

Flat Hierarchies Testing

Tests to determine whether Fedora 4 ingest performance is impacted when there is a lack of hierarchy in the repository structure. There are two sets of tests here.

The first hypothesis was that we'd see a slow down in ingest time at the ten thousand [container](#) mark. To test this, we ran tests of 30k records using [bench tool](#) and a [Bash script](#) written to alternate between the Infinispan configurations available in the [fcrepo4/fcrepo-configs](#) repository (e.g., [file](#), [leveldb-default](#), [leveldb](#), and [ram](#)). The tests were run on an [EC2 m3.medium instance](#). The initial tests, documented below, did not reveal a problem with a lack of hierarchy. However, there was an uptick at the end of the 30k 2 MB container tests that indicated more testing was warranted. When 60k container tests were run with benchtool, there were exceptions thrown and the tests ended prematurely.

The second round of tests were designed to load a larger number of containers. The initial Bash script [was modified](#) to use curl, instead of benchtool, to submit containers to Fedora 4. These tests were single threaded and attempted to load 100k, and then 500k, containers with 1 MB [binary](#) into a flat fcrepo4 structure. These tests ran without exception, but showed a definite slow down in the containers' ingest times as the number of containers ingested continued to grow. The second test (of 500k containers) was discontinued at 175k containers when the increase in ingest times was determined to be growing to the point of being prohibitively slow. The graphs for these tests are near the bottom of the page (the last two red graphs).

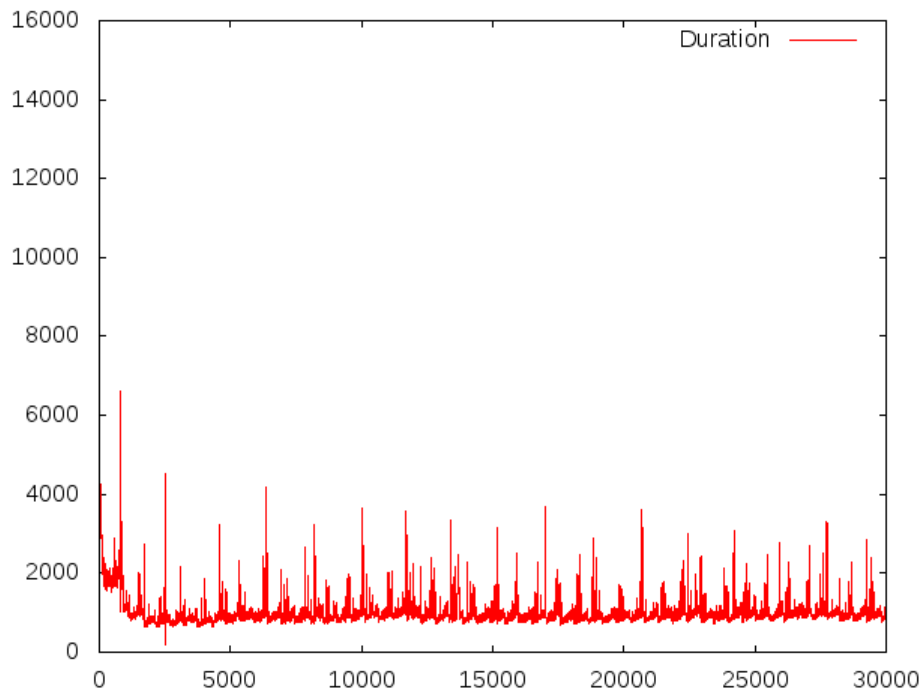
At the very bottom of this page are graphs from a test run by Ben Pennell, which shows a similar pattern in the timing of ingests of over 100k containers. He experienced the repository becoming non-responsive near the 200k container mark.

Updated (at the bottom of the page) with results from a hierarchical test (same architecture, but using a 64/64/64 structure).

Results from the First Round of Testing

| Configuration | Results Graph | Logging Output |
|---|---------------|---|
| leveldb-default (30k 1 KB containers / 15 threads) - tested using benchtool | | Running new test [config: config /infinispan /leveldb-default /infinispan.xml] [Jetty pid: 30298] 18:30:11 INFO Found Fedora 4 at http://localhost:8080 18:30:11 INFO Running 30000 INGEST action(s) against FCREPO4 with a binary size of 1.0 KB using 15 thread(s) 18:30:15 INFO The Fedora cluster has 0 node(s) before the benchmark 18:30:15 INFO preparing 30000 objects 18:32:30 INFO creating 30000 objects |

| | | |
|---|--|--|
| | | <p>took 125639 ms 18:33:27 INFO scheduling 30000 actions 19:08:14 INFO purging 30000 objects and datastreams</p> <p>19:10:38 INFO Completed 30000 INGEST action(s) executed in 31292451 ms 19:10:38 INFO The Fedora cluster has 0 node(s) after the benchmark 19:10:38 INFO Throughput was 0.01 MB/sec 19:10:38 INFO Throughput per thread was 0 MB /sec 19:10:38 INFO Condensed results: 19:10:38 INFO 30000 1024 15 INGEST 31292451 9.3622814 E-4 no-tx 19:10:38 INFO All operations completed in 2426697 ms</p> |
| leveldb (30k 1 KB containers / 15 threads) - tested using benchtool | | <p>Running new test [config: config /infinispan /leveldb /infinispan. xml] [Jetty pid: 3870] 19:12:40 INFO Found Fedora 4 at http://localhost:8080 19:12:40</p> |



INFO
Running
30000
INGEST
action(s)
against
FCREPO4
with a
binary size
of 1.0 KB
using 15
thread(s)
19:12:44
INFO The
Fedora
cluster has
0 node(s)
before the
benchmark
19:12:44
INFO
preparing
30000
objects
19:15:01
INFO
creating
30000
objects
took
126888 ms
19:16:04
INFO
scheduling
30000
actions
19:51:06
INFO
purging
30000
objects
and
datastreams
19:53:28
INFO
Completed
30000
INGEST
action(s)
executed
in
31517411
ms
19:53:29
INFO The
Fedora
cluster has
0 node(s)
after the
benchmark
19:53:29
INFO
Throughput
was 0.01
MB/sec
19:53:29
INFO
Throughput
per thread
was 0 MB
/sec
19:53:29
INFO
Condensed
results:
19:53:29
INFO
30000

1024 15
INGEST
31517411
9.295457E-
4 no-tx
19:53:29
INFO All
operations
completed
in
2448012
ms

Running
new test
[config:
config
/infinispan
/file
/infinispan.
xml] [Jetty
pid: 7558]
20:02:31
INFO
Found
Fedora 4
at <http://localhost:8080>

20:02:31
INFO
Running
30000
INGEST
action(s)
against
FCREPO4
with a
binary size
of 1.0 KB
using 15
thread(s)

20:02:34
INFO The
Fedora
cluster has
0 node(s)
before the
benchmark

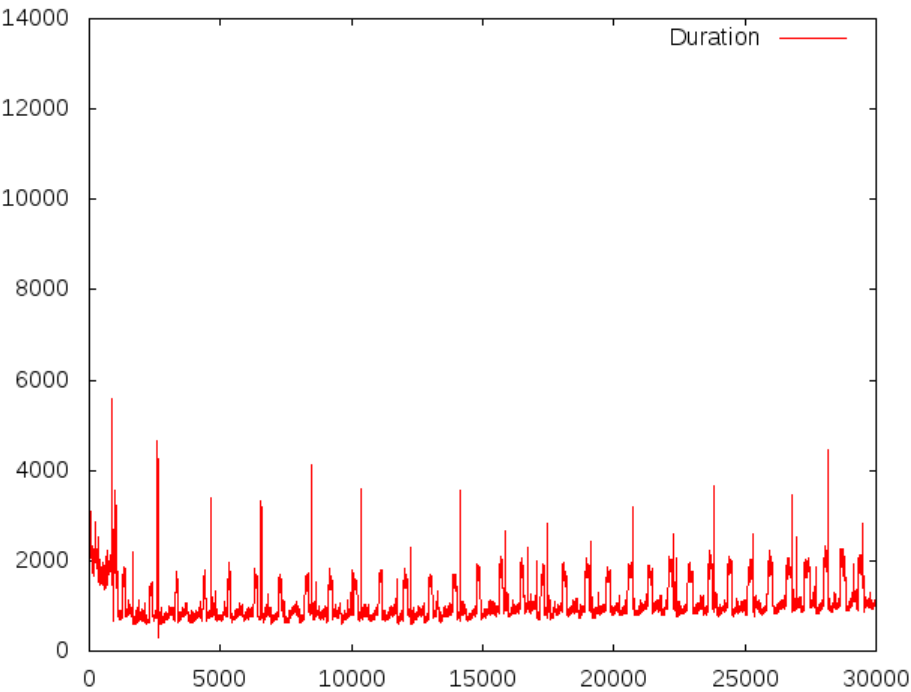
20:02:34
INFO
preparing
30000
objects
20:04:47
INFO
creating
30000
objects
took

124113 ms
20:08:28
INFO
scheduling
30000
actions
20:45:14
INFO

purging
30000
objects
and
datastreams

20:50:51
INFO
Completed
30000
INGEST
action(s)

file (30k 1 KB containers /
15 threads) - tested using
benchtool

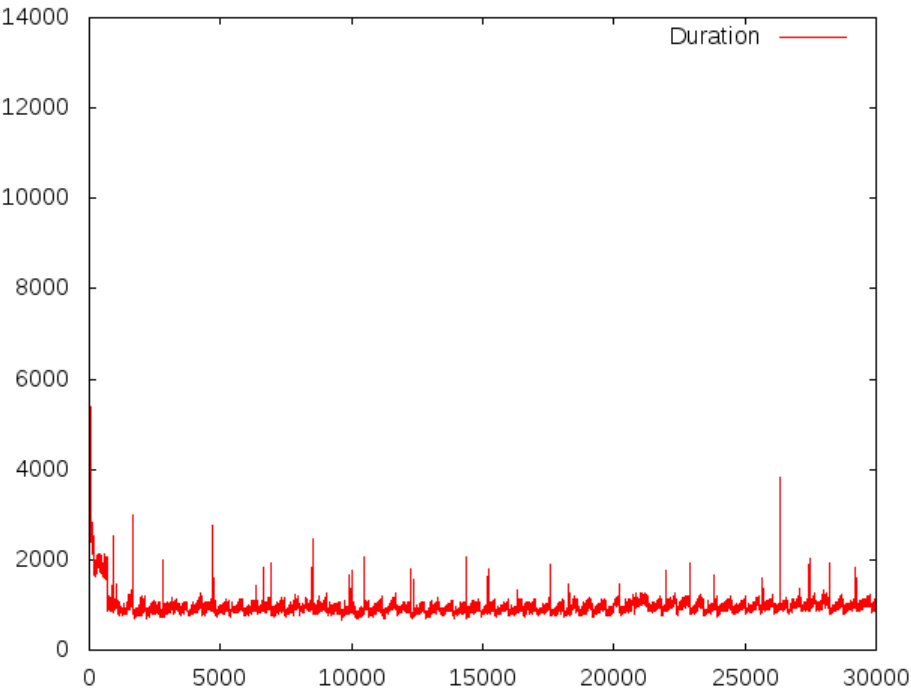


executed
in
33069648
ms
20:50:52
INFO The
Fedora
cluster has
0 node(s)
after the
benchmark
20:50:52
INFO
Throughput
was 0.01
MB/sec
20:50:52
INFO
Throughput
per thread
was 0 MB
/sec
20:50:52
INFO
Condensed
results:
20:50:52
INFO
30000
1024 15
INGEST
33069648
8.859143E-
4 no-tx
20:50:52
INFO All
operations
completed
in
2900327
ms

Running
new test
[config:
config
/infinispan
/ram
/infinispan.
xml] [Jetty
pid: 11324]
20:52:45
INFO
Found
Fedora 4
at <http://localhost:8080>

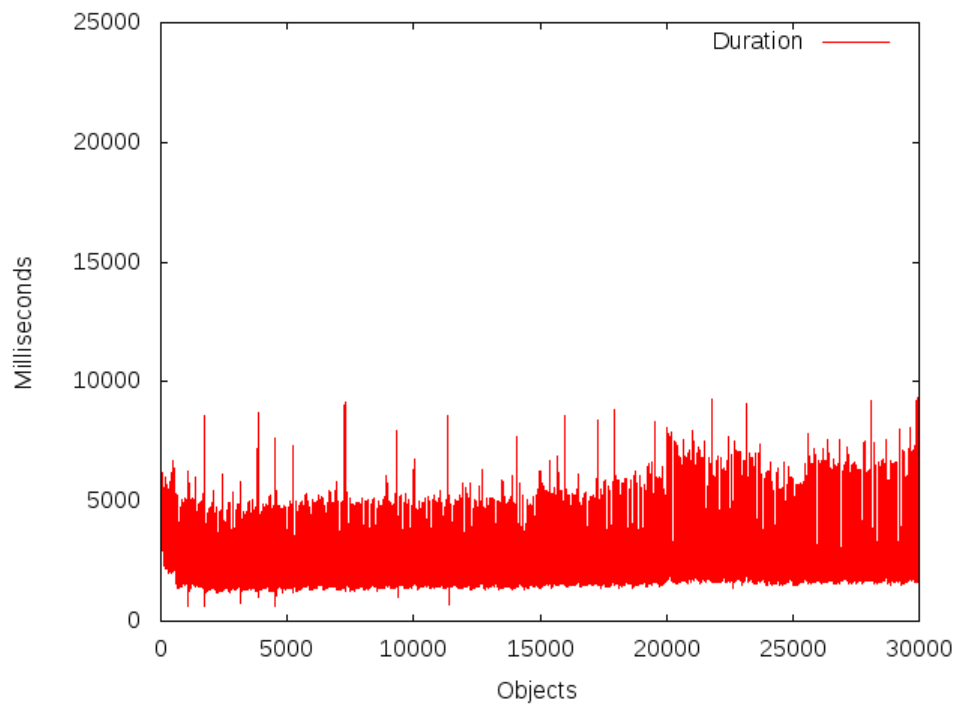
20:52:45
INFO
Running
30000
INGEST
action(s)
against
FCREPO4
with a
binary size
of 1.0 KB
using 15
thread(s)
20:52:48
INFO The
Fedora
cluster has
0 node(s)
before the
benchmark
20:52:48
INFO

ram (30k 1 KB containers /
15 threads) - tested using
benchtool



| | | |
|--|--|---|
| | | <pre>preparing 30000 objects 20:55:00 INFO creating 30000 objects took 123502 ms 20:55:45 INFO scheduling 30000 actions 21:28:38 INFO purging 30000 objects and datastreams 21:30:23 INFO Completed 30000 INGEST action(s) executed in 29588918 ms 21:30:23 INFO The Fedora cluster has 0 node(s) after the benchmark 21:30:23 INFO Throughput was 0.01 MB/sec 21:30:23 INFO Throughput per thread was 0 MB /sec 21:30:23 INFO Condensed results: 21:30:23 INFO 30000 1024 15 INGEST 29588918 9.9013E-4 no-tx 21:30:23 INFO All operations completed in 2257950 ms</pre> |
|--|--|---|

| Configuration | Results Graph | Logging Output |
|---|---------------|----------------------------------|
| leveldb-default (30k 2 MB containers / 15 threads) - tested using benchtool | | Running new test [config: config |



```
/infinispan
/leveldb-
default
/infinispan.
.xml] [Jetty
pid: 1370]
14:10:27
INFO
Found
Fedora 4
at http://localhost:8080

14:10:27
INFO
Running
30000
INGEST
action(s)
against
FCREPO4
with a
binary size
of 2.0 MB
using 15
thread(s)
14:10:29
INFO The
Fedora
cluster has
0 node(s)
before the
benchmark
14:10:29
INFO
preparing
30000
objects
14:12:40
INFO
creating
30000
objects
took
121853 ms
14:13:34
INFO
scheduling
30000
actions
15:48:36
INFO
purging
30000
objects
and
datastreams

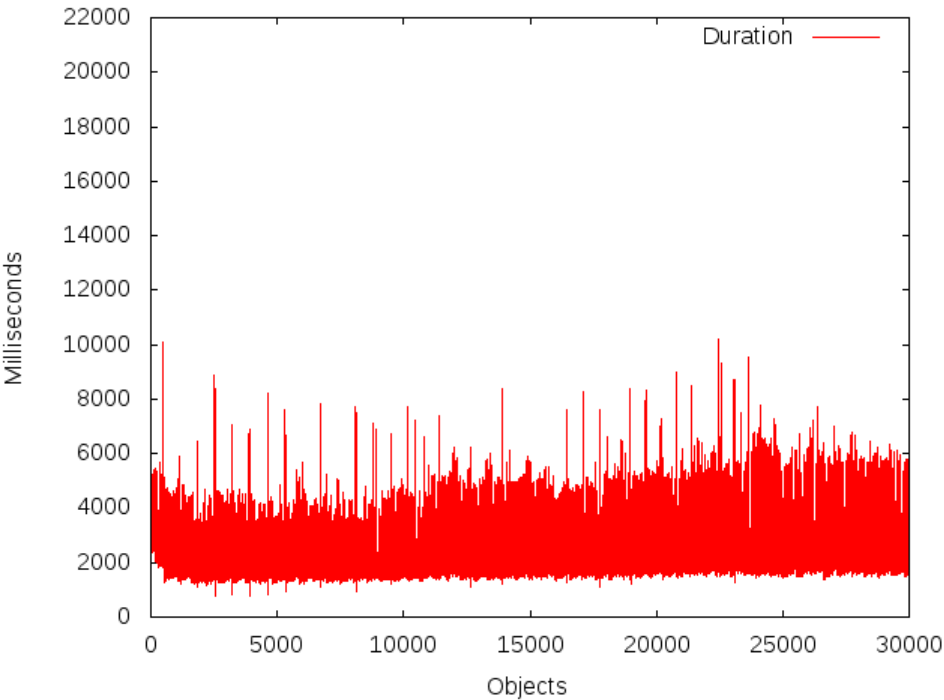
15:51:10
INFO
Completed
30000
INGEST
action(s)
executed
in
85507949
ms
15:51:10
INFO The
Fedora
cluster has
0 node(s)
after the
benchmark
15:51:10
INFO
Throughput
was 10.53
```

MB/sec
15:51:10
INFO
Throughput
per thread
was 0.7
MB/sec
15:51:10
INFO
Condensed
results:
15:51:10
INFO
30000
2097152
15
INGEST
85507949
0.7016891
no-tx
15:51:10
INFO All
operations
completed
in
6042901
ms

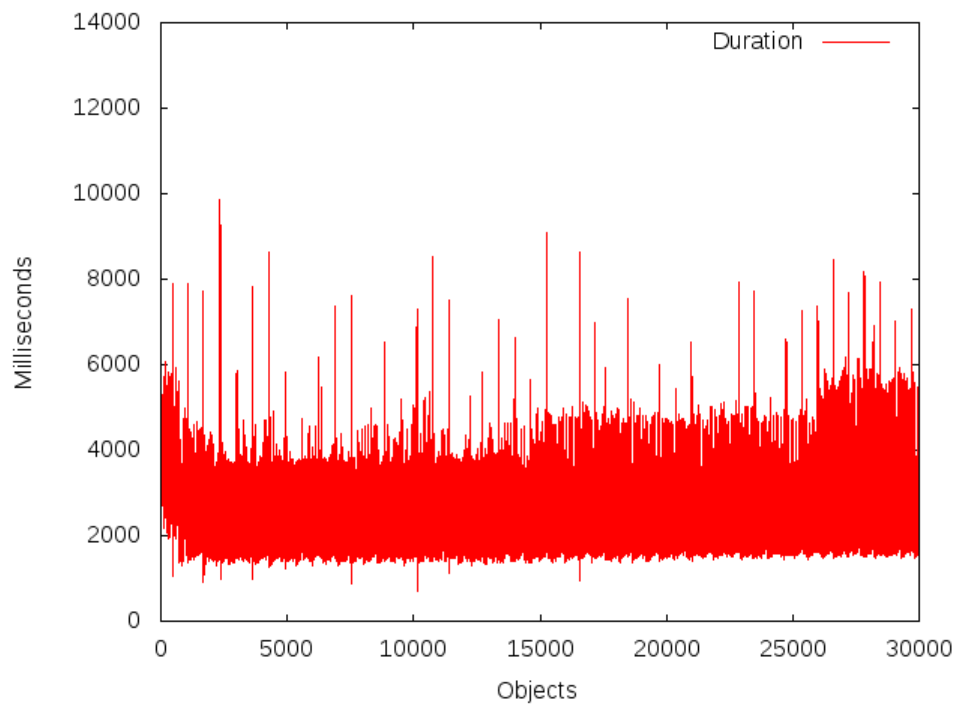
Running
new test
[config:
config
/infinispan
/leveldb
/infinispan.
xml] [Jetty
pid: 7694]
18:06:30
INFO
Found
Fedora 4
at <http://localhost:8080>

18:06:30
INFO
Running
30000
INGEST
action(s)
against
FCREPO4
with a
binary size
of 2.0 MB
using 15
thread(s)
18:06:32
INFO The
Fedora
cluster has
0 node(s)
before the
benchmark
18:06:32
INFO
preparing
30000
objects
18:08:41
INFO
creating
30000
objects
took
120016 ms
18:09:34
INFO
scheduling
30000

leveldb (30k 2 MB
containers / 15 threads) -
tested using benchtool



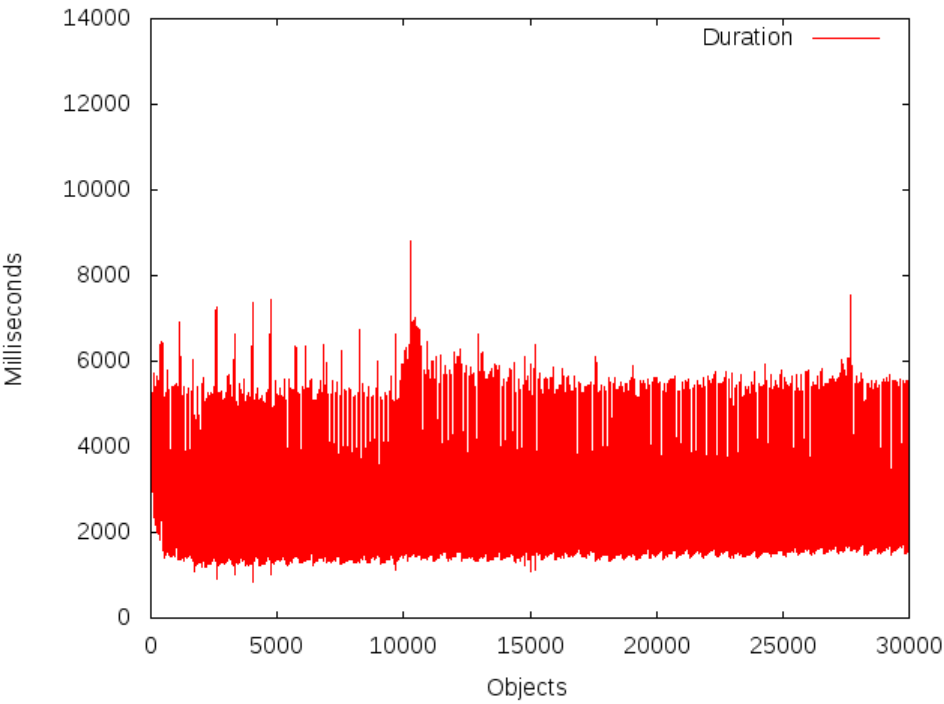
| | | |
|--|--|---|
| | | <p>actions 19:45:21 INFO purging 30000 objects and datastreams</p> <p>19:47:52 INFO Completed 30000 INGEST action(s) executed in 86194079 ms 19:47:52 INFO The Fedora cluster has 0 node(s) after the benchmark 19:47:52 INFO Throughput was 10.44 MB/sec 19:47:52 INFO Throughput per thread was 0.7 MB/sec 19:47:52 INFO Condensed results: 19:47:52 INFO 30000 2097152 15 INGEST 86194079 0.6961035 no-tx 19:47:52 INFO All operations completed in 6081741 ms</p> |
| file (30k 2 MB containers / 15 threads) - tested using benchtool | | <p>Running new test [config: config /infinispan /file /infinispan. xml] [Jetty pid: 14070] 21:45:20 INFO Found Fedora 4 at http://localhost:8080</p> <p>21:45:20 INFO Running 30000 INGEST action(s)</p> |



against
 FCREPO4
 with a
 binary size
 of 2.0 MB
 using 15
 thread(s)
 21:45:22
 INFO The
 Fedora
 cluster has
 0 node(s)
 before the
 benchmark
 21:45:22
 INFO
 preparing
 30000
 objects
 21:47:36
 INFO
 creating
 30000
 objects
 took
 124645 ms
 21:51:06
 INFO
 scheduling
 30000
 actions
 23:26:46
 INFO
 purging
 30000
 objects
 and
 datastreams

 23:31:50
 INFO
 Completed
 30000
 INGEST
 action(s)
 executed
 in
 86083830
 ms
 23:31:51
 INFO The
 Fedora
 cluster has
 0 node(s)
 after the
 benchmark
 23:31:51
 INFO
 Throughput
 was 10.45
 MB/sec
 23:31:51
 INFO
 Throughput
 per thread
 was 0.7
 MB/sec
 23:31:51
 INFO
 Condensed
 results:
 23:31:51
 INFO
 30000
 2097152
 15
 INGEST
 86083830
 0.696995

ram (30k 2 MB containers / 15 threads) - tested using benchtool



no-tx
23:31:51
INFO All
operations
completed
in
6390673
ms

Running
new test
[config:
config
/infinispan
/ram
/infinispan.
xml] [Jetty
pid: 1573]
17:46:22
INFO
Found
Fedora 4
at <http://localhost:8080>

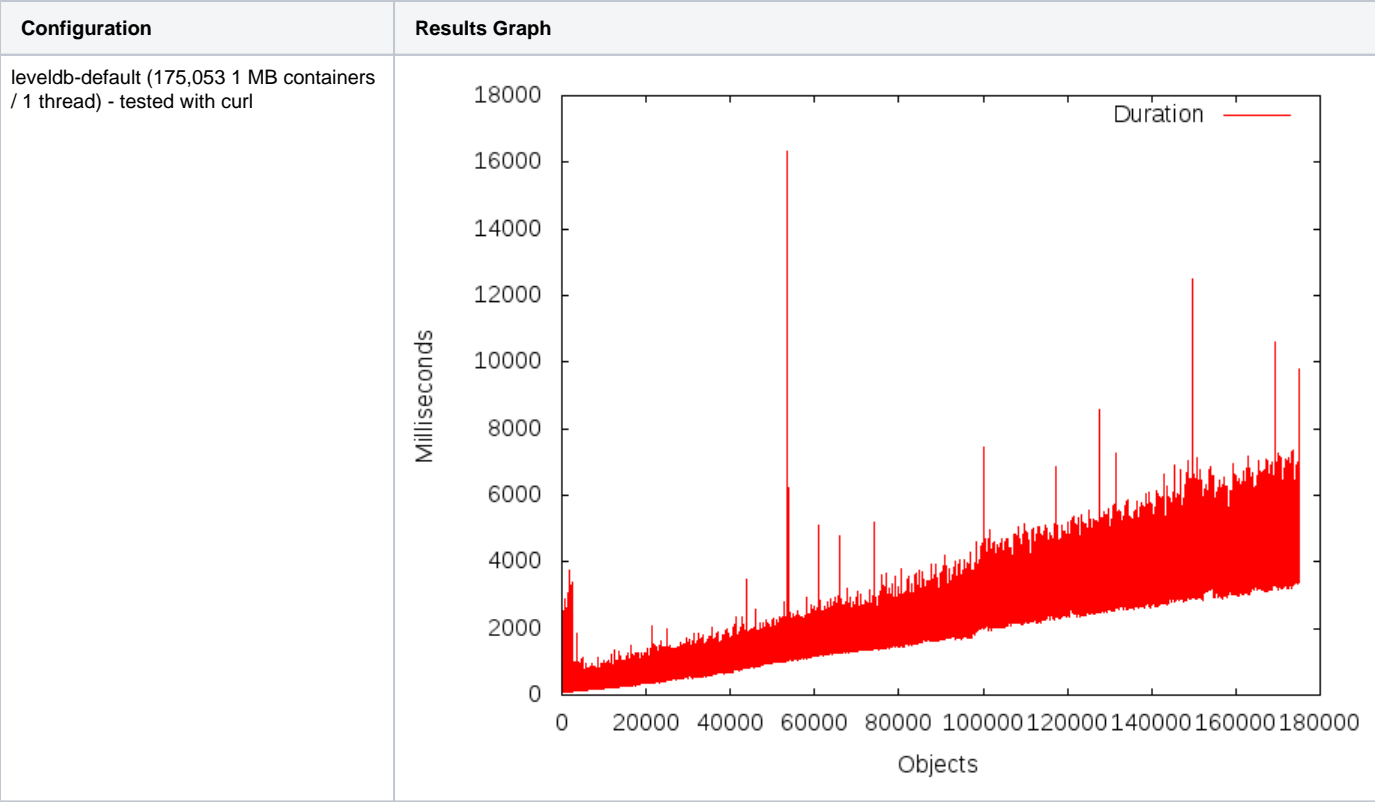
17:46:22
INFO
Running
30000
INGEST
action(s)
against
FCREPO4
with a
binary size
of 2.0 MB
using 15
thread(s)
17:46:24
INFO The
Fedora
cluster has
0 node(s)
before the
benchmark
17:46:24
INFO
preparing
30000
objects
17:48:35
INFO
creating
30000
objects
took
122076 ms
17:49:25
INFO
scheduling
30000
actions
19:18:49
INFO
purging
30000
objects
and
datastreams

19:21:02
INFO
Completed
30000
INGEST
action(s)
executed
in
80453741
ms

| | | |
|--|--|--|
| | | 19:21:02 INFO The Fedora cluster has 0 node(s) after the benchmark 19:21:02 INFO Throughput was 11.19 MB/sec 19:21:02 INFO Throughput per thread was 0.75 MB/sec 19:21:02 INFO Condensed results: 19:21:02 INFO 30000 2097152 15 INGEST 80453741 0.7457701 6 no-tx 19:21:02 INFO All operations completed in 5680046 ms |
|--|--|--|

Results from the Second Round of Testing

| Configuration | Results Graph |
|---|--|
| leveldb-default (100k 1 MB containers / 1 thread) - tested using curl | <p>The graph displays 'Duration' on the y-axis (0 to 10,000) against an unlabeled x-axis (0 to 100,000). A red line represents the data, showing a general upward trend with significant noise and several sharp spikes. The highest spike occurs at approximately x=65,000, reaching a value of about 9,500. Other notable spikes are at x=5,000 (approx. 6,000), x=55,000 (approx. 4,500), x=85,000 (approx. 5,500), and x=95,000 (approx. 6,000). The baseline noise level starts around 1,000 and increases to about 3,000 by the end of the x-axis range.</p> |



Results from Ben Pennell's Testing

