Infrapilot

Security overview

One page, generated from the security model
18 November 2026. The product is in alpha.

Pilot changes production infrastructure without a person reviewing each change. Here is what keeps that safe, and where the guarantees stop.

Pilot proposes; an engine decides

Pilot has no shell, no cloud SDK and no credentials of its own. It can only submit **transactions**, and each one passes through an engine that is ordinary code:

1. **Capability**: only this fleet's resources; never the intents or boundaries.
2. **Rules**: the provider's schema; the spec's bounds, cost ceiling and boundaries; autoscaler ownership; at most half of any capacity removed per change; delete only what Pilot created; irreversible operations, including undoing a create, need a named person.
3. **Preconditions**: if anything changed since Pilot read it, the change is refused.
4. **Change budget**: refills only while objectives are met.
5. **Health gate**: after a change, the resource is checked; on failure the change is undone automatically.

None of this is a prompt. A wrong or manipulated model still can't get past the engine, or beyond the role you created.

The ledger

Every change, refusal, check, undo, outside change, request and answer is appended to a ledger. Each entry carries the SHA-256 hash of the one before. `GET /api/workspaces/<id>/ledger/verify` recomputes the chain and replays it. `GET /api/workspaces/<id>/export` gives you all of it.

One sandbox per workspace

Each workspace's agent runs in its own container, on Cloudflare Containers: its own virtual machine.

- a non-root user, with memory and CPU limits
- its disk is discarded whenever it restarts; what it keeps between runs is in that workspace's state, which no other workspace's agent can reach
- nothing listens inside it, and the only parts of Infrapilot it can reach are its own workspace's state and the model gateway
- no internet access of its own: the only outside hosts it can reach are your cloud provider's API hosts (below)
- only that workspace's cloud credentials

It never holds other customers' data, Infrapilot's database, the encryption key, or Infrapilot's model key.

What a sandbox can connect to

Being switched on during the alpha. The list below is in place and every outbound request is checked against it. Until it has run cleanly against each cloud, a request to a host off the list is logged for Infrapilot's operator and forwarded, not refused. This note is removed when refusal is on.

The network enforces it, not the agent's code.

Your cloud	Hosts
AWS	`*.amazonaws.com`, `*.api.aws`
Azure	`login.microsoftonline.com`, `management.azure.com`
Google Cloud	`*.googleapis.com`

A workspace gets one row: an AWS workspace cannot reach Google's APIs.

What this does not do: the list is by host name, so it allows your provider's APIs as a whole, including API hosts of accounts that are not yours.

Cloud credentials

Cloud	Access	Infrapilot stores
AWS	assumes your IAM role with a workspace-unique external id, an hour at a time	the role ARN. No keys. (Access keys are accepted if you choose them, stored encrypted.)
Google Cloud	impersonates a service account named for your workspace, an hour at a time	the account email. No keys. (A key file is accepted if you choose it, stored encrypted.)
Azure	an app registration that trusts Infrapilot's signed tokens for your workspace's subject only (a federated credential)	no secret: the tenant, app and subscription ids

Infrapilot's control plane obtains the hourly credentials, or decrypts the stored secret, and hands them to your workspace's sandbox and to nothing else.

Tokens (sign-in links, sessions, setup and API tokens, answer links) are stored only as hashes.

To cut Pilot off, delete the role, service account or app registration. Everything keeps running as left.

People

- Sign-in is by emailed link.
- Only people change intents and boundaries.
- Only the owner adds or removes members and deletes the workspace.
- Email answer links are personal and expire after a week.
- **The Slack app** keeps one bot token per installation, encrypted and never returned by the API.
- A pause stops the next change, including inside a review that is already running.
- Infrapilot's operator can pause Pilot in your workspace, undo one of its changes, or give up Infrapilot's access to your account, when something has gone wrong.
- An approved operation runs only if the resources are unchanged since Pilot asked.

Limits of these guarantees

- **A role that allows too much.** The engine enforces the spec; your cloud enforces the role.
- **A spec that asks for the wrong thing.** A ceiling set too high is respected as written.
- **Outside changes.** Detected, attributed and kept or reverted, but not prevented.
- **Resources Infrapilot has no price for** count as zero toward a ceiling.
- **Secrets in plain configuration.** Plain environment-variable values and secret-named fields are hidden: they are not recorded, shown to Pilot, or changed when Pilot updates the resource.
- **Text Pilot reads.** Resource names and settings, error messages and notes from people reach the model marked as data, never as instructions, and the engine applies the same rules whatever the model was told.
- **Secrets in a workload's** spec would be recorded.