

RI3009: Cloud Computing – July 14 to Aug 1, 2017

(updated: 12 July 2017)

L00: Overview

- Learning Objectives
- What we cover?
- Course Schedule & Webpage
- Main Text
- Module Assessment

Books

1. *Cloud Computing: Concepts, Technology and Architecture*, Thomas Erl, Zaigham Mahmood and Ricardo Puttini, Prentice-Hall, 2013. [CTA]
2. *Cloud Computing: Theory and Practice*, Dan C. Marinescu, Elsevier, 2013. [TP]

PART A: Principles of Cloud Computing

L01: Introduction [CTA-chapter 3]

- What and Why
- Cost Model
- History
- Key Business Drivers
- Basic Concepts & Terminology
- Goals and Benefits
- Technical and Non-Technical Challenges
- Summary

References

1. [Above the Clouds: A Berkeley View of Cloud Computing](#), 2009.
2. [The NIST Definition of Cloud Computing](#), NIST Report, 2011.

L02: Concepts and Models [CTA – chapter 4]

- NIST Definition
- Cloud Characteristics
- Cloud Service (Delivery) Models
- Conceptual Reference Architecture
 - Actor Roles and Organizational Boundaries
 - Interactions between the Actors
 - Key Functions of Cloud Providers
- Cloud Deployment Models
- Summary

References

1. [NIST Cloud Computing Reference Architecture](#), NIST Report, 2011.

H01: IBM Cloud Services (PaaS, SaaS)

- Objective
- IBM Cloud Platform – Bluemix
 - Overview
 - Target Consumers
 - Technology
- Examples
 - PaaS: Setup a Simple Web Application using Boilerplates to query dashDB database
 - SaaS: Data Analytics using dashDB, SQL Query and R

- Summary

References

1. [NodeJS-DashDB example](#)
2. How-to's & Demos: A tour of Bluemix 2017 - <https://www.youtube.com/watch?v=gq5MvY1nNOM> (24 mins)
3. [IBM Bluemix The Cloud Platform for Creating and Delivering Applications](#), IBM Redbooks, 2015.
4. Bluemix users: <https://www.ibm.com/cloud-computing/bluemix/case-studies>
5. IBM dashDB: <http://www-01.ibm.com/support/knowledgecenter/SS6NHC/com.ibm.swg.im.dashdb.kc.doc/welcome.html>
6. SQL Query: <http://www.w3schools.com/sql/>
7. R Language: <http://cran.r-project.org/doc/manuals/r-release/R-intro.html>
8. Plotting: <http://docs.ggplot2.org/current/>

HO1: IBM Cloud Services (PaaS, SaaS) Lab

- Objective
- IBM Cloud Platform – Bluemix
 - Overview
 - Target Consumers
 - Technology
- Examples
 - PaaS: Setup a Simple Web Application using Boilerplates to query dashDB database
 - SaaS: Data Analytics using dashDB, SQL Query and R
- Summary

References

1. [NodeJS-DashDB example](#)
2. Demo: Getting Started with [Node.js](#) on Bluemix 2016- https://www.youtube.com/watch?v=sHhNoV-sS_I&list=PLJxa6IsF8C5qFbRinR2ZaEZ3AilFtdLqx (13 mins)
3. [IBM Bluemix The Cloud Platform for Creating and Delivering Applications](#), IBM Redbooks, 2015.
4. Bluemix users: <https://www.ibm.com/cloud-computing/bluemix/case-studies>
5. IBM dashDB: <http://www-01.ibm.com/support/knowledgecenter/SS6NHC/com.ibm.swg.im.dashdb.kc.doc/welcome.html>
6. SQL Query: <http://www.w3schools.com/sql/>
7. R Language: <http://cran.r-project.org/doc/manuals/r-release/R-intro.html>
8. Plotting: <http://docs.ggplot2.org/current/>

PART B: Technologies, Programming & Applications

L03: Technologies behind Cloud Computing [CTA – chapter 5 and Appendix D]

- Resource Hosting
- Main Components in a Datacenter
 - Server, storage and network
 - Cooling systems and energy
 - Fire protection
 - Security
- Datacenter Tiers
- Virtualization
- Multitenancy
- Summary

References

1. Chapters 3 & 4, [The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines](#), Second Edition, Synthesis Lectures on Computer Architecture, Luiz André Barroso, Jimmy Clidaras, Urs Hölzle, Morgan & Claypool Publishers, 2013.
2. [Efficiency: How we do it, Google Datacenters](#), July 2017.
3. [Data Centres Shine Amid Property Gloom](#), Straits Times, Jan 12, 2016.
4. [Building Data Centres at Sea An Idea Worth Floating](#), Straits Times, May 2017.

L03S: Datacenter at the School of Computing

- Objective
- SoC Datacenters: CR1 and CR2
- Unix and Windows Servers
- Others
 - Power
 - Cooling
 - Fire
 - Security
- Safety during Tour

L04: Cloud Architecture [CTA – chapter 11]

- Purpose
- How to organize (partition) resources?
 - Workload Distribution
 - Resource Pooling
 - Dynamic Scalability
 - Elastic Resource Capacity
 - Service Load Balancing
- How to operate/manage resources to meet certain objectives?
 - Cloud Bursting
 - Elastic Disk Provisioning
- Summary

L05: Applications and Paradigms ([TP chapter 4](#))

- Cloud Applications
 - Types of Applications
 - Elasticity and Workload
 - Challenges in Developing Applications
 - Architectural Styles
 - Coordination of Multiple Activities – Workflow
- Application Development Models
 - IaaS, PaaS and SaaS
- MapReduce Programming Model
 - Example: Word Count, GrepTheWeb
- Summary

References

1. [Cloud Architectures](#), Jinesh Varia, Amazon, June 2008 (describes Amazon GrepTheWeb production system).

H02: Amazon Web Services (IaaS, PaaS, SaaS)

- Objectives
- Main EC2 Steps

- Sign-in Page Credentials
- Launch Instance
- Connect to Instance
- Run Program on Instance Created
- Close Instance Connection
- Stop Instance
- Examples
 - Word Count on EC2
 - Word Count on Elastic MapReduce
- Summary

References

- AWS EC2 - <https://www.youtube.com/watch?v=TsRBftzZsQo> (4 min)
- AWS S3 - https://www.youtube.com/watch?v=YyraqI9A_Rc (7 min)
- AWS EMR - <https://www.youtube.com/watch?v=Hhj3fOdt7zo> (10 min)
- <https://aws.amazon.com/getting-started/tutorials/launch-a-wordpress-website/>
- <http://searchaws.techtarget.com/definition/Amazon-Elastic-MapReduce-Amazon-EMR>
- http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-launch-instance_linux.html
- <http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-connect-to-instance-linux.html>
- <http://wiki.apache.org/hadoop/AmazonEC2?action=recall&rev=10>
- <http://wiki.apache.org/hadoop/AmazonEC2>
- <https://portal.futuresystems.org/manual/hadoop-wordcount>

Accessing AWS Cloud Services

H02: Amazon Web Services (IaaS, PaaS, SaaS) Lab

- Objective
- Main EC2 steps
 - Sign-in page credentials
 - Launch Wordpress instance
 - Connect to instance (admin)
 - Connect to server using SSH (with key)
 - Close connection
 - Stop instance
- Main steps for Amazon EMR
 - Configure S3 storage
 - Upload application
 - Setup EMR cluster
 - Execute the application
 - Download/view results
 - Terminate cluster and S3 storage
- Summary

References

- <https://aws.amazon.com/getting-started/tutorials/launch-a-wordpress-website/>
- http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-launch-instance_linux.html
- <http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-connect-to-instance-linux.html>
- <http://wiki.apache.org/hadoop/AmazonEC2?action=recall&rev=10>
- <http://wiki.apache.org/hadoop/AmazonEC2>
- <https://portal.futuresystems.org/manual/hadoop-wordcount>

L06: Cloud Infrastructure [TP - chapter 3]

- Cloud Platforms

- Amazon Web Services
 - Regions & Availability Zones
 - Instances
 - Examples
- Google Cloud Platform
- Microsoft Windows Azure
- Open-source Platforms
- Cloud Interoperability and Vendor Lock-in
- Energy use of Data Centers
- Energy-proportional Systems
- Summary

References

1. [Architecting for the Cloud: Best Practices](#), Amazon, 2011.
2. [Overview of Amazon Web Services](#), January 2014.

L07: K-Means Clustering using Elastic MapReduce (IaaS, PaaS)

- Objective
- Algorithm
- K-means in MapReduce
- K-means using Amazon Elastic MapReduce
- Summary

References

- <http://cmj4.web.rice.edu/MapRedKMeans.html>
- <http://www.kau.edu.sa/GetFile.aspx?id=187901&Lng=ar&fn=k-mean-clustering.ppt>
- http://en.wikipedia.org/wiki/K-means_clustering
- <http://aws.amazon.com/elasticmapreduce/>

L08: Building a Video-Sharing SaaS Cloud Application

- Objective
 - Design of Application
 - Upload Process
 - Encoding Process
 - Streaming Process
 - Performance and Scaling
 - Pricing
 - Summary
1. [How AWS Pricing Works](#), July 2014.

PART C: Cloud Management

L09: Cost Metrics, Pricing Models, Service Metrics and TCO [CTA chapters 15 & 16]

- Cost Metrics
 - Business
 - Cloud Usage
 - Cost Management
- Pricing Models
- Service Metrics
 - Service Quality
 - Service Availability
 - Service Reliability
 - Service Resiliency

- Service Level Agreement
- Total Cost of Ownership
- Summary

References

1. [How AWS Pricing Works](#), Amazon, 2012.
2. The Total Cost of (Non) Ownership of Web Applications in the Cloud, Jinesh Varia, August 2012.
3. Chapter 6: Modeling Costs in “The Datacenter as a Computer”, 2013.

L10: Cloud-enabled Data Analytics

- Cloud Analytics
- Analytics Workflow for Big Data
- Four Key Issues
 - Data Management
 - Model Building and Scoring
 - Visualization and User Interaction
 - Business Models
- Summary

References

1. [Big Data Computing and Clouds: Trends and Future Directions](#), Journal of Parallel & Distributed Computing, 79-80 (2015), 3-15.

L11: Summary and Open Issues

- Revisit of Learning Objective
- Topics Covered
- Open Issues
 - 10 Obstacles and Opportunities
 - Open Issues: 5 Areas
- Our Research

References

1. [Cloud Computing Synopsis and Recommendations](#) (Open Issues), NIST Report, May 2012.
2. [Challenges and Opportunities with Big Data](#), Computing Community Consortium, Feb 2012.

Books

1. *Cloud Computing: Concepts, Technology and Architecture*, Thomas Erl, Zaigham Mahmood and Ricardo Puttini, Prentice-Hall, 2013. [CTA]
2. *Cloud Computing: Theory and Practice*, Dan C. Marinescu, Elsevier, 2013. [TP]