1. YARN：Hadoop的资源管理器，负责集群资源管理和调度。
   1. RsourceManager：主要的资源协调者和管理者。ResourceManager负责给用户提交的所有程序分配资源。根据程序优先级、队列容量、数据位置等信息做出决策。
   2. NodeManager：每个具体节点的管理者。负责节点内容器的生命周期管理，监视资源和跟踪节点健康。

启动时向ResoureManager发送心跳、维护Container的生命周期并监控Container的资源使用情况。管理任务运行时的依赖，更具ApplicationMaster的需要，在启动Container之前将需要的程序及依赖拷到本地。

* 1. ApplicationMaster：ApplicationMaster负责协调来自ResoureceManager的资源。并通过NodeManager监视容器内资源的使用情况。

根据应用的运行状态来决定动态计算资源需求，向ResoureceManager申请资源并监控所申请资源的使用情况。跟踪任务状态和进度，报告资源的使用情况和应用的进程信息。

2.4 Container：Container是YARN中对资源的抽象，它封装了多维度资源。如内存、CPU、磁盘、网络等。YARN会为每个任务分配一个Container，该任务只能使用Container中的资源。

3. Yarn任务执行原理：

3.1 Client将作业提交到Yarn

3.2 ResoureceManager选择一个NodeManager，启动Container并运行Application Master。

3.3 ApplicationMaster根据实际需要向ResourceManager请求更多的Container资源

3.4 ApplicationMaster通过获取到的Container资源执行分布式计算。