Object WORM Feature in Ceph Rados Gateway

Zhang Shaowen
China Mobile (Suzhou) software technology Co., Ltd
2019/5
1. What is WORM
2. Why is WORM
3. S3 Object Lock
4. Ceph implementation
What is WORM

Write once read many (WORM)

• Non-rewritable

• Non-erasable

• No restrictions on reading
Types of WORM

Physical WORM

- specific storage media
- CD-R and DVD-R

Logical WORM

- non-WORM storage media
- disk driver
Why is WORM

• Secure access to data
• Compliance
• Governance
Object Lock
- Object or bucket control
- Retention
- Legal Hold

Data Protection and Compliance
- SEC 17a-4, CFTC, and FINRA
- Protection against accidental or malicious delete
Enable Object Lock

Enabled on Bucket
• Only available for new buckets
• Once enabled, no disable

Versioning
• Automatically enabled
• No Suspend
Retention period & target

- Bucket
  - Days
  - Years

- Object
  - Date
Retention Mode

• Governance
  • Intended for data protection
  • Enables privileged delete
  • Retention can be changed to Compliance Mode
  • `s3:BypassGovernanceMode`
  • `x-amz-bypass-governance-retention:true`

• Compliance
  • Intended for Compliance
  • Deletes disallowed
Legal Holds

- Set on object
  - `s3:PutObjectLegalHold`
- Work until removed
- Independent with Retention
Add Object Lock API compatible with S3

- Put/GET Bucket Object Lock
  - set/get retention period on bucket

- Put/GET Object Retention
  - set/get retention date on specific object

- Put/GET Object Legal Hold
  - set/get legal hold on specific object

```xml
<ObjectLockConfiguration>
  <ObjectLockEnabled>Enabled</ObjectLockEnabled>
  <Rule>
    <DefaultRetention>
      <Mode>GOVERNANCE</Mode>
      <Years>1</Years>
    </DefaultRetention>
  </Rule>
</ObjectLockConfiguration>

<Retention>
  <Mode>GOVERNANCE</Mode>
  <RetainUntilDate>2020-01-05T00:00:00.000Z</RetainUntilDate>
</Retention>

<LegalHold>
  <Status>ON</Status>
</LegalHold>
```
WORM in RGW

Reuse flag in RGWBucketInfo

```cpp
RGWBucketFlags{
    BUCKET_SUSPENDED = 0x1,
    BUCKET_VERSIONED = 0x2,
    BUCKET_VERSIONS_SUSPENDED = 0x4,
    BUCKET_DATASYNC_DISABLED = 0x8,
    BUCKET_MFA_ENABLED = 0x10,
    BUCKET_OBJ_LOCK_ENABLED = 0x20,
}

bool obj_lock_enabled()
{
    return (flags&
        BUCKET_OBJ_LOCK_ENABLED)!=0
}
```

Add configuration to RGWBucketInfo

```cpp
class RGWObjectLock
{
    bool enabled;
    bool rule_exist;
    ObjectLockRule rule;
}

class ObjectLockRule
{
    DefaultRetention defaultRetention
}

class DefaultRetention
{
    string mode;
    days;
    years;
}
WORM in RGW

Retention
• mode
• date
• user.rgw.object-retention

Legal holds
• status(ON/OFF)
• user.rgw.object-legal-hold

class RGWObjectRetention
{
  string mode;
  ceph::real_time retain_until_date;
}

class RGWObjectLegalHold
{
  string status;
}
Further Work

- Need test cases, including S3-tests
- Improve working with policy

https://github.com/ceph/ceph/pull/26538
Thanks!