



使用搜索集成服务

StorageGRID 11.8

NetApp
May 10, 2024

目录

使用搜索集成服务	1
用于搜索集成的配置 XML	1
配置搜索集成服务	5
由搜索集成服务生成的 JSON	7
元数据通知中包含的对象元数据	7

使用搜索集成服务

搜索集成服务是三种 StorageGRID 平台服务之一。您可以启用此服务，以便在创建，删除对象或更新其元数据或标记时将对象元数据发送到目标搜索索引。

您可以使用租户管理器将自定义 StorageGRID 配置 XML 应用于存储分段来配置搜索集成。



由于搜索集成服务会将对象元数据发送到目标，因此其配置 XML 称为 *metadata notification configuration xml*。此配置 XML 与用于启用事件通知的 *notification 配置 xml* 不同。

请参见 ["有关实施 S3 客户端应用程序的说明"](#) 有关以下自定义 StorageGRID S3 REST API 操作的详细信息：

- 删除存储分段元数据通知配置
- 获取存储分段元数据通知配置
- PUT 存储分段元数据通知配置

相关信息

["用于搜索集成的配置 XML"](#)

["元数据通知中包含的对象元数据"](#)

["由搜索集成服务生成的 JSON"](#)

["配置搜索集成服务"](#)

["使用S3 REST API"](#)

用于搜索集成的配置 XML

搜索集成服务使用中包含的一组规则进行配置

`<MetadataNotificationConfiguration>` 和 `</MetadataNotificationConfiguration>` 标记。每个规则都指定规则适用场景 所对应的对象以及 StorageGRID 应将这些对象的元数据发送到的目标。

可以按对象名称的前缀筛选对象。例如、您可以发送具有前缀的对象的元数据 `images` 一个目标、并为具有前缀的对象提供元数据 `videos` 另一个。前缀重叠的配置无效、在提交时将被拒绝。例如、一种配置、其中包含一个前缀为的对象规则 `test` 和第二个规则、用于具有前缀的对象 `test2` 不允许。

必须使用为搜索集成服务创建的 StorageGRID 端点的 URN 指定目标。这些端点是指 Elasticsearch 集群上定义的索引和类型。

```

<MetadataNotificationConfiguration>
  <Rule>
    <ID>Rule-1</ID>
    <Status>rule-status</Status>
    <Prefix>key-prefix</Prefix>
    <Destination>
      <Urn>arn:aws:es:region:account-
ID:domain/mydomain/myindex/mytype</Urn>
    </Destination>
  </Rule>
  <Rule>
    <ID>Rule-2</ID>
    ...
  </Rule>
  ...
</MetadataNotificationConfiguration>

```

下表介绍了元数据通知配置 XML 中的元素。

Name	Description	Required
MetadataNotificationConfiguration	用于指定元数据通知的对象和目标的规则的容器标记。 包含一个或多个规则元素。	是的。
规则	用于标识应将其元数据添加到指定索引中的对象的规则的容器标记。 拒绝前缀重叠的规则。 包含在 MetadataNotificationConfiguration 元素中。	是的。
ID	规则的唯一标识符。 包含在 Rule 元素中。	否
Status	状态可以是 " 已启用 " 或 " 已禁用 "。不会对已禁用的规则执行任何操作。 包含在 Rule 元素中。	是的。

Name	Description	Required
前缀	<p>与前缀匹配的对象受此规则的影响，其元数据将发送到指定目标。</p> <p>要匹配所有对象，请指定一个空前缀。</p> <p>包含在 Rule 元素中。</p>	是的。
目标	<p>规则目标的容器标记。</p> <p>包含在 Rule 元素中。</p>	是的。
URN	<p>发送对象元数据的目标的 urn 。必须具有以下属性的 StorageGRID 端点的 URN：</p> <ul style="list-style-type: none"> • es 必须是第三个元素。 • URN必须以存储元数据的索引和类型结尾、格式为 domain-name/myindex/mytype。 <p>端点使用租户管理器或租户管理 API 进行配置。它们的形式如下：</p> <ul style="list-style-type: none"> • arn:aws:es:region:account-ID:domain/mydomain/myindex/mytype • urn:mysite:es:::mydomain/myindex/mytype <p>必须在提交配置 XML 之前配置端点，否则配置将失败并显示 404 错误。</p> <p>urn 包含在目标元素中。</p>	是的。

使用示例元数据通知配置 XML 了解如何构建自己的 XML。

用于适用场景 所有对象的元数据通知配置

在此示例中，所有对象的对象元数据都将发送到同一目标。

```

<MetadataNotificationConfiguration>
  <Rule>
    <ID>Rule-1</ID>
    <Status>Enabled</Status>
    <Prefix></Prefix>
    <Destination>
      <Urn>urn:myes:es:::sgws-notifications/test1/all</Urn>
    </Destination>
  </Rule>
</MetadataNotificationConfiguration>

```

具有两个规则的元数据通知配置

在此示例中，是指与前缀匹配的对象的对象元数据 /images 发送到一个目标、而与前缀匹配的对象的对象元数据则发送到一个目标 /videos 发送到另一个目标。

```

<MetadataNotificationConfiguration>
  <Rule>
    <ID>Images-rule</ID>
    <Status>Enabled</Status>
    <Prefix>/images</Prefix>
    <Destination>
      <Urn>arn:aws:es:us-east-1:3333333:domain/es-
domain/graphics/imagetype</Urn>
    </Destination>
  </Rule>
  <Rule>
    <ID>Videos-rule</ID>
    <Status>Enabled</Status>
    <Prefix>/videos</Prefix>
    <Destination>
      <Urn>arn:aws:es:us-west-1:2222222:domain/es-
domain/graphics/videotype</Urn>
    </Destination>
  </Rule>
</MetadataNotificationConfiguration>

```

相关信息

["使用S3 REST API"](#)

["元数据通知中包含的对象元数据"](#)

["由搜索集成服务生成的 JSON"](#)

配置搜索集成服务

每当创建，删除对象或更新其元数据或标记时，搜索集成服务都会将对象元数据发送到目标搜索索引。

开始之前

- StorageGRID 管理员已为租户帐户启用平台服务。
- 您已创建要为其内容编制索引的S3存储分段。
- 要用作搜索集成服务目标的端点已存在、并且您具有其URN。
- 您属于具有的用户组 **"管理所有分段或root访问权限"**。使用租户管理器配置存储分段时，这些权限会覆盖组或存储分段策略中的权限设置。

关于此任务

为源存储分段配置搜索集成服务后，创建对象或更新对象的元数据或标记会触发要发送到目标端点的对象元数据。如果为已包含对象的存储分段启用搜索集成服务、则不会自动为现有对象发送元数据通知。您必须更新这些现有对象，以确保其元数据已添加到目标搜索索引中。

步骤

1. 使用文本编辑器创建启用搜索集成所需的元数据通知 XML 。
 - 请参见有关用于搜索集成的配置 XML 的信息。
 - 配置 XML 时，请使用搜索集成端点的 URN 作为目标。

```
<MetadataNotificationConfiguration>
  <Rule>
    <Status>Enabled</Status>
    <Prefix></Prefix>
    <Destination>
      <Urn>arn:aws:es:us-east-
1:1111111111111111:domain/mydomain/myindex/mytype</Urn>
    </Destination>
  </Rule>
</MetadataNotificationConfiguration>
```

2. 在租户管理器中，选择 * 存储 (S3) * > * 分段 * 。
3. 选择源存储分段的名称。

此时将显示存储分段详细信息页面。

4. 选择 * 平台服务 * > * 搜索集成 *
5. 选中*启用搜索集成*复选框。
6. 将元数据通知配置粘贴到文本框中，然后选择 * 保存更改 * 。

Bucket options
Bucket access
Platform services

Replication
Disabled
▼

Event notifications
Disabled
▼

Search integration
Disabled
▲

Enable the search integration service to send object metadata to a destination search index whenever an object is created, deleted, or its metadata or tags are updated.

- Platform services must be enabled for your tenant account by a StorageGRID administrator.
- You must have already configured an endpoint for the search integration service.
- You must specify the URN of that endpoint in the search integration configuration XML for the bucket you want to index.

Enable search integration

Clear

```

<MetadataNotificationConfiguration>
  <Rule>
    <Status>Enabled</Status>
    <Prefix></Prefix>
    <Destination>
      <Urn>arn:aws:es:us-east-1:111111111111:domain/mydomain/myindex/mytype</Urn>
    </Destination>
  </Rule>
</MetadataNotificationConfiguration>
          
```

Save changes



StorageGRID 管理员必须使用网格管理器或管理 API 为每个租户帐户启用平台服务。如果保存配置 XML 时发生错误，请联系 StorageGRID 管理员。

7. 验证是否已正确配置搜索集成服务：

- a. 向源存储分段添加一个对象，以满足配置 XML 中指定的元数据通知触发要求。

在前面显示的示例中，添加到存储分段的所有对象都会触发元数据通知。

- b. 确认包含对象元数据和标记的 JSON 文档已添加到端点中指定的搜索索引中。

完成后

根据需要，您可以使用以下任一方法禁用存储分段的搜索集成：

- 选择*storage (S3)*>*Bucbes*并清除*Enable search integration*复选框。
- 如果您直接使用 S3 API ，请使用删除分段元数据通知请求。请参见有关实施 S3 客户端应用程序的说明。

相关信息

["了解搜索集成服务"](#)

["用于搜索集成的配置 XML"](#)

["使用S3 REST API"](#)

["创建平台服务端点"](#)

由搜索集成服务生成的 JSON

为存储分段启用搜索集成服务后，每次添加，更新或删除对象元数据或标记时，系统都会生成一个 JSON 文档并将其发送到目标端点。

此示例显示了使用密钥的对象时可能生成的JSON示例 SGWS/Tagging.txt 在名为的存储分段中创建 test。。 test 存储分段未进行版本控制、因此 versionId 标记为空。

```
{
  "bucket": "test",
  "key": "SGWS/Tagging.txt",
  "versionId": "",
  "accountId": "86928401983529626822",
  "size": 38,
  "md5": "3d6c7634a85436eee06d43415012855",
  "region": "us-east-1",
  "metadata": {
    "age": "25"
  },
  "tags": {
    "color": "yellow"
  }
}
```

元数据通知中包含的对象元数据

下表列出了启用搜索集成后发送到目标端点的 JSON 文档中包含的所有字段。

文档名称包括存储分段名称，对象名称和版本 ID（如果存在）。

Type	项目名称和问题描述
存储分段和对象信息	bucket: 存储分段的名称
key: 对象密钥名称	versionID: 对象版本、用于受版本控制的分段中的对象
region: 例如、Bucket区域 us-east-1	系统元数据
size: HTTP客户端可见的对象大小(以字节为单位)	md5: 对象哈希
用户元数据	metadata: 对象的所有用户元数据、作为键值对 key:value
Tags	tags: 为对象定义的所有对象标记、作为键值对 key:value



对于标记和用户元数据，StorageGRID 会将日期和数字作为字符串或 S3 事件通知传递给 Elasticsearch。要配置 Elasticsearch 以将这些字符串解释为日期或数字，请按照 Elasticsearch 说明进行动态字段映射和映射日期格式。在配置搜索集成服务之前，必须在索引上启用动态字段映射。为文档编制索引后、无法在索引中编辑文档的域类型。

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。