

MYTHOS-CMP-AUTOMATION-PORT-0001

Agentic, Infrastructure, and Automation Comparative Evaluation Portfolio

CMP-004, CMP-007, and CMP-008 with current release deltas, proof gates, and profile-specific decision posture.



CMP-004



CMP-007



CMP-008

PORTFOLIO POSITION / CURRENT EVIDENCE BOUNDARY

Agentic, Infrastructure, and Automation Comparative Evaluation

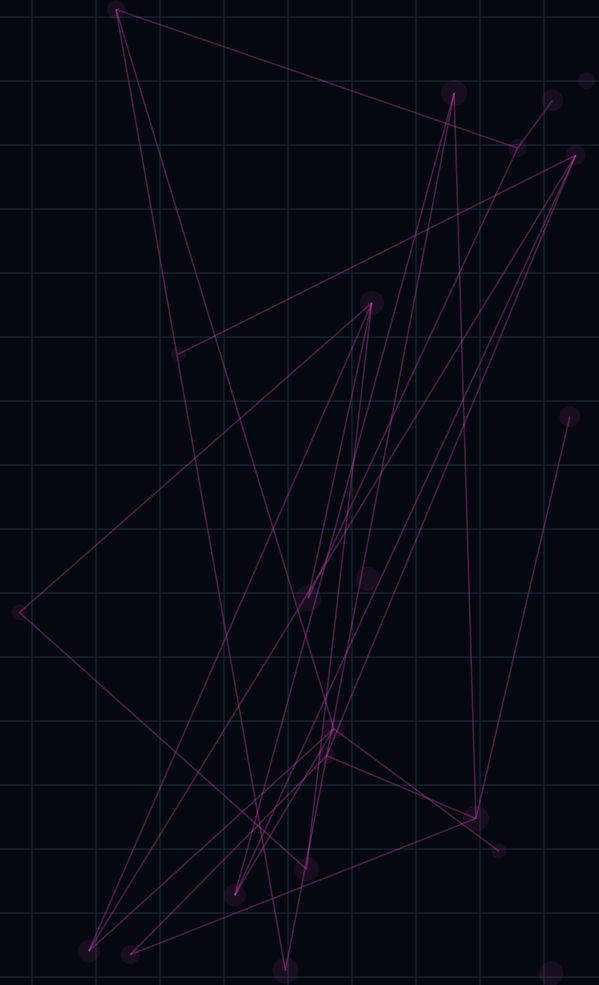
CMP-004, CMP-007, and CMP-008 with current release deltas, proof gates, and profile-specific decision posture.

CMP-004	Agentic Frameworks	CURRENT LAB RERUN PENDING
CMP-007	IaC and Control Planes	CURRENT LAB RERUN PENDING
CMP-008	Configuration Management	CURRENT LAB RERUN PENDING

CMP-004

Agentic Frameworks

Agent runtimes, orchestration and governed infrastructure automation



CURRENT EVIDENCE SNAPSHOT

Official product sources refreshed through 2026-09-17

Controlled Mythos laboratory rerun: NOT ASSERTED

VERSION 1.2.0-candidate

CANDIDATE COMPARATIVE EVALUATION / PUBLIC

BENCHMARK DOCTRINE

Gates before scores. Evidence before certainty. Profile fit before universal ranking.

Agentic Frameworks

Agent runtimes, orchestration and governed infrastructure automation

DOCUMENT ID

CMP-004

VERSION

1.2.0-candidate

STATUS

Candidate Comparative Evaluation

PUBLICATION DATE

2026-09-17

SCHEDULED REVIEW

2026-12-16

CANONICAL ROUTE

/library/evaluations/cmp-004/

3

CANDIDATES

18

ATOMIC CRITERIA

9

MANDATORY GATES

3

WORKLOAD PROFILES

E2

CURRENT MAX EVIDENCE

CURRENT DECISION BOUNDARY

Reopen the comparison. AgentCore and Itential changed materially after the July baseline while LangGraph continued to evolve. No inherited ranking should be promoted to a current recommendation until durable execution, authority, observability, replay, migration, and failure tests are rerun.

NO CURRENT UNIVERSAL WINNER IS PUBLISHED FROM THIS REFRESH.

July 24 baseline scores are preserved as lineage and sensitivity evidence, not silently reissued as September laboratory results.

LangGraph

Python 1.2.11 / JS 1.4.15

REFRESH TRIGGER

Durable graph runtime and checkpoint ecosystem continue to evolve across Python and JavaScript lines.

Amazon Bedrock AgentCore

Managed service / Sept 2026 updates

REFRESH TRIGGER

Batch Evaluations and A/B Testing are now GA, materially strengthening the observe-evaluate-improve proof loop.

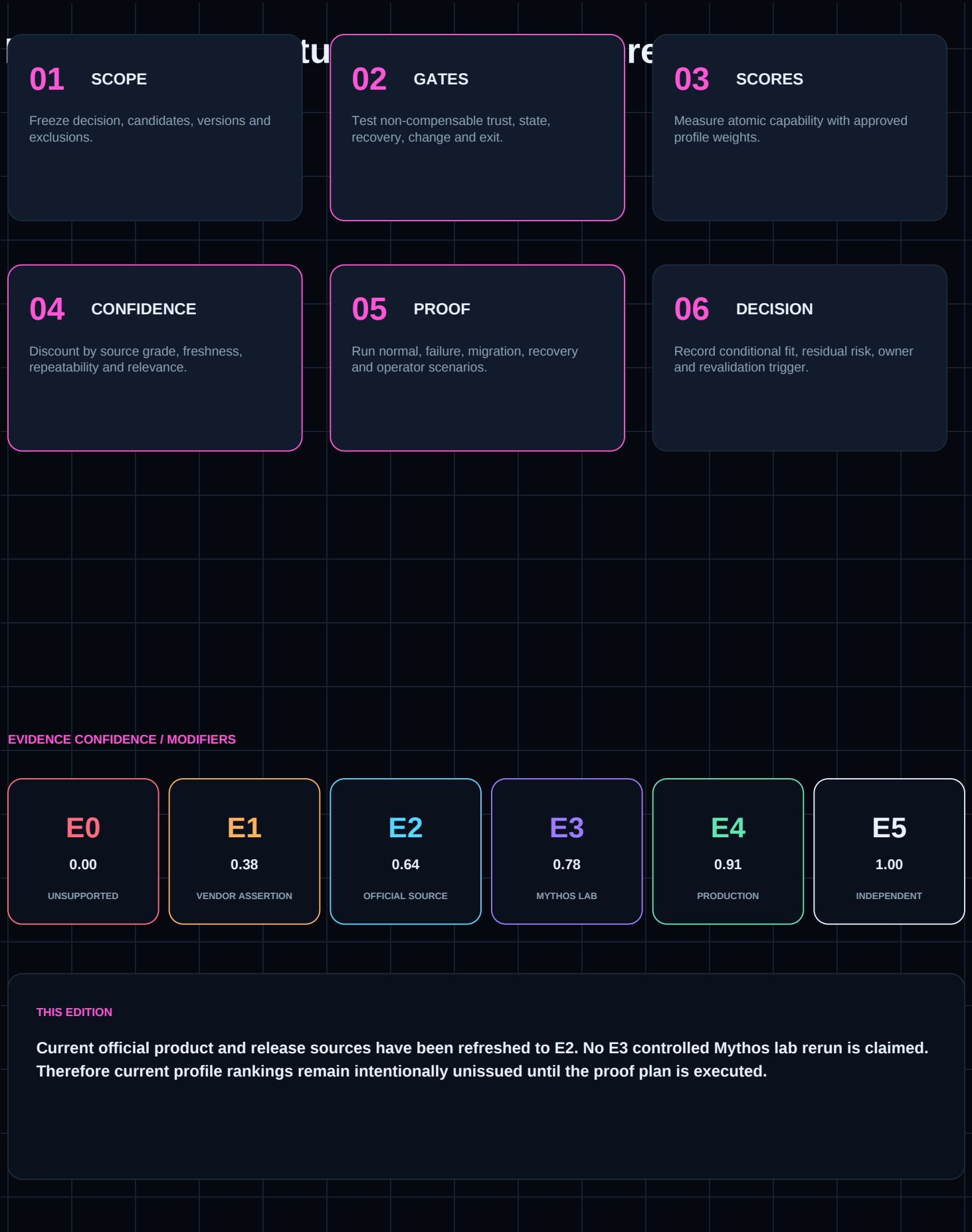
Itential Automation Platform

Platform 6.5 / Cloud Jul-Aug 2026

REFRESH TRIGGER

FlowAI, Work Center, asynchronous workflow support, and auditable agent sessions expand governed infrastructure-agent operation.

FINAL BENCHMARK SYSTEM / DECISION ARCHITECTURE



What changed since the July benchmark snapshot

Changes below are sourced from current official release documentation. They identify revalidation triggers; they are not translated into new numeric scores without controlled proof.

LangGraph

Python 1.2.11 / JS 1.4.15

MATERIAL DELTA

Durable graph runtime and checkpoint ecosystem continue to evolve across Python and JavaScript lines.

CRITERIA TOUCHED

C01 / C04 / C07 / C11 / C12 / C14 MODERATE REVALIDATION

Amazon Bedrock AgentCore

Managed service / Sept 2026 updates

MATERIAL DELTA

Batch Evaluations and A/B Testing are now GA, materially strengthening the observe-evaluate-improve proof loop.

CRITERIA TOUCHED

C01 / C03 / C07 / C08 / C11 / C18 MATERIAL REVALIDATION

Itential Automation Platform

Platform 6.5 / Cloud Jul-Aug 2026

MATERIAL DELTA

FlowAI, Work Center, asynchronous workflow support, and auditable agent sessions expand governed infrastructure-agent operation.

CRITERIA TOUCHED

C01 / C03 / C05 / C06 / C07 / C11 MATERIAL REVALIDATION

MANDATORY GATES / CURRENT REVALIDATION POSTURE

Mandatory gates remain non-compensable

No current release is marked PASS solely from documentation. E2 means the official source supports the capability path; LAB and RETEST identify proof still required. HOLD blocks a current decision.

MANDATORY GATE	LANGGRAPH	AMAZON BEDROCK	ITENTIAL AUTOMA
C01 Functional coverage	RETEST	RETEST	RETEST
C02 Security / trust	E2	E2	E2
C03 Governance / approval	E2	RETEST	RETEST
C04 State integrity	RETEST	E2	E2
C07 Observability / evidence	RETEST	RETEST	RETEST
C09 Resilience / continuity	LAB	LAB	LAB
C11 Change safety / rollback	RETEST	RETEST	RETEST
C16 Portability / exit	LAB	LAB	LAB
C18 Assurance confidence	HOLD	HOLD	HOLD

LEGEND

E2

official-source support only

LAB

controlled proof required

RETEST

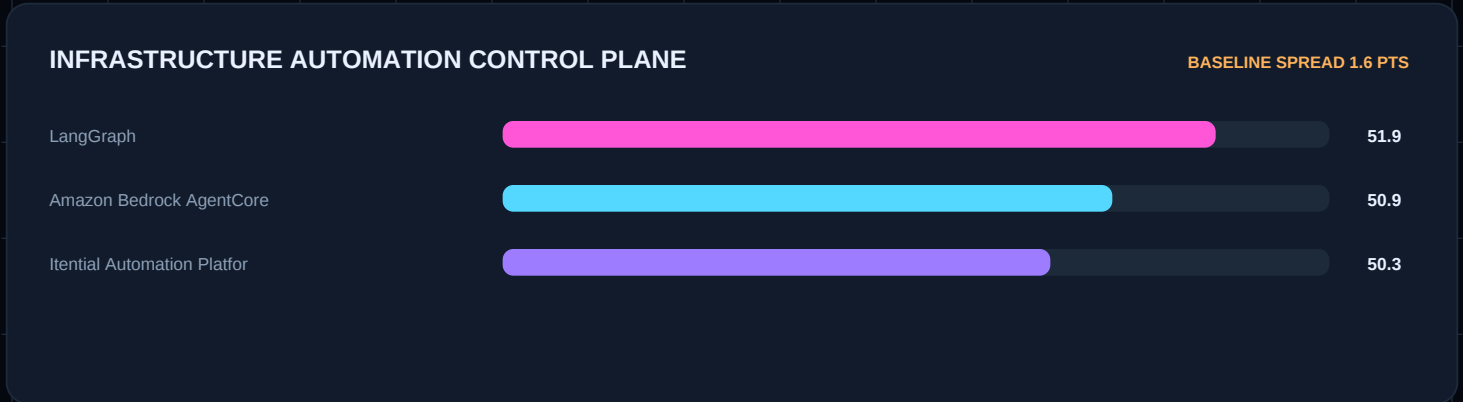
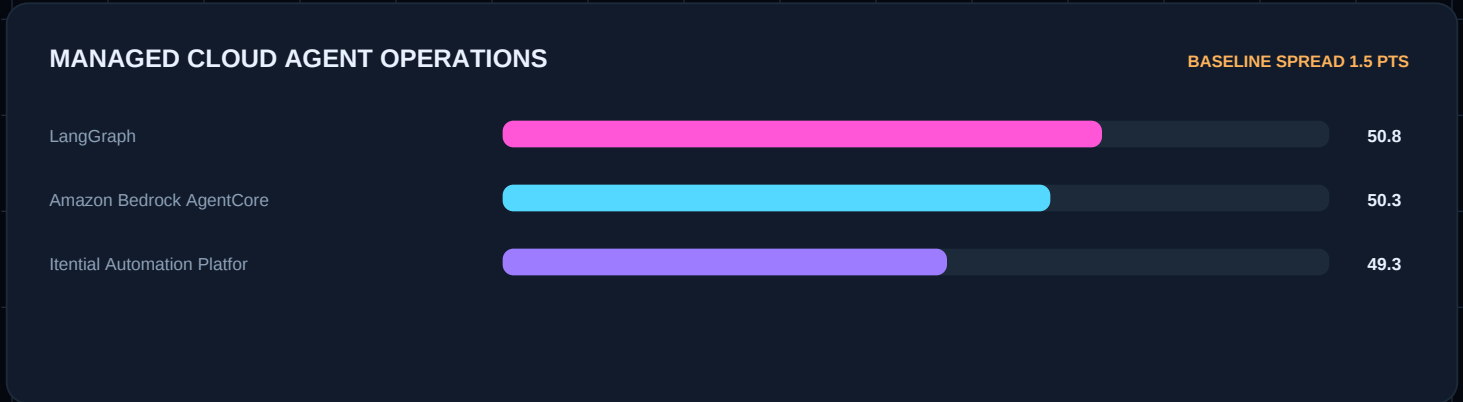
release delta touches this gate

HOLD

decision gate remains closed

Historical profile sensitivity - preserved, not reissued

Values below are the 2026-07-24 evidence-adjusted baseline. They show sensitivity to operating-model weights. The September refresh does not recompute rank because E3 lab proof has not been reproduced.



CURRENT RANK STATUS: WITHHELD PENDING CONTROLLED PROOF.

Evidence confidence is separate from capability

E5

Independent repeatable validation

confidence modifier 1.00

E4

Production evidence in adopting environment

confidence modifier 0.91

E3

Controlled Mythos laboratory result

confidence modifier 0.78

E2

Official documentation / release / source

confidence modifier 0.64

E1

Vendor assertion or marketing

confidence modifier 0.38

E0

Unsupported or unverified

confidence modifier 0.00

CURRENT EVIDENCE STATE

LangGraph

E2 REFRESHED

Historical adjusted confidence: 59.7%

E3 LAB: NOT REPRODUCED

Amazon Bedrock AgentCore

E2 REFRESHED

Historical adjusted confidence: 58.2%

E3 LAB: NOT REPRODUCED

Itential Automation Platform

E2 REFRESHED

Historical adjusted confidence: 58.2%

E3 LAB: NOT REPRODUCED

ASSURANCE RULE

A current release note can change capability context, but it cannot raise assurance beyond E2. Current recommendation strength is therefore bounded until the test plan produces reproducible artifacts.

MIGRATION, LIFECYCLE AND IRREVERSIBLE-COMMITMENT RISK

Lifecycle risk is evaluated independently of features

The benchmark models migration, upgrade, compatibility, support, staffing, data/state export, downtime and exit. Current release changes below are explicit proof triggers.

LangGraph

MODERATE REVALIDATION

Durable graph runtime and checkpoint ecosystem continue to evolve across Python and JavaScript lines.

RETEST: C01 / C04 / C07 / C11 / C12 / C14

Amazon Bedrock AgentCore

MATERIAL REVALIDATION

Batch Evaluations and A/B Testing are now GA, materially strengthening the observe-evaluate-improve proof loop.

RETEST: C01 / C03 / C07 / C08 / C11 / C18

Itential Automation Platform

MATERIAL REVALIDATION

FlowAI, Work Center, asynchronous workflow support, and auditable agent sessions expand governed infrastructure-agent operation.

RETEST: C01 / C03 / C05 / C06 / C07 / C11

IRREVERSIBLE COMMITMENT GATE

Before migration or procurement lock-in, prove rollback, state portability, operational reconstruction, and supplier-exit assumptions.

Controlled proof is the next decision-producing step

REPRODUCTION STATE / NOT EXECUTED IN THIS EVIDENCE-REFRESH RELEASE

01

CONTROLLED SCENARIO

Run a long-lived mission with branching, retry, interruption, approval, restart, replay, and compensation.

02

CONTROLLED SCENARIO

Integrate local and frontier models, MCP tools, secrets, identity, and policy.

03

CONTROLLED SCENARIO

Inject model timeout, malformed tool output, duplicate event, lost checkpoint, and operator cancellation.

04

CONTROLLED SCENARIO

Measure task throughput, state growth, token use, tool latency, and evidence completeness.

05

CONTROLLED SCENARIO

Export mission state, traces, memory, policies, and tool definitions.

06

CONTROLLED SCENARIO

Red-team prompt injection, confused-deputy behavior, excess authority, and unsafe recovery.

EVIDENCE PACKAGE

Preserve exact versions, architecture, configuration, scenario, expected result, instrumentation, raw output, timing, failures, operator actions, deviations, acceptance state, owner, and artifact hashes.

Current decision posture

Reopen the comparison. AgentCore and Itential changed materially after the July baseline while LangGraph continued to evolve. No inherited ranking should be promoted to a current recommendation until durable execution, authority, observability, replay, migration, and failure tests are rerun.

CURRENT RANKING: NOT REISSUED / CONTROLLED LAB REQUIRED

CURRENT OFFICIAL-SOURCE REGISTER / CHECKED 2026-09-17

LangGraph	E2 / OFFICIAL	
LangGraph GitHub releases		github.com

Amazon Bedrock AgentCore	E2 / OFFICIAL	
Amazon Bedrock AgentCore release notes		docs.aws.amazon.com

Itential Automation Platform	E2 / OFFICIAL	
Itential Platform 6.5 release notes		docs.itential.com

REVISION LINEAGE

1.2.0-candidate / 2026-09-17 - current-source evidence refresh; no lab rescore issued.

1.1.0-candidate / 2026-07-24 - baseline benchmark using the final comparative evaluation system.

Trigger earlier review after major release, acquisition, licensing change, critical vulnerability, material outage, failed PoC, or workload change.

BASELINE ANALYTIC RECORD CONTINUES ON PAGE 11 FOR TRACEABILITY.

ATOMIC CRITERIA MODEL

WEIGHTED CAPABILITY WITH EXPLICIT MINIMUM EVIDENCE

ID	CRITERION	WEIGHT	GATE	MINIMUM EVIDENCE
C01	Agent orchestration, execution, memory, tool, and human-control coverage	5	YES	Feature documentation, APIs, configuration exports, and scenario demonstration.
C02	Agent identity, tool authority, isolation, and policy architecture	5	YES	Threat model, identity flows, RBAC tests, audit records, and security configuration.
C03	Governance and approval controls	5	YES	Approval workflow, policy controls, separation-of-duties tests, and decision records.
C04	Durable mission, state, checkpoint, and memory integrity	5	YES	Backup, restore, rollback, drift, and corruption-recovery evidence.
C05	Automation and API quality	4	NO	API specifications, authentication tests, rate limits, event samples, and automation runs.
C06	Integration ecosystem	4	NO	Supported integrations, connector tests, failure behavior, and lifecycle ownership.
C07	Observability and evidence	4	YES	Logs, traces, metrics, reports, export tests, and evidence-retention controls.
C08	Agent, task, model, tool, and concurrency scale	4	NO	Controlled load tests, topology scale, saturation behavior, and capacity model.
C09	Resilience and continuity	5	YES	Failure injection, HA tests, recovery timing, degraded-mode operation, and runbooks.
C10	Usability and operator cognition	3	NO	Task observation, error-rate measures, navigation tests, and operator interviews.
C11	Interrupt, approval, replay, compensation, and rollback safety	5	YES	Plan or preview output, canary tests, rollback execution, and post-change verification.
C12	Extensibility and programmability	3	NO	Plugin, module, provider, workflow, or code-extension tests and upgrade impact analysis.
C13	Deployment and sovereignty options	3	NO	Deployment architecture, data-flow mapping, residency evidence, and offline tests.
C14	Lifecycle and upgrade discipline	4	NO	Release notes, upgrade test, compatibility matrix, support policy, and migration plan.
C15	Economics and cost predictability	3	NO	Bill-of-materials, licensing model, staffing estimates, metering, and sensitivity analysis.
C16	Portability and exit	4	YES	Export tests, format analysis, replacement workflow, and decommission evidence.
C17	Vendor and ecosystem risk	3	NO	License analysis, roadmap evidence, bus-factor indicators, and dependency inventory.
C18	Assurance confidence	5	YES	Source provenance, lab artifacts, production evidence, independent replication, and review date.

SCORE SCALE

0 absent; 1 materially deficient; 2 partial; 3 adequate with limits; 4 strong; 5 leading for the defined profile. Scores remain provisional until the required evidence grade is achieved.

RAW CAPABILITY SCORECARD

DOCUMENTED CAPABILITY BEFORE EVIDENCE DISCOUNT

CRITERION	LANGGRAPH	AGENTCORE	ITENTIAL
C01 Agent orchestration, execution, memory, tool, and human-control coverage	4.5	4.7	4.0
C02 Agent identity, tool authority, isolation, and policy architecture	4.0	4.6	4.4
C03 Governance and approval controls	4.1	4.5	4.7
C04 Durable mission, state, checkpoint, and memory integrity	4.6	4.4	4.2
C05 Automation and API quality	4.5	4.6	4.5
C06 Integration ecosystem	4.3	4.6	4.8
C07 Observability and evidence	4.4	4.7	4.4
C08 Agent, task, model, tool, and concurrency scale	4.2	4.7	4.2
C09 Resilience and continuity	4.1	4.5	4.3
C10 Usability and operator cognition	4.0	4.4	4.3
C11 Interrupt, approval, replay, compensation, and rollback safety	4.8	4.4	4.7
C12 Extensibility and programmability	4.8	4.2	4.2
C13 Deployment and sovereignty options	4.5	3.2	3.8
C14 Lifecycle and upgrade discipline	4.1	4.2	4.1
C15 Economics and cost predictability	4.4	3.5	3.4
C16 Portability and exit	4.7	3.3	3.5
C17 Vendor and ecosystem risk	4.2	3.4	3.6
C18 Assurance confidence	3.5	3.4	3.4

WEIGHTED BASELINE / 100

LangGraph

86.1

AgentCore

84.6

Itential

83.4

EVIDENCE-ADJUSTED SCORECARD

CAPABILITY DISCOUNTED BY CURRENT EVIDENCE CONFIDENCE

CANDIDATE	RAW	EVIDENCE CONFIDENCE	ADJUSTED	CURRENT STATE
LangGraph	86.1	59.7%	51.5	CONTROLLED LAB PENDING
AgentCore	84.6	58.2%	49.8	CONTROLLED LAB PENDING
Itential	83.4	58.2%	49.2	CONTROLLED LAB PENDING

EVIDENCE SCALE

E0 / 0.00

Unsupported or unverified

E1 / 0.38

Vendor assertion or marketing statement

E2 / 0.64

Official documentation, release notes, or source repository

E3 / 0.78

Controlled Mythos laboratory result

E4 / 0.91

Production evidence from the adopting environment

E5 / 1.00

Independent repeatable validation

CURRENT CONFIDENCE BOUNDARY

Most candidate criteria are supported at E2: official documentation or source evidence. Environment-specific laboratory, production, and independent evidence remain future gates.

PROFILE-SPECIFIC RANKINGS

EVIDENCE-ADJUSTED SENSITIVITY ANALYSIS

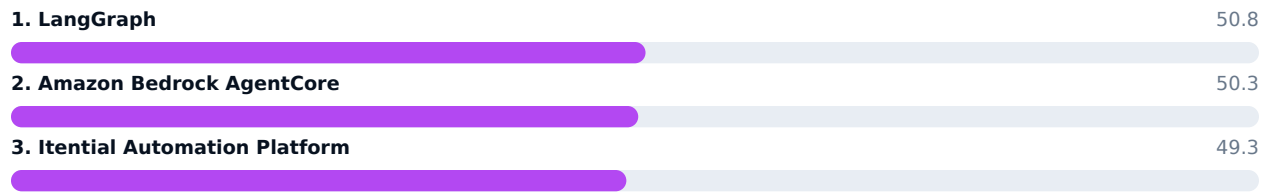
CUSTOM GOVERNED AGENT RUNTIME

Prioritizes deterministic graphs, durable execution, interrupts, model independence, portability, and deep engineering control.



MANAGED CLOUD AGENT OPERATIONS

Prioritizes managed runtime, identity, gateway, memory, observability, scale, and reduced platform burden.



INFRASTRUCTURE AUTOMATION CONTROL PLANE

Prioritizes approvals, integration, deterministic execution, infrastructure workflows, audit, and enterprise operations.



RANK SENSITIVITY

Small score differences are not decision certainty. Treat near-ties as a requirement for deeper PoC, commercial, migration, and operating-model evidence.

CANDIDATE DEEP DIVE / 01

LANGGRAPH

OPERATING MODEL

Low-level open-source orchestration framework and runtime for long-running, stateful agent graphs.

DOCUMENTED STRENGTHS

- Fine-grained control over graph, state, branching, interrupts, and persistence.
- Open-source and model-provider flexibility.
- Strong fit for custom deterministic agent runtimes.
- Developer-centric composability and testability.

CAUTIONS AND PROOF OBLIGATIONS

- Production platform services must be assembled or adopted separately.
- Security, tenancy, identity, policy, and operations remain architectural responsibilities.
- Low-level flexibility can become complexity without strong engineering standards.

BASELINE SCORE PROFILE



GATE DISPOSITION

PASS - custom governed agent runtime | CONDITIONAL - production platform services must be engineered

CANDIDATE DEEP DIVE / 02

AMAZON BEDROCK AGENTCORE

OPERATING MODEL

Managed framework-agnostic agent platform providing runtime, gateway, memory, identity, policy, observability, and operational services.

DOCUMENTED STRENGTHS

- Managed runtime and scaling.
- Integrated gateway, memory, identity, policy, and observability services.
- Framework and model flexibility within the AWS operating boundary.
- Reduced infrastructure burden for production agent operations.

CAUTIONS AND PROOF OBLIGATIONS

- AWS dependency and service maturity must be accepted.
- Portability, data boundaries, and cost require controlled validation.
- Managed abstractions can constrain low-level control and debugging.

BASELINE SCORE PROFILE



GATE DISPOSITION

PASS - managed cloud agent operations | CONDITIONAL - AWS boundary, portability, and cost

CANDIDATE DEEP DIVE / 03

INTENTIAL AUTOMATION PLATFORM

OPERATING MODEL

Enterprise network and infrastructure orchestration platform with integration, workflow, governance, and operational automation capabilities.

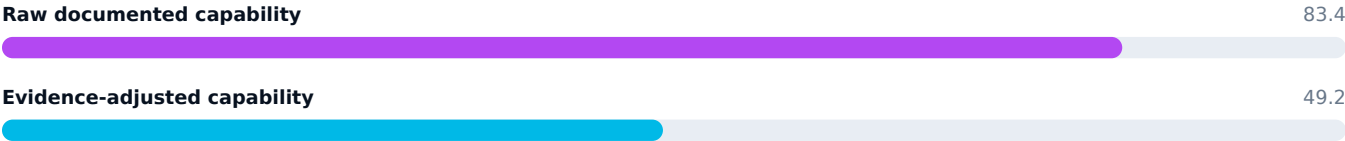
DOCUMENTED STRENGTHS

- Strong network and infrastructure orchestration orientation.
- Enterprise workflow, integration, approval, and operational controls.
- Useful bridge between intent, existing automation, controllers, and ITSM.
- Fits environments where deterministic infrastructure execution is primary.

CAUTIONS AND PROOF OBLIGATIONS

- Not a general-purpose low-level agent runtime.
- Model, memory, reasoning, and multi-agent features require integration or custom architecture.
- Commercial platform coupling and connector lifecycle require governance.

BASELINE SCORE PROFILE



GATE DISPOSITION

PASS - infrastructure automation control-plane profile | NOT EQUIVALENT - general-purpose agent-runtime comparison

CROSS-CANDIDATE CONTRASTS

WHERE ARCHITECTURAL DIFFERENCES MATTER MOST

CONTROL-LOOP MODEL

How authority moves from intent to plan, approval, execution, observation, reconciliation, and recovery.

STATE OWNERSHIP

Which system owns desired state, operational state, evidence, history, and conflict resolution.

MANAGED VERSUS SOVEREIGN BOUNDARY

Which runtime, identity, data, policy, relay, or service dependencies remain outside organizational control.

EXTENSION MODEL

Plugins, providers, modules, applications, workflows, SDKs, APIs, and custom code introduce different lifecycle risks.

FAILURE TRANSPARENCY

Candidates differ in how clearly they expose partial execution, degraded mode, retries, inconsistency, and recovery state.

EXIT MECHANICS

Export formats, state semantics, policy portability, knowledge transfer, and replacement effort are materially different.

CANDIDATE	CURRENT ADVANCEMENT POSITION
LangGraph	PASS - custom governed agent runtime
AgentCore	PASS - managed cloud agent operations
Itential	PASS - infrastructure automation control-plane profile

ARCHITECTURE AND INTEGRATION FIT

THE PRODUCT IS ONLY ONE PART OF THE OPERATING SYSTEM

REFERENCE OPERATING LAYERS

01 Intent, task, and mission model

02 Supervisor, graph, workflow, and durable state

03 Model, memory, context, and knowledge services

04 Tool gateway, identity, policy, approvals, and execution

05 Observation, evaluation, evidence, replay, and recovery

INTEGRATION BOUNDARIES

- Model providers and local runtimes
- MCP, APIs, CLIs, infrastructure controllers, and automation tools
- Identity, secrets, policy, and approval systems
- Vector, graph, object, event, and relational data stores
- Observability, evaluation, incident, and evidence platforms

ARCHITECTURE GATE

Every external dependency requires an owner, identity boundary, failure mode, evidence path, version policy, recovery procedure, and exit strategy.

SECURITY, OPERATIONS, LIFECYCLE, ECONOMICS, AND EXIT

CROSS-DOMAIN DECISION RECORD

SECURITY AND AUTHORITY

Identity, credentials, secrets, roles, privileged actions, signing, approvals, isolation, audit, and incident response must be tested as an integrated system.

RELIABILITY AND RECOVERY

Control-plane, data-plane, dependency, region, database, queue, network, and operator failures require defined degraded modes and recovery objectives.

CHANGE AND LIFECYCLE

Version, compatibility, migration, canary, rollback, content, plugin, provider, firmware, and decommission procedures are part of product fitness.

ECONOMICS AND CAPACITY

Licenses, metering, data, compute, hardware, storage, support, staffing, training, integration, migration, and opportunity cost require sensitivity analysis.

SOVEREIGNTY AND PRIVACY

Data location, support access, telemetry, subprocessors, encryption, key control, disconnected operation, and legal obligations must align with policy.

PORTABILITY AND EXIT

Configuration, policy, data, state, evidence, workflows, identities, and operational knowledge must be exportable and reconstructable.

DECISION OWNERSHIP

Architecture, security, operations, finance, procurement, legal, data, and service owners jointly accept the final profile, evidence, residual risk, and exit obligations.

RISK AND FAILURE REGISTER

SCENARIOS THAT CAN INVALIDATE A FEATURE-BASED SELECTION

ID	SEVERITY	FAILURE SCENARIO	REQUIRED TREATMENT
R-01	CRITICAL	A flexible framework is mistaken for a complete production platform.	PREVENT / DETECT / RECOVER / EVIDENCE
R-02	HIGH	A managed platform is accepted without testing portability, cost, data boundaries, and failure transparency.	PREVENT / DETECT / RECOVER / EVIDENCE
R-03	HIGH	Agent authority exceeds tool, data, or environment permissions.	PREVENT / DETECT / RECOVER / EVIDENCE
R-04	HIGH	State replay duplicates irreversible actions.	PREVENT / DETECT / RECOVER / EVIDENCE
R-05	MEDIUM	Human approval is cosmetic because plans, diffs, and consequences are not understandable.	PREVENT / DETECT / RECOVER / EVIDENCE
R-06	MEDIUM	Model quality is used to compensate for weak deterministic execution and verification.	PREVENT / DETECT / RECOVER / EVIDENCE

RISK RULE

A candidate with an unresolved critical failure path cannot advance because optional capability, market position, or commercial discount cannot compensate for missing safety or recovery evidence.

CONTROLLED PROOF-OF-CAPABILITY PLAN

REPEATABLE SCENARIOS, INSTRUMENTATION, AND ACCEPTANCE

TEST 01

Run a long-lived mission with branching, retry, interruption, approval, restart, replay, and compensation.

TEST 02

Integrate local and frontier models, MCP tools, secrets, identity, and policy.

TEST 03

Inject model timeout, malformed tool output, duplicate event, lost checkpoint, and operator cancellation.

TEST 04

Measure task throughput, state growth, token use, tool latency, and evidence completeness.

TEST 05

Export mission state, traces, memory, policies, and tool definitions.

TEST 06

Red-team prompt injection, confused-deputy behavior, excess authority, and unsafe recovery.

EVIDENCE PACKAGE

Preserve versions, architecture, configuration, data set, scenario, expected result, instrumentation, raw output, timing, failures, operator actions, deviations, acceptance state, owner, and artifact hashes.

CONDITIONAL RECOMMENDATION AND REVALIDATION

DECISION RECORD, ACCEPTANCE, AND NEXT EVIDENCE

RECOMMENDATION POSITION

- Select the architecture layer before selecting the product.
- Require durable-state, authority, replay, approval, and failure-injection tests.
- Preserve model and tool portability through typed contracts and evidence.
- Revalidate monthly because managed agent platforms and framework capabilities are changing rapidly.

CANDIDATE	ADJ. SCORE	GATE POSITION	LAB STATE
LangGraph	51.5	PASS - custom governed agent runtime; CONDITIONAL - production platform services must be engineered	PENDING
AgentCore	49.8	PASS - managed cloud agent operations; CONDITIONAL - AWS boundary, portability, and cost	PENDING
Itential	49.2	PASS - infrastructure automation control-plane profile; NOT EQUIVALENT - general-purpose agent-runtime comparison	PENDING

REVALIDATION SCHEDULE

Review no later than 2026-10-24. Review earlier after a major release, commercial change, critical vulnerability, material outage, architecture transition, failed acceptance test, or change to the target workload profile.

SOURCE AUTHORITY AND REVISION

CURRENT EVIDENCE SNAPSHOT AND PUBLICATION HISTORY

SOURCE ID	PUBLISHER	TITLE	CHECKED
NIST-AI-RMF	NIST	AI Risk Management Framework	2026-07-24
NIST-SP800-53	NIST	Security and Privacy Controls for Information Systems and Organizations	2026-07-24
LANGGRAPH	LangChain	LangGraph Overview	2026-07-24
LANGGRAPH-GH	LangChain	LangGraph Source Repository	2026-07-24
AGENTCORE	Amazon Web Services	Amazon Bedrock AgentCore Overview	2026-07-24
AGENTCORE-GATEWAY	Amazon Web Services	Amazon Bedrock AgentCore Gateway	2026-07-24
ITENTIAL	Itential	Itential Automation Platform Documentation	2026-07-24

SOURCE POSITION

Official documentation and source repositories support the current capability model. Source records preserve publisher, title, URL, purpose, and review date in the machine-readable authority register. Commercial terms, performance, and environment-specific behavior remain unverified until controlled proof.

REVISION HISTORY

1.1.0-candidate / 2026-07-24 - permanent-publication rebuild using the final benchmark system, profile-specific rankings, evidence adjustment, mandatory gates, and controlled PoC requirements.

CMP-007

IaC and Control Planes

Infrastructure as code, declarative delivery and
platform control planes



CURRENT EVIDENCE SNAPSHOT

Official product sources refreshed through 2026-09-17
Controlled Mythos laboratory rerun: NOT ASSERTED

VERSION 1.2.0-candidate

CANDIDATE COMPARATIVE EVALUATION / PUBLIC

BENCHMARK DOCTRINE

Gates before scores. Evidence before
certainty. Profile fit before universal ranking.

IaC and Control Planes

Infrastructure as code, declarative delivery and platform control planes

DOCUMENT ID

CMP-007

VERSION

1.2.0-candidate

STATUS

Candidate Comparative Evaluation

PUBLICATION DATE

2026-09-17

SCHEDULED REVIEW

2026-12-16

CANONICAL ROUTE

/library/evaluations/cmp-007/

4

CANDIDATES

9

MANDATORY GATES

E2

CURRENT MAX EVIDENCE

18

ATOMIC CRITERIA

3

WORKLOAD PROFILES

CURRENT DECISION BOUNDARY

The July benchmark is a near-tie, not a universal result. Current Terraform, Pulumi, Crossplane, and OpenTofu releases change lifecycle and control-plane assumptions. Reproduce state, policy, drift, rollback, provider failure, migration, and exit under the intended operating model.

NO CURRENT UNIVERSAL WINNER IS PUBLISHED FROM THIS REFRESH.

July 24 baseline scores are preserved as lineage and sensitivity evidence, not silently reissued as September laboratory results.

Terraform / HCP Terraform

Terraform 1.16.3 / 2026-09-16

REFRESH TRIGGER

Stable 1.16.3 is current; 1.17 is beta. HCP Terraform continues policy/Stacks evolution.

OpenTofu

v1.12.0 / 2026-05-14

REFRESH TRIGGER

1.12 adds dynamic prevent_destroy, checksum improvements, output changes, and lifecycle controls; WinRM deprecation affects some estates.

Pulumi

v3.263.0 / 2026-09-15

REFRESH TRIGGER

Current line adds continued policy, Terraform-state, cross-language, and HCL integration improvements.

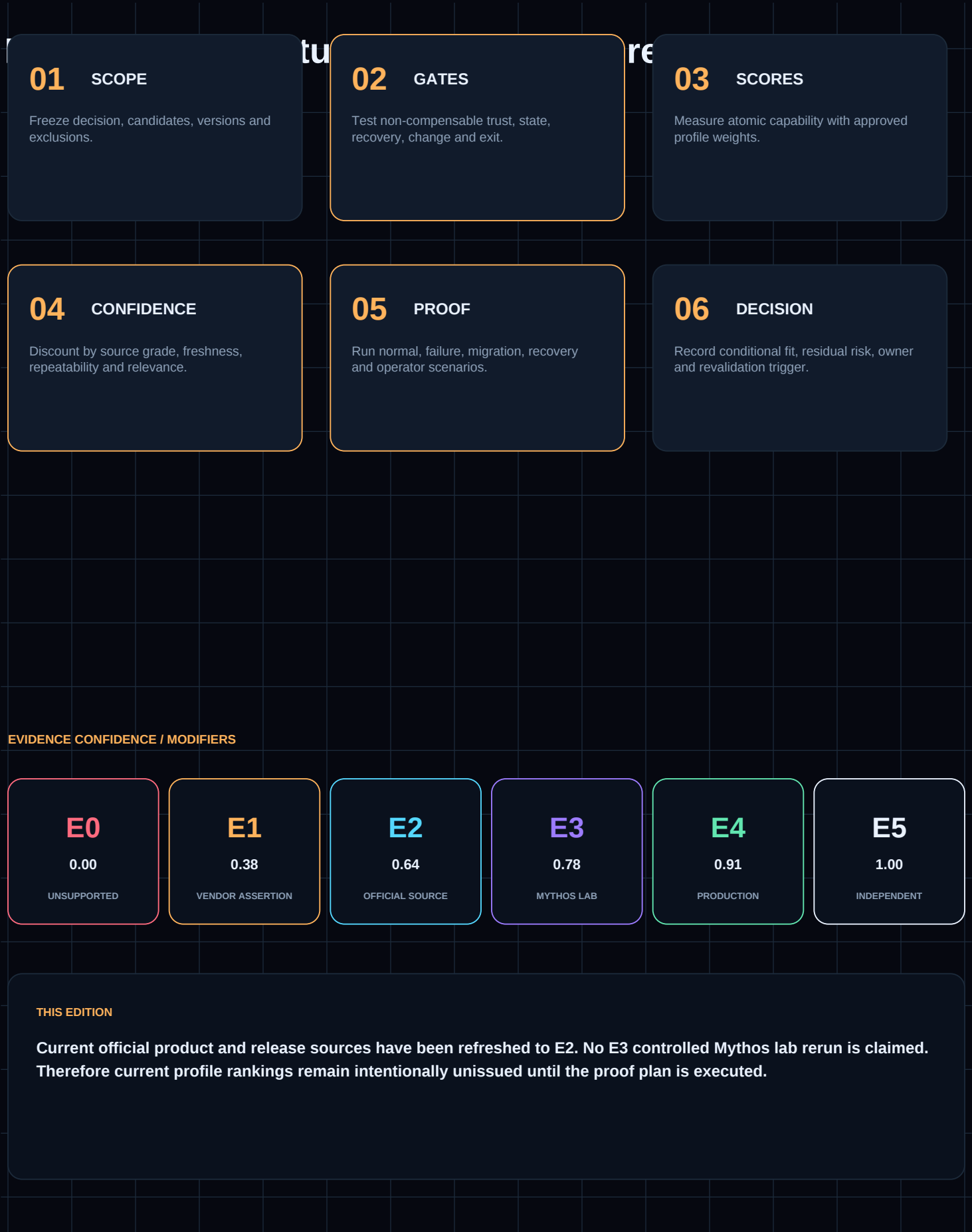
Crossplane

v2.4 / 2026-08-20

REFRESH TRIGGER

v2.4 is a maintained release in the broader v2 control-plane model; upgrade and composition behavior remain material.

FINAL BENCHMARK SYSTEM / DECISION ARCHITECTURE



What changed since the July benchmark snapshot

Changes below are sourced from current official release documentation. They identify revalidation triggers; they are not translated into new numeric scores without controlled proof.

Terraform / HCP Terraform Terraform 1.16.3 / 2026-09-16 MATERIAL DELTA Stable 1.16.3 is current; 1.17 is beta. HCP Terraform continues policy/Stacks evolution. CRITERIA TOUCHED C01 / C03 / C04 / C05 / C11 / C14 MODERATE REVALIDATION	OpenTofu v1.12.0 / 2026-05-14 MATERIAL DELTA 1.12 adds dynamic prevent_destroy, checksum improvements, output changes, and lifecycle controls; WinRM deprecation affects some estates. CRITERIA TOUCHED C04 / C11 / C14 / C16 MODERATE REVALIDATION
Pulumi v3.263.0 / 2026-09-15 MATERIAL DELTA Current line adds continued policy, Terraform-state, cross-language, and HCL integration improvements. CRITERIA TOUCHED C03 / C04 / C05 / C06 / C12 / C14 MODERATE REVALIDATION	Crossplane v2.4 / 2026-08-20 MATERIAL DELTA v2.4 is a maintained release in the broader v2 control-plane model; upgrade and composition behavior remain material. CRITERIA TOUCHED C01 / C04 / C08 / C09 / C11 / C14 HIGH REVALIDATION

MANDATORY GATES / CURRENT REVALIDATION POSTURE

Mandatory gates remain non-compensable

No current release is marked PASS solely from documentation. E2 means the official source supports the capability path; LAB and RETEST identify proof still required. HOLD blocks a current decision.

MANDATORY GATE	TERRAFORM / HCP	OPENTOFU	PULUMI	CROSSPLANE
C01 Functional coverage	RETEST	E2	E2	RETEST
C02 Security / trust	E2	E2	E2	E2
C03 Governance / approval	RETEST	E2	RETEST	E2
C04 State integrity	RETEST	RETEST	RETEST	RETEST
C07 Observability / evidence	E2	E2	E2	E2
C09 Resilience / continuity	LAB	LAB	LAB	RETEST
C11 Change safety / rollback	RETEST	RETEST	LAB	RETEST
C16 Portability / exit	LAB	RETEST	LAB	LAB
C18 Assurance confidence	HOLD	HOLD	HOLD	HOLD

LEGEND

E2

official-source support only

LAB

controlled proof required

RETEST

release delta touches this gate

HOLD

decision gate remains closed

Historical profile sensitivity - preserved, not reissued

Values below are the 2026-07-24 evidence-adjusted baseline. They show sensitivity to operating-model weights. The September refresh does not recompute rank because E3 lab proof has not been reproduced.



CURRENT RANK STATUS: WITHHELD PENDING CONTROLLED PROOF.

Evidence confidence is separate from capability

E5

Independent repeatable validation

confidence modifier 1.00

E4

Production evidence in adopting environment

confidence modifier 0.91

E3

Controlled Mythos laboratory result

confidence modifier 0.78

E2

Official documentation / release / source

confidence modifier 0.64

E1

Vendor assertion or marketing

confidence modifier 0.38

E0

Unsupported or unverified

confidence modifier 0.00

CURRENT EVIDENCE STATE

Terraform / HCP Terraform

E2 REFRESHED

Historical adjusted confidence: 58.2%

E3 LAB: NOT REPRODUCED

OpenTofu

E2 REFRESHED

Historical adjusted confidence: 58.2%

E3 LAB: NOT REPRODUCED

Pulumi

E2 REFRESHED

Historical adjusted confidence: 58.2%

E3 LAB: NOT REPRODUCED

Crossplane

E2 REFRESHED

Historical adjusted confidence: 58.2%

E3 LAB: NOT REPRODUCED

ASSURANCE RULE

A current release note can change capability context, but it cannot raise assurance beyond E2. Current recommendation strength is therefore bounded until the test plan produces reproducible artifacts.

MIGRATION, LIFECYCLE AND IRREVERSIBLE-COMMITMENT RISK

Lifecycle risk is evaluated independently of features

The benchmark models migration, upgrade, compatibility, support, staffing, data/state export, downtime and exit. Current release changes below are explicit proof triggers.

Terraform / HCP Terraform MODERATE REVALIDATION Stable 1.16.3 is current; 1.17 is beta. HCP Terraform continues policy/Stacks evolution. RETEST: C01 / C03 / C04 / C05 / C11 / C14
OpenTofu MODERATE REVALIDATION 1.12 adds dynamic prevent_destroy, checksum improvements, output changes, and lifecycle controls; WinRM deprecation affects some estates. RETEST: C04 / C11 / C14 / C16
Pulumi MODERATE REVALIDATION Current line adds continued policy, Terraform-state, cross-language, and HCL integration improvements. RETEST: C03 / C04 / C05 / C06 / C12 / C14
Crossplane HIGH REVALIDATION v2.4 is a maintained release in the broader v2 control-plane model; upgrade and composition behavior remain material. RETEST: C01 / C04 / C08 / C09 / C11 / C14

IRREVERSIBLE COMMITMENT GATE

Before migration or procurement lock-in, prove rollback, state portability, operational reconstruction, and supplier-exit assumptions.

Controlled proof is the next decision-producing step

REPRODUCTION STATE / NOT EXECUTED IN THIS EVIDENCE-REFRESH RELEASE

01

CONTROLLED SCENARIO

Provision representative cloud, network, identity, data, and SaaS resources with dependencies.

02

CONTROLLED SCENARIO

Exercise import, drift, replacement, partial failure, retry, refresh, taint or repair, and state recovery.

03

CONTROLLED SCENARIO

Validate policy, approval, secret isolation, provider signing, and execution-environment controls.

04

CONTROLLED SCENARIO

Measure plan, apply, reconcile, provider, and state performance at representative scale.

05

CONTROLLED SCENARIO

Upgrade providers, modules, languages, compositions, and the control plane.

06

CONTROLLED SCENARIO

Migrate a controlled stack between approaches and verify data, state, policy, and rollback.

EVIDENCE PACKAGE

Preserve exact versions, architecture, configuration, scenario, expected result, instrumentation, raw output, timing, failures, operator actions, deviations, acceptance state, owner, and artifact hashes.

Current decision posture

The July benchmark is a near-tie, not a universal result. Current Terraform, Pulumi, Crossplane, and OpenTofu releases change lifecycle and control-plane assumptions. Reproduce state, policy, drift, rollback, provider failure, migration, and exit under the intended operating model.

CURRENT RANKING: NOT REISSUED / CONTROLLED LAB REQUIRED

CURRENT OFFICIAL-SOURCE REGISTER / CHECKED 2026-09-17

Terraform / HCP Terraform Terraform releases github.com	E2 / OFFICIAL
OpenTofu OpenTofu 1.12.0 release opentofu.org	E2 / OFFICIAL
Pulumi Pulumi versions and changelog www.pulumi.com	E2 / OFFICIAL
Crossplane Crossplane releases github.com	E2 / OFFICIAL

REVISION LINEAGE

- 1.2.0-candidate / 2026-09-17 - current-source evidence refresh; no lab rescore issued.
- 1.1.0-candidate / 2026-07-24 - baseline benchmark using the final comparative evaluation system.
- Trigger earlier review after major release, acquisition, licensing change, critical vulnerability, material outage, failed PoC, or workload change.
- BASELINE ANALYTIC RECORD CONTINUES ON PAGE 11 FOR TRACEABILITY.

ATOMIC CRITERIA MODEL

WEIGHTED CAPABILITY WITH EXPLICIT MINIMUM EVIDENCE

ID	CRITERION	WEIGHT	GATE	MINIMUM EVIDENCE
C01	Infrastructure lifecycle and control-plane coverage	5	YES	Feature documentation, APIs, configuration exports, and scenario demonstration.
C02	Identity, policy, secret, provider, and execution trust architecture	5	YES	Threat model, identity flows, RBAC tests, audit records, and security configuration.
C03	Governance and approval controls	5	YES	Approval workflow, policy controls, separation-of-duties tests, and decision records.
C04	State, plan, dependency, and reconciliation integrity	5	YES	Backup, restore, rollback, drift, and corruption-recovery evidence.
C05	Automation and API quality	4	NO	API specifications, authentication tests, rate limits, event samples, and automation runs.
C06	Integration ecosystem	4	NO	Supported integrations, connector tests, failure behavior, and lifecycle ownership.
C07	Observability and evidence	4	YES	Logs, traces, metrics, reports, export tests, and evidence-retention controls.
C08	Resource, stack, workspace, provider, and control-loop scale	4	NO	Controlled load tests, topology scale, saturation behavior, and capacity model.
C09	Resilience and continuity	5	YES	Failure injection, HA tests, recovery timing, degraded-mode operation, and runbooks.
C10	Usability and operator cognition	3	NO	Task observation, error-rate measures, navigation tests, and operator interviews.
C11	Plan, approval, policy, canary, rollback, and reconciliation safety	5	YES	Plan or preview output, canary tests, rollback execution, and post-change verification.
C12	Extensibility and programmability	3	NO	Plugin, module, provider, workflow, or code-extension tests and upgrade impact analysis.
C13	Deployment and sovereignty options	3	NO	Deployment architecture, data-flow mapping, residency evidence, and offline tests.
C14	Lifecycle and upgrade discipline	4	NO	Release notes, upgrade test, compatibility matrix, support policy, and migration plan.
C15	Economics and cost predictability	3	NO	Bill-of-materials, licensing model, staffing estimates, metering, and sensitivity analysis.
C16	Portability and exit	4	YES	Export tests, format analysis, replacement workflow, and decommission evidence.
C17	Vendor and ecosystem risk	3	NO	License analysis, roadmap evidence, bus-factor indicators, and dependency inventory.
C18	Assurance confidence	5	YES	Source provenance, lab artifacts, production evidence, independent replication, and review date.

SCORE SCALE

0 absent; 1 materially deficient; 2 partial; 3 adequate with limits; 4 strong; 5 leading for the defined profile. Scores remain provisional until the required evidence grade is achieved.

RAW CAPABILITY SCORECARD

DOCUMENTED CAPABILITY BEFORE EVIDENCE DISCOUNT

CRITERION	TERRAFORM	OPENTOFU	PULUMI	CROSSPLANE
C01 Infrastructure lifecycle and control-plane coverage	4.8	4.6	4.7	4.5
C02 Identity, policy, secret, provider, and execution trust architecture	4.3	4.3	4.3	4.4
C03 Governance and approval controls	4.4	4.1	4.2	4.3
C04 State, plan, dependency, and reconciliation integrity	4.5	4.6	4.4	4.3
C05 Automation and API quality	4.7	4.6	4.7	4.5
C06 Integration ecosystem	4.7	4.5	4.6	4.5
C07 Observability and evidence	4.4	4.2	4.4	4.4
C08 Resource, stack, workspace, provider, and control-loop scale	4.7	4.5	4.5	4.5
C09 Resilience and continuity	4.3	4.1	4.2	4.3
C10 Usability and operator cognition	4.4	4.2	4.3	4.0
C11 Plan, approval, policy, canary, rollback, and reconciliation safety	4.6	4.5	4.5	4.2
C12 Extensibility and programmability	4.6	4.4	4.8	4.8
C13 Deployment and sovereignty options	4.0	4.7	4.5	4.4
C14 Lifecycle and upgrade discipline	4.4	4.1	4.2	4.0
C15 Economics and cost predictability	3.7	4.3	3.8	3.7
C16 Portability and exit	4.1	4.6	4.2	4.0
C17 Vendor and ecosystem risk	3.7	4.0	3.7	3.7
C18 Assurance confidence	3.6	3.5	3.5	3.4

WEIGHTED BASELINE / 100

Terraform

86.9

OpenTofu

86.3

Pulumi

86.1

Crossplane

84.4

EVIDENCE-ADJUSTED SCORECARD

CAPABILITY DISCOUNTED BY CURRENT EVIDENCE CONFIDENCE

CANDIDATE	RAW	EVIDENCE CONFIDENCE	ADJUSTED	CURRENT STATE
Terraform	86.9	58.2%	51.3	CONTROLLED LAB PENDING
OpenTofu	86.3	58.2%	50.8	CONTROLLED LAB PENDING
Pulumi	86.1	58.2%	50.8	CONTROLLED LAB PENDING
Crossplane	84.4	58.2%	49.7	CONTROLLED LAB PENDING

EVIDENCE SCALE

E0 / 0.00

Unsupported or unverified

E1 / 0.38

Vendor assertion or marketing statement

E2 / 0.64

Official documentation, release notes, or source repository

E3 / 0.78

Controlled Mythos laboratory result

E4 / 0.91

Production evidence from the adopting environment

E5 / 1.00

Independent repeatable validation

CURRENT CONFIDENCE BOUNDARY

Most candidate criteria are supported at E2: official documentation or source evidence. Environment-specific laboratory, production, and independent evidence remain future gates.

PROFILE-SPECIFIC RANKINGS

EVIDENCE-ADJUSTED SENSITIVITY ANALYSIS

ENTERPRISE DECLARATIVE DELIVERY

Prioritizes broad provider coverage, plans, approvals, policy, collaboration, audit, and established operating patterns.

1. Terraform and HCP Terraform	51.9
2. Pulumi	51.4
3. OpenTofu	51.3
4. Crossplane	50.2

OPEN SOVEREIGN IAC

Prioritizes open governance, compatibility, state control, encryption, exportability, and self-managed operation.

1. Terraform and HCP Terraform	51.4
2. OpenTofu	51.3
3. Pulumi	51.1
4. Crossplane	50.0

INTERNAL DEVELOPER PLATFORM CONTROL PLANE

Prioritizes reusable platform APIs, continuous reconciliation, self-service, policy, Kubernetes integration, and lifecycle.

1. Terraform and HCP Terraform	51.7
2. Pulumi	51.2
3. OpenTofu	51.0
4. Crossplane	50.1

RANK SENSITIVITY

Small score differences are not decision certainty. Treat near-ties as a requirement for deeper PoC, commercial, migration, and operating-model evidence.

CANDIDATE DEEP DIVE / 01

TERRAFORM AND HCP TERRAFORM

OPERATING MODEL

Declarative infrastructure provisioning using HCL with local, remote, and managed workflow options.

DOCUMENTED STRENGTHS

- Large provider and module ecosystem.
- Widely understood declarative workflow and mature operational patterns.
- Strong plan-and-apply model with managed collaboration options.
- Broad cloud, network, SaaS, and on-premises coverage.

CAUTIONS AND PROOF OBLIGATIONS

- State and provider trust require strict protection.
- Licensing and commercial-platform boundaries must be accepted.
- Rollback is not universally transactional and must be engineered by resource type.

BASELINE SCORE PROFILE

Raw documented capability

86.9

Evidence-adjusted capability

51.3

GATE DISPOSITION

PASS - enterprise declarative delivery | CONDITIONAL - licensing, state, and provider trust

CANDIDATE DEEP DIVE / 02

OPENTOFU

OPERATING MODEL

Community-governed open-source Terraform-compatible infrastructure-as-code engine with state and plan encryption capabilities.

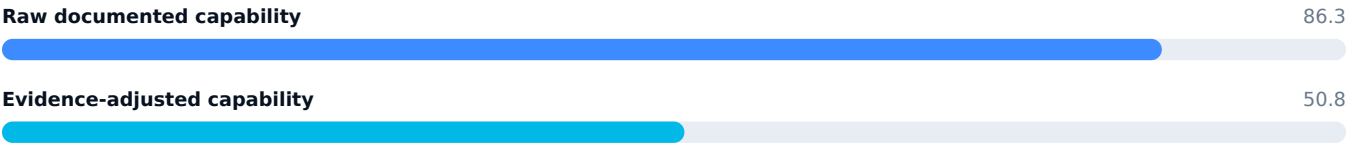
DOCUMENTED STRENGTHS

- Strong compatibility with Terraform workflows and providers.
- Open governance and permissive open-source posture.
- Native state and plan encryption options.
- Lower migration friction for many existing HCL estates.

CAUTIONS AND PROOF OBLIGATIONS

- Commercial ecosystem and enterprise platform depth differ from HCP Terraform.
- Compatibility and provider behavior require release-specific validation.
- Operational tooling, policy, registry, and support model must be selected explicitly.

BASELINE SCORE PROFILE



GATE DISPOSITION

PASS - open sovereign IaC | CONDITIONAL - enterprise workflow and support model

CANDIDATE DEEP DIVE / 03

PULUMI

OPERATING MODEL

Infrastructure as code using general-purpose languages, component resources, managed or DIY state, and policy capabilities.

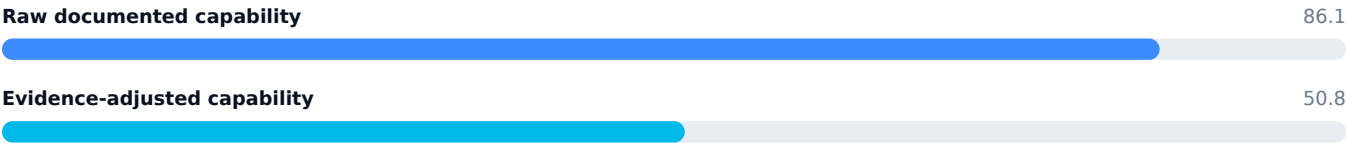
DOCUMENTED STRENGTHS

- General-purpose languages improve abstraction and reuse for software-oriented teams.
- Strong multi-language SDK and component model.
- Managed and self-managed backend choices.
- Rich testing and software-engineering integration potential.

CAUTIONS AND PROOF OBLIGATIONS

- Language flexibility can increase inconsistency and dependency risk.
- State and provider behavior still require infrastructure discipline.
- Team skill, package lifecycle, and code review standards are critical.

BASELINE SCORE PROFILE



GATE DISPOSITION

PASS - software-engineering-centric IaC | CONDITIONAL - language and dependency governance

CANDIDATE DEEP DIVE / 04

CROSSPLANE

OPERATING MODEL

Kubernetes-native control-plane framework using composite resources, providers, compositions, and continuous reconciliation.

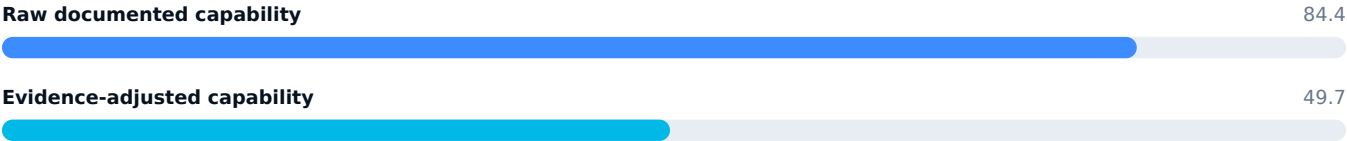
DOCUMENTED STRENGTHS

- Creates organization-specific cloud and infrastructure APIs.
- Continuous reconciliation supports platform control-plane patterns.
- Strong Kubernetes-native integration and composition model.
- Useful for self-service platform engineering and policy-bound abstractions.

CAUTIONS AND PROOF OBLIGATIONS

- Kubernetes becomes a critical control-plane dependency.
- State and failure semantics differ materially from plan-and-apply IaC.
- Provider and composition lifecycle require specialized platform engineering.

BASELINE SCORE PROFILE



GATE DISPOSITION

PASS - internal developer platform control plane | NOT EQUIVALENT - continuous reconciliation changes the operating model

ARCHITECTURE AND INTEGRATION FIT

THE PRODUCT IS ONLY ONE PART OF THE OPERATING SYSTEM

REFERENCE OPERATING LAYERS

01 Intent and platform-product interfaces

02 Configuration or API definitions and policy

03 Plan-and-apply engine or reconciliation control loop

04 Providers, credentials, state, dependency, and execution environment

05 Evidence, drift, recovery, migration, and retirement

INTEGRATION BOUNDARIES

- Cloud and infrastructure provider APIs
- Git, CI/CD, artifact, and module registries
- Identity, secrets, policy, approval, and signing
- Kubernetes, service catalog, developer portal, and platform APIs
- Cost, observability, ITSM, security, and evidence systems

ARCHITECTURE GATE

Every external dependency requires an owner, identity boundary, failure mode, evidence path, version policy, recovery procedure, and exit strategy.

SECURITY, OPERATIONS, LIFECYCLE, ECONOMICS, AND EXIT

CROSS-DOMAIN DECISION RECORD

SECURITY AND AUTHORITY

Identity, credentials, secrets, roles, privileged actions, signing, approvals, isolation, audit, and incident response must be tested as an integrated system.

RELIABILITY AND RECOVERY

Control-plane, data-plane, dependency, region, database, queue, network, and operator failures require defined degraded modes and recovery objectives.

CHANGE AND LIFECYCLE

Version, compatibility, migration, canary, rollback, content, plugin, provider, firmware, and decommission procedures are part of product fitness.

ECONOMICS AND CAPACITY

Licenses, metering, data, compute, hardware, storage, support, staffing, training, integration, migration, and opportunity cost require sensitivity analysis.

SOVEREIGNTY AND PRIVACY

Data location, support access, telemetry, subprocessors, encryption, key control, disconnected operation, and legal obligations must align with policy.

PORTABILITY AND EXIT

Configuration, policy, data, state, evidence, workflows, identities, and operational knowledge must be exportable and reconstructable.

DECISION OWNERSHIP

Architecture, security, operations, finance, procurement, legal, data, and service owners jointly accept the final profile, evidence, residual risk, and exit obligations.

RISK AND FAILURE REGISTER

SCENARIOS THAT CAN INVALIDATE A FEATURE-BASED SELECTION

ID	SEVERITY	FAILURE SCENARIO	REQUIRED TREATMENT
R-01	CRITICAL	A green plan conceals provider-side behavior or destructive replacement.	PREVENT / DETECT / RECOVER / EVIDENCE
R-02	HIGH	State compromise exposes secrets or permits unauthorized control.	PREVENT / DETECT / RECOVER / EVIDENCE
R-03	HIGH	Provider or module supply-chain failure changes behavior across many environments.	PREVENT / DETECT / RECOVER / EVIDENCE
R-04	HIGH	Continuous reconciliation fights emergency or manual recovery actions.	PREVENT / DETECT / RECOVER / EVIDENCE
R-05	MEDIUM	Abstractions hide cost, topology, failure, or compliance consequences from users.	PREVENT / DETECT / RECOVER / EVIDENCE
R-06	MEDIUM	Migration proves syntax compatibility but loses state lineage, policy, workflows, and evidence.	PREVENT / DETECT / RECOVER / EVIDENCE

RISK RULE

A candidate with an unresolved critical failure path cannot advance because optional capability, market position, or commercial discount cannot compensate for missing safety or recovery evidence.

CONTROLLED PROOF-OF-CAPABILITY PLAN

REPEATABLE SCENARIOS, INSTRUMENTATION, AND ACCEPTANCE

TEST 01

Provision representative cloud, network, identity, data, and SaaS resources with dependencies.

TEST 02

Exercise import, drift, replacement, partial failure, retry, refresh, taint or repair, and state recovery.

TEST 03

Validate policy, approval, secret isolation, provider signing, and execution-environment controls.

TEST 04

Measure plan, apply, reconcile, provider, and state performance at representative scale.

TEST 05

Upgrade providers, modules, languages, compositions, and the control plane.

TEST 06

Migrate a controlled stack between approaches and verify data, state, policy, and rollback.

EVIDENCE PACKAGE

Preserve versions, architecture, configuration, data set, scenario, expected result, instrumentation, raw output, timing, failures, operator actions, deviations, acceptance state, owner, and artifact hashes.

CONDITIONAL RECOMMENDATION AND REVALIDATION

DECISION RECORD, ACCEPTANCE, AND NEXT EVIDENCE

RECOMMENDATION POSITION

- Select the control-loop model before comparing feature scores.
- Treat state, provider, secret, and execution trust as mandatory gates.
- Require resource-specific failure, rollback, and migration testing.
- Revalidate providers, licensing, state formats, and control-plane releases quarterly.

CANDIDATE	ADJ. SCORE	GATE POSITION	LAB STATE
Terraform	51.3	PASS - enterprise declarative delivery; CONDITIONAL - licensing, state, and provider trust	PENDING
OpenTofu	50.8	PASS - open sovereign IaC; CONDITIONAL - enterprise workflow and support model	PENDING
Pulumi	50.8	PASS - software-engineering-centric IaC; CONDITIONAL - language and dependency governance	PENDING
Crossplane	49.7	PASS - internal developer platform control plane; NOT EQUIVALENT - continuous reconciliation changes the operating model	PENDING

REVALIDATION SCHEDULE

Review no later than 2026-10-24. Review earlier after a major release, commercial change, critical vulnerability, material outage, architecture transition, failed acceptance test, or change to the target workload profile.

SOURCE AUTHORITY AND REVISION

CURRENT EVIDENCE SNAPSHOT AND PUBLICATION HISTORY

SOURCE ID	PUBLISHER	TITLE	CHECKED
NIST-CSF-20	NIST	Cybersecurity Framework 2.0	2026-07-24
NIST-SP800-53	NIST	Security and Privacy Controls for Information Systems and Organizations	2026-07-24
TERRAFORM	HashiCorp	Terraform Documentation	2026-07-24
TERRAFORM-STATE	HashiCorp	Terraform State Documentation	2026-07-24
OPENTOFU	OpenTofu	OpenTofu Documentation	2026-07-24
OPENTOFU-ENC	OpenTofu	OpenTofu State and Plan Encryption	2026-07-24
PULUMI	Pulumi	Pulumi Infrastructure as Code Documentation	2026-07-24
PULUMI-STATE	Pulumi	Pulumi State and Backends	2026-07-24
CROSSPLANE	Crossplane	Crossplane Compositions	2026-07-24

SOURCE POSITION

Official documentation and source repositories support the current capability model. Source records preserve publisher, title, URL, purpose, and review date in the machine-readable authority register. Commercial terms, performance, and environment-specific behavior remain unverified until controlled proof.

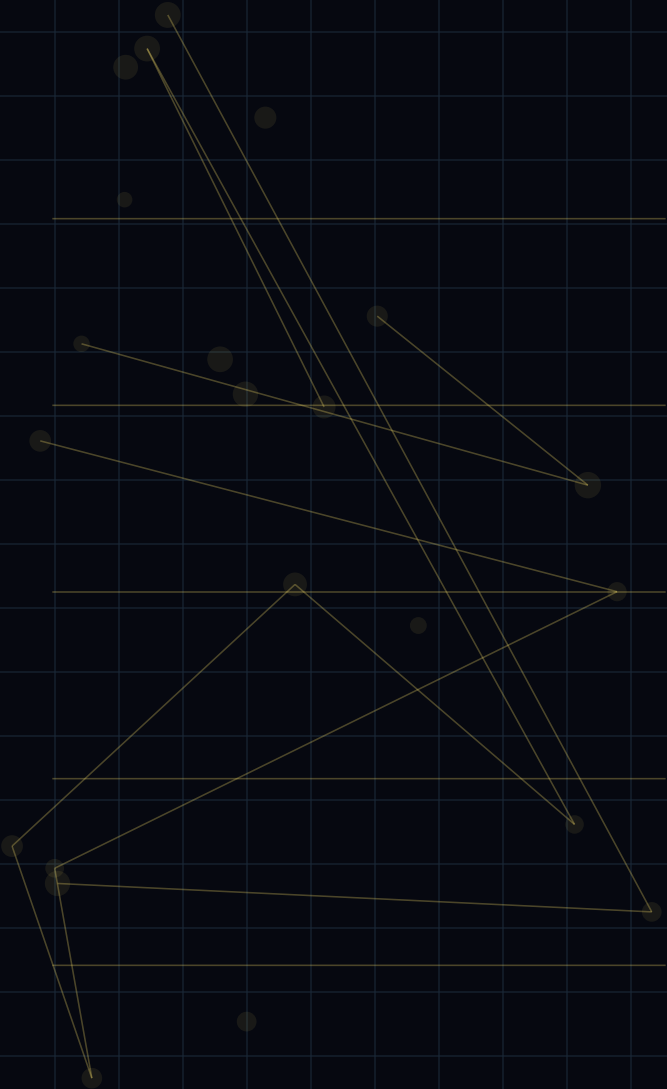
REVISION HISTORY

1.1.0-candidate / 2026-07-24 - permanent-publication rebuild using the final benchmark system, profile-specific rankings, evidence adjustment, mandatory gates, and controlled PoC requirements.

CMP-008

Configuration Management

Fleet configuration, convergence and orchestration platforms



CURRENT EVIDENCE SNAPSHOT

Official product sources refreshed through 2026-09-17
Controlled Mythos laboratory rerun: NOT ASSERTED

VERSION 1.2.0-candidate

CANDIDATE COMPARATIVE EVALUATION / PUBLIC

BENCHMARK DOCTRINE

Gates before scores. Evidence before certainty. Profile fit before universal ranking.

Configuration Management

Fleet configuration, convergence and orchestration platforms

DOCUMENT ID

CMP-008

VERSION

1.2.0-candidate

STATUS

Candidate Comparative Evaluation

PUBLICATION DATE

2026-09-17

SCHEDULED REVIEW

2026-12-16

CANONICAL ROUTE

/library/evaluations/cmp-008/

4

CANDIDATES

9

MANDATORY GATES

E2

CURRENT MAX EVIDENCE

18

ATOMIC CRITERIA

3

WORKLOAD PROFILES

CURRENT DECISION BOUNDARY

Ansible remains a strong July documentation baseline, but that is not a universal selection. Puppet lifecycle changes, Chef 19 migration, and Salt runtime evolution make current lab evidence mandatory. Prove drift, content trust, secrets, fleet scale, rollback, recovery, and upgrade behavior.

NO CURRENT UNIVERSAL WINNER IS PUBLISHED FROM THIS REFRESH.

July 24 baseline scores are preserved as lineage and sensitivity evidence, not silently reissued as September laboratory results.

Red Hat Ansible Automation P

AAP 2.6 patch / 2026-08-04

REFRESH TRIGGER

Current 2.6 patch line includes security fixes and updates across controller, hub, Event-Driven Ansible, and Receptor.

Puppet Enterprise

PE 2025.11.3 / 2026-09-10

REFRESH TRIGGER

Puppet is transitioning support lifecycle models while updating certificate, patching, PostgreSQL, security, and platform support.

Chef Infra

Client 19.3.15 LTS / 2026-05-22

REFRESH TRIGGER

Client 19 LTS moves to Habitat-based packaging; migration and platform compatibility are first-class selection risks.

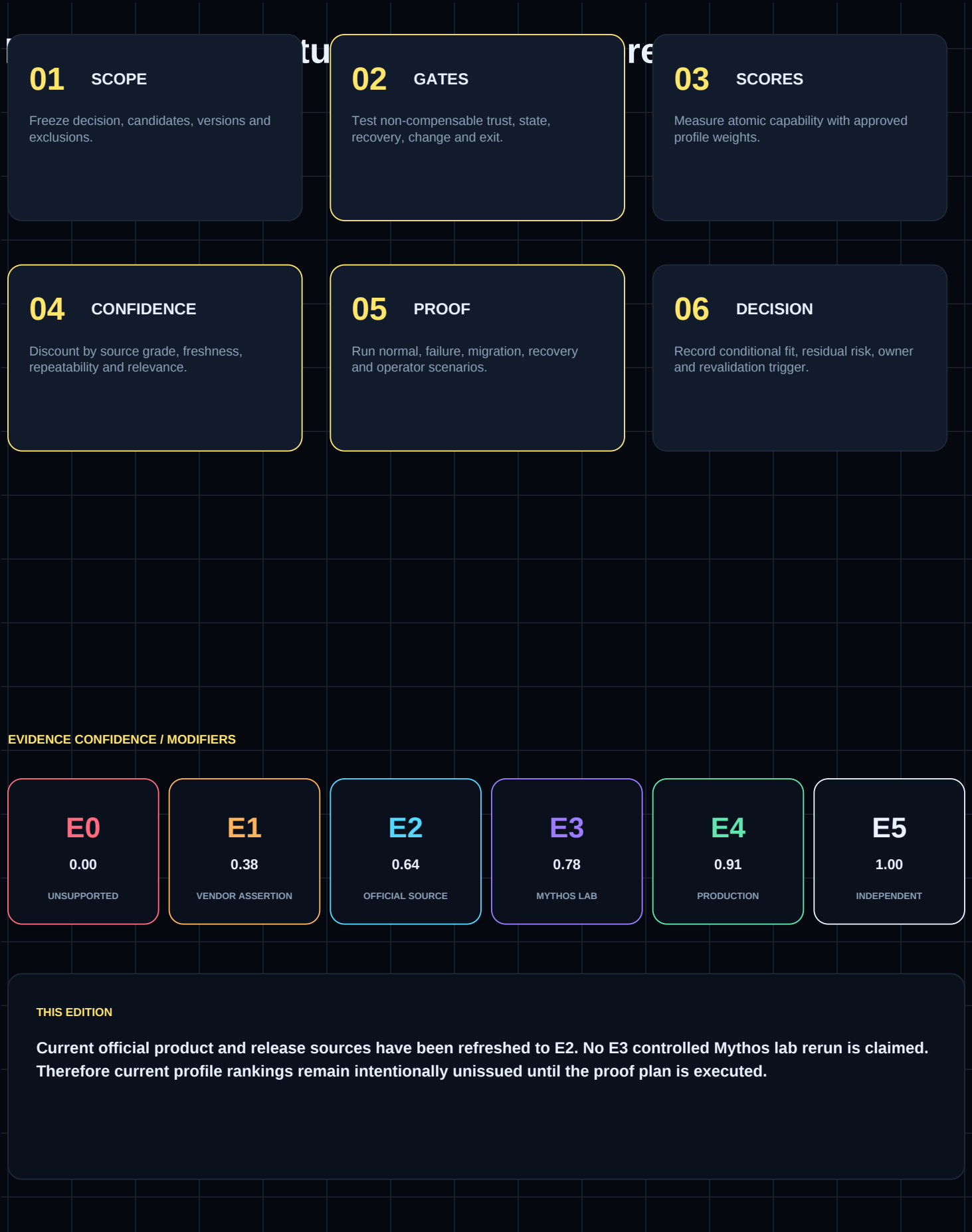
Salt

3008.2 line / current docs

REFRESH TRIGGER

Current 3008 line includes packaging/runtime changes while 3006.x remains an important long-lived branch.

FINAL BENCHMARK SYSTEM / DECISION ARCHITECTURE



What changed since the July benchmark snapshot

Changes below are sourced from current official release documentation. They identify revalidation triggers; they are not translated into new numeric scores without controlled proof.

Red Hat Ansible Automation Platfor AAP 2.6 patch / 2026-08-04 MATERIAL DELTA Current 2.6 patch line includes security fixes and updates across controller, hub, Event-Driven Ansible, and Receptor. CRITERIA TOUCHED C02 / C05 / C07 / C09 / C14 MODERATE REVALIDATION	Puppet Enterprise PE 2025.11.3 / 2026-09-10 MATERIAL DELTA Puppet is transitioning support lifecycle models while updating certificate, patching, PostgreSQL, security, and platform support. CRITERIA TOUCHED C02 / C04 / C09 / C14 / C17 HIGH REVALIDATION
Chef Infra Client 19.3.15 LTS / 2026-05-22 MATERIAL DELTA Client 19 LTS moves to Habitat-based packaging; migration and platform compatibility are first-class selection risks. CRITERIA TOUCHED C04 / C09 / C11 / C14 / C16 / C17 HIGH REVALIDATION	Salt 3008.2 line / current docs MATERIAL DELTA Current 3008 line includes packaging/runtime changes while 3006.x remains an important long-lived branch. CRITERIA TOUCHED C04 / C08 / C09 / C14 / C17 MODERATE REVALIDATION

MANDATORY GATES / CURRENT REVALIDATION POSTURE

Mandatory gates remain non-compensable

No current release is marked PASS solely from documentation. E2 means the official source supports the capability path; LAB and RETEST identify proof still required. HOLD blocks a current decision.

MANDATORY GATE	RED HAT ANSIBLE	PUPPET ENTERPRI	CHEF INFRA	SALT
C01 Functional coverage	E2	E2	E2	E2
C02 Security / trust	RETEST	RETEST	E2	E2
C03 Governance / approval	E2	E2	E2	E2
C04 State integrity	E2	RETEST	RETEST	RETEST
C07 Observability / evidence	RETEST	E2	E2	E2
C09 Resilience / continuity	RETEST	RETEST	RETEST	RETEST
C11 Change safety / rollback	LAB	LAB	RETEST	LAB
C16 Portability / exit	LAB	LAB	RETEST	LAB
C18 Assurance confidence	HOLD	HOLD	HOLD	HOLD

LEGEND

E2	official-source support only	LAB	controlled proof required	RETEST	release delta touches this gate	HOLD	decision gate remains closed
----	------------------------------	-----	---------------------------	--------	---------------------------------	------	------------------------------

Historical profile sensitivity - preserved, not reissued

Values below are the 2026-07-24 evidence-adjusted baseline. They show sensitivity to operating-model weights. The September refresh does not recompute rank because E3 lab proof has not been reproduced.



CURRENT RANK STATUS: WITHHELD PENDING CONTROLLED PROOF.

Evidence confidence is separate from capability

E5

Independent repeatable validation

confidence modifier 1.00

E4

Production evidence in adopting environment

confidence modifier 0.91

E3

Controlled Mythos laboratory result

confidence modifier 0.78

E2

Official documentation / release / source

confidence modifier 0.64

E1

Vendor assertion or marketing

confidence modifier 0.38

E0

Unsupported or unverified

confidence modifier 0.00

CURRENT EVIDENCE STATE

Red Hat Ansible Automation P

E2 REFRESHED

Historical adjusted confidence: 58.2%

E3 LAB: NOT REPRODUCED

Puppet Enterprise

E2 REFRESHED

Historical adjusted confidence: 58.2%

E3 LAB: NOT REPRODUCED

Chef Infra

E2 REFRESHED

Historical adjusted confidence: 58.2%

E3 LAB: NOT REPRODUCED

Salt

E2 REFRESHED

Historical adjusted confidence: 58.2%

E3 LAB: NOT REPRODUCED

ASSURANCE RULE

A current release note can change capability context, but it cannot raise assurance beyond E2. Current recommendation strength is therefore bounded until the test plan produces reproducible artifacts.

MIGRATION, LIFECYCLE AND IRREVERSIBLE-COMMITMENT RISK

Lifecycle risk is evaluated independently of features

The benchmark models migration, upgrade, compatibility, support, staffing, data/state export, downtime and exit. Current release changes below are explicit proof triggers.

Red Hat Ansible Automation Platform MODERATE REVALIDATION Current 2.6 patch line includes security fixes and updates across controller, hub, Event-Driven Ansible, and Receptor. RETEST: C02 / C05 / C07 / C09 / C14
Puppet Enterprise HIGH REVALIDATION Puppet is transitioning support lifecycle models while updating certificate, patching, PostgreSQL, security, and platform support. RETEST: C02 / C04 / C09 / C14 / C17
Chef Infra HIGH REVALIDATION Client 19 LTS moves to Habitat-based packaging; migration and platform compatibility are first-class selection risks. RETEST: C04 / C09 / C11 / C14 / C16 / C17
Salt MODERATE REVALIDATION Current 3008 line includes packaging/runtime changes while 3006.x remains an important long-lived branch. RETEST: C04 / C08 / C09 / C14 / C17

IRREVERSIBLE COMMITMENT GATE

Before migration or procurement lock-in, prove rollback, state portability, operational reconstruction, and supplier-exit assumptions.

Controlled proof is the next decision-producing step

REPRODUCTION STATE / NOT EXECUTED IN THIS EVIDENCE-REFRESH RELEASE

01

CONTROLLED SCENARIO

Manage representative Linux, Windows, network, cloud, middleware, and application targets.

02

CONTROLLED SCENARIO

Execute configuration, patch, compliance, orchestration, secret, and remediation workflows.

03

CONTROLLED SCENARIO

Test dry-run, check or noop mode, canary, concurrency, failure, retry, rollback, and recovery.

04

CONTROLLED SCENARIO

Inject controller loss, certificate expiry, repository outage, partial fleet reachability, and conflicting state.

05

CONTROLLED SCENARIO

Measure convergence time, drift detection, job throughput, operator effort, and evidence completeness.

06

CONTROLLED SCENARIO

Rebuild the control plane and representative nodes from versioned content and backups.

EVIDENCE PACKAGE

Preserve exact versions, architecture, configuration, scenario, expected result, instrumentation, raw output, timing, failures, operator actions, deviations, acceptance state, owner, and artifact hashes.

Current decision posture

Ansible remains a strong July documentation baseline, but that is not a universal selection. Puppet lifecycle changes, Chef 19 migration, and Salt runtime evolution make current lab evidence mandatory. Prove drift, content trust, secrets, fleet scale, rollback, recovery, and upgrade behavior.

CURRENT RANKING: NOT REISSUED / CONTROLLED LAB REQUIRED

CURRENT OFFICIAL-SOURCE REGISTER / CHECKED 2026-09-17

Red Hat Ansible Automation P	E2 / OFFICIAL	
Red Hat AAP 2.6 Aug 4 release		docs.redhat.com
Puppet Enterprise	E2 / OFFICIAL	
Puppet Enterprise release notes		help.puppet.com
Chef Infra	E2 / OFFICIAL	
Chef Infra Client release notes		docs.chef.io
Salt	E2 / OFFICIAL	
Salt 3008.2 release notes		docs.saltproject.io

REVISION LINEAGE

1.2.0-candidate / 2026-09-17 - current-source evidence refresh; no lab rescore issued.

1.1.0-candidate / 2026-07-24 - baseline benchmark using the final comparative evaluation system.

Trigger earlier review after major release, acquisition, licensing change, critical vulnerability, material outage, failed PoC, or workload change.

BASELINE ANALYTIC RECORD CONTINUES ON PAGE 11 FOR TRACEABILITY.

ATOMIC CRITERIA MODEL

WEIGHTED CAPABILITY WITH EXPLICIT MINIMUM EVIDENCE

ID	CRITERION	WEIGHT	GATE	MINIMUM EVIDENCE
C01	Configuration, patch, compliance, orchestration, and remediation coverage	5	YES	Feature documentation, APIs, configuration exports, and scenario demonstration.
C02	Controller, agent, credential, content, and execution trust architecture	5	YES	Threat model, identity flows, RBAC tests, audit records, and security configuration.
C03	Governance and approval controls	5	YES	Approval workflow, policy controls, separation-of-duties tests, and decision records.
C04	Desired-state, inventory, content, report, and run-state integrity	5	YES	Backup, restore, rollback, drift, and corruption-recovery evidence.
C05	Automation and API quality	4	NO	API specifications, authentication tests, rate limits, event samples, and automation runs.
C06	Integration ecosystem	4	NO	Supported integrations, connector tests, failure behavior, and lifecycle ownership.
C07	Observability and evidence	4	YES	Logs, traces, metrics, reports, export tests, and evidence-retention controls.
C08	Node, job, event, content, and controller scale	4	NO	Controlled load tests, topology scale, saturation behavior, and capacity model.
C09	Resilience and continuity	5	YES	Failure injection, HA tests, recovery timing, degraded-mode operation, and runbooks.
C10	Usability and operator cognition	3	NO	Task observation, error-rate measures, navigation tests, and operator interviews.
C11	Dry-run, change window, canary, rollback, and recovery safety	5	YES	Plan or preview output, canary tests, rollback execution, and post-change verification.
C12	Extensibility and programmability	3	NO	Plugin, module, provider, workflow, or code-extension tests and upgrade impact analysis.
C13	Deployment and sovereignty options	3	NO	Deployment architecture, data-flow mapping, residency evidence, and offline tests.
C14	Lifecycle and upgrade discipline	4	NO	Release notes, upgrade test, compatibility matrix, support policy, and migration plan.
C15	Economics and cost predictability	3	NO	Bill-of-materials, licensing model, staffing estimates, metering, and sensitivity analysis.
C16	Portability and exit	4	YES	Export tests, format analysis, replacement workflow, and decommission evidence.
C17	Vendor and ecosystem risk	3	NO	License analysis, roadmap evidence, bus-factor indicators, and dependency inventory.
C18	Assurance confidence	5	YES	Source provenance, lab artifacts, production evidence, independent replication, and review date.

SCORE SCALE

0 absent; 1 materially deficient; 2 partial; 3 adequate with limits; 4 strong; 5 leading for the defined profile. Scores remain provisional until the required evidence grade is achieved.

RAW CAPABILITY SCORECARD

DOCUMENTED CAPABILITY BEFORE EVIDENCE DISCOUNT

CRITERION	ANSIBLE AAP	PUPPET	CHEF	SALT
C01 Configuration, patch, compliance, orchestration, and remediation coverage	4.8	4.4	4.3	4.5
C02 Controller, agent, credential, content, and execution trust architecture	4.4	4.4	4.3	4.2
C03 Governance and approval controls	4.4	4.2	4.1	4.0
C04 Desired-state, inventory, content, report, and run-state integrity	4.2	4.5	4.4	4.3
C05 Automation and API quality	4.8	4.1	4.0	4.6
C06 Integration ecosystem	4.7	4.2	4.1	4.2
C07 Observability and evidence	4.5	4.4	4.3	4.2
C08 Node, job, event, content, and controller scale	4.6	4.6	4.4	4.6
C09 Resilience and continuity	4.3	4.4	4.2	4.1
C10 Usability and operator cognition	4.5	4.0	3.9	4.0
C11 Dry-run, change window, canary, rollback, and recovery safety	4.5	4.2	4.1	4.2
C12 Extensibility and programmability	4.8	4.3	4.4	4.6
C13 Deployment and sovereignty options	4.4	4.3	4.2	4.4
C14 Lifecycle and upgrade discipline	4.3	4.1	3.9	3.8
C15 Economics and cost predictability	3.8	3.7	3.6	4.0
C16 Portability and exit	4.4	4.0	3.9	4.2
C17 Vendor and ecosystem risk	4.1	3.7	3.5	3.6
C18 Assurance confidence	3.7	3.5	3.4	3.4

WEIGHTED BASELINE / 100

Ansible AAP

88.0

Puppet

83.7

Chef

81.4

Salt

83.2

EVIDENCE-ADJUSTED SCORECARD

CAPABILITY DISCOUNTED BY CURRENT EVIDENCE CONFIDENCE

CANDIDATE	RAW	EVIDENCE CONFIDENCE	ADJUSTED	CURRENT STATE
Ansible AAP	88.0	58.2%	51.8	CONTROLLED LAB PENDING
Puppet	83.7	58.2%	49.2	CONTROLLED LAB PENDING
Chef	81.4	58.2%	47.9	CONTROLLED LAB PENDING
Salt	83.2	58.2%	48.9	CONTROLLED LAB PENDING

EVIDENCE SCALE

E0 / 0.00

Unsupported or unverified

E1 / 0.38

Vendor assertion or marketing statement

E2 / 0.64

Official documentation, release notes, or source repository

E3 / 0.78

Controlled Mythos laboratory result

E4 / 0.91

Production evidence from the adopting environment

E5 / 1.00

Independent repeatable validation

CURRENT CONFIDENCE BOUNDARY

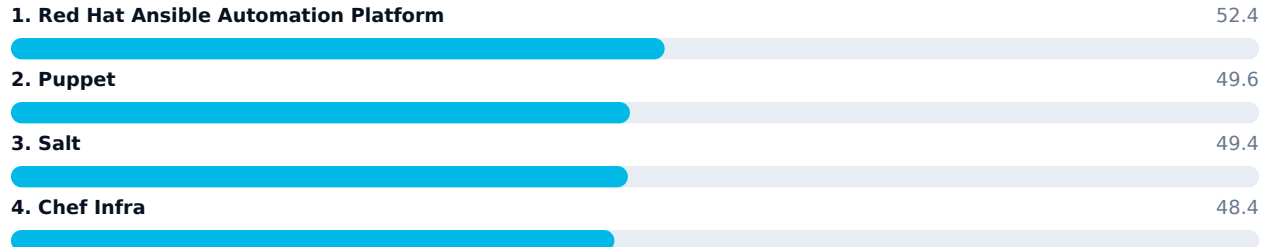
Most candidate criteria are supported at E2: official documentation or source evidence. Environment-specific laboratory, production, and independent evidence remain future gates.

PROFILE-SPECIFIC RANKINGS

EVIDENCE-ADJUSTED SENSITIVITY ANALYSIS

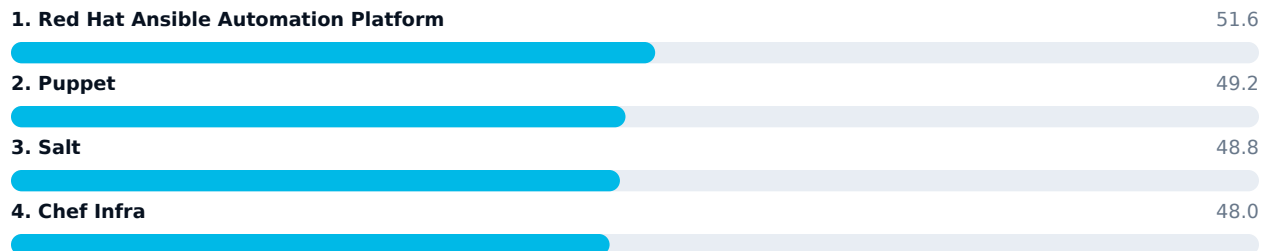
AGENTLESS ENTERPRISE AUTOMATION

Prioritizes broad automation, workflows, execution portability, network and cloud coverage, and low endpoint footprint.



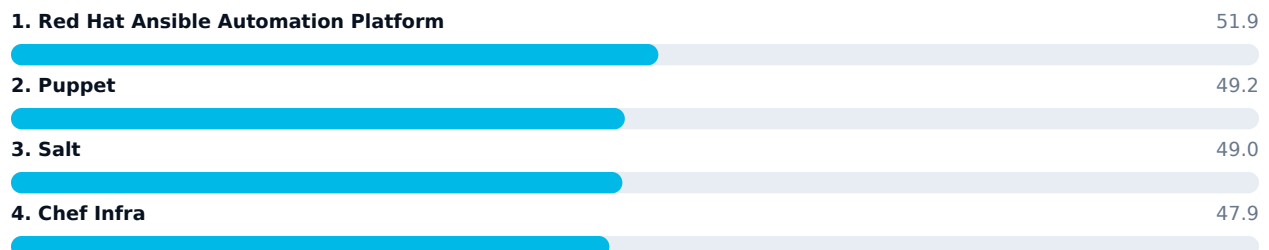
CONTINUOUS DESIRED-STATE ENFORCEMENT

Prioritizes periodic convergence, drift correction, reporting, fleet scale, and stable server configuration.



EVENT-DRIVEN OPERATIONS FABRIC

Prioritizes high-speed remote execution, event reactions, targeting, scale, extensibility, and operational control.



RANK SENSITIVITY

Small score differences are not decision certainty. Treat near-ties as a requirement for deeper PoC, commercial, migration, and operating-model evidence.

CANDIDATE DEEP DIVE / 01

RED HAT ANSIBLE AUTOMATION PLATFORM

OPERATING MODEL

Agentless automation platform using playbooks, inventories, collections, execution environments, controller, API, and event-driven capabilities.

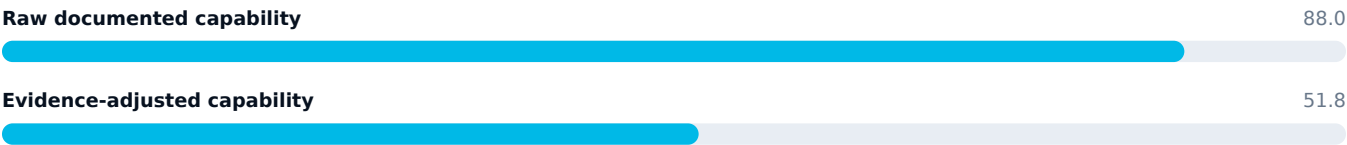
DOCUMENTED STRENGTHS

- Broad automation ecosystem and strong network, cloud, server, and application coverage.
- Agentless SSH and API execution reduces endpoint footprint.
- Execution environments improve reproducibility.
- Controller, RBAC, workflow, and API support enterprise operations.

CAUTIONS AND PROOF OBLIGATIONS

- Idempotence and rollback depend on module and playbook engineering.
- Push execution can miss drift between runs without additional controls.
- Content sprawl and inventory quality require governance.

BASELINE SCORE PROFILE



GATE DISPOSITION

PASS - agentless enterprise automation | CONDITIONAL - content quality and drift architecture

CANDIDATE DEEP DIVE / 02

PUPPET

OPERATING MODEL

Agent-server desired-state configuration management using catalogs, facts, modules, reports, and periodic enforcement.

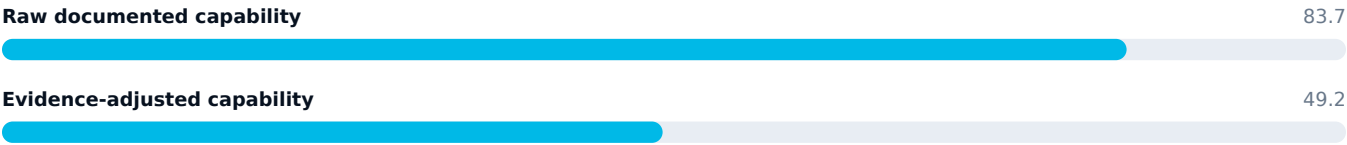
DOCUMENTED STRENGTHS

- Strong desired-state convergence and drift correction.
- Mature agent-server architecture and reporting.
- Rich module and policy ecosystem.
- Suitable for large persistent server fleets.

CAUTIONS AND PROOF OBLIGATIONS

- Agent and server lifecycle add operational responsibility.
- General orchestration and network/cloud breadth may require complementary tools.
- Language, module, certificate, and upgrade governance are specialized.

BASELINE SCORE PROFILE



GATE DISPOSITION

PASS - continuous desired-state server fleet | CONDITIONAL - agent and platform lifecycle

CANDIDATE DEEP DIVE / 03

CHEF INFRA

OPERATING MODEL

Agent-based configuration management using recipes, cookbooks, Policyfiles, Chef Infra Server, and client convergence.

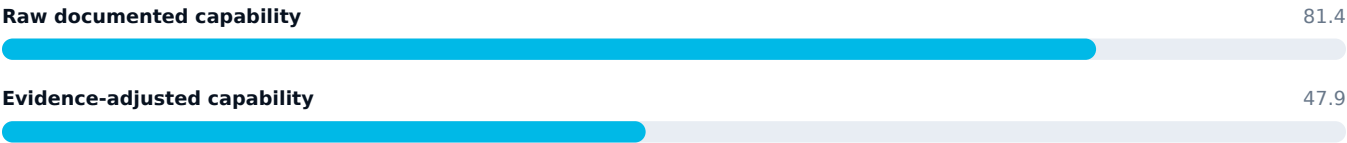
DOCUMENTED STRENGTHS

- Powerful code-driven configuration model.
- Policyfiles support versioned dependency and run-list control.
- Mature convergence architecture and ecosystem.
- Strong fit for organizations with existing Chef engineering capability.

CAUTIONS AND PROOF OBLIGATIONS

- Ruby and cookbook engineering skill are required.
- Platform and product transitions must be evaluated.
- Agent, server, policy, and content lifecycle add complexity.

BASELINE SCORE PROFILE



GATE DISPOSITION

PASS - code-driven convergence for established Chef estates | CONDITIONAL - skills and product trajectory

CANDIDATE DEEP DIVE / 04

SALT

OPERATING MODEL

Event-driven remote execution and configuration-management platform using master/minion, states, pillars, reactors, and execution modules.

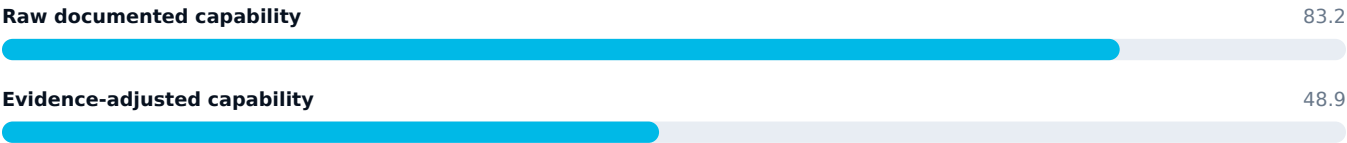
DOCUMENTED STRENGTHS

- Fast remote execution and event-driven operations.
- Flexible states, modules, targeting, and orchestration.
- Strong fit for large-scale operational control and reactive automation.
- Extensible Python-based architecture.

CAUTIONS AND PROOF OBLIGATIONS

- Master, key, event, and minion architecture require strong security and operations.
- Content and state complexity can become difficult to govern.
- Enterprise support, packaging, and lifecycle expectations need validation.

BASELINE SCORE PROFILE



GATE DISPOSITION

PASS - event-driven operations fabric | CONDITIONAL - security, content, and support model

ARCHITECTURE AND INTEGRATION FIT

THE PRODUCT IS ONLY ONE PART OF THE OPERATING SYSTEM

REFERENCE OPERATING LAYERS

01 Inventory, facts, targeting, and desired-state authority

02 Controller, server, master, agent, or execution environment

03 Content, modules, cookbooks, states, playbooks, policy, and secrets

04 Execution, convergence, orchestration, events, and remediation

05 Reports, evidence, recovery, content supply chain, and lifecycle

INTEGRATION BOUNDARIES

- Identity, secrets, PKI, SSH, and privilege management
- Git, artifact, collection, module, cookbook, and package registries
- CMDB, source of truth, ITSM, CI/CD, and policy
- Cloud, network, operating system, application, and endpoint platforms
- Observability, compliance, vulnerability, and security operations

ARCHITECTURE GATE

Every external dependency requires an owner, identity boundary, failure mode, evidence path, version policy, recovery procedure, and exit strategy.

SECURITY, OPERATIONS, LIFECYCLE, ECONOMICS, AND EXIT

CROSS-DOMAIN DECISION RECORD

SECURITY AND AUTHORITY

Identity, credentials, secrets, roles, privileged actions, signing, approvals, isolation, audit, and incident response must be tested as an integrated system.

RELIABILITY AND RECOVERY

Control-plane, data-plane, dependency, region, database, queue, network, and operator failures require defined degraded modes and recovery objectives.

CHANGE AND LIFECYCLE

Version, compatibility, migration, canary, rollback, content, plugin, provider, firmware, and decommission procedures are part of product fitness.

ECONOMICS AND CAPACITY

Licenses, metering, data, compute, hardware, storage, support, staffing, training, integration, migration, and opportunity cost require sensitivity analysis.

SOVEREIGNTY AND PRIVACY

Data location, support access, telemetry, subprocessors, encryption, key control, disconnected operation, and legal obligations must align with policy.

PORTABILITY AND EXIT

Configuration, policy, data, state, evidence, workflows, identities, and operational knowledge must be exportable and reconstructable.

DECISION OWNERSHIP

Architecture, security, operations, finance, procurement, legal, data, and service owners jointly accept the final profile, evidence, residual risk, and exit obligations.

RISK AND FAILURE REGISTER

SCENARIOS THAT CAN INVALIDATE A FEATURE-BASED SELECTION

ID	SEVERITY	FAILURE SCENARIO	REQUIRED TREATMENT
R-01	CRITICAL	Automation changes systems successfully but cannot prove intended service outcome.	PREVENT / DETECT / RECOVER / EVIDENCE
R-02	HIGH	Agent or controller compromise enables fleet-wide execution.	PREVENT / DETECT / RECOVER / EVIDENCE
R-03	HIGH	Content dependencies change without controlled locking or signing.	PREVENT / DETECT / RECOVER / EVIDENCE
R-04	HIGH	Dry-run behavior diverges from actual enforcement.	PREVENT / DETECT / RECOVER / EVIDENCE
R-05	MEDIUM	Desired-state engines fight manual incident recovery or application controllers.	PREVENT / DETECT / RECOVER / EVIDENCE
R-06	MEDIUM	Migration loses inventory semantics, encrypted data, schedules, reports, exceptions, and operator knowledge.	PREVENT / DETECT / RECOVER / EVIDENCE

RISK RULE

A candidate with an unresolved critical failure path cannot advance because optional capability, market position, or commercial discount cannot compensate for missing safety or recovery evidence.

CONTROLLED PROOF-OF-CAPABILITY PLAN

REPEATABLE SCENARIOS, INSTRUMENTATION, AND ACCEPTANCE

TEST 01

Manage representative Linux, Windows, network, cloud, middleware, and application targets.

TEST 02

Execute configuration, patch, compliance, orchestration, secret, and remediation workflows.

TEST 03

Test dry-run, check or noop mode, canary, concurrency, failure, retry, rollback, and recovery.

TEST 04

Inject controller loss, certificate expiry, repository outage, partial fleet reachability, and conflicting state.

TEST 05

Measure convergence time, drift detection, job throughput, operator effort, and evidence completeness.

TEST 06

Rebuild the control plane and representative nodes from versioned content and backups.

EVIDENCE PACKAGE

Preserve versions, architecture, configuration, data set, scenario, expected result, instrumentation, raw output, timing, failures, operator actions, deviations, acceptance state, owner, and artifact hashes.

CONDITIONAL RECOMMENDATION AND REVALIDATION

DECISION RECORD, ACCEPTANCE, AND NEXT EVIDENCE

RECOMMENDATION POSITION

- Select by enforcement model and target estate, not language preference alone.
- Require content-locking, secret, privilege, dry-run, and recovery evidence.
- Define which system owns each configuration domain.
- Revalidate quarterly and before major controller, agent, content, or packaging changes.

CANDIDATE	ADJ. SCORE	GATE POSITION	LAB STATE
Ansible AAP	51.8	PASS - agentless enterprise automation; CONDITIONAL - content quality and drift architecture	PENDING
Puppet	49.2	PASS - continuous desired-state server fleet; CONDITIONAL - agent and platform lifecycle	PENDING
Chef	47.9	PASS - code-driven convergence for established Chef estates; CONDITIONAL - skills and product trajectory	PENDING
Salt	48.9	PASS - event-driven operations fabric; CONDITIONAL - security, content, and support model	PENDING

REVALIDATION SCHEDULE

Review no later than 2026-10-24. Review earlier after a major release, commercial change, critical vulnerability, material outage, architecture transition, failed acceptance test, or change to the target workload profile.

SOURCE AUTHORITY AND REVISION

CURRENT EVIDENCE SNAPSHOT AND PUBLICATION HISTORY

SOURCE ID	PUBLISHER	TITLE	CHECKED
NIST-CSF-20	NIST	Cybersecurity Framework 2.0	2026-07-24
NIST-SP800-53	NIST	Security and Privacy Controls for Information Systems and Organizations	2026-07-24
CIS-CONTROLS	Center for Internet Security	CIS Critical Security Controls v8.1	2026-07-24
ANSIBLE-AAP	Ansible	Red Hat Ansible Automation Platform	2026-07-24
ANSIBLE-EE	Ansible	Ansible Execution Environments	2026-07-24
PUPPET	Perforce	Puppet Platform Components	2026-07-24
CHEF	Progress Chef	Chef Policyfiles	2026-07-24
SALT	Salt Project	Salt States	2026-07-24
SALT-TEST	Salt Project	Salt State Testing	2026-07-24

SOURCE POSITION

Official documentation and source repositories support the current capability model. Source records preserve publisher, title, URL, purpose, and review date in the machine-readable authority register. Commercial terms, performance, and environment-specific behavior remain unverified until controlled proof.

REVISION HISTORY

1.1.0-candidate / 2026-07-24 - permanent-publication rebuild using the final benchmark system, profile-specific rankings, evidence adjustment, mandatory gates, and controlled PoC requirements.