

3-LAYER X-RAY AUDIT — ANONYMIZED SAMPLE

Cloud cost, reliability, and delivery diagnostic for a Series B SaaS scale-up

A redacted sample of the audit deliverable produced from the 14-day diagnostic. Real client identities, exact figures, and stack-specific names are replaced with anonymized placeholders. The structure, methodology, finding density, and prioritization framework are unchanged.

ENGAGEMENT 3-Layer X-Ray Audit	DURATION 14 business days, read-only
CLOUD SURFACE AWS multi-region	ENGINEERING TEAM ~22 engineers across 4 squads
MONTHLY CLOUD SPEND Mid-six-figures USD (anonymized)	ARR \$15-25M (anonymized band)

Anonymization notice. Client identity, account IDs, exact spend figures, service names, and team member names removed. Patterns and methodology preserved. This sample exists to give buyers an accurate sense of what the actual audit deliverable looks like.

Executive Summary

The team came in to the 3-Layer X-Ray with three stated pains: (1) cloud bill outpacing revenue growth, (2) on-call rotation burning out senior engineers, (3) deploy frequency stalled despite headcount growth. The audit identified **seventeen findings** across the three layers, eleven actionable in <30 days, six requiring >30-day execution windows.

TOP THREE QUANTIFIED OPPORTUNITIES

1. **Reserved capacity coverage gap** — current effective coverage 41%, sweet spot 78%. Sized monthly savings: *roughly 1.7x the audit fee, recurring.*
2. **Cross-AZ data transfer leakage** — three microservices unintentionally crossing AZs at peak. Sized monthly savings: *strong four-figure recurring.*
3. **Observability vendor pricing model misalignment** — custom-metric cardinality at 4x the contracted tier. Recommendation includes a renegotiation script.

HIGHEST-RISK RELIABILITY FINDING

Single-AZ database for a P0 customer-facing service, missed during the original migration. Recommended fix in §L2.4. Effort: 2-3 weeks; risk reduction: removes the single largest residual SPOF in the platform.

HIGHEST-LEVERAGE DELIVERY FINDING

CI pipeline median is 24 minutes; 70th percentile is 47 minutes. Cause: redundant test stages and an unindexed cache warmup. §L3.2 contains a sized refactor proposal returning the median to under 9 minutes.

Layer 1 — FinOps

Methodology: tag audit, RI/SP coverage analysis, idle-resource sweep, environment sprawl census, observability tax review.

ID	FINDING	SEV	EFFORT	SIZED IMPACT
L1.1	Reserved Instance / Savings Plans coverage at 41%	HIGH	1 wk	~17% of compute spend
L1.2	Cross-AZ data transfer between core services	HIGH	2-3 wks	Strong four-figure / mo
L1.3	Observability vendor cardinality 4x contracted tier	MED	2 wks	Renegotiation lever
L1.4	Preview environments running prod-tier instances	MED	1 wk	Mid-four-figure / mo
L1.5	S3 lifecycle missing on log buckets >90d	LOW	2 days	Low-four-figure / mo
L1.6	Detached EBS volumes from prior cluster migration	LOW	1 day	Three-figure / mo

DAY 4 QUICK-WIN MEMO (DELIVERED SEPARATELY)

The Day-4 deliverable focused on L1.1, L1.4, and L1.6 — three interventions executable inside the audit window with no production-side change. Engineering team picked up L1.6 internally on Day 7.

Layer 2 — Reliability & SRE

Methodology: SLO discipline review, on-call rotation health metrics, incident history walk-through, DR posture sweep, runbook coverage audit.

ID	FINDING	SEV	EFFORT	RISK
L2.1	SLO definitions missing on 3 of 7 P0 services	HIGH	1 wk	Implicit-vs-explicit gap
L2.2	Alert fatigue: 47% of pages closed as no-action	HIGH	2 wks	On-call burnout & missed signal
L2.3	MTTR P95 >90 min, runbook coverage at 38%	MED	3 wks	Time-to-recovery
L2.4	Single-AZ database on a P0 service	HIGH	2-3 wks	Largest residual SPOF
L2.5	Backup restore not validated in >180d	MED	3 days	Untested DR posture
L2.6	Tribal-knowledge concentration on 1 senior engineer	MED	Ongoing	Bus factor

SPOTLIGHT: L2.4 SINGLE-AZ DATABASE

FIX RECIPE · ESTIMATED 18-22 ENG-DAYS · ZERO DOWNTIME PATH

Add multi-AZ replica → promote to multi-AZ cluster → repoint application config → decommission single-AZ instance. The team had migrated four other databases this way; the runbook from the L2 spotlight is reusable.

Layer 3 — DevOps & Delivery

Methodology: pipeline maturity scoring, IaC drift detection, K8s upgrade posture, observability stack coverage, branching/release strategy review, lead-time-to-deploy benchmarking.

ID	FINDING	SEV	EFFORT	OUTCOME
L3.1	CI median 24 min, P70 47 min	MED	1.5 wk	Target <9 min median
L3.2	Manual rollback procedure, no automation	HIGH	2 wks	Push-button rollback
L3.3	K8s control plane 2 minor versions behind support window	HIGH	3-4 wks	Support compliance
L3.4	Terraform drift on 6 modules	MED	1 wk	Auditable infra state
L3.5	Secrets rotation undefined for 9 services	MED	2 wks	Compliance + security

LEAD-TIME-TO-DEPLOY DISTRIBUTION

Median commit-to-prod is 4h22m. The bottleneck is not technical — it is a manual approval step on every PR. Recommendation: gate the manual approval on PR size + service tier, allowing 80% of PRs to bypass with appropriate quality gates. This brings median down to ~38 minutes without weakening control on high-blast-radius deploys.

Prioritized Roadmap

NEXT 14 DAYS (START DURING THE 30-DAY ASYNC SPARRING WINDOW)

1. L1.1 — Move RI/SP coverage from 41% → 70%+ via 1-year No-Upfront Compute Savings Plan
2. L2.4 — Begin single-AZ DB migration (sized 18-22 days)
3. L3.2 — Rollback automation runbook (sized 2 weeks)
4. L1.5, L1.6 — Quick-win cleanup (3 days total)

30-90 DAYS

1. L1.2 — Cross-AZ traffic remediation (3 services)
2. L2.2 — Alert hygiene + SLO discipline (3 services)
3. L3.3 — K8s minor-version upgrade
4. L3.5 — Secrets rotation policy

90+ DAYS

1. L1.3 — Observability vendor renegotiation (use cardinality data as leverage)
2. L2.3 — Runbook coverage to $\geq 80\%$
3. L2.6 — Reduce bus factor on the senior engineer
4. L3.1 — CI median to < 9 minutes

Methodology

READ-ONLY ACCESS

Audit ran entirely with read-only IAM roles + read-only Terraform Cloud workspace token + read-only observability dashboards. Zero changes to production during diagnosis. The team was not blocked on a single deploy or rollback for the 14-day window.

CADENCE

- Day 0 — 90-min kickoff: stack walkthrough, scope confirmation, access provisioning
- Day 1-3 — L1 deep-dive (async questions in shared Slack)
- Day 4 — Quick-Win Memo delivered (PDF)
- Day 5-9 — L2 deep-dive (60-min incident-history walkthrough)
- Day 10-13 — L3 deep-dive (60-min consolidation call)
- Day 14 — Full report + 90-min recorded debrief
- Day 14-44 — 30-day async sparring channel

TEAM TIME INVESTMENT

Total client engineering time: 3h12m across the 14-day window. The bulk of the audit work happens async on the auditor's side reading IaC, console state, and observability data. The client team is a sounding board, not a source of work.

About this sample. CloudWarrior produces this exact deliverable shape for every 3-Layer X-Ray Audit. Your report would replace the anonymized placeholders with real findings sized against your stack and your bill. The methodology, structure, and prioritization framework are identical.

Want the full audit applied to your stack? cloudwarrior.io/intake — 5-minute qualification, fit verdict in 24 hours.