

Global Software Provider Boosts Spark Performance and Reduces Waste

A global provider of software products and services needed to optimize the resource consumption of their Apache Spark jobs, increase Spark performance, enhance cluster utilization, and significantly reduce costs.

Pepperdata Capacity Optimizer and Pepperdata Application Spotlight enabled the software vendor to efficiently manage their workloads, keep costs down, and troubleshoot small issues before they grew into major complications.

- The company specializes in enterprise software solutions that cater to a broad range of industries, including engineering, design, and architecture.
- To ensure customer satisfaction and success, the software company constantly analyzes data using Spark to measure product performance, customer subscriptions, sales, customer journeys, and more.
- Spark had become costly and was frequently running out of capacity, forcing them to add more resources to continuously drive its processes.

Visibility and Optimization Boost Performance and Reduce Costs

This software products and services company is a global leader in design and manufacturing software with primary markets in the engineering, architecture, construction, manufacturing, media, and entertainment industries.

“Pepperdata drove a lot of growth. It stretched our system in many ways. Without the spotlight and optimization components, we probably wouldn’t have been able to meet our needs.”

— Chief Data Architect, Data Platforms and Insights

It is crucial for the software company to keep its products and services performing to meet SLAs. Every piece of software they roll out must satisfy the needs of their customers. To achieve this, the software vendor leverages Apache Spark to constantly analyze large volumes of data. This enables them to derive the insights required to improve their existing products and create new solutions.

With Pepperdata, the company is able to optimize Spark performance. In addition, they managed to keep their Spark-related costs down, enabling them to run more data processes and positively impact their bottom line.

Prior to implementing Pepperdata, the software enterprise had to increase memory capacity and deploy more resources whenever they experienced Spark performance issues. They wanted to reduce costs by 50%. With Pepperdata, the company is well on its way to reaching that target.

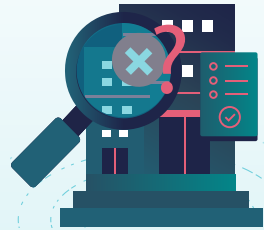
The Situation: Spark Complications

Spark allows the company to process and analyze large sets of big data and turn them into insights. While Spark processes large quantities of resources, it suffers from performance issues when unoptimized.

The company has been struggling to enhance the performance of its Spark applications because they don't have the in-house expertise to cope with the challenges of Spark tuning. As their Chief Data Architect put it:

"We didn't have an automated way to identify potential problems or make our systems more efficient."

They resorted to the [autoscaling](#) provided by their cloud vendor with the goal of improving performance and cutting costs. But they soon found that autoscaling did not work quickly enough. This forced the data team at the software firm to increase capacity and continually add more resources to ensure smooth performance.



The Impact: Costs Spiraling Out of Control

Developers at the software company had to resort to using the shotgun approach to tune Spark performance. However, nothing worked. As their Chief Data Architect recalls:

"Spark is notoriously hard to tune correctly and make efficient. People don't have time to go into every job. As a result, our entire platform just wasn't as efficient as it could have been."

As the company expanded, so did their workloads and Spark issues. The increased compute consumption was using up most of their budget. Each cluster was consuming two or three times the planned capacity.

In 2020, the software vendor saw its data processing needs increase 10x over the previous year. Had they used the same approach of doubling down their capacity and overprovisioning resources, the company would have been overwhelmed by costs, lags, and downtime.

Resolution: Visibility and Optimization in One Package

"We needed observability and insights."

Turning to Pepperdata, the software firm found a comprehensive solution that gives them visibility into their Spark framework, improves performance, and effectively reduces their compute consumption.

Pepperdata Capacity Optimizer allows the company to implement a very sophisticated approach to Spark performance tuning. The software automatically optimizes the resources in their clusters, resulting in the significant recapture of compute waste. With more resources available, the company can run more applications.

Pepperdata Application Spotlight gives the software firm a complete picture of their big data architecture and its processes, which is instrumental to all their Spark performance tuning efforts. With everything visible and accessible in one place, the company's development team can spot issues and determine if the issues are application-based or a symptom of cluster performance.

Pepperdata Platform Spotlight provides the software enterprise with tuning recommendations they can quickly apply to ensure all applications are running smoothly and efficiently—while eliminating applications and processes deemed wasteful by the system.

"We've had cases where we just couldn't figure out what was happening with a particular job or system. Platform Spotlight helped us dive into those things."

With Pepperdata, the company immediately saw results. The visibility and observability the solution delivers empower developers to immediately see and resolve issues to ensure everything is running efficiently, resources are fully optimized, and costs are significantly reduced.



FREE DEMO

Sign up for a free interactive demo of the Pepperdata solution to see how you can improve your Spark performance.

©2022 Pepperdata Inc. All rights reserved. Pepperdata and the Pepperdata logo are trademarks or registered trademarks of Pepperdata Inc. All other trademarks are the property of their respective owners. Pepperdata reserves the right to change this document without notice. To ensure you have the latest version of this document, visit www.pepperdata.com.



3945 Freedom Circle, Suite 920
Santa Clara, CA 9505

Start a Free Trial
www.pepperdata.com/trial

Send an Email
eval@pepperdata.com