

**CLOUDERA**

# Three Steps to Successfully Migrate to the Public Cloud

Streamline workload migration and burst to  
cloud with Cloudera Data Platform (CDP)



---

# Table of Contents

<b>A Smarter Approach to Realizing the Benefits of Public Cloud</b>	<b>3</b>
<b>Create a Strategic Plan to Move to Public Cloud</b>	<b>4</b>
Step 1: Evaluate Clusters and Workloads	5
Step 2: Enable Migration Activities	6
Step 3: Launch Production Workloads	7
<b>Create a Unified Hybrid Experience</b>	<b>8</b>
<b>Adopt a Strategic Approach and Learn from the Best</b>	<b>10</b>
<b>Take the Next Steps</b>	<b>11</b>

---

# A Smarter Approach to Realizing the Benefits of Public Cloud

Organizations continue to optimize how they gain value and insight from their data in order to make better business decisions. More often than not, they've done this by getting smarter about the public cloud.

While many organizations have enjoyed the initial fruits of cloud success, they've also had negative experiences such as spiraling cloud charges, unexpected security and governance risks, and performance.

In hindsight, the first wave of public cloud adoption was rushed. As a result, according to IDC, many organizations have backtracked on their cloud strategy and repatriated a number of their workloads, citing security, cost, performance, regulatory compliance, and complexity concerns.<sup>1</sup>

Organizations are instead leveraging a cloud-first strategy to achieve the flexibility and agility of public cloud.

The way forward starts with a strategic approach that carefully considers workload placement but also employs a data platform that delivers a seamless experience across hybrid environments, provides consistent security and governance, and ensures interoperability between clouds.

With that, you will realize the original benefits you've been looking for, including:

- Increased business agility
- Speed of deployment
- Enhanced workload elasticity
- Greater scalability
- Retirement of legacy hardware
- Pay-as-you-go pricing
- Improved control and management
- Consistent security and governance

## The Dangers of Lock-in: Cloud Concentration Risk

Cloud Concentration Risk is a real concern for global and regional regulators of the Financial Services Industry (FSI). But it is also a worry to other industries with regard to ensuring uninterrupted services to their customers.

That concern stems from an institution's over-reliance on one service provider to support key services.

To mitigate cloud lock-in, embrace open source solutions, industry standard formats and open architecture and open architecture that ensures interoperability between clouds and portability of data and applications across any platform.

For more information, read the article, [Minimizing Cloud Concentration Risk for Financial Services Institutions, Regulators and Cloud Service Providers](#)

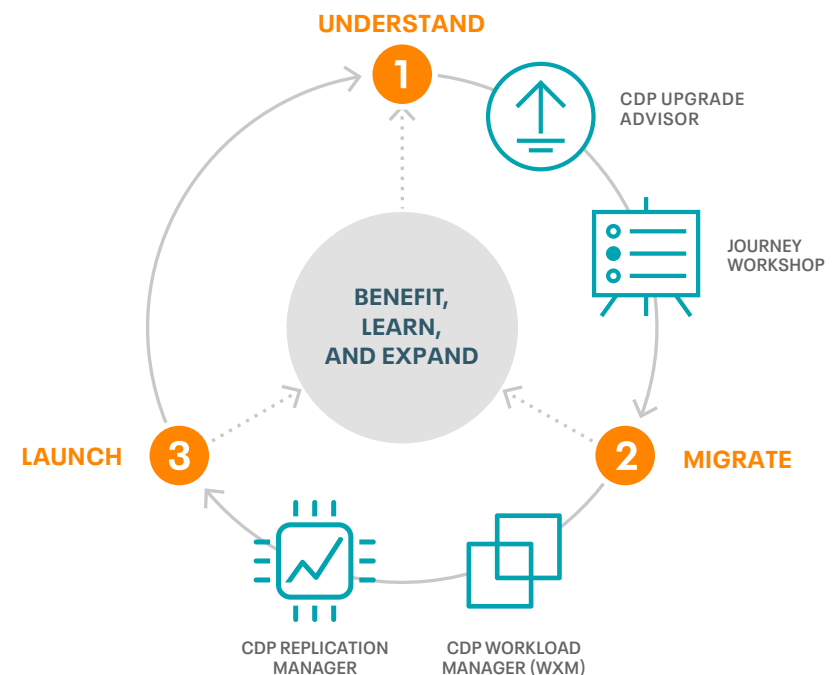
# Create a Strategic Plan to Move to Public Cloud

It is well known that modernizing a data platform is a complex process marred with challenges unique to how an individual organization operates. Cloudera Data Platform (CDP) reduces those complexities by optimizing the data lifecycle across both public cloud and on-premises with consistent data security and governance.

Cloudera de-risks the process with a customer-driven step by step approach that capitalizes on the collective expertise acquired across many industries. Best practices are enhanced with CDP automated tools so that your valuable resources are focused on the intricacies of your organization.

- **Step 1—Evaluate Clusters and Workloads:** CDP Upgrade Advisor is a self-service tool that evaluates your legacy assets to recommend upgrade paths, highlight risks, and explain why. Armed with that knowledge, the Journey Workshops provide a constructive expert forum to carefully consider optimal placement for both existing and new workloads.
- **Step 2—Enable Migration Activities:** CDP Workload Manager and CDP Replication Manager are tools that automate how you replicate or migrate your data and workloads to the public cloud.
- **Step 3—Launch Production Workloads:** Seamlessly extend to production, leveraging the unified data experience of CDP. This strategic and progressive plan lets you benefit and learn from small wins and expand into more sophisticated and value driven use cases.

## STEPS TO PUBLIC CLOUD MIGRATION



## Step 1: Evaluate Clusters and Workloads

Successful migrations require preparation and planning. Cloudera provides the comprehensive tools and expertise needed to understand, plan, and execute on their objectives.

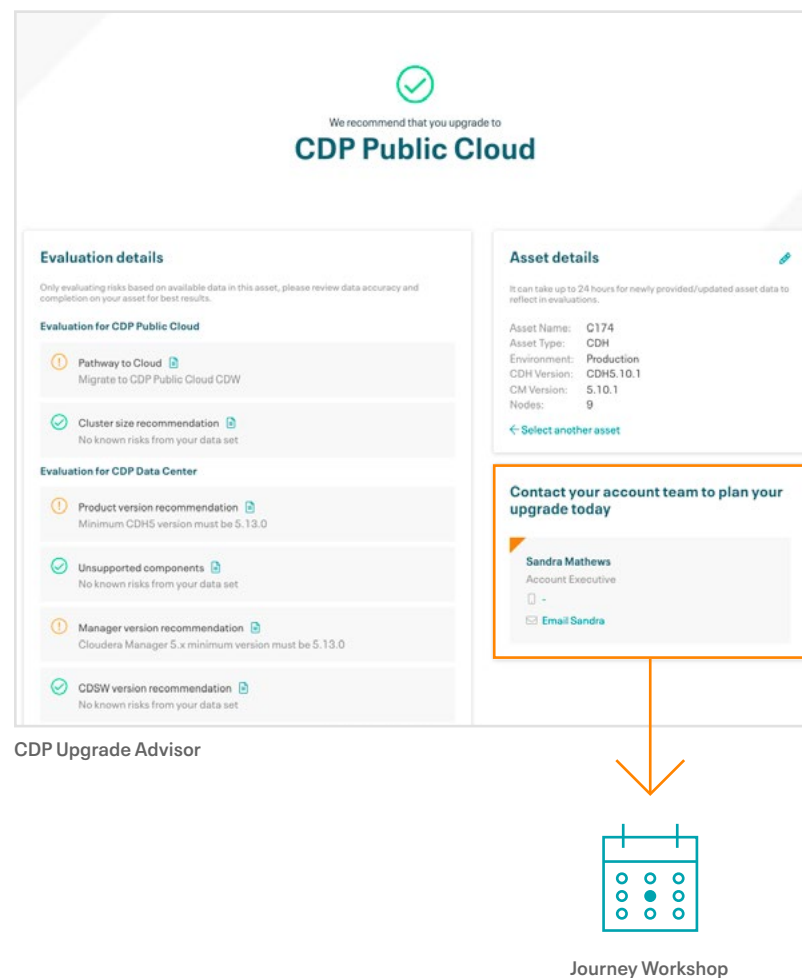
**CDP Upgrade Advisor** is a tool used to evaluate your cluster environment in preparation for a migration to CDP Public Cloud. Using diagnostic data, it determines the CDP readiness of your current CDH or HDP clusters and raises awareness of assets that may present risks during an upgrade. Running this utility is one of your first steps to adopting CDP, followed by an in-depth conversation of its results with your Cloudera account team.

**The Journey Workshop** leverages Cloudera's expertise while making the best use of your time and resources to identify suitable workloads for public cloud, such as:

- Batch or intermittent workloads with low cluster utilization rates.
- Elastic workloads that scale to handle volatile utilization and resource demands.
- Workloads that spur innovation but are not yet mission critical.

The Journey Workshop will prepare your team to successfully implement and launch your workloads onto CDP Public Cloud and recommend upgrade paths that often provide the quickest ROI, including:

- **Shift to Cloud:** Migrate existing on-premises workloads to the cloud.
- **Burst to Cloud:** Augment existing, on-premises workloads by extending them temporarily to the cloud.



## Step 2: Enable Migration Activities

Whether you are shifting existing on-premises workloads or bursting them to the cloud, implementing and launching workloads into production is an iterative process made simple with automated tools.

**CDP Workload Manager (WXM)** is Cloudera's intelligent workload management tool that enables:

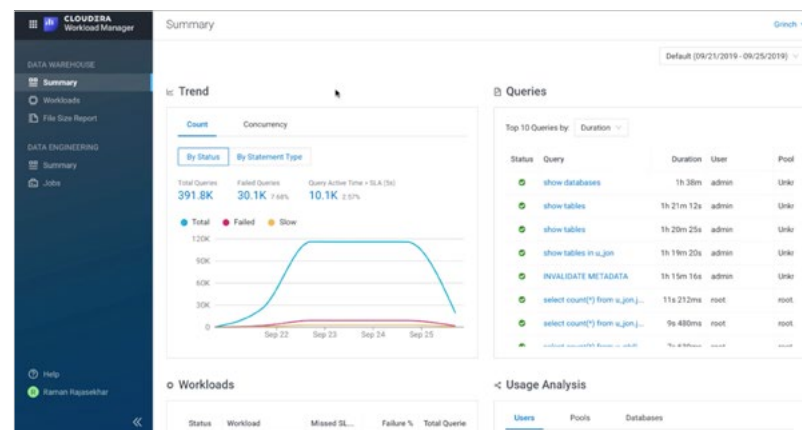
- Cloud bursting scenarios for Hive and Impala workloads.
- Evaluating the cloud readiness of a Hive or Impala workload.
- Generating a replication plan and capacity plan based on deeper analysis of the workload.

**CDP Replication Manager** is a centralized management tool for replicating and migrating Hive, Impala, and HDFS data and metadata between environments. It is a simple, easy-to-use, and feature-rich data movement capability, providing:

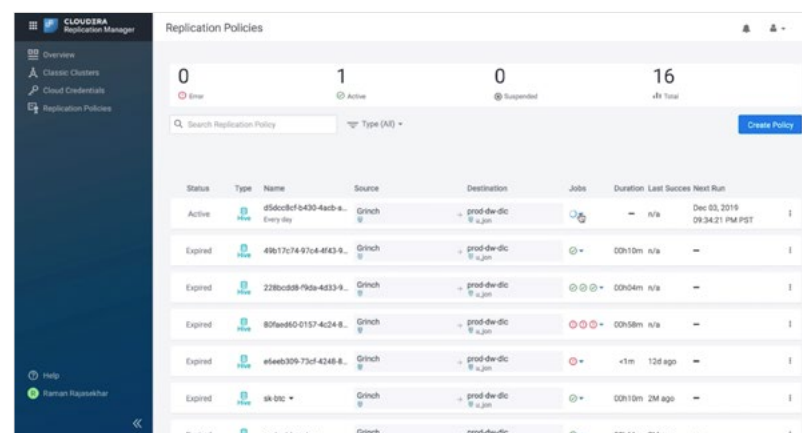
- Controlled data and replication.
- Integration with WXM to simplify workload migration to the public cloud.

For HBase workloads, use **Operational Database Data Hub cluster definitions**. **HBase replication** can be used as your data import method. Simply test your application in the new cluster and then enable HBase replication between the two clusters.

For stream processing and real-time analytics workloads, use **Data Hub cluster definitions** (Kafka, NiFi, and Flink) to create streaming clusters in the cloud. **Streams Replication Manager** can then be used to replicate Kafka data to the public cloud.



CDP Workload Manager



CDP Replication Manager

### Step 3: Launch Production Workloads

At this point in the process, workload identification and data migration is complete, you understand the complexities of your data environment, and you have a plan that mitigates the risks of migrating to public cloud because you started with lower risk workloads. Now you can seamlessly launch to production, leveraging the unified data experience of CDP.

Cloudera will work with you to move your workloads to CDP Public Cloud, whether that be to traditional clusters on Data Hub or through the containerized experiences of Cloudera Machine Learning, Data Engineering, or Data Warehouse.

Much preparation went into having a clear and detailed understanding of your workloads as they relate to use cases, datasets, schemas, component architecture, dependent systems, and SLA requirements. Once the migration is complete, you will be able to test, evaluate, and iterate to optimize the workloads in public cloud. An engagement with Cloudera progresses on that work by helping with required code changes along with testing and validating the execution of the migrated workloads.

The mechanism by which all this happens is CDP. You are ensured a seamless move to the public cloud because of the common set of tools and processes provided through CDP.

To identify new workloads with CDP Public Cloud, Cloudera offers Accelerated Cloud Execution (ACE) program that will help get you from planning to pilot for cloud use cases. The program includes a series of workshops for Data Flow, Data Warehouse and Machine Learning to understand how CDP Public Cloud can drive your business objectives. Once you've identified the target workloads, Cloudera ACE will support delivery of new workloads to public cloud.

### A CDP Public Cloud Migration Success Story

A US-based telecommunications company had deployed workloads using a legacy data platform on IaaS. However, with that configuration, they were unable to get the benefits of a true public cloud deployment.

Leveraging the hybrid cloud data architecture of Cloudera Data Platform (CDP), the organization was able to make significant savings in cloud storage and compute resource costs, amounting to 40% on an annual basis.

The shift from on-premises workloads to CDP Public Cloud resulted in the following:

- Improved cost-effectiveness and control.
- More complex multi-stage and multi-tenant use cases and pipelines.
- Increased security and governance through a unified data experience.
- Future proofed data analytics by leveraging open source architecture.



Annual savings in cloud storage  
and compute resource costs



---

# Create a Unified Hybrid Experience

While public cloud is a great first step on your path to modernizing your data platform with CDP, most enterprises currently (or plan to) distribute their data workloads across hybrid environments, including on-premises, private cloud, and public cloud.

CDP, the world's first enterprise data cloud, ensures success with hybrid cloud because it is easy to deploy, manage, and use. It separates data management from infrastructure so that you can move data and apps from one environment to another without rewriting applications and retraining personnel.

Extending CDP from your data center to public cloud provides a seamless and end-to-end data cloud that provides the following:

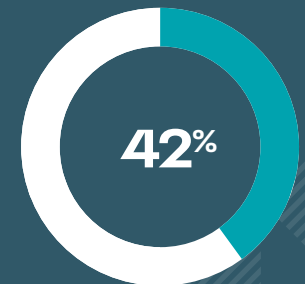
- Reduce infrastructure, operational, and administrative costs.
- Scalability for more users, use cases, and workloads.
- Improved security and governance.
- A future-proof architecture for data and analytics.

CDP creates a true hybrid environment that provides equivalent data management and cloud-native analytic experiences across your data center and the public cloud of your choice, including AWS, Azure, and GCP (coming soon).

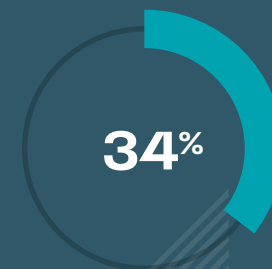
## The Future is Hybrid—for Most<sup>2</sup>



Enterprises moving toward a hybrid environment.



Enterprises considering or developing a hybrid IT strategy.



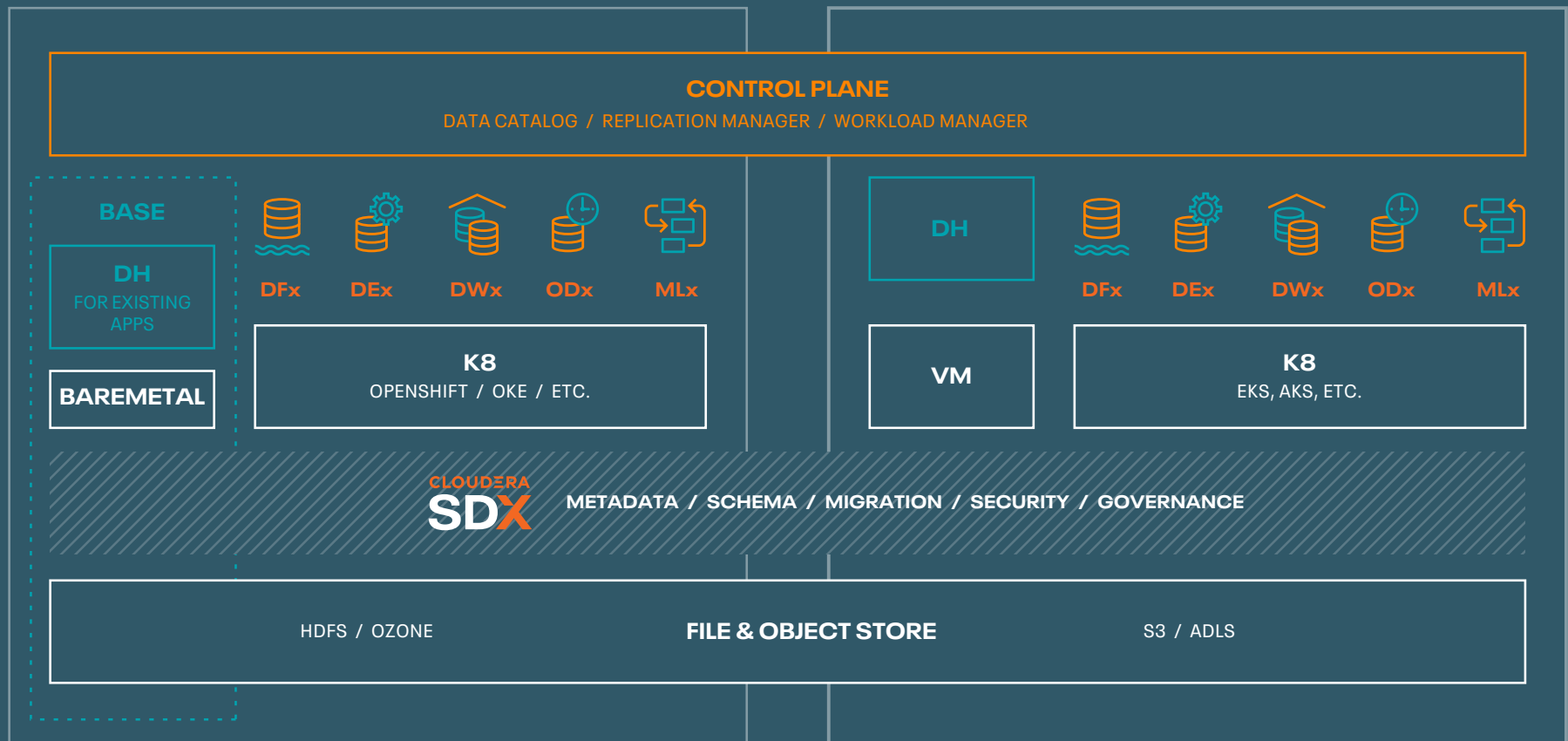
Enterprises actively executing a hybrid IT strategy.



# CDP—A Unified Hybrid Experience

## CDP PRIVATE CLOUD

## CDP PUBLIC CLOUD



---

# Adopt a Strategic Approach and Learn from the Best

Cloudera Professional Services and Training (PS&T) has worked with over 1,000 organizations to drive value from their data. We've seen common themes in our customers' challenges as they modernize their platforms and, by helping them find solutions, have learned how to best help companies in their journey to public cloud.

From that experience, we've developed CLOUDSMART, our unique approach that streamlines how you move on-premises workloads to the public cloud.

Our structured process and automated tools enable your organization to learn with quick wins from which to expand into more complex use cases.

We strive to de-risk the process with faster time to value while fostering the organizational confidence and expertise

you need for repeatable success. PS&T helps you lay the groundwork for a hybrid environment and accelerates your time to value through CDP. We ensure that your organization is informed and well prepared for success with access to the right expertise when you need it.

Our comprehensive training portfolio provides a mix of OnDemand and instructor-led training, including many free courses to help get your team ready for CDP.

## The Advantages of CDP —The World's First Enterprise Data Cloud

**Hybrid & multi-cloud:** Operates across all major public clouds and the private cloud with a public cloud experience everywhere.

**Multi-function analytics:** Integrates data management and analytic experiences across the data lifecycle for data anywhere.

**Secure & governed:** Delivers security, compliance, migration, and metadata management across all environments.

**Open platform:** Open source, open integrations, extensible, and open to multiple data stores and compute architectures.

[See this video](#) to learn how CDP powers data-driven decision by easily, quickly, and safely connecting and securing the entire data lifecycle.

---

# Take the Next Steps

Cloudera can help you simplify your move to CDP on public cloud and go from pilot to production quickly, painlessly, at lower cost, and with peak performance that is underpinned with consistent data security, governance, and management.

If you are considering migrating to the public cloud, a CDP Journey Advisor workshop can be scheduled with your account team to help you determine which path is best for your organization.

## Your Action Steps

- Test drive [Cloudera Data Platform Public Cloud](#): Begin your interactive guided tour and experience the future of the enterprise data cloud.
- Discover your custom upgrade path to CDP with the [Upgrade Advisor](#) on My Cloudera.

## Learn More

- [Eliminate the pitfalls on your path to public cloud](#)
- [Top 10 Ways to Master Performance for Your Cloud Migration](#) (Forrester Research)
- Skim the public cloud environment prerequisites for [AWS](#) or [Azure](#) environments
- [Extend Your On-premises Streaming to the Public Cloud](#)

---

### Sources

<sup>1</sup> IDC's CloudPulse Q1, June 2019

<sup>2</sup> 451 Research, "Voice of the Enterprise—Cloud, Hosting & Managed Services, Vendor Evaluations", 2019

© 2021 Cloudera, Inc. All rights reserved. Cloudera and the Cloudera logo are trademarks or registered trademarks of Cloudera Inc. in the USA and other countries. All other trademarks are the property of their respective companies. Information is subject to change without notice. 4264-001 December 23, 2020

[Privacy Policy](#) | [Terms of Service](#)

# CLOUDERA

## About Cloudera

At Cloudera, we believe that data can make what is impossible today, possible tomorrow. We empower people to transform complex data into clear and actionable insights. Cloudera delivers an enterprise data cloud for any data, anywhere, from the Edge to AI. Powered by the relentless innovation of the open source community, Cloudera advances digital transformation for the world's largest enterprises.

Learn more at [cloudera.com](https://cloudera.com)

US: +1 888 789 1488

Outside the US: +1 650 362 0488