



AWS COST · SAMPLE DELIVERABLE

Architecture Changes and Guardrails

Week 5 of 6: Architecture and guardrails · prepared for
Copperwick Therapeutics, Inc.

Prepared by Charles Sieg · charles.sieg@vantalect.com · vantalect.com

March 27, 2026 · AWS Cost Analysis & Optimization: The Optimization (six weeks)

SAMPLE DELIVERABLE

This is a sample deliverable. Copperwick Therapeutics, Inc. is a fictional company: its people, programs, products, accounts, and figures are invented for this sample, and any resemblance to a real company, person, program, or medicine is coincidental. AWS prices are list prices from the AWS Price List API retrieved on October 13, 2026, and from the public AWS Price List bulk API on October 17, 2026; Spot prices are 30-day means for September 14 to October 14, 2026.

Contents

Summary	2
What happened this week	4
The guardrail register	6
Guardrail ownership.....	9
Where the guardrails land, by account family	9
The data-class lifecycle standard	10
The deferred GxP changes, sent to the board	13
The deferral calendar	18
Cost avoidance on planned growth.....	18
Honest limits, as of March 27, 2026	20
What next week does	21
The exact mechanism	21
The VPC template (GR-1)	21
The volume-type policy (GR-2)	22
Spot-by-default Batch (GR-3)	22
Log class and retention (GR-4).....	23
The lifecycle module (GR-5)	24
The Lustre review (GR-6)	25
Idle-GPU alarms (GR-7)	26
Notebook and workstation idle-stop (GR-8).....	26
Budgets and anomaly monitors (GR-9).....	27

Summary

Week 5 counted no savings, and that is by design. Every counted change was in place by Thursday, March 19: the 22 changes in place are estimated at **\$2,526,309 a year, net of Copperwick's 6% private pricing discount**, 13.8% of the in-scope bill's \$18,315,196, and next week measures them on Copperwick's Cost and Usage Report. This week built what keeps the bill down and lands planned growth right: **9 guardrails**, GR-1 to GR-9, each with one owner and built as code ([The guardrail register](#)); the **data-class lifecycle standard**, which says where each kind of data may be stored and who may change it ([The data-class lifecycle standard](#)); and **8 deferred**

GxP changes, of which 7 went to the GxP change advisory board as change requests on Tuesday, March 24, each with its validation impact assessment, and E-2b is a standard change after topline ([The deferred GxP changes, sent to the board](#)).

The guardrails matter because Copperwick’s 2026 plan adds about \$260,000 a month to the in-scope bill by March 2027: a long-read sequencing core in June, a second cryo-EM microscope in August, Feldspar’s Phase 2 biomarker work, ML growth. The patterns that produced this engagement’s savings are the ones that would recur in that growth, in the new Batch environments, instrument buckets, file systems, GPUs, and logs. We priced what the guardrails should prevent as **\$250,000 to \$410,000 a year**, 8% to 13% of the planned growth ([Cost avoidance on planned growth](#)). That figure is **not measured, not counted toward the gate, and not a saving on the bill**. There is no "before" on the Cost and Usage Report for a workload that does not exist yet, so nothing in this report adds it to a saving, a total, or the gate.

One rule runs through all of it: no guardrail shortens a GxP record’s retention. GxP buckets keep Object Lock and versioning, GxP log groups keep their SOP retention, and GxP raw data is never moved to colder storage by a default. Its only path is a change request, co-signed by Dr. Lorraine Osei.

Week 5 in numbers. The last row is a model of the future, not a result, and is not part of any total.

Week 5 at a glance		Result
Savings counted this week		\$0
In place, carried from week 4, estimated, net		\$2,526,309 (22 changes)
Guardrails built, each with one owner		9
Deferred GxP changes sent to the board as change requests		7 of 8 (\$336,844 a year, net, deferred; E-2b is a standard change)
Modeled avoidance on planned growth, a year (not measured, not counted toward the gate, and not a saving on the bill)		\$250,000 to \$410,000

Dates that need an owner

Adaeze Whitfield: no decision is needed this week. Three dates are set and each has an owner. The board votes on N-1b and S3-3 on Tuesday, March 31 (Daniel Achterberg and Grace Mbeki attend). The guardrails for the 9 GMP accounts go in on Monday, March 30, the first business day after the GMP campaign freeze, as a GxP standard change with notice to QA. The 7 clinical accounts follow on Monday, April 27, after topline.

What happened this week

The week ran Monday, March 23 to Friday, March 27. The GMP campaign freeze (Monday, March 16 to Friday, March 27) and the clinical freeze (to Friday, April 24, topline) both applied, and a freeze stops every change to the systems it names, with no exception for cost. A freeze is scoped by system, not by family: the GMP freeze names gmp-lims-prod, gmp-mes-prod, gmp-cds-prod, gmp-historian-prod, and the clinical freeze names clin-sce-prod, clin-sce-val, clin-cdr-prod, clin-biomarker-prod. So the rule for the week was simple: build everything that observes, or that touches a system no freeze names, and leave the 8 frozen systems alone.

Week 5. The nine guardrails went live between Monday, March 23 and Friday, March 27. A change needs seven full days of finalized data after it to be measured, so only changes made by March 22 can be measured on April 1; every counted change was in place by March 19.

Date	What happened
Monday, March 23	Live: GR-1, the VPC template
Tuesday, March 24	Live: GR-2, an SCP that refuses new gp2 volumes; GR-3, spot-by-default Batch compute environments
Tuesday, March 24	The GxP change advisory board receives 7 of the 8 deferred GxP changes as change requests, each with its validation impact assessment; E-2b is a standard change after topline
Wednesday, March 25	Live: GR-4, log class and retention in the account baseline; GR-6, a throughput and capacity review before any new Lustre file system
Thursday, March 26	Live: GR-5, the data-class lifecycle standard; GR-7, idle-GPU alarms on every GPU instance; GR-8, notebook auto-stop and workstation idle-stop defaults
Thursday, March 26	The seven-day window to return C-1 closes; the plan is fixed for the year from here
Friday, March 27	Live: GR-9, budgets and anomaly monitors per account family
Friday, March 27	The GMP campaign freeze ends; the last guardrails for the GMP accounts follow on the next business day
Friday, March 27	This report

Allowed during the freezes: budgets and anomaly monitors, which observe and change nothing a workload runs; a change to any system a freeze does not name, including every development account in both families and the GMP validation environments; and runbooks, templates, and

tickets that wait for a window. **Not allowed:** a change to a frozen system, including a policy that could refuse one of its launches. A change elsewhere in the same family is allowed. The guardrails still wait for March 30 (GMP) and April 27 (clinical) because each is one policy or one StackSet for the whole organizational unit, and an organizational-unit policy reaches the frozen production systems along with the rest of the family.

The guardrail register

The guardrail register, as of March 27, 2026. A guardrail is a small piece of automation with one owner that catches the expensive thing on the first day instead of the first quarterly review.

ID	Guardrail	Owner	Kind	Live	Answers
GR-1	The VPC template with S3 gateway endpoints	Grace Mbeki	Template	Mon, Mar 23	N-1a, N-1b
GR-2	An SCP that refuses new gp2 volumes	Grace Mbeki	Preventive control	Tue, Mar 24	E-1, E-2a, E-2b
GR-3	Spot-by-default Batch compute environments; On-Demand needs an exception tag	Priya Okonkwo	Module and detective control	Tue, Mar 24	B-1, B-1b
GR-4	Log class and retention in the account baseline, with SOP retention on every GxP-tagged log group in any account	Grace Mbeki	Baseline	Wed, Mar 25	L-1, L-2
GR-5	The data-class lifecycle standard, co-signed by Dr. Lorraine Osei for its GxP class	Tomasz Reyes	Standard and module	Thu, Mar 26	S3-1, S3-2, S3-3, S3-4, S3-5
GR-6	A throughput and capacity review before any new Lustre file system	Priya Okonkwo	Review gate	Wed, Mar 25	F-1, F-3
GR-7	Idle-GPU alarms on every GPU instance	Mateo Varga	Detective control	Thu, Mar 26	W-3, B-3
GR-8	Notebook auto-stop and workstation idle-stop as account defaults	Mateo Varga	Account default	Thu, Mar 26	W-2, S-1
GR-9	Budgets and anomaly monitors per account family, routed to each owner	Rafael Quintero	Detective control	Fri, Mar 27	Every account family

How each guardrail is built, where it is live, and how it is undone. The code is in [The exact mechanism](#).

ID	How it is built	Where it applies	Evidence it works	How to undo it
GR-1	A VPC template that creates the S3 gateway endpoint and routes it into every private route table, so a new VPC never sends S3 traffic through a NAT gateway at \$0.045 a GB	Merged; applies to every new VPC. Existing validated VPCs get the endpoint through N-1b by change request (April 14)	A scratch VPC built from the template had the endpoint and route entries, and a read from S3 in it left no NAT gateway bytes	Revert the template; existing endpoints stay
GR-2	A service control policy that denies creating or modifying a volume to gp2	Live in 32 accounts, 6 of them qualified, where it goes in as a GxP standard change with notice to QA; the 9 GMP and quality accounts on Monday, March 30; the 7 clinical accounts on Monday, April 27	A request for a gp2 volume in a sandbox account was refused with an explicit deny	Detach the policy from the organizational unit
GR-3	The Batch compute environment module defaults to Spot with diversified pools; On-Demand needs an exception tag, and a Config rule lists any environment that has neither	Live in 32 accounts, 6 of them qualified, where it goes in as a GxP standard change with notice to QA; the 9 GMP and quality accounts on Monday, March 30; the 7 clinical accounts on Monday, April 27. The biomarker pipelines carry the exception tag until B-1b on May 11	The rule listed the On-Demand genomics and biomarker environments; both now carry the exception tag	Set the module's compute type back to EC2 and delete the rule
GR-4	The account baseline creates log groups in the Infrequent Access class for job logs with their retention set, and a function sets retention on any group created without one; it never shortens a retention that exists. A log group tagged GxP gets the SOP retention in any account, never the research default	Live in 32 accounts, 6 of them qualified, where it goes in as a GxP standard change with notice to QA; the 9 GMP and quality accounts on Monday, March 30; the 7 clinical accounts on Monday, April 27	A test log group created in a sandbox account had its retention set within a minute, and a test group tagged GxP got the SOP value	Delete the rule and the function; existing groups keep their settings
GR-5	A bucket module that takes a required data class and writes the lifecycle rules for it, and a Config rule that lists any bucket without a class; the GxP class writes no transition rule at all	Merged and co-signed by Dr. Lorraine Osei for the GxP class; applies to every new bucket. Existing buckets were mapped to a class by S3-1 to S3-5	A bucket built with each class had the expected rules, and the GxP bucket had Object Lock, versioning, and no transition	Revert the module; existing buckets keep their rules

ID	How it is built	Where it applies	Evidence it works	How to undo it
GR-6	A one-page review that sizes throughput and capacity from measured data, and a service control policy that refuses to create a Lustre file system without a review tag	Live in 32 accounts, 6 of them qualified, where it goes in as a GxP standard change with notice to QA; the 9 GMP and quality accounts on Monday, March 30; the 7 clinical accounts on Monday, April 27	A request for a Lustre file system without the tag was refused; the review was run on the genomics file system and reproduced F-1	Detach the policy; the review stays as guidance
GR-7	A CloudWatch alarm on GPU utilization for every GPU instance, created when the instance starts and sent to the instance's owner; it notifies and never stops an instance, because p4d capacity is hard to get back	Live in the 4 GPU accounts	The alarm fired on a test instance left idle for the evaluation window and reached its owner	Delete the alarm template
GR-8	Every SageMaker notebook starts with the 60-minute auto-stop lifecycle configuration, and every visualization workstation starts with an idle-stop alarm; an exemption tag, approved by the owner, opts one out	Live in the 4 GPU accounts	A new notebook stopped itself after the idle hour, and an exempt one stayed up	Remove the lifecycle configuration and the alarm template
GR-9	One budget and one anomaly monitor per account family, created in the management account, each routed to its owners	Live in all 48 accounts	Each family's first budget notification reached its owners	Delete the budget and monitor resources

Guardrail ownership

A guardrail with no owner decays. Each has one owner who reads the alert, decides whether the cause is expected, and either fixes it or changes the threshold with a reason in the pull request. An exception tag is a request for a decision, not a loophole: the guardrail's owner approves each one, and the monthly cost review lists them.

- **Grace Mbeki, Senior Director, Cloud Platform** (GR-1, GR-2, GR-4). Maintains the VPC template, the gp2 policy, and the log baseline, and approves their exceptions.
- **Priya Okonkwo, Director, Research Computing** (GR-3, GR-6). Approves each On-Demand exception tag on a Batch environment and runs the Lustre review before any new file system.
- **Tomasz Reyes, Director, Scientific Data Platform** (GR-5). Owns the lifecycle module and the data-class standard; the GxP class is co-signed by Dr. Lorraine Osei.
- **Mateo Varga, Director, Computational Design** (GR-7, GR-8). Reads the idle-GPU alarms and approves exemption tags for notebooks and workstations.
- **Rafael Quintero, Vice President, Financial Planning and Analysis** (GR-9). Sets and revisits budget behavior and anomaly thresholds at the monthly review and receives the immediate alerts.

The same people attend the monthly cost review in the handover report.

Where the guardrails land, by account family

The 48 in-scope accounts by family; the frozen systems are those named by a freeze, and no other account is frozen. Budgets and anomaly monitors (GR-9) and the idle-GPU and idle-stop defaults (GR-7,

GR-8) are the exceptions: GR-9 changes nothing in any account, and GR-7 and GR-8 apply only in the 4 accounts with GPUs, none of them frozen.

Account family	Accounts	Validated or Frozen qualified systems	Guardrails land
Platform and shared	7	5 None	Week 5
Research computing	9	0 None	Week 5
Scientific data platform	6	1 None	Week 5
Lab sandboxes	10	0 None	Week 5
Clinical development	7	44 of 7	Monday, April 27, after topline
GMP and quality systems	9	74 of 9	Monday, March 30, after the GMP freeze
Total	48	178	

In the GMP and clinical families, a guardrail that refuses a launch or sets retention is a change to environments that are validated or sit beside validated systems, and it reaches the frozen production systems with the rest of the family, so each family waits for the end of its freeze. The GMP set goes in as a GxP standard change: like-for-like, reversible, with an executed checklist and notice to QA, because a deny that stops a volume type nobody should be creating changes no running resource. The clinical set follows topline.

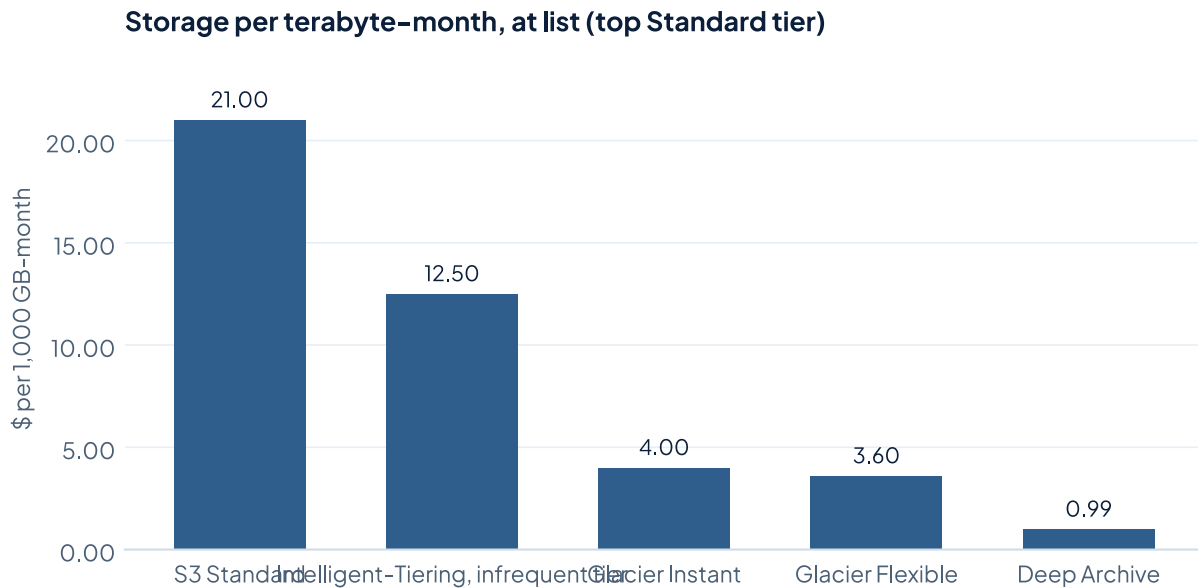
The 6 qualified accounts (cw-log-archive, cw-security-tooling, cw-network, cw-shared-services, cw-backup, sdp-lake-prod) are not frozen and sit in the first wave, but Rule 1 does not cover them. GR-2, GR-3, GR-4, and GR-6 go live there as Rule 2 standard changes with notice to QA, on the dates in the register, and no date moves. In sdp-lake-prod the qualification covers the GxP raw-data buckets and the account controls that protect them; a change to its research buckets follows Rule 1.

The data-class lifecycle standard

Copperwick's data lake holds data with very different obligations: ten-year research raw data, GxP records under Object Lock, results nobody reads after the first month, and files a pipeline can regenerate in an afternoon. Before this week the lifecycle rules were set bucket by bucket, by whoever made the bucket. The standard (GR-5, owned by Tomasz Reyes and co-signed by Dr. Lorraine Osei for its GxP class) fixes one rule per class and makes the class a required parameter of every new instrument bucket.

The five data classes. Standard is the top tier, \$0.021 a GB-month, because Copperwick stores far more than 500 TB there, so every gigabyte that leaves comes off that tier. Intelligent-Tiering and the replica rule are priced by their tiers and rules, not one rate. GxP raw data has no default rule at all.

Data class	Holds	Rule	Storage class	A GB-month, at list	Against S3Restore Standard	Change path
Research raw data	Sequencer run folders (thousands of small files each)	Tar each run with a checksum manifest, verify it, and only then delete the originals by version ID, with the data owner's sign-off. Transition at 18 months	Glacier Deep Archive	\$0.00099	95.3% lessAbout 12 hours (standard retrieval)	Normal change process
Research raw data, large files	Cryo-EM raw movies, one to four gigabytes each	Lifecycle transition at 180 days; reprocessing restores in bulk	Glacier Flexible Retrieval	\$0.00360	82.9% lessBulk restore is free per GB; standard is \$0.01 a GB	Normal change process
GxP raw data	QC instrument data and GLP study data	No default rule. Only a change request moves it, after two years, with Object Lock and versioning kept	Glacier Instant Retrieval, by change request only	\$0.00400	81.0% lessMilliseconds (meets the record-retrieval SOP)	GxP change request, QA approval
Analysis results	Pipeline outputs read often at first and rarely after	Transition to Intelligent-Tiering (the automatic tiers only)	S3 Intelligent-Tiering	By tier	By tierMilliseconds; objects return to the frequent tier on access	Normal change process
Regenerable intermediates	Files that can be recomputed from raw data	Not replicated to the disaster-recovery Region	Replication rule excludes them	Not stored	NotRecompute from raw replicateddata	Normal change process, after the disaster-recovery test



What 1,000 GB costs a month in each storage class the standard uses, at list.

What each class does, and why.

- Research raw data: tar, then Deep Archive.** Sequencer run folders hold thousands of small files each, and Deep Archive adds 40 KB of overhead per object (8 KB billed at the Standard rate and 32 KB at the Deep Archive rate) and charges \$0.05 per 1,000 transitions. An object smaller than about 10 KB costs more archived than left in Standard, before the request fee. So a Batch job tars each run with a checksum manifest, verifies the tar, and only then deletes the originals by version ID, with the data owner's sign-off, so no noncurrent copy stays billed. Research raw data keeps its 10-year retention, and a 12-hour restore is acceptable for it. Standard retrieval is \$0.02 a GB and bulk retrieval \$0.0025. S3-1 is in place; S3-1b, the remaining runs, follows the tar job's second pass.
- Cryo-EM movies: Glacier Flexible Retrieval at 180 days.** Each movie is one to four gigabytes, so there is no small-object problem and no tar job. Reprocessing restores in bulk, which is free per GB; standard retrieval is \$0.01 a GB. The 90-day minimum storage duration never applies to a ten-year hold.
- GxP raw data: Glacier Instant Retrieval, only by change request, never automatically.** QC instrument data and GLP study data keep Object Lock and versioning, and the module writes no transition rule for the GxP class. A move to colder storage is S3-3, a change request with a documented retrieval test and a retrieval at \$0.03 a GB. QA declined Deep Archive for GxP records because a 12-hour restore does not meet the record-retrieval SOP; Glacier Instant Retrieval serves a read in milliseconds, which an inspector's request needs.

- **Analysis results: Intelligent-Tiering.** Access logs show 55% of the objects untouched for 30 days; 20% of all objects, inside that 55%, untouched for 90. So 35% bill at the Infrequent Access tier and 20% at the Archive Instant Access tier. The monitoring fee on about 34 million objects is \$85 a month, and objects move tiers only after 30 days without access, which is why S3-4 is in place and counts for \$0 until its first measurement window, May 11 to May 17, 2026 (the result is ready May 19). That window measures the infrequent tier only; the first objects reach the archive-instant tier on June 9.
- **Regenerable intermediates: not replicated.** The disaster-recovery replicas are research data kept at Standard today. The replication rule's destination class becomes Deep Archive for raw data and results, and about 0.25 PB of intermediates are expired in the replica bucket and stop replicating. That changes the recovery time in the disaster-recovery plan for research data, so it waits for the disaster-recovery test (S3-5). The GxP replicas sit in the Standard-IA line under Object Lock and are not touched.

Restore costs and times are part of the standard, not an afterthought. A class is only as good as its restore, so each row names how long it takes and what it costs, and the owner confirms the restore is acceptable before the class is applied to a bucket.

The deferred GxP changes, sent to the board

15 of the 41 changes touch validated or qualified systems. The validated accounts are about 8.7% of the in-scope bill, so a small share of the savings sits there and this report says so plainly.

GxP changes by status, net of the discount. Only the 4 in place can count toward the gate; deferred, declined, and approved-then-declined changes count for nothing.

The 15 GxP changes	Changes	Annualized, net
In place (S-3a, E-2a, L-2, L-3)	4	\$190,678
Approved, deferred (S-3b, R-2, R-3, R-4, B-1b, E-2b, N-1b, S3-3)	8	\$349,550
Declined at the readout (R-5, D-2)	2	\$134,451
Approved at the readout, then declined (L-4)	1	\$35,194
Total	15	\$709,873

On Tuesday, March 24, the GxP change advisory board received 7 of the 8 deferred changes as change requests (E-2b is a standard change and goes to no board), each with a validation impact assessment, a risk-based test plan, and a release window. A change request is the right path when a change could affect a GxP function: it needs the assessment, testing proportional to the risk (unscripted where the risk is low, scripted operational qualification where a GxP function could be

affected), QA approval from Dr. Lorraine Osei, and a window. The board set each vote and window in the table, one board calendar for all of them. N-1b and S3-3 come back for the vote on Tuesday, March 31, to run on Tuesday, April 14; for S3-3 the board asked for a documented retrieval test first. The rest vote at the board meeting on the last Tuesday at least 14 days before their window. E-2b has no vote: it is a GxP standard change (Rule 2), run in its window with notice to QA. A vote is an approval, not a change, so a freeze does not stop it.

The 8 deferred GxP changes in window order. Votes after March 31 are the meetings before each window, and none has been held. E-2b has no vote: it is a standard change on its window, after topline.

ID	System	Function affected	Risk	Testing	Path	Board vote	Release window	Net a year
N-1b	Validated VPCs (8 TB a month to S3)	A new network path for a qualified system	Low to medium	Scripted verification that S3 traffic takes the endpoint and that access is unchanged, from flow logs and a read-write check	Change request	Tue, Mar 31	Tue, Apr 14	\$6,006
S3-3	GxP raw-data buckets in the data lake (0.42 PB)	The storage class of GxP records, with Object Lock and versioning kept	Medium	Scripted documented retrieval test: a sample of records retrieved through the record-retrieval SOP path, with checksums and Object Lock retention confirmed, and the audit trail read from S3 server access logs and S3 Inventory	Change request	Tue, Mar 31	Tue, Apr 14	\$84,451
E-2b	The clinical systems in the freeze: SCE production and validation, and the biomarker pipelines (55 TB)	None: online and like-for-like, with the same or better IOPS and throughput	Low	Unscripted: each volume's modification state read back, IOPS and throughput confirmed, application health checked	GxP standard change (Rule 2) after topline, with notice to QA and no board vote	None: standard change	Mon, Apr 27	\$12,706
S-3b	SCE validation environment	None: the val environment holds no GxP records, the same evidence as S-3a	Low	Unscripted: the first stop and start watched by the owner; the hold tag honored	Change request	Tue, Apr 7	Mon, Apr 27	\$45,819
B-1b	GCP-validated biomarker pipelines, in their own Batch environment in clin-biomarker-prod	An instance-type change is a revalidation under the pipelines' validation plan	Medium	Scripted revalidation under the validation plan, with a reference dataset run on both environments and compared	Change request, with the revalidation	Tue, Apr 21	Mon, May 11	\$42,020
R-3	LIMS production database (Oracle)	Database performance and failover of a validated system	Medium	Scripted operational qualification regression and a performance test at month-end load, in the LIMS release	Change request	Tue, Apr 28	Mon, May 18	\$30,698

ID	System	Function affected	Risk	Testing	Path	Board vote	Release window	Net a year
R-4	MES production database (SQL Server)	Database performance and failover of a validated system	Medium	Scripted operational qualification regression of the MES database interfaces and a performance test, in the maintenance window	Change request	Tue, May 26	Mon, Jun 15	\$50,065
R-2	SCE production grid	Capacity of a validated statistical computing grid; the queue never held more than six nodes' work outside the database-lock weeks	Medium	Scripted operational qualification regression of job scheduling, run first in the SCE validation environment after topline, plus a performance test at representative load	Change request, after the regression and the post-topline analyses (board June 2)	Tue, Jun 2	Mon, Jun 22	\$77,785
Total, 8 changes								\$349,550

- **N-1b, Grace Mbeki.** S3 gateway endpoints in validated VPCs (8 TB a month). Window: Tuesday, April 14; waits on the board's vote on Tuesday, March 31. Rollback: remove the route.
- **S3-3, Daniel Achterberg.** GxP raw data older than two years to Glacier Instant Retrieval, Object Lock kept (0.42 PB). Window: Tuesday, April 14; waits on the board's vote on Tuesday, March 31, with a documented retrieval test. Rollback: copy back to standard.
- **E-2b, Ines Halvorsen.** gp2 to gp3, the clinical systems in the freeze (SCE production and val, and the biomarker pipelines: 55 TB). Window: Monday, April 27; waits on the clinical freeze: a GxP standard change after topline, with no board vote. Rollback: modify back.
- **S-3b, Ines Halvorsen.** SCE val environment: the same schedule. Window: Monday, April 27; waits on topline: S-3b by change request, E-2b as a standard change. Rollback: remove the schedule tag.
- **B-1b, Ines Halvorsen.** Clinical biomarker pipelines (GCP-validated), in their own Batch environment in clin-biomarker-prod: the same Spot move, 40% of their work. Window: Monday, May 11; waits on biomarker pipeline revalidation. Rollback: point the job queue back.
- **R-3, Daniel Achterberg.** LIMS database: db.r5.8xlarge to db.r6i.4xlarge, Multi-AZ (Oracle BYOL). Window: Monday, May 18; waits on the LIMS release, with operational qualification regression. Rollback: modify back to db.r5.8xlarge in a maintenance window.
- **R-4, Daniel Achterberg.** MES database: db.r5.4xlarge to db.r5.2xlarge, Multi-AZ (SQL Server SE). Window: Monday, June 15; waits on the MES maintenance window after the periodic review. Rollback: modify back in a maintenance window.
- **R-2, Ines Halvorsen.** SCE grid: 8 to 6 r6i.16xlarge nodes after topline. Window: Monday, June 22; waits on the post-topline analyses, the scripted regression in SCE validation, and the board's vote on Tuesday, June 2. Rollback: add the nodes back.

Two things the table does not show. R-2's scripted operational qualification regression runs in the SCE validation environment after topline, and only then does its change request go to the board, so the grid is cut after the post-topline analyses. S3-3's audit evidence is the S3 server access logs, which record each lifecycle transition, and S3 Inventory, which shows each object's storage class: a lifecycle transition creates no CloudTrail data event, so CloudTrail cannot show it.

Five other deferred changes are not GxP: **B-3, F-3, S3-1b, S3-5, G-1**, together \$843,187 a year net. They wait on a cryo-EM campaign, a benchmark, the tar job's capacity, and the disaster-recovery test ([The deferral calendar](#)). 6 of the 8 tests are scripted and 2 are unscripted, in proportion to the risk. None of the 8 shortens a GxP record's retention.

The deferral calendar

Every deferral has one date and one reason, told the same way in every report. 13 changes are deferred, worth \$1,192,737 a year net, and they count for nothing toward the gate until they are in place and measured.

The deferral calendar. Five of the 13 are not GxP; the others are the 8 GxP changes above, 7 change requests and E-2b, a standard change.

Changes	Date	Waiting on
F-3, B-3	Monday, April 6	The next cryo-EM campaign
N-1b, S3-3	Tuesday, April 14	Approved change requests (S3-3 with a documented retrieval test)
S3-1b	Monday, April 20	The tar job's second pass
G-1	Monday, April 20	The molecular dynamics benchmark
S3-5	Tuesday, April 21	The disaster-recovery test
S-3b, E-2b	Monday, April 27	Topline: S-3b by change request, E-2b as a standard change
B-1b	Monday, May 11	Biomarker pipeline revalidation
R-3	Monday, May 18	The LIMS release, with operational qualification regression
R-4	Monday, June 15	The MES maintenance window after the periodic review
R-2	Monday, June 22	After the post-topline analyses, by change request

Cost avoidance on planned growth

Not a saving

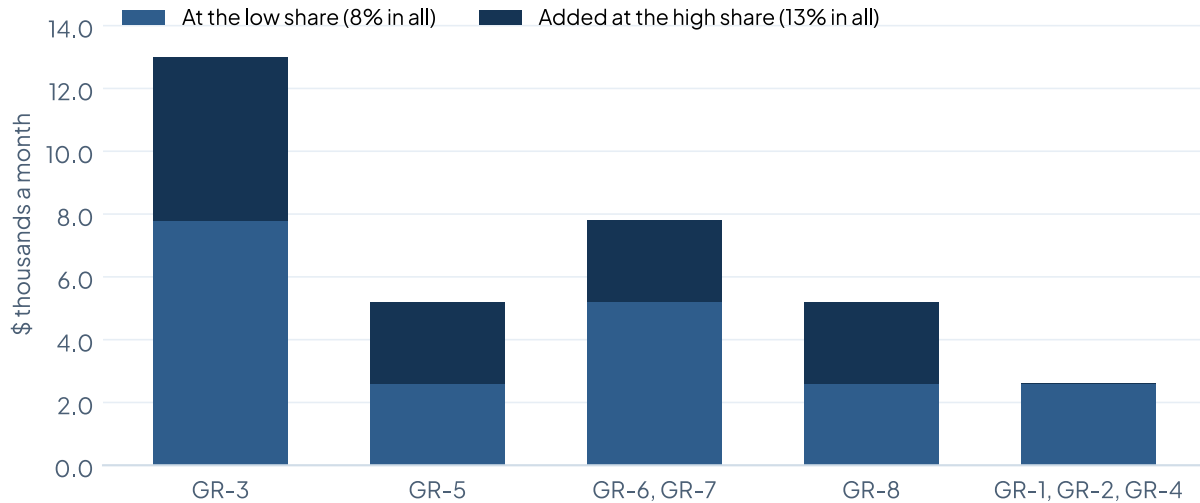
Everything in this section is not measured, not counted toward the gate, and not a saving on the bill. There is no "before" on Copperwick's Cost and Usage Report for a workload that does not exist yet, so the guarantee cannot count it, and this engagement does not. It is a model of what the guardrails should prevent, against Copperwick's own plan.

Copperwick's 2026 plan adds about \$260,000 a month to the in-scope bill by March 2027: a long-read sequencing core in June, a second cryo-EM microscope in August, Feldspar's Phase 2 biomarker work, ML growth. The model applies to that growth the share this engagement's findings cut from the estate it found, lever by lever, and every lever is a guardrail built this week.

The avoidance model's levers: not measured, not counted toward the gate, and not a saving on the bill. The total is a range from a plan, not an observation.

Lever, built in week 5	How it prevents cost	Guardrails	Owners	Share of planned growth	Monthly, on \$260,000
New Batch compute environments default to Spot with diversified pools	Batch lands on Spot	GR-3	Priya Okonkwo	3% to 5%	\$7,800 to \$13,000
Lifecycle rules by data class on every new instrument bucket	New results start in Intelligent-Tiering and intermediates expire (raw data moves only at 18 months, after March 2027)	GR-5	Tomasz Reyes	1% to 2%	\$2,600 to \$5,200
A throughput and capacity review before any new Lustre file system, and an idle-GPU alarm on every GPU instance	No standing peak capacity	GR-6, GR-7	Mateo Varga, Priya Okonkwo	2% to 3%	\$5,200 to \$7,800
Notebook auto-stop and workstation idle-stop as account defaults	GPUs bill when used	GR-8	Mateo Varga	1% to 2%	\$2,600 to \$5,200
The VPC template, the gp2 SCP, and log class and retention in the baseline	Storage, network, and logging defaults	GR-1, GR-2, GR-4	Grace Mbeki	1%	\$2,600
Total				8% to 13%	\$20,800 to \$33,800

Modeled avoidance by lever, a month: not a saving on the bill



Modeled avoidance by lever, a month: not measured, not counted toward the gate, and not a saving on the bill.

At 8% to 13% of the \$260,000 a month the plan adds, avoidance at the March 2027 run rate is **\$20,800 to \$33,800 a month, or \$250,000 to \$410,000 a year** (the monthly range times twelve, rounded to \$10,000). It arrives as each piece of growth lands: a new Batch environment defaults to Spot, a new instrument bucket gets its class's rules, a new file system faces the review, and a new GPU instance gets its alarm. The range is written the same way everywhere in these reports, and it carries the label above wherever it appears.

What is deliberately not claimed. GxP data is never auto-archived by a default, so no lever includes it. The Oracle processor licenses R-3 frees are not on the AWS bill. The fall in Enterprise Support that follows any saving is not counted. None of these is in the model.

Honest limits, as of March 27, 2026

- **No savings, by design.** This week's work does not change what the bill falls by. The 22 changes in place are measured next week, change by change, and nothing here is added to them.
- **The guardrails are new.** Each has a test result, not a record of catching real cost. The first quarter's alerts show whether the thresholds are right: the GPU alarm at 10% for 6 hours, the notebook auto-stop at 60 idle minutes, and the anomaly alerts at \$1,000 and \$10,000.
- **Two account families are still unprotected.** The 9 GMP accounts get their guardrails on March 30 and the 7 clinical accounts on April 27; until then they rely on review.

- **The board's votes on the deferred changes are still ahead.** N-1b and S3-3 are scheduled for March 31; none of the others has been voted on, and E-2b needs none.
- **The avoidance model rests on the plan.** If the long-read core, the second microscope, Feldspar's Phase 2 work, or ML growth arrives larger, smaller, or later, the range moves with it. It is not measured, not counted toward the gate, and not a saving on the bill.

What next week does

Week 6 measures every change on Copperwick's Cost and Usage Report, seven days before against seven days after, net of the discount, and annualizes it. The data runs through March 30, 2026, the last day the report will have finalized when the measurement runs on Wednesday, April 1. It then decides the second invoice (\$17,500) and hands the playbook over on Thursday, April 2. The measurement counts only the 22 changes in place and L-4 at its estimate; the guardrails and the avoidance model are not in it.

The exact mechanism

Everything here is in `cw-landing-zone` and deployed as CloudFormation. The organization-level pieces (the service control policies and the budgets and anomaly monitors) live in one stack in the management account. The account-level pieces (the Config rules, the baseline functions, the alarms, and the modules) go out as a StackSet to the accounts in [the families table](#), in the order it gives. For the validated and qualified accounts the StackSet is the same pipeline that installed the read-only role in week 1, which is a GxP standard change with notice to QA. Placeholders in angle brackets are values the pull request fills from the account list.

The VPC template (GR-1)

The template creates the S3 gateway endpoint and associates it with every private route table, so S3 traffic stays on the AWS network and off the NAT gateway. A gateway endpoint carries only same-Region traffic, so reads from the `us-west-2` replica bucket still use a NAT gateway.

YAML

```
# cw-landing-zone/network/vpc-template.yaml (excerpt)
Parameters:
  PrivateRouteTableIds: {Type: "List<String>"}
Resources:
  S3GatewayEndpoint:
    Type: AWS::EC2::VPCEndpoint
    Properties:
      VpcId: !Ref Vpc
      ServiceName: !Sub com.amazonaws.${AWS::Region}.s3
      VpcEndpointType: Gateway
      RouteTableIds: !Ref PrivateRouteTableIds
```

The volume-type policy (GR-2)

A service control policy denies creating a volume as gp2, modifying one to gp2, and launching an instance whose block-device mapping asks for gp2. gp3 gives 3,000 IOPS and 125 MiB/s included and scales to 80,000 IOPS and 2,000 MiB/s per volume, so a volume that needs more is an io2 Block Express decision made in the design review. A service control policy does not apply to service-linked roles, so an Auto Scaling group launching from an old launch template is not stopped by it; that is the gap the Config rule in the baseline watches.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "DenyNewGp2",
      "Effect": "Deny",
      "Action": ["ec2:CreateVolume", "ec2:ModifyVolume", "ec2:RunInstances"],
      "Resource": "arn:aws:ec2:*:*:volume/*",
      "Condition": {"StringEquals": {"ec2:VolumeType": "gp2"}}
    }
  ]
}
```

Spot-by-default Batch (GR-3)

The Batch module takes an `OnDemandException` parameter. Left empty, the compute environment is Spot with the price-capacity-optimized strategy; set to a ticket reference, it is On-Demand and carries that reference as a tag. A Config custom policy rule lists any Batch compute environment that is On-Demand without the tag, so a hand-made environment is found the day it is made.

YAML

```
# cw-landing-zone/batch/compute-environment.yaml (excerpt)
Parameters:
  OnDemandException: {Type: String, Default: ""} # a ticket reference, for example B-1b
Conditions:
  IsOnDemand: !Not [!Equals [!Ref OnDemandException, ""]]
Resources:
  ComputeEnvironment:
    Type: AWS::Batch::ComputeEnvironment
    Properties:
      Type: MANAGED
      ComputeResources:
        Type: !If [IsOnDemand, EC2, SPOT]
        AllocationStrategy: !If [IsOnDemand, BEST_FIT_PROGRESSIVE,
SPOT_PRICE_CAPACITY_OPTIMIZED]
        InstanceTypes: !Ref InstanceTypes # defaults to [c7i.8xlarge, c6i.8xlarge]
        MinvCpus: 0
        MaxvCpus: !Ref MaxVcpus
        Subnets: !Ref Subnets
        SecurityGroupIds: !Ref SecurityGroupIds
        InstanceRole: !Ref InstanceProfile
        Tags: !If [IsOnDemand, {OnDemandException: !Ref OnDemandException}, !Ref
"AWS::NoValue"]
```

Log class and retention (GR-4)

A log group's class is set when it is created, so the baseline creates job-log groups in the Infrequent Access class with 30-day retention, and research log groups get 90 days. A function sets retention on any group created without one, and it never shortens a retention that exists. On any log group tagged GxP, in any account, the function sets the SOP retention value from the baseline's parameter, and the research default never applies to it; GxP audit groups keep that value.

YAML

```
# cw-landing-zone/baseline/log-groups.yaml (excerpt)
JobLogs:
  Type: AWS::Logs::LogGroup
  Properties:
    LogGroupName: !Sub /batch/${Workload}/jobs
    LogGroupClass: INFREQUENT_ACCESS
    RetentionInDays: 30
```

PYTHON

```
# cw-landing-zone/baseline/set_retention.py (runs on CreateLogGroup)
import os
from typing import Any

import boto3

logs = boto3.client("logs")

def handler(event: dict[str, Any], _context: object) -> None:
    params: dict[str, Any] = event["detail"]["requestParameters"]
    name: str = params["logGroupName"]
    gxp = (params.get("tags") or {}).get("DataClass") == "gxp" # a GxP-tagged group, in any
    account
    groups = logs.describe_log_groups(logGroupNamePrefix=name)["logGroups"]
    group = next(g for g in groups if g["logGroupName"] == name)
    if "retentionInDays" not in group: # never shorten a retention that exists
        days = os.environ["SOP_RETENTION_DAYS"] if gxp else os.environ["RETENTION_DAYS"]
        logs.put_retention_policy(logGroupName=name, retentionInDays=int(days))
```

The lifecycle module (GR-5)

The bucket module takes a required `DataClass`. The research-raw rule applies to the archive/ prefix, where the tar job writes verified tars; the GxP class writes no transition rule, only Object Lock and versioning. A Config rule lists any bucket with no `DataClass` tag.

YAML

```
# cw-landing-zone/data/instrument-bucket.yaml (excerpt)
Parameters:
  DataClass: {Type: String, AllowedValues: [research-raw, cryoem-raw, gxp, results,
intermediate]}
Conditions:
  IsResearchRaw: !Equals [!Ref DataClass, research-raw]
  IsCryoEm: !Equals [!Ref DataClass, cryoem-raw]
  IsResults: !Equals [!Ref DataClass, results]
  IsGxp: !Equals [!Ref DataClass, gxp]
Resources:
  Bucket:
    Type: AWS::S3::Bucket
    Properties:
      Tags: [{Key: DataClass, Value: !Ref DataClass}]
      VersioningConfiguration: {Status: Enabled}
      ObjectLockEnabled: !If [IsGxp, true, !Ref "AWS::NoValue"]
      LifecycleConfiguration:
        Rules:
          - !If [IsResearchRaw, {Id: tarred-runs-to-deep-archive, Status: Enabled, Prefix:
archive/, Transitions: [{StorageClass: DEEP_ARCHIVE, TransitionInDays: 0}]}, !Ref
"AWS::NoValue"]
          - !If [IsCryoEm, {Id: movies-to-flexible, Status: Enabled, Transitions:
[{StorageClass: GLACIER, TransitionInDays: 180}]}, !Ref "AWS::NoValue"]
          - !If [IsResults, {Id: results-to-intelligent-tiering, Status: Enabled, Transitions:
[{StorageClass: INTELLIGENT_TIERING, TransitionInDays: 0}]}, !Ref "AWS::NoValue"]
          - Id: abort-incomplete-uploads
            Status: Enabled
            AbortIncompleteMultipartUpload: {DaysAfterInitiation: 7}
```

The research-raw tar job is a Batch job that writes archive/<run>.tar and archive/<run>.tar.sha256, reads the tar back and checks the manifest, and only then, under the data owner's sign-off, deletes the originals by version ID so no noncurrent copy stays billed. Replication of the intermediate class is excluded by a replication-rule filter on the class tag.

The Lustre review (GR-6)

The review sizes a new file system from measured data. Throughput is the lowest tier whose total is at least 1.25 times the measured 90-day peak, and capacity is set so used space is at most 70% at creation, because Lustre storage cannot shrink. Run on the genomics file system, it reproduces F-1: a 14.8 GB/s peak needs 18.5 GB/s, so 250 MB/s/TiB at 48 TiB (12 GB/s) fails and 500 MB/s/TiB (24 GB/s) is the lowest that passes. The 29 TiB in use is 60% of 48 TiB.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "LustreNeedsReview",
      "Effect": "Deny",
      "Action": "fsx:CreateFileSystem",
      "Resource": "*",
      "Condition": {"Null": {"aws:RequestTag/LustreReview": "true"}}
    }
  ]
}
```

The `LustreReview` tag carries the review's ticket reference, so the requester must have a review on file before the file system can be created.

Idle-GPU alarms (GR-7)

An event rule on instance start in the 4 GPU accounts creates this alarm for every instance in a GPU family. It notifies the instance's owner and never stops the instance, because p4d capacity is hard to get back.

YAML

```
# cw-landing-zone/gpu/idle-alarm.yaml (excerpt)
IdleGpuAlarm:
  Type: AWS::CloudWatch::Alarm
  Properties:
    AlarmName: !Sub idle-gpu-${InstanceId}
    Namespace: CWAgent
    MetricName: nvidia_smi_utilization_gpu
    Dimensions: [{Name: InstanceId, Value: !Ref InstanceId}]
    Statistic: Average
    Period: 3600
    EvaluationPeriods: 6
    Threshold: 10
    ComparisonOperator: LessThanThreshold
    TreatMissingData: notBreaching
    AlarmActions: [!Ref OwnerTopic]
```

Notebook and workstation idle-stop (GR-8)

Every notebook starts with the 60-minute auto-stop lifecycle configuration, and every visualization workstation starts with an alarm that stops it. An exemption tag, approved by Mateo Varga, opts one out: the notebook's lifecycle configuration is not attached, and the workstation's alarm is not created.

YAML

```
# cw-landing-zone/sagemaker/auto-stop.yaml (excerpt)
NotebookAutoStop:
  Type: AWS::SageMaker::NotebookInstanceLifecycleConfig
  Properties:
    NotebookInstanceLifecycleConfigName: auto-stop-60
    OnStart:
      - Content: !Base64 |
          #!/bin/bash
          # Stops the notebook after 3600 seconds without kernel activity (the AWS auto-stop
script)
          IDLE_TIME=3600
          aws s3 cp s3://<baseline-bucket>/autostop.py /home/ec2-user/autostop.py
          (crontab -l 2>/dev/null; echo "*/5 * * * * /usr/bin/python3 /home/ec2-user/
autostop.py --time $IDLE_TIME --ignore-connections") | crontab -
WorkstationIdleStop:
  Type: AWS::CloudWatch::Alarm
  Properties:
    Namespace: AWS/EC2
    MetricName: CPUUtilization
    Dimensions: [{Name: InstanceId, Value: !Ref InstanceId}]
    Statistic: Average
    Period: 300
    EvaluationPeriods: 12
    Threshold: 5
    ComparisonOperator: LessThanThreshold
    AlarmActions: [!Sub "arn:aws:automate:${AWS::Region}:ec2:stop"]
```

Budgets and anomaly monitors (GR-9)

One budget and one anomaly monitor per account family, created in the management account with a filter on the family's linked accounts, so a new account is a new line in a list. Each budget adjusts itself from the trailing 3 months of actual cost, so the limits need no hand-set dollar figure, and it alerts the family's owners at 80% and 100% of actual cost and at 100% of forecast. Monitors send each owner a daily summary of anomalies with an impact of \$1,000 or more, and the on-call topic gets an immediate alert at \$10,000, about 0.7% of the in-scope monthly bill. There are 6 family monitors and one by service, 7 in all.

YAML

```
# cw-landing-zone/finops/budgets.yaml (management account; the Research computing family shown)
ResearchBudget:
  Type: AWS::Budgets::Budget
  Properties:
    Budget:
      BudgetName: family-research-computing
      BudgetType: COST
      TimeUnit: MONTHLY
      BudgetLimit: {Amount: <set-by-auto-adjust>, Unit: USD}
      AutoAdjustData:
        AutoAdjustType: HISTORICAL
        HistoricalOptions: {BudgetAdjustmentPeriod: 3}
```

```

    CostFilters:
      LinkedAccount: ["<rc-account-ids>"]
    NotificationsWithSubscribers:
      - Notification: {NotificationType: ACTUAL, ComparisonOperator: GREATER_THAN, Threshold:
80, ThresholdType: PERCENTAGE}
        Subscribers: [{SubscriptionType: EMAIL, Address: "<owner-email>"}]
      - Notification: {NotificationType: FORECASTED, ComparisonOperator: GREATER_THAN,
Threshold: 100, ThresholdType: PERCENTAGE}
        Subscribers: [{SubscriptionType: EMAIL, Address: "<owner-email>"}]
    ResearchMonitor:
      Type: AWS::CE::AnomalyMonitor
      Properties:
        MonitorName: family-research-computing
        MonitorType: CUSTOM
        MonitorSpecification: '{"Dimensions":{"Key":"LINKED_ACCOUNT","Values":["<rc-account-
ids>"],"MatchOptions":["EQUALS"]}}'
    ServiceMonitor:
      Type: AWS::CE::AnomalyMonitor
      Properties: {MonitorName: by-service, MonitorType: DIMENSIONAL, MonitorDimension: SERVICE}
    ImmediateAlert:
      Type: AWS::CE::AnomalySubscription
      Properties:
        SubscriptionName: immediate-over-10000
        Frequency: IMMEDIATE
        MonitorArnList: [!Ref ServiceMonitor]
        Subscribers: [{Type: SNS, Address: "<on-call-sns-topic-arn>"}]
        ThresholdExpression: '{"Dimensions":{"Key":"ANOMALY_TOTAL_IMPACT_ABSOLUTE","MatchOptions":
["GREATER_THAN_OR_EQUAL"],"Values":["10000"]}}'

```

© 2026 Vantalect, LLC. Sample deliverable; Copperwick Therapeutics, Inc. is fictional.

AWS, Amazon Web Services, AWS Graviton, and the AWS service names are trademarks of Amazon.com, Inc. or its affiliates. Oracle and Java are registered trademarks of Oracle Corporation and its affiliates. Microsoft, Windows, and SQL Server are trademarks of the Microsoft group of companies. Red Hat and Red Hat Enterprise Linux are trademarks of Red Hat, Inc. PostgreSQL is a registered trademark of the PostgreSQL Community Association of Canada. Kubernetes is a registered trademark of The Linux Foundation. Nextflow is a trademark of Seqera Labs. Vantalect is not affiliated with or endorsed by any of them. Copperwick Therapeutics, Inc. is fictional.