SECURING SENSITIVE DATA WITHIN AMAZON WEB SERVICES EC2 AND EBS

The Challenges and the Solutions
This paper offers a detailed look at how organizations can protect data when using servers within Amazon Web Services (AWS) environments. This paper examines why securing data in AWS environments is so critical, particularly as the nature of cyber threats continues to evolve, and it details the specific challenges organizations face in establishing effective data protection. The paper then describes the core requirements of a data protection solution for AWS implementations, and then reveals how Vormetric Transparent Encryption for AWS addresses these requirements.

THE NEED FOR EC2 AND EBS DATA PROTECTION

AWS offers reliable, scalable, and cost-effective services that have fueled the growth of public cloud adoption among businesses, cloud service providers, and developers worldwide. With a keen understanding of the needs of developers, software-as-a-service (SaaS) providers, and IT organizations, AWS has been a pioneer in a number of new markets. Further, the organization continues to deliver new public cloud solutions that address the growing demands of customers.

AWS offers its customers optimal flexibility, delivering a broad array of cloud-based infrastructure services. By far the most popular approach is for customers to run server instances within Amazon Elastic Compute Cloud (EC2) and to store persistent data associated with those servers using Amazon Elastic Block Store (EBS).

Amazon’s Security Focus: Management, Network, and Identity

AWS management and network environments offer layered security, but they don't provide a customer controlled, on-premises key management and data access policy option for EC2 and EBS data. AWS environments feature a customized Xen hypervisor that isolates instances within the network and offers a hardened host management plane for administration. Amazon offers multi-factor authentication for cloud administrators and it logs and audits all management actions. The hosted systems are also protected by a mandatory inbound network firewall. Amazon even allows you to limit access to instances based on the AWS Identity and Access Management (IAM) service, the Amazon Virtual Private Cloud (VPC), or your own internal directory service. However, within your AWS instances, the responsibility for protecting the data is yours.

Data stored in EBS is usually directly linked to instances within AWS, appearing as a volume within the local system environment. In EBS environments, Amazon's focus is on offering access control through IAM and ensuring availability, but it does not offer specific protections for data stored within the EBS volume.

Drivers for Data Protection within AWS Environments

Whether your organization is an early-stage startup, a SaaS provider, or an established enterprise, AWS enables your business to take advantage of significant flexibility and economic benefits. All organizations share common requirements for securing their data in AWS environments:

- Establishing compliance with industry and government regulations. Compliance with industry and government regulations is a core driver. Regulations such as the Payment Card Industry Data Security Standard (PCI DSS), the Health Insurance Portability and Accountability Act (HIPAA) and the Health Information Technology for Economic and Clinical Health (HITECH) require specific data access controls, protection from privileged users, separation of duties, audit reporting, and in some cases data encryption. Failure to meet the requirements of a compliance audit can result in the loss of certification and the loss of business.
• Reducing exposure associated with data breach disclosure requirements. Most personally identifiable information (PII), including names, addresses, medical records, and financial statements, must be protected. Data breach laws worldwide—such as the UK Data Protection Act, the EU Data Protection Directive, South Korea's Personal Information Protection Act (PIPA), and US Federal and State data protection laws—impose fines and notification requirements in the event of a data breach, and outline specific protections and safe harbor criteria for encrypted data.

• Safeguarding intellectual property (IP). Many enterprises and government organizations also have a substantial set of intellectual property in the form of planning documents, manufacturing methods, designs, user profiles, source code, and other data that needs to remain confidential. If publicly disclosed, or acquired by a business or government rival, this information can expose an organization to financial losses, loss of customer trust, or, in the case of government agencies, even jeopardize national security interests.

“...in the end, customers are responsible. Customers will always be responsible in the public cloud for their applications and their data.”

Kyle Hilgendorf
Principal Research Analyst, Gartner
Blog: Cloud Security Configurations: Who is Responsible?

Expanded Threats to Data across Both Enterprises and AWS

The last few years have seen a well-documented change in the advanced threats organizations are encountering. For years, the vast majority of hackers were motivated by a desire for fame, recognition, or support for a specific cause. Increasingly, cyber attacks are routinely being waged by sophisticated criminal organizations and government-sponsored agencies. These well-organized and well-funded groups are looking for specific data sets that can further their financial or geo-political agendas. Major financial organizations and other business institutions are being targeted, with the goal of destabilizing entire economies. Manufacturers have become prime targets of criminals looking to steal critical intellectual property, formulas, product plans, and other information. Criminal organizations target game sites to gain credential sets that can be used to compromise accounts within financial or other organizations. While attacks and attackers continue to evolve, one fact remains constant: sensitive data is the target.

Attacks continue to grow more sophisticated in evading detection. Advanced persistent threats (APTs) are top of mind for every security organization. According to Mandiant investigations, it took victims of these attacks an average of over 200 days to discover their perimeter security had been penetrated—in spite of the fact these organizations all had up-to-date antivirus software. In addition, 100% of the breaches involved attackers stealing credentials in order to bypass defenses and gain access to sensitive systems. Given these threats, many decision makers are understandably reluctant to have their organizations introduce another potential set of risks by relying on critical infrastructure outside their perimeter—including in cloud environments such as AWS.

“...sending 10 phishing e-mails approaches the point where most attackers would be able to slap a ‘guaranteed’sticker on getting a click.”

Verizon Data Breach Report
The Inevitability of "The Click"
Privileged users represent another threat that raises concerns in AWS environments. The disclosures of Edward Snowden highlight the risks privileged users represent. As a system administrator, Snowden had access to massive amounts of highly sensitive data, even though he didn’t need this data to fulfill his role within the organization. Given the unknown number of cloud administrators and other privileged users working in AWS environments, this threat can pose significant concerns for executives considering moving sensitive data into these environments.

In addition, AWS snapshots create another area of exposure. Many privileged users also can gain access to snapshots of EC2 instances, which can include critical data in local storage. If these privileged users decide to exploit this access, or if their credentials are compromised, these data snapshots can lead to a significant compromise.

These risks prompt critical questions decision makers need to answer when considering a move to AWS.

- Is it possible to meet compliance requirements when using AWS?
- How can my organization maintain control of our sensitive data?
- Will using AWS increase my risks of exposure to a data breach?
- How can I protect my data from cloud administrators and other privileged users in AWS?
- Will using AWS increase my exposure to APTs?
- How will I maintain and prove data residency in AWS?

**CORE SOLUTION ELEMENTS**

To address the concerns outlined above, it is critical to employ a data-centric security solution in your AWS instances. A data-centric solution places the security controls and protections directly around the data that attackers are targeting. Protections must reside at the file-system level and be accessible by your EC2 instances, both those that reside locally and those in EBS. The solution should also protect data wherever it resides, including snapshots, backup repositories, and disaster recovery (DR) locations.

**Integrated Encryption and Key Management.** Using strong industry-standard algorithms to encrypt data is the first step. Integrated, centralized encryption key management should be seamless and simple—and should offer deployment options that match your operational needs. Solutions should support the secure storage of all keys within the AWS cloud, or hybrid cloud implementations, such as an AWS VPC implementation in which computing assets reside both in AWS and on the customer premises. VPC implementations should offer key management in the cloud or within your data center. To establish maximum security, keys must stored separately from data and never revealed, even to security administrators.

This combination of strong encryption with integrated key management is required to meet baseline compliance requirements, provide a safe harbor from data breach disclosures, and adhere to best practices for protecting critical information.
Access Policies with Privileged User Access Controls to Encrypted Data. When used with EC2 and EBS, the AWS IAM service is focused on controlling network access to instances, but it does not provide access controls for data residing within the AWS instance. To prevent unauthorized users from gaining access to the data within instances, there must be strong, centrally managed access control policies employed at the file system level that strictly enforce when data can be decrypted. To control data access, users and groups should be linked to system users as defined within the instance. Also, if an Amazon VPC is used, access controls should link to your organization’s directory services solution.

The principle of least privilege should apply as the basis for data access, which means a person should have access only to what they need in order to perform assigned tasks. A trial period that “audits” data access should be available to make the process of setting policy simple. By default, policy controls should deny any user or process access to sensitive data when there is no business need for that access. Privileged users, such as system administrators or root users of Linux systems, need access to file metadata without gaining access to sensitive information.

By integrating access controls with encryption capabilities, organizations can establish strong privileged user access controls—enabling privileged users to do system management, updates, and other standard functions, without seeing protected database tables or other files in the clear. Data should only be decrypted for authorized users and processes.

Especially for Linux systems, access controls should be able to track how a user obtained their role. For instance, in a Linux environment, a root user can escalate their own privileges to become a database user. Solutions should be able to track the chain of privilege and authority changes, recognize and identify users based on their original role, and then block access if the user’s original role does not have authorization to access data.

Establishing these strong controls around sensitive data access are essential elements for many compliance mandates—including the PCI DSS and HIPAA/HITECH. These access controls represent a set of best practices for preventing data breaches and protecting intellectual property. When applied properly, these controls can enable organizations to establish secure shared storage scenarios in which users in individual departments like human resources, finance, and research and development can see their own group’s data, but never access data from other departments.

“I have been a systems engineer, systems administrator...When you are in positions of privileged access like a systems administrator for the intelligence community, you are exposed to a lot more information on a broader scale than the average employee.”

Edward Snowden
Former Infrastructure Analyst at the NSA
Policy-based access controls to encrypted data also solve the problems associated with sensitive data residing on snapshots, backup locations, and DR implementations. For snapshots and backups, encrypted data resides within the image or backup file, which makes the data inaccessible until it is restored and the appropriate policy is applied. In DR environments, bringing up a DR location should be as simple as applying the same policy used to access the source AWS instance.

**Separation of Roles for Security and Systems Management.** Strict segregation of security management and systems management roles is another key factor in a comprehensive data security solution. Data protection security roles should be separate from network security roles. Users with network security roles should only be allowed to make changes to network policies. Within AWS, users with security roles should only have control over data access policies relating to system instances. Likewise, administrators with systems management roles should not have visibility into security management, or have the ability to change data access policies.

**Security Intelligence: Detailed Access Logging and Auditing.** Policy-based access controls to encrypted data address many risks by ensuring only authorized users can access data, and preventing privileged users from accessing data. However, when an APT compromises an authorized user’s account or a malicious insider gains access to data, these situations require additional security analysis to detect. A two-step process is required to identify these threats. The first step is to establish which users and processes are accessing data, and the second is to analyze data access patterns to recognize unusual or anomalous activity.

Solutions should enable users to select specific data log collections. For instance, when collecting file system access information for a database table, accesses by the database process represent normal operations that should be excluded from logs. This filtering enables security analysts to focus on the data logs that may point to a real threat.

Audit reports should be a base feature of a solution in this area. These reports should address the access reporting requirements of compliance auditors.

Data collected should also include access to the security management infrastructure. This is essential in giving you the detailed information you need to track attempts to access the security management infrastructure and the usage patterns of security administrators.

> Many organizations understand that data and system access is the first step to understanding and isolating potential data breaches due to malicious insiders or advance threats. Splunk and Vormetric together can quickly and efficiently help businesses protect intellectual property and other data the business wishes to be kept private.

**Bill Gaylord**  
**Splunk**  
**SVP of Business Development**
Security Intelligence: SIEM Analysis, Alerts, and Reports. It is also vital to leverage the detailed information held within access logs. One use of this information is for generating immediate alerts when unauthorized data access occurs. Log data collected should show when unauthorized users attempt to access either protected data or the security management infrastructure. Security information and event management (SIEM) solutions enable analysis of this information so monitoring can be conducted and alerts can be generated when these events occur. The second major use is for establishing usage patterns associated with authorized users and processes. SIEM systems enable you to generate top user information and establish baselines for typical usage—and generate alerts when a deviation from normal behavior is detected.

One example of anomaly detection and reporting is to monitor user access rates and identify the most active users. When a privileged user who doesn't typically access financial information, suddenly begins to appear as a top user accessing the finance department's records, this behavior change could indicate that an account has been compromised by an APT or that a malicious insider incident is in process. Baseline patterns are another use case that should be addressed. Baselines capture typical access patterns over a period of time (typically over a weekly or monthly period). This allows alert notifications based on unexpected behavior patterns. For example, if an account that typically accesses only certain classes of data at month end suddenly begins to continuously access much more diverse data sets over multiple weekends, this activity may indicate a problem.

Integration Capabilities. In addition to integrating directory services, it is also important to integrate policy control, deployment, governance risk and compliance, and other tools. This integration is important whether your organization employs an “all-in-the-cloud” implementation or an Amazon VPC/VPN hybrid model integration. Web-based APIs or command line options are required to make the required connections. In addition, integration capabilities should enable data protection solutions to dynamically adjust policy based on real-time threat analysis. This is a core requirement for SaaS providers as well, giving them the flexibility they need to scale infrastructure to align with customer demand.

Multi-tenancy and Business Unit Segregation. For SaaS providers that implement their infrastructure within AWS, multi-tenancy is a core requirement, allowing them to segment data access and management by customer. Multi-tenancy allows for multiple customers to use a common infrastructure, while keeping customer data and access separate. For enterprises that have a “data-across-borders” requirement, or the need to simply isolate business units from each other (as
frequently occurs in large-multinational organizations as well as in defense and aerospace enterprises), the capability to isolate management and data access allows the use of common applications and infrastructures in order to reduce resource requirements.

**Scalability and Performance.** Data protection solutions should easily support scalability in AWS environments, from supporting deployments with a few servers to large hybrid deployments that use Amazon VPC or elastic computing and on-premises resources. These solutions should offer the performance required to minimize any impact on transaction times, and support existing SLAs.

**VORMETRIC TRANSPARENT ENCRYPTION FOR AWS ENVIRONMENTS**

Vormetric Transparent Encryption for AWS delivers all the capabilities required to establish strong data protection within AWS environments. The solution includes:

- Integrated encryption and key management that protects data at the file system level within EC2 and EBS instances
- Access policies with privileged user access controls that ensure data is only decrypted for authorized users and processes, while allowing system and cloud administrators to do their work without accessing sensitive data
- Separation of roles for security and systems management that supports best practices for data protection
- Security intelligence with data event logs that enable integration with SIEM tools to support event analysis and compliance reporting
- APIs and command line interfaces for integration with other infrastructure tools
- Multi-tenancy and business unit segregation support within the management infrastructure
- Highly scalable management infrastructure
- High performance operation that supports existing SLAs and operations
**Transparent.** The solution is completely transparent, which means it can be deployed without requiring any process, application, database, or infrastructure changes. The file system level encryption solution allows administrators to see the metadata and file system structure, but reveals only encrypted data to those accounts. At the same time, processes and users that legitimately require data can access it in clear text.

**Strong.** Vormetric Transparent Encryption protects your data using a policy-driven approach, which is linked to LDAP and system accounts to provide granular access to protected structured information in databases or unstructured data in file systems. The solution enables policies to be instituted by process, user, time, and other parameters. The solution even monitors and prevents access by tracking how users obtain their role. If a root user creates a new account with data access rights, and then logs in using the new account, Vormetric will still associate actions performed by this new account with the root user and prevent unauthorized access to cleartext data. As a result of these capabilities, organizations can enable privileged users to manage systems, without risking the exposure of protected information.

**Efficient.** Vormetric Transparent Encryption for AWS is a high-performance, low-overhead solution that has minimal impact on the response times of operational processes.

**Easy to Deploy.** The solution can be deployed in AWS in minutes. More extensive enterprise and cloud deployments can be deployed in days to weeks, not weeks to months.

**Vormetric Transparent Encryption for AWS: Offerings and Components**

There are two core Vormetric components in the AWS data protection solution: The Vormetric Data Security Manager (DSM) and Vormetric Transparent Encryption for AWS Client Systems.

The DSM for AWS provides the secure management of encryption, keys, access controls, and integration across client systems within your AWS environment. The DSM includes in-depth data access policy control, auditing and reporting, management, and integration with LDAP and other tools.

The Vormetric Transparent Encryption for AWS Client agent is available for AWS CentOS 6.3 64bit, and is ready for immediate connection and use with DSM.

Vormetric Transparent Encryption is available as an AWS Marketplace offering and a bring-your-own-license (BYOL) product. The AWS Marketplace offering offers a fast and simple way to get started. BYOL is flexible, allowing for additional clients, more client system versions beyond CentOS, and extended deployment scenarios that include highly scalable deployments within AWS as well as hybrid deployment capabilities for managing keys and policies from VPC environments or on-premises locations.

**AWS Deployment Models**

Vormetric’s comprehensive data protections solutions support a number of deployment models, including all-in-the-cloud approaches and hybrid environments comprised of AWS and private clouds.

The AWS Marketplace offering is available for immediate activation with a pay-as-you-go model that makes it easy to implement. Client software is easy to install on CentOS clients.
SaaS vendors and larger enterprise customers will typically deploy larger numbers of client instances within AWS, and may support multiple customers with independent infrastructure sets, while centrally managing data security.

Lastly, enterprise customers that want to control their keys and manage data security for their AWS cloud instances on premises can use Amazon’s VPC with a VPN link to AWS server instances.

Regardless of the deployment scenario, Vormetric Transparent Encryption for AWS offers a data-centric security approach that reduces risk by limiting attack surfaces. This approach allows organizations to take advantage of the convenience, business flexibility, and scalability of AWS environments—while addressing their compliance requirements, guarding against data breaches, and protecting critical IP.

ABOUT VORMETRIC

A leader in data security solutions, Vormetric (@Vormetric) protects data-at-rest in physical, virtual, big data, and cloud environments. Trusted by businesses and governments for over a decade, the Vormetric Data Security Platform secures the data of more than 1,500 global enterprises—including 17 of the Fortune 30. With Vormetric, a single infrastructure and management environment protects data wherever it resides with file, volume, and cloud storage encryption, tokenization with dynamic data masking, field-level application encryption, sophisticated access control policies, and third party and integrated encryption key management. For more information, please visit: www.vormetric.com.