

UNIT - IV

4

Cloud Platforms and Cloud Applications

Syllabus

Amazon Web Services (AWS) : Amazon Web Services and Components, Amazon Simple DB, Elastic Cloud Computing (EC2), Amazon Storage System, Amazon Database services (Dynamo DB). **Microsoft Cloud Services :** Azure core concepts, SQL Azure, Windows Azure Platform Appliance. **Cloud Computing Applications :** Healthcare : ECG Analysis in the Cloud, Biology : Protein Structure Prediction, Geosciences : Satellite Image Processing, Business and Consumer Applications : CRM and ERP, Social Networking, Google Cloud Application : Google App Engine. Overview of OpenStack architecture.

Contents

- | | | | |
|-----|--|--------------------------|---------|
| 4.1 | Amazon Web Services | Dec.-19, | Marks 8 |
| 4.2 | Elastic Cloud Computing | June-19, Dec.-19, | Marks 9 |
| 4.3 | Amazon Storage System | | |
| 4.4 | Amazon Database Services | | |
| 4.5 | Microsoft Cloud Services : Azure | | |
| 4.6 | Cloud Computing Applications | | |
| 4.7 | Google Cloud Application : Google App Engine | | |
| 4.8 | Overview of OpenStack Architecture | | |
| 4.9 | Multiple Choice Questions | | |

4.1 Amazon Web Services

SPPU : Dec.-19

- Amazon Web Services (AWS) is a cloud computing platform from Amazon that provides customers with a wide array of cloud services.
- Amazon first debuted its Amazon Web Services in 2006 as a way to enable the use of online services by client-side applications or other web sites via HTTP, REST or SOAP protocols.
- Amazon bills customers for Amazon AWS based on their usage of the various Amazon Web Services.
- In 2012, Amazon launched the AWS Marketplace to accommodate and grow the emerging ecosystem of AWS offerings from third-party providers that have built their own solutions on top of the Amazon Web Services platform.
- The AWS Marketplace is an online store for Amazon Web Services customers to find, compare and begin using AWS software and technical services.
- Amazon Web Services is a secure cloud services platform, offering compute power, database storage, content delivery and other functionality to help businesses scale and grow.
- In 2017, AWS comprised more than 90 services spanning a wide range including computing, storage, networking, database, analytics, application services, deployment, management, mobile, developer tools, and tools for the Internet of Things.
- Today, Amazon Web Services provides a highly reliable, scalable, low-cost infrastructure platform in the cloud that powers hundreds of thousands of businesses in 190 countries around the world.
- In 2016 AWS partnered with Digital Currency Group to create a laboratory environment allowing companies to experiment with block chain technologies.
- In January 2018, Amazon launched an autoscaling service on AWS.

What is Amazon Web Services ?

- Amazon Web Services (AWS) is a collection of remote computing services (web services) that together make up a cloud computing platform, offered over the Internet by Amazon.com.
- The AWS Cloud infrastructure is built around Regions and Availability Zones (AZs). A Region is a physical location in the world where we have multiple AZs. AZs consist of one or more discrete data centers, each with redundant power, networking, and connectivity, housed in separate facilities.

- These AZs offer you the ability to operate production applications and databases that are more highly available, fault tolerant, and scalable than would be possible from a single data center.
- The AWS cloud operates 42 AZs within 16 geographic regions around the world, with five more availability zones and two more regions coming online in 2017.
- Each availability zone is designed as an independent failure zone. This means that availability zones are physically separated within a typical metropolitan region and are located in lower risk flood plains.

4.1.1 Components

- AWS consists of many cloud services that you can use in combinations tailored to your business or organizational needs.
- With Amazon Web Services you will find a complete cloud platform ready to use for virtually any workload.
- The user requests to the server by the method such as e-mail either to register or to transfer the domain.
- Your request which includes all information will be sent to Amazon API Gateway restful service.
- API Gateway will transfer the collected user information to an AWS Lambda function.
- AWS Lambda function will generate an e-mail and forward it to the 3rd party mail server using Amazon SES.
- Components of Amazon Web Service architecture are Amazon API Gateway, AWS Lambda, Amazon Simple Email Service.
- API Gateway is a front-door to access data, business logic and functionality. API Gateway will provide a restful API endpoint for our AWS Lambda function.
- API works at small as well as large-scale and helps developers to manage, spectate, create and provide security to the APIs.

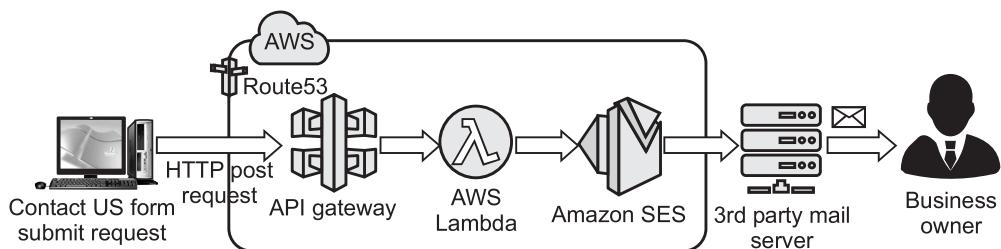


Fig. 4.1.1 AWS

- AWS Lambda is a compute service that runs your back-end code and responds to events such as object uploads to Amazon S3 bucket, Dynamo DB or in-app activity. The Lambda function will get all the information from a user through API Gateway.
- Amazon Simple email service helps us to send e-mail with minimal setup and maximum deliverability. It is integrated with AWS management console so that you can monitor your sending activity. Amazon Simple Email Service helps us by monitoring insecurity.

4.1.2 Advantages and Disadvantages of AWS

Advantages :

1. Easy to use.
2. No capacity limits : Organizations launch different projects and the guess what capacity they will need.
3. Provides speed and agility.
4. Secure and reliable : AWS provides security and also helps to protect the privacy as it is stored in AWS data centers.

Disadvantages :

1. Limitations of Amazon EC2 : AWS sets default limits on resources which vary from region to region. These resources consist of images, volumes, and snapshots.
2. Technical support fee : AWS charges you for immediate support.
3. Security Limitations.

Review Question

1. Write a note on services offered by Amazon.

SPPU : Dec.-19, End Sem, Marks 8

4.2 Elastic Cloud Computing

SPPU : June-19, Dec.-19

- Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. It is designed to make web-scale computing easier for developers and system administrators.
- The Amazon EC2 simple web service interface allows you to obtain and configure capacity with minimal friction. It provides you with complete control of your computing resources and lets you run on Amazon's proven computing environment.

- Amazon EC2 reduces the time required to obtain and boot new server instances (called Amazon EC2 instances) to minutes, allowing you to quickly scale capacity, both up and down, as your computing requirements change.
- Amazon EC2 changes the economics of computing by allowing you to pay only for capacity that you actually use. Amazon EC2 provides developers and system administrators the tools to build failure resilient applications and isolate themselves from common failure scenarios.
- EC2 allows creating Virtual Machines (VM) on-demand. Pre-configured template Amazon Machine Image (AMI) can be used get running immediately. Creating and sharing your own AMI is also possible via the AWS marketplace.
- Amazon Machine Image (AMI) is a template for software configuration (Operating System, Application Server, and Applications). Fig. 4.2.1 shows AMI and instance.
- Instance is a AMI running on virtual servers in the cloud. Each instance type offers different compute and memory facilities. Create an Amazon Machine Image (AMI) containing your applications, libraries, data and associated configuration settings. Or use pre-configured, templated images to get up and running immediately.
- Auto scaling allows automatically scale of the capacity up seamlessly during demand spikes to maintain performance and scales down during demand lulls to minimize costs.
- Elastic load balancing automatically distributes incoming application traffic across multiple Amazon EC2 instances. It provide tools to build failure resilient applications by launching application instances in separate availability zones.
- Pay only for resources actually consume, instance-hours. VM Import/Export enables you to easily import virtual machine images from your existing environment to Amazon EC2 instances and export them back at any time.
- Boto is a Python package that provides programmatic connectivity to Amazon Web Services.

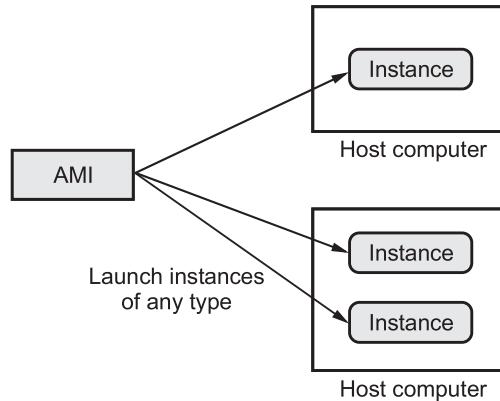


Fig. 4.2.1 AMI and instance

launching an EC2 instance :

```
#/usr/bin/python
import boto.ec2
```

```
conn = boto.ec2.connect_to_region("us-west-2")
conn.run_instances(
    'ami-6ac2a85a',
    key_name='nitheesh_oregon',
    instance_type='t1.micro',
    security_groups=['nitheesh_oregon']
)
```

Stop instances :

```
#!/usr/bin/python
import boto.ec2
conn = boto.ec2.connect_to_region("us-west-2")
conn.stop_instances(instance_ids=['instance-id-1',
'instance-id-2'])
```

- Boto supports more than fifty Amazon services, running the whole range from compute, database, application and payments and billing.
- EC2 functions :
 1. Load variety of operating system.
 2. Install custom applications.
 3. Manage network access permission.
 4. Run image using as many/few systems as you desire.
- EC2 advantages :
 1. Amazon EC2 enables you to increase or decrease capacity within minutes.
 2. User have complete control of your Amazon EC2 instances.
 3. Support flexible cloud hosting services
 4. Secure : Amazon EC2 works in conjunction with Amazon VPC to provide security and robust networking functionality.
 5. Reliable : Amazon EC2 offers a highly reliable environment where replacement instances can be rapidly and predictably commissioned.

4.2.1 Configuring Amazon EC2 Linux Instances

- Let's get started with Amazon Elastic Compute Cloud (Amazon EC2) by launching, connecting to and using a Linux instance. An instance is a virtual server in the AWS cloud. With Amazon EC2, you can setup and configure the operating system and applications that run on your instance.
- When you sign up for AWS, you can get started with Amazon EC2 using the AWS Free Tier.
- The instance is an Amazon EBS-backed instance (meaning that the root volume is an EBS volume). You can either specify the availability zone in which your

instance runs or let Amazon EC2 select an availability zone for you. When you launch your instance, you secure it by specifying a key pair and security group. When you connect to your instance, you must specify the private key of the key pair that you specified when launching your instance.

- Various steps to configure Amazon EC2 Linux instance is shown in Fig. 4.2.2.

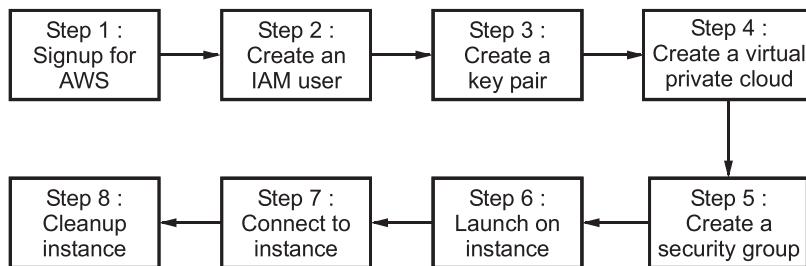


Fig. 4.2.2 Steps to signup for EC2

Step 1 : SignUp for AWS

- When you signup for Amazon Web Services (AWS), your AWS account is automatically signed up for all services in AWS, including Amazon EC2. You are charged only for the services that you use.
- With Amazon EC2, you pay only for what you use. If you are a new AWS customer, you can get started with Amazon EC2 for free.

Step 2 : Create an IAM user

- Services in AWS, such as Amazon EC2, require that you provide credentials when you access them, so that the service can determine whether you have permission to access its resources. The console requires your password.
- You can create access keys for your AWS account to access the command line interface or API. However, we don't recommend that you access AWS using the credentials for your AWS account; we recommend that you use AWS Identity and Access Management (IAM) instead.
- Create an IAM user and then add the user to an IAM group with administrative permissions or grant this user administrative permissions. You can then access AWS using a special URL and the credentials for the IAM user. If you signed up for AWS but have not created an IAM user for yourself, you can create one using the IAM console.

Step 3 : Create a key pair

- AWS uses public-key cryptography to secure the login information for your instance. A Linux instance has no password; you use a key pair to log in to your instance securely. You specify the name of the key pair when you launch your instance, then provide the private key when you log in using SSH.

- If you haven't created a key pair already, you can create one using the Amazon EC2 console. Note that if you plan to launch instances in multiple regions, you'll need to create a key pair in each region.

Step 4 : Create a Virtual Private Cloud (VPC)

- Amazon VPC enables you to launch AWS resources into a virtual network that you've defined, known as a Virtual Private Cloud (VPC). The newer EC2 instance types require that you launch your instances in a VPC. If you have a default VPC, you can skip this section and move to the next task, create a security group. To determine whether you have a default VPC, open the Amazon EC2 console and look for default VPC under account attributes on the dashboard.

Step 5 : Create a security group

- Security groups act as a firewall for associated instances, controlling both inbound and outbound traffic at the instance level. You must add rules to a security group that enable you to connect to your instance from your IP address using SSH. You can also add rules that allow inbound and outbound HTTP and HTTPS access from anywhere. Note that if you plan to launch instances in multiple regions, you'll need to create a security group in each region.

Step 6 : Launch an instance

- You can launch a Linux instance using the AWS management console as described in the following procedure.
 1. Open the Amazon EC2 console at <https://console.aws.amazon.com/ec2/>.
 2. From the console dashboard, choose **Launch Instance**.
 3. The **Choose an Amazon Machine Image (AMI)** page displays a list of basic configurations, called Amazon Machine Images (AMIs), that serve as templates for your instance. Select an HVM version of Amazon Linux 2. Notice that these AMIs are marked "Free tier eligible."
 4. On the **Choose an Instance Type** page, you can select the hardware configuration of your instance. Select the t2.micro type, which is selected by default. Notice that this instance type is eligible for the free tier.
 5. Choose **Review and Launch** to let the wizard complete the other configuration settings for you.
 6. On the **Review Instance Launch** page, under security groups, you'll see that the wizard created and selected a security group for you. You can use this security group or alternatively you can select the security group that you created when getting setup using the following steps.
 - a) Choose **Edit security groups**.

- b) On the **Configure Security Group** page, ensure that **Select an existing security group** is selected.
- c) Select your security group from the list of existing security groups and then choose **Review and Launch**.
7. On the **Review Instance Launch** page, choose **Launch**.
 8. When prompted for a key pair, select **Choose an existing key pair**, then select the key pair that you created when getting setup. When you are ready, select the acknowledgement check box and then choose launch instances.
 9. A confirmation page lets you know that your instance is launching. Choose **View Instances** to close the confirmation page and return to the console.
 10. On the **Instances** screen, you can view the status of the launch. It takes a short time for an instance to launch. When you launch an instance, its initial state is pending. After the instance starts, its state changes to running and it receives a public DNS name.
 11. It can take a few minutes for the instance to be ready so that you can connect to it. Check that your instance has passed its status checks; you can view this information in the status checks column.

Step 7 : Connect to your Instance

Several ways to connect to your Linux instance is shown in Table 4.2.1.

Your computer OS	Topic
Linux	Connecting to your Linux instance using SSH.
Windows	Connecting to your Linux instance from Windows using PuTTY.
	Connecting to your Linux instance from Windows using Windows Subsystem for Linux.
Other	Connecting to your Linux instance using MindTerm

Table 4.2.1 Ways to connect to Linux instance

Step 8 : Cleanup your instance

- After you've finished with the instance, you should cleanup by terminating the instance.
- Terminating an instance effectively deletes it; you can't reconnect to an instance after you've terminated it.
- If you launched an instance that is not within the AWS free tier, you'll stop incurring charges for that instance as soon as the instance status changes to shutting down or terminated. If you'd like to keep your instance for later, but not incur charges, you can stop the instance now and then start it again later.

- To terminate your instance following steps can be used :
 - 1) In the navigation pane, choose instances. In the list of instances, select the instance.
 - 2) Choose actions, instance state, terminate.
 - 3) Choose yes, terminate when prompted for confirmation.
- Amazon EC2 shuts down and terminates your instance. After your instance is terminated, it remains visible on the console for a short while and then the entry is deleted.

4.2.2 Amazon S3

- Amazon S3 has a simple web services interface that you can use to store and retrieve any amount of data, at any time, from anywhere on the web. S3 can serve as a raw data store for IoT systems for storing raw data, such as sensor data, log data, audio and video data.

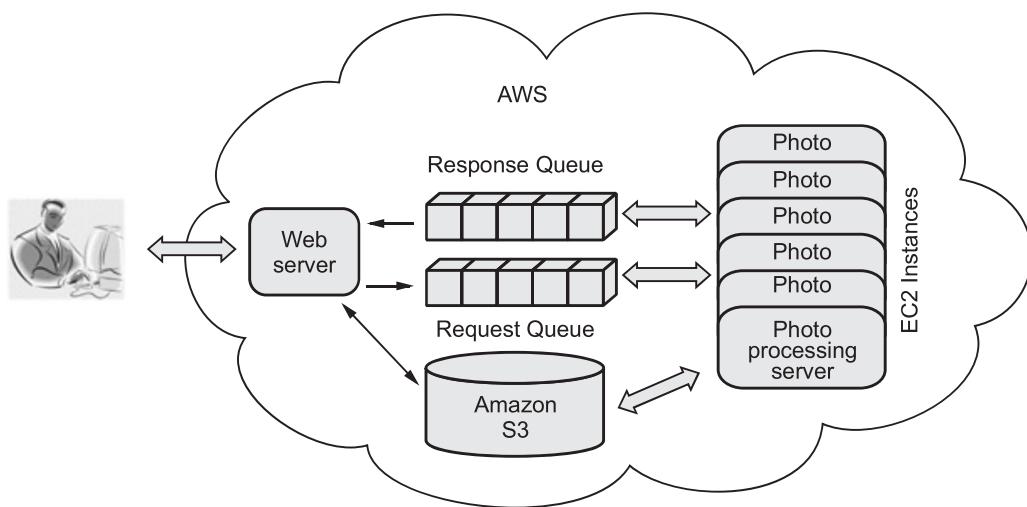


Fig. 4.2.3 Amazon S3 working

- Features :
 1. Unlimited storage
 2. Highly scalable : In terms of storage, request rate and concurrent users.
 3. Reliable : Store redundant data in multiple facilities and on multiple devices.
 4. Secure : Flexibility to control who / how / when / where to access the data.
 5. Performance : Choose region to optimize for latency / minimize costs.
- Example : Online photo processing service.

Procedure :

1. Web server receive request.
 2. Put request message in the queue.
 3. Pictures stored in S3.
 4. Multiple EC2 instances run photo processing.
 5. Put back in the queue.
 6. Return
- Store data on Amazon's distributed system containing multiple servers within Amazon's data center locations. Amazon doesn't offer you a GUI based tool to access your data. You can use one of the several tools online or build one through APIs.
 - Amazon EC2 provides three type of storage option : Amazon EBS, Amazon S3 and Instance Storage. Amazon EBS (Elastic Block Store) provides with persistent, block-level storage. Basically additional hard disk that you can attach to instance. It suitable for apps which require database, filesystem, block level storage.
 - A **bucket** is a container for objects stored in Amazon S3. Every object is contained in a bucket. For example, if the object named "photos/puppy.jpg" is stored in the rakshita bucket, then it is addressable using the URL <http://rakshita.s3.amazonaws.com/photos/puppy.jpg>
 - Buckets serve several purposes : They organize the Amazon S3 namespace at the highest level, they identify the account responsible for storage and data transfer charges, they play a role in access control and they serve as the unit of aggregation for usage reporting.
 - **Objects** are the fundamental entities stored in Amazon S3. Objects consist of object data and metadata. The data portion is opaque to Amazon S3. The metadata is a set of name-value pairs that describe the object. These include some default metadata, such as the date last modified and standard HTTP metadata, such as content-type. You can also specify custom metadata at the time the object is stored.
 - A **key** is the unique identifier for an object within a bucket. Every object in a bucket has exactly one key. Because the combination of a bucket, key and version ID uniquely identify each object, Amazon S3 can be thought of as a basic data map between "bucket + key + version" and the object itself. Every object in Amazon S3 can be uniquely addressed through the combination of the web service endpoint, bucket name, key and optionally, a version.

- **Regions :** You can choose the geographical region where Amazon S3 will store the buckets you create. Objects stored in a region never leave the region unless you explicitly transfer them to another region.

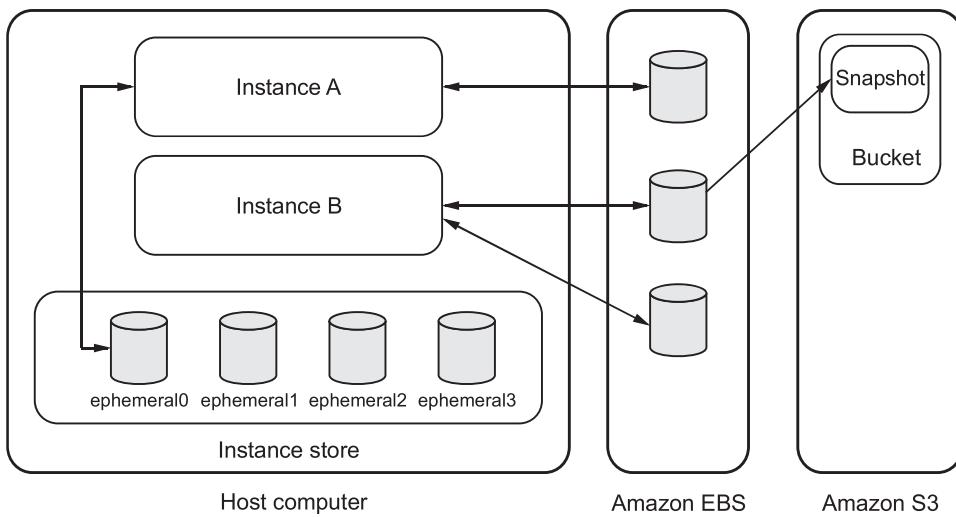


Fig. 4.2.4 Amazon EBS and S3

4.2.3 Amazon CloudWatch

- It is a monitoring service for AWS cloud resources and the applications user run on AWS.
- User can use Amazon CloudWatch to collect and track metrics, collect and monitor log files and set alarms.
- Amazon CloudWatch can monitor AWS resources such as Amazon EC2 instances, Amazon DynamoDB tables and Amazon RDS DB Instances, as well as custom metrics generated by user applications and services and any log files user applications generate.
- User can use Amazon CloudWatch to gain system-wide visibility into resource utilization, application performance and operational health. User can use these insights to react and keep user application running smoothly.
- CloudWatch stores its data for two weeks, making it possible to track metrics across an extended period.
- CloudWatch is enabled when you create an account. Thereafter, you simply select (or define) the metrics to track and then use the metrics that are generated as you choose.
- CloudWatch provides a lot of monitoring for free. For example, for any EC2 instances you have, you get ten metrics at 5-minute intervals with no cost.

Review Questions

1. Explain the steps for configuring a server for EC2. **SPPU : June-19, End Sem, Marks 9**
2. Explain the steps to create an Amazon S3 Bucket and managing associated objects ? **SPPU : Dec.-19, End Sem, Marks 9**
3. Explain steps to configure server for EC2. **SPPU : Dec.-19, End Sem, Marks 9**

4.3 Amazon Storage System

- Amazon S3 defines a bucket name as a series of one or more labels, separated by periods, that adhere to the following rules : The bucket name can be between 3 and 63 characters long, and can contain only lower-case characters, numbers, periods, and dashes
- Amazon S3 defines a bucket name as a series of one or more labels, separated by periods, that adhere to the following rules :
 1. The bucket name can be between 3 and 63 characters long, and can contain only lower-case characters, numbers, periods, and dashes.
 2. Each label in the bucket name must start with a lowercase letter or number.
 3. The bucket name cannot contain underscores, end with a dash, have consecutive periods, or use dashes adjacent to periods.
 4. The bucket name cannot be formatted as an IP address (198.51.100.24).
- A bucket is owned by the AWS account that created it. By default, you can create up to 100 buckets in each of your AWS accounts. If you need additional buckets, you can increase your bucket limit by submitting a service limit increase
- The following are the rules for naming S3 buckets in all AWS Regions :
 1. Bucket names must be unique across all existing bucket names in Amazon S3.
 2. Bucket names must comply with DNS naming conventions.
 3. Bucket names must be at least 3 and no more than 63 characters long.
 4. Bucket names must not contain uppercase characters or underscores.
 5. Bucket names must start with a lowercase letter or number.
 6. Bucket names must be a series of one or more labels. Adjacent labels are separated by a single period (.). Bucket names can contain lowercase letters, numbers, and hyphens. Each label must start and end with a lowercase letter or a number.
 7. Bucket names must not be formatted as an IP address (for example, 192.168.5.4).

8. When you use virtual hosted-style buckets with Secure Sockets Layer (SSL), the SSL wildcard certificate only matches buckets that don't contain periods. To work around this, use HTTP or write your own certificate verification logic. We recommend that you do not use periods (".") in bucket names when using virtual hosted-style buckets.

4.4 Amazon Database Services

- Amazon Relational Database Service (RDS) is a web service that makes it easy to set up, operate and scale a relational database in the cloud. RDS gives access to the capabilities of a familiar MySQL, Oracle or Microsoft SQL Server database engine. Code, applications and tools already used with existing databases can be used with RDS.
- Amazon RDS is just a replacement to running your own database server. With a simple sequence of commands at the console, you can choose from two commercial DBMS (Oracle or Microsoft SQL Server) or two open source DBMS (MySQL or PostgreSQL).
- Amazon RDS automatically patches the database software and backs up the database, storing the backups for a user-defined retention period and enabling point-in-time recovery.
- Amazon RDS provides scaling the compute resources or storage capacity associated with the Database Instance. Pay only for the resources actually consumed, based on the DB Instance hours consumed, database storage, backup storage and data transfer.
- On-Demand DB Instances let you pay for compute capacity by the hour with no long-term commitments. Reserved DB Instances give the option to make a low, one-time payment for each DB Instance and in turn receive a significant discount on the hourly usage charge for that DB Instance.
- Amazon RDS provides you six familiar database engines to choose from, including Amazon Aurora, PostgreSQL, MySQL, MariaDB, Oracle and Microsoft SQL Server.

- Create the actual RDS instance :

```
instance = conn.create_dbinstance ( id=identifier, instance_class="db." + instance_class,
allocated_storage=allocated_storage,
engine=engine, db_name=db_name, master_username=master_username,
master_password=master_password,
db_subnet_group_name=identifier, vpc_security_groups=security_groups, port=port)
```

Advantages :

1. **Fast and easy to administer** : Amazon RDS makes it easy to go from project conception to deployment.

2. **Available and durable** : Amazon RDS runs on the same highly reliable infrastructure used by other Amazon Web Services.
3. **Highly scalable** : You can scale your database's compute and storage resources.
4. **Secure** : Amazon RDS makes it easy to control network access to your database.

4.4.1 Dynamo DB

- DynamoDB is a fast, fully managed **NoSQL database service** that makes it simple and cost-effective to store and retrieve any amount of data and serve any level of request traffic.
- All data items are stored on **Solid State Drives (SSDs)** and are **replicated** across **3 Availability Zones** for high availability and durability.
- DynamoDB **tables do not have fixed schemas** and each item may have a **different number of attributes**.
- DynamoDB has **no upfront costs** and implements a **pay as you go** plan as a **flat hourly rate** based on the **capacity reserved**.
- With DynamoDB, you can create database tables that can store and retrieve any amount of data and serve any level of request traffic. You can scale up or scale down your tables' throughput capacity without downtime or performance degradation and use the AWS management console to monitor resource utilization and performance metrics.
- DynamoDB allows you to delete expired items from tables automatically to help you reduce storage usage and the cost of storing data that is no longer relevant.
- The first step in accessing DynamoDB is to create a connection to the service :

```
import boto.dynamodb
conn = boto.dynamodb.connect_to_region('us-west-2',
aws_access_key_id='<YOUR_AWS_KEY_ID>',
aws_secret_access_key='<YOUR_AWS_SECRET_KEY>')
```

- Python program for creating a DynamoDB table with other operation :

```
import boto.dynamodb
from boto.dynamodb.condition import *
connection = boto.dynamodb.connect_to_region('eu-west-1')
table = connection.get_table('table')

id = '1'
timestamp = 1234
atrs = {
    'key1': 'value1',
    'key2': set(['value2', 'value3'])
}
```

```

# create
item = table.new_item(hash_key=id, range_key=timestamp, attrs=attrs)
item.put()

# read
item = table.get_item(hash_key=id)
key2 = list(item['key2'])

# update
item['key1'] = 'foo'
item['key3'] = 'bar'
item.put()

# query
table.query(hash_key=id, range_key_condition=LT(1500))

# scan
table.scan(scan_filter={'key1': EQ('foo')})

# delete
item = table.get_item(hash_key=id)
item.delete()

```

Advantages :

1. **Flexible** : Amazon DynamoDB supports both document and key-value data structures.
2. **Fully managed** : Amazon DynamoDB is a fully managed cloud NoSQL database service.
3. **Highly scalable** : When you create a table, simply specify how much request capacity you require.
4. **Event-driven programming** : Amazon DynamoDB integrates with AWS lambda to provide triggers

4.4.2 Difference between DynamoDB and Amazon S3

DynamoDB	Amazon S3
DynamoDB is database	S3 is file storage
Used in semi structured data	It is used for unstructured data
Size limit is 400 kb	Size limit is 5 TB

DynamoDb supports two kinds of primary keys, Partition Key and Partition key and Sort key.	S3 uses unique Ids called Keys to retrieve files from the bucket
DynamoDb is used to store key-value. It uses items and attributes for its tables	S3 stores files in a flat organisation of containers called Buckets

4.5 Microsoft Cloud Services : Azure

- Windows Azure is a cloud computing platform and infrastructure, created by Microsoft, for building, deploying and managing applications and services through a global network of Microsoft - managed data centers.
- Azure queue storage is a service for storing large numbers of messages that can be accessed from anywhere in the world via authenticated calls using HTTP or HTTPS. A single queue message can be up to 64 KB in size, and a queue can contain millions of messages, up to the total capacity limit of a storage account.
- Azure is a virtualized infrastructure to which a set of additional enterprise services has been layered on top, including, a virtualization service called Azure AppFabric that creates an application hosting environment. AppFabric is a cloud-enabled version of the .NET framework.
- Windows Azure is Microsoft's application platform for the public Cloud. Applications can be deployed on to Azure in various models
- Windows Azure is used to :
 1. Build a web application that runs and stores its data in Microsoft data centers.
 2. Store data while the applications that consume this data run on premise (outside the public Cloud).
 3. Create virtual machines to develop and test, or run SharePoint and other out-of-the-box applications.
 4. Develop massively scalable applications with many users.
 5. Offer a wide range of services
- Azure has three components : compute, storage and fabric
 1. **Compute** : Windows Azure provides a hosting environment for managed code. It provides a computation service through roles. Windows Azure supports three types of roles :
 - a) Web roles used for web application programming and supported by IIS7.
 - b) Worker roles are also used for background processing of web roles.
 - c) Virtual Machine (VM) roles are generally used for migrating windows server applications to Windows Azure in an easy way.

- 2. **Storage** : Windows Azure provides storage in the cloud. It provides four different types of storage services :
 - a) Queues for messaging between web roles and worker roles.
 - b) Tables for storing structural data.
 - c) BLOBs (Binary Large Objects) to store text, files or large data.
 - d) Windows Azure Drives (VHD) to mount a page blob. They can easily be downloaded and uploaded via blobs.
- 3. AppFabric provides infrastructure services for developing, deploying and managing Windows Azure application. It provides five services: Service bus, Access, Caching, Integration and Composite.
- Fig. 4.5.1 shows Windows Azure platform architecture.

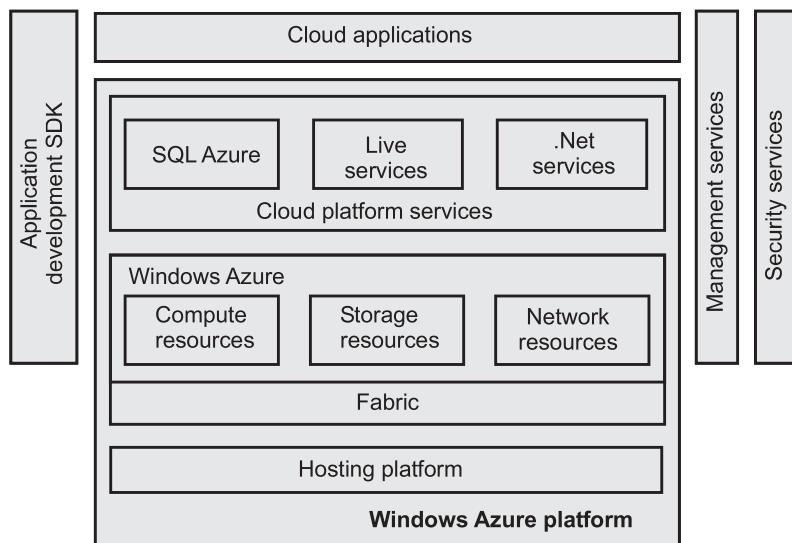


Fig. 4.5.1 Windows Azure platform architecture

- Microsoft Azure is a cloud computing service created by Microsoft for building, testing, deploying and managing applications and services through a global network of Microsoft-managed data centers.
- It provides software as a service (SaaS), platform as a service and infrastructure as a service and supports many different programming languages, tools and frameworks, including both Microsoft-specific and third-party software and systems.
- Windows Azure provides resources and services for consumers. For example, hardware is abstracted and exposed as compute resources.

- Physical storage is abstracted as storage resources and exposed through very well-defined interfaces.
- A common windows fabric abstracts the hardware and the software and exposes virtual compute and storage resources.
- Each instance of an application is automatically managed and monitored for availability and scalability.
- If an application goes down, the Fabric is notified and a new instance of the application is created. Because virtualization is a key element in cloud computing, no assumption must be made on the state of the underlying hardware hosting the application.
- Advantages of Microsoft Azure
 1. Microsoft Azure offers high availability
 2. It offers you a strong security profile
 3. It is a cost-effective solution for an IT budget.
 4. Azure allows you to use any framework, language, or tool.
 5. Azure allows businesses to build a hybrid infrastructure.

4.6 Cloud Computing Applications

- Cloud Adoption is a strategic move by organisations of reducing cost, mitigating risk and achieving scalability of data base capabilities. Cloud adoption may be up to various degrees in an organisation, depending on the depth of adoption. In fact the depth of adoption yields insight into the maturity of best practices, enterprise-ready cloud services availability.
- A variety of industries benefit from cloud adoption, including healthcare, marketing and advertising, retail, finance and education. Benefits include :
- Healthcare : Fueled by digital and social consumer behaviors and the need for secure and accessible Electronic Health Records (EHRs), hospitals, clinics and other medical organizations are using cloud computing for document storage, marketing and human resources.
- Marketing and Advertising : In an industry dependent on social media, as well as the quick creation and publishing of customer - relevant content, agencies are using hybrid cloud adoption strategies to deliver critical client messages to their local and worldwide audiences.
- Retail : A successful e-commerce strategy requires a sound Internet strategy. With the help of cloud adoption, internet-based retail is able to effectively market to customers and save their product data for less money.

- Finance : Efficient expense management, human resources and customer communications are three of the most important business needs of today's finance organizations. For these reasons, financial services institutions are now placing their email platforms and marketing tools in the cloud.

4.6.1 Healthcare : ECG Analysis in the Cloud

- Electrocardiograph (ECG) analysis brings a lot of technical concerns because ECG is one of the tools frequently used in the diagnosis of cardiovascular disease.
- In the cloud system, medical data can be gathered and distributed automatically to medical practitioners anywhere in the world. From there, doctors in the field have the capability of returning input to specific patients.
- Cloud computing technologies allows the remote monitoring of a patient's heart beat data. Through this way the patient at risk can be constantly monitored without going to the hospital for ECG analysis. At the same time the Doctor's can instantly be notified with cases that need's their attention. Fig. 4.6.1 shows cloud ECG process.

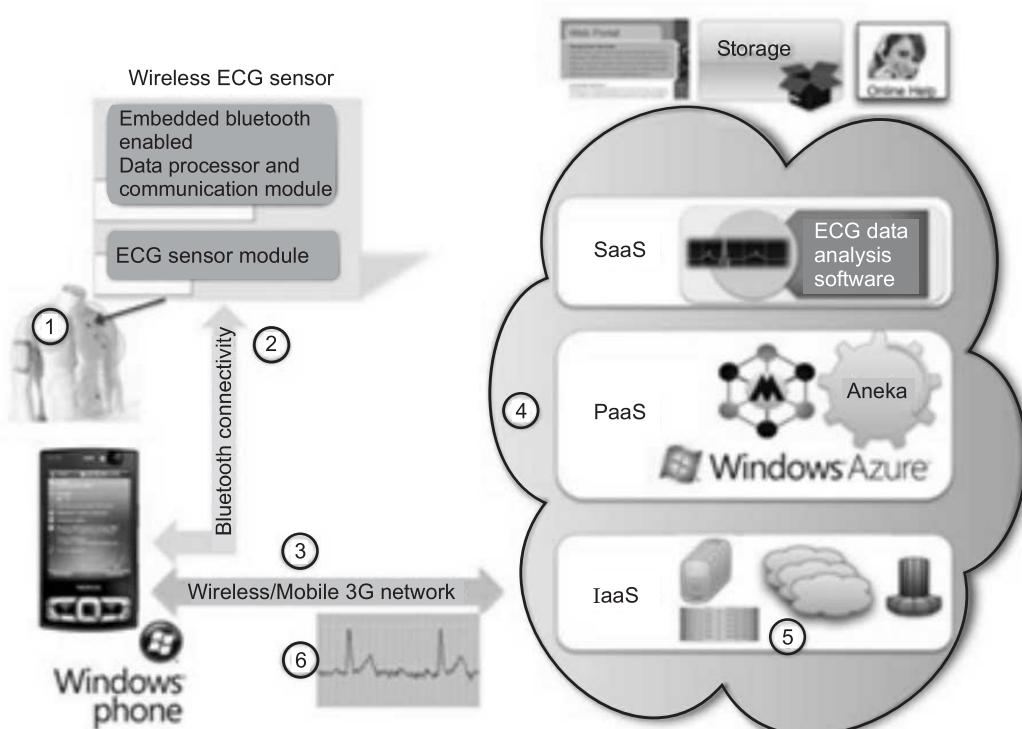


Fig. 4.6.1 Cloud ECG process

- The respective information is transmitted to the patient's mobile device that will immediately forwarded to the cloud- hosted web services for analysis. The entire web services from the front end of a platform that is completely hosted in the cloud that consist of three layers : SaaS, PaaS, IaaS.

4.6.2 Biology : Protein Structure Prediction

- The prediction of the protein structure is the inference from its amino acid sequence of the three - dimensional structure of a protein, that is, the prediction of its secondary and tertiary structure from the primary structure.
- Fig. 4.6.2 shows protein structure prediction using cloud.

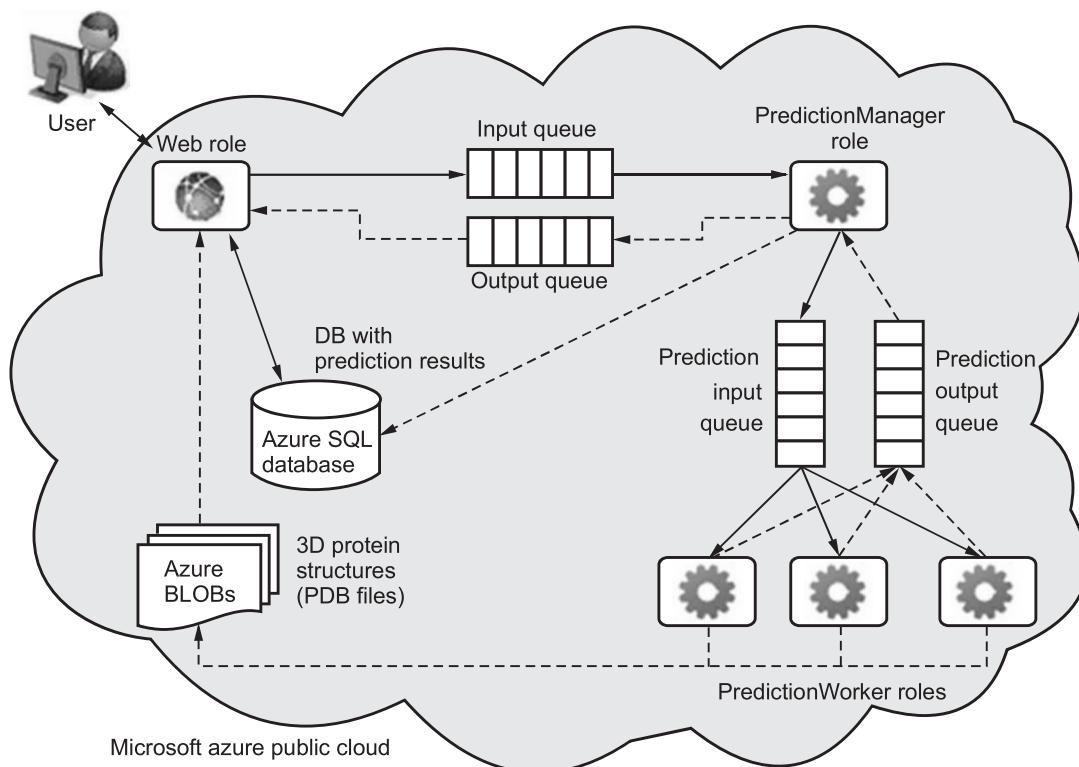


Fig. 4.6.2

- The Predict Protein cloud solution builds upon the open source operating system Debian and provides its functionality as a set of free software packages. Bio-Linux is an operating system for bioinformatics and computational biology.

- The latest Bio-Linux release provides more than 500 bioinformatics programs on an Ubuntu Linux base. Ubuntu is a "derivative" operating system based on Debian, with its own additions.
- Cloud BioLinux is a comprehensive cloud solution that is derived from Bio-Linux and Ubuntu. Debian derivatives can easily share packages between each other. For example, Debian packages are automatically incorporated in Ubuntu and are also usable in Cloud BioLinux.

4.6.3 Geosciences : Satellite Image Processing

- Satellite image processing plays a vital role for research and developments in Astronomy, Remote Sensing, GIS, Agriculture Monitoring, Disaster Management and many other fields of study. Satellite images are recorded in digital forms and then processed by the computers to extract information.
- Satellite remote sensing generates hundreds of giga - bytes of raw images that need to be further processed to become the basis of several different GIS products. Fig. 4.6.3 shows cloud environment for satellite data processing.

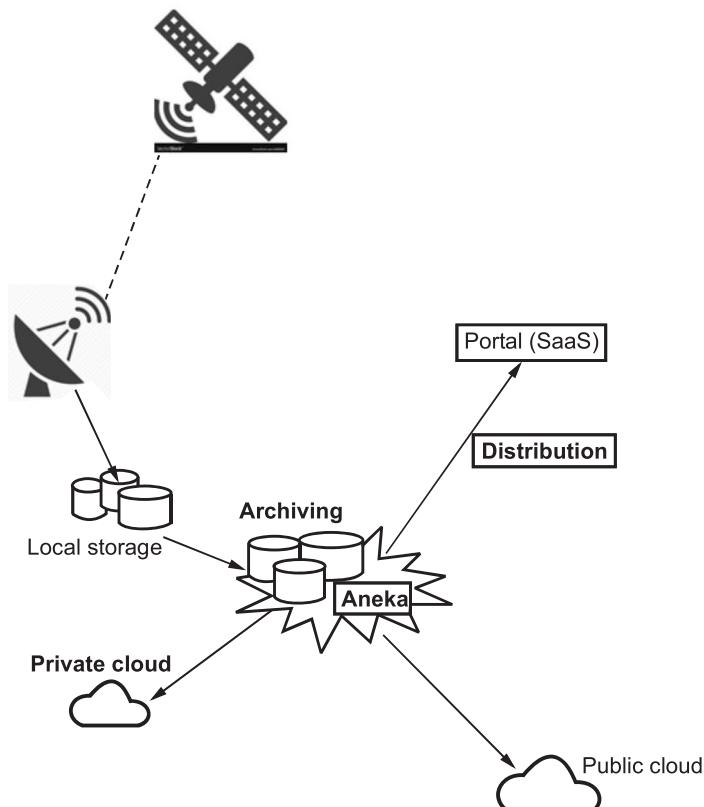


Fig. 4.6.3 Cloud environment for satellite data processing

- This process requires both I/O and compute-intensive tasks. Large images need to be moved from a ground station's local storage to compute facilities, where several transformations and corrections are applied.
- Cloud computing provides the appropriate infrastructure to support such application scenarios.
- Satellite data processing consists of various methods to correct the radiometric errors and geometric distortions in the basic data generated by the sensor; this data is termed as Level-0.
- The procedures like georeferencing and registration applied on the Level-0 data to generate the products such as;
 1. Level 1 - Radio metrically corrected and geometrically corrected only for earth rotation (Browse product)
 2. Level 2 - Both radiometric and geometrically corrected (Standard product)

4.6.4 Business and Consumer Applications : CRM and ERP

- Customer Relationship Management (CRM) is a strategy for managing all your company's relationships and interactions with your customers and potential customers. It helps you improve your profitability.
- Cloud ERP is Software as a Service that allows users to access Enterprise Resource Planning (ERP) software over the Internet. Cloud ERP generally has much lower upfront costs, because computing resources are leased by the month rather than purchased outright and maintained on premises. Cloud ERP also gives companies access to their business-critical applications at any time from any location.
- While technically the only difference between Cloud ERP and on-premises ERP is where the software is physically located, there are other significant differences. Here we explain some of the key characteristics and advantages of Cloud ERP software.
- The Cloud is particularly valuable to Small and Medium - Size Businesses (SMB's) because it provides access to full-function applications at a reasonable price without a substantial upfront expenditure for hardware and software. Using the right cloud provider, a company can rapidly scale their business productivity software as their business grows or a new company is added.
- Cloud ERP has been proven to reduce costs in many ways because it :
 - a) Avoids upfront costs for all computing infrastructure such as hardware and data servers.
 - b) Reduces IT support services because IT support is provided by the data center.

- c) Eliminates paying upfront for application software licenses in favor of a monthly fee.
- d) Shrinks the cost of maintaining and supporting those applications since the cloud vendor handles the updates and upgrades.

4.6.5 Social Networking

- Social Network Analysis (SNA) is an important and valuable tool for knowledge extraction from massive and un-structured data. Social network provides a powerful abstraction of the structure and dynamics of diverse kinds of inter-personal connection and interaction.
- Facebook is a social networking service and website that connects people with other people and share data between people. A user can create a personal profile, add other users as friends, exchange data, create and join common interest communities.
- Twitter is a social net-working and microblogging service. The users of Twitter can exchange text-based posts called tweets. A tweet is a maximum 140 characters long but can be augmented by pictures or audio recording. The main concept of Twitter was to build a social network formed by friends and fol-lowers. Friends are people who you follow, followers are those who follow you.
- The role of social networks in labor markets deserves attention for at least two reasons : First, because of the central role networks play in disseminating information about job openings they place a critical role in determining whether labor markets function efficiently and second, because network structure ends up having implications for things like human capital investment as well as inequality.
- Social Network Analysis (SNA) primarily focuses on applying analytic techniques to the relationships between individuals and groups and investigating how those relationships can be used to infer additional information about the individuals and groups.
- SNA is used in a variety of domains. For example, business consultants use SNA to identify the effective relationships between workers that enable work to get done; these relationships often differ from connections seen in an organizational chart.
- Law enforcement personnel have used social networks to analyze terrorist networks and criminal networks. The capture of Saddam Hussein was facilitated by social network analysis : Military officials constructed a network containing Hussein's tribal and family links, allowing them to focus on individuals who had close ties to Hussein

- Collecting social network data used to be a tedious, labor - intensive process. In fact, several notable dissertations came out of the researcher's being at the right place and the right time to be able to observe a social conflagration and gather data on it.
- Social network data collection is, by nature, more invasive and harder to anonymize; survey instruments had to be approved by Institutional Review Boards (IRBs) and administration of the surveys was tedious manual labor.
- Some of key challenges in this kind of data collection are :
 - 1 Network boundaries are difficult to define.
 - 2 People do not easily recall their network members and need appropriate "prompts" to elicit them. In addition, networks are very large in general and different social network members may have different importance depending on the phenomenon studied.
 - 3 Information about the network members needs to balance detail and interviewee's burden.
- Most social network data collection can be divided into "whole" and "egocentric" networks. Whole network studies examine actors "that are regarded for analytical purposes as bounded social collectives"; actors in these studies are named in closed lists, usually pre-defined and known a priori.
- Since these boundaries are very difficult to define in urban settings with large populations, whole network studies are unpractical, making egocentric data collection the only feasible method.
- Egocentric network studies concentrate in specific actors or egos and those who have relations with them, called alters. That is, from the participant's perspective, egocentric networks constitute a "network of me" or a network of actors with whom the participant has some relationship.
- Egocentric network data is thus composed by two levels :
 - i) An ego-network level, constituted by the ego's characteristics and overall network features; and
 - ii) An ego-alter level, constituted by the characteristics of each alter and alter-ego ties.

4.7 Google Cloud Application : Google App Engine

- Google App Engine (GAE) is a Platform as a Service cloud computing platform for developing and hosting web applications in Google-managed data centers.

- Google App Engine is a way to write your own web applications and have them hosted on Google servers. It enables developers to build their web applications on the same scalable system that power Google applications.
- An app is a piece of software which can run on the computer, internet, phone or any other electronic device. Google refers to their online services as Apps. They also sell a specific suite of services known as Google Apps.
- Google's providing both SaaS and PaaS solutions in cloud computing. Some of the examples for SaaS solutions including Google Apps which including Gmail, Doc, etc. and PaaS includes Google App engine.
- Services provided by App engine includes :
 - a) Platform as a Service (PaaS) to build and deploy scalable applications
 - b) Hosting facility in fully-managed data centers
 - c) A fully-managed, flexible environment platform for managing application server and infrastructure.
 - d) Support in the form of popular development languages and developer tools.
- **Major feature of Google App Engine :**
 1. Automatic scaling and load balancing.
 2. Authentication using Google Accounts API.
 3. Provides dynamic web services based on common standards.
 4. Integration with other Google Cloud Services and API.
 5. Support persistent storage, with query access sorting and transaction management features.
- Google App engine offers users the ability to build and host web applications on Google's infrastructure.

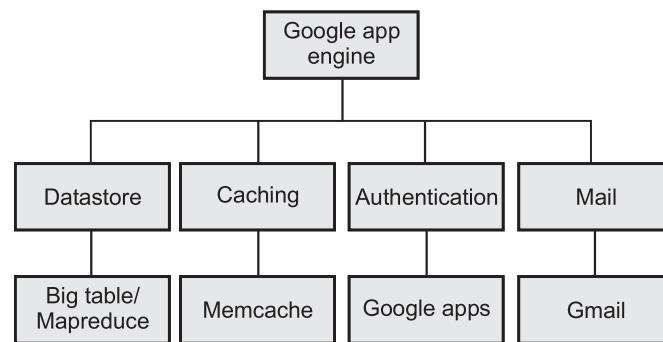


Fig. 4.7.1

- The App Engine offers a number of services that enable you to perform several common operations when managing your application. The following APIs are available to access these services :
 1. **Mail** : Using the mail API, the developers can send email messages.
 2. **Memcache** : The Memcache service gives the users the benefit of working efficiently by providing high retrieval speed, even when multiple users access the same application at the same instance of time.
 3. **Image manipulation** : The Image service allows you to manipulate images of your application. With the use of this API, you can resize, crop, rotate and flip images in JPEG and PNG formats.
- In the PaaS space Google is a key player. App Engine is a platform to create, store and run applications on Google's servers using development languages as java and python.
- App Engine includes tools for managing the data store, monitoring the site and its resource consumption and debugging and logging. A user can serve the app from his own domain name using Google Apps.
- **Key features of GAE programming mode using java and python.**
- The Google App engine Software Development Kit (SDK) provides Java and Python programming languages.
- The languages have their own web server application that contains all Google App Engine services on a local computer. The web server also simulates a secure sandbox environment.
- The Google App engine SDK has APIs and libraries including the tools to upload applications. The architecture defines the structure of applications that run on the Google App engine.

1. Python :

- The Google App engine allows implementation of applications using python programming language and running them on its interpreter.
- The Google App engine provides rich APIs and tools for designing web applications, data modeling, managing, accessing apps data, support for mature libraries and frameworks like Django.
- The main characteristics of Google App engine are its DataStore, configuration file app.yaml and how it serves an application.

2. Java :

- The Google App engine provides tools and APIs required for the development of web applications that run on the Google App engine Java run time.

- The application interacts with the environment using servlets and web technologies like Java Server Pages (JSPs) which can be developed using Java6.
- The GAE environment uses Java SE Runtime JRE platform 6 and libraries which the applications can access using APIs.
- Java SDK has implementations for Java Data Objects (JDO) and Java Persistence (JPA) interface.
- To exchange email messages with Google App engine, it provides the Google App Engine mail service through the Java Mail API.
- Support for other languages like JavaScript, Ruby or Scala is also provided by Google App engine with the use of JVM compatible compilers and interpreters.
- When Google App engine gets a web request that corresponds to the URL mentioned in the applications deployment descriptor it invokes a servlet corresponding to the request and uses Java Servlets API to provide requested data and accepts response data.
- Google App engine makes it easy to build an applications that runs reliably, even under heavy load and with large amounts of data.
- App engine includes the below features :
 - a) Dynamic web serving, with full support for common web technologies.
 - b) Persistent storage with queries, sorting and transactions.
 - c) Automatic scaling and load balancing.
 - d) APIs for authenticating users and sending email using Google accounts.
 - e) Scheduled tasks for triggering events at specified times and regular intervals.

4.8 Overview of OpenStack Architecture

- OpenStack is a recently open-sourced, IaaS cloud-computing platform founded by Rackspace Hosting and NASA, and is used widely in industry
- OpenStack is an open-source cloud platform. OpenStack software controls large pools of compute, storage, and networking resources throughout a data center, all managed by a dashboard that gives administrators control while empowering their users to provision resources through a web interface.
- To produce the ubiquitous Open-Source cloud computing platform that will meet the needs of public and private cloud providers regardless of size, by being simple to implement and massively scalable.

- Components of OpenStack are as follows :
 1. Horizon - Dashboard : It provides a modular web-based user interface for all the OpenStack services. With this web GUI, user can perform most operations on your cloud like launching an instance, assigning IP addresses and setting access controls.
 2. Keystone is a framework for authentication and authorization for all the OpenStack services. It handles API requests as well as providing configurable catalog, policy, token and identity services. Keystone is a framework for authentication and authorization for all the OpenStack services.
 3. Nova : It provides virtual servers upon demand. Nova is the most complicated and distributed component of OpenStack. A large number of processes cooperate to turn end user API requests into running virtual machines.
 4. Glance - Image Store : It provides discovery, registration and delivery services for disk and server images.
 5. Quantum - Network : It provides " network connectivity as a service " between interface devices managed by other OpenStack services. The service works by allowing users to create their own networks and then attach interfaces to them. Quantum has a pluggable architecture to support many popular networking vendors and technologies.
 6. Cinder allows block devices to be exposed and connected to compute instances for expanded storage & better performance.
 7. Object store allows you to store or retrieve files. It provides a fully distributed, API-accessible storage platform that can be integrated directly into applications or used for backup, archiving and data retention.

4.9 Multiple Choice Questions

- Q.1** Amazon EC2 is a computing service, whereas Amazon SQS and Amazon S3 are _____ services.
- | | |
|-------------------------------------|--------------------------------------|
| <input type="checkbox"/> a) cloud | <input type="checkbox"/> b) support |
| <input type="checkbox"/> c) storage | <input type="checkbox"/> d) platform |
- Q.2** Google Cloud Storage is a RESTful online _____ storage web service for storing and accessing one's data on Google's infrastructure.
- | | |
|-----------------------------------|---|
| <input type="checkbox"/> a) data | <input type="checkbox"/> b) information |
| <input type="checkbox"/> c) block | <input type="checkbox"/> d) file |

- Q.3** Google Cloud Print is a service that extends the printer's function to any device that can connect to the _____.
- [a] Intranet [b] Internet
[c] WAN [d] LAN
- Q.4** The AWS global infrastructure consists of multiple geographical locations which are called _____.
- [a] division [b] block
[c] regions [d] cluster
- Q.5** Amazon Elastic Block Store provides persistent _____ storage volumes for use with Amazon EC2 instances in the AWS Cloud.
- [a] file [b] block
[c] directory [d] all of these
- Q.6** Amazon _____ is a low-cost cloud storage service for data with longer retrieval times offered by Amazon Web Services.
- [a] Cloudnet [b] RedShift
[c] DynamoDB [d] Glacier
- Q.7** Amazon EC2 provides virtual computing environment, known as _____.
- [a] chunks [b] instances
[c] block [d] messages
- Q.8** What does S3 stand for ?
- [a] Super Storage Service [b] Storage Simple Service
[c] Single Storage Service [d] Simple Storage Service
- Q.9** Amazon S3 is a _____.
- [a] Content Delivery Network [b] Domain System
[c] Key-Based object store [d] Relational database
- Q.10** Which of these services are used to distribute content to end users using a global network of edge locations ?
- [a] CloudFront [b] Route 53
[c] Virtual Private Cloud [d] CloudWatch

Q.11 DynamoDB is a fast, fully managed _____ database service that makes it simple and cost-effective to store and retrieve any amount of data and serve any level of request traffic.

- | | |
|-----------------------------------|---|
| <input type="checkbox"/> a PL/SQL | <input type="checkbox"/> b NoSQL |
| <input type="checkbox"/> c SQL | <input type="checkbox"/> d All of these |

Q.12 S3 stores files in a flat organisation of containers called _____.

- | | |
|------------------------------------|--|
| <input type="checkbox"/> a chunks | <input type="checkbox"/> b block |
| <input type="checkbox"/> c Buckets | <input type="checkbox"/> d none of these |

Q.13 Microsoft offers its own online collaboration tool called _____.

- | | |
|---------------------------------------|---------------------------------------|
| <input type="checkbox"/> a Azer | <input type="checkbox"/> b salespoint |
| <input type="checkbox"/> c Sharepoint | <input type="checkbox"/> d Aneka |

Answer Keys for Multiple Choice Questions

Q.1	b	Q.2	d	Q.3	b	Q.4	c
Q.5	b	Q.6	d	Q.7	b	Q.8	d
Q.9	c	Q.10	a	Q.11	b	Q.12	c
Q.13	c						



Notes

UNIT - V

5

Security in Cloud Computing

Syllabus

Risks in Cloud Computing : Risk Management, Enterprise-Wide Risk Management, Types of Risks in Cloud Computing. **Data Security in Cloud :** Security Issues, Challenges, advantages, Disadvantages, Cloud Digital persona and Data security, Content Level Security.
Cloud Security Services : Confidentiality, Integrity and Availability, Security Authorization Challenges in the Cloud, Secure Cloud Software Requirements, Secure Cloud Software Testing.

Contents

- 5.1 Risks in Cloud Computing
- 5.2 Enterprise-wide Risk Management
- 5.3 Types of Risks in Cloud Computing
- 5.4 Data Security in Cloud : Security Issues and Challenges **March-20,** Marks 5
- 5.5 Content Level Security
- 5.6 Cloud Security Services
- 5.7 Security Authorization Challenges in the Cloud
- 5.8 Secure Cloud Software Requirements
- 5.9 Secure Cloud Software Testing
- 5.10 Multiple Choice Questions