

## Ceph - Feature #619

### objecter: optionally read from replicas

11/30/2010 12:23 PM - Sage Weil

<b>Status:</b> Resolved	<b>% Done:</b> 0%
<b>Priority:</b> Normal	<b>Spent time:</b> 0.00 hour
<b>Assignee:</b> Colin McCabe	
<b>Category:</b>	
<b>Target version:</b> v0.25	
<b>Source:</b>	<b>Reviewed:</b>
<b>Tags:</b>	<b>Affected Versions:</b>
<b>Backport:</b>	<b>Pull request ID:</b>
<b>Description</b> Add a read flag to allow reads to come from a random replica. If a replica replies with EAGAIN, retry the request, but always on the primary. (This may happen if the replica is missing the object due to a recovery in progress.)	
<b>Related issues:</b>	
Related to Ceph - Feature #618: osd: allow reads from replicas	<b>Resolved</b> 11/30/2010
Blocked by Ceph - Cleanup #650: objecter: refactor request tracking to be per...	<b>Resolved</b> 12/14/2010

### History

#### #1 - 01/15/2011 04:04 PM - Sage Weil

- Assignee set to Colin McCabe

see objecter\_balance\_reads branch for untested client-side implementation

#### #2 - 01/24/2011 02:41 PM - Colin McCabe

- Status changed from New to Resolved

I ran test\_rw.sh with the following patch:

```
diff --git a/src/osdc/Objecter.h b/src/osdc/Objecter.h
index 37a23fb..84afa9d 100644
--- a/src/osdc/Objecter.h
+++ b/src/osdc/Objecter.h
@ -309,6 +309,7 @ public:
tid(0), attempts(0),
paused(false), objver(ov) {
ops.swap(op);
flags |= CEPH_OSD_FLAG_BALANCE_READS;
}
};
```

It succeeded when reading from the replicas.

One thing to note. If an object is unfound on the OSD we try to fetch it from, we block until it becomes found on that OSD. We do not ask other OSDs if they already have a copy of that same object. This is probably acceptable because eventually, we will eliminate all unfound items on the different replicas through PG recovery.

### #3 - 01/24/2011 02:42 PM - Colin McCabe

Sigh. Reposting patch for clarity:

```
--- a/src/osdc/Objecter.h
+++ b/src/osdc/Objecter.h
@@ -309,6 +309,7 @@ public:
     tid(0), attempts(0),
     paused(false), objver(ov) {
     ops.swap(op);
+    flags |= CEPH_OSD_FLAG_BALANCE_READS;
     }
};
```