Cloud Platforms in Industry Amazon web services: - Compute services, Storage services, Communication services, Additional services. Google AppEngine: - Architecture and core concepts, Application life cycle, Cost model, Observations.

Cloud Applications: Scientific applications: - HealthCare: ECG analysis in the cloud, Biology: gene expression data analysis for cancer diagnosis, Geoscience: satellite image processing. Business and consumer applications: CRM and ERP, Social networking, media applications.

Module-5					
Q. 09	a	Explain the core components of Google app engine.	L1,L2,L3	3, 4	10
	b	Discuss in detail the following media applications of cloud computing technologies. i) Animoto ii) Maya Rendering with Aneka iii)Video encoding on cloud.	L1,L2,L3	3, 4	10
OR					
Q. 10	a	Explain in detail about the application of cloud computing in i)Healthcare: ECG analysis in the cloud ii)Geoscience: satellite image processing	L1,L2,L3	3, 4	10
	b	Explain Amazon web services(AWS) in detail.	L1,L2,L3	3, 4	10

1. What is AWS ,Different types of services offered by AWS?

Amazon Web Services (AWS) is a comprehensive cloud computing platform provided by Amazon. It offers a wide range of services to help businesses and developers build scalable, reliable, and cost-effective applications. AWS operates on a pay-as-you-go pricing model and is known for its flexibility, scalability, and global presence. It provides resources through APIs and web interfaces such as SOAP and REST.

Detailed Services Offered by AWS

1. **Compute Services**
    - **Amazon EC2 (Elastic Compute Cloud)**: Provides virtual servers, called instances, with customizable configurations (CPU, memory, and storage). It is a key service for deploying and managing applications.
    - **AWS Lambda**: Serverless computing that allows you to run code without provisioning servers. It automatically scales based on the request rate.
    - **Elastic Beanstalk**: A platform-as-a-service (PaaS) for deploying and managing web applications. It abstracts the infrastructure management.

- **Auto Scaling**: Ensures applications scale dynamically based on demand.

2. **Storage Services**
   - **Amazon S3 (Simple Storage Service)**: An object storage service offering secure and scalable storage for any type of data. Ideal for backups, archives, and application data.
   - **Amazon EBS (Elastic Block Store)**: Provides persistent block storage volumes for EC2 instances, supporting applications requiring high performance and low-latency storage.
   - **Amazon Glacier**: Cost-effective storage for long-term archival and backup data with retrieval times ranging from minutes to hours.
   - **Amazon Elastic File System (EFS)**: A scalable and managed shared file system for EC2 instances.
3. **Database Services**
   - **Amazon RDS (Relational Database Service)**: Managed database service for MySQL, PostgreSQL, SQL Server, Oracle, and more. Includes automatic backups, patching, and scaling.
   - **Amazon DynamoDB**: A fully managed NoSQL database designed for applications requiring high throughput and low latency.
   - **Amazon ElastiCache**: Provides in-memory caching for applications, enhancing performance for read-heavy workloads.
   - **Amazon Redshift**: A fully managed data warehouse for fast analytics on structured data.
4. **Networking Services**
   - **Amazon VPC (Virtual Private Cloud)**: Allows you to create isolated networks within AWS with granular control over IP ranges and subnets.
   - **Elastic Load Balancer (ELB)**: Distributes incoming application traffic across multiple EC2 instances for fault tolerance and scalability.
   - **Amazon Route 53**: A scalable DNS web service for managing domain names and routing users to applications.
5. **Developer and Deployment Tools**
   - **AWS CodeCommit**: A secure and scalable Git-based source control service.

- **AWS CodeDeploy**: Automates software deployment to EC2, on-premises servers, or Lambda functions.
- **AWS CodePipeline**: Helps automate CI/CD workflows, including building, testing, and deploying code.

6. **Machine Learning and Artificial Intelligence**
   - **Amazon SageMaker**: A fully managed service for building, training, and deploying machine learning models at scale.
   - **AWS Rekognition**: A service for image and video analysis, including object recognition and facial analysis.
   - **Amazon Polly**: Converts text to lifelike speech, enabling text-to-speech capabilities in applications.
7. **Analytics and Big Data**
   - **Amazon EMR (Elastic MapReduce)**: Processes big data using Apache Hadoop and Spark.
   - **Amazon Kinesis**: Real-time data streaming for collecting, processing, and analyzing data.
   - **AWS Glue**: A fully managed ETL service for data preparation and integration.
8. **Security and Identity Services**
   - **AWS IAM (Identity and Access Management)**: Manages user access and permissions for AWS resources.
   - **AWS Key Management Service (KMS)**: Centralized control for creating and managing encryption keys.
   - **AWS Shield**: Provides protection against Distributed Denial of Service (DDoS) attacks.
9. **Content Delivery and CDN**
   - **Amazon CloudFront**: A global content delivery network (CDN) for distributing data, videos, and APIs with low latency and high transfer speeds.

2. Describe Google AppEngine runtime environment and features

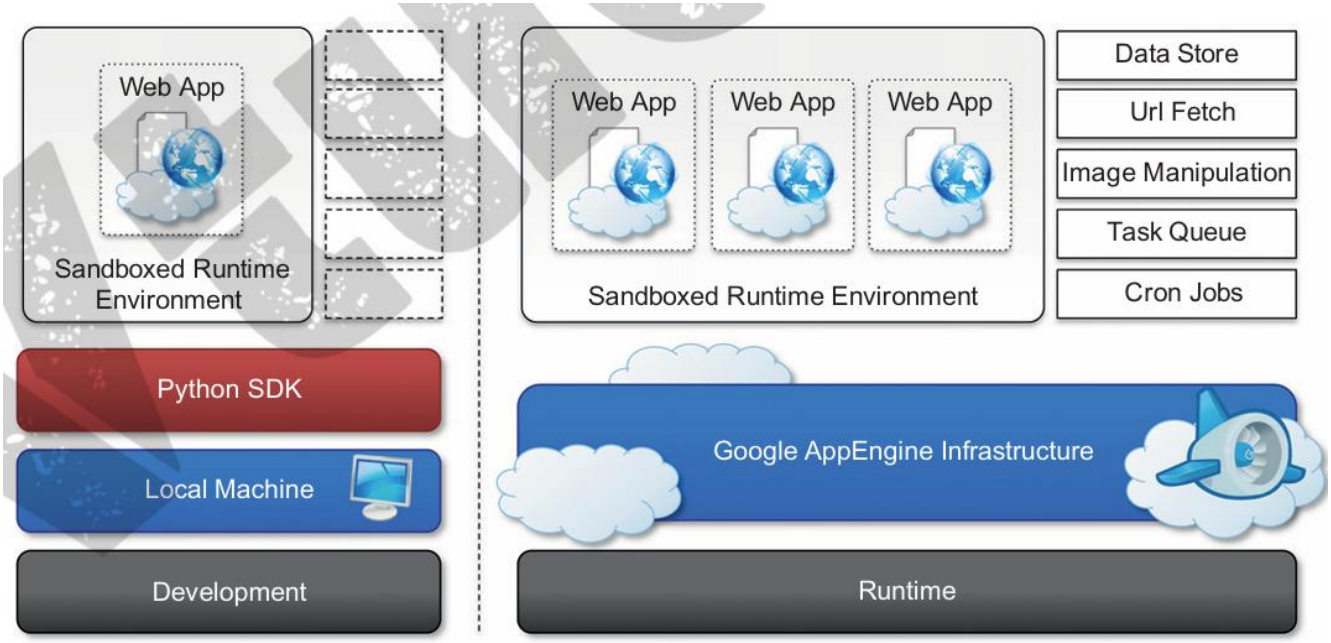

FIGURE 9.2

Google AppEngine platform architecture.

Components of Google AppEngine Runtime Environment

1. **Sandboxed Runtime Environment**
   - **Purpose**: The sandbox isolates web applications, ensuring they run securely without interfering with each other.
   - **Execution**: Each web application runs within its isolated environment with restricted access to system resources.
2. **Google AppEngine Infrastructure**
   - This layer manages the underlying hardware, networking, and scaling.
   - It ensures that applications have access to the required infrastructure, such as compute resources and storage.
3. **Development Environment**
   - **Python SDK**: Developers can create and test applications locally before deploying them to AppEngine.
   - **Local Machine**: Applications can be developed on the developer's system using tools provided by AppEngine SDKs.

4. **Runtime Environment**
    - Handles the actual execution of web applications deployed on the Google AppEngine infrastructure.

Built-in Services for Applications

Google AppEngine provides several built-in services for efficient web application development:

- **Data Store**: A highly scalable NoSQL database to store and retrieve application data.
- **URL Fetch**: Enables applications to make HTTP/HTTPS requests to external services.
- **Image Manipulation**: Provides tools for image processing (resizing, cropping, etc.).
- **Task Queue**: Allows background processing for long-running or asynchronous tasks.
- **Cron Jobs**: Schedules recurring tasks like cleanup operations or batch data processing.

Key Features of the Google AppEngine Runtime Environment

1. **Language Support**
    - Supports **Java**, **Python**, **Go**, and other languages with a sandboxed environment.
    - Includes SDKs for local development and tools for deployment.
2. **Sandboxed Environment**
    - Applications run in an isolated environment, preventing interference with the underlying infrastructure or other applications.
    - Certain operations, such as writing to the file system or arbitrary network access, are restricted to enhance security.
3. **Automatic Scalability**
    - AppEngine scales automatically by allocating more resources (e.g., virtual instances) as traffic increases.
    - Ensures efficient request handling by redirecting traffic to the least-loaded server or spinning up additional instances when required.
4. **Integrated Services**
    - **DataStore**: A NoSQL database optimized for web applications with high performance.
    - **MemCache**: An in-memory caching service for frequently accessed data to reduce latency.

- **Google Accounts**: Integrated user authentication and account management using Google credentials.

5. **Static File Hosting**
   - Optimized servers for hosting static content like CSS, JavaScript, HTML, and media files.
6. **Task Execution**
   - **Task Queues**: Enable asynchronous task processing to handle long-running or delayed operations.
   - **Cron Jobs**: Schedule recurring tasks to execute at specific times.
7. **Support for APIs and Services**
   - Includes APIs for messaging (XMPP), email, and image manipulation.
   - Allows integration with Google Cloud services like BigQuery, Pub/Sub, and Machine Learning APIs.
8. **High Availability and Fault Tolerance**
   - Applications are distributed across Google's data centers, ensuring high uptime and resilience to failures.
9. **Customizable Runtime Settings**
   - Developers can configure runtime environments through flexible YAML or JSON configuration files.

Explain in detail about the application of cloud computing in i)Healthcare: ECG analysis in the cloud ii)Geoscience: satellite image processing

i) Healthcare: ECG Analysis in the Cloud

1. **Overview**:
   - Cloud computing enables remote monitoring and analysis of electrocardiogram (ECG) data.
   - ECG measures the heart's electrical activity and detects arrhythmias, a common indicator of heart disease.

2. **Workflow**:
    - **Wearable Devices**: Equipped with ECG sensors to collect the patient's heartbeat data continuously.
    - **Mobile Devices**: Transmit the collected data to cloud-hosted Web services.
    - **Cloud Infrastructure**:
        - **SaaS (Software-as-a-Service)**: Web service stores data in platforms like Amazon S3 and serves as the user interface.
        - **PaaS (Platform-as-a-Service)** and **IaaS (Infrastructure-as-a-Service)**:
            - Scalable platforms like Aneka and Amazon EC2 handle processing tasks.
            - Workflow engines dynamically allocate resources for analyzing ECG waveform data.

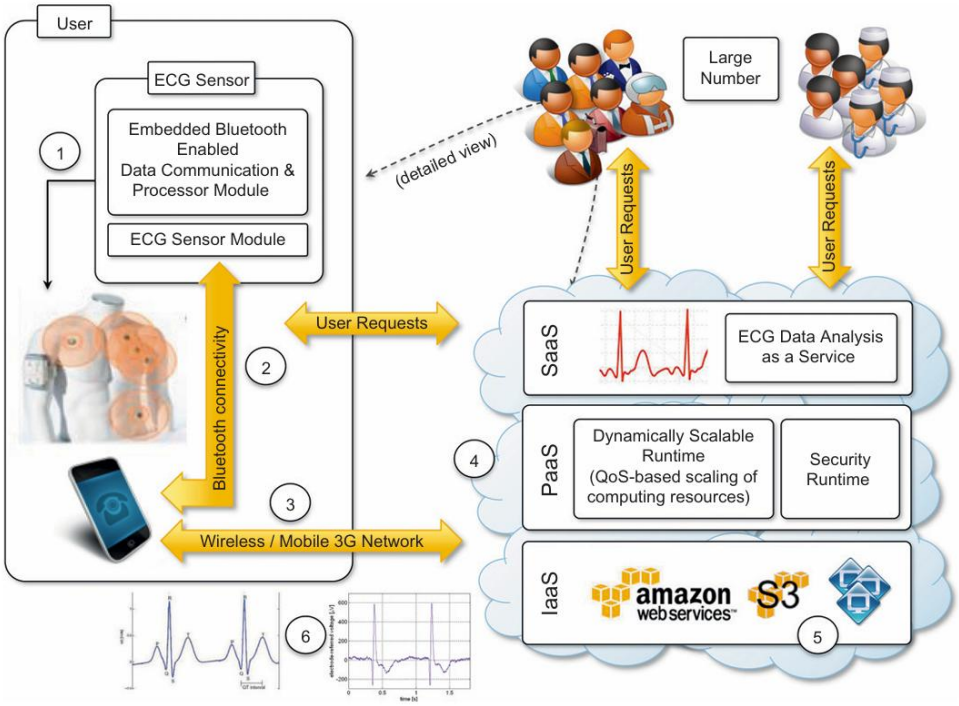

FIGURE 10.1
An online health monitoring system hosted in the cloud.

3. **Advantages**:
    - **Real-Time Monitoring**:
        - Continuous analysis and anomaly detection (e.g., arrhythmias).
        - Immediate alerts to doctors and emergency responders.

- **Elastic Infrastructure**:
  - Scales up or down based on demand, reducing operational costs for hospitals.
- **Accessibility**:
  - Data and services are available via any Internet-enabled device.
  - Seamless integration with hospital systems for diagnostics.
- **Cost Efficiency**:
  - Pay-per-use pricing eliminates the need for upfront investments in infrastructure.

4. **Impact**:
   - Enables constant monitoring of high-risk patients without requiring frequent hospital visits.
   - Ensures timely interventions during cardiac emergencies.

---

ii) Geoscience: Satellite Image Processing

1. **Overview**:
   - Cloud computing simplifies the storage, processing, and analysis of massive satellite datasets used in geoscience applications.
   - Applications include environmental monitoring, disaster management, and land use analysis.
2. **Workflow**:
   - **Data Collection**: Satellites generate large volumes of data, including high-resolution images and other geospatial information.
   - **Data Storage**:
     - Cloud platforms (e.g., Amazon S3, Google Cloud Storage) provide scalable storage for the data.
     - Distributed storage enables efficient access and redundancy.
   - **Data Processing**:
     - Cloud platforms leverage parallel computing frameworks like MapReduce and Apache Spark for processing data at scale.

- Machine learning models hosted on the cloud are used for advanced analytics, such as pattern detection and anomaly identification.

- **Visualization and Analysis**:
  - Processed data is made available through web-based dashboards or APIs for researchers and policymakers.

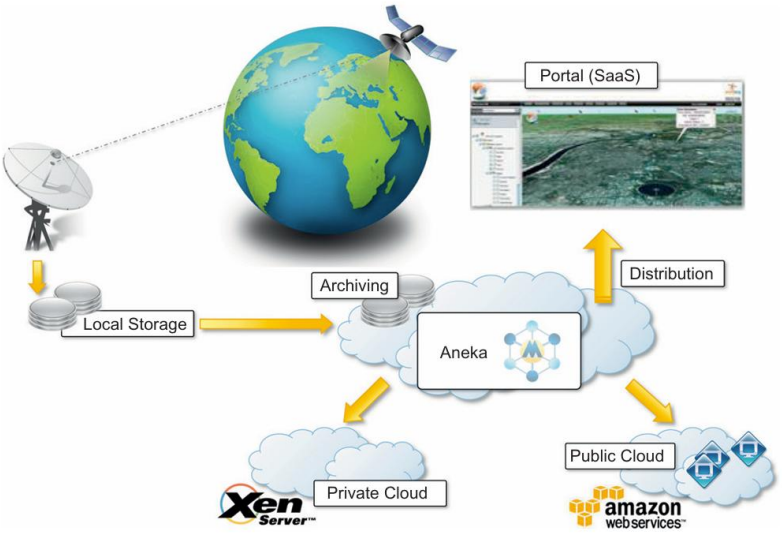

FIGURE 10.4
A cloud environment for satellite data processing.

3. **Advantages**:
   - **Scalability**:
     - Handles vast amounts of data from multiple satellites efficiently.
     - Dynamically adjusts resources based on computational requirements.
   - **Collaboration**:
     - Facilitates sharing of datasets and insights across research teams globally.
   - **Cost-Effectiveness**:
     - Eliminates the need for expensive in-house high-performance computing systems.
     - Offers pay-per-use and subscription models.
   - **Speed**:
     - Accelerates the analysis of critical data, aiding in disaster response and decision-making.

4. **Impact**:
    - Improves precision in predicting weather patterns, monitoring deforestation, and managing natural resources.
    - Enhances disaster management by quickly identifying affected regions and assessing damage.

Discuss in detail the following media applications of cloud computing technologies. i) Animoto ii) Maya Rendering with Aneka iii)Video encoding on cloud.

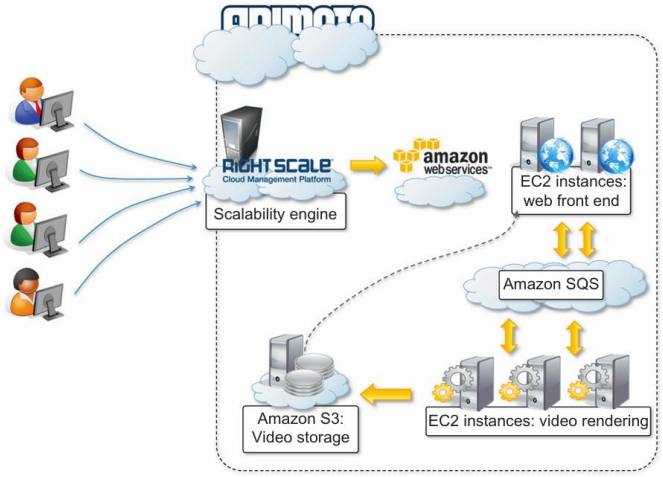

FIGURE 10.8
Animoto reference architecture.

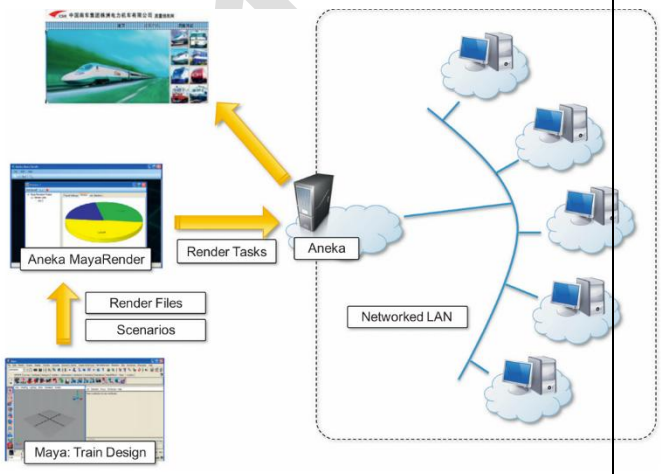

FIGURE 10.9
3D rendering on private clouds.

i) Animoto

1. **Overview**:
    - Animoto is a cloud-based video creation platform that automates the process of producing professional-quality videos.
    - It leverages cloud computing to handle resource-intensive tasks like video rendering and processing.
2. **Workflow**:
    - **User Input**:
        - Users upload photos, videos, and music to the platform.
        - They select templates and customize settings (e.g., transitions, text, effects).
    - **Cloud Processing**:
        - Animoto's backend distributes tasks across a cloud infrastructure.

- This parallel processing speeds up rendering and ensures scalability.

- **Output Delivery**:
  - The processed video is delivered to the user in various formats suitable for social media, presentations, or downloads.

3. **Advantages**:
   - **Scalability**: Handles varying workloads effectively, allowing simultaneous processing for multiple users.
   - **Ease of Use**: Provides non-technical users with professional video creation tools.
   - **Cost-Effectiveness**: Users pay for the service without needing high-end hardware or software.
   - **Speed**: Quick video rendering enabled by distributed computing.
4. **Applications**:
   - Marketing and branding videos.
   - Personal slideshows and social media content.

---

ii) Maya Rendering with Aneka

1. **Overview**:
   - Maya is a professional 3D computer graphics software used for animation, modeling, simulation, and rendering.
   - Rendering is computationally expensive, and integrating Maya with Aneka, a cloud computing framework, accelerates the process.
2. **Workflow**:
   - **Scene Creation**:
     - Users design and develop 3D scenes using Maya.
   - **Task Submission**:
     - Maya projects are uploaded to the Aneka framework, which distributes rendering tasks across cloud resources.

- **Distributed Rendering**:
  - Aneka divides the rendering process into smaller tasks.
  - These tasks are executed simultaneously on multiple virtual machines, leveraging parallel computing.
- **Result Compilation**:
  - The rendered frames are collected, combined, and delivered to the user.

3. **Advantages**:
   - **High Performance**: Reduces rendering time significantly by using distributed resources.
   - **Flexibility**: Supports multiple rendering engines and workflows.
   - **Scalability**: Handles complex rendering projects by allocating more resources dynamically.
   - **Cost-Efficiency**: Pay-per-use pricing reduces the need for expensive hardware investments.
4. **Applications**:
   - Animated movies and visual effects in films.
   - Architectural visualizations and gaming industry assets.

---

iii) Video Encoding on the Cloud

1. **Overview**:
   - Video encoding is the process of converting raw video files into compressed formats (e.g., MP4, H.264) for storage and streaming.
   - Cloud platforms like AWS, Azure, and Google Cloud offer encoding services to process video efficiently.
2. **Workflow**:
   - **Input**:
     - Users upload raw video files to the cloud encoding service.
   - **Encoding**:
     - The cloud platform utilizes distributed computing to process video in parallel.

        - Encoding options include resolution adjustment, format conversion, and adding subtitles.
    - **Output**:
        - Encoded videos are stored in the cloud for on-demand streaming or download.

3. **Advantages**:
    - **Efficiency**:
        - Parallel processing reduces encoding time, even for large video files.
    - **Scalability**:
        - Can handle large-scale encoding for live streaming or batch processing.
    - **Cost Savings**:
        - Eliminates the need for dedicated on-premises encoding hardware.
    - **Customization**:
        - Offers various encoding presets and advanced options for specific needs.
4. **Applications**:
    - Online streaming platforms (e.g., Netflix, YouTube).
    - Video marketing campaigns and OTT (Over-The-Top) media services.

CRM and ERP in Cloud Computing

Customer Relationship Management (CRM) and **Enterprise Resource Planning (ERP)** are two major segments leveraging cloud computing technologies. CRM applications, being more mature, offer accessible and cost-effective solutions for small enterprises and start-ups, while ERP applications are still evolving due to challenges in transitioning from established in-house systems.

Cloud CRM applications allow businesses to manage customer relationships efficiently without the need for significant upfront investment. They provide access to customer and business data from anywhere, fostering their widespread adoption. Conversely, cloud-based ERP solutions face hurdles such as high migration costs and competition from existing in-house ERP systems. These solutions integrate enterprise-wide operations

like finance, HR, and supply chain management, aiming to offer a unified view of organizational processes. However, their adoption is slower as organizations often hesitate to shift from well-established systems.

Examples of Cloud-Based CRM and ERP Solutions

1. Salesforce.com

Salesforce.com is one of the most popular cloud CRM solutions, used by over 100,000 customers. It is built on the Force.com platform, which supports the entire lifecycle of cloud applications.

- **Features**:
  - Provides scalable, high-performance CRM solutions.
  - Enables customization through its metadata architecture, where application logic and data are stored as metadata.
  - Offers full-text search capabilities for handling large datasets efficiently.
  - Allows customization using visual tools or APIs in popular programming languages.
  - Includes APEX, a Java-like programming language, for advanced scripting and querying.
- **Advantages**:
  - Multi-tenant architecture ensures uniform application execution.
  - Customizable workflows and scalable infrastructure.

---

2. Microsoft Dynamics CRM

Microsoft Dynamics CRM provides customer relationship management solutions either as on-premise installations or cloud-based subscriptions.

- **Features**:
  - Hosted in Microsoft's global datacenters with a 99.9% SLA.
  - Offers functionalities for marketing, sales, and advanced customer relationship management.
  - Accessible via web browsers or APIs using SOAP and RESTful Web services.
  - Allows development of plug-ins for custom behaviors and integration with Windows Azure for new feature development.
- **Advantages**:

- Seamless integration with other Microsoft products.
- Scalability and flexibility to meet diverse business needs.

---

3. NetSuite

NetSuite provides a comprehensive suite of applications for managing various aspects of a business.

- **Features**:
  - Major offerings include NetSuite Global ERP, Global CRM, and Global Ecommerce.
  - Offers an integrated solution, NetSuite OneWorld, combining ERP, CRM, and ecommerce functionalities.
  - Powered by two datacenters in the US, ensuring 99.5% uptime.
  - Provides an online development environment, SuiteFlex, for creating and integrating new applications.
  - Includes SuiteBundler for packaging and distributing applications.
- **Advantages**:
  - High reliability due to redundant infrastructure.
  - Comprehensive business management capabilities in a single platform.

Productivity Applications of Cloud Computing

Cloud computing enables productivity applications by replicating common desktop tasks like document storage, office automation, and full desktop environments on the cloud. These applications allow users to work seamlessly across devices and locations, significantly enhancing accessibility and efficiency.

Dropbox and iCloud

Dropbox and iCloud are prominent cloud-based document storage and synchronization solutions. Dropbox provides a folder-based synchronization system across platforms like Windows, Mac, Linux, and mobile devices. Changes in the Dropbox folder are automatically synchronized, allowing users to access updated content on any device seamlessly. This platform's key advantage lies in its platform-agnostic approach and ease of use.

iCloud, designed for Apple's ecosystem, automatically synchronizes documents, photos, and videos across iOS and Mac devices. It eliminates manual operations by ensuring that changes on one device are instantly reflected on others. While iCloud is highly efficient for Apple users, its limitations include a lack of support for non-iOS platforms and the absence of a web interface for broader access.

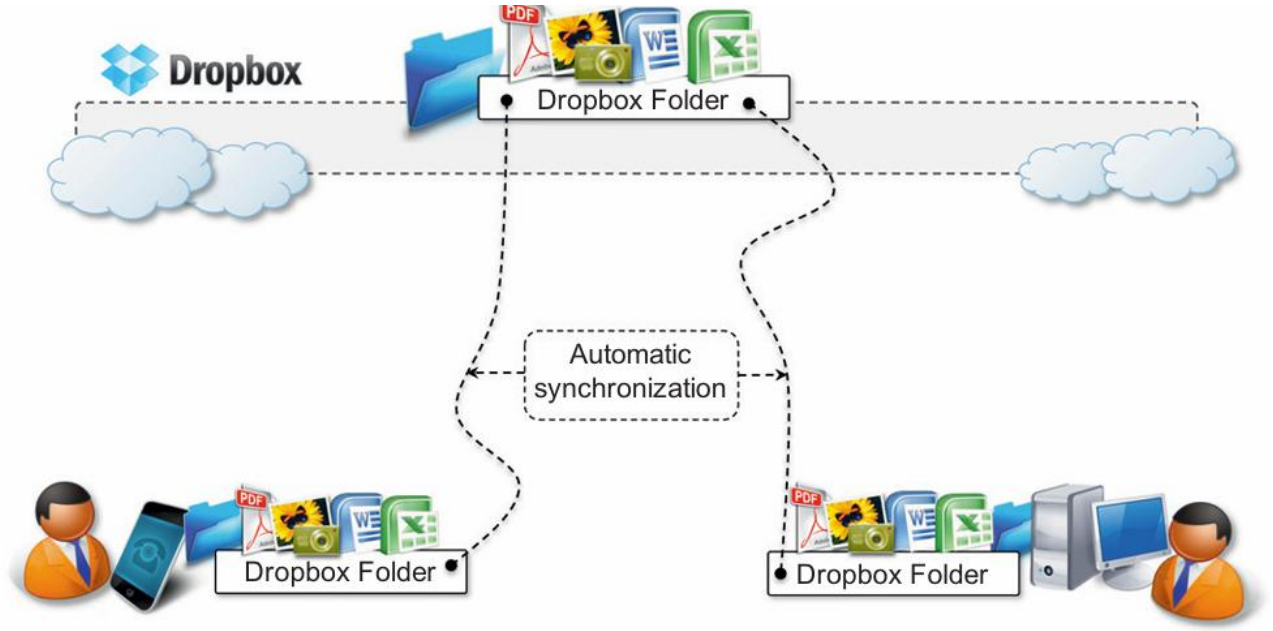

FIGURE 10.6

Dropbox usage scenario.

Google Docs

Google Docs is a SaaS application that delivers office automation capabilities like text editing, spreadsheets, and presentations. Built on Google's distributed computing infrastructure, it scales dynamically to accommodate user demand.

Key features include collaborative editing, eliminating the need for emailing and synchronizing files among team members. Documents are stored on Google's infrastructure, making them accessible from any internet-connected device. Additionally, Google Docs supports offline work and various file formats, allowing easy import/export from popular desktop office suites. The platform demonstrates the benefits of cloud computing by providing ubiquitous access, cost-effectiveness, and minimal maintenance.

Cloud Desktops: EyeOS and XIOS/3

EyeOS and XIOS/3 replicate desktop environments in the cloud, making them accessible via web browsers. EyeOS uses AJAX technologies to create a desktop-like experience, complete with pre-installed applications for document and file management. Users can access their virtual desktops from any device, while

organizations can deploy private EyeOS clouds to centralize employee desktop environments. EyeOS supports application development through APIs and simplifies user interaction by loading JavaScript-based interfaces.

XIOS/3, part of CloudMe, leverages XML extensively for rendering user interfaces, managing file systems, and developing applications. The client-side handles most operations, while the server manages core tasks like document transactions and collaborative editing. XIOS/3 offers a development environment (XIDE) for creating applications using visual tools and XML-based logic, making it a robust platform for service integration and peer collaboration.

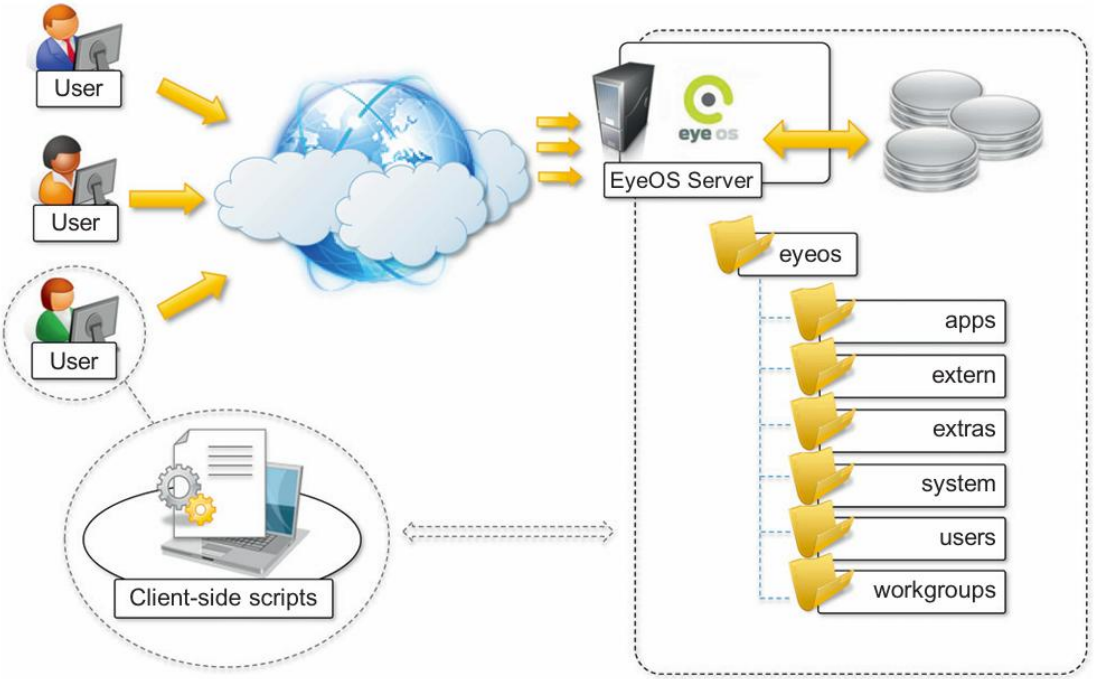

FIGURE 10.7
EyeOS architecture.

Social Networking Applications in Cloud Computing

Social networking applications have grown exponentially, becoming some of the most visited sites on the web. Platforms like Facebook and Twitter rely on cloud computing technologies to handle their vast user bases and increasing traffic. The ability to scale dynamically while maintaining system uptime is a critical feature that supports their growth and operations.

Facebook

Facebook serves over 800 million users, requiring a robust and scalable infrastructure. Its ability to continuously add capacity while delivering a seamless user experience is a testament to its reliance on cloud technologies.

1. **Infrastructure**:
    - Facebook operates two custom-built, highly optimized data centers designed to reduce costs and environmental impact.
    - These data centers use inexpensive hardware paired with customized, open-source software systems to achieve scalability and efficiency.
2. **Technology Stack**:
    - The backbone of Facebook is the LAMP stack (Linux, Apache, MySQL, PHP), accompanied by additional in-house developed services.
    - These services are written in various programming languages and handle functionalities like search, news feeds, and notifications.
    - User interactions and data are managed through a **social graph**, which links interrelated user information. MySQL clusters store key-value pairs that are cached for fast retrieval.
3. **Tools and Frameworks**:
    - **Thrift**:
        - A framework facilitating cross-language development.
        - Handles data serialization, deserialization, communication, and client/server boilerplate code.
        - Allows developers to create new services quickly and integrate existing ones.
    - **Scribe**:
        - Aggregates streaming log feeds for better system monitoring and analytics.
    - Additional internal tools support alerting and monitoring for efficient system management.
4. **APIs for Third-Party Integration**:
    - Facebook provides APIs to enable the integration of third-party applications, such as social games and quizzes.
    - These APIs leverage the core infrastructure to extend functionality and user engagement.