

# E-MapReduce

常见问题

# 常见问题

## EMR 常见问题

### Q：作业和执行计划的区别

A：在Aliyun E-MapReduce中，要运行作业，需要有分成两个步骤，分别是：

- 创建作业

在E-MapReduce产品中，说创建一个“作业”，实际上是创建一个“作业运行配置”，它并不能被直接运行。既如果在E-MapReduce中创建了一个“作业”，实际上只是创建了一个“作业如何运行的配置”，这份配置中包括该作业要运行的jar包，数据的输入输出地址，以及一些运行参数。这样的一份配置创建好后，给它命一个名，既定义了一个“作业”。当你需要调试运行作业的时候就需要执行计划了。

- 创建执行计划

执行计划，是将作业与集群关联起来的一个纽带。通过它，我们可以把多个作业组合成一个作业序列，通过它我们可以为作业准备一个运行集群（或者自动创建出一个临时集群或者关联一个已存在的集群），通过它我们可以为这个作业序列设置周期执行计划，并在完成任务后自动释放集群。我们也可以在他的执行记录列表上查看每一次执行的执行成功情况与日志。

### Q：如何查看作业日志

A：在E-MapReduce系统里，系统已经将作业运行日志按照jobid的规划上传到OSS中（路径由用户在创建集群时设置），用户可以直接在网页上点击查看作业日志。如果用户是登录到master机器进行作业提交和脚本运行等，则日志根据用户自己的脚本而定，用户可以自行规划。

### Q:如何登录Core节点

A：按照如下步骤：

1. 首先在Master节点上切换到Hadoop账号：

```
su hadoop
```

2. 然后即可免密码ssh登录到对应的core节点：

```
ssh emr-worker-1
```

3. 通过sudo 可以获得root权限

```
sudo vi /etc/hosts
```

## Q：直接在OSS上查看日志

A：用户也可以直接从OSS上直接查找所有的日志文件，并下载。但是因为OSS不能直接查看，使用起来会比较麻烦一些。如果用户打开了日志，并且指定了一个OSS的日志位置，那么作业的日志要如何找到呢？例如对下面这个保存位置OSS://mybucket/emr/spark

1. 首先来到执行计划的页面，找到对应的执行计划，点击“运行记录”进入运行记录页面。
2. 在运行记录页面找到具体的哪一条执行记录，比如最后的一条执行记录。然后点击它对应的“执行集群”查看这个执行集群的ID。
3. 然后再OSS://mybucket/emr/spark目录下寻找OSS://mybucket/emr/spark/集群ID 这个目录
4. 在OSS://mybucket/emr/spark/集群ID/jobs 目录下会按照作业的执行ID存放多个目录，每一个目录下存放了这个作业的运行日志文件。

## Q：集群、执行计划以及运行作业的计时策略

A：三种计时策略如下：

### 集群的计时策略

在集群列表里可以看到每个集群的运行时间，该运行时间的计算策略为  $\text{运行时间} = \text{集群释放时刻} - \text{集群开始构建时刻}$ 。即集群一旦开始构建就开始计时，直到集群的生命周期结束。

执行计划的计时策略：在执行计划的运行记录列表，可以看到每次执行记录运行的时间，该时间的计时策略总结为两种情况：

1. 如果执行计划是按需执行的，每次执行记录的运行过程涉及到创建集群、提交作业运行、释放集群。所以按需执行计划的运行时间计算策略为， $\text{运行时间} = \text{构建集群的时间} + \text{执行计划包含所有作业全部运行结束的总耗时} + \text{集群释放的时间}$ 。
2. 如果执行计划是关联已有集群运行的，整个运行周期不涉及到创建集群和释放集群，所以其运行时间 = 执行计划包含所有作业全部运行结束的总耗时。

作业的计时策略：

这里的作业指的是被挂载到执行计划里面的作业。在每条执行计划运行记录右侧的查看作业列表点击进入可以看到该作业。这里每个作业的运行时间的计算策略为，运行时间 = 作业运行结束的实际时间 - 作业开始运行的实际时间。作业运行开始（结束）的实际时间指的是作业被Spark或hadoop集群实际开始调度运行或运行结束的时间点。

## Q：第一次使用执行计划时没有安全组可选

A：因为一些安全的原因，EMR目前的安全组并不能直接选择用户的所有已有安全组来使用，所以如果你还没有在EMR中创建过安全组的话，在执行计划上将无法选择到可用的安全组。我们推荐您先手动创建一个按需集群来进行作业的测试，手动创建集群的时候可以创建一个新的EMR安全组，等到测试都通过了以后，再设置您的执行计划来周期调度。这个时候之前创建的安全组也会出现在这里可供选择。

## Q：读写ODPS时，抛出java.lang.RuntimeException.Parse responded failed: '<!DOCTYPE html>...'

A：检查odps tunnel endpoint是否正确，如果写错会出现这个错误。

## Q：多个ConsumerID消费同一个Topic时出现TPS不一致问题

A：有可能这个Topic在公测或其他环境创建过，导致某些Consumer组消费数据不一致。请在工单系统中将对应的Topic和ConsumerID提交到ONS处理。

## Q：E-MapReduce中能否查看作业的Worker上日志

A：可以。前置条件：是创建集群时发开“保存日志”选项。查看日志位置：执行计划列表->点击“运行记录”->执行记录->点击“查看作业列表”->作业列表->点击“查看作业Worker实例”。

## Q：Hive 创建外部表，没有数据

A：例如：

```
CREATE EXTERNAL TABLE storage_log(content STRING) PARTITIONED BY (ds STRING)

ROW FORMAT DELIMITED
FIELDS TERMINATED BY '\t'
STORED AS TEXTFILE
LOCATION 'oss://log-124531712/biz-logs/airtake/pro/storage';

hive> select * from storage_log;

OK

Time taken: 0.3 seconds

创建完外部表后没有数据
```

实际上Hive并不会自动关联指定目录的partitions目录，您需要手动操作，例如：

```
alter table storage_log add partition(ds=123);
OK
Time taken: 0.137 seconds
hive> select * from storage_log;
OK
abcd 123
efgh 123
```

## Q：Spark Streaming作业运行一段时间后无故结束

A：首先检查Spark版本是否是1.6之前版本。Spark1.6修复了一个内存泄漏的BUG，这个BUG会导致container内存超用然后被kill掉（当然，这只是可能的原因之一，不能说明Spark1.6不存在任何问题）。此外，检查自己的代码在使用内存上有没有做好优化。

## Q：Spark Streaming作业已经结束，但是E-MapReduce控制台显示作业还处于“运行中”状态

A：检查Spark Streaming作业的运行模式是否是“yarn-client”，若是建议改成“yarn-cluster”模式。E-MapReduce对“yarn-client”模式的Spark Streaming作业的状态监控存在问题，会尽快修复。

## Q：“Exception in thread “main” java.lang.RuntimeException: java.lang.ClassNotFoundException: Class com.aliyun.fs.oss.nat.NativeOssFileSystem not found”

A：在Spark作业中读写OSS数据时，需要将E-MapReduce提供的SDK打进作业Jar包中，具体操作请参考：开发手册->Spark->开发准备。

## Q：“Error: Could not find or load main class”

A：检查作业配置中作业jar包的路径协议头是否是“ossref”，若不是请改为“ossref”。

## Q：集群机器分工使用说明

A：E-MapReduce中包含一个Master节点和多个Slave（或者Worker）节点。其中Master节点不参与数据存储和计算任务，Slave节点用来存储数据和计算任务。例如3台4核8G机型的集群，其中一台机器用来作为Master节点，另外两台用来作为Slave节点，也就是集群的可用计算资源为2台4核8G机器。

## Q：Spark接Flume时出现内存超用问题

A：检查是否是以Push-based方式接收数据，若不是，请尝试改成Push-based方式接收数据。参考文档。

**Q : Caused by: java.io.IOException: Input stream cannot be reset as 5242880 bytes have been written, exceeding the available buffer size of 524288**

A : ( OSS ) 网络连接重试时缓存不足的BUG请, 使用1.1.0版本以上的emr-sdk。

**Q : " Failed to access metastore. This class should not accessed in runtime.org.apache.hadoop.hive.ql.metadata.HiveException: java.lang.RuntimeException: Unable to instantiate org.apache.hadoop.hive.ql.metadata.SessionHiveMetaStoreClient"**

A : 首先, Spark处理Hives数据需要作业的执行模式为yarn-client ( local也行 ), 不能为yarn-cluster, 否则会报上述异常。其次, 作业jar中引入一些第三方的包也有可能导致Spark运行期间报上述异常。

**Q : 如何在MR作业中使用本地共享库**

A : 方法有很多, 这里给出一种方式。修改mapred-site.xml文件, 例如 :

```
<property>
<name>mapred.child.java.opts</name>
<value>-Xmx1024m -Djava.library.path=/usr/local/share/</value>
</property>

<property>
<name>mapreduce.admin.user.env</name>
<value>LD_LIBRARY_PATH=$HADOOP_COMMON_HOME/lib/native:/usr/local/lib</value>
</property>
```

只要加上你所需的库文件即可。

**Q : 如何在MR/Spark作业中指定OSS数据源文件路径**

A : 如下, OSS URL : oss://[accessKeyId:accessKeySecret@]bucket[.endpoint]/object/path

用户在作业中指定输入输出数据源时使用这种URI, 可以类比hdfs://。 用户操作OSS数据时 :

1. ( 建议 ) EMR提供了MetaService服务, 支持免AK访问OSS数据, 直接写 oss://bucket/object/path。
2. ( 不建议 ) 可以将accessKeyId, accessKeySecret以及endpoint配置到Configuration ( Spark作业是SparkConf, MR类作业是Configuration ) 中, 也可以在URI中直接指定 accessKeyId, accessKeySecret以及endpoint。具体请参考开发准备一节。

**Q : Spark SQL抛出 "Exception in thread "main"**

## java.sql.SQLException: No suitable driver found for jdbc:mysql:xxx” 报错

A :

1. 低版本mysql-connector-java有可能出现类似问题，更新到最新版本。
2. 作业参数中使用 “—driver-class-path ossref://bucket/.../mysql-connector-java-[version].jar” 来加载mysql-connector-java包，直接将mysql-connector-java打进作业jar包也会出现上述问题。

## Q : Spark SQL连RDS出现ConnectionException

A : 检查RDS数据库地址是否为内网地址，如不是请在RDS控制台将相应的数据库地址切换成内网地址。

## Q : Spark SQL连RDS出现 “Invalid authorization specification, message from server: ip not in whitelist”

A : 检查RDS的白名单设置，将集群机器的内网地址加到RDS的白名单中。

## Q : Spark程序中使用OSS SDK出现

“java.lang.NoSuchMethodError:org.apache.http.conn.ssl.SSLConnectionSocketFactory.init(Ljavax/net/ssl/SSLContext;Ljavax/net/ssl/HostnameVerifier)”

A : OSS SDK依赖的http-core和http-client包与Spark和Hadoop的运行环境存在版本依赖冲突，不建议在代码中使用OSS SDK，否则需要手动解决依赖冲突问题，比较麻烦。如果有需要对OSS中的文件做一些基础操作例如list等等，可以参照这里的用法进行操作。

## Q : 创建低配置机型集群注意事项

A : 1. 若Master节点选择2核4G机型，则Master节点内存非常吃紧，很容易造成物理内存不够用，建议调大Master内存。 2. 若 Slave节点选择2核4G机型，在运行MR作业或者Hive作业时，请调节参数。MR作业添加参数-D yarn.app.mapreduce.am.resource.mb=1024 ; Hive作业设置参数 set yarn.app.mapreduce.am.resource.mb=1024; 避免作业hang住。

## Q : java.lang.IllegalArgumentException: Wrong FS: oss://xxxxx, expected: hdfs://ip:9000

A : 因为在操作oss数据的时候，使用hdfs的默认的fs，所以在初始化的时候，要使用oss的路径来初始化fs，这样在后续的操作中，才能使用这个fs来操作oss源上的数据。

```
Path outputPath = new Path(EMapReduceOSSUtil.buildOSSCompleteUri("oss://bucket/path", conf));
```

```
org.apache.hadoop.fs.FileSystem fs = org.apache.hadoop.fs.FileSystem.get(outputPath.toUri(), conf);
if (fs.exists(outputPath)) {
    fs.delete(outputPath, true);
}
```

## Q: 作业读OSS中的Snappy文件出错

A: 作业读OSS中的Snappy出错，提示OOM。一般这种情况出现在作业处理LogService归档到OSS的压缩文件。这是因为LogService归档到OSS的Snappy文件是标准的Snappy文件，区别于Hadoop中分块压缩的Snappy文件。而作业底层读取Snappy文件是走的标准Hadoop接口，所以就会出错。EMR从底层支持了标准Snappy文件，需要您在MR（包括Hive，Pig等类型）作业中配置参数

“io.compression.codec.snappy.native”为“true”。相应地，在Spark作业中我们需要配置

“spark.hadoop.io.compression.codec.snappy.native”为“true”。

## 常见问题

## 版本相关

## 集群创建相关

## 执行计划使用

## 集群端口配置

### Hadoop HDFS

服务	限定	端口	访问要求	配置	描述
NameNode	-	9000	外部	fs.default.na	fs.default.na

				me 或者 fs.defaultFS	me已经过期 但仍能使用
NameNode	-	50070	外部	dfs.http.add ress 或者 dfs.nameno de.http- address	dfs.http.add ress已过期但 仍能使用

## Hadoop YARN (MRv2)

服务	限定	端口	访问要求	配置	描述
JobHistory Server	-	10020	内部	mapreduce.j obhistory.ad dress	-
JobHistory Server	-	19888	外部	mapreduce.j obhistory.w ebapp.addr ess	-
ResourceMa nager	-	8025	内部	yarn.resourc emanager.r esource- tracker.addr ess	-
ResourceMa nager	-	8032	内部	yarn.resourc emanager.a ddress	-
ResourceMa nager	-	8030	内部	yarn.resourc emanager.s cheduler.ad dress	-
ResourceMa nager	-	8088	内部	yarn.resourc emanager.w ebapp.addr ess	-

## Hadoop MapReduce (MRv1)

服务	限定	端口	访问要求	配置	描述
JobTracker	-	8021	外部	mapreduce.j obtracker.a ddress	-

# 技术运维问题

## 集群服务管理页面

您可以购买一台1C2GB配置的ubuntu系统的ECS机器放在内网环境中，作为集群的管理页面访问使用。

集群中的所有服务访问地址一览

软件	服务	访问地址
hadoop		
	yarn resourcemanager	k masternode1_private_ip:8088 ,masternode2_private_ip:8088
	jobhistory	masternode1_private_ip:19888
	timeline server	masternode1_private_ip:8188
	hdfs	masternode1_private_ip:50070,masternode2_private_ip:50070
spark		
	spark ui	masternode1_private_ip:4040
	history	masternode1_private_ip:18080
tez		
	tez-ui	masternode1_private_ip:8090/tez-ui2
hue		
	hue	masternode1_private_ip:8888
zeppelin		
	zeppelin	masternode1_private_ip:8080
hbase		
	hbase	masternode1_private_ip:16010
presto		
	presto	masternode1_private_ip:9090
oozie		
	oozie	masternode1_private_ip:1100

		0
ganglia		
	ganglia	masternode1_private_ip:8085 /ganglia

## 附录

## 错误代码表

### 常见错误码

错误码	说明
4001	请求参数不符合要求,如参数缺少,参数格式不满足要求
4005	没有权限访问其他用户的资源
4006	该集群处于异常状态,不能提交job,请您检查执行计划关联的集群是否已经释放
4007	安全组名称不能为空
4009	请检查您的账户是否欠费或者被冻结
4011	该集群处于异常状态,不能恢复调度,请您检查执行计划关联的集群是否已经释放
5012	用户创建的安全组数量超过上限,您可以前往安全组页面删除一些无用的安全组
5038	作业在运行或待运行的执行计划中,暂不能修改,需等待相关联的执行计划运行完成才能修改,您可以克隆一个新的作业进行修改使用
5039	集群角色锁定失败, E-MapReduce的运行需要一些权限,关于EMR的角色说明,请前往进行创建
5050	访问数据库错误,请稍后重试
6002	更新集群状态失败
8002	没有操作的权限,请使用RAM授权
8003	没有PassRole操作的权限,请使用RAM授权

9006	集群ID不存在,请确认
9007	创建集群时设置master节点的登陆密码,必须包含大小写字母以及数字,长度为8~30

## ECS相关错误

错误码	说明
指定的 InstanceType 未授权使用	创建集群使用的机型尚未申请。请在ECS的购买页面提交申请开通高配。
zone 或者 Cluster 的库存不够了	ECS在该区域的库存不足

## 状态表

### 集群状态表

备注：在集群列表和集群详情页面会查看到集群状态

状态名	状态代码	状态说明
创建中	CREATING	集群正在构建，包括两个阶段：一是物理ECS机器的创建，二是Spark集群的启动，稍等片刻即可达到运行中状态
创建失败	CREATE_FAILED	创建过程中遇到异常，已经创建的ECS机器会自动回滚，在集群列表页面点击状态右边的问号，可查看异常明细
运行中	RUNNING	计算集群处于正常运行状态
空闲	IDLE	集群目前没有运行执行计划
释放中	RELEASING	点击集群状态列表的释放按钮可达到此状态，此状态表示集群正在努力释放，用户稍等片刻，即可达到已释放状态
释放失败	RELEASE_FAILED	释放集群过程中出现异常，在集群列表页面点击状态右边的问号，可查看异常明细，出现此状态需要用户重新点击释放按钮
已释放	RELEASED	计算集群以及托管计算集群的

		ECS机器均处于释放状态
状态异常	ABNORMAL	当计算集群中的一个或多个计算节点出现不可恢复的错误时，则集群处于此状态，用户可点击释放集群按钮，释放该集群

## JOB状态列表

备注：在JOB状态列表里会查看到JOB状态

状态名	状态说明
任务就绪	JOB创建信息完整正确，并被成功的保存，正在准备提交到系统的调度队里中，稍等片刻会处于提交中状态
提交中	JOB在计算集群的提交队列中排序等待，还未提交给计算集群进行计算
提交失败	JOB在提交到计算集群的过程中，发生异常，如需再次执行该JOB，用户需要克隆JOB重新提交
运行中	JOB正在计算集群中拼命的运算，请稍等片刻，点击JOB列表中相应的日志按钮可实时查看输出日志
运行成功	JOB在计算集群中被成功的执行并执行完毕，点击JOB列表中相应的日志按钮可查看相关日志
运行失败	JOB在计算集群中执行出现异常，点击JOB列表中相应的日志按钮可查看相关日志