aws startups

Boost your startup application while lowering costs

8 ways to reduce cloud costs and innovate more for virtually every application

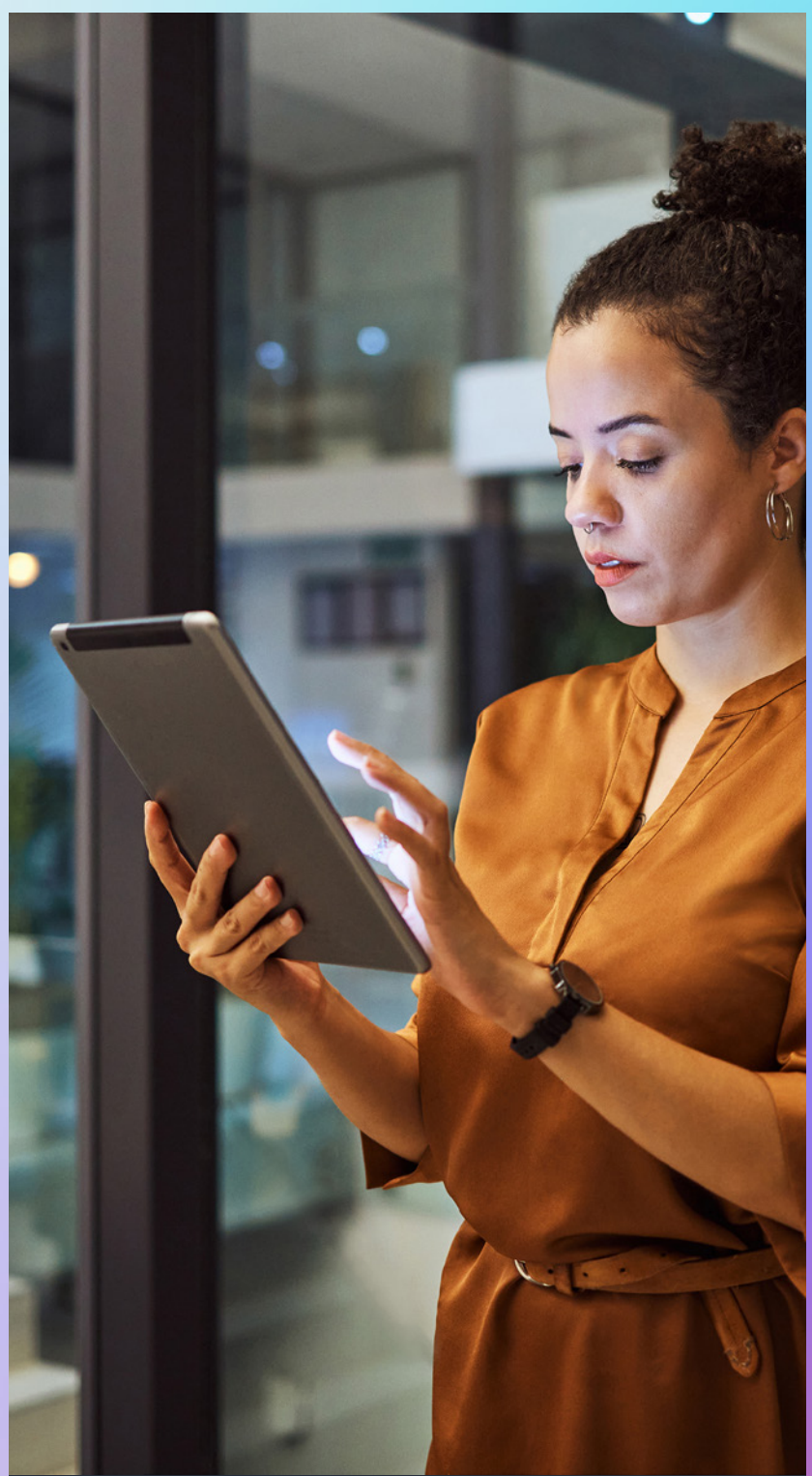


Table of contents

INTRODUCTION

Priority number one: Optimize costs and accelerate innovation

When starting a new venture, the investment that is needed to get off the ground can quickly become overwhelming. Founders are facing enormous pressure to do more with less. They are tasked with finding solutions that reduce costs, increase agility, accelerate innovation, and, ultimately, deliver transformational business value. In addition, security remains a top priority for startups at every stage, with increasing demands around data protection and privacy and a lack of cloud security professionals globally. Founders are discovering that a secure cloud infrastructure can support and even accelerate the speed and innovation needs of their businesses.

The cloud offers nearly unlimited opportunities to meet these goals—and beyond. From streamlining operations to delivering better user experiences securely, cloud migration enables startups to yield substantial cost savings while presenting new business opportunities. In a recent Gartner survey, infrastructure and operations leaders ranked cost savings among their top priorities.[1] Not surprisingly, business leaders and startup founders, in particular, are looking for ways to speed up their digital transformation so that they can improve operational efficiencies and invest in innovation. Amazon Web Services (AWS) offers support and services to enable startups to solve the big challenges, turn ideas into reality, and prove what's possible.

Wherever you are in your cloud journey, whether it's migrating your applications to AWS or building cloud-native applications to innovate faster and create new customer experiences, enabling startups to achieve more while lowering costs is central to how AWS builds cloud technology.

[1] "Gartner Leadership Vision for 2022: Top 3 Strategic Priorities for CIOs," Gartner, 2021

Get started with saving costs

Where does cost optimization start? Thinking about cloud migration as it applies to your startup's unique challenges, goals, and opportunities is a place to begin.

AWS enables you to take control of costs and continuously optimize your startup spend while building modern, scalable applications to meet your needs. The breadth of services and pricing options offers the flexibility to effectively manage your costs while maintaining the performance and capacity you require. AWS is dedicated to helping startups achieve the highest savings potential.

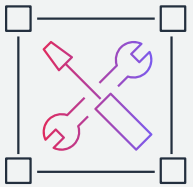
AWS can help you wherever you are in your cloud journey—with eight ways to optimize costs—so you can focus on innovating to gain competitive advantage.

The AWS Well-Architected Framework: A guide for best practices

We have created the **AWS Well-Architected Framework** to help you understand the pros and cons of the decisions you make while building systems on AWS. Using the AWS WA Framework helps you learn architectural best practices for designing and operating secure, reliable, efficient, cost-effective, and sustainable workloads in the AWS Cloud.

The **Cost Optimization pillar** of the AWS WA Framework includes the ability to run systems to deliver business value at the lowest price point. There are five design principles for cost optimization in the cloud:

- **Implement Cloud Financial Management:** To achieve financial success and accelerate business value realization in the cloud, you need to invest in Cloud Financial Management/cost optimization. Your startup needs to dedicate time and resources to build capability in this new domain of technology and usage management. Like your security or operational excellence capability, you need to build capability through knowledge building, programs, resources, and processes to become cost-efficient.
- **Adopt a consumption model:** Pay only for the computing resources that you require and increase or decrease usage depending on business requirements, not by using elaborate forecasting.
- **Measure overall efficiency:** Measure the business output of the workload and the costs associated with delivering it. Use this measure to know the gains you make from increasing output and reducing costs.
- **Stop spending money on undifferentiated heavy lifting:** AWS does the heavy lifting of data center operations, such as racking, stacking, and powering servers. AWS also removes the operational burden of managing operating systems and applications with managed services. This allows you to focus on your customers and business projects rather than on IT infrastructure.
- **Analyze and attribute expenditure:** The cloud makes it easier to accurately identify the usage and cost of systems, which then allows transparent attribution of IT costs to individual workload owners. This helps measure return on investment (ROI) and gives workload owners an opportunity to optimize their resources and reduce costs.



Access a potential cost savings of 75% (40 hours versus 168 hours) when you stop development and test environments when they are not in use. Typically, they are only used 8 hours a day during the week.

The AWS WA Framework is designed to help your team build secure, high-performing, resilient, and efficient infrastructure for a variety of applications and workloads. To help you apply best practices, we have created **AWS Well-Architected Labs**, which provide a repository of code and documentation to give you hands-on experience implementing best practices. We also have teamed up with select **AWS Partner Network Partners** that are members of the AWS Well-Architected Partner Program. These AWS Partners have deep AWS knowledge and can help you review and improve your workloads.

AWS: A cloud pioneer from the beginning

With more than a million active customers and a global cloud presence, AWS has experience in helping startups at every stage migrate workloads to the AWS Cloud.

AWS pioneered cloud computing in 2006, years before any other cloud provider. Since the beginning, we have helped customers across industries migrate their workloads to the cloud, offering more performance at lower costs than any other cloud provider.

Cost-efficiency and cost-effectiveness enable startups like yours to experiment at lower costs and accelerate innovation. Maximizing savings so startups can focus on innovation is at the core of what we do.

AWS relentlessly optimizes the efficiency of AWS services—since its inception, it has reduced prices

115x

This eBook guides you through eight ways you can reduce costs and make room for more innovation by optimizing your infrastructure investment.

METHOD #1

Migrate your applications to AWS

Familiarity and low friction are common reasons behind a startup's choice of a first development platform. Attractive entry pricing, functional advantage, and a founder's preference or brand bias are others. As products and teams grow, so does the cloud landscape. Infrequently but occasionally, we see startups that run IT operations in a multi-cloud environment. This is often the result of team turnover or acquisition activity. It can also be the choice of startup engineering teams that prefer to develop and run core applications on premises or in collocated environments—for security or other IP-related reasons.

AWS recommends choosing a primary cloud provider to get the best experience, performance, and cost at your cloud-native startup. The multi-cloud approach, while understandable in concept, typically generates undesirable growth in overhead—person-hours, expense, and technical debt—as a startup grows in size and operating history.

Startups often must use solutions from multiple providers to provision, manage, and govern IT resources to monitor the health of their applications and collect and analyze data stored in multiple locations. To help customers overcome these challenges, AWS has extended its services over the past several years to help create, manage, and govern infrastructure and applications hosted in hybrid and multi-cloud environments.

Migrate any workload—applications, websites, databases, storage, physical or virtual servers, and even entire data centers—from an on-premises environment, hosting facility, or other public cloud to AWS. At every step, leverage our years of experience to build your organizational, operational, and technical capabilities to start gaining business benefits from the beginning of the migration process.

Save on third-party licensing costs

AWS provides an option to save on third-party licensing costs and run your resources more efficiently: the **AWS Optimization and Licensing Assessment** (AWS OLA).

The AWS OLA is a free program for new and existing customers. It gives you a clear overview of your current costs and potential savings by assessing your current on-premises IT environment, third-party licenses, and projections of running these workloads on AWS.

The AWS OLA will provide you with a report that will model your deployment options using existing licensing entitlements. These results can help you explore available cost savings across our flexible licensing options. The AWS OLA can also be used in combination with the **AWS Migration Acceleration Program** (AWS MAP) **for Windows**, providing you with tools, support, and resources during your cloud migration.

Innovate faster, reduce costs, and enhance security

Take advantage of better performance at lower costs with AWS. AWS collaborates with major processor manufacturers to offer you the latest-generation processors, and for nearly a decade, AWS has invested in designing and producing silicon optimized for the cloud. Our custom silicon enables us to offer you industry-leading performance, enhanced security, and faster innovation—all at lower costs.

Designed virtualization platform: The AWS Nitro System

The **AWS Nitro System** is the underlying platform for our modern **Amazon Elastic Compute Cloud** (Amazon EC2) instances. It delivers the best price performance of any cloud provider, along with enhanced security and speed of innovation:

Best price performance: With the AWS Nitro System, Amazon EC2 instances can deliver more than 15 percent higher throughput performance on some workloads as compared to other major cloud providers running the same CPU. Dedicated Nitro Cards enable high-speed networking, high-speed **Amazon Elastic Block Store** (Amazon EBS), and I/O acceleration. Not having to hold back resources for management software means more savings passed on to end users. Additionally, the **Elastic Network Adapter** (ENA) **Express for Amazon EC2 instances**, introduced in 2022, delivers up to 25 Gbps of single-flow bandwidth—up from 5 Gbps—at no additional cost to customers. This latest Nitro-based innovation improves price performance by providing customers with more bandwidth without incurring additional costs.

Enhanced security: The Nitro Security Chip provides a simple hardware-based root of trust, enabling the most secure cloud platform with a minimized attack surface. Virtualization and security functions are offloaded to dedicated hardware and software. Additionally, a locked-down security model prohibits all administrative access, including those of Amazon employees, eliminating the possibility of human error and tampering.

Speed of innovation: With the AWS Nitro System, functions are modularized. It breaks the Amazon EC2 architecture into smaller blocks by offloading the virtualization functions onto dedicated hardware. These blocks can be assembled in many ways, which delivers the flexibility to design and rapidly deliver EC2 instances with an ever-broadening number of compute, storage, memory, and networking options. The results are lower costs and the ability to bring more workloads to the cloud, enabling you to increase your pace of innovation.

AWS firsts—innovation for your digital business

First to offer AMD, Arm, and macOS instances in the cloud
First with custom-built silicon for general purpose, HPC, and ML workloads in the cloud
First with the Nitro System
First to offer Intel's latest, third-generation, Intel Xeon Scalable processors
First to offer DDR5 memory in the cloud

METHOD #2

Choose the instance type that matches your application needs and budget

AWS instance types

AWS has more than 600 instance types, exceeding any other cloud provider. Each instance type provides a choice of processor, storage, networking, and operating system, so you can choose the instance configuration that best fits your startup-specific workload. And each instance type includes one or more instance sizes, allowing you to scale your resources to the requirements of your target workload.

AWS gives you the flexibility to change your instance type as quickly as your needs change, eliminating overhead costs for unused resources.

AWS has more than

600

instance types

Amazon EC2 instances fall into 6 categories:

1

General purpose

Our most popular instances provide a balance of CPU, memory, and network resources and are ideal for running web servers, containerized microservices, caching fleets, and development environments. One of the main distinctions within this class is between instances with fixed performance (**Amazon EC2 M7g**, for example) versus burstable (**Amazon EC2 T4g**, for example) performance.

2

Compute optimized

Good for compute-intensive, CPU-bound, demanding applications, such as frontend fleets for high-traffic websites, on-demand batch processing, distributed analytics, video encoding, dedicated gaming servers, and high-performance science and engineering applications. These instances offer the highest ratio of virtual CPUs to memory than the other families and the lowest cost-per-virtual CPU of all the EC2 instance types.

3

Memory optimized

These instances are ideal for memory-intensive applications, such as real-time big data analytics, in-memory databases, enterprise-class applications that require significant memory resources, or general analytics, such as Hadoop or Spark.

4

Accelerated computing

Instances in this category include additional accelerators and GPUs, FPGAs, and machine learning (ML) chips that provide massive amounts of parallel processing for tasks such as graphics processing, **ML training**, **ML inference**, and **high performance computing** (HPC).

5

Storage optimized

Ideal for tasks requiring local access to very large amounts of storage, extreme storage performance, or both. Instances are available that include both large-capacity HDD and extreme low-latency local NVMe SSDs. Choose from the industry's broadest portfolio of storage solutions optimized for your block, file, and object data.

6

HPC optimized

HPC instances are purpose-built to offer the best price performance for running HPC workloads at scale on AWS. HPC instances are ideal for applications that benefit from high-performance processors, such as large, complex simulations and deep learning workloads.

METHOD #3

Optimize costs for Windows applications

Are you looking to boost the reliability and availability of your Windows Server workloads? Or maybe you've reached the end-of-service (EOS) cutoff, where your SQL Server workloads will lose their patches and security updates.

You're not alone. Customers across multiple industries have been running Windows workloads on AWS since 2009, long before it was possible with other cloud providers. AWS supports everything you expect to build and run on Windows, including Active Directory, .NET, SQL Server, Windows desktop as a service (DaaS), and supported versions of Windows Server.

With AWS as the foundation for your Windows environment, there's no limit to the range of business benefits you can achieve, including:

Lower total cost of ownership (TCO): You can reduce your five-year cost of operations by 56 percent and gain 37 percent lower infrastructure costs, delivering up to 442 percent ROI over three years.

Licensing options: AWS has helped many customers break free from restrictive licensing scenarios. By purchasing license-included instances of Amazon EC2 or **Amazon Relational Database Service** (Amazon RDS), you get new, fully compliant SQL Server licenses from AWS.

License optimization: An AWS OLA can help reduce the number of Windows Server cores requiring a license by 77 percent and the number of SQL Server cores requiring a license by 45 percent when modeling workloads from on premises to AWS.

Better performance: Achieve two times better performance when you run your SQL Server workloads on an Amazon EC2 R5b.8xlarge instance.

Higher reliability: By moving your Windows workloads to AWS, you can realize a 98 percent reduction in unplanned downtime.

Can you bring your Microsoft licenses to AWS?

The answer is a resounding YES! There are three ways to do this:

Option 1: Bring your own license (BYOL) to AWS

You can often deploy eligible Microsoft applications on AWS by bringing your own licenses. AWS is a Microsoft Authorized License Mobility Partner, so if you have licenses with License Mobility—such as SQL Server with active Software Assurance—you can bring them to Amazon EC2 with default (shared) tenancy. Choose BYOL if you want to:

- Take advantage of cloud efficiencies and leverage your existing licensing investments
- Extend the lifecycle of your software without additional hardware costs
- Expedite your migration to the cloud by using existing virtual machine images

Option 2: Get your Microsoft licenses from AWS

You can use AWS instances that include the cost of licensing. By using AWS-provided licenses on Amazon EC2 or Amazon RDS instances, your Windows Server, SQL Server, Office, and Visual Studio licenses will always be fully compliant. You have the option to purchase **Amazon Machine Images** (AMIs) from AWS or through the **AWS Marketplace**. Choose AWS-provided licenses if you want to:

- Pay as you go—with no upfront costs or long-term commitments
- Fully control when you are being billed for license-included instances
- Get out of the business of managing licensing compliance and leave it up to AWS—we will handle it!

Option 3: Replace your Microsoft-licensed products and applications

You can speed up innovation with specialized built-for-the-cloud alternatives by replacing your Microsoft-licensed products and applications. Your modernization journey could include moving to Linux, porting applications from .NET Framework to .NET, decomposing monoliths into microservices, implementing DevOps techniques with container and serverless technologies, and transitioning your data tier to **Amazon Aurora** and other purpose-built databases. Choose this option if you want to:

- Move off legacy licensing for ultimate freedom and savings
- Break free from the punitive terms and high costs of Microsoft licenses

METHOD #4

Migrate to AWS Graviton for the best price performance for a broad set of applications

AWS designed **AWS Graviton processors** to deliver the best performance at the lowest cost for your cloud workloads running on Amazon EC2. AWS Graviton–based instances deliver up to 40 percent better performance at lower costs versus comparable x86-based EC2 instances. They are also highly energy efficient, using up to 60 percent less energy for the same performance as comparable Amazon EC2 instances.

AWS Graviton is available in over 25 AWS Regions, and migrating to AWS Graviton can help you increase performance, reduce costs, lower latency, and achieve better scalability.

The **AWS Graviton Fast Start** program helps you quickly and easily move your workloads to AWS Graviton in less than four hours. Or, accelerate your adoption of AWS Graviton with the help of **AWS Graviton Partners**.

AWS Graviton–powered managed services

AWS Graviton–based instances are also available in more than 25 popular managed AWS services. These services deliver the price performance benefits of AWS Graviton processors while providing a fully managed experience. AWS Managed Services using AWS Graviton include serverless solutions, such as **AWS Lambda** and **AWS Fargate**, and AWS Graviton–based databases, such as Amazon Aurora, Amazon RDS, and **Amazon ElastiCache**.

METHOD #5

Lower costs with AWS machine learning accelerators and services

AWS designed **AWS Inferentia** accelerators to deliver high performance at the lowest cost in Amazon EC2 for ML inference applications. The first-generation AWS Inferentia accelerator powers Amazon EC2 Inf1 instances, which deliver up to 2.3 times higher throughput and up to 70 percent lower cost per inference than comparable GPU-based instances. AWS Inferentia2 accelerator is purpose-built to deploy ML models, such as large language models (LLMs) and diffusion models, with more than 100 billion parameters. Inferentia2-based Amazon EC2 Inf2 instances deliver up to four times higher throughput and up to 10 times lower latency compared to Inf1 instances. They also deliver up to 70 percent better price performance than comparable GPU-based instances in Amazon EC2.

AWS Trainium is an ML accelerator that AWS purpose-built for high-performance, low-cost deep learning training. Amazon EC2 Trn1 instances powered by Trainium chips offer up to 50 percent cost-to-train savings over comparable GPU-based EC2 instances. Trn1 instances are the first EC2 instances with up to 800 Gbps of Elastic Fabric Adapter (EFA) network bandwidth. They are deployed in EC2 UltraClusters that enable scaling up to 30,000 Trainium accelerators, which are interconnected with a non-blocking petabit-scale network to provide up to 6.3 exaflops of compute.

Amazon SageMaker is a fully managed service that offers high-performance, low-cost ML at scale. SageMaker enables startups of all sizes to build, train, and deploy ML models with a choice of tools and workflows—integrated development environments (IDEs) for data scientists and engineers, no-code interfaces for business analysts, and support for the leading ML frameworks, toolkits, and programming languages for MLOps engineers.

To make it easier for your startup to get started, **Amazon SageMaker JumpStart** is an ML hub that provides built-in algorithms with pretrained models from model hubs, pretrained foundation models to help you perform tasks such as article summarization and image generation, and pre-built solutions to solve common use cases. In addition, you can share ML artifacts, including ML models and notebooks, within your organization to accelerate ML model building and deployment.

Visit the **AWS Machine Learning Infrastructure** webpage to learn more about how you can choose your ML infrastructure, tools, and frameworks to accelerate ML innovation.

Tens of thousands of customers are using SageMaker to generate more than

1 Trillion

predictions per month

METHOD #6

Select the compute purchase model that best fits your budget

On-Demand Instances

On-Demand Instances let you pay for compute capacity by the hour or second, depending on which instances you run. No long-term commitments or upfront payments are needed. On-Demand Instances are ideal for applications with short-term, spiky, or unpredictable workloads that cannot be interrupted and applications being developed or tested on Amazon EC2 for the first time.

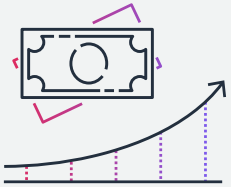
Savings Plans

Take advantage of Savings Plans with flexible pricing models to achieve cost-efficiency. AWS offers three types of plans: Compute Savings Plans, EC2 Instance Savings Plans, and SageMaker Savings Plans. Compute Savings Plans apply to usage across Amazon EC2, Lambda, and Fargate. The EC2 Instance Savings Plans apply to EC2 usage, and the SageMaker Savings Plans apply to SageMaker usage. Savings Plans automatically and simultaneously apply to eligible AWS usage. They enable you to innovate faster by leveraging the newest instances, families, generations, and regions while staying on the same plan. Since the launch of Savings Plans in 2019, customers have saved more than $15 billion.

Amazon EC2 Spot Instances

EC2 Spot Instances let you take advantage of unused EC2 capacity in the AWS Cloud. Spot Instances are available at up to a 90 percent discount compared to On-Demand prices. You can use Spot Instances for various stateless, fault-tolerant, or flexible applications, such as big data, containerized workloads, continuous integration and continuous delivery (CI/CD), web servers, HPC, and test and development workloads.

You can easily combine Spot Instances with other purchase models, giving you the flexibility to grow and change over time while still getting the lowest cost available on AWS. Since 2015, Spot Instances have helped our customers save more than $8 billion.



Reduce your bill by up to 72% compared to On-Demand prices in exchange for a 1- or 3-year hourly spend commitment. Choose savings plans with flexible pricing models.

The **Amazon CloudFront security savings bundle** is a flexible, self-service pricing plan that helps you save up to 30 percent on your CloudFront bill in exchange for a monthly spend commitment of a one-year term. These savings are not limited to data delivered by CloudFront but apply to all CloudFront usage types, including CloudFront Functions and Lambda@Edge. The CloudFront security savings bundle also includes free **AWS WAF** usage up to 10 percent of your committed amount.

TextNow saves 93% on data transfer fees using AWS PrivateLink

TextNow used **AWS PrivateLink** to connect its **Amazon Virtual Private Cloud** (Amazon VPC) with Datadog, a popular software-as-a-service (SaaS) application, while keeping all data on the AWS network. By eliminating the need to send data over the public internet, AWS PrivateLink reduced TextNow's data transfer fees by 93 percent and improved user data security.

Read the case study ›

15B+

Since 2019, customers have saved $15 billion with Savings Plans

8B+

Since 2015, customers have saved $8 billion with EC2 Spot Instances

METHOD #7

Optimize costs with a choice of storage services

You can minimize your TCO with AWS Storage services. These services eliminate on-premises capital equipment investment, management complexity, and infrastructure maintenance. With AWS Storage, you get the right mix of price and performance for your workloads while paying only for the storage that you use.

AWS gives you the broadest array of storage services of any cloud provider. This includes more ways to optimize storage costs through a choice of storage classes and intelligently tier data to lower cost storage and data reduction capabilities, such as compression and data deduplication.

For block-based workloads, **Amazon EBS** provides easy-to-use, high-performance block storage at any scale. You select the storage that best fits your workload, service level, and budget. Amazon EBS scales fast for your most demanding, high-performance workloads, including Microsoft, Oracle, and SAP products. **Amazon EBS Snapshots** provide a solution to maintain compliance and further reduce snapshot storage costs by up to 75 percent by using **Amazon Data Lifecycle Manager** to automatically move seldom-used snapshots to the **Amazon EBS Snapshots Archive**.

Amazon Simple Storage Service (Amazon S3) is the lowest-cost object storage in the cloud. No matter the size of your startup, Amazon S3 lets you store and protect any amount of data for virtually any use case, such as data lakes, cloud-native applications, and mobile apps. Amazon S3 storage classes provide the lowest-cost storage for specific access patterns. And **Amazon S3 Intelligent-Tiering** has saved customers more than $750 million in storage costs compared to Amazon S3 Standard since the launch of S3 Intelligent-Tiering in 2018. With cost-effective storage classes and easy-to-use management features, you can optimize costs, organize data, and configure fine-tuned access controls to meet specific business, organizational, and compliance requirements. AWS offers the industry's widest portfolio of fully managed file storage. We handle all infrastructure—provisioning, patching, and backups—allowing you to choose the right file system technology to meet your workload requirements.

AWS offers the industry's widest portfolio of fully managed file storage. We handle all infrastructure—provisioning, patching, and backups—allowing you to choose the right file system technology to meet your workload requirements.

Amazon Elastic File System (Amazon EFS) is a serverless elastic file system that scales automatically as files are added, removed, and burst to higher throughput levels when necessary. You can also reduce costs by up to 92 percent by automatically tiering infrequently accessed files. The Amazon FSx file system family includes **Amazon FSx for NetApp ONTAP**, **Amazon FSx for OpenZFS**, **Amazon FSx for Windows File Server**, and **Amazon FSx for Lustre**. Amazon FSx enables you to optimize your price and performance to support a broad spectrum of use cases, from small user shares to the most demanding compute-intensive workloads. Amazon FSx file systems support a rich set of storage efficiency features, including data deduplication, compression, and usage quotas.

Amazon S3 Glacier storage classes are purpose-built for data archiving, providing you with the highest performance, most retrieval flexibility, and lowest-cost archive storage in the cloud.

75%

EBS Snapshots Archive reduces snapshots storage costs by up to 75%

92%

Amazon EFS reduces storage costs by 92%

$750M

Since its launch in 2018, S3 Intelligent-Tiering has saved customers more than $750 million in storage costs compared to S3 Standard

50%+

Amazon FSx family data reduction technology reduces storage costs by up to 65%

METHOD #8

Optimize your compute resources with intelligent AWS tools

AWS Compute Optimizer and AWS Auto Scaling allow you to provision with precision. These services help lower your costs by responding in real time to changes in demand—and they're free to use:

Compute Optimizer

Over-provisioning resources can lead to unnecessary infrastructure costs, while under-provisioning can lead to poor application performance. Using ML to analyze historical utilization metrics, Compute Optimizer recommends the optimal AWS resources for your workloads, further reducing your infrastructure costs. With just a few clicks, it automatically generates recommendations based on current utilization data, eliminating the need to invest time and resources to set up rule-based thresholds. Since its launch in December 2019, Compute Optimizer has generated recommendations for more than 80 percent of Amazon EC2 usage and provided more than 10 billion recommendations, resulting in reduced costs and improved performance for a variety of workloads.

AWS Auto Scaling

AWS Auto Scaling monitors your applications and automatically adjusts capacity to maintain steady, predictable performance at the lowest possible cost. You can easily set up application scaling for multiple resources across multiple services with a single intuitive interface and maintain optimal application performance and availability even when workloads are periodic, unpredictable, or continuously changing. When demand spikes, AWS Auto Scaling automatically increases the capacity of constrained resources, so you can maintain a high quality of service. When demand drops, it removes any excess resource capacity to help keep you from overspending. There's no need to add AWS Auto Scaling as a separate tool—it is already built into AWS solutions.

Compute Optimizer has provided more than

10 billion

recommendations since launch, resulting in reduced costs and improved performance for a variety of workloads.

Customer stories

CUSTOMER STORIES

Ula builds a cost-effective, scalable purchasing app for small businesses

About Ula

Ula is a B2B ecommerce marketplace app that serves Indonesia's microenterprises. Customers using Ula can order the products they need on their mobile phones and expect delivery within two days.

The challenge

Ninety-eight percent of Indonesia's economy comprises microenterprises, which continuously suffer from inventory management, cash flow issues, inconsistent deliveries, or lack of delivery options. Ula set out to build a cost-effective, scalable purchasing app and technology stack that could overcome the challenges of basic mobile devices and poor network connections to serve small businesses.

The AWS solution

Leveraging the **AWS Well-Architected Tool**, Ula accesses high availability and reliability from the get-go. With **Amazon Elastic Container Service** (Amazon ECS), the startup can automate orchestration at any scale to enable business expansion. It achieves cost-efficiency by leveraging **EC2 Spot Instances** in its development, staging, and production environments.

The results

The app is at least one-third lighter than its competitors', making it easier for MSMEs to download and use in any setting. High availability ensures at least 99 percent uptime, while scalable capabilities support a 300 times increase in business volume. Leveraging AWS services, the startup saves 70 percent on monthly infrastructure costs.

Read the full story ›

aws startups

ula

CUSTOMER STORIES

Fork Media achieves 40% lower costs on AWS databases

About Fork Media

Fork Media is an India-based media tech company that offers “brand safety.” It deploys artificial intelligence (AI) to understand the overall sentiment of a webpage before placing an ad for its customers.

The challenge

Previously using two different cloud computing platforms, the company was challenged with latency issues in servicing its media assets. Ranging from 500 milliseconds to one second, this came in higher than its internal service level agreement (SLA) of 200 milliseconds. The engineering team also faced challenges managing high query volumes for its mission-critical database.

The AWS solution

To maximize ad visibility, Fork Media now consolidates its ad server databases on AWS. It deploys **Amazon RDS** and **Amazon EBS** to automate database administration at scale. **CloudFront** is utilized as a low-latency content delivery network (CDN), and it has adopted **Amazon EC2 Auto Scaling** with **Spot Instances** to optimize costs and dynamically add capacity when traffic spikes.

The results

It took Fork Media under three months to migrate two databases, including developing the proof of concept and testing. With the integration of cloud-native solutions on AWS, the latency between endpoints for serving any static advertising asset has dropped from 500 milliseconds to only 50 milliseconds. In addition, Fork Media’s CPU consumption has decreased by 45 percent since migrating to AWS Graviton2 processors.

Read the full story ›

aws startups

CUSTOMER STORIES

Inmagine migrates to AWS to enhance operational efficiency

About Inmagine

Inmagine is a global creative ecosystem specializing in creative content, supplying designers with ready-to-use stock content.

The challenge

Having experienced unpredictable web traffic patterns, unexpected distributed-denial-of-service (DDoS) attacks, site reliability issues, and challenges with managing storage capacity, Inmagine was seeking a new IT partner. The challenges it faced also escalated operational costs and resulted in an overinvestment in compute capacity.

The AWS solution

To solve for these challenges, Inmagine migrated to AWS. To reduce operational uptime, the startup leverages **Application Load Balancers** and **Amazon EC2 Auto Recovery**. **Amazon S3** allows the startup to offload storage management and access the capacity, availability, and reliability of AWS infrastructure. To protect dynamic content traffic, **CloudFront** and **AWS WAF** solutions are deployed.

The results

The migration to AWS has resulted in a reduction of application performance degradation by 34 percent. Instead of spending value resources on compute capacity, funding has been reallocated to optimize and innovate new features for customers. With a 25 percent improvement in loading times for dynamic content pages, Inmagine has seen an overall increase in its site's performance.

Read the full story ›

aws startups

INMAGINE

CUSTOMER STORIES

Ramp scales its fast-growing finance automation platform

About Ramp

Ramp, a B2B fintech startup, is a tech-first finance automation platform. Its serverless modern application, in conjunction with its corporate card, allows businesses to manage their finances more efficiently.

The challenge

To succeed, the startup needed a scalable modern architecture, high developer productivity, multi-region availability, and optimized cloud costs. A bare-bones approach to cloud computing was not going to work. With rising demand, it required sub-millisecond latency for Ramp's caching needs.

The AWS solution

Ramp turned to AWS to build its platform's core infrastructure utilizing **Aurora** database clusters and **ElastiCache for Redis** to accelerate application and database performance. To improve developer productivity, web servers run on **Amazon ECS** and **Fargate**. The startup uses **Amazon Aurora Global Database** across regions with Aurora, which enables support across regions to meet multi-region requirements. AWS Graviton processors are leveraged for Ramp's databases to performance relative to cost.

The results

AWS solutions provide the flexibility to meet demand and add components to increase system robustness. Developer velocity across the organization has significantly increased with AWS. Cost-optimization tools have allowed Ramp to decrease its cloud spend. **AWS Activate** is integrated into the go-to-market strategy and has been instrumental from a product perspective.

Read the full story ›

aws startups

ramp

CUSTOMER STORIES

DataFarming processes data from 35 million acres

About DataFarming

DataFarming builds simple, low-cost precision agriculture solutions that allow farmers to access data and gain insights. These can be incorporated into everyday farming activities.

The challenge

DataFarming was seeking to deliver high-quality data from satellite imagery faster and more effectively to help farmers optimize their crop growth. Farmers would need to view their data on mobile devices out in the field, where very limited bandwidth is common.

The AWS solution

DataFarming leveraged **Amazon ECS** to make it simple to deploy, manage, and scale containerized applications. The company stores most of its image-based data in **Amazon S3**, which offers industry-leading scalability, data availability, security, and performance. It leverages **Spot Instances** to scale cost-effectively.

The results

Using AWS services, DataFarming built a solution that includes 28,000 farms and processes data for 35 million acres of farmland. In just four years, the organization has achieved a 900 percent increase in the use of satellite imagery in the Australian grains market. It successfully achieved 70 percent savings using Spot Instances compared to On-Demand pricing.

Read the full story ›

DATA FARMING

CUSTOMER STORIES

Mayden increases reach and reliability of healthcare software

About Mayden

Mayden is a UK healthcare technology company that provides technology for mental healthcare services as part of the National Health Service (NHS).

The challenge

Mayden was seeking more stability from a new cloud provider. It required expertise and support to migrate efficiently without disrupting service for patients and care providers.

The AWS solution

Mayden migrated using **AWS Application Migration Service**, which minimizes time-intensive, error-prone manual processes by automatically converting customer source servers to run natively on AWS. **Cloud Migration Factory on AWS** was leveraged to coordinate and automate the large-scale migration. By deploying AWS Managed Services, including **Amazon Route 53** and **Amazon Elastic Load Balancing** (ELB), its team is now free to focus on building and supporting its applications.

The results

Mayden migrated 300 servers to AWS in just six weeks with minimal downtime during cross-cloud replication of servers. Since migrating to AWS, Mayden has not experienced any major service issues. Partnering with AWS, the IT team can perform faster, building the 75 database servers that make up a key part of Mayden's infrastructure in just two minutes and making it 98 percent faster to build database servers on AWS.

Read the full story ›

Mayden

CONCLUSION

Start maximizing your savings now

With innovations in silicon and serverless technologies and a variety of flexible pricing options, AWS is committed to helping your startup get the most from its cloud infrastructure spend. We provide you with the capability to optimize your costs while building modern, scalable applications to meet your startup needs. The breadth of services and pricing models demonstrates our commitment to giving you all the compute performance and capacity you require at the lowest cost—now and as your startup grows.

Start improving your application performance today and reach your highest savings potential with AWS.

Discover all the ways AWS can help you focus on innovation while lowering your costs.

Explore our startup cost-optimization hub ›

Gain access to free tools, resources, and expert support.

Become a member of AWS Activate ›