

Load Testing and Benchmarks

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

QoS Management Activities

- Benchmarking
- Load Testing
- Application Performance Management

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

Benchmarking

- Process used to compare the performance of a hardware or software system, the System Under Test (SUT).
 - Workload specification
 - Specification of metrics
 - Specification of measurement procedures

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

Benchmarking

- Workload specification: type of requests and frequency of submission of requests to the SUT.
 - Download the home page of a Web site every minute, 24 hours a day, from agent programs.
 - Mimic the execution of a customer buying books from an online bookstore according to some CBMG (see TPC-W).
 - Execute an online transaction on an online brokerage site.

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

Benchmarking

- Specification of metrics: determines what to measure.
 - Average download time of the home page and all of its components.
 - TPC-W: WIPS (Web Interactions per Second), \$/WIPS, WIPSb, WIPSo.

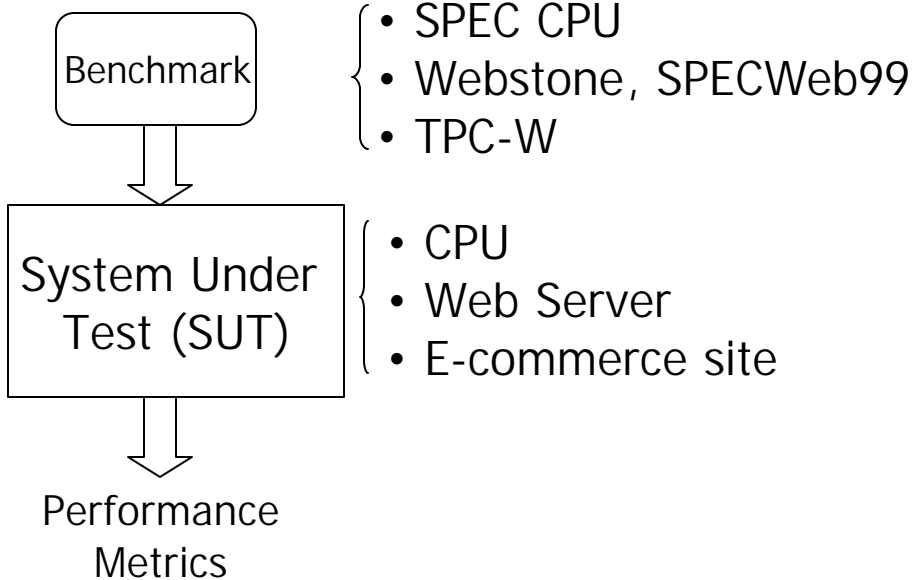
© 2004 D. A. Menascé. All Rights Reserved.

Benchmarking

- Specification of the measurement procedure: how the values of the metrics are obtained.
 - Average the home page download obtained by each agent.
 - Divide the total number of Web pages downloaded during the experiment by the measurement interval.

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

Benchmarking



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

Web and E-commerce Benchmarks

- Standard Performance Evaluation Corporation: SPECWeb99
 - <http://www.specbench.org/osg/web99/>
- Mindcraft: Webstone
 - <http://www.mindcraft.com/webstone/>
- Transaction Processing Council: TPC-W
 - <http://www.tpc.org/>

8

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

SPECweb99

- Successor to SPECweb96.
- Standard workload includes:
 - Connections at low speed
 - Dynamic and static GETs and POSTs
 - HTTP 1.0 and 1.1
 - Dynamic ad rotation using cookies and table lookup.
 - File access pattern more closely matching actual web workloads.
- A new extension of SPECweb99, called SPEC web99 SSL, tests secure web servers over SSL connections.

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

SPECweb99 Workload

- The workload file set consists of a number of directories. Each directory contains 9 files per class, 36 files in total.
- The files in Class 0 are in increments of 0.1K, those in Class 1 are in increments of 1K, those in Class 2 are in increments of 10K, and those in Class 3 are in increments of 100K in size.
- A Zipf distribution is used to access files in each directory.

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

SPECweb99 Workload (cont'd)

- File Size Distribution (four classes):

Class	File Size (KB)	Access Percentage
0	0-1	35%
1	1-10	50%
2	10-100	14%
3	100-1000	1%

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

SPECweb99 metric

- SPECweb99 measures the maximum number of simultaneous connections, requesting the predefined benchmark workload that a web server is able to support while still meeting specific throughput (40-50KB/sec) and error rate requirements.
- The connections are made and sustained at a specified maximum bit rate with a maximum segment size (1460 bytes) intended to more realistically model conditions that will be seen on the Internet during the lifetime of this benchmark.

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

SPECWeb Example

- An estimate for the number of concurrent customers, N , during the peak period is 10,000. The maximum avg. response time is 4 sec. The average think time per customer is estimated at 3 seconds. SPECWeb99 is deemed to be representative of the workload.

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

SPECWeb Example (cont'd)

- From the Response Time Law:

$$X_0 \geq \frac{N}{R + Z} = \frac{10,000}{4 + 3} = 1,429 \text{ requests/sec}$$

- Avg. number of simultaneous connections?
Let the avg. network time = 1.2 sec.
Therefore, the Web site time = 2.8 sec (= 4.0 – 1.2).
- From Little's Law:

$$N_{conn} = X_0 \times R_{site} = 1429 \times 2.8 = 4,001$$

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

SPECWeb99 Results

System	Conforming Connections	Throughput (req/sec)	Response Time (sec)	Kbps
A	1890	5190.1	351.9	341.1
B	3222	9020.4	358.5	335.9
C	8710	24,334.1	359.6	340.2

Select system C since it meets both the throughput and avg. no. connections requirements.

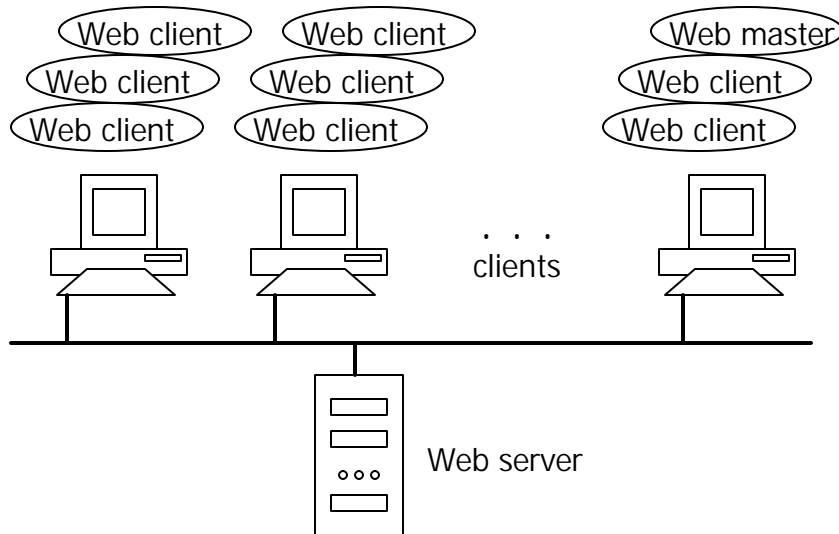
© 2004 D. A. Menascé. All Rights Reserved.

WebStone Workload

- Originally developed by SGI
- Workload specified in files read by clients.
- Requests supported:
 - GET HTML files
 - Run CGI Scripts
 - Run applications through server APIs (e.g., ISAPI and NSAPI).

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

Webstone



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

TPC-W: an E-commerce benchmark by the TPC

- www.tpc.org
- Designed to mimic operation of an e-commerce site (e-tailer).
- Scalable in number of concurrent users and in the database size.
- Transactions generated by TPC-W include:
 - Browsing activities (e.g., browse, search, select, view product detail)
 - Product order activities (e.g., shopping cart, login, register, buy request, and buy confirm)
- Database transactions must be ACID.
- Security through SSL is used for authentication.

18

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

The TPC defines transaction processing and database benchmarks and delivers trusted results to the industry.

- Home
- Results
- Benchmarks
 - TPC-C
 - Results
 - Description
 - FAQ
 - TPC-H
 - TPC-R
 - TPC-W
 - Obsolete
 - TPC-A
 - TPC-B
 - TPC-O
- Technical Articles
- Related Links
- What's New
- About the TPC
- Who We Are
- Member Login

Top Ten TPC-W Results by Performance

Item Count 10,000

Rank	Company	System	WIPS	\$/WIPS	System Availability	Database	Operating System	HTTP Server	Date Submitted
1	IBM	IBM eServer xSeries 440 with xSeries 330	21,139	32.42 US \$	12/31/02	Microsoft SQL Server 2000 Enterprise Edition	Microsoft Windows .NET Enterprise Server	Microsoft Internet Information Server 5.0	09/12/02
2	DELL	PowerEdge 6650 1.6GHz w/PowerEdge 1650 1.4GHz	10,449	27.40 US \$	08/22/02	Microsoft SQL Server Enterprise Edition SP2	Microsoft Windows 2000 Advanced Server	Microsoft Internet Information Server 5.0	08/22/02
3	DELL	Dell PowerEdge 6400/700 with PowerApp Web 1.0	7,783	24.50 US \$	01/28/02	Microsoft SQL Server 2000 Standard Ed.	Microsoft Windows 2000 Advanced Server	Microsoft Internet Information Server 5.0	01/28/02
4	IBM	IBM eServer xSeries 350 with IBM xSeries 330	7,073	31.77 US \$	12/17/01	Microsoft SQL Server 2000 Standard Ed.	Microsoft Windows 2000 Advanced Server	Microsoft Internet Information Server 5.0	12/17/01
5	DELL	Dell PowerEdge 6400/700 with PowerApp Web 1.0	6,422	25.70 US \$	12/19/01	Microsoft SQL Server 2000 Standard Ed.	Microsoft Windows 2000 Advanced Server	Microsoft Internet Information Server 5.0	12/19/01
6	UNISYS	Unisys e-@ton Enterprise Server ES508SR	5,745	69.00 US \$	03/05/01	Microsoft SQL Server 2000 Standard Ed.	Microsoft Windows 2000 Advanced Server	Microsoft Internet Information Server 5.0	03/05/01
7	DELL	PowerEdge 6400/700MHz w/PA. Web 1.0 & PA. Cache 200	3,130	67.50 US \$	03/01/01	Microsoft SQL Server 2000 Standard Ed.	Microsoft Windows 2000 Advanced Server	Microsoft Internet Information Server 5.0	03/01/01
8	UNISYS	Unisys e-@ton Enterprise Server ES508SR (4P)	3,008	81.77 US \$	11/14/00	Microsoft SQL Server 2000 Standard Ed.	Microsoft Windows 2000 Advanced Server	Microsoft Internet Information Server 5.0	11/14/00
9	IBM	IBM Netfinity 5600 with Netfinity 6600R	1,262	277.08 US \$	08/01/00	Microsoft SQL Server 2000 Standard Ed.	Microsoft Windows 2000 Advanced Server	Microsoft Internet Information Server 5.0	07/01/00

Item Count 100,000

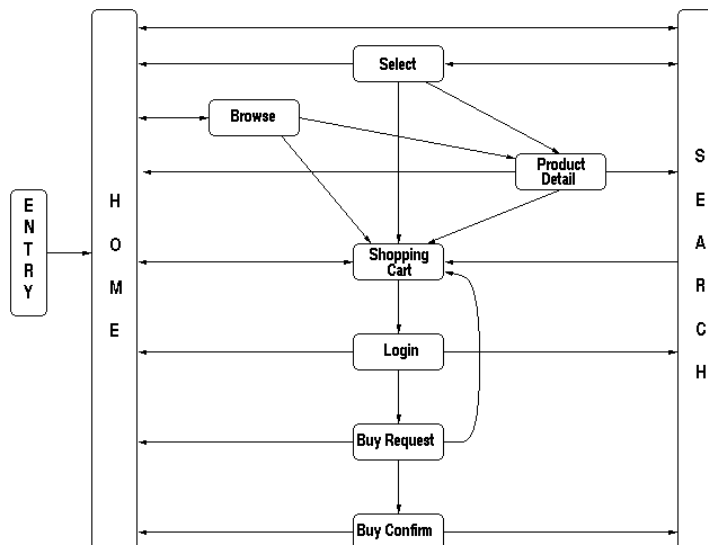
Rank	Company	System	WIPS	\$/WIPS	System Availability	Database	Operating System	HTTP Server	Date Submitted
1	UNISYS	e-@ton Enterprise Server ES 7000 (16P)	10,439	106.73 US \$	07/10/01	Microsoft SQL Server 2000 Enterprise Edition	Microsoft Windows 2000 Datacenter Server	Microsoft Internet Information Server 5.0	07/10/01
2	DELL	PowerEdge 6650	9,708	34.40 US \$	05/31/02	Microsoft SQL Server 2000 Enterprise Edition	Microsoft Windows 2000 Advanced Server	Microsoft Internet Information Server 5.0	05/31/02
3	IBM	IBM eServer xSeries 430 with IBM Netfinity 4500R	7,554	136.80 US \$	04/08/01	IBM DB2 UDB 7.2	IBM ptx 4.6.1	Microsoft Internet Information Server 5.0	05/01/01
4	IBM	NUMA-Q E410	6,272	195.59 US \$	02/02/01	IBM DB2 UDB 7.1	DYNIX/ptx 4.5.1	Microsoft Internet Information Server 5.0	02/02/01
5	IBM	IBM eServer xSeries 370 with xSeries 330	6,045	76.47 US \$	05/31/01	Microsoft SQL Server 2000 Enterprise Edition	Microsoft Windows 2000 Advanced Server	Microsoft Internet Information Server 5.0	05/31/01

A top

Home Results Benchmarks Technical Articles Related Links What's New About the TPC Who We Are Privacy Policy About Pricing

©2001 - 2002 TPC. All rights reserved.

TPC-W CBMG



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

TPC-W Types of Sessions

- Browsing mix: 95% browse interactions and 5% ordering interactions – 0.69% buy to visit ratio.
- Shopping mix: 80% browse interactions and 20% ordering interactions – 1.2% buy to visit ratio.
- Ordering mix: 50% browse interactions and 50% ordering interactions – 10.18% buy to visit ratio.

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

TPC-W Metrics: throughput and cost/throughput

- WIPS (Web Interactions Per Second) during shopping mix sessions. Specified as WIPS@number_items
- WIP Sb – Web Interactions Per Second during browsing mix sessions.
- WIP So – Web Interactions Per Second during ordering mix sessions.
- Cost/Performance

$$\frac{\text{Hdw Cost} + \text{Softw Cost} + \text{Maint. Cost}}{\text{WIPS}}$$

WIPS

22

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

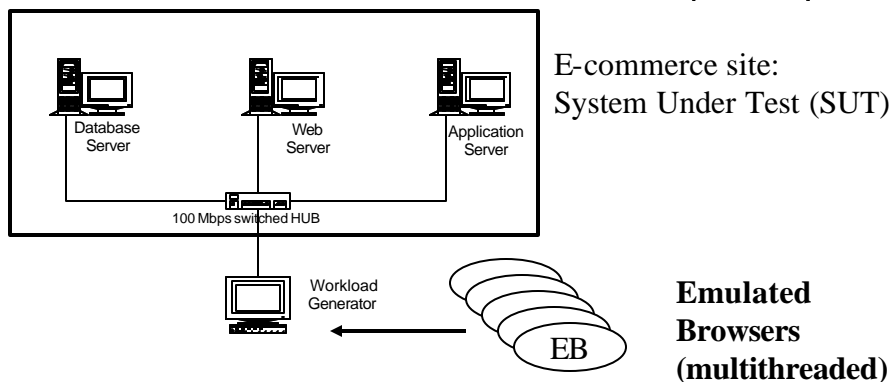
Example of TPC-W for 10,000 Items in the Catalog

Rank	System	WIPS	\$/WIPS
1	A	5,745	\$ 69.00
2	B	3,130	\$ 67.50
3	C	3,008	\$ 81.77
4	D	1,262	\$277.08

- ❑ the total price of System A is \$396,405, i.e., $5,745 \times \$69.00$.
- ❑ system D costs almost the same, i.e., \$349,675 but can only deliver 22% of the maximum throughput measured in WIPS.

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

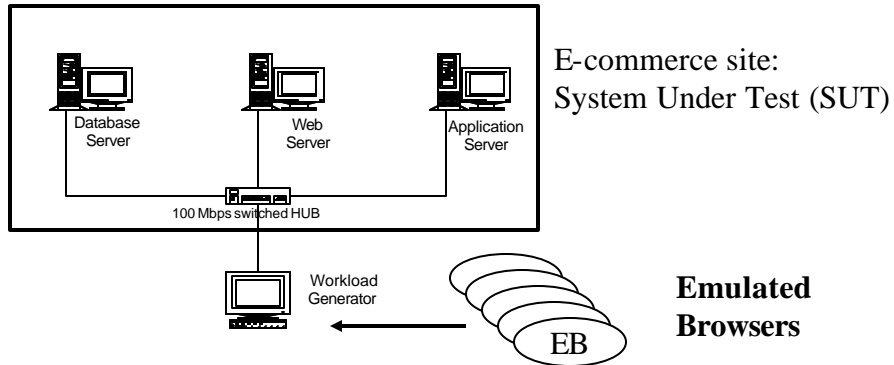
TPC Emulated Browsers (EBs)



- ❑ Each EB starts a session and generates all requests of that session.
- ❑ The minimum duration of a session (USMD) is exponentially distributed with mean 15 minutes, truncated at 60 minutes.

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

TPC Emulated Browsers (EBs)



- ❑ Requests are separated by user think times (Z), which are exponentially distributed with mean 7 sec truncated at 70 sec.
- ❑ Response Time Law:

$$R = (\text{No. EBs}) / \text{WIPS} - Z$$

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

Use of Interactive Response Time Law to TPC-W Results

Rank	System	WIPS	\$/WIPS
1	A	5,745	\$ 69.00
2	B	3,130	\$ 67.50
3	C	3,008	\$ 81.77
4	D	1,262	\$277.08

$$R = \frac{\text{No. EBs}}{\text{WIPS}} - Z$$

- ❑ Assume 50,000 concurrent users.
- ❑ System A: $R = 50,000 / 5,745 - 7 = 1.7 \text{ sec.}$
- ❑ System D: $R = 50,000 / 1,262 - 7 = 32.6 \text{ sec.}$

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

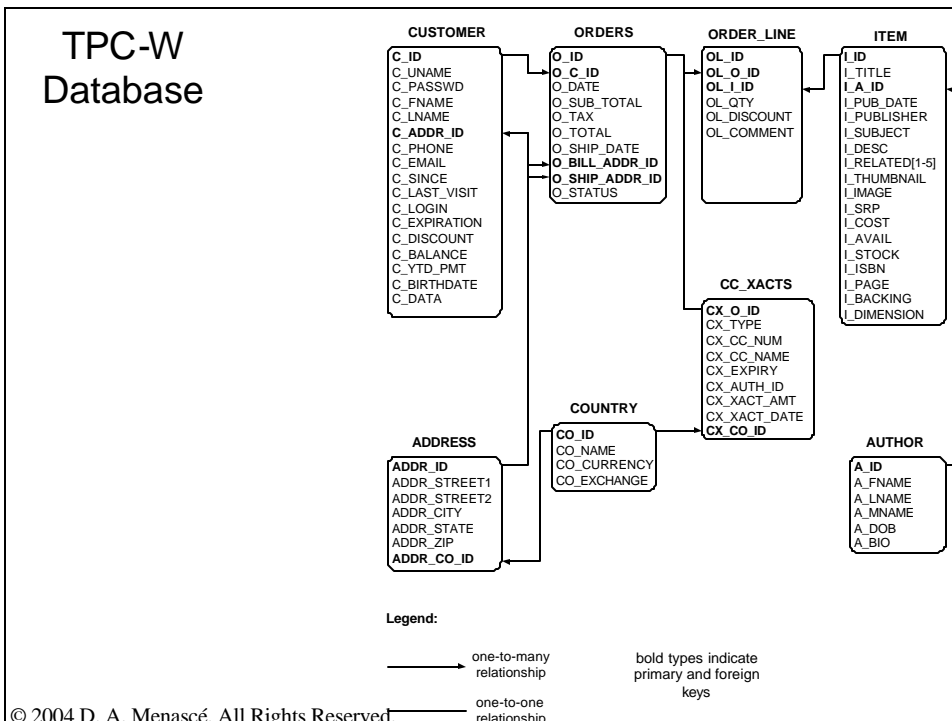
TPC-W Database

❑ Minimum of 8 tables with defined minimum number of fields:

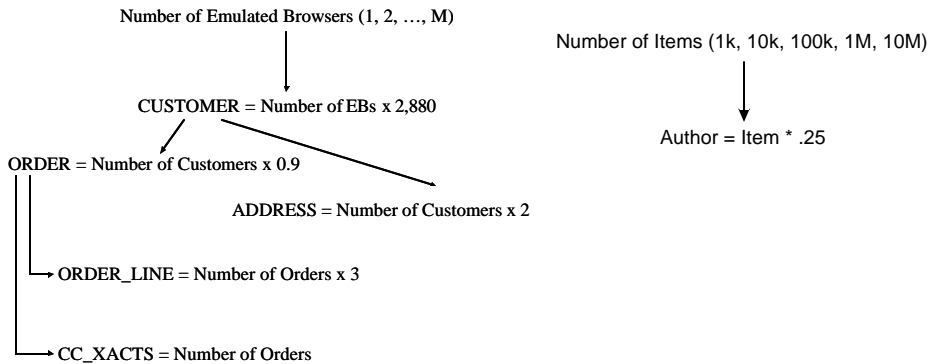
- Customer: Customer name and ID information,
- Address: Customer address data,
- Country: Country name and exchange rate information,
- Order: Order total and shipping information,
- Order line: Order line item data,
- Credit card: Credit card data,
- Item: Book information, and
- Author: Author data.

❑ TPC-W provides a function (WGEN) to generate the item title and the author last name in the database. We wrote a program to populate all other fields.

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



TPC-W Scalability



□ Assume 100 EBs:

- 288,000 rows in CUSTOMER table
- 259,200 rows in the ORDER table
- 576,000 rows in the ADDRESS table
- 777,600 rows in the ORDER_LINE table
- 259,200 rows in the CC_XACTS table

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

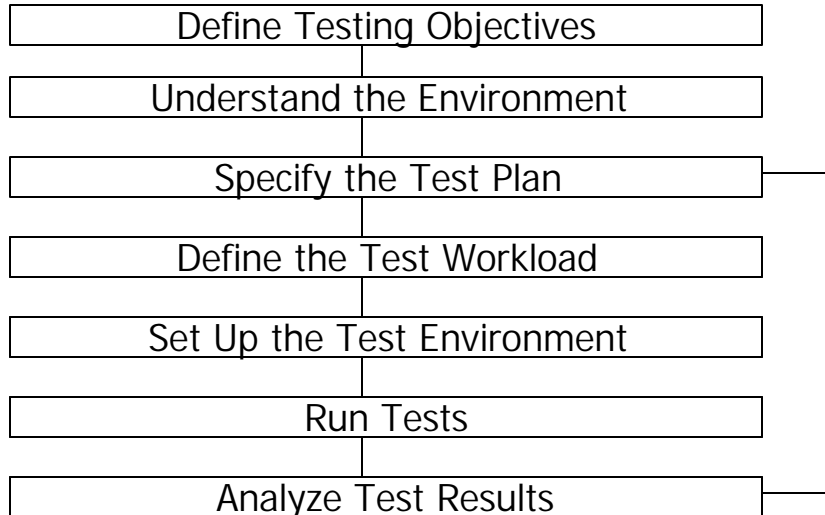
Example of TPC-W for 10,000 Items in the Catalog

Rank	System	WIPS	\$/WIPS
1	A	5,745	\$ 69.00
2	B	3,130	\$ 67.50
3	C	3,008	\$ 81.77
4	D	1,262	\$277.08

- the total price of System A is \$396,405, i.e., 5,745 x \$69.00.
- system D costs almost the same, i.e., \$349,675 but can only deliver 22% of the maximum throughput measured in WIPS.

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

A Methodology for Load Testing



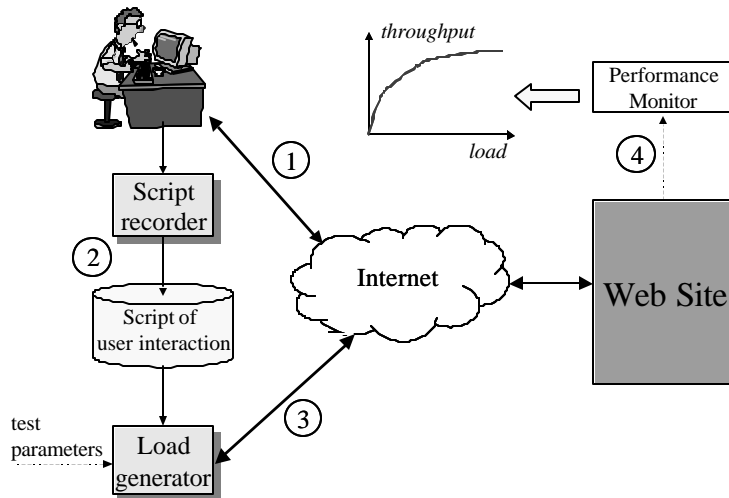
© 2004 D. A. Menascé. All Rights Reserved.

Load Testing

- Virtual user: emulates a real user.
- Virtual users need to have realistic think times.
- Virtual users should act like frustrated users and abandon the site in the case of excessive response time.
- Session abandonment is important for the correct assessment of a site's revenue throughput.

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

Load Testing

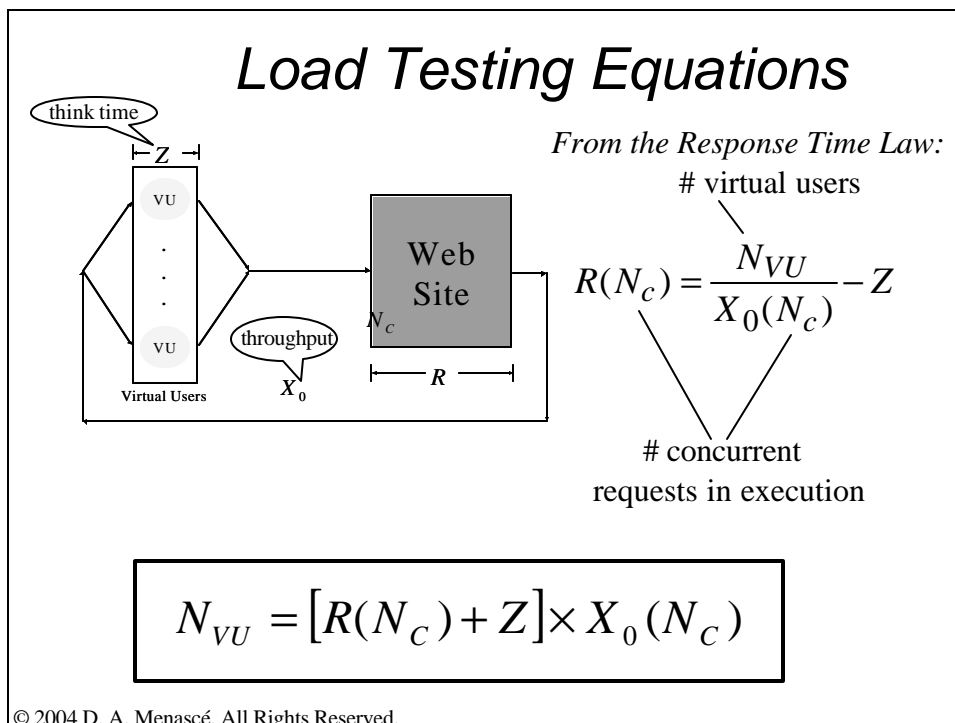
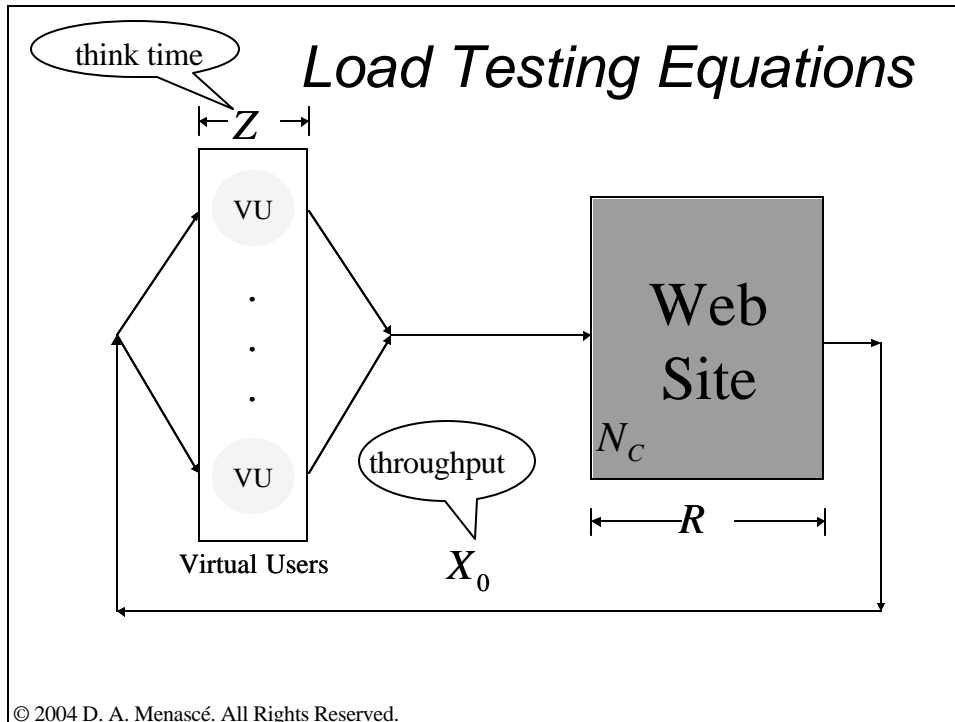


© 2004 D. A. Menascé. All Rights Reserved.

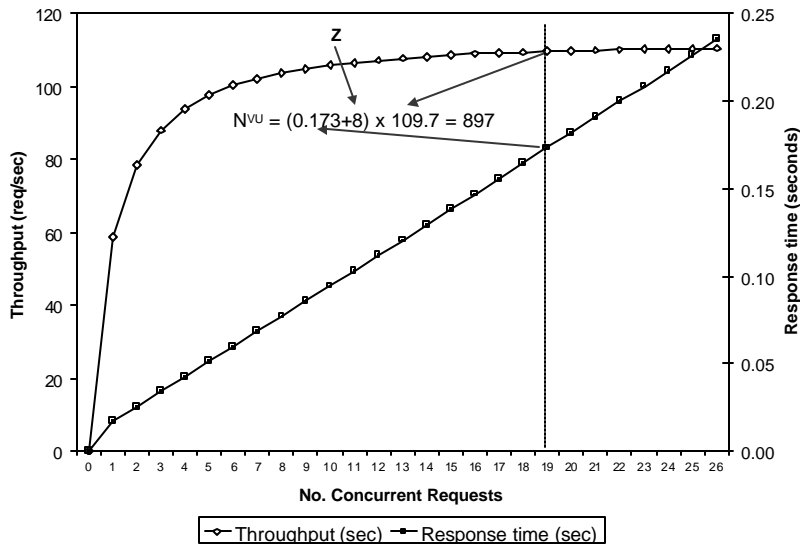
Load Testing

- What parameters to vary:
 - Workload intensity (session starts/hr)
 - Workload mix (combination of scripts)
 - Abandonment threshold and think times.
- Results of a load test:
 - Number of completed and abandoned sessions/hr as a function of the number of started sessions/hr
 - Revenue throughput and potential lost revenue throughput as a function of the number of started sessions/hr
 - Individual page download times and transaction completion times as a function of the number of started sessions/hr

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



Throughput and Response Time vs. No. Concurrent Users



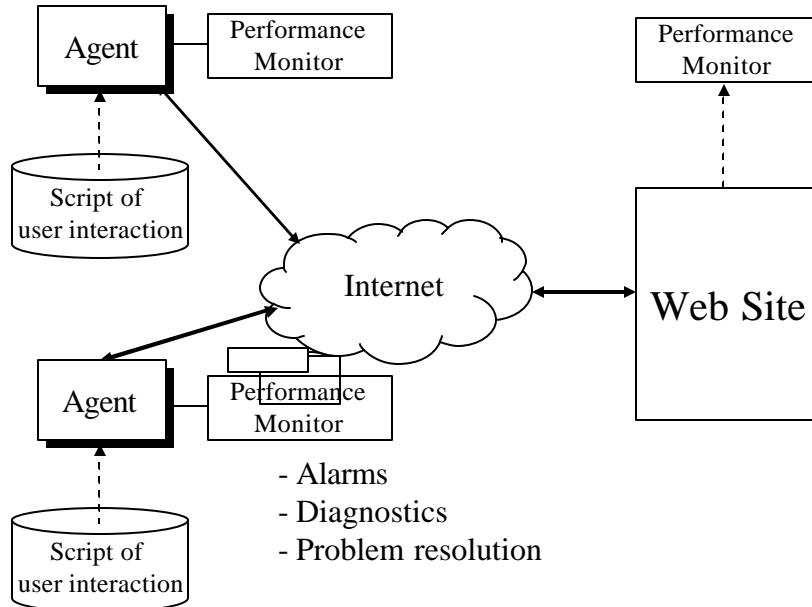
© 2004 D. A. Menascé. All Rights Reserved.

Application Performance Management (APM)

- Management processes aimed at guaranteeing that the QoS of e-business applications meet their business goals.
- Approaches to APM:
 - reactive (fire-fighting): monitor QoS and react to problems
 - proactive (preferred): try to reduce occurrences of poorly performing applications.

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

APM Framework



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