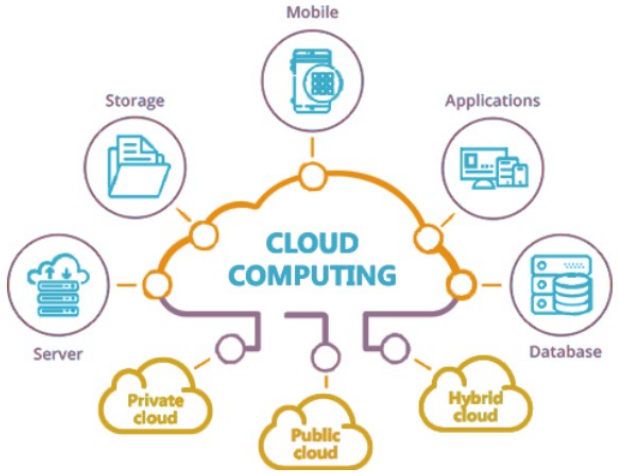


Key Essesntials for Building Apps in Cloud

Ravindu Nirmal Fernando

SLIIT | March 2025

Shared Responsibility Model in Public Cloud

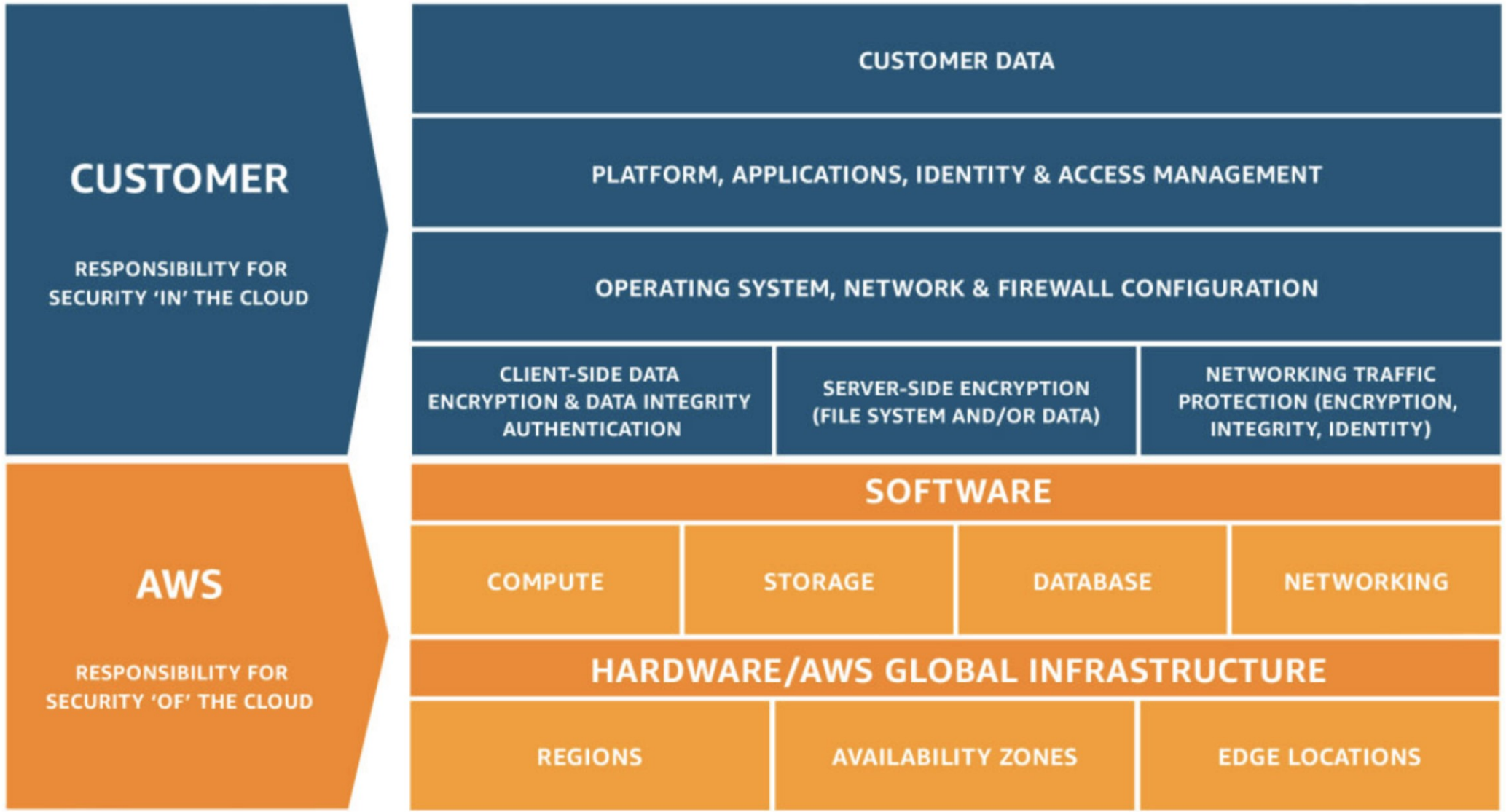
- A framework that outlines how security responsibilities are divided between the cloud service provider and the cloud user.
- **Security and compliance** is a shared responsibility between AWS and the customer. Cloud service provider manages the infrastructure, while customers are responsible for managing their data and applications.
- Cloud Service Provider (CSP) Responsibility: Known as "**Security of the Cloud."** CSP is in charge of the infrastructure, including hardware, software, networking, and physical security.
- Customer Responsibility: Termed "**Security in the Cloud."** Customers handle the guest operating system, application software, and AWS-provided firewall configuration.

Key aspects of Shared Responsibility Model

- **Service/ Delivery Models**: Responsibilities vary depending on whether the service is IaaS (like EC2), PaaS, or SaaS.
- **IT Controls**: Shared management of IT controls between CSP and customers. CSP manages physical infrastructure controls, while customers handle specific application-level controls.

Control types in Shared Responsibility Model

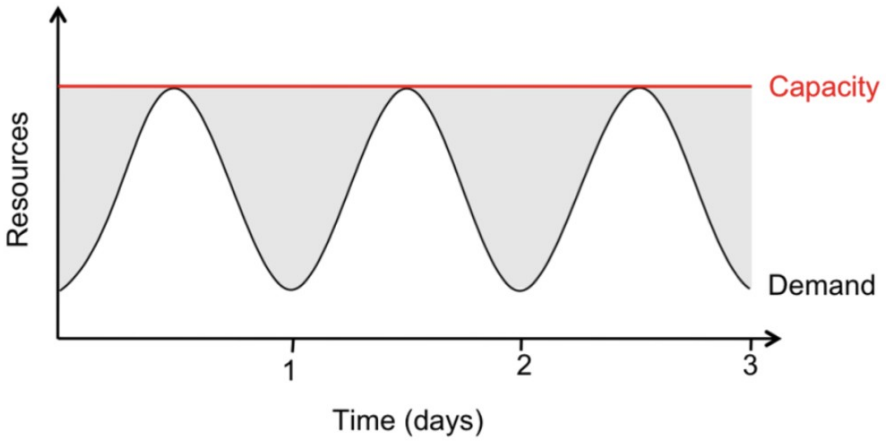
- **Inherited Contr**ols: Controls fully managed by CSP (e.g., physical and environmental controls).
- **Shared Controls**: Controls that apply to both CSP and customers but in different contexts (e.g., patch management, configuration management).
  - Patch Management – CSP is responsible for patching and fixing flaws within the infrastructure, but customers are responsible for patching their guest OS and applications.
  - Configuration Management – CSP maintains the configuration of its infrastructure devices, but a customer is responsible for configuring their own guest operating systems, databases, and applications.
  - Awareness & Training - CSP trains CSP's employees, but a customer must train their own employees.
- **Customer Specific Controls**: Controls solely managed by the customer, depending on their applications and use of CSP services.


CUSTOMER
RESPONSIBILITY FOR SECURITY 'IN' THE CLOUD
CUSTOMER DATA
PLATFORM, APPLICATIONS, IDENTITY & ACCESS MANAGEMENT
OPERATING SYSTEM, NETWORK & FIREWALL CONFIGURATION
CLIENT-SIDE DATA ENCRYPTION & DATA INTEGRITY AUTHENTICATION
SERVER-SIDE ENCRYPTION (FILE SYSTEM AND/OR DATA)
NETWORKING TRAFFIC PROTECTION (ENCRYPTION, INTEGRITY, IDENTITY)
AWS
RESPONSIBILITY FOR SECURITY 'OF' THE CLOUD
SOFTWARE
COMPUTE
STORAGE
DATABASE
NETWORKING
HARDWARE/AWS GLOBAL INFRASTRUCTURE
REGIONS
AVAILABILITY ZONES
EDGE LOCATIONS

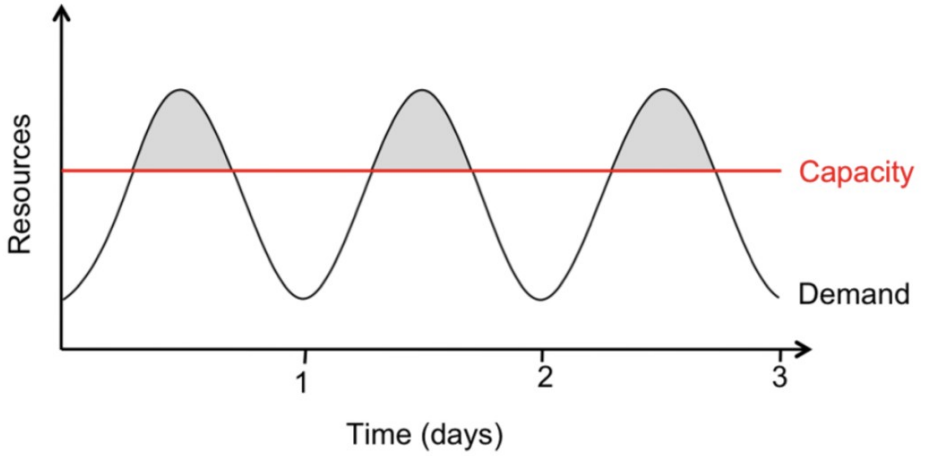
How to use Shared Responsibility Model practically…

- Understanding the Model: Customers need to comprehend the CSP Shared Responsibility Model and its general application in cloud operations.
- Application to Use Case: Determine the model's relevance to their specific use case.
- Variability in Responsibility: Customer responsibility changes based on:
  - The choice of CSP services and geographical locations. (e.g: AWS EC2 in specified AWS region)
  - How these services integrate into their IT environment.
- Legal and Regulatory Considerations: Consideration of laws and regulations that apply to their organization and workload.

Resource Provisioning



Provisioning for Peak Load

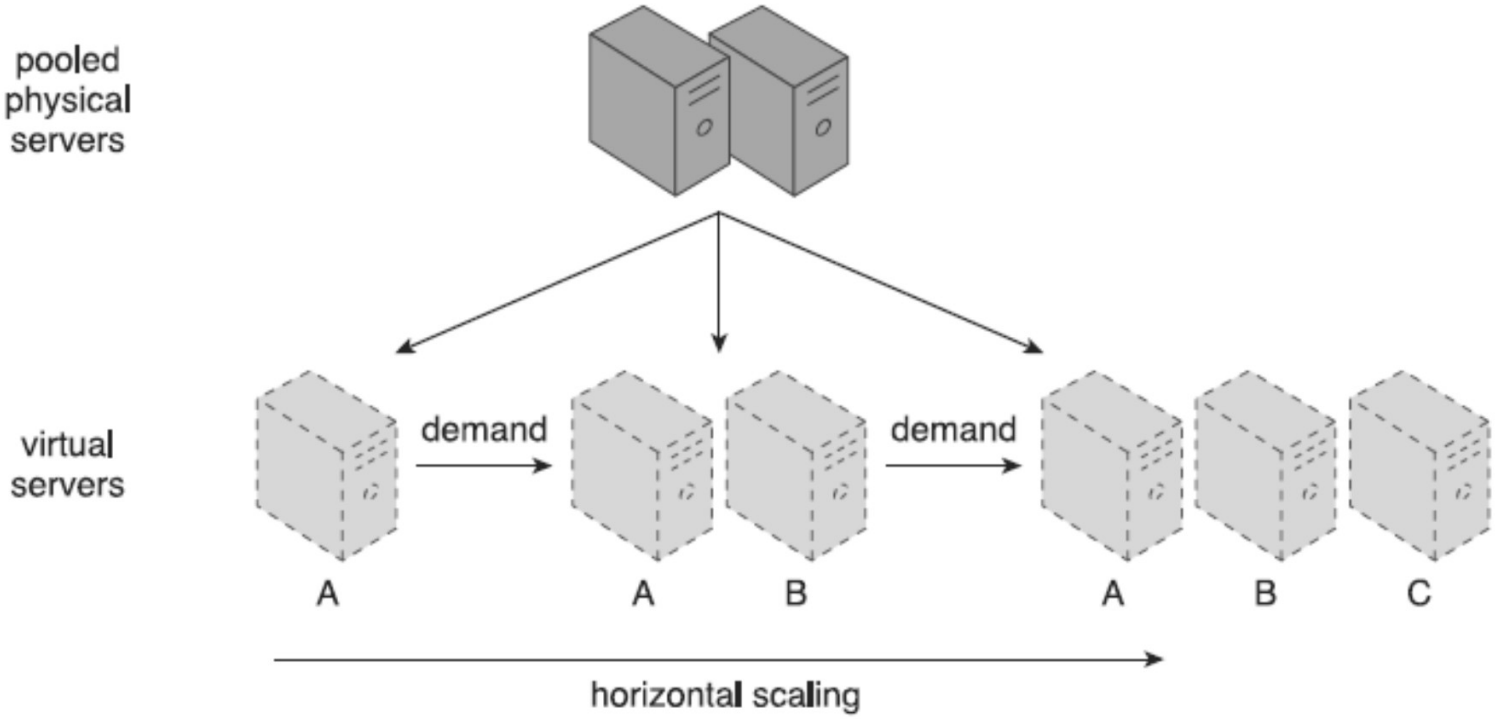


Under-Provisioning

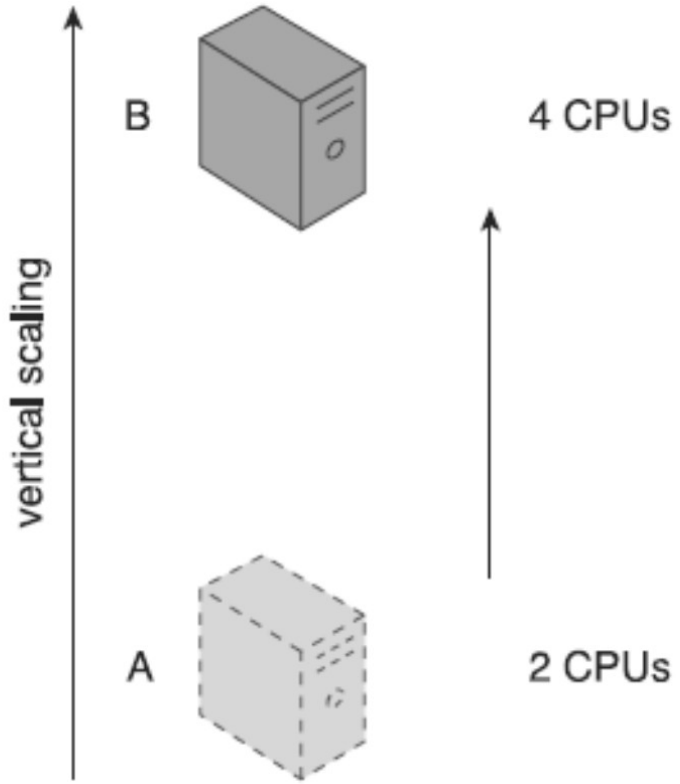
Reference: Above the Clouds: A Berkeley View of Cloud Computing, 2009.

Scalability in Cloud

- Horizontal scaling
  - The allocating or releasing of IT resources that are of the same type
  - Scaling in and out



- Vertical scaling
  - Existing IT resource is replaced by another with higher or lower capacity
  - Scaling up & down



Horizontal vs Vertical Scaling

Aspect	Horizontal Scaling (Scaling Out/In)	Vertical Scaling (Scaling Up/Down)
Definition	Adding or removing servers to adjust capacity.	Increasing or decreasing the capacity of a server.
Cost	Can be more cost-effective with pay-as-you-go models.	May involve higher costs due to high-end hardware.
Downtime	Often allows scaling with no downtime.	May require downtime for hardware upgrades.
Resource Limits	Limited by the number of servers you can add.	Limited by the maximum capacity of a single server.
Complexity	Can increase architectural complexity.	Simpler, as it involves a single resource.
Availability	Improved, as load is distributed across multiple servers.	Risk of a single point of failure.
Use Case	Ideal for distributed systems and microservices.	Suited for applications with fixed or known peaks.

- Reactive scaling
  - Once something (e.g., workload) happen
- Proactive scaling
  - Based on predictions (e.g., workload)

Based on

- Rules
  - Spawn a new VM if ave. CPU util. > 80%
- Models based on QoS/SLA targets
  - No of VMs to maintain latency < 300 ms

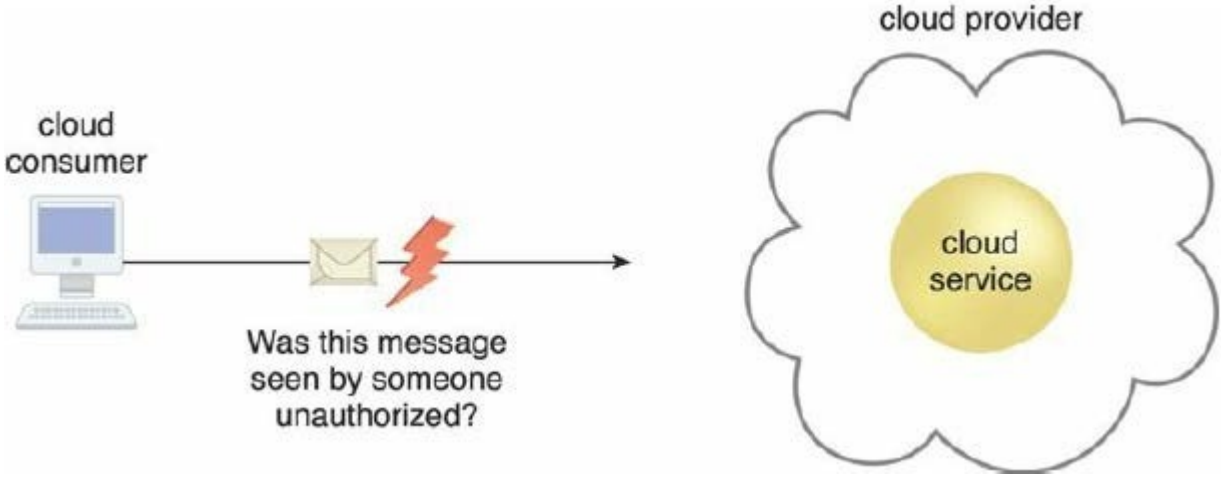
Cloud Security Basics and Countermeasures

Based on Cloud Computing: Concepts, Technology Architecture, Thomas Erl, et al., Prentice-Hall, 2013,

Concepts

Confidentiality

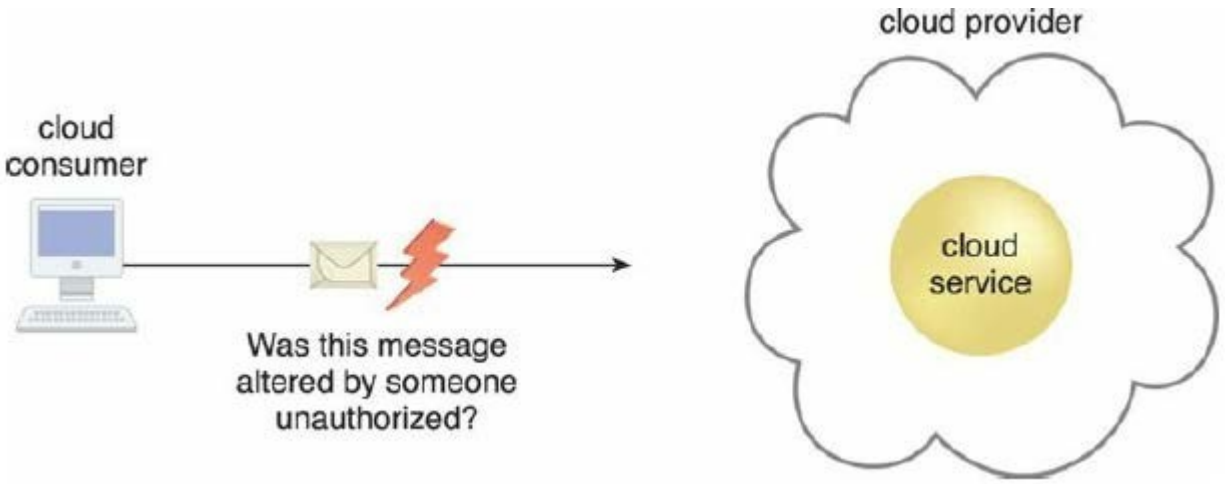
- Accessible only to authorized parties
- Within cloud environments, confidentiality targets to restricting access to data in transit and storage.



Concepts

Integrity

- Not having been altered by an unauthorized party
- Can cloud consumer be guaranteed transmitted data to matches the data received.
- Extends to how data is stored, processed, and retrieved.



Concepts

Authenticity

- Ensuring something has been provided by an authorized source.
- Can cloud consumer guarantee the authentication of an interaction and no other party can deny or challenge that.

Availability

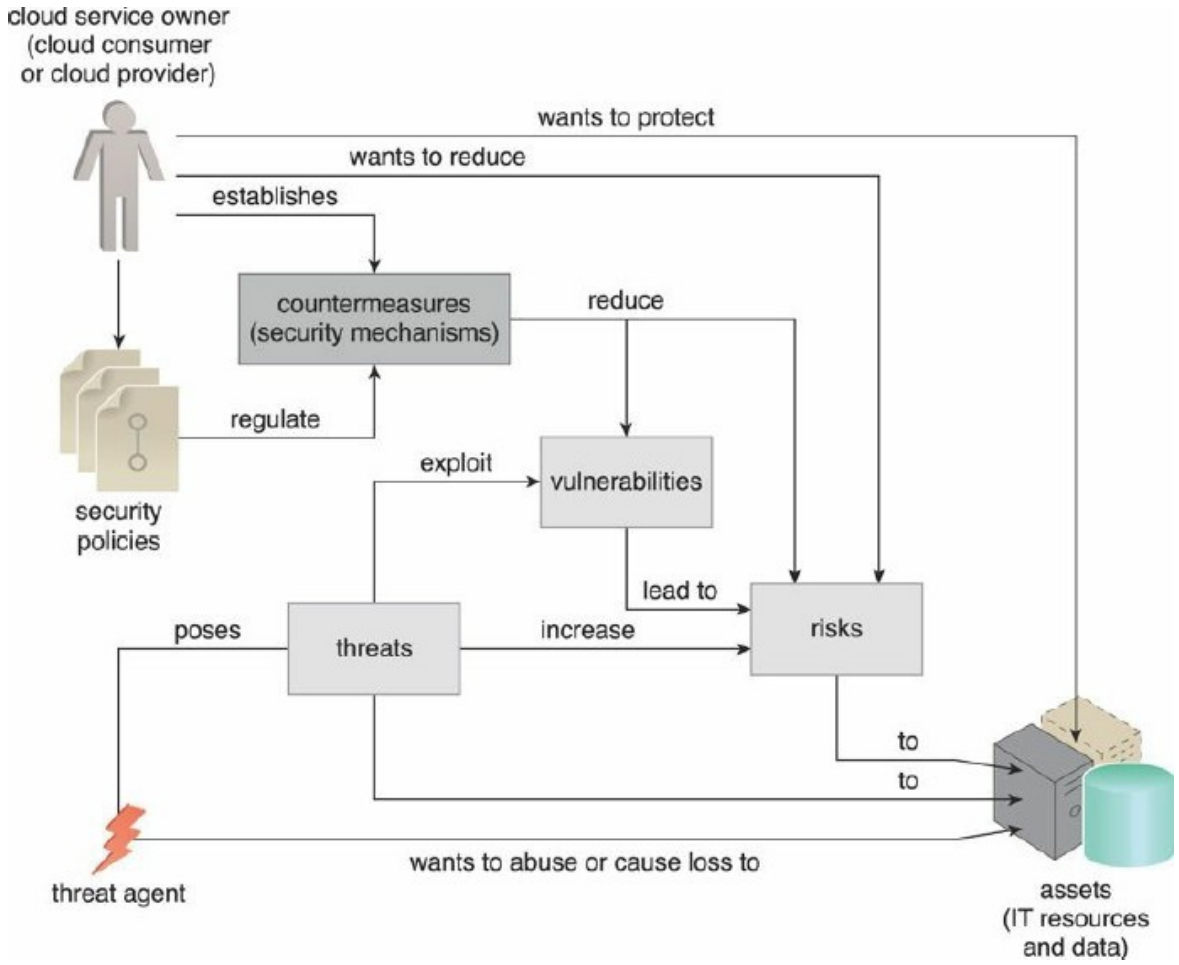
- Being accessible, available and usable within defined time period.
- In cloud the availability of cloud services can be a responsibility that is shared by the cloud provider and the cloud carrier. The availability of a cloud-based solution that extends to cloud service consumers is further shared by the cloud consumer.

Threat Agents

An entity that poses a threat because it is capable of carrying out an attack.

- Can originate either internally or externally.
- Human or Software.

Threat Agents



Threat Agents

Anonymous Attacker

- Non-trusted cloud service consumer without permissions in the cloud. Attempts attacks from outside cloud permission boundary, mostly using public networks.

Malicious Service Agent

- Able to intercept and forward the network traffic that flows within a cloud. Then to maliciously use and augment the data.

Trusted Attacker

- Shares IT resources in the same cloud environment as the cloud consumer and attempts to exploit legitimate credentials to target cloud providers and the cloud tenants.

Malicious Insider

- Human threat agents acting on behalf of or in relation to the cloud provider. Typically current or former employees or third parties with access to the cloud provider’s premises.

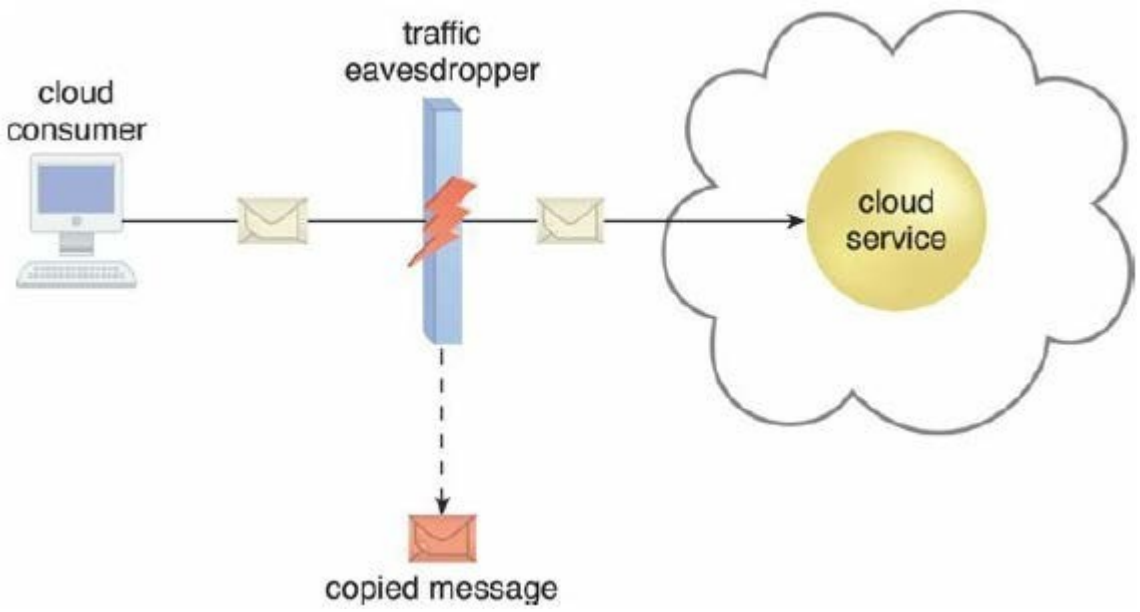
Cloud Security Threats

- Traffic Eavesdropping
- Malicious Intermediary
- Denial of Service
- Insufficient Authorization
- Virtualization Attack
- Overlapping Trust Boundaries

Traffic Eavesdropping

Data transferred to or within a cloud is passively intercepted by a malicious service agent for information gathering purposes.

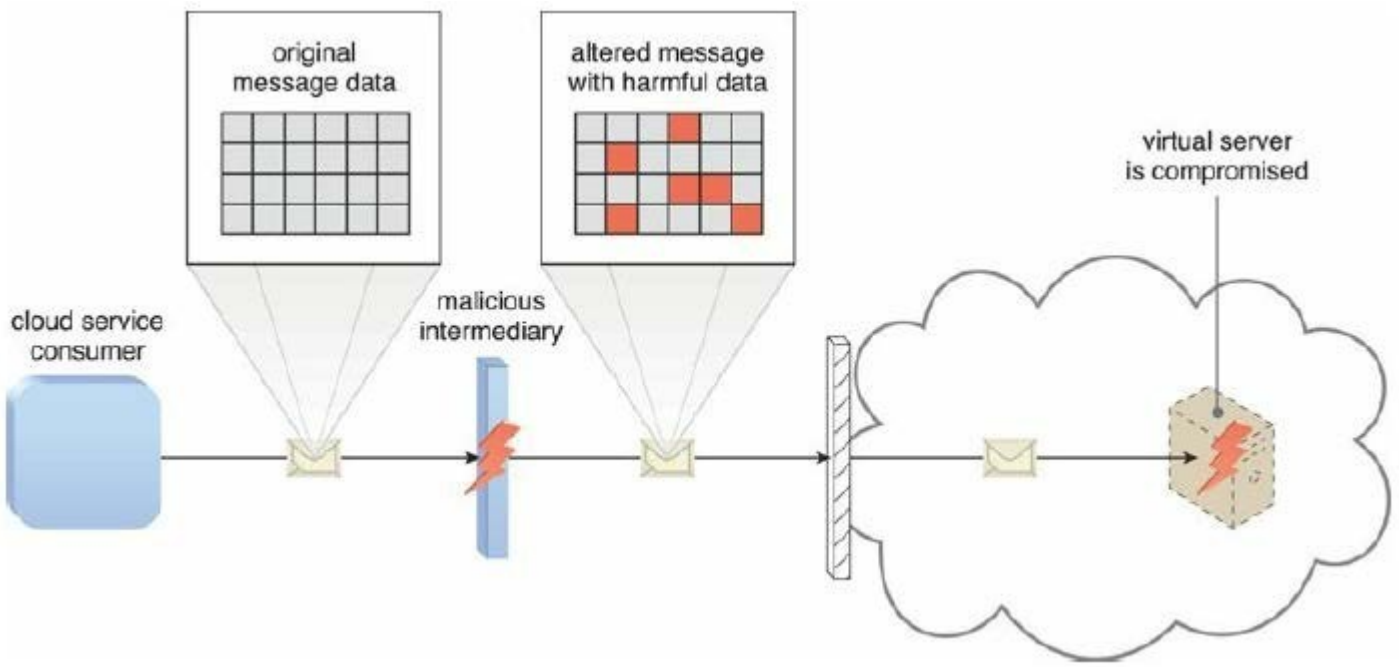
- Aim to compromise the confidentiality of the data.
- Due to passive nature of the attack, it can take place undetected for extended periods of time.



Malicious Intermediary

Messages are intercepted and altered by a malicious service agent.

- Potentially compromising the message's confidentiality and/or integrity.
- May insert harmful data into the message.

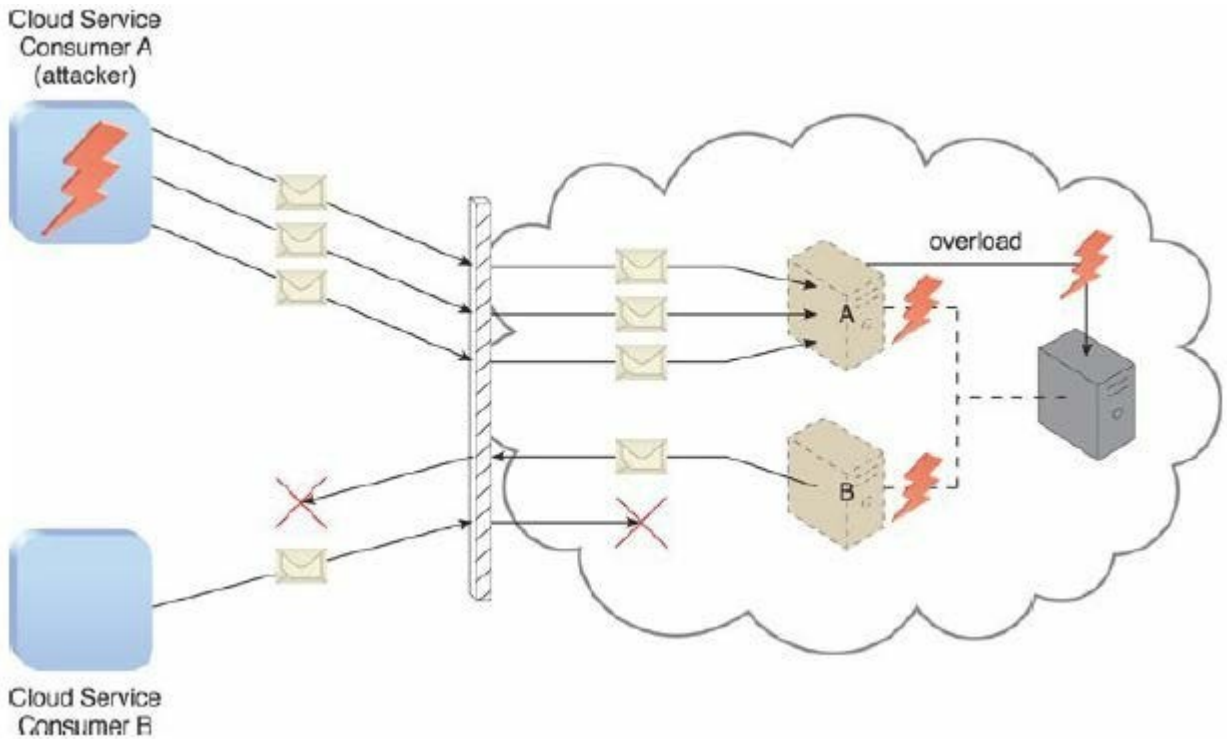


Denial of Service

Overload IT resources to the point where they cannot function properly.

- The workload on cloud services is artificially increased with imitation messages or repeated communication requests.
- The network is overloaded with traffic to reduce its responsiveness and cripple its performance.
- Multiple cloud service requests are sent, each of which is designed to consume excessive memory and processing resources.

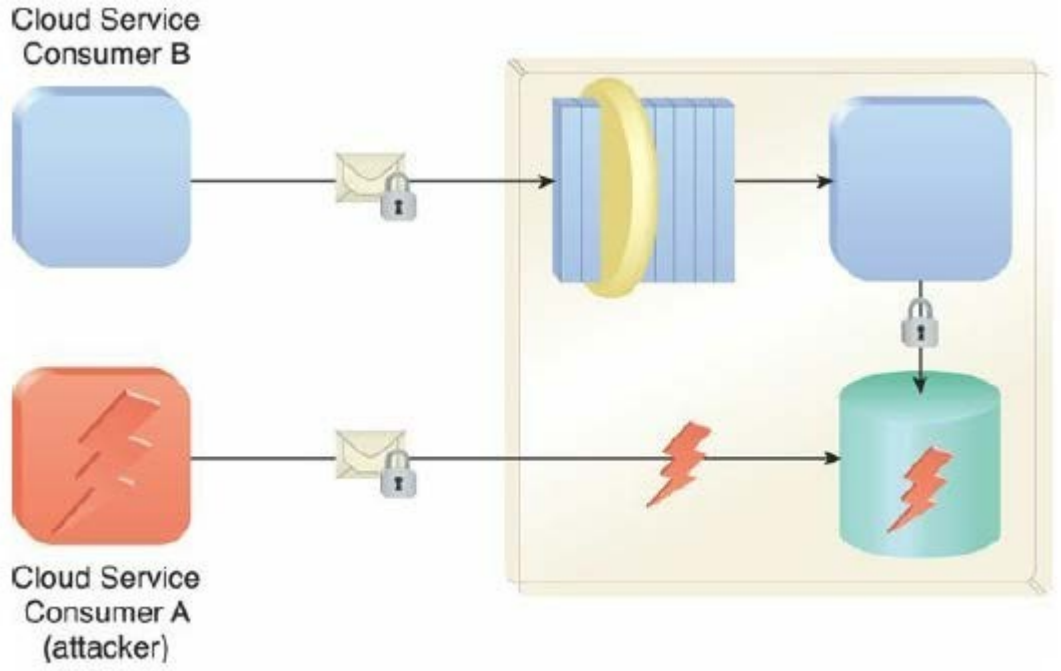
Denial of Service



Insufficient Authorization

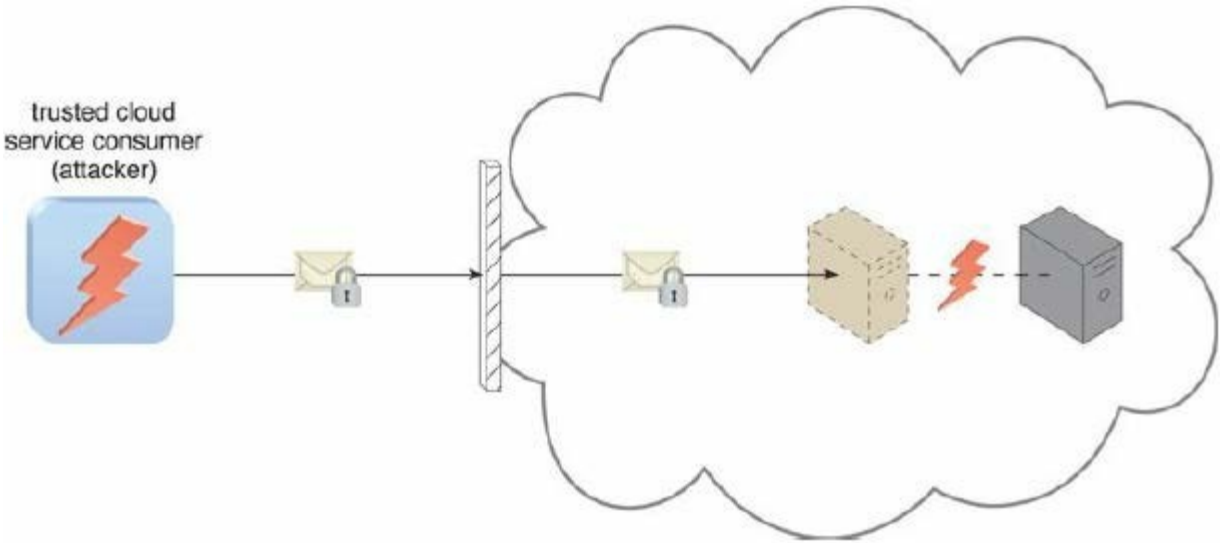
Occurs when access is granted to an attacker erroneously or too broadly to IT resources that are normally protected.

- Result of the attacker gaining direct access to IT resources that were implemented to be accessed by trusted consumer programs.



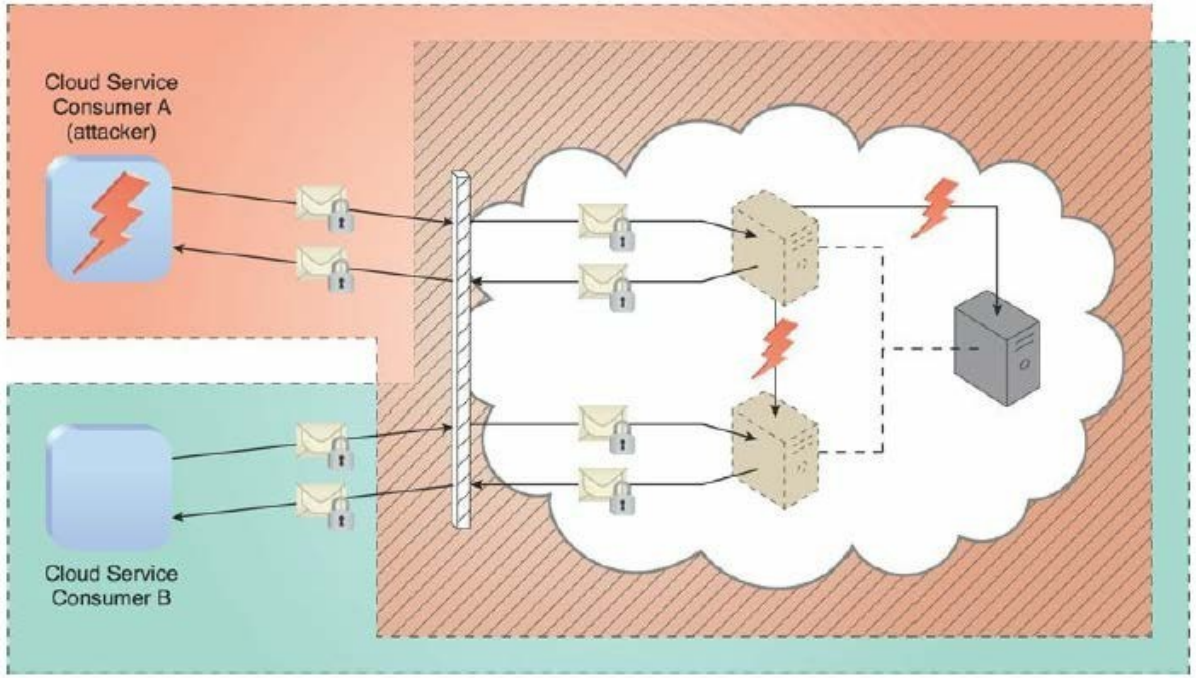
Virtualization Attack

Exploits vulnerabilities in the virtualization platform to jeopardize its confidentiality, integrity, and/or availability.



Overlapping Trust Boundaries

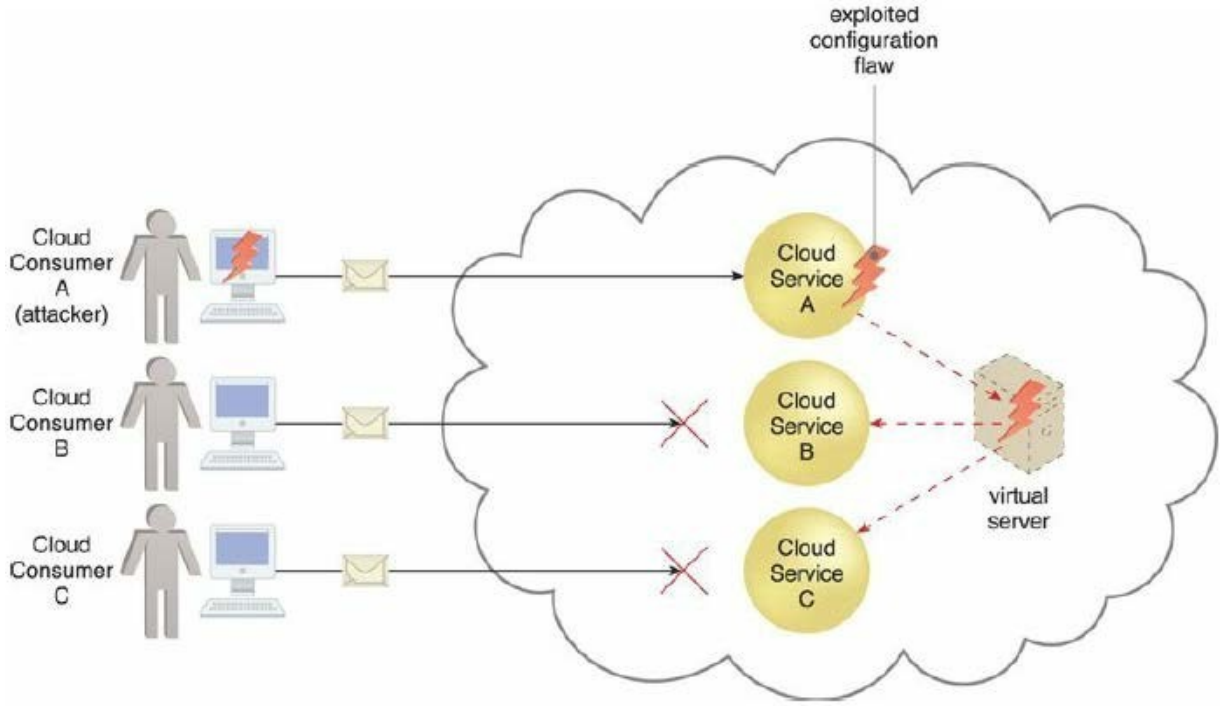
Target shared IT resources with the intention of compromising cloud consumers or other IT resources that share the same trust boundary



Checklist on Cloud Security

Substandard design, implementation or configuration

- Security issues may arise beyond runtime exception or system failures
- Attackers may exploit these vulnerabilities



Checklist on Cloud Security (Cont)

Security policy checks

- Check for any disparity as majority of the IT resources are now managed by cloud service providers

Contracts

- Examine contracts and SLAs put forth by cloud providers to ensure that security policies, and other relevant guarantees, are satisfactory when it comes to security

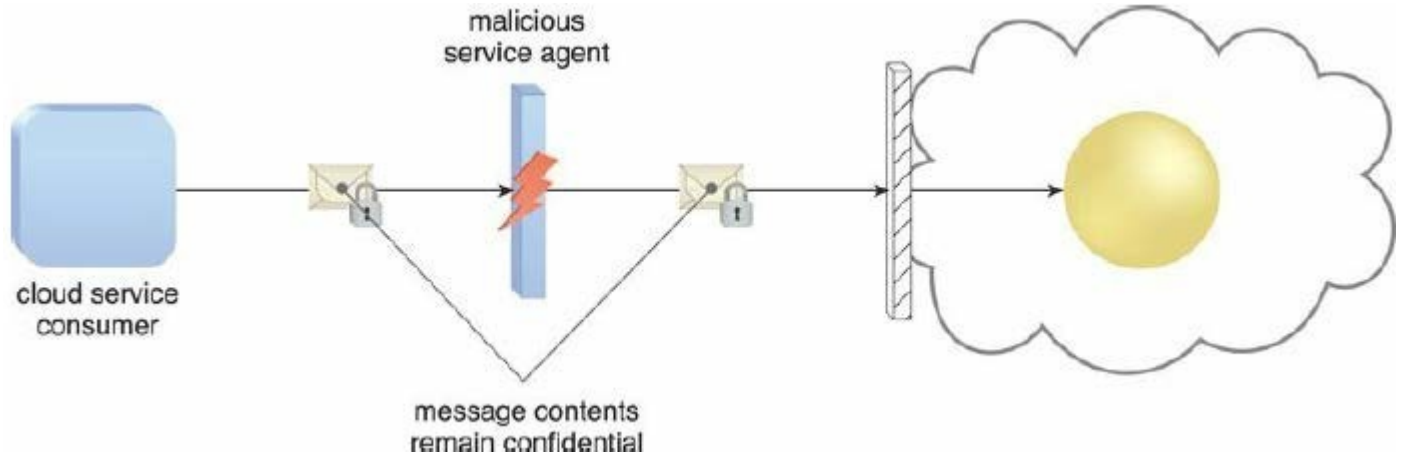
Risk Management

- Continuous risk assessment as part of risk management strategy

Cloud Security Countermeasures

Encryption

Secret key based encryption mechanisms to counter traffic eavesdropping, malicious intermediary, insufficient authorization, and overlapping trust boundaries security threats.



Encryption

Symmetric Encryption

- Secret key cryptography, uses the same key for both encryption and decryption
- Does not have the characteristic of non-repudiation, cannot determining which party performed the message encryption or decryption

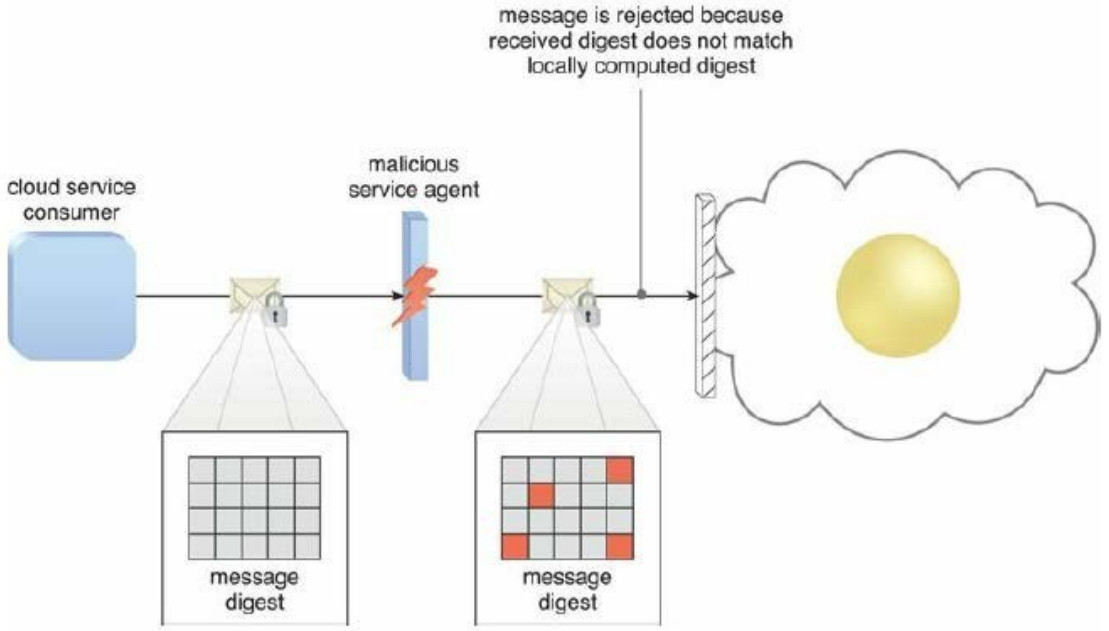
Asymmetric Encryption

- Relies on the use of two different keys, private key and a public key
- private key is known only to its owner while the public key is commonly available
- Doesn’t provide message integrity or authenticity protection due to the communal nature of the public key

Hashing

When non-reversible form of data protection is required

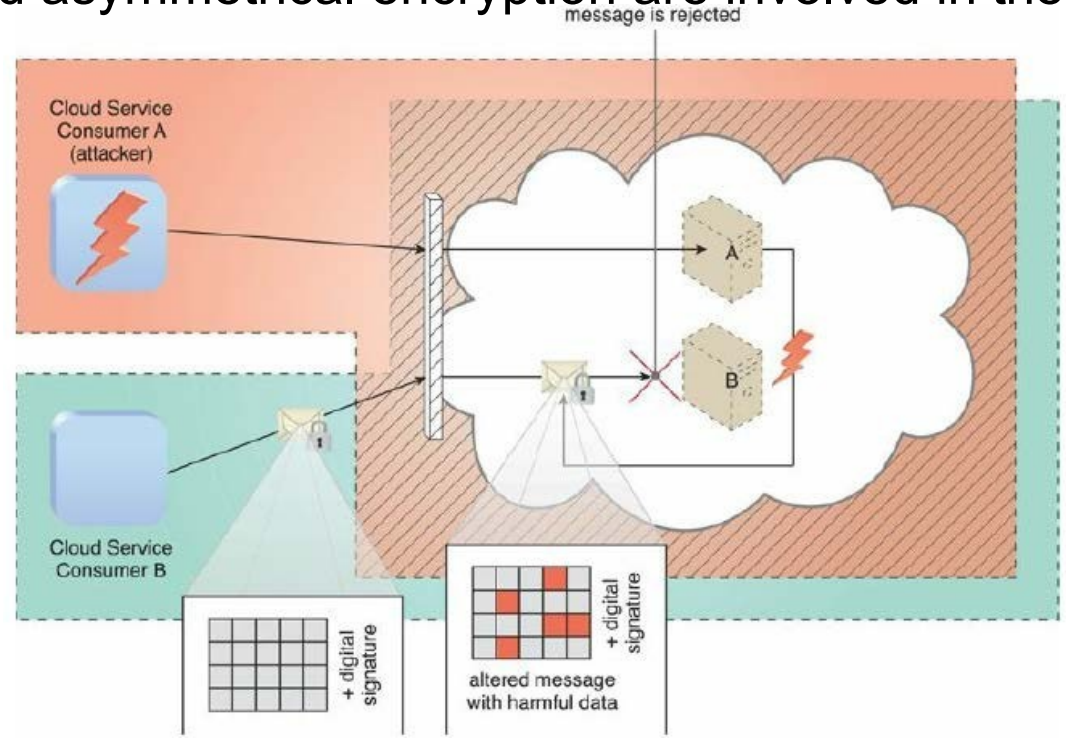
- Derive a hashing code or *message digest* from a message
- Sender attach message digest to the message
- Recipient applies the same hash function to the message to verify that the produced message digest is identical to the one that accompanied the message



Digital Signature

Messages are assigned a digital signature prior to transmission, which is then rendered invalid if the message experiences any unauthorized modifications

- Both hashing and asymmetrical encryption are involved in the creation of a digital signature



Public Key Infrastructure (PKI)

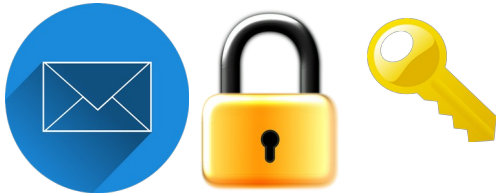
- Used to associate public keys with their corresponding key owners
- Rely on the use of digital certificates, which are digitally signed data structures that bind public keys to certificate owner identities with a validity time period
- Digital certificates are usually digitally signed by a third- party certificate authority

DECRYPTION

Request

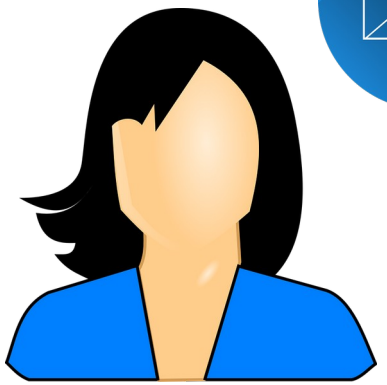
Sends her public key to encrypt the message

Encrypts and sends the Secure message



Public Key

Private Key

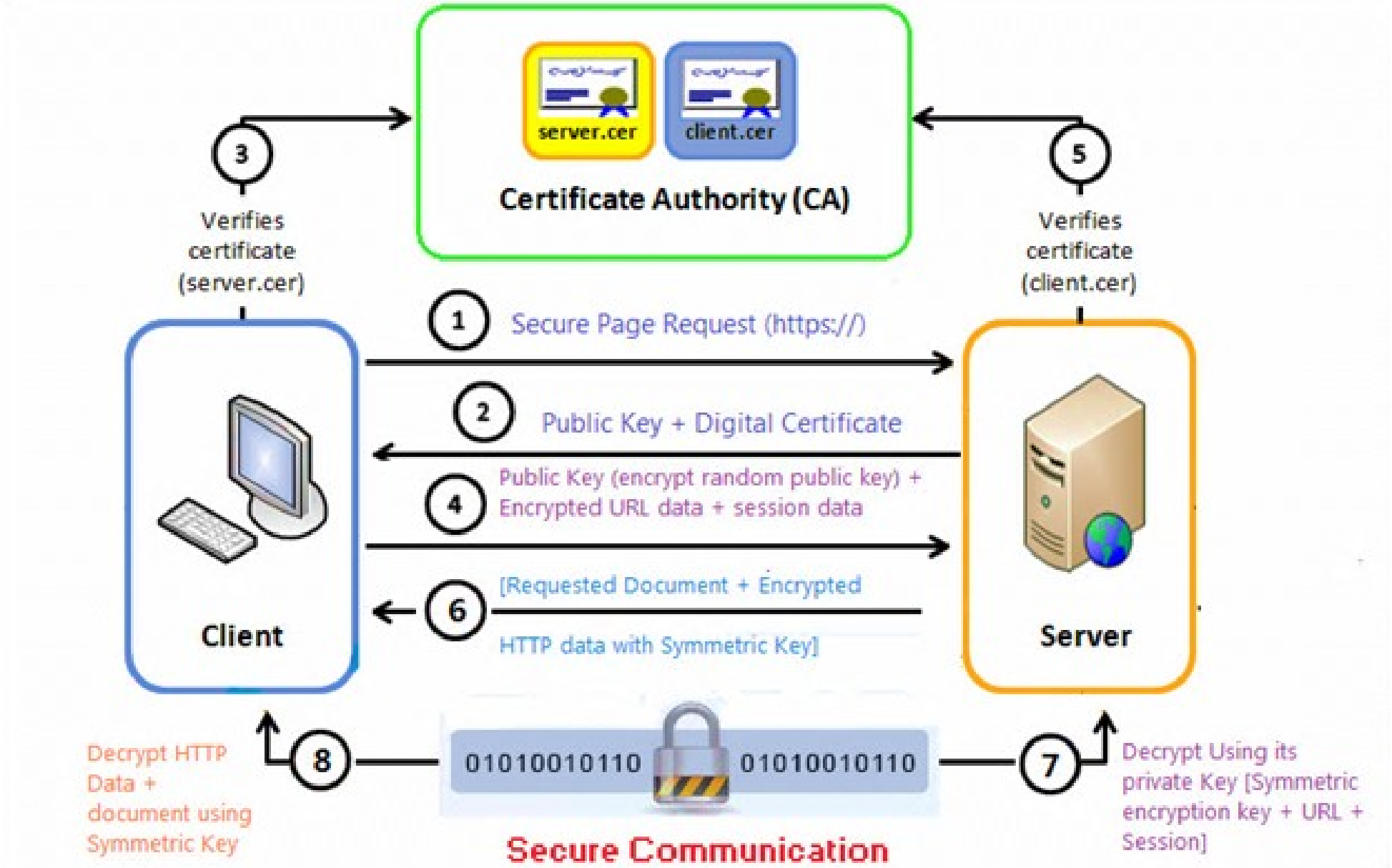


Public Key

Can be distributed to anyone

Private Key

Will be always private


server.cer
client.cer
Certificate Authority (CA)
3
Verifies certificate (server.cer)
5
Verifies certificate (client.cer)
1
Secure Page Request (https://)
2
Public Key + Digital Certificate
4
Public Key (encrypt random public key) + Encrypted URL data + session data
6
[Requested Document + Encrypted HTTP data with Symmetric Key]
Client
Server
8
Decrypt HTTP Data + document using Symmetric Key
01010010110
01010010110
7
Decrypt Using its private Key [Symmetric encryption key + URL + Session]
Secure Communication

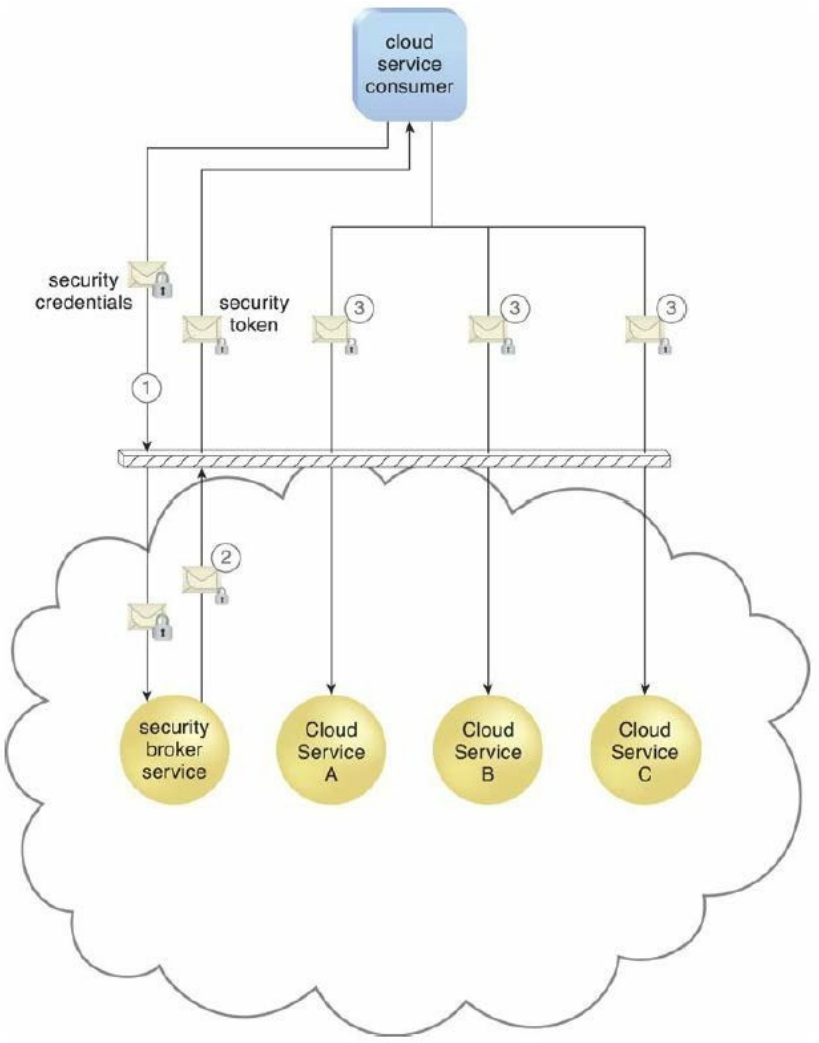
Identity Access Management (IAM)

- IAM encompasses controlling and tracking user identities and access in IT environments.
  - **Authentication**: Manages credentials like usernames, passwords, digital signatures, biometric data, and binds accounts to specific hardware or software identifiers.
  - **Authorization**: Defines access control levels and manages relationships between user identities, access rights, and resource availability.
  - **User Management**: Involves administrative tasks like creating user accounts, resetting passwords, setting password policies, and managing privileges.
  - **Credential Management**: Establishes and manages access rules for user accounts to prevent unauthorized access.

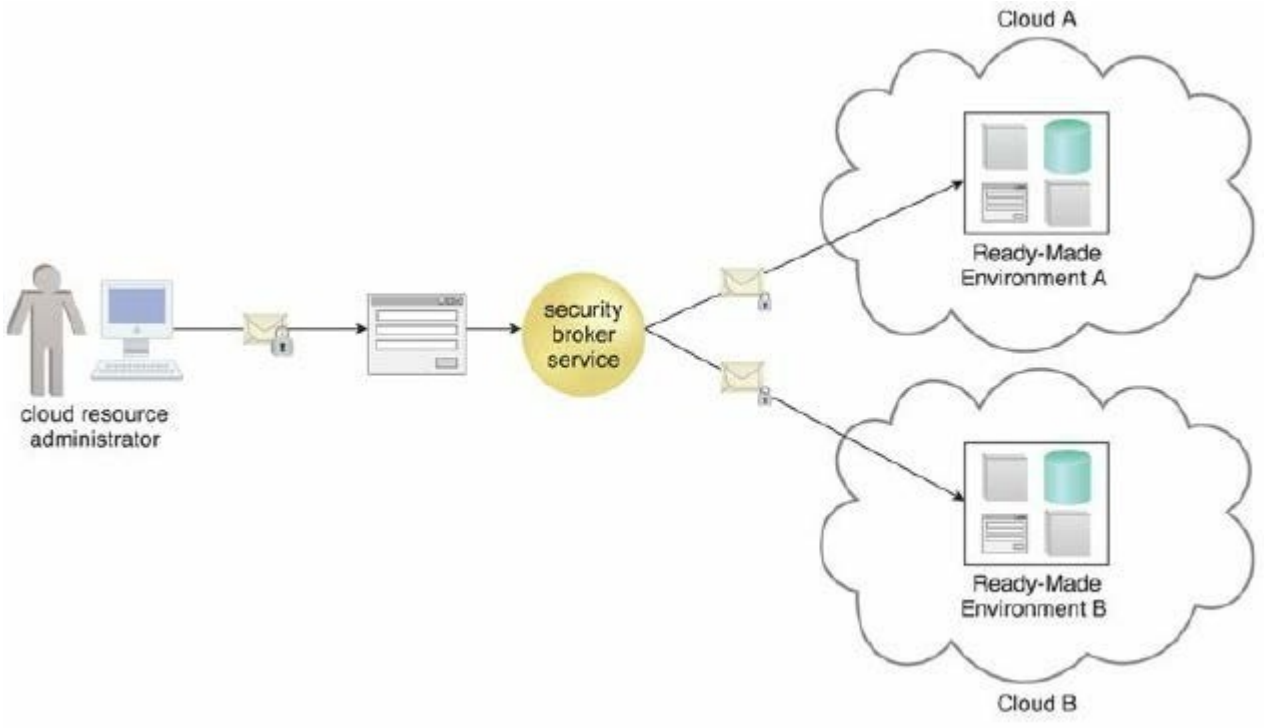
Single Sign-On (SSO)

Propagating the authentication information across multiple cloud services can be a challenging

- Enables authentication by a security broker that establish a persisted security context during consumer accesses to cloud services

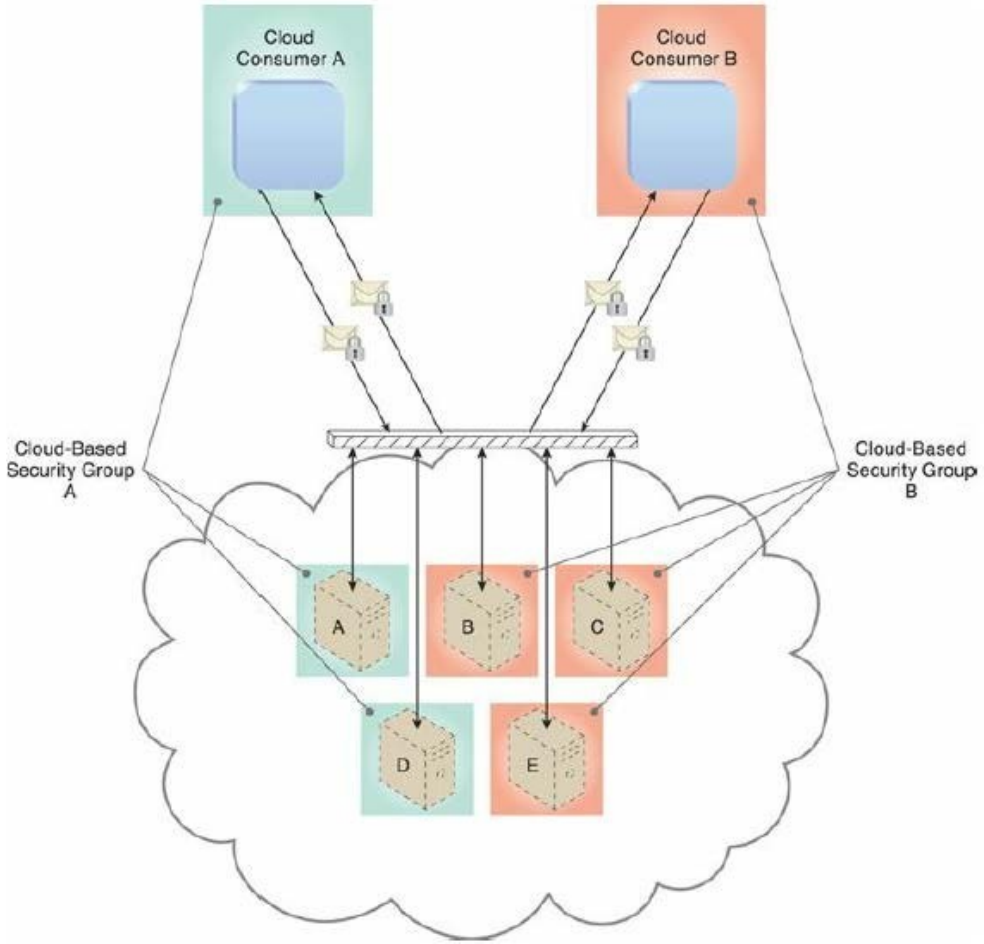


Single Sign-On (SSO)



Cloud-Based Security Groups

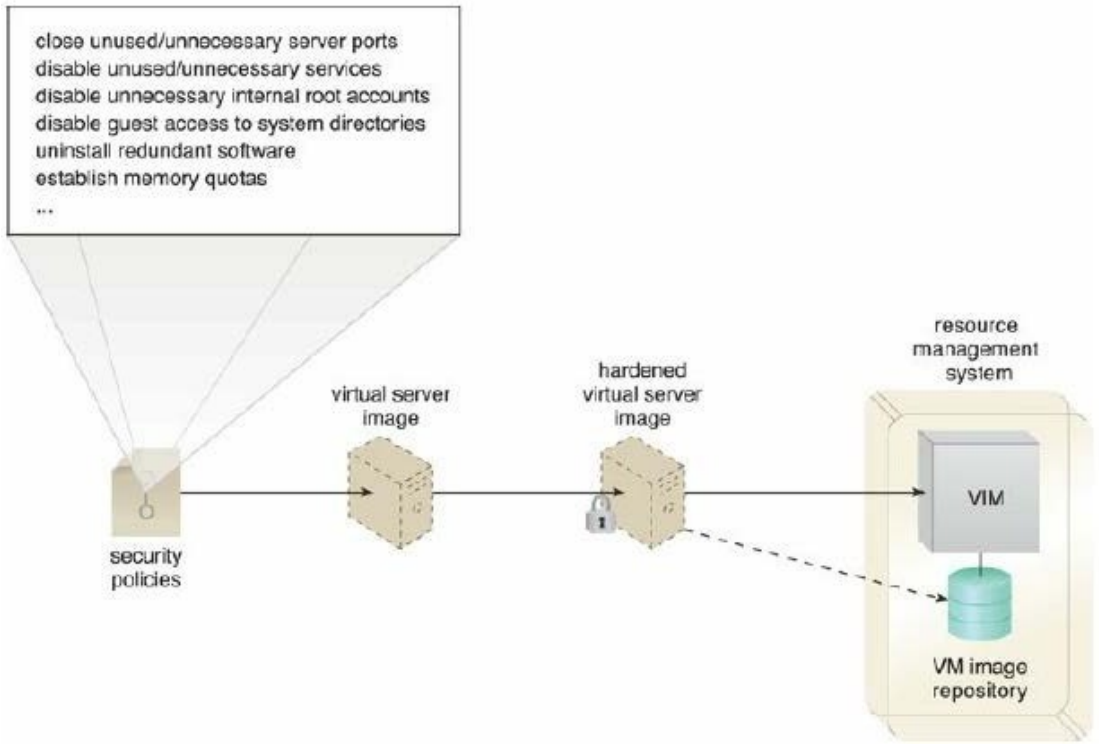
- Cloud resource segmentation is a process by which separate physical and virtual IT environments are created for different users and groups



Hardened Virtual Server Images

Process of stripping unnecessary software to limit potential vulnerabilities that can be exploited by attackers

- Removing redundant programs, closing unnecessary server ports, and disabling unused services, internal root accounts, and guest access



References

- https://aws.amazon.com/compliance/shared-responsibility-model/
- Chapter 6 and 10, Cloud Computing: Concepts, Technology & Architecture, Thomas Erl, et al., Prentice- Hall, 2013