



# **StorageGRID 如何管理数据**

## **StorageGRID 11.7**

NetApp  
April 12, 2024

# 目录

- StorageGRID 如何管理数据 ..... 1
  - 什么是对象 ..... 1
  - 对象的生命周期 ..... 3
  - 载入数据流 ..... 4
  - 副本管理 ..... 5
  - 检索数据流 ..... 7
  - 删除数据流 ..... 8
  - 使用信息生命周期管理 ..... 10

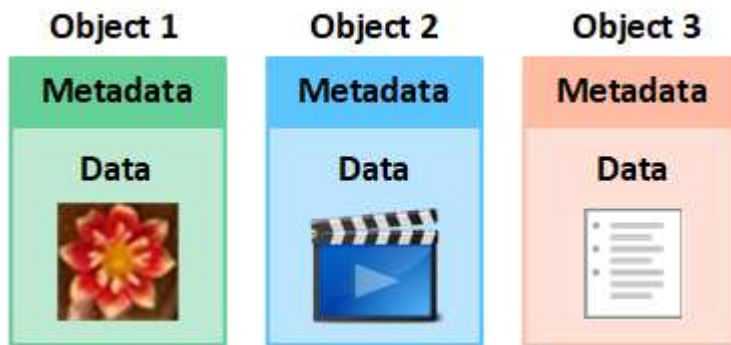
# StorageGRID 如何管理数据

## 什么是对象

对于对象存储，存储单元是对象，而不是文件或块。与文件系统或块存储的树状层次结构不同，对象存储以非结构化的平面布局对数据进行组织。

对象存储可将数据的物理位置与用于存储和检索数据的方法分离。

基于对象的存储系统中的每个对象都有两部分：对象数据和对象元数据。



## 什么是对象数据？

对象数据可以是任何内容；例如，照片，电影或病历。

## 什么是对象元数据？

对象元数据是指描述对象的任何信息。StorageGRID 使用对象元数据跟踪网格中所有对象的位置，并管理每个对象的生命周期。

对象元数据包括以下信息：

- 系统元数据，包括每个对象的唯一 ID（UUID），对象名称，S3 存储分段或 Swift 容器的名称，租户帐户名称或 ID，对象的逻辑大小，首次创建对象的日期和时间，以及上次修改对象的日期和时间。
- 每个对象副本或纠删编码片段的当前存储位置。
- 与对象关联的任何用户元数据。

对象元数据可自定义并可扩展，因此应用程序可以灵活地使用。

有关 StorageGRID 如何以及在何处存储对象元数据的详细信息，请转到 ["管理对象元数据存储"](#)。

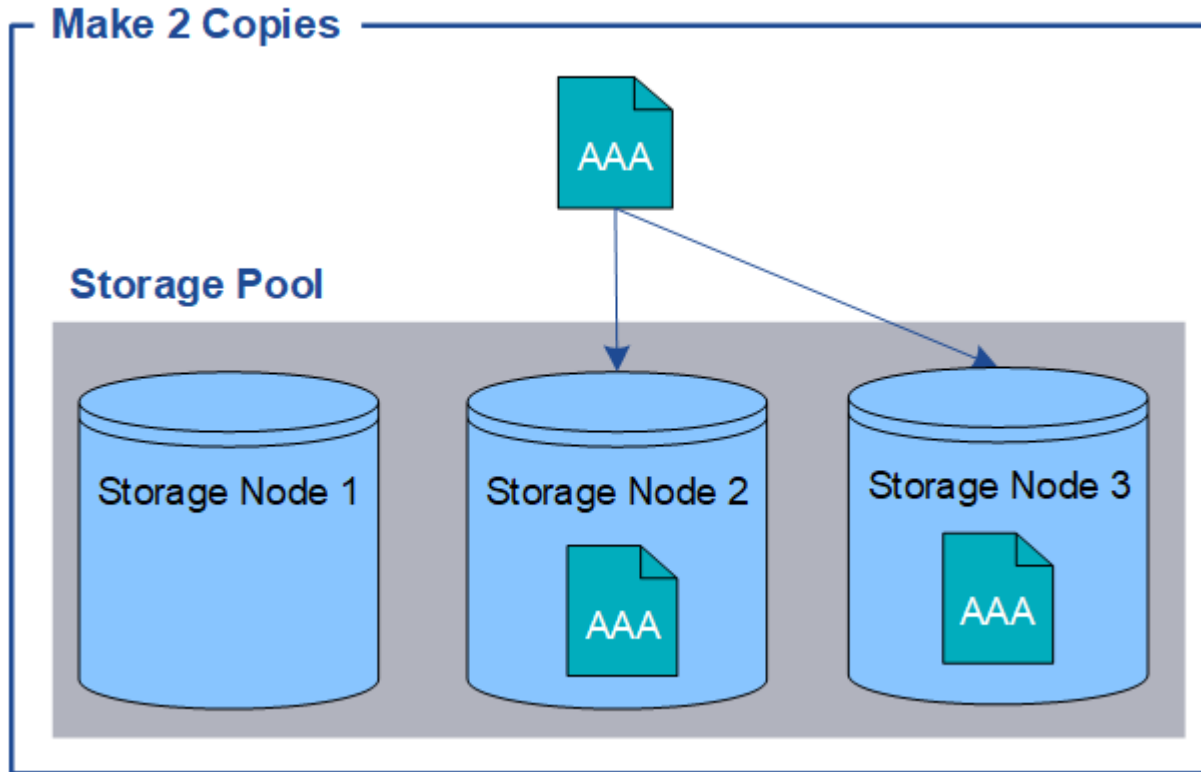
## 如何保护对象数据？

StorageGRID 系统提供了两种机制来防止对象数据丢失：复制和纠删编码。

## Replication

如果 StorageGRID 将对象与配置为创建复制副本的信息生命周期管理（ILM）规则匹配，则系统会创建对象数据的精确副本，并将其存储在存储节点，归档节点或云存储池中。ILM 规则规定了创建的副本数量，这些副本的存储位置以及系统保留这些副本的时间长度。例如，如果由于存储节点丢失而导致副本丢失，则如果 StorageGRID 系统中的其他位置存在该对象的副本，则该对象仍可用。

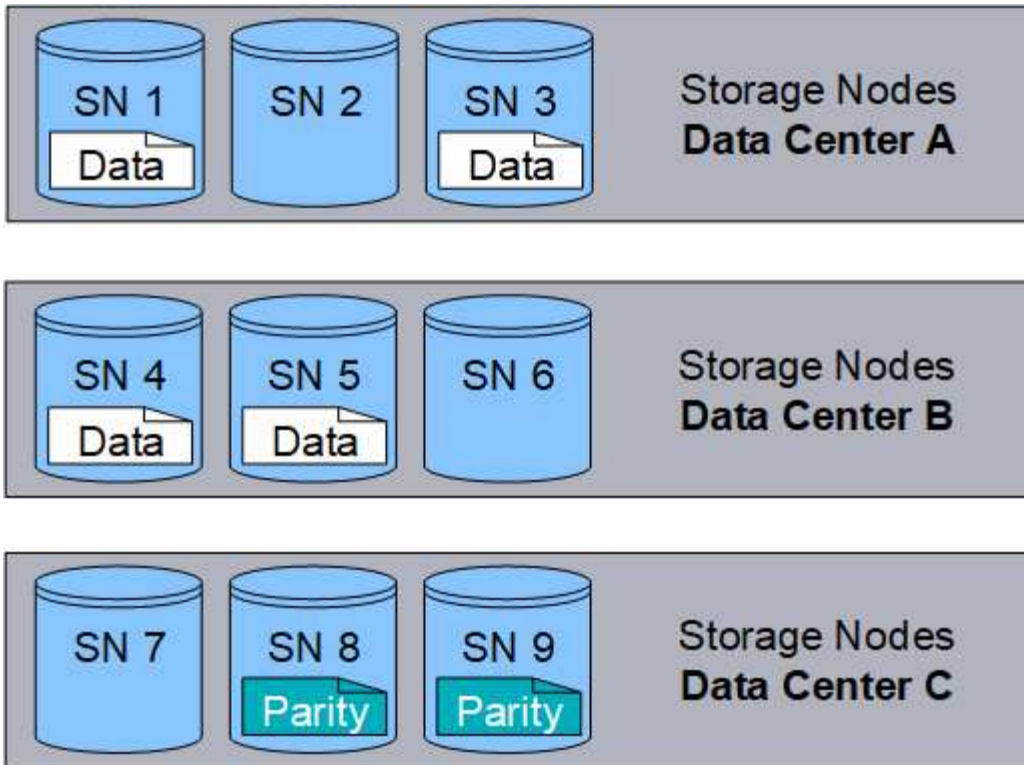
在以下示例中，make 2 copies 规则指定将每个对象的两个复制副本放置在包含三个存储节点的存储池中。



## 纠删编码

如果 StorageGRID 将对象与配置为创建纠删编码副本的 ILM 规则匹配，则会将对象数据分段为数据片段，计算额外的奇偶校验片段，并将每个片段存储在不同的存储节点上。访问某个对象时，系统会使用存储的片段重新组合该对象。如果数据或奇偶校验片段损坏或丢失，纠删编码算法可以使用剩余数据和奇偶校验片段的子集重新创建该片段。ILM 规则和纠删编码配置文件决定了所使用的纠删编码方案。

以下示例说明了如何对对象数据使用纠删编码。在此示例中，ILM 规则使用 4+2 纠删编码方案。每个对象都会被划分为四个相等的数据片段，并根据对象数据计算两个奇偶校验片段。六个片段中的每个片段都存储在三个数据中心的不同存储节点上，以便为节点故障或站点丢失提供数据保护。



#### 相关信息

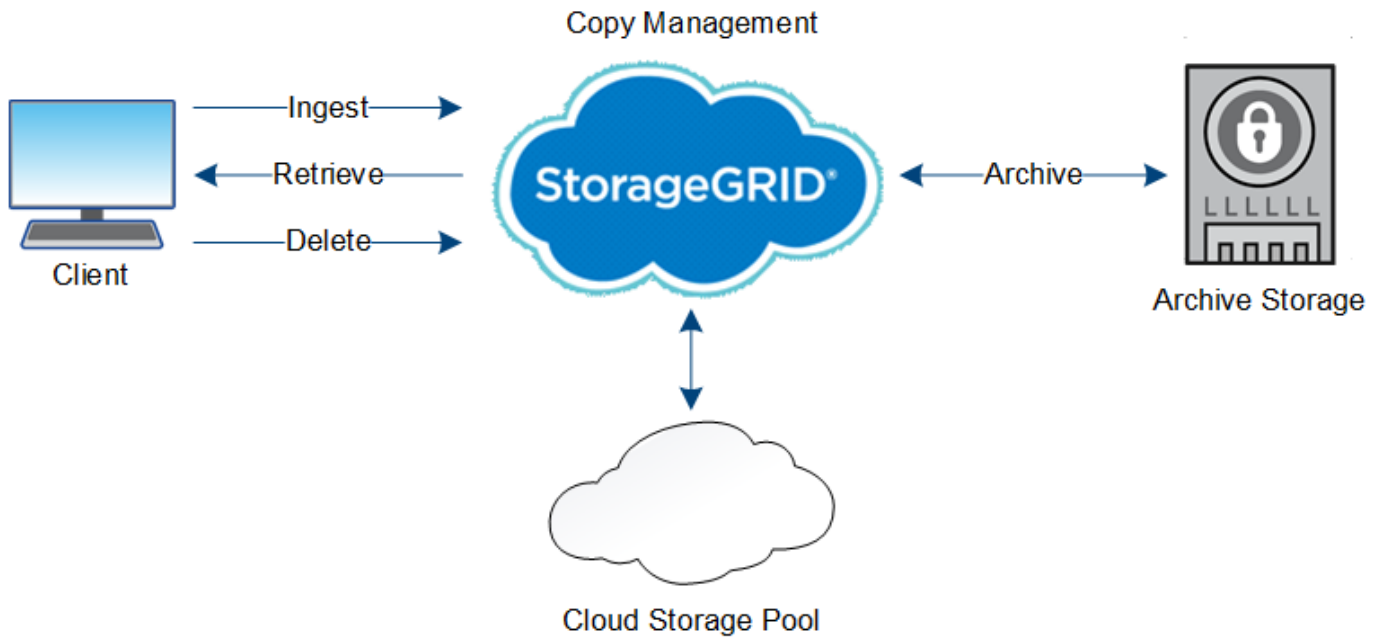
- ["使用 ILM 管理对象"](#)
- ["使用信息生命周期管理"](#)

## 对象的生命周期

对象的生命周期由多个阶段组成。每个阶段都表示对象发生的操作。

对象的生命周期包括载入，副本管理，检索和删除操作。

- **\* 载入 \***：S3 或 Swift 客户端应用程序通过 HTTP 将对象保存到 StorageGRID 系统的过程。在此阶段，StorageGRID 系统将开始管理此对象。
- **\* 副本管理 \***：在 StorageGRID 中管理复制的副本和经过纠删编码的副本的过程，如活动 ILM 策略中的 ILM 规则所述。在副本管理阶段，StorageGRID 通过在存储节点，云存储池或归档节点上创建和维护指定数量和类型的对象副本来保护对象数据不会丢失。
- **\* 检索 \***：客户端应用程序访问 StorageGRID 系统存储的对象的过程。客户端读取从存储节点，云存储池或归档节点检索到的对象。
- **\* 删除 \***：从网格中删除所有对象副本的过程。如果客户端应用程序向 StorageGRID 系统发送删除请求，或者由于 StorageGRID 在对象的生命周期到期时自动执行过程，则可以删除对象。



相关信息

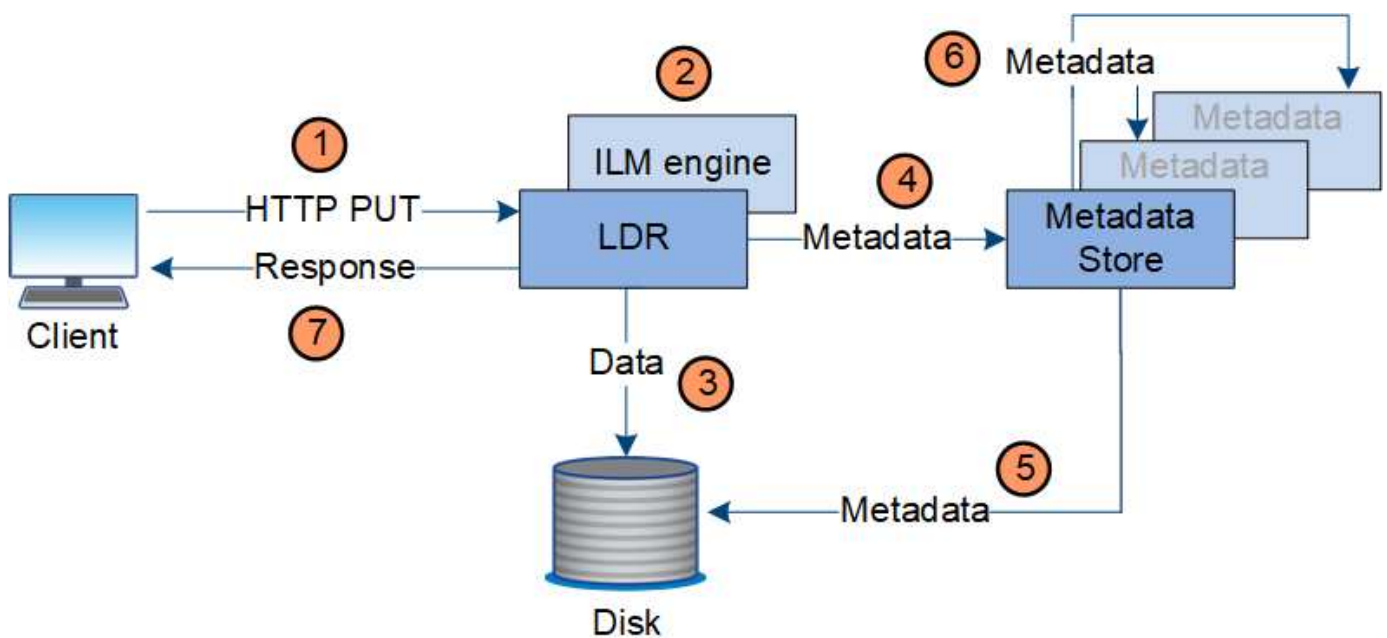
- ["使用 ILM 管理对象"](#)
- ["使用信息生命周期管理"](#)

## 载入数据流

载入或保存操作由客户端和 StorageGRID 系统之间定义的数据流组成。

### 数据流

当客户端将对象载入 StorageGRID 系统时，存储节点上的 LDR 服务将处理此请求并将元数据和数据存储到磁盘。



1. 客户端应用程序将创建此对象，并通过 HTTP PUT 请求将其发送到 StorageGRID 系统。
2. 将根据系统的 ILM 策略评估对象。
3. LDR 服务将对象数据保存为复制副本或纠删编码副本。（图中显示了将复制副本存储到磁盘的简化版本。）
4. LDR 服务将对象元数据发送到元数据存储。
5. 元数据存储将对象元数据保存到磁盘。
6. 元数据存储会将对象元数据的副本传播到其他存储节点。这些副本也会保存到磁盘中。
7. LDR 服务向客户端返回 HTTP 200 OK 响应，以确认已载入对象。

## 副本管理

对象数据由活动 ILM 策略及其 ILM 规则管理。ILM 规则可创建复制的或经过纠删编码的副本，以防止对象数据丢失。

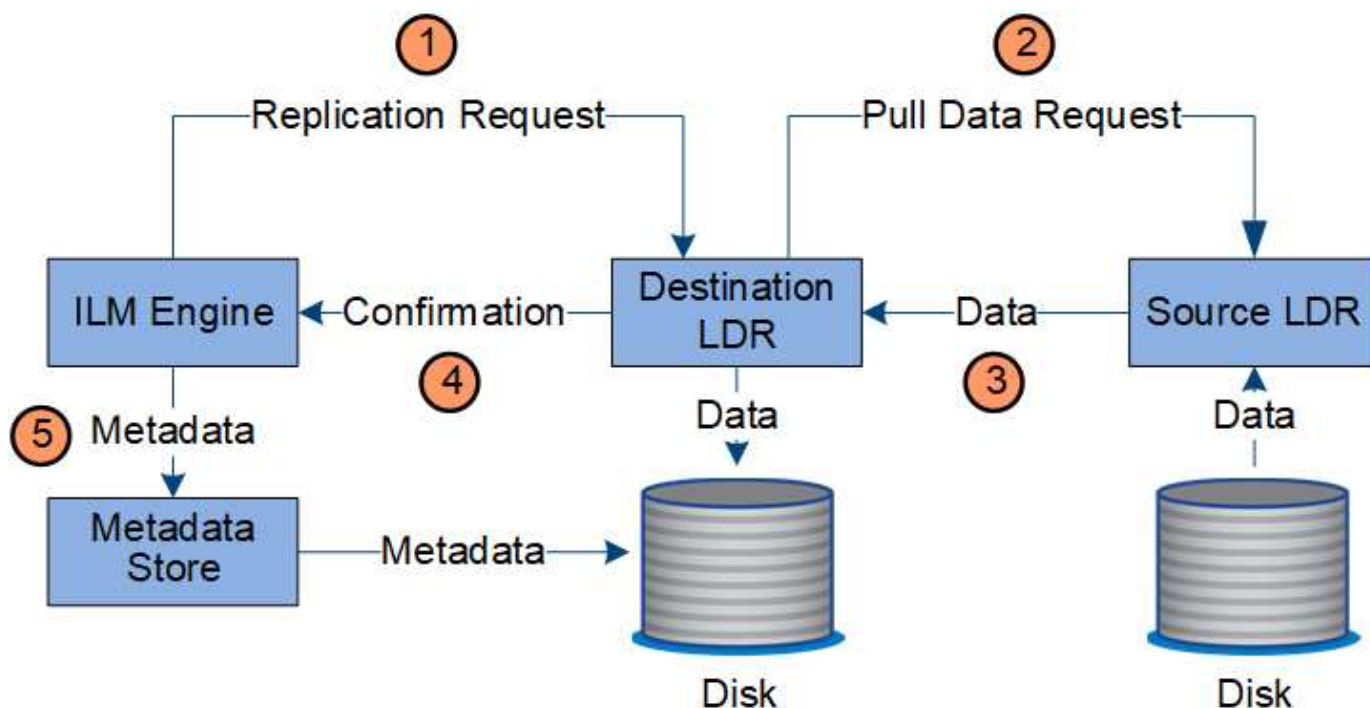
在对象生命周期的不同时间，可能需要不同类型或位置的对象副本。系统会定期评估 ILM 规则，以确保根据需要放置对象。

对象数据由 LDR 服务管理。

### 内容保护：复制

如果 ILM 规则的内容放置说明要求复制对象数据的副本，则构成已配置存储池的存储节点会创建副本并将其存储到磁盘中。

LDR 服务中的 ILM 引擎可控制复制，并确保将正确数量的副本存储在正确的位置和正确的时间内。



1. ILM 引擎会查询此 ADC-LDR 服务，以确定 ILM 规则指定的存储池中的最佳目标 LDR 服务。然后，它会向该 LDR 服务发送一个命令以启动复制。

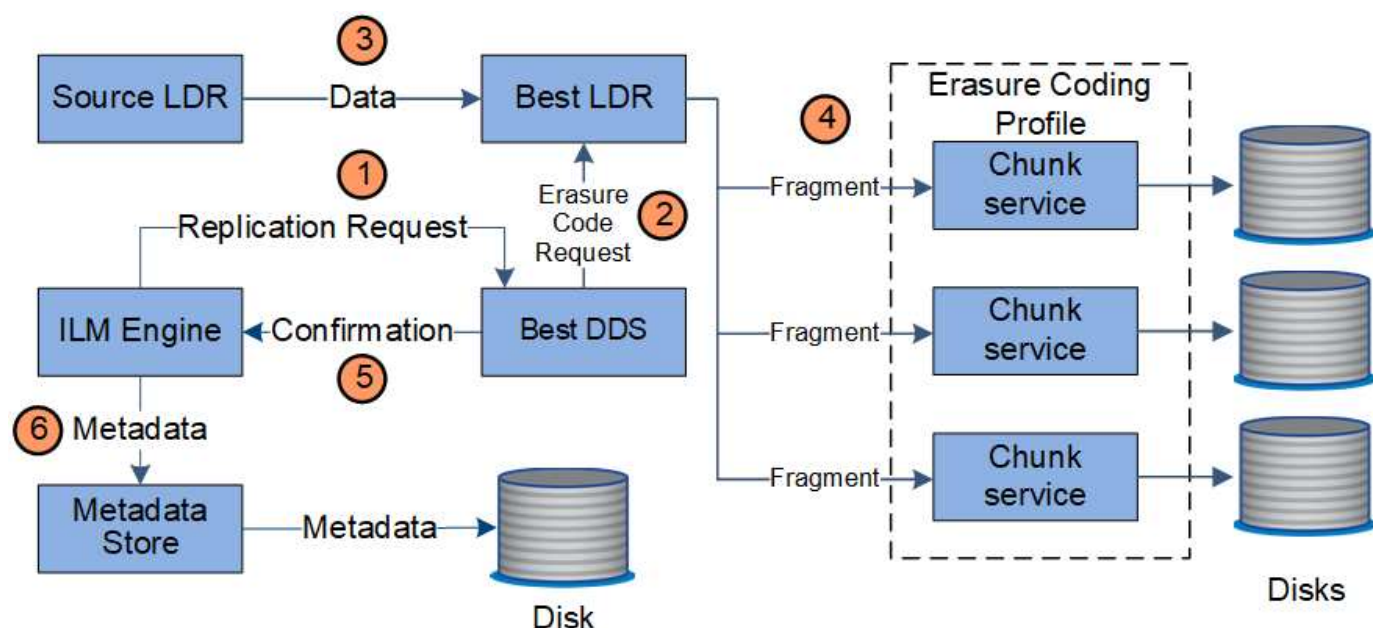


2. 目标 LDR 服务会向此 ADC-Service 查询最佳源位置。然后，它会向源 LDR 服务发送复制请求。
3. 源 LDR 服务会向目标 LDR 服务发送一份副本。
4. 目标 LDR 服务通知 ILM 引擎已存储对象数据。
5. ILM 引擎使用对象位置元数据更新元数据存储。

## 内容保护：纠删编码

如果 ILM 规则包含为对象数据创建纠删编码副本的说明、则适用的纠删编码方案会将对象数据拆分为数据和奇偶校验片段、并将这些片段分布在纠删编码配置文件中配置的存储节点上。

ILM 引擎是 LDR 服务的一个组件、用于控制纠删编码、并确保将纠删编码配置文件应用于对象数据。



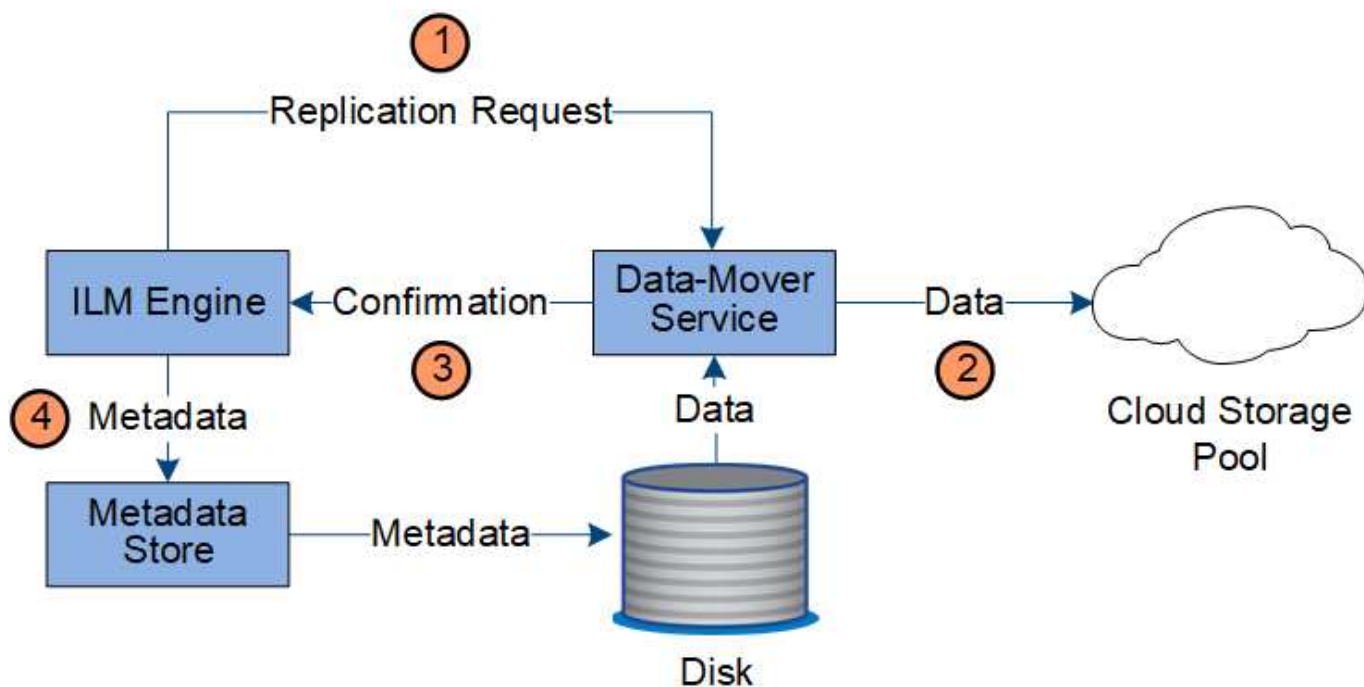
1. ILM 引擎会查询此 ADC-Service，以确定哪种 DDS 服务能够以最佳方式执行纠删编码操作。一旦确定，ILM 引擎就会向该服务发送 "启动" 请求。
2. DDS 服务指示 LDR 对对象数据进行纠删编码。
3. 源 LDR 服务会向选定用于纠删编码的 LDR 服务发送一份副本。
4. 一旦细分为适当数量的奇偶校验和数据片段、LDR 服务就会将这些片段分布在构成纠删编码配置文件存储池的存储节点(区块服务)之间。
5. LDR 服务通知 ILM 引擎，确认对象数据已成功分发。
6. ILM 引擎使用对象位置元数据更新元数据存储。

## 内容保护：云存储池

如果 ILM 规则的内容放置说明要求将对象数据的复制副本存储在云存储池中，则对象数据将复制到云存储池指定的外部 S3 存储分段或 Azure Blob 存储容器。

ILM 引擎是 LDR 服务的一个组件，Data Mover 服务可控制对象到云存储池的移动。





1. ILM 引擎选择要复制到云存储池的数据转换服务。
2. Data Mover 服务会将对象数据发送到云存储池。
3. Data Mover 服务会通知 ILM 引擎已存储对象数据。
4. ILM 引擎使用对象位置元数据更新元数据存储。

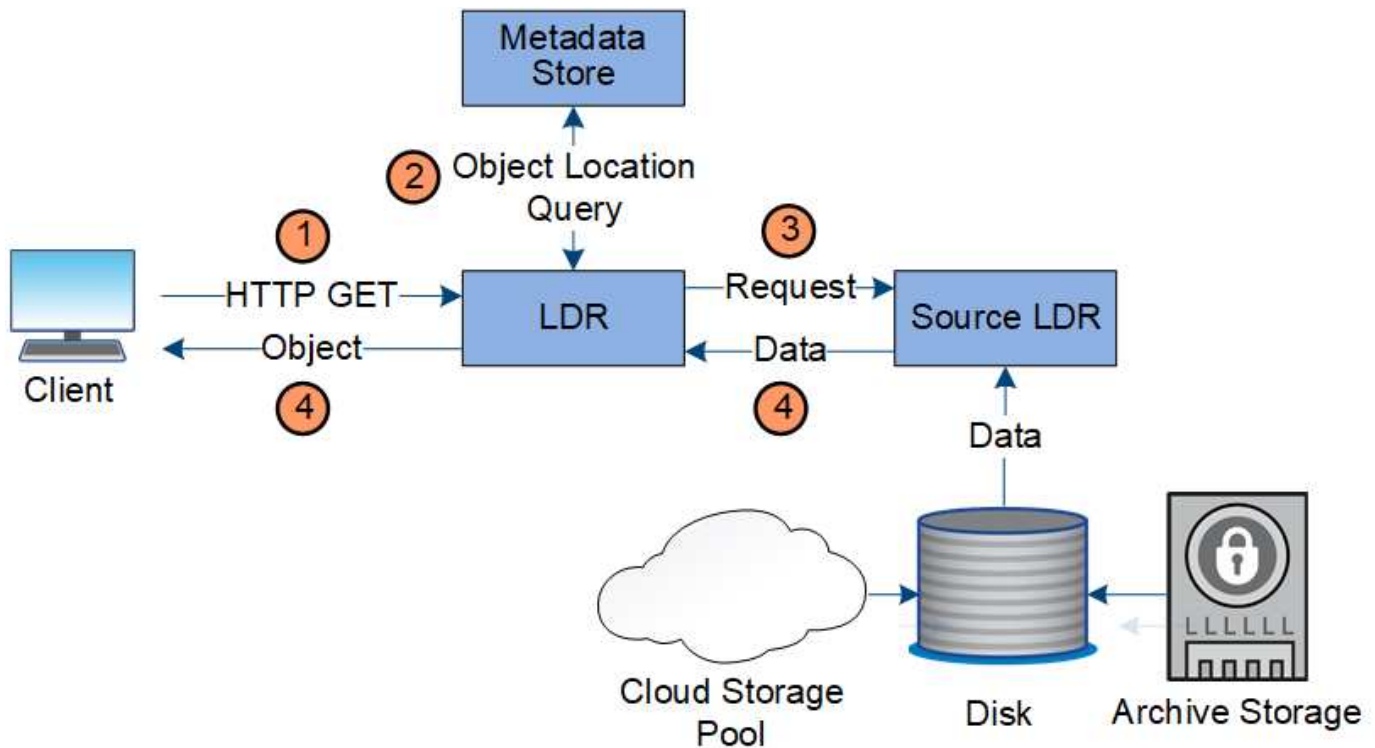
## 检索数据流

检索操作由 StorageGRID 系统与客户端之间定义的数据流组成。系统使用属性跟踪从存储节点或云存储池或归档节点检索对象的情况（如有必要）。

存储节点的 LDR 服务会在元数据存储中查询对象数据的位置，并从源 LDR 服务中检索这些数据。首选情况下，从存储节点检索。如果对象在存储节点上不可用，则检索请求将定向到云存储池或归档节点。



如果唯一的对象副本位于 AWS Glacier 存储或 Azure 归档层上，则客户端应用程序必须对 S3 后对象还原请求进行问题描述，才能将可检索的副本还原到云存储池。



1. LDR 服务从客户端应用程序接收检索请求。
2. LDR 服务会在元数据存储库中查询对象数据位置和元数据。
3. LDR 服务将检索请求转发到源 LDR 服务。
4. 源 LDR 服务从查询的 LDR 服务返回对象数据，系统将对象返回给客户端应用程序。

## 删除数据流

当客户端执行删除操作或对象的生命周期到期时，所有对象副本都会从 StorageGRID 系统中删除，从而触发自动删除。已定义用于删除对象的数据流。

### 删除层次结构

StorageGRID 提供了多种方法来控制何时保留或删除对象。可以根据客户端请求删除对象，也可以自动删除对象。StorageGRID 始终将任何 S3 对象锁定设置优先于客户端删除请求，而客户端删除请求优先于 S3 存储分段生命周期和 ILM 放置说明。

- \* S3 对象锁定 \*：如果为网格启用了全局 S3 对象锁定设置，则 S3 客户端可以在启用了 S3 对象锁定的情况下创建存储分段，然后使用 S3 REST API 为添加到存储分段的每个对象版本指定保留日期和合法保留设置。
  - 无法通过任何方法删除处于合法保留状态的对象版本。
  - 在达到对象版本的保留截止日期之前、任何方法都无法删除该版本。
  - 启用了 S3 对象锁定的存储分段中的对象由 ILM "Forever" 保留。但是，在达到保留截止日期后，可以通过客户端请求或存储分段生命周期到期来删除对象版本。
  - 如果 S3 客户端对存储分段应用默认的保留截止日期、则无需为每个对象指定保留截止日期。

- \* 客户端删除请求 \*：S3 或 Swift 客户端可以问题描述 删除对象请求。当客户端删除某个对象时，该对象的所有副本都会从 StorageGRID 系统中删除。
- 删除存储分段中的对象：租户管理器用户可以使用此选项从StorageGRID 系统中永久删除选定存储分段中的对象和对象版本的所有副本。
- \* S3 存储分段生命周期 \*：S3 客户端可以将生命周期配置添加到指定到期操作的存储分段中。如果存储分段生命周期存在，则在满足到期操作中指定的日期或天数时，StorageGRID 会自动删除对象的所有副本，除非客户端先删除该对象。
- \* ILM 放置说明 \*：假设存储分段未启用 S3 对象锁定，并且没有存储分段生命周期，则 StorageGRID 会在 ILM 规则中的最后一个时间段结束且没有为此对象指定其他放置时自动删除对象。

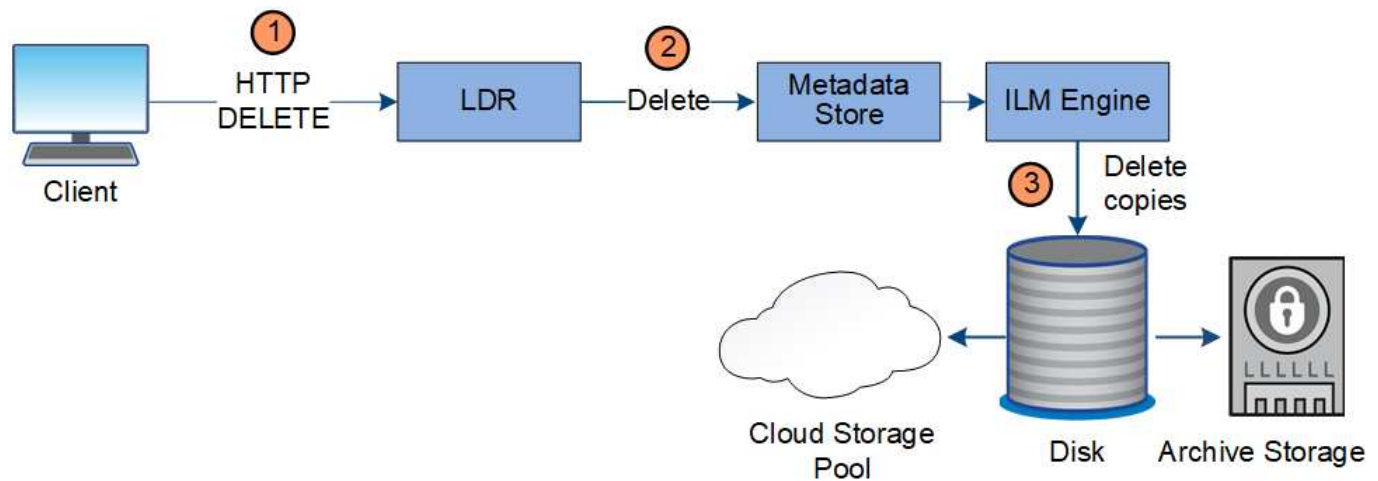


S3 存储分段生命周期中的到期操作始终会覆盖 ILM 设置。因此，即使有关放置对象的任何 ILM 指令已失效，该对象也可能会保留在网格中。

## 如何删除S3删除标记

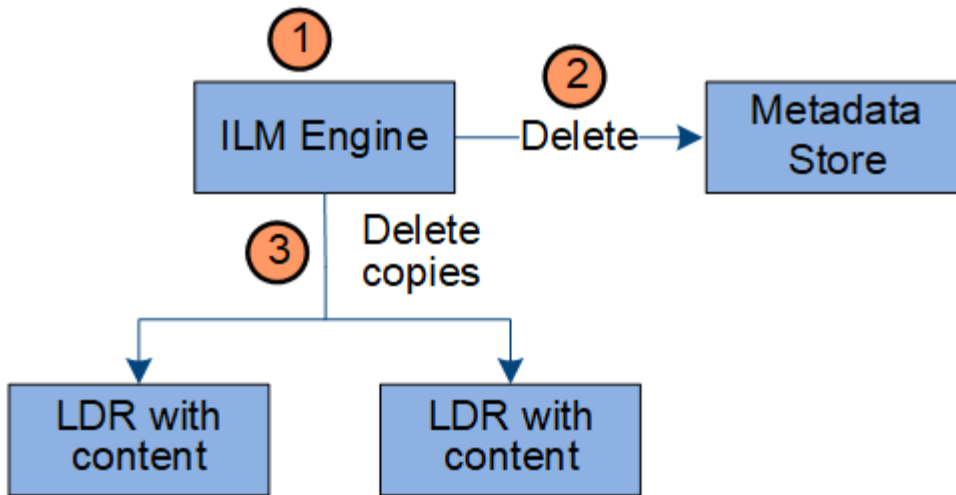
删除受版本控制的对象后、StorageGRID 会创建一个删除标记作为对象的当前版本。要从存储分段中删除零字节删除标记、S3客户端必须明确删除对象版本。ILM、存储分段生命周期规则或删除存储分段操作中的对象不会删除标记。

## 用于客户端删除的数据流



1. LDR 服务从客户端应用程序接收删除请求。
2. LDR 服务会更新元数据存储，使对象在客户端请求时看起来已被删除，并指示 ILM 引擎删除对象数据的所有副本。
3. 对象将从系统中删除。元数据存储已更新，以删除对象元数据。

## 用于 ILM 删除的数据流



1. ILM 引擎确定需要删除此对象。
2. ILM 引擎会通知元数据存储。元数据存储可更新对象元数据，以便在客户端请求中删除此对象。
3. ILM 引擎会删除对象的所有副本。元数据存储已更新，以删除对象元数据。

## 使用信息生命周期管理

您可以使用信息生命周期管理(ILM)控制StorageGRID 系统中所有对象的放置、持续时间和加载行为。ILM 规则可确定 StorageGRID 在一段时间内如何存储对象。您可以配置一个或多个 ILM 规则，然后将其添加到 ILM 策略中。

一个网格一次只有一个活动策略。一个策略可以包含多个规则。

ILM 规则定义：

- 应存储哪些对象。规则可以应用于所有对象，也可以指定筛选器来标识规则适用场景 中的对象。例如，规则只能应用于与特定租户帐户，特定 S3 分段或 Swift 容器或特定元数据值关联的对象。
- 存储类型和位置。对象可以存储在存储节点，云存储池或归档节点上。
- 创建的对象副本的类型。可以复制副本或对副本进行纠删编码。
- 对于复制的副本，为创建的副本数。
- 对于纠删编码副本，使用纠删编码方案。
- 对象的存储位置和副本类型会随时间发生变化。
- 在将对象载入网格时如何保护对象数据（同步放置或双提交）。

请注意，对象元数据不受 ILM 规则管理。而是将对象元数据存储于 Cassandra 数据库中，该数据库称为元数据存储。每个站点会自动维护三个对象元数据副本，以防止数据丢失。

### ILM 规则示例

例如、ILM规则可以指定以下内容：

- 仅应用于属于租户A的对象

- 为这些对象创建两个复制副本、并将每个副本存储在不同的站点上。
- 保留两个副本“forever、”、这意味着StorageGRID 不会自动删除它们。相反， StorageGRID 将保留这些对象，直到客户端删除请求或存储分段生命周期到期时将其删除为止。
- 使用均衡选项进行加载行为：租户A将对象保存到StorageGRID 后立即应用双站点放置指令、除非无法立即创建两个所需的副本。

例如，如果租户 A 保存对象时无法访问站点 2，则 StorageGRID 将在站点 1 的存储节点上创建两个临时副本。一旦站点 2 可用， StorageGRID 就会在该站点创建所需的副本。

## ILM 策略如何评估对象

StorageGRID 系统的活动ILM策略用于控制所有对象的放置、持续时间和入射行为。

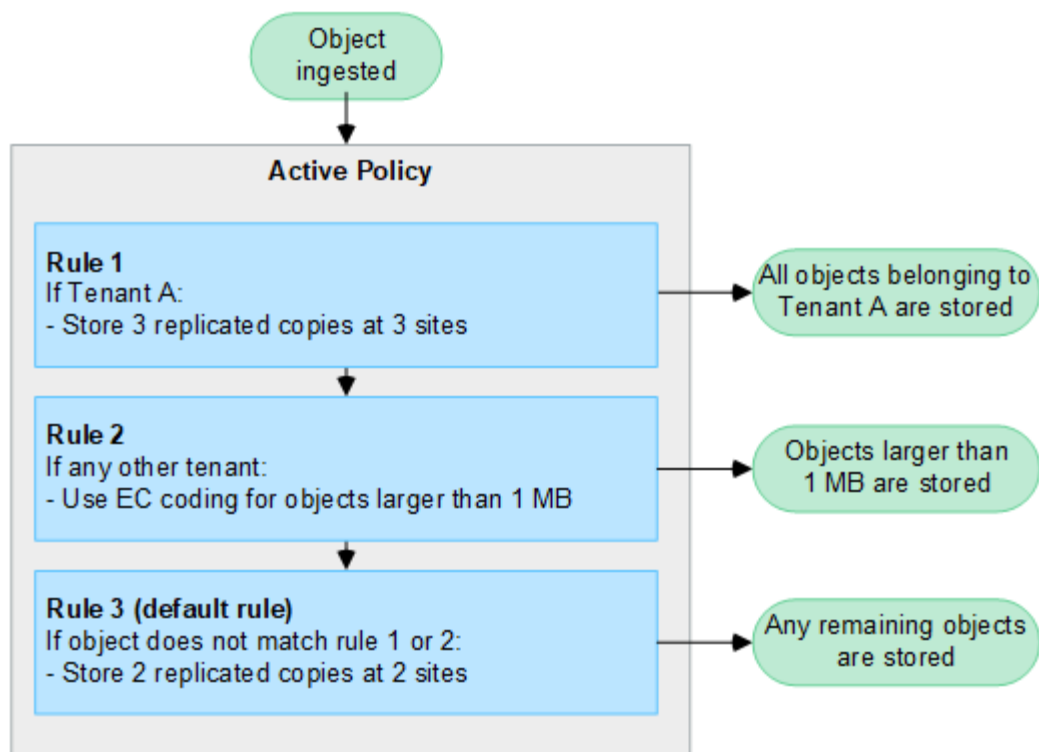
当客户端将对象保存到 StorageGRID 时，系统会根据活动策略中按顺序排列的一组 ILM 规则对这些对象进行评估，如下所示：

1. 如果策略中第一个规则的筛选器与某个对象匹配，则会根据该规则的载入行为载入该对象，并根据该规则的放置说明进行存储。
2. 如果第一个规则的筛选器与对象不匹配、则系统将根据策略中的每个后续规则评估对象、直到进行匹配为止。
3. 如果没有与对象匹配的规则，则会应用策略中默认规则的载入行为和放置说明。默认规则是策略中的最后一个规则、不能使用任何筛选器。它必须应用于所有租户，所有分段和所有对象版本。

## ILM 策略示例

例如、一个ILM策略可以包含三个ILM规则、这些规则可指定以下内容：

- **\*规则1：为租户A\*复制的副本**
  - 匹配属于租户A的所有对象
  - 将这些对象作为三个复制副本存储在三个站点上。
  - 规则1不匹配属于其他租户的对象、因此会根据规则2对其进行评估。
- **\*规则2：对大于1 MB\*的对象进行纠删编码**
  - 匹配其他租户的所有对象、但前提是这些对象大于1 MB。这些较大的对象在三个站点上使用 6+3 纠删编码进行存储。
  - 与小于或等于1 MB的对象不匹配、因此将根据规则3评估这些对象。
- **规则3：2个副本2个数据中心(默认)**
  - 是策略中的最后一个默认规则。不使用筛选器。
  - 为规则1或规则2不匹配的所有对象创建两个复制副本(不属于租户A且小于或等于1 MB的对象)。



#### 相关信息

- ["使用 ILM 管理对象"](#)

## 版权信息

版权所有 © 2024 NetApp, Inc.。保留所有权利。中国印刷。未经版权所有者事先书面许可，本文档中受版权保护的任何部分不得以任何形式或通过任何手段（图片、电子或机械方式，包括影印、录音、录像或存储在电子检索系统中）进行复制。

从受版权保护的 NetApp 资料派生的软件受以下许可和免责声明的约束：

本软件由 NetApp 按“原样”提供，不含任何明示或暗示担保，包括但不限于适销性以及针对特定用途的适用性的隐含担保，特此声明不承担任何责任。在任何情况下，对于因使用本软件而以任何方式造成的任何直接性、间接性、偶然性、特殊性、惩罚性或后果性损失（包括但不限于购买替代商品或服务；使用、数据或利润方面的损失；或者业务中断），无论原因如何以及基于何种责任理论，无论出于合同、严格责任或侵权行为（包括疏忽或其他行为），NetApp 均不承担责任，即使已被告知存在上述损失的可能性。

NetApp 保留在不另行通知的情况下随时对本文档所述的任何产品进行更改的权利。除非 NetApp 以书面形式明确同意，否则 NetApp 不承担因使用本文档所述产品而产生的任何责任或义务。使用或购买本产品不表示获得 NetApp 的任何专利权、商标权或任何其他知识产权许可。

本手册中描述的产品可能受一项或多项美国专利、外国专利或正在申请的专利的保护。

有限权利说明：政府使用、复制或公开本文档受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中“技术数据权利 — 非商用”条款第 (b)(3) 条规定的限制条件的约束。

本文档中所含数据与商业产品和/或商业服务（定义见 FAR 2.101）相关，属于 NetApp, Inc. 的专有信息。根据本协议提供的所有 NetApp 技术数据和计算机软件具有商业性质，并完全由私人出资开发。美国政府对这些数据的使用权具有非排他性、全球性、受限且不可撤销的许可，该许可既不可转让，也不可再许可，但仅限在与交付数据所依据的美国政府合同有关且受合同支持的情况下使用。除本文档规定的情形外，未经 NetApp, Inc. 事先书面批准，不得使用、披露、复制、修改、操作或显示这些数据。美国政府对国防部的授权仅限于 DFARS 的第 252.227-7015(b)（2014 年 2 月）条款中明确的权利。

## 商标信息

NetApp、NetApp 标识和 <http://www.netapp.com/TM> 上所列的商标是 NetApp, Inc. 的商标。其他公司和产品名称可能是其各自所有者的商标。