
AWS
re:Invent
DECEMBER 2 – 6, 2024 | LAS VEGAS, NV

DAT318

Powering the grid: GE's 600 TB migration to Amazon Keyspaces

Yogini Parkhi
Director of Software Engineering
GE Digital

Steve Mayszak
Sr. Software Development Manager
Amazon

Speakers

Yogini Parkhi

Director: Analytics Strategy
GE Digital

Steve Mayszak

Sr Manager, Software Engineering
AWS Keyspaces

Agenda

01 Apache Cassandra:
- Introduction
- Primary components
- Architecture
- Common operations

02 Amazon Keyspaces
- Introduction
- Architecture
- Key differences

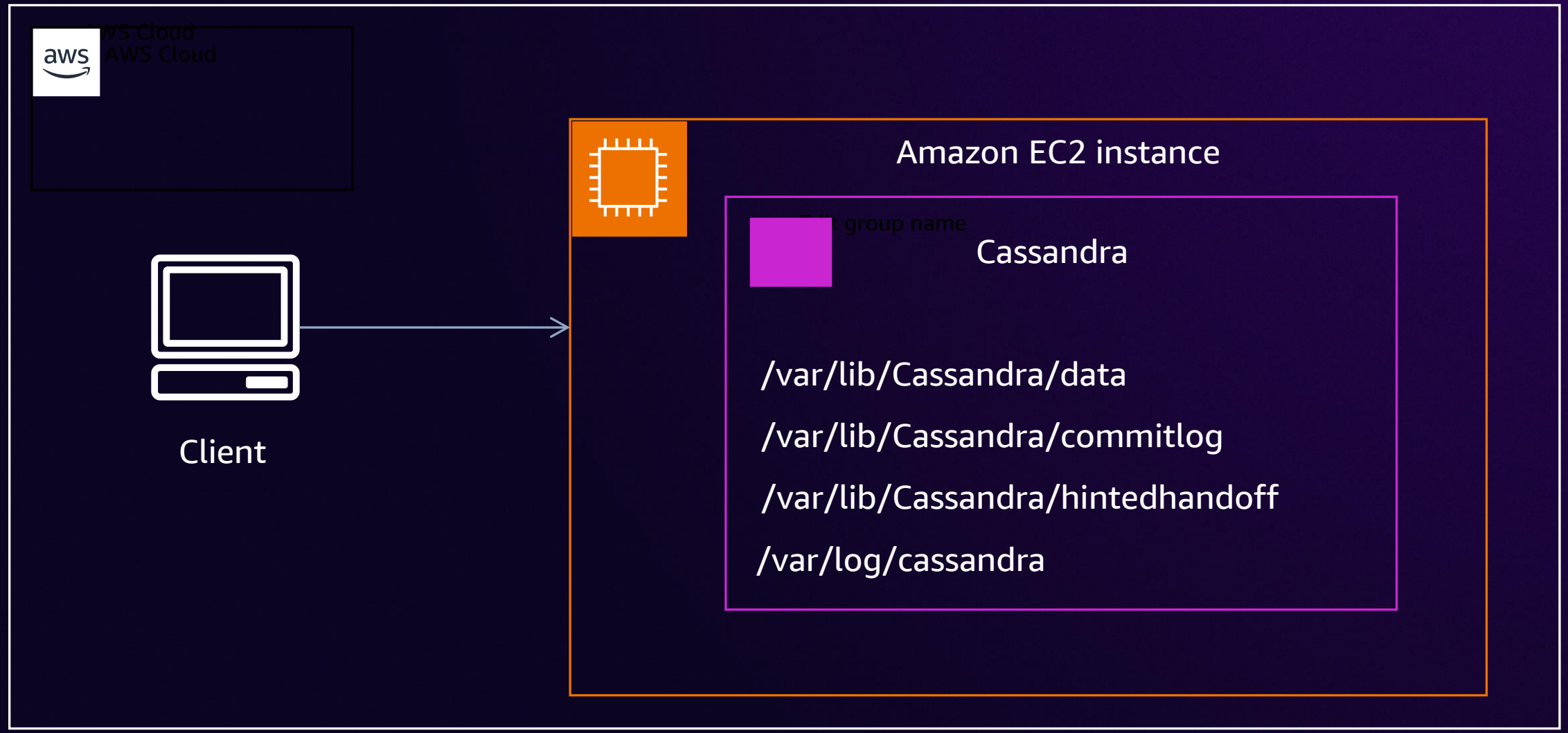
03 Introduction to GE Vernova
- Who we are
- Mission
- Opportunity

04 Migration
- Approach
- Challenges
- Learnings
- Benefits

Apache Cassandra overview

- Open source
- Distributed / wide-column
- High throughput
- No single point of failure
- Commodity hardware

Cassandra application – Simple view

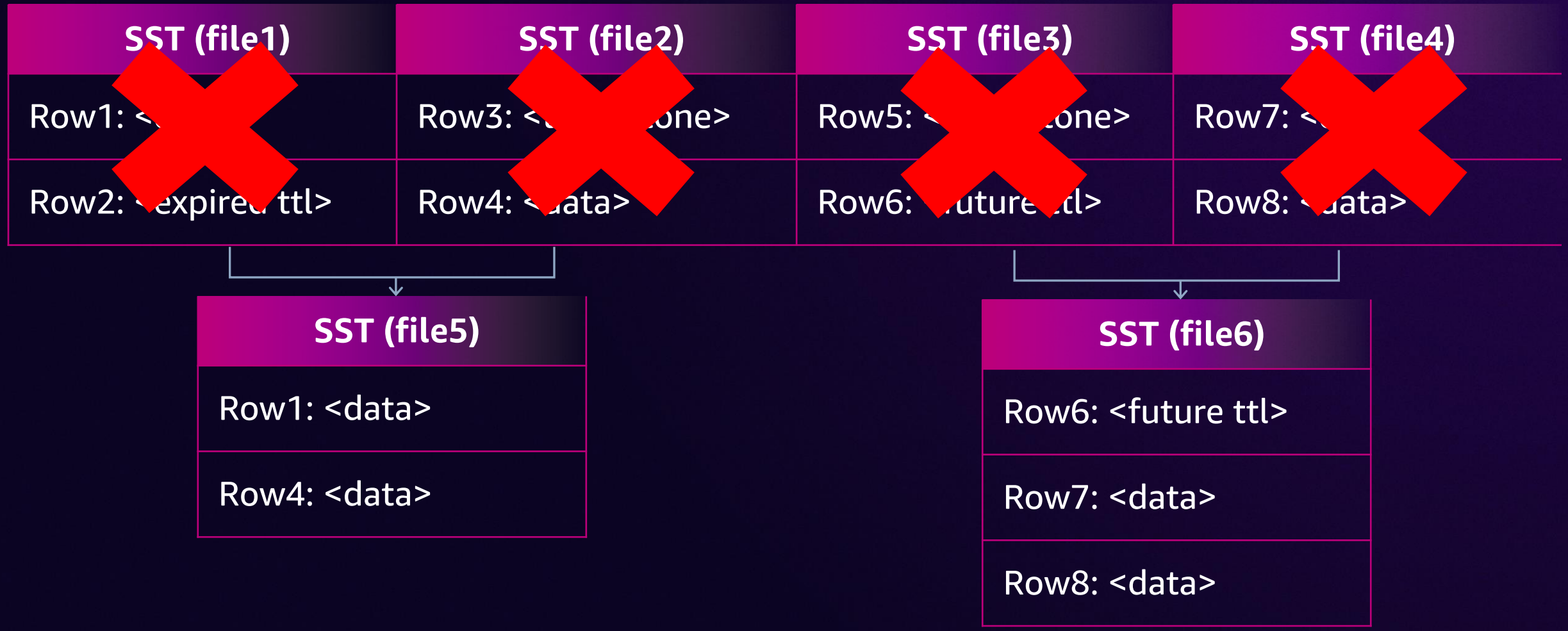


Data on a Cassandra Node

New write request

- Append to /var/lib/Cassandra/commitlog
- Write to memtable (in memory data structure)
- Write to replica node, or to /var/lib/cassandra/hintedhandoff
- When memtable == configured max, flush to disk
- Write an SSTable (/var/lib/Cassandra/data/<sstables>

TTL, tombstone, and compaction



Cassandra's token ring

Token
Determined by output of hash function
Used to locate data in C*
Value: -2^{63} to $+2^{63}-1$(128 bit)

The output of a hash function, known as the *token*, determines the node in which the data resides

Table:

Partition Key: 1
<row data>

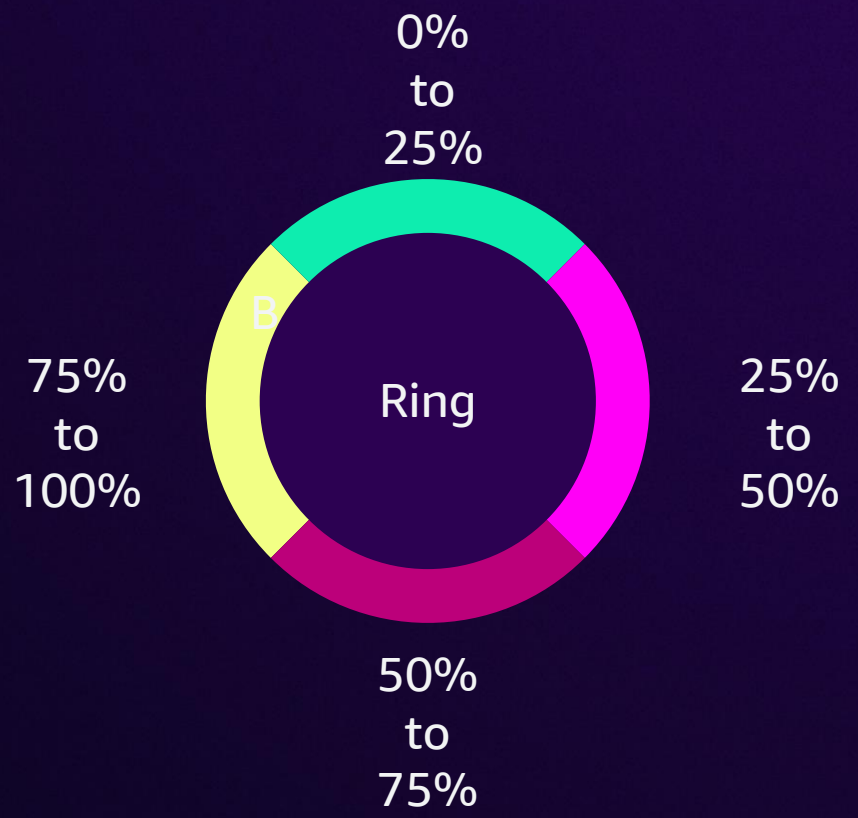
Partition Key: 2
<row data>

Partition Key: 3
<row data>

Hash(1) = -4.5e18

Hash(2) = 2.3e18

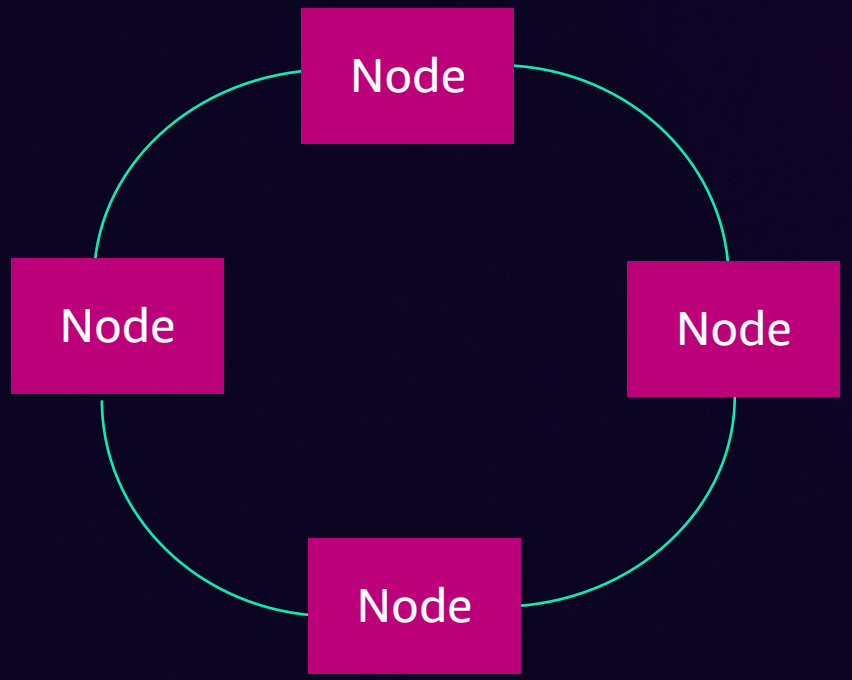
Hash(3) = 4.6e18



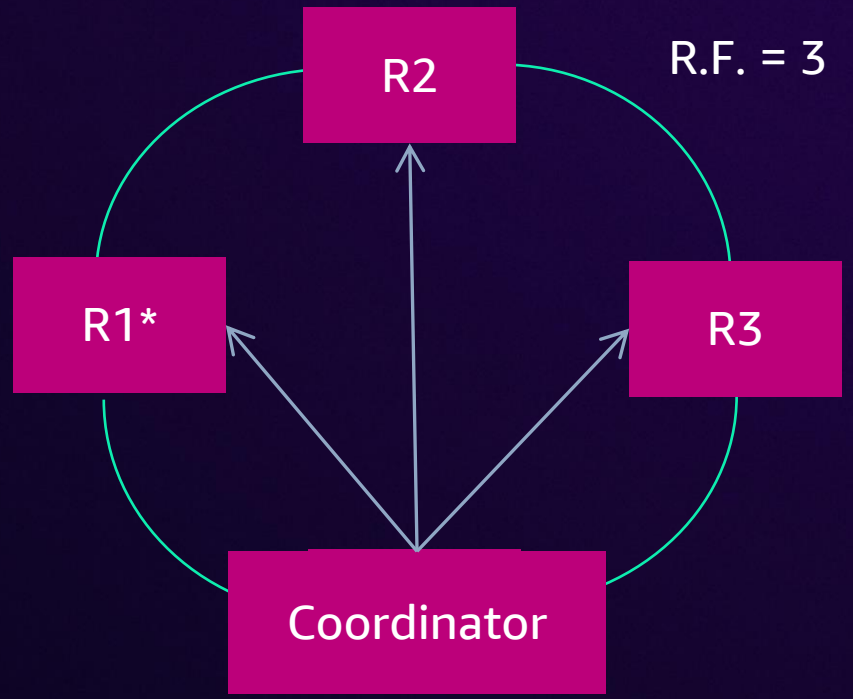
Replication factor

R* = Replica
A replica is a copy of the data.
The R.F. determines the number of replicas

All C* nodes are equal and nodes work in a cluster



Replication factor (R.F.) determines duplication



Recap of key terms

CQL (Cassandra Query Lang)	Drivers (Java, Go, etc…)	Node (a Cassandra instance)
WAL (Write ahead log)	Memtable (in-memory struct)	SSTable (sorted string table)
TTL (Time to live)	Tombstone (marker for delete)	Compaction (Cleanup)
Token ring (collection of nodes)	Cluster (a set of Cassandra instances working together)	Token (64 bit integer – ring position)
Partition Key (row identifier)	Consistent hash (algorithm to minimize data movement)	Replication factor (how many copies)

Cassandra DBA activities

- Cluster management
- Data management
- Performance tuning
- Security and access control
- Capacity planning and scaling
- Disaster recovery
- Monitoring and altering
- Documenting and automation

Amazon Keyspaces (for Apache Cassandra)

Amazon Keyspaces (for Apache Cassandra)

Apache Cassandra–compatible

Use the same Cassandra drivers and tools

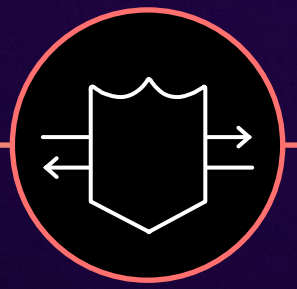
No servers, tombstones, or compaction strategies to manage

No need to provision, configure, and operate large Cassandra clusters

Single-digit-millisecond performance at any scale

Auto scale tables up and down; virtually unlimited throughput and storage

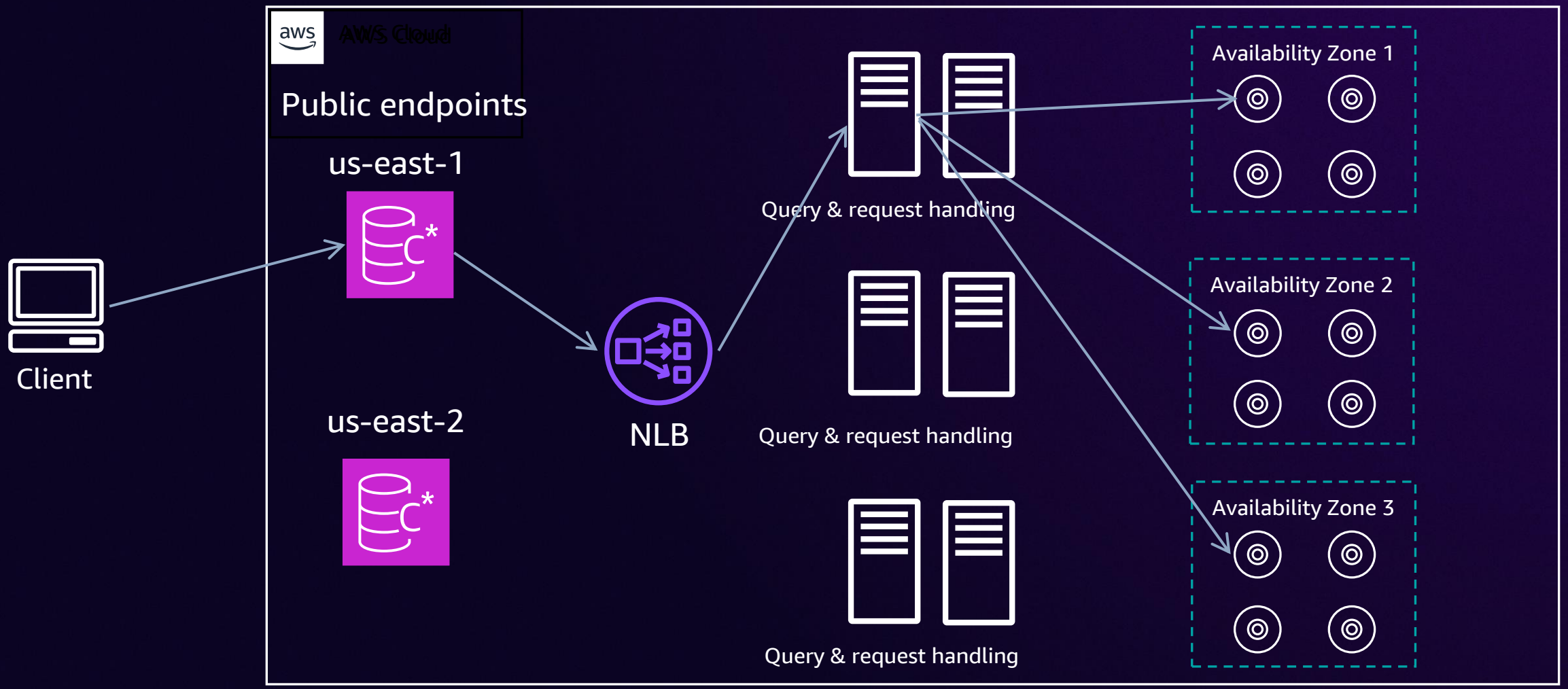
Highly available and secure



99.999% SLA with multi-Region replication; data encrypted at rest

Amazon Keyspaces architecture

Amazon Keyspaces



Key takeaways

- Serverless
- Separated concerns
- Storage scales independently
- Compatibility
- Pay for use

About GE Vernova

WE ARE

GE VERNOVA

For the new era of energy … a new company with full focus on the energy transition

•75K

global employees

•100+

countries

POWER

Gas power, hydro power, nuclear, steam power

WIND

LM wind power, onshore wind, offshore wind,

ELECTRIFICATION

Electrification software, grid solutions, power conversion, solar & storage solutions

ACCELERATORS

Advanced research, consulting services, financial services

GE VERNOVA

- OUR CHALLENGE:

ELECTRIFY THE WORLD WHILE DECARBONIZING IT

GE VERNOVA

- OUR OPPORTUNITY:

DRIVE THE ENERGY TRANSITION FORWARD

GE VERNOVA

Uniquely positioned to help lead the energy transition

ENERGY SOURCE

Nuclear

Gas

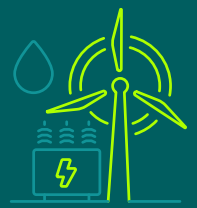
Wind

Solar

Water

GENERATE

Our technology helps to generate approximately 30% of the world's electricity



TRANSFER

Grid solutions to transfer power more reliably and efficiency

ORCHESTRATE

Automated, digital power management systems

CONVERT

Science, systems, and advanced technology to improve efficiencies in power conversion

STORE

Intelligent solar and storage solutions to address the variability of renewable energy supplies

ENERGY DESTINATION

Homes

Business

Industry

GE VERNOVA

OUR GE VERNOVA WAY

THE ENERGY TO CHANGE THE WORLD

*We **drive INNOVATION** in everything we do to **electrify and decarbonize** the world.*

We operate with a founder's mindset. We deliver innovation with passion, speed, and courage. We continuously challenge our thinking and are empowered to dream big and take smart risks.

*We **serve our CUSTOMERS** with pride and a focus on mutual success and long-term impact.*

We listen with humility, and act with urgency. We focus on priorities that create customer value. We anticipate the future and proactively identify solutions for customers and society.

*We challenge ourselves to be better everyday. **LEAN is how we work**.*

We embrace Safety, Quality, Delivery and Cost as our compass. We engage each other to identify and solve problems using data. We balance long-term breakthroughs with continuous improvements that drive sustainable growth.

*We break boundaries and cross borders to **win as ONE TEAM.***

We create an inclusive environment where we can be our authentic selves, respecting, promoting, and valuing our diversity. We support each other's growth. We care about each other and work together to succeed as a global team.

*We are **ACCOUNTABLE individually and collectively** to deliver on our **purpose and commitments**.*

We build trust and credibility by acting with integrity and communicating with transparency. We take ownership in everything we do, continuously learning from our experiences. We uphold our commitment to social responsibility.

WITH THE GE VERNOVA WAY, WE CREATE VALUE FOR OUR PEOPLE, CUSTOMERS, SHAREHOLDERS AND PLANET.

The challenge

Current state

•600 TB

* Data migrated

•300+

* SaaS subscribers

99.999%

* Availability

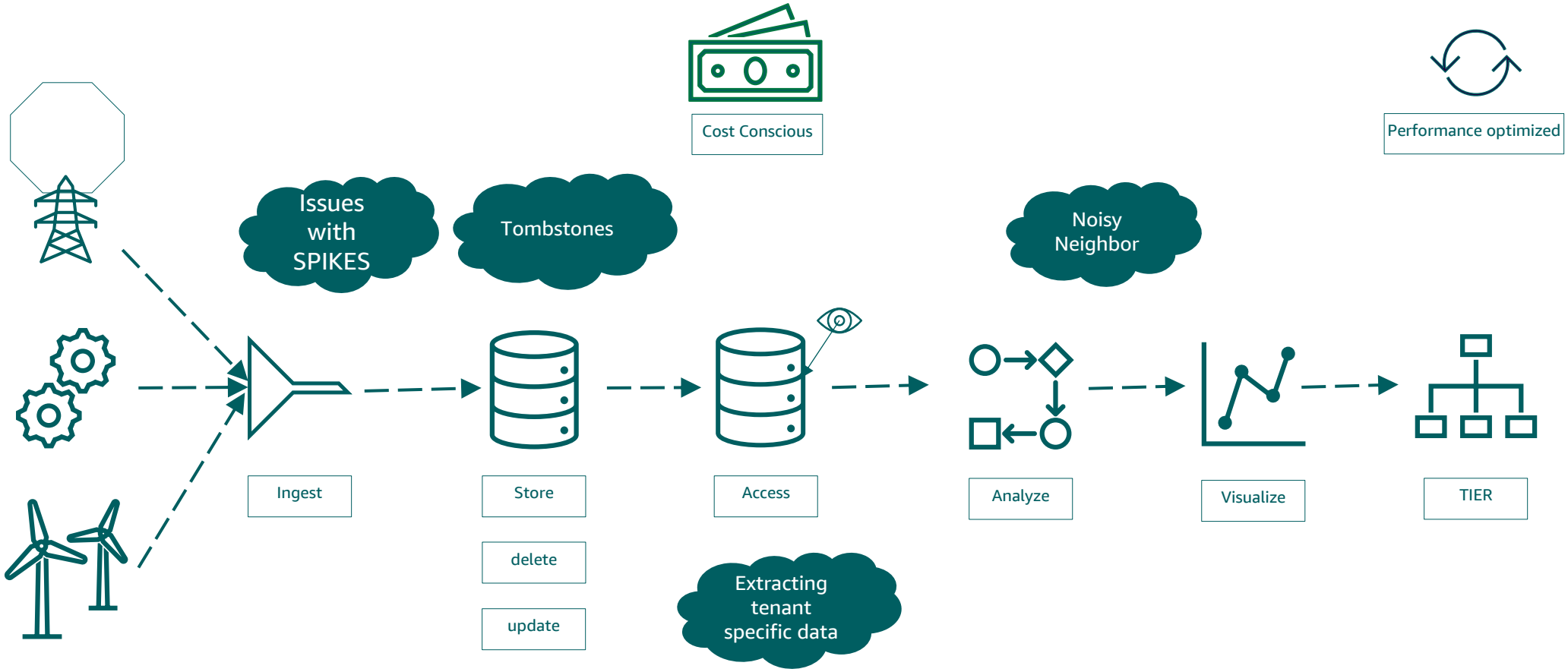
•Challenge

- 6 months timeline
- 300+ SaaS enterprise customers with countless cloud users
- 99.999% availability target throughout the transformation
- Existing system had scaling challenges when reading data
- Reading must be rate-limited to eliminate any impact on operations
- Read from shared model and sort into siloed tenancy model
- Fully backward compatible for query syntax and performance
- Cost savings of ~50%

GE VERNOVA

How our customers consume timeseries services

ACCESS, ANALYZE, VISUALIZE



Why migrate?

- We wanted better cost to performance balance
- Unable to delete stale data easily
- Tombstones made cost run up
- Unable to customize TTL for customers (tenants)
- Noisy neighbor issues in query and ingestion

The lean experiment

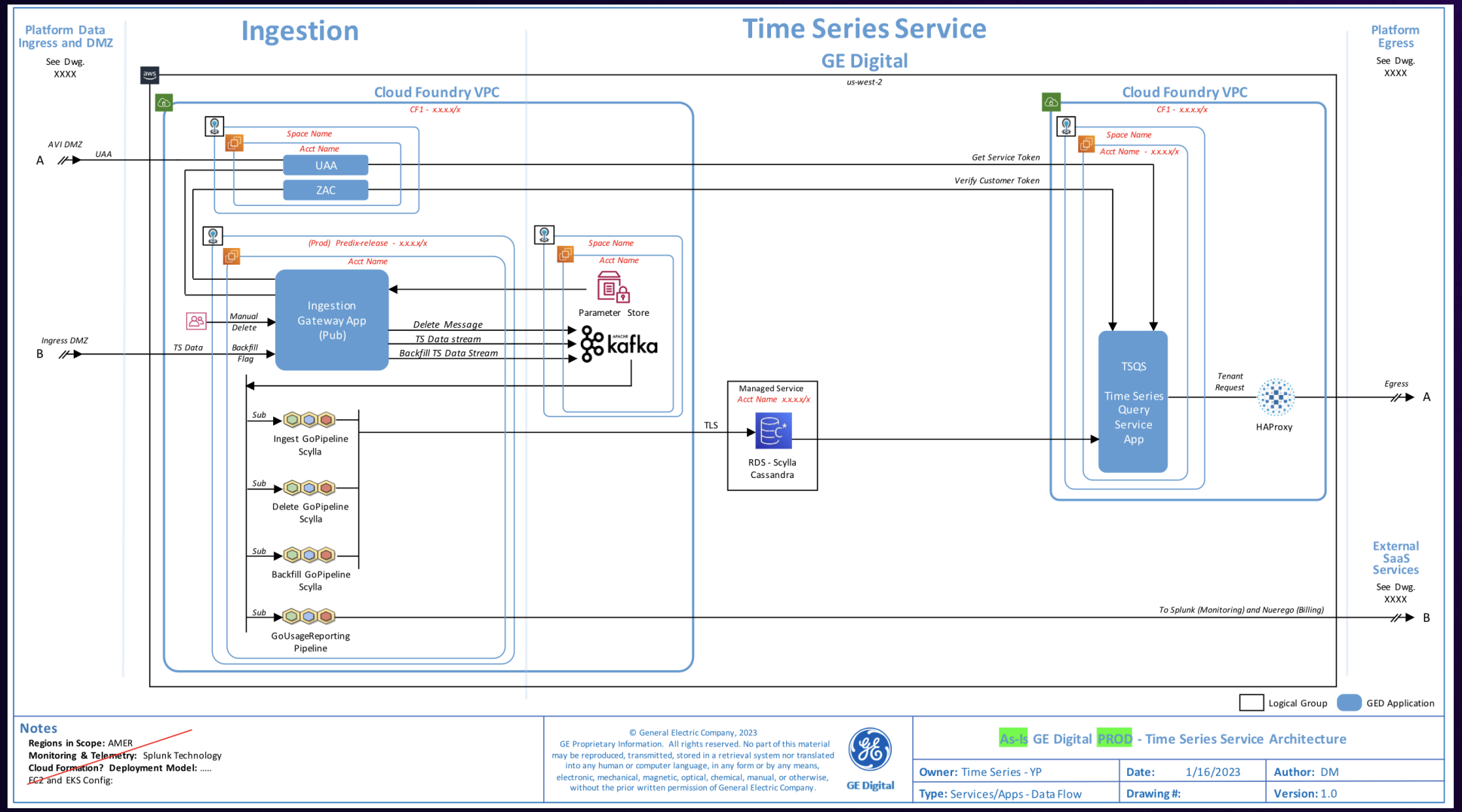
Prior attempts to migrate

- Challenge of impact on live system
- Limits on reads to avoid noisy neighbor
- Out-of-order writes
- Spikes in ingestion/query
- Not enough observability for debugging

Small lean experiment

- Send some test data into Amazon Keyspaces
- We routed one of our integration testing tenants to Keyspaces
- Data was inspected by Proserv peers, SMEs on both sides
- Using excellent observability built on both sides, we looked at the usage metrics
- We took into account cost – we design with economic sense

Architecture before



Learning and planning

- 5x size challenge – unique data model optimized for performance
- Non-uniform reads and writes
- This increased our cost of ingestion by 5 times initially
- Some latency to reads could be seen
- Down the line the cost of deletes would be high too
- We needed to go back to a drawing board

Optimizing in partnership with ProServ

Lean experiments identified a need to optimize storage

Optimizing for size meant optimizing for cost

Team optimized the data model and created a new hash

ProServ and GE product team worked hand in hand in some fantastic collaboration to optimize schema to get most out of storage

MichaelRaney-estimates

Throughput		Provisioned Target Utilization	Request Per Second	Average Hourly Provisioned Capacity Units	Strong Consistency	Eventual Consistent		Request Per Second	Total OnDemand Request	Strong Consistency	Eventual Consistent
	reads	50.00%	66,600.00	133200	14,585.40	7,292.70		7,400.00	175,024,800,000	5,639.69	$2,819.84
	writes	50.00%	40,318.00	80636	44,148.21	44,148.21		4,480.00	105,955,704,000	17,071.49	$17,071.49
Storage		Size In GB	Storage cost								
	Raw Size	425000	127,500.00				60,000	60,000			
	PITR Back up	425000	102,000.00				127,500.00	7,292.70			
							102,000.00	44,148.21			
							14585.4	$2,819.84			
							44148.21	$17,071.49			
							5,639.69	127,500.00			
							17,071.49	102,000.00			
							370,944.79	360,832.24			

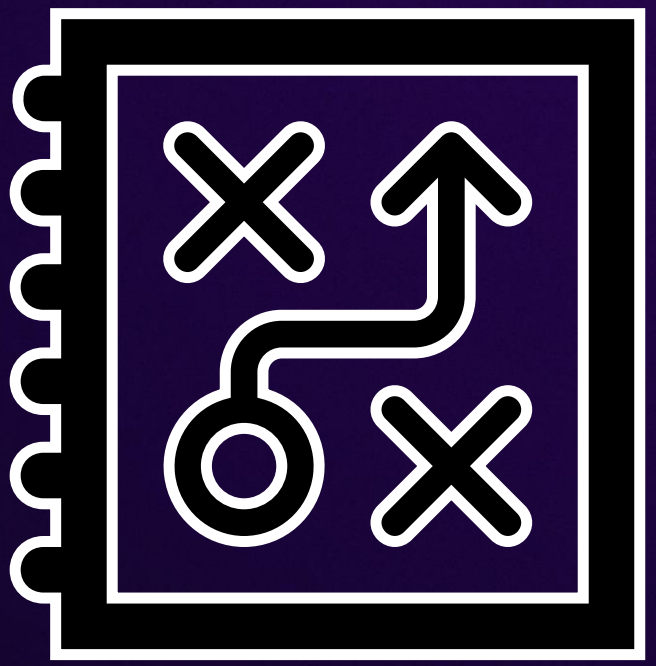
AWS Keyspace Schema optimization

Created by Pylvas, Simo, last modified on Mar 22, 2023

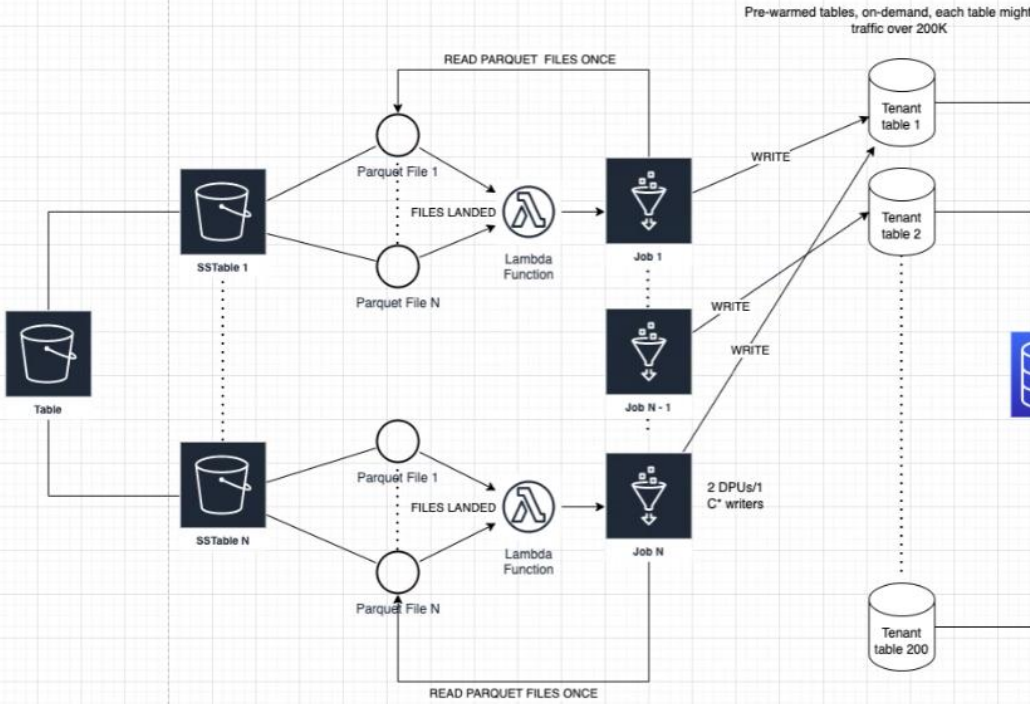
- Option 1. Hashing partition key
- Option 2. Grouping rows by interval with LZ4 compression
- Option 3. Tenant separation
- TTL
- Querying

Key learnings

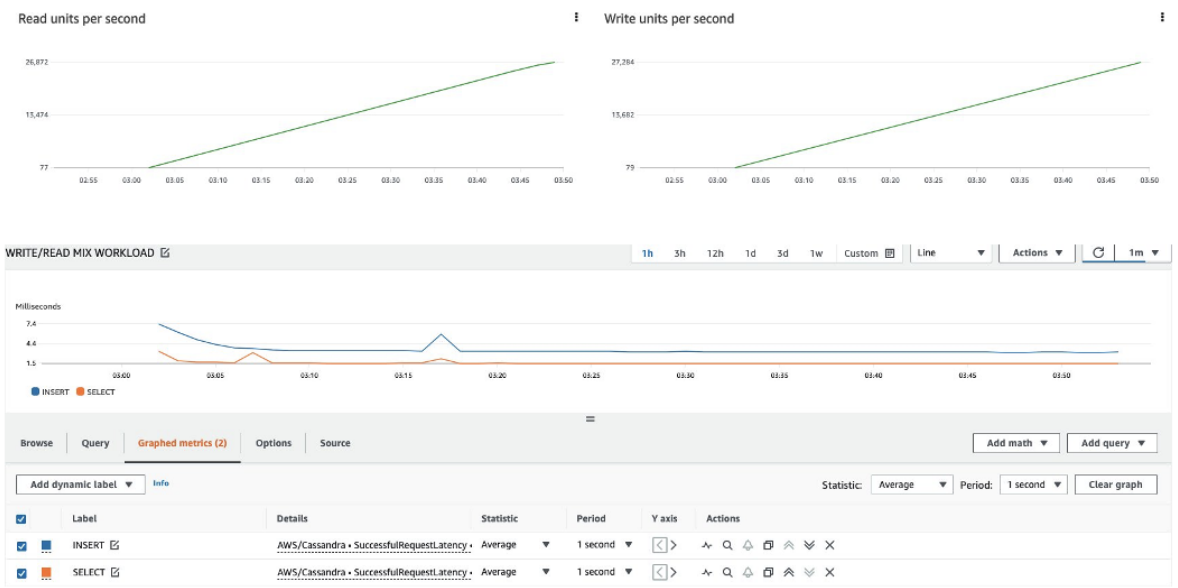
- Start with an experiment
- Always have a plan B
- While you are problem solving also brainstorm a plan B
- Take a breath, take a step back, and find out-of-the-box solutions



Migration architecture – ProServ – Iteration 1

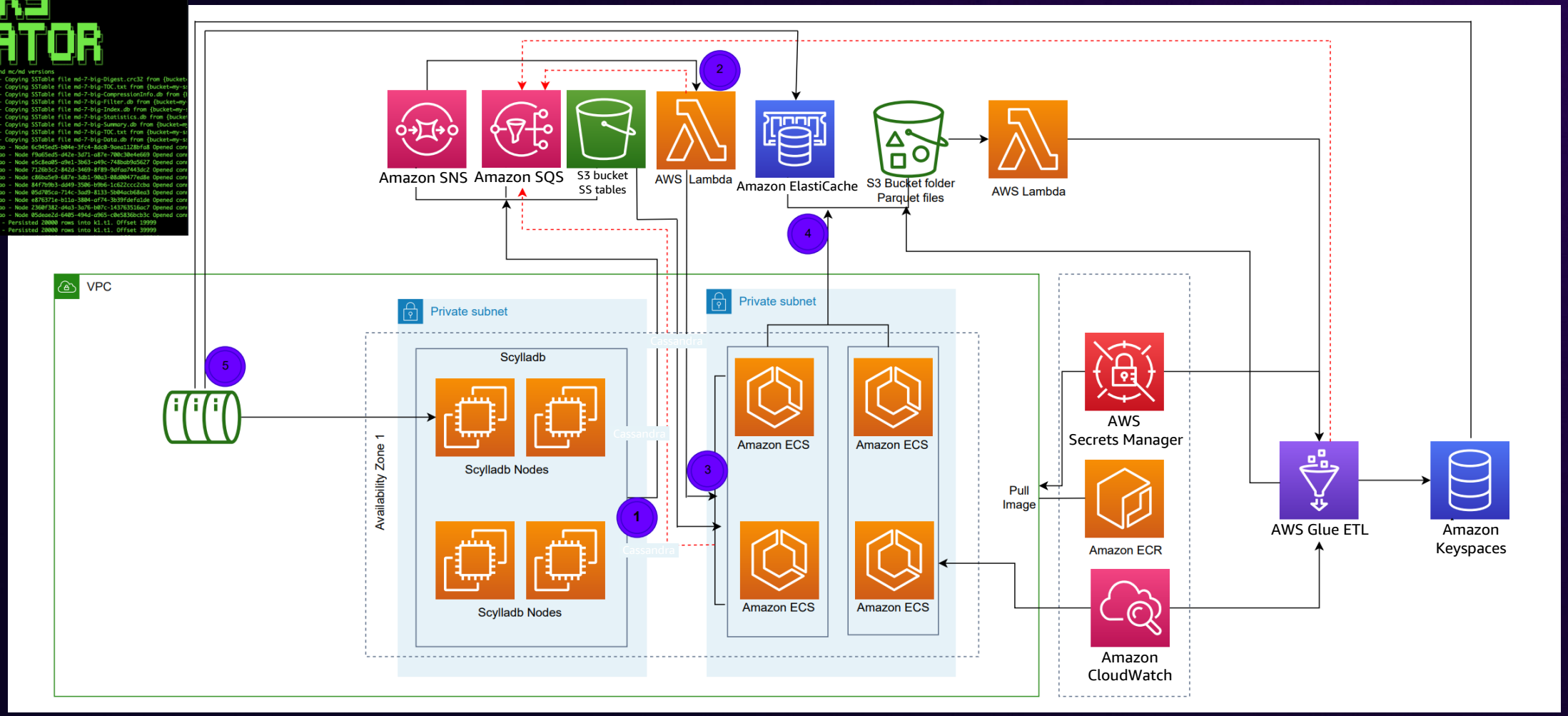


Mixed workload (50%/50% READ/WRITE)

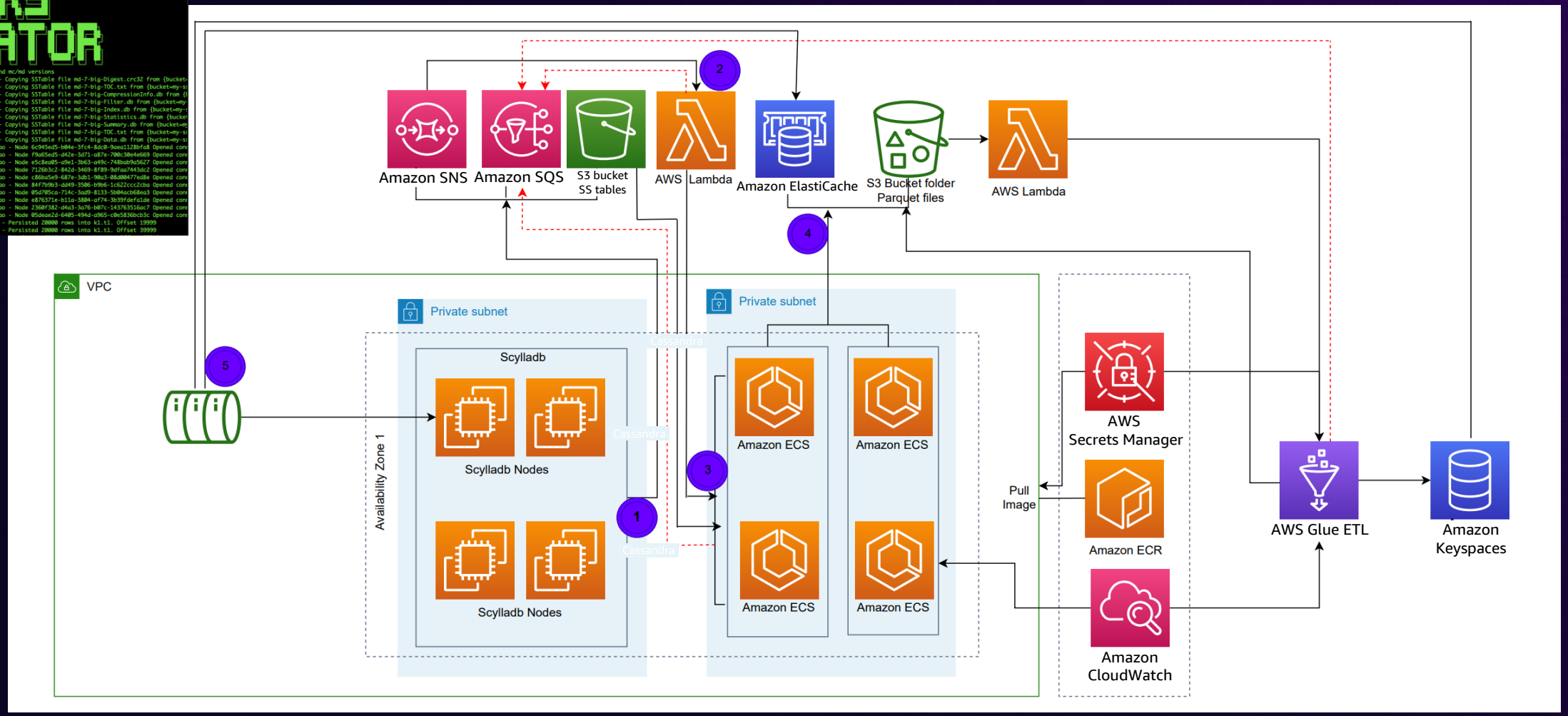


The migration plan and execution

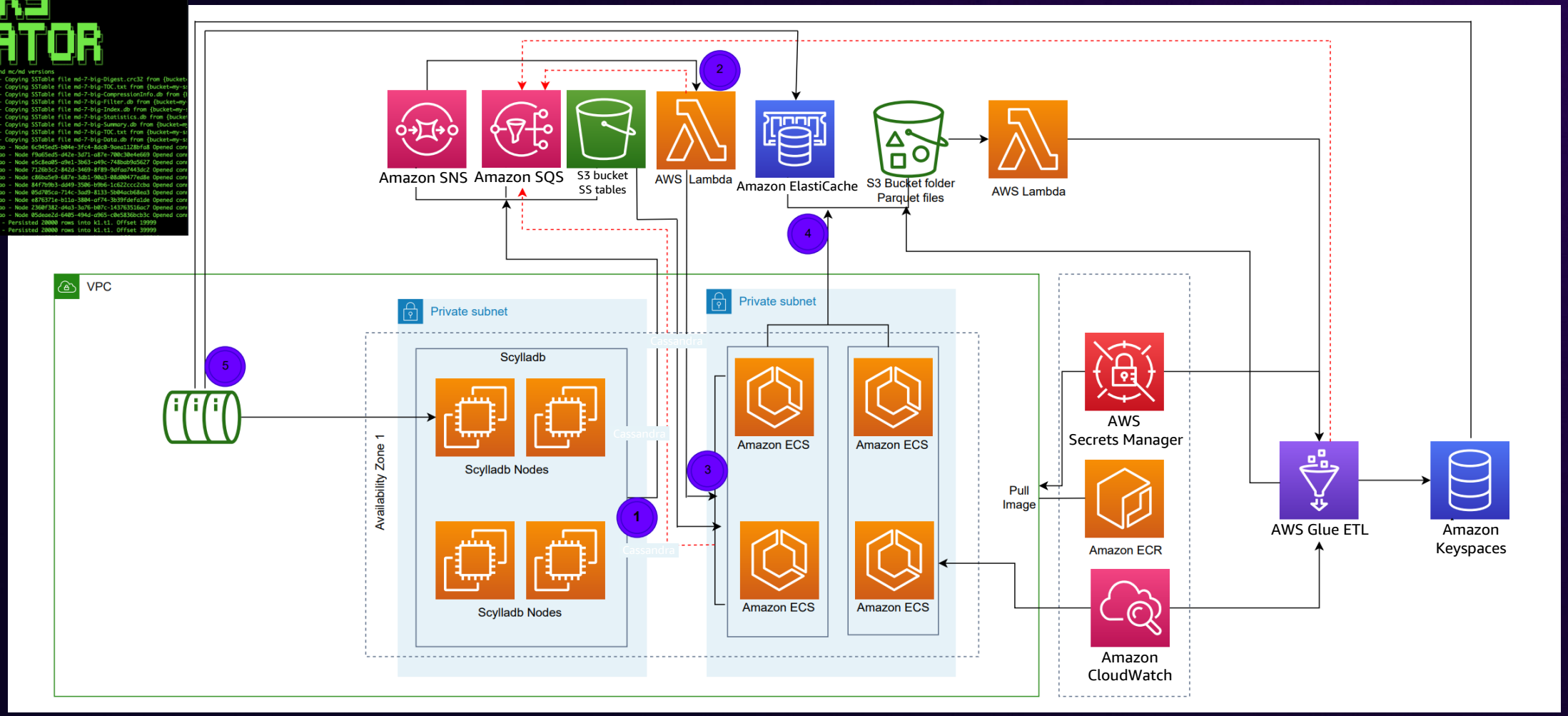
The architecture – Migration utility



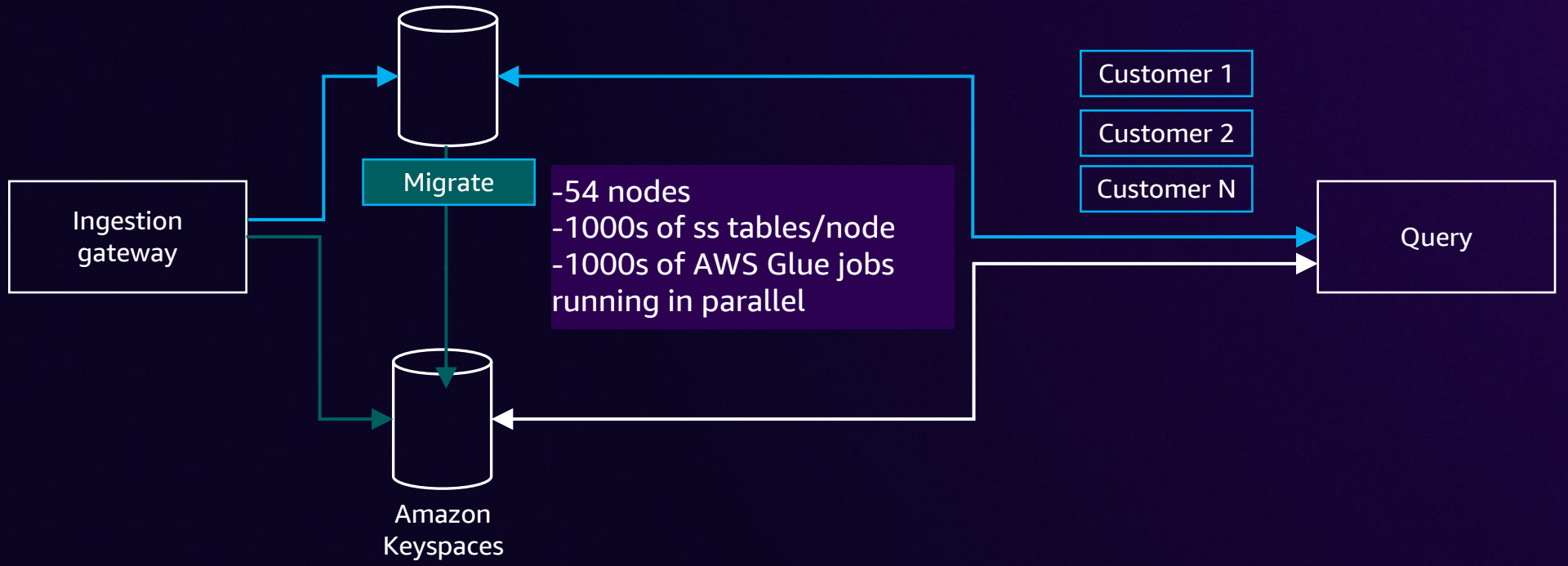
The architecture – Migration utility



The architecture – Migration utility

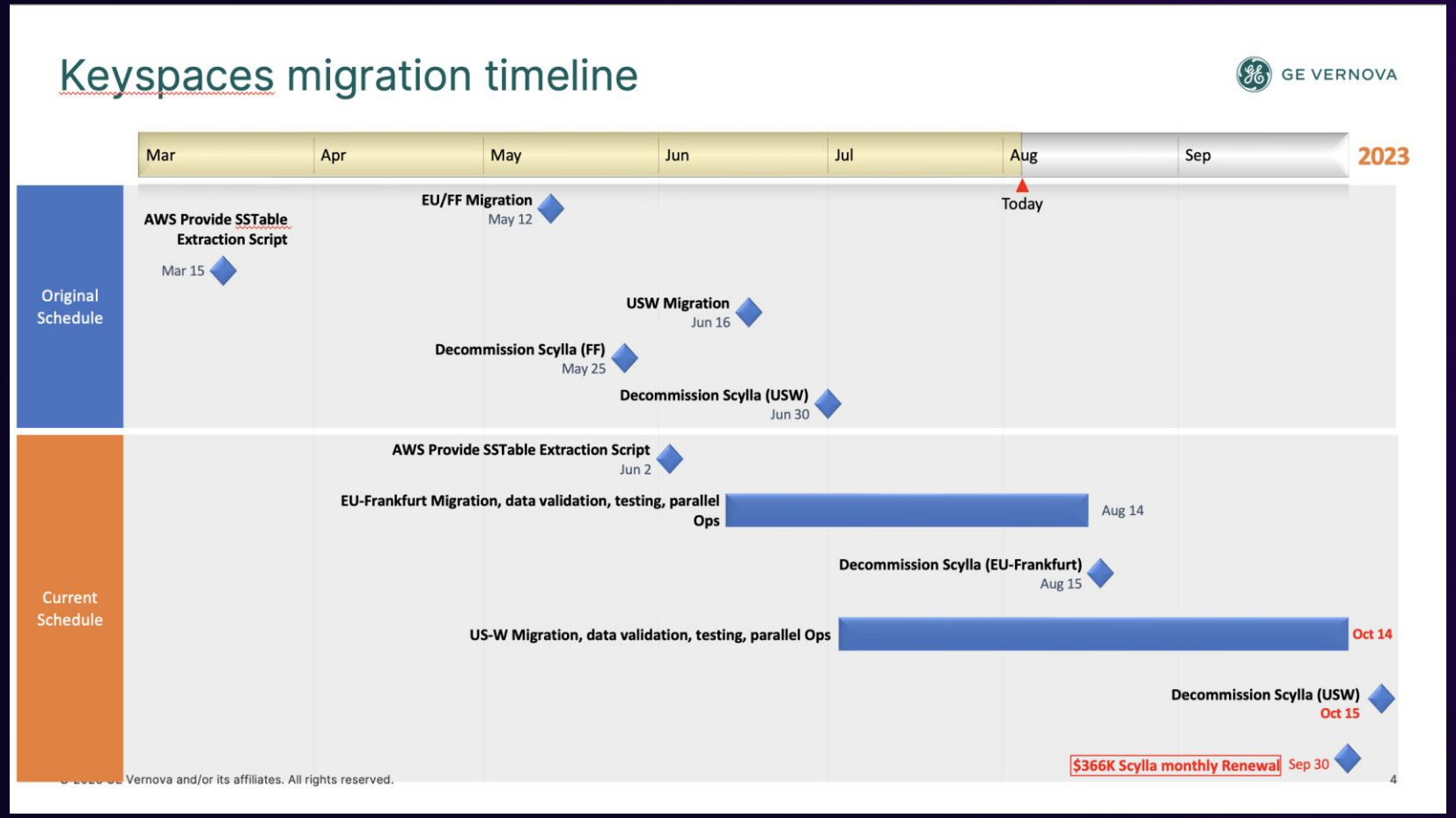


A live swap!



The daily management

Keeping up with Keyspaces schedule



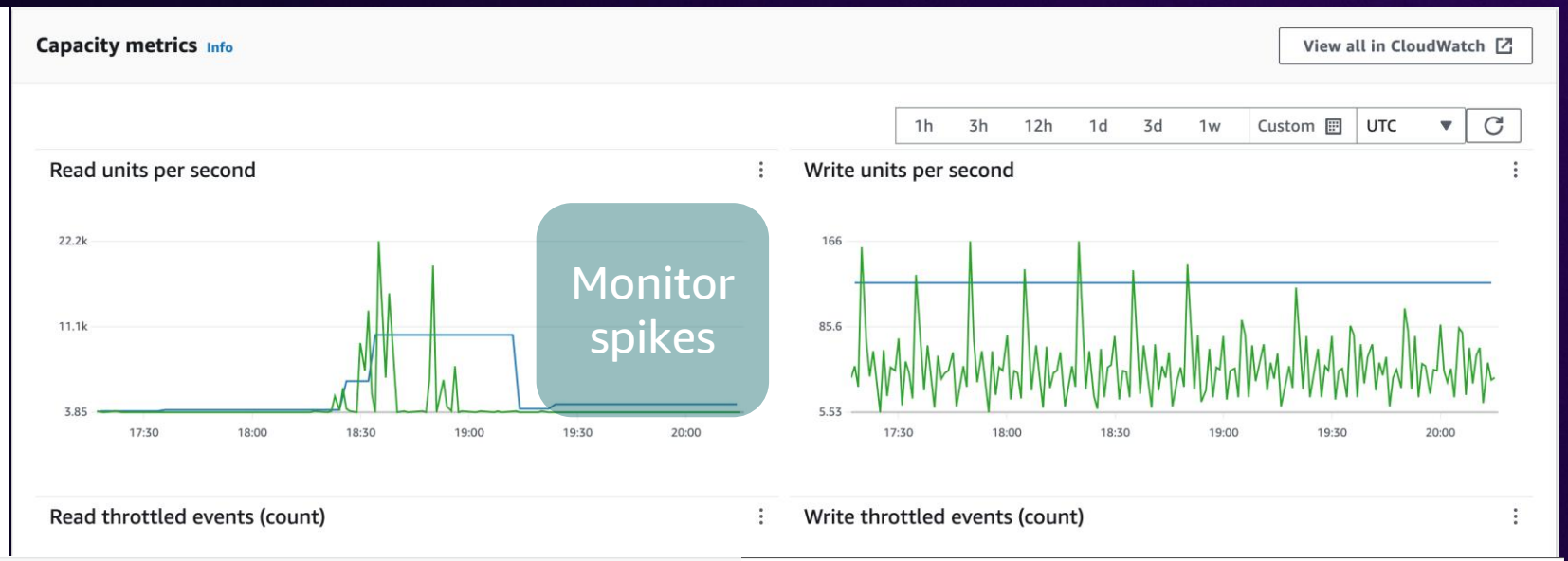
Combined daily management

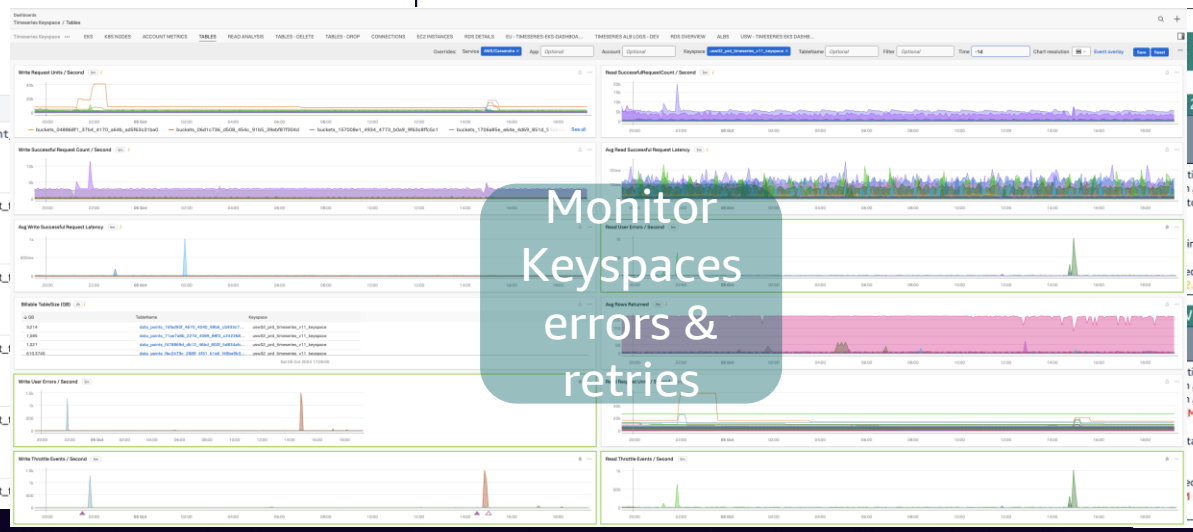
Glue Performance Optimization

Created by Pylvas, Simo, last modified on Jun 14, 2023

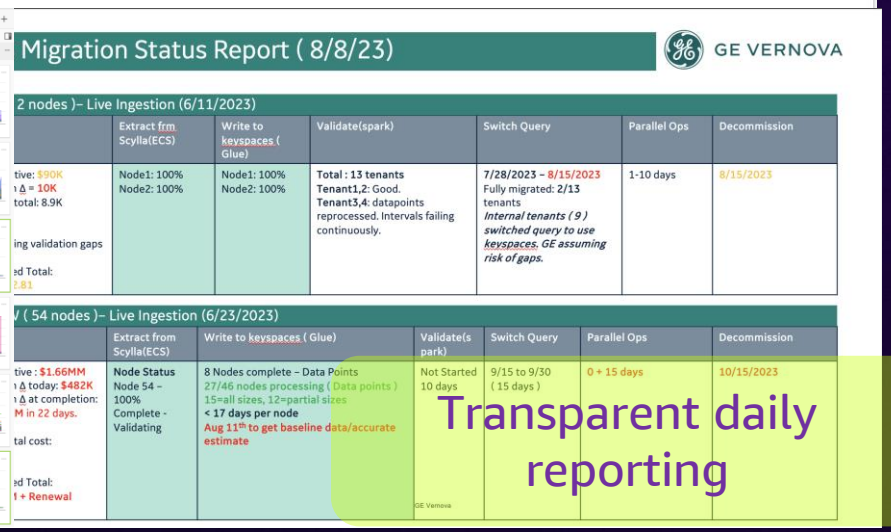
Workers	Worker type	Writers	PDUs	Processing time
5	G1.x	5	5	16-17m
6	G1.x	5	6	13-15m
7	G1.x	5	7	11 - 14m
8	G1.x	4	8	12 -13m
9	G1.x	4	9	~12m
10	G1.x	4	10	11m 45s

Monitor AWS Glue jobs





Monitor Keyspaces errors & retries



Transparent daily reporting

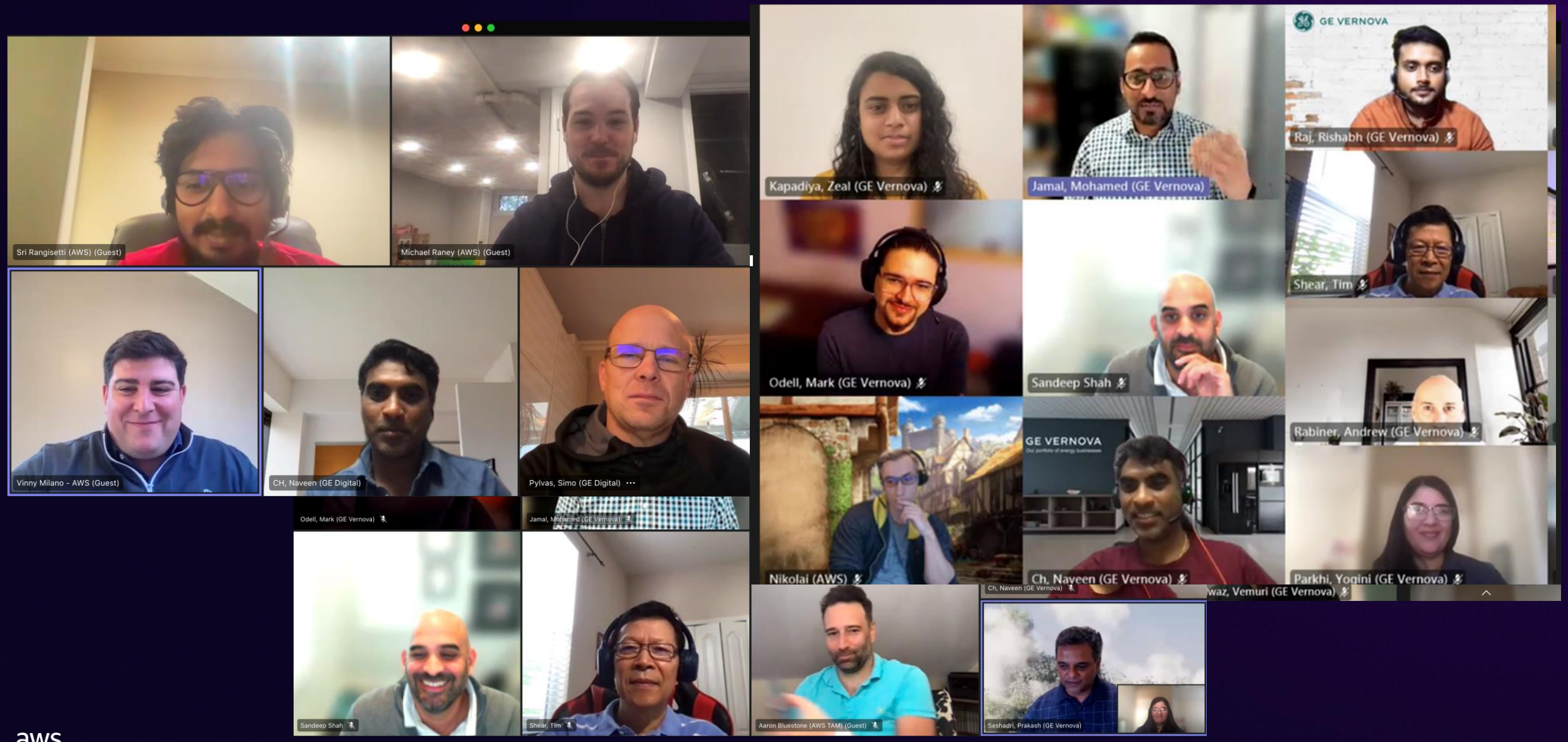
Challenges and solutioning as one team

We found that our data was disparate across customers which created lead time	GE Vernova + Proserv innovated together to read in metadata and use it for intelligent processing
In Europe – We ran into JVM Memory issues	Proserv experimented and found the right JVM configuration to surpass this
In USW we had varying sizes of tables which made for some very large files	GE Vernova + Proserv worked together to break down the workloads by size into XS, S, M, L, XL and process accordingly
We tested limits of elasticity – ran into service limits	AWS Support team jumped in for prompt problem solve to lift soft limits as per our request
We ran into unique challenges with AWS Glue jobs that tried to repeatedly process files	Leadership team helped to mitigate cost of the runaway jobs. GE team wrote a quick Spark job to finish the last remaining workload in one go.

Key learnings

- Expect some twists and turns
- Bake time into roadmap for validation and iteration
- Parallel workstreams can make issues visible early
- Make sure team is feeling like the challenges are something they can solve

Teamwork makes the dream work



Key learnings

- People are key
- Get to know the participants on both sides
- How are they feeling?
- Trust is key to success – Buildup of trust takes diligent work
- Do not get sick! :D Take care of health

The case of the missing data

Missing data?

- GE Vernova team ran validation against the migrated data in old store
- We ran various cross-checks and tried to identify specific tags that seem to have this issue
- I worked with product management and specific customers to determine the root cause
- We found that the data was from time when devices were in error and needed to be discarded

Missing data was never sent to Keyspaces during migration

Key learnings

- If you are migrating historic data – write a comparator utility
- Understand the transformations and have a workstream for validation
- We lost some time because we did not have a validation workstream from the start; we repurposed an existing tool midway

Observability

Observability for cost and monitoring

- The scale makes debugging complex
- Needle-in-a-haystack scenarios are common
- Building out good observability as part of your design and development process is key
- Development with economic sense saved us from discovering runaway cost too late

To finish line

To finish line!

•600 TB

Data migrated

•300+

SaaS subscribers

0

Issues/outages

•Highlights

Just plain beautiful:

- Fully backward compatible query and ingestion
- Zero downtime for customers from the switch
- Monthly usage metrics report per tenant
- PITR backups

Cultural highlights:

- Global engineering and product team working around the clock
- AWS Proserv collaboration
- Building in observability
- Cost-conscious every day
- Lean in action (daily management, small experiments, many pivots, continuous improvement)

GE VERNOVA

Benefits

- Point in time recovery enabled
- Siloed tenancy model helps us optimize for unique customer needs
- Performance benefits
- Spike resiliency due to autoscaling
- Excellent observability for continuous improvement
- Helps GE Vernova team focus on our customers’ unique needs

Wir sind

Nosotros somos

نحن

Мы

CHÚNG TÔI

Nous sommes

LÀ

Είμαστε

我们是

TÁIM

NÓS

SOMOS

우리는

GE VERNOVA

אנחנו

ID

WE ARE

ние сме

私たちです

Wij zijn

MES

Jesteśmy

Mi vagyunk

My jsme

ESAME

اسان آهیون

BIZ

Hum hai

Amra achi

Thank you!

Please complete the session survey in the mobile app

Yogini Parkhi

Director: Analytics Strategy
GE Digital

Steve Mayszak

Sr Manager, Software Engineering
AWS Keyspaces