ON-PREMISE VS. HOSTED STORAGE

DOES IT REALLY MATTER WHERE MY ORGANIZATION'S ESI IS KEPT?

By Bill Tolson

Bill Tolson

Where organizations store their electronically stored information (ESI) is of no concern to attorneys … right? The answer is, yes and no. There's an on-going debate over the question of where the "best" place is to store a company's ESI for legal reasons; in the company's own facility (on-premise) or in someone else's facility (hosted—also known as "storage as a service" or SaaS).

There are three main questions to ask yourself when considering the question: *From a legal perspective, where's the best place to store the organization's ESI?* First, is my ESI secure and can I prove it has not been altered in any way? Second, can I access my ESI quickly enough to place legal holds and perform searches? Third, do I have access to the full ESI data set?

Let's review some of the main points to keep in mind when dealing with ESI in litigation. First, when litigation is reasonably obvious, you have a responsibility to immediately protect all ESI that *could* be responsive in the approaching civil case. This responsibility is an absolute requirement in U.S. Federal court and many state courts. There are few if any excuses a judge will accept for inadvertently deleting potentially responsive ESI after your legal hold responsibility has been triggered. Second, the timeframe you'll have to fully respond to an eDiscovery request is generally much shorter now than in the past (e.g., pre December 2006). Quick access to all potentially responsive data is extremely important to make these judicial deadlines. Third, good intentions, proactive eDiscovery planning and documentation will mean something to the judge (possibly) if you have an inadvertent ESI deletion.

If we can't find it, it doesn't exist … right?

Before we jump into the question of the best place to store your organization's ESI and the advantages and disadvantages of on-premise verses hosted ESI storage (archiving), let's take a look at how many organizations currently handle their own ESI, and what that means for litigation support.

Most employees store their ESI haphazardly; that is, without a great deal of thought as to making it easier to find later, sometimes much later. Employees have many options when it comes to ESI storage; desktop/laptop hard disks, external hard disks, USB thumb drives, CDs, DVDs, personal email accounts, and third party storage facilities are some of the better known examples.

The above-mentioned employee ESI storage possibilities dramatically increase risk and complicate eDiscovery due to the simple fact that there are so many more locations responsive ESI can exist. Now multiply those possible storage locations for one employee by tens, hundreds or thousands of employees and you begin to see the enormity (and cost) of the problem.

The first key to effective ESI management is to know what you have and where it's stored … it's as simple as that. Giving employees carte blanche on storage possibilities may be the accepted corporate culture but is not the most cost conscious strategy for litigation support. With that in mind, what storage strategies should organizations consider to reduce their litigation risk and support costs? The simple but difficult answer is to control those locations in which employees can store their ESI. For example, I worked with a large bank who limited the

continued on page 6

locations employees could store their information to networked file system locations (shared drives). This strategy allowed the bank to control the number of locations employee ESI could exist.

Archiving systems are an excellent way to control and manage employee ESI, especially for litigation support.

Archiving verses Storage: What is archiving and why is it important in eDiscovery?

The difference between storing your ESI and archiving your ESI is straight-forward: storing a file saves a copy of the file as it was when you hit the "save" button. The individual has to remember where they saved it and what the file contained. If the file is never accessed again, it could remain on the storage device indefinably.

An archive is a separate system of storage and control. It works with your existing infrastructure such as your email system, your SharePoint system, and your file system to add additional functionality and control over your ESI. For example the email system is not meant to be a long-term storage system even though most employees use it that way. An email system becomes slow and problematic when overloaded with email and attachments. An archive system will offload email from the email server and de-duplicate it, while keeping the email available for employee use. The archive system should also completely index all content improving the search capability. And lastly, the archive system will control and manage the archived email with retention policies insuring email is removed from the archive when it has reached the end of its retention period.

Why is archiving important for eDiscovery? The obvious answer is, if set up and managed correctly, the archive can become the single point of discovery for ESI within the infrastructure. For example, by archiving your email system and prohibiting the use of PSTs (personal archives) by employees, the archive can become the email repository of record.

On-Premise ESI Archiving

Capturing and storing your organization's ESI within your own facility, especially for the purposes of litigation preparedness, has several advantages:

1. Sensitive Company data is kept within your data center (usually the main advantage mentioned by General Counsel).
2. Government agency access is controlled by your organization, not by some third party.
3. Legal holds can be placed immediately.
4. You have more granular control of the archived data including granular retention policies.
5. PSTs can be migrated into the archive to create a more historic archive.
6. Single instancing of ESI ensures there is only one object to keep track of verses thousands or millions.

There are also some perceived disadvantages of on-premise ESI archiving including:

1. Upfront costs. You will have to purchase the software, servers, storage, data center space and potentially additional IT personnel to operate the archiving system.
2. Time to implement the archiving solution will take longer simply because all the above resources have to be ordered and installed first.
3. Providing for annual growth requirements, especially for storage.

The diagram in Figure 1 shows the typical on-premise email archiving solution architecture.

In a typical on-premise email archiving system, email comes in through your firewall and spam filters to your data center and arrives at your email servers where employees have access to it through local applications such as Microsoft Outlook. The archive solution interfaces with your email servers and captures and indexes email, attachments, and other objects throughout the workday. The email within the email system and archived email is stored within your organization's control inside your data center and is always ready to search. Employees can read, store, and delete email from within their mailbox without affecting the "copy of record" emails being managed by the archive system. In discovery, the attorney queries the email archive directly to search for responsive ESI, place legal holds, and start the review process of the potentially responsive ESI.

continued on page 7

Hosted ESI Archiving

A hosted ESI archive is a pretty obvious concept; however, there are two basic flavors. Let's take a look at the email example again. In the first scenario the email server and archive server are both at a third party (SaaS) provider (Figure 2). Email sent to the organization actually travels to the third party's servers, where it is recognized as belonging to the organization and is made available to the specific employee it was addressed to through a client or Web application on the employee's workstation. The email is always controlled within the third parties data center. The email archive will also be set up in the third party's data center available also via a client application.

The second hosted email archiving scenario (Figure 3) is where the email server is within your data center; your employees control and manage day to day email operations but email is also sent to an offsite third party where the archiving is performed. The ESI held within the archive is available via a client or Web application.

The differences between the two scenarios are just a matter of degree. Let's look at the advantages and disadvantages of the two.

Advantages

1. Upfront costs are less than the on-premise solution. There are no capital costs such as archive servers, storage, or additional employees. Usually these hosted solutions are based on a contract for a specified period of time (2-3 years) and are also based on the amount of storage used.
2. The time to implement a hosted solution is much less than an on-premise solution.
3. You usually don't have to hire additional employees to help run the solution.

Disadvantages

1. Your organization's sensitive ESI is held by a third party outside of your direct control.
2. Depending on how you define your retention policies, the cost of a hosted archive could be more than an on-premise archive.
3. Direct access to your ESI at the third party facility may not always be immediate.

continued on page 8

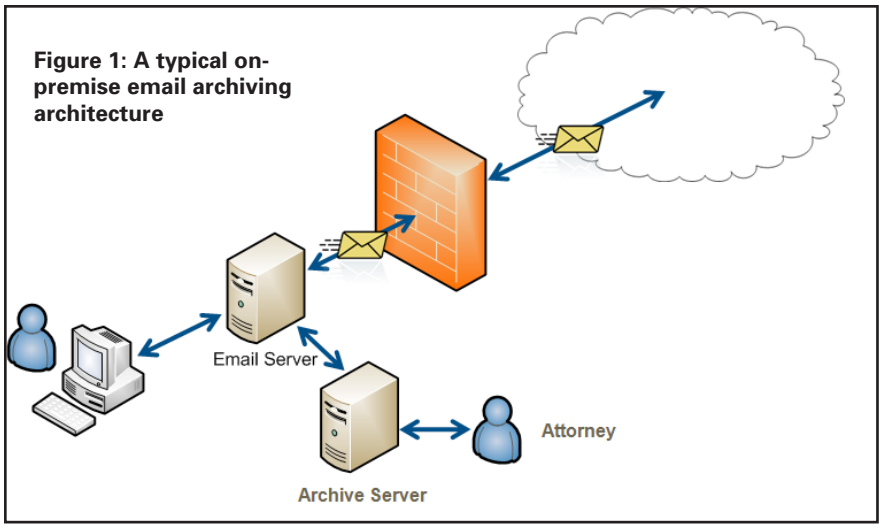

Figure 1: A typical on-premise email archiving architecture

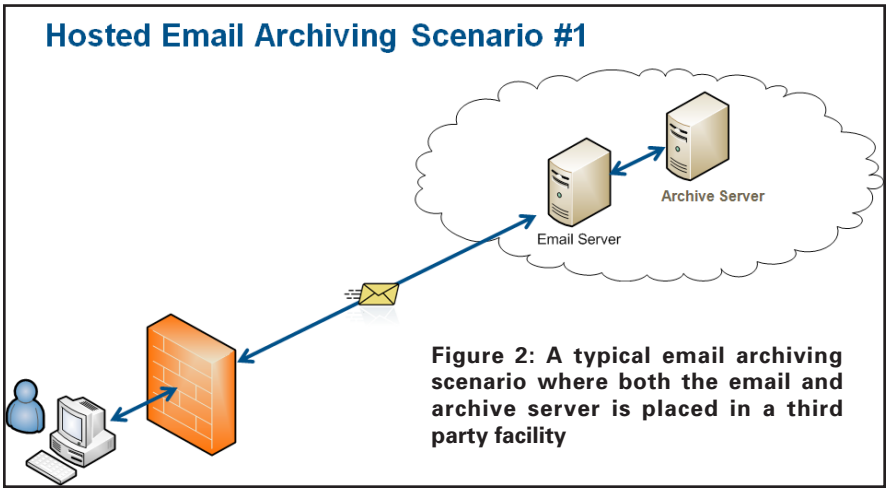

Figure 2: A typical email archiving scenario where both the email and archive server is placed in a third party facility

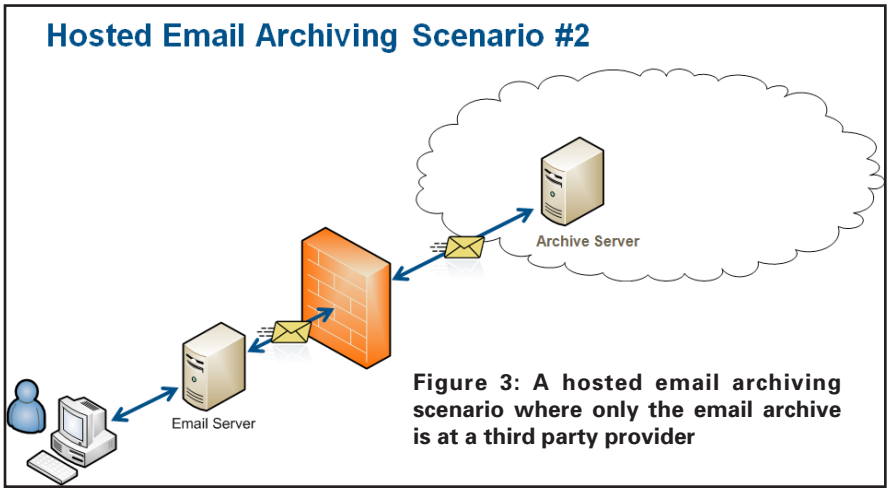

Figure 3: A hosted email archiving scenario where only the email archive is at a third party provider

 CONTINUED FROM PAGE 7

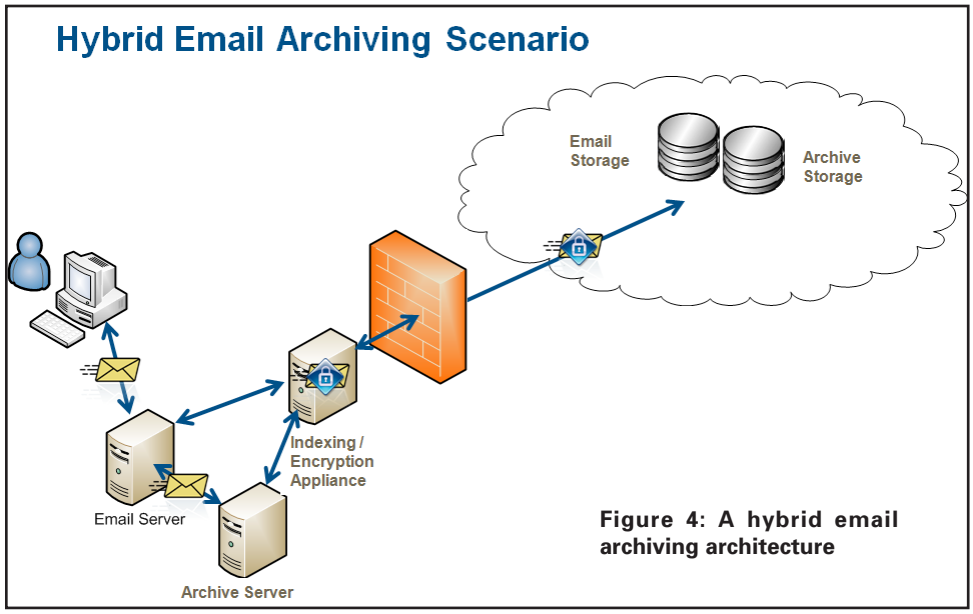

Figure 4: A hybrid email archiving architecture

Hybrid ESI Archiving

The hybrid ESI archiving scenario is relatively new and was created to address the biggest disadvantage of the hosted archiving solution: third party suppliers having access to your organization's sensitive data. In the hybrid model, the email server and archive server are both in your data center controlled and owned by you (Figure 4). But email archive data is encrypted with encryption keys known only by you before the data is sent to a third party for storage; this is usually accomplished by a dedicated appliance that encrypts and controls the transactions.

The other advantages of the hybrid model include local mailbox management and reduced storage costs.

The most obvious question to ask yourself is, which archiving scenario is best for your organization especially considering eDiscovery requirements? The answer is, any one of them can be a great solution for your organization depending on your specific individual requirements and budgets. The main questions to consider are:

1. Is the archiving solution capturing a full ESI data set, meaning that if you choose to archive your email, are you capturing and managing every object type you could be asked for in eDiscovery?
2. Is the archived ESI completely indexed so that when you run a search you find every occurrence?
3. Can you run the searches yourself or do you have to have the third party conduct the search?
4. Do you have access to your archived data immediately?
5. Is the archived ESI managed to your retention policies?
6. When you have to stop deletions because of anticipated litigation, can you do so quickly?

The bottom line is, your organization is responsible for your ESI in litigation/eDiscovery so be sure you're comfortable with the solution your organization chooses.

Bill Tolson *is currently a Director of Product Marketing for Archiving Solutions at Iron Mountain. Bill has more than 20 years experience in product marketing and consulting in both the storage and archiving solutions markets. Previously, Bill was a principal consultant and practice manager for Contoural Inc. where he led the eDiscovery and compliance consulting business specializing in storage solutions, email archiving, enterprise content management, and information lifecycle management. Bill has been a featured speaker at many archiving events including the Government Technology Conferences, AIIM 2009, ARMA, ARMA Canada, LegalTech West and TechTarget's Email Archiving Series. Bill is the author of two eBooks* The Know IT All's Guide to eDiscovery *and* The Bartenders Guide to eDiscovery *as well as the book* Email Archiving for Dummies. *Bill has held senior management positions at Hewlett-Packard, Hitachi Data Systems, StorageTek, and Iomega.*