Acronis Storage 2.4

Installation Guide

January 15, 2019

Contents

CHAPTER 1

Introduction

To support the growing demand for both high performance and high data availability, modern data centers need a fast, flexible storage solution. Existing solutions, however, are often difficult to manage and maintain, or not flexible enough (e.g., local RAID arrays), or too expensive (e.g., storage area networks).

Acronis Storage is designed to solve these issues. It can run on commodity hardware, so no significant infrastructure investments are needed. It also is easy to set up and grow on demand.

1.1 About Acronis Storage

Acronis Storage is a software-defined storage solution that allows you to quickly and easily transform low-cost commodity storage hardware and network equipment into protected enterprise-grade storage like SAN or NAS.

Acronis Storage is optimized for storing large amounts of data and provides data redundancy (replication and erasure coding), high availability, self-healing, and storage sharing.

In Acronis Storage, user data is stored on organized clusters of servers in the form of fixed-size chunks. These chunks are automatically replicated and distributed across available servers in the cluster to ensure high availability of user data.

Cluster storage space can be exported through access points like iSCSI, S3, NFS, or Acronis Backup Gateway.

1.2 Deployment Overview

To deploy Acronis Storage for evaluation purposes or in production, you will need to do the following:

1. Plan the infrastructure.
2. Install and configure Acronis Storage on each server in the planned infrastructure.
3. Create the storage cluster.
4. Set up data export for the cluster.
5. Populate the cluster with user data.

CHAPTER 2

Planning Acronis Storage Infrastructure

To plan your Acronis Storage infrastructure, you will need to decide on the hardware configuration of each server, plan the Acronis Storage networks, decide on the redundancy method (and mode) to use, and decide which data will be kept on which storage tier.

Information in this chapter is meant to help you complete all of these tasks.

2.1 Understanding Acronis Storage Architecture

The fundamental component of Acronis Storage is a cluster: a group of physical servers interconnected by network. Each server in a cluster is assigned one or more roles and typically runs services that correspond to these roles:

- storage role: chunk service or CS
- metadata role: metadata service or MDS
- network roles:
  - iSCSI access point service (iSCSI)
  - Acronis Backup Gateway access point service (ABGW)
  - S3 gateway (access point) service (GW)

    - S3 name service (NS)
    - S3 object service (OS)
    - Admin panel
    - SSH
- supplementary roles:
    - management,
    - SSD cache,
    - system

Any server in the cluster can be assigned a combination of storage, metadata, and network roles. For example, a single server can be an S3 access point, an iSCSI access point, and a storage node at once.

Each cluster also requires that a web-based admin panel be installed on one (and only one) of the nodes. The panel enables administrators to manage the cluster.

2.1.1 Storage Role

Storage nodes run chunk services, store all the data in the form of fixed-size chunks, and provide access to these chunks. All data chunks are replicated and the replicas are kept on different storage nodes to achieve high availability of data. If one of the storage nodes fails, remaining healthy storage nodes continue providing the data chunks that were stored on the failed node.

Only a server with disks of certain capacity can be assigned the storage role.

2.1.2 Metadata Role

Metadata nodes run metadata services, store cluster metadata, and control how user files are split into chunks and where these chunks are located. Metadata nodes also ensure that chunks have the required amount of replicas and log all important events that happen in the cluster.

To provide system reliability, Acronis Storage uses the Paxos consensus algorithm. It guarantees fault-tolerance if the majority of nodes running metadata services are healthy.

To ensure high availability of metadata in a production environment, at least five nodes in a cluster must be running metadata services. In this case, if up to two metadata services fail, the remaining ones will still be

controlling the cluster.

2.1.3 Network Roles (Storage Access Points)

Storage access points enable you to access data stored in storage clusters via the standard iSCSI and S3 protocols and use the clusters as backend storage for Acronis Backup Cloud.

To benefit from high availability, access points should be set up on multiple node.

The following access points are currently supported:

iSCSI Allows you to use Acronis Storage as a highly available block storage for virtualization, databases, office applications, and other needs.

S3 A combination of scalable and highly available services that allows you to use Acronis Storage as a modern backend for solutions like OpenXchange AppSuite, Dovecot, and Acronis Access. In addition, developers of custom applications can benefit from a Amazon S3-compatible API and compatibility with the S3 libraries for various programming languages, S3 browsers, and web browsers.

ABGW Acronis Backup Gateway allows you to connect Acronis Storage to Acronis Backup Cloud via Acronis FES API.

NFS Allows you to create redundant NFS exports in Acronis Storage that can be mounted like regular directories shared over NFS.

The following remote management roles are supported:

Admin panel Allows you to access the web-based user interface from an external network.

SSH Allows you to connect to Acronis Storage nodes via SSH.

2.1.4 Supplementary Roles

Internal management Provides a web-based admin panel that enables administrators to configure, manage, and monitor storage clusters. Only one admin panel is needed to create and manage multiple clusters (and only one is allowed per cluster).

SSD cache Boosts chunk read/write performance by creating write caches on selected solid-state drives (SSDs). It is recommended to also use such SSDs for metadata, see *Metadata Role* on page 4. The use of write journals may speed up write

operations in the cluster by two and more times.

System One disk per node that is reserved for the operating system and unavailable for data storage.

2.2 Planning Node Hardware Configurations

Acronis Storage works on top of commodity hardware, so you can create a cluster from regular servers, disks, and network cards. Still, to achieve the optimal performance, a number of requirements must be met and a number of recommendations should be followed.

2.2.1 Hardware Requirements

The following table lists the minimal and recommended hardware for a single node in the cluster:

Type	Minimal	Recommended
CPU	Dual-core CPU	Intel Xeon E5-2620V2 or faster; at least one CPU core per 8 HDDs
RAM	4GB	16GB ECC or more, plus 0.5GB ECC per each HDD
Storage	System: 100GB SATA HDD Metadata: 100GB SATA HDD (on the first five nodes in the cluster) Storage: 100GB SATA HDD	System: 250GB SATA HDD Metadata+Cache: One or more recommended enterprise-grade SSDs with power loss protection; 100GB or more capacity; and 75 MB/s sequential write performance per serviced HDD. For example, a node with 10 HDDs will need an SSD with at least 750 MB/s sequential write speed (on the first five nodes in the cluster) Storage: Four or more HDDs or SSDs; 1 DWPD endurance minimum, 10 DWPD recommended
Disk controller	None	HBA or RAID
Network	1Gbps or faster network interface	Two 10Gbps network interfaces; dedicated links for internal and public networks

Type	Minimal	Recommended
Sample configuration		Intel Xeon E5-2620V2, 32GB, 2xST1000NM0033, 32xST6000NM0024, 2xMegaRAID SAS 9271/9201, Intel X540-T2, Intel P3700 800GB

2.2.2 Hardware Recommendations

The following recommendations explain the benefits added by specific hardware in the hardware requirements table and are meant to help you configure the cluster hardware in an optimal way:

2.2.2.1 Cluster Composition Recommendations

Designing an efficient cluster means finding a compromise between performance and cost that suits your purposes. When planning, keep in mind that a cluster with many nodes and few disks per node offers higher performance while a cluster with the minimal number of nodes (5) and a lot of disks per node is cheaper. See the following table for more details.

Design considerations	Minimum nodes (5), many disks per node	Many nodes, few disks per node
Optimization	Lower cost.	Higher performance.
Free disk space to reserve	More space to reserve for cluster rebuilding as fewer healthy nodes will have to store the data from a failed node.	Less space to reserve for cluster rebuilding as more healthy nodes will have to store the data from a failed node.
Redundancy	Fewer erasure coding choices.	More erasure coding choices.
Cluster balance and rebuilding performance	Worse balance and slower rebuilding.	Better balance and faster rebuilding.
Network capacity	More network bandwidth required to maintain cluster performance during rebuilding.	Less network bandwidth required to maintain cluster performance during rebuilding.
Favorable data type	Cold data (e.g., backups).	Hot data (e.g., virtual environments).
Sample server configuration	Supermicro SSG-6047R-E1R36L (Intel Xeon E5-2620 v4 CPU, 32GB RAM, 36 x 12TB HDDs, 1 x 500GB system disk).	Supermicro SYS-2028TP-HC0R-SIOM (4 x Intel E5-2620 v4 CPUs, 4 x 16GB RAM, 24 x 1.9TB Samsung SM863a SSDs).

Note:

1. These considerations only apply if failure domain is host.
2. The speed of rebuilding in the replication mode does not depend on the number of nodes in the cluster.
3. Acronis Storage supports hundreds of disks per node. If you plan to use more than 36 disks per node, contact sales engineers who will help you design a more efficient cluster.

2.2.2.2 General Hardware Recommendations

- At least five nodes are required for a production environment. This is to ensure that the cluster can survive failure of two nodes without data loss.
- One of the strongest features of Acronis Storage is scalability. The bigger the cluster, the better Acronis Storage performs. It is recommended to create production clusters from at least ten nodes for improved resiliency, performance, and fault tolerance in production scenarios.
- Even though a cluster can be created on top of varied hardware, using nodes with similar hardware in each node will yield better cluster performance, capacity, and overall balance.
- Any cluster infrastructure must be tested extensively before it is deployed to production. Such common points of failure as SSD drives and network adapter bonds must always be thoroughly verified.
- It is not recommended for production to run Acronis Storage in virtual machines or on top of SAN/NAS hardware that has its own redundancy mechanisms. Doing so may negatively affect performance and data availability.
- To achieve best performance, keep at least 20% of cluster capacity free.
- During disaster recovery, Acronis Storage may need additional disk space for replication. Make sure to reserve at least as much space as any single storage node has.
- If you plan to use Acronis Backup Gateway to store backups in the cloud, make sure the local storage cluster has plenty of logical space for staging (keeping backups locally before sending them to the cloud). For example, if you perform backups daily, provide enough space for at least 1.5 days' worth of backups. For more details, see the Administrator's Guide.

2.2.2.3 Storage Hardware Recommendations

- It is possible to use disks of different size in the same cluster. However, keep in mind that, given the same IOPS, smaller disks will offer higher performance per terabyte of data compared to bigger disks. It is recommended to group disks with the same IOPS per terabyte in the same tier.
- Using the recommended SSD models may help you avoid loss of data. Not all SSD drives can withstand enterprise workloads and may break down in the first months of operation, resulting in TCO spikes.
  - SSD memory cells can withstand a limited number of rewrites. An SSD drive should be viewed as a consumable that you will need to replace after a certain time. Consumer-grade SSD drives can withstand a very low number of rewrites (so low, in fact, that these numbers are not shown in their technical specifications). SSD drives intended for Acronis Storage clusters must offer at least 1 DWPD endurance (10 DWPD is recommended). The higher the endurance, the less often SSDs will need to be replaced, improving TCO.
  - Many consumer-grade SSD drives can ignore disk flushes and falsely report to operating systems that data was written while it in fact was not. Examples of such drives include OCZ Vertex 3, Intel 520, Intel X25-E, and Intel X-25-M G2. These drives are known to be unsafe in terms of data commits, they should not be used with databases, and they may easily corrupt the file system in case of a power failure. For these reasons, use to enterprise-grade SSD drives that obey the flush rules (for more information, see the PostgreSQL documentation). Enterprise-grade SSD drives that operate correctly usually have the power loss protection property in their technical specification. Some of the market names for this technology are Enhanced Power Loss Data Protection (Intel), Cache Power Protection (Samsung), Power-Failure Support (Kingston), Complete Power Fail Protection (OCZ).
  - Consumer-grade SSD drives usually have unstable performance and are not suited to withstand sustainable enterprise workloads. For this reason, pay attention to sustainable load tests when choosing SSDs. We recommend the following enterprise-grade SSD drives which are the best in terms of performance, endurance, and investments: Intel S3710, Intel P3700, Huawei ES3000 V2, Samsung SM1635, and Sandisk Lightning.
- The use of SSDs for write caching improves random I/O performance and is highly recommended for all workloads with heavy random access (e.g., iSCSI volumes).
- Running metadata services on SSDs improves cluster performance. To also minimize CAPEX, the same SSDs can be used for write caching.

- If capacity is the main goal and you need to store non-frequently accessed data, choose SATA disks over SAS ones. If performance is the main goal, choose SAS disks over SATA ones.
- The more disks per node the lower the CAPEX. As an example, a cluster created from ten nodes with two disks in each will be less expensive than a cluster created from twenty nodes with one disk in each.
- Using SATA HDDs with one SSD for caching is more cost effective than using only SAS HDDs without such an SSD.
- Use HBA controllers as they are less expensive and easier to manage than RAID controllers.
- Disable all RAID controller caches for SSD drives. Modern SSDs have good performance that can be reduced by a RAID controller's write and read cache. It is recommend to disable caching for SSD drives and leave it enabled only for HDD drives.
- If you use RAID controllers, do not create RAID volumes from HDDs intended for storage (you can still do so for system disks). Each storage HDD needs to be recognized by Acronis Storage as a separate device.
- If you use RAID controllers with caching, equip them with backup battery units (BBUs) to protect against cache loss during power outages.

2.2.2.4 Network Hardware Recommendations

- Use separate networks (and, ideally albeit optionally, separate network adapters) for internal and public traffic. Doing so will prevent public traffic from affecting cluster I/O performance and also prevent possible denial-of-service attacks from the outside.
- Network latency dramatically reduces cluster performance. Use quality network equipment with low latency links. Do not use consumer-grade network switches.
- Do not use desktop network adapters like Intel EXPI9301CTBLK or Realtek 8129 as they are not designed for heavy load and may not support full-duplex links. Also use non-blocking Ethernet switches.
- To avoid intrusions, Acronis Storage should be on a dedicated internal network inaccessible from outside.
- Use one 1 Gbit/s link per each two HDDs on the node (rounded up). For one or two HDDs on a node, two bonded network interfaces are still recommended for high network availability. The reason for this recommendation is that 1 Gbit/s Ethernet networks can deliver 110-120 MB/s of throughput, which is close to sequential I/O performance of a single disk. Since several disks on a server can deliver higher throughput than a single 1 Gbit/s Ethernet link, networking may become a bottleneck.

- For maximum sequential I/O performance, use one 1Gbit/s link per each hard drive, or one 10Gbit/s link per node. Even though I/O operations are most often random in real-life scenarios, sequential I/O is important in backup scenarios.
- For maximum overall performance, use one 10 Gbit/s link per node (or two bonded for high network availability).
- It is not recommended to configure 1 Gbit/s network adapters to use non-default MTUs (e.g., 9000-byte jumbo frames). Such settings require additional configuration of switches and often lead to human error. 10 Gbit/s network adapters, on the other hand, need to be configured to use jumbo frames to achieve full performance.

2.2.3 Hardware and Software Limitations

Hardware limitations:

- Each physical server must have at least two disks with the assigned three roles: System, Metadata, and Storage. The System role can be combined with the Metadata or Storage role, if the system disk capacity is greater than 100GB.

  Note:

  1. It is recommended to assign the System+Metadata role to an SSD. Assigning both these roles to an HDD will result in mediocre performance suitable only for cold data (e.g., archiving).
  2. The System role cannot be combined with the Cache and Metadata+Cache roles. The reason is that is I/O generated by the operating system and applications would contend with I/O generated by journaling, negating its performance benefits.

- Five servers are required to test all the features of the product.
- The system disk must have at least 100 GBs of space.

Software limitations:

- The maintenance mode is not supported. Use SSH to shut down or reboot a node.
- One node can be a part of only one cluster.
- Only one S3 cluster can be created on top of a storage cluster.
- Only predefined redundancy modes are available in the management panel.

- Thin provisioning is always enabled for all data and cannot be configured otherwise.

Note: For network limitations, see *Network Limitations* on page 19.

2.2.4 Minimum Configuration

The minimum configuration described in the table will let you evaluate Acronis Storage features:

Node #	1st disk role	2nd disk role	3rd and other disk roles	Access points
1	System	Metadata	Storage	iSCSI, Object Storage private, S3 public, NFS, ABGW
2	System	Metadata	Storage	iSCSI, Object Storage private, S3 public, NFS, ABGW
3	System	Metadata	Storage	iSCSI, Object Storage private, S3 public, NFS, ABGW
4	System	Metadata	Storage	iSCSI, Object Storage private, ABGW
5	System	Metadata	Storage	iSCSI, Object Storage private, ABGW
5 nodes in total		5 MDSs in total	5 or more CSs in total	Access point services run on five nodes in total

Note: SSD disks can be assigned metadata and cache roles at the same time, freeing up one more disk for the storage role.

Even though five nodes are recommended even for the minimal configuration, you can start evaluating Acronis Storage with just one node and add more nodes later. At the very least, a storage cluster must have one metadata service and one chunk service running. However, such a configuration will have two key limitations:

1. Just one MDS will be a single point of failure. If it fails, the entire cluster will stop working.
2. Just one CS will be able to store just one chunk replica. If it fails, the data will be lost.

Important: If you deploy Acronis Storage on a single node, you must take care of making its storage persistent and redundant to avoid data loss. If the node is physical, it must have multiple disks so you

can replicate the data among them. If the node is a virtual machine, make sure that this VM is made highly available by the solution it runs on.

Acronis Backup Gateway works with the local object storage in the staging mode. It means that the data to be replicated, migrated, or uploaded to a public cloud is first stored locally and only then sent to the destination. It is vital that the local object storage is persistent and redundant so the local data does not get lost. There are multiple ways to ensure the persistene and redundancy of the local storage. You can deploy your Acronis Backup Gateway on multiple nodes and select a good redundancy mode. If your gateway is deployed on a single node in Acronis Storage, you can make its storage redundant by replicating it among multiple local disks. If your entire Acronis Storage installation is deployed in a single virtual machine with the sole purpose of creating a gateway, make sure this VM is made highly available by the solution it runs on.

2.2.5 Recommended Configuration

The recommended configuration will help you create clusters for production environments:

Node #	1st disk role	2nd disk role	3rd and other disk roles	Access points
Nodes 1 to 5	System	SSD; metadata, cache	Storage	iSCSI, Object Storage private, S3 public, ABGW
Nodes 6+	System	SSD; cache	Storage	iSCSI, Object Storage private, ABGW
5 or more nodes in total		5 MDSs in total	5 or more CSs in total	All nodes run required access points

Even though a production-ready cluster can be created from just five nodes with recommended hardware, it is still recommended to enter production with at least ten nodes if you are aiming to achieve significant performance advantages over direct-attached storage (DAS) or improved recovery times.

Important: To ensure high availability of metadata, at least five metadata services must be running per cluster in any production environment. In this case, if up to two metadata service fail, the remaining metadata services will still be controlling the cluster.

Following are a number of more specific configuration examples that can be used in production. Each configuration can be extended by adding chunk servers and nodes.

2.2.5.1 HDD Only

This basic configuration requires a dedicated disk for each metadata server.

Nodes 1-5 (base)

Disk No.	Disk Type	Disk Role(s)
1	HDD	System
2	HDD	MDS
3	HDD	CS
...		
N	HDD	CS

Nodes 6+ (extension)

Disk No.	Disk Type	Disk Role(s)
1	HDD	System
2	HDD	CS
3	HDD	CS
...		
N	HDD	CS

2.2.5.2 HDD + System SSD (No Cache)

This configuration is good for creating capacity-oriented clusters.

Nodes 1-5 (base)

Disk No.	Disk Type	Disk Role(s)
1	SSD	System, MDS
2	HDD	CS
3	HDD	CS
...		
N	HDD	CS

Nodes 6+ (extension)

Disk No.	Disk Type	Disk Role(s)
1	SSD	System
2	HDD	CS
3	HDD	CS
...		
N	HDD	CS

2.2.5.3 HDD + SSD

This configuration is good for creating performance-oriented clusters.

Nodes 1-5 (base)

Disk No.	Disk Type	Disk Role(s)
1	HDD	System
2	SSD	MDS, cache
3	HDD	CS
...		
N	HDD	CS

Nodes 6+ (extension)

Disk No.	Disk Type	Disk Role(s)
1	HDD	System
2	SSD	Cache
3	HDD	CS
...		
N	HDD	CS

2.2.5.4 SSD Only

This configuration does not require SSDs for cache.

When choosing hardware for this configuration, have in mind the following:

- Each Acronis Storage client will be able to obtain up to about 40K sustainable IOPS (read + write) from

the cluster.

- If you use the erasure coding redundancy scheme, each erasure coding file, e.g., a single VM's or container's HDD disk, will get up to 2K sustainable IOPS. That is, a user working inside a VM or container will have up to 2K sustainable IOPS per virtual HDD at their disposal. Multiple VMs and containers on a node can utilize more IOPS, up to the client's limit.
- In this configuration, network latency defines more than half of overall performance, so make sure that the network latency is minimal. One recommendation is to have one 10Gbps switch between any two nodes in the cluster.

Nodes 1-5 (base)

Disk No.	Disk Type	Disk Role(s)
1	SSD	System, MDS
2	SSD	CS
3	SSD	CS
...		
N	SSD	CS

Nodes 6+ (extension)

Disk No.	Disk Type	Disk Role(s)
1	SSD	System
2	SSD	CS
3	SSD	CS
...		
N	SSD	CS

2.2.5.5 HDD + SSD (No Cache), 2 Tiers

In this configuration example, tier 1 is for HDDs without cache and tier 2 is for SSDs. Tier 1 can store cold data (e.g., backups), tier 2 can store hot data (e.g., high-performance virtual machines).

Nodes 1-5 (base)

Disk No.	Disk Type	Disk Role(s)	Tier
1	SSD	System, MDS	

Disk No.	Disk Type	Disk Role(s)	Tier
2	HDD	CS	1
3	SSD	CS	2
...			
N	HDD/SSD	CS	1/2

Nodes 6+ (extension)

Disk No.	Disk Type	Disk Role(s)	Tier
1	SSD	System	
2	HDD	CS	1
3	SSD	CS	2
...			
N	HDD/SSD	CS	1/2

2.2.5.6 HDD + SSD, 3 Tiers

In this configuration example, tier 1 is for HDDs without cache, tier 2 is for HDDs with cache, and tier 3 is for SSDs. Tier 1 can store cold data (e.g., backups), tier 2 can store regular virtual machines, and tier 3 can store high-performance virtual machines.

Nodes 1-5 (base)

Disk No.	Disk Type	Disk Role(s)	Tier
1	HDD/SSD	System	
2	SSD	MDS, T2 cache	
3	HDD	CS	1
4	HDD	CS	2
5	SSD	CS	3
...			
N	HDD/SSD	CS	1/2/3

Nodes 6+ (extension)

Disk No.	Disk Type	Disk Role(s)	Tier
1	HDD/SSD	System	

Disk No.	Disk Type	Disk Role(s)	Tier
2	SSD	T2 cache	
3	HDD	CS	1
4	HDD	CS	2
5	SSD	CS	3
...			
N	HDD/SSD	CS	1/2/3

2.2.6 Raw Disk Space Considerations

When planning the Acronis Storage infrastructure, keep in mind the following to avoid confusion:

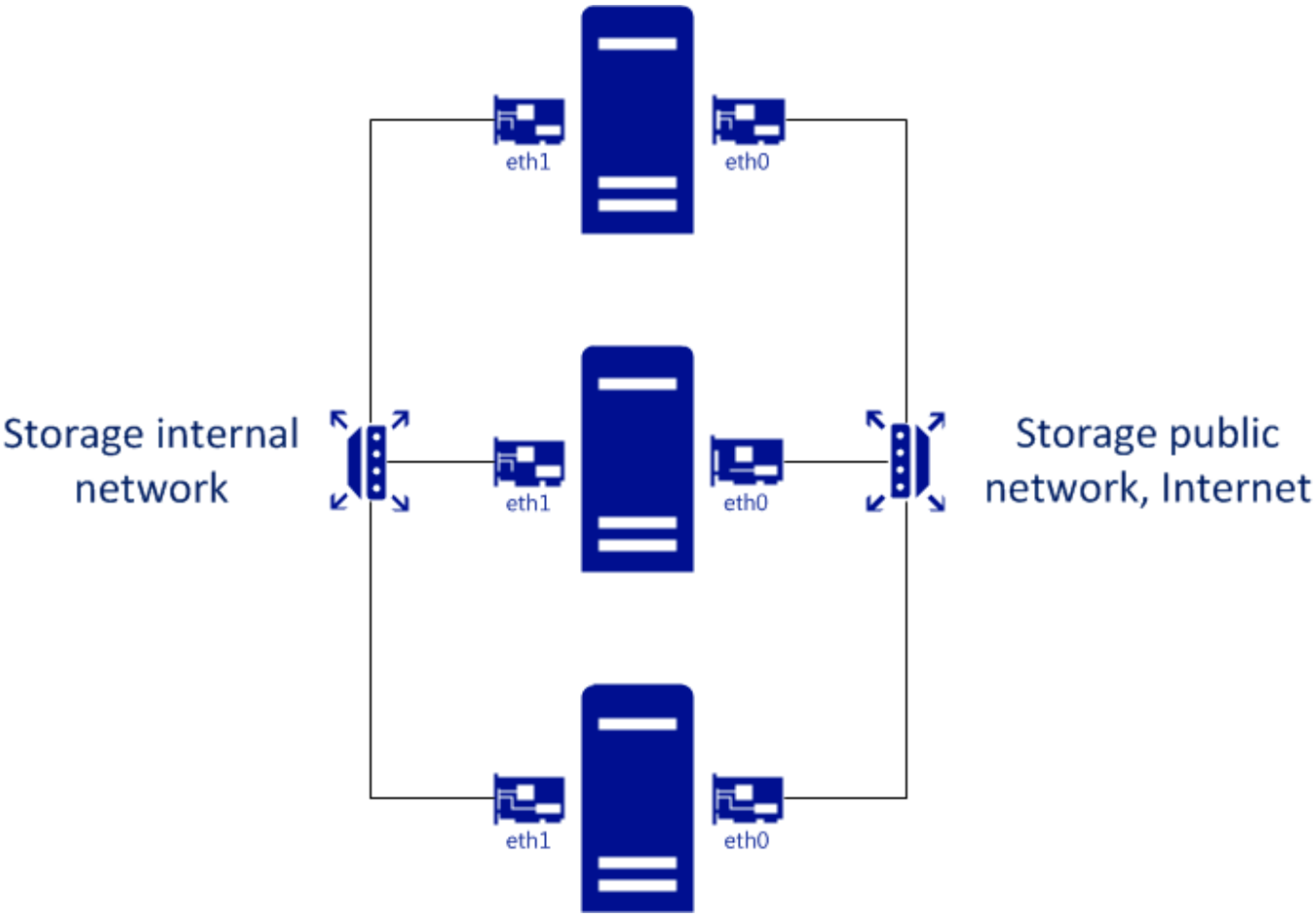
- The capacity of HDD and SSD is measured and specified with decimal, not binary prefixes, so “TB” in disk specifications usually means “terabyte”. The operating system, however, displays drive capacity using binary prefixes meaning that “TB” is “tebibyte” which is a noticeably larger number. As a result, disks may show capacity smaller than the one marketed by the vendor. For example, a disk with 6TB in specifications may be shown to have 5.45 TB of actual disk space in Acronis Storage.
- Acronis Storage reserves 5% of disk space for emergency needs.

Therefore, if you add a 6TB disk to a cluster, the available physical space should increase by about 5.2 TB.

2.3 Planning Network

Acronis Storage uses two networks: (a) an internal network that interconnects nodes and combines them into a cluster, and (b) a public network for managing the cluster via the admin panel and SSH, exporting stored data to users, and providing external access from virtual machines.

The figure below shows a top-level overview of the internal and public networks of Acronis Storage.



2.3.1 General Network Requirements

Make sure that time is synchronized on all nodes in the cluster via NTP. Doing so will make it easier for the support department to understand cluster logs.

2.3.2 Network Limitations

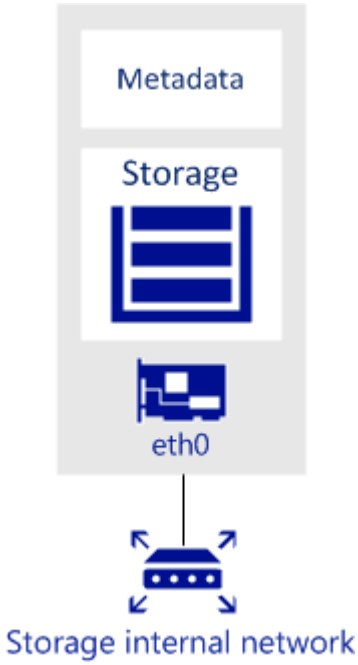
- Nodes are added to clusters by their IP addresses, not FQDNs. Changing the IP address of a node in the cluster will remove that node from the cluster. If you plan to use DHCP in a cluster, make sure that IP addresses are bound to the MAC addresses of nodes' network interfaces.
- Fibre Channel and InfiniBand networks are not supported.
- Each node must have Internet access so updates can be installed.

- MTU is set to 1500 by default.
- Network time synchronization (NTP) is required for correct statistics.
- The management traffic type is assigned automatically during installation and cannot be changed in the admin panel later.
- Even though the management node can be accessed from a web browser by the hostname, you still need to specify its IP address, not the hostname, during installation.

2.3.3 Per-Node Network Requirements

Network requirements for each cluster node depend on services that will run on this node:

- Each node in the cluster must have access to the internal network and have the port 8888 open to listen for incoming connections from the internal network.
- Each storage and metadata node must have at least one network interface for the internal network traffic. The IP addresses assigned to this interface must be either static or, if DHCP is used, mapped to the adapter's MAC address. The figure below shows a sample network configuration for a storage and metadata node.

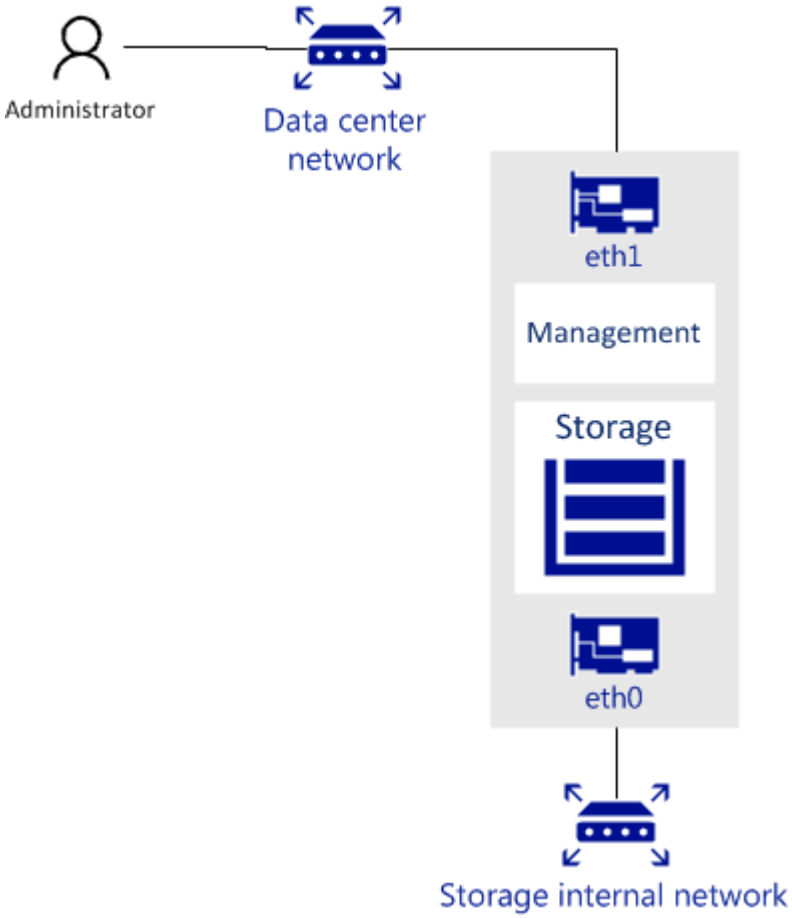


- The management node must have a network interface for internal network traffic and a network interface for the public network traffic (e.g., to the datacenter or a public network) so the admin panel

can be accessed via a web browser.

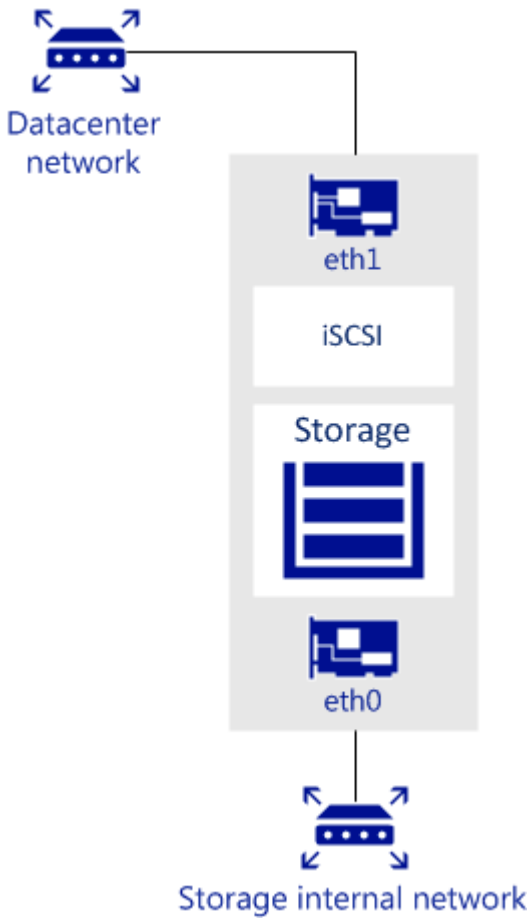
The management node must have the port 8888 open by default to allow access to the admin panel from the public network and to the cluster node from the internal network.

The figure below shows a sample network configuration for a storage and management node.



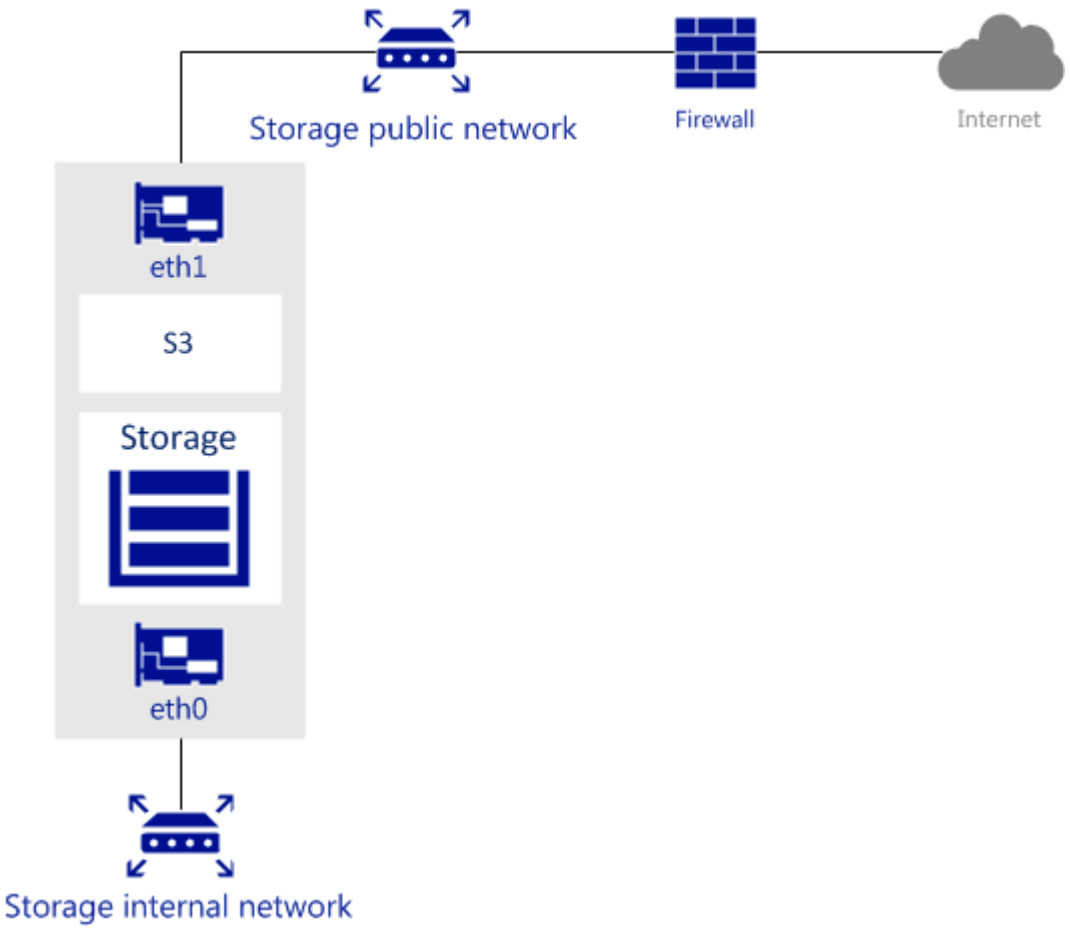
- A node that runs one or more storage access point services must have a network interface for the internal network traffic and a network interface for the public network traffic.

  The figure below shows a sample network configuration for a node with an iSCSI access point. iSCSI access points use the TCP port 3260 for incoming connections from the public network.



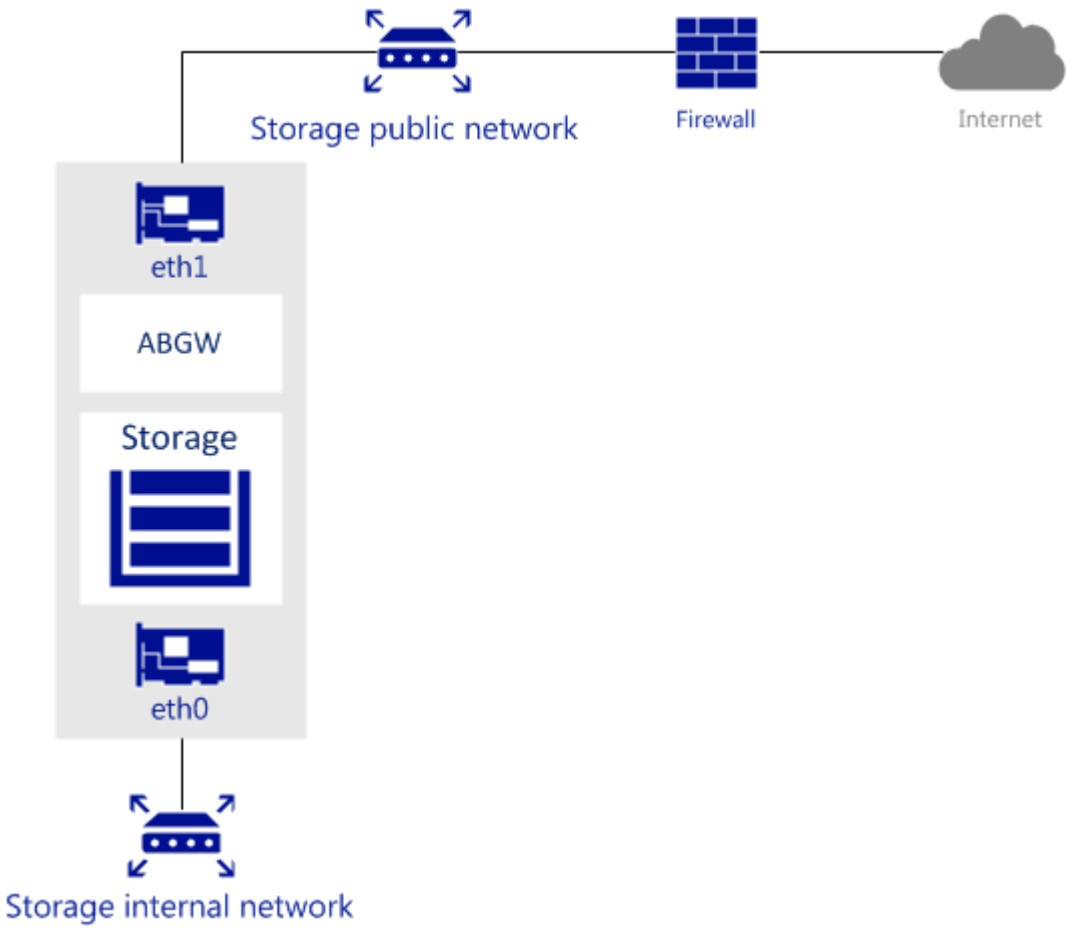
The next figure shows a sample network configuration for a node with an S3 storage access point. S3 access points use ports 443 (HTTPS) and 80 (HTTP) to listen for incoming connections from the public network.

2.3. Planning Network



In the scenario pictured above, the internal network is used for both the storage and S3 cluster traffic.

The next figure shows a sample network configuration for a node with an Acronis Backup Gateway storage access point. Acronis Backup Gateway access points use port 44445 for incoming connections from both internal and public networks and ports 443 and 8443 for outgoing connections to the public network.



2.3.4 Network Recommendations for Clients

The following table lists the maximum network performance a client can get with the specified network interface. The recommendation for clients is to use 10Gbps network hardware between any two cluster nodes and minimize network latencies, especially if SSD disks are used.

Storage network interface	Node max. I/O	VM max. I/O (replication)	VM max. I/O (erasure coding)
1 Gbps	100 MB/s	100 MB/s	70 MB/s
2 x 1 Gbps	~175 MB/s	100 MB/s	~130 MB/s
3 x 1 Gbps	~250 MB/s	100 MB/s	~180 MB/s
10 Gbps	1 GB/s	1 GB/s	700 MB/s
2 x 10 Gbps	1.75 GB/s	1 GB/s	1.3 GB/s

2.3.5 Sample Network Configuration

The figure below shows an overview of a sample Acronis Storage network.

In this network configuration:

- The Acronis Storage internal network is a network that interconnects all servers in the cluster. It can be used for the management, storage, and S3 internal services. Each of these services can be moved to a separate dedicated internal network to ensure high performance under heavy workloads.

  This network cannot be accessed from the public network. All servers in the cluster are connected to this network.

  > **Important:** Acronis Storage does not offer protection from traffic sniffing. Anyone with access to the internal network can capture and analyze the data being transmitted.

- The Acronis Storage public network is a network over which the storage space is exported. Depending on where the storage space is exported to, it can be an internal datacenter network or an external public network:
  - An internal datacenter network can be used to manage Acronis Storage and export the storage space over iSCSI to other servers in the datacenter.
  - An external public network can be used to export the storage space to the outside services through S3 and Acronis Backup Gateway storage access points.

2.4 Understanding Data Redundancy

Acronis Storage protects every piece of data by making it redundant. It means that copies of each piece of data are stored across different storage nodes to ensure that the data is available even if some of the storage nodes are inaccessible.

Acronis Storage automatically maintains the required number of copies within the cluster and ensures that all the copies are up-to-date. If a storage node becomes inaccessible, the copies from it are replaced by new ones that are distributed among healthy storage nodes. If a storage node becomes accessible again after downtime, the copies on it which are out-of-date are updated.

The redundancy is achieved by one of two methods: replication or erasure coding (explained in more detail in the next section). The chosen method affects the size of one piece of data and the number of its copies

that will be maintained in the cluster. In general, replication offers better performance while erasure coding leaves more storage space available for data (see table).

Acronis Storage supports a number of modes for each redundancy method. The following table illustrates data overhead of various redundancy modes. The first three lines are replication and the rest are erasure coding.

Redundancy mode	Min. number of nodes required	How many nodes can fail without data loss	Storage overhead, %	Raw space needed to store 100GB of data
1 replica (no redundancy)	1	0	0	100GB
2 replicas	2	1	100	200GB
3 replicas	3	2	200	300GB
Encoding 1+0 (no redundancy)	1	0	0	100GB
Encoding 1+2	3	2	200	300GB
Encoding 3+2	5	2	67	167GB
Encoding 5+2	7	2	40	140GB
Encoding 7+2	9	2	29	129GB
Encoding 17+3	20	3	18	118GB

Note: The 1+0 and 1+2 encoding modes are meant for small clusters that have insufficient nodes for other erasure coding modes but will grow in the future. As redundancy type cannot be changed once chosen (from replication to erasure coding or vice versa), this mode allows one to choose erasure coding even if their cluster is smaller than recommended. Once the cluster has grown, more beneficial redundancy modes can be chosen.

You choose a data redundancy mode when configuring storage access points and their volumes. In particular, when:

- creating LUNs for iSCSI storage access points,
- creating S3 clusters,
- configuring Acronis Backup Gateway storage access points.

No matter what redundancy mode you choose, it is highly recommended is to be protected against a simultaneous failure of two nodes as that happens often in real-life scenarios.

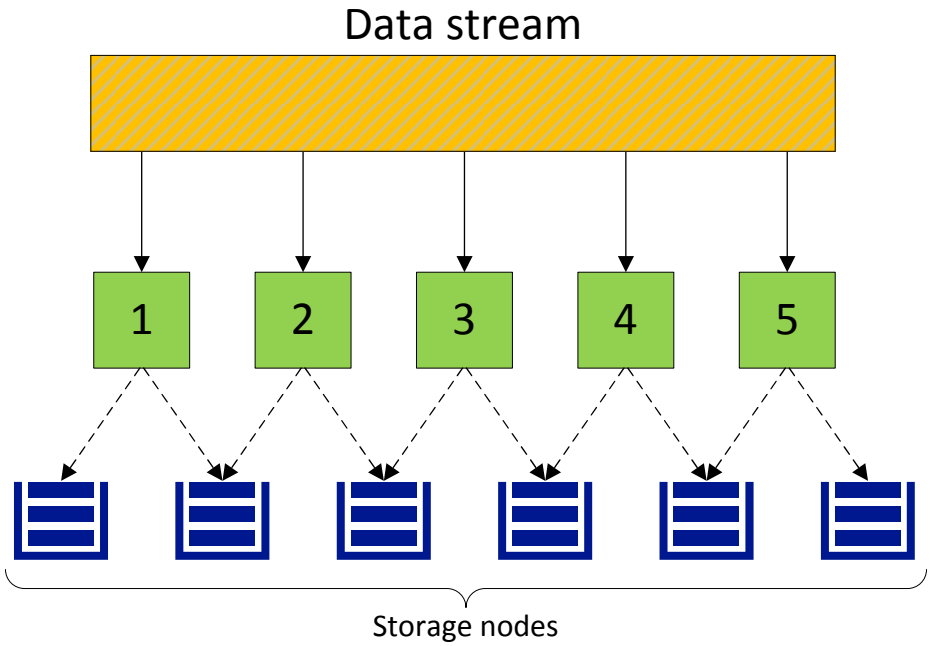
2.4. Understanding Data Redundancy

All redundancy modes allow write operations when one storage node is inaccessible. If two storage nodes are inaccessible, write operations may be frozen until the cluster heals itself.

2.4.1 Redundancy by Replication

With replication, Acronis Storage breaks the incoming data stream into 256MB chunks. Each chunk is replicated and replicas are stored on different storage nodes, so that each node has only one replica of a given chunk.

The following diagram illustrates the 2 replicas redundancy mode.



Replication in Acronis Storage is similar to the RAID rebuild process but has two key differences:

- Replication in Acronis Storage is much faster than that of a typical online RAID 1/5/10 rebuild. The reason is that Acronis Storage replicates chunks in parallel, to multiple storage nodes.
- The more storage nodes are in a cluster, the faster the cluster will recover from a disk or node failure.

High replication performance minimizes the periods of reduced redundancy for the cluster. Replication performance is affected by:

- The number of available storage nodes. As replication runs in parallel, the more available replication sources and destinations there are, the faster it is.
- Performance of storage node disks.

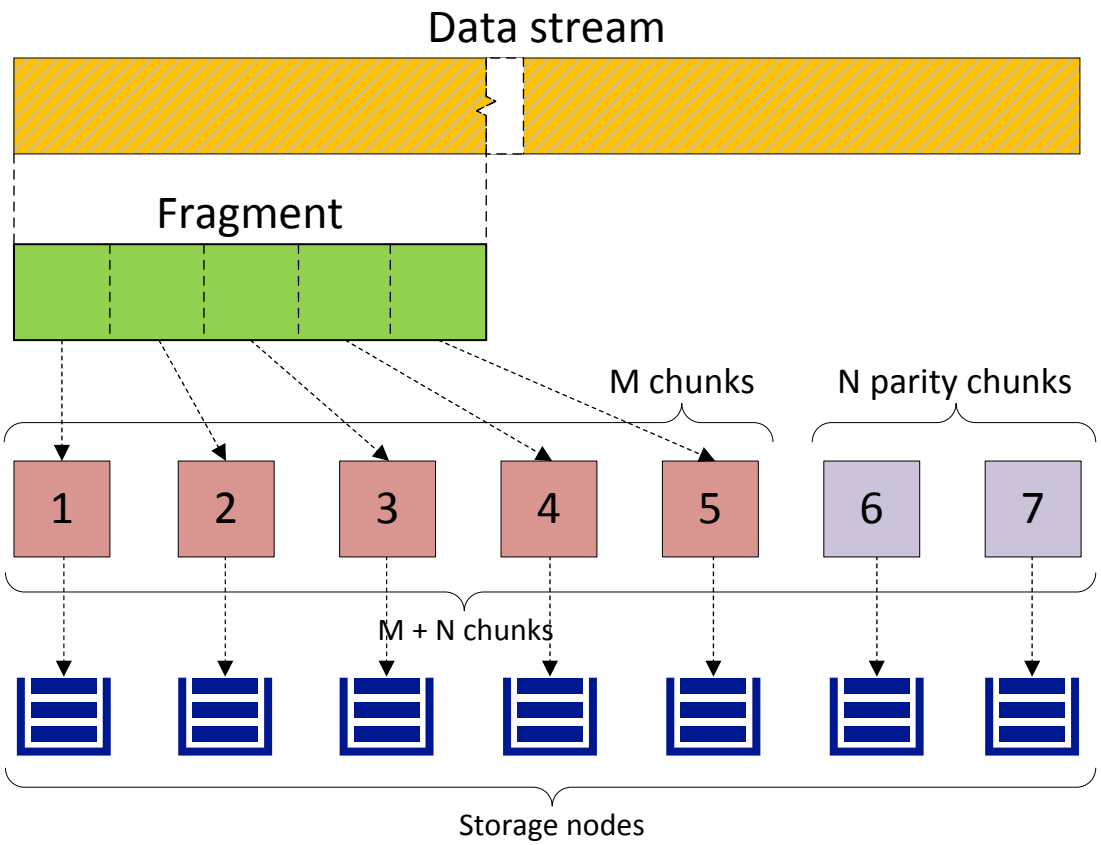
- Network performance. All replicas are transferred between storage nodes over network. For example, 1 Gbps throughput can be a bottleneck (see *Per-Node Network Requirements* on page 20).
- Distribution of data in the cluster. Some storage nodes may have much more data to replicate than other and may become overloaded during replication.
- I/O activity in the cluster during replication.

2.4.2 Redundancy by Erasure Coding

With erasure coding, Acronis Storage breaks the incoming data stream into fragments of certain size, then splits each fragment into a certain number (M) of 1-megabyte pieces and creates a certain number (N) of parity pieces for redundancy. All pieces are distributed among M+N storage nodes, that is, one piece per node. On storage nodes, pieces are stored in regular chunks of 256MB but such chunks are not replicated as redundancy is already achieved. The cluster can survive failure of any N storage nodes without data loss.

The values of M and N are indicated in the names of erasure coding redundancy modes. For example, in the 5+2 mode, the incoming data is broken into 5MB fragments, each fragment is split into five 1MB pieces and two more 1MB parity pieces are added for redundancy. In addition, if N is 2, the data is encoded using the RAID6 scheme, and if N is greater than 2, erasure codes are used.

The diagram below illustrates the 5+2 mode.



2.4.3 No Redundancy

Warning: Danger of data loss!

Without redundancy, singular chunks are stored on storage nodes, one per node. If the node fails, the data may be lost. Having no redundancy is highly not recommended no matter the scenario, unless you only want to evaluate Acronis Storage on a single server.

2.5 Understanding Failure Domains

A failure domain is a set of services which can fail in a correlated manner. To provide high availability of data, Acronis Storage spreads data replicas evenly across failure domains, according to a replica placement policy.

The following policies are available:

- Host as a failure domain (default). If a single host running multiple CS services fails (e.g., due to a power

outage or network disconnect), all CS services on it become unavailable at once. To protect against data loss under this policy, Acronis Storage never places more than one data replica per host. This policy is highly recommended for clusters of five nodes and more.

- Disk, the smallest possible failure domain. Under this policy, Acronis Storage never places more than one data replica per disk or CS. While protecting against disk failure, this option may still result in data loss if data replicas happen to be on different disks of the same host and it fails. This policy can be used with small clusters of up to five nodes (down to a single node).

2.6 Understanding Storage Tiers

In Acronis Storage terminology, tiers are disk groups that allow you to organize storage workloads based on your criteria. For example, you can use tiers to separate workloads produced by different tenants. Or you can have a tier of fast SSDs for service or virtual environment workloads and a tier of high-capacity HDDs for backup storage.

When assigning disks to tiers (which you can do at any time), have in mind that faster storage drives should be assigned to higher tiers. For example, you can use tier 0 for backups and other cold data (CS without SSD cache), tier 1 for virtual environments—a lot of cold data but fast random writes (CS with SSD cache), tier 2 for hot data (CS on SSD), caches, specific disks, and such.

This recommendation is related to how Acronis Storage works with storage space. If a storage tier runs out of free space, Acronis Storage will attempt to temporarily use the space of the lower tiers down to the lowest. If the lowest tier also becomes full, Acronis Storage will attempt to use a higher one. If you add more storage to the original tier later, the data, temporarily stored elsewhere, will be moved to the tier where it should have been stored originally. For example, if you try to write data to the tier 2 and it is full, Acronis Storage will attempt to write that data to tier 1, then to tier 0. If you add more storage to tier 2 later, the aforementioned data, now stored on the tier 1 or 0, will be moved back to the tier 2 where it was meant to be stored originally.

Inter-tier data allocation as well as the transfer of data to the original tier occurs in the background. You can disable such migration and keep tiers strict as described in the Administrator's Command Line Guide.

> **Note:** With the exception of out-of-space situations, automatic migration of data between tiers is not supported.

2.7 Understanding Cluster Rebuilding

The storage cluster is self-healing. If a node or disk fails, a cluster will automatically try to restore the lost data, i.e. rebuild itself.

The rebuild process involves the following steps. Every CS sends a heartbeat message to an MDS every 5 seconds. If a heartbeat is not sent, the CS is considered *inactive* and the MDS informs all cluster components that they stop requesting operations on its data. If no heartbeats are received from a CS for 15 minutes, the MDS considers that CS *offline* and starts cluster rebuilding (if prerequisites below are met). In the process, the MDS finds CSs that do not have pieces (replicas) of the lost data and restores the data—one piece (replica) at a time—as follows:

- If replication is used, the existing replicas of a degraded chunk are locked (to make sure all replicas remain identical) and one is copied to the new CS. If at this time a client needs to read some data that has not been rebuilt yet, it reads any remaining replica of that data.
- If erasure coding is used, the new CS requests almost all the remaining data pieces to rebuild the missing ones. If at this time a client needs to read some data that has not been rebuilt yet, that data is rebuilt out of turn and then read.

Self-healing requires more network traffic and CPU resources if replication is used. On the other hand, rebuilding with erasure coding is slower.

For a cluster to be able to rebuild itself, it must have at least:

1. as many healthy nodes as required by the redundancy mode;
2. enough free space to accomodate as much data as any one node can store.

The first prerequisite can be explained on the following example. In a cluster that works in the 5+2 erasure coding mode and has seven nodes (i.e. the minumum), each piece of user data is distributed to 5+2 nodes for redundancy, i.e. each node is used. If one or two nodes fail, the user data will not be lost, but the cluster will become degraded and will not be able to rebuild itself until at least seven nodes are healthy again (that is, until you add the missing nodes). For comparison, in a cluster that works in the 5+2 erasure coding mode and has ten nodes, each piece of user data is distributed to the random 5+2 nodes out of ten to even out the load on CSs. If up to three nodes fail, such a cluster will still have enough nodes to rebuild itself.

The second prerequisite can be explained on the following example. In a cluster that has ten 10 TB nodes, at least 1 TB on each node should be kept free, so if a node fails, its 9 TB of data can be rebuilt on the remaining nine nodes. If, however, a cluster has ten 10 TB nodes and one 20 TB node, each smaller node should have at

least 2 TB free in case the largest node fails (while the largest node should have 1 TB free).

Two recommendations that help smooth out rebuilding overhead:

- To simplify rebuilding, keep uniform disk counts and capacity sizes on all nodes.
- Rebuilding places additional load on the network and increases the latency of read and write operations. The more network bandwidth the cluster has, the faster rebuilding will be completed and bandwidth freed up.

CHAPTER 3

Installing Acronis Storage

After planning out the infrastructure, proceed to install Acronis Storage on each server included in the plan.

Acronis Storage is installed in a similar way on all required servers. One exception is the first server where you must also install the management panel (only one is allowed per cluster).

> **Note:** On all nodes in the same cluster, time needs to be synchronized via NTP. Make sure the nodes can access the NTP server.

3.1 Preparing for Installation

Acronis Storage can be installed from

- DVD discs (burn the distribution ISO image onto a DVD disc),
- PXE servers (see the *Installation via PXE Server* guide for information on installing Acronis Storage over the network).

  > **Note:** Time synchronization via NTP is enabled by default.

- USB drives

3.1.1 Preparing for Installation from USB Storage Drives

To install Acronis Storage from a USB storage drive, you will need a 2 GB or higher-capacity USB drive and the Acronis Storage distribution ISO image.

Make a bootable USB drive by transferring the distribution image to it with dd.

> **Important:** Be careful to specify the correct drive to transfer the image to.

For example, on Linux:

```
# dd if=storage-image.iso of=/dev/sdb
```

And on Windows (with dd for Windows):

```
C:\>dd if=storage-image.iso of=\\?\Device\Harddisk1\Partition0
```

3.2 Starting Installation

To start the installation, do the following:

1. Configure the server to boot from a DVD or USB drive.
2. Boot the server from the chosen media and wait for the welcome screen.
3. On the welcome screen, choose **Install Acronis Storage**. After the installation program loads, you will see the **Installation Summary** screen. On it, you need to specify a number of parameters required to install Acronis Storage.

3.3 Setting Date and Time

If you need to set the date and time for your Acronis Storage installation, open the **DATE & TIME** screen and make the necessary changes. Make sure that NTP is enabled to synchronize time on each node.

3.4 Configuring Network

Acronis Storage requires at least one network interface per server for management (that is, the management role). You will specify this network interface on the **Component Installation** screen and will not be able to remove the management role from it later (you will, however, be able to add more network roles to it).

Usually network is configured automatically (via DHCP) by the installation program. If you need to modify

network settings, you can do so on the **NETWORK & HOST NAME** screen.

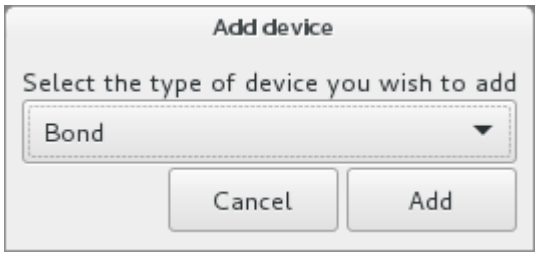
If manual configuration is required, specify the necessary parameters for at least one network card and provide a hostname: either a fully qualified domain name (hostname, domainname) or a short name (hostname).

3.4.1 Creating Bonded Connections

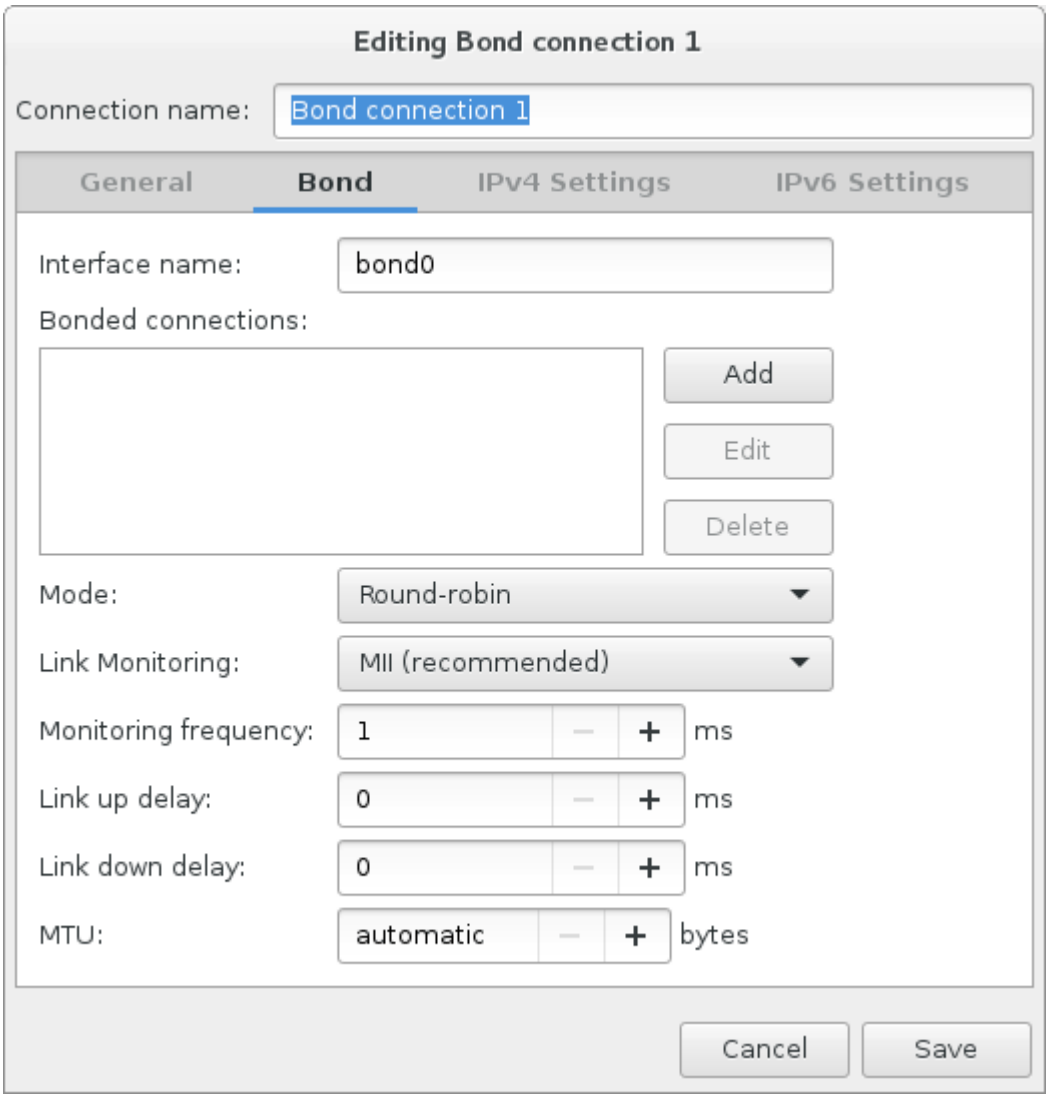
Bonded connections offer increased throughput beyond the capabilities of a single network card as well as improved redundancy.

While installing Acronis Storage, you can configure bonding on the **NETWORK & HOSTNAME** screen as described below.

1. To add a new bonded connection, click the plus button in the bottom, select **Bond** from the drop-down list, and click **Add**.

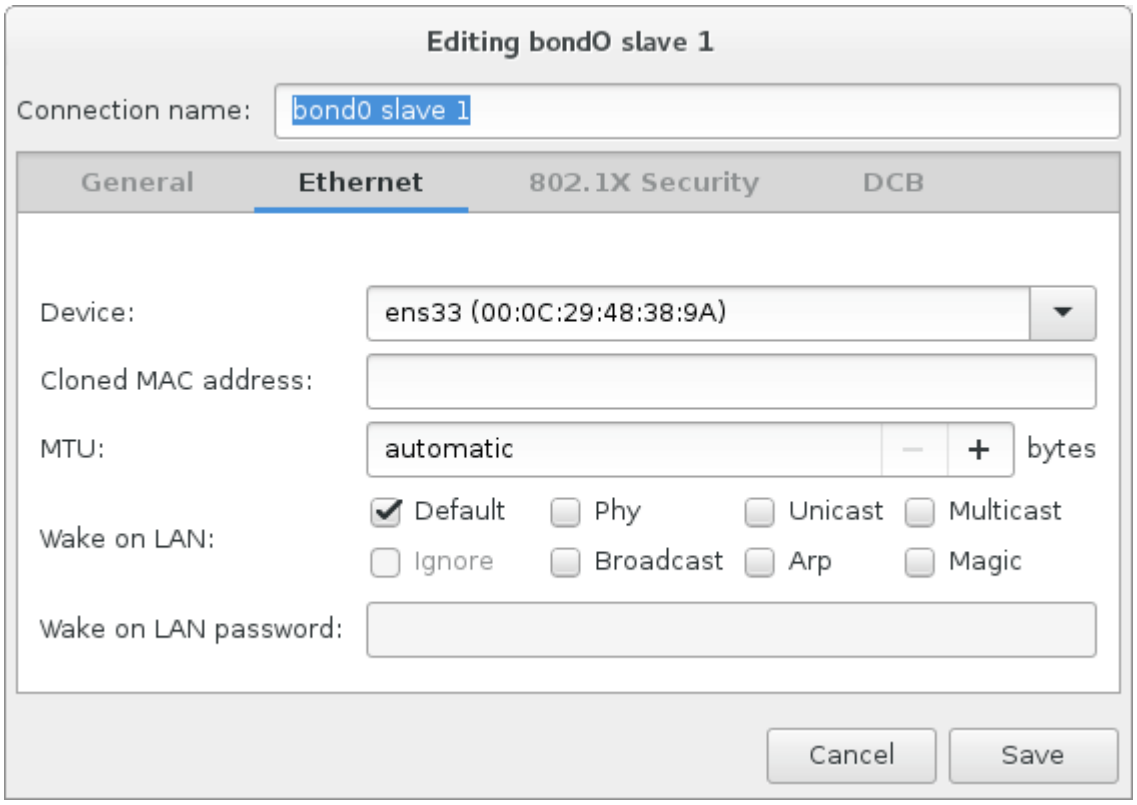


2. In the **Editing Bond connection...** window, click **Add**.



3. In the **Choose a Connection Type** window, select **Ethernet** from the in the drop-down list, and click **Create**.

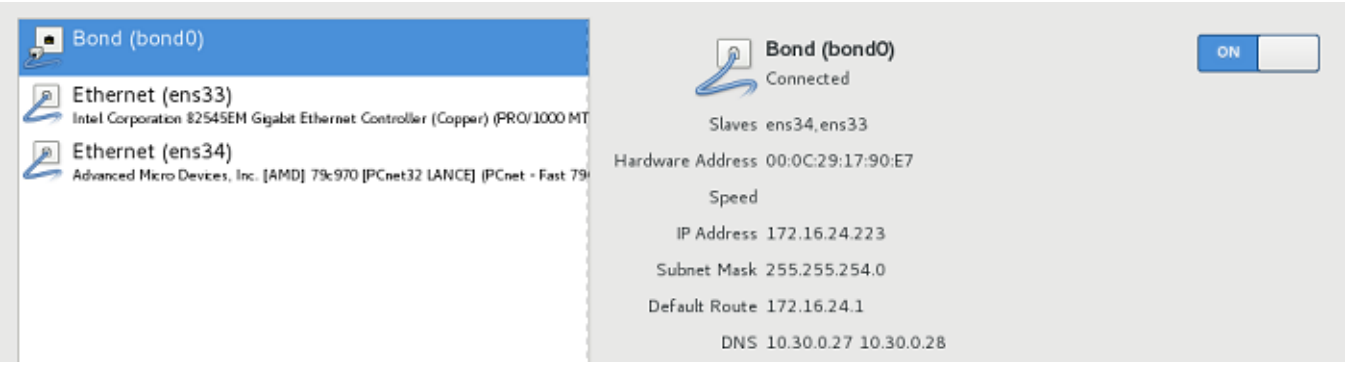
4. In the **Editing bond slave...** window, select a network interface to bond from the **Device** drop-down list.



5. Configure other parameters if required.
6. Click **Save**.
7. Repeat steps 3 to 7 for each network interface you need to add to the bonded connection.
8. Configure other parameters if required.

9. Click **Save**.

The connection will appear in the list on the **NETWORK & HOSTNAME** screen.



3.5 Choosing Components to Install

To install Acronis Storage on a server, you need to choose a component to install on the Acronis Storage screen:

The following options are available:

- **Management Panel.** Install the web-based user interface for managing Acronis Storage clusters.
- **Storage.** Turn the server into a node ready to run Acronis Storage services related to data storage.
- **Management Panel and Storage.** Install both above components at once.

You will need to install **Management Panel** or **Management Panel and Storage** on the first server and **Storage** on all other servers. The detailed instructions are provided in the following sections.

Note:	The management panel will be installed on the system disk.

3.5.1 Choosing Components to Install on the First Server

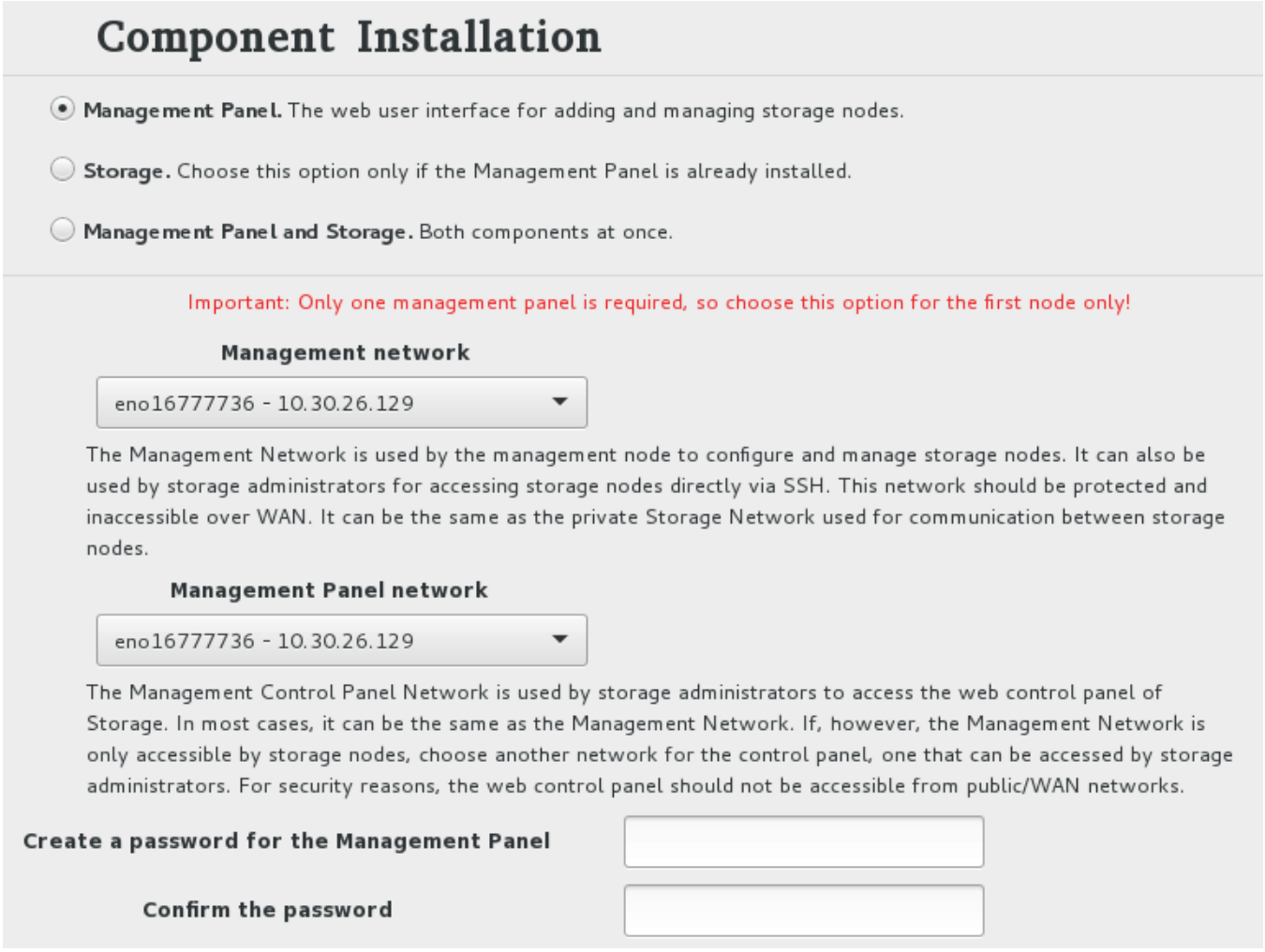
On the first server, you will need to install the management panel (with or without storage, as per your plan).

Do the following on the **Acronis Storage** screen:

1. Choose **Management Panel** or **Management Panel and Storage**.
2. In the **Management Panel network** drop-down list, select a network interface that will provide access

to the management panel.

3. In the **Management network** drop-down list, select a network interface and specify a port for internal management and configuration purposes (the port 8888 is used by default).
4. Create a password for the `admin` account of the management panel and confirm it in the corresponding fields.
5. Click **Done**.



After completing the steps above, proceed to *Selecting Destination Partition* on page 42.

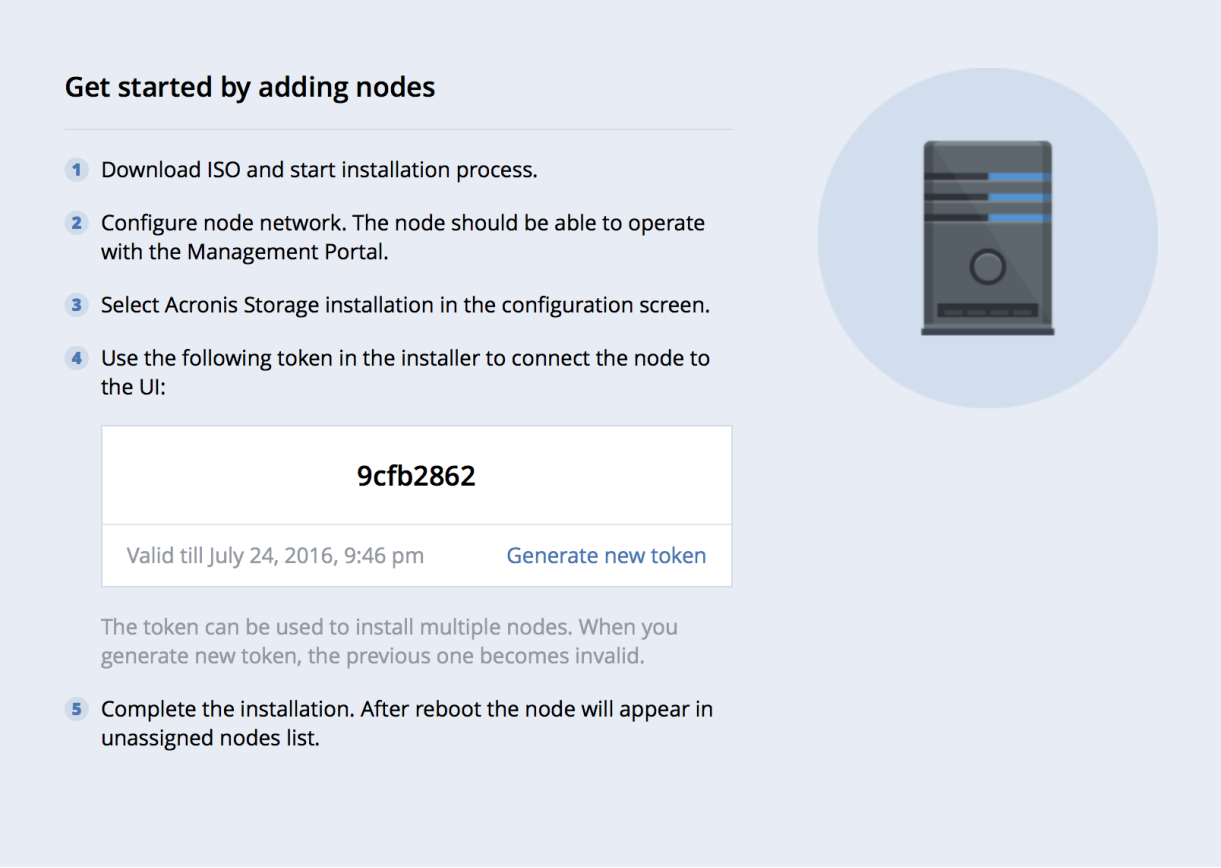
3.5.2 Choosing Components to Install on the Second and Other Servers

On the second and other servers, you will need to install the **Storage** component only. Such servers will run services related to data storage and will be added to the Acronis Storage infrastructure during installation.

For security reasons, you will need to provide a token that can only be obtained from the management panel you have installed on the first server. A single token can be used to install the storage component on multiple servers in parallel.

To obtain a token:

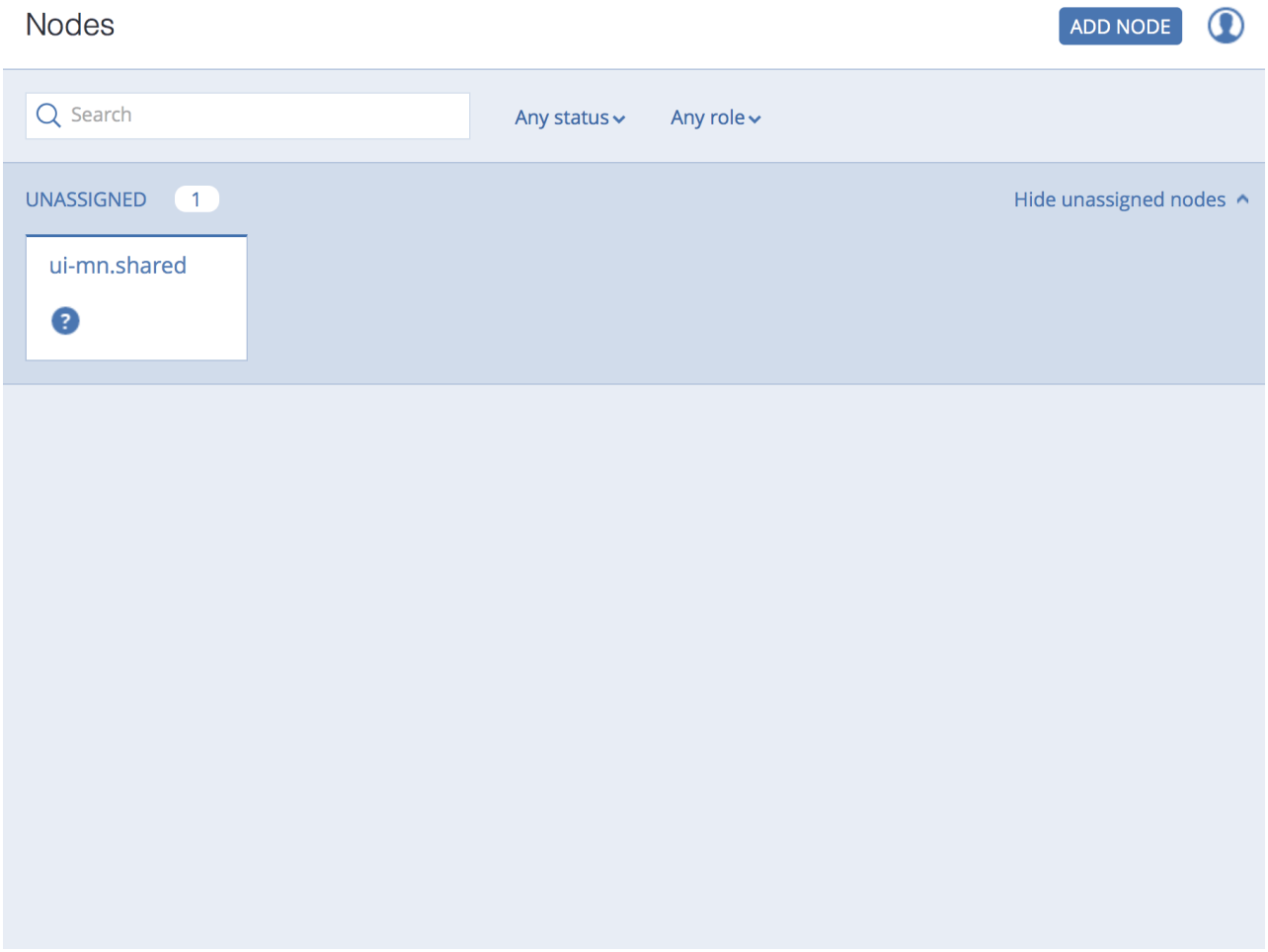
1. Log in to the Acronis Storage management panel. On any computer with access to the management panel network, open a web browser and visit the management node IP address on port 8888: https://<management_node_IP_address>:8888. If prompted, add the security certificate to browser's exceptions.
2. In the management panel:
   - If you only installed the management component on the first server, you will see the welcome screen where a token will be shown (you can generate a new one if needed; generating a new token invalidates the old one).



   - If you installed both the **Management** and **Storage** components on the first server, you will see the **NODES** screen where the only node will be shown in the **UNASSIGNED** list.

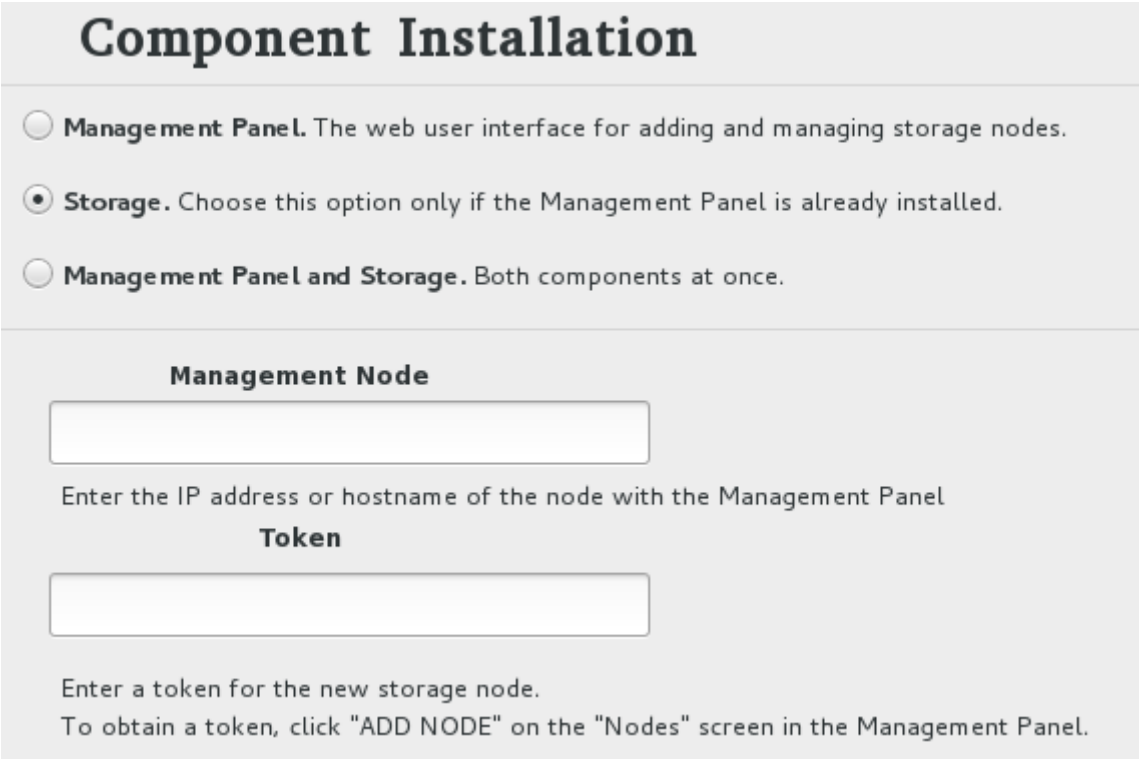
3.5. Choosing Components to Install

Click **ADD NODE** and a screen similar to open a screen similar to the welcome one. On it, a token will be shown (you can generate a new one if needed; generating a new token invalidates the old one).



Having obtained the token, do the following on the **Acronis Storage** screen:

1. Choose **Storage**.



2. In the **Management node** field, specify the IP address of the node with the management panel.
3. In the **Token** field, specify the acquired token.
4. Click **Done** and proceed to *Selecting Destination Partition* on page 42.

3.6 Selecting Destination Partition

You need to choose on which server disk the operating system will be installed. This disk will have the system supplementary role and will not be used for data storage. To choose a system disk, open the **INSTALLATION DESTINATION** screen and select a device in the **Device Selection** section. Configure other options if required.

3.7 Finishing Installation

On the **ROOT PASSWORD** screen, create a password for the root account. Installation cannot be started until the password is created.

3.7. Finishing Installation

Having configured everything necessary on the **INSTALLATION SUMMARY** screen, click **Begin Installation**.

Once the installation is complete, the server will automatically reboot.

Your next steps depend on which server you installed Acronis Storage:

- If you installed the management component on the first server (with or without the storage component), proceed to install the storage component on the second and other servers.
- If you installed the storage component on a server and need to install it on more servers, repeat the installation steps. When on the **Acronis Storage** screen, follow the instructions in *Choosing Components to Install on the Second and Other Servers* on page 39.
- If you installed the storage component on the last server, log in to the management panel and make sure that all the storage nodes are present in the **UNASSIGNED** list on the **NODES** screen.

With the management panel ready and with all the nodes present in the **UNASSIGNED** list, you can start managing your Acronis Storage infrastructure as described in the Administrator's Guide.