Production Readiness Testing Using Apache Spark

Jag Jayaprakash
Ganesh Tiwari

Background: Salesforce App Cloud

FORCE

Model-driven development platform

LIGHTNING

Visual Development Platform

HEROKU

Polyglot platform for elastic scale

THUNDER

Stream & event based primitives

APP EXCHANGE

Enterprise App Marketplace

Background: Apex

STRONGLY TYPED

Direct references to schema objects

OBJECT ORIENTED

Visual Development Platform

EASY TO TEST

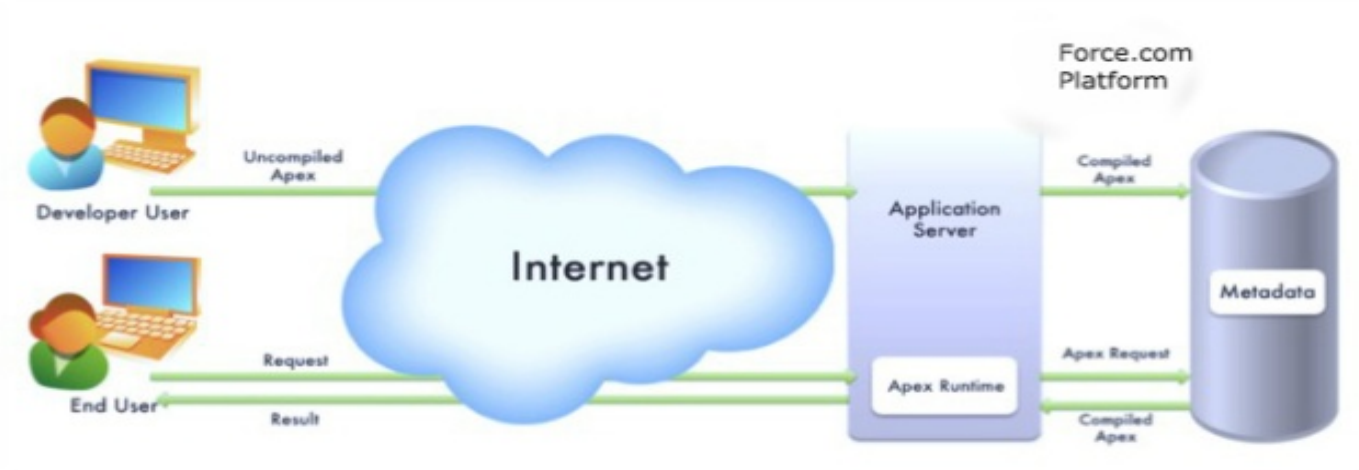
Built-in support for creation & execution of unit tests

LOOKS LIKE JAVA

Acts like database stored procedures

CLOUD HOSTED COMPILER

Interpreted & executed on on multitenant environment



Background: Hammer Process
(Aka Production Readiness Testing)

CUSTOMER UNIT TESTS ARE EXECUTED TWICE

in Salesforce secured environment in data centers

1st EXECUTION CALLED THE **BASELINE**
on current production version

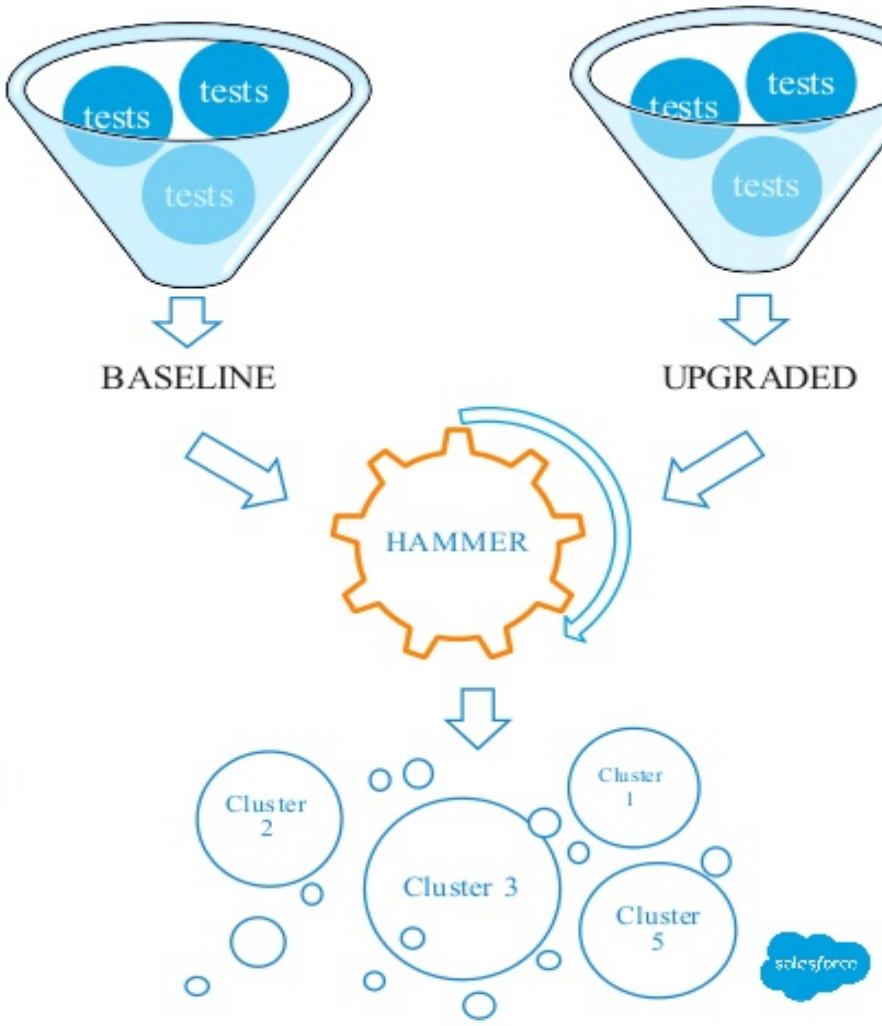
2nd EXECUTION CALLED THE **UPGRADED**
on release candidate version

UPGRADED AND BASELINE RESULTS ARE COMPARED

When a test passes, it should pass in both versions.

When a test fails, it should fail in both versions.

Any other outcome is a potential bug in release candidate version!



Challenges

150+ million customer tests and growing

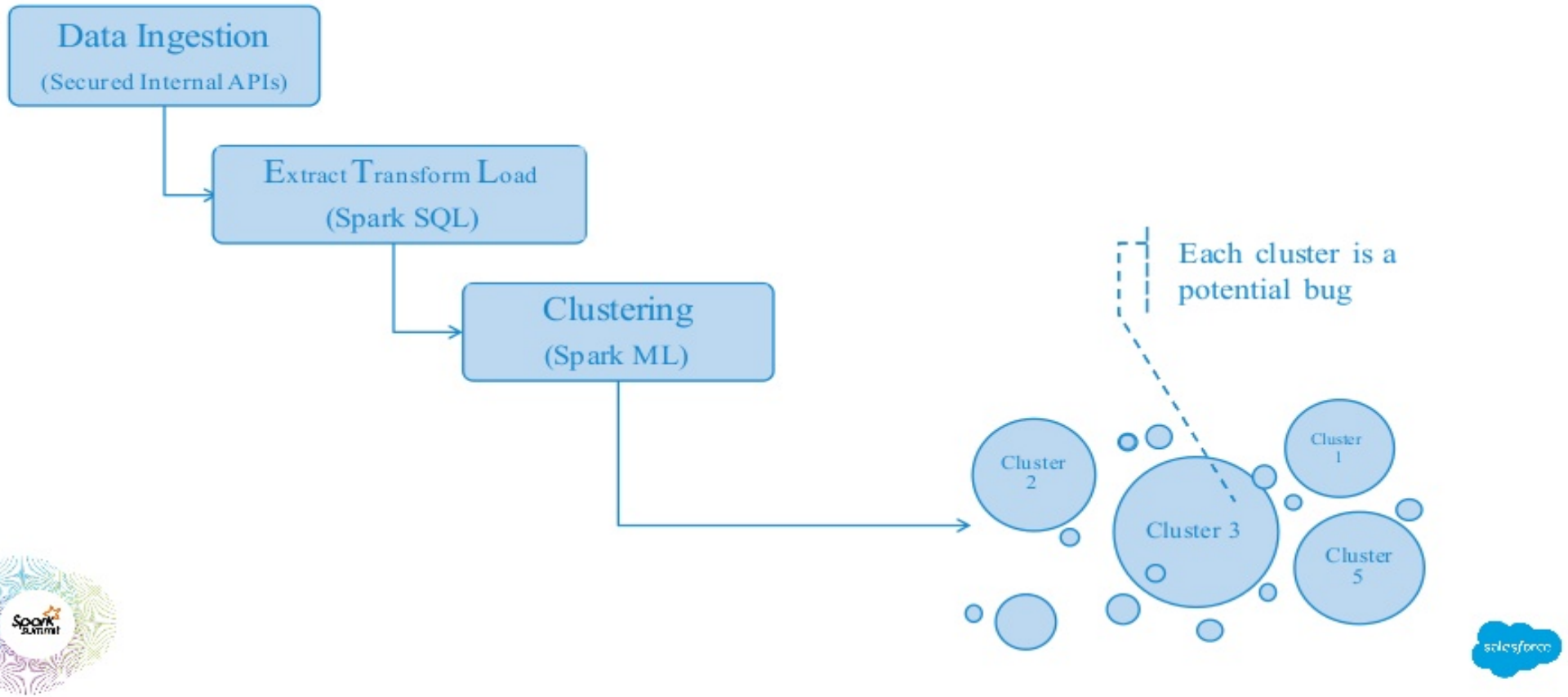
Infrastructure setup to execute hammer on two different platform versions

Persist and compare test execution results

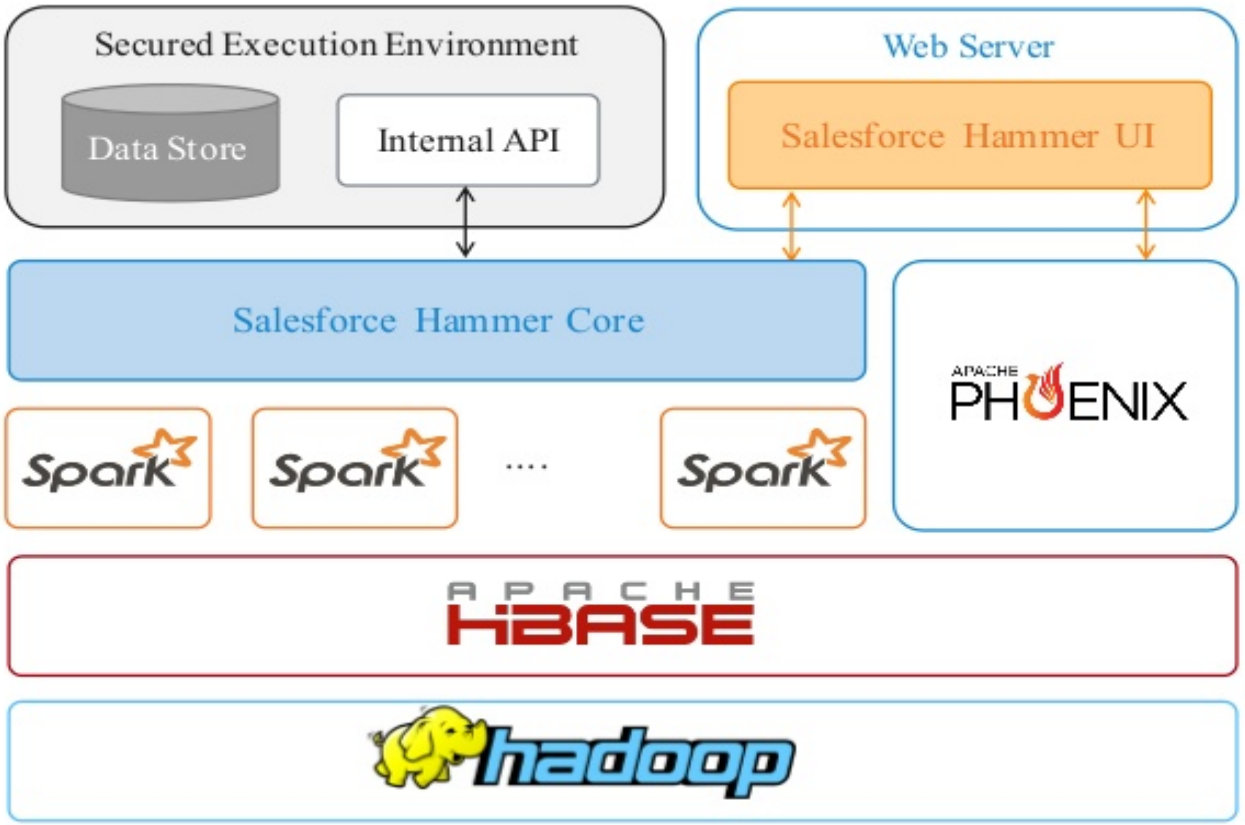
Tests are executed highly secured data centers

Internal SLA to keep these mammoth efforts to under 3 weeks

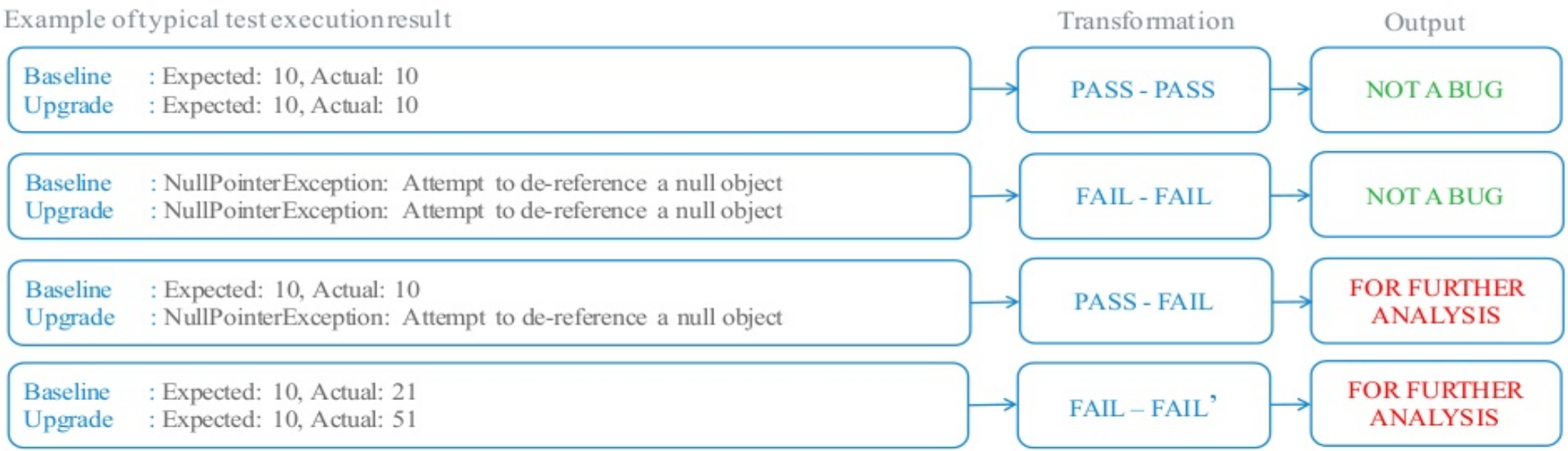
Hammer Core: A Functional Overview



Hammer: The Architecture



Data Preparation (ETL): Using Apache SparkSQL



Spark Machine Learning Pipeline

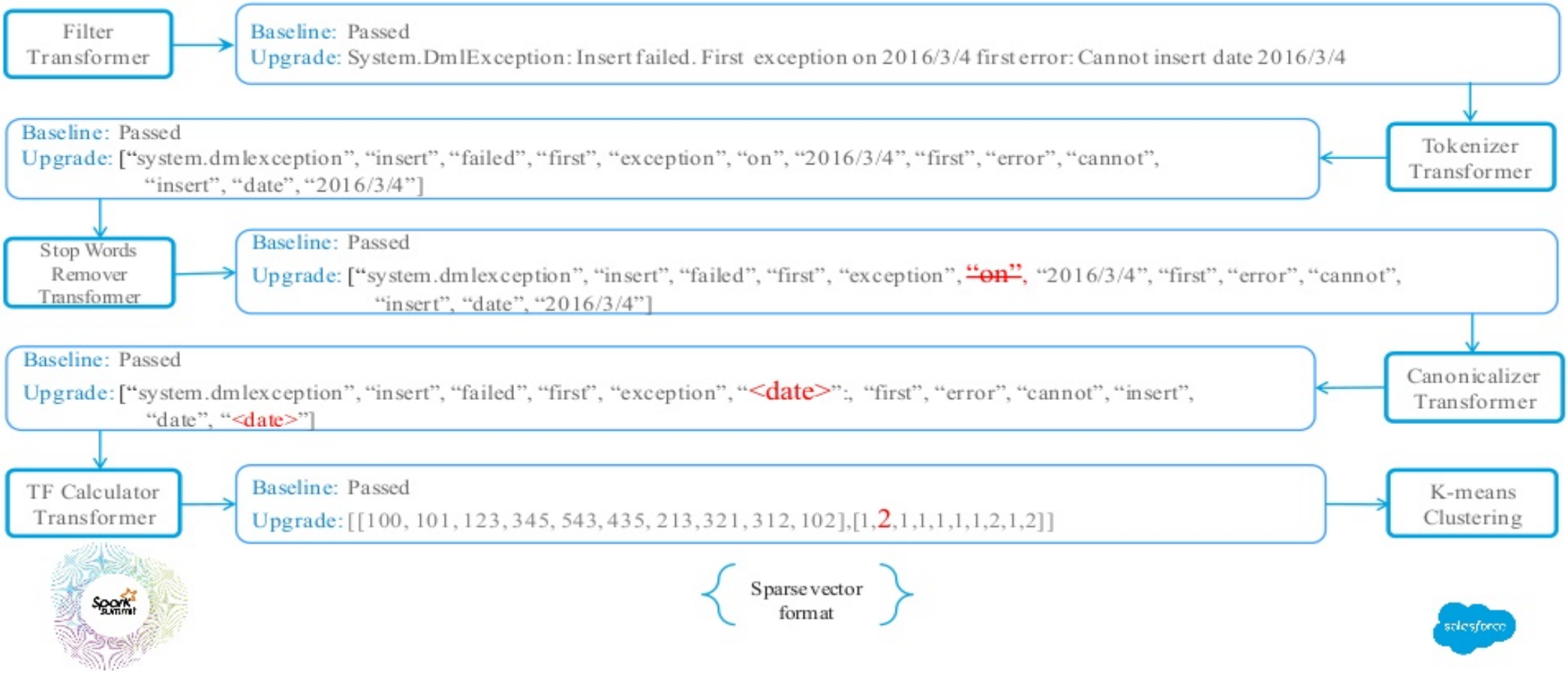
Designed to operate on records marked "FOR FURTHER ANALYSIS"

Group test failure records into K Clusters

Each cluster is a potential bug in Salesforce App Cloud platform

Enables inspection of cluster to determine if it's a bug or not

Clustering Using Apache Spark Machine Learning



Accomplishments

In Extract Transform Load process -

Test Records Analyzed	Old Hammer Engine (hours)	New Hammer with Spark (minutes)	Speed Improvement
241K	7.5	9	97.9 %
562K	7.8	13	97.2 %
269K	8	11.5	97.6 %
242K	11.2	10	97.9 %
394K	14.2	20	97.0 %
374K	12.5	12	98.4 %

Accomplishments

In Clustering Analysis -

- Fewer clusters to analyze
- Well formed clusters yielding to good quality bugs
- Speed – On an average clustering took 40 minutes to complete for 100K+ records

Q & A

thank you