

构建实时垂直搜索网站经验分享

刘连春@去哪儿网

大纲

- 垂直搜索
- 实时搜索
- 监控系统

垂直搜索

- 什么是垂直搜索？
 - 与业务紧密耦合
- 高访问量
- 高可用性
- 快速响应

垂直搜索-高访问量

- 构建集群
- 动静分离
- 转发策略
- 缓存系统
- 消息系统

高访问量-构建集群

Nginx

AppServer

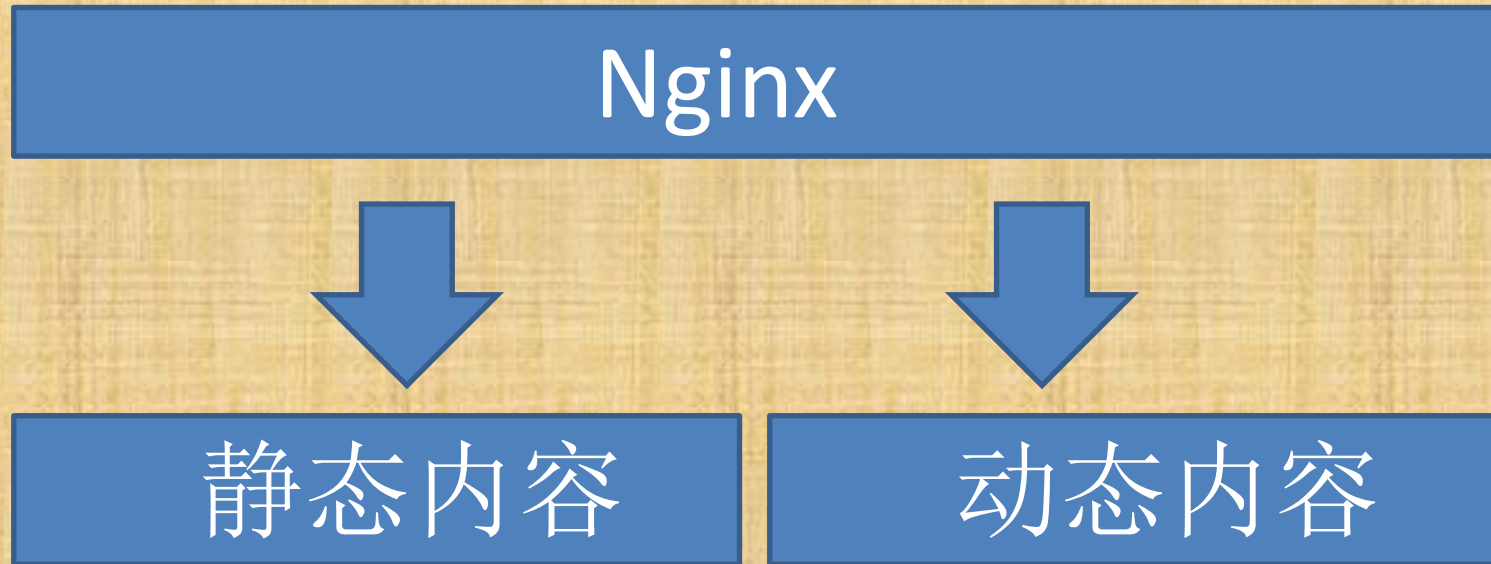
AMQ

Memcached

Spider

添加机器即可扩容

高访问量-动静分离



高访问量-转发策略

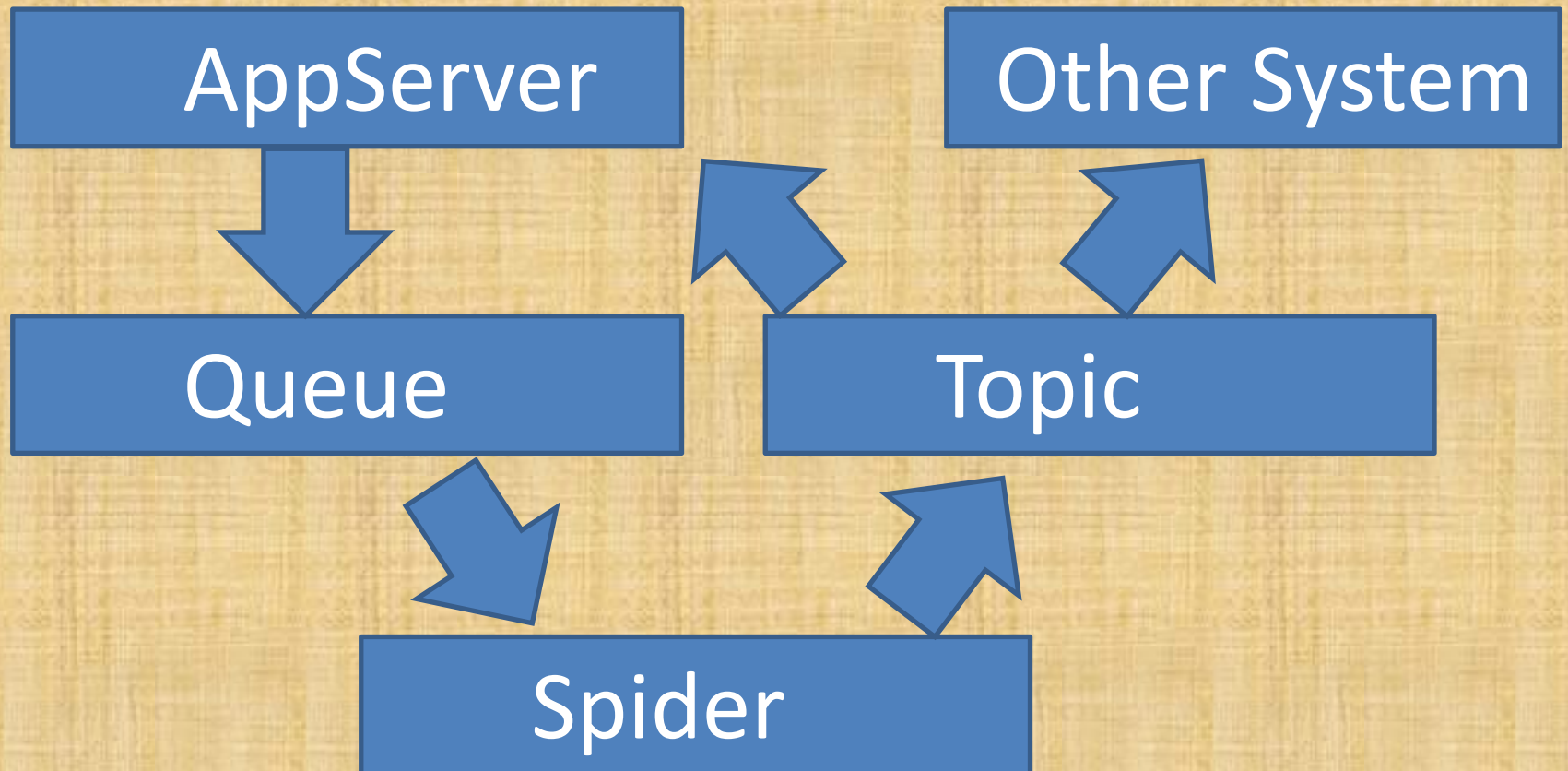
- 静态内容轮叫（Round-Robin）
- 动态内容转发
 - 按用户转发
 - 按搜索条件转发

高访问量-缓存系统

- 基本信息缓存
 - 配置文件
 - 基础数据
- 分布式缓存
 - Memcached

高访问量-消息系统

- AMQ (ActiveMQ)



垂直搜索-高可用性

- 消除单点
- 自动故障转移
- 系统尽量小而多
- 控制后端压力
- 使用外部接口经验
- 留有备用系统

高可用性-消除单点

- 为什么要消除单点
 - 软硬件故障是常态
- 设计上的单点
 - 双机是基本要求
- 隐藏的单点
 - 容量单点
 - 配置单点

高可用性-自动故障转移

- Failover
- Health check

高可用性-系统尽量小而多

- 降低影响用户比例
- 快速启动
- 快速添加服务器

高可用性-控制后端压力

- 分航线配置抓取
- 有效航线
- 限制搜索量
- 自动上下线
- 协助对方优化系统

高可用性-使用外部接口经验

- 假定不稳定
 - WebService, Socket, Mail
- 设置超时
- 添加监控
- 用日志界定边界
 - 所有的输入输出

高可用性-留有备用系统

- 缓冲系统压力

垂直搜索-快速响应

- 有效利用缓存数据
- Ajax多次回数
- 优化静态内容

快速响应-有效利用缓存

- 快速返回未过期数据
- 主动更新缓存
 - 首页低价

快速响应-Ajax多次回数

- 只返回必要的的数据
- 动态调整刷新时间

快速响应-优化静态内容

- 利用CDN网络
- 缓存静态内容

实时搜索

- 缓存更新策略
- Memcached使用经验分享

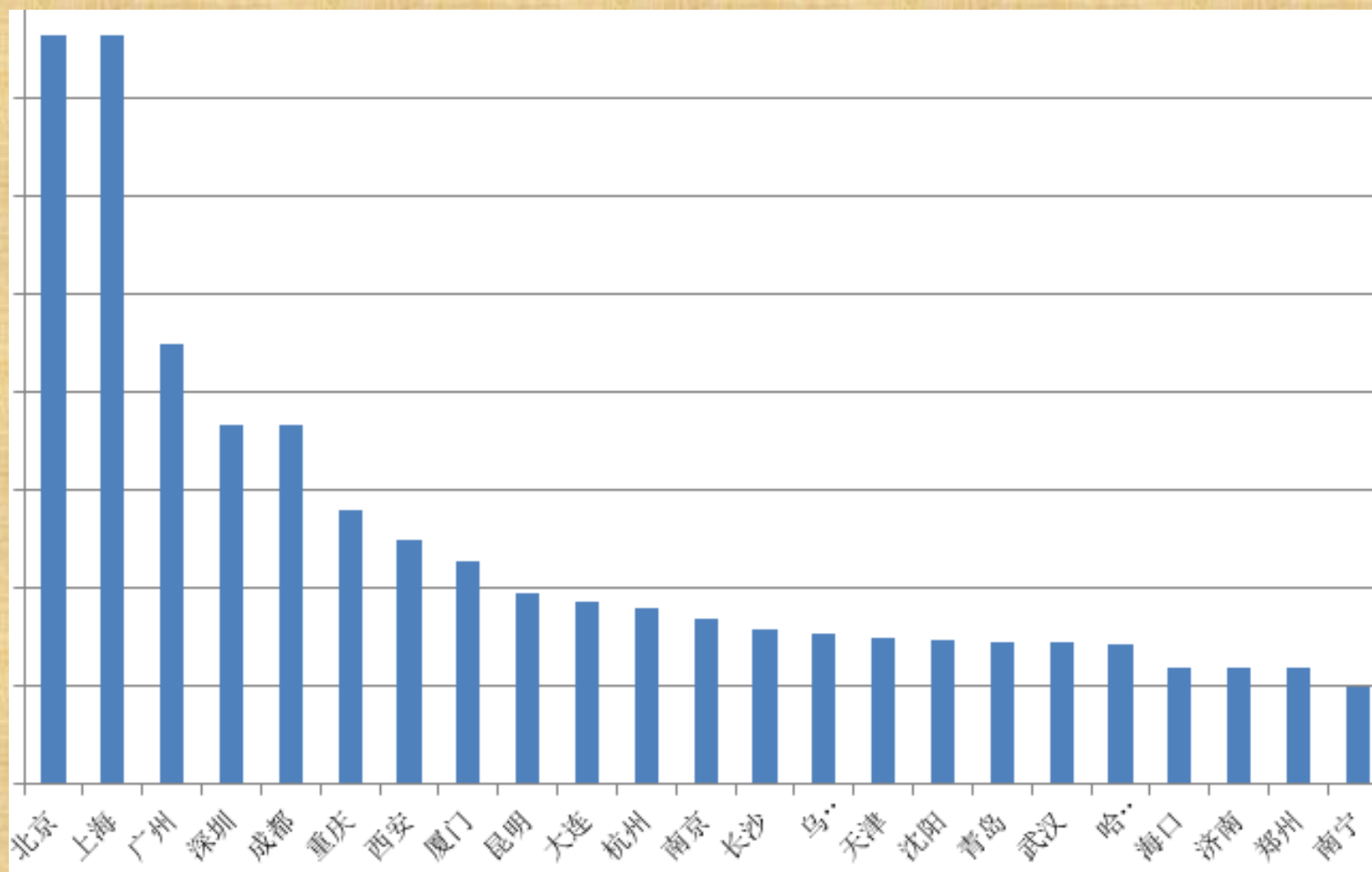
实时搜索-缓存更新策略

- 设置过期时间
- 反馈式更新
- 热门点击易变化

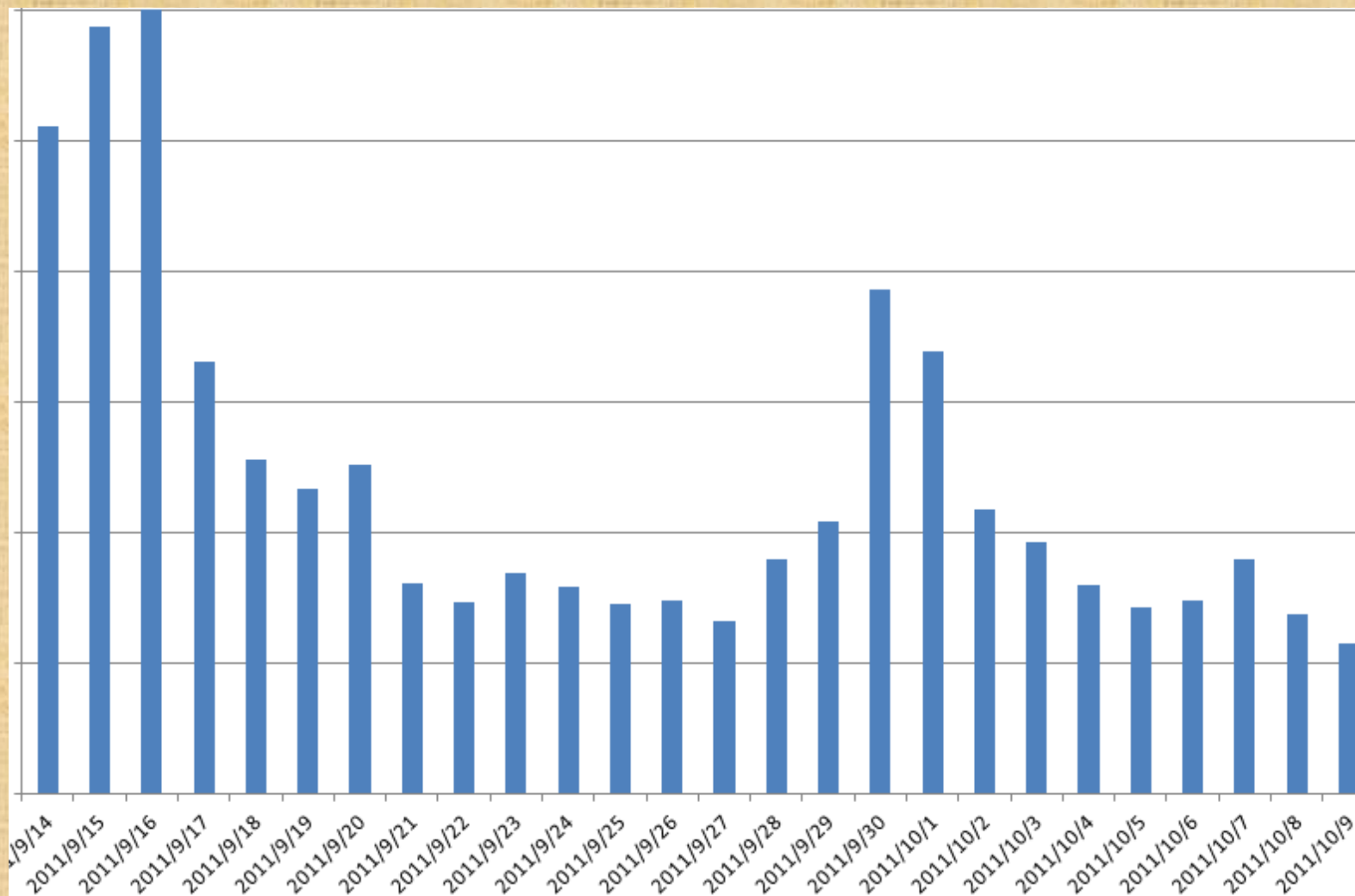
缓存更新策略-设置过期时间

- 根据业务特点设置

热门出发城市

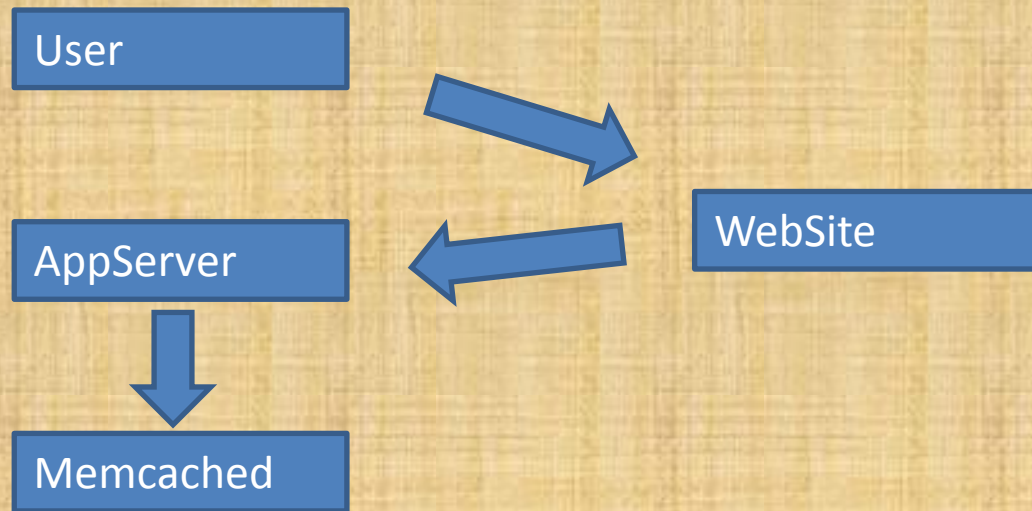


热门日期



缓存更新策略-反馈式更新

- 构建反馈回路



缓存更新策略-热门点击易变化

- 最低的价格
- 最好的时段

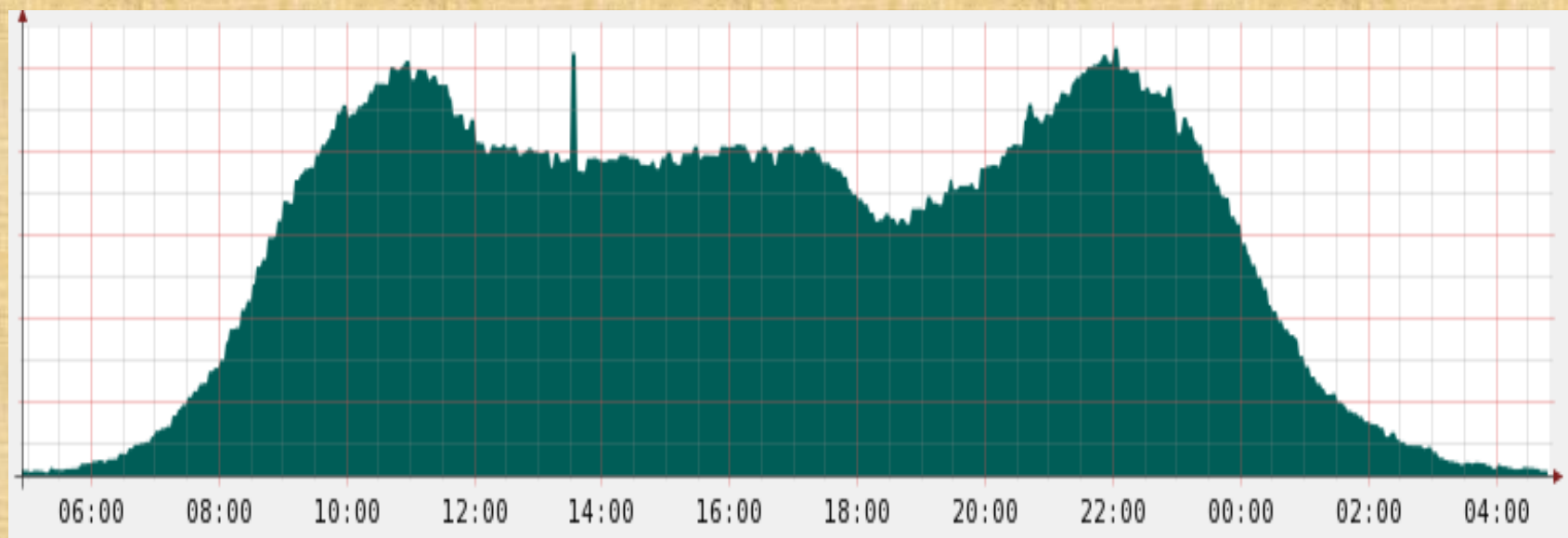
实时搜索-Memcached

- 独立的LRU
- 设置合理的过期时间
- 缩减Key长度
- 设置客户端的网络超时

监控系统

- Cacti+Nagios
- 深入应用的探针
 - 了解应用特性
 - 快速定位故障
 - 指导优化
- 分时报警
 - 不同时间段，不同的预警值
- 设置合理阈值

业务监控，重中之重！



9月17日搜索的结果



去哪儿网
Qunar.Com
聪明你的旅行

stevenliu | 我的机票▼ | 退出 | 关于订票 | 收藏

首页 | **机票** | 酒店 | 团购^新 | 度假 | 火车票 | 博客 | 更多 >

手机版

☒ 单程

城市 从 到

☐ 往返

日期

搜索

搜索返程

我喜欢的网站 | 搜索历史▼

去哪儿提供的链接

北京特价机票010-51665188
Tel:010-5166 5188

北京特价机票010-57798738
www.bjxsjip.com

十一上海自由行1100元起
www.5i7u.cn

10-18 周二 ¥446	10-19 周三 ¥446	10-20 周四 ¥446	10-21 周五 ¥325	10-22 周六 ¥325	10-23 周日 ¥325	10-24 周一 ¥325	价格 日历
------------------	------------------	------------------	------------------	------------------	------------------	------------------	----------

搜索120家网站，其中99家有直飞报价，105家有联程报价，正在搜索**机票吧**

起飞时间 ☐ 上午 ☐ 中午 ☐ 下午 ☐ 晚上 机型 ☐ 大型机 ☐ 中型机 ☐ 其他机型 [更多筛选条件▼](#)

航空公司/航班 机型	起降时间	机场	准点率 平均延时	推荐网站	最低报价
 东方航空MU5166 波音737(中)	07:20 09:20	首都机场T2 萧山机场	95% 小于5分钟		¥325 2.8折 订票
 联合航空KN2287 波音737(中)	08:00 10:00	南苑机场 萧山机场	82% 14分钟		¥490 4.3折 订票
 东方航空MU8087 波音737(中)	08:00 10:00	南苑机场 萧山机场	代码共享 同KN2287		¥656 5.7折 订票
 东方航空MU5574 空客320(中)	22:55 01:00 ^{+1天}	首都机场T2 萧山机场	--		¥713 6.2折 订票



单程 人民币 **1379** 起

100%航协认证
去哪儿网有  的机票代理100%获得
中国航空运输协会资格认证

去哪儿通行证预订国内机票有“**赔**”计划

1. 付款后100%可以出票；
2. 付款后100%不加收额外费用；
3. 所购机票100%可以使用。

客户服务中心：010-57603866

9月17日搜索的结果



去哪儿? **Qunar.Com**
聪明 你的 旅行

[首页](#)[机票](#)[酒店](#)[团购](#)[度假](#)[火车票](#)[博客](#)[更多 >](#)

stevenliu | [我的机票](#) | [退出](#) | [关于订票](#) | [收藏](#)

[手机版](#)

☒ 单程

城市 从

到

☐ 往返

日期

搜索

搜索返程

[我喜欢的网站](#) | [搜索历史](#)

去哪儿提供的链接

百事通-特价机票预订专线
4001-118-114

特价机票预定热线
4006 113119

杭州-北京0571-85343801
www.hzlhjp.com

10-20 周四 ¥448	10-21 周五 ¥528	10-22 周六 ¥325	10-23 周日 ¥325	10-24 周一 ¥325	10-25 周二 ¥325	10-26 周三 ¥325	价格 日历
------------------	------------------	------------------	------------------	------------------	------------------	------------------	----------

搜索117家网站, 其中96家有直飞报价, 113家有联程报价, 共5341个报价信息, 搜索结束

起飞时间 ☐上午 ☐中午 ☐下午 ☐晚上 机型 ☐大型机 ☐中型机 ☐其他机型 [更多筛选条件](#)

航空公司/航班 机型	起降时间	机场	准点率 平均延时	推荐网站	最低报价
东方航空MU5163 波音737(中)	10:10 12:25	萧山机场B楼 首都机场	80% 19分钟	东航银联旗舰店 ¥350 直接订	¥325 2.8折 订票
东方航空MU5131 空客320(中)	07:50 09:50	萧山机场B楼 首都机场	94% 10分钟	东航银联旗舰店 ¥350 直接订	¥350 3.0折 订票
东方航空MU8088 波音737(中)	11:00 13:00	萧山机场B楼 南苑机场	代码共享 同KN2288	东航银联旗舰店 ¥430 直接订	¥430 3.7折 订票



成都/天津/杭州
→ 吉隆坡
豪华
单程
人民币1866起
AirAsia.com

100%航协认证

去哪儿网有的机票代理100%获得
中国航空运输协会资格认证

去哪儿通行证预订国内机票有“赔”计划

1. 付款后100%可以出票;
2. 付款后100%不赔付额外费用;

欢迎加入Qunar！

Email: lianchun.liu@qunar.com

微博: <http://weibo.com/lianchunliu>