CS 672 - Capacity Planning for Web and E-commerce Examples

Daniel A. Menascé, Ph. D.

menasce@cs.gmu.edu

www.cs.gmu.edu/faculty/menasce.html

Exercise 1

- The following measurements were taken from a Web server. Compute the service demands, maximum throughput, and response times for arrival rates equal to 5, 10, and 15 requests/sec. What would you do to double the maximum throughput?

Measurement Period	1 hour
CPU Utilization	0.25
Disk 1 Utilization	0.35
Disk 2 Utilization	0.30
Number of HTTP requests processed	21,600

Exercise 2

- The following measurements were taken from a Web server. Compute the service demands and response times for HTML and image files for the current load and for a load 5 times bigger.

Measurement Period	1	hour
Number of HMTL files	14040	
Number of Image files	1034	

CPU time per KB/read	0.002	sec
Avg. Size of HTML file	3	KB
Avg. Size of an Image File	15	KB
Avg. Disk Time per KB/read	0.012	sec
File independent CPU Time/HTTP Request	0.008	sec

Exercise 3

- Consider the Web Server of Exercise 1. Draw a graph of the throughput and response time as a function of the number of concurrent requests in execution at the server. What do you observe?

Exercise 4

- Consider the Web Server of Exercise 2. What is the throughput and response time of HTML and image requests when there are 14 concurrent HTML and 6 image requests being executed concurrently?

Exercise 5

- A Web server has an arrival rate of 10 requests/sec and it has a processing capacity of 8 requests/sec. A maximum of 10 requests can be in execution at any time. Requests that arrive and find 10 requests in execution are rejected. What is the average response time, fraction of lost requests, server throughput, and server utilization?

Exercise 6

- Consider the Web server of exercise 5. Plot the response time versus the arrival rate L for L = 10, 20, 30, 40 requests/sec. What do you observe?

Exercise 7

- An e-commerce site has three Web servers, 2 application servers, and one DB server. Each box has one CPU and one disk. An analysis of the logs revealed that the majority of the requests are for searching items in the catalog.
- A CVM for the site reveals that each session has an average of 4.5 executions of the Search function.

Exercise 7

- Average sessions/week (historical data):

Week	Number of Sessions/week
1	284,600
2	340,012
3	411,431
4	483,798
5	601,567

- Peak hour shows arrival rate 3 x average
- Want to do CP for week 6.

Exercise 7

- More data:

Measurement Period	600 sec	
Completed sessions	1290	
	Utilizations	
	Ucpu	Udisk
WS	0.10	0.15
AS	0.25	0.18
DS	0.35	0.40

- What is the maximum throughput in sessions/sec?
- What is the average response time of the Search function at peak hours?