DATA SECURITY AND STORAGE

Data security is the practice of protecting digital information from unauthorized access, corruption or theft throughout its entire lifecycle.

Data storage security is a segment of the cybersecurity field specifically focused on protecting data and storage infrastructure against unauthorized disclosure, modification, or destruction, and assuring the data is only accessible to authorized users. This takes many forms, and a successful data storage security strategy includes hardware, software, risk assessments, administrative controls, and policy components. This article provides an overview of data storage security and how it fits in with data protection and DevSecOps, and looks at storage security vulnerabilities and key drivers in the data storage security field.

Types of Data Security

To enable the confidentiality, integrity and availability of sensitive information, organizations can implement the following data security measures:

1. Encryption
2. Data erasure
3. Data masking
4. Data resiliency

Encryption

By using an algorithm to transform normal text characters into an unreadable format, encryption keys scramble data so that only authorized users can read it. File and database encryption software serve as a final line of defense for sensitive volumes by obscuring their contents through encryption or tokenization. Most encryption tools also include security key management capabilities.

Data erasure

Data erasure uses software to completely overwrite data on any storage device, making it more secure than standard data wiping. It verifies that the data is unrecoverable.

Data masking

By masking data, organizations can allow teams to develop applications or train people that use real data. It masks personally identifiable information (PII) where necessary so that development can occur in environments that are compliant.

Data resiliency

Resiliency depends on how well an organization endures or recovers from any type of failure—from hardware problems to power shortages and other events that affect data availability. Speed of recovery is critical to minimize impact.

Data Security Strategies

A comprehensive data security strategy incorporates people, processes and technologies. Establishing appropriate controls and policies is as much a question of organizational culture as it is of deploying the right tool set. This means making information security a priority across all areas of the enterprise.

Consider the following facets in your data security strategy:

- Physical security of servers and user devices
- Access management and controls
- Application security and patching
- Backups
- Employee education
- Network and end point security monitoring and controls

The cloud data protection and security strategy must also protect data of all types. This includes:

Data in use: Securing data being used by an application or endpoint through user authentication and access control.

Data in motion: Ensuring the safe transmission of sensitive, confidential or proprietary data while it moves across the network through encryption and/or other email and messaging security measures.

Data at rest: Protecting data that is being stored on any network location, including the cloud, through access restrictions and user authentication.

Data Security Mitigation

If prospective customers of cloud computing services expect that data security will serve as compensating controls for possibly weakened infrastructure security, since part of a customer's infrastructure security moves beyond its control and a provider's infrastructure security may (for many enterprises) or may not (for small to medium- size businesses, or SMBs) be less robust than expectations, you will be disappointed. Although data-in-transit can and should be encrypted, any use of that data in the cloud, beyond simple storage, requires that it be decrypted. Therefore, it is almost certain that

in the cloud, data will be unencrypted. And if you are using a PaaS-based application or SaaS, customer-unencrypted data will also almost certainly be hosted in a multitenancy environment (in public clouds). Add to that exposure the difficulties in determining the data's lineage, data provenance—where necessary—and even many providers' failure to adequately address such a basic security concern as data remanence , and the risks of data security for customers are significantly increased.

Provider Data and Its Security

In addition to the security of your own customer data, customers should also be concerned about what data the provider collects and how the CSP protects that data. Specifically, with regard to your customer data, what metadata does the provider have about your data, how is it secured, and what access do you, the customer, have to that metadata? As your volume of data with a particular provider increases, so does the value of that metadata.

Additionally, your provider collects and must protect a huge amount of security-related data. For example, at the network level, your provider should be collecting, monitoring, and protecting firewall, intrusion prevention system (IPS), security incident and event management (SIEM), and router flow data. At the host level your provider should be collecting system log files, and at the application level SaaS providers should be collecting application log data, including authentication and authorization information.

Storage

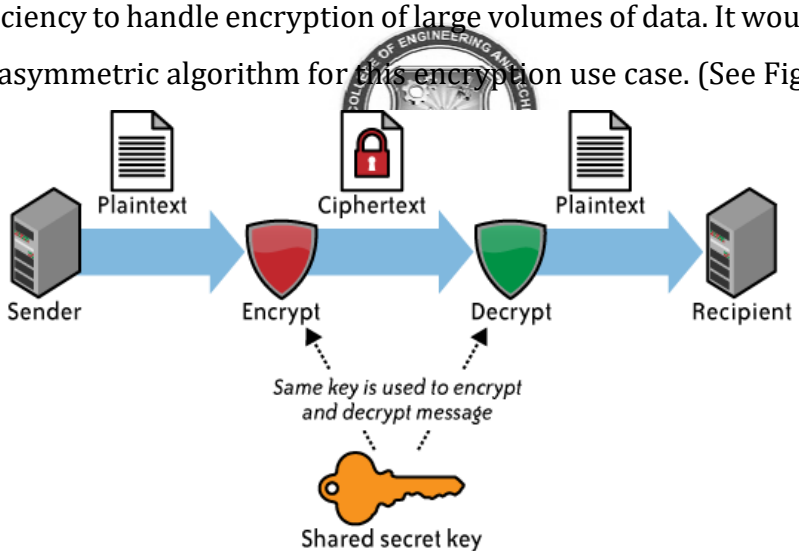
For data stored in the cloud (i.e., storage-as-a-service), we are referring to IaaS and not data associated with an application running in the cloud on PaaS or SaaS. The same three information security concerns are associated with this data stored in the cloud (e.g., Amazon's S3) as with data stored elsewhere: confidentiality, integrity, and availability.

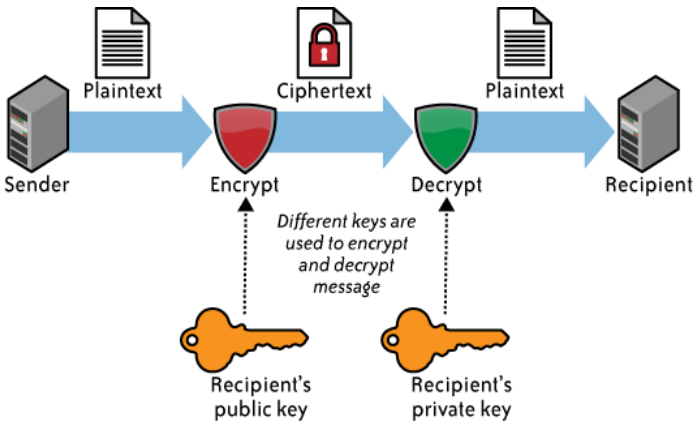
Confidentiality

When it comes to the confidentiality of data stored in a public cloud, you have two potential concerns. First, what access control exists to protect the data? Access control consists of both authentication and authorization. As we will discuss further in CSPs generally use weak authentication mechanisms (e.g., username + password), and the authorization ("access") controls available to users tend to be quite coarse and not very granular. For large organizations, this coarse authorization presents significant

security concerns unto itself. Often, the only authorization levels cloud vendors provide are administrator authorization (i.e., the owner of the account itself) and user authorization (i.e., all other authorized users)—with no levels in between (e.g., business unit administrators, who are authorized to approve access for their own business unit personnel).

If a CSP does encrypt a customer's data, the next consideration concerns what encryption algorithm it uses. Not all encryption algorithms are created equal. Cryptographically, many algorithms provide insufficient security. Only algorithms that have been publicly vetted by a formal standards body (e.g., NIST) or at least informally by the cryptographic community should be used. Any algorithm that is proprietary should absolutely be avoided. Note that we are talking about symmetric encryption algorithms here. Symmetric encryption (see Figure 4-1) involves the use of a single secret key for both the encryption and decryption of data. Only symmetric encryption has the speed and computational efficiency to handle encryption of large volumes of data. It would be highly unusual to use an asymmetric algorithm for this encryption use case. (See Figure 4-2.)





Integrity

In addition to the confidentiality of your data, you also need to worry about the integrity of your data. Confidentiality does not imply integrity; data can be encrypted for confidentiality purposes, and yet you might not have a way to verify the integrity of that data. Encryption alone is sufficient for confidentiality, but integrity also requires the use of message authentication codes (MACs). The simplest way to use MACs on encrypted data is to use a block symmetric algorithm (as opposed to a streaming symmetric algorithm) in cipher block chaining (CBC) mode, and to include a one-way hash function. This is not for the cryptographically uninitiated—and it is one reason why effective key management is difficult. At the very least, cloud customers should be asking providers about these matters. Not only is this important for the integrity of a customer's data, but it will also serve to provide insight on how sophisticated a provider's security program is—or is not. Remember, however, that not all providers encrypt customer data, especially for PaaS and SaaS services.

Availability

Assuming that a customer's data has maintained its confidentiality and integrity, you must also be concerned about the availability of your data. There are currently three major threats in this regard—none of which are new to computing, but all of which take on increased importance in cloud computing because of increased risk. The first threat to availability is network-based attacks under Infrastructure Security: The Network Level. The second threat to availability is the CSP's own availability. No CSPs offer the sought-after "five 9s" (i.e., 99.999%) of uptime. A customer would be lucky to get "three 9s" of uptime.

Percentage of uptime

	Total downtime (HH:MM:SS)		
Availability	**Per day**	**Per month**	**Per year**
99.999%	00:00:00.4	00:00:26	00:05:15
99.99%	00:00:08	00:04:22	00:52:35
99.9%	00:01:26	00:43:49	08:45:56
99%	00:14:23	07:18:17	87:39:29