阿里云
E-MapReduce

快速入门

文档版本：20181130

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的”现状“、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含”阿里云”、Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 **禁止：** 重置操作将丢失用户配置数据。
	该类警示信息可能导致系统重大变更甚至故障，或者导致人身伤害等结果。	 **警告：** 重启操作将导致业务中断，恢复业务所需时间约10分钟。
	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 **说明：** 您也可以通过按**Ctrl** + **A**选中全部文件。
>	多级菜单递进。	设置 > 网络 > 设置网络类型
粗体	表示按键、菜单、页面名称等UI元素。	单击 **确定**。
`courier` 字体	命令。	执行 `cd /d C:/windows` 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	`bae log list --instanceid` *`Instance_ID`*
[]或者[a|b]	表示可选项，至多选择一个。	`ipconfig` *`[-all|-t]`*
{}或者{a|b}	表示必选项，至多选择一个。	`swich` *`{stand | slave}`*

目录

1 准备工作

在创建 E-MapReduce 之前，您需要先完成以下准备工作：

1. 注册阿里云账号

   在申请 E-MapReduce 集群之前，您需要一个阿里云的云账号用于标识您在整个阿里云生态系统中的身份。该账号不仅可以用来申请 E-MapReduce 集群，同时还能够开通阿里云的对象存储服务 OSS、云数据库 RDS等服务。

   如果您还没有阿里云的云账号，请参见注册云账号进行申请。

2. 创建 AccessKey（可选）

   由于 E-MapReduce 调用访问的需要，您至少需要创建一个 AccessKey，创建步骤如下：

   a. 登录阿里云官网。

   b. 登录管理控制台。

   c. 单击 **AccessKeys**。

   > 注意：
   > 若出现如下提示框，请单击继续使用 **AccessKey**。

   d. 单击创建 **AccessKey**。

   e. 输入短信校验码，单击确定。AccessKey 创建成功。

3. 开通阿里云 OSS 服务

   E-MapReduce 会将您的作业日志和运行日志保存在您的阿里云 OSS 存储空间中，所以需要您开通阿里云 OSS 服务，操作步骤请参见开通 OSS 服务。并在您期望创建集群的相同地域创建Bucket，参见创建Bucket。

4. 开通高配机型（可选）

   如果您需要在按量的集群中使用8核及8核以上的机型时，需要先在ECS处申请开通。申请高配机型

5. 准备足够的余额

   目前根据阿里云 ECS 的规则，用户在购买按量付费 ECS 的时候，要保证阿里云账户中至少有 100 元的现金（注意：代金券无效）。因此，在创建按量集群前，请确认您的账户中已至少充值 100 元，否则会创建失败。前往充值。

   当您使用完成并释放集群以后，在没有ECS或者其他按量产品在使用的情况下，您可以将这100元提现，回到您自己的原有账户中。

2 创建 E-MapReduce

2.1 E-MapReduce 快速开始

通过本教程，用户能够基本了解E-MapReduce中集群、作业的作用和使用方法。能够创建一个Spark Pi的作业在集群上运行成功，并最后在控制台页面上看到圆周率Pi的近似计算结果。

> 注意：
> 请确认您已经完成了必选的准备工作。

1. 创建集群。
   a. 在阿里云 *E-MapReduce* 控制台单击上方的**集群管理**页签，进入集群列表页面，并单击右上**创建集群**。
   b. 软件配置。
      A. 选择最新的EMR产品版本，比如**EMR-3.13.0**。
      B. 使用默认软件配置。
   c. 硬件配置
      A. 选择按量付费。
      B. 若没有安全组，请输入新的安全组名称新建。
      C. 选择 Master 4核8G。
      D. 选择 Core 4核8G， 两台。
      E. 其他保持默认。
   d. 基础配置。
      A. 填写集群名称。
      B. 选择日志路径保存作业日志，请务必开启运行日志服务。在集群对应的地域，创建*OSS*的*Bucket*。
      C. 填写密码。
   e. 确认。
2. 创建作业。
   a. 单击上方的**数据开发**页签，进入项目列表页面，并单击右上**新建项目**。
   b. 在**新建项目**对话框输入项目名称和项目描述，单击**创建**。

c. 单击对应项目右侧的工作流设计，进入作业编辑页面。

d. 在页面左侧，在需要操作的文件夹上单击右键，选择新建作业。

e. 填写作业名称和作业描述。

f. 选择Spark类型。

g. 单击确定。

> 说明：
> 您还可以通过在文件夹上单击右键，进行创建子文件夹、重命名文件夹和删除文件夹操作。

h. 参数填写，使用如下。

```
--class org.apache.spark.examples.SparkPi --master yarn-client
 --driver-memory 512m --num-executors 1 --executor-memory 1g
 --executor-cores 2 /usr/lib/spark-current/examples/jars/spark-
examples_2.11-2.1.1.jar 10
```

> 说明：
> 这个`/usr/lib/spark-current/examples/jars/spark-examples_2.11-2.1.1.jar`, 需要根据实际集群中的 Spark 版本来修改这个jar包，比如 Spark 是2.1.1的， 那么就是`spark-examples_2.11-2.1.1.jar`,如果是2.2.0的，那么就是`spark-examples_2.11-2.2.0.jar`。

i. 单击运行。

3. 查看作业日志并确认结果。

   作业运行后，您可以在页面下方的运行记录页签中查看作业的运行日志。单击详情跳转到运行记录中该作业的详细日志页面，可以看到作业的提交日志、YARN Container日志。

2.2 创建集群

本文将介绍创建集群的操作步骤和相关配置。

进入创建集群页面

1. 登录阿里云 *E-MapReduce* 控制台。
2. 完成 RAM 授权，操作步骤请参见角色授权。
3. 在上方选择所在的地域（Region），所创建集群将会在对应的地域内，一旦创建后不能修改。
4. 单击创建集群，进行创建。

创建集群流程

要创建集群，您需要继续完成以下三个步骤：

- 软件配置
- 硬件配置
- 基础配置

步骤1：软件配置

配置项说明：

- **产品版本**：选择默认最新的软件版本。
- **集群类型**：目前的EMR提供了。
  - Hadoop集群，提供半托管的Hadoop、Hive、Spark离线大规模分布式数据存储和计算，SparkStreaming、Flink、Storm流式数据计算，Presto、Impala交互式查询，Oozie、Pig等Hadoop生态圈的组件，具体的组件信息可以在选择界面的列表中查看。
  - Kafka集群，是半托管分布式的、高吞吐量、高可扩展性的消息系统。提供一套完整的服务监控体系，保障集群稳定运行，用户无需部署运维，更专业、更可靠、更安全。广泛用于日志收集、监控数据聚合等场景，支持离线或流式数据处理、实时数据分析等。
  - Druid集群，提供半托管式实时交互式分析服务，大数据查询毫秒级延迟，支持多种数据摄入方式。可与 EMR Hadoop、EMR Spark、OSS、RDS 等服务搭配组合使用，构建灵活稳健的实时查询解决方案。
  - Data Science集群，主要面向大数据+AI场景，提供了Hive、Spark离线大数据ETL，TensorFlow模型训练，用户可以选择CPU+GPU的异构计算框架，利用英伟达GPU对部分深度学习算法进行高性能计算。
- **必选服务**：默认的服务组件，后期可以在管理界面中添加和启停服务。
- **可选服务**：根据需求，您可选择不同的组件，被选中的组件会默认启动相关的服务进程。
- **安全模式**：是否开启集群的 Kerberos 认证功能。一般的个人用户集群无需该功能，默认关闭它。
- **软件自定义配置**：可以指定一个json文件修改软件配置，详细使用方法请参见软件配置。

步骤2：硬件配置

配置项说明：

- 付费配置

- **付费类型**：测试的场景下使用按量开始，测试都正常了以后。可以新建一个包月的生产集群正式使用。

- 集群网络配置
  - **集群可用区**：一般使用默认的可用区即可。
  - **网络类型**：默认使用VPC。若还未创建，可前往*VPC*控制台进行创建。
  - **可用区:**：可用区为在同一地域下的不同物理区域，可用区之间内网互通。
  - **VPC**：选择在该地域的VPC。如没有，单击**创建 VPC / 子网(交换机)** 前往新建。
  - **交换机**：选择在对应的VPC下的在对应可用区的交换机，如果在这个可用区没有可用的交换机，那么就需要前往去创建一个新的使用。
  - **安全组**：一般用户初次来到这里还没有安全组，用户可以输入新的安全组名称新建一个安全组。若已经有在使用的安全组可以直接这里选择使用。
- 集群节点配置
  - **高可用**：打开后，Hadoop集群会有2个master来支持Resource Manager和Name Node的高可用。HBase集群原来就支持高可用，只是另一个节点用其中一个core节点来充当，如果打开高可用，会独立使用一个master节点来支持高可用，更加的安全可靠。后续正式集群如果是使用高可用的，测试情况下也打开高可用。
  - Master节点：主要负责Resource Manager，Name node等控制进程的部署。
    - **Master配置**：根据需要选择实例规格，请参考实例规格族。
    - **系统盘配置**：根据需要选择高效或者是SSD云盘。
    - **系统盘大小**： 根据需要调整磁盘容量，推荐至少120G。
    - **数据盘配置**：根据需要选择高效或者是SSD云盘。
    - **数据盘大小**：根据需要调整磁盘容量，推荐至少80G。
    - **Master数量**：默认1台。
  - Core节点：主要负责集群所有数据的存储，可以按照需要进行扩容。
    - **Core配置**：根据需要选择实例规格，请参考实例规格族。
    - **系统盘配置**：根据需要选择高效或者是SSD云盘。
    - **系统盘大小**： 根据需要调整磁盘容量，推荐至少80G。
    - **数据盘配置**：根据需要选择高效或者是SSD云盘。
    - **数据盘大小**：根据需要调整磁盘容量，推荐至少80G。

- **Core数量**：默认2台，根据需要调整。

— **任务实例组**：不保存数据，调整集群的计算力使用。默认关闭，需要的时候再追加。

步骤3：基础配置

配置项说明：

- 基础信息
  - **集群名称**：集群的名字，长度限制为 1-64 个字符，仅可使用中文、字母、数字、中划线（-）和下划线（_）。
- 运行日志
  - **运行日志**：是否保存作业的日志，日志保存默认是打开的。开启后会需要您选择用来保存日志的 OSS 目录位置，会将您的作业的日志保存到该 OSS 存储目录上。当然，您要使用这个功能必须先开通 OSS，同时上传的文件会按照使用的量来计算用户的费用。强烈建议您打开 OSS 日志保存功能，这会对您的作业调试和错误排查有极大的帮助。
  - **日志路径**：保存日志的 OSS 路径。
  - **统一Meta数据库**：Hive使用统一的集群外部的meta数据库，集群释放后meta信息仍然存在。推荐先关闭。
- **权限设置**：无需调整，使用默认即可。
- 登录设置
  - **远程登录**：是否打开安全组22端口，默认开启。
  - **登录密码**：设置 master 节点的登录密码。8 - 30 个字符，且必须同时包含大写字母、小写字母、数字和特殊字符!@#$%^&*。
- **引导操作（可选）**： 您可以在集群启动 Hadoop 前执行您自定义的脚本，详细使用说明请参见引导操作。

配置清单和集群费用

在配置清单上确认配置和对应的费用。

确认创建

当所有的信息都有效填写以后，**创建**按钮会亮起，确认无误后单击**创建**将会创建集群。

注意：

- 若是按量付费集群，集群会立刻开始创建。页面会返回集群列表页，就能看到在**集群概览**中有一个**初始化**中的集群。请耐心等待，集群创建会需要几分钟时间。完成之后集群的状态会切换为**空闲**。
- 若是包年包月集群，则会先生成订单，在支付完成订单以后集群才会开始创建。

创建失败

如果创建失败，在集群列表页上会显示**集群创建失败**，将鼠标移动到红色的感叹号上会看到失败原因，如下图所示。

创建失败的集群可以不用处理，对应的计算资源并没有真正的创建出来。这个集群会在停留3天以后自动隐藏。

2.3 创建作业

本文介绍老版E-MapReduce作业调度创建作业流程。

要运行一个计算任务，首先需要定义一个作业，其步骤如下：

1. 登录阿里云 *E-MapReduce* 控制台。
2. 选择地域（Region），则作业将会创建在对应的地域内。
3. 单击上方的**老版作业调度**页签，进入作业列表页面。
4. 单击该页右上角的**创建作业**，进入创建作业页面，如下图所示：

5. 填写作业名称。
6. 选择作业类型。
7. 填写作业的应用参数。应用参数需要完整填写该作业运行的 jar 包、作业的数据输入输出地址以及一些命令行参数，也就是将用户在命令行的所有参数填写在这里。如果有使用到 OSS 的路径，可以单击下方的选择 **OSS** 路径选择 OSS 资源路径。关于各作业类型的参数配置，请参见《用户指南》中的《作业》章节。
8. 实际执行命令。这里会显示作业在 ECS 上实际被执行的命令。用户如果把这个命令直接复制下来，就能够在 E-MapReduce 集群的命令行环境中直接运行。
9. 失败重试。可设定重试次数与重试间隔，默认否。

10.选择执行失败后策略。暂停当前执行计划会在这个作业失败后，暂停当前整个执行计划，等待用户处理。而继续执行下一个作业在这个作业失败以后，会忽略这个错误继续执行后一个作业。

11.单击确定完成创建。

作业示例

这是一个 Spark 类型的作业，应用参数中设置了相关的参数，输入输出路径等。

> 注意：
> 本作业仅仅示例，不能实际运行。

oss 与 ossref

oss:// 的前缀代表数据路径指向一个 OSS 路径，当要读写该数据的时候，这个指明了操作的路径，与 **hdfs://** 类似。

ossref:// 同样是指向一个 OSS 的路径，不同的是它会将对应的代码资源下载到本地，然后将命令行中的路径替换为本地路径。它是用于更方便地运行一些本地代码，而不需要登录到机器上去上传代码和依赖的资源包。

上面的例子中， **ossref://xxxxxx/xxx.jar** 这个参数代表作业资源的jar，这个jar存放在OSS上，在运行的时候，E-MapReduce会自动下载到集群中运行。而跟在jar后面的2个 **oss://xxxx** 以及另外两个值则是作为参数出现，他们会被作为参数传递给jar中的主类来处理。

注意：

ossref 不可以用来下载过大的数据资源，否则会导致集群作业的失败。

2.4 创建执行计划

本文介绍老版E-MapReduce执行计划创建流程。

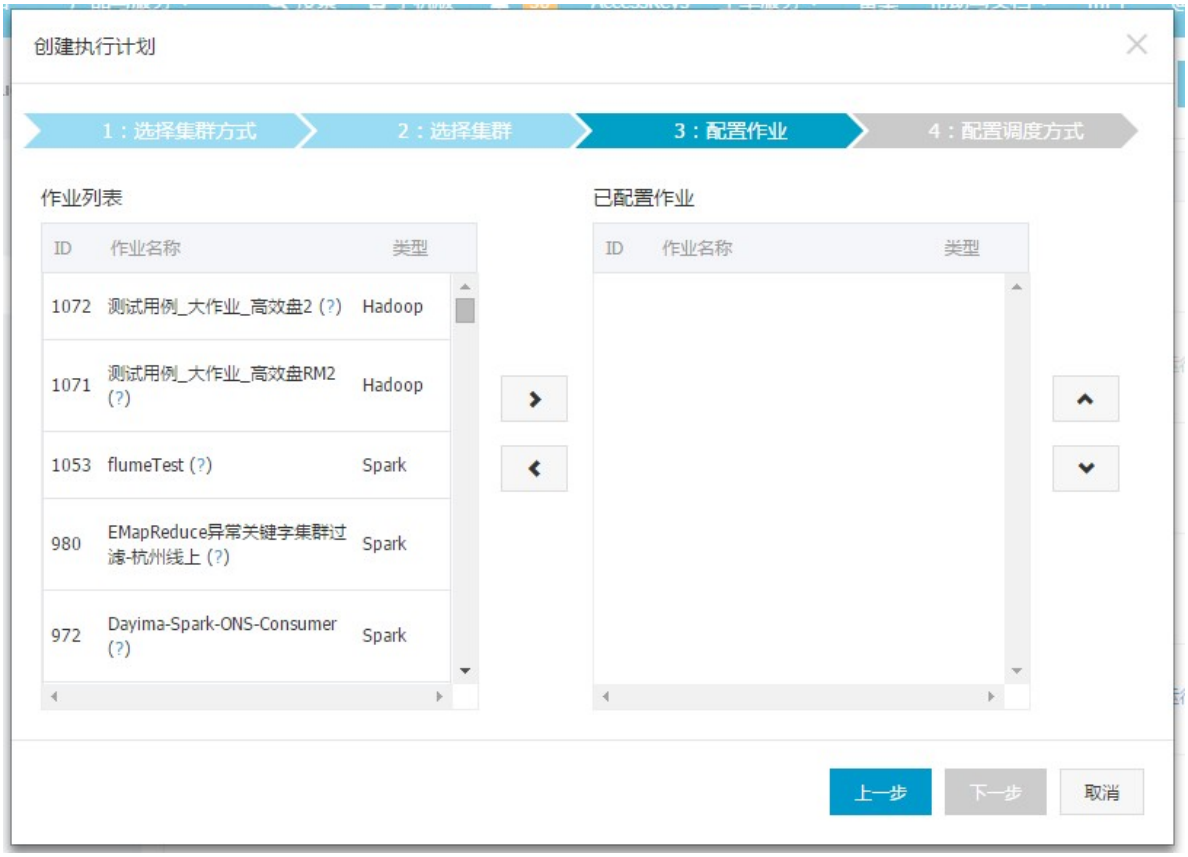
创建完作业后，若要让定义的作业到集群上运行，就需要创建一个执行计划。一个执行计划可以包含多个作业，用户也可自定义其先后顺序。例如，假设用户的一个场景是：产生数据 -> 处理数据 -> 清理数据，则用户可以分别定义三个名为**prepare-data**、**process-data**和**cleanup-data**的作业，然后创建一个执行计划来包含这三个作业。

创建执行计划的步骤如下：

1. 登录阿里云 *E-MapReduce* 控制台。
2. 选择地域（Region）。
3. 单击上方的**老版作业调度**。
4. 单击左侧的**执行计划**页签，进入执行计划页面。
5. 单击右上角的**创建执行计划**，进入创建执行计划页面。
6. 在选择集群方式页面上，有两个选项，分别是**按需创建**和**已有集群**。其中**按需创建**表示目前用户还没有集群，打算用一个临时集群运行该执行计划，并在运行完后将该临时集群自动释放。而**已有集群**表示用户目前已有集群在运行，该执行计划要提交到已有集群中运行。
   a. 如果选择**按需创建**，则步骤跟创建集群一样进行，选择完这个按需集群的配置以后确定即可。
   b. 如果选择**已有集群**，则进入选择集群页面。用户可选择要将该执行计划关联到的集群

目前只有运行中和空闲这 2 个状态的集群可以被提交执行计划。

7. 单击下一步，进入到配置作业页面。在该页面中，左边为先前已经定义好的作业列表，右边是该新创建的执行计划要运行的作业列表。将左边的作业按照执行顺序选择到右边，即可完成执行计划的定义。可以单击问号，查看作业的详细参数。完成后，单击下一步。



8. 设置执行计划名称。

9. 选择调度策略。

   - 手动执行，只有在用户手动单击的情况下才会执行。
   - 周期调度，定义周期调度的频率与启动调度的时间。

10. 单击确认提交，完成执行计划的创建。

3 选型配置

选择配置合适的Hadoop集群是E-MapReduce产品使用的第一步。E-MapReduce配置选型要考虑企业大数据使用场景，估算数据量、服务可靠性要求，又要考虑企业预算。

大数据使用场景

E-MapReduce产品当前主要满足企业的以下大数据场景：

- 批处理场景，要求具有高磁盘吞吐、高网络吞吐，实时性要求低。要处理的数据量大，但对处理的实时性要求不高，可采用MapReduce、Pig、Spark组件。对内存要求不高，选型时重点关注大作业对CPU和内存的需求，以及shuffle对网络的需求。
- Ad-Hoc查询，数据科学家或数据分析师利用即席查询工具检索数据。要求查询实时性高、高磁盘吞吐、高网络吞吐，可以选择使用E-MapReduce的Impala、Presto组件，内存要求高，选型时要考虑数据和并发查询的数量。
- 流式计算、高网络吞吐、计算密集型场景下，可以选择使用E-MapReduce Flink、Spark Streaming、Storm组件。
- 消息队列，高磁盘吞吐，高网络吞吐，内存消耗大，存储不依赖于HDFS，可以选择使用E-MapReduce Kafka。为避免对Hadoop的影响，E-MapReduce将Kafka与Hadoop分为两个集群。
- 数据冷备场景，计算和磁盘吞吐要求不高，但要求冷备成本低，推荐使用EMR D1实例做数据冷备，D1本地盘实例存储成本为0.02元/月/GB。

EMR节点

EMR有3种实例类型：主实例（Master）、核心实例（Core）和计算实例（Task）。

EMR存储可以采用高效云盘、SSD云盘和本地盘。磁盘效能为SSD云盘>本地盘>高效云盘。详细性能参数可以查看：块存储性能和本地盘。

EMR底层存储支持OSS（仅标准型OSS）和HDFS。OSS相对HDFS数据可用性更高，OSS的数据可用性为99.99999999%，HDFS为99.99999%。

存储价格大致估算如下：

- 本地盘实例存储为0.02元/GB/月
- OSS标准型存储为0.14元/GB/月
- 高效云盘存储为0.35元/GB/月

- SSD云盘存储为1元/GB/月

EMR选型

- Master节点选型
  - Master节点主要部署Hadoop的Master进程，如NameNode、ResourceManager等。
  - 生产集群建议打开高可用HA，E-MapReduce的HDFS、YARN、Hive、HBase等组件均已实现HA。生产集群建议在“节点配置”位置开启高可用。如果购买时不开启高可用，集群将无法在后续使用过程中开启高可用功能。
  - Master节点主要用来存储HDFS元数据和组件Log文件，属于计算密集型，对磁盘IO要求不高。HDFS元数据存储在内存中，建议根据文件数量选择16GB以上内存空间。
- Core节点选型
  - Core节点主要用来存储数据和执行计算的节点，运行DataNode、Nodemanager。
  - HDFS(3备份)数据量大于60TB，建议采用本地盘实例（ECS.d1，ECS.d1NE），本地盘的磁盘容量为：(CPU核数/2)*5.5TB*实例数量，如购买4台8核D1实例，磁盘容量为：8/2*5.5*4台=88TB。因为HDFS采用3备份，所以本地盘实例最少购买3台，考虑到数据可靠性和磁盘损坏因素，建议最少购买4台。
  - HDFS数据量小于60TB，可以考虑高效云盘和SSD云盘。
- Task节点选型
  - Task节点主要用来补充Core节点CPU和内存计算能力的不足，节点并不存储数据，不运行DataNode，用户可以根据CPU和内存需求的估算实例个数。

EMR生命周期

EMR支持弹性扩展，可以快速的扩容，灵活调整集群节点配置，或者对ECS节点升降配。

可用区选择

为保证效率，EMR最好与业务系统部署在同一地域的同一个可用区。