

阿里巴巴在线技术峰会  
Alibaba Online Technology Summit

# 基于大数据的全球电商系统性能优化

A Holistic Attribution Model for  
Performance Optimization

Dongbai Guo , PhD



# 商业背景

# Some Background

# Alibaba's Marketplace Partition

## Domestic

## International

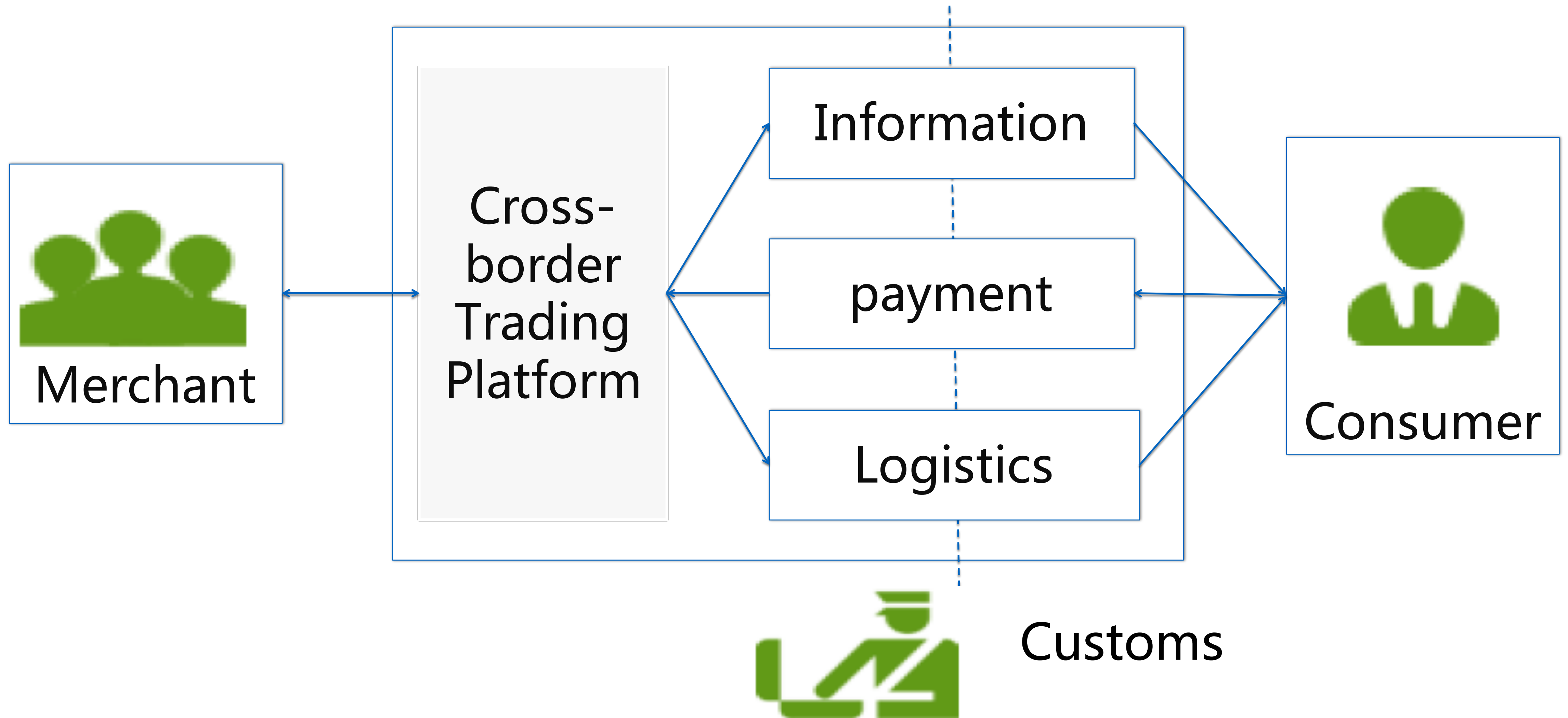
2C



2B



# Cross Border and Local Trade



# Some Numbers

**Global Ranking**  
Alexa #25

**2015.11.11**

**Volume**

2200 W orders

**Daily Variety**

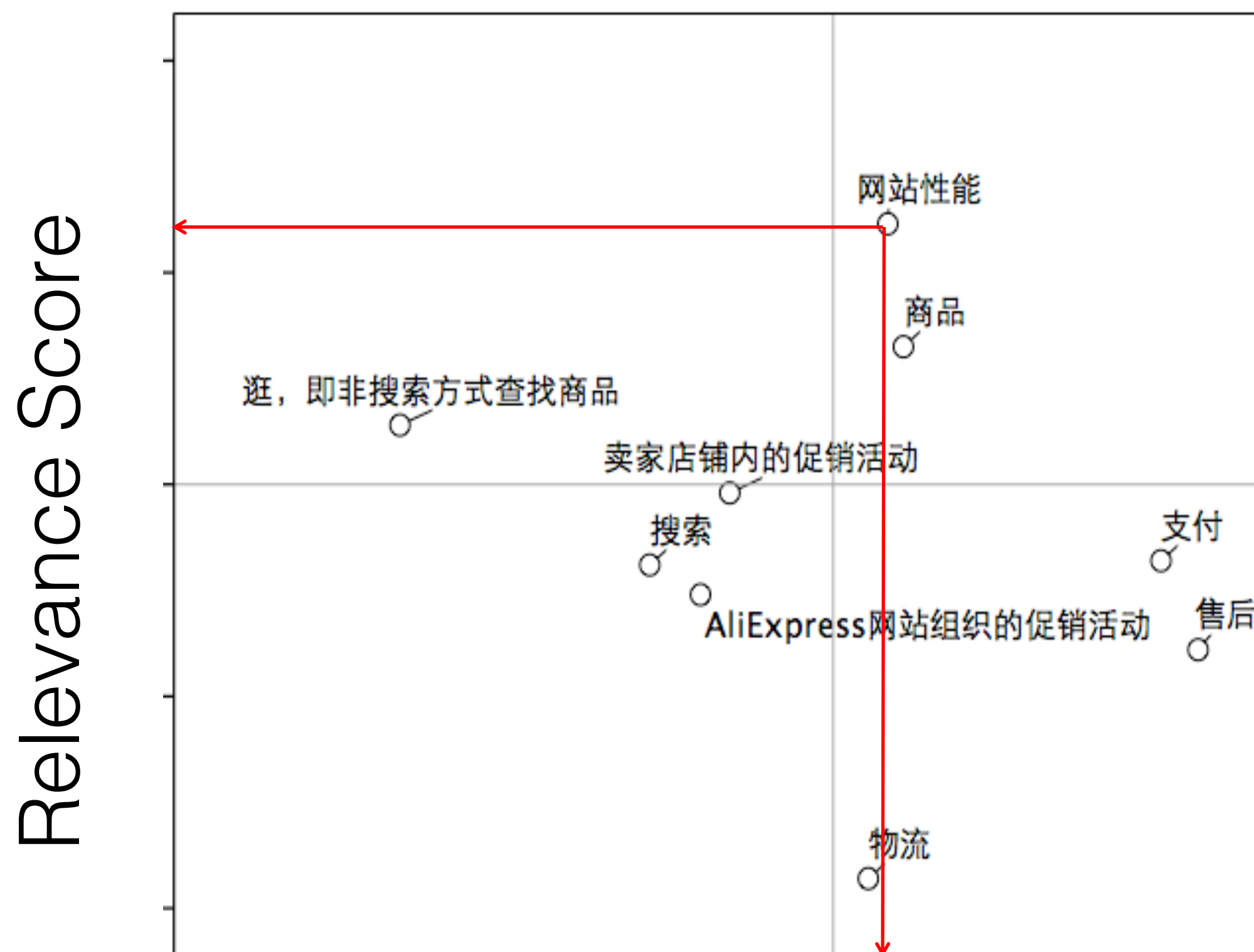
9000 devices

214 countries

16 languages

The image displays the AliExpress mobile application interface across three devices: iPhone, iPad, and Android. The iPhone screen shows a 'NEW LOOK' banner with 'up to 70% off' and a 'NEW TO THE APP?' notification. The iPad screen features a 'BUNDLE UP IN STYLE' banner for winter fashion with 'From 40% off'. The Android screen displays a list of clothing items, including a '9 Colors 2015 Spring Autumn Women Blouse Pullover' for US \$7.90. Below the device screens are download instructions for the App Store and Google Play. The main desktop interface includes a search bar, navigation tabs (SuperDeals, Featured Brands, etc.), a category list on the left, and a 'TRENDING STYLES' banner with a 'Shop Now' button. A 'LUXURY WATCHES FOR HIM' advertisement for a GUANQIN watch is also visible, with a 'Shop Now' button and social media icons.

# Why Do We Care Performance ?



NPS score from AliExpress Spain

## Landing page load time will soon be incorporated into Quality Score

Thursday, March 06, 2008

As part of our continuing efforts to improve the user experience, we will soon incorporate an additional factor into **Quality Score**: landing page load time. Load time is the amount of time it takes for a user to see the landing page after clicking an ad.

*Why are we doing this?*

Two reasons: first, users have the best experience when they don't have to wait a long time for landing pages to load. **Interstitial pages**, multiple redirects, excessively slow servers, and other things that can increase load times only keep users from getting what they want: information about your business. Second, users are more likely to abandon landing pages that load slowly, which can hurt your conversion rate.

*When are we making this change?*

In the next few weeks, we will add load time evaluations to the **Keyword Analysis page** (we'll notify you when they are available). You will then have one month to review your site and **make necessary adjustments**.

After the one month review period, this load time factor will be incorporated into your keywords' Quality Scores. Keywords with landing pages that load very slowly may get lower Quality Scores (and thus higher **minimum bids**). Conversely, keywords with landing pages that load very quickly may get higher Quality Scores and lower minimum bids.

To learn more about the upcoming change, please see [this article](#) in the AdWords Help Center.

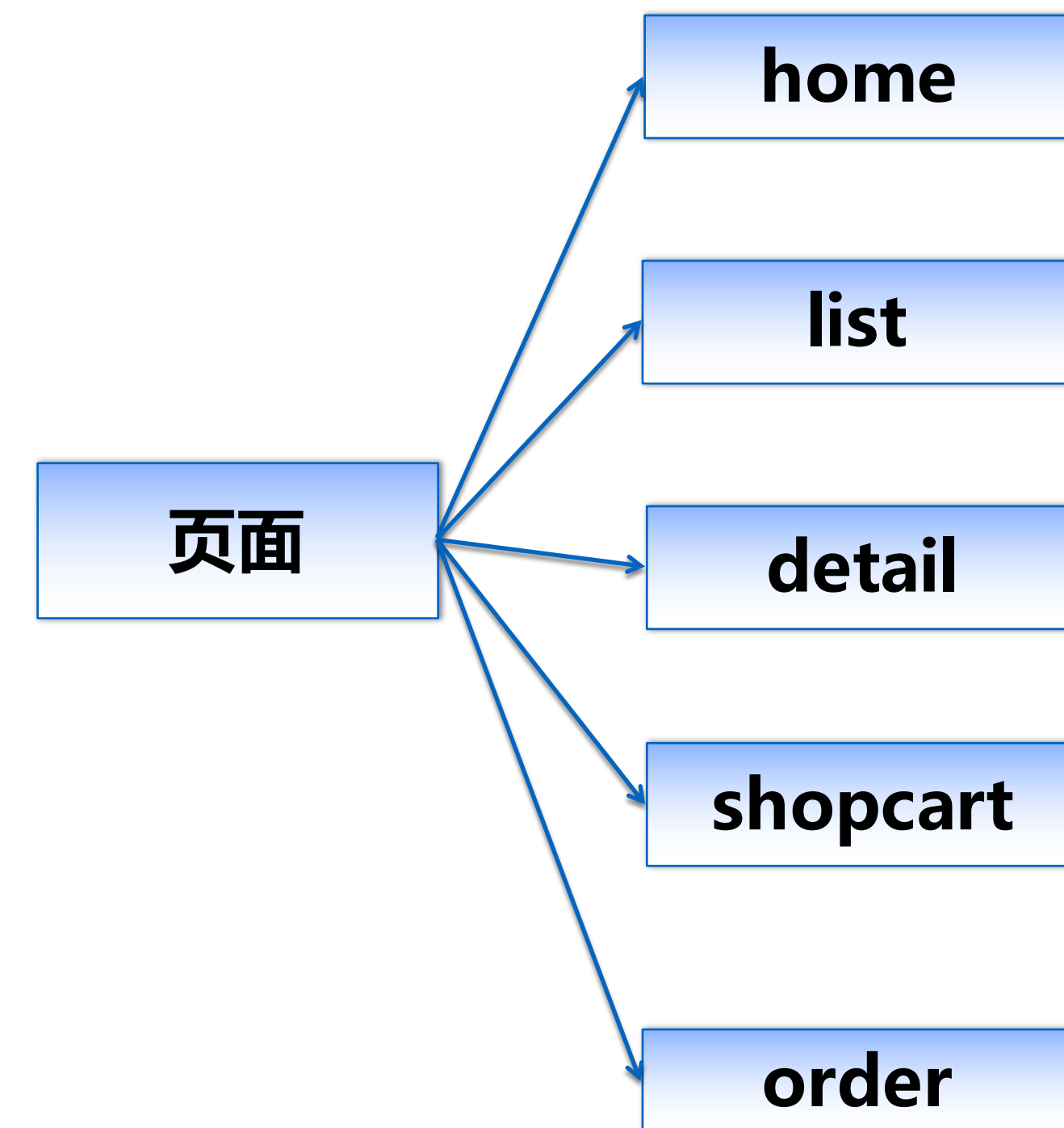
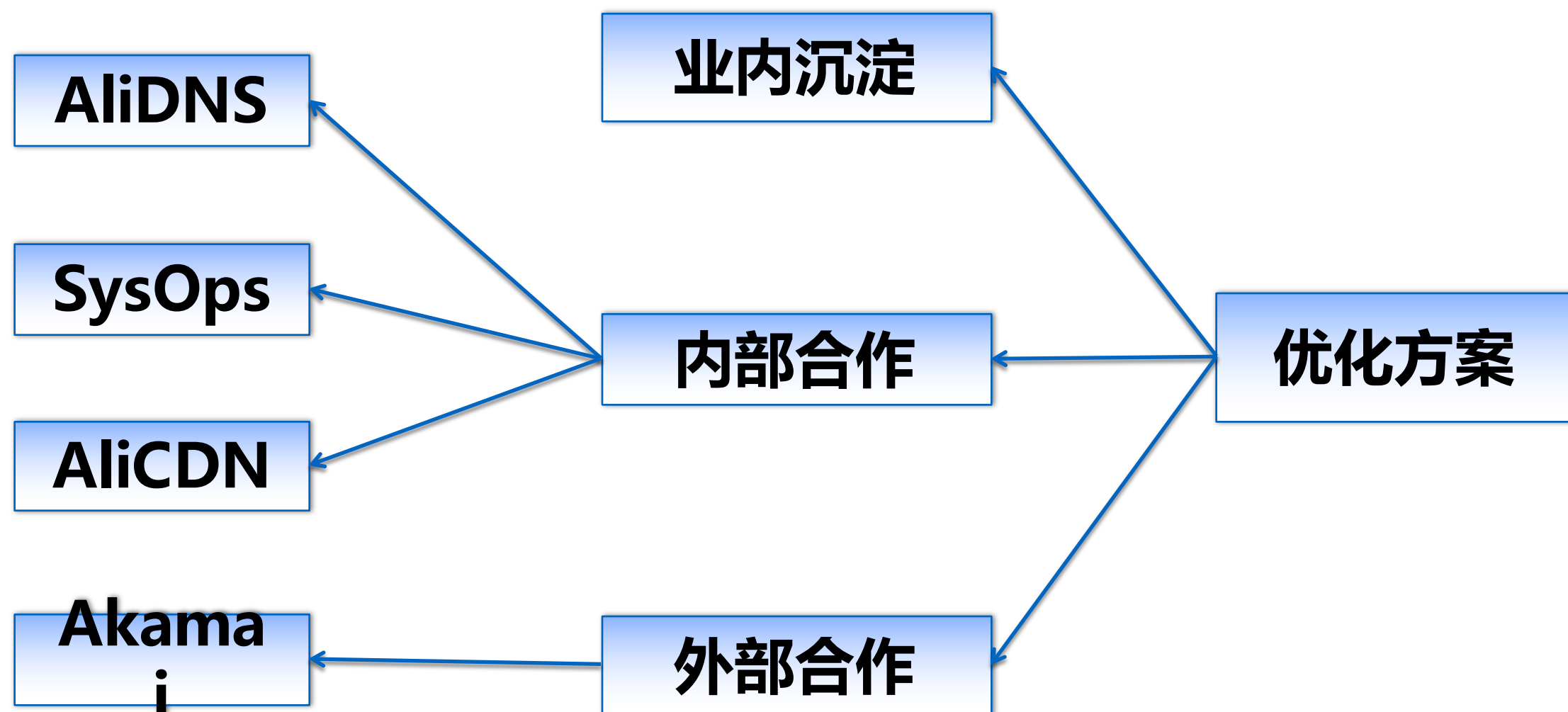
<https://adwords.googleblog.com/2008/03/landing-page-load-time-will-soon-be.html>

# Performance Battle – Ancient Warfare



Source : <http://pic2.duowan.com/dt2/1004/134144850170/134144935174.jpg>

# The Challenge



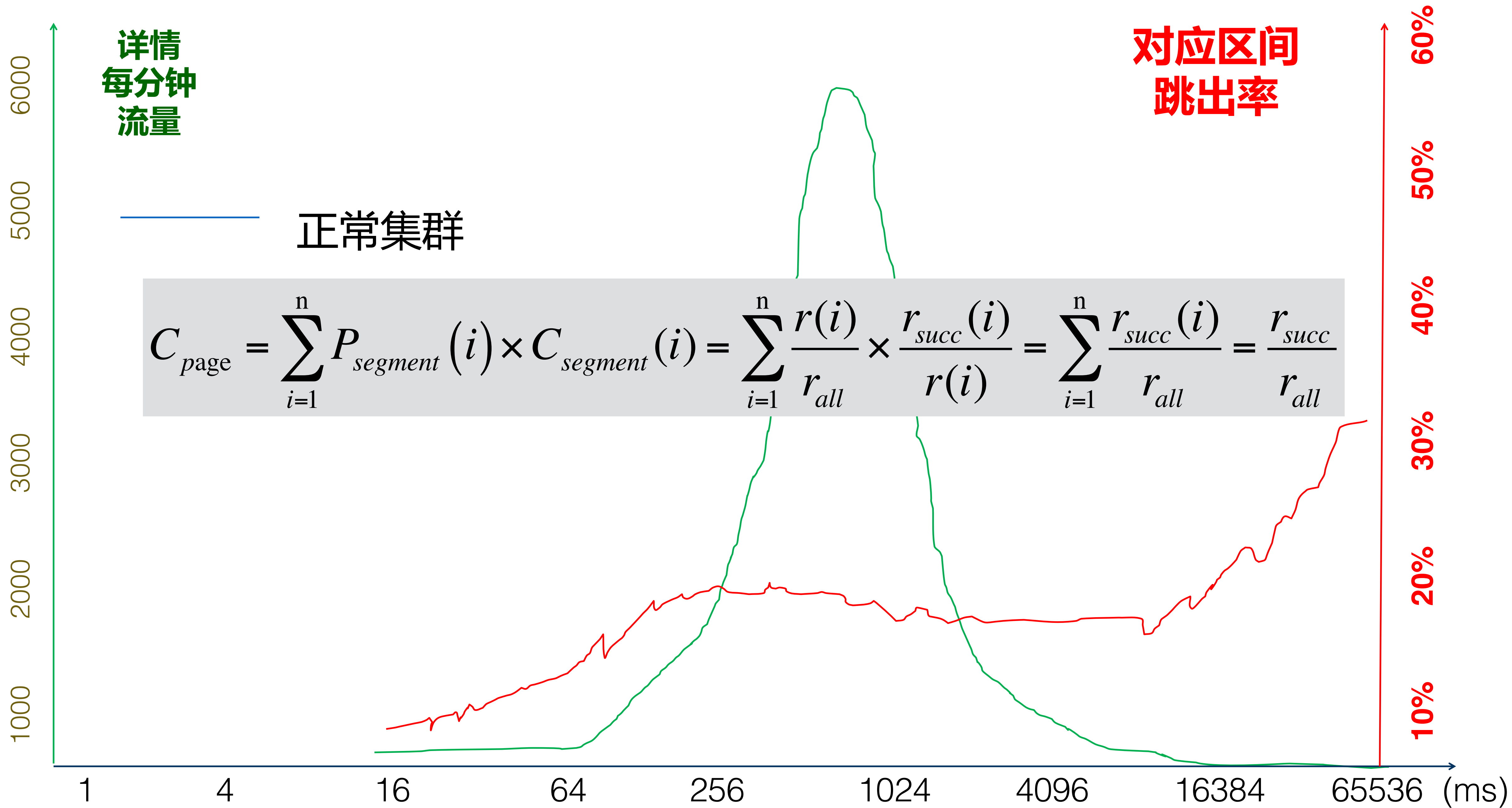
✓

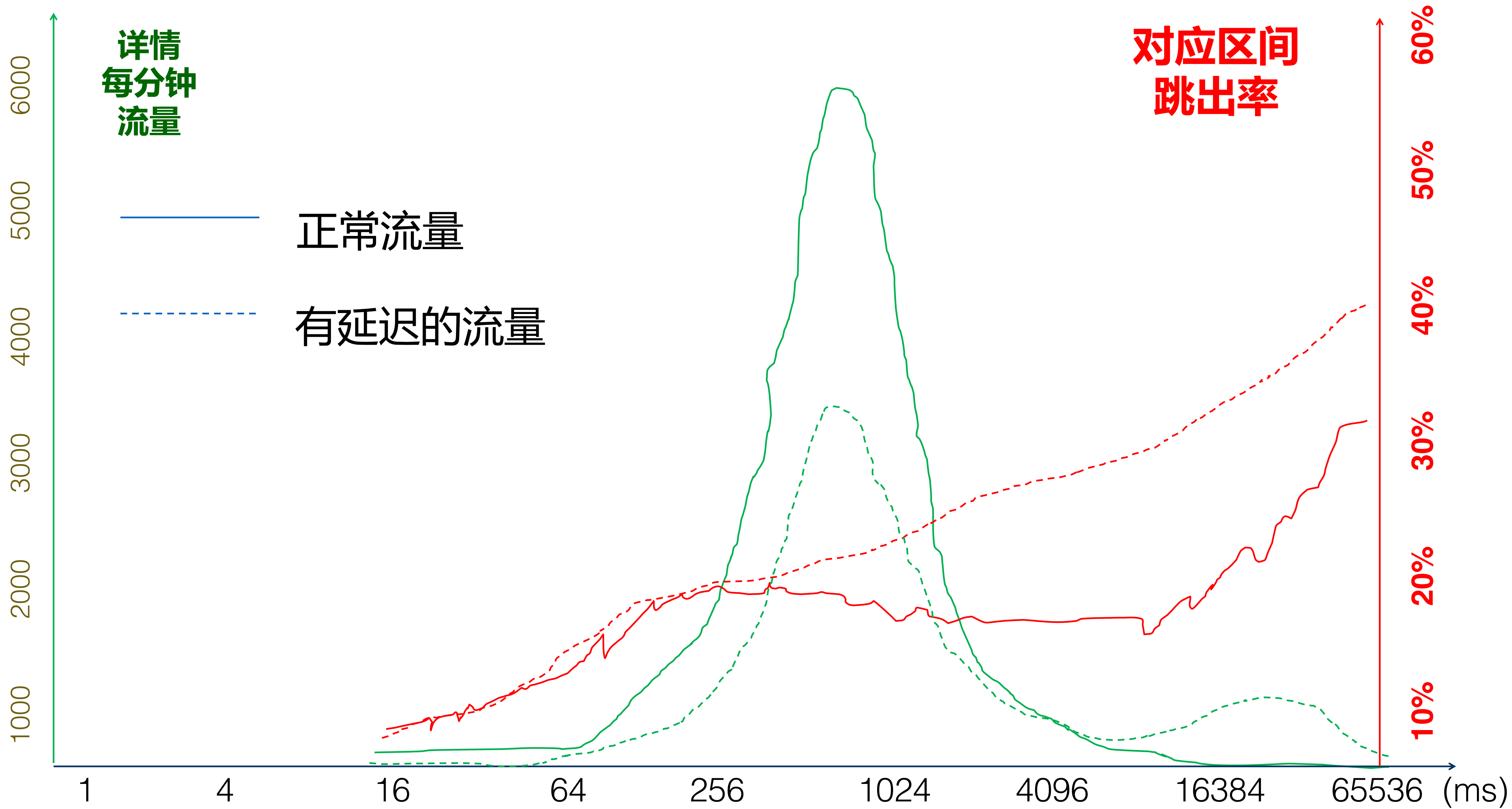
X

✓



# 理论基础 The Theory





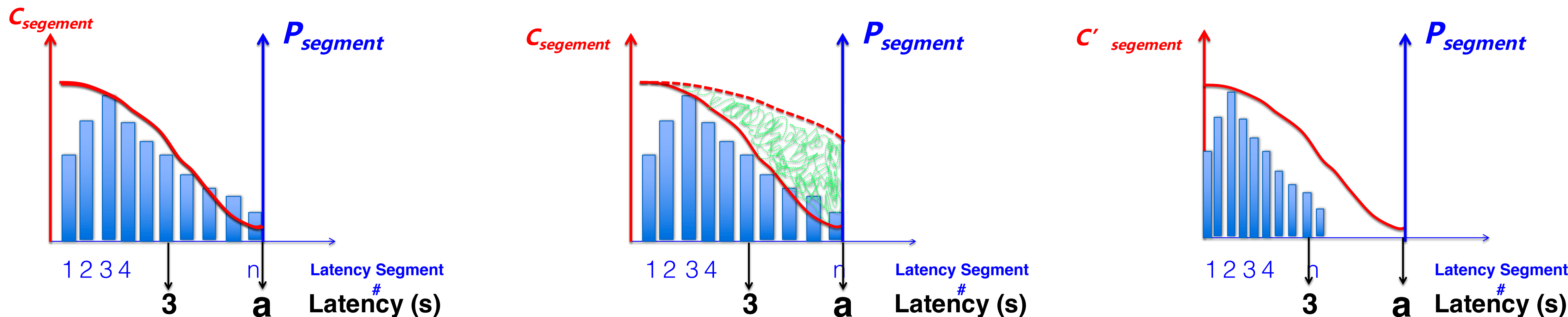
# Page Performance Loss : $L_{page}$

On any page, each latency segment  $P_{segment}(i)$  has its own conversion rate  $CR_{segment}(i)$ , the conversion rate of the entire page  $CR_{page}$  is the weighted sum:

$$C_{page} = \sum_{i=1}^n P_{segment}(i) \times C_{segment}(i)$$

If we uniformly compress all segments to within 3 seconds, each new latency segment will have the same amount of traffic  $P_{segment}(i)$ , but belongs to a new latency segment  $i' = [3i/a]$ . Therefore the new conversion rate  $CR'_{page}$  is:

$$C'_{page} = \sum_{i=1}^n P_{segment}(i) \times C_{segment}\left(\frac{3i}{a}\right)$$



$$L_{page} = C'_{page} - C_{page} = \sum_{i=1}^n P_{segment}(i) \times [C_{segment}\left(\frac{3i}{a}\right) - C_{segment}(i)] = \sum_{i=1}^n P_{segment}(i) \times L_i$$

The entire performance optimization based on the assuming that  $C(i) < C(j)$  if  $i < j$ , which is intuitively correct because people don't want to waste their life on page loading

# Site Performance Loss(全站性能损耗) $L_{site}$

## Traffic Compensation

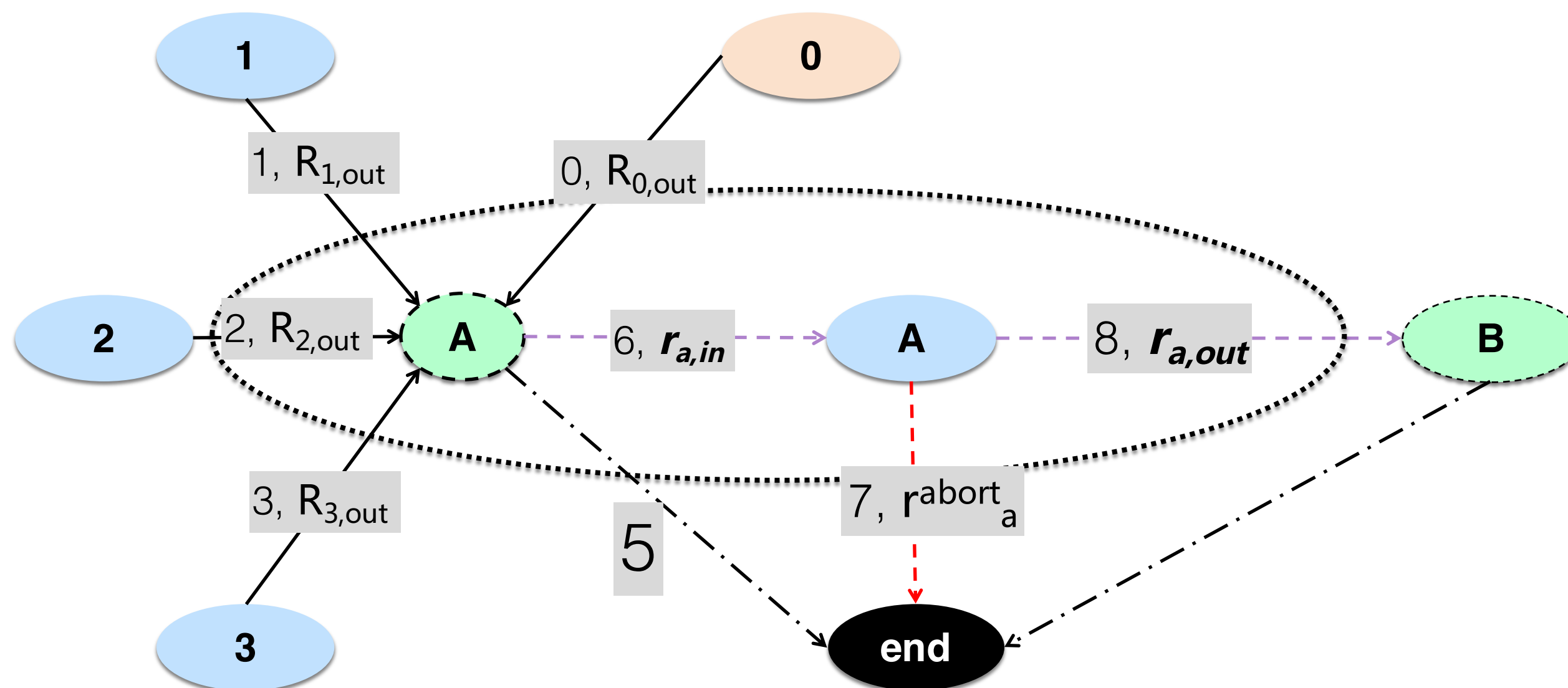
The total actual traffic for a page being  $r_{a,in}$  after compensating for performance loss, the new compensated incoming traffic for page  $a$  is  $r_{a,in}^c$ :

$$r_{a,in} = \sum_{i=0}^{inPageNum} r_{i,out} \times C_{i-a}$$

$$r_{a,in}^c = \sum_{i=0}^{inPageNum} r_{i,out}^c \times C'_{i-a} = \sum_{i=0}^{inPageNum} r_{i,out}^c \times (C_{i-a} + L_{i-a})$$

$$r_{a,out}^c = r_{a,in}^c - r_a^{abort}$$

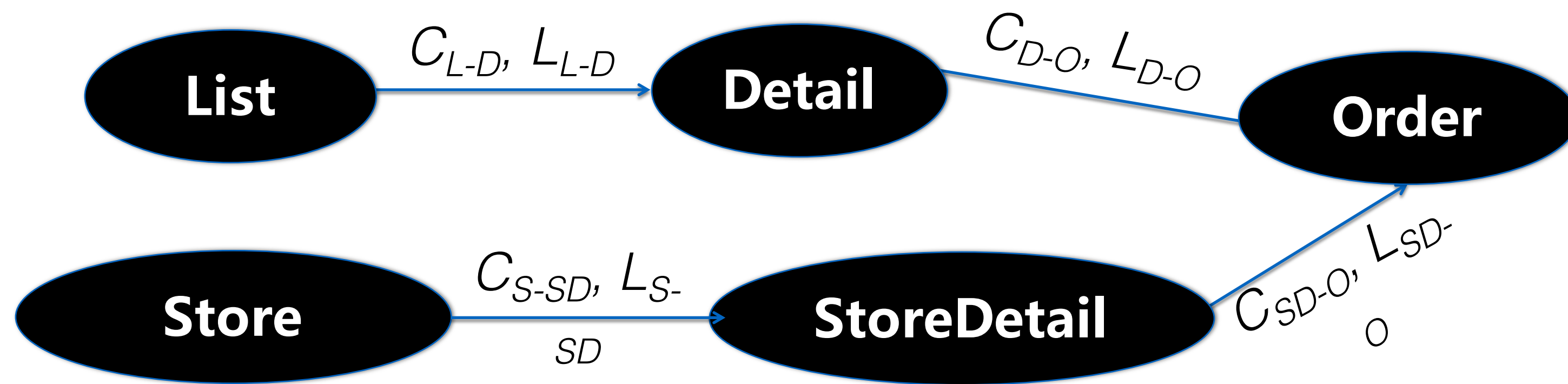
$$where : r_a^{abort} = r_{a,curPageLoad} - \sum_{requestNextPage=a} r_{i,in}$$



**Where**  $r_{a,in}^c$  being compensated incoming traffic for the page  $a$ ;  $r_a^{abort}$  being the bounce traffic for  $a$ ;  $r_{a,out}^c$  is the compensated outgoing traffic for page  $a$ .

## Site Performance Loss(全站性能损耗) $L_{site}$

Assuming a site's traffic flow is the following:



$$r_O^c = r_D^c C'_{D-O} - r_D^{abort} + r_{SD}^c C'_{SD-O} - r_D^{abort}$$

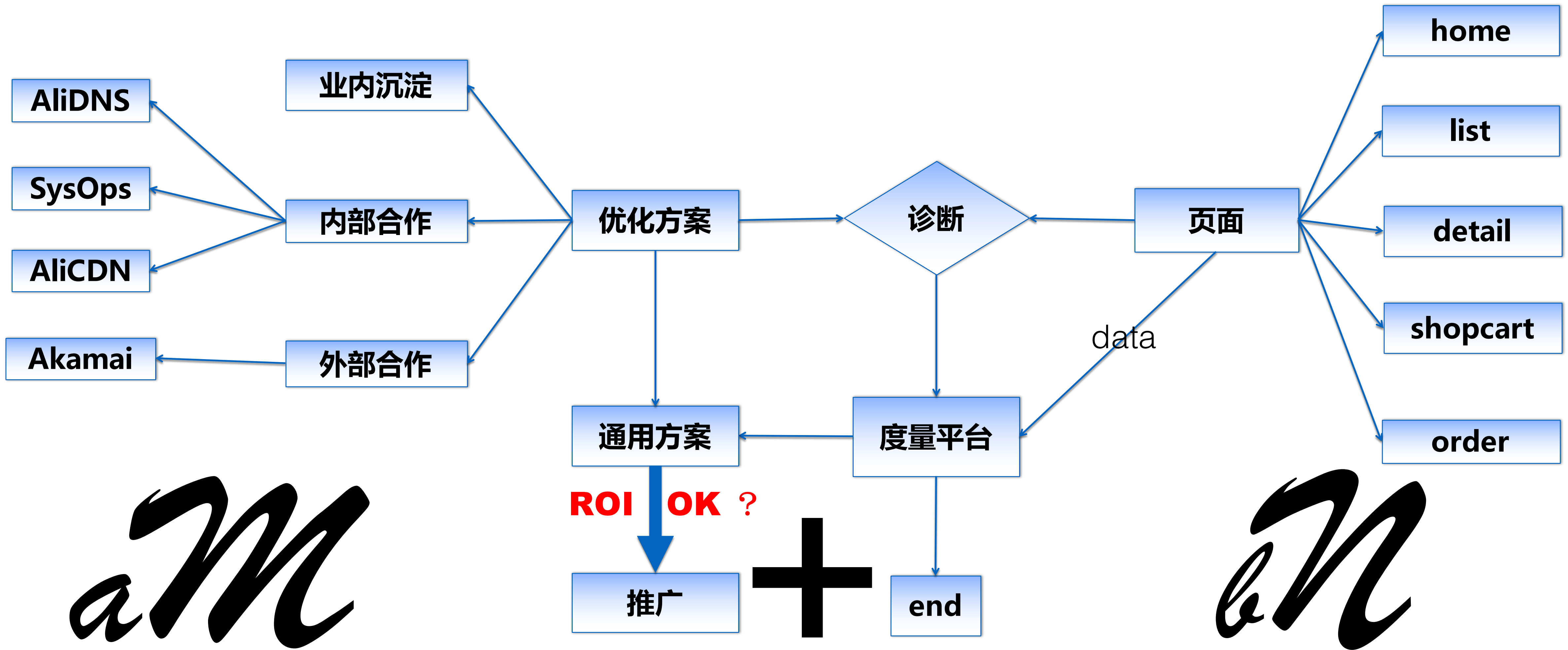
$$r_O^c = r_D^c (C_{D-O} + L_{D-O}) - r_D^{abort} + r_{SD}^c (C_{SD-O} + L_{SD-O}) - r_{SD}^{abort}$$

$$r_D^c = r_L \times C'_{L-D} - r_L^{abort} = r_L \times (C_{L-D} + L_{L-D}) - r_L^{abort}$$

$$r_O = r_L \times C_{L-D} \times C_{D-O} + r_S \times C_{S-SD} \times C_{SD-O}$$

$$L_{site} = \frac{r_O^c - r_O}{r_O}$$

# The Solution



# Performance Battle – Modern Warfare

**All  
forces**



Picture Credit:  
[https://incgames.files.wordpress.com/2011/11/xl\\_modern\\_warfare\\_3\\_screenshot\\_nyc\\_624.jpg](https://incgames.files.wordpress.com/2011/11/xl_modern_warfare_3_screenshot_nyc_624.jpg)  
[http://www.zacharyquinto.com/assets\\_c/2009/04/war%20room-thumb-376x282-452.jpg](http://www.zacharyquinto.com/assets_c/2009/04/war%20room-thumb-376x282-452.jpg)

**coordinated by a**

**single  
platform**



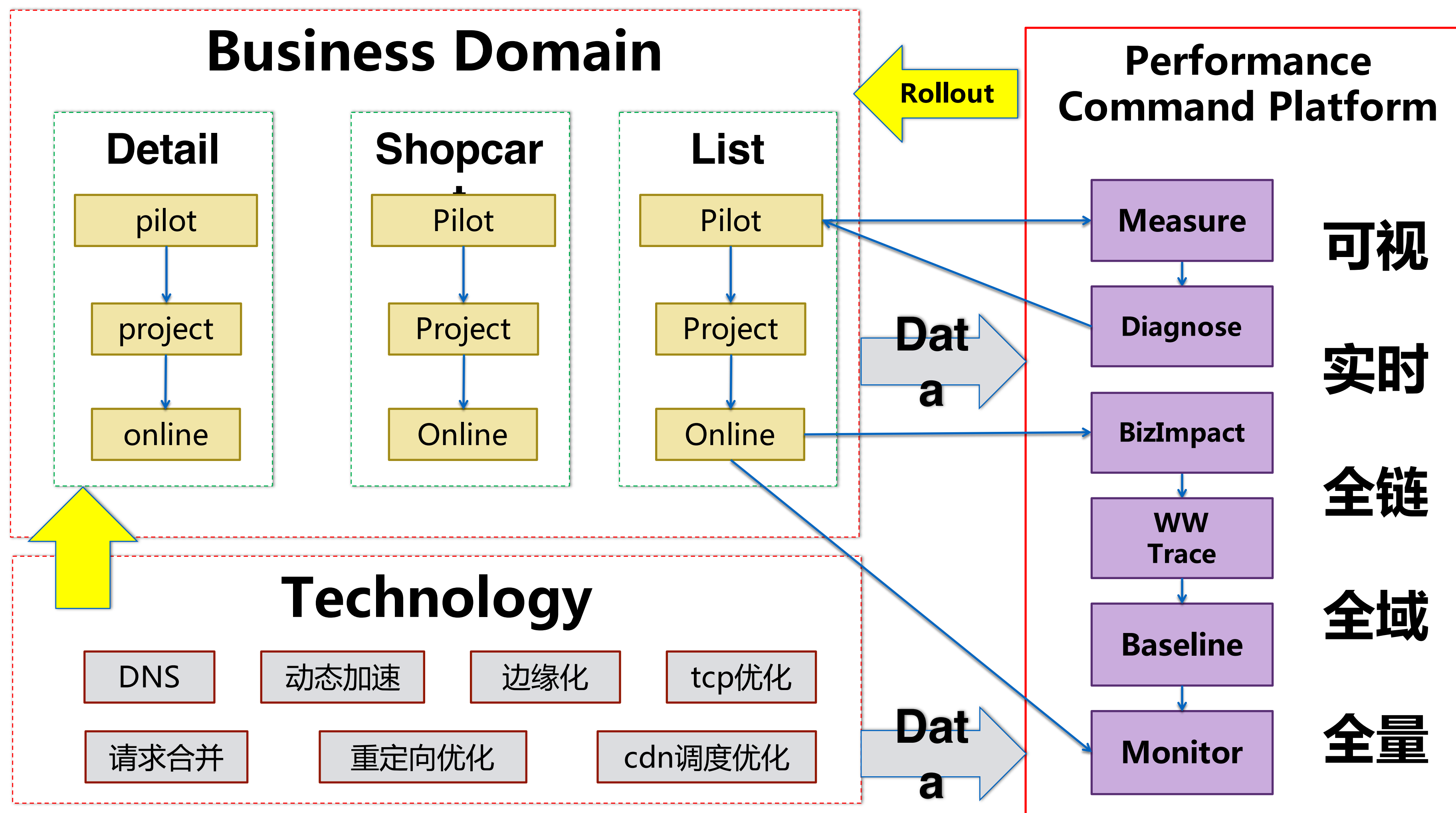
**to win the performance battle.**



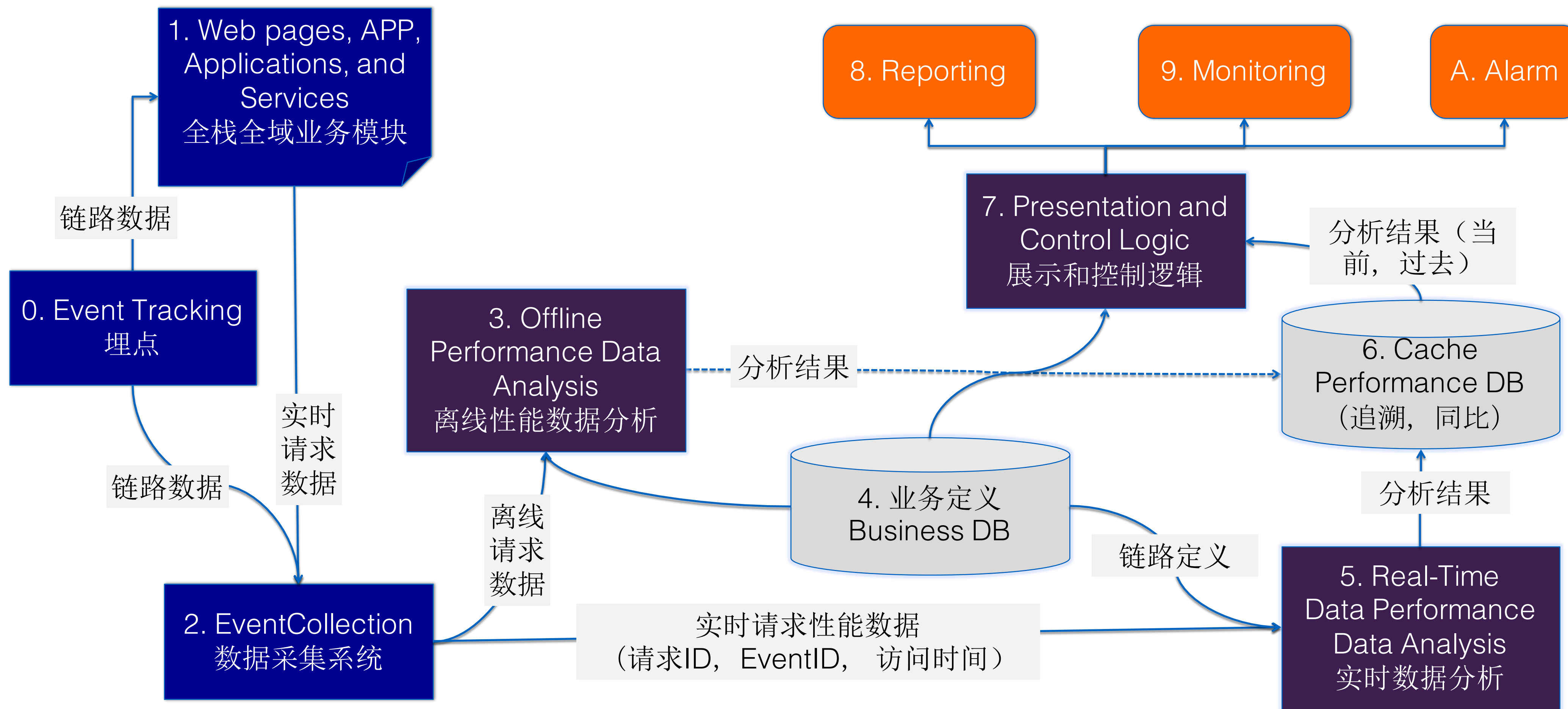
# 平台设计

# Platform Design

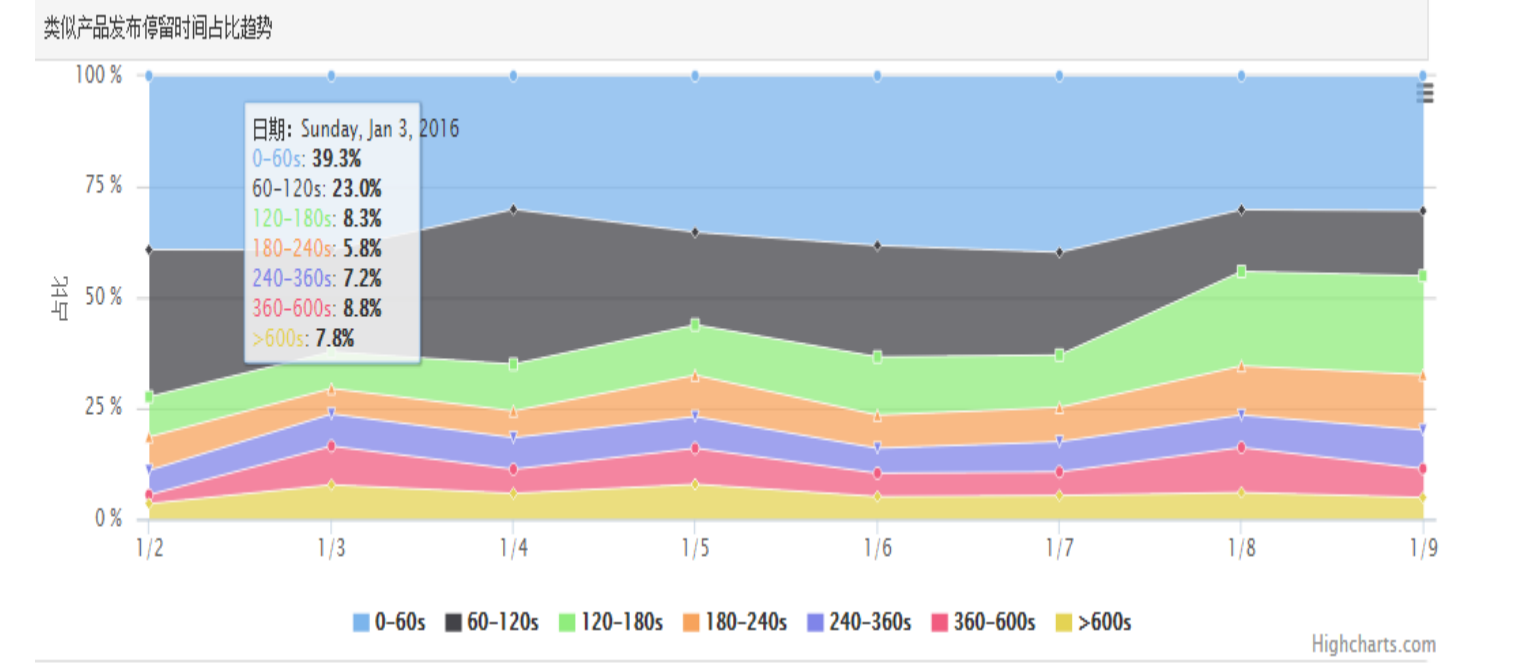
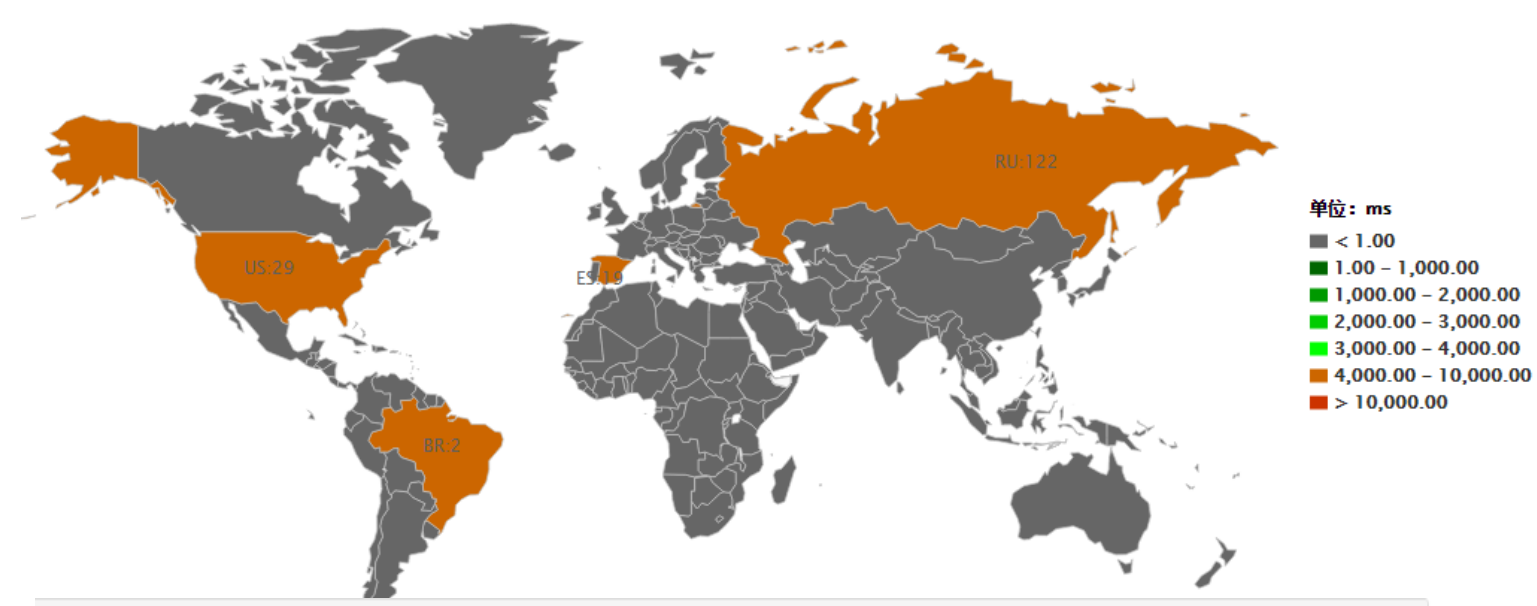
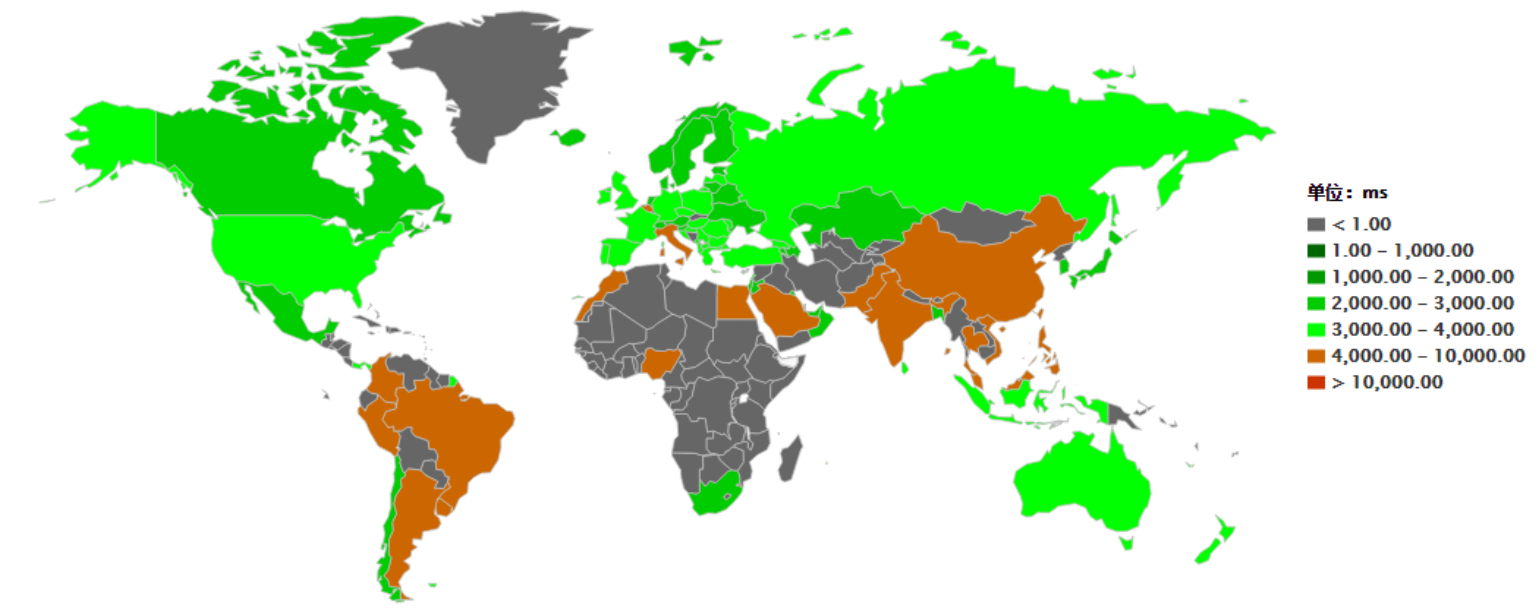
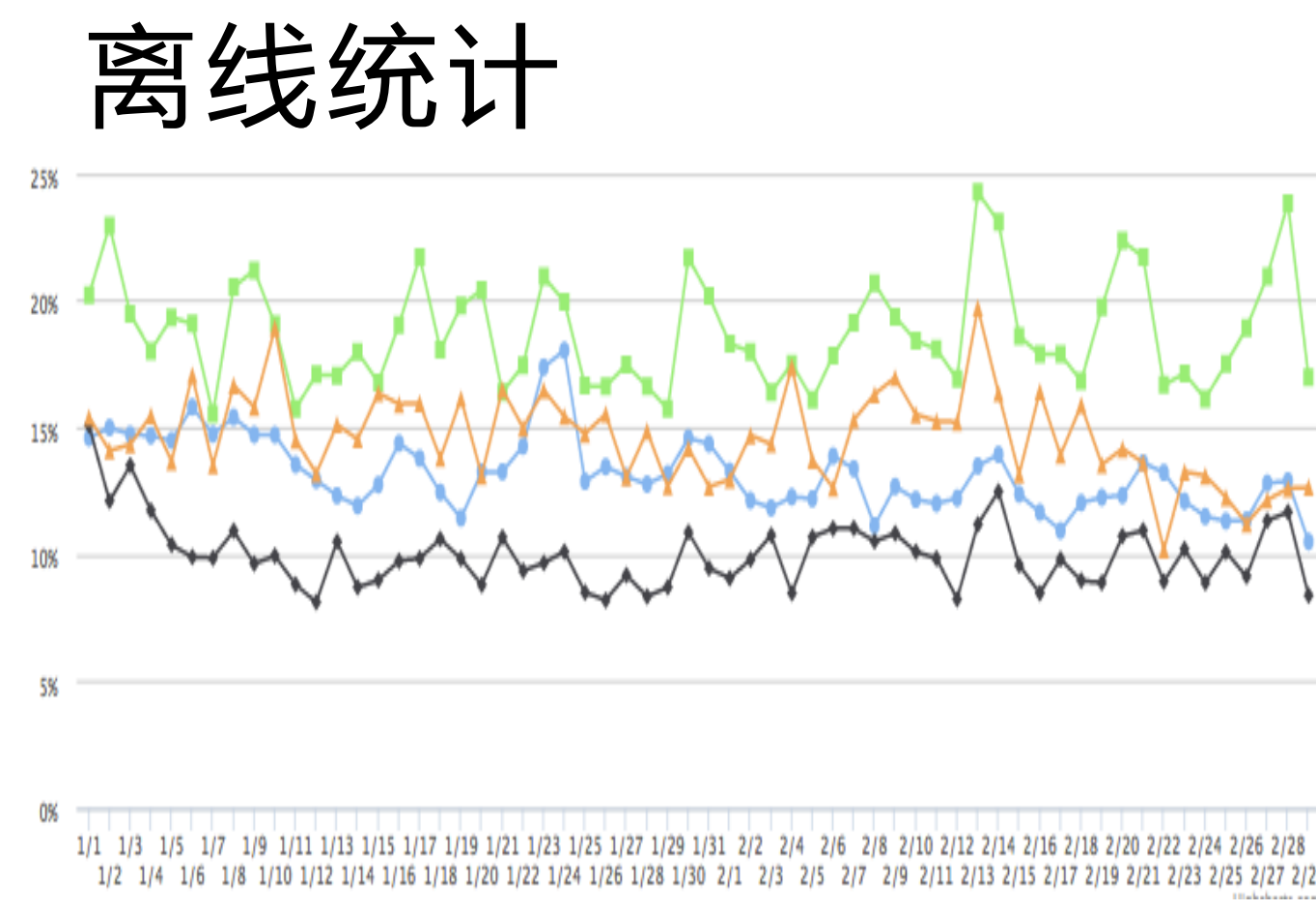
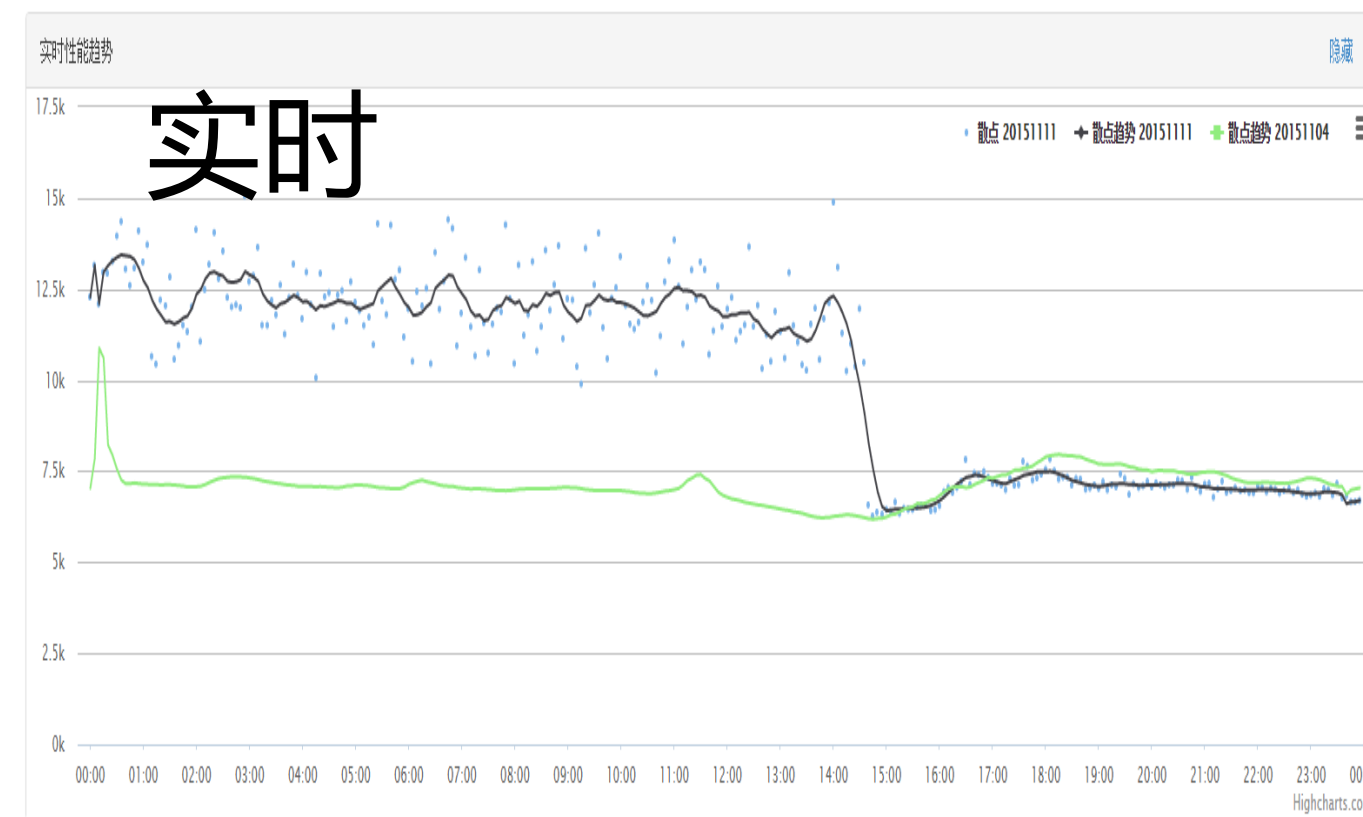
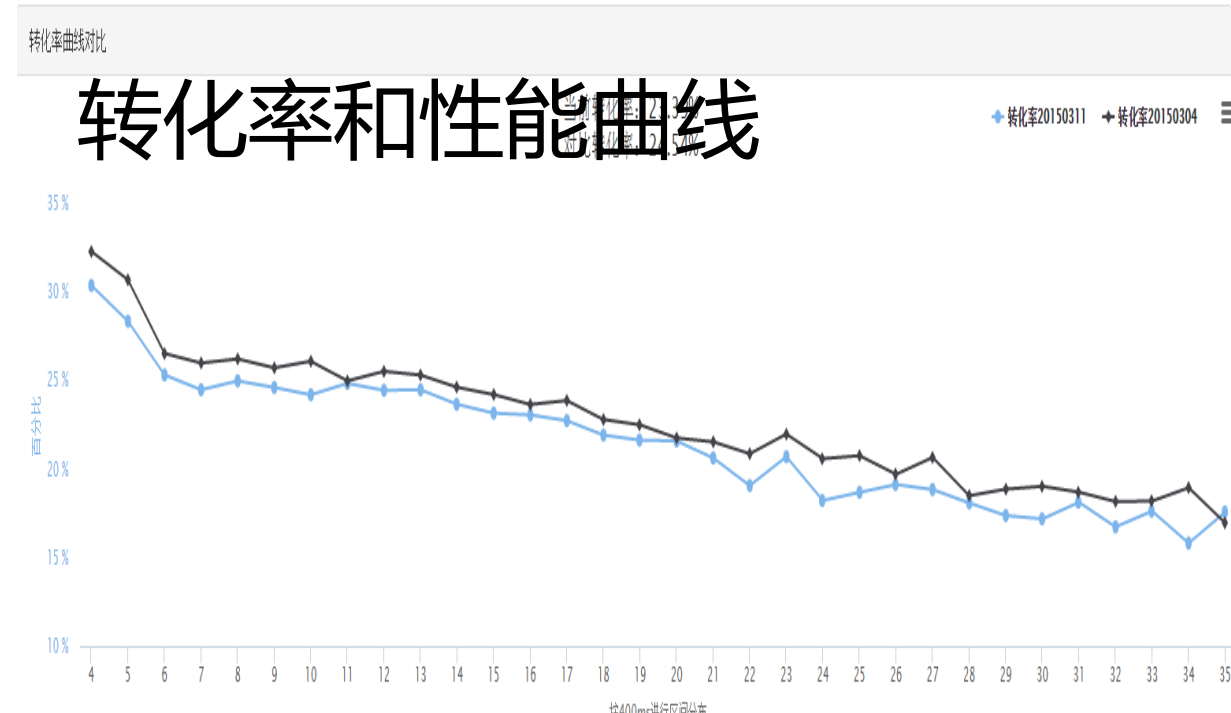
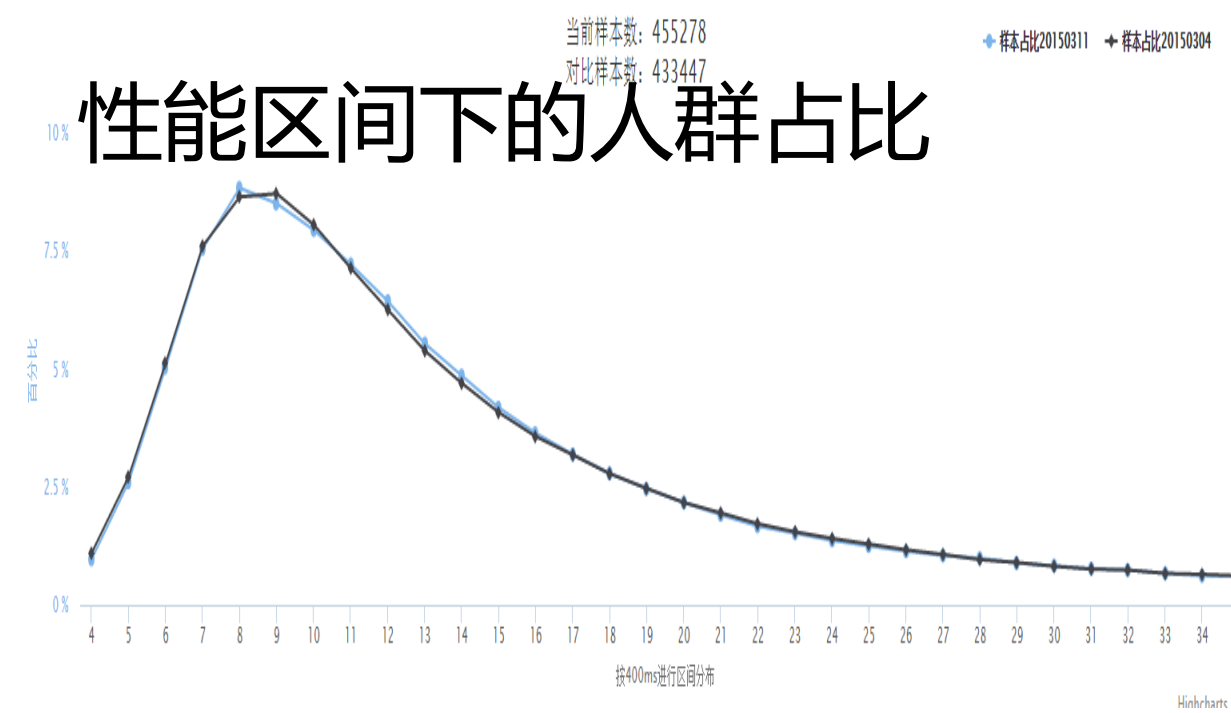
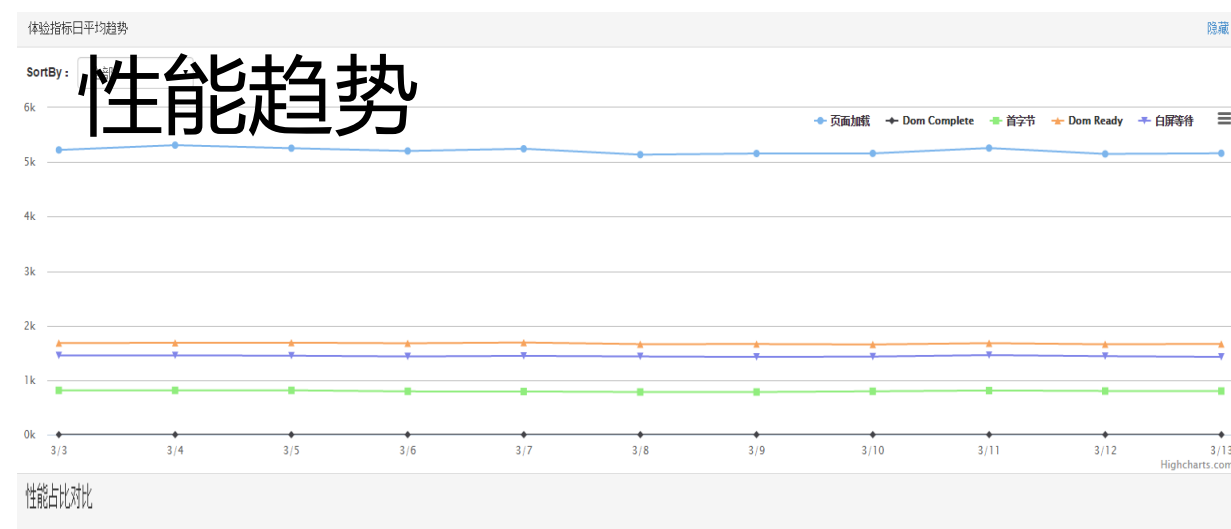
# Conceptual System Diagram



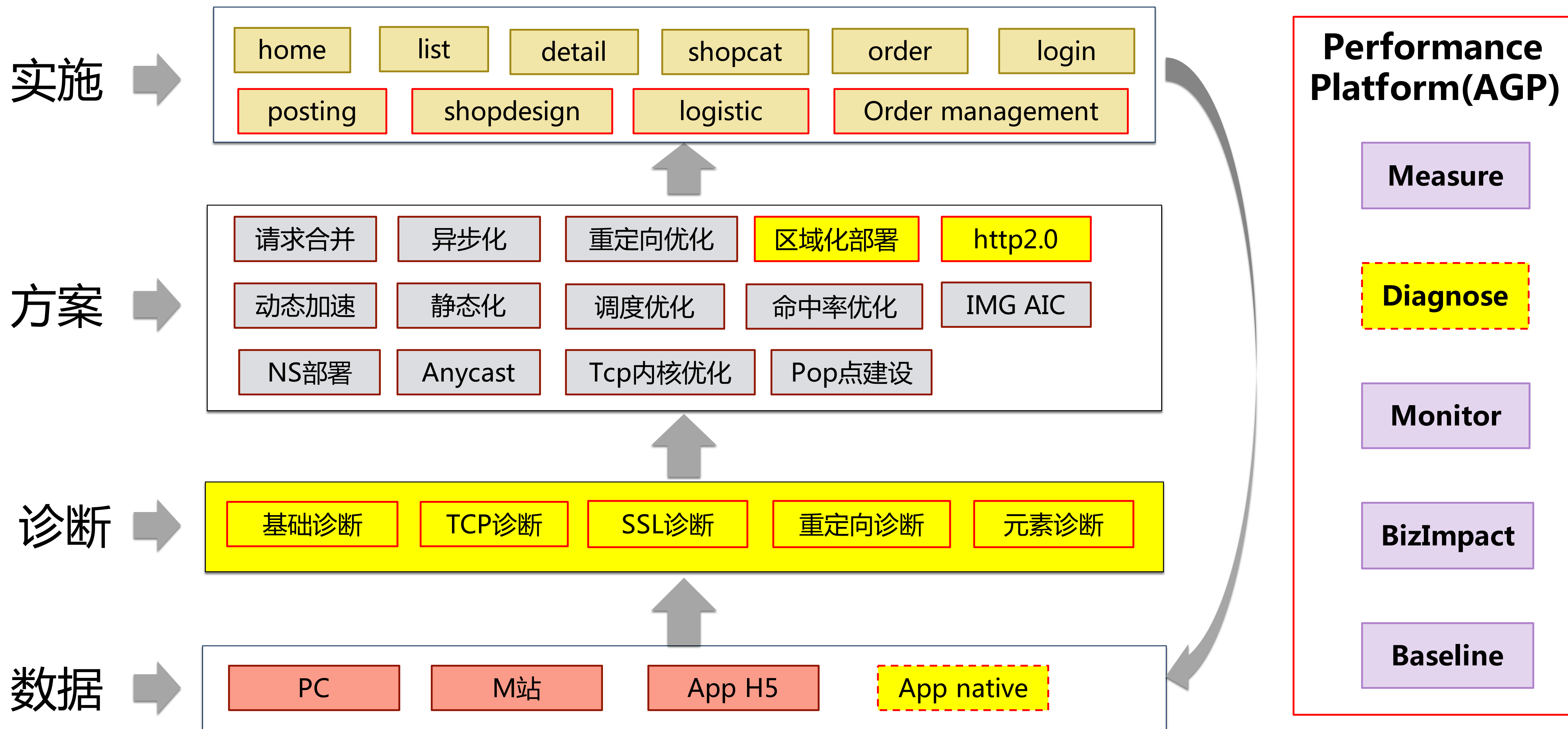
# Performance Platform (性能优化平台) Design



# System User Interface



# Platform Future



# 优化实施 Execution

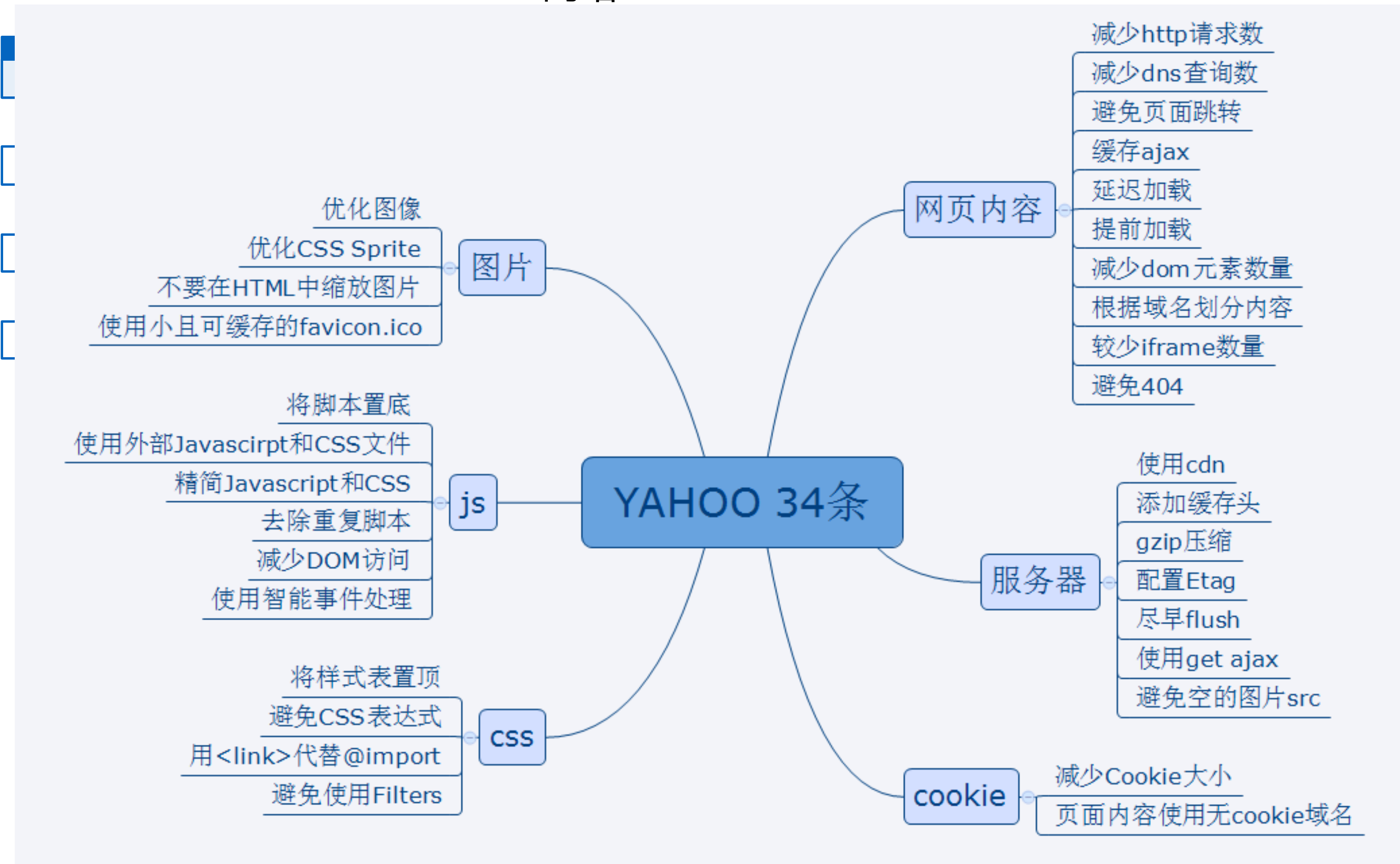
# 优化-方案

## DNS

## 网络

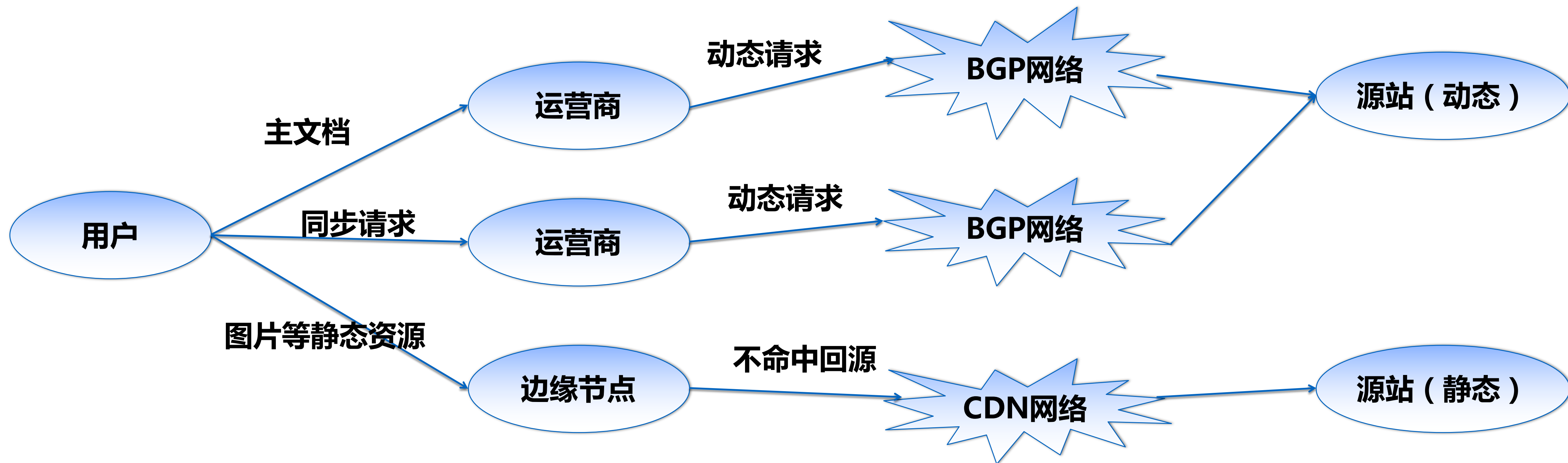
## CDN

## 业内沉淀



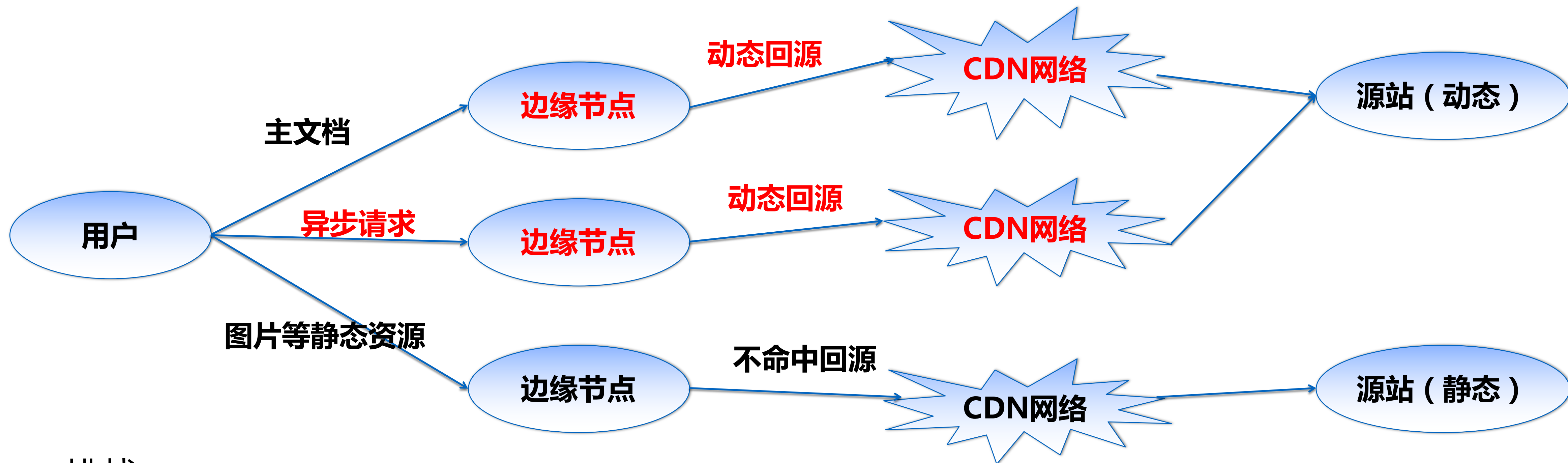
- YAHOO 34条
- HTTP2.0
- Prefetch
- prerender
- bigpipe
- 异步化
- 请求合并
- 其它

## 优化策略：动态加速（优化前）





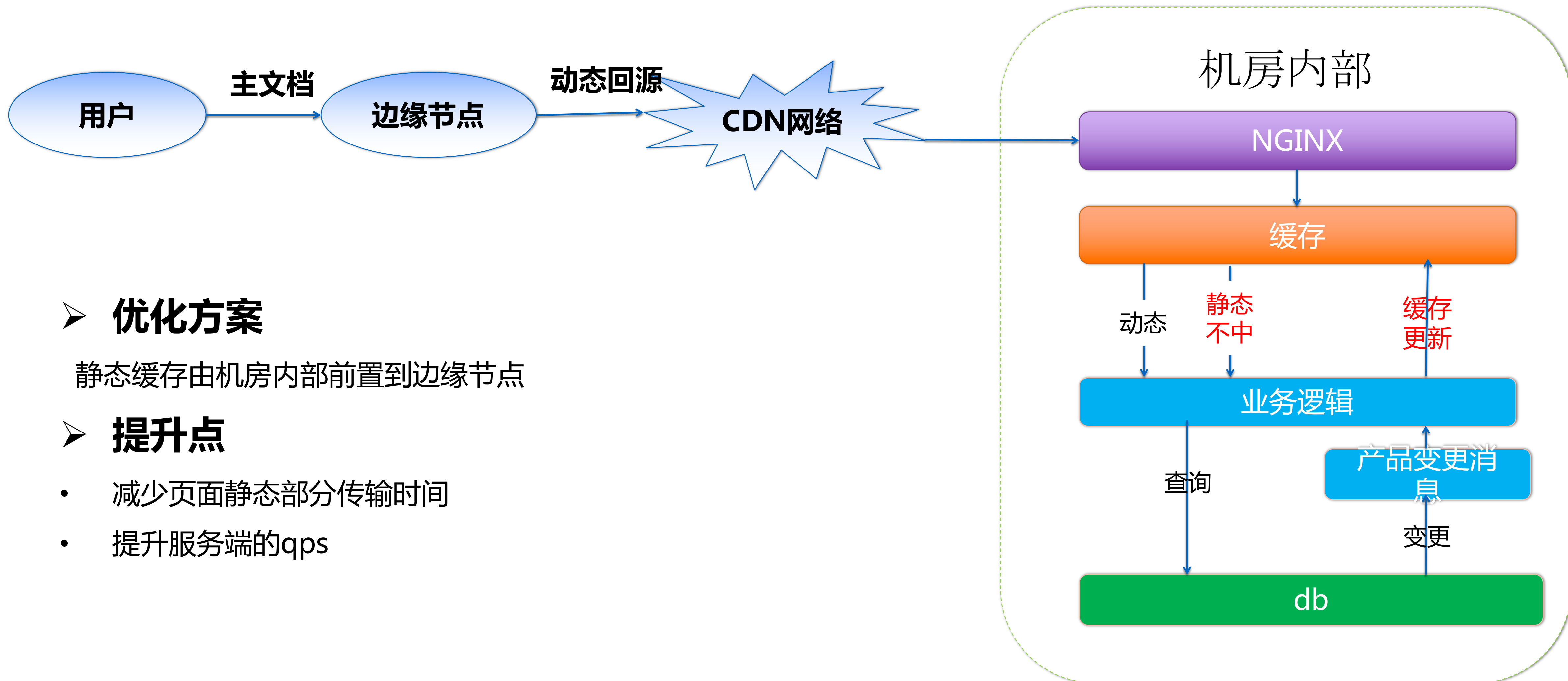
## 优化策略：动态加速（优化后）



### 挑战

- 获取真实用户ip
- 源站请求拦截
- 跨域调用

## 优化策略：静态化+ESI



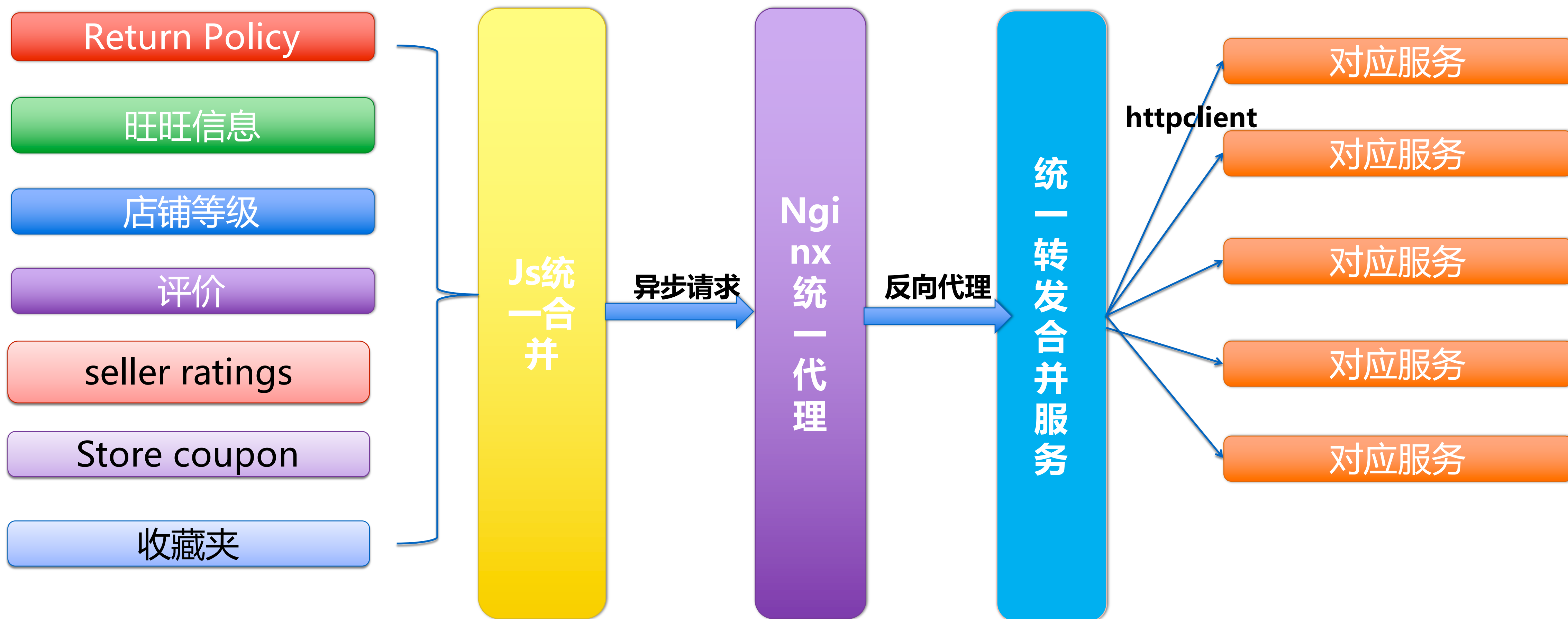
### ➤ 优化方案

静态缓存由机房内部前置到边缘节点

### ➤ 提升点

- 减少页面静态部分传输时间
- 提升服务端的qps

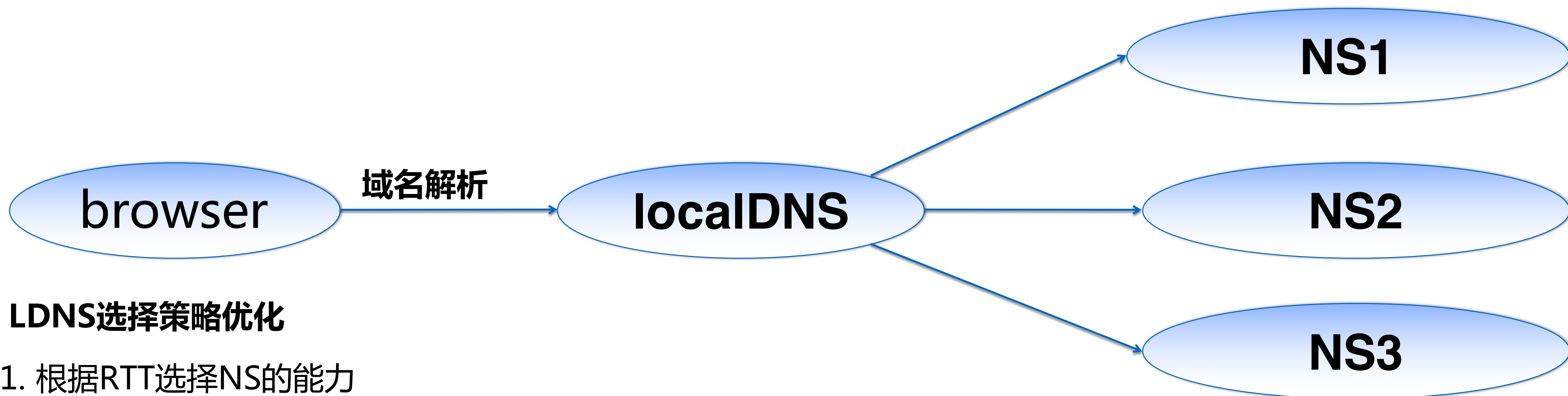
## 优化策略：元素请求合并



### 性能提升点

- a) 多个域名合并成主页面域名，dns=0，tcp=0
- b) 小内容合并，提高传输效率

## 优化策略：权威DNS优化



### ● LDNS选择策略优化

1. 根据RTT选择NS的能力
2. 存在一些选择RTT较大ns的情况

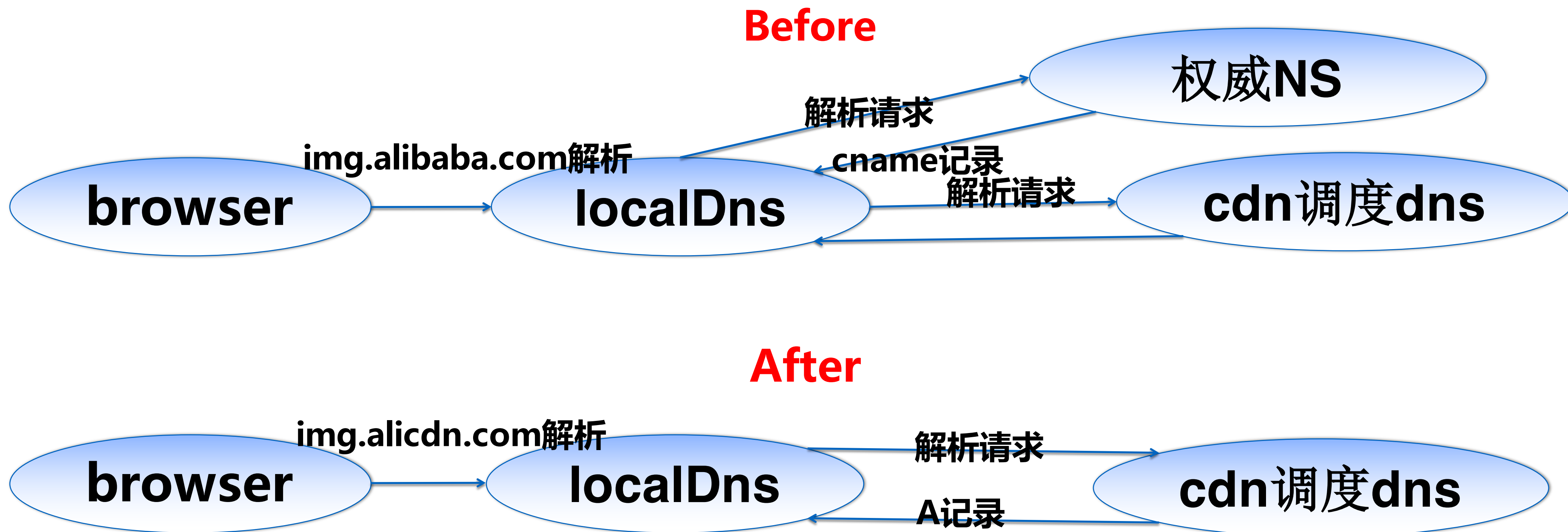
### ● 实施方案

选取流量较高且DNS性能差的国家部署权威NS

### ● 实施效果

1. 某些国家dns\_top99性能提升明显，某些国家反而下降
2. 整体页面加载时间基本没变化

## 优化策略：图片DNS优化



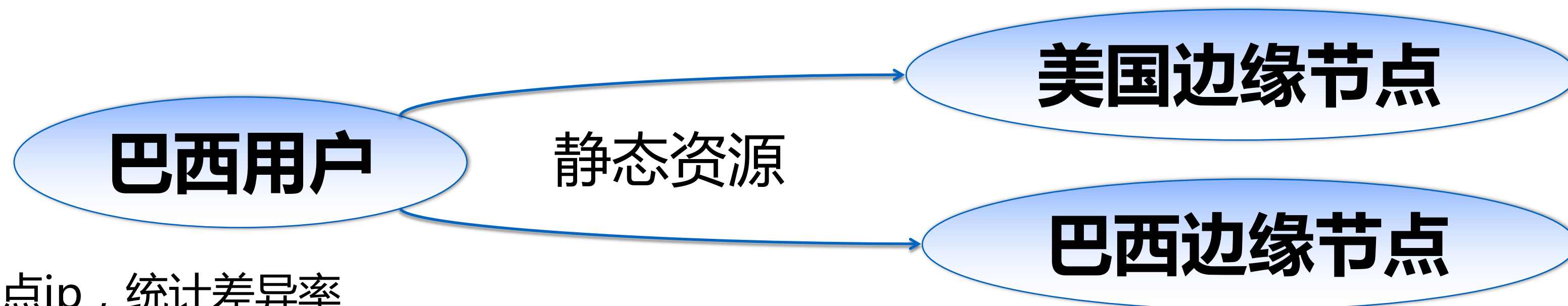
### ●实施效果

1. dns\_top99性能提升10%左右， dns\_top50基本没提升
2. 整体页面加载时间基本没变化

## 优化策略：CDN调度优化

### 统计方法

1. 采集客户端ip和边缘节点ip，统计差异率
2. 采集每个请求的耗时，统计不同差异下的耗时



客户端所在国家	边缘节点所在国家	相对耗时
BRAZIL	BRAZIL	4
BRAZIL	UNITED STATES	6
BRAZIL	ARGENTINA	7
BRAZIL	MEXICO	7

# 业务结果

# Business Impact

# Business Impact

## Time to diagnose a perform problem :

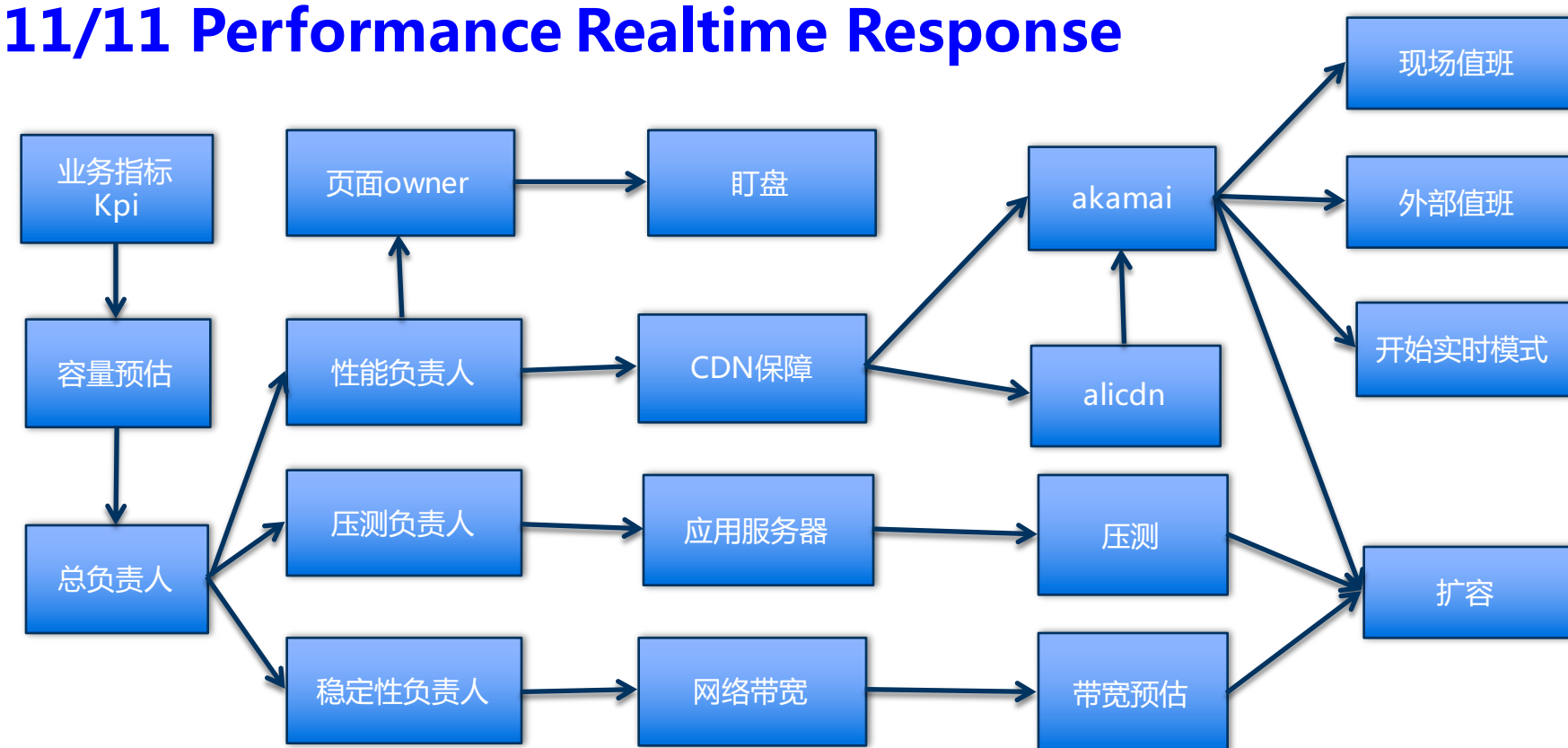
Before: 3 days

Now: < 1 min

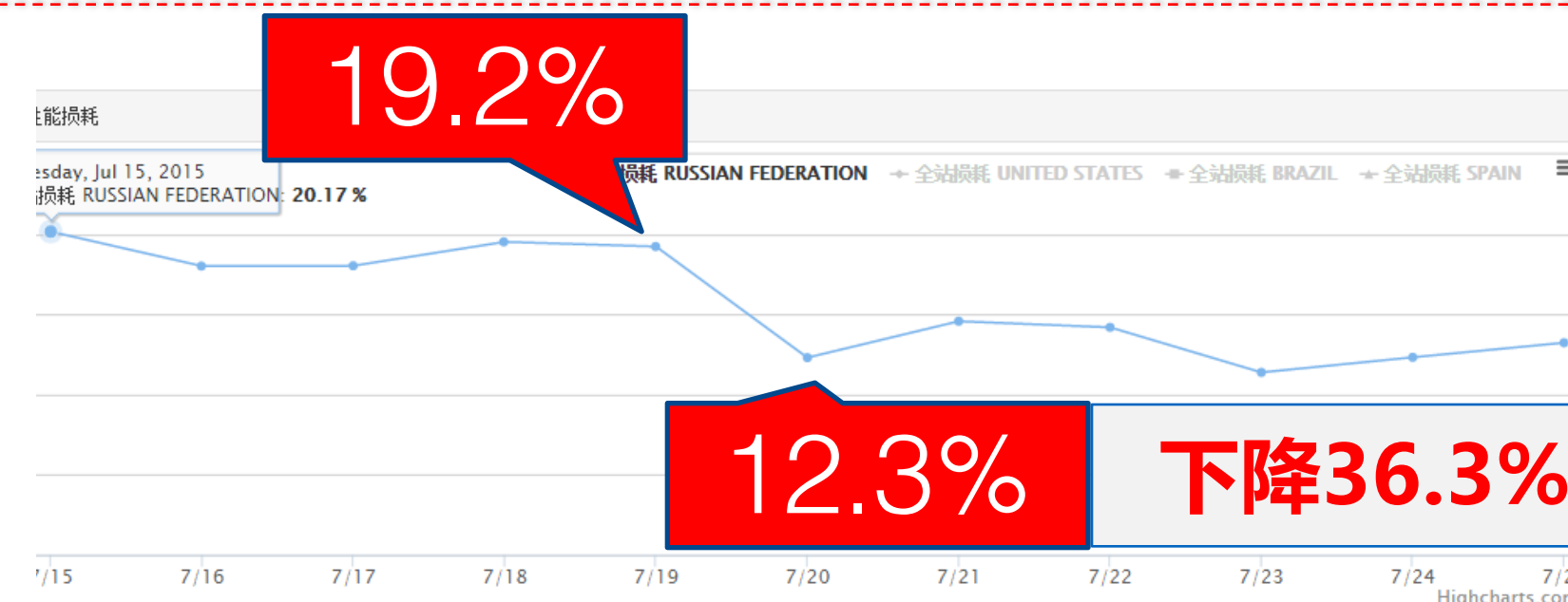
Replaced Gomez, saving \$600,000/year ,  
with TB/day of data analyzed.

Added 5.07% orders in absolute value.  
Verified during an outage: 8% drop of GMV.

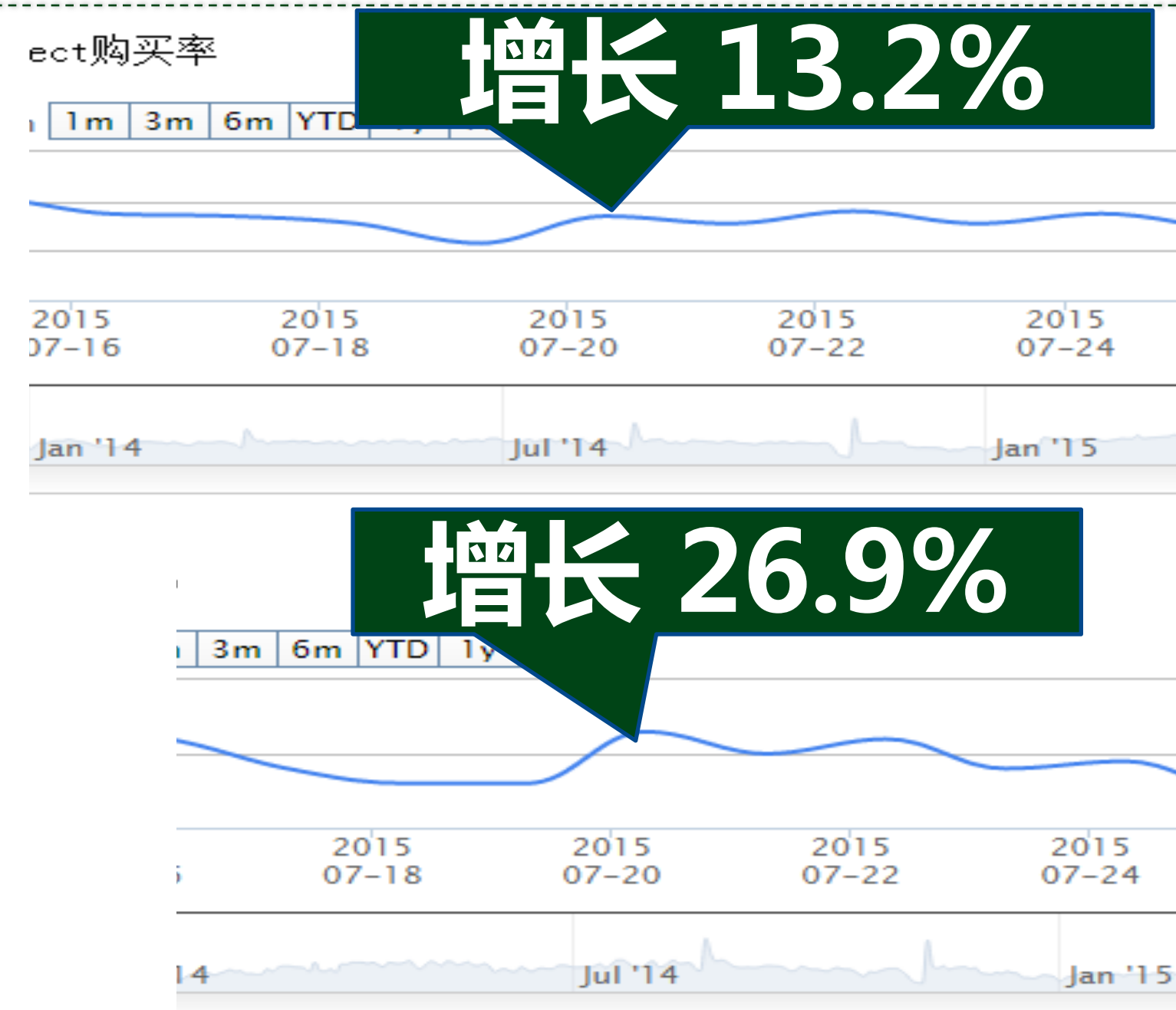
## 11/11 Performance Realtime Response



## 性能损耗

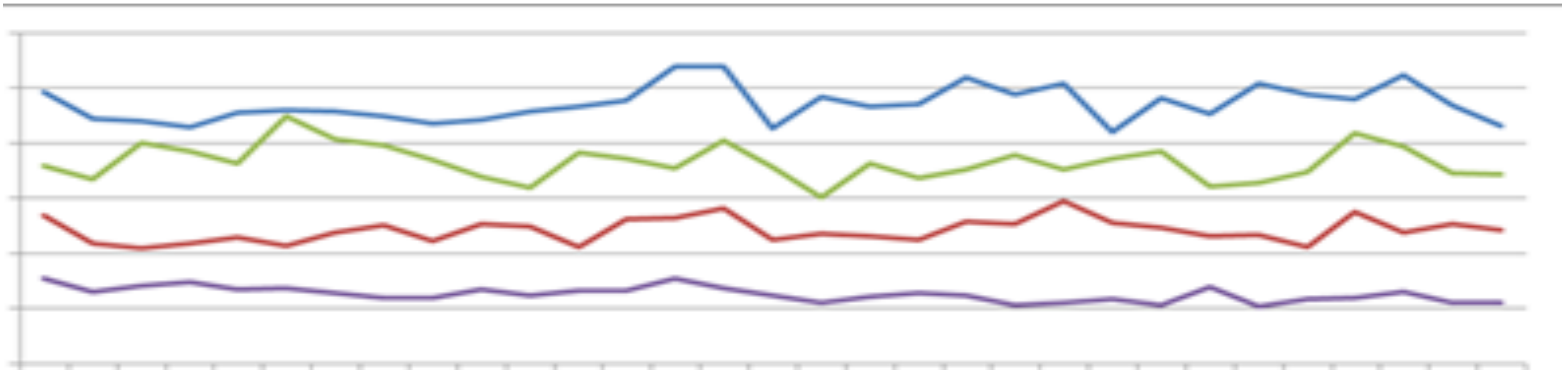


## 业务回报

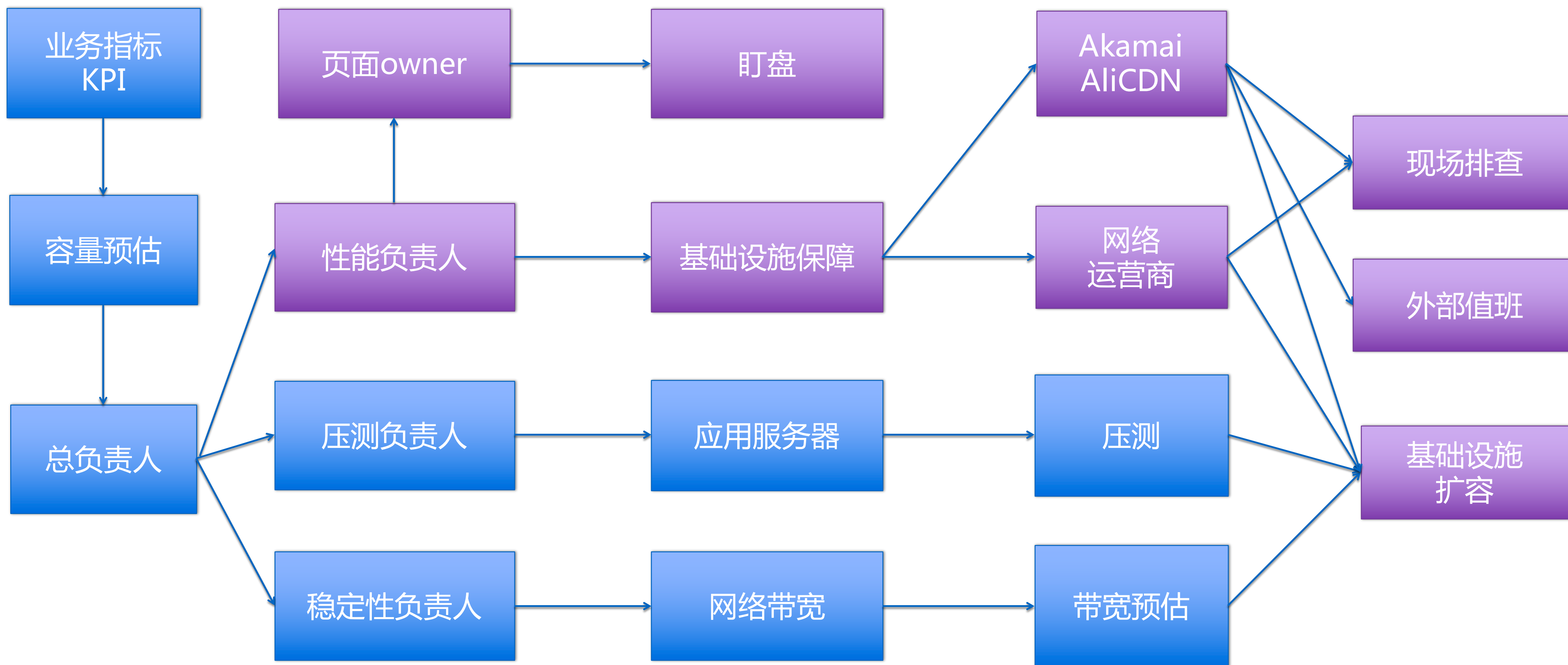




## KPI：性能损耗率

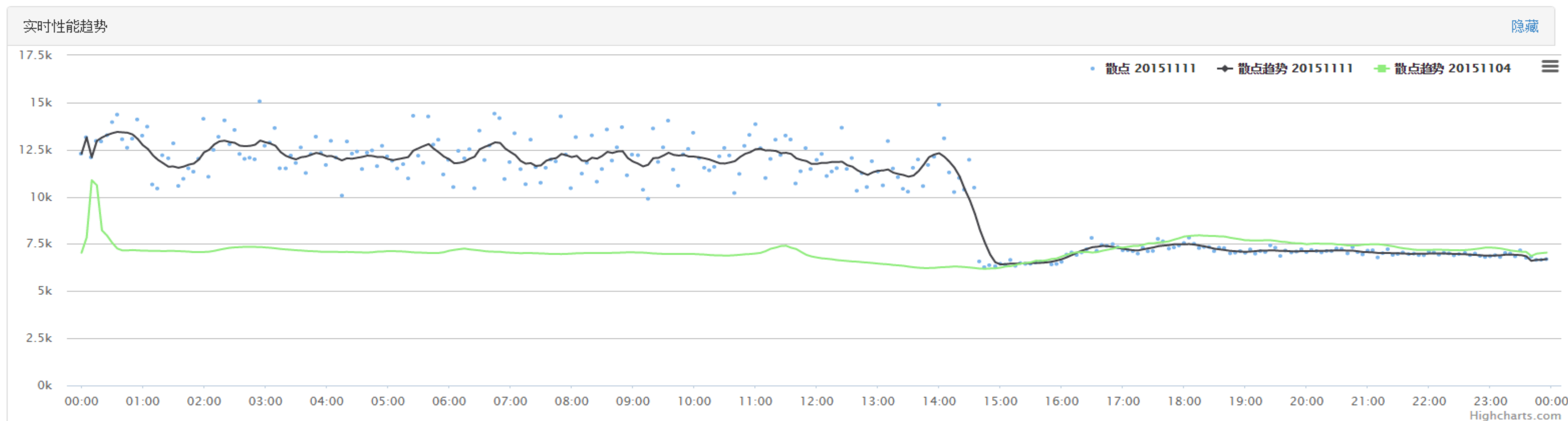


# 双11性能保障流程

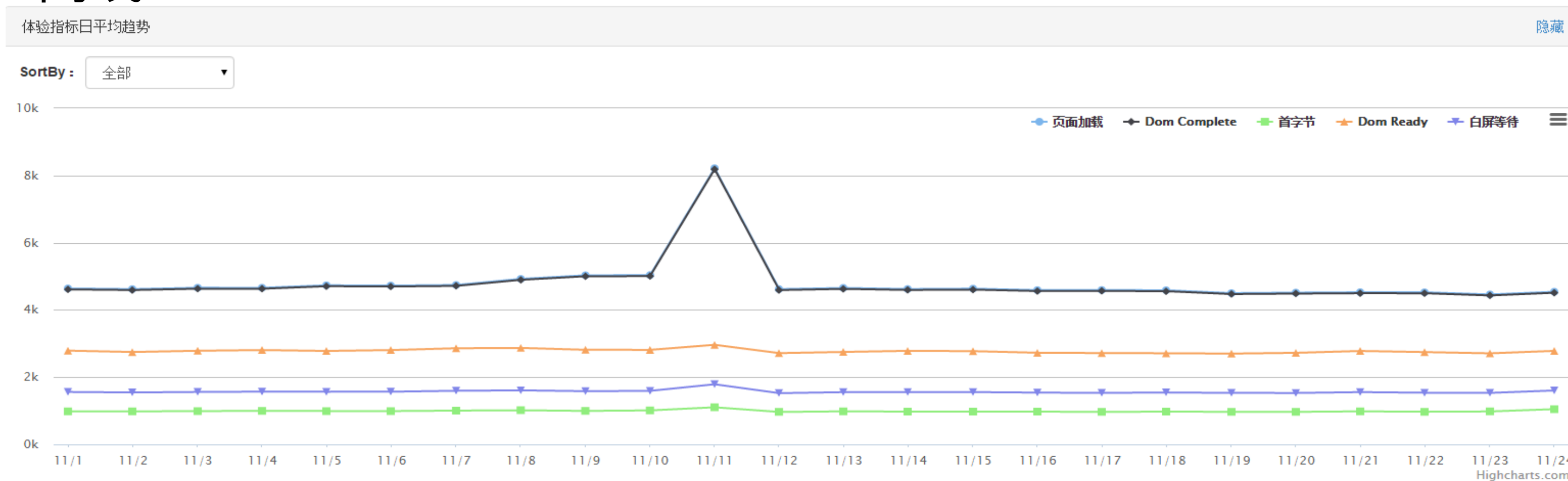


# 性能问题发现（双11页面性能下降）

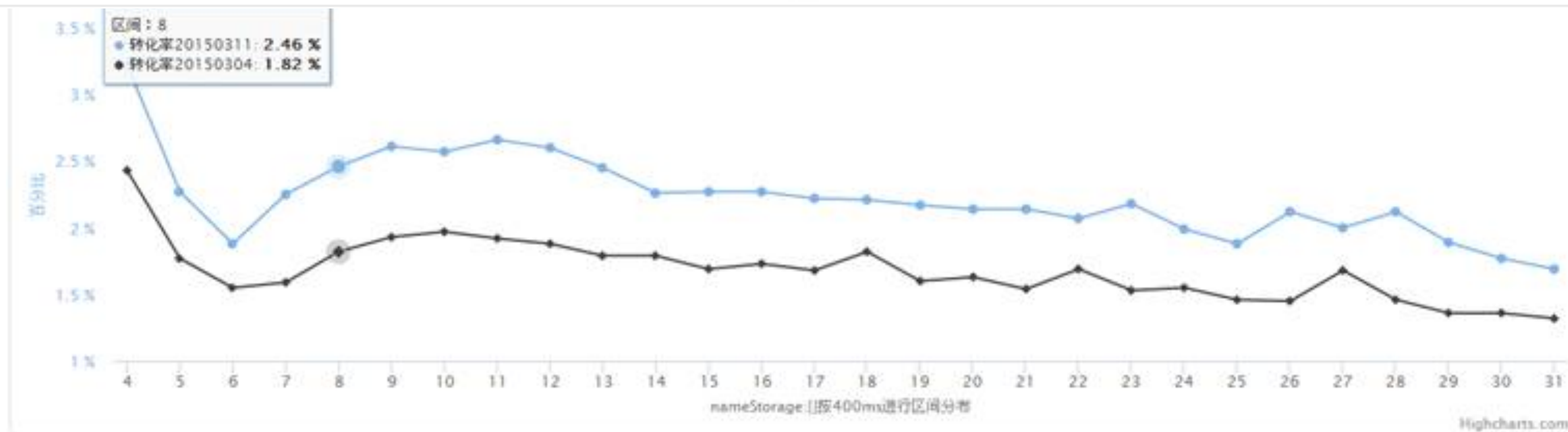
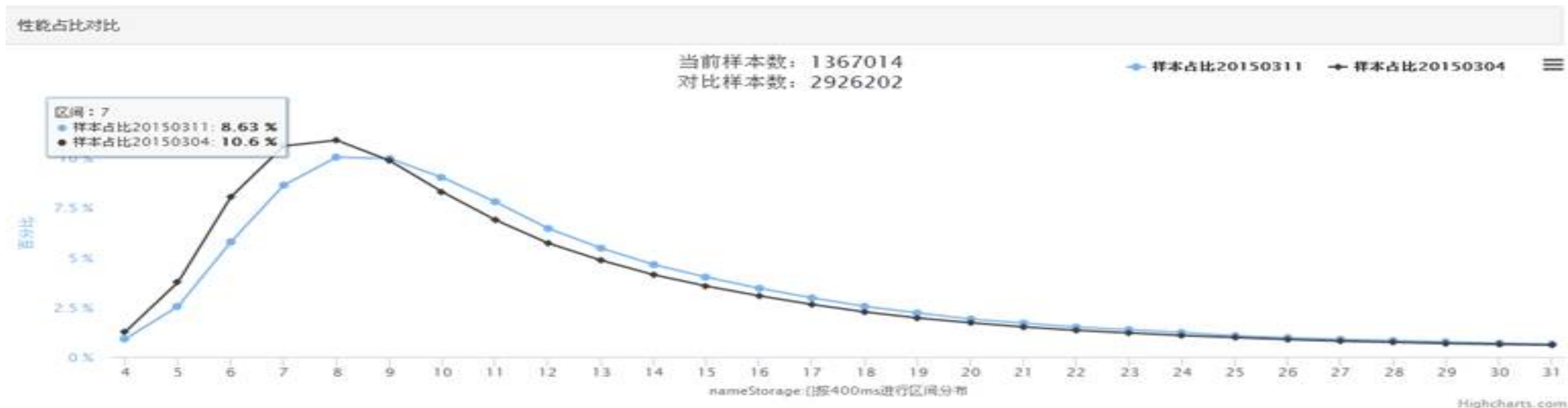
## 实时



## 离线

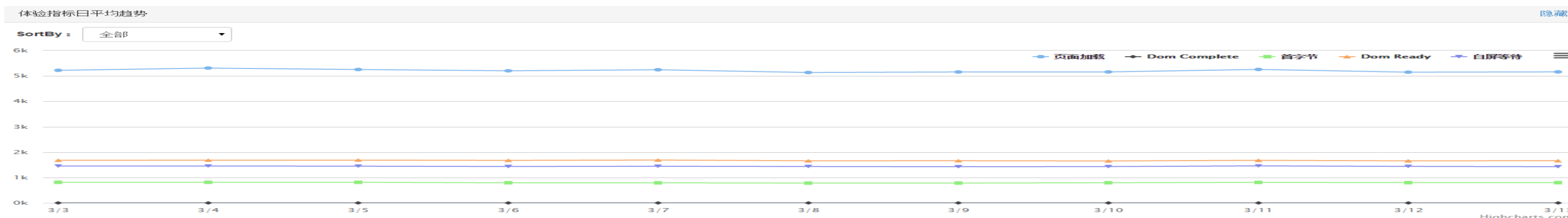


# 页面性能分析

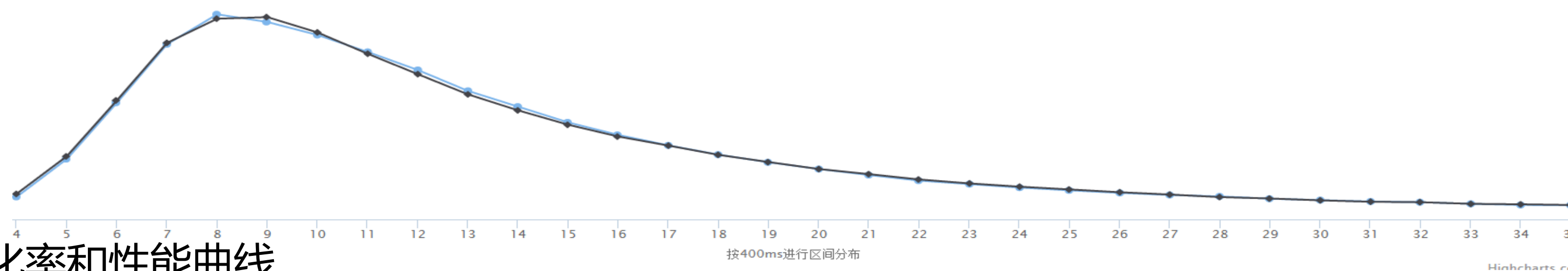


# 业务下降，性能造成？

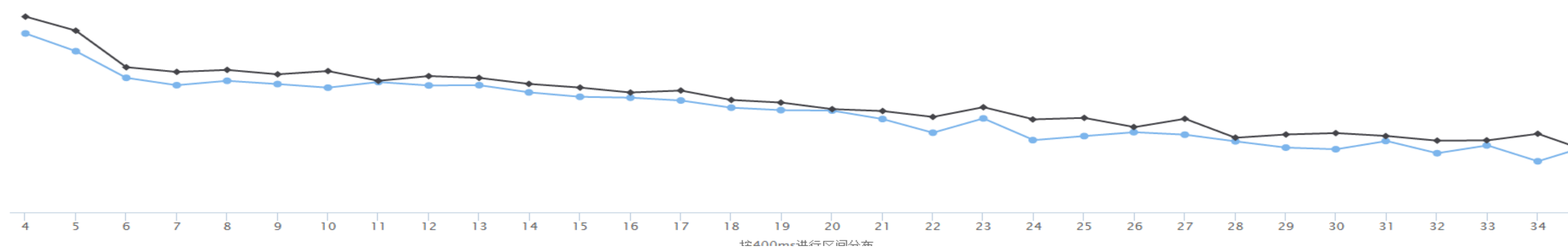
- 性能趋势



- 性能区间下的人群占比



- 转化率和性能曲线



# 架构思考

# Architecture Takeaways

## Architecture Takeaways

Does your design:

Encourage innovations?

Democratize your team?

Measure and magnify business impact?

Earn trust step by step?

AliExpress is Hiring!



<http://tinyurl.com/zm6x818>