

组网及说明

正常DE组网

问题描述

客户DataEngine大数据管理平台下发任务卡住，执行命令yarn application -list查看yarn任务发现任务进度条倒退的情况，比如增长到70%然后下降到5%再继续轮回。现场反馈一个任务基本需要两个小时左右，大概过上一天一夜所有任务都会被卡住。

如下图所示任务，前面查询已经到23.94%，再查询进度下降到23.75%。

```

application_1541494108616_119418 MAP_DW_TB_4G_PROTOCOL_201920 MAPREDUCE root default
RUNNING UNDEFINED 78.7% http://czbdhbase6.hde.h3c.com:46040
application_1541494108616_119419 MAP_PS_4G_TB_PROTOCOL_201919 MAPREDUCE root default
RUNNING UNDEFINED 42.48% http://czbdhbase5.hde.h3c.com:58834
application_1541494108616_119420 DW_TB_4G_PROTOCOL_201920 MAPREDUCE root default
RUNNING UNDEFINED 23.94% http://czbdhbase5.hde.h3c.com:34393
application_1541494108616_119318 MAP_PS_4G_TB_PROTOCOL_201920 MAPREDUCE root default
RUNNING UNDEFINED 49.81% http://czbdhbase4.hde.h3c.com:60145
application_1541494108616_119319 DW_TB_4G_PROTOCOL_201921 MAPREDUCE root default
RUNNING UNDEFINED 17% http://czbdhbase5.hde.h3c.com:49246
application_1541494108616_119334 PS_4G_TB_PROTOCOL_201920 MAPREDUCE root default
RUNNING UNDEFINED 48.88% http://czbdhbase3.hde.h3c.com:34284
application_1541494108616_119335 MAP_DW_TB_4G_PROTOCOL_201921 MAPREDUCE root default
RUNNING UNDEFINED 45.12% http://czbdhbase6.hde.h3c.com:39471
root@czbdhbase1 ~# yarn application -list
/05/20 10:20:39 INFO impl.TimelineClientImpl: Timeline service address: http://czbdhbase2.hde.h3c.com:8188/ws/v1/timeline/
/05/20 10:20:40 INFO client.ConfiguredRMFailoverProxyProvider: Failing over to rm2
total number of applications (application-types: [] and states: [SUBMITTED, ACCEPTED, RUNNING]):7
Application-Id Application-Name Application-Type User Queue State
Final-State Progress Tracking-URL
application_1541494108616_119418 MAP_DW_TB_4G_PROTOCOL_201920 MAPREDUCE root default
RUNNING UNDEFINED 84.23% http://czbdhbase6.hde.h3c.com:46040
application_1541494108616_119419 MAP_PS_4G_TB_PROTOCOL_201919 MAPREDUCE root default
RUNNING UNDEFINED 51.76% http://czbdhbase5.hde.h3c.com:58834
application_1541494108616_119420 DW_TB_4G_PROTOCOL_201920 MAPREDUCE root default
RUNNING UNDEFINED 23.75% http://czbdhbase5.hde.h3c.com:34393
application_1541494108616_119318 MAP_PS_4G_TB_PROTOCOL_201920 MAPREDUCE root default
RUNNING UNDEFINED 49.81% http://czbdhbase4.hde.h3c.com:60145
application_1541494108616_119319 DW_TB_4G_PROTOCOL_201921 MAPREDUCE root default
RUNNING UNDEFINED 17% http://czbdhbase5.hde.h3c.com:49246

```

过程分析

1、了解现场业务情况：

了解现场提交处理的任务是bulkload把文件写到hbase。查看DataEngine管理平台各项服务状态都是正常的，查看磁盘利用率一直维持在80%左右。

点击快速链接进入ResourceManager UI管理界面查看资源任务情况。



2、查看资源任务运行情况

在ResourceManager UI页面查看VCores Total是146，而VCores Used已经达到了180，VCores Reserved为0表示虚拟内核资源不足。对比验证其他普通业务（vcores和memory都在总资源内）运行正常。

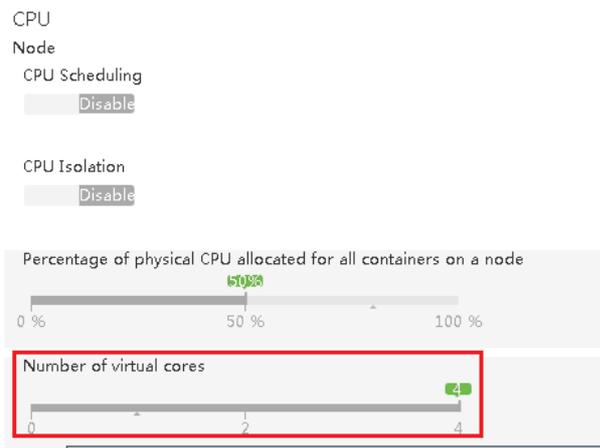
Cluster Metrics										
Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved
119418	0	6	119412	180	810 GB	1.17 TB	0 B	180	146	0

解决方法

问题定位在业务所需资源超过Yarn总资源量。此时可以选择将集群资源增大或者控制任务提交数量。

1、增大集群virtual core操作：

在Yarn“配置”下面的CPU部分，调整“Number of virtual cores”参数，然后保存、点击重启相应服务。



2、需要在提交任务的时候控制下map reduce的个数，减少map reduce负载。（其他计算组件类似。）

说明：对于Yarn任务卡住相关，第一先去UI页面查看各项资源是否有不足的情况，查看不出原因可以去看下Yarn日志。

附件下载：[\[经验案例\]virtual cores实际使用超过设置参数导致Yarn任务卡住_厉梦如.doc](#)