

4 Reasons

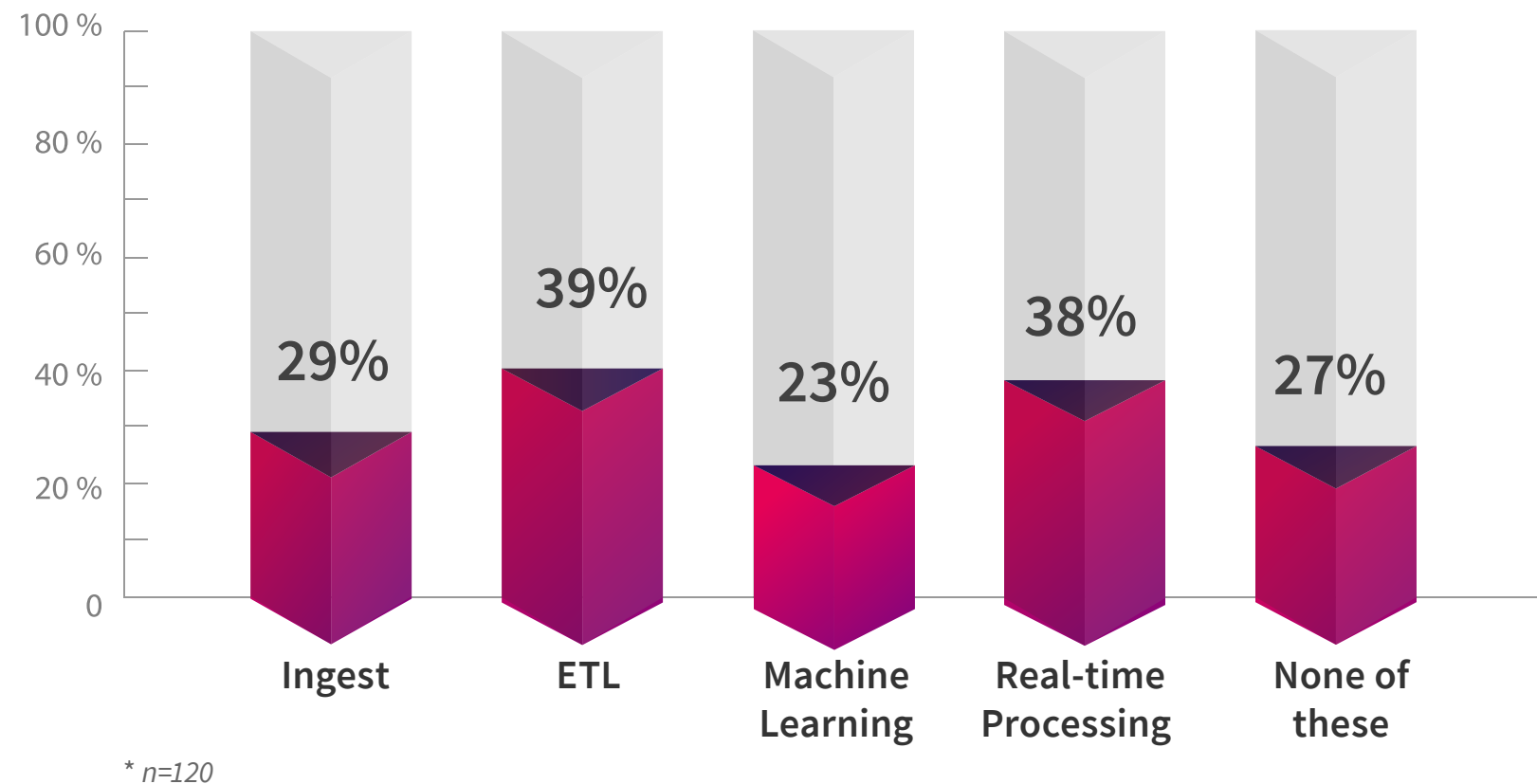
to choose Apache Spark over multi-vendor approach for all data processing needs

Open source systems are being widely adopted and deployed across enterprises. As more and more organizations take the open source route, Apache Spark stands out as the leading big data analytics platform for large scale batch processing, stream processing, and machine learning. Let's find out what makes Apache Spark the enterprise backbone for all types of data processing workloads.

How are enterprises currently using Apache Spark?

In a recent survey of large enterprises, we found that 98% respondents* believe Apache Spark is a strong candidate for end-to-end enterprise data processing.

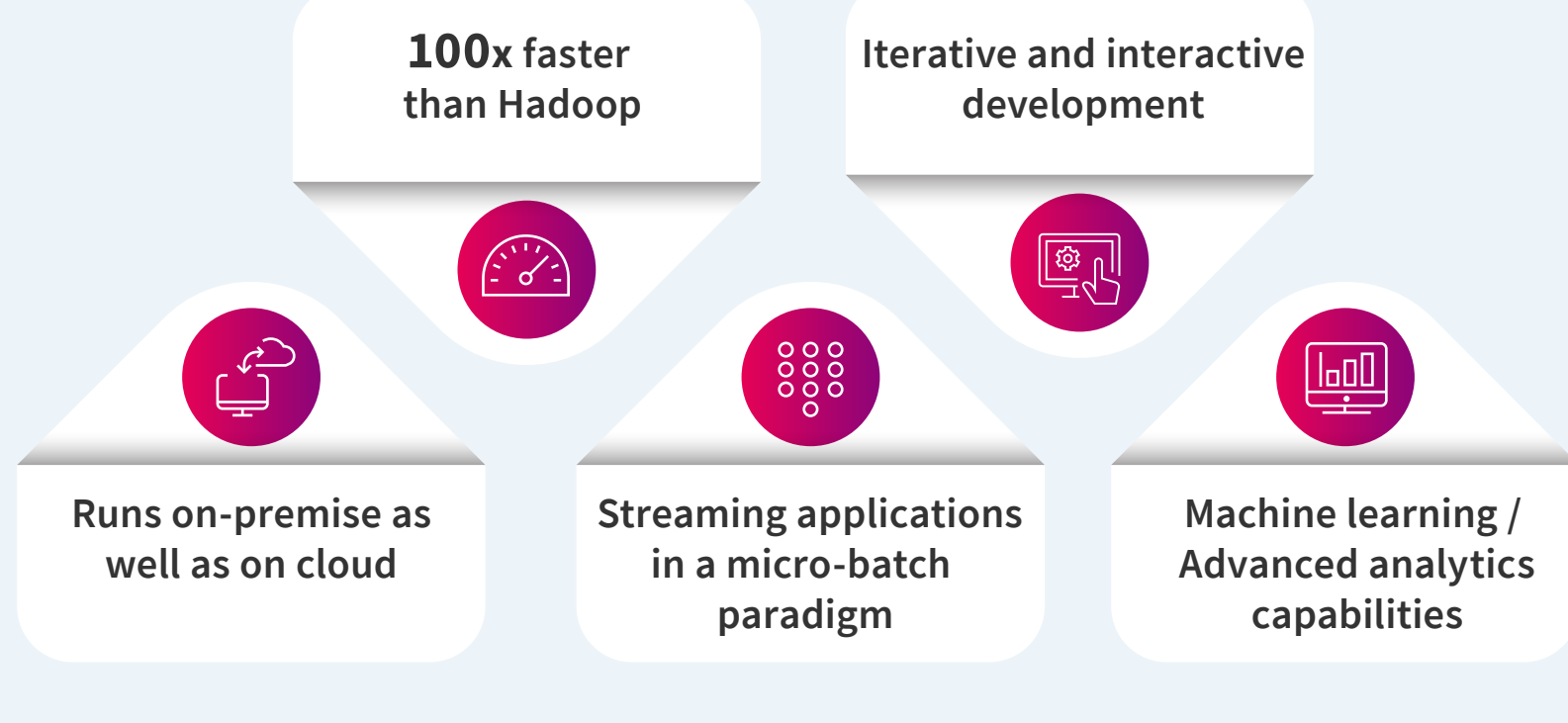
Following are the top use cases of Apache Spark in the enterprise today:



Let's dig deeper to find the reasons for the widespread adoption of Apache Spark.

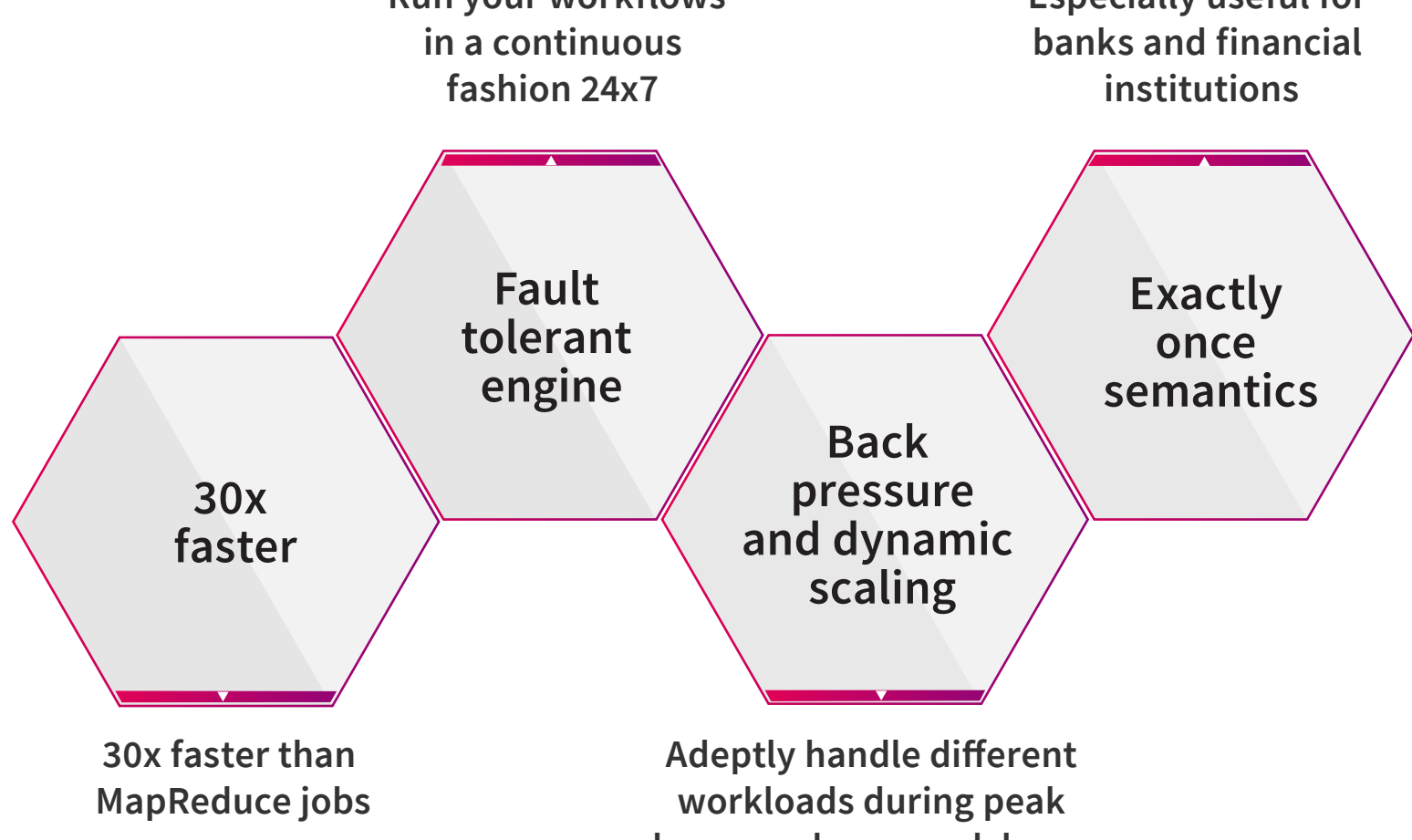
Reason 1: A Highly Capable Platform

Apache Spark provides a host of features and capabilities besides covering a wide range of workloads such as batch applications, iterative algorithms, interactive queries, and streaming. This makes Spark ideal for multiple data processing needs.



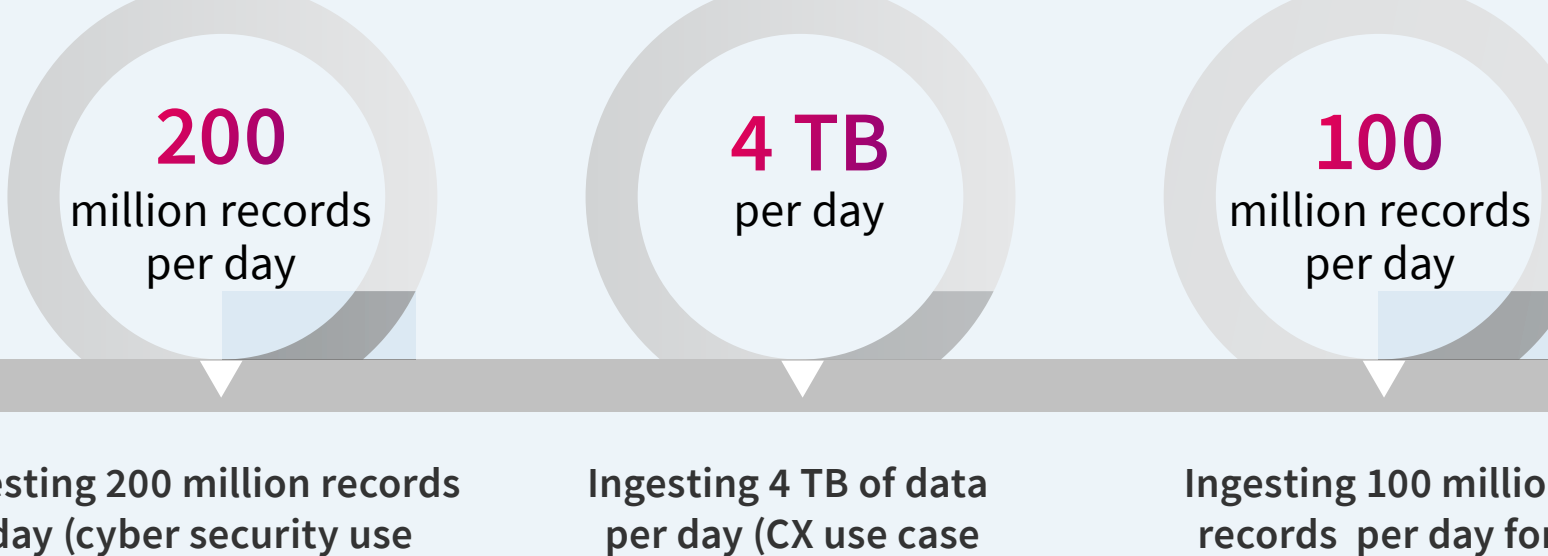
Reason 2: Apache Spark is Ready for the Enterprise

Enterprises across industry sectors, from hi-tech and telecommunications to financial services and others, are rapidly adopting and deploying Spark all the way up to production. Here are some compelling reasons.



Reason 3: Apache Spark is Tested in the Enterprise

Our Fortune 1000 customers have successfully tested and deployed Apache Spark for its versatility and strengths as a distributed computing framework that can completely handle all needs for data processing, analytics, and machine learning workloads.



Ingesting 200 million records / day (cyber security use case for a major US bank)

Ingesting 4 TB of data per day (CX use case for a US airline)

Ingesting 100 million records per day for a Tier 1 US telco

Reason 4: Visual IDE for Accelerated Development: Ingest, Cleanse, Blend, Transform, and Load – All on One UI

Enterprise grade tools like StreamAnalytix offer a visual integrated development environment (IDE) for Apache Spark, and serve all streaming and batch data processing and analytics needs. You can ingest, blend, and process high velocity big data streams as they arrive, run machine learning models, visualize results on real-time dashboards, and train and refresh models in real-time or in batch mode. Enterprise teams can build and operationalize Apache Spark applications 5 - 10x faster using drag-and-drop interface with more than 150 built-in Spark operators, and a full application lifecycle support.

A deep dive into StreamAnalytix – A Visual IDE for Apache Spark



Watch our solution architects enact a typical business scenario to walk you through the capabilities of Apache Spark and provide demos of real-life use cases to give you a hands-on experience.

[WATCH THE WEBINAR](#)

StreamAnalytix is an enterprise grade, visual, big data analytics platform for unified streaming and batch data processing based on best-of-breed open source technologies. It supports the end-to-end functionality of data ingestion, enrichment, machine learning, action triggers, and visualization. StreamAnalytix offers an intuitive drag-and-drop visual interface to build and operationalize big data applications five to ten times faster, across industries, data formats, and use cases.

Visit www.streamanalytix.com or write to us at inquiry@streamanalytix.com

© 2018 Impetus Technologies, Inc. All rights reserved. Product and company names mentioned herein may be trademarks of their respective companies.

STREAMANALYTIX