

# Advanced Queuing Models Theory

## 1. M/M/c/K Queue (Finite Capacity)

Description: Multi-server queue with finite system capacity ( $K$ ). Customers arriving when the system is full are blocked.

Parameters:  $\lambda$  (arrival rate),  $\mu$  (service rate per server),  $c$  (number of servers),  $K$  (system capacity).

Performance Measures:

- $P_0$ : Probability system is empty.
- $P_k$ : Probability system has  $k$  customers.
- Blocking probability:  $P(K)$ .
- Effective arrival rate =  $\lambda * (1 - P(K))$ .

Applications: Parking lots, hospital beds, call centers with fixed lines.

## 2. M/M/Infinity Queue (Infinite Servers)

Description: Each arrival gets immediate service, no queue forms.

Parameters:  $\lambda$  (arrival rate),  $\mu$  (service rate per server).

Performance Measures:

- Number in system follows Poisson distribution with mean =  $\lambda / \mu$ .
- Little's Law applies:  $L = \lambda * W$ .

Applications: Cloud computing, phone switching centers, large-scale resource systems.

## 3. Priority Queues (M/M/1 with Priority)

Description: Customers divided into classes with different service priorities.

Types:

- Preemptive: High-priority customers interrupt service of lower-priority ones.
- Non-preemptive: High priority served first but ongoing service is not interrupted.

Performance Measures depend on priority discipline and arrival/service rates for each class.

Applications: Hospital ER triage, network packet handling, emergency dispatch systems.