CLOUDERA

SOLUTION BRIEF

A Modern, Cloud-Native Data Platform for On-Prem

# CLOUDERA DATA PLATFORM (CDP) PRIVATE CLOUD

**What differentiates CDP Private Cloud?**

- Elastic analytics running on a containerized compute cloud
- Highly scalable object storage separated from the compute layer
- An "always on" secure and governed data lake that combines data across the enterprise
- Consistent management services across all clouds, enabling hybrid cloud

Cloudera Data Platform (CDP) Private Cloud is a next-generation, hybrid data platform with cloud-native, self-service analytic experiences bringing the speed, scale, and economics of the public cloud to your data center. CDP Private Cloud enables organizations to deliver powerful data analytics and machine learning in minutes with 50% less data center infrastructure. At its core, CDP Private Cloud brings cloud architecture to on-premises deployments.

## Business Trends and Challenges

With the speed of today's business, users are expecting information technology (IT) departments to instantly provision resources and scale on-demand for a variety of highly specialized data needs. This includes both data practitioners (ranging from data engineers to BI analysts and data scientists) as well as technology and operations related roles (covering data architects, infrastructure admins, and infosec/compliance managers). This shift in IT consumption has led to innovations in data architecture. Instead of "shadow IT" – point products that bypass IT and create a security risk for the company – organizations require an enterprise data platform that both centralizes control and provides agility and flexibility to provision resources for custom data workloads.

Moreover, as many organizations come to terms with the fact that public cloud options are not always the best fit or most cost effective option, they realize the need for a data platform that has the capabilities of the cloud but can be deployed on-premise, enabling agility for departments, groups, and lines-of-business, while facilitating efficiency and security within IT.

To better understand how it all fits together, a simplified view of the CDP architecture consists of four constructs:

**Management Console** – A single pane of glass to manage one or more environments and the services that run within each environment.

**Environment** – A logical encapsulation of a customer network and the services that run within that network.

**Cluster** – A distributed computing service that has access to shared data lake and runs on either virtual machines (Cloudera Data Hub) or containers (Cloudera Analytics Experiences)

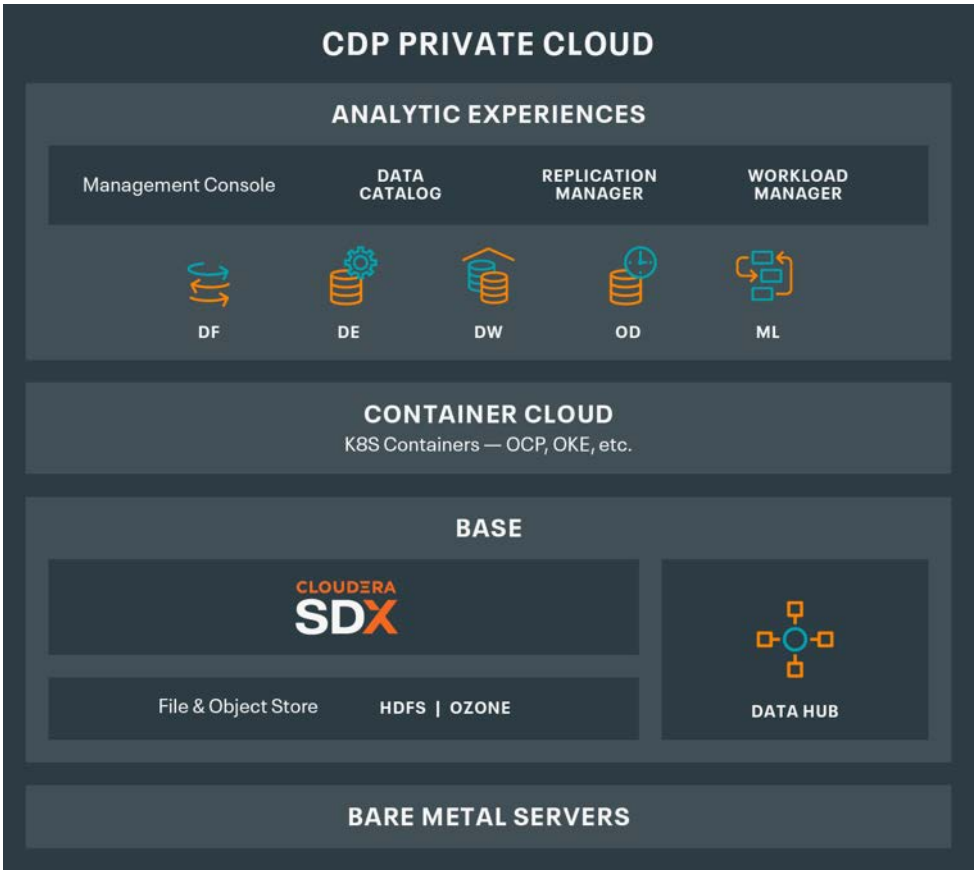
**SDX** – The data access control layer that sits on top of the backend object store and provides coherent data security and governance for all the applications running within the environment.

**Containers** - A software container is a means of packaging up application code and everything it needs to run properly in any environment, from a developer's laptop, to a bare-metal server, to a private or public cloud.

Because they are lightweight and immutable, containers at scale can help organizations manage hundreds or even thousands of containerized workloads in production.

**Kubernetes** - an open source platform for container orchestration. Kubernetes automates the "deployment, scaling, and management of containerized applications." This declarative form of automation simplifies operational complexity and reduces manual effort.

CDP Private Cloud offers data warehouse and machine learning services, a data hub service for building custom business applications, and a unified control plane to manage infrastructure, data, and analytic workloads across hybrid and multi-cloud environments. This includes consistent data security, governance, and control with Shared Data Experience (SDX) to safeguard data privacy, ensure regulatory compliance, and prevent cybersecurity threats. All of this is built on an open source foundation that avoids vendor lock-in and accelerates speed to insight.




CDP Private Cloud will have a range of capabilities to enable a multitude of analytics experiences (which run on Kubernetes), including:

- Cloudera Data Warehouse (CDW) – delivers self-service analytics on massive amounts of data to thousands of users without compromising cost, speed or security
- Cloudera Machine Learning (CML) – optimizes ML workflows for deploying, serving and monitoring models.
- Cloudera Data Engineering (CDE) – (coming soon) features an all-inclusive data engineering toolset to orchestrate and automate complex data pipelines securely at any scale.
- Cloudera DataFlow (CDF) – (coming soon) provides real-time streaming data analysis at high volume and high scale.

The advantage of CDP architecture is that it is modular, and the same constructs can be transposed from a private data center to the public cloud and everything in between for a seamless hybrid experience.

10X

Faster time to value.

On-board in minutes not weeks.

100%

Isolation of tenants.

Eliminate noisy neighbours & meet SLAs

50%

Lower data center infrastructure costs.

## Business and Technology Benefits

CDP Private Cloud offers a number of business and technological benefits for enterprises. These include:

| Technology Benefit | Business Impact |
|---|---|
| Agile, scalable, multi-tenant infrastructure | Reduce data center costs as much as 50% by drastically improving operational efficiency of clusters, utilizing compute infrastructure, and eliminating data replication. |
| Complete tenant isolation | Easily meet the SLAs for mission-critical workloads with custom environments eliminating resource bottlenecks. |
| Self-service provisioning of analytics | Delegate responsibility to those closest to active workloads while still maintaining centralized oversight, removing the bottleneck on IT to enable faster business value |
| Flexible workloads deployment (public, private and true hybrid cloud) | When demand for computing capacity spikes, deploy and move applications across on-premises, private cloud, and public cloud environments as needed. With a consistent user and IT admin experience across environments, organizations can stay focused on their business' data strategy, not on infrastructure or incompatibility challenges. |
| Cloudera Runtime is 100% open source | Prevents vendor lock-in and ensures interoperability with a broad range of third party applications |
| Security by design via Cloudera's SDX layer | Ensures consistent data security, governance, and control across the data lifecycle and across all environments while mitigating risk and costs |

CDP Private Cloud delivers a cloud-native hybrid data architecture, letting organizations deploy data and analytics on-premises with the agility, flexibility and cost efficiency of cloud architectures.

Compared to legacy on-premises data platforms, CDP Private Cloud:

- Is **easy to operate**, allowing new use cases to be on-boarded in minutes rather than weeks, resulting in a faster time to value.
- Delivers **rock-solid service**, through complete tenant isolation thereby eliminating resource contention and delivering consistent end user performance.
- **Reduces overall costs** through more efficient use of existing data center infrastructure.

Research shows that organizations will continue to maintain their data center infrastructure in addition to public cloud adoption. In short, traditional data centers will not go away because of Azure, AWS and Google.

**85%**

of enterprises continue to rank hybrid cloud as the "ideal" IT operating model.

**73%**

of respondents are migrating applications away from the public cloud back to on-premises infrastructures.

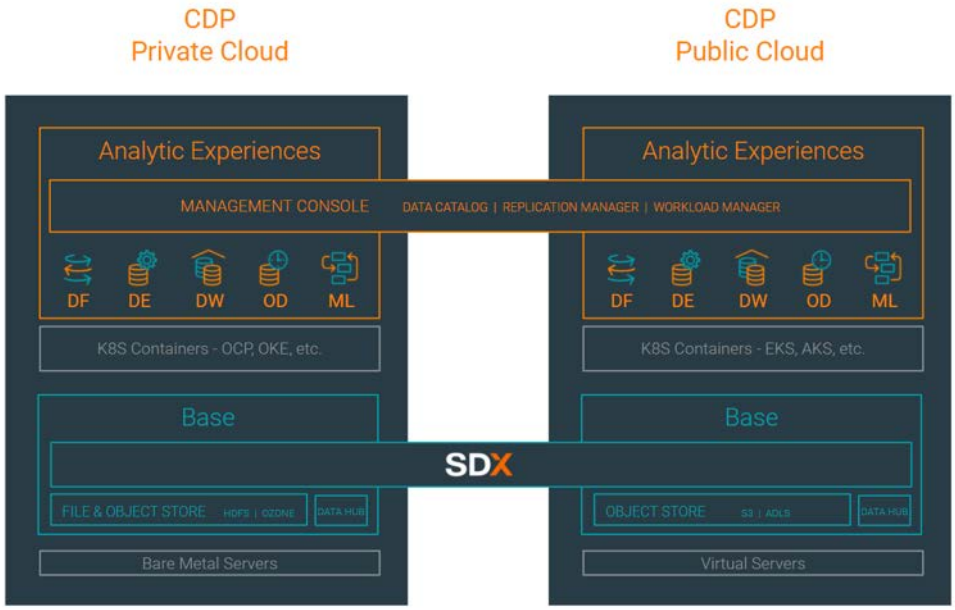
**49%**

cited hybrid cloud as the model meeting all their needs.

## The Future of Data Platforms: Hybrid Cloud

A number of studies conducted by 451 Research confirm a similar trend. Hybrid/multi-cloud is now emerging as the predominant strategic posture for managing IT and business transformation. In fact, 62% of enterprises today are pursuing a hybrid IT strategy.

These findings are echoed by IDC, which predicts that by 2022, over 90% of enterprises worldwide will rely on a mix of on-premises/dedicated private clouds, multiple public clouds, and legacy platforms to meet their infrastructure needs.




CDP Private Cloud is designed to take advantage of today's hybrid environments and allow organizations to effectively utilize their existing on-premises infrastructure while effortlessly bursting into public cloud when required.

**Services to enable your path to CDP Private Cloud**

In addition to **SMART**UPGRADE, here are additional services available from Cloudera to help with the transition to CDP Private Cloud:

- **SMART**ARCHITECTURE - Begin defining a hybrid strategy and architecture to modernize your analytics with CDP
- **SMART**DEPLOY - Quickly install and secure your first cluster on CDP Private Cloud based on best practices
- **SMART**PRIVATE - Begin your analytics journey with secure platform deployment on PVC Experiences and pilot workload on CDW or CML

## Paths from Traditional Big Data Platforms

As customers migrate from legacy platforms, they can choose either an in-place upgrade, leveraging existing infrastructure, or a migration to a new environment. In order to prepare for an upgrade or migration, enterprise data platform migrations have to be carefully planned, executed and supported to avoid business disruptions or SLA violations.

Cloudera's SMARTUPGRADE service enables customers to move to CDP Private Cloud from existing Cloudera Distributed Hadoop (CDH) and Hortonworks Data Platform (HDP) deployments with minimal risk and maximum confidence.


Cloudera helps organizations by offering necessary skills, expertise and processes to supplement customer teams during migrations, resulting in reduced disruption and downtime of critical data analytics. Cloudera contributes deep experience and best practices from professional services, support, and engineering to ensure a successful migration to CDP. High-level activities include the following:

- Check cluster health and make optimization recommendations.
- Design, install and secure new CDP Private Cloud cluster.
- Plan and execute migration or upgrade (workloads, data, users and security).
- Application testing and refactoring.
- Ensure production operation readiness.
- Monitor baseline performance and resolve any issues.
- Manage documentation and knowledge transfer.
- Understand future use cases and open issues.

PRICING METRICS

**Cloudera Compute Unit** - One physical core and 8GB of RAM and addressed storage (TB) under management.

Pricing assumes typical server configuration of 32 cores, 256GB RAM, and 96TB addressed storage.

CDP Private Cloud pricing reflects business-level support.

---

**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

## CDP Private Cloud Pricing

With Base Edition or Plus options, CDP Private Cloud is cloud-native data management for lightning-fast operations, requiring up to 50% less compute and storage costs, delivering analytics that both data engineers and data scientists will love. (Note that CDP Data Center is now CDP Private Cloud Base Edition.)

CDP Private Cloud Base Edition offers bare metal workloads and cluster management updated with the latest open-source technologies. It features SDX security, governance, and metadata; HDFS file storage, and massively scalable Ozone object storage.

Plus Edition includes Base Edition, as well as easy-to-use containerized machine learning and data warehousing analytics and a hybrid management control plane for a better user experience and lower data center costs.

| Base Edition | Plus Edition |
|---|---|
| Enterprise analytics and data management, optimized for bare metal and ready for private cloud | Next-generation data analytics experiences, running on a container cloud with fast provisioning and scaling |
| Includes CDP Private Cloud Base installable software; SDX for security, governance, and metadata; HDFS file storage and Ozone object storage. | Includes Base Edition, containerized machine learning and data warehousing analytics running on Kubernetes, and a hybrid cloud management control plane. |
| Annual subscription, priced per Cloudera Compute Unit and storage (TB). Refer to pricing page on Cloudera website for details. | Annual subscription, priced per Cloudera Compute Unit and storage (TB). Refer to pricing page on Cloudera website for details. |

## References


1. Nutanix, The Nutanix Enterprise Cloud Index 2019, https://www.nutanix.com/enterprise-cloud-index.
2. 451 Research, Cloud Trends in 2020, https://go.451research.com/2020-mi-cloud-trends-year-of-complexity-and-its-management.html
3. IDC, IDC Expects 2021 to Be the Year of Multi-Cloud as Global COVID-19 Pandemic Reaffirms Critical Need for Business Agility, 31 March 2020, https://www.idc.com/getdoc.jsp?containerId=prMETA46165020


CLOUDERA

Cloudera Inc. 5470 Great America Pkwy, Santa Clara, CA 95054 USA cloudera.com