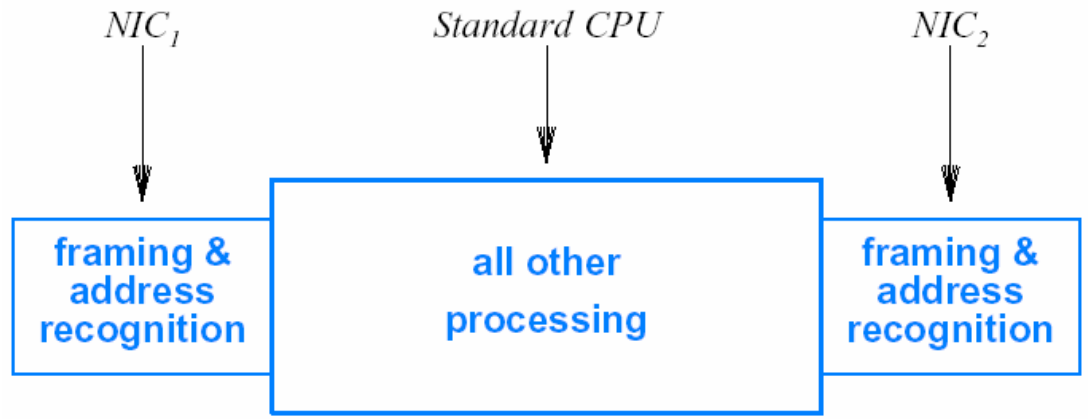
ECE 697J – Advanced Topics in Computer Networks

Hardware-Based Router Architectures
9/30/03

Traditional Software Router

- System architecture so far:



- Problem: single CPU bottleneck
  - Limited aggregated data rate

Aggregate Data Rate

- How is aggregate data rate defined?

  "The aggregate data rate is defined to be the sum of rates at which traffic enters or leaves a system. The maximum aggregate data rate of a system is important because it limits the type and number of network connections the system can handle"

- Examples?
  - Workstation is limited by PCI bus
  - Firewall by processing with limited processing capability
- Is this really the most important metric?
  - Aggregate Packet Rate just as important

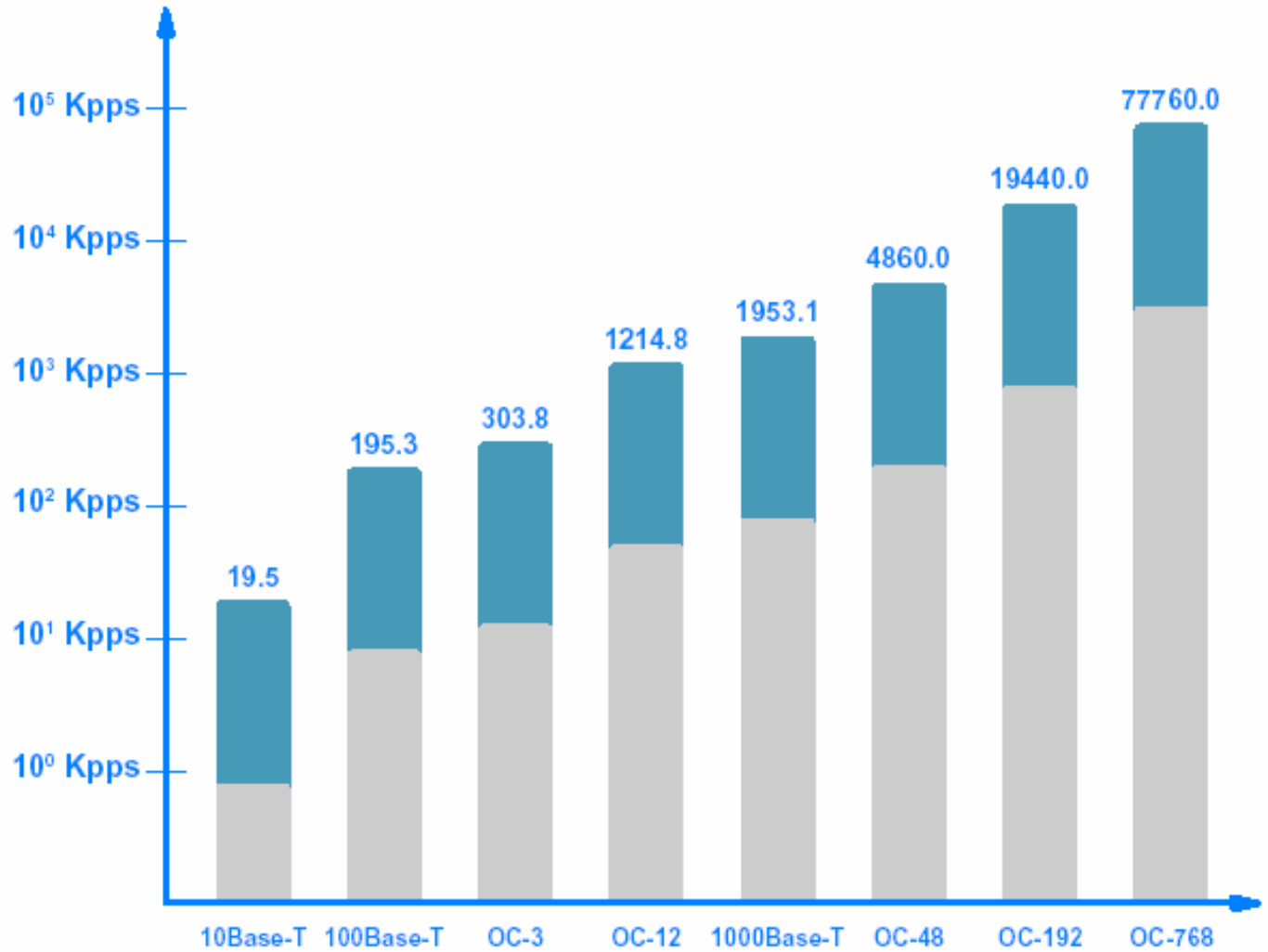
Aggregate Packet Rate

"For protocol processing tasks that have a fixed cost per packet, the number of packets processed is more important than the data aggregate rate"

- Packet rate depends on packet size:

Technology	Network Data Rate In Gbps	Packet Rate For Small Packets In Kpps	Packet Rate For Large Packets In Kpps
10Base-T	0.010	19.5	0.8
100Base-T	0.100	195.3	8.2
OC-3	0.156	303.8	12.8
OC-12	0.622	1,214.8	51.2
1000Base-T	1.000	1,953.1	82.3
OC-48	2.488	4,860.0	204.9
OC-192	9.953	19,440.0	819.6
OC-768	39.813	77,760.0	3,278.4

Aggregate Packet Rate



Time Per Packet

- Aggregate packet rate determines time per packet (on single CPU)

Technology	Time per packet for small packets (in µs)	Time per packet for large packets (in µs)
10Base-T	51.20	1,214.40
100Base-T	5.12	121.44
OC-3	3.29	78.09
OC-12	0.82	19.52
1000Base-T	0.51	12.14
OC-48	0.21	4.88
OC-192	0.05	1.22
OC-768	0.01	0.31

- Packet processing requires in the order of 100s of instructions per packet
- Single CPU router lacks **scalability**

Scalability

- What does scalability mean?
  - A system (design) is scalable if it can easily be extended in "size" and performance
  - E.g., router design can be used with more ports or faster links
- This is an extremely important criteria
  - Designing a system is expensive
  - Being able to extend to new requirements easily is important
  - Performance requirement increase really fast (related to Moore's Law)
  - Systems will eventually be used in a different context than they were designed
- How can we make a network system (router) more scalable?

Advanced Router Architectures

- Changing requirements:
  - Increasing link speed
  - Increasing number of ports
  - Increasing routing tables
  - Increasing processing complexity
- Scalable system design:
  - Exploit parallelism wherever possible
    - Per-port, per-flow, per-packet, instruction-level
  - One Processing engine per port (instead of single CPU)
  - Multiple processors per port
  - “Better” processors
- Note: Comer focuses mostly on processing on a single port

Processing Power

- Overcoming processing bottlenecks:
  - Specialized hardware (ASICs)
  - Fine-grained parallelism
  - Symmetric coarse-grain parallelism
  - Asymmetric coarse-grain parallelism
  - Special-purpose coprocessors
- Other improvements:
  - NICs with onboard processing
  - Smart NICs
  - => basically same as per-port processing engines

Parallelism in Processors

- Fine-grained parallelism
  - Exploits instruction-level parallelism
  - Examples: VLIW, SMT, etc.
  - Limited due to workload
- Symmetric coarse-grain parallelism
  - Multiple parallel identical CPUs
  - Inter-processor communication can limit performance
- Asymmetric coarse-grain parallelism
  - Multiple parallel different CPUs
  - E.g., one processor for layer 2, one for layer 3
- Special-purpose coprocessors
  - Custom logic for lookups, checksums, etc.
  - High-performance but not (fully) programmable
- Key question becomes how such a system can be programmed

Next Class

- We can achieve high processing power on port
- How can we interconnect ports in a scalable way?
- Read Chapter 10 (Switching Fabrics)