



全球软件案例研究  
峰会

# 100

TOP 100 CASE STUDIES OF THE YEAR

全球软件案例研究峰会



TOP 100 CASE STUDIES  
OF THE YEAR

[www.top100summit.com](http://www.top100summit.com)



全球软件案例研究  
峰会

# 摘要

- 前言
- 什么时候做性能优化
- QPS相关公式推导
- Web开发性能优化案例
- 优化效果
- 总结



TOP 100 CASE STUDIES  
OF THE YEAR

[www.top100summit.com](http://www.top100summit.com)



全球软件案例研究  
峰会

# 分享者

80后，屌丝程序员  
来自阿里巴巴杭州  
爱好：自行车，摩托车



TOP 100 CASE STUDIES  
OF THE YEAR

[www.top100summit.com](http://www.top100summit.com)

# 什么时候做性能优化



## 魏文王 问 扁鹊



全球软件案例研究  
峰会

# 性能跟哪些指标相关呢？



TOP 100 CASE STUDIES  
OF THE YEAR

[www.top100summit.com](http://www.top100summit.com)

公式一：  $RT = CPU\ Time + Wait\ Time$

公式二： 单线程的  $QPS = 1000ms / RT$

举例推导

公式三： 最佳线程数 =  $((wait\ time + cpu\ time) / cpu\ time) * cpu\ cores * CPU\ 利用率 = RT / cpu\ time * cpu\ cores * CPU\ 利用率$



## QPS相关推导二

$$\text{最佳线程数} = \left[ \frac{\text{Wait Time} + \text{CPU Time}}{\text{CPU Time}} \right] \times \text{CPU 核数} \times \text{CPU 利用率}$$

单线程QPS=1000ms/RT

$$\text{QPS} = \frac{1000}{\text{CPU Time}} \times \text{CPU 核数} \times \text{CPU 利用率}$$

那么如何降低  
CPU Time呢





全球软件案例研究  
峰会

# CPU利用率低的几种情况

## ➤ IO能力

- 网络IO
- 磁盘IO
- FD

## ➤ 内存不足

## ➤ 数据库连接池

## ➤ BIO的RPC要考虑RPC的连接池

## ➤ ○ ○ ○ ○



TOP 100 CASE STUDIES  
OF THE YEAR

[www.top100summit.com](http://www.top100summit.com)



# CPU Time的降低取决于数据结构和算法，还有架构

## ➤ 数据结构降低CPU Time的例子

- List->Hash
- Contains和Get
- 均摊

## ➤ 架构改进降低CPU Time的例子

- 请看案例





全球软件案例研究  
峰会

# 影响性能的架构案例



TOP 100 CASE STUDIES  
OF THE YEAR

[www.top100summit.com](http://www.top100summit.com)

# 一般网站技术架构

今天我们的案例在灰色部分

## 反向代理及Web缓存

软件应用

业务应用集合

OpenAPI平台

统计系统

平台服务

分布式消息队列

分布式缓存

等等

分布式任务处理

权限系统

RPC框架

基础设施

分布式计算框架

分布式数据存储

分布式文件存储

虚拟云架构

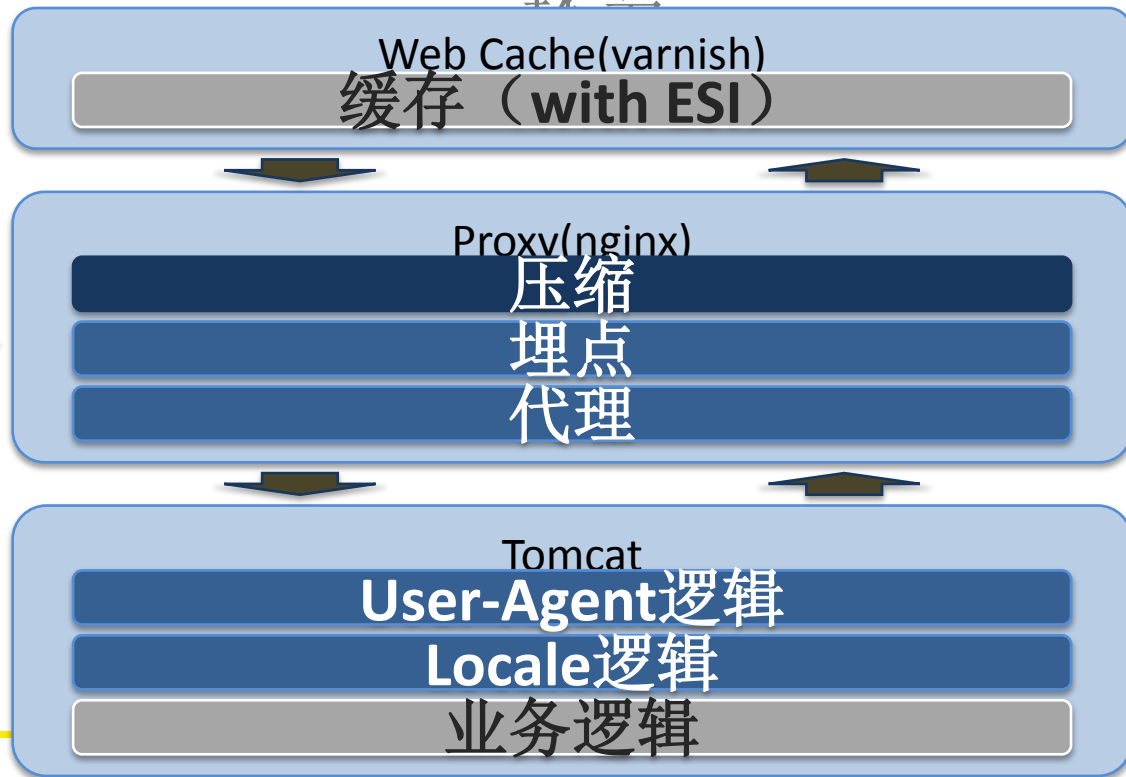
监控系统

日志数据的采集

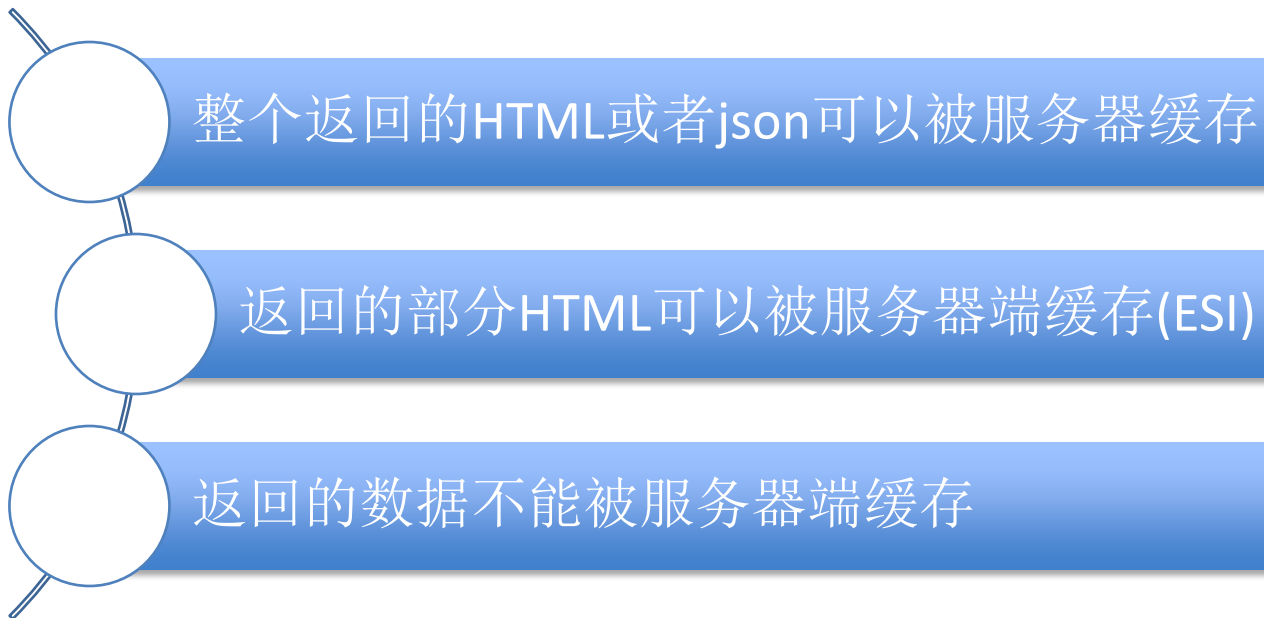


# 一个web页面打开涉及到的各环境

埋点的实现方法和sitemesh的实现类似



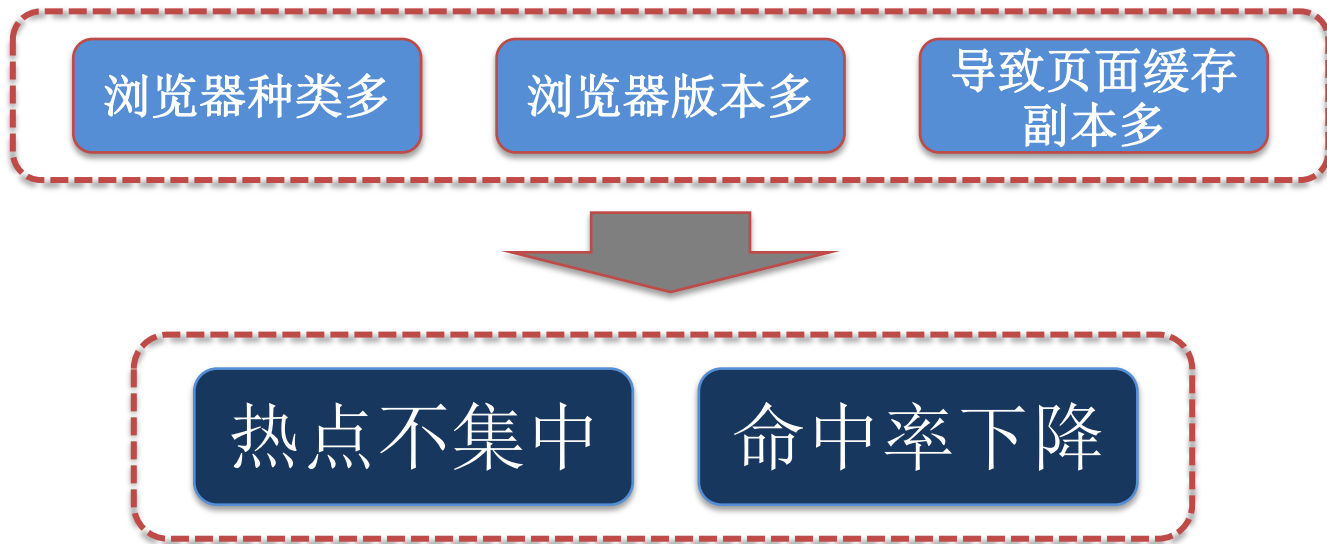
# 这套模型的适用场景



# 这个架构解决不了的问题

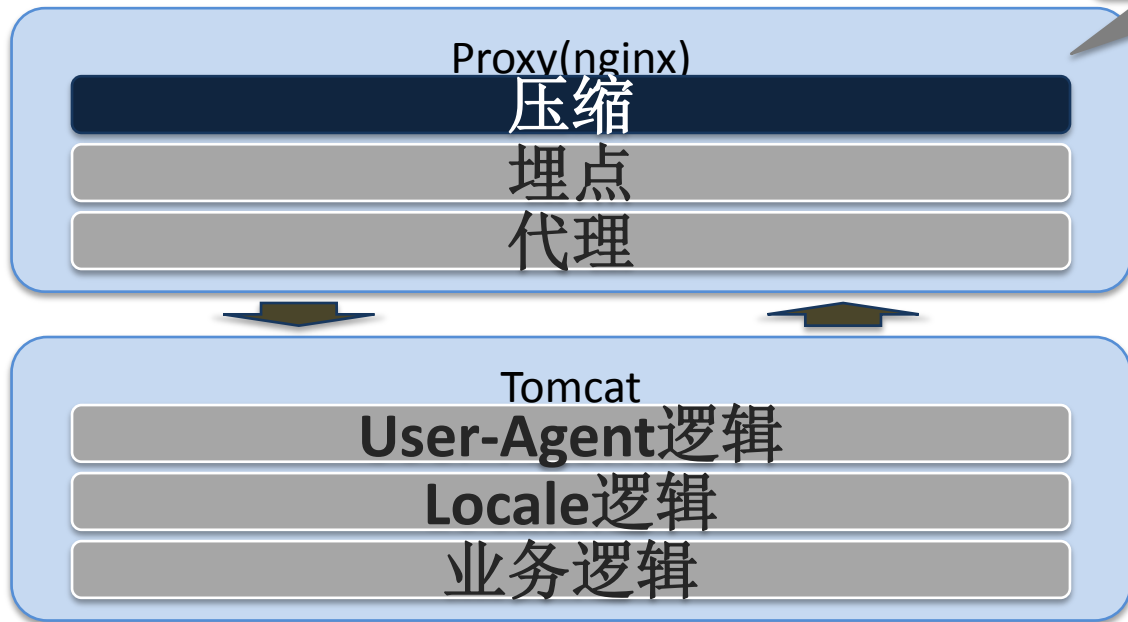


# Web Cache + HTTP Vary:user-agent带来的问题



# 方案一

为每次请求都做压缩操作





但是这个压缩在我们场景里就是CDU Time最大的占比者

压缩级别从6到3

6->3	QPS	RT	Band Width
92KB	↑51%	↓32%	↑13%
138KB	↑53%	↓37%	↑12%
182KB	↑60%	↓45%	↑13%
248KB	↑65%	↓39%	↑16%
295KB	↑61%	↓38%	↑16%

压缩级别从6到1

6->1	QPS	RT	Band Width
92KB	↑60%	↓37%	↑19%
138KB	↑63%	↓40%	↑17%
182KB	↑70%	↓45%	↑20%
248KB	↑68%	↓40%	↑25%
295KB	↑70%	↓42%	↑25%



# 降低压缩级别ROI分析

## 成本分析



原始方案

降低压缩级别方案

■ 机器成本 ■ 带宽成本

## 用户体验分析

可能会多出RTT，在国际环境下，可能会导致客户端RT增加





全球软件案例研究  
峰会

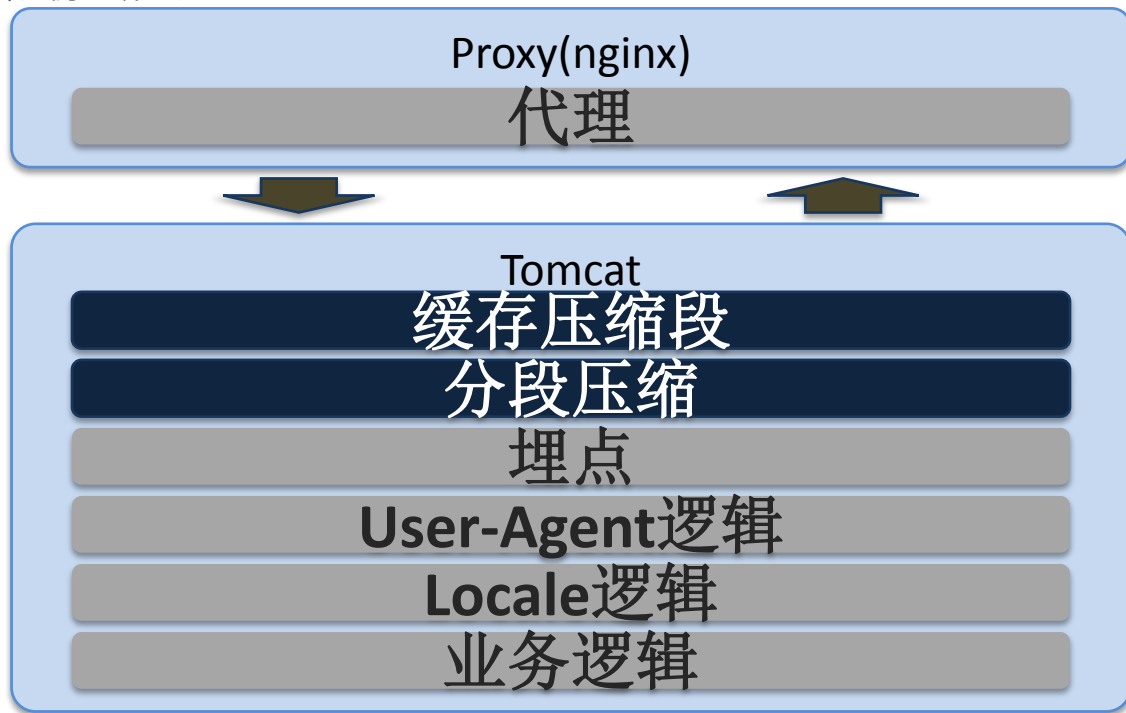
# 那 we 能把压缩去掉吗？



TOP 100 CASE STUDIES  
OF THE YEAR

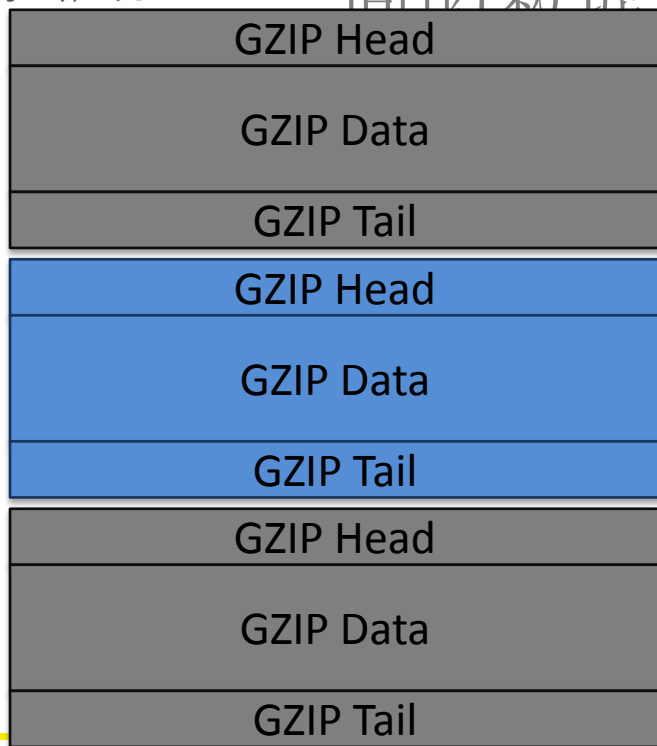
[www.top100summit.com](http://www.top100summit.com)

# 方案二：动静分离，分段预压缩



# HTTP RFC规定，HTTP server返回

## 的数据应该是一个压缩包



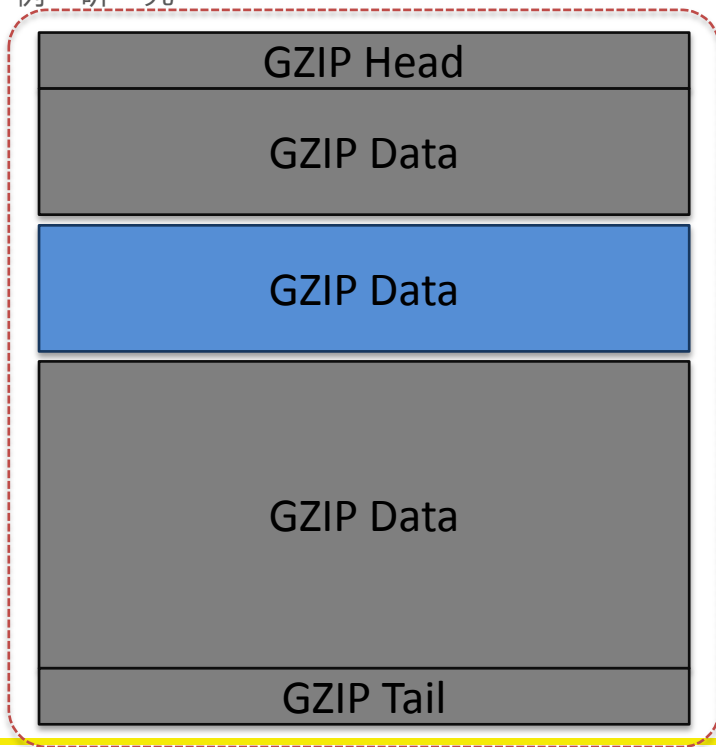
HTML片段A压缩后的数据，  
可被缓存

HTML片段B压缩后的数据，  
这个片段是动态数据

HTML片段C压缩后的数据，  
可被缓存

# 所以返回的压缩包应该是一个

整体



HTML片段A压缩后的数据，可被缓存

HTML片段B压缩后的数据，这个片段是动态数据

HTML片段C压缩后的数据，可被缓存

# 优化后效果

原始页面大小	压缩后大小	优化前QPS	优化后QPS	优化前RT	优化后RT
92KB	17KB	164	2024	60.7ms	4.9ms
138KB	8.7KB	143	1859	69.8ms	3.3ms
182KB	11.4KB	121	2083	82.3ms	4.8ms
248KB	32KB	77	1977	129.6ms	5.0ms
295KB	34KB	70	1722	141.4ms	5.8ms





全球软件案例研究  
峰会

# 案例 RIO分析

可能的话，进行投入产出分析

- 是十几个manday的人力投入
- 200台机器的需求下降到20台机器
- 回报率是相当高的



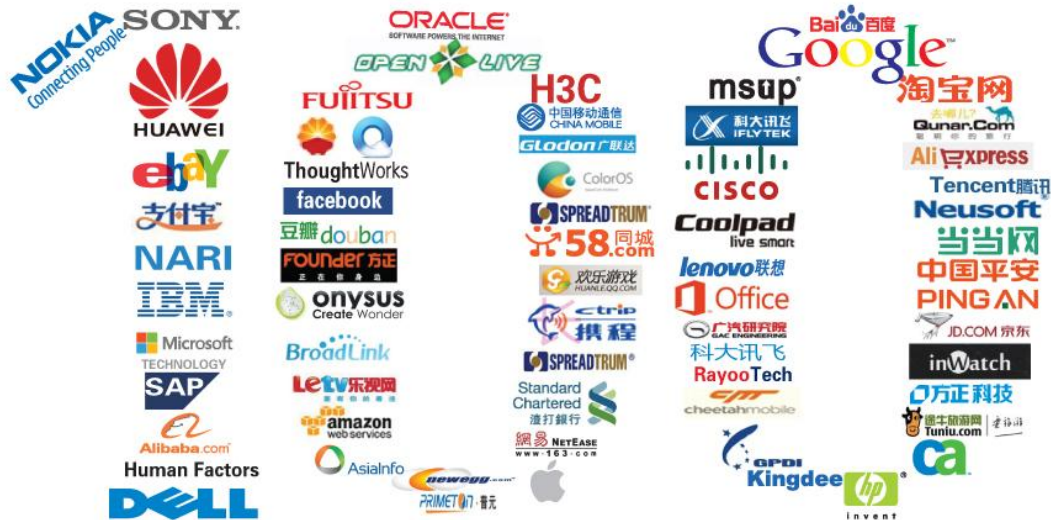
TOP 100 CASE STUDIES  
OF THE YEAR

[www.top100summit.com](http://www.top100summit.com)





全球软件案例研究  
峰会



TOP100Summit案例研究峰会 官网: [www.top100summit.com](http://www.top100summit.com)



TOP 100 CASE STUDIES  
OF THE YEAR

[www.top100summit.com](http://www.top100summit.com)

## 附录：根据公式和测试数据推导CPU TIME

$$\text{CPU Time} = \frac{1000}{\text{QPS}} \times \text{CPU 核数} \times \text{CPU 利用率}$$

优化前QPS	优化后QPS	优化前RT	优化后RT	优化前CPU Time	优化后CPU Time
164	2024	60.7ms	4.9ms	12.2ms	0.988ms
143	1859	69.8ms	3.3ms	14ms	1.1ms
121	2083	82.3ms	4.8ms	16.5ms	0.96ms
77	1977	129.6ms	5.0ms	26ms	1.0ms
70	1722	141.4ms	5.8ms	28.6ms	1.2ms

