

rgw - Bug #22982

Objects only serving first 512K

02/12/2018 07:48 AM - Robin Johnson

Status:	Resolved	% Done:	0%
Priority:	High	Spent time:	0.00 hour
Assignee:			
Category:			
Target version:			
Source:	Community (dev)	Reviewed:	
Tags:	rgw	Affected Versions:	v12.0.0, v12.1.0, v12.2.0, v12.2.1, v12.2.2, v12.2.3
Backport:	luminous	ceph-qa-suite:	rgw
Regression:	Yes	Pull request ID:	
Severity:	2 - major	Crash signature:	

Description

This has been seen in previous releases, but it's back again in Luminous! Too late for 12.2.3 tag, but I'll apply the fix to our internal build when it's available.

```
$ wget ...
--2018-02-11 23:44:57-- https://objects-us-west-1.dream.io/BUCKET-CENSORED/FILENAME-CENSORED?Expires=1518424203&Signature=CENSORED&AWSAccessKeyId=CENSORED
Resolving objects-us-west-1.dream.io... 2607:f298:4:143:acce:55:2:1, 64.90.32.8
Connecting to objects-us-west-1.dream.io|2607:f298:4:143:acce:55:2:1|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 11127936 (11M) [audio/x-wav]
Saving to: '/dev/null'

/dev/null                                4%[====>
                                           ] 512.00K  1.90MB/s   in 0.3s

2018-02-11 23:44:58 (1.90 MB/s) - Connection closed at byte 524288. Retrying.
```

JSON of object stat

```
$ radosgw-admin --bucket BUCKET-CENSORED --object 'FILENAME-CENSORED' object stat
{
  "name": "FILENAME-CENSORED",
  "size": 11127936,
  "policy": {
    "acl": {
      "acl_user_map": [
        {
          "user": "USER-CENSORED",
          "acl": 15
        }
      ],
      "acl_group_map": [],
      "grant_map": [
        {
          "id": "USER-CENSORED",
          "grant": {
            "type": {
              "type": 0
            }
          }
        }
      ]
    }
  }
}
```

```

        "id": "USER-CENSORED",
        "email": "",
        "permission": {
            "flags": 15
        },
        "name": "USER-CENSORED",
        "group": 0,
        "url_spec": ""
    }
}
]
},
"owner": {
    "id": "USER-CENSORED",
    "display_name": "USER-CENSORED"
}
},
"etag": "722e7c40151930b38cdf285659a92d94",
"tag": "default.52038511.39424258",
"manifest": {
    "objs": [
        0,
        {
            "loc": {
                "bucket": {
                    "name": "BUCKET-CENSORED",
                    "marker": "default.51695731.142363",
                    "bucket_id": "default.51695731.142363",
                    "tenant": "",
                    "explicit_placement": {
                        "data_pool": ".rgw.data.1",
                        "data_extra_pool": "",
                        "index_pool": ".rgw.data.1"
                    }
                },
                "key": {
                    "name": "FILENAME-CENSORED",
                    "instance": "",
                    "ns": ""
                }
            },
            "loc_ofs": 0,
            "size": 524288
        },
        524288,
        {
            "loc": {
                "bucket": {
                    "name": "BUCKET-CENSORED",
                    "marker": "default.51695731.142363",
                    "bucket_id": "default.51695731.142363",
                    "tenant": "",
                    "explicit_placement": {
                        "data_pool": ".rgw.data.1",
                        "data_extra_pool": "",
                        "index_pool": ".rgw.data.1"
                    }
                },
                "key": {
                    "name": "__JbcJZZ4eXDtuKeFN8SakI46t0Erh4T5_1",
                    "instance": "",
                    "ns": "shadow"
                }
            },
            "loc_ofs": 0,
            "size": 4194304
        },
    ],
}

```

```

4718592,
{
  "loc": {
    "bucket": {
      "name": "BUCKET-CENSORED",
      "marker": "default.51695731.142363",
      "bucket_id": "default.51695731.142363",
      "tenant": "",
      "explicit_placement": {
        "data_pool": ".rgw.data.1",
        "data_extra_pool": "",
        "index_pool": ".rgw.data.1"
      }
    },
    "key": {
      "name": "__JbcJZZ4eXDtuKeFN8SakI46t0Erh4T5_2",
      "instance": "",
      "ns": "shadow"
    }
  },
  "loc_ofs": 0,
  "size": 4194304
},
8912896,
{
  "loc": {
    "bucket": {
      "name": "BUCKET-CENSORED",
      "marker": "default.51695731.142363",
      "bucket_id": "default.51695731.142363",
      "tenant": "",
      "explicit_placement": {
        "data_pool": ".rgw.data.1",
        "data_extra_pool": "",
        "index_pool": ".rgw.data.1"
      }
    },
    "key": {
      "name": "__JbcJZZ4eXDtuKeFN8SakI46t0Erh4T5_3",
      "instance": "",
      "ns": "shadow"
    }
  },
  "loc_ofs": 0,
  "size": 2215040
}
],
"obj_size": 11127936,
"explicit_objs": "true",
"head_size": 524288,
"max_head_size": 524288,
"prefix": "",
"rules": [],
"tail_instance": "",
"tail_placement": {
  "bucket": {
    "name": "",
    "marker": "",
    "bucket_id": "",
    "tenant": "",
    "explicit_placement": {
      "data_pool": "",
      "data_extra_pool": "",
      "index_pool": ""
    }
  }
},
"placement_rule": ""

```

```
}
},
"attrs": {
  "user.rgw.content_type": "audio/x-wav"
}
}
```

Related issues:

Copied to rgw - Backport #23102: luminous: Objects only serving first 512K

Resolved

History

#1 - 02/12/2018 04:09 PM - Robin Johnson

- Backport set to luminous

#2 - 02/12/2018 04:20 PM - Matt Benjamin

@robin, can you point out the distinguishing characteristics of this? how is our CI testing missing this?

Matt

#3 - 02/12/2018 05:06 PM - Yehuda Sadeh

it's an old object with explicit placement, so maybe pre-firefly (?). The tail objects are starting with underscore, so maybe there's some regression there.

#4 - 02/12/2018 09:01 PM - Robin Johnson

It doesn't start with an underscore in this case, but the data is from at least Dumpling, maybe older (Dumpling->Firefly upgrade happened 2015/Jan)

```
# rados -p .rgw.data.1 stat default.51695731.142363__shadow__JbcJZZ4eXDtuKeFN8SakI46t0Erh4T5_1
.rgw.data.1/default.51695731.142363__shadow__JbcJZZ4eXDtuKeFN8SakI46t0Erh4T5_1 mtime 2014-03-31 03:52:59.00000
0, size 4194304
```

#5 - 02/12/2018 09:08 PM - Robin Johnson

Ok, Yehuda's underscore could still be correct. This is what the RGW is looking for:

- default.51695731.142363__shadow__JbcJZZ4eXDtuKeFN8SakI46t0Erh4T5_1 - looking for
- default.51695731.142363__shadow__JbcJZZ4eXDtuKeFN8SakI46t0Erh4T5_1 - actual

```
2018-02-12 21:01:40.932636 7f3771123700 1 -- [XXXX::7590]:0/382522710 --> [XXXX::8880]:6809/3072030 -- osd_op
(unknown.0.0:69623801 5.cb53 5:cad336d1:::default.51695731.142363__shadow__JbcJZZ4eXDtuKeFN8SakI46t0Erh4T5_1:
head [read 0~4194304] snapc 0=[] ondisk+read+known_if_redirected e1289786) v8 -- 0x56283b8a59c0 con 0
2018-02-12 21:01:40.933380 7f37be3c3700 1 -- [XXXX::7590]:0/382522710 <== osd.186 [XXXX::8880]:6809/3072030 5
444 ==== osd_op_reply(69623801 default.51695731.142363__shadow__JbcJZZ4eXDtuKeFN8SakI46t0Erh4T5_1 [read 0~419
4304] v0'0 uv0 ondisk = -2 ((2) No such file or directory)) v8 ==== 211+0+0 (880413640 0 0) 0x56283c198580 con
0x56283eb10800
```

#6 - 02/12/2018 09:52 PM - Yehuda Sadeh

Can you provide the raw manifest somehow? I want to know if the issue is the way we stored the original objects (that might have been buggy) or the new decoding. To do that you will need to read the manifest off the manifest xattr in the head object.

#7 - 02/13/2018 04:24 AM - Robin Johnson

- *File Dupont.wav-user.rgw.manifest added*

#9 - 02/13/2018 04:25 AM - Robin Johnson

Private attachment made w/ manifest & hexdump of manifest.

#10 - 02/13/2018 04:55 PM - Yehuda Sadeh

Manifest has appropriate raw rados objects there, so we do an unneeded conversion when reading it.

#11 - 02/13/2018 06:22 PM - Yehuda Sadeh

a possible fix:

<https://github.com/ceph/ceph/pull/20425>

#12 - 02/14/2018 06:44 AM - Robin Johnson

This change has been deployed at Dreamhost, seems to work so far, watching for other breakage

#13 - 02/15/2018 06:43 PM - Casey Bodley

- *Status changed from New to 12*

#14 - 02/22/2018 07:08 PM - Casey Bodley

- *Status changed from 12 to Pending Backport*

#15 - 02/23/2018 01:50 PM - Nathan Cutler

- *Copied to Backport #23102: luminous: Objects only serving first 512K added*

#17 - 08/27/2019 08:52 AM - Nathan Cutler

- *Status changed from Pending Backport to Resolved*

While running with --resolve-parent, the script "backport-create-issue" noticed that all backports of this issue are in status "Resolved".

Files

Dupont.wav-user.rgw.manifest	1.06 KB	02/13/2018	Robin Johnson
------------------------------	---------	------------	---------------