



Buyers' Guide Insight Brief



in association with



HubStor™

Taming your information sprawl with cloud data management

An Evaluation Guide for Business

What is information sprawl and how can it be tamed?

Modern web applications and subscription-based software are attractive, but they all need to store data somewhere...



Our business information is not just in storage systems on-site nowadays. It's in the likes of Box and Dropbox, and in our other web-based applications, such as Salesforce, Office365, Slack or ServiceNow. It's even in the laptops, smartphones, tablets and other mobile devices that we and our colleagues travel with.

Like it or not, we all have information sprawl, and that reinforces risks and challenges that many IT teams still have problems with for on-site systems, such as backup and recovery, and data governance.

For example, just because we store data 'in the cloud', that does not mean it is backed up. Similarly, data maliciously deleted from a web application may well not be recoverable unless we have made specific provision for that – few services provide such a capability unasked. And yes, we can put a second copy of a 'live' database in cloud storage, but unless it is specifically implemented as a backup, it is very unlikely that this will allow the relevant application to be restored to a coherent, working state.

Then there is the need for consistent visibility into all our data. This is both for governance and management, and to enable us to analyze and cross-index the data to get information from it - and of course so that we can confirm that we have the necessary consents and permissions to do that analysis.

And lastly, there is the requirement to protect all that data, wherever it may reside. It must be kept secure, and it must also be auditable and verifiable. Let's look at these key requirements in more detail.

Is our data protected?

Our first concern is likely to be whether that data is safe from accidental damage or deletion, natural disasters, or man-made malice such as ransomware. This requires not just a second copy, but hardened, air-gapped backups, proof against unauthorized access. This can be used to restore files, databases and even complete systems.

Is our data useful?

If important information is spread across multiple systems and locations, it is harder to derive value from it. It may also lack coherence, with different versions of the same file existing in different places. We need to bring it all together – but this becomes ever-harder as we add more remote apps and services, each with its own data store and format.

Is our data legal?

For the 21st century organization, effective data governance is a must. But before we can comply with the relevant regulations and laws, we need to know what data we have, where it is stored, where it came from, what permissions or consents it is subject to, who is responsible for it - and for access to it - and so on.

Is our data in the right place?

We sometimes speak of 'hot' and 'cold' data, the former being in active use and the latter being inactive. Clearly, cold data should not be taking up space on expensive primary storage systems or being backed-up daily. We therefore need some way to ensure data is stored appropriately, transparently moving it to and fro as it warms and cools.

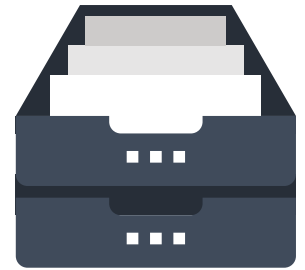
A backup is not an archive, and vice-versa

You might think that, because your backups contain all your old data, they can also be regarded as archives, but sadly, this is not the case. The two are performed and organized in quite different ways to serve different purposes. A backup is composed of files, application data-sets, system images and so on, and is designed so they can be restored or recovered. Each backup is time-specific and is usually immutable.

An archive stores historical data and changes as data gets added. In effect, it keeps notes on what happened and when, who did it, and so on. It might tell you how many widgets we sold each year for the last decade, say. A backup set cannot tell you that, nor can you restore a failed database server from an archive.

Backup and archive consolidation

For legacy reasons, many organizations have two long-term data stores: files and system data stored as backups, and simple data archives. However, just as other elements of IT are turning from physical to virtual, and from purpose-built to software-defined, we can - in a logical sense, at least - view backup and archiving as two different software-defined 'views' onto the same giant pool of data. This allows considerable efficiency savings, and also opens up the data for other uses, such as historical analysis.



There are challenges though, most notably how to create and manage that data pool, how to ingest data from all our different applications and sources, and how to use the pool to build and restore 'virtual backups'. Let's look in more detail at the main alternatives and options here, and consider some of the additional opportunities that consolidation can present.



Do-it-yourself

Historically, we would have built and maintained our data stores ourselves, using specialist skills and software tools, and on-site storage systems.

This can be complex and costly, however, especially when data volumes are constantly growing yet we want to analyze ever-more data from an ever-wider range of sources.



Moving to cloud

A modern alternative is to use cloud infrastructure. This offers the benefit of elasticity, both in terms of adding more storage capacity and of applying analytics to the data store. It can be implemented as privately-owned cloud infrastructure or by buying infrastructure-as-a-service (IaaS) from a cloud service provider.



Software-as-a-Service

Building and maintaining your own data store might be OK if you have the software and storage skills and it brings you significant competitive advantage. Otherwise, it may make more sense to look at pre-built online applications of this kind, typically offered as SaaS or a managed service. These will already have connectors for popular applications and their data.



Picking a price model

Whether it is for cloud storage, IaaS or SaaS, there are several pricing models in use. Per user pricing can be good for smaller organizations. Consumption-based pricing is straightforward once your initial uploads are complete. And egress-based pricing is good for data you might never need again, but expensive once you want to retrieve or analyze archived data.



Single or multi-tenant

In a SaaS or IaaS model, you typically have the option to choose a shared service, with your data software-partitioned from other tenants, or a private single-tenant one. Multi-tenant with per-user pricing can be cheaper for smaller users, but is often less flexible than single-tenant services in areas such as performance, security and data sovereignty.



Data tiering

It is usual for primary storage systems to contain a mix of frequently accessed 'hot' data and older 'cold' data. Tiering the latter to cloud storage frees up primary storage space, cuts the cost of storing that cold data, and reduces the time and capacity needed to backup those primary systems. Some tiering schemes can also keep cold data available for analysis and auditing.

UNDERSTANDING THE OPTIONS

Creating a converged data platform for backup and archiving offers opportunities, but how do you get there? As with so many IT projects, there are now several options, from acquiring off-the-shelf software solutions to public cloud-based managed services and SaaS (software-as-a-service) applications.

The list below therefore approaches the challenge from a functional perspective. For completeness, you could first assess how confident you are in your ability to answer these questions positively today. Then move on to answer them for the backup and archiving scheme (or schemes) that you're evaluating.

	<i>Level of confidence</i>		
	High	Medium	Low
Could I recover my remote application data in the event of a disaster at my SaaS or cloud provider?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Is my archived data accessible online for analysis, research, auditing, e-discovery, etc?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
How granular is my data sovereignty, can I control data location to at least app and country level?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Do I know what data my organization has in cloud and remote storage, and can I readily find it all?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Can I move data effectively both between different remote locations and between local and remote?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Can I do storage tiering for any data, not just for Windows shares?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Do I have full governance controls, e.g. retention dates, data classification, subject access requests?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Can my tools see the full range of data? Can they detect new file types and personal/sensitive data?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Do I have integrated analytics for full, yet secure, visibility into storage usage and consumption?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Can I set storage policies or rules, and classify and manage data dynamically?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Do I have full visibility of how my costs are likely to change with usage, searches, data recovery, etc?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

About Freeform Dynamics

Freeform Dynamics is an IT industry analyst firm. Through our research and insights, we aim to help busy IT and business professionals get up to speed on the latest technology developments, and make better-informed investment decisions.

For more information, please visit www.freeformdynamics.com.

About HubStor

HubStor helps organizations protect their data, simplify their IT infrastructure, and achieve compliance by uniquely converging data protection, archival, cloud storage, and content indexing through a simple and reliable software service.

For more information, please see: www.hubstor.net.

Terms of Use

This document is Copyright 2020 Freeform Dynamics Ltd. It may be freely duplicated and distributed in its entirety on an individual one to one basis, either electronically or in hard copy form. It may not, however, be disassembled or modified in any way as part of the duplication process. Hosting of the entire report for download and/or mass distribution by any means is prohibited unless express permission is obtained from Freeform Dynamics Ltd or HubStor. The contents contained herein are provided for your general information and use only, and neither Freeform Dynamics Ltd nor any third party provide any warranty or guarantee as to its suitability for any particular purpose.