**为什么要使用分布式文件系统呢？**

访问流量的增多，项目的扩大，出现了两个明显的问题。

1. **文件代码的耦合，存放混乱2.静态文件的访问占用资源，影响正常的业务的进行。**

**分离后的优势所在**：Web/App服务器可以更专注发挥动态处理的能力；独立存储，更方便做扩容、容灾和数据迁移；方便做图片访问请求的负载均衡，方便应用各种缓存策略（HTTP Header、Proxy Cache等），也更加方便迁移到CDN。

**分布式存储系统的三个重要的特点：**

1.扩展能力：容量的扩展，处理速度的扩展

2.高可用性:一是整个文件系统的可用性，二是数据的完整和一致性；

3.弹性存储: 可以根据业务需要灵活地增加或缩减数据存储以及增删存储池中的资源，而不需要中断系统运行

FastDFS分布式文件系统

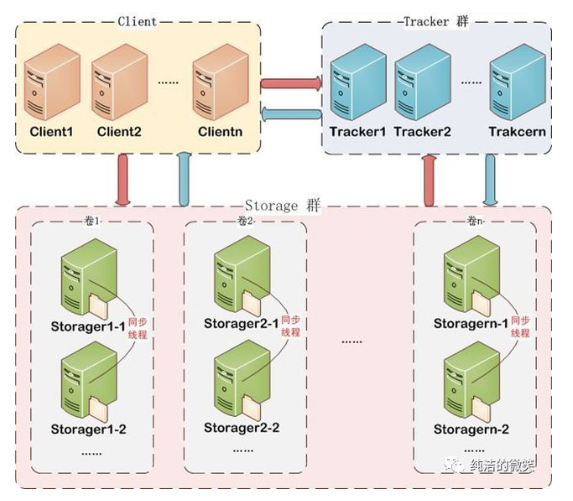
适合文件中小文件（建议范围：4KB~500MB）的在线服务，如相册文件、视频网站

FastDFS支持Linux、UNIX系统，提供了JAVA、C、PHP语言的API

**FastDFS服务端的三个角色：**跟踪服务器（**tracker server**）、存储服务器（**storage server**）和客户端（**client**）

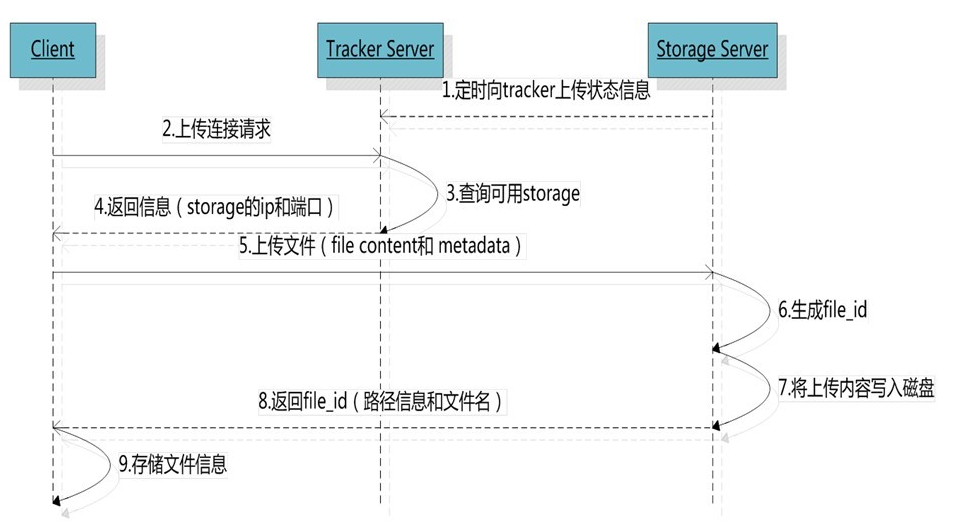
1.Tracker：跟踪服务器，主要做调度工作，起负载均衡的作用。在内存中记录集群中所有存储组和存储服务器的状态信息，是客户端和数据服务器交互的枢纽。

2.Storage：文件和文件属性（meta data）都保存到存储服务器上。

3.Client：请求的发起方，通过专有接口，使用TCP/IP协议与跟踪器服务器或存储节点进行数据交互。

**group** ：同组内服务器上的文件是完全相同的 ，同一组内的storage server之间是对等的， 文件上传、 删除等操作可以在任意一台storage server上进行 。

**上传机制：**



生成的文件名由group、存储目录、两级子目录、fileid、文件后缀名拼接而成。

Storage在启动后会会保持周期性的心跳机制，向tracker告知自己的状态信息。