

客户端网站性能 及用户行为数据采集

携程旅行网 储诚栋



储诚栋 高级研发经理

携程旅行网 ， 框架研发部

- 前端框架
- 客户端性能优化
- 用户行为数据采集

cdchu@ctrip.com



Agenda

- **Overview of Data Collection**
- **Best Practices in Ctrip**
 - Tracking Architecture and Deploy
 - Data Visualization
 - Alerting
 - Trouble Shooting
- **More Usage**
- **Q&A**



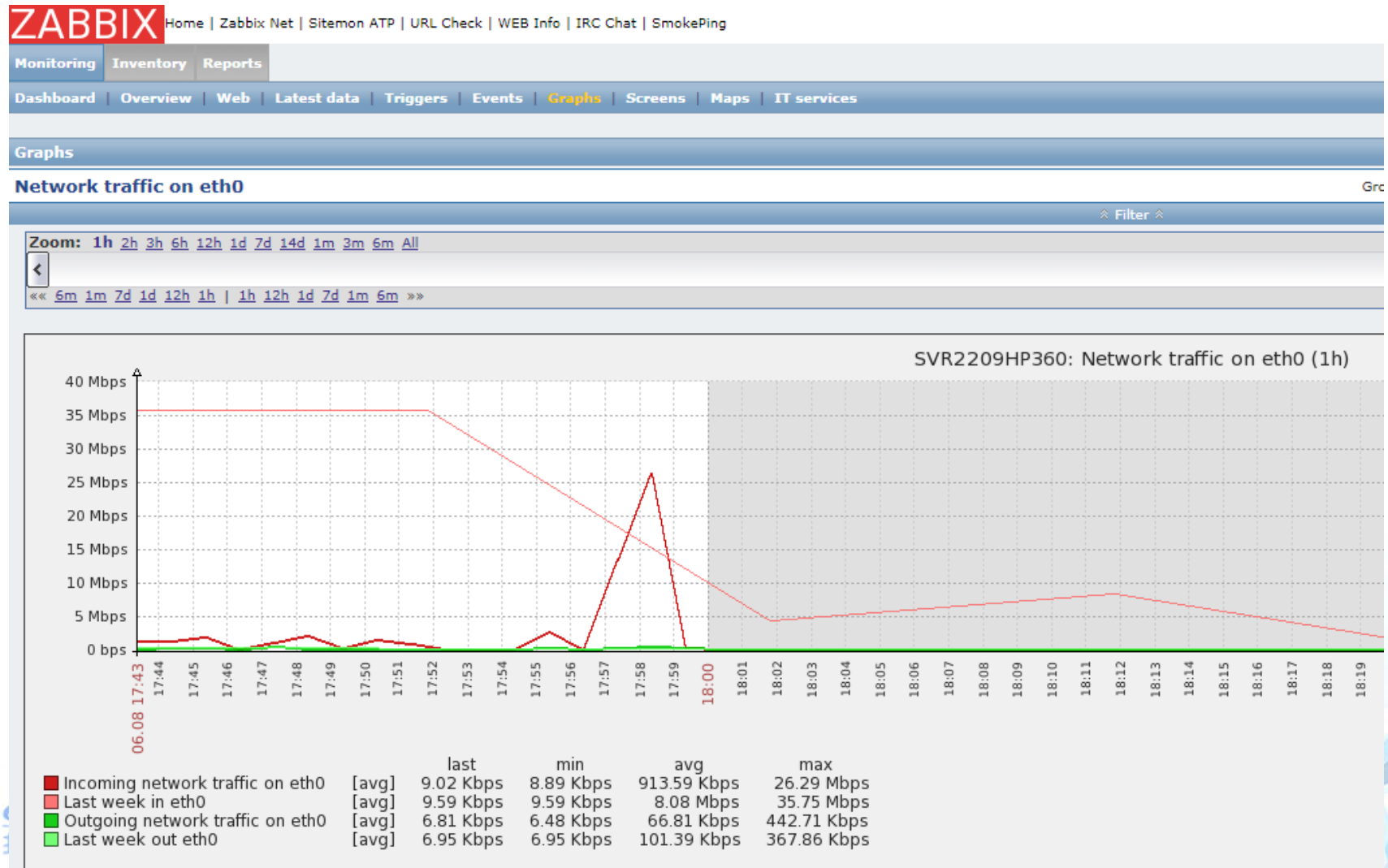


Overview of Data Collection



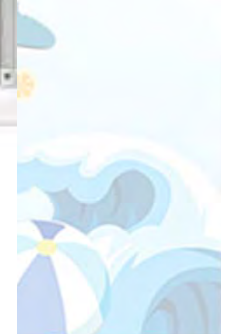
Server-side

zabbix



Pseudo Client

Gomez



Pseudo Client

Ctrip Page Performance Test

The screenshot displays the 'Page Performance Test' interface. The main content area shows a report for 'Ctrip-QA规范' (Ctrip-QA Standard) for the '国内酒店首页' (Domestic Hotel Homepage). The report includes the following information:

- 国内酒店首页 创建人: shaohanwang 创建时间: 13/07/31
- 地址: <http://hotels.ctrip.com/>
- 浏览器选择: IE6, IE7, IE8, IE9, Chrome, Firefox
- 总评分: **74**
- 高优先级 (High Priority): 图片需要压缩 (Images need compression)

The '高优先级' section details the image compression issue:

- 有28个图片需要压缩 (28 images need compression)
- 请使用此工具进行优化 http://192.168.81.77:3000/#css_compress
- 示例链接:
 - <http://images3.ctrip.com/ad/bigbutton/2013/06/summer.png>
 - <http://images3.ctrip.com/ad/banner/2012/11/3006jingq.png>

On the right side, a list of optimization suggestions is provided:

- 图片需要压缩 (Images need compression)
- HTML标签符合W3C规范 (HTML tags conform to W3C standards)
- jshint校验js代码 (jshint validates js code)
- 混淆CSS和JS代码 (Minify CSS and JS code)
- 使用CSS和JS外链 (Use external CSS and JS)
- 减少DNS的查询 (Reduce DNS queries)
- 避免URL重定向 (Avoid URL redirection)
- 使用无Cookie的域 (Use cookieless domains)
- 使favicon小且缓存 (Make favicon small and cacheable)
- 外链资源使用了Gzip (External resources use Gzip)
- 无404请求 (No 404 requests)
- CSS放置于顶部 (Place CSS at the top)
- JS放置于底部 (Place JS at the bottom)
- 删除重复的JS和CSS (Remove duplicate JS and CSS)

Real Client

Google Analytics





Tracking Architecture and Deploy

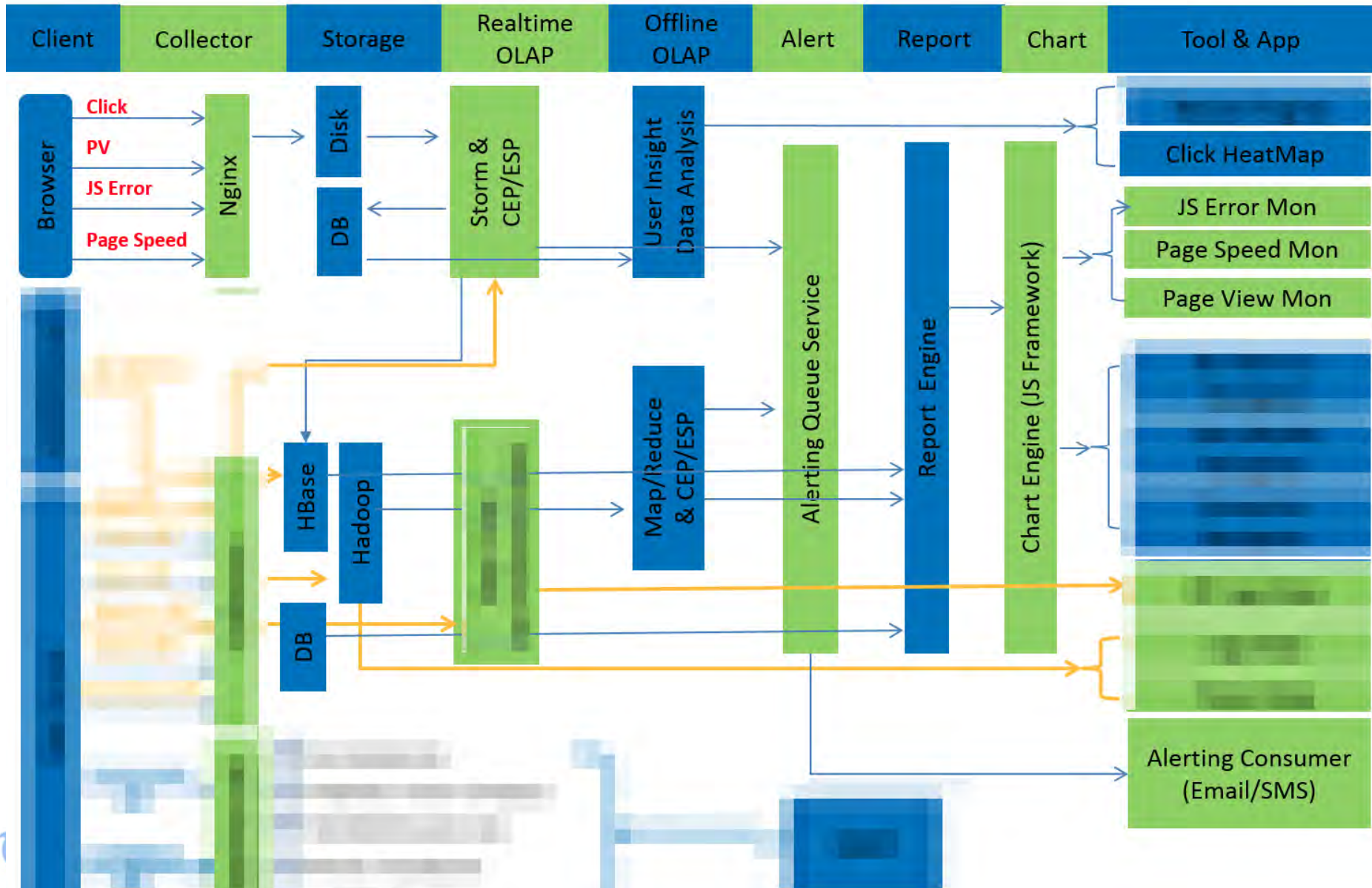
Best Practices in Ctrip

UBT (User Behavior Tracking)

- **Basic PVs**
- **User Actions**
- **Page Performance**
- **Error Tracking**
- **User Blocks**
- **Page View Flow**



Data Stream



Client-side Tracking - Methods

- **Transfer page** - weibo, edm

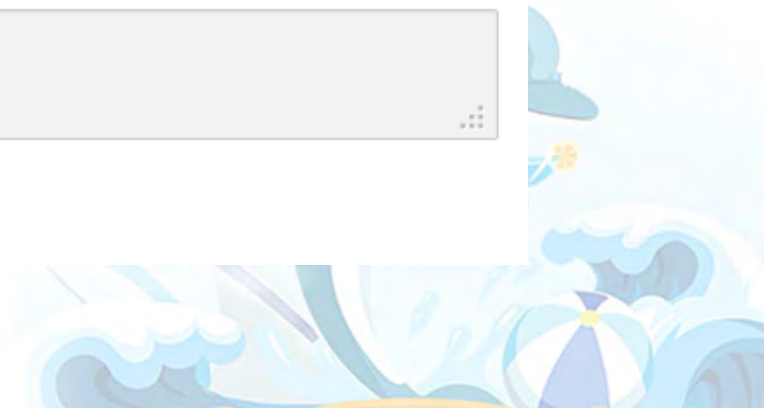
Url Generator

Namespace:

Jmpurl:

Url:

Weibo Url:



Client-side Tracking - Difficulties

- **Data Loss**

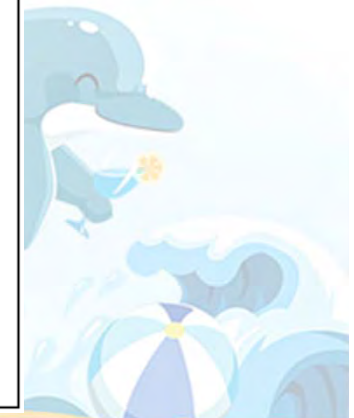
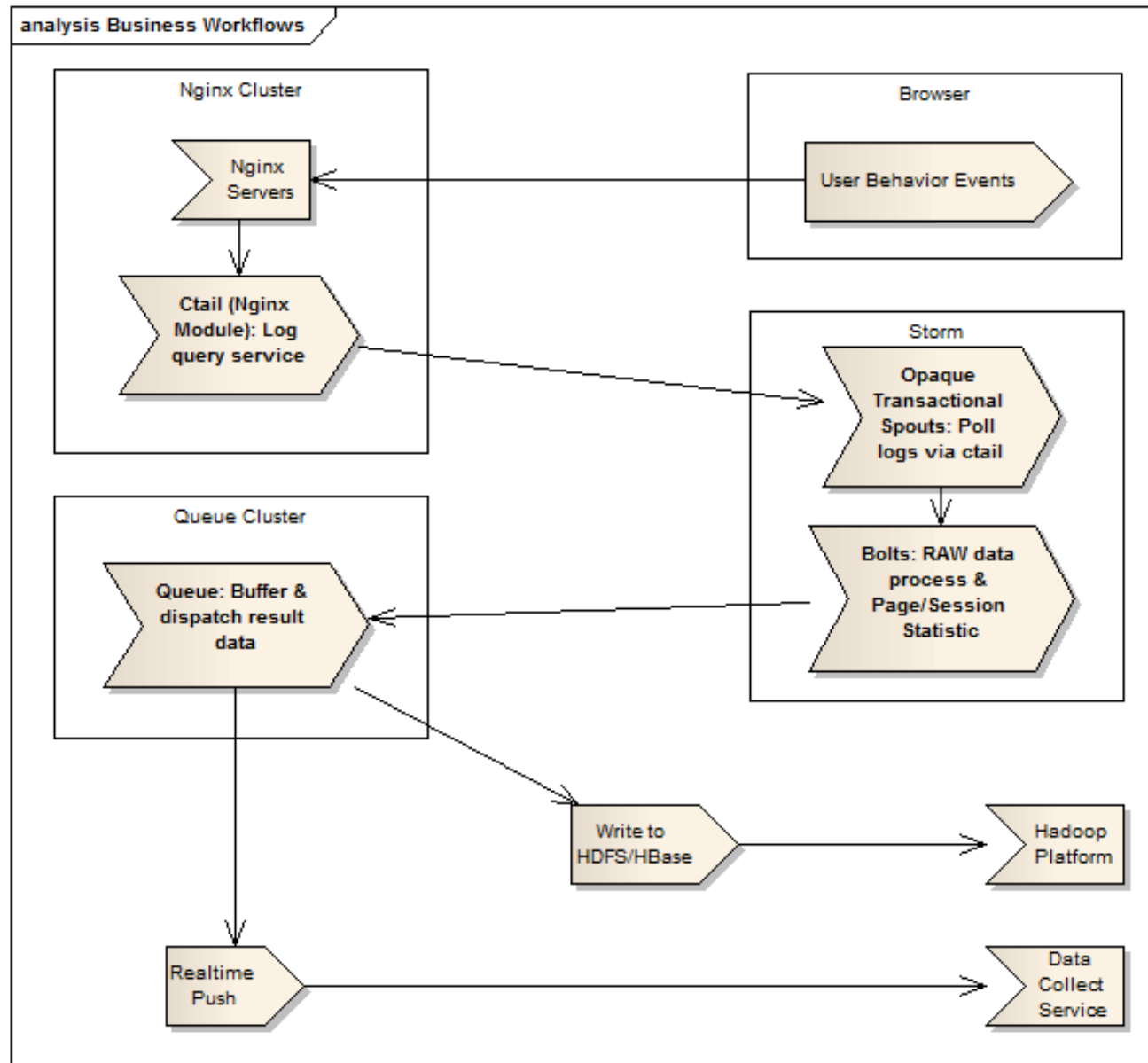
- 15% -> 5% -> 2%
- onload send data -> more triggers -> send queue

- **Identify**

- dom: marks + xpath
- position: center, left or auto adjust

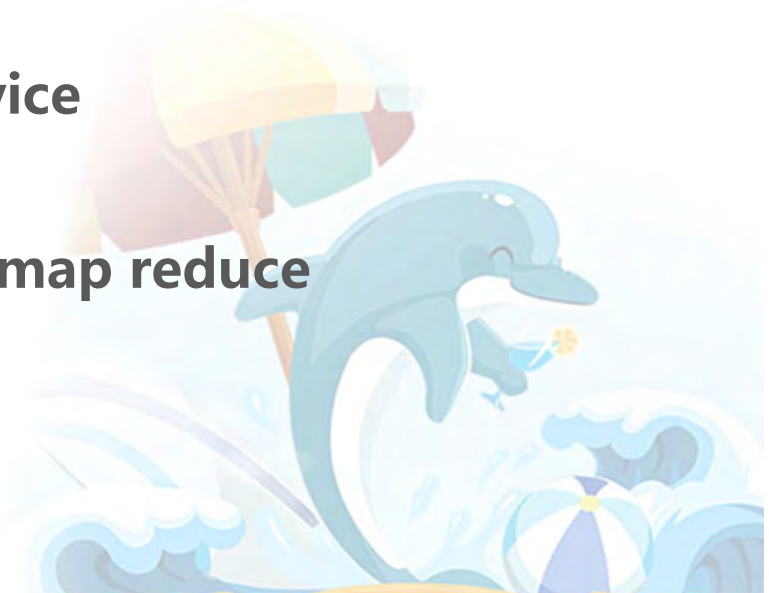


Server-side Tracking



Server-side Tracking

- **Data Landing** - nginx log
- **Queue & Classify** - activemq -> ctail
- **Data Analysis**
- **Real-time** - storm & push service
- **Non-real-time** - hadoop & map reduce





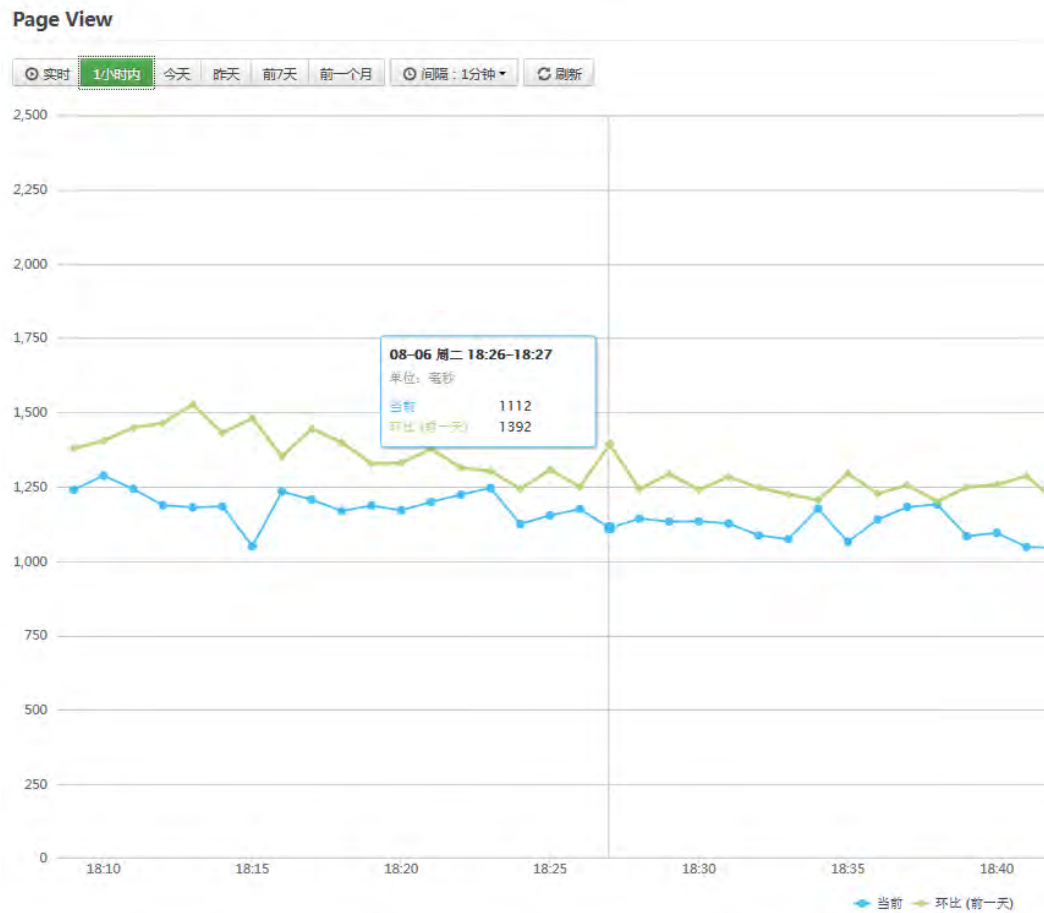
Data Visualization

Best Practices in Ctrip



Base PV

- PV
- UV
- CR
- Etc.



转化率

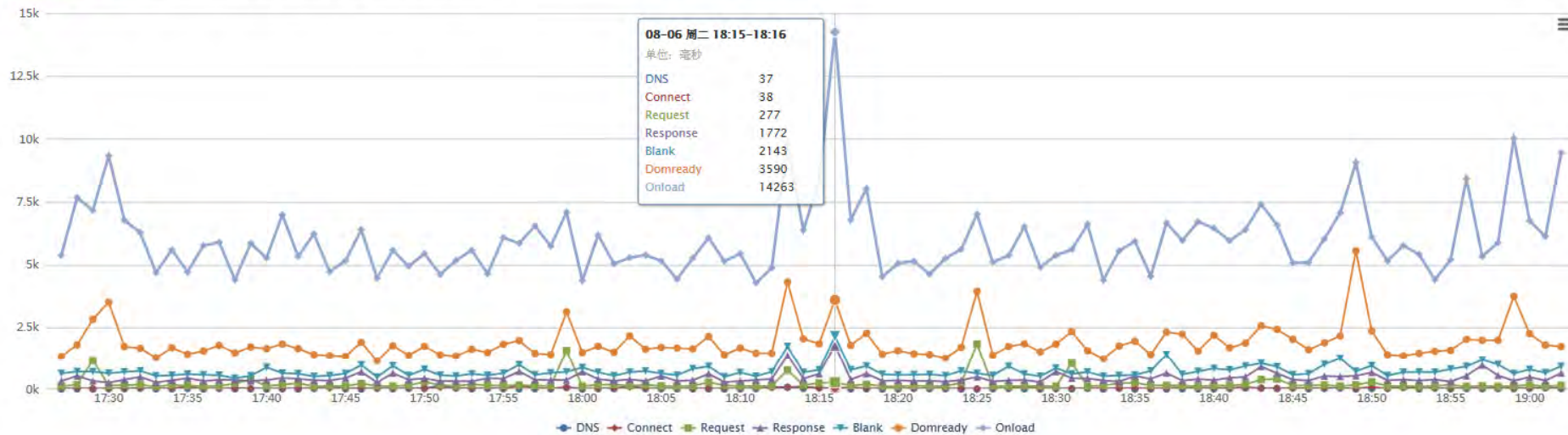


Performance

Performance

名词帮助

实时 1小时内 今天 昨天 前7天 前一个月 间隔: 1分钟 刷新



网络	延迟
移动	7286ms
网通	6045ms
铁通	5976ms
电信	5732ms
有线通	5316ms
广电网	5304ms
联通	4986ms
科技网	4743ms
长城宽带	3961ms
教育网	1958ms

地区	延迟
西北地区	7427ms
西南地区	6649ms
华南地区	6443ms
东北地区	6340ms
华北地区	6048ms
华中地区	5638ms
港澳台地区	4906ms
华东地区	4815ms

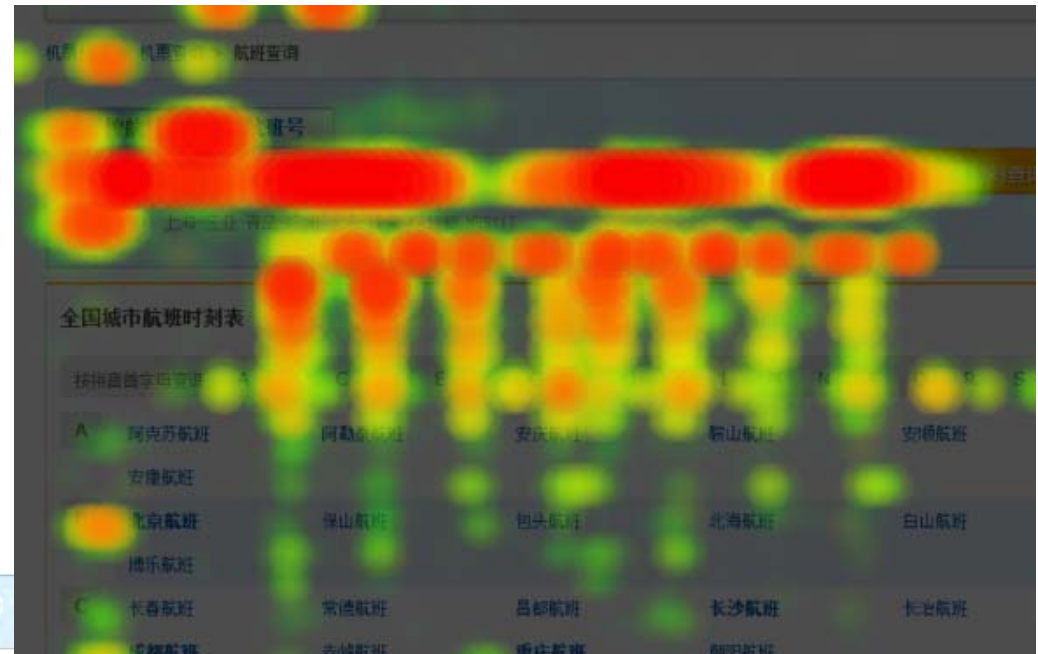


柳州	36087ms	景德镇	11614ms	遵义	9125ms
自贡	28453ms	三亚	11098ms	澳门	9123ms
百色	25010ms	佳木斯	10330ms	大连	9057ms
呼伦贝尔	19902ms	博乐	9997ms	伊犁	9038ms
庆阳	13304ms	克拉玛依	9599ms	唐山	8867ms
库尔勒	12519ms	敦煌	9523ms	张掖	8841ms
延吉	12282ms	湛江	9481ms	十堰	8732ms
蚌埠	12275ms	乐山	9374ms	中卫	8625ms
银川	12151ms	邢台	9302ms	乌鲁木齐	8554ms
通辽	11722ms	哈密	9255ms	运城	8467ms



User Actions

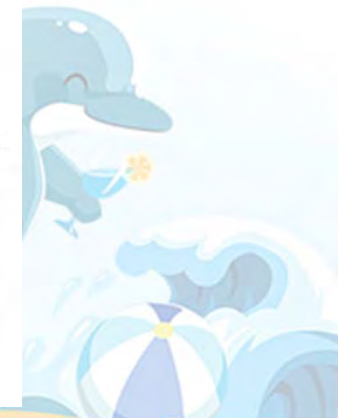
- **Types**
 - Click
 - Key press
 - Touch
 - Tab change



The screenshot shows a flight search interface with two data analysis overlays. The main interface displays a search bar, a grid of flight options categorized by letter (A, B, C), and a '高级查询' (Advanced Search) button. The first overlay, titled '搜航线', shows data for a search on '航线' (Route). The second overlay, titled '安澜航班', shows data for a specific flight.

搜航线	航线	click	click_pv	display_pv(page)
	航线	2,087	1,788	14,377
		4.30%	15.06%	12.43%

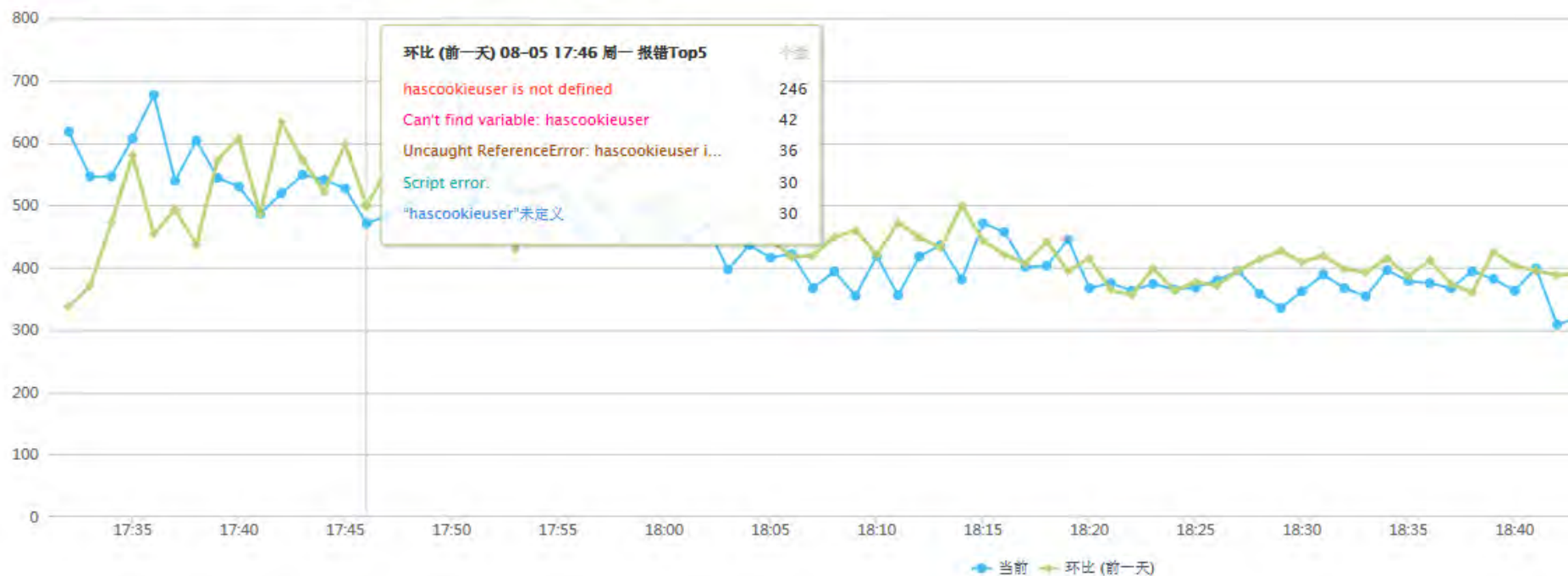
安澜航班	click	click_pv	display_pv(page)
	7	7	14,377
	0.01%	0.05%	0.04%



Error Tracking

JS Error

实时
1小时内
今天
昨天
前7天
前一个月
间隔: 1分钟
刷新



浏览器	错误文件	错误信息	网络
Chrome 64.57%	http://flights.ctrip.com/WebResourceOnline/JavaScript/ShowFare... 74.39%	hascookieuser is not defined 55.65%	电信 55.73%
Safari 11.09%	http://webresource.c-ctrip.com/code/cquery/LABjs/LAB.js 13.19%	Can't find variable: hascookieuser 12.40%	网通 38.81%
ie8.0 8.70%	http://flights.ctrip.com/WebResourceOnline/JavaScript/FlightSear... 8.43%	Uncaught ReferenceError: hascookieuser i... 10.57%	移动 3.07%



User Blocks

- Form tips
- Business constraints
- Security problems
- Third-party issues
- Server maintenance
- Etc.



Page View Flow

返回 Performance 用户浏览跟踪

用户信息 vid : 1348477974968.sex1qx sid : 2 浏览器 : MSIE 9.0

浏览时间 : 08-07 00:34:16 至 08-07 01:36:39 时长 : 62.38m 期间 4 次新打开页面

返回 JS Error 用户浏览跟踪

用户信息 vid : 1375236596217.lflvx4 sid : 1 未登陆 浏览器 : MSIE 8.0 城市 : 成都 网络 : 电信

浏览时间 : 07-31 10:09:58 至 07-31 10:11:02 时长 : 1.07m 期间 1 次新打开页面

10:09:58 新打开, 浏览 4 页面

1 新版全站首页

访问时间	行	登录	时长	错误文件	错误信息
2013-07-31 10:10:07	1	否	8841	http://uma.gtags.net/s?id=208&t=1&r=http%3A%2F%2Fwww.ctrip.com%2F%3Fallianceid%3D5341%26sid%3D128924%26oid%3D000401app-&cids=&_=0.1897871673950508	对象不支持此属性或方法

16.00s

频道首页

34.00s

第一程航班搜索结果页

1.07m

第一程航班搜索结果页

8.33m

新版全站首页

Response: 3758ms

8.78m

第一程航班搜索结果页

9.15m

填写预订信息页

9.83m

第一程航班搜索结果页

11.63m

频道首页

12.43m

搜索无结果页





Alerting

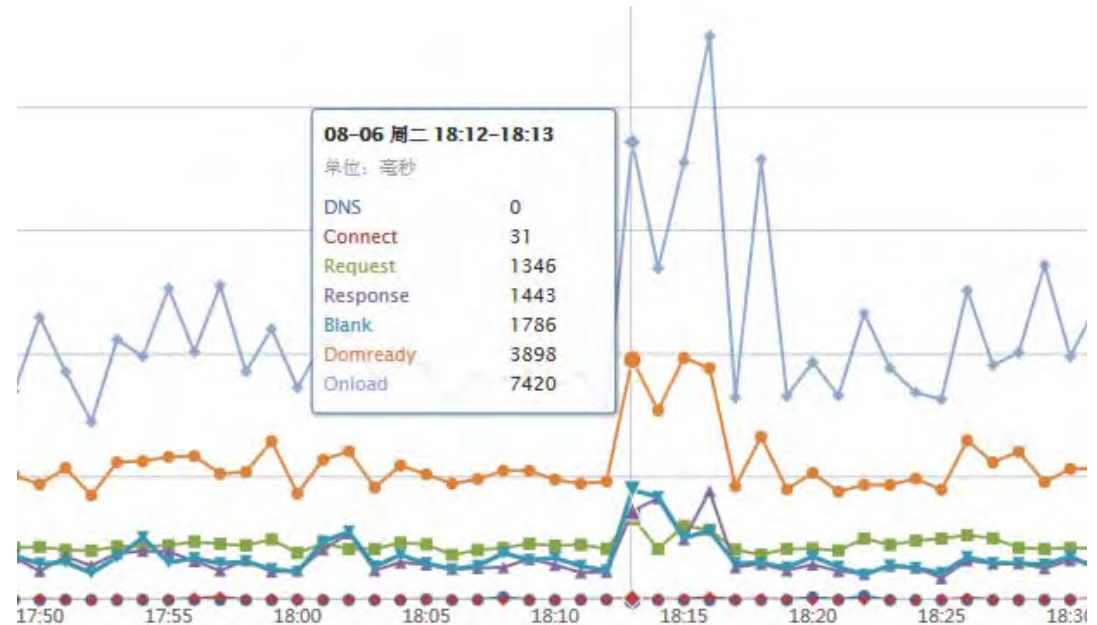
Best Practices in Ctrip



Unseasonal

- **Examples**

- Page speed
- Js errors per page
- User blocks per page



- **Solutions**

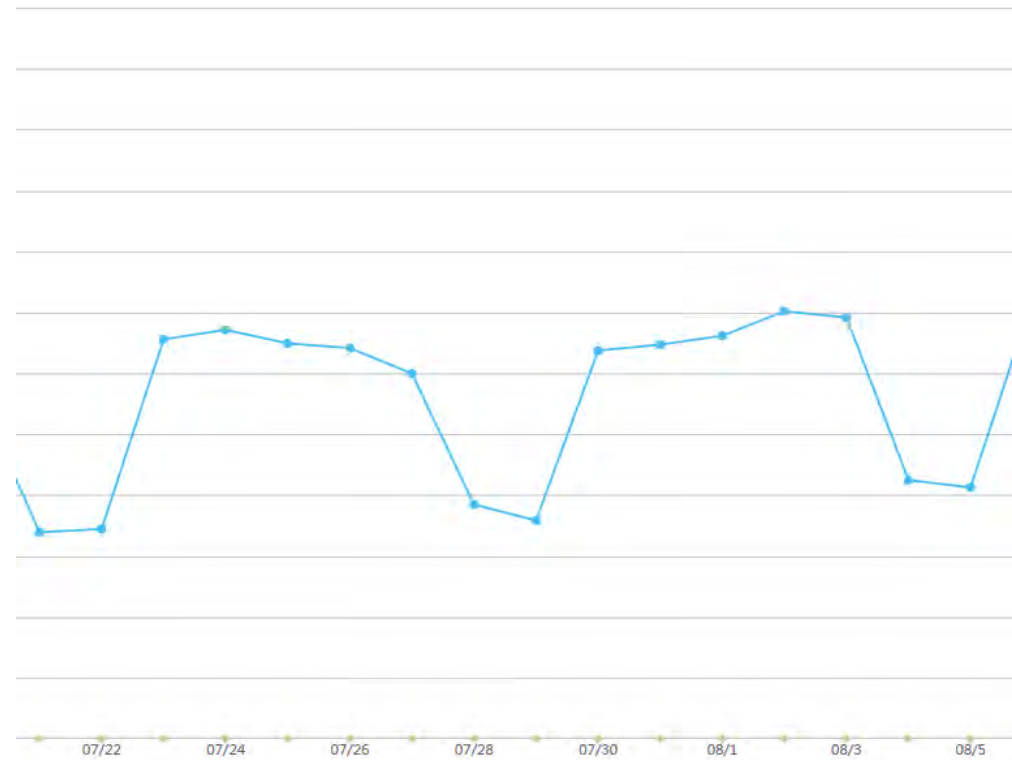
- Base line
- Moving average
- Control chart / CPK (Process capability index)



Seasonal Fluctuations

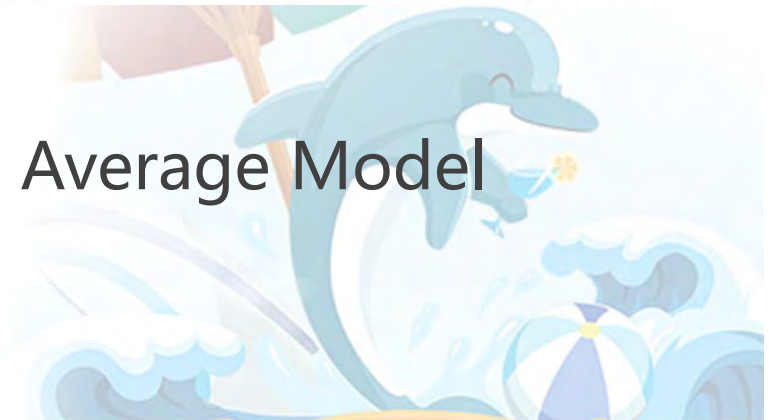
- **Examples**

- PV
- UV
- CR



- **Solutions**

- Month-on-month
- Year-on-year
- Auto-Regressive and Moving Average Model



Site Health and Page Rank

- **Examples**
 - Rank: A / B / C / D / E
 - Score: 0 – 100
- **Solutions**
 - Weighted arithmetic mean

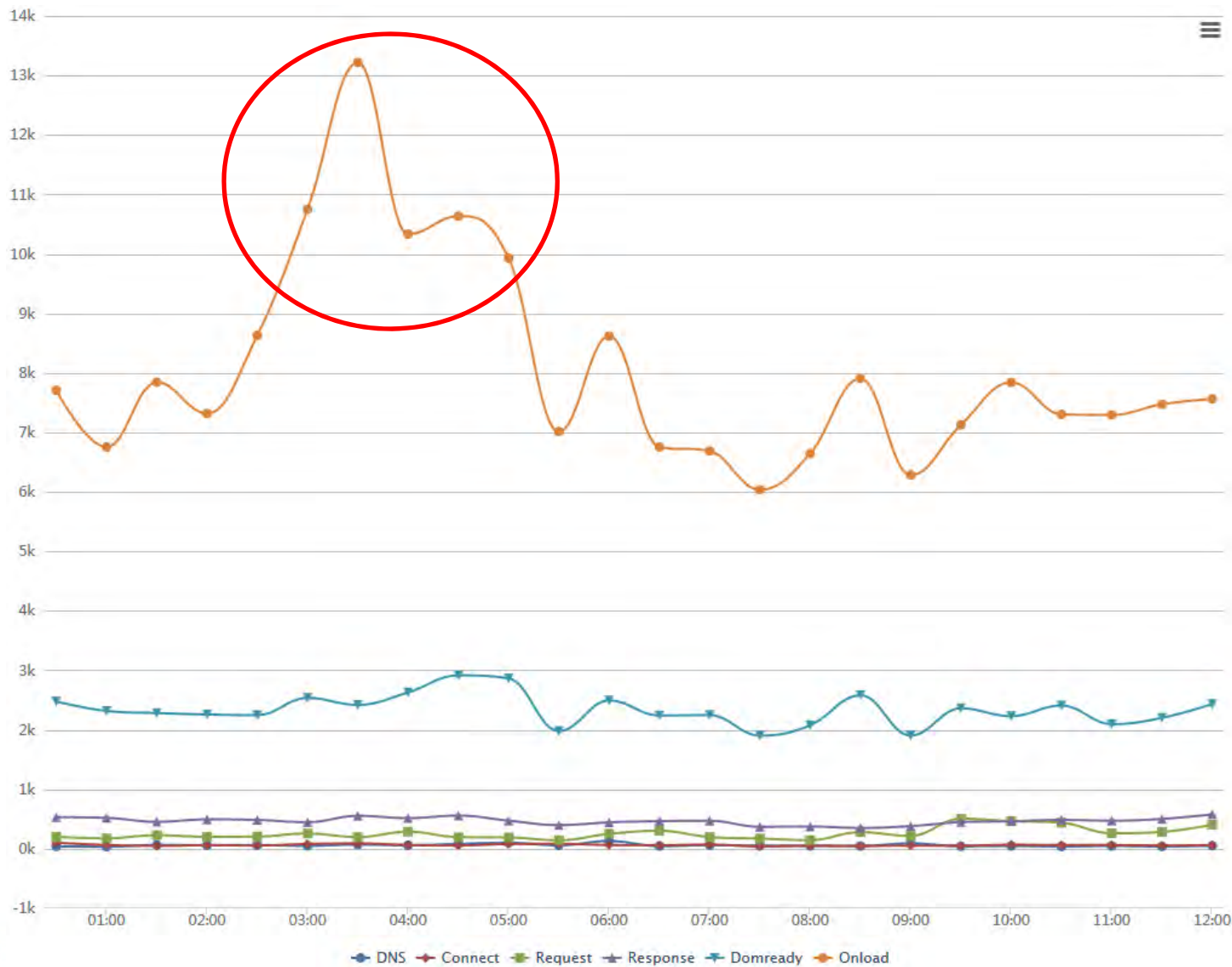




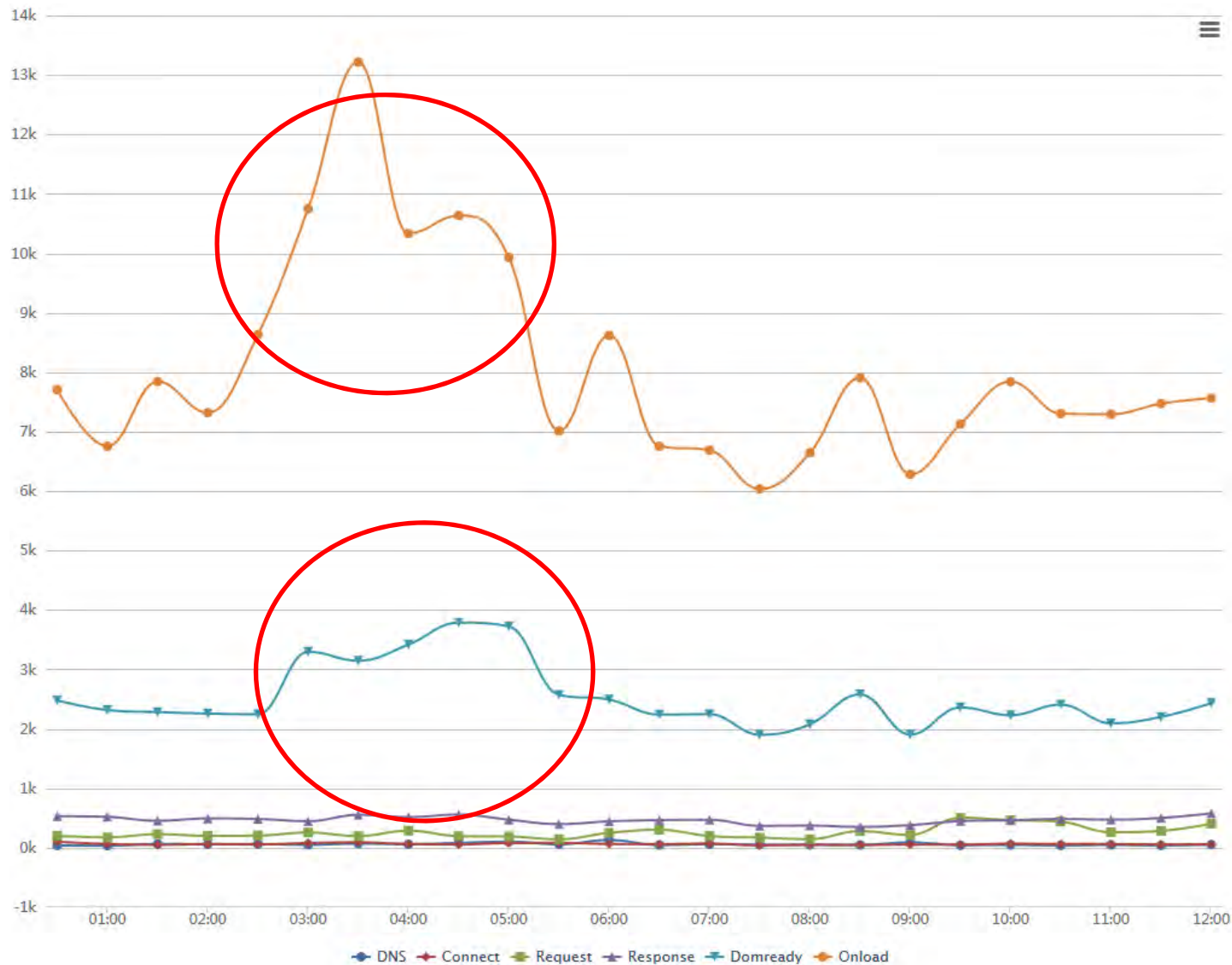
Trouble Shooting

Best Practices in Ctrip

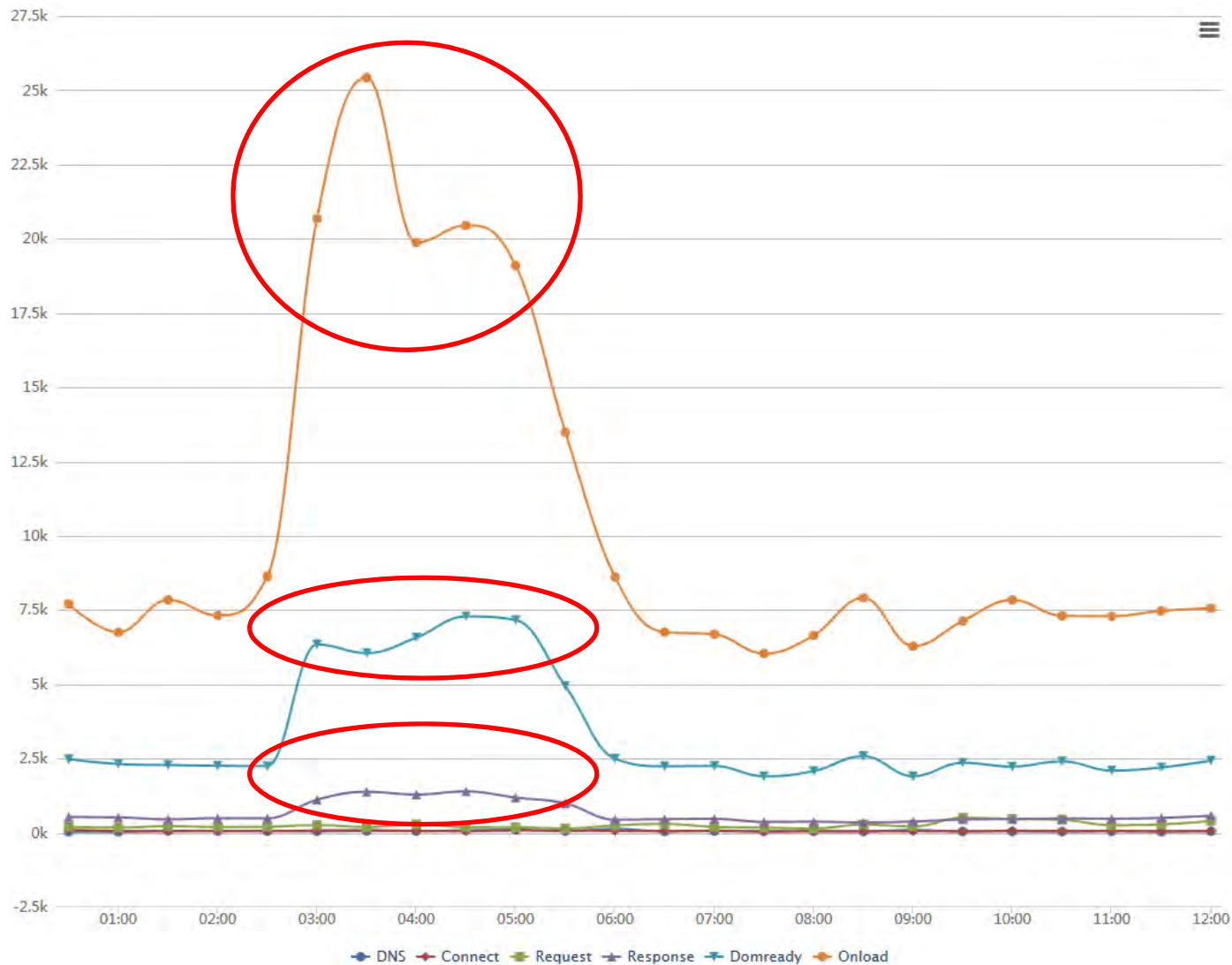
Performance (Case 1)



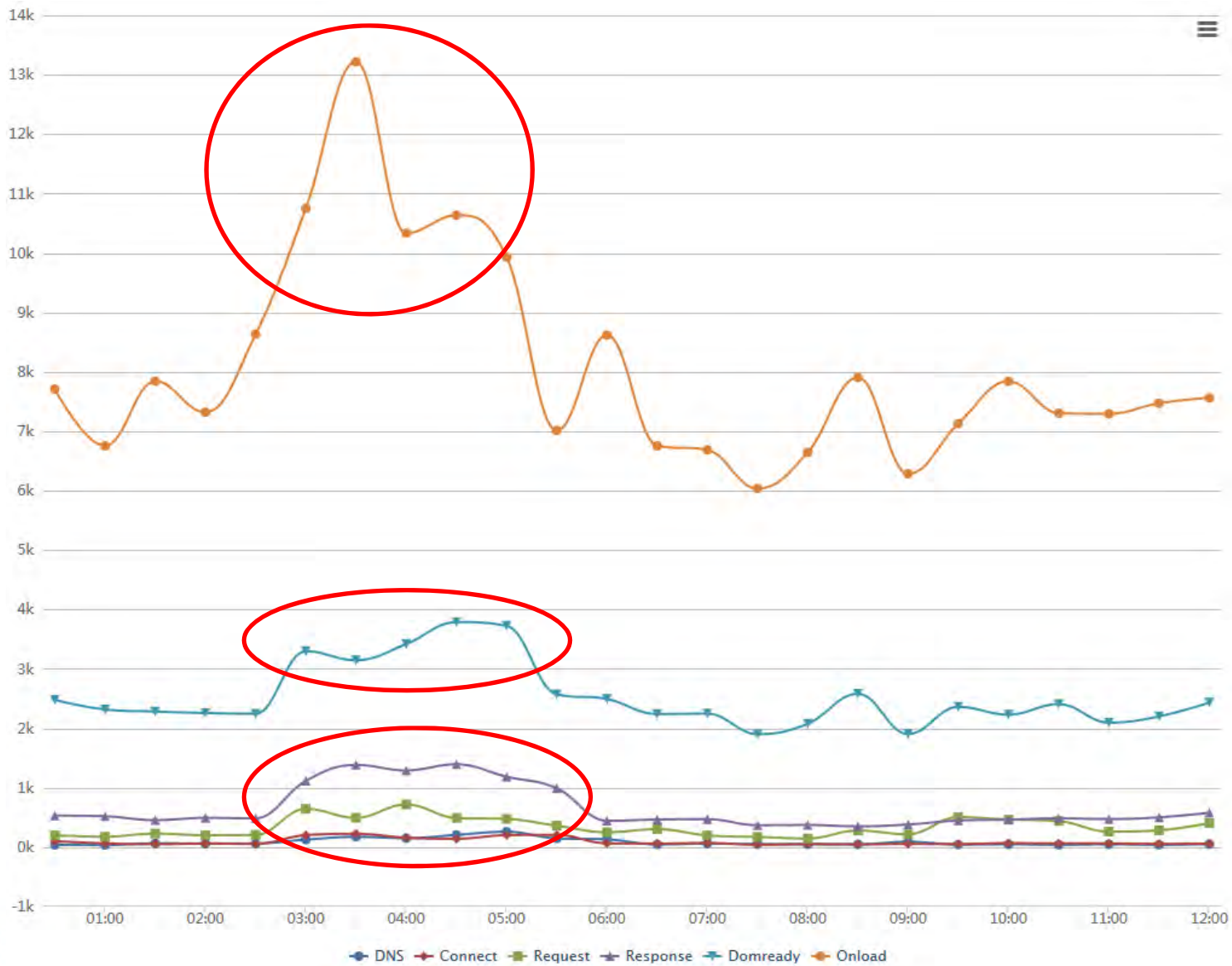
Performance (Case 2)



Performance (Case 3)



Performance (Case 4)





More Usage

More Usage

- **Client data associate with server data**
- **Real-time product recommendation**
- **Intelligent resource preload**
- **Data combine**
- **Etc.**



Q&A

储诚栋 高级研发经理
携程旅行网， 框架研发部

- 前端框架
- 客户端性能优化
- 用户行为数据采集

cdchu@ctrip.com

