Securing Amazon Web Services (AWS) and Simple Storage Service (Amazon S3) Security
David Boland

The Cloud

Whether a company is using Facebook for marketing purposes or migrating their entire organization's productivity suite to Google Apps, most - if not all - enterprises are operating in the cloud to some degree whether they know it or not. Given the cost savings and collaboration benefits of cloud productivity offerings, it's easy to see why companies are making the switch.

Even the United States federal government is using the cloud. It has moved recovery.gov, the Web site people can use to track spending under 2009’s $787 billion economic stimulus package, to Amazon's Elastic Compute Cloud infrastructure-as-a-service platform.

"As the world's largest consumer of information technology and as stewards of taxpayer dollars, the U.S. federal government has a duty to be a leader in pioneering the use of new technologies that are more efficient and economical," Kundra said in a blog post aimed squarely at federal agencies. "By using cloud services, the federal government will gain access to powerful technology resources faster and at lower costs. This frees us to focus on mission-critical tasks instead of purchasing, configuring, and maintaining redundant infrastructure."[1]

Cloud Computing is becoming next stage platform in the evolution of the internet. It provides customers an enhanced and efficient way to store data in the cloud with different ranges of capabilities and applications.

The one downside to this migration is that organizations are trusting cloud platforms to sensitive information without a full understanding of the data security implications.
Security has been and remains one of the main concerns holding federal agencies back from considering cloud computing. Before moving Recovery.gov to Amazon EC2, the US officials sought and received assurances from Amazon that none of the Recovery.gov data would be stored in foreign countries, and went through the certification and accreditation required to be compliant with the Federal Information Security Management Act, which regulates federal cyber security. The press release provided by the U.S. officials says that by running the site on EC2, recovery.gov's security has actually improved by adding "greater protection against network attacks and real-time detection of system tampering."[2]

While NASA and other agencies have been testing EC2, Vivek Kundra the 1st Chief Information Officer of the United States said that Recovery.gov is the first production system running on Amazon Web Services.

There are multiple cloud providers, both commercial and open source, including Amazon Web Services (AWS), Rackspace, GoGrid, Nimbus, and Eucalyptus.

**Amazon Web Services**

In 2006, Amazon Web Services (AWS) began offering IT infrastructure services to businesses in the form of web services - now commonly known as cloud computing. One of the key benefits of cloud computing is the opportunity to replace up-front capital infrastructure expenses with low variable costs that scale with a business. With the Cloud, businesses no longer need to plan for and procure servers and other IT infrastructure weeks or months in advance. Instead, they can instantly spin up hundreds or thousands of servers in minutes and deliver results faster for hundreds of thousands of businesses in 190 countries around the world.

Amazon Web Services (AWS) is a collection of remote computing services that together make up a cloud computing platform - offered over the Internet by Amazon.com. The most central and well-known of these services are Amazon S3 and Amazon EC2. AWS provides the necessary computing environment, including CPUs, storage, memory (RAM), networking, and operating system, and is an example of “infrastructure as a service” (IaaS).
AWS (Amazon Web Services) Security

AWS is a secure, durable technology platform with industry-recognized certifications and audits: PCI DSS Level 1, ISO 27001, FISMA Moderate, HIPAA, and SAS 70 Type II. Their services and data centers have multiple layers of operational and physical security to ensure the integrity and safety of every company’s data.

This paper will describe AWS’s approach to cloud computing privacy and security using several methods including: Identity and Access Management (IAM), encryption and decryption services, firewalls, Access Control Lists (ACL) and bucket policies, versioning and redundancy.

Encryption

Use Public/Private Key Pair

In this cryptographic scheme, a pair of keys is constructed, a public encryption key and a private decryption key. The public key is made available to anyone, but the data encrypted by it can only be decrypted by its paired private key. The private key should never be shared because it represents digital proof of the user's identity. Public/private keys provide a more secure login authentication method than usernames and passwords. To grant someone access to an instance, the public key must be copied into the list of authorized keys.

For security of data and encryption of data transfers, it is imperative that access to an instance is made via a secure protocol such as Secure Shell (ssh) or Secure Copy (scp). Both the AWS Web console and command line tools provide a simple interface to generate key pairs when launching an instance. The public key is automatically installed onto an instance, and the private key can then be used on a computer that will ssh into that instance.

Create Restricted User Accounts

Recently, AWS introduced Identity and Access Management (IAM) to offer greater control and management of multiple users. Each user has their own set of security credentials to access cloud resources, eliminating the need to share login information and keys for the master AWS account owner. This is important because the master account contains the personal billing information, which, for obvious reasons, should not be accessible to all users. IAM can restrict services based
on specific users or group policies. For example, it is possible to restrict a user to a specific S3 bucket between 9 A.M. and 5 P.M. from a specific IP address.

**Control Access Using Firewalls**
A security group defines a set of rules that govern how traffic (data or communication) reaches the AWS instance. By default, the security group restricts all inbound traffic, allows all outbound traffic, and allows other instances within the group to communicate. These rules can be completely customized; for example, it is possible to restrict access to a specific IP address on a specific port address and not allow that instance to communicate with other instances within the AWS account.

**Additional Security**
All security keys should be replaced with new ones every 30–90 days. Installing regular software updates is essential to protect the operating system and third-party software from vulnerabilities. There are many additional security features such as private clouds (http://aws.amazon.com/vpc/), encrypted file systems, and encrypted data volumes that may be used by those who have security needs beyond these basic best practices.

**Amazon S3**
Amazon S3 (Simple Storage Service) is an online storage web service offered by Amazon Web Services. Amazon launched S3, its first publicly-available web service, in the United States in March 2006 and in Europe in November 2007.

At its inception, Amazon charged end users US$0.15 per gigabyte-month, with additional charges for bandwidth used in sending and receiving data, and a per-request (get or put) charge. As of November 1, 2008, pricing moved to tiers where end users storing more than 50 terabytes receive discounted pricing. Amazon claims that S3 uses the same scalable storage infrastructure that Amazon.com uses to run its own global e-commerce network.
Amazon S3 is reported to store more than 762 billion objects as of December 2011. S3 uses include web hosting, image hosting, and storage for backup systems. S3 comes with a 99.9% monthly uptime guarantee which equates to approximately 43 minutes of downtime per month.

Amazon S3 is storage for the Internet. It’s a simple storage service that offers software developers a highly-scalable, reliable, and low-latency data storage infrastructure at very low costs.

It provides a simple web services interface that you can use to store and retrieve any amount of data, at any time, from anywhere on the web. Using this web service, developers can easily build applications that make use of Internet storage. Since Amazon S3 is highly scalable and you only pay for what you use, developers can start small and grow their application as they wish, with no compromise on performance or reliability. It is designed to be highly flexible: Store any type and amount of data that you want; read the same piece of data a million times or only for emergency disaster recovery; build a simple FTP application, or a sophisticated web application such as the Amazon.com retail web site. Amazon S3 frees developers to focus on innovation, not figuring out how to store their data.

For example you can host your entire static website on Amazon S3. Customers commonly store images, video, and other content for their website on Amazon S3. You can also take advantage of root and index document support as well as custom HTML errors to host your entire website from your Amazon S3 bucket.

You can store virtually any kind of data in any format.

The total volume of data and number of objects you can store are unlimited. Individual Amazon S3 objects can range in size from 1 byte to 5 terabytes. The largest object that can be uploaded in a single PUT is 5 gigabytes. For objects larger than 100 megabytes, customers should consider using the Multipart Upload capability.
You can use the Object Expiration feature to remove objects from your buckets after a specified number of days. You can define the expiration rules for a set of objects in your bucket through the Lifecycle Configuration policy that you apply to the bucket. Each Object Expiration rule allows you to specify a prefix and an expiration period.

**Securing S3**

**Data Protection**

**Redundancy**

Figure 1: Objects are redundantly stored on multiple devices across multiple facilities.

**Versioning**

Versioning allows you to preserve, retrieve, and restore every version of every object in an Amazon S3 bucket. Once you enable Versioning for a bucket, Amazon S3 preserves existing objects anytime you perform a PUT, POST, COPY, or DELETE operation on them. By default, GET requests will retrieve the most recently written version. Older versions of an overwritten or deleted object can be retrieved by specifying a version in the request. Versioning offers an additional level of protection by providing a means of recovery when customers accidentally
overwrite or delete objects. This allows you to easily recover from unintended user actions and application failures.

Figure 2: You can use versioning to preserve, retrieve, and restore every version of every object stored in your Amazon S3 bucket.

Checksums

Figure 3: At the firewall Amazon S3 also regularly verifies the integrity of data stored using checksums.
**Data Encryption**

Data can be encrypted using either Server Side Encryption (SSE) or using a client library like the Amazon S3 Encryption Client. If you are using a client library, you retain control of keys used to encrypt. Some customers prefer this additional control of keys; others prefer not to incur the overhead required to manage and protect keys. If you are using SSE, AWS handles key management and key protection for you. You should choose SSE if you prefer AWS to manage your keys. SSE uses one of the strongest block ciphers available, 256-bit Advanced Encryption Standard (AES-256). 256-bit is the largest key size defined for AES. Both client side encryption and Server Side Encryption are supported for objects stored in Standard Storage and Reduced Redundancy Storage (RRS).

**Server-Side Encryption**

Server-side encryption is about data encryption at rest, that is, Amazon S3 encrypts your data as it writes it to disks in its data centers and decrypts it for you when you access it. As long as you authenticate your request and you have access permissions, there is no difference in the way you access encrypted or unencrypted objects. Amazon S3 manages encryption and decryption for you. For example, if you share your objects using a pre-signed URL, the pre-signed URL works the same way for both encrypted and unencrypted objects.

Amazon S3 Server Side Encryption employs strong multi-factor encryption. Amazon S3 encrypts each object with a unique key. As an additional safeguard, it encrypts the key itself with a master key that it regularly rotates. Amazon S3 Server Side Encryption uses one of the strongest block ciphers available, a 256-bit Advanced Encryption Standard (AES-256), to encrypt your data.

You can specify data encryption at the object level. Therefore your bucket might contain both encrypted and unencrypted objects.
Client-Side Encryption

Important - Your private encryption keys and your unencrypted data are never sent to AWS; therefore, it is important that you safely manage your encryption keys. If you lose your encryption keys, you won't be able to unencrypt your data.

AWS uses a process called envelope encryption. In envelope encryption, you provide your encryption key to the Amazon S3 encryption client and the client takes care of the rest of the process. The process works like this:

1) The Amazon S3 encryption client generates a one-time-use symmetric key that the client uses to encrypt your data.

2) The client encrypts the envelope symmetric key using your private encryption key.

3) The client then uploads the encrypted envelope key along with your encrypted data to Amazon S3.

Retrieving and decrypting the client-side encrypted data from Amazon S3 is the reverse of the encryption mentioned in the previous slide.

Your private encryption key that the client uses in the envelope encryption process can be an asymmetric key pair (composed of public and private keys) or it can be a symmetric key.

Amazon recommends using asymmetric keys for the increased level of security that they provide in the envelope encryption process.

Access Control

Identity and Access Management (IAM)

AWS Identity and Access Management (IAM) enables you to securely control access to AWS services and resources for your users. IAM enables you to create and manage users in AWS's identity management system (referred to as "IAM users"), and it also enables you to grant access
to AWS resources for users managed outside of AWS in your corporate directory (referred to as "federated users"). IAM offers greater security, flexibility, and control when using AWS.

IAM makes it easy to provide multiple users secure access to your AWS account and resources. IAM enables you to:

1) Manage IAM users and their access - You can create users in AWS's identity management system, assign users individual security credentials (i.e. Access Keys, password, Multi Factor Authentication devices) or request temporary security credentials to provide them access to AWS services and resources. You can manage permissions to control which operations a user can perform.

2) You can manage access for federated users - You can request security credentials with configurable expirations for users that you manage in your corporate directory, allowing you to provide your employees and applications secure access to resources in your AWS account, without creating them an IAM user. You specify the permissions for these security credentials, to control which operations a user can perform.

Figure 4: Notice in this figure only the Account Holder has access to the cloud.
IAM Policies

Permissions are managed via IAM policies. These policies are attached to the users, enabling centralized control of permissions for users under your AWS account.

Figure 5: With IAM Policies, you can only grant users within your own AWS Account permission to your Amazon S3 resources.

IAM supports the ability to:

1) Create IAM user identities - Add IAM users to your AWS account.
2) Organize IAM users in groups - Create groups to easily manage permissions for multiple IAM users under your AWS account.
3) Centralize control of user access - Control which operations your users can perform, such as accessing specific AWS service APIs and resources.
4) Conditional user access - Add conditions to control how your users can use AWS, such as time of day, their originating IP address, or whether they are using SSL.
5) Create and assign security credentials - Assign security credentials to your IAM users, and rotate and/or revoke these credentials as desired.
6) Create temporary security credentials - Request temporary security credentials with configurable expiration and permissions for your IAM users, federated users or applications.

IAM supports multiple methods to create, delete and list IAM users, manage group membership, manage user security credentials, and assign permissions. You can create and manage users, groups and permissions via IAM APIs, Command Line Tools, or via the IAM console which gives you a point-and-click, web-based interface. You can also use the AWS Policy Generator to create policies.

An IAM user can have any combination of credentials that AWS supports - AWS Access Key, X.509 certificate, password for web app logins, or Multi Factor Authentication (MFA) device. This allows users to interact with AWS in any way that makes sense for them - an employee might have both an AWS Access Key and a password; a software system might have only an AWS Access Key to make programmatic calls; and an outside contractor might have only an X.509 certificate to use the EC2 command line interface.

Resource-based Permissions on Objects and Buckets
Amazon S3 supports resource-based permissions on objects and buckets using access control list (ACL) and bucket policies. The ACLs and bucket policies are attached to buckets and objects defining which AWS accounts have access and the type of access.

With ACLs, you can only grant other AWS Accounts access to your Amazon S3 resources.
Figure 6: With ACLs, you can only grant other AWS Accounts access to your Amazon S3 resources.

Bucket Policies

With Bucket Policies, you can do both. You can grant users within your own AWS Account permission to your Amazon S3 resources and you can grant other AWS Accounts access to your Amazon S3 resources.

Figure 7: With Bucket Policies, you can do both.
Figure 8: Bucket Policies let you manage access based on specific resources i.e. buckets.

With any shared storage system, the biggest question is whether unauthorized users can access information either intentionally or by mistake. To ensure that customers have the utmost in flexibility to determine how, when, and to whom they wish to expose the information they store in AWS, Amazon S3 APIs provide both bucket- and object-level access controls, with defaults that only permit authenticated access by the bucket and/or object creator. Write and Delete permission is controlled by an Access Control List (ACL) associated with the bucket. Permission to modify the bucket ACLs is itself controlled by an ACL, and it defaults to creator-only access. Therefore, the customer maintains full control over who has access to their data. Amazon S3 access can be granted based on AWS Account ID, DevPay Product ID, or open to everyone.

An Amazon S3 account starts with a clean root account. On the root is where you create buckets. Bucket is Amazon S3’s terminology for root folder. You can create multiple buckets, and inside buckets is where you place your folders and images.
Amazon S3 releases a set of API and developers around the world that allows your Amazon S3 account to talk to your local computer so you can do all the file uploading, synchronization, back-up etc.
References

http://search.proquest.com.jproxy.lib.ecu.edu/docview/223222108

*Biomedical Cloud Computing With Amazon Web Services, Vincent A. Fusaro, Prasad Patil, Erik Gafni, Dennis P. Wall, and Peter J. Tonellato, Center for Biomedical Informatics, Harvard Medical School, Boston, Massachusetts, August 7, 2011
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3161908/


Amazon Web Services
http://docs.amazonwebservices.com/AmazonS3/latest/dev/UsingAuthAccess.html

Five Ways to Protect Company Information in the Cloud
http://cloudcomputing.sys-con.com/node/1851360

Cloud Security Is Looking Overcast, Aaron Ricadela, September 1, 2011
http://www.businessweek.com/magazine/cloud-security-is-looking-overcast-09012011.html

Being Smart about Cloud Security, Brian Bergstein, October 4, 2011
http://www.technologyreview.com/business/38714/

Is the cloud secure? marcy hoffman , June 27, 2011
http://smallbusinessblog.infostreet.com/2011/06/is-the-cloud-secure/