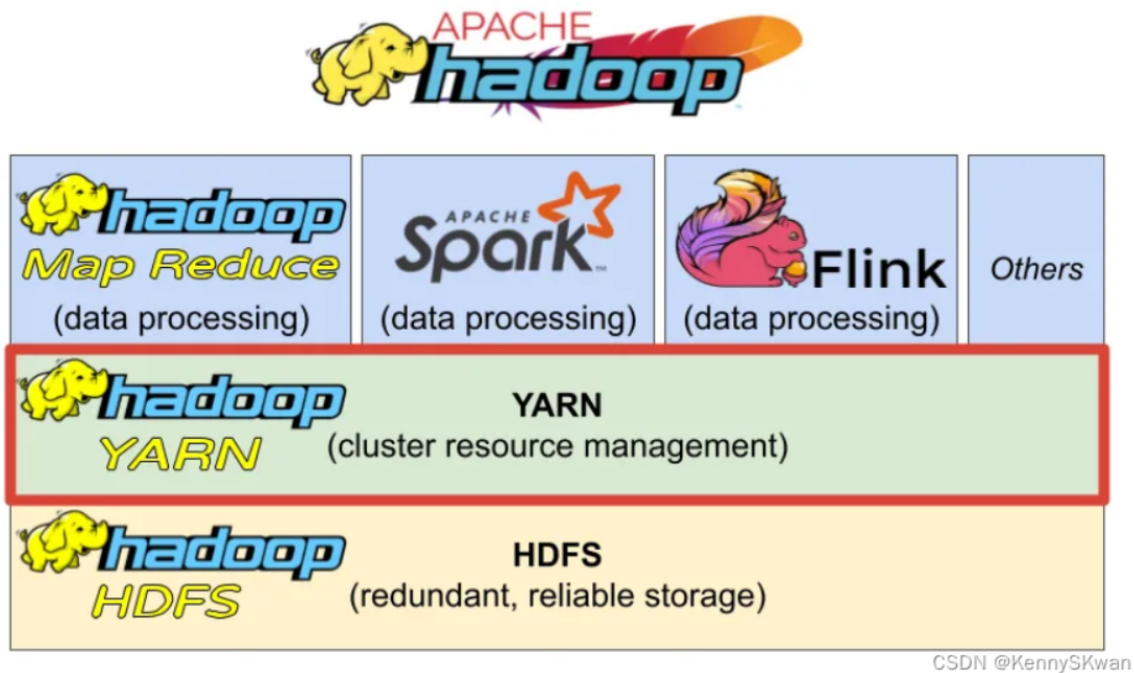


hadoop

主体构成



HDFS

Hadoop Distributed File System的简称

Namenode

负责构建命名空间，管理文件的元数据等

Datanode

Datanode负责实际存储数据，负责读写工作。

MapReduce

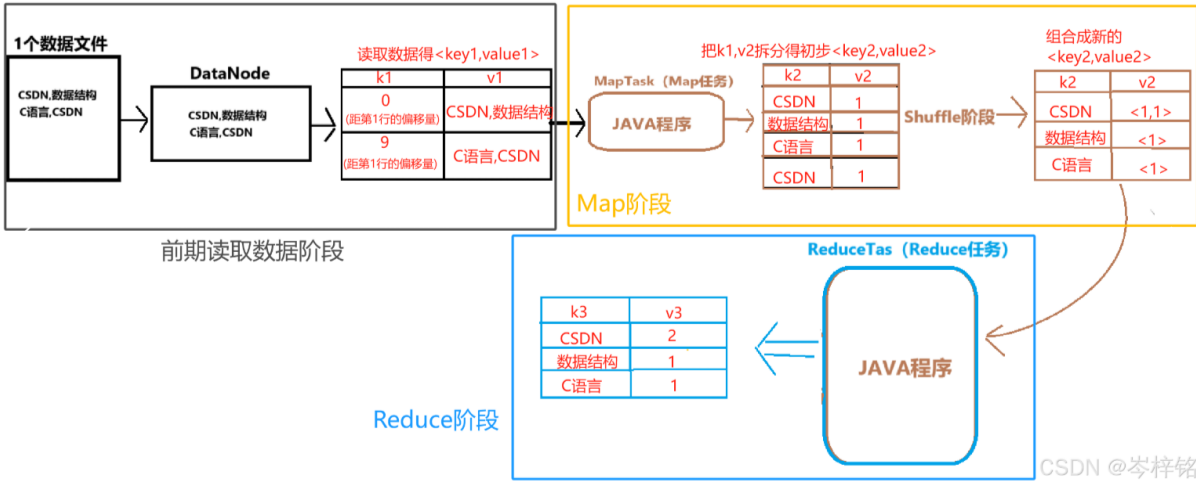
Map

就是将海量大数据拆分成若干份小数据，多个程序同时并行计算出自己的结果

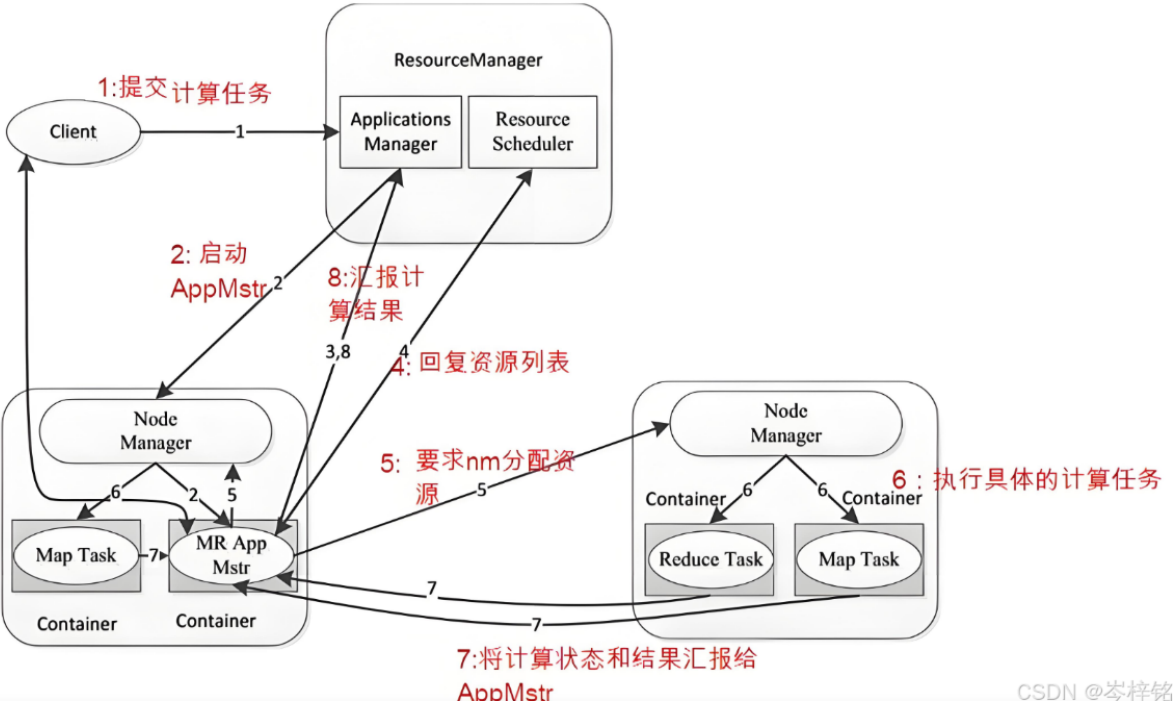
Reduce

将【Map】阶段求得的所有结果汇总，最终得到想要的的结果

完整流程

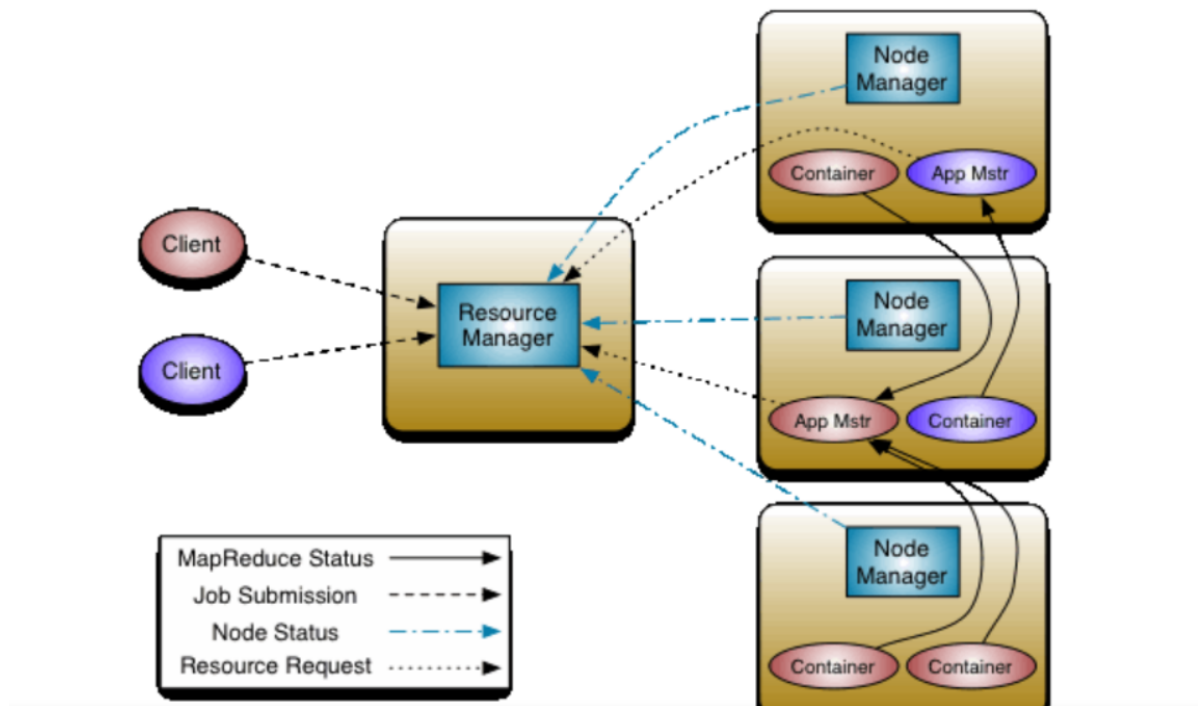


【YARN】的集群[MapReduce工作流程]

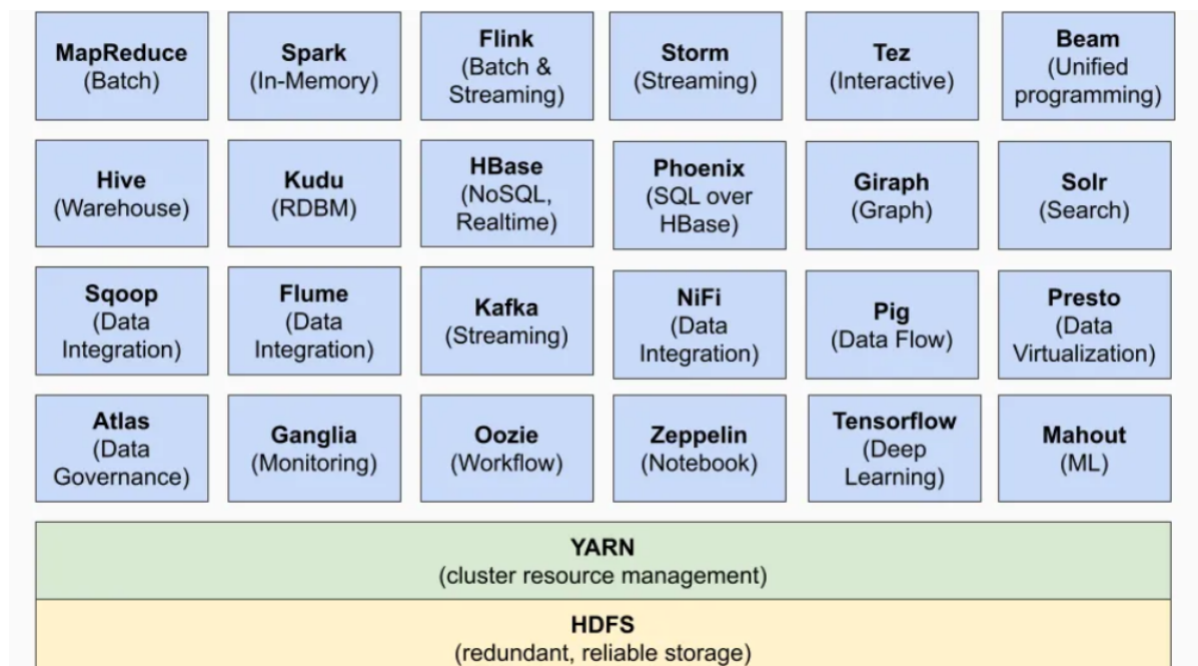


YARN

YARN的基本核心思想是将资源管理和作业调度从特定的计算框架（如MapReduce）中分离出来使其成为单独的守护进程，使得Hadoop集群能够更通用地支持多种类型的应用程序和工作负载。



示例



总结

Hadoop中可以用于利用的知识包括 MapReduce 和Yarn的相关部分

其中， MapReduce很好的展示了一种可行的分布式计算方案，先分配任务，最后做个汇总。

而Yarn管理着全部的资源，当有任务过来的时候，根据所掌握的资源进行分配。

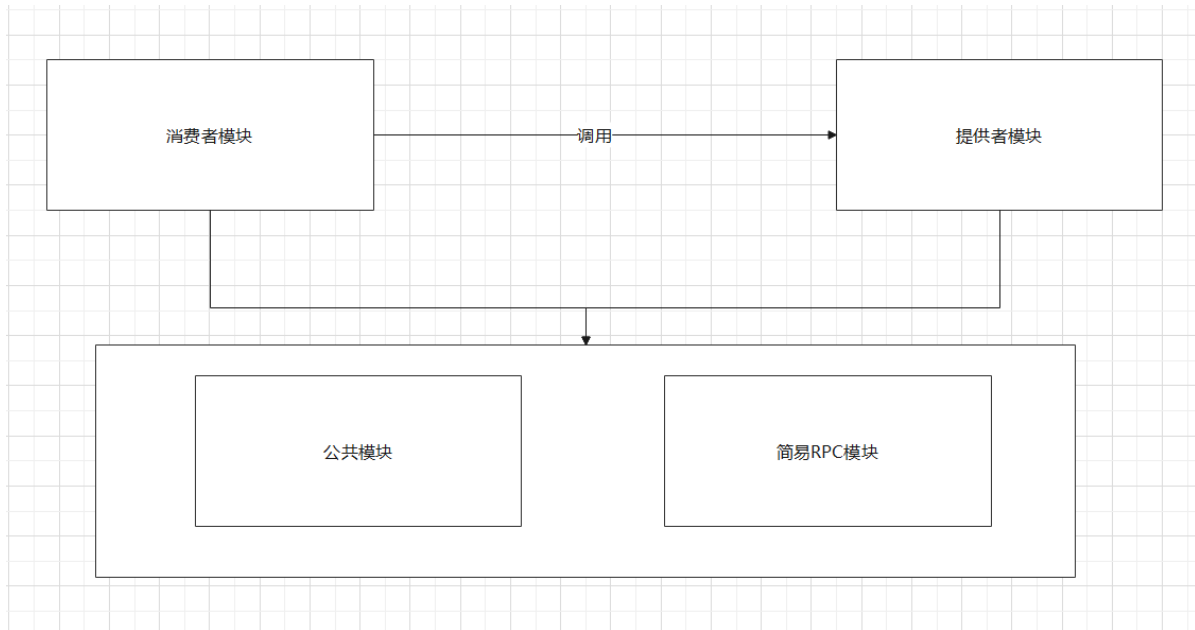
经验

在本RPC项目中，可以参考Yarn中的资源管理，设置多台机器，灵活增加或减少对应任务所分配的机器数量。而MapReduce相关内容暂时用不上，因为MapReduce对应的是大量数据中进行计算处理，而一般RPC中的任务不会设计相关内容。待后续考察

RPC

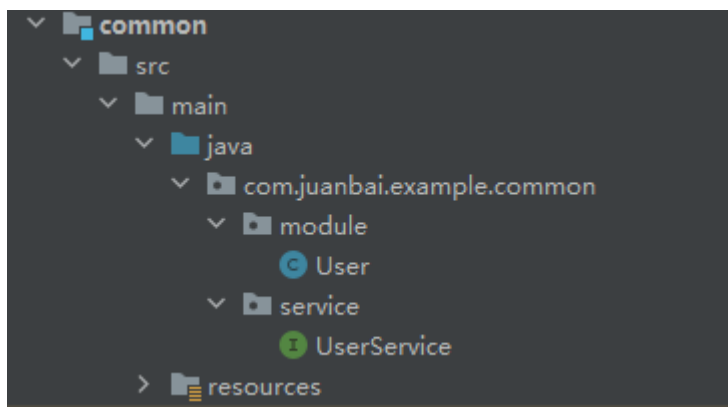
本地架构

结构



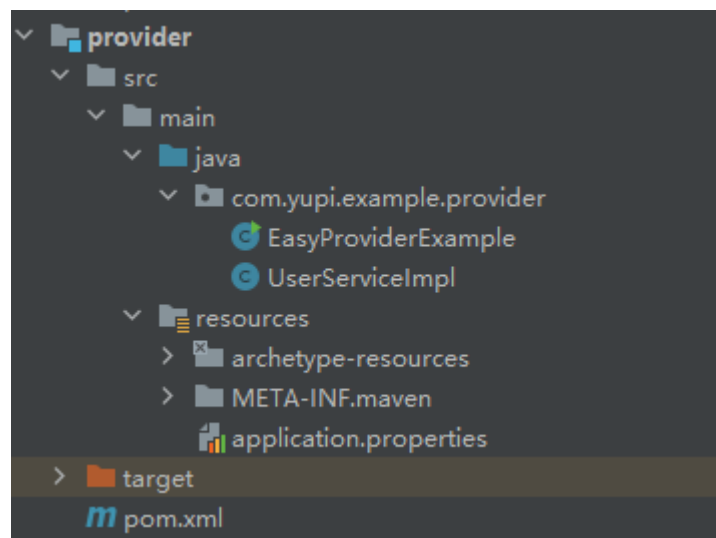
公共模块

提供公共内容，给服务提供者及消费者提供基础的功能。包括一些工具



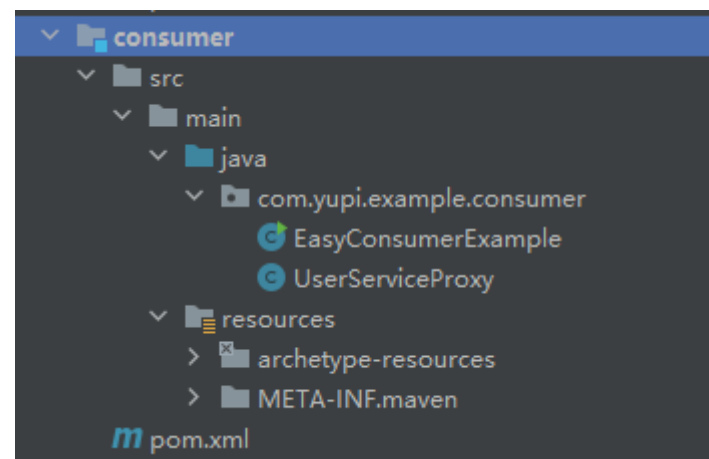
服务提供者

提供服务

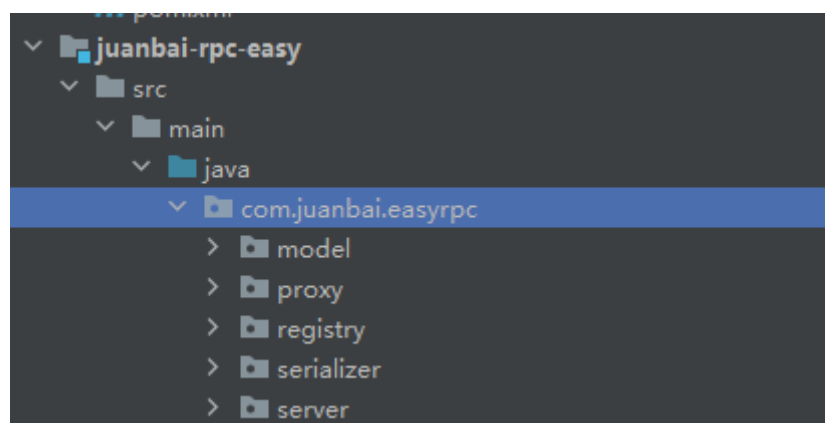


服务消费者

消费服务



rpc模块



提供RPC功能

包括:

web服务器

本地服务注册器

序列化器

提供者调用请求处理器

消费者发起调用代理

