Lecture #25

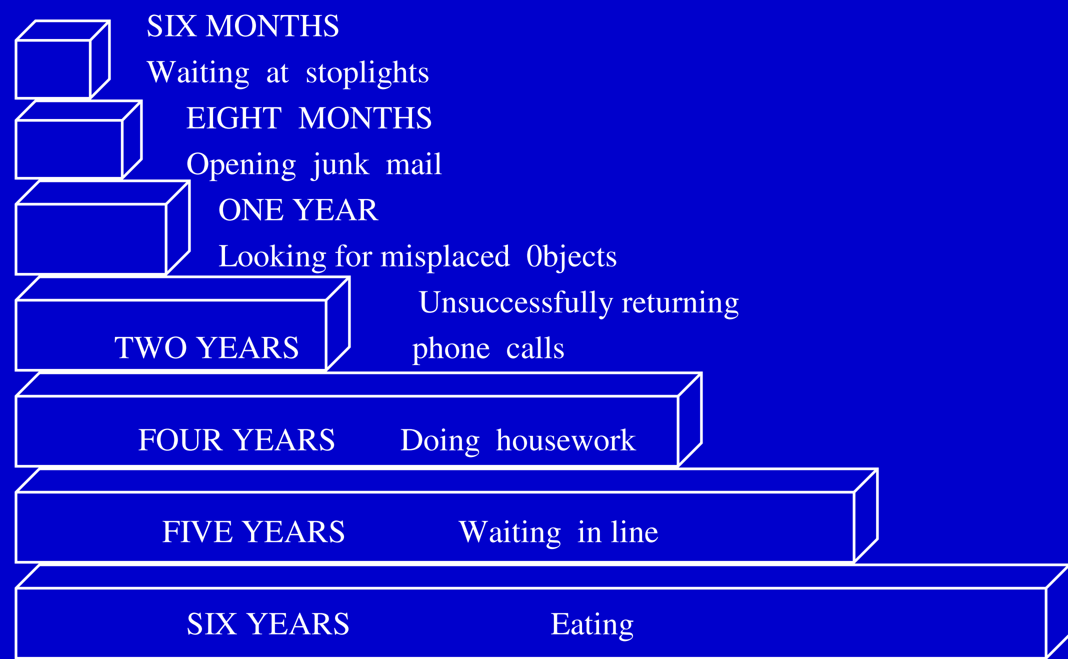
Prof. John W. Sutherland

March 20, 2006

Service Processes & Systems
Dept. of Mechanical Engineering - Engineering Mechanics
Michigan Technological University

Where the Time Goes

In a life time, the average
American will spend--



Service Processes & Systems
Dept. of Mechanical Engineering - Engineering Mechanics
Michigan Technological University

Waiting Realities

- **Inevitability of Waiting:** Waiting results from variations in arrival rates and service rates

- **Economics of Waiting:** High utilization purchased at the price of customer waiting. Make waiting productive (salad bar) or profitable (drinking bar).

Laws of Service

- **Maister's First Law:**
  - Customers compare expectations with perceptions.
- **Maister's Second Law:**
  - Hard to play catch-up ball.
- **Skinner's Law:**
  - The other line always moves faster.
- **Jenkin's Corollary:**
  - When you switch to another line, the line you left moves faster.

Remember Me

- I am the person who goes into a restaurant, sits down, and patiently waits while the wait-staff does everything but take my order.
- I am the person that waits in line for the clerk to finish chatting with his buddy.
- I am the one who never comes back and it amuses me to see money spent to get me back.
- I was there in the first place, all you had to do was show me some courtesy and service.

The Customer

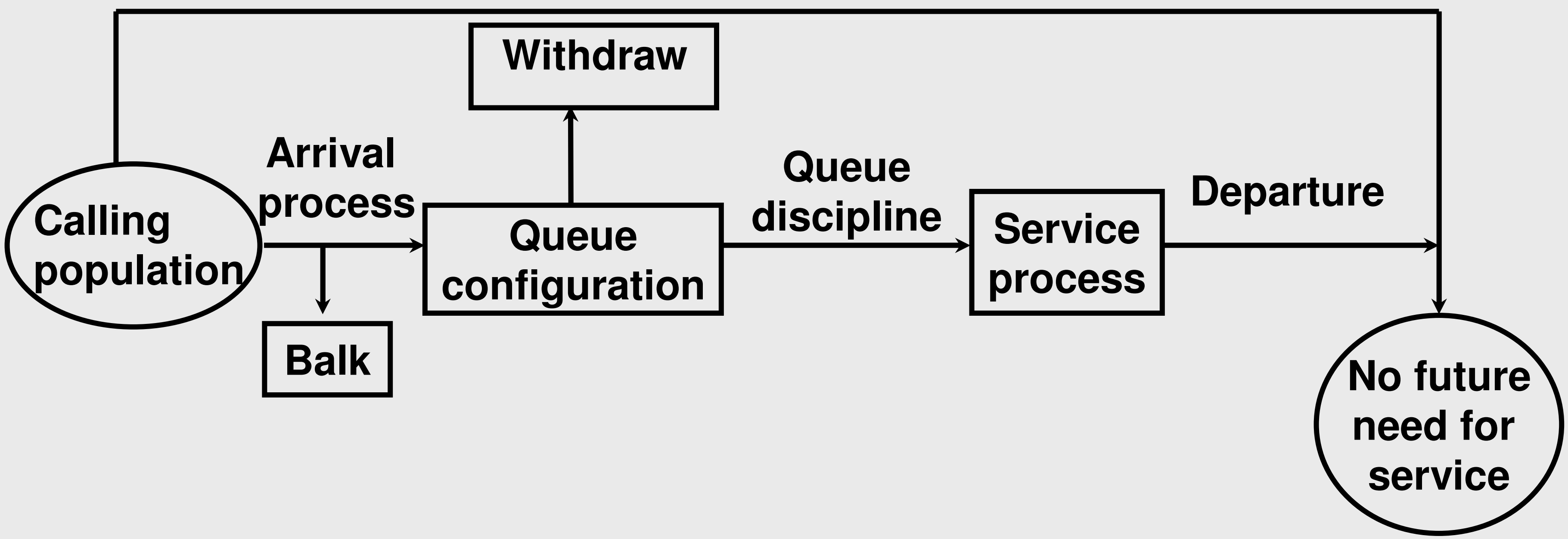
Psychology of Waiting

- That Old Empty Feeling: Unoccupied time goes slowly
- A Foot in the Door: Pre-service waits seem longer that in-service waits
- The Light at the End of the Tunnel: Reduce anxiety with attention
- Excuse Me, But I Was First: Social justice with FCFS queue discipline
- They Also Serve, Who Sit and Wait: Avoids idle service capacity

Controlling Customer Waiting

- Animate: Disneyland distractions, elevator mirror, recorded music
- Discriminate: Avis frequent renter treatment (out of sight)
- Automate: Use computer scripts to address 75% of questions
- Obfuscate: Disneyland staged waits (e.g., Haunted Mansion)

Features of Queuing Systems

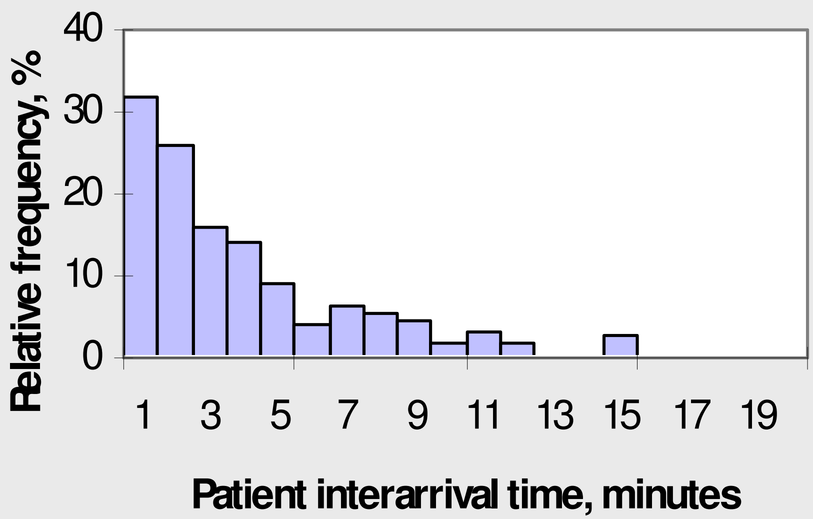


Arrival Process

- Any analysis of a service system must begin with a complete understanding of the temporal and spatial distribution of the demand for that service.
- Typically, data are collected by recording the actual times of arrivals.
- These data then are used to calculate interarrival times.

Arrival Process

- Many empirical studies indicate that the distribution of interarrival times will be exponential



Distribution of Patient Interarrival Times for a University Health Clinic.

Exponential Distribution

- It has a continuous probability density function of the form
  - $f(t) = \lambda e^{-\lambda t} \quad t \geq 0$
  - where λ = average arrival rate within a given interval of time
  - t = time between arrivals
- The cumulative distribution function is
  - $F(t) = 1 - e^{-\lambda t} \quad t \geq 0$
  - This equation gives the probability that the time between arrivals will be t or less.

Poisson Distribution

- It has a unique relationship to the exponential distribution.
- It is a discrete probability function of the form
  - $f(n) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}$ $n = 0,1,2,3,\ldots$

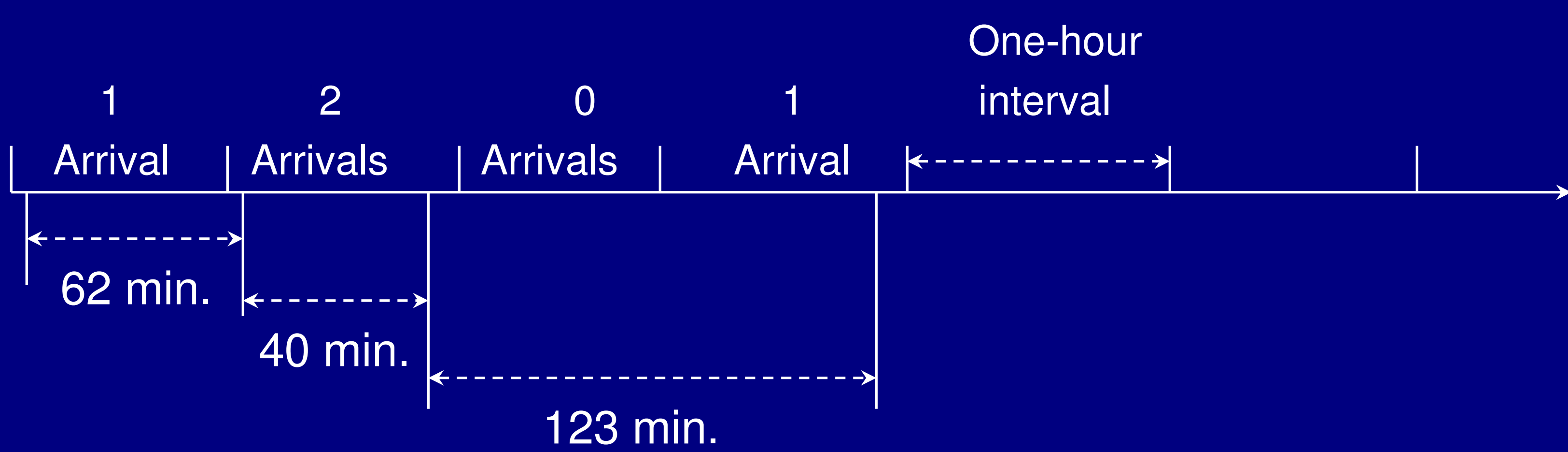
Where λ = average arrival rate within a given interval of time

t = number of time periods of interest

n = number of arrivals (0,1,2,…)

Poisson – Exponential Relation

Poisson distribution for number of arrivals per hour (top view)



Exponential distribution of time between arrivals in minutes (bottom view)

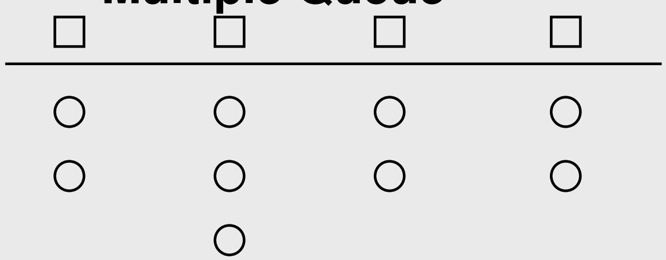
Dept. of Mechanical Engineering - Engineering Mechanics
Michigan Technological University

Queue Configuration

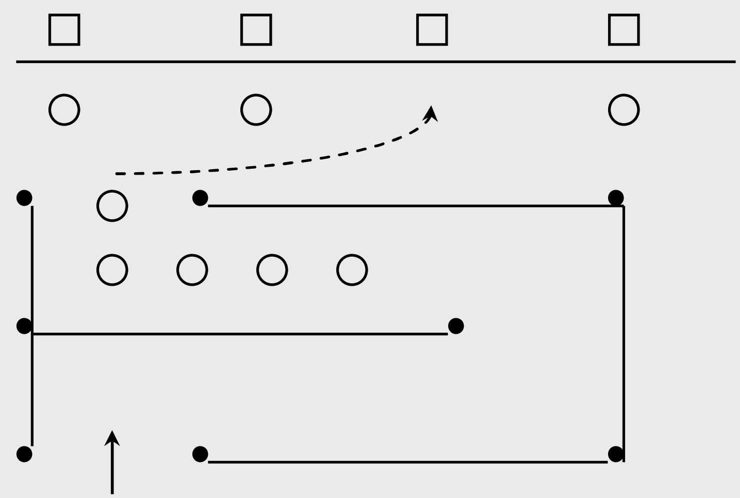
- It refers to the number of queues, their locations, their spatial requirements, and their effects on customer behavior.

Queue Configuration

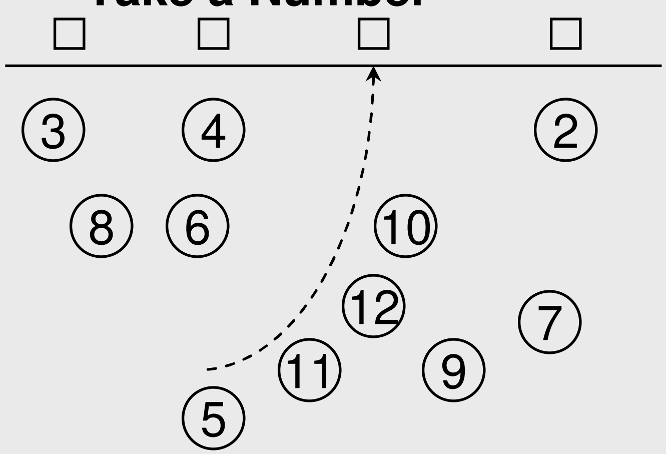
Multiple Queue



Single queue

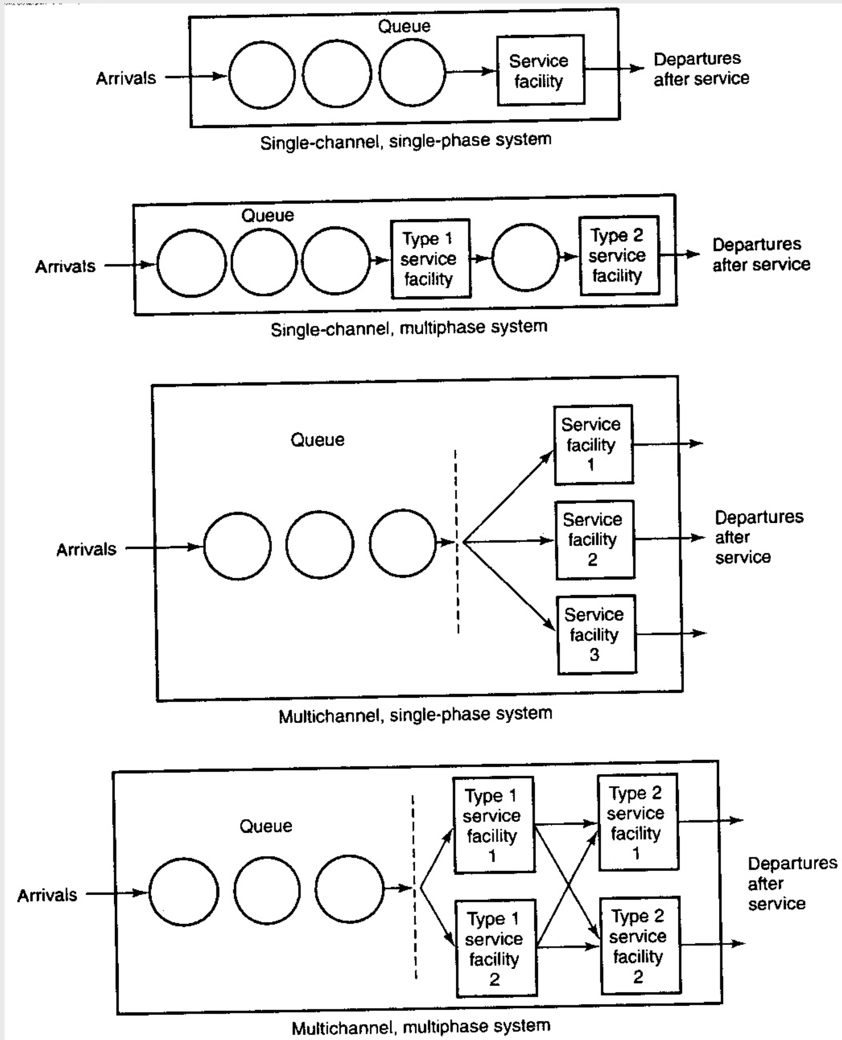


Take a Number



Fitzsimmons, 2006

Other Queuing Representations



Haksever, 2000

Queue Discipline

- It is a policy established by management to select the next customer from the queue for service.
- The most popular service discipline is the first-come, first-served (FCFS/FIFO) rule.
  - LIFO rule
- This represents "fairest" approach to serving waiting customers, because all customer are treated alike
- This rule is considered to be static because no information other than position in line is used to identify the next customer for service

Queue Discipline

- Dynamic queue disciplines are based on some attribute of the customer or status of the waiting line.
  - For example, computer installations typically give first priority to waiting jobs with very short processing times.
  - This shortest-processing-time (SPT) rule has the important feature of minimizing the average time that a customer spends in the system.

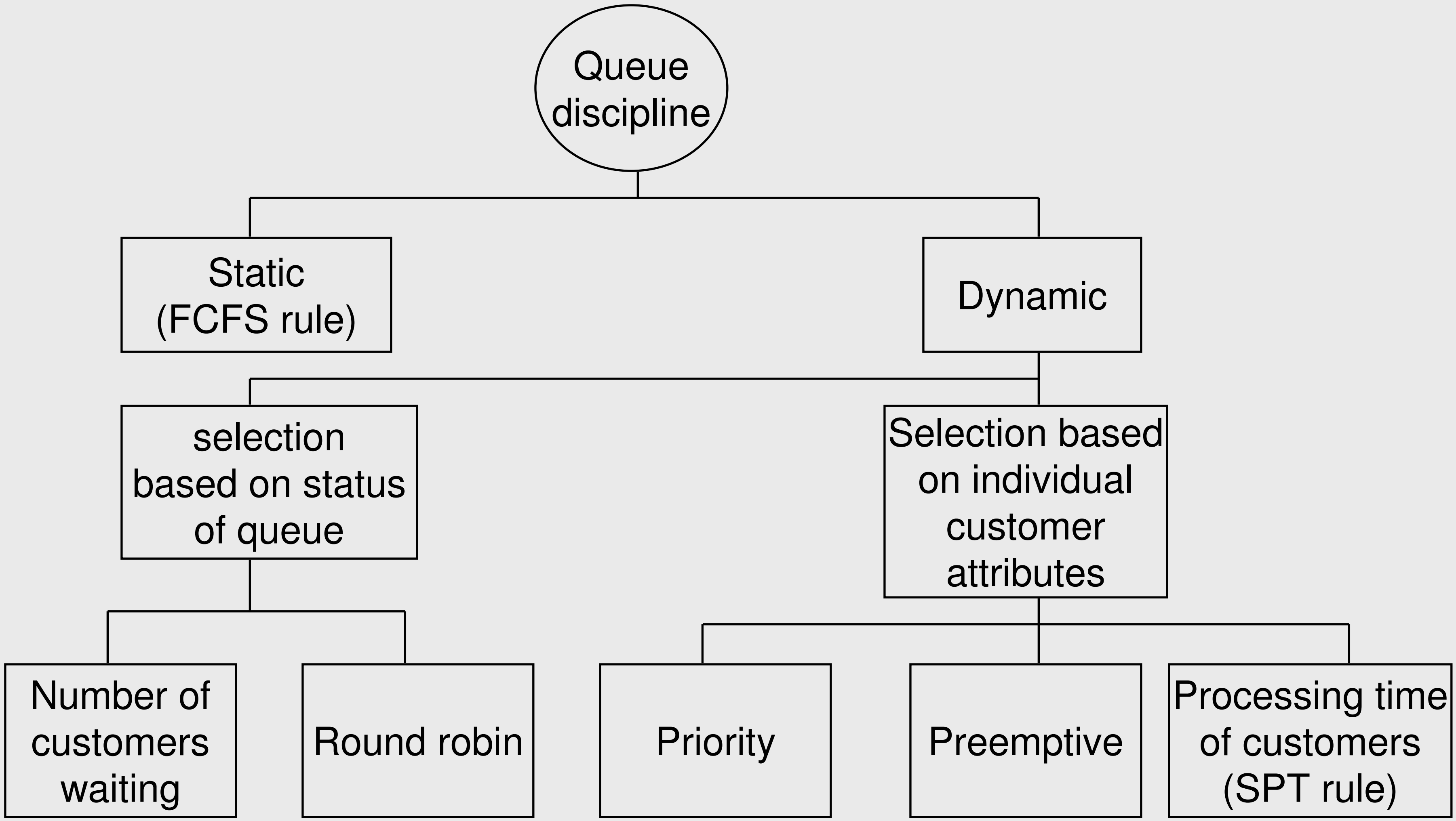
Queue Discipline

- Typically, arrivals are placed in priority classes on the basis of some attribute, and the FCFS rule is used within each class.
  - An example is the express check-out counter at supermarkets, where orders of 10 or fewer items are processed.
  - In a medical setting, the procedure known as triage is used to give priority to those who would benefit most from immediate treatment.

Queue Discipline

- The most responsive queue discipline is the preemptive priority rule.
  - Under this rule, the service currently in process for a person is interrupted to serve a newly arrived customer with higher priority. This rule usually is reserved for emergency services – e.g., fire or ambulance service.

Queue Discipline



Service Facility Arrangements

Service facility	Server arrangement
Parking lot	Self-serve
Cafeteria	Servers in series
Toll booths	Servers in parallel
Supermarket	Self-serve, first stage; parallel servers, second stage
Hospital	Many service centers in parallel and series, not all used by each patient

Waiting-Line Performance Measures

- Average queue time, W_q
- Average queue length, L_q
- Average time in system, W_s
- Average number in system, L_s
- Probability of idle service facility, P_0
- System utilization, ρ
- Probability of k units in system, $P_{n > k}$

Assumptions of the Basic Simple Queuing Model

- Arrivals are served on a first come, first served basis
- Arrivals are independent of preceding arrivals
- Arrival rates are described by the Poisson probability distribution, and customers come from a very large population
- Service times vary from one customer to another, and are independent of one and other; the average service time is known
- Service times are described by the negative exponential probability distribution
- The service rate is greater than the arrival rate

Types of Queuing Models

- Simple (M/M/1)
  - Example: Information booth at mall
- Multi-channel (M/M/S)
  - Example: Airline ticket counter
- Constant Service (M/D/1)
  - Example: Automated car wash
- Limited Population
  - Example: Department with only 7 drills

Simple (M/M/1) Model Characteristics

- Type: Single-channel, single-phase system
- Input source: Infinite; no balks, no reneging
- Arrival distribution: Poisson
- Queue: Unlimited; single line
- Queue discipline: FIFO (FCFS)
- Service distribution: Negative exponential
- Relationship: Independent service & arrival
- Service rate > arrival rate

Simple (M/M/1) Model Equations

Average number of units in queue $$L_s = \frac{\lambda}{\mu - \lambda}$$

Average time in system $$W_s = \frac{1}{\mu - \lambda}$$

Average number of units in queue $$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

Average time in queue $$W_q = \frac{\lambda}{\mu(\mu - \lambda)}$$

System utilization $$\rho = \frac{\lambda}{\mu}$$

Simple (M/M/1) Probability Equations

Probability of 0 units in system, i.e., system idle:

$$P_0 = 1 - \rho = 1 - \frac{\lambda}{\mu}$$

Probability of more than k units in system:

$$P_{n>k} = \left(\frac{\lambda}{\mu}\right)^{k+1}$$

Where n is the number of units in the system

Remember: λ & μ Are Rates

- λ = Mean number of arrivals per time period
  - e.g., 3 units/hour
- μ = Mean number of people or items served per time period
  - e.g., 4 units/hour
    - $1/\mu$ = 15 minutes/unit