Waiting-Line Models

21

PowerPoint presentation to accompany
Heizer, Render, Munson
Operations Management, Twelfth Edition, Global Edition
Principles of Operations Management, Tenth Edition, Global Edition

PowerPoint slides by Jeff Heyl

Outline

- Queuing Theory
- Characteristics of a Waiting-Line System
- Queuing Costs
- The Variety of Queuing Models
- Other Queuing Approaches

Learning Objectives

When you complete this chapter you should be able to:

D.1 ***Describe*** the characteristics of arrivals, waiting lines, and service systems

D.2 ***Apply*** the single-server queuing model equations

D.3 ***Conduct*** a cost analysis for a waiting line

Learning Objectives

When you complete this chapter you should be able to:

- **D.4** ***Apply*** the multiple-server queuing model formulas
- **D.5** ***Apply*** the constant-service-time model equations
- **D.6** ***Perform*** a finite-population model analysis

Queuing Theory

- The study of waiting lines
- Waiting lines are common situations
- Useful in both manufacturing and service areas

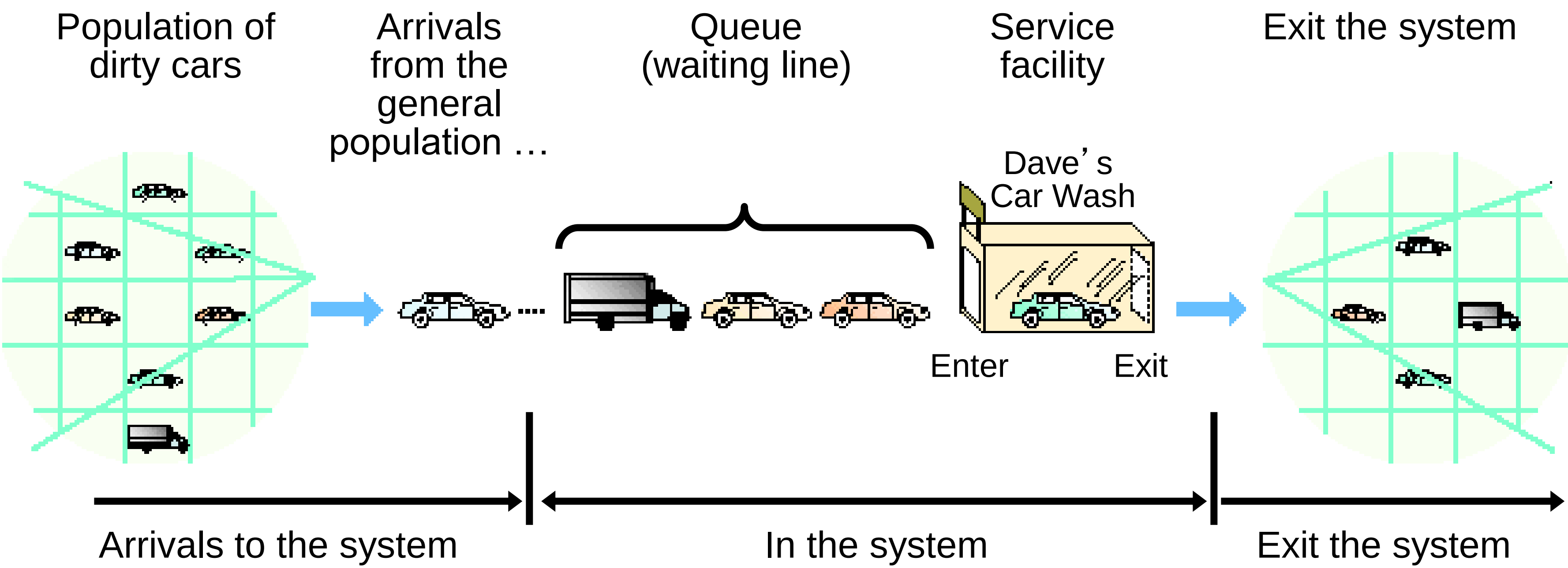
Common Queuing Situations

TABLE D.1	Common Queuing Situations	
SITUATION	**ARRIVALS IN QUEUE**	**SERVICE PROCESS**
Supermarket	Grocery shoppers	Checkout clerks at cash register
Highway toll booth	Automobiles	Collection of tolls at booth
Doctor's office	Patients	Treatment by doctors and nurses
Computer system	Programs to be run	Computer processes jobs
Telephone company	Callers	Switching equipment forwards calls
Bank	Customer	Transactions handled by teller
Machine maintenance	Broken machines	Repair people fix machines
Harbor	Ships and barges	Dock workers load and unload

Characteristics of Waiting-Line Systems

1. *Arrivals or inputs to the system*
   - Population size, behavior, statistical distribution
2. *Queue discipline, or the waiting line itself*
   - Limited or unlimited in length, discipline of people or items in it
3. *The service facility*
   - Design, statistical distribution of service times

Parts of a Waiting Line



Arrival Characteristics
- Size of the population
- Behavior of arrivals
- Statistical distribution of arrivals

Waiting-Line Characteristics
- Limited vs. unlimited
- Queue discipline

Service Characteristics
- Service design
- Statistical distribution of service

Figure D.1

Arrival Characteristics

1. *Size* of the arrival population
   - **Unlimited** (**infinite**) or **limited** (**finite**)
2. *Pattern* of arrivals
   - Scheduled or random, often a **Poisson distribution**
3. *Behavior* of arrivals
   - Wait in the queue and do not switch lines
   - No *balking* or *reneging*

Poisson Distribution

$$P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \quad \text{for } x = 0, 1, 2, 3, 4, \ldots$$

where $P(x)$ = probability of x arrivals

x = number of arrivals per unit of time

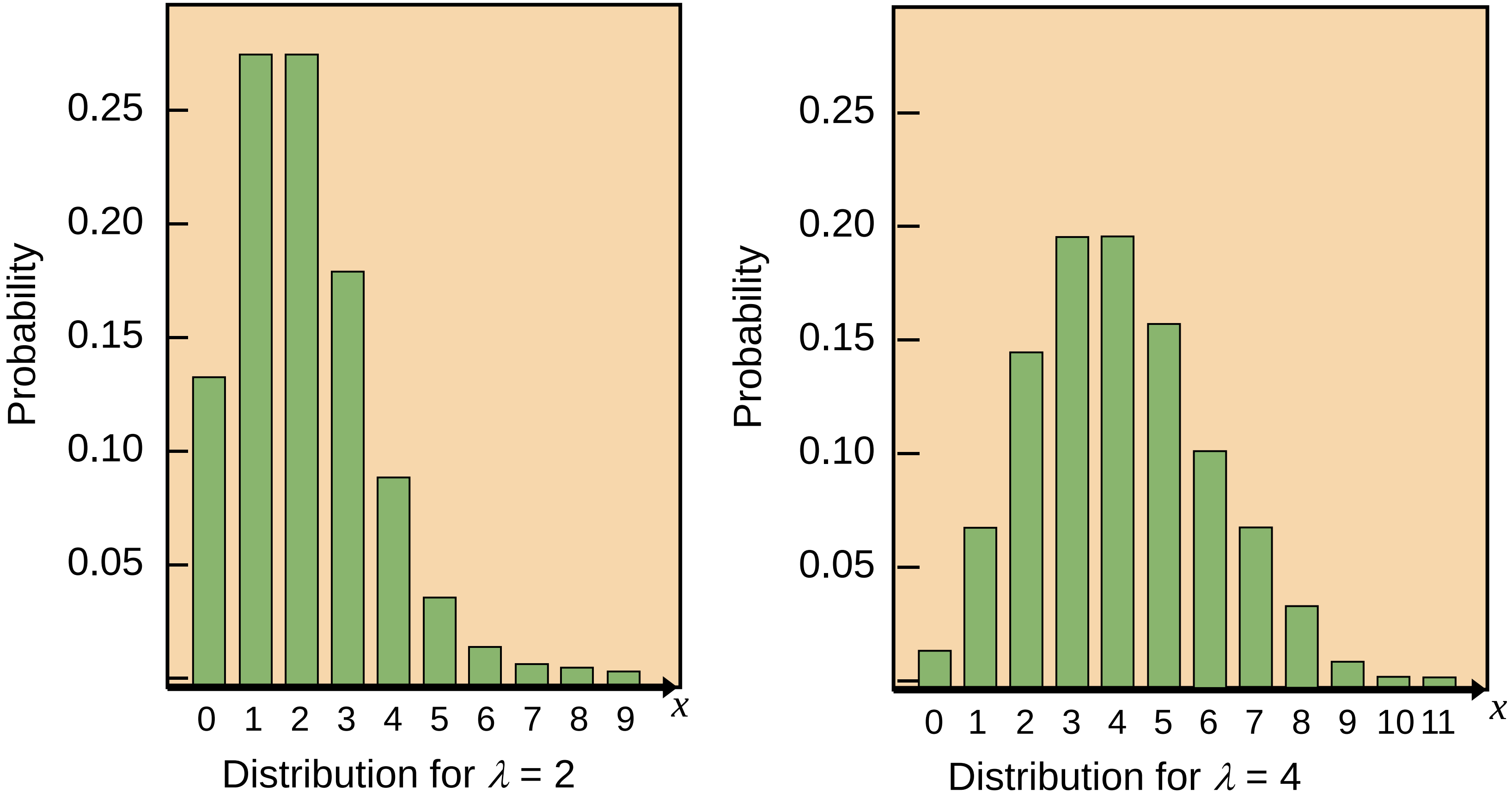
λ = average arrival rate

e = 2.7183 (which is the base of the natural logarithms)

Poisson Distribution

Figure D.2

$$\text{Probability} = P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$



Waiting-Line Characteristics

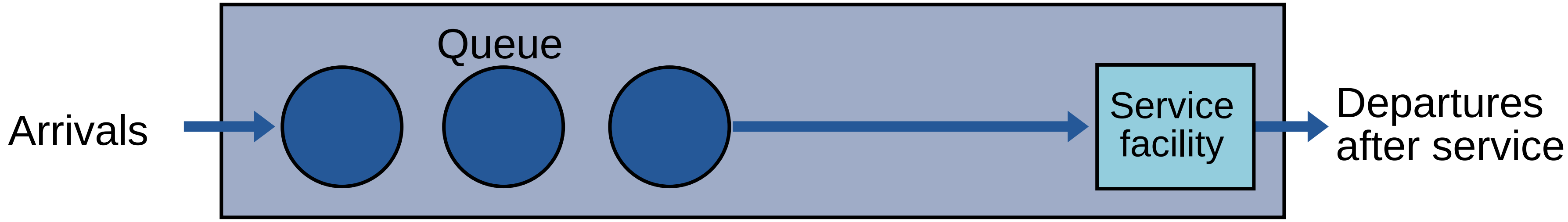
- *Limited* or *unlimited* queue length
- *Queue discipline* - **first-in, first-out (FIFO)** is most common
- Other priority rules may be used in special circumstances

Service Characteristics

1. Queuing system designs
   - **Single-server system**, **multiple-server system**
   - **Single-phase system**, **multiphase system**
2. Service time distribution
   - Constant service time
   - Random service times, usually a **negative exponential distribution**

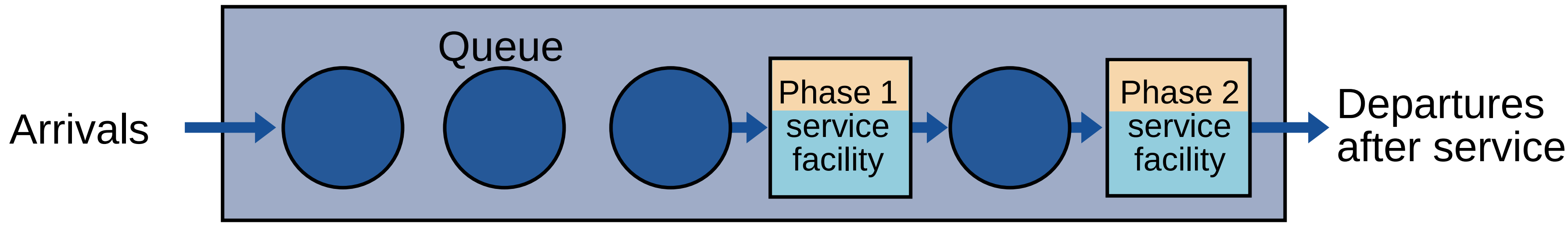
Queuing System Designs

A family dentist's office



Single-server, single-phase system

A McDonald's dual-window drive-through

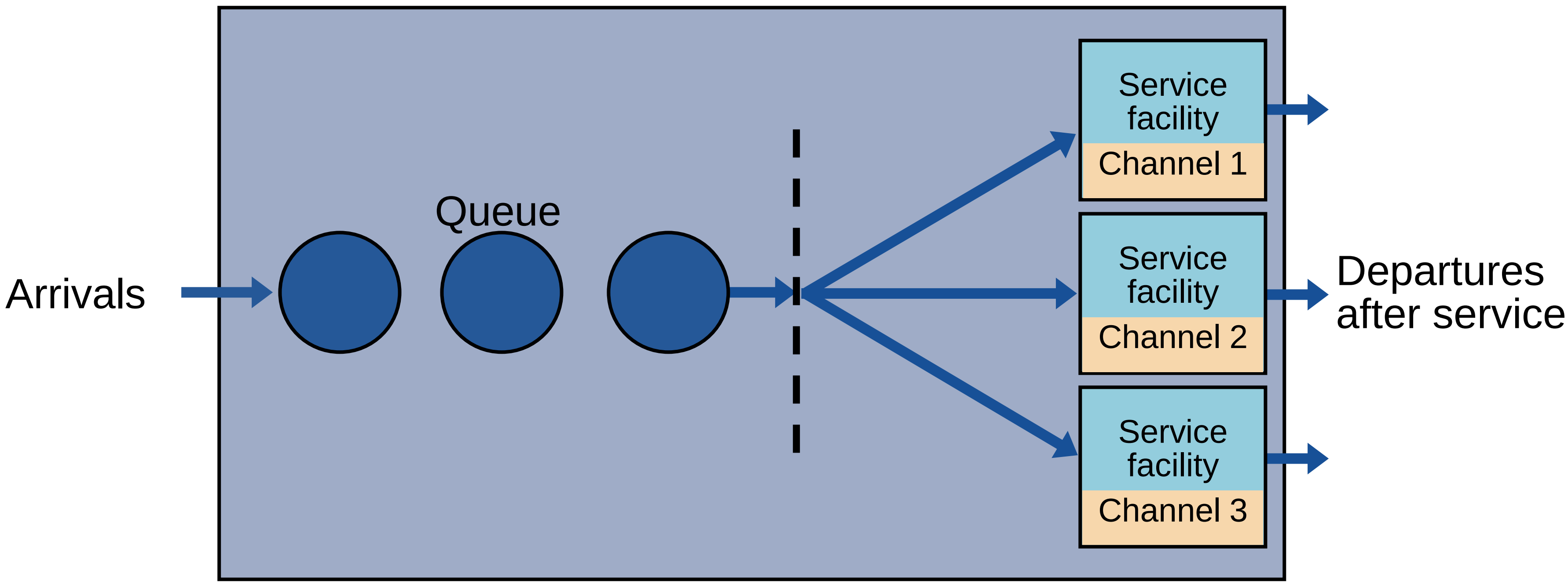


Single-server, multiphase system

Figure D.3

Queuing System Designs

Most bank and post office service windows

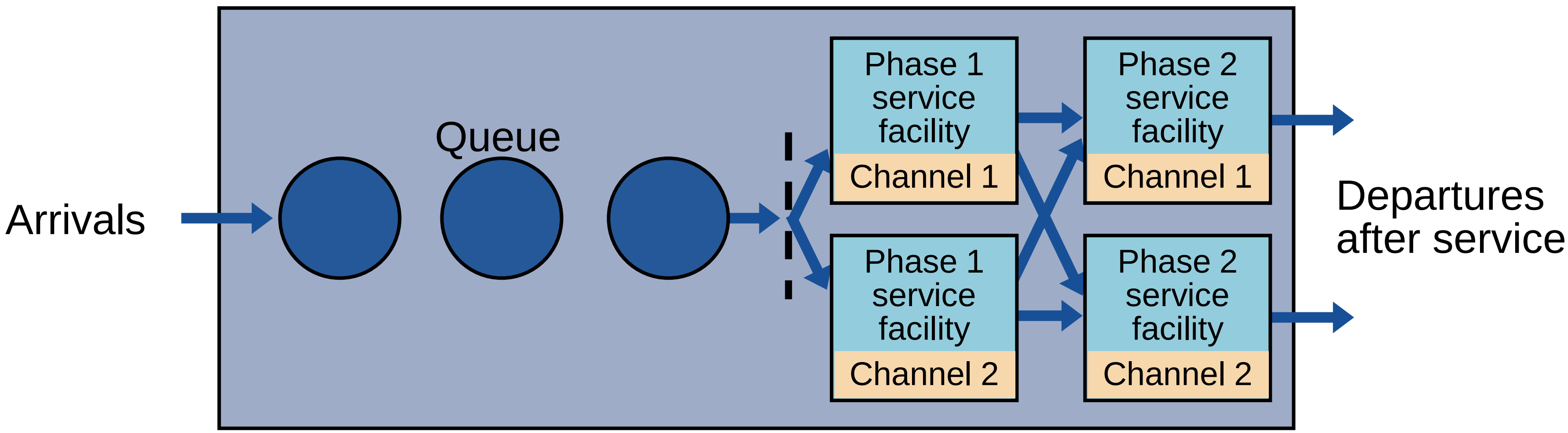


Multi-server, single-phase system

Figure D.3

Queuing System Designs

Some college cafeterias

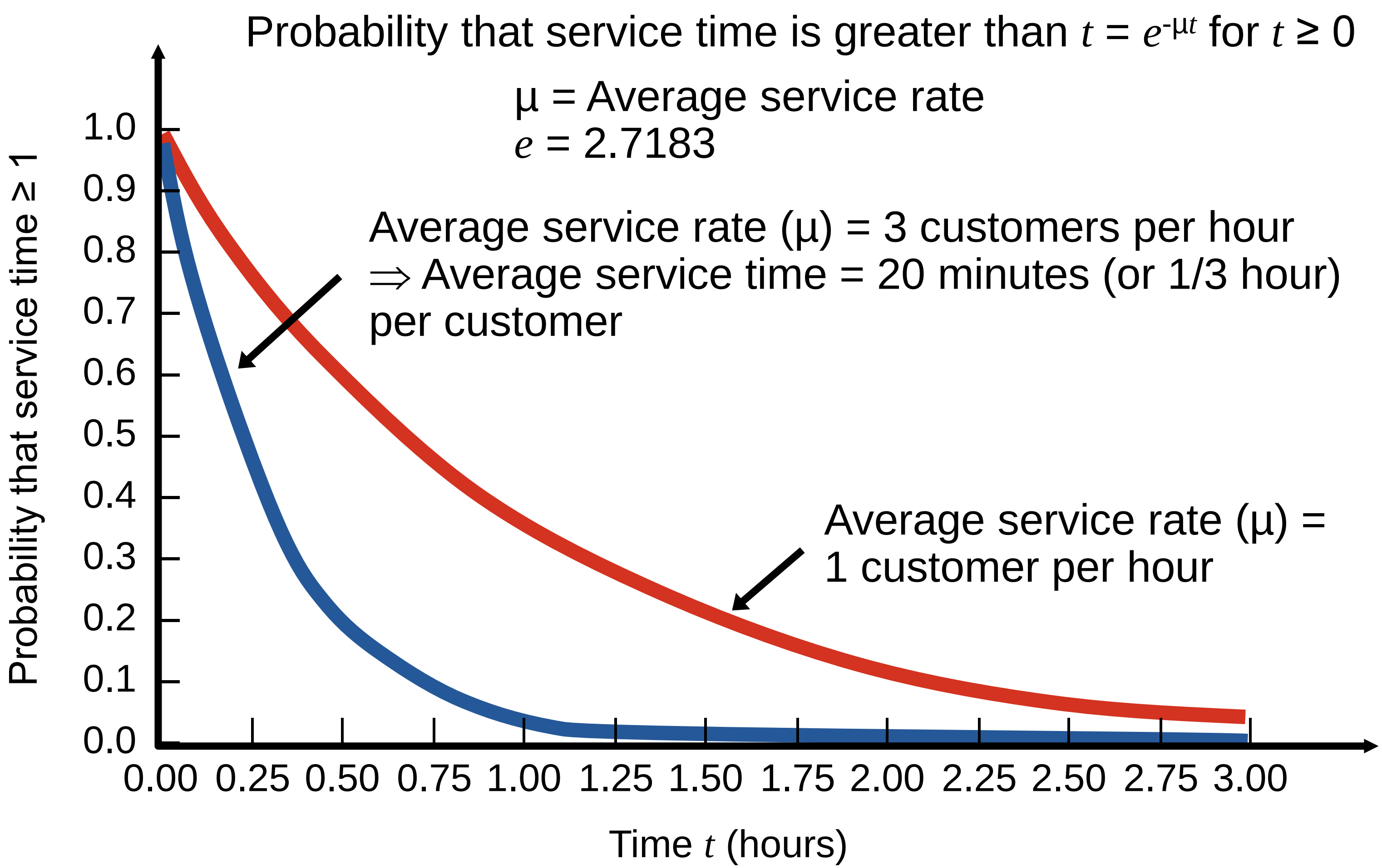


Multi-server, multiphase system

Figure D.3

Negative Exponential Distribution

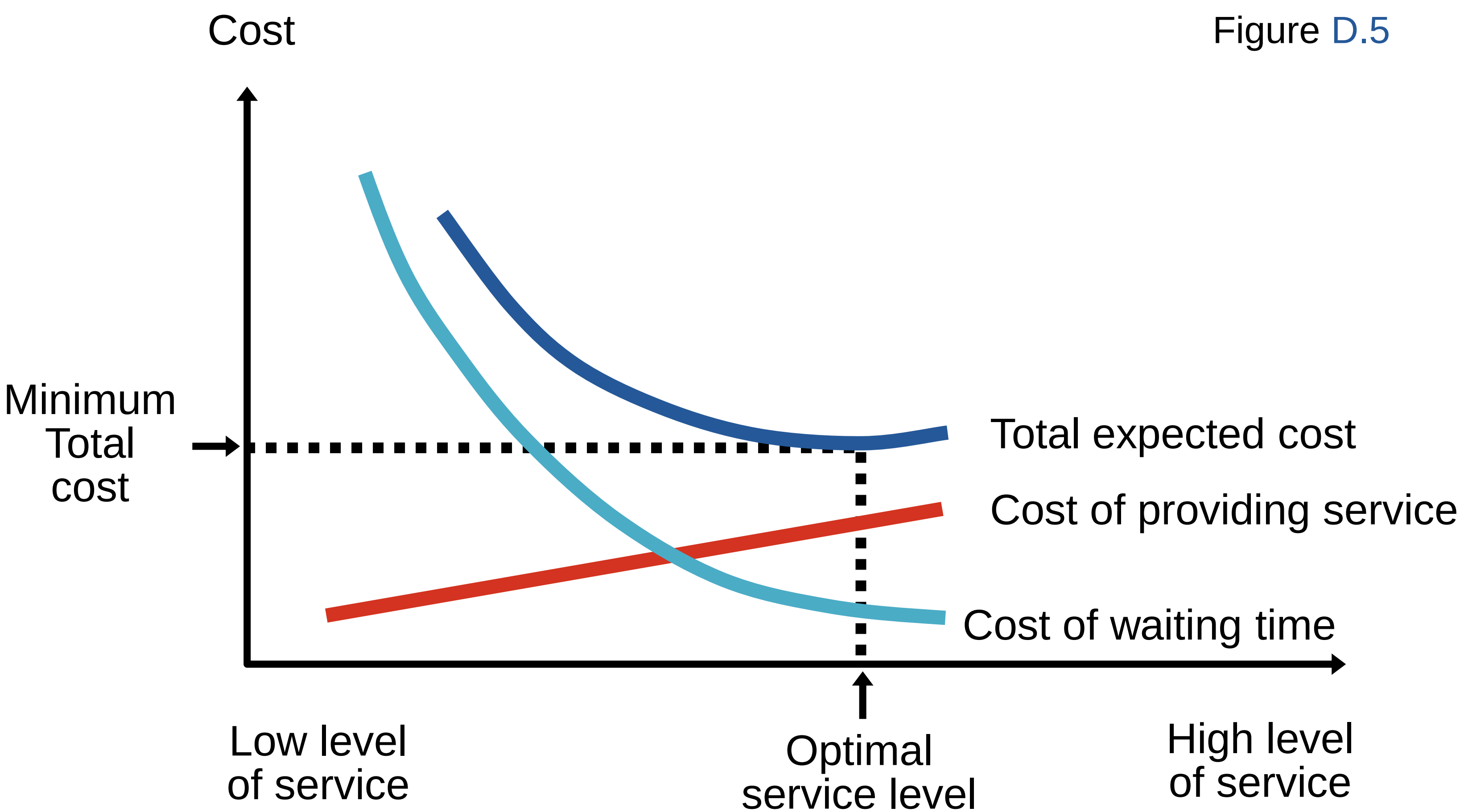
Figure D.4



Measuring Queue Performance

1. Average time that each customer or object spends in the queue
2. Average queue length
3. Average time each customer spends in the system
4. Average number of customers in the system
5. Probability that the service facility will be idle
6. Utilization factor for the system
7. Probability of a specific number of customers in the system

Queuing Costs



Queuing Models

The four queuing models that follow all assume:

1. Poisson distribution arrivals
2. FIFO discipline
3. A single-service phase

Queuing Models

TABLE D.2	Queuing Models Described in This Chapter	
MODEL	**NAME**	**EXAMPLE**
A	Single-server system (M/M/1)	Information counter at department store

NUMBER OF SERVERS (CHANNELS)	NUMBER OF PHASES	ARRIVAL RATE PATTERN	SERVICE TIME PATTERN	POPULATION SIZE	QUEUE DISCIPLINE
Single	Single	Poisson	Negative exponential	Unlimited	FIFO

Queuing Models

TABLE D.2	Queuing Models Described in This Chapter	
MODEL	**NAME**	**EXAMPLE**
B	Multiple-server (M/M/S)	Airline ticket counter

NUMBER OF SERVERS (CHANNELS)	NUMBER OF PHASES	ARRIVAL RATE PATTERN	SERVICE TIME PATTERN	POPULATION SIZE	QUEUE DISCIPLINE
Multi-server	Single	Poisson	Negative exponential	Unlimited	FIFO

Queuing Models

TABLE D.2	Queuing Models Described in This Chapter	
MODEL	**NAME**	**EXAMPLE**
C	Constant-service (M/D/1)	Automated car wash

NUMBER OF SERVERS (CHANNELS)	NUMBER OF PHASES	ARRIVAL RATE PATTERN	SERVICE TIME PATTERN	POPULATION SIZE	QUEUE DISCIPLINE
Single	Single	Poisson	Constant	Unlimited	FIFO

Queuing Models

TABLE D.2	Queuing Models Described in This Chapter	
MODEL	**NAME**	**EXAMPLE**
D	Finite population (M/M/1 with finite source)	Shop with only a dozen machines that might break

NUMBER OF SERVERS (CHANNELS)	NUMBER OF PHASES	ARRIVAL RATE PATTERN	SERVICE TIME PATTERN	POPULATION SIZE	QUEUE DISCIPLINE
Single	Single	Poisson	Negative exponential	Limited	FIFO

Model A – Single-Server

1. Arrivals are served on a FIFO basis, every arrival waits to be served regardless of the length of the queue
2. Arrivals are independent of preceding arrivals, the average number of arrivals does not change over time
3. Arrivals are described by a Poisson probability distribution and come from an infinite population

Model A – Single-Server

4. Service times vary from one customer to the next and are independent of one another, but their average rate is known
5. Service times occur according to the negative exponential distribution
6. The service rate is faster than the arrival rate

Model A – Single-Server

TABLE D.3	Queuing Formulas for Model A: Single-Server System, also Called M/M/1
λ =	average number of arrivals per time period
μ =	average number of people or items served per time period (average service rate)
L_s =	average number of units (customers) in the system (waiting and being served)
=	$\dfrac{\lambda}{\mu - \lambda}$
W_s =	average time a unit spends in the system (waiting time plus service time)
=	$\dfrac{1}{\mu - \lambda}$

Model A – Single-Server

TABLE D.3	Queuing Formulas for Model A: Single-Server System, also Called M/M/1
L_q	= average number of units waiting in the queue
	$= \dfrac{\lambda^2}{\mu(\mu - \lambda)}$
W_q	= average time a unit spends waiting in the queue
	$= \dfrac{\lambda}{\mu(\mu - \lambda)} = \dfrac{L_q}{\lambda}$
ρ	= utilization factor for the system
	$= \dfrac{\lambda}{\mu}$

Model A – Single-Server

TABLE D.3	Queuing Formulas for Model A: Single-Server System, also Called M/M/1
P_0	= Probability of 0 units in the system (that is, the service unit is idle) $= 1 - \frac{\lambda}{\mu}$
$P_{n>k}$	= probability of more than k units in the system, where n is the number of units in the system $= \left[\frac{\lambda}{\mu}\right]^{k+1}$

Single-Server Example

λ = 2 cars arriving/hour μ = 3 cars serviced/hour

$$L_s = \frac{\lambda}{\mu - \lambda} = \frac{2}{3-2} = 2 \text{ cars in the system on average}$$

$$W_s = \frac{1}{\mu - \lambda} = \frac{1}{3-2} = 1 \text{ hour average waiting time in the system}$$

$$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{2^2}{3(3-2)} = 1.33 \text{ cars waiting in line}$$

Single-Server Example

λ = 2 cars arriving/hour μ = 3 cars serviced/hour

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)} = \frac{2}{3(3-2)}$$ = 2/3 hour = 40 minute average waiting time

$$\rho = \frac{\lambda}{\mu} = \frac{2}{3}$$ = 66.6% of time mechanic is busy

$$P_0 = 1 - \frac{\lambda}{\mu}$$ = .33 probability there are 0 cars in the system

Single-Server Example

Probability of more than *k* Cars in the System

K	$P_{n>k} = (2/3)^{k+1}$
0	.667 ← Note that this is equal to 1 – P_0 = 1 – .33
1	.444
2	.296
3	.198 ← Implies that there is a 19.8% chance that more than 3 cars are in the system
4	.132
5	.088
6	.058
7	.039

Single-Channel Economics

Customer dissatisfaction and lost goodwill = \$15 per hour

W_q = 2/3 hour

Total arrivals = 16 per day

Mechanic's salary = \$88 per day

Total hours customers spend waiting per day $= \frac{2}{3}(16) = 10\frac{2}{3}$ hours

Customer waiting-time cost $= \$15\left(10\frac{2}{3}\right) = \$160$ per day

Total expected costs = \$160 + \$88 = \$248 per day

Multiple-Server Model

TABLE D.4 Queuing Formulas for Model B: Multiple-Server System, also Called M/M/S

M = number of servers (channels) open

λ = average arrival rate

μ = average service rate at each server (channel)

The probability that there are zero people or units in the system is:

$$P_0 = \frac{1}{\left[\sum_{n=0}^{M-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n\right] + \frac{1}{M!}\left(\frac{\lambda}{\mu}\right)^M \frac{M\mu}{M\mu - \lambda}} \quad \text{for } M\mu > \lambda$$

Multiple-Server Model

TABLE D.4	Queuing Formulas for Model B: Multiple-Server System, also Called M/M/S

The average number of people or units in the system is:

$$L_S = \frac{\lambda\mu(\lambda/\mu)^M}{(M-1)!(M\mu-\lambda)^2}P_0 + \frac{\lambda}{\mu}$$

The average time a unit spends in the waiting line and being serviced (namely, in the system) is:

$$W_S = \frac{\mu(\lambda/\mu)^M}{(M-1)!(M\mu-\lambda)^2}P_0 + \frac{1}{\mu} = \frac{L_S}{\lambda}$$

Multiple-Server Model

TABLE D.4	Queuing Formulas for Model B: Multiple-Server System, also Called M/M/S

The average number of people or units in line waiting for service is:

$$L_q = L_S - \frac{l}{m}$$

The average time a person or unit spends in the queue waiting for service is:

$$W_q = W_S - \frac{1}{m} = \frac{L_q}{l}$$

Multiple-Server Example

$\lambda = 2 \qquad \mu = 3 \qquad M = 2$

$$P_0 = \frac{1}{\left[\sum_{n=0}^{1} \frac{1}{n!}\left(\frac{2}{3}\right)^n\right] + \frac{1}{2!}\left(\frac{2}{3}\right)^2 \frac{2(3)}{2(3) - 2}}$$

$$= \frac{1}{1 + \frac{2}{3} + \frac{1}{2}\left(\frac{4}{9}\right)\left(\frac{6}{6-2}\right)} = \frac{1}{1 + \frac{2}{3} + \frac{1}{3}} = \frac{1}{2}$$

$= .5$ probability of zero cars in the system

Multiple-Server Example

$$L_S = \frac{(2)(3)\left(2/3\right)^2}{\left(1\right)!\left(2(3)-2\right)^2}\left(\frac{1}{2}\right) + \frac{2}{3} = \frac{8/3}{16}\left(\frac{1}{2}\right) + \frac{2}{3} = \frac{3}{4}$$

$$= .75 \text{ } \textit{average number of cars in the system}$$

$$W_S = \frac{L_S}{\lambda} = \frac{3/4}{2} = \frac{3}{8} \text{ } \textit{hour}$$

$$= 22.5 \text{ } \textit{minutes average time a car spends in the system}$$

$$L_q = L_S - \frac{\lambda}{\mu} = \frac{3}{4} - \frac{2}{3} = \frac{9}{12} - \frac{8}{12} = \frac{1}{12}$$

$$= .083 \text{ } \textit{average number of cars in the queue (waiting)}$$

$$W_q = \frac{L_q}{\lambda} = \frac{.083}{2} = .0415 \text{ } \textit{hour}$$

$$= 2.5 \text{ } \textit{minutes average time a car spends in the queue (waiting)}$$

Multiple-Server Example

	SINGLE SERVER	TWO SERVERS (CHANNELS)
P_0	.33	.5
L_s	2 cars	.75 cars
W_s	60 minutes	22.5 minutes
L_q	1.33 cars	.083 cars
W_q	40 minutes	2.5 minutes

Waiting Line Tables

TABLE D.5	Values of L_q for M = 1-5 Servers (channels) and Selected Values of λ/μ				
POISSON ARRIVALS, EXPONENTIAL SERVICE TIMES					
	NUMBER OF SERVICE CHANNELS, M				
λ/μ	**1**	**2**	**3**	**4**	**5**
.10	.0111				
.25	.0833	.0039			
.50	.5000	.0333	.0030		
.75	2.2500	.1227	.0147		
.90	8.1000	.2285	.0300	.0041	
1.0		.3333	.0454	.0067	
1.6		2.8444	.3128	.0604	.0121
2.0			.8888	.1739	.0398
2.6			4.9322	.6581	.1609
3.0				1.5282	.3541
4.0					2.2164

Waiting-Line Table Example

Bank tellers and customers

$\lambda = 18$, $\mu = 20$

Ratio $\lambda/\mu = .90$

$$W_q = \frac{L_q}{\lambda}$$

From Table D.5

NUMBER OF SERVERS	*M*	NUMBER IN QUEUE	TIME IN QUEUE
1 window	1	8.1	.45 hrs, 27 minutes
2 windows	2	.2285	.0127 hrs, ¾ minute
3 windows	3	.03	.0017 hrs, 6 seconds
4 windows	4	.0041	.0002 hrs, <1 second

Constant-Service-Time Model

TABLE D.6	Queuing Formulas for Model C: Constant Service, also Called M/D/1
Average length of queue:	$L_q = \dfrac{\lambda^2}{2\mu(\mu - \lambda)}$
Average waiting time in queue:	$W_q = \dfrac{\lambda}{2\mu(\mu - \lambda)}$
Average number of customers in the system:	$L_s = L_q + \dfrac{\lambda}{\mu}$
Average time in the system:	$W_s = W_q + \dfrac{1}{\mu}$

Constant-Service-Time Example

Trucks currently wait 15 minutes on average
Truck and driver cost \$60 per hour
Automated compactor service rate (μ) = 12 trucks per hour
Arrival rate (λ) = 8 per hour
Compactor costs \$3 per truck

Current waiting cost per trip = (1/4 hr)(\$60) = \$15 /trip

$$W_q = \frac{8}{2(12)(12-8)} = \frac{1}{12} \text{ hour}$$

Waiting cost/trip with compactor	= (1/12 hr wait)(\$60/hr cost)	= \$ 5 /trip
Savings with new equipment	= \$15 (current) – \$5(new)	= \$10 /trip
Cost of new equipment amortized		= \$ 3 /trip
Net savings		= \$ 7 /trip

Little's Law

- A queuing system in *steady state*

$$L_s = \lambda W_s \text{ (which is the same as } W_s = L_s/\lambda)$$
$$L_q = \lambda W_q \text{ (which is the same as } W_q = L_q/\lambda)$$

  - Once two of the parameters is known, the other can be easily found
  - It makes no assumptions about the probability distribution of arrival and service times
  - Applies to all queuing models except the finite population model

Little's Law Example

$\lambda = 20$ per hour $\qquad L_q = 5$

$$W_q = L_q/\lambda$$
$$= 5/20 = 0.25 \text{ hours}$$

(0.25 hours)(60 min/hour) = 15 minutes

Finite-Population Model

- Assumptions
  1. There is only one server
  2. The population of units seeking service is finite
  3. Arrivals follow a Poisson distribution, service times are negative exponentially distributed
  4. Customers are served on a first-come, first-served basis

Finite-Population Model

TABLE D.7	Queuing Formulas and Notation for Model D: Finite-Population, also called M/M/1 with Finite Source
λ = average arrival rate μ = average service rate N = size of population	Average waiting time in the queue: $$W_q = \frac{L_q}{(N - L_s)\lambda}$$
Probability that the system is empty: $$P_0 = \frac{1}{\sum_{n=0}^{N} \frac{N!}{(N-n)!}\left(\frac{\lambda}{\mu}\right)^n}$$	Average time in the system: $$W_s = W_q + \frac{1}{\mu}$$

Finite-Population Model

TABLE D.7	Queuing Formulas and Notation for Model D: Finite-Population, also called M/M/1 with Finite Source
Average length of the queue: $$L_q = N - \left(\frac{\lambda + \mu}{\lambda}\right)(1 - P_0)$$	Probability of *n* units in the system: $$P_n = \frac{N!}{(N-n)!}\left(\frac{\lambda}{\mu}\right)^n P_0 \quad \text{for } n = 0, 1, \ldots, N$$
Average number of customers (units) in the system: $$L_s = L_q + (1 - P_0)$$	

Finite-Population Example

Laser printer breakdown analysis

λ = 1/20 = 0.05 printers/hour μ = ½ = 0.50 printers/hour

1. $$P_0 = \frac{1}{\sum_{n=0}^{5} \frac{5!}{(5-n)!} \left(\frac{0.05}{0.5} \right)^n} = 0.564$$

2. $$L_q = 5 - \left(\frac{0.05 + 0.5}{0.05} \right)(1 - P_0) = 5 - (11)(1 - 0.564) = 5 - 4.8 = 0.2 \text{ printers}$$

Finite-Population Example

Laser printer breakdown analysis

λ = 1/20 = 0.05 printers/hour μ = ½ = 0.50 printers/hour

3. $L_s = 0.2 + (1 - 0.564) = 0.64$ printers

4. $W_q = \dfrac{0.2}{(5 - 0.64)(0.05)} = \dfrac{0.2}{0.22} = 0.91$ hours

5. $W_s = 0.91 + \dfrac{1}{0.50} = 2.91$ hours

Finite-Population Example

Laser printer breakdown analysis

λ = 1/20 = 0.05 printers/hour μ = ½ = 0.50 printers/hour

3. $L_s = 0.2 + (1 - 0.564) = 0.64$ printers

4. $W_q = \frac{0.2}{(5 - 0.64)(}$

5. $W_s = 0.91 + \frac{1}{0.5}$

Total hourly cost = (Average number of printers down)(Cost per downtime hour) + Cost per technician hour

= (0.64)($120) + $25

= $76.80 + $25.00

= $101.80

Other Queuing Approaches

- The single-phase models cover many queuing situations
- Variations of the four single-phase systems are possible
- Multiphase models exist for more complex situations