How ILM operates throughout an object's life

StorageGRID 11.5

NetApp
January 04, 2024

Table of Contents

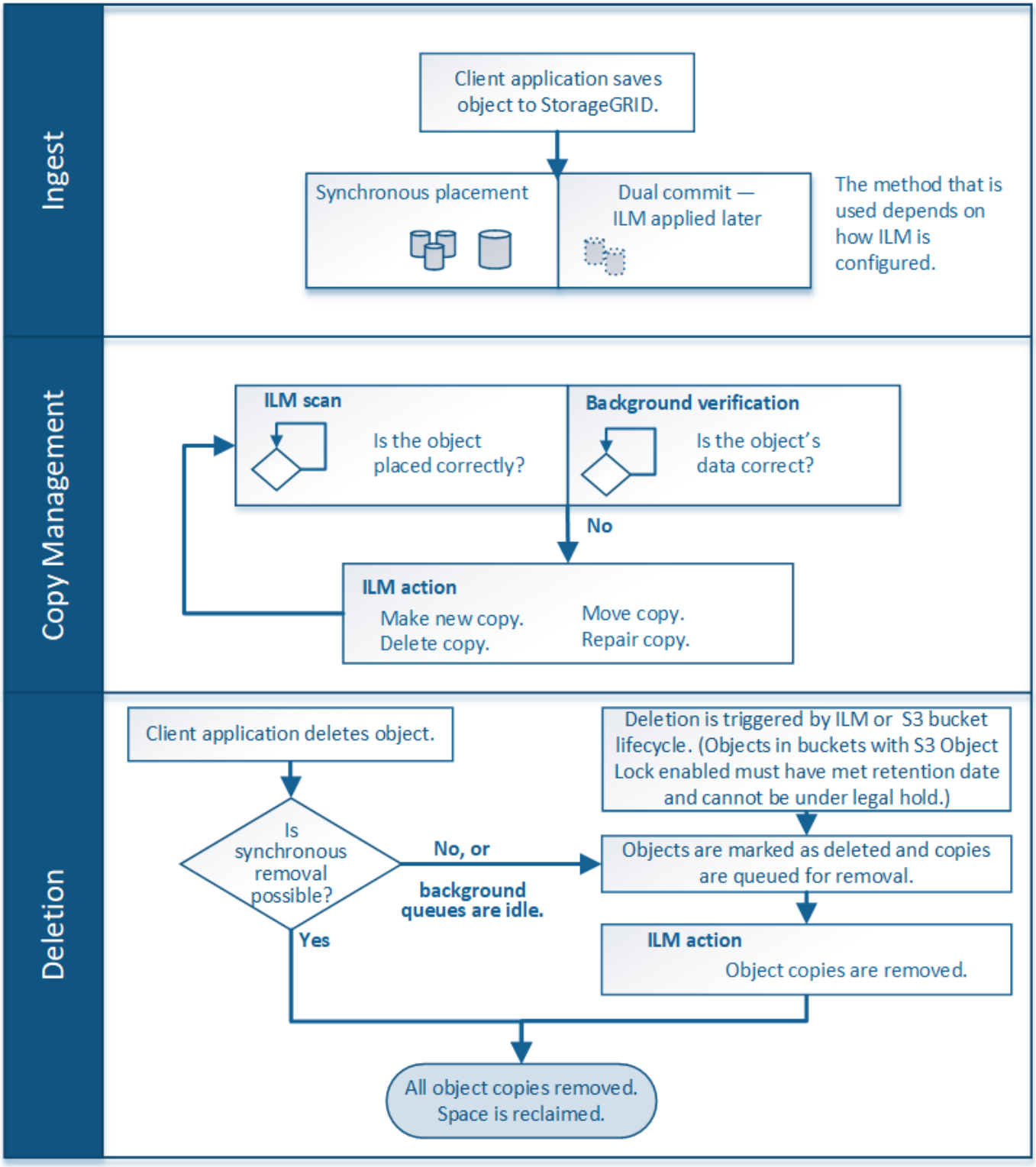
How ILM operates throughout an object's life

Understanding how StorageGRID uses ILM to manage objects during every stage of their life can help you design a more effective policy.

- **Ingest**: Ingest begins when an S3 or Swift client application establishes a connection to save an object to the StorageGRID system, and is complete when StorageGRID returns an "ingest successful" message to the client. Object data is protected during ingest either by applying ILM instructions immediately (synchronous placement) or by creating interim copies and applying ILM later (dual commit), depending on how the ILM requirements were specified.
- **Copy management**: After creating the number and type of object copies that are specified in the ILM's placement instructions, StorageGRID manages object locations and protects objects against loss.
  - ILM scanning and evaluation: StorageGRID continuously scans the list of objects stored in the grid and checks if the current copies meet ILM requirements. When different types, numbers, or locations of object copies are required, StorageGRID creates, deletes, or moves copies as needed.
  - Background verification: StorageGRID continuously performs background verification to check the integrity of object data. If a problem is found, StorageGRID automatically creates a new object copy or a replacement erasure-coded object fragment in a location that meets current ILM requirements. See the instructions for monitoring and troubleshooting StorageGRID.
- **Object deletion**: Management of an object ends when all copies are removed from the StorageGRID system. Objects can be removed as a result of a delete request by a client, or as a result of deletion by ILM or deletion caused by the expiration of an S3 bucket lifecycle.

> Objects in a bucket that has S3 Object Lock enabled cannot be deleted if they are under a legal hold or if a retain-until-date has been specified but not yet met.

The diagram summarizes how ILM operates throughout an object's lifecycle.



Related information

Monitor & troubleshoot

How objects are ingested

StorageGRID protects objects during ingest either by performing synchronous placement or by performing dual commit, as specified in the ILM rule that matches the objects.

When an S3 or Swift client stores an object to the grid, StorageGRID ingests the object using one of these two methods:

- **Synchronous placement**: StorageGRID immediately creates all object copies that are needed to meet ILM requirements. StorageGRID sends an “ingest successful” message to the client when all copies are created.

  If StorageGRID cannot immediately create all object copies (for example, because a required location is temporarily unavailable), it either sends an “ingest failed” message to the client, or it falls back to creating interim object copies and evaluating ILM later, depending on the choice you made when you created the ILM rule.

- **Dual commit**: StorageGRID immediately creates two interim copies of the object, each on a different Storage Node, and sends an “ingest successful” message to the client. StorageGRID then queues the object for ILM evaluation.

  When StorageGRID performs the ILM evaluation, it first checks to see if the interim copies satisfy the placement instructions in the ILM rule. For example, the two interim copies might satisfy the instructions in a two-copy ILM rule, but they would not satisfy the instructions in an erasure-coding rule. If the interim copies do not satisfy the ILM instructions, StorageGRID creates new object copies and deletes any interim copies that are not needed.

  If StorageGRID cannot create two interim copies (for example, if a network issue prevents the second copy from being made), StorageGRID does not retry. Ingest fails.

> S3 or Swift clients can specify that StorageGRID create a single interim copy at ingest by specifying `REDUCED_REDUNDANCY` for the storage class. See the instructions for implementing an S3 or Swift client for more information.

By default, StorageGRID uses synchronous placement to protect objects during ingest.

Related information

Data-protection options for ingest

Use S3

Use Swift

Data-protection options for ingest

When you create an ILM rule, you specify one of three options for protecting objects at ingest: Dual commit, Balanced, or Strict. Depending on your choice, StorageGRID makes interim copies and queues the objects for ILM evaluation later, or it uses synchronous placement and immediately makes copies to meet ILM requirements.

Dual commit

When you select the Dual commit option, StorageGRID immediately makes interim object copies on two different Storage Nodes and returns an “ingest successful” message to the client. The object is queued for ILM evaluation, and copies that meet the rule’s placement instructions are made later.

When to use the Dual commit option

Use the Dual commit option in either of these cases:

- You are using multi-site ILM rules and client ingest latency is your primary consideration. When using Dual commit, you must ensure your grid can perform the additional work of creating and removing the dual-commit copies if they do not satisfy ILM. Specifically:
  - The load on the grid must be low enough to prevent an ILM backlog.
  - The grid must have excess hardware resources (IOPS, CPU, memory, network bandwidth, and so on).
- You are using multi-site ILM rules and the WAN connection between the sites usually has high latency or limited bandwidth. In this scenario, using the Dual commit option can help prevent client timeouts. Before choosing the Dual commit option, you should test the client application with realistic workloads.

Strict

When you select the Strict option, StorageGRID uses synchronous placement on ingest and immediately makes all object copies specified in the rule's placement instructions. Ingest fails if StorageGRID cannot create all copies, for example, because a required storage location is temporarily unavailable. The client must retry the operation.

When to use the Strict option

Use the Strict option if you have an operational or regulatory requirement to immediately store objects only in the locations outlined in the ILM rule. For example, to satisfy a regulatory requirement, you might need to use the Strict option and a Location Constraint advanced filter to guarantee that objects are never stored at certain data center.

Example 5: ILM rules and policy for Strict ingest behavior

Balanced

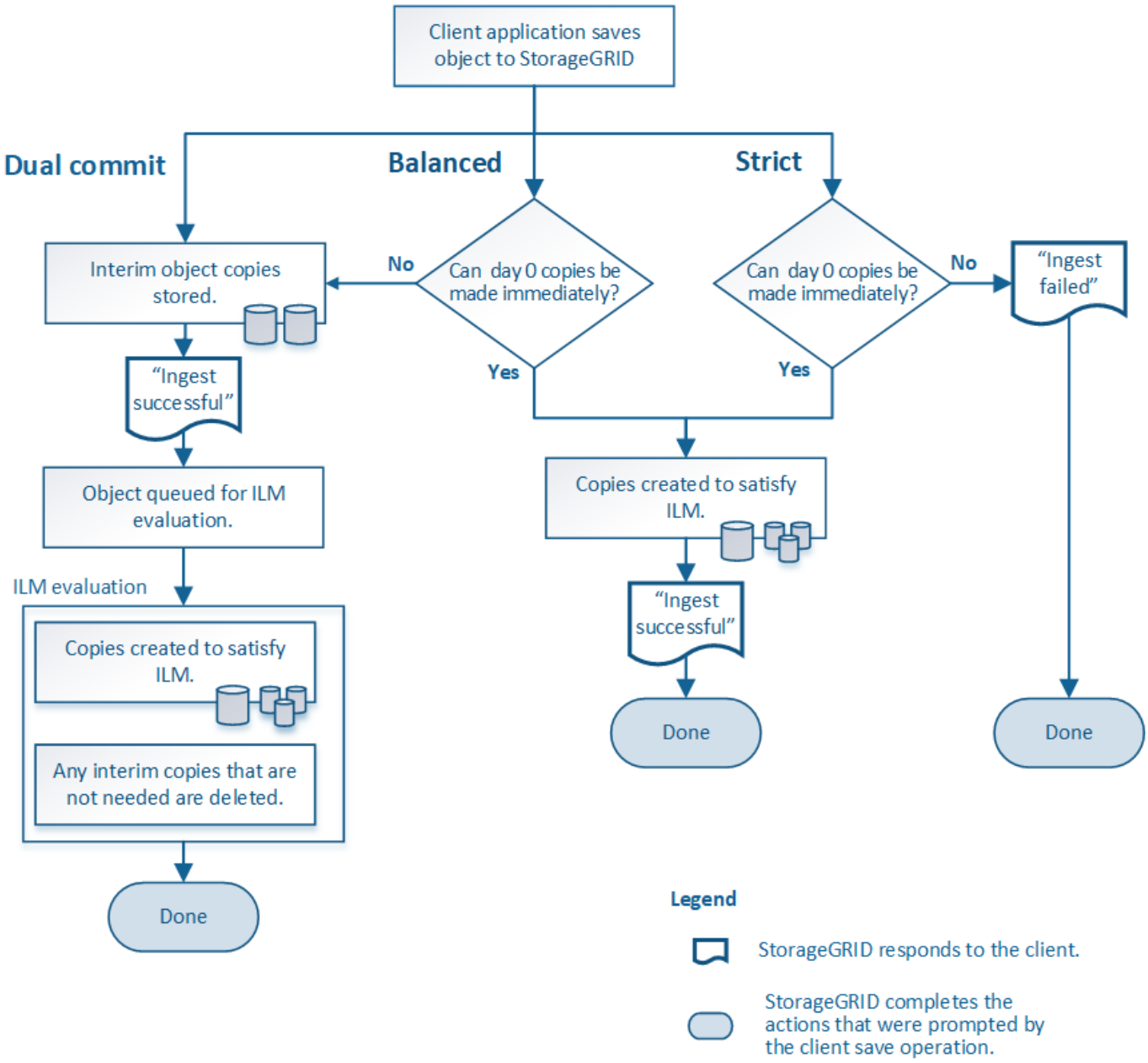
When you select the Balanced option, StorageGRID also uses synchronous placement on ingest and immediately makes all copies specified in the rule's placement instructions. In contrast with the Strict option, if StorageGRID cannot immediately make all copies, it uses Dual commit instead.

When to use the Balanced option

Use the Balanced option to achieve the best combination of data protection, grid performance, and ingest success. Balanced is the default option in the ILM rule wizard.

Flowchart of three ingest options

The flowchart shows what happens when objects are matched by an ILM rule that uses one of these ingest options.



Related information

How objects are ingested

Advantages, disadvantages, and limitations of the data-protection options

Understanding the advantages and disadvantages of each of the three options for protecting data at ingest (Balanced, Strict, or Dual commit) can help you decide which one to select for an ILM rule.

Advantages of the Balanced and Strict options

When compared to Dual commit, which creates interim copies during ingest, the two synchronous placement options can provide the following advantages:

- **Better data security**: Object data is immediately protected as specified in the ILM rule's placement instructions, which can be configured to protect against a wide variety of failure conditions, including the

failure of more than one storage location. Dual commit can only protect against the loss of a single local copy.

- **More efficient grid operation**: Each object is processed only once, as it is ingested. Because the StorageGRID system does not need to track or delete interim copies, there is less processing load and less database space is consumed.
- **(Balanced) Recommended**: The Balanced option provides optimal ILM efficiency. Using the Balanced option is recommended unless Strict ingest behavior is required or the grid meets all of the criteria for using for Dual commit.
- **(Strict) Certainty about object locations**: The Strict option guarantees that objects are immediately stored according to the placement instructions in the ILM rule.

Disadvantages of the Balanced and Strict options

When compared to Dual commit, the Balanced and Strict options have some disadvantages:

- **Longer client ingests**: Client ingest latencies might be longer. When you use the Balanced and Strict options, an “ingest successful” message is not returned to the client until all erasure-coded fragments or replicated copies are created and stored. However, object data will most likely reach its final placement much faster.
- **(Strict) Higher rates of ingest failure**: With the Strict option, ingest fails whenever StorageGRID cannot immediately make all copies specified in the ILM rule. You might see high rates of ingest failure if a required storage location is temporarily offline or if network issues cause delays in copying objects between sites.
- **(Strict) S3 multipart upload placements might not be as expected in some circumstances**: With Strict, you expect objects either to be placed as described by the ILM rule or for ingest to fail. However, with an S3 multipart upload, ILM is evaluated for each part of the object as it ingested, and for the object as a whole when the multipart upload completes. In the following circumstances this might result in placements that are different than you expect:
  - **If ILM changes while an S3 multipart upload is in progress**: Because each part is placed according to the rule that is active when the part is ingested, some parts of the object might not meet current ILM requirements when the multipart upload completes. In these cases, ingest of the object does not fail. Instead, any part that is not placed correctly is queued for ILM re-evaluation, and is moved to the correct location later.
  - **When ILM rules filter on size**: When evaluating ILM for a part, StorageGRID filters on the size of the part, not the size of the object. This means that parts of an object can be stored in locations that do not meet ILM requirements for the object as a whole. For example, if a rule specifies that all objects 10 GB or larger are stored at DC1 while all smaller objects are stored at DC2, at ingest each 1 GB part of a 10-part multipart upload is stored at DC2. When ILM is evaluated for the object, all parts of the object are moved to DC1.
- **(Strict) Ingest does not fail when object tags or metadata are updated and newly required placements cannot be made**: With Strict, you expect objects either to be placed as described by the ILM rule or for ingest to fail. However, when you update metadata or tags for an object that is already stored in the grid, the object is not re-ingested. This means that any changes to object placement that are triggered by the update are not made immediately. Placement changes are made when ILM is re-evaluated by normal background ILM processes. If required placement changes cannot be made (for example, because a newly required location is unavailable), the updated object retains its current placement until the placement changes are possible.

Limitations on object placements with the Balanced or Strict options

The Balanced or Strict options cannot be used for ILM rules that have any of these placement instructions:

- Placement in a Cloud Storage Pool at day 0.
- Placement in an Archive Node at day 0.
- Placements in a Cloud Storage Pool or an Archive Node when the rule has a User Defined Creation Time as its Reference Time.

These restrictions exist because StorageGRID cannot synchronously make copies to a Cloud Storage Pool or an Archive Node, and a User Defined Creation Time could resolve to the present.

How ILM rules and consistency controls interact to affect data protection

Both your ILM rule and your choice of consistency control affect how objects are protected. These settings can interact.

For example, the ingest behavior selected for an ILM rule affects the initial placement of object copies, while the consistency control used when an object is stored affects the initial placement of object metadata. Because StorageGRID requires access to both an object's metadata and its data to fulfill client requests, selecting matching levels of protection for the consistency level and ingest behavior can provide better initial data protection and more predictable system responses.

Here is a brief summary of the consistency controls that are available in StorageGRID:

- **all**: All nodes receive object metadata immediately or the request will fail.
- **strong-global**: Object metadata is immediately distributed to all sites. Guarantees read-after-write consistency for all client requests across all sites.
- **strong-site**: Object metadata is immediately distributed to other nodes at the site. Guarantees read-after-write consistency for all client requests within a site.
- **read-after-new-write**: Provides read-after-write consistency for new objects and eventual consistency for object updates. Offers high availability and data protection guarantees.
- **available** (eventual consistency for HEAD operations): Behaves the same as the "read-after-new-write" consistency level, but only provides eventual consistency for HEAD operations.

> Before selecting a consistency level, read the full description of these settings in the instructions for creating an S3 or Swift client application. You should understand the benefits and limitations before changing the default value.

Example of how the consistency control and ILM rule can interact

Suppose you have a two-site grid with the following ILM rule and the following consistency level setting:

- **ILM rule**: Create two object copies, one at the local site and one at a remote site. The Strict ingest behavior is selected.
- **Consistency level**: "strong-global" (Object metadata is immediately distributed to all sites.)

When a client stores an object to the grid, StorageGRID makes both object copies and distributes metadata to both sites before returning success to the client.

The object is fully protected against loss at the time of the ingest successful message. For example, if the local

site is lost shortly after ingest, copies of both the object data and the object metadata still exist at the remote site. The object is fully retrievable.

If you instead used the same ILM rule and the “strong-site” consistency level, the client might receive a success message after object data is replicated to the remote site but before object metadata is distributed there. In this case, the level of protection of object metadata does not match the level of protection for object data. If the local site is lost shortly after ingest, object metadata is lost. The object cannot be retrieved.

The inter-relationship between consistency levels and ILM rules can be complex. Contact NetApp if you require assistance.

Related information

What replication is

What erasure coding is

What erasure-coding schemes are

Example 5: ILM rules and policy for Strict ingest behavior

Use S3

Use Swift

How objects are stored (replication or erasure coding)

StorageGRID can protect objects against loss either by storing replicated copies or by storing erasure-coded copies. You specify the type of copies to create in the placement instructions of ILM rules.

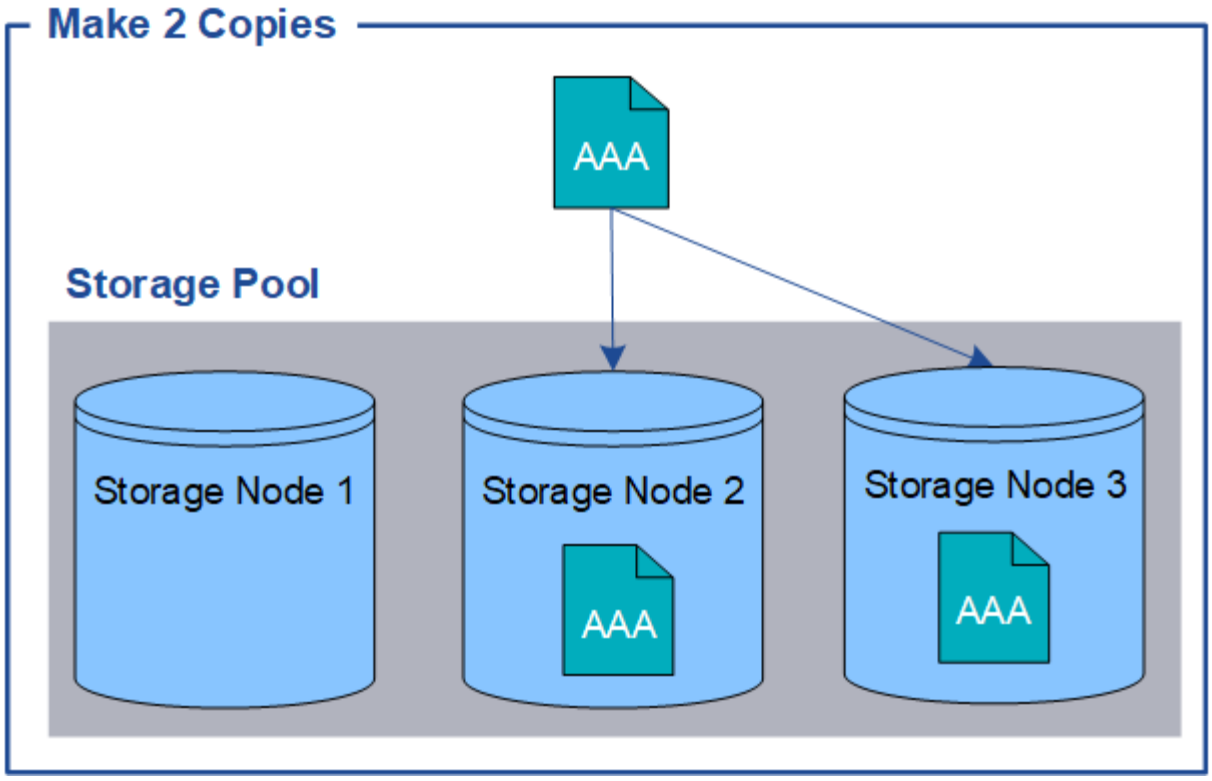
- What replication is
- Why you should not use single-copy replication
- What erasure coding is
- What erasure-coding schemes are
- Advantages, disadvantages, and requirements for erasure coding

What replication is

Replication is one of two methods used by StorageGRID to store object data. When objects match an ILM rule that uses replication, the system creates exact copies of object data and stores the copies on Storage Nodes or Archive Nodes.

When you configure an ILM rule to create replicated copies, you specify how many copies should be created, where those copies should be placed, and how long the copies should be stored at each location.

In the following example, the ILM rule specifies that two replicated copies of each object be placed in a storage pool that contains three Storage Nodes.



When StorageGRID matches objects to this rule, it creates two copies of the object, placing each copy on a different Storage Node in the storage pool. The two copies might be placed on any two of the three available Storage Nodes. In this case, the rule placed object copies on Storage Nodes 2 and 3. Because there are two copies, the object can be retrieved if any of the nodes in the storage pool fails.

> StorageGRID can store only one replicated copy of an object on any given Storage Node. If your grid includes three Storage Nodes and you create a 4-copy ILM rule, only three copies will be made—one copy for each Storage Node. The **ILM placement unachievable** alert is triggered to indicate that the ILM rule could not be completely applied.

Related information

What a storage pool is

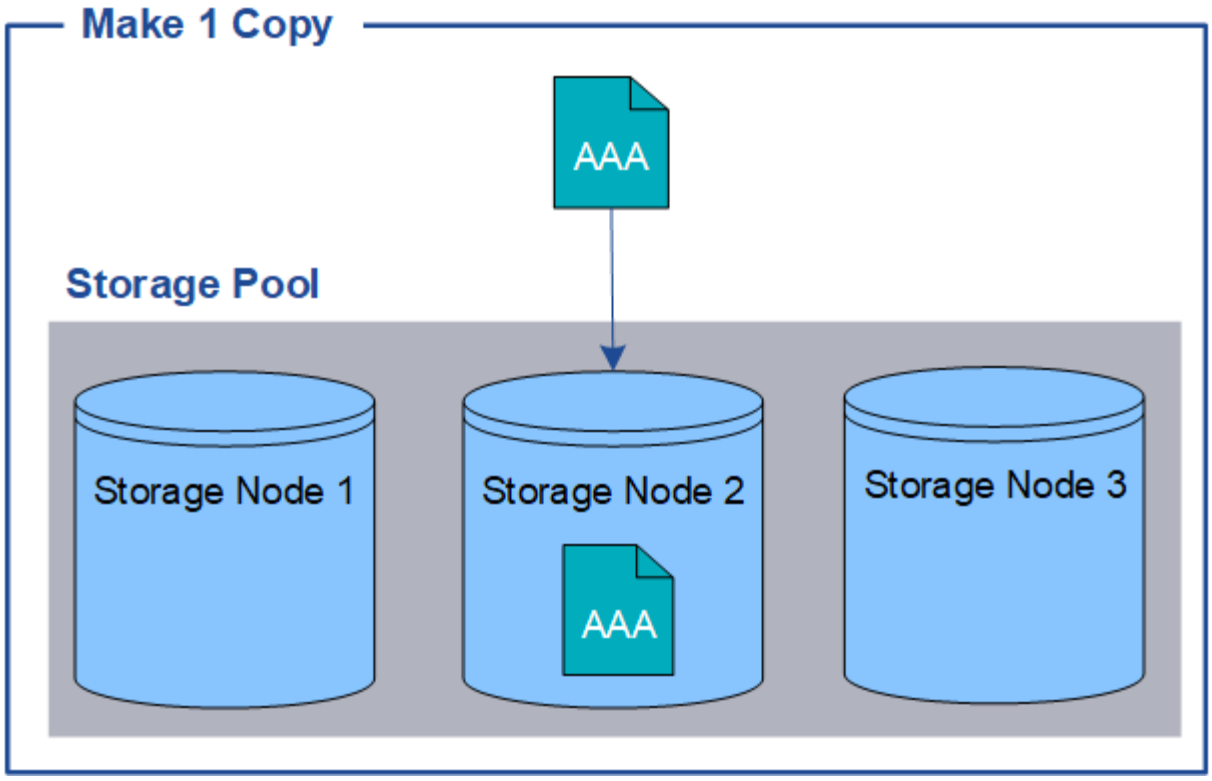
Using multiple storage pools for cross-site replication

Why you should not use single-copy replication

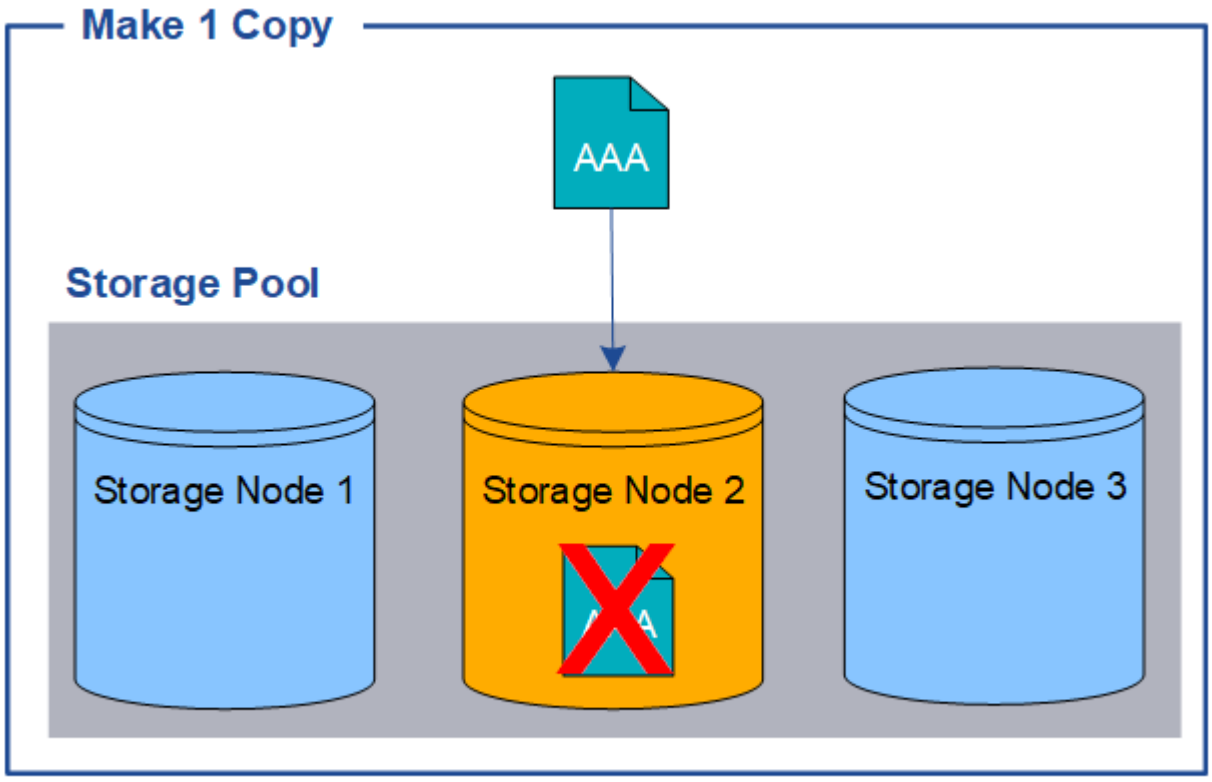
When creating an ILM rule to create replicated copies, you should always specify at least two copies for any time period in the placement instructions.

> Do not use an ILM rule that creates only one replicated copy for any time period. If only one replicated copy of an object exists, that object is lost if a Storage Node fails or has a significant error. You also temporarily lose access to the object during maintenance procedures such as upgrades.

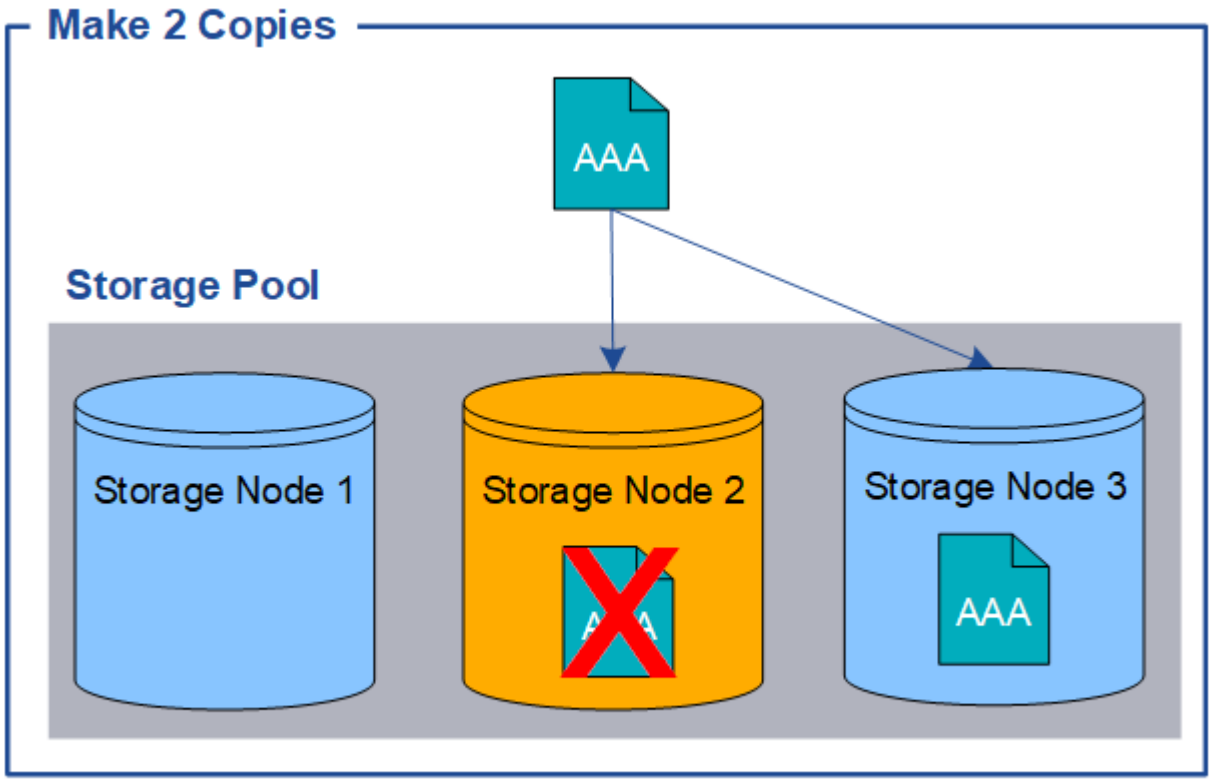
In the following example, the Make 1 Copy ILM rule specifies that one replicated copy of an object be placed in a storage pool that contains three Storage Nodes. When an object is ingested that matches this rule, StorageGRID places a single copy on only one Storage Node.



When an ILM rule creates only one replicated copy of an object, the object becomes inaccessible when the Storage Node is unavailable. In this example, you will temporarily lose access to object AAA whenever Storage Node 2 is offline, such as during an upgrade or other maintenance procedure. You will lose object AAA entirely if Storage Node 2 fails.



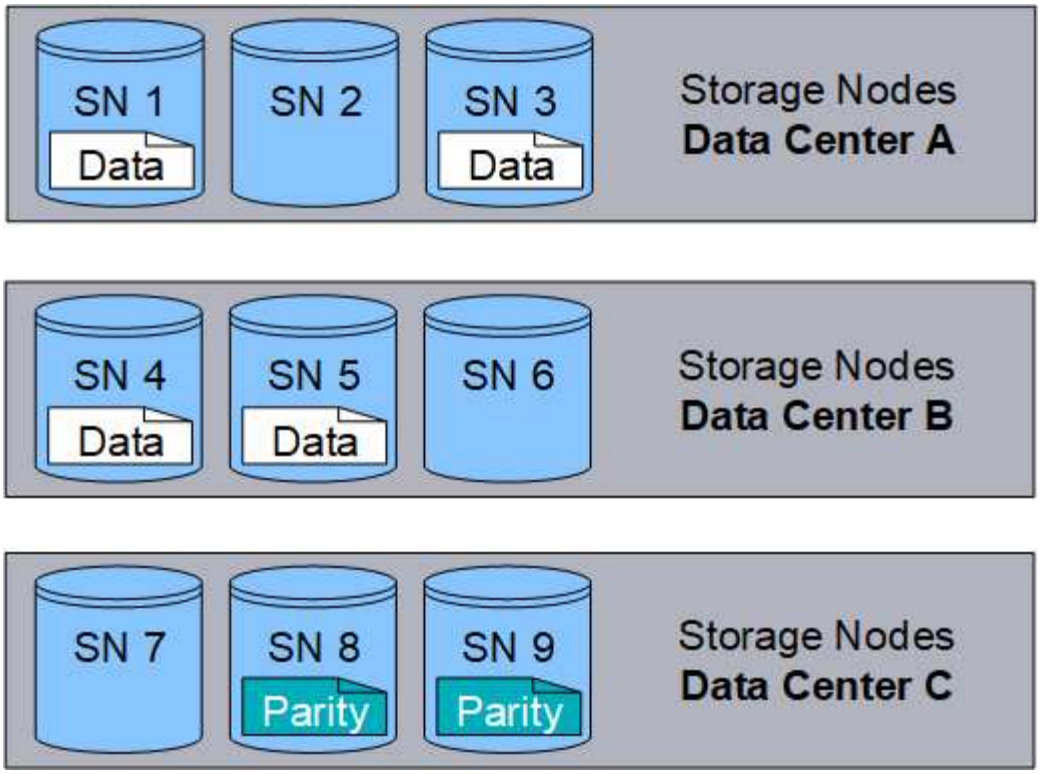
To avoid losing object data, you should always make at least two copies of all objects you want to protect with replication. If two or more copies exist, you can still access the object if one Storage Node fails or goes offline.



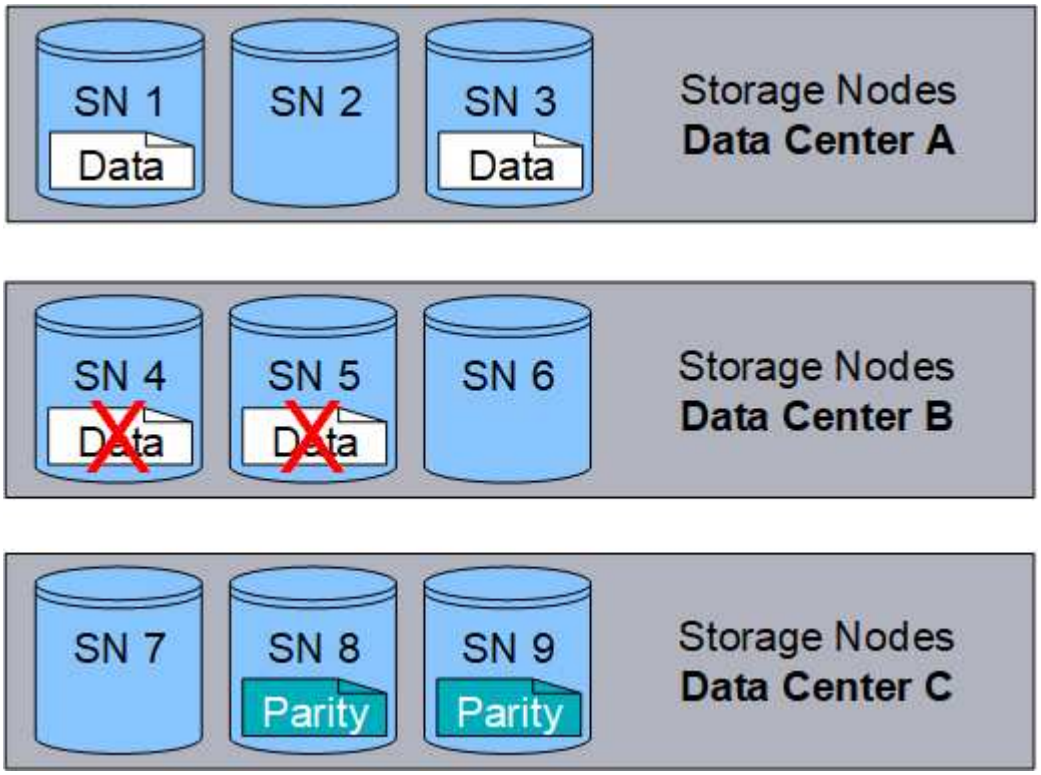
What erasure coding is

Erasure coding is the second method used by StorageGRID to store object data. When StorageGRID matches objects to an ILM rule that is configured to create erasure-coded copies, it slices object data into data fragments, computes additional parity fragments, and stores each fragment on a different Storage Node. When an object is accessed, it is reassembled using the stored fragments. If a data or a parity fragment becomes corrupt or lost, the erasure-coding algorithm can recreate that fragment using a subset of the remaining data and parity fragments.

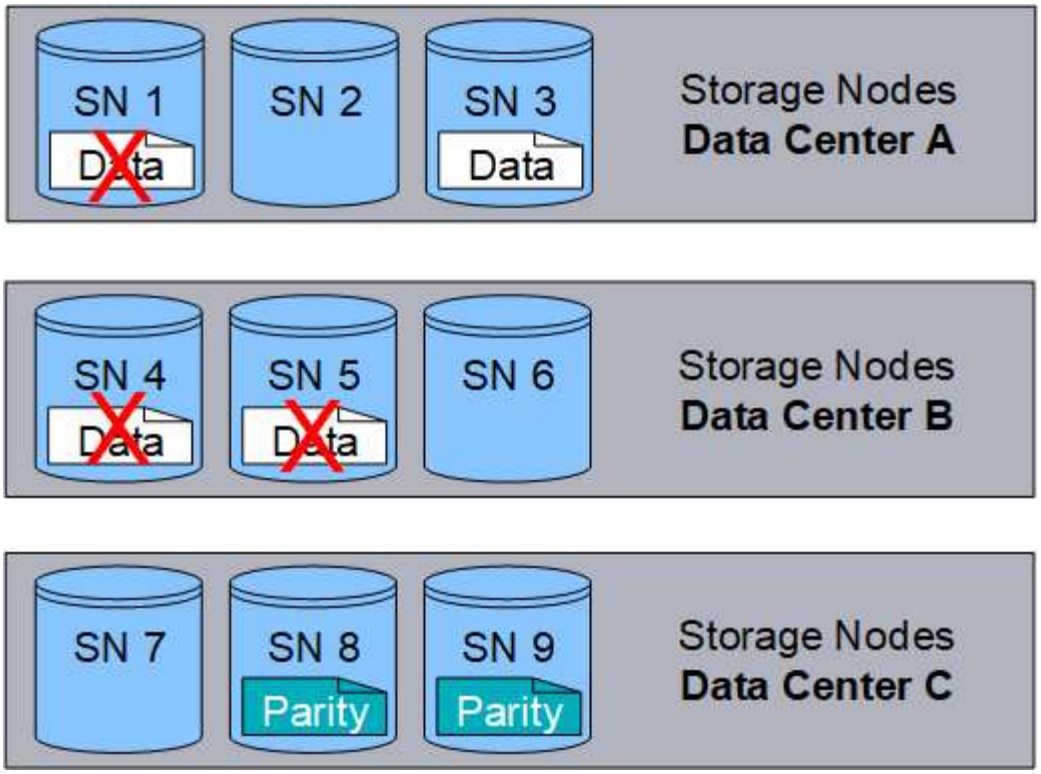
The following example illustrates the use of an erasure-coding algorithm on an object's data. In this example, the ILM rule uses a 4+2 erasure-coding scheme. Each object is sliced into four equal data fragments, and two parity fragments are computed from the object data. Each of the six fragments is stored on a different node across three data center sites to provide data protection for node failures or site loss.



The 4+2 erasure-coding scheme requires a minimum of nine Storage Nodes, with three Storage Nodes at each of three different sites. An object can be retrieved as long as any four of the six fragments (data or parity) remain available. Up to two fragments can be lost without loss of the object data. If an entire data center site is lost, the object can still be retrieved or repaired, as long as all of the other fragments remain accessible.



If more than two Storage Nodes are lost, the object is not retrievable.



Related information

What a storage pool is

What erasure-coding schemes are

Configuring Erasure Coding profiles

What erasure-coding schemes are

When you configure the Erasure Coding profile for an ILM rule, you select an available erasure-coding scheme based on how many Storage Nodes and sites make up the storage pool you plan to use. Erasure-coding schemes control how many data fragments and how many parity fragments are created for each object.

The StorageGRID system uses the Reed-Solomon erasure-coding algorithm. The algorithm slices an object into k data fragments and computes m parity fragments. The k + m = n fragments are spread across n Storage Nodes to provide data protection. An object can sustain up to m lost or corrupt fragments. k fragments are needed to retrieve or repair an object.

When configuring an Erasure Coding profile, use the following guidelines for storage pools:

- The storage pool must include three or more sites, or exactly one site.

  You cannot configure an Erasure Coding profile if the storage pool includes two sites.

  - Erasure-coding schemes for storage pools containing three or more sites
  - Erasure-coding schemes for one-site storage pools
- Do not use the default storage pool, All Storage Nodes, or a storage pool that includes the default site, All Sites.

- The storage pool should include at least *k*+*m* +1 Storage Nodes.

  The minimum number of Storage Nodes required is *k*+*m*. However, having at least one additional Storage Node can help prevent ingest failures or ILM backlogs if a required Storage Node is temporarily unavailable.

The storage overhead of an erasure-coding scheme is calculated by dividing the number of parity fragments (*m*) by the number of data fragments (*k*). You can use the storage overhead to calculate how much disk space each erasure-coded object requires:

```
disk space = object size + (object size * storage overhead)
```

For example, if you store a 10 MB object using the 4+2 scheme (which has 50% storage overhead), the object consumes 15 MB of grid storage. If you store the same 10 MB object using the 6+2 scheme (which has 33% storage overhead), the object consumes approximately 13.3 MB.

Select the erasure-coding scheme with the lowest total value of *k*+*m* that meets your needs. erasure-coding schemes with a lower number of fragments are overall more computationally efficient, as fewer fragments are created and distributed (or retrieved) per object, can show better performance due to the larger fragment size, and can require fewer nodes be added in an expansion when more storage is required. (See the instructions for expanding StorageGRID for information on planning a storage expansion.)

Erasure-coding schemes for storage pools containing three or more sites

The following table describes the erasure-coding schemes currently supported by StorageGRID for storage pools that include three or more sites. All of these schemes provide site loss protection. One site can be lost, and the object will still be accessible.

For erasure-coding schemes that provide site loss protection, the recommended number of Storage Nodes in the storage pool exceeds *k*+*m*+1 because each site requires a minimum of three Storage Nodes.

Erasure-coding scheme (*k*+*m*)	Minimum number of deployed sites	Recommended number of Storage Nodes at each site	Total recommended number of Storage Nodes	Site loss protection?	Storage overhead
4+2	3	3	9	Yes	50%
6+2	4	3	12	Yes	33%
8+2	5	3	15	Yes	25%
6+3	3	4	12	Yes	50%
9+3	4	4	16	Yes	33%
2+1	3	3	9	Yes	50%
4+1	5	3	15	Yes	25%

Erasure-coding scheme (*k*+*m*)	Minimum number of deployed sites	Recommended number of Storage Nodes at each site	Total recommended number of Storage Nodes	Site loss protection?	Storage overhead
6+1	7	3	21	Yes	17%
7+5	3	5	15	Yes	71%

StorageGRID requires a minimum of three Storage Nodes per site. To use the 7+5 scheme, each site requires a minimum of four Storage Nodes. Using five Storage Nodes per site is recommended.

When selecting an erasure-coding scheme that provides site protection, balance the relative importance of the following factors:

- **Number of fragments**: Performance and expansion flexibility are generally better when the total number of fragments is lower.
- **Fault tolerance**: Fault tolerance is increased by having more parity segments (that is, when *m* has a higher value.)
- **Network traffic**: When recovering from failures, using a scheme with more fragments (that is, a higher total for *k*+*m*) creates more network traffic.
- **Storage overhead**: Schemes with higher overhead require more storage space per object.

For example, when deciding between a 4+2 scheme and 6+3 scheme (which both have 50% storage overhead), select the 6+3 scheme if additional fault tolerance is required. Select the 4+2 scheme if network resources are constrained. If all other factors are equal, select 4+2 because it has a lower total number of fragments.

If you are unsure of which scheme to use, select 4+2 or 6+3, or contact technical support.

Erasure-coding schemes for one-site storage pools

A one-site storage pool supports all of the erasure-coding schemes defined for three or more sites, provided that the site has enough Storage Nodes.

The minimum number of Storage Nodes required is *k*+*m*, but a storage pool with *k*+*m*+1 Storage Nodes is recommended. For example, the 2+1 erasure-coding scheme requires a storage pool with a minimum of three Storage Nodes, but four Storage Nodes is recommended.

Erasure-coding scheme (*k*+*m*)	Minimum number of Storage Nodes	Recommended number of Storage Nodes	Storage overhead
4+2	6	7	50%
6+2	8	9	33%
8+2	10	11	25%

Erasure-coding scheme (*k*+*m*)	Minimum number of Storage Nodes	Recommended number of Storage Nodes	Storage overhead
6+3	9	10	50%
9+3	12	13	33%
2+1	3	4	50%
4+1	5	6	25%
6+1	7	8	17%
7+5	12	13	71%

Related information

Expand your grid

Advantages, disadvantages, and requirements for erasure coding

Before deciding whether to use replication or erasure coding to protect object data from loss, you should understand the advantages, disadvantages, and the requirements for erasure coding.

Advantages of erasure coding

When compared to replication, erasure coding offers improved reliability, availability, and storage efficiency.

- **Reliability**: Reliability is gauged in terms of fault tolerance—that is, the number of simultaneous failures that can be sustained without loss of data. With replication, multiple identical copies are stored on different nodes and across sites. With erasure coding, an object is encoded into data and parity fragments and distributed across many nodes and sites. This dispersal provides both site and node failure protection. When compared to replication, erasure coding provides improved reliability at comparable storage costs.
- **Availability**: Availability can be defined as the ability to retrieve objects if Storage Nodes fail or become inaccessible. When compared to replication, erasure coding provides increased availability at comparable storage costs.
- **Storage efficiency**: For similar levels of availability and reliability, objects protected through erasure coding consume less disk space than the same objects would if protected through replication. For example, a 10 MB object that is replicated to two sites consumes 20 MB of disk space (two copies), while an object that is erasure coded across three sites with a 6+3 erasure-coding scheme only consumes 15 MB of disk space.

> Disk space for erasure-coded objects is calculated as the object size plus the storage overhead. The storage overhead percentage is the number of parity fragments divided by the number of data fragments.

Disadvantages of erasure coding

When compared to replication, erasure coding has the following disadvantages:

- An increased number of Storage Nodes and sites is required. For example, if you use an erasure-coding scheme of 6+3, you must have at least three Storage Nodes at three different sites. In contrast, if you simply replicate object data, you require only one Storage Node for each copy.
- Increased cost and complexity of storage expansions. To expand a deployment that uses replication, you simply add storage capacity in every location where object copies are made. To expand a deployment that uses erasure coding, you must consider both the erasure-coding scheme in use and how full existing Storage Nodes are. For example, if you wait until existing nodes are 100% full, you must add at least *k+m* Storage Nodes, but if you expand when existing nodes are 70% full, you can add two nodes per site and still maximize usable storage capacity. For more information, see the instructions for expanding StorageGRID.
- There are increased retrieval latencies when you use erasure coding across geographically distributed sites. The object fragments for an object that is erasure coded and distributed across remote sites take longer to retrieve over WAN connections than an object that is replicated and available locally (the same site to which the client connects).
- When you use erasure coding across geographically distributed sites, there is higher WAN network traffic usage for retrievals and repairs, especially for frequently retrieved objects or for object repairs over WAN network connections.
- When you use erasure coding across sites, the maximum object throughput declines sharply as network latency between sites increases. This decrease is due to the corresponding decrease in TCP network throughput, which affects how quickly the StorageGRID system can store and retrieve object fragments.
- Higher usage of compute resources.

When to use erasure coding

Erasure coding is best suited for the following requirements:

- Objects larger than 1 MB in size.

  > Due to the overhead of managing the number of fragments associated with an erasure-coded copy, do not use erasure coding for objects 200 KB or smaller.

- Long-term or cold storage for infrequently retrieved content.
- High data availability and reliability.
- Protection against complete site and node failures.
- Storage efficiency.
- Single-site deployments that require efficient data protection with only a single erasure-coded copy rather than multiple replicated copies.
- Multiple-site deployments where the inter-site latency is less than 100 ms.

Related information

Expand your grid

How object retention is determined

StorageGRID provides options for both grid administrators and individual tenant users to specify how long to store objects. In general, any retention instructions provided by a tenant user take precedence over the retention instructions provided by the grid administrator.

How tenant users control object retention

Tenant users have three primary ways to control how long their objects are stored in StorageGRID:

- If the global S3 Object Lock setting is enabled for the grid, S3 tenant users can create buckets with S3 Object Lock enabled and then use the S3 REST API to specify retain-until-date and legal hold settings for each object version added to that bucket.
  - An object version that is under a legal hold cannot be deleted by any method.
  - Before an object version's retain-until-date is reached, that version cannot be deleted by any method.
  - Objects in buckets with S3 Object Lock enabled are retained by ILM "forever." However, after its retain-until-date is reached, an object version can be deleted by a client request or the expiration of the bucket lifecycle.

    Managing objects with S3 Object Lock
- S3 tenant users can add a lifecycle configuration to their buckets that specifies an Expiration action. If a bucket lifecycle exists, StorageGRID stores an object until the date or number of days specified in the Expiration action are met, unless the client deletes the object first.
- An S3 or Swift client can issue a delete object request. StorageGRID always prioritizes client delete requests over S3 bucket lifecycle or ILM when determining whether to delete or retain an object.

How grid administrators control object retention

Grid administrators use ILM placement instructions to control how long objects are stored. When objects are matched by an ILM rule, StorageGRID stores those objects until the last time period in the ILM rule has elapsed. Objects are retained indefinitely if "forever" is specified for the placement instructions.

Regardless of who controls how long objects are retained, ILM settings control what types of object copies (replicated or erasure coded) are stored and where the copies are located (Storage Nodes, Cloud Storage Pools, or Archive Nodes).

How S3 bucket lifecycle and ILM interact

The Expiration action in an S3 bucket lifecycle always overrides ILM settings. As a result, an object might be retained on the grid even after any ILM instructions for placing the object have lapsed.

Examples for object retention

To better understand the interactions between S3 Object Lock, bucket lifecycle settings, client delete requests, and ILM, consider the following examples.

Example 1: S3 bucket lifecycle keeps objects longer than ILM

ILM

Store two copies for 1 year (365 days)

Bucket lifecycle

Expire objects in 2 years (730 days)

Result

StorageGRID stores the object for 730 days. StorageGRID uses the bucket lifecycle settings to determine whether to delete or retain an object.

> If the bucket lifecycle specifies that objects should be kept longer than specified by ILM, StorageGRID continues to use the ILM placement instructions when determining the number and type of copies to store. In this example, two copies of the object will continue to be stored in StorageGRID from days 366 to 730.

Example 2: S3 bucket lifecycle expires objects before ILM

ILM

Store two copies for 2 years (730 days)

Bucket lifecycle

Expire objects in 1 year (365 days)

Result

StorageGRID deletes both copies of the object after day 365.

Example 3: Client delete overrides bucket lifecycle and ILM

ILM

Store two copies on Storage Nodes “forever”

Bucket lifecycle

Expire objects in 2 years (730 days)

Client delete request

Issued on day 400

Result

StorageGRID deletes both copies of the object on day 400 in response to the client delete request.

Example 4: S3 Object Lock overrides client delete request

S3 Object Lock

Retain-until-date for an object version is 2026-03-31. A legal hold is not in effect.

Compliant ILM rule

Store two copies on Storage Nodes “forever.”

Client delete request

Issued on 2024-03-31.

Result

StorageGRID will not delete the object version because the retain-until-date is still 2 years away.

Related information

Managing objects with S3 Object Lock

Use S3

What ILM rule placement instructions are

How objects are deleted

StorageGRID can delete objects either in direct response to a client request or automatically as a result of the expiration of an S3 bucket lifecycle or the requirements of the ILM policy. Understanding the different ways that objects can be deleted and how StorageGRID handles delete requests can help you manage objects more effectively.

StorageGRID can use one of two methods to delete objects:

- Synchronous deletion: When StorageGRID receives a client delete request, all object copies are removed immediately. The client is informed that deletion was successful after the copies have been removed.
- Objects are queued for deletion: When StorageGRID receives a delete request, the object is queued for deletion and the client is informed immediately that deletion was successful. Object copies are removed later by background ILM processing.

When deleting objects, StorageGRID uses the method that optimizes delete performance, minimizes potential delete backlogs, and frees space most quickly.

The table summarizes when StorageGRID uses each method.

Method of performing deletion	When used
Objects are queued for deletion	When **any** of the following conditions are true: • Automatic object deletion has been triggered by one of the following events: ◦ The expiration date or number of days in the lifecycle configuration for an S3 bucket is reached. ◦ The last time period specified in an ILM rule elapses. **Note:** Objects in a bucket that has S3 Object Lock enabled cannot be deleted if they are under a legal hold or if a retain-until-date has been specified but not yet met. • An S3 or Swift client requests deletion and one or more of these conditions is true: ◦ Copies cannot be deleted within 30 seconds because, for example, an object location is temporarily unavailable. ◦ Background deletion queues are idle.
Objects are removed immediately (synchronous deletion)	When an S3 or Swift client makes a delete request and **all** of the following conditions are met: • All copies can be removed within 30 seconds. • Background deletion queues contain objects to process.

When S3 or Swift clients make delete requests, StorageGRID begins by adding a number of objects to the delete queue. It then switches to performing synchronous deletion. Making sure that the background deletion queue has objects to process allows StorageGRID to process deletes more efficiently, especially for low concurrency clients, while helping to prevent client delete backlogs.

Understanding the impact of how StorageGRID deletes objects

The way that StorageGRID deletes objects can affect how the system appears to perform:

- When StorageGRID performs synchronous deletion, it can take StorageGRID up to 30 seconds to return a result to the client. This means that deletion can appear to be happening more slowly, even though copies are actually being removed more quickly than they are when StorageGRID queues objects for deletion.
- If you are closely monitoring delete performance during a bulk delete, you might notice that the deletion rate appears to slow after a certain number of objects have been deleted. This change occurs when StorageGRID shifts from queuing objects for deletion to performing synchronous deletion. The apparent reduction in the deletion rate does not mean that object copies are being removed more slowly. On the contrary, it indicates that on average, space is now being freed more quickly.

If you are deleting large numbers of objects and your priority is to free space quickly, consider using a client request to delete objects rather than deleting them using ILM or other methods. In general, space is freed more quickly when deletion is performed by clients because StorageGRID can use synchronous deletion.

You should be aware that the amount of time required to free space after an object is deleted depends on a number of factors:

- Whether object copies are synchronously removed or are queued for removal later (for client delete requests).
- Other factors such as the number of objects in the grid or the availability of grid resources when object copies are queued for removal (for both client deletes and other methods).

How S3 versioned objects are deleted

When versioning is enabled for an S3 bucket, StorageGRID follows Amazon S3 behavior when responding to delete requests, whether those requests come from an S3 client, the expiration of an S3 bucket lifecycle, or the requirements of the ILM policy.

When objects are versioned, object delete requests do not delete the current version of the object and do not free space. Instead, an object delete request simply creates a delete marker as the current version of the object, which makes the previous version of the object “noncurrent.”

Even though the object has not been removed, StorageGRID behaves as though the current version of the object is no longer available. Requests to that object return 404 NotFound. However, because noncurrent object data has not been removed, requests that specify a noncurrent version of the object can succeed.

To free space when deleting versioned objects, you must do one of the following:

- **S3 client request**: Specify the object version number in the S3 DELETE Object request (`DELETE /object?versionId=ID`). Keep in mind that this request only removes object copies for the specified version (the other versions are still taking up space).
- **Bucket lifecycle**: Use the `NoncurrentVersionExpiration` action in the bucket lifecycle configuration. When the number of NoncurrentDays specified is met, StorageGRID permanently removes all copies of noncurrent object versions. These object versions cannot be recovered.
- **ILM**: Add two ILM rules to your ILM policy. Use **Noncurrent Time** as the Reference Time in the first rule to match the noncurrent versions of the object. Use **Ingest Time** in the second rule to match the current version. The **Noncurrent Time** rule must appear in the policy above the **Ingest Time** rule.

Related information

Use S3

Example 4: ILM rules and policy for S3 versioned objects

Copyright information

Trademark information