



# **StorageGRID 如何实施S3 REST API**

## **StorageGRID 11.5**

NetApp  
April 11, 2024

# 目录

StorageGRID 如何实施S3 REST API .....	1
客户端请求冲突 .....	1
一致性控制 .....	1
StorageGRID ILM 规则如何管理对象 .....	4
对象版本控制 .....	5
实施S3 REST API的建议 .....	6

# StorageGRID 如何实施S3 REST API

客户端应用程序可以使用S3 REST API调用连接到StorageGRID 来创建、删除和修改存储分段以及存储和检索对象。

- ["客户端请求冲突"](#)
- ["一致性控制"](#)
- ["StorageGRID ILM 规则如何管理对象"](#)
- ["对象版本控制"](#)
- ["实施S3 REST API的建议"](#)

## 客户端请求冲突

冲突的客户端请求(例如、写入同一密钥的两个客户端)会按"latest-WINS"的原则进行解决。

"latest-WINS"评估的时间取决于StorageGRID 系统何时完成给定请求、而不是S3客户端何时开始操作。

## 一致性控制

一致性控制可以根据应用程序的要求，在对象的可用性与这些对象在不同存储节点和站点之间的一致性之间实现平衡。

默认情况下，StorageGRID 保证新创建的对象写入后读一致性。成功完成 PUT 后的任何 GET 都将能够读取新写入的数据。对现有对象的覆盖，元数据更新和删除最终保持一致。覆盖通常需要几秒钟或几分钟才能传播，但可能需要长达 15 天的时间。

如果要在不同的一致性级别执行对象操作，可以为每个分段或每个 API 操作指定一致性控制。

### 一致性控制

一致性控制会影响 StorageGRID 用于跟踪对象的元数据在节点之间的分布方式，从而影响对象用于客户端请求的可用性。

您可以将存储分段或 API 操作的一致性控制设置为以下值之一：

一致性控制	Description
全部	所有节点都会立即接收数据，否则请求将失败。
强大的全局功能	保证所有站点中所有客户端请求的写入后读一致性。
强大的站点	保证站点内所有客户端请求的写入后读一致性。

一致性控制	Description
读后写	<p>(默认) 为新对象提供写入后读一致性，并为对象更新提供最终一致性。提供高可用性和数据保护保证。与 Amazon S3 一致性保证匹配。</p> <ul style="list-style-type: none"> <li>注意：* 如果应用程序对不存在的对象使用 head 请求，则在一个或多个存储节点不可用时，可能会收到大量 500 个内部服务器错误。要防止出现这些错误，请将一致性控制设置为 "Available"，除非您需要类似于 Amazon S3 的一致性保证。</li> </ul>
可用 (机头操作的最终一致性)	<p>与 read-after-new-write 一致性级别相同，但仅为机头操作提供最终一致性。如果存储节点不可用，则为机头操作提供的可用性比 "read-after-new-write" 更高。与 Amazon S3 一致性保证不同，仅适用于机头操作。</p>

## 使用 "read-after-new-write" 和 "available" 一致性控制

当 head 或 get 操作使用 read-after-new-write 一致性控制或 get 操作使用 "available" 一致性控制时，StorageGRID 将执行多个步骤的查找，如下所示：

- 它首先使用低一致性查找对象。
- 如果该查找失败，它会在下一个一致性级别重复执行查找，直到达到最高一致性级别 "All"，这要求对象元数据的所有副本都可用。

如果 head 或 get 操作使用 read-after-new-write 一致性控制，但对象不存在，则对象查找将始终达到 "all" 一致性级别。由于此一致性级别要求对象元数据的所有副本均可用，因此，如果一个或多个存储节点不可用，您可能会收到大量 500 个内部服务器错误。

除非您需要类似于 Amazon S3 的一致性保证，否则可以通过将一致性控制设置为 "Available" 来防止机头操作出现这些错误。当机头操作使用 "Available" 一致性控制时，StorageGRID 仅提供最终一致性。它不会重试失败的操作，直到达到 "所有" 一致性级别为止，因此它不要求对象元数据的所有副本都可用。

## 指定 API 操作的一致性控制

要为单个 API 操作设置一致性控制，此操作必须支持一致性控制，并且必须在请求标题中指定一致性控制。此示例将 GET 对象操作的一致性控制设置为 strong-site。

```
GET /bucket/object HTTP/1.1
Date: date
Authorization: <em>authorization name</em>
Host: <em>host</em>
Consistency-Control: strong-site
```



PUT 对象和 GET 对象操作必须使用相同的一致性控制。

## 指定存储分段的一致性控制

要为存储分段设置一致性控制，您可以使用 StorageGRID PUT 存储分段一致性请求和 GET 存储分段一致性请求。或者，您也可以使用租户管理器或租户管理 API。

为存储分段设置一致性控制时，请注意以下事项：

- 设置存储分段的一致性控制可确定对存储分段中的对象或存储分段配置执行 S3 操作时使用的一致性控制。它不会影响存储分段本身的操作。
- 单个 API 操作的一致性控制将覆盖存储分段的一致性控制。
- 通常，存储分段应使用默认一致性控制 `read-after-new-write`。如果请求无法正常工作，请尽可能更改应用程序客户端的行为。或者，配置客户端以指定每个 API 请求的一致性控制。请仅作为最后一种方法在存储分段级别设置一致性控制。

## 一致性控制和 ILM 规则如何交互以影响数据保护

您选择的一致性控制和 ILM 规则都会影响对象的保护方式。这些设置可以进行交互。

例如，存储对象时使用的一致性控制会影响对象元数据的初始放置，而为 ILM 规则选择的载入行为会影响对象副本的初始放置。由于 StorageGRID 需要访问对象的元数据及其数据来满足客户端请求，因此为一致性级别和载入行为选择匹配的保护级别可以提供更好的初始数据保护和更可预测的系统响应。

ILM 规则可以使用以下载入行为：

- **\* 严格 \***：必须创建 ILM 规则中指定的所有副本，才能将成功返回给客户端。
- **\* 平衡 \***：StorageGRID 尝试在载入时创建 ILM 规则中指定的所有副本；如果无法创建，则创建临时副本并将成功返回给客户端。在可能的情况下，将创建 ILM 规则中指定的副本。
- **\* 双提交 \***：StorageGRID 会立即为对象创建临时副本，并将成功返回给客户端。如果可能，将创建 ILM 规则中指定的副本。



在为 ILM 规则选择载入行为之前，请阅读有关通过信息生命周期管理管理对象的说明中有关这些设置的完整问题描述。

## 一致性控制和 ILM 规则如何交互的示例

假设您有一个双站点网格，其中包含以下 ILM 规则和以下一致性级别设置：

- **\* ILM 规则 \***：创建两个对象副本，一个在本地站点，一个在远程站点。此时将选择严格的载入行为。
- **\* 一致性级别 \***：`strong-global`（对象元数据会立即分发到所有站点。）

当客户端将对象存储到网格时，StorageGRID 会创建两个对象副本并将元数据分发到两个站点，然后再向客户端返回成功。

在载入成功消息时，此对象将受到完全保护，不会丢失。例如，如果本地站点在载入后不久丢失，则远程站点上仍存在对象数据和对象元数据的副本。此对象完全可检索。

如果您改用相同的 ILM 规则和 `strong-site` 一致性级别，则在将对象数据复制到远程站点之后、但在该站点分发对象元数据之前、客户端可能会收到一条成功消息。在这种情况下，对象元数据的保护级别与对象数据的保护级别不匹配。如果本地站点在载入后不久丢失，则对象元数据将丢失。无法检索此对象。

一致性级别和 ILM 规则之间的关系可能很复杂。如需帮助，请联系 NetApp。

相关信息

["使用 ILM 管理对象"](#)

["获取存储分段一致性请求"](#)

["PUT 存储分段一致性请求"](#)

## StorageGRID ILM 规则如何管理对象

网络管理员创建信息生命周期管理（ILM）规则，用于管理从 S3 REST API 客户端应用程序输入到 StorageGRID 系统中的对象数据。然后，这些规则将添加到 ILM 策略中，以确定随着时间的推移对象数据的存储方式和存储位置。

ILM 设置可确定对象的以下方面：

- \* 地理位置 \*

对象数据在 StorageGRID 系统（存储池）或云存储池中的位置。

- \* 存储级别 \*

用于存储对象数据的存储类型：例如闪存或旋转磁盘。

- \* 损失保护 \*

创建的副本数量以及创建的副本类型：复制，纠删编码或两者。

- \* 保留 \*

随着时间的推移，对象数据的管理方式，存储位置以及保护数据免受丢失的方式会发生变化。

- \* 载入期间的保护 \*

用于在载入期间保护对象数据的方法：同步放置（使用平衡或严格的 InGest 行为选项）或创建中间副本（使用双提交选项）。

ILM 规则可以筛选和选择对象。对于使用 S3 载入的对象，ILM 规则可以根据以下元数据筛选对象：

- 租户帐户
- Bucket Name
- 载入时间
- 密钥
- 上次访问时间



默认情况下，对所有 S3 分段禁用上次访问时间更新。如果您的 StorageGRID 系统包含使用上次访问时间选项的 ILM 规则，则必须为该规则中指定的 S3 分段启用对上次访问时间的更新。您可以使用租户管理器中的 PUT 分段上次访问时间请求，\* S3\* > \* 分段 \* > \* 配置上次访问时间 \* 复选框或租户管理 API 启用上次访问时间更新。启用上次访问时间更新时，请注意 StorageGRID 性能可能会降低，尤其是在具有小型对象的系统中。

- 位置限制
- 对象大小
- 用户元数据
- 对象标记

有关 ILM 的详细信息，请参见有关通过信息生命周期管理管理对象的说明。

相关信息

["使用租户帐户"](#)

["使用 ILM 管理对象"](#)

["PUT 分段上次访问时间请求"](#)

## 对象版本控制

您可以使用版本控制来保留一个对象的多个版本，从而防止意外删除对象，并可用于检索和还原对象的早期版本。

StorageGRID 系统实施版本控制，并支持大多数功能，但存在一些限制。StorageGRID 最多支持 1,000 个对象版本。

对象版本控制可以与 StorageGRID 信息生命周期管理（ILM）或 S3 存储分段生命周期配置结合使用。要为每个存储分段启用此功能，您必须明确启用版本控制。存储分段中的每个对象都分配有一个版本 ID，该 ID 由 StorageGRID 系统生成。

不支持使用 MFA（多因素身份验证）Delete。



只能在使用 StorageGRID 10.3 或更高版本创建的存储分段上启用版本控制。

## ILM 和版本控制

ILM 策略将应用于对象的每个版本。ILM 扫描过程会持续扫描所有对象，并根据当前 ILM 策略重新评估这些对象。对 ILM 策略所做的任何更改都会应用于先前载入的所有对象。如果启用了版本控制，则包括先前载入的版本。ILM 扫描会将新的 ILM 更改应用于先前输入的对象。

对于启用了版本控制的分段中的 S3 对象，版本控制支持允许您创建使用非当前时间作为参考时间的 ILM 规则。对象更新后，其先前版本将变为非最新版本。使用非当前时间筛选器可以创建策略，以减少先前版本的对象对存储的影响。



使用多部分上传操作上传新版本的对象时，原始版本对象的非当前时间反映为新版本创建多部分上传的时间，而不是多部分上传完成的时间。在有限情况下，原始版本的非当前时间可能比当前版本的时间早数小时或数天。

有关 S3 版本控制对象的示例 ILM 策略，请参见有关通过信息生命周期管理管理对象的说明。

相关信息

["使用 ILM 管理对象"](#)

## 实施S3 REST API的建议

在实施用于 StorageGRID 的 S3 REST API 时，应遵循以下建议。

### 针对不存在的对象的建议

如果您的应用程序定期检查某个对象是否位于您不希望该对象实际存在的路径上，则应使用 "available" 一致性控制。例如，如果您的应用程序在放置到某个位置之前一直位于某个位置，则应使用 "Available" 一致性控制。

否则，如果 head 操作未找到对象，则在一个或多个存储节点不可用时，可能会收到大量 500 个内部服务器错误。

您可以使用 PUT 存储分段一致性请求为每个存储分段设置 "Available" 一致性控制，也可以在单个 API 操作的请求标题中指定一致性控制。

### 对象密钥建议

对于在 StorageGRID 11.4 或更高版本中创建的分段，不再需要限制对象密钥名称以满足性能最佳实践。例如，现在可以对对象密钥名称的前四个字符使用随机值。

对于在 StorageGRID 11.4 之前的版本中创建的分段，请继续对对象密钥名称遵循以下建议：

- 不应使用随机值作为对象密钥的前四个字符。这与 AWS 以前针对密钥前缀的建议不同。而应使用非随机、非唯一前缀、例如 image。
- 如果您按照以前的 AWS 建议在密钥前缀中使用随机和唯一字符，则应在对象密钥前添加目录名称。也就是说，请使用以下格式：

```
mybucket/mydir/f8e3-image3132.jpg
```

而不是以下格式：

```
mybucket/f8e3-image3132.jpg
```



## 关于“范围读取”的建议

如果选择了\*压缩存储的对象\*选项(配置>\*网格选项\*)、则S3客户端应用程序应避免执行指定要返回的字节数范围的GET对象操作。这些“range read”操作效率低下，因为 StorageGRID 必须有效解压缩对象以访问请求的字节。从非常大的对象请求少量字节的 GET 对象操作效率尤其低下；例如，从 50 GB 压缩对象读取 10 MB 范围的操作效率非常低。

如果从压缩对象读取范围，则客户端请求可能会超时。



如果需要压缩对象，并且客户端应用程序必须使用范围读取，请增加应用程序的读取超时时间。

相关信息

["一致性控制"](#)

["PUT 存储分段一致性请求"](#)

["管理 StorageGRID"](#)

## 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。