

Deployment Architecture

Questions this doc answers

- Which deployment shape fits our risk posture?
- What does private / air-gapped evaluation look like?
- What do we need ready before a pilot?

Deployment modes

Mode	Ownership	Network boundary	Typical fit
Hosted	Reflect Memory cloud	Public internet (TLS)	Fast evaluation, lower infra burden
Isolated-hosted	Dedicated runtime/DB per tenant	Managed / restricted	Stronger isolation without self-ops
Self-host	Your VPC or air-gapped network	Private	Regulated / residency / no-egress requirements

Private-deploy expectations

Self-host and air-gapped pilots typically require:

- Deploy inside your network (container or package delivered under NDA)
- Optional SSO / OIDC for enterprise identity
- Controls to keep model traffic on approved internal endpoints
- Ability to disable public webhook exposure
- Audit export for compliance review

Exact env, networking, and install steps are provided in the private deploy guide during onboarding — not in this public bundle.

Reflect cloud (hosted)

Reflect’s production cloud runs on Reflect-owned dedicated infrastructure (not a consumer PaaS). Buyers evaluating residency should treat hosted as Reflect-operated cloud; isolated-hosted and self-host remain available when that is not enough.

Pilot sequence

1. Scope call — mode, identity, residency, and compliance requirements.
2. Provision pilot (hosted, isolated, or private package).
3. Connect AI tools (MCP and/or REST) with customer-controlled credentials.
4. Validate audit export, Ambient preference, and team sharing against your policy.
5. Expand seats / harden egress controls as the pilot graduates.

Contact

Private deploy packages and NDA technical packs: `Van Mendoza <vm@reflectmemory.com>`