druva

White Paper

THE CLOUD APPS DATA PROTECTION HANDBOOK

Addressing the critical gaps in SaaS backup, recovery and archival

Q418-CON-10860

Why Backup Your Cloud Data?

Cloud-based applications have become critical to businesses and operations around the globe. But do leading Software as a Service (SaaS) providers such as Box, Microsoft, Google, and Salesforce protect their customers' data with equally critical backup options? And can they recover deleted data when needed, or is it just lost?

Asking why you would want additional protection for data that's already in the cloud is becoming a standard practice. It turns out that cloud application providers may only offer some limited levels of retention and recovery, which are primarily in place just to ensure data accessibility and save themselves and their clients from only certain types of data loss.

The Missing Layers of Cloud Data Protection

These retention capabilities are not intended to make all versions of all data, from every point in time, available to customers whenever they need it. The simple fact is that cloud services are not designed for long-term, policy-based data retention, search, management, and access that companies need. The cloud service providers that do have limited backup capabilities may charge customers a sizeable fee to retrieve even the smallest amounts of their data. Ask yourself, would you consider the Recycle Bin on your computer a "backup solution"?

> ***"IT organizations wrongly assume that high-availability and disaster recovery capabilities offered by SaaS providers can recover data loss by user errors or malicious attacks."***
>
> *— Gartner, "Data Backup/Recovery Factors to Consider When Adopting SaaS," 12/22/2016*

Here are few key reasons why having a third-party data availability and governance offering in conjunction with your SaaS tools is critical and provides a major benefit (capabilities and price point) to any organization.

Data Recovery

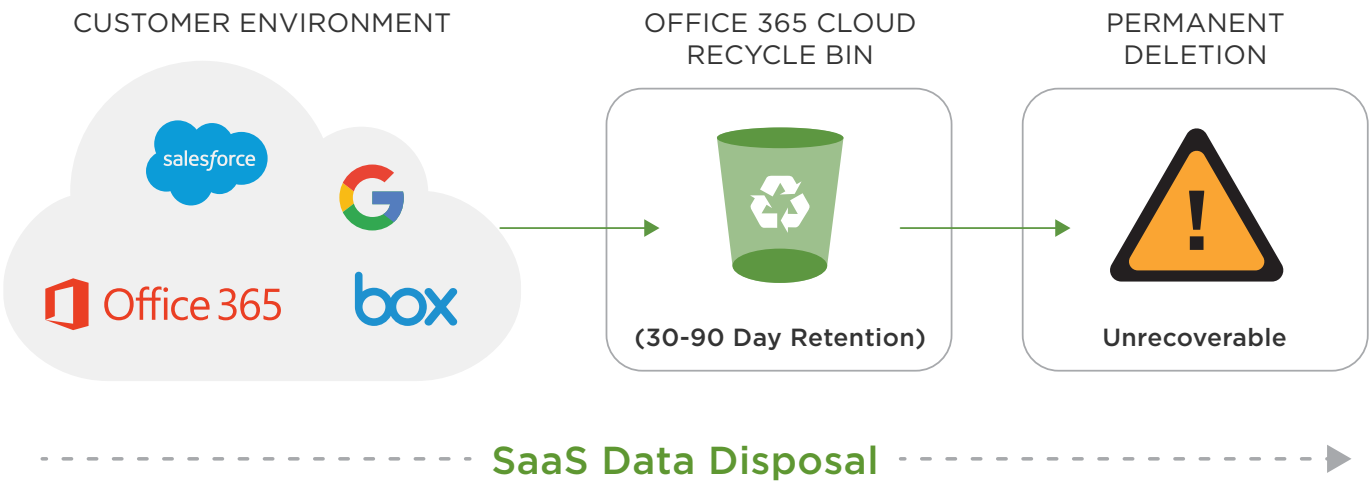
Leading online service providers such as Box, Microsoft, Google and Salesforce offer cloud-based information solutions that are essential to businesses and operations around the globe. But do these major SaaS providers protect their customers' data with backup and recovery? Why would anyone want additional protection for data that's already in the cloud? It turns out that cloud providers do indeed offer different levels of recovery, largely to ensure data accessibility and save themselves and their clients from data loss. But here's the catch: such backups are not intended to make all data available to customers. In fact, cloud solutions are not natively designed for data restoration, and the cloud providers that do have backup capabilities may charge customers a sizeable fee for retrieval. Generally speaking, in most online services, the only backup you have for your organization's data is via the Recycle Bin, which is automatically purged after a fixed period of time. After that, your data is gone forever.

The truth is that once your data is deleted, altered or corrupted — whether accidentally or intentionally — there is very little a system administrator can do to recover it.

File Sharing is Not Data Protection

People often assume that because they're using a cloud-based file-sync-and-share solution like OneDrive for Business, Google Drive or Box, their data is protected as if it was backed up. It's an old argument: "We already have cloud file sharing, can't you just store your files there and call it a day?" The short answer is there are significant and important differences between these types of online services. While file-sharing and data-protection technologies have some overlapping features, they are fundamentally different in their approaches. Here's what you need to know:

1. File sharing is built for real-time collaboration with user content, but it is not designed for data recovery in the case of user error, data corruption or ransomware. Nor does it address archiving or a completely new set of compliance and eDiscovery challenges.
2. Enterprise backup software differs from file sync and sharing in that the software automatically makes a copy of every user's data available for recovery. Endpoint and cloud application data is protected in its entirety, and if a device is lost or stolen, additional features such as remote wipe and geotracking help organizations trace the device and/or remotely delete corporate data. In addition, backing up a user's system and application settings ensures that new or replacement devices can be set up quickly, while preserving the user's familiar working environment.



The Many Causes of Data Loss

While it's extremely unlikely that a major online service provider will lose your data due to a complete service outage, there are a number of other causes of data loss that are very real and occur all too frequently, including the following:

Accidental Deletion and User Error — More often than not, data is deleted by an employee, only for that same employee or their organization to later realize that it is still needed. For example, a collaborator might accidentally delete a shared project or you might delete a scrapped project, and then later learn it is starting up again. Information can also unknowingly be overwritten or corrupted by users and third-party apps.

Malicious Actions — People often delete data before they quit a job if they suspect they are going to be fired or to spite a boss or coworker they are angry at. Hackers can also be the culprits, surpassing security systems to delete, corrupt or lock up data with ransomware. Whether internal or external, untrustworthy people are a reality.

Data Corruption — Applications can hold extremely large data sets that are constantly updated. Overwriting data is a common problem that occurs when large data sets are imported into the application via bulk uploads or when integrated third-party applications are used to manage the data inside the base SaaS application. What if your project management app purges all your calendar events or overloads your inbox with redundant, malformed messages? What if your expense report app paves over your tax-records spreadsheets with garbage data? What if your marketing analytics tool corrupts your CMS database, destroying all your carefully coded web designs?

Ransomware in the Cloud

A few years ago, no one had even heard of ransomware. Today, ransomware is not only commonplace, it's on the rise. What most organizations don't realize is that SaaS applications are equally at risk, with hackers constantly employing new strategies and turning this once-rare form of intrusion into its own mature industry. The ransomware threat is no longer limited to a handful of businesses in a couple of verticals, but now affects all organizations and industries. At the same time, the threat is no longer limited to physical devices, but is now a major concern for users of cloud applications as well. Companies are quickly finding themselves struggling to understand this unsettling new threat and how to adequately plan their response to an attack.

Downtime from ransomware costs small businesses around $8,500 an hour. In the United States, this adds up to a loss of $75B+ per year. And because these criminals continue to operate with zero consequences, it's likely these crimes will not only increase in frequency and severity but will also become a standard part of a company's daily threat landscape. According to the Federal Bureau of Investigation's Internet Crime Complaint Center, there were nearly 2,500 complaints registered in 2015, representing $1.6M+ in damages. But the true numbers are far higher, as fewer than one out of four incidents are actually reported.

Ransomware is on track to become a

$1B

industry

What's at Stake?

Many organizations fail to understand that the cloud is just an extension of a user's operating environment. Data in the cloud is just as susceptible to loss, theft or malicious attack as anywhere else. Enterprises are still responsible for managing data in the cloud, and failure to comply with rules and regulations can result in hefty fines and, worse yet, loss of reputation.

> ***"By 2020, over 50% of all corporate data will reside outside of the corporate data center."***
>
> *— Gartner, "Plan Your Data Exit Strategy Before You Sign a SaaS Contract," published March 2016*

Organizations need to take into account three new challenges and considerations around data availability, compliance and security in order to adequately address the data protection and governance gaps brought about by the rise of cloud apps:

Ensuring Always-On Data Availability — A common misconception among IT leaders and end users alike is that SaaS or cloud data does not need to be protected, because the SaaS vendor is already backing up your sensitive, enterprise information under their Service Level Agreement (SLA). However, many people are not aware that the SLA provided by their SaaS vendor only covers data loss if the provider is at fault (e.g., a service outage). The SLA typically does not cover data lost due to accidental deletion, migration errors, data corruption or malicious attacks. SaaS vendors may not be able to help you recover deleted data older than 30 days, because their service, as a part of their standard, permanently purges the deleted information after that period. Even if the SaaS provider is willing to work with you, and the data still exists, they may charge you a sizeable fee and recommend you use a cloud backup solution. And even if the data is actually recovered, countless hours of productivity will most likely have been lost while trying to get it back.

Meeting Legal Hold Obligations — Today, businesses can face very serious consequences if they fail to produce data stored on SaaS platforms during litigation following a discovery request made by the courts. This requires legal teams within an organization to have immediate access to user data that may be critical for the defense of their case or to avoid serious penalties. In many cases, some or all of this data resides in cloud services like Office 365 or Box, which may not be recoverable or may remain completely unprotected throughout the litigation process and susceptible to deletion or mishandling by the users.

The core of legal discovery is the process of mining through the data to identify and isolate information that is relevant to the litigation. To do so assumes that information is properly indexed and that the search functionality is sufficiently flexible. In addition, during early case assessment, the ability to see results in real time and refine the search based upon the results becomes essential.

Not having timely and easy access to current and historical data for collection and review purposes could cost an organization millions of dollars in legal fees or even the outcome of a lawsuit. Collecting data residing on cloud applications while preserving and handling it in a way that can be defensibly presented in court (no data spoliation) is crucial for every organization and their legal team to address with an effective solution.

Addressing Security and Compliance in the Cloud — A top concern for any Information Security (InfoSec) team is the risk associated with the leakage of sensitive and confidential data. A recent study performed by Dimensional Research indicates that close to 95% of businesses have some form of sensitive data in the cloud. The cost of not protecting this data can be staggering, not just in the form of regulatory fines, but also measured by the effects it would have on a business's reputation and the significant loss of trust as a result.

With privacy laws changing constantly, the regulatory environment is becoming even more complex. The General Data Protection Regulation (GDPR) and Privacy Shield, adopted by the European Union (EU), demonstrates data visibility mandates that go beyond what most organizations have in place today. Sarbanes-Oxley, the Health Insurance Portability and Accountability Act (HIPAA) and new data privacy regulations have likewise forced businesses to drastically change how they capture, store and secure data.

Business Case for Third-Party Apps

These SaaS applications offer a range of valuable capabilities that organizations rely on everyday to be more productive in achieving business goals. However, these powerful tools are not the purpose-built products that are needed to address the key concerns highlighted above. An increasing number of organizations have taken action to address the gaps in end-user data protection, data recovery, legal hold and eDiscovery, as well as third-party management of Office 365 archival data.

Office 365 | Microsoft Office 365 Data Recovery

The core capabilities of Office 365, while powerful, are not necessarily built to be a comprehensive solution for companies' data availability and governance requirements.

- **Office 365 Exchange —** Deleted items are moved to the Deleted Items folder, where they will remain until either manually or automatically deleted based on a retention policy that, by default, is 30 days. Once deleted from the Deleted Items Folder, items will remain in the Recovered Items folder for a minimum of 14 days.
- **Exchange Online Archiving —** Microsoft offers Exchange Online Archiving as part of its E3 and E5 plans, or as a separate add-on at $3 per user for all other plans. This is an email-only archive option that must be set for each individual mailbox, and does not include archiving of Calendar, Contacts or Tasks. Although individual emails can be recovered, this does not provide the ability to restore a mailbox from a specific point in time.
- **SharePoint Online and OneDrive for Business —** Deleted items first go to the site's Recycle Bin, and then they are automatically removed after 90 days. Once the items are automatically or manually purged from the Recycle Bin, they will go to the Site Collection Recycle Bin and remain for a set number of days specified by the system administrator before being completely purged from SharePoint.

According to Gartner, "Because the infrastructure is not unified, and the backup capabilities of the components differ, users need to have a clear understanding of strengths and gaps. For those users who want to have more control and configuration flexibility, they may have to look to a third-party backup tool for Office 365."[1]

G Suite | Google Apps Data Recovery

Google data retention and recovery varies by service, so here's a summary of data retention policies from a variety of Google Help documents:

- **Gmail —** An email is gone forever 30 days after moving it to the trash, or immediately after clicking on "Delete Forever."
- **Google Contacts —** A contact is gone forever 30 days after deleting it.
- **Google Calendar —** Once an event is deleted, its full details can't be recovered.
- **Google Drive —** Once a document has been deleted from the Trash file, a Google Apps administrator can recover the data for up to 25 days. However, after 25 days, it's gone forever.
- **Google Sites —** A deleted site can be recovered from the Deleted Sites folder for 30 days; after that, the site is gone forever.
- **Google Account —** An account can only be recovered "within a short period of time after deletion." If a Google Apps administrator deletes an end user's account, all documents and files owned by that person will no longer be accessible by collaborators and viewers.

Google users can manually download a copy of selected files to their local PC, but according to a recent Gartner study, "The process is not scalable and can't be managed centrally by an organization."[2] Gartner goes on to state that "organizations desiring more robust backup/recovery functions, such as longer retention and backup of all key components of Google Apps for Work, should consider a third-party backup tool to achieve those goals."

[1] Gartner, "You May Need Additional Backup to Prevent Data Loss From Your SaaS Solutions," June 2016

[2] ibid

| Salesforce Data Recovery

Salesforce data recovery is primarily based on the Recycle Bin as follows:

- Deleted Salesforce records can be recovered from the Recycle Bin for 15 days before they are permanently deleted.
- Once the Recycle Bin storage limits have been reached, Salesforce automatically removes the oldest records if they have been in the Recycle Bin for at least two hours.
- Deleting a custom object is unrecoverable, as the data is immediately deleted from the database.
- Salesforce offers an Admin Export function for their Enterprise and Unlimited Editions, but the export can only be run once a week and requires a system administrator to manually download and archive the data to local storage each week.

Salesforce offers a data recovery service if the data has been deleted within the last three months. This service costs a minimum of $10,000 and generally takes 15 business days to recover the data. Metadata, however, is not included, so it's up to you to restore the data back to Salesforce using the CSV file they provide.

Gartner concludes that "for customers who want to have additional backup/recovery functionalities, they may find it worthwhile to evaluate and adopt a third-party backup tool that offers more automation and simpler procedures than the native backup/restore functions."[3]

| Box Data Recovery

Box is a powerful enterprise file sync, share and collaboration platform, but it lacks the ability to recover files that have been accidentally or maliciously destroyed. Box purges items from the Recycle Bin after 30 days by default, or after a time period specified by a system administrator. Once the Trash folder is emptied, data is gone forever. Box also fails to provide visibility into information stored on endpoints unless it's shared, so it cannot satisfy compliance and eDiscovery challenges for information outside of Box.

Closing the Gap in Cloud Data Protection

Up until now, IT has used labor-intensive, costly and complicated processes to access and manage data. SaaS applications have changed much of that old paradigm by revolutionizing the way organizations manage and consume the software that generates the majority of their critical data. However, organizations must change the way cloud data is protected and governed if they want to meet today's real-world business needs.

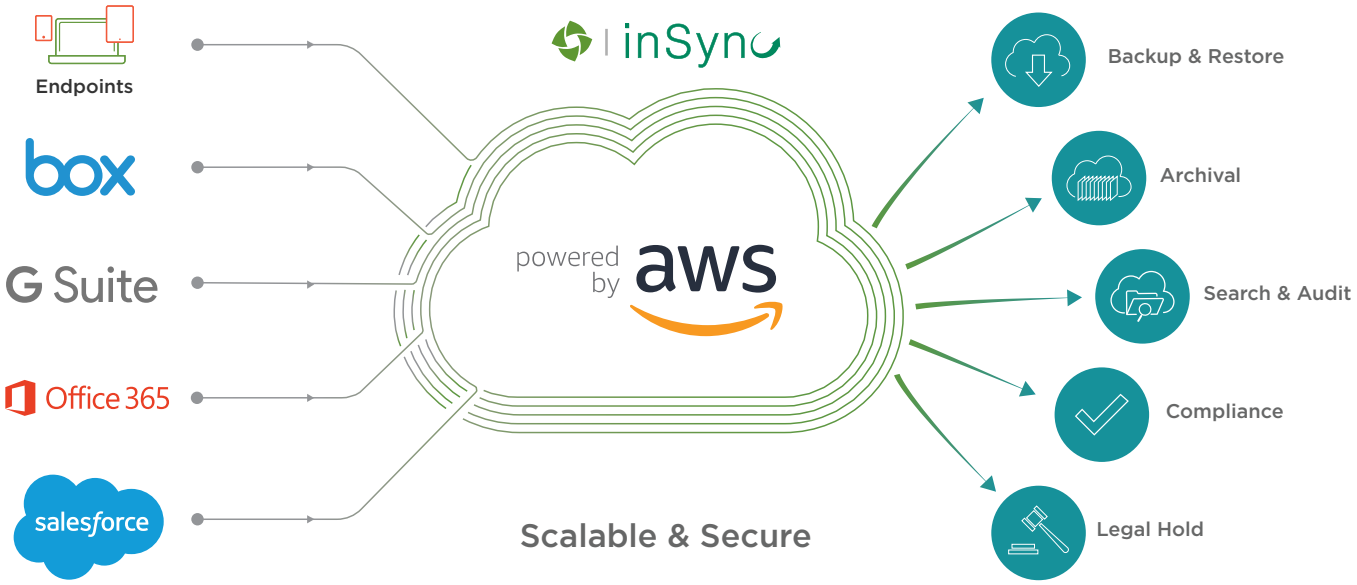
While solutions are available to solve individual challenges, it's essential that you adopt a comprehensive, integrated platform that manages data regardless of device type, service provider or physical location. This integrated platform should also provide a single, centralized view of data that's created and stored across cloud applications, endpoints and server data. This allows organizations to better conduct analysis, assess risks, improve compliance and meet other needs. Thus, the notion of a unified level of data protection across all data sources is emerging today.

How Druva Fits In

A modern and comprehensive data-protection solution should address the above challenges for all user data, irrespective of where it is located — on a laptop, a mobile device, or a cloud service like Office 365, Box or Google Apps. In order to be truly effective, a solution needs to "follow the full user behavior" to collect enterprise data from different data sources in a nimble and unified way. This enables IT and security organizations to meet data availability and information governance demands.

[3] Gartner, "You May Need Additional Backup to Prevent Data Loss From Your SaaS Solutions," June 2016

Druva helps some of the world's largest organizations protect their investment in Microsoft Office 365, Google G Suite and other SaaS environments from data loss and compliance violations. Druva's industry-leading solutions give users a single pane of glass to monitor and protect data no matter where it resides.



Druva is the essential layer of data-protection functionality, where administrators can continuously track and monitor data within cloud applications in addition to mobile and laptop endpoints. It also allows system administrators to be quickly alerted of potential data risks associated with sensitive information such as Personal Healthcare Information (PHI), Personal Credit Information (PCI), Personally Identifiable Information (PII) and Intellectual Property (IP) so they can then take the appropriate actions.

The Big Takeaways

Consider these two critical issues after reading what is offered by these major SaaS providers:

- **Hidden Gaps —** By ignoring the data retention gaps within these products, you are relinquishing control of your organization's business-critical information and putting it entirely

  In the hands of the end users. This puts the burden of regulatory compliance solely on the shoulders of those who may have no understanding of what is needed to manage company data correctly.

- **Legal Pitfalls —** Most litigation takes weeks, if not months, to reach a stage at which custodians are identified and data is put on legal hold. By the time this happens, all relevant data could be lost under every single one of the scenarios outlined above.

Learn how to address the critical gaps in your SaaS services at our Cloud Applications Solution Page.

About Druva

Druva is the global leader in Cloud Data Protection and Management, delivering the industry's first data management-as-a-service solution that aggregates data from endpoints, servers and cloud applications and leverages the public cloud to offer a single pane of glass to enable data protection, governance and intelligence—dramatically increasing the availability and visibility of business critical information, while reducing the risk, cost and complexity of managing and protecting it.

Druva's award-winning solutions intelligently collect data, and unify backup, disaster recovery, archival and governance capabilities onto a single, optimized data set. As the industry's fastest growing data protection provider, Druva is trusted by over 4,000 global organizations, and protects over 40 petabytes of data. Learn more at www.druva.com and join the conversation at twitter.com/druvainc.

Druva, Inc.
Americas: +1 888-248-4976
Europe: +44 (0) 203-7509440
India: +91 (0) 20 6726-3300
Japan: +81-3-6890-8667
Singapore: +65 3158-4985
Australia: +61 1300-312-729
sales@druva.com
www.druva.com