

New Technology Alters the Terrain on Accessibility of Backup Tape Data

Constantine & Aborn Advisory Services, LLC

*Gerard Britton, Director of Investigative and
Compliance Services*

*Richard Davis, Director of Litigation Risk
Management Services*

By now, most litigators are aware that under the amended Federal Rules of Civil Procedure (FRCP) the e-discovery of electronically stored information (ESI) contained within backup tapes is generally considered presumptively unduly burdensome because of the relatively high costs associated with accessing their contents. Courts have considered e-discovery of such tapes a special case, requiring procedures such as sampling to determine the probable existence of relevant information as well as judicial determinations regarding the merits of restoring tapes and the apportioning among parties of the costs associated with these restorations. However, the inaccessibility of backup tapes was never set in stone, and recent advances in technology promise to drastically alter how courts view the accessibility of ESI on backup tapes. Because technological advances may soon fundamentally undermine the what-happens-in-Vegas-stays-in-Vegas notion for backup tapes, attorneys need to stay abreast of these technological changes and evolve their legal practices and advice in response to them.

Backup Tapes are Considered Inaccessible

Consistent with the opinion in the now famous *Zubulake I* decision (*Zubulake v. UBS Warburg LLC*, 217 F.R.D. 309 [S.D.N.Y. 2003]), courts have repeatedly reaffirmed the presumption that ESI contained on backup tapes is inaccessible: "Backup tapes must be restored using a process similar to that previously described [...] That makes such data inaccessible." This is true for several reasons. Data is recorded onto tapes in a linear fashion, sometimes in multiple forward and reverse passes. Generally, tapes must be restored cover-to-cover at slow speeds, then indexed before the data can be examined or searched. Restorations often result in the production of multiple copies of the same documents. In addition, tapes handle data in blocks, which requires a great deal of memory and processor resources in the restoration process. As the court in *U.S. v. Amerigroup*, 2005 WL 3111972 (N.D. Ill.) noted in discussing the difficulties of retrieving data from backup tapes, "...[t]o be sure, one can imagine the use of three dedicated servers to perform each of the six weeks of restoration work concurrently, but the end result is still eighteen weeks of man-power and eighteen weeks of use of the necessary equipment. That burden, which is undeniably substantial, exists independently of the monetary costs entailed." In other words, generating data from tapes to determine the existence of any responsive ESI is a costly, resource-intensive undertaking.

As recent decisions have confirmed, the discovery of backup tape ESI is subject to the balancing analysis, which requires an examination of seven factors: the specificity of the request; the likelihood of availability of ESI from more accessible sources; the failure to produce once-available ESI; the likelihood of finding responsive ESI; the importance of the information; the nature of the issues at stake in the litigation; and the resources of the parties. To overcome the cost of restoring all data to determine if a tape contains responsive ESI, courts have employed creative workarounds. These include sampling methodologies, as in *Zubulake I*, and phased restorations, as in *AAB Joint Venture (AAB Joint Venture v. U.S.)*, 2007 WL 646157 [Fed. Cl.] Both amount to a hit-or-miss endeavor, and neither obviates the cost of restoration when samples uncover relevant ESI.

A review of the relative costs for processing online data versus offline data provides an illustration of the current cost disparities. For a network in which the home and shared directories on the file servers yield 100 MB per user, the cost

could exceed \$750,000 to defensibly identify, gather, and conduct preliminary analysis for a hypothetical 500 custodians.

With offline data, the same processing costs will be incurred once the data is restored. However, the restoration process incurs significant additional cost. The data for the same 500 users may be on four years of tapes. Assuming there is less than 250 GB of data on the backup tapes that constitutes a target set of data, based on the disaster recovery tape backup protocol that was used; the data can be spread out across 5000 tapes. Moreover, IT will probably not know with certainty which of the 5000 backup tapes they have contain the information they need. This presents the following cost parameters: It costs approximately \$500 to restore each tape, assuming a full restore of each tape. If all 5,000 tapes need to be restored, this cost could exceed \$2,500,000 before data processing even starts. Then there's the cost for data processing. Following restoration, there could potentially be 100 instances of the same file restored for all 500 custodians and a de-duplication process must be executed. One can easily see that retrieval from backup tapes can be extraordinarily expensive.

The result of the cost and obstacles to retrieving backup tape ESI has been a prevailing sentiment that the potential smoking guns (or, depending on the viewpoint, the latent liability) will generally remain safely buried.

Emerging Technologies May Soon Make Tape ESI Readily Accessible

Innovation, however, may make backup tape ESI more accessible. Index Engines, an enterprise discovery solution provider in New Jersey, has announced a new tape review technology. In essence, the Index Engine appliance performs full content and metadata indexing at the speed of tape, including email server tapes—without ever restoring them. The resulting index can be keyword-searched, and the system provides de-duplicated index information for all responsive documents found on tape. Tapes can then be surgically restored based on the tape location information provided by the appliance. At the very least, the process provides for a more robust methodology to conduct *Zubulake I* sampling. Moreover, it provides a much less costly and time consuming method to extract responsive ESI from backup tapes. It is not difficult to envision a court, when confronted with the appropriate circumstances, holding that the reduced cost and resource requirements of the appliance render production of tape ESI not significantly more burdensome than active online ESI. The court in *Analog Devices, Inc. v. Michalski*, 2006 WL 3287382 (N.C., Super. Ct.) noted this reality, "As technology changes and new factual situations arise, the law in this area will develop and change."

Even more significant, this development strongly suggests that we are perhaps only a step or two from an even more powerful tape tool, one that will, in a single high-speed pass, record and store only de-duplicated responsive ESI that has been culled automatically by selected concept or keyword search criteria. A tool of this kind would crack open the vault of backup tape ESI.

The arrival of this type of technology requires fundamental changes in the litigation readiness and response strategies of corporations and counsel. The technology enables parties to potentially capture single instances of files merely by passing the tape. It also eliminates the need to do a full restore of an Exchange server to locate relevant email and attachments.

Questions abound. Could class-action plaintiffs' counsel successfully argue against a finding of undue burden and gain court imprimatur to have defendants apply the tool to their backup tapes as part of discovery? Should corporations proactively use the technology to more aggressively subject tape ESI to document retention policies, or should they wait until they are forced to do so in litigation? Will the technology enable responding parties who have been ordered to retrieve otherwise unavailable emails from tape to do so less expensively?

Time will tell whether the technology delivers on its promise to render backup tape ESI susceptible to routine discovery. Clearly, corporations and counsel need to begin addressing the possibility today.

About the Authors

Gerard J. Britton, Director of Investigative and Compliance Services for Constantine & Aborn Advisory Services LLC, manages the design and implementation of monitoring and compliance programs and oversees the firm's investigative services. Mr. Britton also works closely with the firm's police and criminal justice consultancies. Britton is also Of Counsel to Constantine Cannon LLP.

Richard E. Davis, Director of Litigation Risk Management Services, specializes in litigation readiness planning, e-discovery cost containment and data lifecycle management strategies for Constantine & Aborn Advisory Services LLC.