

 Bigdatastack.eu

 @bigdatastackeu



Connected Consumer Technologies for Retailers

BigDataStack

Bernat Quesada / WORLDLINE

Co-funded by the European Commission
Horizon 2020 - Grant # 779747



1. Business context and opportunities

- ✓ Retail business – Highly competitive market, each customer matters!
- ✓ Need for an attractive offer and differentiate from competitors
- ✓ Personalization still a challenge! Today still too simple, based only on:
 - Traditional and outdated customer segments: age, gender and residence
 - Own historical data, not linked to tendencies or context
 - Not linked to external data sources, significantly social media
 - Unable to produce predictions
- ✓ Predictive analysis
 - Adapt most appropriate messages to each customer
 - Send only those promotions that best suits customer needs at the right time
 - Predict potential buyers for products or services
 - Improve shopping experience for consumers
 - Open new business opportunities for retailers

Worldline is collaborating along with **one of the largest Spanish distribution companies.**

More than 3000 physical stores
1.2 million products
1 million customers
3000 million ticket items (2 years history)

Altogether 1TByte of data

The **challenge** is to **build a collaborative-filtering recommender system over BigDataStack** that helped our partner to personalize product predictions to their customers. Thus, increasing customer satisfaction and loyalty.



The analytic algorithm chosen to calculate recommendations has been Alternating Least Squares (collaborative filtering). Concretely, the implementation of ALS in Spark ML Lib.

Prior to invocation of the algorithm, calculation of the input matrix needs to be done.

- Affine customers
- Rating for the products

	Product 1	Product 2	Product ...	Product M
User 1	1	BLANK	BLANK	3
User 2	BLANK	5	BLANK	3
User 3	BLANK	BLANK	1	BLANK
User 4	2	3	BLANK	BLANK
User 5	BLANK	BLANK	1	BLANK
User 6	4	BLANK	5	BLANK
User 7	BLANK	4	BLANK	BLANK
User ...	BLANK	3	BLANK	BLANK
User N	BLANK	BLANK	BLANK	4



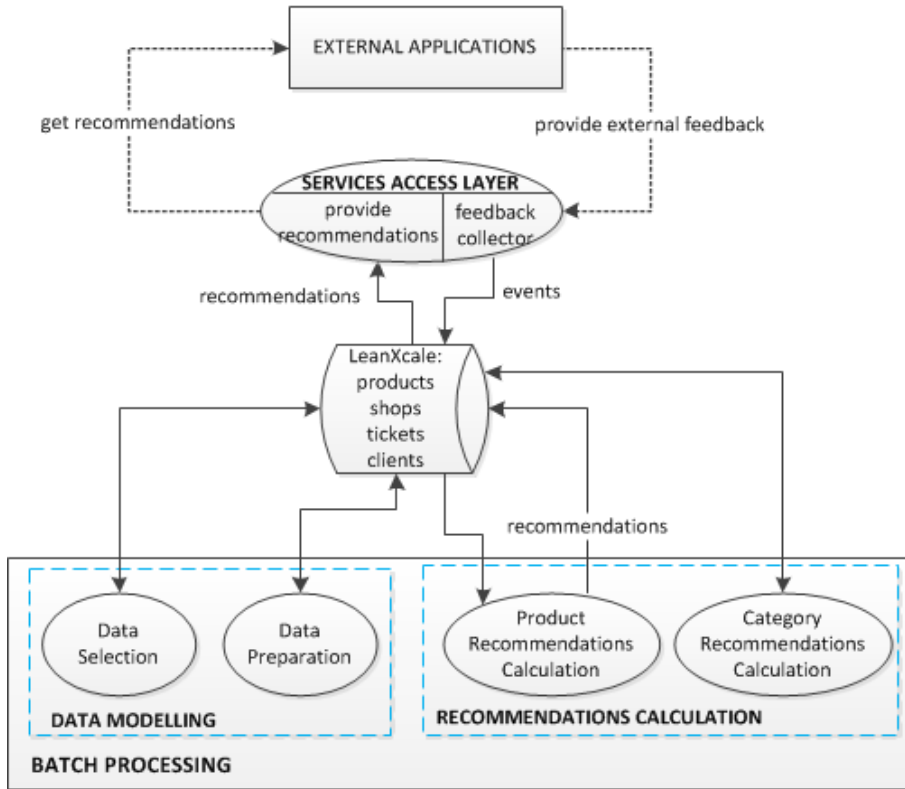
	Product 1	Product 2	Product ...	Product M
User 1	1	4	2	3
User 2	1	5	3	3
User 3	2.5	2.8	1	3.5
User 4	2	3	2	3.5
User 5	2.5	2.8	1	3.1
User 6	4	1.2	5	1.4
User 7	1	4	2.5	3
User ...	2	3	2	3
User N	1	4	2	4

Business constraint: recommendations should be calculated every day before 6:00 a.m.
Business constraint becomes a business objective for the system



3. Use case – Application Deployment

KPIs definition



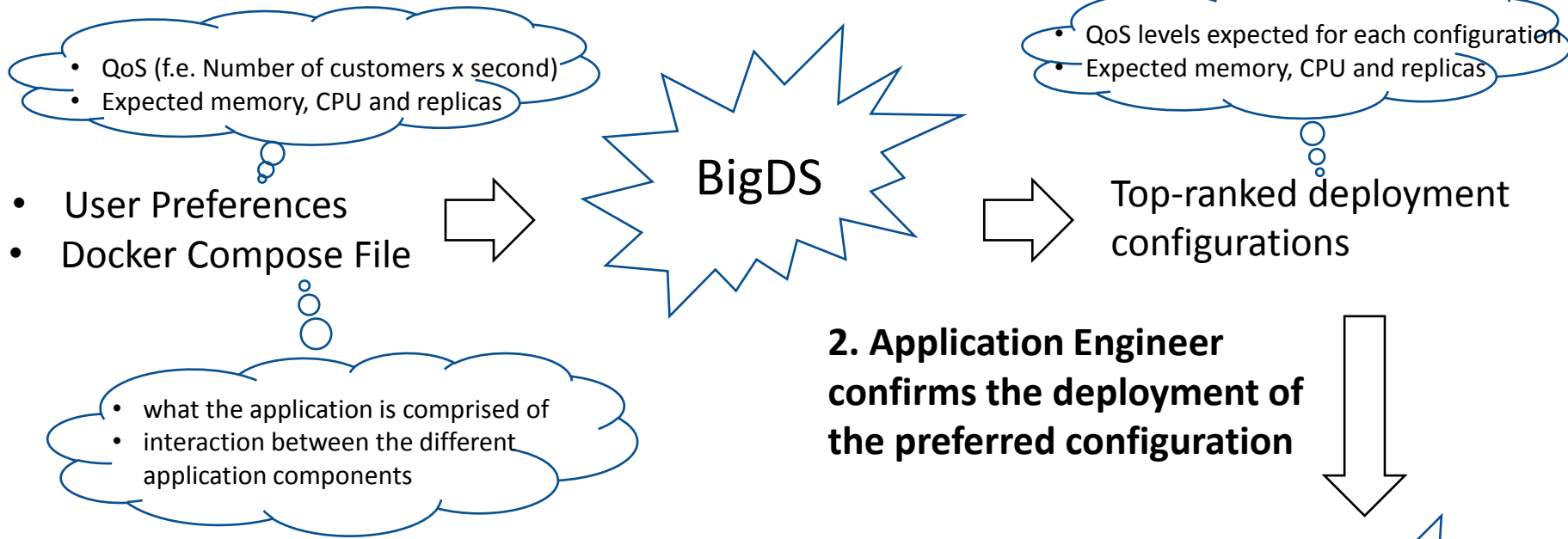
Provide Recommendations	Time response. <200ms x request	REST Web Service
Feedback Collector	Time response. <200ms x request	REST Web Service
Data Selection	Throughput. Number of customers/second	Python application
Data Preparation	Number of customers/second	Python application
Products Recommendation Calculation	Number of recommendations / minute	Python application + Apache Spark
Category Recommendation Calculation	Number of customers/second	Python application

1 Docker Image for each application service



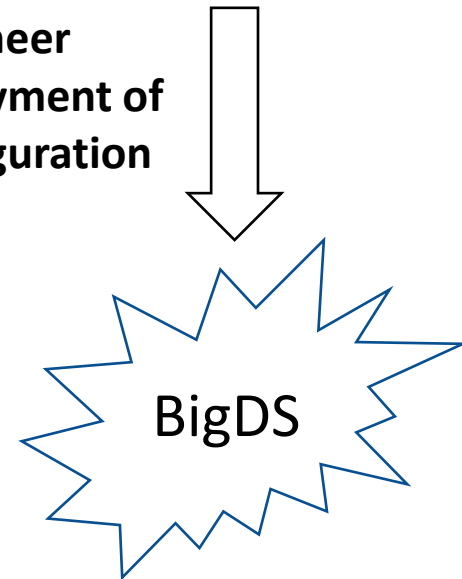
3. Use case – Application Deployment

1. Application Engineer feeds BigDataStack with



2. Application Engineer confirms the deployment of the preferred configuration

Application deployed in OpenShift



4. Solutions provided by BigDS to the retail ecosystem



**Required
resources**

**Provisioning of
resources**

**Run-time
adaptations**

**End Users
tools**



Thanks!

Bernat Quesada
Project Manager
Worldline
bernat.quesada@worldline.com



Use case – Next Steps

■ Business/Functional Objectives

- Introduce a clustering algorithm for Segmentation of customers in the current scenario
- Create a model for prediction of the next customer purchase – predictive shopping list

■ Integration with BigDataStack

- use CEP and gateway for injecting data in real-time from heterogeneous datasources such as programmed events or weather information
- data provided by our partner has proved to be full of mistakes, integration with BigDS data services such as data cleaning will help to save time to the data scientist, ...
- integration with the end-user tools of BigDS such as Process Modeling and Data Toolkit. since we want to be able to inject new analytical models



Thanks!

Bernat Quesada
Project Manager
Worldline
bernat.quesada@worldline.com

