VOLTDB AND FLYTXT PRESENT:
BUILDING A SINGLE TECHNOLOGY PLATFORM FOR REAL-TIME AND ITERATIVE ANALYTICS ON FAST + BIG DATA

OUR SPEAKERS

Ryan Betts
CTO at VoltDB

Prateek Kapadia
CTO at Flytxt

VOLTDB OVERVIEW

Founded by winner of the 2014 ACM Turing Award

Mike Stonebraker

Other Stonebraker Companies

Technology

- In-Memory (but data is durable to disk)
- Scale-Out shared-nothing architecture
- Reliability and fault tolerance

- SQL + Java with ACID
- Hadoop and data warehouse integration
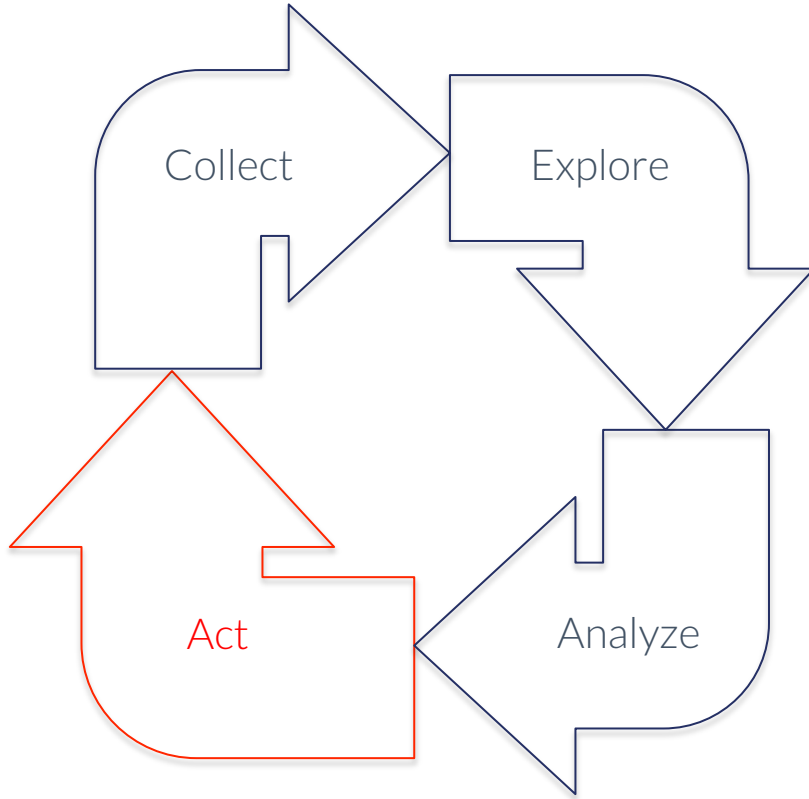- Open source and commercially licensed (24X7)

Customers

FAST

World Record Cloud Benchmark:

YCSB (Yahoo Cloud Serving Benchmark) - 2.4 million tps (transactions per second)
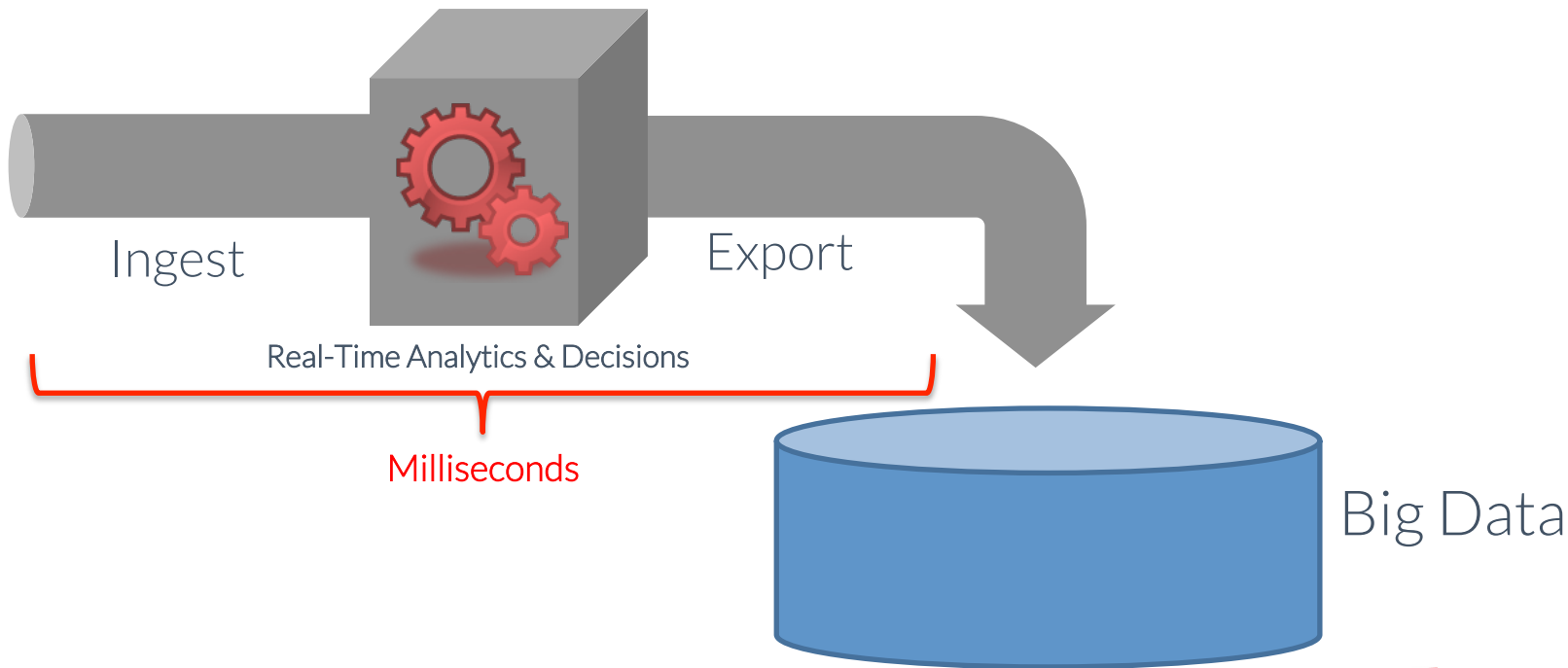
Big Data analytic results:

1. *Discoveries*: seasonal predictions, scientific results, long-term capacity planning

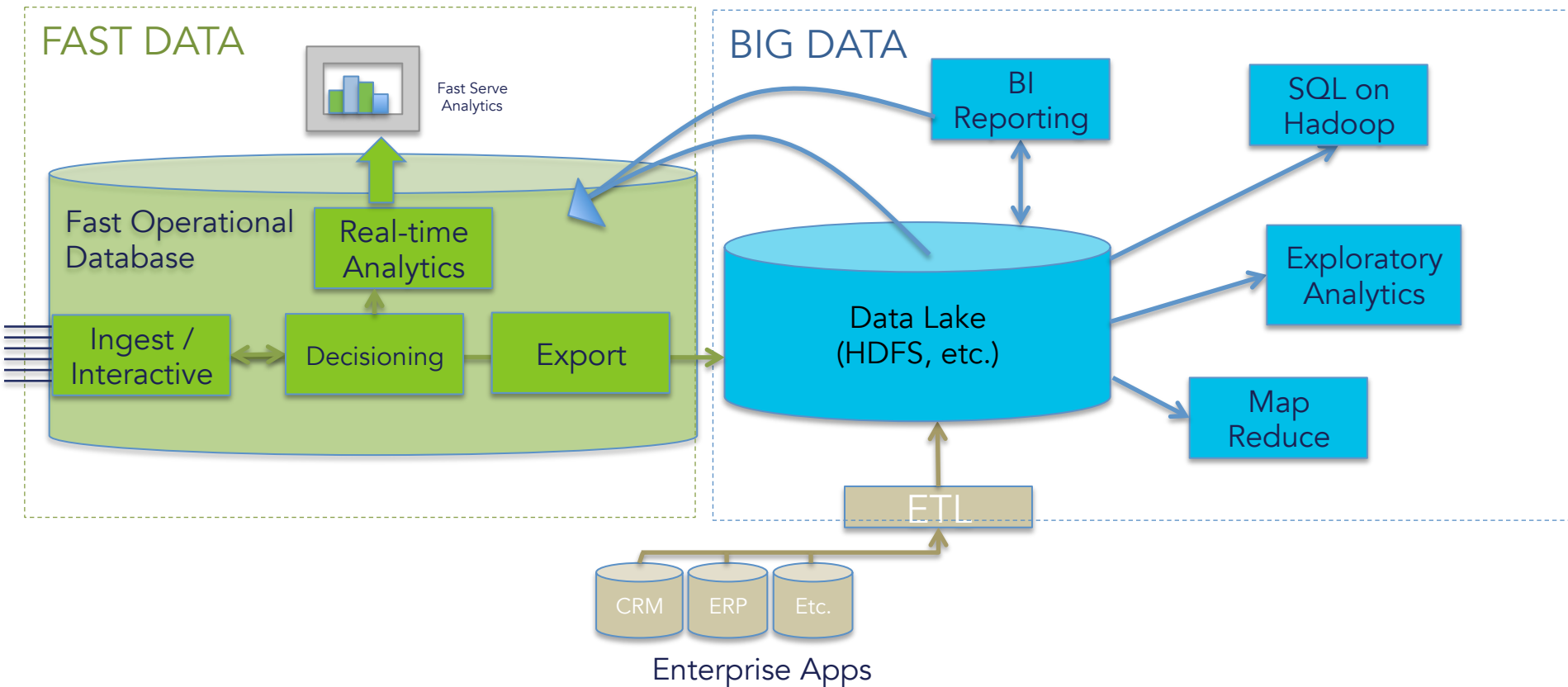2. *Optimizations*: market segmentation, fraud heuristics, optimal customer journey

FAST DATA – BIG DATA

Fast Data Pipeline

Fast Data: the velocity side of Big Data

Ingest

Export

Real-Time Analytics & Decisions

Milliseconds

Big Data

DATA ARCHITECTURE FOR FAST + BIG DATA

Gartner

"89% of marketers surveyed plan to compete primarily on the basis of customer experience by 2016."

Source: Gartner 2014 survey, Companies > $50M in revenue

FAST DATA SOURCES AND DRIVERS

Mobile

IoT

Social

Sensors

Logs

Data is doubling every two years

Mobile

9.5 BILLION

mobile subscriptions
by the end of 2020

90%

**of the world's population
over 6 years old will have
a mobile phone by 2020**

—	Mobile subscriptions	—	Fixed broadband subscriptions
••••	Mobile broadband subscriptions	••••	Mobile PCs, tablets and mobile router subscriptions
		—	Mobile subscribers

IoT

- 26 billion connected devices by 2020 (Gartner 2014)
- 37% of most data will be processed at the edge in milliseconds (Cisco IoT Study 12/11/14)

THE FAST DATA PIPELINE
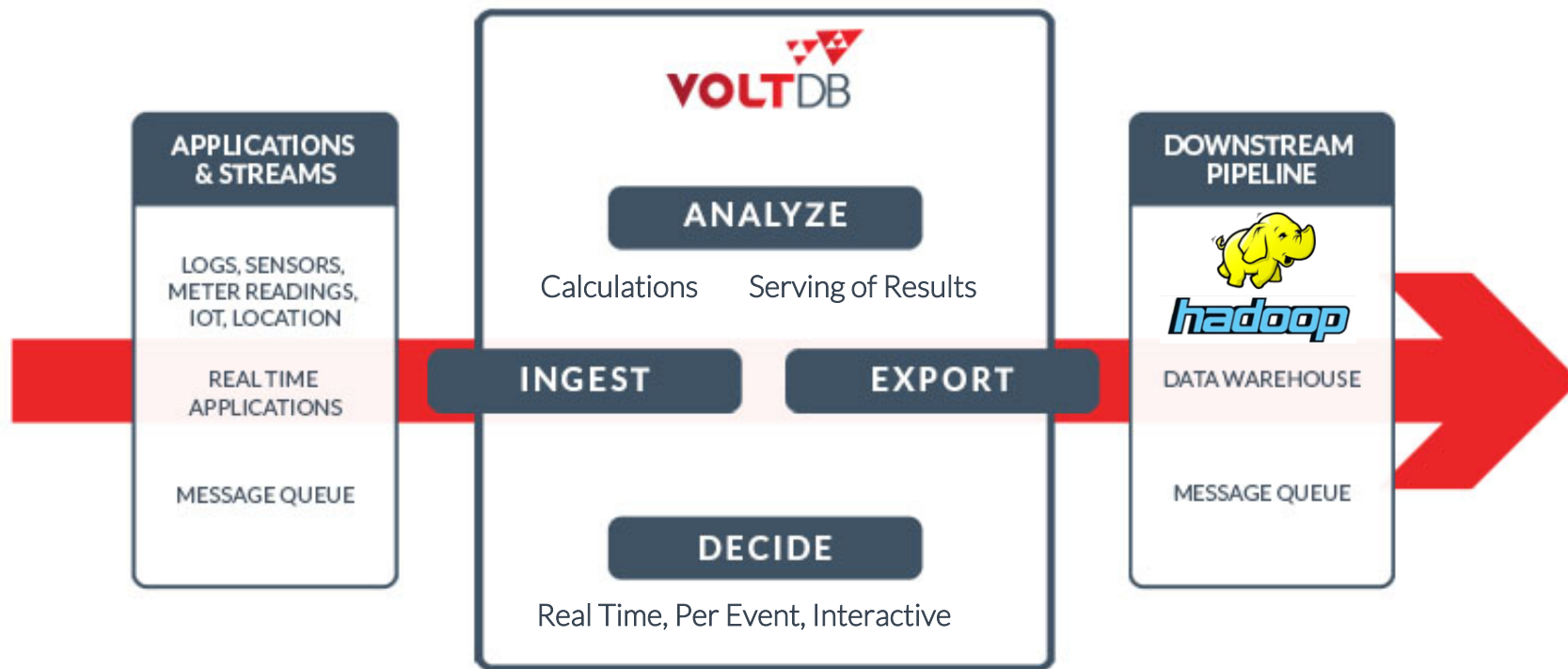
STREAMING: REAL TIME ANALYTICS

- Operational analytics and monitoring
- RT analytics enabling user-facing applications
- KPI for internal BI/Dashboards
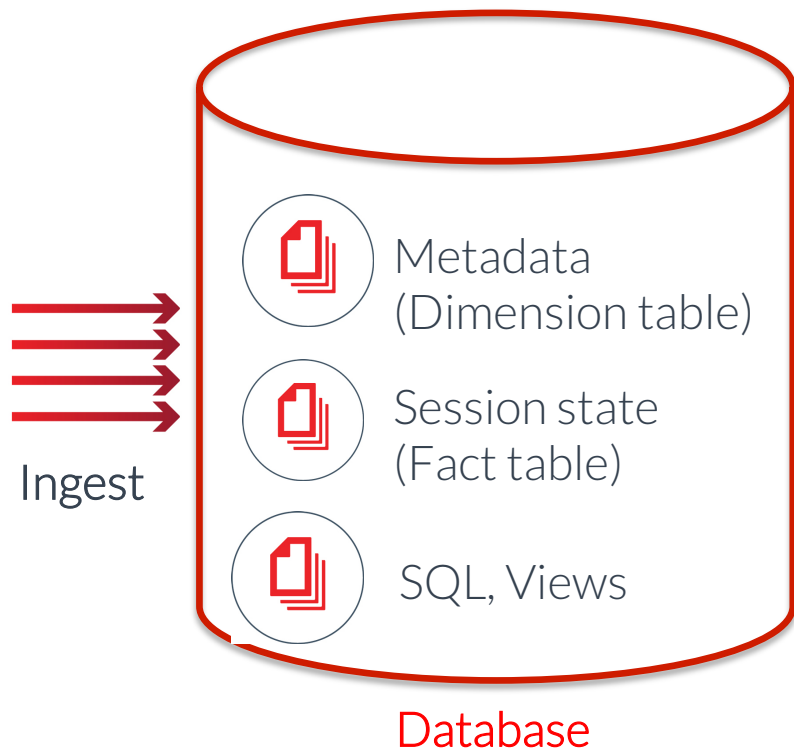
STREAMING OPERATORS NEED STATE

Require State

- Filter
- Join
- Aggregate
- Group By

Stateless

- Partition

REAL-TIME ANALYTICS

Ingest →
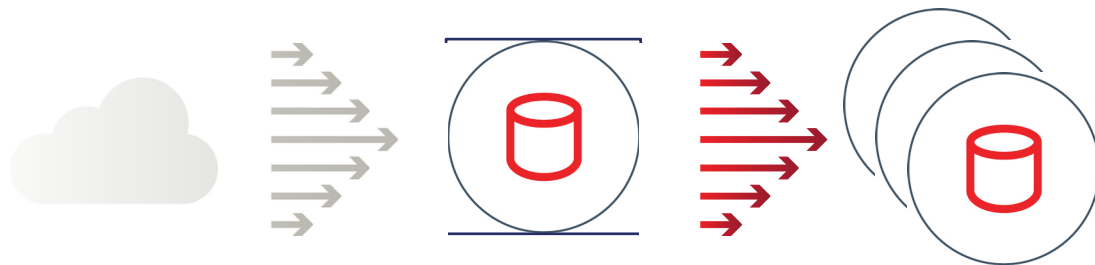
Database

- Metadata (Dimension table)
- Session state (Fact table)
- SQL, Views

- In-memory MPP SQL over ODBC/JDBC
- Cheap + correct materialized views for streaming aggregations

- Operational analytics and monitoring
- RT analytics enabling user-facing applications
- KPI for internal BI/Dashboards

INTEGRATING WITH EXPORT TARGETS

- Local file system export
- JDBC export
- Kafka export
- RabbitMQ export
- HDFS export
- HTTP export
- Extensible API

EXPORT FORMATS

- CSV
- TSV
- Avro container
- Raw data

DATA PIPELINES WITH EXPORT

Database

Metadata (Dimension table)

Session state (Fact table)
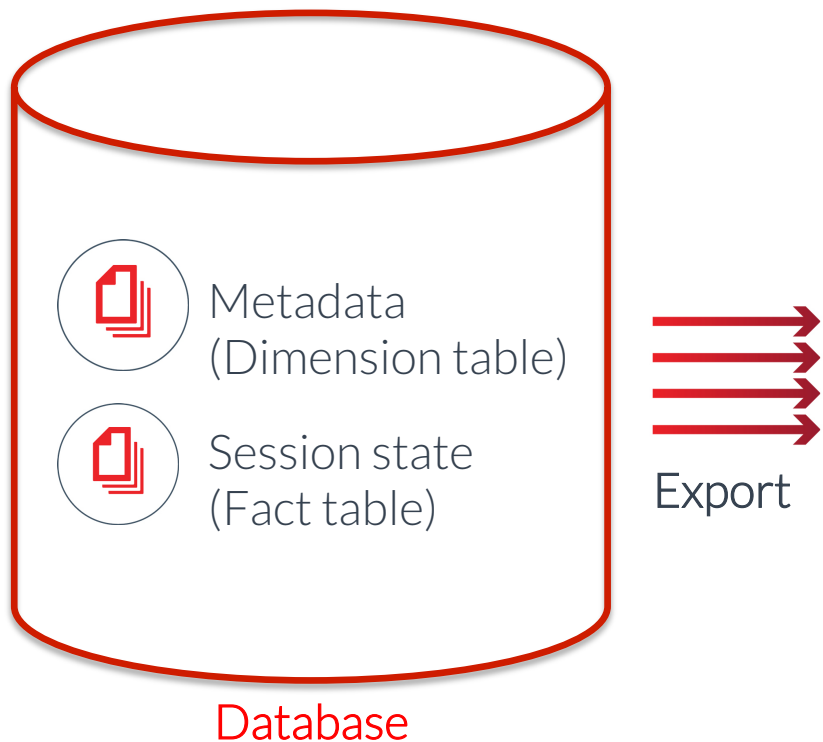
Export

- MPP streaming Export
- Row data, Thrift messages, CSV
- OLAP, HDFS and message queues

- Filtering (ex: only RFID / iBeacon readings that show change from previous location).
- Sessionization
- Common version re-writing
- Data enrichment

FLYTXT AND VOLTDB

FLYTXT OVERVIEW

Vision, Mission & Impact

- Vision: Create >10% measurable economic value for Communication Services Providers through Big Data Analytics

- Flytxt's internal and external monetization solutions increase revenue, reduce churn and improve customer experience

- Dutch company with corporate office in Dubai, global delivery centres in India and regional presence in Mexico City, Johannesburg, Singapore, Dhaka and Nairobi.

Awards & Achievements

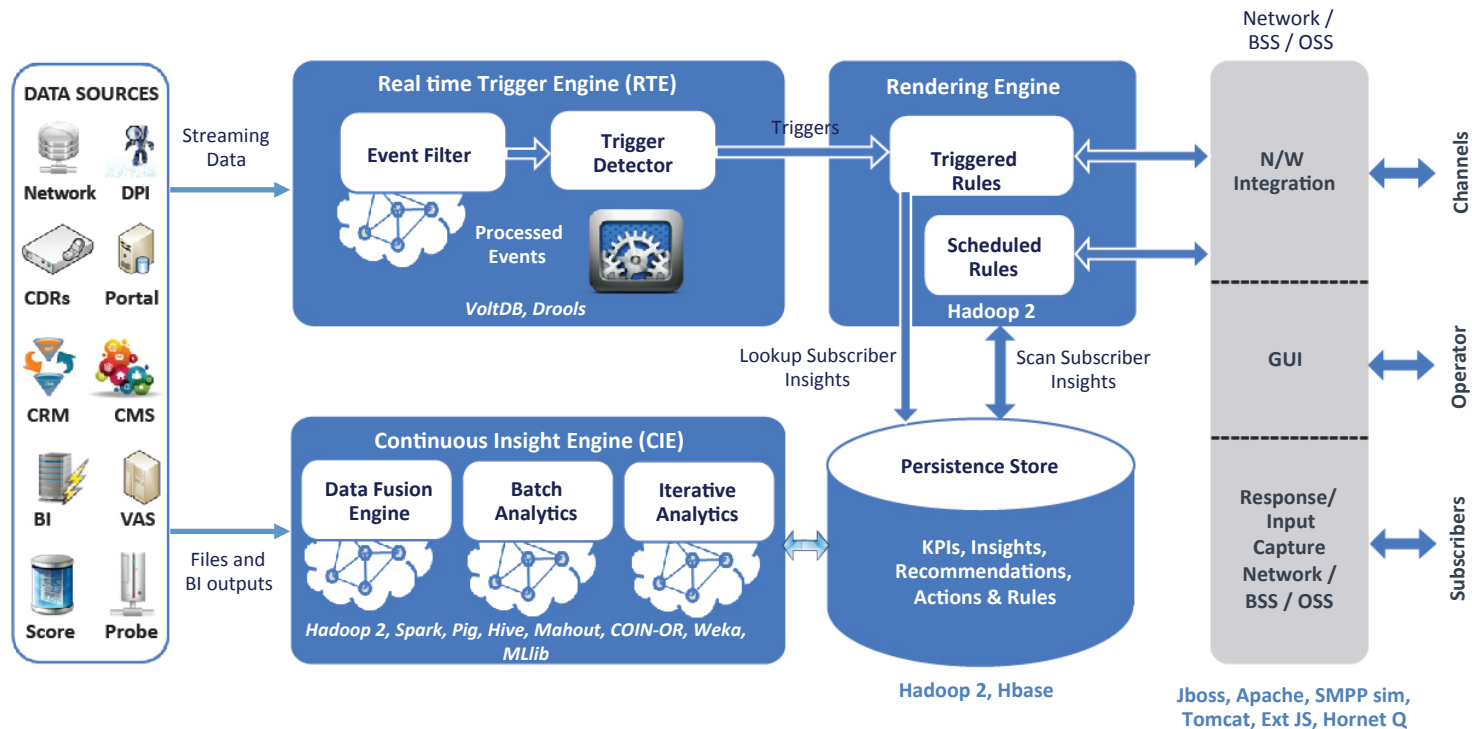## Customers and Partners
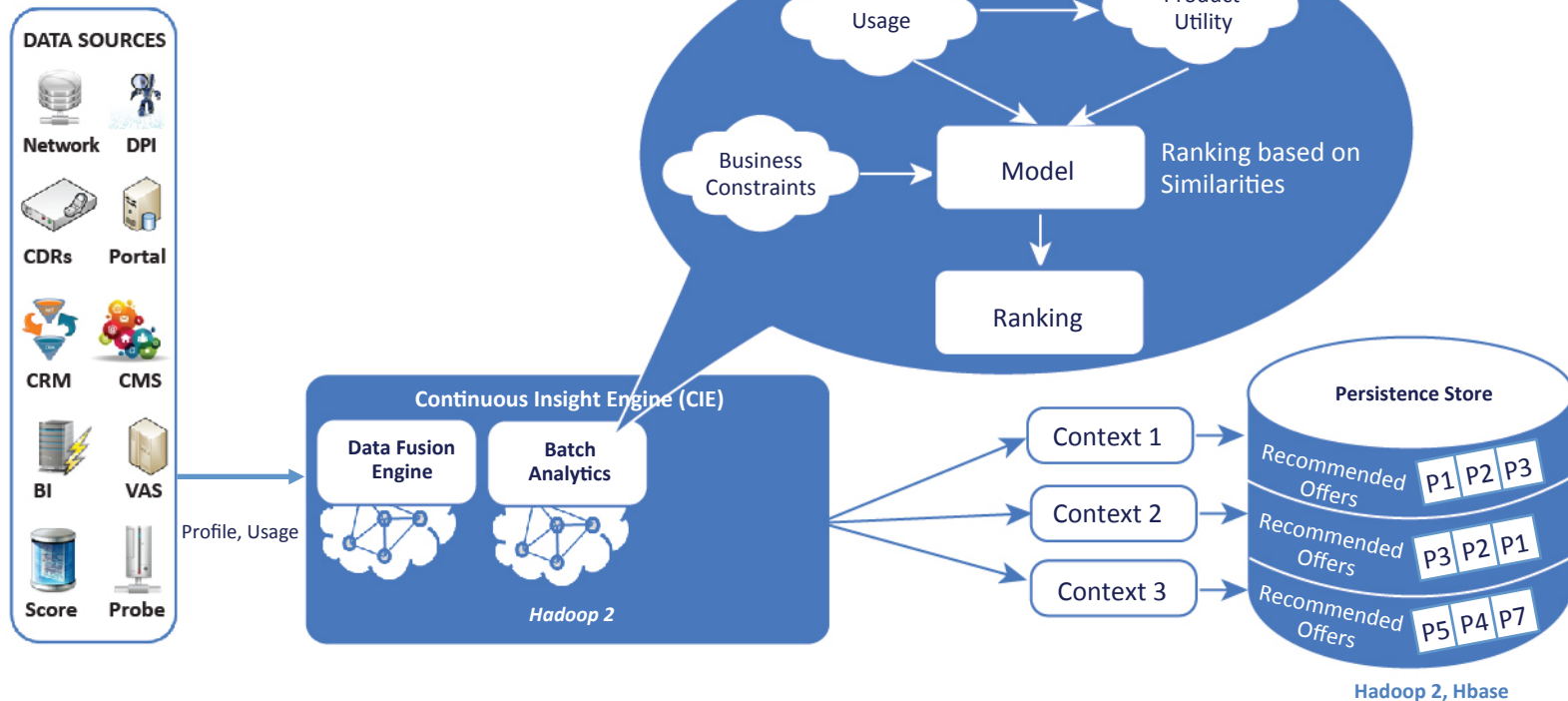
Operators

Brands

Partners

FLYTXT'S INTEGRATED ANALYTICS SOLUTION ARCHITECTURE: BIG DATA, ITERATIVE AND REAL-TIME ANALYTICS

BIG DATA ANALYTICS USE CASE: BEST FIT PRODUCT RECOMMENDATION

Objective: Recommend best fit product to subscribers based on usage and business objectives

CASE STUDY: CONTEXTUAL PRODUCT RECOMMENDATION FOR TIER 1 OPERATOR
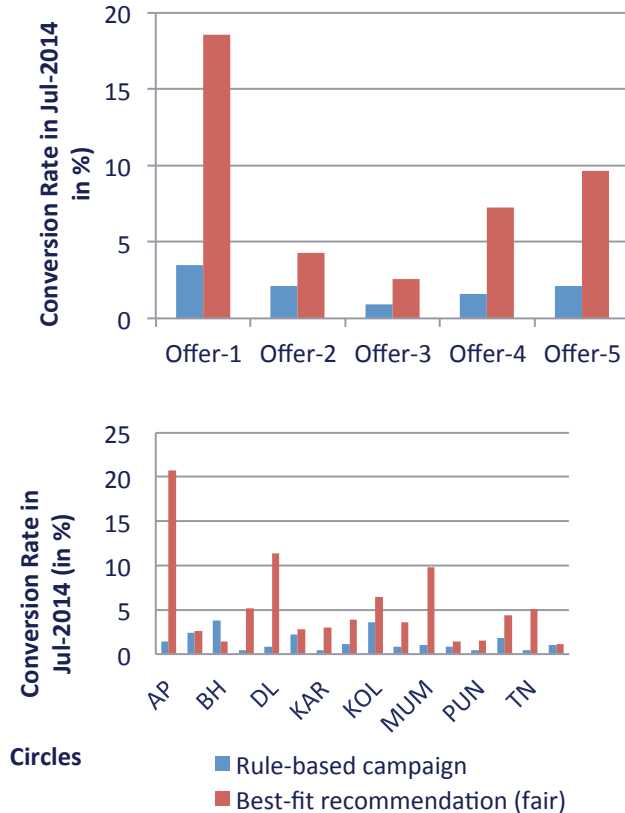
Recommendation Personas:
CLV (HVC, MVC, LVC), Volatile,
Early Adopter, Frequent Handset Changer,
Heavy Data user, Social Media Fan,
Bollywood Fan, Music Fan, Sports Fan,
potential iPad buyer, International Caller
Etc.......

Objectives:
Cross sell, Upsell, Stimulate recharge/
usage/Service adoption
Etc...

Offers:
Data Plan, 3G plan, VAS usage, International
Calling packs, Bundle offers, Recharge
stimulation, Seeding, ebill subscription
etc......

Channels:
IVR, In store, Retailer, WAP portal,
Customer care portal

ITERATIVE ANALYTICS USE CASE: MICRO-SEGMENTATION

DATA SOURCES

Network DPI

CDRs Portal

CRM CMS

BI VAS

Score Probe

Files and BI outputs

Clustering to enable micro segmentation for personalized offers

Hard Clustering

Soft Clustering

Gaussian Mixture Model

$$J_{\mathrm{Kmeans}} = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{\mathrm{nk}}\, dist(x_n - \mu_k)$$

Continuous Insight Engine (CIE)

Data Fusion Engine

Iterative Analytics

Spark, MLlib

Persistence Store

Segment Offers S1 → P3

Segment Offers S2 → P8

Segment Offers S4 → P9

Hadoop 2, Hbase

Micro-segmented Offers

USE CASE: REAL-TIME ANALYTICS SUPPLEMENTED BY BIG DATA, ITERATIVE ANALYTICS

Real-time analytics

DATA SOURCES

Network | DPI
CDRs | Portal
CRM | CMS
BI | VAS
Score | Probe

Streaming Data

Event Filter → **Trigger Detector**

Processed Events

VoltDB, Drools

Balance Threshold
Recharge
Usage
Network

Triggered Rules

Scheduled Rules

Hadoop 2

Real-times Actions

Files and BI outputs

Continuous Insight Engine (CIE)

Data Fusion Engine | **Analytics Engine** | **Iterative Analytics**

Hadoop 2, Spark, Pig, Hive, Mahout, COIN-OR, Weka, MLlib

Behaviour | Preference | Recommendation | Contextual Needs

Persistence Store

KPIs, Insights, Recommendations, Actions & Rules

Big Data+ Iterative analytics
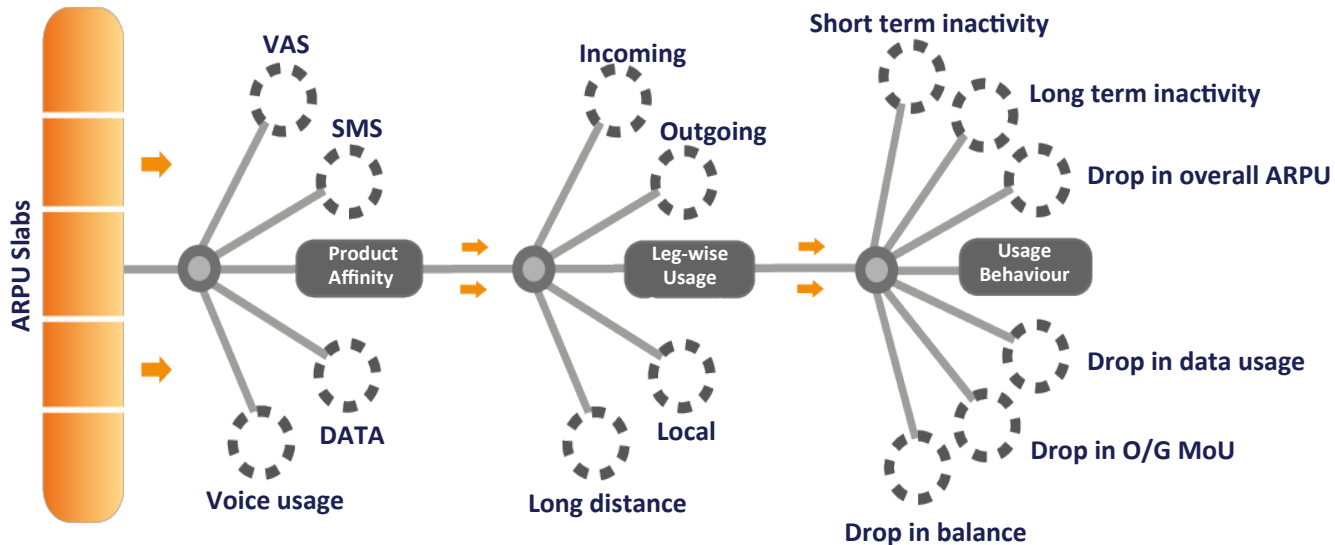
Hadoop 2, Hbase

CASE STUDY: REAL-TIME TRIGGER BASED MICRO-SEGMENTED OFFERS

Client Objective

- Improve customer engagement for ARPU enhancement

Data Analyzed

- Customer usage history, ARPU charts, spend patterns and preferences

ARPU Slabs

VAS
SMS
Product Affinity
DATA
Voice usage

Incoming
Outgoing
Leg-wise Usage
Local
Long distance

Short term inactivity
Long term inactivity
Drop in overall ARPU
Usage Behaviour
Drop in data usage
Drop in O/G MoU
Drop in balance

Solution

- Marketing Program based on usage behaviour driven micro-segmentation and tripwire monitoring

Impact

- 2% increase in month-on-month revenue
- 28% higher revenues & MOU

QUESTIONS?

- Use the chat window to type in your questions or hashtag #VoltDBFlytxt

- Know more about Flytxt
  - Visit www.Flytxt.com

- Try VoltDB yourself:

  ➢ Free trial of the Enterprise Edition:
    - www.voltdb.com/download

  ➢ Try VoltDB in the Cloud
    ➢ Amazon's Cloud Formation

  ➢ Open source version is available on github.com

THANK YOU!