

Resource Sharing

- To increase the Utilization rate resources run on pools of resources
- Avg. utilization can be increased by sharing them among applications, users and servers
- All application do not use their peak demand
- Implementation need Architectural support
- Resources are shared among several VMs

Resource Sharing

- Challenges:
- Quality of service(QoS)
- Performance Isolation
- Sharing may affect run time behavior
 - Multiple application may compete for the same set of resources
- **Resource management strategies required**

Resource Sharing strategies

- Multi Tenancy:
 - Serve different tenants in isolation from each other
 - Available in public cloud
 - No pre-occupy of any resources
 - Temporary basis
 - Ownership free resource sharing
 - Lower computation cost
 - Not limited for IaaS
 - Can be for PaaS and SaaS

Resource Provisioning strategies

- Autonomic resource provisioning:
 - No of VMs for demand
 - Automatic process by AI
 - Rapid resource provision and release with minimum management
- **Static Approach**
 - Once at the beginning, causes problem when demand crosses limit
- **Dynamic Approach:** On demand, elasticity on cloud
- **Hybrid Approach**

Resource Provisioning strategies

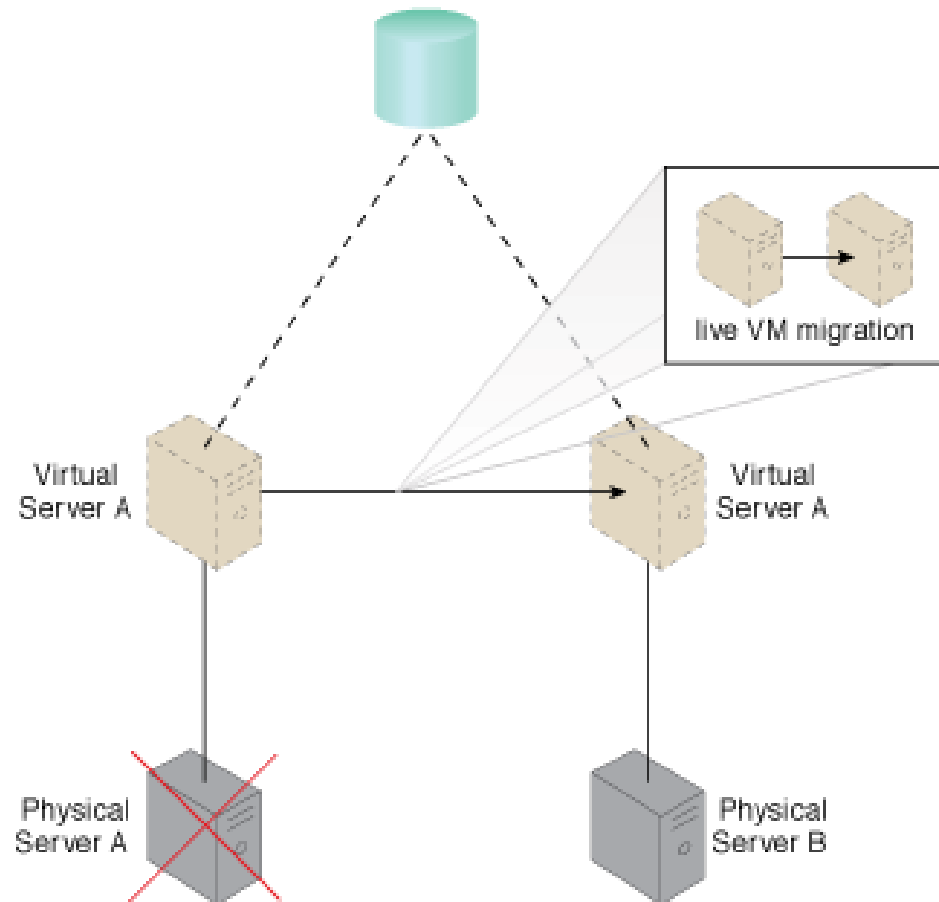
- Under Provisioning
- Over Provisioning
- **Cloud Resource provision plan**
 - **Short term on demand plan**
 - Dynamic approach, pay-per-use
 - **Long term plan**
 - Static approach, charged one time for fixed period

VM Sizing strategies

- Amount of resource for VM
- VM capacity always remain proportionate with load
- Static Approach: fixed at the beginning
- Dynamic: size changes with time and load
- **Individual-VM based**
 - Static at the beginning and allocated as required
- **Joint-VM based**
 - Dynamic

Dynamic Provisioning and Fault Tolerance

- **Zero Downtime Architecture:**
- **VM Migration**
- Physical Server A fails, triggering the live VM migration program to dynamically move Virtual Server A to Physical Server B.



References

- *Cloud Computing*, Sandeep Bhowmik
- https://patterns.arcitura.com/cloud-computing-patterns/design_patterns/zero_downtime