



oscon.com

Consistency or Bust

Breaking a Riak Cluster

Jeff Kirkell

email: jeff.kirkell@gmail.com

twitter: [@jeffkirkell](https://twitter.com/jeffkirkell)

NoSQL Primer



Riak Sales Pitch

- Dynamo inspired
- AP of CAP
- key / value store
- scales consistently
- fault tolerance
- true high availability
- built-in replication
- lot's of clients
- simple to implement
- low \$\$\$ to use
- multiple license options

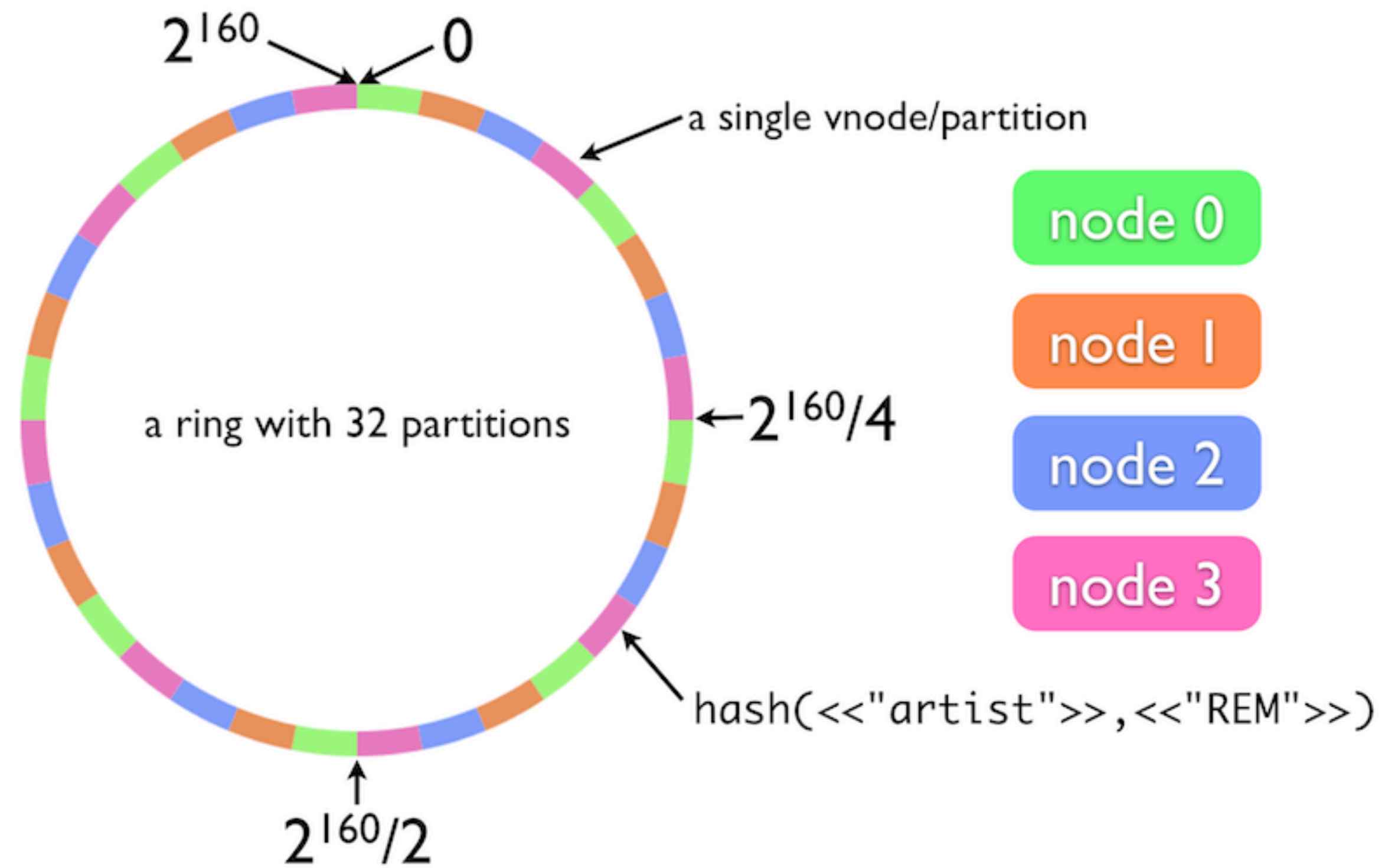
Riak Deep Dive



Consistency or Bust: Breaking a Riak Cluster

Cluster Defined

- cluster
- # nodes
- $64 / \# \text{ nodes}$
- vnodes



Important Bits

- gossiping
- hintful handoff
- n # copies of data
- r # nodes that must respond for read
- w # nodes that must respond for write
- dw # replicas written on write for success
- rw # replicas deleted for success



Tools of the Trade

- terminal
- curl
- rebar
- rekon

The Workshop

setup instructions for cluster

https://github.com/jkirkell/riak_workshop

Start Your Engines

```
$ ./dev/dev1/bin/riak start
```

```
$ ./dev/dev2/bin/riak start
```

```
$ ./dev/dev3/bin/riak start
```

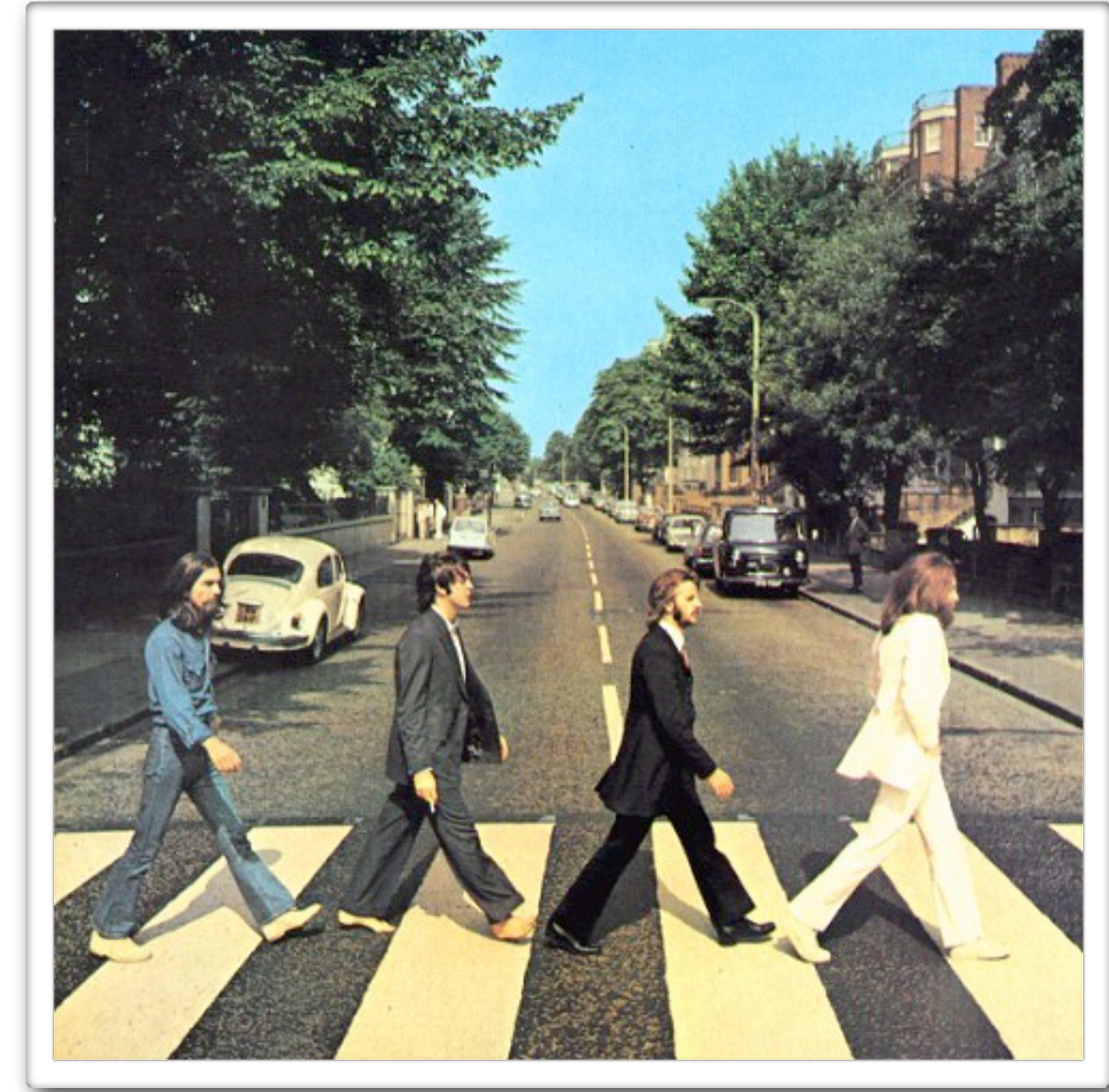
```
$ ./dev/dev4/bin/riak start
```

```
$ ./dev/dev5/bin/riak start
```

```
$ ps aux | grep erl
```

Come Together

```
$ ./dev/dev2/bin/riak-admin join dev1 @127.0.0.1  
$ ./dev/dev3/bin/riak-admin join dev1 @127.0.0.1  
$ ./dev/dev4/bin/riak-admin join dev1 @127.0.0.1  
$ ./dev/dev5/bin/riak-admin join dev1 @127.0.0.1  
$ ./dev/dev3/bin/riak-admin status
```



Default Settings



$$\text{quorum} = \text{floor}(5 / 2 + 1) = 3$$

Add Data

```
$ curl -v -d 'inserted data' -H "Content-Type: text/plain"  
http://127.0.0.1:8091/riak/default
```

```
* About to connect() to 127.0.0.1 port 8091 (#0)  
*   Trying 127.0.0.1... connected  
* Connected to 127.0.0.1 (127.0.0.1) port 8091 (#0)  
> POST /riak/default HTTP/1.1  
> User-Agent: curl/7.21.4 (universal-apple-darwin11.0) libcurl/7.21.4 OpenSSL/0.  
9.8r zlib/1.2.5  
> Host: 127.0.0.1:8091  
> Accept: */*  
> Content-Type: text/plain  
> Content-Length: 13  
>  
< HTTP/1.1 201 Created  
< Vary: Accept-Encoding  
< Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)  
< Location: /riak/default/7j0uqpT0JxkDrAk0sroQsldTZ6e  
< Date: Thu, 28 Jul 2011 07:23:50 GMT  
< Content-Type: application/json  
< Content-Length: 0  
<  
* Connection #0 to host 127.0.0.1 left intact  
* Closing connection #0
```

Confirm Data Exists

```
$ curl -i http://127.0.0.1:8091/riak/default?keys=true
```

```
HTTP/1.1 200 OK
```

```
Vary: Accept-Encoding
```

```
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
```

```
Link: </riak/default/7j0uqpT0JxkDrAk0sroQsldTZ6e>; riaktag="contained"
```

```
Date: Thu, 28 Jul 2011 07:25:28 GMT
```

```
Content-Type: application/json
```

```
Content-Length: 464
```

```
{"props":{"name":"default","allow_mult":false,"basic_quorum":true,"big_vclock":50,"chash_keyfun":{"mod":"riak_core_util","fun":"chash_std_keyfun"},"dw":"quorum","last_write_wins":false,"linkfun":{"mod":"riak_kv_wm_link_walker","fun":"mapreduce_linkfun"},"n_val":3,"notfound_ok":false,"old_vclock":86400,"postcommit":[],"pr":0,"precommit":[],"pw":0,"r":"quorum","rw":"quorum","small_vclock":10,"w":"quorum","young_vclock":20},"keys":["7j0uqpT0JxkDrAk0sroQsldTZ6e"]}
```


Find Your Data

```
$ find ./dev/dev*/data/bitcask/*/*.data
```

```
./dev/dev2/data/bitcask/936274486415109681974235595958868809467081785344/1311837703.bi  
tcask.data  
./dev/dev3/data/bitcask/1050454301831586472458898473514828420377701515264/1311793197.b  
itcask.data  
./dev/dev3/data/bitcask/867766597165223607683437869425293042920709947392/1311837830.bi  
tcask.data
```

Stop Primary Node

```
$ dev[n]/bin/riak stop
```

```
$ curl -i http://127.0.0.1:8091/riak/default/[key]
```

```
HTTP/1.1 200 OK
X-Riak-Vclock: a85hYGBgzGDKBVISjP0sbTKYEhnzWBmCwpYf58sCAA==
Vary: Accept-Encoding
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
Link: </riak/default>; rel="up"
Last-Modified: Fri, 29 Jul 2011 16:05:38 GMT
ETag: "6Reox6nQyagmRlkvsgZ8HY"
Date: Fri, 29 Jul 2011 16:07:58 GMT
Content-Type: text/plain
Content-Length: 13
```

```
inserted data
```

Stop Another Primary Node

```
$ dev[n]/bin/riak stop
```

```
$ curl -i http://127.0.0.1:8091/riak/default/[key]
```

```
HTTP/1.1 503 Service Unavailable
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
Date: Fri, 29 Jul 2011 16:10:29 GMT
Content-Type: text/plain
Content-Length: 25

R-value unsatisfied: 1/2
```


Query Again

```
$ curl -i http://127.0.0.1:8091/riak/default/[key]
```

```
.
HTTP/1.1 200 OK
X-Riak-Vclock: a85hYGBgzGDKBVISjP0sbTKYEhnzWBmCwpYf58sCAA==
Vary: Accept-Encoding
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
Link: </riak/default>; rel="up"
Last-Modified: Fri, 29 Jul 2011 16:05:38 GMT
ETag: "6Reox6nQyagmRlkvsGZ8HY"
Date: Fri, 29 Jul 2011 16:10:43 GMT
Content-Type: text/plain
Content-Length: 13
```

```
inserted data: <[REDACTED]>
```

Check What's Happening

```
$ find ./dev/dev*/data/bitcask/*/*.data
```

```
./dev/dev1/data/bitcask/114179815416476790484662877555959610910619729920/1311955538.bitcask.data  
./dev/dev2/data/bitcask/68507889249886074290797726533575766546371837952/1311955678.bitcask.data  
./dev/dev2/data/bitcask/91343852333181432387730302044767688728495783936/1311955829.bitcask.data  
./dev/dev3/data/bitcask/68507889249886074290797726533575766546371837952/1311955829.bitcask.data  
./dev/dev4/data/bitcask/68507889249886074290797726533575766546371837952/1311955538.bitcask.data  
./dev/dev5/data/bitcask/91343852333181432387730302044767688728495783936/1311955538.bitcask.data
```

```
$ ./dev/dev[n]/bin/riak-admin transfers
```

```
Nodes ['dev4@127.0.0.1', 'dev5@127.0.0.1'] are currently down.  
'dev3@127.0.0.1' waiting to handoff 1 partitions  
'dev2@127.0.0.1' waiting to handoff 2 partitions  
'dev1@127.0.0.1' waiting to handoff 1 partitions
```

Update Data

```
$ curl -v -d 'updated data' -H "Content-Type: text/plain"  
http://127.0.0.1:8091/riak/defaults/[key]
```

```
* About to connect() to 127.0.0.1 port 8091 (#0)  
*   Trying 127.0.0.1... connected  
* Connected to 127.0.0.1 (127.0.0.1) port 8091 (#0)  
> POST /riak/defaults/9IKT4Dg2lnjPzjfb5k5agBQW8Tq HTTP/1.1  
> User-Agent: curl/7.21.4 (universal-apple-darwin11.0) libcurl/7.21.4 OpenSSL/0.9.8r zlib/1.2.5  
> Host: 127.0.0.1:8091  
> Accept: */*  
> Content-Type: text/plain  
> Content-Length: 12  
>  
< HTTP/1.1 204 No Content  
< Vary: Accept-Encoding  
< Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)  
< Date: Fri, 29 Jul 2011 16:39:48 GMT  
< Content-Type: text/plain  
< Content-Length: 0  
<  
* Connection #0 to host 127.0.0.1 left intact  
* Closing connection #0
```


Start Node

```
$ dev[n]/bin/riak start
```

```
$ ./dev/dev[n]/bin/riak-admin transfers
```

```
Nodes ['dev4@127.0.0.1'] are currently down.  
'dev3@127.0.0.1' waiting to handoff 1 partitions  
'dev2@127.0.0.1' waiting to handoff 3 partitions  
'dev1@127.0.0.1' waiting to handoff 2 partitions
```

Review Logs

```
$ less ./dev2/log/*.1
```

```
2011-07-29 12:46:04 [info] <0.19767.0>@riak_core_handoff_sender:start_fold:39 Starting handoff of
partition riak_kv_vnode 91343852333181432387730302044767688728495783936 from 'dev2@127.0.0.1' to
'dev5@127.0.0.1'
2011-07-29 12:46:04 [info] <0.19767.0>@riak_core_handoff_sender:start_fold:77 Handoff of partitio
n_riak_kv_vnode 91343852333181432387730302044767688728495783936 from 'dev2@127.0.0.1' to 'dev5@12
```

```
$ less ./dev5/log/*.1
```

```
2011-07-29 12:46:04 [info] <0.19767.0>@riak_core_handoff_sender:start_fold:39 Starting handoff of
partition riak_kv_vnode 91343852333181432387730302044767688728495783936 from 'dev2@127.0.0.1' to
'dev5@127.0.0.1'
2011-07-29 12:46:04 [info] <0.19767.0>@riak_core_handoff_sender:start_fold:77 Handoff of partitio
n_riak_kv_vnode 91343852333181432387730302044767688728495783936 from 'dev2@127.0.0.1' to 'dev5@12
```

Start Other Node

```
$ dev[n]/bin/riak start
```

```
$ curl -i http://127.0.0.1:8091/riak/default/[key]
```

```
HTTP/1.1 200 OK
X-Riak-Vclock: a85hYGBgz2DKBVISjP0sbTKYEhznzWBmCwpYf54MKM4QGcECFXVLhwmzNSQxtalmyEizsm5v3QoVDUNSzyC
g4YkqwMPFxeUGFY1DUs5tEb0aWyAIA
Vary: Accept-Encoding
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
Link: </riak/default>; rel="up"
Last-Modified: Fri, 29 Jul 2011 17:09:48 GMT
ETag: "3QQ2gn6ZhSeYvLXjCLLFqj"
Date: Fri, 29 Jul 2011 17:11:53 GMT
Content-Type: text/plain
Content-Length: 12
```

```
updated data: [REDACTED]
```


Summary

- consistently fail 1st read with 2 node down
- 2 node outage max.
- hintful-handoff takes time
- default settings good for general use

Clean Your Cluster

```
$ dev1/bin/riak stop
$ dev2/bin/riak stop
$ dev3/bin/riak stop
$ dev4/bin/riak stop
$ dev5/bin/riak stop
$ cd .. && rm -rf ./dev/*
$ make devrel
$ dev1/bin/riak start
$ dev2/bin/riak start
$ dev3/bin/riak start
$ dev4/bin/riak start
$ dev5/bin/riak start
```

Low Value Data



$$n = 2, r = 1, w = 1$$

Set Bucket Properties

```
$ curl -v -X PUT -H "Content-Type: application/json" -d  
'{"props":{"n_val":2,"r":1,"w":1}}' http://127.0.0.1:8091/  
riak/low_value
```

```
* About to connect() to 127.0.0.1 port 8091 (#0)  
*   Trying 127.0.0.1... connected  
* Connected to 127.0.0.1 (127.0.0.1) port 8091 (#0)  
> PUT /riak/low_value HTTP/1.1  
> User-Agent: curl/7.21.4 (universal-apple-darwin11.0) libcurl/7.21.4 OpenSSL/0.9.8r zlib/1.2.5  
> Host: 127.0.0.1:8091  
> Accept: */*  
> Content-Type: application/json  
> Content-Length: 33  
>  
< HTTP/1.1 204 No Content  
< Vary: Accept-Encoding  
< Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)  
< Date: Sun, 31 Jul 2011 14:58:30 GMT  
< Content-Type: application/json  
< Content-Length: 0  
<  
* Connection #0 to host 127.0.0.1 left intact  
* Closing connection #0
```

Add Data

```
$ curl -v -d 'inserted data' -H "Content-Type: text/plain"  
http://127.0.0.1:8091/riak/low_value
```

```
* About to connect() to 127.0.0.1 port 8091 (#0)  
*   Trying 127.0.0.1... connected  
* Connected to 127.0.0.1 (127.0.0.1) port 8091 (#0)  
> POST /riak/low_value HTTP/1.1  
> User-Agent: curl/7.21.4 (universal-apple-darwin11.0) libcurl/7.21.4 OpenSSL/0.9.8r zlib/1.2.5  
> Host: 127.0.0.1:8091  
> Accept: */*  
> Content-Type: text/plain  
> Content-Length: 13  
>  
< HTTP/1.1 201 Created  
< Vary: Accept-Encoding  
< Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)  
< Location: /riak/low_value/AtgiNS6YGoQ0qlvv2XMZLWhB0J5  
< Date: Sun, 31 Jul 2011 14:58:41 GMT  
< Content-Type: application/json  
< Content-Length: 0  
<  
* Connection #0 to host 127.0.0.1 left intact  
* Closing connection #0
```

Confirm Data Exists

```
$ curl -i http://127.0.0.1:8091/riak/low_value?keys=true
```

```
HTTP/1.1 200 OK
```

```
Vary: Accept-Encoding
```

```
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
```

```
Link: </riak/low_value/AtgiNS6YGoQ0qlvv2XMZLWhB0J5>; riaktag="contained"
```

```
Date: Sun, 31 Jul 2011 14:58:58 GMT
```

```
Content-Type: application/json
```

```
Content-Length: 452
```

```
{"props":{"allow_mult":false,"basic_quorum":true,"big_vclock":50,"chash_keyfun":{"mod":"riak_core_util","fun":"chash_std_keyfun"},"dw":"quorum","last_write_wins":false,"linkfun":{"mod":"riak_kv_wm_link_walker","fun":"mapreduce_linkfun"},"n_val":2,"name":"low_value","notfound_ok":false,"old_vclock":86400,"postcommit":[],"pr":0,"precommit":[],"pw":0,"r":1,"rw":"quorum","small_vclock":10,"w":1,"young_vclock":20},"keys":["AtgiNS6YGoQ0qlvv2XMZLWhB0J5"]}
```


Find Your Data

```
$ find ./dev/dev*/data/bitcask/*/*.data
```

```
./dev/dev4/data/bitcask/867766597165223607683437869425293042920709947392/1312124321.bitcask.data  
./dev/dev5/data/bitcask/890602560248518965780370444936484965102833893376/1312124321.bitcask.data
```

Stop Primary Node

```
$ dev[n]/bin/riak stop
```

```
$ curl -i http://127.0.0.1:8091/riak/default/[key]
```

```
HTTP/1.1 200 OK
X-Riak-Vclock: a85hYGBgzGDKBVIslCmcDz0YEhnzWBkWvlx5nC8LAA==
Vary: Accept-Encoding
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
Link: </riak/low_value>; rel="up"
Last-Modified: Sun, 31 Jul 2011 14:58:41 GMT
ETag: "1AUIZPR9pSnCGlgHrLTUBs"
Date: Sun, 31 Jul 2011 15:00:44 GMT
Content-Type: text/plain
Content-Length: 13
```

```
inserted data.
```

Find Your Data Again

```
$ find ./dev/dev*/data/bitcask/*/*.data
```

```
./dev/dev1/data/bitcask/867766597165223607683437869425293042920709947392/1312124444.bitcask.data  
./dev/dev4/data/bitcask/867766597165223607683437869425293042920709947392/1312124321.bitcask.data  
./dev/dev5/data/bitcask/890602560248518965780370444936484965102833893376/1312124321.bitcask.data
```

Stop Another Primary Node

```
$ dev[n]/bin/riak stop
```

```
$ curl -i http://127.0.0.1:8091/riak/default/[key]
```

```
HTTP/1.1 404 Object Not Found
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
Date: Sun, 31 Jul 2011 15:03:19 GMT
Content-Type: text/plain
Content-Length: 10
```

```
not found
```

not found

Query Again

```
$ curl -i http://127.0.0.1:8091/riak/default/[key]
```

```
HTTP/1.1 404 Object Not Found
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
Date: Sun, 31 Jul 2011 15:04:51 GMT
Content-Type: text/plain
Content-Length: 10

not found
```

Check What's Happening

```
$ find ./dev/dev*/data/bitcask/*/*.data
```

```
./dev/dev1/data/bitcask/867766597165223607683437869425293042920709947392/1312124444.bitcask.data  
./dev/dev4/data/bitcask/867766597165223607683437869425293042920709947392/1312124321.bitcask.data  
./dev/dev5/data/bitcask/890602560248518965780370444936484965102833893376/1312124321.bitcask.data
```

```
$ ./dev/dev[n]/bin/riak-admin transfers
```

```
Nodes ['dev4@127.0.0.1', 'dev5@127.0.0.1'] are currently down.  
'dev2@127.0.0.1' waiting to handoff 1 partitions  
'dev1@127.0.0.1' waiting to handoff 2 partitions
```

Add More Data

```
$ curl -v -d 'new data' -H "Content-Type: text/plain" http://127.0.0.1:8091/riak/low_value
```

```
* About to connect() to 127.0.0.1 port 8091 (#0)
*   Trying 127.0.0.1... connected
* Connected to 127.0.0.1 (127.0.0.1) port 8091 (#0)
> POST /riak/low_value HTTP/1.1
> User-Agent: curl/7.21.4 (universal-apple-darwin11.0) libcurl/7.21.4 OpenSSL/0.9.8r zlib/1.2.5
> Host: 127.0.0.1:8091
> Accept: */*
> Content-Type: text/plain
> Content-Length: 8
>
< HTTP/1.1 201 Created
< Vary: Accept-Encoding
< Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
< Location: /riak/low_value/WLoiwv0nFgfGmQKMTQdr4ofimRB
< Date: Sun, 31 Jul 2011 15:10:07 GMT
< Content-Type: application/json
< Content-Length: 0
<
* Connection #0 to host 127.0.0.1 left intact
* Closing connection #0
```

Confirm Data Exists

```
$ curl -i http://127.0.0.1:8091/riak/low_value?keys=true
```

```
HTTP/1.1 200 OK
```

```
Vary: Accept-Encoding
```

```
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
```

```
Link: </riak/low_value/WLoiww0nFgfGmQKMTQdr4ofimRB>; riaktag="contained"
```

```
Date: Sun, 31 Jul 2011 15:11:06 GMT
```

```
Content-Type: application/json
```

```
Content-Length: 452
```

```
{"props":{"allow_mult":false,"basic_quorum":true,"big_vclock":50,"chash_keyfun":{"mod":"riak_core_util","fun":"chash_std_keyfun"},"dw":"quorum","last_write_wins":false,"linkfun":{"mod":"riak_kv_wm_link_walker","fun":"mapreduce_linkfun"},"n_val":2,"name":"low_value","notfound_ok":false,"old_vclock":86400,"postcommit":[],"pr":0,"precommit":[],"pw":0,"r":1,"rw":"quorum","small_vclock":10,"w":1,"young_vclock":20},"keys":["WLoiww0nFgfGmQKMTQdr4ofimRB"]}
```


Start Down Nodes

```
$ dev[n]/bin/riak start
```

```
$ dev[n]/bin/riak start
```

```
$ ./dev/dev1/bin/riak-admin transfers
```

```
Nodes ['dev5@127.0.0.1'] are currently down.  
'dev2@127.0.0.1' waiting to handoff 2 partitions  
'dev1@127.0.0.1' waiting to handoff 2 partitions
```

Confirm Data Exists

```
$ curl -i http://127.0.0.1:8091/riak/low_value?keys=true
```

```
HTTP/1.1 200 OK
Vary: Accept-Encoding
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
Link: </riak/low_value/AtgiNS6YGoQ0qlvv2XMZLWhB0J5>; riaktag="contained", </riak/low_value/WLoiww0nFgfGmQKMTQdr4ofimRB>; riaktag="contained"
Date: Sun, 31 Jul 2011 15:15:40 GMT
Content-Type: application/json
Content-Length: 482
```

```
{"props":{"allow_mult":false,"basic_quorum":true,"big_vclock":50,"chash_keyfun":{"mod":"riak_core_util","fun":"chash_std_keyfun"},"dw":"quorum","last_write_wins":false,"linkfun":{"mod":"riak_kv_wm_link_walker","fun":"mapreduce_linkfun"},"n_val":2,"name":"low_value","notfound_ok":false,"old_vclock":86400,"postcommit":[],"pr":0,"precommit":[],"pw":0,"r":1,"rw":"quorum","small_vclock":10,"w":1,"young_vclock":20},"keys":["AtgiNS6YGoQ0qlvv2XMZLWhB0J5","WLoiww0nFgfGmQKMTQdr4ofimRB"]}
```

Summary

- hintful-handoff does not happen *
- 1 node outage max for reads
- 4 node outage max for writes
- great for fire and forget data like logs

* Look at the logs for the nodes you brought down. It will show data transfer from hand-off even though the available nodes do not serve the read requests. This may or may not happen randomly due to Eventual Consistency. One other note is that $R + W = N$ cannot guarantee consistency.

Clean Your Cluster

```
$ dev1/bin/riak stop
$ dev2/bin/riak stop
$ dev3/bin/riak stop
$ dev4/bin/riak stop
$ dev5/bin/riak stop
$ cd .. && rm -rf ./dev/*
$ make devrel
$ dev1/bin/riak start
$ dev2/bin/riak start
$ dev3/bin/riak start
$ dev4/bin/riak start
$ dev5/bin/riak start
```


Mission Critical Data



$n = 5, r = 1, w = 5, dw = 5$

Set Bucket Properties

```
* About to connect() to 127.0.0.1 port 8091 (#0)
*   Trying 127.0.0.1... connected
* Connected to 127.0.0.1 (127.0.0.1) port 8091 (#0)
> PUT /riak/high_value HTTP/1.1
> User-Agent: curl/7.21.4 (universal-apple-darwin11.0) libcurl/7.21.4 OpenSSL/0.9.8r zlib/1.2.5
> Host: 127.0.0.1:8091
> Accept: */*
> Content-Type: application/json
> Content-Length: 40
>
< HTTP/1.1 204 No Content
< Vary: Accept-Encoding
< Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
< Date: Sun, 31 Jul 2011 16:01:37 GMT
< Content-Type: application/json
< Content-Length: 0
<
* Connection #0 to host 127.0.0.1 left intact
* Closing connection #0
```

Set Bucket Properties

```
$ curl -v -X PUT -H "Content-Type: application/json" -d  
'{"props":{"n_val":5,"r":1,"w":5,"dw":5}}' http://  
127.0.0.1:8091/riak/high_value
```

```
* About to connect() to 127.0.0.1 port 8091 (#0)  
*   Trying 127.0.0.1... connected  
* Connected to 127.0.0.1 (127.0.0.1) port 8091 (#0)  
> PUT /riak/high_value HTTP/1.1  
> User-Agent: curl/7.21.4 (universal-apple-darwin11.0) libcurl/7.21.4 OpenSSL/0.9.8r zlib/1.2.5  
> Host: 127.0.0.1:8091  
> Accept: */*  
> Content-Type: application/json  
> Content-Length: 40  
>  
< HTTP/1.1 204 No Content  
< Vary: Accept-Encoding  
< Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)  
< Date: Sun, 31 Jul 2011 16:01:37 GMT  
< Content-Type: application/json  
< Content-Length: 0  
<  
* Connection #0 to host 127.0.0.1 left intact  
* Closing connection #0  
_
```


Add Data

```
$ curl -v -d 'inserted data' -H "Content-Type: text/plain"  
http://127.0.0.1:8091/riak/high_value
```

```
* About to connect() to 127.0.0.1 port 8091 (#0)  
*   Trying 127.0.0.1... connected  
* Connected to 127.0.0.1 (127.0.0.1) port 8091 (#0)  
> POST /riak/high_value HTTP/1.1  
> User-Agent: curl/7.21.4 (universal-apple-darwin11.0) libcurl/7.21.4 OpenSSL/0.9.8r zlib/1.2.5  
> Host: 127.0.0.1:8091  
> Accept: */*  
> Content-Type: text/plain  
> Content-Length: 13  
>  
< HTTP/1.1 201 Created  
< Vary: Accept-Encoding  
< Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)  
< Location: /riak/high_value/Lh2mPmJjYJ6Q7BUdd86HrGFhJl4  
< Date: Sun, 31 Jul 2011 16:04:06 GMT  
< Content-Type: application/json  
< Content-Length: 0  
<  
* Connection #0 to host 127.0.0.1 left intact  
* Closing connection #0  
_
```


Confirm Data Exists

```
$ curl -i http://127.0.0.1:8091/riak/high_value?keys=true
```

```
HTTP/1.1 200 OK
```

```
Vary: Accept-Encoding
```

```
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
```

```
Link: </riak/high_value/Lh2mPmJjYJ6Q7BUdd86HrGFhJl4>; riaktag="contained"
```

```
Date: Sun, 31 Jul 2011 16:06:50 GMT
```

```
Content-Type: application/json
```

```
Content-Length: 446
```

```
{"props":{"allow_mult":false,"basic_quorum":true,"big_vclock":50,"chash_keyfun":{"mod":"riak_core_util","fun":"chash_std_keyfun"},"dw":5,"last_write_wins":false,"linkfun":{"mod":"riak_kv_wm_link_walker","fun":"mapreduce_linkfun"},"n_val":5,"name":"high_value","notfound_ok":false,"old_vclock":86400,"postcommit":[],"pr":0,"precommit":[],"pw":0,"r":1,"rw":"quorum","small_vclock":10,"w":5,"young_vclock":20},"keys":["Lh2mPmJjYJ6Q7BUdd86HrGFhJl4"]}
```

Find Your Data

```
$ find ./dev/dev*/data/bitcask/*/*.data
```

```
./dev/dev1/data/bitcask/1370157784997721485815954530671515330927436759040/1312128246.bitcask.dat  
a  
./dev/dev2/data/bitcask/1278813932664540053428224228626747642198940975104/1312128246.bitcask.dat  
a  
./dev/dev3/data/bitcask/1301649895747835411525156804137939564381064921088/1312128246.bitcask.dat  
a  
./dev/dev4/data/bitcask/1324485858831130769622089379649131486563188867072/1312128246.bitcask.dat  
a  
./dev/dev5/data/bitcask/1347321821914426127719021955160323408745312813056/1312128246.bitcask.dat  
a
```

Stop Any 2 Nodes

```
$ dev[n]/bin/riak stop
```

```
$ dev[n]/bin/riak stop
```

```
$ curl -i http://127.0.0.1:8091/riak/high_value/[key]
```

```
HTTP/1.1 200 OK
X-Riak-Vclock: a85hYGBgzGDKBVISLE9vsmQwJTLmsTJ8+7Hy0F8WAA==
Vary: Accept-Encoding
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
Link: </riak/high_value>; rel="up"
Last-Modified: Sun, 31 Jul 2011 16:04:06 GMT
ETag: "6sURQsGQKKkUQwKiS1l4xz"
Date: Sun, 31 Jul 2011 16:09:14 GMT
Content-Type: text/plain
Content-Length: 13
```

```
inserted data:
```

Check What's Happening

```
$ find ./dev/dev*/data/bitcask/*/*.data
```

```
./dev/dev1/data/bitcask/1370157784997721485815954530671515330927436759040/1312128246.bitcask.dat
a
./dev/dev2/data/bitcask/1278813932664540053428224228626747642198940975104/1312128246.bitcask.dat
a
./dev/dev2/data/bitcask/1370157784997721485815954530671515330927436759040/1312128554.bitcask.dat
a
./dev/dev3/data/bitcask/1301649895747835411525156804137939564381064921088/1312128246.bitcask.dat
a
./dev/dev4/data/bitcask/1301649895747835411525156804137939564381064921088/1312128554.bitcask.dat
a
./dev/dev4/data/bitcask/1324485858831130769622089379649131486563188867072/1312128246.bitcask.dat
a
./dev/dev5/data/bitcask/1347321821914426127719021955160323408745312813056/1312128246.bitcask.dat
a
```

```
$ ./dev/dev[n]/bin/riak-admin transfers
```

```
Nodes ['dev1@127.0.0.1', 'dev3@127.0.0.1'] are currently down.
'dev4@127.0.0.1' waiting to handoff 1 partitions
'dev2@127.0.0.1' waiting to handoff 1 partitions
```


Stop Another Node

```
$ dev[n]/bin/riak stop
```

```
$ curl -i http://127.0.0.1:8091/riak/high_value/[key]
```

```
HTTP/1.1 200 OK
X-Riak-Vclock: a85hYGBgzGDKBVISLE9vsmQwJTLmSTJ8+7Hy0F8WAA==
Vary: Accept-Encoding
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
Link: </riak/high_value>; rel="up"
Last-Modified: Sun, 31 Jul 2011 16:04:06 GMT
ETag: "6sURQsGQKKkUQwKiS1l4xz"
Date: Sun, 31 Jul 2011 16:12:43 GMT
Content-Type: text/plain
Content-Length: 13
```

inserted data

Check What's Happening

```
$ find ./dev/dev*/data/bitcask/*/*.data
```

```
./dev/dev1/data/bitcask/1370157784997721485815954530671515330927436759040/1312128246.bitcask.dat
a
./dev/dev2/data/bitcask/1278813932664540053428224228626747642198940975104/1312128246.bitcask.dat
a
./dev/dev2/data/bitcask/1301649895747835411525156804137939564381064921088/1312128763.bitcask.dat
a
./dev/dev2/data/bitcask/1370157784997721485815954530671515330927436759040/1312128554.bitcask.dat
a
./dev/dev3/data/bitcask/1301649895747835411525156804137939564381064921088/1312128246.bitcask.dat
a
./dev/dev4/data/bitcask/1301649895747835411525156804137939564381064921088/1312128554.bitcask.dat
a
./dev/dev4/data/bitcask/1324485858831130769622089379649131486563188867072/1312128246.bitcask.dat
a
./dev/dev4/data/bitcask/1347321821914426127719021955160323408745312813056/1312128763.bitcask.dat
a
./dev/dev5/data/bitcask/1347321821914426127719021955160323408745312813056/1312128246.bitcask.dat
a
```

```
$ ./dev/dev[n]/bin/riak-admin transfers
```

```
Nodes ['dev1@127.0.0.1', 'dev3@127.0.0.1', 'dev5@127.0.0.1'] are currently down.
'dev4@127.0.0.1' waiting to handoff 2 partitions
'dev2@127.0.0.1' waiting to handoff 2 partitions
```

Update Data

```
$ curl -v -d 'updated data' -H "Content-Type: text/plain"  
http://127.0.0.1:8091/riak/high_value/[key]
```

```
* About to connect() to 127.0.0.1 port 8092 (#0)  
*   Trying 127.0.0.1... connected  
* Connected to 127.0.0.1 (127.0.0.1) port 8092 (#0)  
> POST /riak/high_value/Lh2mPmJjYJ6Q7BUdd86HrGFhJl4 HTTP/1.1  
> User-Agent: curl/7.21.4 (universal-apple-darwin11.0) libcurl/7.21.4 OpenSSL/0.9.8r zlib/1.2.5  
> Host: 127.0.0.1:8092  
> Accept: */*  
> Content-Type: text/plain  
> Content-Length: 12  
>  
< HTTP/1.1 204 No Content  
< Vary: Accept-Encoding  
< Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)  
< Date: Sun, 31 Jul 2011 16:17:15 GMT  
< Content-Type: text/plain  
< Content-Length: 0  
<  
* Connection #0 to host 127.0.0.1 left intact  
* Closing connection #0
```

Query Data

```
$ curl -i http://127.0.0.1:8091/riak/high_value/[key]
```

```
HTTP/1.1 200 OK
X-Riak-Vclock: a85hYGDgzWDKBVIIsLE9vsmQwJTLmsTJ8+7Hy0B9UmD3iQxpUmPsPXJit0YklWrUVU4KFKcBv01TYGkU9Q
37ucUwJFjburp1QYQ8U9axp2gcxJVgY0xp5ocLRK0oZ102fjSnBwprayQgVLkBRz/whSA5TgoVVSS0eKlyH6v7q+6LIElka
Vary: Accept-Encoding
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
Link: </riak/high_value>; rel="up"
Last-Modified: Sun, 31 Jul 2011 16:19:10 GMT
ETag: "7A09xv8czBX0uoS2NfZsZ0"
Date: Sun, 31 Jul 2011 16:19:27 GMT
Content-Type: text/plain
Content-Length: 12
```

updated data

Start Down Nodes

```
$ dev[n]/bin/riak start
```

```
$ dev[n]/bin/riak start
```

```
$ dev[n]/bin/riak start
```

```
$ ./dev/dev1/bin/riak-admin transfers
```

No transfers active

Confirm Data Exists

```
$ curl -i http://127.0.0.1:8091/riak/high_value?keys=true
```

```
HTTP/1.1 200 OK
```

```
Vary: Accept-Encoding
```

```
Server: MochiWeb/1.1 WebMachine/1.7.3 (participate in the frantic)
```

```
Link: </riak/high_value/Lh2mPmJjYJ6Q7BUdd86HrGFhJl4>; riaktag="contained"
```

```
Date: Sun, 31 Jul 2011 16:27:07 GMT
```

```
Content-Type: application/json
```

```
Content-Length: 446
```

```
{"props":{"allow_mult":false,"basic_quorum":true,"big_vclock":50,"chash_keyfun":{"mod":"riak_core_util","fun":"chash_std_keyfun"},"dw":5,"last_write_wins":false,"linkfun":{"mod":"riak_kv_wm_link_walker","fun":"mapreduce_linkfun"},"n_val":5,"name":"high_value","notfound_ok":false,"old_vclock":86400,"postcommit":[],"pr":0,"precommit":[],"pw":0,"r":1,"rw":"quorum","small_vclock":10,"w":5,"young_vclock":20},"keys":["Lh2mPmJjYJ6Q7BUdd86HrGFhJl4"]}
```

Summary

- data will duplicate on available nodes to ensure 5 copies
- 4 node outage max for reads
- 4 node outage max for writes
- slower due to all data writes
- great for high availability data

Total Destruction



Total Destruction



Summary

- easy to get up and going
- scales as you need it
- flexible configuration for consistency sake
- eventual consistency doesn't always do what you expect
- just scratched the surface

Great Reading

- <http://www.basho.com>
- <http://bit.ly/eG08u0>
- <http://bit.ly/eFfcjb>
- <http://bit.ly/ejCClJ>
- <http://slidesha.re/g3hQfg>
- <http://bit.ly/n6hWMK>