CLOUD COMPUTING: NEW DISCOVERY CHALLENGES AWAIT

BY ANDREW M. HINKES AND GAVIN C. GAUKROGER

"Cloud computing" has begun to revolutionize business practices by allowing users of cloud services to off-load significant overhead and expenses for information technology (IT) functions while obtaining scalable and flexible computing services that do not depend on a specific location. These advantages,[1] however, come with compromises, many of which create new challenges in the context of civil discovery. This article explains the significance of cloud computing and discusses issues of production obligations, access to data from the cloud, and the increased difficulty of managing civil discovery when information is stored in the cloud.

What Is Cloud Computing?

The National Institute of Standards and Technology defines "cloud computing" as "a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."[2] Simply stated, "cloud computing" generally refers to the new paradigm of gaining access to software, platforms, and services over the Internet.[3] These services are on-demand, pooled, and often rapidly scalable to meet surges in customer power or storage need in a way that is transparent.

Cloud computing services can generally be classified as one of a few types of delivery models: Software as a Service (SaaS), Platform as a Service (PaaS), or Infrastructure as a Service (IaaS).

> SaaS can be explained as "when an internet connection delivers hardware power and software functionality to users regardless of where they are or which computer they are using."[4] Web mail (such as Microsoft's Hotmail and Google's Gmail) and

online office applications (such as Microsoft's Office.com and Google's Google Docs) are well-known examples of SaaS. Web mail has revolutionized e-mail by making the location of the user and the machine used to access e-mail irrelevant. Similarly, Google Docs allows a user to access word processing, spreadsheet, and presentations regardless of the user's location or machine. When using SaaS, the user does not know where the data is being held, who is managing the servers holding those data, or how the data is being delivered. The services are delivered through a password-protected website and, typically, the user does not have to install any software locally in order to gain access to the services. The user has no configuration, management, upgrade, or support obligations; all these services are provided transparently.

PaaS simplifies software development by providing preconfigured "back end" infrastructure components and simplified development tools, greatly reducing the time and labor required to develop new software. PaaS generally does not affect end users of computers, but it is of great interest to software developers.

IaaS essentially outsources the server space, computing processor cycles, and hardware/software maintenance to a third party. Instead of acquiring additional equipment, storage, or connectivity to rapidly increase or "scale up" computing capacity, IaaS allows for transparent and immediate expansion of user access, power, or storage (that is, scalability) in exchange for a fee. Typically, the user has no access to or responsibility for the maintenance of the infrastructure. The best-known example of IaaS is the Amazon Cluster, which sells computing power, capacity, and storage to a variety of companies for various purposes.[5]

How is cloud computing different from a third-party vendor's hosting of a company's data? The critical question is what is kept in-house and what is remote. Many enterprises use so-called co-location facilities to host servers or a redundant copy of their data for continuity of business planning. Although the line can be blurry, typically, a third-party vendor may store only data produced by local users on local machines, data that are primarily held on local servers, from which the data are transmitted to remote vendor systems, or host the actual server providing the services for the client. In the cloud computing model, a client's data can be held in any location to which the cloud provider has access; data may be distributed across multiple machines in multiple countries and may be highly dynamic, but that distribution of data is totally transparent to the user.

Cloud computing is a relatively new technology, and many companies have already embraced it and rely on it to store both critical and sensitive data. Despite the many positive attributes of cloud computing, there are substantial risks associated with storage of business data in the cloud. Careful consideration should be paid to the selection of the vendor, and several factors are of crucial importance, such as—

- the vendor's stability,
- the vendor's disaster recovery plan,
- the vendor's policy regarding access to customer data upon the vendor's bankruptcy,
- acquisition of the vendor's business by another vendor,
- dissolution of the vendor's business, and
- the vendor's willingness to negotiate specific terms of service to ensure compliance with data security and privacy laws.

Finally, regardless of the location in the cloud, liability to third parties for the security and integrity of the data held in the cloud remains with the client, not the vendor. Thus, selection of vendors—particularly an understanding of the clients' rights to gain access to data held by a cloud service vendor—is critical.

Why Does Cloud Computing Matter?

Simply stated, cloud computer is an emerging, fast, and often easier way for businesses to add computing services, storage, and power. An enterprise functioning entirely in the cloud would need to maintain a minimal technology infrastructure. (Essentially, only thin clients or terminals to provide users with access and sufficient local infrastructure to provide wide-bandwidth Internet access would be necessary.) The business would not be responsible for upgrading software or increasing storage space to react to users' fluctuating needs. In exchange for off-loading those responsibilities, the enterprise pays a monthly fee. Thus, using cloud computing can avoid the expense and time required to purchase, configure, and upgrade existing infrastructure each time a new user need arises. Users simply use the service; all storage, processing, and infrastructure are stored, maintained, and upgraded elsewhere. Users do not need to know or care about how and where these services are done. However, as discussed below, the lack of control of the data stored in the cloud, insidious limitations on users' access to data, and dubious contractual relationships among the vendors regarding the rights to access data may complicate the decision a business makes about storing critical or sensitive data in the cloud.

Discovery of Information That Is Electronically Stored in the Cloud

Under the Federal Rules of Civil Procedure, a party in litigation in federal court is obligated to provide to the other parties "a copy—or a description by category and location—of all documents, electronically stored information, and tangible things that the disclosing party has in its possession, custody, or control."[6] To satisfy this obligation, lawyers must prepare for the Rule 26 conference and outline where their client's information is stored and how it can be accessed. In addition, the parties must have already

endeavored to preserve documents from destruction as a predicate to the Rule 26 disclosures.[7] As companies and individuals move their data, applications, and workspaces to the cloud, the obligation to preserve electronically stored information (ESI) and to nimbly produce that data can become increasingly complex. The Federal Rules state the following: "Parties may obtain discovery regarding any non-privileged matter that is relevant to any party's claim or defense—including the existence, description, nature, custody, condition, and location of any documents or other tangible things and the identity and location of persons who know of any discoverable matter."[8] This is a broad requirement that can lead to expensive and extensive discovery practice if not properly managed; with the addition of cloud computing, new questions about access to ESI emerge.

In the context of Rule 26 initial disclosures, a request for production,[9] or a subpoena to third parties,[10] the Federal Rules allow for the requesting party to seek the discovery of ESI in a party's "possession, custody, or control."[11] Challenges can arise when a request for production is served upon counsel for a party seeking production of data held by a cloud vendor. Producing data held by cloud service providers can introduces new and unforeseen complexity to production of ESI.

Getting Data Out of the Cloud: Vendor Service Contracts

When confronted with a request for the production of ESI stored in the cloud, the first questions to be answered are likely to be: How will ESI be produced and who will produce it? Cloud service providers typically provide the answers in their terms of service or privacy policies.

Many cloud service providers disclaim ownership of the data and clearly state that the data stored on their servers remain in the control of the user. For users of the Gmail Web mail service, the relevant terms of its service agreements and privacy policies state: "Google does not claim any ownership in any of the content, including any text, data, information, images, photographs, music, sound, video, or other material, that you upload, transmit or store in your Gmail account."[12]

Similarly, Microsoft's free Internet-based e-mail and applications services such as Windows Live, Bing, MSN, Microsoft Office Live, and Office.com contain a common "Microsoft Service Agreement," which provides the following: "Except for material that we license to you, we don't claim ownership of the content you provide on the service. Your content remains your content. We also don't control, verify, or endorse the content that you and others make available on the service."[13]

In short, even though Google or Microsoft may have "possession" and "custody" of the ESI in the user's Gmail or Hotmail account, these services' contract with the user puts control of the data in the hands of the clients.

Gaining access to the ESI in a cloud if the service provider goes out of business is yet another consideration. When general IT services or data storage are outsourced to the cloud, that data is exposed to loss if a cloud service provider (including downstream providers of cloud services of whom the client may not be) ceases operations or declares bankruptcy. A cloud service provider could terminate the terms of service or even delete the content belonging to the user, which could be disastrous.

If the cloud service provider goes out of business, the user is no longer in control of the data unless backups have been retained or alternative storage arrangements have been made.[14] When using cloud-based services like Gmail and Hotmail, users run the risk that access to their own content may be terminated or that their data may be deleted for lack of use.[15] In fact, both Google and Microsoft disclaim liability for deleting a user's data and maintain their right to deny access and delete content. By way of example, Microsoft clarifies its position as follows: "You're responsible for backing up the data that you store on the service. If your service is suspended or canceled, we may permanently delete your data from our servers. We have no obligation to return data to you after the service is suspended or canceled. If data is stored with an expiration date, we may also delete the data as of that date. Data that is deleted may be irretrievable."[16]

Given the risk of deletion of data by the service provider, users of cloud-based services may need to back up their data in anticipation of litigation or issue a valid litigation hold letter to their service provider to mitigate the risks of a spoliation motion down the road.[17]

Legal Protection for Users of Cloud Computing

A party is entitled to the protections of the Federal Rules, which provide: "A party need not provide discovery of electronically stored information from sources that the party identifies as not reasonably accessible because of undue burden or cost."[18] Depending on the scope of the document request, the position of the requesting and producing party, and other variables, backing up data from a third-party cloud service provider may or may not pass the "undue burden or cost" threshold sufficient to withstand sanctions for failing to do so. In those situations, courts may order that discovery not be had or that the parties allocate the costs of that discovery to even out the burden of production. [19]

Under the Federal Rules, further protections are available to litigants who rely on cloud computing for storage or processing of data. When data is destroyed as a matter of course based on the routine operation of an electronic system—as happens frequently when companies archive e-mail or mandate deletion of certain data after a period of time—a party may seek the protection of the "safe harbor" provisions of the Federal Rules.[20] Generally, a court will not sanction a party "for failing to provide electronically stored information lost as a result of the routine, good-faith operation of an electronic information system."[21] This protection is not without limitation, however, and courts have ordered sanctions in cases in which ESI held in storage devices (such as flash drives) that are easily backed up is lost or damaged.[22] Similarly, this safe harbor is expressly limited to the routine operation of electronic information systems and does not apply to the habits or behavior of users of the systems; the court is likely to consider evi-

dence to make a factual finding of a "routine" if the issue of spoliation is raised and the safe harbor provision is offered as a defense.[23]

Retention of Data Held in the Cloud

While cloud-based services can offer users the efficiencies and economies of outsourced IT functions, a party may be liable for sanctions if the cloud service provider loses the data. To that end, one of the ongoing challenges users of cloud services must address is the issue of data retention and destruction. When negotiating with cloud service providers, a company should do the following:

- consider the efficacy and scope of its current document retention plan,
- determine whether the current plan includes terms of retention for data held in the cloud, and
- examine the vendor's data retention and destruction policies.

If the cloud service client's document retention plan is not clearly identified and considered during negotiations, or if the plan is not compatible with the cloud service provider's service agreement, the cloud service provider's document retention plan or limits on data will supplant the company's intended plan (because it has actual custody and possession of the data). Failure to address this issue could have lasting deleterious effects. For example, a company's document retention plan may allow for e-mail data to be stored in short-term backup storage for a period of one year, after which the data is automatically backed up on long-term archival storage that is in an inaccessible format and deleted from the active system by an automatic system rule set up on the e-mail server. Under that data preservation regime, the company may be able to seek the protections of the Federal Rules to avoid the cost of producing inaccessible ESI.[24] However, if that same company moves its data storage to the cloud and the cloud service provider has a different data destruction schedule (six months instead of one year, for example), the company may find that it is unable to access its own data despite maintaining its own data retention schedule. This can be particularly problematic in regulated industries with strict data retention requirements.[25]

When data is stored in the cloud, the information may never be converted or made inaccessible. Although this feature may be considered a benefit to some, the scope of potentially discoverable information that a client may need to review and produce could grow exponentially or be reduced in a problematic manner when the party is served with a request for documents. To prevent unexpected loss of data and to avoid allegations of spoliation, businesses must review the terms of service and retention periods for any cloud vendor that may store data that is critical to the business and must also ensure that the respective data retention and destruction policies do not create conflicts or contain problematic inconsistencies.

Other Concerns

The very nature of cloud computing invites troubling complexity. If a software as a service system is built on a platform as a service system, which itself is built on an infrastructure as a service system designed to scale up the combined system's storage capacity rapidly, which data retention and data destruction policy governs? To answer this question requires a company to analyze each of the respective vendor's terms of use and data retention policies to determine who has the data at any time and what the retention policy for that vendor is. Obviously, with multiple parties handling a company's data, there is an increased risk of data loss resulting from miscommunication or a risk of delay in acquiring data because of the need for additional discussion among vendors.

In addition, users of cloud services should carefully evaluate the ability to enforce a litigation hold on their data when they are stored in the cloud. The ability to access data in a timely manner is essential when responding to discovery requests and should be considered when moving traditional IT functions to the cloud. In negotiations with cloud service providers, the business should treat document retention needs as a foremost priority. As in all situations, companies should use an integrated team of management, legal, and IT personnel to assist with the implementation of litigation holds and document preservation as well as for authentication purposes as if those functions were maintained in-house.

Cloud computing, of course, complicates this further. Who is responsible for cloud service downtime or the loss of data caused by downtime? If the SaaS system is built on other cloud models—for example, an SaaS system that relies on an IaaS system to scale up its storage capacities—are the IaaS vendor's data retention and destruction policies compatible with those of the SaaS vendor? Are the policies compatible with those of the client? In the case of a vendor using a proprietary system, how quickly can an SaaS vendor produce client data to its client in a format that is compliant with the Federal Rules? Is that procedure laid out in advance for the client?

Accommodating the new issues may include negotiating the availability of consultants and IT personnel in advance or requiring indemnification from the cloud service provider for network downtime or data loss. In the same vein, limiting access to privileged and confidential data should be considered when appropriate. Creating a list of go-to individuals who understand where the data is stored, how to access the data, and how to preserve the data for production is essential when the IT functions of a company are outsourced to the cloud service provider. Finally, as in general document management practices, running litigation hold "fire drills" is a critical means of preparing for a litigation hold event and assessing institutional hold competencies; users of cloud computing should involve cloud service vendors in their fire drills in order to assess the vendors' responsiveness and the quality of result.

Conclusion

Business use of cloud computing is on the rise. Vendors continue to parry for market share as new services emerge

and threaten the basic model of the "in-sourced" IT department by providing rapid scalability and distribution of services. In the civil litigation context, the use of cloud computing to store data that are critical to the business invites new complexity and new challenges. However, consultation with cloud service vendors and close attention to the vendors' terms of use, data retention polices, and litigation hold policies can help businesses take advantage of cloud vendors while minimizing risk. **TFL**

Andrew M. Hinkes is an associate with Berger Singerman in Fort Lauderdale and concentrates his practice in complex commercial litigation focusing on business relationships. He also advises clients regarding document retention issues, management of electronically stored information production, and conducting incoming electronic discovery analysis for complex litigation in state and federal litigation. Gavin C. Gaukroger is an associate with Berger Singerman in Fort Lauderdale and concentrates his practice in commercial litigation, in both state and federal courts, and business dispute resolution.

Endnotes

[1]*See* en.wikipedia.org/wiki/Colocation_centre (accessed Dec. 23, 2010). Collocation centers typically offer other benefits, such as more robust power grids, advanced fire protection and disaster protection schemes, enhanced physical security, and fast bandwidth.

[2]*See* csrc.nist.gov/groups/SNS/cloud-computing/index.html (accessed Dec. 4, 2010).

[3]*See* David D. Cross and Emily Kuwahara, *E-Discovery and Cloud Computing: Control of ESI in the Cloud*, 1 EDDE Journal, (Spring 2010), citing Mark L. Austrian and W. Michael Rya, *Cloud Computing Meets E-Discovery*, 14 Cyberspace Lawyer 7 (2009).

[4]Daniel J. Buller and Mark H. Wittow, *Cloud Computing: Emerging Legal Issues, Data Flows and the Mobile User*, Landslide (American Bar Association, Nov. 2009).

[5]*See* aws.amazon.com/what-is-aws/ (accessed Dec. 4, 2010).

[6]*See* Fed. R. Civ. P. 26(a)(1)(A)(ii).

[7]Pursuant to Fed. R. Civ. P. 26(a)(3)(b), parties are obligated to make initial disclosures, including "a copy—or a description by category and location—of all documents, electronically stored information, and tangible things that the disclosing party has in its possession, custody, or control and may use to support its claims or defenses, unless the use would be solely for impeachment;" as required by Fed. R. Civ. P. 26(a)(1)(a)(ii), within 14 days of the Rule 26 conference.

[8]*See* Fed. R. Civ. P. 26(b)(1).

[9]*See* Fed. R. Civ. P. 34 (a)(1)(A); as to electronically stored information, *see* Fed. R. Civ. P. 34 (b)(2)(E).

[10]*See* Fed. R. Civ. P. 45(d)(1).

[11]*See* Fed. R. Civ. P. 26(a)(1)(A)(ii) and Fed. R. Civ. P. 34(a)(1).

[12]www.google.com/mail/help/legal_notices.html (accessed Dec. 6, 2010). *See also* mail.google.com/mail/help/about_privacy.html (Gmail specifically retains multiple backup copies of users' e-mails for recovery and restoration for limited periods of time: "Even if a message has been deleted or an account is no longer active, messages may remain on our backup systems for some limited period of time. This is standard practice in the email industry, which Gmail and other major webmail services follow in order to provide a reliable service for users. We will make reasonable efforts to remove deleted information from our systems as quickly as is practical.")

[13]explore.live.com/microsoft-service-agreement?ref=none (effective Aug. 31, 2010).

[14]Google's terms of use include a provision that, in the case of a merger, acquisition, or sale of assets, Google will "ensure the confidentiality of any personal information involved in such transactions and provide notice before personal information is transferred and becomes subject to a different privacy policy." *See* www.google.com/privacy/privacy-policy.html (accessed Dec. 24, 2010).

[15]Gmail's terms of use include provisions allowing Google to stop providing services to users at Google's sole discretion without notice (see § 4.3) and allowing Google to disable access to services and acknowledging that disabling an account will deprive the user of access to the data in that account (see § 4.4). *See* www.google.com/accounts/TOS (accessed Dec. 24, 2010).

[16]explore.live.com/microsoft-service-agreement?ref=none (Effective Aug. 31, 2010).

[17]However, a user should be aware of the cloud users' retention terms; as discussed herein, conflicts between differing policies may result in exposure to sanctions.

[18]*See* Fed. R. Civ. P. 26(b)(2)(B).

[19]*Zubulake v. USB Warburg, LLC,* 217 F.R.D. 309 (S.D.N.Y. 2003).

[20]*See* Fed. R. Civ. P. 37(e).

[21]*Id.*

[22]*See Wilson v. Thorn Energy LLC*, 2010 WL 1712236 (S.D.N.Y. Mar. 15, 2010) (awarding sanctions where defendants failed to back up financial information stored in a flash drive).

[23]*Doe v. Norwalk Community College*, 2007 WL 2066497 (D. Conn. July 16, 2007).

[24]*See* Fed. R. Civ. P. 26(b)(2)(B).

[25]*See* The Health Care Insurance Portability and Accountability Act of 1996, requiring a six-year retention period for documents relate to health care. *See also* Medicaid Regulations, 42 CFR 482.24, 482.26, and 482.53, which deal with the retention of medical records for hospitals that participate in Medicare, Sarbanes-Oxley Act. § 802, Regulation SX, Rule 2-06, regarding documents used for financial audits and so forth.