

Real Time Big Data Analytic Platform with Oracle GoldenGate Big Data Adapters



Rajit Saha

Vengata(Venky) Guruswamy

Principal BigData Engineer

Principal Database Administrator

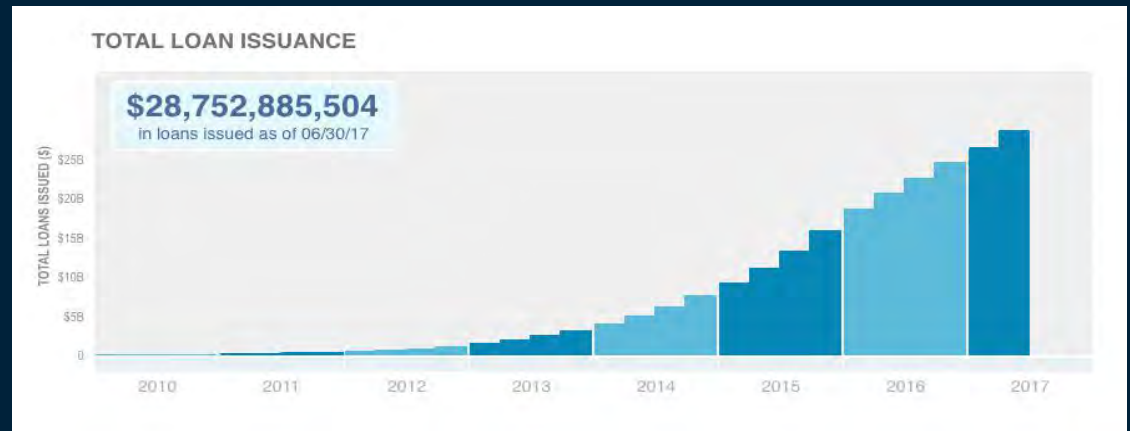
Agenda

- Introduction - LendingClub
- Real-time Big Data Platform Use cases
- Lambda Architecture – On-Premise / AWS
- Data Processing Flow/Algorithm
- GoldenGate Big Data Adapter Implementation Architecture, Configuration and Troubleshooting Scenarios

LendingClub

**LendingClub is
America's
largest online
marketplace
connecting
borrowers and
investors.**

- Headquartered in San Francisco, CA
- Office in Westborough, MA



Product Lines

Personal Loans

Loans up to \$40K

600+ FICO

36 and 60 mo. terms

Small Business

Business Loans up to
\$300K

At least \$75,000 in
annual sales

At least 2 years in
business

Patient Financing

Extended plans up to
\$50K; no-interest
plans up to \$32K

Auto Refinance

Must have an
outstanding balance
of \$5K-\$55K, initiated
in the last 3 months
with 24 months of
remaining payments

Use Cases of Oracle GoldenGate

@

LendingClub

- Maintain Reporting Databases
- Implement transaction history for key tables(Change Data capture)
- Enable real-time feed to Big data platform

Why Big Data Platform ?

Why Real-time Data Needed ?

- LendingClub's Big Data Platform is responsible for generating thousands of reports for the company – Daily, Monthly, Quarterly
 - Investor Reports, Financial Reports, Collection Reports, Risk Reports, Marketing Reports etc.
- Near Real-time Availability of OLTP Data in Hadoop Based OLAP warehouse
- Thousands of Tables and rapidly increasing data volumes towards petabytes
- Thousands of internal adhoc Users

NO Call list is a list of phone numbers we upload to dialer system in order to suppress Collections calls due to loan and/or customer status changes throughout the day.

- Typical changes are:
 - Positive Response from customers
 - Indicated DON'T CALL on a particular phone number from previous contacts.
 - Under certain circumstance, it can also be limited on how many times we can contact borrower by legal or regulatory reason.
 - “NO CALL” list needs be uploaded and fed into dialer in a very timely fashion, and ideally it should be **near real-time at the event/change occurred.**

What is it?

Track the progress of all the marketing channels with respect to quarterly forecast.

Why are we using Real-time?

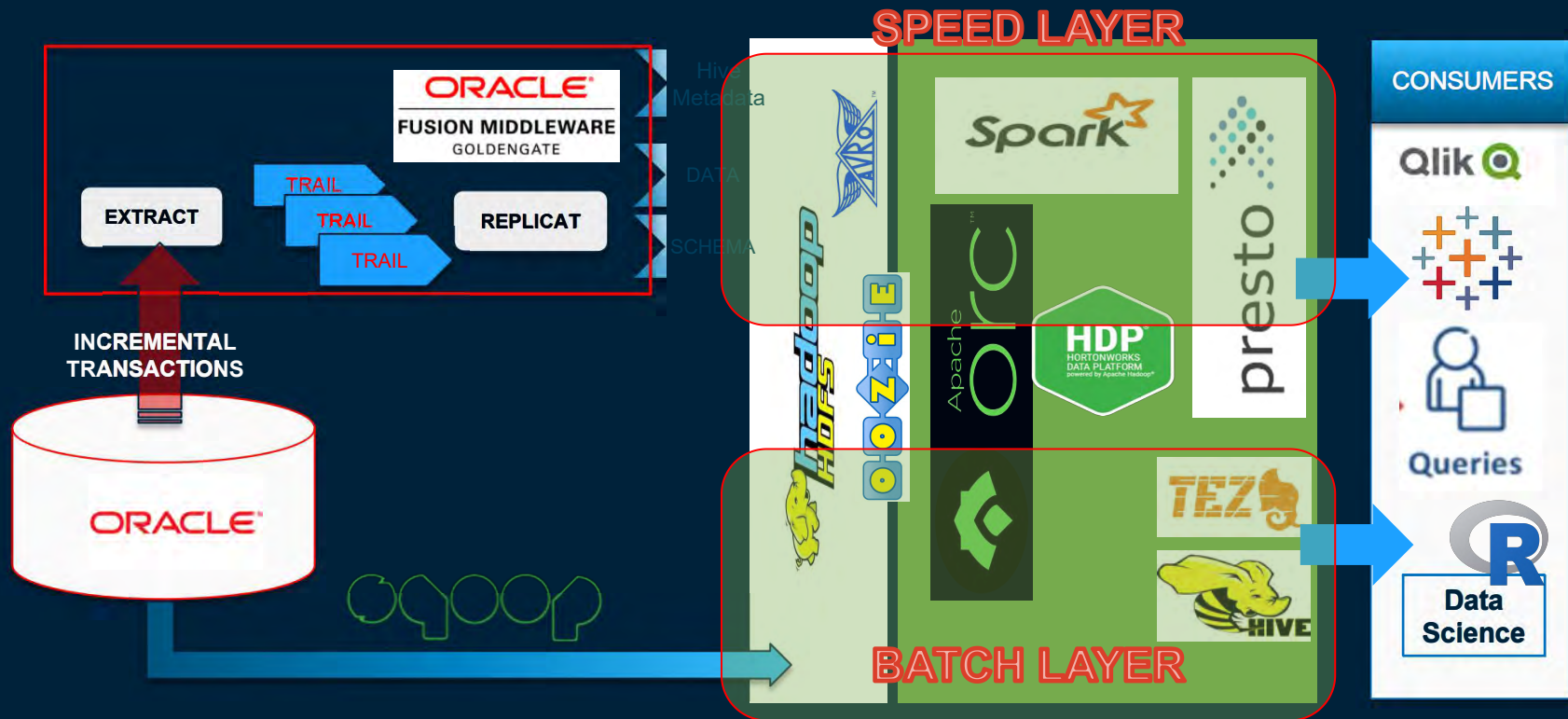
Business needs numbers every hour to change the strategy in case any channel performance compared to target.

Who is it for?

Marketing Team

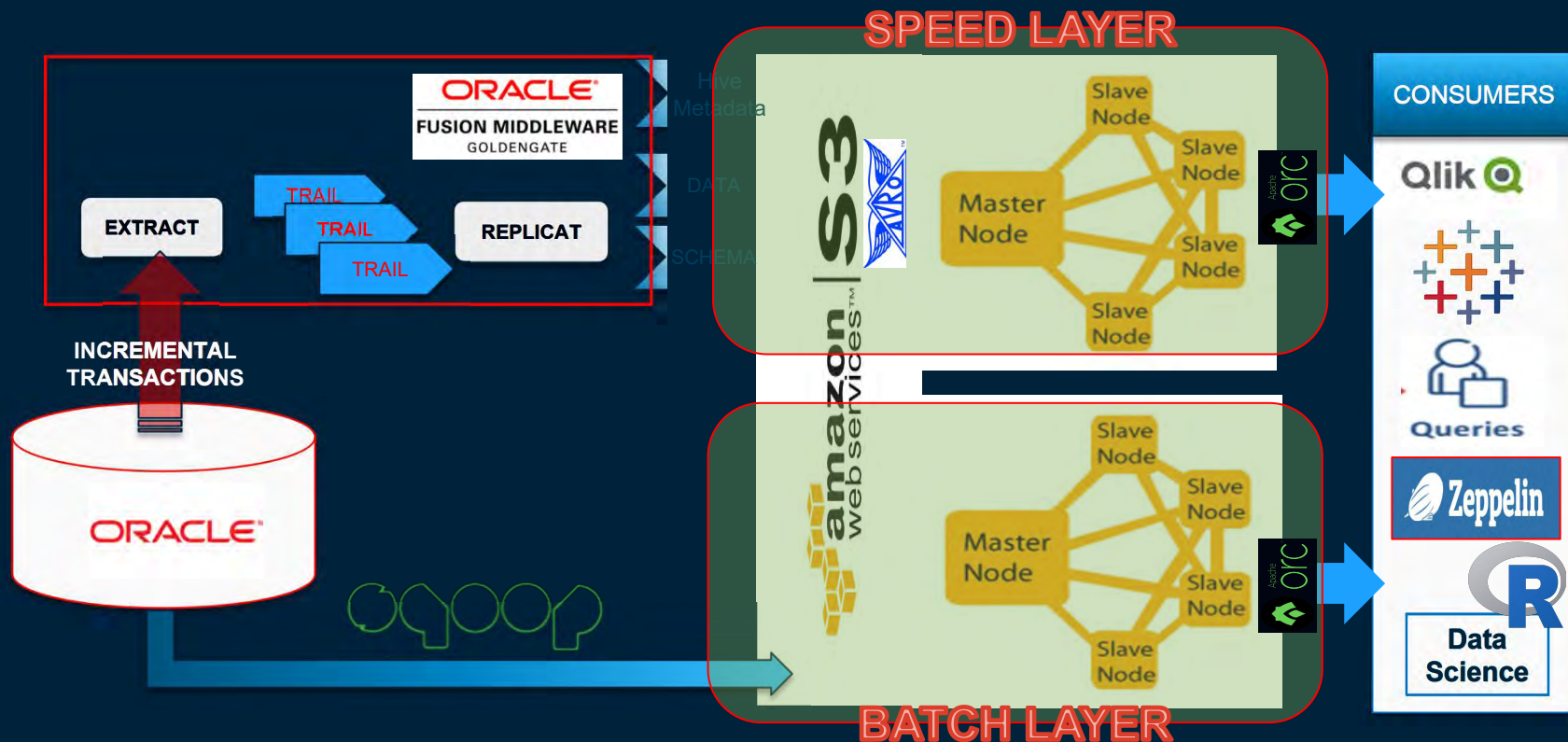
Lambda Architecture

On-Premise HortonWorks Cluster

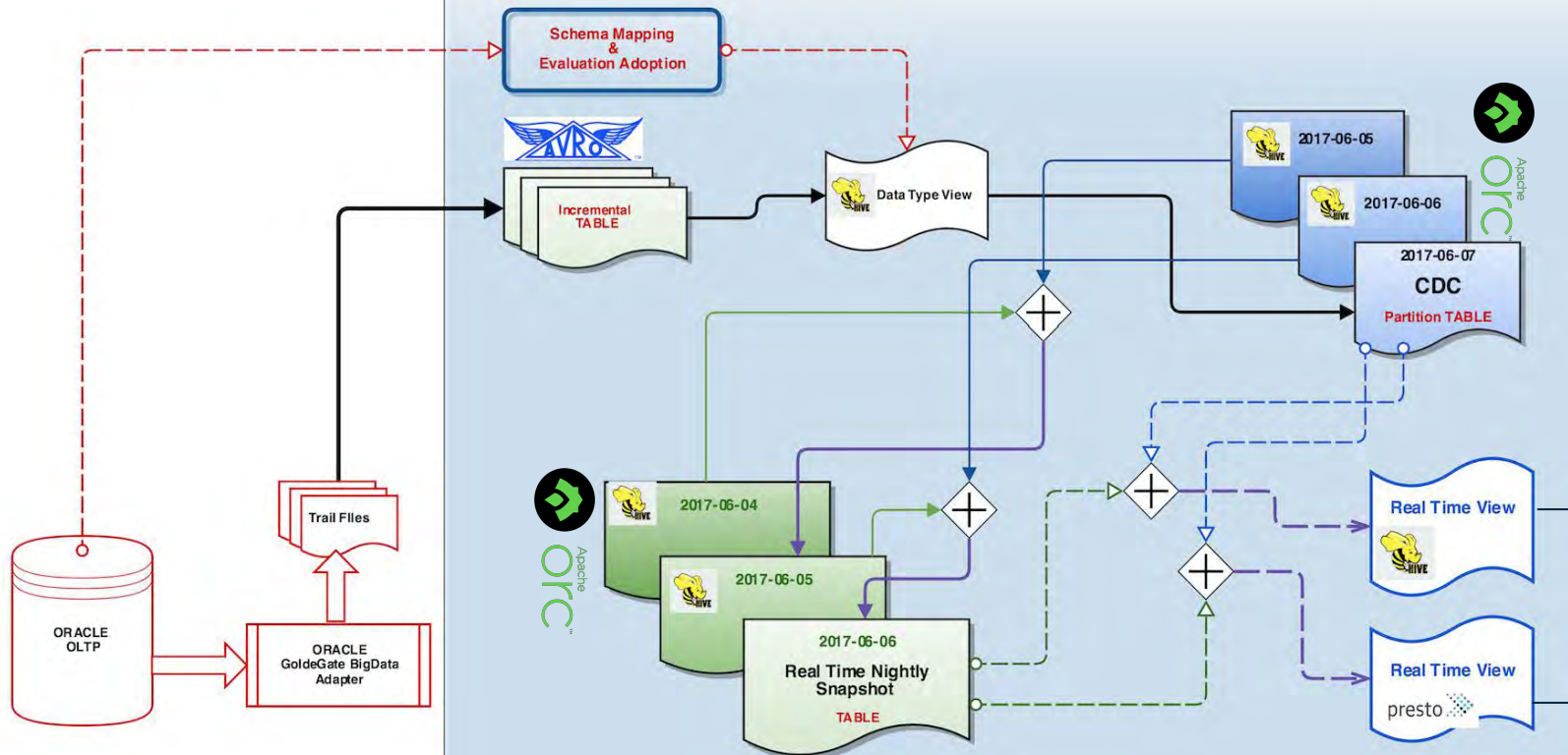


Lambda Architecture

AWS S3/EMR Clusters



SPEED LAYER - DATA FLOW DIAGRAM



Process Flow ..

Few Important Steps in Speed Layer Processing

- DDL Verification from Oracle Source to GG Table before processing starts
- Data Inserted from Avro GG Table to CDC partition table
 - Hive Dynamic Partition on **OP_TS**
 - **ORC** (Optimized Row Columnar) **Format**
- **Nightly Snapshot Creation –**
 - **Latest Updates from CDC tables and rest from previous day Snapshot**
- **Create Hive and Presto Real Time View with CDC and Nightly Snapshot**
- **Data Validation : Full table Checksum with Sqoop Snapshot**

How are we getting latest update

– For a Table , within a transaction for a row

Input Table From GG Adaptor

```
CREATE EXTERNAL TABLE
SCHEMA_GG.FOO (
TABLE STRING,
OP_TYPE STRING,
OP_TS STRING,
CURRENT_TS STRING ,
POS STRING ,
PRIMARY_KEYS ARRAY<STRING> ,
TOKENS MAP<STRING,STRING> ,
PKEY STRING ,
..
..
..
```

```
SELECT TOKENS FROM SCHEMA_GG.FOO LIMIT 1;
```

```
{"TKN-OPTYPE":"INSERT","TKN-RECORDLENGTH":"1400","TKN-LOGRBA":"1774157","TKN-
OBJECTNAME":"SCHEMA.FOO","TKN-FILERBA":"","TKN-LAGMSEC":"54420","TKN-
USERNAME":"XXX","TKN-SCN":"7633890649912","TKN-RSN":"7633890649838","TKN-
TRANSACTIONINDICATOR":"BEGIN","TKN-RECORDTIMESTAMP":"2017-07-30 23:49:14","TKN-
LOGPOSITION":"19784208","TKN-FILESEQNO":"","TKN-ROWID":"AAA46FACCAAD4CtAAS","TKN-
XID":"184.13.1380795","TKN-COMMITTIMESTAMP":"2017-07-30 23:49:15.000000","TKN-REDOTHREAD":"1"}
```

All the transaction records goes to Hive Partitioned ORC
CDC table- **SCHEMA_RT_CDC.FOO**

```
SELECT DISTINCT VW.PKEY,.. FROM (SELECT FOO.ID, .. RANK() OVER (PARTITION BY
FOO.PKEY ORDER BY FOO.TKN-SCN DESC, FOO.TKN-RSN DESC, FOO.POS DESC,
FOO.CURRENT_TS DESC) AS DEDUP_RANK FROM SCHEMA_RT_CDC.FOO WHERE
FOO.OP_TS_DATE >='2017-07-30') VW WHERE VW.DEDUP_RANK = 1
```

Verification of Real Time View definition

Records from CDC

```
SELECT PKEY, OP_TS, TKN_SCN, TKN_RSN, POS, CURRENT_TS, MODIFIED_D,
OP_TYPE FROM SCHEMA_RT_CDC.FOO WHERE PKEY = 2181322736 ORDER BY
TKN_SCN DESC,TKN_RSN DESC,POS DESC,CURRENT_TS DESC
```

PKEY	OP_TS	TKN_SCN	TKN_RSN	POS	CURRENT_TS	MODIFIED_D	OP_TYPE
2181322736	2017-06-29 02:32:33.0	7630245451950	7630245451909	00004193750007692575	2017-06-28 19:33:53.778	2017-06-28 19:32:33.0	U
2181322736	2017-06-29 02:32:33.0	7630245451950	7630245451909	00004193750007692575	2017-06-28 19:33:53.778	2017-06-28 19:32:33.0	U
2181322736	2017-06-29 00:39:16.001	7630236411132	7630236411130	00004191270019251264	2017-06-28 17:41:09.519	2017-06-28 17:39:16.0	U
2181322736	2017-06-29 00:39:16.001	7630236411132	7630236411130	00004191270019251264	2017-06-28 17:41:09.519	2017-06-28 17:39:16.0	U
2181322736	2017-06-28 17:02:59.002	7630205764444	7630205764072	00004187440038392981	2017-06-28 10:04:13.76	2017-06-28 10:02:58.0	I
2181322736	2017-06-28 17:02:59.002	7630205764444	7630205764072	00004187440038392981	2017-06-28 10:04:13.76	2017-06-28 10:02:58.0	I
2181322736	2017-06-28 17:02:59.002	7630205764444	7630205764072	00004187440038392981	2017-06-28 10:04:13.76	2017-06-28 10:02:58.0	I
2181322736	2017-06-28 17:02:59.002	7630205764444	7630205764072	00004187440038392981	2017-06-28 10:04:13.76	2017-06-28 10:02:58.0	I

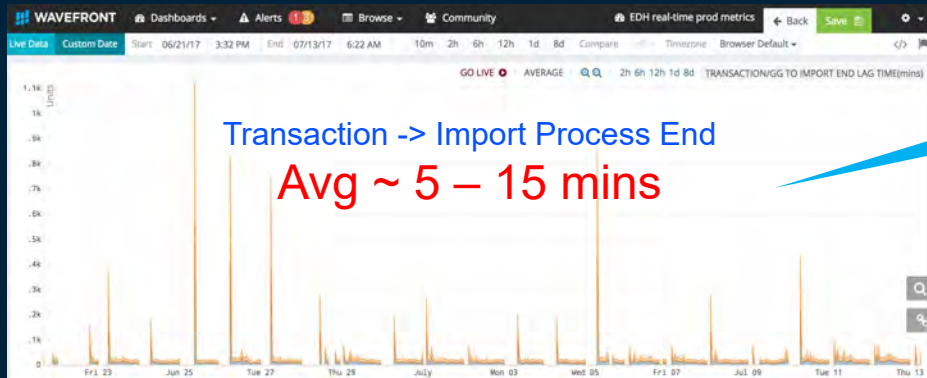
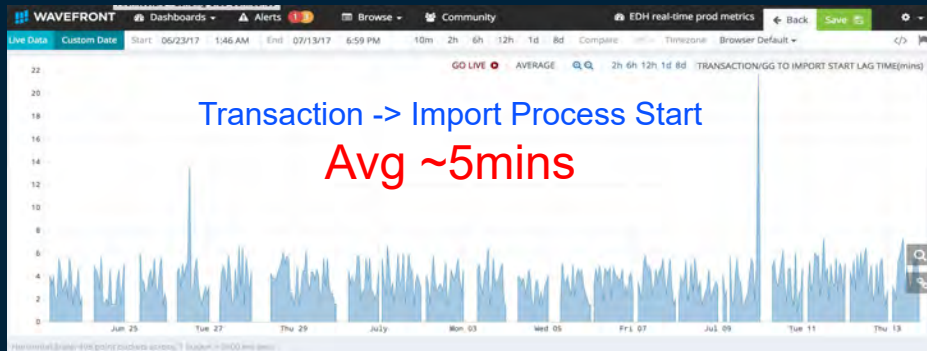
Records from Real Time View

```
SELECT PKEY, MODIFIED_D FROM SCHEMA_RT.FOO WHERE PKEY = 2181322736
```

2181322736

2017-06-28 19:32:33.0

SLAs

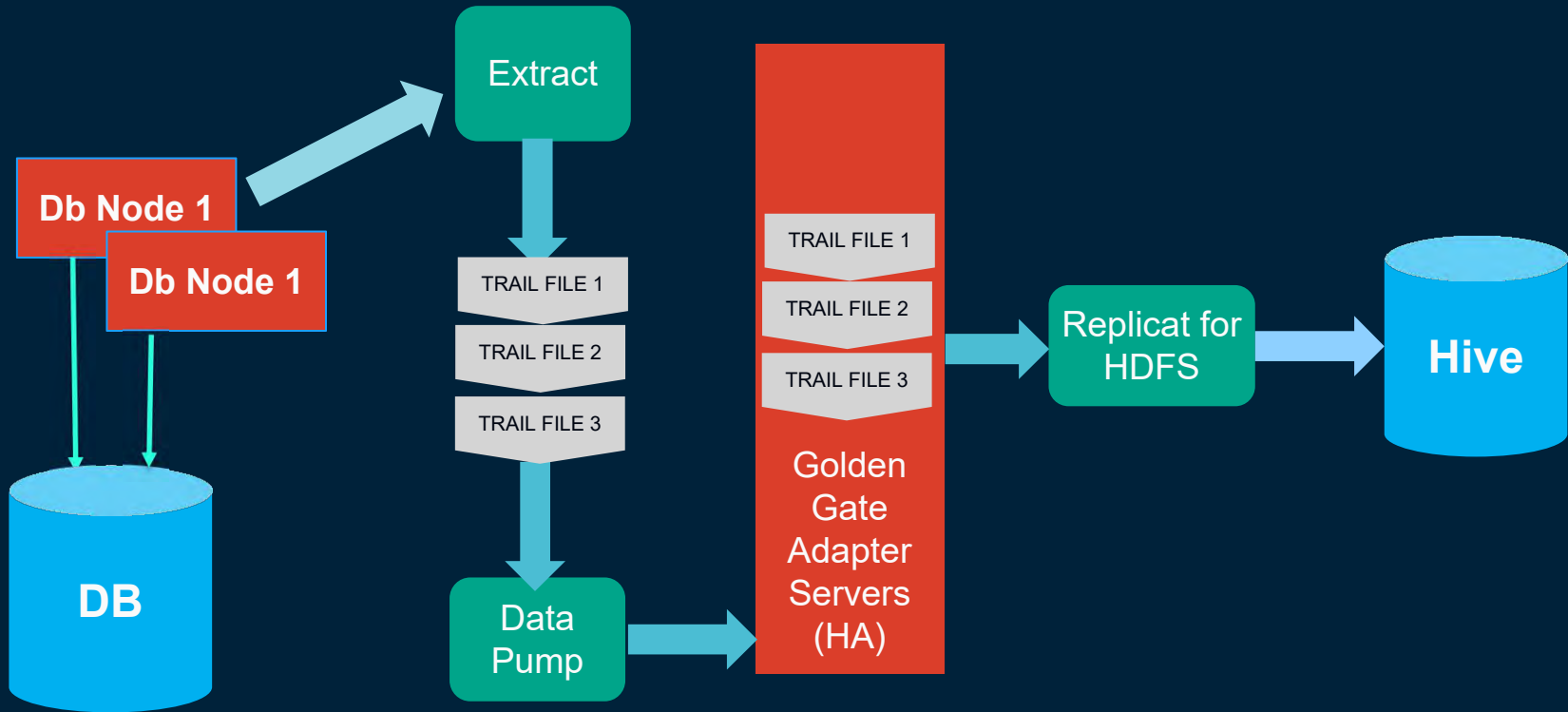


Spikes are for NIGHTLY SNAPSHOT Creation

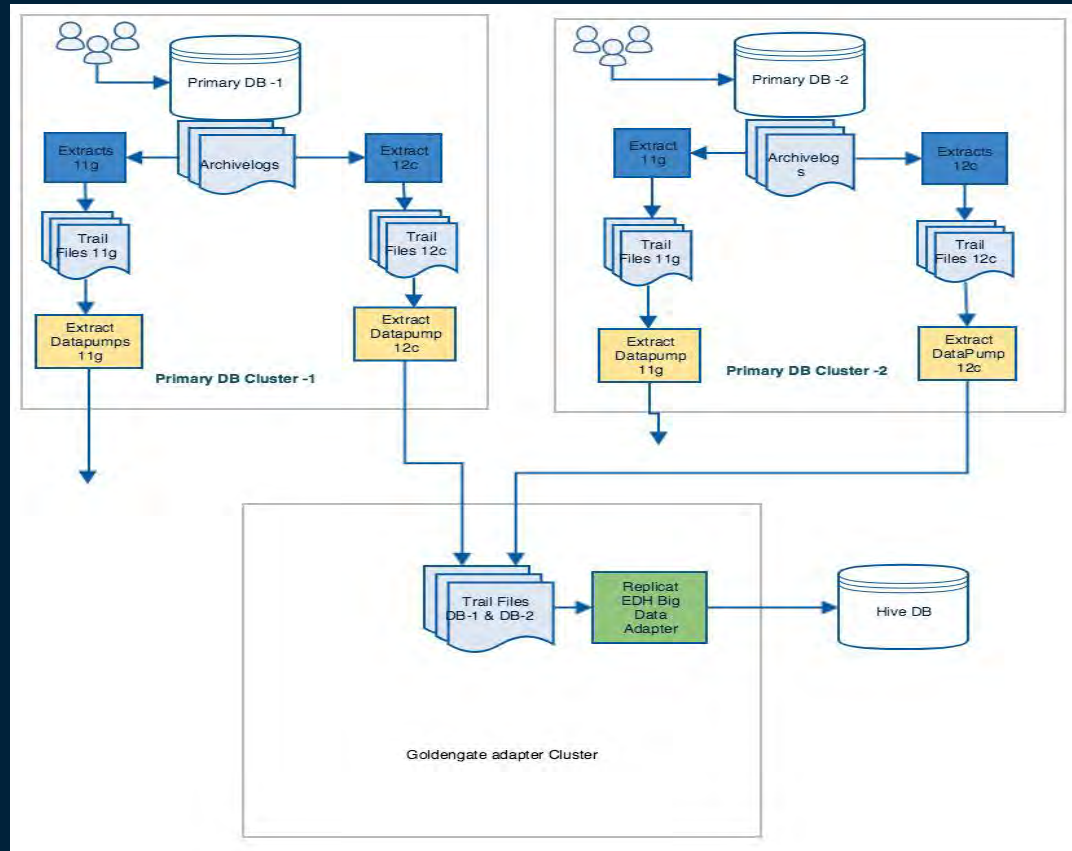
1. Full table creation
2. TEZ is not good for Skew JOIN
3. Cluster Busy

GoldenGate Big Data Adapter Implementation Architecture

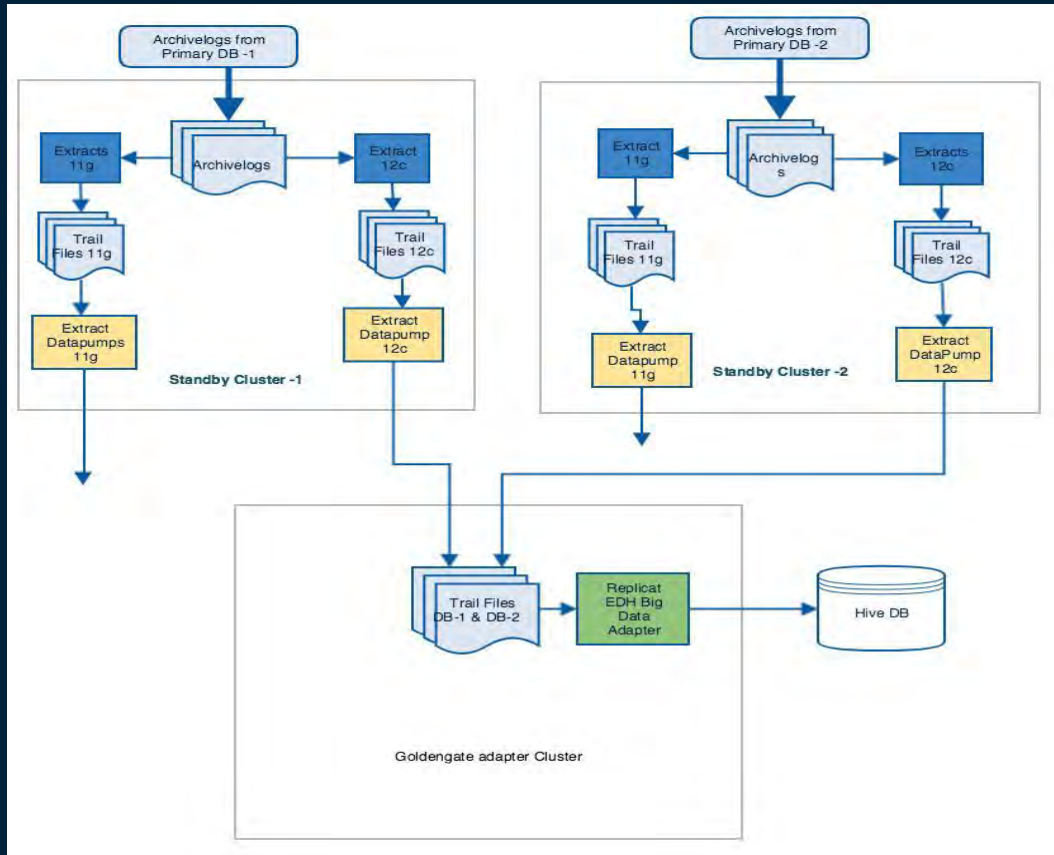
GoldenGate BigData Adapter Pipeline



GoldenGate BigData Infrastructure Architecture



GoldenGate BigData Infrastructure Architecture – Site Failover



GoldenGate Big Data Adapter Configuration

Source Extract Parameter - eprod.prm

```
NOCOMPRESSUPDATES
GETUPDATEBEFORES
TABLE <schema name>.*, tokens (
TKN-COMMITTIMESTAMP = @GETENV('GGHEADER', 'COMMITTIMESTAMP'),
TKN-FILESEQNO = @GETENV('RECORD', 'FILESEQNO'),
TKN-FILERBA = @GETENV('RECORD', 'FILERBA'),
TKN-LAGMSEC = @GETENV('LAG', 'MSEC'),
TKN-LOGPOSITION = @GETENV('GGHEADER', 'LOGPOSITION'),
TKN-OBJECTNAME = @GETENV('GGHEADER', 'OBJECTNAME'),
TKN-OPTYPE = @GETENV('GGHEADER', 'OPTYPE'),
TKN-RECORDLENGTH = @GETENV('GGHEADER', 'RECORDLENGTH'),
TKN-RECORDTIMESTAMP = @GETENV('RECORD', 'TIMESTAMP'),
TKN-ROWID = @GETENV('RECORD', 'ROWID'),
TKN-RSN = @GETENV('RECORD', 'RSN'),
TKN-SCN = @GETENV('TRANSACTION', 'CSN'),
TKN-TRANSACTIONINDICATOR = @GETENV('GGHEADER', 'TRANSACTIONINDICATOR'),
TKN-USERNAME = @GETENV('TRANSACTION', 'USERNAME'),
TKN-XID = @GETENV('TRANSACTION', 'XID')
)
```

Source Data Pump Extract Parameter - pprod.prm

EXTRACT PPROD

PASSTHRU

RMTHOST <Goldengate adapter server>, MGRPORT <port no>

-- Trail file written by Pump to the remote host

RMTRAIL /filesystem/pumptrail_ALL/pt

-- Pass data through without mapping, filtering, conversion:

PASSTHRU

-- Specify tables to be captured:

TABLE *.*;

Target – Replicat Parameter - rhdfs.prm

REPLICAT RHDFS

TARGETDB LIBFILE libggjava.so SET property=dirprm/hdfs.props

REPORTCOUNT EVERY 1 MINUTES, RATE

GROUPTRANSOPS 10000

-Schemas

MAP SCHEMA1.*, TARGET SCHEMA1_GG.*;

MAP SCHEMA2.*, TARGET SCHEMA2_GG.*;

-Heartbeat

MAP DBASHEMA.HEARTBEAT, TARGET DBASHEMA_GG.HEARTBEAT;

Target – Replicat Parameters - hdfs.props

```
gg.handler.hdfs.type=hdfs
```

```
##performance
```

```
gg.handler.hdfs.maxFileSize=256m
```

```
gg.handler.hdfs.fileRollInterval=5m
```

```
gg.handler.hdfs.inactivityRollInterval=5m
```

```
#config
```

```
gg.handler.hdfs.fileSuffix=.avro
```

```
gg.handler.hdfs.partitionByTable=true
```

```
gg.handler.hdfs.rollOnMetadataChange=true
```

```
gg.handler.hdfs.format=avro_row_ocf
```

```
## hive jdbc config
```

```
gg.handler.hdfs.schemaFilePath=/ogg/schema
```

```
gg.handler.hdfs.rootFilePath=/ogg/data
```

```
gg.handler.hdfs.hiveJdbcUrl=jdbc:hive2://@Hive_Server_Name:Hive_Port
```


Oracle GoldenGate to AWS S3 Connector

s3replicat.props

```
gg.handlerlist=hdfs  
gg.handler.hdfs.type=hdfs  
gg.handler.hdfs.rootFilePath=S3://LCBUCKET/OGG
```

Hadoop Client Talks to s3 via Hadoop-AWS package **via s3a protocol**

EMR only understand s3 , so s3a to s3 mapping is done by

Custom Built Hadoop Client from Trunk **hadoop-3.0.0-alpha3-SNAPSHOT**

```
<property>  
  <name>fs.s3a.access.key</name>  
  <value><< AWS KEY >></value>  
</property>  
<property>  
  <name>fs.s3a.secret.key</name>  
  <value><< AWS SECRET KEY >></value>  
</property>  
<property>  
  <name>fs.s3a.proxy.host</name>  
  <value> << PROXY SERVER >> </value>  
</property>  
<property>  
  <name>fs.s3a.proxy.port</name>  
  <value><< PROXY PORT >></value>  
</property>  
<property>  
  <name>fs.s3a.connection.ssl.enabled</name>  
  <value>true</value>
```

core-site.xml

```
</property>  
<property>  
  <name>fs.s3.impl</name>  
  <value>org.apache.hadoop.fs.s3a.S3AFileSystem</value>  
</property>
```

```
<property>  
  <name>fs.s3a.server-side-encryption-algorithm</name>  
  <value>SSE-KMS</value>  
</property>  
<property>  
  <name>fs.s3a.server-side-encryption-key</name>  
  <value>arn:aws:kms:<< KEY >></value>  
</property>
```

Security - Authentication

Kerberos params

gg.handler.name.authType=kerberos

gg.handler.name.kerberosKeytabFile=/etc/security/keytabs/lcapp.headless.keytab

gg.handler.name.kerberosPrincipal=osuser@PROD-DOMAIN.COM

Security - Username/password Encryption

```
gg.handler.hdfs.hiveJdbcUserName=unencrypted_username  
gg.handler.hdfs.hiveJdbcPassword=unencrypted_password
```

```
ORACLEWALLETUSERNAME ggalias ggadapters  
ORACLEWALLETPASSWORD ggalias ggadapters
```

```
gg.handler.hdfs.hiveJdbcUsername=ORACLEWALLETUSERNAME[ggalias ggadapters]  
gg.handler.hdfs.hiveJdbcPassword=ORACLEWALLETPASSWORD[ggalias ggadapters]
```

Security-Credential Store

Password encryption steps:

1. Add credential store:

```
GGSCI > ADD CREDENTIALSTORE  
Credential store created in ./dircrd/.
```

2. Add the credential for these users.

```
ALTER CREDENTIALSTORE ADD USER <Hive_username>, password  
<Hive_password> alias ggadalias domain ggadapters
```

How HDFS Handler handles the Data Flow ?

- The output format chosen is Avro row ocf format for three reasons
 - Integration with Hive
 - Seamless Schema Evolution
 - More compact than Avro op ocf
- There are two files produced namely Avro and Avsc files.

Configure High Availability for GoldenGate Resource

As ROOT user.

```
agctl add goldengate gg1 --gg_home /bdggvol/ggaws --instance_type dual
--nodes node1,node2 -network 1 --ip 99.99.99.99 --user appuser --group oinstall
--filesystems ora.registry.acfs --environment_vars
"LD_LIBRARY_PATH=/opt/java/default/jre/lib/amd64/server:/lib,
PATH=/opt/java/default/bin:/usr/lib64/qt-3.3/bin:/usr/local/bin,
JAVA_HOME=/opt/java/default"
```

```
$GRID_HOME/bin/crsctl setperm resource xag.gg1-vip.vip -u user:oracle:rwx
```

```
$GRID_HOME/bin/crsctl setperm resource xag.gg1-vip.vip -u group:oinstall:rwx
```

```
$GRID_HOME/bin/crsctl setperm resource xag.gg1.goldengate -u group:oinstall:rwx
```

```
$ GRID_HOME/bin/crsctl setperm resource ora.net1.network -u user:appuser:rwx
```

Schema Regex

Avro Doesn't like (\$) symbol ! – No problem :

```
gg.schemareplaceregex=[$:]
```

```
gg.schemareplacestring=_
```

Goldengate Big Data Adapter – Things To Remember

- Upgrade from 12.2 to 12.3 has changed the timestamp format from [YYYY-MM-DD:HH24:MI:SS.FFF] to [YYYY-MM-DD HH24:MI:SS.FFF]
- To avoid the dreaded ERROR OGG-15050 Error loading Java VM runtime library: (2 No such file or directory) ensure to set the following parameter JAVA_HOME, LD_LIBRARY_PATH and PATH properly. Remember manager passes the ENV variables.
- It's a multi-threaded process and ensure to allocate sufficient Unix resources like NPROC ..
- Please monitor the heart beat table from Hive.
- Design the application to be Idempotent.
- Run the HDFS handler from dedicated hadoop client node.

Troubleshooting Scenarios

Rewind – Functionality

```
2> alter hdfstest extseqno 10 extrba 0
```

```
2017-08-02 17:36:52 INFO    OGG-06594 Replicat HDFSTEST has been altered through GGSCI. Even  
the start up position might be updated, duplicate suppression remains active in next startup.  
To override duplicate suppression, start HDFSTEST with NOFILTERDUPTRANSACTIONS option.
```

```
REPLICAT altered.
```

```
3> start hdfstest NOFILTERDUPTRANSACTIONS
```

```
Sending START request to MANAGER ...  
REPLICAT HDFSTEST starting
```

Before Patch - 17995064 ☹️

```
strace -c -p 6170;date
Mon Aug 1 14:34:49 PDT 2016
Process 6170 attached
^CProcess 6170 detached
```

% time	seconds	usecs/call	calls	errors	syscall
74.87	0.062990	143	442		fsync
10.56	0.008886	4	2288		write
6.00	0.005052	1	3690		poll
3.69	0.003108	31	100	1	futex
3.20	0.002693	1	2592		sendto
0.98	0.000825	0	3134		read
0.52	0.000434	0	3686		recvfrom
0.07	0.000062	0	445		lseek
0.07	0.000056	1	90	6	stat
0.03	0.000027	1	28		select
0.00	0.000000	0	6		open
0.00	0.000000	0	7		close
0.00	0.000000	0	6		fstat
0.00	0.000000	0	1		rt_sigreturn
0.00	0.000000	0	1		dup2
0.00	0.000000	0	1		getsockopt
0.00	0.000000	0	1		restart_syscall
100.00	0.084133		16518	7	total

After Patch - 17995064 😊

```
date;strace -c -p 30907
```

```
Process 30907 attached
```

```
^CProcess 30907 detached
```

% time	seconds	usecs/call	calls	errors	syscall
28.76	0.029337	1	24245		poll
25.63	0.026140	9	2987		fsync
20.59	0.021005	1	15101		write
12.74	0.012998	116	112		futex
5.97	0.006091	0	26808		read
3.06	0.003125	0	17360		sendto
1.93	0.001965	0	24222		recvfrom
0.98	0.001000	42	24		open
0.29	0.000298	0	2987		lseek
0.04	0.000043	2	24		close
0.00	0.000000	0	78		stat
0.00	0.000000	0	24		fstat
0.00	0.000000	0	26		select
0.00	0.000000	0	1		restart_syscall
100.00	0.102002		113999		total

```
Mon Aug 1 14:35:20 PDT 2016
```

Lobs-Clobs-Blobs

- Data team needed the CLOBS even though the CLOB is not updated.
- GETUPDATEBEFORES doesn't work for CLOB
- Lobs :Always Include LOBS In Trail File (Doc ID 1639717.1)
- Include the extract parameters NOCOMPRESSUPDATES and FETCHCOLS option in the TABLE parameter, for example:

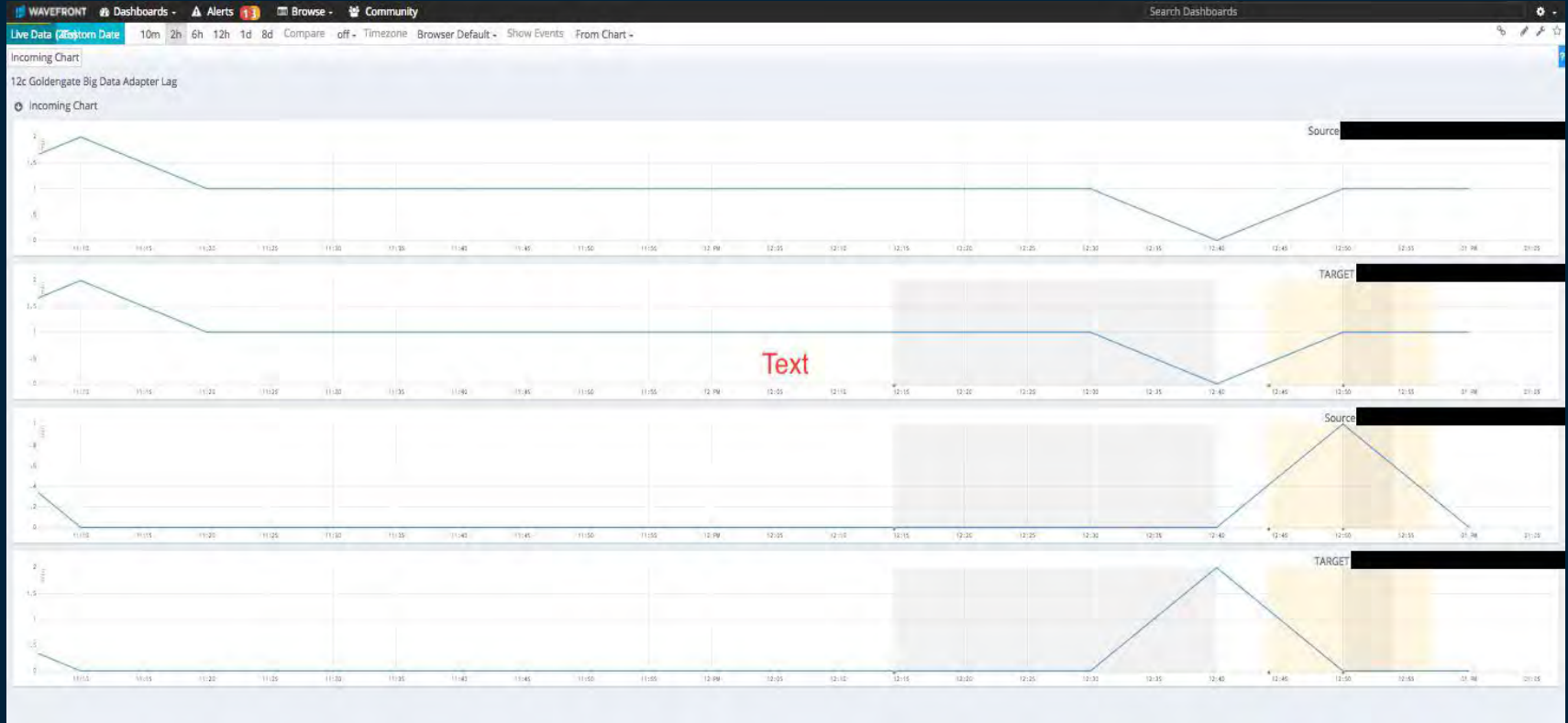
NOCOMPRESSUPDATES

TABLE schema.mytable, fetchcols (mylobs1, mylob2);

Ability to handle hadoop failures

- **SR 3-15688859001 : Data loss after a glitch between goldengate adapter and namenode**
 - Please test the rewind functionality and ability to handle duplicate data
 - Please upgrade to latest release - OGGBD 12.3.1.1.0 which contains the fix for the abend issue.
 - Monitor heart beat table in Hive.
 - Monitor the ERROR messages from logfiles - RHDFS_info_log4j.log [Splunk]
 - java.io.EOFException: Premature EOF: no length prefix available

GoldenGate Big Data Adapter - Monitoring [Wavefront]



Thank You



Q & A