CLOUDERA

Achieving Cloud Anywhere: A Comprehensive Roadmap

Table of Contents

CLOUDERA

Introduction

By now, most businesses have embraced hybrid cloud. As a strategy, hybrid cloud offers greater flexibility to tap into the vast amounts of data necessary to fuel artificial intelligence (AI) and drive advanced analytics projects. Access to that data is a game-changer. It can reveal trends using historical information or offer users a deeper look into the state of their organization and its operations.

Whatever the intent, it's not enough to just bring data together across any combination of environments. The real value of data is its reliability—in other words, can it be trusted? If the answer is no, then those AI and machine learning (ML) models it feeds will yield bad outputs and inaccurate insights, resulting in bias and flawed decision-making. Avoiding these pitfalls requires going beyond the standard hybrid approach to establish 'cloud anywhere'.

But exactly what does it mean to deliver cloud anywhere? How does it differ from the traditional hybrid or multi-cloud approach? And what's needed to get your data, analytics, and AI infrastructure there? In this ebook, we'll answer these questions and explore some of the key advantages of cloud anywhere.

70%

of organizations take a hybrid approach according to Flexera's 2025 State of the Cloud report.

CLOUDERA

Chapter 1: What Does It Take to Deliver Cloud Anywhere?

Cloud anywhere is a mixed infrastructure—inclusive of on-premises and cloud—environment that integrates seamlessly allowing for data, applications, and workloads to move freely across environments and acting as a single cloud and delivering a consistent cloud experience everywhere. From a user perspective, there is no noticeable distinction between where data or where workloads are deployed, maximizing flexibility, and efficiency.

There are several key factors to determine if a hybrid cloud approach meets the underline criteria of cloud anywhere. One of the most significant differentiators is the ability to support all infrastructures—across the data center, cloud, and at the edge—on a single platform while providing identical, portable and a consistent cloud experience across all environments.

According to Cloudera research, nearly all organizations (93%) recognize the importance of hybrid cloud capabilities for data and analytics in helping them adapt to change. For example, a hybrid cloud approach allows authorized personnel across an organization to have access to data from multiple locations, allowing them to analyze all the relevant data available rather than be limited to data stored in a single deployment, infrastructure, or format.

The right data platform for true hybrid will deliver the following:

- A distributed cloud model
- Portable, interoperable data services
- Data stores to address every eventuality
- A common interface to manage infrastructure
- Cross-platform security and governance
- Openness and extensibility

CLOUDERA

A hybrid cloud approach also makes organizations more adaptable to new technologies, including AI. But the aim isn't just to react to today's technology. Hybrid cloud takes it a step further—giving organizations peace of mind knowing that their infrastructure can adapt to whatever new technology is on the horizon.

It's important to remember that while supporting multiple clouds offers significant benefits, that alone is not a complete solution for organizations with a considerable data center investment or who must keep datasets on-premises. Data sovereignty is an ever greater driver for organizations across the globe as they determine the most appropriate infrastructure for their data and AI. A cloud anywhere model enables an organization to keep data and workloads moving multi-directionally and friction-free while having identical and portable functionality available across infrastructures.

We've talked extensively about data thus far, but we haven't delved into its forms—structured, unstructured, and semi-structured.

Understanding, handling, securing, and governing structured, semi-structured, or unstructured data is a core part of what makes a hybrid approach valuable. To seamlessly move data across infrastructures in a cloud anywhere model, security and governance must remain consistent. This leads to true hybrid heterogeneity, allowing organizations to orchestrate data and workloads across the entire IT landscape, including multiple infrastructures.

Now that we understand cloud anywhere, where do we go from here? For anyone looking to redefine their hybrid approach, the next step would be to build a strong foundation of trusted data.

> **Defining data:** structured data is highly organized and formatted in a way that makes it easily searchable in databases and data warehouses. Unstructured data lacks a predefined format or structure, making it more complex to manage and utilize.

CLOUDERA

Chapter 2: The Keys to Establishing Trusted Data

The goal of cloud anywhere and modern technologies is to work with trustworthy data to make better strategic decisions. Accomplishing this requires overcoming several obstacles related to data privacy, data reliability, and data security.

From a privacy perspective, cultivating trusted data means understanding what data is available, who may (or may not) access it, and what environment it will be used in. When you're dealing with large volumes of data, there are a variety of privacy considerations to keep in mind. Specifically, we can pinpoint three areas of consideration: local or regional regulation (predominantly data privacy; GDPR, CCPA), industry regulations (mostly for highly regulated industries like HIPAA, BASEL IV), and self-imposed regulation. Understanding and solving for each requires a deep knowledge of your data landscape—in other words, you need to know it can be trusted.

Not adhering to data privacy guidelines and data protections leaves organizations vulnerable and can result in government penalties and reputational damage. Whether it's a data breach stemming from improper access controls or an intentional breach of systems housing sensitive data, the outcome can be devastating. These scenarios can result not only in financial loss, but serious damage to a brand's reputation. A granular understanding of an organization's data flow—where it's being stored, where it's going, and who has authorized access—can protect against these vulnerabilities and greatly reduce risk.

With so many questions surrounding data at any given time, data security and data management become paramount. Connecting users to data via the cloud introduces more endpoints, increasing access to sensitive information compared to a strictly on-premises environment. As a result, the movement of data to new places demands a heightened standard of security. Consequently, security and compliance remain among the top concerns for IT modernization and organizations transitioning more comprehensively to the cloud. Managing access to sensitive information should always be a priority.

CLOUDΞRA

To ensure users have access to the most current and relevant data, silos must be eliminated. By removing silos, users eliminate the risk of duplicate dataset creation and ensure that up-to-date data isn't located in one environment, while outdated datasets remain elsewhere.

To manage the authentication of those who should have access to specific data sets, organizations should have an access control management solution in addition to strong data management practices. Enterprises and IT teams should follow the Identity, Credential, and Access Management (ICAM) framework to ensure that only the right people access the right data, at the right time. Cloudera's Shared Data Experience (SDX) helps IT teams with ICAM by delivering integrated security and governance across an environment.

A cloud anywhere strategy includes a modern technology stack that unlocks data at scale through a variety of solutions, making it available for use while governing and securing it. More on the role of proven accelerators and blueprints for data architectures later.

52%

of IT leaders said that governance and compliance issues were holding them back from end-to-end data management.

CLOUDΞRA

Chapter 3: Portability of Data and Analytics Workloads Between Infrastructures

Once the foundation is set, organizations must next explore their approach to the data lifecycle.

In any hybrid cloud environment, data can be analyzed and utilized within a single deployment, while also moving between different deployments. This can take place in the cloud, on-premises, or at the edge. To optimize existing hybrid infrastructure, strong data management practices are essential. These practices ensure the portability of necessary workloads across infrastructures and enable data workloads to be created, deployed, or redeployed to other locations.

To facilitate this, strong data management capabilities must include data visibility, access to data aligned with internal governance policies, and the right security capabilities to ensure data remains safe throughout the entire life cycle.

A cloud anywhere approach goes a step further. Cloud anywhere accounts for the complete data lifecycle, spanning from the point of ingestion to transformation and warehousing, all the way through to machine learning and AI. It not only supports the lifecycle of analytics on a single infrastructure, but also supports the same workloads on all

infrastructures, with the ability to seamlessly move between them without the need for redevelopment or refactoring. This flexibility provides organizations with the tools required to adapt to an unpredictable future. It enables them to better respond to changes in technology, business processes, and regulations.

It also reduces complexity but eliminates the need for many additional data services that can make managing automated processes more difficult. A cloud anywhere platform delivers a consistent cloud experience by ensuring the same engine and services are available on all infrastructures, simplifying management and enabling a unified cloud experience to data anywhere. This structure enables one of the most significant benefits of a cloud anywhere approach—its ability to move away from infrastructure dependent deployments. A cloud anywhere approach eliminates a scenario where each infrastructure and process operates in its own closed ecosystems, which demand more resources, reduce flexibility, and result in higher costs and risks. By breaking down these silos, data processing becomes more trustworthy and consistent, empowering data practitioners to access high-quality information, analytics, and models in real time.

According to Cloudera research,

46%

of IT leaders reported that their organization interacts with all stages of the data life cycle process.

CLOUDERA

# Chapter 4: The Role of Proven Accelerators and Blueprints

The modern enterprise must use a variety of technologies to store, process, manage, and analyze data across an organization. Proven accelerators and blueprints provide authorized personnel user-friendly access to query and analyze disparate data exactly where it lives, ensuring full security and control across the entire environment.

The ideal data infrastructure for today's enterprise is built on three key approaches: unified data fabric, scalable data mesh, and the open data lakehouse. Together, they enable an organization to simplify data management across an environment, allowing IT managers to have both a bird's eye view of data flow and granular conditions on data access, all from a single pane. They also eliminate data silos, ensuring information remains up to date and reducing the need for replication.

By leveraging a data fabric, organizations can unlock disparate data sources across hybrid cloud environments, enabling them to fully understand and utilize their contents. The data fabric helps make them available in a safe, compliant, and self-service manner across the organization. Consistent data management delivers a strong foundation for accelerated value, insights, and all

modern data architectures. This enables an enterprise's data architecture to help organizations bring together disparate data into one single source—a critical part of what differentiates hybrid cloud and true hybrid cloud. It also gives businesses a single point of access for data stored securely in multiple systems. That said, depending on the organization, and its approach, there may be a single data fabric serving the whole enterprise or there may be a scenario where each domain has its own data fabric.

A data mesh also serves an important role because it is designed to be decentralized and self-service, giving options to implement data as a product. This may include infrastructure as well as engines that are all tied together through security and governance. The goal is to create a "mesh" of data products connected to various parts of an organization rather than organizing around a central repository. This decentralized approach benefits users because it provides simple authorized access and gives each domain autonomy, agility, and flexibility to craft their data as a product of their organization.

CLOUDERA
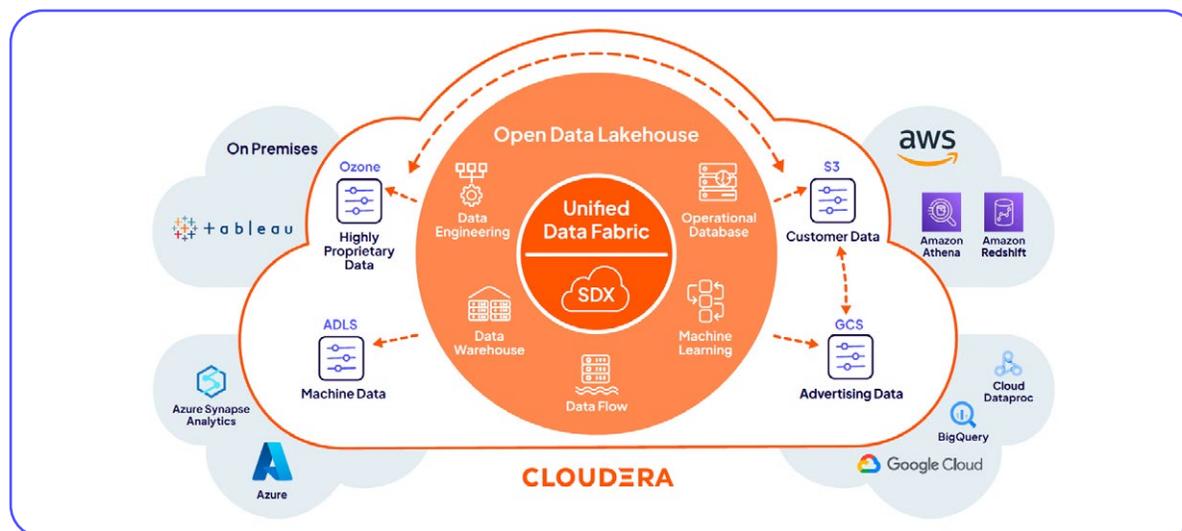
Meanwhile, a <u>data lakehouse</u> offers a single location for storing and accessing data across all analytics engines. It combines the strengths of data lakes and data warehouses to meet the complex needs of a modern hybrid cloud approach. Data lakehouses offer significant benefits, including reducing the need for data duplication and synchronization, which lowers costs. They also enhance trust by providing a central lakehouse that serves as a single source of truth or adopting a lakehouse for each domain.

Together, these three components of a modern data technology stack deliver key functionality in terms of flexibility and scalability of data management and ensure that the right, trusted data can be used to feed AI and analytics effectively. With these proven accelerators and blueprints in place, organizations have the foundation to begin capitalizing on emerging technologies, knowing their data will deliver the best possible outcomes.

> Cloudera's Open Data Lakehouse eliminates data silos and lets data teams collaborate on the same data with the tools of their choice on any cloud and data center.

CLOUDERA
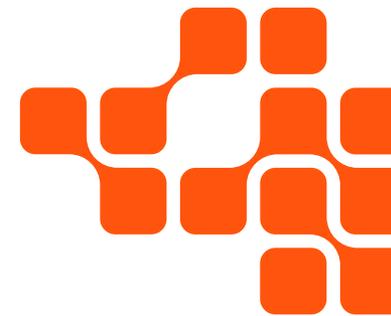
Chapter 5: Emerging Technologies in the World of Cloud Anywhere

By enabling organizations to easily adapt to new technologies and processes, cloud anywhere provides organizations with the flexibility to be prepared for the future. Whether it's AI, GenAI, Agentic AI, or any other number of emerging technologies, each one has its own set of unique needs. A cloud anywhere approach gives the flexibility to move AI or AI lifecycle parts where it will generate the greatest value. That flexibility allows you to bring AI models to the data, rather than pulling the data from where it is stored and moving it to where it can be used. This ensures data remains secure and compliant with data policies and governance guidelines regardless of any new technology.

Organizations can also avoid the need to re-architect their environment when moving from the data center to the cloud, or vice versa, making migration faster and reducing costs.

In a cloud anywhere approach, organizations can split important tasks and objectives across a multitude of infrastructures and move them from one to the other seamlessly. They can also ensure they are ready for the future. This creates an additional degree of freedom previously unattainable.

Cloudera is now partnered with Snowflake, the AI Data Cloud company, to bring enterprises an open, unified hybrid data lakehouse, powered by Apache Iceberg.

CLOUDꓱRA

We've discussed the foundational elements needed to establish a cloud anywhere approach. Next, let's turn our attention to how this approach is applied in the real world. While there are numerous use cases, we'll focus on three common scenarios to illustrate its practical impact:

- **Accelerate Cloud Migration:** A cloud anywhere platform enables a lift-and-shift by providing a like-for-like architecture on-premises and in the cloud. This architecture enables data teams to choose to either modernize before migration or shift to the cloud and then modernize the applications of their choice.

- **Adaptive Use Case Deployment:** With a cloud anywhere platform, customers can deploy use cases composed of multiple workloads, each using the most optimal resource profile, while still maintaining consistency through unified security and governance. They can deploy applications that dynamically adjust to new information through AI and ML abilities.

- **Continuous Optimal Deployment:** This deployment strategy not only guarantees that workloads operate at peak performance and cost efficiency, but also ensures that the architecture is future-proof. Organizations are not locked into a single infrastructure solution, and they can be agile enough to respond to market conditions quickly. Different teams and departments can then deploy workloads where it makes the most sense, whether it's on-premises or in the cloud, and later decide to shift elsewhere. This is a seamless exchange that does not risk compliance thanks to consistent security and governance across an enterprise.

You can learn more about how these scenarios manifest in the real world in this ebook, Three Scenarios Only Cloud Anywhere Unlocks.

CLOUDERA
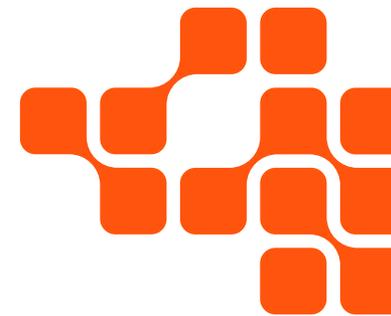
# Cloud Anywhere: A Necessity for the Modern Enterprise

Cloud anywhere is a differentiator for modern enterprises. It is the difference between creating new insight and value with AI and analytics tools or falling behind competitors in a crowded market.

The most advanced AI companies will have the most forward-thinking technology stacks and infrastructures, a true hybrid approach. As discussed, this entails establishing a foundation of trusted data, stronger data management standards for security and compliance, and a technology stack that facilitates proven accelerators, such as unified data fabric, scalable data mesh, and an open data lakehouse.

Ultimately, this allows users to work with and analyze trusted data through seamlessly integrated infrastructure in a revolutionary way—even with hybrid cloud—securely maximizing productivity and efficiency.

Cloudera's data and AI platform brings a deep set of solutions and knowledge to help organizations get the most value out of their data, make more informed decisions, and future-proof operations to scale and evolve as business needs change.

Ready to take the next step in your hybrid cloud journey to deliver a consistent cloud experience to your data anywhere, for AI everywhere? Find out how Cloudera's expertise and solutions can help you make it happen.

CLOUDꟻRA
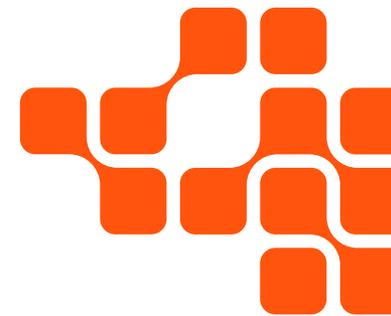
About Cloudera

Cloudera is the only data and AI platform company that large organizations trust to bring AI to their data anywhere it lives. Unlike other providers, Cloudera delivers a consistent cloud experience that converges public clouds, data centers, and the edge, leveraging a proven open-source foundation. As the pioneer in big data, Cloudera empowers businesses to apply AI and assert control over 100% of their data, in all forms, delivering unifi ed security, governance, and real-time predictive insights. The world's largest organizations across all industries rely on Cloudera to transform decision-making and ultimately boost bottom lines, safeguard against threats, and save lives.

To learn more, visit **Cloudera.com** and follow us on **LinkedIn** and **X**.

CLOUDΞRA

Cloudera, Inc. | 6220 America Center Dr, 5th Fl, San Jose, CA 95002 USA | cloudera.com