



# Design Issues for High- Performance Active Routers

---

T. Wolf and J. S. Turner, WU  
IEEE JSAC Mar. 2001

K. C. Shih, 2001/11/1



# Outline

---

- Introduction
- System Organization
- Router Port Design
- Scaling Issues
- Benchmarking
- Commercial Products
- Conclusion



# Introduction

---

- Active and Programmable Networks
  - Customizing routing, open signaling, fully programmable control plane, support new protocols without changes in the underlying hardware
  - Competition among existing and future ISPs may hinge on the speed at which one service provider can respond to new market demands over another



# The problems are ...

---

- Inherently lower performance of software processing compared to hardwired logic
- Active networks implement more complex services than just plain forwarding
- Terabit routers with 10 Gbps links are commercially available now. Can active routers keep pace with the rapid growth?
- Moore's Law is losing
  - Transmission bandwidth has been growing faster than processing bandwidth



# For example,

---

- A single processor

	500 MIPS processor	1 GIPS processor
150Mbps link	25	50
1.5 Gbps link	12.5	<b>25</b> Instr./byte

- JPEG: **81** instr./byte

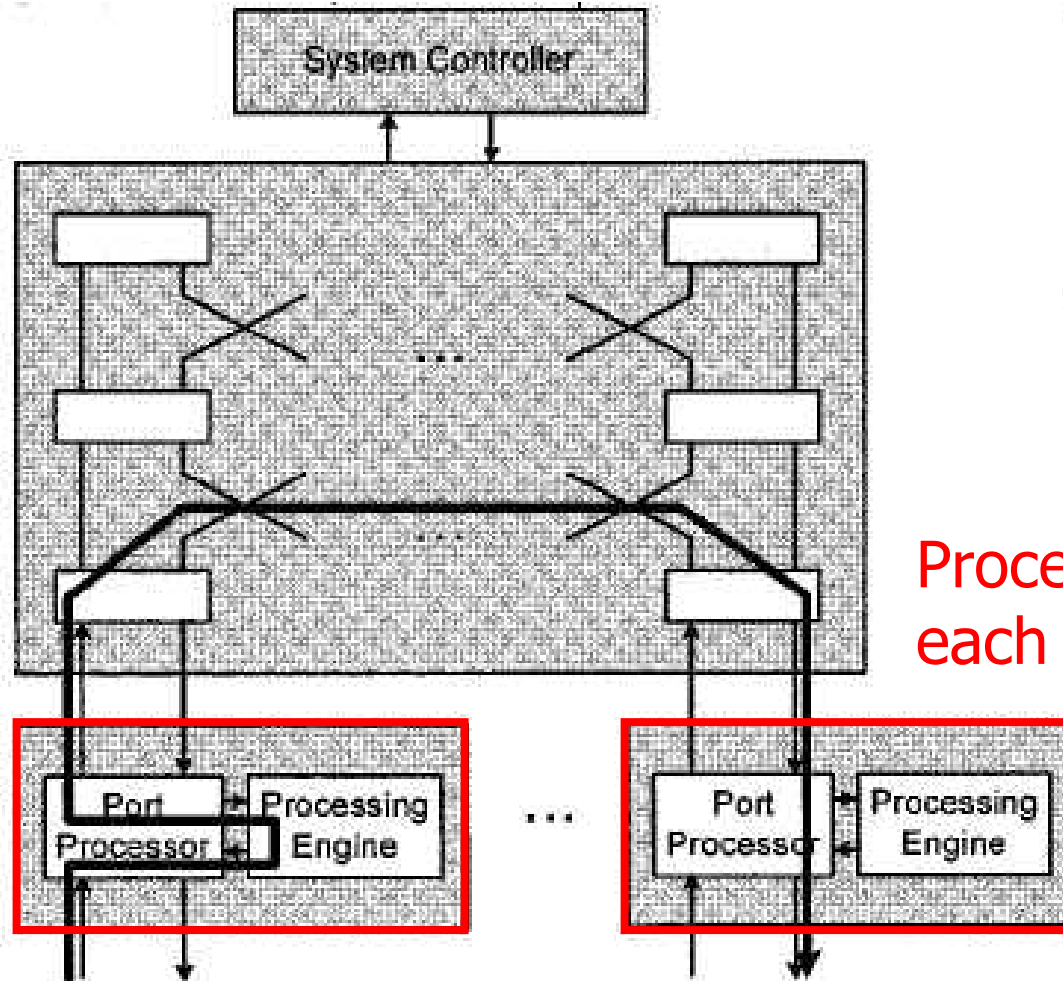


# Focus on

---

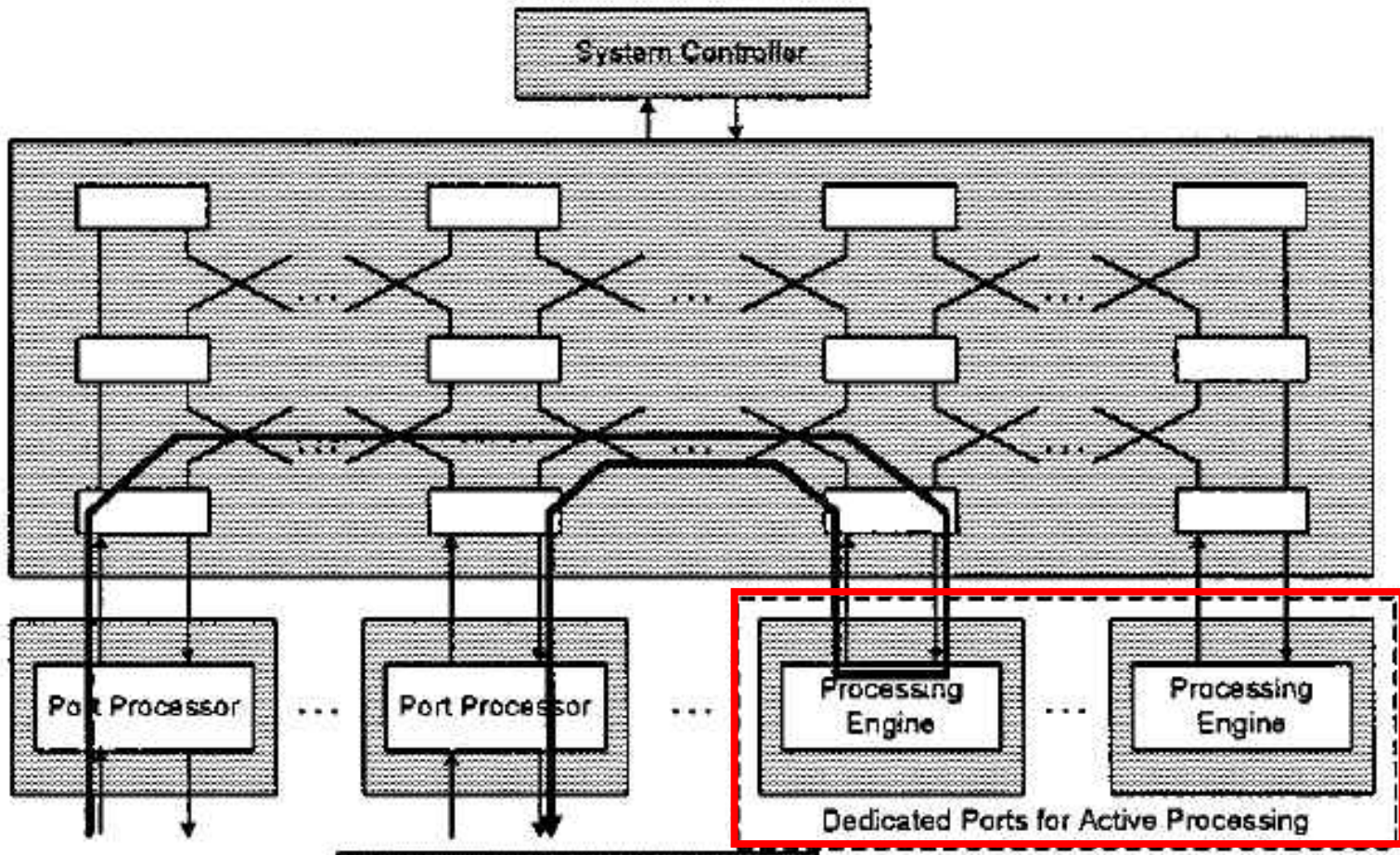
1. Multiple network processors on a single application specific integrated circuit (ASIC)
2. Design a scalable hardware for processing packets at wire speeds of 2.4 Gbps and higher
3. Benchmarking

# System Organization - 1



Processing engine at each router port

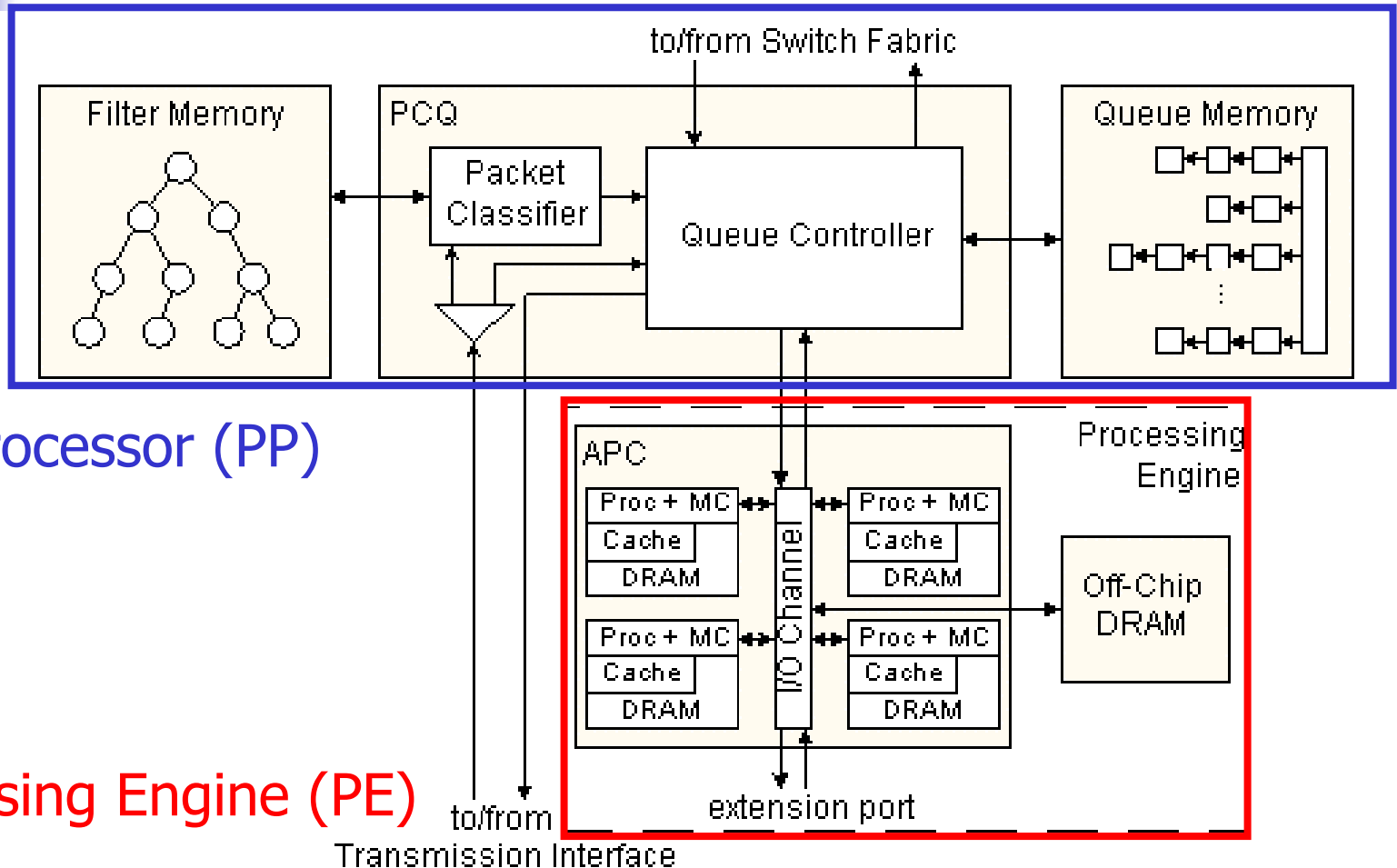
# System Organization - 2



Shared pool of processing engines



# Router Port Design



Port Processor (PP)

Processing Engine (PE)



# Port Processor

---

- Packet classification and queuing (PCQ) chip
- The headers are passed to the packet classifier and the entire packet is passed to the queue controller (QCTL)
- The queues can be rate-controlled to provide QoS guarantees

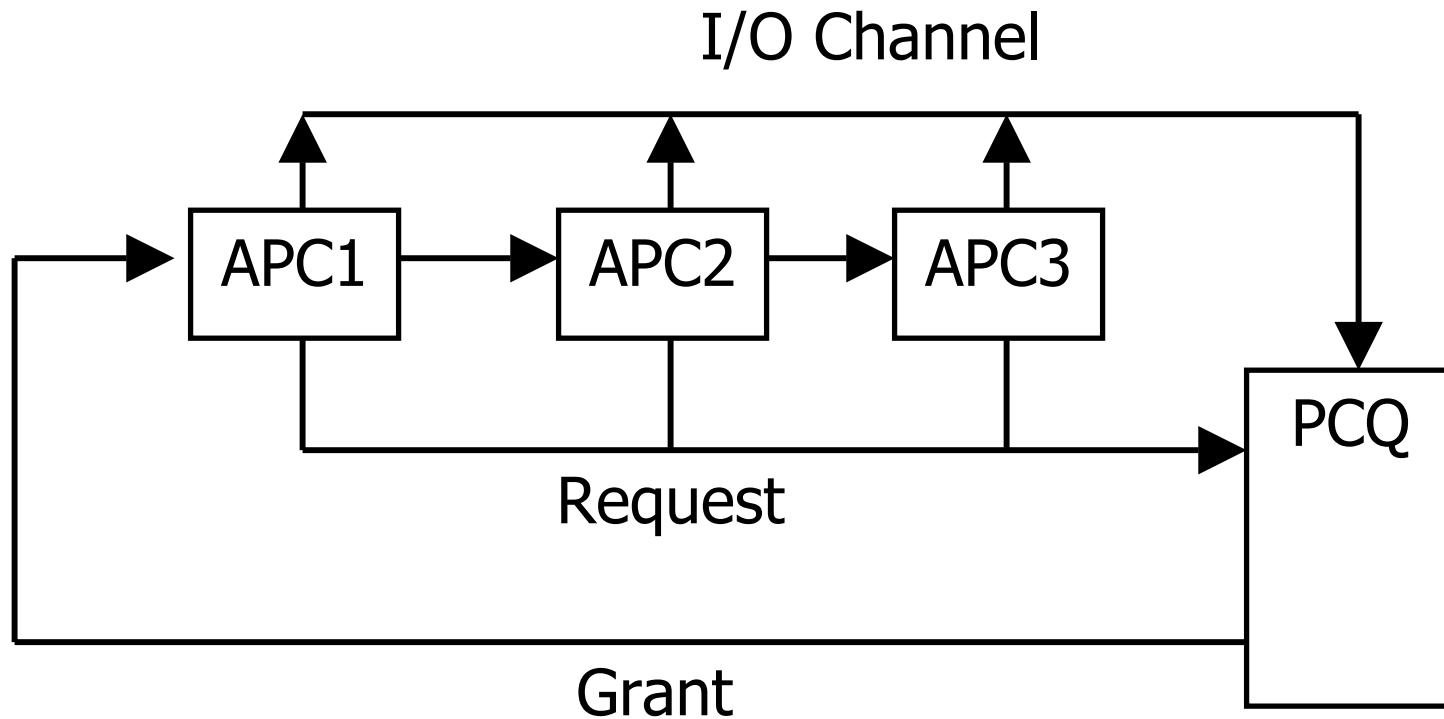


# Processing Engine

---

- Active processing is provided by one or more *active processor chips* (APC)
- Multiple APCs are arranged in a **daisy-chain** configuration to eliminate the need for multiple interfaces to the PCQ.

# Multiple APCs



\* Support priority naturally

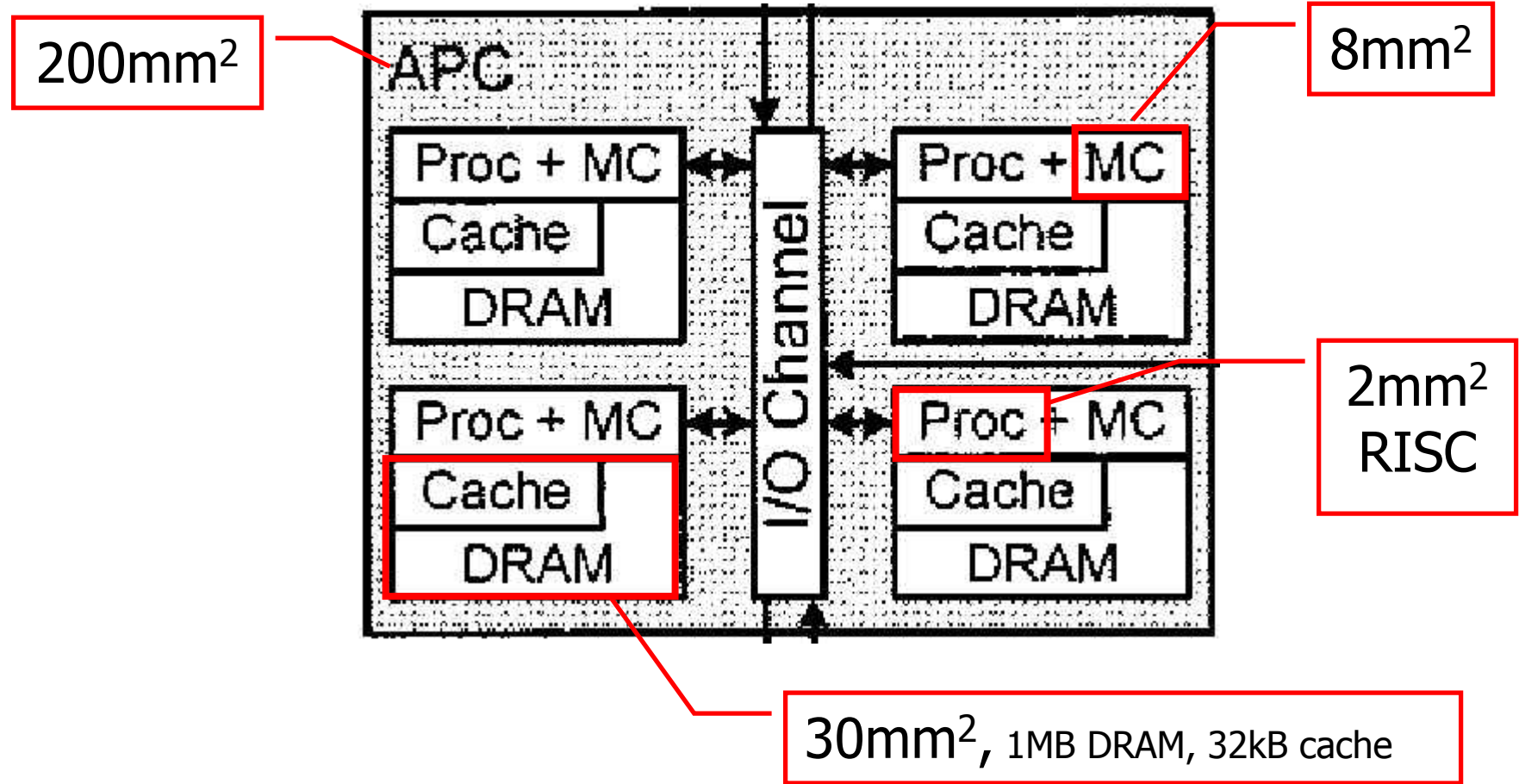


# Design of APCs

---

- Using 0.25-um (deep submicron) CMOS technology

# Design of APCs





# I/O Channel Bandwidth

---

- 32-bit interface in each direction at a clock rate of 250 MHz
  - 8 Gbps
    - packet fragmentation effects
    - bus (I/O channel) contention
    - > 2.4 Gbps (links speed)



## More other consideration...

---

- High-performance memory interface
- Size of embedded memory
  - Operating system kernel
  - Active application codes
  - Per flow state (few hundred flows)
  - Packets currently being processed
- Size of Queue Memory



# Scaling Issues (1/2)

- Technology scaling

## APC TECHNOLOGY SCALING

Year	1999		2002		2005		2008
Feature size ( $\mu\text{m}$ )	0.25	0.18	0.12	0.09	0.06	0.045	0.03
No. of APUs	4	4	8	8	16	16	32
Cache size (kB)	32	64	64	128	128	256	256
DRAM size (MB)	1	2	2	4	4	8	8
Proc + MC area ( $\text{mm}^2$ )	10	5.2	2.3	1.3	0.6	0.3	0.14
SRAM area per MB	175	90	40	23	10	5.7	2.5
DRAM area per MB	25	13	5.8	3.2	1.4	0.8	0.4
Total APU area ( $\text{mm}^2$ )	162	148	131	137	122	132	117
Processor clock frequency (MHz)	400	556	833	1,111	1,667	2,222	3,333
External memory bandwidth (MB/s)	500	694	2,083	2,778	8,333	11,111	33,333
Instructions per byte for 2.4 Gb/s link	5.3	7.4	22	30	89	119	356



## Scaling Issues (2/2)

---

- Multiple APCs
  - Each interface that connects to another APCs acts as a gateway and routes data to other APCs



# Benchmarking

## SIZE AND COMPUTATIONAL COMPLEXITY OF BENCHMARK APPLICATIONS

Application	Description	Object Code (bytes)	Executed Code (bytes)	Complexity (instr. / byte)
RTR	routing table lookup	16,000	15,220	2.1
DRR	packet scheduling	2,500	5,412	4.1
FRAG	packet fragmentation	2,400	5,032	7.7
TCP	traffic monitoring	352,000	29,028	10
JPEG	image compression	260,000	24,620	81
CAST	data encryption	19,500	10,116	104
ZIP	data compression	117,000	14,152	226
REED	forward error correction	6,900	6,040	603



# Suggestions

---

- The large difference of code size (6 –30 kB) suggests that it may be necessary to specialize the different APUs
- For example, CAST, a single APC will be able to encrypt all the data on a 2.4 Gbps link only sometime after 2005 (16 APUs)

# Commercial Products

- Programmable packet processing engines for routers



## Intel IPX1200

- Six processing microengines (6 APUs), one control processor, 200-MHz clock rate, 2.6-Gbps line speed, 6.26-Gbps I/O bus, four threads per processor



# Conclusion

---

- Active networking is an important new direction in networking research and potentially for commercial networks
- Active routers should keep pace with the link speed
- The paper proposed a fundamental design of active routers and related design issues.