

去哪儿网

机票实时搜索引擎的优化

2011.12

“去哪儿网”是中国最早的旅游搜索引擎，是目前全球最大的中文在线旅行网站。

机票搜索是“去哪儿网”核心业务之一。

实时搜索国内293个,国际265个代理商网站。
接口动态请求数 2亿次/天。

第一部分：前端优化分享

前端优化分享

单程

城市 从

上海

到 重庆

搜索

搜索返程

往返

日期

2011-11-19

明天

我喜欢的网站 |

搜索历史 >



lufthansa.com
现在就访问lufthansa.com订购
¥3,300*起
Lufthansa

11-18 周五 ¥349	11-19 周六 ¥340	11-20 周日 ¥340	11-21 周一 ¥340	11-22 周二 ¥310	11-23 周三 ¥297	11-24 周四 ¥297	价格 日历
------------------	------------------	------------------	------------------	------------------	------------------	------------------	----------

搜索134家网站, 其中118家有直飞报价, 123家有联程报价, 正在搜索海航去哪儿旗舰店

起飞时间 上午 中午 下午 晚上

机型 大型机 中型机

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	
 四川航空3U8976 空客321(中)	20:35 23:35	浦东机场T1 江北机场	约100%	四川航空 ¥360	¥340 2.3折	订票
 中国国航CA4542 波音737(中)	19:50 23:00	浦东机场T2 江北机场	约100%	中国国航官网 ¥370 直接订	¥347 2.3折	订票
 中国国航CA1941 空客321(中)	16:30 19:30	浦东机场T2 江北机场	约100%	中国国航官网 ¥370 直接订	¥348 2.3折	热门 订票



lufthansa.com
到欧洲
¥3,300*起
不含任何机场税和其他费用
现在就访问lufthansa.com订购
德国汉莎航空公司

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

前端优化分享

四川航空3U8976 20:35 浦东机场T1 约100% 四川航空 ¥340
空客321(中) 23:35 江北机场 ¥360 2.3折 [订票](#)

飞行3小时0分钟 | 燃油税: 130元 | 机场建设费: 50元 | 有餐食 | 无网上值机 | 近期价格走势: 稳定波动

[上海差旅通旅行网](#) [担保通](#) [¥348](#)+20(保险) [预订](#)
[五星](#) 来自349位用户 [点评](#) [退改签](#) [¥363](#) [预订](#)

[1号店](#) [担保通](#) [¥349](#)+20(保险) [预订](#)
[四星](#) 来自253位用户 [点评](#) [退改签](#) [¥364](#) [预订](#)

[恒通之旅](#) [担保通](#) [¥359](#)+20(保险) [预订](#)
[五星](#) 来自321位用户 [点评](#) [7x24服务](#) [退改签](#) [¥370](#) [预订](#)

[普汇商旅网](#) [担保通](#) [¥340](#)+20(保险) [预订](#)
[四星](#) 来自417位用户 [点评](#) [退改签](#) [¥355](#) [预订](#)

[同程网](#) [7x24服务](#) [¥358](#)+20(保险) [预订](#)
[四星](#) 来自1034位用户 [点评](#) [退改签](#) [2分钟前更新](#)

[拍票网](#) [担保通](#) [¥437](#)+20(保险) [预订](#)
[五星](#) 来自312位用户 [点评](#) [退改签](#) [¥450](#) [预订](#)

[票谋天下](#) [担保通](#) [¥414](#)+20(保险) [预订](#)
[五星](#) 来自773位用户 [点评](#) [退改签](#) [¥429](#) [预订](#)



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

去哪儿提供的链接 [更多](#)

[保](#) 商户, 服务安全双重保障

[上海-重庆 2-5折 021-6206 1260 特价中](#)
各大航空2-5折特价促销中 折扣低 优惠多 信用卡免...
[www.5yaofei.com](#) [保](#)

[上海-重庆 2折起 021-62101085](#)
各大航空2折起 免费送票可刷卡 提供月结
[www.yuxintjipiao.com](#) [保](#)

[上海-重庆【2折起特价】021-51087192](#)
精选国内特价机票 航空公司直销 市内免费送票!
[www.shyh66.com](#)

[【抢订】上海-重庆 2折起 021-51870981](#)
各大航空特价机票2折起售 信用卡支付 市内免费送票
[www.98air.com](#)

[上海-重庆【特价热线】021-3126 2136](#)
最低2折 数量有限 速速抢购 市内免费送票!

JS渲染, 到底做了什么?

- 前端计算最低价、排序、过滤、分页
- 状态维持
- 实时更新列表区域 (约1500个DOM元素)



难题？

- 如何更快？
 - 首屏加速
 - 实时更新带来的重绘，让用户感觉不到
- 如何更易维护？
 - 处理交互及状态维持
 - 数据实时展示

丑陋的TrimPath页面模板

```
{var GROUP="normal"}
{var loopData = getPureData(pureData,currentPage)}
{var recordindex = 0}
{for record in loopData}
{var recordindex = recordindex + 1}
    {var bean = FlightList.getBean(record.key)}
    {if bean}${setPage(bean,currentPage,recordindex)}{/if}
    {if record.groupType == "av-onewayflight"}
        {var info = record.info}
        <ul id="anchor_${record.key}$$${GROUP}" >
            {if paper.quasipointrate}\
                ${qu}\
            {else}\
                <div class="gray">${languageVars._NO_CORRECTNESS}</div>\
            {/if}\
        </ul>
    }
}
.....
```


模板的问题

- 不够快

- 需要先被编译，才能使用（Parse Time）

- 15k 模板文件需要编译约14ms，该时间内浏览器僵死
 - 搜索结果页模板曾经到达70k，僵死时间为约65ms
 - 以上数据为firefox 7获取，IE6约400ms，用户感觉明显

- 模板引擎使用eval, with, etc...（Apply Time）

- 搜索结果页的Apply Time，每次需要约250ms，浏览器僵死。
 - 1分钟内，每隔3至8秒更新一次

- Dom Tree完成后，需要用选择器获取元素

- 未优化的选择器（如包括逗号的多元素选择），响应时间高达200ms

模板的问题

- 解耦/维护困难

- 模板中包含大量逻辑，if / else / for ...
- 模板可以嵌套使用，但是...
 - 在模板本身嵌套，很难将子模板组件化。
 - 以组件方式，必须在父组件完成渲染后嵌套。
- 使用id，必须手工维护，代码丑陋

```
<ul>
  {{foreach list}}
    <li><input type="checkbox" /><span id="item${index}">${name}</span></li>
  {{/foreach}}
</ul>
```

XControl

Qunar的模板解决方案

XControl能够解决？

- 够快

- 实现代码仅3k
- 无需编译
- Dom Tree 完成后，取得元素方法核心是document.getElementById
- 模板生成时，使用Array Join方式拼接模板片段及子组件

	TrimPath或其它模板引擎	XControl
Parse Time	决定于模板规模	0
Apply Time	决定于模板引擎生成的中间代码	join拼接
取得模板渲染后元素	决定于选择器的性能	最快

XControl能够解决？

- 解耦/维护容易
 - 易于理解的业务逻辑与UI混排
 - 支持子组件嵌套调用
 - 自动管理ID

XControl能够解决?

- 解耦/维护容易
 - 开发一个 to do list (HTML)

```
<ul>
```

```
<li><input type="checkbox" /><span id="itemXl32">text 1</span></li>
```

```
<li><input type="checkbox" /><span id="itemXl34">text 2</span></li>
```

```
<li><input type="checkbox" /><span id="itemXl36">text 3</span></li>
```

```
</ul>
```

XControl能够解决?

- 解耦/维护容易
 - 开发一个 to do list (模板实现)

```
var template = '...!';
```

```
var t = TrimPath.parseTemplate( template );
```

```
var html = t.process( data );
```

```
document.body.innerHTML = html;
```

```
document.getElementById('itemX132').innerHTML = 'hello!';
```

```
$(“li:first span”).html('hello!');
```

XControl能够解决?

- 解耦/维护容易
 - 开发一个 to do list (基于XControl 实现)

```
function Item(){}
```

```
function List(){}
```

```
var list = new List();
```

```
list.append( new Item() );
```

```
list.renderTo( document.body );
```

```
list.items[0] .find('item').innerHTML = 'hello!';
```

```
<span id="itemXI32">text 1</span>
```

前端优化分享

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	
 中国国航CA1419 波音737(中)	07:30 10:00	首都机场T3 江北机场	69% 17分钟	中国国航官网 ¥630 直接订	¥555 3.6折	订票
飞行2小时30分钟 燃油税: 130元 机场建设费: 50元 有餐食 有网上值机 近期价格走势: 大幅下跌						
 AIR CHINA 中国国际航空公司 A STAR ALLIANCE MEMBER	进入国航官网, 注册登录 立享10元优惠, 常旅客用 户登录更可享受20元优 惠。		 7×24服务  退改签	¥630 4.0折		预订 2小时前更新
☆☆☆☆☆ 来自4503位用户	点评					
 天泰航空	☆☆☆☆☆ 来自370位用户	点评	 担保通  7×24服务  退改签	¥555+20(保险) ¥570		预订 预订
 360机票网	☆☆☆☆☆ 来自533位用户	点评	 担保通  7×24服务  退改签	¥572+20(保险) ¥587		预订 预订
 逍遥行商旅网	☆☆☆☆☆ 来自1525位用户	点评	 担保通  7×24服务  退改签	¥578+20(保险) ¥593		预订 预订

XControl如何工作?

- 初始化

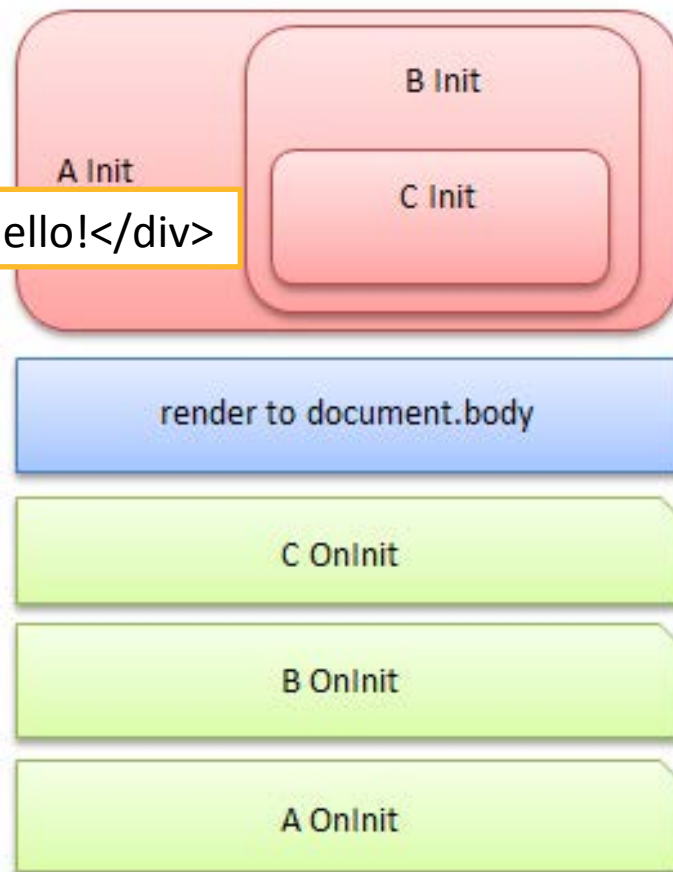
```
A.append(B);  
B.append(C);  
A.append("<div id='myidXI3324'>hello!</div>")
```

- 渲染

```
A.renderTo(document.body);
```

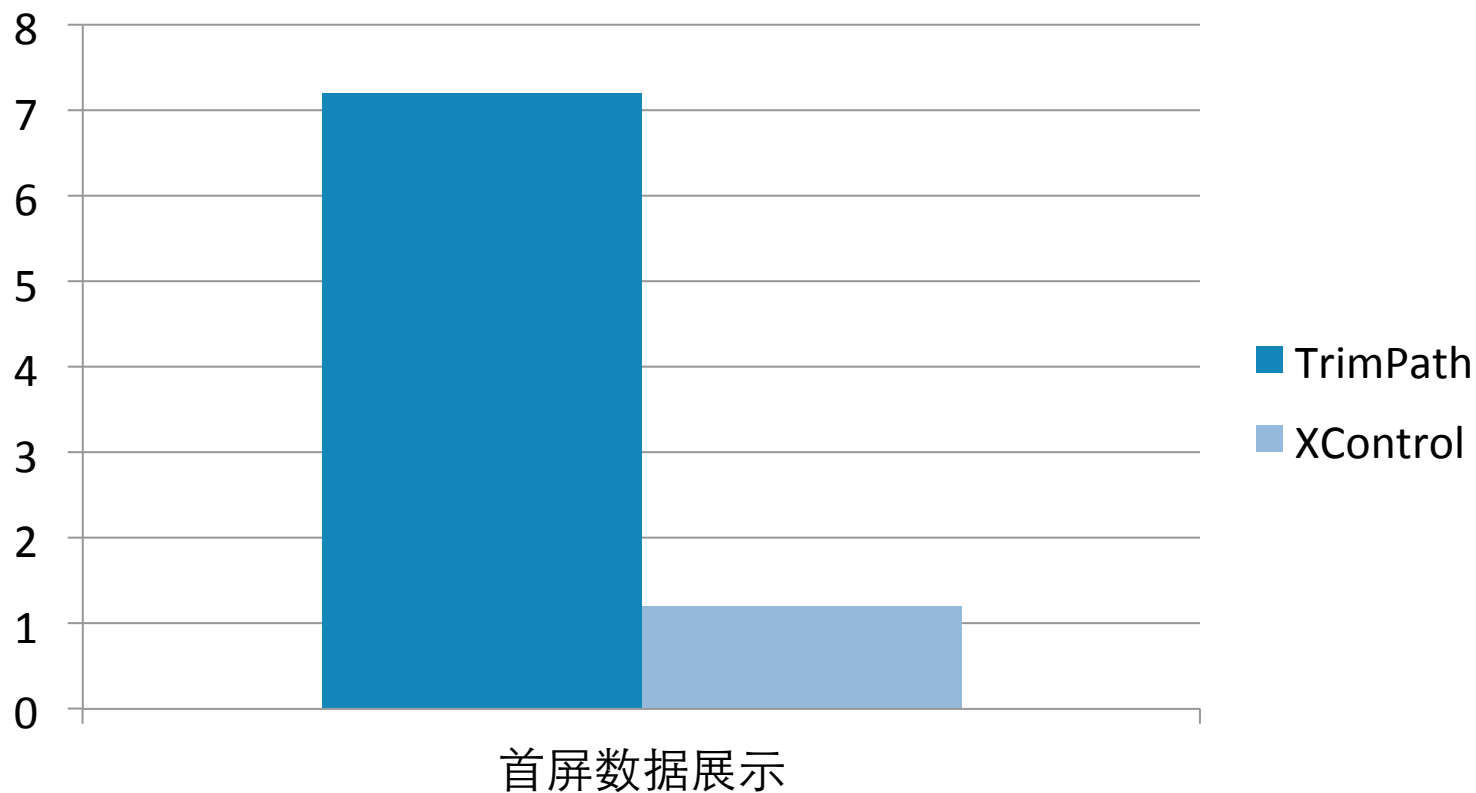
- 触发事件

```
fire C.onInit  
fire B.onInit  
fire A.onInit
```



TrimPath与XControl性能对比 (IE6)

性能提升73%



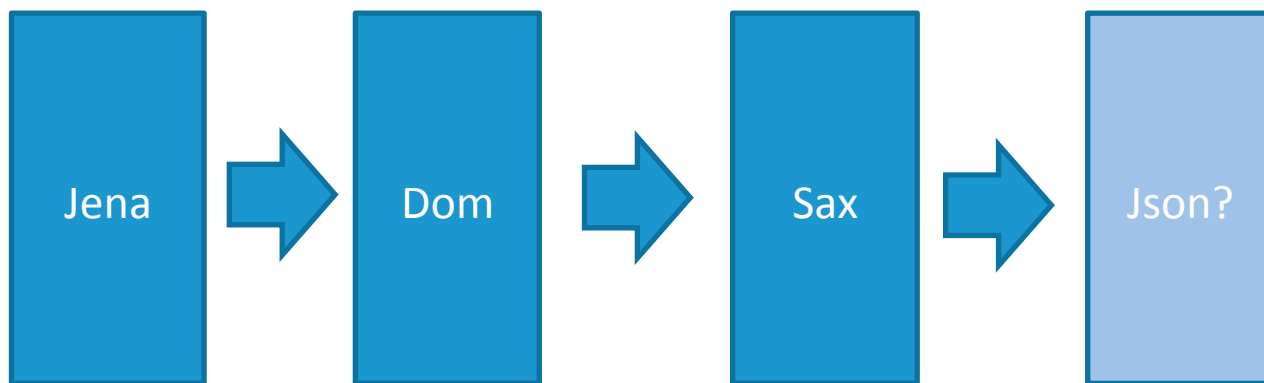
小结

- 技术选型需要结合自身业务特点
- 优化需要重视细节

第二部分： 后端优化分享

- 数据解析
- 内存占用
- 设计优化

RDF(XML)解析



避免过早、过度设计

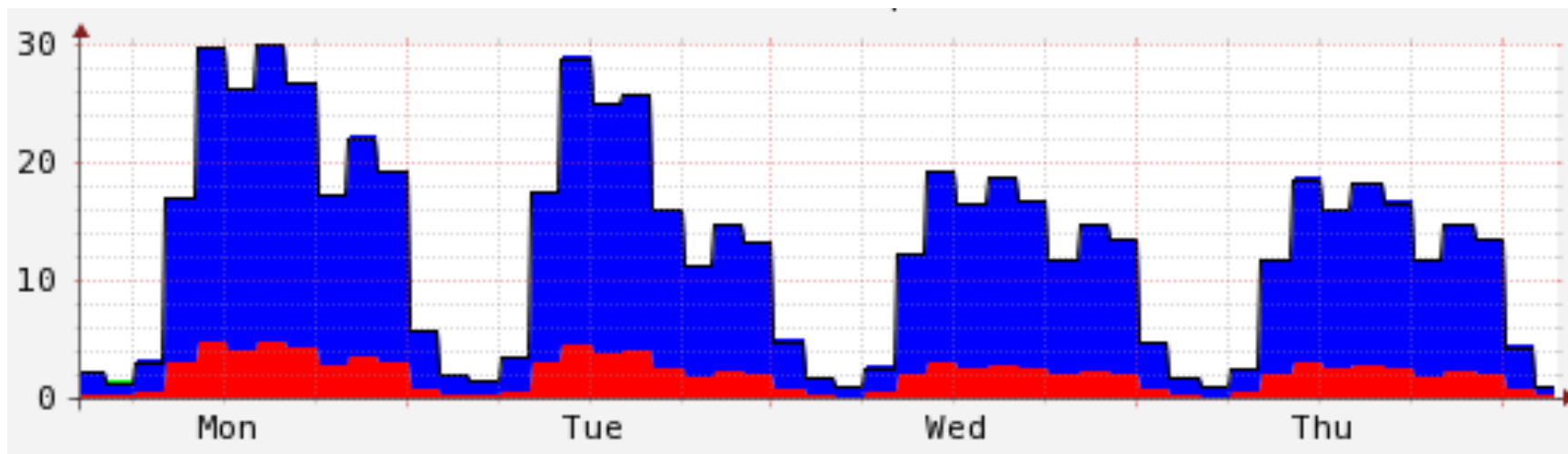
不走寻常路

把XML当做字符串处理

长度72723的XML，Dom处理 335ms

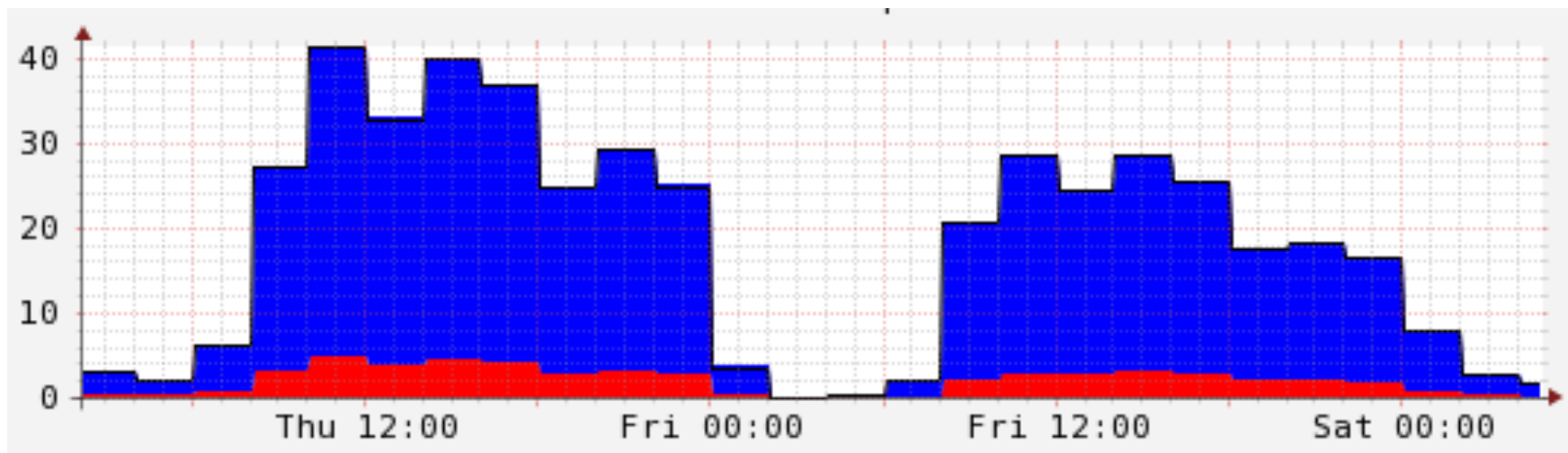
长度72723的XML，字符串处理3ms

1. 适合系统内部
2. 对性能敏感



修改数据格式

使用自定义文本压缩XML (由58k -> 11k, 减少80%)
节省25%的CPU消耗



后端优化 – 内存占用

102e3a68ded,DLC,CGO,2008-12-23

112b82c6c1d,CKG,CAN,2008-11-20

```
class CacheKey {  
    String wrapperId;  
    String depCode;  
    String arrCode;  
    String depDate;  
}
```

1. 每个字段的范围都很小
2. 组合Key节约内存

后端优化 - 内存占用

102e3a68ded,DLC,CGO,2008-12-23

长度: (11 + 3 + 3 + 10 = 27)

```
class String {  
    char[] value;  
    int offset;  
    int count;  
    int hash;  
}
```

在32位系统上，每个普通对象头大小为8字节，数组对象头为12字节，一个char类型占用2个字节，所以一个长度为27的String对象占用内存大小为：

$8 + 4 + (12 + 27 * 2) + 4 + 4 + 4 = 90$ 字节

后端优化 – 内存占用

```
class CacheKey {  
    String wrapperId;  
    String depCode;  
    String arrCode;  
    String depDate;  
}
```

一个CacheKey对象的内存占用为:

$8 + 4 + 4 + 4 + 4 = 24$ 字节

wrapperId集合范围约为300个，depCode, arrCode在同一集合内，范围约为200个，depDate的集合范围为366，也就是这些基础对象占用的内存总量约为：

$58 * 300 = 17400$

$42 * 200 = 8400$

$56 * 366 = 20496$

总计：46296字节

后端优化 - 内存占用

在生产环境这些数据会达到1000万条，采用拼接字符串的方式占用内存为：

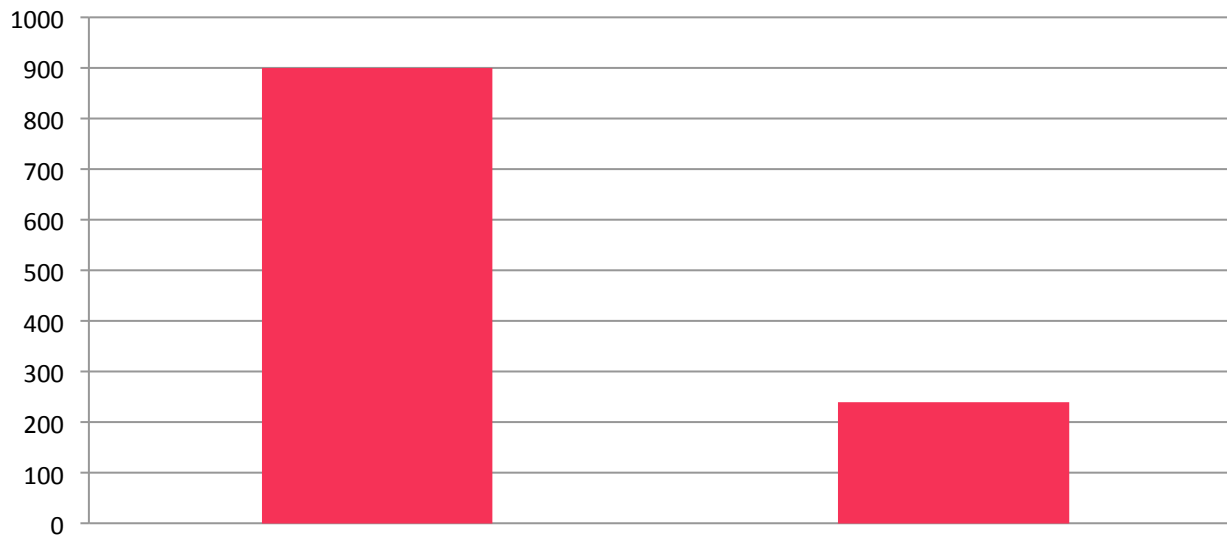
$$90 * 10M = 900M$$

新的方案占用内存为：

$$24 * 10M + 46k = 240M$$

优化后内存占用比优化前减少内存占用减少660M（73%）

集群内有40台服务器，可以减少26.4G的内存占用



后端优化 – 内存占用

- 使用自定义CacheKey可以通过jmap等工具实时查看系统中CacheKey对象的数量
- `jmap -histo $pid`

```
[lianchun.liu@l-tr12.f.cn1 ~]$ sudo jmap -histo 25089
```

num	#instances	#bytes	class name
1:	11068651	442746040	java.util.LinkedList\$Entry
2:	1245386	222646120	[C
3:	514039	197110984	[I
4:	81823	104695240	[B

Memcached优化

- Key是数据的一部分



/
submit=Test&codeBase=ttsgnd00164&codeName=Yoyod&required0=
SHE&required1=PEK&required2=2011-12-11 （长度96）

Flight_ttsgnd00164_SHE_PEK_2011-12-11 （长度37）

节省（61%）

Memcached优化

- 数据是按Slab保存的
 - 内存无法共享
 - 会浪费数据空间



100K Slab

200K Slab

300K Slab

Memcached优化

- 设置合理的过期时间
杜绝不设置超时的用法
- 优化回收过期数据策略

前端数据按需返回

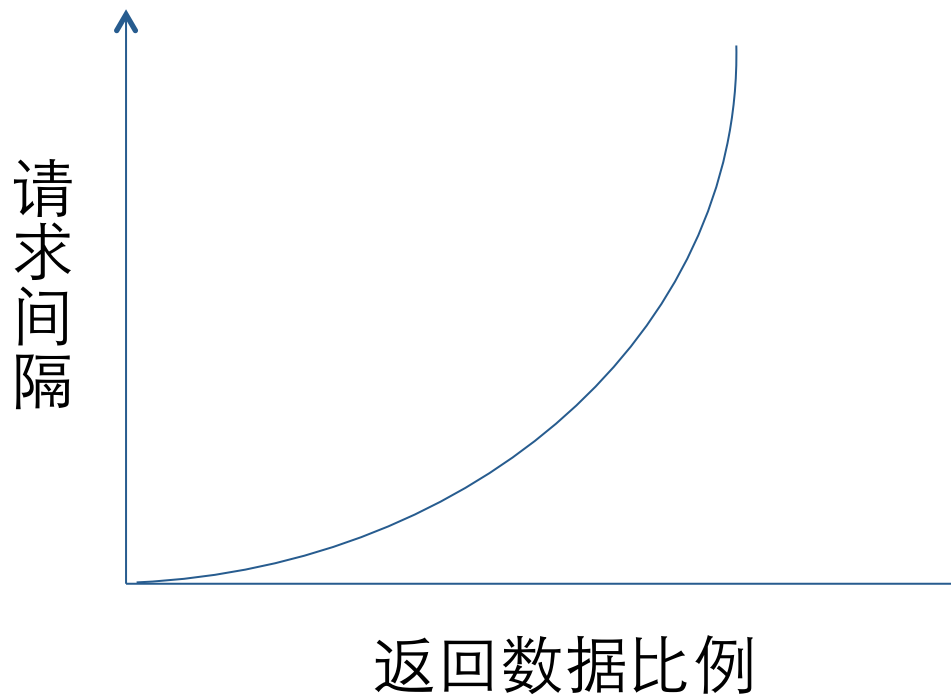
航空公司/航班 机型	起降时间	机场	准点率 平均延时	推荐网站	最低报价	
 春秋航空9C8992 空客320(中)	06:35 09:00	首都机场T1 虹桥机场	--	春秋航空 ¥290	¥283 2.5折	订票
 东方航空MU5138 空客330(大)	07:00 09:10	首都机场T2 虹桥机场	约100%	东航银联旗舰店 ¥450 直接订	¥417 3.7折	订票
 东方航空MU5102 空客330(大)	08:00 10:10	首都机场T2 虹桥机场	89% 9分钟	东航银联旗舰店 ¥450 直接订	¥417 3.7折	订票
 东方航空MU5183 空客321(中)	08:05 10:20	首都机场T2 浦东机场	约100%	东航银联旗舰店 ¥450 直接订	¥417 3.7折	订票

后端优化 – 设计优化

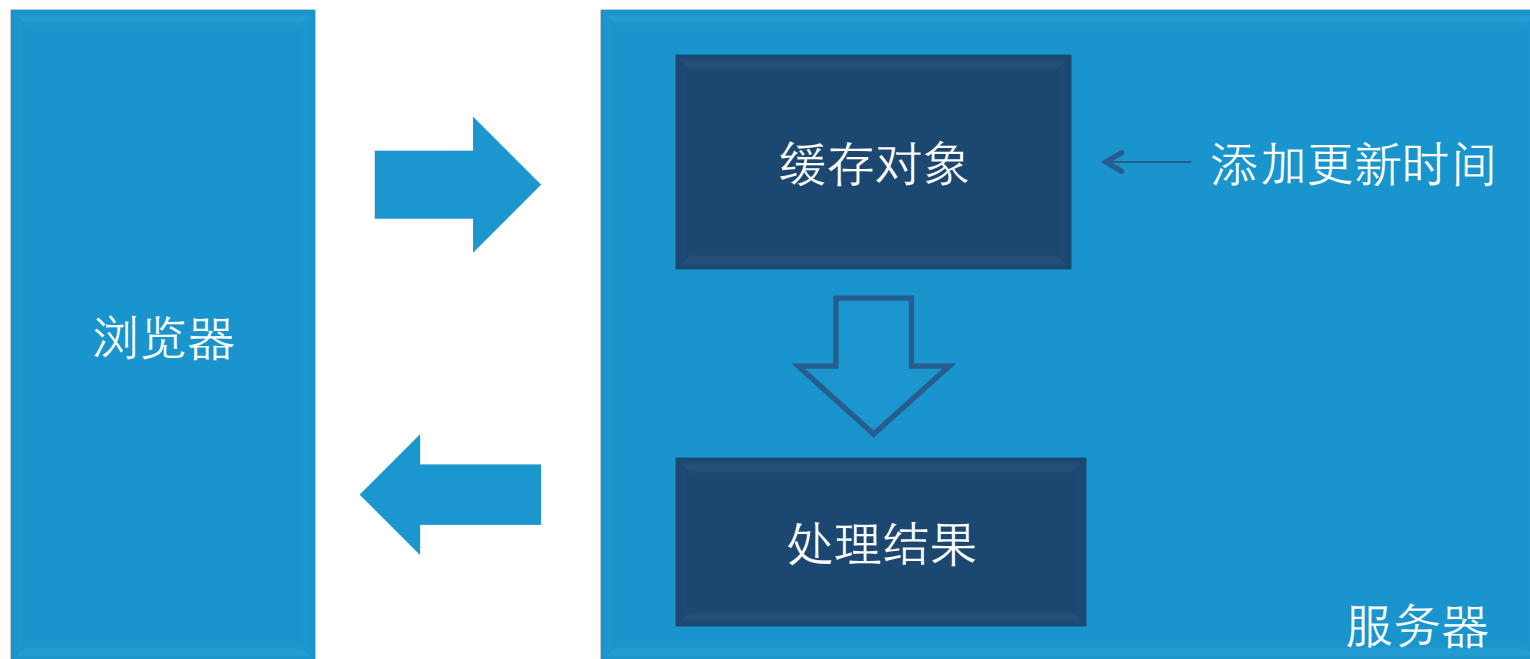
前端数据按需返回

 上海航空FM9108 波音737(中)	07:35 09:30	首都机场T2 虹桥机场	约100%	¥412 3.6折	订票
飞行1小时55分钟 燃油税: 140元 机场建设费: 50元 有餐食 有网上值机					
 芒果网	★★★★☆ 来自2155位用户	点评	 7×24服务  免费邮寄  退改签	¥450 4.0折	预订 55分钟前更新
 恒通之旅	★★★★☆ 来自287位用户	点评	 担保通  7×24服务  退改签	¥429 +20(保险) ¥444	预订 预订
 逍遥行商旅网	★★★☆☆ 来自1493位用户	点评	 担保通  7×24服务  退改签	¥427 +20(保险) ¥442	预订 预订
 阿凡提旅行网	★★★☆☆ 来自654位用户	点评	 担保通  退改签	¥412 +20(保险) ¥427	申请 申请

动态调整获取频率



避免不必要的计算

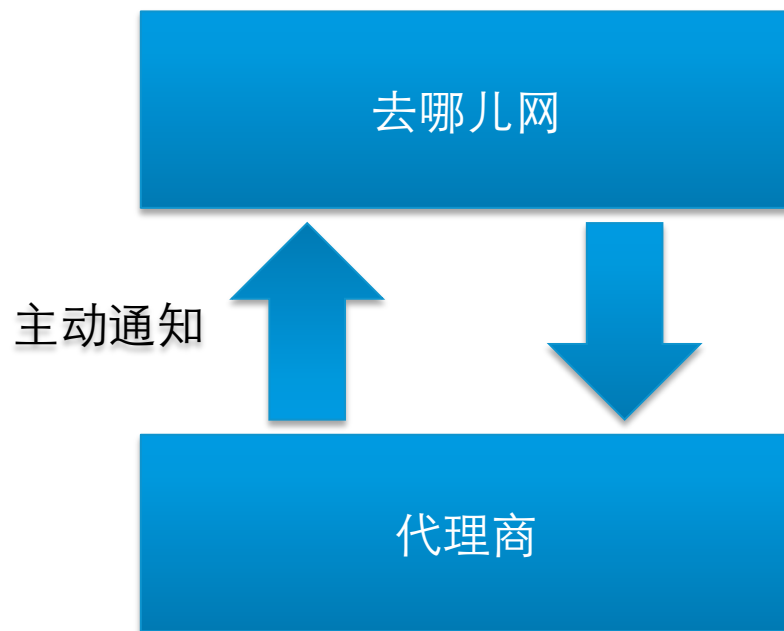


数据闭环

轮询模式



数据闭环



感谢

欢迎加入我们!

Thank You!

hao.lin@qunar.com
lianchun.liu@qunar.com