



如何存储对象（复制或擦除编码） StorageGRID software

NetApp
July 23, 2026

目录

- 如何存储对象（复制或擦除编码） 1
 - 了解 StorageGRID 中的复制 1
 - StorageGRID 中单副本复制的风险 2
 - 了解 StorageGRID 中的纠删编码 4
 - 了解 StorageGRID 中的纠删编码方案 6
 - 数据保护 6
 - 存储开销 6
 - 有关存储池的指南 6
 - 包含三个或更多站点的存储池的擦除编码方案 7
 - 单站点存储池的纠删码方案 8
 - StorageGRID 中纠删编码的优势、劣势和要求 8
 - 纠删码的优势 8
 - 纠删码的缺点 9
 - 何时使用擦除编码 9

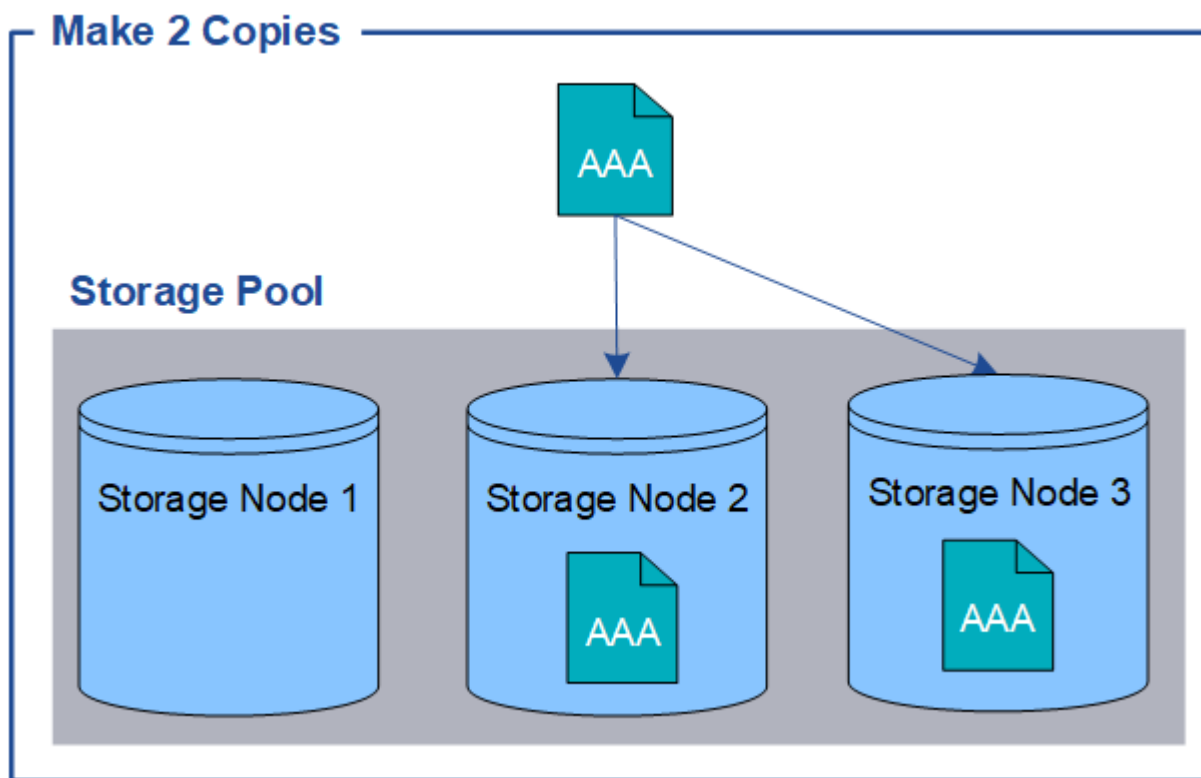
如何存储对象（复制或擦除编码）

了解 StorageGRID 中的复制

复制是 StorageGRID 用于存储对象数据的两种方法之一（纠删码是另一种方法）。当对象与使用复制的 ILM 规则匹配时，系统会创建对象数据的精确副本，并将副本存储在存储节点上。

当您将 ILM 规则配置为创建复制副本时，您需要指定应创建多少个副本、这些副本应放置在何处以及副本应在每个位置存储多长时间。

在以下示例中，ILM 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。



当 StorageGRID 将对象与此规则匹配时，它会创建对象的两个副本，将每个副本放在存储池中的不同存储节点上。这两个副本可以放置在三个可用存储节点中的任意两个节点上。在这种情况下，规则将对象副本放置在存储节点 2 和 3 上。由于存在两个副本，因此如果存储池中的任何节点出现故障，仍可检索该对象。



StorageGRID 只能在任何给定存储节点上存储对象的一个复制副本。如果您的网格包括三个存储节点，并且您创建了 4 副本 ILM 规则，则只会创建三个副本——每个存储节点一个副本。系统将触发 **ILM placement unachievable** 警报，指示 ILM 规则无法完全应用。

相关信息

- ["什么是纠删码"](#)
- ["什么是存储池"](#)
- ["使用复制和纠删码启用站点丢失保护"](#)

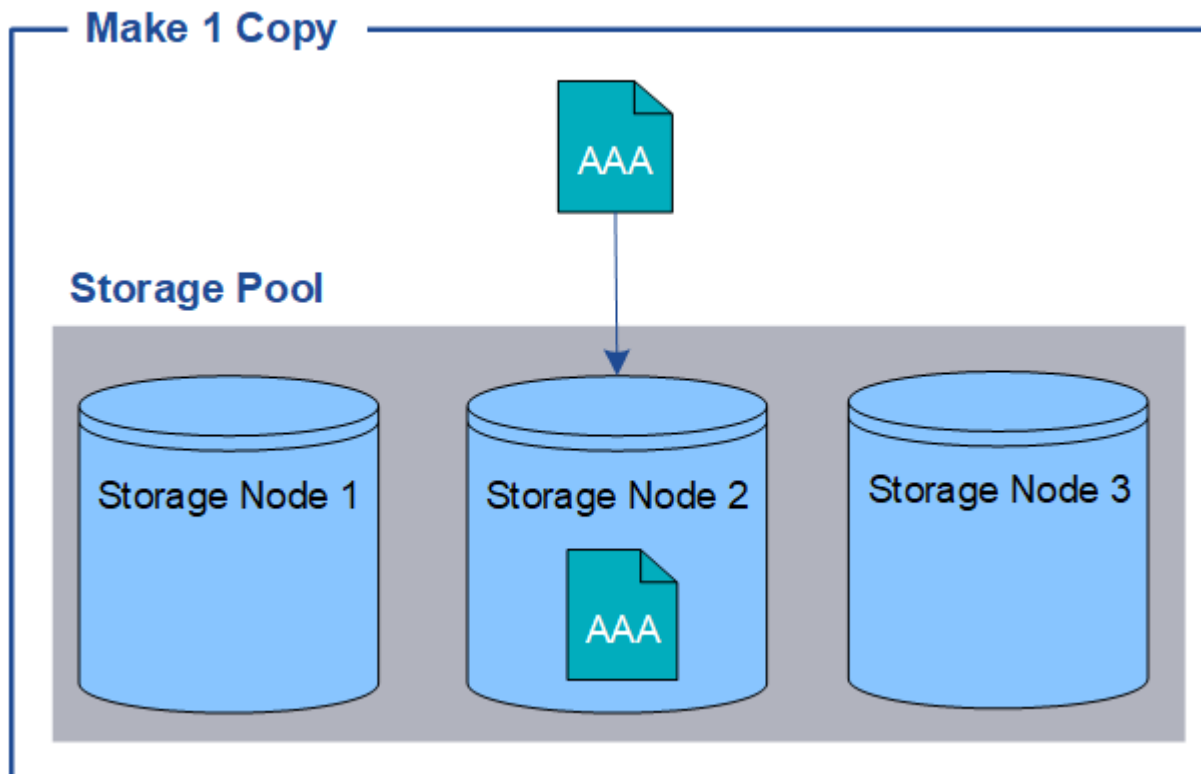
StorageGRID 中单副本复制的风险

在创建 ILM 规则以创建复制副本时，应始终在放置说明中为任何时间段指定至少两个副本。



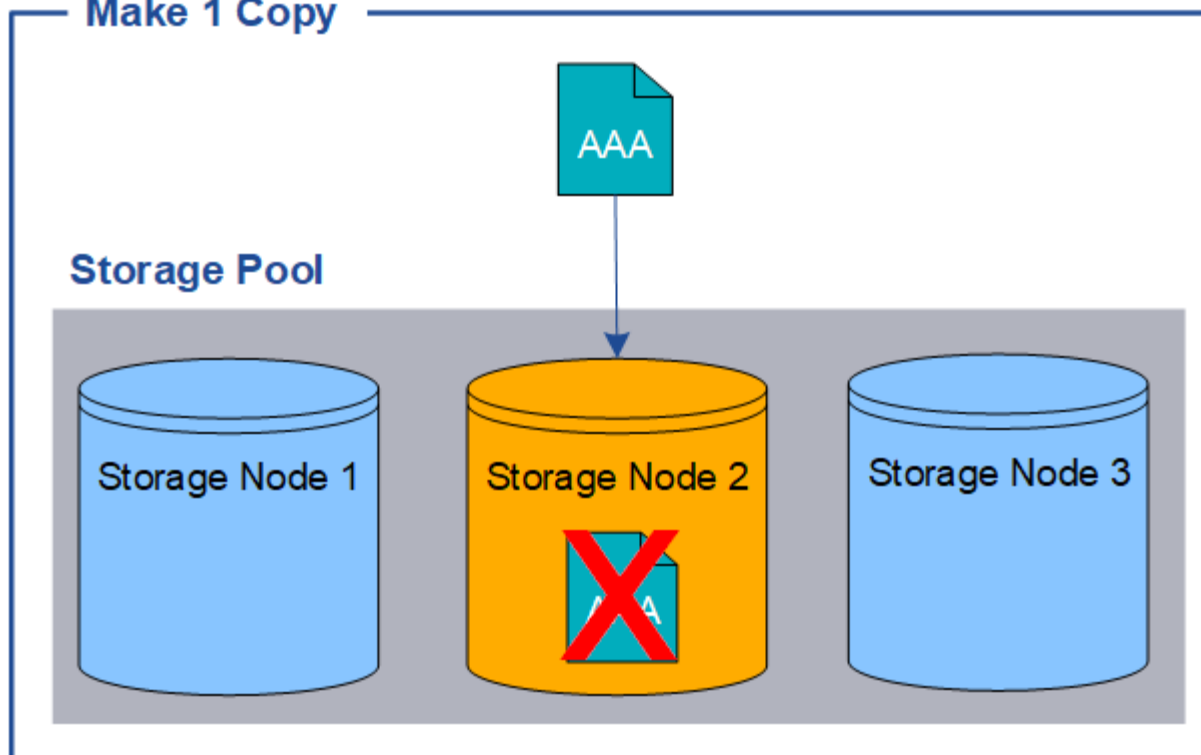
请勿使用在任何时间段仅创建一个复制副本的 ILM 规则。如果仅存在一个对象的复制副本，则如果存储节点出现故障或出现重大错误，该对象将丢失。在升级等维护过程中，您还会暂时失去对该对象的访问权限。

在以下示例中，Make 1 Copy ILM 规则指定将对象的一个复制副本放入包含三个存储节点的存储池中。当接收到与此规则匹配的对象时，StorageGRID 仅在一个存储节点上放置一个副本。



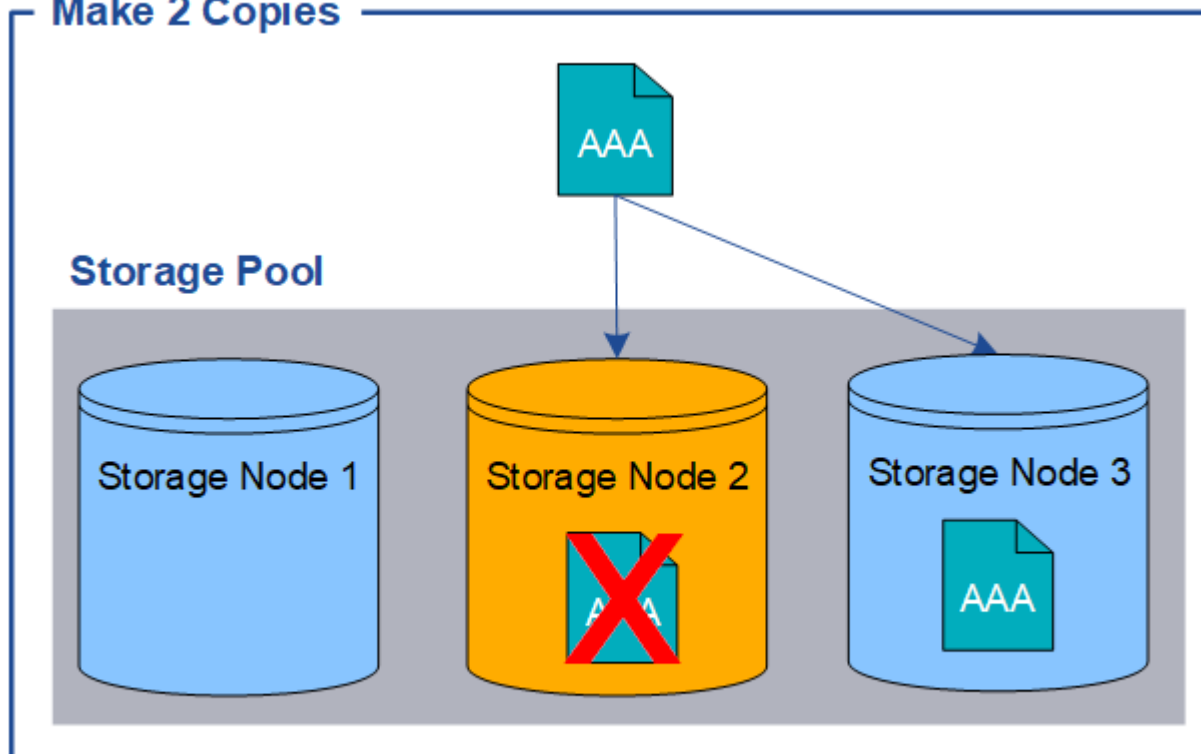
当 ILM 规则仅创建对象的一个复制副本时，当存储节点不可用时，该对象将变得不可访问。在此示例中，每当 Storage Node 2 脱机时（例如在升级或其他维护过程中），您将暂时失去对对象 AAA 的访问权限。如果 Storage Node 2 出现故障，您将完全丢失对象 AAA。

Make 1 Copy



为避免丢失对象数据，应始终至少制作要通过复制保护的所有对象的两个副本。如果存在两个或多个副本，则即使一个 Storage Node 出现故障或脱机，您仍然可以访问该对象。

Make 2 Copies



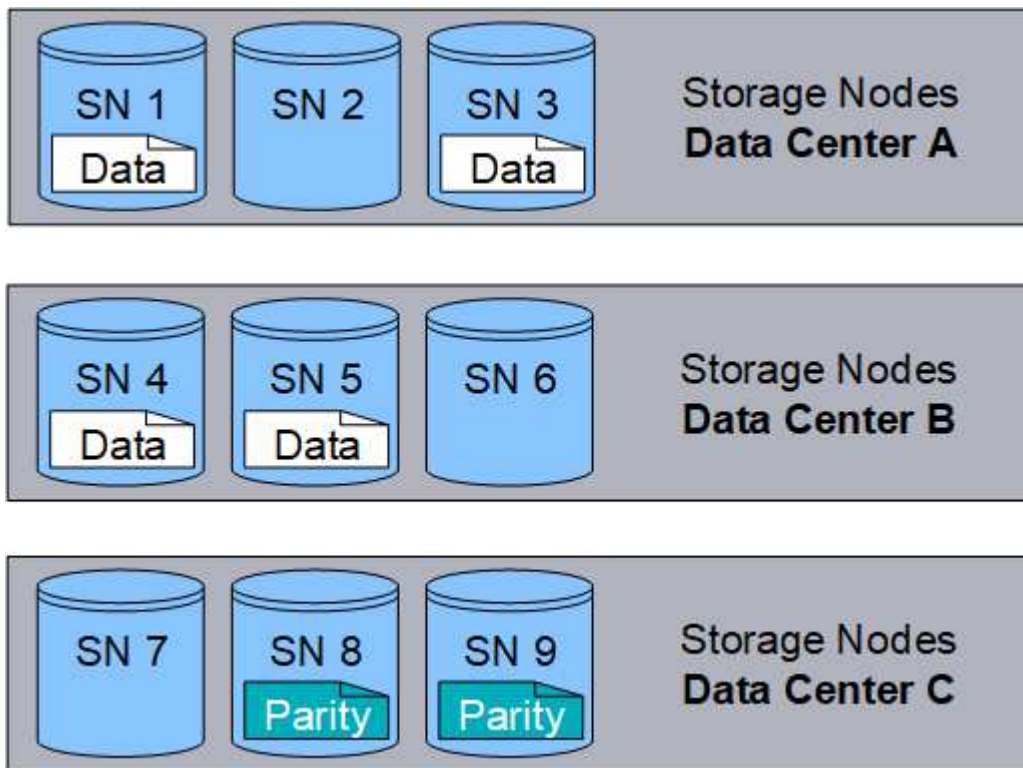
了解 StorageGRID 中的纠删编码

纠删编码是 StorageGRID 用于存储对象数据的两种方法之一（复制是另一种方法）。当对象与使用纠删编码的 ILM 规则匹配时，这些对象将被切分为数据片段，并计算额外的奇偶校验片段，每个片段存储在不同的存储节点上。

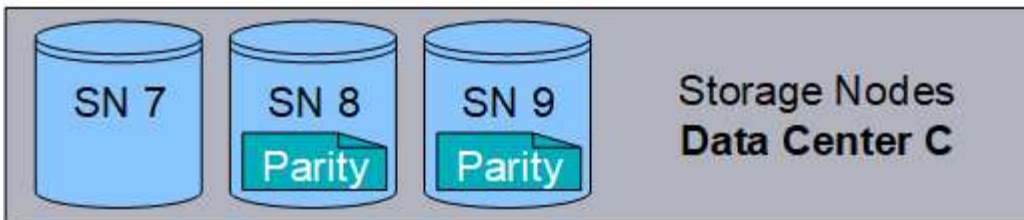
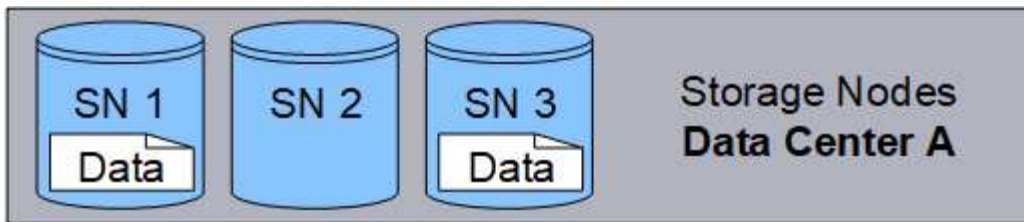
访问对象时，使用存储的片段重新组装对象。如果数据或奇偶校验片段损坏或丢失，则擦除编码算法可以使用剩余数据和奇偶校验片段的子集重新创建该片段。

创建 ILM 规则时，StorageGRID 会创建支持这些规则的纠删码配置文件。您可以查看纠删码配置文件列表、["重命名擦除编码配置文件"](#)或["如果当前未在任何 ILM 规则中使用纠删码配置文件，则停用该配置文件"](#)。

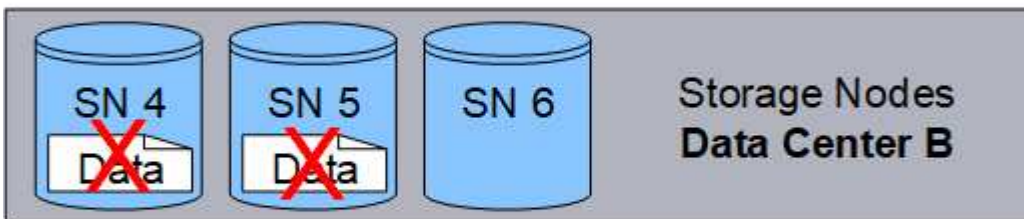
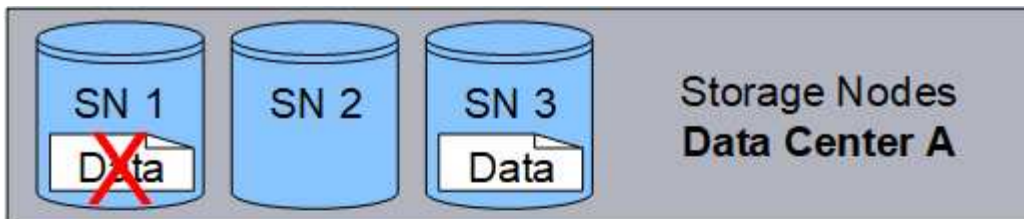
以下示例说明了对对象数据的纠删编码算法的使用。在此示例中，ILM 规则使用 4+2 纠删编码方案。每个对象被切割成四个相等的数据片段，并从对象数据中计算两个奇偶校验片段。六个片段中的每一个都存储在三个数据中心站点的不同节点上，为节点故障或站点丢失提供数据保护。



4+2 擦除编码方案可以通过多种方式进行配置。例如，您可以配置包含六个 Storage Node 的单站点存储池。对于["站点丢失保护"](#)，您可以使用包含三个站点的存储池，每个站点具有三个 Storage Node。只要六个片段中的任意四个片段（数据或奇偶校验）仍然可用，即可检索对象。最多可以丢失两个片段而不会丢失对象数据。如果整个站点丢失，只要所有其他片段仍可访问，仍然可以检索或修复该对象。



如果丢失两个以上的存储节点，则无法检索该对象。



相关信息

- ["什么是复制"](#)
- ["什么是存储池"](#)
- ["什么是擦除编码方案"](#)

- ["重命名擦除编码配置文件"](#)
- ["停用擦除编码配置文件"](#)

了解 StorageGRID 中的纠删编码方案

擦除编码方案控制为每个对象创建多少个数据片段和多少个奇偶校验片段。

创建或编辑 ILM 规则时，您需要选择一个可用的纠删码方案。StorageGRID 会根据构成您计划使用的存储池的存储节点数量和站点数量，自动创建纠删码方案。

数据保护

StorageGRID 系统使用 Reed-Solomon 纠删码算法。该算法将对象切片成 k 个数据片段并计算 m 个奇偶校验片段。

这些 $k + m = n$ 碎片分布在 n 存储节点上，以提供以下数据保护：

- 要检索或修复对象，需要 k 个片段。
- 一个对象最多可以承受 m 个丢失或损坏的片段。 m 的值越高，故障容忍度越高。

最佳数据保护由存储池中具有最高节点或卷故障容忍度的纠删码方案提供。

存储开销

擦除编码方案的存储开销通过将奇偶校验片段数 (m) 除以数据片段数 (k) 来计算。您可以使用存储开销来计算每个擦除编码对象所需的磁盘空间：

$$\text{disk space} = \text{object size} + (\text{object size} * \text{storage overhead})$$

例如，如果使用 4+2 方案存储 10 MB 对象（具有 50% 的存储开销），则该对象会消耗 15 MB 的网格存储。如果使用 6+2 方案（存储开销为 33%）存储相同的 10 MB 对象，则该对象大约会消耗 13.3 MB。

选择符合您需求的总值最低的擦除编码方案。 $k+m$ 片段数量较少的擦除编码方案的计算效率更高，因为：

- 每个对象创建和分发（或检索）的片段更少
- 它们显示出更好的性能，因为片段大小较大
- 它们可能需要在 ["在需要更多存储时进行扩展"](#) 中添加更少的节点

有关存储池的指南

在选择要用于将创建擦除编码副本的规则存储池时，请使用以下存储池指南：

- 存储池必须包含三个或更多站点，或仅包含一个站点。



如果存储池包括两个站点，则无法使用擦除编码。

- [包含三个或更多站点的存储池的擦除编码方案](#)

◦ 单站点存储池的纠删码方案

- 请勿使用包含"所有站点"站点的存储池。
- 存储池应至少包括 $k+m+1$ 个可以存储对象数据的存储节点。



存储节点可以在安装过程中配置为包含对象数据和元数据 ("组合"存储节点)，仅包含对象元数据或仅包含对象数据。有关详细信息，请参阅 ["存储节点类型"](#)。

所需存储节点最小数量为 $k+m$ 。但是，如果所需的存储节点暂时不可用，则至少拥有一个额外的存储节点可以帮助防止接收故障或 ILM 积压。

包含三个或更多站点的存储池的擦除编码方案

下表介绍了 StorageGRID 目前为包含三个或更多站点的存储池支持的纠删码方案。所有这些方案都提供站点丢失保护。一个站点可能会丢失，但该对象仍可访问。

对于提供站点丢失保护的纠删码方案，存储池中存储节点的建议数量超过 $k+m+1$ ，因为每个站点至少需要三个存储节点。

擦除编码方案($k+m$)	最小部署站点数	每个站点上推荐的存储节点数量	建议的存储节点总数	站点丢失保护?	存储开销
4+2	3	3	9	是	50%
6+2	4	3	12	是	33%
8+2	5	3	15	是	25%
6+3	3	4	12	是	50%
9+3	4	4	16	是	33%
2+1	3	3	9	是	50%
4+1	5	3	15	是	25%
6+1	7	3	21	是	17%
7+5	3	5	15	是	71%



StorageGRID 每个站点至少需要三个存储节点。要使用 7+5 方案，每个站点至少需要四个存储节点。建议每个站点使用五个存储节点。

在选择提供站点保护的擦除编码方案时，请权衡以下因素的相对重要性：

- 碎片数量：当碎片总数较低时，性能和扩展灵活性通常更好。

- 容错：通过使用更多的奇偶校验段（即，当 `m` 具有更高的值时）来提高容错能力。
- 网络流量：从故障中恢复时，使用具有更多片段（即 `k+m` 总数更高）的方案会产生更多网络流量。
- 存储开销：开销较高的方案每个对象需要更多的存储空间。

例如，在 4+2 方案和 6+3 方案（两者的存储开销均为 50%）之间进行选择时，如果需要额外的容错能力，请选择 6+3 方案。如果网络资源受限，请选择 4+2 方案。如果所有其他因素相同，请选择 4+2 方案，因为它的片段总数较少。



如果您不确定要使用哪种方案，请选择 4+2 或 6+3，或联系技术支持。

单站点存储池的纠删码方案

一个站点存储池支持为三个或多个站点定义的所有纠删码方案，前提是该站点具有足够的存储节点。

所需存储节点的最小数量为 $k+m$ ，但建议使用含有 ` $k+m+1$ ` 个存储节点的存储池。例如，2+1 纠删码方案需要至少包含三个存储节点的存储池，但建议使用四个存储节点。

擦除编码方案($k+m$)	最小存储节点数	建议的存储节点数	存储开销
4+2	6	7	50%
6+2	8	9	33%
8+2	10	11	25%
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%

StorageGRID 中纠删编码的优势、劣势和要求

在决定是否使用复制或纠删编码来保护对象数据免受丢失之前，您应该了解纠删编码的优点和要求。

纠删码的优势

与复制相比，纠删码提高了可靠性、可用性和存储效率。

- 可靠性：通过在不丢失数据的情况下可以承受的并发故障数量来衡量。
 - 复制：多个相同的对象副本存储在不同节点和站点上。
 - 纠删码：对象被编码为数据和奇偶校验片段，并分布在许多节点和站点上。此分散提供站点和节点故障保护。
- 可用性：在存储节点出现故障或无法访问时获取对象的能力。与复制相比，纠删编码以相当的存储成本提供了更高的可用性。
- 存储效率：为了获得相似的可用性和可靠性，纠删码对象比复制对象使用更少的磁盘空间。例如，复制到两个站点的 10 MB 对象会消耗 20 MB 的磁盘空间（两个副本），而使用 6+3 纠删码方案跨三个站点进行纠删码的对象只会消耗 15 MB 的磁盘空间。



擦除编码对象的磁盘空间计算为对象大小加上存储开销。存储开销百分比是奇偶校验片段数除以数据片段数。

纠删码的缺点

与复制相比，擦除编码具有以下缺点：

- 建议增加存储节点和站点的数量，具体取决于纠删码方案。相比之下，如果复制对象数据，则每个副本只需要一个存储节点。请参阅 ["包含三个或更多站点的存储池的擦除编码方案"](#) 和 ["单站点存储池的纠删码方案"](#)。
- 存储扩展的成本和复杂性增加。若要扩展使用复制的部署，请在生成对象副本的每个位置添加存储容量。要扩展使用擦除编码的部署，必须考虑正在使用的擦除编码方案以及现有存储节点的完整程度。例如，如果等待现有节点 100% 满，则必须至少添加 `k+m` 存储节点，但如果在现有节点 70% 满时进行扩展，则可以为每个站点添加两个节点，并最大限度地提高可用存储容量。有关详细信息，请参见["为纠删码对象添加存储容量"](#)。
- 当您跨地理位置分散的站点使用纠删码时，检索延迟会增加。通过 WAN 连接检索经过纠删码编码并分布在远程站点上的对象片段，比检索已复制且在本地可用的对象（即客户端所连接的同一站点）需要更长的时间。
- 当您跨地理位置分散的站点使用纠删码时，用于检索和修复的 WAN 网络流量使用率较高，特别是对于频繁检索的对象或通过 WAN 网络连接进行的对象修复。
- 当您跨站点使用纠删码时，随着站点之间的网络延迟增加，最大对象吞吐量急剧下降。这种下降是 TCP 网络吞吐量下降的结果，这会影响 StorageGRID 系统存储和检索对象片段的速度。
- 计算资源使用率更高。

何时使用擦除编码

对以下要求使用擦除编码：

- 大于 1 MB 的对象。



纠删码最适合大于 1 MB 的对象。不要对小于 200 KB 的对象使用纠删码，以避免管理极小纠删码片段所带来的开销。

- 针对不常检索的内容进行长期或冷存储。
- 高数据可用性和可靠性。
- 针对完整站点和节点故障的保护。

- 存储效率。
- 需要通过单个纠删码副本而非多个复制副本实现高效数据保护的单站点部署。
- 站点间延迟小于 100 ms 的多站点部署。

版权信息

版权所有 © 2026 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。