

Red Hat OpenShift Data Foundation

Cluster data management and persistent storage for Red Hat OpenShift

Key benefits

- A complete Red Hat OpenShift cluster data management platform to accelerate deployment of modern applications—in the datacenter, public or private cloud, and at the edge
- Simplified solution deployment and hybrid cloud data management for improved Day 1 and Day 2 operations
- Agile and flexible data access supporting common protocols—file, block, and object—for expanding workloads and innovation in less time
- Delivering insights through a consistent experience for developers, data scientists, and data analysts, across hybrid and multicloud environments
- Dynamic scale, allowing organizations to start small and automate rapid data services growth

A data foundation for modern production applications

Modern container-based applications are data driven and deliver important insights that help organizations compete and thrive, powering diverse workloads ranging from databases to analytics, data pipelines, artificial intelligence and machine learning (AI/ML), and more. Unfortunately, fragmented infrastructure can frustrate these efforts by failing to provide production applications with sophisticated cluster data management capabilities.

[Red Hat® OpenShift® Data Foundation](#) is a persistent storage and cluster data management solution integrated with and optimized for Red Hat OpenShift. In addition to a distributed, scalable software-defined storage platform, OpenShift Data Foundation provides sophisticated enterprise cluster data management services, allowing applications to interact with data in a simplified, consistent, and scalable manner. Multicloud data management capabilities extend and federate data across infrastructures. OpenShift Data Foundation runs anywhere Red Hat OpenShift does: on-premise, in public or private clouds, and at the edge. The platform abstracts the details and inconsistencies of different underlying storage infrastructures while delivering cluster data management services that organizations require—whether in the datacenter or to support timely insights and workflows at remote locations and at the edge. OpenShift Data Foundation provides file, block, and object storage classes, enabling a wide range of data modalities and business workloads, including:

- ▶ Databases, warehouses, and data caches.
- ▶ Artificial Intelligence and machine learning.
- ▶ Distributed event-driven workflows and data pipelines.
- ▶ Data ingestion, preparation, and analytics.

Red Hat OpenShift Platform Plus—a comprehensive platform

In early Kubernetes deployments, storage was often an afterthought, with many just using their local storage or cloud provider-specific offerings, limiting scalability and flexibility. As cloud-native applications and workloads running in Kubernetes have evolved, data has played an increasingly important role in both development and deployment. Red Hat views cluster data management as a fundamental component of a comprehensive modern application development platform. To that end, [Red Hat OpenShift Platform Plus](#) includes [Red Hat OpenShift Data Foundation Essentials](#) at no additional charge, along with other vital components, including:

- ▶ [Red Hat OpenShift Container Platform](#) for on-premise deployments.
- ▶ [Red Hat Advanced Cluster Management for Kubernetes](#) for multicluster deployment at scale.
- ▶ [Red Hat Advanced Cluster Security for Kubernetes](#) for built-in security across the software development life cycle.
- ▶ [Red Hat Quay](#) for a security-focused, global container private registry platform.

Red Hat OpenShift allows organizations to experience the breakthrough capabilities of Kubernetes from a consistent and supported enterprise platform—open to any application, team, or infrastructure.

OpenShift Platform Plus provides a complete platform with enhanced security capabilities, a consistent user experience, advanced cluster management, and cluster data management across the hybrid cloud and edge infrastructures. For additional data management services functionality, organizations can deploy OpenShift Platform Plus with Red Hat OpenShift Data Foundation Advanced (Table 1).

Table 1. OpenShift Data Foundation capabilities

Essentials edition provides	Advanced edition adds
<ul style="list-style-type: none"> • Kubernetes RWO (block, file) • Kubernetes RWX (shared file, shared block) • Object storage (s3-compatible) • Internal mode storage (on-cluster) • Volume snapshots • Cluster-wide encryption • Multicloud object gateway 	<ul style="list-style-type: none"> • External mode storage (shared storage cluster) • Mixed usage patterns (off-cluster workloads) • Volume-level encryption • External key management with bring-your-own-key (BYOK) support • Metro disaster recovery (DR) • Regional DR

Red Hat OpenShift Data Foundation

Delivered as an operator-based Kubernetes service, OpenShift Data Foundation is engineered, tested, and qualified to provide cluster data management services for Red Hat OpenShift on any infrastructure (Figure 1). OpenShift Data Foundation can also be managed as an external-mode data store, delivering data for one or multiple Red Hat OpenShift clusters and external workloads.

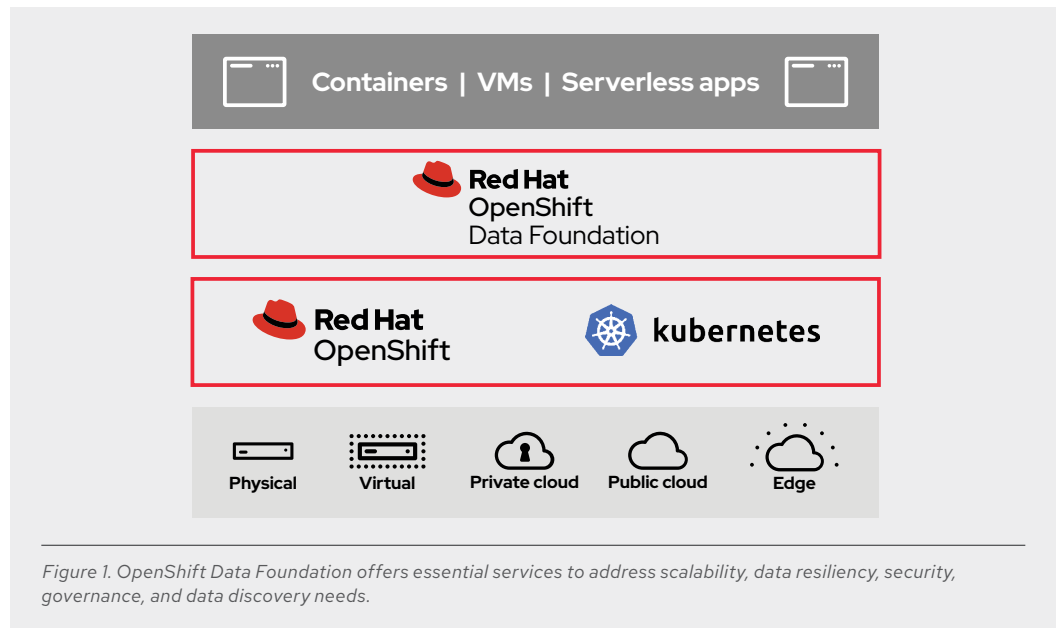


Figure 1. OpenShift Data Foundation offers essential services to address scalability, data resiliency, security, governance, and data discovery needs.

Deployed, consumed, and managed through the Red Hat OpenShift administrator console user interface (UI), the platform is built upon Ceph, Noobaa, and Rook technologies. The integrated technology combination yields tightly integrated, persistent cluster data management services for Red Hat OpenShift in hybrid and multicloud environments, delivering consistent data services for Kubernetes applications running across multiple infrastructure platforms. OpenShift Data Foundation provides cluster data management functionalities, broadened object data management services, data protection options, security capabilities, and automated DR capabilities for business continuity.

Simplified access

OpenShift Data Foundation ignites cross-team collaboration by providing simplified access to a consistent data platform experience. Self-service access from different user, developer, and administrator roles delivers storage on demand with a click, not with a ticket. Easy to install, intuitive to monitor, and simple to use, organizations spend less time managing separate platforms and more time delivering value, allowing them to:

- ▶ Provide accessible data and support for all their Red Hat OpenShift applications.
- ▶ Simplify data management across hybrid and multicloud environments with easy and consistent access, productive innovation, and instant insights.

Consistent user experience

Because OpenShift Data Foundation runs as an operator-based service in Red Hat OpenShift, users benefit from a consistent user experience independent of where the actual data resides. Whether you are building cloud-native applications with Kubernetes or just beginning your journey to cloud technology, your team gets a consistent hybrid cloud environment with intuitive cluster data management services across any infrastructure. Developers can build once and extend their deployment to the most appropriate location and infrastructure. Data scientists can generate insights in less time. Administrators can manage with consistency across cloud footprints. Organizations gain the ability to:

- ▶ Deliver a unified end-user experience for cluster data management services with consistency across clouds and confidence across teams.
- ▶ Provide an all-in-one view of Red Hat OpenShift data with a comprehensive data platform.

Dynamic scale

Moving to cloud technology cannot come at the expense of application performance. OpenShift Data Foundation supports scalability to multiple petabytes with resiliency and peak performance. Organizations can dynamically provision across any environment—bare metal, virtual machines, or hybrid cloud—all through the Red Hat OpenShift administrator console. Native object support in OpenShift Data Foundation dramatically increases input/output (I/O) performance, yielding performance at scale for every workload. Organizations can:

- ▶ Scale data services with the confidence to go anywhere and span any cloud.
- ▶ Innovate at scale while supporting diverse Red Hat OpenShift workloads, allowing easy object data sharing across geographic locations and platforms.

Tight integration with Red Hat OpenShift and Red Hat OpenShift Virtualization

OpenShift Data Foundation was created for container-based environments and integrated tightly with Red Hat OpenShift. With a supported Red Hat OpenShift operator, the platform is simple to install and manage as a part of the container-based application life cycle. This innovation allows Red Hat to provide support for the entire container-based environment, including cloud-native container management, scheduling, and orchestration, yielding:

- ▶ **Storage for trusted enterprise-class Kubernetes.** OpenShift Data Foundation adds support for diverse workloads, multicloud object gateway functionality, and business continuity for workloads running in Red Hat OpenShift.
- ▶ **Data protection and resiliency for Red Hat OpenShift.** High availability for stateful enterprise applications demands robust and highly available data storage with sophisticated capabilities. OpenShift Data Foundation supports important features like replication, allowing application data placement across different availability zones. Native support for data resiliency supports backup and restoration services for Kubernetes applications, including their namespace-related data, custom resources, and state.
- ▶ **A cloud-like experience everywhere.** Circumstances are constantly changing, favoring one cloud provider over another, or in-house deployment versus public cloud. Organizations must react quickly to changing business conditions, such as pricing, security and data governance, new technologies, and evolving workflow modalities. OpenShift Data Foundation provides software-defined storage that lets organizations deploy their applications and storage as needs dictate and adjust as situations change.
- ▶ **Increased developer productivity.** Cloud developers need the ability to innovate without arbitrary limitations. Traditional storage has often been an impediment to cloud development, requiring separate and time-consuming provisioning procedures. OpenShift Data Foundation provides common functionality across all hybrid cloud platforms, simplifying processes for developers and users.

Embracing diverse workloads

Most cloud providers support data storage for diverse workloads but typically use multiple storage technologies. This approach introduces inconsistency, management complexities, and inefficient access and scale. It can also lock applications into a given cloud vendor because other vendors offer different storage technologies and capabilities. In contrast, OpenShift Data Foundation provides:

- ▶ Uniform cluster data management services.
- ▶ Support for diverse workloads with support for file, block, and object storage protocols.
- ▶ A common presentation and management layer delivered within the Red Hat OpenShift console.

As a singular container-native storage solution, OpenShift Data Foundation provides resources for today's expanding workloads like databases and data warehouses, event-driven architectures, analytics, and AI/ML, as well as cluster infrastructure services (Figure 2). It supports multiple cloud platforms and works with a wide range of partners and technologies.

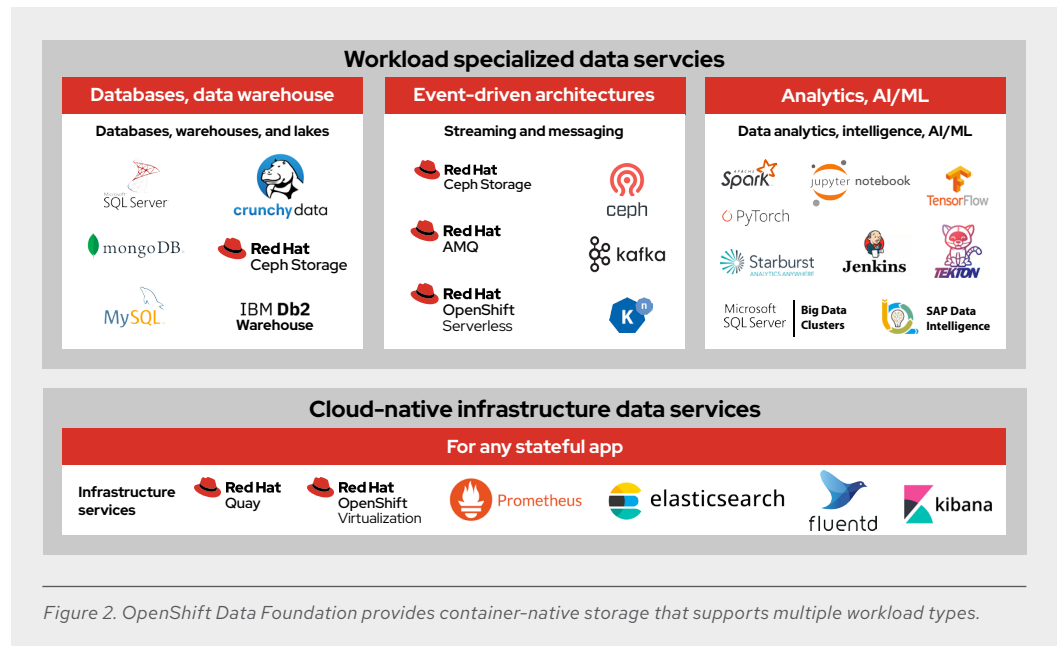


Figure 2. OpenShift Data Foundation provides container-native storage that supports multiple workload types.

Popular use cases for OpenShift Data Foundation include:

- ▶ Data repositories and cloud-native application development, including continuous integration and continuous deployment (CI/CD) models.
- ▶ Structured data, including SQL/NoSQL databases and data warehouses.
- ▶ Big data workloads such as data analytics and AI/ML workloads.
- ▶ Container-aware data protection with backup and data resiliency support for both persistent volumes and Kubernetes namespaces.

Sophisticated cluster data management services

OpenShift Data Foundation provides the functionality that organizations need to support applications and workloads in one complete solution—fully integrated and delivered with Red Hat OpenShift. By engineering cloud-native storage in lock-step with the rest of the Red Hat OpenShift platform, issues are identified in development and testing. The result is a platform with fewer support challenges and risk remediation.

Through its open source origin, OpenShift Data Foundation continues to evolve, powered by community innovation. Features and functionality are not hobbled by proprietary software constraints and concerns. This approach gives Red Hat the agility to rapidly surface emerging functionality while embracing expansive deployment modalities in less time.

Data services at the edge

Edge computing allows organizations to make flexible choices about data processing to reduce latency, decrease bandwidth usage, and improve the user experience. To accelerate business insights and data processing speeds, data can be stored closer to the inception point. By reducing the amount of data that needs to be transferred from the remote site, organizations can decrease cost and reduce latency and reaction time on issues.

Data access can be federated across hybrid and multicloud deployments with namespace buckets so data does not have to be moved. Critical data can be mirrored, spread, and tiered across edge, core datacenter, and cloud deployments for business continuity and further operational processing. OpenShift Data Foundation offers key datacenter functionality in a small footprint and scale to serve essential edge workloads, including:

- ▶ **3-node compact clusters.** Compact mode supports a complete cluster that includes Red Hat OpenShift and OpenShift Data Foundation on just three production nodes. Compact mode includes storage and eliminates distinct compute or worker nodes.
- ▶ **Single node.** OpenShift Data Foundation supports a single-node configuration, allowing snapshots and backups at edge locations. This capability allows virtual machines (VMs) to revert to a previous state or reliably instantiate from a snapshot-based template.¹
- ▶ **Single-node thin provisioning.** Thin provisioning creates the appearance of having more physical resources than are available. Thin provisioning lets workloads use only necessary resources initially, growing over time to accommodate user demand.¹

Business continuity and disaster recovery

Resilience and DR are essential for critical applications in production environments. In addition to commercial and partner data protection solutions, support for OpenShift APIs for Data Protection provides cluster and namespace context for snapshots and backup of data served by OpenShift Data Foundation. With Red Hat OpenShift Platform Plus, the combination of Red Hat Advanced Cluster Management for Kubernetes and OpenShift Data Foundation allows sophisticated DR solutions. These advances let organizations address business continuity needs for modern workloads running in Red Hat OpenShift.

- ▶ **OpenShift APIs for Data Protection** provides cluster-aware backup and data recovery interfaces in the context of applications running within Red Hat OpenShift. It can provide native namespace backup and recovery for individual workloads, applications, and cluster services. This process captures application state and cluster specifics allowing them to be restored onto the same or different clusters. OpenShift APIs for Data Protection can also provide APIs required by enterprise data protection solutions developed by Red Hat ecosystem partners.
- ▶ **Regional DR** operators based on Red Hat Advanced Cluster Management automate stateful application recovery. Block volume protection and asynchronous replication yield application recovery that works across regions to protect against geographic-scale failures. This functionality can be used with higher-latency connections.²
- ▶ **Metropolitan DR** allows no-loss recovery across two DR sites within a campus or metropolitan area. This approach protects against data loss using Red Hat Advanced Cluster Management with multiple clusters and synchronous replication, requiring low-latency network connectivity. Metropolitan DR provides instant protection of business functionalities, with a near zero recovery point objective (RPO).³

¹ This functionality is in Technology Preview status as of this writing.

² This functionality is in Technology Preview status as of this writing. OpenShift Data Foundation internal-mode support only.

³ This functionality is in Technology Preview status as of this writing. An arbiter for the storage cluster is required in a neutral zone.

Multicluster monitoring

Manually managing storage resources across multiple clusters can be a daunting task. Red Hat Advanced Cluster Management provides a single view of storage health and capacity spread across multiple Red Hat OpenShift clusters. This tool offers observability for health and optimization. Multicluster dashboards can also store long-term historical data. The graphical console can identify, isolate, and resolve issues affecting distributed workloads.

Multicloud gateway

Managing multiple cloud-based data resources can quickly become complicated. Organizations now routinely use numerous data resources across multiple public cloud providers and employ a mixture of on-premise and cloud storage. Keeping track of data resources and credentials can be complex and a source of human error. With cloud-native applications built on microservices, individual workloads often need to access multiple data sources that may reside in more than one cloud environment. Replication of large datasets between sites is not always possible or desirable.

The integral [multicloud object gateway](#) in OpenShift Data Foundation allows administrators to address these issues while dramatically reducing complexity. As a lightweight object storage service for Red Hat OpenShift, multicloud object gateway delivers simplified data management, flexibility across multiple storage backends, and a consistent experience for both data consumption and management.

As shown in Figure 3, the multicloud object gateway exposes a consistent Amazon Simple Storage Service (Amazon S3) compatible endpoint to developers and applications. On the back end, the multicloud object gateway connects to multiple kinds of policy-driven data buckets:

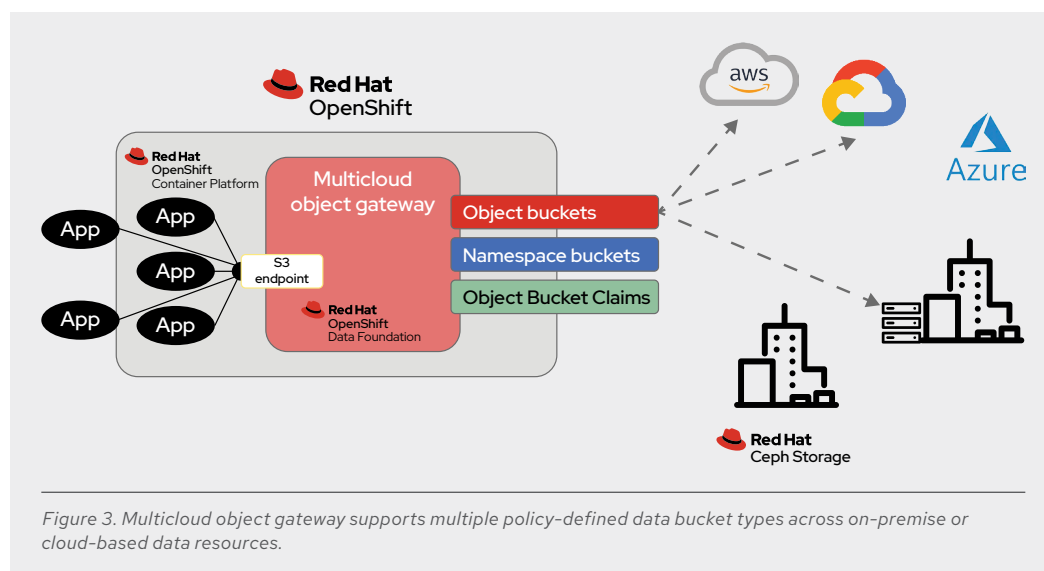
- ▶ **Object buckets** contain object data and offer data safety with available mirroring, spreading, and tiering approaches.
- ▶ **Namespace buckets** provide data federation capabilities, letting administrators and data managers organize, configure, and manage diverse data resources, without having to copy data sets.
- ▶ **Object Bucket Claims** dramatically simplify data creation and management for developers, allowing the dynamic creation of object buckets for workflows.

Enhanced data security

Data security has never been more essential. Organizations must protect their valuable data, whether in transit or at rest. For event-driven architectures, OpenShift Data Foundation clusters support optional end-to-end encryption between Ceph daemons and Ceph clients. Databases and data warehouses can also be protected through available persistent volume level encryption. Ceph can encrypt all data at rest, irrespective of whether the client is a Ceph block device, Ceph File System (CephFS), or a custom application. In addition, the Ceph Object Gateway supports encryption with customer-provided keys. The platform supports external key management services (KMS) integration for maximum organizational flexibility.

Network File System (NFS) support

As organizations modernize their applications and migrate data to the cloud, they need support for traditional access protocols for interconnectivity with existing applications. OpenShift Data Foundation supports NFS services for internal or external applications running in different operating



About Red Hat

Red Hat helps customers standardize across environments, develop cloud-native applications, and integrate, automate, secure, and manage complex environments with [award-winning](#) support, training, and consulting services.

North America
1 888 REDHAT1
www.redhat.com

Europe, Middle East, and Africa
00800 7334 2835
europe@redhat.com

Asia Pacific
+65 6490 4200
apac@redhat.com

Latin America
+54 11 4329 7300
info-latam@redhat.com



f facebook.com/redhatinc
t [@RedHat](https://twitter.com/RedHat)
in linkedin.com/company/red-hat

redhat.com
#0000000_0021

environments. NFS support helps provide data migration to a Red Hat OpenShift environment from storage solutions that use NFS, including Red Hat Gluster Storage.

Embracing application modernization strategies

Organizations engaged in application modernization efforts often face challenges around existing applications. Supporting multiple disparate environments while applications are in transition can be expensive and inefficient. Since not all applications are ready to move to containers, Red Hat OpenShift offers [Red Hat OpenShift Virtualization](#), supporting virtual machine environments on Red Hat OpenShift. With this feature, organizations can continue to run their existing applications alongside their cloud-native application development—all on a single platform. OpenShift Data Foundation delivers an ideal storage solution within the cluster for both modern applications and existing applications running in Red Hat OpenShift Virtualization. An external OpenShift Data Foundation Advanced cluster can [support workloads within Red Hat OpenShift](#) as well as existing Linux® or virtual machine environments.

Conclusion

OpenShift Data Foundation provides simplified access, a consistent experience, and dynamic scale for persistent data services anywhere that Red Hat OpenShift runs—across on-premise infrastructure, public cloud, or hybrid cloud. The platform offers sophisticated enterprise data services that can serve myriad modern applications supporting production workloads. Organizations can use these services to deploy flexible data services at the edge, secure their operations, and implement essential DR scenarios. Because OpenShift Platform Plus includes OpenShift Data Foundation Essentials at no additional cost, effective data cluster management has never been more cost-effective or simpler to achieve.