

S3 Storage RESTful API

Table of Contents

1. Indledning	1
1.1. Forudsætninger	1
1.2. Sikkerhed	2
1.3. Begreber	2
1.4. API endepunkter	3
1.5. HTTP verb	3
2. Upload et objekt PUT	3
2.1. Filstruktur	4
2.2. Requeststruktur	4
2.3. HTTP-statuskoder	4
2.4. Foretag valid upload	4
3. Hent et objekt GET	5
3.1. Requeststruktur	5
3.2. HTTP-statuskoder	6
3.3. Hent med valid objectId	6
3.4. Hent med IKKE valid identifikator	7
4. Hent headere for et objekt HEAD	8
4.1. Requeststruktur	8
4.2. HTTP-statuskoder	8
4.3. Hent med valid objekt id	8
5. Hent liste med information om objekter GET	9
5.1. Requeststruktur	9
5.2. HTTP-statuskoder	9
5.3. Hent en liste med gyldigt SE-nummer	10

1. Indledning

Dette er en detaljeret beskrivelse af web service API'et for overførsel af større filer. Løsningen er baseret på HTTP/REST og er en delvis implementation af [Amazon S3 REST API](#).

Selvom API'et er generelt ligesom S3 API'et så er dette dokument skrevet til Rente systemet, og skal betragtes som et addendum til *Rente RESTful API* som er der hvor man skal starte med at læse REST documentation.

1.1. Forudsætninger

Det forventes at læseren har kendskab til HTTP, REST og har læst *Rente RESTful API*.

Nedenfor er angivet anbefalet læsning til relevant baggrundsstof.

- Architectural Styles and the Design of Network-based Software Architectures, Roy Thomas Fielding PhD Dissertation University of California, Irvine <http://www.ics.uci.edu/~fielding/pubs/dissertation/top.htm>
- RESTful Web Services Leonard Richardson & Sam Ruby O'Reilly 2007, ISBN 978-0-596-52926-0 <http://shop.oreilly.com/product/9780596529260.do>
- RESTful Web APIs Leonard Richardson & Mike Amundsen O'Reilly 2013, ISBN 978-1-449-35806-9 <http://shop.oreilly.com/product/0636920028468.do>
- Amazon S3 REST API: <http://docs.aws.amazon.com/AmazonS3/latest/API/Welcome.html>

1.2. Sikkerhed

Der bruges 2-vejs-ssl til al kommunikation med servicen.

1.3. Begreber

Begreb	Beskrivelse
Indberetter	Den juridiske enhed (CVR-nummer) som rent faktisk udfører indberetningen. Kan også være den indberetningspligtige. (Kaldes også bureau (i forhold til indberetningspligtig) hvis indberetter er forskellig fra pligtig.)
Indberetningspligtig	Den skattemæssige enhed som skal foretage indberetningen. Kan være indberetter.
Bucket, bucketName	<i>Bucket</i> er et begreb fra S3 REST API. <i>bucketName</i> er identifikator/navn for en bucket. I en S3 service bruges bucketName som hostnavn. I rente bruges der 3 buckets; en for indgående filer (in), og 2 for udgående filer (out , val-out). Nedenfor kan man se en tabel med hvilke konkrete URL'er der skal til de 3 buckets for de 2 miljøer TFE og PRD
Rentekategori	Rentetypen som der benyttes i forbindelse med indberetningen, som kan have én af fem værdier: <ul style="list-style-type: none">• udlån• indlån• pantebreve• prioritetslån• pensiondiverse
pligtigSeNummer	SE-nummer(CVR) på Indberetningspligtig, som er den juridiske entitet der har indberetningspligten.
ObjectName	En unik identifikator som identificerer filen på en bucket. Kan betragtes som filnavn (med absolut sti). ObjectName overholder mønsteret {rentekategori}/pligtige/{pligtigSeNummer}/{objectId} .

Begreb	Beskrivelse
ObjectId	En unik identifikator som identificerer filen på en bucket indenfor en bestemt rentekategori og en bestemt pligtig. ObjectId kan betragtes som filnavn (med relativ sti). Det kan fx. være <code>2017-q1/pakke1987654321.zip</code> . Den maksimale længde af objectId er 255 bytes. I den sammenhæng skal man være opmærksom på at de 3 danske bogstaver fylder mere end een byte pr. bogstav.

1.4. API endepunkter

I alle eksempler er der brugt `in.s3.tse3pindberet.local.skat.dk` som `{host}`. Det egentlige host navn der skal bruges afhænger af miljø og bucket. Nedenstående tabel viser den egentlige url der skal bruges i Rente.

Miljø	Bucket	URL
PRD	in	https://in.s3.tse3pindberet.skat.dk/
	out	https://out.s3.tse3pindberet.skat.dk/
	val-out	https://val-out.s3.tse3pindberet.skat.dk/
TFE	in	https://in.s3.tse3pindberet.tfe.skat.dk/
	out	https://out.s3.tse3pindberet.tfe.skat.dk/
	val-out	https://val-out.s3.tse3pindberet.tfe.skat.dk/

1.5. HTTP verb

S3 Storage lægger sig op ad standard HTTP og REST-konventioner i brugen af HTTP verber/metoder og statuskoder.

Verb	Beskrivelse
PUT	Bruges til at gemme en fil.
GET	Bruges til at hente en fil, eller til at liste filer.
HEAD	Henter headers tilsvarende GET forespørgslen
DELETE	Det er ikke muligt at benytte delete, da sletning af filer ikke er understøttet.

Af sikkerhedsgrunde er adgangen til de forskellige buckets/endepunkter begrænset således at **HEAD** virker på alle endepunkter, men **PUT** er kun muligt på `in.s3.*` endepunkter og **GET** kun tillades på `out.s3.*` og `val-out.s3.*` endepunkterne.

2. Upload et objekt **PUT**

Først fremstiller man en zip-fil(objektet) med alle de indberetninger (et kontoId pr. fil). Filnavnet skal kun være kontoId'et, dvs. ingen filtype på. Dernæst overføres denne zip fil ved brug af nedestående requeststruktur.

2.1. Filstruktur

```
├── 11111111_2017_3187892327.zip
│   ├── 311861252797392
│   ├── 311847686493738
│   ├── 311847684655171
│   ├── 311847681003489
│   ├── 311847686552573
│   └── 311847688414680
```

2.2. Requeststruktur

Når der uploades en fil skal der angives rentekategori (f.eks. udlån), den Indberetningspligtigs SE-nummer (CVR-nummer) og Indberetters identifikator (objectId).

URL	Beskrivelse
<code>{scheme}://{host}/{rentekategori}/pligtige/{pligtigSeNummer}/{objectId}</code>	Uploader et enkelt objekt (fil)

2.3. HTTP-statuskoder

Statuskode	Beskrivelse
200 OK	Filen blev overført korrekt.
400 Bad Request	Den angivne MD5 checksum passer ikke til den overførte fil.
401 Unauthorized	Brugeren har ikke angivet brugeroplysninger.
403 Forbidden	Brugeren har ikke læseadgang.
405 Method Not Allowed	Forkert brug af HTTP method (f.eks PUT, hvor der skulle være GET).
411 Length Required	Content-Length headeren skal angives.

2.4. Foretag valid upload

Når der uploades et objekt (fil) som er valid, returneres et http svar (200 OK), men ikke nogen body. Informationerne, som returneres, ligger i svarets header:

- MD5 Checksum på objektet (filen) som er modtaget
- Stien til filen. Denne kan direkte bruges til at hente objektet (filen) ind i Rente-applikationen. Stien kan ligeledes bruges til senere at hente objektet (filen) ned fra S3.

2.4.1. Eksempel-request 200 OK

2.4.2. curl

```
$ curl
'https://in.s3.tse3pindberet.local.skat.dk/pantebreve/pligtige/11111111/masseindlevering_put.zip' -i -X PUT \
-H 'Content-Type: text/plain; charset=ISO-8859-1' \
-d 'DENNE TEKST ER ERSTATNING FOR INDHOLDET AF EN ZIP-FIL'
```

2.4.3. HTTP

```
PUT /pantebreve/pligtige/11111111/masseindlevering_put.zip HTTP/1.1
Content-Type: text/plain; charset=ISO-8859-1
Host: in.s3.tse3pindberet.local.skat.dk
Content-Length: 53
```

DENNE TEKST ER ERSTATNING FOR INDHOLDET AF EN ZIP-FIL

2.4.4. Response-header

Name	Description
<code>x-amz-version-id</code>	Versions-id for den uploadede fil. Forskellige versioner af en fil kan have samme indhold.
<code>ETag</code>	MD5 checksum af objektets indhold
<code>Location</code>	URL for den oprettede resurse

2.4.5. Eksempel-response

```
HTTP/1.1 200 OK
X-Content-Type-Options: nosniff
x-amz-version-id: f6ecd96b-0bfb-31e7-9d01-4f8ed4bd0551
ETag: Ojh14kscAfb6+ITIOpoCzg==
Content-MD5: Ojh14kscAfb6+ITIOpoCzg==
Location: /pantebreve/pligtige/11111111/masseindlevering_put.zip
```

3. Hent et objekt GET

Når der hentes en fil skal der angives rentekategori, den indberetningspligtiges SE-nummer (CVR-nummer) og indberetters selvvalgte objectId.

3.1. Requeststruktur

URL	Beskrivelse
<code>{scheme}://{host}/{rentekategori}/pligtige/{pligtigSeNummer}/{objectId}</code>	Henter et enkelt objekt (fil) med objectId (filnavn).

3.2. HTTP-statuskoder

Statuskode	Beskrivelse
200 OK	Body er indholdet af den gemte fil.
401 Unauthorized	Brugeren har ikke angivet bruger oplysninger.
403 Forbidden	Brugeren har ikke læseadgang.
404 Not Found	Filen kan ikke findes.

3.3. Hent med valid objectId

Et filnavn kan bruges som objectId, eksempel: "2017_01-03_pantebreve.zip"

3.3.1. Eksempel-request 200 OK

3.3.2. curl

```
$ curl
'https://out.s3.tse3pindberet.local.skat.dk/pantebreve/pligtige/11111111/masseindlevering_get_with_objectId.zip' -i -X GET
```

3.3.3. HTTP

```
GET /pantebreve/pligtige/11111111/masseindlevering_get_with_objectId.zip HTTP/1.1
Host: out.s3.tse3pindberet.local.skat.dk
```

3.3.4. Respons header

Name	Description
<code>x-amz-version-id</code>	Versions-id for den uploadede fil. Forskellige versioner af en fil kan have samme indhold.
<code>ETag</code>	MD5 checksum af objektets indhold

3.3.5. Eksempel-respons

```
HTTP/1.1 200 OK
X-Content-Type-Options: nosniff
x-amz-version-id: f6ecd96b-0bfb-31e7-9d01-4f8ed4bd0551
ETag: Ojh14kscAfb6+ITIOpoCzg==
Content-MD5: Ojh14kscAfb6+ITIOpoCzg==
Content-Type: application/zip
```

DENNE TEKST ER ERSTATNING FOR INDHOLDET AF EN ZIP-FIL

3.4. Hent med IKKE valid identifikator

Når der hentes en fil skal der angives rentekategori, den Indberetningspligtigs SE-nummer(CVR) og Indberetters indentifikator. Hvis objektId ikke er korrekt, vil objektet ikke findes og en HTTP Status 404 vil blive returneret.

3.4.1. Eksempel-request 404 Not Found

3.4.2. curl

```
$ curl
'https://out.s3.tse3pindberet.local.skat.dk/pantebreve/pligtige/11111111/masseindlevering_get_404.zip' -i -X GET
```

3.4.3. HTTP

```
GET /pantebreve/pligtige/11111111/masseindlevering_get_404.zip HTTP/1.1
Host: out.s3.tse3pindberet.local.skat.dk
```

3.4.4. Respons felter

Path	Type	Description
Error	Number	Fejl-objekt
Error/Code	Number	Fejlkode
Error/Message	String	Beskrivende fejlbesked
Error/Resource	String	Stien, der er forespurgt efter
Error/RequestId	String	Transaktion id'et på requested der bliver var årsagen til fejlen

3.4.5. Eksempel-respons

```
HTTP/1.1 404 Not Found
X-Content-Type-Options: nosniff
Content-Type: application/xml;charset=utf-8
Transfer-Encoding: chunked

<?xml version="1.0" encoding="UTF-8"?>
<Error>
  <Code>NoSuchKey</Code>
  <Message>The specified key does not exist.</Message>
  <Resource>/pantebreve/pligtige/11111111/masseindlevering_get_404.zip</Resource>
  <RequestId>a68110ef-2861-4172-810c-51462d4d2aa1</RequestId>
</Error>
```

4. Hent headere for et objekt **HEAD**

HEAD metoden bruges til at returnere headerne for et objekt uden at body returneres. Derved spares der tid og båndbredde.

4.1. Requeststruktur

URL	Beskrivelse
<code>{scheme}://{host}/{rentekategori}/pligtige/{pligtigSeNummer}/{objectId}</code>	Henter et enkelt objekt (fil) med objectId (filnavn) og uden body.

4.2. HTTP-statuskoder

Statuskode	Beskrivelse
200 OK	Headerne overførtes korrekt.
401 Unauthorized	Brugeren har ikke angivet brugeroplysninger.
403 Forbiden	Brugeren har ikke læseadgang.
404 Not Found	Filen kan ikke findes.

4.3. Hent med valid objekt id

Når der bruges objectId til at hente en fil, kan objectId'et betragtes som filnavnet. for eksempel "2017_01-03_pantebreve.zip"

4.3.1. Eksempel-request **200 OK**

4.3.2. curl


```
$ curl
'https://in.s3.tse3pindberet.local.skat.dk/pantebreve/pligtige/11111111/masseindlevering_head.zip' -i -X HEAD
```

4.3.3. HTTP

```
HEAD /pantebreve/pligtige/11111111/masseindlevering_head.zip HTTP/1.1
Host: in.s3.tse3pindberet.local.skat.dk
```

4.3.4. Respons header

Name	Description
<code>x-amz-version-id</code>	Versions-id for den uploadede fil. Forskellige versioner af en fil kan have samme indhold.
<code>ETag</code>	MD5 checksum af objektets indhold

4.3.5. Eksempel-respons

```
HTTP/1.1 200 OK
X-Content-Type-Options: nosniff
x-amz-version-id: f6ecd96b-0bfb-31e7-9d01-4f8ed4bd0551
ETag: Ojh14kscAfb6+ITIOpoCzg==
Content-MD5: Ojh14kscAfb6+ITIOpoCzg==
Content-Type: application/zip;charset=ISO-8859-1
```

5. Hent liste med information om objekter

GET

Når der skal findes en liste af objekter (filer), bruges SE-nummeret (CVR-nummer) på den indberetningspligtige som parameter. Listen viser de objekter der er gemt under den specificerede rentekategori for den givne indberetningspligtige.

5.1. Requeststruktur

URL	Beskrivelse
<code>{scheme}://{host}/{rentekategori}/pligtige/{pligtigSeNummer}</code>	Henter en liste med informationer om objekter (filer), som returneres i body.

5.2. HTTP-statuskoder

Statuskode	Beskrivelse
200 OK	Den efterpurgte liste returneres i body.
401 Unauthorized	Brugeren har ikke angivet credentials
403 Forbiden	Brugeren har ikke læseadgang

5.3. Hent en liste med gyldigt SE-nummer

5.3.1. Eksempel-request 200 OK

5.3.2. curl

```
$ curl 'https://out.s3.tse3pindberet.local.skat.dk/pantebreve/pligtige/11111111/' -i
-X GET
```

5.3.3. HTTP

```
GET /pantebreve/pligtige/11111111/ HTTP/1.1
Host: out.s3.tse3pindberet.local.skat.dk
```

5.3.4. Respons felter

Path	Type	Description
ListBucketResult	String	Amazon S3 Bucket list
ListBucketResult/Name	String	Navnet på den bucket som er brugt.
ListBucketResult/Prefix	String	Benyttes ikke, men er 'Keys' som begynder med det prefix som er benyttet. (Er med for at være Amazon S3 Kompatibel)
ListBucketResult/Marker	String	Benyttes ikke, men er en indikator på hvor listen begynder. (Er med for at være Amazon S3 Kompatibel)
ListBucketResult/MaxKeys	Number	Der er sat en max på 1000
ListBucketResult/IsTruncated	Boolean	Viser om der er flere filer, men som bare ikke er vist i den nuværende søgning.
ListBucketResult/Contents	Array	Objekt som indeholde informationer om den specifikke fil

Path	Type	Description
ListBucketResult/Contents/Key	String	Sti'en til filen, som også bliver brugt som locations parameter i retursvaret ved upload.
ListBucketResult/Contents/LastModified	String	Tidsstempel for hvornår filen blev lagt op.
ListBucketResult/Contents/ETag	String	MD5 Checksum af filen
ListBucketResult/Contents/Size	Number	Størrelsen på filen i bytes.
ListBucketResult/Contents/StorageClass	String	Denne er sat til 'STANDARD', for at være Amazon S3 Kombatibel

5.3.5. Eksempel-respons

```

HTTP/1.1 200 OK
X-Content-Type-Options: nosniff
Content-Type: application/xml;charset=utf-8
Transfer-Encoding: chunked

<?xml version="1.0" encoding="UTF-8"?>
<ListBucketResult xmlns="http://s3.amazonaws.com/doc/2006-03-01/">
  <Name>out</Name>
  <Prefix xmlns="http://s3.amazonaws.com/doc/2006-03-01/">
  <Marker xmlns="http://s3.amazonaws.com/doc/2006-03-01/">
  <MaxKeys xmlns="http://s3.amazonaws.com/doc/2006-03-01/">
  <IsTruncated/>
  <Contents xmlns="http://s3.amazonaws.com/doc/2006-03-01/">
    <Key>/pantebreve/pligtige/11111111/masseindlevering_list.zip</Key>
    <LastModified>2017-01-02T10:10:10.123Z</LastModified>
    <ETag xmlns="http://s3.amazonaws.com/doc/2006-03-01/">0jh14kscAfb6+ITIOpoCzg==</ETag>
    <Size>53</Size>
    <StorageClass>STANDARD</StorageClass>
  </Contents>
</ListBucketResult>

```