

Alluxio 效果测试

快速生成随机文本

```
base64 /dev/urandom | head -c 10000000 > example
```

```
base64 /dev/urandom > example
```

```
hadoop fs -put -f example hdfs://192.168.33.101:8020/alluxio/example
alluxio fs copyFromLocal example /example2
time alluxio fs cat /example | grep -c a
time alluxio fs cat /example2 | grep -c a
alluxio fs load /example2      #加载到内存
alluxio fs free /example      #从内存中释放
time hadoop fs -cat hdfs://192.168.33.101:8020/alluxio/example | grep -c a
```

不知道为什么通过Alluxio 并没有加速了数据访问 反而慢了





网上有人测试

<https://www.cnblogs.com/seaspring/p/6186359.html>

他的测试效果和我的一样 想用alluxio来提升单个job的性能，基本上是比较难的

alluxio应用场景：

计算层需要反复访问远程（比如在云端，或跨机房）的数据；

计算层需要同时访问多个独立的持久化数据源（比如同时访问S3和HDFS中的数据）；

多个独立的大数据应用（比如不同的Spark Job）需要高速有效的共享数据；

当计算层有着较为严重的内存资源、以及JVM GC压力，或者较高的任务失败率时，Alluxio作为输入输出数据的Off heap存储可以极大缓解这一压力，

并使计算消耗的时间和资源更可控可预测。