The Security Pro's Guide to Cloud File Storage and Collaboration

Security for enterprise file sync and share services

Version 1.0
Released: September 12, 2014

Table of Contents

Author's Note

The content in this report was developed independently of any sponsors. It is based on material originally posted on the Securosis blog but has been enhanced and professionally edited.

Special thanks to Chris Pepper for additional editing and content support.

Copyright

The Rise of Cloud File Storage and Collaboration

Few technologies have invaded the enterprise as rapidly as cloud file storage and collaboration services. Often called File Sync and Share, these tools originated as consumer services to help people store, sync, and share files across computers and mobile devices. We hate to dip into hyperbole, but calling these tools groundbreaking is an understatement. Users can now access their files from any computer or device and share them with anyone anywhere, with ease and simplicity never seen before. To many consumers that *is* "the cloud".

These services showed their value so quickly that they inevitably made their way into the enterprise. Unfortunately few were architected to support the needs or security requirements of business. In response many organizations simply banned and blocked them, but when a tool provides substantial and demonstrable business benefit, use is inevitable — with or without support.

The category is most often called *enterprise file sync and share*, but we prefer the term *cloud file storage and collaboration* because many of the services and tools offer much more than basic syncing and sharing.

In response, enterprise-class options emerged. But many security professionals still struggle to understand the implications of cloud storage and collaboration, and the differences between consumer and enterprise-grade services. But some of these services offer security superior to traditional on-premise file storage.

The market is evolving incredibly rapidly, with new features and competitors showing up constantly. We see continuous change as everyone scrambles for competitive advantage in this wide-open new market.

Cloud file storage and collaboration services are an unavoidable disruptive innovation.

Security implications, but also significant benefits

The risk is obvious: pick the wrong service, or configure it incorrectly, and it is all too easy to effectively punch a hole in your firewall, offering unfettered access to your files to all the denizens of the Internet. Without centralized visibility and control employees make mistakes and expose sensitive information. Choose an insecure service and you will suffer the consequences of misplaced trust and exposed files.

Practically speaking, file security is something most organizations have struggled with since long before the Internet. But cloud services enable us to fail globally.

But pick the right service *and configure it properly*, and you can realize security benefits impossible with traditional file storage.

By centralizing all file storage, security gains a choke point for complete control and visibility. You can track the full history of access to all files from all users and devices. You can set enterprise-wide policies for how files are managed and

shared, both internally and externally. And unlike many other security approaches, you can do so while providing the business *something they want* and so are highly likely to adopt.

The alternative? Our existing troves of dozens, if not hundreds or thousands, of file repositories — all managed separately, with different policies, and usually without any real monitoring.

This paper delves into the security implications of cloud file storage and collaboration services. It covers the security fundamentals (including risks and benefits), core security features, and some more advanced security features such as encryption. We will separate out what you can expect from an enterprise-class service vs. a consumer offering — to help security professionals evaluate, select, and leverage the right options for their organizations.

Cloud or Local?

You will notice we tend to focus on cloud storage and collaboration services, rather than non-cloud sync and share tools you implement internally. Various products are available to create private file sync and share services, but they don't offer all the benefits of a true cloud service. Covering both would complicate this paper, and most of the concerns and questions we hear are about cloud-native services, not internal tools with similar functionality.

Understanding Cloud File Storage and Collaboration Services

Cloud File Storage and Collaboration (often called Sync and Share) is one of the first things people think of when they hear the term 'cloud', and one of the most popular product/service categories. It tends to be one of the first areas IT departments struggle to manage, because many users and business units want the functionality and personally use a wide variety of free and inexpensive options.

As you might expect from our inability to even agree on a category name, we see a wide range of different features and functions across the various services. We will start by detailing the core features with security implications, then the core security features themselves, and finally more advanced security features we see cropping up at some providers.

This is not merely a feature list — we cover each feature's security implications, what to look for, and how you might want to integrate it (if available) into your security program.

Overview and Core Features

When these services first appeared, the term *Cloud Sync and Share* did a good job of encapsulating their capabilities. You could save a file locally, it would sync and upload to a cloud service, and you could expose a *share link* so someone else on the Internet could download it. The tools offered various mobile agents for different devices, and consistently offered some level of versioning so you could recover deleted files and previous versions.

Most providers now offer much more than basic sync and share. Here are the core features which tend to define these services:

- *Storage:* The cloud provider stores files. This typically includes multiple versions and retention of deleted files. The retention period, recovery method, and mechanisms for reverting to a previous version vary widely. Enterprises need to understand how much is stored, what users can access/recover, and how this affects security. For example, make sure you understand version and deletion recovery so sensitive files you 'removed' don't turn up later.
- *Sync:* A software agent syncs local user directory (or server directory) changes with the cloud provider. Edit a file locally, and it silently syncs up to the server. Update it on one device and it propagates to the rest. The cloud provider handles version conflicts (which can leave version orphans in user folders). Users typically access alternate versions and recover deleted files through the web interface, and sometimes it also manages collisions.
- *Share:* Users can share files through a variety of mechanisms, including sharing directly with another user of the service (inside or outside the organization) which allows the recipient to sync the file or folder just like their own content. Shared items can be restricted to web access only; sharing can be open (public), restricted to registered users, or require a one-off password. This is often handled at the file or folder level, allowing capabilities such as *project rooms* to support

collaboration across organizations without providing direct access to any participant's private data. We will cover security implications of sharing throughout this report, especially how to manage and secure it.

- *View:* Many services now include in-browser viewers for different file types. Aside from convenience and ensuring users can see files, regardless of whether they have Office installed, this can also function as a security control, providing restricted access without allowing users to download files.
- *Collaborate:* Expanding on simple viewers (and the reason *Sync and Share* isn't entirely descriptive any more), some platforms allow users to mark up, comment on, and even edit collaborative documents directly in a web interface. This also ties into the project/share rooms mentioned above.
- *Web and mobile support:* The platform syncs locally with multiple operating systems using local agents (at least Windows, Mac, and iOS), provides a browser-based user interface for access from anywhere, and offers native apps for multiple mobile platforms.
- *APIs:* Most cloud services expose APIs for direct integration into other applications. This is how, for example, Apple is adding various cloud storage providers direct operating system integration in the next version of iOS. On the other hand, you could potentially link into APIs directly to pull security data or manage security settings.

These core features cover the basics offered by most enterprise-class cloud file storage and collaboration services. Most of the security features we are about to cover in the next section are designed to directly manage and secure these capabilities.

And since "Cloud File Storage and Collaboration Service" is a bit of a mouthful, for the rest of this paper we will simply refer to them as *cloud storage providers*.

Cloud Storage Security

Core Security Features

Core security features are those most commonly offered by enterprise-class cloud storage providers. Not every provider supports them, but this is where you should start to evaluate the security of a service. Keep in mind that different providers offer different levels of support for these features — it is important to dig into the documentation to understand how well each feature matches your requirements. Never assume marketure is accurate.

Security Baseline

Few things matter more than starting with a provider that offers strong baseline security. The last thing you want to do is trust your sensitive files to a company that doesn't consider security among their highest priorities. Key areas to look at include:

- *Datacenter security:* The provider should offer exceptional datacenter security — including physical controls, logical controls, and all the other essentials to reduce the risk of physical and technical attacks. You cannot necessarily assess this yourself, so look for up-to-date third-party certifications and attestations such as SOC 2 or ISO 27001. Generally the more the better, but make sure you sign the NDA and get the actual reports, when the assessment occurred, and which organization performed them.
- *Business continuity:* Short of a major asteroid impact, your provider should **never** lose your data. Their business continuity plans should account for multiple catastrophic outages, including complete loss of at least one data center. They should test their plans and provide assurance of their effectiveness, with documentation.
- *Encryption:* All customer data should be encrypted for compliance and to protect data from accidental loss of physical media. Ideally different keys should be used for different customers, and you should review their encryption and key management architecture. Note that this is not client-managed encryption — the provider manages the keys and can see your data — but it provides higher assurance against data spillage and exposure of physical media.
- *Application security:* The web application must be free from vulnerabilities to SQL injection, CSS, CSRF, and other application and business logic attacks. This also applies to direct API access. The provider should offer proof of ongoing security testing and web application security controls.
- *Internal controls:* Providers should have well-documented internal security controls to prevent both external attacks and insider abuse. Don't expect them to provide you with all the details, but one key area to ask about is administrative access and auditing. Essentially who can access your data, how, and how it is monitored for internal or external abuse. Two-factor authentication for administrators is a must.
- *Transparency, staffing, and documentation:* The short version: you want a provider with a dedicated security team, who is transparent about security operations and provides good documentation — both of their inherent security, and how to secure the services you use with their platform features.

This isn't an exhaustive list, but some key areas to focus on in your initial assessment. Without a solid security baseline it really doesn't matter what else the service offers.

Identity and Access Management

Managing users and access are the most important features after the security baseline. The entire security and governance model relies on it. These are the key elements to look for:

- *Service and federated IDM:* The cloud service needs to implement an internal identity model to allow sharing with external parties without requiring those individuals or organizations to register with your internal identity provider. The service must also support federated identity so you can use your internal directory and don't need to manually register all *your* users with the service. SAML is the preferred standard. Both models should support API access, which is key to integrating the service with your applications as back-end storage.
- *Authorization and access controls:* Once you establish and integrate identity, the service should support a robust and granular permissions model. The basics include user and group access at the directory, subdirectory, and file levels. The model should integrate internal, external, and anonymous users. Permissions should include read, write/edit, download, and view (web viewing but not downloading of files). Additional permissions manage who can share files (internally and externally), alter permissions, comment, and delete files.
- *Device control:* Cloud storage services are very frequently used to support mobile users on a variety of devices. Device control allows management of which devices (computers and mobile devices) are authorized for which users, to ensure only authorized devices have access.
- *Two-factor authentication (2FA):* Account credential compromise is a major concern, so some providers can require a second authentication factor to access their services. Today this is typically a text message with a one-time password sent to a registered mobile phone. The second factor is generally only required to access the service from a 'new' (unregistered) device or computer.
- *Centralized management:* Administrators can manage all permissions and sharing through the service's web interface. For enterprise deployment this includes enterprise-wide policies, such as restricting external sharing completely and auto-expiring all shared links after a configurable interval. Administrators should also be able to identify all shared links without having to crawl the directory structure.

> An external authenticated user is one who registers with the cloud provider but isn't part of your organization. This is important for collaborative group shares, such as deal and project rooms. Most services also support public external shares, but these are open to the world. That is why providers need to support *both* their own platform user model and federated identity to integrate with your existing internal directory.

Sharing permissions and policies are a key differentiator between enterprise-class and consumer services. For enterprises central control and management of shares is essential. So is the ability to manage who can share content externally, with what permissions, and to which categories of users (*e.g.,* restricted to registered users vs. via an open file link). You might, for example, only allow employees to share with authenticated users on an enterprise-wide basis. Or only allow certain user roles to share files externally, and even then only for in-browser viewing, with links automatically expiring in 30 days.

Each organization has its own tolerances for sharing and file permissions. Granular controls allow you to align your use of the service with your existing policies. They can also provide a security benefit, providing centralized control over all storage — unlike the traditional model where you need to manage dozens or even thousands of different systems with different authentication methods, authorization models, and permissions.

Audit and Transparency

One of the most powerful security features of cloud storage services is a complete audit log of all user and device activity. Enterprise-class services track **all** activity: which users touch which files from which devices. Features to look for include:

- *Completeness of the audit log:* It should include user, device, accessed file, what activity was performed (download/view/edit, with before and after versions if appropriate), and additional metadata such as location.
- *Log duration:* How much data does the audit log contain? Is it eternal or does it expire in 90 days?
- *Log management and visibility:* How do you access the log? Is the user interface navigable and centralized, or do you need to hunt around and click individual files? Can you filter and report by user, file, and device?
- *Integration and export:* Logs should be externally consumable in a standard format to integrate with existing log management and SIEM tools. Administrators should also be able to export activity reports and raw logs.

These features don't cover everything available, but they are the core security capabilities enterprise and business users should have to start with.

Additional Security Features

The core security features are a baseline which enterprise and business customers should look for when selecting a service for their organization, but the various services offer a plethora of additional security features. Providers see them as a way to entice enterprise users onto their services, show advantages over traditional storage infrastructure, and create competitive differentiation.

So security is used as both a competitive baseline (the minimum capabilities you need to compete for enterprise customers over a consumer offering) and a differentiator, which is why we see new capabilities appearing constantly. The odds are high that this report won't cover everything available by the time you read it.

Universal Search and Investigation Support

As we described earlier, most cloud storage providers track all files, offer content search, and track every user and every device that accesses each file — including who viewed it in a web browser, downloaded a copy, or synced it with a computer or mobile device.

That single central control point enables powerful security capabilities. Worried a document leaked? Find all copies and the entire access history. An obvious caveat applies: once a file leaves the service it isn't tracked, but at least you have a starting point to identify where it went. This is often one of the more difficult first steps in any leak/breach/abuse investigation, because traditional storage products rarely track this level of detail.

Enhancing this is full content indexing and search. This isn't a pure security feature, but it enables you to search your entire cloud storage repository for keywords or specific content. Some providers offer options for more advanced searches, particularly *regular expressions*. This is also quite useful for non-security reasons, so we expect indexing and searching capabilities to evolve over time — but make sure you understand what your provider supports now.

Another limitation is that providers don't support every possible document type. For example, the odds are low that your CAD file format is supported today. Typically standard Office and text formats are supported — check with potential providers.

Client-Managed Encryption

All enterprise-class cloud storage providers encrypt data in their backend, but they manage the keys and can thus technically see your content. There are now third-party security vendors who encrypt cloud data using different approaches, and some cloud storage vendors are adapting their architectures to allow customers to encrypt directly within the service, but to control their own keys.

This is a different approach than a third-party tool. Your cloud provider still handles the encryption in their backend but you have your own encryption keys. There are two major options:

- The cloud platform endpoint agents, or additional, independent encryption agents, handle encryption operations synchronized with your enterprise key store. For this to work they need to include the capability in both workstation and mobile agents, and a mechanism for integrating key distribution.
- The cloud platform manages encryption in their backend, but offers mechanisms for enterprise users to provision and manage their own keys. There are a few ways to handle this but it typically involves a Hardware Security Module (HSM) located in the cloud provider's data center yet managed by the client, in a client data center, or at an infrastructure cloud provider. The important part is that the *customer,* rather than the cloud provider, has exclusive access to encryption keys. Technically they are exposed in the cloud provider's data center during cryptographic operations, but if architected correctly the risk of key exposure can be minimized.

We won't be surprised to see other approaches develop over time, but we know these two are or soon will be on the market. In both cases the customer needs their own key management infrastructure.

One major warning: encrypting data with your own key breaks most or all collaboration features, including indexing/search, because the cloud provider cannot read your content. So this is something you should generally limit to your most sensitive data. Apply it to everything and you may see users circumventing encryption to take advantage of platform features you do not support.

Data Loss Prevention

Full-text indexing and search, combined with a complete audit log of all activity associated with a file, satisfies our definition of basic content-aware DLP.

In addition, cloud providers can offer real-time monitoring of all content based on search terms, and tie queries to enforcement policies. For example a cloud storage provider could quarantine a file and alert an administrator any time a credit card number is found. This enables enterprise-wide content policies for the entire cloud storage platform.

More advanced rules can apply by user or group, restricting only certain activities — perhaps "never share a file with PII or this keyword in it externally", or any other combination of analysis and rules, such as device restrictions. DLP combines full content indexing and search with persistent policies for near-real-time content-aware protection.

The market for integrated DLP is still extremely young, and when available features tend to be limited. Third-party integration can provide more capability, and as with everything else we expect to see capabilities expand at a reasonable pace. In discussions with clients this seems to be a popular requirement which will continue to push the market along.

DRM/IRM

Digital Rights Management, also called Information Rights Management, is encrypting data and then limiting usage according to access (rights) policies. For example you might allow someone to view a file, but not email or print it.

Cloud file storage and collaboration services often include in-browser readers and granular rights policies — together they can be leveraged to provide limited DRM. Set a policy so a file can only be viewed in a browser and never downloaded, and you can restrict activity.

But to truly provide DRM, a service should include in-browser protections against actions such as copy and screen capture. Ideally the platform offers document markup and editing directly or through third-party integration — not merely viewing. On the horizon we expect services that define DRM policies at the file level via integration with on-premise enterprise DRM — not that anyone is really using that today.

Device Security

At the most basic level, the cloud service should support restricting content only to approved devices, and include integration for federated identity. The service should also track all files each device downloads. Content on devices should be encrypted, using either a proprietary mechanism or the device's native app encryption.

For content security some providers allow remote wiping of the service's app, effectively destroying its data. Another option is to remotely log out the application, which also removes access to the data, although depending on the encryption a forensic recovery may still be possible if the data (or encryption key) isn't deliberately wiped.

Providers are also exploring additional mobile device features, predominantly centered around Mobile Device Management-like capabilities, to better characterize device configuration and even detect certain forms of jailbreaking or other compromise that could expose data.

API Support

Nearly all cloud storage providers support a robust API for integration into other applications and services. This is becoming an essential baseline feature in this market. But not all APIs are created equal. From the security standpoint, all security features should be API-enabled to support integration with existing and future tools.

Security Tool Integrations

While APIs provide the hooks, there are a few common categories of security tools we tend to see pre-integrated. This is often enabled by vendor partnerships, the API support mentioned above, and/or making data from the service (especially security audit logs) accessible in standard formats. Established integrations save time and are officially supported, unlike something you code up yourself.

The most common and useful integrations include:

- *Cloud security gateways:* This product category has a number of names, but includes products that identify cloud services used by your organization and add security using multiple techniques — such as more granular controls, inline encryption, better security management, and DLP.
- *Encryption gateways:* While sometimes implemented as a feature of cloud security gateways, these stand-alone products proxy connections to the cloud provider and selectively encrypt data. We aren't big fans of this approach due to logistics considerations and its impact on cloud service features, but it is a viable option when encryption is mandated but not supported by the vendor.
- *Digital Rights Management:* We see very few successful enterprise-wide DRM deployments, but new providers focused on mobile collaboration have emerged to support valuable use cases.
- *Mobile Device Management:* This integration allows greater granularity and control of which devices and users can connect to the cloud service.

- *SIEM/log management:* The SIEM tool collects activity directly from the cloud provider and integrates it into consolidated logging, alerting features, and even correlation with other security events.

We also see integrations developing with electronic discovery tools, standard DLP, business intelligence, and others as new use cases emerge. As enterprise adoption of cloud file storage increases we expect to see many more integrations.

Conclusions, and a Caution

This biggest concern when using a cloud file storage and collaboration service is the worry of entrusting your sensitive information — sometimes your *most* sensitive information — to an outsider who mixes it up with everyone else's data. That can be a tall mental hurdle to leap, especially when your experience is with consumer services with spotty security records.

But using a secure enterprise-class service brings demonstrable security advantages. It reduces file sprawl and provides an overall view of all activity that is effectively impossible with traditional storage paradigms. Add some security features such as content indexing and search, and some creative security solutions emerge. Full visibility and control are very powerful capabilities.

Despite these advantages you need to be careful. One security fumble can expose all your files to the entire Internet. These providers need to be absolutely best-in-class, with extensive transparency, strong governance, continuous deep security testing, and dedicated security teams. Even then you might occasionally need to add additional encryption to provide assurance, especially if you are concerned with government snooping.

Do your homework and choose wisely, and you will gain control over your files which most security professionals never dream of.

Who We Are

About the Analyst

Rich Mogull, Analyst and CEO

Rich has twenty years experience in information security, physical security, and risk management. He specializes in cloud security, data security, emerging security technologies, and security management. Rich is the primary developer of the Cloud Security Alliance CCSK training program. Prior to founding Securosis, Rich was a Research Vice President at Gartner on the security team where he also served as research co-chair for the Gartner Security Summit. Prior to his seven years at Gartner, Rich worked as an independent consultant, web application developer, software development manager at the University of Colorado, and systems and network administrator. Rich is the Security Editor of TidBITS, on the advisory board of DevOps.com, and a frequent contributor to publications ranging from Information Security Magazine to Macworld. He is a frequent industry speaker at events including the RSA Security Conference, Black Hat, and DefCon, and has spoken on every continent except Antarctica (where he is happy to speak for free — assuming travel is covered).

About Securosis

Securosis, L.L.C. is an independent research and analysis firm dedicated to thought leadership, objectivity, and transparency. Our analysts have all held executive level positions and are dedicated to providing high-value, pragmatic advisory services.

We provide services in four main areas:

- Publishing and speaking: Including independent objective white papers, webcasts, and in-person presentations.
- Strategic consulting for end users: Including project accelerator workshops, product selection assistance, technology and architecture strategy, education, security management evaluations, and risk assessments.
- Strategic consulting for vendors: Including market and product analysis and strategy, technology guidance, product evaluations, and merger and acquisition assessments.
- Investor consulting: Technical due diligence including product and market evaluations, available in conjunction with deep product assessments with our research partners.

Our clients range from stealth startups to some of the best known technology vendors and end users. Clients include large financial institutions, institutional investors, mid-sized enterprises, and major security vendors.

Securosis has partnered with security testing labs to provide unique product evaluations that combine in-depth technical analysis with high-level product, architecture, and market analysis.