

EXHIBIT I

TROLLTECH
 Trolltech Software Development
 Vendor Perspectives NetSeminar

Overcoming the Challenge

2000
14-1

VIEW ON DEMAND



Part of the **TopWeb** Business Technology Network

INTERNETWEEK
YOU BUY OF

W
Sponsor

SEARCH

find it!

Free Newsletter Glossary Contact Us About

NEWS | TRENDS | HANDS ON | BLOG | PRODUCT FINDER | INTERNET BUSINESS | SOFTWARE | SECURITY | W

INTERNETWEEK.com
CONNECTING THE ENTERPRISE, CUSTOMERS & SUPPLIERS

Thursday, June 1, 2000

Google Bets The Farm On Linux

Cost, support benefits of open-source OS drive massive server expansion

By MITCH WAGNER

Search engine Google.com is dramatically expanding one of the world's most computationally intensive Web sites by lashing together thousands of Linux PCs—a configuration it says is far less expensive than other options.

Google's 4,000 PC servers constitute one of the largest Linux installations. With the recent addition of a third data center in Herndon, Va., Google now has twice the server capacity it had just six months ago. That expansion is driven by the increased popularity of Google's Web site, which had 3.2 million unique visitors in April.

Why is Google running its business on an operating system supported mostly by legions of independent programmers? For starters: price.

The operating system itself, downloaded from vendor Red Hat, costs nothing, compared with \$500 to \$900 per server for Windows. The hardware is also cheap; Red Hat runs on custom-assembled PCs rather than more expensive RISC Unix servers. Many of the systems are based on Intel Celeron processors, the same chips in cheap consumer PCs.

"Our hypertext analysis is computationally expensive. We need to have an efficient system for doing that," said Google founder and president Sergey Brin. "That's why we use a lot of cheap PCs. The cost per MIPS is better for PCs."

Google's search algorithm requires massive computing power. Google weights each Web page for importance by analyzing the pattern of which pages link to others over all 300 million pages the search engine indexes. Google's process entails 500 million variables and 2 million terms to index every month, resulting in about 1 terabyte of data to index.

Support was another critical factor in Google's choice of Linux. The company has Linux expertise in-house, and it values the ability to delve into the source code to correct problems rather than having to rely on a vendor, Brin said. Where the in-house expertise isn't adequate, Google has found the Linux community responsive with fixes.

Major Commitment

Google may well be the largest Linux installation in the world. Internet content distribution company Akamai claims on its Web site to have "4,000+" Linux servers, but the company didn't respond to requests for comment.

Google's endorsement of Linux should make IT managers feel safer about deploying it for e-business applications, analysts said. But while the site is vast, the technology used to deploy it isn't new.

Like Deja.com, DoubleClick and other Linux-based e-businesses, the Google site uses custom software, conducting a single task and running a "stateless" application—meaning that the app doesn't need to remember anything from one transaction to the next. A typical e-commerce application needs to remember the state of a customer's shopping basket and line of credit from one transaction to the next.

Such business transactions run better on large-scale servers rather than arrays of small computers, which have usually sacrifice robustness for redundancy, said Jim Garden, an analyst at Technology Business Research.

Nonetheless, analysts say the Google strategy is sound. "This type of use of Linux enhances the concept that you can cluster a large number of platforms to solve an important problem and that you can run Linux on just about anything," said Aberdeen Group analyst Bill Claybrook.

But for Linux to spread into brick-and-mortar enterprises, more packaged applications will have to be written for it, Claybrook said.

In deciding on a platform for its search engine, Google made the hardware decision first. It went with PCs because RISC Unix systems from Sun Microsystems and Silicon Graphics are five to 10 times as expensive, Brin said.

Also, PCs are more conducive to high-density, rack-mounted configurations, Brin said. Google's PC servers are made by two small vendors: Rackable Systems and King Star Computer. Two Rackable machines fit in a single 1U high space in a rack, with one behind the other. This economy of space saves Google about \$1 million per year in collocation costs, Brin said.

These off-brand PCs cost less than half what the major PC makers charge, and the major brands are lagging in providing high-density systems, Brin said.

The systems Google uses are typically single-processor with 256 megabytes of memory and 80 gigabytes of storage. Google chose single-processor systems because multiprocessor servers are less stable and harder to manage, Brin said.

For an operating system, Linux seemed the best choice of the PC OSes, he said.

The price was right; Google doesn't pay any significant amount of money to Red Hat. Google downloads the software for free and gets support in-house and from the Linux community. Google actually paid for only about 50 copies of Red Hat, and those purchases were more of a goodwill gesture. "I feel like I should be nice, so when I go to Fry's I pick up a copy," Brin said.

Among Unix options, Sun's Solaris is available on the PC but isn't widely supported, Brin said. The open-source BSD operating system is well supported, but Linux appears to have more development momentum, more application support and a ready supply of personnel trained on the OS.

Windows NT and 2000 are more expensive than Linux, and they aren't stable enough to run Google.com, said Brin, who added that he doesn't trust the quality of Microsoft tech support. "In the Windows case, it's not how many dollars it would cost—it's how much heartburn," he said.

Google chose Red Hat because it's the most commonly used Linux OS, but it isn't relying on off-the-shelf Red Hat. It stripped the operating system of lots of unnecessary functionality: the compiler, the X Window system and the Apache Web server, as well as networking tools such as Telnet that Google thought would leave security vulnerabilities. Google does keep the Emacs text editor on each server so that IT managers can make changes to code on the fly to keep the PC servers up and running, Brin said.

Google developed its own network installation tools to load software remotely on 40 to 80 servers at once. Automatic installation is one area where Sun still has an edge over Linux, he said, noting that the installation tools were difficult to write and configure.

Distributed Network

Google has three data centers. Its first data center—in Santa Clara, Calif., hosted by Exodus Communications—went online two years ago. It opened a data center hosted by Frontier Corp. (now Global Crossing) in Sunnyvale, Calif., later that year to provide redundancy. It opened its third data center in Herndon as protection against an earthquake or other West Coast catastrophe.

Each of the data centers has an identical image of all the data. Google uses round-robin DNS to direct traffic among the data centers. However, it's transitioning to Border Gateway Protocol for better traffic management.

Within the data centers, Google uses its own traffic management and load-balancing software to direct traffic to the best server. The index of the Web is broken down into parts, with each section of the index distributed to a cluster of about 40 servers for redundancy and failover. The traffic management tools get their biggest workout about once a month, when Google updates its index and has to ship tens of terabytes of data across the LAN to update each server.

The servers are connected on 100-Mbps links within the racks; gigabit links connect the racks to one another.

Each server has two disks totaling about 80 GBs of storage. The disks currently used are IBM IDE disks and 40-GB MaxSource disks, though Google will switch to the 75-GB IBM disks when they become available. Google distributes its storage among the servers rather than use a centralized storage area network because it's cheaper and because a large RAID array would be a single point of failure, Brin said.

The company relies on revenue from ads on its search pages and from licensing the search engine to other Web sites, including Netscape, Red Hat, the Virgin Group and The Washington Post.

INTERNETWEEK MARKETPLACE (sponsored links)

On the CIO Agenda with IBM

With business growth back on the agenda, the role of the CIO is changing from manager of technology to C-suite collaborator in enabling innovation that matters for the business. Read an executive summary and register to download the full IBM paper.

Covad T1

Save \$2,006 in 2006 on Covad's Premier T1 Service. Arm your business with blazing-fast access with 24/7/365 service and guaranteed 99.99% monthly uptime. Keep your business a step ahead with Covad T1. Free Equipment and Installation.

Free Bandwidth Monitoring Tool

Monitor WAN bandwidth, find out top talkers, top protocols, and more. FREE software that uses Cisco NetFlow to generate reports on bandwidth usage. Download NetFlow Analyzer now and take control of your WAN.

Vonage High-Speed Internet Phone Service

Make unlimited calls to US and Canada over your high-speed Internet for just \$24.99/month.

Try WebEx Web Conferencing Free

WebEx online meetings and web conferencing solutions: Meet colleagues online, host web events, train your staff, provide live support. Try it out free now.

Buy a Link Now

[Top ten search terms from the TechWeb TechEncyclopedia](#) .

[How does your pay rate? Check the InformationWeek Salary Survey](#) .

[Mobilized Solutions Guide: Find and compare solutions for your business](#) .

[Top Requested White Paper Categories from TechWeb White paper Library](#) .

[Top ten search terms from the TechWeb TechEncyclopedia](#) .

Sponsored Links: [White Papers](#) . [Sponsor Resources](#) . [WebCasts](#) .

[News](#) | [Trends](#) | [Hands On](#) | [Blog](#) | [Product Finder](#) | [Internet Business](#) | [Software](#) | [Security](#) | [Web Dev](#)
[Original Articles](#) | [Free Newsletters](#) | [Internet Glossary](#) | [Contact Us](#) | [About Us](#) | [Privacy](#)

[The TechWeb Pipelines](#) | [TechWeb.com](#) | [InformationWeek](#) | [Optimize](#) | [Network Computing](#) | [IT Architect](#) | [Intelligent Enterprise](#)
[Bank Systems & Technology](#) | [Wall Street & Technology](#) | [Insurance & Technology](#) | [IT Pro Downloads](#) | [CommWeb](#)
[Business Intelligence Pipeline](#) | [Compliance Pipeline](#) | [Desktop Pipeline](#) | [Developer Pipeline](#) | [InternetWeek](#) | [Linux Pipeline](#)
[Messaging Pipeline](#) | [Networking Pipeline](#) | [Personal Tech Pipeline](#) | [Security Pipeline](#) | [Server Pipeline](#)
[Small Business Pipeline](#) | [SOA Pipeline](#) | [Systems Management Pipeline](#) | [Byte and Switch](#) | [Light Reading](#) | [Unstrung](#)

Copyright © 2006 CMP Media LLC. | INTERNETWEEK PIPELINE All rights reserved. [Privacy Policy](#) | [Your California Privacy Rights](#) | [Terms of Service](#)