



分布式文件系统

FastDFS



目 录

CONTENTS



分 布 式 文 件 系 统



F a s t D F S 结 构



FastDFS 实 际 部 署



实 机 演 示



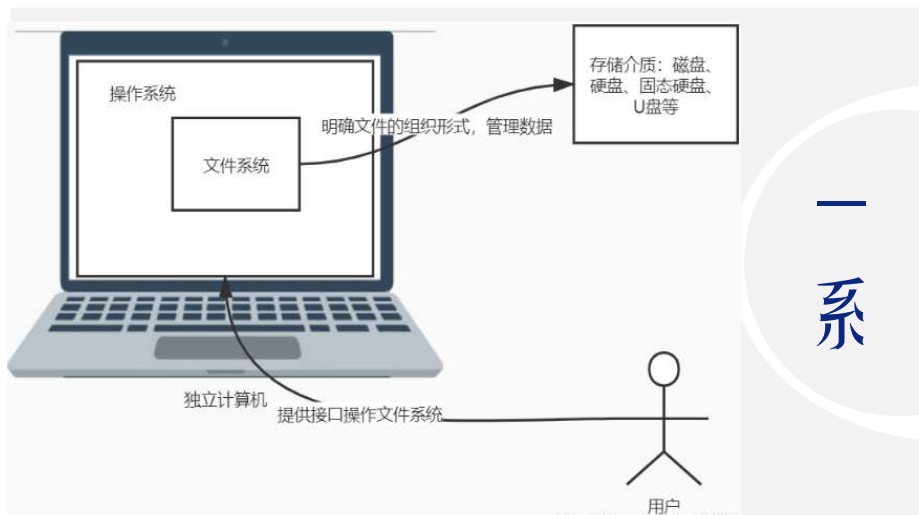
PART ONE

分布式文件系统

分布式文件系统——是什么

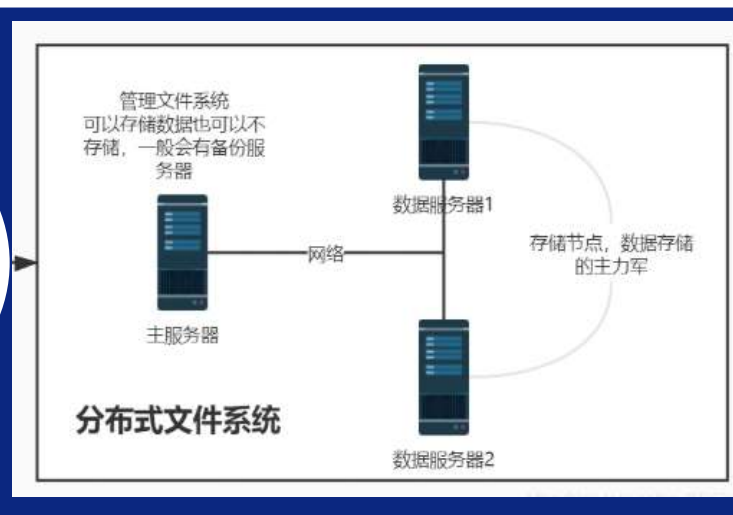
分布式文件系统

分布式文件系统中的**数据存储在台机器上**，这些专门用来存储数据的机器称之为**存储节点**，由**多个节点构成分布式集群**，节点上的小的分布式文件系统组合成总的分布式文件系统，由主服务器对总的文件系统进行管理。用户任意访问某一台主机，都能获取到自己想要的目标文件。



一般系统

分布式



分布式文件系统——为什么

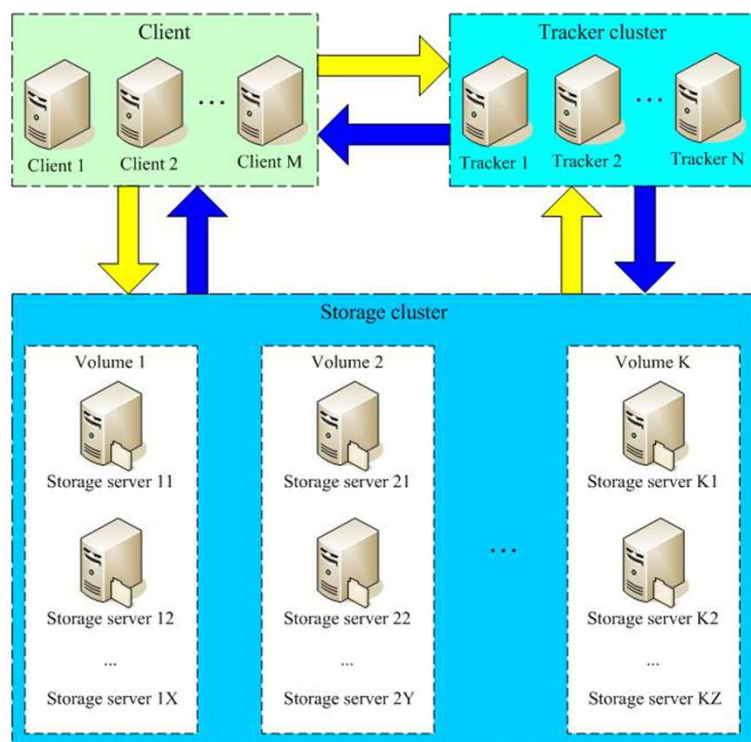
为什么需要分布式文件系统？

从存粹的应用角度而言，就是用**较低的存储介质成本完成海量数据的存储**，且能够满足大部分的业务数据使用需求



分布式文件系统——什么是FastDFS

什么是FastDFS



- FastDFS是一个开源的轻量级分布式文件系统，它对文件进行管理，功能包括：文件存储、文件同步、文件访问（文件上传、文件下载）等，解决了大容量存储和负载均衡的问题。
- FastDFS设计是用来存储小文件的，过大的文件处理方案是拆分为小文件，可跟踪小文件的上传情况。如果应用场景都是处理大文件，可能选择其他分布式文件系统方案会更合适。

分布式文件系统——FastDFS与HDFS

VS

F a s t D F S

FastDFS主要用于互联网网站，为文件上传和下载提供在线服务。所以在负载均衡、动态扩容等方面都支持得比较好，**FastDFS不会对文件进行分块存储。**

FastDFS用于**存储中小文件**都是不错的，比如用户头像，一些较小的音视频文件等等都行，适合处理中小型文件**(4KB<file_size<500MB)**。

H D F S

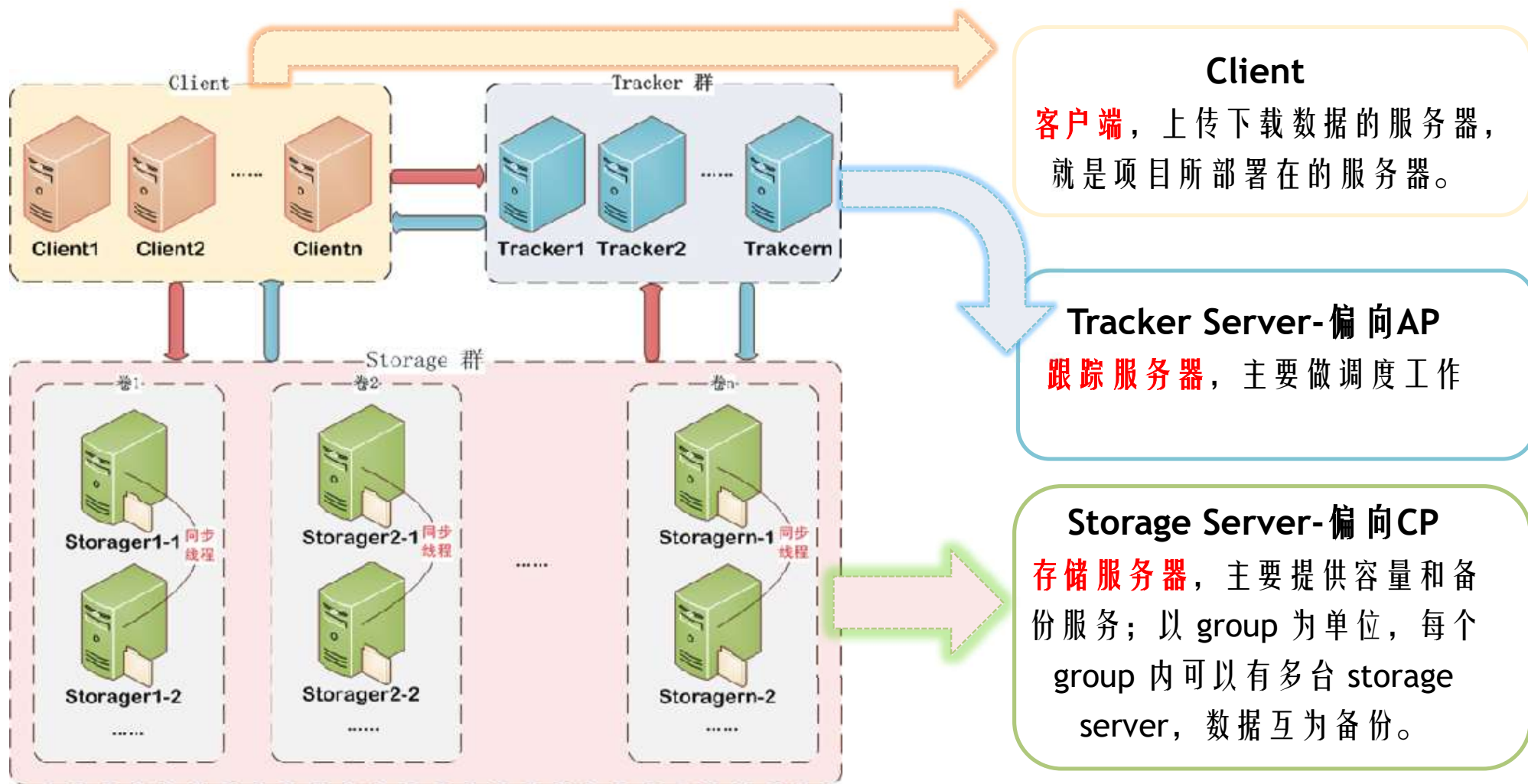
Hadoop中的文件系统HDFS主要解决并行计算中分布式存储数据的问题。其**单个数据文件通常很大**，采用了分块（切分）存储的方式，所以是大数据大文件存储来使用的场景，**不适合实时低延迟的文件读写场景。**



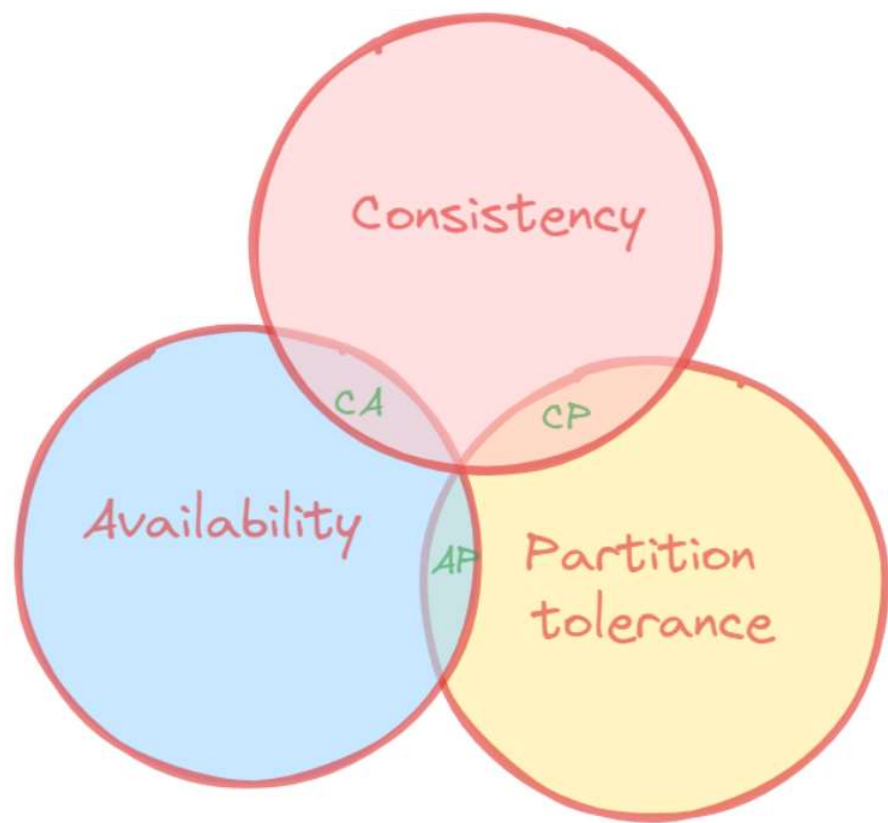
PART TWO

F a s t D F S 结 构

FastDFS 结构——三个角色



FastDFS 结构——CAP 理论



Consistency (一致性)

数据保持一致，在分布式系统中，可以理解为多个节点中数据的值是一致的

C

Availability (可用性)

某一个服务在某台或者多台服务器出现问题后，在其他服务器上依然能够完成用户的操作

A

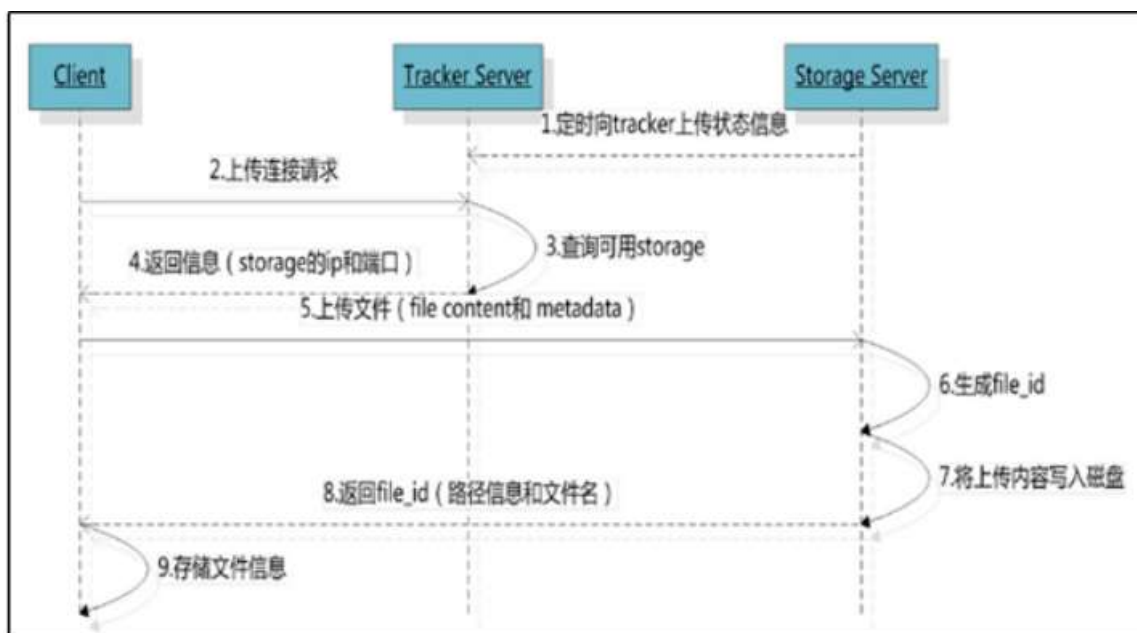
Partition tolerance (分区容错性)

整个系统集群中有单台服务器，或多台服务器出现故障后（不是提供的同一个服务），正常服务的服务器依然能正常提供服务

P

FastDFS 结构——文件上传

FastDFS 文件上传过程



1. Storage会定时的向Tracker发送心跳，告诉Tracker自己还活着，这样Fastdfs就可以工作了；
2. 客户端发送上传请求给Tracker，Tracker会检查是否有可用Storage；
3. 如果有可用的，客户端就可以上传文件数据到Storage上；
4. Storage将文件写入磁盘后，会返回路径信息给客户端；
5. 客户端就可以根据这个路径信息找到上传的文件

FastDFS 结构——文件上传

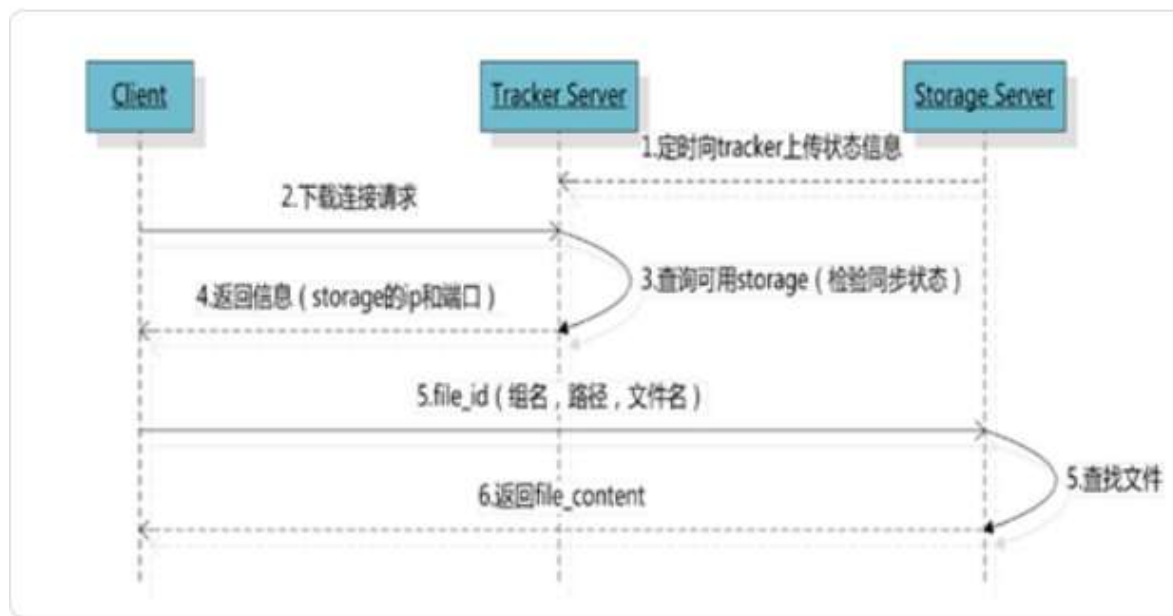
1. 当Tracker收到客户端上传文件的请求时，会为该文件分配一个可以存储文件的group
2. 当选定了group后就要决定给客户端分配group中的哪一个storageserver
3. 当分配好storageserver后，客户端向storage发送写文件请求，storage将会为文件分配一个数据存储目录。
4. 然后为文件分配一个fileid，最后根据以上的信息生成文件名存储文件。

组名 磁盘 目录 文件名
↑ ↑ ↑ ↑
group1/M00/00/00/wKgDCmDun6mATfmlARBlpcz7xUs536.jpg

- 组名：文件上传后所在的storage组名称，在文件上传成功后有storage服务器返回，需要客户端自行保存。
- 虚拟磁盘路径：storage配置的虚拟路径，与磁盘选项store_path*对应。如果配置了store_path0则是M00，如果配置了store_path1则是M01，以此类推。
- 数据两级目录：storage服务器在每个虚拟磁盘路径下创建的两级目录，用于存储数据文件。
- 文件名：与文件上传时不同。是由存储服务器根据特定信息生成，文件名包含：源存储服务器IP地址、文件创建时间戳、文件大小、随机数和文件拓展名等信息。

FastDFS 结构——文件下载

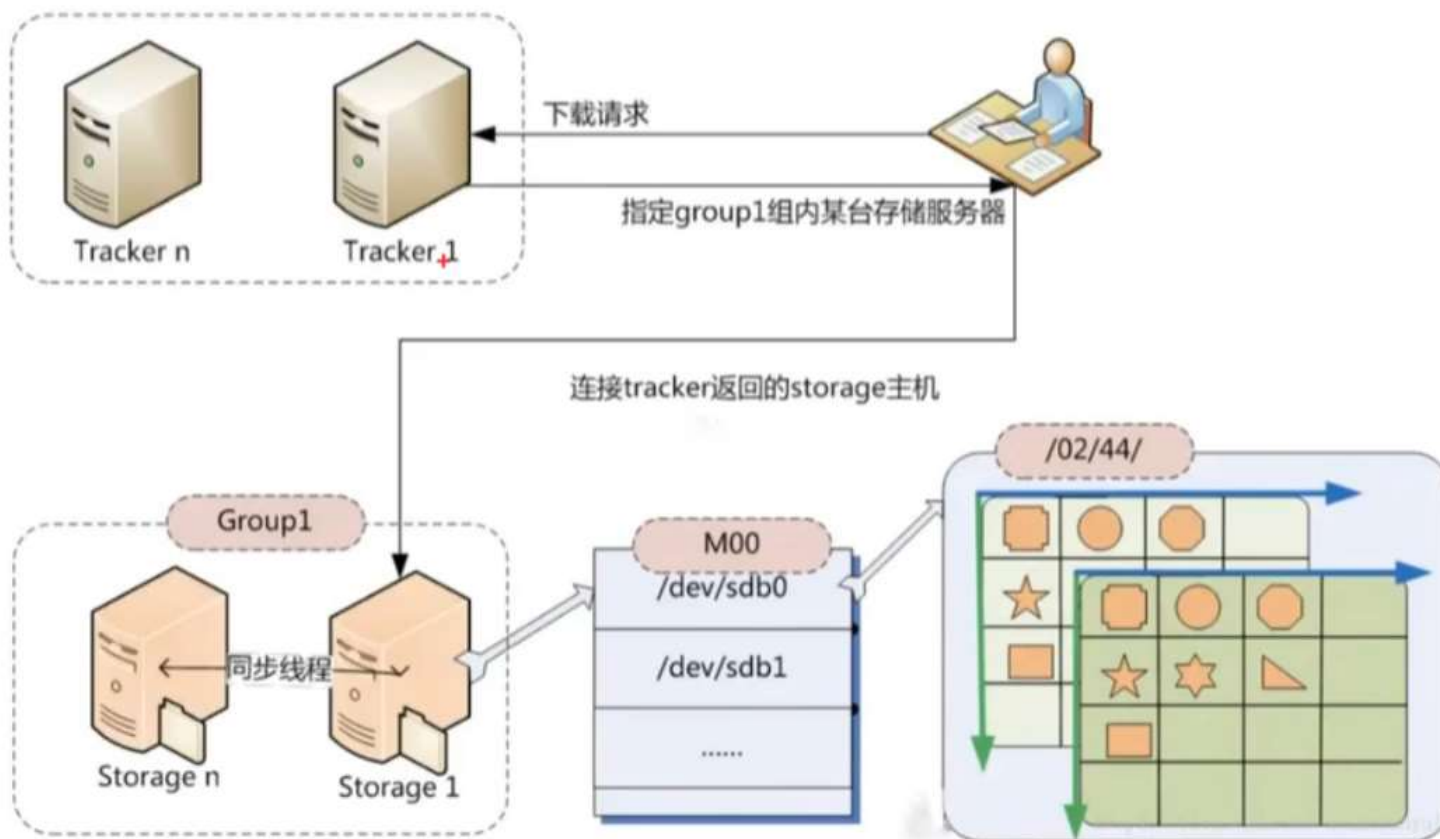
FastDFS 文件下载过程



1. Storage会定时的向Tracker安装发送心跳，告诉Tracker自己还活着，这样Fastdfs就可以工作了；
2. 客户端发送下载请求到Tracker上，Tracker查找到存储的Storage地址后返回给客户端；
3. 客户端拿到Storage地址后，去Storage上找到文件；
4. 把文件返回给客户端

FastDFS 结构——文件下载

FastDFS 文件下载过程——基于Web的HTTP请求





PART THREE

FastDFS 实际部署

FastDFS实际部署

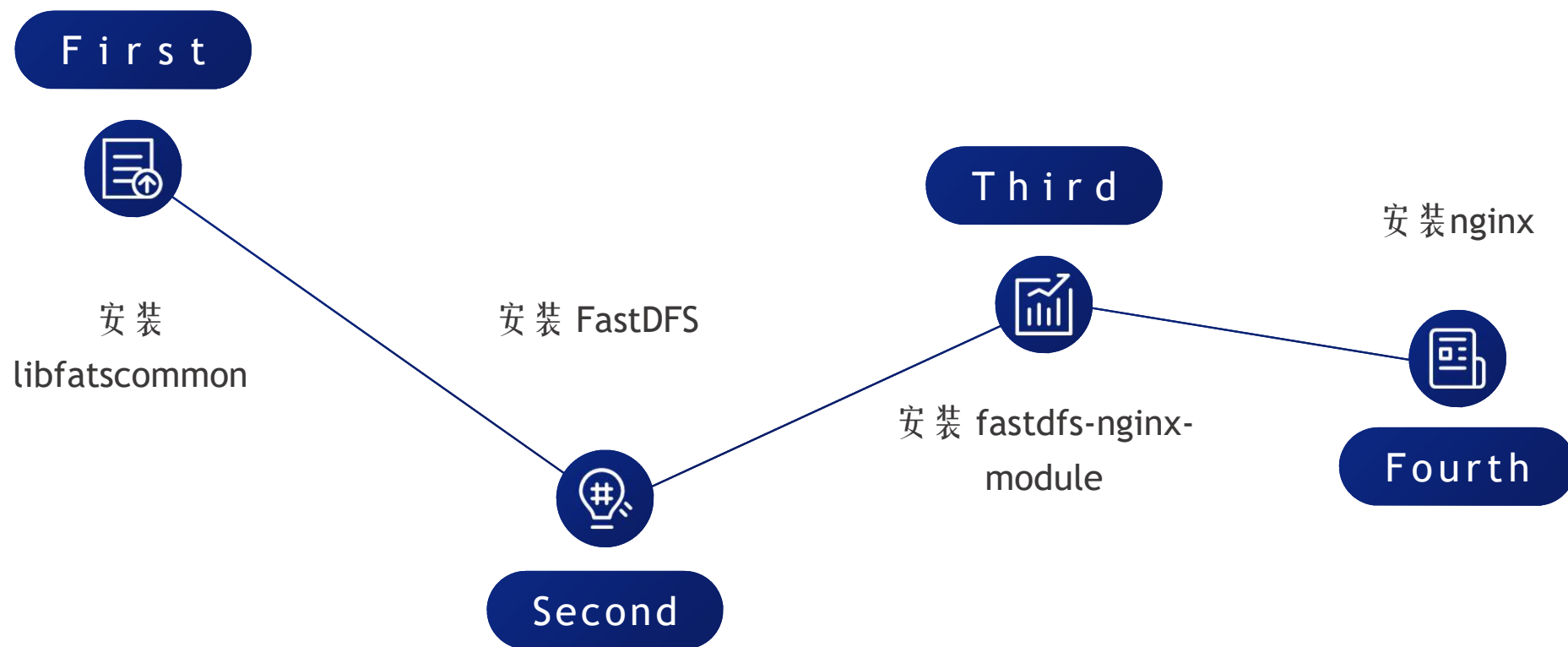
所需软件

名称	说明
Ubuntu	Ubuntu Server 22.04 LTS 64bit
libfastcommon	分离出的公用函数包
FastDFS	本体文件
fastdfs-nginx-module	FastDFS与nginx的关联包
nginx	nginx 1.18.0 (Ubuntu) 用于HTTP访问

数据位置

说明	位置
所有相关安装包	/usr/local/src
数据存储位置	/data/fastdfs

FastDFS 实际部署





PART FOUR

实 机 演 示



致 谢

- 1、[用FastDFS一步步搭建文件管理系统 - bojiangzhou - 博客园](#)
- 2、[一篇文章搞懂FastDfs\(全是干货，建议收藏\) -CSDN博客](#)
- 3、[FastDFS——从入门到入土-阿里云开发者社区](#)
- 4、[FastDFS 安装与使用-腾讯云开发者社区-腾讯云](#)
- 5、[happyfish100/fastdfs: FastDFS is an open source high performance distributed file system \(DFS\).](#)



感谢观看

相关文件请访问
starryv.top