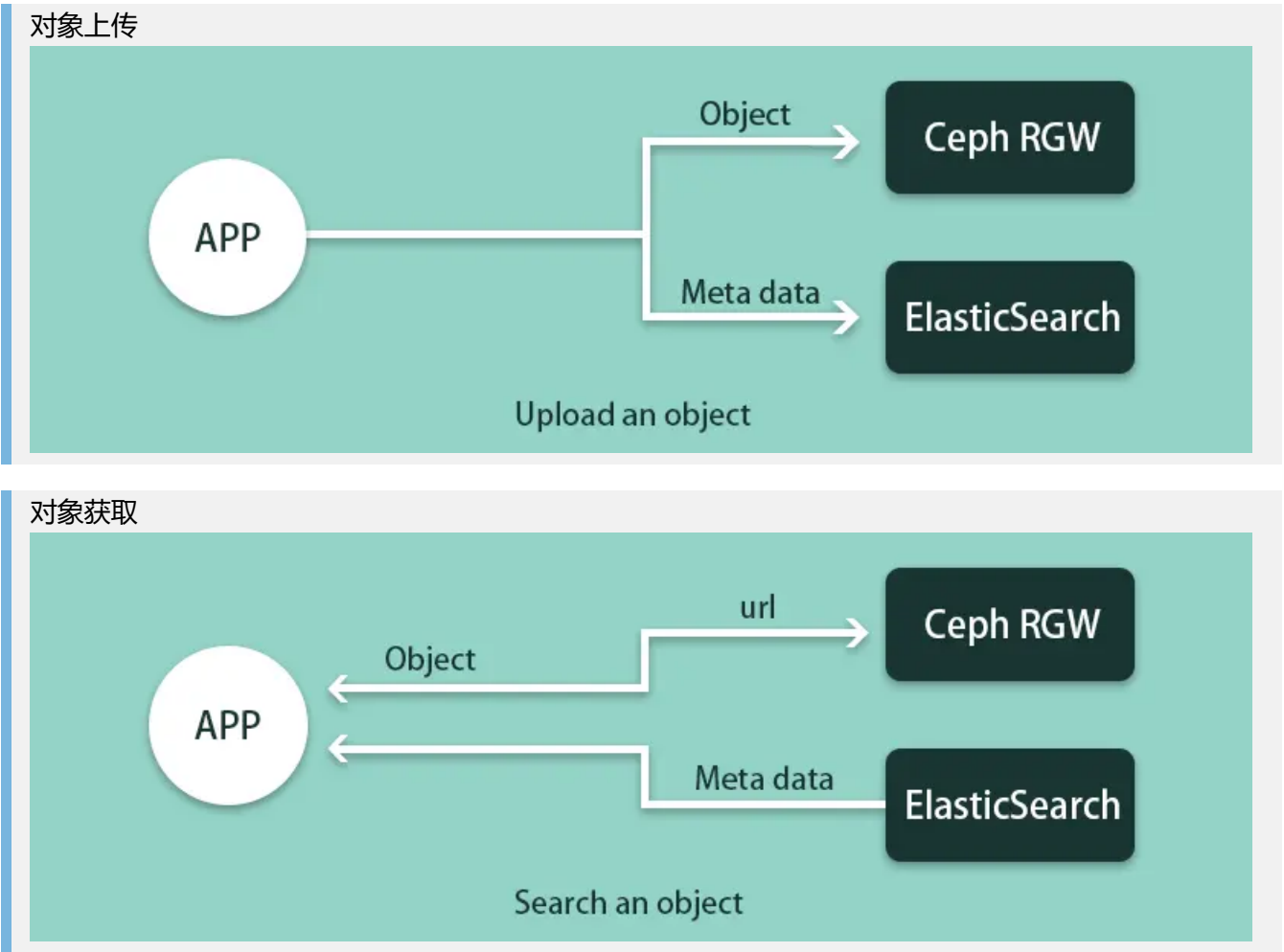


目的

- Ceph 本身并没有元数据搜索的功能,不能索引元数据来访问对象。
- 下面介绍使用ElasticSearch 来实现Ceph的元数据搜索功能。
-

方案一

- 前端应用将对象上传给ceph RGW中，并把这个对象的自定义元数据发送到ElasticSearch集群中，
- 当用户需要获取某个对象时，先向ElasticSeach 发送搜索请求来获取对象地址
 - 该请求可以是对象名或是对象id号，或者是用户自定义的元数据。
 - Elastic返回这个对象的地址后，前端应用通过这个地址来直接向Ceph RGW获取对象本身。

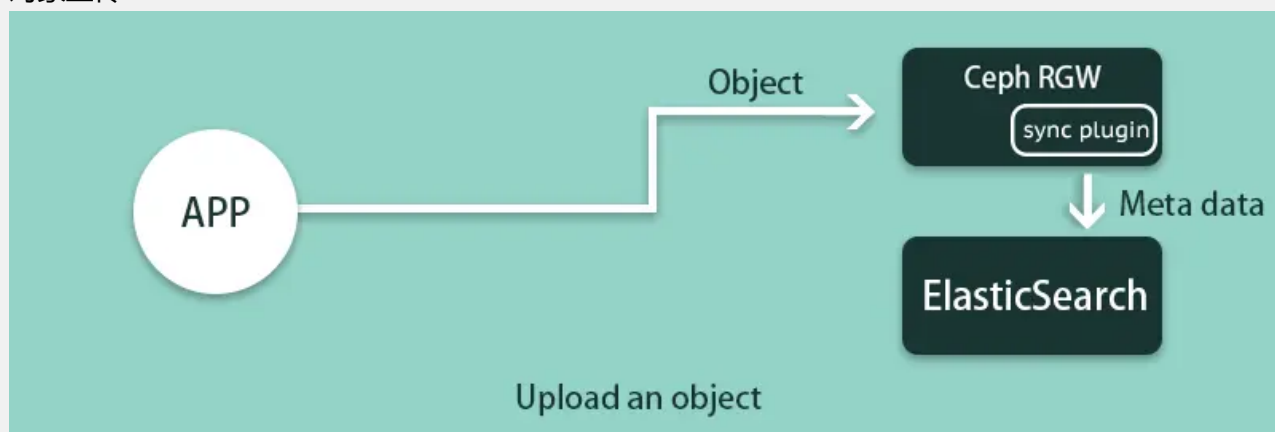


- RGW与ElasticSearch之间的数据一致性完全依赖于前端应用。S3除了SDK以外还有s3cmd这个命令行工具，RGW自身也可以通过HTTP请求来上传对象，用户完全有可能直接在后台上传对象，而不是通过前端应用，这样就造成了ElasticSearch无法同步对应的元数据。

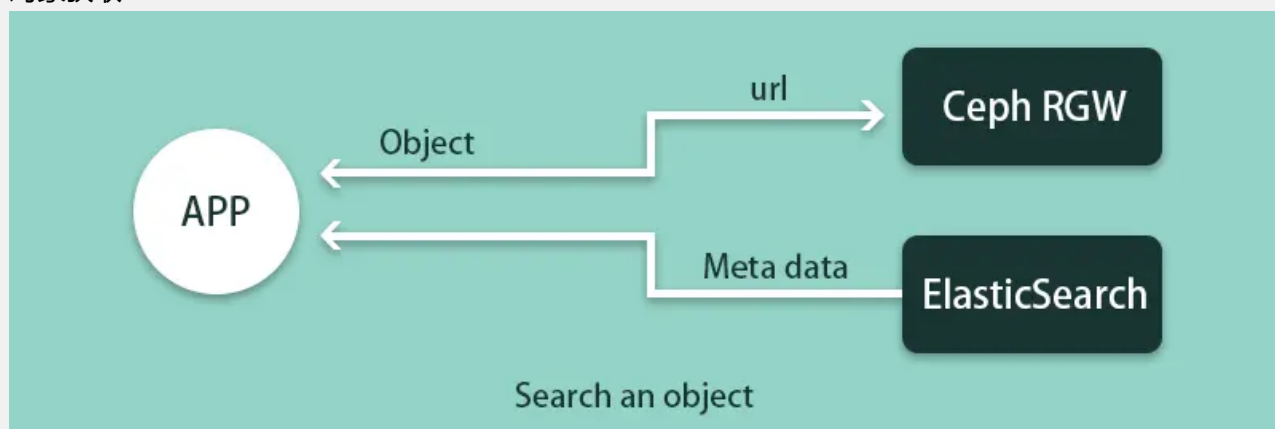
方案二

- Ceph在Jewel版本之后加入对ElasticSearch的支持，我们可以通过定义新的zone类型来实现RGW元数据自动同步到ElasticSearch中，这样一来便保证了RGW和ElasticSearch数据的一致性，让前后端耦合大大降低。

对象上传

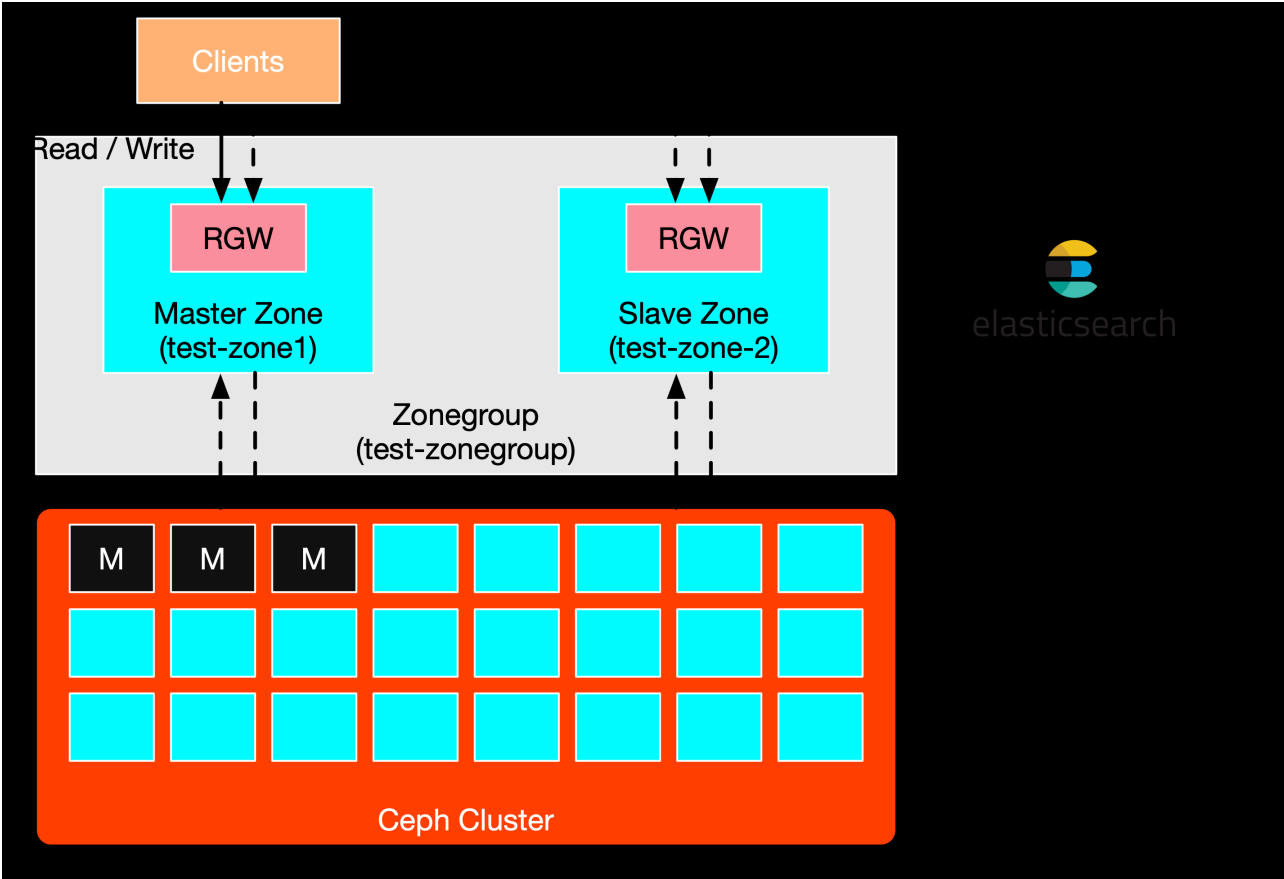


对象获取



- 方案一和方案二的唯一区别在与对象上传：前端上传对象时不再需要上传元数据到ElasticSearch，Ceph RGW中自带的sync plugin可以自动将元数据同步到ElasticSearch中去。
- 对于第二种方案一般和ceph 对象存储的多站点同步中同步扩展模块结合起来进行元数据的同步
 - 同步扩展模块构建在多站点框架的基础上，可将数据和元数据转发到不同的外部层。
 - 利用同步扩展模块，可让系统在每当有数据发生更改（例如执行存储桶或用户创建等元数据操作）时即执行一组操作。
 - 当对象网关多站点更改最终在远程站点上保持一致时，更改将以异步方式传播。因而很多情况下都适合使用同步扩展模块，例如将对象存储备份到外部云集群、使用磁带机的自定义备份解决方案，或在 ElasticSearch 中为元数据编制索引。

- 架构如下



- 建立两个zone, 一个是提供使用者利用 s3 potocal 通过 Radosgw 对 Ceph Object Storage 进行操作, 另一个 Zone 是同步数据使用, 并将Metadata 存储至 Elasticsearch 上。

<https://www.jianshu.com/p/e1b65ca2e095> 参考