..

Overview

Date published: 2025-11-14
Date modified:

https://docs.cloudera.com/

Legal Notice

Contents

Overview

Familiarize yourself with Cloudera Hybrid Environments.

The conversation around cloud adoption has matured significantly. It is no longer a question of **whether** enterprises should use the cloud, but **how** they can strategically blend public cloud agility with the security and control of their on-premise infrastructure. This hybrid approach is now the dominant strategy for modern data-driven organizations.

At Cloudera, we believe a true hybrid cloud platform must deliver a seamless, unified experience. Our strategy is built on four key tenets:

Unified Runtime:

Ensure true **workload portability without any rewrites**, allowing applications to work and feel the same everywhere.

Hybrid Control Plane

Offer a **single pane of glass** for managing all private and public deployments.

Cloudera Hybrid Environments

Provide the ability to federate access to data and run workloads across both on-premises and cloud environments.

Platform Security:

Deliver **centralized security and governance** with hardened, out-of-the-box security.

Cloud Modernization Strategies - Lift-and-shift migration vs temporary burst to cloud

Learn about cloud modernization strategies.

Organizations employ two complementary strategies for hybrid cloud adoption: workload migration and cloud bursting. While traditional migration involves the permanent relocation of applications and datasets to the cloud for modernization, cloud bursting dynamically extends a private data center into a public cloud. This provides temporary, on-demand compute to handle demand spikes, scaling back down as capacity needs subside.

These two strategies co-exist. Migration is a long-term approach for modernizing to cloud-native workloads, whereas bursting provides immediate compute elasticity for workloads that are retained on-premises, bypassing physical hardware procurement cycles.

To enable native cloud-bursting capability, Cloudera is introducing Cloudera Hybrid Environments. This new architecture is designed to provide cloud agility while leveraging existing infrastructure. It facilitates dynamic workload movement—as distinct from planned migration—by natively discovering and accessing datasets on-premises, all while strictly adhering to established on-premises security and governance protocols.

Feature	Cloud Burst / Remote Access Strategy	Data Replication / Migration Strategy
Primary Goal	**Elastic scaling and temporary peak load handling.** Extend existing on-premises resources temporarily.	**Resilience, modernization, and geo-availability.** Move workloads closer to the cloud environment.
Data Residency	**Data stays on-premises.** Only processing or temporary application components move to the cloud.	**Data exists in two or more locations** (on-premise and cloud) simultaneously.
Workload Type	**Stateless or loosely coupled workloads.** Compute-intensive tasks, batch processing, web tier scaling (e.g., VDI, render farms).	**Stateful workloads.** Databases, file systems, applications requiring fast, local data access (e.g., ERP, CRM).
Latency Tolerance	**High tolerance for latency.** Performance depends heavily on the network connection to the on-premises data center.	**Low tolerance for latency.** Data must be accessible quickly by cloud applications (since the data is local to the cloud environment).

Feature	Cloud Burst / Remote Access Strategy	Data Replication / Migration Strategy
Implementation	Setting up **VPNs, dedicated network links** (e.g., AWS Direct Connect, Azure ExpressRoute), and configuring firewalls to allow remote access to on-premise databases.	Implementing **Data sync tools** (e.g., Cloudera Replication Manager), SAN replication, or specialized software-defined storage.
Cost Driver	**Network egress fees.** Moving large amounts of data from the cloud back to on-premise is expensive.	**Cloud storage costs.** Paying for duplicate storage space and the cost of the replication service itself.
Management Complexity	**Network management.** Ensuring a stable, low-latency, high-bandwidth connection between sites.	**Data consistency management.** Ensuring data integrity, conflict resolution, and synchronization lag between the two copies.
Disaster Recovery (DR)	**Poor/non-existent.** The cloud is only used for compute; if the on-premise data center fails, the application fails.	**Excellent.** The replicated data can be used to instantly spin up services in the cloud (Active-Passive or Active-Active DR).

What are Cloudera Hybrid Environments and Cloudera Hybrid Data Hubs?

Cloudera Hybrid Cloud deployments combine cloud-native elasticity, including provisioning and autoscaling, with the built-in capability to securely access datasets directly from an associated on-premises cluster.

This architectural design provides a native solution for temporarily extending an on-premises cluster, allowing organizations to burst workloads to the cloud seamlessly.

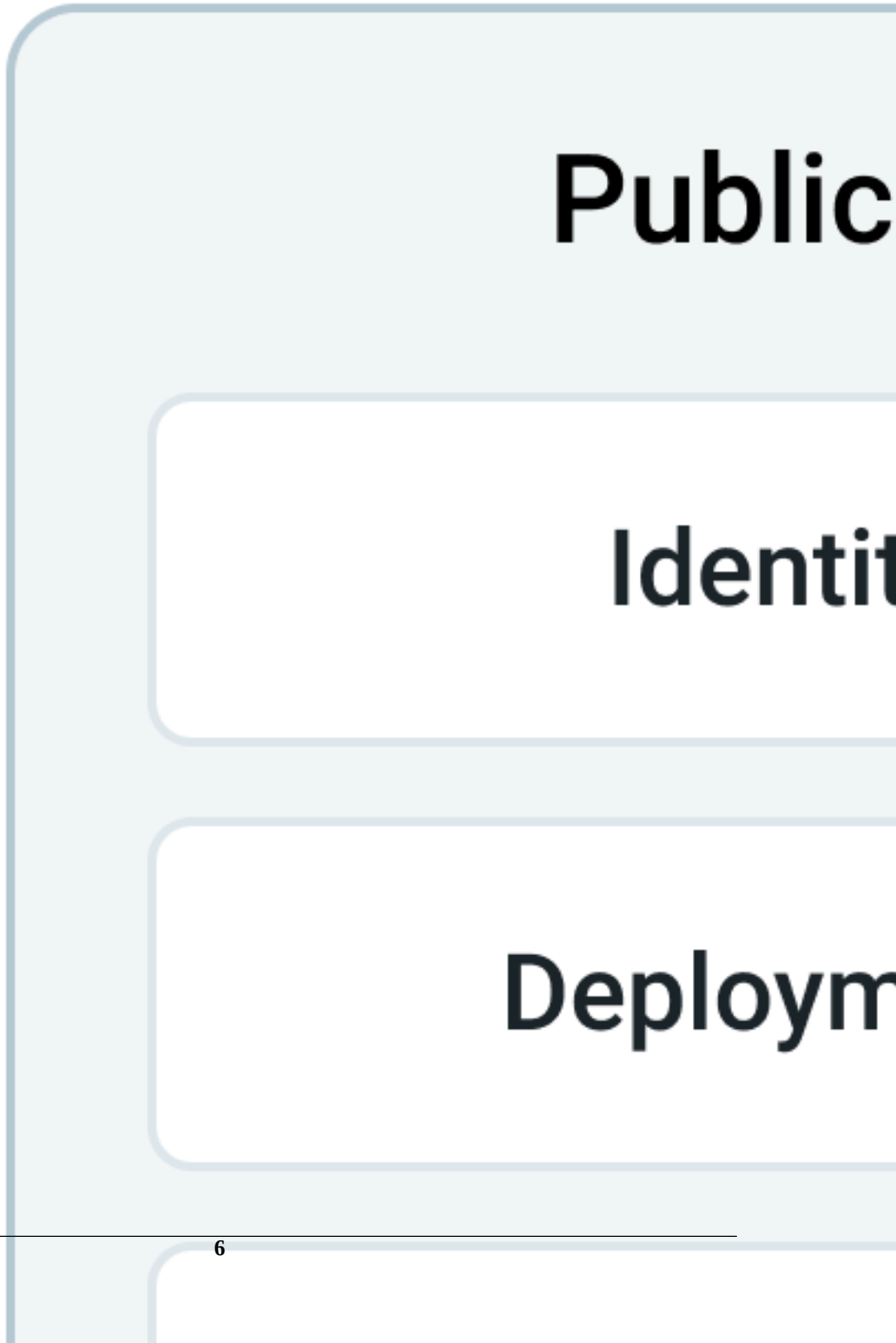
1. **Centralized Governance:** All metadata, access permissions, and governance rules remain centralized on-premises and are enforced consistently, whether the workload runs on-prem or in the cloud.
2. **No Data Migration:** You avoid the cost and complexity of application redesign and data migration.
3. **Dynamic Workload Movement:** You can dynamically "burst" workloads to the cloud to handle temporary needs without ever creating a second, out-of-sync copy of your data.

Besides cloud bursting, Cloudera Hybrid Environments and Data Hubs can also be utilized for the following:

- **Strategic Workload Isolation:** Maintaining critical workload SLAs by offloading additional workloads to the cloud.
- **Accelerating Software Development:** Creating instant development environments that leverage zero-copy data access from your on-prem source.

How Do Cloudera Hybrid Environments Work?

Cloudera Hybrid Environments use a novel **compute-only cloud replica** model based on **in-place data access**.


Public
Identit
Deploym

Workloads submitted to compute services in Cloudera Hybrid Data Hubs access data and metadata directly from an associated Cloudera on premises cluster.

This architecture relies on the following building blocks.

Unified Authentication

Implementing a **two-way Kerberos cross-realm trust** between the hybrid cloud and on-premises clusters. This enables centralized authorization and governance.

Metadata Synchronization

Associating Cloudera on premises cluster with the Cloudera Hybrid Data Hub cluster as its metadata, authorization, and governance context.

Network Connectivity

Ensuring stable, bi-directional network connectivity exists between the on-premises cluster and hybrid cloud environments to support **in-place data read/write operations** for active jobs.

Workload Portability

Maintaining a **unified runtime version** across both the Hybrid Datahub and the on-premises cluster for default workload portability.