

Packet Scheduling

Fundamentals of Computer Networks
Guevara Noubir

F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 1

Outline

- Max-min fairness
- Generalized processor sharing
- Weighted round-robin
- weighted fair queuing

F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 2

First-Come First-Serve Scheduling

- Example:
 - Two streams of 20Kbps and 2Mbps
 - One shared link of 100Kbps < 20Kbps + 2Mbps
 - In a FCFS system: stream 1 gets only 1% of the shared link!
- Unfairness of FCFS:
 - Each stream gets: $requested_share / Total_requested_BW$

F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 3

Max-Min Fairness

- Max-Min fairness:
 - Order the connections by increasing demand
 - Starts by equally dividing the total BW over all connections
 - Allocate to each connection $\text{Min}\{rqst_BW, one\ share\}$
 - If a connection doesn't use all its share divide the remaining BW between the remaining connections
- Example:
 - $100\text{Kbps}/2 = 50\text{Kbps}$.
 - Allocate to connection 1: 20 Kbps = $\text{Min}\{20, 50\}$
 - Allocate to connection 2: 80Kbps = $\text{Min}\{30+50, 2000\}$

F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 4

Max-Min Fairness

- Example:
 - 4 connections with demands: 1, 2, 5, 3. Total capacity: 8

F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 5

Max-Min Fairness with Weights

- Example:
 - 4 connections with demands: 1, 2, 3, 5. Total capacity: 8
 - Connections have weights: 0.5, 2, 1, 0.5
 - The weighted shares per connection are: 1, 4, 2, 1
 - Allocation:
 - Connection 1 is satisfied,
 - Connection 2 is satisfied and 2 remaining units to be shared by the other connections
 - Connection 3 gets: $3 = \text{Min}\{3, 2 + 2*1/(0.5+1)\} = 3$
 - Connection 4 gets: $1 + 2*0.5/(0.5+1) + 1/3 = 2$

F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 6

Fair Scheduling Disciplines

- Generalized Processor Sharing: an ideal realization of max-min fairness
 - Packets are broken into infinitesimal units
- Weighted round robin
- Weighted Fair Queuing

F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 7

Generalized Processor Sharing

- The scheduler divides the resource equally in proportion to weights. The overhead of scheduling/switching is assumed to be zero.
- Example: n connections, weights = 1
 - Scheduler divides the resource into n shares, if a connection doesn't have packets, then this results in $1/(n-1)$ shares for remaining connections

F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 8

Weighted Round Robin

- If the connections have different weights, the scheduler divides the resource into weighted shares. WRR works at the packet level
- Example:
 - 3 connections, with weights 1, 2, 10
 - The scheduler allocates:
 - 1/13 of the time to connection 1 => transmits one packet from 1
 - 2/13 of the time to connection 2 => transmits 2 packets from 2
 - 10/13 of the time to connection 3 => transmits 10 packets from 3
- If the packets have variable packet size => divide the connection weight by the average packet size
- Problems: unknown packets distribution, local unfairness

F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 9

Weighted Fair Queuing

- Emulates GPS (on the side): schedules packets in the order they finish completion according to GPS
- Computes a *finish number* for each packet which is the round in which the packet completes service (GPS)
- A *round* finishes when GPS has served one unit from each nonempty queue (=> variable duration rounds)
- Assumption (to be released later): all connection have the same weight

F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 10

Weighted Fair Queuing (Cont'd)

- $R(t)$: round number at time t (real number)
- $F(i, k, t)$: finish number for k^{th} packet of connection i , arriving at time t
- $P(i, k)$: size of k^{th} packet of connection i
- $F(i, k+1, t) = \text{Max}\{F(i, k, t), R(t)\} + P(i, k+1)$
- Packets are scheduled according to an increasing finish number
- $R(t)$ increases at a rate which is the inverse of the number of active connections (GPS)

F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 11

Weighted Fair Queuing (Cont'd)

- Example:
 - Three connections A, B, C , of equal weight
 - $t = 0$: packets arrive for the connections with length 1, 2, 2
 - $t = 4$: 1 packet of size 2 arrives for connection A
- GPS:
 - From $t=1$ to 3: increase rate 1/3 (three active connections)
 - From $t=3$ to 4: increase rate 1/2 (two active connections)
 - From $t=4$ to 5.5: increase rate 1/3 (three active connections)
 - From $t=5.5$ to 7: increase rate 1 (one active connection)
- WFQ:
 - $F(i, 1, 0) = 1, 2, 2$
 - $F(2, 1, 4) = 2; F(3, 1, 4) = 2;$
 - $F(1, 2, 4) = \text{Max}\{R(4), F(1, 1, 4)\} + P(2, 2) = 1.5 + 2 = 3.5$

F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 12

Weighted Fair Queuing (Cont'd)

- Packets are not preempted
- Example:
 - Three connections *A, B, C*, of equal weight
 - $t = 0$: packets arrive for the connections with length 1, 4, 5
 - $t = 4$: 1 packet of size 1 arrives for connection *A*

F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 13

Weighted Fair Queuing (Cont'd)

- Generalization to weighted connections:
 - $F(i, k+1, t) = \text{Max}\{F(i, k, t), R(t)\} + P(i, k)/w(i)$
 - Round number increases at rate equal to the inverse of the sum of the weights

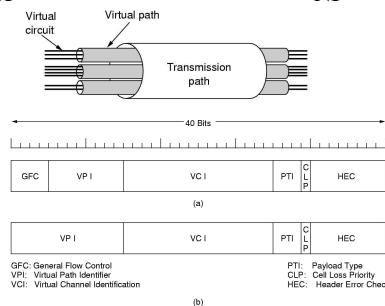
F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 14

Asynchronous Transfer Mode

- Architecture: connection oriented packet switching tech.
 - nodes and switches: ATM Forum ITU-UNI and ITU-NNI
- Virtual Circuits/Paths
 - Q.2931
- Physical Layer
 - Sonet, SDH, Wireless
- Fixed size packets
 - 5 bytes header +
 - 48 bytes data



F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 15

Variable vs Fixed-Length Packets

- No Optimal Length
 - if small: high header-to-data overhead
 - if large: low utilization for small messages
- Fixed-Length Easier to Switch in Hardware
 - simpler
 - enables parallelism

F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 16

Big vs Small Packets

- Small Improves Queue behavior
 - finer-grained pre-emption point for scheduling link
 - maximum packet = 4KB
 - link speed = 100Mbps
 - transmission time = $4096 \times 8/100 = 327.68\mu\text{s}$
 - high priority packet may sit in the queue 327.68us
 - in contrast, $53 \times 8/100 = 4.24\mu\text{s}$ for ATM
 - near cut-through behavior
 - two 4KB packets arrive at same time
 - link idle for 327.68us while both arrive
 - at end of 327.68us, still have 8KB to transmit
 - in contrast, can transmit first cell after 4.24us
 - at end of 327.68us, just over 4KB left in queue

F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 17

Big vs Small (cont)

- Small Improves Latency (for voice)
 - voice digitally encoded at 64Kbps (8-bit samples at 8KHz)
 - need full cell's worth of samples before sending cell
 - example: 1000-byte cells implies 125ms per cell (too long)
 - smaller latency implies no need for echo cancellers
- ATM Compromise: 48 bytes = $(32+64)/2$

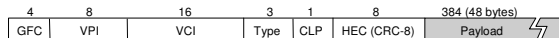
F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 18

Cell Format

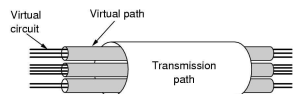
- User-Network Interface (UNI)



- host-to-switch format
- GFC: Generic Flow Control (still being defined)
- VCI: Virtual Circuit Identifier
- VPI: Virtual Path Identifier
- Type: management, congestion control, AAL5 (later)
- CLPL: Cell Loss Priority
- HEC: Header Error Check (CRC-8)

- Network-Network Interface (NNI)

- switch-to-switch format
- GFC becomes part of VPI field



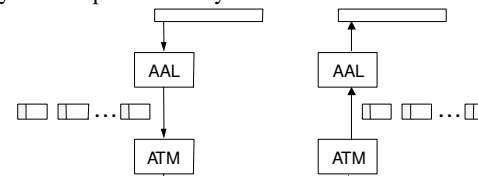
F2003, CSG150

Fundamentals of Computer Networks

Lecture 6, 19

Segmentation and Reassembly

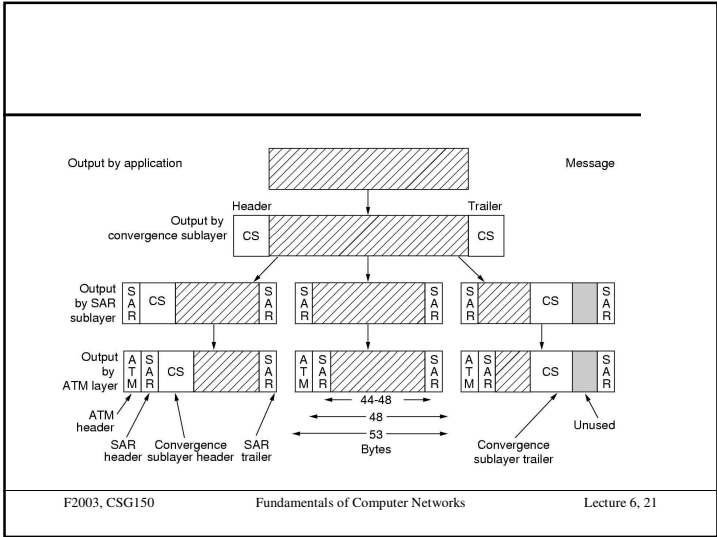
- ATM Adaptation Layer (AAL)
 - AAL 1 (CBR) and 2 (VBR) designed for applications that need guaranteed rate (e.g., voice, video)
 - AAL 3/4 designed for packet data
 - AAL 5 is an alternative standard for packet data. Designed by the computer industry. Most used interface to ATM.



F2003, CSG150

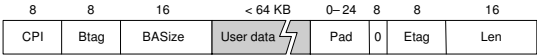
Fundamentals of Computer Networks

Lecture 6, 20



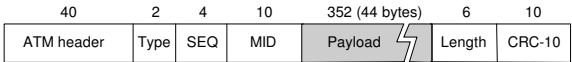
AAL 3/4

- Convergence Sublayer Protocol Data Unit (CS-PDU)



- CPI: common part indicator (version field: currently 0)
- Btag/Etag: beginning and ending tag
- BASize: hint on amount of buffer space to allocate
- Length: size of whole PDU

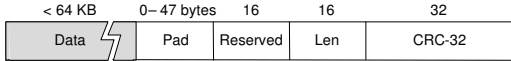
Cell Format



- Type
 - BOM: beginning of message
 - COM: continuation of message
 - EOM: end of message
- SEQ: sequence of number
- MID: message id
- Length: number of bytes of PDU in this cell

AAL5

- CS-PDU Format



- pad so trailer always falls at end of ATM cell
- Length: size of PDU (data only)
- CRC-32 (detects missing or misordered cells)
- Cell Format
 - end-of-PDU bit in Type field of ATM header