

SAMPLE AUDIT EXCERPT

Conventional SaaS / web application estate

Simulated example. Not a client engagement. Built on a constructed AWS environment so a buyer can see the report format. Every figure below is computed from the stated inputs: annualized × 365/30 for the 30-day baseline month, and each bound of a range rounded down to the nearest \$100. None is measured from a live account.

Prepared by Xavier Perera, Founder · Format date: October 2026 · Baseline convention in this sample: June 2026 (30 days).

1. Environment assumptions

Item	Simulated value
Workload	Multi-AZ SaaS API + workers on ECS Fargate, Lambda, RDS Multi-AZ, S3, NAT Gateways
Region	us-east-1
Monthly net AWS spend (baseline month)	\$45,000
Cost metric	Net amortized cost, credits / refunds / tax excluded
Commitment state	0% Savings Plan / RI coverage
Pricing basis	Public us-east-1 list rates (AWS Price List, checked October 2026) unless noted
Access model	Read-only. No production changes.

What this sample is for. Show how Valmora writes a finding: evidence, mechanism, quantified range, assumptions, effort, production risk, next action, validation.

What this sample is not. A prediction or promise of results for any live account. A live report uses the most recent completed billing month (or the SOW replacement rule) and the account’s own CUR.

2. Executive summary

Three recurring, non-overlapping opportunities are developed in this excerpt. A fourth candidate is evaluated and not claimed.

Opportunity	Monthly range	Annualized range	Sequencing
S3 Gateway Endpoints for S3-bound NAT traffic	\$1,300–\$1,700	\$16,600–\$20,800	Lowest blast radius. Do first.
Unattached EBS + unreferenced snapshots	\$300–\$400	\$4,400–\$5,600	Hygiene. Confirm ownership before delete.
Compute Savings Plan on the post-rightsizing floor	\$1,000–\$2,300	\$12,500–\$28,000	Last. Never size a plan against idle capacity.
Excerpt total (non-overlapping)	\$2,700–\$4,400	\$33,600–\$54,500	
Spot on the customer-facing API	\$0 claimed	\$0	Evaluated. Rejected at this maturity.

Annualized method in a live engagement: baseline-period savings × 365 / actual baseline-period days. Here, a 30-day month, so × 365/30. Each bound is rounded down to the nearest \$100 from unrounded figures, so a total can be up to \$100 per row above the sum of its rounded rows. This sample is illustrative and is not covered by the guarantee. The guarantee applies only to live engagements, subject to the terms of the engagement agreement.

FINDING A · DATA TRANSFER NAT processing on a free path (S3 Gateway Endpoint)

Mechanism. Traffic from private subnets to S3 through a NAT Gateway is billed NAT data-processing and still reaches a service that supports a free Gateway VPC Endpoint. The endpoint removes that S3 traffic from the NAT path. It does not remove NAT hourly charges, internet egress, or non-S3 NAT traffic.

Evidence a live audit would use

- Cost Explorer / CUR: NatGateway-Bytes (or equivalent NAT processing line) for the baseline month.
- VPC inventory: presence or absence of com.amazonaws.us-east-1.s3 Gateway Endpoints on the private route tables.
- The S3-destined share is confirmed by your team (for example from VPC Flow Logs, which the audit role cannot read) or estimated from NAT and S3 CloudWatch metrics. Without destination proof, the finding may be stated as a range, not a point.

Simulated inputs (constructed)

NAT processing observed: 38,000 GB (38 TB) in the baseline month. Share attributable to S3: 80–100% (a range, because the audit role does not read Flow Logs). Public us-east-1 NAT processing rate: \$0.045 / GB. Gateway Endpoint for S3: no hourly charge, no per-GB processing charge.

Economics	38,000 GB × \$0.045 = \$1,710 if 100% is S3. At 80% S3-bound: \$1,368. Claimed range, rounded down: \$1,300–\$1,700 / month; annualized × 365/30: \$16,600–\$20,800. Not claimed: NAT hourly charge (\$0.045 per gateway-hour, \$32.40 per gateway in a 30-day month), internet egress, DynamoDB (separate Gateway Endpoint), cross-AZ hop to the NAT.
Assumptions	Destination mix stays similar after the change. Private route tables can associate the endpoint without breaking existing S3 access patterns (some restrictive endpoint policies need review). Public list rate; a live account with private pricing uses that rate instead.
Effort / risk	Effort: low (hours, not days). Risk: low to moderate. Switching routes drops open TCP connections to S3, and bucket or IAM policies that match aws:SourceIp stop matching. The route-table change is reversible. Residual risk is an overly tight endpoint policy blocking a bucket or a tool that expected the NAT path.
Next action (client implements)	Create S3 Gateway Endpoints in each VPC that shows S3-bound NAT traffic. Associate them with the private route tables. Leave NAT Gateways in place for true internet egress.
Validation / rollback	Within 24–48 hours, NAT processing bytes should fall by approximately the S3 share. S3 API error rates should be unchanged. Rollback = detach the endpoint from the route table.
Overlap rule	This saving is independent of compute rightsizing and commitments.

FINDING B · STORAGE**Unattached EBS volumes and snapshots of deleted volumes**

Mechanism. EBS bills allocated GB-month whether or not a volume is attached. Snapshots of deleted volumes continue to bill snapshot-GB until deleted or expired by a lifecycle policy.

Evidence a live audit would use

- ec2:DescribeVolumes where state = available, with creation time; CloudWatch volume metrics for last recorded activity.
- ec2:DescribeSnapshots whose source volume no longer exists and that no AMI references (ec2:DescribeImages).
- CUR EBS volume and snapshot line items for the baseline month.

Simulated inputs (constructed)

38 volumes, state=available, 2,100 GB (2.1 TB) usable, assumed gp3 at the public \$0.08 / GB-month: 2,100 × \$0.08 = \$168 / month on volumes alone. Snapshot sprawl constructed at 4,000–6,000 GB retained (common when snapshots of deleted volumes are never expired) at the public snapshot rate of \$0.05 / GB-month: \$200–\$300 / month. Combined: \$368–\$468 / month. Claimed range, rounded down: \$300–\$400 / month; annualized × 365/30: \$4,400–\$5,600.

The range above is what the inventory supports.

Assumptions	Volumes are not attached to stopped instances that the team still considers in-service (stopped instances keep volumes attached; those are a different finding). No snapshot is the last copy of data the team cannot rebuild. Deletion is permanent.
Effort / risk	Effort: low. Risk: low-to-medium. The failure mode is deleting the only copy of something. Mitigation is a tagged grace window, not same-day delete.
Next action	Tag unattached volumes pending-delete. Notify the engineering channel. Delete only what is unclaimed after the grace window. Install Data Lifecycle Manager so future snapshots expire.
Validation / rollback	Snapshot anything ambiguous before delete. There is no rollback after deletion.

FINDING C · COMMITMENTS**First Compute Savings Plan, sized after the floor is clean**

Mechanism. Compute Savings Plans discount EC2, Fargate, and Lambda in exchange for a 1- or 3-year \$/hour commitment. They do not cover SageMaker. They cannot be cancelled during the term; a narrow return window exists only for plans ≤ \$100/hour, purchased in the last 7 days, and in the same calendar month.

Evidence a live audit would use

- Cost Explorer Savings Plans coverage = 0%.
- Six-month compute trend (Fargate + Lambda) stability.
- Rightsizing candidates from Compute Optimizer / CloudWatch so the commitment is not sized against idle tasks.

Simulated inputs (constructed)

Current compute (Fargate + Lambda) baseline: \$19,400 / month, ±8% for six months. Estimated idle / oversized share to remove before purchase: 15–25%, which gives an optimized floor of \$14,550–\$16,490 / month. Coverage target: 60–70% of that floor, 1-year no-upfront Compute SP. Discount vs on-demand on the covered slice, from the AWS Savings Plans price list (1-year no-upfront Compute Savings Plan, us-east-1, October 2026): 11.8% on Lambda x86 duration to 20.0% on Fargate Linux/x86 vCPU and memory (not the headline “up to 66%,” which is 3-year all-upfront on specific families).

Economics (the lower bound is the figure counted)	Covered slice: $60\% \times \$14,550 = \$8,730$ to $70\% \times \$16,490 = \$11,543$ / month. Monthly save: $11.8\% \times \$8,730 = \$1,030.14$ to $20.0\% \times \$11,543 = \$2,308.60$. Claimed range, rounded down: \$1,000–\$2,300 / month; annualized $\times 365/30$: \$12,500–\$28,000.
Assumptions	Workload stays on AWS compute (not a replatform off Fargate/Lambda mid-term). No material drop in product usage.
Effort / risk	Effort to purchase: low. Effort to unwind: high. After the 7-day / same-month / $\leq \$100$ -hour window, unused commitment is billed whether or not you use it. Risk: medium-high if purchased against today's \$19,400 instead of the optimized floor.
Next action	Do not buy in week 1. Complete rightsizing and NAT/storage hygiene. Re-measure 30–60 days. Then buy a 1-year no-upfront Compute SP at 60–70% of the new floor. Revisit quarterly.
Validation	Coverage and utilization reports in Cost Explorer. If utilization falls below ~80% for 30 days, stop layering more commitment.

Rejected: Spot on the customer-facing API

Evaluated. Not claimed. The API tier is stateless containers. A published Spot discount of 50–70% off on-demand is real as a rate card, not as a reliability promise. This estate has no interruption handling, no capacity-rebalance, and a strict p99 latency SLO. A two-minute Spot interruption notice against a customer-facing SLO is the wrong trade until batch/async workers are a separate tier and drain tooling exists. That saving is left on the table on purpose.

Sequencing

1. Finding A (S3 endpoints) and Finding B (EBS hygiene): week 1, client-implemented.
2. ECS/Lambda rightsizing: before any Savings Plan. Not fully developed in this excerpt; it is a prerequisite for Finding C, not double-counted in the excerpt total.
3. Finding C: after a clean 30–60 day floor.
4. Spot on API: not now.

Methodology, scope, guarantee

Baseline. Most recent completed billing month unless the SOW replaces it. Annualize as baseline-period savings $\times 365$ / days in that period. Default cost metric: net amortized cost. Exclude Credit, Refund, Tax.

Recurring only. One-off credits and other one-time savings are stated separately and excluded. Deleting an orphaned volume or snapshot is a one-time action, but the charge it stops recurs every month, so Finding B counts as recurring.

No double count. Rightsizing and termination reduce the floor that a Savings Plan is allowed to cover.

Read-only. Cost Explorer, CUR, CloudWatch metrics, resource describe APIs, Compute Optimizer (if enabled). Valmora does not start, stop, purchase, or apply changes. No log contents are read. Lambda sizing uses CloudWatch metrics and function configuration.

Guarantee. This sample is illustrative and is not covered by the guarantee. The guarantee applies only to live engagements, subject to the terms of the engagement agreement.

Pricing disclosure. Public us-east-1 list rates (AWS Price List, checked October 2026). A live report uses client net/effective rates where determinable.

Xavier Perera, Founder, Valmora Consulting · xavier@valmoraconsulting.com · valmoraconsulting.com