[.]oduct or feature is in a pre-release state and might change or have limited support. For more information, see the <u>pr</u> <u>1 stages</u> (/products/#product-launch-stages).

This page provides information that you might need before <u>creating a service level objective (SLO)</u> (/service-mesh/docs/observability/create-slo) in the Service Mesh Dashboard.

Anthos Service Mesh supports the following types of service level indicators:

- Latency: How long it takes a service to return a response to a request, measured in milliseconds.
- Availability: The fraction of the time that a service responds successfully.

You also define the compliance target that you want from your service. In general, SLOs shouldn't be higher than is necessary or meaningful for your users. Consider at what point users might notice service degradation. For example, if your users cannot tell the difference between a latency of 300ms or 500ms for your service, use the higher value as the latency threshold in the SLO. The lower value is more expensive to meet, and your users won't notice the difference.

When you set a compliance target, consider the end-user requirements for your service. For example, an internal tool used by employees to book vacation time might be fine with a 99% availability target (~3 days of downtime per year). But a critical service for an online store might need 99.999% availability (~5 minutes of downtime per year).

In addition to defining a target for an SLI, an SLO specifies a period of time in which the SLI is being measured. For example, 99% availability over a single day is different from 99% availability over a month. The first SLO would not permit more than 14 minutes of consecutive downtime (24 hrs * 1%), whereas the second SLO would allow consecutive downtime up to ~7 hours (30 days * 1%)."

The compliance period is particularly important when an SLO is included in a service level agreement (SLA) with your users. An SLA is a contract with the users of your service that typically specifies the consequences of not meeting the SLOs. Whether or not you have an SLA with your users is a product or business decision, but for monitoring purposes, you still need to specify a compliance period for your SLOs when you create them in the Service Mesh Dashboard.

When you configure SLOs in the Service Mesh Dashboard, you choose the type of compliance period:

- **Calendar**: When you select **Calendar** as the **Period Type**, you also specify the **Period Length**, which can be a day, week or month. Periods are non-overlapping and fixed to the calendar start and end dates. Compliance can only be evaluated at the end of the period.
- Rolling: When you select Rolling as the Period Type, you also specify the number of days for the Period Length, for example, 30 days. Unlike Calendar periods, rolling periods don't have fixed start and end dates. Anthos Service Mesh continually evaluates SLOs with a rolling compliance period. The oldest data in the previous calculation drops out of the current calculation as it is replaced by new data. A rolling time period provides more compliance measurements because each day, you get a measure of compliance for the last 30 days, rather than one per month. However, services can hover between compliance and noncompliance as the SLO status changes daily.

Another important monitoring concept is the error budget. An SLO specifies an SLI and a target value that measures success of the service in the compliance period. The error budget for an SLO represents the total amount of time that a service can be noncompliant before it is in violation of its SLO. Thus, an error budget is 100% - SLO%. For example, if you have a rolling 30-day availability SLO with a 99.99% compliance target, your error budget is 0.01% of 30 days: just over 4 minutes of allowed downtime each 30 days. A service required to meet a 100% SLO has no error budget.

Error budgets let you track how many bad SLI measurements are allowed to occur during the remainder of your compliance period before the service violates the SLO. You can use the error budget to help manage maintenance tasks like deployment of new versions. When the error budget is close to depleted, it's not a good time to take risky actions like deploying new updates. Conversely, if you have a full error budget near the end of a compliance period, you might want launch new features since the risk of violating the SLO is lower.

If you are measuring an SLO with a calendar compliance period, Service Mesh starts the error budget at the maximum value and reduces the budget over time, triggering an SLO violation when the error

budget drops below 0. Anthos Service Mesh resets the SLO's error budget at the end of the compliance period.

If you are measuring an SLO over a rolling compliance period, you are effectively always at the end of a compliance period. Rather than starting from scratch, old data points are continuously dropped and new data points are continuously added. If a period of poor compliance rolls out of the compliance window, and if the SLO is compliant, the error budget goes up. At any point in time, an error budget \geq 0 indicates a compliant rolling SLO window, and an error budget < 0 indicates a non-compliant rolling SLO window.

- Learn more about SLOs from Site Reliability Engineering at Google:
 - <u>Site Reliability Engineering</u> (https://landing.google.com/sre/sre-book/chapters/service-level-objectives/)
 - <u>The Site Reliability Workbook</u> (https://landing.google.com/sre/workbook/chapters/implementing-slos/)
- <u>Creating SLOs</u> (/service-mesh/docs/observability/create-slo)
- Monitoring SLOs (/service-mesh/docs/observability/monitor-slo)