Chap-05

Cloud Applications

- Cloud computing has gained huge popularity in industry due to its ability to host applications for which the services can be delivered to consumers rapidly at minimal cost.
- Will discuss applications from a range of domains, from scientific to engineering, gaming, and social networking, are considered.

- 1. Scientific applications
  - Scientific applications are a sector that is increasingly using cloud computing systems and technologies.
  - The immediate benefit seen by researchers and academics is the potentially infinite availability of computing resources and storage at sustainable prices compared to a complete in-house deployment.
  - Cloud computing systems meet the needs of different types of applications in the scientific domain:
    - high-performance computing (HPC) applications,
    - high-throughput computing (HTC) applications,
    - And data-intensive applications.

- The opportunity to use cloud resources is even more appealing because
- minimal changes need to be made to existing applications in order to leverage cloud resources.

- The most relevant option is IaaS solutions, which offer the optimal environment for running bag-of-tasks applications and workflows

- Virtual machine instances are opportunely customized to host the required software stack for
  - running such applications
  - and coordinated together with distributed computing middleware capable of interacting with cloud-based infrastructures

- PaaS solutions have been considered as well.

- They allow scientists to explore new programming models for tackling computationally challenging problems.

- Applications have been redesigned and implemented on top of cloud programming application models and platforms to leverage their unique capabilities.

- For instance, the MapReduce programming model provides scientists with a very simple and effective model for building applications that need to process large datasets.

- Therefore it has been widely used to develop data-intensive scientific applications.

1.1Healthcare: ECG analysis in the cloud

- Healthcare is a domain in which computer technology has found several and diverse applications:
- from supporting the business functions to assisting scientists in developing solutions to cure diseases.

- An important application is the use of cloud technologies to support doctors in providing more effective diagnostic processes.

- The capillary development of Internet connectivity and its accessibility from any device at any time has made cloud technologies an attractive option for developing health-monitoring systems.

- ECG data analysis and monitoring constitute a case that naturally fits into this scenario.

- ECG is the electrical manifestation of the contractile activity of the heart's myocardium.

- This activity produces a specific waveform that is repeated over time and that represents the heartbeat.

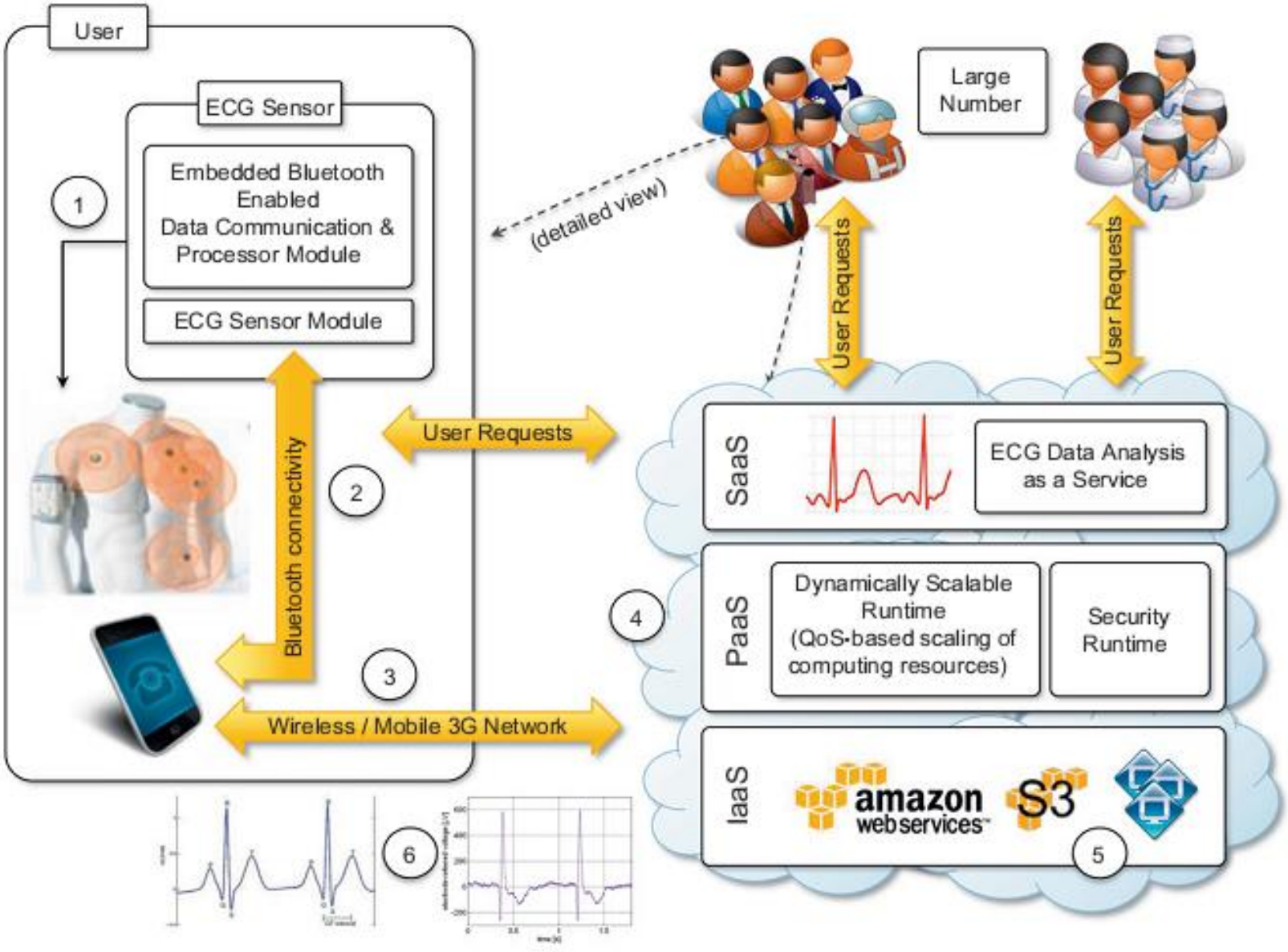
- The analysis of the shape of the ECG waveform is used to identify arrhythmias and is the most common way to detect heart disease.

- Cloud computing technologies allow the remote monitoring of a patient's heartbeat data, data analysis in minimal time,
- and the notification of first-aid personnel and doctors should these data reveal potentially dangerous conditions.

- This way a patient at risk can be constantly monitored without going to a hospital for ECG analysis.

- At the same time, doctors and first-aid personnel can instantly be notified of cases that require their attention.

- ECG monitoring is shown in Figure



An online health monitoring system hosted in the cloud.

- Wearable computing devices equipped with ECG sensors constantly monitor the patient's heartbeat.

- Such information is transmitted to the patient's mobile device, which will eventually forward it to the cloud-hosted Web service for analysis.

- The Web service forms the front-end of a platform that is entirely hosted in the cloud and that leverages the three layers of the cloud computing stack: SaaS, PaaS, and IaaS.

- The Web service constitute the SaaS application that will store ECG data in the Amazon S3 service and issue a processing request to the scalable cloud plat-form.

- The runtime platform is composed of a dynamically sizable number of instances running the workflow engine and Aneka.

- The number of workflow engine instances is controlled according to the number of requests in the queue of each instance,
- while Aneka controls the number of EC2 instances used to execute the single tasks defined by the workflow engine for a single ECG processing job.

- Each of these jobs consists of a set of operations involving the extraction of the waveform from the heartbeat data and the comparison of the waveform with a reference waveform to detect anomalies.

- If anomalies are found, doctors and first-aid personnel can be notified to act on a specific patient.

- The first advantage is the elasticity of the cloud infrastructure that can grow and shrink according to the requests served.

- As a result, doctors and hospitals do not have to invest in large computing infrastructures designed after capacity planning, thus making more effective use of budgets.

- The second advantage is ubiquity.

- Cloud computing technologies have now become easily accessible and promise to deliver systems with minimum or no downtime.

- Computing systems hosted in the cloud are accessible from any Internet device through simple interfaces (such as SOAP and REST-based Web ser-vices).

- This makes these systems not only ubiquitous, but they can also be easily integrated with other systems maintained on the hospital's premises.

- Finally, cost savings constitute another reason for the use of cloud technology in healthcare.

- Cloud services are priced on a pay-per-use basis and with volume prices for large numbers of service requests.

- These two models provide a set of flexible options that can be used to price the service, thus actually charging costs based on effective use rather than capital costs.

1.2 Biology: protein structure prediction

- Applications in biology often require high computing capabilities and often operate on large data-sets that cause extensive I/O operations.

- Because of these requirements, biology applications have often made extensive use of supercomputing and cluster computing infrastructures.

- Similar capabilities can be leveraged on demand using cloud computing technologies in a more dynamic fashion, thus opening new opportunities for bioinformatics applications.

- Protein structure prediction is a computationally intensive task that is fundamental to different types of research in the life sciences.

- Among these is the design of new drugs for the treatment of diseases.

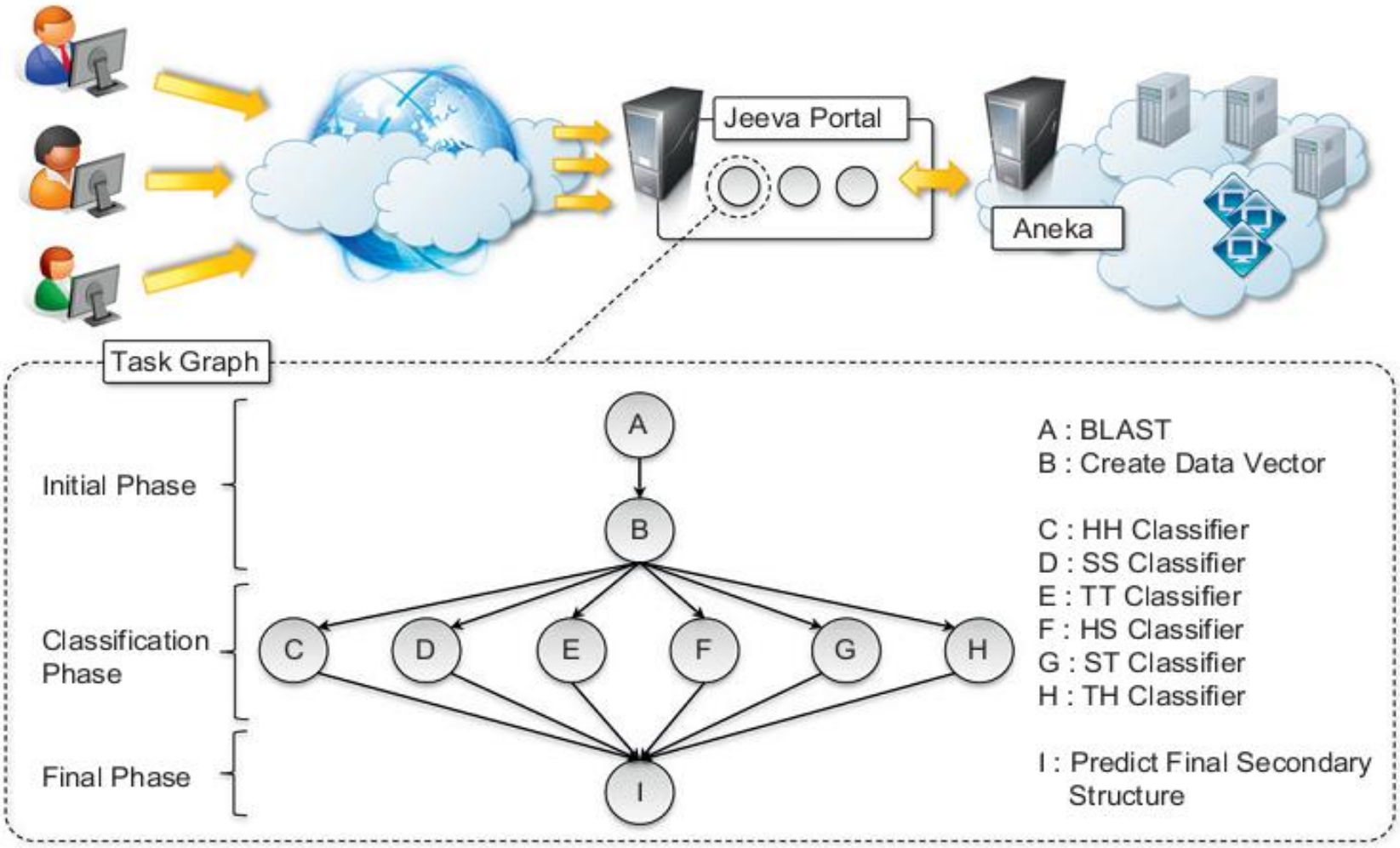
- The geometric structure of a protein can not be directly inferred from the sequence of genes that compose its structure,
- but it is the result of complex computations aimed at identifying the structure that minimizes the required energy.

- This task requires the investigation of a space with a massive number of states, consequently creating a large number of computations for each of these states.

- The computational power required for protein structure prediction can now be acquired on demand, without owning a cluster or navigating the bureaucracy to get access to parallel and distributed computing facilities.

- Cloud computing grants access to such capacity on a pay-per-use basis.

– One project that investigates the use of cloud technologies for protein structure prediction is Jeeva, an integrated Web portal that enables scientists to offload the prediction task to a computing cloud based on Aneka



Architecture and overview of the Jeeva Portal.

- The prediction task uses machine learning techniques (support vector machines) for determining the secondary structure of proteins.

- These techniques translate the problem into one of pattern recognition, where a sequence has to be classified into one of three possible classes (E, H, and C).

- A popular implementation based on support vector machines divides the pattern recognition problem into three phases:
  - initialization,
  - classification,
  - and a final phase.

- Even though these three phases have to be executed in sequence, it is possible to take advantage of parallel execution in the classification phase, where multiple classifiers are executed concurrently.

- This creates the opportunity to sensibly reduce the computational time of the prediction.

- The prediction algorithm is then translated into a task graph that is submitted to Aneka.

- Once the task is completed, the middleware makes the results available for visualization through the portal.

- The advantage of using cloud technologies (i.e., Aneka as scalable cloud middleware) versus conventional grid infrastructures is the capability to leverage a scalable computing infrastructure that can be grown and shrunk on demand.

- This concept is distinctive of cloud technologies and constitutes a strategic advantage when applications are offered and delivered as a service.

1.3 Biology: gene expression data analysis for cancer diagnosis

- Gene expression profiling is the measurement of the expression levels of thousands of genes at once.
- It is used to understand the biological processes that are triggered by medical treatment at a cellular level.
- Together with protein structure prediction, this activity is a fundamental component of drug design, since it allows scientists to identify the effects of a specific treatment.
- Another important application of gene expression profiling is cancer diagnosis and treatment.

- Cancer is a disease characterized by uncontrolled cell growth and proliferation.
- This behavior occurs because genes regulating the cell growth mutate.
- This means that all the cancerous cells contain mutated genes.
- In this context, gene expression profiling is utilized to provide a more accurate classification of tumors.
- The classification of gene expression data samples into distinct classes is a challenging task.
- The dimensionality of typical gene expression datasets ranges from several thousands to over tens of thousands of genes.
- However, only small sample sizes are typically available for analysis.

- This problem is often approached with learning classifiers, which generate a population of condition-action rules that guide the classification process.

- Among these, the eXtended Classifier System (XCS)has been successfully utilized for classifying large datasets in the bioinformatics and computer science domains.

- However, the effectiveness of XCS, when confronted with high dimensional datasets (such as microarray gene expression data sets), has not been explored in detail.

- A variation of this algorithm, CoXCS, has proven to be effective in these conditions.

- CoXCS divides the entire search space into sub domains and employs the standard XCS algorithm in each of these sub domains.

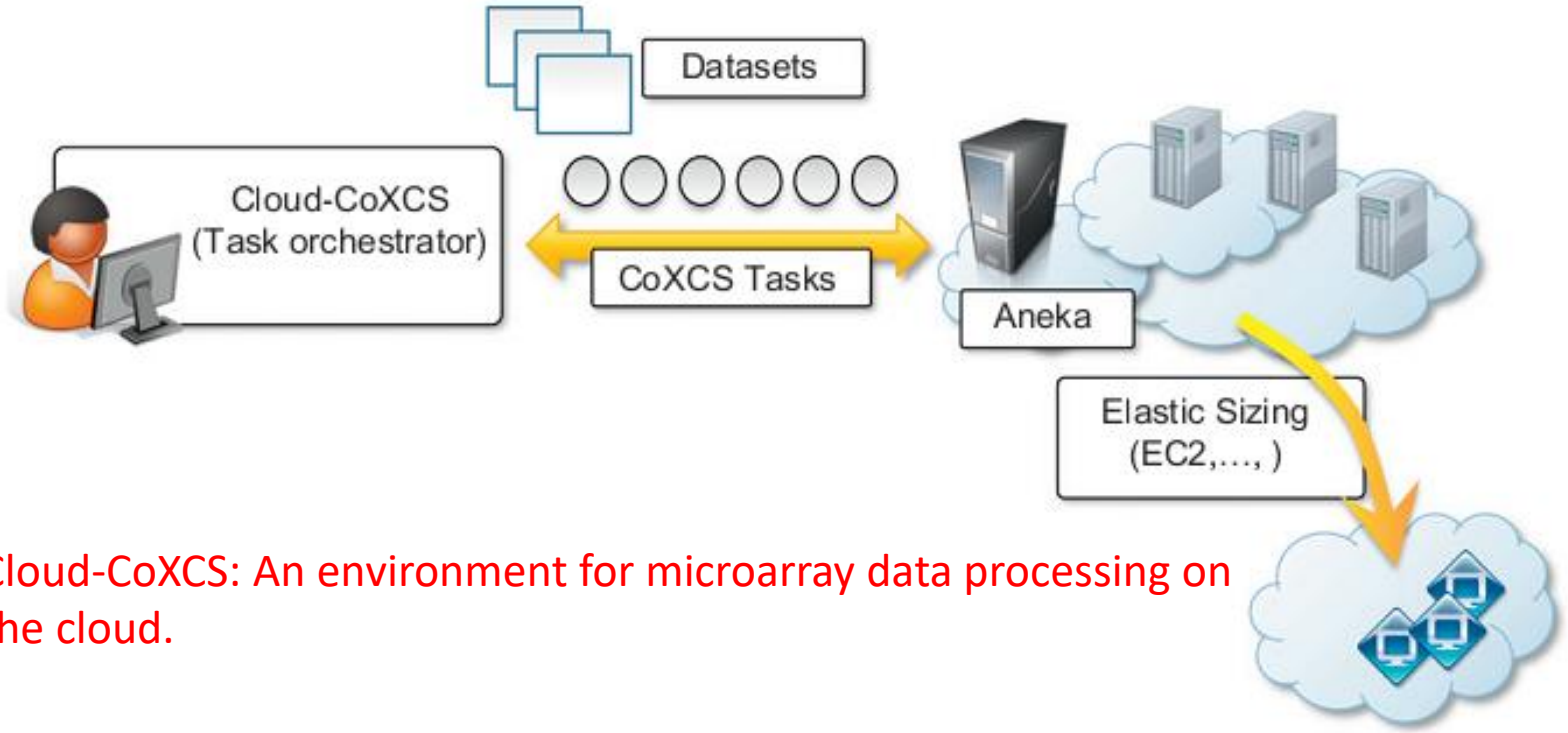
- Such a process is computationally intensive but can be easily parallelized because the classifications problems on the sub domains can be solved concurrently.

- Cloud-CoXCS is a cloud-based implementation of CoXCS that leverages Aneka to solve the classification problems in parallel and compose their outcomes.

- The algorithm is controlled by strategies, which define the way the outcomes are composed together and whether the process needs to be iterated.

– Because of the dynamic nature of XCS, the number of required compute resources to execute it can vary over time.

– Therefore, the use of scalable middleware such as Aneka offers a distinctive advantage.



Cloud-CoXCS: An environment for microarray data processing on the cloud.

1.4 Geoscience: satellite image processing

- Geoscience applications collect, produce, and analyze massive amounts of geospatial and nonspatial data.

- As the technology progresses and our planet becomes more instrumented (i.e., through the deployment of sensors and satellites for monitoring), the volume of data that needs to be processed increases significantly.

- In particular, the geographic information system (GIS) is a major element of geoscience applications.

- GIS applications capture, store, manipulate, analyze, manage, and present all types of geographically referenced data.

- This type of information is now becoming increasingly relevant to a wide variety of application domains:
- from advanced farming to civil security and natural resources management.

- As a result, a considerable amount of geo-referenced data is ingested into computer systems for further processing and analysis.

- Cloud computing is an attractive option for executing these demanding tasks and extracting meaningful information to support decision makers.

- 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.

- 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 pro-vides the appropriate infrastructure to support such application scenarios.
- A cloud-based implementation of such a workflow has been developed by the Department of Space, Government of India.
- The system shown in Figure integrates several technologies across the entire computing stack.

- A SaaS application provides a collection of services for such tasks as geocode generation and data visualization.

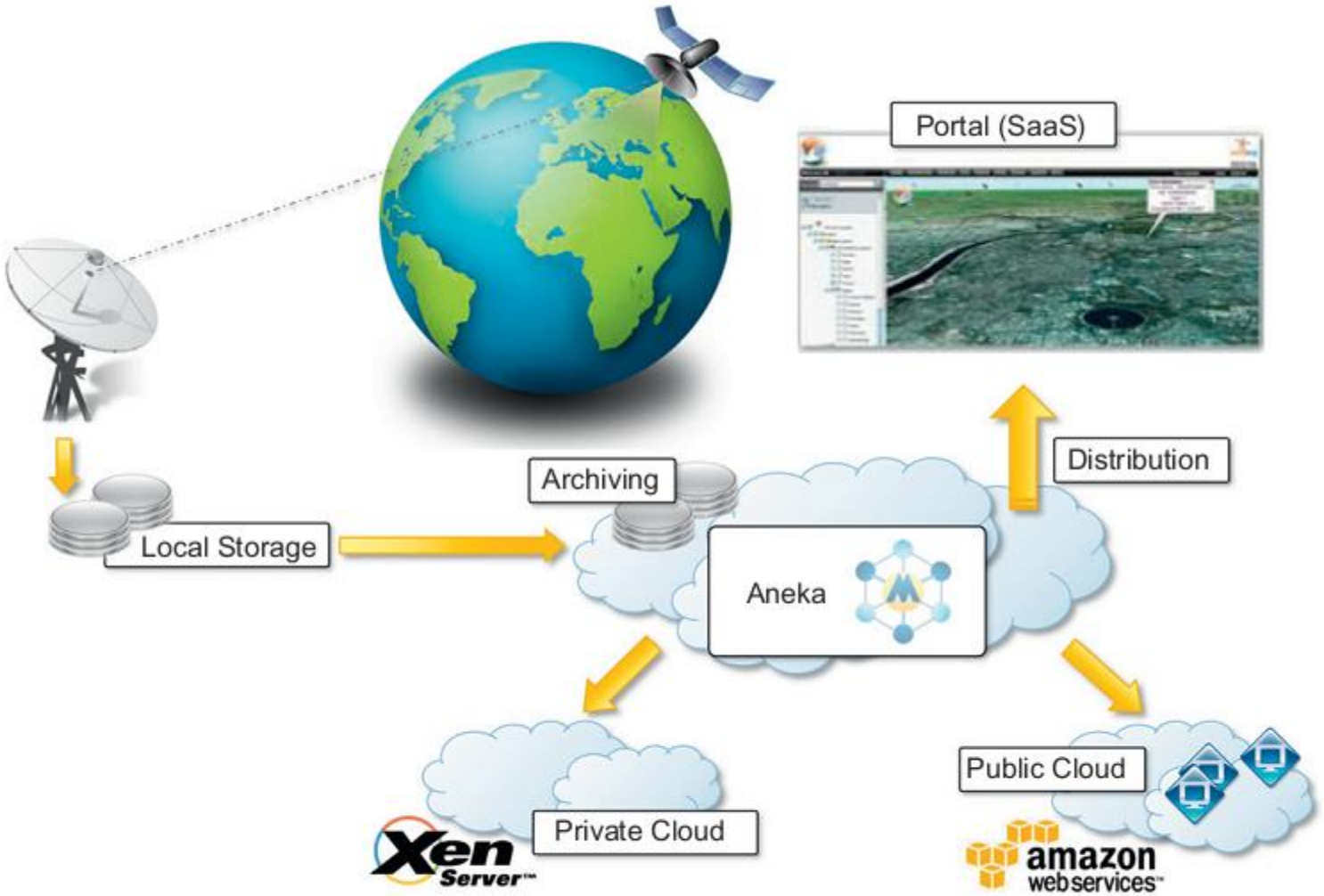


FIGURE 10.4

A cloud environment for satellite data processing.

- At the PaaS level, Aneka controls the importing of data into the virtualized infrastructure and the execution of image-processing tasks that produce the desired outcome from raw satellite images.

- The platform leverages a Xen private cloud and the Aneka technology to dynamically provision the required resources (i.e., grow or shrink) on demand.

- The project demonstrates how cloud computing technologies can be effectively employed to off-load local computing facilities from excessive workloads and leverage more elastic computing infrastructures.

2. Business and consumer applications

- The business and consumer sector is the one that probably benefits the most from cloud computing technologies.

- On one hand, the opportunity to transform capital costs into operational costs makes clouds an attractive option for all enterprises that are IT-centric.

- On the other hand, the sense of ubiquity that the cloud offers for accessing data and services makes it interesting for end users as well.

- Moreover, the elastic nature of cloud technologies does not require huge up-front investments, thus allowing new ideas to be quickly translated into products and services that can comfortably grow with the demand.

- The combination of all these elements has made cloud computing the preferred technology for a wide range of applications,
- from CRM and ERP systems to productivity and social-networking applications.

2.1 CRM and ERP

- Customer relationship management (CRM)and enterprise resource planning (ERP) applications are market segments that are flourishing in the cloud, with CRM applications the more mature of the two.

- Cloud CRM applications constitute a great opportunity for small enterprises and start-ups to have fully functional CRM software without large up-front costs and by paying subscriptions.

- Moreover, CRM is not an activity that requires specific needs, and it can be easily moved to the cloud.

- Such a characteristic, together with the possibility of having access to your business and customer data from everywhere and from any device, has fostered the spread of cloud CRM applications.

- ERP solutions on the cloud are less mature and have to compete with well-established in-house solutions.

- ERP systems integrate several aspects of an enterprise: finance and accounting, human resources, manufacturing, supply chain management, project management, and CRM.

- Their goal is to provide a uniform view and access to all operations that need to be performed to sustain a complex organization.

- Because of the organizations that they target, the transition to cloud-based models is more difficult:
- he cost advantage over the long term might not be clear, and the switch to the cloud could be difficult if organizations already have large ERP installations.

- For this reason cloud ERP solutions are less popular than CRM solutions at this time.

2.1.1 Salesforce.com

- Salesforce.com is probably the most popular and developed CRM solution available today.
- As of today more than 100,000 customers have chosen Safesforce.com to implement their CRM solutions.
- The application provides customizable CRM solutions that can be integrated with additional features developed by third parties.
- Salesforce.com is based on the Force.com cloud development platform.
- This represents scalable and high-performance middleware executing all the operations of all Salesforce.com applications.

- The architecture of theForce.com platform is shown in Figure

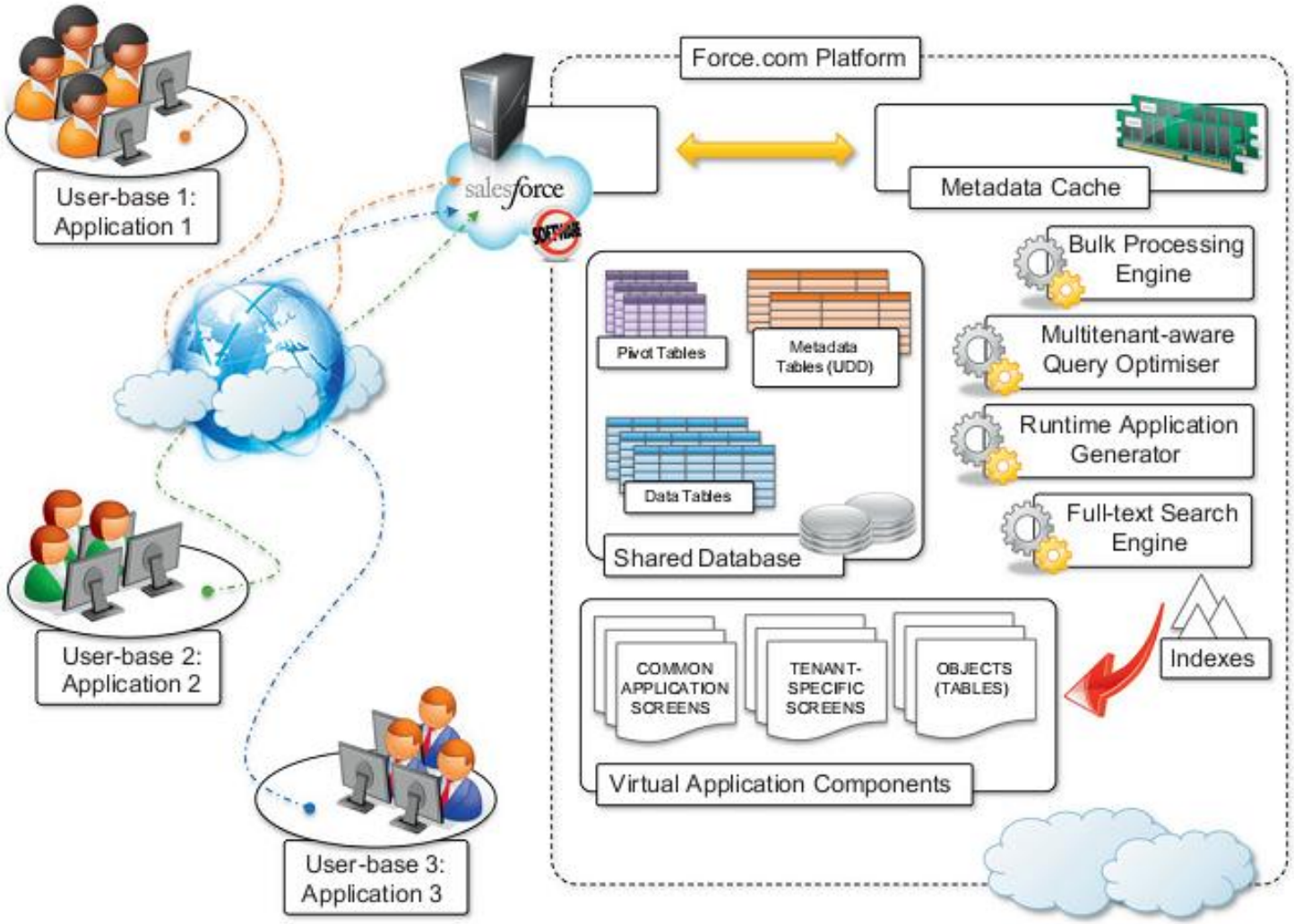


FIGURE 10.5

Salesforce.com and Force.com architecture.

- Initially designed to sup-port scalable CRM applications, the platform has evolved to support the entire life cycle of a wider range of cloud applications by implementing a flexible and scalable infrastructure.

- At the core of the platform resides its metadata architecture, which provides the system with flexibility and scalability.

- Rather than being built on top of specific components and tables, application core logic and business rules are saved as metadata into the Force.com store.

- Both application structure and application data are stored in the store.

- A runtime engine executes application logic by retrieving its metadata and then performing the operations on the data.

- Although running in isolated containers, different applications logically share the same database structure, and the runtime engine exe-cutes all of them uniformly.

- A full-text search engine supports the runtime engine.

- This allows application users to have an effective user experience despite the large amounts of data that need to be crawled.

- The search engine maintains its indexing data in a separate store and is constantly updated by background processes triggered by user interaction.

- Users can customize their application by leveraging the “native”Force.com application frame-work or by using programmatic APIs in the most popular programming languages.

- The application framework allows users to visually define either the data or the core structure of aForce.com application,
- while the programmatic APIs provide them with a more conventional way for developing applications that relies on Web services to interact with the platform.

- Customization of application processes and logic can also be implemented by developing scripts in APEX.

- This is a Java-like language that provides object-oriented and procedural capabilities for defining either scripts executed on demand or triggers.

- APEX also offers the capability of expressing searches and queries to have complete access to the data managed by theForce.com platform.

2.1.2 Microsoft dynamics CRM

- Microsoft Dynamics CRM is the solution implemented by Microsoft for customer relationship management.

- Dynamics CRM is available either for installation on the enterprise's premises or as an online solution priced as a monthly per-user subscription.

- The system is completely hosted in Microsoft's datacenters across the world and offers to customers a 99.9% SLA, with bonus credits if the system does not fulfill the agreement.

- Each CRM instance is deployed on a separate database, and the application provides users with facilities for marketing, sales, and advanced customer relationship management.

- Dynamics CRM Online features can be accessed either through a Web browser interface or programmatically by means of SOAP and RESTful (Representational State Transfer)Web services.

- **Restful Web Services** is a lightweight, maintainable, and scalable **service** that is built on the **REST** architecture.

- This allows Dynamics CRM to be easily integrated with both other Microsoft products and line-of-business applications.

- Dynamics CRM can be extended by developing plug-ins that allow implementing specific behaviors triggered on the occurrence of given events.

- Dynamics CRM can also leverage the capability of Windows Azure for the development and integration of new features.

2.1.3 NetSuite

- NetSuite provides a collection of applications that help customers manage every aspect of the business enterprise.

- Its offering is divided into three major products:NetSuite Global ERP, NetSuite Global CRM1, and NetSuite Global Ecommerce.

- Moreover, an all-in-one solution: NetSuite One World, integrates all three products together.

- The services NetSuite delivers are powered by two large datacenters on the East and West coasts of the United States, connected by redundant links.

- This allows NetSuite to guarantee 99.5% uptime to its customers.

- Besides the prepackaged solutions, NetSuite also provides an infrastructure and a development environment for implementing customized applications.

- TheNetSuite Business Operating System (NS-BOS)is a complete stack of technologies for building SaaS business applica-tions that leverage the capabilities of NetSuite products.

- On top of the SaaS infrastructure, the NetSuite Business Suite components offer accounting, ERP, CRM, and ecommerce capabilities.

- An online development environment, SuiteFlex, allows integrating such capabilities into new Web applications, which are then packaged for distribution bySuiteBundler.

- The entire infrastructure is hosted in the NetSuite datacenters, which provide warranties regarding application uptime and availability.

10.2.2Productivity

- Productivity applications replicate in the cloud some of the most common tasks that we are used to performing on our desktop: from document storage to office automation and complete desktop environments hosted in the cloud.

2.2.1 Dropbox and iCloud

- One of the core features of cloud computing is availability anywhere, at any time, and from any Internet-connected device.

- Therefore, document storage constitutes a natural application for such technology.

- Online storage solutions preceded cloud computing, but they never became popular.

- With the development of cloud technologies, online storage solutions have turned into SaaS applications and become more usable as well as more advanced and accessible.

– Perhaps the most popular solution for online document storage is Dropbox,
– an online application that allows users to synchronize any file across any platform and any device in a seamless manner as shown in Figure .

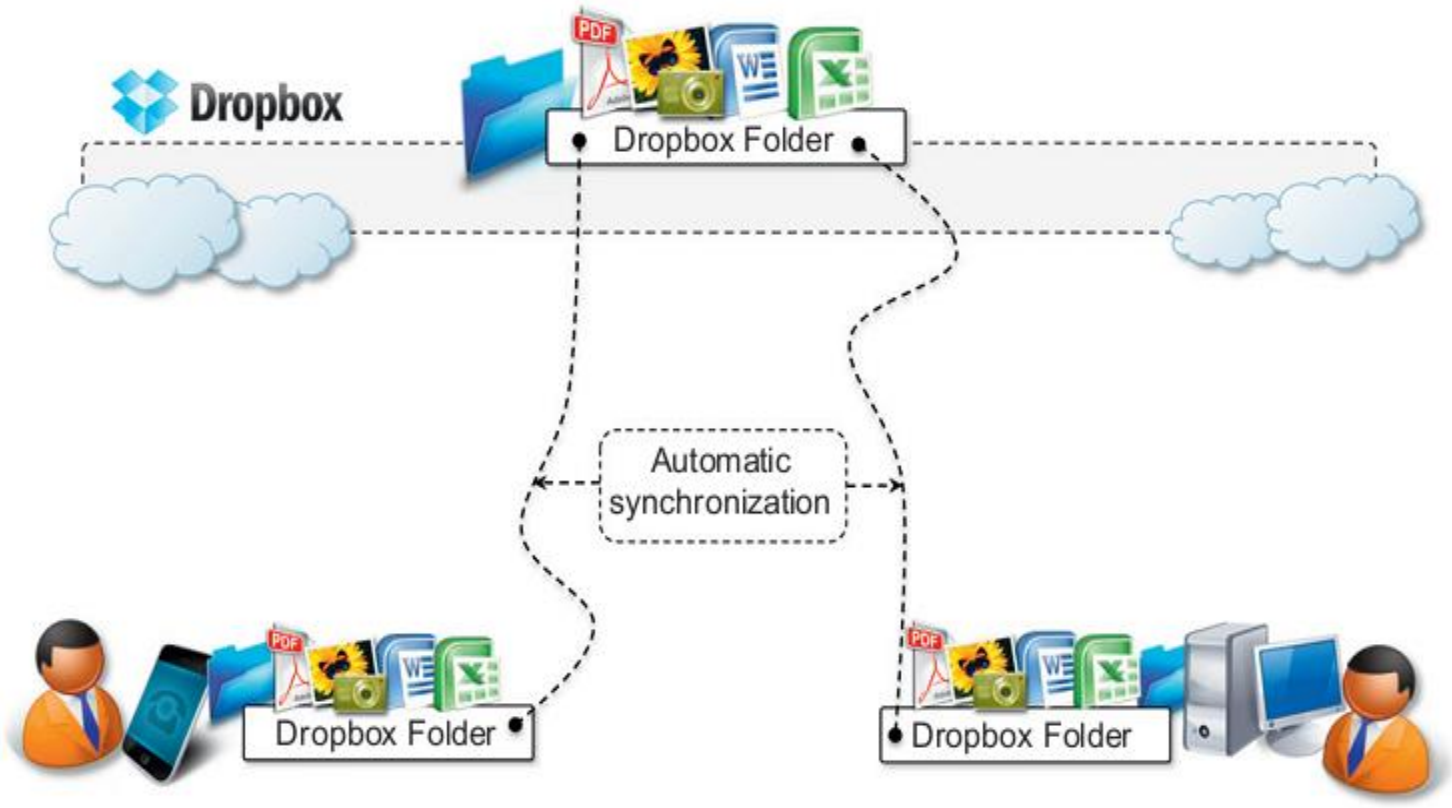


FIGURE 10.6

Dropbox usage scenario.

- Dropbox provides users with a free amount of storage that is accessible through the abstraction of a folder.

- Users can either access their Dropbox folder through a browser or by downloading and installing a Dropbox client, which provides access to the online storage by means of a special folder.

- All the modifications into this folder are silently synched so that changes are notified to all the local instances of the Dropbox folder across all the devices.

- The key advantage of Dropbox is its availability on different platforms (Windows, Mac, Linux, and mobile) and the capability to work seamlessly and transparently across all of them.

- Another interesting application in this area is iCloud, a cloud-based document-sharing application provided by Apple to synchronize iOS-based devices in a completely transparent manner.

- Unlike Dropbox, which provides synchronization through the abstraction of a local folder, iCloud has been designed to be completely transparent once it has been set up.

- Documents, photos, and videos are automatically synched as changes are made, without any explicit operation.

- This allows the system to efficiently automate common operations without any human intervention:
  - taking a picture with your iPhone and having it automatically available in iPhoto on your Mac at home;
  - editing a document on the iMac at home and having the changes updated in your iPad.
- Unfortunately, this capability is limited to iOS devices,
- and currently there are no plans to provide iCloud with a Web-based interface that would make user content accessible from even unsupported platforms.
- There are other solutions for online document sharing, such as Windows Live, Amazon Cloud Drive, and CloudMe, that are popular .

2.2.2 Google docs

- Google Docs is a SaaS application that delivers the basic office automation capabilities with sup-port for collaborative editing over the Web.

- The application is executed on top of the Google distributed computing infrastructure, which allows the system to dynamically scale according to the number of users using the service.

- Google Docs allows users to create and edit text documents, spreadsheets, presentations, forms, and drawings.

- It aims to replace desktop products such as Microsoft Office and Open Office and provide similar interface and functionality as a cloud service.

- It supports collaborative editing over the Web for most of the applications included in the suite.

- This eliminates tedious emailing and synchronization tasks when documents need to be edited by multiple users.

- By being stored in the Google infrastructure, these documents are always available from anywhere and from any device that is connected to the Internet.

- Moreover, the suite allows users to work offline if Internet connec-tivity is not available.

- Support for various formats such as those that are produced by the most popular desktop office solutions allows users to easily import
- and move documents in and out of Google Docs, thus eliminating barriers to the use of this application.

- Google Docs is a good example of what cloud computing can deliver to end users:
- ubiquitous access to resources, elasticity, absence of installation and maintenance costs, and delivery of core functionalities as a service.

2.2.3 Cloud desktops: EyeOS and XIOS/3

- Asynchronous JavaScript and XML (AJAX) technologies have considerably augmented the capabilities that can be implemented in Web applications.

- This is a fundamental aspect for cloud computing, which delivers a considerable amount of its services through the Web browser.

- Together with the opportunity to leverage large-scale storage and computation, this technology has made possible the replication of complex desktop environments in the cloud
- and made them available through the Web browser.

- These applications, called cloud desktops, are rapidly gaining in popularity.

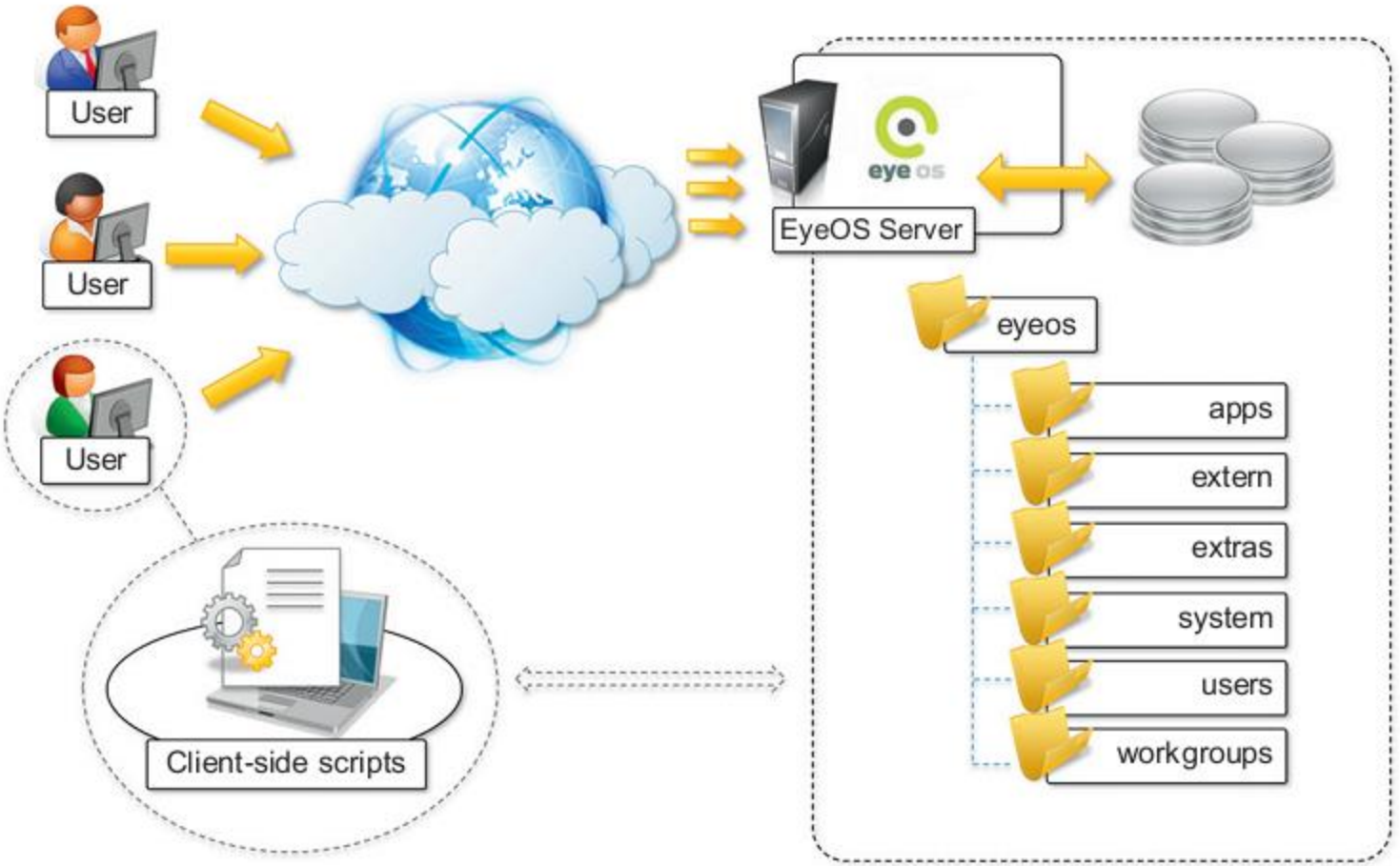


FIGURE 10.7

EyeOS architecture.

- EyeOS is one of the most popular Web desktop solutions based on cloud technologies.

- It replicates the functionalities of a classic desktop environment
- and comes with preinstalled applications for the most common file and document management tasks as shown in Figure.

- Single users can access the EyeOS desktop environment from anywhere and through any Internet-connected device,

- whereas organizations can create a private EyeOS Cloud on their premises to virtualize the desktop environment of their employees and centralize their management.

- The EyeOS architecture is quite simple: On the server side, the EyeOS application maintains the information about user profiles and their data,
- and the client side constitutes the access point for users and administrators to interact with the system.

- EyeOS stores the data about users and applications on the server file system.

- Once the user has logged in by providing credentials, the desktop environment is rendered in the client's browser by downloading all the JavaScript libraries required ,
  - to build the user interface
  - and implement the core functionalities of EyeOS.

- Each application loaded in the environment communicates with the server by using AJAX.

- this communication model is used to access user data as well as to perform application operations:
  - editing documents,
  - visualizing images,
  - copying and saving files,
  - sending emails,
  - and chatting.

- EyeOS also provides APIs for developing new applications and integrating new capabilities into the system.

- EyeOS applications are server-side components that are defined by at least two files.

- (stored in the eyeos/apps/app name directory) : appname.php and appname.js

- The first file defines and implements all the operations that the application exposes;
- the JavaScript file contains the code that needs to be loaded in the browser in order to provide user interaction with the application.

- Xcerion XML Internet OS/3 (XIOS/3)is another example of a Web desktop environment.

- The service is delivered as part of the CloudMe application, which is a solution for cloud document storage.

- The key differentiator of XIOS/3 is its strong leverage of XML, used to implement many of the tasks of the OS:
  - rendering user interfaces,
  - defining application business logics,
  - structuring file system organization,
  - and even application development.

- The architecture of the OS concentrates most of the functionalities on the client side
- while implementing server-based functionalities by means of XML Web services.

- The client side renders the user interface, orchestrates processes, and provides data-binding capabilities on XML data that is exchanged with Web services.

- The server is responsible for implementing core functions such as
  - transaction management for documents edited in a collaborative mode
  - and core logic of installed applications into the environment.

- XIOS/3 also provides an environment for developing applications (XIDE), which allows users to quickly develop complex applications by visual tools for the user interface and XML documents for business logic.

- XIOS/3 is released as open-source software and implements a marketplace where third parties can easily deploy applications that can be installed on top of the virtual desktop environment.

- It is possible to develop any type of application and feed it with data accessible through XML Web services:
  - developers have to define the user interface,
  - bind UI components to service calls and operations,
  - and provide the logic on how to process the data.

- XIDE will package this information into a proper set of XML documents, and the rest will be performed by an XML virtual machine implemented in XIOS.

- XIOS/3 is an advanced Web desktop environment that focuses on the integration of services into the environment by means of XML-based services and that simplifies collaboration with peers.

2.3 Social networking

- Social networking applications have grown considerably in the last few years to become the most active sites on the Web.

- To sustain their traffic and serve millions of users seamlessly, services such as Twitter and Facebook have leveraged cloud computing technologies.

- The possibility of continuously adding capacity while systems are running is the most attractive feature for social networks, which constantly increase their user base.

2.3.1 Facebook

- Facebook is probably the most evident and interesting environment in social networking.

- With more than 800 million users, it has become one of the largest Websites in the world.

- To sustain this incredible growth, it has been fundamental that Facebook be capable of continuously adding capacity

  - and developing new scalable technologies and software systems
  - while maintaining high performance to ensure a smooth user experience.

- Currently, the social network is backed by two data centers that have been built and optimized to reduce costs and impact on the environment.

- On top of this highly efficient infrastructure, built and designed out of inexpensive hardware, a completely customized stack of opportunely modified
- and refined open-source technologies constitutes the back-end of the largest social network.

- Taken all together, these technologies constitute a powerful platform for developing cloud applications.

- This platform primarily supports Facebook itself and offers APIs to integrate third-party applications with Facebook's core infrastructure to deliver additional services such as social games and quizzes created by others.

- The reference stack serving Facebook is based on LAMP(Linux, Apache, MySQL, and PHP).

- This collection of technologies is accompanied by a collection of other services developed in-house.

- These services are developed in a variety of languages and implement specific functionalities such as
  - search,
  - news feeds,
  - notifications,
  - and others.

- While serving page requests, the social graphof the user is composed

- The social graph identifies a collection of interlinked information that is of relevance for a given user.
- Most of the user data are served by querying a distributed cluster of MySQL instances, which mostly contain key-value pairs.
- These data are then cached for faster retrieval.
- The rest of the relevant information is then composed together using the services mentioned before.
- These services are located closer to the data and developed in languages that provide better performance than PHP.

- The development of services is facilitated by a set of internally developed tools.
- One of the core elements is Thrift.
- This is a collection of abstractions (and language bindings) that allow cross-language development.
- Thrift allows services developed in different languages to communicate and exchange data.

- Bindings for Thrift in different languages take care of data serialization and deserialization, communication, and client and server boilerplate code.

- This simplifies the work of the developers, who can quickly prototype services and leverage existing ones.

- Other relevant services and tools are Scribe, which aggregates streaming log feeds, and applications for alerting and monitoring.

2.4 Media applications

- Media applications are a niche that has taken a considerable advantage from leveraging cloud computing technologies.

- In particular, video-processing operations, such as
  - encoding,
  - transcoding,
  - composition,
  - and rendering, are good candidates for a cloud-based environment.

- These are computationally intensive tasks that can be easily offloaded to cloud computing infrastructures.

2.4.1 Animoto

- Animoto is perhaps the most popular example of media applications on the cloud.

- The Website provides users with a very straightforward interface for quickly creating videos out of images, music, and video fragments submitted by users.

- Users select a specific theme for a video, upload the photos and videos and order them in the sequence they want to appear, select the song for the music, and render the video.

- The process is executed in the background and the user is notified via email once the video is rendered.

- The core value of Animoto is the ability to quickly create videos with stunning effects without user intervention.

- A proprietary artificial intelligence (AI) engine, which selects the animation and transition effects according to pictures and music, drives the rendering operation.

- Users only have to define the storyboard by organizing pictures and videos into the desired sequence.

- If users don’t like the result, the video can be rendered again and the engine will select a different composition, thus producing a different outcome every time.

- The service allows users to create 30-second videos for free.

– By paying a monthly or a yearly subscription it is possible to produce videos of any length and to choose among a wider range of templates.

– The infrastructure supporting Animoto is complex and is composed of different systems that all need to scale as shown in figure

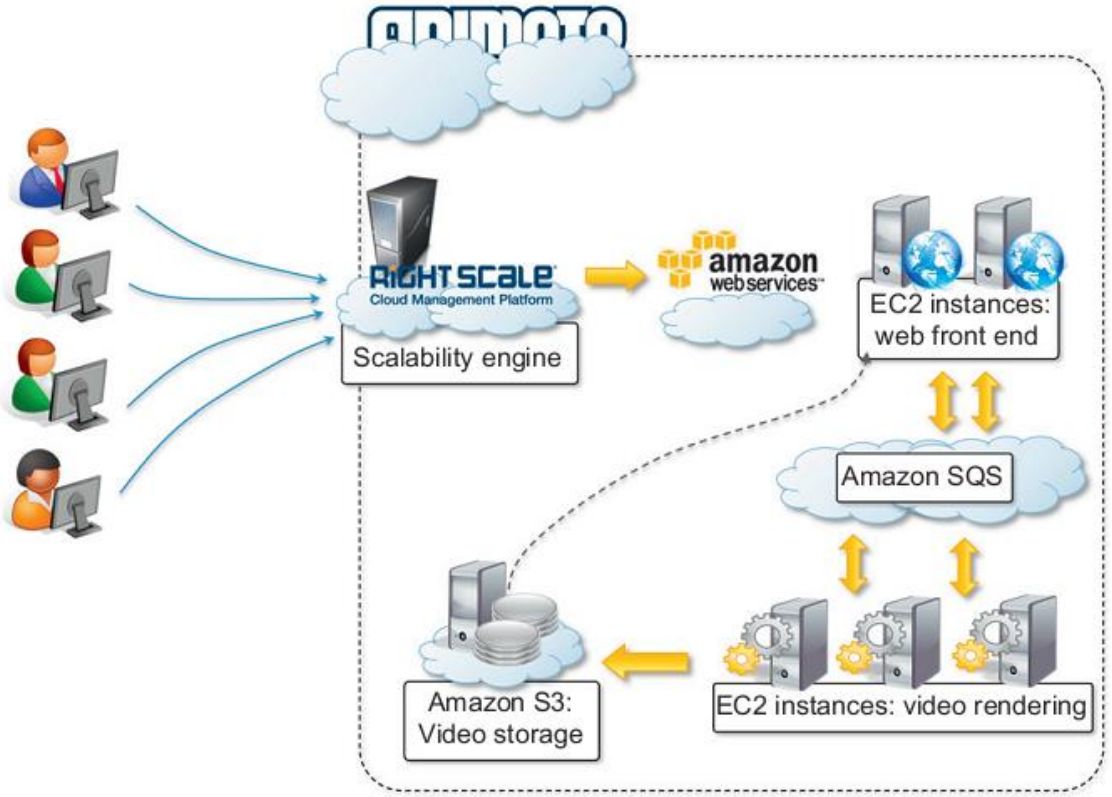


FIGURE 10.8

Animoto reference architecture.

- 2.4.2 Maya rendering with Aneka
  - Interesting applications of media processing are found in the engineering disciplines and the movie production industry.
  - Operations such as rendering of models are now an integral part of the design workflow, which has become computationally demanding.
  - The visualization of mechanical models is not only used at the end of the design process, it is iteratively used to improve the design.
  - It is then fundamental to perform such tasks as fast as possible.
  - Cloud computing provides engineers with the necessary computing power to make this happen.

– A private cloud solution for rendering train designs has been implemented by the engineering department of GoFront group, a division of China Southern Railway as shown in figure

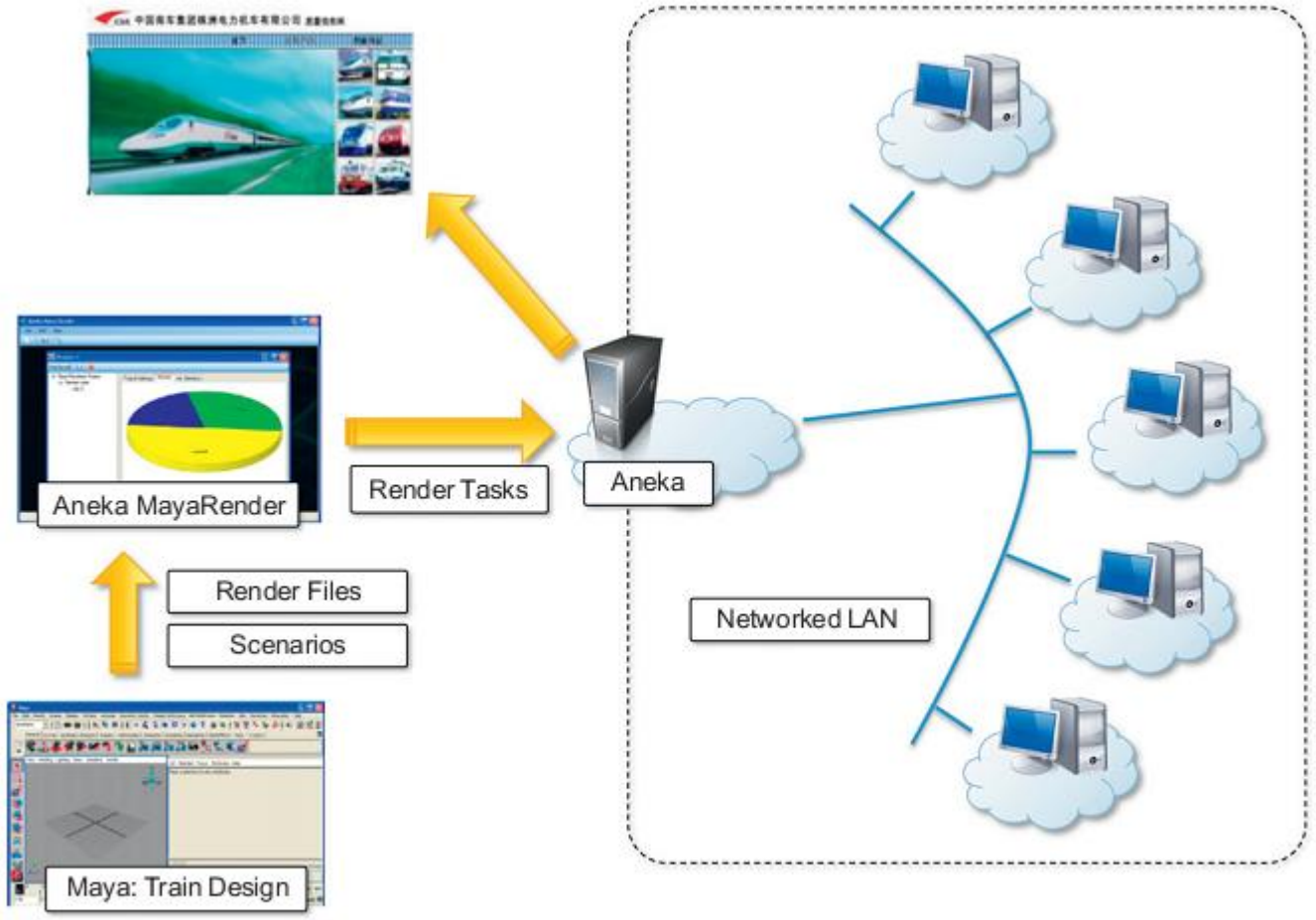


FIGURE 10.9

3D rendering on private clouds.

- The department is responsible for designing models of high-speed electric locomotives, metro cars, urban transportation vehicles, and motor trains.

- The design process for prototypes requires high-quality, three-dimensional (3D) images.

- The analysis of these images can help engineers identify problems and correct their design.

- Three-dimensional rendering tasks take considerable amounts of time,
- especially in the case of huge numbers of frames, but it is critical for the department to reduce the time spent in these iterations.

- This goal has been achieved by leveraging cloud computing technologies, which turned the network of desktops in the department into a desktop cloud managed by Aneka.

- The implemented system includes a specialized client interface that can be used by GoFront engineers to enter all the details of the rendering process (the number of frames, the number of cameras, and other parameters).

- The application is used to submit the rendering tasks to the Aneka Cloud, which distributes the load across all the available machines.

- Every rendering task triggers the execution of the local Maya batch renderer and collects the result of the execution.

- The renders are then retrieved and put all together for visualization.

- By turning the local network into a private cloud, the resources of which can be used off-peak
- (i.e., at night, when desktops are not utilized),

- it has been possible for GoFront to sensibly reduce the time spent in the rendering process from days to hours.

2.4.3 Video encoding on the cloud: Encoding.com

- Video encoding and transcoding are operations that can greatly benefit from using cloud technologies:

- They are computationally intensive and potentially require considerable amounts of storage.

- Moreover, with the continuous improvement of mobile devices as well as the diffusion of the Internet, requests for video content have significantly increased.

- The variety of devices with video playback capabilities has led to an explosion of formats through which a video can be delivered.

- Software and hardware for video encoding and transcoding often have prohibitive costs or are not flexible enough to support conversion from any format to any format.

- Cloud technologies present an opportunity for turning these tedious and often demanding tasks into services
- that can be easily integrated into a variety of workflows or made available to everyone according to their needs.

- Encoding.com is a software solution that offers video-transcoding services on demand
- and leverages cloud technology to provide both the horsepower required for video conversion and the storage for staging videos.

- The service integrates with both Amazon Web Services technologies (EC2, S3, and CloudFront) and Rackspace (Cloud Servers, Cloud Files, and Limelight CDN access).

- Users can access the services through a variety of interfaces: the Encoding.com Website, Web service XML APIs, desktop applications, and watched folders.

- To use the service, users have to specify the location of the video to transcode, the destination format, and the target location of the video.

- Encoding.com also offers other video-editing operations such as the insertion of thumb-nails, watermarks, or logos.

- Moreover, it extends its capabilities to audio and image conversion.

- The service provides various pricing options: monthly fee, pay-as-you-go (by batches), and special prices for high volumes.

- Encoding.com now has more than 2,000 customers and has already processed more than 10 million videos.

- 10.2.5Multiplayer online gaming
  - Online multiplayer gaming attracts millions of gamers around the world
  - who share a common experience by playing together in a virtual environment that extends beyond the boundaries of a normal LAN.
  - Online games support hundreds of players in the same session, made possible by the specific architecture used to forward interactions, which is based on game log processing.
  - Players update the game server hosting the game session, and the server integrates all the updates into a log that is made available to all the players through a TCP port.

- The client software used for the game connects to the log port and, by reading the log, updates the local user interface with the actions of other players.

- Game log processing is also utilized to build statistics on players and rank them.

- These features constitute the additional value of online gaming portals that attract more and more gamers.

- The processing of game logs is a potentially compute-intensive operation that strongly depends on the number of players online and the number of games monitored.

- Moreover, gaming portals are Web applications and therefore might suffer from the spiky behavior of users that can randomly generate large amounts of volatile workloads that do not justify capacity planning.

- The use of cloud computing technologies can provide the required elasticity for seamlessly processing these workloads
- and scale as required when the number of users increases.

- A prototypal implementation of cloud-based game log processing has been implemented by Titan Inc. (now Xfire),
- a company based in California that extended its gaming portal for offload game log proces-sing to an Aneka Cloud.

– The prototype as shown in Figure uses a private cloud deployment that allowed Titan Inc. to process concurrently multiple logs and sustain a larger number of users.

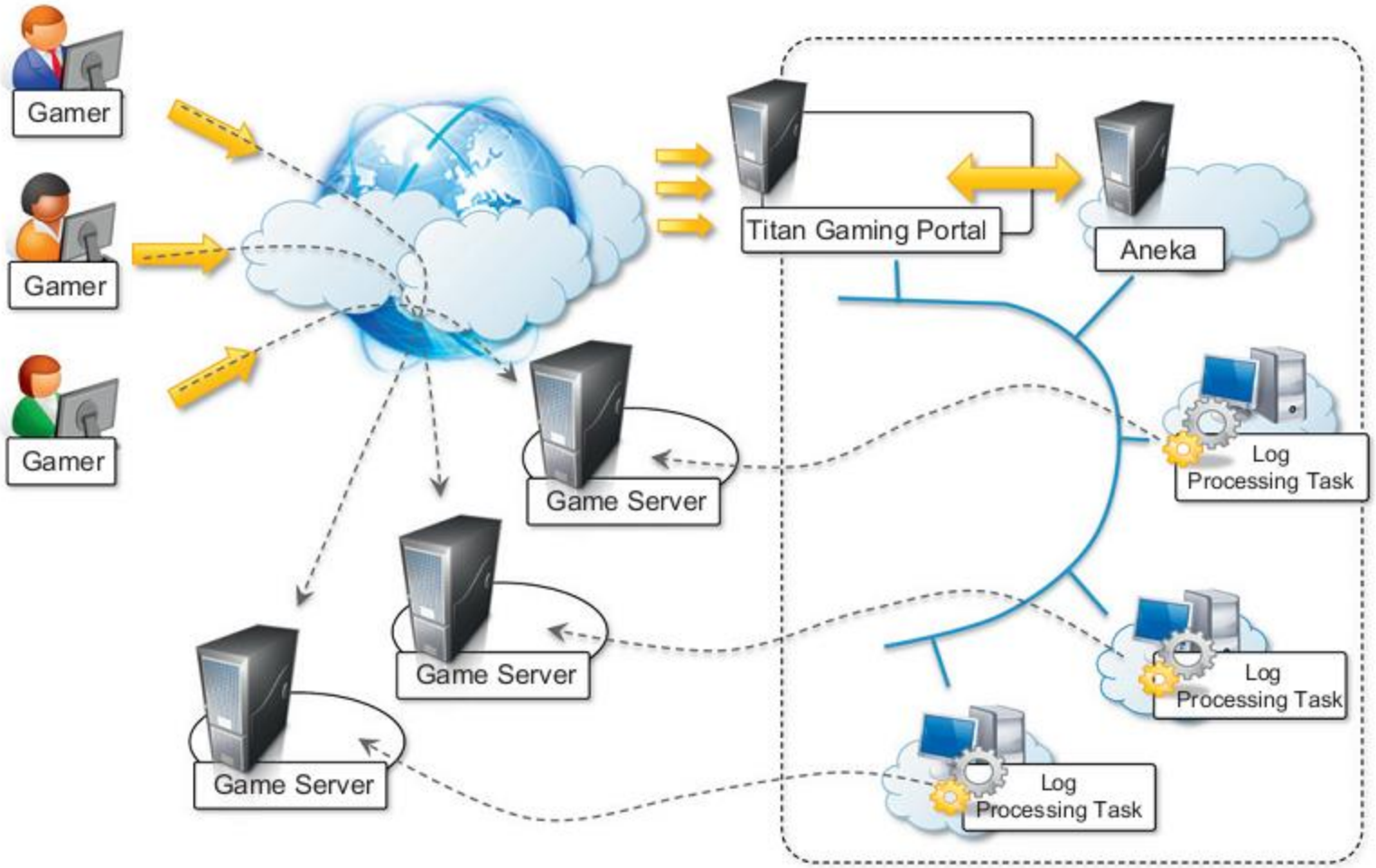


FIGURE 10.10

Scalable processing of logs for network games.

End