UNIT VI

Cloud Applications

Syllabus

- Scientific Applications -
  - Health Care: ECG analysis in the cloud
  - Biology: protein structure prediction
  - Biology: gene expression data analysis for cancer diagnosis
  - Geoscience: satellite image processing
- Business and Consumer Applications -
  - CRM and ERP
  - Social Networking
  - Media Applications
  - Multiplayer Online Gaming

Introduction

- 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.
- In this unit we discuss some application case studies, detailing their architecture and how they leveraged various cloud technologies.
- Applications from a range of domains, from scientific to engineering, gaming, and social networking, are considered.

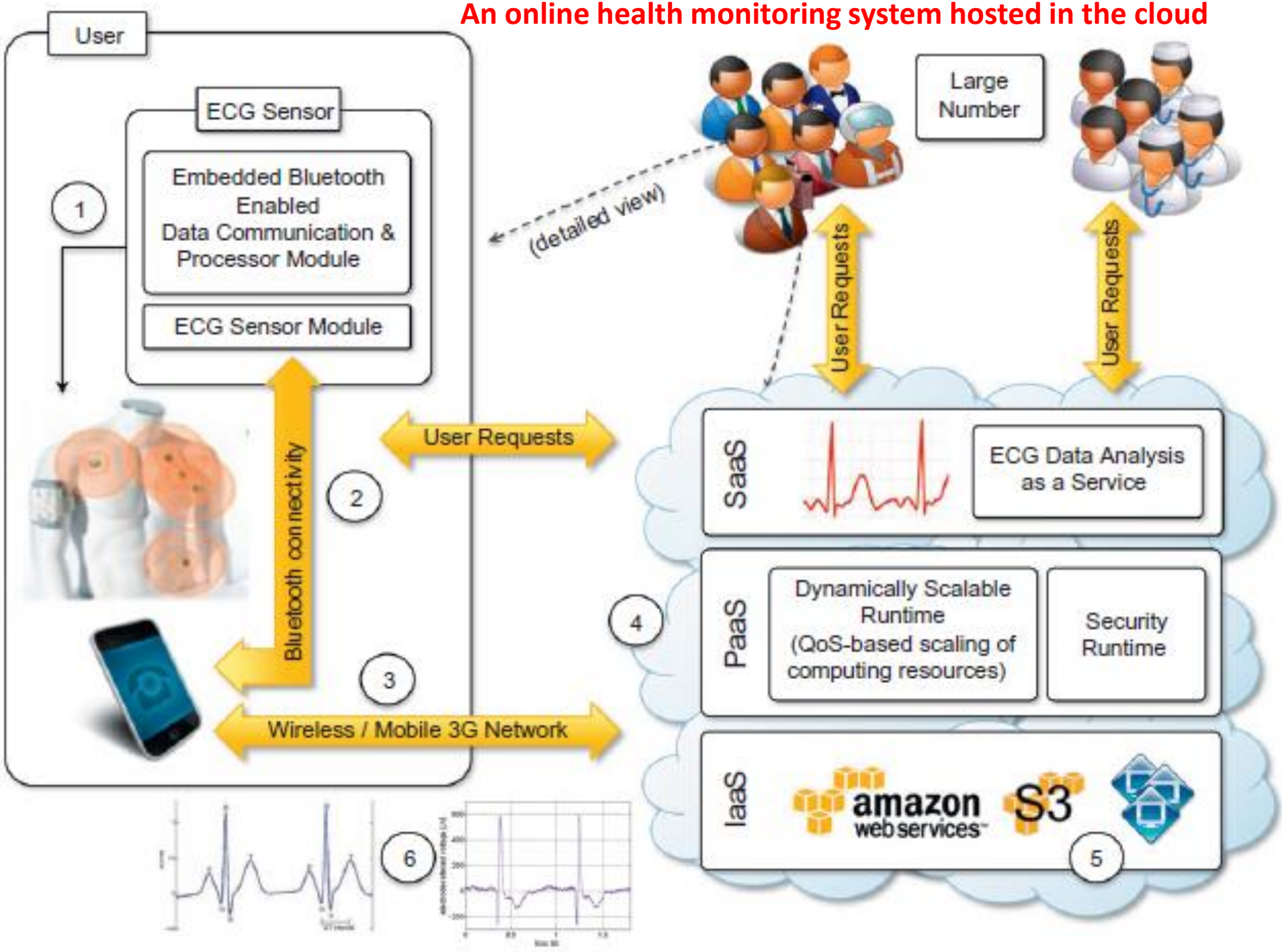
Scientific Applications

- Scientific applications are a sector that is increasingly using cloud computing systems and technologies.
- 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.

Healthcare: 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.
- An illustration of the infrastructure and model for supporting remote ECG monitoring is shown in below figure.

An online health monitoring system hosted in the cloud



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

Healthcare: ECG analysis in the cloud

- 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.

Advantages

- **Elasticity** - The elasticity of 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.
- **Ubiquity** - Cloud computing technologies are easily accessible and promise to deliver systems with minimum or no downtime. Computing systems hosted in cloud are accessible from any Internet device through simple interfaces (such as SOAP and REST-based Web services). This makes systems easily integrated with other systems maintained on hospital's premises.
- **Cost savings** - Cloud services are priced on a pay-per-use basis and with volume prices for large numbers of service requests.

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 made use of supercomputing and cluster computing infrastructures.
- Similar capabilities can be leveraged using cloud computing technologies in a more dynamic fashion, thus opening new opportunities for bioinformatics applications.

Biology: protein structure prediction

- 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 cannot 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.

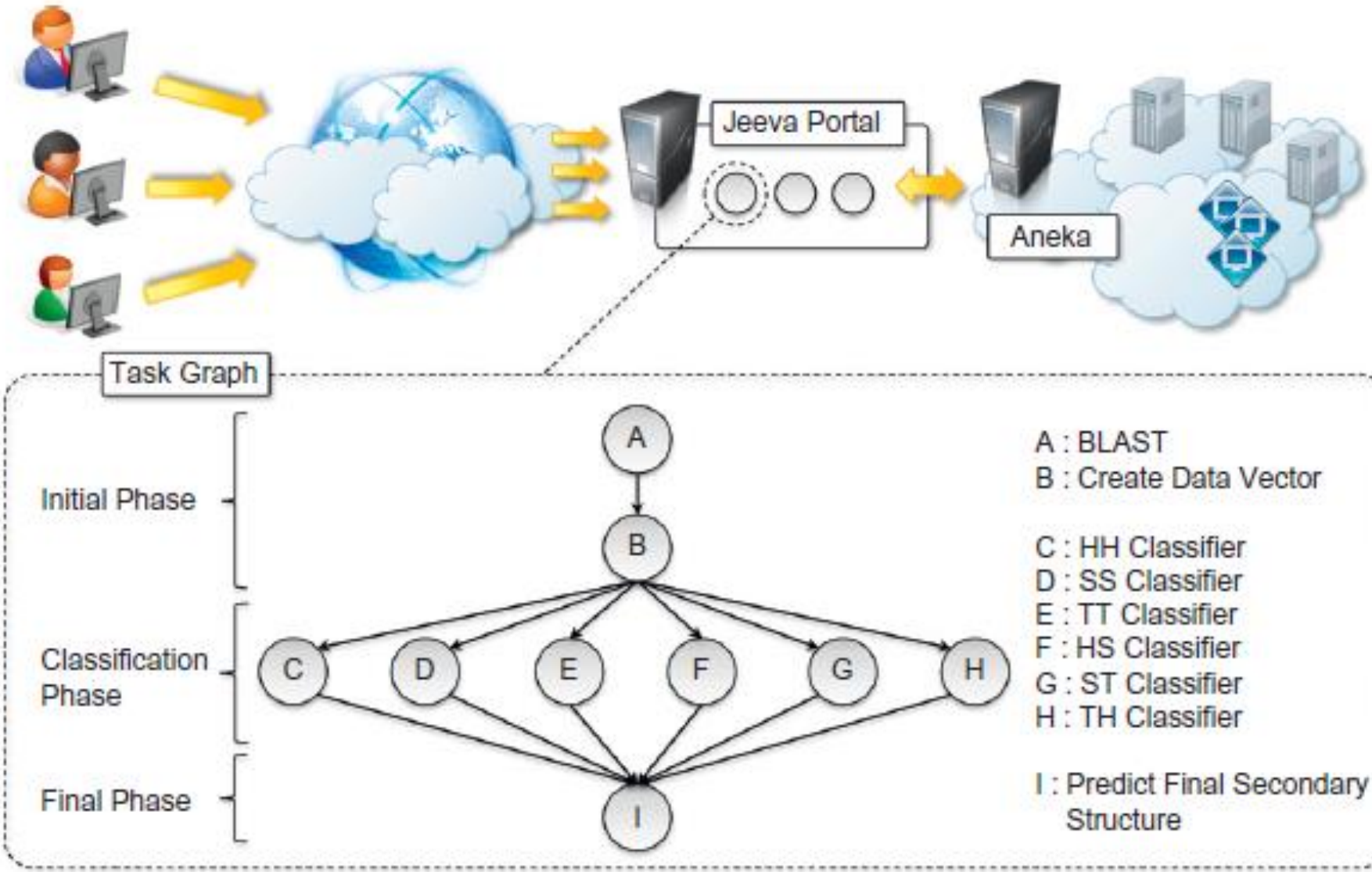
Biology: protein structure prediction

- 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 (see below figure).
- The prediction task uses machine learning techniques (support vector machines) for determining the secondary structure of proteins.
- These techniques translate 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*.

Biology: protein structure prediction

- These **three** phases have to be executed in sequence, we can perform parallel execution in the classification phase, where multiple classifiers are executed concurrently.
- This reduces 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.

Architecture and overview of the Jeeva Portal



Advantage

- 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.

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.

Biology: gene expression data analysis for cancer diagnosis

- The 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.

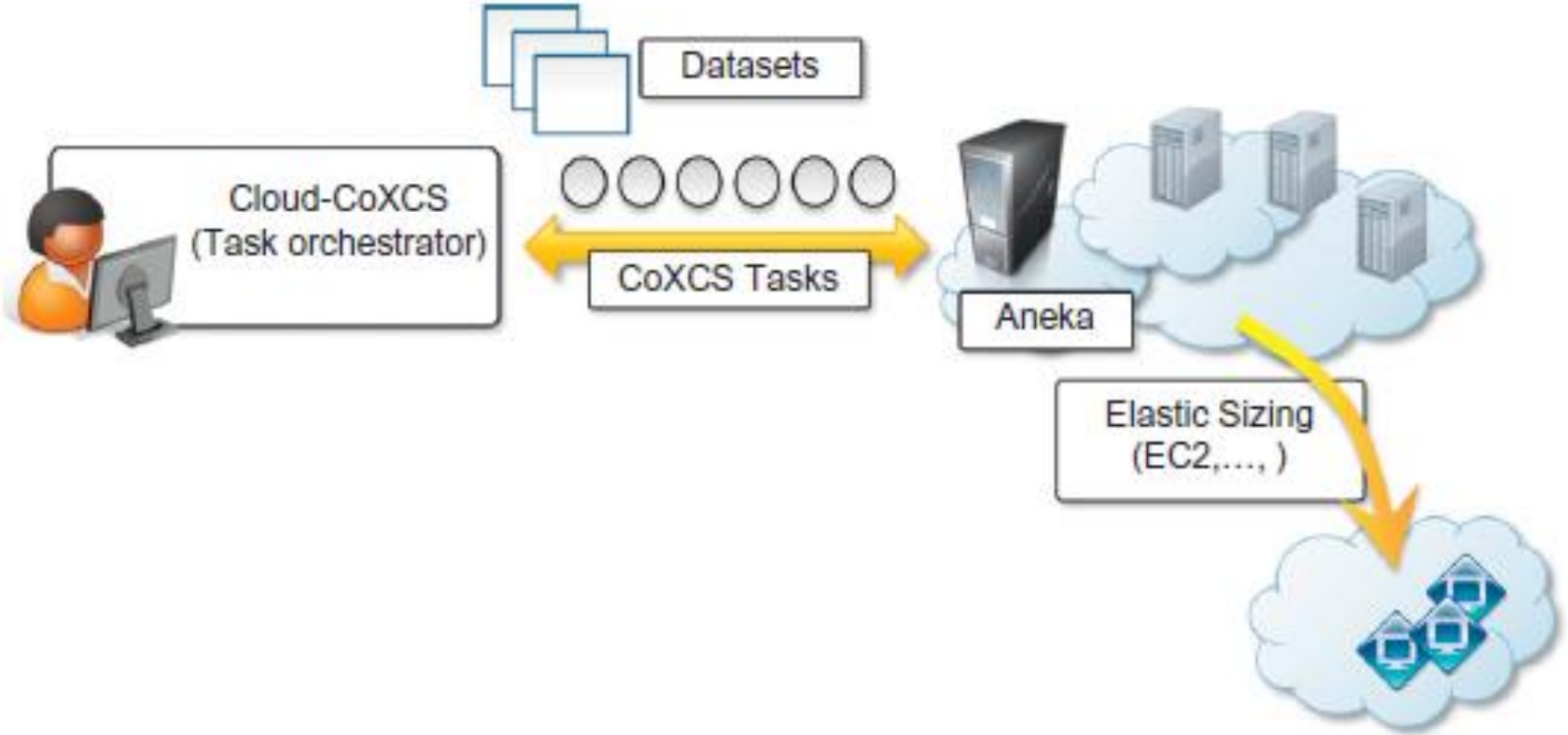
Biology: gene expression data analysis for cancer diagnosis

- 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.

Biology: gene expression data analysis for cancer diagnosis

- This problem is approached with learning classifiers, which generate a population of condition-action rules that guide the classification process.
- The eXtended Classifier System (XCS) has been utilized for classifying large datasets in bioinformatics and computer science domains.
- A variation of this algorithm, CoXCS, has proven to be effective in these conditions. CoXCS divides the entire search space into subdomains and employs the standard XCS algorithm in each of these subdomains.
- Such a process is computationally intensive but can be easily parallelized because the classifications problems on the subdomains can be solved concurrently.

Cloud-CoXCS (see below figure) 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.



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

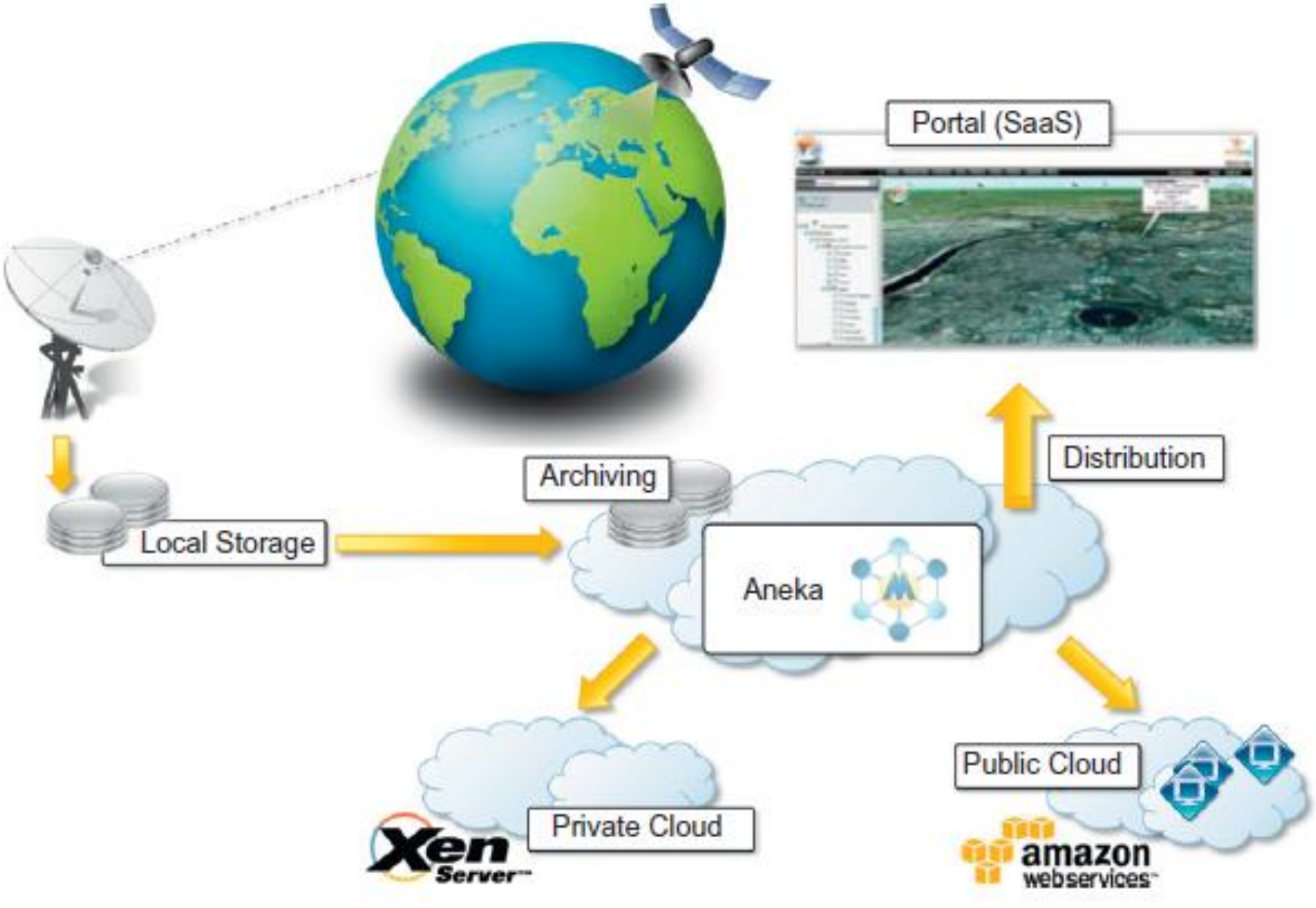
Geoscience: satellite image processing

- Geoscience applications collect, produce, and analyze massive amounts of geospatial and nonspatial data. As the technology progresses and our planet becomes more instrumented, volume of data that needs to be processed increases significantly.
- 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.

Geoscience: satellite image processing

- 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 provides 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 below figure integrates several technologies across the entire computing stack.

A cloud environment for satellite data processing



Geoscience: satellite image processing

- A SaaS application provides a collection of services for such tasks as geocode generation and data visualization.
- 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 on demand.
- The project demonstrates how cloud computing technologies can be effectively employed to offload local computing facilities from excessive workloads and leverage more elastic computing infrastructures.

Business and consumer applications

- CRM and ERP
  - Salesforce.com
  - Microsoft dynamics CRM
  - NetSuite
- Social Networking
  - Facebook
- Media Applications
  - Animoto
  - Maya rendering with Aneka
  - Video encoding on the cloud: Encoding.com
- Multiplayer Online Gaming

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.

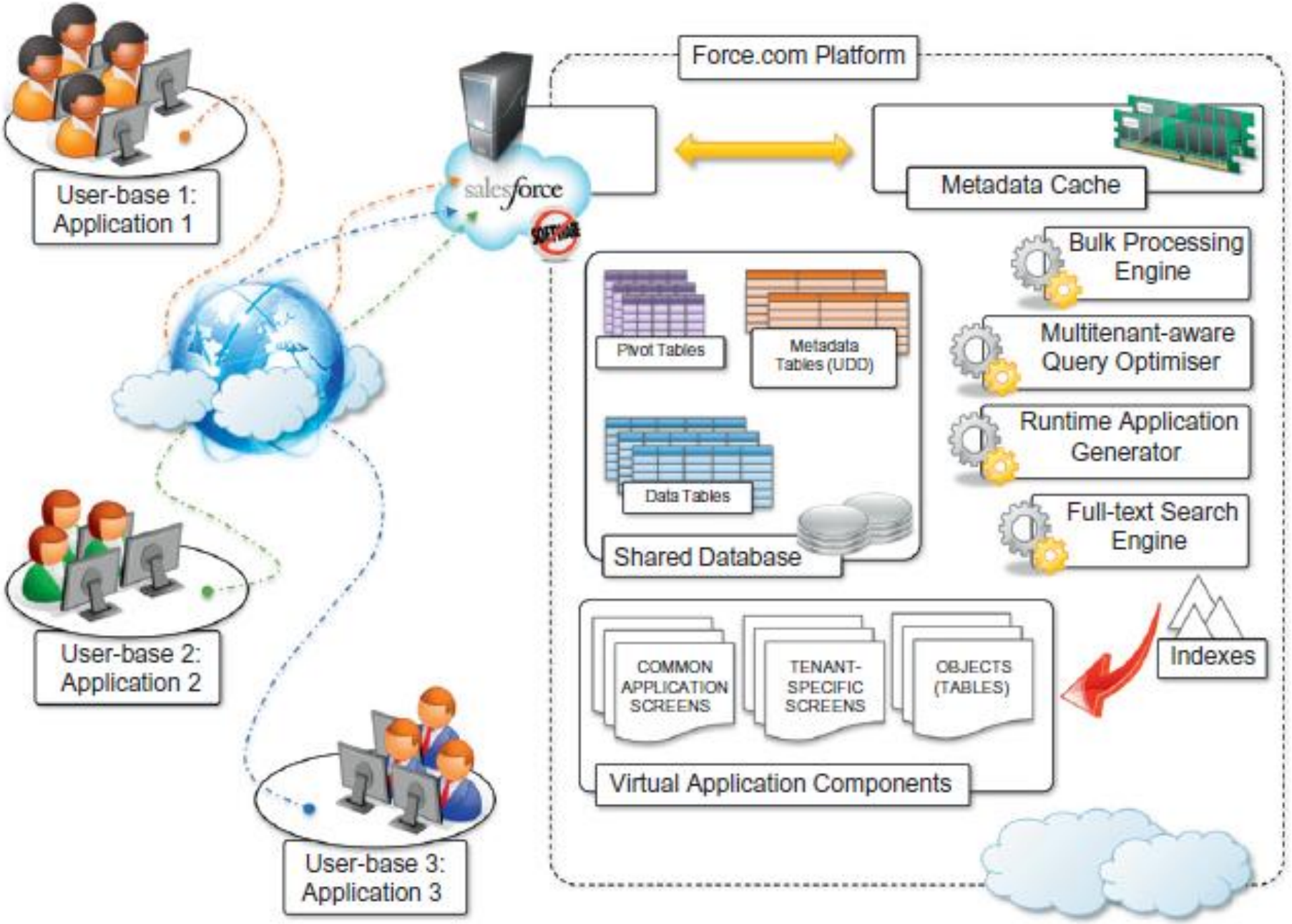
CRM and ERP

- **Customer relationship management (CRM)** and **enterprise resource planning (ERP)** applications are market segments that are flourishing in the cloud 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.
- 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.

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 plat- form.
- This represents scalable and high-performance middleware executing all the operations of all Salesforce.com applications.
- The architecture of the Force.com platform is shown in below figure.

Salesforce.com and Force.com architecture



Salesforce.com

- 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.
- 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.

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.

Microsoft dynamics CRM

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

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 CRM+***, 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.

NetSuite

- The ***NetSuite Business Operating System (NS-BOS)*** is a complete stack of technologies for building SaaS business applications 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 by ***SuiteBundler***.
- The entire infrastructure is hosted in the NetSuite datacenters, which provide warranties regarding application uptime and availability.

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.

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.

Facebook

- 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.

Facebook

- 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 graph of 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.

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.

Animoto

- ***Animoto*** is perhaps the most popular example of media applications on the cloud.
- The Website provides users with a very straight forward 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.

Animoto

- 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.

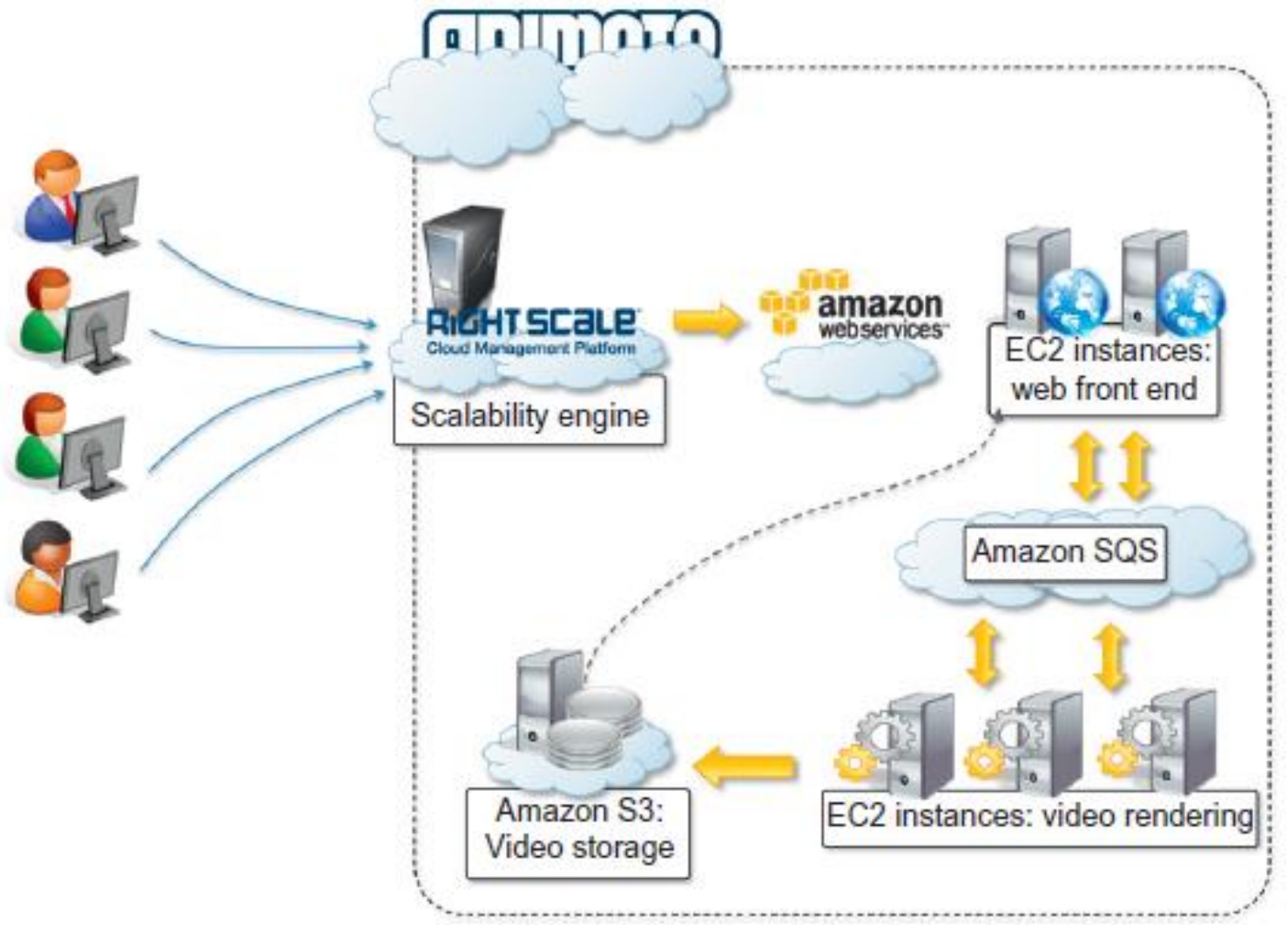
Animoto

- The infrastructure supporting Animoto is complex and is composed of different systems that all need to scale (see below figure).
- The core function is implemented on top of the Amazon Web Services infrastructure.
- In particular, it uses Amazon EC2 for the Web front-end and the worker nodes; Amazon S3 for the storage of pictures, music, and videos; and Amazon SQS for connecting all the components.
- The system's auto-scaling capabilities are managed by Rightscale, which monitors the load and controls the creation of new worker instances as well as their reclaim.

Animoto

- Front-end nodes collect the components required to make the video and store them in S3.
- Once the storyboard of the video is completed, a video-rendering request is entered into a SQS queue. Worker nodes pick up rendering requests and perform the rendering.
- When the process is completed, another message is entered into a different SQS queue and another request is served. This last queue is cleared routinely and users are notified about the completion.
- The life of EC2 instances is controlled by Rightscale, which constantly monitors the load and the performance of the system and decides whether it is necessary to grow or shrink.
- The architecture of the system has proven to be very scalable and reliable by using up to 4,000 servers on EC2 in peak times without dropping requests but simply causing acceptable temporary delays for the rendering process.

Animoto reference architecture



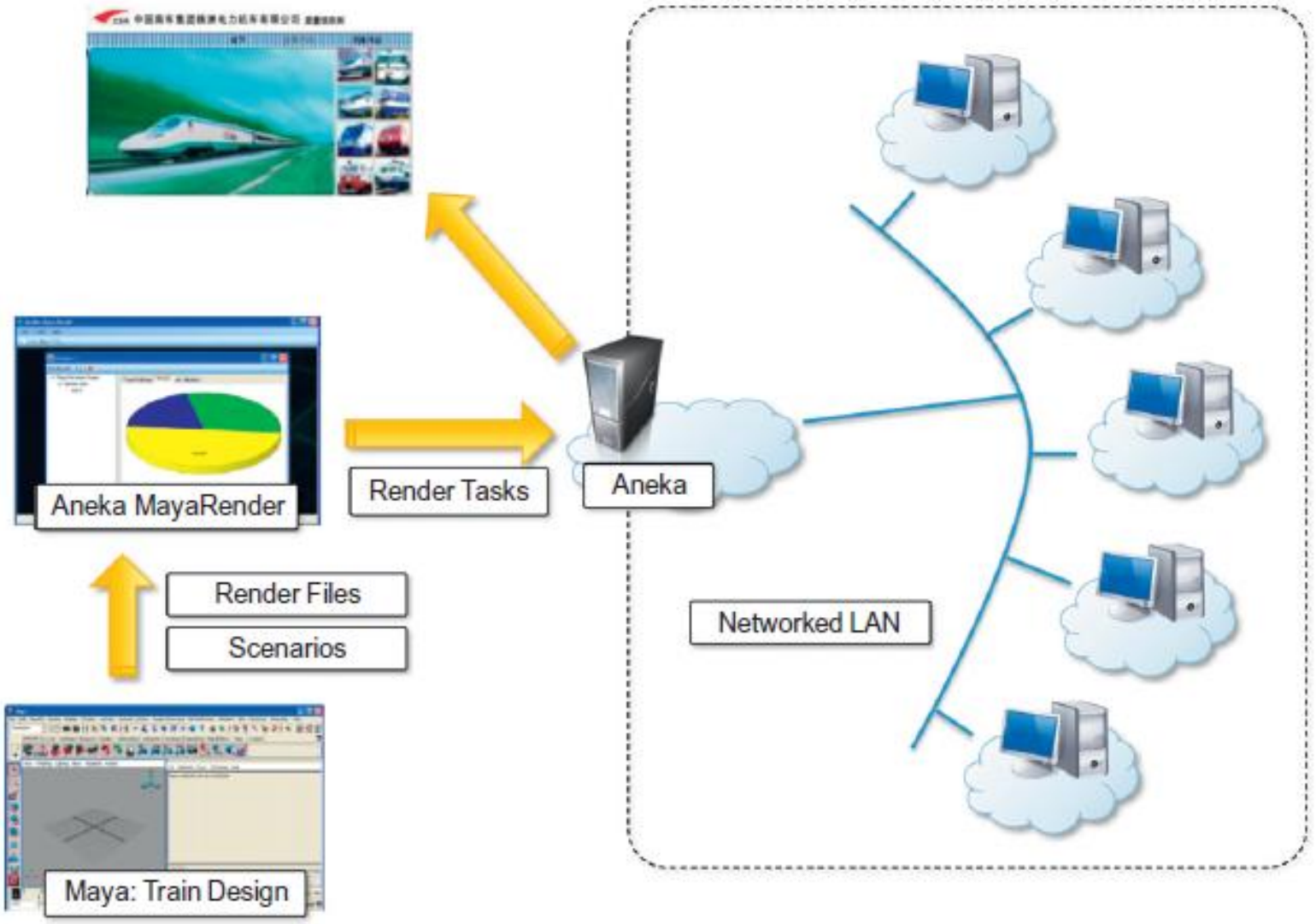
Maya rendering with Aneka

- 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.

Maya rendering with Aneka

- A private cloud solution for rendering train designs has been implemented by the engineering department of GoFront group, a division of China Southern Railway (see below figure).
- 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.

3D rendering on private clouds



Maya rendering with 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.

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.
- Encoding.com is a software solution that offers video-transcoding services on demand and leverages cloud technology to provide both the horse power 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).

Multiplayer 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.

Multiplayer online gaming

- 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.
- 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 prototype implementation of cloud-based game log processing has been implemented by Titan Inc. (see below figure)

Scalable processing of logs for network games