

Cost caps and alerts

DOCUMENT	VERSION	STATUS
SB-GD-112	1.0, October 2026	Public. Checked against the product, October 2026

Who this is for

Workspace owners and admins who want a spending limit on AI use, with alerts as spend approaches it.

Before you start

- You need the Owner or Admin role. Cost caps live in the Admin center, and only owners and admins can create, change or delete a cap.
- Cost caps need the Business plan or above. On a lower plan the page shows "Business plan required" and links to the plans. Check your plan under Admin center > Billing > Subscription.
- A cap compares against the cost figures SynapBridge records for AI calls made through SynapBridge. Those figures are estimates based on a rate table. Your AI provider's invoice is the source of truth for what you owe.
- Decide the scope, the period and the behavior before you begin. Scope and period cannot be changed after a cap is created.

What a cap is

A cap is one dollar limit on one scope for one period, with one behavior.

- Scope: Entire workspace, a single User, or a Group. For User and Group you enter the user ID or group ID in the Target field.
- Period: Daily or Monthly.
- Behavior: Warn or Block.
- Alert thresholds: percentages of the cap that trigger an alert. The default is 50, 75, 90.

There is one cap for each combination of scope, target and period. Creating a second one for the same combination returns an error that points to the existing cap.

A cap has a single behavior. If you want a warning at one amount and a block at a higher amount, create two caps.

Create a cap

- 1 Open Admin center, then Cost management, then Cost caps.
- 2 Select Create cap.
- 3 Choose the Scope. For User or Group, enter the ID in "Target (user or group ID)". Leave the target blank for Entire workspace.
- 4 Choose the Period: Daily or Monthly.
- 5 Enter the "Cap amount (USD)", for example 100.00.
- 6 Enter "Alert thresholds (%)" as a comma-separated list, for example 50,75,90. Each value must be a whole number from 1 to 100.
- 7 Under "When cap is reached", choose Warn or Block.
- 8 Under alerts, add any of these: a Slack webhook URL, a Teams webhook URL, an alert email. All three are optional and independent.
- 9 Leave Enabled on and select Create cap.

The table then shows the cap with its current-period spend, thresholds and behavior.

Edit, disable or delete a cap

- Select Edit on a row to change the amount, thresholds, behavior, enabled state or alert destinations. Saved webhook URLs are shown masked. To replace one, select Change and enter the new URL. If you leave it untouched, the saved URL is kept.
- Use the Enabled control to switch a cap off without deleting it. Disabling a Block cap removes its block until you switch it back on. The page asks you to confirm.
- Select Delete to remove a cap. This stops its alerts and its block, and it cannot be undone.
- Creating, changing and deleting caps is recorded in the audit trail.

How periods are counted

- Daily is a rolling 24 hours ending now. It is not a calendar day.
- Monthly is the calendar month so far, starting at 00:00 UTC on the first day of the month.

What happens at the cap

- Warn: calls continue. The breach is recorded and alerts are sent to the destinations you set.

- **Block:** before a call to an AI provider is dispatched, SynapBridge adds the estimated cost of that call to the spend so far. If the total reaches the cap, the call is refused and never sent.

A blocked call returns the error code `spend_cap_exceeded`. Through the API gateway this is HTTP status 402. In the app, the person sees a message that depends on the feature. Two examples from the product are "Your workspace has reached its AI spend limit. A workspace admin can raise it, then you can retry." and "Your workspace's AI spend cap has been reached, so questions are paused until the cap resets or is raised."

To restore access, raise the cap, switch it off, or wait for the period to roll over.

Group caps apply to the people who are members of the group when the caps are loaded. A group with no members matches nobody.

Alerts

- Destinations per cap: email, a Slack webhook, a Teams webhook. Any combination, or none.
- A Slack webhook URL must start with `https://hooks.slack.com/`. A Teams webhook URL must start with `https://`. Both must be under 512 characters.
- An alert fires once per threshold crossing per period. It does not repeat on each later call.
- Alert sending is switched on per deployment. If your alerts do not arrive, ask SynapBridge support whether cost alerts are enabled for your deployment, then send a test by setting a low threshold on a test cap.

Related pages: Governance > Spend thresholds sets separate soft, hard and kill-switch percentages. Insights > Anomalies & alerts lists spend anomalies and threshold events.

Limits and notes

Note A cap is a check, not a guarantee. Treat it as a control that lowers the chance of overspend, not as a promise that spend cannot exceed the amount.

- Caps are checked before a model call that goes through SynapBridge's governed paths. Some model-call paths are only counted, or are not yet counted. The product's own register lists background and platform jobs, an agent code runner, compare synthesis and a few others as not checked before the call.
- In the product's Budgets and limits information, chat and Compare turns sent to an OpenAI model use a route that runs no spend cap check. Their spend may be counted toward caps on some deployments, but those turns are not stopped. Ask SynapBridge support what applies to your deployment.
- Enforcement must be switched on for the deployment. When it is off, caps are recorded and do not stop any call.

- If the spend check itself cannot run, the default setting allows the call and records the failure. A deployment can be set to stop calls for workspaces with a Block cap, or to stop calls unless the workspace is known to have none. Ask SynapBridge support which setting applies to you.
- The check uses an estimate of the cost of the call before it runs. The final recorded cost can differ.
- Only Daily and Monthly periods exist. There are no weekly, quarterly or custom ranges, and no per-model or per-workspace cap scopes.
- Agent budgets have their own setting and are off unless the deployment turns them on.
- Caps do not cover spend on your AI provider account that happens outside SynapBridge.