BUAD 467/667
Service Management
Spring 2008

Professor Patrick T. Harker

Class 5a

Design in Action: Call Centers

Outline

- Some background on call centers
- Call center economics
- Call center management

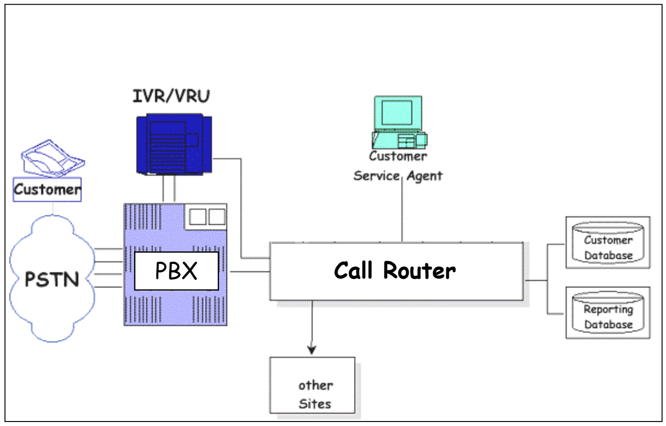
Call Centers are Big Business

- Big Dollars
  - They sold $200 billion in goods and services in 1995
  - Their services earned $80 billion in revenue in 1996
  - The market is estimated to be growing at 8% - 10% per year
- Lots of people
  - 3 million people in 30,000 call centers in the U.S.
  - that's one percent of private-sector employment
  - that's more than employment in the motor-vehicle, electronics, and aerospace industries combined

What Do Call Centers Do?

- "Outbound" Call Centers
  - telemarketing: credit cards, long distance providers
  - marketing research: all kinds of organizations
- "Inbound" Call Centers
  - customer service: banks, other services, manufacturers
  - reservations: airlines, hotels, car rentals, doctors' offices
  - help desks: computer hardware and software
- Blended inbound and outbound operations

How an Inbound Call Center Works



Source: Genesys Telecommunications Laboratories

What I See When I See an Inbound Call Center

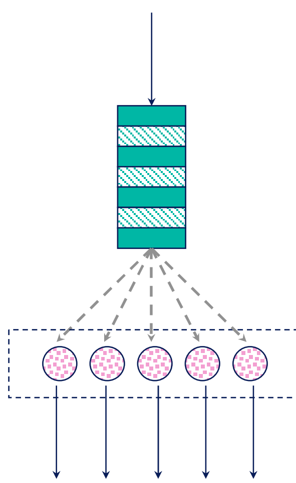
customers arrive

customers queue

customers routed to CSR's

customers served by CSR's

customers exit



Call-Center Economics: The Value Proposition

- More convenience and better service for customers
  - on the phone, rather than travel
  - any time, rather than business hours
  - get someone with a lot of experience
- Lower unit costs for suppliers
  - economies of scale for "bricks and mortar"
  - economies of scale for personnel

Personnel Economies of Scale: Example

- BankZero has 500 branches in 4 states
  - λ_{branch} = 0.2 calls per minute
    (a call arrives every 5 minutes)
  - μ_{branch} = 0.1333 calls per minute
    (7.5 minutes to handle a call)
- BankZero has a service-level target that it wishes to meet
  - Average Speed of Answer (ASA) of 20 seconds
  - what is the queueing term for "ASA"?

Base Case Staffing: 2,000 CSR's

- Suppose each branch answered its own calls
  - requires 2,000 = 4 * 500 people
  - with telephone utilization of 37.5% (they could do other things...)

Inputs:

lambda	0.2	calls / min	Ca^2	1	lambda/mu
mu	0.1333333	calls / min	Cb^2	1	1.5

Outputs:

s	Nq	Ns	Wq	Ws	P(delay)	Utilization
1	infinity	infinity	infinity	infinity	1.000000	1.000000
2	1.928571	3.428571	9.642857	17.142857	0.642857	0.750000
3	0.236842	1.736842	1.184211	8.684211	0.236842	0.500000
4	**0.044751**	**1.544751**	**0.223757**	**7.723757**	**0.074586**	**0.375000**
5	0.008631	1.508631	0.043156	7.543156	0.020139	0.300000

Regional Consolidation Saves More Than 60%

- Suppose there are four call centers, each handling 125 branches
  - $\lambda_{regional}$ = 25 = 125 * 0.2 calls / minute
  - requires 788 = 197 * 4 people
  - savings of 60.6% = (1 - 788/2,000) * 100%

Inputs:

lambda	25	calls / min	Ca^2	1	lambda/mu
mu	0.1333333	calls / min	Cb^2	1	187.5

Outputs:

s	Nq	Ns	Wq	Ws	P(delay)	Utilization
187	infinity	infinity	infinity	infinity	1.000000	1.000000
196	9.462521	196.962521	0.378501	7.878501	0.428968	0.956633
197	**7.573904**	**195.073904**	**0.302956**	**7.802956**	**0.383744**	**0.951777**
198	6.113632	193.613632	0.244545	7.744545	0.342363	0.946970

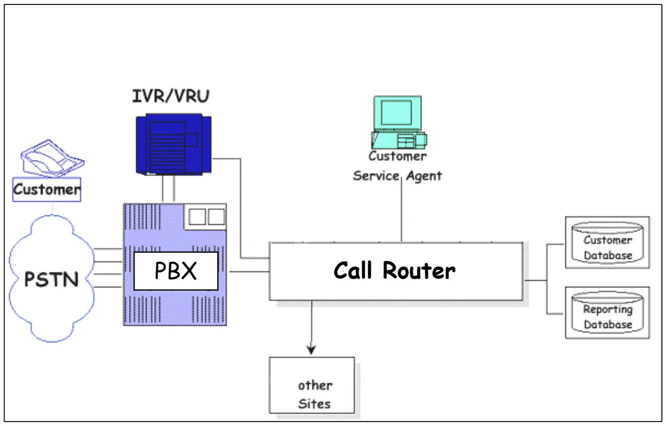
Going National Saves < 5% More Than Regional

- Suppose there is one national call center, handling all 500 branches
  - the queueing calculator bugs out!
  - can we bound the maximum savings over the regional plan?
- Maximum personnel savings is less than 5%
  - λ_{nation} = 100 calls per minute = 500 * 0.2
  - μ_{nation} = 0.1333 calls per minute (same as before...)
  - λ/μ = 750
  - need 750 people working at 100% utilization
  - savings of (1 - 750/788) * 100% = 4.8% at most

Other Factors Affecting Consolidation…

- Telecommunications
  - local infrastructure
  - PSTN charges
- Personnel
  - local wages and (un)employment rates
  - local “service ethic”
  - local language abilities and accent

Technology Allows "Virtual" Consolidation

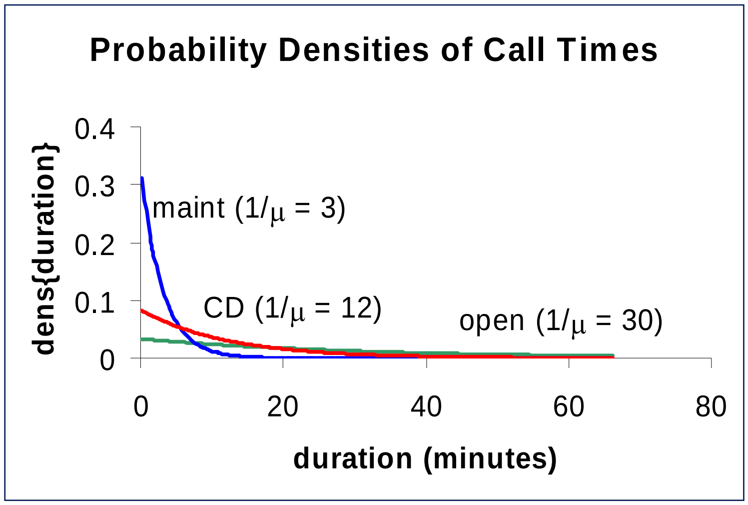


Source: Genesys Telecommunications Laboratories

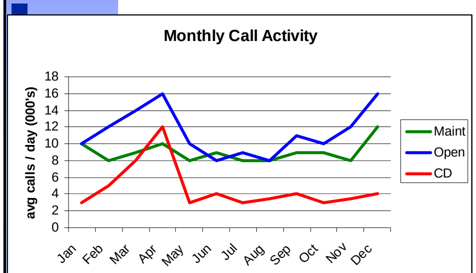
How to Run A Call Center

- Develop the capacity
  - plant: building, telecommunications, computer systems
  - people: hiring, training, coaching
- Understand the demand
  - call duration: by call type
  - call volumes: by time of day, day of week/month, time of year
- Match Supply to Demand
  - schedule CSR's to be available at the right times
  - route the calls to the "best" CSR

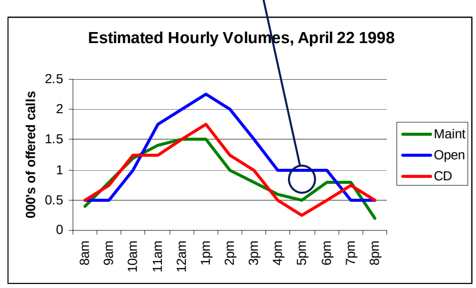
Understanding Demand: Call Duration



Understanding Demand: Call Volumes



$\lambda_{(5\text{-}5{:}15\text{pm, CD})}$ = 250 calls / hr



Also Day-of-Month Volumes

Matching Supply to Demand: Staffing

- The simplest version of the problem
  - one type of call - say CD's
  - simple work rules
    - 4 hours on, one hour break, 4 hours on
    - work 5 days in a row, then 2 off
    - start work same time every day
- Solving the problem
  - calculate minimum number of staff for every 15-minute bucket
  - find minimum-cost staffing plan

The Minimum Staff for a 15-Minute Bucket

- Example: calls about CD's from 5pm to 5:15pm Tuesday
  - λ = 250 calls/hr $\Rightarrow$ λ = 4.167 calls/min
  - $1/\mu$ = 12 min/call $\Rightarrow$ μ = 0.083 calls/min
- Minimum staff for 0.33 minute (20 second) ASA: 58

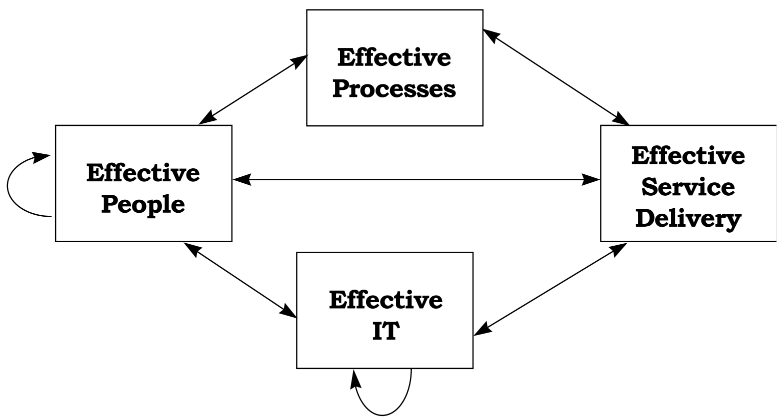
Inputs:

lambda	4.1666667	calls / min	Ca^2 1	lambda/mu
mu	0.0833333	calls / min	Cb^2 1	50

Outputs:

s	Nq	Ns	Wq	Ws	P(delay)	Utilization
50	infinity	infinity	infinity	infinity	1.000000	1.000000
51	41.986374	91.986374	10.076730	22.076730	0.839727	0.980392
57	1.760507	51.760507	0.422522	12.422522	0.246471	0.877193
58	**1.215199**	**51.215199**	**0.291648**	**12.291648**	**0.194432**	**0.862069**

Call Centers in Practice



Banking results...

- Customer-focus goes hand in hand with employee focus.
- Turnover is much more heavily influenced by work environment, rather than compensation.
- Additional capacity results in lower CSR turnover, and a greater potential for increasing customer retention.
- Outbound sales efforts can shift attention away from effective service delivery.
- IT spending needs to be closely monitored to ensure “smart design” versus increased system complexity.
- Outsourcing can result in a shift of focus away from the customer.

The NationsBank Story

- What are the key advantages of moving toward the new call center at Nations?
- What are the key disadvantages?
- How would you put together a business case to justify the investment in the new call center design?
- What problems do you foresee in the implementation of this new design? How can these problems be dealt with?

BUAD 467/667
Service Management
Spring 2008

Professor Patrick T. Harker

Class 5b

Capacity Design III

Copyright P.T. Harker 1999

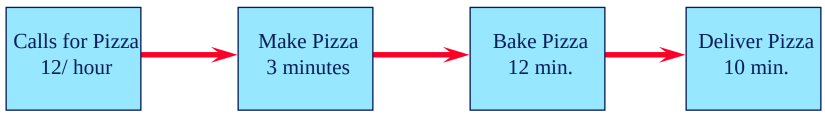
Outline for the Class

- Pronto Pizza
- Hints for Manzana Case
- Next Class

Pronto Pizza

- Delivery within 40 minutes or $2 off
- Pizza maker =$10/hour; 6 delivery people @ $8/hour
- Orders arrive one every 5 minutes
- Make a pizza every 3 minutes; 12 minutes to cook (10 pizza capacity in oven)
- Delivery takes 10 minutes (max) plus 10 minutes to return.

Pronto Process Flow



Deterministic Analysis

20/hour

5*10= 50/hour

6*6= 36/hour

Simple Queuing Analysis:

- bake time is a constant 12 minutes
- pizza is placed immediately in the oven (no capacity limit)
- pizzas are delivered one at a time
- deliveries are uniformly distributed
- delivery is assumed to be 10 minutes (conservative)
- wait time for delivery is negligible

$W_q = \rho\ 1/(\mu - \lambda) = (12/20)/(20-12)(60)$
$= 4.5$ minutes

Expected maximum delivery time =
4.5 + 3 + 12 + 10 = 29.5 minutes

Things look OK!!

However, this is all an *average analysis*.

P(delivery time $\geq$ 40 minutes)
= P(Actual wait for prep. > 40-25 = 15 minutes)
= P(number in queue > 15/5 = 3)
= P(number in system > 4)
= P(n $\geq$ 5)
= $\rho^5 = (12/20)^5 = 0.077$ = 7.7% chance.

Thus, average cost of guarantee = 0.077*$2.00
= $0.15

Number of drivers for wait < 1 min.

$\lambda = 12$ $\mu = 6$

lambda	12		lambda = arrival rate			
mu	6		mu = service rate			
Outputs:						
s	Nq	Ns	Wq	Ws	P(delay)	Utilization
0						
1	infinity	infinity	infinity	infinity	1.000000	1.000000
2	infinity	infinity	infinity	infinity	1.000000	1.000000
3	0.888889	2.888889	0.074074	0.240741	0.444444	0.666667
4	0.173913	2.173913	0.014493	0.181159	0.173913	0.500000
5	0.039801	2.039801	0.003317	0.169983	0.059701	0.400000

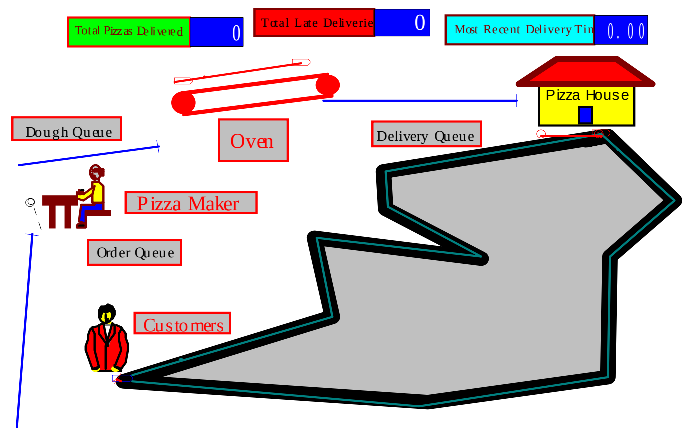
Number of makers for wait < 1 min.

$\lambda = 12$ $\mu = 20$

Outputs:

s	Nq	Ns	Wq	Ws	P(delay)	Utilization
1	0.900000	1.500000	0.075000	0.125000	0.600000	0.600000
2	0.059341	0.659341	0.004945	0.054945	0.138462	0.300000
3	0.006164	0.606164	0.000514	0.050514	0.024658	0.200000

Other Solutions: Simulation as a Tool



 Page 31

Arena Results

TALLY VARIABLES

Identifier	Average	Variation	Minimum	Maximum	Observations
Pizza Maker_R_Q Queue	4.5330	1.5264	.00000	33.478	466
Dispatching Station_R_	.00000	--	.00000	.00000	460

DISCRETE-CHANGE VARIABLES

Identifier	Average	Variation	Minimum	Maximum	Final Value
Pizza Maker_R Busy	.58667	.83936	.00000	1.0000	1.0000
TOTAL_TIME	24.813	.36702	.00000	60.177	45.515
# in Pizza Maker_R_Q	.89611	1.7838	.00000	9.0000	5.0000
# in Dispatching Stati	.00000	--	.00000	.00000	.00000
Pizza Maker_R Availabl	1.0000	.00000	1.0000	1.0000	1.0000
LATE_DELIVERY	8.6306	.84724	.00000	39.000	39.000

 Page 32

TOTAL_TIME(1)

Cell	Cell Limits From	To	Abs. Freq. (Time) Cell	Cumul.	Rel. Freq. Cell	Cumul.
1	-Infinity	12	18.85	18.85	0.007855	0.007855
2	12	14	87.07	105.9	0.03628	0.04413
3	14	16	136.3	242.2	0.0568	0.1009
4	16	18	268.6	510.9	0.1119	0.2129
5	18	20	335.5	846.4	0.1398	0.3527
6	20	22	273.7	1120	0.114	0.4667
7	22	24	210.3	1330	0.08763	0.5543
8	24	26	156.6	1487	0.06527	0.6196
9	26	28	248.8	1736	0.1037	0.7233
10	28	30	166.2	1902	0.06923	0.7925
11	30	32	97.27	1999	0.04053	0.833
12	32	34	80.21	2079	0.03342	0.8665
13	34	36	40.34	2120	0.01681	0.8833
14	36	38	48.56	2168	0.02023	0.9035
15	38	40	46.46	2215	0.01936	0.9229
16	40	42	46.28	2261	0.01928	0.9421
17	42	44	24.65	2286	0.01027	0.9524
18	44	46	15.48	2301	0.00645	0.9589
19	46	48	43.63	2345	0.01818	0.977
20	48	50	3.686	2349	0.001536	0.9786
21	50	52	0.7645	2349	0.0003185	0.9789
22	52	54	44.07	2393	0.01836	0.9973
23	54	56	0	2393	0	0.9973
24	56	58	0.3392	2394	0.0001413	0.9974
25	58	60	4.577	2398	0.001907	0.9993
26	60	+Infinity	1.655	2400	0.0006897	1

What About a Better Guarantee?

Drop the guarantee to 30 minutes:

Case	% Late	Salary	Revenue	Late	Profit	%Change
Base	32.49%	$11600	$18644	$1516	$5548	
1 more maker	18.72	13600	19616	918	5098	-8.11%
Base, $12 pizza	32.49	11600	25663	1516	12547	+126%
$12 + 1 maker	18.72	13600	26972	918	12454	+124%

Hints for Manzana

How to do a queueing analysis with multiple products?

Take two products (x & y), with means μ_x and μ_y and variance var(x) and var(y).

If p% of the products are x and (1-p) are y, then:

$E(mix) = p*\mu_x + (1-p)*\mu_y$

$var(mix) = p*var(x) + (1-p)*var(y) + p(1-p)[\mu_x - \mu_y]^2$

$C.V.(mix)^2 = var(mix)/E(mix)^2$

Example: RUNS & RAPS

Mix is RUNS = 13%, RAPS = 36%

p(RUN) = 13/(13+36) = 27%, p(RAP) = 73%

At the DC, E(RUNS) = 68.5, var(RUNS) = $(30.7)^2$

E(RAPS) = 50.0, var(RAPS) = $(24.9)^2$

E(mix) = 0.27*(68.5) + 0.73*(50.0) = 54.99

$var(mix) = 0.27(30.7)^2 + 0.73(24.9)^2 + 0.27*0.73*[68.5-50.0]^2 = (27.83)^2$

$C.V.(mix)^2 = (27.83)^2 / (54.99)^2 = (0.51)^2$

Next Class

- Manzana Insurance; try to apply the queueing results from the spreadsheet to this case
- Two groups chosen to discuss their analyses and analytical approaches!
- eBay case is optional and will wrap up the first half of the class

