An Investigation into Factors Associated with Web Page Download Delay

Kevin Curran, University of Ulster at Magee, UK
Noel Broderick, University of Ulster at Magee, UK

ABSTRACT

Since the Mosaic browser was introduced with its ability to display images, there has been a proliferation in the number of Web pages using a combination of text and image-heavy design with the use of images contributing heavily to slow-loading sites. Studies have shown that surfers spend a lot of time waiting impatiently for Web pages to emerge on the screen, and HCI guidelines indicate that 10 seconds is the maximum response time before users lose interest. Such delays impact a site’s success and are expensive in terms of lost business opportunity or user productivity. This paper presents research into the observed usage of images by academic institutions and the effects on page retrieval times. Findings of this study will help to devise a strategy reducing potentially harmful effects of slow page retrieval times.

Keywords: image compression; performance measurements; Web page latency

INTRODUCTION

Studies have shown that Web users spend a lot of time waiting impatiently for Web pages to emerge on the screen. Web access is the most popular service, and statistics from NUA speculated that there are an overwhelming 605 million users online (Nua, 2002). Slow page retrieval time is the most widely reported problem, and users prefer Web pages to be presented on a computer screen as quickly as one can turn a page of a book. HCI guidelines indicate that 10 seconds is the maximum response time before users lose interest (Selvidge et al., 2002). Such delays impact the site’s success and are expensive in terms of lost business opportunity or user productivity (Saiedian et al., 2001). Any technique that saves bandwidth and makes browsing more pleasurable should be explored. When a user launches a browser and requests an action to be performed, the browser interprets the request. It sends information to the appropriate site server, where the re-
requested information is stored, and the site server sends the information back. The Internet is actually a packet switching network that sends requests via packets of data (datagrams). Each packet contains the IP address of the sender and the receiver and the information being requested. On one request, there can be more than one packet, because each packet is a fixed size, and some requests may need more than this. This means the request must be broken up into the appropriate number of packets. The route taken to obtain the requested information depends on the sender’s and the receiver’s geographic locations. If there are a lot of packets along a certain route, they will be queued or find a different route until they reach their destination. The destination cannot send any information until all the associated packets have been received. When all the packets have been received, the destination sends the requested information back in packets via routers to the sender (Comer, 2000).

Findings of this study will provide a better understanding and will help to devise a strategy on what one can do to eliminate or at least to reduce potentially harmful effects of very slow page retrieval times. The top-level domains (TLD) (i.e., home pages) of 47 academic Web sites were chosen. They span five different countries: UK, Ireland, Canada, US, and Australia. They have many Web users who typically hit the same home page. There will be potential students checking out the course prospectus. Perhaps commercial users will be looking for consultancy. The five countries under scrutiny have their own private multi-gigabit data communication network reserved specifically for research and education use and are linked to international peer networks. The first study examines network latency by visiting the Web sites for the first time. The efficiency of a cache mechanism in reducing the client latency also was assessed. Latency measurements were obtained from two different sources: a workstation at the author’s institution and a Web performance monitoring service provided by TraceRT (Tracert, 2004). The second study surveys the academics’ Web sites and accounts for variations in page retrieval times, particularly to images that were the interest of this study. The final study explores image compressions and assesses how efficiently Web developers are optimizing images for their Web sites. Another topic awaiting exploration was trimming the file size of images while retaining visual fidelity. The effects that reduced image sizes (in bytes) have on page retrieval times was looked into.

NETWORK DELAY COMPONENTS AND RELATED WORK

What follows is a description of some of the main factors influencing Web page delay.

Maximum Packet Size

A modern server uses Path Maximum Transmission Unit Discovery (PMTUD) heuristics to determine the Maximum Segment Size (MSS), which is the safe packet size that can be transmitted (Mogul et al., 1990). This technique was adopted to address the poor performance and communication failures associated with oversized packets that are fragmented at routers with small MTU (Kent et al., 1987). Today, the PMTUD concept is imperfect, as it uses the Internet Control Message Protocol (ICMP), which some network administrators view as a threat, and block them all, disabling PMTUD, usually without realiz-
Related Content

Multicast and Unicast ODMRP Routing in Ad hoc Networks Based on Quality of Service Support and Local Recovery Approach
[www.igi-global.com/article/multicast-unicast-odmrp-routing-hoc/1467?camid=4v1a](www.igi-global.com/article/multicast-unicast-odmrp-routing-hoc/1467?camid=4v1a)

The Internet's Impact on the Law and Economics Of the US's Motion Picture Industry
[www.igi-global.com/chapter/internet-impact-law-economics-motion/20462?camid=4v1a](www.igi-global.com/chapter/internet-impact-law-economics-motion/20462?camid=4v1a)

A Novel Dynamic Noise-Dependent Probabilistic Algorithm for Route Discovery in MANETs
[www.igi-global.com/article/novel-dynamic-noise-dependent-probabilistic/50481?camid=4v1a](www.igi-global.com/article/novel-dynamic-noise-dependent-probabilistic/50481?camid=4v1a)
Indoor Short Range Wireless Broadband Communications Based on Optical Fiber Distribution
Haymen Shams (2014). Convergence of Broadband, Broadcast, and Cellular Network Technologies (pp. 244-276).
www.igi-global.com/chapter/indoor-short-range-wireless-broadband-communications-based-on-optical-fiber-distribution/108097?camid=4v1a