TECHNICAL WHITE PAPER

Rubrik and Amazon Web Services

Technology Overview and How It Works

TABLE OF CONTENTS

THE UNSTOPPABLE RISE OF CLOUD SERVICES

The digitization of business requires enterprises to move faster and be more agile to survive. Applying new technologies to existing business activities (e.g., leveraging AI to increase customer satisfaction) will continue to fuel the cloud paradigm. According to IDC, enterprises will spend more than $500 billion on cloud and cloud services by 2023, with 80% of application development on cloud platforms. For many enterprises, public cloud represents the ability to rapidly access resources for innovation while operating in a data-rich environment.

CLOUD PARADIGM INTRODUCES DIFFERENT PRINCIPLES

Enterprise IT looking to increase cloud usage will find that marrying non-cloud systems with cloud-native applications and infrastructure offers up new principles.

- **Shift from asset to service consumption.** Traditional IT is largely based on providing finite assets that service relatively stable workloads and predictable business growth. In a cloud model, IT rapidly provisions services according to business demand.
- **Automate service delivery.** With cloud, near-zero time to market can be delivered through automation frameworks. Infrastructure becomes programmable through code by being structured into templates that can be easily versioned and replicated for future deployments.
- **Develop applications based on microservices.** Rapid shifts in business demand require applications to deliver newer capabilities faster, to be resilient to failures, and to scale-out on-demand. Applications built in this new manner can be decomposed into independent components called "microservices", each delivering a single function.

With public cloud playing a greater role in overall enterprise IT strategy, the need for a cloud-scale data management platform becomes paramount to protect and manage data born in the cloud and elsewhere.

WHAT IS RUBRIK CLOUD DATA MANAGEMENT?

As enterprise migrate applications to AWS, IT will need to deliver core data protection (backup, disaster recovery, archival) in the event of service outages, data loss, and natural disaster.

The Rubrik Cloud Data Management platform provides a cloud-native approach to managing the lifecycle of data, from creation to expiration, to drive better performance and operational continuity at lower costs. Rubrik bridges the gap between owned, on-premises infrastructure and the cloud by decoupling data from the data center through a software-defined fabric and offering a single management plane for all data, whether on-prem or in the cloud. Comprehensive data management is delivered through instant access, automated orchestration, and enterprise-class data protection and resiliency.

- **Instant Access:** Rubrik empowers users to find the right data quickly, with predictive global search across applications. That search functionality also enables rapid recovery, with file-level granularity.
- **Automated Orchestration:** Rubrik dramatically reduces daily operational management, providing a stepfunction 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 data center and AWS.
- **Security and Compliance:** Rubrik secures data whether in-flight or at-rest throughout its lifecycle. 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.

DATA MANAGEMENT DESIGNED FOR AWS

Delivering data protection and management for cloud requires a modern approach to accommodate the shift to service consumption, automation of service delivery, and development of modular, scale-out applications. Rubrik Cloud Data Management is designed with the following principles:

- **Master-less, self-healing architecture:** Rubrik distributes data, metadata, and task management across the cluster to deliver predictive scalability and eliminate performance bottlenecks. The system has its own distributed file system (Atlas) built from the ground up to store and manage versioned data at scale. Tasks are divvied up across cluster nodes based on data location and resource availability. Data is also stored efficiently while delivering resiliency (erasure coding).

- **Distributed metadata and namespace:** Rubrik's Distributed Metadata System operates alongside its cloud-scale file system (Atlas), 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. The 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.

- **Data Operations Platform as a single system of record:** Rubrik aggregates and organizes metadata into a universal ledger that surfaces semantic relationships across your on-prem and cloud data silos. Rubrik provides a unified control plane for all data and all workloads under management, regardless of whether they are on-premises or in the cloud. This enables the use of SLA policies across all enterprise locations.

- **Elastic compute for light-weight data operations:** Rubrik employs proprietary exocompute technology that establishes a bi-directional communication path between your data in your private cloud and Rubrik to exchange metadata. Through the use of cloud-native compute best practices, Rubrik spins up resources as necessary to manage and process data. This circumvents the impracticality of data movement and costs associated with managing multiple replicas of petabytes of data. And by avoiding the tight coupling of data and operations, the data never leaves your environment, reducing compromise to data sovereignty and governance.

- **Policy-driven data management:** Rubrik offers a global SLA policy engine in which users can automate protection of cloud applications, databases, and file sets to business requirements. Rubrik pioneers a declarative policy approach to eliminate the minutiae of scheduling data protection jobs–users simply select the desired snapshot frequency, retention duration, etc.

- **Secure access in self-service environment:** Granular control over user access is defined at a platform level, regardless of location. Rubrik allows self-service access (role-based access control) to empower users to perform their own backup, recovery, and archival services.

- **Consumption and compliance analytics:** Rubrik delivers real-time platform insights on data management, compliance, and capacity planning across your cloud environment. Rubrik Polaris GPS provides full-featured SaaS-based monitoring across all Rubrik clusters in all infrastructures.

- **Easy integration with automation frameworks:** Rubrik's API-first architecture enables automation of all types of data management workflows. Automate all aspects of data lifecycle management with a full-featured RESTful API. Move local data to AWS and intelligently manage cloud data to reflect business needs.

HOW CLOUD DATA MANAGEMENT WORKS AND USE CASES

Deploy Rubrik Cloud Data Management on your choice of infrastructure: plug-and-play appliances, certified third-party hardware platforms, or directly in AWS.

MANAGING CLOUD-NATIVE APPLICATIONS AND DATA

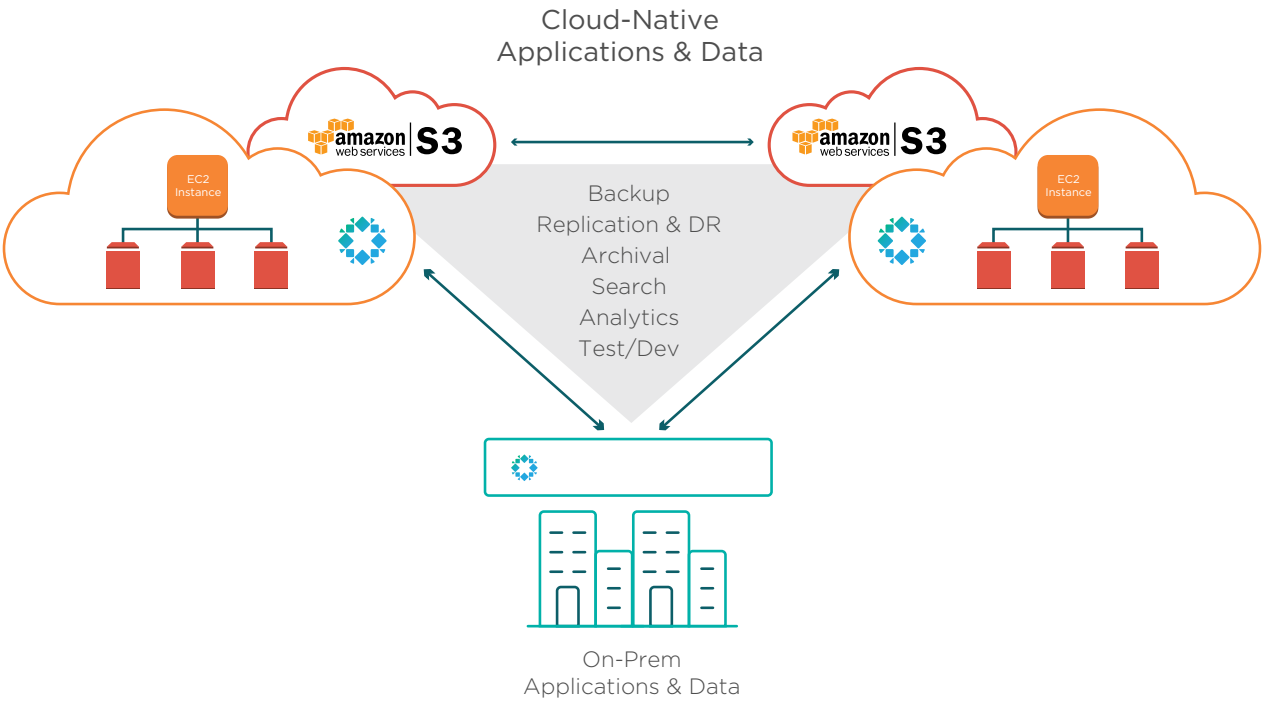
Rubrik Polaris is a SaaS platform that organizes metadata from across your Rubrik deployments and enables global management through a single control plane. Cloud-native backup for Amazon EC2 instances is delivered through Rubrik Polaris. Using AWS APIs, Rubrik automatically spins up a single lightweight Rubrik node in the cloud to index snapshots and relevant metadata in AWS S3. Once complete, the Rubrik node is automatically powered down, saving on both bandwidth and egress costs. Users can protect, index, catalog, and recover data all in AWS, even from an on-premises Rubrik installation.

Alternately, Rubrik can be deployed as a software instance in AWS to orchestrate all critical data management functions. Users can spin up the recommended compute instance and scale easily by growing the Rubrik cloud cluster in lock-step with production cloud data.

Users can instantly locate (with real-time predictive search) and deliver application-consistent recoveries for data born in the cloud, including files, folders, file sets, VMs, and database instances (e.g., Windows, Linux, SQL databases). Users receive actionable insights with Rubrik's rich visual reporting, which allows creation, customization, and sharing of platform analytics on consumption, compliance, and more, across a hybrid cloud environment.

Regardless of how you protect your AWS applications, you will get the exact same user interface and same control plane as for other applications through Polaris. With Rubrik, you have the power to export data across regions, and you can take full advantage of lower-cost S3 and Glacier storage for archival.

Figure 1: Cloud Native Applications & Data



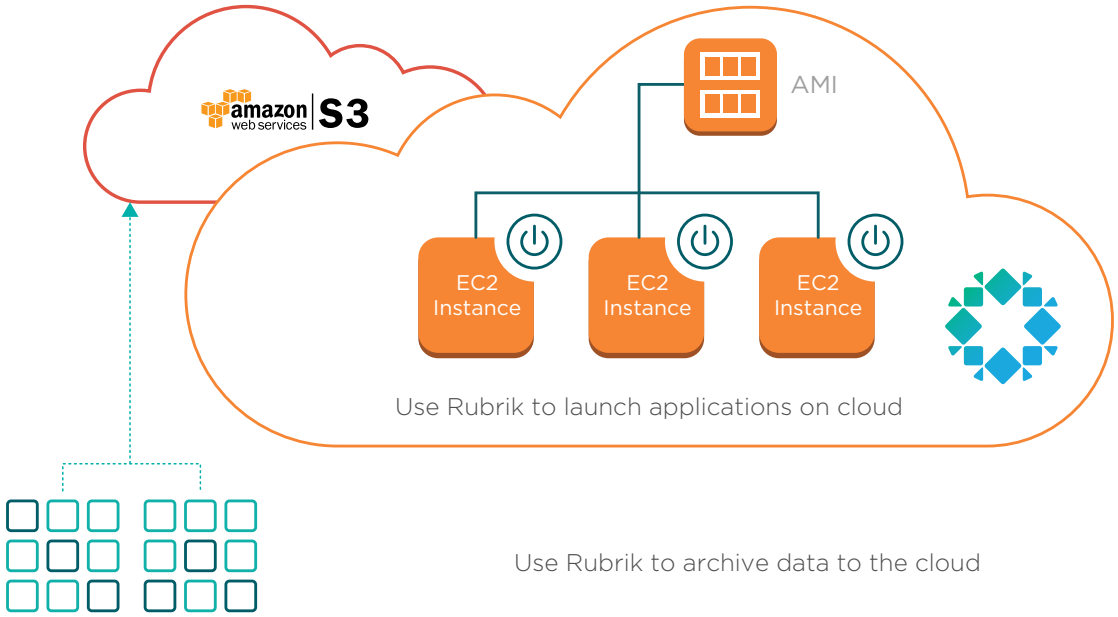
MANAGING HYBRID CLOUD APPLICATIONS AND DATA

Hybrid cloud enterprises can deploy Rubrik to manage applications on-premises while utilizing cloud services for archival, disaster recovery, and test/development.

Enterprises often ease into AWS by utilizing S3 for long-term retention of application data while eliminating tape management complexity. With Rubrik, users can quickly set up archival policies to be applied across their hybrid cloud environment. Rubrik globally indexes all data, no matter where it resides, allowing users to retain quick access to archived data with predictive search. Users can instantly locate a file (rather than the entire data set) and download to restore in any location, saving both bandwidth and egress costs.

Enterprises can also leverage Rubrik to launch applications in AWS for application mobility, and app restores for recovery from downtime and test/dev. To power on applications in AWS initially, users will configure the desired security group and virtual private cloud (VPC) details. Rubrik scans the configuration file of a VM to understand its characteristics (compute, memory, storage, etc.) and recommends a compatible cloud instance type. At this point, Rubrik begins constructing a cloud instance from data stored in the cloud storage service. A single ephemeral, lightweight Rubrik node is automatically created in the target VPC to begin conversion of the VM into a cloud instance. This prevents the need for any data to exit the cloud region, saving both bandwidth and egress costs. Once the conversion is completed, the Rubrik node powers down and is purged until needed again.

Figure 2: App Instantiation in AWS



DATA MANAGEMENT FOR CLOUD-NATIVE APPLICATIONS	
Backup & Recovery of Cloud-Native Applications	Get up and running in minutes with backup for EC2 delivered as-a-service. Protect, index, catalog, search and recover data in AWS. Automate protection across hundreds of cloud accounts. Apply SLAs with granularity down to the VM or tag-level. Leverage an incremental-forever approach to minimize cloud storage costs. Leverage a single, consumer-grade UI to protect, manage and secure data and workloads across AWS, on-premises, and edge locations"
Cloud-Native Data Archival	Archive cloud-native data to AWS S3. Ensure instant accessibility of archived data with Rubrik's real-time predictive search.

DATA MANAGEMENT FOR HYBRID CLOUD APPLICATIONS	
App restores to AWS	Use AWS to recover on-demand from outages, regardless of where you run Rubrik or store data. Rubrik can automate the conversion of VMs, or cloud-based object storage like S3, into compute instances running on EC2. Whether your applications are on-prem or in AWS, you can move on from your largely idle DR site.
Migrating Test/Dev to AWS	Migrate existing on-premises applications to AWS for test or development tasks. Spin up test or dev instances from on-prem VM backups – then get rid of them when you're done.
Replication – On-Prem to Cloud and Cross-cloud	Bi-directional replication is available from an on-premises Rubrik cluster to AWS. Or you can replicate data within a multi-cloud environment.
Data Archival	Send your application data to AWS S3 for long-term retention while retaining immediate access with predictive search.

ENVIRONMENT SUPPORT CONSIDERATIONS	
Amazon Web Services (AWS)	**m4.xlarge instance** Minimum of 4 instances (nodes), starting at 3 TB per node, for overall beginning usable capacity of 8 TB (erasure coding)
Archival Locations	Public Cloud: Amazon S3/S3-IA/S3-RRS/Glacier – including all government cloud options.

Global HQ
1001 Page Mill Rd., Building 2
Palo Alto, CA 94304
United States

1-844-4RUBRIK
inquiries@rubrik.com
www.rubrik.com

20200430_v2