

Virtualization and Cloud Computing

Unit-3

2 Marks Question

Q.1. What is capacity planning?

Ans:

- Capacity planning involves analyzing current resource usage, forecasting future demand, and identifying potential constraints or bottlenecks.
- It is important in cloud computing, where resources are shared and must be allocated dynamically based on demand.
- Capacity planning helps organizations optimize the utilization of their infrastructure and minimize costs while ensuring they have the resources they need to meet their business objectives.
- Capacity planning involves considering factors such as workload patterns, peak usage periods, and growth projections.

Q.2. Explain steps in Capacity planning.

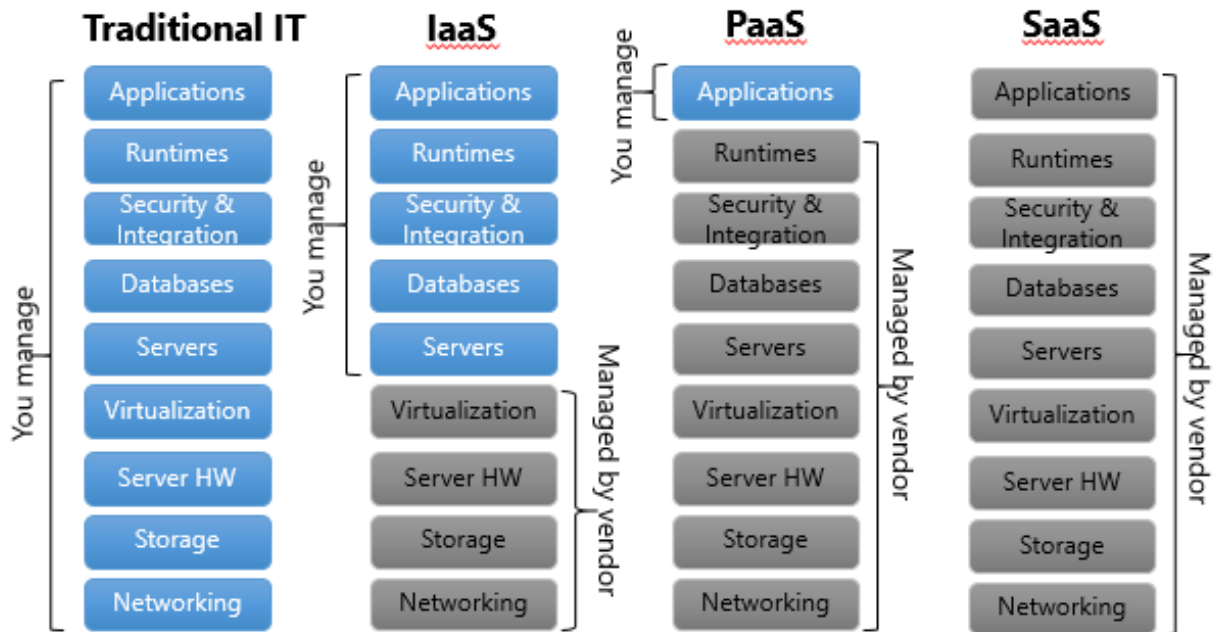
Ans:

- Determine the distinctiveness of the present system.
- Determine the working load for different resources in the system such as CPU, RAM, network, etc.
- Load the system until it gets overloaded; & state what's requiring to uphold acceptable performance.
- Predict the future based on older statistical reports & other factors.
- Deploy resources to meet the predictions & calculations.
- Repeat step (i) through (v) as a loop.

10 Marks Question

Q.1. Explain Cloud Computing Stack.

Ans:



1. Composability:

Composability is a system design principle that deals with the inter-relationships of components. A highly composable system provides components that can be selected and assembled in various combinations to satisfy specific user requirements.

2. Infrastructure:

Virtual servers described in terms of a machine image or instance have characteristics that often can be described in terms of real servers delivering a certain number of microprocessor (CPU) cycles, memory access, and network bandwidth to customers.

3. Platforms:

Provisioning various platforms to users to customize and develop applications. Development, testing, deployments are made easier through this medium.

4. Virtual Appliances:

The machines that are installed in order to run services in cloud. These are platform instances in particular that can be provisioned to cloud users.

5. Communication Protocols:

Common XML based set of protocols used as the messaging format are the Simple Object Access Protocol (SOAP) protocol as the object model, and a set of discovery and description protocols based on the Web Services Description Language (WSDL) to manage transactions in cloud.

6. Applications:

Services running in the cloud

Q.2. Explain Workload Distribution Architecture and its working.

Ans:

- Workload distribution architecture uses IT resources that can be horizontally scaled with the use of one or more identical IT resources.
- This is accomplished through the use of a load balancer that provides runtime logic which distributes the workload among the available IT assets evenly.
- This model can be applied to any IT resource and is commonly used with; distributed virtual servers, cloud storage devices, and cloud services.
- In addition to a load balancer and the previously mentioned resources, the following mechanisms can also be a part of this model:
- Cloud Usage Monitor that can carry out run-time tracking and data processing.
- Audit Monitor used for monitoring the system as may be required to fulfill legal requirements. Hypervisor which is used to manage workloads and virtual hosts that require distribution. Logical network perimeter which isolates cloud consumer network boundaries. Resource clusters commonly used to support workload balancing between cluster nodes.
- Resource replication which generates new instances of virtualized resources under increased workloads.

Working of Workload Distribution Architecture:-

- Resource A and resource B are exact copies of the same resource.
- Inbound requests from consumers are handled by the load balancer which forwards the request to the appropriate resource dependent on workload being handled by each resource.

- In other words, if resource A is busier than resource B, it will forward the resource request to resource B.
- In this manner this model distributes the load among the available IT resources based on workload of each resource.

Q.3. Explain terms in Cloud Architecture.

Ans:

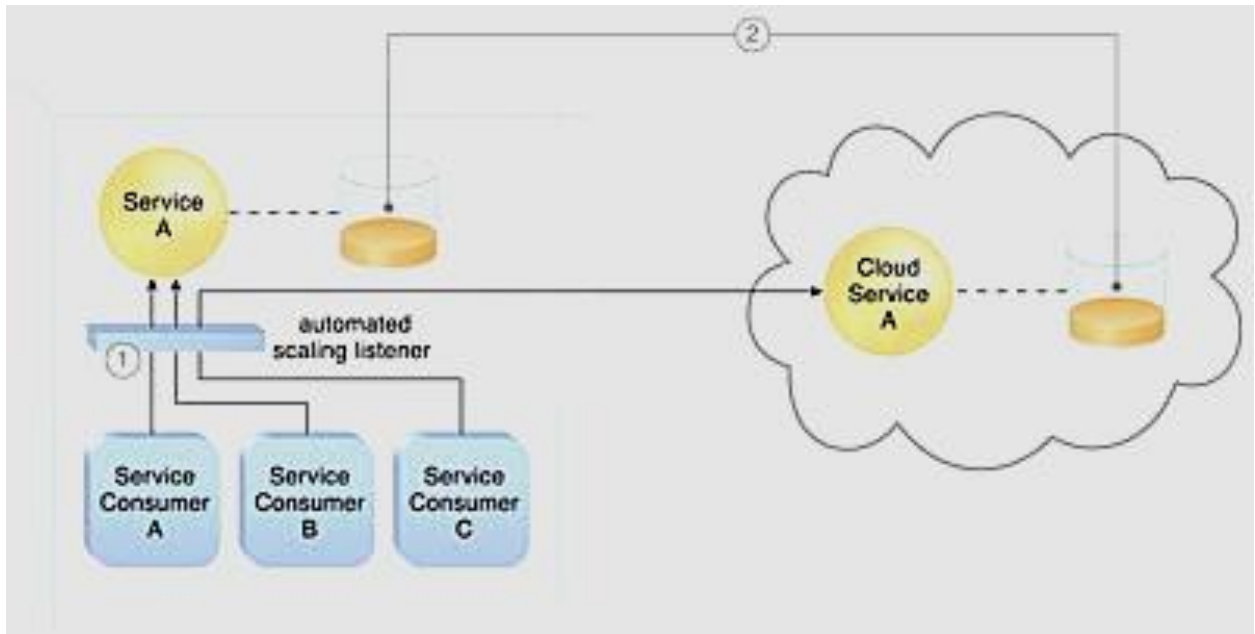
- Audit Monitor – When distributing runtime workloads, the type and geographical location of the IT resources that process the data can determine whether monitoring is necessary to fulfill legal and regulatory requirements.
- Cloud Usage Monitor – Various monitors can be involved to carry out runtime workload tracking and data processing.
- Hypervisor – Workloads between hypervisors and the virtual servers that they host may require distribution.
- Logical Network Perimeter – The logical network perimeter isolates cloud consumer network boundaries in relation to how and where workloads are distributed.
- Resource Cluster – Clustered IT resources in-active/active mode are commonly used to support workload balancing between different cluster nodes.
- Resource Replication – This mechanism can generate new instances of virtualized IT resources in response to runtime workload distribution demands

Q.4. Explain Cloud Bursting Architecture.

Ans:

- The cloud bursting architecture establishes a form of dynamic scaling that scales or “bursts out” on-premise IT resources into a cloud whenever predefined capacity thresholds have been reached. The corresponding cloud-based IT resources are redundantly pre-deployed but remain inactive until cloud bursting occurs. After they are no longer required, the cloud-based IT resources are released and the architecture “bursts in” back to the on-premise environment.

- Cloud bursting is a flexible scaling architecture that provides cloud consumers with the option of using cloud-based IT resources only to meet higher usage demands. The foundation of this architectural model is based on the automated scaling listener and resource replication mechanisms.



Q.5. Explain Elastic Disk Provisioning architecture.

Ans: A request is received from a cloud consumer, and the provisioning of a new virtual server instance begins

As part of the provisioning process, the hard disks are chosen as dynamic or thin-provisioned disks.

The hypervisor calls a dynamic disk allocation component to create thin disks for the virtual server.

(3) Virtual server disks are created via the thin-provisioning program and saved in a folder of near-zero size.

(4) The size of this folder and its files grow as operating applications are installed and additional files are copied onto the virtual server.

(5) The pay-per-use monitor tracks the actual dynamically allocated storage for billing purposes.

The following mechanisms can be included in this architecture in addition to the cloud storage device, virtual server, hypervisor, and pay-per-use monitor:

- Cloud Usage Monitor – Specialized cloud usage monitors can be used to track and log storage usage fluctuations.
- Resource Replication – Resource replication is part of an elastic disk provisioning system when conversion of dynamic thin-disk storage into static thick-disk storage is required.

Q.6. Explain Resource pooling architecture.

Ans: A resource pooling architecture is based on the use of one or more resource pools, in which identical IT resources are grouped and maintained by a system that automatically ensures that they remain synchronized.

- Physical server pools are composed of networked servers that have been installed with operating systems and other necessary programs and/or applications and are ready for immediate use.
- Virtual server pools are usually configured using one of several available templates chosen by the cloud consumer during provisioning. For example, a cloud consumer can set up a pool of mid-tier Windows servers with 4 GB of RAM or a pool of low-tier Ubuntu servers with 2 GB of RAM.
- Storage pools, or cloud storage device pools, consist of file-based or block-based storage structures that contain empty and/or filled cloud storage devices.
- Network pools (or interconnect pools) are composed of different preconfigured network connectivity devices. For example, a pool of virtual firewall devices or physical network switches can be created for redundant connectivity, load balancing, or link aggregation.
- CPU pools are ready to be allocated to virtual servers, and are typically broken down into individual processing cores.
- Pools of physical RAM can be used in newly provisioned physical servers or to vertically scale physical servers.

Q.7. Explain SLA Management System.

Ans:

- SLA management systems are used to manage service-level agreements (SLAs) between cloud providers and their customers.
- SLAs define the expected levels of service quality for cloud services, including factors such as uptime, performance, and support.
- SLA management systems help providers monitor and track their performance against SLA commitments.
- These systems typically include monitoring and reporting tools that provide real-time visibility into service performance and help identify areas where SLA commitments may be at risk.
- SLA management systems also typically include mechanisms for customer feedback and support, allowing customers to report issues and receive assistance in resolving them.
- In addition, these systems may include automated escalation and notification features that alert providers and customers when SLA thresholds are breached.
- SLA management systems are important for maintaining customer trust and ensuring that cloud services meet the needs of the business.
- Providers should regularly review and update SLAs as necessary to reflect changes in their services or customer requirements.
- Proper SLA management can help providers retain customers, improve service quality, and mitigate financial risks associated with SLA breaches.

Q.8. Explain SOA and benefits of SOA.

Ans:

SOA, or service-oriented architecture, defines a way to make software components reusable via service interfaces. These interfaces utilize common communication standards in such a way that they can be rapidly incorporated into new applications without having to perform deep integration each time.

SOA (Service-Oriented Architecture) offers the following benefits:

- Faster application development by reusing service interfaces.
- Ability to use legacy functionality in new markets.
- Improved collaboration between business and IT by defining services in business terms.

- Overall, SOA enables businesses to be more agile and responsive to new opportunities, while improving collaboration and efficiency in software development.

Q.9 Explain Cloud Computing Architecture and its benefits.

Ans:

Cloud Computing Architecture:

- Cloud computing architecture typically consists of front-end, back-end, and cloud-based infrastructure layers.
- The front-end layer includes the user interface and application software.
- The back-end layer includes the cloud storage and computing resources.
- The cloud-based infrastructure layer includes the servers, network, and other components that enable cloud services.

Benefits of Cloud Computing Architecture:

- **Scalability:** Cloud computing architecture allows organizations to easily scale up or down based on their changing resource requirements.
- **Cost-effective:** Cloud computing architecture enables organizations to pay only for the resources they use, rather than investing in and maintaining their own IT infrastructure.
- **Improved efficiency:** Cloud computing architecture can streamline processes and improve efficiency by enabling users to access resources and services from anywhere, at any time.
- **Greater flexibility:** Cloud computing architecture enables users to work remotely, collaborate with others more easily, and access resources from a variety of devices.
- **Enhanced security:** Cloud computing architecture typically includes robust security features and protocols to protect against cyber threats.
- **Reduced environmental impact:** Cloud computing architecture can reduce the environmental impact of IT infrastructure by enabling more efficient use of resources and reducing energy consumption.