METROPOLITAN
CORPORATE COUNSEL®

NOVEMBER 2017

WWW.METROCORPCOUNSEL.COM

IN-HOUSE OPS

Now That Your Data's in the Cloud, How Are You Supposed to Get It Out?

Best practices to do it the right way during litigation

By Marc Cressall & Albert Barsocchini / AccessData & NightOwl Discovery

Data migration has reached a tipping point. The vast majority of technology decision-makers (84 percent) say that their organization invested in cloud services in 2016, according to Insight's 2017 Intelligent Technology Index report. It noted that "while only 15 percent have fully migrated their corporate application workloads to public clouds, 47 percent are more than halfway implemented in the cloud, with large and medium companies leading the way."

Corporations have been migrating their data at a remarkable rate, and it's unlikely that this genie is going back into the bottle. In fact, the question you should be asking yourself assumes that it won't.

Now that your data is in the cloud, when the discovery request arrives or the corporate investigation commences, how do you collect that data from the cloud in a manner that is efficient, comprehensive and done with forensic integrity?

Key Challenges

When your data is behind your own firewall or stored on your premises, you can point right at it and collect it with your preferred software tools. The cloud, however, is a very difficult environment for doing e-discovery because you lose the line of sight with your data. You can no longer point at it, easily harvest it and seamlessly drop it into a database. There are a few common challenges you need to overcome in order to perform forensically sound data collection.

Easy to get in, hard to get out: Corporate IT teams have been drawn to the cloud migration trend for a variety of reasons, but unfortunately many of them don't give much thought in advance to how they'll get their data back when the need arises. The truth is that cloud providers are financially incentivized to keep your data housed on their platform. So perhaps unsurprisingly, they typically make it onerous and inefficient to extract. Specifically, the collection tools offered by cloud providers tend to be clunky and forensically unsound, as they lack features to track the chain of custody in data collection. Part of your cloud migration planning should be to ensure that the right data collection tools are in place when you need to get your data out, rather than finding yourself scrambling after you're hit with an urgent request.

International data protection regulations: The EU's General Data Protection Regulation, Japan's Act on the Protection of Personal Information and China's new Cybersecurity Law have all been adopted this year. These new data protection laws govern the way personal data must be handled in various countries, and they have imposed strict constraints on the collection of electronic information pertaining to identifiable individuals. Importantly, these regulations apply to third-party vendors hosting or managing corporate data, so it's critical that your service providers are in compliance with them as well. If you're not careful about how you extract data from sources covered by these regulations, you run the risk of breaking laws that contain serious civil and criminal penalties.

> ***There are many pitfalls, and companies that inadvertently break laws could face serious civil and criminal penalties.***

The rise of Office 365 and Google Apps: The rapid corporate IT move to the cloud has been fueled by the swift market penetration of Microsoft's Office 365 and Google Apps as the primary software-as-a-service (SaaS) email and collaboration suites. A 2016 Gartner study found that 13 percent of publicly listed companies are using either Office 365 or Google's Apps for Work as their cloud email provider. Of this, 8.5 percent use Office 365 while 4.7 percent use Gmail. The remaining 87 percent of companies surveyed have on-premises, hybrid, hosted or private cloud email managed by smaller vendors.

This creates significant challenges for data collection because the built-in collection tools for Office 365 and Google

Apps are inadequate for large-scale collections, forcing digital forensics teams to use a piecemeal approach involving the use of other software tools to complete the collection. Corporate teams and their outside service providers need to have a good collection process for data residing in Office 365, including the use of tools that are designed to connect with Office 365 databases in a forensically sound manner. In addition, on-premise data sources like SharePoint and File Shares are not easily consolidated by Office 365 to enable efficient collections.

Other cloud data sources: The continued migration of systems and servers to the cloud using technologies such as Amazon Web Services, Microsoft Azure and Google Cloud, not to mention other SaaS-based solutions such as Box, Google Drive and Dropbox – create forensic headaches of their own. There are other challenges as well that corporate teams need to overcome in order to carry out effective data collection, but these issues tend to create the most headaches.

Five Best Practices

Once the initial challenges are confronted and managed, there are a few best practices your corporate team may want to consider to pave the way for effective collection in the cloud.

1. Protect data integrity. It's surprisingly easy to inadvertently alter collected data, sometimes with an error as innocent as a mistaken mouse click. For this reason, the courts require the collection of electronic evidence to be done in a forensically sound manner, which essentially means that the data must be left in an unaltered state. This involves preserving the file and system metadata, maintaining thorough audit logs (e.g. what was collected, by whom, when, what actions were taken, etc.) and protecting the collected files to ensure the integrity of the original data.

When collecting data from the cloud, it is important that the tools being utilized do more than simply export the data to a folder on your system. They also need to preserve and duplicate the data in its unaltered state so that it can be put into a tool and further processed for e-discovery. This protects the data integrity so that all parties involved in the dispute can be assured that no one tampered with the data. Also, be very careful about leading software tools on the market that only offer in-place hold, a feature that prevents file deletion but isn't designed with strict protections against alterations that can impact data integrity. Forensically sound collection and preservation require software tools with robust litigation hold capabilities.

2. Be comprehensive. Many corporate legal teams mistakenly believe the only way to perform data collection in the cloud in a defensible manner is to apply digital forensics practices to every piece of data that is possibly relevant. On the other hand, some teams are tempted to apply a short list of search terms to a cloud-based data set and narrow down the number of collected files as much as possible. Our experience is that the best practice is to be comprehensive in your data collection and not too narrow in what you extract. It's unlikely that you will know with certainty all relevant keywords to search at the time of collection, so it will ultimately cost more time and money to perform a second collection later if the initial collection was insufficient.

Many in-house legal teams have learned this lesson the hard way. Some of the leading software tools on the market were built for mass-market uses, not for the rigors and peculiarities of digital forensics and e-discovery. As a result, these tools tend to lack the precise indexing capabilities needed in order to pull in everything the legal team needs during an e-discovery search. The point is that it's wise to use a tool that pulls in all data, indexes it quickly and accurately, and then processes that data with precision.

3. Use reliable software. To meet the rising challenges associated with forensic collection of data in the cloud, corporate investigators and legal IT professionals need access to better and faster software tools that will help them process complex data collected during litigation and digital investigations. With so much data lurking in so many places, getting that relevant evidence from collection to analysis is crucial. For example, AccessData's new AD eDiscovery allows users to quickly collect data in the cloud from Office 365, SharePoint, OneDrive for Business and Office 365 Exchange.

Make sure that your software tool enables your team to collect data from the cloud *and* from on-premise repositories – anywhere the data lives. Even though the cloud migration is in full motion, your team still requires access to a tool that can collect from other end points as well. The optimal solution is to use reliable tools that can collect from any data source.

4. Have experts on your team. It's important to be realistic about the complexities associated with data collection in the cloud. It's not just a matter of clicking a few buttons and checking a few boxes. For your collections to be conducted effectively and accurately, it's important to have the proper expertise on your team. If you work with a cloud services provider, you may be able to connect with their professional support team and obtain the assistance you need.

Of course, you will likely need access to more than just technical expertise and support. You may also want to have access to experts who can instruct and train your team about when to issue litigation holds, how to perform the cloud-based data collection, etc. This might require the services of an outside consulting firm with specific experience assisting corporate legal and IT departments when it comes to their discovery management requirements.

5. Document the chain of custody. Courts want assurance that electronic evidence presented during litigation is the same as what was originally collected (*U.S. v. O'Keefe*, 537 F.Supp.2d 14 [2008]). It's crucial to document the collection process as thoroughly as possible so that your legal team can clearly and credibly demonstrate respect for the digital chain of custody. This includes the proper audit logs discussed above, as well as the use of collection tools that can establish an electronic fingerprint of each collected document.

Corporate counsel who are lax about the digital chain of custody are vulnerable to problems with litigants or investigators, but a detailed process suggests the execution of a forensically sound and fully defensible cloud-based data collection.

The Bottom Line

The steady movement of corporate data and systems from on-premise servers to cloud computing and storage is not an aberration; it is here to stay. This means that data collection in the cloud is going to be a long-term fixture in the worlds of digital forensics and e-discovery.

For those of us in the trenches, the result of the swift migration of corporate systems and data into the cloud is that we're suddenly confronting the daunting challenge of collecting electronic evidence from sources we can't touch or see. It's important to understand the challenges you're going to confront and use the best available tools to navigate those challenges in a forensically sound manner.

Marc Cressall *is senior director of operations for AccessData, a leading provider of integrated digital forensics and e-discovery software for corporate legal departments. He can be reached at mcressall@accessdata.com.*

Albert Barsocchini *is director of strategic consulting for NightOwl Discovery, a discovery services provider that helps companies reach their discovery, investigations and data analysis objectives. He can be reached at abarsocchini@nightowldiscovery.com.*