

Running Big Data Analytics on AWS: Benefits and Use Cases

The AWS Advantage in Big Data Analytics

As we become a more digital society, the amount of data being created and collected is accelerating significantly. Analysis of this ever-growing data becomes a challenge with traditional analytical tools. Big data tools and technologies offer opportunities and challenges in being able to analyze data efficiently to better understand customer preferences, gain a competitive advantage in the marketplace, and grow your business.

Amazon Web Services (AWS) provides a broad platform of managed services to help you build, secure, and seamlessly scale end-to-end big data applications quickly and with ease. Whether your applications require real-time streaming or batch data processing, AWS provides the infrastructure and tools to tackle your next big data project.

The significant compute capacity that big data workloads require is ideally suited to the pay-as-you-go cloud computing model, where applications can easily scale up and down based on demand. As requirements change, you can easily resize your environment on AWS in a matter of minutes to meet your needs, without having to wait for additional hardware or being required to over invest to provision enough capacity.

These capabilities of the AWS platform make it an ideal fit for solving big data problems, and many customers have implemented successful big data analytics workloads on AWS. Let's explore a few use cases for big data analytics on AWS.

The Key Value Drivers of Big Data Solutions on AWS



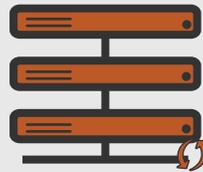
OPERATIONAL COSTS

Lower operational costs with the ability to match supply and demand and create operational transparency.



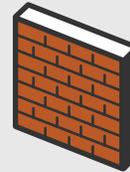
WORKFORCE PRODUCTIVITY

AWS services and automation capabilities help drive operational maintenance efficiency and increase developer productivity.



COST AVOIDANCE

Eliminate maintenance and hardware refresh programs.



OPERATIONAL RESILIENCE

Improve revenue and margins by reducing the cost of risk mitigation.



BUSINESS AGILITY

Reduce your time-to-market, increase operational agility, and drive innovation.



Business Cases for Big Data Analytics in AWS

QUERIES AGAINST AN AMAZON S3 DATA LAKE

Data lakes are an increasingly popular way to store and analyze both structured and unstructured data. Using Amazon S3, AWS Glue, Amazon Athena, and Amazon Redshift Spectrum, your data is immediately available for analytics without being moved and you can directly query your data lake.

A BI tool like Amazon QuickSight enables you to easily build visualizations, perform ad-hoc analysis, and quickly get business insights from your data. You can also use the AWS Glue Catalog as your external Apache Hive Metastore for big data applications running on Amazon EMR.

CAPTURING AND ANALYZING SENSOR DATA

An international air conditioner manufacturer sells large air conditioners to various commercial and industrial companies. They also offer add-on services where you can see real-time dashboards in a mobile app or web browser. Each air conditioner unit sends its sensor information for processing and analysis and this data is used by the manufacturer and its customers. This capability enables the manufacturer to visualize the dataset and spot trends.

If the initial release is successful, the manufacturer would like to expand this offering to their consumer line as well. The Big Data solution needs to be able to handle massive amounts of data without interruption. The solution also needs to have scalability so it can keep up with their business as it grows.

AWS can solve this big data problem by breaking it up into two work streams: 1) A/C unit's current information with near real-time requirements and a large number of customers consuming this information. 2) All historical information on the A/C units to run trending and analytics for internal use.

A near real-time reporting data flow designed for low latency and quick scalability can be achieved using Amazon Kinesis Data Streams, DyanmoDB, and a custom web application run on Amazon EC2 for reporting. The second work stream doesn't have such stringent speed and latency requirements, allowing a different solution stack that can hold and analyze larger amounts of data at a smaller cost per byte. Using an Amazon Kinesis-enabled application on a smaller EC2 instance, Amazon S3, and Amazon Redshift, this architecture can start small and grow as needed.

SENTIMENT ANALYSIS OF SOCIAL MEDIA

A large toy maker wants to use social media to monitor how consumers responding to their products after each new toy release. They want to use social media insights to: 1) understand how consumers are using their products, 2) ensure customer satisfaction, and 3) plan future roadmaps.

The company can analyze and classify social media data in a cost-effective and programmatic way using Amazon EC2, Amazon Kinesis, Amazon S3, Amazon ML, and Amazon SNS. Amazon ML is used to make predictions on the input data, such as customer sentiment analysis about a product. Lambda can also be used to take actions based on predictions from Amazon ML, such as routing a social media post to the customer service team for further review.

By using the combination of Amazon Kinesis Data Streams, Lambda, Amazon ML, and Amazon SNE, the company has created a scalable and easily customizable social listening platform.



The AWS and Pinnacle Solutions Advantage

As data continues to be generated and collected at an exponential rate, data analysis requires scalable, flexible, and high performing tools to provide insights in a timely fashion. However, organizations are facing a growing big data ecosystem with a plethora of new tools, many of which “die” very quickly. This makes it difficult to keep pace and choose the right tools.

This information brief serves as a first step to helping you solve this challenge. With a broad set of managed services to collect, process, and analyze big data, the AWS platform makes it easier to build, deploy, and scale big data applications. Using the AWS platform allows you to focus on business problems instead of updating and managing tools.

AWS provides many solutions to address your big data analytic requirements. Most big data architecture solutions use multiple AWS tools to build a complete solution. This approach helps meet stringent business requirements in the most cost-optimized, performant, and resilient way possible. The result is a flexible, big data architecture that is able to scale along with your business.

Pinnacle Solutions adds our administration and maintenance expertise to the computing infrastructure of AWS with the Pinnacle Shelter™ offering. With Pinnacle Shelter™, our team of certified AWS experts will work with you to execute a smooth, seamless migration to the cloud. There are so many ways your cloud migration journey could go, depending on your goals, budget, culture, and vision. Our consultants will evaluate your business and technical needs and map out your optimum migration strategy for each of your legacy systems - whether to rehost, replatform, repurchase, refactor, retain, or retire. Pinnacle Shelter™ also includes ongoing support and administration services, allowing you to focus on business operations instead of infrastructure.



About Pinnacle Solutions

Pinnacle Solutions is your resource for all your business intelligence and analytic needs. With Pinnacle Shelter™, you can combine the computing infrastructure of Amazon Web Services (AWS) with our administration and maintenance expertise. The result is a service that allows your company to eliminate costly hardware and the administrative challenges that go along with it. Interested? [Visit our website](#) to learn more.

Pinnacle Solutions Resources:

<https://thepinnaclesolutions.com/solutions/shelter/>

<https://thepinnaclesolutions.com/about-amazon-web-services/>

Additional Resources & Info:

<https://aws.amazon.com/big-data/>