

# Cloud Computing Architecture

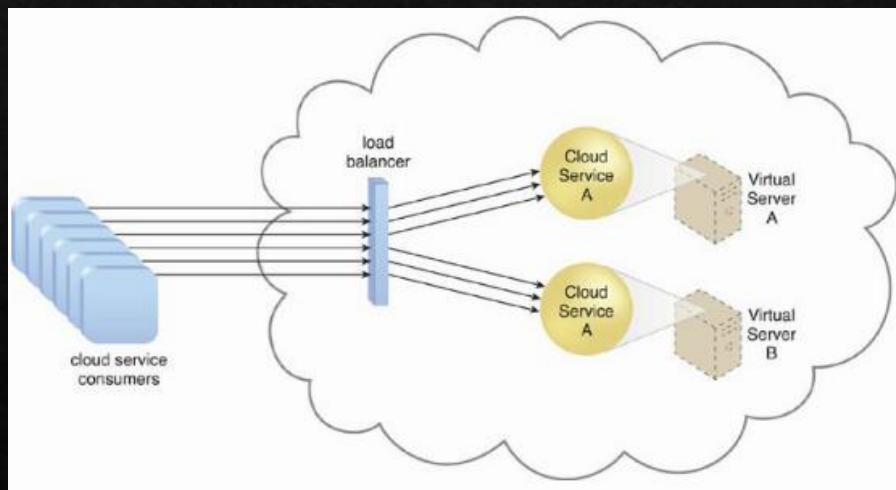
## Fundamental Cloud Architectures

# Fundamental Cloud Architectures

- ❖ Workload Distribution Architecture
- ❖ Dynamic Scalability Architecture
- ❖ Elastic Resource Capacity Architecture
- ❖ Service Load Balancing Architecture
- ❖ Cloud Bursting Architecture
- ❖ Elastic Disk Provisioning Architecture
- ❖ Redundant Storage Architecture

# Workload Distribution Architecture

- ❖ IT resources can be horizontally scaled, adding of one or more identical IT resources and a load balancer that provides runtime logic capable of evenly distributing the workload among the available IT resources.

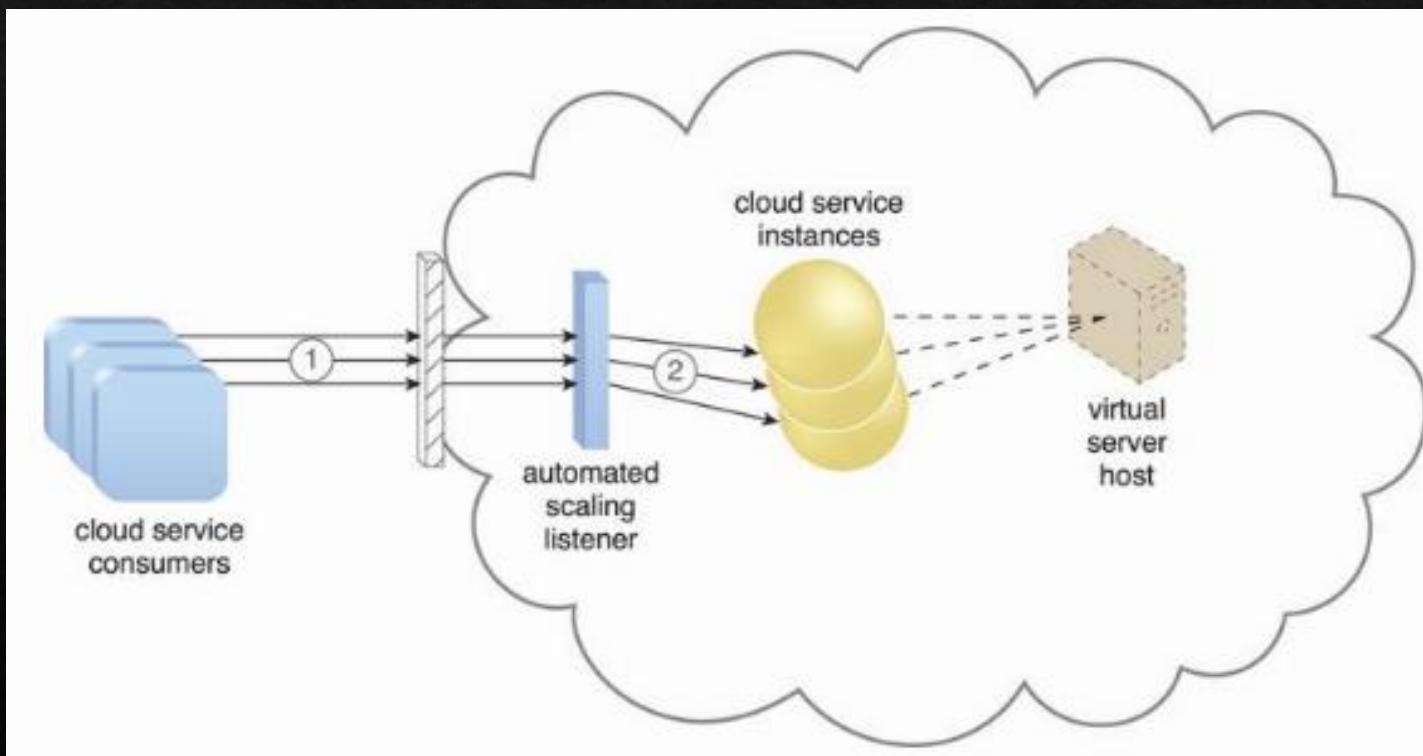


Additional mechanisms:

- Audit monitor
- Cloud usage monitor
- Hypervisor
- Logical network perimeter
- Resource cluster
- Resource replication

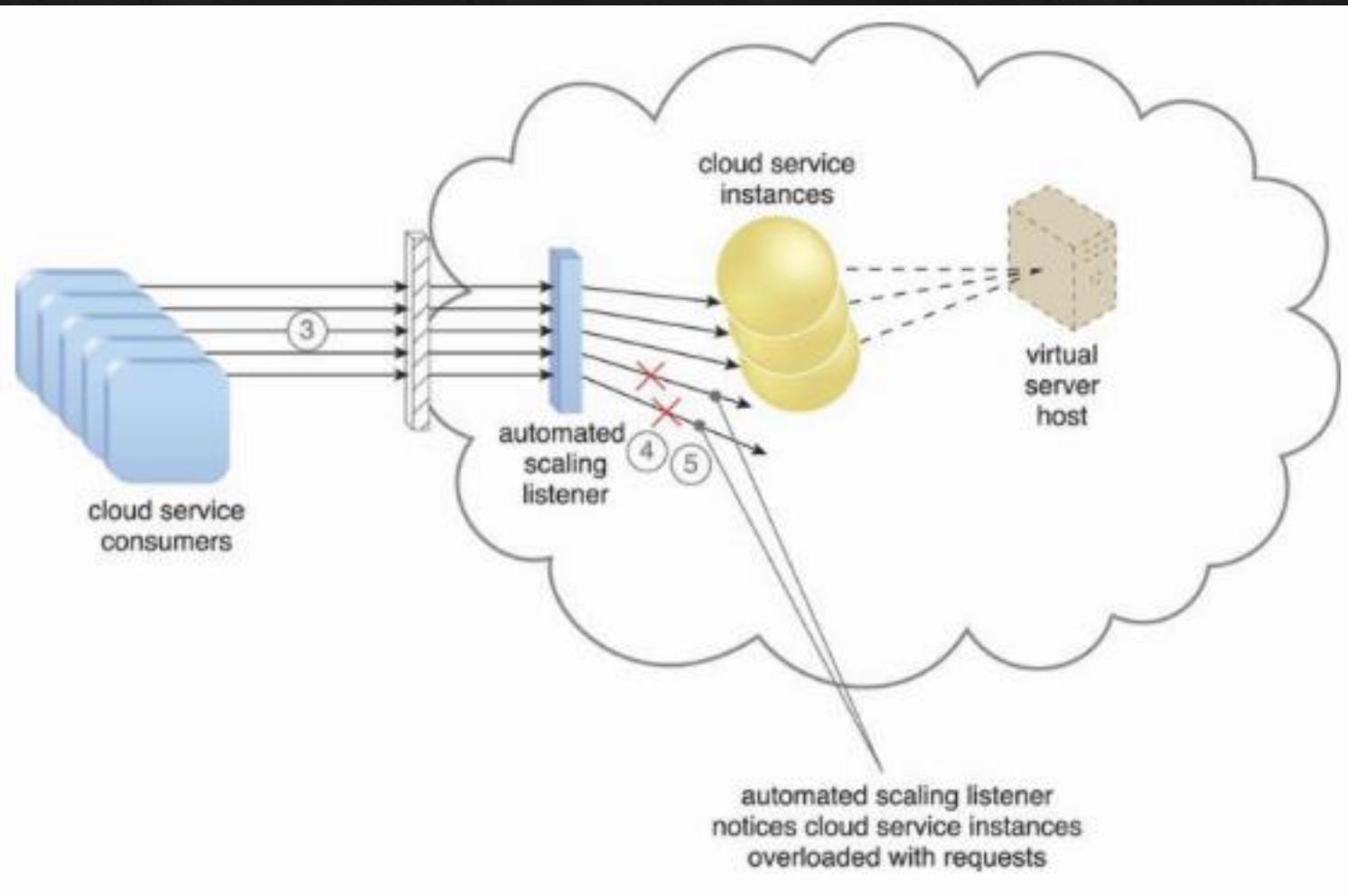
# Dynamic Scalability Architecture

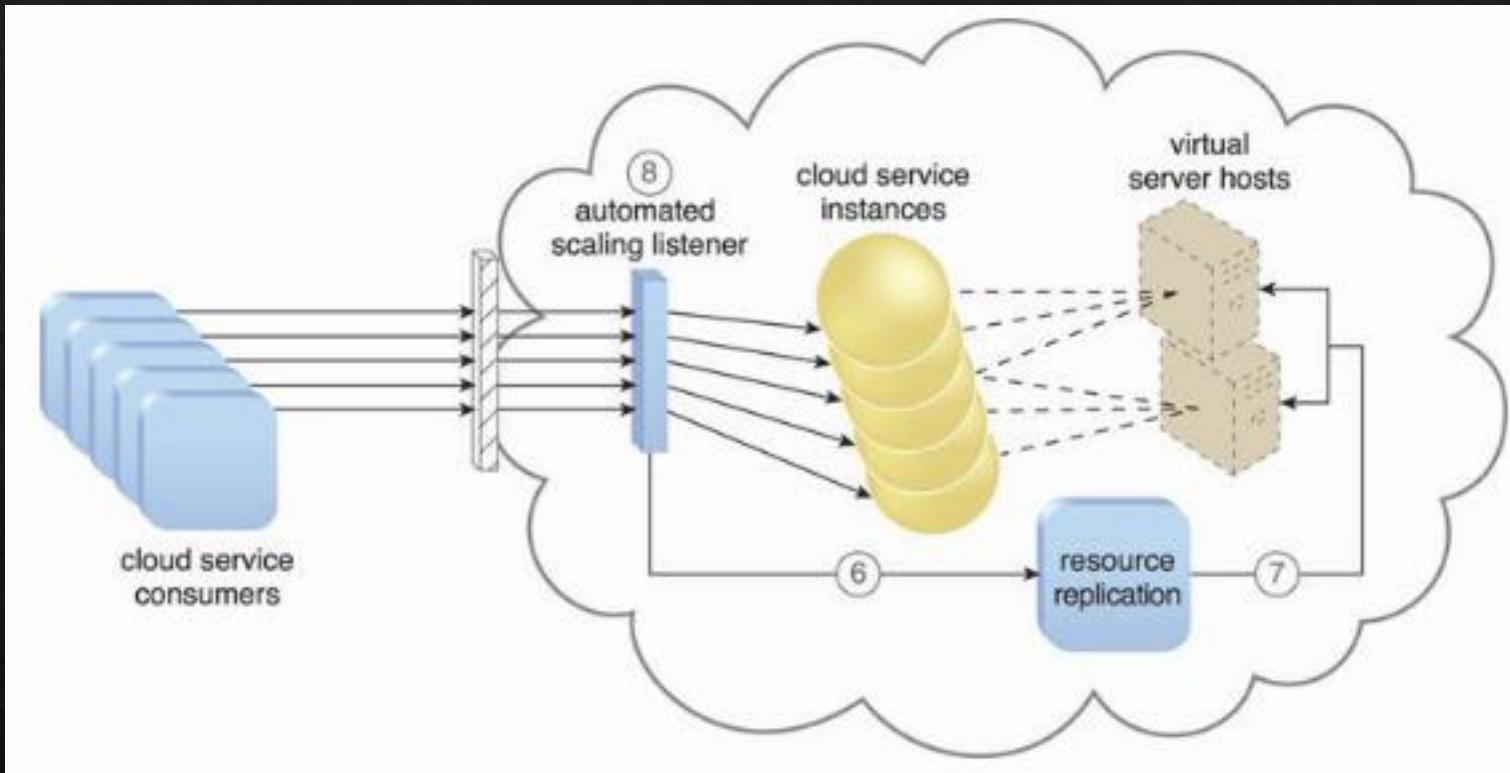
- ❖ An architectural model based on a system of predefined scaling conditions that trigger the dynamic allocation of IT resources from resource pools.



# Dynamic Scaling Mechanisms

- ❖ **Dynamic Horizontal Scaling** – IT resource instances are scaled out and in to handle fluctuating workloads.
- ❖ **Dynamic Vertical Scaling** – IT resource instances are scaled up and down when there is a need to adjust the processing capacity of a single IT resource. For example, a virtual server that is being overloaded can have its memory dynamically increased or it may have a processing core added.
- ❖ **Dynamic Relocation** – The IT resource is relocated to a host with more capacity. For example, a database may need to be moved from a tape-based SAN storage device with 4 GB per second I/O capacity to another disk-based SAN storage device with 8 GB per second I/O capacity.



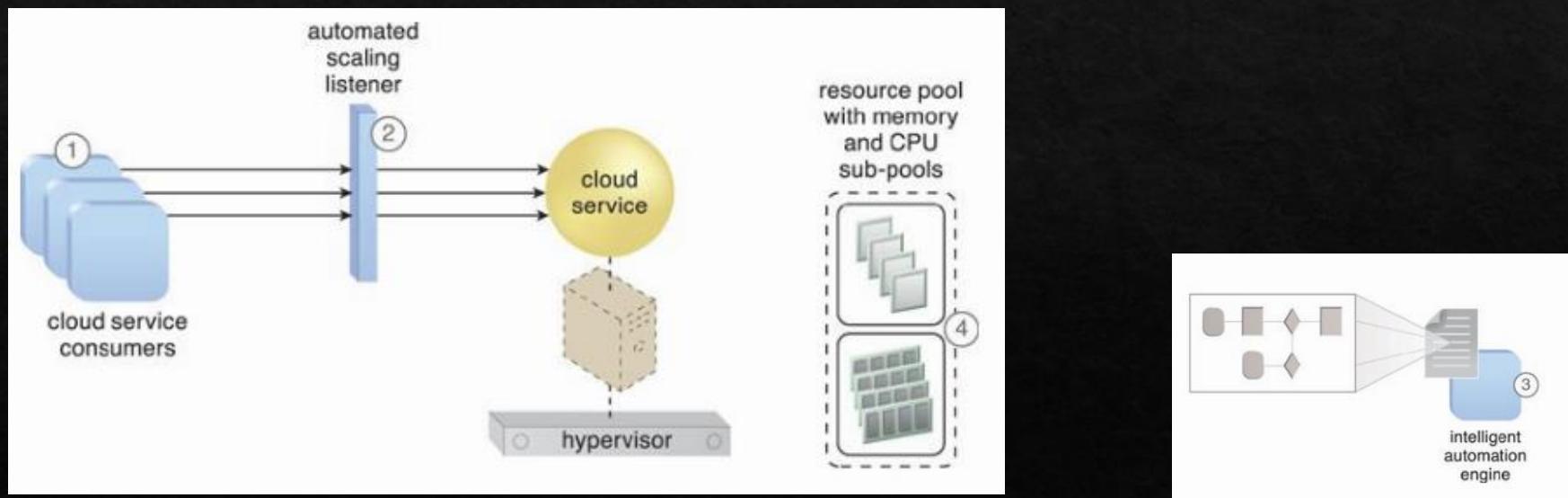


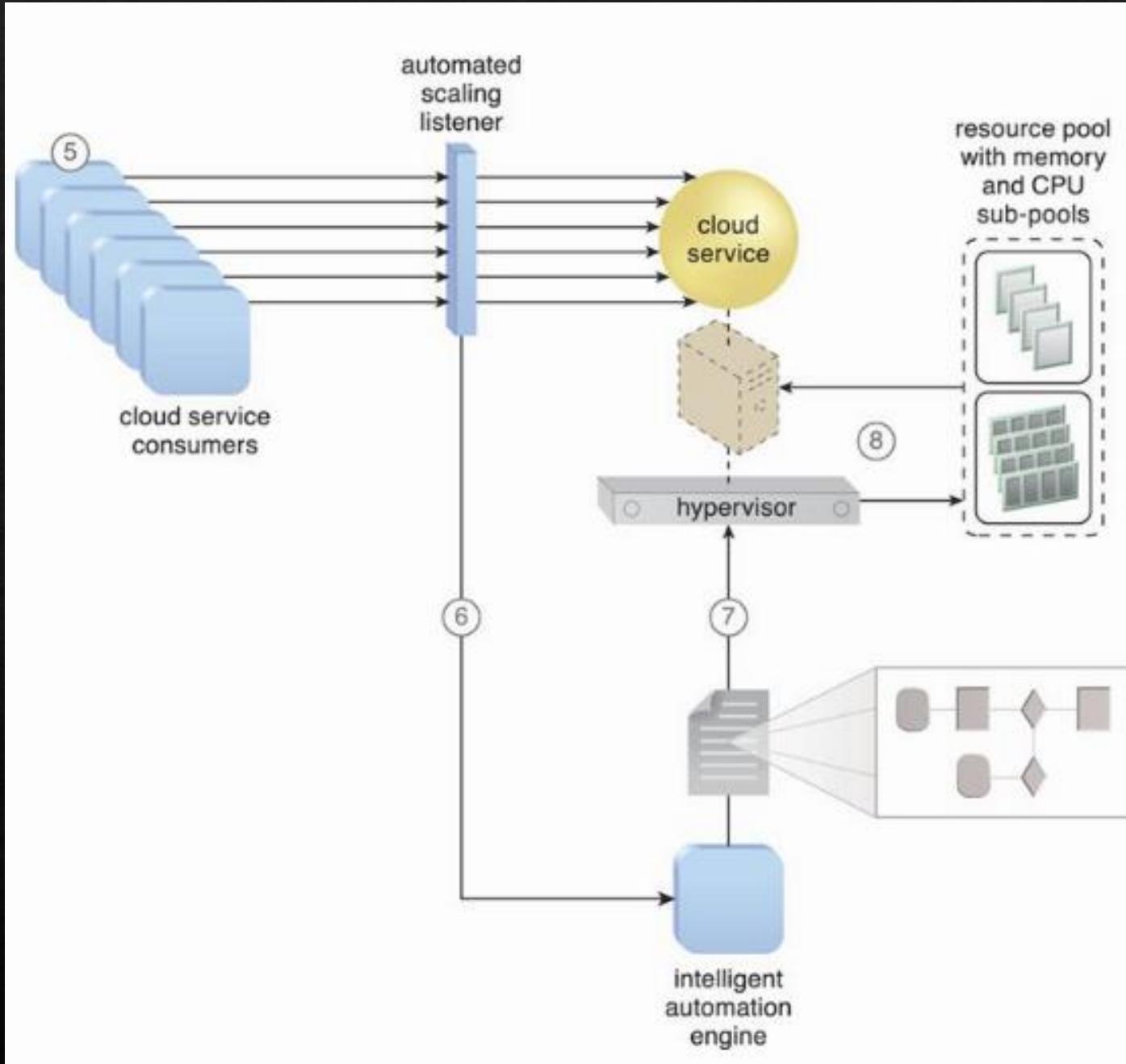
Additional mechanisms to this architecture:

- Cloud usage monitor
- Pay-per-use monitor
- Resource replication
- Hypervisor

# Elastic Resource Capacity Architecture

- ❖ The dynamic provisioning of virtual servers, using a system that allocates and reclaims CPUs and RAM in immediate response to the fluctuating processing requirements of hosted IT resources.



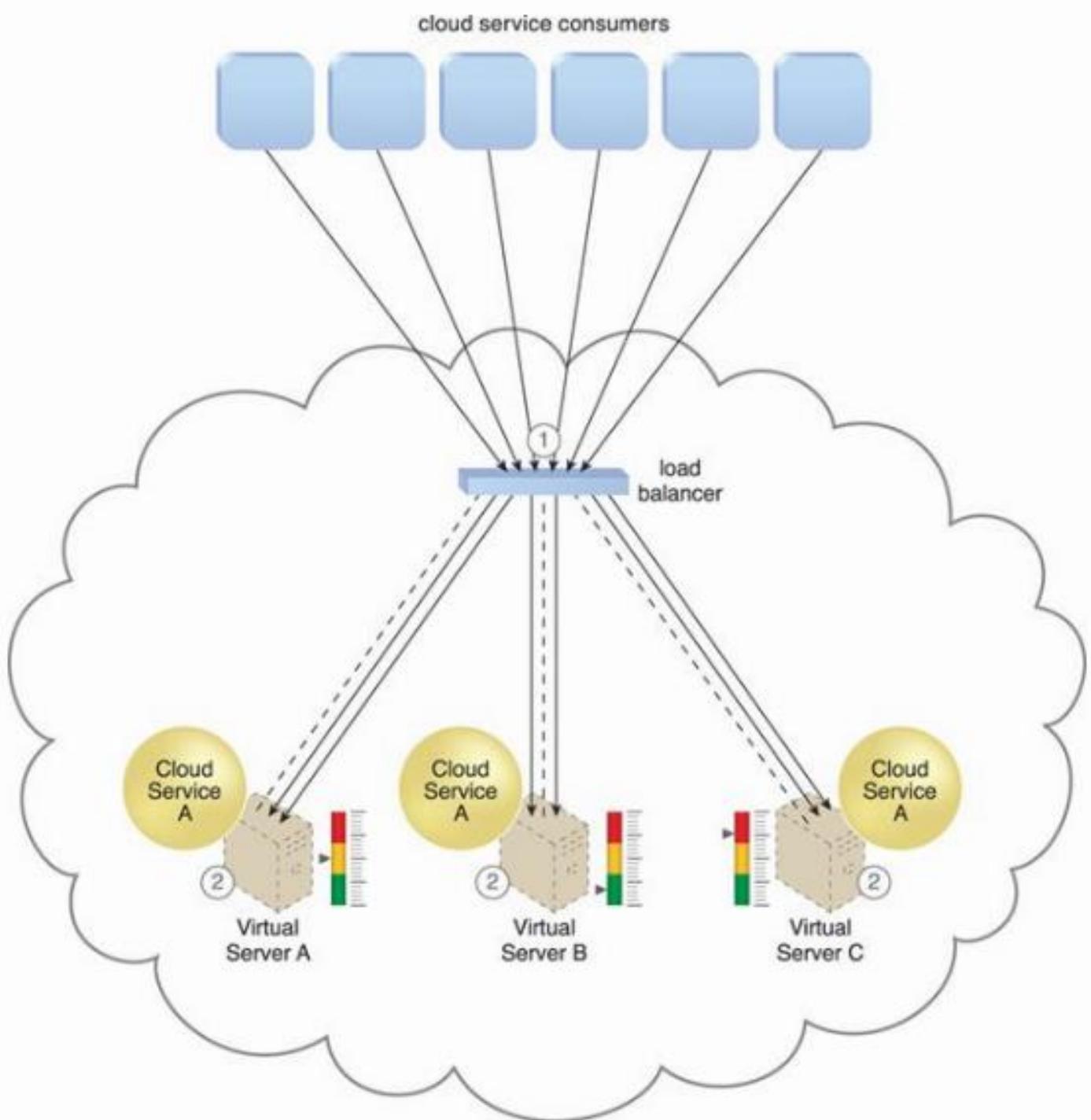


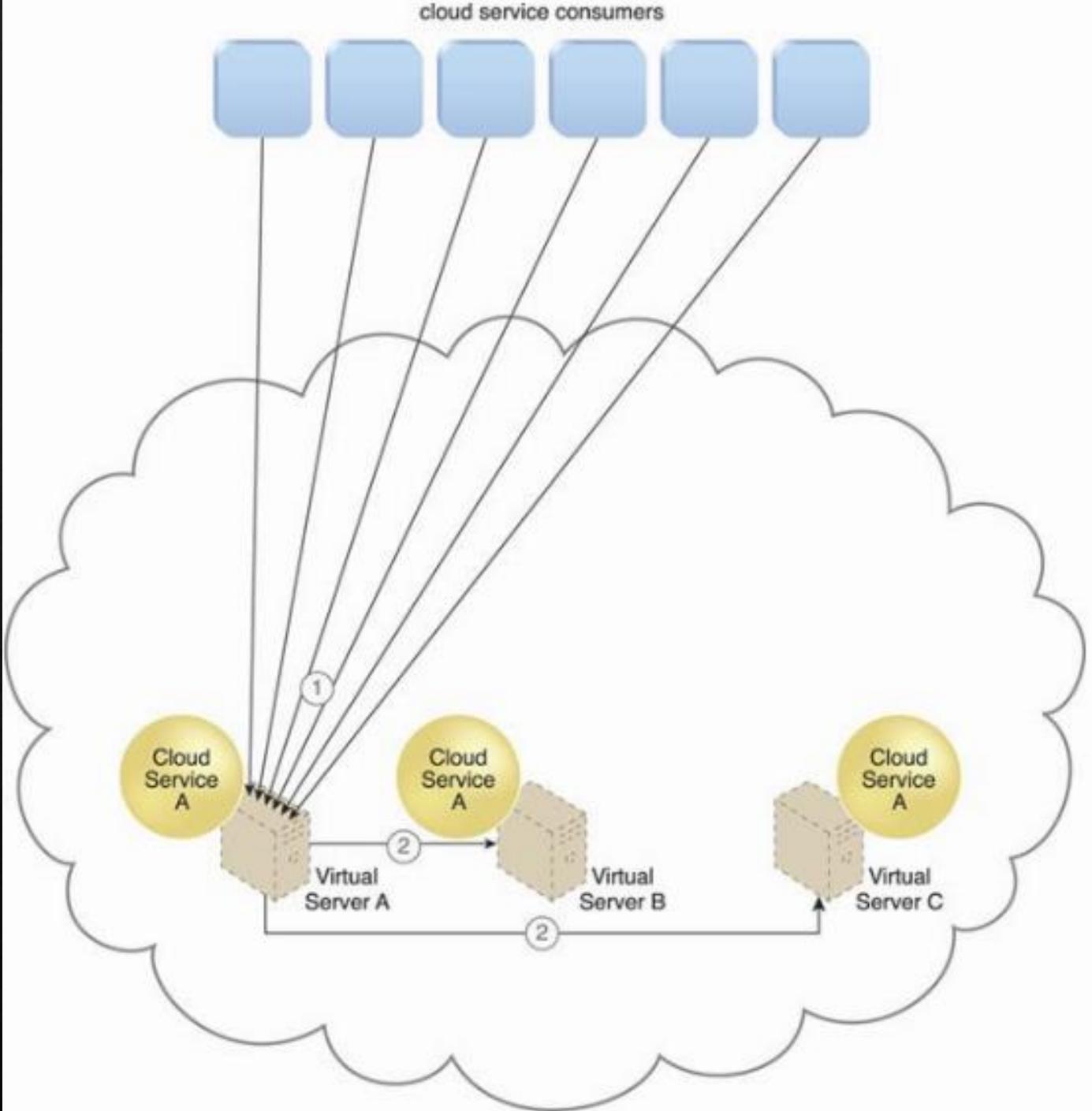
Additional mechanisms to this architecture:

- Cloud usage monitor
- Pay-per-use monitor
- Resource replication
- Hypervisor

# Service Load Balancing Architecture

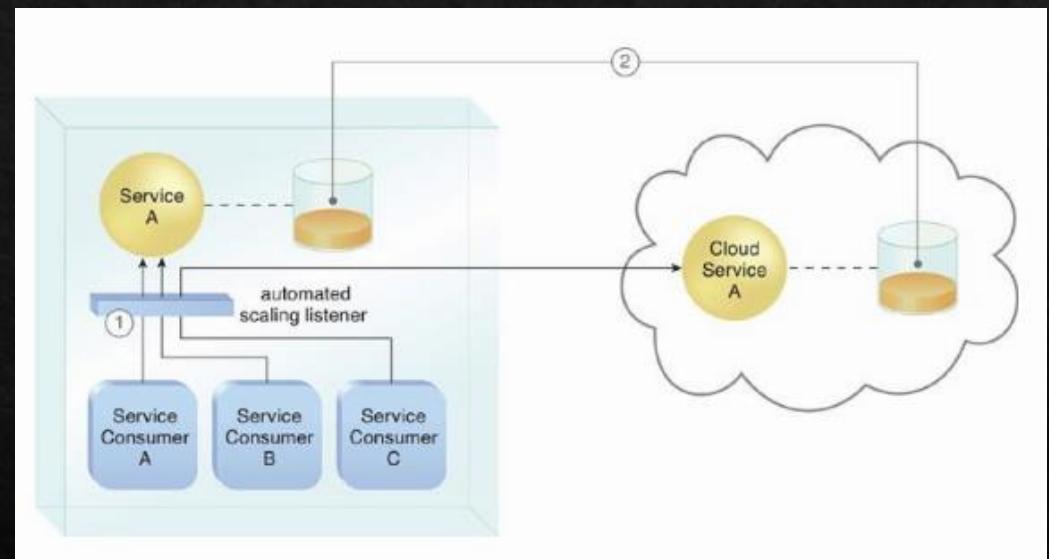
- ❖ A specialized variation of the workload distribution architecture that is geared specifically for scaling cloud service implementations. Redundant deployments of cloud services are created, with a load balancing system added to dynamically distribute workloads.
- ❖ The load balancer can be either independent of the cloud services or built-in part of the application or server's environment.





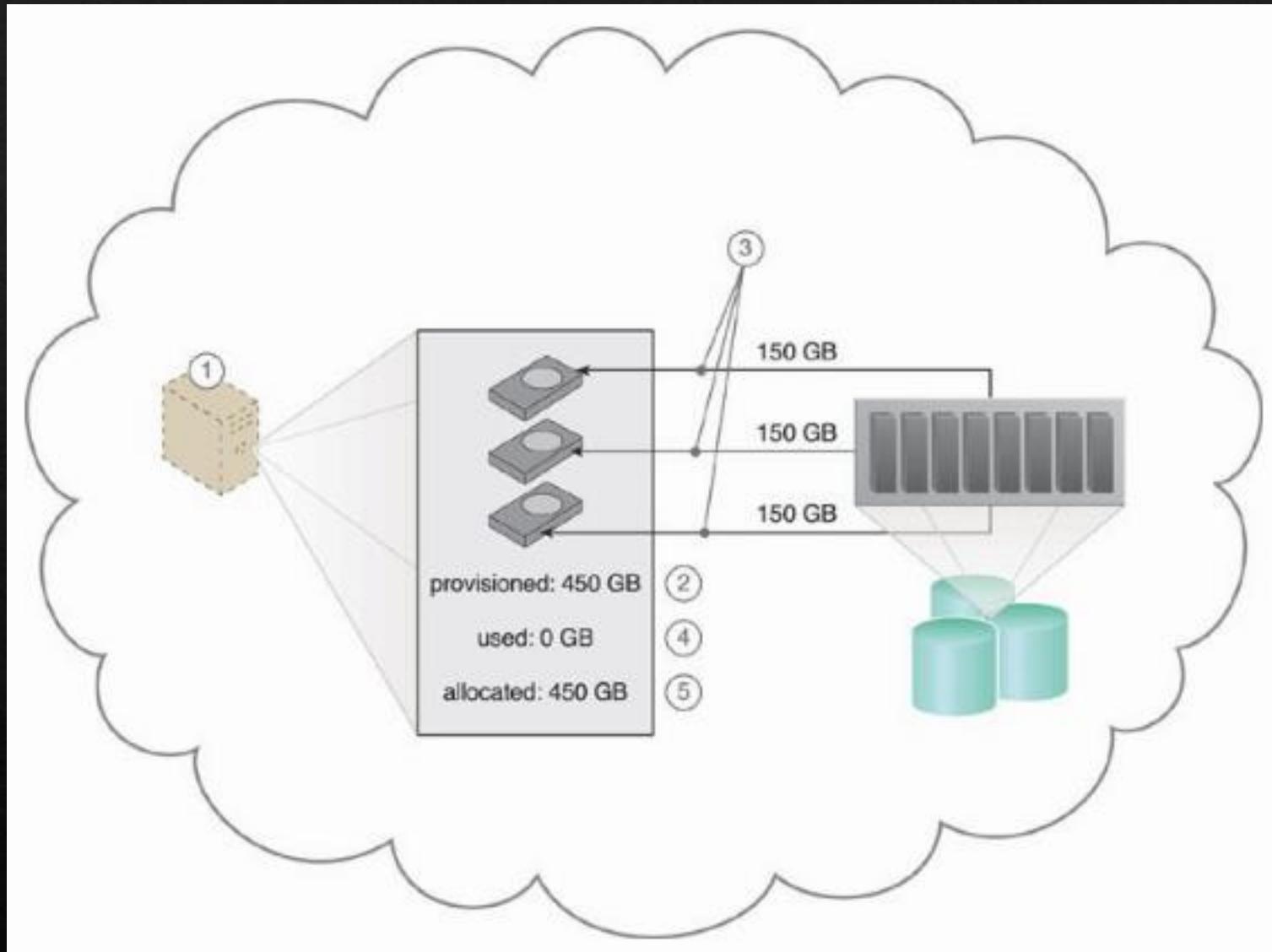
# Cloud Bursting Architecture

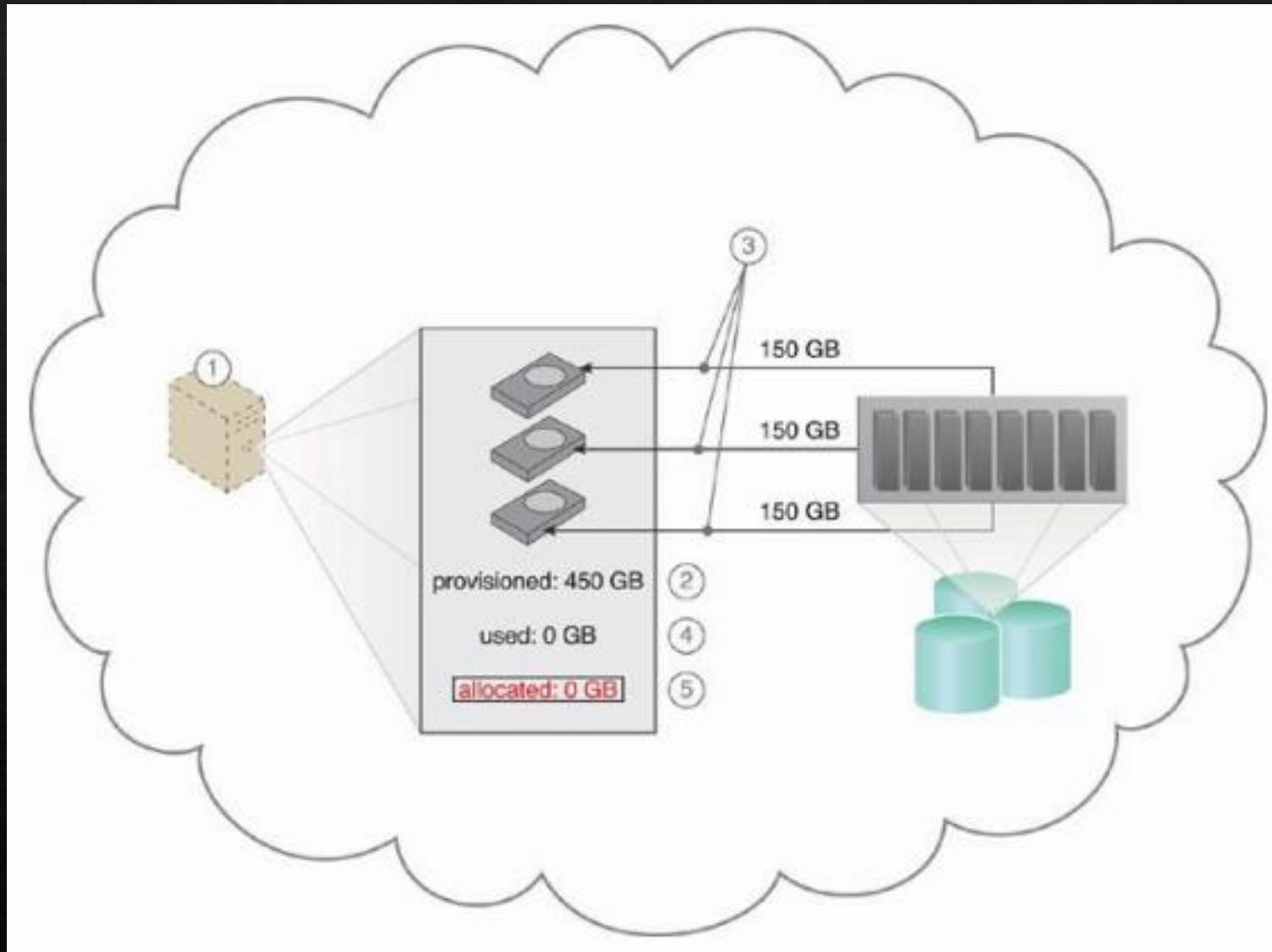
- ❖ A form of dynamic scaling that scales or “bursts out” on-premise IT resources into a cloud whenever predefined capacity thresholds have been reached.
- ❖ The option of using cloud-based IT resources only to meet higher usage demands.
- ❖ The automated scaling listener determines when to redirect requests to cloud-based IT resources, and resource replication is used to maintain synchronicity between on-premise and cloud-based IT resources.



# Elastic Disk Provisioning Architecture

- ❖ Cloud consumers are commonly charged for cloud-based storage space based on fixed-disk storage allocation, meaning the charges are predetermined by disk capacity and not aligned with actual data storage consumption.
- ❖ Thin-provisioning – consumers charged based on real usage.





# Redundant Storage Architecture

- ❖ Failure storages can cause a severe impact to cloud services, cloud consumers' satisfaction, cloud providers' financial and reputation, etc.
- ❖ The redundant storage architecture introduces a secondary duplicate cloud storage device as part of a failover system that synchronizes its data with the data in the primary cloud storage device.

