SMVS
Page No.
Date : / /

Virtual Machines for Cloud Infrastructure

* Infrastructure as a Service Anatomy:

=> A Cloud Infrastructure is a comprehensive system composed of interconnected hardware, software and networking components.

1 Hardware Infrastructure:

Hardware Infrastructure includes servers, Storage, Networking and Power.

Servers are physical or virtual and provides the processing power.

Networking ensures connectivity within the cloud environment and provides communication.

2 Virtualization:

Virtual Machine ~~is~~ is isolated computing environment that run on physical server.

SMVS
Page No.
Date : / /

Hypervisors are software layer that enable the creation and Management of virtual Machine.

3 Network Infrastructure:

It includes Load Balancers that distribute incoming network traffic across multiple server.

Firewalls is a security devices that monitor and control incoming and outgoing traffic.

Virtual Private Network is secure connection that allow remote users to access the cloud.

4 Software Infrastructure:

Software Infrastructure includes,

- Operating System
- Cloud Management Platforms
- DBMS
- Middleware

SMVS
Page No.
Date : / /

=> Types:

1 Public Iaas:

Resources are shared among multiple customers and infrastructure is managed by service provider.

Ex. Amazon Web Service

2 Private Iaas:

Resources are dedicated to a single organization and hosted by third-party provider.

=> Characteristics:

1 Scalability: Ability to adjust resources based on demand.

2 Cost-Efficiency: Pay-as-you-go Pricing model.

3 Flexibility

4 Accessibility

5 Virtualization and Performance

SMVS
Page No.
Date : / /

* Distributed Management for Virtual Infrastructure.

=> Distributed means managing virtual resources across multiple location.

Distributed virtual infrastructure are interconnected virtual environments across various sites.

This system is use Load Balancer to distribute workload to optimize performance.

Distributed system provides continuous operation during failures or disasters.

=> Advantages :

1 Improved Performance

2 Increased Availability

3 Enhanced Scalability

4 Cost Reduction

SMVS
Page No.
Date : / /

5. Simplified Management

6. Robust monitoring and Reporting

7. Clear Defination of Roles

8. Strong Security Measures.

=> Challenges:

1. Network Latency: Potential impact on performance due to geographical sepraction.

2. Complexity: Required Sepecial Skills

3. Security

4. Data Consistency: Essential to maintain data integrity across different sites.

* Load Balancing Techniques in Distributed system:

=> In Distributed system, Load Balancing technique is used to distribute

Date : / /

traffic across multiple servers.

It improves servers performance, enhanced scalability and resource optimization.

=> Types :

1 Static Load Balances:

This Load Balances are based on predefined rules of Load Balances.

This are method used in Load Balance.

- Round Robin: Distributes requests sequentially across servers in a circular manner.
- Weighted Round Robin: Assign weights to servers based on their capacity.
- IP Hash: Uses the client's IP address to consistently route requests to the same server.

Page No.
Date: / /

2 Dynamic Load Balancing:

This Load Balances are work based on servers real-time condition.

· Method:

· Least Connections: Routes traffic to the server with the fewest active connection.

· Weighted Least Connection: Consider server weighte capacity.

· Least Response Time: Directs request to the server with shortest average response time.

· Resource-Based: Monitors servers resources and distribute traffic based on e resource availability.

* Scheduling Method for Advance Reservation of Capacity:

=> Advance Reservation of specific resources ~~gura~~ guarantee to you to access resources predetermined time.

SMVS
Page No.
Date : / /

=> Methods :

1 BackFilling :

It is prioritizes scheduled reservation.

Fills unused time sot slots with other jobs before reservation starts.

This can lead to less efficient use of resources.

2 Preemption :

It makes room for high-priority jobs by pausing lower-priority one

It can disrupt ongoing jobs and affect fairness.

3 Virtual Advance Reservation for Queues :

Reserves resources in advanse based on predicted wait time.

It predicats wait times to better schedule jobs.

4 Planning Based Approaches:

Create detailed schedules to maximize efficiency and meet reservation.

Can be complex and slow.

5 Lease-Based Models:

Allocates resources based on different types of lease.

Manage resources through a negotiation model.

=> Challenges:

1 Resource Utilization

2 Fairness

3 Overhead

4 Job Turnaround Time

SMVS
Page No.
Date : / /

* SLA Commitment in Cloud Computing

=> In Cloud Computing, Capacity Management is essential to ensure that Service Level Agreements are met.

The goal is to balance resource utilzation with service availability.

Capacity Planning involves determining the optimal amount of resources need to met projected demand.

SLA continiously adjusting resource allocation to match current workload.

=> Capacity Management Process :

1 Capacity Planning:

Evaluate how predictable workload pattern are, and defined SLA for service level.

Determining the necessary resources based on workload and SLAs.

Finding the most cost-effective configuration for capacity.

2 Resource Optimization:

Continuous Monitoring and traking usage and performance of resource.

Adjust resources in real-time based on demand.

Evenly distribute workload across resources and assign resources based on priority and demand.

=> Types of SLAs:

1 No SLAs:

Suitable for non-critical workload where downtime or performance variation is acceptable.

Offers the highest Flexbility in resource allocation.

Least expensive SLAs

SMVS
Page No.
Date : / /

2 Probabilistic SLA's:

This SLAs provides strikes a balance between cost and service reliability.

Suitable for moderate priority application where service variation is acceptable.

More affordable than deterministic SLAs.

3 Deterministic SLA:

Suitable for critical business Function where it required high performance and reliability.

Provides strict gurantees on performances.

Most expensive SLA.

=> Advantages:

1 Improved Service Performance and Reliability.

2 Optimized Resources

3 Cost Saving

4 Reduced Risk of SLA Violation