# Basic Queueing Theory (I)

Cheng-Fu Chou



# Outline

- Little result
- M/M/1
- Its variant
- Method of stages



# Queueing System

- Kendall's notations
  - A/B/C/K
  - C: number of servers
  - K: the size of the system capacity; the buffer space including the servers
- A(t): the inter-arrival time dist.
- B(t): the service time dist.
  - M: exponential dist.
  - G: general dist.
  - D: deterministic dist.



# Time Diagram for queues

- $C_n$: the n-th customer to enter the systsem
- N(t): number of customers in the system at time t
- U(t): unfinished work in the system at time t
- $\tau_n$: arrival time for $C_n$
- $t_n$: inter-arrival time between $C_{n-1}$ and $C_n$, i.e., $A(t) = P[t_n \leq t]$
- $x_n$: service time for $C_n$, $B(t) = P[x_n \leq t]$
- $w_n$: waiting time for $C_n$
- $s_n$: system time for $C_n = w_n + x_n$
  - Draw the diagram

# Little Result

- $\alpha(t)$: no. of arrivals in $(0,t)$
- t(t): no. of departures in (0,t)
- lt: the average arrival rate during the interval (0,t)
- r(t): the total time all customers have spent in the system during (0,t)
- Tt: the average system time during (0,t)
  - proof



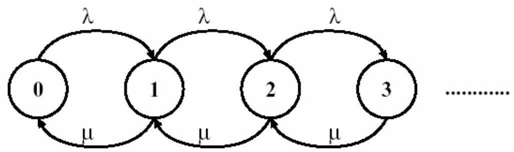
# M/M/1

- The average inter-arrival time is $t = 1/\lambda$ and $t$ is exponentially distributed.
- The average service time is $x = 1/\mu$ and $x$ is exponentially distributed.
- Find out
  - $p_k$ : the prob. of finding k customers in the system
  - N : the avg. number of customers in the system
  - T : the avg. time spent in the system



# M/M/1

- Poisson arrival



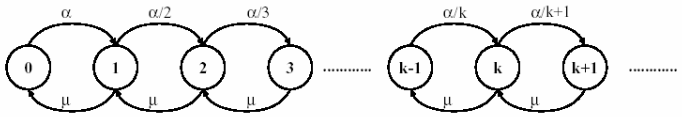



# Discouraged Arrival

- A system where arrivals tend to get discouraged when more and more people are present in the system
  - arrival rate: $\lambda_k = \alpha/(k+1)$ , where $k = 0,1,2,\ldots$
  - service rate: $\mu_k = \mu$, where $k = 1,2,3,\ldots$

# Discouraged Arrival

$$\lambda_k = \frac{\alpha}{k+1} \qquad k = 0, 1, 2, \cdots \qquad \mu_k = \mu$$








# M/M/∞

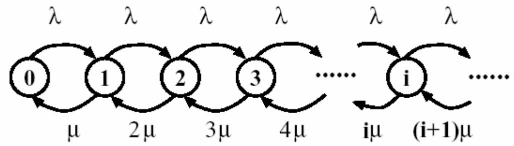
- Infinite number of servers
  - there is always a new server available for each arriving customer.
  - arrival rate : $\lambda$
  - service rate of each server: $\mu$





# M/M/∞

- We know
  - Arrival rate $\lambda_k = \lambda$ , $k = 0, 1, 2, \ldots$
  - Departure rate $\mu_k = k\mu$ , $k = 1, 2, 3, \ldots$





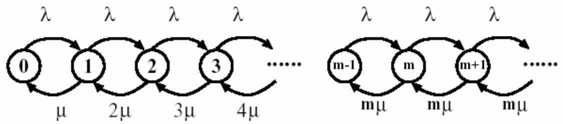



# M/M/m

- The m-server case
  - The system provides a maximum of *m* servers

# M/M/m

- Arrival rate $\lambda_k = \lambda$ and service rate $\mu_k = \min(k\mu, m\mu)$



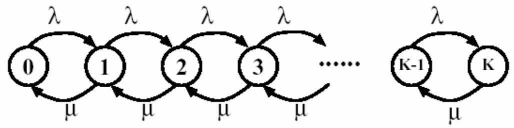



# M/M/1/K

- Finite storage: a system in which there is a maximum number of customers that may be stored ( K customers)



# M/M/1/K



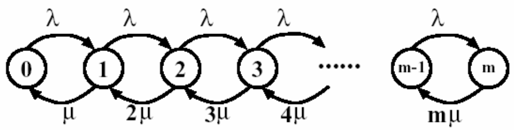



# M/M/m/m

- m-server loss system

# M/M/m/m (m-server loss system)

- ❑ m-server loss systems



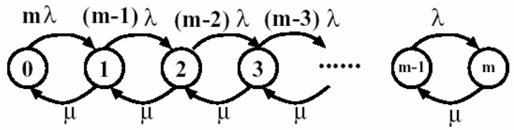



# M/M/1//m

- ❑ Finite customer population and single server
  - A single server
  - There are total m customers



# M/M/1//m (finite customer population)



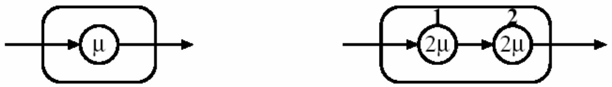



# PASTA

- ❑ Poisson arrival see time average

# Method of stages

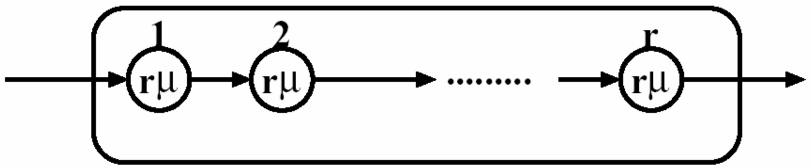
- Erlangian distribution






# Er: r-stage Erlangian Dist.

- r-stage Erlangian dist.






# M/Er/1

- HW: $E_3$/M/1



# Bulk arrival systems