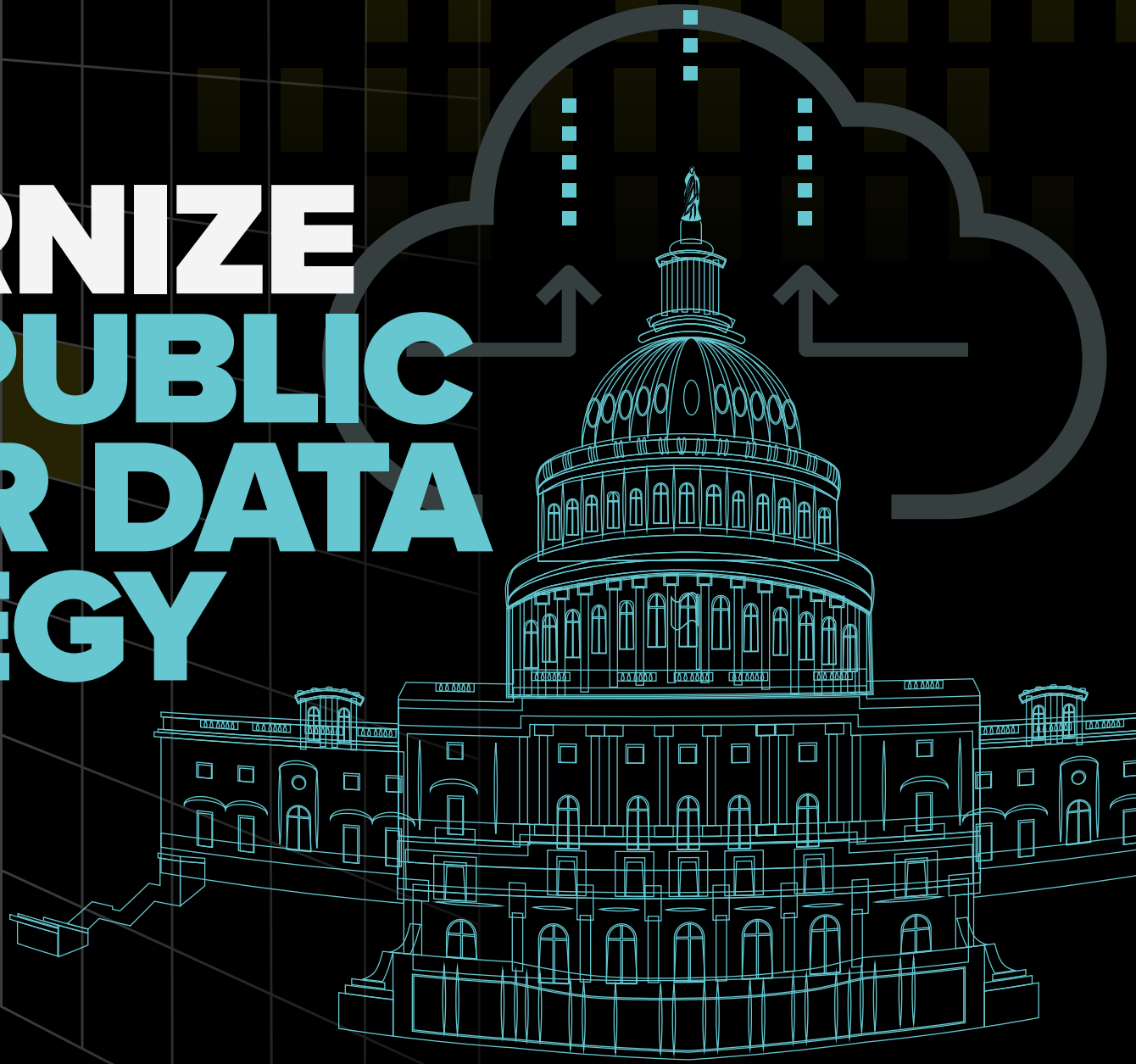




How to

MODERNIZE YOUR PUBLIC SECTOR DATA STRATEGY



Red Hat

CLUSTERA

The Transformative Power of Private Cloud in the Public Sector

Federal government agencies must modernize their data strategies in order to better execute their respective missions, not just for today but for the future. Tight budgets and aging IT infrastructure hamper their pursuit of this critical goal. Naturally, federal agencies – like so many other organizations – are looking to the cloud to help solve their most pressing business and technology challenges. It promises the agility, flexibility, scalability, and efficiency they need to collect and operationalize data from a vast number of sources and turn it into actionable, accessible insights.

The public sector operates under strict compliance and security requirements, however, and its collective mission is too important to risk. How do federal agencies harness the advantages of an enterprise data cloud – powerful modern IT infrastructure and potential cost efficiencies, among other benefits – in a secure, compliant, and reliable manner?

The Answer:

Cloudera Data Platform (CDP) Private Cloud on OpenShift.

CLOUDERA



Public Sector Dynamics – Data and the Cloud

Federal agencies are moving to the cloud to modernize their legacy IT infrastructure and transform how they discover, process, model, analyze, act on, and store massive amounts of data. They're doing so with very specific requirements about data sovereignty – and with increasing urgency to lead the implementation of new data initiatives.

\$3.8 billion

One industry report projects that federal government spending on cloud will increase by \$3.8 billion between 2019 and 2024, reaching an estimated \$9.1 billion at the end of that five-year span.

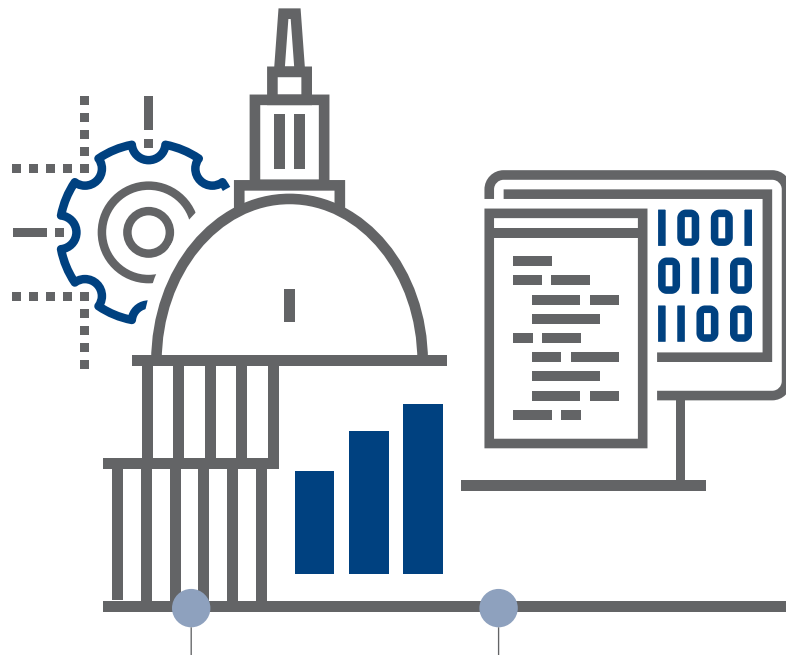
2x

Research firm Gartner has predicted that government agencies will adopt private cloud at twice the rate of public cloud through 2021. The firm attributes this to concerns around security and privacy, a lack of features, and vendor lock-in. Gartner also notes that data sovereignty remains a mandatory requirement for many public agencies.

149 zettabytes

Meanwhile, data is growing by an order of magnitude every year. Statista projects the total amount of data in the world will hit 149 zettabytes in 2024, up from 59 ZB in 2020 and more than 3.5 times the 41 ZB of data “created, captured, copied, and consumed” in 2019. Federal agencies must contend with similar trajectories of data growth given the sheer scope of their constituencies and missions.

<https://www.statista.com/statistics/1175706/worldwide-edge-computing-market-revenue/>



Mission-Critical Business Challenges for Federal Agencies

The public sector faces some important challenges as it looks to modernize its data strategy while ensuring long-term agility, accessibility, flexibility, and efficiency. They include:

Exponential data growth: 82% of feds agree that the amount of streaming data their agency manages is on the rise, according to MeriTalk research. This is essentially an irreversible trend: Agencies must prepare for continuously massive data growth.

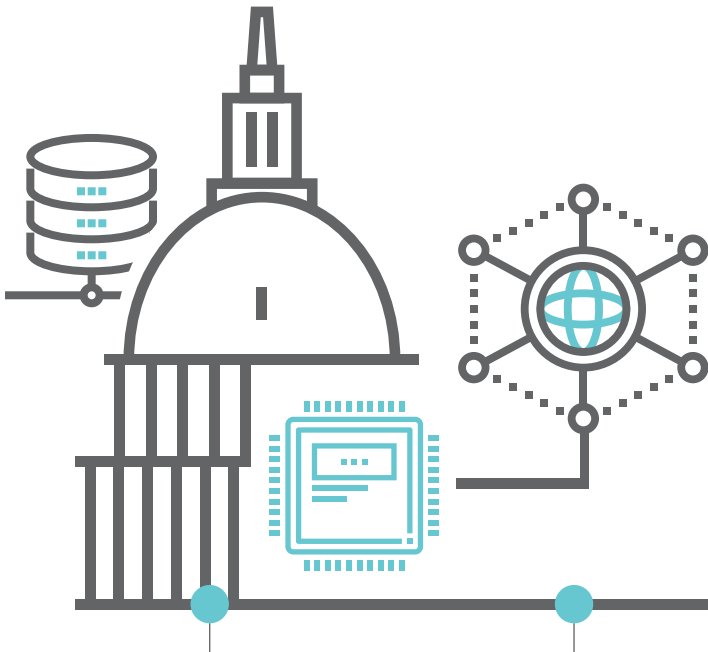
Efficient data storage at scale: Agencies must figure out ways to effectively store ever-increasing volumes of data.

Accessible data: The public sector must make the right data available in the right place at the right time, whether for internal use, citizen consumption, or other reasons.

The need for speedier insights: Analytics engines must be able to deliver actionable insights fast. 85% of feds agree that the ability to derive insights from streaming data is the key to modern decision-making, according to MeriTalk.

Regulatory compliance: There is simply no shirking this obligation in the public sector. MeriTalk found that “ensuring compliance across environments” was the #1 challenge for federal agencies in managing security through the data lifecycle.

Data silos: Too many one-off or “point” solutions produces data siloes that hinder accessibility and transparency.



Key Technical Challenges for Federal IT and Data Teams

Legacy infrastructure exacerbates these mission-critical challenges, as do multi-tenant environments, leading to a series of issues for federal IT pros.

“Noisy neighbors”: Demand spikes for pooled resources can negatively impact other workloads running in multi-tenant or shared clusters, potentially causing performance problems and missed SLAs.

Cluster sprawl: New clusters get spun up for new applications to avoid impacting existing workloads, but this also causes data siloes.

Lack of elasticity: Legacy infrastructure is often brittle, with assigned compute and storage nodes for every workload. Moreover, agencies must over-provision hardware to plan for future needs – including exponential data growth – but this infrastructure sits idle for long stretches of time

Complex upgrades: Needed upgrades to the cluster can become a multi-month process in to properly plan for the needs of different workloads.

Time to value: One reason that IT departments over-provision: Because spinning up new resources can be a multi-week or multi-month project. This also causes shadow IT as different departments acquire and implement their own solutions.

Federal agencies need a platform with the modern technical capabilities required to solve these mission-critical challenges.

A Rock-Solid Foundation for a Robust Private Cloud: Red Hat OpenShift Container Platform

One big misunderstanding about containerization and other cloud-native technologies is that they only run in public cloud. They can run pretty much anywhere, including in hybrid cloud, private cloud, and on-premises environments.

Cloudera Data Protection (CDP) Private Cloud runs on top of Red Hat OpenShift, an enterprise-grade Kubernetes platform for container orchestration. This ensures maximum portability and security across on-premises infrastructure, hybrid cloud, and multi-cloud environments.

44%

OpenShift has 44% of the container market share, and is one of the most trusted hybrid cloud platforms for managing containers and Kubernetes in production.



Months become hours:

With a mix of OpenShift and other solutions, the Government of British Columbia and its BC Developers Exchange cut the time to launch new services from 4-6 months for a single project to a matter of hours.

Containers and Kubernetes, defined



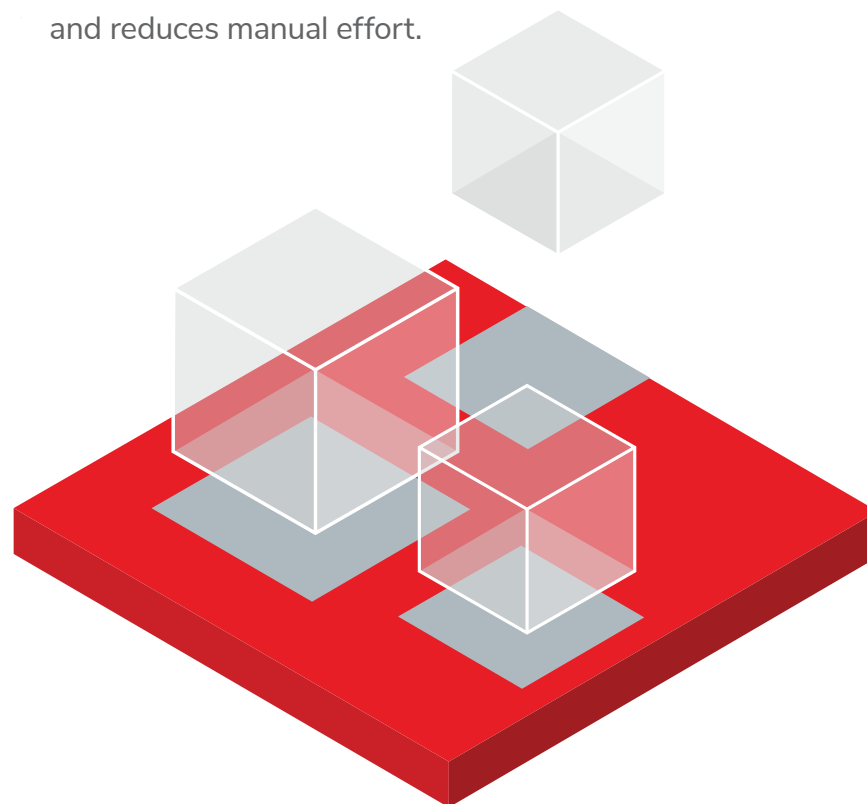
Containers

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, running containers at scale can mean you are managing hundreds or even thousands of containerized workloads in production.

Kubernetes

An open source platform for container orchestration. In its own words, Kubernetes automates the “deployment, scaling, and management of containerized applications.” This declarative form of automation simplifies operational complexity and reduces manual effort.



Solving the Data Management Challenge: The Enterprise Data Cloud

Federal agencies need an enterprise data cloud that helps address maximize the value of the growing data stores while being practical about budget, compliance, and security. That intertwined mix of business and technical challenges facing federal agencies require meaningful solutions that help agencies:

- Process and store massive amounts of data.
- Stay prepared for the continuous exponential growth of that data.
- Provide anytime, anywhere access to that data in a secure, compliant manner.
- Apply data science to unearth new opportunities that enhance their missions.
- Speed up analytics engines to provide those insights much faster than in the past.
- Deploy and move applications across on-premises, private cloud, and public cloud environments as needed.
- Leverage open source and open standards to enhance application portability, reduce costs, and minimize vendor lock-in.

CDP Private Cloud, supported by Red Hat OpenShift, delivers the cloud-native, open source enterprise data cloud that meets the needs of today's federal agencies.



Cloudera Data Platform (CDP) Private Cloud

Don't let legacy infrastructure and budget constraints impede your mission. CDP Private Cloud is designed with a powerful hybrid architecture that separates compute and storage for maximum agility, cost efficiency, ease of use, and resource optimization.

CDP Private Cloud Delivers

10x

faster deployments of analytics and machine learning services than traditional data management solutions.

100%

tenant isolation, solving the noisy neighbor problem and ensuring that you can the SLAs of your mission-critical workloads.

50%

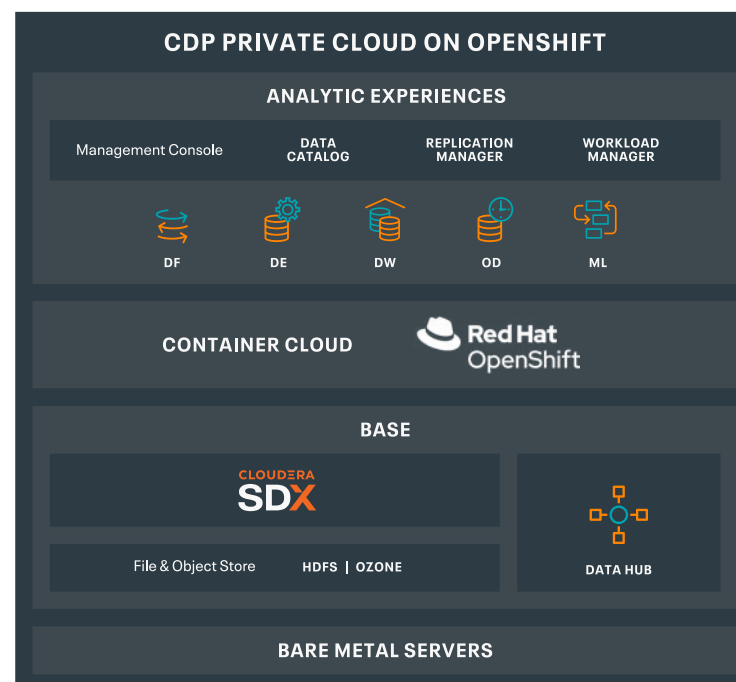
reduced datacenter costs by drastically improving efficiency and utilization of your compute infrastructure and eliminating data replication

0

Zero vendor lock-in, thanks to the open source Cloudera Runtime, interoperability with a wide range of third-party services, and the portability of workloads across hybrid cloud, multi-cloud, and on-premises environments.

CDP Private Cloud is the enterprise data cloud for your complete data lifecycle, solving the business and IT challenges of the public sector today and in the future.

- Speed up time-to-value with a cutting-edge data analytics platform from edge to AI:
 - Deliver self-service analytics on massive amounts of verified data to thousands of users without compromising cost, speed or security.
 - Quickly deploy and monitor ML models with optimized workflows.
 - Orchestrate and automate complex data pipelines securely and at any scale with a rich data engineering toolset.
 - Analyze massive volumes of real-time streaming data.
- Quickly onboard new use cases as they're needed and run them anywhere with OpenShift's robust portability.
- Reduce your costs with a subscription-based cloud model that better utilizes infrastructure – no more overprovisioning.
- Reduce or eliminate data siloes – no more point solutions to bypass slow implementations.
- Maintain complete oversight over governance, compliance, and security with powerful control and audit capabilities.



The Best of Both Worlds for the Public Sector

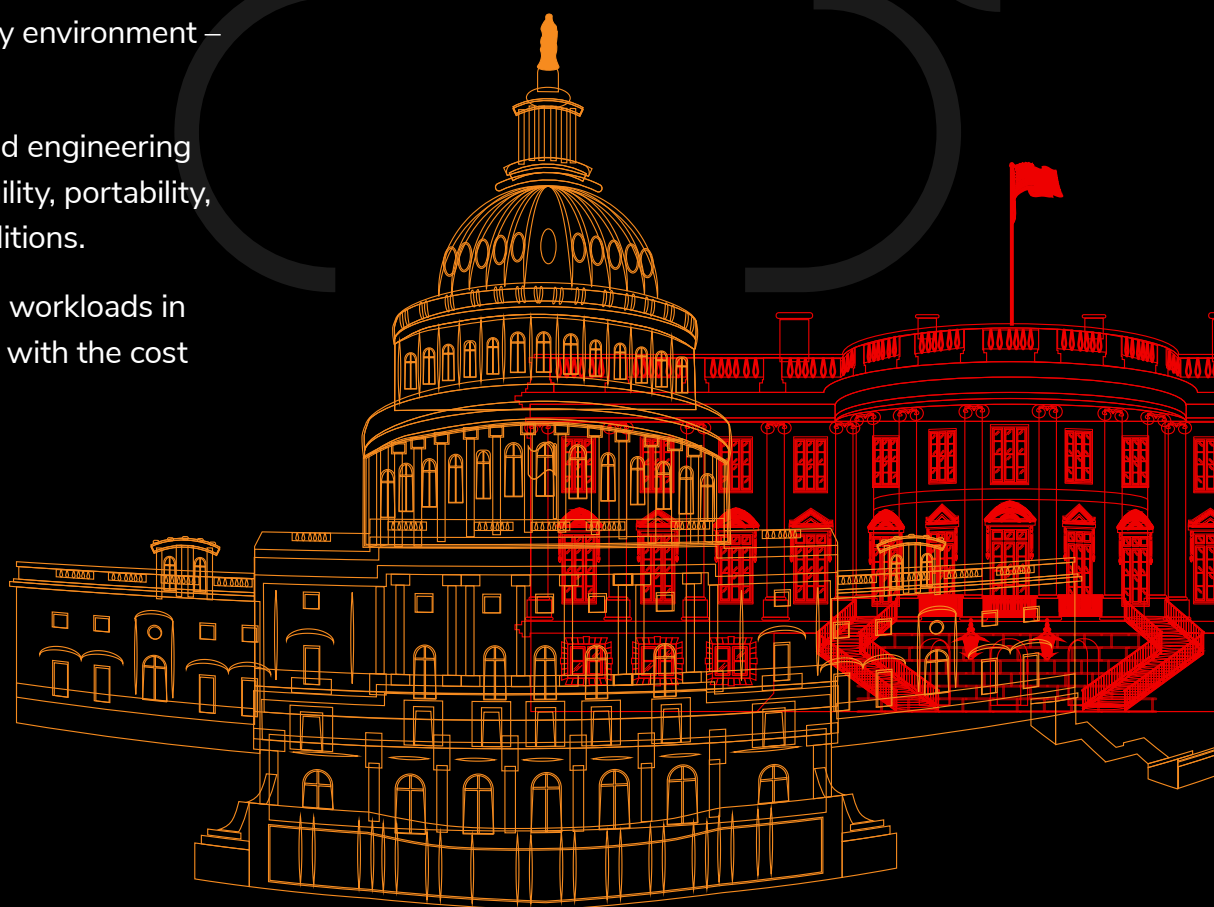
Red Hat OpenShift delivers the most comprehensive container orchestration capabilities across any environment – hybrid cloud, multi-cloud, or on-premises.

This means you get the rich data analytics and engineering capabilities of CDP Private Cloud, plus the agility, portability, and speed needed to adapt to changing conditions.

Federal agencies can rapidly deploy analytics workloads in a private cloud, getting cloud-like capabilities with the cost efficiencies of on-premises infrastructure.

Avoid vendor lock-in and achieve the flexibility needed to move workloads as conditions change, including across hybrid cloud and multi-cloud environments.

Retain complete visibility and control over your data with the security of a private cloud and robust compliance capabilities built for the needs of the public sector.

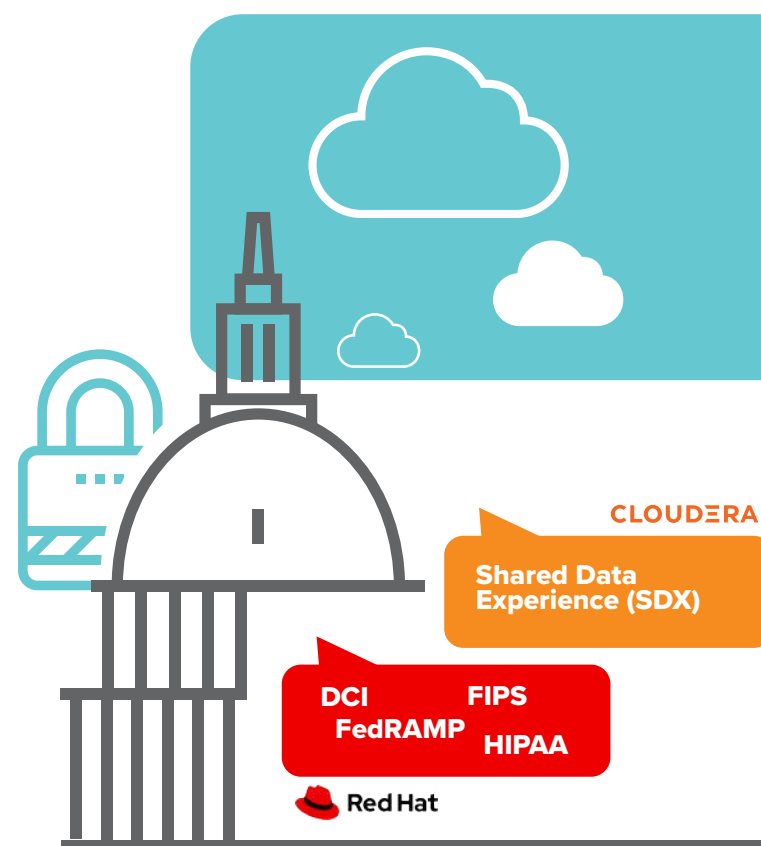


The Security the Public Sector Requires

Transforming your data strategy with a modern platform does not mean you must take unnecessary security or compliance risks. CDP Private Cloud on OpenShift delivers the security and compliance federal agencies require as a core fundamental of how they operate.

Red Hat is DCI, HIPAA, FedRAMP, and FIPS compliant, and it solves more types of security issues than anyone because of its speed and distribution network. That's one of the appeals of OpenShift in the broader cloud-native ecosystem of containers and Kubernetes. Its enterprise-grade platform ensures that security settings are properly configured for your environment.

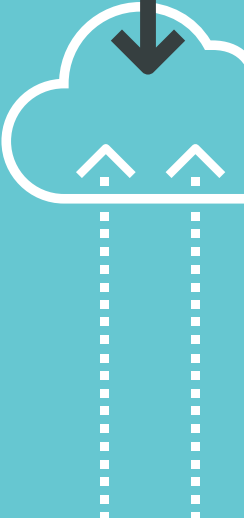
With Cloudera's Shared Data Experience (SDX) embedded into your CDP Private Cloud, you also get an integrated set of security and governance technologies that allow you to apply consistent granular policies to all data across the organization, plus end-to-end encryption. You also get rich compliance capabilities at your fingertips, including metadata search, data lineage, chain of custody, and data access auditing. This allows you to enable self-service data access throughout the organization while remaining fully compliant and prepared for an audit.



A Modern Data Platform in a Cloud You Control

A private cloud approach, paired with a cutting-edge enterprise data platform built on a cloud-native hybrid data architecture, brings the power of cloud on-premises – while maintaining the flexibility and portability required to prepare for growing data volumes and new use cases.

With CDP Private Cloud on OpenShift, federal agencies achieve the speed and scale of today’s cloud environments while maintaining tight control over their data sovereignty to ensure compliance and reduce security risks.



Cloudera Key Capabilities



ANY CLOUD

Deploy a hybrid cloud data platform that seamlessly works across private and multiple public clouds.



DATA LIFECYCLE INTEGRATION

Utilize multiple analytics together on the same data, eliminating costly and inefficient data silos.



SECURITY AND GOVERNANCE

Maintain strict enterprise data security, governance, and control across all environments.



OPEN

Ensure zero vendor lock-in and maximum interoperability with open source software, compute and storage.

CDP Private Cloud on **OpenShift** helps federal agencies manage and secure the data lifecycle, transforming complex data into clear and actionable insights in any cloud or data center.

Learn more at cloudera.com/partners/red-hat



Red Hat

CLUSTERA

