



# UNITE YOUR ENTERPRISE WITH A MODERN CLOUD DATA PLATFORM

One governed, secure, and highly extensible solution  
for all your data, workloads, and applications



# TABLE OF CONTENTS

- 2** A data platform for all your workloads
- 3** A flexible and extensible platform for any cloud
- 4** What to look for in a modern cloud data platform
- 5** A unified data platform architected for every cloud
- 7** Secure collaboration across all data
- 9** One platform for building data applications
- 11** From data platform to enterprise ecosystem
- 13** Moving forward with a data platform built for the cloud

# A DATA PLATFORM FOR ALL YOUR WORKLOADS

In nearly every industry, today's market leaders and aspiring disruptors have one thing in common: they put data at the center of their operations. Younger firms, born in the cloud, build their businesses on data, while established enterprises learn how to expeditiously acquire, analyze, and share their data to discover valuable insights.

But even as data becomes progressively more important, the analytic capabilities that make that data useful have become a moving target. For example, in years past, companies created traditional data warehouses to store and analyze data from enterprise applications. Next came data lakes, driven by steady advancements in data science and a desire to store and examine non-relational data types. Today we commonly see a mix of data processing technologies and analytics techniques, each of which provides limited insight from a unique slice of data.

It's been a wild ride, especially for IT professionals charged with developing, deploying, and managing these solutions. Each analytic endeavor necessitates a unique set of tools for transforming, loading, securing, and managing data, and many require custom infrastructure in the center. Technology specialists labor to rationalize these dissimilar technologies, while business leaders are forced to make critical decisions based more on gut instinct than on data-driven facts.

Whether they are legacy solutions ported to the cloud or legacy on-premises platforms running in a data center, many analytic solutions create and rely on disparate data silos, increasing complexity for IT professionals and delaying time to value for the business community. What's needed is a modern data platform—one that's born in the cloud: a platform that delivers the best of modern data warehousing, the best of data lakes, and much more, so organizations can shift their focus from managing infrastructure to managing data, and to getting every insight possible from that data.

This ebook explores the limitations of traditional data management and analytical practices. It also recommends what a cloud data platform should

provide to enable powerful, versatile, and extensible decision-making capabilities by bringing all of your data into one cohesive platform. Read on to learn how you can simplify your technology stack and unite your data, data management, and analytics endeavors with a platform that addresses your current data needs and is extensible to support new types of workloads in the future.



# A FLEXIBLE AND EXTENSIBLE PLATFORM FOR ANY CLOUD

A properly architected cloud data platform automates many essential tasks, from how data is stored and processed to transaction management, security, governance, and metadata management (see Figure 1).

It streamlines how teams work with data by facilitating secure collaboration across the data, without having to move that data from place to place, and creates new opportunities and revenue streams for securely sharing and even monetizing the data you choose to make available to other organizations. The platform should enable you to

utilize any cloud—including popular offerings from Amazon, Microsoft, and Google—and work freely across them. It should seamlessly support many different workloads, from data warehouses and data lakes to data pipelines and data exchanges, along with many types of business intelligence, data science, and data engineering applications.

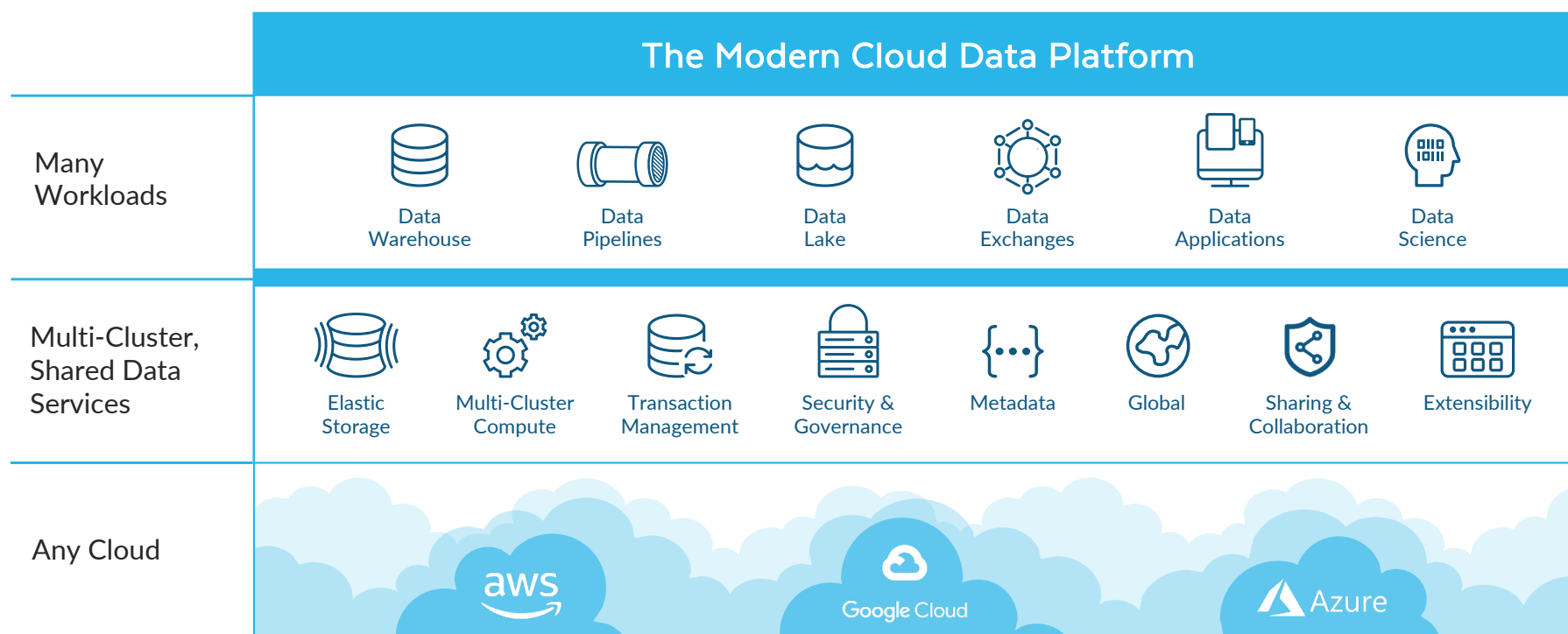


Figure 1: Only a unique multi-cluster architecture that works with any cloud delivers a host of powerful services to enable a number of modern use cases and workloads.

# WHAT TO LOOK FOR IN A MODERN CLOUD DATA PLATFORM

Your modern cloud data platform should:

- Operate across regions or across clouds
- Be delivered as a service, scaling instantly and automatically
- Synchronize concurrent workloads, comprising all data types across multiple clouds
- Work with data where it resides

## CLOUD AGNOSTIC

With a cloud-agnostic platform, you can distribute data and architecture across regions and cloud providers, while maintaining the same experience for all users and workloads—including data warehouses, data lakes, data pipelines, data exchanges, data applications, and data science activities.

## DELIVERED AS A SERVICE; INFINITELY SCALABLE

The service layer, identified as multi-cluster, shared data services in [Figure 1](#), is the brain of this architecture. It should enable you to automatically scale instantly and near-infinitely—up, down, and out—as you load data and simultaneously execute workloads on that data. The platform should also be delivered as a service to simplify administration and eliminate all the maintenance activities that on-premises platforms require, including optimizing performance; minimizing resource contention; and manually attending to security, governance, partitioning, and metadata management.

## SYNCHRONIZE WORKLOADS OF ANY TYPE, ON ANY DATA

The platform should execute a near-unlimited number of concurrent workloads, without any limitations on performance or scale, encompassing all types of data and spanning multiple clouds. These different types of workloads should all get instant access to any amount of dedicated compute power that scales up without contention and automatically scales down the second it's no longer needed. And because the platform scales to the exact amount of compute power your organization

needs at any one time, there should be no need for capacity planning. There should be no one-size-fits-all, and you should pay only for what you use.

The platform should also natively ingest structured data and semi-structured data, including JSON, Avro, and XML, along with open source data types such as Apache Parquet and ORC—all at petabyte scale. The platform should leverage the elasticity, scale, and low cost of blob storage. It also should enable native data loading and analytics on these mixed data formats with complete transactional integrity, and store these diverse data types in their native forms, without creating new data silos.

## WORK WITH DATA WHERE IT RESIDES

As an extensible repository and unified data service, the cloud data platform you choose should work with data where it resides, eliminating the need to move from one data silo to another in order to analyze multiple data sets located in multiple locations. Instead, it should allow you to easily manage data in many disparate places and many diverse workloads from one single service.

In the pages that follow, you will see how a cloud data platform can satisfy all of these important needs, with a foundation built on three fundamental pillars:

- A unified data platform architected for every cloud
- Secure collaboration across all data assets
- One platform for building all types of data applications



# A UNIFIED DATA PLATFORM ARCHITECTED FOR EVERY CLOUD

While many software vendors build cloud data warehouses on revamped on-premises infrastructure, your modern data platform must be architected first and foremost for the cloud. This is important for several reasons, including:

- Improved scalability and performance; no vendor lock-in
- Ability to mix and match clouds
- Global data replication, minimal disruption
- Work with data where it resides

## IMPROVED SCALABILITY AND PERFORMANCE; NO VENDOR LOCK-IN

Traditional data platforms are derived from legacy environments that work on closed networks with local data. These platforms don't scale well. As analytic applications gain popularity, legacy platforms slow down under the weight of many concurrent users and large volumes of diverse data. For better or worse, many of these on-premises systems have been ported to the cloud. Lacking a multi-cluster, shared-data architecture, these systems are still plagued with the limitations of poor scalability, concurrency, and performance.

Furthermore, as these age-old architectures are modified for cloud hosting, they typically work with only one provider. That means customers are locked into a single cloud, such as Amazon, Google, or Microsoft, which inhibits the ability for organizations to be locally or globally cloud-agnostic. This limitation doesn't match the diversity of today's software market, since most organizations are shifting from working with one cloud provider to diversifying with several. 84% of large organizations have a multi-cloud strategy, and 56% spend more than \$1.2 million a year on more than one public cloud.<sup>1</sup>

## ABILITY TO MIX AND MATCH CLOUDS

To maximize versatility, your modern cloud data platform architecture should support multiple clouds and multiple regions, and should allow you to mix and match clouds as you see fit—a capability

<sup>1</sup> <https://info.flexerasoftware.com/SLO-WP-State-of-the-Cloud-2019>

## The Value of a Cloud Data Platform

### Utilize any cloud:

Replicate data across regions and clouds, yet keep your data and apps in place, maximizing opportunities and improving business continuity.

**Gain unlimited scale, performance, and concurrency:** Provision limitless amounts of computing power to deliver analytic capabilities to any number of users and workloads, without impacting performance.

**Turn data into insight:** Unite internal business units and collaborate with partners by easily sharing secure and governed data in real time, without spending all your time managing infrastructure.

**Pay only for the resources you use:** Per-second pricing allows you to economically scale platform resources—up, down, and out—without paying for idle capacity.

known as cross-cloud. It should also separate storage resources from compute resources, with each scaling independently from the other for maximum performance, scale, concurrency, and efficiency.

This type of global platform will ensure a consistent experience regardless of where you store your data or which cloud providers you use. Delivered

as a service, and with consistent functionality across clouds, this advanced architecture delivers a near-unlimited elasticity and allows one cohesive platform to serve all types of users and workloads in a consistent way.

### GLOBAL DATA REPLICATION, MINIMAL DISRUPTION

To maintain business continuity and minimize disruption, your modern cloud data platform should include replication across cloud services. Cross-region and cross-cloud database replication and database failover and failback should provide modern data sharing, high availability, and quick data recovery. This allows organizations to take protective measures to ensure data availability and durability and to seamlessly share data with other consumers across regions and clouds.



# SECURE COLLABORATION ACROSS ALL DATA

Data platforms come in various flavors. When choosing your modern data platform, insist on the following characteristics to ensure secure collaboration across all your data, with all your users, both inside and outside your organization:

- Seamless data sharing; ability to build controlled data services
- Micro-partitioning for improved performance
- Multilayered security

## SEAMLESS DATA SHARING; ABILITY TO BUILD CONTROLLED DATA SERVICES

As data becomes more and more important to business activities, users everywhere need the flexibility to collaborate on mission-critical projects. Unfortunately, legacy data platforms foster the creation of multiple data silos, requiring inconsistent administrative procedures and the frequent movement of data from place to place. In order to integrate business units inside your organization, as well as share data with your customers and business partners, these older systems require you to set up extract, transform, and load (ETL) procedures, file transfer protocols (FTP), hand-coded application programming interfaces (APIs), and shared object stores.

A modern cloud data platform provides secure data sharing. Unlike with traditional methods, this secure data sharing doesn't require data copy or movement and eliminates the need for costly and cumbersome

ETL processes or development and maintenance of APIs. It also allows the creation of centrally managed data exchanges that act as data hubs and allow for secure data sharing at scale. Internal stakeholders can build controlled sets of data services that can be secured and governed for use across multiple users and departments.

Your cloud data platform should also make it easy to create data exchanges and marketplaces in which external parties such as customers, suppliers, and business partners can participate by securely publishing and accessing shared data. When you leverage a modern cloud data platform, you can quickly set up data exchanges to enhance secure collaboration, improve data access, reduce data sharing costs, and monetize data assets.

## MICRO-PARTITIONING FOR IMPROVED SCALE AND PERFORMANCE

To improve query performance for large amounts of structured and semi-structured data, a modern cloud data platform should be transparently divided into micro-partitions—contiguous units of storage that are organized in a columnar fashion. This structure allows for extremely granular pruning of very large tables, which can comprise millions, or even hundreds of millions, of micro-partitions to significantly improve query performance.

## MULTILAYERED SECURITY

Finally, and most importantly, don't skimp on security. All aspects of a cloud data

## The Value of Built-in Secure Data Sharing

### Gain seamless collaboration on mission-critical projects:

Enable stakeholders from various business units and departments to share data without moving the data or creating data silos.

### Monetize your data:

Get the most out of your data by creating secure data exchanges and marketplaces without having to create complex interfaces or deploy cumbersome file transfer procedures.

### Enjoy multilevel security:

Safeguard your data with encryption, access control, comprehensive monitoring, alerts, cybersecurity practices, and more—automatically.

platform—architecture, implementation, and operation—must center around protecting your data, both in transit and at rest. The modern cloud data platform should have a multilayered security strategy baked into the platform that addresses encryption, access control, data storage, and physical

infrastructure, in conjunction with comprehensive monitoring, alerts, and cybersecurity practices.

Work only with a cloud data platform provider that deploys industry-standard, end-to-end security practices. Make sure these security mechanisms are built into the foundation of the data platform,

complementing and extending the security provided by cloud infrastructure providers. You should gain these additional capabilities by default as part of the data platform; you shouldn't have to do anything extra to have the best data security available (see Figure 2 below).

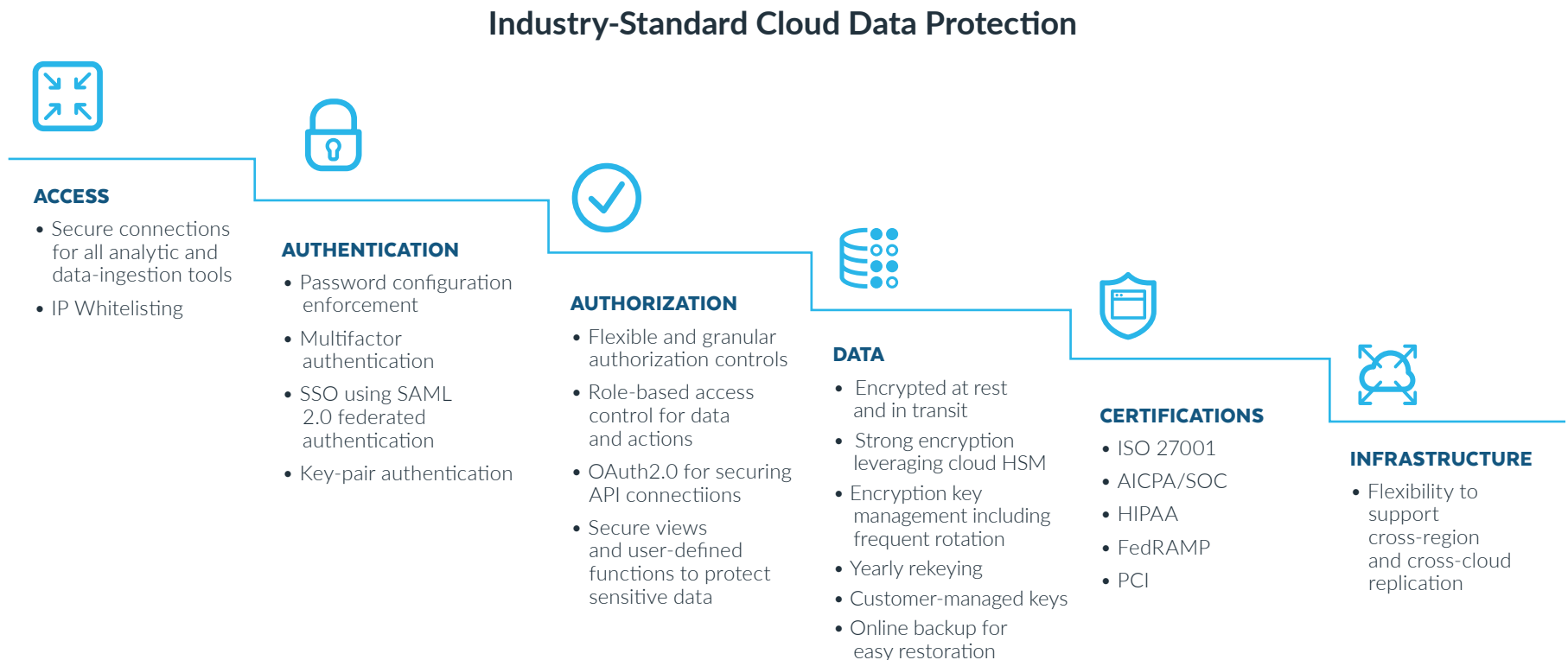


Figure 2: Make sure your cloud data platform employs multilevel protections—complementing and extending the security supplied by the cloud storage vendor.

# ONE PLATFORM FOR BUILDING DATA APPLICATIONS

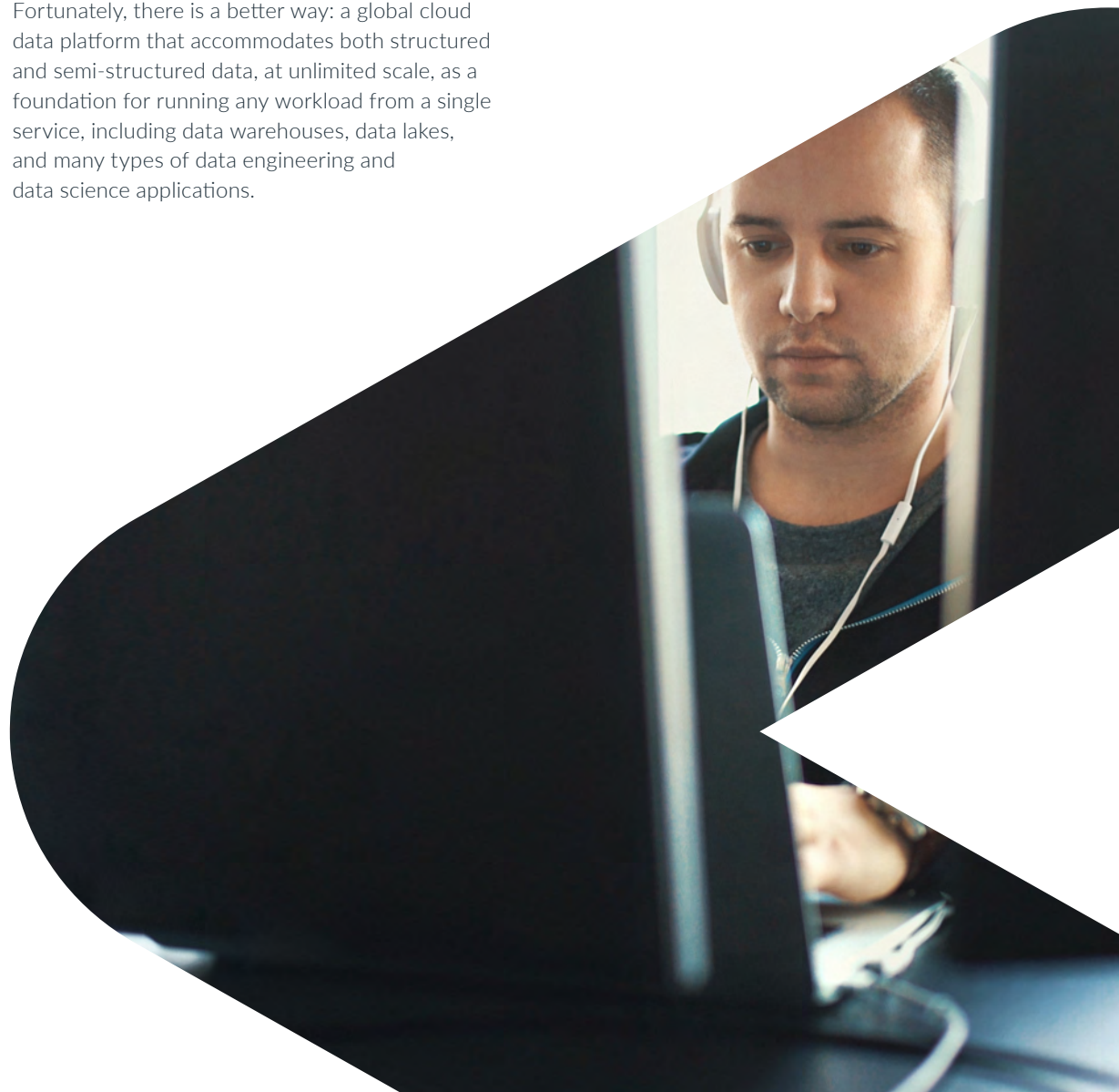
A modern cloud data platform paves the way to create innovative approaches for getting more value from your data. Look for a data platform that lets you:

- Run any workload from a single service
- Expand and scale while paying only for the compute power you use
- Manage your data—not your data platform

## RUN ANY WORKLOAD FROM A SINGLE SERVICE

In an era of rapid technology advancement, complexity is the enemy of progress. It's hard to gain value from your data when each new initiative requires a specialized solution. Each analytic application creates a data silo, which increases costs, reduces performance, and stalls creativity and productivity for end users, who must often wait for data and contend with data quality problems.

Fortunately, there is a better way: a global cloud data platform that accommodates both structured and semi-structured data, at unlimited scale, as a foundation for running any workload from a single service, including data warehouses, data lakes, and many types of data engineering and data science applications.



### EXPAND AND SCALE WHILE PAYING ONLY FOR THE COMPUTE POWER YOU USE

The cloud data platform you choose must be extensible—designed from the ground up to be open and accessible to other solutions. An extensible data platform enables you to maximize the value of your data and all the services that connect to it, now and in the future. Metered, per-second usage lets you expand and scale these applications, yet pay only for the computing power you use.

### MANAGE YOUR DATA—NOT YOUR DATA PLATFORM

To gain this type of simplicity, avoid architectures that have been stitched together using proprietary components that all require specialized training and skills. Remember the goal: near-zero maintenance, no infrastructure to manage, and no knobs to turn, so you can focus on the data and not on managing the platform (see Figure 3).

#### What Your Cloud Data Platform Should Enable



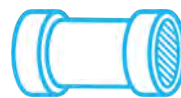
The Modern Data Warehouse



Augmented Data Lakes



Accelerated Analytics



Integrated Data Engineering



Secure Data Exchange



Agile Data App Development



Advanced Data Science

Figure 3: Your platform should enable many use cases and workloads, so your organization can leverage the power and simplicity of a single solution for data analytics.

## The Value of One Platform for Building SaaS Apps

**Innovate without complexity and cost:**  
Build new analytic applications to gain additional value from your data but without creating data silos, increasing costs, or reducing performance.

**Manage your data, not your data platform:**  
A global, extensible cloud data platform frees your organization from proprietary components that require specialized training and skills, so you can focus on your data, not your data platform.

# FROM DATA PLATFORM TO ENTERPRISE ECOSYSTEM

A properly architected cloud data platform becomes the core of an organization's data analytics strategy across and beyond the enterprise. Look for a data platform that is:

- Highly extensible
- Integrates with a broader ecosystem
- Unifies real-time analytics capabilities

To get the most value from your data platform, make sure it is highly extensible and supports an ecosystem of solutions, regardless of the source and type of data each solution uses or how it is loaded, integrated, and transformed (batch, microbatch, or streaming). A properly architected cloud data platform should also enable integration with third-party solutions for added data security, cataloging, or business intelligence and analytics. (see [Figure 4](#) on the following page).



## The Business Value of a Modern Cloud Data Platform

### **Better, faster business decisions:**

Give your business users the deepest possible insights from all your data, with a platform that makes it all look easy.

### **Cost optimized for today's needs:**

Minimize costs with near-zero maintenance, pay-by-the-second data solutions that scale instantly to the size you need.

**Extensible to accommodate upcoming journeys:** Eliminate access barriers by standardizing on a cloud data platform that works with many other types of software applications, services, and components.

## The Modern Enterprise Data Ecosystem

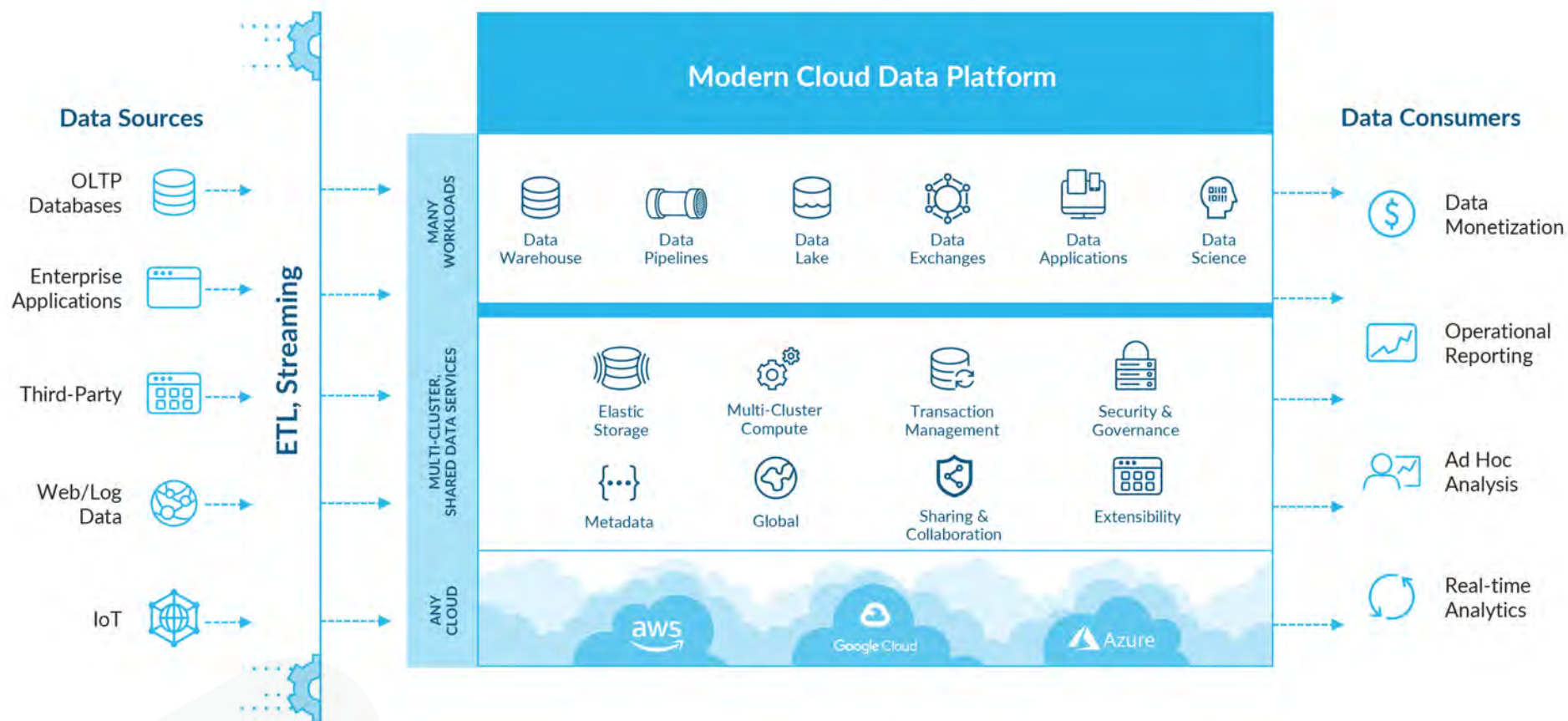


Figure 4: A modern platform should power and extend your data architecture, so you can easily get all your data into a single location and get all the insights from that data.

# MOVING FORWARD WITH A DATA PLATFORM BUILT FOR THE CLOUD

Market leaders rely on data from all types of sources, from point-of-sale solutions to financial applications to social media networks. Due to differences in these data types—and further differences in the ways in which the data is acquired, managed, and used—data professionals have to develop unique data strategies and technologies for each specific workload and each unique analytic need. The result is a complex data architecture that is difficult and costly to manage.

A modern cloud data platform brings these disparate elements together with a unified solution that supports many types of data, workloads, analytics, applications, and clouds. It eliminates the cost and headache of having to learn new ways to ingest, duplicate, and share data. And it allows data to be stored in its native form, supporting a broad range of use cases. This platform has the essential ingredients to position you for success, now and in the future:

- A multi-cluster, shared data architecture that spans multiple clouds
- Integrated capabilities for partitioning, sharing, and accessing all types of data
- Built-in and extensible services that reduce the cost and enhance the value of all types of analytic workloads

As a cloud service, your data platform should eliminate the expense of buying, maintaining, and securing on-premises solutions, and “cloud-washed” solutions, which have many of the same limitations. Your cloud data platform vendor should take responsibility for server provisioning, data governance, data protection, data security, and performance tuning. This frees your team to focus on gaining the most value from your data. And since cloud storage typically costs only a fraction of what you’ve been paying for traditional and limited on-premises storage, you can meet today’s demands for data volumes, velocity, and variety on a scale unimaginable only a few years ago.

To learn more, visit  
<https://www.snowflake.com/product/>.





## ABOUT SNOWFLAKE

Snowflake's cloud data platform shatters the barriers that have prevented organizations of all sizes from unleashing the true value from their data. More than 2,000 customers deploy Snowflake to advance their businesses beyond what was once possible by deriving all the insights from all their data by all their business users. Snowflake equips organizations with a single, integrated platform that offers the only data warehouse built for the cloud; instant, secure, and governed access to their entire network of data; and a core architecture to enable many types of data workloads, including a single platform for developing modern data applications. Snowflake: Data without limits. Find out more at [snowflake.com](https://www.snowflake.com).

