



**INSTRUCTION DIVISION**  
**1<sup>st</sup> SEMESTER 2017-2018**  
**Course Handout Part II**

In addition to part-I (General Handout for all courses appended to the time table) this portion gives further specific details regarding the course

**Course No. :** CS / SS G527  
**Course Title :** Cloud Computing  
**Instructors:** K HARIBABU (email:khari@pilani)

**1. Scope and Objective:**

The primary objective of the course is to introduce the student to cloud computing from architectural and design perspectives. As such the emphasis of the course would be on the underlying infrastructure and architecture of clouds, techniques for enabling services and the quality of such services, as well as issues in designing clouds. Specific research issues in performance, security, and management would also be addressed. Programming on the cloud would be encouraged but not taught in class. Students are expected to learn and understand tools and techniques for using, designing, and implementing clouds and services via assignments and a term project.

**2. Text and References:**

a.Text Book: NONE

b.References:

**R1.** Kai Hwang , Jack Dongarra , Geoffrey C. Fox *Distributed and Cloud Computing: From Parallel Processing to the Internet of Things*. Morgan Kauffman 2011.

**R2.** Jim Smith, Ravi Nair. *Virtual Machines: Versatile Platforms for Systems and Processes*. Morgan Kaufmann. 2005

**AR.** Additional references (papers) to be posted online.





### 3. Course Plan:

#### 3a. Modules

Module	Theme	Learning Objectives
I	Cloud Computing – Introduction, Principles, and Issues.	<ul style="list-style-type: none"><li>To understand the motivation for Cloud Computing.</li><li>To understand the underlying (distributed) computing model.</li></ul>
II	Cloud Architecture – Resources and Virtualization	<ul style="list-style-type: none"><li>To understand how to leverage and provision computing resources at different levels of abstraction.</li><li>To understand virtualization techniques at different levels of abstraction.</li><li>To understand how to architect a cloud to suit different requirements</li></ul>
III	Programming for the Cloud and Application Models	<ul style="list-style-type: none"><li>To understand the execution of applications on the cloud</li><li>To understand how to develop &amp; deploy applications for the cloud and the relevant tools &amp; technologies</li></ul>
IV	Services, Service Models, and QoS.	<ul style="list-style-type: none"><li>To understand how to use the cloud to deliver Software as a Service.</li><li>To understand how to deliver computing Infrastructure (e.g. processors, storage, network) as a Services</li><li>To understand Quality of Service issues and QoS support mechanisms for Services on the Cloud</li></ul>
V	Cloud Management, Performance and Security Issues	<ul style="list-style-type: none"><li>To understand how to manage a cloud platform and a services ecosystem</li><li>To understand performance issues and techniques to enable performance of a cloud at different levels of abstraction</li><li>To understand security issues specific to cloud computing and solutions to address them.</li></ul>

[Note: Module III will not be covered in detail in class. It is to be learnt primarily through assignments and project. End of Note.]





## 3.b Lecture Schedule:

Lecture	Module	Topics	Readings
1	I	Evolution of Computing Systems – Mainframes to PCs to Networked Systems to Clouds. Cloud Computing and Services – User’s Perspective.	R1 Sec. 1.2 and 1.4.1
2		Spectrum of Parallel and Distributed Computers – Programmer’s Perspective. Forms of Parallelism: Data Parallelism vs. Task Parallelism vs. Request Parallelism.	R1 Sec. 1.3.4
3		Distributed Computing Models – Message Passing: Synchrony, Failures, and Buffering.	NONE
4		Parallel and Distributed Computing: Performance: Speedup and Amdahl’s Law. Performance vs. Fault Tolerance: Reliability and Availability. Energy Efficiency	R1 Sec. 1.5
5		Failures: Byzantine Failures: Byzantine Agreement: Upper Bound on faulty entities, Solutions and capabilities.	AR
6		Failures and Consensus: Paxos	AR
7 - 8		Introduction to Computer Clusters: Structure and Components of a cluster – nodes, interconnect, and middleware. Design of a cluster: Single System Image, Design for availability, Fault-tolerance and Fault Recovery.	R1 Sec. 2.1 to 2.3, AR
9		Computer Clusters: Task Management and Resource Management; Scheduling and Load Balancing.	R1 Sec. 2.4, AR
10		High Availability Clusters: Design Issues. Check-pointing and Recovery.	R1. Sec. 2.3.2 to 2.3.4
11-12		III	Distributed Programming: RPC/RMI, Message Passing: Blocking vs. Non-Blocking, Buffered vs. Non-Buffered Communication.
13	Programming on Clusters: MPI, Map-Reduce/Hadoop, Apache Spark		R1 Sec. 1.4.3 and 6.2.2, AR
14	I	Distributed Mutual Exclusion.	AR
15-16	II	Virtualization: Introduction, Different Levels of Abstraction, and Mechanisms for Virtualization. Process vs. System VMs, HLL VMs.	R2 Ch. 1, AR
17-18		Instruction Set Virtualization: Interpretation and Binary Translation.	R2 Ch. 2, AR





# Birla Institute of Technology & Science, Pilani

## Pilani Campus

		Code Discovery and Dynamic Translation. Optimizations. Case Study.	
19		Process VMs: Structure and Behavior.	R2 Sec. 3.1, AR
20-21		Implementation of Process VMs: Emulation of: Memory Architecture, Instruction Execution, and Exception Handling; OS Emulation; Code Cache Management. Case Study.	R2 Ch. 3, AR
22		System VMs : Applications and Implementation Models.	R2 Sec. 8.1
23		System VMs: Resource Virtualization and ISA Virtualizability	R2 Sec. 8.2, AR
24		System VMs: Virtualization of Memory and I/O	R2 Sec. 8.3-8.4
25		Virtualization of Multi-processor systems / Multi-core Processors.	R2 Sec. 9.1-9.3
26-27		Storage Systems and Storage Virtualization: Storage Devices, File Systems and Volumes, Storage Networks – NAS and SAN, NAS Internals, Virtual Storage and Storage on the Cloud.	AR
28-29		Virtualized Networks and Virtual Clusters. Process Migration and VM Migration. Live Migration – Suspend and Resume.	R1 Sec. 3.4, R2 Sec. 10.2, AR
30		Virtualization in Data Centers and Clouds. Cloud OS.	R1 Sec. 3.5, AR
31-32	III	Handling Large Data: Big Data Issues; Programming and Storage Models: GFS/HDFS, NoSQL, Google BigData etc.	AR
33	IV	Services on the Cloud. Software as a Service; SOA; REST; Web Services;	R1 Sec. 5.1, AR
34-35		Platform as a Service; Computing Infrastructure as a Service. Case Studies (EC2 and Azure).	AR
36		Quality of Service – Models, Parameters, and Metrics. SLAs.	AR
37-38	V	Resource Scaling and Capacity Management: Managing VMs Resource Provisioning.	AR
39		Data Centers and Cloud: Performance Measurement and Models. Application Performance on the Cloud – Performance Modeling and Enhancement Techniques	AR
40-41		Cloud-specific Security Issues and Challenges. Access Control and Privacy Issues. Process Isolation and Security	AR
42		Cloud Computing and Energy Consumption.	AR



### 3.c. Practicum

The structured laboratory hours will be used for demonstrations and assessment. Practical work by students is expected to be carried out at the students' own convenience via assignments and term project.

The scope of the assignments would include (i) ***distributed programming at the cluster level***, (ii) ***implementation of virtualization modules***, and (iii) ***implementation of services on the cloud***. The assignments are primarily meant for students to learn implementation skills with some design.

Assignments will require students to work in a team, set up developmental / experimental platform(s) and conduct performance studies on their implementations.

## 4. Evaluation

### 4. a. Evaluation Scheme:

Component	Weight	Date	Remarks
Mid-Term Test	25%	14/10 11:00 - 12:30 PM	Open Book
Comprehensive Exam	35%	13/12 AN	Open Book
Class Participation	5%		In-class
Assignments	20%	Small problems. 2 weeks long.	Programming, Take Home
Term Project	15%	Semester long project	Design and implementation

**Note:** Assignments/Term Project are to be done by teams of at most three students. **End of Note.**



#### 4. b. Make-up Policy:

- **Assignments / Term Project:**
  - No Make-up will be available for assignments or term project. Late submissions will be evaluated at 25% less weight for that component for a delay of up to 24% hours after which no submissions will be accepted.
- **Test:**
  - Prior Permission of the Instructor-in-Charge is usually required to get a make-up for a test.
  - A make-up test shall be granted only in genuine cases where - in the Instructor's judgment - the student would be physically unable to appear for the test. Instructor's decision in this matter would be final. .
- **Comprehensive Exam:**
  - Make-up for the comprehensive exam may be applied only with the Associate Dean of Instruction.

#### 4.c. Fairness Policy:

- Students are expected to work within their team on assignments / project expect where explicitly instructed / permitted otherwise.
- When students are allowed to consult/discuss with other students/teams such consultation/discussion should be explicitly acknowledged and reported to the instructor prior to evaluation.
- When students are expected to collaborate within a team:
  - Individual contributions should be identified and documented in qualitative and quantitative terms by each team member.
  - Instructor may assess and mark each individual in a team separately.
  - The instructor's assessment of the contributions in this matter would be final.

#### 5. Consultation Hours: Tue 4-5PM (6121-M)

**6. Notices:** All notices concerning this course will be displayed online only. If there is a need email would be used on short notice (12 hours) – only BITS Pilani email would be used.

**Instructor –In- Charge  
CS / SS G527**

