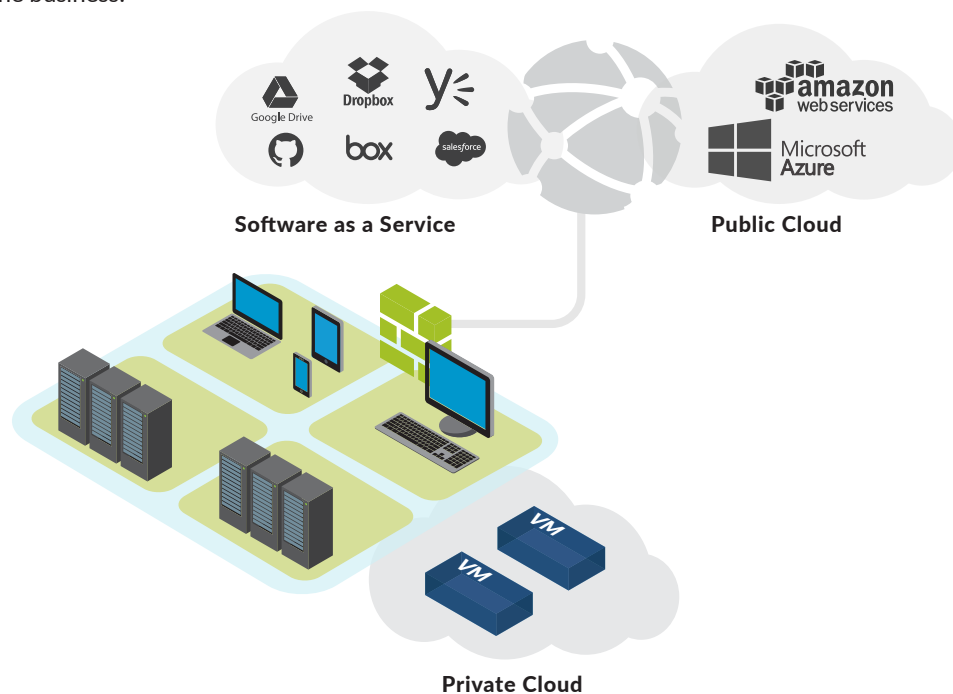


# COMPREHENSIVE DATA SECURITY IN THE CLOUD

The security perimeter once familiar to the enterprise has become incredibly fragmented. Data and applications reside everywhere: on the network, endpoints and in the cloud. The cloud, in particular, is seeing huge growth with enterprises adopting these environments at a rapid pace. According to Gartner, 55 percent of large enterprises will successfully implement an all-in cloud SaaS strategy by 2025.<sup>1</sup> Combined with an increasingly mobile and global workforce, and more importantly, increasingly distributed SaaS cloud environments, organizations are now faced with securing a multitude of applications, users, devices and networks – all hosting sensitive data that is critical to business growth, reputation and customer trust.

## Common Cloud and SaaS Application Threats

The cloud is no longer an “exploration exercise.” It is the single biggest computing paradigm to unfold since the early 2000s. Applications and data in the cloud need to be protected as vigilantly as on-premise applications and data. Why? The cloud, particularly SaaS applications, often hosts businesses’ most sensitive data, such as critical intellectual property (IP) or personally identifiable information (PII) (e.g., Social Security or credit card numbers). A breach or inadvertent exposure of this type of data can result in compliance violations, loss of customer trust and financial impact to the business.



**Figure 1: Protection for various cloud environments**

<sup>1</sup> <http://www.gartner.com/smarterwithgartner/cios-flip-for-cloud-saas/>

When protecting against data loss or exposure in the cloud, there are several risks outlined in detail below that must be secured.

- **Accidental data exposure.** Because cloud and SaaS applications are designed for easy sharing, data runs the risk of becoming unintentionally shared or exposed. All the user needs to do to share a file with a colleague is paste a link into an email, or share directly within an application by inputting the colleague's email address. But it's equally easy to accidentally share confidential information with the wrong person. For example, an email with a shared link could be forwarded outside of the organization, or sent to the wrong person when the system attempts to autocomplete.
- **Insider threats.** Insider threats come from malicious internal users who purposefully share data for theft or revenge. Perpetrators are often current employees with criminal intent, or former employees who recently left the company, whether of their own volition or not, who want to harm the business. Unfortunately, causing harm can be relatively straightforward because data in the cloud falls beyond the purview of traditional perimeter firewalls and data loss prevention techniques.

For example, an employee leaving a company can simply change the folder permissions within an application to be shared publicly or via a personal email address to steal data remotely. This type of data exfiltration can also be executed with malware, by pulling data from a SaaS application or extracting files stored on a device that uses the cloud to send data outside of the organization.

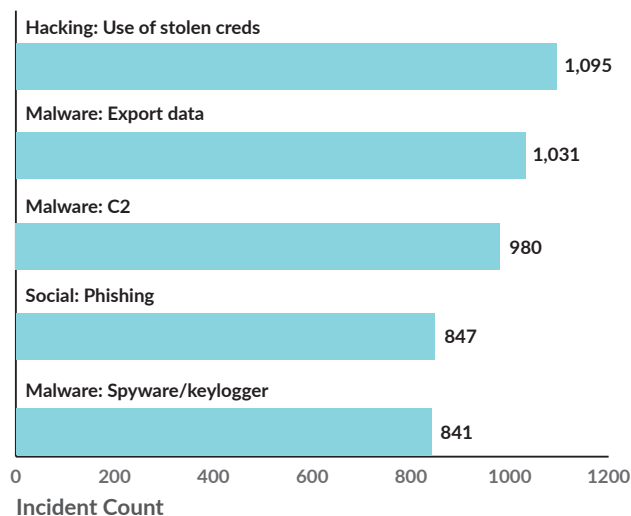
- **Credential theft.** Stolen credentials represent the keys to the digital universe. Once successfully stolen, user credentials provide attackers with the power to access everything the legitimate user can. This can be especially problematic for cloud access.

Cloud services are often protected solely by user credentials – a simple user name and password – and in the instance of credential theft, access provides an attacker a direct route to the valuable data stored in the environment. One of the most common ways attackers steal credentials is via phishing, where an email recipient is lured into logging into an account. Attackers also target credentials through malware. For example, the Sofacy threat actor group has used their [XAgent tool](#) against both Microsoft® Windows® and Apple® macOS™ systems.

As organizations continue to move data off-premise and into the public cloud and SaaS environments, it's important to remember that data security is always the sole responsibility of the enterprise. Security tactics should be implemented to prevent data loss via the use of stolen credentials and unauthorized access.

### Three Steps to Securing Your Data

While the concept of securing data in the cloud seems straightforward, much like securing data on-premise, it can be an overwhelming undertaking for many organizations. The average organization often has little idea of where its sensitive data is stored, let alone how it's being used, making securing it a tricky task. What follows are three steps organizations can follow to effectively secure business-critical data in the cloud.



**Figure 2: Top threat action varieties within incidents involving credentials**

### Step One: Do a Self-Assessment

- **Take inventory.** Assess all cloud migrations and cloud-based applications that your organization has adopted.
  - Understand key business drivers that led to cloud adoption initially.
  - Note how each application is being used, why, and the types of data being stored or transferred.
  - Inventory applications that may not necessarily be sanctioned, but that employees and other approved users still use for work-related or personal reasons.

This comprehensive catalogue of information will provide a complete snapshot of applications and data in the cloud that need to be classified and prioritized for security and compliance.
- **Classify data.** There are a few reasons to do this. First, it's critical to determine the level of data sensitivity so that IT can prioritize resources and develop effective policies to control access. Second, and more importantly, businesses must control the exfiltration of data that should not leave the organization under any circumstances, whether to meet industry compliance mandates or general business regulations.

**Confidential Data Examples:**

- Intellectual property
- Credit card numbers
- Social Security numbers
- Personal health information
- Other personal identifiable details

- INTERNAL** Not explicitly shared
- COMPANY** Explicitly shared with collaborators within the company
- EXTERNAL** Explicitly shared with external collaborators outside the company
- PUBLIC** Accessible by anyone on the web

Simplicity goes a long way when classifying data. Minimize classification patterns to categories like public, internal-only (sometimes defined further as private or proprietary) and confidential data.

- **Define risk and governance.** With data classified, define level of risk and develop a data governance model to intelligently dictate who should have access to what. Access should be based on the combination of risk level and legitimate business need. Then comes enforcement, also referred to as data governance. Data governance accounts for standards, procedures and policies that define how an organization manages the availability, usability, integrity and security of the data employed in the enterprise – in this instance, the data hosted in the cloud.
 

With a clear understanding of how cloud services, such as Box or Office 365®, are being used, access controls can be defined and enforced through the implementation of data-driven business policies. The core principle here, beyond defining exactly how data is to be used within a cloud service, is to ensure that the data is only used by those with authorization.

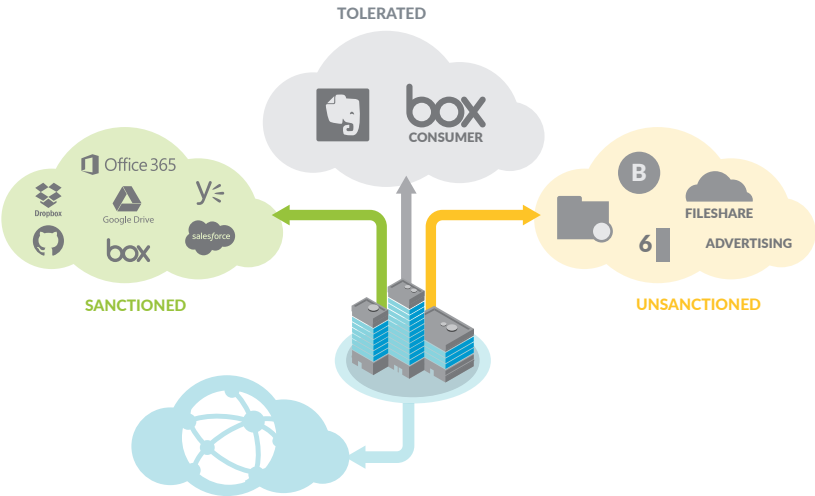


Figure 3: Classification of Cloud Applications

### Step Two: Get the Right Tools to Prevent Data Exposure

Compliance regulations are becoming more stringent and cybercriminals increasingly target assets in the cloud, knowing that's where more and more organizations keep valuable data. It's critical to implement the right data loss prevention tools to keep sensitive, business-critical data in the cloud protected, both from accidental exposure and malicious exfiltration.

Cloud-based data loss prevention technology is designed to reduce an organization's risk posture by preventing end users from sharing data, whether accidentally or maliciously. It enables organizations to protect sensitive information through detailed data classification across three primary use cases: controlling the migration of data from your traditional infrastructure to the cloud; protecting and controlling access to data already in the cloud, done through policy; and locating sensitive data leaked to the cloud. Note, however, data loss prevention tools for the cloud should be deployed inline and integrate with cloud providers' application programming interfaces (APIs). This will help to ensure the visibility and control required to comply with data privacy and protection regulations, as well as identify and defend against inherent risks.

Several controls are available to minimize the risk of data loss and exposure in the cloud:

- **Application discovery.** Businesses can scan all network traffic and detect all applications in use on the network, as well as identify what data is being transferred to the cloud.
- **Content scanning.** Organizations can enforce granular data loss prevention policies based on keywords, file characteristics and other content patterns. Data loss policies allow you to identify, monitor and automatically protect sensitive information. Predefined policy templates are available to help businesses maintain compliance with industry standards and regulations, such as HIPAA, PCI DSS, etc. You can also define custom policies based on business needs, such as policies to classify and control content that includes sensitive information such as credit card numbers, tax IDs or other similar keywords.

- **Matching data patterns:** To help prevent data loss, content can be scanned to identify potential matches based on data and text patterns defined in policies. You can use predefined patterns or develop custom patterns known as regular expressions.

#### Examples of data patterns include:

- PCI data (e.g. credit card numbers, magnetic stripe data)
- IBAN numbers
- Magstripe data
- WIFT codes

```
Mastercard 471632232459327 471
6347094324714 4716483924720670
Mastercard 471632232459327 4716
347094324714 47163924720670
Mastercard 471632232459327 47163
47094324714 4716483924720670
```

- **Machine learning.** It's not always possible to create a pattern for all data types and appearances. Machine learning algorithms can come to understand what sensitive data looks like over time.

**For example,** known machine learning algorithms identify the content of text so that security tools can classify a document as "financial" or "legal" with a certain level of confidence.

- **Block untrusted users.** Organizations should list all trusted domains, inclusive of acquisitions and third-party vendors, such as law firms or marketing firms. Beware of employees with malicious intent and block the sharing of information to untrusted domains, such as Gmail accounts or other personal email addresses.
- **Detect anomalous user activity.** Stolen credentials or malware may trigger unusual behavior such as large data downloads or user login attempts from varying locations within a short time frame. Businesses should implement tools that alert the administrators or users of these types of anomalous activities, as well as suggest possible remediation actions.

- **Find and remove malware.** Cloud applications are increasingly becoming the norm for the insertion of malware into an organization. Businesses must be able to find and remove known and unknown variants of malware sbefore the threats spread throughout their infrastructure.
- **Multiple deployment options.** Data security tools need to be able to enforce policy for internet traffic as it traverses the perimeter (data in motion), but certificate pinning, encryption and remote users can potentially bypass inline security tools. For complete protection, it is essential to have the ability to scan activity and data within the cloud (data at rest).
- **Broader context.** Organizations can reduce false positives and negatives with the availability of broader context from other agents, devices or security tools.

In summary, the classification and control process is entirely driven via policy enforcement, starting with identification. An organization must determine where, when and how to protect its sensitive content by enforcing policies based on a variety of predefined conditions and automatic follow-up actions.

In large enterprises, the sheer number of employees can initially lead to thousands of risks and threats. It is important to address these risks automatically through quarantine; restricting data exposure and sharing permissions; end-user coaching and notification; and logging for audits.

### Step Three: Implement an Ongoing Feedback Loop

New enterprise applications are consistently being consumed by teams and employees to help drive productivity and get the job done. This has become apparent based on data presented in Okta's 2017 [Businesses @ Work](https://www.okta.com/Businesses-at-Work/2017-01/) report.<sup>2</sup> As such, existing data risk and governance models have the potential to become quickly outdated if measures aren't taken to continuously update based on new applications, technologies and data types. As data models and applications evolve, it is important to update your risk criteria to account for the changing threats and attack vectors.

- **Ongoing audit sessions and reporting trails.** These actions provide visibility into information being accessed and transmitted so that necessary adjustments to cloud application and data access rules and policies can be applied. Data loss prevention tools should be able to export all activity to the SIEM tool in use by the organization for a comprehensive summary of all risk and application usage behaviors so that appropriate updates to risk criteria and policies can be made. Security and business leaders should participate in this process to ensure alignment.
- **Ongoing education.** The user community – employees, partners and any third party who may have access to your applications and data in the cloud – must be educated and coached regularly on the risks their dealings with cloud data might pose to the business, as well as procedures and protocols to follow when working with sensitive applications and data in the cloud. Defining processes and advising employees of needed changes to behaviors will ease the burden over time.

**For example,** a user may decide to use Dropbox® to share a file with a client, although Dropbox is not the sanctioned application for use. Rather than deny use altogether, which might upset or demotivate the user, provide a list of rules – a security and compliance checklist – that should be followed when using the application.

### The Rise of the Cloud Access Security Broker

As businesses have become more concerned about the volume and sensitivity of data being transferred, stored and shared in cloud-based SaaS environments, there has been rapid adoption of the cloud access security broker market (CASB). A CASB accesses cloud-based services, primarily focused on addressing security gaps in highly productive and collaborative SaaS applications, such as Box, Dropbox, GitHub®, Google® Drive, Microsoft Office 365 and Salesforce.com®, where traditional security products have not been able to keep pace. A CASB, essentially a data loss prevention provider, supplies organizations with key SaaS security functions: visibility, control, compliance and threat prevention.

<sup>2</sup> <https://www.okta.com/Businesses-at-Work/2017-01/>

However, organizations who prefer to deploy a CASB for SaaS security are often faced with unnecessary network complexity and inefficient security. To avoid such issues, they should approach SaaS security methodically and strategically. Think first about the main SaaS security issues that need to be addressed. It's not necessarily a requirement to deploy another disparate point product – in this case a CASB – to protect your data in the cloud. In fact, it's much easier to extend and enforce your existing enterprise security tools and policies across the cloud to mitigate risk and prevent data loss as users and devices access your cloud resources.

### A Platform Approach to Data Loss Prevention in the Cloud

At Palo Alto Networks®, we're addressing cloud data loss from an entirely new perspective – the platform. The Palo Alto Networks Next-Generation Security Platform approach to prevention begins with visibility into the applications in use on your network, in the cloud and SaaS environments, as well as on endpoints. The platform reduces threat exposure by controlling sanctioned and unsanctioned application usage, preventing known and unknown threats within allowed traffic, and continually strengthening prevention efforts based on ongoing threat analysis.

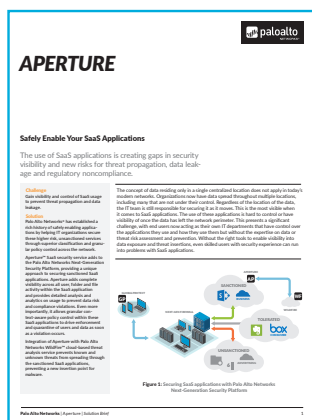
*Data Security in the Cloud Checklist*

You Need:	Palo Alto Networks Platform	Other CASB vendors
Application discovery and visibility	✓	✓
Matching data patterns	✓	✓
Blocking of untrusted collaborators	✓	✓
Detection of anomalous user behavior	✓	✓
Deployment flexibility: in-line and API	✓	✓
Data classification based on machine learning	✓	
Identification and removal of known and unknown malware	✓	
Global threat intelligence	✓	
Credential theft prevention	✓	
Broader context across network, cloud and endpoints	✓	

The Palo Alto Networks Next-Generation Security Platform provides complete cloud protection through detailed visibility and granular control, data governance, automated risk remediation and malware prevention – all at a lower total cost of ownership.

Given the agility and scalability advantages, more and more business-critical workloads and data exist in the public cloud. Simultaneously, the ways users access data in the cloud – smartphones, tablets and other remote access devices – continue to evolve and expand, from the corporate campus to virtually anywhere. Because the risk of enterprise data loss increases in tandem, organizations are demanding security frameworks that can seamlessly and effectively account for data loss prevention requirements in cloud-centric architectures. To learn more, check out the following resources:

## Aperture



**APERTURE**

**Safely Enable Your SaaS Applications**

The use of SaaS applications is creating gaps in security visibility and new risks for threat propagation, data leakage and regulatory noncompliance.

**Challenges**

The concept of data residing only in a single centralized location does not apply to SaaS. Visibility and control of data is fragmented across multiple locations, making it difficult to track and control. Regardless of the location of the data, the data is still accessible for viewing and transfer. This is the exact challenge when it comes to SaaS applications. The use of these applications is hard to control or have visibility of since the data has left the corporate perimeter. This presents a significant challenge for security teams. The use of SaaS applications is hard to control or have visibility of since the data has left the corporate perimeter. This presents a significant challenge for security teams. The use of SaaS applications is hard to control or have visibility of since the data has left the corporate perimeter. This presents a significant challenge for security teams.

**Figure 1** – Securing SaaS applications with Palo Alto Networks Next-Generation Security Platform

## App-ID



**APP-ID**

**A foundation for visibility and control in the Palo Alto Networks Security Platform**

App-ID uses multiple identification techniques to determine the exact identity of applications traversing your network – irrespective of port, protocol, source/destination, or encryption. Identifying the application is the very first task performed by App-ID, providing you with the knowledge and flexibility needed to safely enable applications and secure your organization.

**App-ID is a general purpose identification technique that identifies applications traversing your network by inspecting a variety of parameters, including:**

- **Behavioral analysis** – understanding the behavior of an application by observing its actions and interactions with other applications on the network.
- **Signature analysis** – comparing the application's behavior to known signatures.
- **Deep application visibility** – inspecting the application's behavior at the network layer.
- **Machine learning** – using machine learning to identify unknown applications.

**App-ID is a general purpose identification technique that identifies applications traversing your network by inspecting a variety of parameters, including:**

- **Behavioral analysis** – understanding the behavior of an application by observing its actions and interactions with other applications on the network.
- **Signature analysis** – comparing the application's behavior to known signatures.
- **Deep application visibility** – inspecting the application's behavior at the network layer.
- **Machine learning** – using machine learning to identify unknown applications.

## WildFire



**WILDFIRE**

**Automatically Prevent Highly Evasive Zero-Day Exploits and Malware**

Palo Alto Networks' WildFire™ cloud-based threat analysis service is the industry's most advanced analysis and prevention engine for highly evasive zero-day exploits and malware. The service employs a unique machine-learning approach combining dynamic and static analysis, innovative machine learning techniques, and a groundbreaking base metal analysis environment to detect and prevent even the most evasive threats.

**WildFire threat analysis and prevention service:**

- Detects evasive zero-day exploits and malware.
- Analyzes and classifies threats based on dynamic and static analysis, innovative machine learning techniques, and a groundbreaking base metal analysis environment.
- Provides automatic remediation and prevention for detected threats.
- Integrates with Palo Alto Networks' security platform for real-time threat prevention.

**Find the Unknown With a Unique Machine Learning Approach**

WildFire's unique machine learning approach is designed to identify and prevent threats that are unknown to traditional security engines. WildFire's unique machine learning approach is designed to identify and prevent threats that are unknown to traditional security engines.

## GlobalProtect



**GLOBALPROTECT**

**Prevent Breaches and Secure the Mobile Workforce**

GlobalProtect extends the protection of the Palo Alto Networks Next-Generation Security Platform to the members of your mobile workforce, no matter where they may go.

**Key Usage Scenarios and Benefits:**

- **Secure Remote Access** – Provide secure access to corporate resources from any location.
- **Advanced Threat Prevention** – Protect mobile devices from malware and other threats.
- **USB Filtering** – Control data flow to and from mobile devices.
- **Secure Access to SaaS Applications** – Control access to SaaS applications from mobile devices.
- **BYOD** – Support Bring Your Own Device (BYOD) environments.
- **Strongest Network Security** – Provide the highest level of security for mobile devices.



4401 Great America Parkway  
 Santa Clara, CA 95054  
 Main: +1.408.753.4000  
 Sales: +1.866.320.4788  
 Support: +1.866.898.9087  
 www.paloaltonetworks.com

© 2017 Palo Alto Networks, Inc. Palo Alto Networks is a registered trademark of Palo Alto Networks. A list of our trademarks can be found at <http://www.paloaltonetworks.com/company/trademarks.html>. All other marks mentioned herein may be trademarks of their respective companies. data-security-in-the-cloud-eg-050117