

# nginx+fastdfs 图片系统

王非

2013-06-14

# 小型网站图片存储

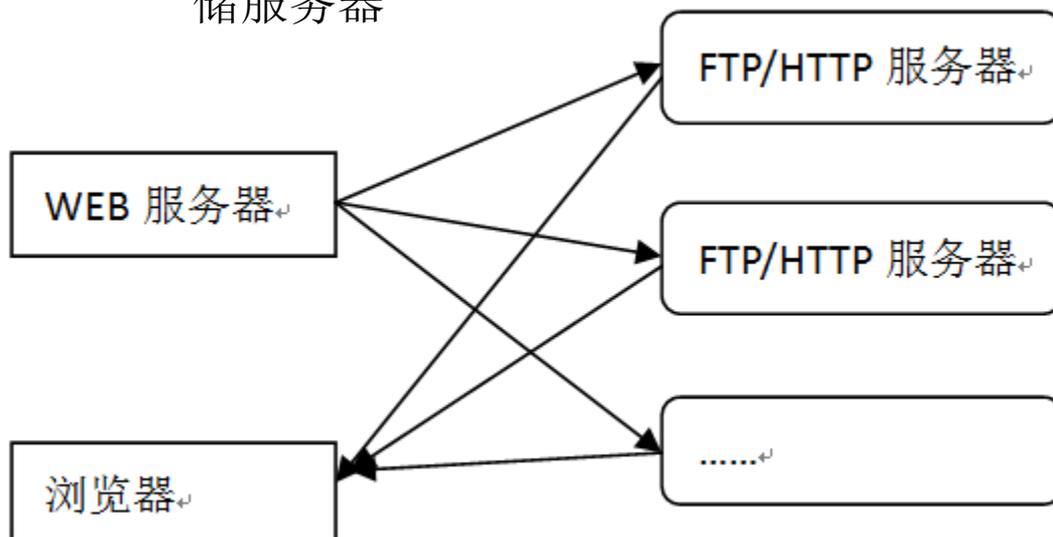
- 单机存储，所有文件放在一个目录下面，根据上传日期创建子目录，文件名使用不重复的唯一编号。
- Upload/<2013>/<0615>/<uniqid>.jpg
- 优点：简单，维护方便。
- 缺点：扩展性差，存在单点故障。

# 中型网站图片存储

- 多台服务器存储，通过FTP或者HTTP等上传到服务器。根据文件名Hash到多个服务器。
- 文件存储在远程服务器，Web服务器可以根据需要把文件同时写入两台服务器。
- 优点：简单，运维难度低。
- 缺点：扩容不方便（手动迁移数据，修改应用服务器的配置）

# 中型网站图片存储

WEB服务器通过一定的规则把文件写入存储服务器



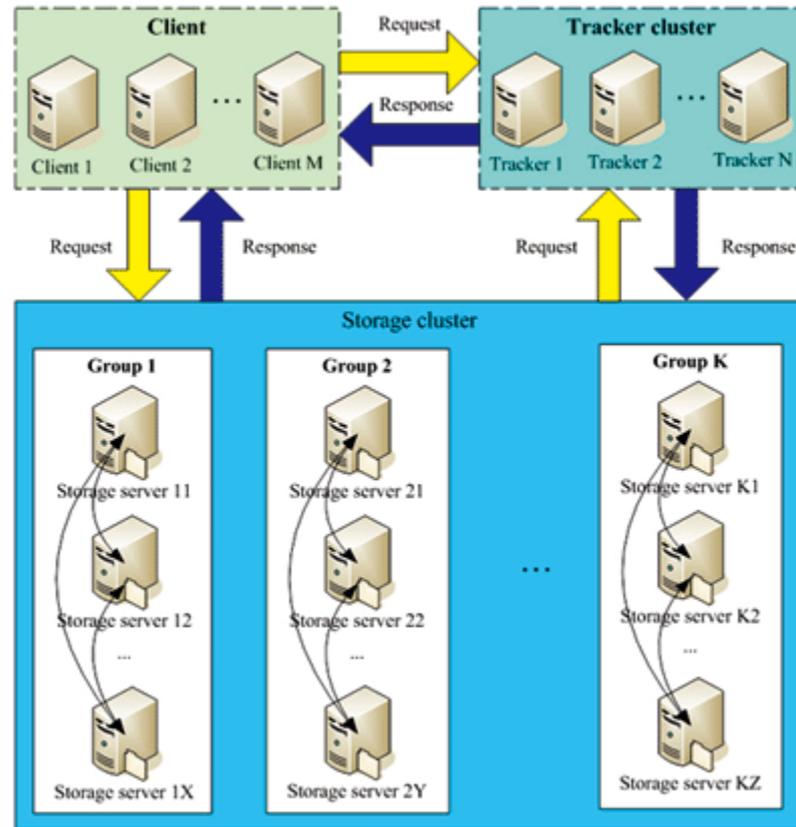
# 大型网站图片存储

- 大型网站对图片系统的需求：
  - 运行稳定
  - 平滑扩容
  - 高性能
  - 易于维护和管理
  - 松耦合

# FastDFS简介

- 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。特别适合以文件为载体的在线服务，如相册网站、视频网站等等。

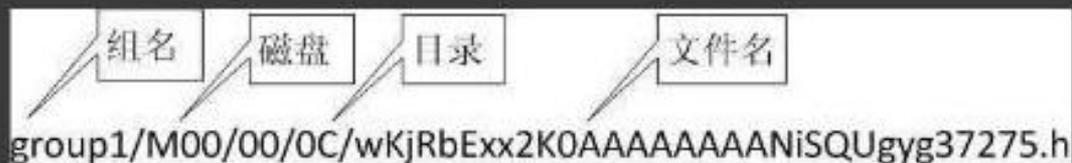
# FastDFS



# FastDFS

## 文件索引解析

- 文件的存储方式和返回



- 文件名解析

文件名包含：源头storage IP地址和文件创建时间戳、文件大小、随机数、文件拓展名。

# FastDFS

- FastDFS的Storage定时向Tracker报告心跳和状态，Tracker的数据来源于storage的报告，本身可以不存储任何数据，只有配置文件。
- 上传文件：client先请求tracker，tracker告诉client文件存入哪个storage，然后client向相应的storage发送文件，storage返回文件名。
- 下载文件：client先请求tracker，tracker告诉client向哪个storage去读，client向指定的storage读文件。

# FastDFS

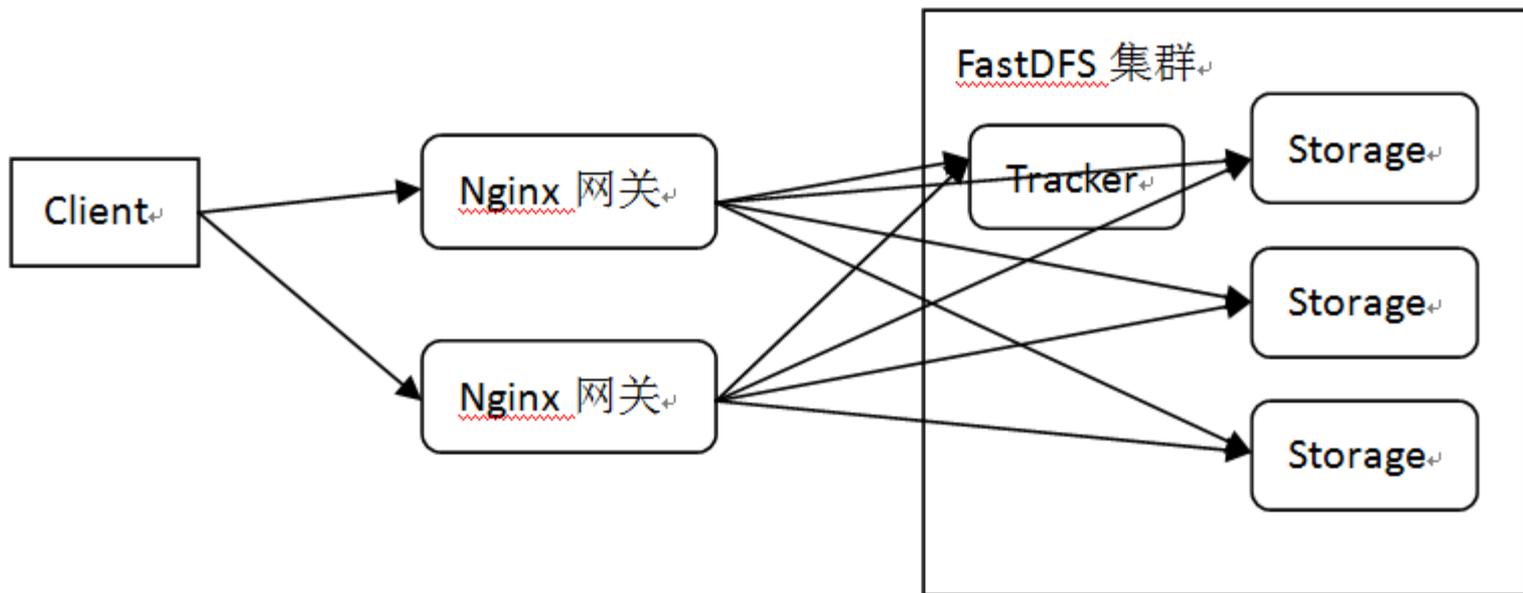
- 文件副本的存储：client上传文件到storage，只写一份，然后storage会异步想同组其他storage推数据，实现文件的多副本。
- 扩容：只要增加新的group就实现的集群的扩容，tracker默认的分配规则是那个group的剩余空间大数据就向那个group写。

# FastDFS

- 优点：
  - 平滑扩容
  - 运行稳定
  - 易于维护和管理（日志、管理工具）
  - Trunk模式可以合并小文件（变复杂）
- 缺点：
  - Php的扩展依赖fastdfs的二进制程序
  - Fastdfs的client和应用程序耦合

# 优化方案

- 搭建一个高性能通用的网关



# 优化方案

- Client使用标准的HTTP协议和网关通讯，不用安装特殊的客户端程序。
- Client和网关之间每个文件是单次请求request、response方式。Tcpdump可以清晰的看到请求的详细信息。
- 网关对client的行为记录日志，便于管理和统计
- 网关之间是无状态的非常容易扩展，也可与在网关前面加一层LVS。

# OpenResty

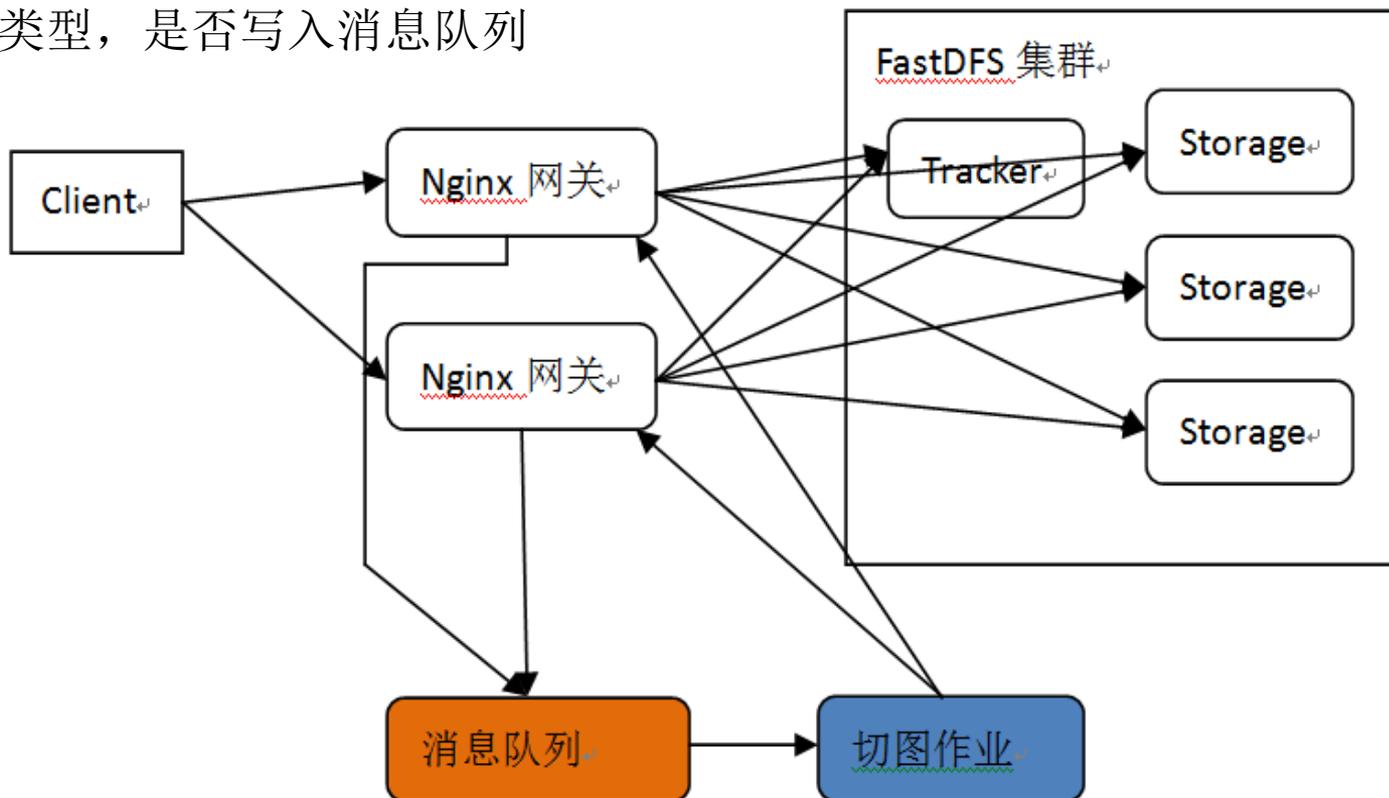
- OpenResty（也称为 ngx\_openresty）是一个全功能的 Web 应用服务器。它打包了标准的 Nginx 核心，很多的常用的第三方模块，以及它们的大多数依赖项。
- 通过众多进行良好设计的 Nginx 模块，OpenResty 有效地把 Nginx 服务器转变为一个强大的 Web 应用服务器，基于它开发人员可以使用 Lua 编程语言对 Nginx 核心以及现有的各种 Nginx C 模块进行脚本编程，构建出可以处理一万以上并发请求的极端高性能的 Web 应用。

# 网关软件包

- [标准的Nginx\(1.2.x\)](#)
- [openresty组件](#)
  - [Nginx\\_Lua\\_Module](#)
  - [lua-resty-mysql](#)
- <https://github.com/azurewang/lua-resty-fastdfs>

# 异步切图

根据HTTP的Header判断图片类型，是否写入消息队列



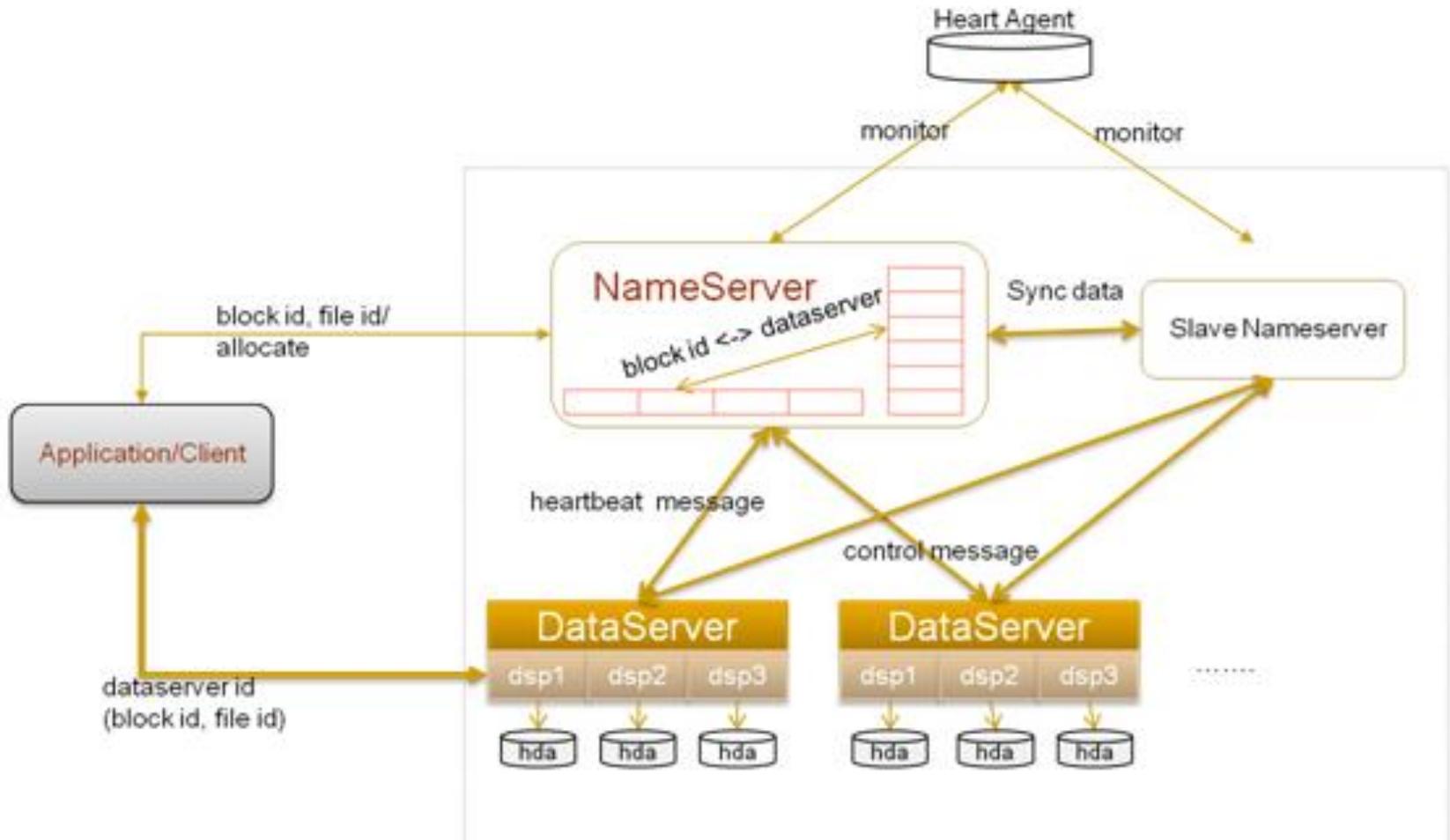
# 测试

- 使用curl测试服务
- Curl -T abc.jpg <http://192.168.1.100/upload>
- group1/M00/11/0E/rBEHaFDZHiEIAAAAAAKY3nnMJn4AADZ-QDEyNEAApj2654.jpg
  
- Curl -H 'Image-Type:1' -T abc.jpg <http://192.168.1.100/upload>
- group1/M00/11/0E/rBEHaFDZHiEIAAAAAAKY3nnMJn4AADZ-QDEyNEAApj2654.jpg
- 同时会向消息队列写入消息
  
- Curl -XDELETE  
<http://192.168.1.100/del?filename=group1/M00/11/0E/rBEHaFDZHiEIAAAAAAKY3nnMJn4AADZ-QDEyNEAApj2654.jpg>
  
- Curl -T abc.jpg  
[http://192.168.1.100/upload\\_slave?filename=xxxxxx&ext=300x300](http://192.168.1.100/upload_slave?filename=xxxxxx&ext=300x300)
- group1/M00/11/0E/rBEHaFDZHiEIAAAAAAKY3nnMJn4AADZ-QDEyNEAApj2654\_300x300.jpg

# TFS

- TFS (Taobao !FileSystem) 是一个高可扩展、高可用、高性能、面向互联网服务的分布式文件系统，主要针对海量的非结构化数据，它构筑在普通的Linux机器集群上，可为外部提供高可靠和高并发的存储访问。TFS为淘宝提供海量小文件存储，通常文件大小不超过1M，满足了淘宝对小文件存储的需求，被广泛地应用在淘宝各项应用中。它采用了HA架构和平滑扩容，保证了整个文件系统的可用性和扩展性。同时扁平化的数据组织结构，可将文件名映射到文件的物理地址，简化了文件的访问流程，一定程度上为TFS提供了良好的读写性能。

# TFS



# TFS

- TFS有点：
  - 高性能
  - 平滑扩容
  - 自动数据迁移
- TFS缺点：
  - 安装环境依赖组件多
  - 对操作系统和GCC版本有要求。
  - 安装后的二进制文件非常大
  - 只能用在ext3和ext4
  - 运维难度高
  - Nameserver要求内存高，单个集群容量有上限。

