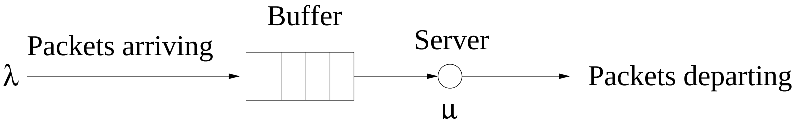
Queueing Delay

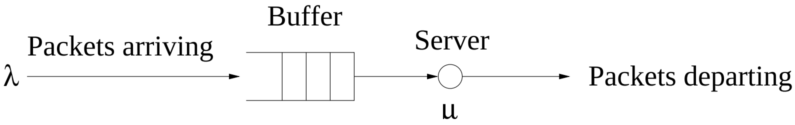


- λ: Packet arrival rate (packets/time)
- μ: Service rate (packets/time)
- $\rho = \lambda/\mu$: traffic intensity

We want to know:

- Average Number of Packets in the System
- Average Delay of a Packet (Queueing + Transmission Delay)

Queueing Delay

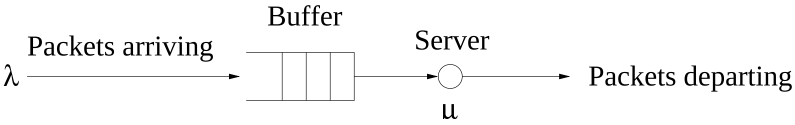


- λ: Packet arrival rate (packets/time)
- μ: Service rate (packets/time)
- $\rho = \lambda/\mu$: traffic intensity

We want to know:

- Average Number of Packets in the System
- Average Delay of a Packet (Queueing + Transmission Delay)

Queueing Delay



- λ: Packet arrival rate (packets/time)
- μ: Service rate (packets/time)
- $\rho = \lambda/\mu$: traffic intensity

We want to know:

- Average Number of Packets in the System
- Average Delay of a Packet (Queueing + Transmission Delay)

Notation

- $A(t)$: number of packets that arrived in $[0, t]$
- $B(t)$: number of packets that departed in $[0, t]$
- $N(t) = A(t) - B(t)$: number of packets in the system (in queue and in service) at time t.
- T_i**:** Time spent in the system by the ith arriving packet.
- $N_t = \frac{1}{t} \int_0^t N(\tau) d\tau$: time average number of packets in the system up to time t.
- $N = \lim_{t \to \infty} N_t$: (time) average number of packets in the system.

Notation

- $A(t)$: number of packets that arrived in $[0, t]$
- $B(t)$: number of packets that departed in $[0, t]$
- $N(t) = A(t) - B(t)$: number of packets in the system (in queue and in service) at time t.
- T_i: Time spent in the system by the ith arriving packet.
- $N_t = \frac{1}{t}\int_0^t N(\tau)d\tau$: time average number of packets in the system up to time t.
- $N = \lim_{t\to\infty} N_t$: (time) average number of packets in the system.

Notation

- $A(t)$: number of packets that arrived in $[0, t]$
- $B(t)$: number of packets that departed in $[0, t]$
- $N(t) = A(t) - B(t)$: number of packets in the system (in queue and in service) at time t.
- T_i: Time spent in the system by the ith arriving packet.
- $N_t = \frac{1}{t} \int_0^t N(\tau) d\tau$: time average number of packets in the system up to time t.
- $N = \lim_{t \to \infty} N_t$: (time) average number of packets in the system.

Notation

- $A(t)$: number of packets that arrived in $[0, t]$
- $B(t)$: number of packets that departed in $[0, t]$
- $N(t) = A(t) - B(t)$: number of packets in the system (in queue and in service) at time t.
- T_i: Time spent in the system by the ith arriving packet.
- $N_t = \frac{1}{t}\int_0^t N(\tau)d\tau$: time average number of packets in the system up to time t.
- $N = \lim_{t\to\infty} N_t$: (time) average number of packets in the system.

Notation

- $A(t)$: number of packets that arrived in $[0, t]$
- $B(t)$: number of packets that departed in $[0, t]$
- $N(t) = A(t) - B(t)$: number of packets in the system (in queue and in service) at time t.
- T_i**:** Time spent in the system by the ith arriving packet.
- $N_t = \frac{1}{t}\int_0^t N(\tau)d\tau$: time average number of packets in the system up to time t.
- $N = \lim_{t\to\infty} N_t$: (time) average number of packets in the system.

Notation

- $A(t)$: number of packets that arrived in $[0, t]$
- $B(t)$: number of packets that departed in $[0, t]$
- $N(t) = A(t) - B(t)$: number of packets in the system (in queue and in service) at time t.
- T_i**:** Time spent in the system by the ith arriving packet.
- $N_t = \frac{1}{t}\int_0^t N(\tau)d\tau$: time average number of packets in the system up to time t.
- $N = \lim_{t\to\infty} N_t$: (time) average number of packets in the system.

Notation

- $A(t)$: number of packets that arrived in $[0, t]$
- $B(t)$: number of packets that departed in $[0, t]$
- $N(t) = A(t) - B(t)$: number of packets in the system (in queue and in service) at time t.
- T_i**:** Time spent in the system by the ith arriving packet.
- $N_t = \frac{1}{t}\int_0^t N(\tau)d\tau$: time average number of packets in the system up to time t.
- $N = \lim_{t\to\infty} N_t$: (time) average number of packets in the system.

Notation (continued)

- $\lambda_t = \frac{A(t)}{t}$: time average arrival rate over $[0, t]$
- $\lambda = \lim_{t \to \infty} \lambda_t$: steady state arrival rate
- $T_t = \frac{\sum_{i=0}^{A(t)} T_i}{A(t)}$: average time spent in the system per packet up to time t
- $T = \lim_{t \to \infty} T_t$: steady-state time average packet delay.

Notation (continued)

- $\lambda_t = \frac{A(t)}{t}$: time average arrival rate over $[0, t]$
- $\lambda = \lim_{t \to \infty} \lambda_t$: steady state arrival rate
- $T_t = \frac{\sum_{i=0}^{A(t)} T_i}{A(t)}$: average time spent in the system per packet up to time t
- $T = \lim_{t \to \infty} T_t$: steady-state time average packet delay.

Notation (continued)

- $\lambda_t = \frac{A(t)}{t}$: time average arrival rate over $[0, t]$
- $\lambda = \lim_{t\to\infty} \lambda_t$: steady state arrival rate
- $T_t = \frac{\sum_{i=0}^{A(t)} T_i}{A(t)}$: average time spent in the system per packet up to time t
- $T = \lim_{t\to\infty} T_t$: steady-state time average packet delay.

Notation (continued)

- $\lambda_t = \frac{A(t)}{t}$: time average arrival rate over $[0, t]$
- $\lambda = \lim_{t \to \infty} \lambda_t$: steady state arrival rate
- $T_t = \frac{\sum_{i=0}^{A(t)} T_i}{A(t)}$: average time spent in the system per packet up to time t
- $T = \lim_{t \to \infty} T_t$: steady-state time average packet delay.

Outline:

- **Little's Theorem**: $N = \lambda T$
- **M/M/1 queue:** $N = \sum_{n=0}^{\infty} np_n$,
  where p_n is the steady-state probability that n packets are in the system.
- Traffic on a rainy day
- Fast-food restaurants

Outline:

- **Little's Theorem**: $N = \lambda T$
- **M/M/1 queue:** $N = \sum_{n=0}^{\infty} np_n$,
  where p_n is the steady-state probability that n packets are in the system.
- Traffic on a rainy day
- Fast-food restaurants

Outline:

- **Little's Theorem**: $N = \lambda T$
- **M/M/1 queue:** $N = \sum_{n=0}^{\infty} n p_n$,
  where p_n is the steady-state probability that n packets are in the system.
- Traffic on a rainy day
- Fast-food restaurants

Outline:

- **Little's Theorem**: $N = \lambda T$
- **M/M/1 queue:** $N = \sum_{n=0}^{\infty} np_n$, where p_n is the steady-state probability that n packets are in the system.

Examples for Little's Theorem:

- Traffic on a rainy day
- Fast-food restaurants

Outline:

- **Little's Theorem**: $N = \lambda T$
- **M/M/1 queue:** $N = \sum_{n=0}^{\infty} np_n$,
  where p_n is the steady-state probability that n packets are in the system.

Examples for Little's Theorem:

- Traffic on a rainy day
- Fast-food restaurants

Probabilistic Formulation of Little's Theorem:

So far: time average - "I observe the system for a long, long time".
Next: ensemble average - "I come back at time t to check on the system "

- $p_n(t)$: Probability that n packets are in the system at time t.
- $\bar{N}(t) = \sum_{n=0}^{\infty} np_n(t)$: expected number of packets in the system at time t
- $p_n = \lim_{t\to\infty} p_n(t)$: steady-state probability that n packets are in the system
- $\bar{N} = \sum_{n=0}^{\infty} np_n$: steady-state expected number of packets in the system
- $\bar{T}_k$: expected delay of the kth packet.
- $\bar{T} = \lim_{k\to\infty} \bar{T}_k$: expected packet delay.

Probabilistic Formulation of Little's Theorem:

So far: time average - "I observe the system for a long, long time".
Next: ensemble average - "I come back at time t to check on the system "

- $p_n(t)$: Probability that n packets are in the system at time t.
- $\bar{N}(t) = \sum_{n=0}^{\infty} n p_n(t)$: expected number of packets in the system at time t
- $p_n = \lim_{t\to\infty} p_n(t)$: steady-state probability that n packets are in the system
- $\bar{N} = \sum_{n=0}^{\infty} n p_n$: steady-state expected number of packets in the system
- $\bar{T}_k$: expected delay of the kth packet.
- $\bar{T} = \lim_{k\to\infty} \bar{T}_k$: expected packet delay.

Probabilistic Formulation of Little's Theorem:

So far: time average - "I observe the system for a long, long time".
Next: ensemble average - "I come back at time t to check on the system "

- $p_n(t)$: Probability that n packets are in the system at time t.
- $\bar{N}(t) = \sum_{n=0}^{\infty} np_n(t)$**:** expected number of packets in the system at time t
- $p_n = \lim_{t\to\infty} p_n(t)$: steady-state probability that n packets are in the system
- $\bar{N} = \sum_{n=0}^{\infty} np_n$: steady-state expected number of packets in the system
- $\bar{T}_k$**:** expected delay of the kth packet.
- $\bar{T} = \lim_{k\to\infty} \bar{T}_k$**:** expected packet delay.

Probabilistic Formulation of Little's Theorem:

So far: time average - "I observe the system for a long, long time".
Next: ensemble average - "I come back at time t to check on the system "

- $p_n(t)$: Probability that n packets are in the system at time t.
- $\bar{N}(t) = \sum_{n=0}^{\infty} np_n(t)$**:** expected number of packets in the system at time t
- $p_n = \lim_{t\to\infty} p_n(t)$: steady-state probability that n packets are in the system
- $\bar{N} = \sum_{n=0}^{\infty} np_n$: steady-state expected number of packets in the system
- $\bar{T}_k$**:** expected delay of the kth packet.
- $\bar{T} = \lim_{k\to\infty} \bar{T}_k$**:** expected packet delay.

Probabilistic Formulation of Little's Theorem:

So far: time average - "I observe the system for a long, long time".
Next: ensemble average - "I come back at time t to check on the system "

- $p_n(t)$: Probability that n packets are in the system at time t.
- $\bar{N}(t) = \sum_{n=0}^{\infty} n p_n(t)$**:** expected number of packets in the system at time t
- $p_n = \lim_{t\to\infty} p_n(t)$: steady-state probability that n packets are in the system
- $\bar{N} = \sum_{n=0}^{\infty} n p_n$: steady-state expected number of packets in the system
- $\bar{T}_k$**:** expected delay of the kth packet.
- $\bar{T} = \lim_{k\to\infty} \bar{T}_k$**:** expected packet delay.

Probabilistic Formulation of Little's Theorem:

So far: time average - "I observe the system for a long, long time".
Next: ensemble average - "I come back at time t to check on the system "

- $p_n(t)$: Probability that n packets are in the system at time t.
- $\bar{N}(t) = \sum_{n=0}^{\infty} np_n(t)$**:** expected number of packets in the system at time t
- $p_n = \lim_{t\to\infty} p_n(t)$: steady-state probability that n packets are in the system
- $\bar{N} = \sum_{n=0}^{\infty} np_n$: steady-state expected number of packets in the system
- $\bar{T}_k$**:** expected delay of the kth packet.
- $\bar{T} = \lim_{k\to\infty} \bar{T}_k$**:** expected packet delay.

Probabilistic Formulation of Little's Theorem:

So far: time average - "I observe the system for a long, long time".
Next: ensemble average - "I come back at time t to check on the system "

- $p_n(t)$: Probability that n packets are in the system at time t.
- $\bar{N}(t) = \sum_{n=0}^{\infty} np_n(t)$**:** expected number of packets in the system at time t
- $p_n = \lim_{t\to\infty} p_n(t)$: steady-state probability that n packets are in the system
- $\bar{N} = \sum_{n=0}^{\infty} np_n$: steady-state expected number of packets in the system
- $\bar{T}_k$**:** expected delay of the kth packet.
- $\bar{T} = \lim_{k\to\infty} \bar{T}_k$**:** expected packet delay.

Ergodic System and Little's Theorem:

Ergodic Systems:

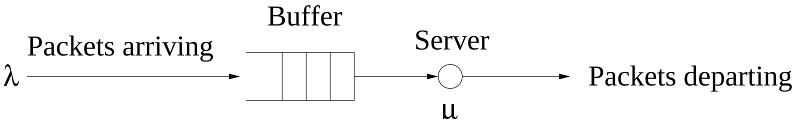
$$N = \bar{N} \qquad \text{and} \qquad T = \bar{T}$$

For ergodic systems, **Little's formula** holds with $N = \bar{N}$ and $T = \bar{T}$,

and with

$$\lambda = \lim_{t \to \infty} \frac{\text{Expected number of arrivals in the interval } [0, t]}{t}$$

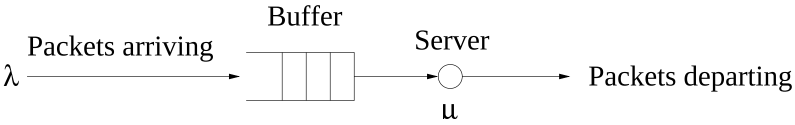
M/M/1 Queue



- Customers (packets) arrive according to a Poisson process
- Service time is exponentially distributed

.

M/M/1 Queue



- Customers (packets) arrive according to a Poisson process
- Service time is exponentially distributed

Goal: want to determine the steady-state probability p_n that n customers are in the system.

Many Applications

- Call Centers
- Traffic planning
- Requests at a Web server