OTM Repository Gateway Specification

The preservation Gateway functions as an aggregating cache for preservation requests originating with a repository and destined for a DDP via the OTM Bridge. This allows repositories to *synchronously* send content preservation in an *asynchronous* deep archive.

Repository Gateway API

GET / (service description)

Description

Retrieve a description of this Gateway, including the api version, and the available storage classes.

Requests/Responses

- Request: `GET / HTTP/1.1`
- Request Body: n/a
- Request Headers:
    - `Accept`
- Response Codes:

    - `200 OK`
- Response Body:

> **JSON response body**
>
> ```
> {
>   api_version: 0.1
>   storage_classes: [{
>     name: ddp1_name
>   },
>   {
>     name: ddp2_name
>   }]
> }
> ```

PUT Object (deposit)

Description

Create or update a gateway object for preservation of an Object from the repository.

Requests/Responses

- Request: `PUT /ObjectId HTTP/1.1`
- Request Body: Compressed Bag
- Request Headers:
    - `Content-Type`
    - `Content-Length`
    - `x-duraspace-storage-class` (DDP target(s) for preservation)
- Response Codes:

    - `201 Created` (on success)
    - `200 OK` (on success; update)
- Response Headers:
    - `ETag`

Examples

Create an Object

Request

```
PUT /af48c3d HTTP/1.1
Host: preservation-gateway.institution.edu
Date: Tue, 02 Jul 2019 20:15:00 GMT
Content-Type: application/zip
Content-Length: 493285
Content-MD5: 4efcb3d98ce0fabfd585eb6c4332859

[493285 bytes of object data]
```

Response

```
HTTP/1.1 201 Created
Date: Tue, 02 Jul 2019 20:15:00 GMT
ETag: "4efcb3d98ce0fabfd585eb6c4332859"
Content-Length: 0
Server: OTM Preservation Gateway
```

Question

Holy Bags? For large file deposit.

 - i don't want to push this content over the network and/or

 - i don't want to store it on a Gateway accessible disk
Write up something about ObjectID semantics and repository reconstruction.

GET Object Audit

Description

Retrieve audit metadata for the object.

Requests/Responses

- Request: `GET /ObjectId/audit HTTP/1.1`
- Request Body: n/a
- Response Code: `200 OK`
- Response Body:

  JSON response body

  ```
  {
    ddp_name: {
      preserved: true
      destroy_requested: true
    }
  }
  ```

DELETE Object (destroy)

Description

Request removal of an Object and all its contents from the archive

Requests/Responses

- Request: `DELETE /ObjectId HTTP/1.1`
- Request Body: n/a
- Response Codes:
    - `204 No Content`

Examples

Destroy an Object

Request

```
DELETE /af48c3d HTTP/1.1
Host: preservation-gateway.institution.edu
Date: Tue, 02 Jul 2019 20:25:00 GMT
Content-Type: text/plain
```

Response

```
HTTP/1.1 204 NoContent
Date: Tue, 02 Jul 2019 20:25:00 GMT
Content-Length: 0
Server: OTM Preservation Gateway
```

POST Object Restore

Description

Request restore of a Bag Object and all its contents for later retrieval

Requests/Responses

- Request: `POST /ObjectId?restore HTTP/1.1`
- Request Body: n/a
- Response Code:
    - `202 Accepted` (if the object is not already restored)
    - `200 OK` (if the object is already restored and available for retrieval)

Examples

Restore an Object

Request

```
POST /af48c3d?restore HTTP/1.1
Host: preservation-gateway.institution.edu
Date: Tue, 02 Jul 2019 20:35:00 GMT
Content-Length: 0
```

Response

```
HTTP/1.1 202 Accepted
Date: Tue, 02 Jul 2019 20:35:00 GMT
Content-Length: 0
Server: OTM Preservation Gateway
```

Requesting a Restore on an Object that is Available

Request

```
POST /af48c3d?restore HTTP/1.1
Host: preservation-gateway.institution.edu
Date: Tue, 02 Jul 2019 20:35:00 GMT
Content-Length: 0
```

Response

```
HTTP/1.1 200 OK
Date: Tue, 02 Jul 2019 20:35:00 GMT
Content-Length: 0
Server: OTM Preservation Gateway
```

ERROR: Restore Already in Progress

Request

```
POST /af48c3d?restore HTTP/1.1
Host: preservation-gateway.institution.edu
Date: Tue, 02 Jul 2019 20:35:00 GMT
Content-Length: 0
```

Response

```
HTTP/1.1 409 Conflict
Date: Tue, 02 Jul 2019 20:35:00 GMT
Content-Type: application/xml
Server: OTM Preservation Gateway
<?xml version="1.0" encoding="UTF-8"?>
<Error>
  <Code>RestoreAlreadyInProgress</Code>
  <Message>Object restore is already in progress.</Message>
  <Resource>/af48c3d</Resource>
</Error>
```

Response

```
HTTP/1.1 409 Conflict
Date: Tue, 02 Jul 2019 20:35:00 GMT
Content-Type: application/xml
Server: OTM Preservation Gateway

<?xml version="1.0" encoding="UTF-8"?>
<Error>
  <Code>RestoreAlreadyInProgress</Code>
  <Message>Object restore is already in progress.</Message>
  <Resource>/af48c3d</Resource>
</Error>
```

Question
ⓘ
Can HTTP Range requests be used to retrieve individual files (e.g. metadata)

GET Object (retrieve)

Description

Retrieve a restored object.

Requests/Responses

- Request: `GET /ObjectId HTTP/1.1`
- Request Body: n/a
- Request Headers
    - `If-Match`
    - `If-None-Match`
    - `Accept` ❓
- Response Code:
    - `200 OK` (if the object is available)
    - `403 Forbidden` (if the object has not been restored)
    - `412 PreconditionFailed`
- Response Body: Compressed bag
- Response Headers:
    - `Content-Type`
    - `ETag`

Examples

Retrieve an Object

Request

```
GET /af48c3d HTTP/1.1
Host: preservation-gateway.institution.edu
Date: Tue, 02 Jul 2019 20:45:00 GMT
Content-Length: 0
```

Response

```
HTTP/1.1 200 OK
Date: Tue, 02 Jul 2019 20:45:00 GMT
ETag: "4efcb3d98ce0fabfd585eb6c4332859"
Content-Length: 493285
Content-Type: application/zip
Server: OTM Preservation Gateway

[493285 bytes of object data]
```

ERROR: Retrieval of an Unavailable Object

Request

```
GET /af48c3d HTTP/1.1
Host: preservation-gateway.institution.edu
Date: Tue, 02 Jul 2019 20:45:00 GMT
Content-Length: 0
```

Response

```
HTTP/1.1 403 Forbidden
Date: Tue, 02 Jul 2019 20:45:00 GMT
ETag: "4efcb3d98ce0fabfd585eb6c4332859"
Content-Type: application/xml
Server: OTM Preservation Gateway

<?xml version="1.0" encoding="UTF-8"?>
<Error>
 <Code>InvalidObjectState</Code>
 <Message>The Object is not available </Message>
 <Resource>/af48c3d</Resource>
</Error>
```

Question

Can HTTP Range requests be used to retrieve individual files (e.g. metadata)

Authentication

Repository Client Authorization

Authentication with the OTM Bridge

Bridge Interactions

GET File (transfer)

Description

Transfer a cached file. This endpoint allows a Bridge to pull content for storage in a preservation service.

Requests/Responses

- Request: `GET /ObjectId/fileName HTTP/1.1`
- Request Headers
    - `If-Match` (checksum)
- Response Code:
    - `200 OK` (if the file is available)
    - `404 NotFound` (if the file is not present)
    - `412 PreconditionFailed` (if no version of the file is present matching the `If-Match` checksum)
- Response Body: file contents
- Response Headers:
    - `Content-Type`
    - `ETag`