CLOUDWALKER

aws

Level Up
your data strategy
with AWS Analytics

Get ready for the Gen AI revolution with
customised cloud and data analytical solutions
that turn your data into a competitive advantage.

Better decisions make better businesses

In the highly competitive industries of gambling and real money gaming, data is the key to staying ahead. And with massive growth predicted in both sectors, the stakes have never been higher:

Between 2024 and 2029 the gambling industry is expected to grow from $244bn to

$397bn

and the real money gaming industry from $98bn to $133bn

But what does this mean for you and your data?

As part of your current priorities, you might be focusing on driving greater efficiency in your operations, or possibly experiencing a period of significant growth. Whatever your challenges or ambitions, it's likely that you're looking to use your data to make better decisions quickly as changes occur.

And here at Cloudwalker, we know that better decisions make for better businesses.

Level up your storage and analytics capabilities

Achieving the agility to make faster, data driven decisions requires a complete view your customers and business operations. This in turn means integrating terabytes, petabytes, or even exabytes of data – with much of it being siloed or unstructured.

Traditional on-premises data analytics solutions can't handle this approach because they don't scale well enough and are too expensive. As a result, there's an acceleration in companies looking to modernize their data and analytics infrastructure by moving to the cloud.

In this guide, we'll look how you can leverage the power of AWS Lakehouse Infrastructure to level up your storage and analytics capabilities to turn your data into a competitive advantage.

Customer data in the real world

To analyze vast amounts of data, many companies are moving all their data from various silos into a single location, often called a data lake, to perform analytics and machine learning (ML).

But with different types of purpose-built data stores feeding into a data lake, a one-size-fits-all approach to data analytics no longer works because it inevitably leads to compromises.

Getting the most from data lakes and these purpose-built stores means moving data between these systems easily. Typically, there are three main scenarios:

Inside-out data movement:

Transferring data from a data lake to a data warehouse, e.g., for daily reporting.
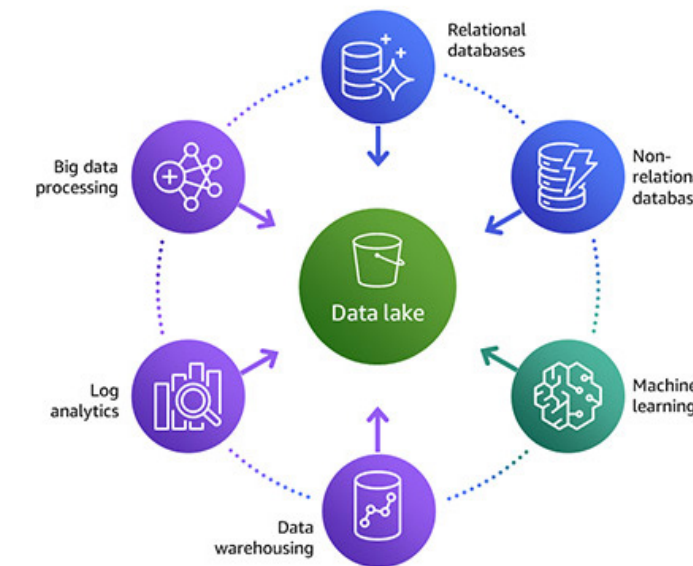
Outside-in data movement:

Bringing data into the data lake from an application or other data base, e.g., importing log analytics or ML output.

Around-the-perimeter movement:

Passing data from one purpose-built data store to another, e.g., for comparing regional user behaviors and running recommendation algorithms.

Customer data in the real world

As data in these data lakes and purpose-built stores continues to grow, it becomes harder to move all this data around.

To make decisions with speed and agility, you need the ability to use a central data lake and a ring of purpose-built data services around that data lake – all while ensuring data can easily move between stores and services in a secure and governed way.

Achieving all this requires a data architecture that supports the following:

Building a scalable data lake rapidly.

A broad and deep collection of purpose-built data services that provide the performance required for interactive dashboards and log analytics.

Moving data seamlessly between the data lakes and purpose-built data services.

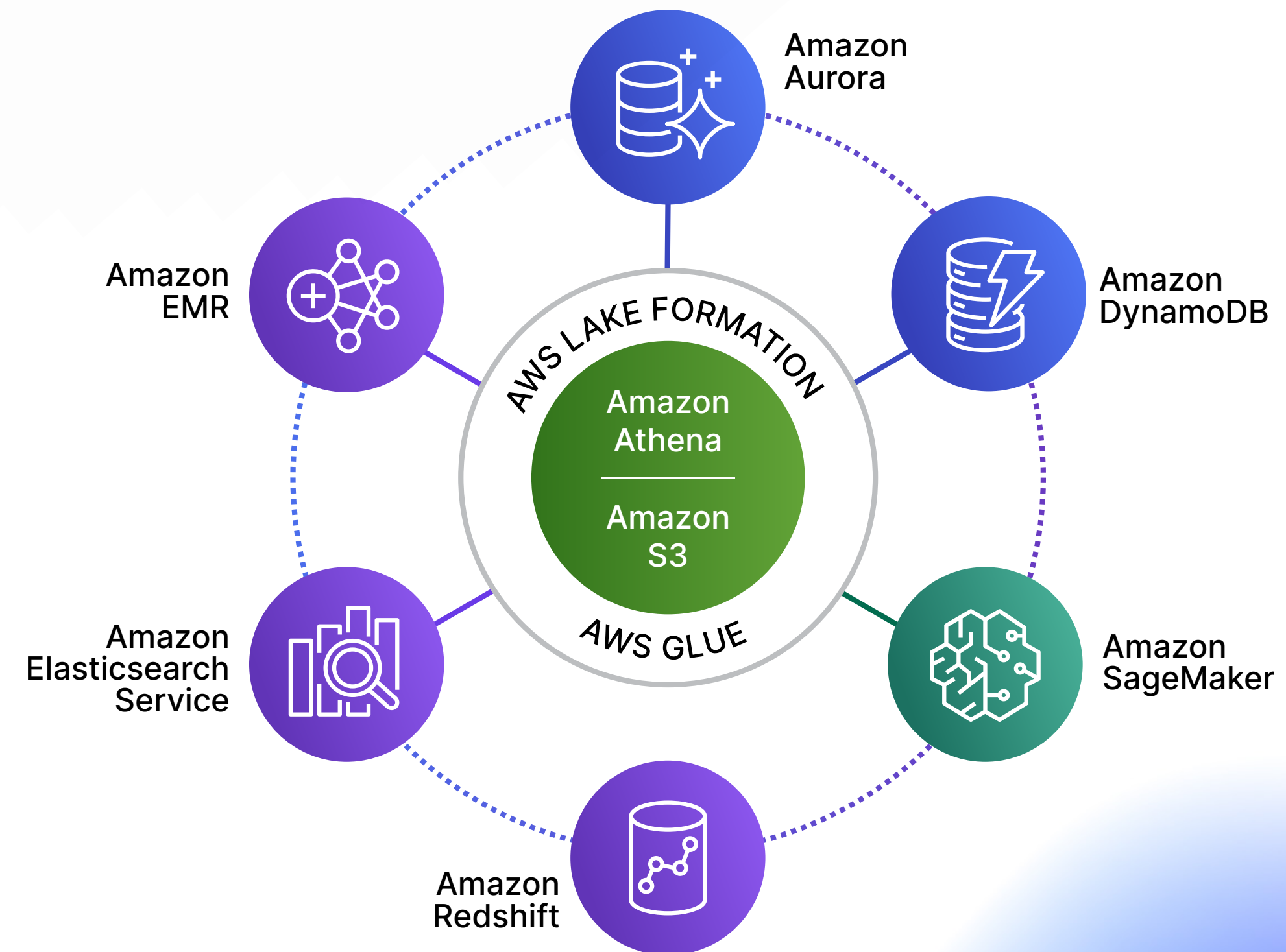Ensuring compliance via unified, secure, and managed access to data.

Scaling systems at low cost without compromising on performance.

This modern approach to analytics is called the Lake House Architecture.

Lake House Architecture on AWS

A Lake House Architecture acknowledges the idea that taking a one-size-fits-all approach to analytics eventually leads to compromises. It is not simply about integrating a data lake with a data warehouse, but rather about integrating a data lake, a data warehouse, and purpose-built stores and enabling unified governance and easy data movement.

Let's take a look at how the Lake House Architecture on AWS and some of its advanced capabilities that can help you turn your data into a competitive advantage.

Amazon Aurora

Amazon DynamoDB

Amazon EMR

AWS LAKE FORMATION

Amazon Athena

Amazon S3

AWS GLUE

Amazon SageMaker

Amazon Elasticsearch Service

Amazon Redshift

Scalable data lakes

Amazon Simple Storage Service (Amazon S3) is the best place to build a data lake because it has unmatched durability, availability, and scalability, with the best security, compliance, and audit capabilities. It also offers the fastest performance at the lowest cost, the most ways to bring data in, and the most partner integrations.

However, setting up and managing data lakes involves a lot of manual and time-consuming tasks such as loading data from diverse sources, monitoring data flows, setting up partitions, turning on encryption and managing keys, reorganizing data into columnar format, and granting and auditing access.
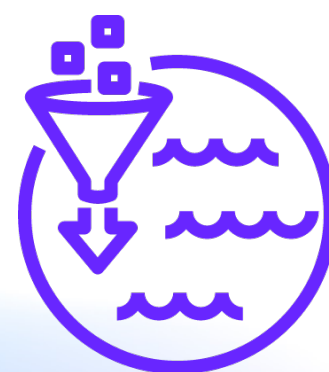
To help make this easier, AWS Lake Formation helps you to build secure data lakes in the cloud in days instead of months. Lake Formation collects and catalogs data from databases and object storage, moves the data into an Amazon S3 data lake, cleans and classifies data using ML algorithms, and secures access to sensitive data.

AWS Lake Formation governs access to S3 buckets by using fine-grained access control policies and integration with AWS Identity and Access Management (IAM). It manages permissions at the table, column, and row levels, rather than just at the bucket or object level in S3.

Tag-Based Access Control: It supports attribute-based access control by tagging data and managing access through tags, making it easier to apply consistent policies across the data lake.

Purpose-built analytics services

AWS offers the broadest and deepest portfolio of purpose-built analytics services that are all built to be best-of-breed, which means you never have to compromise on performance, scale, or cost when using them:
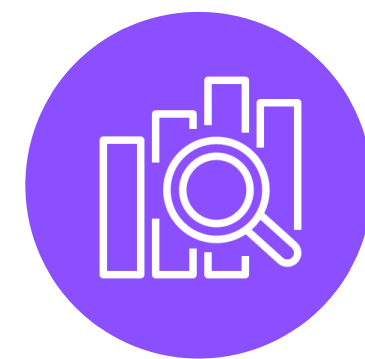
Amazon Athena
Interactive query

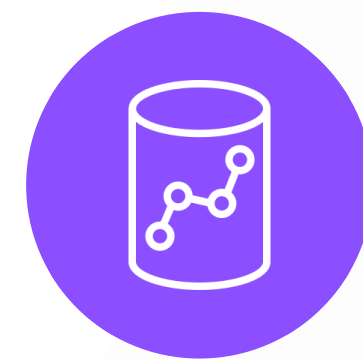**Amazon QuickSight**
Big data processing

Amazon Elasticsearch Service
Log and search analytics

Amazon Kinesis Data Analytics
Real-time analytics

Amazon Redshift
Data warehousing

Automatic Table Optimizations (ATO) for Amazon Redshift simplifies performance tuning of Amazon Redshift data warehouses by using ML to automate optimization tasks such as setting distribution and sort keys to give you the best possible performance without the overhead of manual performance tuning.

Amazon QuickSight Q uses ML to generate a data model that automatically understands the meaning and relationships of business data, enabling you to ask ad hoc questions about your business data and get answers to your questions instantly without having to wait for modeling from your business intelligence (BI) teams.

Seamless data movement

With data stored in many different systems, you need to be able to easily move that data between all of your services and data stores: inside-out, outside-in, and around-the-perimeter. No other analytics provider makes it as easy to move data at scale to where it's needed most.

AWS Glue is a serverless data integration service that allows you to easily prepare data for analytics, machine learning, and application development. AWS Glue provides all the capabilities needed for data integration, so insights can be gained in minutes instead of months.

Amazon Redshift and Athena both support federated queries. This allows you to run queries across data stored in operational databases, data warehouses, and data lakes to provide insights across multiple data sources with no data movement and no need to set up and maintain complex extract, transform, and load (ETL) pipelines.

Amazon Redshift data sharing also provides a secure and easy way to share live data across multiple Amazon Redshift clusters inside the organization and externally without the need to make copies or the complexity of moving around data.

You can use data sharing to run analytics workloads that use the same data in separate compute clusters in order to meet the performance requirements of each workload and track usage by each business group. For example, you can set up a central ETL cluster and share data with multiple BI clusters to provide workload isolation and chargeback.

Unified governance

One of the most important pieces of a modern analytics architecture is the ability to authorize, manage, and audit access to data. Enabling such a capability can be challenging because managing security, access control, and audit trails across all the data stores in an organization is complex and time-consuming. It's also error-prone because it requires manually maintaining access control lists and audit policies across all storage systems, each with different security, data access, and audit mechanisms.

With capabilities like centralized access control and policies combined with column and row-level filtering, AWS gives you the fine-grained access control and governance to manage access to data across a data lake and purpose-built data stores from a single point of control.

Row-level security for AWS Lake Formation makes it even easier to control access for all the people and applications that need to share data. For instance, by filtering and setting data access policies at the row level, you could set a policy that gives a regional sales manager access to only the sales data for their region. This level of filtering eliminates the need to maintain different copies of data lake tables for different user groups, saving you operational overheads and unnecessary storage costs.

Performance and cost-effectiveness

With AWS services from Cloudwalker, you can get the best performance at the lowest cost across all analytics services, with continual innovation improving the price-performance of these services.

In addition to industry-leading price performance for services like Amazon Redshift and AWS Glue, Amazon S3 intelligent tiering can save you up to 40% on storage costs for data stored in a data lake, and enable up to 35% lower costs.

You can also take advantage of AWS Redshift Reserved Nodes, a flexible pricing model that provides savings of up to 72%, while the AQUA (Advanced Query Accelerator) for Amazon Redshift delivers up to ten times faster query performance than other cloud data warehouses.

Cloudwalker: Your data analytics and warehousing partner

At Cloudwalker we're proud to be an AWS Advanced Tier Services Partner, seamlessly delivering Data Analytics & Warehousing solutions tailored to your needs.

We use our expertise in Lake House Architecture on AWS to offer solutions tailored to the challenges and opportunities of the real money gaming and gambling industries.

We'll help you get the best from services like Amazon Redshift and AWS Glue to empower your organization to make data-driven decisions and gain valuable insights, while ensuring your data is clean, organized, and readily available for analysis.

200+
Relational databases migrated to Amazon S3 Data Lake

5x
validations AWS Service

50+
AWS Certifications

Whatever your priorities or ambitions, Cloudwalker can help you harness the broadest and deepest portfolio of data services to realize a Lake House Architecture featuring:

- Scalable data lakes
- Purpose-built analytics services
- Seamless data movement
- Unified governance
- Performance and cost effectiveness

CLOUDWALKER

Let's get started

To start unleashing the power of your data with AWS and Cloudwalker, contact us to discuss how we can deliver seamless solutions tailored to your business needs.

Contact us