EMC CLOUDARRAY BEST PRACTICES: CLOUD STORAGE SECURITY AND DATA INTEGRITY
A multifaceted approach to data security & integrity in the cloud

ABSTRACT
This white paper outlines how EMC CloudArray uses best practices for protecting data in the cloud from unauthorized access and maintaining data integrity. By using these guidelines as a basis, you can make data storage in the cloud as secure or, surprisingly, even more secure than local storage.
To learn more about how EMC products, services, and solutions can help solve your business and IT challenges, contact your local representative or authorized reseller, visit www.emc.com, or explore and compare products at https://store.emc.com/cloudarray

Copyright © 2014 EMC Corporation. All Rights Reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

The information in this publication is provided “as is.” EMC Corporation makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com.

Part Number H13455
# TABLE OF CONTENTS

EXECUTIVE SUMMARY ........................................................................................................... 4

DOESN’T MY CLOUD PROVIDER SECURE THE DATA?.......................................................... 4

ENSURING DATA SECURITY THROUGH GATEWAY ENCRYPTION ...................................... 4

PROPER ENCRYPTION KEY MANAGEMENT ........................................................................... 5

THE ROLE OF ENCRYPTION IN DATA INTEGRITY ............................................................... 6

THE CHALLENGES OF MASS DATA IMPORT .......................................................................... 6

WHAT ABOUT HACKING AND GOVERNMENT SPYING?....................................................... 7

SUMMARY OF BEST PRACTICES ............................................................................................ 7

THE BOTTOM LINE ............................................................................................................... 8
EXECUTIVE SUMMARY

Data security has always been a top consideration for organizations storing sensitive data, but rarely has it garnered attention as immense as in the context of cloud storage. As more companies turn to the cloud as a solution to the rising costs, continued frustration and disruptive maintenance requirements of traditional data storage, IT departments find that legal, risk and compliance executives are taking an even greater interest in data storage than ever before.

One reason for the keen interest in security may be psychological, driven by the thought of storing data outside of the four walls of an organization. Other reasons may include concerns about government spying, hackers and cloud provider security or viability.

If you’re a risk-averse organization, you might choose to banish cloud storage completely as the path of least resistance, but taking such an approach does not come without consequence. In an age of exponential data growth, you and your company must either resign yourself to a comparably exponential expenditure on data storage or seek new storage strategies that are more cost-effective and require less administration and management.

So how do you mitigate the risks of cloud storage? Encryption, certainly, but where, when and what kind? Encryption of data and encrypted communications provide a starting point to counter data security risks and fears. But there are actually many facets of security when storing data in the cloud and encryption alone does not ensure adequate protection.

DOESN’T MY CLOUD PROVIDER SECURE THE DATA?

It’s no secret that government agencies have gained unauthorized access to major data centers. If data is not encrypted at-rest (when data is not being moved or used), unknown parties might be able to access data in the cloud, even if the cloud provider is implicitly trusted and has a strong record of following standard industry practices. It has become clear: All data not encrypted at-rest is exposed.

So let’s start with the bare minimum. The native APIs for storing and recalling data in the cloud almost universally employ in-flight encryption via HTTPS/SSL, which is the industry standard for secure communications used by financial organizations throughout the world. This protects against network intrusion by a third-party during your transmissions to the cloud.

Nearly all cloud storage providers offer some additional level of security for data. For example, many cloud providers have introduced server-side encryption at-rest, using technologies such as 128-bit Advanced Encryption Standard AES-128 or AES-256 bit encryption – the same type of strong encryption used by US government agencies for top secret data. With strong encryption such as these it would theoretically take even a supercomputer billions of years² to decrypt using a brute force attack, making it essentially unbreakable.

Although server-side encryption represents a substantial improvement over unencrypted data, it does not completely close the possibility of a data breach. Since the cloud provider performs the server-side encryption, the provider then owns the encryption keys and can theoretically access the data. Moreover, a third-party (either malicious or legitimate), such as a provider employee or an outside/governmental agency, can theoretically obtain the keys and access the data through the provider alone, without ever seeking authorization from the owner of the data. However unlikely this may seem, this exposure point cannot be ignored when storing classified or confidential data.

To summarize, while the built in security of cloud storage can provide improved protection against unauthorized access, it falls short of providing the level of control necessary to address all possibilities of data breach.

ENSURING DATA SECURITY THROUGH GATEWAY ENCRYPTION

So far, we’ve looked at the built-in security in cloud storage. Better, certainly, than no security at all, but without the level of control required to lock down unauthorized access to data; enough exposure to give security-minded organizations a bad case of heartburn. Now it’s time to strengthen that baseline by increasing your level of control.

By encrypting your data prior to transmission, you add another layer of security – one that you control and cannot be compromised by the cloud provider. One of the easiest ways to do this is through cloud-integrated storage products like CloudArray from EMC.

---

1 This paper is intended to be an overview of the fundamental aspects of securing cloud storage. It is not intended to be an exhaustive reference for cloud security, data security or a guarantee thereof. Keep in mind that securing cloud storage is a necessary aspect of data security when using the cloud, but represents only one element of a holistic corporate security strategy.

CloudArray provides a straightforward way to connect to cloud storage by converting APIs into common storage protocols such as SAN and NAS. In addition, CloudArray functions as a security gate for all the data you've stored in the cloud. That means that all data is encrypted at-rest, using strong AES-256 encryption, prior to transmission to the cloud provider.

Because the encryption occurs in the gateway, you as the customer have exclusive ownership of the keys, and those keys never have to be (nor should be) shared with the cloud provider. Local key management eliminates the single biggest exposure point of a typical cloud storage deployment. The data is rendered unintelligible without these keys, which means that even if someone can physically access data in the cloud (for example, the cloud provider, the cloud provider employees, affiliated parties, or even your own unauthorized employees who happen to have your cloud credentials), they cannot read the data.

To take it one step further, CloudArray encrypts the data and the metadata separately, using two different sets of encryption keys. As a result, the data remains inaccessible unless you have both sets of encryption keys. This double encryption provides customer data even greater protection from unauthorized access.

By storing the encryption keys with the owner of the data, rather than the storage provider, you close the door for third-party data access. Even in the case of a cloud provider being acquired (or their assets being made for sale), the new acquirer does not have the keys and therefore, does not have access to data. And even if the assets storing data are decommissioned or sold, deleting the keys makes the data effectively inaccessible, meeting enterprise requirements for data disposal. There is no need for additional data deletion, scrubbing or physical destruction.

**The most fundamental principle of storing corporate, confidential and sensitive data in the cloud is to ensure that the owner of the data is the sole arbiter of who has access to the encryption keys.**

What’s more, all the data encrypted by CloudArray can be encrypted again by the cloud provider using its own server-side encryption as outlined earlier. The result is that data stored in the cloud is often triply encrypted (both data and metadata by CloudArray, as well as the CSP’s encryption) further ensuring its security.

This essential data security framework is core to enabling compliance with regulatory provisions such as HIPAA, PCI and other regulations that require limiting data access to only authorized individuals. While there are often additional considerations to achieving regulatory compliance, it is nearly impossible without a framework that enables complete control of any party who has access to data.

**PROPER ENCRYPTION KEY MANAGEMENT**

Now that we’ve established the importance of pre-cloud encryption to ensure security, we’ve introduced an additional challenge: If the data owner manages the encryption keys, the onus is on the data owner to ensure that those keys are never lost. Once the keys are lost, the data is unrecoverable and can effectively be considered deleted forever. Should the keys be inadvertently shared with unauthorized parties, they may create new exposure points.

To address this challenge, data owners need to identify a way to back up encryption keys, but limit access to authorized parties – even within the four walls of an organization. A critical component is a key backup strategy that is able to address a site disaster scenario. If critical data resides encrypted in the cloud and a local site disaster loses or deletes those keys, data in the cloud becomes inaccessible.

This is another aspect of security where CloudArray can help. CloudArray customers can take advantage of an optional service that securely backs up keys offsite to the secure CloudArray.com portal. Offered as part of the CloudArray product, this key backup (if selected) provides a built-in disaster protection for the keys at all times. When there is a local site disaster, a user can download a new CloudArray virtual appliance or image a new physical appliance and perform a configuration restore that automatically repopulates the keys necessary for decrypting data stored in the cloud.

While backing up keys off-site to a trusted third party is a good way to ensure keys can be recovered in a disaster, you might naturally wonder if you’re introducing another exposure point into the data supply chain.

CloudArray extends security even further by password-encrypting the keys before they are sent for backup. Passcode encryption ensures that no third parties (EMC CloudArray included) ever have access to unencrypted keys, despite the key backup process. This provides the best of both worlds – peace of mind that you can access your encryption keys even if your primary site goes down, and security in knowing that only users authorized with the password will be able to use those keys to access your data.
With additional passcode encryption for the keys, you do, however, need to ensure that the passcode follows best practices for password creation, including having a mix of numerical and alphabetical characters that cannot easily be guessed or cracked. Unlike your personal passwords, however, there is no "Forgot my password" function – so it’s critical that you ensure the passcode is never lost. We recommend that multiple (authorized) personnel within your organization be provided with the passcode so that there is no single point of failure.

On that note, keep in mind that security is only as strong as the weakest link or endpoint in your organization, meaning compromise can be as simple as an obvious password or an open door. For this reason, cloud security should always be viewed as a necessary component of a holistic corporate security effort.

**THE ROLE OF ENCRYPTION IN DATA INTEGRITY**

In addition to ensuring that your data remains safe and secure from unauthorized access, CloudArray’s robust encryption plays a critical component in our data integrity infrastructure.

CloudArray segments its cache into cache pages. As part of the data integrity process, the encryption algorithm generates, assigns and inserts a unique hash to each cache page. This hash remains with the cache page until that page is retrieved for access by the requesting initiator or share application.

When the page is decrypted, the hash must match the value generated by the decryption algorithm. In the unlikely event the hash does not match, then the page is declared corrupt, and the data cannot be returned to the requesting initiator or application. This process prevents data corruption from propagating to the end users of the data.

**THE CHALLENGES OF MASS DATA IMPORT**

Having established the best way to secure data in the cloud, let’s examine an often forgotten aspect of cloud storage security: the data import process. When moving data from on-premise storage to cloud storage, many organizations require a large initial upload (or seeding) of data into the cloud. A little math can compute how long it takes to upload a large amount of data across a WAN. For instance, an uplink speed of 100Mbit/sec should be able to push nearly 1TB per day. However, that theoretical figure does not account for other users sharing a WAN link, hops/latencies or other overhead that can slow down network throughput. Even at a theoretical maximum of 1TB/day, 100TB of data may take over 3 months to upload – a relatively long and cumbersome process.

If your initial data upload is more than your network bandwidth can handle, then you have the option of shipping your initial data set to the cloud via a cloud provider import service (as offered by Google and Amazon Web Services). Disks containing data go directly to the cloud providers, who can load the data in one of their data centers or at a high-bandwidth access point with the benefit of zero impact to on-premise networks and WANs. For large data sets, this can represent the difference between weeks and months versus days to get data into the cloud.

As we established earlier, encrypting data at-rest before moving it to the cloud is critical to a strong security strategy. This holds true in a bulk import situation where the data is physically moved by unaffiliated third parties (FedEx or UPS, for example) who should not be able to access the data.

With that in mind, **security must be considered an essential aspect of the import process**. Transporting unencrypted data can easily become the weak link in an otherwise tightly secured cloud environment. Major data breaches have occurred because unencrypted tapes have been stolen or lost in transit – a disk-based data import process provides the same kind of threat.

Ideally, a data import process should encrypt and encapsulate data into the same object format stored in the cloud prior to shipping the data from an on-premise environment. Following a consistent security practice used for storing data online in the cloud eliminates having to make any security compromises during the import process.

With CloudArray’s ability to use NFS as a private cloud target, customers can use CloudArray to encrypt and encapsulate data into objects locally in NAS or local storage (as a local cloud), which would then be transported and subsequently loaded into a cloud provider object storage. This offers a unique, secure import service that strictly follows the best practices outlined above, without opening a window for potential unauthorized access. Even if disk media is lost or stolen during transit, the data cannot be accessed without the encryption keys that are stored safely on-premise.
WHAT ABOUT HACKING AND GOVERNMENT SPYING?

Time to tackle the elephant in the room. There has been and continues to be significant press around the NSA’s PRISM surveillance program causing many organizations to wonder how that should affect their business’s decisions to store data in a public cloud. Even if you are not worried about government agencies, rumors that agencies can potentially access your data leads to the obvious concern of malicious third parties not being far behind. So do the measures above adequately protect your data?

The short answer is yes. For starters, keep in mind that much of the data stored in the cloud, especially by consumer-grade products, is not encrypted at all. If your data is stored with strong encryption at-rest, like AES-256, you are well ahead of the game. In spite of unconfirmed allegations that virtually any security can be cracked, the reality remains that it takes billions of years to unlock a single AES-256 encrypted cloud storage object using today’s technology.

It is of great relevance that the NSA itself recommends AES-256 encryption to protect its own top-secret data\(^3\). For many, this may be the strongest reassurance that this type of encryption is more than suitable for protecting corporate data.

By encrypting your data prior to transport and managing the keys locally, you eliminate access to your data aside from the authorized individuals to whom you have given key access. Of course, this does not eliminate the possibility that government or local authorities may request access to your data or alter your legal liabilities. However, it does make the process identical to a request to access data stored within the four walls of your organization.

SUMMARY OF BEST PRACTICES

The tables on the next page summarize the best practices discussed for encrypting corporate and sensitive data stored in the cloud, both by feature and by benefit. They also illustrate how CloudArray addresses security concerns found with point solutions, consumer-grade solutions or solutions that are cobbled together using direct API access to the cloud.

---

### Table 1

<table>
<thead>
<tr>
<th>Access method</th>
<th>Security Features</th>
<th>In-flight security (SSL)</th>
<th>In-cloud encryption</th>
<th>On-premise encryption</th>
<th>Local key management</th>
<th>Key backup</th>
<th>Encrypted key backup</th>
<th>Secure data import</th>
</tr>
</thead>
<tbody>
<tr>
<td>API</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>API cloud encryption</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gateway</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gateway key backup</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMC CloudArray</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Access method</th>
<th>Security Benefits</th>
<th>Granular per-user access controls</th>
<th>No data access via network</th>
<th>No data access in-cloud</th>
<th>No data access by provider</th>
<th>No access to keys by 3rd party</th>
<th>Ability to recover lost keys</th>
<th>No data access during import</th>
</tr>
</thead>
<tbody>
<tr>
<td>API</td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>API cloud encryption</td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Gateway</td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Gateway key backup</td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>EMC CloudArray</td>
<td></td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
</tbody>
</table>

---

\(^3\) [http://www.nsa.gov/ia/programs/suiteb_cryptography/]
THE BOTTOM LINE

Securing your data in the cloud requires a multi-dimensional approach both to seal out unauthorized access and protect against the unexpected. The right approach can make data storage in the cloud as secure, if not more secure, than traditional on-premise data.

CloudArray leverages years of data storage expertise, effectively providing nearly all of the heavy lifting to make cloud storage as secure as the tightest local storage solution – whether it involves meeting regulatory requirements, data disposal requirements or just the added peace of mind your business needs when storing sensitive data. CloudArray can serve as the foundation of a security-centric cloud storage deployment that enables your business to leverage the many economic benefits of the cloud.

Organizations face a number of security risks on a daily basis. Lax security policies threaten data loss and misuse in any storage scenario – not just cloud storage. By building and enforcing strong policies, locking down access to sensitive data within your organization and coupling your storage strategy with EMC CloudArray, you can preserve your security and peace of mind while taking advantage of the cost- and resource-efficiencies of the cloud.