Techniques for Fast Packet Buffers

Sundar Iyer, Nick McKeown
(sundaes,nickm)@stanford.edu
Departments of Electrical Engineering & Computer Science, Stanford University
http://klamath.stanford.edu
Goal
Determine and analyze techniques for building high speed (>40Gb/s) electronic packet buffers.

Example
OC768c = 40Gb/s; RTT*BW = 10Gb; 64 byte packets

Write Rate, \( R \)
1 packet every 12.8 ns

Read Rate, \( R \)
1 packet every 12.8 ns

Use SRAM?
+ fast enough random access time, but
- too expensive, and
- too low density to store 10Gb of data.

Use DRAM?
+ high density means we can store data, but
- too slow (typically 50ns random access time)
Memory Hierarchy

Large DRAM memory holds the middle of FIFOs

Arriving Packets

Small ingress SRAM cache of FIFO tails

Writing b cells

Arbiter or Scheduler

Requests

Departing Packets

Small ingress SRAM cache of FIFO heads

Reading b cells

Stanford University
Questions

How large does the SRAM cache need to be:

1. To guarantee that a packet is immediately available in SRAM when requested, or
2. To guarantee that a packet is available within a maximum bounded time?

What Memory Management Algorithm (MMA) should we use?
Earliest Critical Queue First (ECQF-MMA)

1. **In SRAM**: A Dynamic Buffer of size $Q(b-1)$

2. **Lookahead**: Arbiter requests for future $Q(b-1) + 1$ slots are known.

3. **Compute**: Find out the queue which will run into “trouble” earliest.

4. **Replenish**: “$b$” cells for the “troubled” queue.
Example of ECQF-MMA: $Q=4, b=4$

$t = 0; \text{Green Critical}$

$t = 1$

$t = 2$

$t = 3$

$t = 4; \text{Blue Critical}$

$t = 5$

$t = 6$

$t = 7$

$t = 8; \text{Red Critical}$
Results
Single Address Bus

1. **Patient Arbiter**: ECQF-MMA (earliest critical queue first), minimizes the size of SRAM buffer to Q(b-1) cells; and guarantees that requested cells are dispatched within Q(b-1)+1 cell slots.

2. **Impatient Arbiter**: MDQF-MMA (maximum deficit queue first), with a SRAM buffer of size Qb[2 +ln Q] guarantees zero latency.
Implementation Numbers
(64byte cells, b = 8, DRAM T = 50ns)

1. VOQ Switch - 32 ports
   - Brute Force: Egress. SRAM = 10 Gb, no DRAM
   - Patient Arbiter: Egress. SRAM = 115 kb, Lat. = 2.9 us, DRAM = 10Gb
   - Impatient Arbiter: Egress. SRAM = 787 kb, DRAM = 10Gb
   - Patient Arbiter(MA): No SRAM, Lat. = 3.2us, DRAM = 10Gb

2. VOQ Switch - 32 ports, 16 Diffserv classes
   - Brute Force: Egress. SRAM = 10Gb, no DRAM
   - Patient Arbiter: Egress. SRAM = 1.85Mb, Lat. = 45.9us, DRAM = 10Gb
   - Impatient Arbiter: Egress. SRAM = 18.9Mb, DRAM = 10Gb
   - Patient Arbiter(MA): No SRAM, Lat. = 51.2us, DRAM = 10Gb