New York Law Journal
A NEW YORK LAW JOURNAL SPECIAL SECTION
ALM
E-DISCOVERY
Web address: http://www.nylj.com
MONDAY, NOVEMBER 5, 2007

Know Your Data

Creating a map to help identify ESI.

BY H. CHRISTOPHER BOEHNING, DANIEL J. TOAL AND ROSS GOTLER

ON MARCH 5, 2007, Intel disclosed that "a number of inadvertent mistakes" on its part may have resulted in the destruction of e-mails required for production in an antitrust suit filed by its competitor AMD.[1] Although Intel declared that it had moved quickly to preserve relevant documents after the suit began, some of its employees incorrectly assumed that their e-mails were being automatically archived, resulting in the automatic deletion of those e-mails, while several hundred Intel employees "identified as having potentially relevant information were not instructed to retain those documents."[2]

It also appeared that some of the employees who had been directed to move e-mails to their hard drives failed to comply with these instructions, resulting in further e-mail loss. As a result, Intel indicated that it would implement a new data retention system and that it might be able to retrieve data from its backup tapes. AMD demanded a judicial investigation, citing "a combination of gross communication failures, an ill-conceived plan of document retention and lackluster oversight by outside counsel."[3]

These events highlight the dangers and potential consequences that arise when companies need to respond to large-scale requests for electronically stored information. And these difficulties are often compounded by the fact that most companies and outside counsel enter into litigation with little if any understanding of the company's information systems and data management policies. This is so even though the 2006 amendments to the Federal Rules of Civil Procedure and recent judicial trends make clear that lawyers and their clients are now expected to be familiar with the client's technological infrastructure at the outset of litigation, and preferably before.

Recently created requirements and guidelines from various jurisdictions also highlight the importance of understanding a client's information systems. The U.S. District Court for the District of Delaware now requires that each party in a dispute designate an individual to whom all e-discovery requests are to be made, designating this person "the e-discovery liaison."[4] The liaison is required to be "familiar with the party's electronic systems and capabilities" and is also expected to be well-versed in electronic data storage and formatting.[5] Other jurisdictions have also followed suit, suggesting that such policies may become increasingly common.[6]

As a result, creating and understanding a data map of a company's information systems and related policies can be an excellent first step in managing electronic discovery.

What Is a Data Map?

A data map, in the context of electronic discovery, is a detailed representation of the type and location of all electronically stored information (ESI) throughout a company that may be relevant to electronic discovery (see sample below). Examples of systems that can be mapped are:

- servers—active and dynamic data, such as file servers, e-mail and voice-mail servers;
- data management systems—backup tapes, financial systems and disaster recovery systems;
- endpoints—desktops, laptops, BlackBerry devices and cell phones;
- portable media—flash drives, hard drives, CDs and DVDs;
- data hosted by third party vendors—payroll systems, and junk mail filtering services.

How to Create One

Depending on the size and complexity of the company, creating a data map may be relatively straightforward or be more complicated and require significant investigation and diligence. Information systems can be very complex, and it is important to determine how company employees and agents interact with these systems in terms of storing ESI.

When creating a data map, a good starting point is a company's chief information officer (CIO). The CIO should be able to provide an overview of the company's information systems, including a network architecture diagram (see sample on facing page) that identifies the various systems, servers and data endpoints internal to the company. The CIO should also be able to identify any systems external to the company that contain the company's ESI, the status of ESI generated by previous employees, the handling of ESI located on legacy systems, and the procedures used for storing ESI during technical upgrades.

The next step is to work with the CIO and other company IT professionals to determine what types of ESI are located on each server or system. Common ESI types include e-mail, electronic documents such as word processor or presentation program files, voicemail, instant messages and financial data. A spreadsheet program can be helpful in organizing the information gathered, and will help to begin cross-referencing ESI type to location.

The CIO can also be a useful resource for explaining employee interaction with information systems and the company's standard policies with respect to ESI. It is key both to understand official policies and what employees actually have the ability to do with ESI on a daily basis.

For example, a company's policy might call for a quota on e-mail inbox size and auto-deletion of e-mail messages older than 30 days. However, employees may also have the ability to create their own e-mail archives on the various network and computer locations to which they have access, creating a large pool of potentially relevant e-mail that exists outside of the company's e-mail servers. Adding this information to the data map will help to develop a more complete and realistic picture of the client's ESI and systems.

Finally, an essential element of creating a data map is identifying the active data retention policy and any backup and disaster recovery policies for each ESI type and location.

For example, e-mails may have an active data retention policy of automatic deletion after 30 days unless otherwise archived by an employee. Alternatively, e-mail may be replicated daily to a disaster recovery "hotsite" and backed up onto tapes used pursuant to a detailed rotation policy.

Because these policies can be complicated and may vary by both ESI location and type, it often is useful to speak with the company CIO, the chief compliance officer, or other IT personnel to fully understand them. All relevant information should then be added to the data map.

Conclusions for Outside Counsel

Creating a comprehensive data map can be a daunting and time-consuming task.

And, for many organizations, the scope of the project may be too much to handle at one time. But companies and their outside and inside legal advisors should consider the possible benefits of data mapping, even if the first step is to create a partial data map (e.g., one that focuses exclusively on e-mail systems and retention periods).

H. Christopher Boehning *and* **Daniel J. Toal** *are litigation partners at Paul, Weiss, Rifkind, Wharton & Garrison LLP.* **Ross Gotler** *is the firm's senior manager of practice support, focusing on e-discovery.* **Bilal Faruqi,** *an associate at the firm, assisted in the preparation of this article.*

There is never a good time to prepare a data map, but experience suggests that starting the process before litigation hits can give the company and its legal advisors more time to focus on the merits and legal strategy.

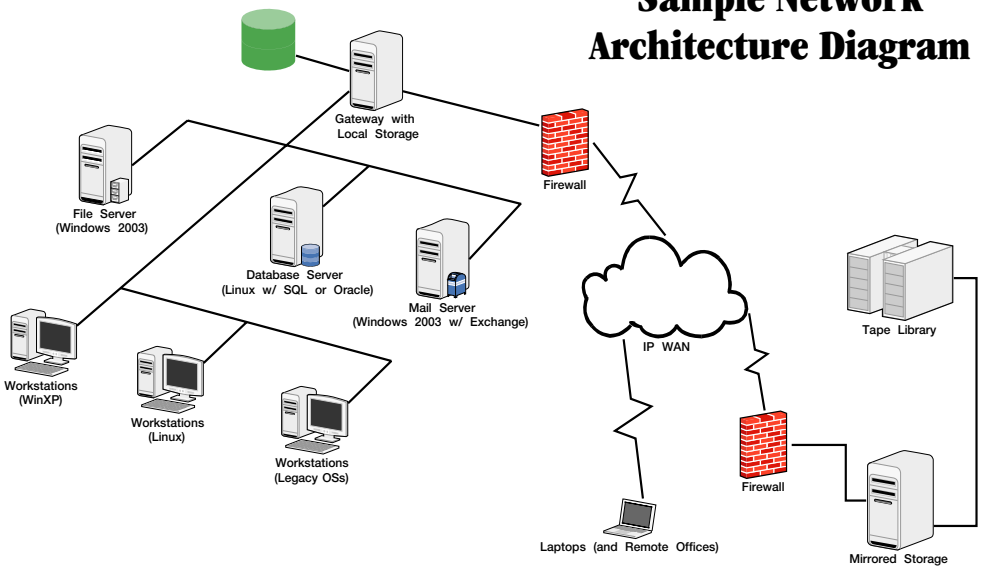
Indeed, the challenges faced by Intel should serve as a reminder to outside counsel of the benefits of working closely with their clients before litigation to create data maps that will allow them to quickly and fully identify and locate any ESI that may be relevant to a litigation or investigation. This will enable the client to proceed in a defensible manner with the next steps of the electronic discovery process—preservation and collection of potentially relevant ESI—and enable outside counsel to represent accurately to judges and regulators the status of discovery in a matter.

As Judge Shira Scheindlin of the Southern District of New York has commented regarding electronic data:

"The one thing you don't want to do is make representations to a court that you must eventually retract."[7]

1. Associated Press, "Intel May Have Lost Emails for AMD Lawsuit" (Mar. 6, 2007), available at http://www.cnbc.com/id/17472365.
2. See id.
3. See id.
4. U.S. District Court for the District of Delaware, Default Standard for Discovery of Electronic Documents, available at http://www.ded.uscourts.gov/Announce/Policies/Policy01.htm.
5. See id.
6. See U.S. District Court for the District of Kansas, Guidelines for Discovery of Electronically Stored Information, available at http://www.ksd.uscourts.gov/guidelines/electronicdiscoveryguidelines.pdf (addressing the specific electronically stored information that should be discussed during a Federal Rule 26(f) conference); see also U.S. District Court for the District of Maryland, Suggested Protocol for Discovery of Electronically Stored Information, available at http://www.mdd.uscourts.gov/news/news/ESIProtocol.pdf (providing detailed suggested guidelines for activities related to electronic discovery).
7. The Sedona Conference, Interview of Judge Shira A. Scheindlin (March 24, 2004), at 4, available at http://www.thesedonaconference.org/content/miscFiles/ScheindlinInterview.pdf.

Sample Network Architecture Diagram



SAMPLE DATA MAP

Items in shaded cells are outside the normal course of business	Data Type									
Data Location	E-mail, Calendar, and Tasks	E-mail Attachments	Address Book	Work Product	Voicemail	Instant Messages	Financial	**Active Data Retention**	**Backup***	**Replication to Disaster Recovery Site**
C: Local hard drive Can include archived e-mail and attachments	X	X	X	X		X		Retained until deleted by user or otherwise destroyed or overwritten in the normal course of business.	None	None
G: Network drive (Group drive)	X	X	X	X		X	X	Retained until deleted by user or otherwise destroyed or overwritten in the normal course of business.	Standard full backup	Every 24 hours
J: Home directory	X	X	X	X		X		Retained until deleted by user or otherwise destroyed or overwritten in the normal course of business.	Standard full backup	Every 24 hours
L: Network drive (Shared drive)	X	X	X	X		X		Retained until deleted by user or otherwise destroyed or overwritten in the normal course of business.	Standard full backup	Every 24 hours
E-mail servers Active user mail files	X	X						Mail automatically deleted after 60 days unless archived. Mail files are deleted 14 days after a user leaves the firm.	Full backup Monday to Friday. Backup tape sets are rotated every seven days.	Every 6 hours
Voicemail system					X			Retained until deleted by user or otherwise destroyed or overwritten in the normal course of business.	None	None
Litigation hold servers	X	X		X				Retained as directed by litigation response committee.	Full backup Monday to Friday. Backup tape sets are rotated every seven days.	None
Litigation hold backup tapes Backup tapes held by vendor	X	X		X				Retained as directed by litigation response committee.	None	None

**Standard full backup: 20 daily full backup tape sets for Monday through Thursday are rotated every five weeks. Four weekly full backup tape sets for Friday are rotated every four weeks. The weekly tape set for the Friday closest to the end of the month is set aside as the monthly backup. The monthly backup is retained for six months and then returns to the rotation or is destroyed.*