Unlocking Cloud Cost Optimization: Best Practices for Amazon EMR

Optimize your cloud application deployment costs on AWS with Pepperdata

pepperdata.

aws PARTNER

Table of contents

The challenges of cloud cost overruns

Over the last few years, the advantages of migrating big data and other computationally intensive workloads to the cloud have become abundantly clear. Enterprises can enjoy near-infinite scalability on demand, agility in deploying new applications, and enhanced security and analytics, all combined with pay-as-you-go pricing. In fact, Gartner predicts that worldwide cloud spend will reach nearly $600 billion in 2023.

At the same time, almost all the factors that make the cloud so appealing for big and smaller workloads alike can lead to cost overruns if not carefully calibrated. It has been estimated that one-third of organizations will exceed their cloud budget by up to 40 percent, a widespread problem that plagues even the most sophisticated IT teams. Further, Flexera's 2023 State of the Cloud Report states that 82 percent of respondents across all organizations indicated that their top cloud challenge is managing cloud spend.

Clearly, cost optimization is key to a successful cloud deployment. Today the market has several approaches and multiple solution offerings, from which you can select the right partner for your technology and business needs.

In this eBook, we will share the different approaches to cloud cost optimization and how you can successfully achieve cost reduction in cloud application deployments on Amazon Web Services (AWS) with Pepperdata, an AWS Partner, and its unique approach.

> Gartner predicts that worldwide cloud spend will reach **nearly $600 billion in 2023**.

Common approaches to optimizing cloud costs

Applications running on cloud-based platforms for big data have the potential to scale quickly. However, rapid scaling, if not calibrated and managed well, can lead to a cost spike. Fortunately, there are several platform level approaches to managing and optimizing cost in the cloud including:

Visibility tools and cost observability dashboards

Observability dashboards refers to tools that allow you to view and understand the source of your cloud costs. Armed with this information, you can develop a list of action items to control and reduce costs.

Instance rightsizing

Instance rightsizing refers to matching instance types and sizes to workload performance and capacity requirements at the lowest possible cost. It's also the process of identifying opportunities to eliminate or downsize without compromising capacity or other requirements.

### Spot instances, reserved instances, and savings plans

Alternative pricing plans can help identify the lowest-cost cloud resources for a particular application. This can include spot instances, reserved instances, and other long-term savings plans. These pricing plans can be combined to achieve the optimal blend of workload performance and cost.

Manual tuning and configuration tweaks

Developers are often tapped to keep track of and adjust for cloud waste problems like overprovisioning, over-scaling, and inefficient resource distribution. Equipped with experience and expertise in their own applications, developers can spend time manually tuning configurations to improve efficiency.

However, getting the best cost configuration from a cloud deployment with any or all of these approaches can require multiple tools and resources to manage them. Most of these savings occur at the platform level, not the application level. What if you could optimize costs without the hassle of 24/7 manual involvement, and you could be sure that not only your platform but your individual applications were running as efficiently as possible?
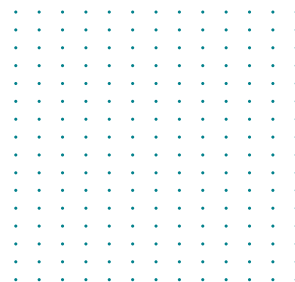
Get real-time cloud cost optimization with Pepperdata

Big data requires big compute. The benefits of running large data workloads on AWS—flexibility, resilience, scalability, sustainability, security, and a pay-as-you-go model—make it the world's most comprehensive and broadly adopted cloud. Two of the most popular cloud-based platforms for big data and other computationally intensive workloads are AWS solutions: Amazon EMR and Amazon Elastic Kubernetes Service (Amazon EKS).

But most IT teams are spread thin enough without having to dedicate resources to manually optimize their ever-changing applications. As a result, most deployments are overprovisioned to ensure applications don't fail due to lack of capacity. This is why Pepperdata exists. It is the only cloud cost optimization solution that delivers dynamic cost savings for your applications—automatically, continuously and in real time—on Amazon EMR and Amazon EKS with no application code changes or manual intervention required. Moreover, it can be installed via a bootstrap script in under an hour in most enterprise environments. Once Pepperdata goes to work, the cost savings are immediate, and ongoing, with no additional effort.

Pepperdata for Amazon EMR and Pepperdata for Amazon EKS provide the only solutions for continuous, automated cost optimization of cloud applications that require no application or code changes, empowering enterprises to extract the most from their cloud investment. The value of this investment can often be negated by runaway cloud deployment costs. On average, typical applications can be overprovisioned by 30 to 50 percent or sometimes more. Pepperdata's internal analysis of randomly selected customer clusters revealed that only 50 percent of allocated resources were used for 50 percent of the time. That's a lot of waste.

Amazon EMR

The industry-leading cloud big data solution for petabyte-scale data processing, interactive analytics, and machine learning using open-source frameworks such as Apache Spark, Apache Hive, and Presto.

Amazon EKS

A managed Kubernetes service to run Kubernetes in the AWS cloud and on-premises data centers.

Get real-time cloud cost optimization with Pepperdata

Pepperdata works by making the decisions of the cloud scheduler usage aware, so that scheduling decisions are based on actual usage instead of allocated usage. This ensures that your cluster uses available resources for each workload before adding new nodes or pods, thus reducing waste and allowing more workloads to run on the same node or pod. Pepperdata also dynamically tunes the cloud autoscaler to respond to your changing application workloads in real time. The result is dramatically improved efficiency, performance maximization, and application cost optimization of up to 47 percent with Amazon EMR and Amazon EKS.

Pepperdata is able to do all of this safely by ensuring that nodes always have some headroom for sudden spikes. On top of that, Pepperdata provides built-in back-off algorithms in place to reduce the optimization levels when the workload increases.

The results speak for themselves. Top enterprises and Fortune 5 organizations whose deployments include the largest, most complex, and highly-scaled clusters in the world have selected Pepperdata as their cloud cost optimization partner of choice. Let's look at what makes Pepperdata offerings unique.

OLD POD RESOURCE REQUEST LEVEL
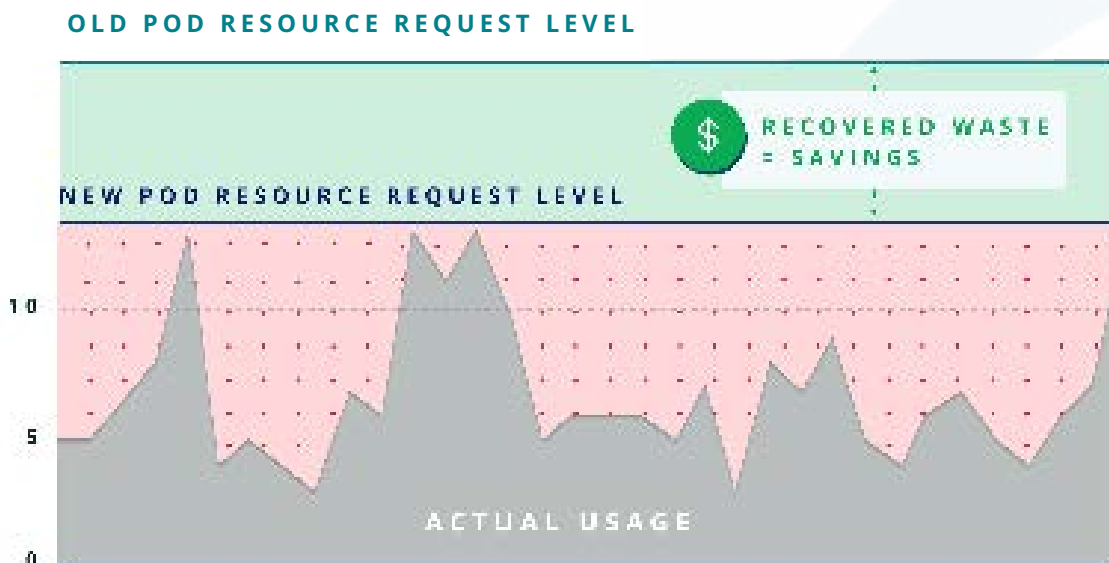
Pepperdata's secret sauce

As with any great recipe, the Pepperdata secret sauce has several key ingredients.

Immediate

As soon as you install Pepperdata, you'll start to see measurable cost savings in your dashboard within hours, not days or weeks.

Real time

Your applications and your cluster environment are dynamic, making it difficult, if not impossible, to predict peak traffic times. Other solutions must be manually trained on typical workload patterns and break if those patterns change. Real-time cluster awareness means that Pepperdata adapts to your ever-changing applications without requiring retraining or manual retuning.

Autonomous

Pepperdata requires no manual intervention and operates on your behalf to automatically eliminate waste as your applications run.

Continuous

Pepperdata works around the clock and never stops optimizing your workloads.

Safe

Pepperdata is safe to deploy because it never touches or makes changes to your running applications and is "node aware", maintaining your nodes and pods at their optimal "sweet spot" at all times.

Complementary

Pepperdata works alongside all other cloud cost savings strategies, most of which optimize at the platform level. Pepperdata operates at the application level to enable a new ongoing set of cost savings. After you've implemented other cloud cost remediation tools, enable Pepperdata. Pepperdata can still save your organization money.

Multi-workload

Pepperdata supports optimization of both batch and microservices workloads.

No application changes required

Pepperdata frees your development teams from applying recommendations or manual application tuning so they can focus on high-value, innovative activities.

Pepperdata in action:
customer savings successes

Global leader in design software cuts cloud deployment costs by 50% with Pepperdata

A global leader in design and manufacturing software was looking to halve its cloud deployment costs to comply with a corporate mandate. With a heavy deployment for constant analyses of large data volumes with Apache Spark on Amazon EMR and critical SLAs to meet, this was a major challenge for the organization.

Equipped with Pepperdata, this global design software leader increased the number of vCores on AWS from 3.5 million to nearly 6 million while maintaining the desired capacity utilization and runtime metrics, successfully reducing overall workload cost by over 50 percent. Pepperdata for Amazon EMR provided automated application optimization, full-stack observability, and continuous, real-time insights across Spark workloads and Amazon EMR instances all in one place. Implementing the Pepperdata solution delivered the desired 50% cost reduction and freed the company's developers to focus on production and innovation.

ISV saves up to 36% on core Amazon EMR clusters with Pepperdata

An Amazon Partner Network (APN) ISV that offers a purpose-built scalable and extensible security analytics platform to deliver next-gen threat detection was looking for cost optimization to meet a corporate cost reduction mandate. Pepperdata provided continuous, autonomous, real-time optimization of both CPU and memory usage, and the ISV now averages cost savings of 18 percent across all clusters, with savings of up to 36 percent on core production Amazon EMR clusters, for a total of $225K in monthly savings.

Consumer electronics brand saves millions annually with Pepperdata

Pepperdata worked with a giant consumer electronics brand to deliver improved node efficiency and improved demand-based scale up via Capacity Optimizer as the company migrated its massive 6,000-node Apache Hadoop-based on-premises data center to Amazon EMR. At the end of two weeks, once all the workloads had been migrated, the customer achieved a net 23.2 percent cost savings, translating into millions of dollars saved annually.

Try Pepperdata in your environment for free

Pepperdata works with a variety of customers across the Fortune 500, including those in the financial services, technology, retail, and healthcare industries. It also offers free Proofs of Value so you can experience the power of automated cost optimization in your own environment and help you consider migrating your big data workloads to Amazon EMR or Amazon EKS.

Your organization can benefit from the scalability and flexibility of AWS, coupled with the ultimate in cost optimization from Pepperdata. To learn more about Pepperdata's capabilities for Amazon EMR and Amazon EKS, download the Amazon EMR and EKS Cloud Migration and Modernization Assessment and/or try it for yourself for free via the AWS Marketplace.

FREE TRIAL

For more information, visit pepperdata.com or contact Dan Marx, Vice President of Sales, at dan@pepperdata.com to schedule a demo or arrange a no-cost savings assessment customized for your environment. Start enjoying the benefits of optimizing your cloud costs today.