How Cisco Unlocks Communications

By **Tony Frazier,** Director of Product Management, Cisco Systems and
David Fishman, VP Marketing, Lucid Imagination

Historically, organizing and finding documents has been at the core of knowledge management and online collaboration—efforts to transform the collective intellect of an organization into a technology-powered asset. But perhaps documents are not the core of the proposition? Today, the best way to find the information you need to do your job may be to look more broadly at the discussions taking place outside of traditional text.

Consider this: you're looking for information and immediately search the documents at your disposal to find the answer. Are you the first person who conducted this search? If you are in a reasonably large organization, given the scope and mix of electronic communications today, there could be more than 10 other employees looking for the same answer. Unearthing documents, one employee at a time, may not be the best way of tapping into that collective intellect and maximizing resources across an organization. Wouldn't it make more sense to tap into existing discussions taking place across the network—over email, voice and increasingly video communications?

The Emerging Technologies Group at Cisco set out to solve these problems using network-based intelligence to find faster ways to close this knowledge gap. The result—a platform called Cisco Pulse.

Solving the knowledge gap begins with enhancing our understanding of who's communicating and what they're communicating about. We also have to take into account the medium—specifically, the explosive growth of online video and social networking applications and their adoption in the enterprise. It's not just about who's who, it's about what content they're creating and determining whether it's relevant to your work.

It turns out that the network is a pretty good place to process this information. Not only can it help you identify who's working with whom and when they're on or off-line, it's also possible to see what topics they're discussing—whether text-based or not.

Cisco's approach to this project centered on vocabulary-based tagging and search.

Every organization has the ability to define keywords for their personalized library. Cisco Pulse then tags a user's activity, content and behavior in electronic communications to match the vocabulary, presenting valuable information that simplifies and accelerates knowledge sharing across an organization. Vocabulary-based tagging makes unlocking the relevant content of electronic communications safe and efficient.

To implement this process of finding and tagging, we turned to open source technology—specifically, Solr/Lucene open source

heterogeneous applications such as people search and video discovery can be managed in a single search server.

With the rapid expansion of audio and video communication in the workplace, it is essential that we be able to handle rich media with our content management technology. Using some of the same search methods that unlock the content, we are able to make rich media a seamless part of the solution.

Putting video to work starts with making the content easier to access and work with, resulting in more useful information for the user. The first step is to eliminate the need to watch a video end-to-end, in real time, in order to find what's in it. By leveraging video metadata and applying text-to-speech technology to the audio track, Cisco Pulse creates an information structure around videos that makes them easier to search and extract information from.

This process of tagging feeds directly into the Solr/Lucene search technology at the core of Pulse. The content of the video—who's talking and what they are talking about—can become part of the social and knowledge flow among workers. It makes videos easier to browse and

> ## *"Unearthing documents, one employee at a time, may not be the best way of tapping into the collective intellect and maximizing resources."*

search to form the foundation of our search architecture. By using Solr, the Lucene Search Server, Cisco Pulse can tag data in real time at a very high rate of high content throughput.

Working with Lucid Imagination, Cisco implemented a high-speed Solr/Lucene search engine within Cisco Pulse that hosts indices as large as 35 million records on a single appliance and yields high-speed queries in a search time ranging in milliseconds. Solr sharding—a mechanism for distributing the index—also makes this architecture easily extensible to support larger volumes of data.

Advantages of Open Source

An important dimension of the use of Solr/Lucene is that it is available as open source. This affords two advantages: 1. the code is publicly available, and can be built upon freely; and 2. its transparency enables us to see, control and optimize how its search operations execute.

Importantly, Solr is fault tolerant and highly available, so it meets the stringent requirements of an enterprise-ready application. With Solr's multi-core architecture,

search, adding tags for topics mentioned in the video to the metadata for the file and resulting in more effective use of content.

When the network plays an active role in connecting people through content in all its forms—be it text, rich media, or online activity—there's yet another frontier. With search technology helping to unearth content and make it useful to the masses, we can now actively match content to end users before they need to look for it: no more searching a database, the content finds you.

When content finds you, it brings the exercise of search and knowledge management full circle. While topics you are working on are indexed and understood by the search system, the same thing is happening at the same time with others, across your organization. The content becomes the connection between people working on similar projects.

By building on the power of Lucene/Solr search, Cisco has transformed content from a passive, accumulating archive to a dynamic network of people and information. ■