

Technology Review: New Tape Indexing Technology May Impact e-Discovery Cost Shifting

By Joe Howie

When responding to discovery requests in U.S. litigation, the producing party normally bears all the costs of gathering and processing his or her own records. However, under the new Federal Rules of Civil Procedure governing electronically stored information, or ESI, the court may order the party seeking discovery to bear some or all the costs of recovering information that is "not reasonably accessible because of undue burden or cost." FR CivP 26(b)(2)(B).

When discussing this rule, commentators often use backup tapes as examples of information that is not reasonably accessible. In fact, in *Zubulake v. UBS Warburg LLC* (S.D. N.Y., No. 02 Civ. 1243), the leading case in this area, the plaintiff Laura Zubulake had requested that the defendant be ordered to restore a number of backup tapes for its e-mail system. Judge Scheindlin first ordered the defendant to restore a small sample set of tapes to determine how useful the information contained on the tapes would be, *Zubulake I*, 217 F.R.D. 309 (S.D.N.Y. 2003). After the results were obtained, the court ordered Zubulake to pay 25 percent of the costs of restoring further backup tapes, *Zubulake III*, 216 F.R.D. 280 (S.D. N.Y. 2003).

Cost shifting for backup tapes is all predicated on the idea that it is costly and burdensome to load, restore, index and search through backup tapes, which has been the case historically. But new technology appears to be lowering the cost and burden of searching backup tapes and, thus, permitting more extensive but far less expensive sampling or even wholesale processing of tapes. It also offers the option for the parties to search files contained on backup without first restoring the full content, choosing what is responsive and then extracting only what's relevant. This new technology is being offered by Index Engines, a New Jersey company that exhibited at the ABA Tech Show in Chicago in March. The Index Engine technology eliminates the need to restore the tape contents using the original backup software, typically very costly and time consuming, to perform discovery on the files and e-mail, said Jim McGann, the company's vice president of marketing.

The Index Engines product includes a computer with a 64-bit Quad-Core AMD Opteron CPU, 16 Gigabyte of RAM, one TB of storage and a variety of the hardware connections for linking the computer to tape hard drives, tape libraries, local area networks or storage area networks to index unstructured files and e-mail. This hardware includes SCSI, fiber and network connectors. Index Engines supports legacy and current versions of backup applications including CA ArcServe, IBM Tivoli Storage Manager, Symantec NetBackup and Backup Exec, and EMC NetWorker.

Because tapes are read by the appliance, Index Engines builds a catalog and index of the metadata values and text content of the objects on the tapes, at tape speed — approximately 30MB/second for LTO-2 tapes, for example. Users can then search the index using the Index Engines Query Builder to identify those records meeting the search criteria. As part of its processing, Index Engines creates a hash value for each object and uses this hash value as a way of suppressing duplicates if the user chooses this option. By using Query Builder, users can identify which objects on which tapes meet the criteria entered in the Query Builder. This information is then available to inform the parties and the court of the costs of doing full restores of the designated tapes.

Another common challenge is that users cannot select discrete objects to restore when using normal backup restoration; the process is normally an all or nothing proposition. With Index Engines, however, users can view a rendering of the file or e-mail from within Query Builder that displays an unformatted version of the content for review before extracting it from tape. Once tagged for restore, the file will be brought back in its original state, formatting, content and metadata all intact.

To continue receiving news from ALSP, become a [member](#) or [contact headquarters](#).

Call for Contributions
Send your article ideas to [ALSP headquarters](#) and make a point to participate in [ALSP Webcasts](#) and member activities.

Index Engines prices its hardware/software solution based on the number of items that can be included in search results at one time. The entry-level pricing is \$75,000 for 2 million files in a single search result. By using duplicate suppression, users could index far more than 2 million files with this entry-level model. McGann said that clients using this more targeted approach to searching and indexing backup tapes can save 50-70 percent of the costs associated with the more traditional tape restoration approach.

Law firms or companies that do not want to put the hardware and personnel infrastructure in place to support the tape indexing and restoration effort can use ONSITE³. "We typically experience an indexing rate of a gigabyte per minute with the newer tape formats like LT03, LT04 or SDLT, said Jeff Fehrman, president of Electronic Evidence Labs (a Division of ONSITE³), which uses Index Engines technology. "Sometimes the older formats like DLT4 will be slower. The Index Engines product is relatively easy to use and well within the capabilities of corporate backup administrators for ongoing operations. We often get involved if there is a large project involving hundreds or thousands of tapes. This may involve a specific case or it may be part of a corporate effort to proactively get control of their backup tapes. We can index thousands of tapes and save only those files which the corporation appears obligated to preserve."

In terms of the economics, Fehrman said: "We have one charge to process a tape and then charge a relatively nominal per gigabyte charge to restore those files that meet the search criteria. We have done projects where the cost savings compared to a normal restoration procedure are in the range of 70-80 percent."

Fehrman especially likes the capability of Index Engines to read Tivoli-formatted tapes as that is typically among the most difficult of tapes to process without Index Engines.

Key Points:

- Can fully index tapes at tape speed, typically from 5 to 40 MB/second
- Can search metadata or text content to identify which tapes contain responsive files
- Can view an unformatted rendering of original file within viewer and can restore original content of from selected files without requiring backup software

Further Information:

- Index Engines Web site: www.indexengines.com
- ONSITE³ Web site: www.onsite3.com

Has your litigation support team adopted a new technology that ALSP members should know about? [Send your technology tip to the ALSP Update managing editor.](#)

[Back to top](#)

[Next article](#)

[Previous article](#)

[Return to ALSP Update Home](#)

[ALSP Charter Sponsors](#)