

# Cloud Applications Exam Notes (with Challenges)

## Scientific Applications

### Healthcare: ECG Analysis in Cloud

#### ECG Basics

- **ECG (Electrocardiogram):** The electrical expression of the heart's contractile movement
- Records electrical activity of heart muscle over time, producing a waveform representing heartbeat
- Analysis of waveform is crucial for identifying arrhythmias and detecting heart disease
- "Arrhythmias" = irregular heartbeats - can be too fast (tachycardia), too slow (bradycardia), or irregular

#### Cloud Integration Architecture

Layer	Components	Function
SaaS	Web service front-end	User interface, analysis results, notifications
PaaS	Scalable cloud platform (e.g., Aneka)	Workflow management, task distribution
IaaS	Amazon S3, EC2	Data storage, computation resources

#### Implementation Flow

1. Wearable ECG sensors monitor patient continuously
2. Data transmitted to mobile device
3. Mobile forwards data to cloud-hosted web service
4. Web service (SaaS) stores ECG data in Amazon S3 (IaaS)
5. Processing requests sent to scalable cloud platform (PaaS)

6. Aneka controls EC2 instances for ECG processing: \*  
Waveform extraction \* Comparison with reference waveforms
7. If anomalies found, notifications sent to medical personnel

## Advantages

- Reduces computational burden on doctors
- Eliminates need for large computing infrastructure investment
- **Elasticity**: Cloud infrastructure scales based on demand
- **Ubiquity**: Services accessible from anywhere with minimal downtime
- Medical data can be gathered and distributed globally

## Challenges

Challenge	Description
<b>Data Security &amp; Privacy</b>	Protecting sensitive patient data from breaches
<b>Signal Quality &amp; Compliance</b>	Ensuring quality from wearable devices and proper patient usage
<b>Interoperability</b>	Integrating with legacy hospital IT systems
<b>Real-time Processing</b>	Minimizing latency for critical alerts
<b>Regulatory Compliance</b>	Adhering to healthcare data regulations (e.g., HIPAA)

# Biology: Protein Structure Prediction

## Protein Fundamentals

- Large molecules essential for cell function, body structure, and regulation
- Found throughout body (muscle, bone, skin, hair); ~20% of total body weight
- Composed of hundreds/thousands of amino acids in chains

- 20 different types of amino acids can be combined
- Sequence determines unique 3D structure and function
- Examples: Keratin (hair, nails), hemoglobin (oxygen transport)

## Protein Structure Hierarchy

Level	Description	Example
<b>Primary</b>	Linear sequence of amino acids	Unique to each protein (e.g., Hemoglobin)
<b>Secondary</b>	3D form of local segments	Alpha-Helix, Beta-Pleated Sheets
<b>Alpha-Helix</b>	H-bonds between N-H and C=O groups	Common secondary structure
<b>Beta-Sheets</b>	H-bonds between adjacent strands	Composed of 3-10 amino acid beta-strands

## Cloud Computing Need

- Predicting geometric structure from gene sequence is computationally intensive
- Requires identifying structure that minimizes energy
- Involves investigating massive number of states with extensive calculations
- Cloud provides on-demand computational power without owning clusters
- Pay-per-use model avoids bureaucratic hurdles for accessing parallel/distributed computing

## JEEVA Portal Implementation

- Integrated web portal for protein structure prediction using cloud technologies
- Offloads prediction to Aneka-based computing cloud
- Uses machine learning (often support vector machines) for secondary structure determination

- Three-class pattern recognition problem (H, E, C)

## Workflow Phases

1. **Initial Phase:** \* BLAST (Basic Local Alignment Search Tool)  
\* PSI-BLAST to obtain PSSM features \* Feature vector creation for classification
2. **Classification Phase:** \* Multiple concurrent classifiers (HH, SS, TT, HS, ST, TH) \* Each reads data vector and generates classification results
3. **Final Prediction Phase:** \* Combines all classification results

## Technology Stack

- Prediction algorithm translated into task graph
- Submitted to Aneka (scalable cloud middleware)
- Aneka harnesses spare CPU cycles from heterogeneous networks/datacenters
- Results visualized through portal interface

## Challenges

Challenge	Description
<b>Computational Complexity</b>	Force field inaccuracies and conformational search limitations
<b>Data Volume &amp; Transfer</b>	Handling massive datasets generated during simulations
<b>Model Accuracy</b>	Ensuring accuracy for proteins with no known homologous structures
<b>Cost Management</b>	Balancing resources for large-scale computations and storage
<b>Software Portability</b>	Deploying complex bioinformatics workflows in cloud environments

# Geoscience: Satellite Image Processing

## Geoscience Data Types

Type	Description	Examples
Spatial Data	Geographic component representing location, size, shape	Satellite imagery, GPS data, aerial photos
Non-Spatial Data	Attribute data without geographical components	Demographic information, temporal measurements

## GIS Applications

- Capture, store, manipulate, analyze, manage, and present geographically referenced data
- Applied in diverse domains: advanced farming, civil security, natural resources management

## Cloud Implementation (Dept. of Space, India)

Layer	Components	Function
SaaS	Application layer	Geocode generation, data visualization
PaaS	Aneka platform	Controls data import, executes image-processing tasks
Infrastructure	Xen cloud private	Dynamically provisions resources on demand

## Challenges

Challenge	Description
Data Volume & Storage	Managing petabytes to exabytes exceeding standalone capacity
I/O Bottlenecks	Transferring and accessing large raster data files efficiently
Processing Complexity	Handling complex transformations for time-series analysis
Format Heterogeneity	Managing various file formats (HDF, GeoTIFF, JP2000)
Signal vs. Noise	

Challenge	Description
	Accurately retrieving geophysical variables from raw data

## Business and Consumer Applications

### CRM and ERP Systems

#### CRM (Customer Relationship Management)

- Cloud-hosted software for tracking business and customer data
- Helps businesses nurture relationships with leads and clients
- Consolidates communications, documents, quotes, purchases, and tasks
- Enables teams to access details at the right time for sales or service

#### Top Cloud CRM Solutions

Solution	Best For
Salesforce	Improving sales functions
Microsoft Dynamics 365	Enterprises
HubSpot	Teams new to CRM
Oracle NetSuite	E-commerce
Zoho CRM	Remote/hybrid sales teams
Pipedrive	Visual sales pipeline tracking

#### Salesforce Architecture (Force.com)

- Core logic and business rules stored as metadata
- Runtime engine executes logic by retrieving metadata and performing operations
- Applications in isolated containers logically share database structure

- Full-text search engine supports runtime engine for effective user experience

## ERP (Enterprise Resource Planning)

- Cloud-hosted software helping enterprises manage business data
- Integrates finance, accounting, HR, manufacturing, supply chain, project management, CRM
- Pay-per-use model based on cloud resource utilization
- Makes traditionally expensive systems accessible to SMBs

## Top Cloud ERP Solutions

Solution	Best For
Microsoft Dynamics 365 Business Central	Best overall
SAP Business One Professional	Customization
SYSPRO	Manufacturing
QT9	Real-time reporting
Epicor Prophet 21 ERP	Distributors
Oracle NetSuite OneWorld	Global companies
Acumatica	Easy pricing

## CRM/ERP Challenges

Challenge	Description
<b>Data Security &amp; Privacy</b>	Protecting sensitive data and compliance with regulations like GDPR
<b>Integration Issues</b>	Connecting with on-premise systems and third-party applications
<b>Customization Complexity</b>	Balancing flexibility with maintainability
<b>Cost Management</b>	Managing long-term subscription costs vs. initial investment
<b>User Resistance</b>	Overcoming employee reluctance to adopt new systems
<b>Connectivity Dependence</b>	Reliance on stable internet connection

Challenge	Description
Vendor Lock-in	Limited flexibility and costly future migrations

## Productivity Applications

### Cloud Storage Solutions

Solution	Key Features
Dropbox	File synchronization across platforms (Windows, Mac, Linux, mobile)
iCloud	Automatic syncing of documents, photos, videos across Apple devices
Windows Live	Microsoft's cloud storage solution
Amazon Cloud Drive	Amazon's storage offering
CloudMe	Cloud storage with XIOS/3 web desktop environment

### Google Docs

- SaaS application offering basic office automation with collaborative editing
- Runs on Google's distributed computing infrastructure for dynamic scaling
- Creates/edits text documents, spreadsheets, presentations, forms, drawings
- Aims to replace desktop products like Microsoft Office and OpenOffice
- Supports collaborative editing to eliminate email-based synchronization

### Cloud Desktops

Solution	Features
EyeOS	Web desktop with pre-installed applications for file/document management
XIOS/3	Web desktop environment leveraging XML for UI rendering, business logic



## Productivity Challenges

Challenge	Description
Security & Privacy	Protecting sensitive documents from unauthorized access
Interoperability	Ensuring seamless operation across different platforms
Service Reliability	Maintaining consistent performance and availability
Connectivity Dependence	Reliance on stable internet access
Feature Limitations	Cloud versions may lack desktop counterpart features
Cost Management	Managing subscription costs for multiple users and storage
Data Ownership	Concerns about control over data on third-party servers

## Social Networking

### Key Characteristics

- Leverages cloud computing to sustain traffic and serve millions of users
- Continuously adds capacity while systems are running
- Examples: Facebook, WhatsApp, Twitter, Instagram

### Social Networking Challenges

Challenge	Description
Data Security & Privacy	Protecting vast amounts of personal user data
Scalability & Performance	Handling massive, fluctuating user traffic with low latency
Content Moderation	Managing user-generated content at scale
Cost Unpredictability	Managing storage and processing costs for immense data volumes
Regulatory Compliance	Adhering to diverse global regulations
Real-time Processing	Handling real-time interactions for millions of concurrent users

Challenge	Description
Expertise Gaps	Finding skilled professionals for large-scale infrastructure

## Media Applications

### Characteristics

- Significantly benefited from cloud computing
- Video processing operations ideal for cloud environments
- Computationally intensive tasks easily offloaded to cloud infrastructures

### Animoto Example

- Creates videos from user-submitted images, music, and video fragments
- Users select theme, upload/order content, choose music, render video
- AI engine selects animation and transition effects based on input

### Animoto Architecture (AWS)

Component	Function
Amazon EC2	Web front-end and worker nodes
Amazon S3	Stores pictures, music, and videos
Amazon SQS	Connects all components
RightScale	Manages auto-scaling capabilities

### Process Flow

1. Front-end nodes collect video components, store in S3
2. Completed storyboard triggers rendering request to SQS queue

3. Worker nodes pick up rendering requests
4. Completion message added to different SQS queue
5. Users notified of completion

## Media Application Challenges

Challenge	Description
Large File Sizes	Transfer and storage of high-resolution media is costly
Processing Power	Video encoding/transcoding requires significant resources
Latency	Critical for live streaming and real-time collaboration
Data Security & DRM	Protection from piracy and unauthorized access
Cost Management	Balancing performance needs with cloud service costs
Workflow Complexity	Integrating production stages into cohesive cloud workflow
Bandwidth	High requirements for both uploading and streaming content

## Multiplayer Online Gaming

### Characteristics

- Attracts millions of gamers worldwide in shared virtual environments
- Supports hundreds of players per session

### Architecture

- Players update game server hosting the session
- Server integrates updates into log available to all players via TCP port
- Client software connects to log port to update local UI
- Log processing tasks can be offloaded (e.g., to Aneka)

## Gaming Challenges

Challenge	Description
Latency	Critical for gameplay; requires edge computing, optimized routing
Scalability	Handling variable loads during peak demand and new launches
Server Optimization	Ensuring proper hardware configuration for smooth gameplay
Security	Protection against DDoS attacks, cheating, account hijacking
Bandwidth Costs	High data transfer between players and servers
Global Distribution	Managing servers across regions for global player base
Connectivity Issues	Server downtime severely impacts player experience