

OLP UPDATE



Technology Beats the Burden

By Jim McGann

When a problem exists you can always count on a technologist to invent a solution. Years ago it was easy to be late to a party and claim you simply got lost. Now that everyone has a GPS, this claim is difficult to pull off. And the magic of Tivo has made it primetime anytime, no longer will you have to miss American Idol. Technology is a good thing, it makes us happier (especially if you are an Idol fan who likes to party) and when applied to our jobs it makes us more productive. The litigation support area has seen many innovations in the past few years, including technology that streamlines the process of identifying and collecting electronically stored information (ESI) in support of legal proceedings. One of the most dramatic changes the industry has seen is the evolution of tools that streamline the collection of ESI. These new tools are also helping to erode the use of the burden argument that was commonly exercised to avoid collection of ESI.

Some signs that the burden argument is eroding include recent rulings and opinions from well respected authorities like Judge Shira A. Scheindlin who has issued fines and sanctions against plaintiffs coming into her court without the data required to support the case. The judge, and others like her, is knowledgeable about changes in discovery technology and that is affecting the courts attitude toward what is or is not burdensome. Judge Scheindlin states in a recent opinion, "... the courts have a right to expect that litigants and counsel will take the necessary steps to ensure that relevant records are preserved, ... and that such records are collected, reviewed and produced..."

Rethinking Backup Tapes

The difficult, laborious and expensive restoration process required to discover data on backup tapes, as recently as 2007, has fallen quite dramatically by the wayside. Automatic tape indexing technology replaced traditional restoration in 2008, making discovery and collection of tape content fast and easy. As this technology has evolved, it has gotten faster and has added support for more files types and backup formats. Technology has even gone so far as to enable proactive tape discovery. Enterprise litigation support teams are processing tape data ahead of a court mandate, so they are armed and ready when the discovery request does come in.

Traditional restoration is a process where the original backup environment is used to restore tape content. Sounds easy, but as time goes by the original tape environment (software, tape drives, expertise, etc) fades away and is difficult to re-implement. This is why third-party restoration specialists came into being. They have the legacy environments, however they charge premium fees to support the old infrastructure and specialized skill set. Even if you are willing to pay these 3rd party fees, you will be required to sift through large quantities of irrelevant ESI to find what you need. Traditional restore has no knowledge of the tape content, so they need to restore it all before culling can begin. Automated direct indexing can “cherry pick” specific files and email from tape based on a query and extract them. Traditional restore moves 100% of the data from tape, while the automated approach extracts what you need, which is typically less than 5% of the total tape contents.

Automated tape discovery today, compared to the lone option of restoration in 2007, has resulted in more than a 90% savings in time. We all know that “Time is Money”. Thus, the impact that this technology has had on the “undue burden” argument, historically used to skirt eDiscovery, has been to render it ineffective.

The Evolution of Tape Discovery

Automated tape discovery solutions deliver direct access to file and email content locked away in proprietary backup formats. No longer do you need to be a restoration specialist with the appropriate software and infrastructure to unlock tape content. Now anyone can perform this process and deliver instant access to relevant data as needed. Technology has not only simplified the discovery process, but has also reduced the time and cost involved as well. Looking at the typical timeframes to access data on tape, the following example shows the dramatic reduction in the time. In this example, the task is to extract ESI from 25 custodians from email data bases backed up on 20 tapes.

AUTOMATED
TAPE
DISCOVERY
TODAY,
COMPARED
TO THE
LONE OP-
TION OF
RESTORA-
TION IN
2007, HAS
RESULTED
IN MORE
THAN A 90%
SAVINGS IN
TIME.

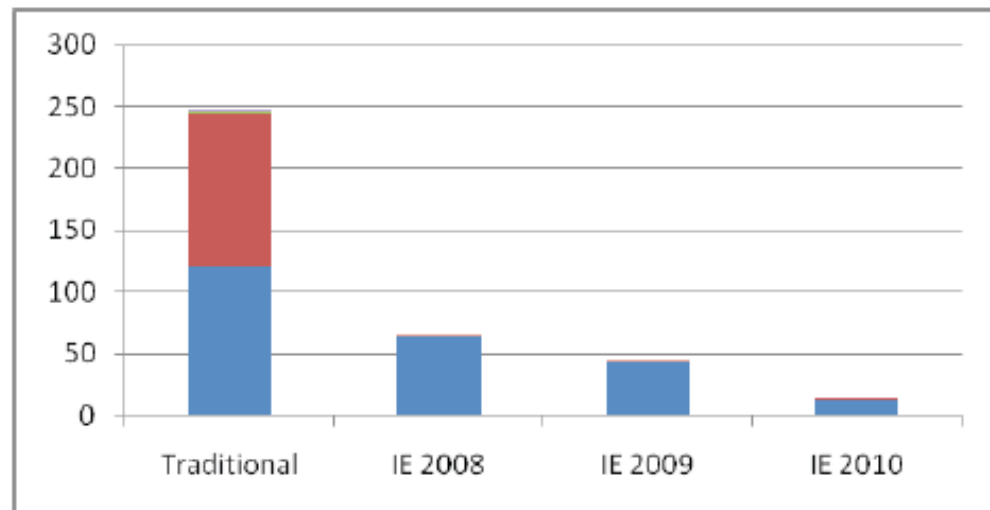
	2007	2008	2009	2010
Approach:	Traditional Tape Restoration	Direct Indexing	Indexing with Caching	Direct Indexing with Multi-Streaming
Step 1:	Restore ESI from tape using original backup software.	Index ESI, dedupe, query (keyword, date) to cull to relevant data set.	Index ESI, dedupe, query (keyword, date) to cull to relevant data set.	Index ESI, dedupe, query (keyword, date) to cull to relevant data set.
	Data: 8 TB	Data: 0 TB	Data: 0 TB	Data: 0 TB
	Time: 120 hours ¹	Time: 63 hours	Time: 43 hours	Time: 13 hours
Step 2:	Collect 25 custodian mailboxes assuming each mailbox is 5GB.	Extract relevant ESI from tape. Extract relevant ESI from tape	Extract relevant ESI from tape.	Extract relevant ESI from tape.
	Data: 125 GB	Data: 6.25 GB	Data: 6.25 GB	Data: 6.25 GB
	Time: 125 hours ²	Time: .17 hours	Time: .08 hours	Time: .05 hours
Step 3:	Load mailboxes to review platform.	Load to relevant ESI to review platform.	Load to relevant ESI to review platform.	Load to relevant ESI to review platform.
	Data: 125 GB	Data: 6.25 GB	Data: 6.25 GB	Data: 6.25 GB
	Time: 3 hours ³	Time: .17 hours	Time: 17 hours	Time: .17 hours
Step 4:	Index ESI, dedupe, query (keyword, date) to cull to relevant data set.			
	Data: 6.25 GB			
	Time: 1 hour			
Total:	Time: 249 hours	Time: 63.34 hours	Time: 43.25 hours	Time: 13.22 hours

1 6 hr/LTO2 tape is based on data from eMag's website - How Long Does it Take to Restore a Tape & Electronic Data Discovery: It's All About Access

2 5 hours/mailbox or 1GB/hour based on using PowerControls or Xmerge (data from Kroll website and Norcross Group)

3 Based on performance numbers of 1TB/day or 41GB/hour from various review platforms.

“THE JUDGE THREW OUT THEIR CLAIM AND MUCH TO THEIR CHAGRIN, REQUIRED THE RELEVANT DATA TO BE PRESENTED.”



As you can see, in the past four years the time involved in this project was reduced from 249 hours to just over 13 hours – or about a 95% reduction.

Now that tape data is easily accessible, it is the most economical source for collection, and also the most reliable since data on backup tapes is not easily tampered with.

As the Burden Erodes

Who wants to be in the position of the three plaintiffs who entered Judge Scheindlin’s court during the Bank of Montreal Pension Plan case? When the plaintiffs claimed burden in collecting ESI for the case the Judge threw out their claim, and much to their chagrin, required the relevant data be presented. Similar is the opinion from the Starbucks vs. ADT Security case. ADT attempted to use the burden argument citing \$834,285 to collect five custodian mailboxes. The court declined to find the information at issue “not reasonably accessible” and ordered the production of the ESI.

The courts are becoming more and more educated when it comes to technology. They know what is possible and have a good idea what it costs to collect. Taking a chance that a case will not have a highly educated judge is a risk. Hence the use of the burden argument is fading away.

About the Author:

Jim McGann is Vice President, Information Discovery – Index Engines. He has extensive experience with the eDiscovery and Information Management and is currently contributing to the Sedona Conference Working Group addressing electronic document retention and production. McGann is also a frequent speaker for industry organizations including ARMA and ILTA, and has authored multiple articles for legal technology and information management publications.