

# 弹性 MapReduce

## 新手指引



腾讯云

**【 版权声明 】**

©2013–2024 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

**【 商标声明 】**

及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

**【 服务声明 】**

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

**【 联系我们 】**

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或 95716。

# 新手指引

最近更新时间：2023-07-12 15:19:40

本文将为刚入门弹性 MapReduce（EMR）的用户提供一条学习的路径。

## 1. 熟悉 EMR 的基础知识

- [为什么选择腾讯云 EMR?](#)
- [EMR 有什么功能?](#)
- [EMR 的各个应用场景介绍。](#)
- [腾讯云 EMR 有哪些节点类型?](#)
- [EMR 的组件版本介绍。](#)

## 2. EMR 的计费模式

腾讯云 EMR 的计费模式分为按量计费和包年包月计费。全面了解 EMR 的计费模式，有利于您选择最优的计费方案，计费详情请参见 [EMR on CVM 计费说明](#)。

## 3. 新手入门

### 3.1 业务评估

在购买集群前，需要根据实际情况对具体的业务进行评估，以确保创建的集群是符合您实际需求的。详情请参见 [业务评估](#)。

### 3.2 选购 EMR 集群

在使用腾讯云 EMR 之前，您需要注册腾讯云账号，然后在官网 [产品介绍页](#) 单击[立即选购](#)创建集群，创建集群可参考 [EMR on CVM 快速入门](#)。

### 3.3 角色授权

集群创建成功后，需要在控制台对各协作者/子账号进行权限授权，详情请参考 [协作者/子账号权限](#) 和 [角色授权](#)。

### 3.4 登录 EMR 集群

弹性 MapReduce 可通过两种方式登录集群，[远程登录软件登录](#)（本地系统为 Windows）和[使用 SSH 登录](#)（本地系统为 Linux/Mac OS），详情请参考 [登录集群](#)。

### 3.5 配置集群

登录集群后，可根据自身情况完成集群的配置，配置集群可参考 [软件配置](#)、[设置引导操作](#)、[设置标签](#) 等。

## 4. 控制台功能概述

如果您想	您可以阅读
在集群列表中查看集群的详细信息	<a href="#">实例信息</a>
查看服务告警记录和集群中改变过状态的服务，例如进程重启、创建集群、扩缩容集群等	<a href="#">操作日志</a>

销毁集群	<a href="#">集群销毁</a>
导出存量集群的软件配置参数，后续在新建集群时可使用这些参数进行 <a href="#">软件配置</a> ，从而快速新建一个熟悉的集群	<a href="#">导出软件配置</a>
在集群中新增未安装的组件	<a href="#">新增组件</a>
修改组件参数配置	<a href="#">配置下发</a>
使用配置组将部署同一组件的不同规格或用途的节点进行分组配置管理	<a href="#">配置组管理</a>
对各组件参数最近的一次配置回滚	<a href="#">配置回滚</a>
将 CHDFS 挂载到 EMR 集群	<a href="#">挂载 CHDFS</a>
实现组件级别的重启和新增组件。使用高级管理，可以针对组件中的服务及节点级别进行重启、暂停、维护的操作	<a href="#">重启服务</a>
通过 Master 节点的外网 IP，快捷访问组件原生 UI	<a href="#">软件 WebUI 入口</a>
未开启 Master 节点公网的集群查看组件原生 WebUI	<a href="#">内网 WebUI 访问操作指南</a>
扩缩容集群中的节点，包括 Master 节点、Core 节点、Task 节点、Common 节点、Router 节点和元数据库 MySQL	<a href="#">集群扩容</a> 和 <a href="#">集群缩容</a>
通过 EMR 控制台变更实例的配置	<a href="#">变更配置</a>
选择设置 CORE、TASK、ROUTER 扩容时的硬件规格	<a href="#">设置节点规格</a>
查看集群运行状态的总体视图，获取集群运行状态、核心服务指标、核心节点指标以及节点负载 TOP10 情况	<a href="#">集群概览</a>
对 HDFS、YARN、HIVE、ZOOKEEPER、SPARK、HBase、PRESTO 等服务进行详细监控	<a href="#">服务状态</a>
查看集群所有节点监控概览和所有节点列表，以及所有节点热点图	<a href="#">节点状态</a>
配置弹性 MapReduce 节点和服务监控指标的告警策略	<a href="#">配置告警</a>
查看监控指标	<a href="#">监控指标</a>
通过关键词搜索集群核心服务日志和节点系统日志，在不登录节点的情况下快速查看服务关键日志	<a href="#">日志</a>

## 5. 数据迁移

腾讯云 EMR 数据迁移目前支持两种数据迁移，第一种是通过[腾讯云对象存储 COS 中转](#)，第二种是自建 HDFS 通过 [DistCp 方式迁移](#)，此模式下需要实现自建集群和腾讯云 EMR 集群的网络互通。

## 5.1 通过 COS 迁移数据

迁移数据包括两种：原始数据为非 HDFS 的数据和原始数据为 HDFS 的数据，详细操作过程可参考 [HDFS 通过对象存储数据迁移实践](#)。

## 5.2 自建 HDFS 通过 DistCp 方式迁移

自建 HDFS 通过 DistCp 方式迁移数据，详细步骤可参考 [DistCp 迁移数据](#)。

## 5.3 数据迁移典型示例

1. HBase 迁移有多种方案，其中基于 Snapshot 的迁移方式是推荐的迁移方案。详情请参考 [HBase 数据迁移实践](#)。
2. Hive 迁移涉及两部分，数据迁移和元数据迁移。Hive 表数据主要存储在 HDFS 上，故数据的迁移主要在 HDFS 层。Hive 的元数据主要存储在关系型数据库，可平滑迁移到云上 TencentDB，并可保障高可用。详情请参考 [Hive 数据迁移实践](#)。

## 6. 新手常见问题

### 6.1 计费相关问题

- [按量计费模式支持转为包年包月吗？](#)
- [集群创建后增加组件是否会额外收费？](#)
- [EMR on CVM 控制台显示节点自动续费，为什么关联的 CVM 控制台显示不是自动续费？](#)
- [EMR on CVM 节点续费是否会续费云硬盘（CBS）？](#)
- [EMR on CVM 是否支持自动续费？](#)

### 6.2 弹性 MapReduce 集群管理常见问题

- [非高可用集群是否可以升级成高可用集群？](#)
- [可以选择的组件以及版本有哪些？](#)
- [同类型节点是否可以选择不机型规格？](#)
- [是否支持降低存量节点配置？](#)
- [Core 节点是否支持缩容？](#)

## 7. 反馈与建议

使用腾讯云 EMR 产品和服务中有任何问题或建议，您可以通过以下渠道反馈，将有专人跟进解决您的问题：

- 如果发现产品文档的问题，如链接、内容、API 错误等，您可以单击文档页右侧[文档反馈](#)或选中存在问题的内容进行反馈。

- 
- 如果遇到产品相关问题，您可咨询 [在线客服](#) 寻求帮助。
  - 如果您有其他疑问，可前往 [腾讯云开发者社区](#) 进行提问。