Case Study IV: An E-Business Service

Prof. Daniel A. Menascé
Department of Computer Science
George Mason University
www.cs.gmu.edu/faculty/menasce.html

© 2004 D. A. Menascé. All Rights Reserved.

Copyright Notice

- Most of the figures in this set of slides come from the book "Performance by Design: computer capacity planning by example," by Menascé, Almeida, and Dowdy, Prentice Hall, 2004. It is strictly forbidden to copy, post on a Web site, or distribute electronically, in part or entirely, any of the slides in this file.

© 2004 D. A. Menascé. All Rights Reserved.

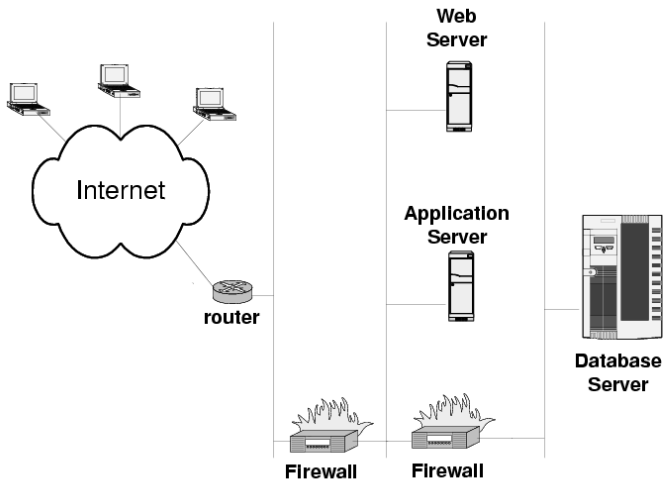
The E-Business Service

- Online auction site.
- One Web Server, one Application Server, and one Database Server.
- Each server has one CPU and one disk.
- Services offered by the site:
  - Create and broker auctions
  - Search for auctions based on categories and keywords
  - Monitor existing bids on open auctions
  - Place bids on open auctions.
  - Login

3

© 2004 D. A. Menascé. All Rights Reserved.

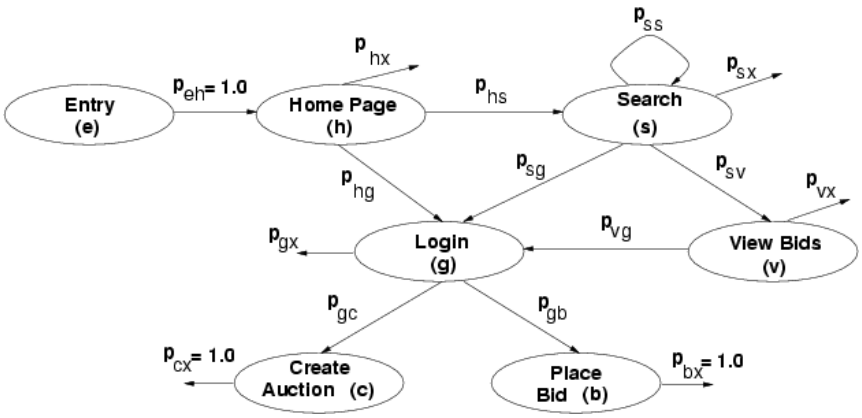
Auction Site's Architecture



4

© 2004 D. A. Menascé. All Rights Reserved.

The Customer Behavior Model Graph (CBMG)



© 2004 D. A. Menascé. All Rights Reserved.

Matrix of Transition Probabilities for Type A Sessions

	Entry (e)	Home (h)	Search (s)	View Bids (v)	Login (g)	Create Auction (c)	Place Bid (b)	Exit (x)
Entry (e)	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
Home (h)	0.00	0.00	0.70	0.00	0.10	0.00	0.00	0.20
Search (s)	0.00	0.00	0.40	0.20	0.15	0.00	0.00	0.25
View Bids (v)	0.00	0.00	0.00	0.00	0.65	0.00	0.00	0.35
Login (g)	0.00	0.00	0.00	0.00	0.00	0.30	0.60	0.10
Create Auction (c)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Place Bid (b)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Exit (x)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

© 2004 D. A. Menascé. All Rights Reserved.

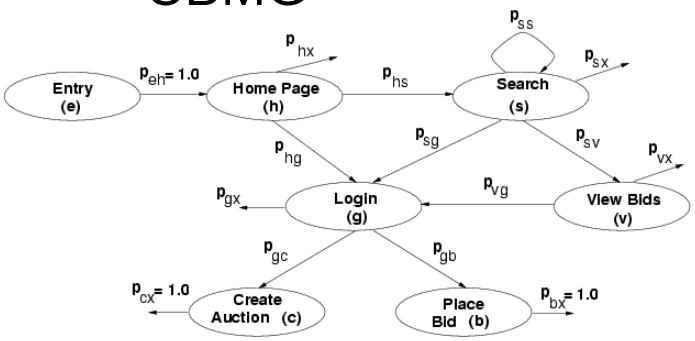
Matrix of Transition Probabilities for Type B Sessions

	Entry (e)	Home (h)	Search (s)	View Bids (v)	Login (g)	Create Auction (c)	Place Bid (b)	Exit (x)
Entry (e)	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
Home (h)	0.00	0.00	0.70	0.00	0.10	0.00	0.00	0.20
Search (s)	0.00	0.00	0.45	0.15	0.10	0.00	0.00	0.30
View Bids (v)	0.00	0.00	0.00	0.00	0.40	0.00	0.00	0.60
Login (g)	0.00	0.00	0.00	0.00	0.00	0.30	0.55	0.15
Create Auction (c)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Place Bid (b)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Exit (x)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

7

© 2004 D. A. Menascé. All Rights Reserved.

Computing Visit Ratios from the CBMG



$$V_e = 1$$

$$V_h = V_e \times p_{eh} = 1$$

$$V_s = V_h \times p_{hs} + V_s \times p_{ss}$$

$$V_v = V_s \times p_{sv}$$

$$V_g = V_h \times p_{hg} + V_s \times p_{sg} + V_v \times p_{vg}$$

$$V_c = V_g \times p_{gb}$$

8

© 2004 D. A. Menascé. All Rights Reserved.

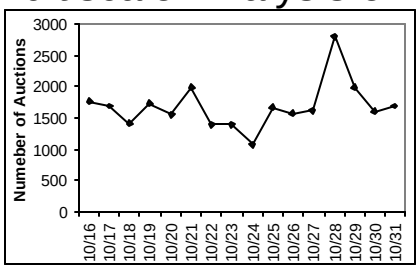
Visit Ratios for Both Session Types

	A	B
Ve	1.000	1.000
Vh	1.000	1.000
Vs	1.167	1.273
Vv	0.233	0.191
Vg	0.427	0.304
Vc	0.128	0.091
Vb	0.256	0.167

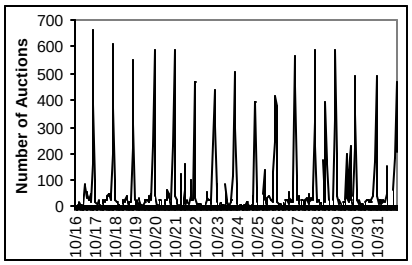
9

© 2004 D. A. Menascé. All Rights Reserved.

Workload Characterization
Multiscale Analysis of Number of Auctions



Time slot = 1 day

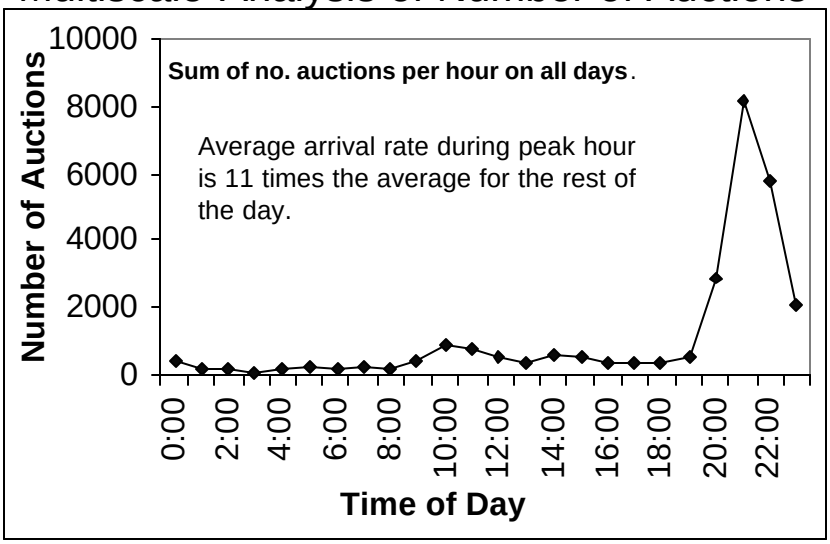


Time slot = 1 hour on each day.

10

© 2004 D. A. Menascé. All Rights Reserved.

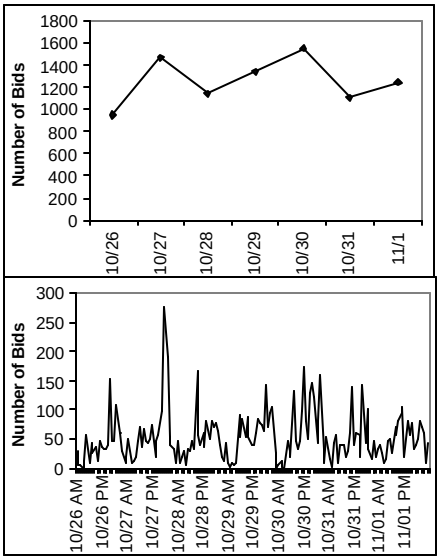
Workload Characterization
Multiscale Analysis of Number of Auctions



11

© 2004 D. A. Menascé. All Rights Reserved.

Workload Characterization
Multiscale Analysis on Number of Bids



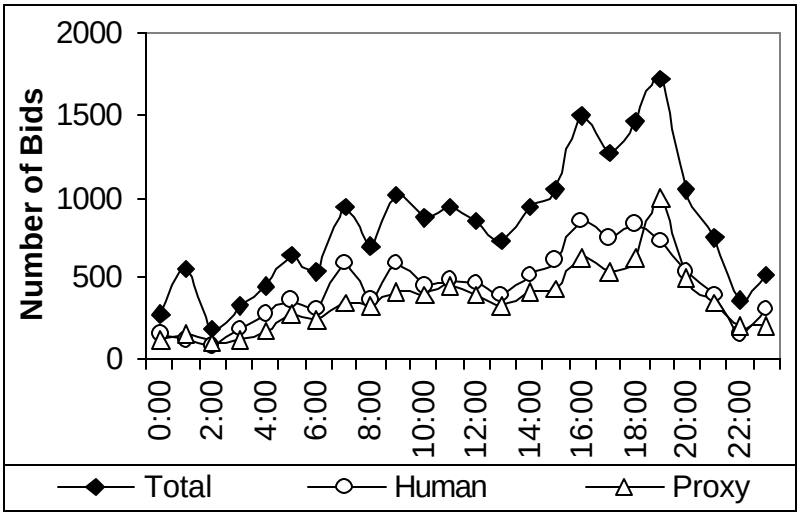
Time slot = 1 day

Time slot = 1 hour on each day.

12

© 2004 D. A. Menascé. All Rights Reserved.

Workload Characterization
Multiscale Analysis on Number of Bids



13

© 2004 D. A. Menascé. All Rights Reserved.

Performance Issues

- A surge in the number of auctions created and bids placed was observed between 8 p.m. and 11 p.m.
- What is the response time of the various types of requests (home page hits, search executions, bid viewings, logins, auction creations, and bid placements)?
- The response time SLA for create auctions and bid placement is 4 seconds.

14

© 2004 D. A. Menascé. All Rights Reserved.

Workload Intensity

γ: total rate at which sessions are started.

$$l_{\text{home}} = g(f_A \times V_h^A + f_B \times V_h^B)$$
$$l_{\text{search}} = g(f_A \times V_s^A + f_B \times V_s^B)$$
$$l_{\text{view}} = g(f_A \times V_v^A + f_B \times V_v^B)$$
$$l_{\text{login}} = g(f_A \times V_g^A + f_B \times V_g^B)$$
$$l_{\text{create}} = g(f_A \times V_c^A + f_B \times V_c^B)$$
$$l_{\text{bid}} = g(f_A \times V_b^A + f_B \times V_b^B)$$

15

© 2004 D. A. Menascé. All Rights Reserved.

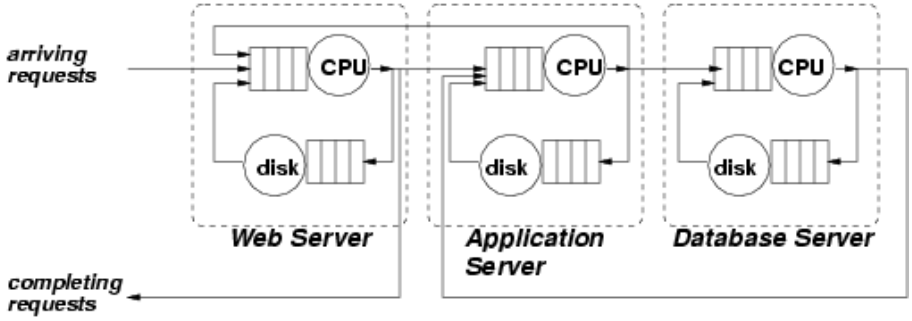
Workload Intensity

Total Session Arrival Rate (sessions/sec)	10.96
Percent of Type A Sessions	0.25
Percent of Type B Sessions	0.75
Arrival of requests (requests/sec)	
Home (h)	10.960
Search (s)	13.658
View bids (v)	2.209
Login (g)	3.665
Create Auction (c)	1.099
Place Bids (b)	2.074

16

© 2004 D. A. Menascé. All Rights Reserved.

Performance Model



Multiclass Open Queuing Network Model

17

© 2004 D. A. Menascé. All Rights Reserved.

Original Configuration

Open Multiclass Queuing Networks

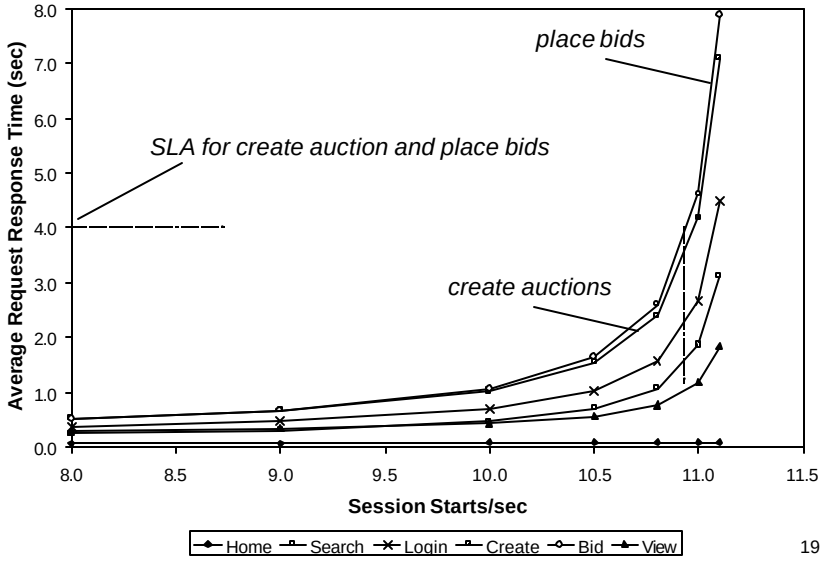
This wokbook comes with the books "Capacity Planning for Web Services" and "Scaling for E-Business"
by D. A. Menascé and V. A. F. Almeida, Prentice Hall, 2002 and 2000.

No. Queues:	6						
No. of Classes:	6						
	Classes ®						
Arrival Rates:		11.200	13.958	2.257	3.745	1.124	2.120
	Service Demand Matrix						
	Classes ®						
Queues ¯	**Type ¯ (LI/D/MPn)**	Home (h)	Search (s)	View bids (v)	Login (g)	Create Auction (c)	Place Bids (b)
WS-CPU	LI	0.008	0.009	0.011	0.060	0.012	0.015
WS-disk	LI	0.030	0.010	0.010	0.010	0.010	0.010
AS-CPU	LI	0.000	0.030	0.035	0.025	0.045	0.040
AS-disk	LI	0.000	0.008	0.080	0.009	0.011	0.012
DS-CPU	LI	0.000	0.010	0.009	0.015	0.070	0.045
DS-disk	LI	0.000	0.035	0.018	0.050	0.080	0.090

18

© 2004 D. A. Menascé. All Rights Reserved.

Response Times per Class



19

© 2004 D. A. Menascé. All Rights Reserved.

Open Multiclass Queuing Networks - Residence Times

This wokbook comes with the books "Capacity Planning for Web Services" and "Scaling for E-Business" by D. A. Menascé and V. A. F. Almeida, Prentice Hall, 2002 and 2000.

	Classes ®					
Queues ¯	Home (h)	Search (s)	View bids (v)	Login (g)	Create Auction (c)	Place Bids (b)
WS-CPU	0.01633	0.01837	0.02245	0.12246	0.02449	0.03061
WS-disk	0.06945	0.02315	0.02315	0.02315	0.02315	0.02315
AS-CPU	0.00000	0.10977	0.12806	0.09147	0.16465	0.14636
AS-disk	0.00000	0.01257	0.12573	0.01414	0.01729	0.01886
DS-CPU	0.00000	0.01640	0.01476	0.02459	0.11477	0.07378
DS-disk	0.00000	11.86563	6.10232	16.95089	27.12143	30.51161
Response Time	0.08578	12.04588	6.41647	17.22671	27.46578	30.80437

The disk at the database server is the bottleneck

20

© 2004 D. A. Menascé. All Rights Reserved.

Upgraded Configuration

Open Multiclass Queuing Networks

This wokbook comes with the books "Capacity Planning for Web Services" and "Scaling for E-Business" by D. A. Menascé and V. A. F. Almeida, Prentice Hall, 2002 and 2000.

No. Queues: 7
No. of Classes: 6

Classes ®

Arrival Rates: 11.200 13.958 2.257 3.745 1.124 2.120

Service Demand Matrix

Classes ®

Queues	**Type (LI/D/MPn)**	Home (h)	Search (s)	View bids (v)	Login (g)	Create Auction (c)	Place Bids (b)
WS-CPU	LI	0.0080	0.0090	0.0110	0.0600	0.0120	0.0150
WS-disk	LI	0.0300	0.0100	0.0100	0.0100	0.0100	0.0100
AS-CPU	LI	0.0000	0.0300	0.0350	0.0250	0.0450	0.0400
AS-disk	LI	0.0000	0.0080	0.0800	0.0090	0.0110	0.0120
DS-CPU	LI	0.0000	0.0100	0.0090	0.0150	0.0700	0.0450
DS-disk1	LI	0.0000	0.0175	0.0090	0.0250	0.0400	0.0450
DS-disk2	LI	0.0000	0.0175	0.0090	0.0250	0.0400	0.0450

21

© 2004 D. A. Menascé. All Rights Reserved.

Results with Two Disks at the DB Server

Open Multiclass Queuing Networks - Residence Times

This wokbook comes with the books "Capacity Planning for Web Services" and "Scaling for E-Business" by D. A. Menascé and V. A. F. Almeida, Prentice Hall, 2002 and 2000.

Classes ®

Queues	Home (h)	Search (s)	View bids (v)	Login (g)	Create Auction (c)	Place Bids (b)
WS-CPU	0.01633	0.01837	0.02245	0.12246	0.02449	0.03061
WS-disk	0.06945	0.02315	0.02315	0.02315	0.02315	0.02315
AS-CPU	0.00000	0.10977	0.12806	0.09147	0.16465	0.14636
AS-disk	0.00000	0.01257	0.12573	0.01414	0.01729	0.01886
DS-CPU	0.00000	0.01640	0.01476	0.02459	0.11477	0.07378
DS-disk1	0.00000	0.03490	0.01795	0.04985	0.07976	0.08974
DS-disk2	0.00000	0.03490	0.01795	0.04985	0.07976	0.08974
Response Time	0.08578	0.25005	0.35004	0.37553	0.50388	0.47223

22

© 2004 D. A. Menascé. All Rights Reserved.

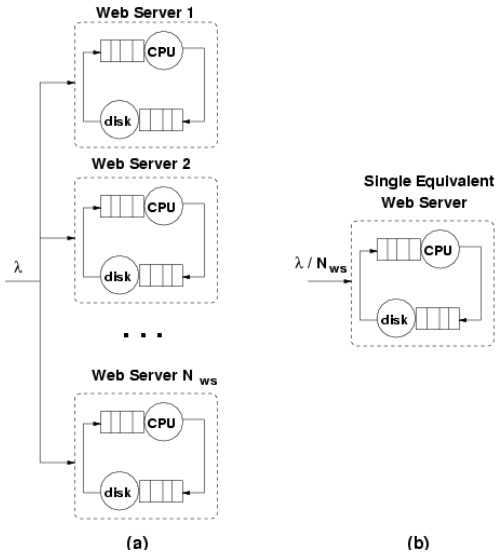
Improvements due to New Configuration

	Response Times (sec)					
	Home (h)	Search (s)	View bids (v)	Login (g)	Create Auction (c)	Place Bids (b)
Arrival Rates (req/sec)	11.200	13.958	2.257	3.745	1.124	2.120
Original Configuration	0.08578	12.04588	6.41647	17.22671	27.46578	30.80437
New Configuration	0.08578	0.25005	0.35004	0.37553	0.50388	0.47223
% Reduction	0.0%	97.9%	94.5%	97.8%	98.2%	98.5%

23

© 2004 D. A. Menascé. All Rights Reserved.

Adding More Identical Servers



24

© 2004 D. A. Menascé. All Rights Reserved.

Adding More Identical Servers

- Response time at the single equivalent server:

$$R = \sum_{i=1}^{K} \frac{D_i}{1-(l / N_{ws})D_i}$$

© 2004 D. A. Menascé. All Rights Reserved.