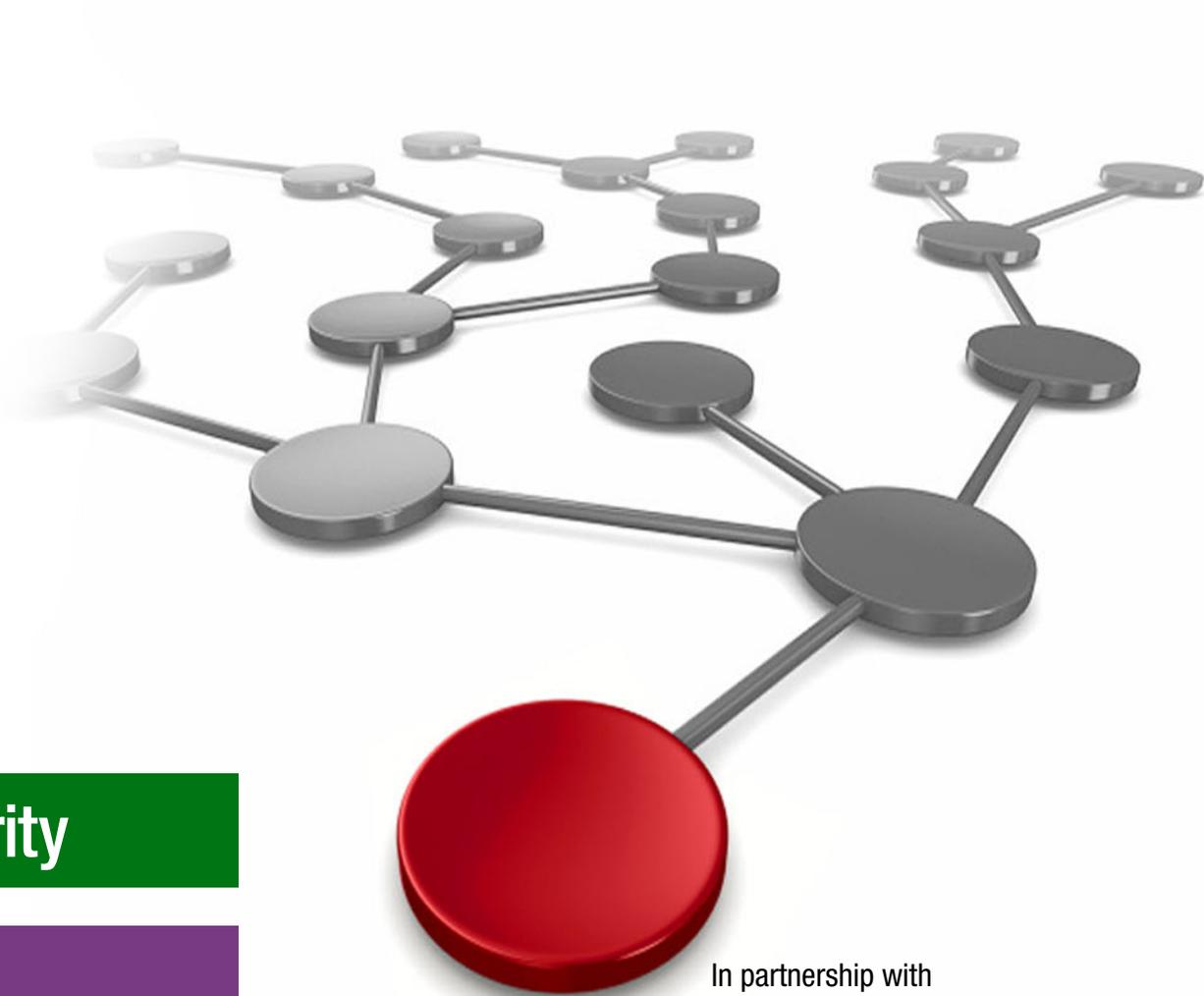


IBM Spectrum Scale Functionality to Support GDPR Requirements

Sandeep R. Patil
Clodoaldo Barrera
Carl Zetie
Felipe Knop
Nils Haustein



 **Security**

Storage

In partnership with
IBM Academy of Technology



IBM Spectrum Scale functionality to support GDPR requirements

European Union (EU) General Data Protection Regulation (GDPR) compliance involves personal data and its protection (article 4, section 1) by any organization that conducts business with personal data of data subjects, in or from the 28 EU member states. GDPR requirements include compliance, data protection, and personal data, including governance, accounting, privacy, data breach procedures, cross border data flow, and other responsibilities across different stakeholders within the organization.

More importantly, compliance requirements start with defined *processing activities* on personal data, which can then require GDPR duties like obtaining consent and restricting data to its permitted use. Organizations cannot achieve compliance by just using specific products or solutions, rather the usual compliance challenge of organizational change across people, policy, and processes is needed.

From an IT point of view, the overall GDPR compliance requirements cover the entire solution stack including applications, middleware, platforms, and infrastructure, especially if any of these are directly or indirectly dealing with personal data. Therefore, there is not going to be a “one size fits all” GDPR solution for businesses. The role of the IT solutions is to enforce the correct handling of personal data per identified processes by the establishment. Each element of the solution stack need to address the objectives that are appropriate to the data it handles.

Typically, personal data resides either in form of structured data (such as databases) or unstructured data (such as files, text, documents, and so on). This paper specifically deals with unstructured data and storage systems that are used to host unstructured data. For the overall approach, the [IBM Pathways for GDPR readiness](#) white paper is a good starting point for businesses to prepare for the implementation of GDPR.

For unstructured data storage in particular, some key attributes enable the overall solution to support compliance with GDPR. Because personal data subject to GDPR is commonly stored in an unstructured data format, a scale-out file system like IBM® Spectrum Scale provides essential functions to support GDPR requirements. The following sections highlight some of the key compliance requirements and explain how IBM Spectrum™ Scale helps to address these concerns:

- ▶ **General Compliance Requirement:** GDPR applies across all relevant unstructured data in an organization. IBM Spectrum Scale™ offers a single global namespace that can store, manage, and help protect unstructured data. This technique helps avoid the creation of independent data islands and consolidates management and compliance activities in one system with a single point of control, reporting, and auditing.

- ▶ **GDPR Requirement Article 32 (Secure personal data):** Securing personal data of EU residents is one of the key requirements of GDPR. One way to accomplish this is using data encryption (which includes secure data at rest and secure data in flight). Furthermore, GDPR article 17 (Right to erasure) requires businesses to address right to erasure of data that is categorized as personal.

IBM Spectrum Scale provides support for file and object encryption at rest and in transit over the supported access protocols. Data at-rest encryption is managed by using encryption keys and encryption policies, and irreversible delete through “crypto-shredding.” This technique involves destruction of the key required to decrypt data. IBM Spectrum Scale encryption software modules to secure data-at-rest are certified according to the Federal Information Processing Standard (FIPS) 140 Publication Series.

- ▶ **GDPR Requirement Article 15 (Right of Access):** This article states the need to control and audit access to data categorized as personal data through mechanisms such as secure authentication, authorization, and audit logging. Controlling data access starts with a proper authentication function that ensures the identity of the user. Authorization of access to data ensures that the authenticated user has rights to access the data. Auditing monitors data access operations and stores these audit trails in a protected fashion.

IBM Spectrum Scale supports authentication with industry standard directory servers over protocols and provides a rich set of ACLs across all its access interfaces to control authorization. These authorization capabilities also help meet certain requirements of Article 25 (Data protection by design and by default). Applications and users can use them to ensure that the personal data is restricted according to the required norms. IBM Spectrum Scale 5.0, introduced in November 2017, has auditing capability to monitor file access to a defined set of data. The resulting audit logs are stored in an immutable file set.

- ▶ **Other GDPR Requirements:** GDPR clearly distinguishes data that is to be categorized as personal data and has a set of rules that needs to be followed to meet compliance. Hence, the unified file and object storage should be able to manage data according to the data categorization to support the compliance needs of the organization. These policies provide key capabilities in support of GDPR compliance, including the following:

- Encrypt selected categories of data
- Irreversible delete of selected categories of data
- Efficiently delete selected data sets based on rules or policies
- Place/Move selected data to specific types of underlying storage
- Mark selected data sets (including audit logs) as immutable
- Retain selected data for a specific period in a non-writable and non-erasable manner, and discard selected data by policy when the retention period expires
- Ensure high availability of access to the data

IBM Spectrum Scale has comprehensive information lifecycle management (ILM) policies to manage the lifecycle of sets of data. Different data sets can be assigned separate rules and policies, allowing separate handling of personal and non-personal data. This technique includes file encryption and secure deletion, transparent movement of files to different storage devices including tape and cloud object storage, backup of files, monitoring and logging of file operations, and retaining files in an immutable manner over a specified retention period. Hence, IBM Spectrum Scale can be used to store all kinds of unstructured data and enforce the most appropriate storage management policies for each category, and ensure high availability of access to the data to support GDPR compliance.

Other resources for more information

The following resources are available for additional information:

- ▶ IBM GDPR readiness journey and GDPR capabilities and offerings
<https://ibm.com/gdpr>
- ▶ How IBM Storage supports GDPR
<https://ibm.biz/nordics-gdpr>
- ▶ The *IBM Pathways for GDPR readiness* white paper
<https://www.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=ASW12436USEN>
- ▶ *IBM Spectrum Scale Security*, REDP-5426
<http://www.redbooks.ibm.com/abstracts/redp5426.html>
- ▶ Understand the GDPR FAQ on the European Union official site
https://europa.eu/european-union/index_en
- ▶ *Impact of General Data Protection Regulation GDPR on systems & storage* white paper
<https://ibm.biz/gdpr-storage-paper>
- ▶ IBM Spectrum Scale
<https://www.ibm.com/us-en/marketplace/scale-out-file-and-object-storage#product-header-top>
- ▶ IBM Spectrum Scale at IBM Knowledge Center
https://www.ibm.com/support/knowledgecenter/en/STXKQY_5.0.0/com.ibm.spectrum.scale.v5r00.doc/b11ins_intro.htm

Notice: Clients are responsible for ensuring their own compliance with various laws and regulations, including the European Union General Data Protection Regulation. Clients are solely responsible for obtaining advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulations that may affect the clients' business and any actions the clients may need to take to comply with such laws and regulations. The products, services, and other capabilities described herein are not suitable for all client situations and may have restricted availability. IBM does not provide legal, accounting or auditing advice or represent or warrant that its services or products will ensure that clients are in compliance with any law or regulation.

Authors

This guide was produced by a team of specialists from around the world working with the International Technical Support Organization (ITSO).

Sandeep R. Patil is a Senior Technical Staff Member who works as a Storage Architect with IBM System Labs. He has over 15 years of product architecture and design experience. Sandeep is an IBM Master Inventor, an IBM developerWorks® Master Author, and a member of the IBM Academy of Technology. Sandeep holds a Bachelor of Engineering (Computer Science) degree from the University of Pune, India.

Clodoaldo Barrera is a Distinguished Engineer and the Chief Technical Strategist for IBM Systems Storage in San Jose California. His responsibilities include development strategy for IBM Flash disk and tape subsystems, SAN, NAS, Object Storage, and software for software-defined storage and storage cloud solutions. Mr. Barrera holds a BS in Mathematics and an MS in Electrical Engineering from Stanford University, and is a Senior Member of the IEEE Computer Society.

Carl Zetie is the offering manager for Spectrum Scale. He is an IT veteran whose roles have included development, pre- and post sales technical support, marketing, and strategy. After spending most of his career working with enterprise software development and release tools, he moved into storage software where he now makes his home.

Felipe Knop is a Senior Technical Staff Member working in software development with the IBM Spectrum Scale team, where he is the File System scrum core architect and also focuses on cluster management and file encryption. In addition, he was the security architect and the technical release lead. He has significant experience in the architecture and development of complex distributed subsystems in IBM Spectrum Scale and previously in Reliable Scalable Cluster Technology (RSCT). He holds a Ph.D. degree in computer science from Purdue University.

Nils Haustein is a Senior Technical Staff Member at IBM Systems group responsible for design and implementation of backup, archiving, file, and object storage solutions in EMEA. He co-authored the book *Storage Networks explained*. As a leading IBM Master Inventor, he has created more than 160 patents for IBM and is a respected mentor for the technical community worldwide.

Thanks to the following people for their contributions to this project:

Larry Coyne
International Technical Support Organization, Tucson Center

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Stay connected to IBM Redbooks

- ▶ Find us on Facebook:
<http://www.facebook.com/IBMRedbooks>
- ▶ Follow us on Twitter:
<http://twitter.com/ibmredbooks>
- ▶ Look for us on LinkedIn:
<http://www.linkedin.com/groups?home=&gid=2130806>
- ▶ Explore new IBM Redbooks® publications, residencies, and workshops with the IBM Redbooks weekly newsletter:
<https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm>
- ▶ Stay current on recent Redbooks publications with RSS Feeds:
<http://www.redbooks.ibm.com/rss.html>

Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at <http://www.ibm.com/legal/copytrade.shtml>

The following terms are trademarks or registered trademarks of International Business Machines Corporation, and might also be trademarks or registered trademarks in other countries.

Redbooks (logo) ®
developerWorks®

IBM®
IBM Spectrum™

IBM Spectrum Scale™
Redbooks®

The following terms are trademarks of other companies:

Other company, product, or service names may be trademarks or service marks of others.



REDP-5489-00

ISBN 0738456764

Printed in U.S.A.

Get connected

