  - little reuse across products and versions
  - require in-house expertise (ASIC designers)
  - ...

Network Processors: The Idea in a Nutshell

- ✓ Devise new hardware building blocks, but make them *programmable*
- ✓ Include support for protocol processing and I/O
  - ➢ General-purpose processor(s) for control tasks
  - ➢ Special-purpose processor(s) for packet processing and table lookup
- ✓ Include functional units for tasks such as checksum computation, hashing, ...
- ✓ Integrate as much as possible onto one chip
- ⇨ Call the result a *network processor*

Designing a Network Processor

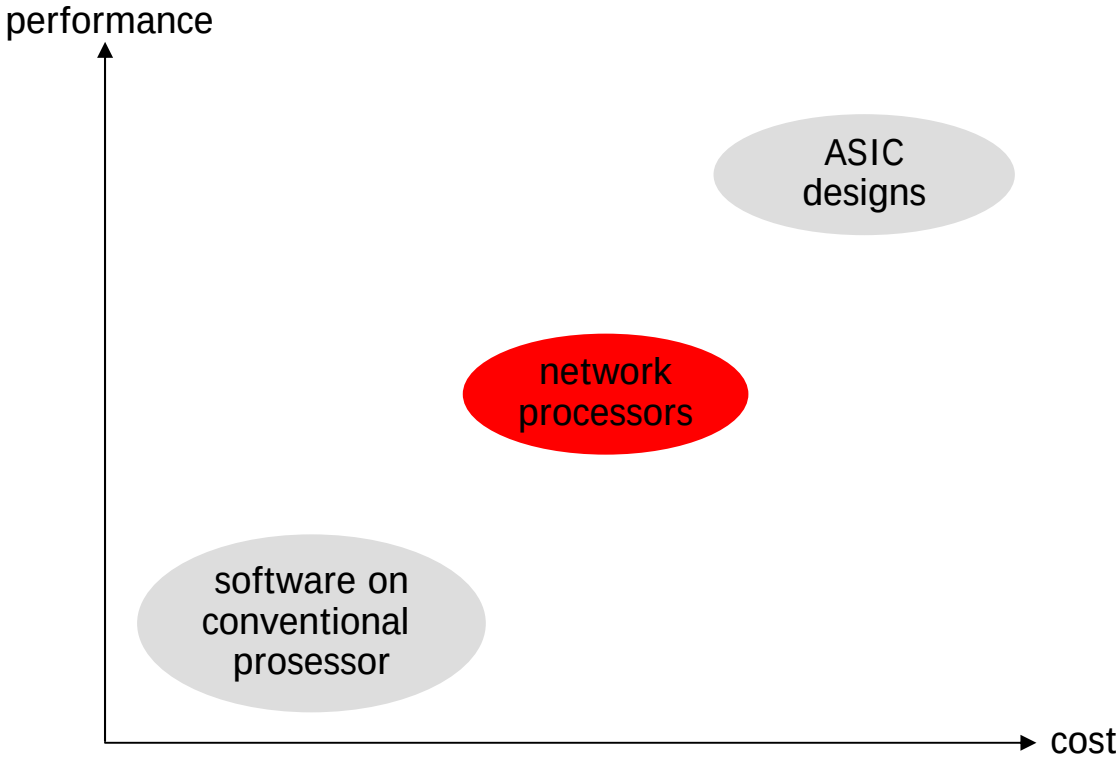
- Depends on
  - operations network processor will perform
  - role of network processor in overall system

- Goals
  - generality: sufficient for all protocols, all protocol processing tasks and all possible networks
  - high speed: scale to high bit rates and high packet rates

- Key point:

  A network processor is not designed to process a specific protocol or part of a protocol. Instead, designers seek a minimal set of instructions that are sufficient to handle an arbitrary protocol processing task at high speed

Where to Place Network Processors



- ✓ Thus, network processors is somewhere in the middle
- ✓ Goal: increase performance and reduce costs
- ✓ **Increase performance:**

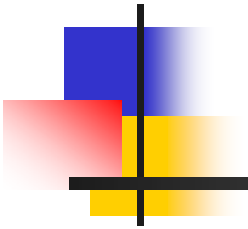
  known issues:
  - must partition packet processing into separate functions
  - to achieve highest speed, must handle each function with separate hardware

  unknown issues:
  - which functions to choose
  - what hardware building blocks to use
  - how to interconnect building blocks
- ✓ **Decrease costs:**

  Economics driving a gold rush
  - NPs will dramatically lower production costs for network systems
  - good NP designs worth lots of $$

Explosion of Commercial Products

- 1990 → 2000: network processors transformed from interesting curiosity to mainstream product
  - used to reduce both overall costs and time to market
  - 2002: over 30 vendors with a vide range of architectures
  - e.g.,
    - Multi-Chip Pipeline (Agere)
    - Augmented RISC Processor (Alchemy)
    - Embedded Processor Plus Coprocessors (Applied Micro Circuit Corporation)
    - Pipeline of Homogeneous Processors (Cisco)
    - Pipeline of Heterogeneous Processors (EZchip)
    - Configurable Instruction Set Processors (Cognigine)
    - Extensive And Diverse Processors (IBM)
    - Flexible RISC Plus Coprocessors (Motorola)
    - Internet Exchange Processor (Intel)
    - ...



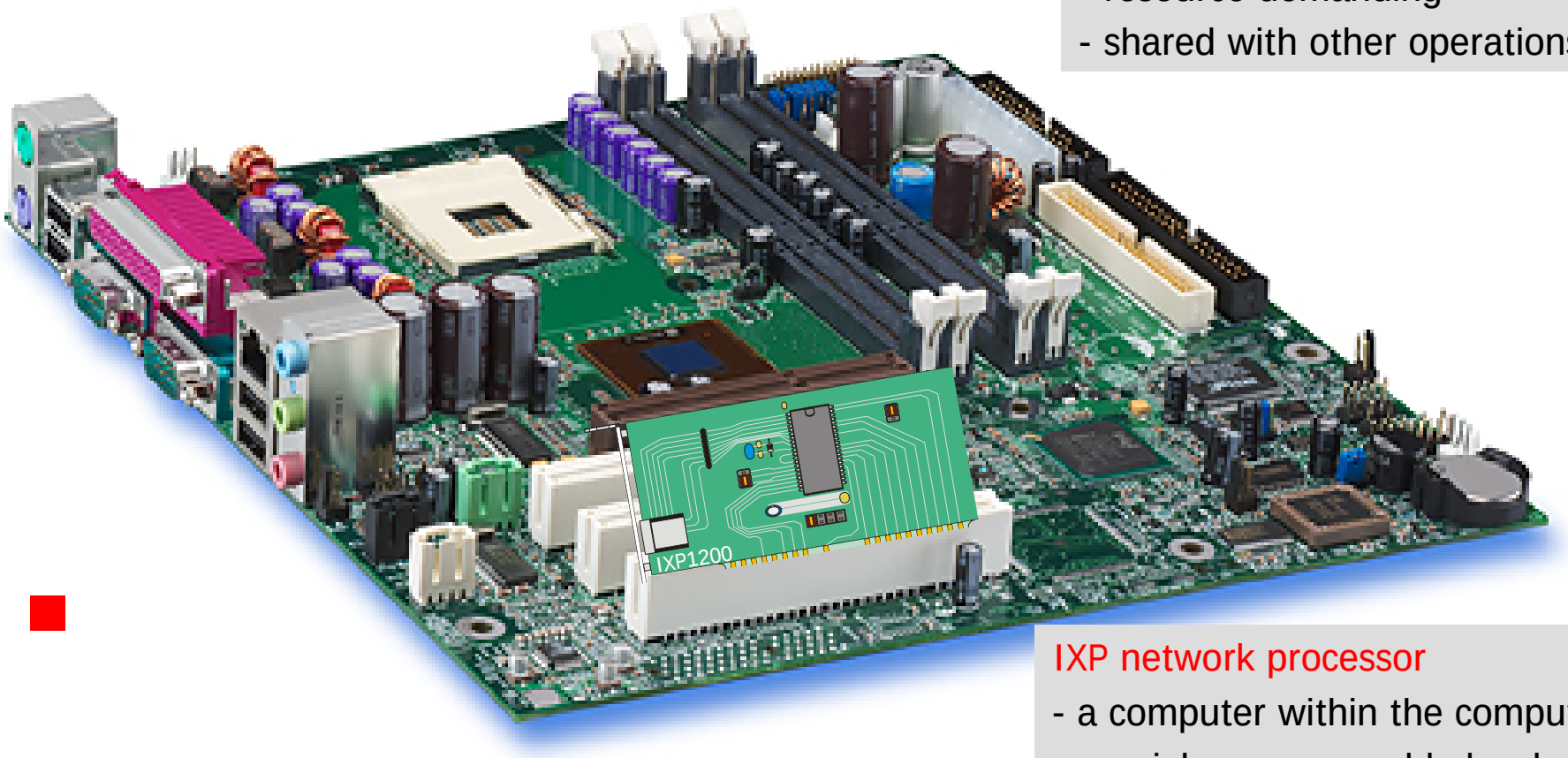
IXP1200:
A Short Overview

IXA: Internet Exchange Architecture

- IXA is a broad term to describe the Intel network architecture (HW & SW, control- & data plane)
- IXP: Internet Exchange Processor
  - processor that implements IXA
  - IXP1200 is the first IXP chip (4 versions)
- IXP1200 basic features
  - 1 embedded RISC processor
  - 6 packet processors (microengines)
  - multiple, independent busses
  - onboard memory (3 types)
  - low-speed serial interface
  - interfaces for external memory and I/O busses
  - ...

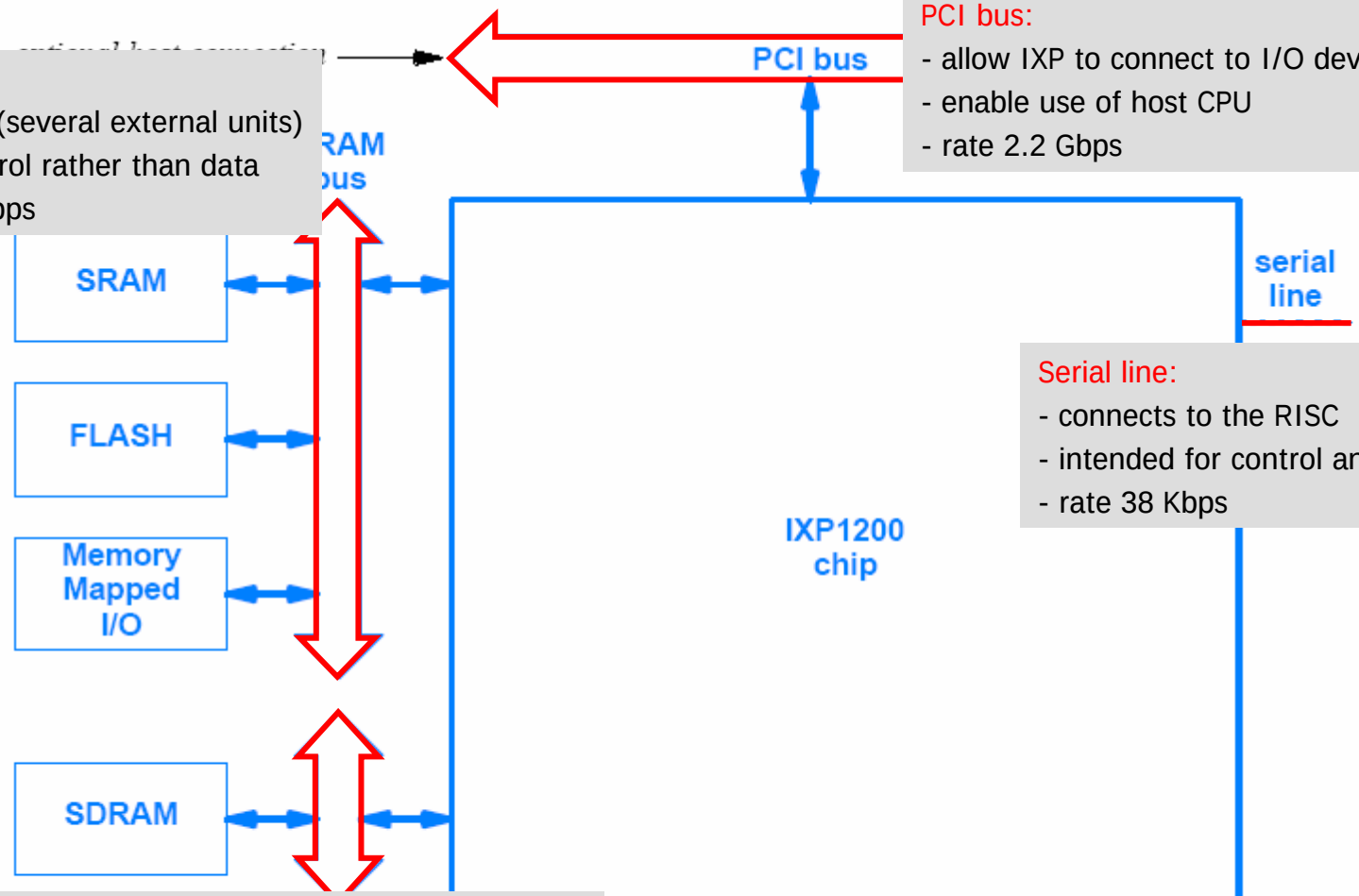
Main Idea

Traditional system:
- slow
- resource demanding
- shared with other operations



IXP network processor
- a computer within the computer
- special, programmable hardware
- offloads host resources

IXP1200 Architecture



SRAM bus:
- shared bus (several external units)
- usually control rather than data
- rate 3.71 Gbps

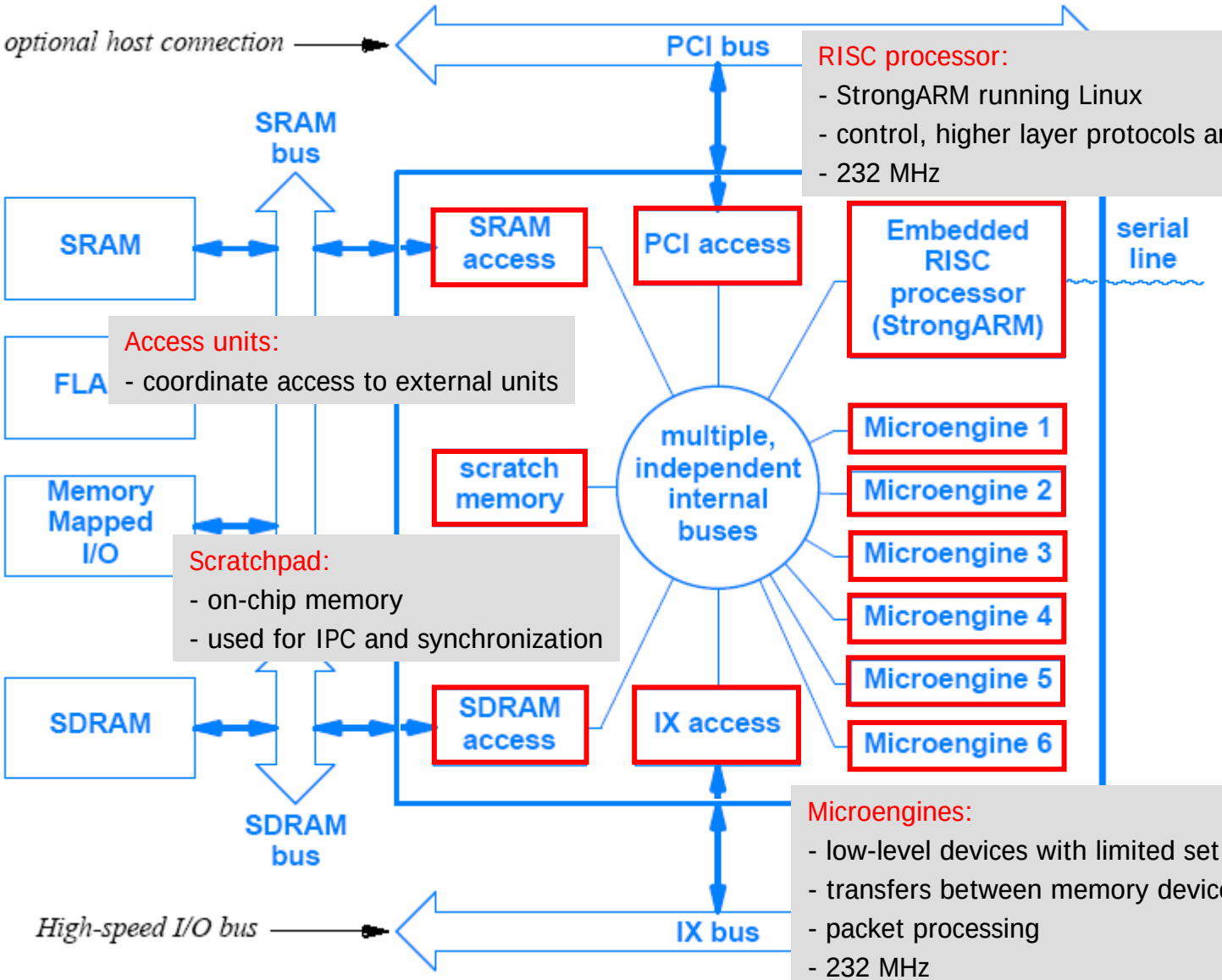
PCI bus:
- allow IXP to connect to I/O devices
- enable use of host CPU
- rate 2.2 Gbps

Serial line:
- connects to the RISC
- intended for control and management
- rate 38 Kbps

SDRAM bus:
- provide access to external SDRAM memory used to store packets
- can also pass addresses, control/store operations, etc.
- rate 7.42 Gbps

IX (Intel eXchange) bus:
- enable higher rates compared to PCI
- form fast path (IXP and high-speed interfaces)
- interface to other IXP cards
- 4.4 Gbps

IXP1200 Architecture



RISC processor:
- StrongARM running Linux
- control, higher layer protocols and exceptions
- 232 MHz

Access units:
- coordinate access to external units

Scratchpad:
- on-chip memory
- used for IPC and synchronization

Microengines:
- low-level devices with limited set of instructions
- transfers between memory devices
- packet processing
- 232 MHz

IXP1200 Processor Hierarchy

Processor Type	Onboard?	Programmable?

General-Purpose Processor:
- used for control and management
- running general applications

RISC processor:
- chip configuration interface (serial line)
- control, higher layer protocols and exceptions

I/O processors (microengines):
- transfers between memory devices
- packet processing

Coprocessors:
- real-time clock and timers
- IX bus controller
- hashing unit
- ...

Physical interface processors:
- implement layer 1 & 2 processing

IXP1200 Memory Hierarchy

Memory Type	Maximum Size	On Chip?	Typical Use
GP Registers	128 regs.	yes	Intermediate computation
Inst. Cache	16 Kbytes	yes	Recently used instructions
Data Cache	8 Kbytes	yes	Recently used data
Mini Cache	512 bytes	yes	Data that is reused once
Write buffer	unspecified	yes	Write operation buffer
Scratchpad	4 Kbytes	yes	IPC and synchronization
Inst. Store	64 Kbytes	yes	Microengine instructions
FlashROM	8 Mbytes	no	Bootstrap
SRAM	8 Mbytes	no	Tables or packet headers
SDRAM	256 Mbytes	no	Packet storage

IXP1200 Memory Hierarchy

Memory Type	Addressable Data Unit (bytes)	Relative Access Time	Special Features

Different **memory** types...

- ✓...are organized into different addressable data units (words or longwords)
- ✓...have different access times
- ✓...connected to different busses

Therefore, to achieve optimal performance, programmers must ***understand the organization*** and ***allocate*** items from the ***appropriate type***

Summary

- ✓ The network challenges are many
- ⇨ Challenge: find ways to improve the design and manufacture of complex networking systems

- ✓ Hope (2002 version)
  lies in the programmers and network processors

- ✓ We will use **Intel IXP1200** as an example which offers...
  - ➢ ...embedded processor plus parallel packet processors
  - ➢ ...connections to external memories and buses

- ✓ Next time: how to start programming these monsters