

| 大数据技术-第五章：YARN资源管理器

# YARN的工作流程



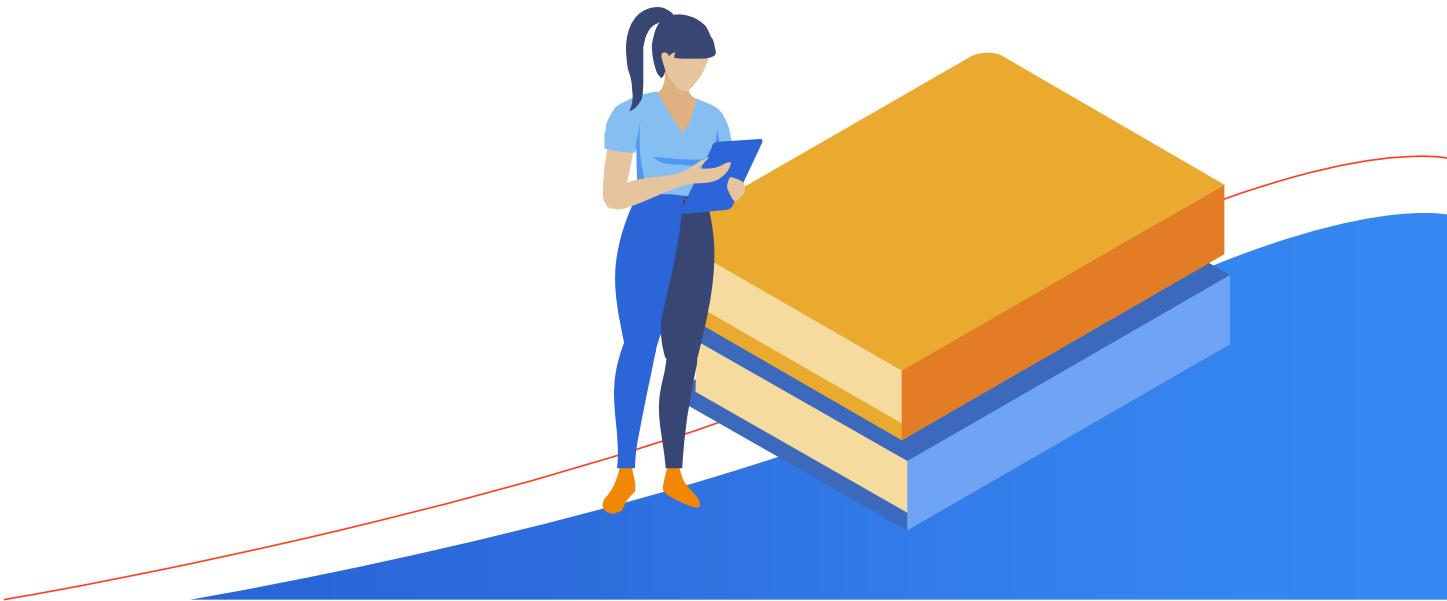
# CONTENTS

---

**01.** YARN上运行的应用类型

**02** YARN的整体工作流程

**03** MapReduce on YARN工作流程



# 01

## YARN上运行的应用类型



运行在YARN上的应用程序主要分为两类：

(1) 短应用程序

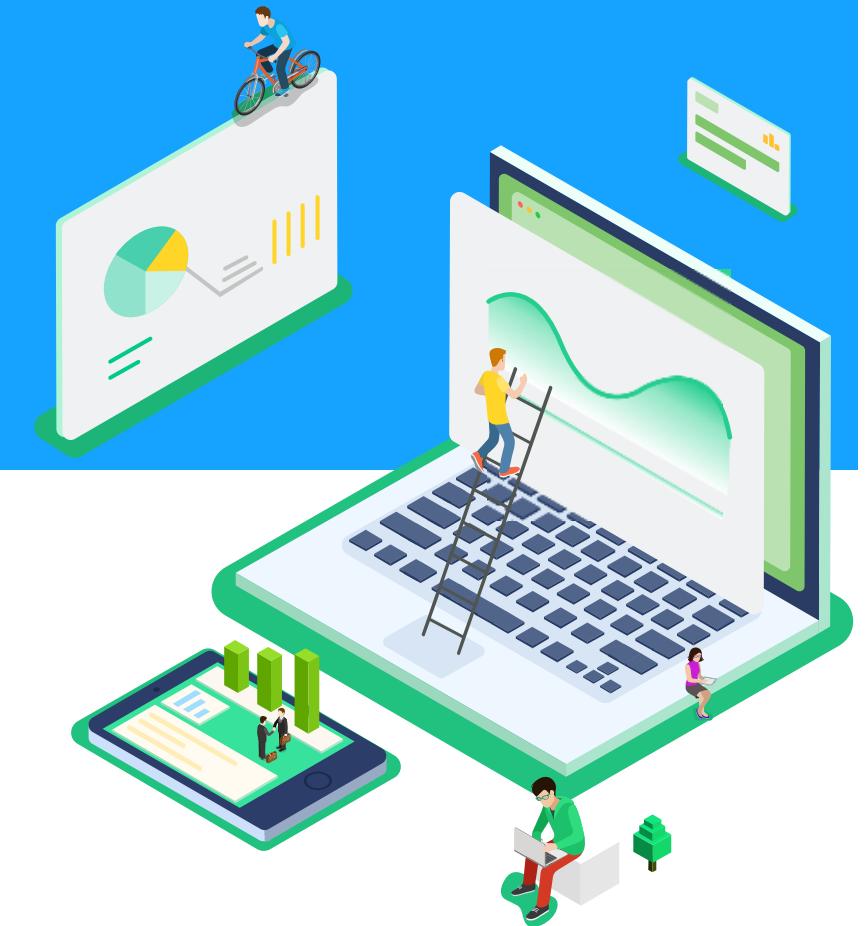
指一定时间内（可能是秒级、分钟级、小时级或天级，尽管天级别或者更长时间的也存在，但非常少）可运行完成并正常退出的应用程序，比如MapReduce作业、Tez DAG作业等。

(2) 长应用程序

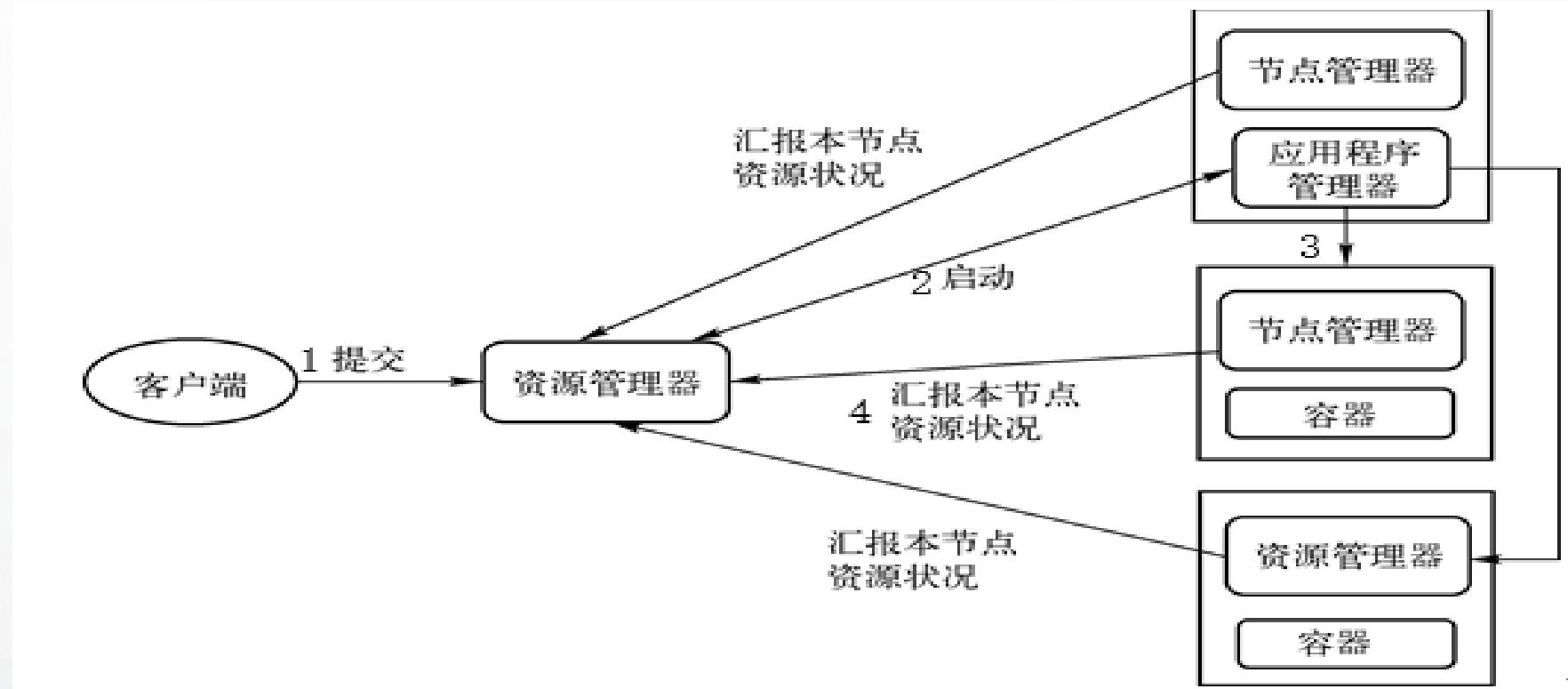
指不出意外，永不终止运行的应用程序，通常是一些服务，比如Storm Service（主要包括Nimbus和Supervisor两类服务），HBase Service（包括Hmaster和RegionServer两类服务）等，而它们本身作为一个框架提供了编程接口供用户使用。

## 02

### YARN的整体工作流程



## » YARN的整体工作流程



1. 客户端 (Client) 向资源管理器 (ResourceManager) 中提交一个作业，作业包括应用程序管理器 (ApplicationMaster) 程序，启动ApplicationMaster的程序和用户自己编写的程序 (比如 MapReduce，因为YARN中可以运行多种应用程序，MapReduce只是其中一种) 。
2. ResourceManager会为该应用程序分配一个容器 (Container) ，它首先会跟Node Manager进行通信，要求它在这个Container中启动应用程序的ApplicationMaster。
3. ApplicationMaster一旦启动以后，它首先会向ResourceManager注册，这样用户可以直接通过ResourceManager查看应用程序的运行状态，然后它将为各个任务申请资源并监控它们的运行状态，直到运行结束，它会以轮询的方式，通过RPC协议向ResourceManager申请和领取资源，一旦ApplicationMaster申请到资源后，它会和Node Manager通信，要求它启动任务并运行任务。
4. 各个任务通过 RPC协议 向 ApplicationMaster汇报自己的状态和进度，这样会让ApplicationMaster随时掌握各个任务的运行状态，一旦任务运行失败，ApplicationMaster就会重启该任务，重新申请资源。应用程序运行完成后，ApplicationMaster就会向ResourceManager注销并关闭此任务。在应用程序整个运行过程中可以用RPC向ApplicationMaster查询应用程序当前的运行状态，在Web上也可以看到整个的作业的运行状态。

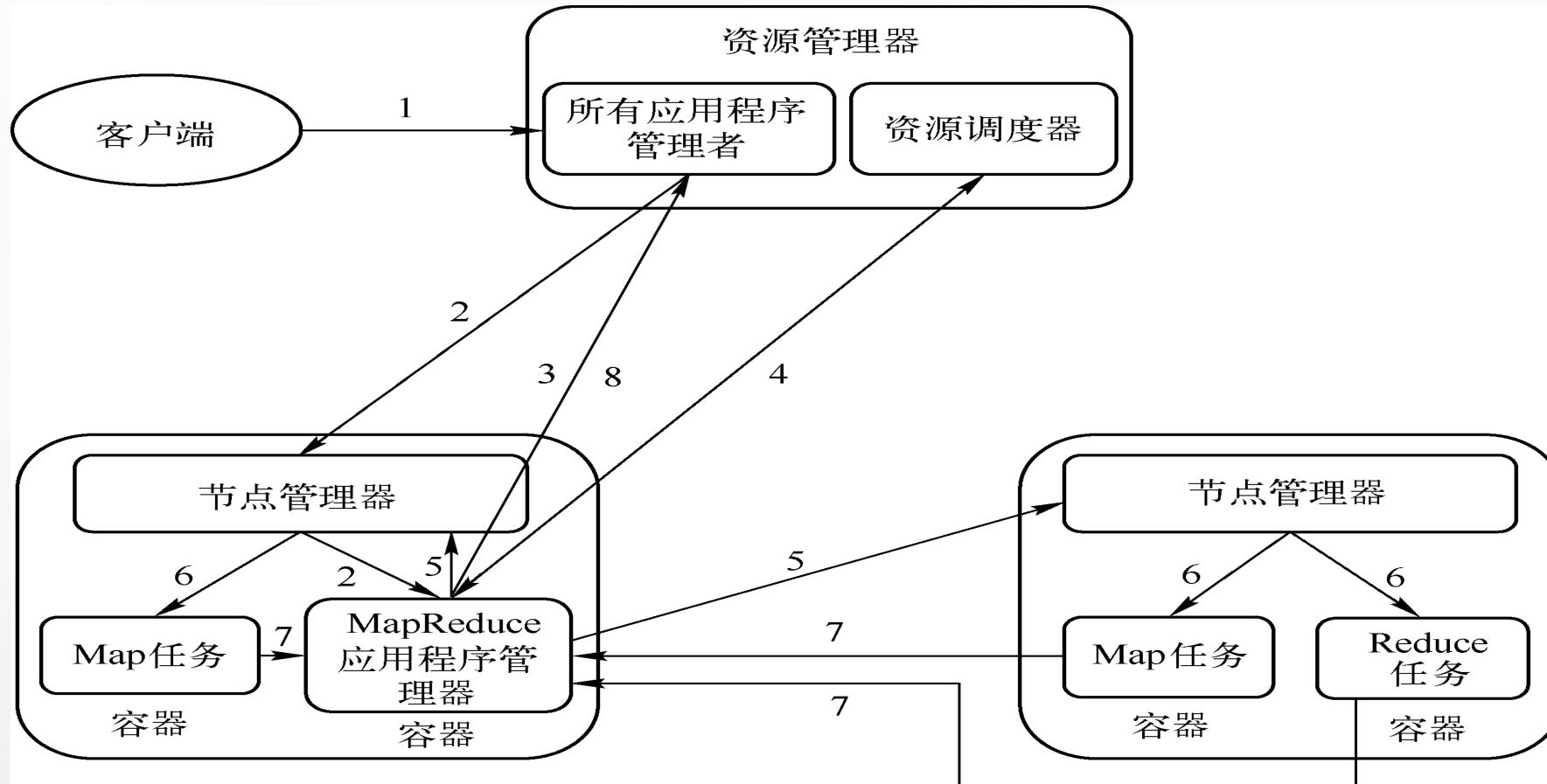


# 03

## MapReduce on YARN工作流程



# >> MapReduce on YARN工作流程



步骤 1：用户向资源管理器（ResourceManager）中提交作业，作业包括MapReduce应用程序管理器，启动MapReduce应用程序管理器的程序和用户自己编写的MapReduce程序。用户提交的所有作业由Applications Manager（所有应用程序管理者）管理。

步骤 2：ResourceManager为该应用程序分配第一个Container，并与对应的NodeManager通信，要求它在这个Container中启动MapReduce的ApplicationMaster。

步骤 3：ApplicationMaster首先向 ResourceManager注册，这样用户可以直接通过 ResourceManager查看应用程序的运行状态，然后它将为各个任务申请资源，并监控他的运行状态，直到运行结束，即要重复步骤4-7。

步骤 4：ApplicationMaster采用轮询的方式通过RPC协议向ResourceManager申请和领取资源。



步骤 5：一旦ApplicationMaster申请到资源后，便与对应的NodeManager通信，要求启动任务。

步骤 6：NodeManager为任务设置好运行环境，包括环境变量、JAR包、二进制程序等，然后将任务启动命令写到另一个脚本中，并通过运行该脚本启动任务。

步骤 7：各个任务通过RPC协议向ApplicationMaster汇报自己的状态和进度，ApplicationMaster随时掌握各个任务的运行状态，从而可以在任务失败时重新启动任务。在应用程序运行过程中，用户可以随时通过RPC协议向ApplicationMaster查询应用程序的当前运行状态。

步骤 8：应用程序运行完成后，ApplicationMaster向ResourceManager注销并关闭自己。



Turing AI 万维  
图灵 | 大数据系列课程

大数据

BIG  
DATA  
智 / 能 / 科 / 技

放 / 眼 / 未 / 来

