

Hermes Agent isolated-live qualification on 11 July 2026

Historical reported result: PASS. The retained narrative gives the execution date as 11 July 2026 but does not retain an exact UTC start or end.

SECTION	Qualification and Evidence
TYPE	Qualification evidence
PROTOCOL	AIP 1.0
SOURCE	docs/testing/hermes-agent-isolated-live.md

This result applies only to the recorded AIP image, Hermes Agent image, upstream revision, two-endpoint topology, model route, and seven assertion groups. It does not qualify the current documentation review revision `97be86e9efedf07ecf1783b03800f683f107fb04`, a different model provider, or a production deployment.

Historical implementation naming: this record documents a pre-v2.0.0 implementation. Executable, crate, configuration, image, and identity names below are retained as evidence; use current reference documentation for present names.

Understand the evidence level

The only retained result is a narrative report. No machine-readable result, raw driver output, command log, signature, full checksum manifest, or cleanup record accompanies it. The report contains exact image digests, execution identifiers, counts, and outcomes, but those observations cannot be independently recomputed from the public record.

Evidence	What it supports	Limitation
Historical narrative	Reported date, status, identities, topology, security controls, assertions, exact execution identifiers, and results	It is a human-readable claim record rather than a signed or machine-generated result
Historical source	The driver and AIP base commit used to frame the campaign	The image included uncommitted qualification work with no retained worktree digest
Current pinned source	The same operator-driver file remains present and byte-identical	Other runtime, deployment, and connector components changed; file equality does not renew the result
Official Hermes source	The recorded upstream commit and exact Git tree	A source commit does not identify the retained Hermes image contents by itself

The historical narrative retained outside this v2 page has SHA-256 digest `0956b7fe3f1c96ee6696d8316d5f2de4f589a72a874f0920ceb055f1a0d19a71`.

Identify the recorded artifacts

Artifact	Recorded identity
AIP source base commit	c51e149ce1df762cb62bc736edfb81a674f56852
AIP base Git tree	e36378e93846cdf6730fd69a80462f9ebc471061
Qualification record and driver commit	f14af77e464625b92bd6c1b86774d1449acd8f3f
AIP qualification image	aipd:local@sha256:00e468c3603941a091e5f9c473b121ef57c9d0f7eaf10f4e07ebfcdf58579786
Hermes Agent upstream commit	7426c09beee73bdf94d916015bac71384f6bc92
Hermes Agent Git tree	583423a976e761d3f076d2cf7f7859c13e40cc29
Hermes Agent image	hermes-agent:aip-lite@sha256:15912eaf6211f3b953ad6ed2268545a521cb0c440604231cc61dabfd9666f32c
Architecture	linux/arm64
Reported model route	cohere/north-mini-code:free through an OpenRouter-compatible provider

The official Hermes Agent checkout confirms the upstream commit and tree. A historical AIP checkout confirms the AIP base commit and tree.

The report says the AIP image also contained then-current connector qualification work, but it records no source digest for that worktree. The direct child commit identified above retains the report and driver, but no recorded digest binds its complete tree to the image. The image digest identifies what ran without making its complete source independently reconstructable.

Review the isolated topology

The historical report describes:

- two isolated Hermes Agent containers;
- one AIP daemon with MCP Streamable HTTP and native AIP access;
- an OAuth 2.0 token-introspection boundary;
- NATS and a deterministic PostgreSQL-backed system of record;
- an external orchestrator and a separately authenticated human approver;
- a configured external model route.

This topology tested governed operator and delegation behavior. The controlled system of record was part of the test boundary, not one of the six maintained public connectors and not a production provider.

Review the reported security boundary

The narrative reports that both Hermes endpoints, the external orchestrator, and the human approver used distinct principals and credentials. An unauthenticated MCP `initialize` request returned HTTP 401 through the token introspection boundary.

It also reports these controls:

- each Hermes endpoint exposed only its selected AIP tools;
- recursive delegation was absent from the model-facing allowlist;
- Hermes received no credentials for the controlled system of record;
- business facts entered the run only through authenticated AIP tool use;

- native lifecycle state, rather than final model prose, supplied the

authoritative delegated result;

- action contract, transport principal, scope, identity, tenant, and result

integrity were checked before accepting that result;

- the report retained no access token, provider key, or introspection secret.

No raw authentication exchange or redaction log is retained with the report, so these are reported controls rather than independently replayable evidence.

Reconstruct the reported procedure

The campaign reportedly followed this sequence:

1. Start the isolated AIP implementation stack, authorization boundary, NATS, controlled record service, and two Hermes Agent endpoints.
2. Verify that unauthenticated MCP initialization fails and that each embedded authenticated Hermes MCP client discovers the expected tool surface.
3. Submit one governed operator action to each endpoint and observe native pending approval.
4. Approve each immutable request through the separately authenticated native principal.
5. Observe structured Hermes runs, incremental stream chunks, and AIP tool start and completion events.
6. Verify exactly one downstream business action and compare its structured result with an independently read expected record.
7. Reject one delegated request outside the allowed scope.
8. Execute one allowed first-class delegation and compare the protocol-assigned child action with the actual downstream action.
9. Repeat the operator and delegation path independently for the second Hermes endpoint.

The report also states that a standalone Hermes MCP check authenticated and discovered 89 tools for each endpoint. This count is historical and topology-specific.

Inspect the required assertions

The driver failed unless every endpoint satisfied all seven boundaries.

Assertion	Required result
Independent approval	The operator action enters pending approval and resumes only after a separately authenticated decision
Structured execution	Hermes starts a structured run and retains incremental stream chunks
AIP tool lifecycle	The stream contains successful AIP tool start and completion events
Trusted data path	Expected business facts arrive through the AIP tool rather than direct database access or model memory

Assertion	Required result
Scope enforcement	A delegated request outside the endpoint allowlist is rejected
Delegation identity	One child action is assigned and the executed downstream action has exactly the same identifier
Independent endpoints	Both isolated Hermes endpoints complete the full operator and delegation path

The equality between the assigned child action and the executed downstream action is the decisive delegation invariant. Model-generated prose or a replacement identifier cannot satisfy it.

Inspect the reported observations

The public v2 record omits controlled-system capability names, disposable business records, and execution identifiers. The retained historical source remains the audit input for authorized reviewers.

Observation	Endpoint A	Endpoint B
Stream chunks	74	76
Pending and authenticated approval	PASS	PASS
AIP tool start and completion	PASS	PASS
Exactly one direct downstream action	PASS	PASS
Out-of-scope delegation rejected	PASS	PASS
Child action equals delegated downstream action	PASS	PASS
Structured business output matches expected record	PASS	PASS

These values are transcribed from the narrative report. No separately retained driver JSON is available to verify them.

Account for missing and excluded evidence

The result does not establish:

- qualification of any connector other than Hermes Agent;
- current behavior of revision
97be86e9efedf07ecf1783b03800f683f107fb04;
- exact reconstruction of the uncommitted AIP qualification worktree;
- independent proof of the model-provider response, raw streams, native state, database reads, or authorization exchange;
- model quality, provider availability, latency, capacity, or cost;
- production TLS, workload isolation, secret management, egress policy, storage, tenancy, or disaster recovery;
- execution of the historically listed deterministic tests and static checks;
- teardown of the isolated stack or cleanup of every disposable effect.

The current qualification standard makes independently verified cleanup part of `PASS`. The historical report has no cleanup record. Its status remains a reported historical `PASS` under the scope used on 11 July 2026, but it cannot be accepted as a new campaign result under the current standard.

Understand the current source relationship

The operator-driver source has the same SHA-256 digest at the historical qualification commit and at the pinned review revision:

```
TEXT
2464b1af2d40ab37280bf184377ceb79991823bac44d660c8247c5a54f19b15d
```

This equality confirms that the seven driver checks did not change in that file. It says nothing about changes in the daemon, runtime, gateway, MCP server, deployment, images, credentials, Hermes configuration, or provider.

Run only a new campaign

Use the current [live-product procedure](#) and the retained operator entry point to create a new result. Record protected credentials outside the command history and supply two endpoint identifiers from the isolated deployment.

```
SH
cargo run -p aip-connector-hermes-agent \
  --bin aipd-hermes-operator-smoke -- \
  --endpoint-id "hermes-qualification-a" \
  --endpoint-id "hermes-qualification-b"
```

The command requires the MCP and native approver tokens described by the procedure. A successful new exit does not repair the missing historical evidence; retain the new driver output, identities, cleanup result, and checksum manifest as a separate campaign.

Related documentation

- [Testing and qualification](#)
- [Live-product qualification procedure](#)
- [Connector upstream baselines](#)
- [Implementation status](#)
- [Hermes Agent connector](#)