Slide supporting material

Lesson 9: Advanced M/G/1 Methods and Examples

Giovanni Giambene

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

M/G/1 with Different Imbedding Options and M/G/1 with Differentiated Service Times

Advanced 'M'/G/1 cases

Advanced 'M'/G/1 cases taking into account:

- Synchronization effects for the service of arrivals occurring at an empty queue
- Compound arrivals
- Batched service

Imbedding options

- Packet level (mean number of packets, N_p)
  - Packet completion instants
  - End of the slot
  - End of the frame
- Message level (mean msg delay, T_m)
  - Message completion instants

Advanced ‘M’/G/1 cases

- These ‘M’/G/1 cases are modeled by **generalized difference equations** as:

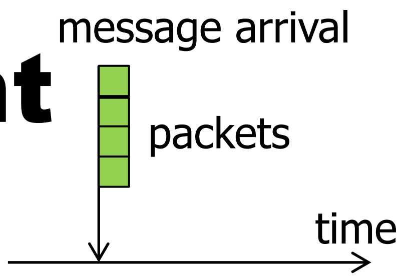
batched service (deterministic or random) per imbedding interval if b > 1

$$n_{i+1} = \begin{cases} \max\{n_i - b, 0\} + a_{i+1}, & \text{if } n_i \geq 1 \\ a_{i+1} + \Delta, & \text{if } n_i = 0 \end{cases}$$

service differentiation if $\Delta > 0$

  - The above is a symbolic difference equation presented to introduce new concepts, but not actually corresponding to a given queuing system.

'M'/G/1 Queue with Different Imbedding Options

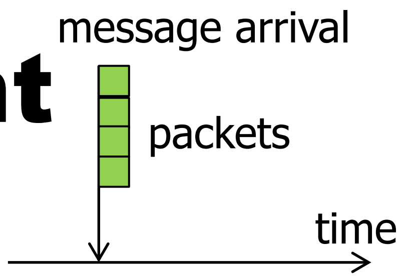


- Let us refer to a queue with a **compound Poisson arrival process.** Different imbedding options are available, also depending on the presence of an **output Time Division Multiplexing (TDM)/TDMA service.**

  1. **Imbedding at the end of the packet transmission time** to study the statistics of the buffer occupancy (like MAC layer performance). This study requires to adopt the service differentiation approach. Notation: $\mathbf{M^{[G]}/D/1}$.

  2. In a TDM output case, we can also imbed the system **at the end of the output slot**, thus avoiding any service differentiation issue. Notation: $\mathbf{M^{[G]}/D/1}$.

'M'/G/1 Queue with Different Imbedding Options

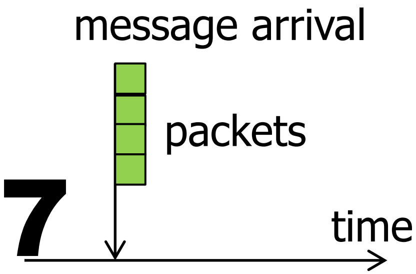


3. In a TDMA case and asynchronous multiplexing, we can imbed the system at the **end of the frame with b slots**, thus having a **batched service** since we can service up to b packets per frame. Notation: $\mathbf{M^{[G]}/D^{[b]}/1}$.

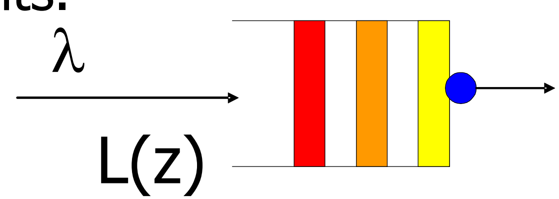
4. **Imbedding at the end of the message transmission time** to study the message delay distribution (like layer 3 performance). Notation: **M/G/1**. We use the Pollaczek-Khinchin formula.

- In the above cases #1, #2, #3, **operating at the packet level, the arrival process is not Poisson, and the Kleinrock principle is not applicable due to the simultaneous arrival of the packets of a message.** Hence, **<u>the 'M'/G/1 solution depends on the imbedding points</u>**, **<u>but in all the cases we must have the same stability limit.</u>**
- Instead, in the above case #4, the arrival process is Poisson and we can apply Kleinrock principle and PASTA property.

Let us Re-examine Exercise #2 of Lesson #7



- We consider a transmission line with a buffer (i.e., we have a single-server queue) where messages arrive according to a Poisson process with mean arrival rate λ.
- The arrival process and the transmission one are continuous-time.
- All the packets of the same message arrive simultaneously: bulk arrival process.
- A message is formed of a random number ℓ of packets, each requiring a time T to be transmitted. Message lengths are iid.
- Let L(z) denote the PGF of the message length in packets that also corresponds to the PGF of the message transmission time in T time-units.



Note that both inputs and outputs are unslotted. We have thus 2 different imbedding options (cases #1 and #4).

Solution #1: Chain Imbedded at pkt Transmission Completion

- Let n_i denote the number of packets in the buffer at the end of the transmission of the i-th packet; let a_i denote the number of packets arrived at the buffer during the service time of the i-th packet.
- To study our $\mathbf{M^{[L(z)]}/G/1}$ queue we write the following difference equation:
  - For $n_i > 0$, $n_{i+1} = n_i - 1 + a_{i+1}$: classical M/G/1 equation.
  - For $n_i = 0$, $n_{i+1} = a_{i+1} + \ell - 1$: when $n_i = 0$ we have to wait for the next group arrival and for the service completion of the first arrived packet of the group of length ℓ in order to go to the next imbedding instant with n_{i+1}. The difference equation in this case is that typical of **differentiated service times** (see next slides). However, considering that queuing phenomena when $n_i = 0$ could have negligible impact on the whole queue behavior, we make the following **approximation**: $n_{i+1} = a_{i+1}$, for $n_i = 0$. This allows to use the same difference equation of the classical M/G/1 theory.
- In order to apply the M/G/1 theory **we need to compute A(z)** that represents the PGF of the number of packets arrived at the buffer in the service time of a packet (time T).

Solution #1… Neglecting Service Differentiation ($n_{i+1} = a_{i+1}$ for $n_i = 0$)

- The mean number of packets N_p is:

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

where: $$A(z) = \sum_{n=0}^{\infty} [L(z)]^n \,\mathrm{Prob}\{n \text{ message arrivals in T}\} = e^{\lambda T(L(z)-1)}$$

z ← L(z)

due to the compound arrival process

$$A'(z=1) = \lambda T L'(1)$$

$$A''(z=1) = [\lambda T L'(1)]^2 + \lambda T L''(1)$$

- The stability of the buffer is assured if $\lambda TL'(1) < 1$ Erlang. The mean packet delay, T_p, is obtained applying the Little theorem and using the **mean packet arrival rate given by $\lambda L'(1)$** packets/second:

$$T_p = \frac{N_p}{\lambda L'(1)} \approx T + \frac{\lambda [T]^2 L'(1) + T \,{}^{L''(1)}\!/\!_{L'(1)}}{2[1 - \lambda T L'(1)]} \quad [\text{seconds}]$$

Solution #1b with Differentiated Service Times ($n_{i+1} = a_{i+1} + \Delta$ for $n_i = 0$)

The chain is again imbedded at the end of packet transmission:

$n_{i+1} = n_i - 1 + a_{i+1}$, if $n_i \neq 0$,

$n_{i+1} = a_{i+1}^*$, if $n_i = 0$ where $a_{i+1}^* = a_{i+1} + w$ and $w = \ell - 1$ (since $a_{i+1}^* > a_{i+1}$, this case is as if the service time was longer when the buffer is empty as with 'differentiated service times'). In terms of PGFs, A(z) does not change with respect to the previous example and W(z) = L(z)/z.

We need to solve the new difference equation in the z-domain and to compute the derivatives of the PGF P(z) at z = 1. We have:

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

Additional term due to differentiation depending on the message length statistics; **this term disappears if the messages are formed of a single packet** L(z) = z.

Classical M/G/1 terms (as in solution #1)

Note:

Random variable L → PGF L(z)

Random variable –1 → PGF z^{-1}

Solution #2: Chain Imbedded at msg Transmission Compl.

This is the same case of Exercise #2 in Lesson #7 !

- Let n_i denote the number of messages in the buffer at the end of the transmission of the i-th message; let a_i denote the number of messages arrived at the buffer during the service time of the i-th message.
- We can write the following difference equation of the classical M/G/1 type:
  - For $n_i > 0$, $n_{i+1} = n_i - 1 + a_{i+1}$
  - For $n_i = 0$, $n_{i+1} = a_{i+1}$.
- In order to apply the M/G/1 theory we need to compute A(z) that represents the PGF of the number of messages arrived at the buffer in the service time of a message.

Of course this case can be solved by **directly applying the Pollaczek-Khinchin formula.** We have however computed A(z) to compare to that of solution 1

$$A(z) = L\left[e^{\lambda T(z-1)}\right] \Rightarrow \begin{aligned} A'(z=1) &= L'(1)\lambda T \\ A''(z=1) &= [\lambda T]^2 [L''(1) + L'(1)] \end{aligned}$$

This composition is exactly the <u>opposite</u> of that used for solution #1

Solution #2...

- Let N_m denote the mean number of messages in the queue:

$$N_m = A'(1) + \frac{A''(1)}{2[1 - A'(1)]} \quad = L'(1)\lambda T + \frac{[\lambda T]^2 [L''(1) + L'(1)]}{2[1 - L'(1)\lambda T]} \quad [\text{pkts}]$$

- The stability condition is $\lambda T L'(1) < 1$ Erlang.
- Since the mean arrival rate of messages is λ, we apply the Little theorem to derive the mean message delay T_m:

$$T_m = \frac{N_m}{\lambda} = L'(1)T + \frac{\lambda [T]^2 [L''(1) + L'(1)]}{2[1 - L'(1)\lambda T]} \quad [\text{s}]$$

- **Let us consider messages with modified-geometric distribution** {so that $L''(1)=2[L'(1)]^2-2L'(1)$}. Hence, the mean message delay T_m results as:

$$T_m = L'(1)T + \frac{\lambda [T]^2 [2L'(1)^2 - L'(1)]}{2[1 - L'(1)\lambda T]} \quad [\text{s}]$$

Under the approximation considered in case #1 for the derivation of T_p, we can prove that $T_p \approx T_m$.

Still on the Comparison Between T_p and T_m

- Under the **common stability condition** $\lambda TL'(1) < 1$ Erlang, we have obtained the following mean packet delay T_p and mean message delay T_m with different imbedding options:

$$T_p \approx T + \frac{\lambda [T]^2 L'(1) + T \, {L''(1)}/{L'(1)}}{2[1 - \lambda T L'(1)]} \quad [\mathrm{s}]$$

Sol. #1 (approximate)

$$T_m = L'(1)T + \frac{\lambda [T]^2 [L''(1) + L'(1)]}{2[1 - L'(1)\lambda T]} \quad [\mathrm{s}]$$

Sol. #2

- **In general, we can see that $T_m \neq L'(1)\, T_p$ due to the queuing terms:**

$$L'(1)T_p = L'(1)T + \frac{\lambda [TL'(1)]^2 + TL''(1)}{2[1 - \lambda TL'(1)]} \neq T_m = L'(1)T + \frac{\lambda [T]^2 [L''(1) + L'(1)]}{2[1 - \lambda TL'(1)]}$$

Still on the Comparison Between T_p and T

- Under the **common stability con** we have obtained the following mean message delay T_m with different im

To remove the approximation we should add a term $L''(1)/[2\lambda L'(1)^2]$.

$$T_p \approx T + \frac{\lambda [T]^2 L'(1) + T\, {L''(1)}/{L'(1)}}{2[1-\lambda T L'(1)]} \quad [\mathrm{s}]$$

Sol. #1 (approximate)

$$T_m = L'(1)T + \frac{\lambda [T]^2 [L''(1) + L'(1)]}{2[1 - L'(1)\lambda T]} \quad [\mathrm{s}]$$

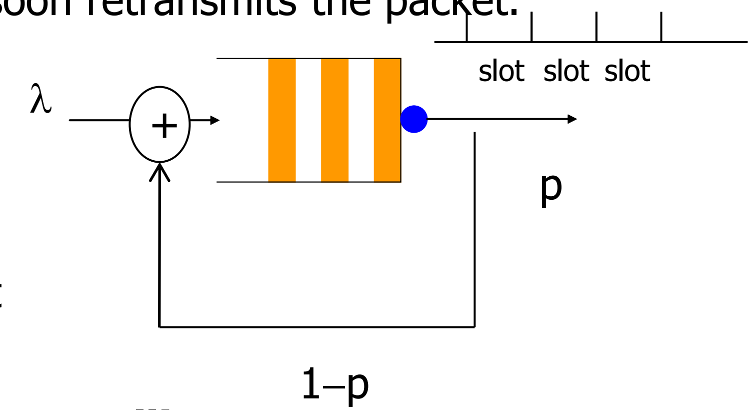
Sol. #2

- **In general, we can see that $T_m \neq L'(1)\, T_p$ due to the queuing terms:**

$$L'(1)T_p = L'(1)T + \frac{\lambda [TL'(1)]^2 + TL''(1)}{2[1-\lambda TL'(1)]} \neq T_m = L'(1)T + \frac{\lambda [T]^2 [L''(1) + L'(1)]}{2[1-\lambda TL'(1)]}$$

M/G/1 Queue with Feedback or Randomly-Available Server

- Let us consider a queue receiving packets according to a Poisson process with mean rate λ. Each packet needs a time T to be transmitted (deterministic service). Service time is slotted (synchronized) with duration T. When a packet is transmitted, the packet is erroneously received with probability 1–p ; in this case, a negative feedback (ARQ) is immediately received by the sender that soon retransmits the packet.



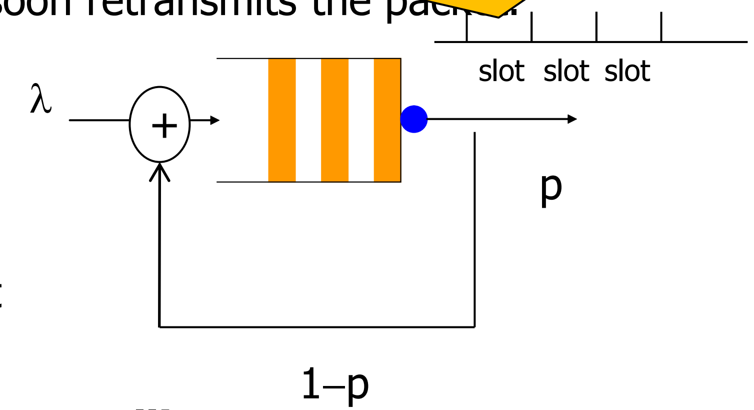
- We **imbed the queue at the slot end instants**.
  - Let n_i denote the number of packets in the queue at the end of the i-th slot;
  - Let a_i denote the number of packets arrived at the queue during the i-th slot.
  - **Variable X representing the service is a Bernoulli random variable** equal to 1 with probability p and equal to 0 with probability 1–p. X(z) = zp+ 1–p. E[X] = p.
- We can thus write the following difference equation modeling this queue:

$$n_{i+1} = \begin{cases} n_i - X + a_{i+1}, & \text{if } n_i \geq 1 \\ a_{i+1}, & \text{if } n_i = 0 \end{cases}$$

M/G/1 Queue with Feedback or Randomly-A

This model is equivalent to consider that the output server is available in a slot to send a packet with probability p and unavailable with probability 1–p (insensitivity property).

- Let us consider a queue recei
  rate λ. Each packet needs a t
  time is slotted (synchronized)
  packet is erroneously receive
  (ARQ) is immediately received by the sender that soon retransmits the pac
- We **imbed the queue at the slot end instants.**
  - Let n_i denote the number of packets in the queue at the end of the i-th slot;
  - Let a_i denote the number of packets arrived at the queue during the i-th slot.
  - **Variable X representing the service is a Bernoulli random variable** equal to 1 with probability p and equal to 0 with probability 1–p. X(z) = zp+ 1–p. E[X] = p.



- We can thus write the following difference equation modeling this queue:

$$n_{i+1} = \begin{cases} n_i - X + a_{i+1}, & \text{if } n_i \geq 1 \\ a_{i+1}, & \text{if } n_i = 0 \end{cases}$$

M/G/1 Queue with Feedback or Randomly-Available Server

- Let us c
rate λ.
time is
probabi
scheme

- We **iml**
  - Let
qu
  - Let
the
  - **Va**
ra
wit

This approach is correct, but **does not permit to model the synchronization delays in the service of packets arrived at an empty buffer**.

Otherwise, we could consider a different model of this system **without slots for output transmissions: transmissions are continuous-time.** We could thus avoid synchronization issues. In this case, we should imbed the study at the packet transmission end, thus using again the classical M/G/1 difference equation.

- We can thus write the following difference equation modeling this queue:

$$n_{i+1} = \begin{cases} n_i \ldots a_{i+1}, & \text{if } n_i \geq 1 \\ a_{i+1}, & \text{if } n_i = 0 \end{cases}$$

M/G/1 Queue … (cont'd)

- We solve the difference equation by transforming in the z-domain and using the same method adopted for the classical M/G/1 study.

- We obtain the following results for the PGF of the packets in the buffer, P(z), and the mean number of packets in the queue:

$$P(z) = P_0 \frac{p\left(1 - z^{-1}\right)A(z)}{1 - X\left(z^{-1}\right)A(z)} = P_0 \frac{p(z-1)A(z)}{z - \left[p + z(1-p)\right]A(z)} \qquad P_0 = \frac{p - A'(1)}{p}, \quad A(z) = e^{\lambda T(z-1)}$$

$$N = A'(1) + \frac{(1-p)A'(1)}{p - A'(1)} + \frac{A''(1)}{2\left[p - A'(1)\right]} \quad [\text{pkts}]$$

Note that for p =1 the feedback is eliminated and this expression yields the classical M/G/1 result.

where the stability condition is now $A'(1) = \lambda T < p$ Erl.

Note:
Random variable X → PGF X(z)
Random variable –X → PGF $X(z^{-1})$

The mean packet delay T can be obtained by applying the Little theorem to the whole queuing system: $T = N/\lambda$

Thank you!

giovanni.giambene@gmail.com