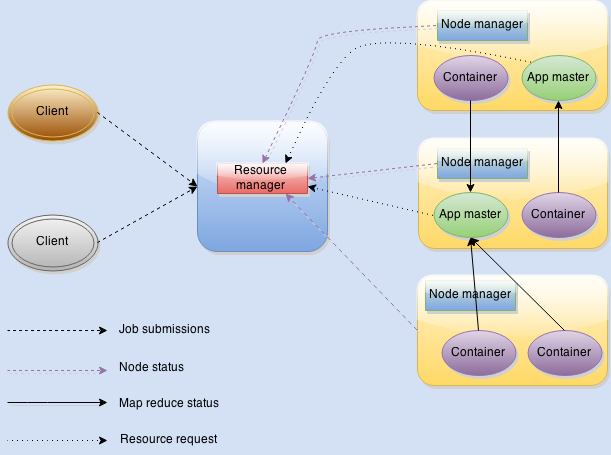
# Hadoop Common

Bazı modüllerin Hadoop’a erişebilmesi için gerekli olan kütüphaneleri sağlar. Mesela Hive yada Hbase HDFS’e erişmek için bu kütüphaneleri kullanır

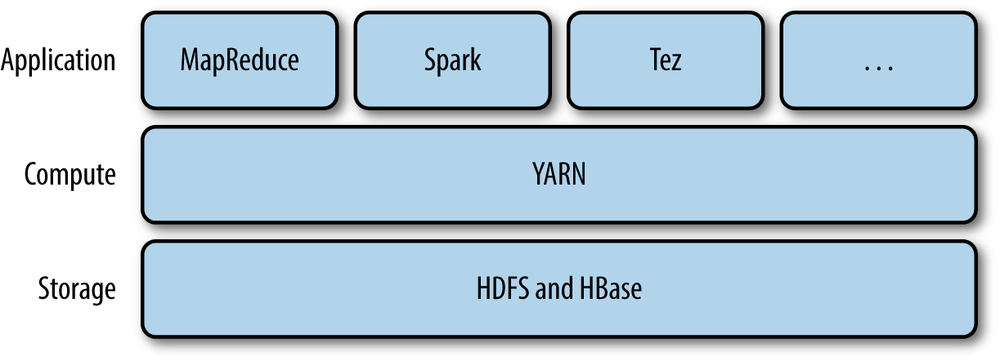
# Hadoop YARN

Hadoop YARN genel olarak çalışan uygulamalara ne kadar ram, CPU ayrılacağını hesaplar ve yönetir.



YARN genel olarak MapReduce gibi dağıtık uygulamalarımız için kaynak yönetimini sağlar(ram,cpu).Fakat MapReduce kodları geliştirirken YARN’a doğrudan müdahe etmemize gerek yoktur.YARN arka planda resource yönetimini bizim için sağlar

Ayrıca Yarn MapReduce dışında Spark ve Tez için kaynak yönetimi’de sağlar



## **ResourceManager** Nedir?

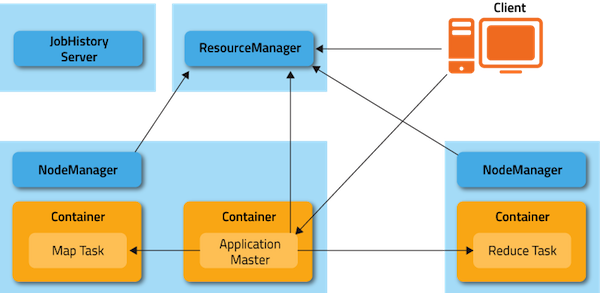
ResourceManager, hadoop cluster içerisindeki kaynakları takip eder ve uygulamaları yönetir

## NodeManager **Nedir**?

NodeManager, hadoop cluster içerisinde bulunan her bir slave üzerinde çalışan uygulamadır.Görevi ise kendisini ResourceManager’a tanıtmak ve belirli periyotlarda bulunduğu makine ile ilgili bilgi mesajı vermektir.Ayrıca bulunduğu makinadaki kaynakları(memory,vcore) yine ResourceManager’a bildirir.

## Container nedir?

HDFS içerisinde uygulamaların çalıştığı,belirli bir kaynağa sahip olan(ram,core) bileşendir.MapReduce job’ı bir yada birden fazla container içerisinde çalışabilir.



## **Hadoop üzerinde bir uygulama nasıl çalışır?**

* Öncelikle Client bir uygulamayı başlatır
* Uygulama ResourceManager’a yönlendirilir
* ResourceManager cluster üzerindeki kaynakları kontrol eder ve node içerisindeki ApplicationMaster’a bir yada birden fazla Container başlatması için bilgi verir.Çalıştırılan Container’lar farklı makinalar üzerinde bulunabilir.
* ApplicationMaster ise uygun sayıda Container başlatarak, bu Container’lar içerisinde MapReduce uygulamalarının çalışmasını sağlar.
* Daha sonra sonuçlar client’a bildirilir

