Cloud Data Lake

Analytic Data Warehouse in the Cloud

John Hitchingham

FINRA Engineering
Managed Cloud Data Lake

A virtual database and processing platform – for all our data
Reporting
Data Science
Machine Learning
Data Management
Data Processing Pipeline

Improved

Cost Reduction
Security
Regulatory Compliance

Simplified

Achieved
Market Regulation – Analytics Pipeline

BDs  Exchanges  Reference Data Providers

- Trade execution records
- Market reference data

Validation

Prepare for Analytics (ETL)

Interactive Analytics

Run Automated Detection Models

Develop Models

Data Scientist

Explore Investigate

Regulatory Analyst

Regulatory Follow up

75B+ events a day  20+ PB of Data  Major Exchange Clients  3Yrs Prod on Cloud
Before Data Lake
Competing workloads on constrained infrastructure

- Daily Order Volume (Billions)
  - Fluctuating Data Volumes
  - Fixed Capacity Data Warehouse Appliances

- Unexpected Processing

- Validation
- ETL
- Interactive Analytics
- Data Science
- Automated Detection Models

- Capacity could not adjust to competing demands
- Adding capacity was time consuming
- Adding fixed capacity was expensive (size for peak)
Analytics and data science was slow

What data do we have?
What format is it in?
Where do I get it?

Here’s your data

Oops, I need more data … Repeat!
I need data in different format … Repeat!

Get this data for them…
Not on disk – pull from tape
Prepare & Format

Wait for tapes from offsite

Business Analysts
Data Scientists
Data Analysts
Data Engineers
Ops
Drivers to change

• Fast-growing data volumes YoY
• High cost of pre-building for peak
• Escalating costs of in-house technology infrastructure
• Long time-to-market for finding insights in data
• Appliance platforms were facing obsolescence and end-of-life as a result of new big data technologies

Keep spending more on legacy infrastructure or redirect dollars to core business of regulation?
The Cloud Data Lake
Manage data consistently

- Define, store and share our data as an enterprise asset
- All data should be enabled for analytics
- Protect data in a holistic manner (At rest and in transit)

Integrate our portfolio

- Shared solutions for common business processes
- All "business" data will be tracked centrally so that FINRA can manage the data lifecycle in a productive and cost effective manner
- All FINRA-developed applications will have service interfaces

Cloud architectural principles

Operational resiliency

- Automatic failover across data centers
- Auto-scaling and load balancing to achieve high availability
- No logon to servers or services for routine operations
- Applications should include automated operations for failure scenarios, recovery, data issues, and notifications

Embrace open source

- Take advantage of rapid innovation in open source tools space
- Balanced against opportunity cost of other alternatives
- Avoid lock in
Where?

On Prem or Cloud

then

or

Cloud A or Cloud B or Cloud C or Cloud D
From data puddles to the Data Lake

![Diagram showing transition from data centers to AWS]

- **Query/Compute**
  - EMR Spark
  - EMR Presto
  - EMR HBase
  - Lambda

- **Storage**
  - Amazon S3
  - Hive metastore

- **Catalog**
  - herd

**CONCLUSION**

FINRA in Data Center → FINRA in AWS

**Key Points**
- Scales
- AWS
- Herd
- Hive metastore
S3 object store simplifies management of storage layer.

**FINRA Main Prod Data Store**

- Incrementally grow data store by petabytes – no operations work

**S3 Price - GB/Mo (1st 50 TB)**

- Periodic price reductions – no operations work

- Multi data center(s) (AZ)
- Versioning
- Encryption
CREATE EXTERNAL TABLE new_orders_prc_bz (oats_roe_id bigint, order_rcvd_dt timestamp)

PARTITIONED BY ( 'oats_prcsg_dt' date)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\u0001'
STORED AS TEXTFILE
LOCATION 's3://finra/oats/oats/prc/bz/new-orders/schm-v3/data-v0';


ALTER TABLE `new_orders_prc_bz` ADD PARTITION (`oats_prcsg_dt`='2016-08-12') LOCATION 's3://finra/oats/oats/prc/bz/new-orders/schm-v3/data-v0/oats-prcsg-dt=2016-08-12'
Catalog for centralized data management

**Unified catalog**
- Schemas
- Versions
- Encryption type
- Storage policies

**Lineage and Usage**
- Track publishers and consumers
- Easily identify jobs and derived data sets

**Shared Metastore**
- Common definition of tables and partitions
- Use with Spark, Presto, Hive, etc.
- Faster instantiation of clusters

[http://finraos.github.io/herd](http://finraos.github.io/herd)
The Analytic Data Processing Pipeline on the Data Lake

- **Ingest**: Broker Dealers, Exchanges, 3rd Party Providers
- **Catalog & Storage**: Data Files
- **ETL**: Normalize, Enrich, Reformat
- **Validation**
- **Detection models (Patterns) Automated Surveillance**
- **Analytics**: Data Scientist, Regulatory User, Human Analytics

**Processing Pipeline**

**Analytics**
Flexible capacity in data processing pipeline

- Raw Data
  - Pre-analytics Processing Pipeline (ETL)
    - Analytics Ready Data
  - Parallel Clusters
    - Parallel workers in cluster

Scaling

- Unexpected processing used to take months – now days
Processing scales to meet demand

Daily Order Volume (Billions)

AWS EMR compute on EC2

Hour of Day

Compute Nodes
Analytics - petabytes of data available to query

Data Analyst

JDBC/ODBC Client

SQL

presto

Metastore

AuthZ

AuthN

Logical “Database”

Table 1

Table 2

Table N
Achieving Interactive Query Speed

<table>
<thead>
<tr>
<th>Query</th>
<th>Table size (rows)</th>
<th>Output size (rows)</th>
<th>ORC</th>
<th>TXT/BZ2</th>
</tr>
</thead>
<tbody>
<tr>
<td>select count(*) from TABLE_1 where trade_date = cast('2016-08-09' as date)</td>
<td>2469171608</td>
<td>1</td>
<td>4s</td>
<td>1m56s</td>
</tr>
<tr>
<td>select col1, count(*) from TABLE_1 where col2 = cast('2016-08-09' as date) group by col1 order by col1</td>
<td>2469171608</td>
<td>12</td>
<td>3s</td>
<td>1m51s</td>
</tr>
<tr>
<td>select col1, count(*) from TABLE_1 where col2 = cast('2016-08-09' as date) group by col1 order by col1</td>
<td>2469171608</td>
<td>8364</td>
<td>5s</td>
<td>2m5s</td>
</tr>
<tr>
<td>select * from TABLE_1 where col2 = cast('2016-08-10' as date) and col3='I' and col4='CR' and col5 between 100000.0 and 103000.0</td>
<td>2469171608</td>
<td>760</td>
<td>10s</td>
<td>2m3s</td>
</tr>
</tbody>
</table>

Key points:
Use ORC (Or Parquet) for performant query

Test Config:
Presto 0.167.0.6t (Teradata) On EMR
Data on S3 (external tables)
Cluster size: 60 worker node x r4.4xlarge
Scale out with multiple clusters…

User A
JDBC Client

User B
JDBC Client

JDBC App

AuthN

Metastore

still one copy of data!

Logical "Database"

Table 1

Table 2

Table N
Data Science on the Lake
Data Lake unifies access all of the data

Before Data Lake

Data Lake

Accelerate discovery through self-service
Data Science Ecosystem on Data Lake

- Data Scientist
- JDBC Client
- EMR Cluster
- Spark Cluster
- DS-in-a-box
- Catalog
- Source Data
- Personal Data Marts
- Logical 'Database'
- Notebook Interface
- Notebook or Shell
- AuthN
- Active Directory
Universal Data Science Platform (UDSP) – DS Tools in “A Box”

Data Lake
RDB
Git Repo
Lib Repo

Other?
Python
Jupyter
LDAP AuthN

/home/me#

Use
Data Scientist

Manage
EC2
CPU | Mem | GPU

Launch New Instance
Select New Instance Details

Launch Node

Other?
Some example benefits - Cloud Data Lake

**Analytics**
- Analysts can now interactively analyze 1000x more market events (billions vs million rows)
- Querying order route detail went from 10s of minutes to seconds
- Quicker turnaround to provide data for
- Machine Learning model development is easier

**Agility**
- Easily reprocess data ... used to take weeks to find capacity now can be done in day/days.
- Cloud makes it very easy to share (even large) data sets with 3\(^{rd}\) parties in Cloud
- Can perform model (pattern) reruns in days not weeks

**Resiliency**
- Market volume changes no longer disruptive events
- Improved system uptime vs in-house

At a run rate 30% less expensive than our previous solution
Want to know more?

Contact: John.Hitchingham@finra.org
FINRA Technology: http://technology.finra.org/
We’re Hiring!