UNIT – IV

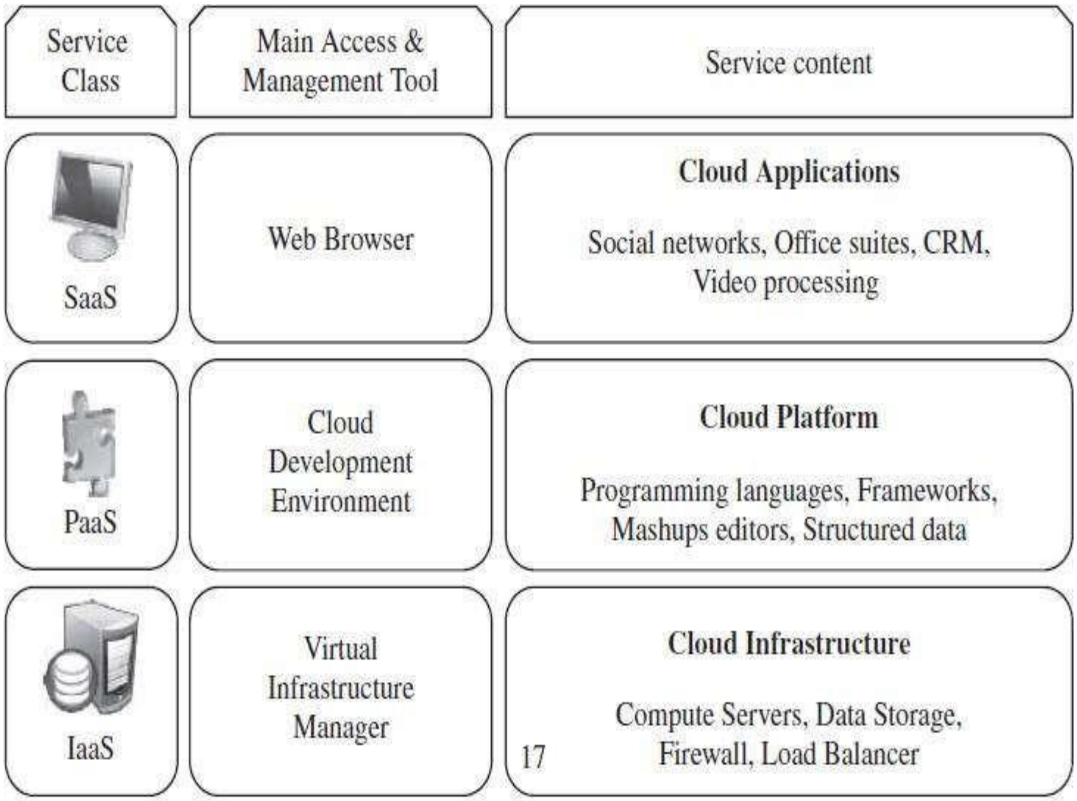
Software as a Service (SAAS) & Data Security in the Cloud: Software as a Service SAAS), Google App Engine – Centralizing Email Communications- Collaborating via Web- Based Communication Tools-An Introduction to the idea of Data Security.

Software as a Service (SAAS) & Data Security in the Cloud

Software as a Service(SAAS)

Traditional desktop applications such as word processing and spreadsheet can now be accessed as a service in the Web. This model of delivering applications, knownas Software as a Service (SaaS), alleviates the burden of software maintenance for customers and simplifies development and testing for providers.

Salesforce.com, which relies on the SaaS model, offers business productivity applications (CRM) that reside completely on their servers, allowing customers to customize and access applications on demand.



Service Class	Main Access & Management Tool	Service content
SaaS	Web Browser	**Cloud Applications** Social networks, Office suites, CRM, Video processing
PaaS	Cloud Development Environment	**Cloud Platform** Programming languages, Frameworks, Mashups editors, Structured data
IaaS	Virtual Infrastructure Manager	**Cloud Infrastructure** Compute Servers, Data Storage, Firewall, Load Balancer

Deployment Models

Although cloud computing has emerged mainly from the appearance of public computing utilities, other deployment models, with variations in physical location and distribution, have been adopted. In this sense, regardless of its service class, a cloud can be classified as public, private, community, or hybrid based on model of deployment as shown figure below.

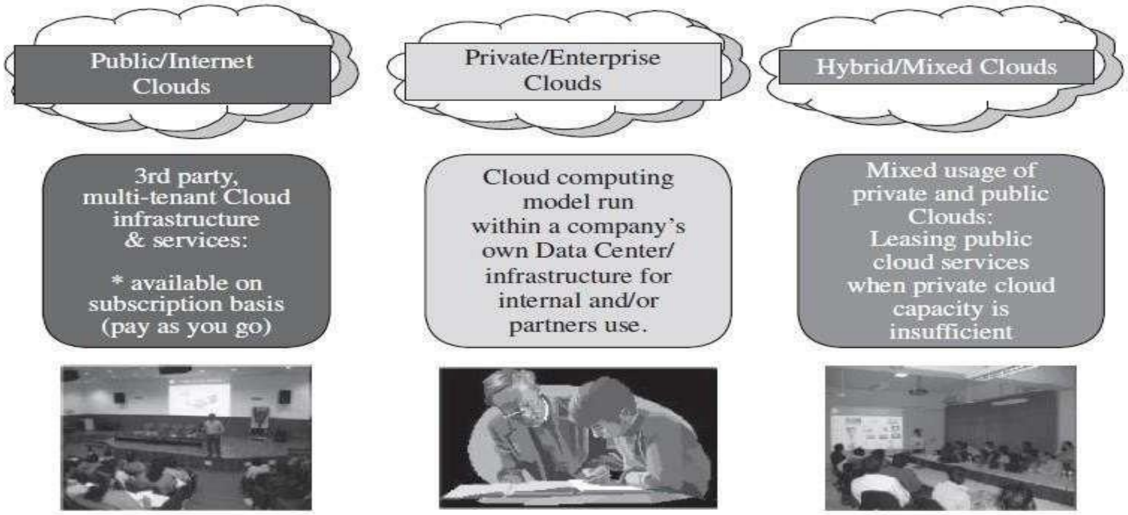

Public/Internet Clouds
Private/Enterprise Clouds
Hybrid/Mixed Clouds
3rd party, multi-tenant Cloud infrastructure & services:
* available on subscription basis (pay as you go)
Cloud computing model run within a company's own Data Center/ infrastructure for internal and/or partners use.
Mixed usage of private and public Clouds: Leasing public cloud services when private cloud capacity is insufficient

FIGURE 1.4. Types of clouds based on deployment models.

Public cloud & Private cloud:

Public cloud as a "cloud made available in a pay-as-you-go manner to the general public". **Private cloud** as "internal data center of a business or other organization, notmade available to the general public."

In most cases, establishing a private cloud means restructuring an existing infrastructure by adding virtualization and cloud-like interfaces. This allows users to interact with the local data center while experiencing the same advantages of public clouds, most notably self-service interface, privileged access to virtual servers, and per-usage metering and billing.

A **community cloud** is "shared by several organizations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations)."

A **hybrid cloud** takes shape when a private cloud is supplemented with computing capacity from public clouds. The approach of temporarily renting capacityto handle spikes in load is known as "cloud-bursting"

DESIRED FEATURES OF A CLOUD

Certain features of a cloud are essential to enable services that truly represent the cloud computing model and satisfy expectations of consumers, and cloud offerings must be having following features:

1. Self-service

2. Per-usage metered and billed

3. Elastic,
4. Customizable.

- **Self-Service**

  Consumers of cloud computing services expect on-demand, nearly instant access to resources. To support this expectation, clouds must allow self-service access so that customers can request, customize, pay, and use services without intervention of human operators.

- **Per-Usage Metering and Billing**

  Cloud computing eliminates up-front commitment by users, allowing them torequest and use only the necessary amount. Services must be priced on a short- term basis (e.g., by the hour), allowing users to release (and not pay for) resources as soon as they are not needed. For these reasons, clouds must implement features to allow efficient trading of service such as pricing, accounting, and billing. Metering should be done accordingly for different types of service (e.g., storage, processing, and bandwidth) and usage promptly reported, thus providing greater transparency.

- **Elasticity**

  Cloud computing gives the illusion of infinite computing resources available on demand. Therefore, users expect clouds to rapidly provide resources in any quantity at any time. In particular, it is expected that the additional resources can be (a) provisioned, possibly automatically, when an application load increases and (b) released when load decreases (scale up and down).

- **Customization**

  In a multi-tenant cloud a great disparity between user needs is often the case. Thus, resources rented from the cloud must be highly customizable. In the case of infrastructure services, customization means allowing users to deploy specialized virtual appliances and to be given privileged (root) access to the virtual servers. Other service classes (PaaS and SaaS) offer less flexibility and are not suitable for general- purpose computing, but still are expected to provide a certain level of customization.

Google APP Engine

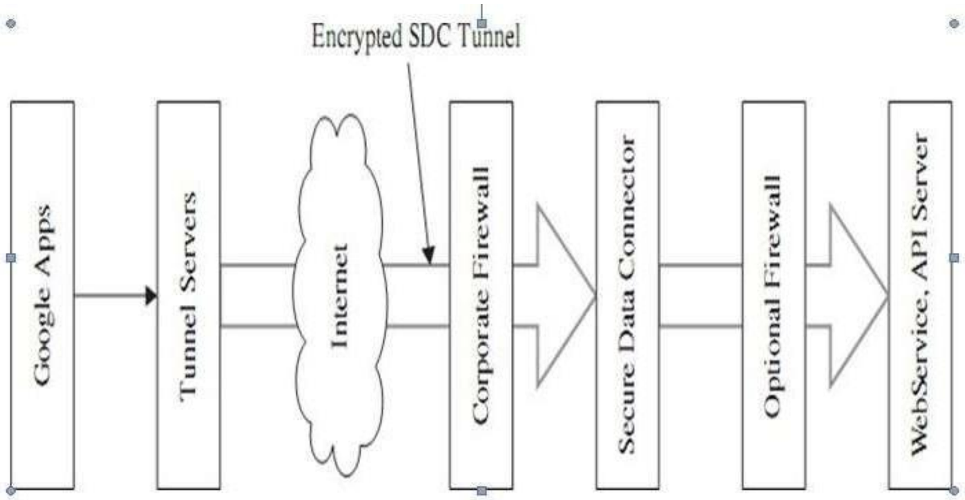
1. The app engine is a Cloud-based platform, is quite comprehensive and combines infrastructure as a service (IaaS), platform as a service (PaaS) and software as a service (SaaS).The app engine supports the delivery, testing and development of software on demand in a Cloud computing environment that supports millions of users and is highly scalable.
2. The company extends its platform and infrastructure to the Cloud through its app engine. It presents the platform to those who want to develop SaaS solutions at competitive costs.

3. It is a platform for hosting web applications in Google managed data centers. It is cloud-computing technology which virtualizes applications across multiple servers and data centers. Running your web application in Google infrastructure and Support different runtime

environments

Java (JRE 6 with limitation, Servlet 2.5, JDO,JPA)Python
(2.5.2)

1. Apps run in sandbox.
2. Automatic scaling and load balancing
3. No server restart, no network issues



1. The SDC constructs an encrypted connection between the data source and Google Apps. As long as the data source is in the Google Apps domain to the Google tunnel protocol servers, when the user wants to get the data, he/she will first send an authorized data requests to Google Apps, which forwards the request to the tunnel server.
2. The tunnel servers validate the request identity. If the identity is valid, the tunnel protocol allows the SDC to set up a connection, authenticate, and encrypt the data that flows across the Internet. At the same time, the SDC uses resource rules to validate whether a user is authorized to access a specified resource.
3. When the request is valid, the SDC performs a network request. The server validates the signed request, checks the credentials, and returns the data if the user is authorized From the perspective of cloud storage services, data integrity depends on the security of operations while in storage in addition to the security of the uploading and downloading sessions. The uploading session can only ensure that the data receivedby the cloud storage is the data that the user uploaded; The downloading session can guarantee the data that the user retrieved is the data cloud storage recorded. Unfortunately, this procedure applied on cloud storage services cannot guarantee data integrity. First, assume that Alice, a company CFO, stores the company financial data at a cloud storage service provided by Eve. And then Bob, the company administration chairman, downloads the data from the cloud.

There are three important concerns in this simple procedure:

1. **Confidentiality.** Eve is considered as an untrustworthy third party, Alice and Bob do not want reveal the data to Eve.
2. **Integrity.** As the administrator of the storage service, Eve has the capability to play with the

data in hand. How can Bob be confident that the data he fetched from Eve are the same as what was sent by Alice? Are there any measures to guarantee that the data have not been tampered by Eve?

3. **Repudiation.** If Bob finds that the data have been tampered with, is there any evidence for him to demonstrate that it is Eve who should be responsible for the fault? Similarly, Eve also needs certain evidence to prove her innocence.

GoogleAPP Engine Solutions for Missing Link

1. Third Authority Certified(TAC)
2. Secret Key Sharing(SKS)Four

Solutions

1. Neither TAC nor SKS
2. With SKS but without TAC
3. With TAC but without SKS O With Both TAC and SKS

Google is a leader in web-based applications,

Google is a leader in web-based applications and leading searching engine in the world. soit’s not surprising that the company also offers cloud development services.

1. These services come in the form of the Google App Engine, which enables developers to build their own web applications utilizing the same infrastructure that powers Google’s powerful applications.
2. The Google App Engine provides a fully integrated application environment. Using Google’s development tools and computing cloud, App Engine applications are easy to build, easy to maintain, and easy to scale. All you haveto do

Features of App Engine

1. These are covered by the depreciation policy and the service-level agreement of the app engine. Any changes made to such a feature are backward-compatible and implementation of such a feature is usually stable. These include data storage, retrieval, and search; communications; process management; computation; app configuration and management.
2. Data storage, retrieval, and search include features such as HRD migration tool, Google Cloud SQL, logs, datastore, dedicated Memcache, blobstore, Memcache and search.
3. Communications include features such as XMPP. channel, URL fetch, mail, and Google Cloud Endpoints.
4. Process management includes features like scheduled tasks and task queue. Computation includes images.

5. App management and configuration cover app identity, users, capabilities, traffic splitting, modules, SSL for custom domains, modules, remote access, and multi- tenancy.

Centralizing email Communications

1. The key here is to enable anywhere/anytime access to email. Precloud computing, your email access was via a single computer, which also stored all your email messages.
2. For this purpose, you probably used a program like Microsoft Outlook or Outlook Express, installed on your home computer.
3. To check your home email from work, it took a bit of juggling and perhaps the use of your ISP's email access web page. That web page was never in sync with the messages on your home PC, of course, which is just the start of the problemswith trying to communicate in this fashion.
4. A better approach is to use a web-based email service, such as Google's Gmail (mail.google.com), Microsoft's Windows Live Hotmail (mail.live.com), or Yahoo! Mail (mail.yahoo.com).
5. These services place your email inbox in the cloud and you can access it from any computer connected to the Internet.

Collaborating via Web-Based Communication Tools-GMAIL

1. Gmail offers a few unique features that set it apart from the web-based email crowd.
2. First, Gmail doesn't use folders. With Gmail you can't organize your mail intofolders, as you can with the other services.
3. Instead, Gmail pushes the search paradigm as the way to find the messages youwant— not a surprise, given Google's search-centric business model.
4. Gmail does, however, let you "tag" each message with one or more labels. This has the effect of creating virtual folders, as you can search and sort your messages by any of their labels.
5. In addition, Gmail groups together related email messages in what Google calls conversationsYahoo! Mail Yahoo! Mail (mail.yahoo.com)
6. There is another web mail service, provided by the popular Yahoo! search site.
7. The basic Yahoo! Mail is free and can be accessed from any PC, using any webbrowser.
8. Yahoo! also offers a paid service called Yahoo! Mail Plus that lets you send larger messages and offers offline access to your messages via POP email clients.

Web Mail Services

1. AOL Mail (mail.aol.com)
2. BigString (www.bigstring.com)
3. Excite Mail (mail.excite.com)

4. FlashMail (www.flashmail.com)
5. GMX Mail (www.gmx.com)
6. Inbox.com (www.inbox.com)
7. Lycos Mail (mail.lycos.com)
8. Mail.com (www.mail.com)
9. Zoho Mail (zoho.mail.com)

An Introduction to Data Security:

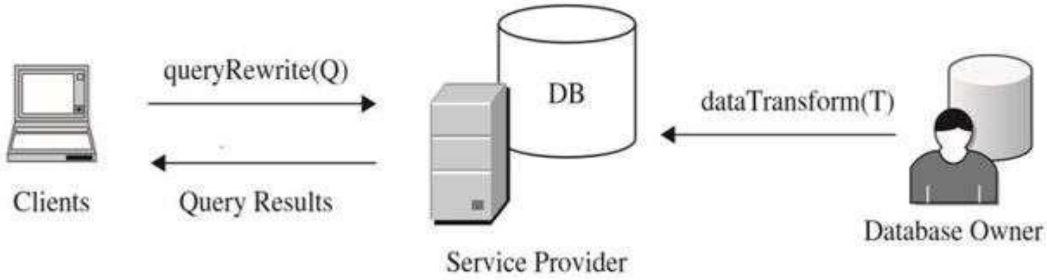
Data Security

1. Data Security defines as Information in a cloud environment has much more dynamism and fluidity than information that is static on a desktop or in a network folder
2. Nature of cloud computing dictates that data are fluid objects, accessible froma multitude of nodes and geographic locations and, as such, must have a datasecurity methodology that takes this into account while ensuring that this fluidity is not compromised.
3. The idea of content-centric or information-centric protection, being an inherent part ofa data object is a development out of the idea of the “de-perimerization” of the enterprise.
4. This idea was put forward by a group of Chief Information Officers (CIOs) who formed an organization called the Jericho Forum

TECHNOLOGIES FOR DATA SECURITY IN CLOUD COMPUTING

Unique issues of the cloud data storage platform from a few different perspectives

1. Database Outsourcing and Query Integrity Assurance
   I. Storing data into and fetching data from devices and machines behind a cloud are essentiallya novel form of database outsourcing



2. Data Integrity in Untrustworthy Storage
   I. The fear of losing data or data corruption

Relieve the users’ fear by providing technologies that enable users to check the integrity of their data

3. Web-Application-Based Security
   - I. Once the dataset is stored remotely, a Web browser is one of the most convenient approaches that end users can use to access their data on remoteservices
   - II. Web security plays a more important role for cloud computing

1. Multimedia Data Security
   1. With the development of high-speed network technologies and large bandwidth connections, more and more multimedia data are being storedand shared in cyber space
   2. The security requirements for video, audio, pictures, or images are different from other applications

CLOUD COMPUTING AND IDENTITY

Digital identity

1. Digital identity holds the key to flexible data security within a cloud Environment
2. A digital identity represents who we are and how we interact with others on-line.
3. **Access, identity, and risk** are three variables that can become inherently connected when applied to the security of data, because access and risk are directly proportional: As access increases, so then risk to the security of the data increases.
4. Access controlled by identifying the actor attempting the access is the most logical manner of performing this operation.
5. Ultimately, digital identity holds the key to securing data, if that digital identity can be programmatically linked to security policies controlling the post-access usage of data.

Identity, Reputation, and Trust

1. Reputation is a real-world commodity; that is a basic requirement of human-to-human relationships
2. Our basic societal communication structure is built upon the idea of reputation andtrust.
3. Reputation and its counter value, trust, is easily transferable to a digital realm:
4. EBay, for example, having partly built a successful business model on the strength of a ratings system, builds up the reputation of its buyers and sellers through successful (or unsuccessful) transactions.
5. These types of reputation systems can be extremely useful when used with a digital identity.
6. They can be used to associate varying levels of trust with that identity, which in turn can be used to define the level (granular variations) of security policy applied to data resources that the individual wishes to access

User-Centric Identity:

1. Digital identities are a mechanism for identifying an individual, particularly within a cloud environment and identity ownership being placed upon the individual is known as user-centric identity
2. It allows users to consent and control how their identity (and the individual identifiers making up the identity, the claims) is used.
3. This reversal of ownership away from centrally managed identity platforms (enterprise-centric) has many advantages.
4. This includes the potential to improve the privacy aspects of a digital identity, by giving an individual the ability to apply permission policies based on their identity and to control which aspects of that identity are divulged
5. An identity may be controllable by the end user, to the extent that the user can then decide what information is given to the party relying on the identity

Information Card:

- Information cards permit a user to present to a Web site or other service (relyingparty) one or more claims, in the form of a software token, which may be used to uniquely identify that user.
- They can be used in place of user name/ passwords, digital certificates, and other identification systems, when user identity needs to be established to control access toa Web site or other resource, or to permit digital signing
- Information cards are part of an identity meta-system consisting of:
  - **Identity providers (IdP**), who provision and manage information cards with specific claims, to users.
  - **Users** who own and utilize the cards to gain access to Web sites and other resources that support information cards.
  - **An identity selector/service**, which is a piece of software on the user's desktop or in the cloud that allows a user to select and manage their cards.
  - **Relying parties.** These are the applications, services & so on, that can use an information card to authenticate a person and to then authorize an action such as logging onto a Web site, accessing a document, signing content, and so on.
  - Each information card is associated with a set of claims which can be used to identify the user. These claims include identifiers such as name, email address post code.

Using Information Cards to Protect Data

1. Information cards are built around a set of open standards devised by a consortiumthat includes Microsoft, IBM, Novell, and so on.
2. The original remit of the cards was to create a type of single sign on system for the Internet, to help users to move away from the need to remember multiple passwords.

3. However, the information card system can be used in many more ways.
4. Because an information card is a type of digital identity, it can be used in the sameway that other digital identities can be used.
5. For example, an information card can be used to digitally sign data and content and to control access to data and content. One of the more sophisticated uses of an information card is the advantage given to the cards by way of the claims system.

Cloud Computing and Data Security Risk

1. Cloud computing is a development that is meant to allow more open accessibility andeasier and improved data sharing.
2. Data are uploaded into a cloud and stored in a data center, for access by users from that data center; or in a more fully cloud-based model, the data themselves are createdin the cloud and stored and accessed from the cloud (again via a data center).
3. The most obvious risk in this scenario is that associated with the storage of that data. A user uploading or creating cloud-based data include those data that are stored and maintained by a third-party cloud provider such as Google, Amazon, Microsoft, and so on.

This action has several risks associated with it:

i. Firstly, it is necessary to protect the data during upload into the data center toensure that the data do not get hijacked on the way into the database.

ii. Secondly, it is necessary to the stores the data in the data center to ensure that theyare encrypted at all times.

iii. Thirdly, and perhaps less obvious, the access to those data need to be controlled;this control should also be applied to the hosting company, including the administrators of the data center.

iv. In addition, an area often forgotten in the application of security to a data resourceis the protection of that resource during its use

Data security risks are compounded by the open nature of cloud computing.

1. Access control becomes a much more fundamental issue in cloud-based systems because of the accessibility of the data
2. Information-centric access control (as opposed to access control lists) can help to balance improved accessibility with risk, by associating access rules with differentdata objects within an open and accessible platform, without losing the Inherent usability of that platform
3. A further area of risk associated not only with cloud computing, but also with traditional network computing, is the use of content after access.

4. The risk is potentially higher in a cloud network, for the simple reason that the information is outside of your corporate walls

Data-centric mashups

1. that are used to perform business processes around data creation and dissemination—by their very nature, can be used to hijack data, leaking sensitive information and/or affecting integrity of that data

2. Cloud computing, more than any other form of digital communication technology, has created a need to ensure that protection is applied at the inception of the information, in a content centric manner, ensuring that a security policy becomes anintegral part of that data throughout its life cycle.

Encryption

1. It is a vital component of the protection policy, but further controls over the access ofthat data and on the use of the data must be met.

2. In the case of mashups, the controlling of access to data resources, can help toalleviatethe security concerns by ensuring that mashup access is authenticated.

3. Linking security policies, as applied to the use of content, to the access control method offer a way of continuing protection of data, post access and throughout thelife cycle; this type of data security philosophy must be incorporated into the use of cloud computing to alleviate security risks.

- **Identity providers (IdP**), who provision and manage information cards with specific claims, to users.
- **Users** who own and utilize the cards to gain access to Web sites and other resources that support information cards.
- **An identity selector/service**, which is a piece of software on the user's desktop or in the cloud that allows a user to select and manage their cards.
- **Relying parties.** These are the applications, services & so on, that can use an information card to authenticate a person and to then authorize an action such as logging onto a Web site, accessing a document, signing content, and so on.
- Each information card is associated with a set of claims which can be used to identify the user. These claims include identifiers such as name, email address post code.