ARCHITECT & CIO WALL CHEAT SHEET

The Enterprise Agentic Software Factory

Level 3 conditional autonomy for AI coding agents: architecture, operating model, 90-day rollout

Yassir Kachri · 2026

1 Why Coding Speed Alone Stalls

Amdahl's Law · triple debt · Level 3 conditional autonomy

Amdahl's Law caps the whole delivery path

Speedup = 1 / ((1 - p) + p / s)

p: coding's share of the time from ready ticket to production; s: how much faster coding gets. Every other stage waits on a person.

Coding share of delivery time	Coding 2x faster	Infinitely fast
20%	1.11x	1.25x
30%	1.18x	1.43x
40%	1.25x	1.67x
50%	1.33x	2.00x

+98% / +91%

Merged pull requests / PR review time in high-AI teams (Faros AI 2025, vendor study)

Amplifier

DORA 2025: AI magnifies an organization's existing strengths and weaknesses

Ungoverned agents run up a triple debt

⚠ Technical debt

Code that works but is hard to change; signal: rework rate. Retired by deterministic checks and four-station verification.

⚠ Intent debt

Behavior nobody wrote down or approved; signal: clarification loops. Retired by durable specifications and the Spec Gate.

⚠ Cognitive debt

Code no current engineer understands; signal: fast approval of large changes. Retired by a Merge Gate on evidence and an independent critic.

Level 3 is the standard for regulated production

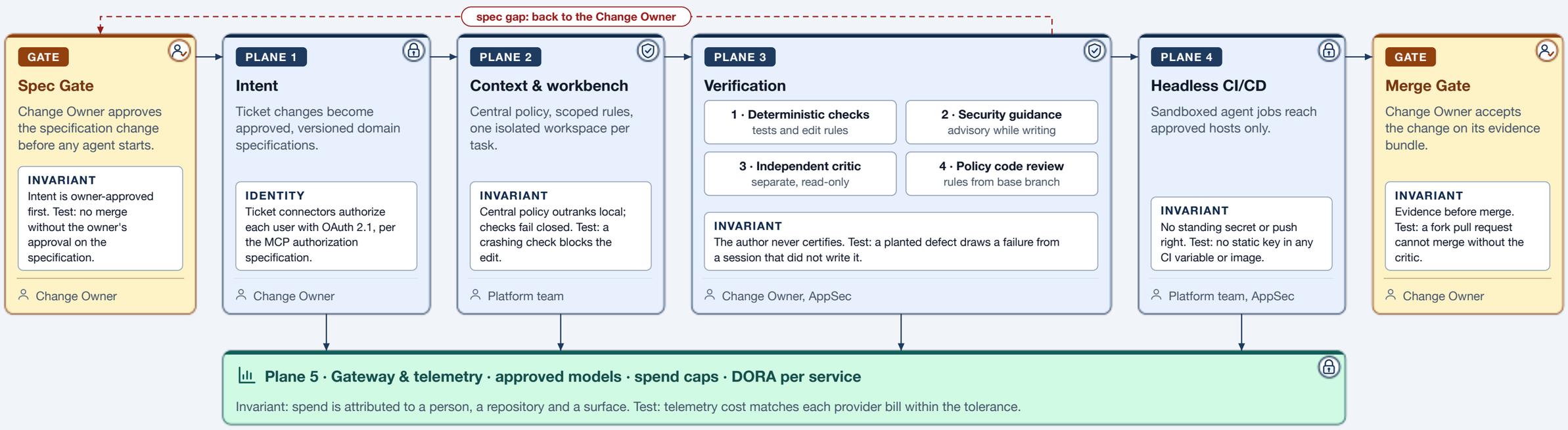
L1 Assisted	Suggestions; a person authors every change
L2 Supervised	Agent edits; a person approves each step
L3 Conditional	**Ticket to pull request; Change Owner approves spec and merge**
L4 High	Merges low-risk classes: no named approver
L5 Full	Ships without approval: out of scope for regulated production

Level 3 holds only while its conditions hold

Non-convergence returns the item to a person at L2; ambiguous intent returns to the Spec Gate; rework above baseline for two weeks (recommended) drops the service to L2.

2 The 5 Planes & 2 Human Gates

End-to-end architecture · control invariants



3 Operating Model

Change management · reversibility governance · tier-1 Day-2 run

Change management: autonomy follows people

Reviewers become Change Owners

Approve the specification change before work; accept on evidence: spec, critic verdict, weakened tests, open threads; the full diff only on a trigger.

Wave 0 Platform → **Wave 1** Two pilots → **Wave 2** Expansion → **Wave 3** Fleet

A service reaches L3 when its people are ready

Two trained Change Owners and a baseline quarter (recommended). It leaves its wave when lead time and rework are at or better than baseline. Brownfield services stay at L2 meanwhile.

Joiner
Identity-provider group grants sign-in

Mover
Old approval right removed the same day

Leaver
Gateway access ends within one hour by default

Governance: decision rights follow reversibility

Two-way	Caps, routing, review depth: one delegated owner
Data flow	L3, egress, connectors: owner of the blast radius
One-way	**Provider, region, failover: Architecture Review Board**

Harder to reverse, decided higher; one accountable role per decision.

Every exception ends

One risk owner, a compensating control, an exit, an end date: 90 days for egress, connector or isolation; 30 days for L3 before readiness (recommended).

⚠ No dispensation exists for

A standing provider key or push right in an agent job · a self-certified change · protection off on an L3 repository · L4 or L5 in regulated production.

Day-2: the factory runs as a tier-1 service

Ring 0 Platform · 1 week → **Ring 1** Canary · 2 weeks → **Ring 2** Fleet

One work item	Change Owner
One service	Platform on-call or Change Owner
Every laptop	Platform on-call, AppSec lead
Every CI agent	Platform on-call, AppSec lead
All laptop traffic	CISO or delegate

99.9% Gateway availability, minutes per month

15 min Longest unmetered window

20 min p90 push to critic verdict

1 day p90 Spec Gate wait

Promoted on replay evidence; kill switches pulled narrowest first and drilled quarterly. Durations and objectives are recommended starting values.

4 90-Day Rollout, Scorecard & Pre-Rollout Checklist

Exit gates · KPIs · executive questions

Ninety days: each window ends at an exit gate

Days 1-30

Rebuild the baseline quarter for two pilots; deploy gateway and telemetry.

EXIT GATE
Every pilot engineer signs in through the gateway; central policy survives it; baselines published.

Days 31-60

Level 3 on the pilots with downgrade triggers live.

EXIT GATE
Readiness criteria met; no unattributed spend; telemetry within 10% of the invoice.

Days 61-90

Add two services; first quarterly readout.

EXIT GATE
Lead time and rework at or better than baseline.

Two pilots, two added services and the 10% tolerance are recommended starting values.

Scorecard: flow and unit cost per service

DORA and FinOps indicators

Lead time · deployment frequency · recovery time
Is delivery faster against the service's baseline?

Change fail rate · deployment rework rate
Is it still stable?

First-pass acceptance · review wait
Do changes merge without being sent back?

Agent cost per merged pull request

(laptop + CI + code review agent spend) ÷ merged pull requests. Never a target alone; lines generated stay off the scorecard.

$150-250 per developer per month

Claude Code docs average, a planning assumption: 250 engineers × $150-250 = $37,500-62,500 per month, before CI and code review.

Before the fleet rollout, each executive can answer

CTO and VP Engineering

- [ ] A full quarter of DORA metrics per service?
- [ ] A Change Owner and deputy per domain?
- [ ] How fast can a service return to L2?

CISO

- [ ] Which agent jobs hold standing keys?
- [ ] Does central policy survive the gateway?
- [ ] Which waivers are open, until when?

CFO and FinOps

- [ ] Does telemetry reconcile with each bill?
- [ ] Is spend attributed and capped per person?
- [ ] What does a merged change cost?

Work item flow · Spec gap: back to a person · Guardrail / sandbox · Identity: OAuth 2.1 / short-lived · Change Owner decision · Human gate · Telemetry & FinOps · Non-dispensable / risk

Evidence standard: every number is a cited external figure, a stated planning assumption, or a recommended starting value to replace with your own; no client results. Sources: G. Amdahl, AFIPS 1967 · Faros AI, The AI Productivity Paradox Report 2025 (vendor study, no control group) · DORA, State of AI-assisted Software Development 2025 and the DORA metrics guide · Claude Code documentation: costs, managed settings, gateway · Model Context Protocol authorization specification (modelcontextprotocol.io).

© 2026 Yassir Kachri · Companion to the strategic whitepaper and the reference guide, The Enterprise Agentic Software Factory · yassirkachri.com