Web Server Benchmarking Using Parallel WAN Emulation

Rob Simmonds
Carey Williamson
Dept of Computer Science
University of Calgary
Calgary, AB, Canada
simmonds,carey@cpsc.ucalgary.ca

Russell Bradford
Dept of Math Sciences
University of Bath
Bath, UK
rjb@maths.bath.ac.uk

Martin Arlitt
Brian Unger
Dept of Computer Science
University of Calgary
Calgary, AB, Canada
arlitt,unger@cpsc.ucalgary.ca

ABSTRACT

This paper discusses the use of a parallel discrete-event network emulator called the Internet Protocol Traffic and Network Emulator (IP-TNE) for Web server benchmarking. The experiments in this paper demonstrate the feasibility of high-performance WAN emulation using parallel discrete-event simulation techniques on shared-memory multiprocessors. Our experiments with the Apache Web server achieve 3400 HTTP transactions per second for simple Web workloads, and 1000 HTTP transactions per second for realistic Web workloads, for static document retrieval across emulated WAN topologies of up to 4096 concurrent Web/TCP clients. The results show that WAN characteristics, including round-trip delays, link speeds, packet losses, packet sizes, and bandwidth asymmetry, all have significant impacts on Web server performance. WAN emulation enables stress testing and benchmarking of Web server performance in ways that may not be possible in simple LAN test scenarios.

1. INTRODUCTION

Network emulation is a hybrid performance evaluation technique that combines experimental implementation with simulation (or even analytical) modeling. This approach has received increasing research attention in recent years, as researchers address large, complex, and challenging Internet performance problems [1, 2, 5].

Network emulation is attractive in performance studies, for several reasons. First, an emulation environment offers a flexible, controllable, and reproducible environment for performance experiments. Second, network emulation enables controlled experimentation with end-user applications (e.g., Internet gaming, video streaming) in their native form, without porting them into a simulator, and without facing the transient behaviours of the Internet. Third, emulation enables experimentation with a wide range of network and workload configurations. As indicated by Nahum *et al.* [5], the properties of a Wide-Area-Network (WAN) environment can have significant impact on Internet protocol behaviours and Web server performance. Rigourous testing of end-user applications in a wide range of WAN scenarios can establish confidence in the robustness of a server or application prior to Internet deployment.

In this paper, we explore the use of network emulation in the evaluation of Internet systems and applications. In particular, we consider Web server benchmarking using the Internet Protocol Traffic and Network Emulator (IP-TNE) [6]. The IP-TNE is built using a parallel discrete-event simulation (PDES) kernel, enabling high-performance WAN emulation.

The IP-TNE enables the testing of real networks and distributed systems under controlled conditions, as provided by network simulation. The IP-TNE provides a detailed simulation model of an arbitrary IP internetwork WAN topology. Internet hosts can send IP packets to other hosts, whether real (on the Internet) or virtual (within the simulated WAN), via the emulator. Similarly, virtual hosts within the emulator can send (real) IP packets to other (real) hosts on the Internet. In other words, the IP-TNE provides the means to "translate" packets between real and simulated Internet environments. This translation is accomplished through a technique similar to "IP masquerading", carefully implemented and tuned to provide high-performance packet reading and writing at Gigabit Ethernet rates [3].

2. OVERVIEW OF PAPER

The purpose of this paper is to show how parallel network emulation using the IP-TNE can be used for Web server benchmarking. The paper demonstrates the capabilities of the IP-TNE as a Web server workload generator, and also serves to validate prior results [5] highlighting the impacts of round-trip delays and WAN packet losses on Web server performance.

Figure 1 provides an illustration of our approach to WAN emulation for the purposes of Web server benchmarking. Rather than following the traditional "centralized" approach in Figure 1(a), we use the approach in Figure 1(b), wherein the clients themselves are within the IP-TNE. This approach has several advantages. First, it eliminates the need for extra equipment in the experimental setup. Second, there is no need for elaborate synchronization of multiple client machines in the experiments. Third, it provides complete control over the client workload: we can model either homogeneous or heterogeneous clients, and we can completely

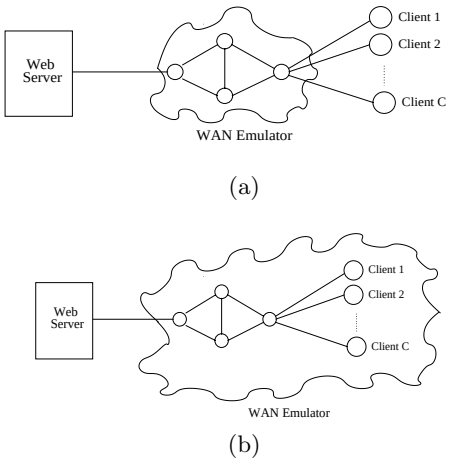

Figure 1: Approaches to WAN Emulation: (a) Traditional Approach; (b) Our Approach Using IP-TNE

specify their HTTP and TCP behaviours. Finally, this approach provides a fuller demonstration of the performance capabilities of the IP-TNE. Nahum *et al.* [5] argue that the approach in Figure 1(a) is not scalable for Web server benchmarking; we demonstrate via the IP-TNE example that the (more aggressive) approach in Figure 1(b) *is* feasible (and scalable enough) for Web server benchmarking.

3. EXPERIMENTAL METHODOLOGY

The experiments were conducted using two Compaq ES40 enterprise servers connected by Gigabit Ethernet on a private LAN. The IP-TNE runs on one ES40 and the Apache Web server on the other. Each ES40 has four Alpha 667 MHz Ev67 processors. Each computer is configured with 4 GB RAM and an 18 GB disk. The host operating system is Compaq's Tru64 (version 5.1A).

We consider six main factors in our WAN emulation experiments: number of clients, number of subnetworks, network link speed, propagation delay (i.e., round-trip time), MTU size, and router queue size. A one-factor-at-a-time experimental design is used. The number of clients and the number of networks are used to change the workload generated by the IP-TNE. The link capacity, RTT, and MTU factors are used to assess the impacts of different WAN configurations on Web server performance. Finally, the router queue size factor is used to assess the impact of packet losses on Web server performance. Synthetic Web workloads are used [4].

4. SUMMARY OF RESULTS

Our WAN emulation experiments demonstrate the importance of WAN characteristics, such as round-trip times, link speeds, MTU sizes, network asymmetry, Internet congestion, and packet losses on Web server performance. Several of our results validate those reported by Nahum *et al.* [5], even though our results were produced with a completely different experimental approach. These results are augmented with new observations regarding the impacts of link speeds, MTU sizes, and network bandwidth asymmetry. All of these factors can have significant impacts on Web server performance in a wide-area network.

5. CONCLUSIONS

This paper discusses the use of a parallel discrete-event IP network emulator called the IP-TNE for Web server benchmarking. The experiments focus on two main issues: the performance of the IP-TNE itself, and the performance of a Web server (Apache) when subjected to workloads from clients in an emulated WAN environment.

This work demonstrates that a "centralized" approach to WAN emulation is feasible, at least in the context of Web server performance testing. Our IP-TNE emulator, using a single physical machine and network interface, can generate adequate client workload to stress a production-quality Web server, while also modeling the packet-level events required for high-fidelity WAN emulation. The performance capabilities of the IP-TNE approach to WAN emulation come from a simulation kernel design that is optimized for parallel execution on shared-memory multiprocessors, and from efficient mechanisms for packet reading and writing at Gigabit Ethernet rates.

This work also reinforces prior observations that wide-area network conditions have an important impact on Internet server performance [5]. The results highlight the importance of WAN testing, or at least WAN emulation, in Web server benchmarking. (See [7] for a full version of this paper.)

6. REFERENCES

[1] G. Banga and P. Druschel, "Measuring the Capacity of a Web Server Under Realistic Loads", *World Wide Web Journal*, Vol. 2, No. 1, pp. 69-83, May 1999.

[2] P. Barford and M. Crovella, "Generating Representative Web Workloads for Network and Server Performance Evaluation", *Proceedings of ACM SIGMETRICS Conference*, Madison, WI, pp. 151-160, June 1998.

[3] R. Bradford, R. Simmonds, and B. Unger, "Packet Reading for Network Emulation", *Proceedings of IEEE MASCOTS Conference*, Cincinnati, OH, pp. 150-157, August 2001.

[4] M. Busari and C. Williamson, "ProWGen: A Synthetic Workload Generation Tool for Simulation Evaluation of Web Proxy Caches", *Computer Networks*, Vol. 38, No. 6, pp. 779-794, June 2002.

[5] E. Nahum, M. Rosu, S. Seshan, and J. Almeida, "The Effects of Wide-Area Conditions on WWW Server Performance", *Proceedings of ACM SIGMETRICS Conference*, Cambridge, MA, pp. 257-267, June 2001.

[6] R. Simmonds, R. Bradford, and B. Unger, "Applying Parallel Discrete Event Simulation to Network Emulation", *Proceedings of the 14th Workshop on Parallel and Distributed Simulation (PADS)*, Bologna, Italy, pp. 15-22, May 2000.

[7] R. Simmonds, C. Williamson, R. Bradford, M. Arlitt, and B. Unger, "A Case Study of Web Benchmarking Using Parallel WAN Emulation", available from http://www.cpsc.ucalgary.ca/~carey/papers/iptne.pdf