1 Introduction

Storage, computing, and networking form the three fundamental building blocks of any information technology (IT) infrastructure. 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 the area of storage system architecture that uses 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 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., local area network – LAN – or wide area network – WAN), while in the other, there is a dedicated network for communicating with all storage resources. 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 network file system (NFS), or server message block / common internet file system (SMB/CIFS). The latter is exemplified by the storage area network (SAN), an implementation of which uses a specialized, high-speed network (e.g., Fibre-Channel) that provides block-level access to storage. Another implementation of a SAN uses the Internet Small Computer Systems Interface (iSCSI) protocol over a possibly shared LAN/WAN infrastructure.

A variation of the traditional enterprise storage infrastructure is the emergence of converged and hyper-converged infrastructure (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. Similar to 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. Unlike a converged system, HCI does not require discrete storage. 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, compute, and networking components are designed to be managed as a single system across all instances of a hyperconverged infrastructure. Furthermore, 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.

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

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, and allow addressing discrete, unstructured data elements by ID or metadata.
- 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 by storage service interface type (e.g., block, file, and object).
- Network-based storage (e.g., NFS, SAN).
- Storage systems that have a layer of software abstraction (e.g., software-defined storage and storage virtualization).
- Storage systems designed exclusively for virtualized server environments (e.g., storage for Virtual Machines (VMs) and containers, converged and hyperconverged storage systems).
- Storage systems designed with Application Programming Interfaces (APIs) for cloud access.

The security recommendations span the following operations:

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

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

Storage infrastructure for Mainframes is out of scope for this document.

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.

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

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 and 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 with 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 800-52 Rev. 2, *Guidelines for the Selection, Configuration, and Use of Transport Layer Security (TLS) Implementations*.
- SP 800-53, Rev. 5, *Security and Privacy Controls for Information Systems and Organizations* (2020).
- SP 800-57 Part 1, *Recommendation for Key Management: Part 1* (2020).
- 800-63A, *Digital Identity Guidelines: Enrollment and Identity Proofing* (2020).
- 800-63B, *Digital Identity Guidelines: Authentication and Lifecycle Management* (2020).
- SP 800-88 Rev. 1, *Guidelines for Media Sanitization* (2014).
- 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 the associated risks. Apart from generic threats (e.g., privilege escalation, credential theft or compromise, cracking

encryption, malware and ransomware), storage infrastructure-specific threats such as unauthorized storage configuration changes, media theft, and insecure storage images are also discussed. The resulting risks to storage infrastructure (e.g., data breach or exposure, unauthorized data alteration and addition, data corruption, data obfuscation and encryption, tampering of storage-related log and audit data, and compromising backup and firmware) 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.
- Appendix A provides a list of acronyms used throughout this document.

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

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 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 is called 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. Furthermore, 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 Transmission Control Protocol/Internet Protocol (TCP/IP), Fibre-Channel over IP [FCIP], Fibre-Channel over Ethernet (FCoE), cloud block storage). In networked storage, the remote storage devices are presented as if they are locally attached to the host system on which storage client software is running.

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).

The data storage media landscape includes many technologies, and is constantly evolving. Common technologies include:

- Magnetic (e.g., spinning disk drives, tapes).
- Optical (e.g., optical drives such as CD-R, DVD-R, Blu-ray), and magneto-optical.
- Semiconductor (e.g., SSD, flash drives, persistent memory devices).

Experimental and less prevalent technologies include molecular memory (e.g., polymer-based), holographic (e.g., crystal-based), DNA-like, and other.

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

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

through a SAN. 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.

2.1.1 Storage Area Network (SAN)

The building block of SAN-based systems typically include (a) host computers (clients); (b) topology, which involves the distribution of switches; and (c) storage devices/arrays, with all three components interconnected using various network stacks. Since SAN is a specialized, high-speed network for block-level network access to storage, a more detailed look at its variants is warranted [1]. The variants are the results of different types of network stacks with different protocols in certain layers of the stack. There are several SAN protocols. Some of the widely deployed ones are:

a. Fibre-Channel SAN (FC SAN) [2]
b. IP SAN
c. Fibre-Channel over Ethernet (FCoE)
d. Protocols of Non Volatile Memory express (NVMe) over fabrics (NVMe-oF)

An FC SAN is a network stack that uses the Fibre-Channel protocol that has five layers (i.e., FC0 through FC4, unlike the seven-layer Open Systems Interconnection (OSI stack)). The logical storage resource addressed by Fibre-Channel SAN is called the logical unit number (LUN). IP SAN is a network stack that uses the IP protocol at the network layer. iSCSI [3] is an example of IP SAN. In FCoE, the Fibre-Channel frame is encapsulated in Ethernet packets.

Although such use is technically inaccurate, a block device in an FC SAN or IP SAN is often referred to as “LUN”. The historical roots of the term stem from SCSI logical unit numbers, or LUNs, that got carried along to SANs.

Before we look at the protocols of NVMe-oF, it is necessary to look at NVMe. NVMe is the standard host controller interface for systems using PCI Express (PCIe)-based SSD. The NVMe over Fabrics (NVMe-oF) specification defines a protocol interface and related extensions that enable the NVMe commands to be transmitted over the network. Thus, NVMe-oF extends the NVMe deployment from a local host to a remote host for a scale-out NVMe storage system.

The protocols of NVMe-oF include Remote Direct Memory Access (RDMA) (a family of protocols), Fibre-Channel, and TCP. RDMA [4] enables server-to-server data movement directly between application memory without any CPU involvement thus allowing a local application to read or write data on a remote computer's memory with minimal demands on memory bus bandwidth and CPU processing overhead, while preserving memory protection semantics. Infiniband [5] is one of the industry standard interconnects that enables RDMA for high performance computing (HPC) environments.

The topology of the nodes together with the various hardware elements in a SAN system is called the SAN fabric. For example, a Fibre-Channel SAN fabric consists of a Name Server (NS) that registers and communicates with switches and end points (Host Bus Adapters – HBAs). There are two kinds of topologies in FC SAN: point-to-point (two devices are directly connected), and switched fabric. In switched fabric, there is a set of hardware switches acting as

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

one big logical switch that connects the host computer and the storage resources. Security recommendations in this document cover only this topology, as it is the one commonly deployed.

2.1.2 Other forms of Networked Block Storage

Other forms of block storage that could be presented to hosts over IP networks include:

- Hyper-Converged storage service (see 2.9, "Converged and Hyper-Converged Storage (Server-based SAN)" below).
- Cloud block storage service, offered in all cloud environments (See 2.10, "Cloud-Based Storage System" below).

2.2 File Storage Service

This type of service presents storage resources in the form of a file system model with files contained in directories within volumes. Behind the scenes, files can be replicated by making redundant copies or be encrypted. The different variations of this service and their associated protocols are:

- Network-attached storage (NAS) with Network File System (NFS) protocol [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. The files or folders can also be shared from either a dedicated appliance (typically referred to as an "NAS device" or "NAS array") or from any host running the NFS server service.
- NAS with SMB protocol connection – This is provided by a LAN-attached file server, 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, and underlies the CIFS file sharing service.
- 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., Access Control List (ACL)/permissions specifications), and some conflicts in access control rights may have to be resolved during access requests.
- NAS with parallel NFS protocol (pNFS) – This is provided through a clustered collection of storage servers (instead of a single NFS server) that slices and/or stripes data and metadata at the back end while providing dynamic, distributed client connections at the front end across the set of clustered hosts. The pNFS 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 clustering) or (b) splitting functionality across servers (called asymmetric clustering) 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].

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

2.3 Object Storage Service

An object storage service presents data as flexible-size discrete buckets or containers storing objects. Unlike the fixed-size blocks offered by a block storage service or the directories and subdirectories of a traditional file system or NAS service, each object can be of arbitrary size and is assigned a unique identifier (an object ID number or OID), which is compiled with other OIDs into a flat index that's used to access the data in each object. In addition, dynamic metadata can also be attached to an object to facilitate flexible search and addressing.

Technically, an object can be of almost any size and could contain multiple files, or only fractional files. A common application for object storage is as an archive for unstructured data, especially large file data such as digital content. But it can be used for more active data as well, like that stored by an online application provider.

The primary benefit of object storage is scalability, since its flat index is more efficiently searched than a traditional file system. With only an OID required to search for data (plus an offset into the object itself), lookups are essentially a two-step process, versus the multiple steps required to walk the directory tree of a traditional file system. The immediate effect of this is a reduction in metadata handling by the storage system, which means faster file access, especially as the system grows to very large proportions.

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-256) that is the identifier of the document it refers to and is used to retrieve the document.

CAS is used for retrieving documents with short- and medium-term retention requirements and is not widely adopted.

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 (i.e., 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., Structured Query Language – or SQL – database clients). These services are available both in enterprise data centers and in the cloud. 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.

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

SQL database services enable the storage and retrieval of structured data that is typically in a tabular format (also called relational tables). The access is enabled through the standardized interactive programming language SQL [International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 9075:2016 Database languages – SQL]. Current SQL databases can not only store data using relational tables and views but also other structures, such as eXtensible Markup Language (XML), JavaScript Object Notation (JSON), and Binary Large Objects (BLOBs).

Messaging queue storage services are specialized 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 integration with stream processing, where events related to multiple message storage and retrieval by distributed system components can be analyzed to discover patterns.

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

2.6 Software-Defined Storage

Software-defined storage (SDS) includes pools of storage with data service characteristics that may be applied to meet the requirements specified through the service management interface [8].

Flavors of SDS can be found in private and public cloud environments, HCI, and various software solutions. It is a storage architecture 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:

- Flexibility to use heterogeneous storage hardware without the issues of interoperability.
- Enabling of functions such as deduplication, replication, snapshots, and thin provisioning using industry-standard server hardware, though this feature is not unique to SDS.
- Automatic and efficient allocation of pooled storage resources to match the application needs of the enterprise.
- Speed of deployment.

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

- 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.
- Ensure that the storage software understands and leverages the capabilities of storage hardware.

2.7 Storage Virtualization

Storage Virtualization is the act of abstracting, hiding, or isolating the internal function of a storage (sub) system or service from applications, compute servers, or general network resources, for the purpose of enabling application and network independent management of

storage or data [9]. 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 in large networks (e.g., SAN) or data centers. Virtualization also allows segregating a storage resource, or a resource pool, to multiple virtual representation of such resources. 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). Furthermore, the composition of physical disk drives in the mirrored volume can be changed, and devices can be written concurrently on both mirrors ("active-active").
- 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.
- Large scale data migrations – that need to be performed transparently to hosts and applications.

The benefits of storage virtualization are scalability, performance, redundancy, and increased storage resource utilization.

2.8 Storage for Virtualized Servers and Containers

A virtualized server is one where a single physical server runs multiple computing stacks (each consisting of an operating system (OS), storage, network, and applications) called virtual machines (VMs) with the use of software called the hypervisor. Storage infrastructure specifically designed for use with virtualized servers is often called virtualization-aware storage or VM-aware storage [11]. In most environments, this infrastructure is managed together with the VMs by the hypervisor rather than as separately managed block devices.

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 quality of service (QoS) requirements for the applications hosted on the VMs. For example, since the VM-aware storage system maps storage to VMs, management tasks like performance monitoring can gauge issues like storage latency to the VM level. To implement VM-aware storage, the storage management system should implement APIs callable from a hypervisor. In other words, a hypervisor integration software layer is placed atop a conventional storage array – where the array itself can use any storage media such as magnetic and/or flash disk [12].

Since the management functions in a VM-aware storage infrastructure are enabled using software, they can be looked upon as 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 [13].

Containers offer a lighter form of packaged compute, network, and storage units. Multiple containers can run on a server, a VM, or specialized clusters that provide orchestration services.

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

Persistent Storage for containers [14] is provided by creating volumes through a newly-created file directory local to the host where the container is running, or through mapping to an external SAN or NAS device using plugins. These volumes can be created ahead of time or at the time the container starts, and can be shared by multiple containers. Plugins are provided by the storage vendors to facilitate the process of volume creations while conforming to the specification of the container engine/orchestrator. Plugins automate the process of LUN/volume creation and mapping to the host and eventually to the container.

2.9 Converged and Hyper-Converged Storage (Server-based SAN)

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 server room 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 [15]. 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). These capabilities introduce a storage abstraction layer for remote disk access.

In HCI, the 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, Dell (VxRail and PowerFlex Rack), Cisco (HyperFlex), and SimpliVity [16]. 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 may need to be shared to perform storage access and management functions. The overall management software stack may include a compute node, a hypervisor, and SDS software, depending on the deployment environment (e.g., virtualized infrastructure, virtual desktops, unstructured data stores, high performance computing) [17]. 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 [18]:

- Optional data reduction features, such as 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 PCIe bus and running NVMe protocol [19]. However, commercial implementations based on this architecture are not currently available.

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

2.10 Cloud-Based Storage System

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 [20]:

- 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-premises storage infrastructure.

Storage systems based in the cloud provide several sophisticated data services [21]:

- 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, and 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) may also be 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.

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

2.11 Storage and Data 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 section 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:

- Storage resource configuration and resource management
- Data Classification or Categorization
- Data Sanitization
- Data Retention
- Data Protection
- Data Reduction

2.11.1 Storage Resource Configuration and Resource Management

Storage resource configuration and resource management involve the complete lifecycle management of storage infrastructure, and include:

Management and control of physical storage devices, such as storage arrays, SAN switches, including software updates, device configuration (including security aspects), and device onboarding and disposal.

- Change orchestration across multiple assets.
- Management of storage resources (e.g., block device, filesystems, pools), including assignment to hosts, replication, point-in-time copy management, data migration, and data tiering.
- Performance management and optimization.
- Capacity management and optimization.
- Inventory management.
- Event management.

2.11.2 Data Classification or Categorization

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

- Sensitivity (e.g., sensitive vs. non-sensitive)
- Frequency (e.g., frequently accessed vs. non-frequently accessed)
- Environment (e.g., Production vs. Development vs. Testing vs. Staging)

Sensitivity classification is required to enable provisioning of appropriate security controls (e.g., authentication, authorization, encryption, key management, sanitization). Furthermore, the sensitivity category may require sub-categories based on regulations applicable to the data, such as personally identifiable information (PII), Health Insurance Portability and Accountability Act (HIPAA)-related, and Payment Card Industry Digital Security Standard (PCI-DSS). Frequency classification is required to provision the appropriate storage media (e.g., SSD vs.

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

HDD). The environment classification may be required for both media selection and security controls. Other classification schemes based on project, application, etc. may exist but are not listed above due to them not necessarily having security control implications.

2.11.3 Data Sanitization

Sanitization is the process of rendering previously written data irretrievable, such that there is reasonable assurance that the data cannot be accessed or reconstructed [10]. There are three methods for sanitization:

- Clear (e.g., overwriting of the existing data),
- Purge (e.g., using a strong magnetic field for magnetic media degaussing, cryptographic erase for encrypted data), and
- Destruct (e.g., physical destruction of the media, such as burning, pulverizing, etc.).

Factors that determine the appropriate type of sanitization include the category of information on the media, the nature of the media (solid-state, magnetic, or optical), and the reuse plans for the media.

Sanitization can be applied to individual storage media, or to logical data (e.g., in the cloud). To achieve its goals, sanitization must consider the type and characteristics of media data. For example, overwriting data is effective on magnetic disks but not on solid state drives using flash memory because those devices do not overwrite data in place.

2.11.4 Data Retention

There may be situations where preserving access to particular data is needed for a short-, medium- (i.e., less than 10 years), or long-term duration. Data retention is usually satisfied by keeping a copy of data on some backup medium. This may be due to operational, legal, regulatory, or statutory requirements.

2.11.5 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 unauthorized disclosure, modification, or destruction.
Data protection involves activities and mechanisms that span the entire storage lifecycle. These phases include [22]:

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

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

The range of objectives and associated activities provides a taxonomy for classifying data protection activities under three facets: storage, privacy, and information assurance/security [22]. 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 dedicated section to discuss their details. Hence, in this section, only storage-related data protection activities and controls are discussed. This category of controls includes:

- Data backup and recovery,
- Archiving,
- Replication technologies,
- Immutability,
- 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 a 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).

Archiving involves the storage of data for long-term retention. While separate techniques and solutions may be used for backup and archiving, there is often a close relationship between the two, as the later frequently involves capturing copies of the former for longer-term storage, and additional classification. Archiving supports:

- Information lifecycle management – where different types of records can be captured, retained, and disposed of, based on organization requirements.
- Record search and data discovery (i.e., detection of records based on identifiable attributes, such as a person's ID).
- Meeting regulatory requirements for data protection, retention, and legal preservation.

Data replication is the process of writing the same data to at least two separate locations [20]. 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 disaster recovery (DR) site). Asynchronous replication involves time delay and may be performed continuously or using a designated 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 recovery time objective (RTO) and recovery point objective (RPO) goals.

Immutability involves the ability to lock data after it has been created, thereby preventing it from alteration or deletion.

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

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.). Additional technologies, such as file and object versioning or journaling (with or without replication of versioned data copies), database log-shipping and other, can also be considered as forms of continuous data protection.

Point-in-time copies, are usually created by the immediate element storing the source data (e.g., storage array, filesystem, database), and are designated for fast recovery and a wide variety of other uses, such as cloning production data for testing purposes. Point-in-time copies can also be taken from remote replicas (and thus can be referred to as "remote Point-in-time copies").

A snapshot is a storage-efficient form of point-in-time copy, that stores only the individual portions of the data changed from a given point in time in reference to the source data. This often means that if the source data is unavailable, the snapshots will not be usable as well.

2.11.6 Data Reduction

Data reduction is the process of reducing the amount of data stored and/or transmitted in an effort to reduce costs and improve efficiency. The two common approaches to data reduction are data deduplication and compression. These two 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 less storage than the unencoded representation [10]. Data compression can be used in many places, but is most commonly used as part of tape backups and during remote data replication in network gateways to reduce the bandwidth requirements for DR and business continuity (BC) operations. Interoperability is a key requirement for data compression where the compression and subsequent decompression may be performed by different entities.

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 [20]. This can take place at the storage device level, the transmission stage, or the filesystem level.

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

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 threats can manifest).

3.1 Threats

A threat is the potential cause of an unwanted incident, which can result in harm to a system or organization. The following sections provides a brief overview of storage infrastructure-related threats.

3.1.1 Credential Theft or Compromise

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. 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, and to protect the sessions in which data at rest or in transit is managed and controlled. Encryption key-generation 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 or bugs, and backdoors [24]. It is not only important to use strong encryption, but to also properly secure the encryption keys. Changing encryption keys proactively can be part of a strategy to protect against compromised or insufficiently strong keys. When it comes to key generation, key strength, quality, and entropy play important roles, and keys should not be reused. Some attacks are based on exploiting random number generator weaknesses, including predictability or limited entropy, and the ability to disrupt it, so that it issues the same random number for key generation twice [25].

3.1.3 Infection of Malware and Ransomware

Malware is the general term for any program that is designed to damage, disrupt, or provide means to compromise a device. Malware compromises a system, slowing down its basic

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

functions and breaching its security. It can be used to steal data, control a device or system, and harvest the system's resources for illegal activities. Malware can infect a system in several ways; 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, or alteration, compromise of future backups, and more. In general, malware will use Operating System (OS) and other software vulnerabilities to install itself and perform various actions. The more common a software package, tool, or OS distribution is, the more likely it is that malware kits have already been published. For this reason, it would be easier to attack the storage management system than the storage device itself. This is not to say that storage devices are not a target – and indeed, the threat of using compromised images, firmware or microcode is of growing concern.

Ransomware is a form of malware that encrypts the stored data, 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 that are ***intentionally*** created by vendors or individual contributors (and, in rare occasions, by nation states or malicious actors) for reasons often considered legitimate by the author (e.g., to improve support, debugging, national security, etc.). Given their potential for damage, 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 software quality assurance (QA) or testing that, if exploited, 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 to close the gap. The timely deployment of such fixes is of paramount importance. A large portion of successful attacks are based on vulnerabilities for which a fix has already been published.

3.1.5 Privilege Escalation

Privilege escalation is the act of exploiting a software vulnerability, design or deployment flaw, or configuration mistake 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)

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

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, 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, or client), the networking devices, or management systems.

3.1.6 Human Error and Deliberate Misconfiguration

Even with the existence of security controls, users may end up performing technically supported storage configuration changes that still present 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 omissions 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. These include:

- Typos.
- Lack of knowledge or an 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 flaws in scripts and manifests.
    - Indirect, such as unrealized software dependencies.

3.1.7 Physical Theft of Storage Media

All data is ultimately stored in one or more copies on physical media, which is susceptible to theft. Media, whether online or offline, can be removed from its designated (stationary) location or while being physically transported between locations (for example, backup media being transported for archiving, or storage equipment being shipped as part of a datacenter relocation project). Theft could be opportunistic, where the content of the stolen media is not known in advance, or targeted – where specific data is of interest to the perpetrators, and they have the means to determine which media to seize.

3.1.8 Network Eavesdropping

Data could be intercepted while being transmitted. Transmission can span many components: network cards (wired or wireless), cables (carrying electricity or light), repeaters, switches, routers, etc. Any of these components can be compromised, and many forms of compromise are difficult or impossible to detect with state-of-the-art tools and methodologies.

While data encryption along the transmission path has an important role in limiting the usability of intercepted data, it might still be compromised (for example, by managing to intercept data at

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

a point prior to its encryption or after its decryption, or by managing to gather enough data to break the encryption).

Certain transmission compromises can extend beyond just interception of the data (also referred to as passive eavesdropping), and involve injection, removal, or alteration of transmitted data, metadata, or control traffic.

3.1.9 Insecure Images, Software and Firmware

Adversaries may attempt to interfere with a storage device's software distribution, update, or installation process in order to introduce incorrect, outdated, or maliciously modified code (e.g., binaries, images, firmware, drivers, etc.). Affected storage components can include disk drives, tape drives and libraries, network cards and controllers (e.g., HBAs, network interface cards - or NICs, FCoE adapters, etc.), switches and other network equipment, storage enclosures and arrays, Storage OS, Client OS storage components, etc. Software update processes can rely on complex delivery chains: an issuer (e.g., vendor, third party, open source community), delivery methods (e.g., transmission or download, shipping of installation media, file copy by vendor employee), local copies kept by an individual organization (e.g., proxies, internal file-servers), and other. Each link in the chain could be targeted to introduce tampered software. Issuers, for example, could be infiltrated to infect source-code libraries, to obtain access to signing software or equipment, to publish altered signed binaries on their download sites or update servers, etc. A variety of other strategies could be devised to compromise other links in the chain.

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 and Data Exposure

A data breach is an incident that involves sensitive, protected information being copied, transmitted, viewed, deliberately exposed to the public, or used by individuals or entities unauthorized to do so. Exposed information may include banking and credit card numbers, personal information (including health-related, home address, phone numbers, dates of birth), session tokens, passwords, customer data, company trade secrets, matters of national security, or any other proprietary or sensitive information.

Data breaches can be exercised by an external source, such as a hacker or cybercriminal, or by an internal one, 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

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

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 or confidential data, resulting in irreparable damage to the reputation and operational health of the organization.

While certain data breaches involve a high degree of sophistication, many are made possible due to simple, inadvertent security settings of data assets. Among the many possible root-causes are weak (or lack of) encryption at rest or in transit, software flaws, loss of custody of removable media, media theft, incorrect or too relaxed access limitations, improper or incomplete data sanitization implementation (including deleted objects, retired or repurposed media, etc.), sending or transmitting information to the wrong recipient, or data being uploaded to an incorrect location, or in an incorrect manner (e.g., uploading protected data to a public data store).

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

3.2.2 Unauthorized Data Alteration and Addition

Data alteration refers to 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 affect future application transactions, or impact other uses of the data. Data alteration and addition can originate from either an external or internal source in either a covert or easily identifiable manner. In certain cases, this type of risk is realized using the "salami attack" method, in which the attacker steals small bits of data or funds over a long period of time from a large number of transactions. The impacts of data alteration and addition can range from the loss of funds to permanent damage to reputation and trust.

3.2.3 Data Corruption

Data corruption refers to damage to, or 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, the object 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, a user may be unable to open it, or it might open with some or all of the data rendered unintelligible. Some types of malware may intentionally corrupt or destroy files as part of their payloads, usually by overwriting them with inoperative or garbage code, or otherwise securely erasing their content. While some forms of data corruption will result in storage device, OS, or software errors upon access, others may be designed to affect data without issuing errors.

3.2.4 Compromising Backups

The backup, or retention and archiving of copies (including replicas and snapshots) 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 way, and accessible in a manner that supports timely restoration. Since these prerequisites closely depend on each other, 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, old or new, will be

unrecoverable. An attacker, therefore, has a high motivation to target not only a "primary" data asset but also its backup and 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, firmware, 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 regain control, or inflict more damage.

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

3.2.5 Malicious 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 recovered using a key held only by an attacker. This type of risk is commonly used in ransomware attacks - a form of malware that encrypts the victim's data and demands a ransom to restore access to the data. Although it originally targeted data or 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 be inflicted by an internal one. These attacks are often meant to be identified and are 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.

3.2.6 Data Unavailability 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 modify or remove 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 altered or deleted), it may cause lengthy 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 storage devices, management interfaces, clients, or network, 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.

3.2.7 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 or 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 rogue log servers, etc.).

3.2.8 Compromising Storage OS or Binaries, Firmware and Images

Compromise of storage software, including storage device OS, firmware, images, etc., could produce a wide range of undesirable outcomes: providing attackers with means for remote access, to read, copy, alter or destroy data and data copies, to change security settings, to expose data, to alter the behavior of the storage infrastructure and more. Storage behavior alteration is of particular concern as it can be used to introduce a variety of latent, hard to detect attack capabilities, including presenting incorrect data to storage clients (even if stored data is intact), providing incorrect status (e.g., falsely reporting on existence or state of snapshots and security settings), bypassing environmental safeguards (e.g., thermal, power consumption, speed limits), stripping or altering encryption, and more.

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.

Threat – Potential for Unwanted Incidents	**Risk - Occurrence Outcomes**
Credential theft or compromise	*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
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

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

Threat – Potential for Unwanted Incidents	Risk - Occurrence Outcomes
Backdoors and unpatched vulnerabilities	Depends on the nature of the vulnerability, but in many cases, all risk outcomes for credential theft or compromise apply to this threat
Privilege escalation	Depending on whether user credentials or administrator credentials are compromised, all risk outcomes for credential theft of compromise 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
Physical theft of storage media	Depending on scale of theft, data breach and data exposure, data corruption, compromising existing backups, data unavailability and denial of service
Insecure images, software and firmware	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 into vehicles transporting physical objects (e.g., hosts, storage arrays, hard drives, tapes). Such intrusion is performed to access, steal, damage, or impact availability of data.

The physical intrusion can involve "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

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

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. In addition, storage OS and access configuration could be modified, and physical tapping and / or transmission devices could be installed to support later, remote access. An intruder can attach or remove media, connect to a storage system port, a management port, a peripheral port that may be used for firmware updates, or to a management terminal. Even if the central storage systems are well-protected by physical restrictions, an attacker can physically target additional components of the storage infrastructure that may be less protected, such as edge switches, exposed network ports, and management workstations. 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 and passwords, or infects the system with malware).

3.3.2 Access to Storage Operating System

This attack surface involves intrusion into a storage device by exploiting operating system vulnerabilities. The term "storage OS" refers to all the operating systems that are related to the storage infrastructure, including storage arrays, switches, data protection appliances, and storage virtualization appliances. In many cases, the operating systems running these devices are based on a proprietary version of the Linux/Unix operating system that is generally more closed or secure than general-purpose OS distributions. However, all operating systems include security vulnerabilities and therefore should be regularly kept current 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 Secure Shell (SSH), 'rshell,' 'telnet,' etc.), through remote login using TCP/IP or by using an OS vulnerability. In HCI the attack surface could be significantly larger, as multiple host OS instances take part, some of which could run arbitrary compute workloads.

3.3.3 Access to Management Hosts

Most storage components are managed or configured through management hosts, which usually run on 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, or 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.

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

3.3.4 Management APIs, Management Software, and In-Band Management

Storage infrastructure components expose management software User Interfaces (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. Simple Object Access Protocol (SOAP) or Representational State Transfer (REST) API) and, in parallel, management software that is installed on a management host. Finally, storage systems often interact with external network services for key management, authentication and authorization, etc. 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., Fibre-Channel paths) - the same connection plane used to provide the storage service (see 4.2, "Data Protection" below for additional discussion of planes). By doing so, it opens yet another attack surface, which can be exploited by impersonating 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. Additionally, 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 host adapters, 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, "man in the middle" (MITM), can apply to a wide range of transport protocols and media types, including Fibre-Channel, iSCSI, NVMeoF, management, etc. The purpose of the MITM is to sniff data, alter it, or bypass encryption and authentication mechanisms.

3.3.7 Compute Environment of Key Individuals

Certain key users are granted the privileges and means to perform remote storage infrastructure management. For example, the storage admin may use one or more workstations to remotely connect to the storage's management host. The compute environment of such key individuals

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

(e.g., laptop, desktops, home network, home computers) can be exploited to gain access to and compromise the storage infrastructure. For example, an attacker can install malware on such a 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 Electrical Network and other Utilities

Since storage infrastructure is connected to the electricity grid, the electrical 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 on 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 converts them to a binary form via demodulation and decoding [33].

Other utilities, security, and environmental control systems (e.g., heating, ventilation and air conditioning – HVAC, fire extinguishing systems, uninterruptible power supply – UPS, sensor systems, surveillance systems, etc.) can also be used to impact data storage, ranging from risk to systems (e.g., overheating, flooding, explosion), through risk of data leak (e.g., tapping into the video surveillance system to intercept password entry, or content of screens, panels, indicator lights, recording auditory signals), to hijacking internal transmission capabilities of environmental systems (e.g., WiFi, Bluetooth) in attempts to circumvent air-gaps and network controls.

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

4 Security Guidelines for Storage Deployments

The following sections 4.1 through 4.12 provide security recommendations and guidelines for storage infrastructure. Each section is dedicated to a specific aspect of storage security and contains a set of recommendations and guidelines, whose naming convention and numbering scheme is designed to allow uniquely identifying each one. Primary unique identifiers take the form 'xx-SS-Ry', where 'xx' is a two-letter combination that relates to the section's heading, 'SS' stands for "Storage Security", and 'y' is a sequential numerical identifier. For example, in section 4.1, "Physical Storage Security", primary identifiers are labeled PS-SS-R1, PS-SS-R2 etc. Secondary alphabetical identifiers ('(a)', '(b)', …) are sometimes used within an individual guideline. This allows addressing an entire compound guideline by its primary identifier, for example 'PS-SS-R1' (that addresses media security), and to also address specific portions thereof, for example 'PS-SS-R1.a' (adherence to NIST SP 800-53, section 3.10), 'PS-SS-R1.b' (supply chain protection), etc.

Finally, bulleted lists are sometimes used as part of a specific guideline when there is no need to address each list item individually.

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 800-53 [28], Rev5 [NIST SP 800-171 [34]. Additional valuable discussion regarding media disposal and destruction is available in ISO 27040 ([10]), and NIST SP 800-88 Rev. 1 [29].

This section provides focused guidance on physical security aspects that are unique to storage infrastructure or on aspects that are less emphasized in other publications.

PS-SS-R1 – Media security measures:

(a) Follow general recommendations, NIST SP 800-53, Rev5, Section 3.10 (including policies, access, marking, storage, transport, sanitization, cryptography, removable media, confidentiality, and disposal)
(b) Lifecycle management should include purchasing media with an adequate supply chain protection.
(c) For sensitive data, physical media for backup should be stored in a location sufficiently distant from that in which the primary data is stored.

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

(d) For sensitive data, a comprehensive inventory of storage media (cataloging) should 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; leakage of passwords, certificates, or encryption keys).
- Mitigation/contingency steps or procedures to employ (e.g., changing passwords, re-issuing keys, re-encrypting data, and notifying relevant stake holders).
- Dependencies between the data and other application.

(e) Consider using advanced tracking controls on sensitive removable media, such as Radio-Frequency Identification (RFID) tags, Global Positioning System (GPS) tracking devices and tamper protection.

(f) For extremely sensitive information consider the use of self-activated and/or remotely controlled self-destruction mechanism. When implemented, carefully consider how to protect such capabilities as they present an attack vector an adversary could use to trigger destruction of the devices, or to harm nearby equipment, data and personnel.

PS-SS-R2 – Protect all sensitive administrative equipment: Sensitive workstations, which can be used to obtain administrative access to storage infrastructure, should be managed using organization-approved security controls for access, surveillance, and auditing, including physical security. The security measures taken to protect the management workstations should be at least as strict as those used to protect the data it manages, and the systems that use that data. This includes workstations located outside of the facility storing the data as well as work-from-home environments, when used.

PS-SS-R3 – Data sanitization approach should cover storage infrastructure in detail, including non-obvious storage components: Certain elements capable of storing sensitive information, beyond the media itself, are sometimes overlooked when disposing of storage equipment, including non-volatile memory and cache objects (often found in storage arrays, SAN switches, routers, etc.), firmware/BIOS settings, and HBA-level settings (which can contain organizationally identifiable addresses, such as IP and SAN Fabric World Wide Names [WWNs], masking configuration, passwords). Validate that all those elements are considered as part of the organization data sanitization policy, including policy definition, operations, and audit.

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,
- Archiving
- Replication technologies,

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

- 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. Additional adjacent requirements are included in sections 4.7, “Isolation”, and 4.8, “Restoration Assurance”.

In discussing various aspects of storage management, access, usage and protection, it is useful to identify and differentiate between different **data planes**. A data plane is a loosely grouped set of access methods, protocols, communication, access control and authorization, and operations that are applicable to a data object (or a set of related data objects). Note that the provided list is by no means binding, some of the mentioned aspects could be omitted, and other – such as network filtering rules, the roles and authorization mechanisms used to allow I/O, etc. can be added. For example, a block-device could be associated with one or more of the following planes:

- “Data consumption plane” - defined in this example as the set of access protocols used to perform I/O operations, the I/O operations themselves, and the physical and logical network connections used to perform such operations.
- “Data management plane” - includes protocols, operations and network access used to create and destroy the device, to configure its attributes, to map it to hosts, etc.
- “Data protection plane” - includes protocols, operations and network access used to replicate, snapshot, backup and archive the content of the device.

The chosen design and implementation can greatly influence the degree of separation between different planes. Separation is a function of multiple variables, including – but not limited to:

- Network Layer 2 separation – e.g., using different Virtual Local Area Networks (VLANs) will increase separation between planes, while using the same ones will reduce it.
- Network logical separation – e.g., using separate IP subnets, will increase separation, while using the same subnet will reduce it.
- Filtering and ACLs – e.g., adding ACLs to prevent management operations on data consumption planes will increase its separation from the data management plane.
- Authorization – e.g., using different users and roles for each plane and restricting each role permissions only to rights associated with that plane.

As a general rule, increasing granularity and separation of data planes can positively impact the security of data assets at a potential cost of increased management and administration overhead, and of additional network-related requirements (additional switch ports, I/O controllers, etc.). In the above example, poor separation between data consumption and data protection planes might allow an adversary that gains control over a host that is mapped to the block device not only to corrupt the content of the device, but also to damage copies and backups. Poor separation between data consumption and data management planes might allow the adversary to impact other storage devices.

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

4.2.1 Data Backup and Recovery, and Archiving

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

(a) Tiering, frequency, and number of copies specification, to meet organization recovery goals. Specification may include more than one tier (e.g., continuous, hourly, daily, weekly, etc.), and should include for each tier:
    a. Frequency and retention - e.g., 48 hourly snapshots, 30 daily backups
    b. Type – e.g., full, incremental, continuous (such as file versioning, journal or log shipping and archiving), replication, point-in-time copies, etc.
(b) Types of media to be used.
(c) Encryption requirements – for data at rest (in particular, encryption methods applied to backup data should be at least as secure as encryption of the protected data), and for data in transit. Encryption key retention and key rotation should also be considered.
(d) Other protection requirements, such as digital signing, archiving, location, facility security (including fire, explosion, and magnetic interference protection), immutability and locking, minimum number of copies per backup set, and the geographic distribution of such copies.
(e) Reference to applicable regulatory frameworks with appropriate controls.
(f) Comprehensive lifecycle management that includes tracking of data copies and backups against protection and retention policies, including affirmative deletion of no longer needed ones.
(g) Restore procedures.

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

(a) Cover all data assets of the enterprise irrespective of where they reside (i.e., on-premises or in the cloud). It is acceptable to refrain from protecting data assets with no significance to the organization, or that can be re-created from other protected data sources within the required RTO, but such omissions should be documented.
(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 should 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 - so that implementation would be based on technology stacks with appropriate characteristics (e.g., disk-based, point-in-time copies such as snapshots or clones, vs. 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. For applications with strict restoration speed requirements, an end-to-end test restore should be performed (e.g., a complete recovery of the dataset to a sandbox recovery environment, simulating real-life restoration scenarios).
(c) An up-to-date recovery catalog should be kept to track each copy (including backup, replication, Point-in-Time copies, etc.) that records which anti-malware tools it has been

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

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. See additional cataloging requirements in **CM-SS-R2**.

(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, or no longer supported media by copying to new one, etc.).

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. This should not be interpreted as excluding redundancy mechanisms that protect against a single-point-of-failure.

4.2.2 Replication and Mirroring

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.

AC-SS-R6 – Eliminate unnecessary replication trust between storage devices: When arrays do not have shared replicated volumes, disable the replication trust relationship between them. When arrays do have shared replicated volumes, their privileges with respect to one another should be limited to the volumes they share.

DP-SS-R7: The confidentiality and integrity of data in transit during replication and mirroring should be protected using encryption. This recommendation can be relaxed when appropriate mitigating controls exist (e.g., mirroring over short distance within the same enclosure, or server room, etc.). See also **EN-SS-R8**.

DP-SS-R8: Automated I/O suspension should be enabled for synchronous replication in all circumstances where the replica cannot be allowed to fall behind the primary data. Enabling this feature implies that the primary storage device will disallow any write operations on the data it stores if its synchronization with the secondary storage server is lost, and it can only resume processing when synchronization is restored. For this reason, enforcing this feature comes at a cost of opening an additional attack vector; an adversary that is aware of or suspects that it is in use, could trigger primary storage denial of service by attacking the replication network paths. This tradeoff should be carefully weighed.

DP-SS-R9: Obsolete replicas should be removed to reduce the attack surface.

4.2.3 Point-in-Time Copies

The term "Point-in-time copies" should be expansively interpreted to cover a variety of storage-technology internal data copy mechanisms that can be used to create a copy of original data as it appeared at a past point-in-time. This include storage-arrays built-in copy-generation tools (often

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

referred to by vendors as “Snapshots”, “Clones”, etc.), filesystem-level copies, database-level copies, cloud storage pool copies, etc.

DP-SS-R10: 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 should 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-R11: Obsolete snapshots and clones should be removed to reduce the attack surface.

4.2.4 Continuous Data Protection

DP-SS-R12 – Security considerations for using continuous data protection: Other than the functional benefits (e.g., improved RPO, finer-grained retention), the use of continuous data protection techniques (e.g., CDP, versioning of source data or of replicas in cloud file- and object-storage, and transaction log shipping) can also assist in improving forensics for sensitive data. While potentially extremely time consuming, replaying to previous versions of data can help learn more about the attack profile, its timing, etc.

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 who manage the 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 contributing 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 and permissions for all user and administrative operations, creation of policies regarding all authenticators (e.g., shared secrets), manage certificate and key stores, manage encryption and key management, manage auditing and logging, 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.

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

- **Storage Auditor** – This operator-like role has "view" rights that allow for the verification of storage parameters and configurations as well as inspections of health or 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 users (including administrators) should have a unique identifier for their personal use only. Identifiers assigned to administrators should, at the minimum, meet the identity assurance level 3 (IAL 3) specified in sections 4.2 and 4.5 of NIST document SP800-63A [35]. The only exception is emergency-use account, whose secure usage is outlined in AC-SS-R16. This principle 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: In a large scale environment, a centralized authentication solution (such as Active Directory, Lightweight Directory Access Protocol [LDAP], single sign-on [SSO], organization approved cloud authentication services) should be deployed to enable the close monitoring and control of user access to storage resources and to ensure uniform enforcement of the organization's authentication policies. The use of built-in authentication and permissions management capabilities 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 state of art protocols such as Transport Layer Security (TLS) 1.2 or higher.

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 with authenticators, at the minimum, meeting the requirements specified in section 5.1.9 of NIST document SP800-63B [36]. This requirement should be made mandatory for access by users assigned to Security Administrator and Storage Administrator roles.

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., Simple Network Management Protocol (SNMP), Network Data Management Protocol (NDMP)) and accounts used by automation tools. The passwords should at the minimum meet the requirements for memorized secrets as outlined in section 5.1.1 of NIST document SP800-63B [36]. In addition, the following requirements should be met.

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

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

AC-SS-R8 – Password complexity: A good password should combine uppercase and lowercase letters, digits, and special characters. It 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. Expiration for administrative accounts should be set shorter than for user accounts.

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 Time to Live (TTL) or an equivalent control mechanism should be employed so 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 should be immediately changed.

4.3.3 Account Management Recommendations

AC-SS-R14 – Use of accounts not associated with system users: Accounts not associated with any system user (e.g., not in 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 locked out after a certain number of unsuccessful login attempts. Certain implementations of account locking include automatic reset (account unlock) after a certain period of time or a power cycle. Automatic reset should not be allowed on sensitive storage systems.

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). In addition, its usage should be allowed from a special physically protected location, and following well documented procedures that include appropriate approval of relevant stakeholders, and notification of use.

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

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, if the feature exists. When the feature to disable or eliminate does not exist or there is a justified reason to keep any of those accounts, the privileges assigned to this account should be kept to the minimum necessary.

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.

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

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] should be implemented for all access to storage resources (i.e., Security Administrator, Storage Administrator, Security Auditor, and Storage Auditor) – see section 4.2 for more details. Storage products that offer only one or two levels or privileges, should not be used for storing sensitive information, unless compensating controls are available to provide equivalent functionality of granular roles. For example, all management traffic could be routed through a management proxy host, or "command gateway" with privilege management tools installed to restrict commands available to each role.

AC-SS-R20 – Adherence to the principle of "Separation of Duty" during assignment of privileges to roles and assignment of roles to users: A critical aspect of storage security is to separate administrative control planes (refer to discussion in 4.2, "Data Protection" above). For example, if attackers gain control over a host or compromise a host admin role, they should not be able to 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 and these two roles should not be assigned to the same user.
(b) The privileges required for *data management* and *host administration* (e.g., tasks such as creating/deleting objects in the storage controller) should be assigned to different roles and these two roles should not be assigned to the same user.

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. In the context of the storage, these permissions pertain to access to storage-specific resources such as block-devices, files, objects, etc.

AC-SS-R22 – Secure Session Management: All sessions between the client and a storage infrastructure system should be managed based on the required authentication assurance level conforming to the requirements in section 7 of [63B] – including termination and automated logout.

AC-SS-R23 – Implement a "message of the day" and "login banner" notice: The "message of the day" or "login banner" notice should appear before every login to any storage infrastructure component or system via UI, Command Line Interface (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.

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. Encryption of data in transit (one of the mechanisms for access control) is covered in Section 4.9.
- Administrative access is discussed in Section 4.10.
- Data-related access control is discussed in this section and covers block device-related access control, implementation of zoning and access control specification for joining the fabric.

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

AC-SS-R24 – Block-device access control: The set of hosts that can access a set of SAN storage devices should be restricted through zoning (software or hardware) and masking to the minimum required access.

AC-SS-R25 – Block-device copy and replica access control: The set of hosts that can access a set of SAN-replicated block-devices, snapshots, and other types of point-in-time copies of such block-devices should be restricted through zoning and masking to the minimum required access. In many cases, a host granted access to a device should not be allowed to access a copy.

AC-SS-R26: The permission for default zone (which may be product-specific) should always be configured as "deny all."

AC-SS-R27: 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, which should be separated to the maximum possible extent, by:

(a) 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-R28: When software zoning is implemented, hosts should only be allowed to 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.

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

AC-SS-R29 Controlling devices that can join fabric: The policy specification feature in SAN that enables the creation of an allowable list of switches, arrays, and hosts that can join the fabric should be leveraged where applicable.

4.3.6 File and Object Access Recommendations

AC-SS-R30 – Restricting access to object storage data of all types (e.g., Files, Objects) to the minimum possible: Follow the "least privilege" principle, including:

(a) Access to object storage data through any protocol (e.g., CIFS, SMB, NFS, and public cloud object storage protocols) 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, Virtual Private Cloud (VPC), VPC endpoints, etc.) should also be used.
(c) Access should only be granted to centrally managed users and roles, such as those found in enterprise directories or approved commercial services, and not to local users of the specific system.
(d) The default access to any share should be set to “deny all” or equivalent.
(e) The default shares should be disabled or removed. If there is a specific purpose for using them, the access rights should be restricted to the minimum required.
(f) The access rights (e.g., read, write, execute, modify, delete, view ACLs, change ACLs), should be individually assigned based on “Need to Know” requirements.
(g) If the feature to define object storage ACLs is available, it should be leveraged in addition to the use of native OS user, group, or admin permission models.
(h) If a policy definition feature to define file level access patterns is available, it should be leveraged and the feature to detect violations to the pattern, which sends notifications, should also be implemented.

AC-SS-R31 Users that do need to authenticate should be disabled (e.g., Anonymous, null, guest, or “public access”). An exception may be provided to allow organization-critical functions such as network discovery, but in such cases these class of users should be mapped to the “nobody” user group and not to "ID 0".

AC-SS-R32 – Regular audits of all the security settings mentioned above for storage data of all types (e.g., Files, Objects) should be performed to ensure that there are no drifts. Audit results should be documented.

AC-SS-R33 – Scan files containing sensitive information with anti-malware tools on-access: Every time a file with sensitive information is accessed, it should first be scanned with organization approved anti-malware tools to ensure that it has not been compromised.

AC-SS-R34 – Granular permission assignment: For file and object sharing systems (e.g., NFS, CIFS, cloud object stores), permissions should be granted at a finer level of granularity rather than a coarser one (e.g., file or object over folder, and label over share or bucket).

AC-SS-R35 – Secure NFS by restricting root access: This includes the use of the “nosuid” option, and avoidance of use of “no_root_squash” to prevent programs from being executed as a

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

root user on the client, and modification of shared files by remote root users. In general, NFS clients should not be allowed to run "suid" and "sgid" programs on exported file systems.

AC-SS-R36: 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-R37 – Export of administrative file systems should not be allowed: This includes the '/' filesystems and restricted OS or storage array system folders.

AC-SS-R38 – 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-R39 – Use object protection to prevent unauthorized deletion – for sensitive information, when supported, use advanced controls to prevent unauthorized object deletion (such as requiring Multifactor Authentication for object deletion, of locking objects against deletion).

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** – e.g., resetting of user passwords, account creation/deletion, modification of privileges, role changes, group membership changes, privileged actions, creation of/changes to configuration. These events are always of interest.
- **Security-Related Events** – e.g., changes to user profiles and security configuration, failed/blocked attempts to storage, blocked logins. These events are most often of interest, though some of them may overlap with management events.
- **Data Access Events** – access information is of interest for incident response in monitoring sensitive information (e.g., determining what an adversary may have touched).

Deficiencies in security logging and analysis allow attackers to hide their location, malicious software, and activities on victims' 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 actual 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

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

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.

AL-SS-R1 – Audit logging should be enabled on all storage infrastructure components using reliable delivery and secure protocols.

AL-SS-R2 – Reliable, external time synchronization: A Network Time Protocol (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. The following are recommendations regarding the deployment and integration of NTP with storage-related devices:

(a) The NTP service should be enabled on all devices (including log servers and storage infrastructure).
(b) All devices should be configured to synchronize time with a time source server, such as an NTP server.
(c) Time synchronization validity should be monitored on all devices (e.g., that the service remains activated, that organization approved time servers are configured in each device), and alerts for detected anomalies should be handled at a high priority.
(d) Redundancy for time source servers should be provided by deploying at least three synchronized time sources that are geographically distributed.
(e) Certificate-based authentication should be used to authenticate time source server.
(f) Access control options, such as “ntpd” access restrictions, should be leveraged to restrict access to the time source servers.

AL-SS-R3 – Collect logs in a centralized fashion: By writing logs to central log servers (e.g., syslog server, cloud logging services), the risk of those logs being lost or altered is lowered since they are more secure within the internal network. The following are recommendations regarding the deployment and integration of central logging with storage-related devices:

(a) Organizational logging standards for storage devices should be defined, to specify required logging level. *Management Events* and *Security-Related Events* (as defined above) should be logged for all types of data. For extremely sensitive data, *Data Access Events* (as defined above) should also be logged. See more specific recommendation in AL-SS-R4.
(b) All devices should be configured to transmit log event data to organization approved central log servers, according to the applicable organizational logging standard.
(c) Central logging configuration validity should be monitored on all devices (e.g., that the logging service remains activated, that logging level configuration matches the organization standards, and that organization approved log servers are configured in each device), and alerts for detected anomalies should be handled at a high priority.
(d) Multiple syslog servers should be deployed to enable continuous logging and to prevent a single-point-of-failure.

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

(e) At least one off-site copy for each log should be maintained.
(f) To prevent the loss of entries in the event that the logging process is stopped and restarted before all entries are written, logging should be configured to be written to disk in real time with no buffers in place and sent over reliable protocols.

AL-SS-R4 – Logging level: The following events related to storage protection of all storage-related objects, sites, and accounts should be included (but not limited to) in the storage audit logging:

(a) Read-only API calls in sensitive environments.
(b) All denied access attempts to services, ports, files, objects, or devices.
(c) Cryptographic key management operations spanning the entire key lifecycle operations (particularly for encryption keys), such as key generation, key deletion, certificate management, etc., especially for events such as key shredding.

AL-SS-R5 – The following measures should be adopted for audit log retention and protection:

(a) Retention of log data for a sufficiently long period of time, as it often takes a while to notice that a compromise has occurred or is occurring.
(b) Allocation of sufficient storage space and proactive monitoring of 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) Archived log data should be protected from tampering (e.g., using WORM or immutable storage, object locking, Multi Factor Authentication (MFA) approval for delete). If supported, the central log servers should also use such storage options.
(d) Restricting access to log data and servers using designated roles and accounts.
(e) Enabling encryption since access to log data can provide attackers with valuable insight into assets and possible attack vectors.

AL-SS-R6 – SIEM integration: If supported, storage infrastructure logs should be integrated with Security Information and Event Management (SIEM) software for potential threat detection.

4.5 Preparation for Data Incident Response and Cyber Recovery

Incident response planning is an important part of Cybersecurity. Comprehensive discussion of the role of Incident response in Cybersecurity program management, and guidance to building a framework for cybersecurity improvement can be found in *NIST Framework for Improving Critical Infrastructure Cybersecurity* [40]. 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 and managing a response plan, testing, and periodical process review and refresh.

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

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

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

(a) Compromise of an entire storage 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, block device)
(d) Compromise of an FC SAN fabric (including individual switches and SAN services)

IR-SS-R2 – Recovery assets immutability during incident management: In conjunction with the recommendations provided in Section 4.7 below regarding the protection of cyber recovery copies, those copies should 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.

4.6.1 FC SAN and NVMEoF

NC-SS-R1 – Host and switch authentication: Every host and storage switch should have a unique identity and should be authenticated before joining the network (e.g., FC-SP-2 AUTH-A).

NC-SS-R2 –The use of an approved PKI mechanism: Use an organization approved and certified centralized PKI system for the management of switch certificates (e.g., Fibre-Channel Certificate Authentication Protocol or FCAP) rather than the devices' self-signed certificates.

NC-SS-R3 – A blended approach to zoning: Implement a zoning approach that blends different types of zoning mechanisms. 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 can be used to

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

control whether the host may interact with any storage device. At the next level, OS capabilities can be used to control which devices the host tries to mount as a storage volume. Finally, the host centralized management software for volume management (e.g., Logival Volume Manager or LVM), clustering tools, and the file system can be utilized to control device access by applications.

(b) In switch-based zoning, the switches (especially the FC switches) have the capability to specify which devices on which ports can access other devices or ports. Port-based zoning uses hardware to enforce zoning and is therefore also called "hard zoning." In other words, switches should support zone control at the port WWN's level 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 block devices 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 create multiple zones dedicated to a particular purpose, such as testing, dynamic reconfiguration, backup, and maintenance.

NC-SS-R4 – Masking Recommendation: Masking refers to making a block device visible or invisible to hosts. **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 backup of switch configuration data: create a backup of the swtches' configuration data, including zone configuration file. The backup 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) Unnecessary configuration management permissions and services, such as 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 may be 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 in the FC switch, which were used for connection. Port World-Wide Name (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 across 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.

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

- **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 Storage Networking

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 layer 3 of the network stack). Sensitive environments should be separated to the maximum possible extent, based on:

(a) Type of traffic: data access protocols vs. management vs. replication vs. backup vs. host and application networking
(b) In sensitive environments, further separate the management traffic of different solutions, vendors, and technologies. For example, if two or more storage solutions are in use (e.g., different array technologies, Server-Based SAN products, switch technologies, storage virtualization, or any combination thereof), and each one has a separate set of management tools, management traffic for each environment should be separated from the others.
(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.

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

NC-SS-R13 – Enable device IP access control: With respect to IP network accessibility, storage device security features that regulate IPs, ports and protocols should be turned on and configured on all storage devices, where applicable. This includes but is not limited to built-in firewall rules, IP filtering, and access lists in order to:

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

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

(a) Between hosts or applications and the storage objects they use, and
(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 non-public storage objects, particularly from the internet.

NC-SS-R16 – Public access: For storage objects that require public access, sufficient controls should be implemented, including:

(a) Minimizing access.
(b) Using physically and logically separate storage subnets and, preferably, separate storage devices and pools from those used for non-public storage objects.
(c) Considering protection from denial-of-service attacks.
(d) Cached copies (e.g., using content delivery network (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).
(f) Any additional applicable security controls (e.g., encryption, authentication).

NC-SS-R17 – Control of IP addresses used for SNMP: When configuring SNMP, all traffic should be directed to valid organization-internal IP addresses as destinations. The validity of the configuration should be periodically reviewed.

NC-SS-R18 – Consider the use of isolated non-routable VLAN for server-based SAN: For more information about server-based SAN deployments, see section 2.9 above. To protect the data storage environment and mitigate security concerns, non-routable VLAN for server-based SAN should be used.

4.6.3 Protocols

NC-SS-R19 – Disable insecure 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, these protocols should be disabled on both the client side and the server side.

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

NC-SS-R20 – SNMP security:

(a) If SNMP if not in use, it should be disabled.
(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) SNMP authentication and encryption (privacy) features should be enforced.
(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 to devices through SNMP.
(h) Validation that SNMP traps are sent to authorized, intended managers should be periodically performed.
(i) Refer to Department of Homeland Security Cybersecurity & Infrastructure Security (DHS CISA) TA17-156A [39] for additional guidance.

NC-SS-R21 – The authenticity of directory, domain, and similar services (e.g., AD, Domain Name System (DNS), LDAP): The service configurations in all storage elements (e.g., devices, switches, management workstations, management software) should be actively and periodically reviewed to make sure that the approved ones are used, and for remediating any discrepancies.

NC-SS-R22 – 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-R23 - FCoE Initialization Protocol **(FIP) snooping filters should be enabled on FCoE VLANs to prevent unauthorized access to data**: FCoE Initialization Protocol (FIP) snooping is a security mechanism that is designed to prevent unauthorized access and data transmission to an FC network. It works by filtering traffic to permit only servers that have logged in to the FC network to access the network. FCoE transit switches connect FC initiators (servers) on the Ethernet network to FCoE forwarders (FCFs) at the FC SAN edge, enabling FIP spoofing on relevant VLANs.

NC-SS-R24 – Limit iSCSI ports: Hosts on the iSCSI network should be prevented from accessing any TCP ports other than those designated for iSCSI on that network.

NC-SS-R25 – Use iSCSI authentication: Use one of the supported methods to authenticate iSCSI initiators upon opening a session (e.g., Challenge-Handshake Authentication Protocol (CHAP), Server Routing Protocol (SRP), Kerberos, Simple Public-Key GSS-API Mechanism (SPKM)1/2). When using CHAP, prefer using two-way authentication over one-way

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

authentication. Note that the use of authentication does not provide encryption or integrity protection to the channel.

NC-SS-R26 – Use of Network Data Management Protocol (NDMP) security features: NDMP provides means for direct transport between storage arrays and backup devices. When used the following security features should be configured, 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-R27 – Use TLS in LDAP: Use TLS to secure LDAP connections when setting up Active Directory options for storage systems.

NC-SS-R28 – Additional protocols: When additional protocols such as SymAPI, Storage Management Initiative Specification (SMI-S), Global Name Server (GNS), and others are used, 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 and UDP ports.
(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 will not have the ability to recover. To improve resilience of backup copies, sufficient isolation should be guaranteed between data assets and their recovery copies.

In this context, organizations should distinguish between at least two separate data protection scenarios:

- **Non-malicious recovery –** requiring data copies that can 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, backups, and long-term archives. The "closer" the copy is to the production environment, the more likely it is to be mapped to compute systems for the purpose of testing and DR.
- **Cyber-attack recovery –** requiring data copes that are hardened, locked, and kept in isolation. The design should strive to achieve a state where these copies could not be impacted by *anything*, including scenarios wherein production volumes or other types of copies they are linked to have been compromised.

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

Keeping copies for Cyber-attack recovery that are completely separate from those used for non-malicious recovery, while recommended for critical and sensitive systems, is not mandatory. It is important, however, to make sure that the data protection scheme in use is suitable to fully support the two scenarios. If existing copy retention mechanisms (e.g., DR or BC copies, offsite archives) are intended to support Cyber-attack recovery, their configuration should be reviewed, and adjusted if necessary, to facilitate data isolation. This includes verification that at least a subset of those recovery copies and systems are inaccessible and independent from the production environment, and that a compromise of production would not allow adversaries to impact those copies.

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

IS-SS-R1 – Separation of 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 archive and backup systems should be separated from production data storage systems.

IS-SS-R2 – Separation of 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 any other system connected to production (including data protection mechanisms). It should not be possible to access such management systems with regular credentials (including production and regular backup). The system should be hosted on a dedicated environment that is only connected to an isolated network.

IS-SS-R3 – Access restriction to cyber-attack recovery system, and long-term archives and backups:

(a) For sensitive information, cyber-attack recovery copies and their systems should not be accessible to regular IT staff but, only to a single person (e.g. CISO), or 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 archives backups should be separate from those used to perform other storage administration duties (e.g., SAN management, storage allocation) and should include the use of separate user IDs, accounts, and credentials.

IS-SS-R4 – Off-site storage: Cyber-attack recovery copies should be stored off-site rather than where the production data is stored. This ensures that if the attackers have 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 copies.

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

IS-SS-R5 – 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 should be accounted for, and sufficient isolation between different types of copies should be maintained. In particular:

(a) Replicated disaster recovery copies should have no dependency on production baseline data.
(b) Cyber-attack recovery copies should have no dependency on production baseline data. Dependency on disaster recovery baseline data is allowed only if those copies are properly isolated from production baseline data, and meet the recommendation in IS-SS-R1, IS-SS-R2, and IS-SS-R3.
(c) Long-term archived data should have no dependency on production and disaster recovery baseline data.

IS-SS-R6 – Disable all unnecessary services and protocols: Unnecessary services and protocols should be disabled on cyber-attack recovery storage systems. In environments where Application Programming Interface (API) or Command Line Interface (CLI) are sufficient for management, it is recommended to also disable any interactive web interfaces.

IS-SS-R7 – Independence from hosts and applications:

(a) Cyber-attack recovery copies should not be mounted, exported, or mapped to a host or application, and they should be restored (pushed) onto an isolated staging (or air-gapped) environment rather than directly onto 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 archived or backed up copies should not be mounted, exported, or mapped directly 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 of sensitive data. Strict implementations of air-gapping should provide full physical and network-level separation. Certain storage technologies also introduce less strict isolation technologies, also referred to as “air-gapping” that enable shutting down data ports and opening them during a limited time for the periodic sync with the production system. It is important to weigh the effectiveness of each of those techniques based on the value of the data and the capabilities of the adversary. When strict implementations are chosen, attention should be paid to circumventing known vulnerabilities of air-gaped systems that include:

(a) Preventing visual, audible, and thermal signal transmission between air-gapped systems and other equipment (e.g., by maintaining sufficient distance, or using dampening and/or sufficient physical distance).
(b) Preventing any potential wireless transmission capabilities in air-gapped equipment.

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

(c) Disabling exposed data ports (e.g., USB, network).
(d) Using power conditioning, or separated power circuits.

IS-SS-R9 – Perform periodic isolation reviews: The above isolation recommendations should be checked at least once per year, as part of a periodic audit to ensure that there are no configuration gaps or drifts that may compromise the isolation of the cyber-attack recovery copy. Sensitive and high-value storage systems may require at least a quarterly audit, and after every major change - whichever comes first. Audit results should be documented.

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).

4.8 Restoration Assurance

To ensure successful recovery from BC or DR events, and a cyber-attack, it is not enough to simply have a process in place. Organizations should 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 these backups 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 recovery copies: All storage elements that contain components of critical data assets should be protected and backed up to support both DR and cyber-attack recovery. This includes storage volumes, critical file systems, databases, software images, certificates, encryption keys, startup files, catalog info, ACLs, virtualization settings, and configuration files.

RA-SS-R2 – Protect all dependent components: Dependent components, such as Active Directory and DNS, external key management systems should be protected to enable full recovery. If automated build-processes are used to configure storage, source-code repositories, build-environment, and build-procedures should be protected as well.

RA-SS-R3 – The availability of all relevant software and hardware components: All of the relevant software and hardware components (e.g., drivers, firmware) used to run the system should be backed up, protected, and available for a restore operation.

RA-SS-R4 – The elected backup and data copies technologies and media should match organizational RTO requirements: Recovery time objective (RTO), is a Key Performance Indicator (KPI) used to define the expected recovery speed. The ability to meet RTOs should be examined holistically, including all dependent and related components (e.g., restoration of 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.

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

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

RA-SS-R6 – Meeting RPO requirements: Set a recovery point objective (RPO) for each data asset, which is the amount of data that can be lost following a failure, expressed in time. The design and implementation of the backup and data copy technologies should support data recovery while meeting this objective.

RA-SS-R7 – Meeting organizational frequency and retention requirements: The data retention and copy frequency requirements for each data asset (refer to **DP-SS-R1** for more details) should be set. The design and implementation of the backup and data copy technologies should support the requirements.

RA-SS-R8 – Ensure that remote replicas backup and data copies are in good health: Periodically validate that the backup copies are in good health. This includes checking that there are no relevant logged errors and that the backup and data copy media is in healthy state. Frequency of validation should match sensitivity and value of protected data, but no less than once per year. Keeping the sample ratio around 1 – 1.5 orders of magnitude lower than the frequency of backup could serve as a solid foundation (e.g., hourly copies validated daily, daily copies by-weekly to monthly, etc.).

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

RA-SS-R10 – Document the DR plan, resources, mapping to production, flow, and test procedures: A disaster recovery plan for storage infrastructure 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 – Cyber hygiene of data copies: 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 tools include antivirus, anti-malware, vulnerability scanning, and security analytics.

RA-SS-R12 – Perform periodic audits: The above recommendations should be reviewed as part of a periodic audit to check completeness of copies, re-evaluate dependencies, software and hardware requirements, suitability of technology to support recovery speed, RPO, retention, health-checking, DR plan, and cyber-hygiene. Gaps should be identified, tracked, and remediated. Frequency of auditing should match sensitivity and value of protected data, but no less than once per year. Sensitive and high-value storage systems may require at least a quarterly audit, and after every major change - whichever comes first. Audit results should be documented.

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

4.9 Encryption

Encryption is the conversion of data from a readable form (i.e., plaintext) into an unreadable 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 unless the entire communication media is within a protected environment such as a data center.
- **Administrative access –** This includes connections through standard and proprietary protocols and APIs to configure or control storage elements, storage networking, and data.

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

Encryption relies on the availability and management of cryptographic keys. All communication parties should have access to the required keys, which need to be generated, distributed, and disposed of. Key management provides the necessary functionality, and is a fundamental requirement in most environments. Detailed recommendation for key management is provided in NIST SP 800-57 Part 1 [41].

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

EN-SS-R1 – TLS, hashing, and encryption: To support encrypted communication between storage clients and servers, Transport Layer Security (TLS) protocol should be used. To prevent the use of insecure or outdated configuration, the selection and configuration of TLS protocol implementations, including the selection of TLS version, and the choice of hashing and encryption algorithms, should be based on the following guidelines (or more current versions when published):

- Guidelines for the Selection, Configuration, and Use of Transport Layer Security (TLS) Implementations (NIST SP800-52 Rev2) [37]
- SNIA TLS Specification for Storage Systems Version 1.1 [38]

EN-SS-R2 – Cleartext protocols (e.g., HyperText Transfer Protocol **(HTTP), Telnet,** File Transfer Protocol **(FTP), or** Remote Shell **(RSH)) should not be used:** 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. In some implementations, HTTP is supported only for the purpose of redirection to HTTPS to deal with the case of mistyped URL. This should not be allowed in sensitive storage environments.

EN-SS-R3 – Encryption for storage management API sessions: Storage management APIs and CLIs are used for administrative access to storage systems. All API and CLI client sessions

should be encrypted leveraging features such as vendor configuration options within the management software or the API/CLI software component.

EN-SS-R4 – Encryption for administrative access sessions: Administrative sessions over HTTP should use TLS (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-R5 – Enable FIPS mode for FIPS-based environments: FIPS 140-3 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 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-R6 – At-rest encryption of sensitive data: 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:

- **Use of infrastructure encryption:** 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 compromises a host already mapped to the storage (or when the storage can be mapped using legitimate means to an unauthorized host).
    - **Privilege escalation attacks** - Administrators or attackers gaining elevated right can either disable encryption or decrypt the data.
- **Use of end-to-end encryption:** Data is encrypted at its source (e.g., an application, database, volume), presenting only ciphertext to the storage infrastructure and administrators. While it significantly increases security, application-level encryption comes at a cost, which can sometimes be considerable:
    - **Data-reduction mechanisms are impacted –** For example, compression and deduplication can become drastically less effective.
    - **Management is more complex**.
- **Use of dual independent layers of encryption:** should be considered for sensitive data storage where possible. This configuration improves resilience to key compromise; especially if diverse crypto services are used.
- **Data retention requirements should be taken into consideration:** If encrypted data is backed up or archived, relevant keys should be protected for a similar duration. Alternatively, the backed-up data should be re-keyed. In either case, data and encryption keys should not be kept together.

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

EN-SS-R7 – Data in transit encryption:

(a) **Block over Fibre-Channel**: Link encryption for FC, while defined in ANSI/ INCITS 545-2019, Information Technology - Fibre-Channel - Framing and Signaling - 5 (FC-FS-5), is currently not supported by most HBAs and storage vendors. For sensitive information use of end-to-end (host to storage) encryption.
(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, or authentication per packet. Similar to link encryption, while specifications for IP storage traffic encryptions exist, current technology does not natively support it. When using block-over IP protocols (e.g., iSCSI, FCIP, proprietary protocols), consider using IPsec tunneling for exposed network segments. In addition, for sensitive information use of end-to-end (host to storage) encryption.
(c) **File and object storage access:** Data encryption in-flight options should be enabled for backup systems and for remote replication whenever supported. For file access, sensitive data should be transmitted and encrypted using mechanisms such as SMB encryption, and available NFS encryption options such as those offered by cloud vendors, or NFS over TLS using tunneling (e.g., 'stunnel'). Ensure that objects are accessed through HTTPS with TLS.
(d) **Extended network communication beyond the boundaries of a physically protected domain:** 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-R8 – Communication between storage system components should be encrypted: Storage system component interaction should be reviewed, and available encryption options should be utilized. Encryption should be used to protect communication between storage nodes and managers, active-active storage nodes with witness devices, communication with policy servers and antivirus servers.

EN-SS-R9 – Requirements for Encryption key management: Follow general recommendations in NIST SP 800-57 Parts 1-3 for key management, in particular lifetime, maximum amount of data that can be protected by a key, key management infrastructure, re-keying, auditing, and key backup and recovery.

4.10 Administrative Access

Administrative access is required to control and manage almost all types of IT infrastructure. This section focuses on the of administrative access recommendations for 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.

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

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 – Limit network access to management ports of SAN switches: Network access to management ports of 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 the components that are authorized have them. Remove unnecessary ones, if found, and debrief.
(b) Remove unnecessary rights and capabilities from authorized devices.

AA-SS-R3 – Implement the principle of least-privilege: Limit the rights of users with administrative rights to the minimum required. This includes the minimum actions that the user can carry out and 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 the access rights of service accounts: Service accounts, such as ones used by monitoring tools, should be limited to read-only and metadata-only access.

AA-SS-R5 – Authenticate and 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-R6 – Favor API access control over CLI/shell access: API access if available should be used instead of CLI/shell access because of the latter's ability to access OS and file system including configuration files. If CLI is the only option available, then it should be used over secure protocols such as SSH as stated in **EN-SS-R4**.

AA-SS-R7 – Restrict management consoles OS privileges: Access to management consoles should only be provided through designated storage accounts, and not as an OS administrative account (see also **AC-SS-R20)**.

AA-SS-R8 – Management web user interface: The web service providing access to management consoles should be hardened to meet or exceed the minimum standards of other web-application servers in the organization.

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

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

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

AA-SS-R10 – Command device or gatekeeper configuration: Certain storage arrays allow in-band administrative control to hosts that have access to special block devices (e.g., devices referred to by certain vendors as “command device”, and ”gatekeeper”). Commands are transferred using I/O operations on those special devices. When used, the following security guidelines are recommended:

(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-R11 – 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, and to perform automated software updates. 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** – Secure protocols such as TLS/SSH/IPSEC using FIPS approved encryption algorithms should be used.
(d) **Limit access with an “allow list”** – Utilize an allow list that limits access by specific IPs and specific users.
(e) **Ensure remote access is 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 the obfuscation of 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, to ensure that it does not contain sensitive information, such as IP addresses, usernames, or the actual content of storage devices. Review that the connection is performed to valid vendor IP addresses.

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

(i) **Authorize each connection** – If possible, implement a mechanism that will ask permission before allowing each connection.
(j) **Restrict access to gateway system** - 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) **Disable software updates over vendor remote access links** - In sensitive environments, download and deployment of software components and updates (manual or automated) should not be allowed through remote vendor connection links.

AA-SS-R12 – Limit network access for management: In addition to separating management from other traffic (see Sections 4.6 and 4.7) in sensitive environments, further access control to management networks is recommended, by using mechanisms such as:

(a) **Virtual Private Network (VPN), IPsec, or one or more "jump servers"** - Using VPN, IPsec, 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 and tracing -** such as session recording.

AA-SS-R13 – 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.

The following controls should be applied to configuration files, scripts, and any other important management-related 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.

AA-SS-R14 - The use of an approved PKI mechanism for management access: Use an organization approved and certified centralized PKI system for the management of storage device management, and for storage management consoles rather than device or software self-signed certificates.

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

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. Comprehensive guidance for IT configuration management is provided in NIST Special Publication (SP) 800-53 [28]. The following paragraphs contain recommendations applicable to storage infrastructure.

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,
- Hosts used to monitor storage remote network connectivity status (e.g. Witness hosts),
- Hosts installed with storage management software or plugins,
- Data protection appliances,
- Backup clients and servers,
- Storage network switches,
- Storage adapters or "Host-Bus Adapters (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, and
- Disk drives and removable media.

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 catalog and access rights.
- Backup sets (on-premises, archived, virtualized in the cloud, on tapes, archive appliances, etc.).
- Users, groups, roles, and rights.

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

- 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).
(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 to storage configuration using logging, comparison of configuration storage assets to past states, or comparison to organization approved baselines.

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

CM-SS-R8 – Software updates and patches:

(a) **Ensure storage software release is updated:** 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) **Ensure important security updates and patches are installed:** There should be a process to proactively and frequently install important and urgent storage security fixes and patches.
(c) **Mitigation plan for missing patches** – storage components with a critical vulnerability, for which the vendor has not issued an update or patch, should be suspended from use, unless an appropriate mitigation plan can be defined.

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

CM-SS-R10 – Audit FC SAN security configuration: Over time, some security changes might not be reliably propagated across all switches in the fabric. FC SANs should be periodically reviewed (Just like IP and Ethernet-based networking), to assess their security, identify and prioritize gaps, and define a remediation plan. Security review should be performed at least annually, and in sensitive environments, at least quarterly, or after any major change – whichever comes first. Audit results should be documented.

4.12 Training

ST-SS-R1 – Storage security training: A storage security training program should be defined, and incorporated into existing organizational training activities and schedules, to accommodate the following audiences:

- **Information Security Professionals** – to provide them the fundamental background of storage security.
- **Storage administrators** – to familiarize them with storage security principles, and organization policies and security baselines.
- **Managers** – to understand the fundamentals of data protection.

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

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 storage resources. It has also 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:

- 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.
- Areas that are specific to storage infrastructures, such as data protection and confidentiality protection using encryption, isolation, and restoration assurance.

Along with compute (encompassing operating system and host hardware) and network infrastructures, storage infrastructure is one of the three fundamental pillars of IT. However, compared to 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 security recommendations for storage infrastructures in this document provide a basis for securing an important element of IT infrastructure.

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 [40] could significantly improve an organization's resilience to different kinds of attacks on data resources.

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