Table of Contents

Announcements

- Analytics
- Artificial Intelligence & Machine Learning
- Business Applications
- Compute and Containers
- Data
- Front End Web & Mobile Development
- Serverless

Index

Index

AI / ML

- ❑ AIM202 – Automate your mortgage document processing with AWS AI/ML
- ❑ AIM203 – Boost user engagement with Amazon Personalize
- ❑ AIM204 – Automate insurance document processing with AI
- ❑ AIM207 – Make better decisions with no-code ML using SageMaker Canvas, feat. Samsung
- ✓ AIM208 – Idea to production on Amazon SageMaker, with Thomson Reuters
- ✓ AIM210 – Solve common business problems with AWS AI/ML services
- ❑ AIM214 – Learn how Black Knight is using AI to accelerate mortgage workflows
- ✓ AIM307 – JPMorganChase real-time agent assist for contact center productivity
- ❑ AIM313 – Scalable data preparation & ML using Apache Spark on Amazon SageMaker Studio
- ✓ AIM314 - Accelerate your ML journey with Amazon SageMaker low-code tools
- ❑ AIM319 – Build, manage, and scale ML development with a web-based visual interface
- ❑ AIM320 - Boost ML development productivity with managed Jupyter notebooks
- ✓ AIM321 – Productionize ML workloads using Amazon SageMaker MLOps, feat. NatWest
- ✓ AIM332 – WaFd bank delivers enhanced self-service with AWS conversational AI
- ❑ AIM333 – How Taco Bell is improving digital availability with ML forecasting
- ❑ AIM337 – Build better ML models for business decisions using Amazon SageMaker Canvas

Analytics

- ✓ ANT201 – What's new with Amazon Redshift
- ❑ ANT204 – Enabling agility with data governance on AWS
- ✓ ANT205 – Achieving your modern data architecture
- ✓ ANT207 – What's new in Amazon Streaming
- ✓ ANT208 – What's new in Amazon Athena
- ❑ ANT209 – Build interactive analytics applications
- ❑ ANT212 – How JPMC and LexisNexis modernize analytics with Amazon Redshift
- ❑ ANT220 – Using Amazon AppFlow to break down data silos for analytics and ML
- ✓ ANT223 – Simplify and accelerate data integration & ETL modernization with AWS Glue
- ❑ ANT301 – Democratizing your organization's data analytics experience
- ✓ ANT302 – What's new with Amazon EMR
- ❑ ANT317 – Self-service analytics with Amazon Redshift Serverless
- ❑ ANT324 - Modernize your data warehouse
- ❑ ANT325 – How Disney processes clickstream data on Amazon EMR Serverless
- ❑ ANT326 – How BMW, Intuit & Morningstar are transforming with AWS and Amazon Athena
- ❑ ANT335 – How Disney used AWS Glue as a data integration and ETL framework
- ✓ ANT336 – Building data mesh architectures on AWS
- ❑ ANT337 – Migrating to Amazon EMR to reduce costs and simplify operations
- ❑ ANT339 – How Samsung modernized architecture for real-time analytics
- ❑ ANT340 – How Sony Orchard accelerated innovation with Amazon MSK
- ❑ ANT341 – How Riot Games processes 20 TB of analytics data daily on AWS
- ❑ ANT342 – How Poshmark accelerates growth via real-time analytics & personalization

Data

- ✓ DAT207 – Choosing the right databases for your application
- ❑ DAT211 – How Capital One accelerates innovation with AWS databases
- ❑ DAT212 – How ADP & Twilio realize business vision with purpose-built databases
- ❑ DAT213 – How Fidelity Investments and Reltio modernized with Amazon DynamoDB
- ❑ DAT217 – Modernize apps with purpose-built databases, featuring Liberty Mutual
- ✓ DAT218 – What's new with database migrations
- ✓ DAT219 – What's new with Amazon DocumentDB
- ❑ DAT221 – Dive deep on AWS managed PostgreSQL databases
- ❑ DAT319 – S&P Dow Jones Indices' AWS journey to Amazon Aurora
- ❑ DAT320 – Deploy modern and effective data models with Amazon DynamoDB
- ❑ DAT325 – Managed Oracle Databases with Amazon RDS: New features & best practices
- ❑ DAT326 – Deep dive into Amazon Aurora and its innovations

< Table of Contents

Compute & Containers

- ✓ CMP225 – What's new in Amazon EC2
- ❑ CMP310 – Best practices for growing your business on Amazon Lightsail
- ✓ CON210 – What's new and what's next with Amazon ECS
- ❑ CON309 – Build your application easily & efficiently with serverless containers
- ❑ CON325 – Building containers for AWS
- ❑ CON326 – Are you ready? Essential strategies for Kubernetes adoption
- ❑ CON327 – Serving 4.8 billion railway customers a year: The story of Deutsche Bahn
- ✓ COM207 – Competition of the modern workloads: Serverless vs. Kubernetes on AWS
- ✓ COM301 – Unleash developer productivity with infrastructure from code
- ❑ STG303 – AWS storage for containers and serverless applications

Serverless

- ✓ SVS204 – What's new with serverless
- ❑ SVS209 – Get started building your first serverless, event-driven application
- ❑ SVS302 – Architecting secure serverless applications
- ✓ SVS303 – Building real-world serverless applications with AWS SAM
- ✓ SVS308 – AWS serverless developer experience: A day in the life of a developer
- ❑ SVS311 – Practical experience with a serverless-first strategy at Capital One
- ❑ SVS312 – Building Serverlesspresso: Creating event-driven architectures
- ❑ SVS401 – Best practices for advanced serverless developers
- ❑ SVS404 – A closer look at AWS Lambda
- ❑ TNC108 – Introduction to serverless computing and AWS Lambda

DevOps

- ✓ DOP201 – AWS infrastructure as code: A year in review
- ✓ DOP202 – Building on AWS with AWS tools, services, and SDKs
- ✓ DOP205 – Build it & run it: Streamline your DevOps capabilities with machine learning
- ✓ DOP206 – Continuous integration and delivery (CI/CD) on AWS
- ❑ DOP207 – AWS Well-Architected best practices for DevOps on AWS
- ❑ DOP301 – The Amazon Builders' Library: 25 years of operational excellence at Amazon
- ❑ DOP302 – How to reuse patterns when developing infrastructure as code
- ❑ DOP314 – Governance and security with infrastructure as code
- ❑ DOP317 – Accessing AWS from your development environment

Front End Web & Mobile Development

- ✓ FWM308 – What's new for frontend web and mobile developers with AWS Amplify
- ✓ FWM309 – Design to code: Build an app frontend UI in minutes with AWS Amplify Studio
- ❑ FWM310 – Building a multi-Region serverless application with AWS AppSync
- ✓ FWM315 – Create real-time, event-driven apps with Amazon EventBridge & AWS AppSync
- ❑ FWM316 – Accelerate GraphQL API app development & collaboration with AWS AppSync

.NET on AWS

- ❑ XNT303 – Building AI-enriched applications with .NET on
- ❑ XNT304 – Purpose-built databases for modern .NET applications on AWS
- ❑ XNT402 – Cloud-native .NET applications on AWS

- ❑ FWM315 – Create real-time, event-driven apps with Amazon EventBridge & AWS AppSync
- ❑ FWM316 – Accelerate GraphQL API app development & collaboration with AWS AppSync

Application Integration

- ❑ API303 – Application integration patterns for microservices
- ❑ API307 – Designing event-driven integrations using Amazon EventBridge
- ❑ API308 – Are you integrating or building distributed applications?
- ❑ API309 – Advanced serverless workflow patterns and best practices
- ❑ API311 – Building next-gen applications with event-driven architectures
- ❑ API314 – What's possible with AWS Step functions.

< Table of Contents

Architecture

- ❑ ARC205 – Establishing your cloud foundation on AWS—beyond your landing zone
- ❑ ARC207 – Modern cloud applications: Do they lock you in?
- ❑ ARC306 – Multi-Region design patterns and best practices
- ✓ ARC311 – Build production-ready prototypes rapidly using serverless patterns
- ✓ ARC313 – Building modern data architectures on AWS
- ❑ ARC314 – Publishing real-time financial data feeds using Kafka
- ❑ ARC315 – How to understand and improve resiliency for serverless systems

- ❑ ENT213 – How to migrate, modernize, and grow using the AWS MAP
- ❑ ENT221 – Building a smarter organization powered by machine learning
- ❑ ENT311 – Migrate and modernize Oracle E-Business Suite on AWS in 5 steps
- ❑ HYB311 – How to simplify your hybrid cloud migration strategy with AWS

Build on AWS / With Purpose

- ❑ ALX201 – 10 things every voice application should do

- ❑ BSI205 – Build stunning customized dashboards with Amazon QuickSight
- ❑ BSI207 – Get clarity on your data in seconds with Amazon QuickSight Q
- ✓ BIZ213 – Serve your customers better with AWS Supply Chain

- ❑ BIZ201 – Improving multichannel communications

- ❑ BOA203 – Evolution of the machine learning development environment
- ❑ BOA309 – Build a managed analytics platform for your ecommerce business
- ❑ BOA310 – Building observable applications with OpenTelemetry

- ❑ BWP102 – The story of headless ecommerce and the future of online storefronts
- ❑ BWP201 – Build with prototypes: Be a customer-obsessed developer
- ❑ BWP302 – How Alloy synced Amazon data in real time

Enterprise & Industry

- ❑ FSI201 - Goldman Sachs: Accelerating time to value in data analytics
- ❑ FSI304 – Your blueprint to secure customer-facing applications on AWS
- ❑ PRT303 – The digital transformation journey: Santander UK
- ❑ PRT222 – Why operationalizing data mesh is critical for operating in the cloud (sponsored by Capital One)
- ❑ SAS305 – SaaS architecture patterns: From concept to implementation

< Table of Contents

Announcements

< Table of Contents

Analytics

Glue

Glue is a serverless, scalable data integration service used to discover, prepare, move, and integrate data from multiple sources.

- *Version 4.0*
  Updated Engines and More Formats, plus native support for the Cloud Shuffle Service Plugin for Spark and Pandas support.
  Glue for Ray
  Data engineers can use Glue for Ray to process large datasets with Python. Ray (ray.io), is a popular new open-source compute framework that helps you scale Python workloads.

OpenSearch

OpenSearch enables you to easily ingest, secure, search, aggregate, view, and analyze data for a number of use cases such as log analytics, application search, enterprise search, and more.

- *OpenSearch Serverless*
  Run large-scale search and analytics workloads without managing clusters. Pay only for the data that is ingest and the queries you run. AWS now has a complete "Serverless" offering for all of its Analytics Services.

QuickSight

- *Paginated Reports*
  Adds a missing capability (for a cost) to print/display/distribute paginated reports.
- *New & Updated APIs*
  Allows programmatic creation and management of dashboards, analysis, and templates.

Redshift

- *Integration with Apache Spark*
  Builds upon an existing open-source connector project and enhances it for performance and security, helping customers gain up to 10x faster application performance.
- *Ingestion from Kinesis and MSK Streams*
  Streaming ingestion benefits from the ability to optimize query performance with materialized views and allows the use of Amazon Redshift more efficiently for operational analytics and as the data source for real-time dashboards.

Artificial Intelligence / Machine Learning

SageMaker

Model Jumpstart

- Share ML Models and Notebooks More Easily Within Your Organization
  An API for programmatically analyzing and classifying the documents contained in mortgage loan application packages and extracting the data they contain.

Studio

- Built-in integration with Amazon EMR
  Starting today, you can now monitor and debug your Apache Spark jobs running on EMR right from SageMaker Studio notebooks with just a click.

Textract

- *Analyze Lending API*
  An API for programmatically analyzing and classifying the documents contained in mortgage loan application packages, and extracting the data they contain

Transcribe

- *Real-Time Analytics During Live Calls*
  Provides APIs for developers to accurately transcribe live calls and at the same time identify customer experience issues and sentiment in real time. Sessions: AIM307

Developer Productivity

CodeWhisperer

- *Support for new languages, administrative controls, and simple signup*
  Support for C# and TypeScript projects. Allows administrators the ability to enable CodeWhisperer for their organization with Single Sign-On authentication.

Security

Security Lake

New Service. Automatically centralizes your organization's security data from cloud and on-premises sources into a purpose-built data lake stored in your account.

Compute

EC2

- ENA (Elastic Network Adapter) Express
  Reduces latency of traffic flows while also increasing maximum bandwidth. You get a lot more per-flow bandwidth and a lot less variability with just a checkbox.
- *New Instance Types for Standard, Network, HPC, and Deep Learning Workloads*
  More compute server types for faster and specialized workloads. AWS specializes in raw compute instances.

ECS

- Service Connect
  Add a layer of resilience to your ECS service communication and get traffic insights with no changes to your application code.Provides an easy network setup and seamless service communication deployed across multiple ECS clusters and virtual private clouds (VPCs).

Lambda

- Lambda SnapStart
  When enabled, makes Java-based Lambda Functions start up to 10x faster, at no extra cost.

Business Applications

Supply Chain

New Service - Supply Chain unifies data, provides ML-powered actionable insights, and offers built-in contextual collaboration. Can connect to your existing enterprise resource planning (ERP) and supply chain management systems, without requiring replatforming, upfront licensing fees, or long-term commitments.

Wickr

New Service (via Acquisition) Enterprise communications service with end-to-end encryption, that allows businesses and public sector organizations to communicate more securely, enabling customers to meet auditing and regulatory requirements like e-discovery, legal hold, and FOIA requests.

Developer Productivity

CodeWhisperer

- *Support for new languages, administrative controls, and simple signup*
  Support for C# and TypeScript projects. Allows administrators the ability to enable CodeWhisperer for their organization with Single Sign-On authentication.

Security

Security Lake

New Service. Automatically centralizes your organization's security data from cloud and on-premises sources into a purpose-built data lake stored in your account.

Analytics

- ANT205 – Achieving your modern data architecture
- ANT207 – What's new in AWS streaming
- ANT208 – What's new in Amazon Athena
- ANT223 – Simplify & accelerate data integration & ETL modernization w/AWS Glue
- ANT302 – What's new with Amazon EMR
- ANT336 – Building data mesh architectures on AWS

< Table of Contents

ANT205

Achieving your modern data architecture

Speakers

- Santosh Chandrachood, General Manager, Amazon

Announcements

- Zero-ETL from Aurora to Redshift
- Redshift integration with Spark
- Data Zone

Takeaways

- AWS's 5 Pillars of Modern Data Architecture and a sampling of how AWS services enable them.
- Data Mesh considered obsolete by Gartner?

Azure

- This was a highly AWS specific take on Modern Data Architecture. No real comparative resource for Azure, maybe the Big Data Architecture concept.

< Section Header
< Table of Contents

Modern Data Architecture

- Challenges
  - Data Silo Challenges – Amount of Data is Growing, Continue to Proliferate, Struggle to combine them, More people want access, need to scale with high costs.
  - People Challenges – Many users with different skillsets, Data Stewards, Analysts, Engineers, scientists, etc. Each has their own tool preferences, each group with their own required set of internal processes.
  - Business Challenges – High performance with low cost, legacy tech, and agility.

Five Pillars

- Unified Analytics
  - Data Access Everywhere, Bringing your data in, Right tools for analytics
  - AWS Examples of Unified Analytics
    - Federated query in Redshift and Athena
    - New - Zero-ETL from Aurora to Redshift
    - New – Redshift integration with Spark
- Highest Performance at Lowest Cost
  - Explains how AWS Service innovations improved performance at same or lower cost.
  - Covered performance improvements in Redshift
  - Glue autoscaling ensures your scaling up compute when you need it, and scaling down when you don't.
- Machine Learning Integration
  - Redshift ML - SageMaker allows integration with Redshift so that SQL queries can be used to train and create models using SageMaker's ML tools to use with Redshift.
  - SageMaker Universal Notebooks – Allows data scientists to have integrated experiences with Glue, S3, Spark, EMR and other AWS services
- Unified Data Access, Security, and Governance
  - Lake Formation unifies data governance. More like "Data Lake" governance.
  - New – Data Zone, Manage Data Access and Information across the enterprise.
- Insights for Everyone
  - OpenSearch Serverless, get started searching without having to understand infrastructure
  - SageMaker Canvas, allows users to prepare, build, and test ML models with limited (low-code) technical knowledge
  - Quicksight – Dashboards, Models, and Reports, and quicksight Q for English query reporting.

End to End Data Cycles

- Overview of Ingest, Store, Transform & Catalog, Analyze & Visualize, Predict, Share.
- Managing Data Access is hard. Central vs Mesh.

The Journey to Modern Data Architecture

- Start Small, Scale Fast, Think Big

Modern data architecture

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Broadest and most cost-effective set of analytics services

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Connecting data services and ML to drive more value

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data Governance starts with business

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

End to end data life cycle

Ingest	Store	Transform & Catalog	Analyze & Visualize	Predict	Share
From any source including on-premises, real-time	Any amount of data at Exabyte scale	Data preparation, transformation, and providing seamless data access	End-to-end analytics & visualization for any use case	Most comprehensive set of ML and AI Services	With just a few lines of code

Unified Security, Governance, and Data Access

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Typical architecture

#1: Ingest

Ingesting data from any source including on-premise sources and data that is generated in real-time.

On-premises, Streams, Databases, Logs

AWS Glue ETL

#3: Transform & Catalog

Making it easy to access their data and keeping their data in sync regardless of where it lives

Amazon S3

AWS Glue Data Catalog

#2: Store

Storing both transactional data in databases and analytical data in data warehouses and data lakes at any scale.

#4: Analyze & Visualize

Analyzing data using any of ad hoc queries, distributed frameworks and search engines, and visualize the data on dashboards

Amazon Redshift, Amazon SageMaker, Amazon Athena, Amazon QuickSight, Amazon EMR

#5: Predict

Adding ML-based intelligence to applications without needing ML skills

#6: Share

Sharing new insights to take intelligent, data-driven actions

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Centralized solution challenges

- Central data team becomes the bottleneck
- Fail to scale data consumers
- Unable to discover and consume the data
- No central Data Governance
- Lack of data auditability

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Why data mesh?

- Treats existing data platforms as **independent** domains
- Improves **data governance** by pushing access policy to data domains
- Establishes a **central** mechanism for **data discovery**
- Provides **self-service data sharing** features

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ANT207

What's new in AWS Streaming

Speakers

- Divya Sridhar, Senior Product Manager Technical, Amazon
- Anusha Dasarakothapalli, Principal Software Engineer, Amazon

Announcements

❖ Lots of incremental announcement. See the What's New Section and 4th Slide in Summary.

Takeaways

❖ Focus on Offloading Capacity Management and Ease of Use.

Azure

❖ Streaming services on Azure are primarily categorized under the Messaging Services category.

Why Streaming

- Data is a perishable resource, streaming data real-time can help you identify problems before they occur.
- Common streaming-data use cases
  - Anomaly and Fraud Detection
  - Tailoring customer experience in real-time
  - Empowering IoT Analytics
  - Nourishing marketing campaigns
  - Real-time personalization
  - Supporting Healthcare and Emergency Services

Streaming Portfolio

- AWS Data Streaming, Processing and Analytics
  - Kinesis Data Streams
  - Kinesis Video Streams
  - Kinesis Data Firehose
  - Kinesis Data Analytics
  - Managed streaming for Apache Kafka

What's New

- Hands Off Capacity Management w/MSK Serverless
  - New – Use MSK Connect with MSK
  - New – MSK Serverless expanded Region Capability
  - New – Hands Off Capacity for MSK Tiered Storage
  - New – Additional Metrics for Kinesis Data Analytics for Apache Flink
- Ease of Use
  - New – Kinesis Data Streams Viewer in AWS Console
  - New – Improved Visibility in Kinesis Data Analytics for Apache Flink with Customized Mtx Windows & 5x Performance Increase in Logging
  - New – MSK Connect now supports private DNS
  - New – Move data directly from Kinesis and MSK to Redshift
  - New – Kinesis Data Analytics launches support for Apache Flink 1.15.2

< Section Header
< Table of Contents

Data streaming workflow

Source
Devices and or applications that produce real-time data at high velocity

Stream ingestion
Data from tens of thousands of data sources can be written to a single stream

Kinesis Data Streams
Amazon MSK

Stream transformation
Records are read in the order they are produced, enabling real-time analytics or streaming ETL

Kinesis Data Analytics
Kinesis Data Firehose

Stream destination
Data warehouse
Data store
Data lake
Analytical tool

Stream insights
Amazon Athena
Presto
Amazon QuickSight

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Common streaming-data use cases

- Anomaly and fraud detection
- Tailoring customer experience in real time
- Empowering IoT analytics
- Nourishing marketing campaigns
- Real-time personalization
- Supporting healthcare and emergency services

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Real-time data streaming, processing, and analytics on AWS

Amazon Kinesis Data Streams
Collect and store data streams for analytics

Amazon Kinesis Video Streams
Collect and store video streams for analytics

Amazon Kinesis Data Firehose
Load data streams into AWS date stores

Amazon Kinesis Data Analytics
Analyze data streams with Apache Flink

Amazon Managed Streaming for Apache Kafka
Fully managed, highly available and secure

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What's new in AWS Streaming

Managed Streaming for Apache Kafka

- ✓ Support for tiered storage
- ✓ MSK Serverless support for MSK Connect
- ✓ MSK Serverless available in 11 AWS Regions
- ✓ MSK Connect support for private DNS
- ✓ Native integration with Amazon Redshift

Kinesis Data Streams

- ✓ Kinesis Data Streams Data Viewer in AWS Management Console
- ✓ Native integration with Amazon Redshift

Kinesis Data Firehose

- ✓ Kinesis Data Firehose supports OpenSearch Serverless as a destination
- ✓ Kinesis Data Firehose supports VPC Flow Logs as a source

Kinesis Data Analytics

- ✓ Enhanced metrics
- ✓ Custom Maintenance Windows
- ✓ Improved application performance with logging
- ✓ Support for Apache Flink 1.15.2

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon MSK Serverless

EASILY STREAM DATA WITH AMAZON MSK WITHOUT MANAGING CLUSTER CAPACITY

"Our banking application gets used mostly during business hours. We have to scale our cluster up and down to stay cost-efficient."

"We have sales days every month, and we have to keep rescaling our clusters based on demand. Having an autoscaling feature will reduce our operational overhead."

Customer

- On-demand streaming capacity
- Fully compatible
- Throughput based pricing
- Same security as Amazon MSK
- Auto partition management
- Same high availability

New

- Use Amazon MSK Connect with Amazon MSK
- MSK Serverless expanded Region availability

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon MSK support for tiered storage

New

VIRTUALLY UNLIMITED AND COST-EFFECTIVE STORAGE TIER FOR LONG-LIVED DATA IN APACHE KAFKA APPLICATIONS

Tiered storage provides a new low-cost option that scales to virtually unlimited storage

- Increase your topics' storage without needing to increase the number of brokers
- Reprocess old data in the exact same order in which it was produced
- Save with a low-cost storage tier

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Kinesis Data Analytics for Apache Flink enhanced metrics

New

Kinesis Data Analytics for Apache Flink task managers

Amazon CloudWatch

Use metrics to scale Kinesis Data Analytics Apache Flink applications

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Kinesis Data Streams data viewer in AWS Console

New

INTERACTIVE, NO-CODE DATA EXPLORER

Use cases

1. Preview data quickly
2. Understand record schema
3. Data debugging

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Kinesis Data Analytics for Apache Flink: Improved usability

New

Customize maintenance windows

5x improvement in application performance with logging turned on

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon MSK Connect

EASILY MOVE DATA BETWEEN APACHE KAFKA CLUSTERS AND OTHER SYSTEMS

Fully managed

Fully compatible

Serverless

Autoscaling

New

Amazon MSK Connect private DNS support

- ✓ Keep databases, data warehouses, and other systems like the secrets manager in your own VPC, inaccessible to the internet
- ✓ Use Amazon MSK Connect with Amazon MSK Serverless

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Move data from MSK and Kinesis directly to Amazon Redshift

New

NATIVE INTEGRATION WITH STREAMING ENGINES FOR FAST AND EASY DATA INGESTION

- Create materialized views on top of streams directly
- Achieve low latency in the seconds
- Support up to hundreds of MBps of throughput
- Entirely configurable with SQL
- Define a schema or ingest semi-structured data with SUPER & VARBYTE data types

Learn more: https://go.aws/3eG5kZF

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon Kinesis Data Analytics launches support for Apache Flink version 1.15.2

Kinesis Data Analytics championed Apache Flink Version 1.15.2 release

Focus feature: Async Sink, an AWS contributed feature

A new framework allowing customers to easily create custom sinks

AWS facilitating community project to create a DynamoDB sink using the framework

Build a custom sink in under 20 minutes: https://bit.ly/3VyAkeJ

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ANT208

What's new in Amazon Athena

Speakers

- Ofer Eliassaf, Director, REM Cloud Infrastructure, Mobileye
- Scott Rigney, Sr. Product Manager, Amazon Web Services

Announcements

- Lots of incremental announcement. See notes and slide 12 for a summary.

Takeaways

- Athena get's faster and new capabilities with SQL v3 engine and Iceberg.

Azure

- I don't believe Azure has Athena like capability. Interestingly Azure Data Lake Storage Gen 2 is a new Athena Connector.

New Features and Capabilities

- New – Athena for Apache Spark.
- New – Athena SQL Engine v3
  - 50+ New SQL Functions, 20% Faster Queries, 100% backwards compatible
  - Faster Queries, on more information, for lower cost
- New – Faster queries with result reuse

How MobileEye Use Athena

- MobileEye is an Autonomous Driving software company. The give an overview of how they chose Athena, and how they use it.

New Data Lake, Connector, and Integration Features

- Presented AWS's view of Modern Data Architecture
  - Data Lakes – Store all your data in open formats. Cost effective storage, Decouple Storage from Compute, Choice of Analytical and ML engines, process data in place.
- New – Capabilities for Apache Iceberg (underlying engine)
  - Create Table as Select (allows you to create new tables as a result of a Select query), Views to hide complex logic, Performance 10x faster queries
  - Merge to sync tables and simplify data prep, Vacuum to manage storage, New formats Avro & ORC.
- New – Expanded fine-grained access controls
- New – Data source connectors

< Section Header
< Table of Contents

NOW AVAILABLE

Amazon Athena for Apache Spark

Get started with interactive analytics on Apache Spark in under a second

Harness Apache Spark for complex, powerful analytics

Spend more time on insights instead of waiting for results

Build applications without managin resources or configuring software

GENERALLY AVAILABLE TODAY

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NOW AVAILABLE

Amazon Athena SQL engine v3

Upgrade today to experience multiple improvements and new features

Faster queries at lower cost

More frequent updates and features

Based on open-source Trino, Presto, and AWS-developed innovations for both

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SQL engine v3 new features and enhancements

FASTER QUERIES AND MORE FREQUENT UPDATES

- NEW — Query result caching
- NEW — Aggregation functions
- NEW — Date & time functions
- UPDATED — Prepared statements
- NEW — Number formatting
- UPDATED — Memory usage
- NEW — More frequent updates
- UPDATED — Geospatial functions
- UPDATED — Syntax enhancements
- UPDATED — AWS Glue metadata
- NEW — Array functions
- UPDATED — Right-join dynamic filter

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NOW AVAILABLE

Faster Queries with Result Reuse

New caching feature to accelerate repeat queries and reduce time to insight

5x

Up to 5x faster queries when result reuse is active

$0

No data scanned, so queries are free to you

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws

The process that led us to Amazon Athena

Design phase

Initial idea of developing two systems for short/long term storage (cost vs. efficiency)

Started with the long-term storage

Research conclusion: Amazon Athena + S3 parquet files for long term

Solution verification with AWS solution architects

Data-lab in Seattle

Implementation phase

Conclusion verification and POC implementation in the data-lab

Finalizing implementation

Production phase

Usage grew exponentially (usage pattern, amount of data, cost)

Iterative optimizations

Up until now, no need for the short-term storage

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Optimizations toward our current state

Iterative approach

- Query-level optimizations
- Partnered with AWS to reach massive concurrency
- Refactored data model to eliminate joins and align with Athena best practices
- Injected partition projection (geospatial + time)

Our latest optimizations led us to get:

90% Query time decrease

90% Cost decrease

0 bottlenecks across different map activities

10s → 100s concurrent map cell runs per minute

0 concurrency issues

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Modern data architecture on AWS

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

New features and capabilities with Iceberg

NEW

CTAS

Easy and fast way to create new tables

NEW

Views

Create views to hide complex logic

NEW

Performance

10x faster queries when using SQL engine v3

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

New features and capabilities with Iceberg

NEW

Merge

Sync tables to simplify data prep and updates

NEW

Vacuum

Manage storage and delete records to meet regulatory requirements

NEW

Avro & ORC

New support for Avro and ORC formats

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NOW AVAILABLE

Expanded fine-grained access controls

Apply AWS Lake Formation access control policies to more of your data

Centrally defined table-, row-, and column-level policies in Lake Formation

Now supports all table formats and file types available with Athena

Filtering operations performed natively by Athena for increased performance

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

New data source connectors

DATA SOURCE CONNECTORS FOR IN-PLACE ANALYTICS

Easy to set up and query

Use the Athena console to create and configure your connectors, then run SQL queries

Secure cross-account access

Deploy and configure one connector, then use flexible IAM policies to enable cross-account access to other analysts

Open-source and fully supported

Athena connectors are open-source, so you can use our code as a starting point to building your own

NEW

Azure Data Lake Storage Gen2
Azure Synapse
Cloudera
Google BigQuery
Hortonworks
IBM Db2
Microsoft SQL Server
Oracle
SAP HANA
Snowflake
Teradata

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws

What's new with Amazon Athena

KEY TAKEAWAYS TO APPLY TO YOUR WORKLOADS TODAY

- **Amazon Athena for Apache Spark** – Available today! Sub-second startup for interactive Spark analytics. Check out **ANT209** to learn more
- Upgrade to **SQL engine v3** for faster queries and more analytics features
- Turn on **Query Result Reuse** for cached results
- Expanded support of **Apache Iceberg** and **AWS Lake Formation**
- **New data source connectors** and **Amazon AppFlow** help you expand analysis and break down data silos
- And many more …

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ANT223

Simplify & accelerate data integration & ETL modernization w/AWS Glue

Speakers

- Roberto Santos Filho, IT Superintendent, Banco Itau Unibanco
- Santosh Chandrachood, General Manager, Amazon
- Moises Nascimento, Director, Itau Unibanco S/A

Announcements

❖ Lots of incremental announcement. See notes and slide 12 for a summary.

Takeaways

❖ Great Overview of Glue. If you don't know what Glue entails this is a great session. Also Itau's approach to Data Strategy is interesting.

Azure

❖ Azure's Data ETL/Pipelining tools is Data Factory.

< Section Header
< Table of Contents

Trends & Challenges in Data Integration

- Business Challenges - Lack of Agility, Lack of Self-Service, Missed SLAs
- Data Engineering Challenges – Infrastructure Management, Multiple Tools, Scalability Issues
- Leadership Challenges – Expensive, No multi-persona support, Lock-in
- Data & Pipeline Challenges (See Slide 2 & 3)

Glue Overview

- Why Glue?
  - Focus on Data not Infrastructure
  - Serverless so scales on Demand
  - Based on Open-Source Engines
  - Works for a variety of workloads and personas
- New – Support for Cloud Shuffle Plugin
  - Improves reliability for Spark jobs by allowing rebalance of storage across clusters
- New – Glue 4.0
  - Upgrades Glue Engine to latest tools for Spark, Python, Scala
  - Spark Connector for Redshift Integration
- New – Glue with Ray
  - Upgrades Glue Engine to latest tools for Spark, Python, Scala
  - Allows Python jobs to scale across multiple nodes for large workload processing.
- New – Native support for Hudi, Delta Lake, Iceberg
  - These technologies bring transaction capability to Data Lake stores
  - Glue now supports this capability natively, no connectors needed
- New – Custom Visual Transforms
  - Build and Share Custom transformations across the organization
- New – Glue Git Integration
  - Manage version and configuration with your favorite Git repository.
- New – Data Catalog Capabilities
  - Makes it easier to track, audit and manage data & access.
- New – Data Catalog Crawler Capabilities
  - Cross Account Capability, Snowflake, MongoDB.
- New – Data Quality
  - Demo.

Why Itau Chose Glue for Data Mesh

- Watch Here.
  - Leveraged AWS and Glue to build a Modern Data Mesh platform.
  - Formulating a Data Mesh Strategy
  - Built a canonical data model to identify domains of data
  - Establish Governance
  - Data Security vs Democratization
  - Data Privacy vs Massive Data Usage
  - Governance vs Agility
  - Centralization vs Scalability
  - Think Product when you think data
  - Example of how Glue helped them achieve visibility of Payments.

Data integration process

1) Motivation (use case examples)

New business report request | Or | Data pipeline between teams | Or | Research using data from several sources for Machine learning

2) Build

Define → Author → Transform → Test

Requirement Specification → Access Data → Structure Data → Clean, aggregate etc. → Visualize & Validate

Debug and Iterate

System failure, data growth, bugs, corrupt data

3) Operationalize

Troubleshoot ← Monitor ← Build workflow & schedule ← Set SLA

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data challenges

- Growing exponentially
- From new sources
- Increasingly diverse
- Analyzed by many applications

- Limited by scale and cost as data volumes grow
- It take months to introduce new data source
- Transformation requires more processing power than companies have
- Challenges to manage security and reliability when used by multiple applications

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data integration and ETL is HARD

Databases
SaaS Apps
3P data
Media
Data Streams

DATA SOURCES

- Complex application code
- Cumbersome retry logic
- Pipeline management
- Need specialized ETL skills

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Glue

Focus on data
Low maintenance serverless solution

Scale on demand
Avoid licensing cost and infrastructure idle time

Powerful open source engines
No lock in, support from wide innovative eco system

All in one
Support all your users personas and workloads

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Glue overview

SERVERLESS DATA INTEGRATION SERVICE

Connectors
- Data warehouse
- Data lakes
- Marketplace
- NoSQL
- Streams

Author
- Visual
- Notebook
- Built-in transformations
- IDE

Operationalize
- Workflow
- Monitoring
- Schedule

Data Management
- Data Catalog
- Data quality
- Sensitive data detection

Engines
- Serverless infrastructure
- Choice of data integration engines
- Compliance and security
- Stream data processing

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

New

Announcing:
Cloud Shuffle Storage Plugin improves fault tolerance

- Shuffles in Spark occur when data needs to be redistributed across the cluster
- Shuffles are constrained by local disk capacity
- Cloud shuffle reduces Spark shuffle failures
- Native support on AWS Glue 3.0+ with Amazon S3
- Open binaries under the Apache 2.0 license

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

New

Announcing Glue 4.0

FAST, PREDICTABLE, AND COST-EFFECTIVE

Upgrades Glue engine

Apache Spark	Apache Spark 3.3.0
python	Python 3.10
Scala	Scala 2.12

Adds more options for scaling, storing, and running your jobs

- Performance optimized Spark 3.3
- Distributed Pandas API on Spark - Improved data processing
- **for Amazon Redshift integration Apache Spark – 10x faster in TPC-DS at 3TB scale**

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Preview

Announcing AWS Glue with Ray

SERVERLESS DATA INTEGRATION IN DISTRIBUTED PYTHON

"Ray is an open-source unified compute framework that makes it easy to scale AI and Python workloads"*

- **Scalable to hundreds of nodes**
- **Superfast start and scale up/down time**
- **Familiar Python Primitives, built-in libraries**

*To learn more - visit Ray.io

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

New

Announcing:
Native support for Hudi, Delta Lake and Iceberg

Apache hudi | Linux Foundation Delta Lake | ICEBERG

Simplify incremental data processing in data lakes built on Amazon S3

- Built in support to read and write (insert, update and delete)
- No setup, minimal configuration and easy to deploy
- Latest framework versions available with Glue v4.0

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

New

Announcing: Custom visual transforms

- (Re)use custom business logic in visual ETL jobs
- Share custom transformations between teams

Benefits:

- Reduces dependence on Spark developers
- Maintain and update jobs more easily
- Custom visual transforms work in both visual and code-based jobs

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

New

AWS Glue Git integration

GitHub

AWS Glue

AWS CodeCommit

- Track changes to your ETL code and Visual ETL jobs
- Deploy AWS Glue jobs with your existing pipelines and DevOps tools
- Supports GitHub and AWS CodeCommit
- Download/upload AWS Glue jobs in AWS Management Console

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

New

Data Catalog – Capabilities

- Audit with CloudWatch Catalog Metrics
- Federate Hive MetaStore and Access with Athena, RedShift and Glue
- Database, Connection Tagging
- Sync Delta Lake Metadata in Catalog to enable column permissions

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Glue Data Catalog - Crawlers

New

Simplified monitoring with properties and metrics associated with past crawler runs

Cross-account crawling with Lake Formation

Faster Incremental crawling for existing Catalog Tables

Table and column stats collection

Catalog SnowFlake, MongoDB Atlas

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Glue Data Quality Features and Benefits

Preview

- Serverless, scalable and performant
- Data Quality Recommendations
- Out-of-the-box data quality rules and actions
- Data Quality Definition Language
- Data Quality for data-at-rest and data-in-transit
- Supports multiple persona

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How it works: Data Quality on AWS Glue Data Catalog

Preview

Data Steward → Selects a dataset in AWS Glue Data Catalog → AWS Glue Data Quality analyzes data and recommends rules → Data Steward refines the rules to create finalized rules → AWS Glue Data Quality Evaluates rules → Data Steward reviews results and alerts, takes appropriate action

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data Platform - Itaú

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ANT302

What's new with Amazon EMR

Speakers

- Neil Mukerje, Principal Product Manager, Amazon Web Services

Announcements

- EMR Serverless, and lots of performance and enhancements. See New links.

Takeaways

- EMR Serverless gives you the power of EMR without having to manage the infrastructure
- Great discussion of Transactional Data Lakes and why they are needed.

Azure

- Azure's Synapse Analytics is the closest service to EMR Serverless. Customers can use Delta Lake or Hudi for Transactional Lakes, although Delta Lake is better supported (IMO)

< Section Header
< Table of Contents

What is EMR

- Runs Big Data Apps using Spark, Hive, Presto, Hbase, Flink
- New – EMR Serverless
- Comparison of EMR On-Premises vs Cloud (blah blah blah)
  - [Editor] If you are still on-prem you are wasting $$$
- Cloud Cost Optimizations through Transient Instances, Spot Instances, or reserved instances.

What's New

- New Enhancements for EMR on EC2
  - Reduced startup time in private subnet up to 30%
  - Reduced task nodes startup time by 30%
  - Better performance with Spark Shuffle awareness
  - Reduced costs, improved performance with EBS GP3 instances
- New Enhancements for EMR on EKS
  - Job Templates – simplifies authoring experience by enforcing common parameters
  - Spark-SQL runner – Execute SQL scripts directly with API
  - DynamoDB connector – Easily Access data in DynamoDB
  - Enhanced job failure messages – show task failure messages
- New - Performance Optimized Runtimes
  - Spark on EMR 3.9x Faster than OSS Spark
  - Spark EMR 6.9 1.3x faster than EMR 6.5
  - Trino up to 4x faster on EMR than OSS
  - Hive S3 Storage Optimizations 15x faster
- New – Engine Improvements (Presto/Trino)
  - Spot Loss Handling – Faster fail on queries when using spot instances
  - Improved join reordering
  - Presto - Enforce fine-grained lake formation policies
  - Trino supports for Iceberg and Delta Lake transactions
  - Trino – HDFS for checkpointing
- New – Hive Engine Improvements
  - Zero-rename Writes, MSCK Optimizations, Parquet Modular Encryption, Fine-grained lake formation policies, 500+ incorporated fixes, Iceberg support in EMR 6.9
  - Glue now supports this capability natively, no connectors needed
- Transactional Data Lakes. More customers investigating or using TDLs (Iceberg, Delta Lake, Hudi) to solve challenges related to
  - New – Iceberg features for Transactional Data Lakes
  - New – Support for OSS Delta Lake 2.1 (Spark3, Trino)

EMR Serverless (New)

- Watch Here.
  - Jobs, Workers, and Pre-Initialized Workers
  - As Jobs are submitted, they acquire workers from a shared pool
  - Support for Multiple Availability Zones for High Availability
  - Can define boundaries for scaling overall, or by job.
  - Spark history service is included to view job history
  - Integrated with Cloud Watch to Monitor EMR Serverless applications
  - Graviton2 Support for lower cost

EMR Studio

- Watch Here.
  - Fully managed IDE for Interactive Notebooks
  - New – Mixed language notebooks, real-time collaboration, SQL Explorer build-in, mount local filesystem into EMR
  - New – Security Enhancements
  - New – Role based access integration with SageMaker

Amazon EMR

EASILY RUN SPARK, HIVE, PRESTO, HBASE, FLINK, AND MORE BIG DATA APPS ON AWS

Latest versions

Updated with latest open source frameworks within 60 days

Support for popular OSS like Spark, Flink, Hudi, Iceberg etc.

Great performance at lower cost

Spark workloads run up to 3x faster compared to open source

50–80% reduction in costs with EC2 Spot, EC2 Savings Plan, Reserved Instances

Per-second billing for flexibility

Use S3 storage

Process data in S3 securely with high performance using the EMRFS connector

Scale Compute and Storage independent of each other

Multiple deployment options

Fully managed, choose multiple deployment options based on your need

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW!

Amazon EMR on EC2 enhancements

Reduce start-up time for Amazon EMR on EC2 cluster in a private subnet by up to 30%

Reduce task nodes start-up time for Amazon EMR on EC2 cluster by up to 30%

Better performance and lower costs with Spark shuffle awareness in managed scaling

Reduce costs and improve performance with EBS GP3 volumes

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW!

Building applications on Amazon EMR on EKS

Job templates

Simplifies job authoring experience for data engineers and scientists by enforcing common parameters

Spark-SQL runner

Execute SQL scripts directly with API without writing any additional code

DynamoDB connector

Easily access data in DynamoDb

Enhanced job failure messages

Show task failure messages in DescribeJobRun API and driver logs for job failures

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW!

Amazon EMR runtime vs. OSS Spark

UP TO 3.9X FASTER PERFORMANCE FOR APACHE SPARK 3.3.0

Spark 3.3.0 on EMR 6.9.0

*Based on TPC-DS 3TB Benchmarking running 6 node C5.9XL cluster and EMR 6.9.0 running Spark 3.3.0

Runtime built on a optimized version of Apache Spark

Best performance

- **Up to 3.9x faster on total time**

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW!

Amazon EMR runtime for Apache Spark: Performance improvements - 2022

Runtime built on a optimized version of Apache Spark

Best performance

- **Up to 1.3X faster on total time**

100% compliant with Apache Spark APIs

Spark 3.3.0 on EMR 6.9.0

*Based on TPC-DS 3TB Benchmarking running 6 node C5.9XL cluster and Amazon EMR 6.5.0 running Spark 3.0

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW!

Amazon EMR runtime vs. OSS Trino

UP TO 3.1X FASTER PERFORMANCE FOR APACHE TRINO 388

Runtime built on a optimized version of Trino

Best performance

- **Up to 3.1x faster on geometric mean**
- **Up to 4.2x faster for total time**

Trino 388 on EMR 6.9.0

*Based on TPC-DS 3TB Benchmarking running 6 node C5.9XL cluster and EMR 6.9.0 running Trino 388

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW!

Amazon EMR runtime vs. OSS Hive

UP TO 15X FASTER PERFORMANCE FOR APACHE HIVE WRITES USING EMRFS S3-OPTIMIZED COMMITTER

Hive write performance with the Amazon EMR Hive zero-rename feature

Best performance:

- **Up to 15x faster writes**

Hive 3.1.2 on EMR 6.5.0

*Based on load phase of TPCx-BB 1 TB Benchmarking running 1 m5d.8xlarge master node, 20 m5d.8xlarge core nodes cluster on EMR 6.5.0 running Hive 3.1.2

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW!

Presto/Trino improvements

Spot loss handling

Improved join reordering, selection and common subexpression removal

Enforce fine-grained lake formation policies (Presto)

Iceberg and Delta support (Trino)

HDFS for checkpointing (Trino)

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Hive improvements

NEW!

Zero-rename Writes

MSCK Optimization

Parquet Modular Encryption

Fine-grained lake formation based policies

Incorporated 300+ improvements, critical fixes into Amazon EMR Hive

Iceberg support in Amazon EMR 6.9

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Transactional data lakes features

TRANSACTIONS, RECORD-LEVEL UPDATES/DELETES, AND CHANGE STREAMS TO DATA LAKES

Ingestion

- Transactions (ACID) - Reader and writer isolation
- Transactions (ACID) - Concurrent write support
- Record level upserts and deletes
- High throughput streaming ingestion
- Spark, Flink, and Java Writer Support
- Automatic compaction of small files
- SQL DML support

Query

- Spark, PrestoDB/Trino, Flink, Hive Support
- Efficient queries across partitions and files
- Incremental query support
- Time travel query support

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Transactional data lakes features

TRANSACTIONS, RECORD-LEVEL UPDATES/DELETES AND CHANGE STREAMS TO DATA LAKES!

Administration

- Async background compaction of files
- Async background sorting and clustering of keys
- Automatically clean up files beyond retention period
- Metrics for past commits or rollbacks

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Apache Iceberg

NEW!

OPEN TABLE FORMAT FOR HUGE ANALYTIC DATASETS

Apache Iceberg 0.14.1 is packaged as a library for Spark3 Runtime, Trino, Flink, and Hive in EMR 6.9.0.

Key new features include:

Apache Iceberg

- Time travel support with Spark SQL and Trino SQL
- Merge on Read (MoR) support
- Optimistic concurrency with AWS Glue Data Catalog
- Disaster recovery with S3 access points
- Flink and Hive integration (EMR 6.9.0)

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW!

OSS Delta Lake

OPEN-SOURCE STORAGE FRAMEWORK THAT ENABLES BUILDING A LAKEHOUSE ARCHITECTURE

DELTA LAKE

OSS Delta Lake 2.1.0

OSS Delta Lake 2.1.0 is packaged as a library in EMR 6.9.0

Engines supported : Spark3 and Trino

To learn more :

https://docs.aws.amazon.com/emr/latest/ReleaseGuide/emr-delta.html

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon EMR Serverless

All the benefits of EMR without managing clusters and servers

- Run frameworks more easily; just pick a version and run
- Automatically scale; don't guess cluster sizes
- Optimize cost; Automatic and fine-grained scaling reduces cost
- Performance-optimized version delivers 2x better performance
- Multi-AZ resiliency from day 1
- Integration with familiar tools like Apache Airflow

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Jobs

Run jobs on applications

Can run multiple jobs on an application

Can control authorization using per-job execution role

Workers

Spark HIVE

Internally used to execute your workloads

Workers run the OSS framework you choose

You can change the size of workers to control performance

Pre-initialized workers

Spark HIVE

Optional feature to pre-initialize workers

Jobs start immediately

Helps you maintain a warm pool

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW!

Amazon EMR Serverless

ALL THE BENEFITS OF EMR – WITHOUT MANAGING CLUSTERS

- AWS Graviton2 Support
- Application capacity and job metrics in Amazon Cloudwatch
- Live debugging of jobs
- Fast Fine-Grained Auto Scaling
- Auto stop Auto start

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

EMR Studio

FULLY MANAGED IDE FOR INTERACTIVE DATA ANALYTICS: DEVELOP, VISUALIZE, AND DEBUG APPLICATIONS

- Single sign-on integration with IdP
- Fully-managed Jupyter Notebooks
- Integrated with Git Repositories
- Simplified debugging with Spark UI and YARN UI
- Browse, create, or delete EMR clusters
- Run Notebooks in workflows using APIs
- Run interactive data analysis using EMR on EKS clusters

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

EMR Studio features

NEW!

FULLY MANAGED IDE FOR INTERACTIVE DATA ANALYTICS: DEVELOP, VISUALIZE, AND DEBUG APPLICATIONS

- IAM authentication and federation support
- Multi-language support (R, PySpark, Scala, SQL)
- Auto-terminate idle clusters
- Real-time co-authoring of notebooks
- SQL explorer with Presto support
- Mount workspace directories to EMR clusters
- Latest JupyterLab, JEG, Livy, SparkMagic

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Security pillars

Isolation	Authentication	Authorization	Encryption	Audit
VPC	LDAP	Cluster IAM Role	Encrpytion at rest	Audit using Ranger
Private subnets	Kerberos	FGAC using Apache Ranger	Encrpytion in transit	Audit using AWS Lake Formation NEW!
Security groups	AWS IAM Identity Center (EMR Studio)	FGAC using AWS Lake Formation NEW!		
	AWS IAM (EMR Studio)	Job runtime role NEW!		

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data access control from Amazon SageMaker

NEW!

Fine-grained access control from interactive SageMaker Studio Notebooks

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ANT336

Building data mesh architectures on AWS

Speakers

- Travis Muhlestein, Chief Data & Analytics Officer, GoDaddy
- Nivas Shankar, Principal Product Manager, AWS
- Ian Meyers, Director of Product Management, Amazon Web Services

Announcements

- None

Takeaways

- Good overview of modern Data Mesh and how it might be realized on AWS
- In 2022 Data Management Hype Cycle Gartner proclaims Data Management Obsolete before plateau.

Azure

- You can build a Data Mesh architecture in Azure too.

< Section Header
< Table of Contents

Modern Data Strategy & Challenges

- See ANT205 – Achieving your modern data architecture

Why Data Mesh

- Acknowledges that data challenges exist and seeks to work within those constraints.
  - Use existing investments in data platforms and treat them as independent domains
  - Improve data governance by pushing policy down to data domains
  - Provides a clear mechanism for centralized data discovery
  - Provides self-service data sharing features to allow owners to grant access to consumers
  - Allows leaders to measure and invest in data products based on usage and business value.

Design Goals & Core Principles - Watch

- Four core principles of a Data Mesh Architecture
  - Data domain ownership
  - Data as a Product
  - Federated computational governance
  - Self-serve Sharing
- Data Mesh Design Goals
  - Enable organizations to get value from data at scale
  - Create business-oriented data products that can support the top strategic goals
  - Allow business domains federated governance through lightweight centralized policy by removing bottlenecks
  - Encourage data-driven agility
  - Support the sharing of data products, with the goal of delighting the experience of data users
- Data Mesh Architecture – How it might look in AWS
  - Decentralized, lightweight federated governance across domain-oriented data systems to drive governed sharing.
  - Lake Formation and Data Zone allow for policies on what the data is and who can access it.

Favorite Quotes

- “Everyone wants to be a data consumer, no one wants to be a producer”

Why GoDaddy built a Data Mesh

- Watch
- Business Outcomes of the Strategy
  - Created hierarchical views of data products at different levels so that business users can analyze information to make quicker business decisions
  - Automated access management framework to enable self-served access to data within and across lines of business
  - Accelerated the data platform adoption to 10+ LOBs and 300+ teams globally
  - Enabled data scientists to find and access data needed to generate ML models across LOBs
  - Achieved 10s of millions of dollars in cost savings from data reuse and better management of purchased datasets.

Why data mesh?

Use existing investments in data platforms and treat them as independent "domains"

Improve data governance by pushing policy down into data domains

Provides a clear mechanism for centralized data discovery

Provides self-service data sharing features to allow domain owners to grant access to consumers

Measure and invest in data products based on usage and business value

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh: Four core principles

Data owner

Data domain ownership

A data mesh features data domains as nodes, which exist in data lake accounts; it is founded in decentralization and distribution of data responsibility to people closest to the data

Data engineer

Data as a product

A data producer contributes one or more data products to a central catalog in a data mesh account; DaaP must be autonomous, discoverable, secure, and correct, and useable

Data steward

Federated computational governance

Federated data governance is how data products are shared – delivering discoverable metadata auditability based on federated decision-making and accountability structures

Data consumer

Self-serve sharing

The platform streamlines the experience of data users to discover, access, and use data products; it streamlines the experience of data providers to build, deploy, and maintain data products

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh design goals

Enable organizations to get value from data at scale

Create a business-oriented data products that can support the top strategic goals

Allow business domains federated governance through lightweight centralized policy by removing bottlenecks

Encourage data-driven agility

Support the sharing of data products, with the goal of delighting the experience of data users

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh architecture

DECENTRALIZED, LIGHTWEIGHT FEDERATED GOVERNANCE ACROSS DOMAIN-ORIENTED DATA SYSTEMS TO DRIVE GOVERNED SHARING

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh principle #1: Data domain ownership

DATA OWNERS ARE ACCOUNTABLE FOR THEIR DATA PRODUCTS TO BE RELIABLE, AVAILABLE, AND ACCURATE

Data owner

Accountability for data domain and consumption of data products

Create
Owns data **created** or **transformed** in my organization

Store
Owns data **stored** in my organization

Share
Owns how my organization **uses** the data

- Register and catalog data assets
- Protect and secure organization data
- Manage data quality
- Maintain ease of use and atomic integrity

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh principle #2: Data as a product

DOMAIN-DRIVEN DESIGN TECHNIQUES TO FORMULATE AND ESTABLISH BOUNDED CONTEXTS FOR DATA PRODUCTS

Data engineer

- Each team manages their data and organizes it as data products
- Each product provides an interface(s) to allow others to interact (e.g., APIs, SQL, reports)
- Remove usability frictions, meet the user where they are
- Provide all supporting metadata, lineage
- Data products are valuable on their own

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh principle #3: Self-serve sharing

ECOSYSTEM OF SELF-SERVE DATA INFRASTRUCTURE WITH OPEN PROTOCOLS

Data consumer

- Design for generalist majority (i.e., make it easy to use and adopt with no specialist skills needed)
- Enable personas to discover, learn, understand, consume, and maintain data products
- Collection of interoperable data products, which enable cross-functional domains to produce and consume data easily and with autonomy and will allow it to scale
- Data products must include data, metadata, code, and policy all as single unit of value
- Abstract complexity through automation

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh principle #4: Federated data governance

GOVERNANCE MODEL THAT EMBRACES DECENTRALIZATION AND DOMAIN SELF-SOVEREIGNTY THROUGH DECISION-MAKING MODEL LED BY FEDERATION OF DATA PRODUCT OWNERS

- Decentralization implementation of governance team and standard authorization
- Governance team = a guild consisting of representatives of all teams taking part in the data mesh
- They create global policies and standardization to achieve interoperability
- Automated execution of policies by the data domains (e.g., data classification and privacy, compliance, security, documentation, and interoperability)

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh architecture pattern: Data lake data products sharing

Multi-domain federated customer actions

Data mesh

GoDaddy conceptual/ domain architecture

Artificial Intelligence & Machine Learning

- AIM307 – JPMorganChase real-time agent assist for contact center productivity
- AIM314 – Accelerate your ML journey with Amazon SageMaker low-code tool
- AIM321 – Productionize ML workloads using Amazon Sagemaker MLOps featuring NatWest
- AIM332 – WaFd bank delivers enhanced self-service with AWS conversational AI

< Table of Contents

AIM307

JPMorganChase real-time agent assist for contact center productivity

Speakers

- Ami Ehlenberger, CTO, Machine Learning & Intelligence Operations, JPMC
- Vafa Ahmadiyeh, Principal Solution Architect, AWS

Relevant Announcements

❖ New Feature - Real-Time Transcribe Call Analytics

Takeaways

❖ CCI is an easy way to add AI/ML capabilities to any conversational audio stream.

Azure

❖ Cognitive Services – Speech to Text but must custom build real-time actionable capability.

Add AI to Your Existing Contact Center with AWS Contact Center Intelligence.

- Supports customers who already have Contact Center Solutions from Genesys, 8x8, talkdesk, CISCO, Avaya, etc.
- Uses AWS AI services to address 3 main use cases.
  - Self-Service virtual agents
  - Real-time call analytics & agent assist
  - Post-call analytics
- Leverages AWS Services to satisfy the use cases
  - Lex – Conversational AI
  - Polly – Text to Speech
  - Transcribe – Speech to Text (*New – Real-Time Analytics during live calls.)
  - Kendra – Intelligent Search (Knowledge Base)
  - Comprehend – Natural Language Processing
  - Translate – Machine Translation
  - Chime SDK – Voice Connector (SIP)
- Ami Ehlenberger.
  - JPMC stood up a product team to specifically address call center operations through artificial intelligence and machine learning capabilities.
  - Discusses Pre-Call / In-Call / Post-Call Workflow
  - How can we consume data from workflow and create a better CX
- Product Team.
  - Due to Pandemic, Agents working from home / hybrid environment
  - Lost ability to Team / Help Each Other / Find Information
  - Needed a Real-time Virtual Agent Assistant (EVEE)
  - Built on-prem solution, that took along time, progress over perfection.
  - Found Transcription Accuracy to be significant. “Chase vs Cheese”
  - Moving to Transcribe improved transcription and outcomes
- Knowledge Management
  - Timeboxed adding Kendra to See if there is value
  - Add content from multiple sources.
  - Incorporate agent feedback (created Agent “Champions”)
  - Collected feature requests from Champions to make better
- Architecture / Build Solution
  - Current architecture replaces most on-prem solutions.
  - Detailed dive into on-prem/cloud solution.
  - Load testing was a challenge
  - You can build this yourself (live walkthrough)
  - Demo of live transcribe analytics on newly built solution.
- Resources
  - https://go.aws/cci

Favorite Quotes

- “Progress over Perfection”

< Section Header
< Table of Contents

Using AI to solve contact center challenges

AWS CCI solutions use cases

ADDRESSING 3 PRIMARY JOBS WITH AI AND AUTOMATION

Self-service virtual agents

- Answer calls without a live agent
- Offload repetitive queries to virtual agents, reduce employee frustration
- Reduce call transfers with intelligent routing to improve agent productivity

Real-time call analytics & agent assist

- Alert supervisor to help improve customer satisfaction
- Provide agents with actionable insights and easy search capabilities from the knowledge base
- Capture insights like customer sentiment, call intent, and more

Post-call analytics

- Identify trends and business improvement opportunities to drive revenue
- Assess agent performance and monitor compliance
- Improve employee retention, reduce operational costs

Powering AWS CCI solutions and Amazon Connect

AWS AI LANGUAGE SERVICES

Conversational AI	Text to speech	Speech to text	Intelligent search	Natural language processing	Machine translation	Voice connector (SIP)
Amazon Lex	**Amazon Polly**	**Amazon Transcribe**	**Amazon Kendra**	**Amazon Comprehend**	**Amazon Translate**	**Amazon Chime SDK**

AMAZON TRANSCRIBE – REAL-TIME ANALYTICS DURING LIVE CALLS

Migration to real-time transcription using Amazon Transcribe

- Improvement in word error rate over incumbent service of ~12%
- Out-of-the-box automated redaction of personally identifiable information
- Redirection of ~3 FTE developers from dedicated support to new feature delivery
- Pilot users called out increase in accuracy in user acceptance tests
- Enabled call intent ML models to analyze calls parallel to transcription API

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

We aim to deliver the right answers to our specialists at the right time

Evaluating Amazon Kendra for retrieving the best knowledge management answer for our specialists resulted in:

- Ability to incorporate agent feedback to improve answer accuracy
- Out-of-the-box filtering capability
- Ability to add content from multiple sources and formats
- Reduction in dev time from weeks to days

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Real-time call analytics and agent assist

REFERENCE ARCHITECTURE

1. Customer audio is streamed real time into AWS
2. Audio is transcribed using Amazon Transcribe
3. The transcribed text is through a NLP model to produce the intent
4. The intent from the model is sent through Amazon Kendra to fetch relevant content
5. The real-time guided content is shown to the agent

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Build transcribe live call analytics with a few clicks

© 2022, Ama

AIM314

Accelerate your ML journey with Amazon SageMaker low-code tool

Speakers

- Vadim Omeltchenko, AI/ML SA, AWS
- Ori Goshen, Co-CEO, AI21 Labs
- Claire OBrien, Product Manager, Amazon Web Services

Relevant Announcements

❖ SageMaker now imports data from EMR

Takeaways

❖ A product demo of SageMaker. How to input, train, and deploy models.

Azure

❖ I don't believe Azure has a similar product. The closest product is Azure Machine Learning capability of Azure Synapse Analytics.

Why Low-Code

- Typical ML Workflow (Prepare, Build, Train & Tune, Deploy & Manage)
  - Challenges: Requires Deep Expertise, Experimentation is expensive, data-scientists are in short supply.
  - Low-Code: Faster ML Workflows, Open & Customizable solutions, great for new and experienced practitioners

SageMaker Low-Code Overview

- Data Wrangler (Prepare)
  - Visually aggregate and prepare data.
  - Exploratory Analysis using, ML powered to diagnose data quality issues, SageMaker Clarify integration to detect and mitigate bias, and Quick Model to estimate quality of model
  - Point & Click Data Transformations and Custom Code Editor and Cod Snippets
  - Can export to Python code or Notebook
- Autopilot (Build, Train, Tune, Deploy)
  - AutoML that to select & tune for best model
  - Feature Selection / Data Split (Train / Test),
  - Choose Training Modes (AutoGluon, Hyper Parameter Optimization
  - Model Visibility, Explainability
  - Deploy to SageMaker Endpoint, can be used to make real-time
- Jumpstart (Prebuilt Models)
  - A model hub of pre-trained and easily tunable models.
  - Using Pre-Trained Models Saves Time
  - Models come from publicly available models
  - Includes Sample Jupyter Notebooks
  - Cloud Formation Templates
  - Ability to bring in 3rd Party Commercial Models

Demos

- Demo of Data Wrangler
- Demo of AutoPilot
- Demo of Jumpstart
- Demo of A21 Labs Language Models

< Section Header
< Table of Contents

Low-code machine learning from AWS

Amazon SageMaker

PREPARE → Data exploration, data preparation, and feature engineering

BUILD → Model development

TRAIN & TUNE → Model training and tuning

DEPLOY & MANAGE → Live or batch predictions: model deployment, hosting, and monitoring

Amazon SageMaker Data Wrangler
A faster, visual way to aggregate and prepare data for machine learning

Amazon SageMaker Autopilot
AutoML capability that automatically prepares your data, as well as builds, trains, and tunes the best machine learning models for your tabular datasets

Amazon SageMaker JumpStart
A model hub of pretrained and easily tunable state-of-the-art models for computer vision and natural language processing tasks

aws

How SageMaker Data Wrangler works

LOW-CODE TOOL FOR PREPARING DATA FOR ML

Amazon SageMaker Data Wrangler
A faster, visual way to aggregate and prepare data for machine learning

Connect & import data
Import data from a wide variety of data sources in various supported formats

Get insights on data quality
Detect data quality issues & bias with automatic analyses & visualizations

Cleanse & enrich
Perform feature engineering with 300+ built-in data transforms
Also bring your own custom transforms

Analyze and visualize
Leverage built-in analyses like target leakage, quick model, and more to understand your data

Export prepared data to train and deploy models
Export prepared data to train & deploy ML models via SageMaker Autopilot, notebooks, or code scripts

How SageMaker Autopilot works

LOW-CODE TOOL FOR BUILDING, TRAINING, AND TUNING MACHINE LEARNING MODELS

Amazon SageMaker Autopilot
A faster, visual way to train, tune, and deploy ML

Upload data and select target
Load tabular data from S3 and select target column for prediction

Automatic model creation
Automatic feature selection, feature engineering, and selection of algorithms and training modes

Model leaderboard
Review model leaderboard to deploy the best model endpoint for inferencing

Full visibility and controls
Data exploration, model insights, and explainability reports

Evaluate, deploy, and register models
Auto-deploy or one-click deploy Autopilot model; easily build training, evaluation, and deployment ML workflows

Experimentation → Customize & optimize → Move to production

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SageMaker Data Wrangler and SageMaker Autopilot in ML workflow

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws

AIM321

Productionize ML workloads using Amazon Sagemaker MLOps featuring NatWest

Speakers

- Shelbee Eigenbrode, Principal AI/ML Specialist, WW MLOps Lead, AWS
- Greig Cowan, Head of Data Science, NatWest Group
- Usman Anwer, Principal Product Manager, AWSRelevant

Announcements

❖ n/a – although new improvements are shown.

Takeaways

❖ NatWest Case Study is best section of this session. How to scale your MLOps w/Vision, Metrics and Outcomes.

Azure

❖ Azure ML standardizes on the Mlflow open-source framework for MLOps.

What is MLOps

- The process of continuously delivering high-performance ML Models at scale. DevOps for Machine Learning.
- People, Process, Tools
- Overall, a weak definition of MLOps. A better overview can be found on Nvidia's site here.

How does Sagemaker Support MLOps

- This session focuses on the Model Lifecycle
  - Project Environment Creation
  - Model Training
  - Model Registry
  - Model CI/CD Pipelines
  - Model Deployment
  - Model Monitoring

What's new in SageMaker for MLOps

- Pipeline (CI/CD) Improvements
  - Local Mode: Helps to run, test pipelines locally, to reduce cycle time and costs during pipeline creation. Deploy to cloud when ready.
  - AutoML Support: Allows the pipelining of AutoML jobs in workflow.
  - Cross-Account Sharing: Coordinate pipelines across enterprise accounts
  - SDK Simplifications: Improved Python SDK Capabilities
- Model Monitor
  - Batch Inference: New capability to Monitor Batch Inference Models
  - Demo of Batch Inference Model Monitoring

Scaling MLOps

- Initial, Repeatable, Reliable, Scalable

NatWest Case Study

- Challenges around People, Process, Data, Technology.
- Established Metrics to solve these for Faster Time to Value
  - Faster delivery of end-to-end solutions (> 12 mos) => (< 3 mos)
  - Simplified discovery and access to data (> 5 Days) => (< 1 day)
  - Simplified ML Model Route to Production (> 6 mos) => (< 2 weeks)
  - End User self-service environment creation (> 50 days) => (< 2 days)

< Section Header
< Table of Contents

MLOps use cases

CONTINUOUSLY DELIVER HIGH-PERFORMANCE ML MODELS AT SCALE, FASTER

Develop

- Provision environments
- Perform experiments

Deploy

- Automate model training
- Package and test models
- Deploy and monitor models

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Press Esc to exit full screen

Scaling MLOps: Maturity framework

	Initial	Repeatable	Reliable	Scalable
KPI	Time to POC	Model deployment time	ML defect rate	MLDC cycle time
Capabilities	• Access to resources • Standardize experimentation environments & processes • Track experiments	• Automate governed access to resources • Automate model-build and deployment tasks • Standardize source-code repositories • Centralize management of models	• Integrate full CI/CD practices into ML pipelines for traceability, improved quality, and lineage tracking	• MLDC is reproducible through standardized templates shared across teams

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NatWest Group challenges before MLOps

NatWest Group

People	Process	Data	Technology
Lack of correct training, capabilities	Manual handovers, governance, and long time to value	Data discovery and access difficult	Fragmented and out-of-date environments

Our MLOps vision

1. Automation & standardized patterns
2. Optimized governance
3. Simplified data access
4. Modern tech stack
5. Faster time to value

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Adoption and engagement

BUILDING A LONG-TERM CAPABILITY ACROSS THE ENTERPRISE

NatWest Group

Training

Trained 100s of data scientists, engineers and leaders across >700 AWS courses and Immersion Days

Dedicated support

Established well-resourced (12 FTE) and capable enablement team…

…gathering continuous feedback

Visible metrics

Continually track and report on KPIs related to speed of setup, training completion, use-case success and data access

User champions

Early adopters identified and given priority support to validate and test the capability

Reach far and wide

Project teams and use cases from across the organization, e.g.,

Fraud
Financial crime
Finance
Audit
Credit
Climate
Pricing
…

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key takeaways

NatWest Group

Hearts & minds

Create a compelling and business-led vision for MLOps

Adoption & engagement support for the user community

Build for complexity

Data scientists & engineers are a smart bunch!

Consider orchestration beyond SageMaker

Bring your own containers & templates

Build for flexibility

MLOps is a journey that you won't take in one step

Deliver value quickly to users and be able to respond promptly to new requests

Operating model

Consider carefully how your users will build ML-driven products

Centralized vs. federated

Standardization frees devs to focus on the value-add

Legacy tech

Build your system to work with legacy estate

Working between cloud and on-premises environments

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

A metric-based approach focused on business value

NatWest Group

Metric	Where NatWest was…	Where NatWest is now
Faster delivery of end-to-end solutions	~12 months	< 3 months
Simplified discovery and access to data	~5 days	< 1 day
Simplified ML-model route to live	3-6 months	< 2 weeks
End user self-service environment creation	> 30 days	< 2 hours

* Other metrics also used during project delivery

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AIM332

WaFd bank delivers enhanced self-service with AWS conversational AI

Speakers

- Fernando Egea, Global Vice President, AI Specialized Sales, Talkdesk
- Dustin Hubbard, CTO, WaFd Bank
- Shivi Gupta, Product Manager, Amazon

Announcements

- n/a

Takeaways

- WaFd is using Lex for both IVR and Chatbots. This demo illustrates interaction w/Talkdesk.

Azure

- Azure has Bot Maker, in addition to Cognitive Services to support Speech to Text, Text to Speech. Azure qna maker is being sunsetted in favor of Cognitive Services for language.

Improve Service with CAI

- Conversational AI Solutions
  - Improve Customer Engagement
  - Enable Omni-Channel Experience
  - Uncover insights to improve quality
  - Lower costs with automation
- How Lex Supports CAI
  - Capabilities: Custom entity recognition, Enhanced speech transcription, Intelligent Dialog, Contextual Resolution
- Lex Builder Tools
  - Visual Builder – Manage complex conversation paths dynamically using a no-code interface
  - Automated Chatbot Designer – Expedite chatbot design processes through automated analysis of transcript data
  - Pre-built Solutions – Kickstart building a personalized, conversational experience for multiple industries.

Talkdesk

- Overview of Talkdesk – Cloud Contact Center Software
  - Uses Voice Biometrics to identify caller
  - Utilizes Lex for Chatbots / Can integrate other bots built by their customers if need..
  - 50% of challenge in Chatbots is after go-live, post training is always needed.

WaFd Case Study

- Migrated traditional IVR to Cloud with Transcribe / Lex / Polly
- Leveraged a lot of Lex Intents built for IVR for Chatbot
- Talkdesk allowed WaFd to bring existing solutions to Contact Center Platform.
- Roadmap is to add more sentiment analysis, support more transaction types and channels (sms, WhatsApp)

< Section Header
< Table of Contents

Use AI to improve customer experience

CHALLENGES

- Long hold and resolution times
- Single communication channel
- Lack of insights on customer pain points
- High operational expenses

Conversational AI solution

BENEFITS

- Improve customer engagement
- Enable omni-channel experience
- Uncover Insights to improve service quality
- Lower costs with automation

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon Lex

CONVERSATIONAL AI FOR VOICE AND TEXT INTERFACES

- Enables sophisticated conversations
- Easy to use
- Provides a scalable omni-channel experience

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Sophisticated conversations

INTERACTION TYPES

Frequently asked questions

- They ask about which credit cards are accepted
- A virtual agent answers their question

Transaction fulfillment

- They ask to renew their driver's license
- They renew it with a self-service virtual agent

Intelligent routing

- The agent assists them in registering a custom vehicle
- The virtual agent routes them to an appropriate agent queue
- The virtual agent verifies their identity

A constituent calls a DMV contact center

They ask to register their custom vehicle

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Popular Amazon Lex use cases

Self-service

Deflect calls with human-like virtual agents and IVRs

Intelligent routing

Identify callers' intent and route to an appropriate agent queue

Information capture

Collect key information to assist agents in resolving questions efficiently

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Talkdesk AI
Jobs-to-be-done
and the customer journey
Increase customer self-service rate
Identify causes of customer issues
Help agents resolve issues correctly and quickly
Customer Journey
CUSTOMER
With a query
AGENT
Assisted by AI automation
Self-service automation
Agent assisted automation
Low
Resolution Cost
RESOLUTION COST
High
Resolution Cost
Amazon Lex
Identity
Virtual Agent Voice
Virtual Agent Digital
Agent Assist
Quality Management
Knowledge Management
Interaction Analytics
AI Trainer Human-in-the-Loop
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Serving customers' banking needs
Customers
Channels
Branches
Call centers
Chat
Mobile app
Website
SMS
Social media
Self-help
Human-help
Common topics
Account opening
Loan origination
Account servicing
Account balances
Money movement
Fraud inquiry
Complaints
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Architecture
AWS Cloud
Agent
Amazon API Gateway
AWS Lambda
Amazon DynamoDB
Contact
Amazon API Gateway
AWS Lambda
Amazon Lex
AWS Lambda
AWS WAF
Browser / Contact Center
QNA Bot (next page)
talkdesk
WaFd Bank
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Business Applications

- BIZ213 – Serve your customers better with AWS Supply Chain

< Table of Contents

BIZ213

Serve your customers better with AWS Supply Chain

Speakers

- Sudarshan Seshadri, Director, Amazon
- Clifford Siegel, EVP Global Supply Chain, Lifetime Brands Inc
- Roberto Espinosa, Sr. Product Manager, Amazon Web Services

Relevant Announcements

- New Service – Supply Chain

Takeaways

- Supply Chain is based on "modern" event-driven architecture that collects the data exhaust for all processes, and uses a "Problem/Solver" method to avoid disruptions.

Azure

- This probably delves into Microsoft Dynamics territory, but it's out of my knowledge zone.

Traditional Supply Chain Challenges

- Siloed, point to point solutions prevent full visibility across all processes and makes it difficult to coordinate actions among different suppliers.

How Supply Chain Accelerates Innovation

- Easily connect data across systems
- Quickly harmonize into a unified view
- View ML-powered insights and risk alerts
- Accelerate migration with recommendations
- Scale with pay-as-you-go pricing.

Overview of Supply Chain

- Supply Chain Data Lake
  - Purpose built to aggregate data form disparate sources using out-of-the-box connectors
- ML Powered Data Association
  - Creates a Harmonized Data Model by using NLP to create associations
- Intelligence Layer for Insights
  - Connects Demand, Supply, Inventory, Shipping information together to provide actionable insights, and risks for potential disruption
- Purpose Built Capabilities
  - Everything is surfaced by persona-based interfaces that allows collaboration across stakeholders to resolve potential risks

Demo of Supply Chain Capabilities

- Supply Chain Insights
  - Aggregates data from disparate systems and consolidates to harmonized supply chain data lake.
  - Generate Insights such as Inventory Risk Alerts or Lead Time deviations
  - Allows execution to resolve potential disruptions.
  - Create "Watchlists" to surface Insights
  - Collaboration is done via a Kanban style task board
  - Resolutions are Scored based on your parameters based on locality, risk, and sustainability.
- Supply Chain Demand Planning
  - ML-powered demand forecasts
  - Demand sensing, actionable insights
  - View, override, and publish forecasts
  - Forecast explainability
  - Automatic accuracy metrics

Lifetime Brands

- Experience with Supply Chain
- How to better position supply and anticipate orders coming from customers. Perform quick scenario planning.
- Found a significant increase in statistical baseline forecast over parallel system.

< Section Header
< Table of Contents

AWS Supply Chain accelerates innovation

- Easily connect data across systems
- Quickly harmonize into a unified view
- View ML-powered insights and risk alerts
- Accelerate mitigation with recommendations
- Scale with pay-as-you-go pricing

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Providing the complete solution

- Purpose-Built Capabilities
- Intelligent Insights
- ML-Powered Data Association
- Supply Chain Data Lake

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Supply Chain Insights

- Enterprise inventory visibility
- Inventory risk and lead time deviations
- Sustainability-aware stock rebalance recommendations
- Notifications and collaboration

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Supply Chain demand planning

- ML-powered demand forecasts
- Demand sensing, actionable insights
- View, override, and publish forecasts
- Forecast explainability
- Automatic accuracy metrics

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Compute & Containers

- CMP225 – What's new in Amazon EC2
- COM207 - Competition of the modern workloads: Serverless vs Kubernetes on AWS
- COM301 – Unleash developer productivity with infrastructure from code
- CON210 – What's new and what's next with Amazon ECS

< Table of Contents

CMP225

What's new in Amazon EC2

Speakers

- Martin Yip, Senior Manager, Product Marketing, AWS
- Art Baudo, Principal Product Marketing Manager, AWS

Relevant Announcements

❖ New instances including those based on Graviton3

Takeaways

❖ AWS Keeps improving silicon and computing power

Azure

Overview

- 20 minutes of marketecture
  - Great Slide covering the types of migration

What's New

- New – Torn Write Protection
  - Improves database performance with EC2 I4i instances and EBS
- New – Nitro Enclaves Enhancements
  - Create isolated compute to protect and process highly sensitive data
- New – R7iz instances
- New – Graviton3

< Section Header
< Table of Contents

Amazon ECS + AWS Fargate

Look ma, no servers!

Serverless
Managed by AWS, no EC2 instances to provision, scale, or manage – simple operations model

Secure by default
Isolated, patched, and compliant for running the most sensitive workloads

Savings
Elastic on-demand pricing model with granular resource allocations giving customers the ability to procure only the vCPU and memory needed when they need it without the excess

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NEW!

AVAILABLE NOW

AWS Nitro Enclaves Enhancements

CREATE ISOLATED COMPUTE ENVIRONMENTS TO FURTHER PROTECT AND PROCESS HIGHLY SENSITIVE DATA

GRAVITON
Nitro Enclaves is now supported across the majority of Graviton, Intel, and AMD-based Amazon EC2 instances

KUBERNETES
Use Kubernetes tools to orchestrate, scale, and deploy enclaves from a Kubernetes pod

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon EC7 R7iz instances

PREVIEW

NEW!

High-frequency memory-optimized instances powered by 4th generation Intel Xeon Scalable processor

- Up to 128 vCPU, up to 1 TiB of memory to provide up to 2.6x more vCPU and memory compared to comparable high frequency instances
- Up to 20% higher performance when compared to comparable high frequency instances
- First x86-based EC2 instance to use DDR5 memory and deliver up to 2.4x higher mem bandwidth over comparable high frequency instances
- Designed for workloads such as front-end Electronic Design Automation (EDA), relational database workloads with high per-core licensing fees, and financial, actuarial, and data analytics simulation workloads

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Graviton3

Improved Performance

Up to 25% higher compute performance and 2x higher floating-point to accelerate compute-intensive workloads

Faster Memory

Supports DDR5 memory to provide 50% more memory bandwidth over DDR4 memory

Energy Efficient

Consumes 60% less power for the same performance compared to other CPUs

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

COM207

Competition of the modern workloads: Serverless vs Kubernetes on AWS

Speakers

- Niv Yungelson, DevOps Leader, Melio
- Shimon Tolts, Co-Founder & CEO, Datree

Relevant Announcements

- N/A

Takeaways

- Depends on your needs. [Editor – Go Serverless].

Azure

- N/A

Overview

- Choosing a serverless approach vs a container approach. This is a fun session for a side-by-side comparison of serverless vs Kubernetes
- Each side debated serverless vs Kubernetes on
  - Maintainability
  - Cost
  - Scalability
  - Developer Experience
  - Ecosystem
  - Monitor and Logging
  - Security

< Section Header
< Table of Contents

Maintainability – Pros and cons

Serverless

⛔ Serverless still mean servers (for someone) to manage

✅ You no longer have to provision and maintain servers to run your applications, databases, and storage systems

✅ No hardware to patch, AWS is taking care of it for you

Kubernetes

⛔ You have to keep up with a short release cycle of 4 month

⛔ Some updates require configuration files changes

✅ You have to be up to date with latest versions that are more secure

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cost – Pros and cons

Serverless

⛔ Pay as you go – you pay for each new function and each run

✅ Pay as you go – you don't pay when the service is not running

✅ Save money in the form of personnel time

⛔ Harder to predict the cost

Kubernetes

✅ Predictable

✅ Pay-per-use

✅ Spot instance supported

⛔ Have to optimize for type/size of workload

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Scalability – Pros and cons

Serverless

⛔ Can't scale over AWS's hard limits

✅ Easy to scale horizontally and vertically

✅ No need to worry about physical servers restrictions

Kubernetes

⛔ Scaling is the admin's responsibility

✅ Automatic autoscalers available by the community and AWS

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Developer experience – Pros and cons

Serverless

⛔ Cannot be agnostic to the infrastructure

✅ Implementing actual DevOps methodology

Kubernetes

⛔ Might be complex to orchestrate a big env

✅ Easy to run locally

✅ Easy to replicate envs

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Ecosystem – Pros and cons

Serverless

- ⛔ Vendor lock-in
- ✅ More and more solutions become compatible with serverless
- ✅ Many serverless services in AWS

Kubernetes

- ✅ Part of the CNCF ecosystem
- ✅ Standardized protocols and open communication

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Pros and cons

Serverless

- ⛔ You cannot connect to a physical server
- ✅ No need to monitor and take care of servers

Kubernetes

- ✅ Supports all standardized monitoring protocols (OpenTelemetry/etc.)
- ✅ Natively integrated with Amazon CloudWatch
- ⛔ Has multiple layers and objects that need to be monitored

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Security – Pros and cons

Serverless

- ⛔ Less defined perimeter
- ✅ No system vulnerabilities to patch and worry about
- ✅ Need to consider mostly package and code vulnerabilities

Kubernetes

- ✅ Vendor neutral
- ✅ Short release cycle

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

COM301

Unleash developer productivity with infrastructure from code

Speakers

- Jeremy Daly, CEO, Ampt Web Services, Inc

Relevant Announcements

- N/A

Takeaways

- Thought provoking talk. Infrastructure from Code helps create a self-learning adaptive infrastructure. [Editor: It's a big idea that will ultimately lead to more configuration and complexity.]

Azure

- N/A

< Section Header
< Table of Contents

How building Apps Changed

- It's no longer applications running on one server, instead modern cloud native architecture is a mesh of services that must be woven together easily, and environments that need to be consistent.
- We create/configure these systems through Infrastructure as Code. There's the Declarative Approach (Terraform, Cloud Formation), need to learn a Domain Specific Language, or the Imperative Approach (Pulumi, CDX)

Benefits / Limitations of Infrastructure as Code

- Benefits
  - Deals with complex environments
  - Consistency, avoids manual human intervention
  - Reproducible environments
  - Improved Transparency and collaboration
- Limitations
  - Need both code and configuration
  - Mismatch between declarative nature of IaC and dynamic nature of systems at runtime
  - Code is coupled to your architecture

Infrastructure from Code

- Idea: Instead of writing infrastructure as code, we infer infrastructure from code based on the application logic and provision the optimal cloud infrastructure.

Favorite Quotes

- "Cloud software is no longer written, it is composed."
- Can move beyond just cloud components, but also be self-auditing, self-managing and more, based on your application code.

The evolving ecosystem IfC and Adaptive Architecture

- Self-learning so the system can modify itself, availability and failover, re-index databases.
- Why people believe this is the right approach

Key takeaways & further reading

- Cloud Solutions are Composed
- IaC tightly couples code to Infrastructure
- IfC is an alternate approach decouples business logic from underlying infrastructure
- Checkout Infrastructure from Code (IfC)

Why Infrastructure as Code?

- Configurations became complex
- Avoid drift from human intervention
- Reproducibility of environments
- Improved transparency and collaboration

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Limitations of Infrastructure as Code

- Duplication and drag as facets are expressed in both code and configuration
- Impedance mismatch between the declarative nature of IaC and the dynamic nature of systems at runtime
- Control-plane specific instructions create static architecture
- Configuration over code (bifurcation of business logic)
- Code is coupled to your architecture

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Write some application code

```
import express from "express";
import cors from "cors";
import { DynamoDBClient } from "@aws-sdk/client-dynamodb"

const app = express();
const ddbClient = new DynamoDBClient();
app.use(cors());

app.get("/user/:id", async (req, res) => {
  const params = {
    TableName: "Users", Key: { pk: { N: req.params.id } }
  };
  const user = await ddbClient.send(new GetItemCommand(params));
  res.send({ user })
})
```

Looks like you're creating an API with a GET route

And you need GetItem permissions to a DynamoDB table named "Users"

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Interpretive approach

- Resources are inferred by the application code
- Permissions are assumed based on service interactions
- Behavioral semantics are declared by event handlers
- Inter- and intra-service communications are automatically mapped
- Code is decoupled from the underlying infrastructure

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

When you understand the code . . .

self-provisioning

self-instrumenting

self-documenting

self-testing

self-auditing

self-managing

self-optimizing

self-upgrading

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Adaptive architecture

- Auto scaling, automated failover, built-in resiliency
- Self-indexing databases
- Reactive telemetry/observability data
  - Memory
  - Timeouts
  - Latency
  - Service dependencies
  - Predictive scaling

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Why people believe in this approach

klotho	Decoupling code from infrastructure enables "adaptive architectures"; code is the spec and doesn't change, but the infrastructure can
shuttle	We are an orchestrator of AWS resources using the Rust API – IaC produces outputs that your code needs to wire together; Rust fills in the gap
encore	IfC enables tools that can understand the structure of your app, and reason about how it interacts with infrastructure and other services
Modal	Everything needs to run in the cloud, not on your laptop – ephemeral environments with fast feedback loops
nitric	It's about portability; we want to help developers delay making decisions until they know what resources their app needs

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key takeaways

- We build modern cloud apps by composing multiple cloud services
- Building multi-service, native cloud applications is really hard
- Infrastructure **as** Code tightly couples your app code to static architecture
- Higher-level abstractions can still provide control while reducing redundancy
- Infrastructure **from** Code can:
  - Accelerate development
  - Decouple business logic from underlying infrastructure
  - Understand the structure of your application
  - Enable platforms to optimize over time

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

CON210

What's new and what's next with Amazon ECS

Speakers

- Akhilesh Thipparthi, Vice President, Goldman Sachs
- Nick Coult, General Manager, Amazon ECS, AWS

Relevant Announcements

❖ New – ECS Service Connect

Takeaways

❖ Amazon runs on ECS, continue to grow.

Azure

❖ Azure Container Services

The Basics

- ECS is a serverless container orchestrator. You can orchestrate containers on EC2 or Fargate.
- ECS Fargate is a serverless compute engine for containers
- ECS Anywhere allows to orchestrate containers on on-premises compute.
- ECS Faster, Better, Cheaper
- Watch – ECS powers Amazon

What's New

- New – ECS Service Connect
  - Built in Service Discovery, Visibility and traffic telemetry for service-to-service communications, higher application availability with resilient network.
- New – Larger tasks on Fargate

How are Customers using it

- Watch – Goldman Sachs
  - Discusses various architecture implementations and best practices with use cases.

What's Next

- Watch – What's Next
  - Touched upon Deployment Improvements like native support for Blue/Green deployments, Elastic File System, Service Connect.

Favorite Quotes

- C: "Are we going to get to big or ECS?"
  A: "Are you ever going to be bigger than Amazon?"

< Section Header
< Table of Contents

Amazon ECS powers Amazon

Amazon ECS is a building block for many services at Amazon

Tested for security, reliability, availability, and scale

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ECS Service Connect simplifies inter-service communication within and across clusters

- Built-in service discovery
- Visibility with traffic telemetry for service-to-service communication
- Higher application availability with resilient network

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Larger tasks on Fargate

More CPU and Memory available.

New! Graviton2

GPU on Roadmap

AWS Fargate

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Upcoming features

Deployments
- ECS native blue/green and canary deployments

Storage
- FSx for Lustre support

Orchestration
- Task scale-in protection (launched!)
- Task def deletion

Performance
- Faster cluster autoscaling
- Ease of managing EC2 capacity and AMI upgrades

Richer Fargate
- GPU workloads
- Improved performance

Developer Experiences
- New ECS console
- AWS Copilot enhancements

Networking
- ECS-A LB support
- Service Connect encryption

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data

- DAT218 – What's new with database migrations
- DAT219 – What's new in DocumentDB

< Table of Contents

DAT218

What's new with database migrations

Speakers

- John Winford, Sr. Manager, AWS
- ryan shevchik, Principal Solutions Architect, AWS

Relevant Announcements

- New – DMS Schema Conversion

Takeaways

- The database migration service is a great tool for data migration and data movement.

Azure

- Azure has the Database Migration Service.

Why Migrate

- Typical Cloud Reason, Faster, Better, Cheaper
  - Great Slide covering the types of migration

Migration Tools & Services

- Assess, Mobilize, Accelerate

Migration Database Service

- A Swiss Army Knife for Database Migrations
- Can also used to migrate from on-prem – cloud – on-prem
- Database stays live even when migrating
- It can also be used for data movement, not just migrations.
- DMS Fleet Advisor for analyzing migrations at scale
- Schema Conversion Tool for migrating between disparate databases or for refactoring

New Capabilities

- New – Since last year
  - DMS – Babelfish target, Db2 mainframe, SQL Server read replica, VPC endpoints, Pre-flight enhancements, Enhanced logging and Time Travel, AL2 and Gen6 instances, IPv6
  - SCT – Babelfish assessments, Snowflake schema conversion, Azure Synapse, Amazon Redshift Serverless, Google BigQuery to Amazon Redshift, SSIS to AWS Glue Studio, Informatica to Amazon Redshift redirector, Improved code conversion
- New – AWS DMS Schema Conversion
  - Allows you to do all your migrations in one place.

Demos

- Demo – DMS Schema Conversion
- Demo – DMS Automation

< Section Header
< Table of Contents

Modernizing leads to maximum innovation velocity and optimal total cost of ownership

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Migration tools and services

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Database Migration Service (AWS DMS)

SECURELY MIGRATE YOUR DATABASES AND DATA WAREHOUSES TO AWS WITH EASE & MINIMAL DOWNTIME

Managed migration

Automated migration service will discover, assess, convert, and migrate your database and analytics workloads to AWS

Just a few clicks to start the process

Breadth of options

Migrate data to and from most widely used commercial and open-source databases

Supports SQL, NoSQL, text-based, and data warehouse targets

High availability

Multi-AZ option; highly resilient and self-healing with continuous replication and monitoring for minimal downtime

Low cost

Self-service tool

Only pay for compute resources and additional log storage used

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS DMS supported sources and targets

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS DMS Fleet Advisor

AUTOMATICALLY INVENTORIES AND ASSESSES ON-PREMISES DATABASE AND ANALYTICS SERVER FLEETS

Identify databases to migrate at scale with minimal effort

Save on costs associated with migration planning and migrating workloads

Build a migration plan in hours instead of weeks

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Schema Conversion Tool (AWS SCT)

CONVERT YOUR COMMERCIAL DATABASE AND DATA WAREHOUSE SCHEMAS TO AWS

Automated inventory and migration recommendations to accelerate migrations

Converts schema and code objects to target engine

Supports a wide variety of source and target databases and data warehouses

Extracts and migrates data warehouses to Amazon Redshift

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Common customer use cases

Migrate

Migrate business-critical applications and data warehouses

Migrate from NoSQL to SQL, SQL to NoSQL, or NoSQL to NoSQL

Upgrade to a minor/major version

Consolidate and archive data

Replicate

Create cross-Region read replicas

Hydrate your data lakes/ data warehouses

Replicate to streaming platform

Modernize

Move to AWS modern managed databases like Aurora, DynamoDB, and Amazon Redshift for faster innovation

Break free from legacy databases to save, grow, and innovate

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

New since last year

AWS DMS

- Babelfish target
- Db2 mainframe
- SQL Server read replica
- VPC endpoints
- Pre-flight enhancements
- Enhanced logging and Time Travel
- AL2 and Gen6 instances
- IPv6

AWS SCT

- Babelfish assessments
- Snowflake schema conversion
- Azure Synapse
- Amazon Redshift Serverless
- Google BigQuery to Amazon Redshift
- SSIS to AWS Glue Studio
- Informatica to Amazon Redshift redirector
- Improved code conversion

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS DMS schema conversion

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

DAT219

What's new in DocumentDB

Speakers

- Vin Yu, Super Product Manager, Amazon
- Rashim Gupta, Senior Manager, Product Management, Amazon Web Services

Relevant Announcements

❖ New – DocumentDB Elastic Clusters

Takeaways

❖ DocumentDB now allows for sharding..

Azure

❖ Azure has CosmosDb which has a MongoDB compatible API and supports sharding.

Overview

- DocumentDB
  - Fast and Scalable
  - Fully Managed
  - MongoDB API Compatible

What's New

- New – Regions and Instance availability
- New – Query Capabilities – Geospatial Indexes, Aggregation operators, Decimal128
- New – Enterprise Management – Dynamic volume resizing, Fast database cloning, Query auditing, Performance Insights (Demo)
  - Great Slide covering the types of migration

DocumentDB Elastic Clusters

- New – DocumentDB Elastic Clusters
  - Compatible with MongoDB APIs for sharding
  - Millions of reads and writes
  - Up to 300,000 connections
  - Easy and Fast scaling
- Pricing – No upfront costs only pay by Compute, Storage, and Backups used.

Demos

- Demo - Getting Started & Product Tour

< Section Header
< Table of Contents

Amazon DocumentDB

SCALABLE, HIGHLY DURABLE, AND FULLY MANAGED DATABASE SERVICE FOR OPERATING MONGODB API-BASED WORKLOADS

Fast and scalable

- Auto scaling storage and I/O
- Storage scales to 64 TiB
- Scale compute in minutes
- Scale out to 15 replicas for millions of reads
- Globally distributed

Fully managed

- Built-in high availability
- Backups enabled by default
- Durable by default
- Security best practices by default
- Automatic patching
- Monitoring and alerting

MongoDB API compatible

- Applications, drivers, and tools can be used with little or no change
- Supports hundreds of APIs, operators, and stages
- Continually working backward from customers to deliver the capabilities they need

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

New Regions and instance availability

Graviton2 instances (t4g, r6g) availability

Asia Pacific (Seoul)

South America (São Paulo)

Europe (Paris)

Region launches

Milan, Italy

Beijing, China

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Query capabilities

Geospatial indexes

$nearSphere, $geoNear, $minDistance, $maxDistance, $geoWithin, $geoIntersects

Aggregation operators

$mergeObjects and $reduce

Decimal128

Provides 128 bits of decimal representation supporting 34 decimal digits of precision

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Enterprise management features

Dynamic volume resizing

Storage space automatically increases up to a maximum size of 64 tebibytes (TiB) when you need more storage and decreases automatically when you delete data.

Fast database cloning

Create clones to enable fast creation of a new cluster that has the same data as the original. Database cloning is faster than restoring a snapshot and requires no additional space at the time of creation.

Query auditing (now with DML)

DocumentDB will record Data Manipulation Language (DML) events to Amazon CloudWatch Logs for monitoring and analysis. This includes events corresponding to insert(), update(), delete(), find(), aggregates, and more.

Performance Insights (preview)

Performance tuning and monitoring feature that helps you quickly assess the load on your database and determine when and where to take action.

With one click, you can add a fully managed performance monitoring solution to your Amazon DocumentDB instances.

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon DocumentDB Elastic Clusters

NEW!

New type of cluster (Elastic Cluster) for Amazon DocumentDB that supports workloads with millions of reads/writes per second and petabytes of storage capacity

Elastic Clusters also simplify how developers interact with Amazon DocumentDB; no need to choose, manage, or upgrade instances!

Compatible with MongoDB APIs for sharding

Millions of reads and writes

Up to 300,000 connections

Easy and fast scaling

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How is it priced?

No upfront costs and only pay for what you use with easy-to-understand predictable pricing

Compute – vCPUs per hour

Storage – GB per month

Backups – GB per month

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Front End Web and Mobile Development

- FWM308 – What's new for frontend and web and mobile developers with AWS Amplify
- FWM309 – Design to code: Build an app frontend UI in minutes with AWS Amplify Studio
- FWM315 – Create real-time, event-driven apps with Amazon EventBridge & AWS AppSync

< Table of Contents

FWM308

What's new for frontend web and mobile developers with AWS Amplify

Speakers

- Mohit Srivastava, Senior Manager Product Management, Amazon
- Abdallah Shaban, Senior Product Manager, Amazon

Relevant Announcements

- Improvements to both Back-End and Front-End components. See New links on Left

Takeaways

- Amplify is a set of tools and framework that allows you to build web & mobile apps from same codebase, as well as easily create skeleton code. It also provide a forms like model to easily create general purpose applications

Azure

- PowerApps may be the closest low code comparison, with dotnet Maui being more code-centric

The State of Web and Mobile Development

- Front end developers have a lot of responsibility

What is AWS Amplify

- A set of tools that can be used together or independently
  - Build AWS Backend
  - Build App UI
  - Host an App (web)
  - Connect to Backend

Building fullstack web and cross-platform app

- Demo – Building out the back-end
  - New – GraphQL in Studio, Real-time data, In-app messages
- Demo – Building out the front-end
  - New – Create cloud connected forms for any API with Studio Form Builder
  - New – Hosting now supports server side rendered apps with Next.js 13
- Demo – Building a Flutter application (mobile)
  - New – Enhancements for Mobile applications

Underlying Architecture

- Brief discussion of underlying architecture

< Section Header
< Table of Contents

Frontend developers have broad responsibility

Agility
Time to market
Rapid innovation
No one-way doors

Great experience
Multidevice
Modern look/feel
Differentiated features

Production-ready
Low latency
Secure
Reliable and scalable

Technical needs

Fullstack development
Features that use device + cloud
Integrated development
Less boilerplate
Best practices built in

Web
Native-like experience
Support for hybrid client/server rendering (e.g., Next.js)

Cross-platform
One codebase
Support for Flutter, React Native, Apple cross-platform

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Amplify

USE THE PIECES TOGETHER OR ON THEIR OWN

FULLSTACK DEVELOPMENT | WEB | CROSS-PLATFORM & WEB

Build AWS backend
Amplify Studio
Amplify CLI

Build app UI (web)
Amplify Studio

Host an app (web)
Amplify Hosting

Connect to backend
Amplify Libraries
UI Components

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Build an AWS backend: What's new

NEW

- GraphQL in Studio: Query auto-generated GraphQL APIs
- Real-time data: Service-side filtering to 1/ send updates only to specific users or groups (collaboration apps) and 2/ selectively sync data (optimize network traffic and local storage)
- In-app messages: Send targeted messages to defined user segments or trigger contextual messages based on user behavior

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amplify Studio Form Builder

CREATE CLOUD CONNECTED FORMS FOR ANY API WITH NEW REACT FORM BUILDER

NEW

Forms for any API
Generate forms from an API definition or create new forms from scratch

Design visually
Save time by visually configuring common validation rules

Extend via code
Customize validation rules and form styling in code

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Fully managed hosting for server side rendered apps built with Next.js 13

NEXT.JS BENEFITS AND DEVELOPER EXPERIENCE MEET THE POWER OF AWS

NEW

Next.js benefits

Hybrid static & server rendering to support both interactivity and faster initial load times

Zero config deployment

Just connect your repo. Single developer experience across browser and server code

Integrate with backends

Works seamlessly with Amplify backends, such as Auth and Data

Amazon CloudWatch integration

Observe, monitor, and troubleshoot SSR apps

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Connect to an AWS backend: What's new

NEW

Flutter: Rewritten using Dart to create mobile and now desktop and web apps from a single Dart-based codebase

iOS: Rewritten using Swift to create iOS and now macOS apps from a single codebase. Now supports idiomatic Swift features like async/await for implementing structured concurrency.

Android: Rewritten using Kotlin with Storage and Auth improvements

Authenticator UI component: Now available for Flutter and React Native

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

FWM309

Design to code: Build an app frontend UI in minutes with AWS Amplify Studio

Speakers

- Christopher Woolum, L6 FEE, Amazon AWS
- Rene Brandel, Principal Product Manager - Technical, Amazon Web Services, Inc.

Relevant Announcements

❖ New – Amplify Studio Form Builder

Takeaways

❖ Easily get from design to code, build interactive forms with form builder, deploy and scale with Amplify hosting.

Azure

❖ PowerApps may be the closest low code comparison, with dotnet Maui being more code-centric

Amplify Studio

- Input Figma designs output React Code. See a Demo.
- Evolution of building an app
  - Build AWS Backend
  - Build App UI
  - Host an App (web)
  - Connect to Backend
- Watch – Building a Full Stack App
  - Build, Ship, and Scale with Amplify
  - New – Amplify Studio Form Builder for connected forms
  - Deploy with Amplify Hosting

< Section Header
< Table of Contents

Introducing AWS Amplify Studio

Amplify Studio

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amplify – Quick tour

Build

Configure UI components and AWS backends fast

Seamlessly connect frontends

Ship

Deploy your app frontend and backend

Scale

Use CDK to override, add, and export infrastructure

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amplify Studio form builder

BUILD, DESIGN, AND RENDER CLOUD CONNECTED FORMS FOR ANY API WITH NEW REACT FORM BUILDER

NEW

Forms for any API

Generate forms from an API definition or create new forms from scratch

Design visually

Save time by visually configuring common validation rules

Extend via code

Customize validation rules and form styling in code

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Ship faster with Amplify Hosting

- Fully managed hosting service for modern web apps
- Full-stack continuous deployments
- Easy custom domain setup
- PR previews

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

FWM315

Create real-time, event-driven apps with Amazon EventBridge & AWS AppSync

Speakers

- Venugopalan Vasudevan, Sr. Specialist SA - FE & Mobile, AWS

Relevant Announcements

- N/A

Takeaways

- If you want to build a real-time event driven app with subscriptions and GraphQL capabilities, this session will get you started

Azure

- Azure has multiple Messaging Products and a GraphQL data store that is compatible with Gremlin

GraphQL Overview

- A query language for your API, and a runtime for servicing the queries on your existing data. (Queries, Mutations, Subscription).
  - Build AWS Backend

AppSync Overview

- A serverless fully managed solution for running graphQL workloads
- Can connect to multiple back-end data sources and APIs
- Customers are using AppSync for APIs, Real-Time Applications, and offline applications.

Type of Applications

- Real-Time Scenarios
- Subscriptions
- Capabilities
  - Filtering and Invalidation
  - Event Driven Apps with Event Bridge

Demo

- Demo building a Real-Time application using Event Bridge and AppSync

< Section Header
< Table of Contents

GraphQL is a **query language** for APIs and a **runtime** for fulfilling those queries with your **existing data**

Describe your data

```
type Project {
  name: String
  tagline: String
  contributors: [User]
}
```

Ask for what you want

```
{
  project(name: "GraphQL") {
    tagline
  }
}
```

Get predictable results

```
{
  "project": {
    "tagline": "A query language for APIs"
  }
}
```

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GraphQL, the GraphQL logo and any related marks are trademarks of Facebook, Inc.

A query language for APIs

Queries

```
query GetPost {
  getPost(id: "1") {
    id
    title
    author
    date
  }
}
```

Mutations

```
mutation CreatePost {
  createPost(input: {…}) {
    id
    title
    author
    date
  }
}
```

Subscriptions

```
subscription OnCreatePost {
  onCreatePost {
    id
    title
    author
    date
  }
}
```

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS AppSync overview

ACCELERATE APPLICATION DEVELOPMENT AND IMPROVE PERFORMANCE WITH GRAPHQL AND PUB/SUB API SERVICE

Build applications faster

Frontend developers can build features without waiting for backend teams to create new APIs

Simplified data access

Interact with data from multiple microservices, data sources, and REST APIs using a single API

Improve application performance

Fetch only required data and reduce the number of network requests to fetch data for a particular view

Real-time experiences

Create engaging applications with real-time subscriptions

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Simplified data access

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What are customers building with AppSync?

Universal APIs
- Enterprise data
- Apps connect to many data sources and microservices

Real-time applications
- Chat functionality
- Fan engagement
- Location-aware notifications

Offline applications
- Consumer
- IoT/connected
- Field service
- Offline support

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Real-time apps

- Sending data (Publishing) – User, event, backend
- Receiving data (Subscribing)
- As it happens
- Push, not polling
- Milliseconds, not seconds
- 1:1, 1:many, and many:many

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Real-time use cases

- Collaboration
- Dashboards
- Leaderboards
- Media and entertainment
- Healthcare
- Live streaming
- Sports
- Geolocation
- Chat and messaging
- Gaming
- Audience interaction
- Social media
- E-learning
- Automation
- Internet of Things (IoT)
- Smart homes
- Trading and fintech
- Ecommerce

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Real-time GraphQL subscriptions

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Real-time GraphQL subscriptions

Connection management | Scalability | Fan-out | Broadcasting | Metrics

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Subscriptions filtering in AWS AppSync

- Basic filtering
  - Up to 5 arguments, defined on the client
  - AND logic only
  - No operators
  - Strict equality
- Enhanced filtering
  - Multiple logical operators (and, or, contains, le, gt, in, between)
  - Defined centrally on the API backend or dynamically from the client
  - JSON format
  - Invalidation capabilities

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon EventBridge

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Use cases

- Send specific events to one client (1:1)
  - Order fulfillment
  - Incident response
  - Location tracking
- Send events to multiple clients (fan out)
  - IoT devices
  - Partner events

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Serverless

- SVS204 – What's new with serverless
- SVS303 – Building real-world serverless applications with AWS SAM
- SVS308 – (Workshop) AWS serverless developer experience: A day in the life of a developer

< Table of Contents

SVS204

What's new with serverless

Speakers

- David Boyne, Senior Developer Advocate, AWS

Relevant Announcements

- Lambda Snap Start, Step Function Distributed Map, EventBridge Pipes

Takeaways

- Three major updates to core serverless components, as well as incremental

Azure

- Azure has Functions, Logic Apps, and Enterprise Messaging

Lambda

- Function URLs – A dedicated HTTPS endpoint for your lambda function. Guidance on when to use.
- New – SnapStart, reduce Cold Start time, currently for Java 11 Runtime
- Telemetry API – Get more information out of our Lambda (platform traces, performance)
- More space in Lambda Storage
- Lambda Power Tools to help discover and incorporate serverless best practices

Step Functions

- Workflow Studio to Visual Step Functions
- Added Intrinsic Functions to help reduce code
- New – Distributed Map
- Added updated console experience

EventBridge

- Inbound WebHooks
- Salesforce Integration with EventBridge (Bi-directional integration with Salesforce)
- New – Pipes
- Create schedules with Scheduler

Serverless Application Model

- Accelerate development with sam build, sam sync, and sam logs
- Connectors –allow developers to define how events flow between two resources as well as permissions.

< Section Header
< Table of Contents

AWS Lambda function URLs

A NEW WAY TO INVOKE AWS LAMBDA FUNCTIONS WITH AN HTTP REQUEST

- A dedicated HTTPS endpoint for your Lambda function
- A separate resource in AWS CloudFormation
- Generates a unique URL to invoke an AWS Lambda Function
- Can point to $LATEST or user-defined alias

https://<url-id>.lambda-url.<region>.on.aws/

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

When to use Lambda URLS?

	Lambda URLS	API Gateway
Auth Type	IAM	API Key, IAM, Amazon Cognito, Lambda (Custom)
HTTPS	Yes	Yes
Max Timeout	15 minutes	29 seconds
Validation	No (In Code)	Yes
SDK generation	No	Yes
Cost	Free	From $3.50/million for first 333 million
AWS WAF	No	Yes

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Lambda SnapStart

OPTIMIZES AWAY THE INITIALIZATION OVERHEAD OF CREATING AN EXECUTION ENVIRONMENT

- New Lambda feature that optimizes away the initialization overhead of creating an execution environment
- Allowing functions to start running handler function code within hundreds of milliseconds
- Initially available for AWS Lambda functions running Amazon Corretto Java 8 (java8.al2) or Amazon Corretto Java 11 (java11)
- Powered by MicroVM Snapshot technology

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Telemetry API (2022)

Receive telemetry data including logs, platform traces, and new performance metrics with new Lambda API

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

10 GB /tmp space – 2022

MORE TEMPORARY SPACE IN YOUR AWS LAMBDA EXECUTION ENVIRONMENT

Features	Description
Maximum size	10,240 MB
Persistence	Ephemeral
Content	Dynamic
Storage type	File system
Operations supported	Any file system operation
Pricing model	Included in Lambda (Charged over 512 MB)
Shared across all invocations	No

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Lambda Powertools

A SUITE OF UTILITIES FOR AWS LAMBDA FUNCTIONS

- Open source library to help discover and incorporate serverless best practices
- Logger, Tracing, Metrics, and Utilities for Lambda
- Python 2.5B invocations in 2021, growing YOY to 6.5B in 2022
- In 2022 released support for .NET (preview) and TypeScript (GA)

2.5B

↓

6.5B

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Step Functions Workflow Studio

OVER 223 SERVICES AND 10,000+ APIS

The **workflows** you build with Step Functions are called **state machines**, and **each step** of your workflow is called a **state**

When you execute your state machine, **each move** from one state to the next is called a **state transition**

You can **reuse components**, easily edit the sequence of steps or swap out the code called by task states as your needs change

AWS Step Functions Workflow Studio

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Step Functions intrinsic functions

Arrays

JSON data manipulation

Encoding and decoding

Math operations

String operations

Unique identifier generation

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Step Functions: New console experience

DEBUGGING STEP FUNCTIONS WITH THE NEW EXECUTION PAGE

- New experience to help you analyze, debug, and optimize Step Functions Standard Workflows
- Events tab that displays all the events for the selected state, with their corresponding timestamp
- Input & Output tab allows you to debug state input to output with transformations in between with Parameters
- Turn on "New execution page" to get started

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Step Functions distributed map state

- Coordinate large-scale parallel workloads in Step Functions
- Iterate over millions of S3 objects such as JSON or CSV files
- Up to 10,000 parallel executions
- Invoke Lambda or Amazon ECS/AWS Fargate for large scale on demand serverless compute
- Choose express or standard workflow

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

EventBridge inbound webhooks

- Save time integrating with GitHub, Stripe, and Twilio
- Forward events from GitHub, Twilio, and Stripe into EventBridge
- Provisions CloudFormation stack with Lambda function URL

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Salesforce integration with EventBridge

- Integrate Salesforce with AWS without any code
- Designed for event driven architecture and real-time applications

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

EventBridge Pipes

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon EventBridge Scheduler: Use case

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS SAM accelerate

sam build

- Parallel builds
- Incremental builds
- Temporary layers

sam sync

- Update code without redeploying
- Watch for changed files

sam logs

- Get aggregated logs in your local terminal
- Filter logs
- Include traces

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Serverless connectors

Connectors allow the developer to describe how data and events flow between two resources and the level of permissions required

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SVS303

Building real-world serverless applications with AWS SAM

Speakers

- Eric Johnson, Principal Developer Advocate, AWS

Relevant Announcements

- Minor announcements, all of these are covered in SVS204

Takeaways

- A great overview of using the Serverless Application Model method for developing serverless apps

Azure

- N/A

Intro to Serverless Application Model (SAM)

- SAM is made up of two parts
  - Transform Templates – Infrastructure as Code
    8 Serverless Resource Types
  - SAM CLI – Tooling for development
    App Development & Stack Management

SAM Local

- Watch and Overview SAM Local commands
  - start-api, invoke, generate-event, start-lambda

SAM Accelerate

- Build, Sync, and Logs to improve development and debugging
  - Filtering and Invalidation
  - Event Driven Apps with Event Bridge

Code Deployment

- Sam depoy, sam pipeline
  - Event Driven Apps with Event Bridge

Best Practices

- Watch Best Practices

< Section Header
< Table of Contents

AWS SAM has two parts

Transform templates

Shorthand syntax expresses resources and event source mappings; it provides infrastructure as code (IaC) for serverless applications

AWS Command Line Interface (AWS CLI)

Provides tooling for local development with rapid iteration, debugging, build, packaging, and deployment for serverless applications

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS SAM serverless resources

8 serverless resource types

- AWS::Serverless::Function
- AWS::Serverless::Api
- AWS::Serverless::HttpApi
- AWS::Serverless::SimpleTable
- AWS::Serverless::LayerVersion
- AWS::Serverless::Application
- AWS::Serverless::StateMachine
- AWS::Serverless::Connector

SAM Version 2016-10-31

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS SAM CLI

App development
- init
- build
- local
- logs
- traces
- sync

Stack management
- validate
- package
- delete
- deploy
- pipeline
- publish*

```
(p3_9_9) ➜  ~ sam --help
Usage: sam [OPTIONS] COMMAND [ARGS]...

  AWS Serverless Application Model (SAM) CLI

  The AWS Serverless Application Model extends AWS CloudFormation to provide
  a simplified way of defining the Amazon API Gateway APIs, AWS Lambda
  functions, and Amazon DynamoDB tables needed by your serverless
  application. You can find more in-depth guide about the SAM specification
  here: https://github.com/awslabs/serverless-application-model.

Options:
  --debug     Turn on debug logging to print debug message generated by SAM
              CLI and display timestamps.

  --version   Show the version and exit.
  --info
  -h, --help  Show this message and exit.

Commands:
  init      Init an AWS SAM application.
  validate  Validate an AWS SAM template.
  build     Build your Lambda function code
  local     Run your Serverless application locally for quick development &...
  package   Package an AWS SAM application.
  deploy    Deploy an AWS SAM application.
  delete    Delete an AWS SAM application and the artifacts created by sam
            deploy.

  logs      Fetch logs for a function
  publish   Publish a packaged AWS SAM template to the AWS Serverless
            Application Repository.

  traces    Fetch AWS X-Ray traces
  sync      Sync a project to AWS
  pipeline  Manage the continuous delivery of the application
(p3_9_9) ➜  ~
```

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS SAM Accelerate

sam build

- Parallel builds
- Incremental builds
- Temporary layers

sam sync

- Update code without redeploying
- Watch for changed files

sam logs

- Get aggregated logs in your local terminal
- Filter logs
- Include traces

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS SAM deploy

Deploys or updates the current stack using AWS CloudFormation

CLI flags
--guided -g
--config-env
--on-failure

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS SAM pipeline

- Generates the resource and configuration required to build a continuous integration and continuous delivery (CI/CD) pipeline
- Supports
  - AWS CodeSuite
  - GitHub Actions
  - Jenkins CI
  - GitLab

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The development cycle

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SVS308

AWS serverless developer experience: A day in the life of a developer

Speakers

- Stephen Liedig , Principal Serverless Specialist Solutions Architect, Amazon Web Services
- Tomas Mihalyi, Serverless Specialist - Ent Support, AWS

Relevant Announcements

- N/A

Takeaways

- A workshop for learning how to build using the Serverless Application Model

Azure

- N/A

Workshop

- This was a workshop, so no recording available
- You can do the workshop here https://s12d.com/aws-sde-workshop

< Section Header
< Table of Contents