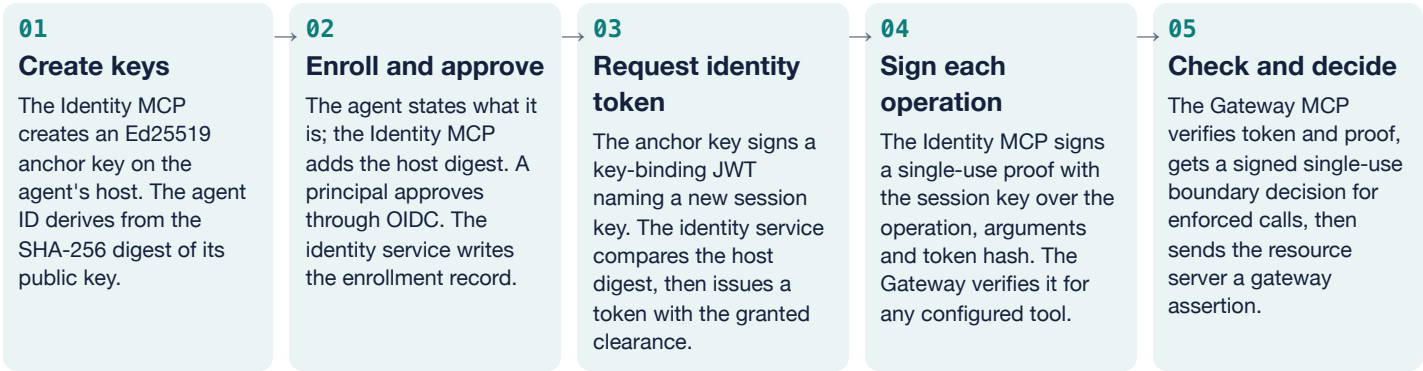


How Oathera identifies and checks an AI agent

How the identity token carries your agent's DNA, which standard supplies what, what Oathera adds, and what is permanent versus per call. Hosted demo-profile services reset when a container sleeps.

LIFECYCLE

How an agent gets, and keeps, an identity



WHAT LASTS, WHAT DOESN'T

Stable for the agent's life on a host, fresh for every request

Stable · created once per agent and host		Per request · short-lived, then gone	
Anchor key	Ed25519 software key in a file only the agent's user can read. Lasts until revoked.	Session key	Ed25519, held in memory, never written to disk. Replaced at every issuance.
Enrollment record	Anchor key, principal, label, platform, host block, NHI attributes, key and revocation epochs.	One token's life	
Enrollment attributes	Agent's self-description, host block and principal: together your agent's DNA, stable on its host.	Identity token	Names the session key in cnf. Reused until 80% of its life has passed.
Host digest	SHA-256 over machine ID, hostname, CPU, memory, MAC addresses, cloud instance ID. Survives reboot.	120 seconds default	
		Operation proof	Signed by the session key. Accepted within 300 seconds and used once.
		One call	
		Boundary decision	Signed allow or deny, bound to one gateway and request. Assertion: 30 seconds, one use.
		15 seconds, one use	

INSIDE THE IDENTITY TOKEN

What a receiving service can read

CLAIM	MEANING
sub	The agent ID, derived from the anchor key digest.
cnf.jwk	The session public key that may use this token.
aud	The single resource the token is minted for, in the audience-bound form.
exp	Expiry. The default lifetime is 120 seconds.
principal	The person who approved the enrollment.
Host posture	Truncated host digest, confidence, attested flag, findings count and drift marker.
clearance	The clearance granted, never more than the evidence supports.
capabilities	Named actions allowed. Privileged ones cap lifetime at 60 seconds.

STANDARDS MAP

What comes from open standards, and what Oathera adds

STANDARD	WHAT IT GIVES	WHAT OATHERA ADDS
JWT, JWS, RFC 7800	JWT format; cnf names the session key; key digest gives the agent ID.	Typed claim sets, and anchor-signed replacement of the session key at every issuance.
RFC 9421	Proof of key possession on HTTP calls.	Signs enrollment, token requests and gateway decision requests. The per-call operation proof is separate.
OIDC with PKCE	The principal signs in at the customer's identity provider.	Sign-in state is bound to the anchor key being enrolled. The hosted demo uses an internal demo IdP.
SPIFFE and SPIRE	Optional enrollment evidence: a SPIRE-issued JWT-SVID.	Identity stays the anchor key. Trust domains are set with TRUSTED_SPIFFE_DOMAINS.
W3C verifiable credentials	Agent, sponsor and boundary credentials at token issuance, with suspend, reinstate and revoke status.	Set with CREDENTIAL_FORMAT=vc; default json. Boundary decision reads credential status. Credentials never authorize resources.
OWASP NHI Top 10	The questionnaire categories.	A risk tier from 1 to 4 and findings for NHI1 to NHI10, when the agent sends answers.

INTEGRATE

Three steps to a provable agent

1 Run the MCP bundle. Runs the Identity MCP, Gateway MCP and a sample resource server in the demo profile, pointed at the hosted identity service. AIS_URL overrides it.

2 Enroll and approve. Call enroll, then have the principal approve at the identity provider. Set IDENTITY_MCP_REQUESTED_CLEARANCE and HOST_BINDING as needed.

3 Front your tools. attachOatheraGateway runs in-process with a Node MCP SDK server. createOatheraGatewayProxy fronts a stdio or HTTP server in any language.

```
# Identity MCP
AIS_URL=https://ais.oathera.org
IDENTITY_STATE_PATH=<state file path>
IDENTITY_MCP_REQUESTED_CLEARANCE=0
HOST_DISCOVERY_LEVEL=<level>
# Identity service
HOST_BINDING=warn
# Gateway MCP
GATEWAY_AIS_DECISION_URL=<decision URL>
GATEWAY_DEFAULT_DENY=true
# In-process: attachOatheraGateway
```

SECURITY PROPERTIES

What holds

Host mismatch can refuse the token

HOST_BINDING=enforce returns 403 host_changed. The default, warn, issues the token with a host-drift claim and an alert.

Every call carries an operation proof

Your agent's fingerprint on every call: a single-use operation proof (not DPoP), key digest and host digest. Copies leave a different one.

Revocation applies at next request

A revoked agent's next token request and boundary decision are refused. Issued tokens stay valid until they expire.

Audience-bound token in production

The production profile issues tokens for one audience: 300 seconds, or 60 if any capability is privileged.