

2

answers

### Q: Hadoop 2.4 maven build OutOfMemoryError: Java heap space

I was trying to build the complete source code of **Hadoop** on my mac. When I try building I constantly run into the error: "java.lang.OutOfMemoryError: Java heap space". I tried to set the maven opts ...

maven   hadoop   heap-memory

asked nov 26 by [Sachin Ananth Navale](#)

2

answers

### Q: Replicate DirectByteBuffers OutOfMemoryError

I am planning to use a DirectByteBuffers to interface a third party lib written in C++. I got worried from the **hadoop** docs saying: DirectByteBuffers are garbage collected by using a phantom ... to a DirectByteBuffer, it's easy to **OutOfMemoryError** yourself using DirectByteBuffers. First I am not sure if all implementations of DirectByteBuffers are the same since the OpenJDK 7 does not seem ...

java   jni

asked jun 7 '12 by [Manuel Schmidt](#)

1

answer

### Q: OutOfMemoryError when reading a local file via DistributedCache

/standalone mode). However, when I execute it in the **Hadoop** cluster (fully distributed mode), I get an "OutOfMemoryError: Java heap space", with the same 40 MB file. I don't understand why this happens, as the file isn't that large. This is the code: Any help would be appreciated, thanks in advance. ... understand why the program works locally but it doesn't distributedly. Thanks for all the answers. I'm using **Hadoop** 1.0.3. The cluster is composed of three machines, all of them running Ubuntu Linux 12.04 ...

hadoop   mapreduce   out-of-memory

asked nov 19 '12 by [iporto](#)

0

answers

### Q: how to determine the size of "mapred.child.java.opts" and HADOOP\_CLIENT\_OPTS in mahout canopy

Is size is about 50M. mapred-site.xml: **hadoop**-env.sh: When running mahout canopy, it always throws **OutOfMemoryError**: Question is, how to determine the size of "mapred.child.java.opts" and HADOOP\_CLIENT\_OPTS? ...

mahout

asked dec 6 by [user3156087](#)

3

answers

### Q: How to use "typedbytes" or "rawbytes" in Hadoop Streaming?

the binary file to HDFS with "**hadoop** fs -copyFromLocal". When I try to use it as input to a map-reduce job, it fails with an **OutOfMemoryError** on the line where it tries to make a byte array ... I have a problem that would be solved by **Hadoop** Streaming in "typedbytes" or "rawbytes" mode, which allow one to analyze binary data in a language other than Java. (Without this, Streaming ...

hadoop   binary   streaming

asked mar 2 '13 by [Jim Pivarski](#)

1

answer

### Q: Mahout - Exception: Java Heap space

I'm trying to convert some texts to mahout sequence files using: But all I get is a **OutOfMemoryError**, as here: I am using Mahout 0.9, **Hadoop** 1.2.1 and OpenJDK Java7u25 defining MAHOUT\_HEAPSIZE ...

hadoop   mahout

asked apr 7 by [user3422072](#)

0

answers

### Q: HBase: How to handle large query results

I have a table with almost 3 million records in it. I want to be allow a user to query large sections of it, or even the whole table. I'm new to hbase/**hadoop** (and fairly inexperienced in db ... function. I could just pass the Scanner, but I want to close it. My problem is when the ArrayList gets filled to about 2 million records, I get an **OutOfMemoryError**: I'm pretty sure my approach ...

java   hadoop   out-of-memory   hbase

asked jul 21 by [Bryany](#)

### A: Amazon EMR Hbase, Thrift server dies out

Did you check the thrift logs for any exception? If you don't find any exception on the logs, then that means that thrift is being killed, and that usually happens on **OutOfMemoryError**. So basically ... you just need to add more memory for the thrift server or open less scanners at the same time (scanners consumes huge amount of resources). As about the JAVA\_HOME error, you have to set it for the user **hadoop** ...

answered jan 7 by [Oussama](#)

### A: Using Amazon MapReduce/Hadoop for Image Processing

in multiple folders for each core and run the bash scripts over it. This should utilize your machine quite well, and better than **Hadoop** ever could. [1] <http://codingwiththomas.blogspot.com/2011/07/dealing-with-outofmemoryerror-in-hadoop.html> ... There are several problems with your task. **Hadoop** does not natively process images as you've seen. But you can export all the file names and paths as a textfile and call some Map function on it. So ...

answered oct 19 '11 by [Thomas Jungblut](#)

### A: Hadoop Pipes: how to pass large data records to map/reduce tasks

**Hadoop** is not designed for records about 100MB in size. You will get **OutOfMemoryError** and uneven splits because some records are 1MB and some are 100MB. By Ahmdal's Law your parallelism will suffer ... greatly, reducing throughput. I see two options. You can use **Hadoop** streaming to map your large files into your C++ executable as-is. Since this will send your data via stdin it will naturally ...

answered oct 26 '10 by [Spike Gronim](#)

### A: sublist for partitioning data set

of records, because there are other problems you might hit: If your program fails (e.g. **OutOfMemoryError**), you probably don't want to have to start over from the beginning. You might want to throw >1 ... . A really common tool for this kind of work is **Hadoop**. You'd copy the data into HDFS, run a map-reduce job (or more than one job) on the data and then copy the data out of HDFS when you're done. A simpler ...

answered jul 8 '12 by [Harold L](#)

4

answers

**Q: Realistic Social Network Model with tens of millions Users. Which technologies should I be u...**

sets, I get an **OutOfMemoryError** due to no more memory in the heap. I now realize that this way of doing it will be impossible if I want to create, say, 30 Area's with 1 millions users per area, and 200 ... . 2. Should I be using a different language all together. I know that technologies such as Lucene, **Hadoop**, etc. were created with Java, and are used for large amounts of data...But I have never used ...

java

database

arrays

memory

data-structures

asked jan 12 '11 by [Eddie](#)