*Source - https://medium.com/cloud-native-daily/microservices-monitoring-and-observability-in-depth-d40aa0795dd3*

*Monitoring is the collection and analysis of data to detect problems or anomalies in a system or application. This data comes from metrics, logs, and alerts and focuses on specific things that can be tracked over time.*

*Key Metrics –*

*Latency and response time: Latency is the time it takes for a request to go from a client to a microservice and back, and response time is the total time it takes a microservice to process a request and generate a response. Usually, this is measured in milliseconds. By constantly monitoring latency and response times, you can identify performance issues and bottlenecks that can slow down your work.*

*Error rate and error metrics: These metrics help you understand the stability and reliability of your microservices. This includes, for example, the number of errors and exceptions encountered in requests and operations. By monitoring error rates, you can identify bugs, configuration issues, and external service failures that may be causing problems. By monitoring failure rates and failure metrics, you can take steps to troubleshoot issues and ensure your microservices are always available and fault tolerant.*

*Throughput and request volume: These metrics measure the number of requests or operations a microservice can handle over a period of time. You can check the scalability and performance of your microservices by monitoring their throughput. You can also identify performance bottlenecks that may be blocking your work.*

*Resource Usage Metrics - These metrics track how much CPU, memory, disk space, and network bandwidth your microservices are using. Monitoring these metrics will give you a better idea of ​​how efficient your microservices are and how much capacity they have. Knowing your resource utilization metrics can help you identify potential problems and optimize your resource allocation.*

*Service Level Indicators (SLIs) and Service Level Objectives (SLOs). These are very important metrics for measuring microservice quality and performance. SLIs are specific metrics that define the desired behavior or performance of a service, such as latency, error rates, or availability. SLOs, on the other hand, are the targets or thresholds set for the SLIs.*