

大型网站及其架构演进过程



什么是大型网站



什么是大型网站

亲, 请登录 免费注册 手机逛淘宝

我的淘宝 购物车 收藏夹 商品分类 卖家中心 联系客服 网站导航



淘宝网 Taobao.com

宝贝 天猫 店铺

男装外套春新款

搜索

高级搜索

手机淘宝



新款连衣裙 四件套 潮流T恤 时尚女鞋 短裤 半身裙 男士外套 墙纸 行车记录仪 新款男鞋 耳机 更多

主题市场

- 女装 / 男装 / 内衣
- 鞋靴 / 箱包 / 配件
- 童装玩具 / 孕产 / 用品
- 家电 / 数码 / 手机
- 美妆 / 洗护 / 保健品
- 珠宝 / 眼镜 / 手表
- 运动 / 户外 / 乐器
- 游戏 / 动漫 / 影视
- 美食 / 生鲜 / 零食
- 鲜花 / 宠物 / 农资
- 房产 / 装修 / 建材
- 家具 / 家饰 / 家纺
- 汽车 / 二手车 / 用品
- 办公 / DIY / 五金电子
- 百货 / 餐厨 / 家庭保健
- 学习 / 卡券 / 本地服务

天猫 聚划算 天猫超市 淘抢购 电器城 司法拍卖 中国质造 特色中国 飞猪旅行 智能生活 苏宁易购

北鼻快买

春夏男装 满149减10

FASHION IS A TERM USED TO DESCRIBE THE STYLE OF CLOTHING WORN BY MOST PEOPLE

运动潮流

春夏强货 迅速来袭

Hi! 你好

领淘金币抵钱 或去 会员俱乐部

登录 注册 开店

公告 规则 论坛 安全 公益

人民日报: 阿里作为新实体经济代表正迅速崛起
阿里掀起打假人民战争 天价学区房现身闲鱼

充值费	游戏	旅行	车险
彩票	电影	酒店	理财
找服务	音乐	水电煤	火车票

阿里APP 更多



内衣配件 19.9元包邮上新

天猫必逛 热门品牌,天天上天猫! 6/6 今日热卖

名店潮搭

新品不断

99%适用雨刷

68元得相同2对

A.O.史密斯 全店抢0元试用

囤货一族 会囤会生活 人气1602437

吃货的后裔 翻滚吧美食君 人气1130667

摩登主妇 拒做黄脸婆 人气5670252

好孕妈咪 一级保护动物 人气743803

单身汪汪汪 有ta不孤单 人气91405

我常逛的 时尚 品质 特色 实惠 猜你喜欢 反馈

淘宝头条 更多

我常逛的

什么是大型网站

成都: 12°C 轻度 102 | 宝箱 换肤 消息

糯米 新闻 hao123 地图 视频 贴吧 loveyouwhyso 设置 更多产品

 百度一下

我的关注 推荐 导航 视频 购物

以下信息根据您的兴趣推荐



外交部：将不参加今年“禁止核武器条约”谈判

中华网 03-21 04:41 热点

英媒：中国城市化的规模惊人 有102个人口超百万城市

凤凰网 03-21 02:03 热点



美国务卿“缓和”口气 中美就朝核问题达成

实时热点

换一换

- 安信支持率暴跌 新 中国维和警察回国 新
- 富豪洛克菲勒去世 新 中国女子在韩遇害 新
- 68万被摸10次脸 新 离异少妇骗大学生
- 陈赫前妻许婧出道 一家六口缅甸被杀
- 女儿趁母做饭跳楼 吴亦凡西装革履
- 韩星转战东南亚 女学生裹被电电竞
- 看电视剧入戏太深 托养中心死20人
- 爬上山头脱衣轻生 女富豪海选丈夫

音乐

什么是大型网站



成都 小雨 16~11°C 3月21日 周二 二月廿四 星座运势

loveyouwhyso 网盘 邮箱 手机版 换肤 猜金像奖最佳导演

今天有雨,记得带伞



网页 你想搜的

百度一下

睡好觉成青年奢望

软件开发需要学什么 学java 网站开发 代理记账 北大青鸟课程 财务代理公司 香港便宜住宿

- 首页 电视剧 最新电影 新闻头条 八卦娱乐 军事热点 热门游戏 小游戏 居家9.9 特价旅游 成都生活
- 权威推荐 人民网 新华网 央视网 国际在线 中国日报网 中国网 中国经济网 光明网 央广网 求是网 中国青年网 中国网信网



- 电视剧 | 综艺 游戏 | 小游戏
- 电影 | 直播 动画 | 漫画
- 新闻 | 军事 旅游 | 音乐
- 彩票 | 竞猜 搞笑 | 小说
- 特价 | 手机 股票 | 理财

这不只是富人的专利 更是门生活艺术

天气 星座 网盘 快递 违章

推荐 社会 娱乐 军事 体育 +



- 国家中心城市抢位战激烈 8城争4名额
- 沙特国王来一趟北京把奔驰都租空了

- 百度·贴吧 新浪·微博 搜狐·热点 腾讯 网易 优酷网
- 天猫 凤凰网 淘宝网 免费游戏 4399游戏 hao123影视
- 京东商城 苏宁易购 爆款五折 今日特价 周二居家 体育·NBA 汽车之家 聚划算
- 东方财富·理财 58同城 房天下 携程旅行网 37游戏 12306·旅游
- 驴妈妈旅游 赶集网 途牛旅游网 超值礼包 去哪儿 头条新闻 哔哩哔哩
- 苏宁手机馆 彩票·双色球 百度地图 亚马逊 爱淘宝 QQ音乐
- 新浪新闻 中关村在线 东方财富 软件中心 中华军事 去哪儿网

聚划算 涨停股推荐 一刀999999 春茶上新 软件开发需要学什么 网站开发 香港便宜住宿 北大青鸟课程 北大青鸟培训课程 代理记账

- 游戏 2017版传奇 一刀999999 蓝月传奇 热血传奇 一刀99999级 传奇单机版 变态999级 换一换
- 热门 60天理财 懒人网活期 微信领牛股 股市高手群 成都二手车 晚上赚外快 在家也赚钱 换一换
- 优选 中小学课程 耐克折扣区 机械设计 泥水分离机 工业品采购 大牌cc霜 二手苹果6 换一换
- 健康 除眼袋方法 去眼袋哪好 除眼袋多钱 检测怀孕 颈椎病治疗 如何除雀斑 肿瘤医院 换一换

电视剧 电影 新闻 娱乐 军事 小游戏 特价 购物 综艺 软件 学习 理财

推荐 教你理财 同城交友 美食菜谱 百度网盘 找工作 电影票预定 四川人事考试 更多>>>



- 视频
- 游戏
- 购物
- 笑话
- 头条新闻

理

网站是用来访问的，访问量要大

大型网站应该有海量的数据

大型网站应该有复杂业务处理的能力



大型网站的架构演进

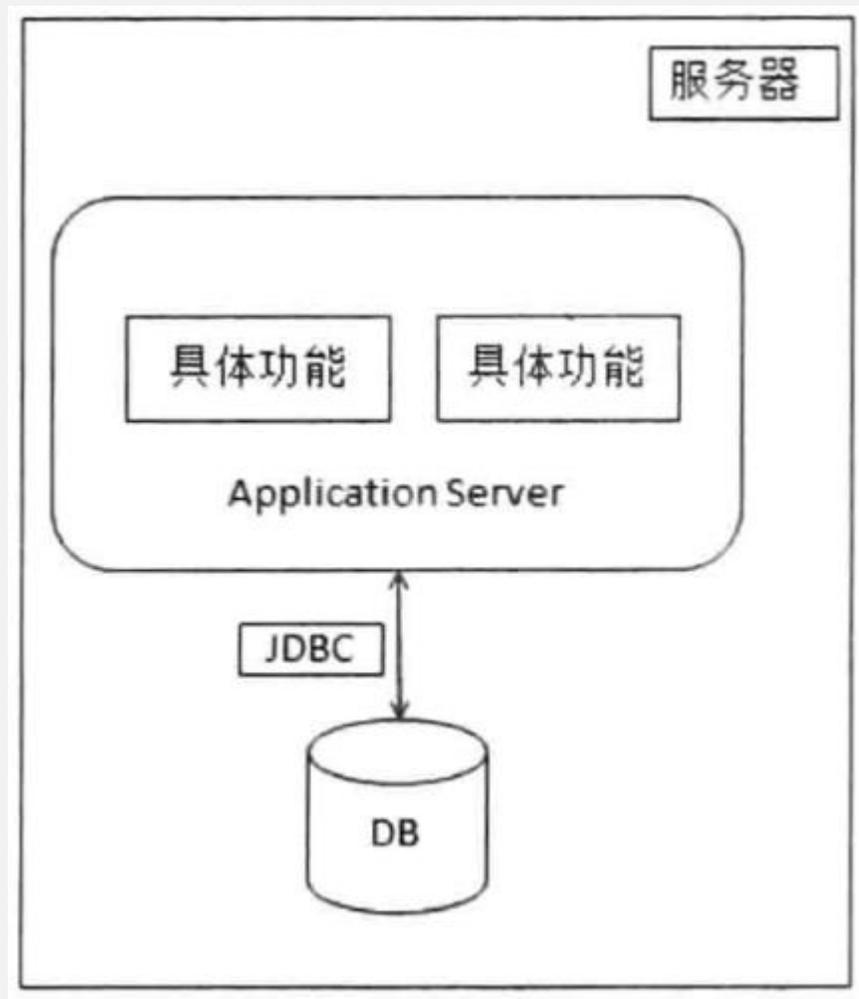
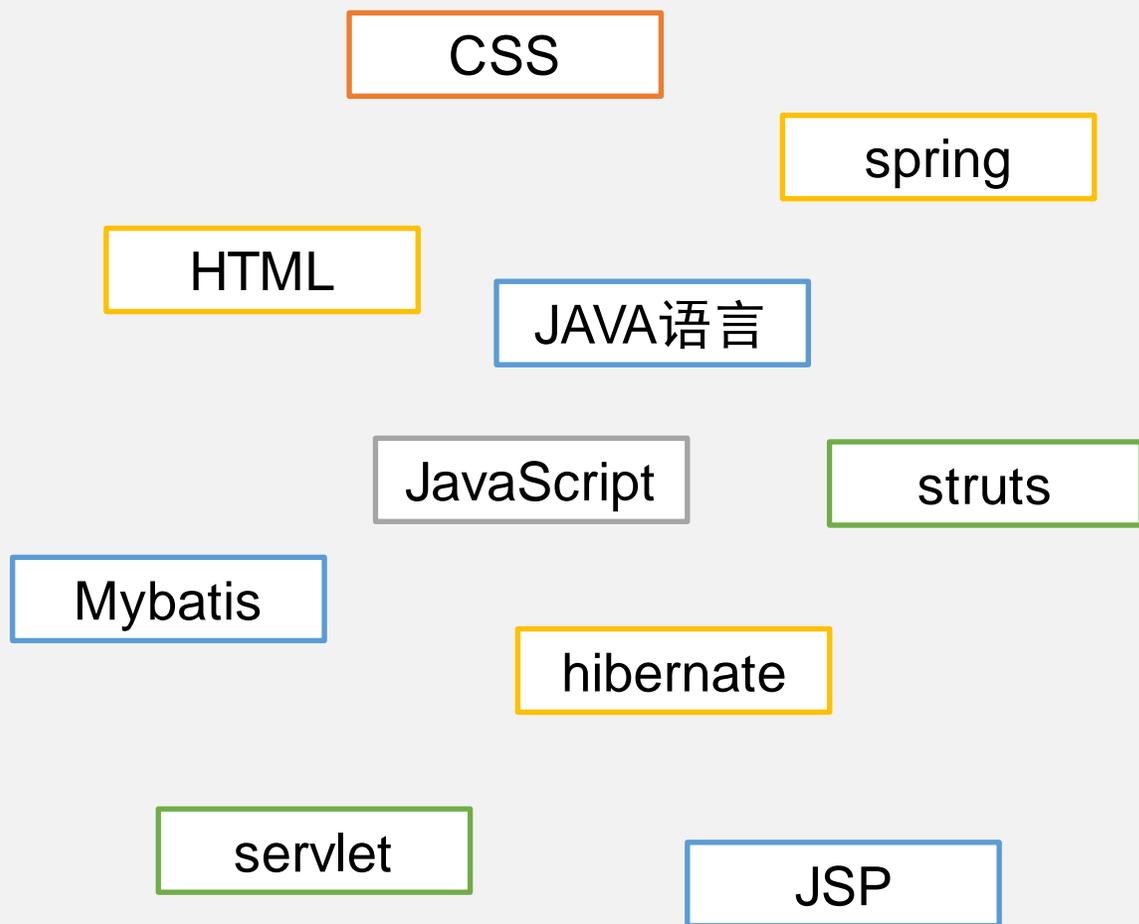


大型网站的架构演进

利用JAVA技术和单机来构建网站



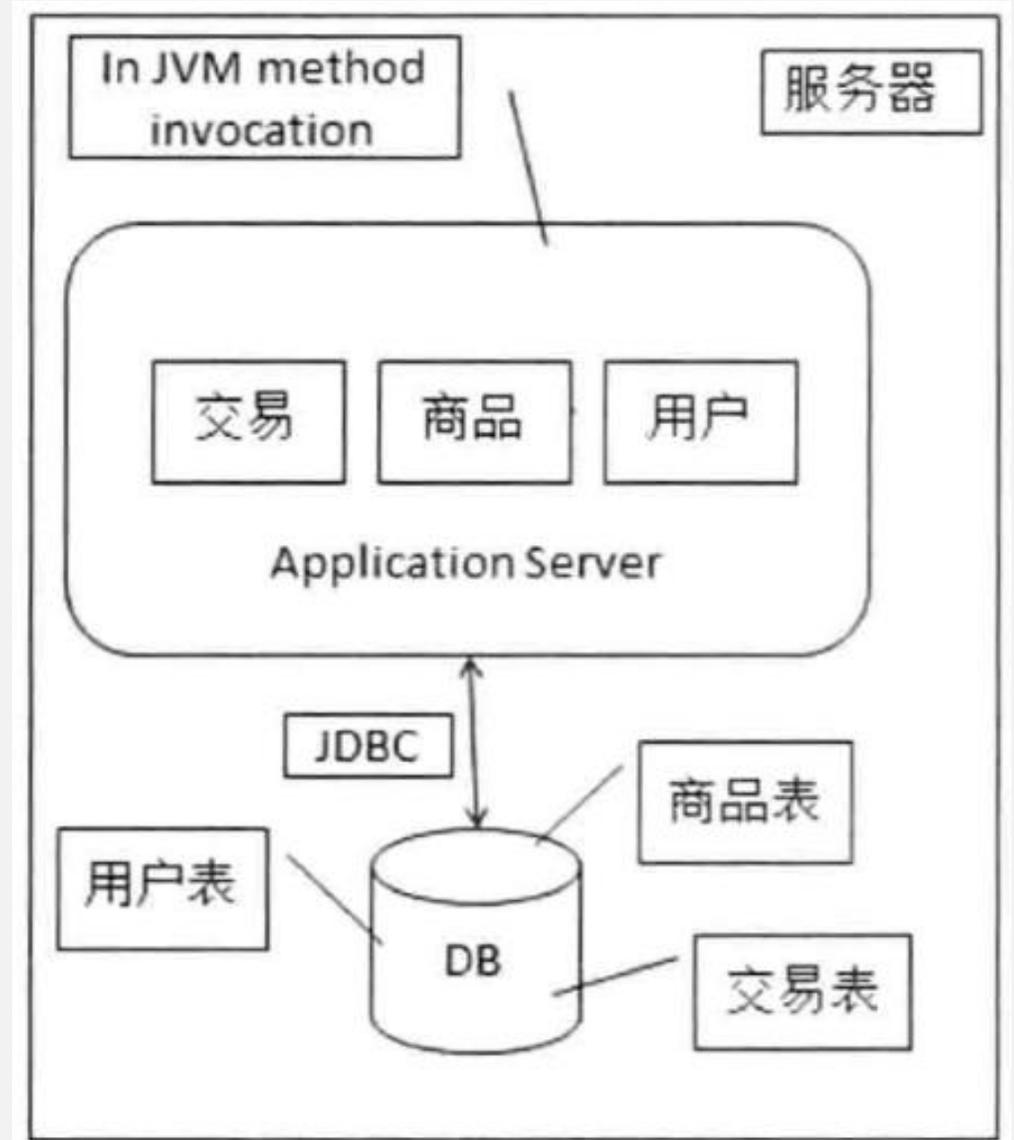
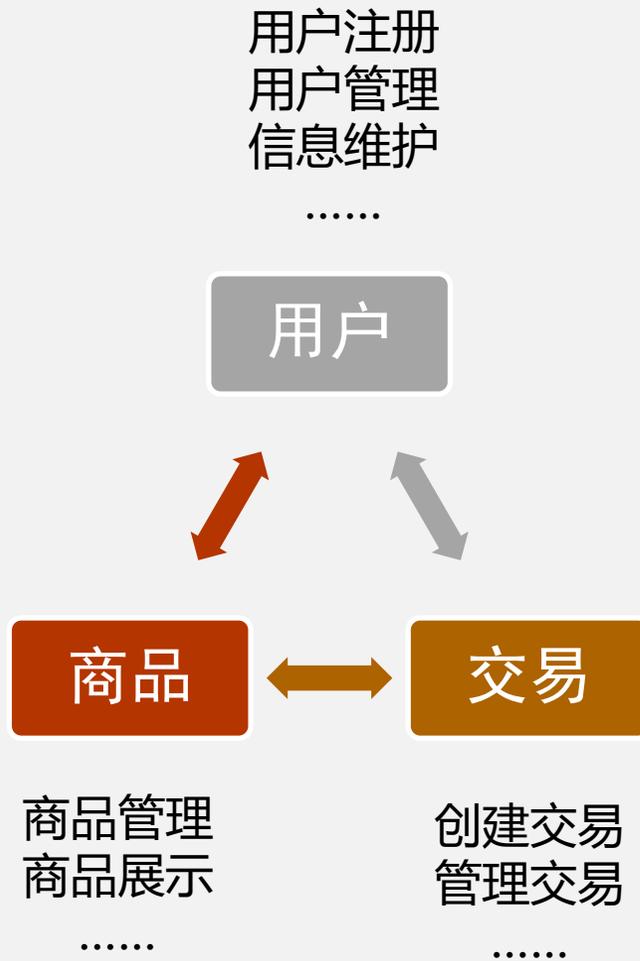
利用**JAVA**技术和单机来构建网站



计算

存储

整理



大型网站的架构演进

单机负载告警，数据库与应用分离



单机负载告警，数据库与应用分离

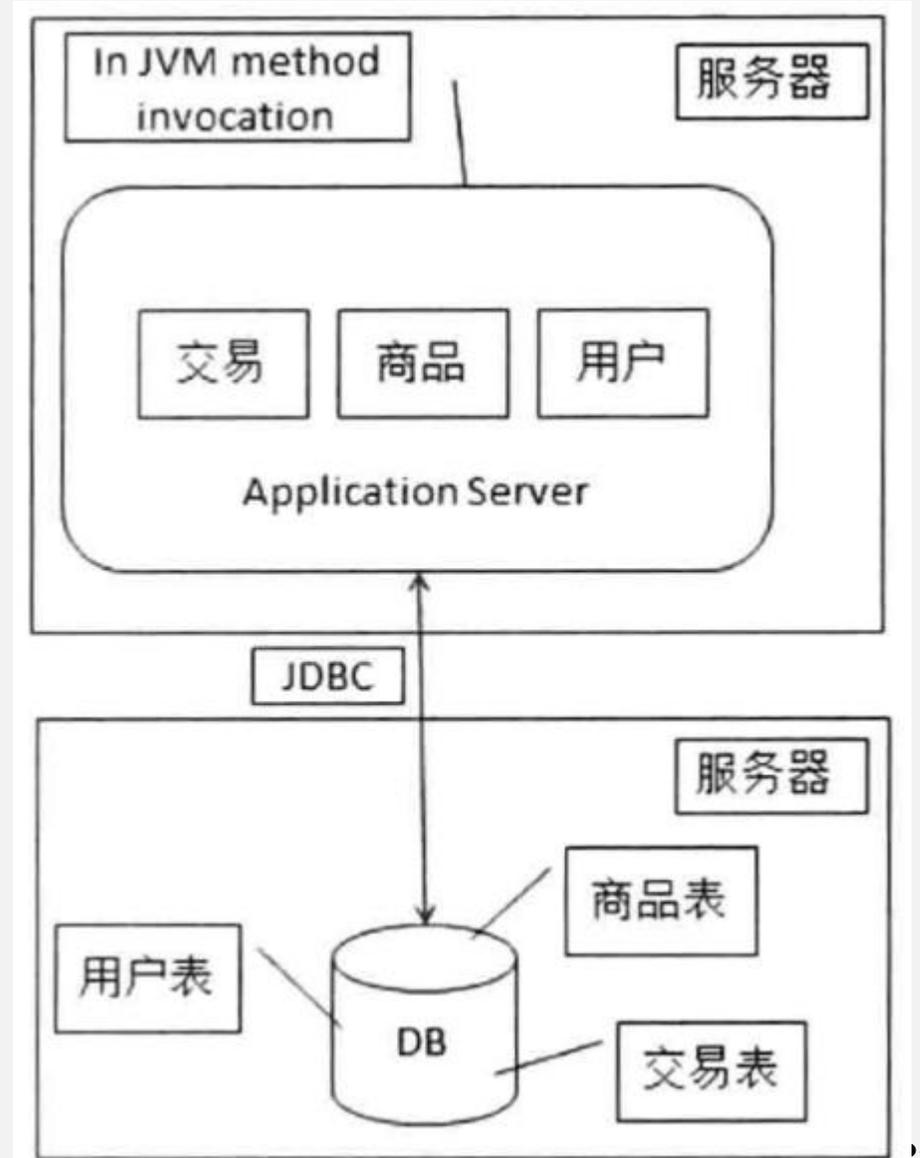
用户访问量增加

网站数据量增加

网站响应变慢

系统宕机

数据库与应用分离



整理



大型网站的架构演进

应用服务器负载告警，应用服务器走向集群

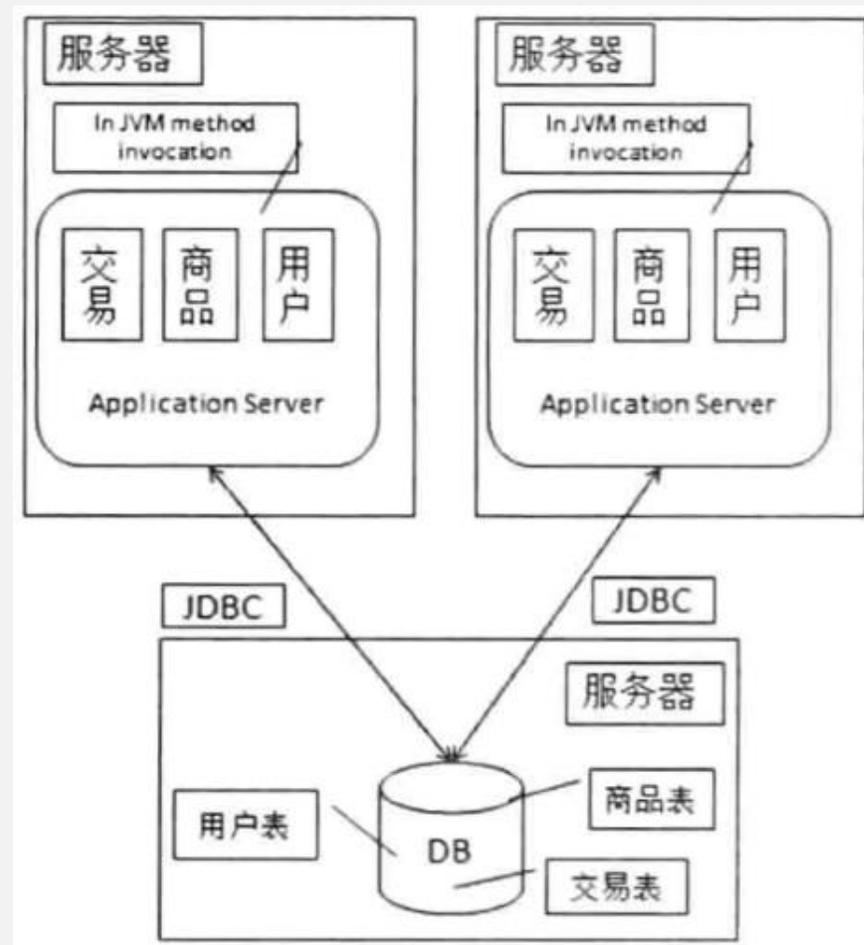


应用服务器负载告警，应用服务器走向集群

用户访问量持续增加

应用服务器达到性能瓶颈

应用服务器走向集群



? 最终用户访问哪一台服务器?

引入负载均衡设备

平衡算法

平衡算法设计的好坏直接决定了集群在负载均衡上的表现，设计不好的算法，会导致集群的负载失衡。一般的平衡算法主要任务是决定如何选择下一个集群节点，然后将新的服务请求转发给它。

轮转法

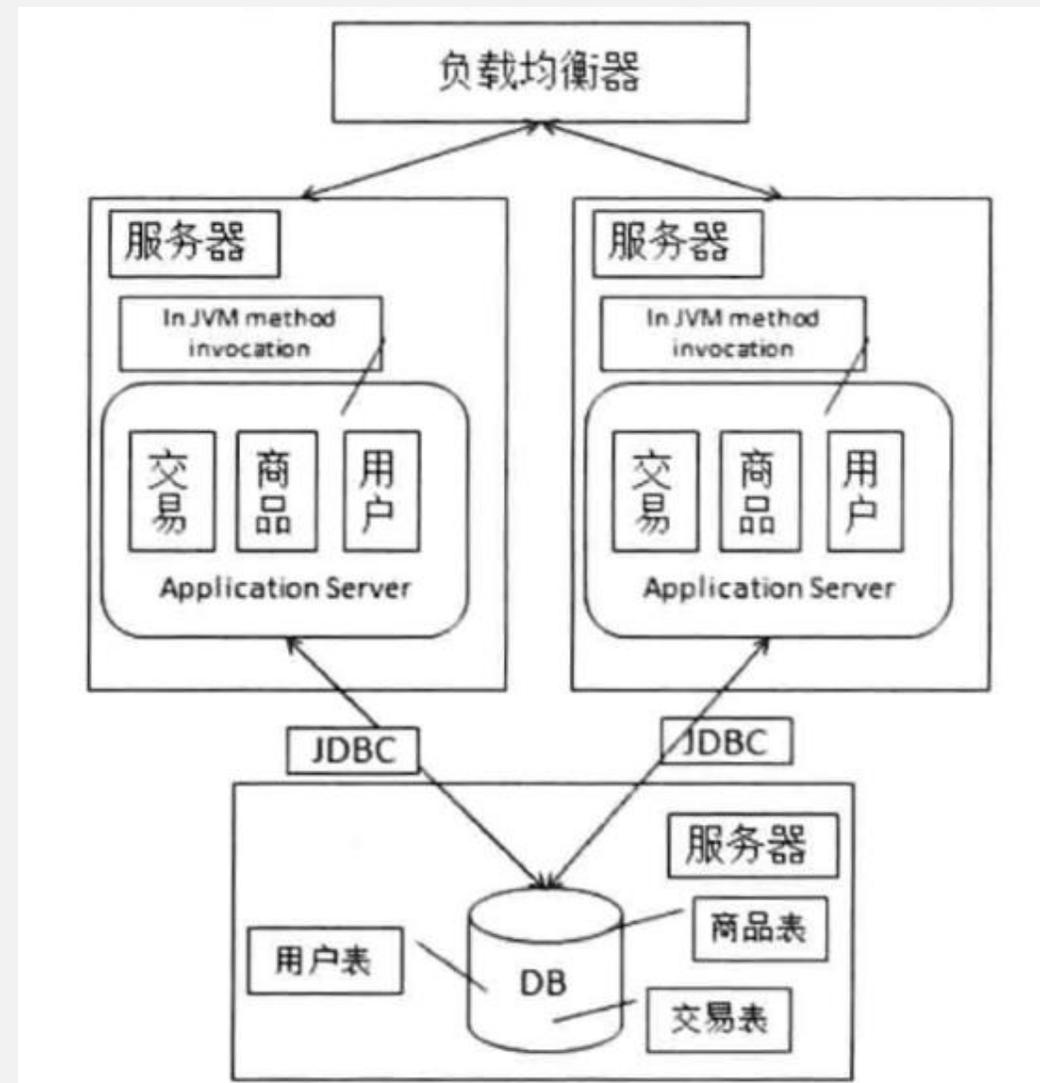
散列法

最少连接法

最低缺失法

最快响应法

加权法

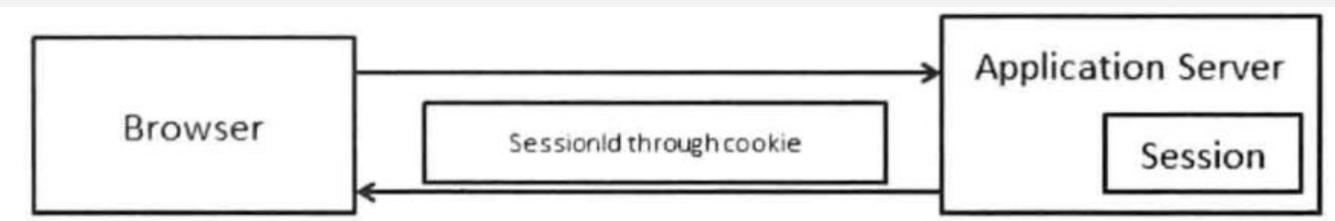
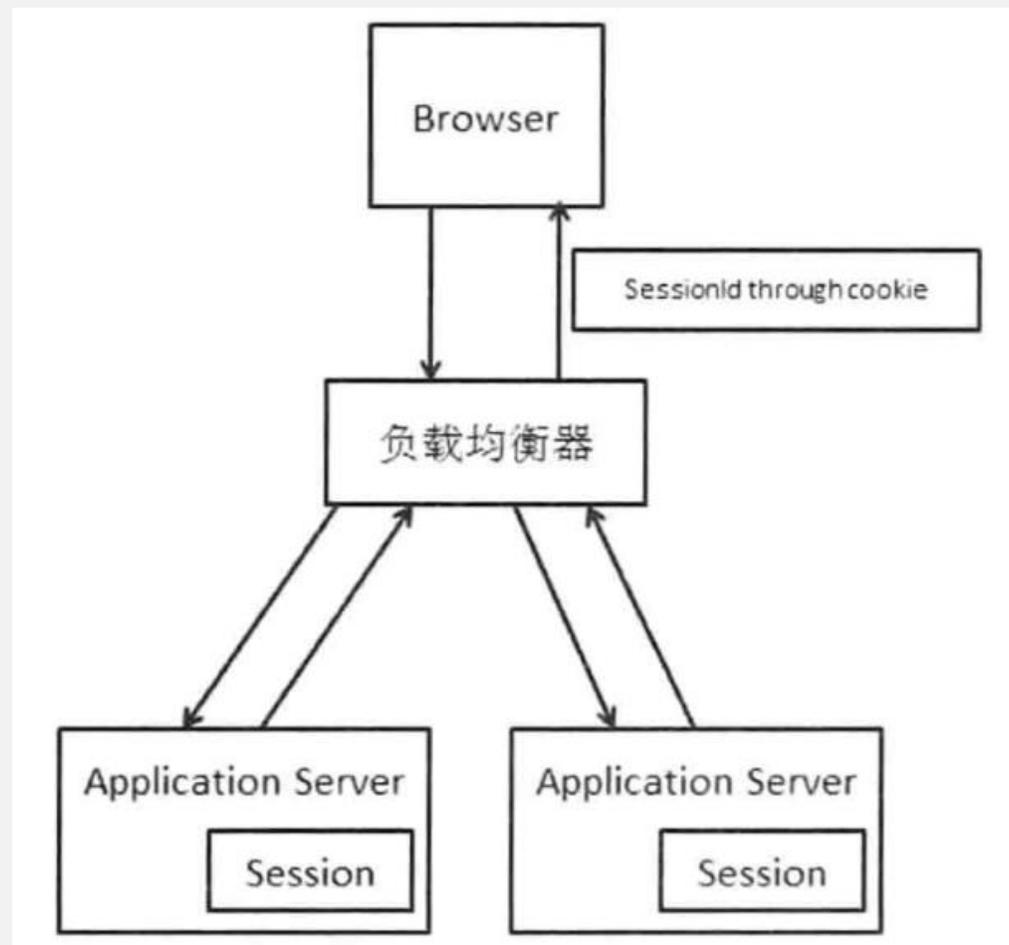


什么是session

http协议本身无状态，需要基于HTTP协议支持会话状态 (session state) 的机制，实现WEB服务器从多次单独的HTTP请求中看到“会话”，也就是知道请求来源于那个会话。

实现方式：

在会话开始时，分配一个唯一会话标识 (sessionId)，通过Cookie把这个标识告诉浏览器，以后每次请求时，浏览器会带上这个会话标识告诉Web服务器请求属于哪个会话，在Web服务器上，各个会话有独立的存储，保存不同会话的信息。如果遇到禁用Cookie的情况，就把这个会话标识放到URL参数中。



问题随即而来

1. 会话保存在单机上
2. 如果第一次请求访问左边服务器，如果不做处理，无法保障每次请求都落在同一服务器

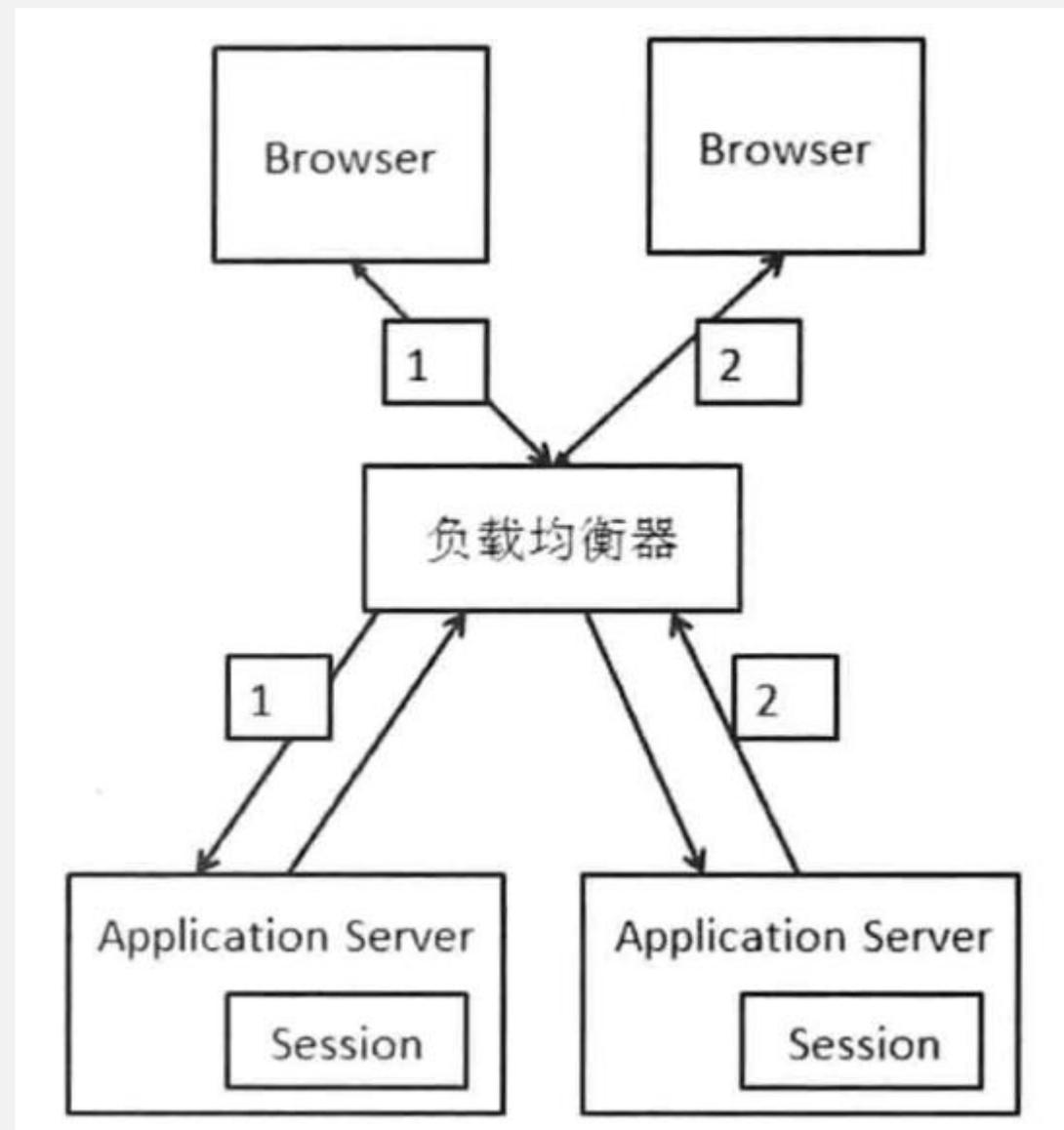
Session Sticky

保证同一会话的请求都在同一个Web服务器上处理

依赖负载均衡器根据每次请求的会话标识来进行请求转发

缺陷

1. 如果一台Web服务器宕机或重启，会话数据丢失；
2. 负载均衡器变成了一个有状态的节点，要讲会话保存到具体Web服务器的映射。和无状态节点相比，内存消耗会更大，容灾麻烦。

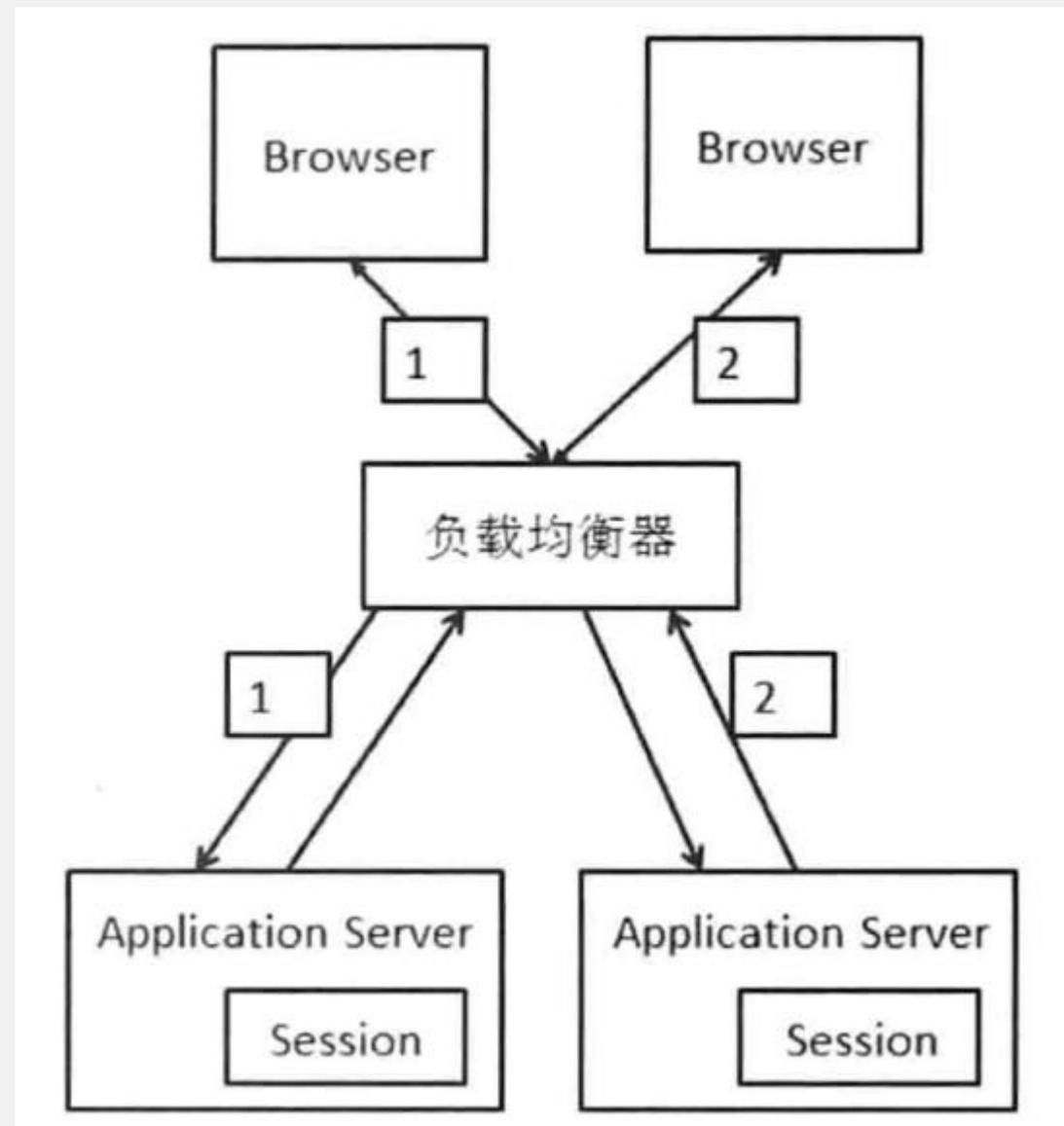


Session Replication

Web服务器之间增加会话数据的同步，保证Web服务器之间Session数据一致

缺陷

1. 同步Session数据造成了网络带宽的开销。
2. 每台Web服务器都保存所有的Session数据，如果整个集群的Session数很多的话，每台机器用于保存Session数据的内容占用会很严重

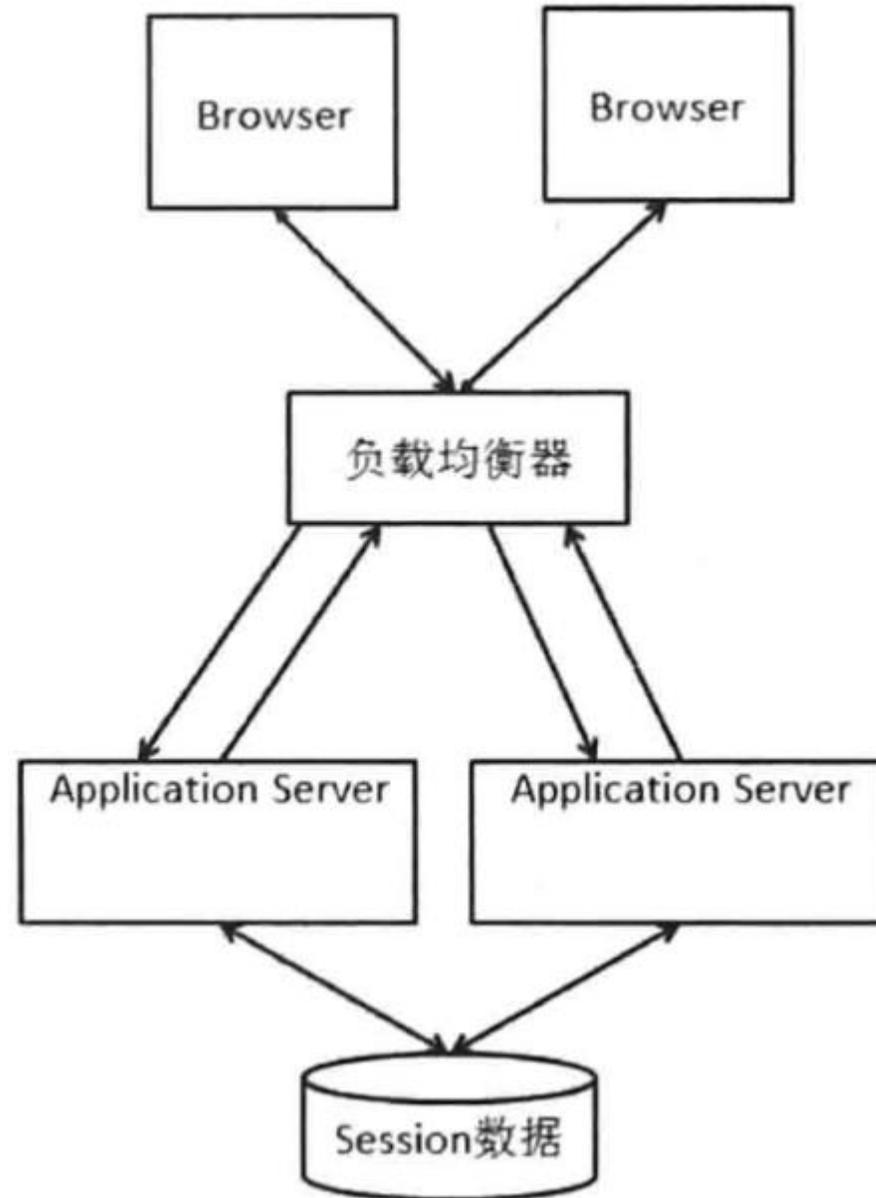


Session 数据集中存储

把Session数据集中存储起来，不同Web服务器从同样地方获取Session

缺陷

1. 读写Session数据引入了网络操作，故而造成延迟和不稳定性；
2. 如果集中存储Session的机器或者集群有问题，会影响应用

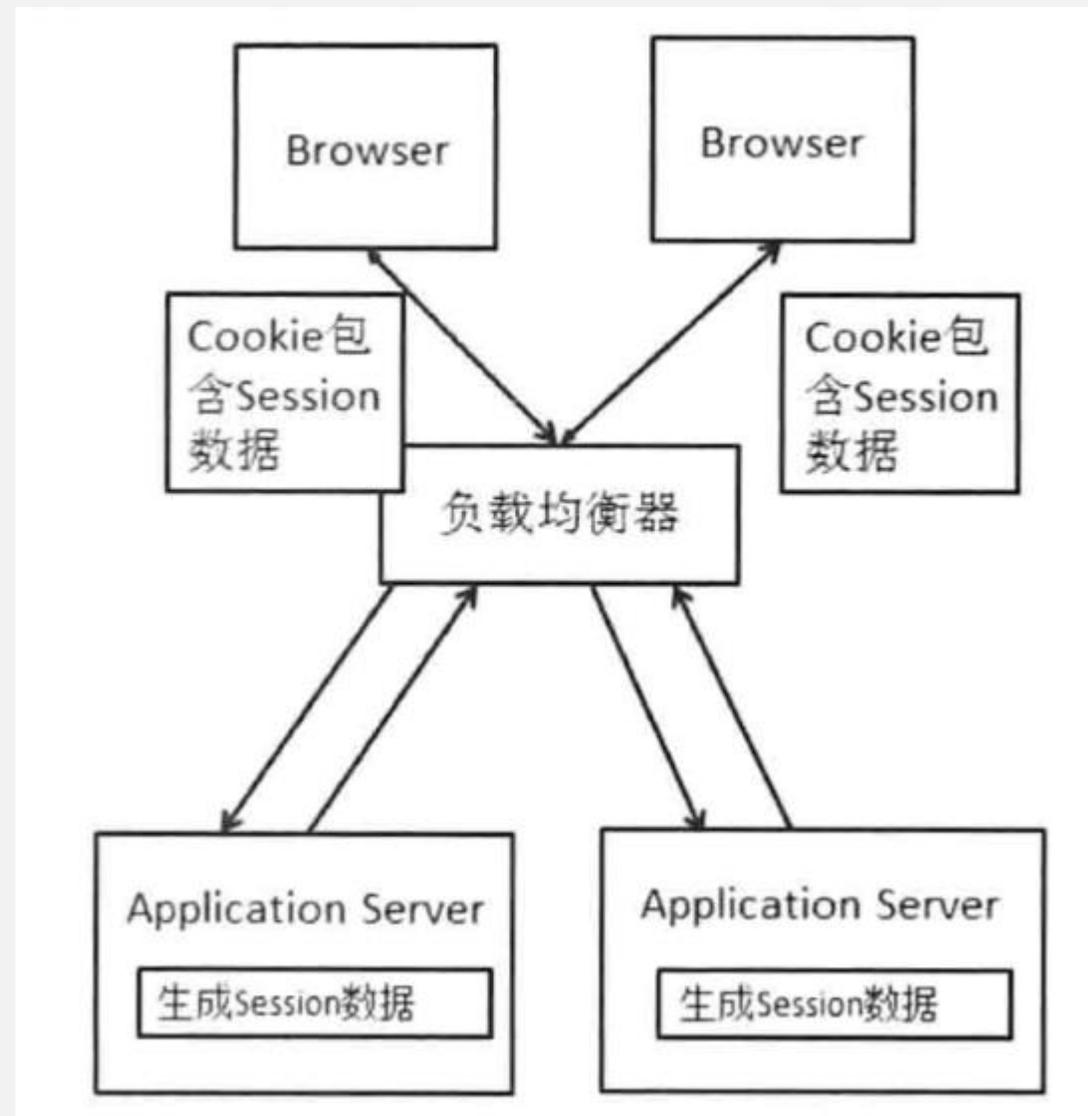


Cookie Based

将Session数据放在Cookie中，然后在Web服务器上从Cookie中生成对应的Session数据

缺陷

1. Cookie长度限制;
2. 安全性;
3. 带宽消耗;
4. 性能影响;



大型网站的架构演进

数据读压力变大，读写分离吧

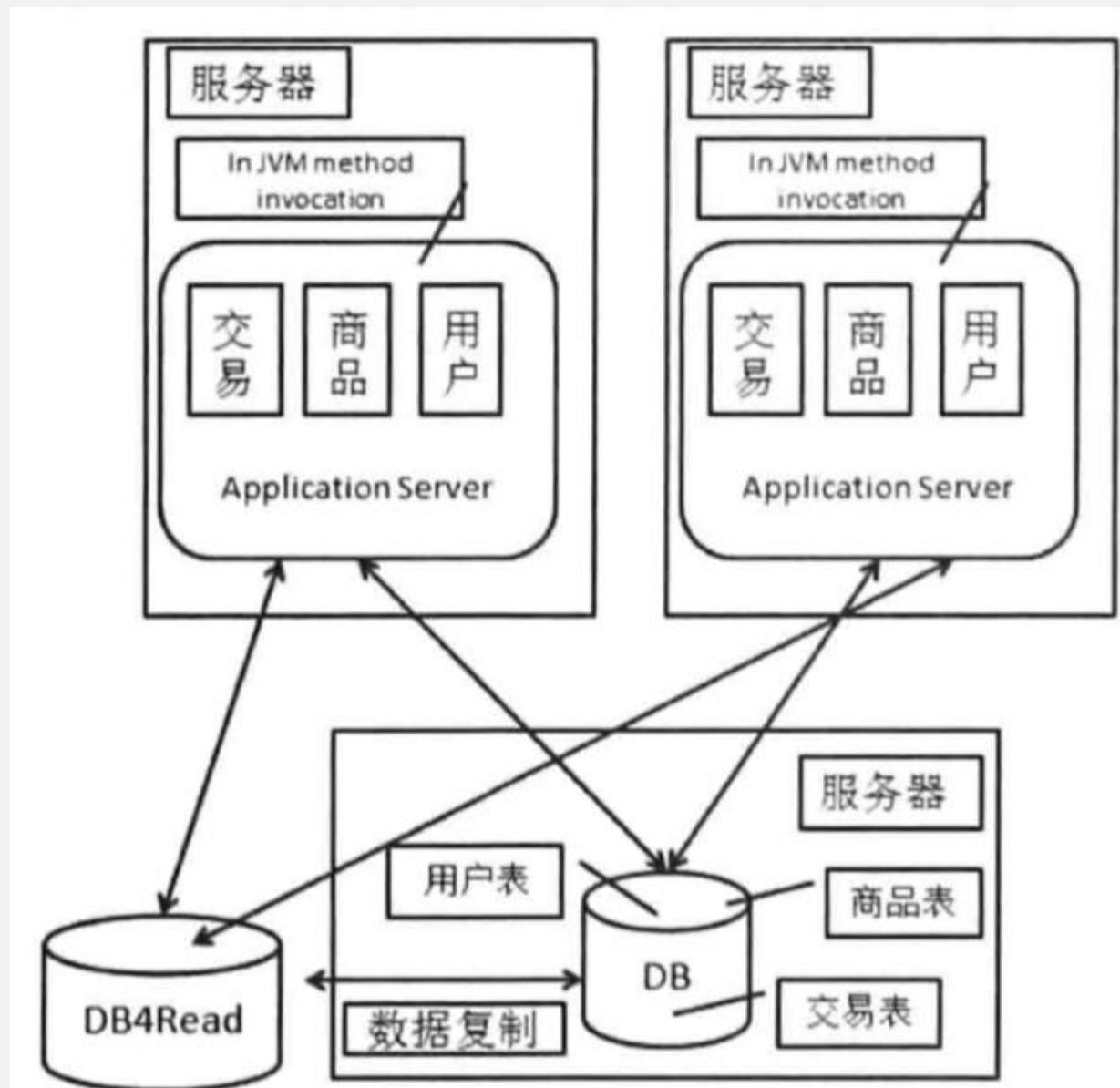


随着业务的发展，数据量和访问量持续增长

对于大型网站来说，不少业务是读多写少的

问题

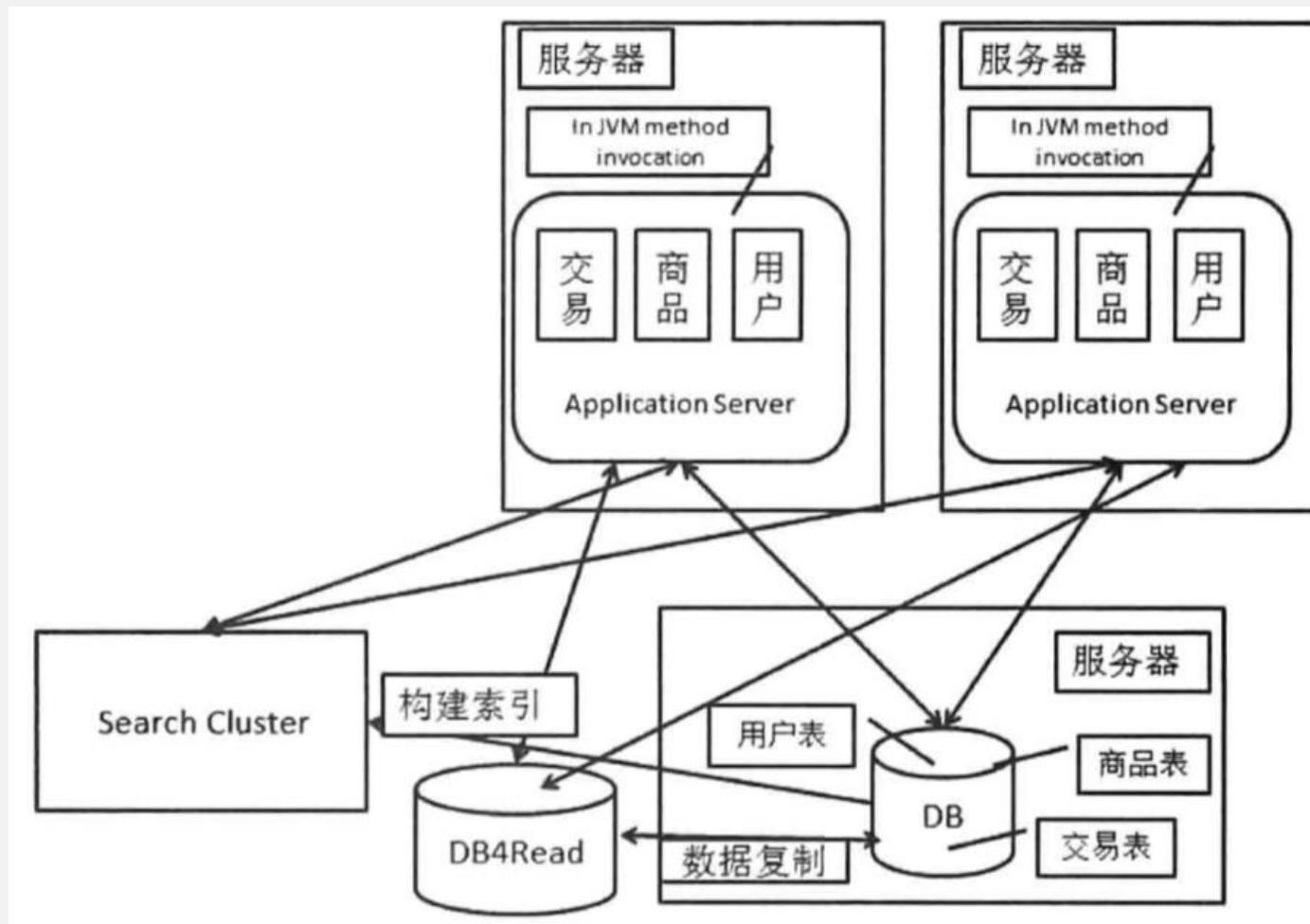
1. 数据复制问题；
2. 短期的数据不一致问题；
3. 应用对于数据源的选择问题



搜索引擎可以看成是一个读库

根据被搜索的数据来构建搜索索引

随着被搜索数据的变化，索引也要进行改变



数据缓存

页面缓存

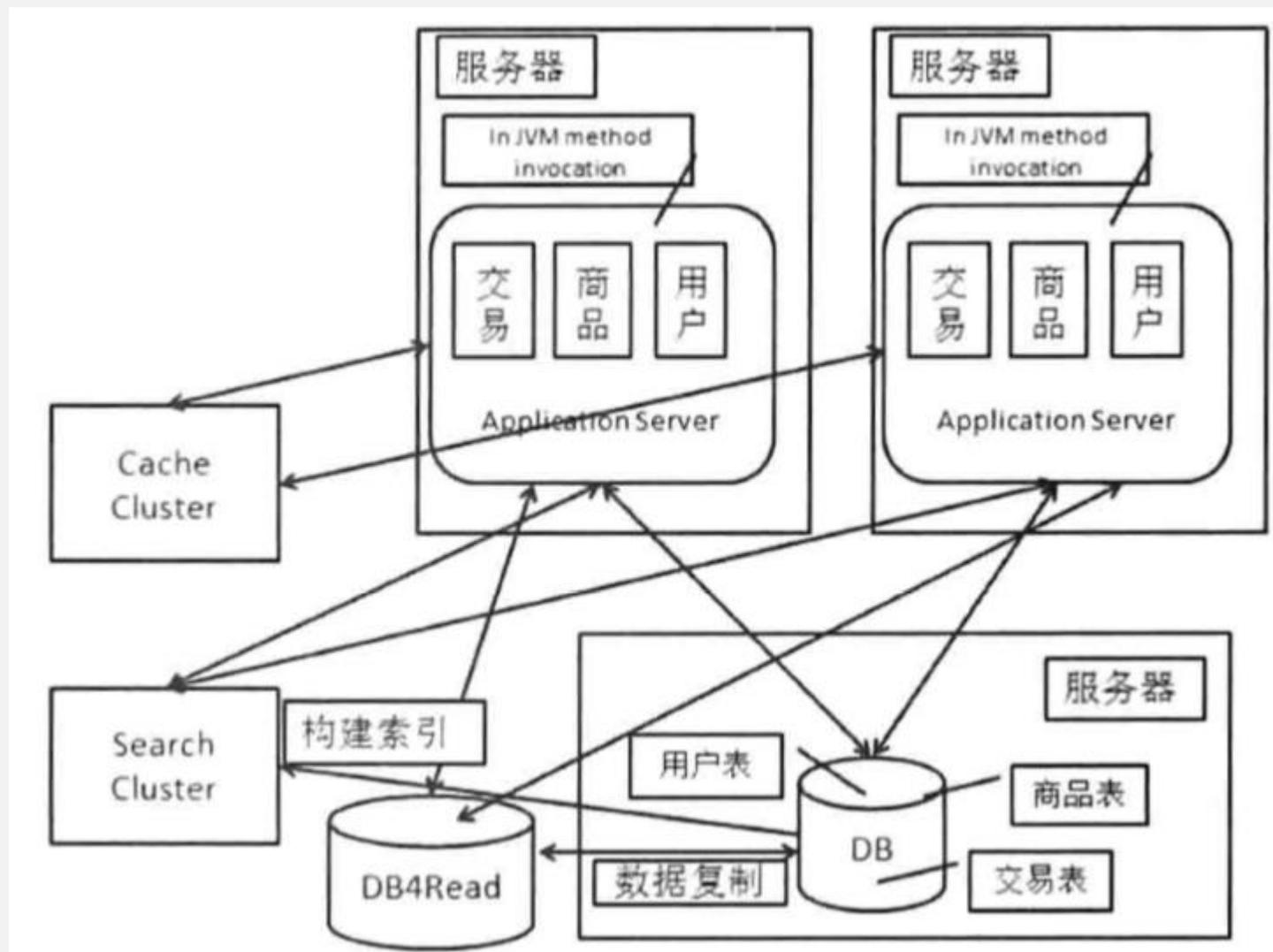
Memcached

EhCache

Static Map

OsCache

JCache



大型网站的架构演进

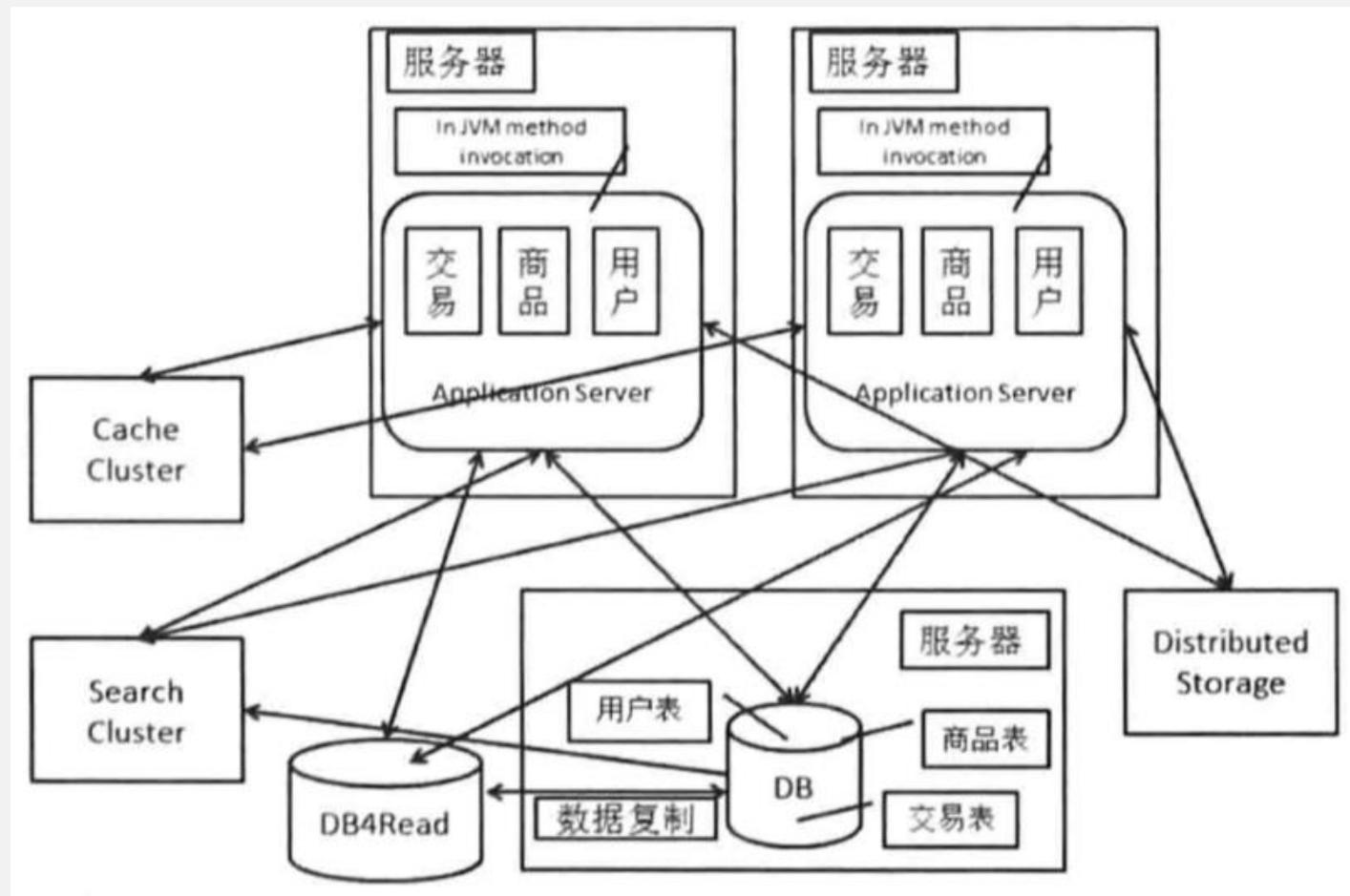
弥补关系型数据库的不足，引入分布式存储系统



弥补关系型数据库的不足，引入分布式存储系统

分布式存储系统通过集群提供一个高容量，高并发访问，数据冗余容灾的支持。

1. 分布式文件系统，解决文件的存储问题；
2. 分布式key-value系统，提供高性能的半结构化的支持；
3. 分布式数据库，提供支持大数据、大并发的数据库系统



大型网站的架构演进

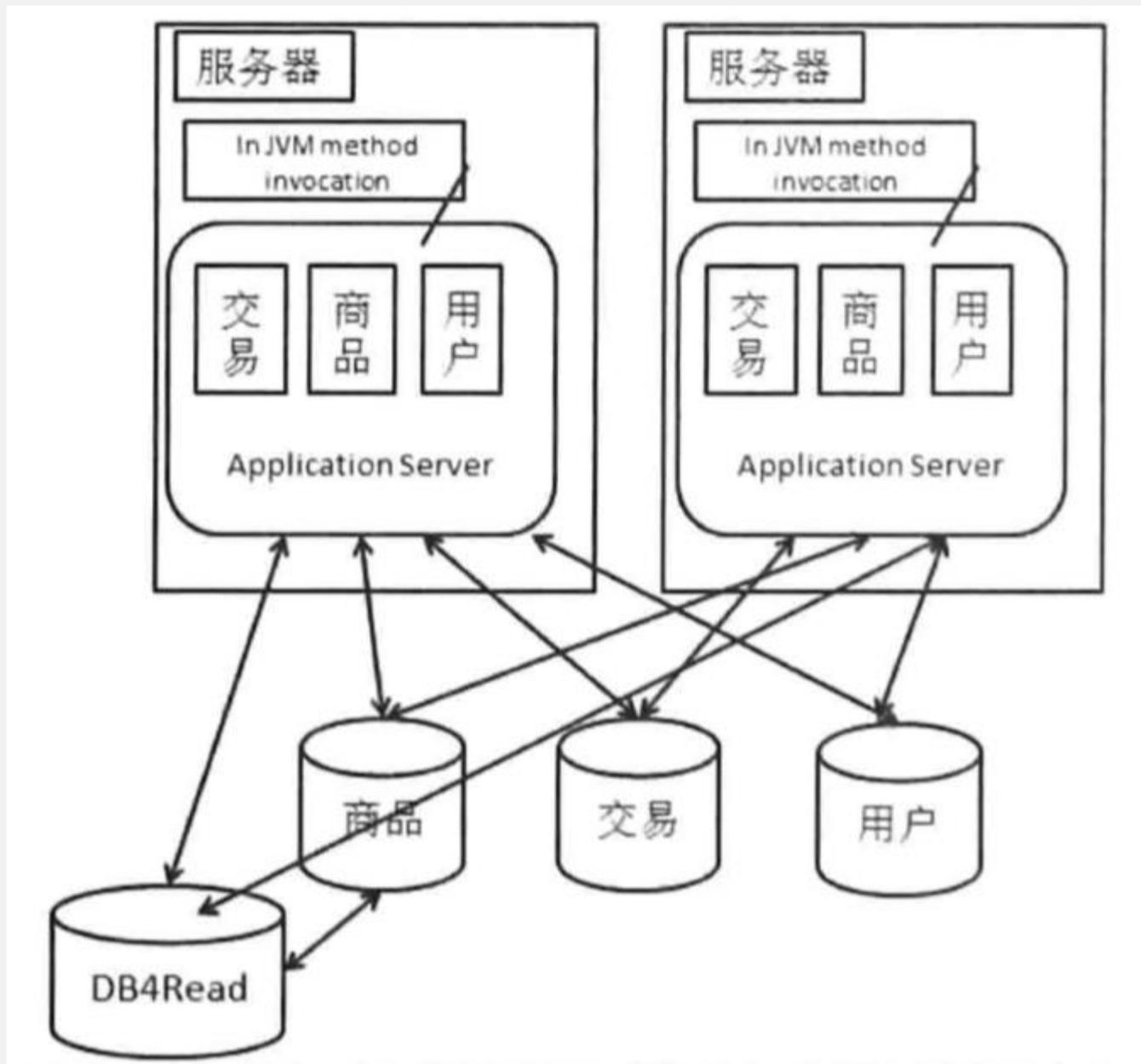
读写分离后，数据库又遇到新的瓶颈



垂直拆分

把数据库中不同的业务数据拆分到不同的数据库中。

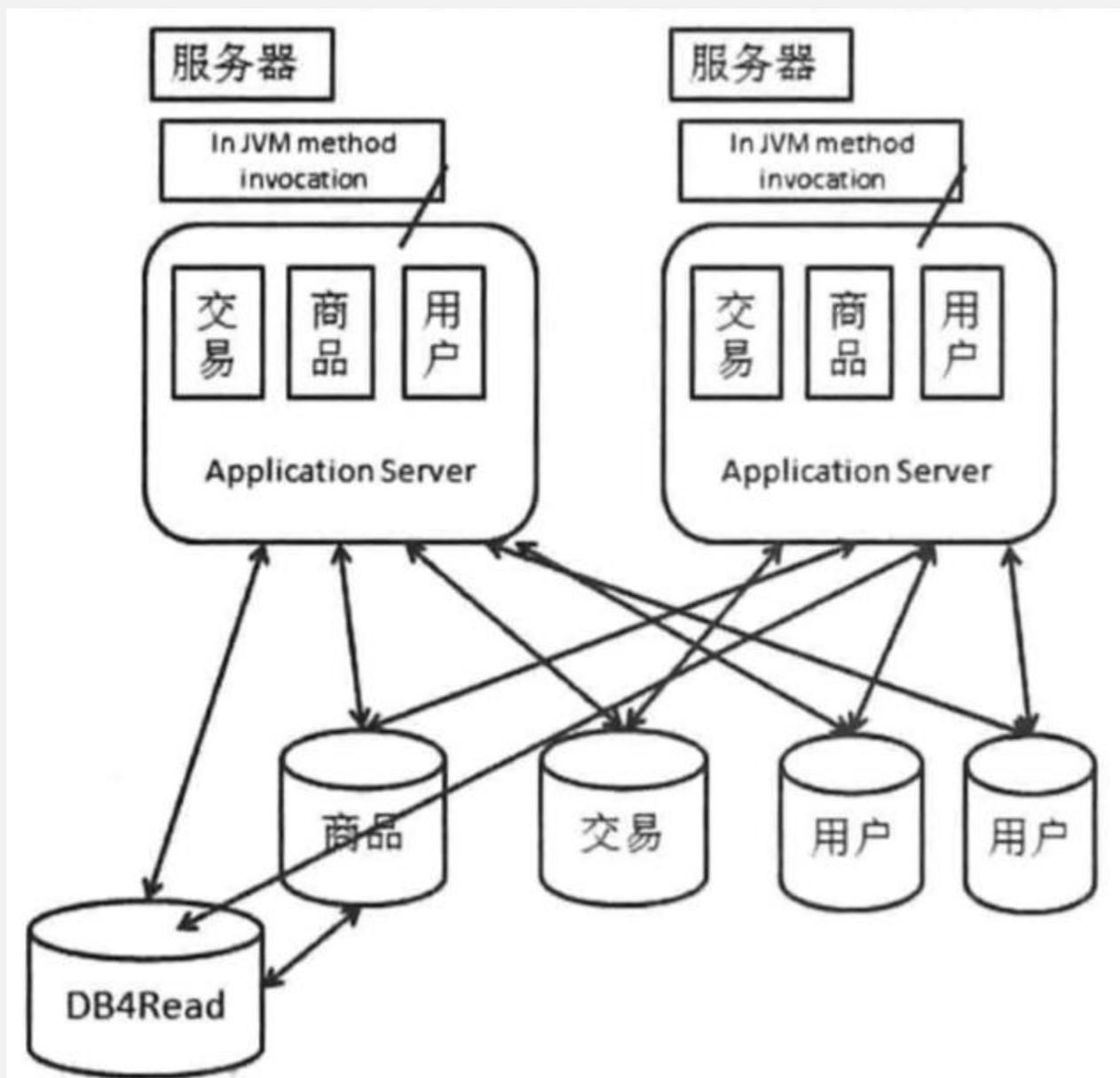
对数据进行垂直拆分后，解决了把所有业务数据放在一个数据库中的压力问题，并且也可以根据不同业务的特点进行更多优化。



水平拆分

某个业务的数据表的数据量或者更新量达到了单个数据库的瓶颈，这时把同一个表的数据拆分到两个数据库中。

一旦完成数据的水平拆分，将能够很好地应对数据量及写入量增长的情况。



大型网站的架构演进

数据库问题解决后，应用面对的新挑战



随着业务的发展，应用的功能会越来越多，应用也会越来越大。需要考虑如何不然应用持续变大，这就需要将应用拆开，从一个应用变为两个甚至多个应用。

第一种方式：根据业务的特性将应用拆分

第二种方式：按功能拆分



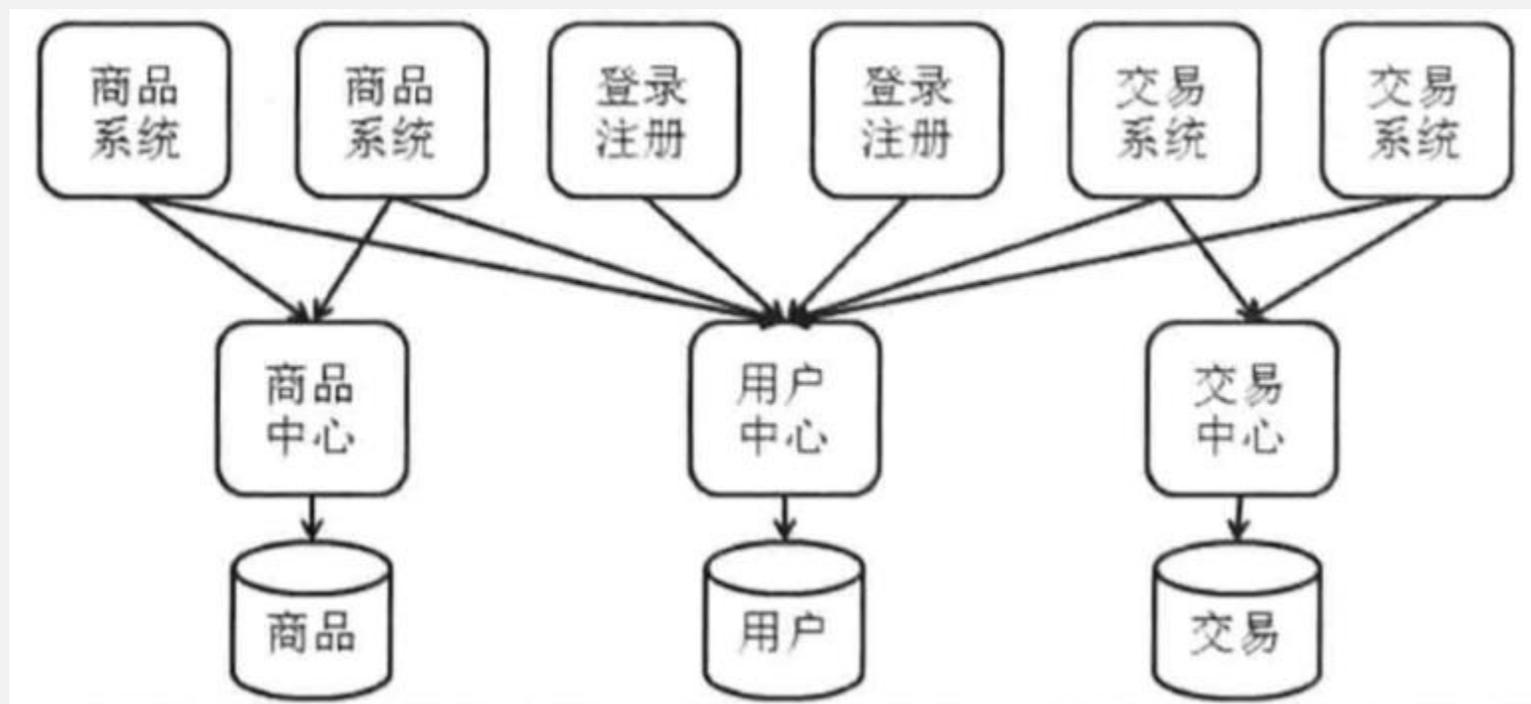
大型网站的架构演进

走服务化的路



走服务化的路

1. 业务功能之间的访问不仅是单机内部的方法调用，还引入远程的服务调用
2. 共享的代码不再散落在不同应用中，这些被放在了各个服务中心。
3. 数据库的连接也发生了一些变化，把与数据库的交互工作放到了服务中心，让前端的Web应用更加注重与浏览器交互的工作，而不必过多关注业务逻辑。
4. 通过服务化，无论前端Web应用还是服务中心，都可以由特定小团队维护。能更好保持稳定性，并能更好控制系统本身发展

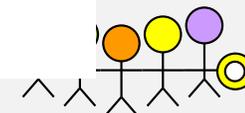
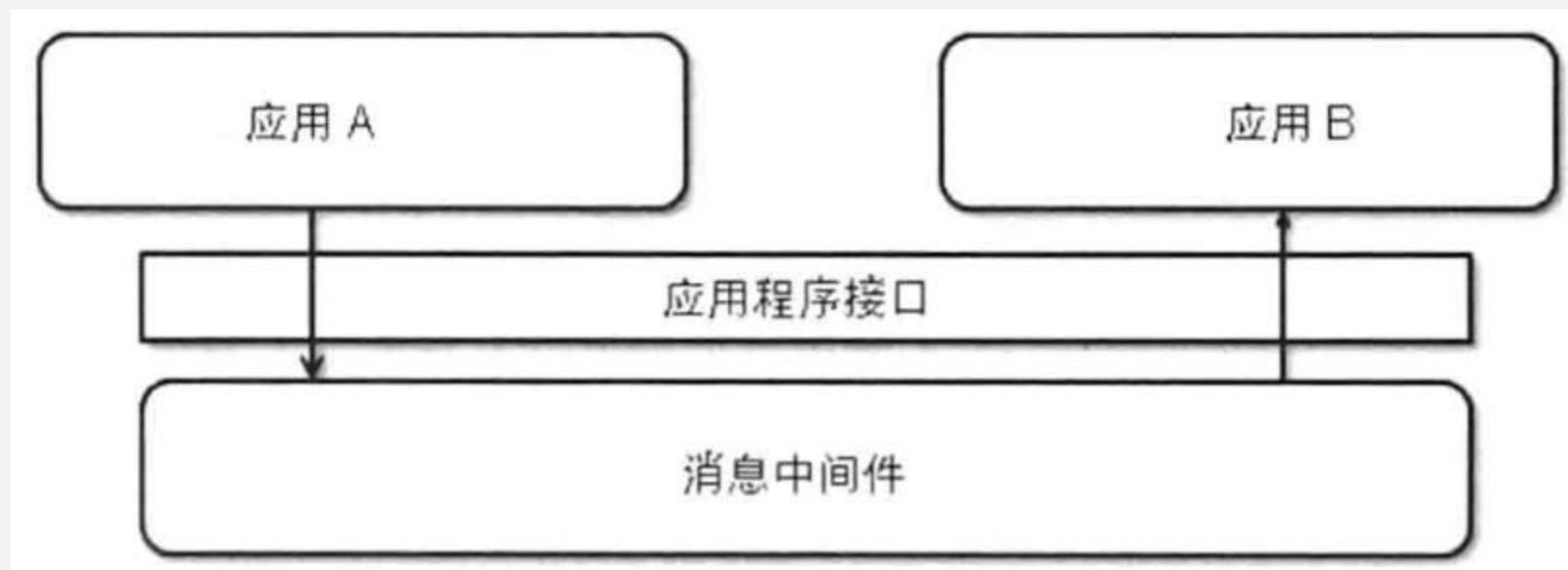


大型网站的架构演进

消息中间件



消息中间件是面向消息的系统
是在分布式系统中完成消息的发送和接收的基础软件
好处是异步和解耦

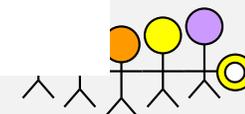
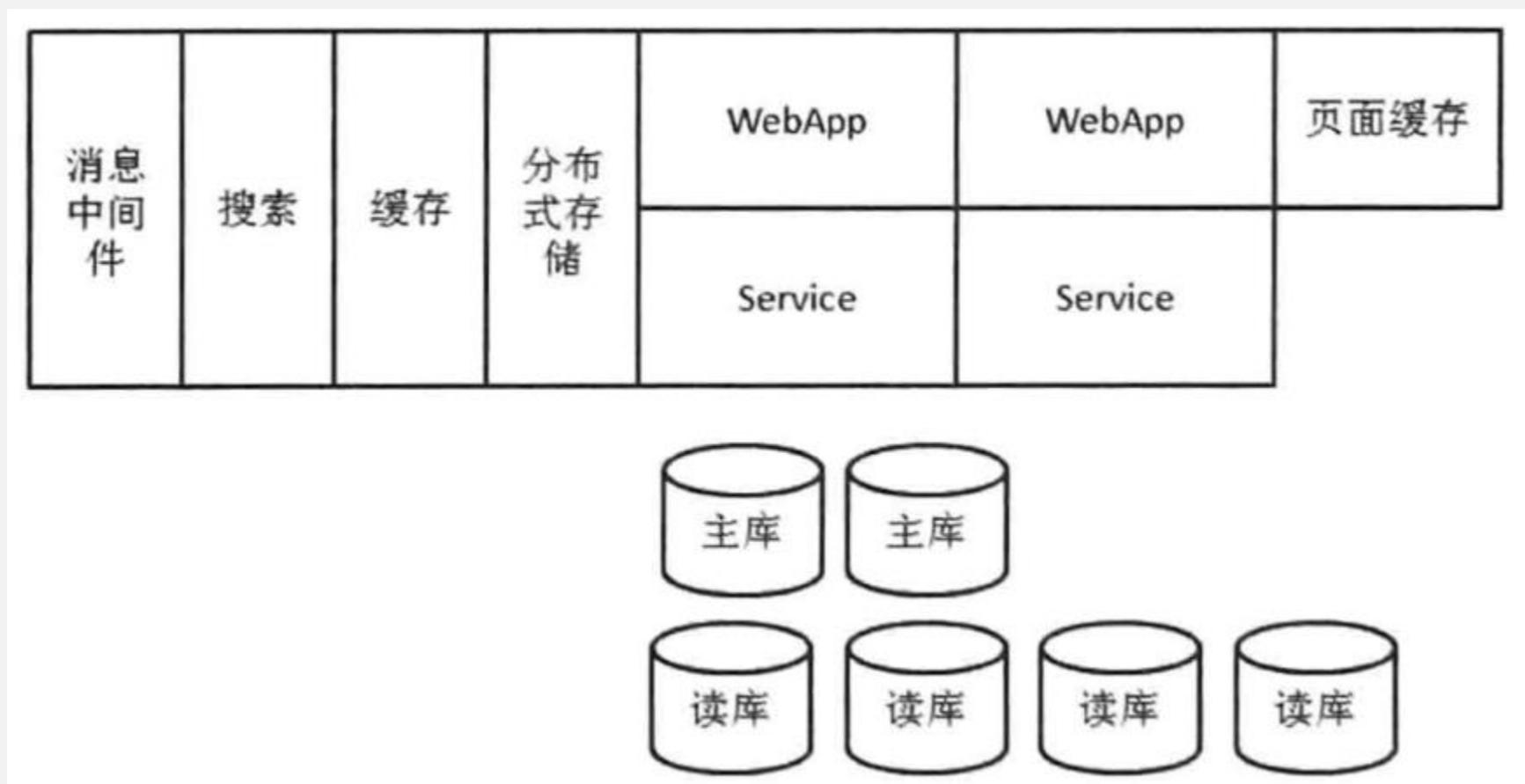


大型网站的架构演进

架构总结



我们的网站变成了什么样？



谢谢!

