Corporate Presentation

Version 1.0 - 2021

xpandit

xpandit
Best 2020
Portuguese SME
ageas seguros
1000 PME
2020
Exame

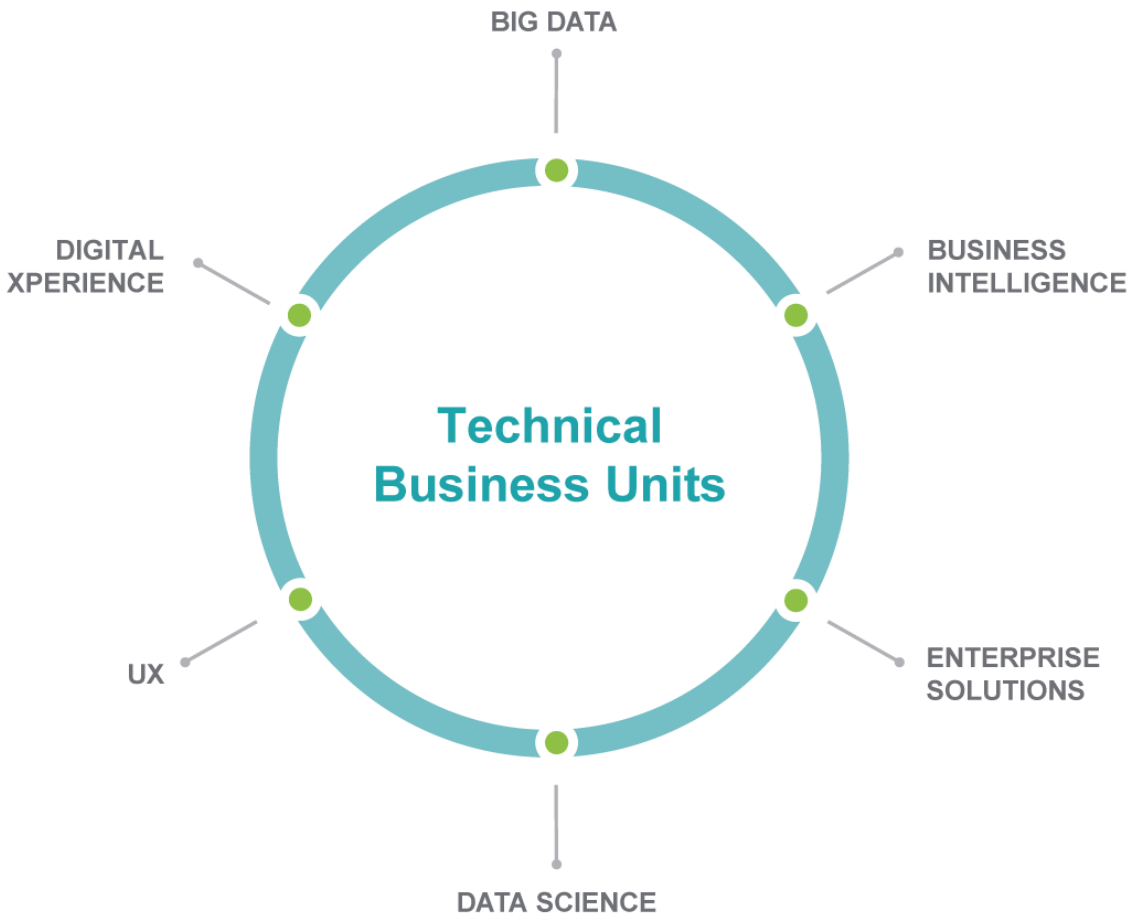
Locations

Numbers

+290 **Technology Experts**

FT1000: Europe´s Fastest Growing Companies

Expertise


BIG DATA
DIGITAL XPERIENCE
BUSINESS INTELLIGENCE
Technical Business Units
UX
ENTERPRISE SOLUTIONS
DATA SCIENCE

Expertise in Data

Complete Data Lifecycle

Big Data

- Infrastructure
- Data Lake
- Data Ingestion
- Realtime Processing

Business Intelligence

- Datawarehousing
- Datamodelling & ETL
- OLAP
- Analytics

Data Science

- Data Wrangling & EDA
- Machine / Deep Learning
- Advanced Analytics
- Optimization & Produtization

Proprietary & Confidential

Experience

Partnerships & Expertise

Proprietary & Confidential

Big Data Team

Partnerships & Certifications

Expertise with flexibility

We improve agility and performance through technology & skills

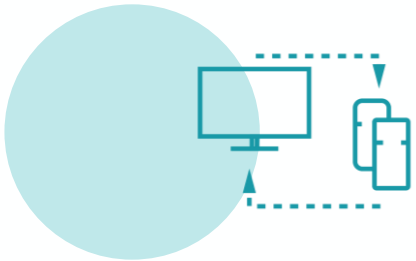
Big Data Consulting

You know your business. We know DATA and we strive to work with you to envision all possible ways to extract value from it, in order to define and develop a Big Data execution plan.

Proprietary & Confidential

Expertise with flexibility

We improve agility and performance through technology & skills



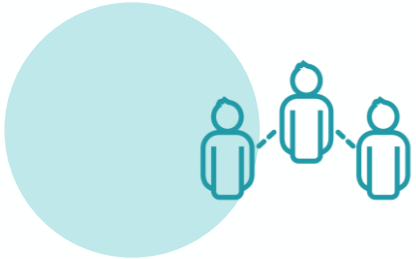
Big Data Development

Our Big Data experts with mastery in **Cloudera, Databricks, Confluent, NoSQL** Databases and Azure Big Data stack, can help you develop NRT (near real-time) and batch data pipelines for all types of data.

Proprietary & Confidential

Expertise with flexibility

We improve agility and performance through technology & skills



Big Data Operations

We help customers define, install, configure, manage and tune a distributed data environment. We can help you install and adopt Big Data software based on **Cloudera**, **Databricks, Confluent** or **Azure**.

Certifications in Data

Gold Microsoft Partner

Microsoft

- Azure Administrator / Security
- Implementing and Designing an Azure Data Solution
- Designing and Implementing a Data Science Solution on Azure
- Analyzing Data with Microsoft Power BI

CONFLUENT

- Confluent Certified Developer for Apache Kafka
- Confluent Certified Administrator for Apache Kafka

Proprietary & Confidential

Certifications in Data

- Cloudera Certified Administrator for Apache Hadoop
- Cloudera Spark and Hadoop Developer

- Databricks Certified Developer

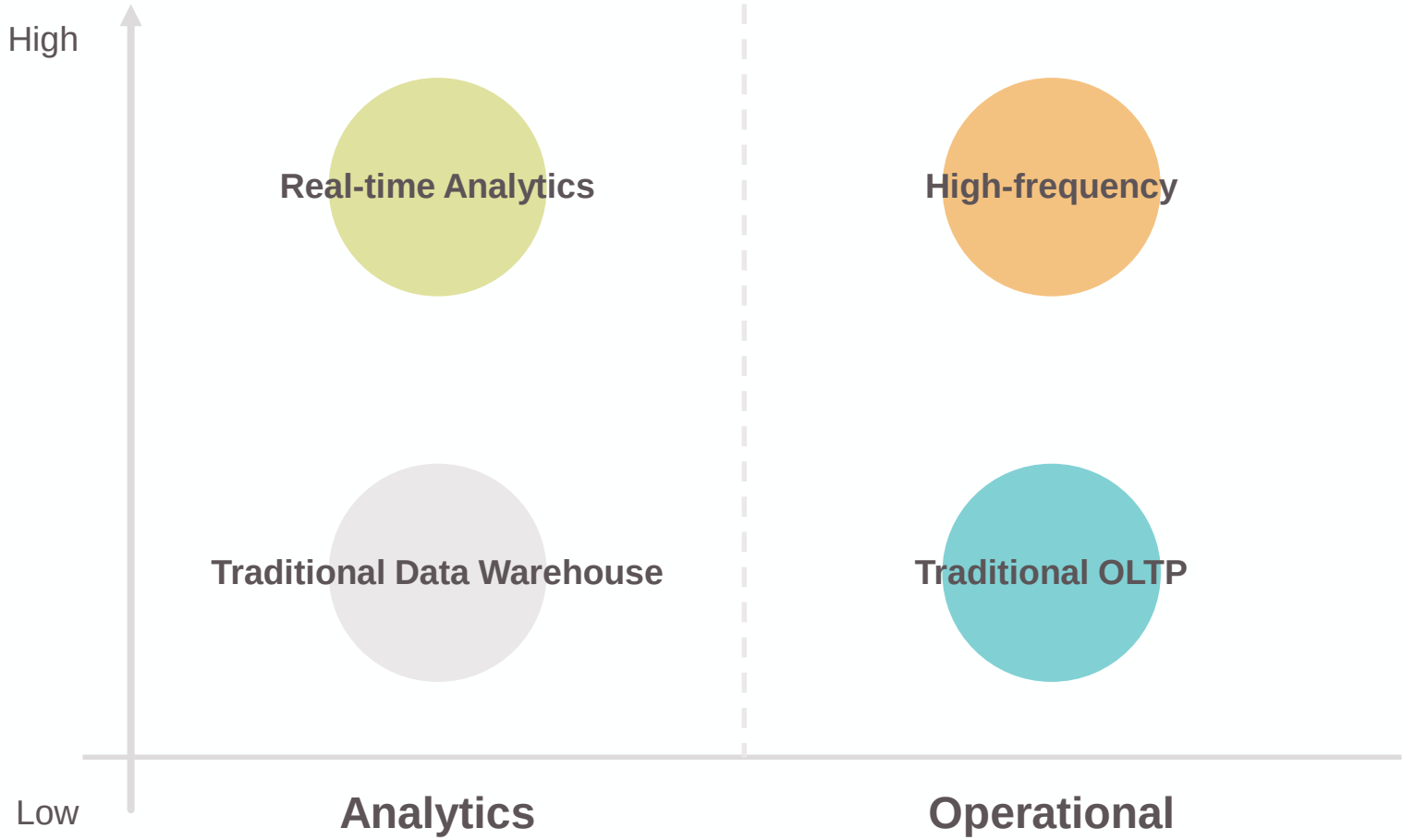
Proprietary & Confidential

What we really are

Spark | Kafka | Analytics | Streaming | Governance | Security | Data Movement

Proprietary & Confidential

Big Data



Proprietary & Confidential

Use Case – Ingestion

CHANGE DATA CAPTURE

DATA BROKERAGE

DATA CONSOLIDATION

DATA STORAGE

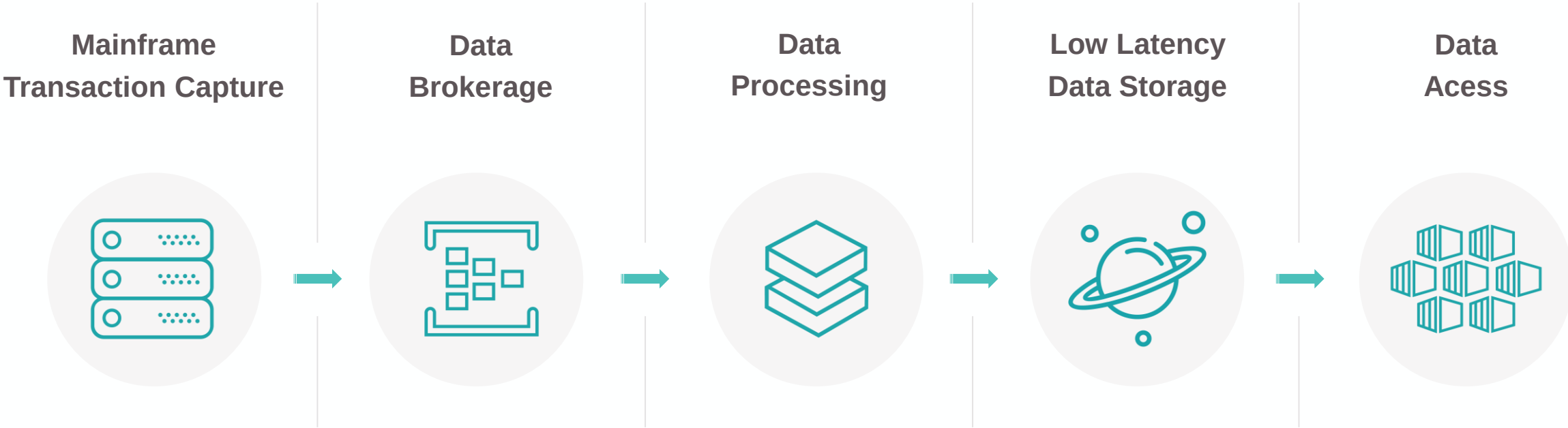
Proprietary & Confidential

Use Case – Data Backend

EVENT GENERATION	DATA BROKERAGE	DATA PROCESSING	DATA DELIVERY
In App Purchase			Invoicing
In App Event			Recommendation
In App Failure			Error Log

Proprietary & Confidential

Use Cases – Mainframe Offloading



Proprietary & Confidential

References in Data

European Agency

European Entity

The challenge was to develop a software solution supported on **Confluent Kafka platform**, based on several micro-services acting as interfaces between low-level/raw formats and higher-level formats, including business logic implementation on **streaming data**. Also, Confluent was used not only as central hub, but also to implement buffering from producers and consumers, built as a **mash of Kafka clusters**. The solution allowed for speed of processing, with low latency end-to-end, resilience and easy scalability (either horizontal or vertical).

Highlight Technologies:

CONFLUENT

Proprietary & Confidential

Insurance

Insurance Leader

The customer wants to implement a platform that can aggregate all relevant and needed information about its customers for search and access of other Applications (mobile apps, website and others), building a Customer Data Hub. Xpand IT was responsible to setup, implement and produtize the platform, based on Confluent, and the ingestion processes that feeds the final data sources.

Highlight Technologies:

CONFLUENT

mongoDB

Proprietary & Confidential

Retail

Major Retail

The challenge was built a full near-real time solution, leveraging the Kafka component. To enable generic applications to produce data a generic ingestion message driven API, was developed and worked a major frontend for data ingestion. To process (filter, enrich, streamline) and deliver (to several different targets, implementing different delivery semantics) the incoming messages a set of Spark streaming application were developed. The solution was developed in a way that new flows could be added by configuring them (without the need of any coding). The customer was then able to achieve the goal of having a robust, high-scale and low-latency solution.

Highlight Technologies:

CLOUDERA

APACHE Spark

SABER MAIS

Proprietary & Confidential

Utilities

Major Utilities

The customer wants to build a Data Lake with an ingestion strategy that allowed for CDC based data to be integrated in fast cycles in the consolidated data. This would allow the customer to build a set of solutions, both traditional and advanced analytics. The challenge was very much tied with the size of some tables (several billions) and the number of changes some tables suffered on a daily basis (one table with 160.000.000 changes per day), which due to their size and RDBMS support couldn’t cope with any analytics type of scenarios.

The solution used CDC data from Golden Gate delivered onto a Kafka cluster to consolidate target table data in fast cycles. The overall solution was able to import historical data (batch) and then merge onto a CDC mode, incorporating DML changes in cycles that could be as short as a few minutes. The Data Lake become a major enable of new use cases and value creation.

Highlight Technologies:

CLOUDERA

SABER MAIS

European Agency

European Entity

The challenge was to develop a **classification solution of events in NRT (near real-time)** and propose an architecture for the deployment and monitoring.

Due to the large volume of messages received every day (~20 million messages from terrestrial and satellite sources), the modelling solution was built using python.

The implemented solution included a **viability analysis** about rich datasets, **determination of feasible models**, **modelling and productization** of a classification solution and **technical and technological support**.

Highlight Technologies:

CONFLUENT

python™

Proprietary & Confidential

Obrigado

Luís Vicente | Marta Gomes | Nuno Barreto