Hadoop的block Size和split Size究竟是什么关系？

原创 2016-04-04 poon BI实验室

学习hadoop map reduce过程的时候，第一步就是split。我们知道，hdfs中的数据是按block来存储的。问题来了，那么split和block之间是什么关系呢？我google到了stackoverflow上的这篇文章，我觉得这个帖子应该把关系说清楚了，翻译出来，欢迎大家批评指正！以下：

问题

hadoop的split size 和 block size 是什么关系？ 是否 split size 应该 n倍于 block size？

概念

在 hdfs 架构中，存在 blocks 的概念。 通常来说，hdfs中的一个block 是 64MB 。当我们把一个大文件导入hdfs中的时候，文件会按 64MB 每个block来分割（如果按默认配置）。
如果你有1GB的文件要存入HDFS中， 1GB/64MB = 1024MB / 64MB = 16 个 blocks 会被分割到不同的datanode上。

目的

数据分割(data splitting)策略是基于文件偏移进行的。文件分割的目的是 有利于数据并行处理， 以及 便于数据容灾恢复。

区别

split 是逻辑意义上的split。 通常在 M/R 程序或者其他数据处理技术上用到。根据你处理的数据量的情况， split size是允许用户自定义的。

split size 定义好了之后，可以控制 M/R 中 Mapper 的数量。如果M/R中没有定义 split size ， 就用 默认的HDFS配置作为 input split。

举例

你有个100MB的文件，block size 是 64MB， 那么就会被split成 2 块。这时如果你你没有指定 input split ， 你的M/R程序就会按2个input split 来处理， 并分配两个 mapper来完成这个job。

但如果你把 split size 指定为 100MB，那么M/R程序就会把数据处理成一个 split，这时只用分配一个mapper 就可以了。

但如果你把split size 指定为 25MB，M/R就会将数据分成4个split，分配4个mapper 来处理这个job。

总结

1. block是物理上的数据分割，而split是逻辑上的分割。
2. 如果没有特别指定，split size 就等于 HDFS 的 block size 。
3. 用户可以在M/R 程序中自定义split size。

4. 一个split 可以包含多个blocks，也可以把一个block应用多个split操作。

5. 有多少个split，就有多少个mapper。

原文:

```
Q: What is relationship be-
tween split size and block size in Hadoop? As I read in this, split siz
e must be n-times of block size (n is an inte-
ger and n > 0), is this correct? Is there any must in relationship be-
tween split size and block size? **


A:
In HDFS architecture there is a concept of blocks. A typi-
cal block size used by HDFS is 64 MB. When we place a large file into H
DFS it chopped up into 64 MB chunks(based on default configura-
tion of blocks), Sup-
pose you have a file of 1GB and you want to place that file in HDFS,the
n there will be 1GB/64MB = 16 split/blocks and these block will be dis-
tribute across the DataNodes. These blocks/chunk will reside on a dif-
ferent DataNode based on your cluster configuration.

Data splitting happens based on file offsets.The goal of split-
ting of file and store it into different blocks is parallel process-
ing and fail over of data.

Difference between block size and split size.

Split is logical split of your data, basically used during data proces-
sing using Map/Reduce program or other processing tech-
niques. Split size is user de-
fined and you can choose your split size based on your data(How much da
ta you are processing).

Split is basically used to control number of Mapper in Map/Reduce pro-
gram. If you have not defined any input split size in Map/Reduce pro-
gram then default HDFS block split will be considered as input split.

Example:

Suppose you have a file of 100MB and HDFS default block configura-
tion is 64MB then it will chopped in 2 split and occu-
```

py 2 blocks. Now you have a Map/Reduce pro-
gram to process this data but you have not specified any in-
put split then based on the number of blocks(2 block) in-
put split will be considered for the Map/Reduce processing and 2 map-
per will get assigned for this job.

But suppose,you have speci-
fied the split size(say 100MB) in your Map/Reduce pro-
gram then both blocks(2 block) will be considered as a sin-
gle split for the Map/Reduce processing and 1 Mapper will get as-
signed for this job.

Suppose,you have specified the split size(say 25MB) in your Map/Re-
duce program then there will be 4 input split for the Map/Reduce pro-
gram and 4 Mapper will get assigned for the job.

Conclusion:

1. Split is a logical division of the in-
put data while block is a physical division of data.
2. HDFS default block size is default split size if in-
put split is not specified.
3. Split is user defined and user can con-
trol split size in his Map/Reduce program.
4. One split can be mapping to multiple blocks and there can be multi-
ple split of one block.
5. The number of map tasks (Mapper) are equal to the number of splits.

ref: http://stackoverflow.com/ques-
tions/30549261/split-size-vs-block-size-in-hadoop

补充

hdfs中如何设置参数

split size 配置：

> mapred.max.split.size

block size 配置：

> dfs.block.size

注意，请不要轻易更改 block size,因为这个是影响整个hdfs的。

split size 计算规则：

针对：

mapreduce.input.fileinputformat.split.maxsize

mapreduce.input.fileinputformat.split.minsize

FileInputFormat.computeSplitSize()

按如下规则取 split size：

Math.max(minSize, Math.min(maxSize, blockSize))

性能

1. 一般来说，block size 和 split size 设置成一致，性能较好。

2. FileInputFormat generates splits in such a way that each split is all or part of a single file. 所以 hadoop处理大文件比处理小文件来得效率高得多。

Read more