TECHNICAL REFERENCE

Technical Practitioners Guide to the GDPR with Rubrik

TABLE OF CONTENTS

Released June 2018 v1

INTRODUCTION

The General Data Protection Regulation (GDPR), Europe's new data privacy regulation, went into effect on May 25th of 2018. GDPR has far reaching consequences on an organisation's processes around the handling of personal data. Compliance with GDPR requires data processors and data controllers to implement appropriate technical and organisational measures. This document is intended to highlight certain aspects of Rubrik's Cloud Data Management platform and its relevance as a technical measure in light of the GDPR.

Please note that we are not suggesting that using Rubrik's Cloud Data Management Platform will make you GDPR compliant, as GDPR compliance is a multifaceted undertaking that requires organization-wide collaboration. While there is no silver bullet for GDPR compliance, however, Rubrik can assist with GDPR compliance efforts. Furthermore, although we reference certain GDPR articles in this document, we are not implying that Rubrik's Cloud Data Management Platform addresses all legal requirements contained in the articles referenced. In addition, this document should not be considered legal advice or a legal guide to compliance with the GDPR overall.

RUBRIK TECHNOLOGY AND THE GDPR

Availability and Resilience

Data controllers and data processors shall implement measures that ensure ongoing availability and resilience.

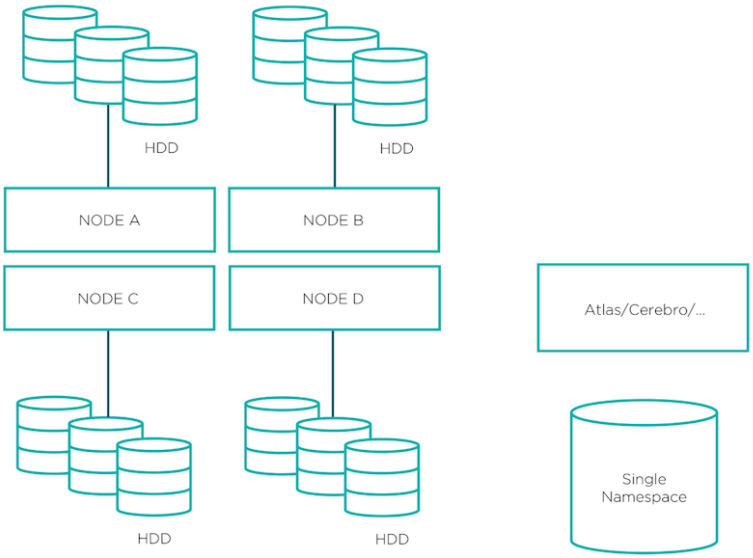
Article 32

HOW RUBRIK CAN HELP »

LOCAL REDUNDANCY

Rubrik's core architecture is a distributed, highly-redundant, masterless-cluster wherein multiple nodes provide local and global resiliency in case of component failures. For the purposes of this discussion we mostly refer to Rubrik's hardware appliance, but the software only approach, which would be implemented with a public cloud provider for example, is very similar in concept. Each Rubrik appliance consists of a minimum number of nodes. The nodes reside in a 2U chassis with no active components and redundant power supplies. Each node has a number of direct attached storage devices. Each storage device contributes to a global storage namespace that is available across the entire cluster. Any node is capable of performing all tasks that are required for operation of the cluster. Therefore, in the event of one or more disk or node failures, the cluster is still operational and data processing can still be performed.

Rubrik uses its own distributed file system called Atlas. Once data is written to Atlas it is immutable, which makes the data immune from ransomware activity. Atlas sits underneath all of the Rubrik Cloud Data Management services, which in turn are running as individual processes distributed across all nodes. The Rubrik solution was designed for data integrity and resiliency.



Rubrik does not rely on traditional RAID architectures, which have become non-viable due to the increasing capacity of hard disks and their steady Unrecoverable Read Error (URE) rate, leading to prolonged rebuild times and increased risk of additional drive failure during rebuild potentially causing the loss of an entire RAID set. We believe this is an unacceptable risk as you backup system is meant to function as the "storage of last resort". Instead of continuing to use RAID or RF2/3 mirroring, Rubrik uses Erasure Coding (4,2) with a specific implementation of Reed-Solomon algorithms to improve performance, provide resiliency, and use space efficiently. In the event of disk failure Erasure Coding enables automatic data rebuild to return the cluster to full protection quickly. Additionally Atlas never updates a single block in a stripe but always performs a full stripe write with verification after the write operation to avoid potential data loss during a power failure.

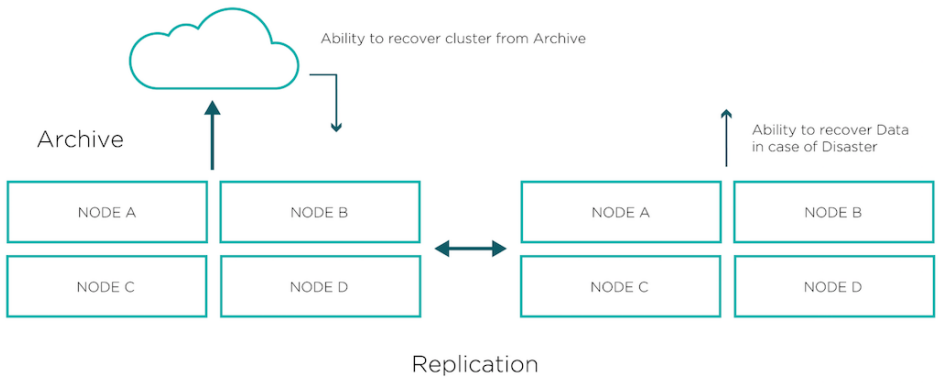
Rubrik protects metadata in multiple ways. For fast processing, metadata is held on the local SSD in each node of the cluster, and for resiliency, metadata is distributed across the cluster via three-way replication and backed up on the local hard drives within in each node. To protect against corruption, multiple copies of versioned metadata are stored and replicated throughout the system.

Atlas also performs continuous validation of data in the cluster; data that enters the cluster is checksummed to verify integrity and the file system uses CRC both at the stripe and chunk level. Stripe level checksumming is used to protect against memory corruption software bugs, and chunk checksumming is used to protect against bit rot. When data is read, the checksum is validated. If verification fails data will be automatically repaired from other copies, additionally, a background scanning process looks for data corruption or inconsistency to prevent unrecoverable read errors.

Fingerprinting algorithms, in which large data items are mapped to shorter bit strings, are employed as a more rigorous end-to-end check. These fingerprints are leveraged during data ingest, replication, and archiving, to ensure the content of the data does not change.

GLOBAL REDUNDANCY

Rubrik offers Replication as part of its core platform; two or more Rubrik clusters can replicate data across multiple sites to protect against local site failure. Additionally, Rubrik can also store data in an archive location and recover from said archive location in the event of a local site failure when data was not replicated to another site.



Protection against Accidental Loss or Damage

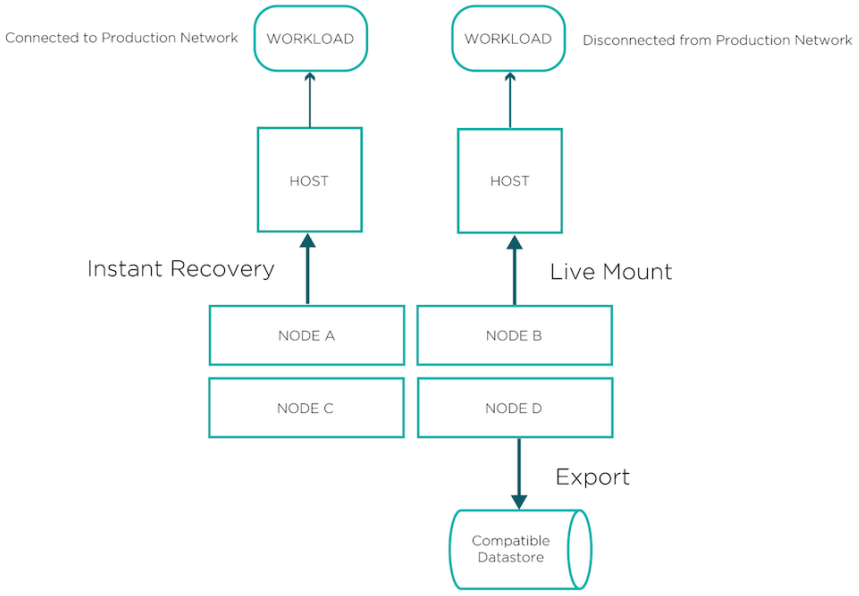
Using appropriate technical or organizational measures, personal data must be protected against accidental loss, destruction or damage.

Article 5

HOW RUBRIK CAN HELP »

INSTANT RECOVERY AND LIVE MOUNT

Besides the redundancy of the platform itself, data managed by Rubrik CDM can be easily and instantly recovered in case of incident by leveraging the Instant Recovery or Live Mount capability.



Instant Recovery replaces the source virtual machine with a fully functional point-in-time copy. The Rubrik cluster powers offs and renames the source virtual machine (in case post-mortem investigation is required) and assigns the name of the source virtual machine to the recovered virtual machine. The Rubrik cluster powers on the recovered virtual machine and connects the recovered virtual machine to the source network. The Rubrik cluster is the datastore for the recovered virtual machine.

A Live Mount creates a new virtual machine from a point-in-time copy of the source virtual machine. The recovered virtual machine uses the Rubrik cluster as its datastore. The Rubrik cluster assigns a new name (consisting of the original vm name appended with the timestamp of the chosen point-in-time copy) and powers it up. The Rubrik cluster does not connect the recovered virtual machine to a network. This ensures that the data is not accessible outside of the administrator preparing the restore of the dataset. The Rubrik cluster sets the protection state of the new virtual machine to Do Not Protect which ensures no copy of previously determined unneeded data is stored again.

In addition this process can be automated. For example, connecting the Live Mounted virtual machine to a specific shielded network and starting certain processes around data deletion.

Encryption

Data controllers and data processors shall implement measures to ensure encryption of personal data.

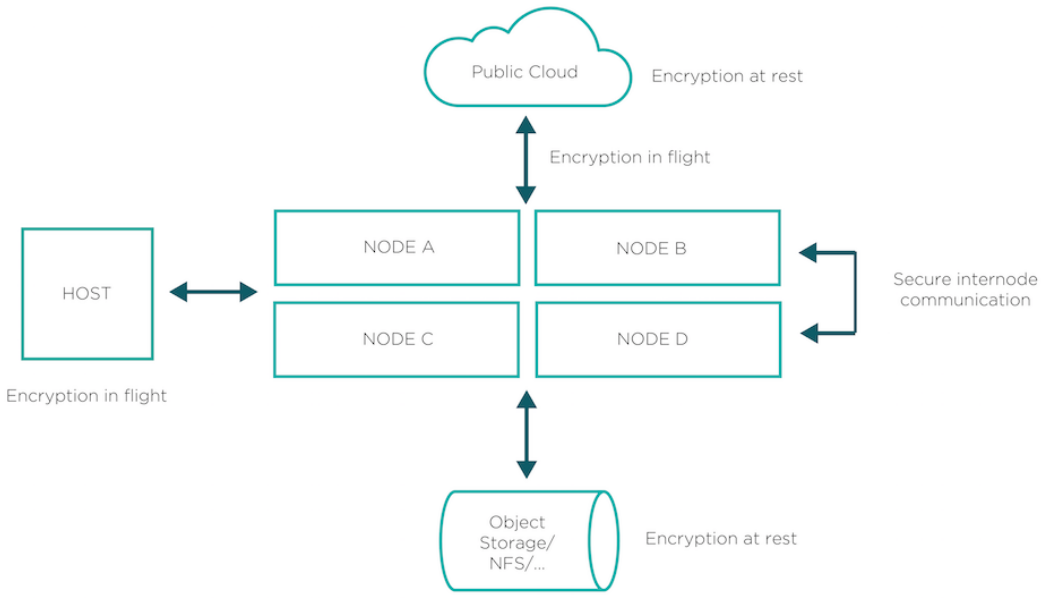
Article 32

HOW RUBRIK CAN HELP »

SECURE BY DESIGN

Rubrik offers a comprehensive portfolio of solutions for security conscious organizations, including those in government, financial, legal, and health care sectors, to address rigid data protection policies regarding classified, confidential, and personally identifiable information (PII). With Rubrik, customers can ensure their data is protected even in the event of a physical theft or breach.

Rubrik implements end-to-end encryption, ingesting data via an encrypted tunnel, storing data using encryption-at-rest while on the cluster itself, offloading data via an encrypted tunnel to archive locations, and again using encryption-at-rest while storing data in said archive location. It also employs secure communication between nodes to avoid eavesdropping.



Protection against Unauthorised Processing

Personal data shall be processed in a way that protects against unauthorised or unlawful processing.

Article 5

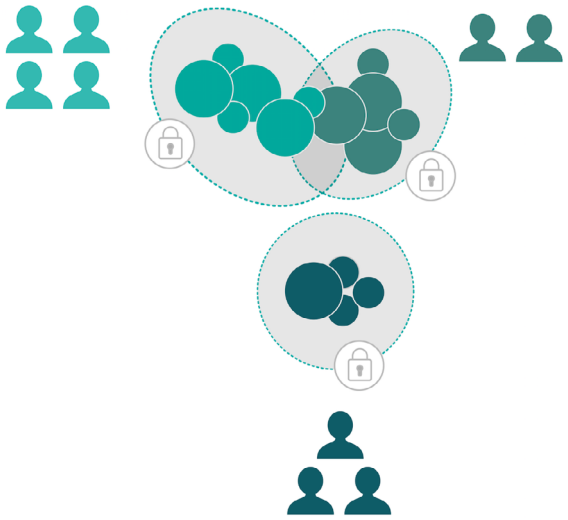
HOW RUBRIK CAN HELP »

RBAC AND SECURE MULTI-TENANCY

Rubrik supports Role Based Access Control and Secure Multi-tenancy to safeguard against unauthorised access to data. Additionally if data is approached outside the control of the Rubrik CDM platform, for example in an archive location, encryption further protects against unwanted access to the data itself.

Rubrik's multi-tenancy is at the object level. You can think of a Rubrik Cluster as a collection of objects: sources from where data is getting backed up, targets where backups are stored, and security principals (users and service accounts) that glue those relationships. As a managed service provider, you can host a Rubrik Cluster that grows linearly with your backup business. With Rubrik CDM's object-level multi-tenancy, you can create virtual instances of Rubrik Cluster for each of your tenants. These virtual instances are aptly named as organizations. An organization can have a dedicated set of sources, targets, and security principals. Or some organizations may share a few security principals, as in the case of MSP staff providing Managed Backup Services. Several organizations might be business subsidiaries sharing a common archival account. All these organizations are secure and isolated while being served from the same Rubrik CDM cluster.

Rubrik's object-level multi-tenancy is extremely useful in large enterprise organizations as well. You may want to delegate VMware vSphere protection to virtual infrastructure admins, Windows systems, and Hyper-V hosts managed by system admins; your SQL Server and Oracle managed by DBAs; and so on. Simply create 'organizations' for them. When they login, they will see only what they should. And unlike legacy solutions, you have the flexibility to delegate entire management responsibility or selected operations. With Rubrik, one UI manages everything; its context changes based on who logs in.



Storage Limitation

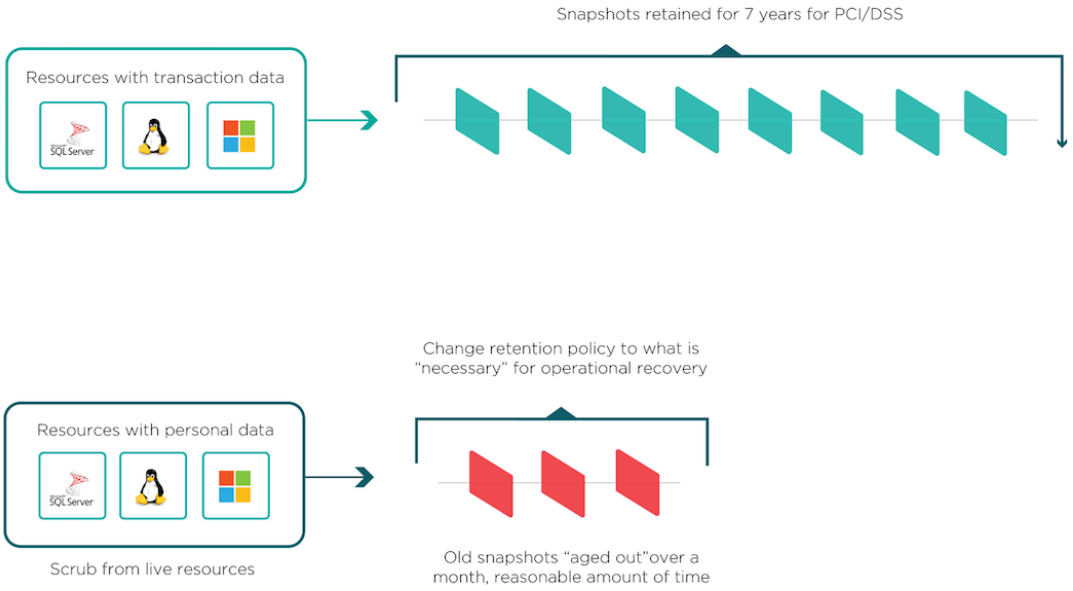
Personal data shall be kept in identifiable form no longer than is necessary for the purposes for which the data are processed.

Article 5

HOW RUBRIK CAN HELP »

SLA BASED POLICIES

Rubrik employs business level SLAs that dictate how data should be managed end-to-end. These SLAs can be easily attached to all supported data sources and ultimately determine how long data is kept. This makes it straightforward to, for example, attach a 30 day SLA to data that has been identified as containing personal information; after 30 days that data is automatically purged from the cluster and no longer available. An additional longer-term retention SLA can be set for data that has a longer legal retention requirement in order to enable longer recovery periods since backup needs to function as storage of last resort in case of calamities.



Another aspect of implementing SLA based policies is verifying adherence to them. This is accomplished on a local level by use of the built-in Rubrik Envision platform that provides monitoring and analytics of events throughout the cluster. If you have deployed Rubrik Clusters in multiple locations, including software only configuration in branch offices or Public Cloud environments, you can monitor compliance with the SLAs via Rubrik Polaris GPS.

Right to be Forgotten

Data subjects have the right to obtain erasure of personal data from the data controller.

Article 17

HOW RUBRIK CAN HELP »

AVOIDING DELETION OF BACKUP DATA

Data classification will determine which datasets contain personal information. By employing the aforementioned shorter-term SLA approach for these specific data sets you can avoid the need to delete data out from your backup.

SELECTIVE RESTORES

In the event that data that needs to be restored contains personal data that was previously removed from the primary systems you can employ Live Mount or Export to selectively recover data.

LIVE MOUNT OF VIRTUAL MACHINES

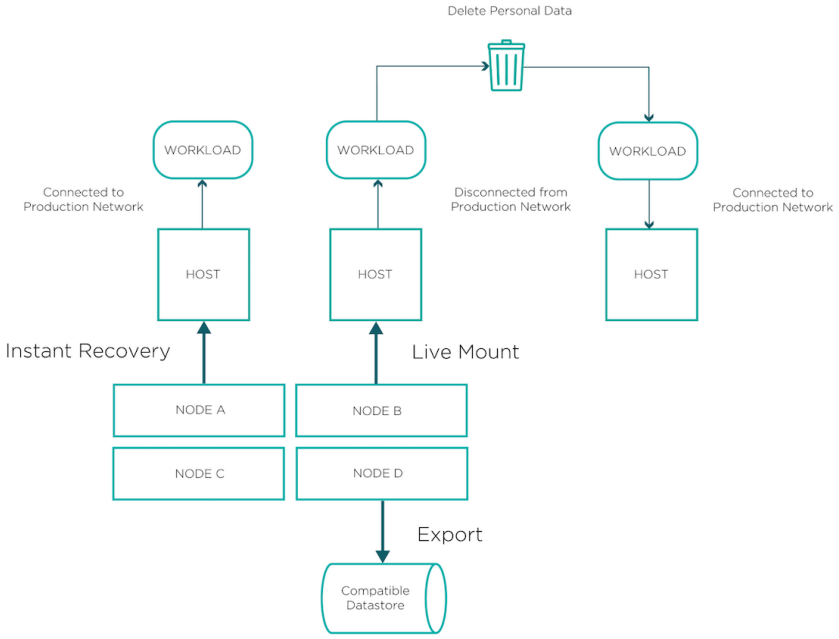
When using Rubrik Live Mount we attach a portion of the Rubrik datastore to the virtual host over NFS but can leave the VM network disconnected. This allows the administrator to first inspect the recovered virtual machine and, if necessary, delete previously identified personal data before starting the final restore back to production.

LIVE MOUNT OF SQL DATA

When using Rubrik Live Mount we attach a portion of the Rubrik datastore to the SQL host over SMBv3. This brings up a copy of the SQL database that can then be selective restored to the production database using SQL commands.

EXPORTING DATA

When using the Export feature of Rubrik, an administrator can export to a controlled environment, and delete previously identified personal data before restoring data back into production.



CONCLUSION

Rubrik Cloud Data Management has several technical capabilities that can assist with certain aspects of GDPR compliance. By leveraging Rubrik's highly resilient and secure platform, data is better protected and more readily available in case of issues. No technical solution alone will provide sufficient coverage for all aspects of GDPR compliance. Instead, this document focuses on practical approaches you can take as part of your overall GDPR compliance strategy.

ABOUT THE AUTHORS

Filip Verloy is Field CTO for EMEA at Rubrik. He is a VMware vExpert, Cisco Champion and maintains his blog at filipv.net. You can find him on Twitter @filipv.

Matt Noe is a Product Manager at Rubrik focused on solving information Governance challenges companies face. Prior to joining the product team, he was an engineer at Rubrik. You can find him on Twitter @m_noe11.