Cloud Environments

Module 4

Google App Engine

Google App Engine (GAE) is a platform-as-a-service product that provides web app developers and enterprises with access to Google's scalable hosting and tier 1 internet service.

GAE requires that applications be written in Java or Python, store data in Google Bigtable and use the Google query language.

Noncompliant applications require modification to use GAE.

GAE provides more infrastructure than other scalable hosting services, such as Amazon Elastic Compute Cloud (EC2).

GAE also eliminates some system administration and development tasks to make writing scalable applications easier.

Google App Engine

Google provides GAE free up to a certain amount of use for the following resources:

•processor (CPU)

•storage

•application programming interface (API) calls

•concurrent requests

Users exceeding the per-day or per-minute rates can pay for more of these resources.

How is GAE used?

GAE is a fully managed, serverless platform that is used to host, build and deploy web applications.

Users can create a GAE account, set up a software development kit and write application source code. They can then use GAE to test and deploy the code in the cloud.

One way to use GAE is building scalable mobile application back ends that adapt to workloads as needed.

Application testing is another way to use GAE. Users can route traffic to different application versions to A/B test them and see which version performs better under various workloads.

What are GAE's key features?

Key features of GAE include the following:

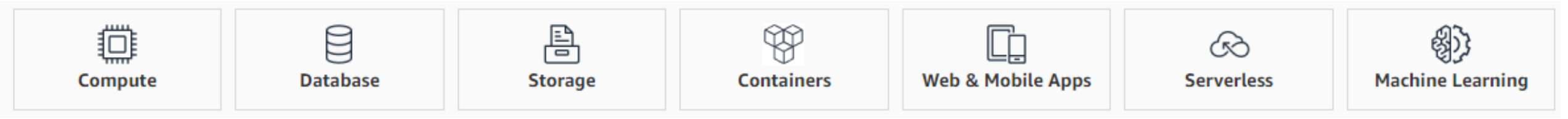
API selection. GAE has several built-in APIs, including the following five:

- **Blobstore** for serving large data objects;
- **GAE Cloud Storage** for storing data objects;
- **Page Speed Service** for automatically speeding up webpage load times;
- **URL Fetch Service** to issue HTTP requests and receive responses for efficiency and scaling; and
- **Memcache** for a fully managed in-memory data store.

Amazon AWS

AWS is designed to allow application providers, ISVs, and vendors to quickly and securely host your applications – whether an existing application or a new SaaS-based application.

You can use the AWS Management Console or well-documented web services APIs to access AWS's application hosting platform.



What is IAM?

•IAM allows you to manage users and their level of access to the aws console.

•It is used to set users, permissions and roles. It allows you to grant access to the different parts of the aws platform.

•AWS Identity and Access Management is a web service that enables Amazon Web Services (AWS) customers to manage users and user permissions in AWS.

•With IAM, Organizations can centrally manage users, security credentials such as access keys, and permissions that control which AWS resources users can access.

•IAM enables the organization to create multiple users, each with its own security credentials, controlled and billed to a single aws account. IAM allows the user to do only what they need to do as a part of the user's job.

Elastic Compute Cloud (EC2)

•Amazon EC2 is a web service that provides resizable compute capacity in the cloud.

•Amazon EC2 reduces the time required to obtain and boot new user instances to minutes rather than in older days, if you need a server then you had to put a purchase order, and cabling is done to get a new server which is a very time-consuming process.

•Now, *Amazon has provided an EC2 which is a virtual machine in the cloud that completely changes the industry.*

•You can scale the compute capacity up and down as per the computing requirement changes.

•Amazon EC2 provides the developers with the tools to build resilient applications that isolate themselves from some common scenarios.

Simple Storage Service (S3)

- S3 provides developers and IT teams with secure, durable, highly scalable object storage.

- It is easy to use with a simple web services interface to store and retrieve any amount of data from anywhere on the web

- The files which are stored in S3 can be from 0 Bytes to 5 TB.

- It has unlimited storage means that you can store the data as much you want.

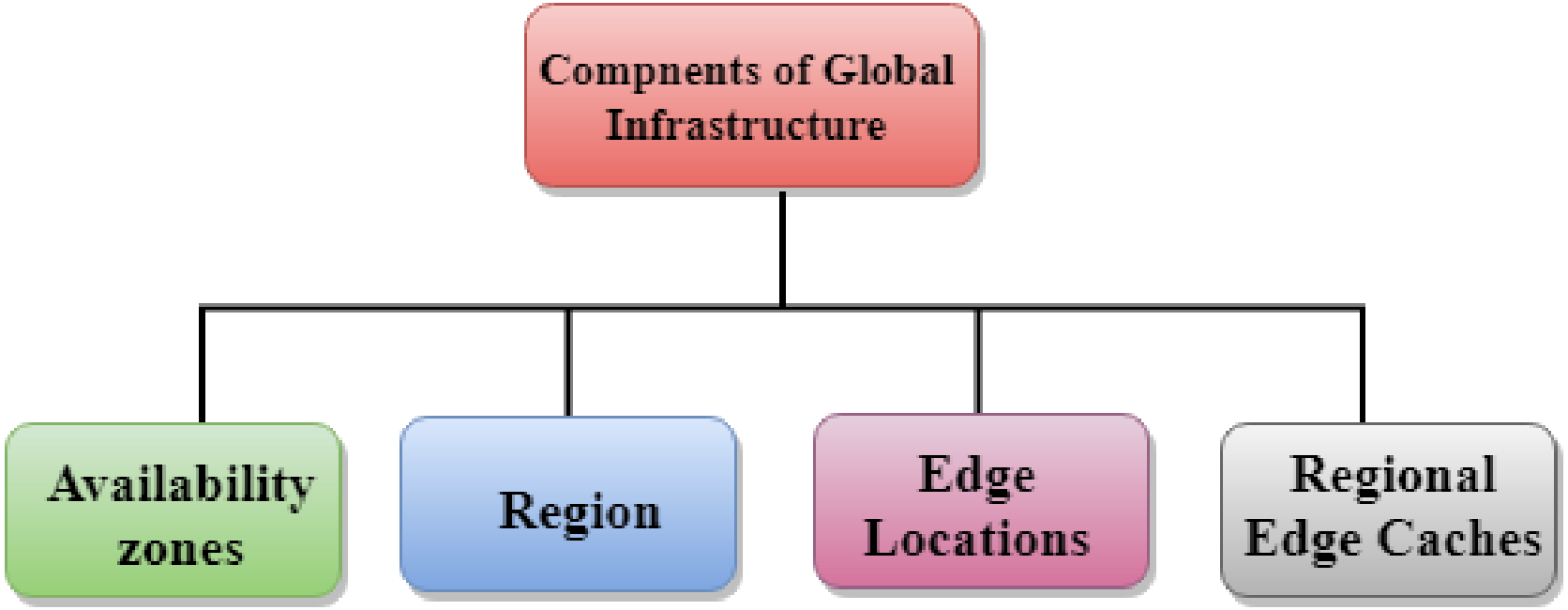
- Files are stored in Bucket. A bucket is like a folder available in S3 that stores the files.

- S3 is a universal namespace, i.e., the names must be unique globally. Bucket contains a DNS address. Therefore, the bucket must contain a unique name to generate a unique DNS address.

AWS Global Infrastructure

The following are the components that make up the AWS infrastructure:

- Availability Zones
- Region
- Edge locations
- Regional Edge Caches



AWS Global Infrastructure

Availability zone as a Data Center

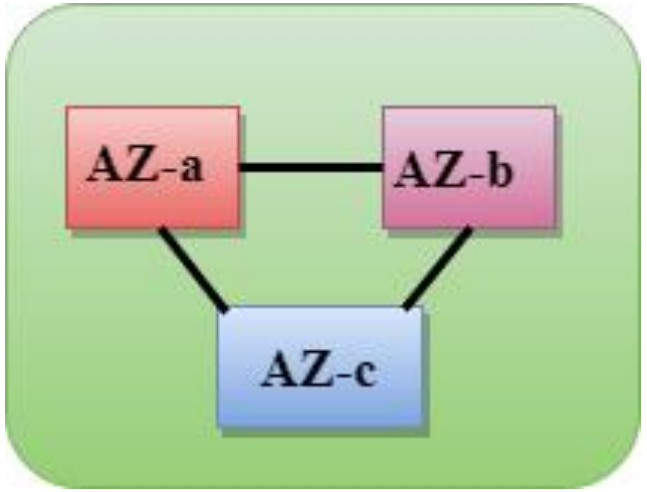
An availability zone is a facility that can be somewhere in a country or in a city. Inside this facility, i.e., Data Centre, we can have multiple servers, switches, load balancing, firewalls. The things which interact with the cloud sits inside the data centers.

An availability zone can be a several data centers, but if they are close together, they are counted as 1 availability zone.

Region

A region is a geographical area. Each region consists of 2 more availability zones.

A region is a collection of data centers which are completely isolated from other regions.



AWS Global Infrastructure

Edge Locations

Edge locations are the endpoints for AWS used for caching content.

Edge locations consist of CloudFront, Amazon's Content Delivery Network (CDN).

Edge location is not a region but a small location that AWS have. It is used for caching the content.

Edge locations are mainly located in most of the major cities to distribute the content to end users with reduced latency.

For example, some user accesses your website from Singapore; then this request would be redirected to the edge location closest to Singapore where cached data can be read.

AWS Global Infrastructure

Regional Edge Cache

Regional Edge cache lies between CloudFront Origin servers and the edge locations.

A regional edge cache has a large cache than an individual edge location.

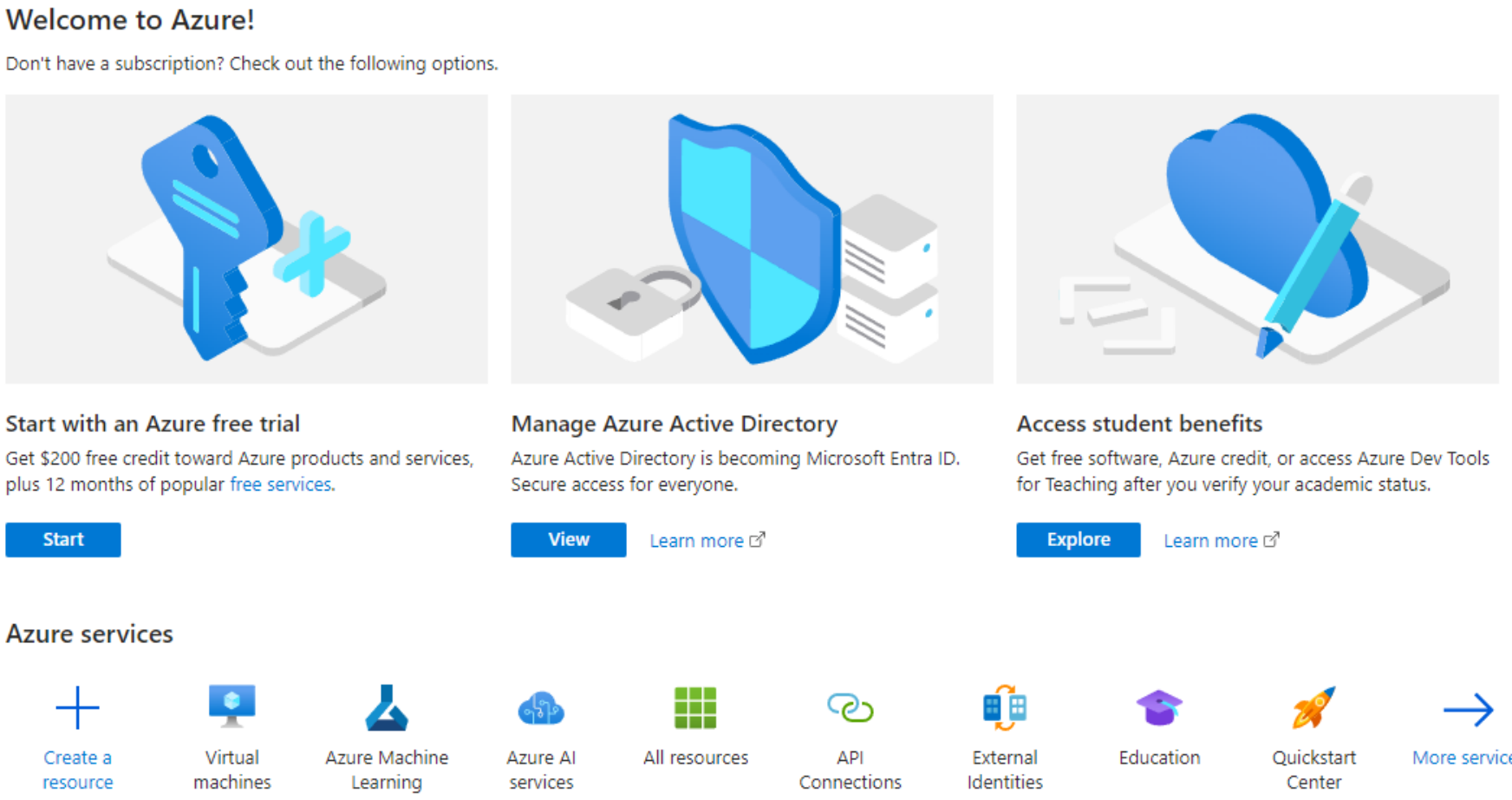
Data is removed from the cache at the edge location while the data is retained at the Regional Edge Caches.

When the user requests the data, then data is no longer available at the edge location.

Therefore, the edge location retrieves the cached data from the Regional edge cache instead of the Origin servers that have high latency.

Microsoft Azure

Azure Management Portal is an interface to manage the services and infrastructure launched in 2012. All the services and applications are displayed in it and it lets the user manage them.



Microsoft Azure - Compute Module

Step 1 − First, login in to your Azure account.
Step 2 − Click ‘New’ at the left bottom corner and drag your cursor to ‘Compute‘.

Create a Web App
Step 1 − Click Web App.
Step 2 − Click Quick Create and enter the URL and choose a service plan from the dropdown list

Windows Azure supports .Net, Java, PHP, Python, Node.js and Ruby.

There are several ways of publishing the code to Azure server. It can be published using FTP, FTPs, Microsoft Web Deploy technology.

Various source control tools such as GitHub, Dropbox and Codeplex can also be used to publish the code. It provides a very interactive interface to keep track of changes that have been published already and also unpublished changes.

Create a Virtual Machine

Step 1 − Click on ‘Virtual Machine’ from the list.
Step 2 − Then click ‘From Gallery’.
Step 3 − Choose the Operating System or Program you want to run.
Step 4 − Choose the configuration and fill in the details.

The Username and Password you set up here will be needed to access the virtual machine every time.

Step 5 − Once the machine is created you can connect to it by clicking on the connect icon displayed at the bottom of the screen. It will save a .rpd file on your machine as shown in the following image. Chose ‘save file’ on the screen and it will save in ‘downloads’ or the in the set location on your machine.

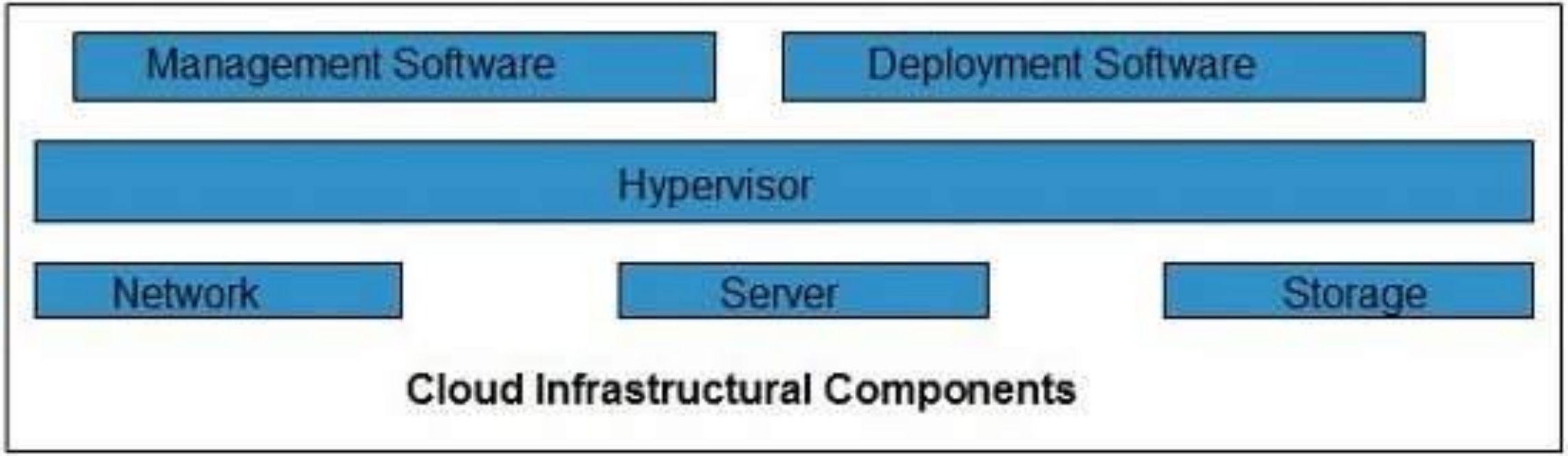
Step 6 − Open that .rpd file and you can connect to the VM by filling in the credentials

open source cloud

- **Cloud Foundry**
- **WSO2**
- **Cloudify**
- **OpenShift**
- **Stackato**
- **Alibaba**
- https://www.openstack.org/
- https://opennebula.io/

Cloud Computing Infrastructure

Cloud infrastructure consists of servers, storage devices, network, cloud management software, deployment software, and platform virtualization.



Hypervisor

Hypervisor is a **firmware** or **low-level program** that acts as a Virtual Machine Manager. It allows to share the single physical instance of cloud resources between several tenants.

Management Software

It helps to maintain and configure the infrastructure.

Deployment Software

It helps to deploy and integrate the application on the cloud.

Network

It allows to connect cloud services over the Internet. It is also possible to deliver network as a utility over the Internet, which means, the customer can customize the network route and protocol.

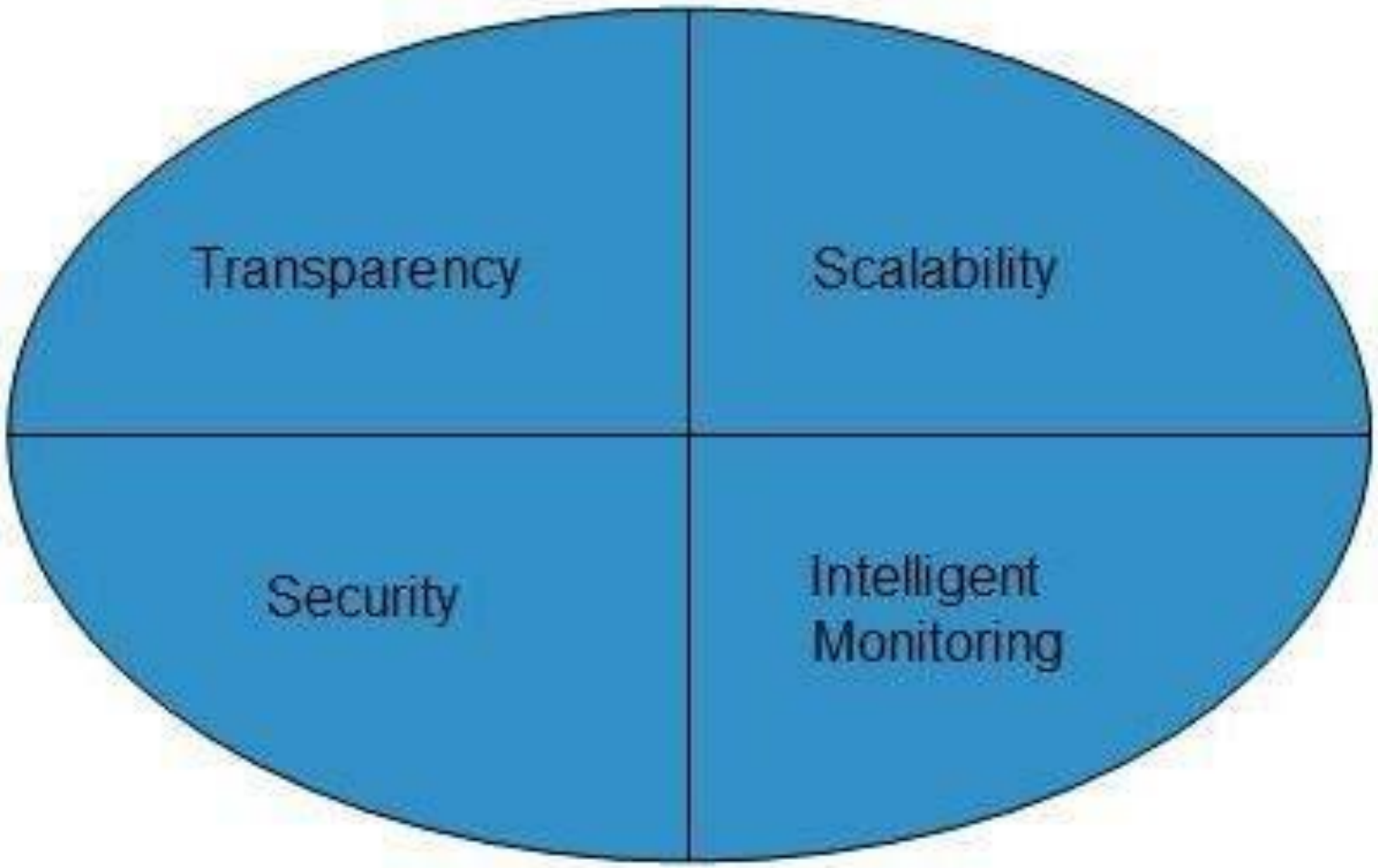
Server

The **server** helps to compute the resource sharing and offers other services such as resource allocation and de-allocation, monitoring the resources, providing security etc.

Storage

Cloud keeps multiple replicas of storage. If one of the storage resources fails, then it can be extracted from another one, which makes cloud computing more reliable.

Infrastructural Constraints



Transparency

Virtualization is the key to share resources in cloud environment. But it is not possible to satisfy the demand with single resource or server. Therefore, there must be transparency in resources, load balancing and application, so that we can scale them on demand.

Scalability

Scaling up an application delivery solution is not that easy as scaling up an application because it involves configuration overhead or even re-architecting the network. So, application delivery solution is need to be scalable which will require the virtual infrastructure such that resource can be provisioned and de-provisioned easily.

Intelligent Monitoring

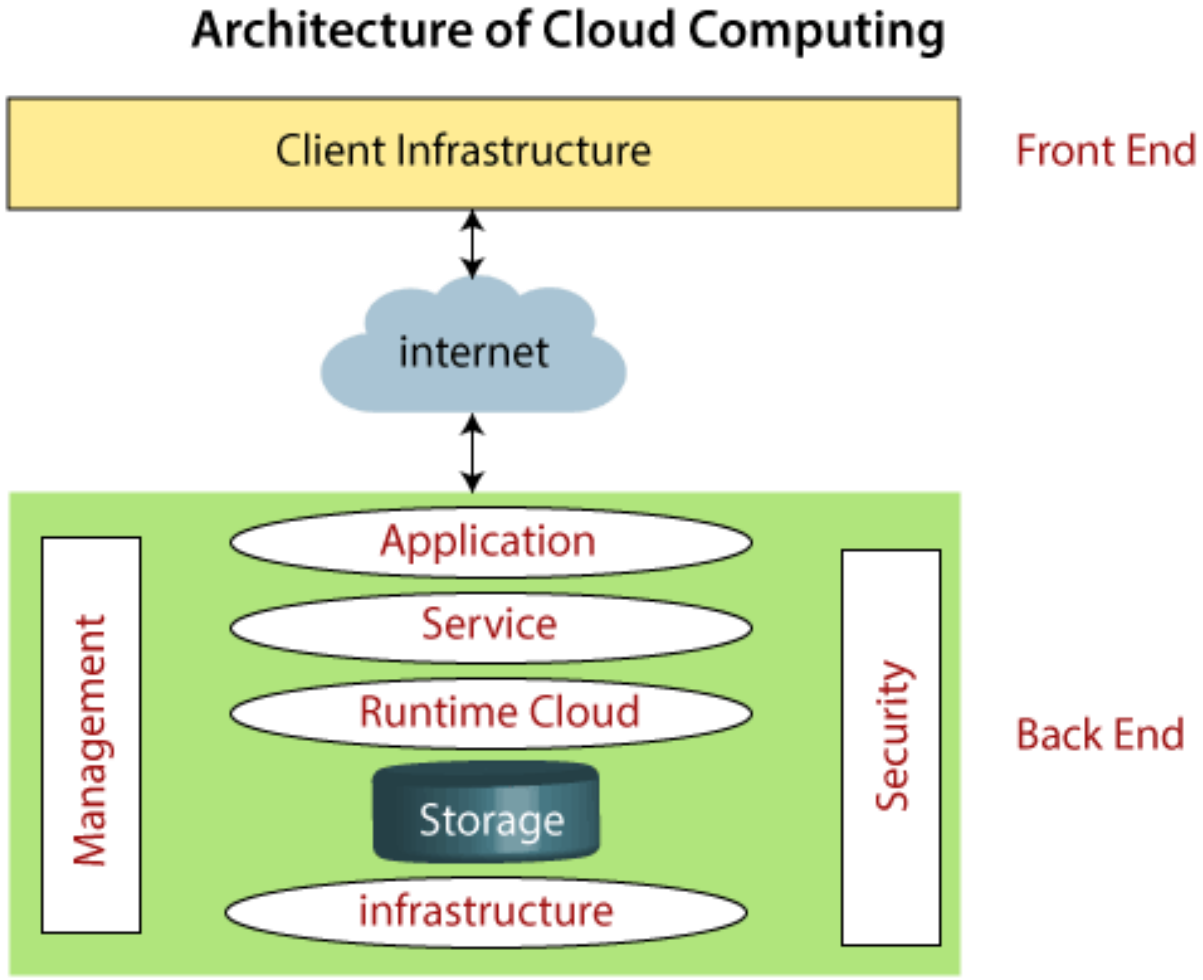
To achieve transparency and scalability, application solution delivery will need to be capable of intelligent monitoring.

Security

The mega data center in the cloud should be securely architected. Also the control node, an entry point in mega data center, also needs to be secure.

Cloud Computing Architecture

Cloud computing architecture is a combination of **service-oriented architecture** and **event-driven architecture**.



Front End

The front end is used by the client. It contains client-side interfaces and applications that are required to access the cloud computing platforms. The front end includes web servers (including Chrome, Firefox, internet explorer, etc.), thin & fat clients, tablets, and mobile devices.

Back End

The back end is used by the service provider. It manages all the resources that are required to provide cloud computing services. It includes a huge amount of data storage, security mechanism, virtual machines, deploying models, servers, traffic control mechanisms, etc.

Components of Cloud Computing Architecture

1. Client Infrastructure

Client Infrastructure is a Front end component. It provides GUI (Graphical User Interface) to interact with the cloud.

2. Application

The application may be any software or platform that a client wants to access.

3. Service

A Cloud Services manages that which type of service you access according to the client's requirement. (SaaS, PaaS, IaaS)

4. Runtime Cloud

Runtime Cloud provides the execution and runtime environment to the virtual machines.

5. Storage

Storage is one of the most important components of cloud computing. It provides a huge amount of storage capacity in the cloud to store and manage data.

Components of Cloud Computing Architecture

6. Infrastructure

It provides services on the host level, application level, and network level. Cloud infrastructure includes hardware and software components such as servers, storage, network devices, virtualization software, and other storage resources that are needed to support the cloud computing model.

7. Management

Management is used to manage components such as application, service, runtime cloud, storage, infrastructure, and other security issues in the backend and establish coordination between them.

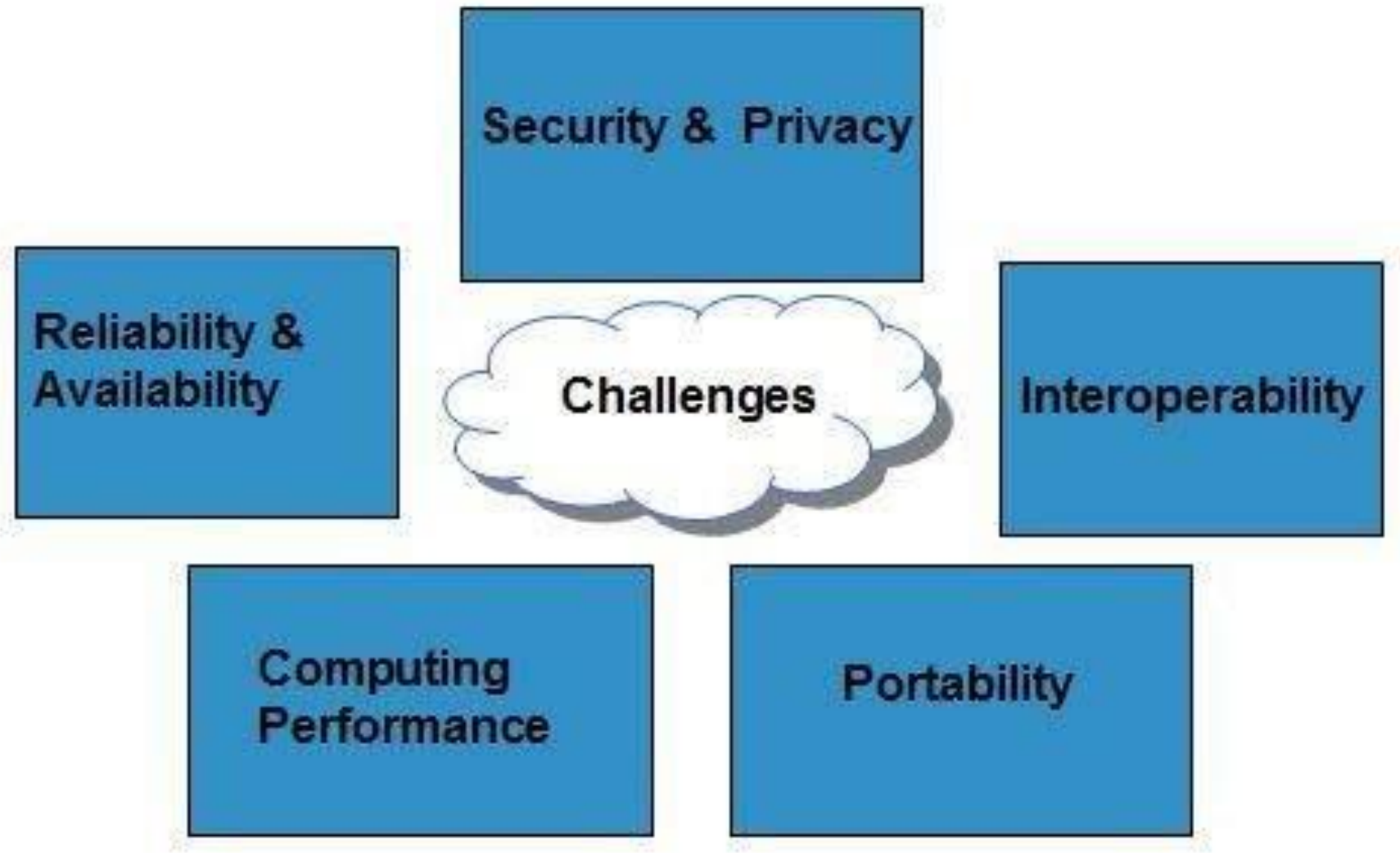
8. Security

Security is an in-built back end component of cloud computing. It implements a security mechanism in the back end.

9. Internet

The Internet is medium through which front end and back end can interact and communicate with each other.

Cloud Computing Challenges



Cloud Computing Challenges

Security and Privacy

Security and Privacy of information is the biggest challenge to cloud computing. Security and privacy issues can be overcome by employing encryption, security hardware and security applications.

Portability

This is another challenge to cloud computing that applications should easily be migrated from one cloud provider to another. There must not be vendor lock-in. However, it is not yet made possible because each of the cloud provider uses different standard languages for their platforms.

Interoperability

It means the application on one platform should be able to incorporate services from the other platforms. It is made possible via web services, but developing such web services is very complex.

Cloud Computing Challenges

Computing Performance

Data intensive applications on cloud requires high network bandwidth, which results in high cost. Low bandwidth does not meet the desired computing performance of cloud application.

Reliability and Availability

It is necessary for cloud systems to be reliable and robust because most of the businesses are now becoming dependent on services provided by third-party.

Can you list out the other challenges?