



RUBRIK CLOUD DATA MANAGEMENT

# Technology Overview & How It Works

# CONTENTS

<b>Introduction to Cloud Data Management</b>	3
Data Management for the Cloud Generation	
Lifecycle Data Management	
The Problem with the Legacy Approach	
<b>Introduction to Cloud Data Management</b>	6
<b>Technology Overview</b>	7
Rubrik Cloud-Scale File System	
Rubrik Distributed Metadata System	
Rubrik Distributed Task Framework	
Rubrik Data Management and Global Search	
Erasure Coding	
Physical Support	
<b>How It Works</b>	9
Rack-and-Go System Setup	
Automated Data Discovery	
Flash-Speed Data Ingestion	
Global Predictive Search	
Instant Recovery	
Single Dynamic Policy Engine Across Clouds	
API-First Architecture	
Security and Compliance	
<b>Business Agility Use Cases</b>	12
RESTful APIs for Integrated Automation	
Rubrik at the Edge – A Software Appliance	
Cloud Mobility Solutions	
Live Storage for DevOps	
<b>Conclusion</b>	15

# INTRODUCTION TO CLOUD DATA MANAGEMENT

*World's first Cloud Data Management platform that delivers data protection, search, analytics, and copy data management to hybrid cloud enterprises.*

---

## DATA MANAGEMENT FOR THE CLOUD GENERATION

Rubrik's Cloud Data Management platform democratizes public cloud for all enterprises by delivering a modern solution to recover, manage, and secure all data, regardless of location. The power within an enterprise to scale organically and respond to shifting business needs depends on instantaneous access, efficient data delivery, and intelligent data management across the entire organization. Rubrik eliminates the barriers to data mobility, providing mission critical functions needed to drive business agility, cost effectiveness, and performance as enterprises shift workloads to the cloud.

Rubrik Cloud Data Management is a single, software fabric that manages all data in the cloud, at the edge, or on-premises for many use cases including backup, disaster recovery, archival, compliance, analytics, and copy data management.

## THIS IS CLOUD DATA MANAGEMENT

---



*"We're in the business of making IT not only easy, but secure, for our customers. Rubrik allows us to deliver Protected Trust's vision of managing data throughout the entire lifecycle from creation to archival to disposal."*

**Chris Bridges, CTO, Protected Trust**

---

## LIFECYCLE DATA MANAGEMENT

Rubrik's Cloud Data Management platform provides a holistic approach to manage the lifecycle of data, from creation to expiration, across hybrid cloud environments to drive better performance and operational continuity at lower costs. Comprehensive data management is delivered through instant access, automated orchestration, and enterprise-class data protection and resiliency.

- **Instant Access:** Rubrik delivers instant application recovery and predictive global search by unifying data locked within disparate application silos into one globally indexed namespace while leveraging zero-byte cloning technology to enable on-demand copy data workflows.
- **Automated Orchestration:** Rubrik dramatically reduces daily operational management, providing a step-function change in simplicity by enabling a single policy engine to orchestrate service level agreements (SLAs) across the entire data lifecycle. The Rubrik programmatic interface automates how data services are created, consumed, and retired across clouds.
- **Security and Compliance:** Rubrik secures data whether in-flight or at-rest throughout its lifecycle, regardless of location. The Rubrik platform delivers granular role-based access control across all cloud data management workflows while providing automated compliance reporting to successfully complete various industry and internal audits.

---

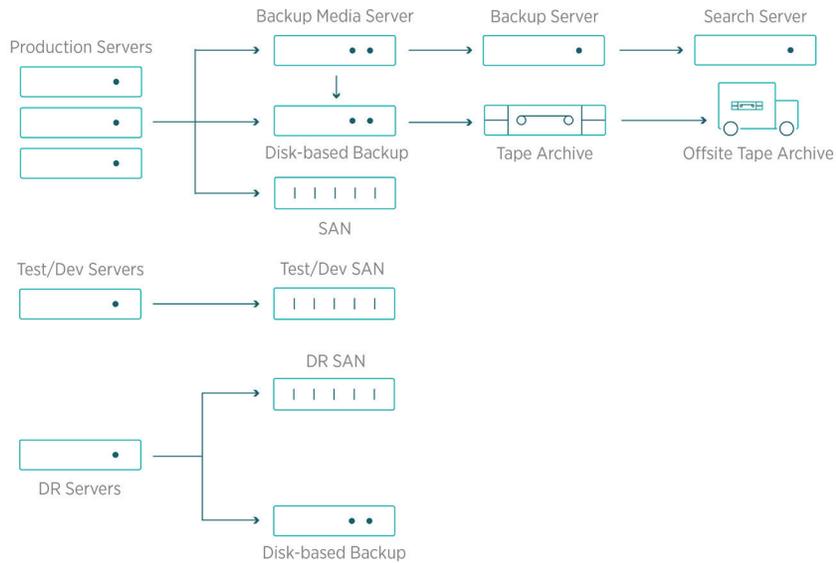
## THE PROBLEM WITH THE LEGACY APPROACH

Enterprises are shedding complex, legacy multi-tiered solutions for simplified data management purpose-built for cloud. The past approach of using multiple point solutions results in brittle infrastructure with challenges such as long deployment cycles, inability to meet SLAs, lack of scalability, architectural complexity, and fragmented management.

Rubrik bridges the gap between owned, on-premises infrastructure and the cloud by decoupling data from the data center. With a software-defined approach, Rubrik delivers the agility and performance of web-scale platforms and cloud providers while automating workflow management to provide resiliency and agility as enterprises on-ramp to the cloud.

## BEFORE RUBRIK

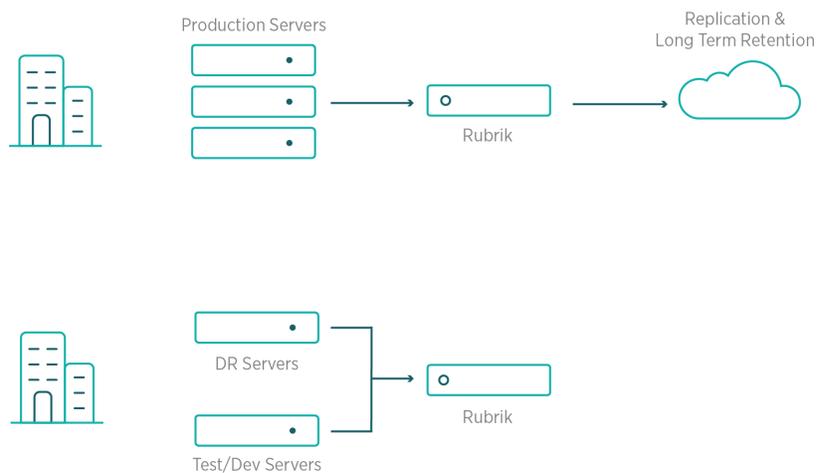
---



The diagram above depicts legacy architecture consisting of multiple points of technology, including media servers, backup storage, tape, and software licenses. This results in brittle infrastructure with challenges such as management complexity, high costs, and lack of scalability.

## AFTER RUBRIK

---



Rubrik simplifies data management by consolidating hardware and software components into a single, software-defined platform built for the cloud.

# THE PRINCIPLES BEHIND OUR TECHNOLOGY

The Rubrik Cloud Data Management platform incorporates the following design principles:

- 1. Software-Defined:** We distill physically disparate hardware and software components underlying the multi-tiered, legacy backup and recovery architecture into one software. We embed rich data services into this software fabric such that captured data can be repackaged for other purposes.
- 2. Simplicity:** We solve for ease of use and bring consumer-grade usability to enterprise software. For example, we design our user interface to display only information that requires user attention, reducing cognitive overload.
- 3. Efficiency:** We build intelligence into our software to help users efficiently manage data without incurring unnecessary costs (e.g., zero-byte cloning to save on storage capacity, sending deduplicated data to the public cloud to reduce data transfer and storage) and labor (e.g., file search across a global index that spans public and private clouds).
- 4. Web-Scale Architecture:** We adopt the same web-scale technologies used by Google, Facebook, and Amazon, allowing our users to easily handle rapidly increasing volumes of information with a linear architecture. Users avoid painful forklift upgrades and continue to easily manage Rubrik as a single system.
- 5. Cloud-Based Infrastructure:** We purpose-built our system for cloud integration and infinite scale to enable hybrid cloud for all enterprises. Extend search, disaster recovery, and archival into the cloud.
- 6. Ecosystem Support:** We have designed our data platform to be vendor-agnostic and to work across modern data center applications and technologies. This includes ability to support a variety of platforms and data sources, such as virtualization, containers, and physical applications. Our software fabric is location-agnostic and can be extended from on-premises to public clouds.

# TECHNOLOGY OVERVIEW

Rubrik Cloud Data Management is a system that distributes data, metadata, and task management across the cluster in order to deliver predictive scalability and eliminate performance bottlenecks. The data management functions as the brains of the system, organizing, removing redundancy, and making data available for search. Our user interface is built on a RESTful API-driven framework with a HTML5 web user interface that supports virtualization, applications, and public cloud technologies.

---

## RUBRIK CLOUD-SCALE FILE SYSTEM

The Rubrik Cloud-Scale File System is a distributed file system built from the ground up to store and manage versioned data. We have designed this system with the following properties:

- **Fault tolerant:** The system is resilient to multiple node and disk failures. We employ an intelligent replication scheme to distribute multiple copies of data throughout the cluster.
- **Flash-optimized:** The system is built with a hybrid flash/disk architecture to maximize I/O throughput.
- **Linearly scalable:** Rubrik is architected to be a web-scale system and infinitely expand. Increasing storage capacity and I/O performance is as simple as adding more nodes to the system.
- **Copy data management:** The system utilizes zero-byte clones to make multiple copies of data from one “golden image,” maximizing storage efficiency for application test and development.

---

## RUBRIK DISTRIBUTED METADATA SYSTEM

The Rubrik Distributed Metadata System operates alongside our Cloud-Scale File System, providing a global index and catalogue that can be accessed at high speeds. It delivers continuous availability, linear scalability, and operational simplicity with no single point of failure in the cluster. Our system is built to handle large amounts of data, distribute replicas of data across nodes (access to metadata is maintained even in the case of node failure), and provide low latency access to facilitate search.

---

## RUBRIK DISTRIBUTED TASK FRAMEWORK

The Rubrik Distributed Task Framework is the engine that globally assigns and executes tasks across the cluster in a fault tolerant and efficient manner. Designed to be masterless, tasks are load balanced across the entire cluster, and tasks are distributed to the nodes that house the impacted data. Employing data affinity, we assign tasks based on data location and node resource availability, ensuring that all relevant data for a group of related data entries is managed on a single node without compromising fault-tolerance.

---

## RUBRIK DATA MANAGEMENT AND GLOBAL SEARCH

Rubrik Data Management and Global Search serves as the “brains” of the system, enabling cradle-to-grave data lifecycle management from data ingest to archive to retirement. Fast, efficient data delivery is made possible by the ability to:

- **Store versions of data:** We use a combination of full snapshot with forward incremental and reverse incremental copies.
- **Ensure data integrity:** We build multiple integrity checks within the file system and data management layers.
- **Apply content-aware global deduplication and compression:** We intelligently apply data reduction at a global level while enabling fast data reconstruction.

---

## ERASURE CODING

Rubrik uses erasure coding to deliver resilience and fault tolerance. Erasure coding encodes and partitions data into fragments and calculates parity such that in the event of a disk or node failure, the original data can be reconstructed. The number of data and parity blocks is configured based on the desired number of failures to withstand. As a result, it maintains fault tolerance while increasing capacity at the same high performance. Erasure coding requires a minimum cluster size of four nodes.

---

## PHYSICAL SUPPORT

For physical Microsoft SQL Server, Windows Servers, and Linux, Rubrik provides comprehensive data management, including backup, replication, recovery, and data archival. Automatic host or database discovery simplifies the traditionally complex, labor-intensive setup posed by legacy backup. Rubrik provides a forever incremental approach, computing changed blocks on the physical host to send back to Rubrik for forever incremental backups. Manage both physical and virtual environments in one user interface with a single SLA policy engine to define backup, replication, and archival irrespective of the source.

For Microsoft applications such as SQL, Exchange, and SharePoint, Rubrik offers built-in integration with Kroll Ontrack PowerControls™ for object-level restores (such as MS SQL tables, MS Exchange mailbox and emails, and MS Sharepoint objects and sites) to eliminate the complexity of doing a database rollback or restore. Additionally, always-on database group support allows the user to switch protection among replicas of a database within the always-on group without taking a duplicated full copy. Point-in-time restore allows recovery at the relevant point-in-time snapshot and applies transaction logs to restore back to the desired point. Rubrik tracks the changed blocks in a database, and the transaction logs they are associated with for granular RPOs.

# HOW IT WORKS

**RACK-AND-GO SYSTEM SETUP:** Once racked, the Rubrik system setup is easily and quickly completed in under an hour for virtual and physical environments. We invoke multicast DNS protocols to automatically discover and self-configure each of the nodes within the cluster. The user assigns IP addresses to each of the nodes and login credentials for the environment to be managed by Rubrik. To expand cluster size, the user simply assigns new IP addresses through the management dashboard. To reduce cluster size, the user selects the nodes to remove. Thereafter, the cluster automatically self-adjusts and re-balances to deliver fault tolerance against node and disk failures.

**AUTOMATIC DATA DISCOVERY:** Once the user enters the credentials for its virtual and physical environment (e.g., vCenter username/password for VMware vSphere environments), Rubrik auto-discovers details of the entire virtual and physical environments, such as hosts and applications. Auto-discovery happens a variety of ways, depending on the user environment. Rubrik utilizes VMware APIs (vStorage APIs for Data Protection) to discover VMware environments. For Microsoft SQL Server, once registered, the Rubrik connector runs on the physical host to automatically discover databases, servers, and physical applications.

**FLASH-SPEED DATA INGESTION:** We have designed Rubrik as a high-speed data ingestion engine that can easily handle large volumes of data for both virtual and physical environments. Rubrik pioneers the usage of flash in data protection, resulting in extremely fast data extraction and minimizing performance impact to the production environment. All data enters Rubrik through the flash tier. In addition, we have built an intelligent distributed workflow management system to maximize the number of parallel data streams processed. Since Rubrik is architected to be a web-scale system, performance for every dimension (such as network and disk throughput) increases predictably at a linear pace as more nodes are added to the cluster.

For VMware environments, we utilize VMware's Changed Block Tracking to identify and copy only the changed blocks from the previous operation. We apply intelligent global deduplication and compression before the data is stored in our Cloud-Scale File System. All metadata is stored in the flash tier for rapid access in a search pulldown. Data is distributed across multiple nodes to deliver a fault tolerant file system. For physical applications, we automate discovery of changed blocks from the previous operation and the transaction logs they are associated with.

---

## INSTANT ACCESS

Instantaneous access and effective data delivery across the organization enables businesses to ensure data availability and recovery real-time without impacting production. We unify data locked within disparate application silos into one globally indexed namespace while facilitating on-demand copy data workflows.

**GLOBAL PREDICTIVE SEARCH:** With Google-like search, Rubrik eliminates the file search complexity inherent in legacy solutions by introducing consumer-grade file search that delivers query results instantly. As the user types the query, Rubrik expedites the query by displaying suggested search results with auto-complete functionality. The user can instantly locate specific versions of files across time and locations for all VMs, applications, and file systems - no matter where the data resides (on-premises or in the cloud).

**INSTANT RECOVERY:** Rubrik radically simplifies the recovery process down to object-level and file-level recovery for virtualized environments. With just a click, users can instantly recover the VM by booting the virtual machine disk file (VDMK) directly on the Rubrik system. Rubrik serves as a storage endpoint for users to recover as many VMs as needed, eliminating the complexity and time wasted in transferring data back into the production system for recovery, thus providing a near zero RTO. Post-recovery, users can either choose to Storage vMotion the VMDK to the primary storage environment or continue using Rubrik as a storage endpoint. Rubrik's flash usage delivers fast IO performance. Writes and reads are gathered on the flash tier to deliver performance required by the recovered application.

---

## AUTOMATED ORCHESTRATION

Rubrik eliminates daily operational management with a single policy engine and programmatic interface that automates how data services are created, consumed, and retired across cloud. The core of Rubrik is its software that captures data from both physical and virtual environments and orchestrates applications anytime, anywhere.

**SINGLE DYNAMIC POLICY ENGINE ACROSS CLOUDS:** From the list of discovered virtual machines (VMs), physical databases, and applications, the user selects which VMs, physical databases, and applications to protect and what SLA policies to apply for recovery. To ease management, we have pre-configured SLA policy templates based on industry standard SLAs. The user has the ability to create new SLA domain policies by specifying the desired snapshot capture frequency and data retention policy. Users can select where data is stored, whether on-premises or in a public or private cloud service (e.g., Amazon S3 or Microsoft Azure). The user simply slides the bar to the time at which data should be archived in the public cloud (e.g., 30 days). Rubrik provides a cost-effective alternative to tape for long-term data retention. Rubrik allows users to intelligently and safely leverage the cloud. For example, Rubrik transmits deduplicated data to the cloud to reduce network and storage costs.

---

## SECURITY AND COMPLIANCE

**API-FIRST ARCHITECTURE:** Rubrik is a vendor-agnostic platform with the ability to support any third-party ecosystem technology by building additional modules to the integration layer. This layer exposes the API set for building custom integration points into applications, hypervisors, containers, and protocols. Our REST APIs are designed to integrate seamlessly with third party solutions to create end-to-end workflow and integrates multiple data sources into one simplified workflow. Please see **Business Agility Use Cases** for common workflows used to automate task management, simplify data management, and enable cloud migration.

The Rubrik platform takes a multidimensional approach coupling data plane security with management plane security from creation to expiration across hybrid cloud environments. Data in-flight and at-rest in the cloud utilizes military-grade AES 256-bit (FIPS 140-2 validated) encryption. Rubrik leverages client-side encryption libraries supported by public cloud providers and all archived data undergoes envelope encryption. Rubrik is in process with Common Criteria, a global security stand, at Evaluation Assurance Level 2 with augmentation of security development lifecycle processes.

Rubrik secures all data through the following methods:

- **Role-based access control:** Granular control over user access is defined at a platform level, regardless of location. Rubrik software allows self-service access to empower users to perform own backups, restores, and archival without impacting production.
- **End-to-end encryption:** Data is secured with HDD and SSD encryption with the r528 appliance and from the source to target location whether at-rest or in-transit, on-prem or in the cloud. Rubrik also supports software-based encryption at rest on all the r300 appliances. All data is encrypted before leaving the appliance for the cloud.
- **Centralized compliance reporting:** Track whether backup, replication and archival jobs are successful and meet the defined policy with automated compliance reporting.
- **Flexible key management:** Protect cryptographic keys with the embedded Trusted Platform Modules (TPMs).
- **Instant Recovery:** Protect from physical breaches by reverting to a point-in-time to determine breach or for recovery from ransomware. Our users have recovered from ransomware attacks and resumed operations in under an hour with zero data lost due to the ability to recover instantly and even utilize Rubrik's API-first architecture for further recovery automation.

# BUSINESS AGILITY USE CASES

---

## RESTFUL APIs FOR INTEGRATED AUTOMATION

Our set of APIs enables a rich ecosystem of infrastructure automation including the following widely used workflows:

- **Simplification of data management:** Our users have built end-to-end workflows with Rubrik's REST APIs to simplify orchestration of daily data management tasks including backup, replication, archival, and recovery. Users commonly use VMware vRealize Orchestrator (vRO) and vRealize Automation to automate delivery of Backup-as-a-Service. These automation workflows enable IT to be easily consumed as a service.
- **Orchestration of tasks:** Rubrik's REST APIs plug into configuration management tools, such as Puppet, Chef, Salt, and Ansible. By integrating Rubrik with these orchestration tools, users can simplify deployments of Rubrik across hundreds of servers or VMs. Additionally, when VMs are created on demand, they are automatically associated with a data protection policy (scheduled backups, replication, archival).
- **Migration to the cloud:** Rubrik provides an easy entry point into major cloud providers, such as Amazon Web Services (AWS) or Microsoft Azure, with our easy-to-consume APIs. Create custom workflows that quickly move local data to the cloud and intelligently manage data in the cloud according to business needs.
- **Customizable intelligent alerting:** Use Rubrik's built-in reporting capabilities or customize your own reports within third party tools by selecting desired system health monitoring and capacity alerts.



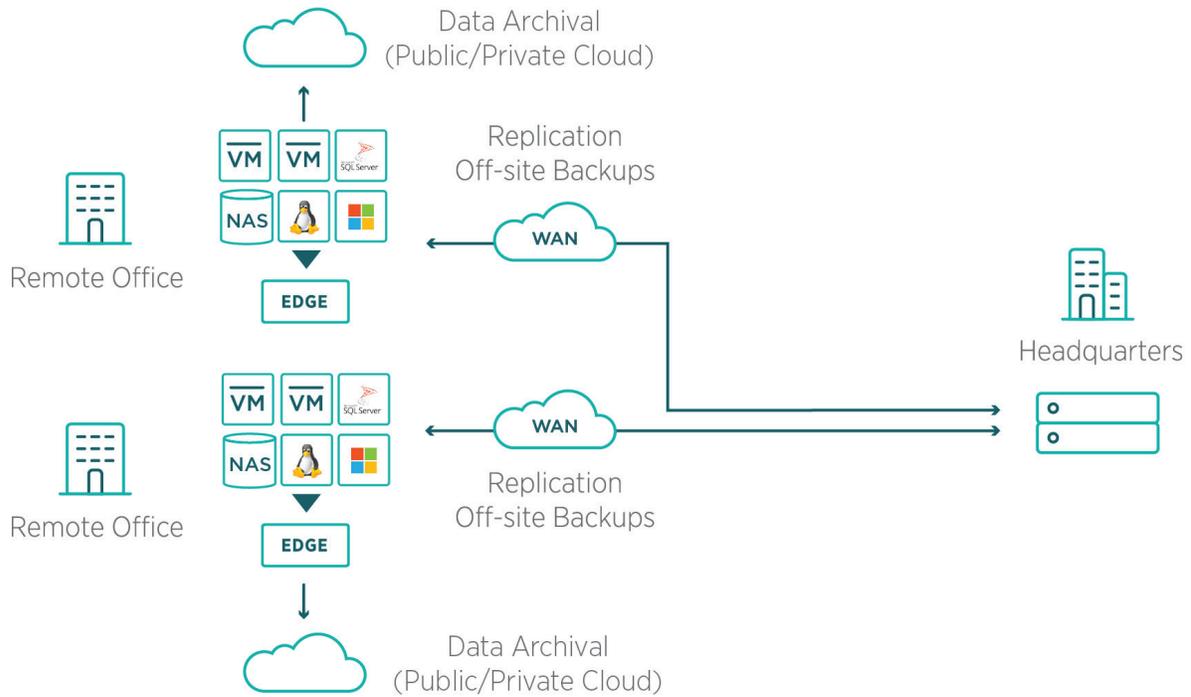
## AUTOMATED EFFICIENCY

“We’ve been able to meet our SLAs for RTOs and RPOs. With Rubrik, we can spin up a VM in minutes rather than hours or days. Restoring data becomes quicker, helping us become more agile. The steps required to recover go from 20 to 2 with Rubrik.”

**Bob Love, Director of IT, Bottomline Technologies**

## RUBRIK AT THE EDGE- A SOFTWARE APPLIANCE

Rubrik introduces Edge, a full-featured software appliance that extends data protection and management to virtual and physical environments in remote and branch offices. Deploy Rubrik Edge at remote locations to backup locally, replicate to a central data center, and archive to the cloud with an easy-to-use interface. All of Rubrik’s capabilities for efficient data management are available at remote sites, including policy-based management, global real-time search, quick recovery, and compliance reporting and alerting. The solution delivers the same, easy-to-use software features in a unified console to manage all your data through a HTML5- based and responsive interface. Reduce transfer and storage costs by sending deduplicated data for replicating back to the central data center. Our software appliance allows you to bring your own hardware with all hardware requirements determined by the capabilities needed from the software.



*Deploy Rubrik Edge as a software appliance at your remote and branch office locations.*



### ELIMINATING TAPE COMPLEXITY

“With Rubrik, I can eliminate tape and shift to using AWS for our long-term data retention. I can instantly locate specific data sets by searching an index that catalogues data on and off-site, eliminating the pain and time spent on taping out and waiting for recalled tapes.”

**Jake Warren, Systems Administrator, Red Hawk Casino**

---

## CLOUD MOBILITY SOLUTIONS

Rubrik enables cloud tiering to cost-effective locations over time by delivering scalable and efficient data management in the cloud. By eliminating the barriers to data mobility, Rubrik allows all enterprises to shift workloads in the cloud. Data archival, disaster recovery, and long-term retention is easily managed with Rubrik and public cloud (such as Amazon AWS and Microsoft Azure) or private cloud object storage (such as Cleversafe, Cloudian, Scality, etc.). Users can easily set up archival locations, and once active, they can use the same interface to modify and update if there are any changes to their archival services. Users can have insight into the efficiency, performance, and usage of their archival services and ensure that their SLA requirements are being met effectively from one single screen. Easily access the current active archive location as well as previous archives for quick data recovery using Rubrik's global search capability. Rubrik ensures secure data archival, encrypting all data that is archived before it leaves the Rubrik appliance.



### PUBLIC CLOUD MOBILITY

“We run systems in a hybrid cloud environment. If we experience some type of DR event, we see Rubrik as a path to restoring these systems in the public cloud. From a capital and time perspective, this poses significant savings. Rubrik's product vision coincides nicely with how we are evolving our business.”

**Todd Shutts, Senior Vice President, Balance Innovations**

---

## LIVE STORAGE FOR DEVOPS

Rubrik pioneers the concept of Live Storage in which any copy of data can be mounted directly on Rubrik as a storage endpoint for virtualized environments. As a result, Rubrik can be used to accelerate application development by providing multiple copies to developers from just one “golden image.” Our Cloud-Scale File System has built-in native linked-cloning capabilities to allow any number of mounts to be created without requiring additional storage capacity. Users can provision as many copies to developers as needed without impacting storage capacity and within a sandbox environment to prevent any network conflicts. As developers alter the provisioned data set, Rubrik stores the deltas by forking to a new branch. Our journaled Cloud-Scale File System provides an extremely efficient mechanism for accelerating and provisioning the latest data for application development. For medium-sized workloads, users receive all-flash performance comparable to a primary system of similar capacity. Rubrik intelligently allocates the flash tier for all writes and hot reads when utilizing Live Storage.



## PROTECTION FROM RANSOMWARE

“Having a top-notch data management solution in place means I can go about my day-to-day job without worrying about data loss. I know I have it covered.”

**Matthew Day, ICT and Support Manager, Langs Building Supplies**

# CONCLUSION

Rubrik Cloud Data Management is a disruptive approach to enterprise data management through a simple, run-anywhere platform that spans across clouds. Rubrik brings new value to your data with data management built to drive cost-effectiveness and resiliency combined with cloud-native support to enable business agility. Its innovative technology redefines how enterprises recover, manage, and secure their data, enabling cloud for all enterprises.