对象的存储方式（复制或纠删编码）

StorageGRID 11.5

NetApp
April 11, 2024

目录

对象的存储方式（复制或纠删编码）

StorageGRID 可以通过存储复制的副本或存储经过纠删编码的副本来防止对象丢失。您可以在ILM规则的放置说明中指定要创建的副本类型。

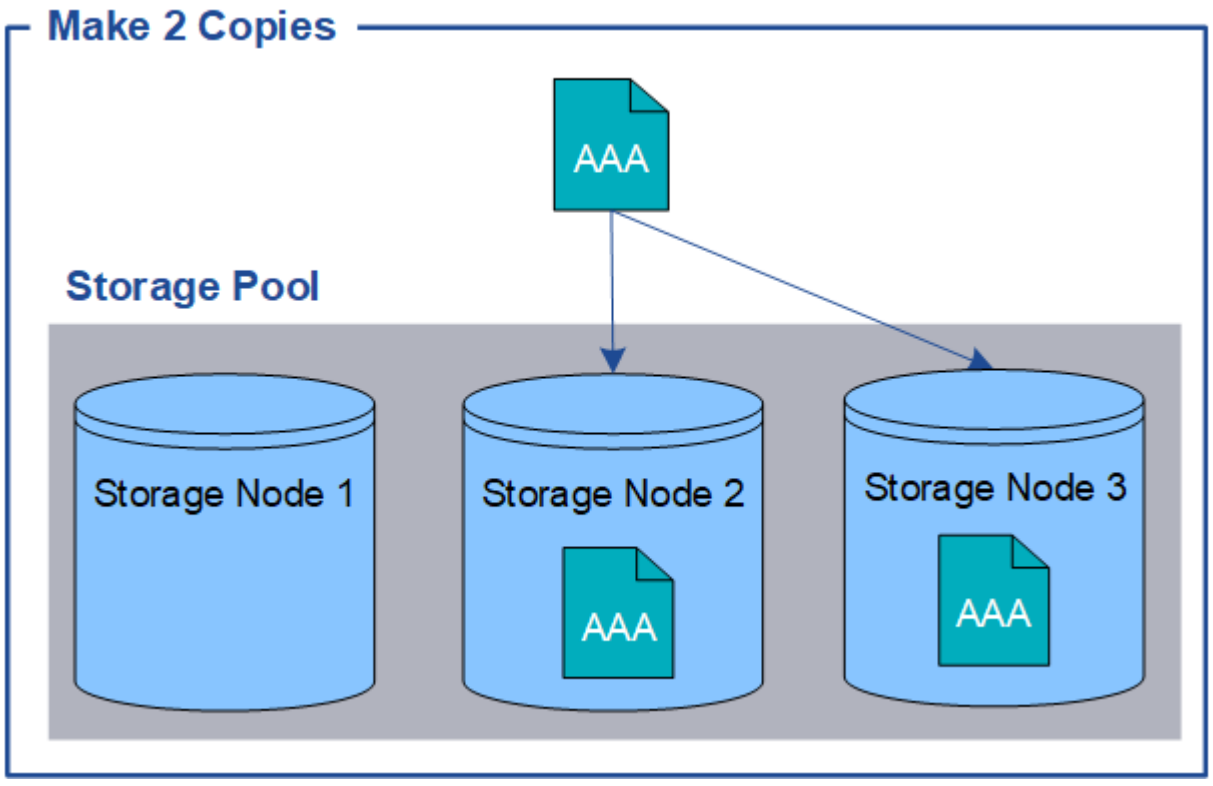
- "什么是复制"
- "为什么不应使用单副本复制"
- "什么是纠删编码"
- "什么是纠删编码方案"
- "纠删编码的优势，劣势和要求"

什么是复制

复制是 StorageGRID 用于存储对象数据的两种方法之一。当对象与使用复制的 ILM 规则匹配时，系统会为对象数据创建精确副本，并将这些副本存储在存储节点或归档节点上。

在配置创建复制副本的 ILM 规则时，您可以指定应创建多少个副本，这些副本应放置在何处以及应将这些副本存储在每个位置的时间长度。

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，它会为该对象创建两个副本，并将每个副本放置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上。在这种情况下，规则会将对象副本放置在存储节点 2 和 3 上。由于有两个副本，因此，如果存储池中的任何节点出现故障，可以检索此对象。

StorageGRID 只能在任何给定存储节点上存储一个对象的一个复制副本。如果您的网格包含三个存储节点，并且您创建了一个 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 * 无法实现 ILM 放置 * 警报，以指示无法完全应用 ILM 规则。

相关信息

"什么是存储池"

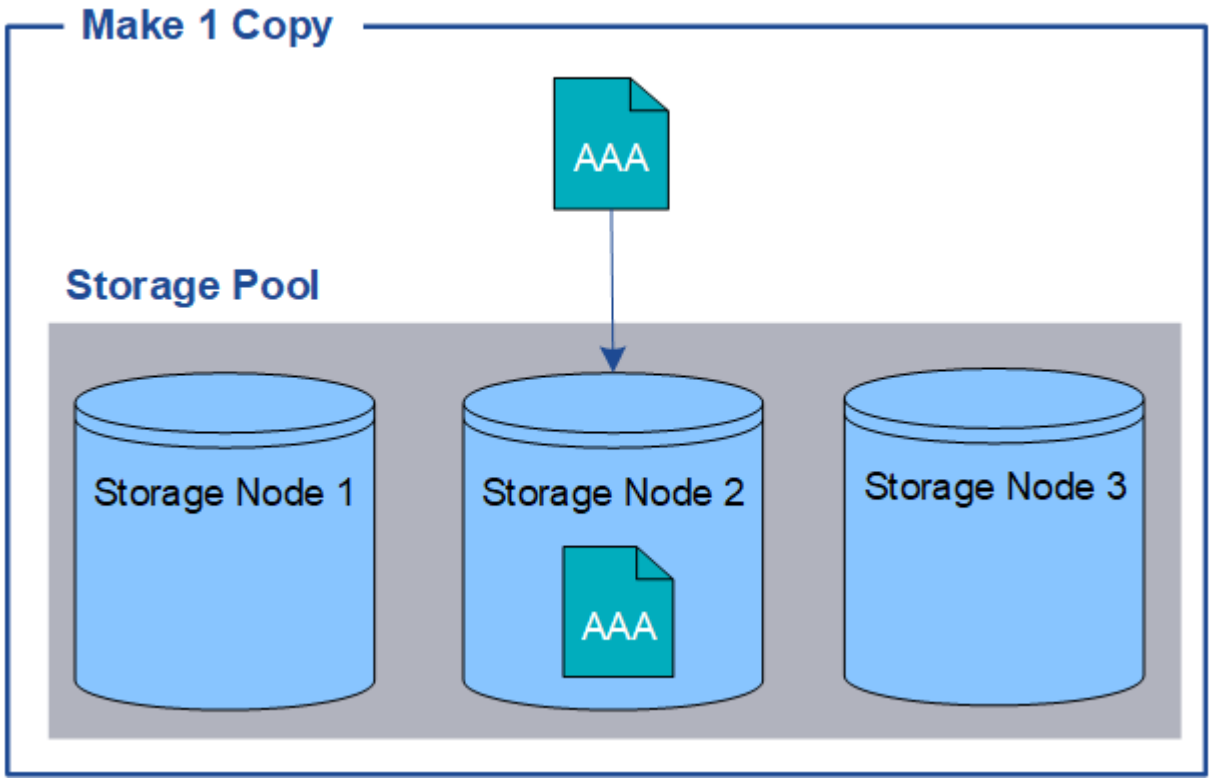
"使用多个存储池进行跨站点复制"

为什么不应使用单副本复制

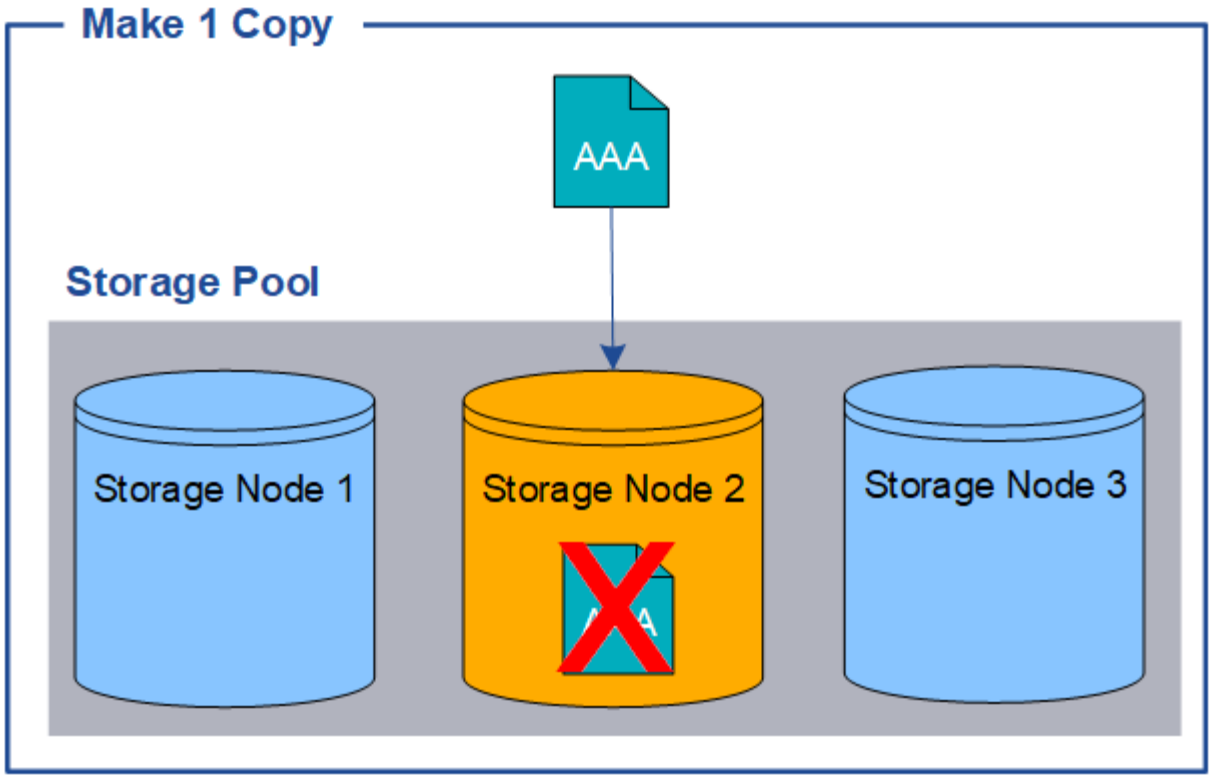
在创建 ILM 规则以创建复制副本时，您应始终在放置说明中指定至少两个副本，以便在任意时间段内使用。

请勿使用仅在任意时间段创建一个复制副本的 ILM 规则。如果某个对象只存在一个复制副本，则在存储节点出现故障或出现严重错误时，该对象将丢失。在升级等维护过程中，您还会暂时失去对对象的访问权限。

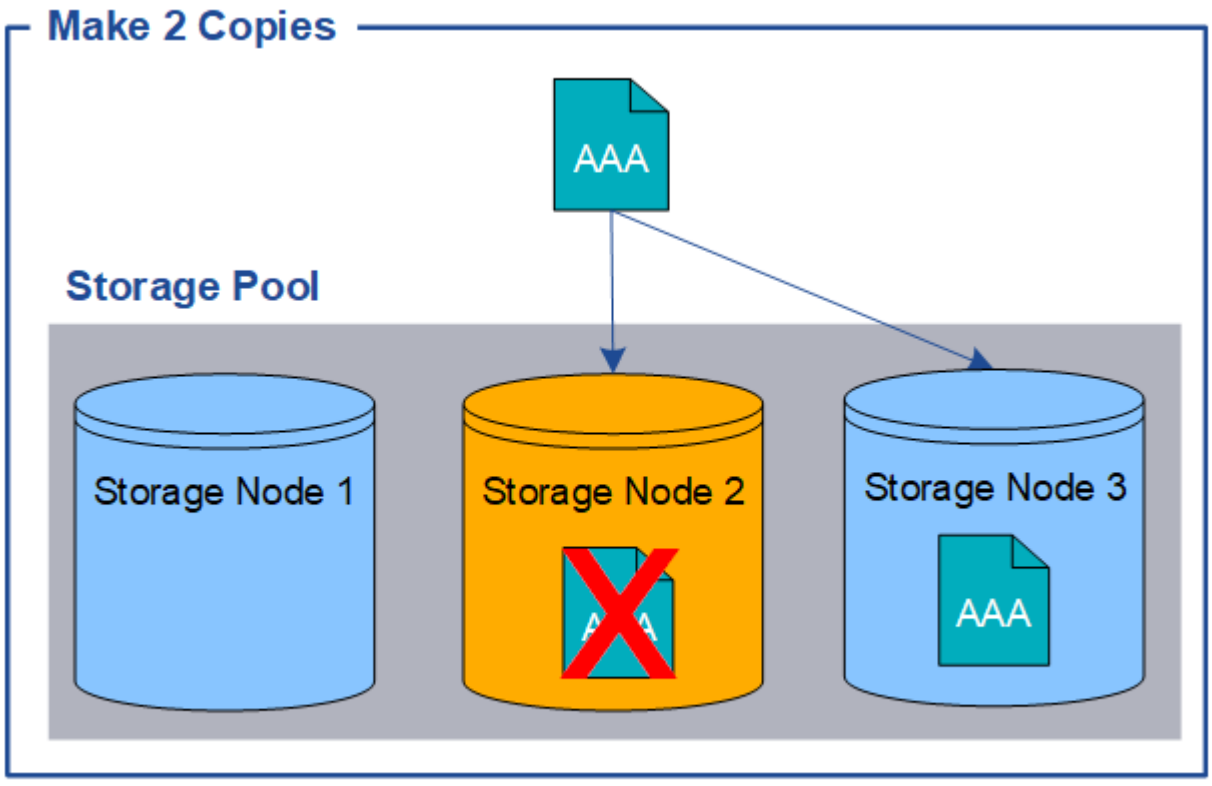
在以下示例中， Make 1 Copy ILM 规则指定将对象的一个复制副本放置在包含三个存储节点的存储池中。如果载入的对象与此规则匹配，则 StorageGRID 仅会在一个存储节点上放置一个副本。



如果 ILM 规则仅创建一个对象的一个复制副本，则在存储节点不可用时，此对象将无法访问。在此示例中，只要存储节点 2 脱机，例如在升级或其他维护操作步骤 期间，您将暂时无法访问对象 AAA 。如果存储节点 2 发生故障，您将完全丢失对象 AAA 。



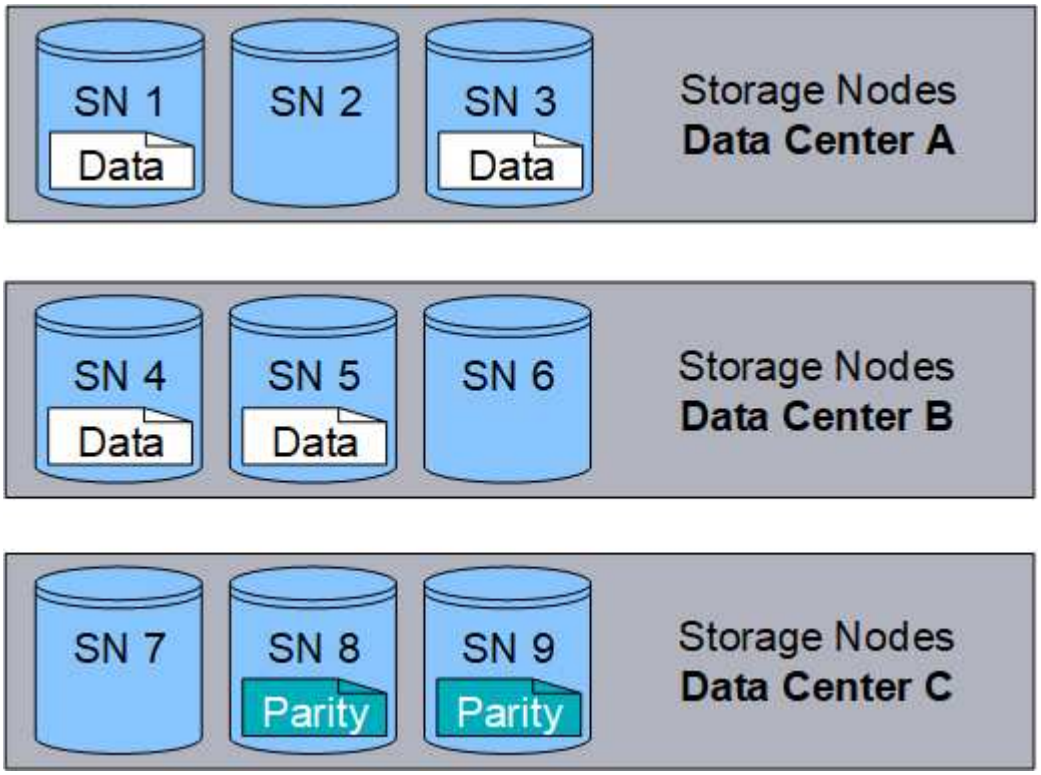
为了避免丢失对象数据，您应始终为要通过复制保护的所有对象创建至少两个副本。如果存在两个或更多副本，则在一个存储节点出现故障或脱机时，您仍可以访问此对象。



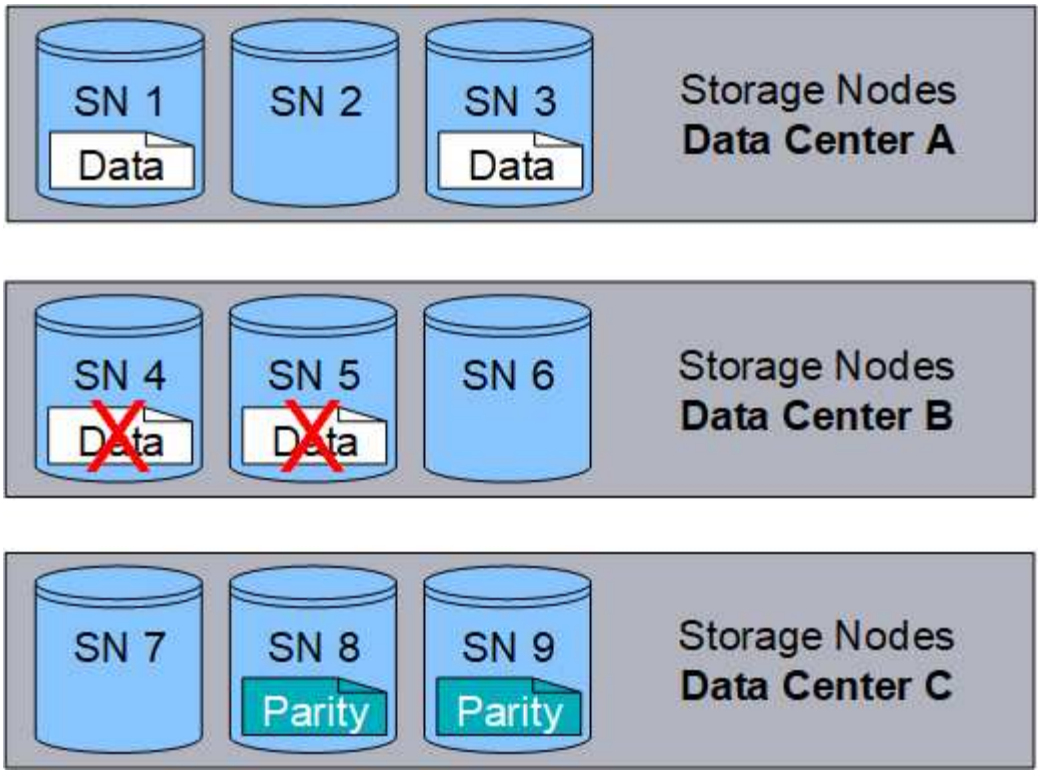
什么是纠删编码

纠删编码是 StorageGRID 存储对象数据的第二种方法。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，则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。

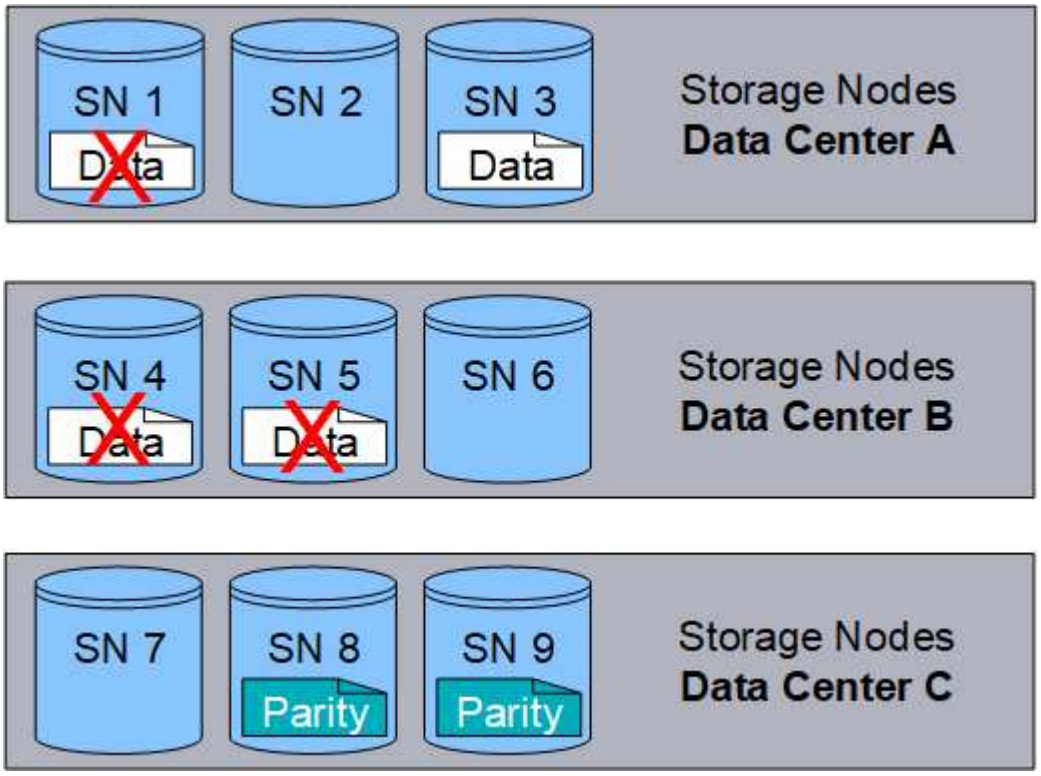
以下示例说明了如何对对象数据使用纠删编码算法。在此示例中， ILM 规则使用 4+2 纠删编码方案。每个对象都会被划分为四个相等的数据片段，并根据对象数据计算两个奇偶校验片段。六个片段中的每个片段都存储在三个数据中心站点的不同节点上，以便为节点故障或站点丢失提供数据保护。



4+2 纠删编码方案至少需要九个存储节点，三个不同站点中的每个站点各有三个存储节点。只要六个片段中的任意四个（数据或奇偶校验）仍然可用，就可以检索对象。最多可以丢失两个片段，而不会丢失对象数据。如果整个数据中心站点丢失，只要所有其他片段仍可访问，仍可检索或修复对象。



如果丢失两个以上的存储节点，则无法检索此对象。



相关信息

"什么是存储池"

"什么是纠删编码方案"

"配置擦除编码配置文件"

什么是纠删编码方案

为 ILM 规则配置纠删编码配置文件时，您可以根据计划使用的存储池中的存储节点和站点数量选择可用的纠删编码方案。纠删编码方案可控制为每个对象创建的数据片段数量和奇偶校验片段数量。

StorageGRID 系统使用 Reed-Solomon 纠删编码算法。该算法会将对象分段为 k 数据片段，并计算 m 奇偶校验片段。k + m = n 个片段分布在 n 个存储节点上，以提供数据保护。一个对象最多可承受丢失或损坏的碎片。检索或修复对象需要 K 个片段。

在配置擦除编码配置文件时，请对存储池遵循以下准则：

- 存储池必须包含三个或更多站点，或者只包含一个站点。

  > 如果存储池包含两个站点，则无法配置擦除编码配置文件。

  - 包含三个或更多站点的存储池的纠删编码方案
  - 单站点存储池的纠删编码方案
- 请勿使用默认存储池，所有存储节点或包含默认站点的存储池所有站点。
- 存储池应至少包含 *k*+*m*+1 个存储节点。

  所需的最小存储节点数为 *k*+*m* 。但是，如果所需的存储节点暂时不可用，则至少添加一个存储节点有助于防止载入失败或 ILM 回退。

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

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

例如，如果使用 4+2 方案存储一个 10 MB 的对象（存储开销为 50% ），则该对象将占用 15 MB 的网格存储。如果使用 6+2 方案存储同一个 10 MB 对象（存储开销为 33% ），则该对象将占用大约 13.3 MB 的空间。

选择总值最低的纠删编码方案 *k*+*m* ，以满足您的需求。碎片数量较少的纠删编码方案在计算方面总体上效率更高，因为每个对象创建和分布（或检索）的碎片数量较少，由于碎片大小较大，性能可能会更高，并且在需要更多存储时，扩展中添加的节点可能会更少。（有关规划存储扩展的信息，请参见有关扩展 StorageGRID 的说明。）

包含三个或更多站点的存储池的纠删编码方案

下表介绍了 StorageGRID 当前支持的纠删编码方案，该方案适用于包含三个或更多站点的存储池。所有这些方案均可提供站点丢失保护。一个站点可能会丢失，但对象仍可访问。

对于提供站点丢失保护的纠删编码方案，存储池中的建议存储节点数超过 *k*+*m*+1 ，因为每个站点至少需要三个存储节点。

纠删编码方案（ *k+m* ）	已部署站点的最小数量	每个站点的建议存储节点数	建议的存储节点总数	站点丢失保护？	存储开销
4+2	3.	3.	9	是的。	50%
6+2	4.	3.	12	是的。	33%
8+2	5.	3.	15	是的。	25%
6+3.	3.	4.	12	是的。	50%
9+3.	4.	4.	16.	是的。	33%
2+1	3.	3.	9	是的。	50%
4+1	5.	3.	15	是的。	25%
6+1	7.	3.	21	是的。	17%
7+5.	3.	5.	15	是的。	71%

StorageGRID 要求每个站点至少有三个存储节点。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纠删编码方案时，请平衡以下因素的相对重要性：

- * 碎片数量 * ： 当碎片总数减少时，性能和扩展灵活性通常会提高。
- * 容错 * ： 容错可通过包含更多奇偶校验分段来提高（即，当 *m* 的值较高时）。
- * 网络流量 * ： 从故障中恢复时，使用包含更多片段的方案（即 *k+m* 的总数更高）会创建更多网络流量。
- * 存储开销 * ： 开销较高的方案需要每个对象更多的存储空间。

例如，在选择 4+2 方案和 6+3 方案（两者都有 50% 的存储开销）时，如果需要额外的容错功能，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等，请选择 4+2 ，因为其碎片总数较低。

如果您不确定要使用的方案，请选择 4+2 或 6+3 ，或者联系技术支持。

单站点存储池的纠删编码方案

单站点存储池支持为三个或更多站点定义的所有纠删编码方案，但前提是该站点具有足够的存储节点。

所需的最小存储节点数为 *k+m* ，但建议使用具有 *k+m*+1 存储节点的存储池。例如， 2+1 纠删编码方案要求一个存储池至少包含三个存储节点，但建议使用四个存储节点。

纠删编码方案（ *k+m* ）	存储节点的最小数量	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%

相关信息

"扩展网格"

纠删编码的优势，劣势和要求

在决定是使用复制还是纠删编码来保护对象数据不会丢失之前，您应了解纠删编码的优点，缺点和要求。

纠删编码的优势

与复制相比，纠删编码可提高可靠性，可用性和存储效率。

- * 可靠性 * ： 可靠性通过容错来衡量—即，在不丢失数据的情况下可以同时发生的故障数量。通过复制，多个相同的副本会存储在不同的节点上以及不同的站点上。通过纠删编码，对象会编码为数据和奇偶校验片段，并分布在多个节点和站点上。这种分散方式可同时提供站点和节点故障保护。与复制相比，纠删编码可提高可靠性，而存储成本相当。
- * 可用性 * ： 可用性可定义为在存储节点出现故障或无法访问时检索对象的功能。与复制相比，纠删编码可以以相当的存储成本提高可用性。
- * 存储效率 * ： 对于相似级别的可用性和可靠性，通过纠删编码保护的对象比通过复制保护的相同对象占用的磁盘空间更少。例如，复制到两个站点的 10 MB 对象会占用 20 MB 的磁盘空间（两个副本），而采用 6+3 纠删编码方案在三个站点之间进行纠删编码的对象只会占用 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。

纠删编码的缺点

与复制相比，纠删编码具有以下缺点：

- 需要增加存储节点和站点的数量。例如，如果使用纠删编码方案 6+3 ，则必须在三个不同站点至少有三个存储节点。相比之下，如果只复制对象数据，则每个副本只需要一个存储节点。
- 存储扩展的成本和复杂性增加。要扩展使用复制的部署，只需在创建对象副本的每个位置添加存储容量即可。要扩展使用纠删编码的部署，您必须同时考虑使用的纠删编码方案以及现有存储节点的容量。例如，如果您等待现有节点达到 100% 全满，则必须至少添加 *k+m* 存储节点，但如果在现有节点达到 70% 全满时进行扩展，则可以为每个站点添加两个节点，同时仍可最大程度地提高可用存储容量。有关详细信息、请参见有关扩展StorageGRID 的说明。
- 在分布在不同地理位置的站点之间使用纠删编码时，检索延迟会增加。与在本地复制并提供的对象（客户端连接的同一站点）相比，通过 WAN 连接检索经过纠删编码并分布在远程站点上的对象的对象片段所需时间更长。
- 在地理位置分散的站点之间使用纠删编码时，检索和修复的 WAN 网络流量使用率较高，尤其是频繁检索的对象或通过 WAN 网络连接进行对象修复。
- 当您在站点间使用纠删编码时，最大对象吞吐量会随着站点间网络延迟的增加而急剧下降。这一减少是由于 TCP 网络吞吐量相应减少，从而影响 StorageGRID 系统存储和检索对象片段的速度。
- 提高计算资源的利用率。

何时使用纠删编码

纠删编码最适合以下要求：

- 大于1 MB的对象。

> 由于管理与纠删编码副本关联的片段数量会产生开销、因此请勿对200 KB或更小的对象使用纠删编码。

- 长期或冷存储，用于存储不经常检索的内容。
- 高数据可用性和可靠性。
- 防止发生完整的站点和节点故障。
- 存储效率。
- 需要高效数据保护的单站点部署，只需一个纠删编码副本，而不是多个复制副本。
- 站点间延迟小于 100 毫秒的多站点部署。

相关信息

"扩展网格"