Management and Monitoring

What is Cloud Monitoring?

Cloud monitoring is the process of reviewing, monitoring and managing the operational workflow and processes within a cloud-based IT asset or infrastructure.

It is the use of manual or automated IT monitoring and management techniques to ensure that a cloud infrastructure or platform performs optimally.

What is Cloud Monitoring?

- Cloud monitoring is primarily part of cloud security and management processes, and it is generally implemented through automated monitoring software that provides central access and control over cloud infrastructure. Cloud administrators can review the operational status and health of any cloud-based device or component.
- In addition to monitoring and ensuring cloud infrastructure/solution/service availability, cloud monitoring data also helps in evaluating the performance of the entire infrastructure on a modular level. Properties such as server uptime and response rate report can help in evaluating customer/user experience.

What does *Cloud Management* mean?

Cloud management is the process of evaluating, monitoring and optimizing cloud computing based solutions and services to produce the desired efficiency, performance and overall service level required.

Cloud management is the practice of end-to-end supervision of the cloud environment by an organization, cloud service vendor or both. It ensures that the cloud computing services are delivered and operated in the most optimal form.

What does *Cloud Management* mean?

- As an IT service, cloud management incorporates most of the underlying tasks and approaches from IT service management.
- It includes very basic to complex management tasks such as maintaining the availability of resources, providing completely functional software/systems and implementing standardized security controls and procedures.
- Some companies are also providing vendor-neutral cloud management software/services to effectively manage and operate cloud services.
- Although the customer or end user is also responsible for their part, cloud management is primarily a vendor end process and includes every task that directly or indirectly affects the cloud environment.

What does *Cloud Management* mean?

- Cloud management is the exercise of administrative control over public, private and hybrid clouds.

- A well-implemented cloud management strategy allows users to maintain control over these dynamic and scalable environments.

Need for Monitoring

- In clouds, monitoring is essential for the health of the system and is important for both providers and consumers
- Primarily, monitoring is a key tool for
- i) managing software and hardware resources;

  And
- ii)providing continuous information for those resources as well as for consumers' hosted applications on the cloud.

Need for Monitoring

- Monitoring of Cloud is a task of paramount importance for both Providers and Consumers.
- On the one side, it is a key tool for controlling and managing hardware and software infrastructures;
- on the other side, it provides information and Key Performance Indicators (KPIs) for both platforms and applications.
- The continuous monitoring of the Cloud and of its SLAs (for example, in terms of availability, delay, etc.) supplies both the Providers and the Consumers with information such as the workload generated by the latter and the performance and QoS offered through the Cloud, also allowing to implement mechanisms to prevent or recover violations (for both the Provider and Consumers).

Need for Monitoring

Cloud Computing involves many activities for which monitoring is an essential task

- Capacity and resource planning
- Capacity and resource management
- Data center management
- SLA management
- Billing
- Troubleshooting
- Performance management
- Security management

Need for Monitoring

1. Capacity and resource planning

- ✓ Cloud Service Providers usually offer guarantees in terms of QoS and thus of resources and capacity for their services as specified in SLAs [23], and they are in charge of their resource and capacity planning so that service and application developers do not have to worry about them [24].

- ✓ To this end, monitoring becomes essential for Cloud Service Providers to predict and keep track of the evolution of all the parameters involved in the process of QoS assurance [25] in order to properly plan their infrastructure and resources for respecting the SLAs.

Need for Monitoring

2. Capacity and resource management

- ✓ The first step to manage a complex system like a Cloud consists in having a monitoring system able to accurately capture its state [26].
- ✓ monitoring is necessary to cope with volatility of resources [31] and fast-changing network conditions (which may lead to faults)
- ✓ when adopting Cloud infrastructures, companies and people expect such services to have 100% uptime. Thus, a resilient and trustworthy monitoring of the entire Cloud infrastructures is needed to provide availability

Need for Monitoring

3. Data center management

- ✓ Data center management activities (e.g. data center control) imply two fundamental tasks: (i) monitoring, that keeps track of desired hardware and software metrics; and (ii) data analysis, that processes such metrics to infer system or application states for resource provisioning, troubleshooting, or other management actions [33].
- ✓ In order to properly manage such data centers,both monitoring and data analysis tasks must support real-time operation and scale up to tens of thousands of heterogeneous nodes, dealing with complex network topologies and I/O structures.

Need for Monitoring

4. SLA management

- ✓ Service level management involves comparing actual performance with pre-defined expectations, determining appropriate actions, and producing meaningful reports.
- ✓ monitoring may allow Cloud Providers to formulate more realistic and dynamic SLAs and better pricing models by exploiting the knowledge of user-perceived performance

5. Billing

- ✓ For each of the reported pricing models and service models, monitoring is necessary both from the Provider side for billing, and from the Consumer side for verifying his own usage and to compare different Providers, a nontrivial process requiring monitoring functionalities and tools

Need for Monitoring

6. Troubleshooting

- ✓ The complex infrastructure of a Cloud represents a big challenge for troubleshooting (e.g. root cause analysis), as the cause of the problem has to be searched in several possible components (e.g. network, host, etc.), each of them made of several layers (e.g. real and virtual hardware, host and guest OS, etc.).
- ✓ A comprehensive, reliable and timely monitoring platform is therefore needed for Providers to understand where to locate the problem inside their complex infrastructure and for Consumers to understand if any occurring performance issue or failure is caused by the Provider, network infrastructure, or by the application itself.

Need for Monitoring

7. Performance management

- ✓ If a Consumer adopts a public Cloud to host a mission-critical service or for a scientific application, performance variability and availability become a concern. Therefore, from a Consumer's perspective, monitoring the perceived performance is necessary to adapt to the changes or to apply corrective measures.
- ✓ Monitoring is then necessary since it may considerably improve the performance of real applications [39] and affect activity planning and repeatability of experiments.

Need for Monitoring

7. Security management

- ✓ For managing the security in Cloud infrastructures and services, proper monitoring systems are needed. Moreover, for hosting critical services for public agencies, Clouds have to satisfy strict regulations and prove it. And this can be done through a monitoring system that enables auditing (e.g. to certify the compliance to regulations and obligations, such as keeping data of a user inside country.

Types of Monitoring

- In Cloud Computing, we can have both high- and low level monitoring, and both are required [46].

1. High-level monitoring
   - It is related to information on the status of the virtual platform.
   - This information is collected at the middleware, application and user layers by Providers or Consumers through platforms and services operated by themselves or by third parties.
   - In the case of SaaS, high-level monitoring information is generally of more interest for the Consumer than for the Provider (being closely related to the QoS experienced by the former).

2. low-level monitoring
   - is related to information collected by the Provider and usually not exposed to the Consumer, and it is more concerned with the status of the physical infrastructure of the whole Cloud (e.g. servers and storage areas, etc.).
   - In the context of IaaS, both levels are of interest for both Consumers and Providers

Tests and metrics

- Monitoring tests can be divided in two main categories:
- Computation-based and Network-based [47].
- **Computation based tests** are concerned about the status of the real or virtualized platforms running cloud applications.
- Data metrics considered in such test include
  - CPU speed,
  - CPU utilization,
  - disk throughput,
  - VM acquisition/release time and
  - system up-time.
- **Network based tests** focus on network layer data related metrics like
  - jitter,
  - round-trip time RTT,
  - packets loss,
  - traffic volume etc. [20][21][23][24].

Monitoring across different applications and Layers

- As mentioned previously, application components (e.g., streaming server, web server, indexing server, compute service, storage service, and network) are distributed across cloud layers including PaaS and IaaS.

- Thus, in order to guarantee the achievement of QoS targets for the application as a whole, monitoring QoS parameters should be performed across all the layers of cloud stack including
  - Platform-as-a-Service (PaaS) (e.g., web server, streaming server, indexing server, etc.) and
  - Infrastructure-as-a-Service (IaaS) (e.g., compute services, storage services, and network).

Brief Process of Monitoring

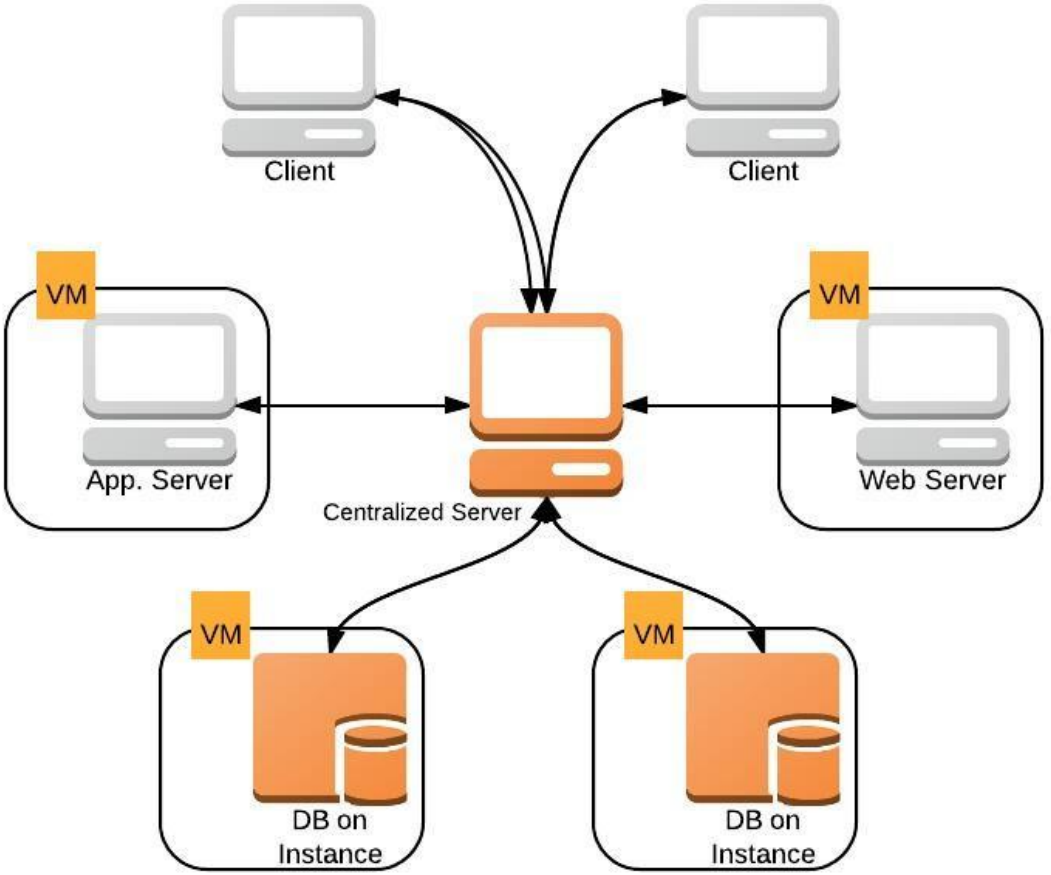
- Cloud monitoring system is a self-adjusting and typically multi-threaded system that is able to support monitoring functionalities [11].
- It comprehensively monitors pre-identified instances/resources on the cloud for abnormalities.
- On detecting an abnormal behavior, the monitoring system attempts to auto-repair this instance/resource if the corresponding monitor has a tagged auto-heal action [11].
- In case of auto-repair failure or an absence of an auto-heal action, a support team is notified.
- Technically, notifications can be sent by different means such as email, or SMS [11].

Monitoring Architectures

Centralized

- In centralized architecture shown in figure, the PaaS and IaaS resources send QoS status update queries to the centralized monitoring server.
- In this scheme, the monitoring techniques continuously pull the information from the components via periodic probing messages.
- In [11], the authors show that a centralized cloud monitoring architecture allows better management for cloud applications.
- Nevertheless, centralized approach has several design issues, including:
    - Prone to a single point of failure;
    - Lack of scalability;
    - High network communication cost at links leading to the information server (i.e., network bottleneck,congestion); and
    - Possible lack of the required computational power to serve a large number of monitoring requests.

Monitoring Architectures

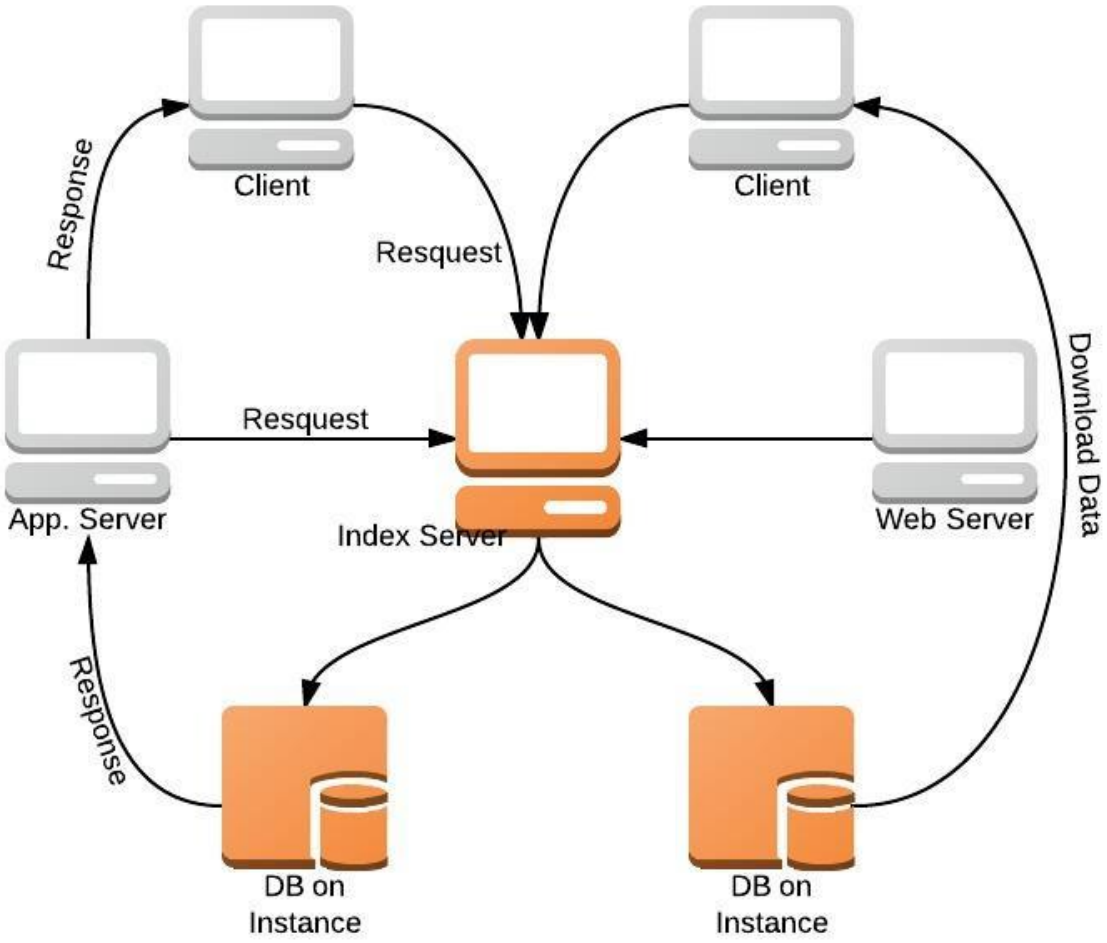


Monitoring Architectures

decentralized

- Recently, proposals for **decentralized cloud monitoring** tools have gained momentum. Figure shows the broad schematic design of decentralized cloud monitoring system.
- The decentralization of monitoring tools can overcome the issues related to current centralized systems.
- A monitoring tool configuration is considered as decentralized if none of the components in the system is more important than others.
- In case one of the components fails, it does not influence the operations of any other component in the system.

Monitoring Architectures



Cloud Monitoring Tools

- **Monitis**
- **RevealCloud**
- **LogicMonitor**
- **Nimsoft**
- **Nagios**
- **CloudWatch**
- **OpenNebula**
- **CloudHarmony**

Cloud Monitoring Tools

CloudWatch

- CloudWatch [41] is one of the most popular commercial tools for monitoring the cloud.
- It is provided by Amazon to enable its consumers monitoring their resources residing on EC2.
- Hence, it does not support multicloud infrastructure monitoring.
- The technical approaches used in CloudWatch to collect data are implicit and of exposed to users.
- CloudWatch is limited in monitoring resources across cloud layers.
- However, an API is provided for users to collect metrics at any cloud layer but requires the users to write additional code.

Cloud monitoring: properties

- Scalability
- Elasticity
- Adaptability
- Timeliness
- Autonomicity
- Comprehensiveness, Extensibility and Intrusiveness
- Resilience, Reliability, and Availability
- Accuracy

Objective of Cloud Provider and Consumer

- The central objectives of cloud provider are profit maximization and increase customer satisfaction level (CSL) i.e. the market sharing maximization.
- To achieve these objectives, the cloud provider need to reduce cost, SLA violations, response time, and power consumption;
- and deploy services in different prices (Dynamic pricing) based on the current consumer's requirements as well as the level of the offered QoS.
- On the contrary, the main objectives of cloud consumers are minimizing cost (price) and access services with high quality of services (QoS).
- To achieve all of these objectives - provider's and consumer's objectives - , the negotiation between them would be established through service level agreement (SLA).

Choose Provider

Consumers can assess the providers depending on the following factors:

- **Pricing Scheme:** the mechanism in which service price could be determined, such as pay-as-you-go model.
- **Service Customizability:** the way by which provider customizes his SaaS services to meet service consumer's requirements.
- **Leasing period:** the period, at which customers can consume services, some examples of this period are subscription, pay-per-use, and perpetual.
- **Service QoS:** the mechanism by which service requirements could be specified; such as scalability, availability, and security.

Management and Monitoring

Pricing

Pricing-Importance

- Pricing represents an important indicator for success business companies which provide services or products [15]. Cloud provider uses several pricing models to specify the price. This pricing model would be established properly to define the fair price for both stakeholders (providers and consumers). A good pricing model supports cloud providers to achieve their objectives such as profit maximization; meanwhile it considers the cloud consumers.
- Pricing is a critical factor for organizations offering services or products [30]. How the price is set affects customer behavior, loyalty to a provider, and the organization's success. Therefore, developing an appropriate pricing model will help achieve higher revenues. The price determined for a service or product must consider the manufacturing and maintenance costs, market competition, and how the customer values the service or product offered.

Pricing Definition

- Pricing is the process of determining what a service provider will receive from an end user in exchange for their services.
- Weinhardt *et al.*, claimed that cloud computing success in the IT market can be obtained only by developing adequate pricing techniques.

Factors that influence pricing

1. **Initial costs**

   This is the amount of money that the service provider spends annually to buy resources.

2. **Lease period**

   This is the period in which the customer will lease resources from the service provider. Service providers usually offer lower unit prices for longer subscription periods.

3. **QoS**

   This is the set of technologies and techniques offered by the service provider to enhance the user experience in the cloud, such as data privacy and resource availability. The better QoS offered, the higher the price will be.

4. **Age of resources**

   This is the age of the resources employed by the service provider. The older the resources are, the lower the price charged will be. This is because resources can sustain wear over time, which reduces their financial value.

5. **Cost of maintenance.**

   This is the amount of money that the service provider spends on maintaining and securing the cloud annually.

Pricing Models

Pricing Models Classification

the two common types of pricing models are:

1. **Fixed Pricing Model:**
   - ✓ Here the price charging doesn't change, and the cloud provider is someone who determines the price to the resource type in advance.
   - ✓ customer is charged the same amount all the time;
   - ✓ For example, Amazon provides disk space for $0.15/GB, and service consumers have the same services at all time, such as Payper- use model. According to Yeoa et al. [2], fixed pricing model is more straightforward and easy to understand, but it is unfair for all customers because they are not having the same needs.

Pricing Models Classification

2. Dynamic Pricing Model:

- ✓ In this model the price charging changes dynamically according to market status quo.
- ✓ The service price could be calculated for each request according to the pricing mechanism that is used.
- ✓ In this case, service consumer requests and receives several types and levels of services in need, such as Market- dependant pricing model.
- ✓ dynamic pricing implies that the price changes dynamically according to the service features, customer characteristics, amount of purchased volumes, or customer preferences. Market-dependent pricing, however, depends on the real-time market conditions such as bargaining, auctioning, demand behavior, and yield management.

Pricing Models Classification

Pricing Model	Advantages	Disadvantages
Fixed pricing model	• It supports assurances for Consumers . • Consumers know how • much they will pay. • More consistent. • It reduces risks. • Make profit estimation • easy.	• Unfair for consumer: If the user doesn't consume the Resource extensively, he/she may pay more than his/her real utilization. • It does not allow provider to change price at any account. • Unfair for provider: During proper resource utilization consumer may pay less than his/her real utilization.

Pricing Models Classification

Pricing Model	Advantages	Disadvantages
Dynamic pricing model	• It supports provider to Maximize profits with each consumer. • Fair for consumer as it enables him to pay according to the offered QoS. • It supports provider to set price based on current state of the market (season or supply and demand)	• Some consumers are not interested in this model as they prefer a fixed price to dynamic price. • In some environments such • as entertainment sites consumers do not prefer dynamic pricing.