HDFS性能测试工具

|  |  |  |  |
| --- | --- | --- | --- |
| 日期 | 版本 | 修订者 | 修订描述 |
| 2018.01.19 | v1 | 林 清 | 初稿 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

[1. HDFS性能&压力测试 1](#_Toc504142730)

[1.1 基准测试工具 1](#_Toc504142731)

[1.1.1 TestDFSIO 1](#_Toc504142732)

[1.1.2 SliveTest 2](#_Toc504142733)

[1.1.3 Terasort 3](#_Toc504142734)

# HDFS性能&压力测试

## 基准测试工具

Hadoop自带的测试工具包括：Terasort、SliveTest、TestDFSIO。通过这些工具可以测试HDFS的I/O吞吐率、NameNode处理RPC请求的能力、同步率等等。

### TestDFSIO

位于hadoop-mapreduce-client-jobclient-x.x.x-tests.jar中（CDH有专门的版本，位于/opt/cloudera/parcels/CDHxxx/jars/）。

这个任务基于MapReduce框架进行读写测试，通过指定Map数量来指定读写的线程数目。

使用例子：

# 写测试，在文件系统中写入10个5GB的文件

yarn jar hadoop-mapreduce-client-jobclient-2.6.0-cdh5.13.1-tests.jar TestDFSIO -write -nrFiles 10 -size 5GB -resFile /tmp/dfsio.txt

# 读测试，在文件系统中读取10个5GB的文件

yarn jar hadoop-mapreduce-client-jobclient-2.6.0-cdh5.13.1-tests.jar TestDFSIO -read -nrFiles 10 -size 5GB -resFile /tmp/dfsio.txt

# 清除测试工具

yarn jar hadoop-mapreduce-client-jobclient-2.6.0-cdh5.13.1-tests.jar TestDFSIO –clean

执行目录下会生成简易的测试报告，下面的报告是在一个三副本节点上的写入测试：

----- TestDFSIO ----- : write

Date & time: Fri Jan 19 01:28:06 EST 2018

Number of files: 10 #写入的文件数目

Total MBytes processed: 51200.0 #写入的总字节数

Throughput mb/sec: 185.3091804050019 # IO吞吐量

Average IO rate mb/sec: 186.26731872558594 # 所有MAP任务的平均吞吐量

IO rate std deviation: 13.518420526488379 #上面的标准差

Test exec time sec: 346.818

### SliveTest

SliveTest和TestDFSIO位于同样的jar包中，也是基于MapReduce对HDFS进行性能测试。但是主要是测试Namenode的性能，原理是通过大量map制造多种rpc请求。我们可以设定map数量，每个map发起的rpc请求次数，每一种rpc操作占总操作的百分比，以及读写数据量、block size等配置。

SliveTest可以支持以下RPC操作，SliveTest每次测试时在Map中平均（也可指定百分比）生成下面这些RPC操作：

* ls:列出路径下所有文件和目录
* append:追加写文件操作
* create:创建文件操作
* delete:删除文件操作
* mkdir:创建目录操作
* rename:重命名文件操作
* read:读取文件内容操作

所有参数包括：

|  |  |
| --- | --- |
| maps | 一共运行多少个mapper，默认值为10 |
| ops | 每个map跑多少个操作，默认值为1000 |
| duration | 每个map task的持续时间，默认值为MAX\_INT，也就是无限制 |
| exitOnError | 遇到第一个Error是否要立即退出，默认不退出 |
| files | 最大生成文件数，默认为10 |
| dirSize | 每个文件夹最多允许生成多少个文件，默认为32 |
| baseDir | SliveTest运行后默认存放的文件根目录，默认为“/test/slive” |
| resFile | 结果文件名，默认为“part-0000” |
| replication | 备份数，可设置最小，最大备份数，默认为3 |
| blockSize | 设置文件的大小上限 |
| readSize | 读入大小可设置为最小值，最大值形式，例如“-readSize 100,1000”，默认无限制（min=max=MAX\_LONG=read entire file） |
| writeSize | 写入大小，最小，最大形式，默认等于blockSize（min=max=blocksize） |
| sleep | 在不同次操作之间随机的插入sleep，这个参数用于定义sleep的时间范围，设置同样是最小，最大，单位是毫秒，默认为0） |
| appendSize | 追加写大小，默认等于blockSize（min=max=blocksize） |
| seed | 随机数种子 |
| cleanup | 执行完所有操作并报告之后，清理目录 |
| queue | 指定队列名，默认为“default” |
| packetSize | 指定写入的包大小 |

使用例子：

# 生成10个Maper，每个maper进行100次RPC操作

yarn jar hadoop-mapreduce-client-jobclient-2.6.0-cdh5.13.1-tests.jar SliveTest -maps 10 -ops 100

生成的报告如下，SliverTest会为每一种操作生成一种报告，记录具体的数据：

Basic report for operation type AppendOp

-------------

Measurement "bytes\_written" = 2751463424

Measurement "failures" = 16

Measurement "files\_not\_found" = 83

Measurement "milliseconds\_taken" = 52737

Measurement "op\_count" = 140

Measurement "successes" = 41

Rate for measurement "bytes\_written" = 49.756 MB/sec

Rate for measurement "op\_count" = 2.655 operations/sec

Rate for measurement "successes" = 0.777 successes/sec

### Terasort

这个工具可以用来生成随机数，主要还是测试MapReduce性能的，是SortBenchmark的标准程序。