Overview of AWS Security - Storage Services

June 2016

(Please consult **http://aws.amazon.com/security/** for the latest version of this paper)

Notices

This document is provided for informational purposes only. It represents AWS' current product offerings and practices as of the date of issue of this document, which are subject to change without notice. Customers are responsible for making their own independent assessment of the information in this document and any use of AWS' products or services, each of which is provided "as is" without warranty of any kind, whether express or implied. This document does not create any warranties, representations, contractual commitments, conditions or assurances from AWS, its affiliates, suppliers or licensors. The responsibilities and liabilities of AWS to its customers are controlled by AWS agreements, and this document is not part of, nor does it modify, any agreement between AWS and its customers.

Storage Services

Amazon Web Services provides low-cost data storage with high durability and availability. AWS offers storage choices for backup, archiving, and disaster recovery, as well as block and object storage.

Amazon Simple Storage Service (Amazon S3) Security

Amazon Simple Storage Service (S3) allows you to upload and retrieve data at any time, from anywhere on the web. Amazon S3 stores data as objects within buckets. An object can be any kind of file: a text file, a photo, a video, etc. When you add a file to Amazon S3, you have the option of including metadata with the file and setting permissions to control access to the file. For each bucket, you can control access to the bucket (who can create, delete, and list objects in the bucket), view access logs for the bucket and its objects, and choose the geographical region where Amazon S3 will store the bucket and its contents.

Data Access

Access to data stored in Amazon S3 is restricted by default; only bucket and object owners have access to the Amazon S3 resources they create (note that a bucket/object owner is the AWS Account owner, not the user who created the bucket/object). There are multiple ways to control access to buckets and objects:

- **Identity and Access Management (IAM) Policies**. AWS IAM enables organizations with many employees to create and manage multiple users under a single AWS Account. IAM policies are attached to the users, enabling centralized control of permissions for users under your AWS Account to access buckets or objects. With IAM policies, you can only grant users within your own AWS account permission to access your Amazon S3 resources.
- **Access Control Lists (ACLs)**. Within Amazon S3, you can use ACLs to give read or write access on buckets or objects to groups of users. With ACLs, you can only grant other AWS accounts (not specific users) access to your Amazon S3 resources.
- **Bucket Policies**. Bucket policies in Amazon S3 can be used to add or deny permissions across some or all of the objects within a single bucket. Policies can be attached to users, groups, or Amazon S3 buckets, enabling centralized management of permissions. With bucket policies, you can grant users within your AWS Account or other AWS Accounts access to your Amazon S3 resources.

Type of Access Control	AWS Account-Level Control?	User-LevelControl?
IAM Policies	No	Yes
ACLs	Yes	No
Bucket Policies	Yes	Yes

You can further restrict access to specific resources based on certain conditions. For

example, you can restrict access based on request time (Date Condition), whether the request was sent using SSL (Boolean Conditions), a requester's IP address (IP Address Condition), or based on the requester's client application (String Conditions). To identify these conditions, you use policy keys. For more information about action-specific policy keys available within Amazon S3, refer to the Amazon Simple Storage Service Developer Guide.

Amazon S3 also gives developers the option to use query string authentication, which allows them to share Amazon S3 objects through URLs that are valid for a predefined period of time. Query string authentication is useful for giving HTTP or browser access to resources that would normally require authentication. The signature in the query string secures the request.

Data Transfer

For maximum security, you can securely upload/download data to Amazon S3 via the SSL encrypted endpoints. The encrypted endpoints are accessible from both the Internet and from within Amazon EC2, so that data is transferred securely both within AWS and to and from sources outside of AWS.

Data Storage

Amazon S3 provides multiple options for protecting data at rest. For customers who prefer to manage their own encryption, they can use a client encryption library like the Amazon S3 Encryption Client to encrypt data before uploading to Amazon S3. Alternatively, you can use Amazon S3 Server Side Encryption (SSE) if you prefer to have Amazon S3 manage the encryption process for you. Data is encrypted with a key generated by AWS or with a key you supply, depending on your requirements. With Amazon S3 SSE, you can encrypt data on upload simply by adding an additional request header when writing the object. Decryption happens automatically when data is retrieved.

Note that metadata, which you can include with your object, is not encrypted. Therefore, AWS recommends that customers not place sensitive information in Amazon S3 metadata.

Amazon S3 SSE uses one of the strongest block ciphers available – 256-bit Advanced Encryption Standard (AES-256). With Amazon S3 SSE, every protected object is encrypted with a unique encryption key. This object key itself is then encrypted with a regularly rotated master key. Amazon S3 SSE provides additional security by storing the encrypted data and encryption keys in different hosts. Amazon S3 SSE also makes it possible for you to enforce encryption requirements. For example, you can create and apply bucket policies that require that only encrypted data can be uploaded to your buckets.

For long-term storage, you can automatically archive the contents of your Amazon S3 buckets to AWS' archival service called Amazon Glacier. You can have data transferred at specific intervals to Glacier by creating lifecycle rules in Amazon S3 that describe which objects you want to be archived to Glacier and when. As part of your data management strategy, you can also specify how long Amazon S3 should wait after the objects are put into Amazon S3 to delete them.

When an object is deleted from Amazon S3, removal of the mapping from the public name

to the object starts immediately, and is generally processed across the distributed system within several seconds. Once the mapping is removed, there is no remote access to the deleted object. The underlying storage area is then reclaimed for use by the system.

Data Durability and Reliability

Amazon S3 is designed to provide 99.999999999% durability and 99.99% availability of objects over a given year. Objects are redundantly stored on multiple devices across multiple facilities in an Amazon S3 region. To help provide durability, Amazon S3 PUT and COPY operations synchronously store customer data across multiple facilities before returning SUCCESS. Once stored, Amazon S3 helps maintain the durability of the objects by quickly detecting and repairing any lost redundancy. Amazon S3 also regularly verifies the integrity of data stored using checksums. If corruption is detected, it is repaired using redundant data. In addition, Amazon S3 calculates checksums on all network traffic to detect corruption of data packets when storing or retrieving data.

Amazon S3 provides further protection via Versioning. You can use Versioning to preserve, retrieve, and restore every version of every object stored in an Amazon S3 bucket. With Versioning, you can easily recover from both unintended user actions and application failures. By default, requests will retrieve the most recently written version. Older versions of an object can be retrieved by specifying a version in the request. You can further protect versions using Amazon S3 Versioning's MFA Delete feature. Once enabled for an Amazon S3 bucket, each version deletion request must include the six-digit code and serial number from your multi-factor authentication device.

Access Logs

An Amazon S3 bucket can be configured to log access to the bucket and objects within it. The access log contains details about each access request including request type, the requested resource, the requestor's IP, and the time and date of the request. When logging is enabled for a bucket, log records are periodically aggregated into log files and delivered to the specified Amazon S3 bucket.

Cross-Origin Resource Sharing (CORS)

AWS customers who use Amazon S3 to host static web pages or store objects used by other web pages can load content securely by configuring an Amazon S3 bucket to explicitly enable cross-origin requests. Modern browsers use the Same Origin policy to block JavaScript or HTML5 from allowing requests to load content from another site or domain as a way to help ensure that malicious content is not loaded from a less reputable source (such as during cross-site scripting attacks). With the Cross-Origin Resource Sharing (CORS) policy enabled, assets such as web fonts and images stored in an Amazon S3 bucket can be safely referenced by external web pages, style sheets, and HTML5 applications.

Amazon Glacier Security

Like Amazon S3, the Amazon Glacier service provides low-cost, secure, and durable storage. But where Amazon S3 is designed for rapid retrieval, Amazon Glacier is meant to be used as an archival service for data that is not accessed often and for which retrieval times of several hours are suitable.

Amazon Glacier stores files as archives within vaults. Archives can be any data such as a photo, video, or document, and can contain one or several files. You can store an unlimited number of archives in a single vault and can create up to 1,000 vaults per region. Each archive can contain up to 40 TB of data.

Data Upload

To transfer data into Amazon Glacier vaults, you can upload an archive in a single upload operation or a multipart operation. In a single upload operation, you can upload archives up to 4 GB in size. However, customers can achieve better results using the Multipart Upload API to upload archives greater than 100 MB. Using the Multipart Upload API allows you to upload large archives, up to about 40 TB. The Multipart Upload API call is designed to improve the upload experience for larger archives; it enables the parts to be uploaded independently, in any order, and in parallel. If a multipart upload fails, you only need to upload the failed part again and not the entire archive.

When you upload data to Amazon Glacier, you must compute and supply a tree hash. Amazon Glacier checks the hash against the data to help ensure that it has not been altered en route. A tree hash is generated by computing a hash for each megabyte-sized segment of the data, and then combining the hashes in tree fashion to represent ever-growing adjacent segments of the data.

As an alternate to using the Multipart Upload feature, customers with very large uploads to Amazon Glacier may consider using the AWS Import/Export service instead to transfer the data. AWS Import/Export facilitates moving large amounts of data into AWS using portable storage devices for transport. AWS transfers your data directly off of storage devices using Amazon's high-speed internal network, bypassing the Internet.

You can also set up Amazon S3 to transfer data at specific intervals to Amazon Glacier. You can create lifecycle rules in Amazon S3 that describe which objects you want to be archived to Amazon Glacier and when. You can also specify how long Amazon S3 should wait after the objects are put into Amazon S3 to delete them.

To achieve even greater security, you can securely upload/download data to Amazon Glacier via the SSL-encrypted endpoints. The encrypted endpoints are accessible from both the Internet and from within Amazon EC2, so that data is transferred securely both within AWS and to and from sources outside of AWS.

Data Retrieval

Retrieving archives from Amazon Glacier requires the initiation of a retrieval job, which is generally completed in 3 to 5 hours. You can then access the data via HTTP GET requests. The data will remain available to you for 24 hours.

You can retrieve an entire archive or several files from an archive. If you want to retrieve only a subset of an archive, you can use one retrieval request to specify the range of the archive that contains the files you are interested or you can initiate multiple retrieval requests, each with a range for one or more files. You can also limit the number of vault inventory items retrieved by filtering on an archive creation date range or by setting a maximum items limit. Whichever method you choose, when you retrieve portions of your

archive, you can use the supplied checksum to help ensure the integrity of the files provided that the range that is retrieved is aligned with the tree hash of the overall archive.

Data Storage

Amazon Glacier automatically encrypts the data using AES-256 and stores it durably in an immutable form. Amazon Glacier is designed to provide average annual durability of 99.999999999% for an archive. It stores each archive in multiple facilities and multiple devices. Unlike traditional systems which can require laborious data verification and manual repair, Amazon Glacier performs regular, systematic data integrity checks and is built to be automatically self-healing.

Data Access

Only your account can access your data in Amazon Glacier. To control access to your data in Amazon Glacier, you can use AWS IAM to specify which users within your account have rights to operations on a given vault.

AWS Storage Gateway Security

The AWS Storage Gateway service connects your on-premises software appliance with cloud-based storage to provide seamless and secure integration between your IT environment and AWS' storage infrastructure. The service enables you to securely upload data to AWS' scalable, reliable, and secure Amazon S3 storage service for cost-effective backup and rapid disaster recovery.

AWS Storage Gateway transparently backs up data off-site to Amazon S3 in the form of Amazon EBS snapshots. Amazon S3 redundantly stores these snapshots on multiple devices across multiple facilities, detecting and repairing any lost redundancy. The Amazon EBS snapshot provides a point-in-time backup that can be restored on-premises or used to instantiate new Amazon EBS volumes. Data is stored within a single region that you specify.

AWS Storage Gateway offers three options:

- Gateway-Stored Volumes (where the cloud is backup). In this option, your volume data is stored locally and then pushed to Amazon S3, where it is stored in redundant, encrypted form, and made available in the form of Elastic Block Storage (EBS) snapshots. When you use this model, the on-premises storage is primary, delivering low-latency access to your entire dataset, and the cloud storage is the backup.
- Gateway-Cached Volumes (where the cloud is primary). In this option, your volume data is stored encrypted in Amazon S3, visible within your enterprise's network via an iSCSI interface. Recently accessed data is cached on- premises for low-latency local access. When you use this model, the cloud storage is primary, but you get low- latency access to your active working set in the cached volumes on premises.
- Gateway-Virtual Tape Library (VTL). In this option, you can configure a Gateway-VTL with up to 10 virtual tape drives per gateway, 1 media changer and up to 1500 virtual tape cartridges. Each virtual tape drive responds to the SCSI command set, so your existing on-premises backup applications (either disk-to-tape or disk-to-disk-to- tape) will work without modification.

No matter which option you choose, data is asynchronously transferred from your on-premises

storage hardware to AWS over SSL. The data is stored encrypted in Amazon S3 using Advanced Encryption Standard (AES) 256, a symmetric- key encryption standard using 256-bit encryption keys. The AWS Storage Gateway only uploads data that has changed, minimizing the amount of data sent over the Internet.

The AWS Storage Gateway runs as a virtual machine (VM) that you deploy on a host in your data center running VMware ESXi Hypervisor v 4.1 or v 5 or Microsoft Hyper-V (you download the VMware software during the setup process). You can also run within EC2 using a gateway AMI. During the installation and configuration process, you can create up to 12 stored volumes, 20 Cached volumes, or 1500 virtual tape cartridges per gateway. Once installed, each gateway will automatically download, install, and deploy updates and patches. This activity takes place during a maintenance window that you can set on a per-gateway basis.

The iSCSI protocol supports authentication between targets and initiators via CHAP (Challenge-Handshake Authentication Protocol). CHAP provides protection against man-in-the-middle and playback attacks by periodically verifying the identity of an iSCSI initiator as authenticated to access a storage volume target. To set up CHAP, you must configure it in both the AWS Storage Gateway console and in the iSCSI initiator software you use to connect to the target.

After you deploy the AWS Storage Gateway VM, you must activate the gateway using the AWS Storage Gateway console. The activation process associates your gateway with your AWS Account. Once you establish this connection, you can manage almost all aspects of your gateway from the console. In the activation process, you specify the IP address of your gateway, name your gateway, identify the AWS region in which you want your snapshot backups stored, and specify the gateway time zone.

AWS Import/Export Security

AWS Import/Export is a simple, secure method for physically transferring large amounts of data to Amazon S3, EBS, or Amazon Glacier storage. This service is typically used by customers who have over 100 GB of data and/or slow connection speeds that would result in very slow transfer rates over the Internet. With AWS Import/Export, you prepare a portable storage device that you ship to a secure AWS facility. AWS transfers the data directly off of the storage device using Amazon's high-speed internal network, thus bypassing the Internet. Conversely, data can also be exported from AWS to a portable storage device.

Like all other AWS services, the AWS Import/Export service requires that you securely identify and authenticate your storage device. In this case, you will submit a job request to AWS that includes your Amazon S3 bucket, Amazon EBS region, AWS Access Key ID, and return shipping address. You then receive a unique identifier for the job, a digital signature for authenticating your device, and an AWS address to ship the storage device to. For Amazon S3, you place the signature file on the root directory of your device. For Amazon EBS, you tape the signature barcode to the exterior of the device. The signature file is used only for authentication and is not uploaded to Amazon S3 or EBS.

For transfers to Amazon S3, you specify the specific buckets to which the data should be loaded and ensure that the account doing the loading has write permission for the buckets. You should also specify the access control list to be applied to each object loaded to Amazon S3.

For transfers to EBS, you specify the target region for the EBS import operation. If the storage device is less than or equal to the maximum volume size of 1 TB, its contents are loaded directly into an Amazon EBS snapshot. If the storage device's capacity exceeds 1 TB, a device image is

stored within the specified S3 log bucket. You can then create a RAID of Amazon EBS volumes using software such as Logical Volume Manager, and copy the image from S3 to this new volume.

For added protection, you can encrypt the data on your device before you ship it to AWS. For Amazon S3 data, you can use a PIN-code device with hardware encryption or TrueCrypt software to encrypt your data before sending it to AWS. For EBS and Amazon Glacier data, you can use any encryption method you choose, including a PIN-code device. AWS will decrypt your Amazon S3 data before importing using the PIN code and/or TrueCrypt password you supply in your import manifest. AWS uses your PIN to access a PIN-code device, but does not decrypt software-encrypted data for import to Amazon EBS or Amazon Glacier.

AWS Import/Export Snowball uses appliances designed for security and the Snowball client to accelerate petabyte-scale data transfers into and out of AWS. You start by using the AWS Management Console to create one or more jobs to request one or multiple Snowball appliances (depending on how much data you need to transfer), and download and install the Snowball client. Once the appliance arrives, connect it to your local network, set the IP address either manually or with DHCP, and use the client to identify the directories you want to copy. The client will automatically encrypt and copy the data to the appliance and notify you when the transfer job is complete.

After the import is complete, AWS Import/Export will erase the contents of your storage device to safeguard the data during return shipment. AWS overwrites all writable blocks on the storage device with zeroes. If AWS is unable to erase the data on the device, it will be scheduled for destruction and our support team will contact you using the email address specified in the manifest file you ship with the device.

When shipping a device internationally, the customs option and certain required subfields are required in the manifest file sent to AWS. AWS Import/Export uses these values to validate the inbound shipment and prepare the outbound customs paperwork. Two of these options are whether the data on the device is encrypted or not and the encryption software's classification. When shipping encrypted data to or from the United States, the encryption software must be classified as 5D992 under the United States Export Administration Regulations.

Further Reading

https://aws.amazon.com/security/security-resources/

Introduction to AWS Security Processes
Overview of AWS Security - Storage Services
Overview of AWS Security - Database Services
Overview of AWS Security - Compute Services
Overview of AWS Security - Application Services
Overview of AWS Security - Analytics, Mobile and Application Services
Overview of AWS Security – Network Services