

Cloud Platforms in Industry

Table 9.1 Some Example Cloud Computing Offerings

Vendor/Product	Service Type	Description
Amazon Web Services	IaaS, PaaS, SaaS	Amazon Web Services (AWS) is a collection of Web services that provides developers with compute, storage, and more advanced services. AWS is mostly popular for IaaS services and primarily for its elastic compute service EC2.
Google AppEngine	PaaS	Google AppEngine is a distributed and scalable runtime for developing scalable Web applications based on Java and Python runtime environments. These are enriched with access to services that simplify the development of applications in a scalable manner.
Microsoft Azure	PaaS	Microsoft Azure is a cloud operating system that provides services for developing scalable applications based on the proprietary Hyper-V virtualization technology and the .NET framework.
SalesForce.com and Force.com	SaaS, PaaS	SalesForce.com is a Software-as-a-Service solution that allows prototyping of CRM applications. It leverages the Force.com platform, which is made available for developing new components and capabilities for CRM applications.
Heroku	PaaS	Heroku is a scalable runtime environment for building applications based on Ruby.
RightScale	IaaS	Rightscale is a cloud management platform with a single dashboard to manage public and hybrid clouds.

Amazon web services (AWS)

- Amazon Web Services (AWS) provides elastic infrastructure scalability, messaging, and data storage.
- The platform is accessible through SOAP (Simple Object Access Protocol) or RESTful (Representational State Protocol)Web service interfaces and provides a Web-based console.
- Expenses computed on a pay-as-you-go basis
- Amazon Elastic Compute (EC2) and Amazon Simple Storage Service(S3).

Services available in the AWS ecosystem.

Compute Services

Amazon Elastic Compute Cloud (EC2)

Amazon Elastic MapReduce

AWS Elastic Beanstalk

AWS Cloudformation

Autoscaling

Storage Services

Amazon Simple Storage Service (S3)

Amazon Elastic Block Store (EBS)

Amazon ElastiCache

Amazon SimpleDB

Amazon Relational Database Service (RDS)

Amazon CloudFront

Amazon Import/Export

Communication Services

Amazon Simple Queue Service (SQS)

Amazon Simple Notification Service (SNS)

Amazon Simple Email Service (SES)

Amazon Route 53

Amazon Virtual Private Cloud (VPC)

Amazon Direct Connect

Amazon Elastic Load Balancing

Additional Services

Amazon GovCloud

Amazon CloudWatch

Amazon Flexible Payment Service (FPS)

Amazon DevPay

Amazon Fulfillment Web Service (FWS)

Amazon Mechanical Turk

Alexa Web Information Service

Alexa Top Sites

Amazon AWS Platform

AWS: Compute services

- The fundamental service in this space is Amazon EC2, which delivers an IaaS solution.
- Amazon EC2 allows deploying servers in the form of virtual machines created as instances of a specific image.
- **Amazon machine images:**
 - Amazon Machine Images(AMIs) are templates from which it is possible to create a virtual machine.
 - They are stored in Amazon S3 and identified by a unique identifier in the form of **ami-xxxxxx**.
 - An AMI contains a physical file system layout with a predefined operating system installed.
 - Amazon Ram disk Image (ARI, id: ari-yyyyyyy) and the Amazon Kernel Image (AKI, id: aki-zzzzzz),
 - Once an AMI is created, it is stored in an S3 bucket

AWS: EC2 instances

- EC2 instances represent **virtual machines**, **created using AMI** as templates, by selecting the number of cores, their computing power, and the installed memory.
- The processing power is expressed in terms of virtual cores and EC2 Compute Units (ECUs).
- One ECU is defined as giving the same performance as a 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor.

Available configurations for EC2 instances

Instance Type	ECU	Platform	Memory	Disk Storage	Price (U.S. East) (USD/hour)
Standard instances					
Small	1(1 × 1)	32 bit	1.7 GB	160 GB	\$0.085 Linux \$0.12 Windows
Large	4(2 × 2)	64 bit	7.5 GB	850 GB	\$0.340 Linux \$0.48 Windows
Extra Large	8(4 × 2)	64 bit	15 GB	1,690 GB	\$0.680 Linux \$0.96 Windows
Micro instances					
Micro	< = 2	32/64 bit	613 MB	EBS Only	\$0.020 Linux \$0.03 Windows
High-Memory instances					
Extra Large	6.5(2 × 3.25)	64 bit	17.1 GB	420 GB	\$0.500 Linux \$0.62 Windows
Double Extra Large	13(4 × 3.25)	64 bit	34.2 GB	850 GB	\$1.000 Linux \$1.24 Windows
Quadruple Extra Large	26(8 × 3.25)	64 bit	68.4 GB	1,690 GB	\$2.000 Linux \$2.48 Windows
High-CPU instances					
Medium	5(2 × 2.5)	32 bit	1.7 GB	350 GB	\$0.170 Linux \$0.29 Windows
Extra Large	20(8 × 2.5)	64 bit	7 GB	1,690 GB	\$0.680 Linux \$1.16 Windows
Cluster instances					
Quadruple Extra Large	33.5	64 bit	23 GB	1,690 GB	\$1.600 Linux \$1.98 Windows
Cluster GPU instances					
Quadruple Extra Large	33.5	64 bit	22 GB	1,690 GB	\$2.100 Linux \$2.60 Windows

AWS:EC2 environment

- The EC2 environment is in charge of **allocating addresses, attaching storage volumes, and configuring security** in terms of access control and network connectivity.
- instances are created with an internal IP address, which makes them capable of communicating within the EC2 network and accessing the Internet as clients.
- EC2 instances are also given a domain name that generally is in the form ***ec2-xxx-xxx-xxx.compute-x.amazonaws.com***,
- where xxx-xxx-xxx normally represents the four parts of the external IP address separated by a dash.
- compute-x gives information about the availability zone where instances are deployed.

AWS: EC2:Advanced compute services

- **AWS Cloud Formation:** Cloud Formation introduces the concepts of templates, which are formatted text files that **describe the resources needed to run an application.**
- **Services:** S3, SimpleDB, SQS, SNS, Route 53, Elastic Beanstalk
- **AWS elastic beanstalk:** AWS Elastic Beanstalk constitutes a simple and easy way to **package applications and deploy them on the AWS Cloud.**
- **Amazon elastic MapReduce :** cloud computing platform for MapReduce applications. It utilizes Hadoop as the MapReduce engine,

AWS: Storage services: Simple Storage Service(S3): **S3 Key Concepts**

- S3 has been designed to provide a simple storage service that's accessible through a Representational State Transfer (REST) interface.
- The storage is organized in a two-level hierarchy .(**Buckets, Objects**)
- Stored objects cannot be manipulated (renaming, modifying, or relocating) like standard files.
- Content is not immediately available to users .
- Requests will occasionally fail.

AWS:S3 Key Concepts: **Resource naming**

- Buckets, objects are represented by uniform resource identifiers(URIs) under the **s3.amazonaws.com** domain.
- **three different ways of addressing a bucket:**
 - Canonical form: http://s3.amazonaws.com/bukect_name/
 - Sub domain form: <http://bucketname.s3.amazonaws.com/>
 - Virtual hosting form: <http://bucket-name.com/>
- **Object ACL:**
[http://s3.amazonaws.com/bukect_name/object_name
?acl](http://s3.amazonaws.com/bukect_name/object_name?acl)
- **Bucket server logging:**
http://s3.amzonaws.com/bucket_name?logging

AWS:S3 Key Concepts: Buckets

- A bucket is a container of objects.
- Buckets are top- level elements of the S3 storage architecture and do not support nesting. That is, it is not possible to create “subbuckets”.
- A bucket is located in a specific geographic location.
- Once a bucket is created, all the objects that belong to the bucket will be stored in the same availability zone of the bucket
- Users create a bucket by sending a PUT request to <http://s3.amazonaws.com/> with the name of the bucket.
- The content of a bucket can be listed by sending a GET request specifying the name of the bucket.
- Once created, the bucket cannot be renamed or relocated.
- Deletion of a bucket is performed by a DELETE request .

AWS:S3 Key Concepts: Objects and metadata

- Objects constitute the content elements stored in S3.
- An object is identified by a name that needs to be unique within the bucket.
- The name cannot be longer than 1,024 bytes when encoded in UTF-8, and it allows almost any character.
- Users create an object via a PUT request that specifies the name of the object together with the bucket name, its contents.
- Maximum size of an object is 5 GB.
- It cannot be modified, renamed, or moved into another bucket.
- Deleting an object is performed via a DELETE request.
- Objects can be tagged with metadata.

AWS:S3 Key Concepts: Access control and security

- Amazon S3 allows controlling the access to buckets and objects by means of Access Control Policies(ACPs).
- An ACP is a set of grant permissions are attached to a resource expressed by means of an XML configuration file.
- A policy allows defining up to 100 access rules.
- READ allows the grantee to retrieve an object.
- WRITE allows the grantee to add an object to a bucket .
- READ_ACP allows the grantee to read the ACP of a resource.
- WRITE_ACP allows the grantee to modify the ACP of a resource.
- FULL_CONTROL grants all of the preceding permissions

AWS:S3 Key Concepts: **Amazon elastic block store**

- The Amazon Elastic Block Store (EBS) allows AWS users to provide EC2 instances with persistent storage.
- They accommodate up to 1 TB of space.
- Currently, Amazon charges \$0.10/GB/month of allocated storage and \$0.10 per 1 million requests made to the volume.

AWS:S3 Key Concepts: **Amazon ElastiCache**

- ElastiCache is an implementation of an elastic in-memory cache based on a cluster of EC2.
- It provides fast data access from other EC2 instances through a Memcached-compatible protocol.

AWS:S3 Key Concepts: Structured storage solutions

- RDBMS have been the common data back-end for a wide range of applications.
- Amazon provides applications with structured storage: **Amazon Relational Data Storage(RDS)**, and **Amazon SimpleDB**.

Table 9.5 Amazon SimpleDB Data Transfer Charges, 2011–2012

Instance Type	Price (U.S. East) (USD)
Data Transfer In All data transfer in	\$0.000
Data Transfer Out 1st GB/month	\$0.000
Up to 10 TB/month	\$0.120
Next 40 TB/month	\$0.090
Next 100 TB/month	\$0.070
Next 350 TB/month	\$0.050
Next 524 TB/month	Special arrangements
Next 4 PB/month	Special arrangements
Greater than 5 PB/month	Special arrangements

AWS:S3 Key Concepts: **Amazon CloudFront**

- CloudFront is an implementation of a **content delivery network**.
- AWS provides users with simple Web service APIs to manage CloudFront.
- DNS domain under the Cloudfront.net domain name(i.e.,my-distribution.Cloudfront.net).
- The content that can be delivered through CloudFront is static (HTTP and HTTPS) or streaming (Real Time Messaging Protocol, or RTMP).

AWS: Communication services

- Amazon provides the communication among existing applications and services residing within the AWS infrastructure.
- Two major categories: virtual networking and messaging.

AWS: Communication services: **Virtual networking**

- comprises a collection of services that allow AWS users to **control the connectivity to and between compute and storage services.**
- Amazon *Virtual Private Cloud(VPC)* and Amazon *DirectConnect* provide connectivity solutions.
- **VPC:** Prepared templates include **public subnets, isolated networks, private networks** accessing Internet through network address translation (NAT), and hybrid networks.
- Amazon Direct Connect allows AWS users to **create dedicated networks** between the user private network and Amazon Direct Connect locations, called ports.
- Amazon Route 53 implements **dynamic DNS** services that allow *AWS resources to be reached through domain names different from the amazon.com domain.*

AWS: Communication services:

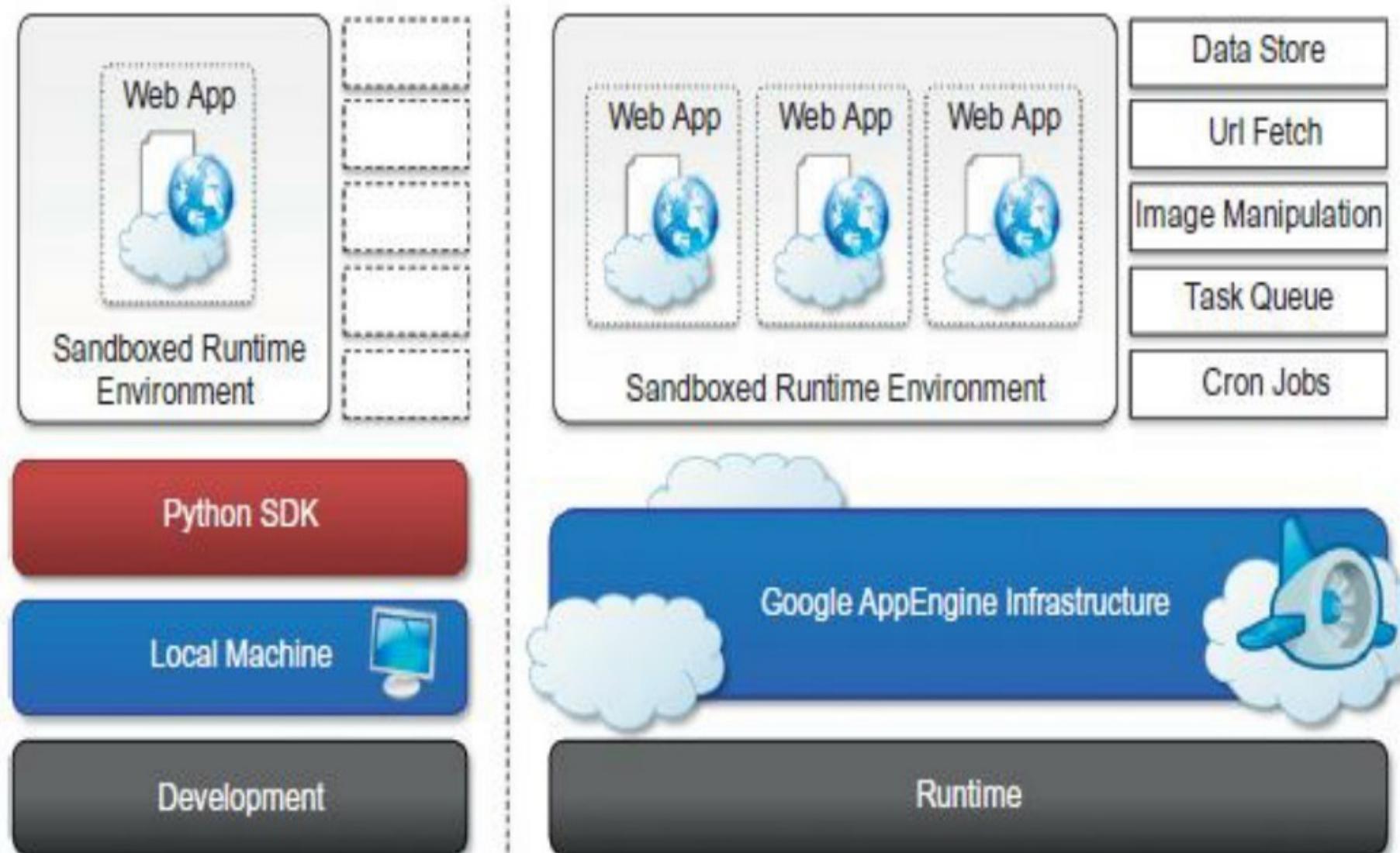
Messaging

- The three different types of messaging services offered are **Amazon Simple Queue Service (SQS)**, **Amazon Simple Notification Service(SNS)**, and **Amazon Simple Email Service(SES)**.
- **SQS**: exchanging messages between applications by means of **message queues**, hosted within the AWS infrastructure.
- **Amazon SNS** provides a **publish-subscribe** method for connecting heterogeneous applications.
- **Amazon SES** provides AWS users with a scalable email service.

Google AppEngine

- Google AppEngine is a PaaS implementation that provides services for developing and hosting scalable Web applications.

Google AppEngine: Architecture



Google AppEngine: Runtime environment

- The runtime environment represents the execution context of applications hosted on AppEngine.
- **Supported runtimes** : Java, Python, and Go.
- Java ServerPages(JSP),
- Python 2.5.2 interpreter
- Python Web application framework, called **webapp**, simplifying the development of Web applications.
- Go that is supported by AppEngine is r58.1.

Google AppEngine: **Storage**

- AppEngine provides in memory-cache, storage for semistructured data, and long-term storage for static data.
- **Static file servers:** Static data of web applications can be hosted on static file servers, since they are not frequently modified.
- **DataStore** : DataStore is a service that allows developers to **store semistructured data**.
- DataStore can be considered as a large object database in which to store objects that can be retrieved by a specified key.

Google AppEngine: Application services

- **UrlFetch** : provide developers with the capability of **retrieving a remote resource through HTTP/HTTPS** .
- **MemCache**: optimized for **fast access** and provides developers with a **volatile store** for the objects that are frequently accessed.
- **Mail and instant messaging**: AppEngine provides developers with the ability **to send and receive mails through Mail**. The service allows sending email on behalf of the application to specific user accounts.
- AppEngine provides also another way to communicate with the external world: the **Extensible Messaging and Presence Protocol (XMPP)** ,Google Talk .
- **Account management** : allowing developers to leverage Google account management by means of Google Accounts.
- **Image manipulation** : AppEngine allows applications to perform **image resizing, rotation, mirroring, and enhancement** by means of Image Manipulation.

Google AppEngine: Compute services

- Task queues:
 - Task Queues allow applications to submit a task for a later execution.
 - The service allows users to have up to 10 queues that can execute tasks at a configurable rate.
- Cron jobs
 - It is possible to *schedule the required operation at the desired time* by using the CronJobs service.

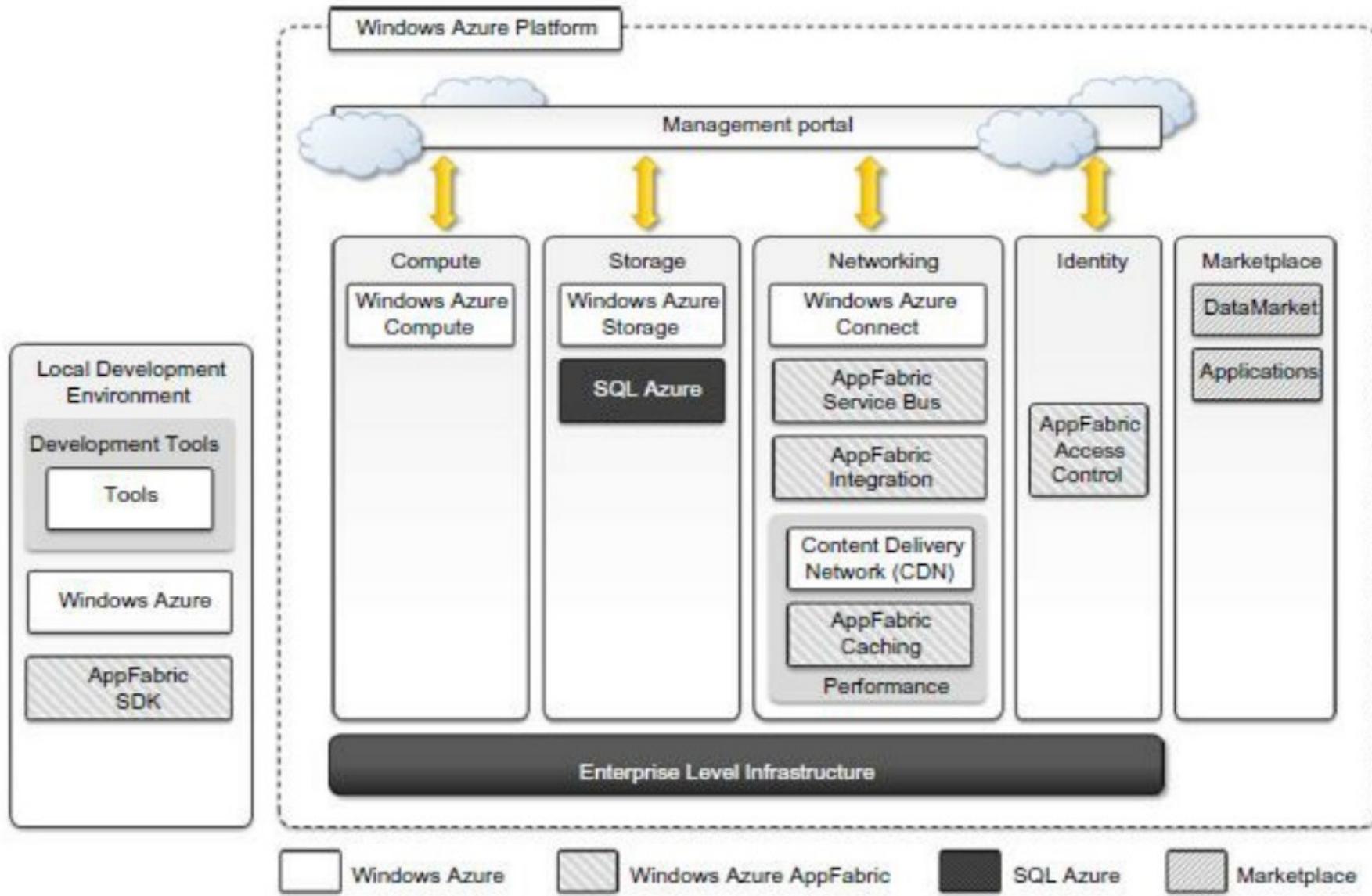
Google AppEngine: Application life cycle

- AppEngine provides support for almost all the phases characterizing the life cycle of an application: **testing and development, deployment, and monitoring.**
- Java SDK , Google Web Toolkit, and Google AppEngine plug-ins into Eclipse, servlet
- Python SDK:
 - The Python SDK allows developing Web applications for AppEngine with Python 2.5.
 - It provides a stand alone tool,called **GoogleApp EngineLauncher**, for managing Web applications locally and deploying them to AppEngine.

Microsoft Azure

- Microsoft Windows Azure is a cloud operating system built on top of Microsoft datacenters' infrastructure .
- **Services** range from **compute, storage, and networking** to **application connectivity, access control, and business intelligence**.
- integrates the scalability features into the common Microsoft technologies such as **Microsoft Windows Server 2008, SQL Server, and ASP.NET**.

Azure core concepts



Microsoft Azure :**Compute services**

- **Virtual Machine role:**
 - The Virtual Machine role is based on the Windows Hyper-V virtualization technology.

Microsoft Azure :**Compute services**

- **Virtual Machine role:**
 - The Virtual Machine role is based on the Windows Hyper-V virtualization technology.

Microsoft Azure :**Compute services**

- **Virtual Machine role:**
 - The Virtual Machine role is based on the Windows Hyper-V virtualization technology.

Microsoft Azure :**Storage services**

- **Queues** :Queue storage allows **applications** to communicate by exchanging messages through durable queues.
- Applications enter messages into a queue, and other applications can read them in a first-in, first-out (FIFO) style.

Microsoft Azure :**Storage services**

- **Queues** :Queue storage allows **applications** to communicate by exchanging messages through durable queues.
- Applications enter messages into a queue, and other applications can read them in a first-in, first-out (FIFO) style.

Microsoft Azure :**Storage services**

- **Queues** :Queue storage allows **applications** to communicate by exchanging messages through durable queues.
- Applications enter messages into a queue, and other applications can read them in a first-in, first-out (FIFO) style.

Microsoft Azure: AppFabric

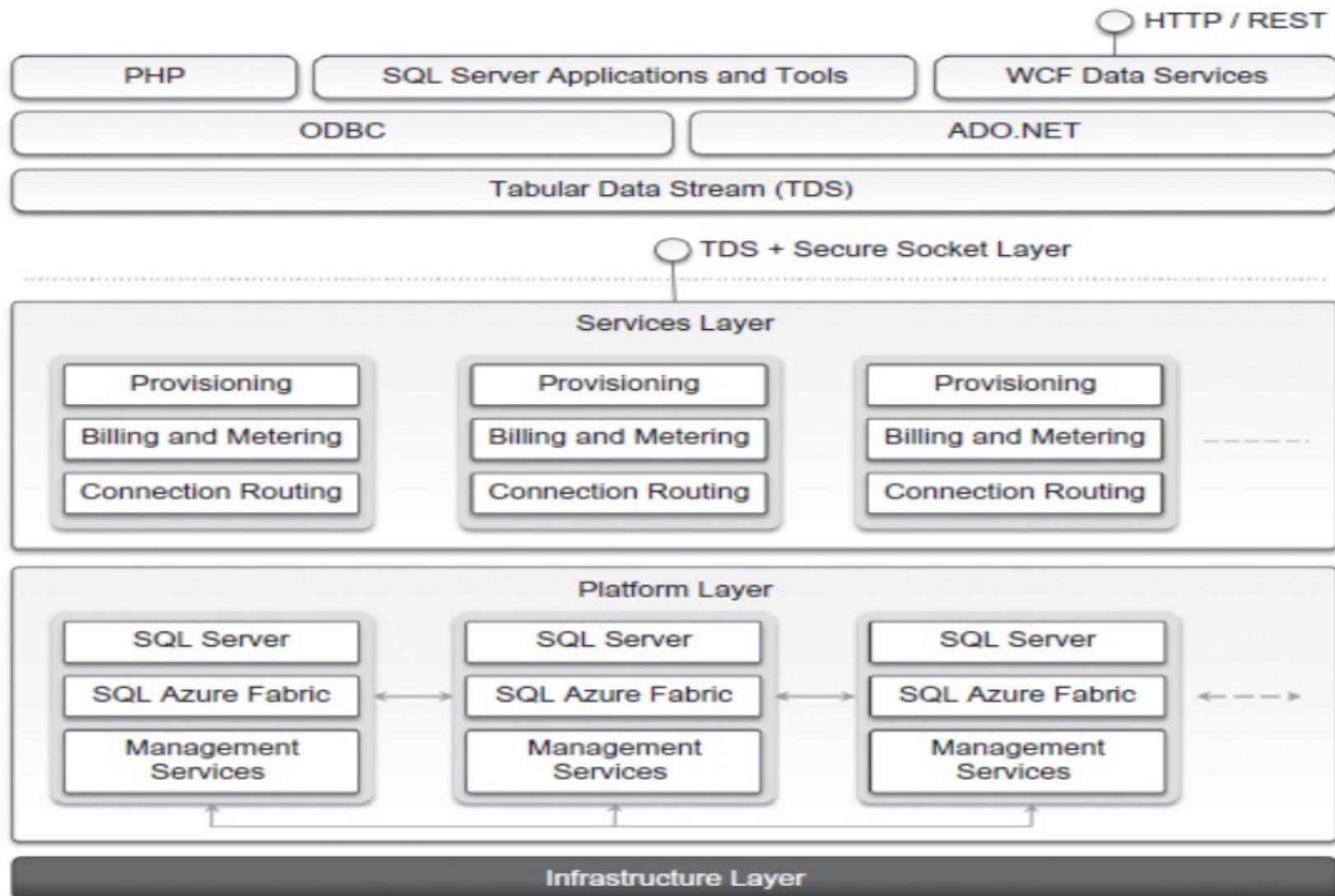
- AppFabric is a comprehensive middleware for **developing, deploying, and managing applications on the cloud.**
- AppFabric implements an optimized infrastructure supporting scaling out and high availability;
- it also provides **communication, authentication and authorization, and data access.**

Windows Azure virtual network

- Networking services for applications are offered under the name Windows Azure Virtual Network, which includes **Windows Azure Connect** and **Windows Azure Traffic Manager**.
- Windows Azure Connect allows easy setup of IP-based network connectivity.
- Windows Azure Traffic Manager provides **load-balancing** features for services listening to the HTTP or HTTPS ports and hosted on multiple roles.

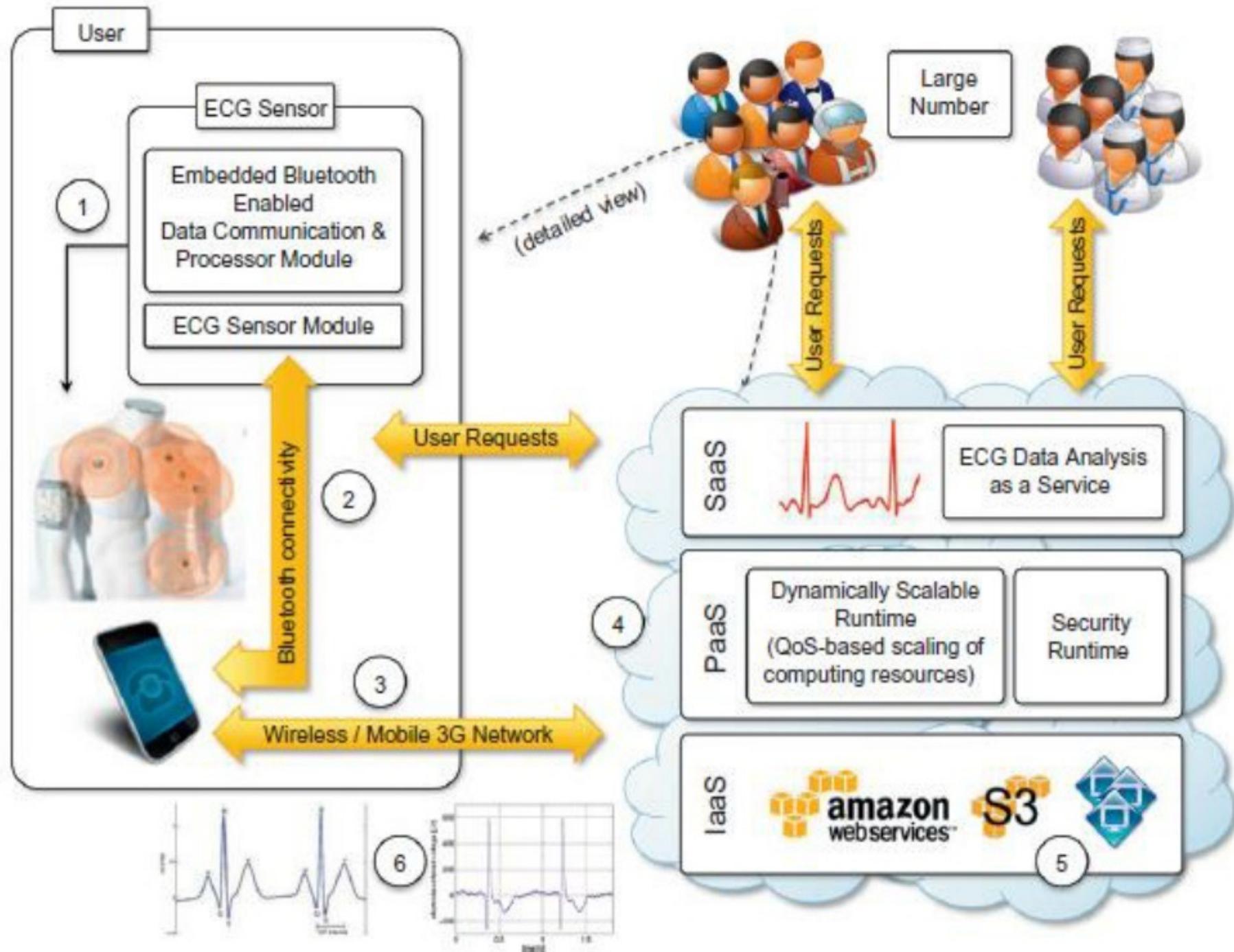
SQL Azure

- SQL Azure is a relational database service hosted on Windows Azure and built on the SQL Server technologies.



Cloud Applications: Healthcare: ECG analysis in the cloud

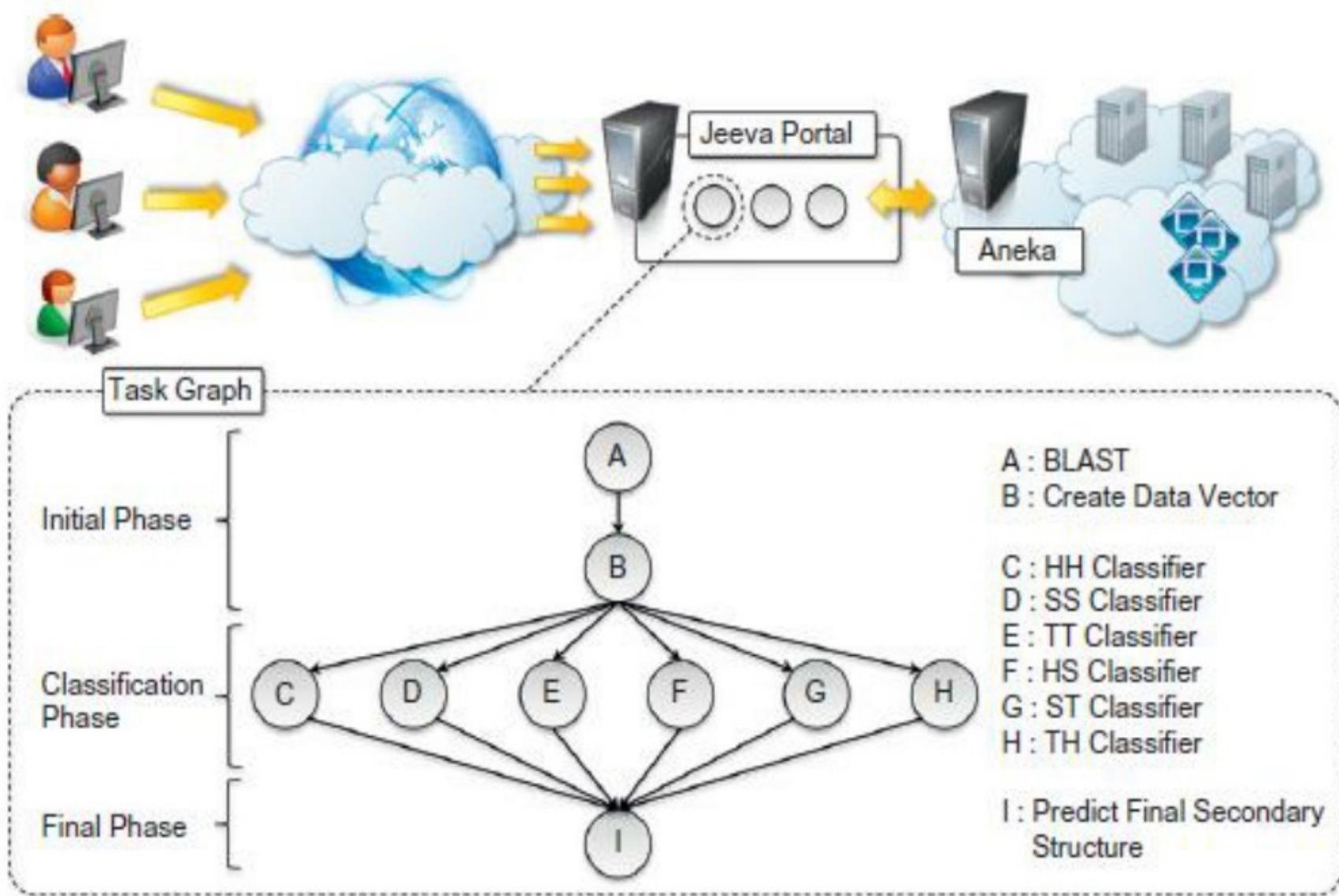
- ECG activity produces a specific waveform that is repeated over time and that represents the heartbeat.
- Cloud computing technologies allow the **remote monitoring of a patient's heartbeat data, data analysis** in minimal time.
- This way a patient at risk can be constantly monitored without going to a hospital for ECG analysis.



Biology: protein structure prediction

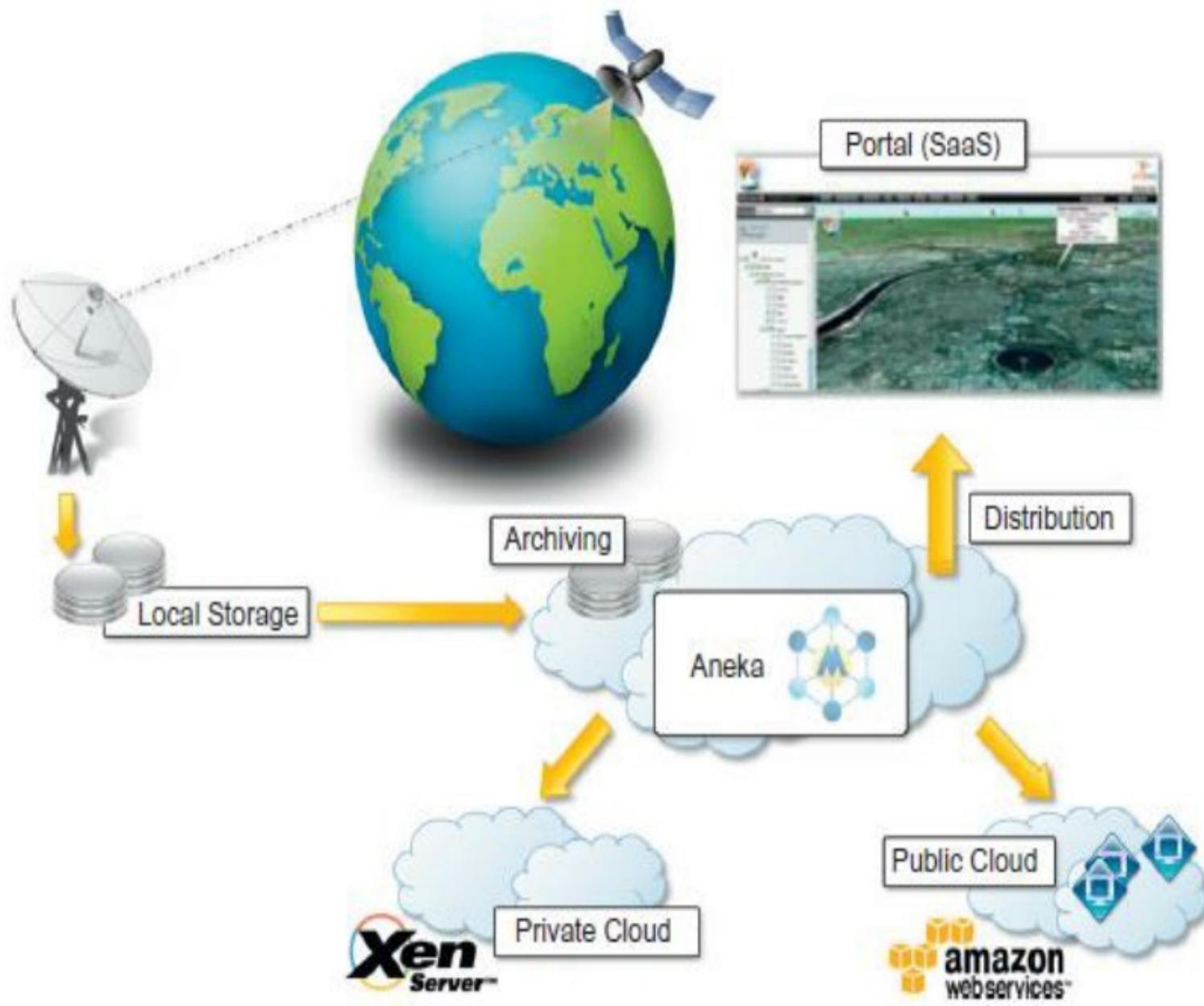
- Protein structure prediction is a computationally intensive task for the **design of new drugs for the treatment of diseases.**
- The **computational power required** for protein structure prediction can now be acquired on demand, without owning a **cluster, parallel and distributed computing facilities.**

Protein structure prediction: JEEVA



Geoscience: satellite image processing

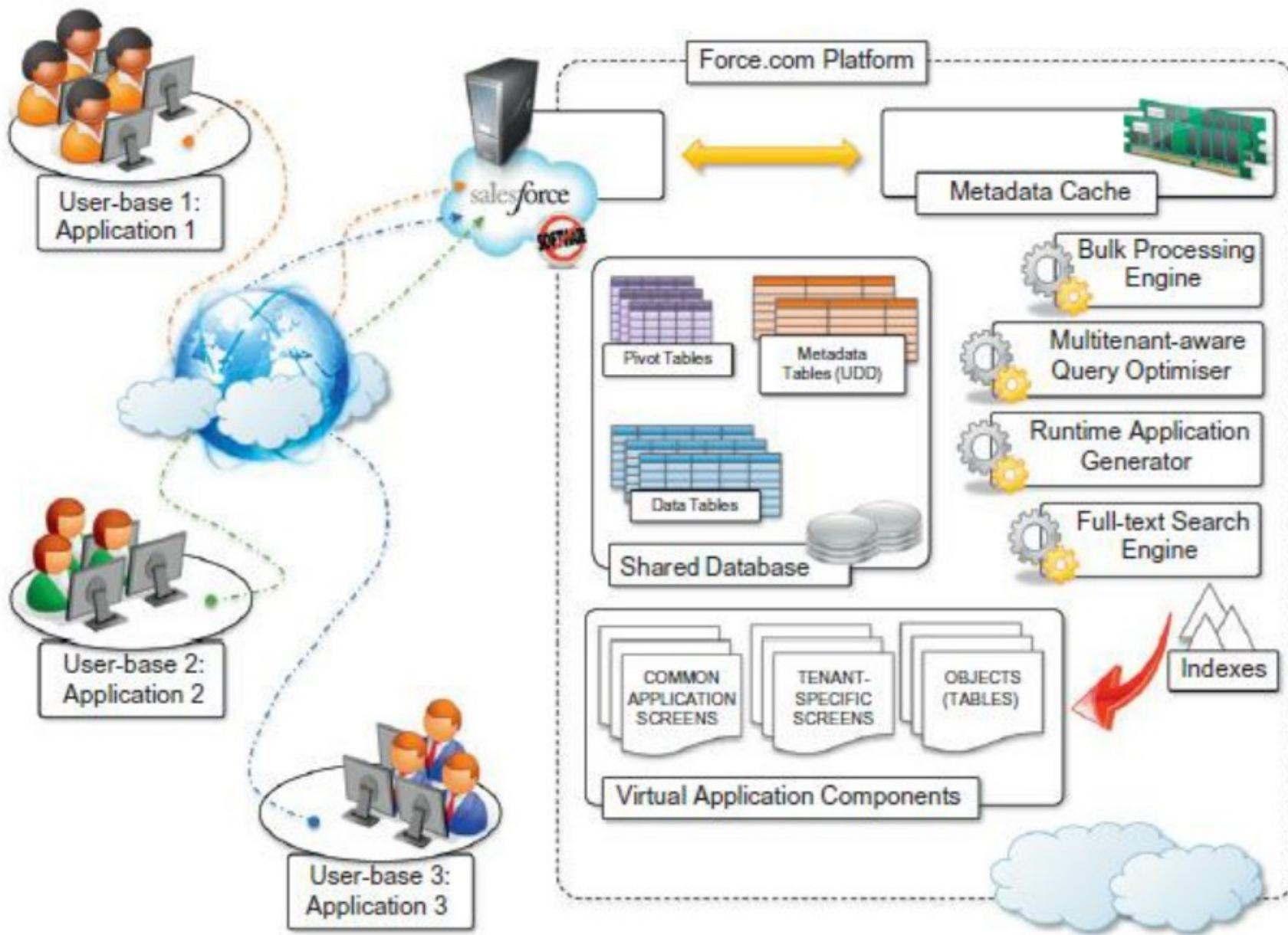
- Geoscience applications collect, produce, and analyze massive amounts of geospatial and non-spatial data.
- **Satellite remote sensing** generates hundreds of gigabytes of raw images that need to be further processed to become the basis of several different GIS products.
- Large images need to be moved 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.



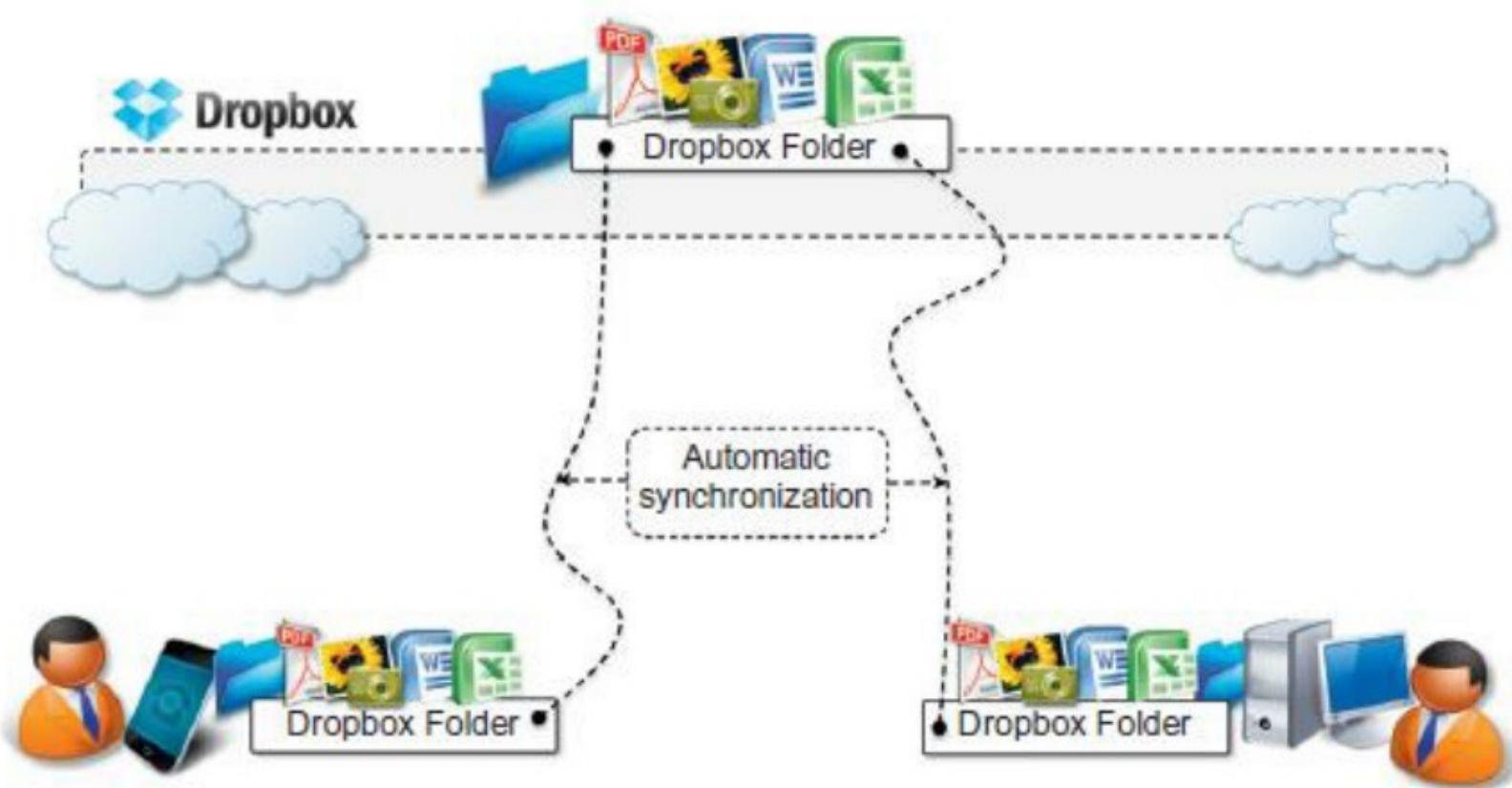
Business and consumer applications :

- Customer Relationship Management(CRM)
- Enterprise Resource Planning(ERP): finance and accounting, human resources, manufacturing, supply chain management, project management

Salesforce.com



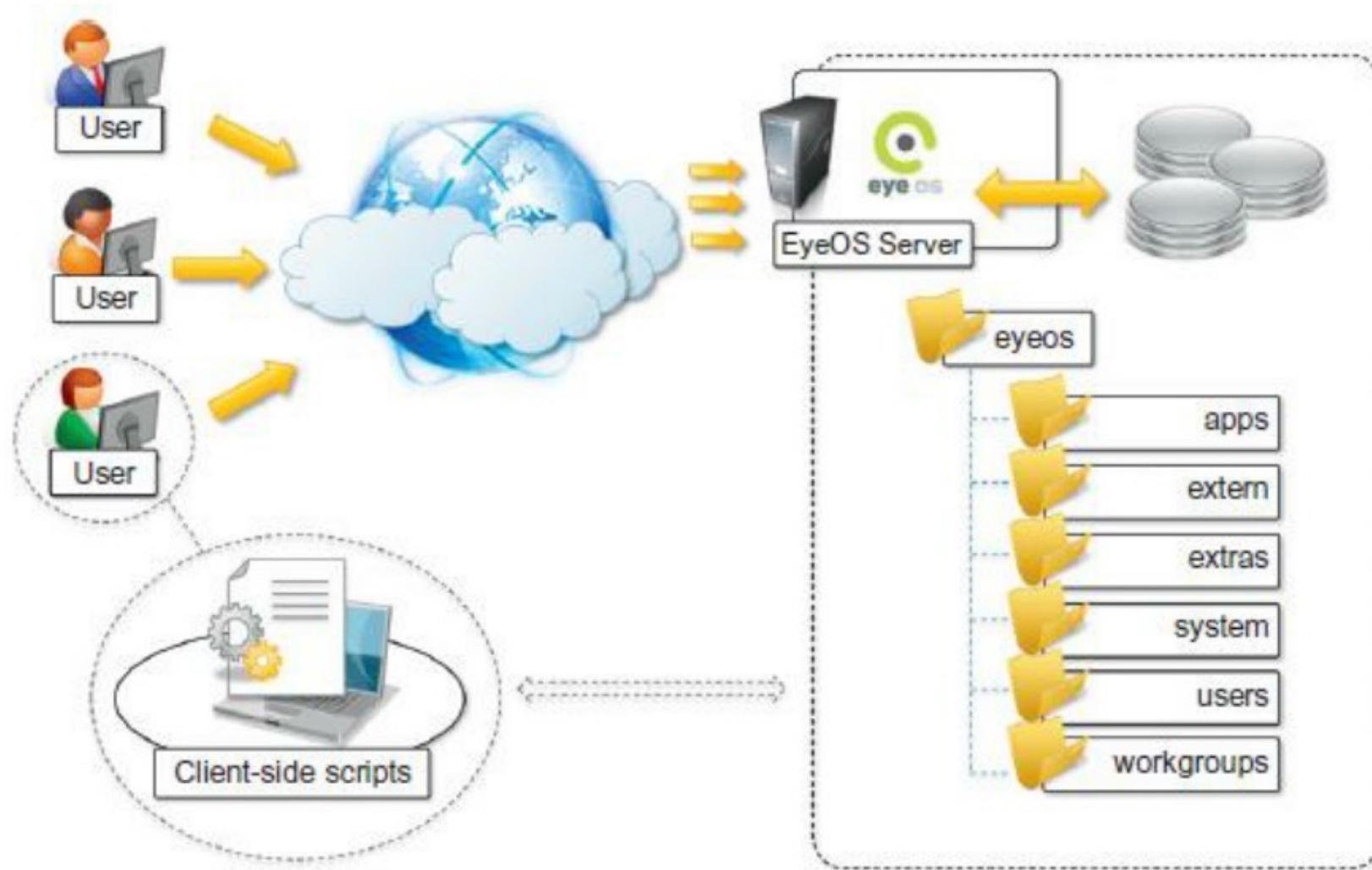
Dropbox and iCloud



Google docs

- GoogleDocs is a SaaS application that delivers the **basic office automation** capabilities with support for collaborative editing over the Web.
- Google Docs allows users to **create and edit text documents, spreadsheets, presentations, forms, and drawings**.

Cloud desktops: EyeOS and XIOS/3



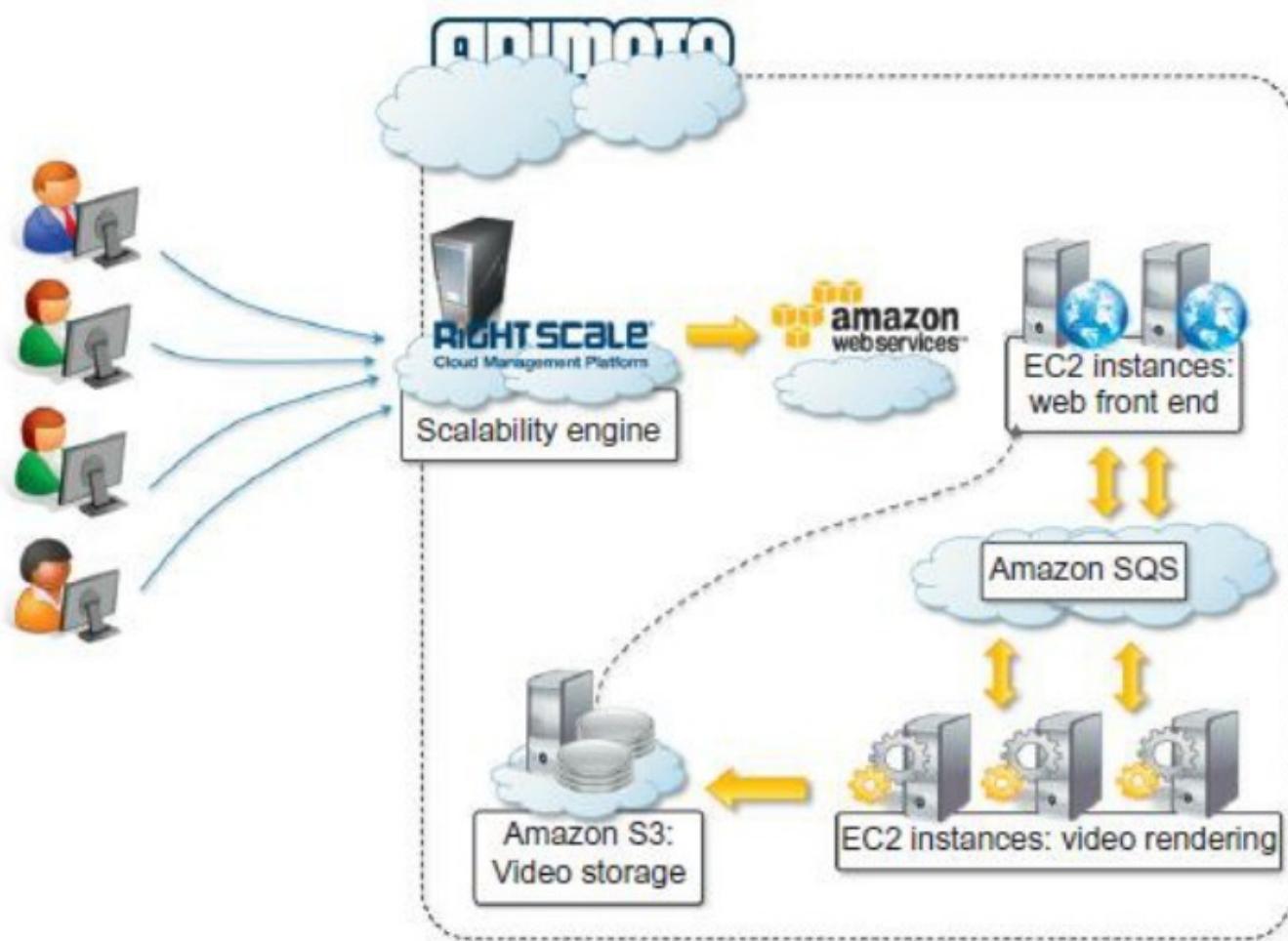
Social networking : Facebook

- Facebook is based on LAMP (Linux, Apache, MySQL, and PHP).

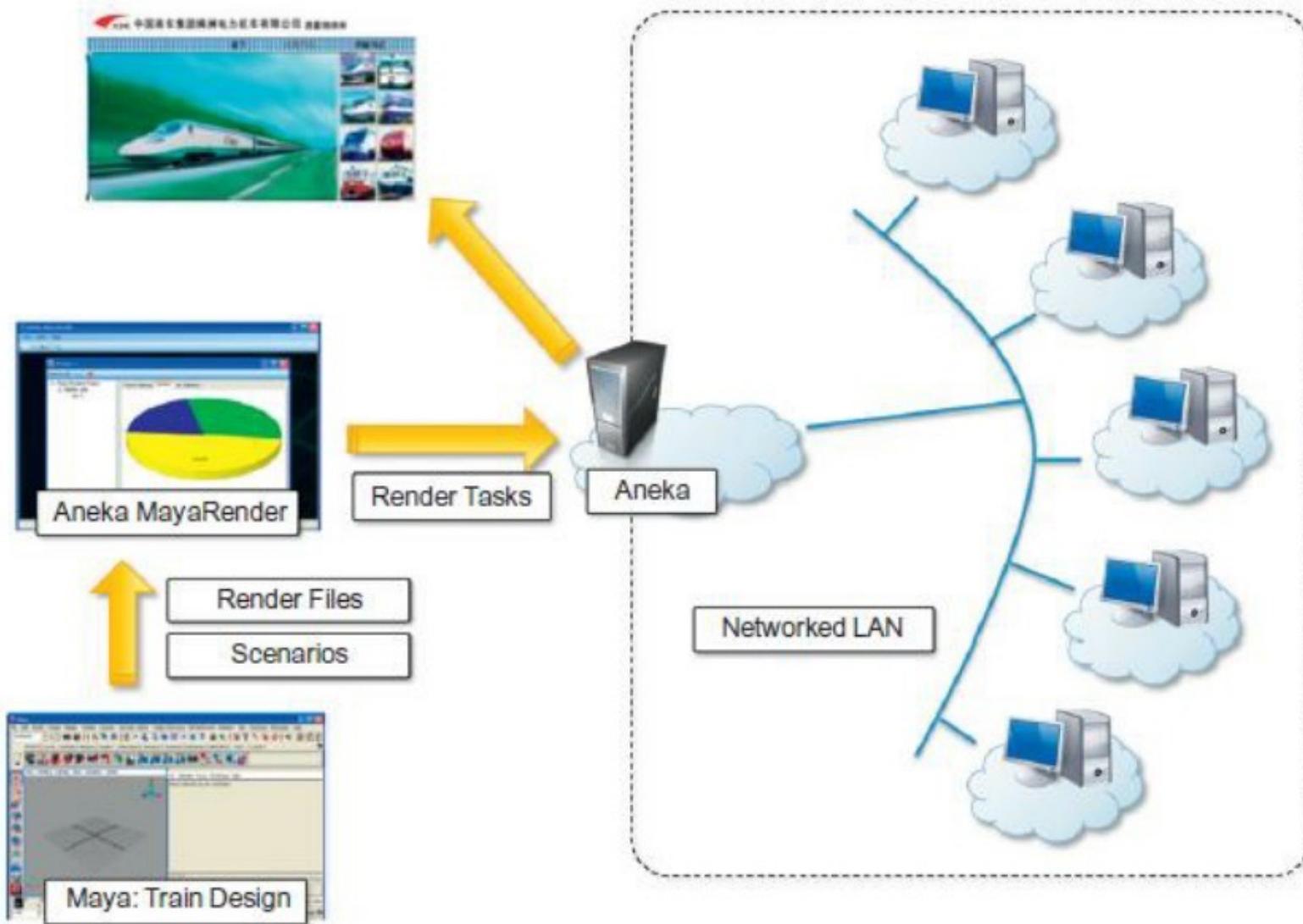
Media applications

- **Animoto:** The Website provides users with a very straightforward interface for quickly creating videos out of images, music, and video fragments submitted by users.
- A proprietary artificial intelligence (AI) engine, which **selects the animation and transition effects according to pictures and music**, drives the rendering operation.

Animoto reference architecture.



Maya rendering with Aneka



- Video encoding on the cloud: Encoding.com
- Multiplayer online gaming