proofpoint

Addressing Legal Discovery & Compliance Requirements

A Comparison of Proofpoint Enterprise Archive and Microsoft Office 365 Archiving

In today's digital landscape, the legal, regulatory and business requirements for email archiving continue to grow in scale and complexity. Relying on basic archiving capabilities can be costly and risky. While Office 365 does include basic functionality, these simple capabilities may not be enough for many organizations, especially for those trying to respond to eDiscovery requests. According to Osterman Research, "Office 365 email archiving capability does provide many basic archiving capabilities, but does not meet the additional eDiscovery and regulatory requirements with which many organizations are faced." Alternatively, Proofpoint Enterprise Archive was designed to meet the most stringent and complex email archiving requirements. Below is a comparison of key features.

Retention Management

The uncontrolled growth of information represents both risk and cost for legal discovery. Having a retention policy that is consistently enforced is becoming more critical than ever. A policy that is not consistently applied is very difficult to defend. It is hard to prove that critical data was not maliciously destroyed if there is no systematic process for elimination of data at the end of its useful lifecycle.

Microsoft Office 365	Proofpoint Enterprise Archive
Retention Tags: Tags that are applied by users identify when items should be removed, and the system will automatically delete item on that date. Tags can be changed at any time and they do not prevent a user from deleting items before their stated retention age.	**Retention Policies:** Retention is rule-based and automatic. Customers can configure as many retention policies as they wish to identify specific types of messages or groups of users that require unique retention periods. Messages not matching any of these specific retention periods are assigned the retention period that the customer configures as the "default".

Search and Investigations

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

Microsoft Office 365	Proofpoint Enterprise Archive
Mailbox-based Index Structure: Exchange maintains separate indexes for each mailbox. This is the optimum model to allow end-users to search within their mailbox, but it creates performance challenges for searching across mailboxes. In addition, because Exchange's indexing is a background process, and is a "best efforts" model, not everything that exists in mailboxes may be fully indexed at any given time. Depending upon your retention strategy, if a user deletes their copy of a message, yet it exists in other mailboxes, focusing on searching only their mailbox might miss relevant items that you still possess.	**Entire Repository Index Structure:** Proofpoint maintains a unified index structure for the entire archive. A single-instance copy of each message exists in the archive, with metadata about which mailboxes each message belongs to. This allows for searching within specific mailboxes, or across the entire repository – all in near real-time. Proofpoint indexes the content before it is added to the archive. This ensures that every item within the archive is fully searchable.
Limited File Types Indexed: Office 365 indexes Office documents, PDF files and textual documents. Other document types are not indexed, so keyword searches will miss relevant content.	**Over 500 File Types Indexed:** Proofpoint indexes over 500 file types, including items nested within zip files. This ensures that you won't miss key relevant documents just because they were in a less common file format.
Batch Search Experience: Due to the mailbox-based indexing model, Office 365 does not execute searches in real-time. Instead, you create a search job and get notified when the search is complete. You can either copy the results to another "discovery mailbox" (in which case your search results are limited to the 50 GB maximum mailbox size) or you can create an Export job directly to PST files. While this model may work if you are extracting a whole mailbox, it does not allow for investigation, or search refinement. The end result is that downstream discovery costs may be higher, as more data needs to be processed.	**Real-time Search Experience:** Proofpoint is the only cloud vendor that offers an SLA on search performance – regardless of how big the archive becomes or how many mailboxes you search. This real-time search experience allows you to easily scan through search results to identify opportunities to refine your search. Not only is this critical for investigational activity, where you are trying to understand what was going on, it allows you to narrow the scope of data exported to third party tools/vendors to reduce the next steps in legal discovery.
Search times grow with number of mailboxes: Office 365 executes searches on a mailbox by mailbox basis. As a result, the more mailboxes that you search within, the longer it will take for the batch search to be completed. You can't preview results or initiate an export task until the search is complete, so there can be significant wasted time during key discovery tasks if you don't initiate follow on steps immediately after the search completes.	**Can Search Entire Archive:** By default, discovery searches are performed across the entire archive. This allows you to use the archive as part of the identification phase to determine who potential custodians for a matter might be.
Cannot Search within Legal Hold: Legal Holds in Office 365 are designed solely to protect the data from being disposed. They do not represent a logical container of matter specific data. As such, while searches include data that is subject to legal hold, you cannot focus your search activity within the data that has already been identified for a matter.	**Can Search Within Legal Hold:** Proofpoint's legal holds not only protect data beyond their standard retention period, they also represent a logical repository of matter-relevant data. While you can search across the entire archive, including holds, collection activities are often easiest if you start with the data that you have already identified as needing to be preserved for the matter.

Data Integrity

To be useful for legal discovery, the data must be of evidentiary quality. It must maintain all of the metadata of the original message and the processes around data management must ensure that data is never lost or corrupted.

Microsoft Office 365	Proofpoint Enterprise Archive
Best-effort Data Loss Management: Office 365 is designed, first and foremost, to meet the business needs of communication and collaboration. To that end, data is replicated in near real-time between datacenters to ensure very high availability. No snapshots or backups are performed. The drawback of this approach is that any corruption is also replicated, with no roll-back possible. For legal discovery, this means that messages can be lost – including those that are on legal hold.	**Zero Data Loss Design:** Proofpoint designed the archiving service to ensure that messaging data is never lost. Redundant copies of the data exist at every stage of the lifecycle – with the vast majority of the data maintained as part of static long-term storage. Because these data stores are not in a constant state of flux, the system can validate that the redundant copies are a true representation of the original data. For recently archived data, XML transaction logs and backups allow us to reprocess items, should a server failure occur.

Legal Hold

As soon as you become aware of the potential for litigation, you have an obligation to preserve data that might be relevant to the matter. As each person may be subject to many legal holds, hold management needs to be organized by matter, allowing for the removal of a legal hold for the matter without risking the releasing overlapping content subject to another matter.

Microsoft Office 365	Proofpoint Enterprise Archive
Matter-based legal holds: Legal holds can be defined for a set of mailboxes, and optionally be constrained based upon date or content.	**Matter-based legal holds:** Legal holds can be defined for a set of mailboxes, and optionally be constrained based upon date or content.
Limits on Legal Holds: While many holds can be applied to a mailbox, if more than 5 holds apply to a mailbox, the entire mailbox is preserved – even all of the holds are constrained by dates or keywords. When a user deletes a message, it is moved to a hidden folder. A nightly process cleans up this folder. A legal hold in Office 365 is basically just a rule that prevents this clean-up process from removing items from the hidden folder. A given legal hold can only apply to 10,000 mailboxes. Should more mailboxes be required, multiple holds must be created. Microsoft's contract terms limit the percentage of users that can be subject to legal hold, so it may not be possible to preserve all of the data that you are legally obligated to.	**No limits on legal holds:** An unlimited number of holds can be applied to a given mailbox, and the rules will be properly applied. There is no limit to the number of mailboxes that a given legal hold can apply to.

Export

The ultimate output of any archiving solution is data that is fed into another legal review system or passed directly to opposing counsel. Export performance and workflows are key attributes of an effective solution.

Microsoft Office 365	Proofpoint Enterprise Archive
Exports run on discovery user's desktop: A client-side app is used to export search results. As a result, the discovery user's machine resources are occupied while the export is performed. For large exports, this could be several hours of processing time. For the export to complete, the machine must remain connected to the network, so the user can't take their laptop home. Files are created on the user's local hard drive, so sufficient disk space to hold the entire search result set is required. Performance of export is based upon the speed of the discovery user's machine.	**Exports run in the background:** Export jobs run on infrastructure in Proofpoint's datacenter – no discovery users' machines. As such, they do not depend on user's computer resources being available.
Only one export can be run at a time: Only one instance of the export client app can be run on a machine at a time. As such, any given discovery user can only perform one export job at a time. In addition, you can't queue multiple search jobs to be exported for processing overnight or on the weekend, resulting in lost processing time. Customers must monitor export jobs and create support tickets if there are issues – Microsoft does not proactively monitor exports as part of the service.	**Multiple Jobs run Concurrently:** Multiple export jobs can be queued for processing and several can run at a time (depending upon the number of export processing machines configured for the customer). As queued jobs are picked up automatically, there is no lost processing time between jobs. Proofpoint support activity monitors progress of export jobs to react quickly if there are issues.
Export Results cannot be pushed to Third Parties: Results must be downloaded from eDiscovery Center to the user's local machine, then uploaded to legal service providers.	**Export Jobs can Automatically Upload Results via Secure FTP:** Export jobs can be configured to upload the results to a Secure FTP location automatically upon completion. This removes manual steps, avoids chain of custody issues and avoids delays.

A checklist of issues for organizations to consider before fully relying on Office 365 for Email Archiving:

Has your organization ever been involved in eDiscovery and do you want more in-house control?
With only the basic eDiscovery features available in 0365, the majority of the eDiscovery process will need to be handed off to outside counsel at a much higher cost.

Has your organization ever been involved in litigation where your email system was part of the collection/discovery process?
The email system is the most targeted data repository in eDiscovery. The ability to search across both the email system and email archive with advanced search capabilities and without limits (i.e., limits to search results size) can mean the difference between an incomplete or late discovery response and winning the case.

Does your organization want/need an absolute copy of record for all archived content?
A copy of record can exist only in an unalterable repository. The Office 365 email archive is not an unalterable repository in its default state.

Do you envision a need to export archived email in a file format other than .PST?
The default file format for email data in discovery is the .PST format. However, there are circumstances in which a different file format is needed. To respond to a regulatory information request, separate PDFs or .CAB files may be the format requested. The Office 365 In-Place eDiscovery search results are exportable only in the .PST format, causing at least one more, potentially costly, step to be included in the process of responding to an information request.

Does your organization employ licensed brokers or traders subject to SEC and FINRA regulations?
Office 365 does not have a Journaled archive mailbox, so to meet SEC and FINRA requirements, a third-party email archive would be required. Office 365 also lacks mailbox-monitoring capability to meet the FINRA surveillance requirements that many third-party email archives do provide.

Does your organization need/want to manage archived email and attachments beyond archive and delete granularity?
Many organizations have moved beyond the keep it forever or delete it records retention policy. Organizations now want their archive automation to be able to make retention/disposition decisions at a more granular capability. Office 365 email archiving provides only for delete it or archive it records retention policies and instead relies on individual employees to manually drag and assign more granular retention policies.

Do you see a need to produce audit logs of archived mailbox activity?
Because Office 365 does not, by default, protect archived email data from deletion or alteration, the ability to produce audit activity reports on specific employee mailboxes would be a must have to be able to represent archived email as unaltered in legal and regulatory events. Office 365 provides for mailbox audit logging that allows customers to track access to the mailbox by people other than the mailbox owner, but does not audit mailbox owner activities.

about proofpoint
Proofpoint Inc. (NASDAQ:PFPT) is a leading security-as-a-service provider that focuses on cloud-based solutions for threat protection, compliance, archiving & governance, and secure communications. Organizations around the world depend on Proofpoint's expertise, patented technologies and on-demand delivery system to protect against phishing, malware and spam, safeguard privacy, encrypt sensitive information, and archive and govern messages and critical enterprise information.

proofpoint™
892 Ross Drive
Sunnyvale, CA 94089
1.408.517.4710
www.proofpoint.com