CrewAI connector configuration

This reference describes the product settings accepted by aip-host-crewai and aip-crewai-sidecar at the reviewed source revision. A complete deployment combines them with the shared connector-host identity, PostgreSQL, routing, signing, lif

SECTION	Connectors
TYPE	Connector
PROTOCOL	AIP 1.0
SOURCE	docs/connectors/crewai/reference/configuration.md

All CrewAI product settings are read when their process starts. Only the shared non-secret credential-revision policy is re-read before provider side effects.

Configuration identity

Property	Value
Rust host binary	`aip-host-crewai`
Python package	`aip-crewai-sidecar` `1.0.0`
AIP workspace version	`1.0.0`
Reviewed AIP source	`97be86e9efedf07ecf1783b03800f683f107fb04`
Connector ID	`crewai`
Connector profile	`aip.connector.crewai.v1`
CrewAI package in default product image	`1.15.5`
Optional source-overlay revision	`bfa652a7be8637562cc9b0833f75d927a64552d1`, source version `1.15.2`
Python range	`>=3.10,<3.14`

The sidecar artifact also pins FastAPI, Pydantic, Uvicorn, the Python base image, the package installer, and the complete Python lockfile.

Configure the Rust host

Environment variable	Command-line flag	Required	Default	Meaning
`AIP_CREWAI_SIDECAR_URL`	`--sidecar-url`	Yes	None	Fixed base URL of the separately deployed sidecar
`AIP_CREWAI_SIDECAR_TOKEN_FILE`	`--sidecar-token-file`	Yes	None	Owner-controlled bearer-token file used by the host
`AIP_CREWAI_CREWS_FILE`	`--crews-file`	Yes	None	JSON array of admitted crew descriptors
`AIP_CREWAI_MAX_RESPONSE_BYTES`	`--max-response-bytes`	No	`67108864`	Maximum blocking response or complete sidecar event stream

The source rejects credentials, a query, and a fragment. HTTPS is required except for explicit loopback HTTP, and redirects are disabled.

A path prefix is accepted but discarded by absolute `/health` and `/jobs` joins. Configure an origin-only value to keep reviewed and effective routes identical.

The host reads a regular owner-only file with a 16 KiB ceiling, trims surrounding ASCII whitespace, and requires non-empty UTF-8. The connector rejects an empty value or any control character; it does not reject an interior plain space.

`AIP_CREWAI_MAX_RESPONSE_BYTES` accepts `1` through `268435456`. The same configured value bounds one blocking sidecar response and the accumulated bytes of one event stream.

Configure admitted crew descriptors

The descriptor file is a JSON array of one to 1,000 entries and may not exceed 4 MiB.

```
JSON
[
  {
    "id": "support-primary",
    "name": "Support primary",
    "description": "Reviewed support workflow",
    "allowed_operations": [
      "run",
      "status",
      "events",
      "cancel"
    ]
  }
]
```

Field	Required	Constraint
`id`	Yes	Non-empty after trimming, unique, no control characters, valid inside capability IDs
`name`	Yes	Non-empty after trimming; no field-specific byte or control-character bound at this revision
`description`	No	No field-specific byte or control-character bound at this revision
`allowed_operations`	No	Non-empty known-operation set; omission uses `run`, `status`, `events`, and `cancel`

The 4 MiB file limit remains the combined payload ceiling. Apply a stricter deployment schema before descriptor text reaches a public manifest.

The ten accepted operation names are `run`, `status`, `events`, `cancel`, `batch_run`, `replay`, `train`, `test`, `knowledge_query`, and `memory_reset`.

Configure the sidecar listener and authentication

Environment variable	Required	Default	Accepted value
`AIP_CREWAI_HOST`	No	`0.0.0.0`	Uvicorn bind host
`AIP_CREWAI_PORT`	No	`8090`	Integer Uvicorn port
`AIP_CREWAI_BEARER_TOKEN_FILE`	Production	None	Protected UTF-8 file, 1 byte through 16 KiB
`AIP_CREWAI_BEARER_TOKEN`	No	None	Mutually exclusive inline token for controlled development
`AIP_CREWAI_ALLOW_UNAUTHENTICATED`	Development only	`false`	`1` or `true`, effective only on a loopback bind host

The bearer file must be a regular non-symlink with no group or world permission bits. The sidecar checks the inspected and opened inode and uses constant-time comparison for the complete Bearer header.

The controlled fleet places both components on a private network and uses one shared mounted token file. It still treats the host and sidecar as separate process identities.

Configure the registry and operation set

Environment variable	Required	Default	Meaning
`AIP_CREWAI_REGISTRY`	Yes	None	Python `module:attribute` resolving to a non-empty crew mapping
`AIP_CREWAI_ALLOWED_OPERATIONS`	No	`run`	Comma-separated startable operation set

Registry values are zero-argument factories or cloneable Crew instances. The sidecar creates a new result or deep clone for every job.

The allowed sidecar set can contain:

```
TEXT
run,batch_run,replay,train,test,knowledge_query,memory_reset
```

`status`, `events`, and `cancel` are control routes and appear in sidecar health automatically. Supplying a control operation in the environment value is invalid.

The host descriptor may be narrower than sidecar policy. Host readiness fails if any descriptor operation is absent from the authenticated sidecar health response.

Configure durable state

Environment variable	Required	Default	Range or role
`AIP_CREWAI_STATE_FILE`	Production	None	Durable JSON journal path
`AIP_CREWAI_ALLOW_EPHEMERAL`	Development only	`false`	Permits no journal only on a loopback bind host
`AIP_CREWAI_MAX_RETAINED_RUNS`	No	`2000`	`max_concurrency` through `100000`
`AIP_CREWAI_MAX_EVENTS_PER_RUN`	No	`1000`	`2` through `100000`
`AIP_CREWAI_MAX_EVENT_BYTES`	No	`262144`	`1024` through `4194304`
`AIP_CREWAI_MAX_STATE_BYTES`	No	`268435456`	`1048576` through `4294967296`
`AIP_CREWAI_STREAM_CHECKPOINT_EVENTS`	No	`128`	`1` through `10000`

The journal parent must be a non-symlink directory with no group or world write bit. The sidecar holds an exclusive non-blocking lock in a regular `0600` companion file.

An existing journal must be a regular non-symlink `0600` file within the configured byte limit. Restore rejects duplicate Action IDs, excessive runs, excessive events, and oversized events.

State writes use a same-directory temporary file, file sync, atomic replace, and directory sync. A configured maximum is a retention and integrity bound, not permission to delete records inside the AIP replay window.

The default product image places `CREWAI_STORAGE_DIR` and `HOME` under the sidecar volume. Persist that volume with the journal; losing the durable home can make locally encrypted provider credentials unreadable.

The optional source-overlay Dockerfile does not set those two paths at this revision. A deployment using it must configure and persist an equivalent CrewAI home explicitly.

Configure execution and training

Environment variable	Required	Default	Range or role
`AIP_CREWAI_MAX_CONCURRENCY`	No	`4`	`1` through `1024` concurrent jobs
`AIP_CREWAI_TRAINING_DIR`	When `train` is enabled	None	Protected directory for controlled training artifacts

The training directory is created with owner access when absent. It must be a non-symlink directory and must not be group- or world-writable.

The Action's `timeout_ms` accepts `1` through `3600000`. Omission uses the sidecar default of `600000` milliseconds.

Telemetry defaults are set before the registry loads:

```
TEXT
CREWAI_DISABLE_TELEMETRY=true
CREWAI_TRACING_ENABLED=false
OTEL_SDK_DISABLED=true
```

An explicit deployment environment can override these existing variables. Review content and destination exposure before doing so.

Apply relevant shared host settings

The global host reference owns the exhaustive common matrix. These groups are required for CrewAI behavior.

Identity and route

Setting	Role
`AIP_CONNECTOR_HOST_PUBLIC_ENDPOINT`	Reachable native endpoint ending in `/aip/v1/messages`
`AIP_CONNECTOR_HOST_TYPE_ID`	Admitted CrewAI connector type
`AIP_CONNECTOR_HOST_VERSION_ID`	Immutable admitted version
`AIP_CONNECTOR_HOST_INSTANCE_ID`	Logical runtime and task-store boundary
`AIP_CONNECTOR_HOST_REPLICA_ID`	Concrete process identity
`AIP_CONNECTOR_HOST_TENANT_ID`	Verified tenant owned by the instance
`AIP_CONNECTOR_HOST_MEMBERSHIP_ID`	Stable membership assertion
`AIP_CONNECTOR_HOST_ARTIFACT_DIGEST`	Exact Rust-host artifact digest
`AIP_CONNECTOR_HOST_SECRET_PROVIDER_REF`	Non-secret secret-provider reference

An optional external-account ID and system must appear together. The reviewed host does not compare that metadata automatically with a crew descriptor.

PostgreSQL, signing, and lifecycle

Setting	Required or default
`AIP_CONNECTOR_HOST_BIND`	Default `0.0.0.0:8081`
`AIP_CONNECTOR_HOST_DATABASE_URL_FILE`	Required owner-controlled PostgreSQL URL
`AIP_CONNECTOR_HOST_SIGNING_SEED_FILE`	Required owner-controlled Ed25519 seed
`AIP_CONNECTOR_HOST_CONTROL_ENDPOINT`	Required lifecycle endpoint
`AIP_CONNECTOR_HOST_CONTROL_DID` or `_FILE`	Exactly one control-plane DID source
`AIP_CONNECTOR_HOST_GATEWAY_PRINCIPAL_ID`	Required trusted gateway principal
`AIP_CONNECTOR_HOST_GATEWAY_DID` or `_FILE`	Exactly one gateway DID source
`AIP_CONNECTOR_HOST_TRUST_DOMAIN`	Required trust domain
`AIP_CONNECTOR_HOST_TLS_CA_FILE`	Optional private-PKI root
`AIP_CONNECTOR_HOST_MAX_IN_FLIGHT`	Default `32`
`AIP_CONNECTOR_HOST_LEASE_TTL_MS`	Default `30000`
`AIP_CONNECTOR_HOST_HEARTBEAT_MS`	Default `10000`
`AIP_CONNECTOR_HOST_DRAIN_TIMEOUT_MS`	Default `30000`
`AIP_CONNECTOR_HOST_HEALTH_TIMEOUT_MS`	Default `2000`
`AIP_CONNECTOR_HOST_CONTROL_TIMEOUT_MS`	Default `5000`

Shared credential ID, issuer, scopes, revision, and revision-policy settings are optional as one complete unit. They fence route authorization but do not replace sidecar or CrewAI provider credentials.

Stream callback

Setting	Default
`AIP_CONNECTOR_HOST_CALLBACK_ENDPOINT`	None
`AIP_CONNECTOR_HOST_CALLBACK_ALLOW_HTTP`	`false`
`AIP_CONNECTOR_HOST_CALLBACK_ALLOW_PRIVATE`	`false`

Configure callback behavior for the selected gateway topology. The product settings do not make a callback mandatory for blocking operations.

Use the effective limits

Boundary	Reviewed value
Native AIP request body, common default	4 MiB
Crew descriptor file	4 MiB
Crew descriptors	1,000
Host and sidecar bearer file	16 KiB
Sidecar-validated operation input	1 MiB
Sidecar output	4 MiB
Sidecar event, default	256 KiB
Sidecar events per run, default	1,000
Sidecar retained runs, default	2,000
Sidecar journal, default	256 MiB
Connector SSE frame	1 MiB
Connector SSE events	100,000
Host response, default	64 MiB
Host response, maximum	256 MiB
Sidecar execution timeout, default	600,000 ms
Sidecar execution timeout, maximum	3,600,000 ms
Rust HTTP connect timeout	10 seconds

The effective ceiling is the smallest applicable bound. Raising the host response limit does not enlarge the input, sidecar output, event, native envelope, or journal limits.

The reviewed Rust client has no explicit total HTTP request deadline. The sidecar's operation timeout bounds its wait for a job, but it does not replace a transport-wide client deadline.

Understand readiness

The Rust connector sends an authenticated `GET /health`. It reports ready only when the response has all of these properties:

- `status` is `ready`;
- `durable_journal` is `true`;
- every descriptor crew ID appears in `crews`;
- every descriptor operation appears in `operations`.

Common host readiness also requires PostgreSQL, a valid lease, correct lifecycle state, and other shared controls.

Readiness does not run a crew, model, tool, training task, evaluator, knowledge query, memory reset, cancellation race, or event recovery sequence.

Know startup and reload ownership

Value	Read
Sidecar base URL, token, descriptors, and response bound	Rust-host startup
Listener, sidecar token, registry, operations, and runtime limits	Sidecar startup
Journal and training paths	Sidecar startup
Shared identity, storage, endpoints, and lifecycle limits	Rust-host startup
Connector task-store scope	Constructed from the instance ID at startup
Credential-revision policy file	Startup, then before provider side effects

Change startup-owned values through a controlled replacement. Editing a mounted file does not update a running registry, descriptor set, bearer token, manifest, instance scope, or execution limit.

Diagnose configuration failures

Failure	First decision
Misleading sidecar path prefix	Supply an HTTPS origin-only value; absolute route joins discard the prefix
Token file is rejected	Check regular-file type, permissions, UTF-8, size, and content without printing it
Descriptor file is rejected	Check its 4 MiB bound, count, non-empty IDs and names, ID controls, uniqueness, capability syntax, and operation set
Registry is rejected	Verify importability, `module:attribute`, non-empty mapping, and factory or clone shape
Durable state is required	Configure a protected state path, not the loopback-only ephemeral exception
Journal is already owned	Stop the competing sidecar or assign an independent state volume
Training directory is required	Configure its protected path before enabling `train`
Runtime integer is outside range	Correct the exact setting instead of weakening validation
Host remains unready	Compare authenticated sidecar crews and operations with every descriptor
Journal retention is full	Archive only terminal records after the required replay-retention window

Related documentation

- CrewAI connector overview
- Authentication and admitted crews
- Connector host configuration
- Connector host lifecycle
- Environment variables