

原

Spark---JVM调优之调节executor离堆内存与连接等待时长

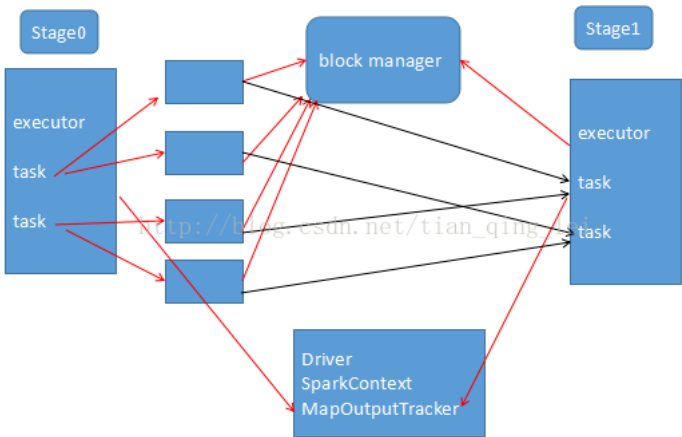
2017年09月15日 23:43:51 缘定三生 阅读数：872 标签： JVM调优 executor 连接等待时长 更多

版权声明：本文为博主原创文章，未经博主允许不得转载。 https://blog.csdn.net/tian_qing_lei/article/details/77996853

1、executor堆外内存

有时候，如果你的spark作业处理的数据量特别特别大，几亿数据量；然后spark作业一运行，时不时的报错，shuffle file cannot find, executor lost, out of memory (内存可能是说executor的堆外内存不太够用，导致executor在运行的过程中，可能会内存溢出；然后可能导致后续的stage的task在运行的时候，可能要从一些executor中去拉取shuffle文件，但是executor可能已经挂掉了，关联的block manager也没有了；所以可能会报shuffle output file not found; resubmitting task; executor上述情况下，就可以去考虑调节一下executor的堆外内存。也许就可以避免报错；此外，有时，堆外内存调节的比较大的时候，对于性能会带来一定的提升。

2、executor的堆外内存不太够用，导致Spark作业停止原理图



分析：1）Stage的executor跑着跑着，突然内存不足了，堆外内存不足了，可能会OOM，挂掉。block manager也没有了，数据也丢失掉了。

2）如果此时，stage0的executor挂了，block manager也没有了；此时，stage1的executor的task，虽然通过Driver的MapOutputTracker获取到了自己数据的地址；但是实际block manager获取数据的时候，是获取不到的

此时，就会在spark-submit运行作业（jar），client（standalone client、yarn client），在本机就会打印出log shuffle output file not found。。。

DAGScheduler，resubmitting task，一直会挂掉。反复挂掉几次，反复报错几次，整个spark作业就崩溃了

4、参数设置

--conf spark.yarn.executor.memoryOverhead=2048

<div><div>APACHE SPARK 2.1.0</div><div>Overview Programming Guides API Docs Deploying More</div></div> <div>section.ack.wait.timeout 第 0 条</div>		
spark.yarn.dist.jars	(none)	Comma-separated list of jars to be placed in the working directory of each executor.
spark.executor.instances	2	The number of executors for static allocation. With spark.dynamicAllocation.enabled, the initial set of executors will be at least this large.
spark.yarn.executor.memoryOverhead	executorMemory * 0.10, with minimum of 384	The amount of off-heap memory (in megabytes) to be allocated per executor. This is memory that accounts for things like VM overheads, interned strings, other native overheads, etc. This tends to grow with the executor size (typically 6-10%).
spark.yarn.driver.memoryOverhead	driverMemory * 0.10, with minimum of 384	The amount of off-heap memory (in megabytes) to be allocated per driver in cluster mode. This is memory that accounts for things like VM overheads, interned strings, other native overheads, etc. This tends to grow with the container size (typically 6-10%).
spark.yarn.am.memoryOverhead	AM memory * 0.10, with minimum of 384	Same as spark.yarn.driver.memoryOverhead, but for the YARN Application Master in client mode.

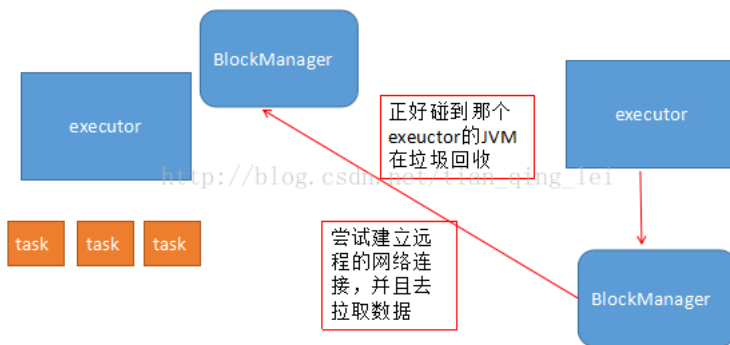
定要在spark-submit脚本中去设置。

spark.yarn.executor.memoryOverhead (看名字, 顾名思义, 针对的是基于yarn的提交模式)

默认情况下, 这个堆外内存上限大概是300多M; 后来我们通常项目中, 真正处理大数据的时候, 这里都会出现问题, 导致spark作业反复崩溃, 无法运行; 此时就会去调节至少1G (1024M), 甚至说2G、4G

通常这个参数调节上去以后, 就会避免掉某些JVM OOM的异常问题, 同时呢, 会让整体spark作业的性能, 得到较大的提升。

5、连接等待超时导致作业失败原理图



分析: 1) executor, 优先从自己本地关联的BlockManager中获取某份数据

如果本地block manager没有的话, 那么会通过TransferService, 去远程连接其他节点上executor的block manager去获取

2) JVM调优: 垃圾回收

处于垃圾回收过程中, 所有的工作线程全部停止; 相当于只要一旦进行垃圾回收, spark / executor停止工作, 无法提供响应

3) task创建的对象特别大, 特别多, 频繁的让JVM堆内存满溢, 进行垃圾回收。

导致现象:

此时呢, 就会没有响应, 无法建立网络连接; 会卡住; ok, spark默认的网络连接的超时时长, 是60s; 如果卡住60s都无法建立连接的话, 那么就宣告失败了。

碰到一种情况, 偶尔, 偶尔, 偶尔!!! 没有规律!!! 某某file。一串file id. uuid (dsfsfd-2342vs--sdf--sdfsd)。not found. file lost。

这种情况下, 很有可能是有那份数据的executor在jvm gc。所以拉取数据的时候, 建立不了连接。然后超过默认60s以后, 直接宣告失败。

报错几次, 几次都拉取不到数据的话, 可能会导致spark作业的崩溃。也可能会导致DAGScheduler, 反复提交几次stage。TaskScheduler, 反复提交几次task。大大延长我们的时间。

6、参数设置

可以考虑调节连接的超时时长。Spark 2.1.0版本

--conf spark.network.timeout=300

spark-submit脚本, 切记, 不是在new SparkConf().set()这种方式来设置的。

以前版本采用: spark.core.connection.ack.wait.timeout (spark core, connection, 连接, ack, wait timeout, 建立不上连接的时候, 超时等待时长)

调节这个值比较大以后, 通常来说, 可以避免部分的偶尔出现的某某文件拉取失败, 某某文件lost掉了。。。

<div><div>APACHE</div><div>spark</div><div>2.1.0</div></div> <div>OverviewProgramming Guides▼API Docs▼Deploying▼More▼</div>		
executors or external systems. This is useful, for example, when running containers with bridged networking. For this to properly work, the different ports used by the driver (RPC, block manager and UI) need to be forwarded from the container's host.		
spark.driver.host	(local hostname)	Hostname or IP address for the driver. This is used for communicating with the executors and the standalone Master.
spark.driver.port	(random)	Port for the driver to listen on. This is used for communicating with the executors and the standalone Master.
spark.network.timeout	120s	Default timeout for all network interactions. This config will be used in place of spark.core.connection.ack.wait.timeout, spark.storage.blockManagerSlaveTimeoutMs, spark.shuffle.io.connectionTimeout, spark.rpc.askTimeout, spark.rpc.lookupTimeout if they are not configured.
spark.port.maxRetries	16	Maximum number of retries when binding to a port before giving up. When a port is given a specific value (non 0), each subsequent retry will increment the port used in the previous attempt by 1 before retrying. This essentially allows it to try a range of ports from the start port specified to port + maxRetries.

7、为什么在这里讲这两个参数呢？

因为比较实用，在真正处理大数据（不是几千万数据量、几百万数据量），几亿，几十亿，几百亿的时候。很容易碰到executor堆外内存，以及gc引起的连接超时的问题。filexecutor lost , task lost。
调节上面两个参数，还是很有帮助的。

一个退役操盘手肺腑之言，写给无数正在亏钱的散户

唯木家金融 · 熾焱

想对作者说点什么？

我来说一句

Spark中executor-memory参数详解

Spark中executor-memory参数详解

7957

来自：wisgood的专栏

Spark中各个角色的JVM参数设置

1)Driver的JVM参数：-Xmx，-Xms，如果是yarn-client模式，则默认读取spark-env文件中的SPARK_DRIVER_ME...

6377

来自：wuxb_2000的专栏

Spark异常处理与调优（更新中~）

资源调优 http://blog.csdn.net/u011239443/article/details/52127689 内存 Memory Tuning，Java对象会占用原始数...

9470

来自：Soul Joy Hub

转型人工智能可以吗？一个小测试就让你知道怎么学

机器学习|深度学习|图像处理|自然语言处理|无人驾驶，这些技术都会吗？看看真正的人工智能师都会那些关键技术？年薪比你高多少！

Spark超时问题一例(Connection has been quiet)

这是用Spark Mllib ALS算法做离线推荐时遇到的问题。我们对历史日志作统计和评分，然后用ALS来训练模型，最...

3063

来自：wesleyflagon的专栏

Spark排错与优化

一. 运维 1. Master挂掉,standby重启也失效 Master默认使用512M内存，当集群中运行的任务特别多时，就会挂掉，...

5.8万

来自：偷闲小苑

spark常见问题处理

1、spark thriftserver报以下错误，其他诸如hive/sparksql等方式均正常 ERROR ActorSystemImpl: Uncaught fatal er...

2206

来自：赵英超的博客

Spark JVM调优之调节executor堆外内存与连接等待时长

executor堆外内存的调优 有时候，如果你的spark作业处理的数据量特别特别大，大约在几亿的数据量，然后spark...

133

来自：mn_kw的博客

spark调优和JVM运行流程

算子继续：MapPartitions:每次处理一个分区数据，常用，经常代替map进行使用，用于调优。调优的作用是减少ta...

483

来自：Addroc_Sue的博客

一个退役操盘手肺腑之言，写给无数正在亏钱的散户

Spark作业执行之JVM参数设置

spark作业提交后，在driver上运行init()方法时报错：java.lang.OutOfMemoryError: GC overhead limit exceeded报...

来自： [zjwcsdn123的博客](#)

Spark_Spark调优-设置executor 数量 和task 并行数

原文地址：https://blog.csdn.net/wanguohe/article/details/80645978—.指定spark executor 数量的公式executor 数...

来自： [迎难而上](#)

博主推荐



chbxw

关注

545篇文章



Evankaka

关注

285篇文章



段智华

关注

827篇文章

spark堆外内存的设置

MemoryOverhead是JVM进程中除Java堆以外占用的空间大小，包括方法区（永久代）、Java虚拟机栈、本地方法...

来自： [西贝没有横](#)

spark学习-42-Spark的driver理解和executor理解

1.看了很多网上的图，大多是dirver和executor之间的图，都不涉及物理机器 如下图，本人觉得这些始终有些抽象 ...

来自： [九师兄-梁川川](#)

Spark性能优化总结

近期优化了一个spark流量统计的程序，此程序跑5分钟小数据量日志不到5分钟，但相同的程序跑一天大数据量日志...

来自： [天道酬勤](#)

青羊大道股王8年追涨停铁律“1272”曝光，震惊众人

第六 · 燦燦

Spark On YARN内存分配

本文主要了解Spark On YARN部署模式下的内存分配情况，因为没有深入研究Spark的源代码，所以只能根据日志...

来自： [wisgood的专栏](#)

spark JVM调优之调节executor堆外内存与连接等待时长

executor堆外内存 有时候，如果你的spark作业处理的数据量特别特别大，几亿数据量；然后spark作业一运行，时...

来自： [涛涛的专栏](#)

Spark 堆外内存

1.堆外内存有哪些前面提到spark中的堆内存溢出，除了堆内存，还有堆外内存。该部分内存主要用于程序的共享库...

来自： [bitcarmanlee的博客](#)

关于spark连接elasticSearch集群失败问题Connection error (check network and/or proxy settings)- al...

在用spark 向elasticSearch集群写入数据时爆出 Connection error (check network and/or proxy settings)- all nodes f...

来自： [a1031891160的博客](#)

记一次spark任务调优过程

最近跑了一个spark任务(逻辑回归训练建模的),数据量是3000多万左右,在跑的过程中出现了各种错误,跑了8h左右,...

来自： [Aochen的博客](#)

青羊大道女股民巨亏后一举成名 买卖技巧惊呆股市高手

潮望 · 燦燦

Spark源码走读8——NetWork

网络管理，由于分布式集群，那么无论master还是worker都离不开网络通讯。Network包位于核心源码org.apache.s...

来自： [峰哥的专栏](#)

spark-一些参数优化

Spark程序优化所需要关注的几个关键点——最主要的是数据序列化和内存优化问题1：reduce task数目不合适解决...

来自： [不管大小写的博客](#)

spark常见错误及调优

一. 运维 1. Master挂掉,standby重启也失效 Master默认使用512M内存，当集群中运行的任务特别多时，就会挂掉， ...

来自： [赵英超的博客](#)

Spark 学习: spark 原理简述与 shuffle 过程介绍

Spark学习: 简述总结 Spark 是使用 scala 实现的基于内存计算的大数据开源集群计算环境.提供了 java,scala, pytho...

来自： [databatman的工厂](#)

Spark程序运行常见错误解决方法以及优化

一.org.apache.spark.shuffle.FetchFailedException 1.问题描述 这种问题一般发生在有大量shuffle操作的时候,task不...

来自： [sdujava2011](#)

青羊大道退休女教师一年18中签，神奇操作意外曝光

合美投资 · 熯熯

- spark排错与优化

一. 运维 1. Master挂掉,standby重启也失效 Master默认使用512M内存，当集群中运行的任务特别多时，就会挂掉， ...

来自： [u013013024的博客](#)
- spark Lost **executor** on YARN

执行脚本出现： 15/07/30 10:18:13 ERROR cluster.YarnScheduler: Lost executor 8 on myhost1.com: remote R...

来自： [赵英超的博客](#)
- 56.JVM**调优之调节executor堆外内存与连接等待时长**

本文为《Spark大型电商项目实战》系列文章之一，主要介绍在实际大型Spark项目中通过调节 executor 堆外内存...

来自： [k_ly的博客](#)
- 调节executor堆外内存与连接等待时长**

注意下面需要配置两个参数：spark-submit脚本里面，去用--conf的方式，去添加配置；一定要注意！！！切记，不...

来自： [城堡之家](#)
- JVM**堆内存**参数优化，让性能飞起来

JVM堆内存参数优化，让性能飞起来 堆内存是Java进程的重要组成部分，几乎所有与应用相关的内存空间都和堆有...

来自： [烂笔头-D的专栏](#)

青羊大道45岁卖菜大妈，玩手机赚钱3个月后惊人存款被发现！

天傲网络投资 · 熯熯

- 深入JVM 原理（五）Java**堆内存**调整参数（**调优**关键）

目录 深入JVM 原理（一）Java内存模型：http://blog.csdn.net/qq_34707744/article/details/79278169 深入JVM原理...

来自： [海蓝星上](#)
- JVM中的堆以及**调优**参数

我们知道对象存放到JVM堆中，更加具体是存放到堆中新生区的伊甸区。而GC也是集中在堆中，一个JVM中的堆内...

来自： [ycd500756的博客](#)

缘定三石

关注

原创184

粉丝32

喜欢23

评论2

等级： 博客 5

访问：7万+

积分：2598

排名：2万+



最新文章

10以太坊Token详解

09以太坊开发工具、框架和库

08智能合约

07交易

06钱包

个人分类

Java实战	28篇
Mysql实战	7篇
Hadoop实战	46篇
Hive实战	26篇
Flume实战	6篇
展开	

归档	
2018年8月	10篇
2018年7月	3篇
2018年6月	2篇
2018年5月	1篇
2018年2月	20篇
展开	

热门文章	
大数据分析整体技术流程及架构	阅读量：5862
新一代大数据技术架构	阅读量：4551
TensorFlow框架--Keras使用	阅读量：3209
解决通过ip能够访问网络，通过域名无法访问的问题	阅读量：2522
流量分析常见指标	阅读量：1858

最新评论	
监督学习	qq_42760276：此段引自网易云吴恩达机器学习视频，第一章课时3！直接看视频就可以
IPFS介绍及安装	u014454120：我配置OCRS时报错，是什么原因？？？ PS C:\Users\py\.ipfs> ipf...

国网天津市电力公司营销金和秀、金证精英量管及可视化平台



数据可视化

联系我们



扫码联系客服



下载CSDN APP



QQ客服



kefu@csdn.net



客服论坛



400-660-0108


工作时间 8:00-22:00

关于我们

招聘

广告服务

网站地图

 百度提供站内搜索 京ICP证09002463号
©2018 CSDN版权所有

[网络110报警服务](#) [经营性网站备案信息](#)
[北京互联网违法和不良信息举报中心](#)
[中国互联网举报中心](#)