set A

Page No.
Date

Countineous Assessment 2

CAP 470

course Title: Cloud computing

Name: Jayshri Lal Pandit

Section: D2112

ROII NO.: RD2112 A03

Reg. No. 12111670

(L) Which platforms are used for large Scale cloud computing? Explain it briefly with example? Ans:- The Platforms for large-scale cloud computing are :-(1) Appriche Hadoop:-It is a open source framework cositten 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 single server to thousands of machines, each offering local computation and storage. Hadoop Distributed File System The Hadoop pistorbuted file system (HDFS) is based on the Google File System (GFS) and provides a distributed file system that is designed to run on commodity Mardware. It has many similarities

with existing distributed file systems.

However, the differences from

Page No.

other distributed file systems

are significant. It is highly
fault - tolerant and is designed
to be deployed on low-cost
hardware. It provides hight
throughput access to application
dato 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 jara

 libraries and utilities required

 by other Hadoop modules.
- · Hadoop YARN: This is a framework for job scheduling and chuter resource of management.

How, Does Hadoop, Work

1+ is quite expensive to build
bigger servers with heavy configurations
that handle large scale processing
that handle large scale processing
but as an alternative, you can
tie together many commodity computers
with single-CPV, as a single
functional distributed system and

1

3

10

3

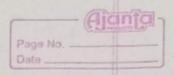
-

-

-

1

practically, the clustered machines can read the dataset in parallel and provide a much higher throughput. Moreover, it is cheaper than one high-end server, so mis is the first motivational factor behind using Hadoop that it runs across clustered and low-cost machines. Hadoop runs code across a cluster of computers. This process included the following core tasks that Hadoop performs: · Data is initially divided into directories and files. Files are divided into Uniform Sized blocks of 128 M and 64M (per preferably 128 M). · These files are then distributed across various cluster nodes for further processing. · HDFS, being on top of the local file system, supervises the processing. · Blocks are reglicented for handling hardware failure. · checking that the code was executed successfully.



- Place between the map and reduce stages.
- · Sending the Sorted data to a certain computer.
- · writing the debugging loss for each job.
- 2.) Map Reduce:

 MapReduce is a programming
 model or pattern cossp cuithin

 the Hadoop framework that is

 wed to access big data stored
 in the Hadoop file System (HDFS)

 . It is a core component,
 integral to the functioning of

 the Hadoop framework.

For example, a Hadoop cluster coith 20,000 inexpensive commodity servers and 256 MB block of data in each, can process around 5 TB of data at the same time. This reduces the processing time as compared to sequential processing of such a large data set.

With mappeduce, rather than sending data to where the

depplication or logic resides, the logic is executed on the serves where the data already resides, to expedite processing.

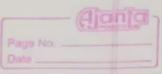
How MapRedyce works

At the Crux of MapReduce are two function: map and Reduce. They are sequenced one after the other.

- The Map function takes input
 from the disk as (key value)
 pairs, processing them, and
 produces another set et
 intermediate (key, value) pairs
 on output.
- inputs as (key, value) pairs, and produces (key, value) pairs as output

 The types of keys and values differ based on the use case.

 All inputs and outputs are stored in the HDFS. While the map is a mandatory step to fifter and sort the initial data, the reduce function is optional.



What is meant by Edge computing cloud ? Ans: In the beginning, there was the Unix era, we learned how to connect to that computer using & dumb terminals Next we had personal computers, which was the first time regular people really owned the hardware that did the work. Right Now, in 2018, we're firmly in the cloud computing. personal computers, but we mostly use them to access centralized services like Dropbox, Comail, Office 365, and Store Stack. Additionally, devices like Amazon Echo, Google chromecast, and

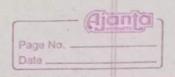
the apple TV are powered by content and itelligence that's in the cloud - as opposed to the DVD box set of little House on the Prairie or CD-ROM copy of Encarta you might ve enjoyed in the personal computing era.

As confordized as this all sounds, the touly amazing thing about cloud computing is that a sessionsly large percentage of all companies in the coord now rely on the infractructure, hosting, machine learning, and compute power of a very select few cloud providers: Amazon, Microsoft, Cropgle, and IBM.

Edge devices can contribute to a cloud, if the Storage and computing capabilities provided by those devices at the endpoints of a newt of network are abstracted, pooled, and shared across a network—essentially becoming part of a larger cloud infrastructure. Edge Computing is not part of a cloud. What makes edge computing so useful is that it is purposefully separate from clouds and cloud computing.

Here's how we see it:

be stored or applications can run.



They are software - defined environments created by datacenters or server farms.

- edges are also places where days is collected. They are physical environments made up to of hardwar outside a dayacenter.
- · Cloud computing is an act:
 the act of running workloads
 in a cloud.
- the act of running workloads on edge devices.

An edge (location) is not the same thing as edge computing (action). Collecting data at the edge of a network and transferring it to a Cloud with minimal (it any) modification is not edge computing - it's just networking. But, if that data is collected and processed at the edge, then it's edge computing.

Edge computing is separate from

clouds 2 main reasons:

10

30

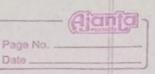
-

30

30

- Dime semsitivity: The rate at which a decision needs to be made doesn't allow for the lay that evould normally take place as data is collected by an edge device, transferred to a central cloud coithout modification and then processed before a decision is sent back to the edge device for execution.
- 2) Data Volume: The sheer volume of data collected is too much to send unaltered to a cloud.
- 3 What is the difference between cloud and traditional datacenters?

Ansi
A Cloud can be described as a term used to describe a group of services either a global or individual network of servers, which processed a unique function. Cloud is not a physical entity, but they are a group or network of remote servers which are as



a single entity for an assigned task.

Types of cloud: -

- methodology that is open to all with the Internet on pay -per - usage method.
- Organisations to build their date centers that are accessible only with the permission of the organisation.
- Combination et public and private clouds. It serves different need et an organisation.
- Services to a group of people in an organisation of a single community.

Data Center

A delta center can be described as a facility / space of networked computers and absociated component (like telecommunications and storage) which helps business and.

Organisations to function a large amount of data These Data Centers allow the data to Organise , process, store and disseminate upon the application wed by businesses.

Types of Data center

- Telecom Data center: 9t is a

 type of data center which are
 operated by telecommunications
 or service providers. It requires
 high-speed connectivity to function
- e Enterprise Data center: 9t is a type of data center which is built and owned by a company that may or may not be onsite.
- * Colocation Data Center: that
 consists of one data centere

 owner place which provides

 cooling to multiple enterprises

 and hyper-scale their customers.

 Hyperscale pata center: data

 center which are owned by and

 operated by the Company itself.

(DIDE)

	Alanta
Page No	

_	Date	Date
	Differences between cloud and	
	Data centera-	
August 1	- X	c — × — — ~
	Cloyd	Data center
(i)	cloud is virtual	Data Center is 9
	resorce that helps	physical resource
	businesses to store,	that helps businesses
	Organize, and operate	to store, organice,
	data efficiently.	and operate and
	3	edficiently.
(11)	The Scalability of	The scalability of
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	the Good required	Data center is
1000	less amount of	hage in investment
(281711	investment.	as compared to
1200,	Corner Strains Start	the cloud.
(111)	The maintenance	The maintenance
27 26	cost is less than	cost is high because
19 NO	Service providers	developers of the
TEANING OF	maintain it.	organisation do
CAN		maintenance.
(14)	Third - Party needs	The Organisation's
	to be trusted	developerà dre
++0	for the organisations	trusted for the
1	data to be stored.	data stored in
(21)	2006	data centers.
(4)	performance es huge	performance ?s
Carlo Carlo Tol	as compared with	loss than compared
Con	investment.	to investment.
(vi)	9+ requires a plan	It is pasily co
	to customize the	customizable without
	Cloud.	any hard plan.