

# EE549: Problem Set #5

## Due: Wednesday March 12, 2008

### I. RINGS AND PRIORITIES

Consider the 3 queue network shown below. There are three input streams  $X_1(t)$ ,  $X_2(t)$ ,  $X_3(t)$ , and each has a well defined rate given by  $r_1, r_2, r_3$ , respectively (in units of kbits/sec). Assume a fluid flow departure model with preemptive priority as follows:  $Priority(1) > Priority(2) > Priority(3)$ , so that data from stream  $X_1(t)$  preempts all other data, and data from  $X_2(t)$  preempts  $X_3(t)$  data. Service is FIFO amongst the same priority class. We assume that this same priority scheme is upheld at all three network nodes.

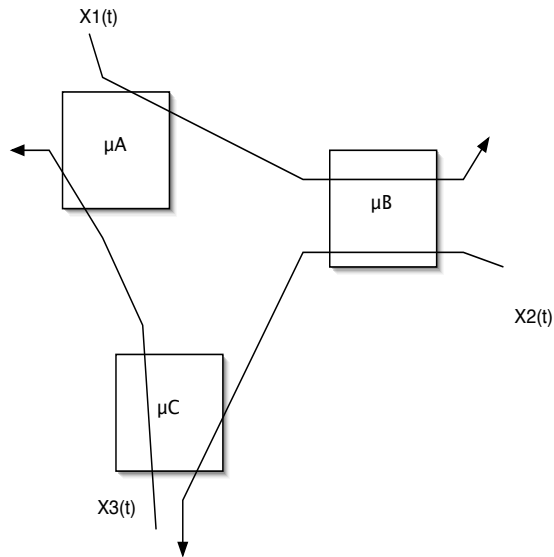


Fig. 1. This figure is not artistically perfect. It simply illustrates which sessions use which servers (as described below). The “direction” that each arrow crosses through each box has nothing to do with the problem.

Data from  $X_1(t)$  passes through the servers at nodes  $A$  and  $B$ . Data from  $X_2(t)$  passes through the servers at nodes  $B$  and  $C$ . Data from  $X_3(t)$  passes through the servers at nodes  $C$  and  $A$ . All servers have constant rates  $\mu_A, \mu_B, \mu_C$ , and we assume that:

$$r_1 + r_3 \leq \mu_A$$

$$r_1 + r_2 \leq \mu_B$$

$$r_2 + r_3 \leq \mu_C$$

- a) Prove that the system is rate stable.
- b) Write a paragraph (at least two or three sentences) about why we cannot conclude the system is rate stable if we do not have the priority service, such as if all service was FIFO (see also ungraded exercise VIII for intuition).
- c) Suppose now that each input  $X_i(t)$  is leaky bucket constrained with parameters  $(r_i, \sigma_i)$  for  $i \in \{1, 2, 3\}$ , where  $r_i$  has units of kbits/sec and  $\sigma_i$  has units of kbits. What are the leaky bucket parameters for the departure process  $Y_1^{(A)}(t)$ ? (i.e., the process of  $X_1$  bits departing from node  $A$ ). What about  $Y_2^{(B)}(t)$  and  $Y_3^{(C)}(t)$ ?
- d) Give worst case end-to-end network delay bounds for traffic from each of the individual streams  $X_1$ ,  $X_2$ , and  $X_3$ , assuming the leaky bucket model of part (c).

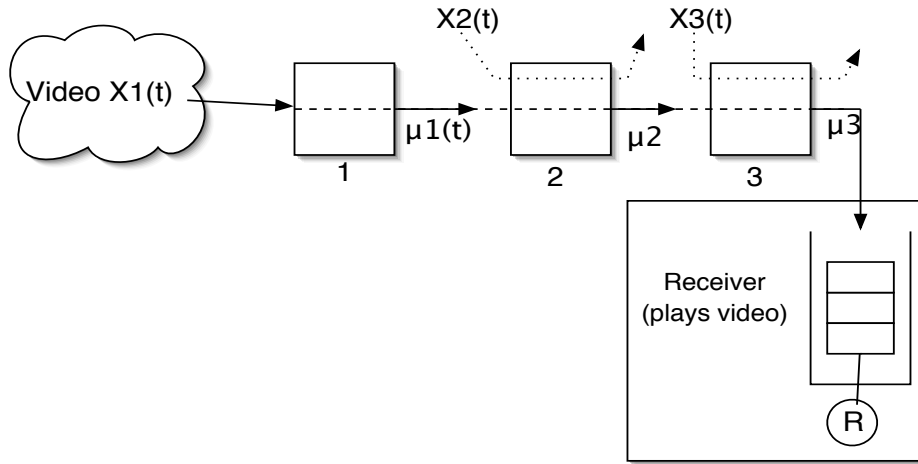


Fig. 2. Video over a network.

## II. REAL TIME VIDEO

A video input stream  $X_1(t)$  produces packets that travel through a network consisting of the following: A single wireless node with time-varying processing rate  $\mu_1(t)$ , a node of constant processing rate  $\mu_2$ , and a node of constant processing rate  $\mu_3$  (see Figure 2). Assume packetized flow, so that a packet does not depart until its last bit departs.

Packets from all input streams have fixed size  $B$  kilobits.

- The video stream  $X_1(t)$  has periodic packet arrivals every  $T$  seconds.
- $\mu_1(t)$  has a time-varying service rate with mean and lag parameters  $(\bar{\mu}, \gamma)$  (where  $B/T \leq \bar{\mu}$ ).
- In node 2, packets from  $X_1(t)$  have *non-preemptive priority* over  $X_2(t)$  packets, so that they are placed first in the buffer but cannot interrupt a low priority packet that is already in the server (assume  $B/T \leq \mu_2$ ).
- Input  $X_3(t)$  is leaky bucket constrained with parameters  $(r, \sigma)$  (assume  $B/T + r \leq \mu_3$ ). All packets from the  $X_1(t)$  and  $X_3(t)$  stream are treated equally in node 3, and are served in FIFO order.

The receiver wants to continuously watch the video from the  $X_1(t)$  stream by processing these packets at a rate of 1 video packet every  $T$  seconds. Thus, it keeps a receive buffer to store up packets that arrive. The receiver cannot process a packet for viewing until it has fully arrived to the receive buffer. Because of the desired continuous operation, the receiver does not start watching the video until the receive buffer fills up with exactly  $K$  packets, for some value of  $K$ .

a) Suppose that  $X_2(t) = 0$  (so that no packets ever arrive from  $X_2(t)$ ). Compute a value of  $K$  that ensures continuous operation with no stalls or outages during the video. (Recall that Fact 3c applies only to constant server rates, and does not apply to queues with time-varying service  $\mu(t)$ . Thus, you will need to add  $B$  to the the packet-based leaky bucket departure process from node 1).

b) How would the value of  $K$  change if the output signal from the first node requires  $D$  seconds to propagate to the second node?

c) How would the value of  $K$  change if  $X_2(t)$  is non-zero?

d) Suppose that  $\mu_3$  is large so that  $B/T + 2r < \mu_3$  (where  $r$  is the leaky bucket parameter of  $X_3(t)$ ). Suppose that  $r$  is doubled from  $r$  to  $2r$ , but the burst size  $\sigma$  is kept the same. How does this impact your  $K$  value?

## III. DYNAMIC POWER ALLOCATION WITH TWO QUEUES

Consider a two-queue system with initial backlog  $U_1(0) = U_2(0) = 11$  kbits, and rate-power curves  $\phi_1(p) = \log_2(1 + p)$ ,  $\phi_2(p) = 2 \log_2(1 + p)$  (in units of kbits/sec). Suppose that  $P_{max} = 3$ , so that  $P_1(t) + P_2(t) \leq 3$  for all  $t$ . Consider the following strategy:  $P_1(t) = 3$  for  $0 \leq t \leq 2$ ,  $P_2(t) = 3$  for  $2 < t \leq 4$ , and  $P_1(t) = P_2(t) = 1.5$  for  $t > 4$ .

a) Compute the emptying time  $T_{empty}$  under this particular strategy, and compute  $\bar{P}_1$ ,  $\bar{P}_2$ , the average powers expended by links 1 and 2 over the duration  $0 \leq t \leq T_{empty}$ .

b) Observe that, indeed,  $\bar{P}_1 + \bar{P}_2 \leq 3$ . Show that the alternate strategy of choosing  $P_1(t) = \bar{P}_1$ ,  $P_2(t) = \bar{P}_2$  for all time  $t$  such that  $0 \leq t \leq T_{empty}$  yields a *strictly better* clearance time.

c) Find a (nonlinear) equation for the *optimal* emptying time  $T_{min}$ , and solve this equation (for example, via matlab, or in closed form) to yield the optimal emptying time and the resulting optimal power allocations. Verify that  $T_{min}$  is lower than either of the emptying times in parts (a) or (b).

#### IV. MINIMUM CLEARANCE TIME

Consider a 3-user wireless downlink with identical rate-power functions  $f_1(p) = f_2(p) = f_3(p) = f(p)$ , where  $f(p)$  is given as follows:

$$f(p) = \min[p, 1] \text{ kbits/sec}$$

The total power constraint is  $P_{tot} = 2$  Watts, so that  $p_1(t) + p_2(t) + p_3(t) \leq 2$  Watts for all time  $t$ . Let  $\vec{U} = (U_1, U_2, U_3)$  represent an initial backlog vector in the three queues (in units of kbits). Define  $T_{min}(\vec{U})$  as the minimum time required to clear the backlog vector  $\vec{U}$ .

Throughout this problem, we assume that  $U_1 \leq U_2 \leq U_3$ .

a) Show that if  $U_1 \leq U_2 \leq U_3$ , then  $T_{min}(\vec{U}) \geq \max\{\frac{U_1+U_2+U_3}{2}, U_3\}$ .

b) Compute  $T_{min}(\vec{U})$ . Hint: you might break the problem into the following two cases:

- $U_1 \leq U_2 \leq U_3$  and  $U_1 + U_2 \leq U_3$
- $U_1 \leq U_2 \leq U_3$  and  $U_1 + U_2 > U_3$

#### V. IMET WITH A PARTICULAR $T_{min}(\vec{U})$ FUNCTION

Consider a queueing network (not necessarily a downlink power allocation problem) with two leaky bucket inputs  $X_1(t)$  and  $X_2(t)$  with parameters  $(r_i, \sigma_i)$  (for  $i \in \{1, 2\}$ ). Suppose that  $T_{min}(U_1, U_2) = \max[5, U_1 + 2U_2, 3U_1 + U_2]$ . Suppose that we use the IMET algorithm with constant frames of size  $T$ , where backlog that arrived in the previous frame is cleared (using the minimum clearance time algorithm) in the current frame. Suppose that  $\max[r_1 + 2r_2, 3r_1 + r_2] \leq 1 - \epsilon$  for some value  $\epsilon > 0$ . Compute a frame size  $T$  such that the IMET algorithm ensures stability with no lost data.

#### VI. VARIABLE FRAME IMET

Consider the downlink power allocation problem with the following modified version of IMET that uses variable frame sizes  $\{T_m\}$ , and that never drops any data. Specifically:

- If there is no data in the system at the beginning of a frame, we *sleep* for  $z$  units of time (so that our current frame is of size  $z$ , and the system is idle during this time).
- If there is an amount of data  $\vec{U}$  at the beginning of a frame, clear this data in minimum time (ignoring any new arrivals), so that the current frame size is  $T_{min}(\vec{U})$ .

We thus have variable frame sizes  $T_1, T_2, T_3, \dots$ , where  $T_1 = z$ .

a) Suppose  $X_i(t) \sim (r_i, \sigma_i)$  for all inputs  $i$ . Suppose  $\vec{r} \in \rho\Lambda$  for some value  $\rho$  that satisfies  $0 < \rho < 1$ . Define  $T$  as follows:

$$T \triangleq \max \left[ \frac{T_{min}(\vec{\sigma})}{(1 - \rho)}, z \right]$$

Show that the frames in the above algorithm are never larger than  $T$ , so that the worst case delay is at most  $2T$ . Hint: Assume that  $T_m \leq T$  for a given frame  $m$ , and prove that this implies  $T_{m+1} \leq T$ .

b) Write a list of advantages and disadvantages of this variable frame length IMET algorithm (particularly comparing it to fixed-frame IMET).

VII. LITTLE’S THEOREM

Consider a queueing network of three queues. The external input rate to the first queue is  $\lambda_1$  packets/second. The packets departing the first queue leave the system with probability 1/2, and join the second queue with probability 1/2. The second queue has an external arrival stream of  $\lambda_2$  packets/second. Packets departing the second queue go to the third queue with probability 1/2, and go back to the first queue with probability 1/2. The third queue has no external arrivals, and all packets leaving the third queue also leave the system.

a) The average total number of packets in the three queues is measured and found to be  $\bar{L}_1 = 2.2, \bar{L}_2 = 4.6, \bar{L}_3 = 12.0$ . What is the average delay of a packet? (where delay considers the time the packet spends in the queueing system, averaged over packets from both stream 1 and stream 2).

b) For the same system, we do a more careful measurement and find that, considering only packets from stream 1, we have:  $\bar{L}_1^1 = 2.0, \bar{L}_2^1 = 2.4, \bar{L}_3^1 = 6$ . What is the average delay in the system of packets from stream 1? What is the average delay of packets from stream 2?

VIII. UNGRADED EXERCISE — PRIORITY SERVICE FOR CYCLIC NETWORKS

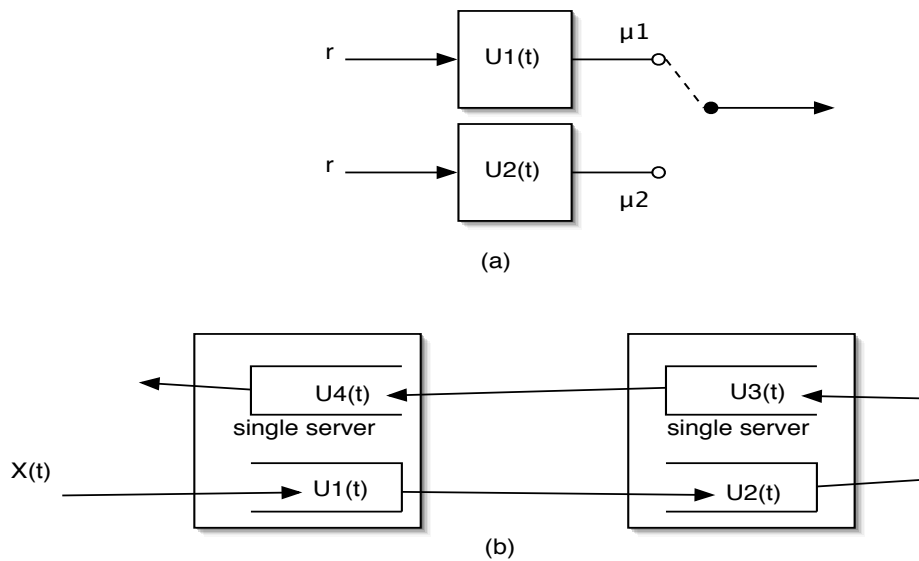


Fig. 3. Example systems.

This problem (problem VIII) will not be graded. You can do it if you like, but do not turn in your work for this problem.

a) Consider a two queue system with a single server operating in *continuous time* (Fig. 3a). If the server is attached to queue 1, then queue 1 transmits at rate  $\mu_1$  (bits/second), while if the server is attached to queue 2, then queue 2 transmits at rate  $\mu_2$  (bits/second). Suppose two separate input streams enter the queues, each with a well defined time average rate  $r$  bits/second. What is the maximum possible value of  $r$  (call it  $r_{max}$ ) for which the system is rate stable? Show that *any* work conserving scheduling policy (where a server is allocated to a non-empty queue whenever there is a non-empty queue) makes the system rate stable provided  $r \leq r_{max}$ .

We now consider a four queue cyclic system shown in Fig. 3b, and demonstrate that rate stability can depend on the type of service, even if service is always work conserving. This example comes from:

P. R. Kumar and T. I. Seidman, “Dynamic Instabilities and Stabilization Methods in Distributed Real-Time Scheduling of Manufacturing Systems,” IEEE Transactions on Automatic Control, 1990.

A single input process  $X(t)$  is injected into the system at queue 1, and this data must travel to queue 2, queue 3, and queue 4, before leaving the system. Let  $U_1(t), U_2(t), U_3(t), U_4(t)$  represent the unfinished work in each queue

at time  $t$ . Let  $\mu_1, \mu_2, \mu_3, \mu_4$  represent the transmission rates of each queue. However, a single server exists to serve both queues 1 and 4 (see figure), and another single server exists to serve both queues 2 and 3. Each queue  $i$  can transmit at rate  $\mu_i$  when a server is allocated to it, and otherwise serves at rate 0. Assume fluid flow.

b) Suppose that the input  $X(t)$  is leaky bucket with parameters  $(r, \sigma)$ . Assume that  $r/\mu_1 + r/\mu_4 < 1$ , and that  $r/\mu_2 + r/\mu_3 < 1$ . Suppose that server scheduling is *priority based*: Queue 1 has preemptive priority over queue 4 (so that queue 1 is always served over queue 4 if there is data in queue 1). Also, queue 2 has preemptive priority over queue 3. Show that the system is rate stable, and compute worst case congestion bounds for  $U_1(t), U_2(t), U_3(t), U_4(t)$  (assuming the system is initially empty).

Suppose now that the server is scheduled as follows: Once a server starts serving a queue, it keeps serving the same queue until that queue empties. If the queue is empty, it continues to serve that queue until another queue that it can serve receives a non-zero amount of data. Consider the case when  $X(t) = rt$ , that is, the input is a constant fluid input with rate  $r$  (note that  $X(t) \sim (r, 0)$ , and this is the only possible fluid type input). Assume that  $U_1(0) = V, U_2(0) = U_3(0) = U_4(0) = 0$ . Assume that:  $r = 1, \mu_1 = 3, \mu_2 = 1.8, \mu_3 = 3, \mu_4 = 2$ .

c) During the interval  $[0, t_1]$ , a server is allocated to queues 1 and 2. Compute the time  $t_1$  at which queue 1 first empties. Compute  $U_2(t_1)$ .

d) During the interval  $[t_1, t_2]$ , the servers are still allocated to queues 1 and 2. Queue 1 is serving at rate  $r$  (not the full rate  $\mu_1$ ), while queue 2 is serving at the full rate  $\mu_2$ . Note that  $U_1(t) = 0$  for  $t_1 \leq t \leq t_2$ . Compute  $t_2$ , the time at which queue 2 becomes empty.

e) Compute  $U_3(t_2)$ . During the interval  $[t_3, t_4]$ , the servers are allocated to queues 3 and 4, as these queues start to have non-zero amount of data. Compute the time  $t_4$  at which queue 4 becomes empty.

f) Let  $T_1 = t_4$ . What is the value of  $U_1(T_1)$ ? Note that  $U_2(T_1) = U_3(T_1) = U_4(T_1) = 0$ , and hence the system starts out at time  $T_1$  with a similar state as at time 0, and the same schedules will be repeated (with the exception that time is scaled according to the initial backlog  $U_1(T_1)$  rather than  $V$ ). Compute  $T_2, T_3, \dots, T_n, \dots$ , the times at which this cycle is again repeated.

g) Show that  $U_1(T_n) \rightarrow \infty$  as  $n \rightarrow \infty$ . Show that  $U_1(T_n)/T_n$  does not converge to zero as  $N \rightarrow \infty$ , and so the system is not rate stable.