Slide supporting material

Lesson 11: Solved M/G/1 Exercises

Giovanni Giambene

Queuing Theory and Telecommunications: Networks and Applications
2nd edition, Springer

Exercise #1

- We consider an **ATM multiplexer** receiving 2 synchronous input time-division traffic flows that have different priorities:
  - Each slot of the high-priority line carries an ATM cell with probability p;
  - Each slot of the low-priority line carries one message with probability q; each message is composed of a random number of cells according to the PGF L(z). The packet arrival process on the low-priority line is **compound Bernoulli.**
- The ATM multiplexer stores the cells before transmission in a buffer of infinite capacity.
- The output line is **synchronous** with the input lines: input and output slot durations are equal; each output slot is used to convey one input cell.
- We have to study **the mean delay experienced by the cells of the low-priority line** due to the presence of the cells served of the high-priority line.

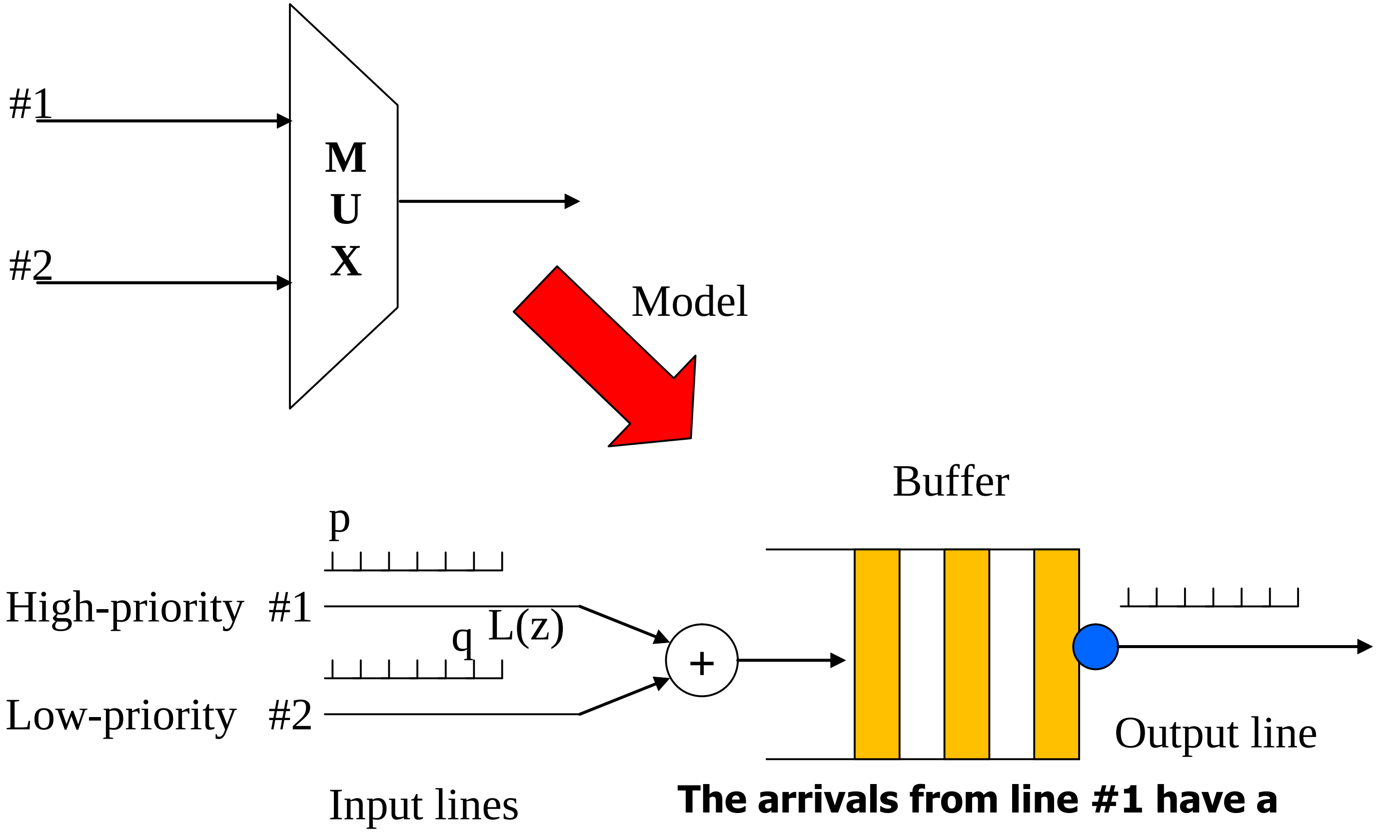
Exercis

This exercise could also be applied to any time-division transmission (e.g., downlink transmissions of wireless systems).

- We consider an **ATM multiplexer** receiving 2 synchronous input time-division traffic flows that have different priorities:
  - Each slot of the high-priority line carries an ATM cell with probability p;
  - Each slot of the low-priority line carries one message with probability q; each message is composed of a random number of cells according to the PGF L(z). The packet arrival process on the low-priority line is **compound Bernoulli.**
- The ATM multiplexer stores the cells before transmission in a buffer of infinite capacity.
- The output line is **synchronous** with the input lines: input and output slot durations are equal; each output slot is used to convey one input cell.
- We have to study **the mean delay experienced by the cells of the low-priority line** due to the presence of the cells served of the high-priority line.

Exercise #1 (cont'd)

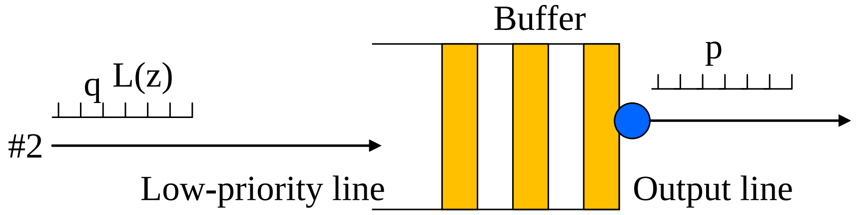
- This system admits a queuing model as depicted below:



The arrivals from line #1 have a non-preemptive priority with respect to those from line #2.

Solution of Exercise #1

- The presence of the high-priority traffic causes that output line slots are available for the low-priority traffic with probability 1 – p and unavailable with probability p. Hence, the equivalent service model for low-priority traffic is shown below:

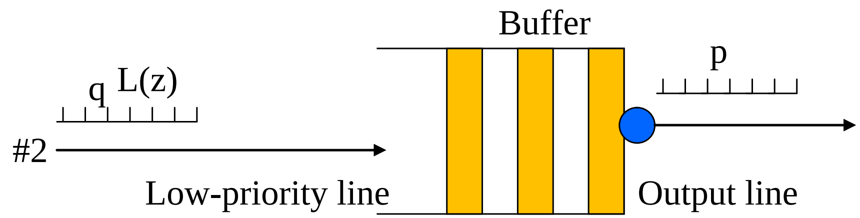


- **Three different imbedding choices can be made**, depending on the performance metric we need to measure. In the different cases, we have different meanings for n_i and a_i modeling the system. Since it is requested to determine the mean cell delay, we **imbed the system at the end of the slots of the output TDM line**.

Solution of Exercise #1

The service of the high-priority traffic from line #1 is unaffected by the service of the lower priority traffic from line #2. The high-priority traffic has no waiting delay, since it is immediately served at the arrival [A''(1) is equal to zero for this traffic].

- Th … pts ar … un … el fo



- **Three different imbedding choices can be made**, depending on the performance metric we need to measure. In the different cases, we have different meanings for n_i and a_i modeling the system. Since it is requested to determine the mean cell delay, we **imbed the system at the end of the slots of the output TDM line**.

Solution: Imbedding at the Slot End of the Output Line

- Let n_i denote the number of ATM cells in the buffer (from the low-priority line) at the end of the i-th slot of the output line.
- Let a_i denote the number of ATM cells (from the low-priority line) arrived at the buffer during the i-th slot.

$$n_{i+1} = \begin{cases} n_i - m + a_{i+1}, & n_i > 0 \\ a_{i+1}, & n_i = 0 \quad (*) \end{cases}$$

where m is a random variable defined as:

$$m = \begin{cases} 1, & \text{with Prob.} 1 - p \\ 0, & \text{with Prob. } p \end{cases}$$

- **We have obtained the same difference equation of the queue with feedback solved at the end of Lesson No. 9** (in that case, however, the arrival process is different, continuous time).

(*) At the i-th imbedding instant ξ_i^+, the queue is empty, $n_i = 0$. Hence, during the next slot no cell is transmitted and at the end of the next slot (instant ξ_{i+1}^-) the system contains the new requests a_{i+1}, arrived in the current slot. **With this type of imbedding instants, no service differentiation is needed for the case $n_i = 0$.**

Solution...

- Let A(z) denote the PGF of **number of cells arrived at the buffer in a slot from the low priority line**:

$$A(z) = 1 - q + qL(z)$$

- We achieve the following expression for the PGF of the number of cells in the queue from the low-priority line, P(z):

$$P(z) = P_0 \frac{(1-p)(z-1)A(z)}{z - [1 - p + zp]A(z)} \qquad P_0 = \frac{1 - p - A'(1)}{1-p}$$

Stability limit according to the condition $P_0 > 0$.

- Since P(z) has a singularity at z = 1, we can derive the mean number of cells in the buffer from the low-priority line, N_p, by **multiplying both sides of P(z) by the denominator and by differentiating twice:**

Traffic intensity in cells/slot

$$N_p = P'(1) = A'(1) + \frac{pA'(1)}{1 - p - A'(1)} + \frac{A''(1)}{2[1 - p - A'(1)]} \quad [\text{cells}]$$

Additional waiting term due to the fact that resources are not always available (prob. p).

Due to the availability of resources with prob. 1−p

Solution…

- **Stability condition for the high-priority line**: $p < 1$.
- **Stability condition for the low-priority line**: $1 - p > qL'(1)$. The low priority cells 'see' the output slot available with probability $1 - p$; this quantity must be bigger than the mean number of cells arrived per slot, $qL'(1)$.
- For $p = 0$ we re-obtain the classical M/G/1 solution:

$$N_p = A'(1) + \frac{A''(1)}{2[1 - A'(1)]} \quad [\text{cells}]$$

- By means of the Little theorem we can derive the mean packet delay T_p dividing N_p by the **mean packet arrival rate of $qL'(1)$ cells/slot that is equal to $A'(1)$**:

$$T_p = \frac{N_p}{A'(1)} = 1 + \frac{p}{1 - p - qL'(1)} + \frac{L''(1)/L'(1)}{2[1 - p - qL'(1)]} \quad [\text{slots}]$$

Exercise #2

- Messages arrive at a node of a telecommunication network to be transmitted on an output line. From measurements we know that the arrival process and the service process are characterized as follows:
  - Interarrival times υ are distributed so that $E[\upsilon^2] \approx 2E[\upsilon]^2$.
  - The message service time, τ, is characterized by a distribution so that $E[\tau^2] \approx E[\tau]^2$.
- We have to determine the mean delay experienced by a message to cross the node.

Solution of Exercise #2

- The interarrival times have mean square value and mean value that fulfill the typical relation of an **exponential distribution** with mean rate $1/E[\upsilon]$. Hence, we can assume that the message arrival process is Poisson.
- The message service time has mean square value and mean value that fulfill the typical relation of a **deterministic distribution** (i.e., Var = 0).
- We can study the node of the telecommunication network according to the **M/D/1 theory by imbedding the chain at the instants of message transmission completion**.

Solution (cont'd)

- We can express the mean message delay by means of the Pollaczek-Khinchin formula:

$$T_m = x + \frac{\lambda x^2}{2[1 - \lambda x]}$$

where $\lambda = 1/E[\upsilon]$ and $x = \tau$. System stability is assured if $\lambda x = \tau/E[\upsilon] < 1$ Erl.

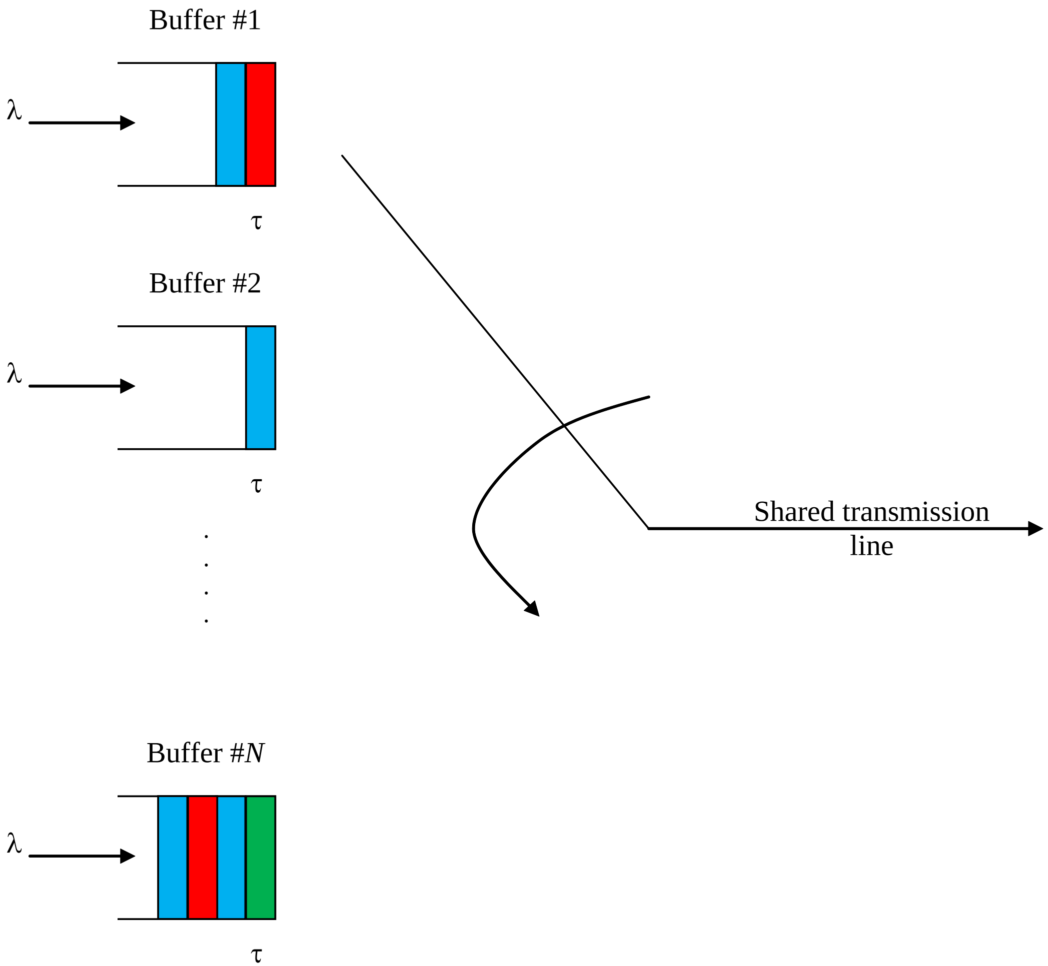
Exercise #3

- **Let us consider a scheduler for multiple flows sharing an output line as follows.** We refer to the transmission system outlined in the following Figure with N traffic flows (each modeled as an independent Poisson arrival of packets with mean rate λ), which correspond to distinct buffers served by a shared transmission line. Let τ denote the packet transmission time.

- The transmission line cyclically serves the different buffers according to a type of **Round Robin (RR) limited scheme**: the line transmits one packet from a buffer (if it is not empty) and then instantaneously switches to service the next buffer (**zero switch-over times**) according to a fixed service cycle.

- We have to determine the mean delay experienced by a packet from its arrival at the system to its departure.

Exercise #3 (cont'd)



Solution of Exercise #3

- Since the server (i.e., the transmission line) instantaneously switches from one buffer to the next one, **we can model the entire system as a single equivalent global queue with a specific service discipline** for the packets.

- The arrival process to this 'global' (virtual) queue is the sum of independent Poisson arrivals; hence, it is still Poisson with mean rate $N\lambda$.

- The transmission time of a packet is deterministic and equal to τ.

- **Therefore, the equivalent global queue admits an M/D/1 model**.

Solution of Exercise #3

- Since the server (i.e., the transmission line) instantaneously switches from one buffer to the next one, **we can model the entire system as a single equivalent global queue with a specific service dis**

- The arrival process to independent Poisson rate $N\lambda$.

- The transmission time of a pa rministic and equal to T.

- **Therefore, the equivalent global queue admits an M/D/1 model**.

Only if the **switching times** of the server from one queue to the other are null, we can model the whole system by means of an M/D/1 queue.

Solution (cont'd)

- We imbed the queue at the instants of packet transmission completions and we adopt the Pollaczek-Khinchin formula to express the mean packet delay T as:

$$T = \tau + \frac{\lambda N \tau^2}{2[1 - \lambda N \tau]}$$

- This system is stable if $\lambda N \tau < 1$ Erl.

- Note that an **M/G/1 queuing model with vacations** is needed to study the case with non-zero-switch-over times from the service of a queue to the service of the next queue.

Thank you!

giovanni.giambene@gmail.com