

秒开

-----安全，快速，可靠

许霞 | 百度运维部
xuxia@baidu.com

个人简介

- 浙江大学
- 08年加入百度运维部
 - 搜索架构
 - 接入
 - 可用性
 - 搜索性能

目录

□ 速度的度量

□ 速度的优化

- ✓ 接入质量优化
- ✓ 云端服务优化
- ✓ 前端渲染优化

□ 一些无线特点

速度的度量

数据收集 - 监测方法

	客户端	JS埋点	第三方监测
样本规模	一般	大	一般
准确性	较好	准确	一般
竞品监测	可以	无法覆盖	可以定制
指标覆盖	比较固定	可以定制	固定
成本	中	低	高

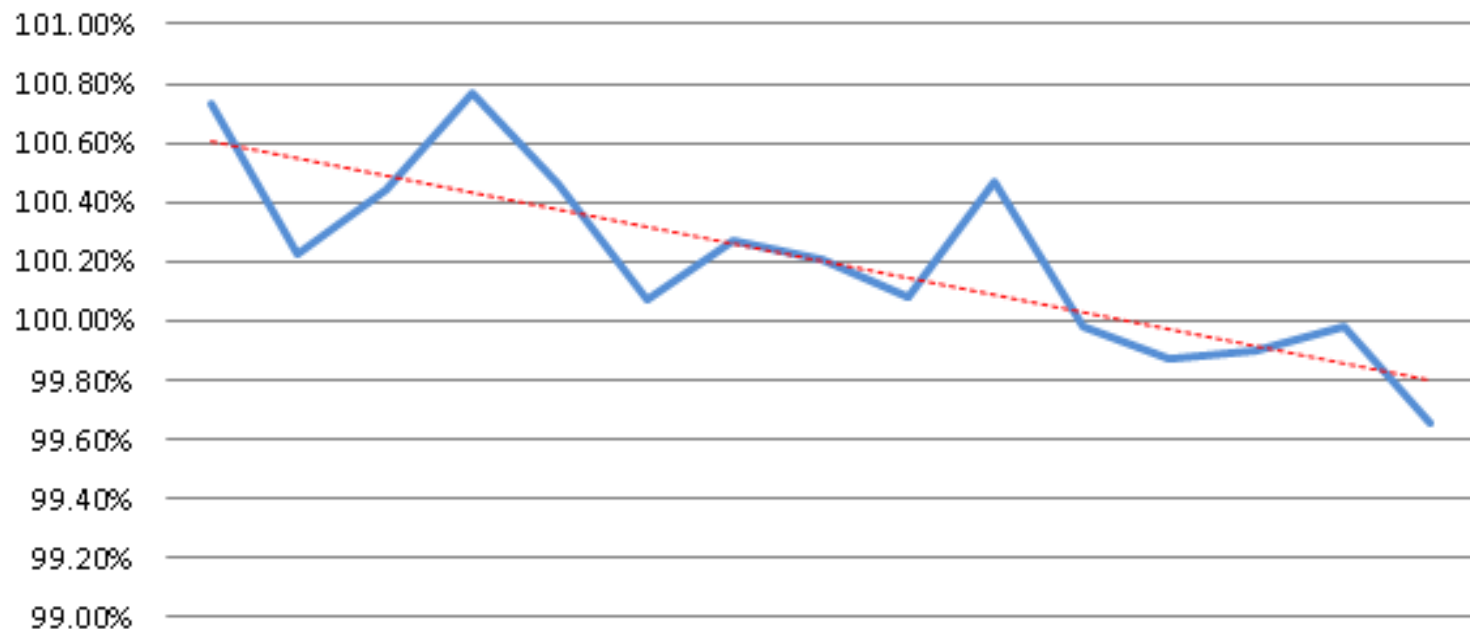
主指标竞品
监测

自身细指标
监测

竞品补充监
测

数据收集 - 优化的意义

首屏速度与PV关系



如何做优化

如何做优化

□接入质量提升

□后端处理优化

□前端渲染优化



搜索引擎的特点

full stack

如何做优化

□ 接入质量提升

- RTT

- 传输效率

□ 后端处理优化

□ 前端渲染优化

接入质量提升 | RTT

降低，减少

识别

- Localdns, clientIP, 真实信息
- IP的归属
- 速度优先？

分配

- 区域接入
- BGP
- 动态代理



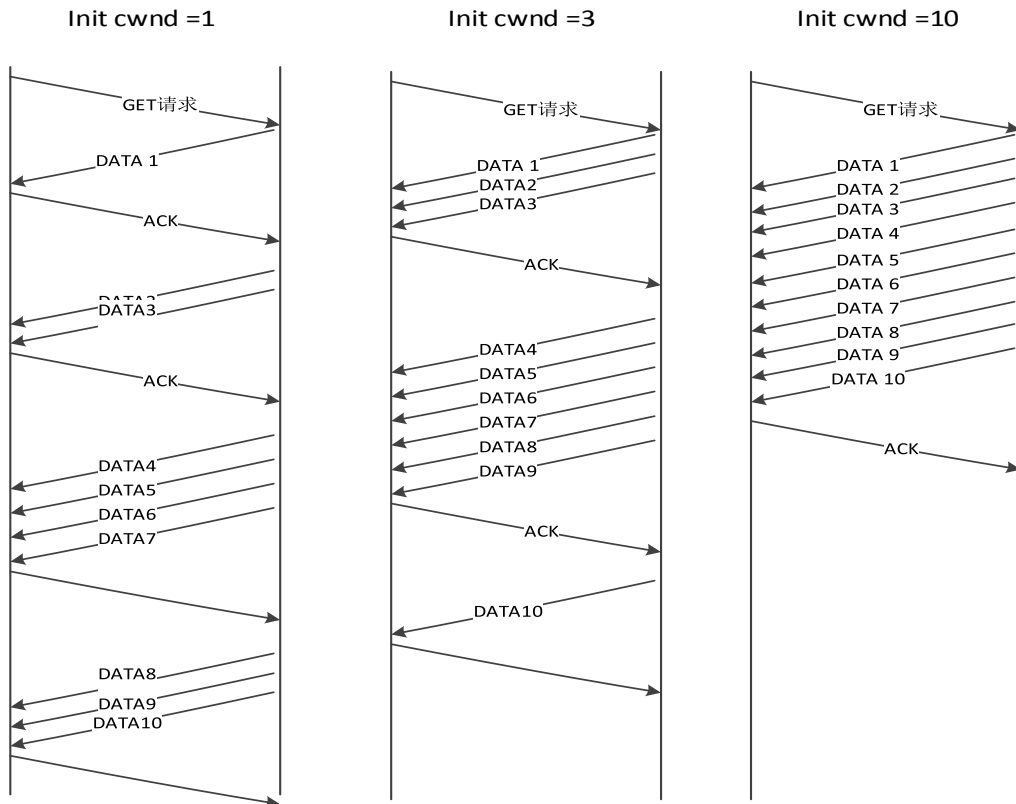
接入质量提升 | 传输

TCP

- INIT_CWND
- Tail loss probe
- TCP_FASTOPEN
- reno,cubic, westwood

SPDY

NonTCP? QUIC?



后端处理优化

Cache提升

缓存

主动更新机制

主动预测

检索优化

模块层

代码优化

架构层

部署与调度

物理层

机型、新硬件

后端处理优化 | 缓存

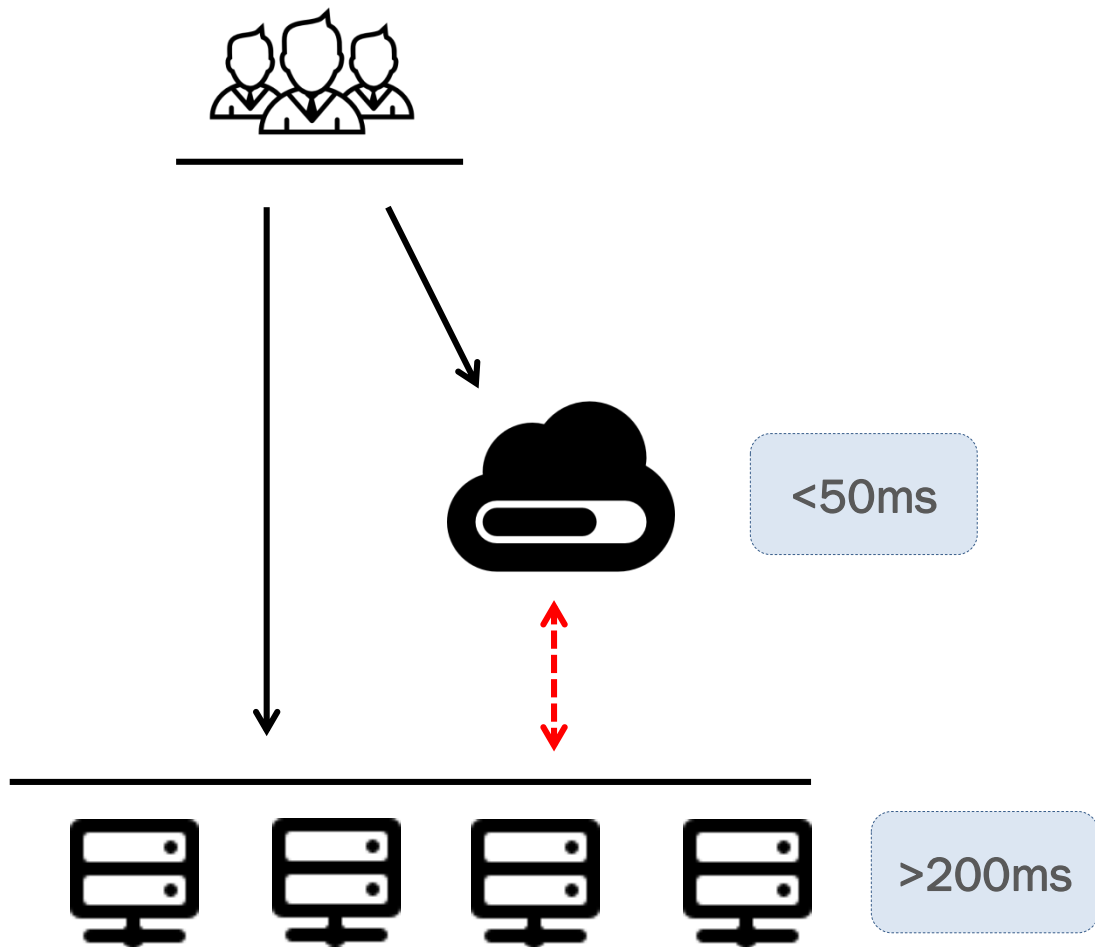
进入、淘汰 ...

主动更新

- Cache VS 时效性

主动预充

- 实时
- 非实时

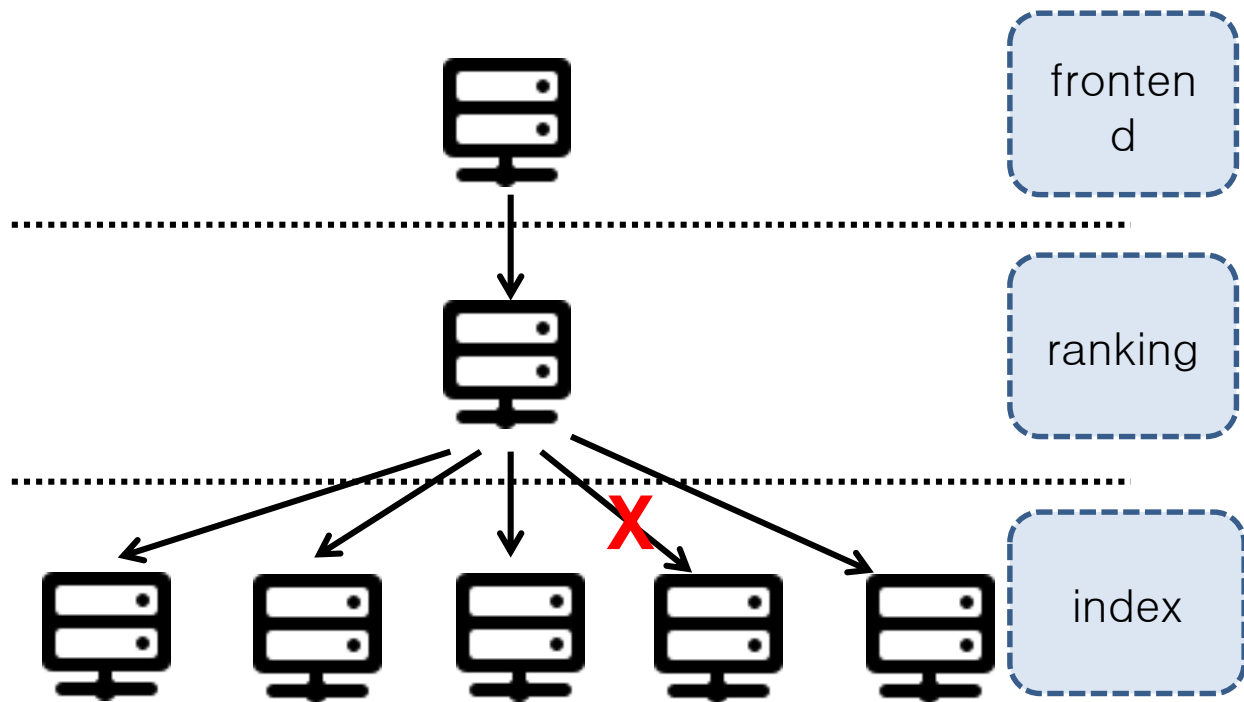


后端处理优化 | 检索

高主频 VS 高吞吐

部署调度合理性

- 快速重试
- 主动容错

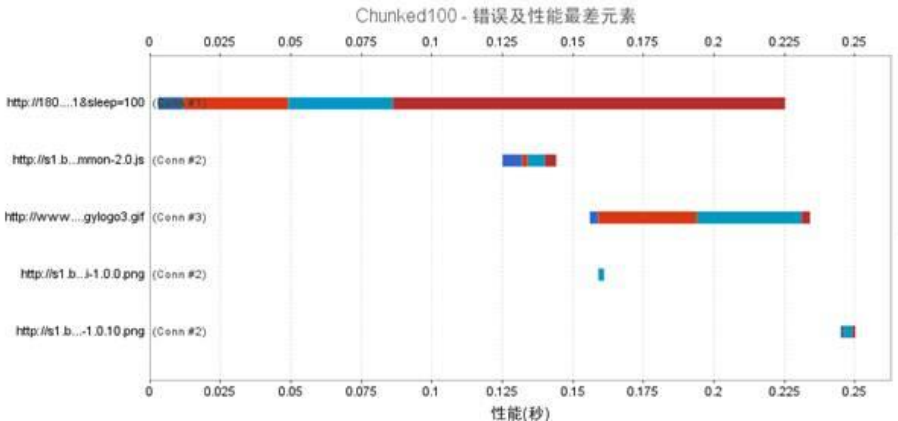


前端渲染优化

Balabala ...

Chunked

预取



[新闻](#) [网页](#) [贴吧](#) [知道](#) [音乐](#) [图片](#) [视频](#) [地图](#)

abc

百度一下

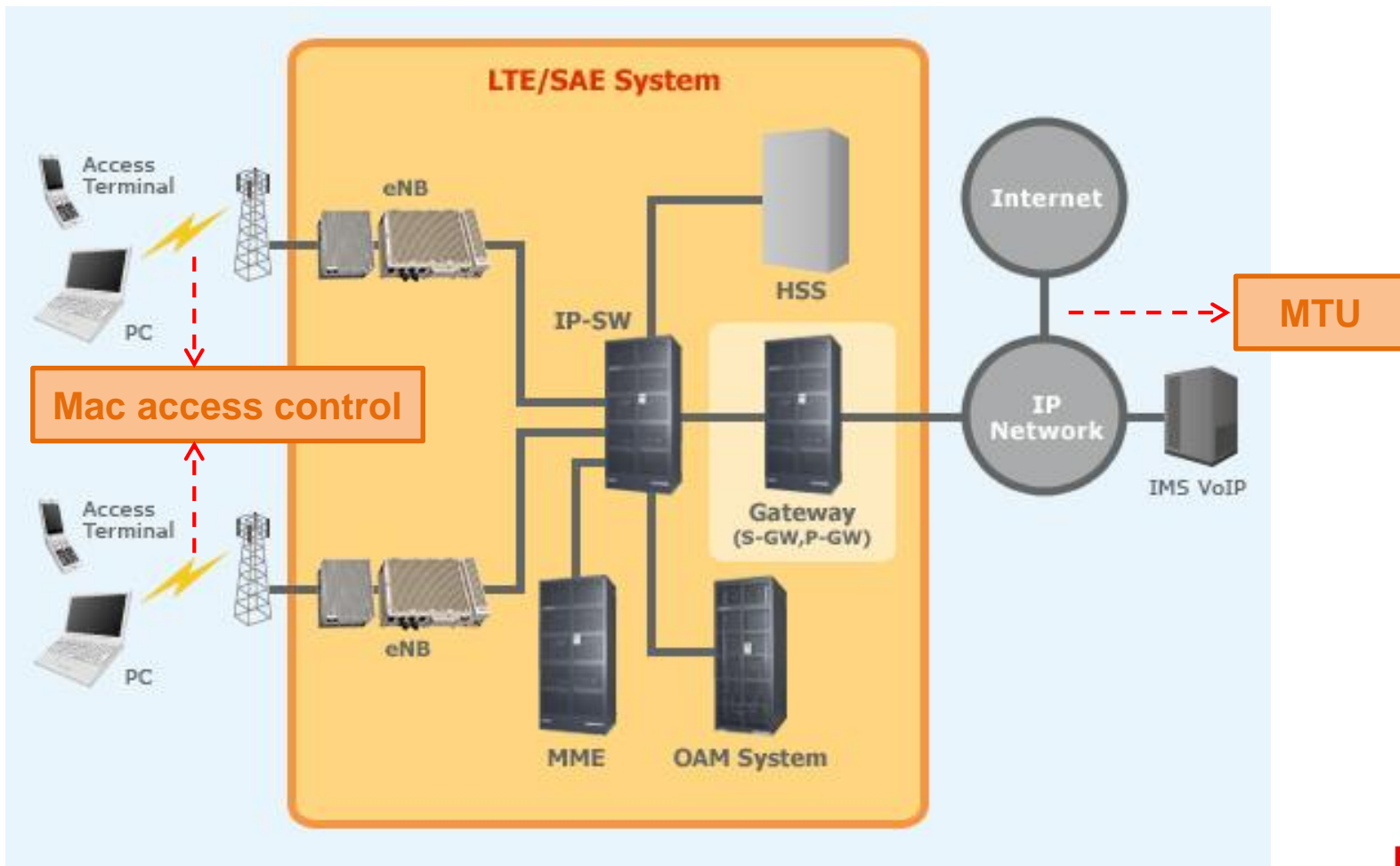
[百科](#) [文库](#) [hao123](#) | [更多>>](#)

Elements Network Sources Timeline Profiles Resources Audits Console						
e						
su?wd=ab&json=1&p=3&sid=5610_4395_5230... suggestion.baidu.com	G...	200 OK	jquery-1.10.2.min f2f...	418 B 161 B	51 ... 50 m:	
s?ie=utf-8&mod=1&sid=EAED549BC0B66666&...	G...	200 OK	jquery-1.10.2.min f2f...	1.1 KB 1.6 KB	70 ... 70 m:	
su?wd=abc&json=1&p=3&sid=5610_4395_5230... suggestion.baidu.com	G...	200 OK	jquery-1.10.2.min f2f...	449 B 207 B	60 ... 59 m:	
s?ie=utf-8&mod=1&sid=EAED549BC0B66666&...	G...	200 OK	jquery-1.10.2.min f2f...	39.5... 146 KE	59... 588 n	



关于无线

关于无线 | 控制请求头部大小



关于无线 | 控制请求头部大小

MTU & MAC(MAC access control)

•MAC: 帧长度**不受限**

运营商会根据基站负载调整调度参数

实践

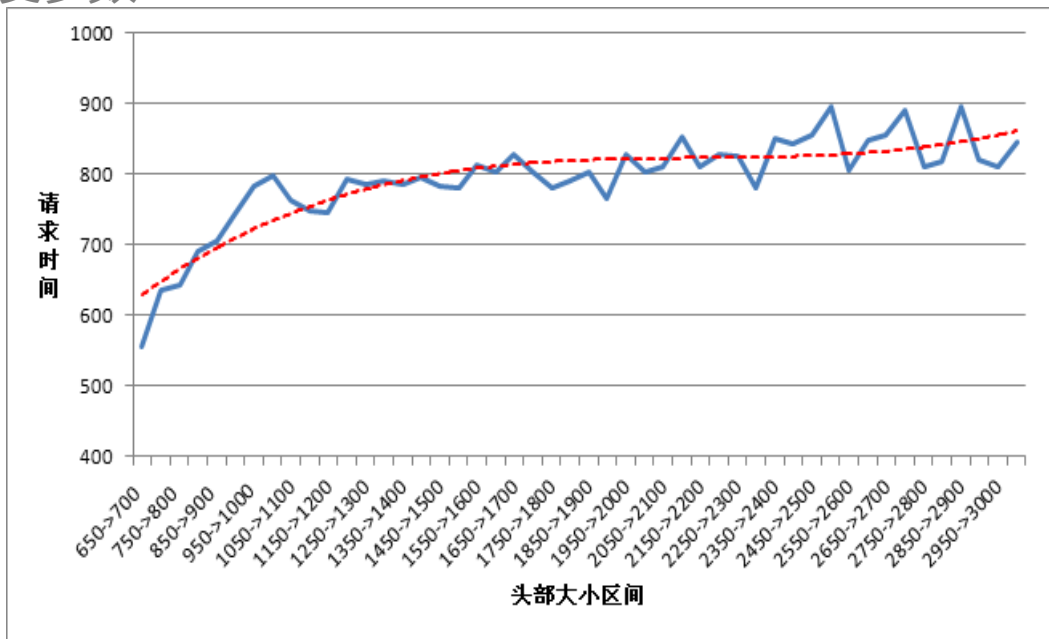
•100byte ~ **20ms**

(cookie/url/refer...)

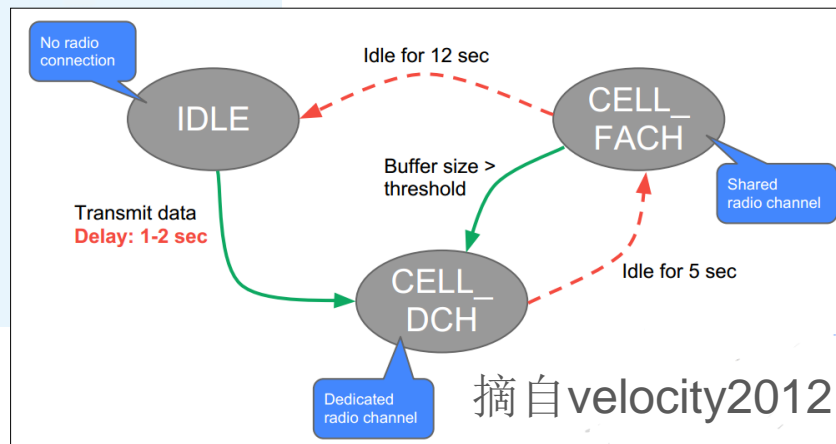
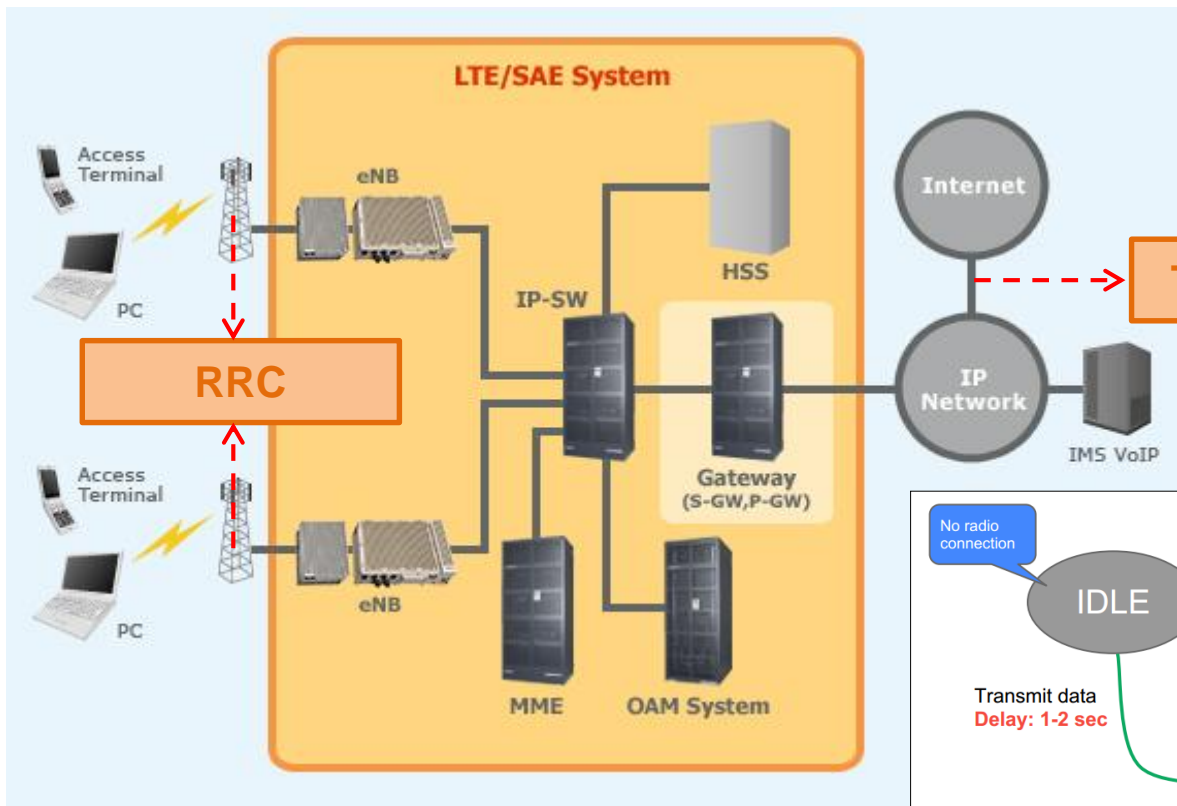
•我们的成果：100ms+

并不是**越小越好**

•微信



关于无线 | 连接的维持



关于无线 | 连接的维持

TCP: keep-alive

- 无需在终端做(PDN/GGSN , ex: facebook)
- 浏览器代理

同时，由于RRC的存在...

- 终端与基站的连接成本很高(运营商看的是吞吐)

我们的尝试

- 针对部分网络制式进行Dummy Request
- 200ms+ VS 电池

关于无线 | 大数据的使用

无线与TCP的恩怨

- 丢包 VS RTT随机波动

大数据

- 历史数据

初始窗口/慢启动节奏/丢包识别...

- 实时数据

设备识别与适配

RTT识别与适配(HTTPS/CONTENT)

...

究竟能有多快

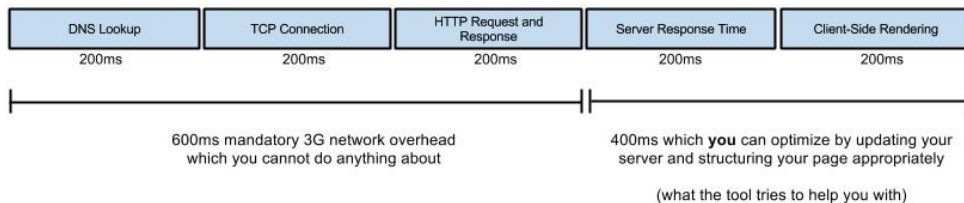
究竟能有多快

ABF: Above the fold

Kryptoners • 2013/08/09 16:19

Google要把移动端网页平均加载时间降至
秒

Rendering a mobile page in 1 second



Baidu : 80% < 1秒

