

# Separating the signal from the noise:

Building next-generation compliance surveillance programs on AWS

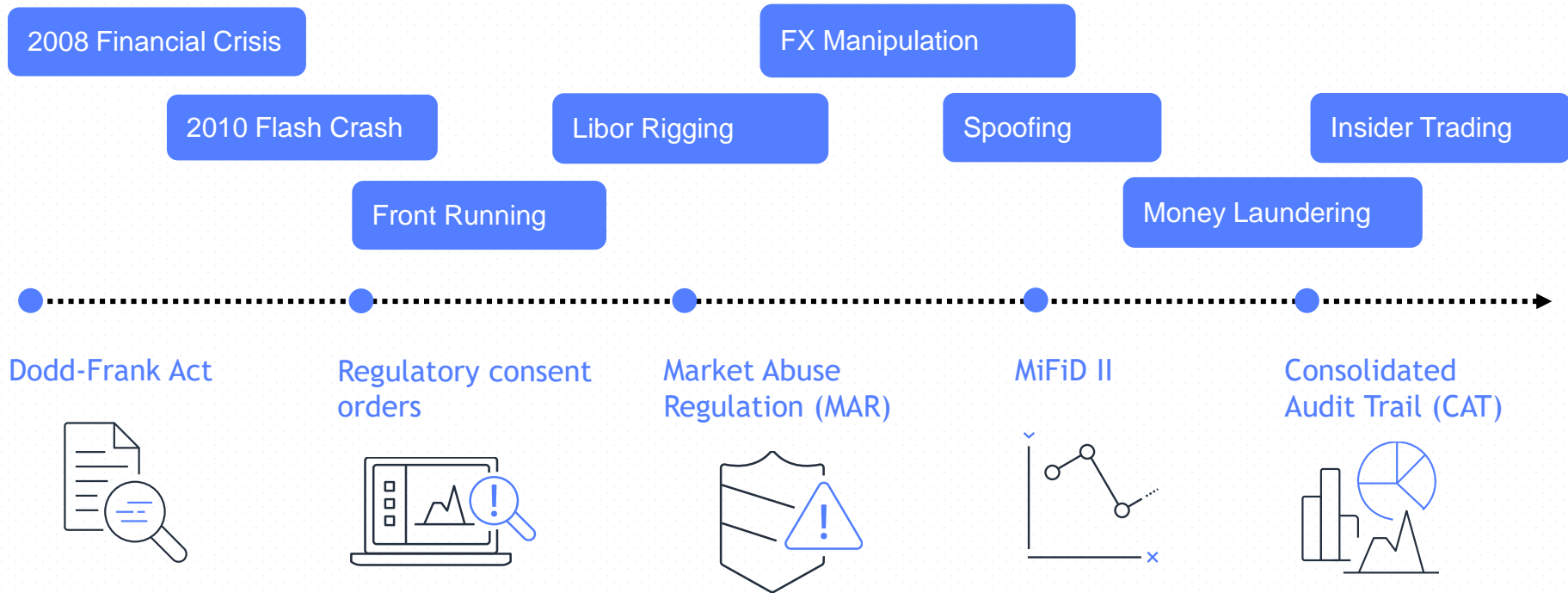
October 2018





# Today's surveillance landscape

# The Financial Services industry has seen a steady stream of regulation



# These regulations require enhancements to current surveillance programs

Longer retention periods

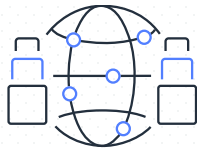
Event reconstruction

Post-trade transparency

Cross product surveillance

Alert logic validation

Global Consistency



Integrated / Holistic Surveillance



Transaction Surveillance



E-Communications Surveillance

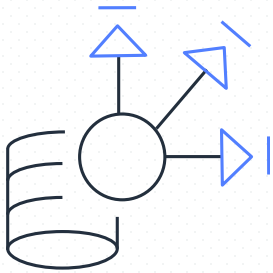


Voice Surveillance



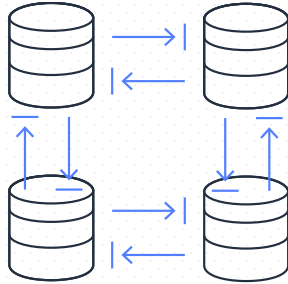
# However, many firms face challenges due to these increased requirements

**Challenge:** Process massive amounts of data



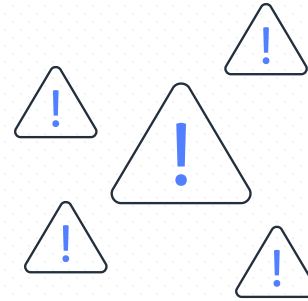
Root Cause: Infrastructure **does not scale** to meet variable and increasing volumes of data

**Challenge:** Build sophisticated surveillance reports



Root Cause: Data stored in **multiple disconnected data silos** with inconsistent extract, transform, and load (ETL) processes

**Challenge:** Large number of unproductive alerts



Root Cause: Rigid rules-based approach **does not adapt** to changing market conditions

**Challenge:** Create an integrated surveillance program



Root Cause: Agility and experimentation **limited by legacy infrastructure**

# AWS helps organizations manage these challenges



## Storage Compute

Dynamically scalable storage and compute services to process all of your data



## Data Catalog Data Warehouse

Optimized environments and frameworks to handle a broad set of big data use cases using consistent ETL processes



## Machine Learning

ML tools and frameworks to optimize alert output and extract hidden relationships and insights in the data



## Sandbox Graph Database

Easily spin up sandbox environments for testing and innovation using tools like graph databases

# FINRA enhances its markets surveillance algorithms with AWS



The Financial Industry Regulatory Authority, (FINRA) is a U.S. not-for-profit self-regulatory organization (SRO) authorized by federal law to help protect investors and ensure the fair and honest operation of financial markets. Under the supervision of the Securities and Exchange Commission (SEC), FINRA regulates the activities of U.S. broker-dealers and performs market regulation.

“ Prior to our move to the public cloud, there simply was not enough storage and processing capacity to have all the different, varying views of the data we require. But today, we can look at all of that data—every single market trade, order or price change for the past five years in seconds to minutes—without having to do any staging or manipulation. The massive scale that we have really requires us to have the flexibility to do things much quicker, with much less labor involved in the process. ”

– Steve Randich, CIO, FINRA

- FINRA ingests, stores, and analyzes data from 12 exchanges and 3,700 broker dealers in the U.S.
- It monitors 99% of equities & 70% of options activity and needed an infrastructure that could process 36 billion market events on average each day and dynamically scale to process 100 billion records on a peak day.
- One of FINRA’s 2018 regulatory priorities is cross-product surveillance of the listed equities and options markets.
- FINRA turned to AWS to solve these challenges, and the SRO is applying ML to build better surveillance algorithms that will more accurately identify potential market manipulation activities and produce fewer false positives.

# Effective surveillance architectures share common traits

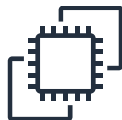
1

Scalable storage and compute



Amazon  
S3

Data lake serving as the single source of truth



Amazon  
EC2

Elastic compute that scales based on market conditions



AWS  
Glue

Centrally managed data catalog and consistent ETL processes



Amazon  
EMR

Scalable and customizable environments for development, testing, and production



Amazon  
Redshift

Data warehouse optimized for big data use cases



Amazon  
Sagemaker

Platform to easily build, train, and deploy machine learning models at any scale

2

Data lineage and effective querying

3

Data analytics



# Customer architecture: FINRA data processing

## Data lineage

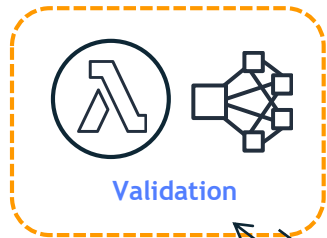
- Ability to track data provenance and end-to-end changes
- Ability to trace errors back to the root cause, or re-create and re-run processes



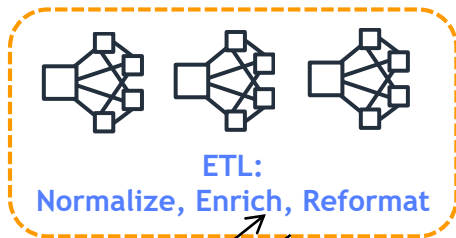
Broker Dealers Exchanges



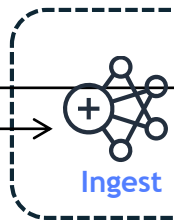
Data Files



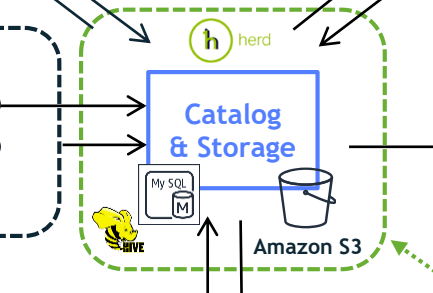
Validation



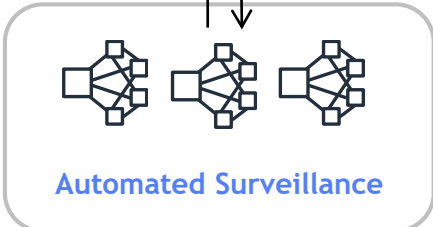
ETL:  
Normalize, Enrich, Reformat



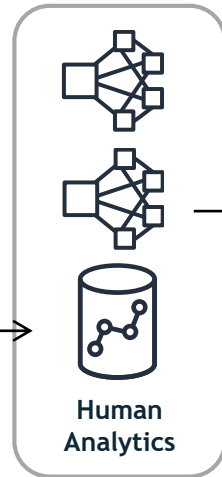
Ingest



Catalog & Storage



Automated Surveillance



Human Analytics



Analyst  
Data Scientist  
Regulatory User

## Data transformation and optimization

- Multiple, scalable on demand Amazon EMR clusters adjust to highly variable market activity

## Machine Learning and analytics

- Amazon S3 acts as a data lake feeding traditional business analytics tools but also enabling data scientists to leverage Machine Learning to provide new, more effective surveillance



# Working with the AWS Partner Network

# APN Partners can help build or enhance existing solutions



partner  
network

Whether your surveillance solutions are built in-house or through a vendor, the AWS Partner Network (APN) can help you build or enhance your surveillance program.

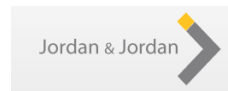
## Consulting Partners

Support assessing compliance requirements, designing surveillance solutions, and developing implementation plans.



## Technology Partners

Technology solutions designed to meet compliance surveillance requirements.



# APN Partners are exploring new approaches to old challenges

Jordan & Jordan



Jordan & Jordan's **Execution Compliance and Surveillance** ("ECS") platform helps broker-dealers and buy-side firms comply with best execution and regulatory reporting requirements such as SEC rules 605 and 606, and also monitors for potential market manipulation activity such as spoofing and layering. ECS also offers integrated case management functionality with easy search and retrieval to assist with regulatory inquiries.



## Best execution

### International Broker-dealer

Uses the ECS platform running on AWS to monitor for compliance with best execution in U.S. and foreign securities

### Regional U.S. Broker-dealer

Uses the ECS platform running on AWS to monitor for compliance with trading regulations including best execution.



## Reporting

### International Broker-dealer

Uses ECS on AWS to generate and publish required SEC rules 605 and 606 reports.



## Market manipulation

### NYSE floor broker

Monitors market manipulations scenarios defined under SEC rule 15c3-5 with ECS running on AWS.

### Regional U.S. Broker-dealer

Monitors for market manipulation including customized reports for OTC securities and market-on-open matching.

# APN Partners are exploring new approaches to old challenges



Eventus Systems, Inc. is a software firm providing regulatory technology (regtech) solutions for capital markets organizations including Tier 1 banks, brokerages, FCMs, proprietary trading firms, exchanges, and buy-side firms. The firm's flagship Validus surveillance and risk software platform is available as a real-time or T+1 solution, and provides risk management and surveillance solutions.



## Top Crypto Exchange

This real-time implementation is utilizing a private AWS cloud solution to securely deliver full trade-lifecycle data to Validus via an Amazon Machine Image (AMI). Extremely fast and cost effective to deploy while maintaining a high level of security.



## Non-clearing FCM

Utilizing a hybrid cloud model where Validus is monitoring risk in real-time while managing all other market manipulation surveillance on a T+1 basis. This allows the customer to meet the real-time RTS-6 requirements while maintaining the benefits of a distributed AWS model.



## Tier 1 Bank

Global Tier 1 Bank that utilizes Validus for T+1 trade surveillance on futures and equities products. Validus consumes customer trade-lifecycle data via exchange drops and overlays market data sourced from a third party. This all takes place in a dedicated AWS instance that is easily accessible to the customer and extremely secure.

# APN Partners are exploring new approaches to old challenges

Partner: **INAWISDOM** 

**Inawisdom** works with organizations to determine their specific business opportunities and then tailors their Rapid Analytics and Machine-Learning Platform (RAMP) to meet the exact requirements of their customers' Machine Learning applications.

Customer:   
**IPSWICH**  
BUILDING SOCIETY

**Ipswich Building Society**, a Financial Services firm that provides mortgages and savings accounts in the U.K., uses Inawisdom's RAMP platform on AWS to comply with FCA anti-money laundering regulations.

# APN Partners also provide underlying support to surveillance solutions



**KSF Global Services'** surveillance platform helps broker dealers comply with USA FINRA, EU MIFID2, UK FCA and Swiss FIDLEG financial supervision regulations. The Arkivy Record Keeping Operations System archives and ties all financial data together providing a centralized repository that facilitates fast search and insightful financial surveillance across all content types.



**IT SOFTWARE S.p.A.** is a leading provider of financial software on the Italian market. IT Software's "EasyTrade" suite of products is an order execution and management system used by more than 30 financial firms globally. "Arkivy on AWS helped to archive EasyTrade orders along with reporting documentation required by MIFID2, while enabling the ability for order reconstruction and surveillance - thus solving the compliance challenges for IT SOFTWARE's financial industry customers worldwide." said Mirko Marcadella, Managing Director, IT Software S.p.A.

# Ready to start building?

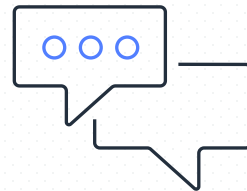
Efficiently transform your data infrastructure and enhance your compliance surveillance program with AWS.



Work with your AWS account team to gather requirements and develop a framework for building a customized data lake on AWS.



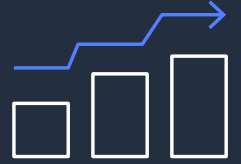
Contact the Amazon ML Solutions Lab and investigate opportunities to accelerate incorporation of machine learning into the surveillance reports.



Work with an APN Partner to implement surveillance solutions on AWS.



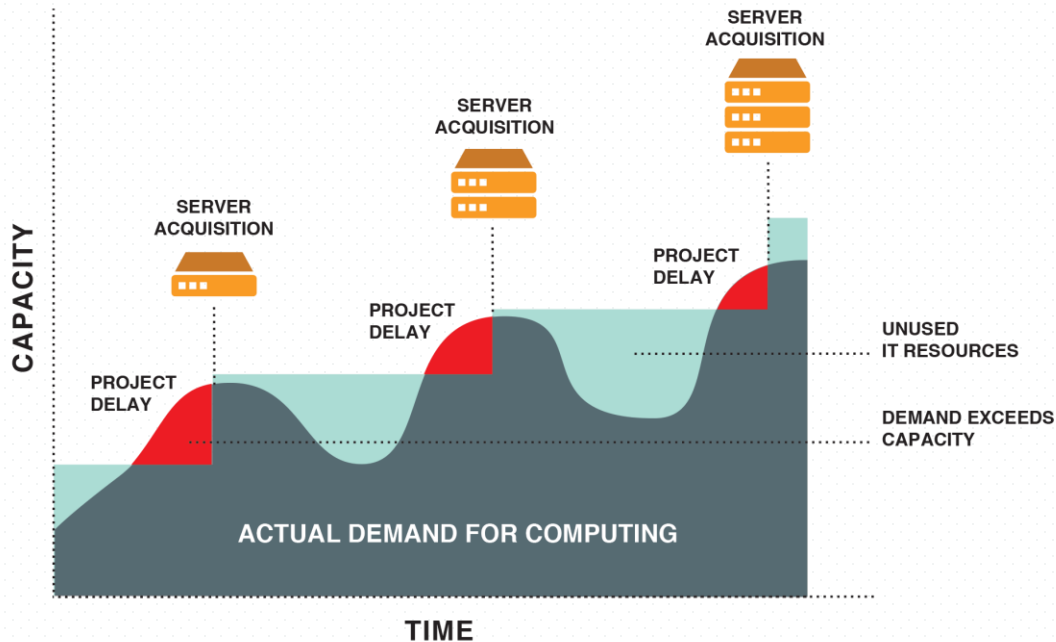
# Appendix 1: Addressing common surveillance challenges



# Capacity and scale

# Challenge: Ability to handle variable and increasing amounts of data

On-premise static capacity requires time and effort to provision the appropriate resources. However, unanticipated market events or business requirements can result in the actual demand exceeding the available capacity and adversely impact SLAs and project timelines.

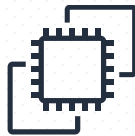


# Solution: Scalable, elastic compute and storage



## Amazon S3

With Amazon S3, you never need to worry about provisioning **storage resources**.



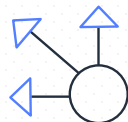
## Amazon EC2

With Amazon EC2, you receive secure, scalable **compute capacity** in the cloud.



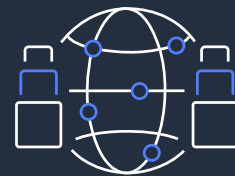
### Elasticity

Scale up or down quickly, as needed, provisioning capacity in minutes, making it easy to meet changing requirements while reducing costs.



### Flexibility

Establish flexible data and overall technology architectures that can meet evolving demands of business and external regulatory entities.



# Data infrastructure

# Challenge: Ability to build more sophisticated surveillance reports



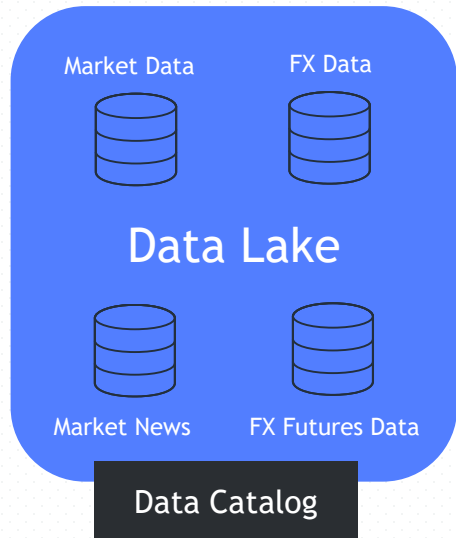
- Legacy infrastructure stores the data for different asset classes, products, and regions in multiple **disconnected data silos**
- Each data silo has **distributed ETL processes** at multiple levels and inconsistent data transformation between silos, obscuring lineage back to source data and resulting in downstream data quality issues
- **Duplication of effort** in data management (data lifecycle, retention, versioning, etc.)

# Solution: Data lakes provide the foundation for enhanced surveillance reports

1. Incorporate cross asset/  
cross product activity

2. Integrate market news  
with alert logic

3. Augment alerts with market  
/ third party data



ETL

Consistent ETL  
process minimizes  
data quality issues



Comprehensive coverage of firm  
and client activity

Generate more substantive alerts

Provide more insight and context  
around the activity for the analyst



# Data analytics



# Challenge: Large volumes of unproductive alerts

## Impact



**Increases costs:**  
Requires resources to review alerts



**Increases risk:**  
High number of false positive alerts may result in an actual issue being overlooked



**Perception risk:**  
Regulators question why the firm has a high number of market misconduct alerts

## Underlying issues



Without a well articulated approach, the parameters for the alert generation logic is often set conservatively to “cast a wide net”

The current rules based approach is rigid and does not easily adapt to changing market conditions.

## Challenges



Regulators require documentation and rationale on how the threshold values were determined

The threshold values in production need to be reproducible using the documented process

# Solution: Use machine learning to optimize surveillance logic

## Optimize alert output:



Calibrate thresholds in transaction surveillance alert logic



Refine keywords and phrases used for e-communications surveillance lexicons

## New scenario analysis:



Extract hidden relationships and insights in the data



Drive “what if” analysis on risks which are not covered by existing metrics/analysis



# Innovation

# Challenge: Building an integrated surveillance program

At many financial institutions, the ability to experiment and innovate is limited due to onerous data pre-processing efforts.

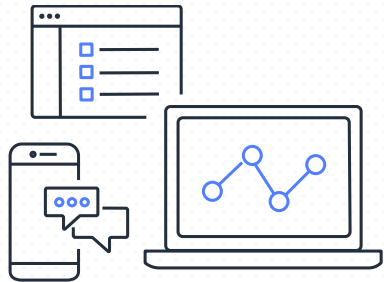


## Common challenges

- Siloed infrastructure without common data framework; inconsistent ETL processes
- Workflows with complex dependencies, operationally difficult to sustain
- Large amount of time and effort spent on preparing the data and sand box environment

# Solution: Easy access to resources that facilitate experimentation

The flexibility and agility of AWS allows the user to spin up sand box environments on demand and easily ingest the required data for experimentation and innovation.



Overlay market events with emails, chats, and phone conversations to provide a holistic view of the activity that includes the full context for the analyst.



Enhance surveillance alerts with network diagrams, using **Amazon Neptune**, to visually link a trader's activity with other employees or clients, providing insights on the extent of the market conduct issue.

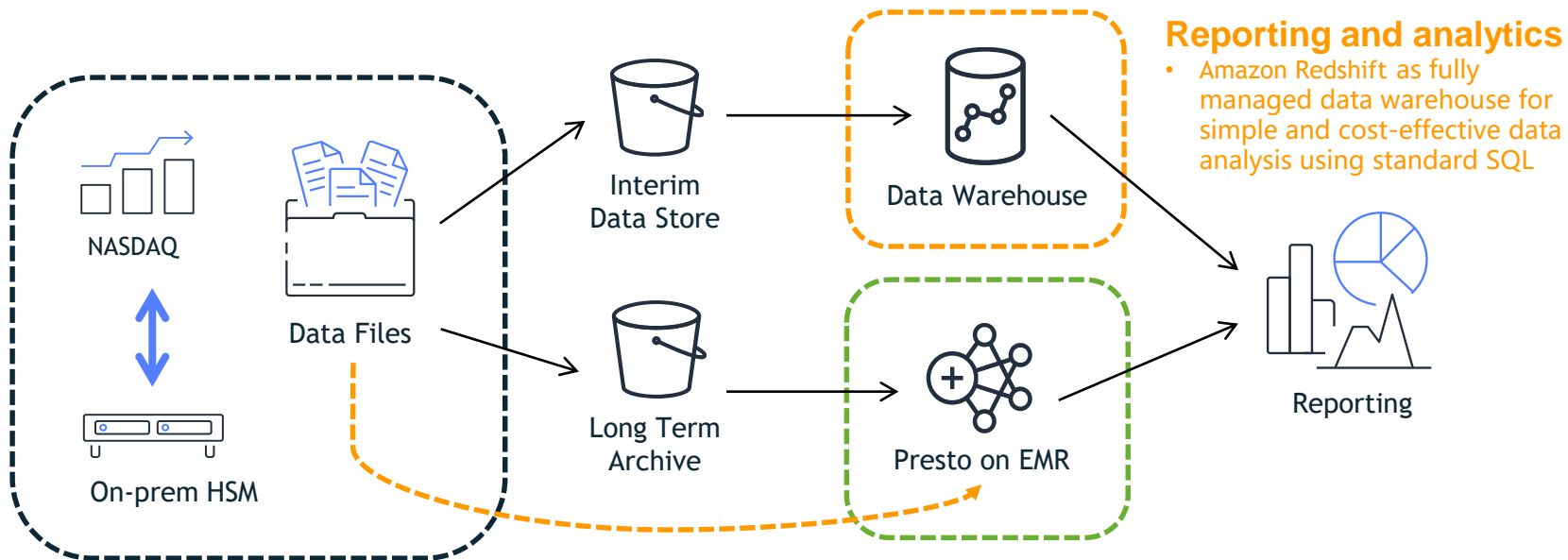


Amazon Neptune



# Appendix 2: Reference Architectures

# Customer architecture: Nasdaq data warehouse



## Reporting and analytics

- Amazon Redshift as fully managed data warehouse for simple and cost-effective data analysis using standard SQL

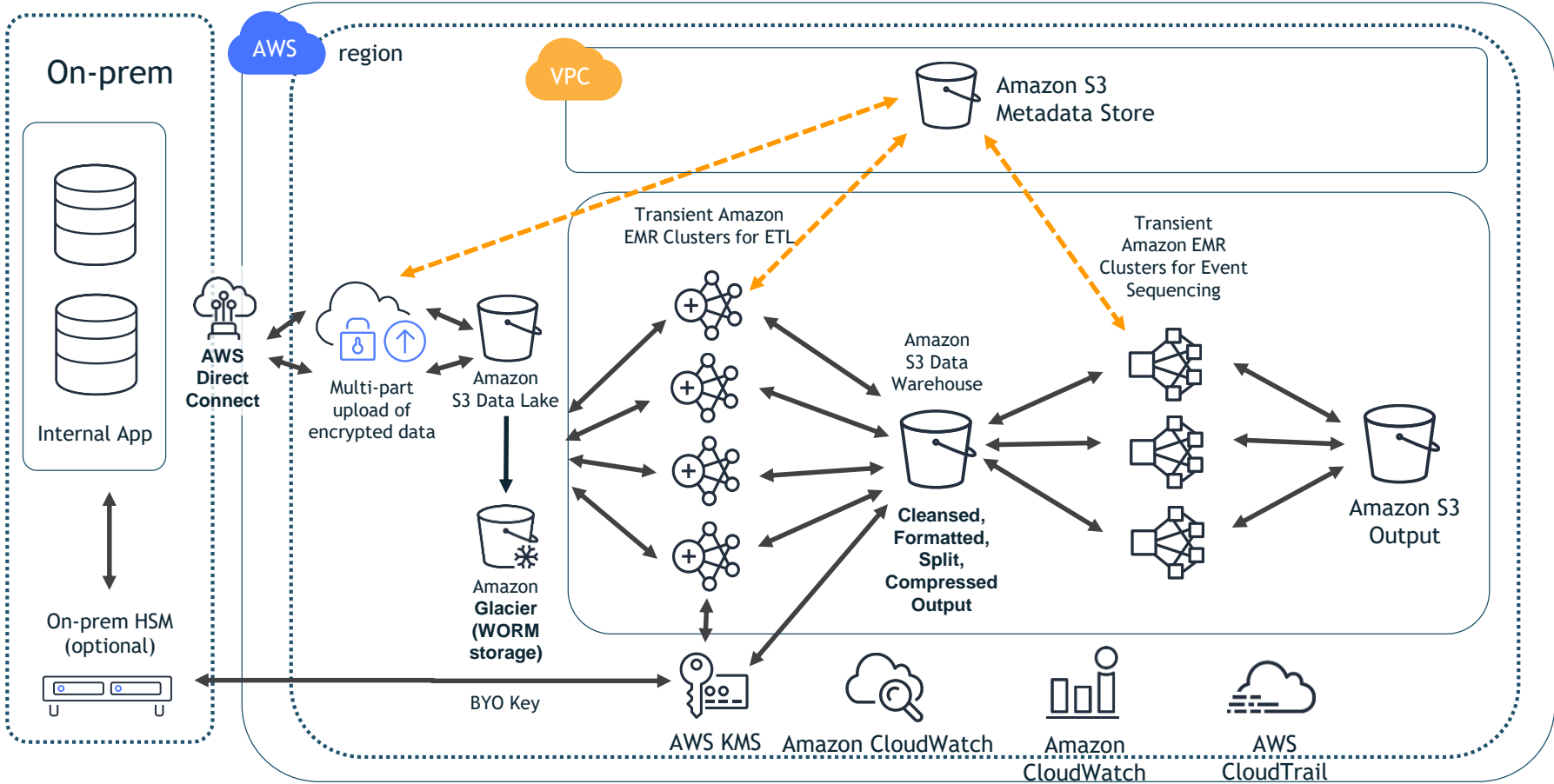
## Security

- Customer maintains control of data
- Designed to encrypt highly sensitive data efficiently

## Queryable archive

- Ability to respond to infrequent customer or regulatory requests for historic data using SQL tools at a fraction of the price of moving data into a database
- Natively available via Amazon Redshift Spectrum

# Representative CAT reporting architecture





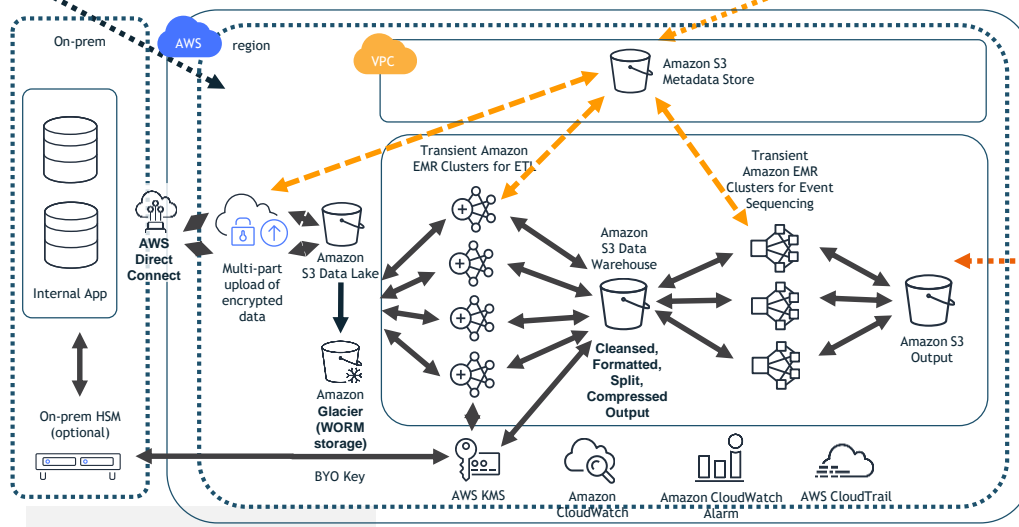
# CAT reporting architecture functional highlights

## Data transformation and optimization

- Highly redundant, scalable, available Amazon S3 data lake
- Multiple, scalable on demand Amazon EMR clusters can be transient or permanent
- Custom security configuration

## Security

- Customer maintains control of data
- Designed to encrypt highly sensitive data efficiently
- Continuous transparency and auditing of controls



## Data lineage

- Ability to track data provenance and end-to-end changes
- Ability to trace errors back to the root cause, or re-create and re-run processes

## Reporting and analytics

- Amazon EMR clusters enable conversion of data to specific reporting format
- Ability to perform cost-effective ad-hoc analytics on a per-query basis