PCAC2011 CLOUD INFRASTRUCTURE & APPLICATIONS (3-0-0)

OVERALL COURSE OBJECTIVES: Students will synthesize and apply comprehensive knowledge on distributed and cloud computing concepts, cloud services, infrastructure, data analytics, storage services, machine learning, and real-time streaming technologies, leveraging major service providers to build, manage, and optimize scalable, effective cloud solutions using cutting-edge tools & techniques.

LEARNING OUTCOMES: On successful completion of the course the students shall be able to:

1. Understand and apply key distributed computing concepts for cloud computing including clouds, MapReduce, and NoSQL stores.

2. Utilize modern cloud technologies and the architecture behind the Cloud Systems and Infrastructure, and how virtualization and containers play a role.

3. Grasp how cloud service providers like Amazon, Google, and Microsoft organize their services and offerings.

4. Apply knowledge of cloud middleware technologies such as RPC and REST, JSON and load balancing for cloud-based applications.

5. Comprehend the operational aspects of cloud storage services such as Hive, HDFS, Ceph, cloud object storage systems, and more.

6. Implement data analytics applications on cloud, understand large scale data storage, consensus algorithms, real-time streaming technologies, and machine learning in cloud computing.

COURSE CONTENT:

Module 1: Cloud Computing Concepts: Part 2 [20 Hours]

Cloud computing systems today, whether open-source or used inside companies, are built using a common set of core techniques, algorithms, and design philosophies – all centered around distributed systems. Learn about such fundamental distributed computing "concepts" for cloud computing. Some of these concepts include: clouds, MapReduce, key-value/NoSQL stores, classical distributed algorithms, widely-used distributed algorithms, scalability, trending areas, and much, much more! Know how these systems work from the inside out. Get your hands dirty using these concepts with provided homework exercises. In the programming assignments, implement some of these concepts in template code (programs) provided in the C++ programming language. Prior experience with C++ is required. The course also features interviews with leading researchers and managers, from both industry and academia. This course builds on the material covered in the Cloud Computing Concepts, Part 1 course.

Sub-Topics
Classical Distributed Algorithms Continued
Concurrency and Replication Control
Emerging Paradigms
Classical Systems
Real-Life Behaviors
Formative   **Assessments:**
6 quizzes and 1 coding/lab assignment.

Module 2: [Cloud Computing Applications, Part 1: Cloud Systems and Infrastructure](#) [15 Hours]

Welcome to the Cloud Computing Applications course - an ensemble of informative components providing detailed insights on the world of Cloud Computing and Big Data. The course begins with an introduction to cloud computing, its economic foundations, and the concept of big data. It progresses to software-defined architectures, cloud virtualization, and cloud provider offerings. Further focus is given on virtualization and container technologies including Docker, JVM, and Kubernetes. The course also deep dives into cloud offerings like platform-as-a-service, mobile backend-as-a-service, and serverless architectures with a focus on cloud middleware technologies. The topic of Metal-as-a-Service (MaaS) is also covered. The final week focuses on higher-level cloud services, especially cloud storage services, with introductions to Hive, HDFS, and Ceph, cloud object storage systems, virtual hard drives, and archival storage options. The course wraps up with a discussion on Dropbox's cloud solution.

Sub-Topics
Introduction to Cloud Computing
Cloudonomics
Foundations: Containers, Virtual Machine, JVM
MAAS, PAAS, Web Services
Storage: Ceph, SWIFT, HDFS, NAAS, SAN, Zookeeper

Formative Assessments:
 4 quizzes.

Module 3: [Cloud Computing Applications, Part 2: Big Data and Applications in the Cloud](#) [20 Hours]

The Cloud Computing Applications course provides comprehensive knowledge on Cloud Computing and Big Data. Starting with an overview of cloud computing, its economic underpinnings, and the concept of big data, the course then delves into software-defined architectures, cloud virtualization, and offerings from cloud service providers. Emphasizing virtualization and containers, it includes lectures on Docker, JVM, and Kubernetes. The course explores higher-level cloud offerings like platform-as-a-service, mobile backend-as-a-service, and serverless architectures, along with cloud middleware technologies. The module on Metal-as-a-Service (MaaS) is also included. The final week concentrates on high-level cloud services - particularly cloud storage services introducing Hive, HDFS, Ceph, cloud object storage systems, virtual hard drives, and archival storage options, concluding with a discussion on Dropbox's cloud solution.

Sub-Topics
Spark, Hortonworks, HDFS, CAP
Introduction to Distros
Large Scale Data Storage
Streaming Systems
Big Data Machine Learning Introduction
Graph Processing and Machine Learning
Formative Assessments:
4 quizzes.
ASSESSMENT:

For summative assessments, Coursera will provide question banks for which exams can be conducted on the Coursera platform or the faculty will create their own assessments.

Note: If a Course or Specialization becomes unavailable prior to the end of the Term, Coursera may replace such Course or Specialization with a reasonable alternative Course or Specialization.