Creating HIPAA-Compliant Medical Data Applications with Amazon Web Services

April 2009

Executive Summary

In the U.S., certain organizations that transmit an individual's protected health information (PHI) across Internet applications or electronic systems are required to meet Health Insurance Portability and Accountability Act of 1996 (HIPAA) requirements. HIPAA is a set of established federal standards, implemented through a combination of administrative, physical and technical safeguards, intended to ensure the security and privacy of PHI. These standards affect the use and disclosure of PHI by certain covered entities (such as healthcare providers engaged in electronic transactions, health plans and healthcare clearinghouses) and their business associates.

Healthcare businesses subject to HIPAA can utilize the secure, scalable, low-cost, IT infrastructure provided by Amazon Web Services (AWS) as part of building HIPAA-compliant applications. Amazon Elastic Compute Cloud (Amazon EC2) provides resizable compute capacity in the cloud, and Amazon Simple Storage Service (Amazon S3) provides a virtually unlimited cloud-based data object store. With no minimum fees, no term-based contracts, and pay-as-you-use pricing, AWS is a reliable and effective solution for growing healthcare industry applications.

This paper briefly outlines how companies can use Amazon Web Services to power HIPAA-compliant information processing systems. We will focus on the HIPAA sections *The Privacy Rule* and *The Security Rule,* and how to encrypt and protect your data in the AWS cloud. For additional information on HIPAA, visit http://www.hhs.gov/ocr/hipaa.

What is HIPAA and Why is it Important?

HIPAA provides national minimum standards to protect an individual's health information. HIPAA was originally created to streamline healthcare processes and reduce costs, while ensuring individual consumer privacy. The U.S. Department of Health and Human Services (HHS) manages and enforces these standards.

HIPAA covers protected health information (PHI) which is any information regarding an individual’s physical or mental health, the provision of healthcare to them, or payment of related services. PHI also includes any personally identifiable information, including for example Employer Identification Number, social security number, name, address, phone number, medical condition when linked to a patient, and some types of billing information.

In order to be compliant, organizations must design their systems and applications to meet HIPAA’s privacy and security standards and related administrative, technical, and physical safeguards.

Privacy & Security Rules

HIPAA’s *Privacy Rule* requires that individuals’ health information is properly protected by covered entities. Among other requirements, the privacy rule prohibits entities from transmitting PHI over open networks or downloading it to public or remote computers without encryption.

The *Security Rule* requires covered entities to put in place detailed administrative, physical and technical safeguards to protect electronic PHI. To do this, covered entities are required to implement access controls, encrypt data, and set up back-up and audit controls for electronic PHI in a manner commensurate with the associated risk.

Privacy Controls: Encrypting Data in the Cloud

HIPAA’s *Privacy Rule* regulations include standards regarding the encryption of all PHI in transmission (“in-flight”) and in storage (“at-rest”). The same data encryption mechanisms used in a traditional computing environment, such as a local server or a managed hosting server, can also be used in a virtual computing environment, such as Amazon EC2 and Amazon S3.

Amazon EC2 provides the customer with full root access and administrative control over virtual servers. To ensure data security during electronic transmission, files containing PHI should be encrypted using technologies such as 256 bit AES algorithms. Furthermore, to reduce the risk of exposing PHI and to reduce bandwidth usage, any data not required by applications running in the cloud, including PHI, should be removed prior to transmission.

Using AWS, customer's system administrators can utilize token or key-based authentication to access their virtual servers. Amazon EC2 creates a 2048 bit RSA key pair, with private and public keys and a unique identifier for each key pair to help facilitate secure access. Administrators can also utilize a command-line shell interface, Secure Shell (SSH) keys, or sudo to enable additional security and privilege escalation.

A complete firewall solution can be created in the cloud by utilizing Amazon EC2's default *deny-all* mode which automatically denies all inbound traffic unless the customer explicitly opens an EC2 port. Administrators can create multiple security groups in order to enforce different ingress policies as needed. They can control each security group with a PEM-encoded X.509 certificate and restrict traffic to each EC2 instance by protocol, service port, or source IP address. For more information on encryption and firewalls, see the AWS Security Whitepaper.

Similar to Amazon EC2, when sending data to Amazon S3 for either short term or long term storage, we highly recommend encrypting data before transmission. We also recommend against putting any PHI or other sensitive data, including keys, in Amazon S3 metadata. Amazon S3 can be accessed via Secure Socket Layer (SSL)-encrypted endpoints over the Internet and from within Amazon EC2. Following these practices ensures that PHI and other sensitive data remain highly secure.

Security Controls: High-Level Data Protection

While data flowing to and from the AWS cloud should be safeguarded with encryption, data that comes in contact with administrators or third-party partners may require different control mechanisms. To help you comply with HIPAA's *Security Rule*, this section discusses AWS security policies and processes regarding data and how you can implement authentication, access consent processes, and audit controls to reduce the risk of outside compromise. These controls, among others, ensure that you can restrict access to your system, can carefully and constantly monitor it, and can quickly lock it down in case of threat or attack.

AWS Security Policies

For Amazon EC2, AWS employees do not look at customer data, do not have access to customer EC2 instances, and cannot log into the guest operating system. AWS internal security controls limit data access. Only in rare cases of customer-requested support or maintenance, can select AWS employees use their individual, cryptographically-strong SSH keys to gain access to the *host* (as opposed to the *guest*) operating system. This access is only available from the Amazon network and requires two-factor authentication. Keys are revoked when access is no longer relevant.

For Amazon S3, AWS employees' access to customer data is highly restricted and not necessary for customer support or maintenance. Despite these internal AWS controls, we strongly suggest that customers encrypt all sensitive data.

Amazon.com's Information Security Policies, followed by AWS, are guided by the fundamental principle of *least privilege*. Least privilege protects customer information assets by requiring that no individual, program or system is granted more access privileges than are necessary to perform the task. Any employee found to have violated this policy may be subject to disciplinary action, including termination.

Access Control Processes

The customer's system administrator should set user and computer access controls to restrict data access and ensure security. AWS provides a number of mechanisms to control access to data while in-flight and at-rest in the AWS cloud. Using Amazon EC2, SSH network protocols can be used to authenticate remote users or computers through public-key cryptography. Public-key cryptography or key pairs are used to ensure confidentiality by issuing a private key for decryption and a public key for encryption. The administrator can also allow or block access at the account or instance level and can set security groups, which restrict network access from instances not residing in that same group.

Using Amazon S3, access can be easily controlled down to the object level. The system administrator maintains full control over who has access to the data at all times and the default setting only permits authenticated access to the creator. Read, write and delete permissions are controlled by an Access Control List (ACL) associated with each object.

For both Amazon S3 and EC2, each account has a secret key that is crucial for maintaining the security of customer accounts. We recommend keeping your keys and your account credentials in a secure location. As such, do not embed your secret key in a web page or other publicly accessible source code and do not transmit it over insecure channels. You should use Secure HTTP (HTTPS) connections for web applications running in the cloud to ensure any PHI presented in the interface is protected as it travels from AWS to the users' browsers.

Auditing, Back-Ups, & Disaster Recovery

HIPAA's security safeguards also require in-depth auditing capabilities, data back-up procedures and disaster recovery mechanisms. AWS services contain many features that help customers address these requirements.

In designing a HIPAA-compliant system, customers should put auditing capabilities in place to allow security analysts to drill down into detailed activity logs or reports to see who had access, IP address entry, what data was accessed, etc. This data should be tracked, logged, and stored in a central location for extended periods of time in case of an audit. Using Amazon EC2, customers can run activity log files and audits down to the packet layer on their virtual servers, just as on traditional hardware. They can also track any IP traffic that reaches their virtual server instance. Customer's administrators can back up the log files into Amazon S3 for long-term, reliable storage.

HIPAA's safeguards also require compliant companies to create and implement a data backup plan. Under HIPAA, covered entities must have a contingency plan to protect data in case of an emergency, and must create and maintain retrievable exact copies of electronic protected health information. To implement a data back-up plan on AWS, Amazon Elastic Block Store

(EBS) offers persistent storage for Amazon EC2 virtual server instances. These volumes can be exposed as standard block devices and offer off-instance storage that persists independently from the life of an instance. To adhere to HIPAA guidelines, customers can create point-in-time snapshots of EBS volumes, which are automatically stored in Amazon S3 and are replicated across multiple Availability Zones. These snapshots can be accessed at any time and can protect data for long-term durability. Amazon S3, also, provides a highly available solution for data storage and automated back-ups. By simply loading a file or image into Amazon S3, multiple redundant copies are automatically created and stored in separate data centers. These files can be accessed at any time, from anywhere (based on permissions) and are stored until intentionally deleted by the customer's system administrator.

Disaster recovery, the process of protecting an organization's data and IT infrastructure in times of disaster, is typically one of the more expensive HIPAA regulations to comply with. It involves maintaining highly available systems, keeping both the data and system replicated off-site, and enabling continuous access to both. AWS inherently offers a variety of disaster recovery mechanisms. With Amazon EC2, administrators can start server instances very quickly and can use an Elastic IP address (a static IP for the cloud computing environment) for elegant failure from one machine to another. Amazon EC2 also offers Availability Zones, which are distinct locations engineered to be insulated from failures in other zones. Administrators can launch Amazon EC2 instances in multiple Availability Zones to create geographically diverse, fault tolerant systems that are highly resilient in the event of network failures, natural disasters, and most other probable sources of downtime. Using Amazon S3, customer's data is replicated and automatically stored in separate data centers to ensure reliable data storage with a service level of 99.9% availability and no single points of failure.

The AWS Solution

Amazon Web Services (AWS) provides a reliable, scalable, and inexpensive computing platform "in the cloud" that can be used to facilitate healthcare customers' HIPAA-compliant applications. This platform is built on the same robust technology that Amazon.com uses to run its global web properties. Amazon EC2 offers a flexible computing environment with root access to virtual machines and the ability to scale computing resources up or down depending on demand. Amazon S3 offers a simple, reliable storage infrastructure for data, images, and back-ups. These services change the way organizations deploy, manage, and access computing resources by utilizing simple API calls and pay-as-you-use pricing. To learn more about the AWS solutions, visit http://aws.amazon.com.

Disclaimer

This white paper is not intended to constitute legal advice. You are advised to seek the advice of counsel regarding compliance with HIPAA and other laws that may be applicable to you and your business. Amazon Web Services LLC. and its affiliated entities make no representations or warranties that your use of Amazon Web Services will assure compliance with applicable laws, including but not limited to HIPAA.