pepperdata

Pepperdata for Amazon EKS

Immediately reduce the cost of running Apache Spark and Amazon EMR on Amazon EKS up to 47%—autonomously, continuously, and without application code changes.

Pepperdata for Amazon EKS

Kubernetes has become the de facto platform for agile enterprises managing applications in the cloud. Many businesses now run their Kubernetes workloads on Amazon Elastic Kubernetes Service (EKS) to gain the deployment and scaling benefits of a fully managed service, including streamlined container management and simplified integration with solutions from the Kubernetes open-source community.

With cost savings the top metric used to measure the ROI of a Kubernetes investment, organizations are turning to Pepperdata for Amazon EKS to maximize their Kubernetes investment.

Real-Time Capacity Optimization at Scale

Pepperdata Capacity Optimizer—a cost-optimization software solution battle-tested for over a decade in some of the world's largest and most complex computing environments—delivers immediate, automatic cost control for Apache Spark and Amazon EMR on EKS, without any need for manual application tuning. Capacity Optimizer augments and is complementary to all existing resource optimization efforts, including Karpenter.

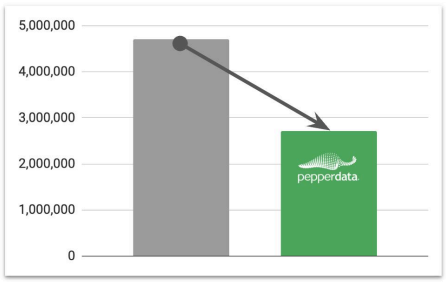
Capacity Optimizer continuously and autonomously reduces resource waste and cost of workloads running on Amazon EKS by 30-47%. By enabling the scheduler to launch more pods on the same nodes in real time—a process known as Continuous Intelligent Tuning—Capacity Optimizer increases node-level utilization without the need for manual tuning.

Because Capacity Optimizer is aware of actual hardware utilization levels, it immediately detects unused resources inside Spark pods and can provide that data to the native Kubernetes or Apache YuniKorn scheduler to enable pending tasks to be run on existing nodes. By safely launching more workloads on existing nodes, Capacity Optimizer enables the scheduler to make better use of node resources.

Dramatic Savings Results in Benchmarking and Customer Environments

According to a recent benchmarking study, Capacity Optimizer **reduced the cost of running a standard Spark workload at scale by 41.8 percent**, as measured by reduced instance hour consumption.

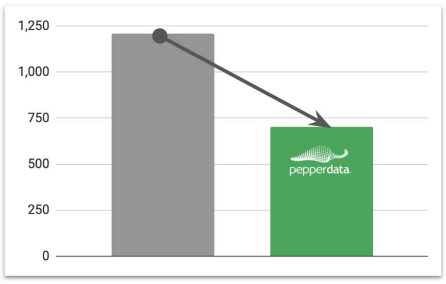
Customers have achieved similar results in their Amazon EKS production environments. An enterprise software company recently realized a **42.5 percent reduction** in normalized instance hours with Capacity Optimizer. This reduction translates directly into lower costs, in this case **over $60,000 per month in reduced instance hours**.



Normalized instance hours consumed decreased by 42.5%

Actual customer data from October 2023

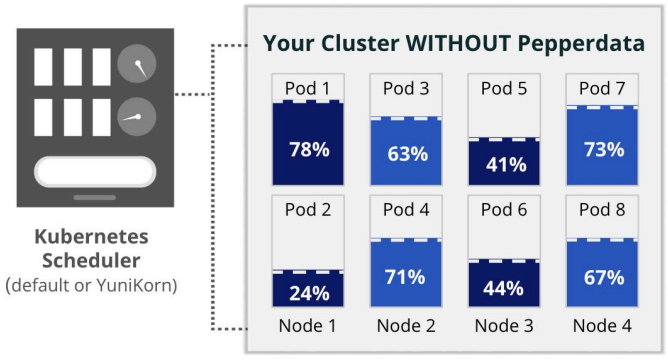
Figure 1: Capacity Optimizer reduced instance hours in Amazon EKS by 41.8% in a benchmarking study from October 2023.



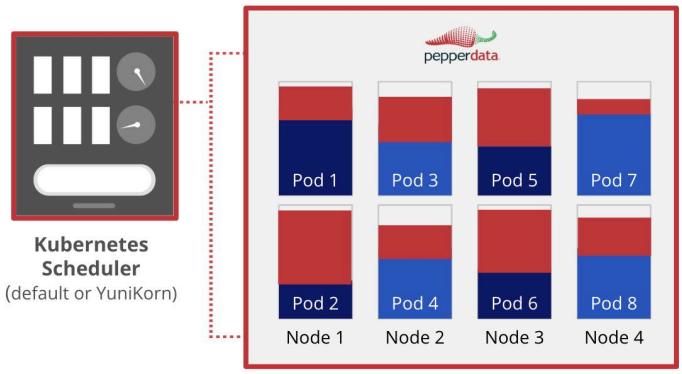
Total instance hours consumed decreased by 41.8%

Benchmarking data from October 2023

Figure 2: An enterprise software company achieved a 42.5% reduction in normalized instance hours on Amazon EKS, translating into a monthly savings of over $60,000.



Without Pepperdata, nodes may run at a fraction of their potential utilization. This is because the scheduler is only aware of requested resources and is not aware of what is actually being used, and therefore it cannot pack additional workloads onto nodes that appear to be fully utilized.



With Pepperdata, the blinders from the scheduler are removed, enabling it to pack nodes according to their actual utilization. The scheduler can make full use of existing nodes, increasing capacity and throughput and reducing cost without having to spin up new nodes.

Figure 3: Pepperdata enables nodes to run at the greatest capacity and efficiency by automatically eliminating waste and maximizing resource utilization.

Pepperdata Augments the Benefits of Karpenter and Prometheus

Karpenter is one of the fastest growing and most widely adopted Kubernetes add-ons for optimizing resource utilization and minimizing infrastructure costs. However, Karpenter cannot address the problem that Spark applications tend to waste the resources that are provisioned for them. Even with the most efficient, Karpenter-optimized infrastructure in the world, overprovisioned Spark applications will use infrastructure resources inefficiently. Capacity Optimizer uniquely addresses this inefficiency and is fully complementary to Karpenter.

Pepperdata makes use of Prometheus as a bundled component to provide key optimization metrics to Capacity Optimizer. Pepperdata developed a specialized configuration for Prometheus that **reduces the memory footprint by nearly tenfold**. Even in massive-scale Amazon EKS environments, Capacity Optimizer with Prometheus as a bundled component exhibits a **very small memory footprint**, supporting stability and overall reliability.

Try Pepperdata Today

Pepperdata pays for itself by immediately decreasing waste, increasing Spark utilization, and freeing developers from constant manual tuning to focus on revenue-impacting innovation. Pepperdata's 100 percent ROI guarantee means you don't pay if you don't save. Request a live demo or experience the savings directly in your Amazon EKS environment with a free trial.

Supported and Compatible Technologies

- Apache Spark on Amazon EMR
- Apache Spark on Amazon EKS
- Amazon EMR on Amazon EKS
- Apache Yunikorn
- Karpenter
- Prometheus

About Pepperdata

Pepperdata is the only cost optimization solution that delivers up to 47% greater cost savings—continuously and in real-time—on Amazon EMR and EKS with no application changes or manual tuning. Our customers include the largest, most complex, and highly-scaled clusters in the world, at top enterprises such as Citibank, Autodesk, Royal Bank of Canada, and those in the Fortune 5. For more information, visit pepperdata.com.

Pepperdata, Inc.
530 Lakeside Drive
Suite 170
Sunnyvale, CA 94085

Start a Free PoV
www.pepperdata.com

Send an Email
eval@pepperdata.com