September 2007

FindLaw®

Electronic Discovery: It's Not Just About E-mail

by Craig S. Mullins

A recent survey of corporate counsel found that the typical U.S. company faces an average of 305 lawsuits and spends $12 million a year on litigation alone, not including settlements or judgments.

In this technological age, businesses have migrated to storing corporate data purely electronically. As such, there has been a strong focus on e-discovery, led by the recent changes in the Federal Rules of Civil Procedure (FRCP). Companies understand the importance of this data and the need to preserve it, however many are not properly retaining and securing structured data for business and legal protection.

Most organizations focus on enacting policies and procedures to retain their e-mail documents, because many of the highly publicized e-discovery cases have relied on information discovered in e-mail communications. As such, it is wise to tackle e-mail first, but it is critical not to stop there. Electronic evidence could encompass any type of data stored anywhere (ref. Section V of the FRCP).

An e-mail archiving and retention program is important, but is insufficient on its own to ensure your company is prepared for an e-discovery request. You should consider data located on the network, on various servers, hard drives, shared drives, laptops and PDAs, smart phones, and even on backup tapes.

With this in mind, consider for a moment what type of data is likely to be the most important to the on-going business of your company – the transactional data stored in database management systems. It is this data that must be readily accessible in the event of an e-discovery request. Unfortunately, little focus is given to database

data, at least in terms of preparation for e-discovery .This is the data that runs the business – from tracking insurance claims and managing brokerage trades, to handling deposits and withdrawals from bank accounts.

Indeed, lawsuits can hinge on the accuracy of this data. It is not sufficient to save statements and reports generated from transaction data, because the accuracy of the data on the statement can bequestioned by the suit. It is imperative that additional steps be taken to archive and retain transaction data for the same reasons we archive and retain e-mail documents.

Of course, the amount of transaction data to be managed can be quite voluminous. Maintaining data in operational databases over long periods of time creates problems— transactional performance problems as data volumes expand, and data authenticity problems for data that must be maintained for legal purposes. As such, data must be periodically archived from operational databases.

But what do we mean by the term "database archiving?" Database archiving is the process of removing selected data records from operational databases when it is likely they won't be referenced again, and placing them into an archive data store where they can be retrieved if needed.

As simple as this may sound, there are many significant challenges and requirements posed by database archiving. Perhaps the most important consideration is that archived data must be hardware and software independent. When data retention requirements span over decades, the production system from which the data was archived may no longer exist – at least not in the same form, and perhaps not at all. For example, data, or music on an eight track would be nearly unusable today. Many people would have issues even accessing data on a cassette tape.

The archive also must be able to store a large amount of data. As we store more data, we will archive more data, but combining this with long mandated data retention periods provides some significant challenges.

The archive must be able to manage data for very long time periods. Many data retention requirements are stated in decades, and as a result, the archived data may outlive the systems and the programmers that generated them. Because data structures change over time, the archive must be able to support multiple variations of the data structure as it changes.

And importantly, data must remain unchanged once it is archived. The archive must be able to protect against data modification. Only "read" access should be available to the archived data, and the data must be guaranteed to be authentic. Mechanisms to prevent surreptitious modification are necessary too.

Finally the archive requires metadata to be useful. The ability to understand the type of transaction data being managed by the archive is vitally important because the types of data that could be archived are as varied as the number of individual businesses and applications in use. Furthermore, chain of evidence metadata is required to prove the authenticity of the data in the archive.

But we create the archive for transaction data because we may someday need to access it for discovery. As you know, culling data for e-discovery can be challenging when trying to locate just a small subset of the applicable data among gigabytes or even petabytes of information. The archive solution must allow for the review of each type of business transaction in context with its metadata.

Taking all of these considerations into account, a secure, durable archive data store must be used to retain data that is no longer needed for operational purposes, and it must enable query retrieval of the archived data in order to cull transactions for discovery.

From FindLaw, September 2007.

DB2PORTAL.com