

Scopeware Vision*: Real-time Information Management with Hyper-Threading Technology

Overview

The most perplexing problem facing knowledge workers is not the lack of information, but how to find the right information efficiently. A typical work day can involve wading through vast oceans of poorly organized information looking for the key materials needed for an ongoing project. The importance of improving the ease of information retrieval is growing daily as users continue to be inundated with ever increasing quantities of email, files, images, multimedia, and web pages. Complexity is compounded when all these different file types are stored in different locations.

Bringing all this information together into a single integrated view is a problem that traditional operating systems do not address. Today's standard information retrieval solutions are electronic versions of the centuries-old desktop and file cabinet metaphors. These models turn the user into a file clerk who is required to sort and classify information by hand, and remember its exact storage location for later retrieval. The scalability of these archaic systems has clearly begun to break down as overwhelming quantities of information continue to flood the knowledge worker's environment.

Rising to the challenge of solving the information retrieval problem is Scopeware Vision*. Vision takes a fresh look at this widespread (and time-wasting) problem. It ushers in a new information management paradigm that allows for the efficient retrieval of information by abstracting away the messy and irrelevant details of where content is stored. Just like search engines help users find and sort millions of documents on the Internet, Vision brings order to the chaotic jumble of email and files that are commonly strewn about the desktop and across network drives.

Breaking New Ground

Based on pioneering research by Scopeware founder, Dr. David Gelernter (a professor of computer science at Yale University and long-time technologist), Vision introduces a unique approach for finding and organizing all the files and email on a PC. "Vision shows you an integrated picture of all the files and email on your PC -- the electronic story of your life -- and lets you zero in quickly on what you need," according to Jim McGann, senior vice president of sales and marketing for Scopeware.

Vision presents search results in a stream of small images that provide a quick visual cue of what's in a document without requiring the user to actually open it. These thumbnails are organized in a time-ordered manner related to the query words. Instead of looking through thousands of unrelated items in Windows Explorer*, Microsoft Outlook*, and across the corporate intranet (or Internet), Vision users enter a search phrase to obtain a cohesive and comprehensive view of all relevant information on the desktop or network. Vision's time-ordered presentation gives this information additional context by arranging it as a natural story with a past, present, and future.



Up until now, users have been required to constantly fret about the best ways to organize the information flowing into their environment. The introduction of Vision turns the tables as it uses automation to further the goal of what computers are supposed to be doing in the first place: freeing people from mundane tasks. Vision does this by constantly monitoring the flow of information coursing through the desktop and shared network drives, and indexing it for later retrieval.

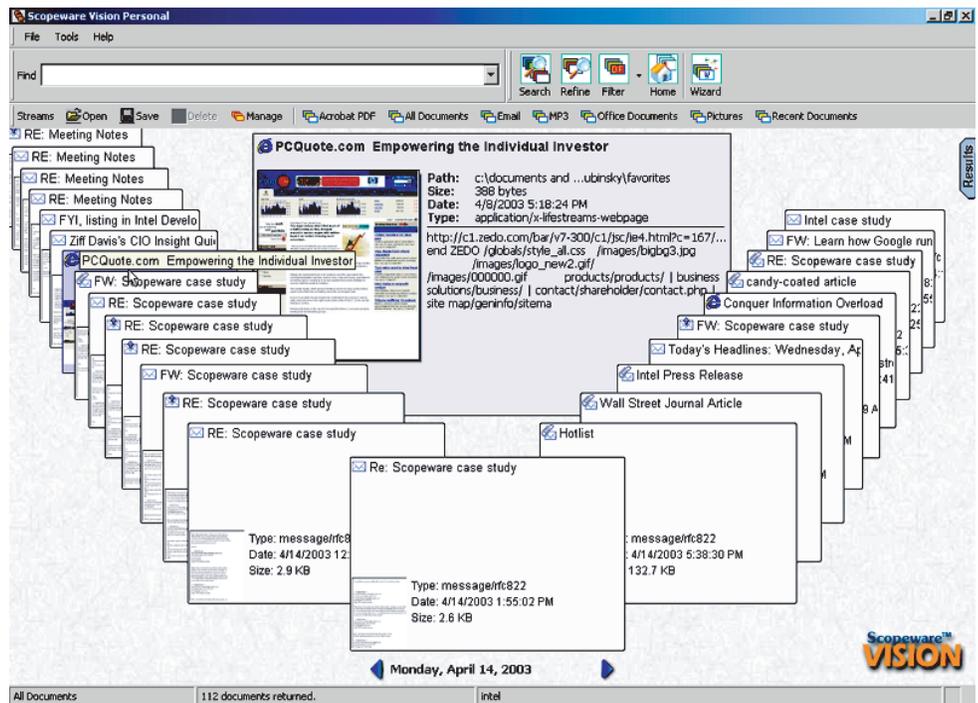


Figure 1. Scopeware Vision Advanced Stream View
(To view a higher resolution screenshot, go to http://www.scopeware.com/products/prod_v_stream.html)

Overcoming Challenges

To the user, Vision appears to be a simple application. All one sees is a cleanly designed interface where search phrases are entered and related results are graphically displayed in a time-ordered manner. The Vision interface also provides straightforward tools for facilitating time-based navigation and customizable filters that can limit search results to a user-specified subset of the indexed content. For instance, a user could limit the search results to matching MP3 files and Microsoft Word* documents. The real complexity in Vision, however, is behind the scenes -- where it makes clever use of internally developed and system provided services.

"The development challenge we faced," explained Peter Sparago, principal software architect, "was how to enable the real-time processing of incoming information without dragging down overall system performance." Unlike traditional Windows* applications that only consume processor resources when they are actively in-use, such as Microsoft Word, many components of Vision are designed to run quietly in the background. This requires the consumption of processor resources be moderated so ongoing foreground applications can still appear responsive.

Exploiting Concurrency

Although the Scopeware team designed Vision to run transparently in idle mode on older PCs, they still wanted a way for Vision to deliver a real-time user experience without interfering with other applications. The key to obtaining the performance needed to support real-time, yet unobtrusive processing of information was adopting a multi-threaded component software architecture that makes optimal use of the concurrency features of Intel processors -- specifically Hyper-Threading Technology (now available with 3.06GHz and faster Intel® Pentium® 4 processors).¹

This technology enables threaded applications to see improved performance due to additional concurrency and better utilization of on-chip execution resources. Each physical processor is viewed as two logical processors each with its own state.

Deciding on highly concurrent software architecture is one thing, making it happen is another. The first challenge was determining how to get the most out of Microsoft's Indexing Service* (a standard Windows service that monitors the file system for updates and indexes the contents of files). "Unfortunately, it was not designed with high levels of concurrency in mind," Sparago commented.

To work around this problem the Vision team created separate index catalogs for data from local files, email stores, and network files. Concurrency is obtained because the Indexing Service manages each catalog with a separate multi-threaded process.

A key part of Vision's feature set is the generation of document thumbnails. Although conceptually quite simple, the task of generating thumbnails without bogging down the system turned out to be problematic. "We found it had a much bigger impact on system responsiveness than expected," Sparago stated. The problem turned out to be an unfortunate interaction between the Microsoft Windows scheduler and the thumbnail generation code.

The Windows scheduler was designed with a bias towards boosting the priority of applications that make GDI calls such as to render an image or draw a font. Normally this is not a problem as only interactive applications typically make these calls. However, Vision must make the same calls to generate the document's thumbnail image. Even though the thumbnail generation code runs as a background process, the Windows scheduler still boosts its priority after the first GDI call, thereby affecting system responsiveness as the thumbnail generation thread steals additional processor resources away from other similarly high-priority interactive applications. On older PCs without Hyper-Threading Technology, Vision circumvents the effects of this issue by limiting the generation of thumbnails to times of user inactivity.

Running the thumbnail generation code on an Intel® Pentium® 4 processor - based system with Hyper-Threading Technology¹ revealed a startling change in system behavior. "Hyper-Threading enables the processor to deal smoothly with competing high priority threads," Sparago emphasized. To make the most of the additional concurrency provided by Hyper-Threading Technology, the team implemented a multi-threaded design within the thumbnailing process itself. This allows Vision to render thumbnail documents as they come into the system to provide the immediate benefits of visual cues in understanding the full story about a client, project, or research topic.

The design of Vision as a highly concurrent multi-threaded application has allowed it to fully benefit from the performance advantages offered by Intel processors with Hyper-Threading Technology. Scopeware ran tests that show that Hyper-Threading Technology provides an average performance gain of 25% when indexing content compared to the same processor with Hyper-Threading disabled. Scopeware also found the concurrency improvements offered by Hyper-Threading Technology resulted in the thumbnail generation process showing performance gains of up to 100% (effectively twice the throughput). Commenting on the significance of Hyper-Threading Technology, Sparago noted that, "It provides the additional concurrency that allows Vision to deal with information in a real-time way unlike any application before."

Summary

In the end, it's all about improving the user's productivity. Vision does this by utilizing the time when your machine might otherwise be idle to intelligently process information for ease of retrieval. System resources that were wasted before are now put to good use improving the daily lives of knowledge workers. Similarly, Hyper-Threading Technology provides the opportunity to further optimize resource utilization by making sure the processor is kept as busy as possible. Working together, Vision and Intel processors with Hyper-Threading Technology offer a solution that is carefully designed to improve productivity by making the most of a user's time and system resources.

For additional information on Scopeware Vision, visit <http://www.scopeware.com>



The logo for Scopeware Vision features the word "Scopeware" in a blue, sans-serif font with a trademark symbol (TM) to its upper right. Below it, the word "VISION" is written in a large, bold, orange, sans-serif font.

About the Author

Allan Mc Naughton is the principal analyst at Triangle Technology Advisors LLC (<http://www.triangletechadvisors.com>), a firm specializing in the composition of high-technology white papers. He can be reached at allan@triangletechadvisors.com.

Additional Resources

Intel® Developer Services provides online resources to help software developers accelerate the development and delivery of their applications, tools and solutions.

Find out more at the Intel Developer Services website at: www.intel.com/ids

1. Hyper-Threading Technology requires a computer system with an Intel® Pentium® 4 processor at 3.06 GHz or higher, a chipset and BIOS that utilize this technology, and an operating system that includes optimizations for this technology. Performance will vary depending on the specific hardware and software you use. See www.intel.com/info/hyperthreading for information.

Information in this document is provided in connection with Intel® products. Except as provided in Intel's terms and conditions of sale for such products, INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY RELATING TO SALE AND/OR USE OF INTEL PRODUCTS, INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT, OR OTHER INTELLECTUALPROPERTY RIGHT. Intel products are not intended for use in medical, life-saving, life-sustaining, critical control or safety systems, or in nuclear facility applications. Intel may make changes to specifications, product descriptions, and plans at any time, without notice.

Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing. For more information on performance tests and on the performance of Intel products, reference www.intel.com/procs/perf/limits.htm, or call (U.S.) 1-800-628-8686 or 1-916-356-3104.

Intel, Itanium, and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

*Other names and brands may be claimed as the property of others.

Copyright © 2003 Intel Corporation
All rights reserved.

