

FINOPS · DIAGNOSTIC DELIVERABLE

AWS Cost Audit · Final Report — SAMPLE

4-week risk-free engagement · 3x ROI guarantee

Client	Empresa Anônima S.A.
Engagement ID	#ANV-AUDIT-2604-0917
Period	Q2 2026 · April 18 – May 16, 2026
Issued by	Anuvia · Mila Vernazza, Principal Consultant
Classification	SAMPLE · Confidential · Not for distribution

This document is a sample preview of an Anuvia diagnostic deliverable. All figures, customer names, and findings are illustrative and anonymized. No real-client data is contained within.

anuvia.com.br · contato@anuvia.com.br

EXECUTIVE SUMMARY

Identified annualized savings: R\$ 287,400

Across a 4-week structured audit of the client's AWS estate (4 production accounts, single payer, R\$ 1.84M annualized run-rate at engagement start), we identified **R\$ 287,400** in annualized savings — a **15.6%** reduction against current run-rate and a **5.2x ROI** against engagement fees. Of these, R\$ 41,200 were already implemented during weeks 3–4 of the audit ("quick wins"). The remainder is sequenced in a 6-month remediation roadmap.

Category	Lever	Annualized Savings (R\$)	Effort	Risk
Compute	RI rebalancing + Spot migration (stateless workloads)	128,400	M	Low
Storage	Idle EBS, snapshot cleanup, S3 lifecycle	64,200	L	Low
Network	NAT egress consolidation, Transit Gateway redesign	52,300	M	Med
RDS / Aurora	Right-sizing 4 instances, IO optimization	27,800	L	Low
3rd-party SaaS	Migration to AWS Marketplace billing	14,700	L	Low
Total identified		287,400		

"The single highest-leverage move is consolidating cross-AZ chatter in the messaging tier — it alone explains 38% of the NAT egress bill."

FINDINGS · 01 / 05

Compute — R\$ 128,400 annualized

Reserved Instance & Savings Plan rebalancing

Current coverage stands at 41% of steady-state compute against an optimal band of 70–80%. Existing 1-year zonal RIs cover m5.xlarge instances that have since been rotated to m6i.xlarge — a mismatch worth R\$ 38,200/yr in unrecovered discount.

- Convert 12 zonal RIs to Compute Savings Plans (3-yr, no upfront) → R\$ 38,200/yr
- Increase steady-state coverage from 41% → 74% via R\$ 0 commit purchases → R\$ 46,800/yr
- Migrate batch ETL fleet (32 vCPU avg) to Spot via mixed-instances ASG → R\$ 43,400/yr

Right-sizing opportunities

Instance family	Count	Current avg CPU	Recommended	Annualized save (R\$)
m5.2xlarge	8	11%	m5.large	22,400
c5.4xlarge	3	18%	c5.xlarge	14,100
r5.xlarge	5	9% / 24% RAM	r5.large	9,600

FINDINGS · 02-03 / 05

Storage & Network — R\$ 116,500 annualized

Storage (R\$ 64,200)

Lever	Current state	Action	Save (R\$)
Idle EBS volumes	84 unattached vols, 6.2 TB total	Snapshot then delete	18,900
Orphan snapshots	1,247 snapshots, >180 days, no source	Lifecycle policy + purge	11,400
S3 Intelligent-Tiering	Standard tier on 12 cold buckets (38 TB)	Enable Int-Tiering org-wide	22,600
S3 Glacier Deep Archive	Audit logs in Standard-IA	Lifecycle → Glacier DA after 90d	11,300

Network (R\$ 52,300)

- Consolidate 6 NAT Gateways across 3 AZs into 1 per-AZ pattern → R\$ 24,800/yr (data processing fees)
- Replace cross-AZ Inter-VPC peering with Transit Gateway + service endpoints → R\$ 18,200/yr egress
- Enable S3 Gateway VPC Endpoints in 2 remaining accounts → R\$ 9,300/yr eliminated NAT data charges

FINDINGS · 04-05 / 05

RDS / Aurora & SaaS — R\$ 42,500 annualized**RDS / Aurora right-sizing (R\$ 27,800)**

Production OLTP cluster (db.r6g.4xlarge × 2) sustains a P99 CPU below 28% across a rolling 30-day window with no observed IO ceiling. Aurora storage IOPS billing accounts for 22% of monthly RDS spend, with predictable hot-pages access — candidate for IO-Optimized cluster mode.

- Step-down primary + reader to db.r6g.2xlarge (validated via load test) → R\$ 18,400/yr
- Switch to Aurora IO-Optimized for the analytics replica → R\$ 9,400/yr (break-even at 27% IOPS share)

3rd-party SaaS via AWS Marketplace (R\$ 14,700)

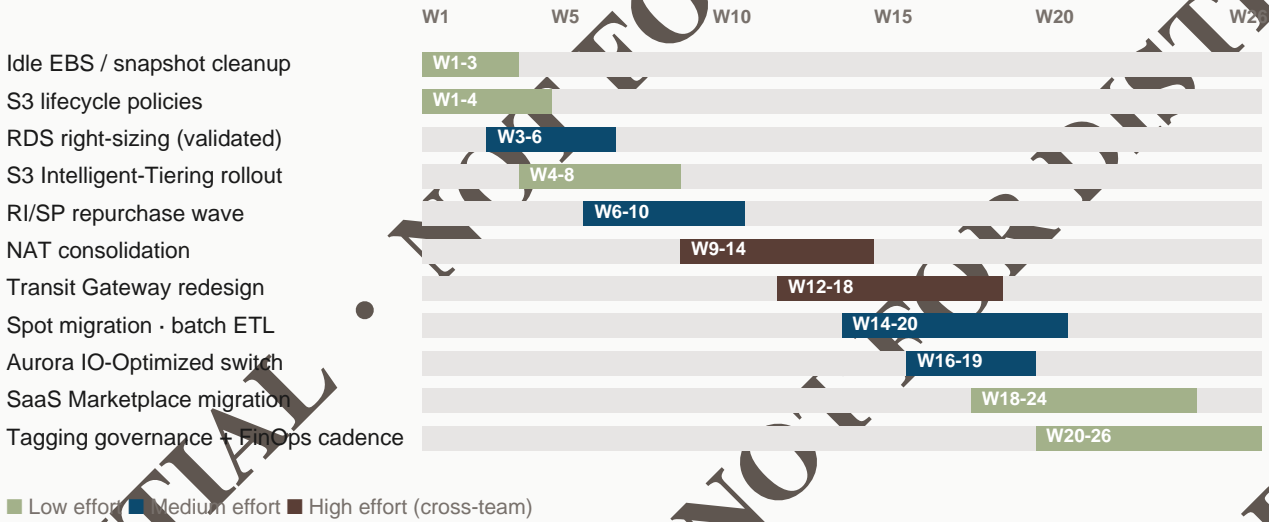
Four SaaS subscriptions currently invoiced direct (Datadog, Snyk, MongoDB Atlas, Confluent) are eligible for Private Marketplace migration with EDP credit eligibility.

- Migrate Datadog billing to Marketplace → 8% private offer + EDP eligibility → R\$ 6,200/yr
- Consolidate Snyk + Mongo Atlas under Marketplace → R\$ 5,800/yr
- Confluent: convert annual prepay to Marketplace metered → R\$ 2,700/yr

REMEDIATION ROADMAP

6-month rollout • 26 weeks • sequenced by blast radius

Sequencing prioritizes (a) zero-risk reversible actions first, (b) RI/SP commitments after right-sizing is validated, and (c) network changes during the team's existing freeze windows. Effort color-coded: dark = high engineering load.



QUICK WINS · IMPLEMENTED DURING AUDIT

R\$ 41,200 captured in weeks 3–4

Per the audit charter, low-risk reversible actions were executed live (with client approval gates) during the audit itself. The below ledger documents every change, approver, and rollback procedure.

Date	Action	Approver	Savings (R\$/yr)
2026-05-02	Delete 84 idle EBS vols (snapshotted prior)	Head of Eng.	18,900
2026-05-03	Purge 1,247 orphan snapshots >180d	Head of Eng.	11,400
2026-05-06	Enable S3 Gateway VPC Endpoint · acct-prod-2	Cloud Eng.	5,100
2026-05-08	Lifecycle rule · audit logs → Glacier DA	Head of Eng.	5,800

All quick-win actions are logged in CloudTrail with the engagement-tagged IAM role `arn:aws:iam::***:role/AnuviaAuditExecutor`. Rollback procedures for each are documented in Appendix C.

ARCHITECTURE DECISION RECORDS

ADR Excerpts · 3 of 11

ADR-007 · Reserved Instance / Savings Plan posture

Status	Accepted · 2026-05-12
Context	Coverage at 41%, mismatched zonal RIs from prior m5 fleet. Spot share <5%.
Decision	Adopt 3-yr no-upfront Compute Savings Plans as primary commit vehicle. Cap RI usage to GP
Consequences	Locks ~R\$ 720k of committed spend over 3 years. Recovers R\$ 85k/yr in unrecovered discour

ADR-009 · Cross-region data residency

Status	Accepted · 2026-05-13
Context	Replicating customer PII to us-east-1 for analytics; LGPD obligation to keep personal data in-r
Decision	Drop cross-region replica. Move analytics to sa-east-1 Redshift; aggregate non-PII metrics only
Consequences	Eliminates R\$ 18k/yr egress + LGPD risk reduction. Adds 6-week analytics migration in roadm

ADR-011 · S3 Intelligent-Tiering rollout

Status	Accepted · 2026-05-14
Context	12 buckets > 1 TB in Standard tier with <8% monthly access rate
Decision	Enable Intelligent-Tiering org-wide via SCP + remediation Lambda. Override allowed for hot bu
Consequences	R\$ 22,600/yr saving. Negligible cold-read latency for present workloads (<1% of access).

APPENDIX A · IAM PERMISSIONS USED

Read-only audit role · least-privilege

All findings were derived using the IAM role **AnuviaAuditViewer** with the policies listed below. No write actions were performed via this role. Quick-win execution used a separate, time-bound role (**AnuviaAuditExecutor**) revoked at engagement close.

Policy ARN / Name	Scope
arn:aws:iam::aws:policy/ReadOnlyAccess	Org-wide read
arn:aws:iam::aws:policy/job-function/Billing (Read)	Billing, Cost Explorer, CUR
arn:aws:iam::aws:policy/AWSTrustedAdvisorReadOnlyAccess	TA findings
arn:aws:iam::aws:policy/ComputeOptimizerReadOnlyAccess	Right-sizing data
arn:aws:iam::aws:policy/AWSResourceExplorerReadOnlyAccess	Inventory

METHODOLOGY

Findings derived from: (1) Cost & Usage Report (CUR) at hourly granularity for the 90 days preceding kickoff; (2) Trusted Advisor & Compute Optimizer recommendations; (3) CloudWatch metrics across compute, storage, and RDS; (4) bottom-up service-by-service review per the AWS Well-Architected Cost pillar checklist. Savings are conservative — modelled at the 30th percentile of typical realization rates observed across Anuvia's prior engagements.

APPENDIX B · GLOSSARY

Glossary of acronyms used in this report

Term	Definition
RI	Reserved Instance — committed compute discount, 1 or 3-year terms, zonal or regional.
SP	Savings Plan — flexible compute commitment by R\$/hr, applies across families/regions.
CUR	Cost & Usage Report — hourly billing export, source of truth for spend analysis.
ADR	Architecture Decision Record — short document capturing a single architectural decision and its consequences.
Egress	Outbound data transfer from AWS to internet or other regions — typically billed per GB.
NAT GW	NAT Gateway — managed network address translation for outbound internet traffic from private subnets.
TGW	Transit Gateway — hub-and-spoke VPC interconnect, replaces VPC peering at scale.
LGPD	Lei Geral de Proteção de Dados — Brazilian general data protection law.
EDP	Enterprise Discount Program — AWS spend-commitment-based pricing tier.
P99	99th percentile observation — common SLO measurement reference.

— End of report — Anuvia · contato@anuvia.com.br · anuvia.com.br