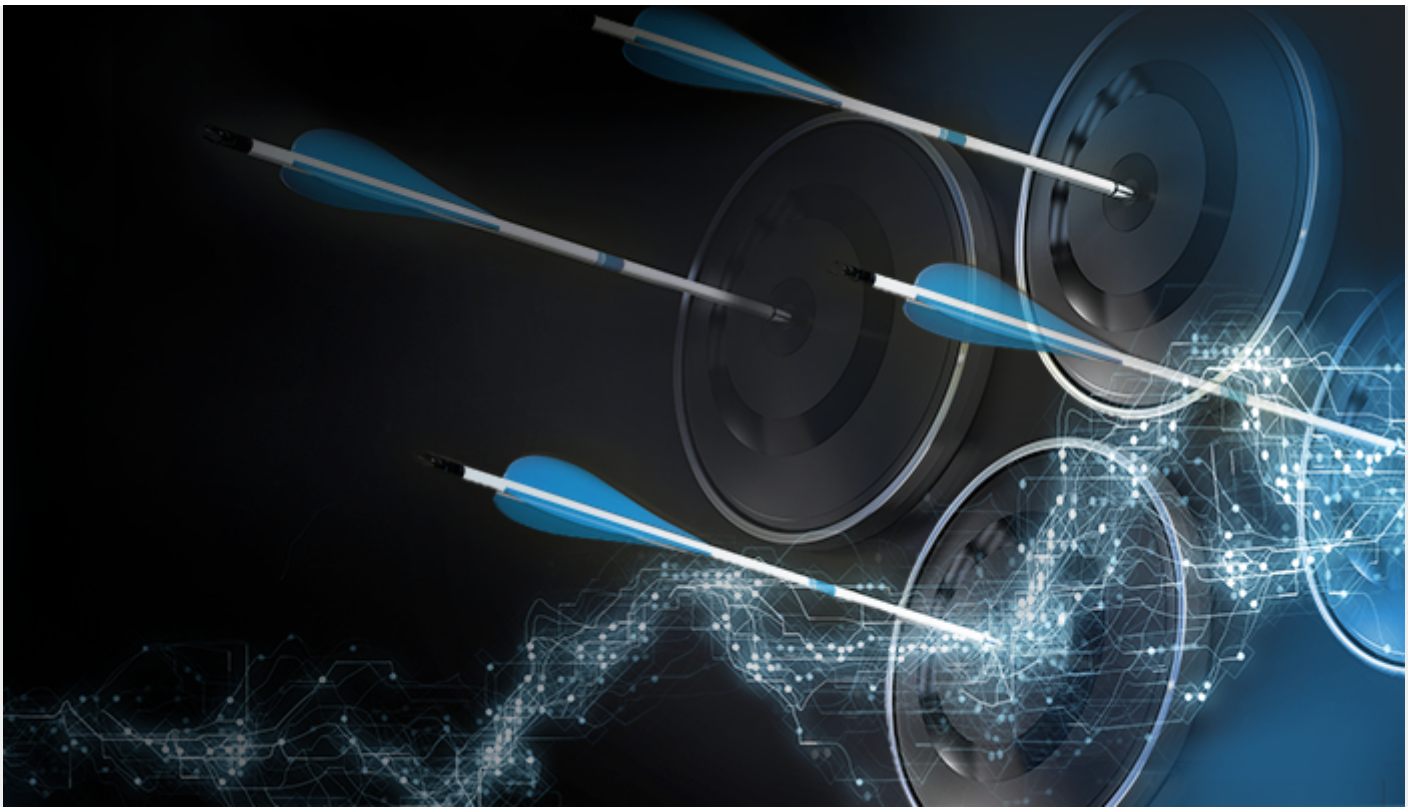
SERVICE LEVEL OBJECTIVES (SLOS) EXPLAINED



When IT workloads are moved from on-premise IT infrastructure to off-site cloud data centers, it's important to ensure that service levels are consistent with your business requirements.

The parameters and metrics defining service levels for each element of the cloud solution should meet those requirements. The service provider should maintain high performance, security, and compliance standards while still offering an affordable cost structure that entices customers.

In order to meet these goals, organizations must understand, measure, and evaluate the service behavior based on well-defined objectives. The topic of Service Level Agreements (SLAs) is widely covered to discuss the responsibilities of cloud service providers. However, a better understanding of Service Level Objectives (SLOs) can help evaluate measurements of service performance that actually matter to end users.

So, let's look at these SLOs.

What are service level objectives?

A Service Level Objective (SLO) serves as a benchmark for indicators, parameters, or metrics defined with specific service level targets. The objectives may be an optimal range or a specific value for each service function or process that constitutes a cloud service.

The SLOs can also be referred to as measurable characteristics of an SLA, such as Quality of Service (QoS) aspects that are achievable, measurable, meaningful, and acceptable for both service

providers and customers. An SLO is agreed within the SLA as an obligation, with validity under specific conditions (such as time period) and is expressed within an SLA document.

(Read our explainer on SLAs vs SLOs.)

How SLOs work

While the end-goal of defining effective SLOs is to deliver reliable services to end-users, the cost and complexity of getting closer to a 100% reliability increases exponentially. (See how this pays out with service availability.)

bmc

The 9s of Availability

Availability percentages vs service downtime

Availability %	Downtime per year	Downtime per month	Downtime per week
90% (1 nine)	36.5 days	72 hours	16.8 hours
99% (2 nines)	3.65 days	7.20 hours	1.68 hours
99.5%	1.83 days	3.60 hours	50.4 minutes
99.9% (3 nines)	8.76 hours	43.8 minutes	10.1 minutes
99.95%	4.38 hours	21.56 minutes	5.04 minutes
99.99% (4 nines)	52.56 minutes	4.32 minutes	1.01 minutes
99.999% (5 nines)	5.26 minutes	25.9 seconds	6.05 seconds
99.9999% (6 nines)	31.5 seconds	2.59 seconds	0.605 seconds
99.99999% (7 nines)	3.15 seconds	0.259 seconds	0.0605 seconds

Every component of the cloud service causes different impact on the service performance as perceived by customers. For instance, an app may require responsiveness at a specific performance level, beyond which customers can no longer feel a difference.

The measure of responsiveness and app performance may be defined through numeric indicators such as:

- Request latency

- Batch throughput
- Failures per seconds
- Other metrics

These indicators describe the service level at any moment in time.

To understand the overall performance in context of the agreed SLA contract or availability requirements, you'll have to analyze these numbers over a longer time period. Mathematically, the SLO analysis involves:

1. Aggregating the service level indicator performance over a long time.
2. Comparing the result with a numerical target for system availability.

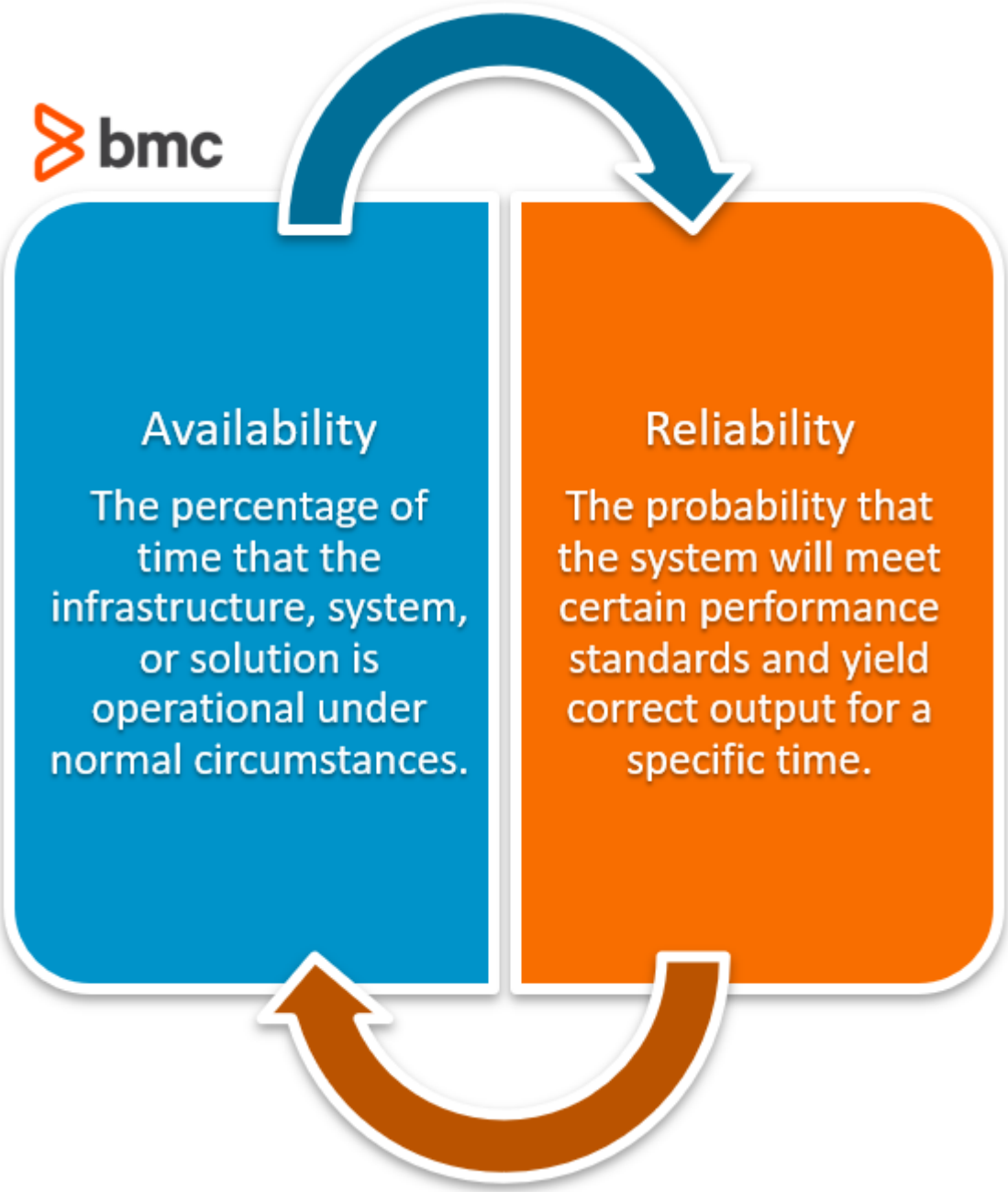
An SLO is a range

An SLO is not intended to define the best performance level. Instead, it should define a range of best possible and least acceptable performance standards.

Imagine a scenario where a cloud service purchased with an SLA of 99% uptime, translating into 7.31 hours of downtime per month. Several months pass by and the systems are maintained at the upper bound of the SLOs delivering 99.9% of uptime, or less than one hour of downtime per month.

Suddenly when the systems do go down for several continuous hours, end-users are unpleasantly surprised for the service to perform below their usual expectations. At the same time, the service provider may not be obligated to offer support if it doesn't intend or commit to delivering the best possible SLO.

Since SLO involves the measurement of several quantifiable metrics describing reliability of the system, you should carefully understand the difference between the terms Reliability and Availability. In practice, the SLOs are defined by the lowest acceptable reliability standards.



SLO best practices

Put simply, Service Level Objectives describe how good the service reliability was during a specific duration of time, based on the measurements of specific service level indicators.

The following best practices can help you achieve these goals:

- **Identify the right metrics & indicators** to accurately describe system reliability as perceived, expected, and required by your organization and end-users.
- **Make sure the right people understand the SLOs.** SLO should be well understood by the technical team and organizational leaders. Organizations should devise SLOs based on the business requirements as well as the technical capacity and expertise available to the organization.
- **Align the technical team & business stakeholders on SLO targets.** If engineers cannot deliver on the SLO targets, the organization risks failure to comply with its SLAs to customers.

- **Use an independent SLO for each logical component of the system.** Every system component may impact or contribute to the overall system differently. Therefore, it's important to define optimum SLOs for every system component based on cost, complexity and other associated business and technical challenges.
- **Measure several service level indicators collectively to evaluate a single SLO target.** For instance, the latency, errors and other QoS metrics may be required to evaluate a complete system performance with respect to specific objectives.
- **Document & communicate SLOs for all stakeholders.** This information is often critical for technical teams or business leaders to make relevant decisions.
- **Prioritize SLOs for certain customers.** Paying customers with stringent availability requirements may require a higher SLO baseline than freemium users.
- **Consider SLOs an ongoing commitment** to deliver optimum system performance across various service level indicators. SLOs evolve over time; they cannot be considered as static targets. IT workloads and end-user expectations change on a continual basis. An SLO designed for the workload requirements at present may not be equally valid for its future performance requirements.
- **Keep SLOs simple, few & realistic.** Avoid absolute numbers that are unrealistic. You may set an internal SLO that acts as a safety margin or buffer to deliver a lower SLO target agreed with the end-users.

Perfection isn't necessary

It may not be possible to meet service level objectives 100% of the time. Cloud service providers need to innovate, add features, and update systems, which likely involve temporary downtime across several data center instances.

Consider this a tradeoff to achieve SLOs to deliver better services

Service Level Objectives are all about the targets beyond which the service level is not acceptable to customers and end-users. Setting expectations realistic and not overachieving the targets may be the first step toward delivering a cloud service with acceptable end-user experience at an affordable budget.

Related reading

- BMC Service Management Blog
- What Is High Availability? Concepts & Best Practices
- System Reliability & Availability Calculations
- 6 SLA Best Practices for Service Management Success
- Availability Management & The Role of The Availability Manager
- A Primer on Service Level Indicator (SLI) Metrics
- Service Level Management in ITIL® 4, part of our ITIL 4 Guide