

Performance evaluation on AWS

The following numbers were gathered on an AWS cluster

Node network I/O performance

When using netcat to transfer a 5 gb file from one VM to another VM

```
ubuntu@ip-10-73-143-203:~$ time sudo bash -c "(nc -w 3 ec2-50-17-13-245.compute-1.amazonaws.com 1234 < /mnt/bf)"
```

```
real 2m56.656s
user 0m1.612s
sys 0m15.088s
```

the throughput was 5000mb/176sec = 28.04 mb/s

Node hdd performance

Root partition

```
ubuntu@ip-10-73-143-203:~$ sync;time sudo bash -c "(dd if=/dev/zero of=bf bs=8k count=500000; sync)"
500000+0 records in
500000+0 records out
4096000000 bytes (4.1 GB) copied, 513.869 s, 8.0 MB/s
```

Partition at /mnt

```
ubuntu@ip-10-73-143-203:~$ sync;time sudo bash -c "(dd if=/dev/zero of=/mnt/bf bs=8k count=500000; sync)"
500000+0 records in
500000+0 records out
4096000000 bytes (4.1 GB) copied, 34.3706 s, 119 MB/s
```

We have to make sure to use the /mnt partition since it writes much faster than the root one

CPU

Following is the output of 'cat /proc/cpuinfo' on one VM

```
processor : 0
vendor_id : GenuineIntel
cpu family : 6
model : 45
model name : Intel(R) Xeon(R) CPU E5-2650 0 @ 2.00GHz
stepping : 7
microcode : 0x70d
cpu MHz : 1795.672
cache size : 20480 KB
physical id : 1
siblings : 2
core id : 2
cpu cores : 1
apicid : 36
initial apicid : 36
fpu : yes
fpu_exception : yes
cpuid level : 13
wp : yes
flags : fpu de tsc msr pae cx8 sep cmov pat clflush mmx fxsr sse sse2 ss ht syscall nx lm constant_tsc rep_good nopl nonstop_tsc pni pclmulqdq ssse3
cx16 pcid sse4_1 sse4_2 x2apic popcnt tsc_deadline_timer aes hypervisor lahf_lm arat epb xsaveopt pln pts dtherm
bogomips : 3591.34
clflush size : 64
cache_alignment : 64
address sizes : 46 bits physical, 48 bits virtual
power management:
```

processor : 1
vendor_id : GenuineIntel
cpu family : 6
model : 45
model name : Intel(R) Xeon(R) CPU E5-2650 0 @ 2.00GHz
stepping : 7
microcode : 0x70d
cpu MHz : 1795.672
cache size : 20480 KB
physical id : 1
siblings : 2
core id : 2
cpu cores : 1
apicid : 36
initial apicid : 36
fpu : yes
fpu_exception : yes
cpuid level : 13
wp : yes
flags : fpu de tsc msr pae cx8 sep cmov pat clflush mmx fxsr sse sse2 ss ht syscall nx lm constant_tsc rep_good nopl nonstop_tsc pni pclmulqdq ssse3
cx16 pcid sse4_1 sse4_2 x2apic popcnt tsc_deadline_timer aes hypervisor lahf_lm arat epb xsaveopt pln pts dtherm
bogomips : 3591.34
clflush size : 64
cache_alignment : 64
address sizes : 46 bits physical, 48 bits virtual
power management:

Load balancing

Load balancing is done by the AWS internal loadbalancing service: <http://aws.amazon.com/elasticloadbalancing/>

Results

Test Utility

BenchTool: <https://github.com/futures/benchtool>

A. Ingest bench using fcrepo3

Created 100 objects with one datastream of 50mb size

Size: 100 * 52428800 bytes = 5gb

B. Ingest bench using single node fcrepo4 with single thread and clustered config

Created 100 objects with one datastream of 50mb size

Size: 100 * 52428800 bytes = 5gb

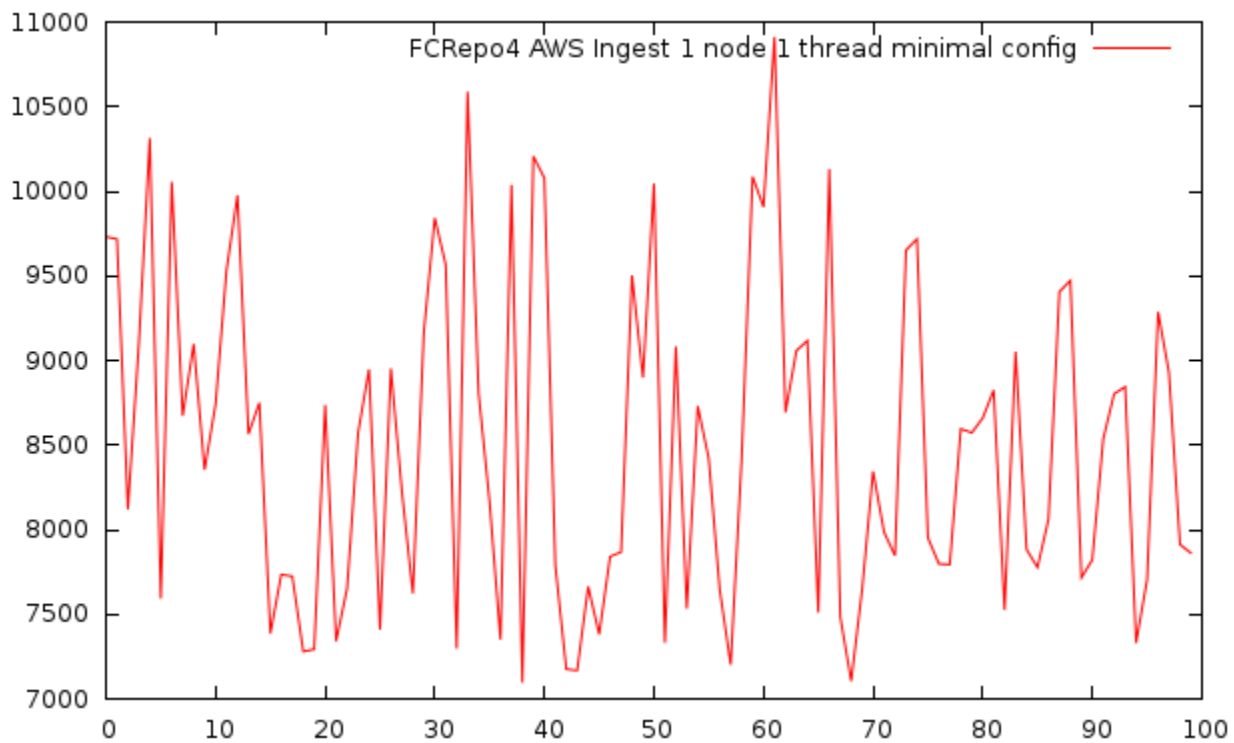
C. Ingest bench using single node fcrepo4 with single thread and minimal config

Created 100 objects with one datastream of 50mb size

Size: 100 * 52428800 bytes = 5gb

Duration: 881883 ms

Throughput was 5.81 mb/s



D. Ingest bench using seven nodes fcrepo4 with seven threads and clustered config

Created 100 objects with one datastream of 50mb size

Size: $100 * 52428800$ bytes = 5gb

E. Ingest bench using seven nodes fcrepo4 with seven threads and clustered config, no indexing

Created 100 objects with one datastream of 50mb size

Size: $100 * 52428800$ bytes = 5gb

F. Ingest bench using seven nodes fcrepo4 with eleven threads and clustered config, no indexing

Created 100 objects with one datastream of 50mb size

Size: $100 * 52428800$ bytes = 5gb