Backup and Archiving for Office 365

White Paper

Summary

There is often confusion between the two complementary processes of data backup and information archiving. In this white paper, we compare them and discuss the specific needs addressed by each one. We review why organizations continue to need both a backup and an archiving solution after they have moved to Office 365. We will also explain why it's not advisable to use a single solution for both processes.

Comparing Backup and Archiving

Backups and archives both store copies of data from the production environment, but the use cases for each are different, which means that different solutions are appropriate for each.

A backup enables recovery from a situation where data has been lost, corrupted or becomes inaccessible, so it's primarily a safeguard to facilitate data recovery. A backup stores multiple copies that are each associated with a specific revision of data, and it provides recovery back to a known good state from a specific point of time.

An archive enables compliance with legal and business data retention policies, as well as supporting eDiscovery. An archive preserves a single copy in secure immutable storage for a finite time period and provides ongoing end user access to historical business information.

This table summarizes some of the key differences between backup and archiving:

Function	Backup	Archiving
Primary Purpose	Protect current and revision data.	Preserve historical data.
What it Enables	Point in time data recovery.	Data retention and discovery.
Business Need	Restore data after loss, corruption or unintentional deletion, to a specific point in time .	Produce complete and accurate evidence to meet legal, regulatory, and policy obligations.
Who Uses It	Administrators (IT)	Business Users (Legal, HR, etc.)
What it Stores	Multiple point in time copies of data (revisions).	A single fully indexed copy of all data.
Optimized For	Point in time restore of production environment.	Item level preservation, search, retrieval, analysis, and export.
Data Disposition	Source data is left in place.	Source data may be deleted.
Search	Point in time revision and basic metadata search.	Full-text search of files, attachments and metadata (custodians, keywords, etc.) plus eDiscovery capability.
Retention Polices	Primarily based on the age of data that would need to be restored.	Primarily based on age, location, content and metadata. Includes overriding Legal Hold capability.

For effective data protection and preservation, organizations need both a backup and an archiving strategy. They may attempt to use a backup solution as an archive (and vice versa), but as will be discussed below, there are significant limitations and deficiencies with this approach that make it inadvisable.

Backup is for Recovery

The primary purpose of backup is to allow recovery from the situation where the original version of data is lost due to unintentional or accidental deletion, or where files have been corrupted in some way to make it unusable.

A backup system achieves this by taking copies of the data on a regular basis to create a series of revisions. Each one of these revisions reflects the data at a specified point in time and can be restored back as needed.

Most backup copies are retained only for a few days or weeks with later copies superseding previous versions. However, it is common for one version to be retained semi-permanently on a weekly or monthly basis to allow data to be recovered from a much earlier point in time. For email data in particular, backup solutions are typically used to protect the most recent data, as this tends to be the most relevant for end users.

Why you need to backup Office 365

Accidental deletion by users is the most frequent cause of data loss in a SaaS environment such as Microsoft Office 365; however, there are other ways where data can be lost. Application errors or mistakes in processing can cause data to be lost or overwritten, and there is always the risk of malicious deletion by employees with access to the data. A threat that has come to prominence recently is ransomware, which encrypts data and renders it inaccessible.

Office 365 itself focuses primarily on ensuring that service and data availability is not disrupted, but Microsoft does provide customers two options for data recovery:

1. **Recycle Bin:** Data that has been recently deleted by users can be recovered from the Recycle Bin (for OneDrive), or the Deleted Items and Recoverable Items folders (for Exchange Online). However, these are subject to retention periods as listed below, after which data is permanently deleted and no longer available.

MICROSOFT SOLUTION	RECOVERY FEATURE	RETENTION PERIOD
SharePoint Online	Site Recycle Bin	93 days
	Second Stage Recycle Bin	93 days
Exchange Online	Deleted Items	Configurable
	Recoverable Items	Up to 30 days
OneDrive for Business	Recycle Bin	93 days
	Second Stage Recycle Bin	93 days

2. **Document Versioning:** If the Document Versioning feature is turned on, OneDrive for Business will retain a number of previous versions of each document that has been amended, and end users are then able to restore back to any of these previous versions. However, this does not provide protection against unintentional or accidental deletion, as all versions of a document are removed when the current version is deleted.

Organizations adopting Office 365 should recognize that there are serious limitations inherent in both of these options:

- These options operate at an individual item level, so they are not suitable for the recovery of larger quantities of data such as entire folders or mailboxes.
- It is not feasible to do a point in time data restore using either of these options without considerable additional processing.

Retaining data when an employee leaves an organization is another problem area with Office 365. All data stored by a user in Exchange Online will be permanently deleted 30 days after their account is deleted, so this data must be backed up first. Email data may be retained for a longer period if a retention policy is applied before their account is deleted, but this is a premium feature only available in Office 365 E3 plans and above.

As a result, an increasing number of organizations are now starting to use third-party backup solutions with Office 365. These solutions provide an additional layer of protection, as well as much longer retention periods and more comprehensive recovery options.

Problems with using a backup as an archive:
There are a number of issues using backup as an archiving solution:

- A backup will capture all data that exists at a specific point in time, but it will not capture volatile data such as new email messages that may be amended or deleted before the next backup is taken. This means that a backup cannot support fully compliant data retention and eDiscovery for an organization.
- Legal requirements and business regulations mean that organizations typically need to enforce a set of retention policies covering different types of information, and these are often based on the data content itself or the metadata associated with each data item. Indexing allows archiving to support a wide range of policy-based retention rules to meet the most demanding of requirements, whereas the retention policies within backup solutions do not usually provide this granular level of control.
- Archive solutions are designed to support search and retrieval, so all data content and metadata will be indexed, and advanced techniques such as full-text search and data tagging are provided to enable business users to undertake discovery exercises easily without IT involvement. In contrast, backup solutions are designed for use by administrators working at the file level or higher, and as data is typically not fully indexed, attempting to use a backup for eDiscovery can be an extremely complex and time-consuming task for the IT department.
- Backup solutions retain multiple point-in-time revisions to enable organizations to meet recovery point objectives. But searching across multiple revisions will return multiple versions of individual data items which then need to be reconciled in order to provide a single set of accurate and verifiable results for eDiscovery. In contrast, an archive provides a single verifiable copy of each data item.
- As backup solutions are not accessible by end users, meeting the needs of end users for access to their historical data can be a significant ongoing overhead for the IT department. In contrast, archive solutions provide the ability for end users to search and retrieve their own archived data, and allow them to restore individual items as needed, all without the involvement of their IT department.

Archiving is for Discovery

The primary purpose of archiving is the long-term preservation and retention of current and historical data, to enable the organization to undertake legal and other eDiscovery requests on this data as well as meeting their compliance and business requirements for data retention and deletion.

An archiving system achieves this by capturing and securing a copy of every item of data as it is created. In the case of email, this must be captured as soon as the message is sent or received, and

before an end user has time to amend or delete the message.

The archive builds up over time to reflect every item of data that has ever existed during that time, even if it has since been deleted from the original location. Retention policies ensure that the archive copy of each item of data is retained for as long as required, and deleted after that time.

End users are able to search and retrieve their own archived data whenever needed, while auditors and other administrative users are provided with a range of advanced search and export features to enable them to undertake complex organization-wide eDiscovery requests.

Why you need to archive from Office 365

Microsoft has improved the compliance features within Office 365, and also provides an archive mailbox within Exchange Online, but there are still a number of limitations. Together with the use of "In-Place Archiving" instead of a separate dedicated archive, this means Office 365 is unlikely to meet the wider needs of those organizations with specific data retention, policy enforcement, and e-discovery requirements. It's also important to note that archiving and compliance features require the more expensive Office 365 E3 and E5 plans.

Data security: Office 365 retains all data (including archived data) in the operational environment where it co-exists with more transient data and is at risk of amendment or deletion. This contrasts with third-party solutions that take the accepted "best practice" approach to retaining a separate immutable copy of every email outside the operational email environment in a separate secure repository.

Data retention: Retention policies for email in Office 365 are limited to just age or location, and do not provide the flexibility or granularity many organizations require to meet their compliance requirements, such as rules pertaining to custodians or content.

Data preservation: Retention policies within Office 365 use a complex process with multiple folders to secure email data against modification or deletion. The Discovery Hold and Versions subfolders within the Recoverable Items folder are used to store original copies of items that have been deleted or modified, whereas unmodified items remain in the user's Inbox or Archive Mailbox. This means that the original copies of emails can be spread across multiple folders and there can be multiple versions of the same email within a mailbox, so it is not easy to ensure and demonstrate you are retaining a complete and accurate copy of every email sent or received.

These and other limitations mean that Office 365 is unlikely to meet the wider needs of organizations that have specific data retention, policy enforcement and eDiscovery requirements, and many are now implementing third party archiving solutions to enhance Office 365.

Problems with using an archive for recovery

Some organizations may be tempted to use data stored in their archive as a backup solution, but this approach has a number of limitations that make it unsuitable and problematic:

- An archive captures a copy of all data items created over a period of time in a single version, compared to a backup where each revision captures each data item within its current context at that specific point in time. It may not be feasible to do a point in time data restore from an archive without considerable additional processing.
- The directory structure for each user's data in the archive will be based on historical information and possibly not reflect their current live usage, making it difficult to restore data back to the correct location.

- An archive will retain (subject to organizational retention policies) a copy of all data that a user has ever owned over the period of time. This will include items that have been intentionally deleted and that should not be restored unless specifically requested.
- An archive is optimized for search and retrieval, and may be appropriate for end users to recover individual items. However, it is likely to be inefficient when used for the recovery of larger quantities of data such as entire folders or mailboxes.

Backup and Archives Should Work Together

As we have seen, backup and archiving perform separate functions for different reasons, but they are complementary, and the capabilities of each one can help the other work more effectively.

Storing data within an archive is the best way to retain a secure copy of all historical information, and will enable an organization to meet its business requirements for compliance and discovery.

Archiving older or inactive data and removing it from operational storage will also improve the operation of backup processes. By reducing the volume of data to be backed up, backups will run more quickly, and with less data to manage, it will be easier to restore data when needed.

Barracuda Solutions for Backup and Archiving from Office 365

Barracuda Cloud-to-Cloud Backup

This protects Exchange Online and OneDrive for Business data by backing it up directly to Barracuda Cloud Storage. Barracuda Cloud-to-Cloud Backup can be used as an add-on to an on-premises Barracuda Backup appliance, or as a standalone subscription without an appliance.

For Exchange Online, Barracuda Cloud-to-Cloud Backup protects all email messages, including all attachments, as well as the complete folder structure of each users' mailbox. In OneDrive for Business, all files under the Documents Library, including the entire folder structure, are protected.

Barracuda Cloud Archiving Service

Cloud-based archiving enables Office 365 customers to meet demanding compliance requirements and address eDiscovery requests easily and effectively. The Barracuda Cloud Archiving Service ensures email is stored securely in a separate repository for as long as needed without risk of amendment or deletion.

Essentials for Office 365

Barracuda Essentials for Office 365 combines Cloud-to-Cloud Backup and the Cloud Archiving Service with the Barracuda Email Security Service to provide an integrated multi-layer security, archiving, and backup solution for Office 365. With complete protection of email, data, and cloud infrastructures, Barracuda Essentials gives Office 365 customers peace of mind.

Conclusion

You still own your data even though it is in Office 365. Although Microsoft will do their best to manage your data effectively, you ultimately remain responsible for the protection, backup, and compliance of that data—just as you did before you moved to Office 365. Therefore, you need to ensure you have effective backup and archiving solutions in place.

We have seen that backup and archiving solutions each have their own characteristics and features that enable them to meet two distinct sets of needs. Although organizations may be tempted to

use their archive as a backup (or vice versa), this is not an effective approach due to the number of issues that it raises, so they are advised to implement both an archive and backup solution.

Management and support overheads can be minimized by choosing an established vendor like Barracuda to supply an integrated solution such as Essentials for Office 365, which includes both backup and archiving, as well as email security.

About Barracuda Networks, Inc.

Barracuda (NYSE: CUDA) simplifies IT with cloud-enabled solutions that empower customers to protect their networks, applications, and data regardless of where they reside. These powerful, easy-to-use, and affordable solutions are trusted by more than 150,000 organizations worldwide, and are delivered in appliance, virtual appliance, cloud, and hybrid configurations. Barracuda's customer-centric business model focuses on delivering high-value, subscription-based IT solutions that provide end-to-end network security and data protection. For additional information, please visit barracuda.com.

Queens Park CyberSecurity Solutions Limited

8a Bravington Road
Maida Vale
London
London
W9 3AH
United Kingdom

02039710740
info@qpcsl.com
www.qpcsl.com