CC Week 3

Week 3

SLA (Service level Agreement) : Formal contract between client and cloud service provider. it is the foundation of consumer's trust in provider

SLO (service level objectives)

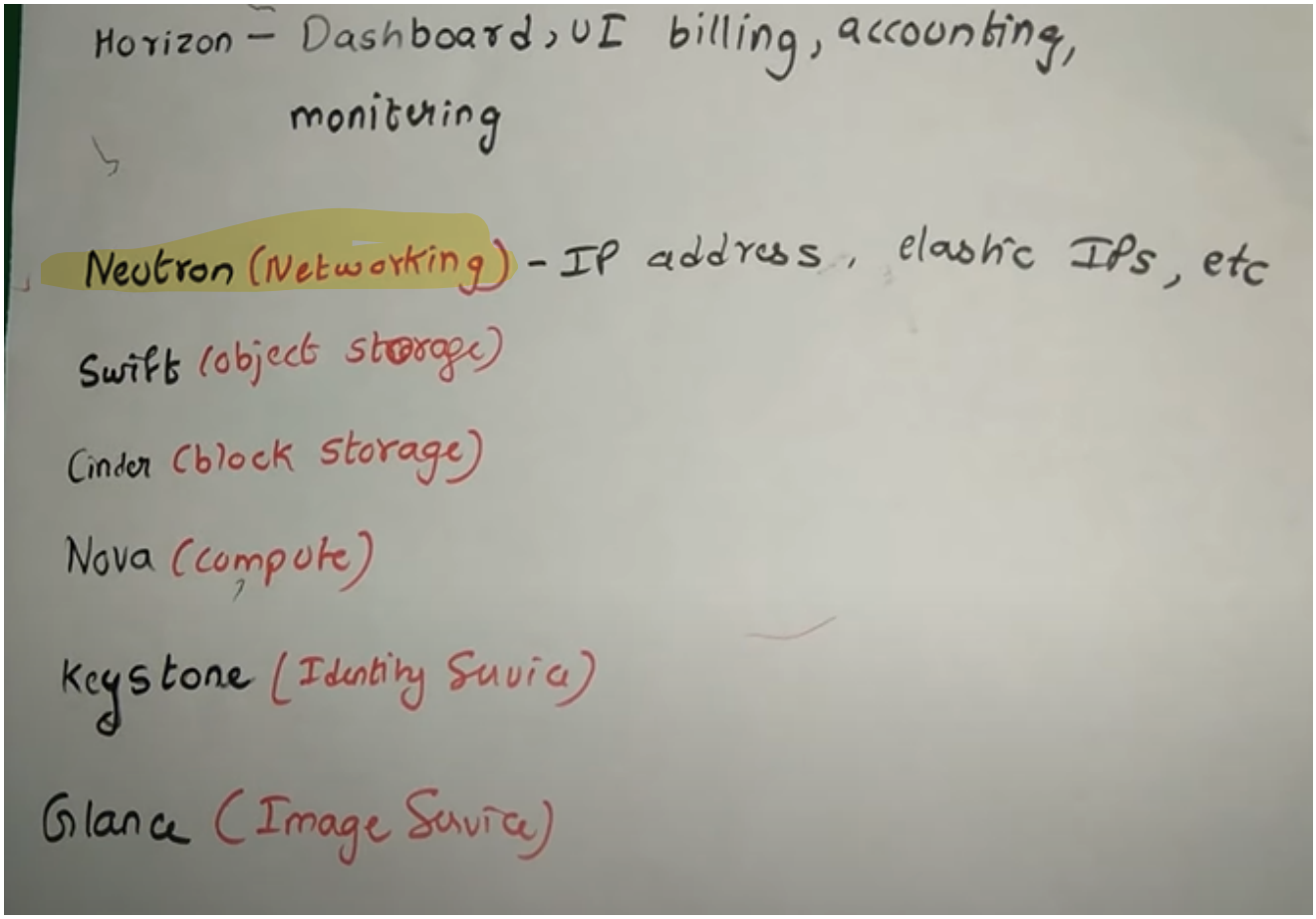
The types of SLAs in the present marketplace are as follows:

1. **Off-the-shelf SLA (Non-negotiable SLA or Direct SLA):** These SLAs are pre-defined by the service provider and are typically not open to negotiation.
2. **Negotiable SLA:** In contrast to off-the-shelf SLAs, negotiable SLAs allow for negotiation between the service provider and the customer.

Cloud Economics:

1. **Value of Common Infrastructure:**
   - Cloud computing shares resources like computers and networks among many users, reducing individual costs.
   - Users benefit from access to powerful computing resources without having to invest heavily in dedicated infrastructure.
2. **Value of Utility Pricing:**
   - Cloud services are often priced like utilities, where users pay based on their usage.
   - This pay-as-you-go model allows users to avoid large upfront investments and instead pay for what they use, similar to how they pay for electricity or water.
3. **Value of on-Demand Services:**
   - Cloud computing offers instant access to a variety of services, such as storage and software, whenever needed.
   - Users can quickly provision resources without the delays associated with traditional procurement processes, enabling faster innovation and adaptation.
4. **Penalty Costs for Exponential Demand:**

- During sudden spikes in demand, traditional IT infrastructure may struggle to scale, leading to downtime or performance issues.
- Cloud computing provides elasticity, allowing resources to scale automatically to meet demand fluctuations.
- While there may be additional costs associated with scaling during peak periods, it's often more cost-effective than dealing with downtime or poor performance.



QUESTION 4:

Horizon is a _____ self-service portal to interact with underlying OpenStack services

a) mobile based
b) OS based
c) web based
d) None of the above

Correct Option: c
Detailed Solution: Horizon provides a web-based self-service portal to interact with underlying OpenStack services, such as launching an instance, assigning IP addresses and configuring access controls

(a) **Google File System (GFS):**

- distributed file system developed by Google to provide efficient storage for large-scale distributed data processing applications.

- GFS is designed for high-throughput access to large files and is optimized for sequential I/O operations rather than random access.
- It employs a master-slave architecture with a single master node coordinating file operations across multiple chunk servers.

(b) **BigTable:**

- Bigtable is a distributed storage system developed by Google for managing structured data.
- It is designed to handle massive datasets with high scalability and low latency.
- It employs a distributed storage model where data is partitioned and stored across multiple servers.

(c) **Dynamo:**

- Dynamo is a highly available and scalable distributed data store developed by Amazon.
- It is designed to provide predictable performance and seamless scalability for large-scale applications.
- Dynamo uses the Quorum protocol for managing concurrent reads and writes, ensuring consistency and availability in the face of network partitions and node failures.
- The Quorum protocol in Dynamo involves a form of decentralized consensus among replicas to determine the validity of read and write operations.

Dynamo

- Developed by Amazon
- Supports large volume of concurrent updates, each of which could be small in size
  - Different from BigTable: supports bulk reads and writes

QUESTION 1:

Which of the following system/ architecture follow(s) Quorum protocol for a large number of concurrent reads & writes?

(a) Google File System (GFS)
(b) BigTable
(c) Dynamo
(d) None of the above

Correct Option: c

Detailed Solution: Dynamo follows Quorum protocol for a large number of concurrent reads & writes.

OpenStack Storage Concepts

- **Ephemeral storage:**
  - Persists until VM is terminated
  - Accessible from within VM as local file system
  - Used to run operating system and/or scratch space
  - Managed by Nova
- **Block storage**:
  - Persists until specifically deleted by user
  - Accessible from within VM as a block device (e.g. /dev/vdc)
  - Used to add additional persistent storage to VM and/or run operating system
  - Managed by Cinder
- **Object storage:**
  - Persists until specifically deleted by user
  - Accessible from anywhere
  - Used to add store files, including VM images
  - Managed by Swift

persistent

QUESTION 2:

Statement 1: In ephemeral storage, the stored objects persist until the VM is terminated.

Statement 2: The ephemeral storage is managed by Cinder in OpenStack.

(a) Statement 1 is TRUE, Statement 2 is FALSE

(b) Statement 2 is TRUE, Statement 1 is FALSE

(c) Both statements are TRUE

(d) Both statements are FALSE

Correct Answer: a

Detailed Solution: Ephemeral storage is managed by NOVA in OpenStack.

horizon dashboard
nova compute
cinder block
swift object
neutron networking
glance image
placement resource
manila file

Data Storage Techniques

- Row-oriented storage
  - Optimal for write-oriented operations viz. transaction processing applications
  - Relational records: stored on contiguous disk pages
  - Accessed through indexes (primary index) on specified columns
  - Example: B^+- tree like storage
- Column-oriented storage
  - Efficient for data-warehouse workloads
    - Aggregation of **measure** columns need to be performed based on values from **dimension** columns
    - Projection of a table is stored as sorted by dimension values
    - Require multiple "join indexes"
      - If different projections are to be indexed in sorted order

QUESTION 3:

Column-oriented storage is efficient for data-warehouse workloads.

(a) TRUE
(b) FALSE

Correct Answer: a
Detailed Solution: From the definition of data storage techniques. (Slide no. 6 of Cloud Computing: Managing Data)

The correct formula for calculating parallel efficiency (Eff) of an algorithm when a task takes time T in a uniprocessor system, P is the number of processors, and M is the time taken by each processor is: `Eff = T / (P * M)`

- measure how effectively an algorithm can utilize multiple processors in a parallel computing environment.

QUESTION 5:

What is the parallel efficiency (Eff) of an algorithm, when a task takes time T in uniprocessor system, P is number of processors, M is time taken by each processor?

(a) Eff = (T*P)/M

(b) Eff = T*(M/P)

(c) Eff = T*P*M

(d) Eff = T/(P*M)

Correct Answer: d
Detailed Solution: Eff = T/(P*M) is the parallel efficiency(Eff) of an algorithm.

The formula to calculate uptime given the downtime and availability of a service is:

$$\text{Uptime} = \frac{\text{Downtime}}{1-\text{Availability}}$$

QUESTION 6:

In cloud, service downtime is 30 minutes and availability of the service is 0.80. What is the service uptime?

(a) 120 minutes
(b) 60 minutes
(c) 150 minutes
(d) 135 minutes

Correct Option: c
Detailed Answer: Availability = 1 – (downtime/uptime).
Uptime = Downtime/(1-Availability) = 30/(1-0.8) = 150 minutes

QUESTION 7:

Which of the following is/are NOT SLA requirement(s) of PaaS cloud delivery model?

a. Data Retention and Deletion
b. Privacy
c. Machine-Readable SLAs
d. Certification

Correct Answer: a,c
Detailed Solution: Data Retention and Deletion and Machine-Readable SLAs are not SLA requirements with respect to Paas cloud delivery model.

Metrics for Monitoring and Auditing

- **Throughput** – How quickly the service responds
- **Availability** – Represented as a percentage of uptime for a service in a given observation period.
- **Reliability** – How often the service is available
- **Load balancing** – When elasticity kicks in (new VMs are booted or terminated, for example)
- **Durability** – How likely the data is to be lost
- **Elasticity** – The ability for a given resource to grow infinitely, with limits (the maximum amount of storage or bandwidth, for example) clearly stated
- **Linearity** – How a system performs as the load increases

donot get confused between availablilty and reliability

QUESTION 8:

What does the 'availability' metric represent in the monitoring and auditing of SLAs?

a) The speed at which a service responds

b) How often the service is available

reliability = how often

c) The ability for a resource to grow infinitely

d) The percentage of uptime for a service

Correct Answer: d
Detailed Solution: availability is represented as a percentage of uptime for a service in a given observation period.

Parallel Database Architectures

- Shared memory
  - Suitable for servers with multiple CPUs
  - Memory address space is shared and managed by a symmetric multi-processing (SMP) operating system
  - SMP:
    - Schedules processes in parallel exploiting all the processors
- Shared nothing
  - Cluster of independent servers each with its own disk space
  - Connected by a network
- Shared disk
  - Hybrid architecture
  - Independent server clusters share storage through high-speed network storage viz. NAS (network attached storage) or SAN (storage area network)
  - Clusters are connected to storage via: standard Ethernet, or faster Fiber Channel or Infiniband connections

---

QUESTION 9:

What architecture is used in a parallel database for the efficient execution of SQL queries?

a) Shared memory architecture
b) Shared disk architecture
c) Shared nothing architecture
d) Shared cache architecture

Correct Answer: c
Detailed Solution: For shared-nothing architecture in the parallel database, tables are partitioned and distributed across multiple processing nodes and SQL optimizer handles distributed joins

OpenStack:

NASC NGIC MP

- **N**: Nova (Compute)
- **A**: Keystone (Identity)
- **S**: Swift (Object Storage)
- **C**: Cinder (Block Storage)
- **N**: Neutron (Networking)
- **G**: Glance (Image Service)
- **I**: Ironic (Bare Metal Provisioning)
- **C**: Ceilometer (Telemetry)
- **M**: Manila (Shared File System)
- **P**: Placement (Resource Allocation)

horizon = dashboard

QUESTION 10:

_____ is used for networking services in OpenStack.

a) Keystone
b) Neutron
c) Cinder
d) Swift

Correct Answer: b
Detailed Solution: Neutron is used for networking services in OpenStack.

QUESTION 4:

Horizon is a _____ self-service portal to interact with underlying OpenStack services

a) mobile based
b) OS based
c) web based
d) None of the above

Correct Option: c
Detailed Solution: Horizon provides a web-based self-service portal to interact with underlying OpenStack services, such as launching an instance, assigning IP addresses and configuring access controls

QUESTION 8:

What is/ are the expected SLA parameters in Storage-as-a-Service?

(a) Geographic Location

(b) Fault tolerance

(c) Customizability

(d) Response time

Correct Answer: a, b

Detailed Solution: *Expected SLA parameters: Geographic location, scalability, storage space, storage billing, security, privacy, backup, fault tolerance/resilience, recovery, system throughput, transferring bandwidth, data life cycle management*

QUESTION 9:

In a system, 2500 unit workloads have been added. What will be the penalty?

(a) 25

(b) 100

(c) 50

(d) 2499

penalty percent is sqrt(N)

Correct Answer: c

Detailed Solution: *Adding n independent demands reduces the C_v by 1/sqrt(n) which in turn results penalty to 1/sqrt(n).*

Cv is ratio of SD to Mean

Adding n independent demands reduces the C_v by $\frac{1}{\sqrt{n}}$

- Penalty of insufficient/excess resources grows smaller
- Aggregating 100 workloads bring the penalty to 10%

When is cloud cheaper than owning? U is lessss than P/A

$C_T < B_T \rightarrow A \times U \times B \times T < P \times B \times T$

$\rightarrow U < \frac{P}{A}$

- When utility premium is less than ratio of peak demand to Average demand

QUESTION 10:

When utility premium is greater than the ratio of peak demand to Average demand, then the cloud is cheaper than owing.

(a) TRUE

(b) FALSE

Correct Answer: b

Detailed Solution: *When utility premium is less than ratio of peak demand to Average demand, the cloud is cheaper than owing.*

QUESTION 1:

Which of the following is false about SLA contents

a) A complete, specific definition of each service
b) An auditing mechanism to monitor the services
c) The responsibilities of the service provider and consumer
d) None of these

Correct Answer: d

Detailed Solution: All of these are part of SLA

QUESTION 2:

In cloud, computing resources are allocated

a) Through a centralized service discovering agent
b) By service level agreement
c) Globally without any centralized agent
d) All of them

Correct Answer: c

Detailed Solution: Cloud related resources are allocated and distributed globally without any central directory or agent

QUESTION 4:

Statement 1: Coefficient of variance represents the ratio of standard deviation to mean.

Statement 2: Coefficient of variance is a useful statistic for comparing the degree of uniformity from one data series to another.

a) Only statement 1 is true
b) Only statement 2 is true
c) Both the statements are true
d) Both the statements are False

Correct Answer: a

Detailed Solution: Coefficient of variance is a useful statistic for comparing the degree of variation from one data series to another. Hence, the correct option is (a).

QUESTION 5:

Say a cloud service is well maintaining it's 95% availability as mentioned in Service Level Agreement. However, still the services are not performing as expected when the load increases abruptly. The problem can be referred to as an issue of

a) Load balancing
b) Linearity
c) Reliability
d) Durability

Correct Answer: b

Detailed Solution: Linearity ensures how system performs when load increases.

D(t) – Instantaneous Demand at time t

R(t) – Resources at time t

$$\text{Penalty Cost } \alpha \int |D(t) - R(t)|\,dt$$

- *If demand is flat, penalty = 0*
- *If demand is linear periodic provisioning is acceptable*

QUESTION 6:

Which of the following is true about penalty, where D(t) is instantaneous demand at time t and R(t) is response at time t?

a) Penalty $\propto \int |R(t) * D(t)|\,dt$
b) Penalty $\propto \int |D(t) + R(t)|\,dt$
c) Penalty $\propto \int |D(t) - R(t)|\,dt$

QUESTION 7:

It is absolutely beneficial to purchase a cloud service when

a) Utility premium < the peak to average demand ratio
b) Utility premium > the peak to average demand ratio
c) Utility premium <= the peak to average demand ratio
d) Utility premium >= the peak to average demand ratio

Correct Answer: a

Detailed Solution: It is always better to purchase cloud service when the utility premium is less than the peak demand to average demand ratio. In case they are equal, we can go for either cloud or self-owned system.

Parallel Efficiency

- If a task takes time *T* in uniprocessor system, it should take *T/p* if executed on *p* processors
- Inefficiencies introduced in distributed computation due to:
  - Need for synchronization among processors
  - Overheads of message communication between processors
  - Imbalance in the distribution of work to processors
- *Parallel efficiency* of an algorithm is defined as:

$$\epsilon = \frac{T}{p\,T_p}.$$

parallel eff

t/(p*m)

t/(p*tp)

QUESTION 8:

If R_t is the ratio of execution time taken by a process in sequential and parallel architecture, and p is the number of parallel computing nodes available, then the parallel efficiency is

a) R_t / p Rt = T/Tp so
b) p / R_t
c) None of the above
d) $p \times R_t$

Correct answer: a

Detailed Solution: Parallel efficiency is (time taken in for sequential execution) / (no of processor x time taken in parallel execution)

QUESTION 9:

Statement 1: Hadoop Distributed File System is an open source implementation of GFS architecture.

Statement 2: In Google File System, the master maintains regular communication with the chunk servers.

a) Only statement 1 is true
b) Only statement 2 is true
c) Both the statements are true
d) Both the statements are false

Correct Answer: c

Detailed Solution: Hadoop Distributed File System is an open source implementation of GFS architecture. In the Google File System, the master maintains regular communication with the chunk server. This is better known as heartbeat messages. In case of a failure, chunk server metadata gets updated to reflect failure.

QUESTION 10:

Which of the following is true?

a) IaaS can be built on PaaS
b) PaaS can be built on SaaS
c) IaaS can be built on SaaS
d) None of the above

Correct Answer: d

Detailed Solution: PaaS can be built on IaaS, and SaaS can be built on PaaS.