Draft NIST Special Publication 800-209

Security Guidelines for Storage Infrastructure

Ramaswamy Chandramouli
Doron Pinhas

This publication is available free of charge from:
https://doi.org/10.6028/NIST.SP.800-209-draft

COMPUTER SECURITY

National Institute of Standards and Technology
U.S. Department of Commerce

Draft NIST Special Publication 800-209

Security Guidelines for Storage Infrastructure

Ramaswamy Chandramouli
Computer Security Division
Information Technology Laboratory

Doron Pinhas
Continuity Software
New York, NY

This publication is available free of charge from:
https://doi.org/10.6028/NIST.SP.800-209-draft

July 2020

U.S. Department of Commerce
Wilbur L. Ross, Jr., Secretary

National Institute of Standards and Technology
Walter Copan, NIST Director and Under Secretary of Commerce for Standards and Technology

Authority

This publication has been developed by NIST in accordance with its statutory responsibilities under the Federal Information Security Modernization Act (FISMA) of 2014, 44 U.S.C. § 3551 *et seq.*, Public Law (P.L.) 113-283. NIST is responsible for developing information security standards and guidelines, including minimum requirements for federal information systems, but such standards and guidelines shall not apply to national security systems without the express approval of appropriate federal officials exercising policy authority over such systems. This guideline is consistent with the requirements of the Office of Management and Budget (OMB) Circular A-130.

Nothing in this publication should be taken to contradict the standards and guidelines made mandatory and binding on federal agencies by the Secretary of Commerce under statutory authority. Nor should these guidelines be interpreted as altering or superseding the existing authorities of the Secretary of Commerce, Director of the OMB, or any other federal official. This publication may be used by nongovernmental organizations on a voluntary basis and is not subject to copyright in the United States. Attribution would, however, be appreciated by NIST.

National Institute of Standards and Technology Special Publication 800-209
Natl. Inst. Stand. Technol. Spec. Publ. 800-209, 65 pages (July 2020)
CODEN: NSPUE2

This publication is available free of charge from:
https://doi.org/10.6028/NIST.SP.800-209-draft

Public comment period: *July 21, 2020* through *August 31, 2020*

National Institute of Standards and Technology
Attn: Computer Security Division, Information Technology Laboratory
100 Bureau Drive (Mail Stop 8930) Gaithersburg, MD 20899-8930
Email: sp800-209-comments@nist.gov

All comments are subject to release under the Freedom of Information Act (FOIA).

Reports on Computer Systems Technology

The Information Technology Laboratory (ITL) at the National Institute of Standards and Technology (NIST) promotes the U.S. economy and public welfare by providing technical leadership for the Nation's measurement and standards infrastructure. ITL develops tests, test methods, reference data, proof of concept implementations, and technical analyses to advance the development and productive use of information technology. ITL's responsibilities include the development of management, administrative, technical, and physical standards and guidelines for the cost-effective security and privacy of other than national security-related information in Federal information systems. The Special Publication 800-series reports on ITL's research, guidelines, and outreach efforts in information system security, and its collaborative activities with industry, government, and academic organizations.

Abstract

Storage technology, just like its computing and networking counterparts, has evolved from traditional storage service types, such as block, file, and object. Specifically, the evolution has taken two directions: one along the path of increasing storage media capacity (e.g., tape, HDD, SSD) and the other along the architectural front, starting from direct attached storage (DAS) to the placement of storage resources in dedicated networks accessed through various interfaces and protocols to cloud-based storage resource access, which provides a software-based abstraction over all forms of background storage technologies. Accompanying the evolution is the increase in management complexity, which subsequently increases the probability of configuration errors and associated security threats. This document provides an overview of the evolution of the storage technology landscape, current security threats, and the resultant risks. The main focus of this document is to provide a comprehensive set of security recommendations that will address the threats. The recommendations span not only security management areas that are common to an information technology (IT) infrastructure (e.g., physical security, authentication and authorization, change management, configuration control, and incident response and recovery) but also those specific to storage infrastructure (e.g., data protection, isolation, restoration assurance, and encryption).

Keywords

storage area network; network attached storage; storage array; file storage service; block storage service; object storage service; storage virtualization; software-defined storage; hyper-converged storage; data protection; cloud storage; backup; replication.

Acknowledgements

[TBD]

Call for Patent Claims

This public review includes a call for information on essential patent claims (claims whose use would be required for compliance with the guidance or requirements in this Information Technology Laboratory (ITL) draft publication). Such guidance and/or requirements may be directly stated in this ITL Publication or by reference to another publication. This call also includes disclosure, where known, of the existence of pending U.S. or foreign patent applications relating to this ITL draft publication and of any relevant unexpired U.S. or foreign patents.

ITL may require from the patent holder, or a party authorized to make assurances on its behalf, in written or electronic form, either:

a) assurance in the form of a general disclaimer to the effect that such party does not hold and does not currently intend holding any essential patent claim(s); or

b) assurance that a license to such essential patent claim(s) will be made available to applicants desiring to utilize the license for the purpose of complying with the guidance or requirements in this ITL draft publication either:

    i. under reasonable terms and conditions that are demonstrably free of any unfair discrimination; or

    ii. without compensation and under reasonable terms and conditions that are demonstrably free of any unfair discrimination.

Such assurance shall indicate that the patent holder (or third party authorized to make assurances on its behalf) will include in any documents transferring ownership of patents subject to the assurance, provisions sufficient to ensure that the commitments in the assurance are binding on the transferee, and that the transferee will similarly include appropriate provisions in the event of future transfers with the goal of binding each successor-in-interest.

The assurance shall also indicate that it is intended to be binding on successors-in-interest regardless of whether such provisions are included in the relevant transfer documents.

Such statements should be addressed to: sp800-209-comments@nist.gov
with Subject: sp800-209 PATENT CLAIMS

Executive Summary

Storage, computing, and networking form the three fundamental building blocks of an information technology infrastructure. Just like computing and network technologies, the storage technology has also evolved over the years. Higher capacity storage media and storage system architecture are the two fronts on which the storage technology has evolved. The developments on the second front have enabled the storage services to support many new and evolving computing use cases but have also introduced storage management complexity and many security challenges.

Just like computing and networking, the current landscape of storage infrastructure consists of a mixture of legacy and advanced systems. With this in mind, this document provides an overview of the storage technology landscape, including traditional storage services (e.g., block, file, and object storage), storage virtualization, storage architectures geared for virtualized server environments, and storage resources hosted in the cloud. Descriptions of various threats to the storage resources are also included, as well as analysis of the risks to storage infrastructure and the impacts of these threats.

The primary purpose of this document is to provide a comprehensive set of security recommendations for the current landscape of the storage infrastructure. The security focus areas span those that are common to the entire IT infrastructure, such as physical security, authentication and authorization, change management, configuration control, incident response, and recovery. Within these areas, security controls that are specific to storage technologies, such as network-attached storage (NAS) and storage area networks (SAN), are also covered. In addition, security recommendations specific to storage technologies are provided for the following areas of operation in the storage infrastructure:

- Data protection
- Isolation
- Restoration assurance
- Encryption

Table of Contents

1 Introduction

Storage, computing, and networking form the three fundamental building blocks of any information technology (IT) infrastructure. The storage infrastructure has evolved over the years to become feature-rich in areas such as performance and efficiency due to developments on two fronts. One is the area of storage media, which features solid-state drives (SSD) with high capacity and storage efficiency features (e.g., deduplication, compression, etc.) compared to hard-disk drives (HDD). The second is in the area of storage system architecture using concepts such as storage virtualization. However, development on this second front has also introduced a great deal of management complexity, including the task of providing security assurance.

Briefly tracing the history of storage system architecture shows that the earliest form of digital storage infrastructure is direct-attached storage (DAS) where the storage element or device (e.g., tape, hard disk) is directly attached to the host server without any intervening network. The next evolution of the storage infrastructure is one where the storage resources are intelligently pooled, located across the network, accessed through networking protocols, and accessible by multiple hosts or servers. This type of storage infrastructure is the only way that the data access needs of distributed systems can be supported since application components that need to share data are in different nodes of a network. In this stage of evolution, the storage infrastructure has taken on two forms, depending on the type of networking protocol. In one form, the storage resource is simply a node in a network using common networking technology (e.g., LAN, WAN), while in the other, there is a dedicated network for communicating with all storage resources (e.g., Fibre Channel). An example of the former is network-attached storage (NAS), which provides file-level access to heterogeneous clients across a network using higher level protocols, such as NFS or SMB/CIFS. The latter is exemplified by the storage area network (SAN), which uses a specialized, high-speed network (e.g., Fibre-Channel) that provides block-level access to storage.

A variation of the traditional enterprise storage infrastructure is the emergence of converged and hyper-converged systems (HCI). A converged system involves a preconfigured package of software and hardware in a single hardware chassis for simplified management. However, with a converged infrastructure, the compute, storage, and networking components are discrete and can be separated. Just like a converged system, an HCI combines storage, computing, and networking into a single hardware unit or chassis and has built in a layer of abstraction for managing all three components. In fact, it includes a common software console or management tool for managing all three components. It also includes a hypervisor for virtualized computing, software-defined storage, and virtualized networking bundled together to run on standard, off-the-shelf hardware. The integrated storage systems, hosts, and networking switches are designed to be managed as a single system across all instances of a hyperconverged infrastructure. Further, each hardware unit can be configured to be a node of a cluster to create pools of shared storage resources, thus providing the advantage of a centralized enterprise storage infrastructure.

The next wave of storage evolution involves the introduction of cloud storage, which offers a highly scalable and durable set of storage services that are completely software-defined. Cloud storage services often include:

- Block storage services, which expose software-defined block devices that can be presented to virtual hosts running in the cloud
- Object storage services, which can be mapped to hosts, applications, or even other cloud services
- Scalable shared filesystems, which can allow a scalable set of hosts to access the same file system at a high speed
- A variety of replication, caching, archiving, mirroring, and point-in-time copy services to all of the above

Additional cloud services—such as managed database services, data lakes, memory caches, and messages queues—are also offered, all of which can store stateful and transient data. However, experts are divided over whether to classify them as storage services in and of themselves.

Another type of storage infrastructure is the one that contains interfaces to support the data storage needs of emerging stateful applications that are designed using microservices-based architecture and deployed using containers organized into clusters with container orchestration platforms. These platforms have a standard plug-in mechanism by way of a container storage interface (CSI) that connects the clusters configured by them to different types of persistent storage implementations.

1.1 Scope

This document provides security recommendations for the following storage technologies:

- Traditional enterprise storage technologies classified based on storage service interface type (e.g., block, file, and object)
- Storage systems that have a layer of software abstraction (e.g., software-designed storage and storage virtualization)
- Storage systems designed exclusively for virtualized server environments (eg., storage for VMs and containers, converged and hyperconverged storage systems)
- Storage systems designed with APIs for cloud access

The security recommendations span the following operations:

- Operations that are carried out for other infrastructures (e.g., computing and networking) but the specific tasks are applicable to storage infrastructure, such as physical security, authentication and authorization, audit logging, network configuration, isolation, configuration control, and change management
- Operations that are unique to storage infrastructure, such as data protection and restoration assurance

1.2 Target Audience

The target audience for the security recommendations discussed in this document includes:

- Chief Security Officer (CSO) or Chief Technology Officer (CTO) of an IT department in a private enterprise, government agency, or a cloud service provider who wishes to formulate enterprise- or data center-wide policies for the entire infrastructure, including storage infrastructure
- System or storage administrators who have to set up specific deployment configurations for storage, converged, or virtualized systems

1.3 Relationship to other NIST Guidance Documents

This guidance document focuses on a particular type of infrastructure that provides access to all data resources/services, similar to how the computing infrastructure provides access to computing services and the networking infrastructure provides access to communication services. Hence, some of the security guidance and recommendations related to computing and networking are relevant security strategies for the storage infrastructure discussed in this document. These common recommendations are either included here through a brief description or incorporated by reference. The relevant NIST documents containing recommendations that span all infrastructures (i.e., computing, networking, and storage) are:

- SP 800-125A, Revision 1, *Security Recommendations for Server-based Hypervisor Platforms* (2018)
- SP 800-125B, *Secure Virtual Network Configuration for Virtual Machine (VM) Protection* (2016)

1.4 Organization of this Document

The organization of this document is as follows:

- Section 2 provides an overview of traditional enterprise storage technologies, storage access technologies that provide a level of abstraction, storage architectures tailored for virtualized server environments, and APIs for accessing storage resources in the cloud. This section also provides an overview of certain general principles of storage administration.
- Section 3 explores the threats to storage infrastructure and associated risks. Apart from generic threats such as privilege escalation, credential theft, cracking encryption, malware, and ransomware, storage infrastructure-specific threats such as unauthorized storage configuration changes are also discussed. The resulting risks to storage infrastructure (e.g., data breach, data exposure, unauthorized data alteration and addition, data corruption, data obfuscation and encryption, and tampering of storage-related log and audit data) are also analyzed based on the possibility of the realization of these threats and their impacts.
- Section 4 provides the core material for the publication. It details security recommendations for all facets of storage infrastructure management.

2 Data Storage Technologies: Background

Data storage technology encompasses the devices, objects (e.g., storage elements, storage arrays, storage network switches or storage media), and processes (e.g., protocols and interfaces) used to store computer data in non-volatile (durable) form. Hence, this technology can be viewed upon from the following two taxonomies:

- **Based on location of storage resource:** The storage device is directly attached to the storage client or host computer and called the direct-attached storage (DAS), or there is a network separating the host computer and the storage device (networked storage).
- **Based on storage type (access type):** This classification is based on the service interface offered by the storage system that is used by the client software. Examples include block-based storage (block storage service), file-based storage (file storage service), and object-based storage (object storage service).

In DAS, the storage device can be either an integral part of the computer (attached to the bus) or external storage (attached to a computer port, such as serial or USB).

Networked storage is broadly classified based on the type of access, such as network-attached storage (NAS), which provides file-level access across the network, and storage area network (SAN) whose protocols provide block-level access across the network. Further, in SAN, either the entire network stack can be comprised of storage-specific protocols (e.g., fibre channel), or it may consist of storage-specific protocols running over (or encapsulated within) common networking protocols (e.g., iSCSI by design running over TCP/IP, Fibre channel over IP [FCIP], Fibre channel over Ethernet [FCoE]).

Since SAN is a specialized, high-speed network for block-level network access to storage, a more detailed look at its variants is warranted. The variants are the results of different types of network stacks with different protocols in certain layers of the stack. The building block of SAN-based systems typically include (a) host computers (clients); (b) topology, most of which involve switches (called SAN Fabric); and (c) storage devices/arrays, with all three components interconnected using various network stacks.

Based on the above descriptions, SANs can be broadly classified as:

a. Fibre Channel SAN
b. IP SAN
c. Fibre channel over Ethernet (FCoE)
d. NVMe over Fabrics

A Fibre Channel SAN is a network stack with five layers (i.e., FC0 through FC4, unlike the seven-layer OSI stack). The logical storage resource addressed by Fibre Channel SAN is called the logical unit number (LUN).

iSCSI [1,2] and Fibre Channel over TCP/IP (FCIP) are examples of IP SANs since they use the IP protocol at the network layer, as in the former, or to encapsulate the Fibre Channel frame in an IP packet, as in the latter.

In FCoE, the Fibre Channel frame is encapsulated in Ethernet packets. The type of traffic (data and commands) carried in all three categories of SAN is SCSI.

NVMe is the standard host controller interface for systems using PCI Express (PCIe)-based solid state drives (SSDs). The NVMe over Fabrics (NVMe-oF) specification defines a protocol interface and related extensions that enable the NVMe commands to be transmitted over other interconnects, including RDMA, Fibre Channel, and TCP. NVMe-oF extends the NVMe deployment from a local host to a remote host for a scale-out NVMe storage system.

The hardware that connects the host computer (client) to storage devices is called the SAN fabric. The type of connection is called a topology. There are three kinds of topologies in SAN: point-to-point (two devices are directly connected), arbitrated loop, and switched fabric. In switched fabric, there is a set of hardware switches (acting as one big logical switch) separating the host computer and the storage resources. Since all variants of SAN run Fibre Channel protocol, the SAN fabric is synonymous with Fibre Channel fabric.

The taxonomy based on storage, access, or service type includes block, file, and object types or services. Since the choice of a storage service is dictated by the specific IT system use case (e.g., volume of data, control required over data, required performance, nature of data representation), this document will present an overview of storage technology in terms of these services (synonymous with storage/access types).

2.1 Block Storage Service

A block storage service offers an interface that reads and writes fixed-size blocks of data, typically offering high bandwidth, low latency access to storage devices at the block level [3]. Each storage device in a block-level storage system can be controlled as an individual hard drive, and the blocks are managed by the host OS. Block storage protocols like SCSI, SATA, iSCSI, Fibre Channel, and FCoE (Fibre Channel over Ethernet) are utilized to transport the storage blocks from the storage resources to the client system [4].

This type of service is offered by DAS. Across the network, this service is provided by SAN protocols, such as iSCSI and Fibre Channel [5]. All variations of SAN—such as FC, FCIP, FCoE, iSCSI, and FC-VNMe—provide block storage service. Here, the remote storage devices are presented as if they are locally attached to the host system on which storage client software is running.

2.2 File Storage Service

This type of service projects storage resources in the form of file system model with files contained in directories within volumes. The different variations of this service and their associated protocols are:

- Network-attached storage (NAS) with NFS protocol (current version is NFS 4.2 [6]) – A module that is part of the protocol implementation system, called the NFS client driver, mounts the volumes that are relevant for the client in its environment. The volume can be shared by multiple clients. Behind the scenes, files can be encrypted and replicated by making redundant copies. The files or folders can also be shared from either a dedicated appliance (typically referred to as a "NAS device" or "NAS array") or from any host running the NFS server service.

- NAS with SMB/CIFS protocol connection – This is provided by a LAN-attached file server, just like those that provide NFS protocol connection, but with the standard SMB protocol that is found in the network stack of operating systems used in personal computers and workstations.
- NAS with multi-protocol support – There are file service storage offerings that support multi-protocol exports of a folder or filesystem (e.g., both NFS and CIFS, concurrently). Each of these may have slightly different access control structures (i.e., ACL/permissions specifications), and some conflicts in access control rights may have to be resolved during access requests.
- NAS with parallel NFS protocol – This is provided through a clustered collection of storage servers (instead of a single NFS server) that slices and/or strips data and metadata at the back end while providing dynamic, distributed client connections at the front end across the set of clustered hosts. The parallel NFS system is implemented either by (a) partitioning filesystem namespace and assigning storage resources (i.e., files) that belong to different namespaces to different servers (called symmetric cluster) or (b) splitting functionality across servers (called asymmetric cluster) by having a primary fileserver provide the directory information for the location of secondary storage servers, the data contained in them, and the method to access them.

This service is used for large-scale content repositories (because of its scalability), media stores, and development environments [7].

2.3 Object Storage Service

An object storage service presents data as flexible-size data containers called objects, unlike the fixed-size blocks offered by block storage service. The interface it provides is called the object-based storage device (OSD) interface. Each object has both data (a linear sequence of bytes) and metadata (an extensible set of attributes describing the object) as well as a unique object identifier (OID). The OSD interface contains commands to create and delete objects, read and write bytes to and from individual objects, and set and get attributes (metadata) on objects. Hence, the object storage model is also called a key-value model, with the value as the object. The user-specified metadata can be arbitrary in number and potentially far higher than what is possible with standard file systems. The objects are organized in a flat, non-hierarchical namespace (unlike file systems with a hierarchical namespace) called a storage pool, bucket, or container.

An OSD interface is agnostic to the type of storage hardware and can span multiple storage devices, thus making it highly scalable. The interface is typically a network-accessible REST API, so the client applications have to be designed to make API calls.

The OSD interface is useful for accessing and modifying unstructured data, such as images, where modification involves replacing the whole object with a new version.

2.4 Content-addressable Storage (CAS) Service

This is a specialized form of object-based storage that is intended for storing the content digests of documents to enable users to retrieve those documents without having to know the location of the actual data or the number of copies. Hence, a CAS service exposes the digest generated by a cryptographic hash function (e.g., SHA-1 or SHA-256) pertaining to the document it refers to.

CAS is used for retrieving documents with short- and medium-term retention requirements.

2.5 Higher-level Data Access Service

There are data access services that provide data at a higher level of abstraction than that of basic storage types (files, blocks, or objects). These services can only be accessed through clients specifically built to access data at the same level of abstraction (e.g., SQL database clients). These services are available both in enterprise data centers and in the clouds. The following are some of these higher-level data access services:

- NoSQL database services
- SQL database services
- Messaging queue storage services

NoSQL database services enable the storage and retrieval of unstructured data, such as images, videos, documents, and large binary objects. Unstructured data has higher logical structures and representations than basic storage types in order to facilitate faster storage and retrievals. They include key-value store, multi-modal database, graph database, and others.

SQL database services enable the storage and retrieval of structured data that is typically in a tubular format (also called relational tables). The access is enabled through a standardized interactive programming language called Structured Query Language (SQL) [ISO/IEC 9075:2016 Database languages – SQL]. Current SQL databases can not only store data using relational tables/views but also other structures, such as XML, JSON, and BLOBs.

Messaging queue storage services are specialized services for the storage and retrieval of data from messaging queue infrastructures. These infrastructures are used by distributed applications whose components communicate asynchronously through subscription to a message queueing system. In addition to providing access to persistent data, this service also facilitates specialized operations, such as stream processing where events relating to multiple message storage and retrieval by distributed system components can be analyzed to discover patterns.

2.6 Software-defined Storage

Software-defined storage (SDS) is a storage architecture [8] that separates the storage hardware from the software that manages the storage infrastructure and automates its configuration. In other words, the storage capabilities and services are separated from the storage hardware. Advantages of this separation include the following:

- Use of heterogeneous storage hardware without the issues of interoperability

- Enabling of functions such as deduplication, replication, snapshots, and thin provisioning using industry-standard server hardware
- Automatic and efficient allocation of pooled storage resources to match the application needs of the enterprise

The following service capabilities are expected of the software managing the hardware storage resources in a software-define storage system [9]:

- Decouple storage policy management from the storage hardware;
- Support heterogeneous storage environments;
- Allow for the ability to add new storage capabilities across all platforms and not just to individual arrays; and
- Ensure that the storage software understands and leverages the capabilities of all storage hardware.

2.7 Storage Virtualization

Storage virtualization allows the capacity of multiple storage devices or arrays to be pooled (abstracted) so that they can be managed as one entity. Virtualization can aggregate and manage storage resources as logical storage across a wide range of physical storage devices (physical storage) in large networks (e.g., SAN) or data centers. This technique provides the flexibility to change the logical to physical relationship over time and mask the details of physical storage resources [10]. The following are some scenarios where storage virtualization is deployed:

- Portions of multiple physical disk drives can be presented as a single mirrored logical volume (using a logical volume manager in a host or storage array). Further, the composition of physical disk drives in the mirrored volume can be changed.
- Sensing changes to access patterns, drives on which data is stored can be changed (e.g., store frequently accessed information on high performance drives), thus providing an automatic tiering functionality.

In addition to logical volumes and masking, other techniques for storage virtualization include zoning, use of host-bus adapters, RAID, and the use of distributed file systems or objects [11]. The benefits of storage virtualization are scalability, performance, redundancy, and increased storage resource utilization.

2.8 Storage for Virtualized Servers

A virtualized server is one where a single physical server runs multiple computing stacks (each consisting of an OS and applications) called virtual machines (VMs) with the use of a software called the hypervisor. Storage infrastructure specifically designed to support virtualized servers is often called virtualization-aware storage or VM-aware storage. In most environments, this infrastructure is managed together with the VMs rather than as separately managed LUNs.

A key driver for building this VM-aware storage is to enable policy-based provisioning of storage resources at the VM-level through the hypervisor (which controls the allocation of all resources to VMs) so as to meet data access QoS requirements for the applications hosted on the

VMs. Another feature of this architecture is the ability to decouple VM storage from the individual hypervisor, thereby enabling an advanced virtual infrastructure function, such as live migration of VMs from one hypervisor to another, as well as enabling automated failover of a VM between hypervisors.

If the storage network infrastructure used to connect virtual machines to storage resources is implemented as a SAN, such a VM-aware storage system can be called a server-based SAN or virtual SAN (VSAN), although there is a specific commercial product offering by that name. A VSAN can be implemented on a cluster of virtualized servers by aggregating the direct-attached storage in the various nodes (virtualized servers) within the cluster and then treating that storage as a shared SAN resource. Each virtualized server that participates in the VSAN must have at least one SATA or SAS hard drive that is dedicated to SAN storage. Additionally, some implementations may have at least one solid-state drive (SSD) in the virtualized server used as a read/write cache or for tiering storage.

Since the management functions in a VM-aware storage infrastructure are enabled using software, they can be looked upon as a subset of SDS tailored for virtualized server environments. The key factor in a VM-aware storage environment is that the storage components are managed together with the VMs rather than as separately managed volumes or LUNs (logical unit numbers) [12]. One of the limitations of VSAN compared to general-purpose SAN is that it is only used for storing data related to VMs and is not multi-purpose. However, some commercial implementations also allow physical hosts to use VSAN volumes.

2.9 Converged and Hyper-Converged Storage

In a converged architecture, the storage, memory, networking, and virtualization software are preconfigured and pre-installed for fast deployment in a single box (e.g., a Rack containing one or more physical hosts, storage resources [DAS or storage arrays], and network components). A hyperconverged architecture takes the level of abstraction one step further where the individual storage components associated with the physical hosts are virtualized to build up a common storage pool, which is shared among all of the VMs or containers through the software-defined storage (SDS) management software [13]. Therefore, a VM or a container hosted on one physical host, say H(i), may use the storage associated with a different physical host, say H(j). As a consequence, the SDS over a shared storage hyperconverged platform introduces a storage network where the storage access is carried out over the network in the form of remote disk access through popular storage networking protocols, such as the iSCSI.

A hyperconverged storage or hyperconverged integrated system (HCI) is one where hardware required for compute, network, and storage are tightly coupled. All primary storage management functions—together with other functional capabilities such as backup, recovery, replication, deduplication, and compression—are delivered via the management software layer of the HCI vendor and/or hardware along with compute provisioning. Examples include Nutanix, Scale Computing, Cisco (HyperFlex), and SimpliVity [14]. The tight integration of the hardware comes about due to HCI vendors working with storage device manufacturers to create a storage solution that is tailored to their software stack as original equipment or as part of an industry-accepted reference architecture.

In this system, some of the CPU used for computing has to be shared for performing storage management functions. The overall management software stack may include a compute node, a hypervisor, and Virtual SAN (VSAN) software, depending on the deployment environment (e.g., virtualized infrastructure, virtual desktops, unstructured data stores, high performance computing) [15]. A common deployment scenario is one where the application environment consists of microservices-based applications implemented using VMs and/or containers. The expected features in an HCI solution include [16]:

- Support for multiple hypervisors
- Data protection features, such as always-on deduplication and compression across primary storage and backup
- Management control through a single pane of glass or a central dashboard
- Ability to provide QoS storage requirements based on application needs

Offering application processing capabilities in the storage controller of the storage device (e.g., NVMe SSD) using a system on a chip (SoC) is one approach for hyperconverged storage architecture. Another approach is to provide an add-in storage card (that can provide SSD or raw flash storage) with an embedded CPU for running applications with them connected directly to the hosting server's PIC bus and running NVMe protocol [17].

2.10 Storage Infrastructure in Cloud

Storage systems in the cloud may be either standards-based or proprietary and may include object-, block-, or file-based services. The technical reasons for enterprises using storage systems in the cloud are [18]:

- To accommodate new demand for storage resources without building an additional data center
- To respond to changes in demand for storage, such as peaks and valleys
- The need for immediate storage capacity
- Increasing management complexity of on-premise storage infrastructure

Storage systems based in the cloud provide several sophisticated data services [19]:

- Collaboration capability – Includes features such as (a) notifications when files are changed by others, (b) file sharing with the ability to set editing and view-only permissions, (c) simultaneous editing, and (d) change tracking and versioning
- Data integration and analytics capability – Ability to integrate data resident on several cloud sources, perform complex analytics, and either instantly serve the extracted information or store it in a persistent storage for later access by cloud service customers
- Advanced data protection services, including replication, mirroring, archiving, auditing, encryption

Cloud storage services often include:

- Block storage services that expose software-defined block devices that can be presented to virtual hosts running in the cloud

- Object storage services that can be mapped to hosts, applications, or even other cloud services
- Scalable shared filesystems that can allow a scalable set of hosts to access the same filesystem at a high speed
- A variety of replication, caching, archiving, mirroring, and point-in-time copy services to all of the above

Additional cloud services (e.g., managed database services, data lakes, memory caches, messages queues) are also offered, all of which can store stateful and transient data. However, experts are divided over whether to classify them as storage services in and of themselves.

2.11 Storage Management

Storage management refers to all activities geared toward ensuring reliability, resilience, performance, and the security of storage resources through the use of management tools and processes. Since storage security is the central focus of this document, this chapter will focus on all activities not related to security controls (and associated recommendations), which are deferred to Chapter 4. The non-security control-related activities that are followed as state of practice are:

- Data Classification or Categorization
- Data Sanitization
- Data Retention
- Data Protection
- Performance Enhancement – Data Reduction
- Security Controls

2.11.1 Data Classification or Categorization

Enterprise data can be classified along several dimensions, such as:

- Sensitive vs. non-sensitive
- Frequently accessed vs. non-frequently accessed
- Productions vs. development

The sensitive vs. non-sensitive classification is required to enable provisioning of appropriate security controls (e.g., authentication, authorization, encryption, key management, sanitization etc.) Further, the sensitive category may require sub-categories based on regulations applicable to the data, such as PII, HIPPA-related, and PCI. The frequently accessed vs. non-frequently accessed classification is required to provision the appropriate storage media (e.g., SSD vs HDD). The production vs. development classification may be required for both media selection and security controls.

2.11.2 Data Sanitization

Sanitization is the process of rendering previously written data in the storage media irretrievable, such that there is reasonable assurance that the data cannot be easily retrieved or reconstructed

[10]. There are three methods for sanitization: (a) clear (i.e., overwrite of the existing data), (b) purge (i.e., using a strong magnetic field for magnetic media degaussing, cryptographic erase for encrypted data), and (c) destroy (i.e., physical destruction of the media, such as burning, pulverizing, etc.). Factors that determine the type of sanitization include the category of information in the media, the nature of the media (magnetic or optic), and the reuse plans for the media.

2.11.3 Data Retention

While data sanitization is the last activity in the data governance framework, there may be situations where data needs to be retained for a short-, medium- (i.e., less than 10 years), or long-term duration. This may be due to operational, legal, regulatory, or statutory requirements.

2.11.4 Data Protection

Data protection is an umbrella term for all activities that ensure that data is accessible, usable, uncorrupted, and available for all authorized purposes with an acceptable level of performance and is handled in accordance with compliance requirements, including privacy and all physical, administrative, and technical means to provide assurance against accidental or intentional disclosure, modification, or destruction.

The range of objectives and associated activities provides a taxonomy for classifying data protection activities under three facets: storage, privacy, and information assurance/security [20]. The activities related to privacy are outside of the scope of this document since privacy-related laws and regulations differ by countries and communities of interest. The activities related to information assurance/security are predominantly technical controls, and each of them needs a dedication session for discussion of their details. Hence, in this section, only storage-related data protection activities/controls are discussed. These controls are:

- Data backup and recovery,
- Replication technologies,
- Continuous data protection, and
- Point-in-time copies and snapshots.

Backup is an operation wherein data stored in storage devices is accessed by production systems and periodically copied to another set of storage devices (some of which may be offline). Because of the changing nature of data content, a backup taken at an earlier time is often made obsolete by the backup taken at a later time. Backups can either be “file backups,” which back up a select portion of the data in a storage device (often based on logical data structures, such as files, directories, data under a database schema, etc.), or “image backups,” which contain the entire content of a particular device (e.g., an individual LUN).

Data replication is the process of writing the same data to two separate locations (i.e., two separate storage systems) [18]. Replication is often used as part of the data recovery process and involves copying data from one site to another. Generally, there are two types of replication: synchronous and asynchronous. Synchronous replication involves the real-time copying of data from site A (e.g., a production platform) to site B (e.g., a specially designated DR site). Asynchronous replication involves time delay and may be performed continuously or using a designed frequency for writing data from site A to site B. The time delay and frequency are dictated by the enterprise's disaster recovery policy and are described in terms of specific RTO (recovery time objective) and RPO (recovery point objective) goals. Specific point-in-time replicas are designated as archives for fast recovery and a wide variety of other uses, such as cloning production data for testing purposes.

Data protection involves activities and mechanisms that span the entire storage lifecycle. These phases include [21]:

- Data at rest/at the endpoint – on a server or client device
- Data in transit – between storage devices, client to server, server to server
- Data in use – during viewing, modifying, or synchronizing between devices
- Data traveling outside of the security perimeter – during downloads, etc.

Continuous data protection (CDP) is a form of backup that supports a fine-grained recovery and improved RPO. Unlike traditional backup, where copies of the data are performed periodically, changed blocks in CDP are continually transmitted to the target storage environment, which captures or journals the changes over time. In this respect, CDP resembles replication. However, unlike replication, CDP will typically allow "playback" of the copied data to previous points in time using a variety of techniques (e.g., byte-by-byte, pre-determined bookmarks, past versions, etc.)

A snapshot is a point-in-time copy of a defined collection of data. Snapshots are a way to create distinct "point-in-time" views of a data set, capturing the state of the data and ignoring those that are in a state of flux. Snapshots are often "thin," meaning that they store only the individual blocks changed from a given point in time in reference to the source data. This often means that if the primary data is unavailable, the snapshots will not be usable.

2.11.5 Enhancing Performance – Data Reduction

Data reduction is the process of reducing the amount of data stored and transmitted in an effort to reduce costs and improve efficiency. The two common approaches to data reduction are data deduplication and compression. Both of these approaches can be used together.

Data compression (sometimes performed in hardware) seeks to reduce the amount of data by encoding it with a known algorithm to produce a representation of data that uses fewer bits of storage than the unencoded representation [10]. Data compression is used in tape backups and during remote data replication in network gateways to reduce the bandwidth requirements for disaster recovery (DR) and business continuity (BC) operations. Interoperability is a key requirement for data compression since the encoding and subsequent reading system may be different.

Data deduplication attempts to replace multiple copies of data with references to a shared copy. It works by eliminating identical blocks of storage. For example, if a storage system has 500 identical blocks, the storage array will store just one copy, thereby eliminating the need to store the other 499 copies [18]. This can take place at the storage device level, the transmission stage, and the file system level.

2.11.6 Security Controls

Security controls—such as encryption, authentication, and authorization—are applicable to many of the storage management and protection activities described in the previous sections. As stated earlier, these security controls form part of the security recommendations of this document and are detailed in Chapter 4.

3 Threats, Risks, and Attack Surfaces

This section provides background information regarding storage system security threats, risks, and attack surfaces (where risks are the possible outcomes or goals of threats, and attack surfaces are the possible means through which risks can manifest).

3.1 Threats

A threat is the adverse potential of an insecure state. Threats will often have a 1:1 correlation to *types of attack* or *types of exploitable vulnerability*. The following sections provides a brief overview of storage infrastructure-related threats.

3.1.1 Credential Theft

Credentials are used to verify the identity of users, authenticate them, and grant access to storage systems and tools. Different forms of credentials exist, including physical keys, tokens and cards, passwords, digital private keys, session cookies, digital certificates on websites, and more. However, all of them are vulnerable to hackers using the right tools or techniques. The most widely used and easily compromised are login-password credentials, which generate a significant amount of risk to any organization. Credential theft is a growing industry within the cybercriminal ecosystem. The market for credential theft is extremely broad with very high potential as a result of the proliferation of cheap malware kits available online, a global increase in active stealer campaigns, and ever more sophisticated tactics, techniques, and procedures implemented by cybercriminals [22]. Password length and complexity alone are often insufficient protection against an attack. In fact, almost all effective methods of credential theft (other than password spray and brute force cracking) involve stealing the user's exact password rather than randomly guessing it. Modern ransomware often scrapes passwords from the data sets it has captured. Along with phishing and list cleaning via ransomware, keystroke logging—in which malware virtually watches a user type in their password—is another method of credential theft that works regardless of password complexity [23]. In many cases, login credentials are stored within storage infrastructure. If this data is not properly encrypted at rest, and the storage infrastructure is compromised, a hacker can gain access to a multitude of user credentials.

3.1.2 Cracking Encryption

Encryption is used to secure data at rest and in transit. In addition, encryptions are also used to secure the sessions in which data at rest or in transit is managed and controlled. Encryption algorithms make use of randomness to create keys or other key components. Encryptions can have a range of weaknesses, from weak encryption algorithms and weak key generators to server-side vulnerabilities, leaked keys, fundamental design flaws of bugs, and backdoors [24]. It is not only important to use strong encryption but to also properly secure the encryption keys. When it comes to key generation, the same key should not be created twice. Some attacks aim to disrupt the random number generator so that it issues the same random number for key generation twice in a row [25].

3.1.3 Infection of Malware and Ransomware

Malware is the general term for any program that is designed to damage, disrupt, or hack a device, whereas ransomware is a particular type of malware that blocks access to data until a ransom fee is paid to its creator. Malware compromises a system, slowing down its basic functions and breaching its security. It can be used to steal data, control a device/system, and harvest the system's resources for illegal activities. Malware can infect a system in several ways. Similar to viruses, it can be transmitted via file sharing, downloading free software, email attachments, using compromised portable storage devices, and visiting infected websites [26]. Malware can be mistakenly installed on a storage management host and consequently cause harm such as credential theft, privilege escalation, data corruption/loss/alteration, compromise future backups, and more. In general, malware will use OS vulnerabilities to install itself and perform various actions, meaning that more common operating systems will be more likely attacked. For this reason, it would be easier to attack the storage management system than the storage device itself.

Ransomware is a form of malware that encrypts the data in the storage systems, rendering it unusable. The attacker then demands a ransom to restore access to the data upon payment. In some cases, the attacker will publish confidential data that was collected from the storage system to create urgency.

3.1.4 Backdoors and Unpatched Vulnerabilities

Backdoors and unpatched vulnerabilities could be used either directly or indirectly to bypass other security controls.

Backdoors are software mechanisms or capabilities ***intentionally*** created by vendors or individual contributors (and, in rare occasions, by states or malicious actors) for reasons often considered legitimate by the author (e.g., to improve support, debugging, national security, etc.). Given their dangerous potential, backdoors are not officially documented and are meant to be known to a restricted set of individuals. However, over time, their existence could be intentionally or unintentionally leaked or discovered by the public.

Unpatched vulnerabilities are ***unintended*** software side effects or dependencies not caught by QA that present a security risk.

Once vulnerabilities are known—and especially if they are discovered in software versions that are still publicly supported—vendors typically issue a software fix in the form of a patch or a new version that closes the gap.

3.1.5 Privilege Escalation

Privilege escalation is the act of exploiting a bug, design flaw, or configuration oversight to gain elevated access to resources that are normally protected from an application or user [27]. It is highly linked to backdoors and vulnerabilities, and some might even consider it a sub-case. Privilege escalation occurs in two forms: 1) vertical privilege escalation (also known as privilege elevation), where a lower privilege user or application accesses functions or content reserved for higher privilege users or applications, and 2) horizontal privilege escalation, where a normal user accesses functions or content reserved for other normal users. In storage systems, this type of threat can result in a wide variety of risks, including data corruption, data alteration, data loss, and more. For example, an attacker can use elevated privileges to gain access to a storage system and delete storage volumes and modify access configuration. The attack can also compromise backup copies of the data (e.g., synchronous/asynchronous copies, snapshots) or the generation of future backups. The privilege escalation itself can occur on various levels, such as the storage components (e.g. storage array, host/client), the networking devices (e.g. the switch), or the management systems (e.g. storage management systems).

3.1.6 Human Error and Deliberate Misconfiguration

Even with the existence of security controls, users may end up with a technically allowed storage configuration that still presents an unacceptable exposure (e.g., mapping a restricted object storage pool to a public network, stopping replication or backup for maintenance without reenabling it afterwards). Such omission could be unintentional (i.e., an error) or deliberate (i.e., a sabotage).

Human errors take different forms, and some are significantly more difficult to identify or prevent than others:

- Typos
- Lack of knowledge or unfamiliarity with internal security baselines and vendor best practices
- Miscommunication between individuals or teams
- Errors related to the orchestration or automation of storage infrastructure
  - Direct, such as bugs in scripts and manifests
  - Indirect, such as unrealized software dependencies

3.2 Risks to Storage Infrastructure

Security risk is defined as:

> "…the extent to which an entity is threatened by a potential circumstance or event. Risk typically is a function of: (i) the adverse impacts that would arise if the circumstance or event occurs; and (ii) the likelihood of occurrence. Information system-related security risks arise from the loss of confidentiality, integrity, or availability of information or information systems. These risks reflect the potential adverse impacts to organizational operations (including mission, functions, image, or reputation), organizational assets, individuals, other organizations, and the Nation." [28]

3.2.1 Data Breach

An incident that involves sensitive, protected, or confidential information being copied, transmitted, viewed, stolen, or used by an individual unauthorized to do so. Exposed information may include credit card numbers and associated data, personal health information, customer data, company trade secrets, matters of national security, or any other proprietary or sensitive information.

Data breaches can originate from an external source, such as a hacker or cybercriminal, or from an internal source, such as a malicious insider or disgruntled employee. Data breaches can be performed in a covert manner with traces being concealed or entirely removed or in a manner that can be easily identified, whether this was deliberate or due to a lack of sophistication. The impact of data breaches can span a wide range, from inconvenience to users to the devastating exposure of sensitive/confidential data, resulting in irreparable damage to the reputation and operational health of the organization.

3.2.2 Data Exposure

An incident that involves the inadvertent exposure of otherwise confidential information. Sensitive data exposure occurs as a result of not adequately protecting a data asset (e.g., a file, database, network share, object store) where information is stored. This might be the result of a multitude of root causes, such as weak (or lack of) encryption, software flaws, loss of custody of removable media, incorrect or too relaxed access limitations, or user data upload to an incorrect location. Different types of data can be exposed in a sensitive data exposure. Banking account numbers, credit card numbers, healthcare data, session tokens, Social Security numbers, home addresses, phone numbers, dates of birth, and user account information like usernames and passwords are some of the types of information that can be left exposed [29]. Unlike *Data Breach*, which involves an active action by a malicious actor, data exposure can occur by mistake, such as when data is left exposed in a database, host, or other storage infrastructure for anyone or specific unauthorized users to see. Examples of data exposure include transmitting confidential information to the wrong recipient, mistakenly making sensitive data available for search on a public search engine, mistakenly configuring access control to allow read permission to sensitive information to unauthorized users/user groups, placing data in publicly available object stores, using weak cryptographic algorithms or keys, not implementing hashed and salted password practices (which is a form of cryptography similar to encryption), and other unsecure data storage practices [29].

3.2.3 Unauthorized Data Alteration and Addition

Unauthorized data alteration or addition is an incident that involves the illegal, unauthorized, or fraudulent alteration of data. It is the process of modifying data before or after it is entered into the system. In this case, the attacker gains access to the data storage infrastructure and modifies the data in a way that will force future transactions to use inaccurate information. Data alteration and addition can originate from both an external and internal source in a covert and/or easily identifiable manner. In certain cases, this type of risk is realized using the “salami attack” method, in which the attacker steals a little bit at a time over a long period of time from a large number of transactions (e.g., by rounding up small sums). The impacts of data alteration and addition can range from the loss of funds to permanent damage to reputation and trust.

3.2.4 Data Corruption

Data corruption refers to errors in data that occur during writing, reading, storage, transmission, or processing and which introduce unintended changes to the original data. In general, when data corruption occurs, a file containing that data will produce unexpected results when accessed by the system or the related application. Results could range from a minor loss of data to a system crash. For example, if a document file is corrupted, when a person tries to open that file with a document editor, they may get an error message, and the file might not be opened or might open with some or all of the data rendered unintelligible. Some types of malware may intentionally corrupt files as part of their payloads, usually by overwriting them with inoperative or garbage code, while a non-malicious virus may also unintentionally corrupt files when it accesses them.

3.2.5 Compromising Backups

The backup, or archiving of copies, of data assets is important to enable the recovery of said assets when they are damaged or lost. Satisfactory recovery is possible only if the backup copies are generated correctly with an appropriate retention and currency, stored in a secure manner, and accessible in a manner that allows timely restoration. Since these prerequisites are codependent, backup is sensitive to multiple failures. For example, incorrect configuration could involve a live database backup performed without applying techniques to ensure consistency or write-order fidelity. Insufficient currency or retention could mean that at least some portion of the data, past or new, will be unrecoverable. An attacker, therefore, has a high motivation to target not only a "primary" data asset but also its copies. When existing copies cannot be compromised, another viable attack strategy could be to interfere with the backup process itself, thereby gradually "poisoning" future copies. When enough time has passed, the attacker can return to the original goal of compromising the primary data assets, knowing that the only available copies for recovery are too old.

Another type of "poisoning" strategy is to specifically infect backup copies of compute or application assets, such as OS images, software packages, or even source code repositories. This way, when an individual component or even an entire environment is rebuilt in an attempt to battle an infection, at least some portions of the malware will be included in the restored environment, allowing the attacker to quickly inflict more damage.

3.2.6 Data Obfuscation and Encryption

The reversable obfuscation and/or encryption of data results in data becoming unavailable to the user or organization unless it is decrypted using an encryption key. This type of risk is commonly used in ransomware attacks. Ransomware is a form of malware that encrypts the victim's data while demanding a ransom to restore access to the data upon payment. Originally targeting data/files on users' computers or enterprise servers, ransomware has evolved to also include other storage components, such as NAS and backup appliances [30]. Data obfuscation and encryption typically originate from an external source but could also potentially originate from an internal one. Since it is reversable and commonly used as part of ransomware attacks, it is meant to be identified and is commonly accompanied by a threat and ransom instructions. The impact of data obfuscation and encryption can range from the loss of funds to permanent damage to reputation and trust.

3.2.7 Data Availability and Denial of Service

In a data availability or denial-of-service incident, the data client cannot gain access to some or all of their data. A data availability disruption risk can occur due to purposeful or unintended damage to the communication path or access configuration. The damage can be physical, such as a disconnection along the communication path, or logical, such as the misconfiguration of an endpoint of network components. For example, an attacker can delete the SAN masking settings of a block storage device or suspend the export setting in NFS so that clients will be unable to access their data. Although the damage may be reversible (e.g., by restoring the settings that were deleted), it may cause temporary disruptions and downtime for the system or service. A denial-of-service (DoS) attack will also achieve disruption to data availability by flooding the targeted machine or resource with superfluous requests in an attempt to overload systems and prevent some or all legitimate requests from being fulfilled. DoS attacks could potentially impact not only individual data assets and clients but also an entire fabric.

3.2.8 Tampering of Storage-Related Log and Audit Data

The tampering of storage-related log and audit data is where an attacker deletes or modifies log data to prevent an effective audit trail in an effort to conceal the attack (in real-time or afterwards) or to mislead the people investigating attacks with false information. The logs can be partially modified, such as by modifying the timestamp. The impact of this risk is that the attacker/attack can remain unnoticed by security systems that rely on log data. During this period, the attacker can perform additional lateral movements that may jeopardize the data and/or service. For example, a brute force attack to log into a sensitive system may be concealed by deleting the login attempts from the logs. Another form of this risk involves tampering with the logging mechanism itself (e.g., disabling it, filling up all free space with synthetic messages, convincing clients to send log data to rouge log servers, etc.).

3.2.9 Mapping of Threats to Risks

The following table provides a mapping of threats discussed in Section 3.1 to the risk outcomes discussed in Section 3.2.

Threats – Insecure States and Adversary Capabilities	***Occurrence* – Risk Outcomes**
Privilege escalation	*Application system* – Data breach, data exposure, unauthorized data alteration, data corruption *Administrative system* – Compromise of existing and future backups, ransomware attack, DoS attack, tamper storage-related log and audit data, unsafe storage configuration parameters
Credential theft	Depending on whether user credentials or administrator credentials are compromised, all risk outcomes for privilege escalation apply to this threat.
Cracking encryption	Data breach and exposure of (a) data at rest, (b) data in transit, and (c) user/administrator session data
Infection of malware and ransomware	Malware can enable other threats – Privilege escalation, credential theft Malware, depending on where it is present – application systems or administrative systems can impact all risk outcomes in Section 3.2
Backdoors and unpatched vulnerabilities	Depends on the nature of the vulnerability, but in many cases, all risk outcomes for privilege escalation apply to this threat
Human error and deliberate misconfiguration	Depending on its type and scope, misconfiguration can impact all risk outcomes in Section 3.2

3.3 Attack Surfaces

Attack surfaces are defined as “the sum of the different points (the “attack vectors”) where an unauthorized user (the “attacker”) can try to enter data into or extract data from an environment” [31]. This section will list common digital and physical attack surfaces that are related to storage infrastructure.

3.3.1 Physical Access

Physical access to storage infrastructure involves physical intrusion into the data center, its perimeter, communication infrastructure (including cabling), or physical objects in transit (e.g., physical hosts, storage arrays, hard drives, tapes). Such intrusion is performed in order to access, steal, or damage the data or prevent its availability. The physical intrusion can occur by “overt access” in which the attacker will masquerade as someone who belongs in the situation (e.g., by playing the part of a cleaning employee, technician, or building maintenance personnel).

“Tailgating” is another way to access restricted areas in the data center. For example, an intruder gains entry to a network operations center by carrying a tray of food. Although the data center is protected by biometrics, the staff may open the door for the intruder and the food. Other intruders may simply follow employees in. Physical access protection is essentially the last line of defense. An intruder who gains access to storage infrastructure can ultimately steal, duplicate, harm, or destroy media and data. An intruder can also connect to the storage system port, insert a removable media to a physical port that may be used for firmware updates or to a management terminal, access data, and damage it. However, even if the storage system is well-protected by physical restrictions, an attacker can physically target additional components of the storage infrastructure that may be less protected, such as switches and management terminals. The attacker can target the data ports to intercept the data itself or the management ports. Communication cables are also a vulnerability. A sophisticated attacker can potentially tap into the storage communication by physically accessing the cables. Another physical access method involves replacing peripheral components, such as the keyboard and mouse, with infected components (e.g., infiltrating an infected keyboard that includes a “keylogger” component that transmits sensitive data, such as usernames/passwords, or infects the system with malware). An additional form of physical access is accessing removable media that is transported between storage sites. In some cases, companies back up large amounts of data on removable media and then transport this removable media to a remote DR site. The removable media can potentially be compromised or intercepted in transit.

3.3.2 Access to Storage OS

Access to a storage OS attack surface involves intrusion into a storage device by exploiting operating system vulnerabilities. The term “storage OS” refers to all of the operating systems that are related to the storage infrastructure. This includes storage arrays, switches, data protection appliances, and storage virtualization appliances. In many cases, the operating systems running these devices are based on a version of the Linux/Unix operating system, which is generally more closed or secure than general-purpose OS distributions. However, all operating systems include security vulnerabilities and are therefore regularly updated with security updates and patches. In addition, any operating system has configurations that may influence its security. An attacker can gain access to the storage OS by a variety of methods, from a local login process (using a standard protocol, such as SSH, ‘rshell,’ ‘telnet,’ etc.) through remote login using TCP-IP or through an OS vulnerability.

3.3.3 Access to Management Hosts

Most storage components are managed or configured through computing devices called management hosts, which run an operating system that is usually some variation of a commercial OS. By infiltrating the management host with malware or through an OS vulnerability, an attacker can, for example, hack an executable, read cached data, install a memory tap that reads data from the memory, install malware, and gain access to the related storage array and/or its configuration. Consequently, through the management host, an attacker can realize most related risks, including data corruption, data loss, data alteration, compromising of future backups, tampering log and audit data, and more. Access to management hosts provides the attacker with the ability to cause almost unlimited damage to the entire domain that is being managed by the management host.

3.3.4 Management APIs, Management Software, In-band Management

Storage infrastructure components expose management software UI, APIs, and other in-band or out-of-band management protocols for administering the devices and managing data storage. In some cases, the device has a management interface (e.g. SOAP or REST API) and, in parallel, management software that is installed on a management host. All of these interfaces create a variety of attack surfaces. For example, an attacker can access a storage device by impersonating the management host or software through the management API. In this case, the attacker does not need to infiltrate the management software in order to gain access to the management capabilities. Some equipment allows in-band access via the data links (e.g., Fiber Channel paths). The storage device allows in-band management access through the same connection plane used for providing the storage service. By doing so, it opens yet another attack surface, which can be exploited by attackers that can impersonate a storage client while sending management commands.

3.3.5 Storage Clients

Storage clients are compute components, or applications installed on compute components, that use the storage protocol to read/write data from a storage object or network. If a storage client is compromised, the attacker can potentially read the data that is consumed by the storage client, write data to the storage device or object, and encrypt data. In addition, if in-band access to the storage is enabled, the attacker impersonating the storage client can send management commands. Archiving systems may sometimes use a storage client to gain access to data in order to create backups. If the storage client is compromised, the attacker can also harm future backups. In this scenario, the attacker can then wait for a while, harming the ability of the organization to defend itself because it will not be able to use its compromised backups.

3.3.6 Storage Network (Tap Into, Alter to Gain Access)

When storage clients consume data from the storage systems, the data is transferred through a variety of storage network components (i.e., data in transit), such as storage switches, cables, and extenders. If such components are compromised, the attacker can tap into the data path and copy, view, reroute, or steal data. In addition, the attacker can read configuration data, management traffic, or other metadata (e.g., if the data in transit may include user credentials, encryption keys, and more). By compromising a network component, an attacker can also potentially perform data corruption, alteration, or addition by modifying the payload. Another form of attack is “man in the middle” (MITM), specifically Fibre Channel man-in-the-middle attacks. The purpose of the MITM is to sniff data, alter it, or bypass encryption and authentication mechanisms. A switch will only transmit information to the correct port, not allowing any other ports to see any communication that is not theirs. An entity using IP, such as a switch or an operating system, will send out ARP (Address Resolution Protocol) requests when it is trying to communicate with other entities. The issue with ARP is that any malicious entity could send out an ARP reply instead of the actual server. Since there is no authentication with ARP, similar to how there is no authentication with PLOGI in Fibre Channel fabrics, an entity receiving an ARP reply from an attacker would update their routing table with the incorrect information. Furthermore, even if a node did not send out an ARP request, which would request the MAC address of a specific IP address, it could potentially receive an ARP reply and update its own routing table. For example, an attacker could send out ARP replies to the entire network segment, telling each entity that the MAC address of the router, which is 172.16.1.1, is actually the MAC address of the malicious entity. When one node tries to communicate to any other node by going through the default router, it will actually be going to the malicious entity first since it is using the MAC address of the malicious entity for layer 2 routing [32].

3.3.7 Compute Environment of Key Individuals – Storage Admins

Storage administrators sometimes have remote access to the storage infrastructure. For example, the storage admin may have a computer that is remotely connected to the storage’s management host. The compute environment of such key individuals can be exploited to gain access to and compromise the storage infrastructure. For example, an attacker can install malware on this compute environment that will, in turn, install a key logger that allows for the interception of login credentials. This compute environment is therefore a potential attack surface.

3.3.8 Electricity Network

Since storage infrastructure is connected to the electricity grid, the electricity network may potentially become an attack surface. A huge spike in electrical current, such as the kind caused by lightning, can potentially damage and even erase data that is stored in electromagnetic discs. Voltage fluctuations that correspond to keystrokes create noise in the ground line. The ground line noise can be intercepted by a hacker connected to a nearby power socket. Another method is through a malware dubbed *PowerHammer*, which can stealthily exfiltrate data from air-gapped computers using power lines. This malware exfiltrates data from a compromised machine by regulating its power consumption, which can be controlled through the workload of the device’s CPU. Sensitive pieces of information, such as passwords and encryption keys, can be stolen one bit at a time by modulating changes in the current flow. In the line level variant of this attack, the attacker intercepts the bits of data exfiltrated by the malware by tapping the compromised computer’s power cable. In the phase level attack, the attacker collects the data from the main electrical service panel. The data can be harvested using a non-invasive tap that measures the emissions on power cables and then converted to a binary form via demodulation and decoding [33].

4 Security Guidelines for Storage Deployments

4.1 Physical Storage Security

Physical security is fundamental to the overall safeguarding of any IT infrastructure. Most software-based security controls can be compromised if an attacker gains access to the physical facility and equipment.

In many regards, physical security requirements for storage infrastructure are identical to those of other infrastructure elements like computers and network equipment (e.g., facility security, surveillance, transportation, etc.). These are well covered by multiple publications, including NIST SP 800-171 [34]. Additional valuable discussion regarding media disposal and destruction is available in ISO 27040 ([10]).

It is, therefore, beyond the scope of this document to cover all physical storage security aspects. Rather, focused guidance is provided for physical security aspects that are unique to storage infrastructure or are less emphasized in other publications.

PS-SS-R1 – Media security measures:

(a) Follow general recommendations, NIST SP 800-171, Section 3.8 (including protection, access restriction, sanitization, marking, transportation, cryptography, removable media, confidentiality, disposal).

(b) Lifecycle management should include purchasing media from a trusted source.

(c) Physical media and backups of sensitive data, which are used for data protection, should be stored in sufficiently distant locations, away from the primary storage.

(d) For sensitive information, a comprehensive inventory of storage media (cataloging) must be kept to track its location, ownership, capacity, and other relevant configuration attributes. Particular attention should be paid to tracking the actual content of media, including:
- Sensitivity level
- Classification (what type of data it stores, what applications and business services it relates to)
- Encryption level
- Potential impact if compromised or stolen (e.g., compromise of financial or medical information; leak of passwords, certificates, or encryption keys).
- Mitigation/contingency steps or procedures to employ (e.g., including changing passwords, re-issue keys, re-encrypt data, notify relevant stake holders)
- Dependencies between the data and other application

(e) Consider using advanced tracking controls on sensitive removable media, such as RFID tags, GPS tracking devices, tamper protection, and for extremely sensitive information triggered self-destruction mechanism (self-activated and/or remotely controlled).

PS-SS-R2 – Protect all sensitive administrative equipment: Sensitive workstations, which can be used to obtain administrative access to storage infrastructure, must be managed using corporate-approved security controls for access, surveillance, and auditing, including physical security. If possible, the same security measures required to protect the data itself should apply to the management workstations. This includes workstations located outside of the facility storing the data as well as work-from-home environments, when used.

PS-SS-R3 – Ensure that the data sanitization approach is sufficiently broad: Certain elements capable of storing sensitive information are sometimes overlooked when disposing of storage equipment, including non-volatile memory and cache objects (often found in storage arrays), firmware/BIOS settings, and HBA-level settings (which can contain WWNs, masking, IP addresses, passwords). Ensure that all of those elements are considered.

4.2 Data Protection

Section 2.11.4 discusses the objectives and associated activities of data protection, the three facets based on the range of objectives, and primary controls from the point of the view of the storage facet. To reiterate, these controls are:

- Data backup and recovery,
- Replication technologies,
- Continuous data protection, and,
- Point-in-time copies and snapshots.

The security recommendations in this section provide the due diligence aspects associated with implementing each of the controls above. Each recommendation has a unique identifier with format DP-SS-Rx, where DP stands for data protection, SS stands for secure storage, and Rx for the recommendation sequence.

4.2.1 Data Backup and Recovery

DP-SS-R1: The backup plan or policy should be established prior to deployment and should include, at the minimum, the following:

(a) Type of backup (e.g., full or incremental/differential, the use of CDP, versioning, replication, and point-in-time copies as part of the backup scheme)
(b) Frequency of backup (in particular, to meet RPOs)
(c) Retention period
(d) Types of media to be used
(e) Encryption requirements (in particular, backup encryption methods used should be at least as secure as those applicable to protected data)
(f) Other protection requirements, such as digital signing, location, facility security (including fire, explosion, and magnetic interference protection), immutability and locking, and a minimum number of copies per backup set and their geographic distribution
(g) Reference to applicable regulatory frameworks with appropriate controls
(h) Restore procedures

DP-SS-R2: The backup plan or policy should be comprehensive enough to:

(a) Cover all data assets of the enterprise irrespective of where they reside (i.e., on-premise or in the cloud)
(b) Be organized by the type of data involved (e.g., Tier 1, Tier 2, etc.)
(c) Consider data integrity at the application and business process levels (e.g., if two components must be recovered to the same point in time to function properly, then federated consistency mechanisms or equivalent should be planned and implemented)
(d) Consider the required restoration speed to meet business or regulatory requirements (e.g., for mission critical data with requirements for fast recovery or RTO, consider the use of point-in-time copies, such as snapshots of clones, as opposed to tape or over-WAN recovery)

DP-SS-R3: In addition to a backup plan or policy, standard operating procedures relating to the backup should:

(a) Monitor the execution of backups based on policy and associated notification mechanisms.
(b) Periodically test backups (at least monthly for critical data) to verify their integrity and their ability to be restored.
(c) Ensure copy hygiene. An up-to-date recovery catalog should be kept for each copy that records which anti-malware tools it has been scanned with and what the results of the scans were. For sensitive data, it is further recommended to periodically scan at least a subset of past copies with current anti-malware tools to identify “poisoned” copies.
(d) Periodically review (at least annually) the backup plan and operations procedures.
(e) Maintain an audit trail that provides the information necessary to ensure conformance of the backup operations consistent with the policy.
(f) Employ special controls when necessary (e.g., refreshing old, at risk, retired media by copying to new one).

DP-SS-R4: The data protection configuration management (including backup, point-in-time copies, and replication) should be centrally managed and separated from the data consumption plane. In particular, servers and clients should not be allowed to change their own data protection configuration.

4.2.2 Replication

DP-SS-R5: In both synchronous and asynchronous replication, the same level of data protection (e.g., encryption of data at rest, access restrictions) that is used in the primary storage should be carried over to the secondary storage.

DP-SS-R6: The confidentiality of data in transit during replication should be protected using encryption.

DP-SS-R7: When synchronous replication is critical, the primary storage server should have a feature to disallow any writes (or accept any write transaction request) on the data it stores if its synchronization with the secondary storage server is lost, and it should only resume processing when synchronization is restored.

DP-SS-R8: Obsolete replicas should be removed to reduce the attack surface.

4.2.3 Point-in-Time Copies and Snapshots

DP-SS-R9: When point-in-time copies, such as snapshots, are used as part of the backup scheme, they should be configured accordingly:

(a) To meet the recovery point objective (RPO) requirements of the target data sets in the snapshot. For example, if the business or compliance standards require that no more than five minutes of committed data could be lost in recovery, then the snapshot interval must be five minutes or less.

(b) To meet retention requirements. For example, if hourly copies are required for at least 48 hours, ensure that a sufficient number of hourly snapshots is preserved.

DP-SS-R10: Obsolete snapshots and clones should be removed to reduce the attack surface.

4.2.4 Continuous Data Protection

DP-SS-R11 – Security considerations for using CDP: Other than the functional benefits (e.g., improved RPO, finer-grained retention), the use of CDP or similar techniques (e.g., versioning of source data or replicas in AWS S3, Azure Blob) can also assist in improving forensics, when applicable. Replaying to previous versions of data can help one learn more about the attack profile and, in particular, better estimate attack times.

4.3 Authentication and Data Access Control

Storage infrastructure systems are administered by designated users who use various accounts to access these systems. The administrative users and their management hosts constitute an important attack surface that can be exploited by attackers. Since the individuals managing storage systems and infrastructure are generally privileged users, the allocation and use of privileged access rights should be restricted and controlled. Inappropriate use of system administration privileges can be a major contributory factor to the failures or breaches of storage systems.

A least privilege model that leverages specific roles should be implemented. According to ISO Standard ISO/IEC 27040 [10], the following roles should be implemented and used within storage technologies:

- **Security Administrator** – This role has view and modify rights to establish and manage accounts, create and associate roles/permissions for audit logging configurations and contents (audit log event entries can never be changed), establish trust relationships with IT infrastructure (e.g., shared secrets for RADIUS), manage certificate and key stores, manage encryption and key management, and set access controls.
- **Storage Administrator** – This role has view and modify rights for all aspects of the storage system. No access is granted to security-related elements or data.
- **Security Auditor** – This role has view rights that allow for entitlement reviews, verification of security parameters and configurations, and inspections of audit logs. No access is granted to the storage, configuration, or data.

- **Storage Auditor** – This operator-like role has view rights that allow for the verification of storage parameters and configurations and inspections of health/fault logs. No access is granted to security-related elements or data.

4.3.1 Authentication Recommendations

AC-SS-R1 – Unique Identifier for all users: All administrators should have a unique identifier for their personal use only. This requirement is important for accountability and audit purposes as well as for the ability to control access on the individual user level.

AC-SS-R2 – A centralized authentication solution: A centralized authentication solution (e.g., such as Active Directory, Remote Authentication Dial-In User Service [RADIUS], single sign-on [SSO]) should be deployed to enable the close monitoring and control of user access and to ensure uniform enforcement of the organization's authentication policies. The use of the authentication and permissions management module that comes with the storage product should be avoided and preferably disabled.

AC-SS-R3 – Configuration of authentication servers:

(a) The designation of servers to perform authentication services should be strictly controlled, and their validity should be periodically checked to detect and prevent the introduction of any rogue or unauthorized authentication servers.
(b) There should be multiple authentication servers to ensure availability and avoid single points of failure.

AC-SS-R4 – Secure connection to centralized authentication server: All communication between the centralized authentication server and the authenticating clients should be secured through protocols such as TLS.

AC-SS-R5 – Use of multi-factor authentication: Access configuration to storage infrastructure components that store mission-critical data should be protected using a minimum of two-factor authentication.

4.3.2 Password Recommendations

AC-SS-R6 – Secure password policies should cover service accounts: The secure password policies should be applied not only to individual accounts but also to service accounts (e.g., SNMP, NDMP) and accounts used by automation tools. Some vendor publications suggest at least 16 randomized characters for NDMP and 20 for SNMP.

AC-SS-R7 – Password Length: A good passphrase should have at least 15, preferably 20, characters.

AC-SS-R8 – Password Complexity: A good passphrase should combine uppercase and lowercase letters, digits, and special characters. They should not be similar to usernames and should not include repeated character sequences.

AC-SS-R9 – Password expiration: Expiry times should be set for all passwords. Passwords for administrative accounts should be set shorter than user words.

AC-SS-R10 – Password reuse: Users should be prohibited from reusing at least the four previous passwords (or more) based on organizational risk factors.

AC-SS-R11 – Password caching:

(a) Passwords should not be cached on the server, desktop, or any other system.
(b) Sufficiently short TTL or an equivalent control mechanism should be employed to guarantee that changes are propagated quickly throughout the network.

AC-SS-R12 – Saving passwords: Passwords should not be saved anywhere in cleartext (e.g., not in files) or in scripts. Furthermore, enabling storage management applications to locally remember users and passwords for automatic login should never be used, even if passwords are stored encrypted, unless managed through an authorized central authentication service, such as LDAP SSO.

AC-SS-R13 – Eliminate or change default passwords: The default passwords that come with system installation or deployment must be immediately changed.

4.3.3 Account Management Recommendations

AC-SS-R14 – Use of accounts unassociated with system users: Accounts not associated with any system user (e.g., not in an active directory, such as "guest," "anonymous," "nobody") should be disabled. In situations where they need to be used, they should not be mapped to any system user, and all of their default configurations (e.g., password, privileges) should be changed to conform to organization-wide policies.

AC-SS-R15 – Account lockout: Users should be temporarily or permanently locked out after a certain number (preferably three) of unsuccessful login attempts.

AC-SS-R16 – A local user account for emergency purposes: A single local user account should be maintained for access to storage resources in order to provide emergency-only access if the centralized authentication system is down. This account should conform to all organizational policies (e.g., password length, complexity).

AC-SS-R17 – Eliminate or disable default user accounts: The default user accounts that come with the storage system installations should be eliminated or disabled immediately. When there is a justified reason to keep any of those accounts, review and change its privileges to the minimum required.

AC-SS-R18 – Limit local and default user accounts: As much as possible, eliminate the use of local and default accounts. In situations where this is not possible:

(a) Limit the use of such accounts and the privileges they have.
(b) Password policies should apply to all user, local, and default accounts, including those with administrative rights.
(c) An exception for password expiration may be considered for the emergency account if its use is only allowed during a real emergency and providing that the password changes after each incident.

4.3.4 Privilege and Session Management Recommendations

AC-SS-R19 – Roles and responsibilities configuration: At a minimum, the four roles in the ISO Standard ISO/IEC 27040 [10] must be implemented for all access to storage resources (i.e., Security Administrator, Storage Administrator, Security Auditor, Storage Auditor) See Section 4.2 for more details.

AC-SS-R20 – Separate roles of administrative account types: A critical aspect of storage security is to separate administrative control planes. (For example, if attackers gain control over a host or compromise a host admin role, they could not trivially compromise its data assets, backups, and replicas.) At a minimum, this includes:

(a) The privileges required for *data management* (e.g., create and map a volume or share) and *data protection* (e.g., configure, stop, and delete backup) should be assigned to different roles.
(b) The privileges required for *data management* and *host administration* should be assigned to different roles.

AC-SS-R21 – The privileges assigned to any role should adhere to the principle of "least privilege": The permissions assigned to a role should be no more than what is required to perform the functions designated for that role.

AC-SS-R22 – Enable session expiration/timeout: All inactive open sessions between the client and a storage infrastructure system should be terminated through an automatic logout. This recommendation is applicable to all accounts that access the storage infrastructure (e.g., CLIs, APIs, etc.) but is especially important and mandatory for admin accounts.

AC-SS-R23 – Implement a "message of the day" and "login banner" notice: The "message of the day" or "login banner" notice should appear on every login to any storage infrastructure component or system via UI, CLI, or API (if applicable). The message should include a legal notice and a warning that the user is accessing a restricted system with sensitive data, as well as any additional warnings and meaningful messages, according to the organization's security and privacy policies.

AC-SS-R24 – Eliminate unnecessary replication trust between storage devices: When arrays do not have shared replicated volumes, disable the replication trust relationship between them.

4.3.5 SAN-Specific Recommendations

The topic of SAN-related access control involves multiple aspects. Some overlap with *Network Configuration* and *Administrative Access*, which are covered in other sections.

To eliminate repetition:

- Access control recommendations closely related to the *network infrastructure* (e.g., switch, port, HBAs, and NICs configuration; additional zoning guidelines) and *protocols* are discussed in Section 4.6.
- Administrative access is discussed in Section 4.10.
- Data-related access control is discussed in this section.

For a complete appreciation of all access control aspects, please refer to all three sections.

AC-SS-R25 – LUN access control: The set of hosts that can access a set of SAN storage devices must be restricted through zoning (software or hardware) and masking to the minimum required access.

AC-SS-R26 – LUN copy and replica access control: The set of hosts that can access a set of SAN-replicated LUNs, Snapshots, and other types of point in time copies of LUNs should be similarly restricted through zoning and masking to the minimum required access. Note that in many cases, a host granted access to a device should not be allowed to access a copy.

AC-SS-R27: The default zone should always be configured as "deny all."

AC-SS-R28: The zoning should be implemented in a switched SAN fabric based on sound logic, particularly as it relates to the separation of environments and traffic type, such as the following which should be separated <u>to the maximum possible extent</u>:

(a) Separation based on environment: *development* vs. *test* vs. *production*, etc.
(b) Type of traffic: data access vs. management vs. replication vs. backup
(c) Type of hosts: *virtualized* vs. *physical*
(d) Storage device type: *tape* vs. *disk*

AC-SS-R29: When software zoning is implemented, care should be taken to ensure that a host can only connect to storage devices provided by the simple name server (SNS) (by looking it up at the software zoning table) and not directly using device discovery.

AC-SS-R30 – Ensure allowlisting of devices that can join fabric: The policy specification feature in SAN that enables the creation of an allowlist of switches, arrays, and hosts that can join the fabric must be leveraged where applicable and carefully configured.

4.3.6 File and Object Access Recommendations

AC-SS-R31 – Restricting access to storage objects to the minimum possible: Follow the principle of least privilege, including:

(a) Access to storage objects for all protocols (e.g., CIFS, SMB, NFS, and object storage such as Amazon S3 and Azure Blob) should be restricted based on client IPs and/or relevant subnets, and the ports/protocols should be required.
(b) If supported, finer-grained access control mechanisms (e.g., by role, ID, labels, accounts, VPC, VPC endpoints, etc.) must also be used.
(c) Prefer granting access to centrally managed users and roles only, such as Active Directory and IAM, and not to local users of the specific system.
(d) Make sure that the default access to any share is set to "deny all" or equivalent.
(e) Disable or remove default shares unless they have a specific purpose, in which case review and adjust access rights to the minimum required.
(f) Provide the minimal access rights (e.g., read, write, execute, modify, delete, view ACLs, change ACLs, etc.), which could typically be adjusted individually.
(g) Prefer using ACLs in addition to the native OS user, group, or admin permission models.

AC-SS-R32 – Avoid permitting anonymous, null, guest, or "public access" users: Such users are typically able to perform network discovery without the need to authenticate. If it is absolutely essential to have this type of user, make sure it is mapped to the "nobody" user group and not to ID 0.

AC-SS-R33 – Regularly audit file and object security settings: Perform regular audits of all the security settings mentioned above.

AC-SS-R34 – Scan files with anti-malware tools on-access: Every time a file is accessed, it should first be scanned with anti-malware tools to ensure that it has not been compromised.

AC-SS-R35 – Granular permission assignment: For file and object sharing systems (e.g., NFS, CIFS, S3 buckets, etc.), prefer granting permissions at a finer level of granularity rather than a coarser one (e.g., file or object over folder or label, over share or bucket).

AC-SS-R36 – In NFS, avoid sharing with execute-as-root allowed: While executing mounting, the "nosuid" option should be used to prevent programs from being executed as a root user on the client. In general, NFS clients should not be allowed to run "suid" and "sgid" programs on exported file systems.

AC-SS-R37: In NFS, for files that are to be used in the "read only" mode, the mount configuration for corresponding NFS shares should always have the "noexec" option.

AC-SS-R38 – Export of administrative file systems should not be allowed: This includes the '/' filesystems, restricted OS/storage array system folders, etc.

AC-SS-R39 – When CIFS is used, "Full Control" permissions should not be granted to any user since the recipient can use it to modify the permissions, thus resulting in the leakage of privileges.

AC-SS-R40 – Enable multi-factor authentication delete/lock, when supported, to sensitive objects in object storage.

4.4 Audit Logging

Storage infrastructure components generate event log entries for a wide range of transactions or events. These event log entries have to be recorded in some manner for event logging. From a security or compliance perspective, it is important to capture those event log entries that are necessary to demonstrate proof of operations (e.g., encryption and retention), enforcement of accountability and traceability, meeting evidentiary requirements, and adequate monitoring of systems. This subset of general event logging is commonly called audit logging.

The following audit logging events are relevant for security purposes:

- **Management events** (i.e., what a human did) are always of interest.
- **Blocked attempts to grant** access (to storage, login sessions, etc.) are most often of interest.
- **Data access events** are usually of limited interest, except in situations where critical files and directories need to be tightly monitored.
- **Control events and data access events** are typically of the least interest (they can provide useful information during root cause analysis after an incident and are important in extremely sensitive environments).

Deficiencies in security logging and analysis allow attackers to hide their location, malicious software, and activities on victim machines. Even if the victims know that their systems have been compromised, without protected and complete logging records, they are blind to the details of the attack and to subsequent actions taken by the attackers. Without solid audit logs, an attack may go unnoticed indefinitely, and the particular damages done may be irreversible. Sometimes, logging records are the only evidence of a successful attack. Many organizations keep audit records for compliance purposes, but attackers rely on the fact that such organizations rarely look at the audit logs, and they do not know that their systems have been compromised. Because of poor or nonexistent log analysis processes, attackers sometimes control victim machines for months or years without anyone in the target organization knowing, even though evidence of the attack can be obtained in unexamined log files.

Based on the criticality of event log data for attack detection and forensic investigation, the following are the security recommendations for implementing audit logging capabilities. Each recommendation has a unique identifier with format AL-SS-Rx, where AL stands for audit logging, SS stands for secure storage, and Rx stands for the recommendation sequence.

AL-SS-R1 – Enable logging of all storage infrastructure components: Storage systems and devices should participate in audit logging, and all significant storage management events should be collected.

AL-SS-R2 – Ensure that all of the device's time is synchronized with a reliable, external time source, such as an NTP service: An NTP service is critical for time synchronization. If the NTP service is disabled, dependent systems may suffer from inaccurate timestamps on messages, events and alerts, inconsistent time across different devices, and subsequent failure to perform log analysis, correlation, anomaly detection or forensics. Establishing and using a common, accurate time source across the environment helps ensure that event records from different sources can be correlated.

(a) Ensure that an NTP service is enabled on all devices.
(b) In particular, log servers, monitor time synchronization validity, and handle alerts at a high priority.
(c) Ensure that devices are configured to synchronize time with a time source server, such as an NTP server.
(d) Ensure that the configured time source servers for each device are secure and approved for use by information security.
(e) Ensure time source server redundancy by using at least three synchronized time sources. It is also important to distribute the servers across multiple geographies so that a localized outage will not impact the entire service.
(f) Use authentication for a time source client and server communication to ensure that the server is a trusted server.
(g) Use access control options, such as "ntpd" access restrictions, to restrict access to the time source servers.

AL-SS-R3 – Collect logs in a centralized fashion: For example, utilize syslog, AWS CloudTrail, or Azure Operational Insights. By writing logs to central log servers, the risk of those logs being lost or altered is lowered since they are more secure within the internal network.

(a) Ensure that log servers are approved. To ensure that storage infrastructure components transmit their log event data to the desired syslog server(s), ensure that the syslog server IP address is correct and that the configured syslog server is authorized/approved.
(b) Ensure log server redundancy. Deploy multiple syslog servers to ensure continuous logging and prevent a single point of failure.
(c) Maintain at least one off-site copy for each log.

AL-SS-R4 – Ensure complete audit logging: Ensure that all types of events and all storage-related objects, sites, accounts, etc., are included in the audit logging:

(a) Ensure audit logging for read-only API calls in sensitive environments.
(b) To prevent loss of entries if the logging process is stopped and restarted before all entries are written, ensure that logging is configured to be written in real time to disk with no buffers in place and sent over TCP (not UDP).
(c) Ensure that all denied access attempts to services, ports, files, objects, or devices are logged.
(d) For sensitive information, enable full logging of all data access activity.

AL-SS-R5 – Ensure sufficient audit log retention and protection:

(a) Retain historical logs for a sufficiently long period of time, as it often takes a while to notice that a compromise has occurred or is occurring.

(b) Ensure sufficient storage space and proactively monitor free space and unusual growth rates of log data to prevent log destinations from filling up. A known attack pattern involves filling up logs first to disrupt forensics, and appropriate monitoring can help identify such attacks in real time.
(c) Ensure that archived log data is safe from tampering (e.g., using immutable storage, object locking, MFA delete).
(d) Restrict access to log data and servers. Consider using separate roles or accounts to manage them.
(e) Enable encryption since access to log data can provide attackers with valuable insight into assets and possible attack vectors.

4.5 Preparation for Data Incident Response and Cyber Recovery

Storage-related incidents should be handled as an integral part of the organization incident response process, including aspects such as isolation, root-cause-analysis, defining a response plan, testing, periodical process review and refresh, etc.

The following recommendations incorporate specific aspects that should be considered with respect to storage infrastructure and data assets.

IR-SS-R1 – Develop a response plan for storage component compromise: Consider the following elements in organizational analysis, isolation, remediation, restoration, and testing procedures:

(a) Compromise of an entire array or an entire cloud-based storage asset (e.g., SAN, NAS, object store, elastic file system)
(b) Compromise of a backup system
(c) Compromise of an individual storage element (e.g., share, LUN)
(d) Compromise of a SAN fabric switch

IR-SS-R2 – Ensure immutability of recovery assets during incident management: In conjunction with the guidance provided in Section 4.7 below regarding the protection of cyber recovery copies, ensure that those copies remain isolated during incident management.

IR-SS-R3 – Validate the hygiene of recovered compute components: Ensure that recovered executables, applications, containers, and OS images are free from infection prior to deploying them in production.

4.6 Guidelines for Network Configuration

As previously mentioned, the topic of storage-related networking involves multiple aspects, some of which overlap with *Data Access Control*, *Administrative Access*, and *Encryption*, which have been covered in other sections. To eliminate repetition, this document discusses:

- Certain network recommendations closely related to data access control in Section 4.3,
- Network- and protocol-related encryption recommendations in Section 4.9,
- Certain network recommendations closely related to administrative access in Section 4.10, and

- *Network infrastructure* (e.g., switch, port, HBA and NICs configuration, zoning guidelines, etc.) and *protocols* in this section.

For a complete appreciation of all network configuration aspects, please refer to all sections.

Each recommendation has a unique identifier with the format NC-SS-Rx, where NC stands for Network Configuration, SS stands for secure storage, and Rx stands for the recommendation sequence.

4.6.1 SAN

NC-SS-R1 – Ensure that switches are authenticated: Ensure that storage switches are authenticated before joining the network.

NC-SS-R2 – Ensure the use of an approved PKI mechanism: Use an approved and certified central key management (PKI) system for the management of switch certificates rather than the devices' self-signed certificates.

NC-SS-R3 – A blended approach to zoning: This is preferable to simply zoning using a single type (i.e., host, switch, and storage device):

(a) Host-based zoning mechanisms control what storage resources or devices are visible to an application on a host as well as the devices that it can access. At the lowest level, the masking capability in a host bus adapter's (HBA) firmware or driver must be used to control whether the host may interact with any storage device. At the next level, OS capabilities must be used to control which devices the host tries to mount as a storage volume. Finally, the centralized management software for volume management, clustering, and the file system must be utilized to control device access by applications.

(b) In switch-based zoning, the switches (especially the FC switches) should have the capability to specify which devices on which ports can access other devices or ports. Port-based zoning uses hardware to enforcing zoning and is therefore also called "hard zoning." In other words, switches must support zone control at the level of port WWNs rather than at the switch (node) WWN level.

(c) In storage device-based zoning, the storage array is configured with a list that shows which hosts (even more specifically, which HBA ports) can access which LUNs and on which ports. Access requests from unlisted hosts or HBA ports are ignored or rejected.

(d) If the zone set feature is available, it should be leveraged. This will help to create multiple zones dedicated to a particular purpose, such as testing, dynamic reconfiguration, testing, backup, and maintenance.

NC-SS-R4 – Prefer masking as close to the data as possible and as far from the data consumer or client as possible (e.g., favor array over switch masking, core switch over edge switch, and switch over HBA).

NC-SS-R5 – A copy of the zone configuration file should be kept outside of the SAN switches to enable redeployment upon erroneous or malicious corruption or deletion.

NC-SS-R6 – Limit switch management capabilities to the minimum necessary:

(a) When implementing the SAN fabric, there should be well-defined policies that specify and minimize the set of switches that are authorized to distribute configuration data (while providing acceptable redundancy).

(b) Any unnecessary configuration management permissions and services, password distribution, should not be enabled.

NC-SS-R7 – Considerations for using soft vs. hard zoning:

- **Soft zoning –** Soft zoning uses filtering implemented in Fibre Channel switches to prevent ports from being seen from outside of their assigned zones. The security vulnerability in soft zoning is that the ports are still accessible if the user in another zone correctly guesses the Fibre Channel address. In this case, the FC switch will place a host WWN in a zone without evaluating the port numbers it is connected to in the FC switch. PWWN identification is considered more secure than port number identification (used in hard zoning) because any device physically connected to a port could grant storage access to an unauthorized host. If the SAN spans facilities with different physical security controls, and if there is a risk that physical ports could be accessed by unauthorized individuals, soft zoning may be preferable.
- **Hard zoning –** Hard zoning uses physical port numbers on SAN switches, thereby physically blocking access to a zone from any device outside of the zone. This type of zoning protects from WWN spoofing attacks as it does not rely on host identity. If the organization's physical access is thoroughly protected (i.e., it is improbable that an intruder will access a physical port), this method may be preferable.

NC-SS-R8 –Limit which SAN Fibre Channel physical and logical ports can be used for management on all SAN switches and storage arrays.

NC-SS-R9 – Limit communication between switches: Limit communication between SAN switches based on security policies while ensuring that switches can only communicate with switches that are necessary.

NC-SS-R10 – Persistently disable unused SAN ports to prevent the accidental or deliberate connection of unauthorized equipment

4.6.2 IP Network

NC-SS-R11 – IP storage network separation: When it comes to storage-related communication over IP networks, sound logic should be applied to the separation of environments and traffic type (at both layer 2 and later 3) <u>to the maximum possible extent</u>.

(a) Type of traffic: data access protocols vs. management vs. replication vs. backup vs. host and application networking
(b) In sensitive environments, further separate management traffic of different solutions, vendors, and technologies (e.g., use separate layer 2 and 3 subnets for managing each array technology, Server-Based SAN, switch technology, storage virtualization, etc.).
(c) Data access protocols (e.g., iSCSI vs. NFS vs. proprietary vendor protocols, such as Server-based SAN).
(d) Type of servers or hosts accessing data: virtualized hosts vs. physical hosts

NC-SS-R12 – IP or Ethernet management ports of SAN switches should reside in an isolated subnet, including separation from subnets used for data access between hosts and storage and for host-to-host communication.

NC-SS-R13 – Enable device IP access control: With respect to IP network accessibility, security features should be turned on for all storage devices regulating IPs and port/protocols, where applicable. This includes but is not limited to built-in Firewall rules, IP Filtering, and access lists in order to:

(a) Control data access between hosts or applications and the storage objects they use.
(b) Separately control management IP traffic between management hosts and applications and the relevant storage management interfaces they manage.

NC-SS-R14 – Enable network IP access control: Restriction should be applied at the network level (e.g., routing, firewall, access lists, VPC security groups, server-based SAN clients) to allow all traffic types (e.g., data-access and management traffic) only to allowed IP addresses and TCP/UDP ports and protocols:

(a) Between hosts or applications and the storage objects they use.
(b) Between management hosts and applications and the relevant storage management interfaces of storage objects they manage.

NC-SS-R15 – Block any public access to storage objects, particularly from the internet. Additionally, ensure that sufficient controls are implemented, such as:

(a) Minimizing access
(b) Using physically and logically separate storage subnets and, preferably, separate storage devices and pools
(c) Considering protection from denial-of-service attacks
(d) Cached copies (e.g., using CDN, replicas, and proxies) retaining at least the same security characteristics as the source data
(e) Considering regulatory requirements (e.g., confidentiality, storage location restrictions, etc.)
(f) Any additional applicable security controls (e.g., encryption, authentication, etc.)

NC-SS-R16 – Ensure that internal IP addresses for SNMP: When configuring SNMP, ensure that all traffic is directed to valid and internal IP addresses as destinations.

NC-SS-R17 – Consider the use of isolated non-routable VLAN for server-based SAN: To protect the data storage environment and mitigate security concerns, consider using non-routable VLAN for server-based SAN.

4.6.3 Protocols

NC-SS-R18 – Disable unsecure versions of file access protocols: Outdated, unrecommended, or unsecured protocol versions, such as SMB v1 or NFS 1 and 2, should be blocked. If possible, disable these protocols on both the client side and the server side.

NC-SS-R19 – SNMP security:

(a) Disable SNMP if not required.
(b) Change the default, known community strings, even if SNMP is not enabled. The configured strings should meet the organizational password policy.
(c) Use different community strings for devices that differ in levels of confidentiality.
(d) Use at least SNMP version 3.
(e) Enforce the use of SNMP authentication and privacy (encryption) features.
(f) Do not configure SNMP with read-write access unless it is absolutely needed. In this case, limit and control the use of read-write SNMP.
(g) Use access control lists to control access through SNMP to devices.
(h) Ensure that SNMP traps are sent to authorized, intended managers.
(i) Refer to DHS CISA TA17-156A [35] for additional guidance.

NC-SS-R20 – Ensure the authenticity of directory, domain, and similar services (e.g., AD, DNS, LDAP): Actively and periodically review service configurations in all storage elements (e.g., devices, switches, management workstations, management software) to make sure that the approved ones are used, and remediate any discrepancies.

NC-SS-R21 – Considerations for using standard and non-standard TCP/IP or UDP ports: Most applications and services have a default TCP/IP or UDP port that is used to connect to the application or service. However, since it is usually possible to configure which logical ports will be used by the various applications and services, the pros and cons of using non-standard ports should be considered.

- **Pros** – Using non-standard ports helps obfuscate the application or service as hackers will not know which port to use.
- **Cons** – Alternately, using non-standard ports can make it difficult for security scanning tools to identify suspicious activities since they are designed to expect specific behaviors on standard ports.

NC-SS-R22: Enable FIP snooping filters on FCoE VLANs to prevent unauthorized access to data.

NC-SS-R23 – Limit iSCSI ports: Hosts on the iSCSI network should be prevented from accessing any TCP ports other than the those designated for iSCSI on that network.

NC-SS-R24 – Use iSCSI authentication: Use one of the supported methods to authenticate iSCSI initiators upon opening a session (e.g., CHAP, SRP, Kerberos, SPKM-1/2). When using CHAP, prefer using two-way authentication over one-way authentication.

NC-SS-R25 – Use of NDMP security features: When NDMP is used to transport data from storage arrays to backup systems, ensure that NDMP security features are used, including:

(a) Access control over which hosts can initiate NDMP sessions
(b) The challenge response authentication (do not use the plaintext authentication option)
(c) Log NDMP connection attempts
(d) An NMDP password that meets the organizational password policy (e.g., length, complexity, etc.)
(e) Restricted NDMP-related rights that require user only
(f) Encrypted NDMP control connections
(g) NDMP throttling per session or per server

NC-SS-R26 – Use of LDAP SSL: Use LDAP over SSL when setting up Active Directory options for storage systems.

NC-SS-R27 – Additional protocols: When additional protocols such as SymAPI, SMI-S, GNS, and others are used, further consider adapting the recommendations in Sections 4.6.2 and 4.6.3 for their use. In particular:

(a) Isolate traffic for data access and management from other environments;
(b) Limit TCP / UDP ports; and
(c) Enable encryption.

4.7 Isolation

When production data is damaged or lost, organizations should be able to recover it using replicated or backed up data copies. If the damage is the result of a malicious attack, and the attackers were also able to compromise the backup data copies, the attack on the production environment can have a devastating effect since the organization does not have the ability to recover. To provide wide support for recovery from various scenarios, sufficient isolation must be guaranteed between data assets, classes, and storage systems holding recovery data, in particular.

In this context, organizations should maintain at least two separate types of data protection copies of their data:

- **Non-malicious recovery copies –** To be used in the event of a natural disaster, hardware failure, human error, etc. These can include local copies (e.g., snapshots taken before performing maintenance), DR copies, long-term backups, and others. The "closer" the copy is to the production environment, the more likely it is to be mapped to systems for the purpose of testing and DR.
- **Cyber-attack recovery copies –** Reserved for the event of a cyber-attack and should be hardened, locked, and kept in isolation. These copies should not be impacted by *anything*,

including cases wherein production volumes or other types of copies have been compromised.

The purpose of isolation is to make the cyber-attack recovery copies and systems inaccessible and independent from the production environment. Other levels of isolation (e.g., between long-term backup and production/DR) may be highly advantageous.

The following security recommendations apply to the creation of these copies and the associated management system.

IS-SS-R1 – Ensure separated storage systems:

(a) Cyber-attack recovery copies should be created on designated separated storage environments. In private clouds, this implies physically separated storage systems. In public clouds, this implies separate accounts (or equivalent).
(b) Long-term backup systems should be separated from production data storage systems.

IS-SS-R2 – Ensure separate management systems: Storage systems that store cyber-attack recovery copies should be managed from designated management systems, which are separated from the production environment and other data protection mechanisms. It should not be possible to access them with regular credentials (including production and regular backup). The system should be hosted on a dedicated host that is only connected to an isolated network.

IS-SS-R3 – Ensure restricted access:

(a) Cyber-attack recovery copies and their systems should not be accessible to regular IT staff, only to a single person (e.g. CISO) or to a very narrow group of executives or security managers who use credentials separate from those used for other day-to-day duties. This ensures that if the credentials of an IT admin are compromised, the attacker cannot use those credentials to access the cyber-attack recovery copies. This restricted team can have access to the cyber-attack recovery copies, but an even smaller subset should have administrative rights that include granting permissions to other users.
(b) Access rights to long-term backup should be separate from those used to perform other storage administration duties (e.g., SAN management, storage allocation, etc.) and should include the use of separate user IDs, accounts, and credentials.

IS-SS-R4 – Ensure off-site storage: Cyber-attack recovery copies should be stored off-site and not where the production data is stored. This ensures that if the attacker has physical access to the production site or manages to compromise the physical site, they would not be able to access or compromise the cyber-attack recovery copy.

IS-SS-R5 – Ensure the use of an independent, full baseline copy: Backup systems often make use of incremental backups that capture changes to the data relative to a baseline copy. These incremental copies cannot be used during recovery without the baseline copy. For certain types of backup schemes, such as snapshots, only incremental copies are used (i.e., the baseline copy is the production data itself).

To handle a recovery scenario properly, dependencies between copies must be accounted for, and sufficient isolation between different types of copies must be maintained. In particular:

(a) Replicated disaster recovery copies should have no dependencies on production baseline data;
(b) Long-term backups should have no dependency on production and disaster recovery baseline data; and
(c) Cyber-attack recovery copies should have no dependency on production and disaster recovery baseline data.

IS-SS-R6 – Disable all unneeded services and protocols: Unneeded services and protocols should be disabled on cyber-attack recovery storage systems. Disabling all web access and relying only on API or CLI for management are recommended.

IS-SS-R7 – Ensure independence from hosts and applications:

(a) Cyber-attack recovery copies should not be mounted, exported, or mapped to a host or application. During recovery, if needed, Cyber-attack recovery copies should preferably be restored (pushed) into an isolated staging (or "air-gapped") environment and not directly to the target hosts or applications. A less secure option is to allow the target hosts or applications limited read-only access (e.g., mapping or mounting) during restore only, and remove such access as soon as restore is complete.
(b) Long-term backups should not be mounted, exported, or mapped to a host or application.

IS-SS-R8 – Consider setting up an air gap: Organizations should consider setting up an air gap around the cyber-attack recovery copies. For example, certain storage vendors enable shutting down data ports and opening them during a limited time for the periodic sync with the production system.

IS-SS-R9 – Perform periodic audits: The above recommendations should be checked as part of a periodic audit to ensure that there are no configuration gaps/drifts that may compromise the isolation of the cyber-attack recovery copy.

IS-SS-R10 – Consider the use of immutable storage, which could help further isolate and protect recovery data (e.g., retention locking, vault locking, immutability policies, etc.).

4.8 Restoration Assurance

In order to ensure successful recovery from a cyber-attack, it is not enough to have a process in places. Organizations must also verify that all components of critical data assets are protected and can be restored faithfully, consistently, and completely and that the speed and currency of restoration are aligned with business and regulatory requirements. In many cases, organizations have backups of their critical systems but do not regularly check whether this backup can actually be used to restore the system. However, due to configuration drifts, changes in the environment, or even a malicious attack that compromises the backups, they are faced with a reality in which they cannot use the backed-up data to recover. The following security recommendations apply for obtaining restoration assurance.

RA-SS-R1 – Ensure the completeness of cyber-attack recovery copies: All storage elements that contain components of critical data assets should be protected and backed up in a cyber-attack recovery copy. This includes components such as storage volumes, critical file systems, databases, software images, certificates, encryption keys, startup files, catalog info, ACLs, and configuration files.

RA-SS-R2 – Protect all dependent components: Dependent components, such as Active Directory or DNS, should be protected to enable full recovery.

RA-SS-R3 – Ensure the availability of all relevant software and hardware components: In order to recover, make sure that all of the relevant software and hardware components (e.g., drivers, firmware, etc.) used to run the system are backed up, protected, and available for a restore operation.

RA-SS-R4 – Ensure that selected backup technology and media matches RTO requirements: Recovery speed (RTO) should be examined holistically, including all dependent and related components (e.g., data, configuration files, encryption keys) while also balancing the actual recovery speed that is required with the cost that it would take to align all of the dependent components to enable this expected recovery speed.

RA-SS-R5 – Test restore to ensure required RTO: Perform a test restore to ensure that it is completed successfully and that it meets the required timeframe.

RA-SS-R6 – Ensure that remote replicas meet the RPO: Set a recovery point objective (RPO), which is the amount of data that can be lost following a failure, and ensure that remote replicas meet this objective.

RA-SS-R7 – Ensure that remote replicas meet the retention requirements: Set the data retention requirement, which is the amount of time that data will be backed up. Based on this requirement, ensure that the system creates the relevant number of copies with the relevant refresh rate.

RA-SS-R8 – Ensure that remote replicas are in good health: Periodically ensure that the remote replicas are in good health. This includes checking that there are no relevant errors in the log and that they are in a healthy state.

RA-SS-R9 – Enable the separate restoration of data and application: Separating the data from the applications will allow for the data to be restored without restoring infected code or software.

RA-SS-R10 – Document DR plan, resources, mapping to production, flow, and test procedures: A disaster recovery plan should be written, including all of the resources, its mapping to production, flows, and test procedures. These documents should be backed up as well.

RA-SS-R11 – Ensure cyber hygiene: For mission critical information, cyber-attack recovery copies should be scanned with various anti-malware scanning tools for known vulnerabilities and anomalies. Ideally, all copies should be scanned. If that is not possible, scan a subset of the copies, and keep a record of those copies scanned and secure. Cyber hygiene includes antivirus, anti-malware, vulnerability scanning, and security analytics.

RA-SS-R12 – Perform periodic audits: The above recommendations should be checked as part of a periodic audit to ensure that systems can be restored according to the defined RTO/RPO requirements.

4.9 Encryption

Encryption is the conversion of data from a readable form (i.e., plaintext) into an obfuscated form (i.e., ciphertext) that cannot be easily understood by unauthorized people. In storage systems, the encryption of sensitive information should be implemented end to end, including:

- **Data at rest –** Data that is physically or logically stored in the storage infrastructure (e.g., tapes, disks, optical media) should be encrypted. A comprehensive approach should be taken that incorporates not only the data itself but also metadata, which can include access permissions, labels, paths, and journaling information.
- **Data in transit –** When the data is transferred between storage elements (e.g., read or written by a client, replicated between storage devices or pools, transmitted in server-based SAN, Storage vMotion) and in transit throughout the network, it should be encrypted.
- **Administrative access –** This includes connections through standard and proprietary protocols and APIs to configure or control storage elements, storage networking, and data.

There are a few different types of encryption algorithms that encrypt information and facilitate the encryption process. Asymmetric, symmetric, and hashing formulas are the common methods to enable encryption, with a few variations.

The following encryption guidance is applicable to storage infrastructure and should be used:

EN-SS-R1 – Ensure that secure TLS and SSL levels are used: Certain versions are considered insecure (e.g., TLS 1.0 includes a means by which a TLS implementation can downgrade the connection to SSL 3.0, thus weakening security and exposing it to the POODLE vulnerability):

(a) Enable TLS 1.3 and 1.2, which are the most recent versions of TLS.
(b) Ensure that SSL 2.0, SSL 3.0, TLS 1.0, and TLS 1.1 are disabled across all storage infrastructure components.

EN-SS-R2 – Ensure robust hash algorithms and message authentication codes (MAC): Do not use weak hash algorithms, such as MD5 and SHA1. Strong algorithms, such as PBKDF2 or HMAC-SHA-256, should be used instead. Refer to NIST Special Publication 800-107 and NIST Special Publication 800-132 for additional guidance.

EN-SS-R3 – Disable the use of cleartext protocols (e.g., HTTP, Telnet, FTP, or RSH): Cleartext protocols are vulnerable to sniffing, interception, and other attacks as they do not encrypt traffic or logon details, making it easy for an eavesdropper to intercept this information.

EN-SS-R4 – Ensure that storage management API sessions are encrypted: Storage management APIs and CLIs are used for administrative access to storage systems. For some of the storage systems, the encryption of API and CLI client sessions is controlled in specific per-vendor configuration options within the management software or the API/CLI software component.

EN-SS-R5 – Ensure that administrative access sessions are encrypted: Administrative sessions over HTTP should use SSL (HTTPS). CLI access should be encrypted using SSH rather than Telnet. The authentication during API access should not use cleartext, and the session itself should be encrypted.

EN-SS-R6 – Enable FIPS mode for FIPS-based environments: FIPS 140-2 specifies that a cryptographic module should be a set of hardware, software, firmware, or some combination of those that implements cryptographic functions or processes, including cryptographic algorithms and, optionally, key generation, and is contained within a defined cryptographic boundary. FIPS specifies certain crypto algorithms as secure, and it also identifies which algorithms should be used if a cryptographic module is to be called FIPS-compliant. Organizations that are FIPS-compliant should ensure that FIPS mode is enabled in their FIPS-compliant storage infrastructure components.

EN-SS-R7 – Ensure that sensitive data is encrypted at rest: At-rest encryption protects against a variety of data-related risks (including unauthorized access, compromise in case of media loss or theft, etc.) and should be enabled for sensitive data. Certain considerations should be applied:

- **Infrastructure encryption** (e.g., the use of built-in encryption capabilities provided by a drive, storage array, or cloud storage, whether using the vendor keys or organization-provided keys) can protect against device loss, misplacement, or theft but is not considered an effective control against:
  - **In-band attacks** – When an attacker can compromise a host already mapped to the storage (or when the storage can be mapped using legitimate means to an unauthorized host).
  - Administrators or attackers gaining elevated right – Who can remove encryption.
- **Application-level encryption** – Data is encrypted at its source, presenting ciphertext only to the storage infrastructure and administrators. While generally considered more secure, application-level encryption comes at a (sometimes considerable) cost:
  - **Data-reduction mechanisms are impacted –** For example, compression and dedupe can be drastically less effective.
  - Management is more complex.

EN-SS-R8 – Ensure that data in transit is encrypted:

(a) **Block over Fibre Channel:** Use Fiber Channel link encryption to encrypt data between SAN ports. When supported, use end-to-end (host to storage) encryption (FC-SP-2 compliant).

(b) **Block over IP:** IP storage traffic is subject to the same security risks as regular IP networks. By default, block-over IP protocols do not provide data confidentially, integrity, and authentication per packet. When using block-over IP protocols (e.g., iSCSI, FCIP, iFCP or iSNS, proprietary protocols), configure IPsec to ensure per-packet authentication, integrity, and confidentiality.

(c) **File and object storage access:** Ensure that data encryption in-flight options are enabled (for backup systems) and that data remote replication is encrypted. For file access, ensure that data is transmitted and encrypted using mechanisms such as SMB encryption and NFS Kerberos. Ensure that objects are accessed through HTTPS with TLS.

(d) Particular attention should be paid to enable encryption on all connectivity segments that extend network communication beyond the boundaries of a physically protected domain (e.g., an ISL link between two physically separated datacenters, IP traffic over WAN or the internet).

EN-SS-R9 – Ensure that internal storage communication is encrypted: Storage systems often have subsystems and components that communicate with each other, such as nodes, managers, and witness devices. Ensure that internal communication between subsystems is properly encrypted. This can also be extended to communication with gateway servers, policy servers, and antivirus servers.

4.10 Administrative Access

Administrative access is required to control and managed a wide spectrum of storage elements, including arrays, network and fabric, management tools, backup, replication, and cloud storage. Administrative access can be based on a direct connection to the storage component and through a management software. Both connection types can involve various interfaces, including a management UI, CLI, and API.

Securing administrative access is critical, as most storage risks discussed in Section 3.2 above, including the most devastating, could materialize if not well-controlled.

Certain other sections in this chapter include aspects that overlap with Administrative Access. To eliminate repetition, additional relevant recommendations can be found in:

- Section 4.9 above, regarding encryption
- Section 4.3 above, regarding data-related access controls, part of which may also apply to administration

The following security guidelines are recommended for the configuration of administrative access.

AA-SS-R1 – Network access to management ports of the SAN switches should be limited to devices and administrators specifically assigned to manage the switches through a mechanism such as an access control list (ACL).

AA-SS-R2 – Control and limit the devices and components that have administrative capabilities to the minimum required: This includes CLI servers, management consoles, API gateways, witness hosts, and storage devices with control permissions. In particular:

(a) Actively discover components that have storage administration capabilities to make sure that only authorized ones have them. Remove unnecessary ones, if found, and debrief.
(b) Remove unnecessary rights and capabilities from authorized devices.

AA-SS-R3 – Implement the least-privileges approach: Limit the rights of users with administrative rights to the minimum required. This includes the minimal actions that the user can carry out while also limiting the scope of these permissions to include only the relevant systems or regions. Full administrative rights should only be granted to users who require these rights.

AA-SS-R4 – Limit monitoring tools' access rights: Service accounts, such as monitoring tools, should be limited to read-only and metadata-only access.

AA-SS-R5 – Use IP filtering on storage systems: Most storage systems offer the ability to manage the list of hosts that are allowed to administer the system. This capability should be used to explicitly restrict management rights to designated and documented hosts using IP filters.

AA-SS-R6 – Authenticate/authorize all CLI/API access: CLI/API usage should be subject to authentication and authorization processes. In cases where it is not possible to perform authentication or authorization, secure the unauthorized access with additional security measures, such as using privilege management tools to restrict control to the minimum required commands and objects.

AA-SS-R7 – Favor API access control over CLI/shell access: API access is more restricted than CLI/shell access. For example, through the CLI access, it is potentially possible to access the system's underlying operating system and file system, which can be used to access configure files. CLI also often includes features that are not documented yet can be found through research on the web.

AA-SS-R8 – Restrict management consoles OS privileges: Management consoles should not run as root users but rather as storage-designated accounts (see also **AC-SS-R20)**. Their web service should be hardened to meet the minimum standards of other web-application servers in the organization.

AA-SS-R9 – Restrict host storage control privileges: In certain shared data configurations (e.g., clusters, geo-clusters, scale-sets, or storage virtualization infrastructure), hosts are granted administrative access to storage to control shared cluster data resource allocation and behavior. When such administrative access is necessary, restrict the scope and privileges granted to hosts to the maximum possible extent:

(a) Only to the particular elements (e.g., LUNs, shares, files, objects) that the hosts need to control
(b) Only to the specific actions that the hosts need to perform

AA-SS-R9 – Command device or gatekeeper configuration: Certain storage arrays allow in-band administrative control to hosts that have access to special block devices (e.g., "command device," "gatekeeper"). Commands are transferred using I/O operations on those special devices. When used, the following security guidelines are recommended apply:

(a) **Limit the use of control devices to the minimum possible** – If feasible, eliminate the use of such devices completely (e.g., using API access instead). If not, ensure that they are mapped to required hosts only (e.g. management hosts).
(b) **Scan for control devices** – Perform network scanning to discover control devices, and ensure that they are mapped to the required and authorized hosts only.

AA-SS-R10 – Disable or limit call home or remote access: Storage infrastructure systems may have the ability to send certain telemetry and diagnostic data back to the manufacturer, such as logs. In some cases, they even enable remote connection to the system by the manufacturer with administrative rights. These mechanisms are in place to allow the manufacturer to investigate and resolve technical issues. These capabilities could potentially be exploited by hackers and should be disabled if they are not required. However, if they are required, they should be limited and controlled by implementing the following settings:

(a) **Change the default credentials** – Modify the default credentials used for the remote connection.
(b) **Limit permissions** - Limit access permissions to only the minimal level required.
(c) **Enforce encryption** – The remote session should be encrypted.
(d) **Limit access with an "allow list"** – Manage access with an access list that limits access by specific IPs and specific users.
(e) **Fully logged** – All remote access should be fully logged for auditing purposes.
(f) **Enable built-in data obfuscation features** – This is applicable for those storage devices that allow obfuscating sensitive data, such as IP addresses, WWNs, device names, and usernames.
(g) **Limit the scope of data sent** to the minimum required
(h) **Review and approve** – Periodically evaluate the data that is occasionally or automatically being sent to the vendor, and ensure that it does not contain sensitive information, such as IP addresses, usernames, or the actual content of storage devices.
(i) **Authorize each connection** – If possible, implement a mechanism that will ask permission before allowing each connection.
(j) When remote access by the vendor is performed through a gateway device, server, or appliance, take particular care to secure and restrict access to the gateway system.
(k) Consider the use of dedicated (private) links to the vendor over the use of the internet.

AA-SS-R11 – Limit network access for management: In addition to separating management from other traffic (see Sections 4.6 and 4.7), in sensitive environments, it is recommended to further control access to management networks, including mechanisms such as:

(a) Using VPN, IP-SEC, or one or more "jump servers" (or "login proxies," which are dedicated servers in the management network that are the only ones accessible from outside of the network and can serve to connect to other servers after proper authentication and authorization).
(b) Enhanced logging, tracing, and session recording.

AA-SS-R12 – Secure and protect core storage management files and binaries: Storage management software often includes configuration files that present various options to control how the storage system would operate, including undocumented options. Such sensitive directories and files should be kept with appropriate limited permission and with correct ownership and group membership. This includes:

- Configuration files outlining users and roles, network settings, consistency groups, device groups, and other storage options. The configuration files that define consistency and device groups are often automatically propagated from central management hosts to other hosts that are attached to the managed storage system. Thus, if compromised, it can affect multiple systems.
- Scripts to control starting, monitoring, and stopping storage management services and daemons as well as the binaries themselves should be kept in a secure way.

Apply the following controls to critical files:

(a) Restrict access and permissions, and control ownership of key folders and files.
(b) For sensitive environments, consider monitoring for content changes in such files to prevent unauthorized ones.

4.11 Configuration Management

The purpose of configuration management is to provide visibility and control over settings, behavior, and the physical and logical attributes of storage assets throughout their life cycle. In the context of storage security, this involves:

- Maintaining comprehensive and current inventory,
- Managing change, and
- Ensuring that the configuration continually meets the organization's security baselines and current industry best-practices and that it is free of known risks.

To this end, appropriate controls, policies, processes, and tools are required. The following paragraphs contain guidelines applicable to achieving those ends.

CM-SS-R1 – Create a comprehensive inventory of all storage devices: This includes identifying the name, address, location, and software, firmware, or driver versions for all storage components, including:

- Arrays
- Storage virtualization systems
- Management consoles
- Witness hosts
- Hosts installed with storage management software and/or plugins
- Data protection appliances
- Backup clients and servers
- Storage network switches
- Storage adapters or "HBA"
- I/O multipathing software

- Pairing of primary and (replication) destination storage systems
- Designated backup servers for hosts or off-site backup
- Tape libraries and drives

CM-SS-R2 – Create a comprehensive inventory of all data and configuration assets: This includes identifying logical data components and data access configurations through the following assets:

- Storage pools, LUNs, masking, and zoning
- Initiators and initiator groups
- File shares and ACLs
- Object storage pools, buckets, etc.
- Replicas and snapshots
- Backup catalogue and access rights
- Backup sets (on-premises, archived, virtualized in the cloud, on tapes, archive appliances, etc.)
- Users, groups, roles, and rights
- Host access configuration to storage assets (e.g., LUNs, file shares, global file systems, object storage)
- Images of storage software, virtual appliances, etc.

CM-SS-R3 – Create a comprehensive storage security policy, either as a dedicated policy or as part of the organization's security policy. It should include configuration baselines for storage systems and could be based on:

- Recommendations from this publication and cited sources
- Storage-related security standards internal to the organization
- Relevant vendor security-best practices

CM-SS-R4 – Keep the storage security policy current: The storage security policy should be reviewed and updated periodically (at least annually). The security baseline should be updated with the latest vendor and industry recommendations available for storage systems and/or specific storage devices (preferably on a quarterly basis, at least).

CM-SS-R5 – Periodically and proactively assess configuration compliance to storage security policy:

(a) Make sure that the actual configuration meets the storage security baselines, and identify gaps.
(b) Track the remediation of gaps in a timely manner.
(c) Consider developing KPIs to track the compliance to storage security baselines based on types of data, their organization function, and their sensitivity.

CM-SS-R6 – Create a storage change management process as a dedicated process or as part of the organization's general change management process. It should cover:

(a) Planning, reviewing, and approving storage configuration changes;
(b) Updating environment documentation and inventory (e.g., infrastructure, data, configuration); and
(c) Assessing compliance to relevant security baselines following any change to the sensitive storage environment.

CM-SS-R7 – Detect unauthorized storage security changes: There should be a process for detecting unauthorized changes, prompt remediation, and thorough debriefing

CM-SS-R8 – Software updates and patches:

(a) **Release updates** – There should be a process for periodically updating storage software to the latest stable and secure storage release available. This includes management software, API and CLI packages, array and HBA firmware versions, and OS drivers.
(b) **Important security updates and patches** – There should be a process to proactively and frequently install important and urgent storage security fixes and patches.

CM-SS-R9 – Network topology documentation: Maintain current storage-related network documentation, including drawings (SAN and IP).

CM-SS-R10 – Ensure the propagation of SAN security configuration changes: Many security changes are not automatically or reliably propagated across all switches in the fabric. There should be a process for enforcement and validation that all such changes are distributed and activated throughout the fabric.

5 Summary and Conclusions

Starting with an overview of the storage technology landscape, this document has discussed the threats and resulting risks to the safe utilization of resources. It then provided detailed security recommendations for the secure deployment, configuration, and operation of storage resources in various security focus areas. These focus areas spanned the following:

- Focus areas that are common to all IT infrastructures, such as physical security, authentication and authorization, audit logging, network configuration, change management, incidence response and recovery, administrative access, and configuration management.
- Focus areas that are specific to storage infrastructures, such as data protection, confidentiality protection using encryption, isolation, and restoration assurance.

Along with compute (encompassing OS and host hardware) and network infrastructures, storage infrastructure is one of the three fundamental pillars of IT. However, compared with its counterparts, it has received relatively limited attention when it comes to security, even though data compromise can have as much of a negative impact on an enterprise as security breaches in compute and network infrastructures. The comprehensive security recommendations for storage infrastructures in this document seek to close that gap.

Building an effective risk management program for storage infrastructure based on the security controls described in this document and tightly integrating it with existing cybersecurity frameworks [36] could significantly improve an organization's resilience to data breaches.

References

[1] Storage Networking Industry Association (2020) *What is iSCSI?* Available at https://www.snia.org/education/what-is-iscsi

[2] Chadalapaka M, Satran J, Meth K, Black D (2014) Internet Small Computer System Interface (iSCSI) Protocol (Consolidated). (Internet Engineering Task Force (IETF) Network Working Group), IETF Request for Comments (RFC) 7143. https://tools.ietf.org/html/rfc7143

[3] International Organization for Standardization/International Electrotechnical Commission (2020) ISO/IEC TS 23167:2020 – *Information technology – Cloud Computing – Common technologies and techniques* (ISO, Geneva, Switzerland). Available at https://www.iso.org/standard/74805.html

[4] Webopedia (2020) *Block-level storage*. Available at https://www.webopedia.com/TERM/B/block-level-storage.html

[5] TechTarget (2018) *Fibre Channel*. Available at https://searchstorage.techtarget.com/definition/Fibre-Channel

[6] Haynes T (2016) Network File System (NFS) Version 4 Minor Version 2 Protocol. (Internet Engineering Task Force (IETF) Network Working Group), IETF Request for Comments (RFC) 7862. https://tools.ietf.org/html/rfc7862

[7] Black D, Glasgow J, Faibish S (2012) Parallel NFS (pNFS) Block Disk Protection. (Internet Engineering Task Force (IETF) Network Working Group), IETF Request for Comments (RFC) 6688. https://tools.ietf.org/html/rfc6688

[8] Webopedia (2020) *SDS – software defined storage*. Available at https://www.webopedia.com/TERM/S/software-defined_storage_sds.html

[9] The Channel Company (2013) *Software-defined Storage Strategy*. Available at https://www.crn.com/news/storage/240150933/emc-outlines-software-defined-storage-strategy-plans-product-release-this-year.htm?itc=xbodyjk

[10] International Organization for Standardization/International Electrotechnical Commission (2020) ISO/IEC 27040:2015 – *Information technology – Security Techniques – Storage Security* (ISO, Geneva, Switzerland). Available at https://www.iso.org/standard/44404.html

[11] SDxCentral (2015) *What is Storage Virtualization?*. Available at https://www.sdxcentral.com/data-center/storage/definitions/what-is-storage-virtualization/

[12] Webopedia (2020) *VM-aware storage*. Available at https://www.webopedia.com/TERM/V/vm-aware_storage.html

[13] Azeem SA, Sharma S (2019) Study of Converged Infrastructure & Hyper Converged Infrastructure As Future of Data Centre. *International Journal of Advanced Computer Research* 8(5):900. https://doi.org/10.26483/ijarcs.v8i5.3476

[14] Gartner (2016) *Magic Quadrant for Integrated Systems*. Available at https://www.gartner.com/document/3471517?ref=ddisp&refval=3500917

[15] Gartner (2019) *Magic Quadrant for Hyperconverged Infrastructure*. Available at https://www.gartner.com/document/3975501?ref=ddisp&refval=3975577

[16] Gartner (2016) *Critical Capabilities for Integrated Systems*. Available at https://www.gartner.com/document/3500917?ref=ddisp&refval=3471517

[17] Gartner (2018) *Cool Vendors in Storage Technologies*. Available at https://www.gartner.com/document/3893182?ref=solrAll&refval=240504932

[18] ActualTechMedia (2017) 2017 *State of Storage in Virtualization*. Available at https://www.actualtechmedia.com/wp-content/uploads/2018/01/ActualTech-Media-Survey-Report-finalv1.pdf

[19] Trust Radius (2020) *Cloud Storage Systems*. Available at https://www.trustradius.com/cloud-storage

[20] Storage Networking Industry Association (2018) *Storage Security: Data Protection* Available at https://www.snia.org/sites/default/files/security/SNIA-Data-Protection-TechWhitepaper.pdf

[21] Galibus T, Krasnoproshin VV, De Oliveira Albuquerque R, Pignaton de Freitas E (2016) Elements of Cloud Storage Security (Springer Nature, Switzerland AG)

[22] Blueliv (2018) *Credential theft: the business impact of stolen credentials*. https://www.blueliv.com/cyber-security-and-cyber-threat-intelligence-blog-blueliv/credential-theft/credential-theft-blog-news-and-articles-blueliv/

[23] JumpCloud (2019) *Credential Theft: How It Works and How to Mitigate It*. Available at https://jumpcloud.com/blog/credential-theft-mitigation/

[24] MalwarebytesLABS (2018) *Encryption 101: How to break encryption*. Available at https://blog.malwarebytes.com/threat-analysis/2018/03/encryption-101-how-to-break-encryption/

[25] TechNadu (2018) *How is Encryption cracked?*. https://www.technadu.com/how-is-encryption-cracked/36616/

[26] Available at https://www.vpnmentor.com/blog/difference-between-malware-ransomware/

[27] Wikipedia (2020) *Privilege escalation*. Available at https://en.wikipedia.org/wiki/Privilege_escalation

[28] Joint Task Force (2020) Security and Privacy Controls for Information Systems and Organizations. (National Institute of Standards and Technology, Gaithersburg, MD), Draft NIST Special Publication (SP) 800-53, Rev. 5. Available at https://doi.org/10.6028/NIST.SP.800-53r5-draft

[29] Norton (2020) *Sensitive data exposure: What is it and how it's different from a data breach*. Available at https://us.norton.com/internetsecurity-privacy-sensitive-data-exposure-how-its-different-from-data-breach.html

[30] MSSPAlert (2019) *Ransom Attacks Target Backups, NAS (Network Attached Storage)*. Available at (https://www.msspalert.com/cybersecurity-news/attacks-target-nas-backups/)

[31] Wikipedia (2020) *Attack Surface*. Available at https://en.wikipedia.org/wiki/Attack_surface

[32] TechTarget-SearchITChannel (2007) *Fibre Channel man-in-the-middle attacks*. Available at https://searchitchannel.techtarget.com/feature/Fibre-Channel-man-in-the-middle-attacks

[33] Security Week (2018) *Hackers Can Stealthily Exfiltrate Data vis Power Lines*. Available at https://www.securityweek.com/hackers-can-stealthily-exfiltrate-data-power-lines

[34] Ross RS, Pillitteri V, Dempsey KL, Riddle M, Guissanie G (2020) Protecting Controlled Unclassified Information in Nonfederal Systems and Organizations. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 800-171, Rev. 2. https://doi.org/10.6028/NIST.SP.800-171r2

[35] Department of Homeland Security (2017) *Reducing the Risk of SNMP Abuse*. Available at https://us-cert.cisa.gov/ncas/alerts/TA17-156A

[36] National Institute of Standards and Technology (2018) Framework for Improving Critical Infrastructure Cybersecurity, Version 1.1. (National Institute of Standards and Technology, Gaithersburg, MD). https://doi.org/10.6028/NIST.CSWP.04162018