



(An Autonomous Institution - AFFILIATED TO ANNA UNIVERSITY, CHENNAI)

S.P.G.Chidambara Nadar - C.Nagammal Campus

S.P.G.C. Nagar, K.Vellakulam – 625 701 (Near VIRUDHUNAGAR).

DEPARTMENT OF INFORMATION TECHNOLOGY

LABORATORY MANUAL CUM RECORD

(AS PER OUTCOME BASED EDUCATION - OBE)

Course Code

VIT312

Course Name

Cloud Computing





(An Autonomous Institution) - AFFILIATED TO ANNA UNIVERSITY, CHENNAI)

S.P.G.C. Chidambara Nadar - C.Nagammal Campus
S.P.G.C. Nagar, K.Vellakulam – 625 701 (Near VIRUDHUNAGAR).

DEPARTMENT : _____

NAME : _____

REGISTER NUMBER : _____

ROLL NO (SECTION) : _____

BRANCH : _____

YEAR : _____

SEMESTER : _____

Bona-fide Record of work done in the VIT312 Cloud Computing of KAMARAJ College of Engineering and Technology, Near Virudhunagar, during the Academic Year 2024 - 2025.

COURSE INCHARGE

HEAD OF THE DEPARTMENT

Submitted for the Practical Examination held on _____ at
KAMARAJ College of Engineering and Technology, Near Virudhunagar.

INTERNAL EXAMINER

EXTERNAL EXAMINER

CONTENT

Section: 1 Course Introductions

- i. Introduction of Course
- ii. Syllabus –List of Experiment - Reference books

Section: 2 Outcome Based Education (OBE)

- i. Vision, Mission of Institution, Department, PEO's
- ii. OBE- Outcome Based Education - Introduction
- iii. Program Outcome - PO, PSO
- iv. Course Outcome - Course Objectives, Course Outcome (CO)
Mapping of CO vs PO - CO-PO Matrix (Calculation of
Expected POs value)
- v. Demonstration of CO Attainment (Internal Assessment),
Action Plan
- vi. Rubrics for Experiments

Section: 3 Safety Instructions

- i. Principles of Safety
- ii. Laboratory specific safety instruction

Section: 4 Laboratory Experiments Skill Development

- i. Experiments
*(Planning - Aim; Description; Procedure; Program; Viva;
Assessment - Rubrics- Mark allotted and awarded;
Result; Observations - by student)*

INDEX

Ex.No	Date	Title of the Exercise	Page No	Mark (100)	CO Attained? (Yes/No)	Sign
1		Installing Guest OS on Virtual Machine				
2		Installation of C-Compiler on Guest OS				
3		Create and Deploy Java/Python Web Application in AWS				
4		Simulation of VM Scheduling Using CloudSim				
5		Installation of Single Node Hadoop Cluster				
6		Word Count Program using MapReduce				
7		Create Virtual Machine in OpenStack Cloud Platform				
8		Mini Project				

Laboratory Specific Instructions to Students

- 1) Students are required to remove their footwear outside the center.
- 2) Students are asked to register their Name, Roll No.& System No. in the Log register before logging into the system.
- 3) Students should leave their belongings outside the lab except their Lab Manuals and Lab File.
- 4) Students are requested not to place their legs on the wall or on the table.
- 5) Students should refrain from leaning on the table and sitting on it.
- 6) Before logging in to a particular terminal, if there is something wrong in the terminal, the student should report the same immediately to the concerned staff.
- 7) Students should not use any disks brought from outside without prior permission from the concerned staff.
- 8) Students should not move from their terminal and ask any clarification from their friends/neighbour without any proper permission from the concerned staff.
- 9) Students should collect their printouts before leaving the lab for that particular session.
- 10) Before leaving the Terminal, the students should logout properly and leave their chairs in position.
- 11) Students are not allowed to use Pen drive in any of the systems in the lab.
- 12) Students are not allowed to take any manual outside the center.
- 13) Edibles are strictly prohibited in the center.
- 14) Mobile Phones usage is strictly prohibited inside the Lab. In case of any violation, the mobile phone will be seized or broken.
- 15) No internet browsing is allowed during the lab hours.

Section: 1 Course Introductions

Introduction of Lab Course

Course Code (As per Curriculum)	:	VIT312
Course Code (As per NBA)	:	21ITC306
Course Name	:	Cloud Computing

Importance of Course (150 Words write up)

- 1) Scalability:** Cloud computing allows you to use as many or as few resources as you need. Therefore, depending on your business needs or projected traffic to your business you can choose to increase or decrease your investment in IT infrastructure
- 2) Saving Costs:** Cloud computing helps businesses to reduce costs in various ways. Companies only pay for the resources they use, making this process a more economical option than having to buy and manage their own resources. Cloud computing also results in considerable savings in Capital Expenditure and Operating Expenditure because companies do not have to invest in expensive hardware, storage devices, software, etc.
- 3) Disaster Recovery:** With all data stored in the ‘cloud’ backup and recovery of data and applications is quicker and more reliable. This applies to all sizes of organizations and volumes of data. 20% of cloud users claim disaster recovery in four hours or less as opposed to only 9% of non-cloud users.
- 4) Security:** It is the duty and responsibility of the cloud service providers to carefully monitor security. Compare this against an in-house I.T. department, for example, which is tasked with so many internal processes to manage; security is just one of the many items on the list.
- 5) Flexibility:** When using cloud computing the number of options is vast. Depending on the size of the organization, business needs, workloads, etc. companies can choose cloud infrastructure and services accordingly.

Lab Course syllabus – List of Experiments

Course Code (As per Curriculum)	:	VIT312
Course Code (As per NBA)	:	21ITC306
Course Name	:	Cloud Computing

Aim : • This course enables the students to understand the basics of cloud computing. This course helps the students to learn the evolution of cloud from the existing technologies. This course familiarizes the lead players in cloud. This course makes the students to have knowledge on the various issues in cloud computing. This course disseminates various cloud platforms.

Objectives :

- 1) Installing Guest OS on Virtual Machine
- 2) Installation of C-Compiler on Guest OS
- 3) Create, deploy, and manage Web Application in AWS
- 4) Simulation of VM Scheduling Using CloudSim
- 5) Installation of Single Node Hadoop Cluster
- 6) Word Count Program using MapReduce
- 7) Create Virtual Machine in OpenStack Cloud Platform

List of Experiment :

Reference Book : Kai Hwang, Geoffrey C. Fox, Jack G. Dongarra, "Distributed and Cloud Computing, From Parallel Processing to the Internet of Things", Morgan Kaufmann Publishers, 2012.

Section: 2 Outcome Based Education (OBE)

VISION & MISSION & QUALITY POLICY OF THE COLLEGE

- Vision** : To Make this Institution the unique of its kind in the field of **Research and Development activities** in this part of world
- Mission** : To impart highly innovative and technical knowledge to the urban and unreachable rural student folks through "**Total Quality Education**".
- Quality Policy** : Committed to impart Quality Technical Education imbued with **proficiency, human values and continual improvement**.

VISION & MISSION OF THE DEPARTMENT

- Vision** : To make the department of information technology the unique of its kind in the field of research and development activities in this part of world
- Mission** : To impart highly innovative and technical knowledge in the field of information technology to the urban and unreachable rural student folks through total quality education

PROGRAM EDUCATIONAL OBJECTIVES (PEO's)

- PEO -1** : Graduates of the programme will exhibit expertise in technical knowledge by applying distinctive skills in various fields of Information Technology
- PEO -2** : Graduates will become pioneers in the field of Information Technology by working collaboratively and providing solutions to meet societal needs through persistent learning
- PEO -3** : Graduates will be able to adopt innovative practices and contribute towards research and technological development in the field of Information Technology through Total Quality Education

OBE- Outcome Based Education - Introduction

Aim:

- i. Outcomes are more **like signboards and roadmaps to help the learners** reach where **they're supposed to reach** and **contribute to progress**.
- ii. **Outcome measurements** provide the basis for **continuous improvement in the quality of learning**

Terminology

Accreditation

- | | Description |
|------|---|
| i. | With the introduction of the washington accords , our higher education demands better, intelligent workflows for the sake of improving quality |
| ii. | The accreditation and regulatory organisations including NBA, NAAC, AICTE etc. has clearly made their intentions clear about introducing a skill-based, competency nurturing learning experience |
| iii. | As of the NBA, their accreditation methods and assessment parameters are based on OBE |

OBE- Outcome Based Education

- | | |
|------|---|
| i. | Deciding outcomes for academic achievements and it's attainment for assessment and formulation is based on a learning theory called Outcome Based Education(OBE) |
| ii. | OBE is an educational theory that bases each part of an educational system around goals (outcomes) . |
| iii. | Student role: By the end of the educational experience , each student should have achieved the goal . There is no single specified style of teaching or assessment in OBE; instead, classes, opportunities, and assessments should all help students achieve the specified outcomes . |
| iv. | Faculty role: The role of the faculty adapts into instructor, trainer, facilitator, and/or mentor based on the outcomes targeted . |

OBE emphasizes on

- | | |
|------|---|
| i. | Stating what you want your students to be able to do at the end of the program, |
| ii. | Assessing the students whether they are able to do what they are expected to do to do what they are expected to do, |
| iii. | Orienting teaching and other academic processes to facilitate students to do what they are expected to do.

Note: Outcomes are different from your course objectives. |

Course Objectives

- | | |
|-----|---|
| i. | An objective is more like a plan, or a road map to reach/attain the outcomes. |
| ii. | A course objective describes what a faculty member will cover in a course. They are generally less broader than desirable goals and more broader than student learning outcomes . |

Course Outcomes (COs)

- | | |
|------|--|
| i. | Outcome — A detailed description of what a student must be able to do at the conclusion of a course. |
| ii. | POs are attained through program specific Core Courses, which has their own, previously set outcomes to attain. These course-specific outcomes are called Course Outcomes. |
| iii. | No.of CO statement: Each course is designed to meet (about 5–6) Course Outcomes. |

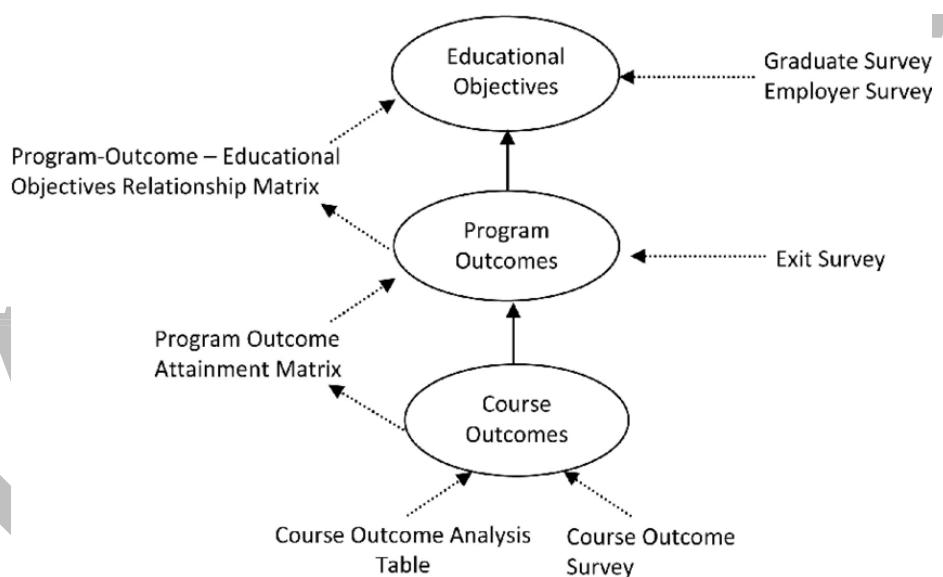
- Program Outcomes (POs)**
- iv. **Structure of COs Statements:** The Course Outcomes are stated in such a way that they can be actually measured. (**SMART** –“S- Specific; M-Measurable; A-Achievable; R-Realistic; T-Time bound”)
 - v. **Framing of COs:** COs are set by the institution, by consulting with the department heads, faculty, students and other stakeholders.
- Mapping of CO vs PO**
- i. POs **are statements** about the knowledge, skills and attitudes (attributes) the graduate of a formal engineering program should have.
 - ii. POs **deal with** the general aspect of graduation for a particular program, and the **competencies and expertise a graduate will possess after completion of the program.**
 - iii. These are broad and covers a wider area than of COs. the NBA has set **12 Program Outcomes**, or Graduate Attributes **for the sake of unity and quality assurance.**
 - i. Map the objective to the **outcomes** to analyse and **document their attainment**
 - ii. Each CO can be identified to **address a subset of POs**
 - iii. Based on the number of COs and sessions dedicated to them, it is possible to identify the **strength of mapping (1,2 or 3) to POs**
 - iv. Based on these strength of selected POs a CPO matrix can be established. **PO matrix can be established**
- | Course | PO1 | PO2 | PO3 | | | PO11 | PO12 |
|--------|-----|-----|-----|---|---|------|------|
| CO1 | 1 | 0 | 3 | 1 | 2 | 3 | 2 |
- Program Educational Objectives (PEOs)**
- i. Program Educational Objectives (PEO) **are statements** that describe the career and professional accomplishments that the **program is preparing the graduates to achieve.**
 - ii. **Measuring PEOs:** PEO's are measured **4–5 years after graduation.**
 - iii. They are set in order to measure the effectiveness of the program, and to check **whether it has prepared the students to deal with the real world**, where they could apply and use the skills and knowledge they've learned to good use.
 - iv. **No.of PEOs:** Each program shall specify 2–4 program specific outcomes for the accreditation by the NBA.
- CO Assessment:**
- i. The assessments should be in alignment with the COs
 - ii. Question paper should be so set to assess all COs
 - iii. The average marks obtained in assessments against items for each CO will indicate the CO attainment
- CO Attainment gaps:**
- iv. Instructors can set targets for each CO of his/her course
 - v. Attainment gaps can therefore be identified
 - vi. Instructor can plan to reduce the attainment gaps or enhance attainment targets enhance attainment targets
- CO Attainment:**
- vii. If the assessment is in alignment with COs, the performance of the students **indicates the CO attainment**

	Program Outcome (PO) <i>(GraduateAttributes)</i>		
PO/ PSO	Engineering Graduate will be able to:	LearningLevel	At end of the laboratory course, attained the Outcomes/Skill of the student - (Remark by students)
PO 1	Engineering knowledge: Apply the knowledge of mathematics, science, engineering fundamentals and an engineering specialization for the solution of complex engineering problems.	K3 - Apply	
PO 2	Problem analysis: Identify, formulate, research literature and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences and engineering sciences.	K4 - Analyze	
PO 3	Design/development of solutions: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for public health and safety, and cultural, societal, and environmental considerations.	K5 - Evaluate	
PO 4	Conduct investigations of complex problems: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.	K5 - Evaluate	
PO 5	Modern tool usage: Create, select and apply appropriate techniques, resources and modern engineering and IT tools, including prediction and modeling to complex engineering activities, with an understanding of the limitations.	K6 - Create	
PO 6	The engineer and society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent	K3 - Apply	

	Program Outcome (PO) <i>(GraduateAttributes)</i>		
PO/ PSO	Engineering Graduate will be able to: responsibilities relevant to the professional engineering practice.	Learning Level	At end of the laboratory course, attained the Outcomes/Skill of the student - (Remark by students)
PO 7	Environment and sustainability: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.	K2 Understand	
PO 8	Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.	K3 - Apply	
PO 9	Individual and team work: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.	A3 - Value	
PO 10	Communication: Communicate effectively on complex engineering activities with the engineering community and with the society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.	A3 - Value	
PO 11	Project management and finance: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.	K3 - Apply	
PO 12	Life-long learning: Recognize the need for and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.	A5 Characterize	

	Program Outcome (PO) <i>(GraduateAttributes)</i>		
PO/ PSO	Engineering Graduate will be able to: At the end of the program, the student	Learning Level	At end of the laboratory course, attained the Outcomes/Skill of the student - (Remark by students)
	Program Specific Outcome (PSOs) At the end of the program, the student		
PSO 1	Demonstrate technical and interpersonal skills to design and develop IT enabled solutions to meet the real time industrial and societal needs	K3 - Apply	
PSO 2	Exhibit an ability to adapt to the evolutionary changes in computing	K3 - Apply	

Flow Chart -Attainment of Program Educational Objectives (PEO's) through PO & CO



Course Outcome (COs)

Course Code (As per Curriculum)	:	VIT312
Course Code (As per NBA)	:	21ITC306
Course Name	:	Cloud Computing

Course Objectives

- This course enables the students to understand the basics of cloud computing. This course helps the students to learn the evolution of cloud from the existing technologies. This course familiarizes the lead players in cloud. This course makes the students to have knowledge on the various issues in cloud computing. This course disseminates various cloud platforms.

Course Outcomes

On successful completion of this course, students will be able to

CO	CO-Statements	Learning Level
CO1	Create and manage a virtual machine instance in a virtualized environment	K3 - Apply
CO2	Deploy a web application in a PaaS environment	K3 - Apply
CO3	Apply a resource scheduling algorithm in cloud environment	K3 - Apply
CO4	Construct a Hadoop framework to store and process large datasets	K3 - Apply
CO5	Use MapReduce program to perform parallel processing	K3 - Apply

Laboratory Assessment Tool

S.No	Assessment Tool	Weightage (%)
1	Laboratory Experiments - Rubrics	75%
2	Model Exam	25%

Target Value = **80 %** of Marks

CO –PO Mapping															
	Program outcomes													Program Specific outcomes	
CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	
CO1	L	L	L	L	L	-	-	-	H	-	-	L	M	L	
CO2	M	M	M	M	L	-	-	-	H	-	-	L	M	L	
CO3	M	M	M	M	L	-	-	-	H	-	-	L	M	L	
CO4	M	M	M	M	L	-	-	-	H	-	-	L	M	L	
CO5	M	M	M	M	L	-	-	-	H	-	-	L	M	L	

CO –PO Matrix															
	Program outcomes													Program Specific outcomes	
CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	
CO1	1	1	1	1	1	-	-	-	3	-	-	1	2	1	
CO2	2	2	2	2	1	-	-	-	3	-	-	1	2	1	
CO3	2	2	2	2	1	-	-	-	3	-	-	1	2	1	
CO4	2	2	2	2	1	-	-	-	3	-	-	1	2	1	
CO5	2	2	2	2	1	-	-	-	3	-	-	1	2	1	
PO Expected (Round off Average)	1.8	1.8	1.8	1.8	1	-	-	-	3	-	-	1	2	1	

Note: Target Value = 80 %

If CO Attainment,

YES = Laboratory Mark (%) >= Target Value (%); **NO**= Laboratory Mark (%) < Target Value (%)

LIST OF EXPERIMENT & ITS CO

S.No	EXPERIMENT NAME	Course Outcome	Blooms Taxonomy level	Marks Obtained Out of 100	CO Attained? (Y/N)
1	Installing Guest OS on Virtual Machine	CO1	K3		
2	Installation of C-Compiler on Guest OS	CO1	K3		
3	Create and Deploy Java/Python Web Application in AWS	CO2	K3		
4	Simulation of VM Scheduling Using CloudSim	CO3	K3		
5	Installation of Single Node Hadoop Cluster	CO4	K3		
6	Word Count Program using MapReduce	CO5	K3		
7	Create Virtual Machine in OpenStack Cloud Platform	CO1	K3		
8	Mini Project	CO5	K3		

CO ATTAINMENT - Internal Assessment EVALUATION SHEET				
Course Outcome	EXPERIMENT NAME	Marks Obtained Out of 100	Average of each CO	CO Attained? (Y/N)
CO1	1. Installing Guest OS on Virtual Machine			
	2. Installation of C-Compiler on Guest OS			
	3. Create Virtual Machine in OpenStack Cloud Platform			
CO2	1. Create, deploy, and manage Web Application in AWS			
CO3	1. Simulation of VM Scheduling Using CloudSim			
CO4	1. Installation of Single Node Hadoop Cluster			
CO5	1. Word Count Program using MapReduce			
	2. Mini Project			
Overall Total				
Average Mark (Out of 100) – (A)				

Co Attainment (Internal Assessment) Evaluation				
Experiment		Model Exam		CO Attained
A	B	C	D	
Mark Obtained (100)	Assessment Weightage (75% of 100 Marks Obtained)	Mark Obtained (100)	Assessment Weightage 25% of Mark Obtained	$E=((B+D) *0.25)$

Measuring Tools for CO attainment (Components & Weight age)			
Internal Experiment - 75% Model Exam - 25%	25%	External University Exam	25%

Action Plan:
(If CO Not Attained)

CO	Reason for not Attaining CO	Remedial Action Plan	Remarks	Signature
'				

Rubrics of Laboratory Experiments

Course Code : VIT312
(As per Curriculum)
Course Code : 21ITC306
(As per NBA)
Course Name : Cloud Computing

Rubrics

Category	Rubrics
Procedural Steps	96% - 100% of Procedural Steps Execution [41 - 50]
	80% - 95% of Procedural Steps Execution [31 - 40]
	41% - 79% of Procedural Steps Execution [21 - 30]
	20% - 40% of Procedural Steps Execution [11 - 20]
	Below 20% of Procedural Steps Execution [1 - 10]
Viva	91% to 100% answers are correct [21 - 25]
	76% to 90% answers are correct [16 - 20]
	51% to 75% answers are correct [11 - 15]
	26% to 50% answers are correct [6 - 10]
	Attempt all the questions with few mistakes [1 - 5]
Record	Submitted on the next lab session [21 - 25]
	2 days after the next lab [16 - 20]
	4 days after the next lab [11 - 15]
	5 days after the next lab [6 - 10]
	7 days after the next lab [1 - 5]

Section:3 Safety Instructions

The Institute works with several hazardous materials and equipment. **Safety** is an important part of any training in research. Potential job givers, be it industry or academia, **expect a certain awareness about safety.**

With guidance of AICTE - **Office of Laboratory Safety and Environmental Health (OLSEH)**, Our College has embarked on a concerted effort to improve the safety culture in our academics and Research lab.

1. Safety Principles

Follow Rules

OLSEH suggested that the **four essential principles of safety** must be follow for academics labs of institutions:
To prevent confusion, this manual clearly defines standards for safe work practices. These **rules need to be followed** by everyone.
"Remember practicing safety means doing things the right way, not the quick way"

Personal Responsibility

The primary responsibility for safety rests with the individual. A responsible, considerate worker with an understanding of the working of the laboratory, its equipment, basic chemistry, common sense,
"safety is your responsibility"

Trust structures more than people

No matter how careful they are, people often make mistakes. An effective safety policy does not rely on people but relies on systems to reduce the probability of accidents.
"I will be careful with chemicals" is a not an **"idiot-proof" safety** precaution, chemical-resistant gloves are.

Respond to Emergencies

Everyone must be prepared to respond quickly and effectively in an emergency. Become familiar with the work area, available exits, and safety equipment.
"training could save a life during an emergency"

Safety Related Documents

here are links for some important Safety related documents shared by IISc: Safety Manual:
https://olseh.iisc.ac.in/wpcontent/uploads/2019/03/IIScSafetyManual_Ver1_01.pdf

Safety - Photo Gallery

Fire Extinguisher

To operate an extinguisher:

Pull
Aim
Squeeze
Sweep

Know your extinguisher
Use the correct extinguisher

(Check your own extinguisher's label
for detailed instructions.)



Section: 4 Laboratory Experiments Skill Development

Ex. No: 1

Date:

Installing Guest OS on Virtual Machine

AIM

To install Guest OS on Virtual Machine using Oracle VirtualBox.

SCENARIO

You are part of an IT support team tasked with setting up a development environment for a new project. Your manager asks you to create a Virtual Machine using Oracle VirtualBox and install Ubuntu Linux as the Guest OS.

DESCRIPTION

VIRTUALIZATION

- Virtualization is the creation of virtual servers, infrastructures, devices and computing resources.
- Virtualization changes the hardware-software relations and is one of the foundational elements of cloud computing technology that helps utilize the capabilities of cloud computing to the full.
- Virtualization techniques allow companies to turn virtual their networks, storage, servers, data, desktops and applications.

Hypervisor or Virtual Machine Monitor (VMM)

A **hypervisor** or **virtual machine monitor (VMM)** is a piece of computer software, firmware or hardware that creates and runs virtual machines. A computer on which a hypervisor is running one or more virtual machines is defined as a *host machine*. Each virtual machine is called a *guest machine*. The hypervisor presents the guest operating systems with a virtual operating platform and manages the execution of the guest operating systems. Multiple instances of a variety of operating systems may share the virtualized hardware resources.

Types of Virtualization

- **Operating System Level Virtualization:** Server virtualization method where the kernel of an operating system allows for multiple isolated user-space instances, instead of just one. Such instances (sometimes called **containers**, **software containers**, virtualization engines (VE), virtual private servers (VPS)) may look and feel like a real server from the point of view of its owners and users.

- **Platform / Hardware Virtualization:** Software executed on these virtual machines is separated from the underlying hardware resources. For example, a computer that is running Microsoft Windows may host a virtual machine that looks like a computer with the Ubuntu Linux operating system; Ubuntu-based software can be run on the virtual machine.
- **In hardware virtualization**, the host machine is the actual machine on which the virtualization takes place, and the guest machine is the virtual machine. The words *host* and *guest* are used to distinguish the software that runs on the physical machine from the software that runs on the virtual machine. Different types of hardware virtualization include:
 - **Full Virtualization:** Almost complete simulation of the actual hardware to allow software, which typically consists of a guest operating system, to run unmodified.
 - **Partial Virtualization:** Some but not all of the target environment is simulated. Some guest programs, therefore, may need modifications to run in this virtual environment.
 - **Para Virtualization:** A hardware environment is not simulated; however, the guest programs are executed in their own isolated domains, as if they are running on a separate system.
- **Application Virtualization:** A software technology that encapsulates computer programs from the underlying operating system on which it is executed. A fully virtualized application is not installed in the traditional sense, although it is still executed as if it were.

PROCEDURAL STEPS

1. Download and Install Oracle Virtual Box latest version & Extension Package.

Link 1: <https://www.virtualbox.org/wiki/Downloads>

Link 2:

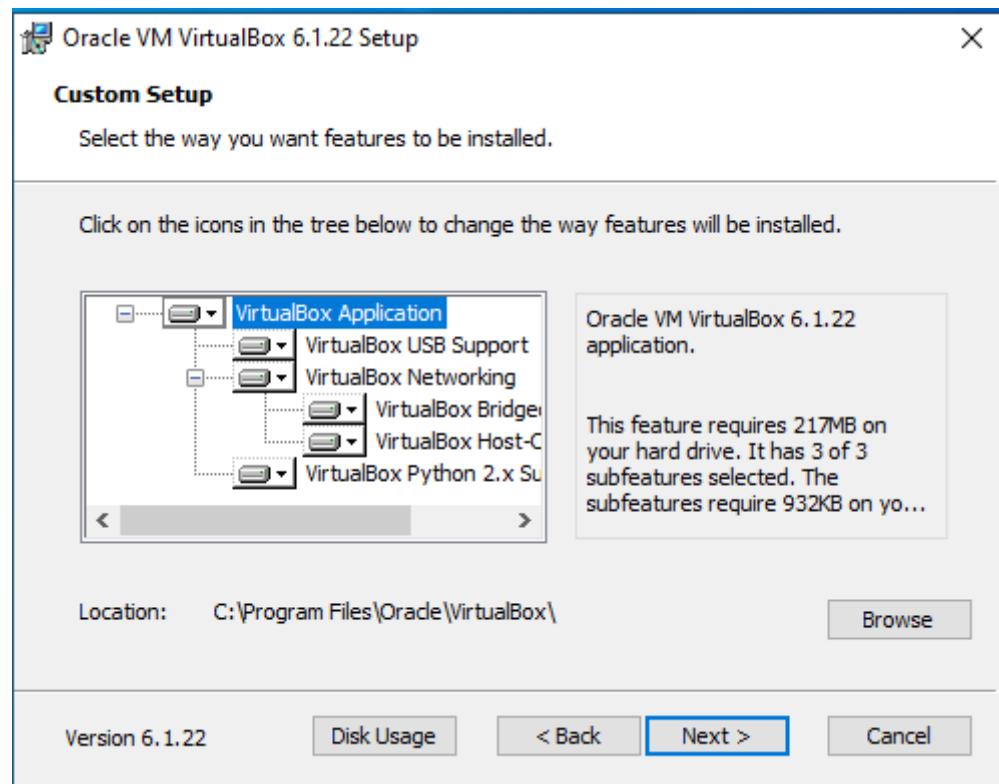
<https://www.oracle.com/in/virtualization/technologies/vm/downloads/virtualbox-downloads.html>

2. Download Ubuntu 14.4 OVA (Open Virtual Appliance)
Link: <https://www.osboxes.org/ubuntu/>

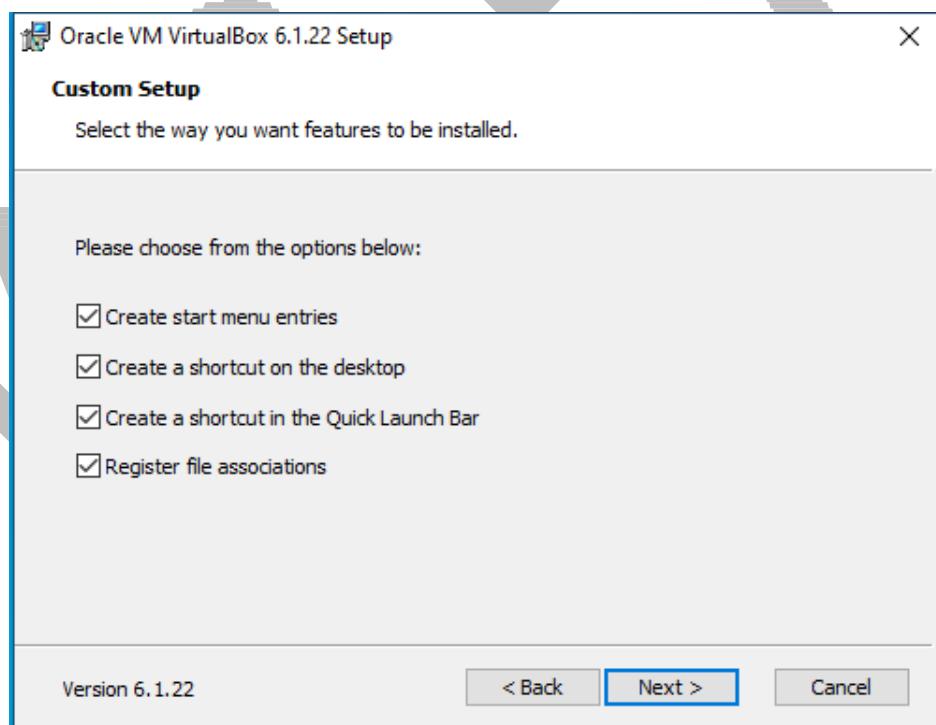
3. The files are downloaded in Local Machine → Click the Oracle VM VirtualBox 6.0.8 & Setup Wizard Move to run time environment (Open Terminal)



4. Custom Setup → Check VirtualBox Application



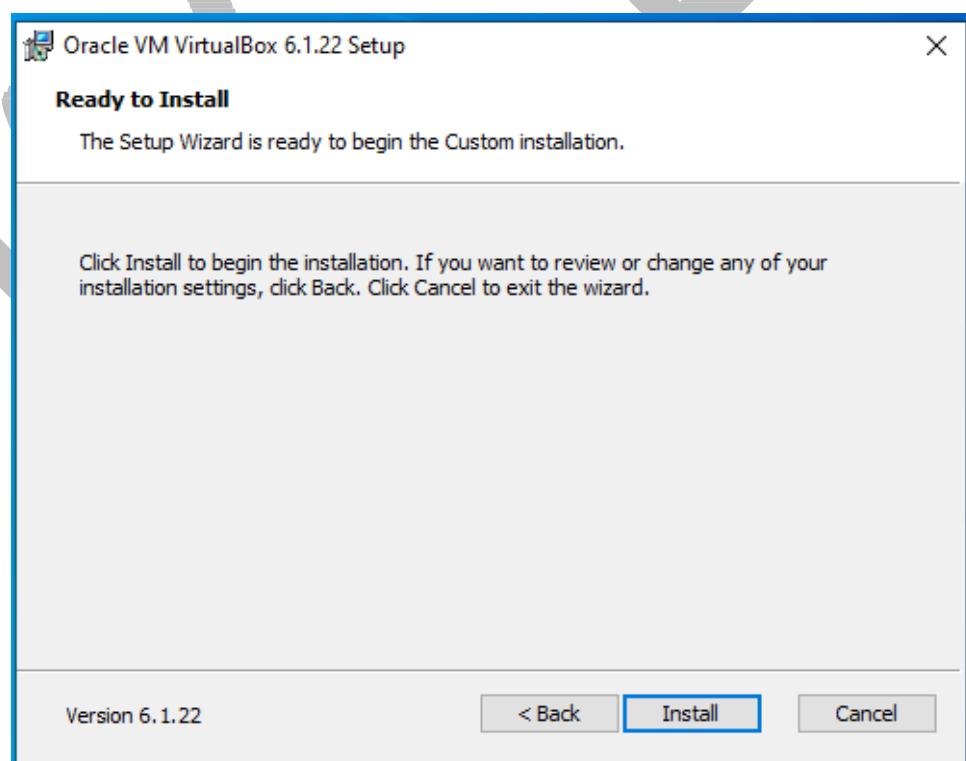
5. Custom Setup → Select the features to be installed



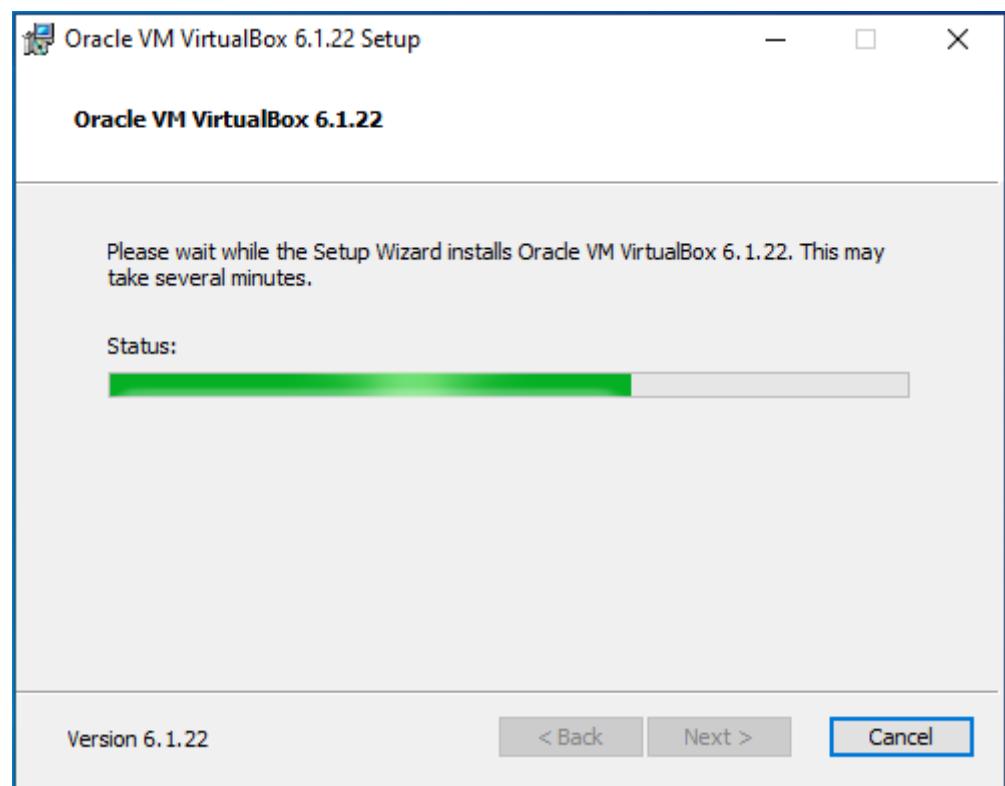
6. Warning: Network Interfaces → Click 'Yes' & Proceed to install



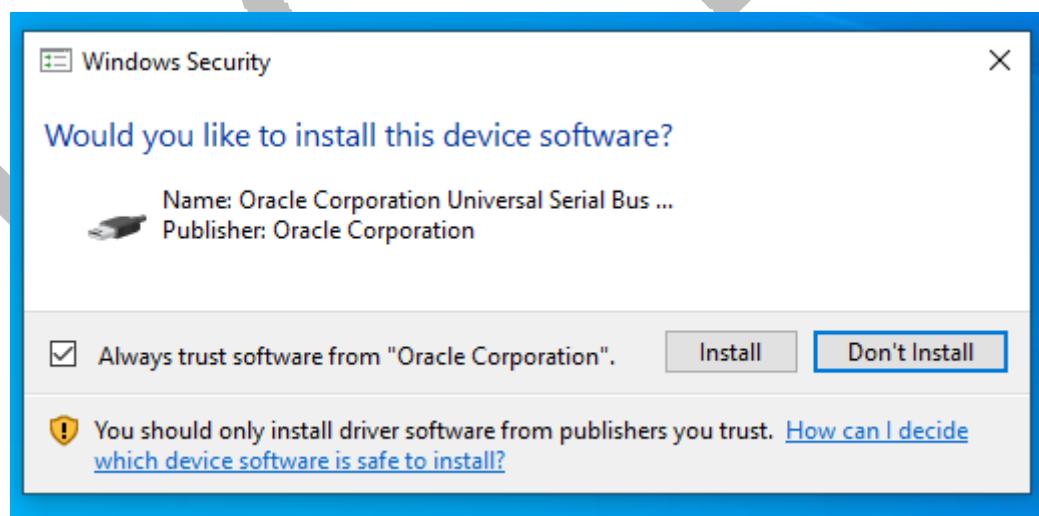
7. Ready to install → Click 'Install'



8. Oracle VM VirtualBox 6.1.22 installing



9. To install device software → Click 'Install'



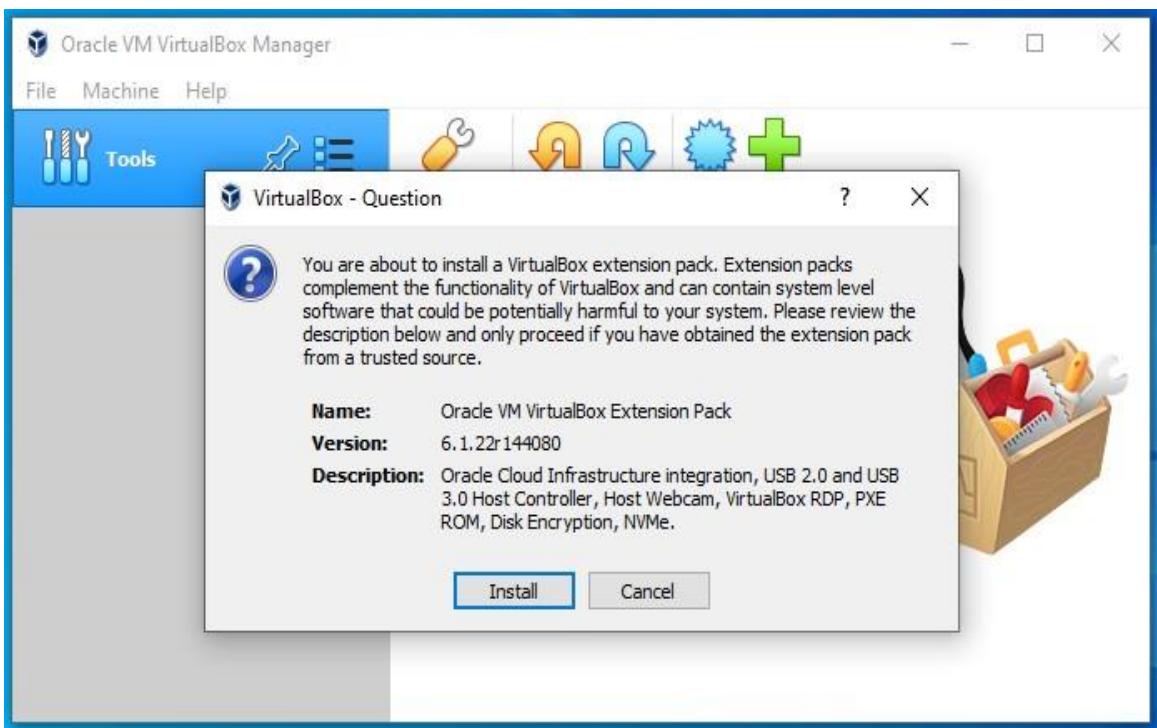
10. Oracle VM VirtualBox 6.1.22 installation is complete → Click ‘Finish’



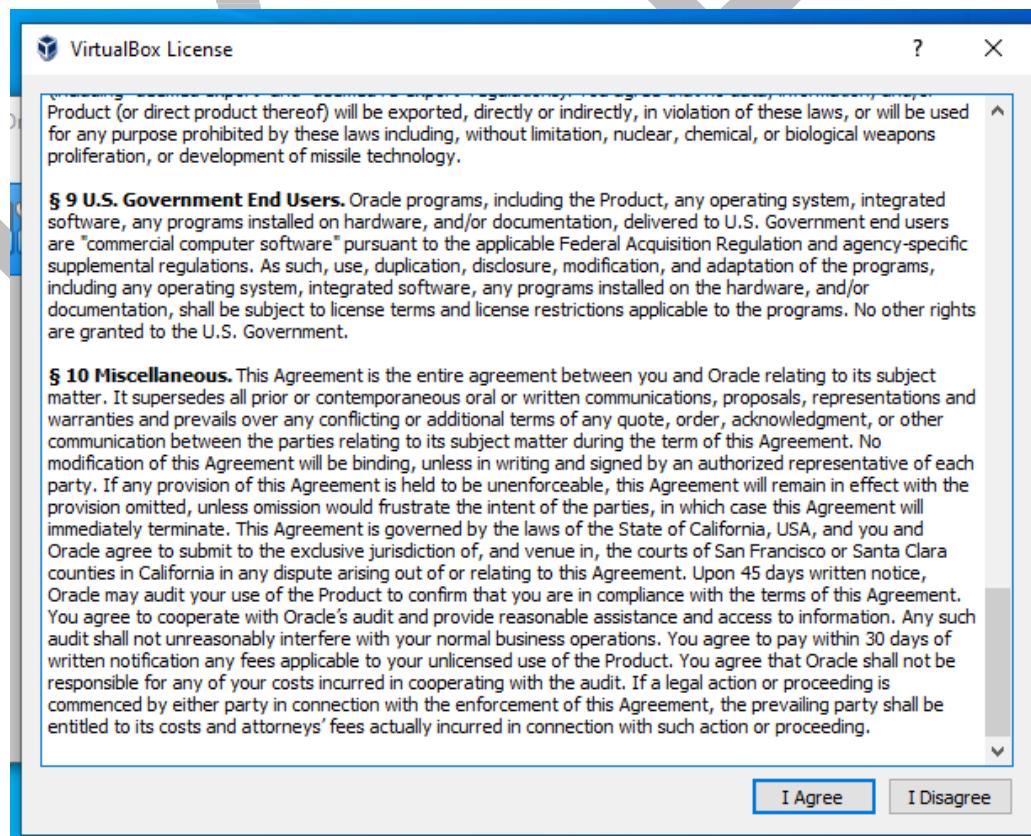
11. Import the Oracle VM Virtual Extension Pack into Oracle VirtualBox



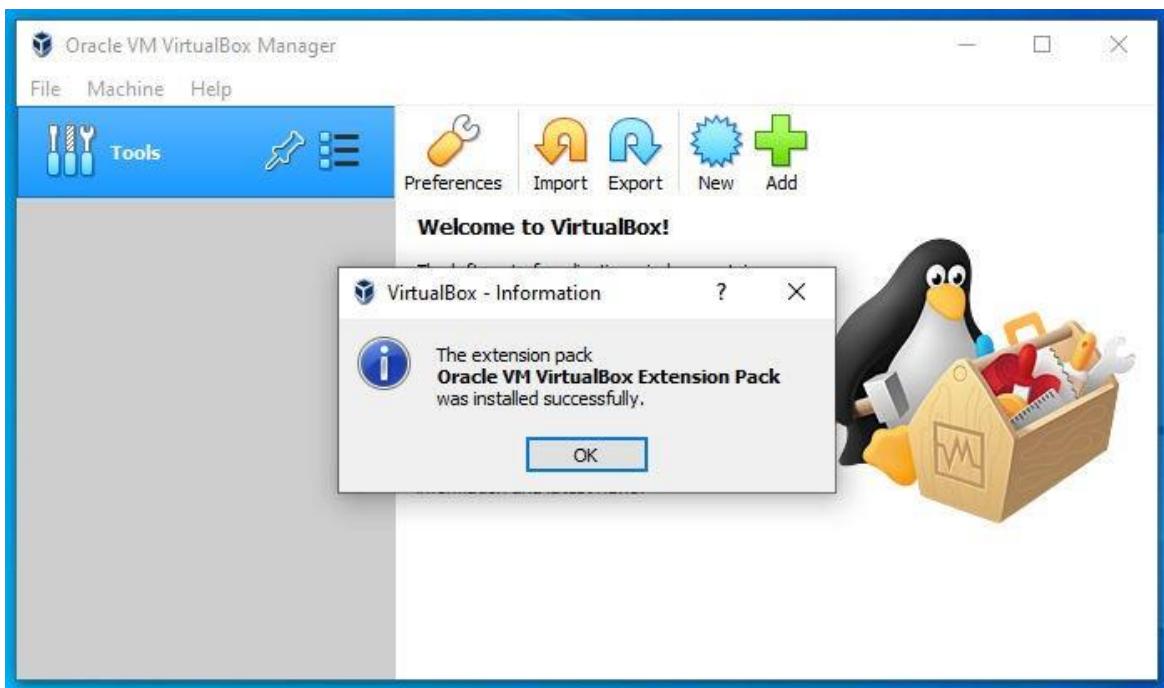
12. Click 'Install' Oracle VM Virtual Extension Pack



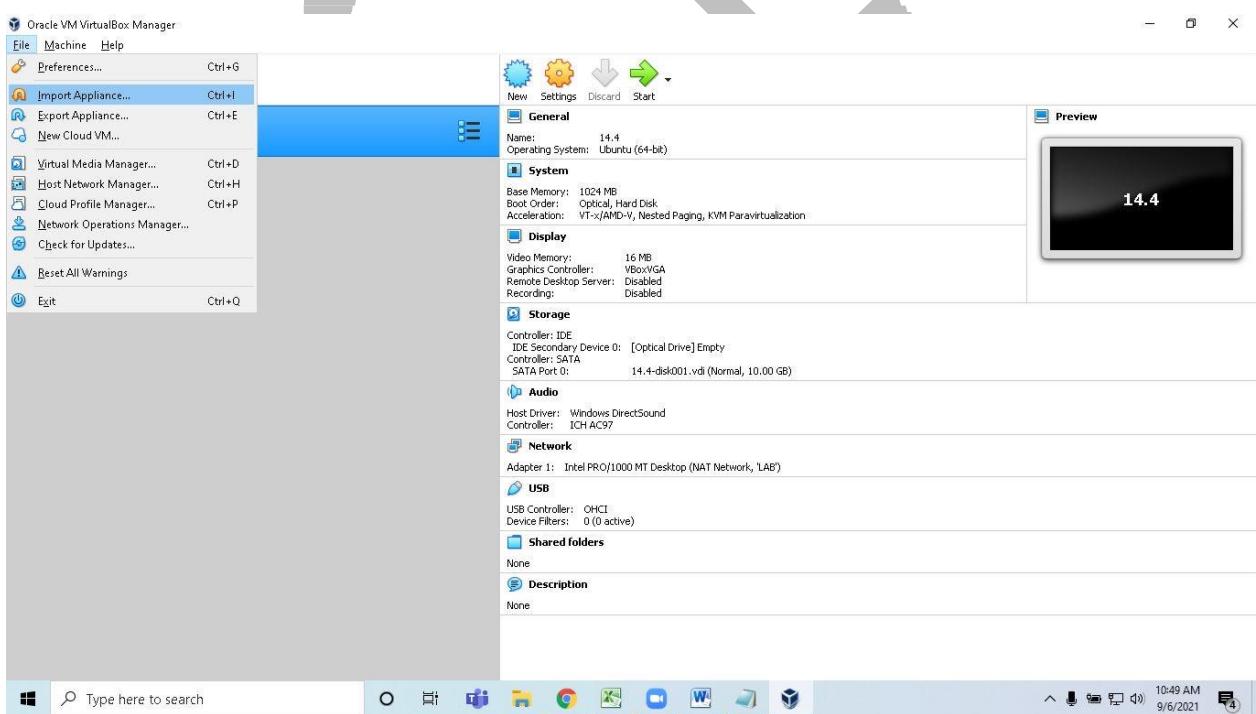
13. VirtualBox License → Click 'I Agree'

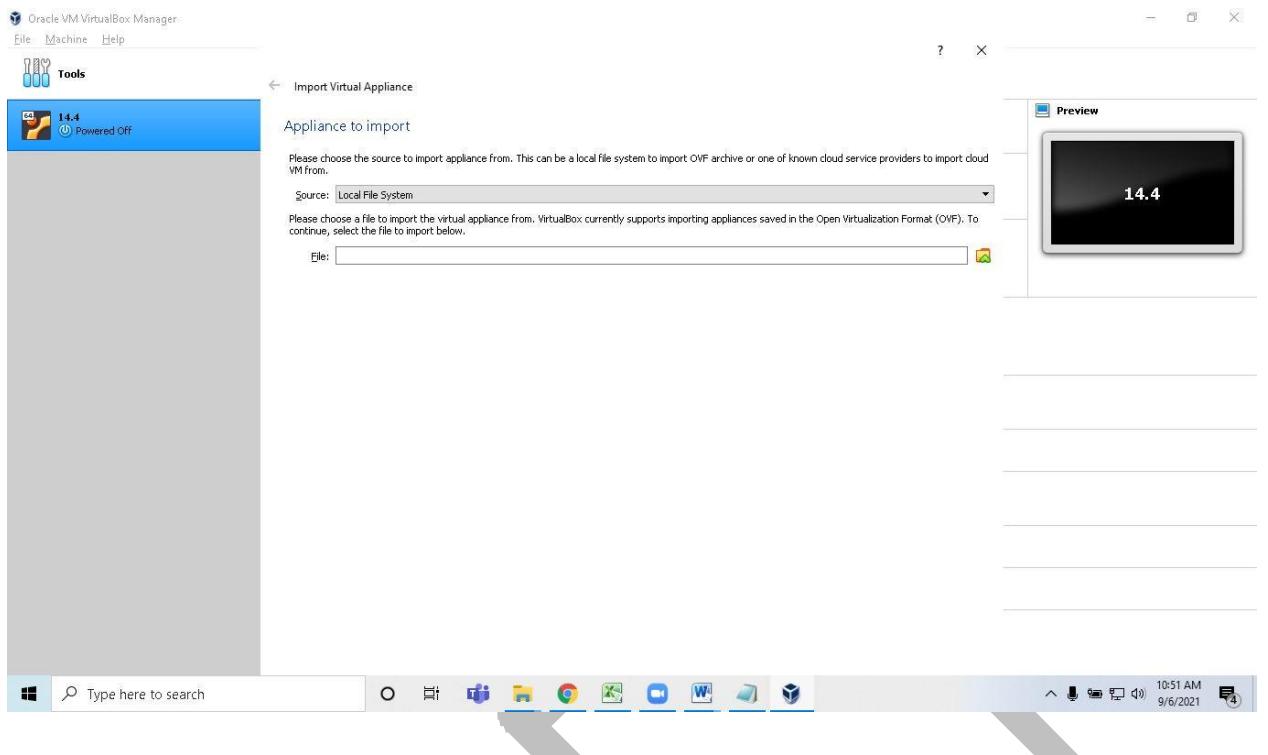


14. Extension Pack installation successful → Click ‘OK’

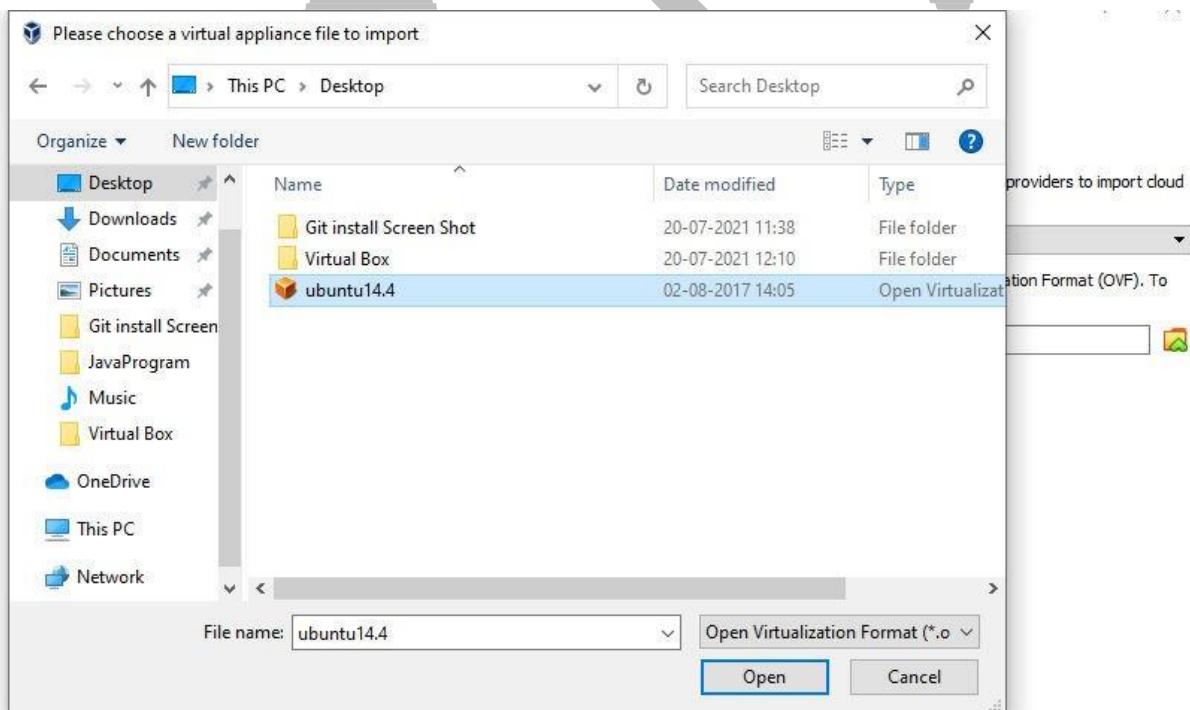


15. To install Virtual Machine: File → Import Appliance

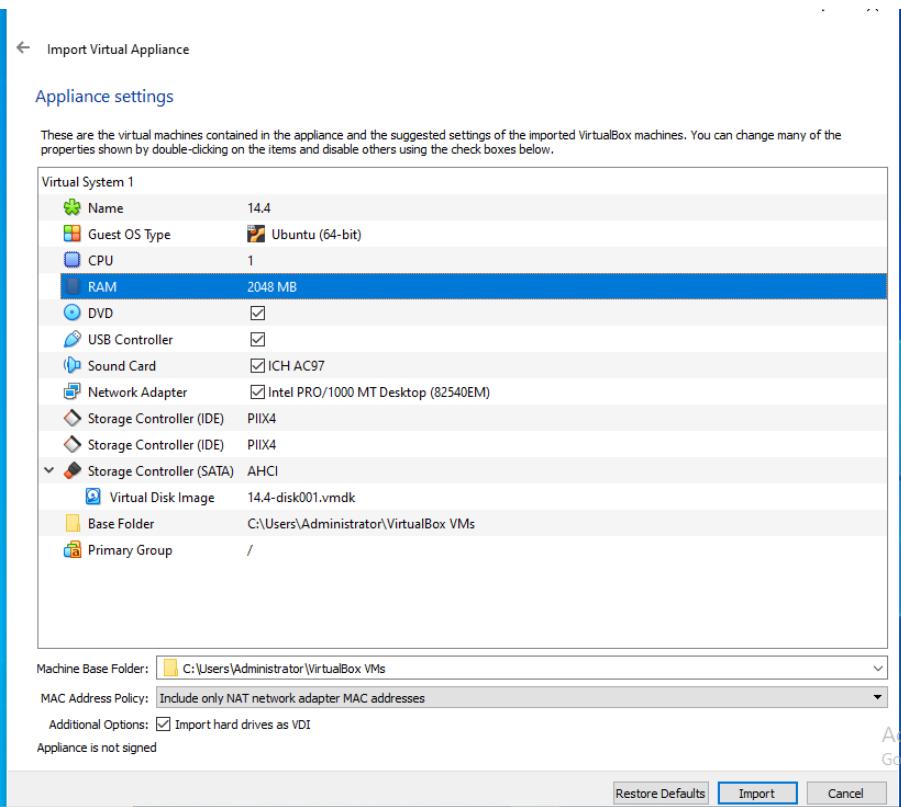




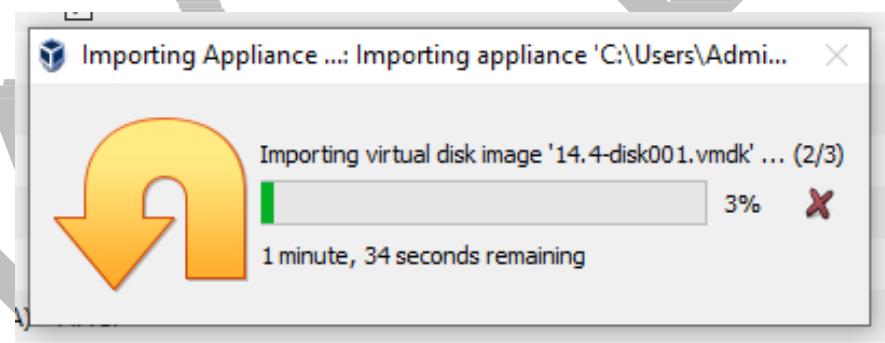
16. Select Ubuntu 14.4 OVA from directory



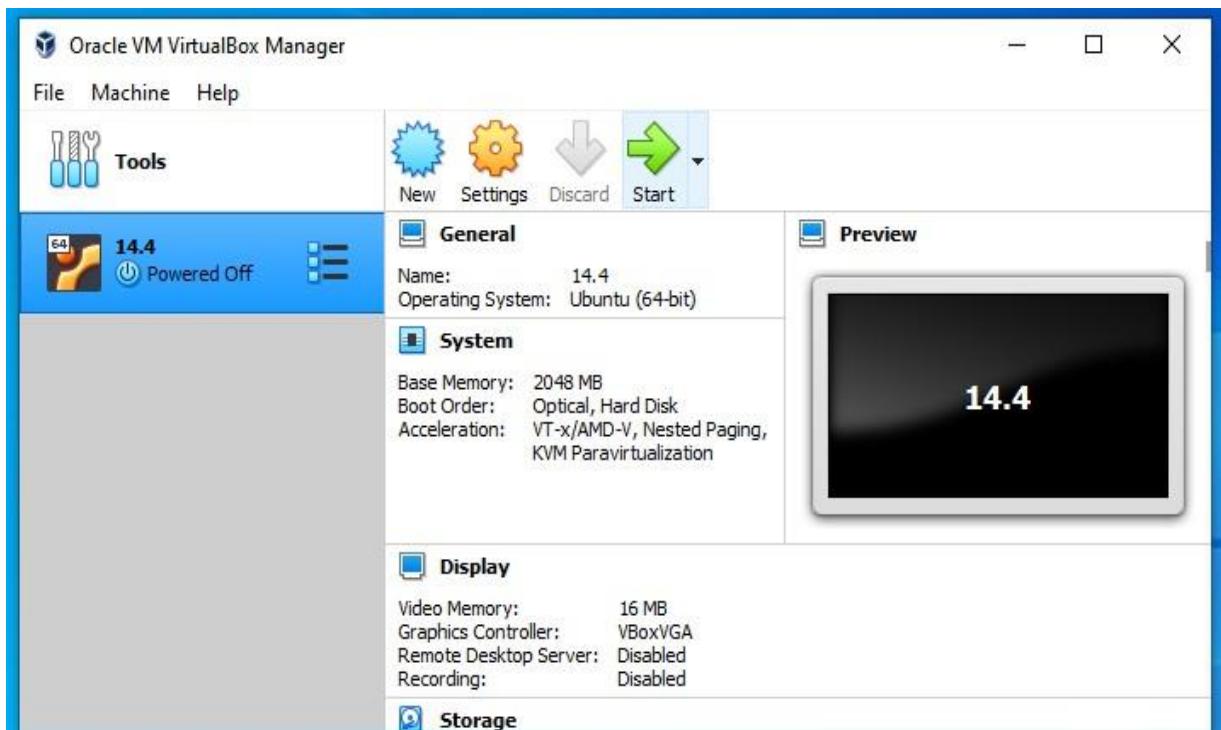
17. Appliance Settings → Choose ‘RAM’ Size → Click ‘Import’



18. Importing the Virtual Disk Image



19. Guest OS ‘Ubuntu 14.4’ is installed successfully and Click ‘Start’ button to launch the virtual machine



20. Login to Ubuntu 14.4

Login Details:

User name: hadoop

Password: Test1234

VIVA QUESTIONS

1. What is virtualization in cloud computing?

2. What is Host OS in Virtualization?

3. What is Guest OS in Virtualization?

4. List some Virtual Machine Software.

5. Name the Virtual Machine Software used in our lab exercise.

RESULT

Thus, the Guest OS is installed on Virtual Machine using Oracle VirtualBox.

Observation by students: (what student able to?)

1.

CATEGORY	MAX. MARKS ALLOTED	MARKS AWARDED
PROCEDURAL STEPS	50	
VIVA	25	
RECORD	25	
TOTAL	100	

Ex. No: 2

Date:

Installing C Complier on Guest OS

AIM

To install C Complier in the virtual machine created using a virtual box and execute Simple Programs.

PROCEDURAL STEPS

1. To download package information from all configured sources → **\$ sudo apt-get update**
2. To install C Compiler on Ubuntu 14.4 → **\$ sudo apt-get install gcc**
3. To install C++ Compiler on Ubuntu 14.4 → **\$ sudo apt-get install g++**

```
user@vmub-hadoop1:~  
Preparing to unpack .../libstdc++6_4.8.4-2ubuntu1-14.04.4_amd64.deb ...  
Unpacking libstdc++6:amd64 (4.8.4-2ubuntu1-14.04.4) over (4.8.4-2ubuntu1-14.04.3) ...  
Setting up libstdc++6:amd64 (4.8.4-2ubuntu1-14.04.4) ...  
Processing triggers for libc-bin (2.19-0ubuntu6.9) ...  
Selecting previously unselected package libstdc++4.8-dev:amd64.  
(Reading database ... 166759 files and directories currently installed.)  
Preparing to unpack .../libstdc++4.8-dev_4.8.4-2ubuntu1-14.04.4_amd64.deb ...  
Unpacking libstdc++4.8-dev:amd64 (4.8.4-2ubuntu1-14.04.4) ...  
Selecting previously unselected package g++-4.8.  
Preparing to unpack .../g++-4.8_4.8.4-2ubuntu1-14.04.4_amd64.deb ...  
Unpacking g++-4.8 (4.8.4-2ubuntu1-14.04.4) ...  
Selecting previously unselected package g++.  
Preparing to unpack .../g++_4.8_4.8.4-2ubuntu6_amd64.deb ...  
Unpacking g++ (4:4.8.4-2ubuntu6) ...  
Processing triggers for man-db (2.6.7.1-1ubuntu1) ...  
Setting up libitm1:amd64 (4.8.4-2ubuntu1-14.04.4) ...  
Setting up libomp1:amd64 (4.8.4-2ubuntu1-14.04.4) ...  
Setting up libasan0:amd64 (4.8.4-2ubuntu1-14.04.4) ...  
Setting up libatomic1:amd64 (4.8.4-2ubuntu1-14.04.4) ...  
Setting up libquadmath0:amd64 (4.8.4-2ubuntu1-14.04.4) ...  
Setting up libgcc-4.8-dev:amd64 (4.8.4-2ubuntu1-14.04.4) ...  
Setting up cpp-4.8 (4.8.4-2ubuntu1-14.04.4) ...  
Setting up libstdc++4.8-dev:amd64 (4.8.4-2ubuntu1-14.04.4) ...  
Setting up g++-4.8 (4.8.4-2ubuntu1-14.04.4) ...  
Setting up g++ (4:4.8.4-2ubuntu6) ...  
update-alternatives: using /usr/bin/g++ to provide /usr/bin/c++ (c++) in auto mode  
Processing triggers for libc-bin (2.19-0ubuntu6.9) ...  
user@vmub-hadoop1:~$ sudo gedit add.c  
  
(gedit:3811): Gtk-WARNING **: Calling Inhibit failed: GDBus.Error:org.freedesktop.DBus.Error.ServiceUnknown: The name org.gnome.SessionManager was not provided by any .service files  
  
>_(gedit:3811): Gtk-WARNING **: Calling Inhibit failed: GDBus.Error:org.freedesktop.DBus.Error.ServiceUnknown: The name org.gnome.SessionManager was not provided by any .service files  
user@vmub-hadoop1:~$ gcc add.c  
user@vmub-hadoop1:~$ ./a.out  
Enter the values for a and b:  
10 10  
The result is: 20user@vmub-hadoop1:~$
```

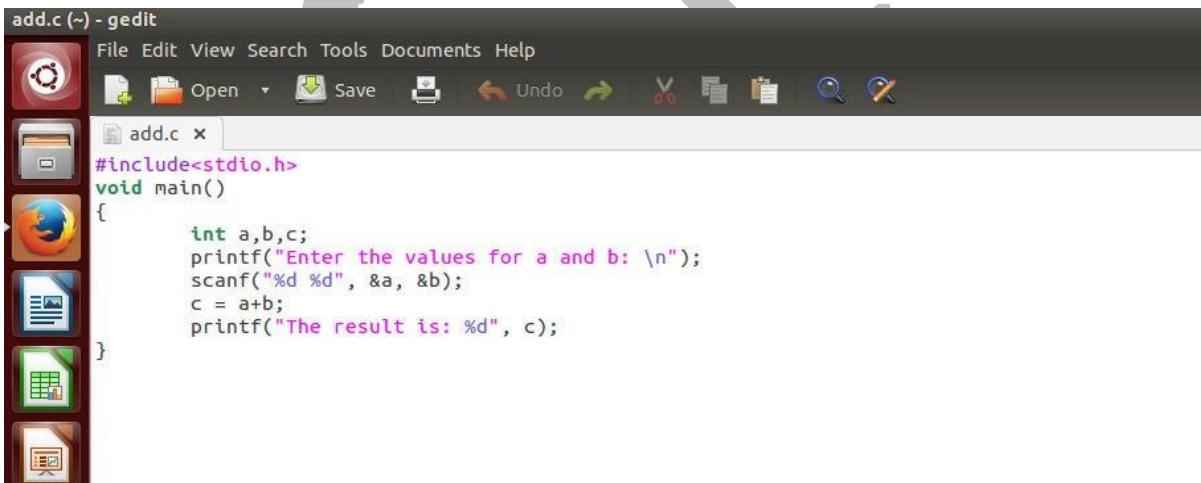
```
user@vmub-hadoop1:~  
Ign http://in.archive.ubuntu.com trusty/restricted Translation-en_IN  
Ign http://in.archive.ubuntu.com trusty/universe Translation-en_IN  
Reading package lists... Done  
user@vmub-hadoop1:~$ sudo apt-get install g++  
Reading package lists... Done  
Building dependency tree  
Reading state information... Done  
The following extra packages will be installed:  
  cpp-4.8 g++-4.8 gcc-4.8-base libasan0 libatomic1 libgcc-4.8-dev  
  libomp1 libitm1 libquadmath0 libstdc++4.8-dev libstdc++6 libtsan0  
Suggested packages:  
  gcc-4.8-locales g++-multilib g++-4.8-multilib gcc-4.8-doc libstdc++4.8-dbg  
  libgcc-4.8-multilib libgcc1-dbg libomp1-dbg libitm1-dbg libatomic1-dbg  
  libasan0-dbg libtsan0-dbg libquadmath0-dbg libstdc++4.8-dbg  
The following NEW packages will be installed:  
  g++ g++-4.8 libstdc++4.8-dev  
The following packages will be upgraded:  
  cpp-4.8 gcc-4.8-base libasan0 libatomic1 libgcc-4.8-dev libomp1  
  libitm1 libquadmath0 libstdc++6 libtsan0  
11 upgraded, 3 newly installed, 0 to remove and 445 not upgraded.  
Need to get 30.9 MB/30.9 MB of archives.  
After this operation, 40.1 MB of additional disk space will be used.  
Do you want to continue? [Y/n] y  
Get:1 http://in.archive.ubuntu.com/ubuntu/ trusty-updates/main libitm1 amd64 4.8.4-2ubuntu1-14.04.4 [28.6 kB]  
Get:2 http://in.archive.ubuntu.com/ubuntu/ trusty-updates/main libomp1 amd64 4.8.4-2ubuntu1-14.04.4 [23.1 kB]  
Get:3 http://in.archive.ubuntu.com/ubuntu/ trusty-updates/main libasan0 amd64 4.8.4-2ubuntu1-14.04.4 [63.1 kB]  
Get:4 http://in.archive.ubuntu.com/ubuntu/ trusty-updates/main libatomic1 amd64 4.8.4-2ubuntu1-14.04.4 [8,630 B]  
Get:5 http://in.archive.ubuntu.com/ubuntu/ trusty-updates/main libtsan0 amd64 4.8.4-2ubuntu1-14.04.4 [94.8 kB]  
Get:6 http://in.archive.ubuntu.com/ubuntu/ trusty-updates/main libquadmath0 amd64 4.8.4-2ubuntu1-14.04.4 [126 kB]  
Get:7 http://in.archive.ubuntu.com/ubuntu/ trusty-updates/main libgcc-4.8-dev amd64 4.8.4-2ubuntu1-14.04.4 [1,688 kB]  
Get:8 http://in.archive.ubuntu.com/ubuntu/ trusty-updates/main gcc-4.8 amd64 4.8.4-2ubuntu1-14.04.4 [5,040 kB]  
Get:9 http://in.archive.ubuntu.com/ubuntu/ trusty-updates/main cpp-4.8 amd64 4.8.4-2ubuntu1-14.04.4 [4,552 kB]  
Get:10 http://in.archive.ubuntu.com/ubuntu/ trusty-updates/main libgcc-4.8-base amd64 4.8.4-2ubuntu1-14.04.4 [16.7 kB]  
Get:11 http://in.archive.ubuntu.com/ubuntu/ trusty-updates/main libstdc++4.8-dev amd64 4.8.4-2ubuntu1-14.04.4 [260 kB]  
Get:12 http://in.archive.ubuntu.com/ubuntu/ trusty-updates/main libstdc++4.8-dev amd64 4.8.4-2ubuntu1-14.04.4 [1,051 kB]  
Get:13 http://in.archive.ubuntu.com/ubuntu/ trusty-updates/main g++-4.8 amd64 4.8.4-2ubuntu1-14.04.4 [18.0 MB]  
Fetched 30.9 MB (927 kB/s)  
(Reading database ... 166759 files and directories currently installed.)  
Preparing to unpack .../libitm1_4.8.4-2ubuntu1-14.04.4_amd64.deb ...  
Unpacking libitm1:amd64 (4.8.4-2ubuntu1-14.04.4) over (4.8.4-2ubuntu1-14.04.3) ...  
Preparing to unpack .../libomp1_4.8.4-2ubuntu1-14.04.4_amd64.deb ...
```

```
user@vmub-hadoop1:~$ sudo apt-get install gcc
[sudo] password for user:
Reading package lists... Done
Building dependency tree
Reading state information... Done
gcc is already the newest version.
0 upgraded, 0 newly installed, 0 to remove and 296 not upgraded.
user@vmub-hadoop1:~$ sudo apt install g++
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following extra packages will be installed:
  g++-4.8 libstdc++-4.8-dev
Suggested packages:
  g++-multilib g++-4.8-multilib gcc-4.8-doc libstdc++6-4.8-dbg
  libstdc++-4.8-doc
The following NEW packages will be installed:
  g++ g++-4.8 libstdc++-4.8-dev
0 upgraded, 3 newly installed, 0 to remove and 296 not upgraded.
Need to get 19.2 MB of archives.
After this operation, 40.0 MB of additional disk space will be used.
Do you want to continue? [Y/n]
```

4. Create a file to write C program → \$ sudo gedit <<file_name>>.c

To compile C program → \$ gcc <<file_name>>.c

To run C program → \$./a.out



```
add.c (~) - gedit
File Edit View Search Tools Documents Help
Open Save Undo Redo Cut Copy Paste Find Replace
add.c x
#include<stdio.h>
void main()
{
    int a,b,c;
    printf("Enter the values for a and b: \n");
    scanf("%d %d", &a, &b);
    c = a+b;
    printf("The result is: %d", c);
}
```

```
user@vmub-hadoop1:~$ sudo gedit add.c
(gedit:3811): Gtk-WARNING **: Calling Inhibit failed: GDBus.Error:org.free
was not provided by any .service files

(gedit:3811): Gtk-WARNING **: Calling Inhibit failed: GDBus.Error:org.free
was not provided by any .service files
user@vmub-hadoop1:~$ gcc add.c
user@vmub-hadoop1:~$ ./a.out
Enter the values for a and b:
10 10
The result is: 20user@vmub-hadoop1:~$
```

arm.c (~) - gedit

The screenshot shows a Linux desktop environment. On the left, there is a vertical file manager sidebar with icons for various applications like the Dash, Home, and specific files. The main workspace contains a terminal window at the bottom and a gedit code editor window above it. The terminal window has a dark background and displays the user's command-line session.

```
#include<stdio.h>

int main()
{
    int n,r,sum=0,temp;
    printf("enter the number = ");
    scanf("%d",&n);
    temp=n;
    while(n>0)
    {
        r=n%10;
        sum=sum+(r*r*r);
        n=n/10;
    }
    if(temp==sum)
        printf("armstrong number \n");
    else
        printf("not armstrong number \n");
    return 0;
}
```

user@vmub-hadoop1: ~

```
user@vmub-hadoop1:~$ gcc arm.c
user@vmub-hadoop1:~$ ./a.out
enter the number = 153
armstrong number
user@vmub-hadoop1:~$ ./a.out
enter the number = 150
not armstrong number
user@vmub-hadoop1:~$
```

odd.c (~) - gedit

The screenshot shows a desktop environment with a window titled "odd.c (~) - gedit". The window contains a C program named "odd.c" with the following code:

```
#include<stdio.h>
int main()
{
    int num;

    printf("Enter an integer: ");
    scanf("%d",&num);

    if ( num%2 == 0 )
        printf("%d is an even number \n", num);
    else
        printf("%d is an odd number \n", num);

    return 0;
}
```

user@vmub-hadoop1: ~

The screenshot shows a terminal window with the following session:

```
user@vmub-hadoop1:~$ gcc odd.c
user@vmub-hadoop1:~$ ./a.out
Enter an integer: 5
5 is an odd number
user@vmub-hadoop1:~$ ./a.out
Enter an integer: 10
10 is an even number
user@vmub-hadoop1:~$ █
```

VIVA QUESTIONS

1. What is Parallel Computing?

2. What is Distributed System?

3. List the compilers to run C program.

4. Name the command to compile C program in Linux.

5. Which command is used to run C Program in Linux?

RESULT

Thus, the C Complier is installed on Guest OS ‘Ubuntu 14.4’ and simple programs are executed.

Observation by students: (what student able to?)

1.

CATEGORY	MAX. MARKS ALLOTED	MARKS AWARDED
PROCEDURAL STEPS	50	
VIVA	25	
RECORD	25	
TOTAL	100	

Ex. No: 3

Date:

Create, Deploy and Manage Java Web Application in AWS

AIM

To create, deploy and manage Java web application in Amazon Web Services.

SCENARIO

You are a junior cloud engineer in a tech company, and your manager has assigned you the task of creating, deploying, and managing a Java web application on Amazon Web Services (AWS). The application will be used by the development team for testing new features before production deployment.

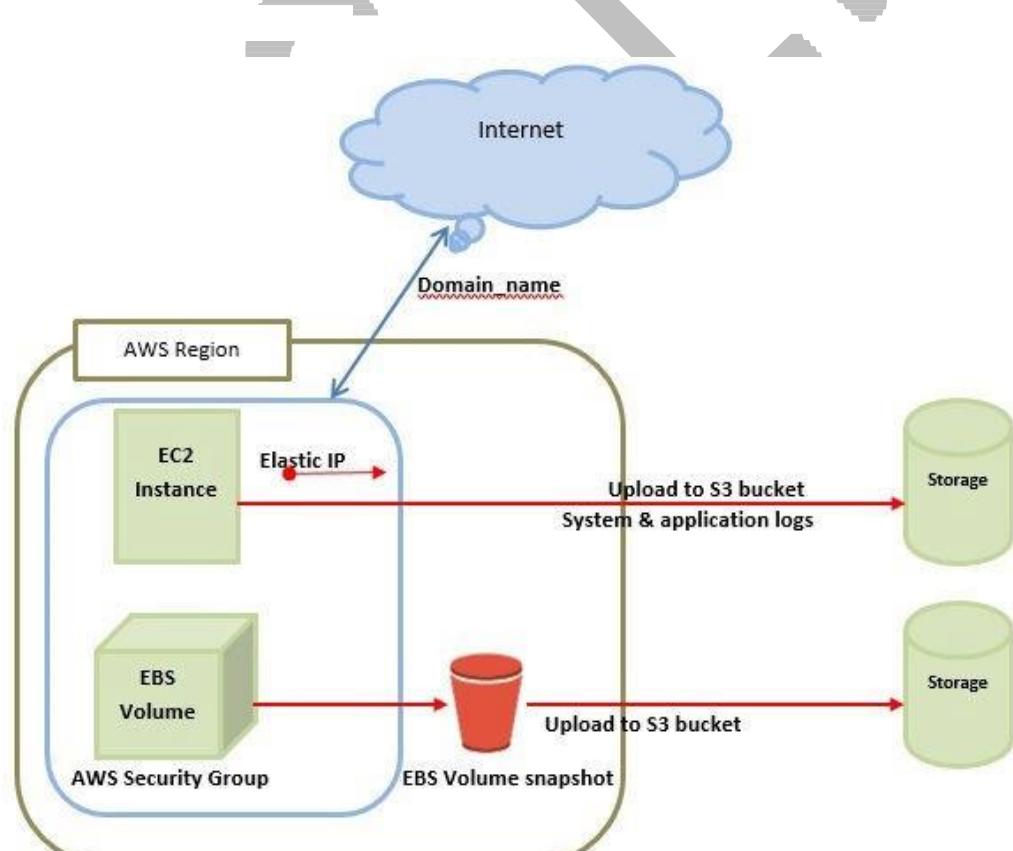
DESCRIPTION

AMAZON WEB SERVICES (AWS)

Amazon Web Services (AWS) is Amazon's cloud web hosting platform that offers flexible, reliable, scalable, easy-to-use, and cost-effective solutions. This tutorial covers various important topics illustrating how AWS works and how it is beneficial to run your website on Amazon Web Services.

AWS Basic Architecture

AWS EC2, where EC2 stands for Elastic Compute Cloud. EC2 allows users to use virtual machines of different configurations as per their requirement. It allows various configuration options, mapping of individual server, various pricing options, etc. We will discuss these in detail in AWS Products section.



Note – In the above diagram **S3** stands for Simple Storage Service. It allows the users to store and retrieve various types of data using API calls. It doesn't contain any computing element. We will discuss this topic in detail in AWS products section.

AWS Elastic Beanstalk

AWS Elastic Beanstalk is an easy-to-use service for deploying and scaling web applications and services developed with Java, .NET, PHP, Node.js, Python, Ruby, Go, and Docker on familiar servers such as Apache, Nginx, Passenger, and IIS.

User can simply upload the code and Elastic Beanstalk automatically handles the deployment, from capacity provisioning, load balancing, auto-scaling to application health monitoring. At the same time, user can retain full control over the AWS resources powering the application and can access the underlying resources at any time.

There is no additional charge for Elastic Beanstalk - user can pay only for the AWS resources needed to store and run your applications.

AWS Elastic Beanstalk is an orchestration service offered by Amazon Web Services for deploying applications which orchestrates various AWS services, including EC2, S3, Simple Notification Service (SNS), CloudWatch, autoscaling, and Elastic Load Balancers. Deployment requires several components to be defined: an 'application' as a logical container for the project, a 'version' which is a deployable build of the application executable, a 'configuration template' that contains configuration information for both the Beanstalk environment and for the product. Finally, an 'environment' combines a 'version' with a 'configuration' and deploys them. Executables themselves are uploaded as archive files to S3 beforehand and the 'version' is just a pointer.

PROCEDURAL STEPS

1. To download all necessary setups use following links

JDK 8 Download:

<https://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html>

Eclipse IDE for Java EE:

<https://www.eclipse.org/downloads/packages/release/neon/3/eclipse-ide-java-ee-developers>

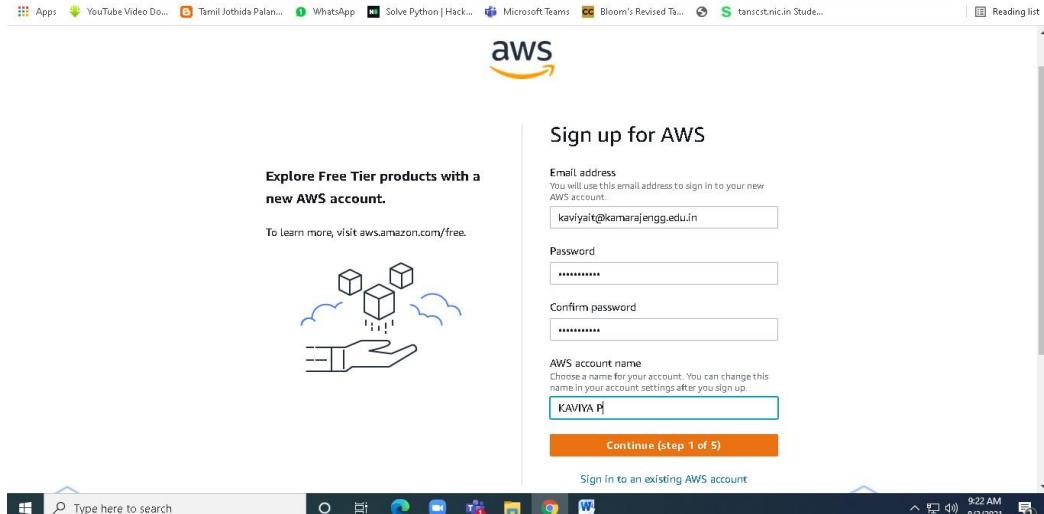
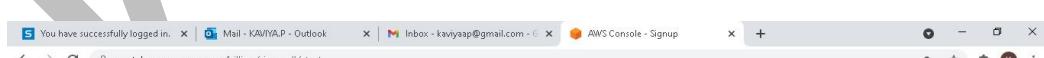
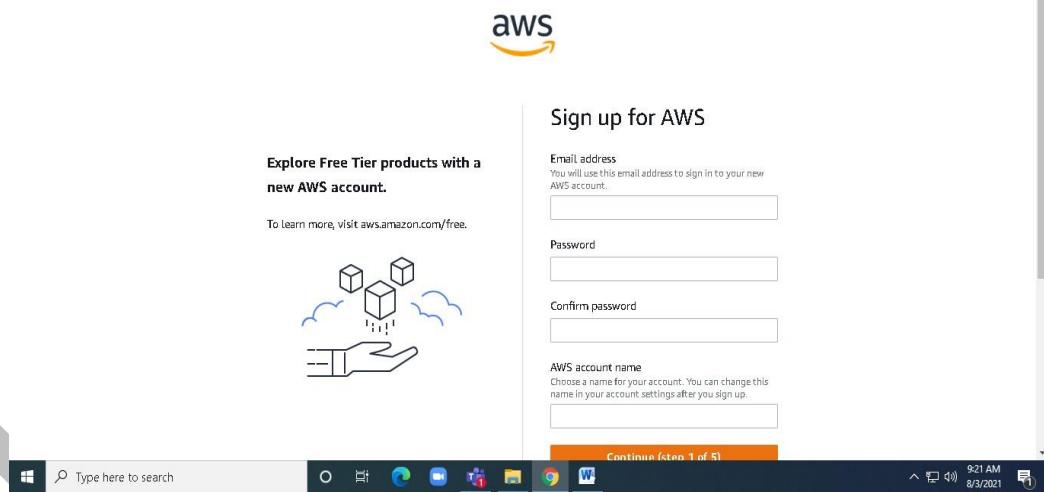
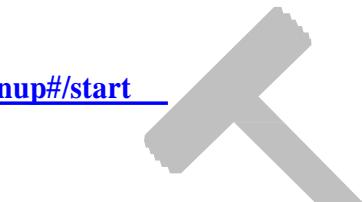
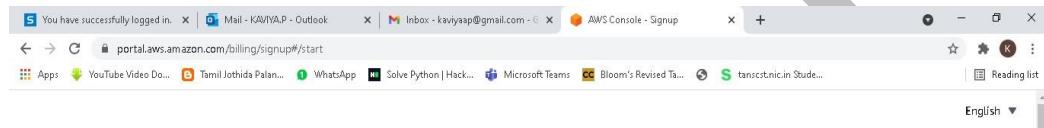
Apache Tomcat 8.0: <https://tomcat.apache.org/download-80.cgi>

2. Account Creation in AWS

2.1 Create an AWS account using

<https://portal.aws.amazon.com/billing/signup#/start>

Enter all necessary details



Enter Billing Information

The screenshot shows the 'Sign up for AWS' page. On the left, there's a 'Secure verification' section with a note: 'We will not charge for usage below AWS Free Tier limits. We temporarily hold INR 2 as a pending transaction for 3-5 days to verify your identity.' Below this is a shield icon with a checkmark. On the right, there's a 'Billing Information' section with fields for 'Credit or Debit card number' (with VISA, MasterCard, AMEX, and Discover logos), 'Expiration date' (Month and Year dropdowns), 'Cardholder's name' (text input), 'CVV' (text input), and 'Billing address' (radio button for 'Use my contact address'). The status bar at the bottom shows it's 9:29 AM on 8/3/2021.

Enter OTP for payment (Message received in registered Mobile Number)

The screenshot shows the 'acs2.onlinesbi.com/bdacs/SBI/Validate/M' page. It displays payment details: Merchant: AMAZON INTERNET SERVICES, Transaction Amount: ₹ 2.00, SBI Debit Card: 5103xxxx xxxx xx59. Below this is an 'Authenticate Payment' section with the message 'OTP sent to your mobile number ending 90xx9xxx14'. It has a 'Enter One Time Password (OTP)' input field, a 'Make Payment' button, and a 'Resend OTP' link. At the bottom, there's a note about activating the SBI debit card and a 'PCI DSS Certified' badge. The status bar at the bottom shows it's 9:32 AM on 8/3/2021.

Confirm your Identity

The screenshot shows a Microsoft Edge browser window with the AWS Console Signup page. The URL is <https://portal.aws.amazon.com/billing/signup?ppw-messageProcessedResponse=APPROVED&ppw-widgetEvent=MessageProcessedEvent&ppw-widgetState=5-7aCz9pUr8itr-wzEKA...>. The page title is "Sign up for AWS". A large icon of a user profile with a checkmark is displayed. The main text says: "Before you can use your AWS account, you must verify your phone number. When you continue, the AWS automated system will contact you with a verification code." Below this, there are two radio buttons: "Text message (SMS)" (selected) and "Voice call". A dropdown menu for "Country or region code" shows "United States (+1)". A text input field for "Mobile phone number" is empty. A "Security check" section contains a CAPTCHA image. The Windows taskbar at the bottom shows various pinned apps like Mail, Outlook, and Teams.

Confirm your Identity using Verification Code (Code received in registered Mobile Number)

The screenshot shows the same Microsoft Edge browser window on the "Sign up for AWS" page. The URL has changed to <https://portal.aws.amazon.com/billing/signup?ppw-messageProcessedResponse=APPROVED&ppw-widgetEvent=MessageProcessedEvent&ppw-widgetState=5-7aCz9pUr8itr-wzEKA...>. The "Confirm your identity" section now includes a "Verify code" input field and a prominent orange "Continue (step 4 of 5)" button. Below the button, a note reads: "Having trouble? Sometimes it takes up to 10 minutes to retrieve a verification code. If it's been longer than that, return to the previous page and try again." The Windows taskbar at the bottom remains the same.

Select a support plan → Basic support - Free

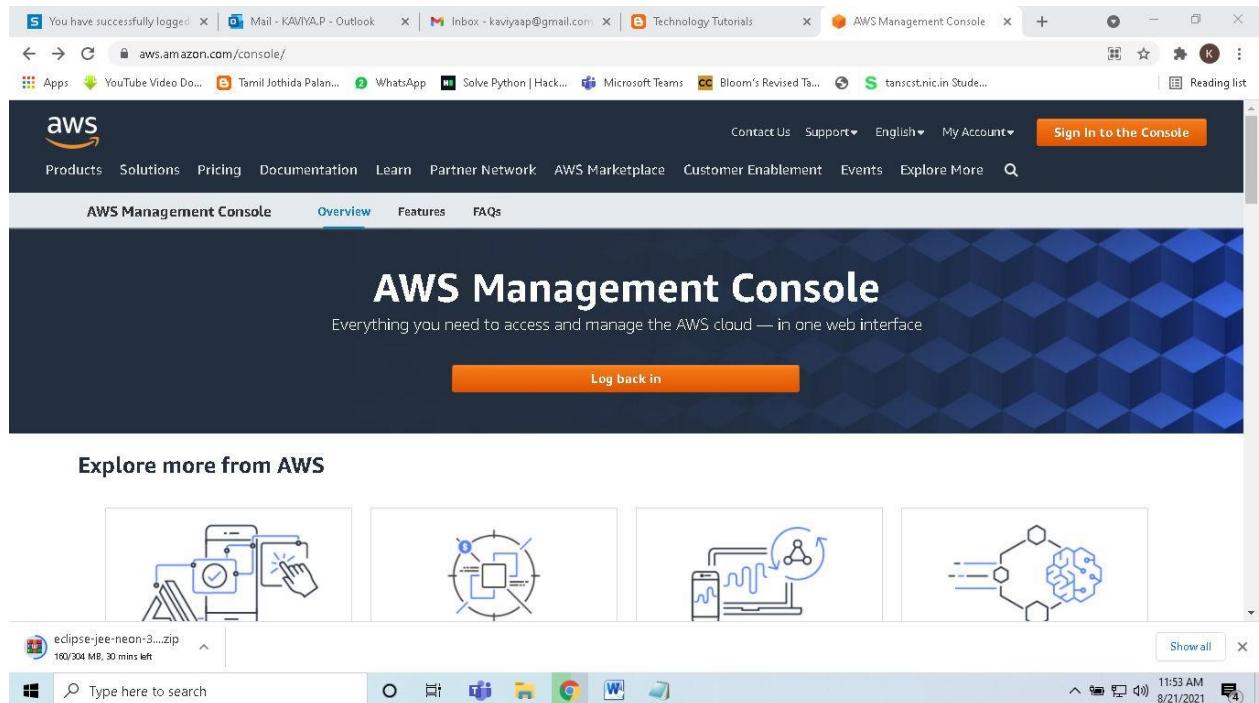
The screenshot shows the AWS Signup page. At the top, it says "Sign up for AWS" and "Select a support plan". Below that, it says "Choose a support plan for your business or personal account. Compare plans and pricing examples" and "You can change your plan anytime in the AWS Management Console." There are three options: "Basic support - Free" (selected), "Developer support - From \$29/month", and "Business support - From \$100/month". Each option has a description and an icon. Below the options, it says "Need Enterprise level support? From \$15,000 a month you will receive 15-minute response times and concierge-style experience with an assigned Technical Account Manager. Learn more" with a link icon.

AWS account creation is successful

The screenshot shows the AWS Signup page with a large "aws" logo at the top. Below it is a rocket launching from a cloud icon. The text "Congratulations" is displayed, followed by "Thank you for signing up for AWS." and "We are activating your account, which should only take a few minutes. You will receive an email when this is complete." A "Go to the AWS Management Console" button is shown. The Windows taskbar at the bottom indicates the date and time as 8/3/2021 9:34 AM.

2.2 AWS – Sign in to the console

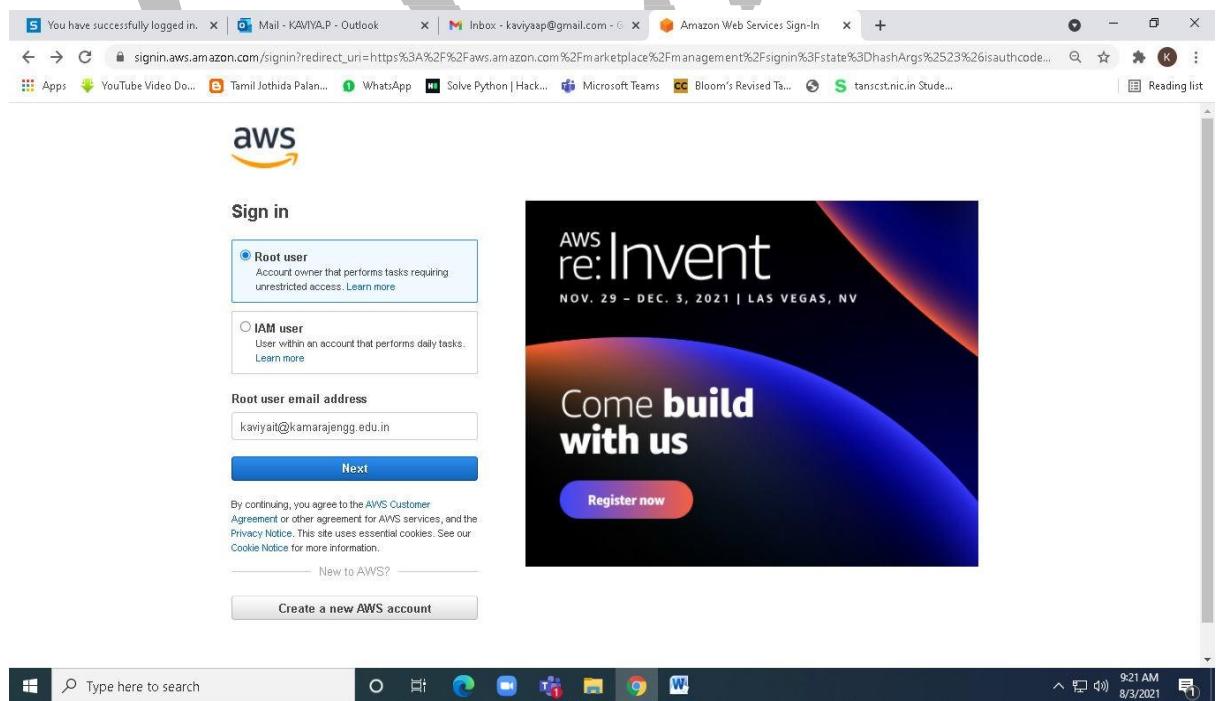
Link: <https://aws.amazon.com/console/>



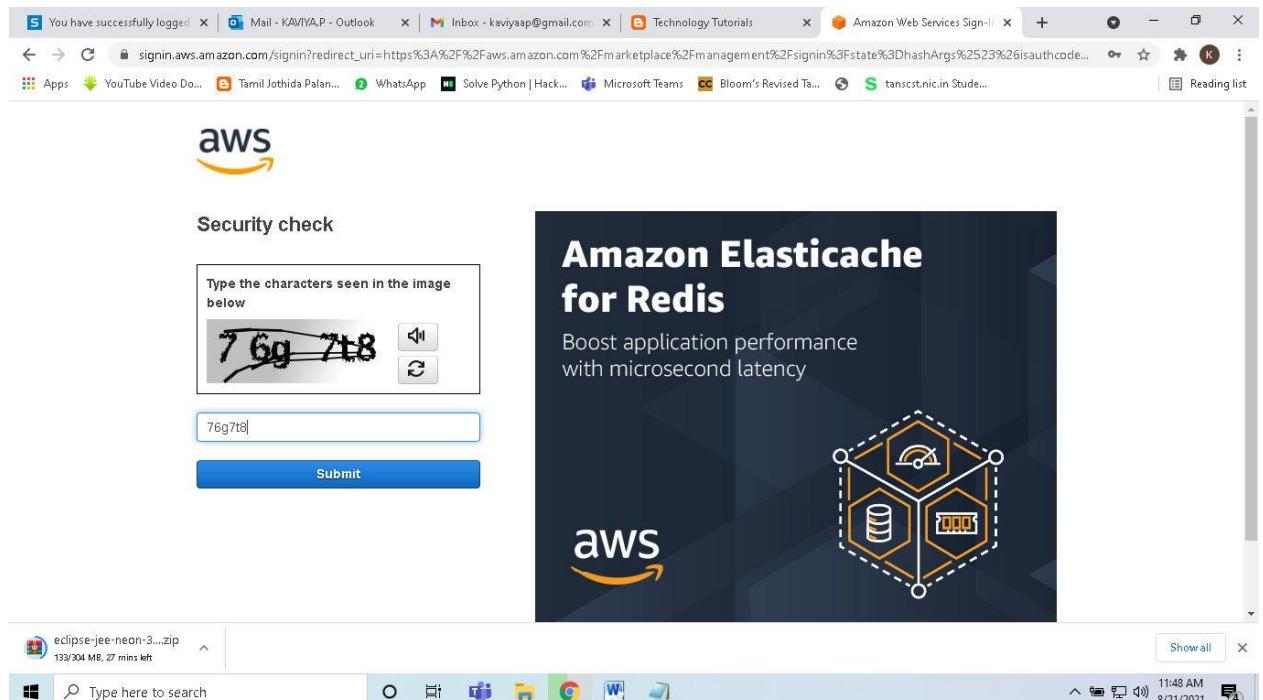
Sign in as “Root User”

Give necessary credentails (E-Mail address, Security check and Password)

Enter E-Mail Address



Enter Security check



You have successfully logged in | Mail - KAVIYAP - Outlook | Inbox - kavyaap@gmail.com | Technology Tutorials | Amazon Web Services Sign-in | + | - | ×

signin.aws.amazon.com/signin?redirect_uri=https%3A%2F%2Faws.amazon.com%2Fmanagement%2Fsign-in%3Fstate%3DhashArgs%2523%26isauthcode... | Reading list

Apps YouTube Video Do... Tamil Jothida Palan... WhatsApp Solve Python | Hack... Microsoft Teams Bloom's Revised Ta... S tanscstrnic.in Stud...

Security check

Type the characters seen in the image below

76g7t8

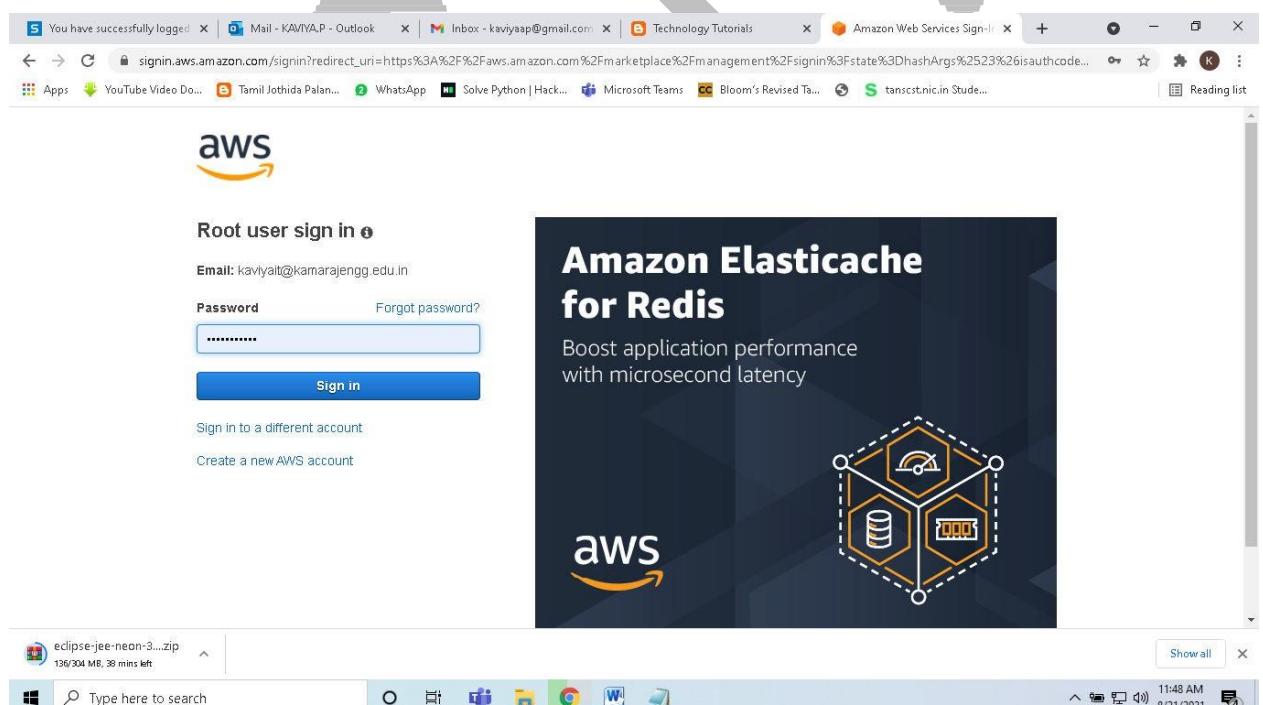
Submit

Amazon Elasticache for Redis

Boost application performance with microsecond latency

aws

Enter Password and Click “Sign in”



You have successfully logged in | Mail - KAVIYAP - Outlook | Inbox - kavyaap@gmail.com | Technology Tutorials | Amazon Web Services Sign-in | + | - | ×

signin.aws.amazon.com/signin?redirect_uri=https%3A%2F%2Faws.amazon.com%2Fmanagement%2Fsign-in%3Fstate%3DhashArgs%2523%26isauthcode... | Reading list

Apps YouTube Video Do... Tamil Jothida Palan... WhatsApp Solve Python | Hack... Microsoft Teams Bloom's Revised Ta... S tanscstrnic.in Stud...

aws

Root user sign in

Email: kavyait@kamarajengg.edu.in

Password [Forgot password?](#)

Sign in

Sign in to a different account

Create a new AWS account

Amazon Elasticache for Redis

Boost application performance with microsecond latency

aws

AWS Management Console & Click “Elastic Beanstalk”

The screenshot shows the AWS Management Console homepage. The top navigation bar includes links for Mail, Inbox, Technology Tutorials, AWS Management Console, and other services like AWS Marketplace Subscriptions and Elastic Beanstalk. The main content area features a sidebar titled "AWS services" with "Recently visited services" showing "AWS Marketplace Subscriptions" and "Elastic Beanstalk". Below this is a section titled "Build a solution" with a link to "Get started with simple wizards and automated workflows". To the right, there's a "Stay connected to your AWS resources on-the-go" section about the AWS Console Mobile App, and an "Explore AWS" section with a link to "Lower Data Processing Charges for AWS PrivateLink". At the bottom, there's a search bar, a taskbar with various icons, and a status bar showing the date and time.

In Elastic Beanstalk – Click “Create Application”

The screenshot shows the Amazon Elastic Beanstalk landing page. The top navigation bar is identical to the previous screenshot. The main content area features a large banner with the text "Amazon Elastic Beanstalk" and "End-to-end web application management.". Below the banner, there's a "Get started" section with a "Create Application" button, a "How it works" section explaining the deployment process, and a "Pricing" section stating that there's no additional charge. The bottom of the page includes a search bar, a taskbar, and a status bar.

AWS Services

The screenshot shows the AWS Management Console with the URL <https://us-east-2.console.aws.amazon.com/elasticbeanstalk/home?region=us-east-2#/welcome>. The interface includes a top bar with 'Presenting...', 'Give control', 'Stop presenting', and 'Talk Manager' buttons. Below the bar is a navigation bar with links to 'Mail - KAVIYA.P - Online', 'YouTube Video Do...', 'Tamil Jothida Palan...', 'WhatsApp', 'Solve Python | Hack...', 'Microsoft Teams', 'Bloom's Revised Ta...', 'tancstrnic.in Stud...', and a 'Reading list'. The main content area has a dark background with white text. On the left, there are sections for 'Favorites' (with a note to add favorites by clicking stars) and 'Recently visited' (listing 'Elastic Beanstalk', 'Console Home', and 'AWS Marketplace Subscriptions'). The central part is titled 'All services' and lists various AWS services under categories: Compute (EC2, Lightsail, Lambda, Batch, Elastic Beanstalk, Serverless Application Rep..., AWS Outposts, EC2 Image Builder, AWS App Runner), Developer Tools (CodeStar, CodeCommit, CodeArtifact, CodeBuild, CodeDeploy, CodePipeline, Cloud9, CloudShell, X-Ray, AWS FIS), Machine Learning (Amazon SageMaker, Amazon Augmented AI, Amazon CodeGuru, Amazon DevOps Guru, Amazon Comprehend, Amazon Forecast, Amazon Fraud Detector, Amazon Kendra, Amazon Lex, Amazon Personalize, Amazon Polly), AWS Cost Management (AWS Cost Explorer, AWS Budgets, AWS Marketplace Subscript..., AWS Application Cost Pro...), Front-end Web & Mobile (AWS Amplify, Mobile Hub, AWS AppSync, Device Farm, Amazon Location Service), AR & VR (Amazon Sumerian), and Application Integration (Step Functions, Amazon AppFlow). A search bar at the top right says 'Search for services, features, marketplace products, and docs [Alt+S]'. At the bottom, there are links for 'Feedback', 'English (US)', 'Privacy Policy', 'Terms of Use', and 'Cookie preferences'. The status bar at the bottom right shows '12:07 PM 8/21/2021'.

AWS Services → All Services → Compute → EC2

This screenshot is identical to the one above, showing the AWS Management Console homepage. However, the 'Compute' category in the 'All services' section is highlighted with a blue border, and the 'EC2' service is also highlighted with a blue border, indicating it is the selected service. The rest of the interface, including the sidebar, search bar, and footer, remains the same.

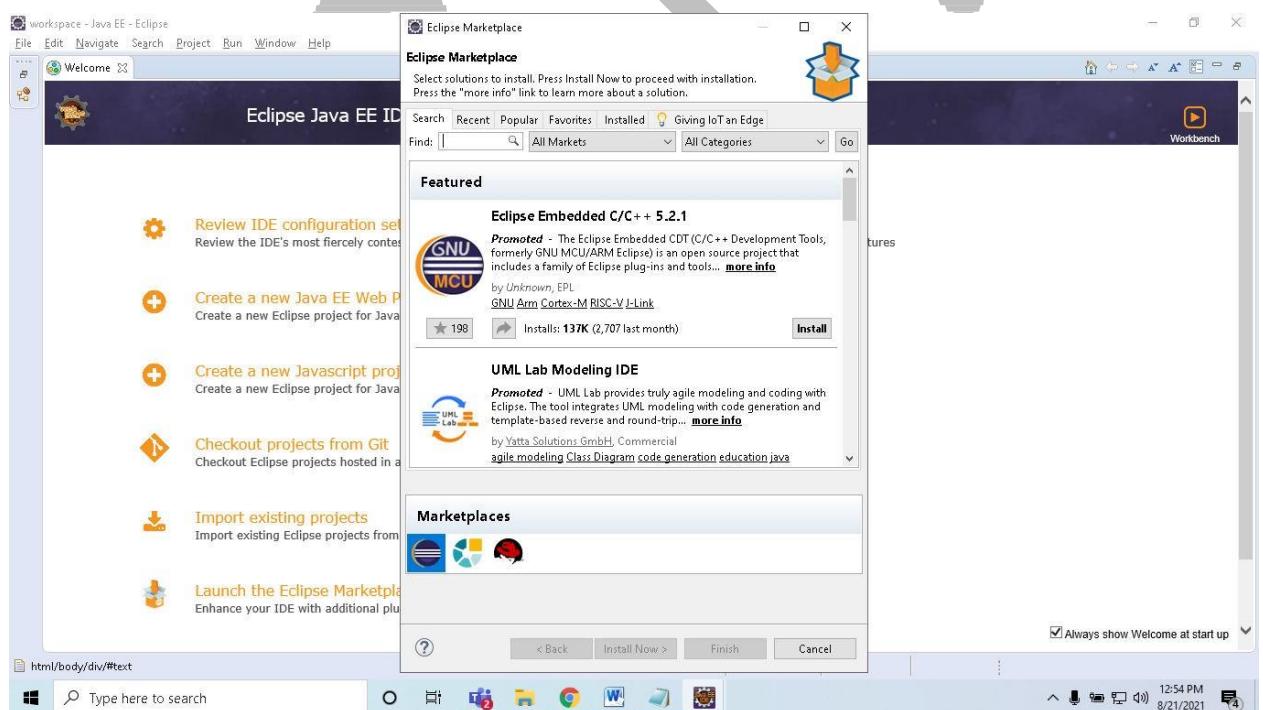
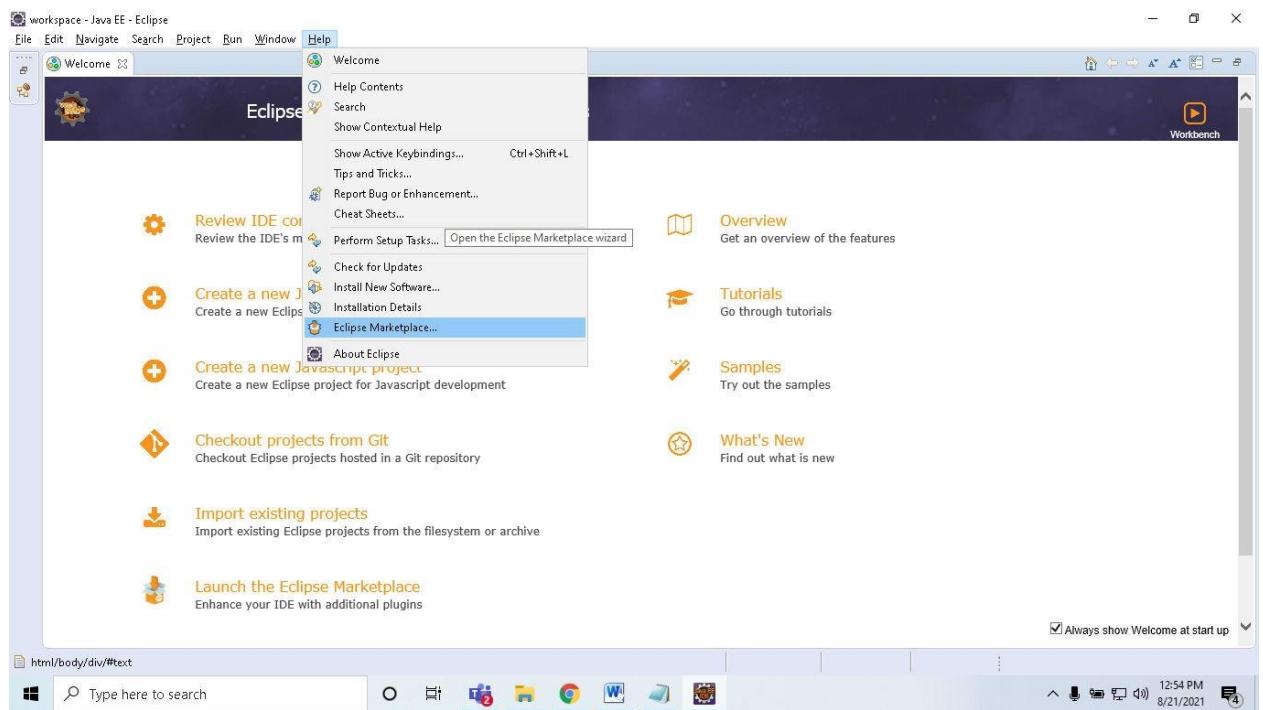
Initially “no instance” is running

The screenshot shows the AWS EC2 Dashboard. On the left, there's a sidebar with options like EC2 Dashboard, Instances, and Images. The main area is titled "Resources" and displays a table of EC2 resources. The table shows 0 instances running, 0 elastic IPs, 0 key pairs, 0 placement groups, 0 snapshots, 0 dedicated hosts, 0 instances, 0 load balancers, 1 security group, and 0 volumes. A tooltip message at the bottom of the table says: "Easily size, configure, and deploy Microsoft SQL Server Always On availability groups on AWS using the AWS Launch Wizard for SQL Server. Learn more". To the right, there's a panel titled "Account attributes" with sections for Supported platforms (VPC), Default VPC (vpc-c83b59a3), Settings, EBS encryption, Zones, EC2 Serial Console, Default credit specification, and Console experiments. At the bottom, there's an "Explore AWS" section and a search bar.

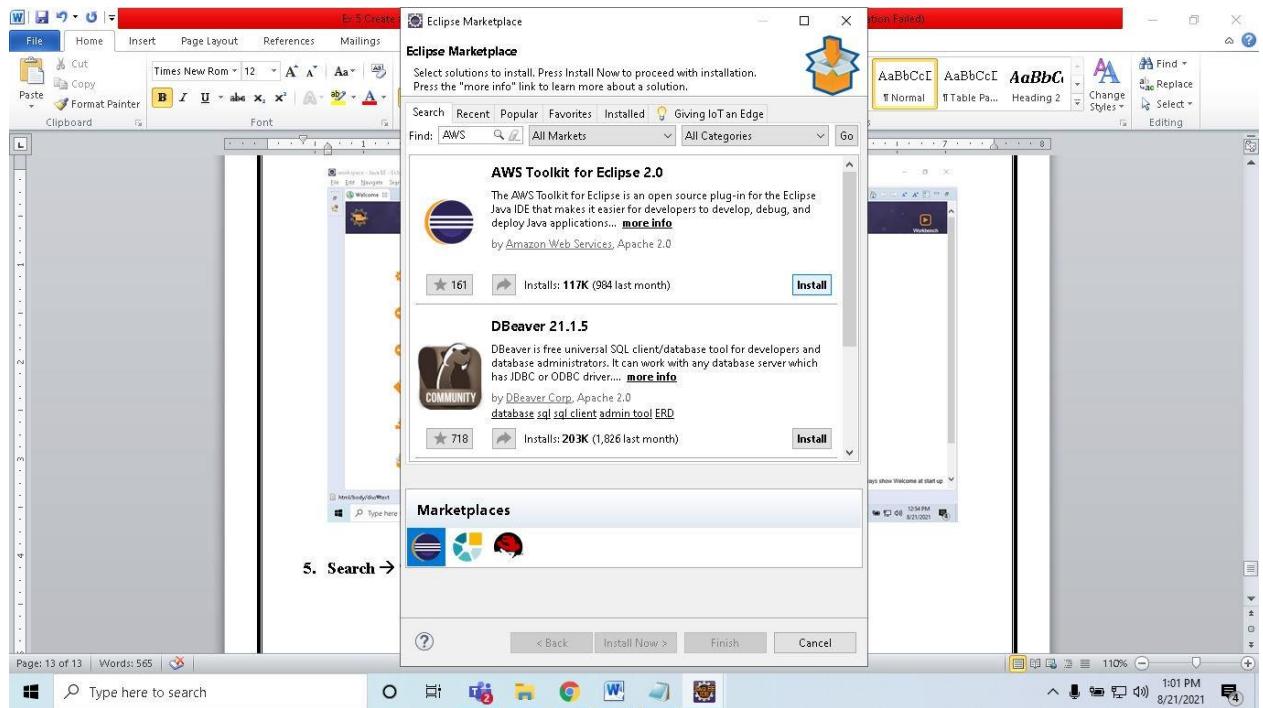
3. Open Eclipse IDE

The screenshot shows the Eclipse Java EE IDE for Web Developers. The window title is "workspace - Java EE - Eclipse". The interface includes a "Welcome" view with several items: "Review IDE configuration settings", "Create a new Java EE Web Project", "Create a new Javascript project", "Checkout projects from Git", "Import existing projects", and "Launch the Eclipse Marketplace". There are also links for "Overview", "Tutorials", "Samples", and "What's New". At the bottom of the "Welcome" view, there's a checkbox for "Always show Welcome at start up". The status bar at the bottom shows the URL "html/body/div/#text" and the date/time "12:52 PM 8/21/2021".

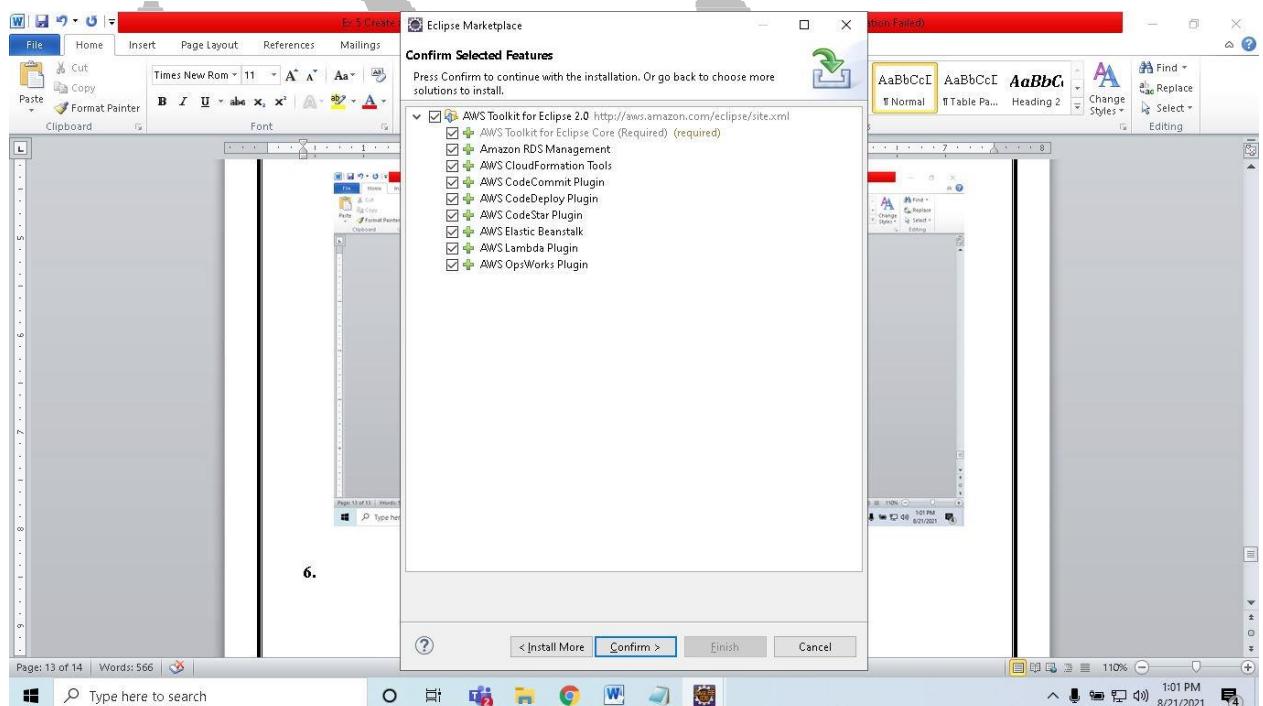
3.1 Go to “Help” option → Eclipse Marketplace



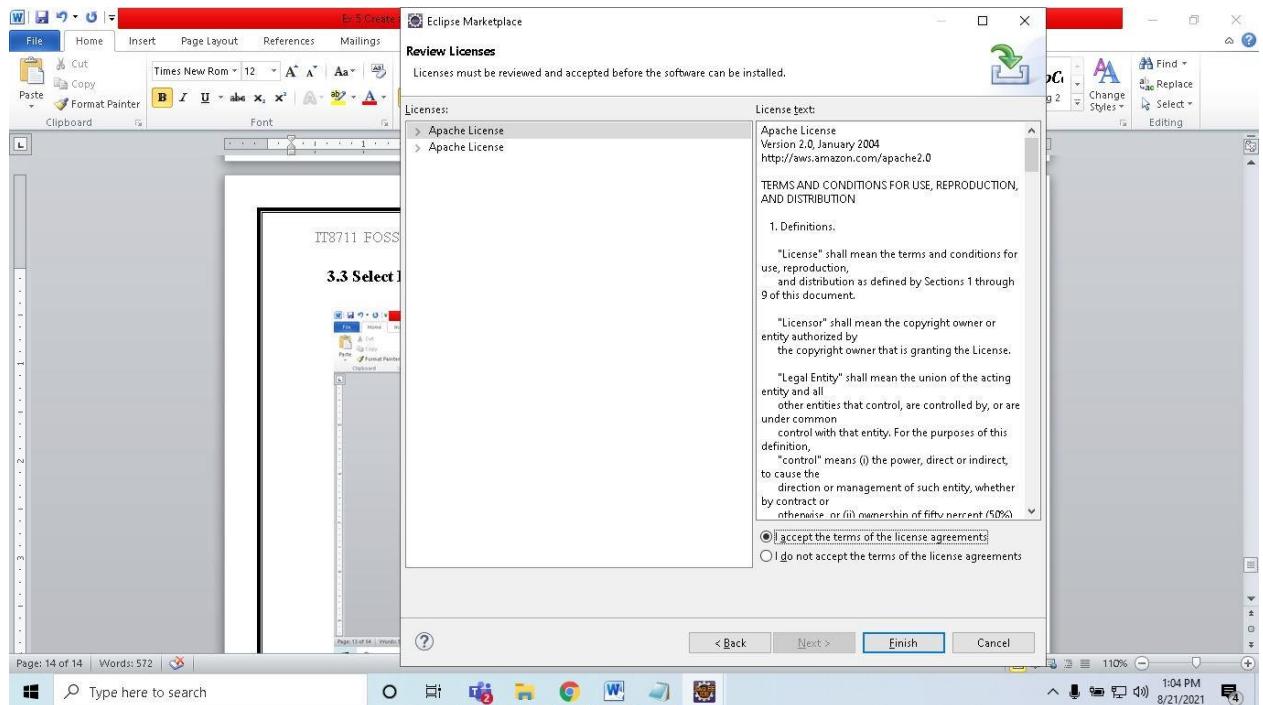
3.2 Search → “AWS Toolkit for Eclipse 2.0” → Click “Install”



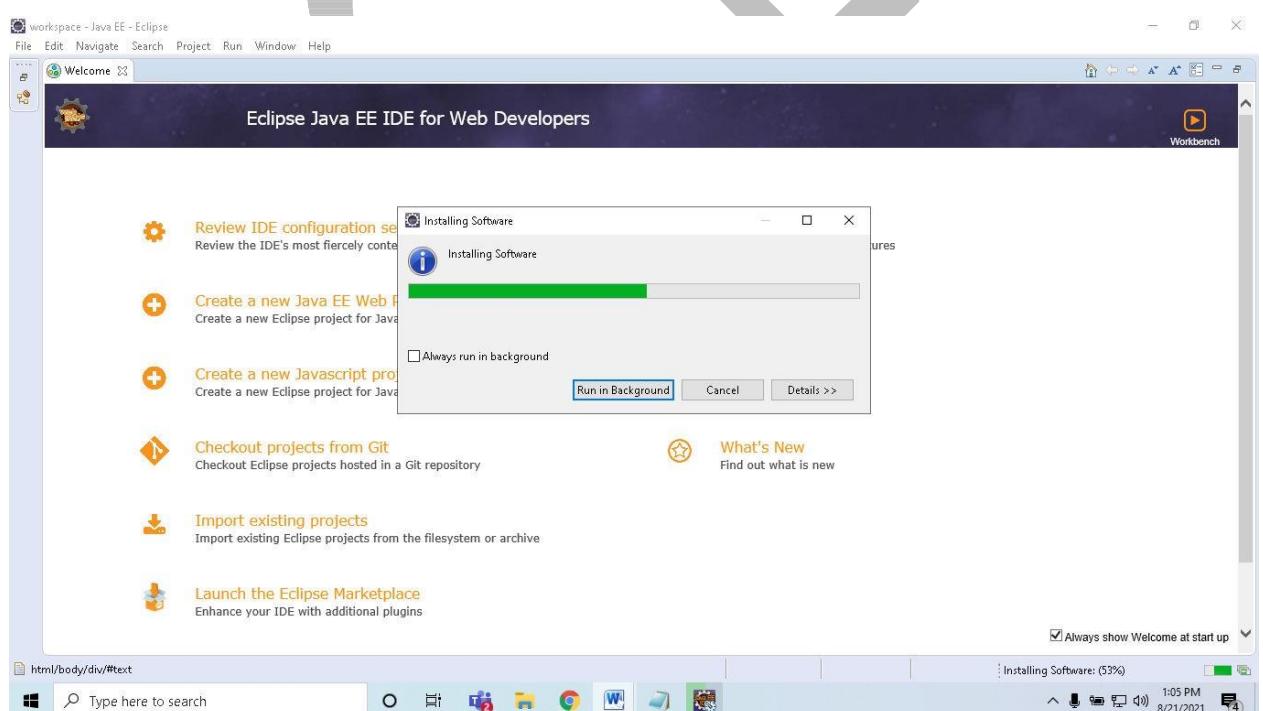
3.3 Select Features → Click “Confirm”



3.4 Review Licenses → Select “Accept the terms” → Click “Finish”

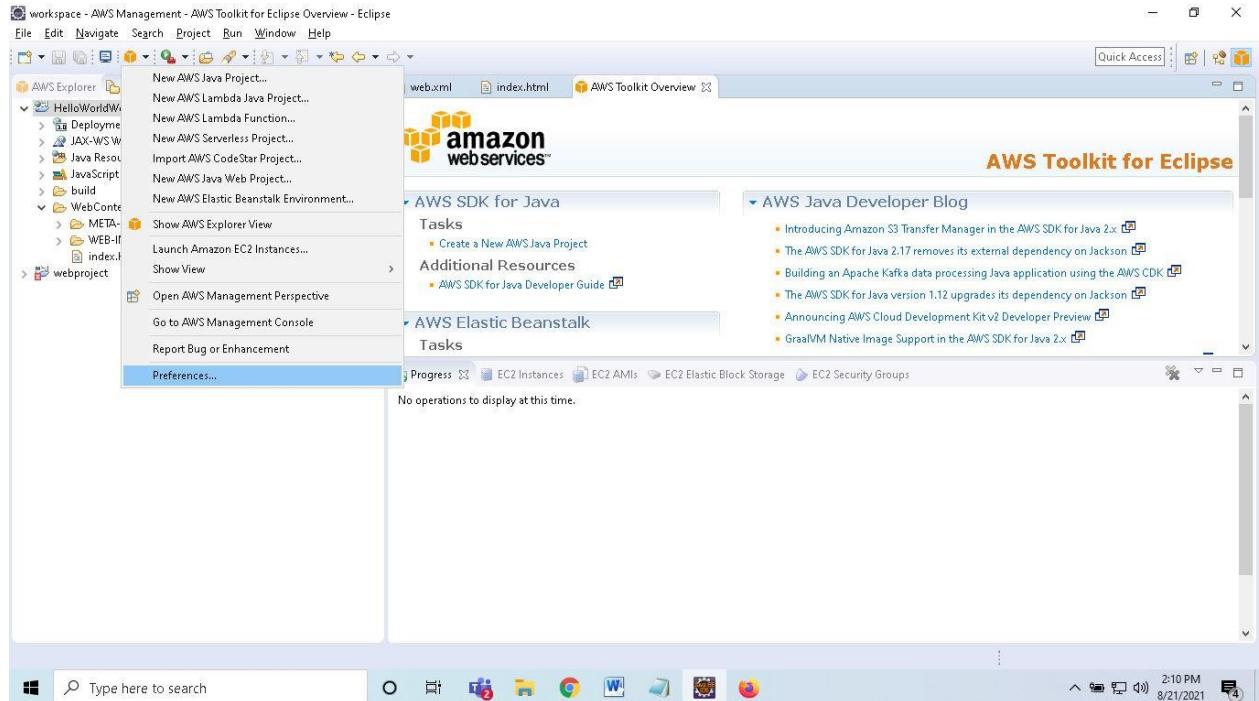


3.5 Installing Software

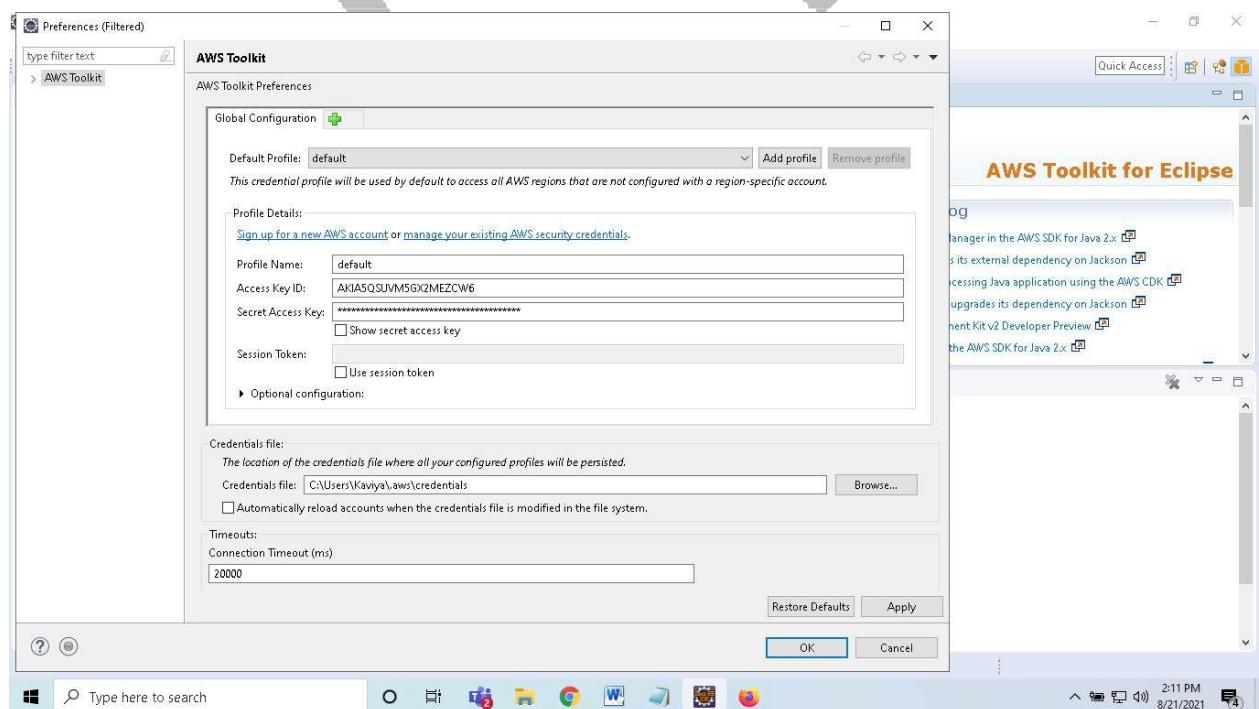


4. Connect Eclipse with AWS Account

4.1 Click “AWS” icon → Preferences



4.2 Click “manage your existing AWS security Credentials”



4.3 IAM Management Console is open in browser → Click “Access keys”

The screenshot shows the AWS IAM Management Console. The left sidebar has 'Identity and Access Management (IAM)' selected. Under 'Access management', 'Access keys (access key ID and secret access key)' is expanded. A table lists one access key:

Created	Access Key ID	Last Used	Last Used Region	Last Used Service	Status	Actions
Aug 17th 2021	AKIA6QSUVM5GX2MEZCW6	2021-08-21 14:13 UTC+0530	us-west-2	elasticbeanstalk	Active	Make Inactive Delete

A button labeled 'Create New Access Key' is visible. A note below the table says: 'Root user access keys provide unrestricted access to your entire AWS account. If you need long-term access keys, we recommend creating a new IAM user with limited permissions and generating access keys for that user instead.' A tooltip provides more information about root user access keys.

4.4 Click “Create New Access Key” → Download Key File

The screenshot shows the same AWS IAM Management Console as above, but with a 'Create Access Key' dialog box open over the 'Access keys' table. The dialog contains the following text:

Your access key (access key ID and secret access key) has been created successfully.
Download your key file now, which contains your new access key ID and secret access key. If you do not download the key file now, you will not be able to retrieve your secret access key again.

To help protect your security, store your secret access key securely and do not share it.

▶ Show Access Key [Download Key File](#) [Close](#)

A note at the bottom of the dialog says: 'Root user access keys provide unrestricted access to your entire AWS account. If you need long-term access keys, we recommend creating a new IAM user with limited permissions and generating access keys for that user instead.' A tooltip provides more information about root user access keys.

The screenshot shows the AWS IAM Management Console interface. On the left, there's a navigation sidebar with options like Dashboard, Access management, and Access reports. The main area displays a section titled "Multi-factor authentication (MFA)" and "Access keys (access keys)". A modal dialog box is open in the center, titled "Opening rootkey.csv". It contains the following text:
You have chosen to open:
rootkey.csv
which is: Text Document (90 bytes)
from: https://console.aws.amazon.com

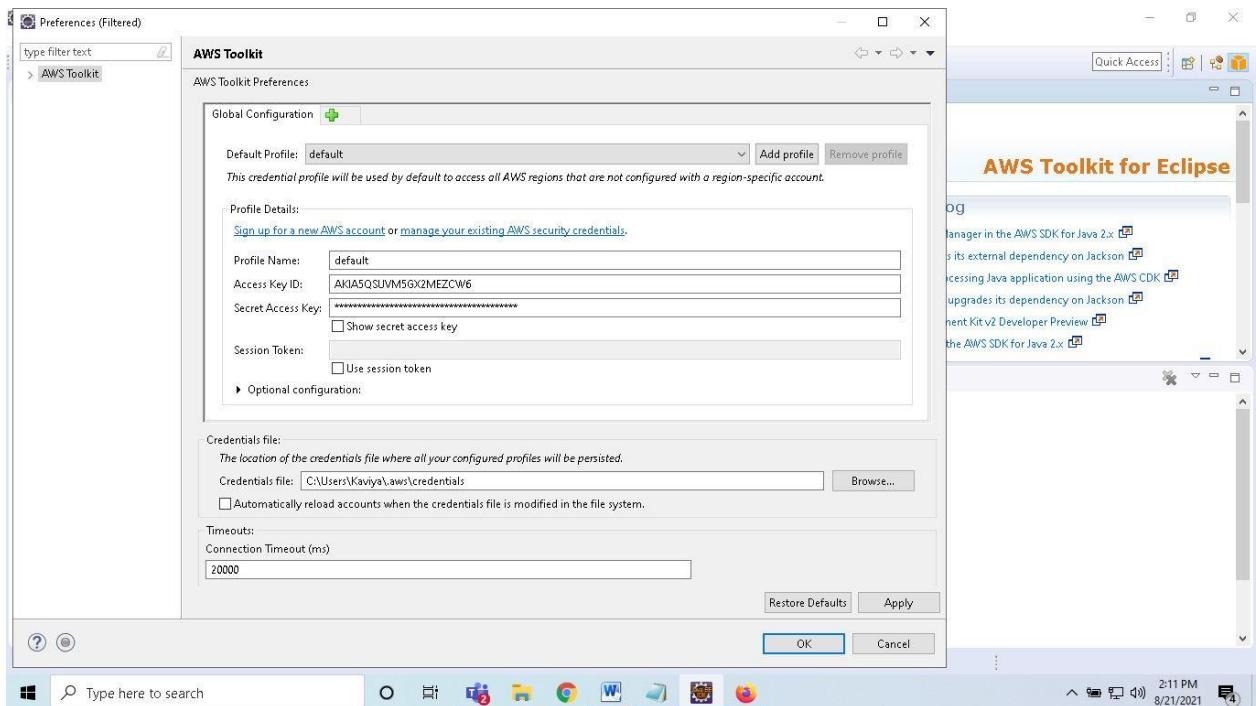
What should Firefox do with this file?
 Your access key
Download your key if you do not download again.
 Open with Notepad (default)
 Save File
 Do this automatically for files like this from now on.

At the bottom of the modal, there are "OK" and "Cancel" buttons. In the background, the IAM service page shows a table with columns "Used", "Service", "Status", and "Actions". One row is highlighted in green, showing "awsbatch" as the service, "Active" as the status, and "Make Inactive | Delete" as the actions. The status bar at the bottom right shows "IT8711 FOSS & Cloud Com... 17:58".

4.5 Open “rootkey.csv”

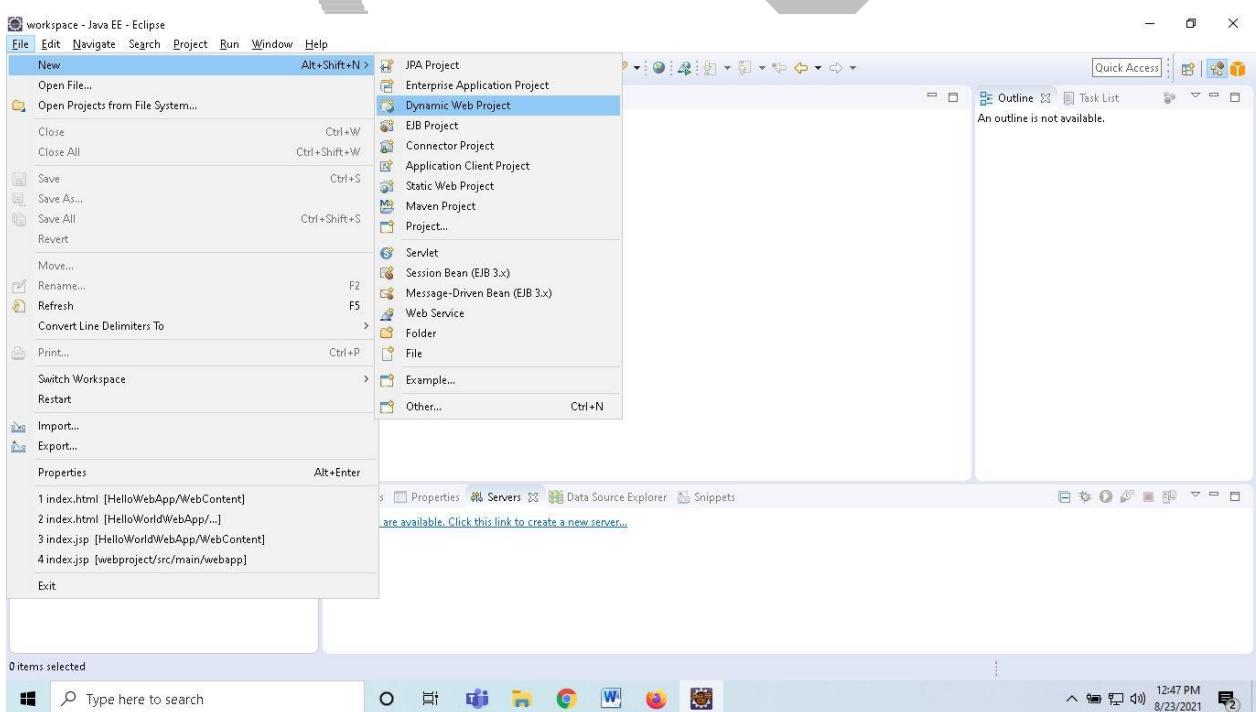
The screenshot shows a Microsoft Excel spreadsheet titled "rootkey(1) - Microsoft Excel (Product Activation Failed)". The spreadsheet has a single sheet with data starting from cell D5. The data consists of two rows of text:
1 AWSAccessKeyId=AKIA5QSUVM5G5J3LHA7Y
2 AWSSecretKey=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

4.6 Paste the AWS Access Key ID & Secret Access Key in AWS Toolkit Preferences from “rootkey.csv”

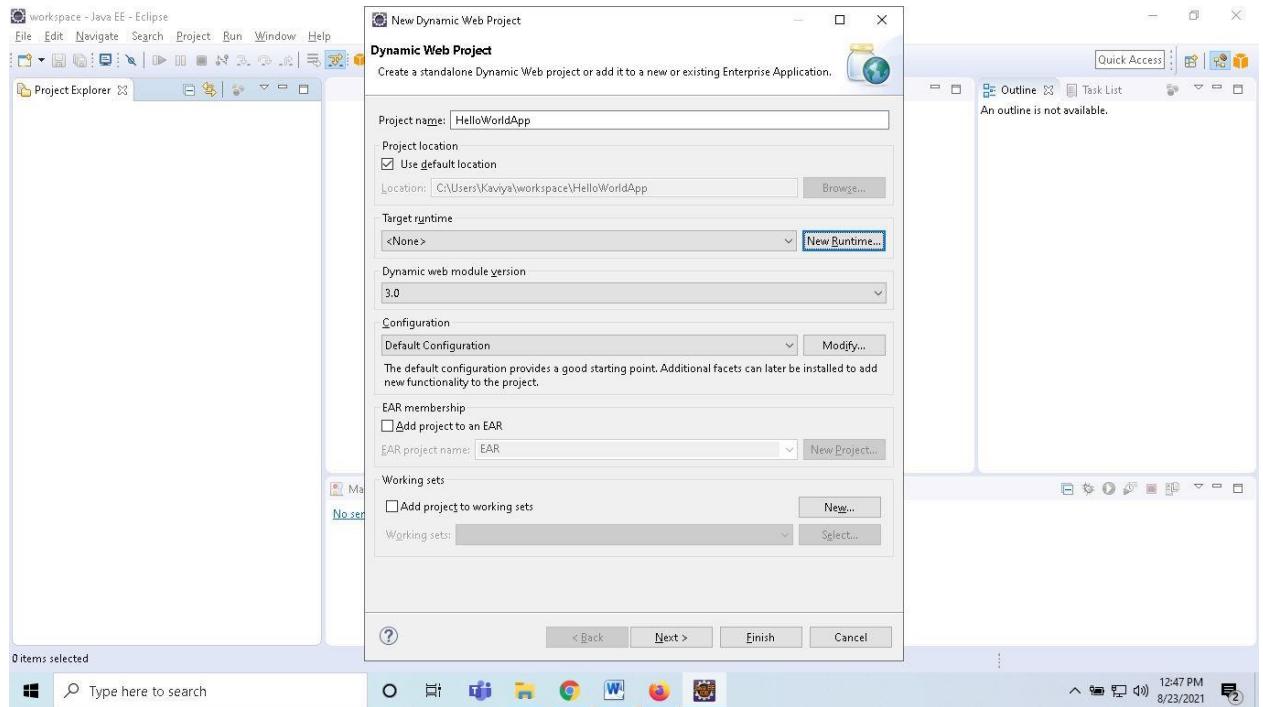


5. Develop Java Web Application in Eclipse IDE

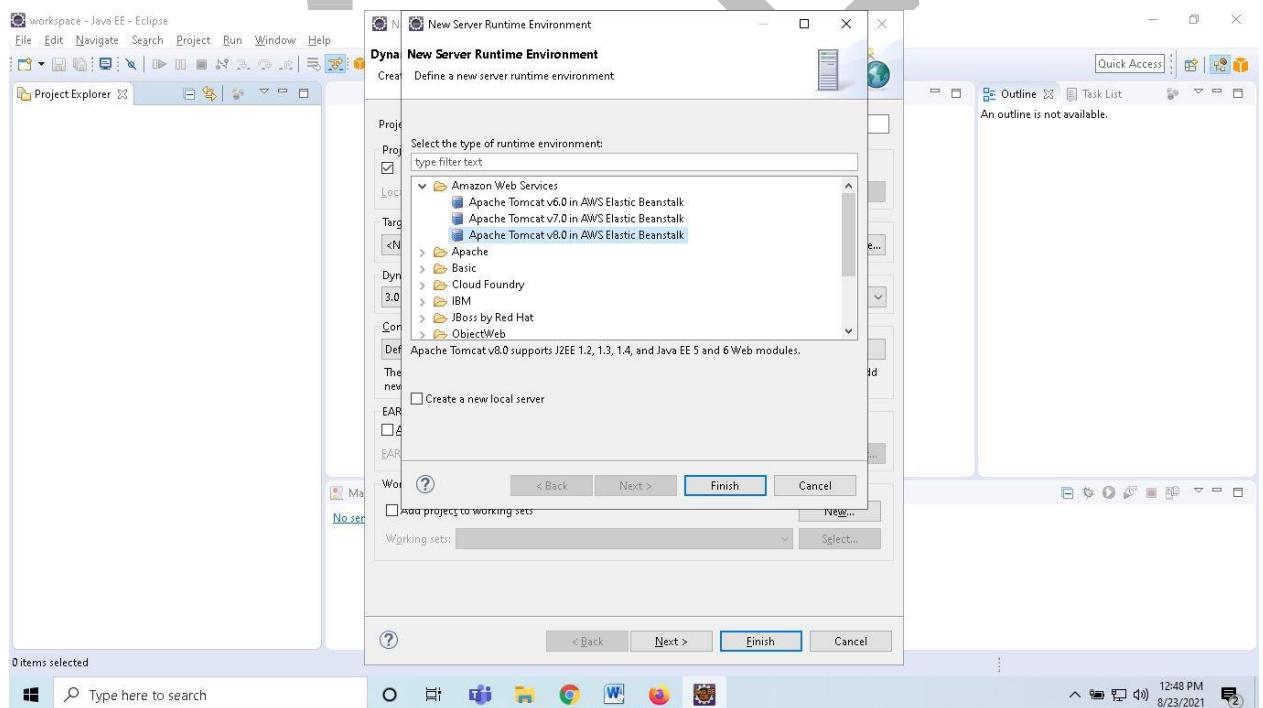
5.1 File → New → Dynamic Web Project



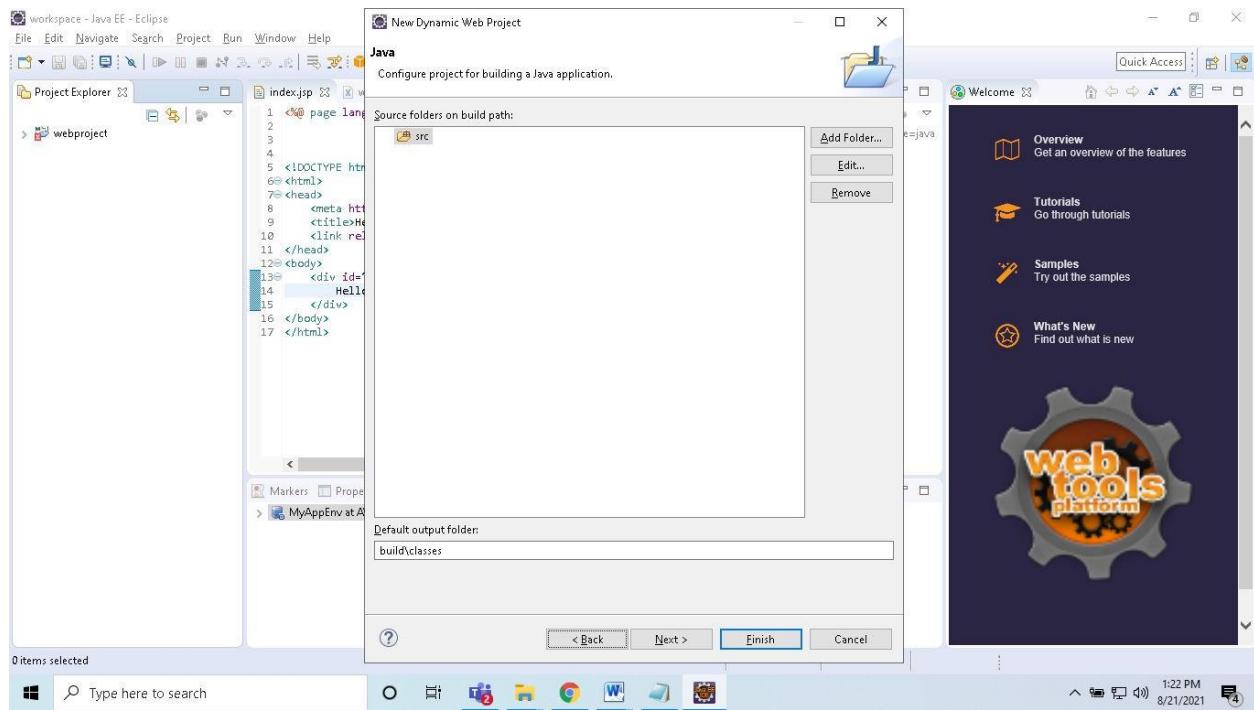
5.2 Give Project Name “HelloWorldApp”



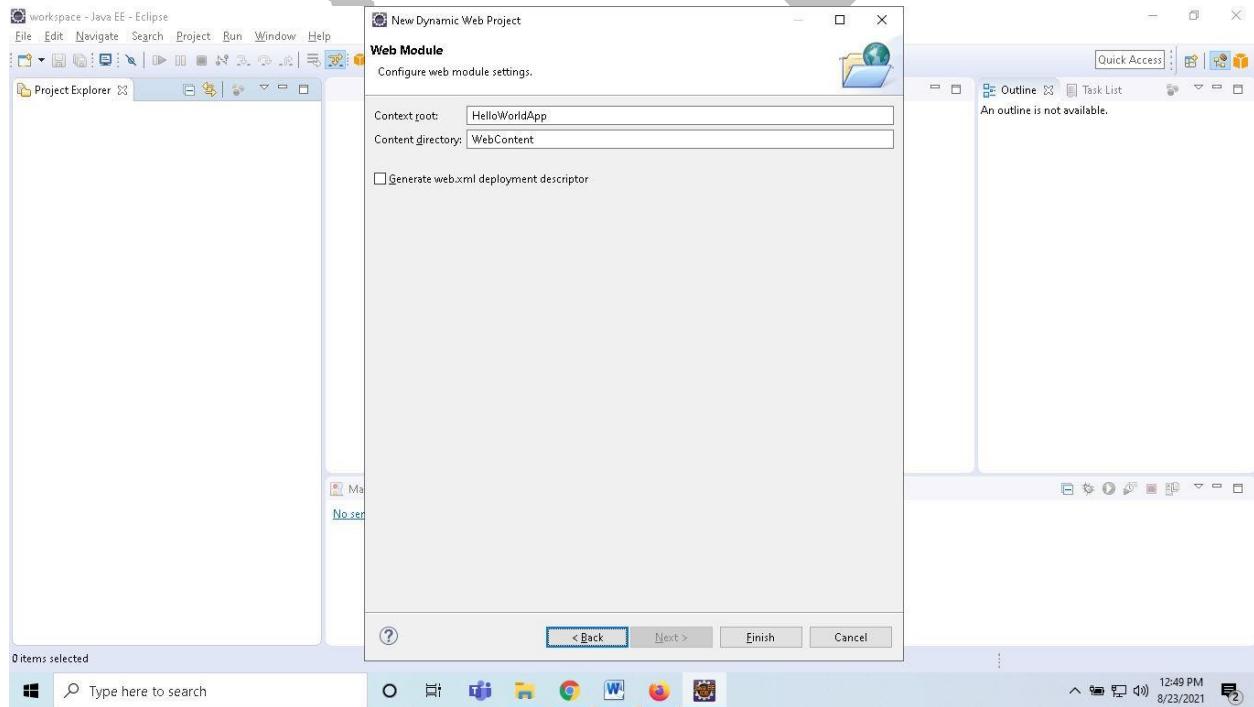
5.3 Choose Target Runtime → Choose “Apache Tomcat v8.0 in AWS Beanstalk” → Click “Finish” → Click “Next”



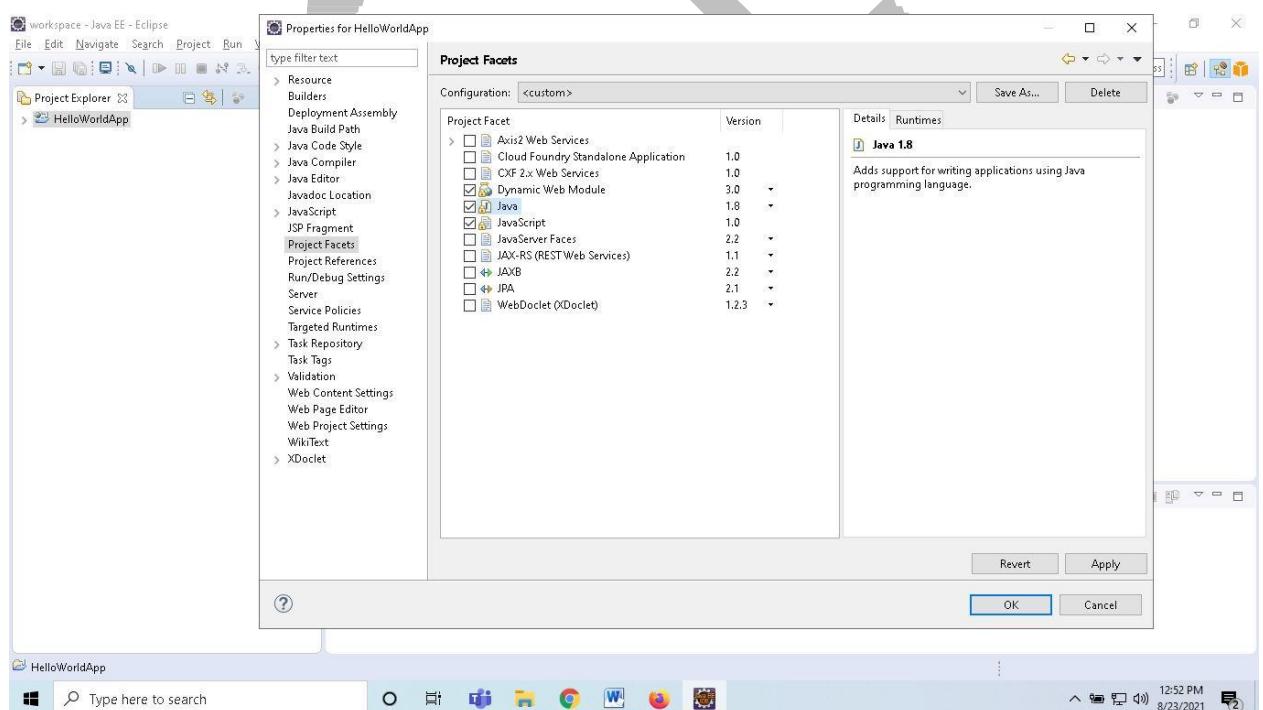
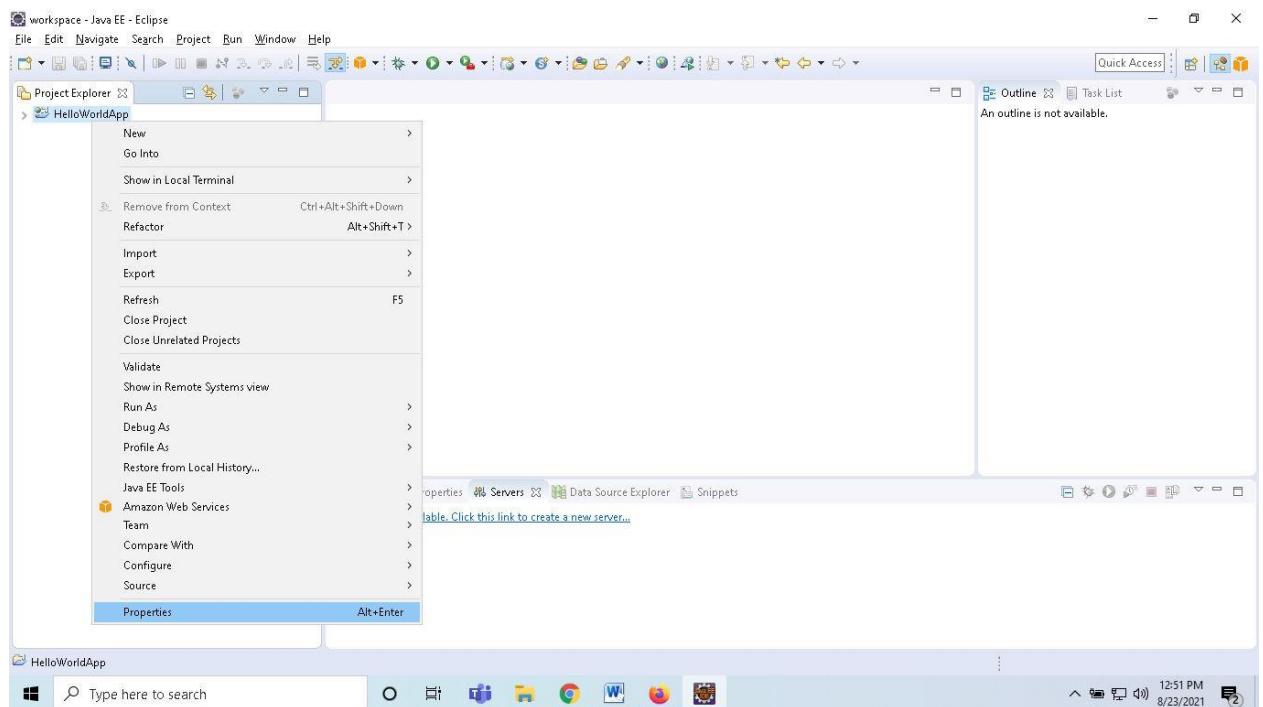
5.4 Configure project for building a Java application → Click “Next”



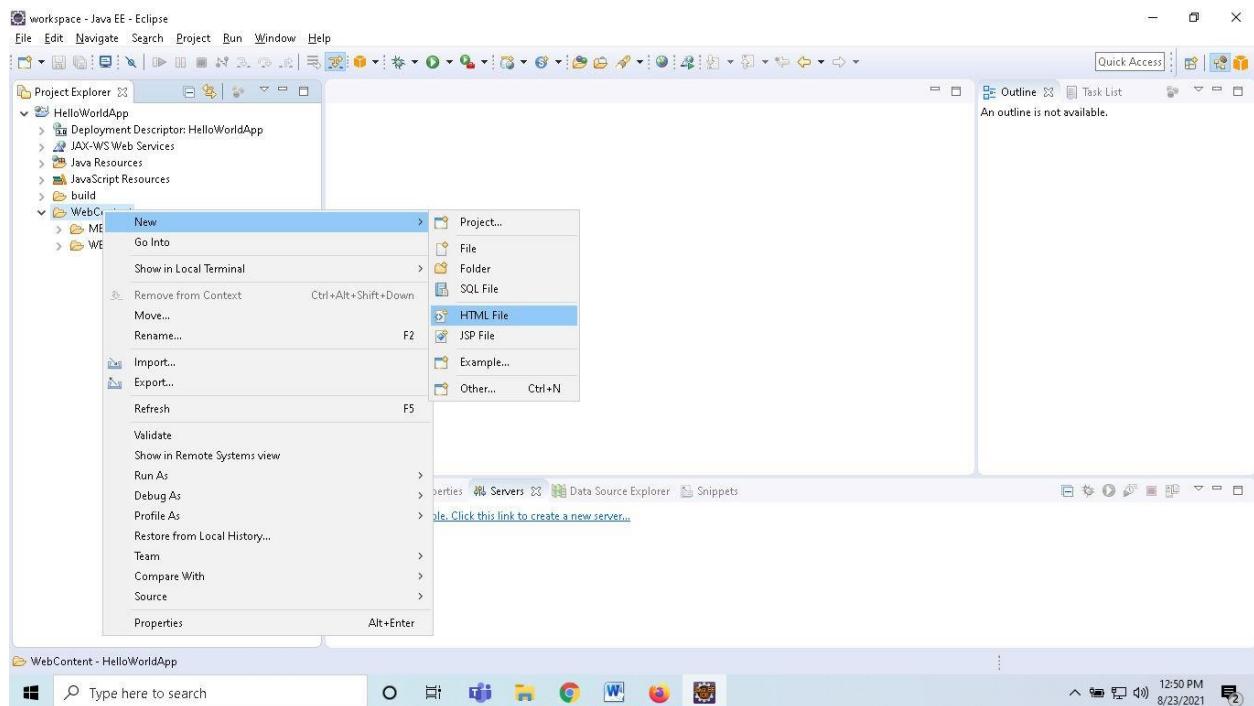
5.5 Configure web module settings → Click “Finish”



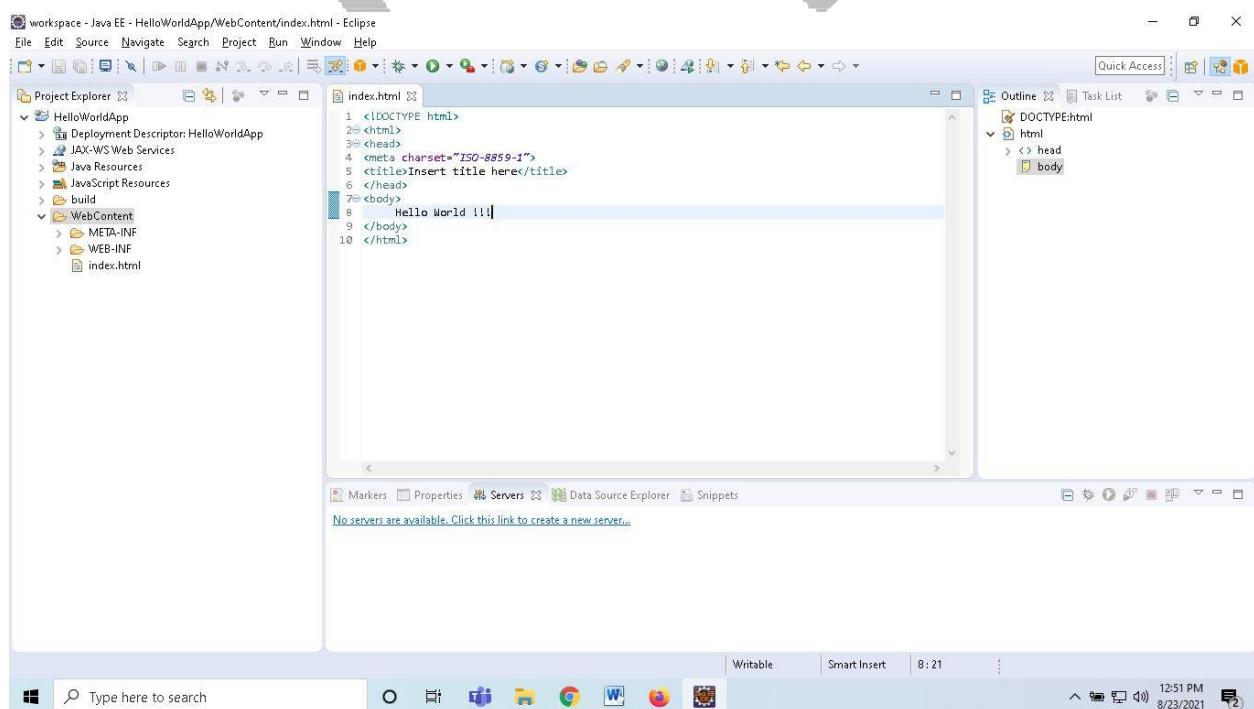
5.6 Check Project Facets → Dynamic Web Module & Java Version



5.7 Create a File “index.html”

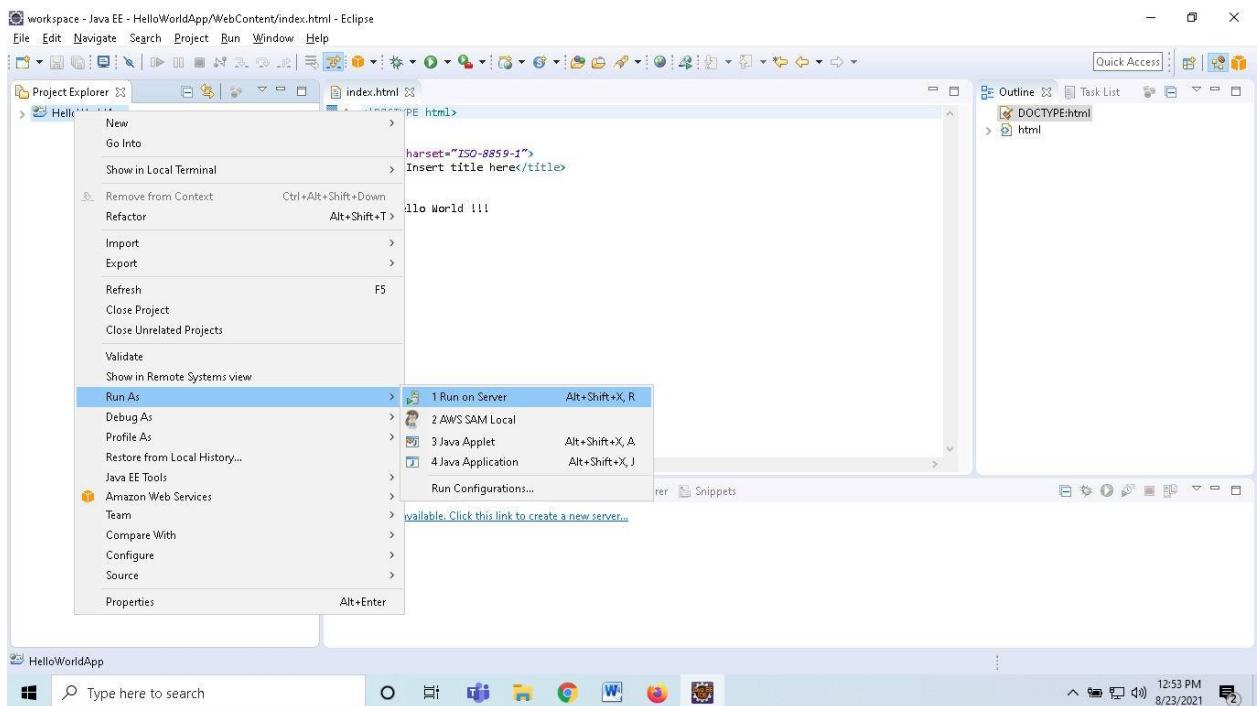


5.8 Add a code in “index.html”



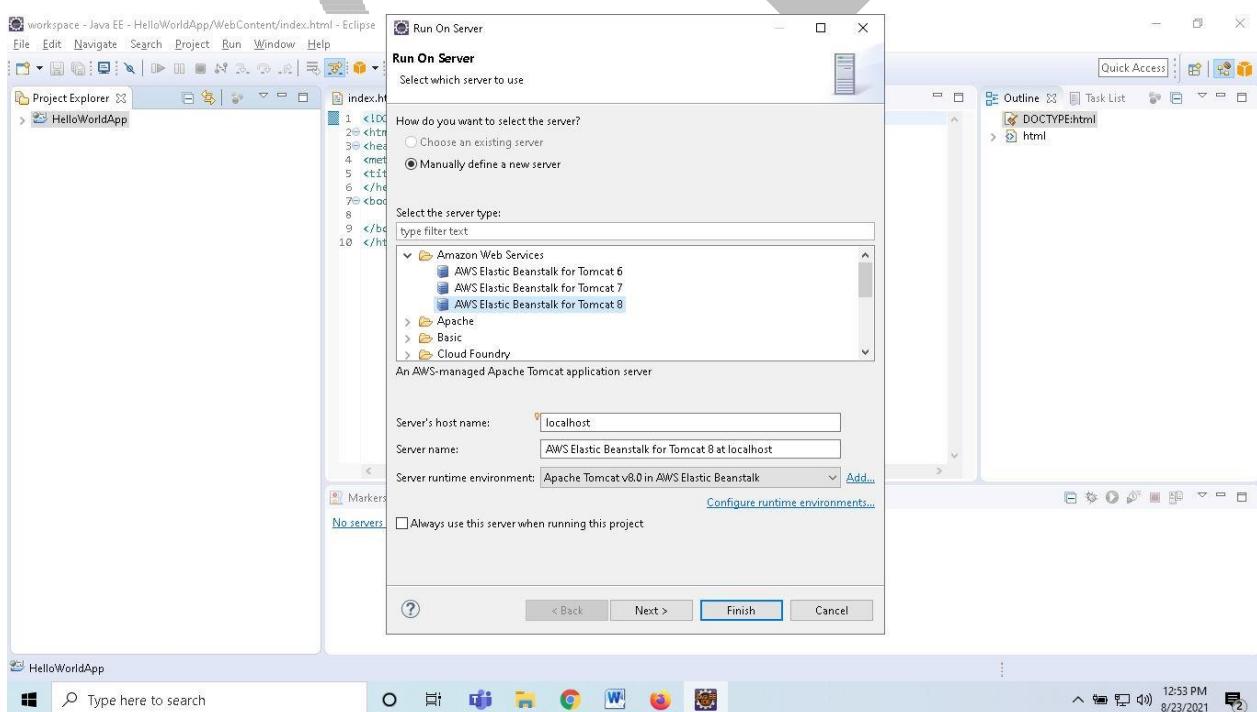
6. Deploy the application in AWS

6.1 Run the application → Choose “Run As” → Choose “Run on Server”

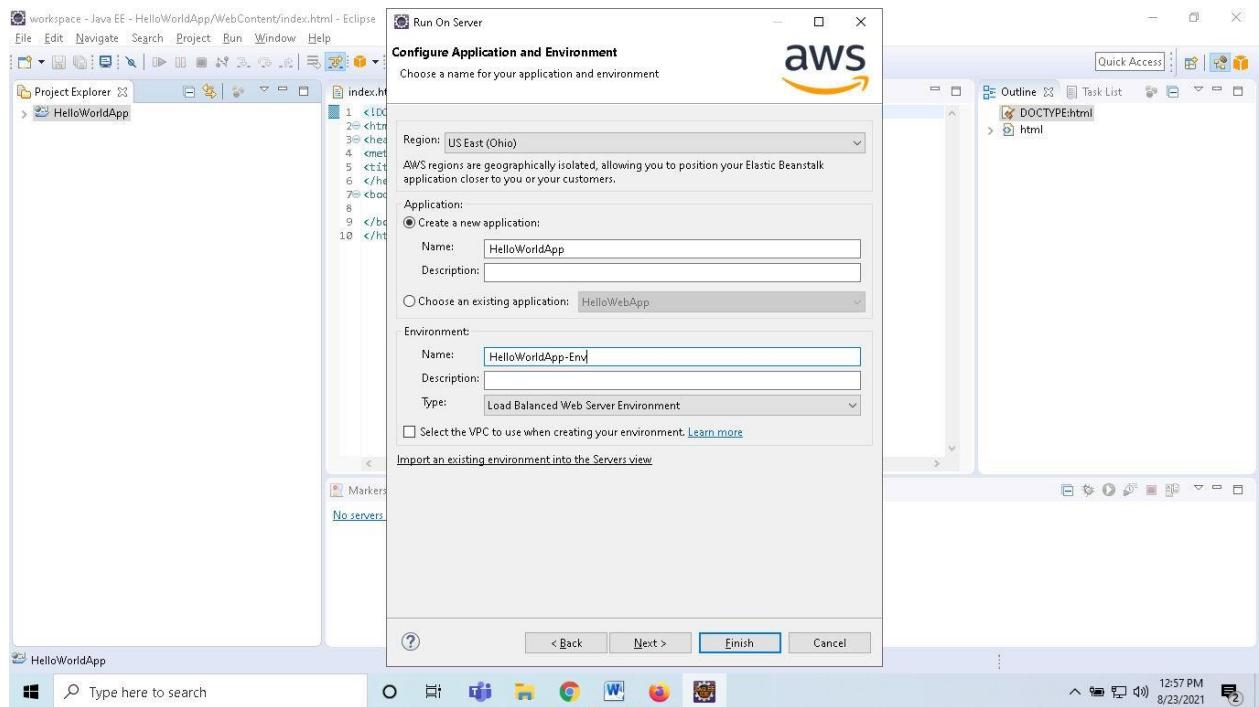


6.2 Run on Server → Select the server type as “AWS Elastic Beanstalk for Tomcat 8”

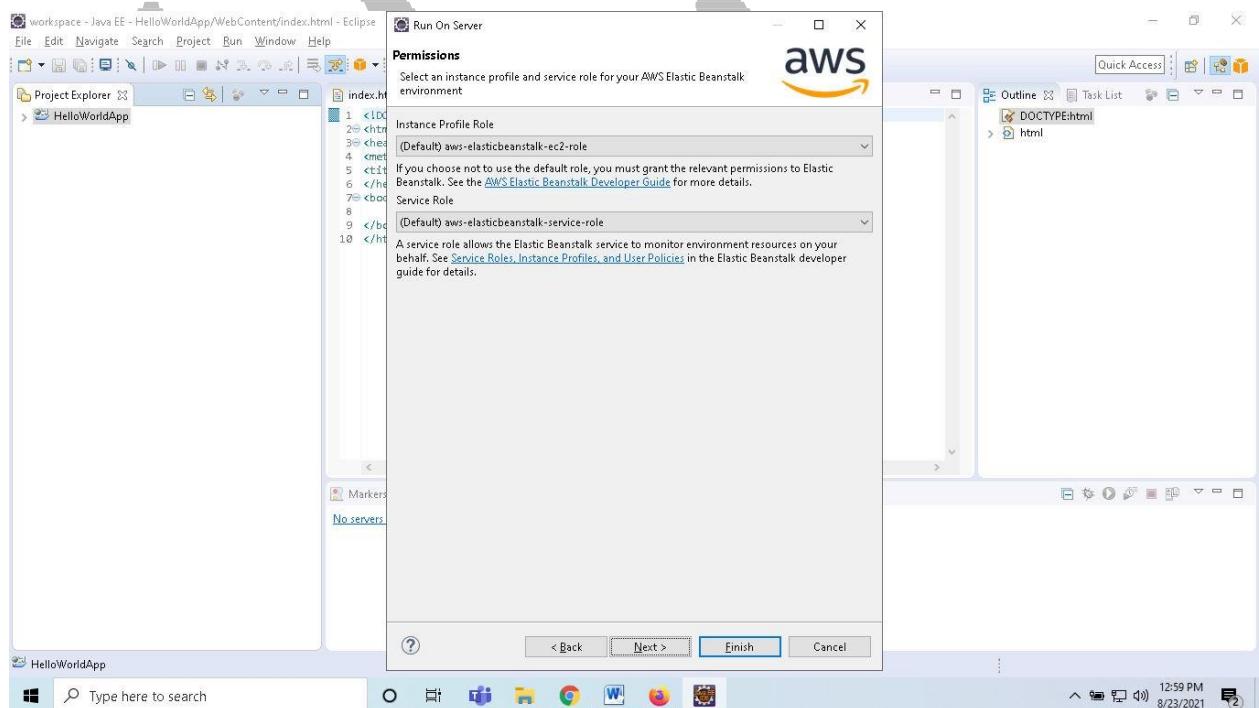
→ Click “Next”



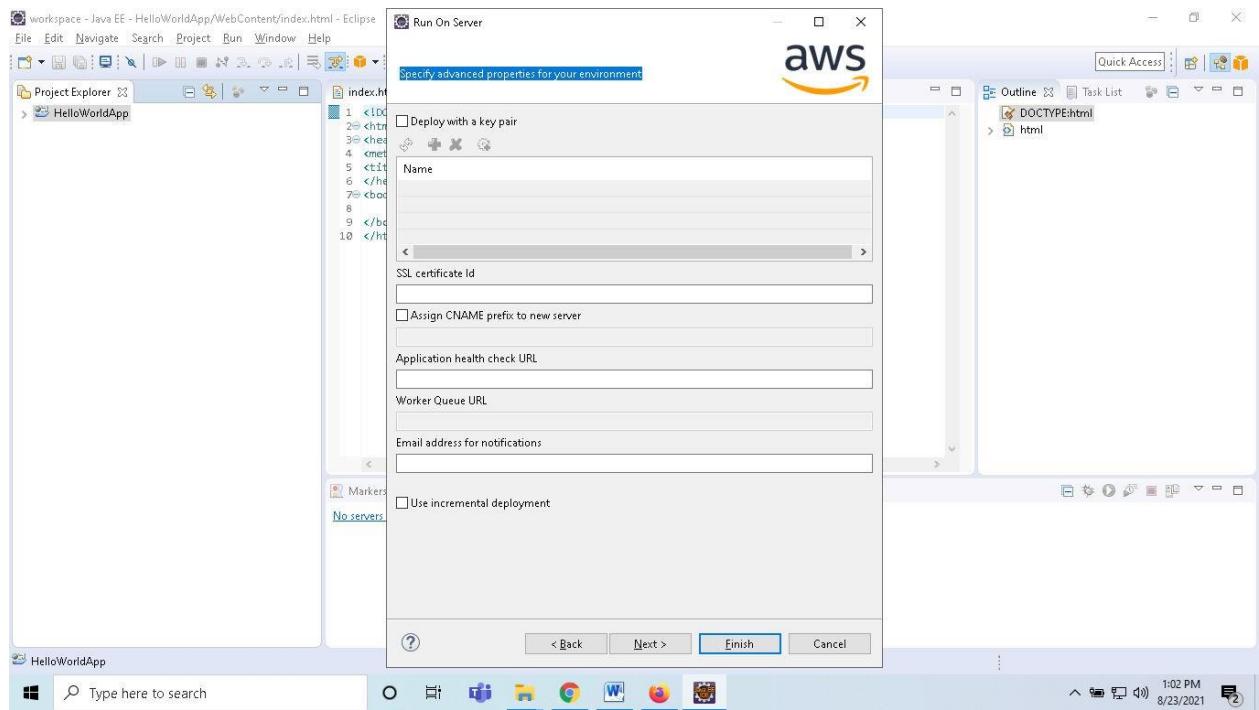
6.3 Configure Application and Environment → Give Region, Application and Environment details



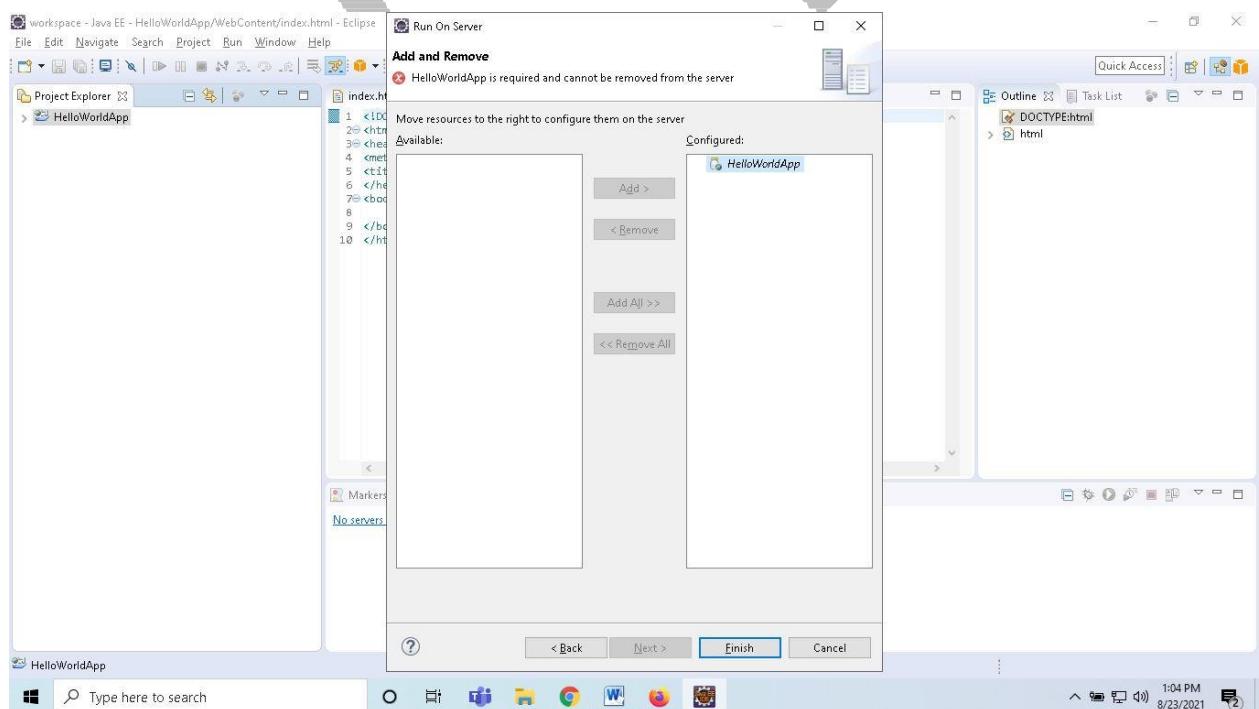
6.4 Permission → Click "Next"



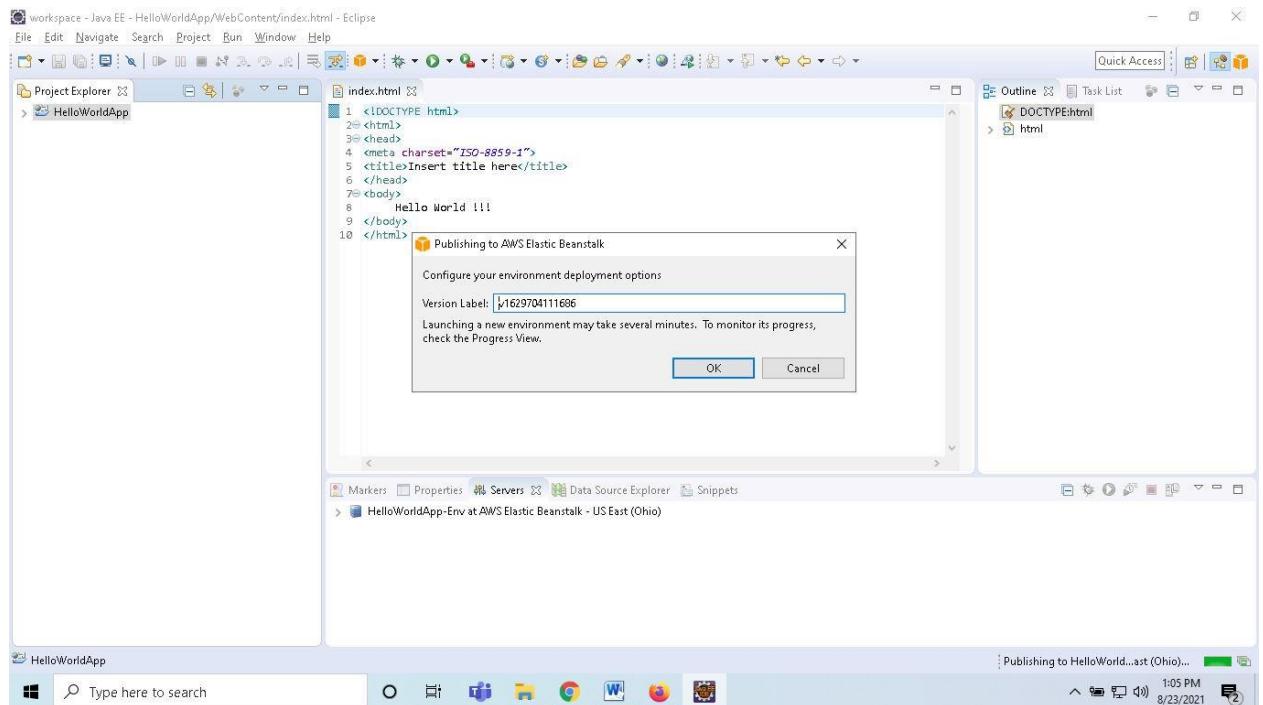
6.5 Specify advanced properties for your environment → Click “Next”



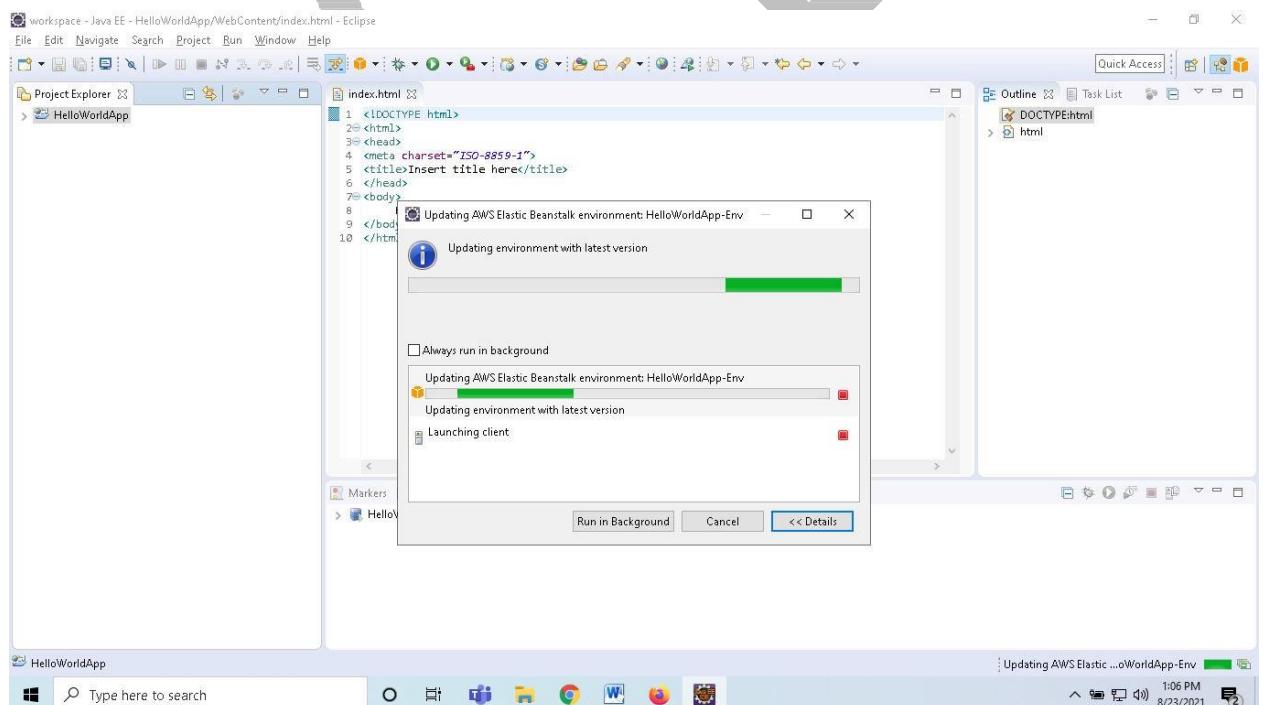
6.6 Add and Remove → Move resources to the right to configure them on the server → Click “Finish”



6.7 Publishing to AWS Elastic Beanstalk → Click “OK”



6.8 Updating AWS Elastic Beanstalk Environment



6.9 In AWS, Environment Creation is started [Health → Pending state]

The screenshot shows the AWS Elastic Beanstalk console interface. The top navigation bar includes tabs for 'Elastic Beanstalk Environments' and 'Dashboard | EC2 Management'. The main title is 'Elastic Beanstalk' under 'Environments'. A prominent alert message states: 'Deprecated/Retired Platform(s) One or more of your environments is configured with a deprecated or retired platform. Learn more'.

The central area displays a table titled 'All environments' with the following columns: Environment name, Health, Application name, Date created, Last modified, URL, and Running versions. A single row is listed:

Environment name	Health	Application name	Date created	Last modified	URL	Running versions
HelloWorldApp-Env	Pending	HelloWorldApp	2021-08-23 13:06:32 UTC+0530	2021-08-23 13:06:52 UTC+0530	HelloWorldApp-Env.eba-alzpvfai.us-east-2.elasticbeanstalk.com	-

Below this, another browser window shows the 'Create environment' process for the 'HelloWorldApp-Env'. The title bar says 'Creating HelloWorldApp-Env This will take a few minutes...'. The log output shows the following steps:

```
1:07pm Created Auto Scaling launch configuration named: awseb-e-7vunt5w2ms-stack-AWSEBAutoScalingLaunchConfiguration-U9XVPC6NRSYB
1:07pm Created security group named: awseb-e-7vunt5w2ms-stack-AWSEBSecurityGroup-RDQHD723LWXU
1:07pm Environment health has transitioned to Pending. Initialization in progress (running for 19 seconds). There are no instances.
1:06pm Created load balancer named: awseb-e-7-AWSEBLos-JEXRAWXQ5QB7
1:06pm Created security group named: sg-0d4e3d58447bf991
1:06pm Using elasticbeanstalk-us-east-2-928964241229 as Amazon S3 storage bucket for environment data.
1:06pm createEnvironment is starting.
```

The bottom of the screen shows a Windows taskbar with various pinned icons and the system clock indicating 1:06 PM on 8/23/2021.

6.10 In AWS, Environment is created successfully [Health → OK state]

The screenshot shows the AWS Elastic Beanstalk Environments page. On the left, there's a sidebar with 'Elastic Beanstalk' and tabs for 'Environments' (which is selected), 'Applications', and 'Change history'. Below that is a 'Recent environments' section with 'HelloWorldApp-Env'. The main content area has a heading 'All environments' with a search bar and a table. The table columns are 'Environment name', 'Health', 'Application name', 'Date created', 'Last modified', 'URL', and 'Running versions'. One row is visible for 'HelloWorldApp-Env', which is in the 'Ok' state. At the top of the main content area, there's a warning box about deprecated/retired platforms.

Environment name	Health	Application name	Date created	Last modified	URL	Running versions
HelloWorldApp-Env	Ok	HelloWorldApp	2021-08-23 13:06:32 UTC+0530	2021-08-23 13:08:56 UTC+0530	HelloWorldApp-Env.eba-eizpvfai.us-east-2.elasticbeanstalk.com	v162970

6.11 Java Web Application “HelloWorldApp” is launched successfully

The screenshot shows the AWS Elastic Beanstalk Applications page. The sidebar is identical to the previous one, with 'Environments' selected. The main content area has a heading 'All applications' with a search bar and a table. The table columns are 'Application name', 'Environments', 'Date created', 'Last modified', and 'ARN'. Two rows are listed: 'HelloWebApp' and 'HelloWorldApp'. The 'HelloWorldApp' row shows it is associated with the 'HelloWorldApp-Env' environment.

Application name	Environments	Date created	Last modified	ARN
HelloWebApp		2021-08-21 15:08:07 UTC+0530	2021-08-21 15:08:07 UTC+0530	arn:aws:elasticbeanstalk:us-east-2:928964241229:application/HelloWebApp
HelloWorldApp	HelloWorldApp-Env	2021-08-23 13:06:28 UTC+0530	2021-08-23 13:06:28 UTC+0530	arn:aws:elasticbeanstalk:us-east-2:928964241229:application/HelloWorldApp

6.12 Services → Compute → EC2 Dashboard → Check “one instance” is running

The screenshot shows the AWS EC2 Dashboard. On the left, a sidebar lists options like New EC2 Experience, EC2 Dashboard, Instances, and Images. The main area displays a summary of resources: 1 Instance (running), 0 Dedicated Hosts, 0 Elastic IPs, 1 Instances, 0 Key pairs, 1 Load balancers, 0 Placement groups, 3 Security groups, 0 Snapshots, and 1 Volumes. A callout box highlights the "Instances (running)" section. To the right, the "Account attributes" pane shows supported platforms (VPC), default VPC (vpc-c83b59a3), and other settings. At the bottom, there's an "Explore AWS" section and a status bar indicating the date and time.

7. Java Web Application “HelloWorldApp” is launched in AWS

7.1 In AWS, Environment Creation is started [Health → Pending state]

The screenshot shows the AWS Elastic Beanstalk Environments page. The sidebar has an "Environments" tab selected. A warning message states: "Deprecated/Retired Platform(s) One or more of your environments is configured with a deprecated or retired platform. Learn more". The main table lists environments: "HelloWorldApp-Env" with a "Pending" health status, created on 2021-08-23 at 13:06:32 UTC+0530, and the URL HelloWorldApp-Env.eba-alzpvfai.us-east-2.elasticbeanstalk.com. The table has columns for Environment name, Health, Application name, Date created, Last modified, URL, and Running versions. A "Create a new environment" button is visible at the top right of the table area.

Elastic Beanstalk > Environments > HelloWorldApp-Env

Creating HelloWorldApp-Env
This will take a few minutes. ..

1:07pm Created Auto Scaling launch configuration named:
awseb-e-7vunt5w2ms-stack-AWSEBAutoScalingLaunchConfiguration-U9XVPC6NRSYB

1:07pm Created security group named:
awseb-e-7vunt5w2ms-stack-AWSEBSecurityGroup-RDQHD723LWXU

1:07pm Environment health has transitioned to Pending. Initialization in progress (running for 19 seconds). There are no instances.

1:06pm Created load balancer named:
awseb-e-7-AWSEBLba-JEXRAWXQ5QB7

1:06pm Created security group named:
sg-0d4e5d38447beff91

1:06pm Using elasticbeanstalk-us-east-2-928964241229 as Amazon S3 storage bucket for environment data.

1:06pm createEnvironment is starting.

7.2 In AWS, Environment is created successfully [Health → OK state]

Elastic Beanstalk > Environments

All environments

Environment name	Health	Application name	Date created	Last modified	URL	Running versions
HelloWorldApp-Env	OK	HelloWorldApp	2021-08-23 13:06:32 UTC+0530	2021-08-23 13:08:56 UTC+0530	HelloWorldApp-Env.eba-alzpvfai.us-east-2.elasticbeanstalk.com	v162970

7.3 Java Web Application “HelloWorldApp” is launched successfully

The screenshot shows the AWS Elastic Beanstalk Applications dashboard. On the left, a sidebar lists 'Environments' and 'Recent environments'. Under 'Recent environments', 'HelloWorldApp-Env' is listed. The main area is titled 'All applications' and displays a table with the following data:

Application name	Environments	Date created	Last modified	ARN
HelloWebApp		2021-08-21 15:08:07 UTC+0530	2021-08-21 15:08:07 UTC+0530	arn:aws:elasticbeanstalk:us-east-2:928964241229:application>HelloWebApp
HelloWorldApp	HelloWorldApp-Env	2021-08-23 13:06:28 UTC+0530	2021-08-23 13:06:28 UTC+0530	arn:aws:elasticbeanstalk:us-east-2:928964241229:application>HelloWorldApp

7.4 Services → Compute → EC2 Dashboard → Check “one instance” is running

The screenshot shows the AWS EC2 Dashboard. The left sidebar includes 'EC2 Dashboard' with sections for Events, Tags, Limits, Instances, Images, and AMIs. The main 'Resources' section displays the following data:

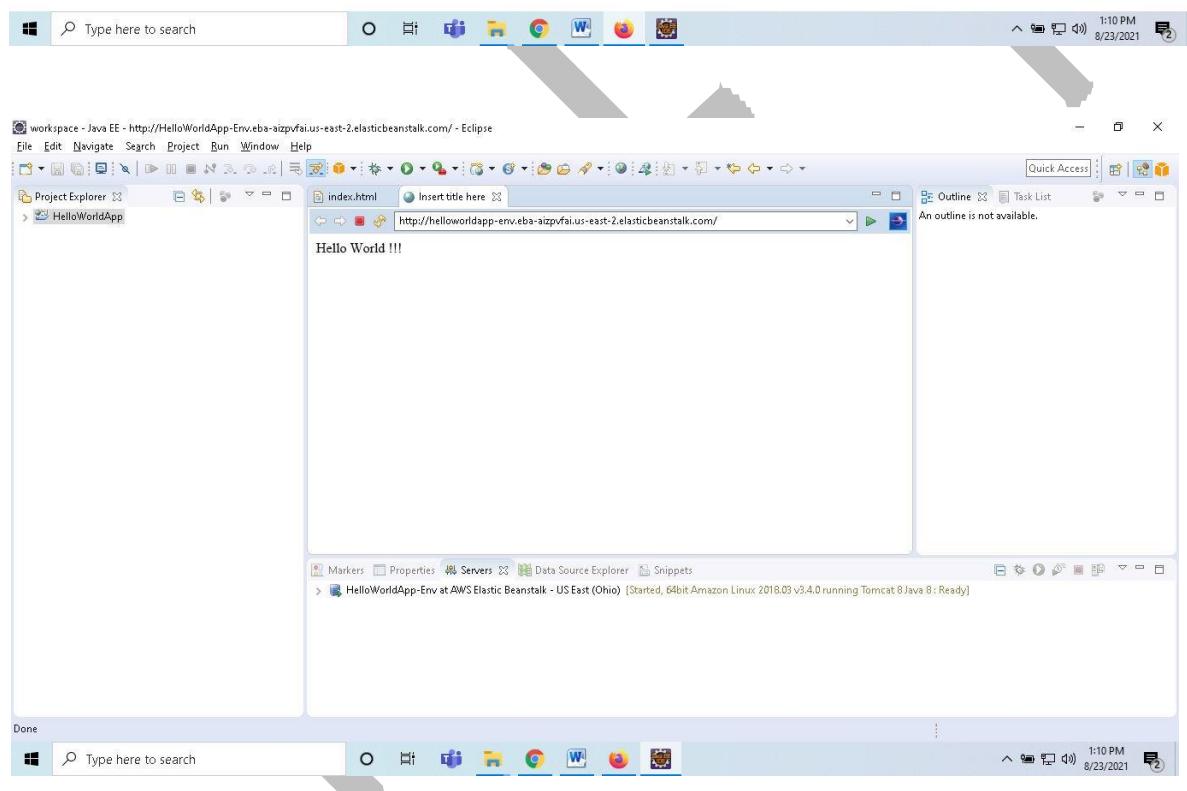
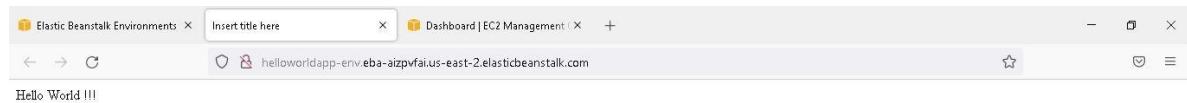
Instances (running)	1	Dedicated Hosts	0
Elastic IPs	0	Instances	1
Key pairs	0	Load balancers	1
Placement groups	0	Security groups	3
Snapshots	0	Volumes	1

A note at the bottom of this section states: "Easily size, configure, and deploy Microsoft SQL Server Always On availability groups on AWS using the AWS Launch Wizard for SQL Server. [Learn more](#)".

The right sidebar contains 'Account attributes' with 'Supported platforms' (VPC), 'Default VPC' (vpc-c83b59a3), and other settings like EBS encryption, Zones, and EC2 Serial Console.

The bottom navigation bar includes links for Feedback, English (US), Privacy Policy, Terms of Use, and Cookie preferences.

OUTPUT



VIVA QUESTIONS

1. What is Web Service?

2. What is Platform as a Service (PaaS)?

3. List some PaaS examples.

4. _____ web service that provides secure, resizable compute capacity in the cloud. It is designed to make web-scale cloud computing easier for developers.

5. _____ is an easy-to-use service for deploying and scaling web applications and services developed with Java, .NET, PHP, Node.js, Python, etc.,,

RESULT

Thus, the Java Web Application is created, deployed and managed in Amazon Web Services.

Observation by students: (what student able to?)

1.

CATEGORY	MAX. MARKS ALLOTED	MARKS AWARDED
PROCEDURAL STEPS	50	
VIVA	25	
RECORD	25	
TOTAL	100	

Ex. No: 4

Date:

Simulation of VM Scheduling Using CloudSim

AIM

To simulate VM Scheduling using CloudSim.

DESCRIPTION

CloudSim

- A Framework for modeling and simulation of Cloud Computing Infrastructures and services
- Originally built at the Cloud Computing Distributed Systems (CLOUDS) Laboratory, The University of Melbourne, Australia
- It is completely written in JAVA

Main Features of CloudSim

- Modeling and simulation
- Data center network topologies and message-passing applications
- Dynamic insertion of simulation elements
- Stop and resume of simulation
- Policies for allocation of hosts and virtual machines

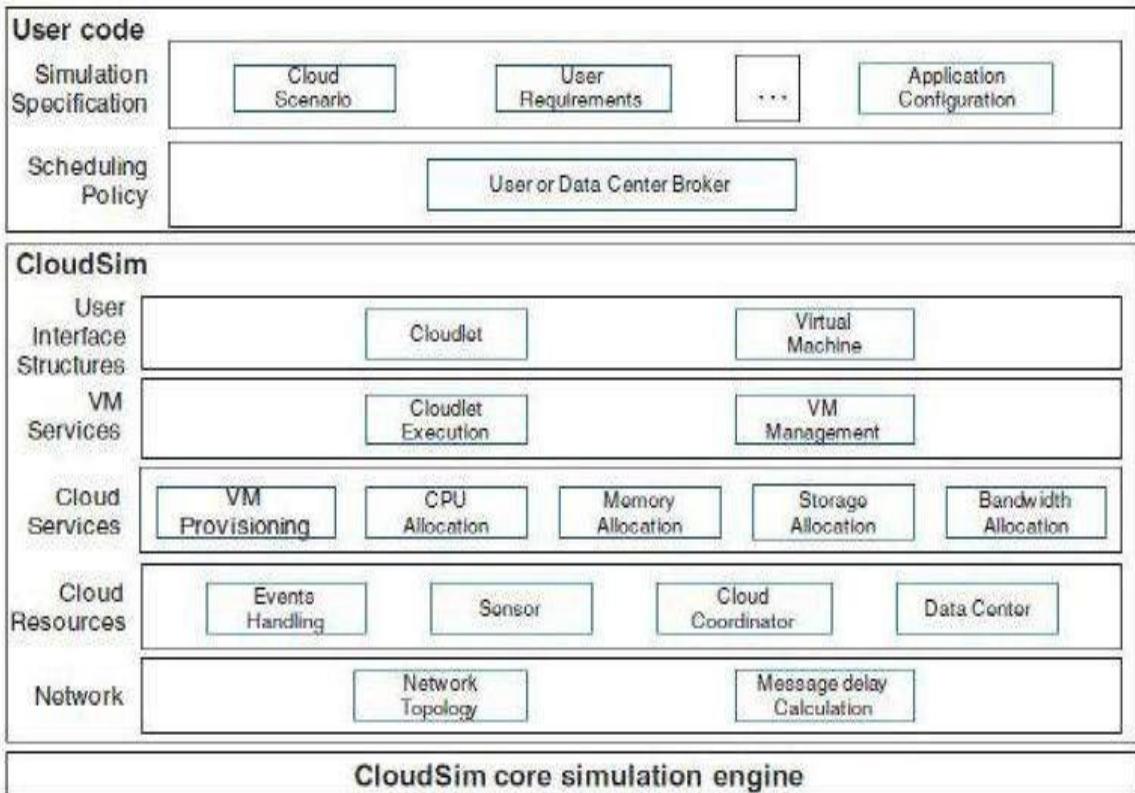
CloudSim – Essentials

- JDK 1.6 or above <http://tinyurl.com/JNU-JAVA>
- Eclipse 4.2 or above <http://tinyurl.com/JNU-Eclipse>
- Alternatively NetBeans <https://netbeans.org/downloads>
- Up & Running with cloudsim guide: <https://goo.gl/TPL7Zh>

CloudSim – Directory structure

- cloudsim/ -- top level CloudSim directory
- docs/ -- CloudSim API Documentation
- examples/ -- CloudSim examples
- jars/ -- CloudSim jar archives
- sources/ -- CloudSim source code

CloudSim - Layered Architecture



- **Core Functionalities** are queuing and processing of events, creation of Cloud system entities (services, host, data center, broker, VMs), communication between components, and management of the simulation clock.
- **CloudSim simulation layer** provides support for modeling and simulation of virtualized Cloud-based data center environments including dedicated management interfaces for VMs, memory, storage, and bandwidth.
- A Cloud host can be concurrently allocated to a set of VMs that execute applications based on SaaS provider's defined QoS levels.
- Top-most layer in the CloudSim stack is the User Code that exposes **basic entities for hosts** (number of machines, their specification, and so on), **applications** (number of tasks and their requirements), **VMs, number of users and their application types, and broker scheduling policies**.

Datacenter Entity

- The infrastructure-level services (IaaS) related to the clouds can be simulated by **extending the data center entity** of CloudSim.
- Data center entity **manages** several **host entities**.
- Hosts are assigned to one or more VMs based on a VM allocation policy that should be defined by the Cloud service provider.
- VM policy stands for the operations control policies related to VM life cycle such as provisioning of a host to a VM, VM creation, VM destruction, and VM migration.

Host Entity

- Represents a physical computing server in a Cloud.
- It assigns a pre-configured processing capability (expressed in millions of instructions per second—MIPS), memory, storage, and a provisioning policy for allocating processing cores to VMs.
- The Host component implements interfaces that support modeling and simulation of both single-core and multi-core nodes.

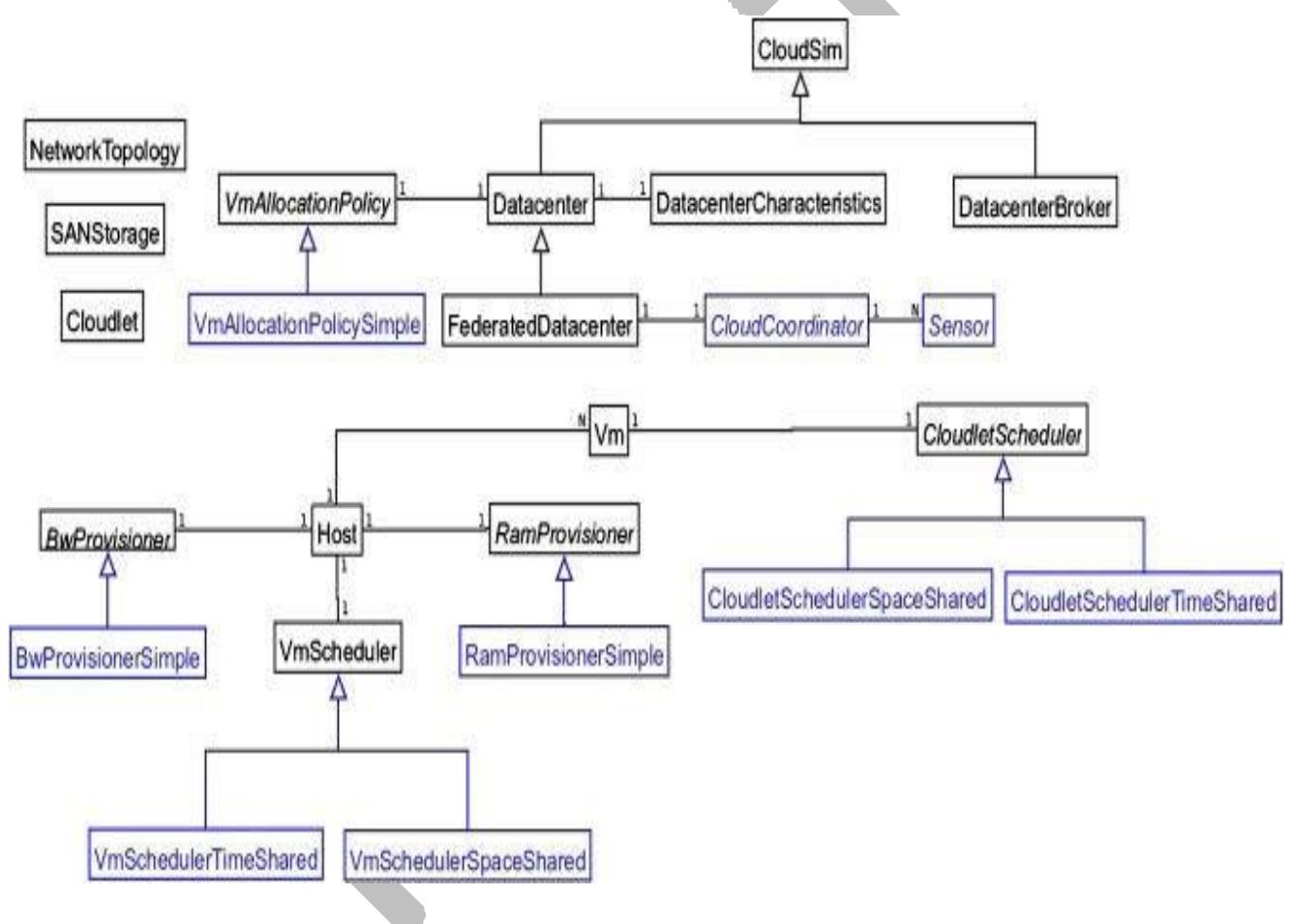
VM Allocation

- Process of creating VM instances on hosts that match the critical characteristics (storage, memory), configurations (software environment), and requirements (availability zone) of the SaaS provider.
- Once an application service (Cloudlet) is defined and modeled, it is assigned to one or more pre-instantiated VMs through a service-specific allocation policy.
- VmAllocationPolicy - VM Allocation controller component to allocate application-specific VMs to hosts in a data center. new policies can be created by the user by extending VmAllocationPolicy Entity.
- Default VmAllocationPolicy is FCFS.
- Hardware requirements, such as the number of processing cores, memory, and storage, form the basis for such provisioning.
- Other policies, including the ones likely to be expressed by Cloud providers, can also be easily simulated and modeled in CloudSim.
- Policies used by public Cloud providers (Amazon EC2, Microsoft Azure) are not publicly available

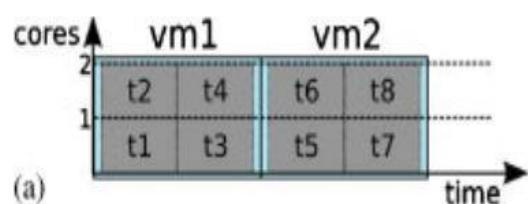
VM scheduler

- Allocation of processing cores to VMs is done by VM Scheduler using host allocation policy.
- Hardware characteristics, such as number of CPU cores, CPU share, and amount of memory (physical and secondary), that are allocated to a given VM instance
- Types:
 - Space Shared:** Assign specific CPU cores to specific VMs
 - Time Shared:** Dynamically distribute the capacity of a core among VMs

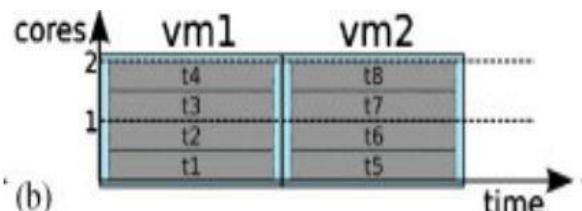
CloudSim – Component Model Classes



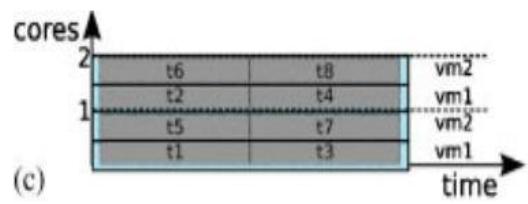
Modelling VM Allocation



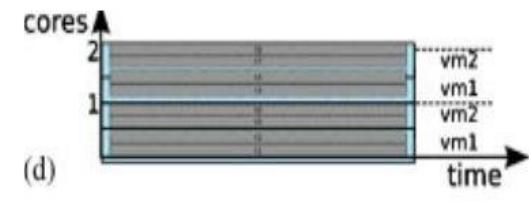
Space-shared provisioning for
VMs and tasks



Space-shared provisioning for VMs
and time-shared provisioning for tasks

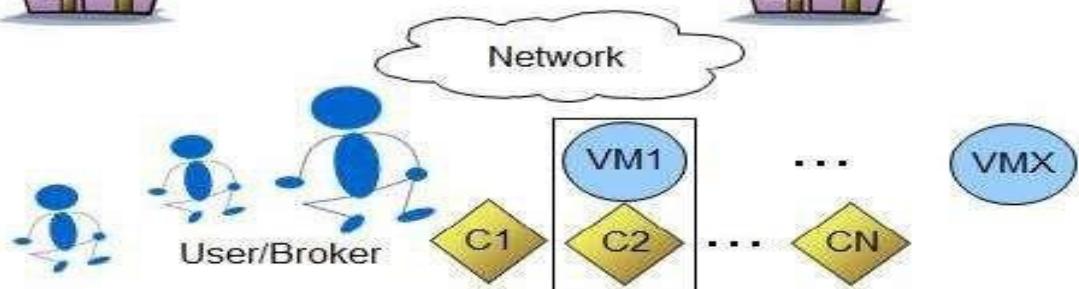
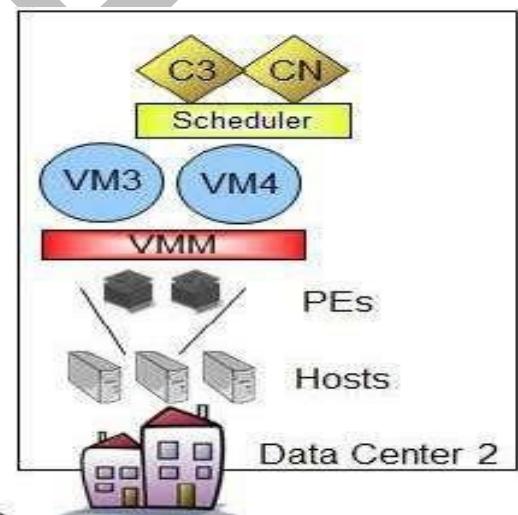
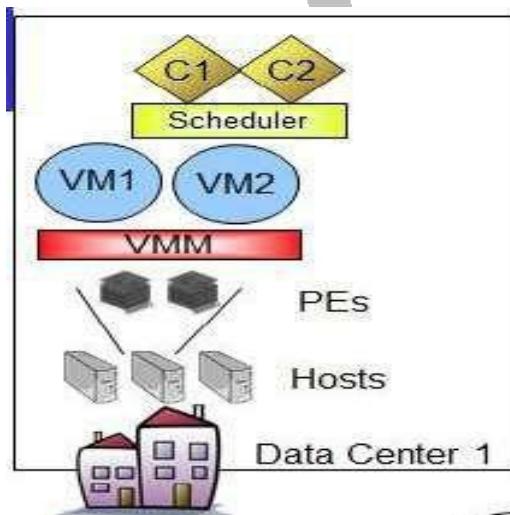


Time-shared provisioning for VMs, space-shared
provisioning for tasks



Time-shared provisioning for VMs and tasks

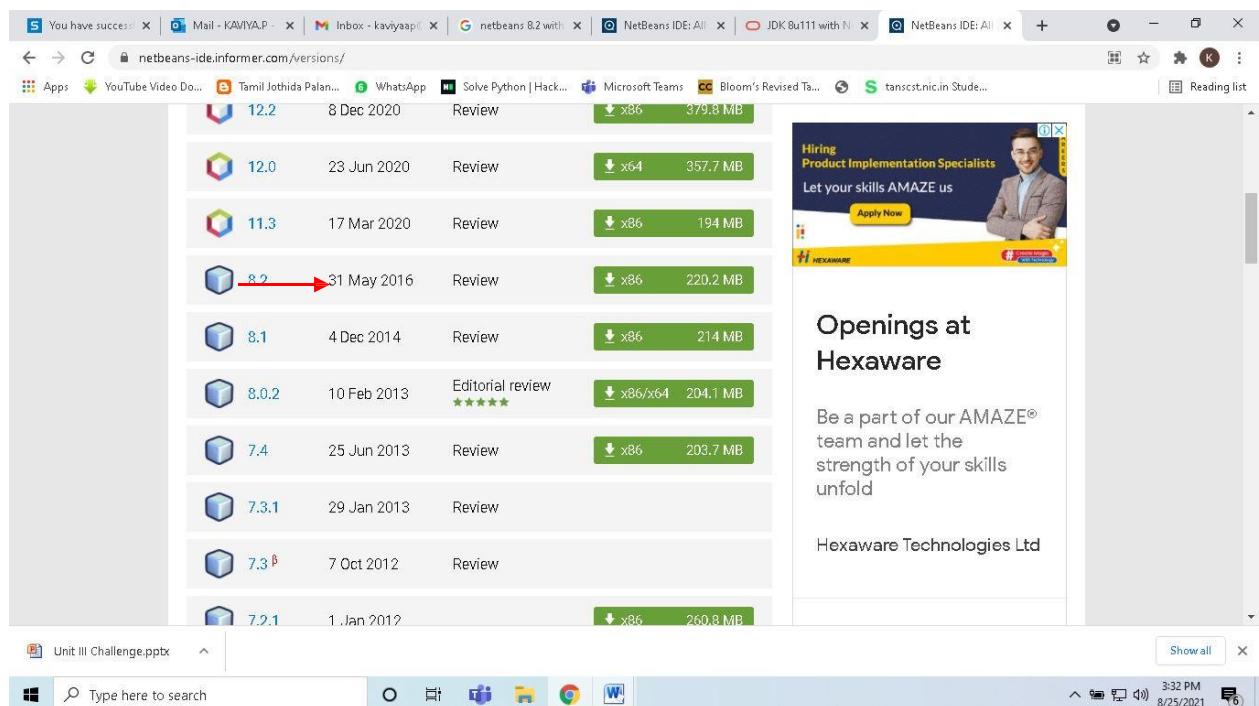
CloudSim Elements / Components



PROCEDURAL STEPS

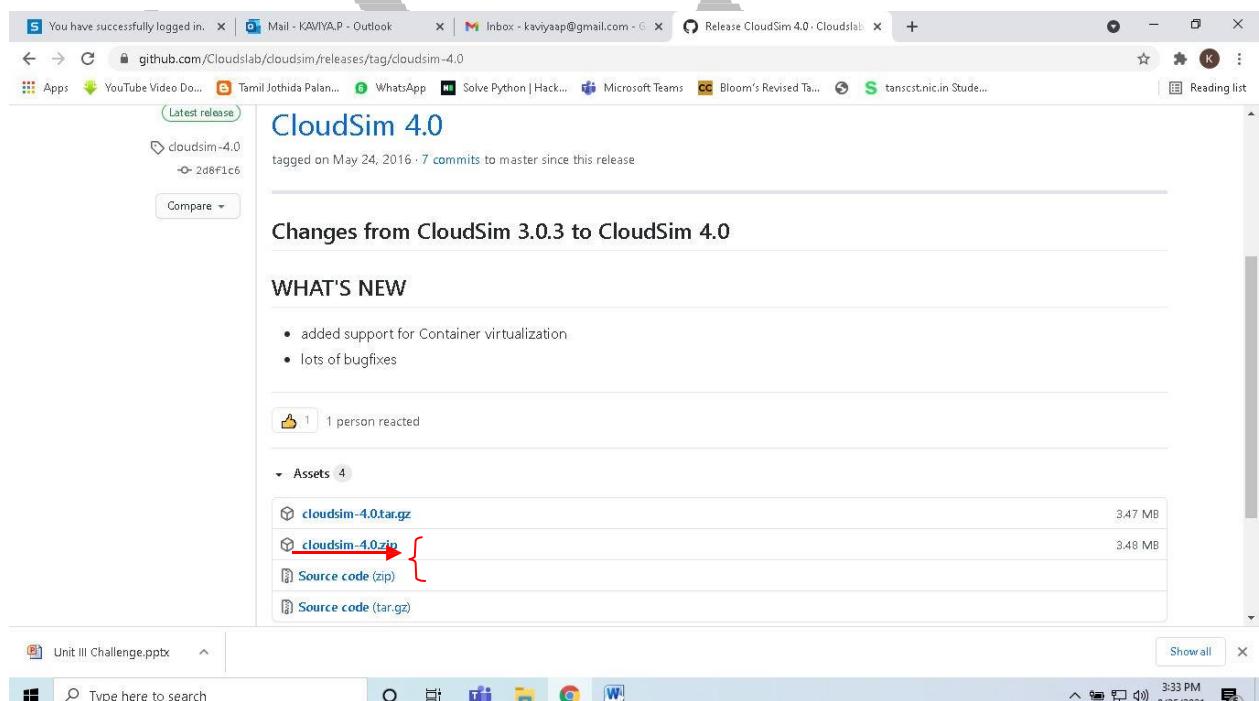
1. Download NetBeans and intall it.

Link: <https://netbeans-ide.informer.com/versions/>

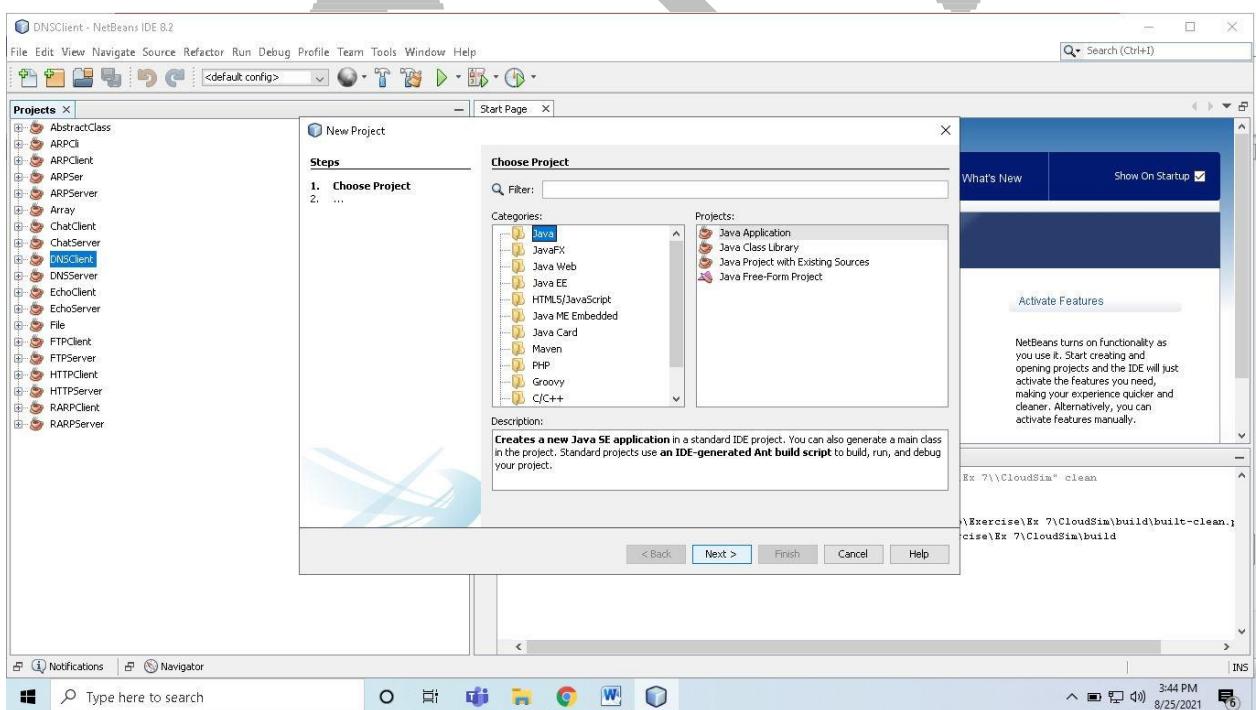
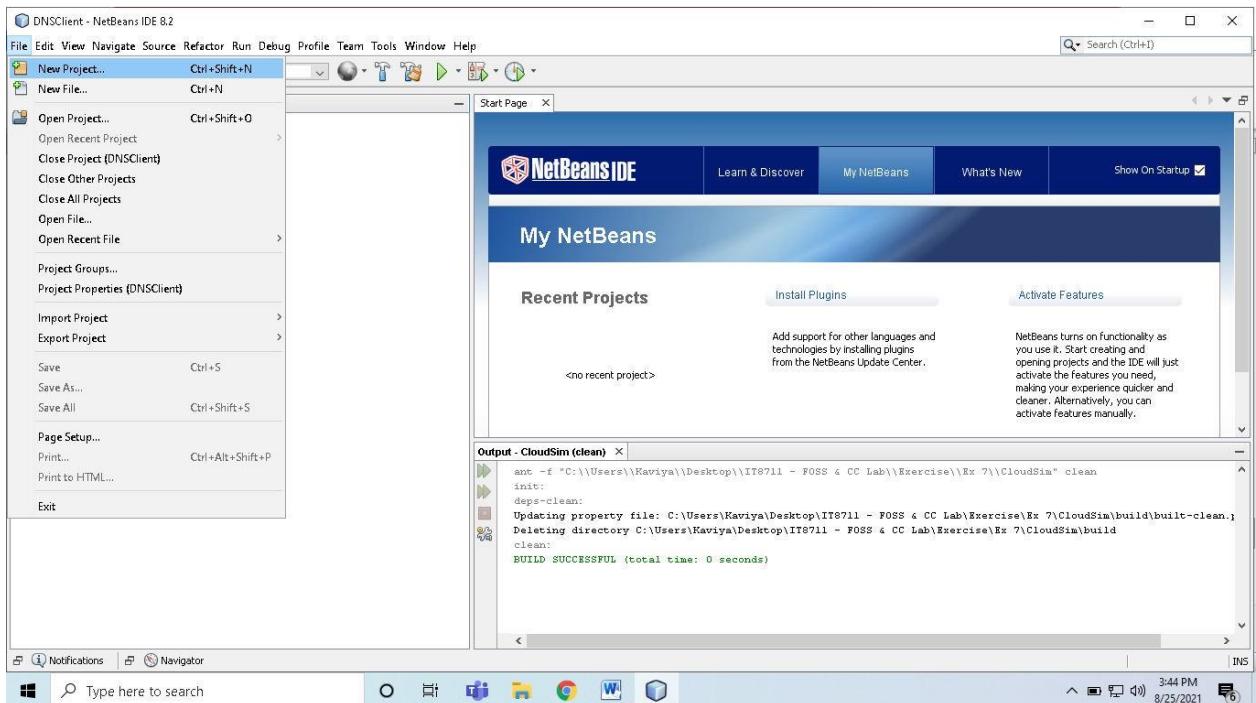


2. Download CloudSim 4.0 (jar and source code), extract and place it in respective folders.

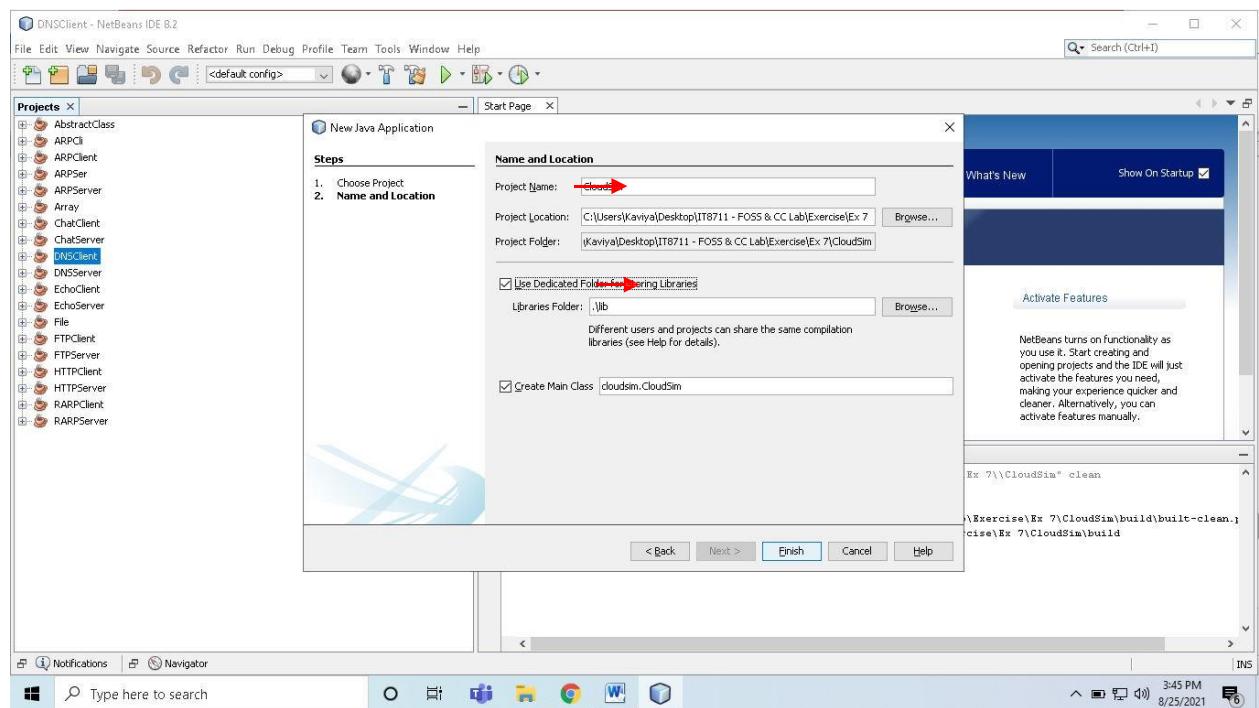
Link: <https://github.com/Cloudslab/cloudsim/releases/tag/cloudsim-4.0>



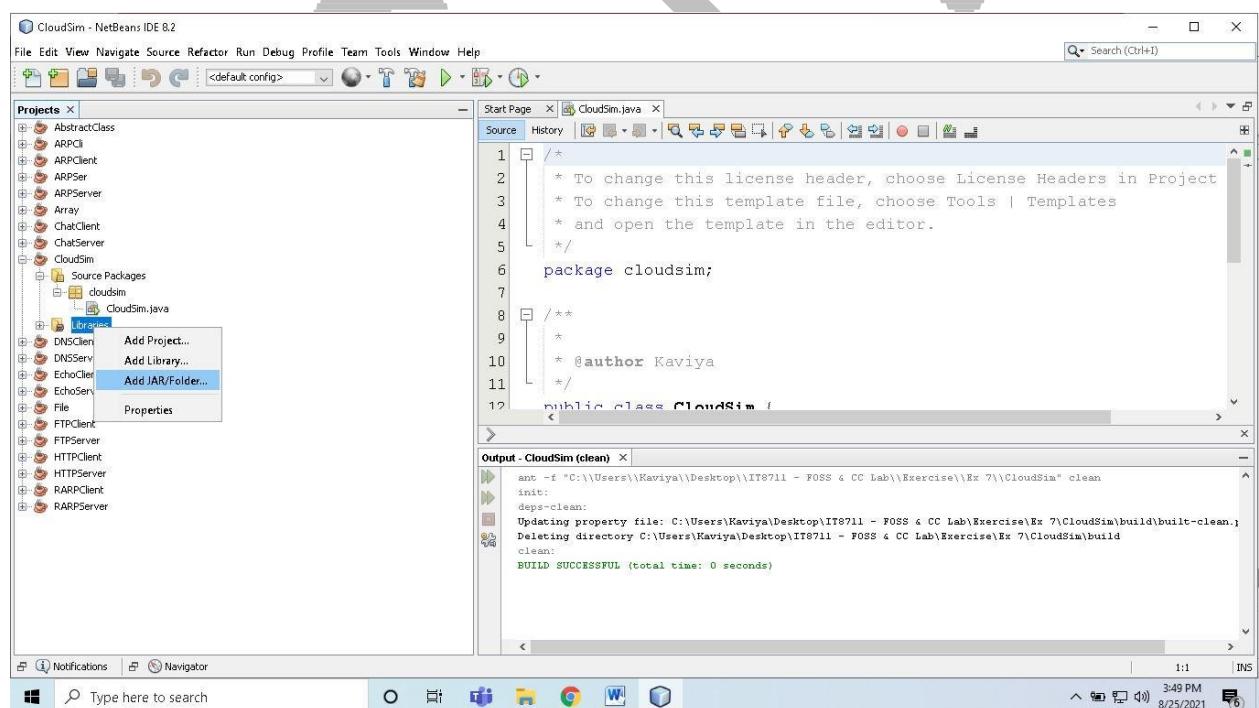
3. Open NetBeans & Create a project “CloudSim”

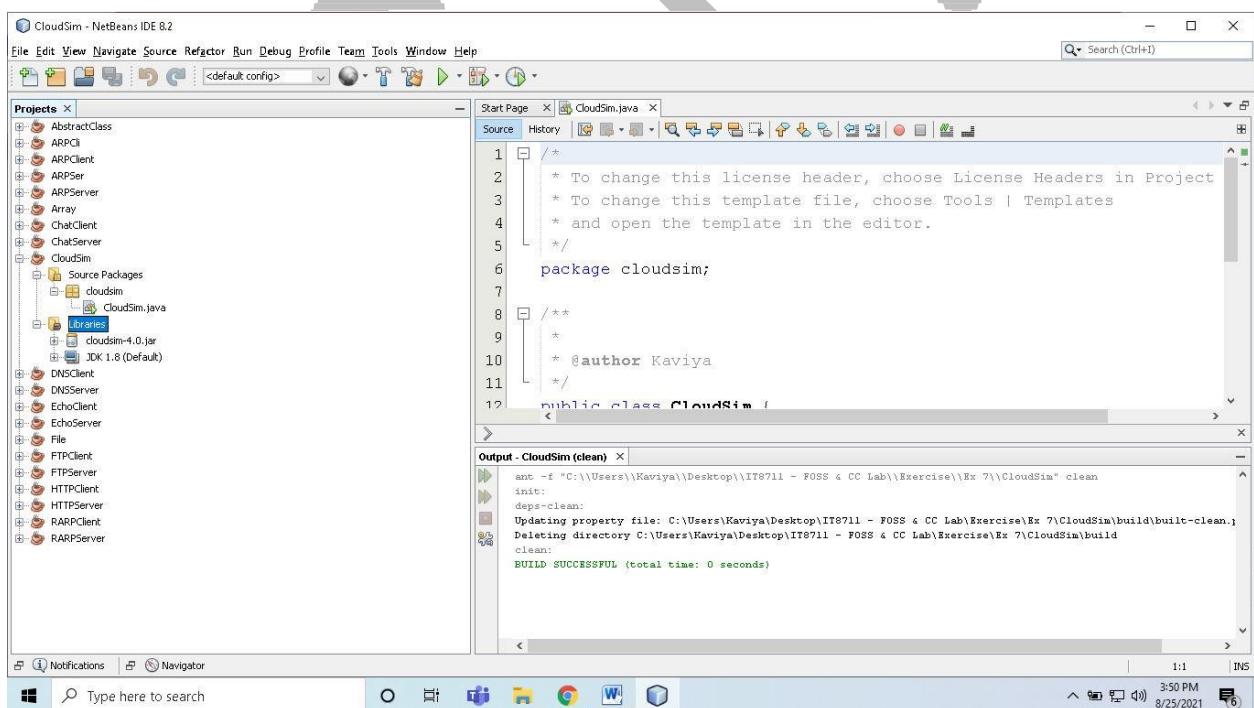
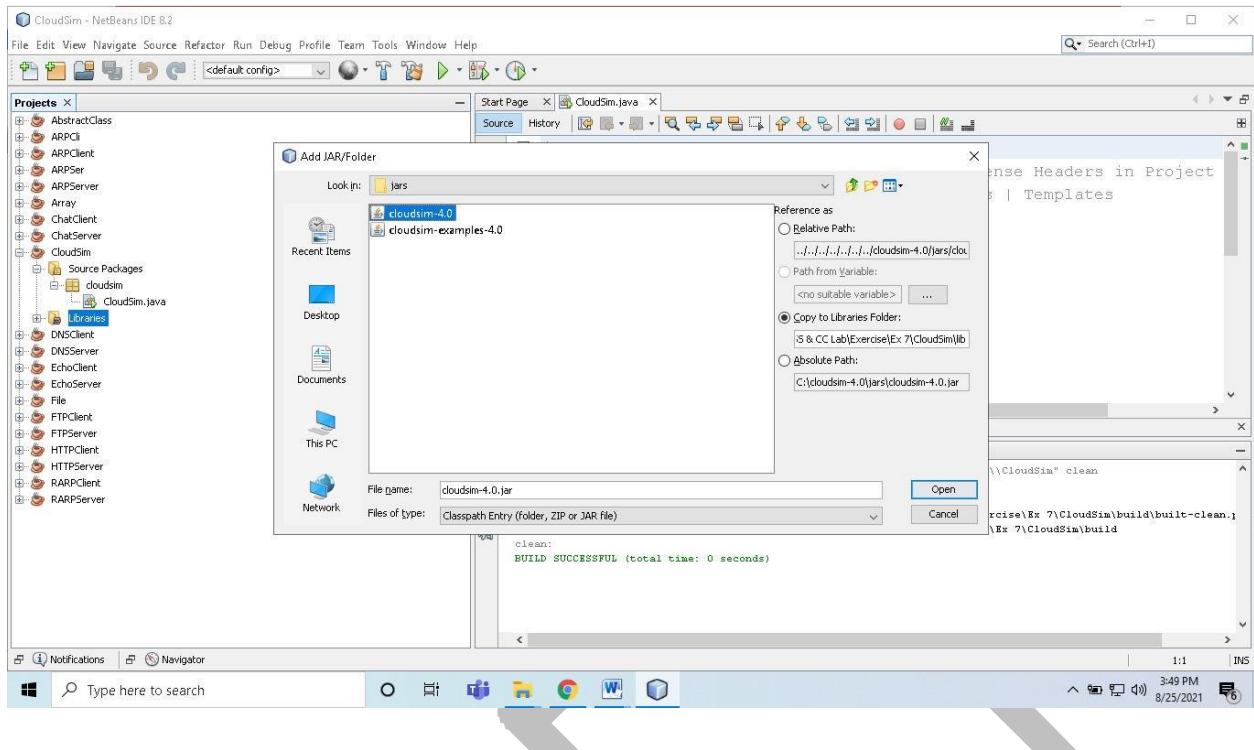


Check “Use dedicated folder for Storing Libraries” → Click “Finish”

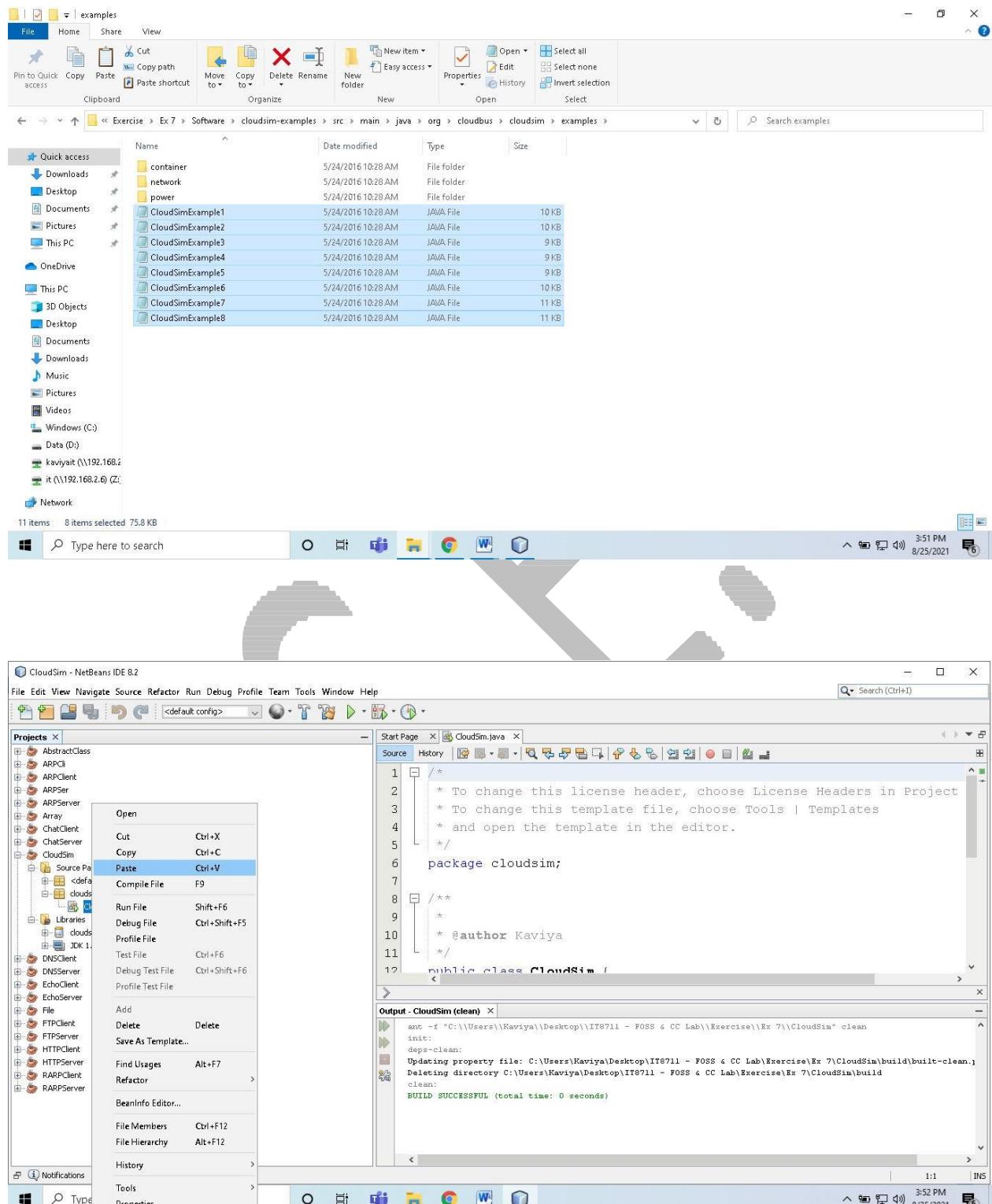


4. Add jar file “cloudsim-4.0”





5. Add “cloudsim examples” in CloudSim Project



CloudSim - NetBeans IDE 8.2

File Edit View Navigate Source Refactor Run Debug Profile Team Tools Window Help

Search (Ctrl+F)

Projects

- AbstractClass
- ARPCli
- ARPClient
- ARPSer
- ARPServer
- Array
- ChatClient
- ChatServer
- CloudSim
- Source Packages
 - cloudsim
 - CloudSim.java
 - CloudSimExample1.java
 - CloudSimExample2.java
 - CloudSimExample3.java
 - CloudSimExample4.java
 - CloudSimExample5.java
 - CloudSimExample6.java
 - CloudSimExample7.java
 - CloudSimExample8.java
- Libraries
 - cloudsim-4.0.jar
 - JDK 1.8 (Default)
- DNSClient
- DNSServer
- EchoClient
- EchoServer
- File
- FTPClient
- FTPServer
- HTTPClient
- HTTPServer
- Random

Output - CloudSim (clean) ×

```
ant -f "C:\\\\Users\\\\Kaviya\\\\Desktop\\\\IT8711 - FOSS & CC Lab\\\\Exercise\\\\Ex 7\\\\CloudSim" clean
init:
deps=clean:
Created dir: C:\\\\Users\\\\Kaviya\\\\Desktop\\\\IT8711 - FOSS & CC Lab\\\\Exercise\\\\Ex 7\\\\CloudSim\\\\build
Updating property file: C:\\\\Users\\\\Kaviya\\\\Desktop\\\\IT8711 - FOSS & CC Lab\\\\Exercise\\\\Ex 7\\\\CloudSim\\\\build\\\\built-clean.properties
Deleting directory C:\\\\Users\\\\Kaviya\\\\Desktop\\\\IT8711 - FOSS & CC Lab\\\\Exercise\\\\Ex 7\\\\CloudSim\\\\build
clean:
BUILD SUCCESSFUL (total time: 0 seconds)
```

Notifications Navigator

Type here to search

10:1 IN5

3:58 PM 8/25/2021

CloudSim - NetBeans IDE 8.2

File Edit View Navigate Source Refactor Run Debug Profile Team Tools Window Help

Search (Ctrl+F)

Projects

- AbstractClass
- ARPCli
- ARPClient
- ARPSer
- ARPServer
- Array
- ChatClient
- ChatServer
- CloudSim
- Source Packages
 - cloudsim
 - CloudSim.java
 - CloudSimExample1.java
 - CloudSimExample2.java
 - CloudSimExample3.java
 - CloudSimExample4.java
 - CloudSimExample5.java
 - CloudSimExample6.java
 - CloudSimExample7.java
 - CloudSimExample8.java
- Libraries
 - cloudsim-4.0.jar
 - JDK 1.8 (Default)
- DNSClient
- DNSServer
- EchoClient
- EchoServer
- File
- FTPClient
- FTPServer
- HTTPClient
- HTTPServer
- Random

Start Page

CloudSim.java × CloudSimExample1.java × CloudSimExample2.java ×

Source History

```
52  * @see vmList. */
53  * @static List<Vm> vmList;
54
55  /**
56   * Main method. It calls the main() method of the example class.
57   */
58  static void main(String[] args) {
59
60      Log.printLine("Starting CloudSimExample2...");
61
62      try {
63          // First step: Initialize the CloudSim package. It should be called
64          // before creating any entities.
65          int num_user = 1; // number of cloud users
66          Calendar calendar = Calendar.getInstance();
67          boolean trace_flag = false; // mean trace events
68
69          // Initialize the CloudSim library
70          CloudSim.init(num_user, calendar, trace_flag);
71
72          // Second step: Create Datacenters
73          // Datacenters are the resource providers in CloudSim. We need at least one of them to
74          // run.
    }
```

Output - CloudSim (run) ×

```
>>> run:
```

10:1 IN5

4:01 PM 8/25/2021

CloudSim - NetBeans IDE 8.2

File Edit View Navigate Source Refactor Run Debug Profile Team Tools Window Help

Start Page CloudSim.java CloudSimExample1.java CloudSimExample2.java

Projects

- AbstractClass
- ARPCli
- ARPClient
- ARPSer
- ARPServer
- Array
- ChatClient
- ChatServer
- CloudSim
 - Source Packages
 - cloudsim
 - CloudSim.java
 - CloudSimExample1.java
 - CloudSimExample2.java
 - CloudSimExample3.java
 - CloudSimExample4.java
 - CloudSimExample5.java
 - CloudSimExample6.java
 - CloudSimExample7.java
 - CloudSimExample8.java
 - Libraries
 - cloudsim-4.0.jar
 - JDK 1.8 (Default)
- DNSClient
- DNSServer
- EchoClient
- EchoServer
- File
- FTPClient
- FTPServer
- HTTPClient
- HTTPServer
- MANet

Source History

```
    // VM DESCRIPTION
    int vmid = 0;
    int mips = 250;
    long size = 10000; //image size (MB)
    int ram = 512; //vm memory (MB)
    long bw = 1000;
    int pesNumber = 1; //number of cpus
    String vmm = "Xen"; //VMM name

    //create two VMs
    Vm vm1 = new Vm(vmid, brokerId, mips, pesNumber, ram, bw, size, vmm, new CloudletSchedulerRoundRobin());
    vmid++;
    Vm vm2 = new Vm(vmid, brokerId, mips, pesNumber, ram, bw, size, vmm, new CloudletSchedulerRoundRobin());

    //add the VMs to the vmList
    vmList.add(vm1);
    vmList.add(vm2);

    //submit vm list to the broker
    broker.submitVmList(vmList);

    //Fifth step: Create two Cloudlets
}
```

Output - CloudSim (run) >> run:

10:1 4:02 PM 8/25/2021 IN5

CloudSim - NetBeans IDE 8.2

File Edit View Navigate Source Refactor Run Debug Profile Team Tools Window Help

Start Page CloudSim.java CloudSimExample1.java CloudSimExample2.java

Projects

- AbstractClass
- ARPCli
- ARPClient
- ARPSer
- ARPServer
- Array
- ChatClient
- ChatServer
- CloudSim
 - Source Packages
 - cloudsim
 - CloudSim.java
 - CloudSimExample1.java
 - CloudSimExample2.java
 - CloudSimExample3.java
 - CloudSimExample4.java
 - CloudSimExample5.java
 - CloudSimExample6.java
 - CloudSimExample7.java
 - CloudSimExample8.java
 - Libraries
 - cloudsim-4.0.jar
 - JDK 1.8 (Default)
- DNSClient
- DNSServer
- EchoClient
- EchoServer
- File
- FTPClient
- FTPServer
- HTTPClient
- HTTPServer
- MANet

Source History

```
    cloudletList = new ArrayList<Cloudlet>();

    //Cloudlet properties
    int id = 0;
    pesNumber=1;
    long length = 250000;
    long fileSize = 300;
    long outputSize = 300;
    UtilizationModel utilizationModel = new UtilizationModelFull();

    Cloudlet cloudlet1 = new Cloudlet(id, length, pesNumber, fileSize, outputSize, utilizationModel);
    cloudlet1.setUserId(brokerId);

    id++;
    Cloudlet cloudlet2 = new Cloudlet(id, length, pesNumber, fileSize, outputSize, utilizationModel);
    cloudlet2.setUserId(brokerId);

    //add the cloudlets to the list
    cloudletList.add(cloudlet1);
    cloudletList.add(cloudlet2);

    //submit cloudlet list to the broker
    broker.submitCloudletList(cloudletList);

}
```

Output - CloudSim (run) >> run:

10:1 4:02 PM 8/25/2021 IN5

CloudSim - NetBeans IDE 8.2

File Edit View Navigate Source Refactor Run Debug Profile Team Tools Window Help

Search (Ctrl+F)

Projects x

- AbstractClass
- ARPClient
- ARPServer
- Array
- ChatClient
- ChatServer
- CloudSim
- Source Packages
 - cloudsim
 - CloudSim.java
 - CloudSimExample1.java
 - CloudSimExample2.java
 - CloudSimExample3.java
 - CloudSimExample4.java
 - CloudSimExample5.java
 - CloudSimExample6.java
 - CloudSimExample7.java
 - CloudSimExample8.java
- Libraries
 - cloudsim-4.0.jar
 - JDK 1.8 (Default)
- DNSClient
- DNSServer
- EchoClient
- EchoServer
- File
- FTPClient
- FTPServer
- HTTPClient
- HTTPServer
- Log

Start Page x CloudSim.java x CloudSimExample1.java x CloudSimExample2.java x

Source History

```

133 //bind the cloudlets to the vms. This way, the broker
134 // will submit the bound cloudlets only to the specific VM
135 broker.bindCloudletToVm(cloudlet1.getId(),vm1.getId());
136 broker.bindCloudletToVm(cloudlet2.getId(),vm2.getId());

137 // Sixth step: Starts the simulation
138 CloudSim.startSimulation();

139 // Final step: Print results when simulation is over
140 List<Cloudlet> newList = broker.getCloudletReceivedList();

141 CloudSim.stopSimulation();

142 printCloudletList(newList);

143 Log.printLine("CloudSimExample2 finished!");

144 catch (Exception e) {
145     e.printStackTrace();
146     Log.printLine("The simulation has been terminated due to an unexpected error");
147 }
148
149
150
151
152
153
154
155

```

Output - CloudSim (run) x

```

>>> run:

```

10:1 402 PM 8/25/2021

6. Run the CloudSim examples

CloudSim - NetBeans IDE 8.2

File Edit View Navigate Source Refactor Run Debug Profile Team Tools Window Help

Search (Ctrl+F)

Projects x

- AbstractClass
- ARPClient
- ARPServer
- Array
- ChatClient
- ChatServer
- CloudSim
- Source Packages
 - cloudsim
 - CloudSim.java
 - CloudSimExample1.java
 - CloudSimExample2.java
 - CloudSimExample3.java
 - CloudSimExample4.java
 - CloudSimExample5.java
 - CloudSimExample6.java
 - CloudSimExample7.java
 - CloudSimExample8.java
- Libraries
 - cloudsim-4.0.jar
 - JDK 1.8 (Default)
- DNSClient
- DNSServer
- EchoClient
- EchoServer
- File
- FTPClient
- FTPServer
- HTTPClient
- HTTPServer
- Log

Start Page x CloudSim.java x CloudSimExample1.java x CloudSimExample2.java x

Source History

Open

- Cut Ctrl+X
- Copy Ctrl+C
- Paste Ctrl+V
- Compile File F9
- Run File Shift+F6**
- Debug File Ctrl+Shift+F5
- Profile File
- Test File Ctrl+F6
- Debug Test File Ctrl+Shift+F6
- Profile Test File

```

1 /*
2  * Title: CloudSim Toolkit
3  * Description: CloudSim (Cloud Simulation) Toolkit for Modeling a
4  *             of Clouds
5  * Licence: GPL - http://www.gnu.org/copyleft/gpl.html
6  *
7  * Copyright (c) 2009, The University of Melbourne, Australia
8  */
9
10 package cloudsim;
11
12
13
14
15
16
17

```

Output - CloudSim (clean) x

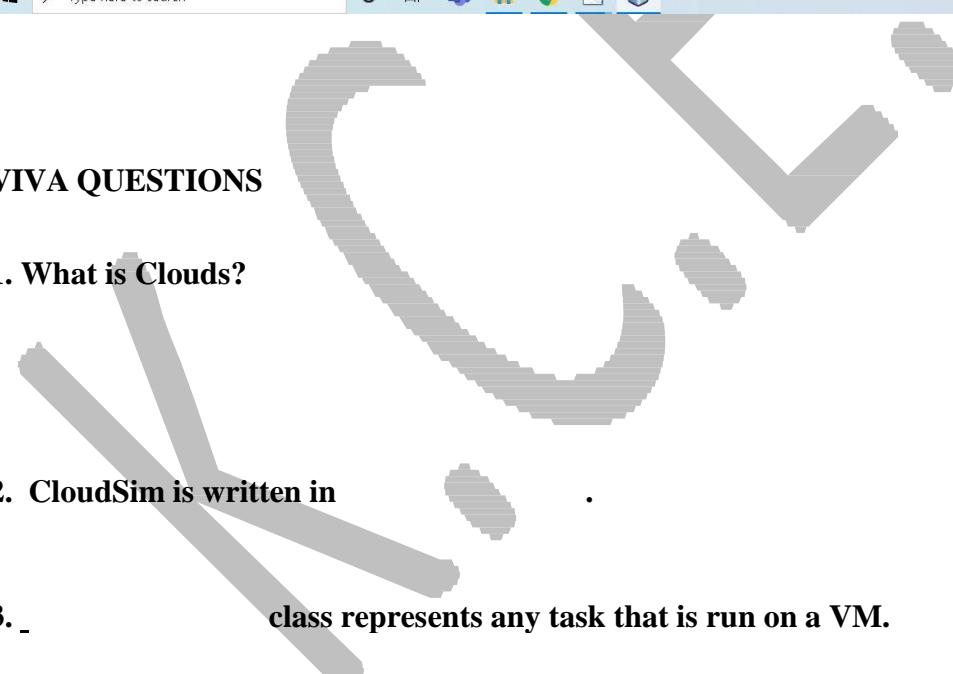
```

>>> ant -f "C:\\Users\\Kaviya\\Desktop\\IT8711 - FOSS & CC Lab\\Exercise\\Ex 7\\CloudSim" clean
init:
deps-clean:
Created dir: C:\\Users\\Kaviya\\Desktop\\IT8711 - FOSS & CC Lab\\Exercise\\Ex 7\\CloudSim\\build
Updating property file: C:\\Users\\Kaviya\\Desktop\\IT8711 - FOSS & CC Lab\\Exercise\\Ex 7\\CloudSim\\build\\built-clean.l
Deleting directory C:\\Users\\Kaviya\\Desktop\\IT8711 - FOSS & CC Lab\\Exercise\\Ex 7\\CloudSim\\build
clean:
BUILD SUCCESSFUL (total time: 0 seconds)

```

10:1 3:59 PM 8/25/2021

7. Output for CloudSimExample2.java



```
CloudSim - NetBeans IDE 8.2
File Edit View Navigate Source Refactor Run Debug Profile Team Tools Window Help
<default config> Search (Ctrl+I)
Projects x Start Page x CloudSim.java x CloudSimExample1.java x CloudSimExample2.java x
Source History ...
Output - CloudSim (run) x
run:
Starting CloudSimExample2...
Initialising...
Starting CloudSim version 3.0
Datacenter_0 is starting...
Broker is starting...
Entities started.
0.0: Broker: Cloud Resource List received with 1 resource(s)
0.0: Broker: Trying to Create VM #0 in Datacenter_0
0.0: Broker: Trying to Create VM #1 in Datacenter_0
0.1: Broker: VM #0 has been created in Datacenter #2, Host #0
0.1: Broker: VM #1 has been created in Datacenter #2, Host #0
0.1: Broker: Sending cloudlet 0 to VM #0
0.1: Broker: Sending cloudlet 1 to VM #1
1000.1: Broker: Cloudlet 0 received
1000.1: Broker: Cloudlet 1 received
1000.1: Broker: All Cloudlets executed. Finishing...
1000.1: Broker: Destroying VM #0
1000.1: Broker: Destroying VM #1
Broker is shutting down...
Simulation: No more future events
CloudInformationService: Notify all CloudSim entities for shutting down.
Datacenter_0 is shutting down...
Broker is shutting down...
Simulation completed.
Simulation completed.

=====
CloudSimExample2 finished!
```

VIVA QUESTIONS

1. What is Clouds?

2. CloudSim is written in _____.

3. _____ class represents any task that is run on a VM.

4. _____ is responsible for functioning of VMs, including VM creation, management, destruction, and submission of cloudlets to the VM.

5. _____ is the default VM allocation policy.

RESULT

Thus, the VM scheduling algorithms in CloudSim are studied and executed.

Observation by students: (what student able to?)

1.

CATEGORY	MAX. MARKS ALLOTED	MARKS AWARDED
PROCEDURAL STEPS	50	
VIVA	25	
RECORD	25	
TOTAL	100	

Ex. No: 5

Date:

Installation of Single Node Hadoop Cluster

AIM

To find procedure to set up the one node Hadoop cluster.

SCENARIO

You are a data engineering intern working for a startup that handles large datasets. Your team is exploring Apache Hadoop for big data processing, and you have been assigned the task of setting up a one-node Hadoop cluster on a Linux machine.

DESCRIPTION

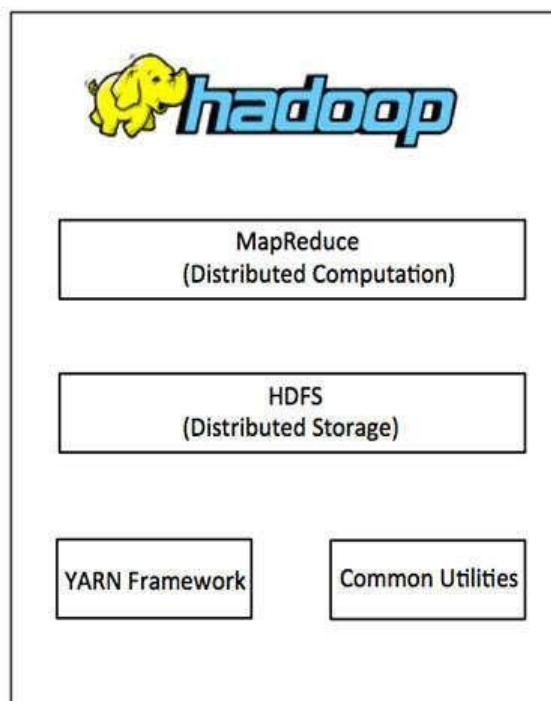
Hadoop is an Apache open source framework written in java that allows distributed

processing of large datasets across clusters of computers using simple programming models. The Hadoop framework application works in an environment that provides distributed *storage* and *computation* across clusters of computers. Hadoop is designed to scale up from single server to thousands of machines, each

Hadoop Architecture

Hadoop has two major layers namely

- Processing/Computation layer (MapReduce),
- Storage layer (Hadoop Distributed File



MapReduce

MapReduce is a parallel programming model for writing distributed applications devised at Google for efficient processing of large amounts of data (multi-terabyte datasets), on large clusters (thousands of nodes) of commodity hardware in a reliable, fault-tolerant manner.

The MapReduce program runs on Hadoop which is an Apache open-source framework.

Hadoop Distributed File System

The Hadoop Distributed File System (HDFS) is based on the Google File System (GFS) and provides a distributed file system that is designed to run on commodity hardware. It has many similarities with existing distributed file systems. However, the differences from other distributed file systems are significant. It is highly fault-tolerant and is designed to be deployed on low-cost hardware. It provides high throughput access to application data and is suitable for applications having large datasets.

Apart from the above-mentioned two core components, Hadoop framework also includes the following two modules –

- **Hadoop Common** – These are Java libraries and utilities required by other Hadoop modules.
- **Hadoop YARN** – This is a framework for job scheduling and cluster resource management.

Advantages of Hadoop

- Hadoop framework allows the user to quickly write and test distributed systems. It is efficient, and it automatically distributes the data and work across the machines and in turn, utilizes the underlying parallelism of the CPU cores.
- Hadoop does not rely on hardware to provide fault-tolerance and high availability (FTHA), rather Hadoop library itself has been designed to detect and handle failures at the application layer.
- Servers can be added or removed from the cluster dynamically and Hadoop continues to operate without interruption.
- Another big advantage of Hadoop is that apart from being open source, it is compatible on all the platforms since it is Java based.

PROCEDURE STEPS

Step 1: Download the following Packages

1. hadoop-3.2.2.tar.gz

Link: <https://www.apache.org/dyn/closer.cgi/hadoop/common/hadoop-3.2.2/hadoop-3.2.2.tar.gz>

2. 7zip to unzip the tar.gz file

Link: <https://www.7-zip.org/download.html>

3. Java 8 (JDK-8U333 & JRE-8U333)

Link: <https://www.oracle.com/java/technologies/javase/javase8u211-later-archive-downloads.html>

4. Hadoop dll Files & hadoop-hdfs-3.2.2.jar

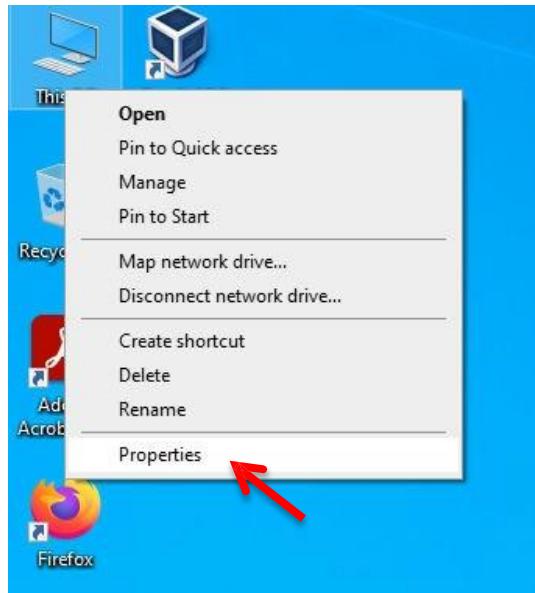
Link: <https://github.com/paranikumar/Hadoop-Dll-JAR-Files.git>

Step 2: Install / Extract the following Packages:

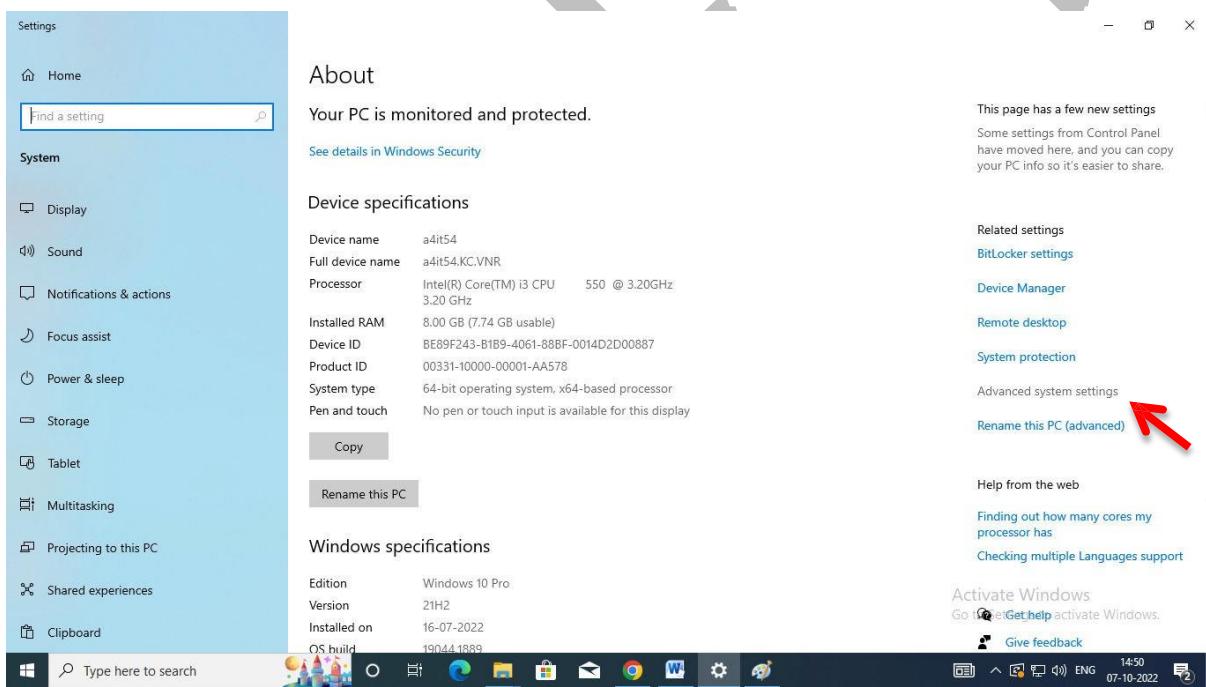
1. Uninstall the existing java and Install Java 8 in the location **C:\JAVA**. Before installing create two folders namely “**jdk1.8.0_333**” and “**jre1.8.0_333**” inside the “**C:\JAVA**” Folder.
2. Create a folder “**hadoop-env**” in **D:**
3. Paste the downloaded “hadoop-3.2.2.tar.gz” into “**d:\ hadoop-env**”
4. Unzip the “hadoop-3.2.2.tar.gz” using 7zip (do the unzip function two times)
5. Copy the downloaded Hadoop dll files to the location “**D:\hadoop-env\hadoop-3.2.2\bin**”
6. Copy the downloaded hadoop-hdfs-3.2.2.jar to “**D:\hadoop-env\hadoop-3.2.2\share\hadoop\hdfs**”
7. Edit the **JAVA_HOME = C:\JAVA\jdk1.8.0_333** in the path “**D:\hadoop-env\hadoop-3.2.2\etc\hadoop\hadoop-env.cmd**

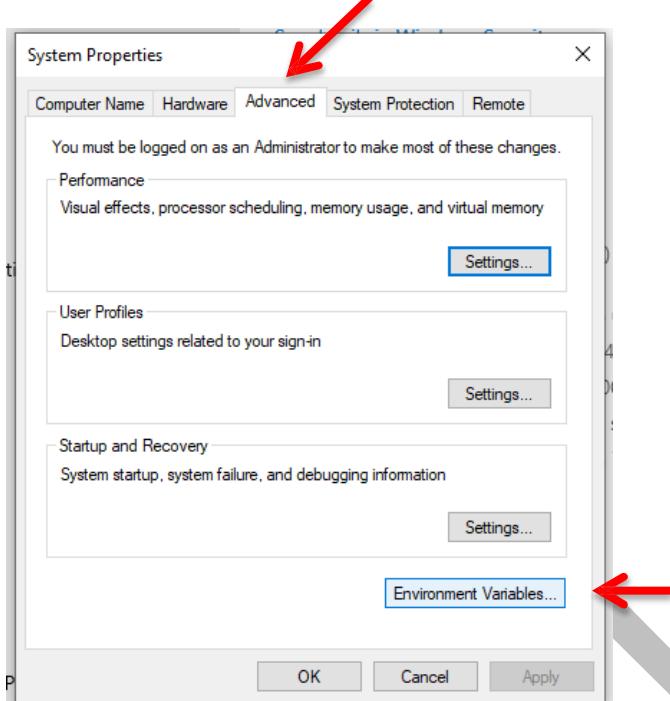
Step 3: Set the path for Java and Hadoop in system environment variables

To edit environment variables, go to Control Panel → System and Security → System (or) right-click “This PC” → Properties (My Computer icon) and click on the “Advanced system settings” link.

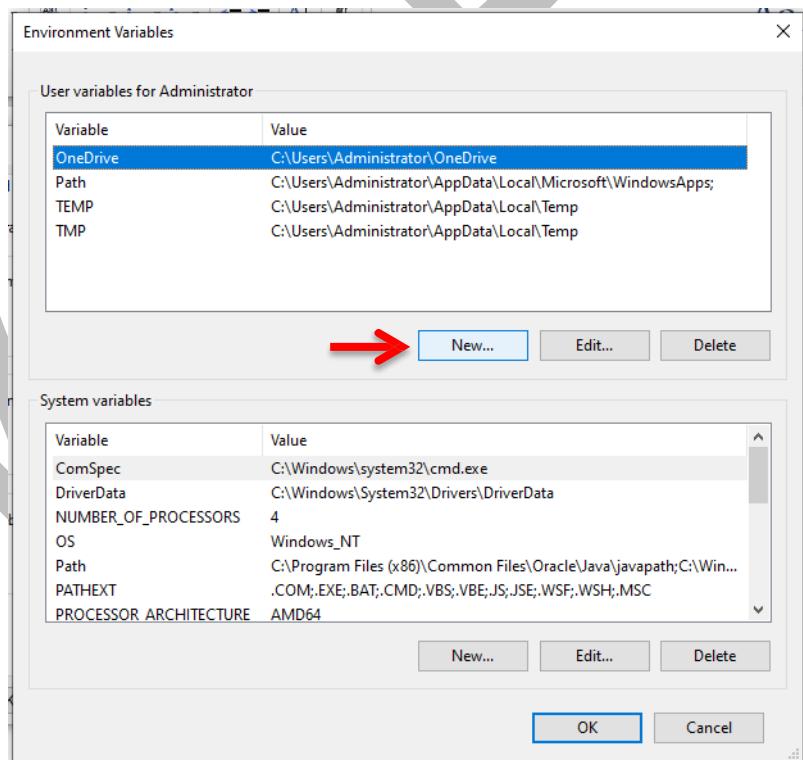


When the “Advanced system settings” dialog appears, go to the “Advanced” tab and click on the “Environment variables” button located on the bottom of the dialog.



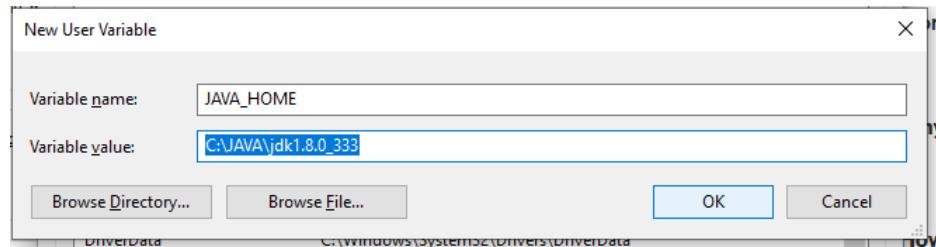


In the “Environment Variables” dialog, press the “New” button to add a new variable.

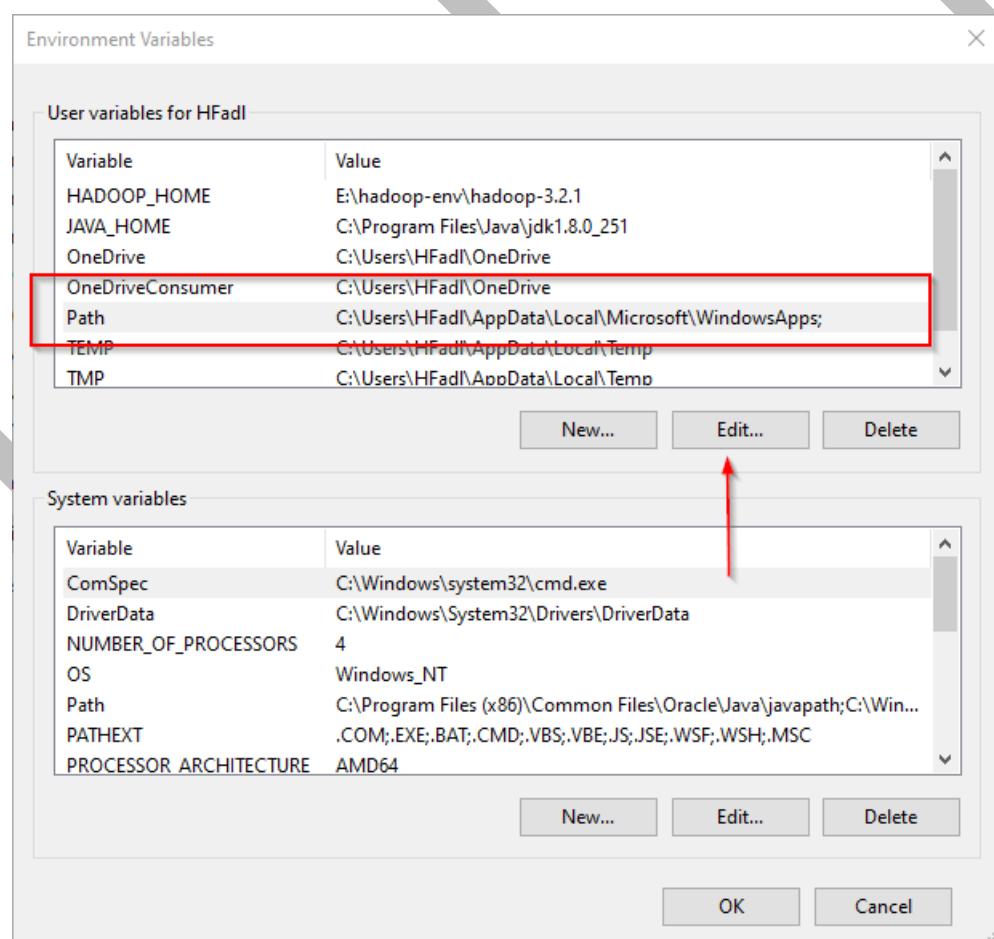
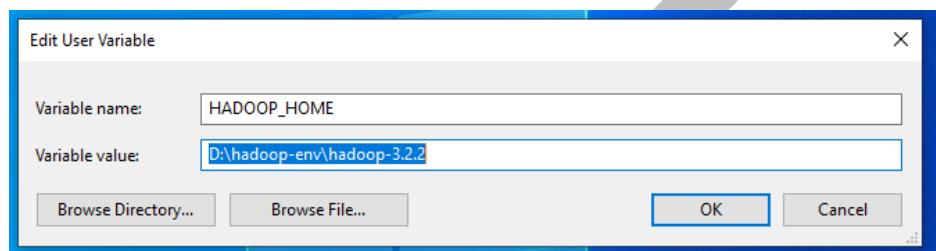


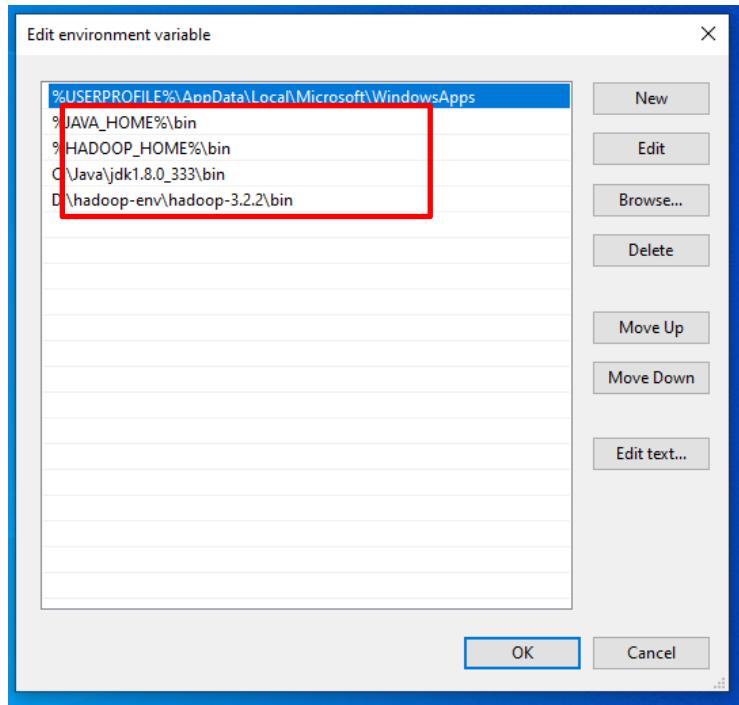
Now, Edit the PATH variable to add the Java and Hadoop binaries paths as shown in the following screenshots.

JAVA_HOME: JDK installation folder path



HADOOP_HOME: Hadoop installation folder path





Step 4: Open Command Prompt as Administrator and run the following command:

hadoop –version

```
Administrator: Windows PowerShell
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

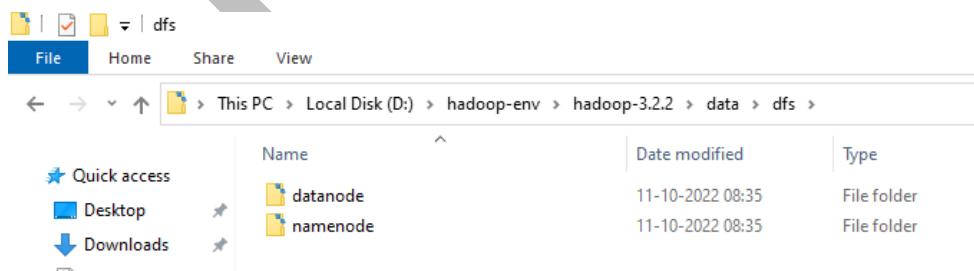
Try the new cross-platform PowerShell https://aka.ms/pscore6

PS C:\Users\Administrator> hadoop -version
java version "1.8.0_333"
Java(TM) SE Runtime Environment (build 1.8.0_333-b02)
Java HotSpot(TM) 64-Bit Server VM (build 25.333-b02, mixed mode)
PS C:\Users\Administrator>
```

Step 5: Create the following folders

D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode

D:\hadoop-env\hadoop-3.2.2\data\dfs\datanode



Step 6: Configuring Hadoop cluster

There are **four files** to configure Hadoop cluster:

Location of the File: “**D:\hadoop-env\hadoop-3.2.2\etc\hadoop**”

File Name: “**hdfs-site.xml**”

Add the following properties within the `<configuration></configuration>` element:

```
<property>
<name>dfs.replication</name>
<value>1</value>
</property>
<property>
<name>dfs.namenode.name.dir</name>
<value>file:///D:/hadoop-env/hadoop-3.2.2/data/dfs/namenode</value>
</property>
<property>
<name>dfs.datanode.data.dir</name>
<value>file:///D:/hadoop-env/hadoop-3.2.2/data/dfs/datanode</value>
</property>
```

File Name: “**core-site.xml**”

Add the following properties within the `<configuration></configuration>` element:

```
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:9820</value>
</property>
```

File Name: “**mapred-site.xml**”

Add the following properties within the `<configuration></configuration>` element:

```
<property>
<name>mapreduce.framework.name</name>
<value>yarn</value>
<description>MapReduce framework name</description>
</property>
```

File Name: “yarn-site.xml”

Add the following properties within the <configuration></configuration> element:

```
<property>
<name>yarn.nodemanager.aux-services</name>
<value>mapreduce_shuffle</value>
<description>Yarn Node Manager Aux Service</description>
</property>
```

Step 7: Format the namenode

After finishing the configuration, format the name using the following command

hdfs namenode –format

```
S C:\Users\Administrator> hdfs namenode -format
202-10-11 08:21:17,190 INFO namenode.NameNode: STARTUP_MSG: *****
*****
STARTUP_MSG: Starting NameNode
STARTUP_MSG: host = a4it55/172.16.8.55
STARTUP_MSG: args = [-format]
STARTUP_MSG: version = 3.2.2
STARTUP_MSG: classpath = D:\hadoop-env\hadoop-3.2.2\etc\hadoop;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\accessors-smart-1.2.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\animal-sniffer-annotations-1.17.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\asm-5.0.4.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\commo
dop-env\hadoop-3.2.2\share\hadoop\common\lib\audience-annotations-0.7.4.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\avro-1.7.7.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\checker-qual-2.5.2.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\commons-beanutils-1.9.4.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\commons-cl1-1.2.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\commons-collections-3.2.2.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\commons-compress-1.19.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\commons-configuration-2.1.1.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\commons-io-2.5.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\commo
202-10-11 08:21:18,894 INFO snapshot.SnapshotManager: SkipList is disabled
202-10-11 08:21:18,904 INFO util.GSet: Computing capacity for map cachedBlocks
202-10-11 08:21:18,904 INFO util.GSet: VM type = 64-bit
202-10-11 08:21:18,905 INFO util.GSet: 0.25% max memory 889 MB = 2.2 MB
202-10-11 08:21:18,905 INFO util.GSet: capacity = 2^18 = 262144 entries
202-10-11 08:21:18,910 INFO metrics.Metrics: NNTop conf: dfs.namenode.top.window.num.buckets = 10
202-10-11 08:21:18,918 INFO metrics.Metrics: NNTop conf: dfs.namenode.top.num.users = 10
202-10-11 08:21:18,919 INFO metrics.Metrics: NNTop conf: dfs.namenode.top.windows.minutes = 1,5,25
202-10-11 08:21:18,925 INFO namenode.FSNamesystem: Retry cache on namenode is enabled
202-10-11 08:21:18,926 INFO namenode.FSNamesystem: Retrying cache will use 0.03 of total heap and retry cache entry expiry time is 60000 millis
202-10-11 08:21:18,928 INFO namenode.FSNamesystem: Capacity cache will use 0.03 of total heap and map NameNodeRetryCache
202-10-11 08:21:18,930 INFO util.GSet: VM type = 64-bit
202-10-11 08:21:18,931 INFO util.GSet: 0.0399999999339447746% max memory 889 MB = 273.1 KB
202-10-11 08:21:18,931 INFO util.GSet: capacity = 2^15 = 32768 entries
Re-Format filesystem in Storage Directory root: D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode; location= null ? (Y or N) y
202-10-11 08:21:22,338 INFO namenode.FSImage: Allocated new image file: D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode\current\fsimage\fsimage_00000000000000000000000000000000
202-10-11 08:21:22,339 INFO common.Storage: Will remove files: []
202-10-11 08:21:22,436 INFO common.Storage: Storage director: D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode has been successfully formatted.
202-10-11 08:21:22,494 INFO namenode.FSImageFormatProtobuf: Saving image file D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode\current\fsimage\fsimage_00000000000000000000000000000000 of size 408 bytes saved in 0 seconds
202-10-11 08:21:22,625 INFO namenode.FSImageFormatProtobuf: 
202-10-11 08:21:22,678 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
202-10-11 08:21:22,688 INFO namenode.FSImage: FSImageSaver clean checkpoint: txid=0 when meet shutdown.
202-10-11 08:21:22,688 INFO namenode.NameNode: SHUTDOWN_MSG: *****
*****
SHUTDOWN_MSG: Shutting down NameNode at a4it55/172.16.8.55
*****
C:\Users\Administrator>
```

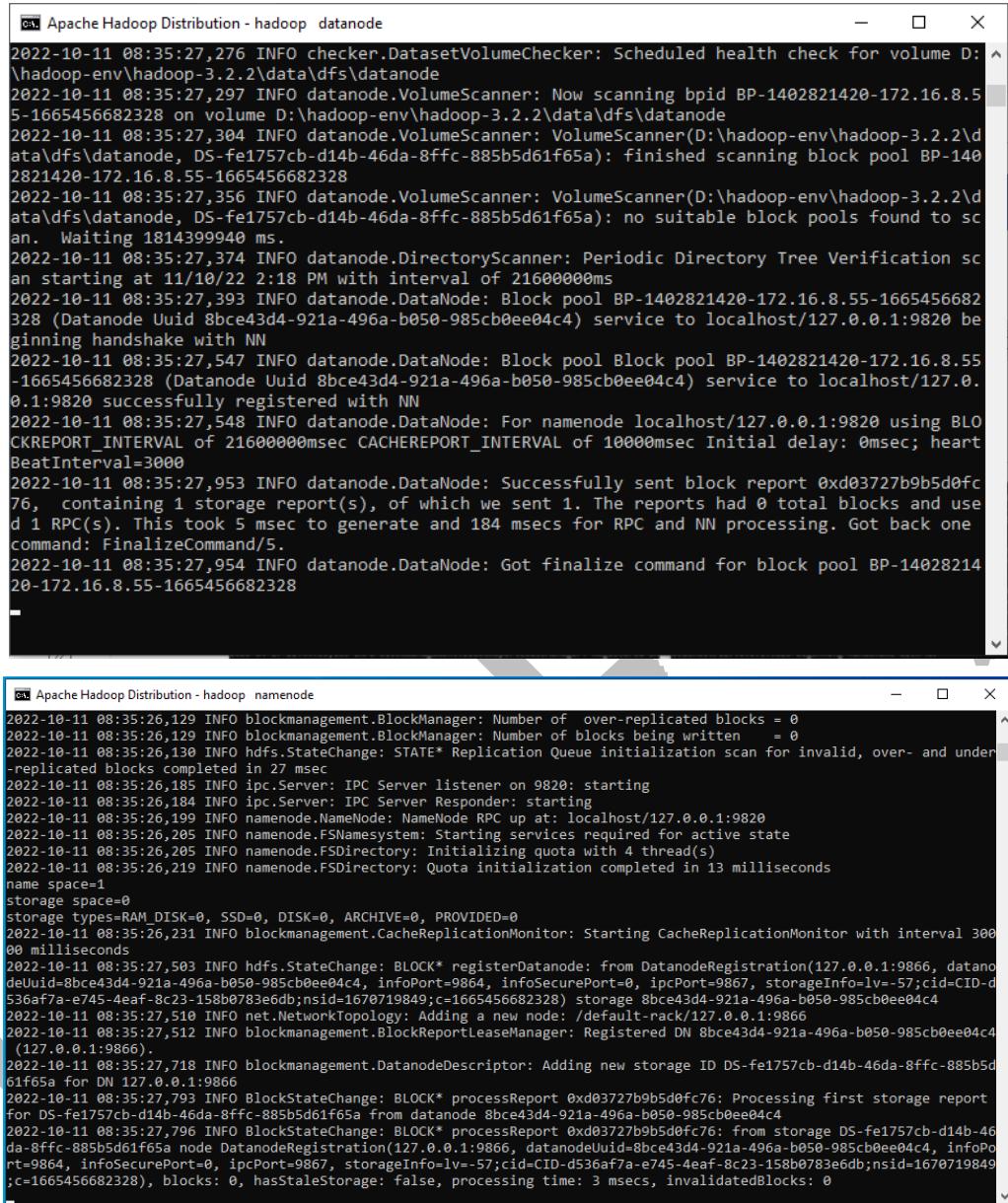
Step 8: Start the Hadoop nodes

Now navigate to the location “D:\hadoop-env\hadoop-3.2.2\sbin” in powershell and then run the following command to start the Hadoop nodes:

.\start-dfs.cmd

```
Administrator: Windows PowerShell
PS D:\hadoop-env\hadoop-3.2.2\sbin> .\start-dfs.cmd
PS D:\hadoop-env\hadoop-3.2.2\sbin>
```

Two command prompt windows will open (one for the **namenode** and one for the **datanode**) as follows:



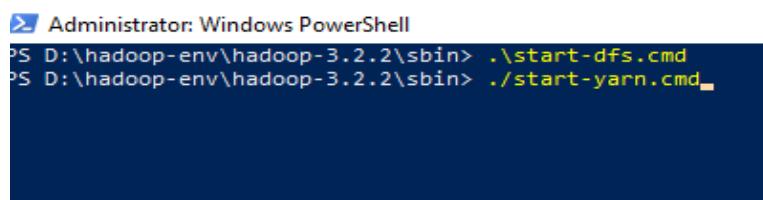
The image shows two separate command-line windows. The top window is titled "Apache Hadoop Distribution - hadoop datanode" and displays log messages from a DataNode. The bottom window is titled "Apache Hadoop Distribution - hadoop namenode" and displays log messages from a NameNode. Both windows show standard Hadoop startup and operational logs.

```
Apache Hadoop Distribution - hadoop datanode
2022-10-11 08:35:27,276 INFO checker.DatasetVolumeChecker: Scheduled health check for volume D:\hadoop-env\hadoop-3.2.2\data\dfs\datanode
2022-10-11 08:35:27,297 INFO datanode.VolumeScanner: Now scanning bpid BP-1402821420-172.16.8.55-1665456682328 on volume D:\hadoop-env\hadoop-3.2.2\data\dfs\datanode
2022-10-11 08:35:27,304 INFO datanode.VolumeScanner: VolumeScanner(D:\hadoop-env\hadoop-3.2.2\data\dfs\datanode, DS-fe1757cb-d14b-46da-8ffc-885b5d61f65a): finished scanning block pool BP-1402821420-172.16.8.55-1665456682328
2022-10-11 08:35:27,356 INFO datanode.VolumeScanner: VolumeScanner(D:\hadoop-env\hadoop-3.2.2\data\dfs\datanode, DS-fe1757cb-d14b-46da-8ffc-885b5d61f65a): no suitable block pools found to scan. Waiting 1814399940 ms.
2022-10-11 08:35:27,374 INFO datanode.DirectoryScanner: Periodic Directory Tree Verification scan starting at 11/10/22 2:18 PM with interval of 21600000ms
2022-10-11 08:35:27,393 INFO datanode.DataNode: Block pool BP-1402821420-172.16.8.55-1665456682328 (Datanode Uuid 8bce43d4-921a-496a-b050-985cb0ee04c4) service to localhost/127.0.0.1:9820 beginning handshake with NN
2022-10-11 08:35:27,547 INFO datanode.DataNode: Block pool Block pool BP-1402821420-172.16.8.55-1665456682328 (Datanode Uuid 8bce43d4-921a-496a-b050-985cb0ee04c4) service to localhost/127.0.0.1:9820 successfully registered with NN
2022-10-11 08:35:27,548 INFO datanode.DataNode: For namenode localhost/127.0.0.1:9820 using BLOCKREPORT_INTERVAL of 21600000msec CACHEREPORT_INTERVAL of 10000msec Initial delay: 0msec; heartBeatInterval=3000
2022-10-11 08:35:27,953 INFO datanode.DataNode: Successfully sent block report 0xd03727b9b5d0fc76, containing 1 storage report(s), of which we sent 1. The reports had 0 total blocks and used 1 RPC(s). This took 5 msec to generate and 184 msecs for RPC and NN processing. Got back one command: FinalizeCommand/5.
2022-10-11 08:35:27,954 INFO datanode.DataNode: Got finalize command for block pool BP-1402821420-172.16.8.55-1665456682328

Apache Hadoop Distribution - hadoop namenode
2022-10-11 08:35:26,129 INFO blockmanagement.BlockManager: Number of over-replicated blocks = 0
2022-10-11 08:35:26,129 INFO blockmanagement.BlockManager: Number of blocks being written = 0
2022-10-11 08:35:26,130 INFO hdfs.StateChange: STATE* Replication Queue initialization scan for invalid, over- and under-replicated blocks completed in 27 msec
2022-10-11 08:35:26,185 INFO ipc.Server: IPC Server listener on 9820: starting
2022-10-11 08:35:26,184 INFO ipc.Server: IPC Server Responder: starting
2022-10-11 08:35:26,199 INFO namenode.NameNode: NameNode RPC up at: localhost/127.0.0.1:9820
2022-10-11 08:35:26,205 INFO namenode.FSNamesystem: Starting services required for active state
2022-10-11 08:35:26,205 INFO namenode.FSDirectory: Initializing quota with 4 threads(s)
2022-10-11 08:35:26,219 INFO namenode.FSDirectory: Quota initialization completed in 13 milliseconds
name space=1
storage space=0
storage types=RAM_DISK=0, SSD=0, DISK=0, ARCHIVE=0, PROVIDED=0
2022-10-11 08:35:26,231 INFO blockmanagement.CacheReplicationMonitor: Starting CacheReplicationMonitor with interval 30000 milliseconds
2022-10-11 08:35:27,503 INFO hdfs.StateChange: BLOCK* registerDataNode: from DatanodeRegistration(127.0.0.1:9866, datanodeUuid=8bce43d4-921a-496a-b050-985cb0ee04c4, infoPort=9864, infoSecurePort=0, ipcPort=9867, storageInfo=lv=-57;cid=d536af7a-e745-4eaf-8c23-158b0783e6db;nsid=1670719849;c=1665456682328) storage 8bce43d4-921a-496a-b050-985cb0ee04c4
2022-10-11 08:35:27,510 INFO net.NetworkTopology: Adding a new node: /default-rack/127.0.0.1:9866
2022-10-11 08:35:27,512 INFO blockmanagement.BlockReportLeaseManager: Registered DN 8bce43d4-921a-496a-b050-985cb0ee04c4 (127.0.0.1:9866).
2022-10-11 08:35:27,718 INFO blockmanagement.DatanodeDescriptor: Adding new storage ID DS-fe1757cb-d14b-46da-8ffc-885b5d61f65a for DN 127.0.0.1:9866
2022-10-11 08:35:27,793 INFO BlockStateChange: BLOCK* processReport 0xd03727b9b5d0fc76: Processing first storage report for DS-fe1757cb-d14b-46da-8ffc-885b5d61f65a from datanode 8bce43d4-921a-496a-b050-985cb0ee04c4
2022-10-11 08:35:27,796 INFO BlockStateChange: BLOCK* processReport 0xd03727b9b5d0fc76: from storage DS-fe1757cb-d14b-46da-8ffc-885b5d61f65a node DatanodeRegistration(127.0.0.1:9866, datanodeUuid=8bce43d4-921a-496a-b050-985cb0ee04c4, infoPort=9864, infoSecurePort=0, ipcPort=9867, storageInfo=lv=-57;cid=d536af7a-e745-4eaf-8c23-158b0783e6db;nsid=1670719849;c=1665456682328), blocks: 0, hasStaleStorage: false, processing time: 3 msecs, invalidatedBlocks: 0
```

Next, Start the Hadoop Yarn service using the following command:

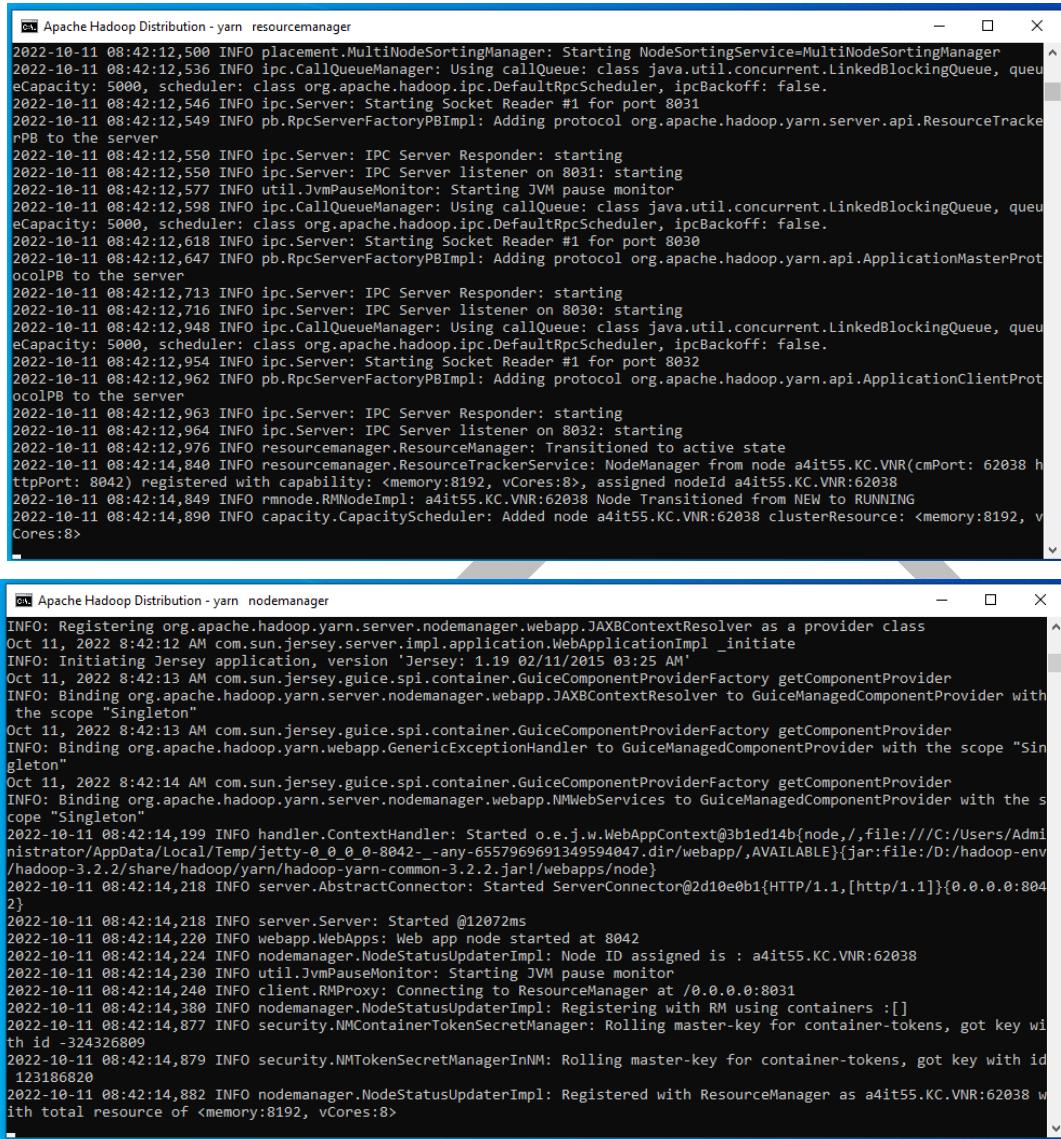
/start-yarn.cmd



The image shows a Windows PowerShell window running as Administrator. It contains the command `./start-yarn.cmd`, which is being executed. The command is part of a larger sequence where the user has already run `./start-dfs.cmd`.

```
Administrator: Windows PowerShell
PS D:\hadoop-env\hadoop-3.2.2\sbin> ./start-dfs.cmd
PS D:\hadoop-env\hadoop-3.2.2\sbin> ./start-yarn.cmd
```

Two command prompt windows will open (one for the **resource manager** and one for the **node manager**) as follows:



```
Apache Hadoop Distribution - yarn resourcemanager
2022-10-11 08:42:12,500 INFO placement.MultiNodeSortingManager: Starting NodeSortingService=MultiNodeSortingManager
2022-10-11 08:42:12,536 INFO ipc.CallQueueManager: Using callQueue: class java.util.concurrent.LinkedBlockingQueue, queueCapacity: 5000, scheduler: class org.apache.hadoop.ipc.DefaultRpcScheduler, ipcBackoff: false.
2022-10-11 08:42:12,546 INFO ipc.Server: Starting Socket Reader #1 for port 8031
2022-10-11 08:42:12,549 INFO pb.RpcServerFactoryPBImpl: Adding protocol org.apache.hadoop.yarn.server.api.ResourceTrackerPB to the server
2022-10-11 08:42:12,550 INFO ipc.Server: IPC Server Responder: starting
2022-10-11 08:42:12,550 INFO ipc.Server: IPC Server listener on 8031: starting
2022-10-11 08:42:12,577 INFO util.JvmPauseMonitor: Starting JVM pause monitor
2022-10-11 08:42:12,598 INFO ipc.CallQueueManager: Using callQueue: class java.util.concurrent.LinkedBlockingQueue, queueCapacity: 5000, scheduler: class org.apache.hadoop.ipc.DefaultRpcScheduler, ipcBackoff: false.
2022-10-11 08:42:12,618 INFO ipc.Server: Starting Socket Reader #1 for port 8030
2022-10-11 08:42:12,647 INFO pb.RpcServerFactoryPBImpl: Adding protocol org.apache.hadoop.yarn.api.ApplicationMasterProtocolPB to the server
2022-10-11 08:42:12,713 INFO ipc.Server: IPC Server Responder: starting
2022-10-11 08:42:12,716 INFO ipc.Server: IPC Server listener on 8030: starting
2022-10-11 08:42:12,948 INFO ipc.CallQueueManager: Using callQueue: class java.util.concurrent.LinkedBlockingQueue, queueCapacity: 5000, scheduler: class org.apache.hadoop.ipc.DefaultRpcScheduler, ipcBackoff: false.
2022-10-11 08:42:12,954 INFO ipc.Server: Starting Socket Reader #1 for port 8032
2022-10-11 08:42:12,962 INFO pb.RpcServerFactoryPBImpl: Adding protocol org.apache.hadoop.yarn.api.ApplicationClientProtocolPB to the server
2022-10-11 08:42:12,963 INFO ipc.Server: IPC Server Responder: starting
2022-10-11 08:42:12,964 INFO ipc.Server: IPC Server listener on 8032: starting
2022-10-11 08:42:12,976 INFO resourcemanager.ResourceManager: Transitioned to active state
2022-10-11 08:42:14,840 INFO resourcemanager.ResourceTrackerService: NodeManager from node a4it55.KC.VNR(cmPort: 62038 httpPort: 8042) registered with capability: <memory:8192, vCores:8>, assigned nodeID a4it55.KC.VNR:62038
2022-10-11 08:42:14,849 INFO rmnode.RMNodeImpl: a4it55.KC.VNR:62038 Node Transitioned from NEW to RUNNING
2022-10-11 08:42:14,890 INFO capacity.CapacityScheduler: Added node a4it55.KC.VNR:62038 clusterResource: <memory:8192, vCores:8>
```



```
Apache Hadoop Distribution - yarn nodemanager
INFO: Registering org.apache.hadoop.yarn.server.nodemanager.webapp.JAXBContextResolver as a provider class
Oct 11, 2022 8:42:12 AM com.sun.jersey.server.impl.application.WebApplicationImpl _initiate
INFO: Initiating Jersey application, version 'Jersey: 1.19 02/11/2015 03:25 AM'
Oct 11, 2022 8:42:13 AM com.sun.jersey.guice.spi.container.GuiceComponentProviderFactory getComponentProvider
INFO: Binding org.apache.hadoop.yarn.server.nodemanager.webapp.JAXBContextResolver to GuiceManagedComponentProvider with the scope "Singleton"
Oct 11, 2022 8:42:13 AM com.sun.jersey.guice.spi.container.GuiceComponentProviderFactory getComponentProvider
INFO: Binding org.apache.hadoop.yarn.webapp.GenericExceptionHandler to GuiceManagedComponentProvider with the scope "Singleton"
Oct 11, 2022 8:42:14 AM com.sun.jersey.guice.spi.container.GuiceComponentProviderFactory getComponentProvider
INFO: Binding org.apache.hadoop.yarn.server.nodemanager.webapp.NMWebServices to GuiceManagedComponentProvider with the scope "Singleton"
2022-10-11 08:42:14,199 INFO handler.ContextHandler: Started o.e.j.w.WebAppContext@3b1ed14b{node/,file:///C:/Users/Administrator/AppData/Local/Temp/jetty-0_0_0-8042_-any-6557969691349594047.dir/webapp/,AVAILABLE}{jar:file:/D:/hadoop-env/hadoop-3.2.2/share/hadoop/yarn/hadoop-yarn-common-3.2.2.jar!/webapps/node}
2022-10-11 08:42:14,218 INFO server.AbstractConnector: Started ServerConnector@2d10e0b1{HTTP/1.1,[http/1.1]}{0.0.0.0:8042}
2022-10-11 08:42:14,218 INFO server.Server: Started @12072ms
2022-10-11 08:42:14,220 INFO webapp.WebApps: Web app node started at 8042
2022-10-11 08:42:14,224 INFO nodemanager.NodeStatusUpdaterImpl: Node ID assigned is : a4it55.KC.VNR:62038
2022-10-11 08:42:14,230 INFO util.JvmPauseMonitor: Starting JVM pause monitor
2022-10-11 08:42:14,240 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8031
2022-10-11 08:42:14,380 INFO nodemanager.NodeStatusUpdaterImpl: Registering with RM using containers :[]
2022-10-11 08:42:14,877 INFO security.NMContainerTokenSecretManager: Rolling master-key for container-tokens, got key with id -324326809
2022-10-11 08:42:14,879 INFO security.NMTokenSecretManagerInNM: Rolling master-key for container-tokens, got key with id 123186820
2022-10-11 08:42:14,882 INFO nodemanager.NodeStatusUpdaterImpl: Registered with ResourceManager as a4it55.KC.VNR:62038 with total resource of <memory:8192, vCores:8>
```

To make sure that all services started successfully, Run the following command:

jps

It should display the following services:

14560 DataNode

4960 ResourceManager

5936 NameNode

768 NodeManager

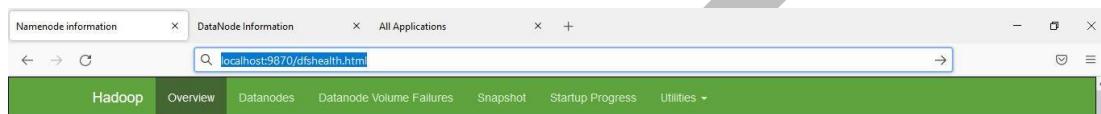
14636 Jps

```
PS D:\hadoop-env\hadoop-3.2.2\sbin> jps
4288 Jps
4896 ResourceManager
4708 NameNode
5852 NodeManager
6860 DataNode
PS D:\hadoop-env\hadoop-3.2.2\sbin>
```

Output

Hadoop Web UI

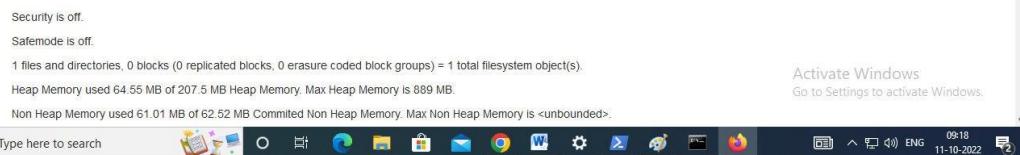
<http://localhost:9870/dfshealth.html>



Overview 'localhost:9820' (active)

Started:	Tue Oct 11 08:35:24 +0530 2022
Version:	3.2.2, r7a3bc90b05f257c8ace2f76d74264906f0f7a932
Compiled:	Sun Jan 03 14:56:00 +0530 2021 by hexiaoqiao from branch-3.2.2
Cluster ID:	CID-d536af7a-e745-4eaf-8c23-158b0783e6db
Block Pool ID:	BP-1402821420-172.16.8.55-1665456682328

Summary



<http://localhost:9864/datanode.html>



DataNode on a4it55.KC.VNR:9866

Cluster ID:	CID-d536af7a-e745-4eaf-8c23-158b0783e6db
Version:	3.2.2, r7a3bc90b05f257c8ace2f76d74264906f0f7a932

Block Pools

Namenode Address	Block Pool ID	Actor State	Last Heartbeat	Last Block Report	Last Block Report Size (Max Size)
localhost:9820	BP-1402821420-172.16.8.55-1665456682328	RUNNING	2s	19 minutes	0 B (64 MB)

Volume Information



<http://localhost:8088/cluster>

The screenshot shows the Hadoop Cluster Management UI at <http://localhost:8088/cluster>. The main title is "All Applications". On the left, there's a sidebar with sections like "Cluster Metrics", "Cluster Nodes Metrics", "Scheduler Metrics", and "Tools". The "Tools" section has a dropdown menu with items: NEW, NEW_SAVING, SUBMITTED, ACCEPTED, RUNNING, FINISHED, FAILED, KILLED, and Scheduler. The main content area displays various metrics tables. One table for "Cluster Metrics" shows: Apps Submitted 0, Apps Pending 0, Apps Running 0, Apps Completed 0, Containers Running 0, Used Resources <memory:0, vCores:0>, Total Resources <memory:8192, vCores:8>, Reserved Resources <memory:0, vCores:0>. Another table for "Scheduler Metrics" shows: Active Nodes 1, Decommissioning Nodes 0, Decommissioned Nodes 0, Lost Nodes 0, Unhealthy Nodes 0. A third table for "Scheduler Metrics" shows: Scheduler Type Capacity Scheduler, Scheduling Resource Type <memory:mb (unit=M), vcores>, Minimum Allocation <memory:1024, vCores:1>, Maximum Allocation <memory:8192, vCores:4>. Below these tables is a table header for "Show 20 entries" with columns: ID, User, Name, Application Type, Queue, Application Priority, StartTime, LaunchTime, FinishTime, State, FinalStatus, Running Containers, Allocated CPU vCores, Allocated Memory MB, Allocated GPUs, Reserved CPU vCores, Reserved Memory MB. At the bottom, it says "Showing 0 to 0 of 0 entries". The status bar at the bottom right shows: 09:20, 11-10-2022.

VIVA QUESTIONS

1. What is Hadoop?

2. State the types of data.

3. Name three components of Hadoop.

4. In Hadoop, _____ is used for processing / computing the task.

5. In Hadoop, _____ provides a distributed file system that is designed to run on commodity hardware.

6. Expansion of YARN.

7. What is YARN?

8. What are the different Hadoop configuration files?

9. Mention different features of HDFS.

10. What are the three modes that Hadoop can run?

RESULT

Thus, the procedure to set up the one node Hadoop cluster was successfully done and verified.

Observation by students: (what student able to?)

1.

CATEGORY	MAX. MARKS ALLOTTED	MARKS AWARDED
PROCEDURAL STEPS	50	
VIVA	25	
RECORD	25	
TOTAL	100	

Ex. No: 6

Date:

Word Count Program using MapReduce

AIM

To count the number of words using JAVA for demonstrating the use of Map and Reduce tasks.

DESCRIPTION

MapReduce is a programming model for writing applications that can process Big Data in parallel on multiple nodes. MapReduce provides analytical capabilities for analysing huge volumes of complex data.

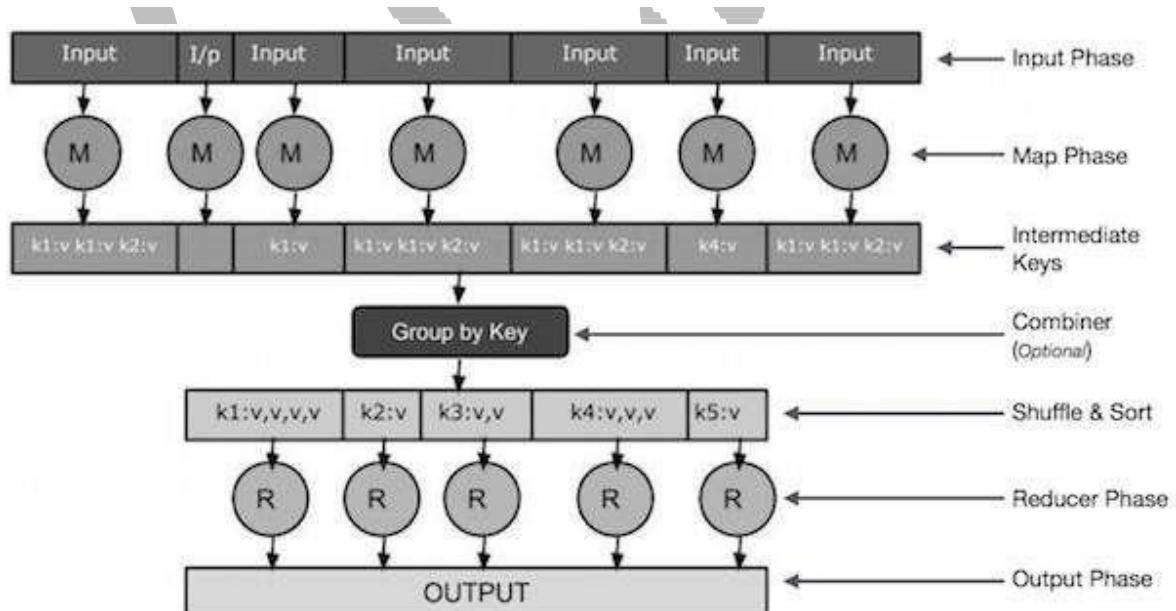
How MapReduce Works?

The MapReduce algorithm contains two important tasks, namely Map and Reduce.

- The Map task takes a set of data and converts it into another set of data, where individual elements are broken down into tuples (key-value pairs).
- The Reduce task takes the output from the Map as an input and combines those data tuples (key-value pairs) into a smaller set of tuples.

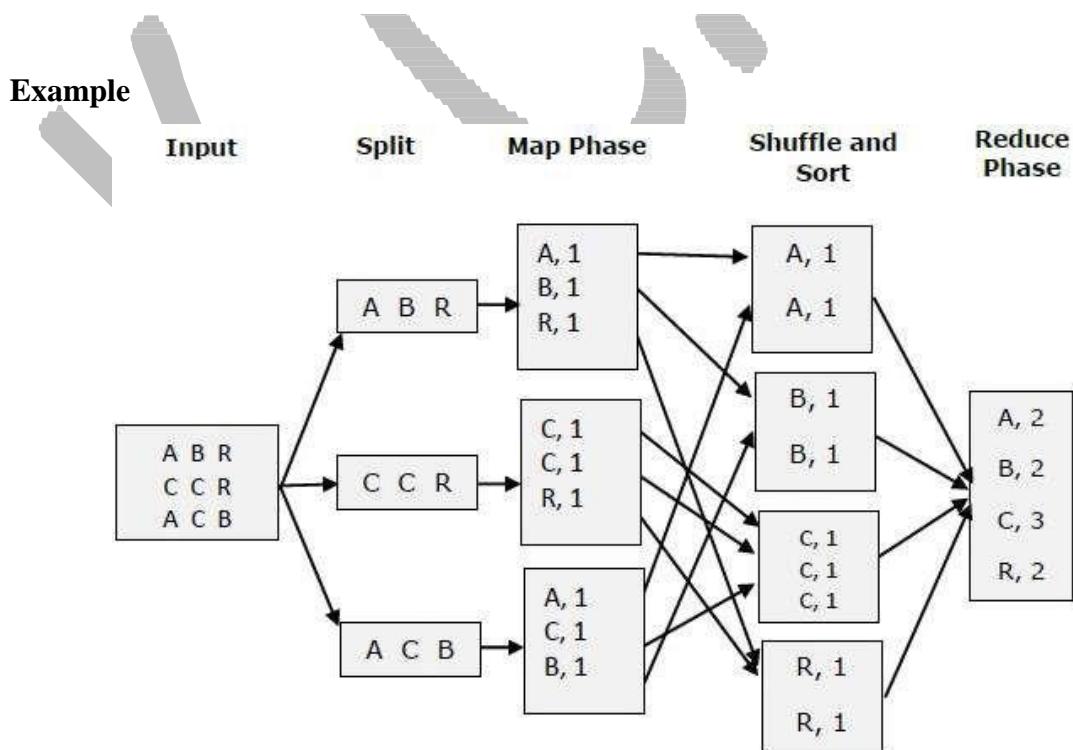
The reduce task is always performed after the map job.

Let us now take a close look at each of the phases and try to understand their significance.



- **Input Phase:** A Record Reader that translates each record in an input file and sends the parsed data to the mapper in the form of key-value pairs.

- **Map:** Map is a user-defined function, which takes a series of key-value pairs and processes each one of them to generate zero or more key-value pairs.
- **Intermediate Keys:** They key-value pairs generated by the mapper are known as intermediate keys.
- **Combiner:** A combiner is a type of local Reducer that groups similar data from the map phase into identifiable sets. It takes the intermediate keys from the mapper as input and applies a user-defined code to aggregate the values in a small scope of one mapper. It is not a part of the main MapReduce algorithm; it is optional.
- **Shuffle and Sort:** The Reducer task starts with the Shuffle and Sort step. It downloads the grouped key-value pairs onto the local machine, where the Reducer is running. The individual key-value pairs are sorted by key into a larger data list. The data list groups the equivalent keys together so that their values can be iterated easily in the Reducer task.
- **Reducer:** The Reducer takes the grouped key-value paired data as input and runs a Reducer function on each one of them. Here, the data can be aggregated, filtered, and combined in a number of ways, and it requires a wide range of processing. Once the execution is over, it gives zero or more key-value pairs to the final step.
- **Output Phase:** In the output phase, an output formatter that translates the final key-value pairs from the Reducer function and writes them onto a file using a record writer.



PROGRAM

WordCount.java

```
import java.io.IOException;
import java.util.StringTokenizer;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.Mapper;
import org.apache.hadoop.mapreduce.Reducer;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
public class WordCount {
    public static class TokenizerMapper
        extends Mapper<Object, Text, Text, IntWritable>{
        private final static IntWritable one = new IntWritable(1);
        private Text word = new Text();
        public void map(Object key, Text value, Context context
            ) throws IOException, InterruptedException {
            StringTokenizer itr = new StringTokenizer(value.toString());
            while (itr.hasMoreTokens()) {
                word.set(itr.nextToken());
                context.write(word, one);
            }
        }
    }
    public static class IntSumReducer
        extends Reducer<Text,IntWritable,Text,IntWritable> {
        private IntWritable result = new IntWritable();
```

```
public void reduce(Text key, Iterable<IntWritable> values,
                  Context context
) throws IOException, InterruptedException {
    int sum = 0;
    for (IntWritable val : values) {
        sum += val.get();
    }
    result.set(sum);
    context.write(key, result);
}
```

```
public static void main(String[] args) throws Exception {
    Configuration conf = new Configuration();
    Job job = Job.getInstance(conf, "word count");
    job.setJarByClass(WordCount.class);
    job.setMapperClass(TokenizerMapper.class);
    job.setCombinerClass(IntSumReducer.class);
    job.setReducerClass(IntSumReducer.class);
    job.setOutputKeyClass(Text.class);
    job.setOutputValueClass(IntWritable.class);
    FileInputFormat.addInputPath(job, new Path(args[0]));
    FileOutputFormat.setOutputPath(job, new Path(args[1]));
    System.exit(job.waitForCompletion(true) ? 0 : 1);
}
```

PROCEDURE STEPS

Step 1: Create a text file “D:/data.txt”

Step 2: Create a directory in HDFS, where to kept text file.

hdfs dfs -mkdir /user

```
Administrator: Windows PowerShell
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -mkdir /user
PS D:\hadoop-env\hadoop-3.2.2\sbin>
```

Step 3: Upload the data.txt file on HDFS in the specific directory

hdfs dfs -put D:/data.txt /user

```
Administrator: Windows PowerShell
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -mkdir /user
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -put D:/data.txt /user
PS D:\hadoop-env\hadoop-3.2.2\sbin>
```

Step 4: List the files or directories in hdfs

hdfs dfs -ls /user/

```
Administrator: Windows PowerShell
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -mkdir /user
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -put D:/data.txt /user
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -ls /user/
Found 1 items
-rw-r--r-- 1 Administrator supergroup      50 2022-10-11 11:34 /user/data.txt
PS D:\hadoop-env\hadoop-3.2.2\sbin>
```

Step 5: To view the content of the file “/user/data.txt”

hdfs dfs -cat /user/data.txt

```
Administrator: Windows PowerShell
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -mkdir /user
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -put D:/data.txt /user
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -ls /user/
Found 1 items
-rw-r--r-- 1 Administrator supergroup      50 2022-10-11 11:34 /user/data.txt
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -cat /user/data.txt
Cloud and Grid Lab. Cloud and Grid Lab. Cloud Lab.
PS D:\hadoop-env\hadoop-3.2.2\sbin>
```

Step 6: Run the jar file

```
hadoop jar D:/hadoop-env/hadoop-3.2.2/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.2.jar wordcount /user /out
```

```
Administrator: Windows PowerShell
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -mkdir /user
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -put D:/data.txt /user
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -ls /user/
Found 1 items
-rw-r--r-- 1 Administrator supergroup      50 2022-10-11 11:34 /user/data.txt
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -cat /user/data.txt
Cloud and Grid Lab. Cloud and Grid Lab. Cloud Lab.
PS D:\hadoop-env\hadoop-3.2.2\sbin> hadoop jar D:/hadoop-env/hadoop-3.2.2/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.2.jar wordcount /user /out
2022-10-11 11:36:37,212 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
2022-10-11 11:36:38,133 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/Administrator/.staging/job_1665467943708_0001
2022-10-11 11:36:38,559 INFO InputFormat: Total input files to process : 1
2022-10-11 11:36:38,864 INFO mapreduce.JobSubmitter: number of splits:1
2022-10-11 11:36:39,157 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1665467943708_0001
2022-10-11 11:36:39,159 INFO mapreduce.JobSubmitter: Executing with tokens: []
2022-10-11 11:36:39,483 INFO conf.Configuration: resource-types.xml not found
2022-10-11 11:36:39,484 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2022-10-11 11:36:48,298 INFO impl.YarnClientImpl: Submitted application application_1665467943708_0001
2022-10-11 11:36:48,328 INFO mapreduce.Job: The url to track the job: http://a4it55:8088/proxy/application_1665467943708_0001/
2022-10-11 11:36:48,322 INFO mapreduce.Job: Running job: job_1665467943708_0001
2022-10-11 11:36:54,677 INFO mapreduce.Job: Job job_1665467943708_0001 running in uber mode : false
2022-10-11 11:36:54,688 INFO mapreduce.Job: map 0% reduce 0%
2022-10-11 11:37:00,861 INFO mapreduce.Job: map 100% reduce 0%
2022-10-11 11:37:07,981 INFO mapreduce.Job: map 100% reduce 100%
2022-10-11 11:37:09,016 INFO mapreduce.Job: Job job_1665467943708_0001 completed successfully
2022-10-11 11:37:09,239 INFO mapreduce.Job: Counters: 54
```

Step 7: To view the output in “/out/*”

```
hadoop fs -cat /out/*
```

```
Bytes Written=28
PS D:\hadoop-env\hadoop-3.2.2\sbin> hadoop fs -cat /out/*
Cloud      3
Grid       2
Lab.       3
and        2
PS D:\hadoop-env\hadoop-3.2.2\sbin>
```

Output

The screenshot shows the HDFS Health Overview page at localhost:9870/dfshealth.html#tab-overview. A red arrow points to the 'Utilities' dropdown menu, which includes options like 'Logs', 'Log Level', 'Metrics', 'Configuration', and 'Process Thread Dump'.

Browse Directory

The screenshot shows the HDFS Browser at localhost:9870/explorer.html/. It displays a list of files in the '/user' directory:

File	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
	drwxr-xr-x	Administrator	supergroup	0 B	Oct 11 11:37	0	0 B	out
	drwx-----	Administrator	supergroup	0 B	Oct 11 11:36	0	0 B	tmp
	drwxr-xr-x	Administrator	supergroup	0 B	Oct 11 11:34	0	0 B	user

The screenshot shows a web-based Hadoop file explorer interface. At the top, there's a navigation bar with links for Overview, Datanodes, Datanode Volume Failures, Snapshot, Startup Progress, and Utilities. Below the navigation bar is a search bar and a toolbar with icons for file operations like copy, move, and delete.

The main area is titled "Browse Directory" and shows a list of files in the "/out" directory. There are two entries:

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	Administrator	supergroup	0 B	Oct 11 11:37	1	128 MB	_SUCCESS
-rw-r--r--	Administrator	supergroup	28 B	Oct 11 11:37	1	128 MB	part-r-00000

Below the table, it says "Showing 1 to 2 of 2 entries". At the bottom right, there are buttons for Previous, Next, and a page number (1).

A modal window titled "File information - part-r-00000" is open. It contains the following details:

- Download
- Head the file (first 32K)
- Tail the file (last 32K)
- Block information - Block 0
- Block ID: 1073741832
- Block Pool ID: BP-789191111-172.16.8.55-1665467846441
- Generation Stamp: 1008
- Size: 28
- Availability:
 - a4it55.KC.VNR
- File contents:

Cloud 3
Grid 2
Lab 3
and 2

VIVA QUESTIONS

1. What is Hadoop MapReduce?

2. What are the operations performed in MapReduce for computation process?

3. State the benefits of MapReduce.

4. What are the main components of MapReduce?

5. Define shuffling in MapReduce.

RESULT

Thus, the numbers of words were counted successfully using Map and Reduce tasks.

Observation by students: (what student able to?)

1.

CATEGORY	MAX. MARKS ALLOTED	MARKS AWARDED
PROCEDURAL STEPS	50	
VIVA	25	
RECORD	25	
TOTAL	100	

Ex. No: 6

Date:

Create Virtual Machine in OpenStack Cloud Platform

AIM

To create Virtual Machine in Online OpenStack Cloud Platform.

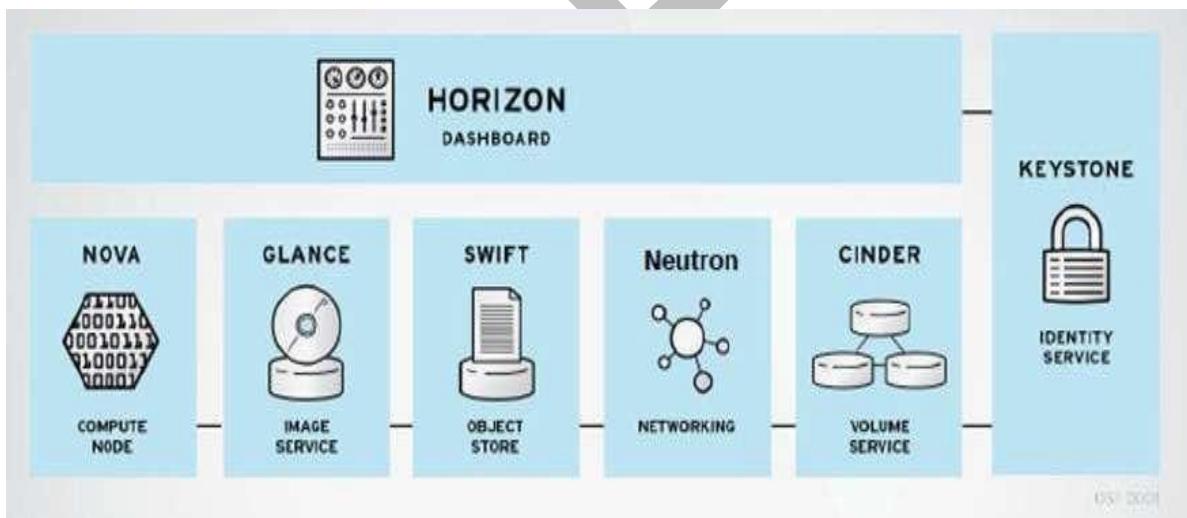
SCENARIO

You are a cloud intern tasked with creating a Virtual Machine (VM) on the OpenStack cloud platform for testing a new application. You will use the OpenStack Horizon Dashboard to set up the VM and configure its resources.

DESCRIPTION

- OpenStack was introduced by Rackspace and NASA in July 2010.
- OpenStack is an Infrastructure as a Service known as Cloud Operating System, that take resources such as Compute, Storage, Network and Virtualization Technologies and control those resources at a data center level
- The project is building an open-source community to share resources and technologies with the goal of creating a massively scalable and secure cloud infrastructure.
- The software is open source and limited to just open-source APIs such as Amazon.

OpenStack Architecture



- It is modular architecture
- Designed to easily scale out
- Based on (growing) set of core services

OPEN STACK COMPONENTS

KEYSTONE

- Identity service
- Common authorization framework
- Manage users, tenants and roles
- Pluggable back-ends (SQL,PAM,LDAP, IDM etc)

NOVA

- Core compute service comprised of
 - Compute Nodes: Hypervisors that run virtual machines
- Supports multiple hypervisors KVM, Xen, LXC, Hyper-V and ESX
 - Distributed controllers that handle scheduling, API calls, etc.
- Native OpenStack API and Amazon EC2 compatible API

GLANCE

- Image service
- Stores and retrieves disk images (Virtual machine templates)
- Supports RAW, QCOW, VHD, ISO, OVF & AMI/AKI
- Back-end Storage: File System, Swift, Gluster, Amazon S3

SWIFT

- Object Storage service
- Modeled after Amazon's Service
- Provides simple service for storing and retrieving arbitrary data
- Native API and S3 compatible API

NEUTRON

- Network service
- Provides framework for Software Defined Network
- Plugin architecture
 - Allows integration of hardware and software-based network solutions
 - Open vSwitch, Cisco UCS, Standard Linux Bridge, NiCira NVP

CINDER

- Block Storage (Volume) service
- Provides block storage for Virtual Machines (persistent disks)
- Like Amazon EBS service

- Plugin architecture for vendor extensions
- NetApp driver for cinder

HORIZON

- Dashboard
- Provides simple self-service UI for end-users
- Basic cloud administrator functions
 - Define users, tenants, and quotas
 - No infrastructure management

HEAT OpenStack Orchestration

- Provides template driven cloud application orchestration
- Modeled after AWS Cloud Formation
- Targeted to provide advanced functionality such as high availability and auto scaling
- Introduced by Redhat

CEILOMETER

- OpenStack Monitoring and Metering
- **Goal:** To Provide a single infrastructure to collect measurements from an entire OpenStack Infrastructure; Eliminate need for multiple agents attaching to multiple OpenStack Projects
- Primary targets metering and monitoring: Provided extensibility

PROCEDURAL STEPS

Step 1: Switch to “root” user

```
user@vmub-hadoop1:~$ sudo su -  
root@vmub-hadoop1:~# sudo apt-get update
```

Step 2: Install “git” and logout from “root user”

```
root@vmub-hadoop1:~# sudo apt-get install git  
root@vmub-hadoop1:~# exit  
logout  
user@vmub-hadoop1:~$
```

Step 3: Create/Add a user “stack” with “no password”

```
user@vmub-hadoop1:~$ sudo useradd -s /bin/bash -d /opt/stack -m stack  
user@vmub-hadoop1:~$ echo "stack ALL=(ALL) NOPASSWD: ALL" | sudo tee  
/etc/sudoers.d/stack  
stack ALL=(ALL) NOPASSWD: ALL
```

Step 4: Move to “stack” user

```
user@vmub-hadoop1:~$ sudo su - stack
```

Step 5: Clone the “devstack” from the specified link

```
stack@vmub-hadoop1:~$ git clone https://github.com/openstack-dev/devstack
```

```
Cloning into 'devstack'...  
remote: Enumerating objects: 49330, done.  
remote: Counting objects: 100% (2556/2556), done.  
remote: Compressing objects: 100% (836/836), done.  
remote: Total 49330 (delta 1831), reused 2226 (delta 1709), pack-reused 46774  
Receiving objects: 100% (49330/49330), 15.56 MiB | 2.33 MiB/s, done.  
Resolving deltas: 100% (34462/34462), done.  
Checking connectivity... done.
```

Step 6: Move to “/devstack/samples”

```
stack@vmub-hadoop1:~$ cd devstack  
stack@vmub-hadoop1:~/devstack$ ls  
clean.sh      doc      functions      gate      lib      openrc      roles      stackrc  tools  
CONTRIBUTING.rst extras.d  functions-common HACKING.rst LICENSE  playbooks  
run_tests.sh  stack.sh  tox.ini  
data          files    FUTURE.rst     inc      Makefile README.rst samples    tests  
unstack.sh  
stack@vmub-hadoop1:~/devstack$ cd samples  
stack@vmub-hadoop1:~/devstack/samples$ ls  
local.conf local.sh
```

Step 7: Copy “local.conf” file to “devstack”

```
stack@vmub-hadoop1:~/devstack/samples$ cp local.conf ../  
stack@vmub-hadoop1:~/devstack/samples$ cd ..  
stack@vmub-hadoop1:~/devstack$ ls  
clean.sh      doc      functions      gate      lib      Makefile  README.rst  samples  
tests        unstack.sh  
CONTRIBUTING.rst  extras.d  functions-common  HACKING.rst  LICENSE  openrc  
roles        stackrc  tools  
data         files    FUTURE.rst     inc      local.conf  playbooks  run_tests.sh  
stack.sh  tox.ini
```

Step 8: Open “local.conf” and edit the lines

```
stack@vmub-hadoop1:~/devstack$ nano local.conf
```

```
ADMIN_PASSWORD=p1  
DATABASE_PASSWORD=p1  
RABBIT_PASSWORD=p1  
SERVICE_PASSWORD=p1
```

```
HOST_IP=10.0.2.4  
FLOATING_RANGE=10.0.2.224/27
```

Step 9: Remove the following files to Lock the error

```
stack@vmub-hadoop1:~/devstack$ sudo rm /var/lib/dpkg/lock  
stack@vmub-hadoop1:~/devstack$ sudo rm /var/lib/apt/lists/lock  
stack@vmub-hadoop1:~/devstack$ sudo rm /var/cache/apt/archives/lock  
stack@vmub-hadoop1:~/devstack$ sudo rm -rf /var/lib/apt/lists/*
```

Step 10: To install stack

```
stack@vmub-hadoop1:~/devstack$ FORCE=yes ./stack.sh
```

```
        print a[2]
    }
' /opt/stack/devstack/local.conf
+./stack.sh:main:1489          set +o xtrace
=====
DevStack Component Timing
(times are in seconds)
=====
run_process      53
test_with_retry   2
apt-get-update    1
osc              177
wait_for_service 21
dbsync           56
pip_install       149
apt-get            7
-----
Unaccounted time 418
=====
Total runtime     884

This is your host IP address: 10.128.0.8
This is your host IPv6 address: ::1
Horizon is now available at http://10.128.0.8/dashboard
Keystone is serving at http://10.128.0.8/identity/
The default users are: admin and demo
The password: StrongAdminSecret

WARNING:
Using lib/neutron-legacy is deprecated, and it will be removed in the future

Services are running under systemd unit files.
For more information see:
https://docs.openstack.org/devstack/latest/systemd.html

DevStack Version: train
Change: 16d11d27f375b8c027bbc3a1db1885e90ce6c604 Merge "Option "lock_path" from group "DEFAULT"
OS Version: Ubuntu 18.04 bionic

2019-06-04 12:19:19.207 | stack.sh completed in 884 seconds.
```

Accessing OpenStack on a web browser

To access OpenStack via a web browser browse your Ubuntu's IP address as shown.

<https://server-ip/dashboard>

This directs you to a login page as shown.

The image shows two screenshots of the OpenStack interface. The top screenshot is the 'Log in' page, featuring the OpenStack logo and fields for 'User Name' and 'Password'. The bottom screenshot is the 'Overview' page for the 'admin' user, displaying resource usage statistics for Compute, Volume, and Network services.

Log in

User Name:

Password:

Sign In

Overview

Limit Summary

Compute	Instances	VCPUs	RAM
Used 0 of 10	Used 0 of 20	Used 0Bytes of 50GB	

Volume	Volume Snapshots	Volume Storage
Used 0 of 10	Used 0 of 10	Used 0Bytes of 1000GB

Network

Network
Used 0 of 10

VIVA QUESTIONS

1. What is OpenStack?

2. List the key components of OpenStack.

3. List the hypervisors supported by OpenStack.

4. What is “Cinder” in OpenStack service?

5. What are the benefits of using OpenStack?

RESULT

Thus, the Virtual Machine is created successfully in Online OpenStack Cloud Platform.

Observation by students: (what student able to?)

1.

CATEGORY	MAX. MARKS ALLOTED	MARKS AWARDED
PROCEDURAL STEPS	50	
VIVA	25	
RECORD	25	
TOTAL	100	
