HIPAA Compliance for Healthcare Workloads on IBM Spectrum Scale

Sandeep R. Patil
Sandeep Zende

In partnership with IBM Academy of Technology

Security

Storage
HIPAA Background

When technology workloads process healthcare data, it is important to understand HIPAA compliance and what it means for the technology infrastructure in general and storage in particular. The Health Insurance Portability and Accountability Act (HIPAA) is US legislation that was signed into law in 1996. HIPAA was enacted to protect the health insurance coverage, but was later extended to ensure protection and privacy of electronic health records and transactions. In simple terms, it was instituted to modernize the exchange of healthcare information and how the Personally Identifiable Information (PII) maintained by the healthcare and healthcare related industries should be safeguarded.

From a technology point of view, one of the core requirements of HIPAA is the protection of Electronic Protected Health Information (ePHI\(^1\)) through physical, technical and administrative defenses. From a non compliance point of view, the Health Information Technology for Economic and Clinical Health Act (HITECH) added extra protections to HIPAA and increased the penalties ranging from $100 USD to $50,000 USD per violation. Today, HIPAA compliant solutions have become a norm in healthcare industry worldwide.

This Redpaper will describe HIPPAA compliance requirements for storage and how security enhanced Software Defined Storage is designed to help meet those requirements. We will correlate how Software Defined IBM® Spectrum Scale security features address the safeguards specified by the HIPAA Security Rule.

\(^1\) Per HIPAA compliance is “any electronic information that is created or received by a health care provider that relates to the past, present or future physical or mental health of an individual and that identifies the individual”.
HIPAA Compliance and Storage

HIPAA compliance has overarching requirements that span the entire healthcare solution stack that deals with health related information. One of the important building blocks of this solution stack is the underlying storage where this sensitive data (i.e. ePHI) resides. Hence, it is important to understand what are the HIPAA requirements for storage solutions used to host ePHI related data. Moreover, many institutes are deploying artificial intelligent (AI) workloads that are mining the health data to get meaningful insights for enhancement in healthcare and precision medicine. These AI workloads drive the need for Secure AI solutions which starts with addressing the HIPAA requirements at the storage level and then building up the Secure AI solution stack from there.

What are these requirements? Basically, the Privacy and Security Rule of HIPAA mandates that health care based technology solutions ensure privacy for health information and must also comply with different types of safeguards defined by HIPAA. These rules are the primary set that one needs to understand from a storage perspective.

Security Enhanced Software Defined Storage

According to the Storage Network Industry Association (SNIA), data security in the context of storage systems is responsible for safeguarding the data against theft, prevention of unauthorized disclosure of data, prevention of data tampering, and accidental corruption. This process ensures accountability, authenticity, business continuity, and regulatory compliance.

Security for storage systems can be classified as follows:
- Data storage (data at rest, which includes data durability and immutability)
- Access to data
- Movement of data (data in flight)
- Management of data

IBM Spectrum Scale

IBM Spectrum® Scale is a software-defined storage system for high performance, large-scale workloads on-premises or in the cloud. It has been developed and enhanced by IBM for 30 years to address data management at massive scale for large, global organizations. It simplified management, helps to lower capital and operational costs, is easy to grow, and has important, enterprise class features. Its most common use areas are in AI and Deep Learning, Big Data Analytics, Content Repository, Private Cloud, and Compute Clusters. It is also commonly used for data optimization and resiliency for archive, high speed back-up, and disaster recovery, and Information lifecycle management.

IBM Spectrum Scale addresses HIPAA requirements by:
- securing data at rest by protecting data at rest with snapshots, backups, and immutability features
- secure access to data by providing secure management of data, and secure access to data by using authentication and authorization across multiple supported access protocols. Supported protocols include POSIX, NFS, SMB, Hadoop, and Object (REST).
- securing data in flight by providing encryption of management as well as user data over wire on supported protocols.
automated data management, it is equipped with powerful information Lifecycle Management (ILM) tools that can help administer unstructured data by providing the correct security for the correct data.

Figure 1 shows the IBM Spectrum Scale Security Features.

![IBM Spectrum Scale Security Features](image)

This Redpaper explains the different safeguards specified by the HIPAA Security Rule and describes which IBM Spectrum Scale security features can be mapped to them. The paper correlates IBM Spectrum Scale security features to the needs of the healthcare sector and discusses why it is the most suitable file and object storage for healthcare workloads.

**HIPAA Security Rules and IBM Spectrum Scale Security**

The HIPAA Security Rule specifies three types of controls for compliance:

- **Administrative Safeguards (Section - 164.308):** To document formal policies and practices for data protection, including the organization's security management process and implementation specification.

- **Physical Safeguards (Section - 164.310):** To protect data from the hazards of fire, weather, the environment, or intrusion.

- **Technical Safeguards (Section - 164.312):** To protect data and control access to information by individuals and guard against unauthorized access through a network.

Overall, these safeguards require the health care industry to measure and keep personal health information secure as well as to decrease the means of tampering with, destruction of, or inappropriate access to ePHI. Let us filter and shortlist the specific safeguards that apply for storage systems while the adherence of the overall safeguards is responsibility of the overall solution and the practitioners owning the solution.

Applicability of safeguards to storage:

- **Administrative Safeguards (164.308) - Semi-applicable to storage components**
Physical Safeguards (164.310) - Minimal applicability to storage components

Technical Safeguards (164.312) - High applicability to storage components

Table 1 shows the filtered safeguards based on their applicability to file storage and the corresponding IBM Spectrum Scale security features that can be used by healthcare solution architects to address the safeguards from the storage point of view.

**Important:** Depending upon the overall architecture of the healthcare solution, implementation and responsibility of the safeguards can be addressed at different levels of the solution stack. Hence many of the safeguards may not be directly applicable for underlying storage, even if the storage is hosting ePHI. So, while below table gives a high level mapping of the applicable safeguards to the security features of IBM Spectrum Scale, the actual applicability depends on the overall architecture of the healthcare solution and policies set by the practitioners.

<table>
<thead>
<tr>
<th>Parts and sub-parts of Section 164</th>
<th>Section heading and sub-heading</th>
<th>IBM Spectrum Scale security features</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIPAA 164.308(a)(4)(ii)(B,C)</td>
<td>Information access management</td>
<td>IBM Spectrum Scale provides a rich set of POSIX as well as NFSV4 ACL for data authorization that can be used by solutions to address this safeguard.</td>
</tr>
<tr>
<td>HIPAA 164.308(a)(5)(ii)(B)</td>
<td>Protection from malicious software</td>
<td>IBM Spectrum Scale supports antivirus as well as integration with threat detection software like IBM QRadar® which can be used by solutions to address this safeguard.</td>
</tr>
<tr>
<td>HIPAA 164.308(a)(5)(ii)(D) and 164.312(a)(2)(i)</td>
<td>Password management and unique user identification (Procedures for centralized password management and centralized, unique ID for tracking)</td>
<td>IBM Spectrum Scale supports Active Directory/LDAP-based and Kerberos-based centralized authentication and ID management system for comprehensive password management and unique identification which can be leveraged by solutions to address these safeguards.</td>
</tr>
<tr>
<td>HIPAA 164.308(a)(1)(ii)(D), 164.308(a)(5)(ii)(C), and 164.312(b)</td>
<td>Information system activity review, logs, and audit control</td>
<td>IBM Spectrum Scale supports auditing of management activity via CLI as well as GUI. Also supports file audit logging which can be integrated with SIEM like IBM QRadar as well as Elasticsearch. Solutions can leverage these features to address the mentioned safeguards.</td>
</tr>
<tr>
<td>Parts and sub-parts of Section 164</td>
<td>Section heading and sub-heading</td>
<td>IBM Spectrum Scale security features</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>---------------------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>HIPAA 164.312(d)</td>
<td>Person or entity authentication</td>
<td>IBM Spectrum Scale supports authentication via corporate directory servers like AD/LDAP. This allows that the data being accessed is authenticated. Also supports authorization via POSIX and NFSV4 ACL. These features can be used by the solution to address these safeguards.</td>
</tr>
<tr>
<td>HIPAA 164.312 (a)(2)(iii)</td>
<td>Automatic logoff (Session termination mechanisms)</td>
<td>IBM Spectrum Scale is Software Defined storage. One needs to configure the auto logoff on the operating systems on which the IBM Spectrum Scale server and client are installed. IBM Spectrum Scale GUI supports auto logoff. IBM Elastic Storage Server (ESS) also supports auto logoff for its command line and GUI interface.</td>
</tr>
<tr>
<td>HIPAA 164.312(a)(2)(iv)</td>
<td>Encryption and decryption (Mechanism for encryption of stored PHI) Secure Data at Rest</td>
<td>IBM Spectrum Scale supports file system level encryption (encrypting data on disk) powered with policies which allows administrators to select what data should be encrypted and what data should not to be encrypted when stored on the backend disks. This feature can be used by the solution to address this safeguard.</td>
</tr>
<tr>
<td>HIPAA 164.312 (c) (i, ii)</td>
<td>▶ Integrity of electronic protected health information ▶ Mechanism to authenticate and maintain systems integrity</td>
<td>IBM Spectrum Scale supports synchronous mirroring which can be deployed to protect/recover from health data being altered or destroyed without human intervention, such as by electronic media errors or failures. IBM Spectrum Scale authentication as well as authorization ensures the required mechanism to authenticate access to electronic protected information. Additionally, IBM Spectrum Scale supports file audit logging which can be used to corroborate the access details to ePHI. File audit logging can be coupled with integration with IBM QRadar which monitors the access and alteration to the ePHI data and alarms if the access was against the defined policy.</td>
</tr>
</tbody>
</table>
Table 1 explains how IBM Spectrum Scale can be used by solution architects to create a healthcare solution that addresses safeguards mandated by HIPAA and why it should be a preferred choice for file storage.

Additionally, there is a strong need for Cyber Resilience solution to safeguard the ePHI and related data from cyber attacks like ransomware. IBM Spectrum Scale along with IBM Spectrum Protect provides a complete Cyber Resilience solution required for such workloads. For more details refer to IBM Redbooks® Blueprint Cyber Resiliency Solution for IBM Spectrum Scale, REDP-5559.

Continuous threat detection and monitoring of malicious access of ePHI data is another concern that administrators need to take care. IBM Spectrum Scale integration with IBM QRadar provides enhanced cyber security and ensures any access to health data violating the compliance or business policies can trigger real time alerts and actions. For more details refer to IBM Redpaper Enhanced Cyber Security with IBM Spectrum Scale and IBM QRadar, REDP-5560.

Regular snapshot based backup supported by IBM Spectrum Scale is another feature that needs to be considered by the healthcare solution to ensure backup and recovery of the ePHI and related information.

IBM Spectrum Scale software defined storage is available as on-premises deployment, as an IBM Elastic Storage® Server hardware integrated solution as well as for cloud deployment. It excels in hybrid cloud deployment leveraging IBM Spectrum Scale features like Active File management (AFM). In addition it supports containerized workloads with Red Hat OpenShift and Kubernetes. The above mentioned HIPAA safeguards and analysis is applicable for these deployment models thus giving healthcare solution various choices based on their requirements.
The healthcare solution based on IBM Spectrum Scale can also be integrated with IBM Spectrum Discover which allows pluggable deep inspection to enable automate cataloging of healthcare as well as classifying sensitive data falling under regulatory compliance. It also supports tiering data to IBM Cloud™ Object Storage as well as workload scheduling software like IBM Spectrum Computing required in healthcare workloads like genomics.

**Note:** By deploying the IBM Spectrum Scale and using its security features does not mean that your solution will automatically comply to HIPAA. However, it can help the solution get there. Also, most of the safeguards stated in the 164-Section of the Security Rule concentrate on security of ePHI, while the stated features in this paper are generic IBM Spectrum Scale features and not specific to ePHI (though can be applied on it as explained). Hence, it's advised to consider the stated mapping only as reference information. The previous checklist should not be treated as a complete list and may not necessarily be comprehensive from compliance perspective.

**Conclusion**

This article presented relevant portions of the HIPAA Security Rule and explained how healthcare solutions can leverage IBM Spectrum Scale/IBM Elastic Storage Server security features to address the security rule safeguards from a storage perspective. Thus IBM Spectrum Scale known for its superlative performance can not only help healthcare workloads run faster but also run securely and help achieve HIPAA compliance.

**Note:** Clients are responsible for ensuring their own compliance with various laws and regulations, including HIPAA. 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.
Related publications

The publications that are listed in this section are considered suitable for a more detailed description of the topics that are covered in this paper.

IBM Redbooks

The following IBM Redbooks publications provide more information about the topics in this document. Some publications that are referenced in this list might be available in softcopy only:

- *IBM Spectrum Scale Security*, REDP-5425
  http://www.redbooks.ibm.com/abstracts/redp5426.html
- *IBM Spectrum Scale Best Practices for Genomics Medicine Workloads*, REDP-5479
  http://www.redbooks.ibm.com/abstracts/redp5479.html
- *Cyber Resiliency Solution for IBM Spectrum Scale*, REDP-5559
  http://www.redbooks.ibm.com/abstracts/redp5559.html
- *Enhanced Cyber Security with IBM Spectrum Scale and IBM QRadar*, REDP-5560
  http://www.redbooks.ibm.com/abstracts/redp5560.html
- *IBM Spectrum Scale Functionality to Support GDPR Requirements*, REDP-5489
  http://www.redbooks.ibm.com/abstracts/redp5489.html
- *IBM Spectrum Discover: Metadata Management for Deep Insight of Unstructured Storage*, REDP-5550
  http://www.redbooks.ibm.com/abstracts/redp5550.html
- *IBM Hybrid Solution for Scalable Data Solutions using IBM Spectrum Scale*, REDP-5549
  http://www.redbooks.ibm.com/abstracts/redp5549.html

You can search for, view, download, or order these documents and other Redbooks, Redpapers, Web Docs, draft and other materials, at this website:
http://www.redbooks.ibm.com/

Online resources

The following websites are also relevant as further information sources:

- Health Information Privacy
  https://www.hhs.gov/hipaa/for-professionals/security/guidance/index.html
- QRadar: Health Insurance Portability and Accountability Act (HIPAA) Reporting Extension
- Health Insurance Portability and Accountability Act of 1996
- IBM Spectrum Scale Knowledge Center

IBM Elastic Storage Server Knowledge Center

Help from IBM

IBM Support and downloads:
http://www.ibm.com/support

IBM Global Services:
http://www.ibm.com/services

Authors

This paper was produced by a team of specialists from around the world working with the IBM Redbooks, Tucson Center.

Sandeep R. Patil is a Senior Technical Staff Member who works as a Storage Architect with IBM System Labs. He has over 18 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. He is recognized and listed by Wikipedia in the World Wide Prolific Inventors list.

Sandeep Zendee is a storage solutions architect in IBM Systems Group primarily for solutions related to healthcare industry. He has architected and delivered storage migrations, disaster recovery solutions, and software-defined solutions for several clients, including healthcare clients.

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

Larry Coyne
IBM Redbooks, Tucson Center

Dave McDonnell
John Sing
Prashant Sodhiya
Carl Zetie
IBM Systems

David Jenkins
IBM Global Markets
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 IBM Redbooks 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 Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:
- 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) ®
devolverWorks®
IBM®
IBM Cloud™
IBM Elastic Storage®
IBM Spectrum®
QRadar®
Redbooks®

The following terms are trademarks of other companies:

OpenShift, Red Hat, are trademarks or registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries.

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