

DAT332

Why GE Aviation Migrated from Cassandra to Amazon DynamoDB

Corey Green
Senior Director – Software
Engineering
GE Aviation

Ariana Lopez
Staff Software Architect
GE Aviation

aws
re:Invent

© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.



GE Aviation sells \$1.7B annually in parts and services through a **customer portal** called **myGEAviation**. To enhance its customers' experience, the portal enables users to input specific variables and build custom reports for later viewing. **These applications, plotting, and data-query were experiencing issues with cost, scalability, and performance.** In this session, GE discusses how its re-architected plotting and data-query to resolve those issues by using Amazon DynamoDB.

Agenda

Who is GE Aviation?

Legacy database solutions

Transitioning to DynamoDB

Outcomes and Lessons Learned

Who is GE Aviation?

If you flew to Las Vegas there is a 2 out of 3 chance a GE engine helped

Every 2 seconds an aircraft with GE engine technology takes off

2200 aircraft carrying 300,000 people in the air right now with GE engine technology



GE Aviation's Customer

60,000 installed engines

300+ airlines who operate these engines

Airlines use our external sites (Portals) for a variety of functions

Functionality allows airlines to ensure reliable and safe operation of engines



Engine Performance Data

Engines report back sensor data to be parsed and made available

Customers plot and download sensor data

Over 40TB (175B rows) of data on demand

Expectation of <50ms query response time

Data continues to grow in density



How do customers use data: time series plots



These are sensor data from an engine plotted against time, this is a primary functionality of our customer portal

How do customers use data: reporting

Contractual and government regulations require access to data for reporting

Data query functionality provides the customer ability to download data on demand

Reports can be many GBs of data packaged for download

New Report Query

▼ Report Fields

Display: Description ▼

Category: Public - Standard Parameters

CCC Selection Status →

Data EGT →

Thrust Derate →

Thrust Derate Smoothed →

Data Oil Pressure Smoothed →

ETDPS EGT Margin - Climb →

EGT Hot Day Margin Smoothed →

EGT Thermocouple Inner Lower Case →

Default Columns

Operator

Aircraft Type

Aircraft Tail

Engine Family

Engine Type

ESN

Flight Phase

Flight Datetime

Engine Install Date

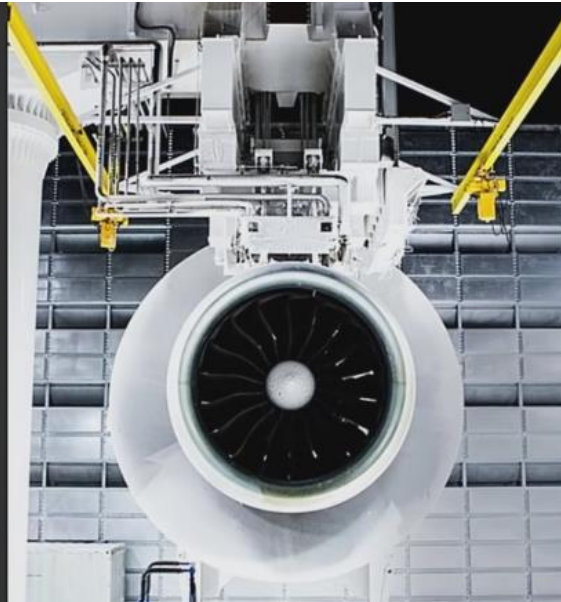
▼ Query Criteria - Default Columns

Parameter	Expression	Value
Operator	=	[JAMS] [AEROS]
Aircraft Type	=	[A3 Available Selected - Search and select to filter]
Aircraft Tail	=	[A3 Available Selected - Search and select to filter]
Engine Family	=	[A3 Available Selected - Search and select to filter]
Engine Type	=	[A3 Available Selected - Search and select to filter]
Installed / Uninstalled Engines	=	Installed
ESN	=	[A3 Available Selected - Search and select to filter]
Flight Phase	=	[A3 Available Selected - Search and select to filter]
Date Range	Last N...	Enter # of days ...

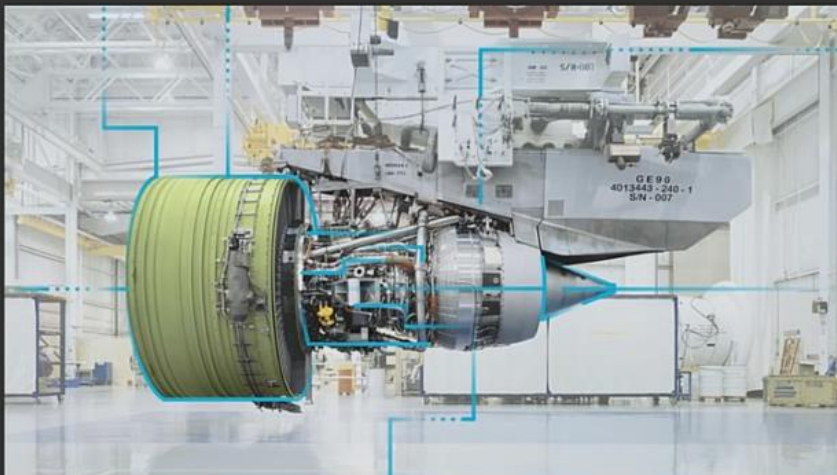
We also use our portal to provide reporting for an airline on the entire history of their data, they can also download the data if they wanted using the portal.

Why move to AWS?

- Ability to manage scale
- Ability to manage cost
- Flexibility to look at new technologies



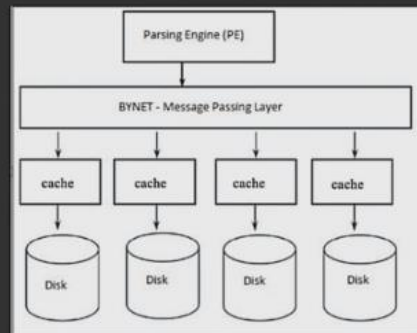
Legacy Database Solutions



Legacy database solution #1: Traditional data warehouse

Issues:

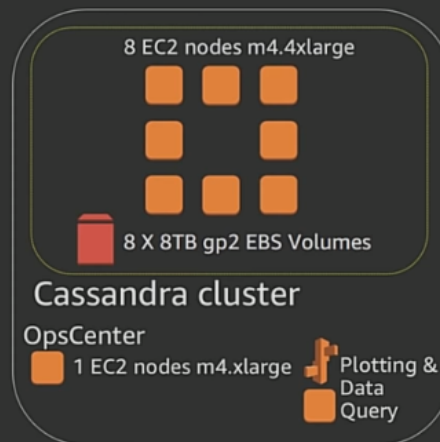
- Performance issues
- Monolithic
- Expensive
- Didn't own the technology



Legacy database solution #2: Cassandra

Benefits of Cassandra

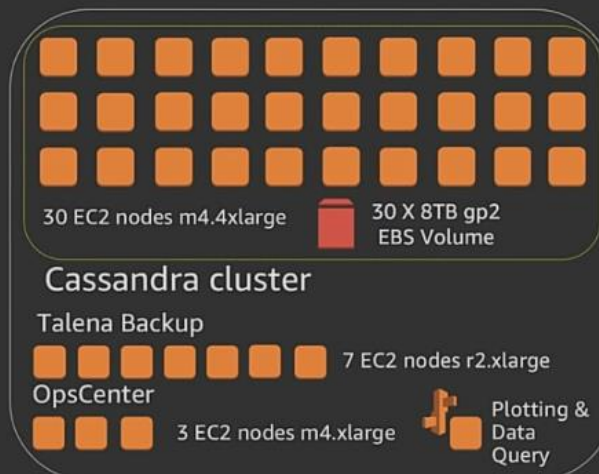
- Flexibility
- Scalable performance
- Faster reads & writes
- Autonomy



Legacy database solution #2: Cassandra

Challenges of Cassandra

- Hot spots
- Expensive
- High overhead
- Query latency
- Architect for max volume



Transitioning to DynamoDB

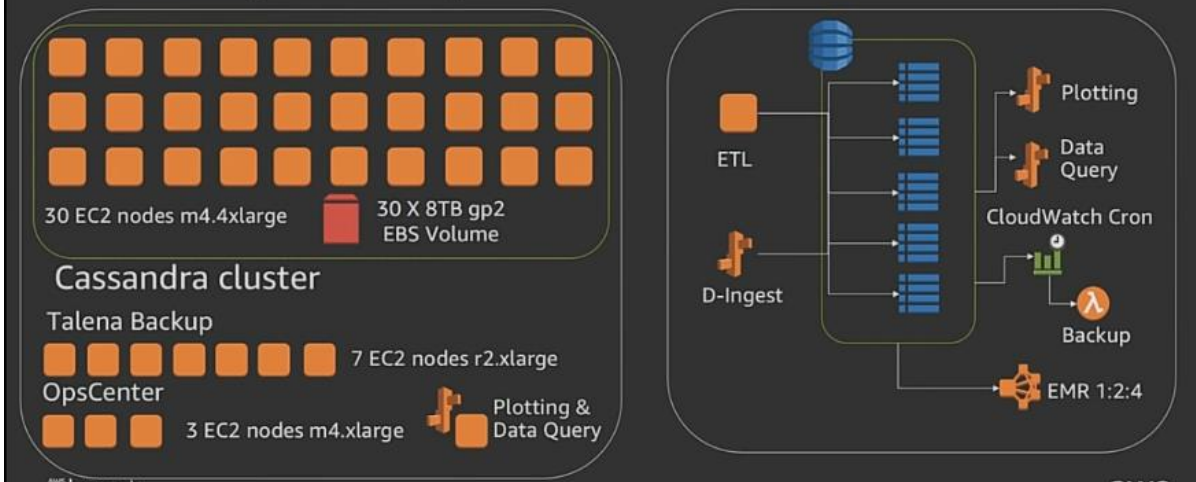
Key DynamoDB selection criteria

- Capacity Autoscaling (reads and writes)
 - Storage optimization
- Low overhead
 - Reduced admin hours
 - No compliance issues
 - No OS patching, Chef
- Smaller footprint



Transitioning to DynamoDB

Cassandra vs. DynamoDB footprint

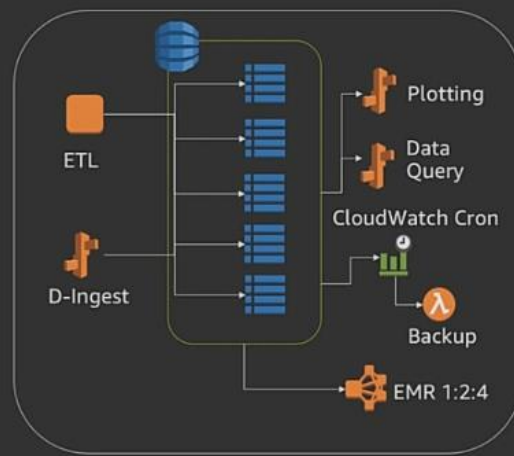


We replaced the Cassandra cluster architecture with 5 DynamoDB tables consisting of a time series table, a flight report table that contains summarized computed results for any flight, then we have 3 master tables.

Transitioning to DynamoDB

Realized benefits

- DynamoDB workload scalability
- DynamoDB storage scalability
- DynamoDB maintenance overhead

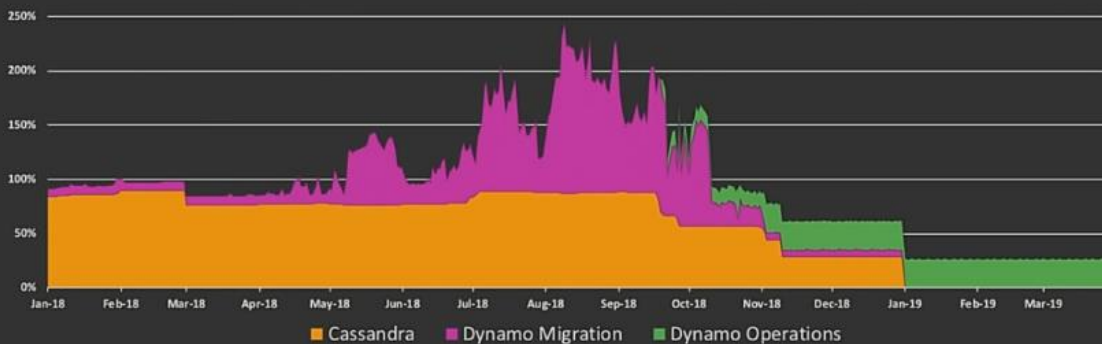


DynamoDB Outcomes

- Better application response
 - 11ms-16ms query response
- Code improvements delivered
 - Streamline reporting code
- Backlog of continued improvements
 - Use of other storage facilities + Athena



Outcome – 72% cost reduction



2018 start...

- Cassandra scaled for peak
- Overhead for backups and DBAs
- Incremental scaling planned

...2018 Migration...

- EC2 ETL servers
- Build ingestion process
- Ingestion at max WCUs

...2018 Completion into 2019

- Run rate 72% lower
- All aspects scalable on demand
- Visions of further efficiencies

Needs + lessons learned

- Data Access
- Autoscaling
- Data loading
- Backup & Point-In-Time Recovery
- AWS is moving fast



Thank you!

Corey Green
corey@qe.com

Ariana Lopez
Ariana.Lopez@qe.com

aws
re:Invent

© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

