4 User Behaviors: Dangers of Closed versus Open Workload Generators

The problem We all use workload generators in our empirical studies of computer system performance. Many workload generators assume a *closed system model* (see Figure 1(left)), where new job arrivals are only triggered by job completions, following a think time, e.g., [2, 4, 13, 7, 6, 14, 9, 11, 5, 18, 19, 20, 21, 17, 1]. However, others assume an *open model* (see Figure 1(right)), where job arrivals occur independently of completions, e.g. according to a stochastic process or fixed trace, e.g., [15, 8, 10, 3, 22, 16]. Unfortunately most systems builders pay little if any attention to whether the workload generator is open or closed. This fact is often left out of the documentation, since it is not assumed to be important. We ask a question that surprisingly hasn't been asked:

> *What is the impact on measured system performance when using an open versus a closed workload generator, given that both are run under the same system load?*

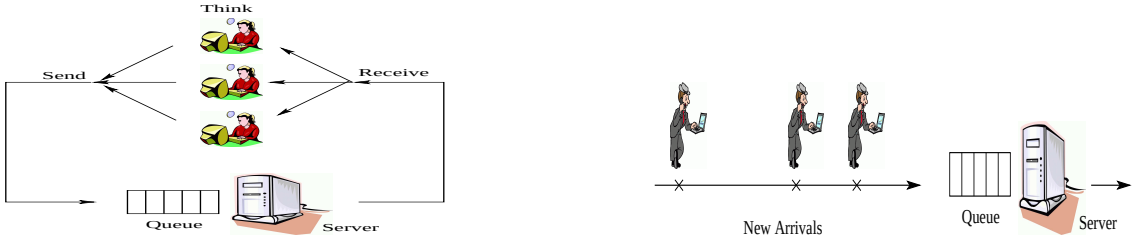

Figure 1: *Illustrations of a (left) closed versus (right) open system model.*

Surprising results In [12], we study performance of closed versus open workload generators under many applications, including static and dynamic web servers, database servers, auctioning sites, and supercomputing centers. The impact is huge! For example, under a fixed load, the mean response time for an open system model can exceed that for a closed system model by an order of magnitude or more, even when the MPL (multiprogramming level) of the closed system is high. We find that while scheduling to favor short jobs is extremely effective in reducing response times in an open systems, it has very little effect in a closed system model; this is tied to the fact that variability in the job sizes (service demands) has a much bigger effect in an open system than a closed one. These differences between open and closed models motivate the need for system designers to accurately determine whether an open, closed, or partly-open model best fits their system. We provide a simple recipe for making this choice, and for how to parameterize the model with respect to think time, MPL, and arrival and service rates.

Impact/Funding Although very recent (2006), our results are already been discussed in many computer systems reading groups. Funding was provided by an IBM graduate fellowship and a Siebel award.

References

[1] Webjamma world wide web (www) traffic analysis tools. http://research.cs.vt.edu/chitra/webjamma.html.

[2] C. Amza, E. Cecchet, A. Chanda, A. Cox, S. Elnikety, R. Gil, J. Marguerite, K. Rajamani, and W. Zwaenepoel. Specification and implementation of dynamic web site benchmarks. In *Workshop on Workload Characterization*, 2002.

[3] G. Banga and P. Druschel. Measuring the capacity of a web server under realistic loads. *World Wide Web*, 2(1-2):69–83, 1999.

[4] P. Barford and M. Crovella. The surge traffic generator: Generating representative web workloads for network and server performance evaluation. In *In Proc. of the ACM SIGMETRICS*, 1998.

[5] J. Fulmer. Siege. http://joedog.org/siege.

[6] P. E. Heegaard. Gensyn - generator of synthetic internet traffic. http://www.item.ntnu.no/ poulh/GenSyn/gensyn.html.

[7] K. Kant, V. Tewari, and R. Iyer. GEIST: Generator of e-commerce and internet server traffic. In *Proc. of Int. Symposium on Performance Analysis of Systems and Software*, 2001.

[8] Z. Liu, N. Niclausse, and C. Jalpa-Villanueva. Traffic model and performance evaluation of web servers. *Performance Evaluation*, 46(2-3):77–100, 2001.

[9] B. A. Mah, P. E. Sholander, L. Martinez, and L. Tolendino. Ipb: An internet protocol benchmark using simulated traffic. In *MASCOTS 1998*, pages 77–84, 1998.

[10] D. Mosberger and T. Jin. HTTPERF: A tool for measuring web server performance, 1998.

[11] E. Nahum, M. Rosu, S. Seshan, and J. Almeida. The effects of wide-area conditions on www server performance. In *Proceedings of* ACM SIGMETRICS '01, pages 257–267, 2001.

[12] B. Schroeder, A. Wierman, and M. Harchol-Balter. Closed versus open system models: a cautionary tale. In *Proceedings of Networked Systems Design and Implementation (NSDI)*, 2006.

[13] sourceforge.net. Deluge - a web site stress test tool. http://deluge.sourceforge.net/.

[14] sourceforge.net. Hammerhead 2 - web testing tool. http://hammerhead.sourceforge.net/.

[15] S. P. E. C. (SPEC). SPECweb99 benchmark. `http://www.specbench.org/osg/web99/`.

[16] Standard Performance Evaluation Corporation. SPECmail2001. `http://www.specbench.org/osg/web99/`.

[17] M. TechNet. Ms web application stress tool (wast). http://www.microsoft.com/technet/itsolutions/intranet/downloads/webstres.mspx.

[18] Transaction Processing Performance Council. TPC benchmark C. Number Revision 5.1.0, December 2002.

[19] Transaction Processing Performance Council. TPC benchmark W (web commerce). Number Revision 1.8, February 2002.

[20] G. Trent and M. Sake. Webstone: The first generation in http server benchmarking. http://www.mindcraft.com/webstone/paper.html.

[21] VeriTest. Webbench 5.0. http://www.etestinglabs.com/benchmarks/webbench/.

[22] M. Yuksel, B. Sikdar, K. S. Vastola, and B. Szymanski. Workload generation for ns simulations of wide area networks and the internet. In *Proc. of Comm. Net. and Dist. Sys. Mod. and Sim.*, 2000.