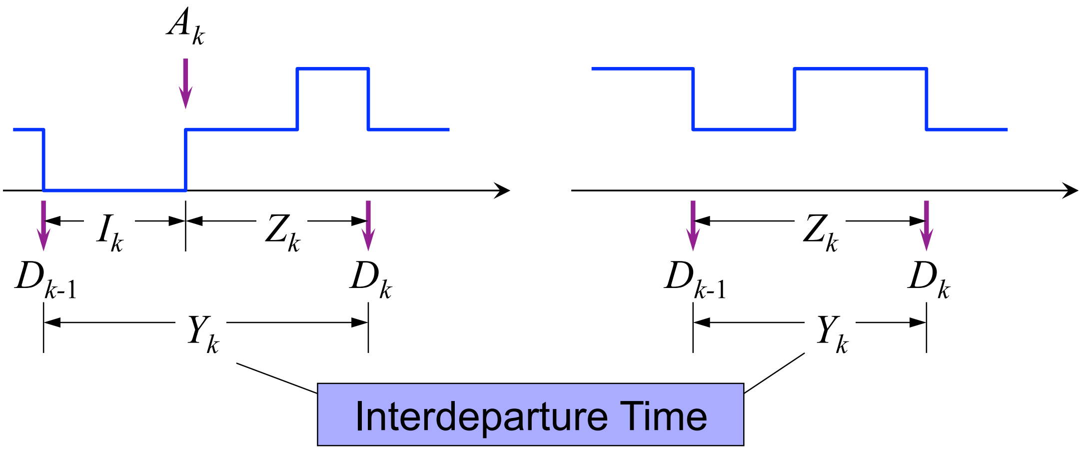
Queueing Theory II

Summary

- M/M/1 Output process
- Networks of Queue
- Method of Stages
  - Erlang Distribution
  - Hyperexponential Distribution
- General Distributions
  - Embedded Markov Chains

M/M/1 Output Process

- **Burke's Theorem**:
- The Departure process of a stable M/M/1 queueing system with arrival rate λ is also a Poisson process with rate λ.



Burke's Theorem

$$\Pr\{Y_k \le y\} = \Pr\{Y_k \le y \mid X(A_k^-) = 0\}\Pr\{X(A_k^-) = 0\} + \Pr\{Y_k \le y \mid X(A_k^-) > 0\}\Pr\{X(A_k^-) > 0\}$$

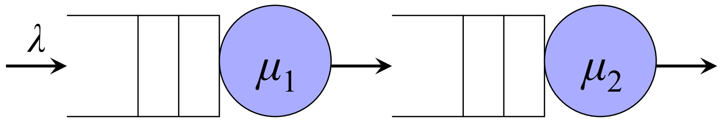
$$\Pr\{X(A_k^-) = 0\} = \pi_0 = 1 - \rho$$ PASTA

$$\Pr\{X(A_k^-) > 0\} = 1 - \pi_0 = \rho$$ PASTA

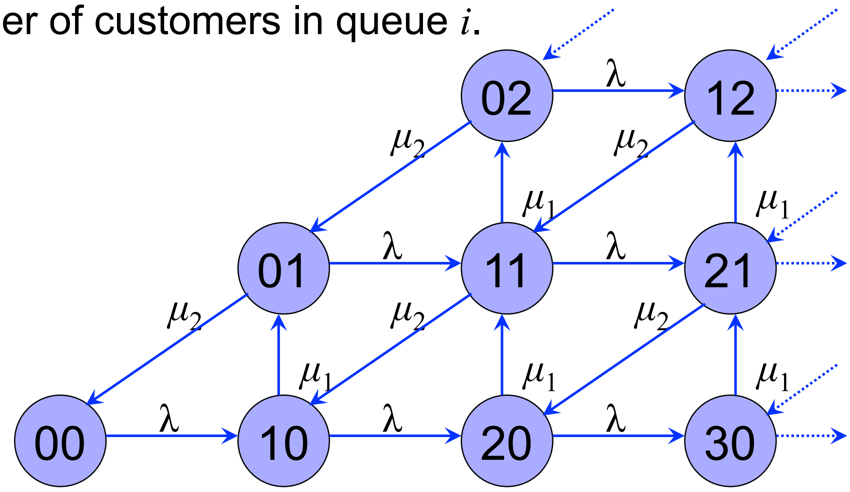
$$\Pr\{Y_k \le y \mid X(A_k^-) > 0\} = \Pr\{Z_k \le y\} = 1 - e^{-\mu y}$$

$$\Pr\{Y_k \le y \mid X(A_k^-) = 0\} = \Pr\{I_k + Z_k \le y\} = \frac{\mu}{\mu - \lambda}[1 - e^{-\lambda y}] - \frac{\lambda}{\mu - \lambda}[1 - e^{-\mu y}]$$

Two Queues in Series



- Let the state of this system be (X_1,X_2) where X_i is the number of customers in queue i.



Two Queues in Series

- Balance Equations

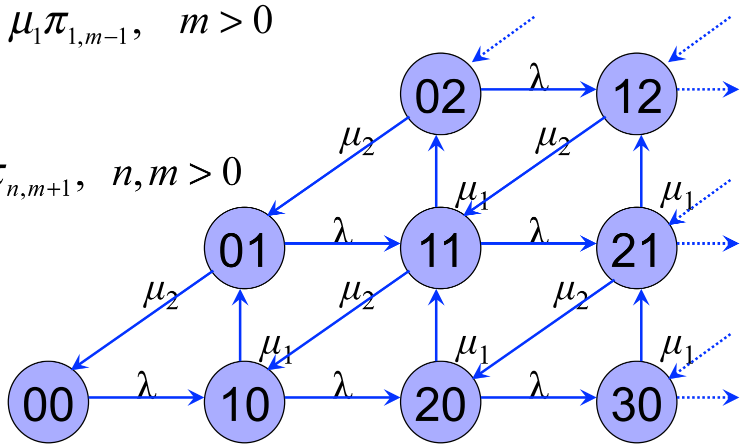
$$\lambda\pi_{00} = \mu_2\pi_{01}$$

$$(\lambda+\mu_1)\pi_{n0} = \mu_2\pi_{n,1} + \lambda\pi_{n-1,0}, \quad n>0$$

$$(\lambda+\mu_2)\pi_{0m} = \mu_2\pi_{0,m+1} + \mu_1\pi_{1,m-1}, \quad m>0$$

$$(\lambda+\mu_1+\mu_2)\pi_{nm} =$$
$$\lambda\pi_{n-1,m} + \mu_1\pi_{n+1,m-1} + \mu_2\pi_{n,m+1}, \quad n,m>0$$

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\pi_{nm} = 1$$



Product Form Solution

- Let $\rho_1 = \lambda/\mu_1$, $\rho_2 = \lambda/\mu_2$

$$\pi_{nm} = (1-\rho_1)\rho_1^n (1-\rho_2)\rho_2^m$$

- Recall the M/M/1 Results

$$\pi_n = (1-\rho)\rho^n$$

- Therefore the two queues can be "decoupled" and studied in isolation.

$$\pi_{nm} = \pi_n^1 \pi_m^2$$

Jackson Networks

- The product form decomposition holds for all open queueing networks with Poisson input processes that do not include feedback
- Even though customer feedback causes the total input process (external Poisson and feedback) become non-Poisson, the product form solution still holds!
- These types on networks are referred to as Jackson Networks
- The total input rate to each note is given by

Aggregate Rate

External Arrival Rate

Routing Prob

$$\Lambda_i = \lambda_i + \sum_{j=1}^{n} \Lambda_j r_{ji}$$

Closed Networks

- The product form decomposition holds for also for closed networks

Aggregate Rate

$$\Lambda_i = \sum_{j=1}^{n} \Lambda_j r_{ji}$$

Routing Prob

Non-Poisson Processes

- Assume that the service time distribution can be approximated by the sum of m iid exponential random variables with rate $m\mu$. That is

$$Z = \sum_{j=1}^{m} Y_j \qquad \text{where} \qquad Y_j \sim F_Y(y) = 1 - e^{-m\mu y}$$

- You can show that the density of Z is given by

$$f_Z(t) = \frac{m\mu (m\mu t)^{m-1}}{(m-1)!} e^{-m\mu t}, \qquad t \geq 0$$

- Z is an Erlang random variable with parameters (m, μ).

Erlang Distribution

- You can show that the distribution function of Z

$$F_Z(t) = 1 - e^{-m\mu t} \sum_{j=0}^{m-1} \frac{(m\mu t)^j}{j!}, \qquad t \geq 0$$

- The expected value of Z is given by

$$E[Z] = E\left[\sum_{j=1}^{m} Y_j\right] = \sum_{j=1}^{m} E\left[Y_j\right] = \sum_{j=1}^{m} \frac{1}{m\mu} = \frac{m}{m\mu} = \frac{1}{\mu}$$

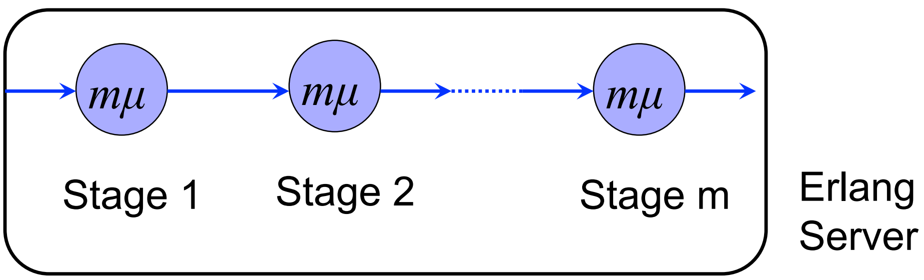
- The variance of Z is given by

$$\text{var}[Z] = \text{var}\left[\sum_{j=1}^{m} Y_j\right] = \sum_{j=1}^{m} \text{var}\left[Y_j\right] = \sum_{j=1}^{m} \frac{1}{(m\mu)^2} = \frac{m}{(m\mu)^2} = \frac{1}{m\mu^2}$$

- Note that the variance of an Erlang is always less than or equal to the variance of the exponential random variable

M/Er$_m$/1 Queueing System

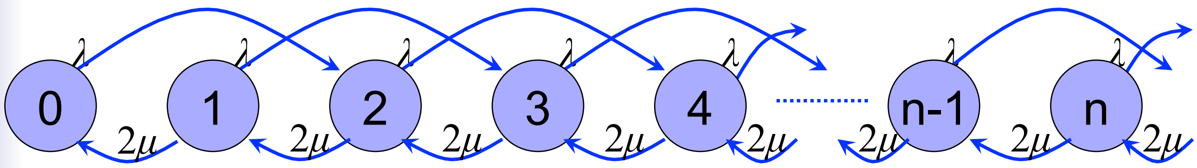
- **Meaning**: Poisson Arrivals, Erlang distributed service times (order *m*), single server and infinite capacity buffer.



- Only a single customer is allowed in the server at any given time.
- The customer has to go through all m stages before it is released.

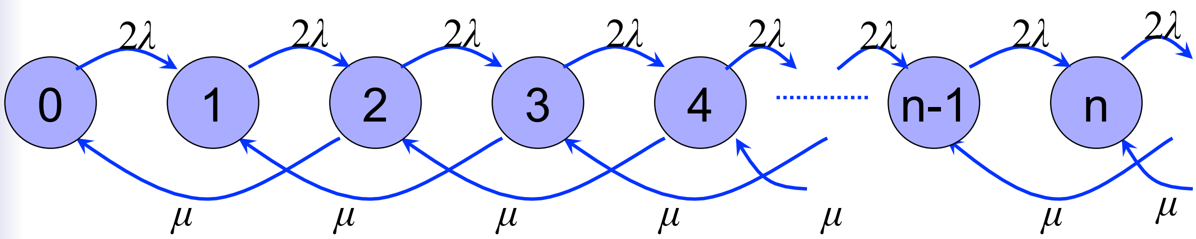
M/Er$_m$/1 Queueing System

- The state of the system needs to take into account the stage that the customer is in, so one could use (x,s) where x=0,1,2… is the number of customers and s=1,…,m is the stage that the customer in service is currently in.
- Alternatively, one can use a single variable y to be the total number of stages that need to be completed before the system empties.
- Let m=2



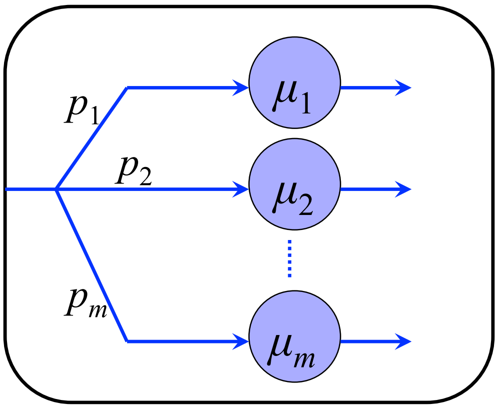
Er$_m$/M/1 Queueing System

- Similar to the M/Er$_m$/1 System we can let the state be the number of arrival stages in the system, so for example, for m=2



Hyperexponential Distribution

- Recall that the variance of the Erlang distribution is always less than or equal than the variance of the exponential distribution.
- What if we need service times with higher variance?



Hyperexponential Server

Hyperexponential Distribution

- Expected values if the Hyperexponential distribution

$$E[Z] = \sum_{j=1}^{m} p_j E\left[Y_j\right] = \sum_{j=1}^{m} \frac{p_j}{\mu_j}$$

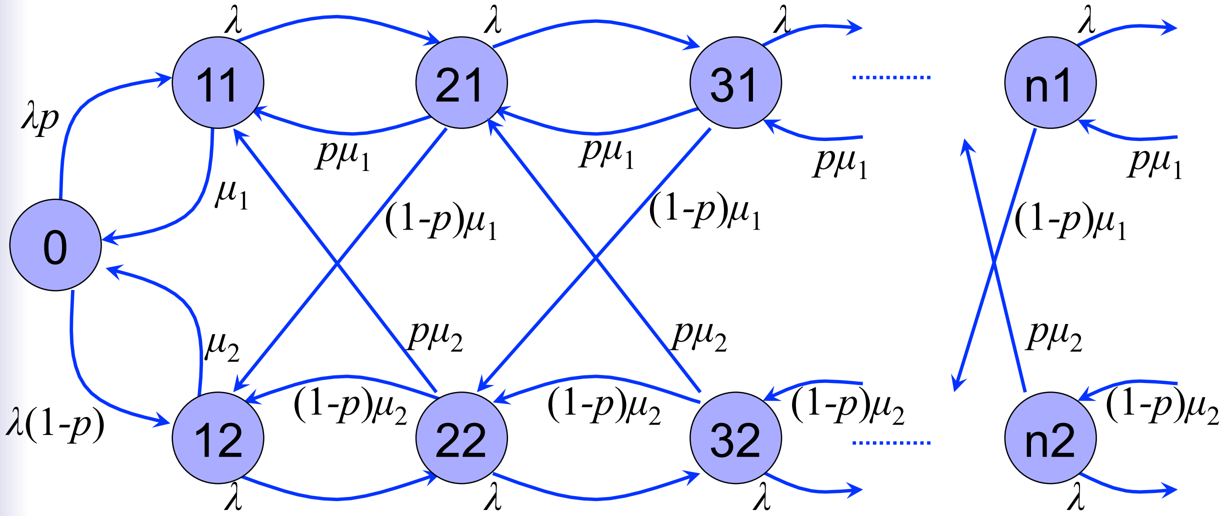
- The variance of Z is given by

$$E[Z^2] = \sum_{j=1}^{m} p_j E\left[Y_j^2\right] = 2\sum_{j=1}^{m} \frac{p_j}{\mu_j^2}$$

$$\Rightarrow \operatorname{var}[Z] = E[Z^2] - (E[Z])^2 = 2\sum_{j=1}^{m} \frac{p_j}{\mu_j^2} - \left(\sum_{j=1}^{m} \frac{p_j}{\mu_j}\right)^2$$

M/H_m/1 Queueing System

- In this case, the state of the system should include both, the number of customers in the system and the stage that the customer in service is in. For example, for m=2

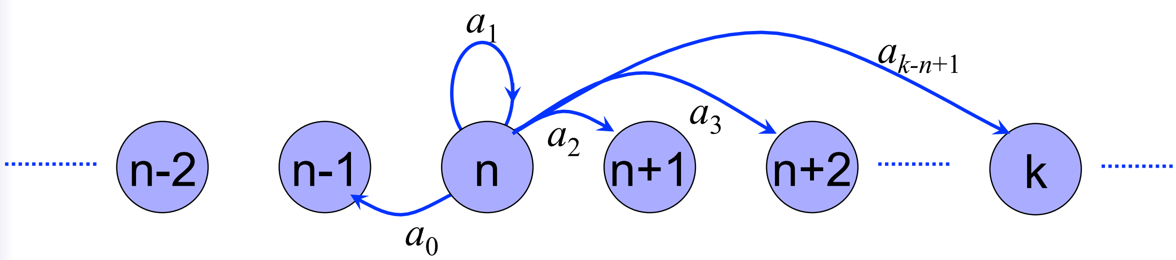


M/G/1 Queueing System

- When the arrival and service processes do not possess the memoryless property, we are back to the GSMP framework.
- In general, we will need to keep track of the age or residual lifetime of each event.
- Embedded Markov Chains
  - Some times it may be possible to identify specific points such that the Markov property holds
  - For example, for the M/G/1 system, suppose that we choose to observe the state of the system exactly after each customer departure.
  - For this problem, it doesn't matter how long ago the previous arrival occurred since arrivals are generated from Poisson processes. Furthermore, at the point of a departure, we know that the age of the event is always 0.

Embedded Markov Chains

- So, right after each departure we observe the following chain (only the transitions from state n are drawn)
- Let a_j be the probability that j arrivals will occur during the interval Y defined by two consecutive departures



- Let N be a random variable that indicates the number of arrivals during Y, then $a_j = \Pr\{N=j\}$.

Embedded Markov Chains

- Suppose we are given the density of Y, $f_Y(y)$ then

$$a_j = \Pr\{N = j\} = \int_0^\infty \Pr\{N = j \mid Y = y\} f_Y(y)dy$$

- Since we have a Poisson process

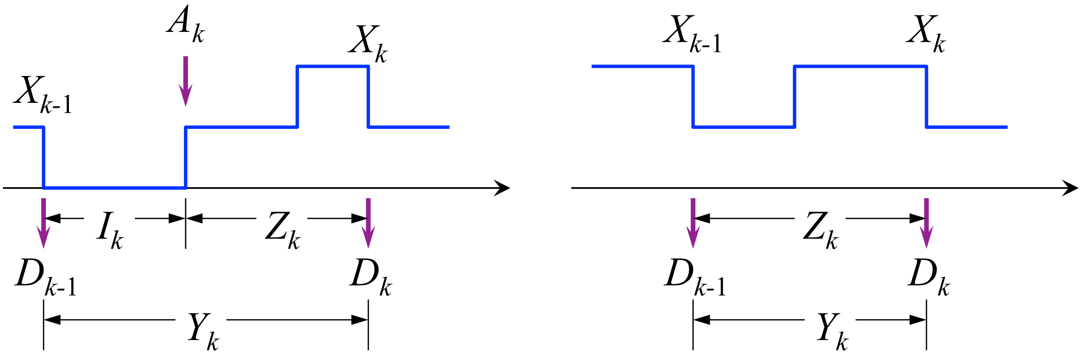
$$\Pr\{N = j \mid Y = y\} = \frac{(\lambda y)^j}{j!} e^{-\lambda y}$$

- Therefore

$$a_j = \int_0^\infty \frac{(\lambda y)^j}{j!} e^{-\lambda y} f_Y(y)dy$$

State Iteration

- Let X_k be the state of the system just after the departure of the kth customer



$$X_k = X_{k-1} + N_k \qquad\qquad X_k = X_{k-1} + N_k - 1$$

Arrivals During Y_k

$$X_k = X_{k-1} + N_k - \mathbf{1}\{X_{k-1} > 0\}$$

M/G/1 Queueing System

- Pollaczek-Khinchin (PK) Formula

$$E[X] = \frac{\rho}{1-\rho} - \frac{\rho^2}{2(1-\rho)}\left(1-\mu^2\sigma^2\right)$$

where

- X is the number of customers in the system
- $1/\mu$ is the average service time
- σ^2 is the variance of the service time distribution
- λ is the Poisson arrival rate
- $\rho=\lambda/\mu$ is the traffic intensity

M/G/1 Example

- Consider the queueing systems M/M/1 and M/D/1
- Compare their average number in the system and average system delay when the arrival is Poisson with rate λ and the mean service time is $1/\mu$.
- For the M/M/1 system, $\sigma^2=1/\mu^2$, therefore

$$E[X]_{M/M/1} = \frac{\rho}{1-\rho} \qquad E[S]_{M/M/1} = \frac{1/\mu}{1-\rho}$$

- For the M/D/1 system, $\sigma^2=0$, therefore

$$E[X]_{M/D/1} = \frac{\rho}{(1-\rho)}\left(1-\frac{\rho}{2}\right) \qquad E[S]_{M/D/1} = \frac{1/\mu}{1-\rho}\left(1-\frac{\rho}{2}\right)$$