

# 改善弱网络

探索移动互联网下  
弱网络处理方式

聚焦前沿技术 传递实践经验

主办方 **Geekbang**  **InfoQ**  
极客邦科技



促进软件开发领域知识与创新的传播



关注InfoQ官方信息  
及时获取移动大会演讲  
视频信息



[深圳站] 2016年07月15-16日



[上海站] 2016年10月20-22日

# 个人介绍

07年华南理工大学毕业

10年加入腾讯，开发S60v3 手中邮

10年末成为微信的第一批程序员， 开发s60v3版本微信

12年末专注于跨平台中间件开发

信令网络(STN)， 内容分发网络(CDN)

统计/监控， 日志， 网络协议， 网络安全等

(兼容iOS/MAC, Android/Linux, Windows/Windows Phone/UWP, BB10)



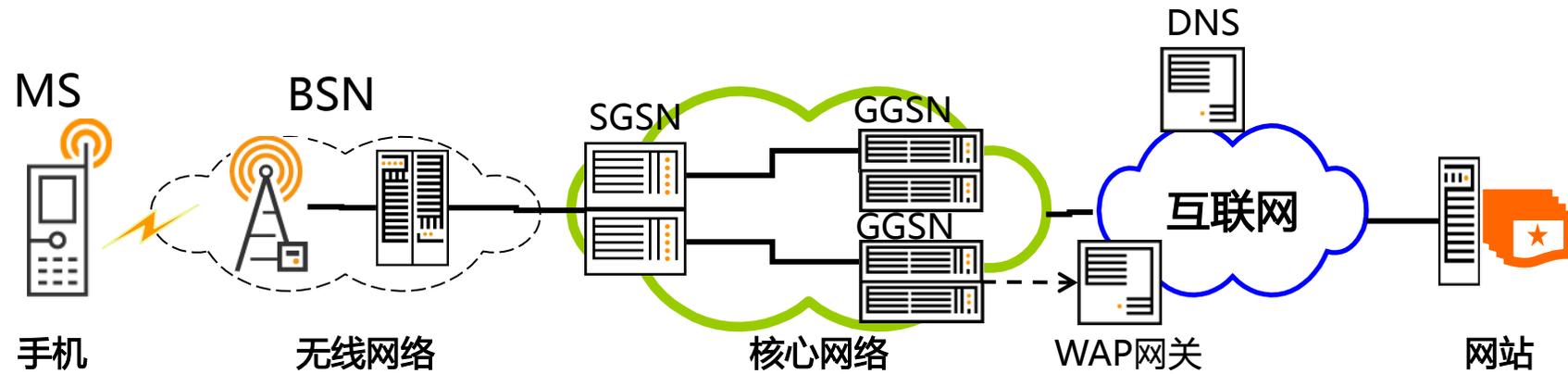
**叶润桂**

微信客户端  
基础组件负责人

# 大纲



# 物理层架构

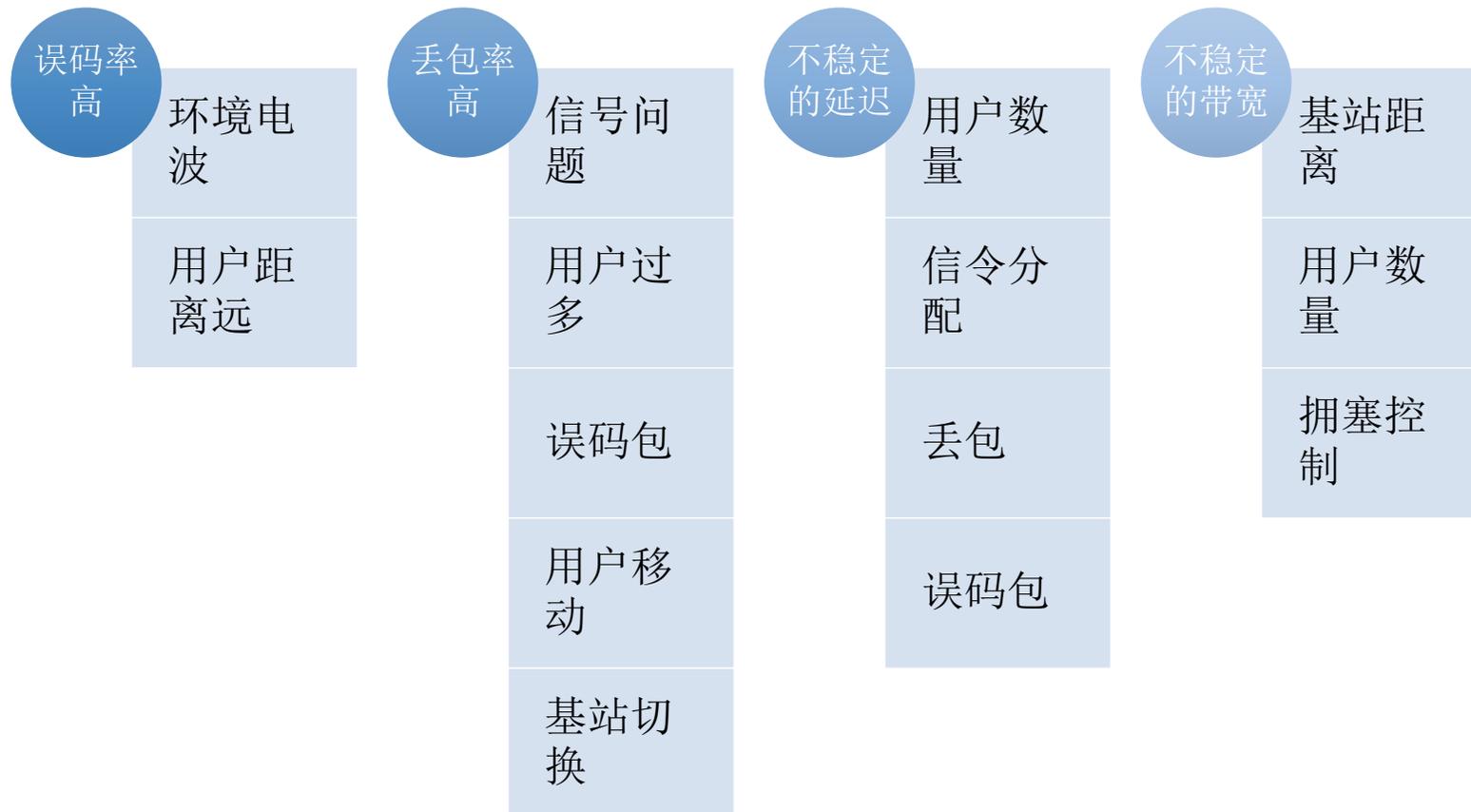


BSN 负责处理一个移动电话和网络交换子系统之间的通信流量和信令

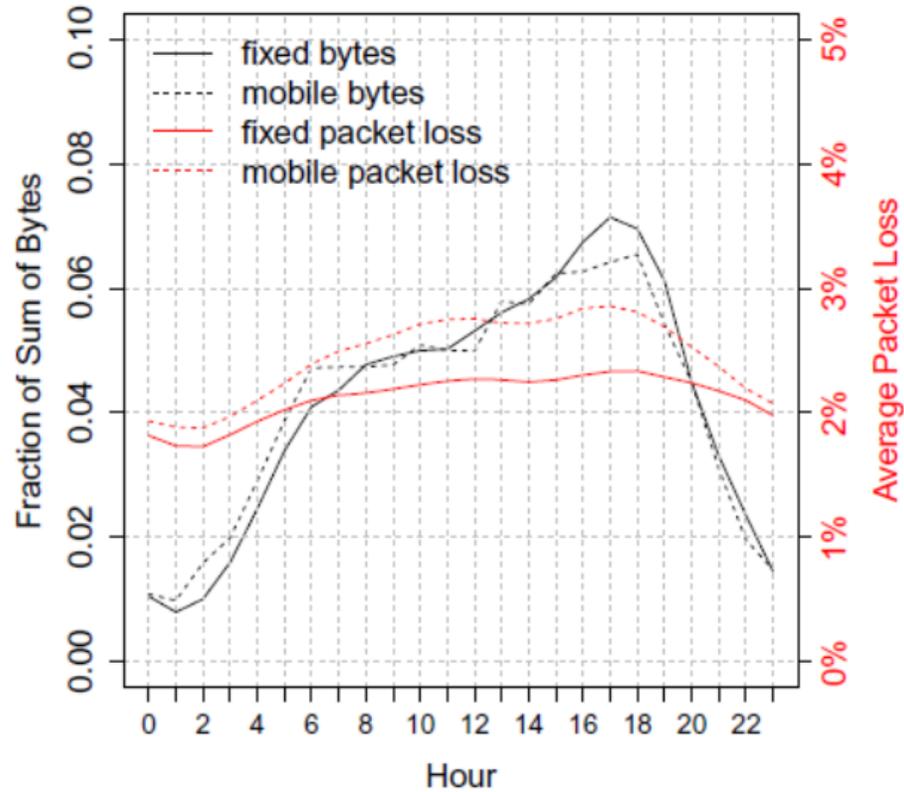
SGSN 负责在它的地理位置服务区域内从移动台接收或向其发送数据包

GGSN 负责在 GPRS 网络和外部包交换网络之间的互联

# MS和BSN间的问题



# Akimai数据分析 July 2010



✓ TCP丢包差别在1%

丢包的定义：  
SACK、DUP ACK、超时重传

统计数据连接成功后和实际使用情况有差别

Fig. 3. Daily distribution for size of connections and packet loss, all countries

# 误码率 Bit Error Rate

	BER = $10^{-5}$	BER = $10^{-6}$
Throughput (pkts/sec)	39.439	87.455
Success Probability	0.9892	0.999
Transfer time of 5000 pkts. in secs.	123.847	58.032

Wireless link has a bandwidth of 0.8 Mb and delay of 100 ms.

Source: "HALA ELAARAG" - "Improving TCP Performance over Mobile Networks" - "ACM Computing Surveys, Vol. 34, NO 3, Sep 2002, pp 357-374"

# 误码率 Bit Error Rate

✓移动互联网：  $10^{-4}$  到  $10^{-6}$

✓有线以太网：  $10^{-12}$

✓光纤：  $10^{-15}$

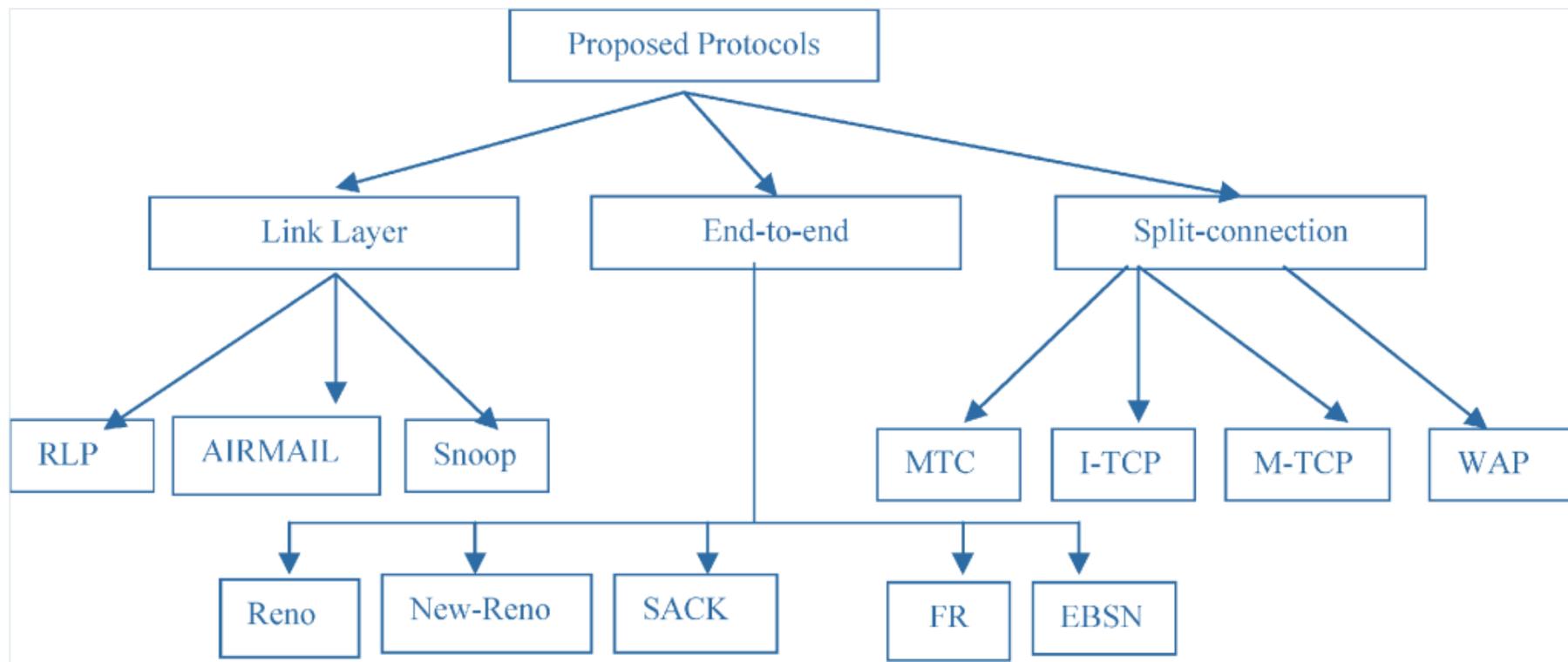
✓ BER:  $10^{-5}$  100k数据出现1bit错误

✓ 设定MTU为1500 PER: 1.5%

✓ TCP/IP直接丢弃误码包

✓ 原始无线链路的BER为  $10^{-4}$

# 业界方案



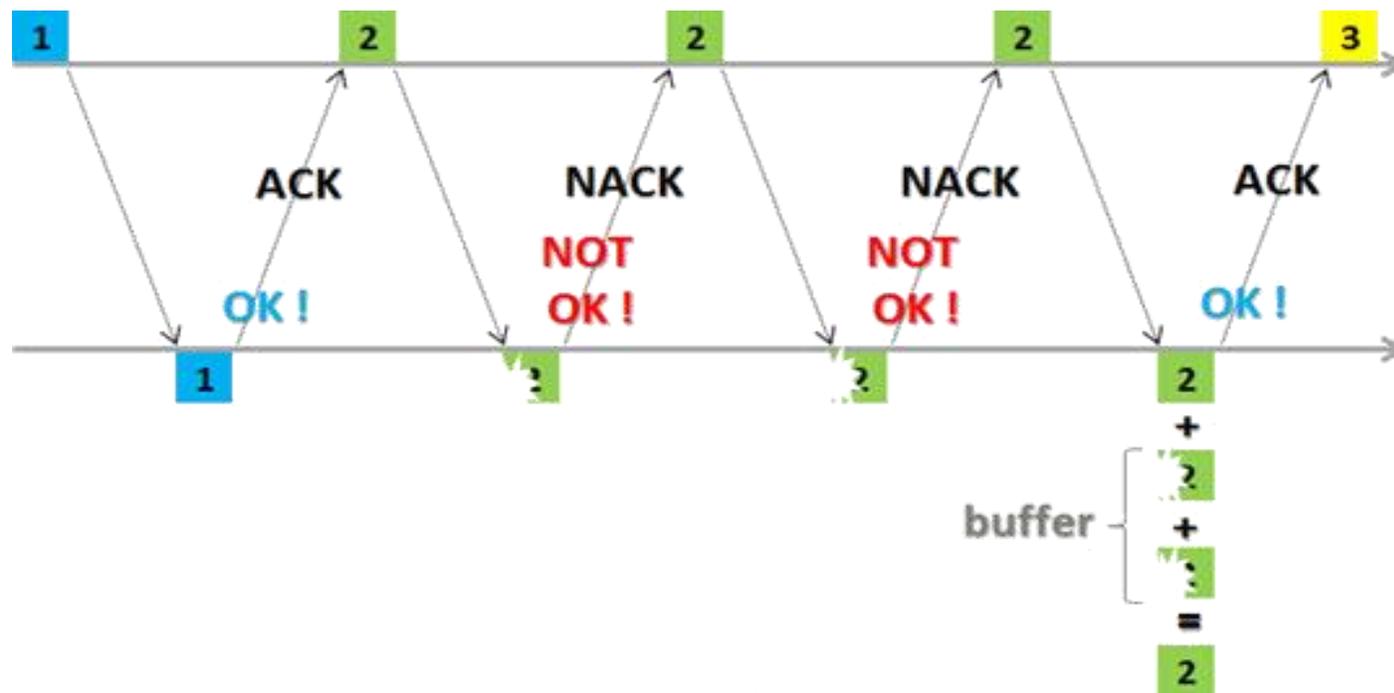
# HARQ Hybrid Automatic Repeat reQuest

✓ ARQ

自动重传请求

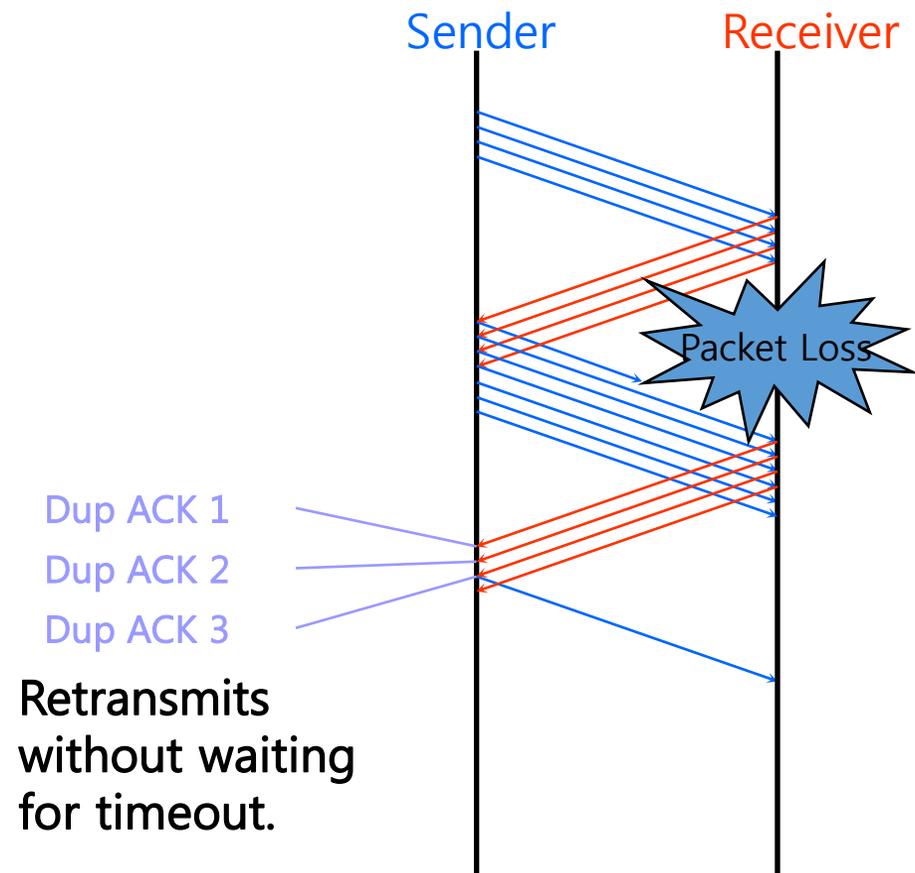
✓ FEC

前向纠错编码



✓ HARQ是3GPP的标准方案

# TCP快速重传



FR & FACK算法触发条件

$SND.FACK - SND.UNACK > 3 * MSS$

或者

$SND.DUP\_ACKS == 3$

# 总结

- ✓减小重传成本 (SACK, FEC)
- ✓尽早发现重传 (DUP ACK, FACK, RTO, NACK)
- ✓增加并发度
- ✓尽量准确避免拥堵 (丢包和拥堵的区分)

## 开发者可控的部分

- ✓ 服务器-传输层-TCP参数
- ✓ 服务器-应用层-协议和程序
- ✓ 客户端-应用层-协议和程序

# TCP可优化参数

✓ TCP Tail Loss Probe & Early

retransmit

✓ TCP Westwood+等拥塞协议

✓ TCP F-RTT

✓ TCP Hybrid Slow Start

✓ TCP FAST OPEN

✓ TCP INIT CWND

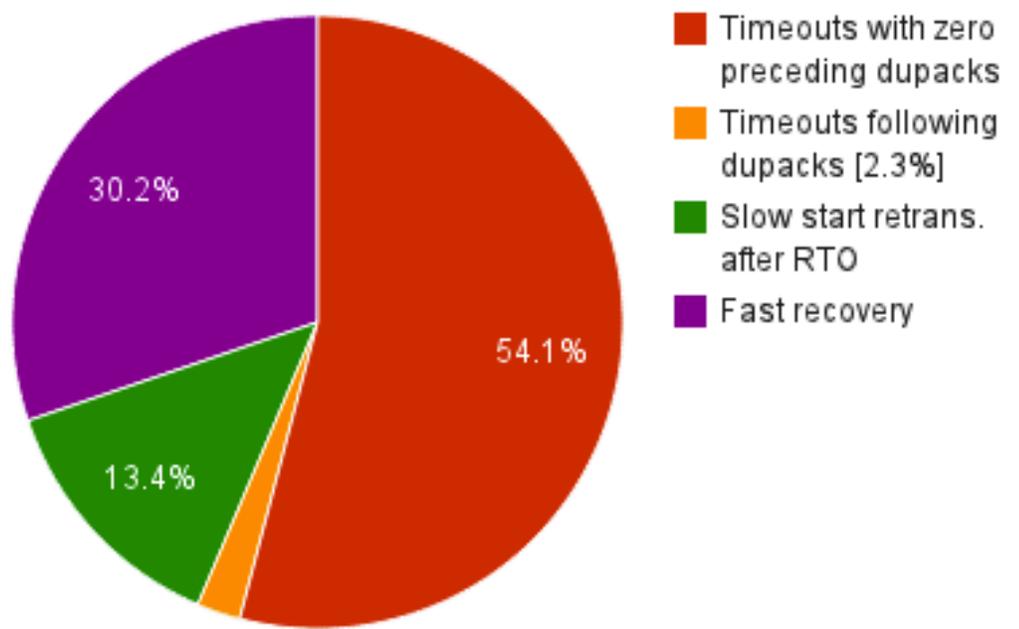
✓ TCP INIT RTT

✓ TCP Thin Stream(和ER冲突)

✓ TCP PRR

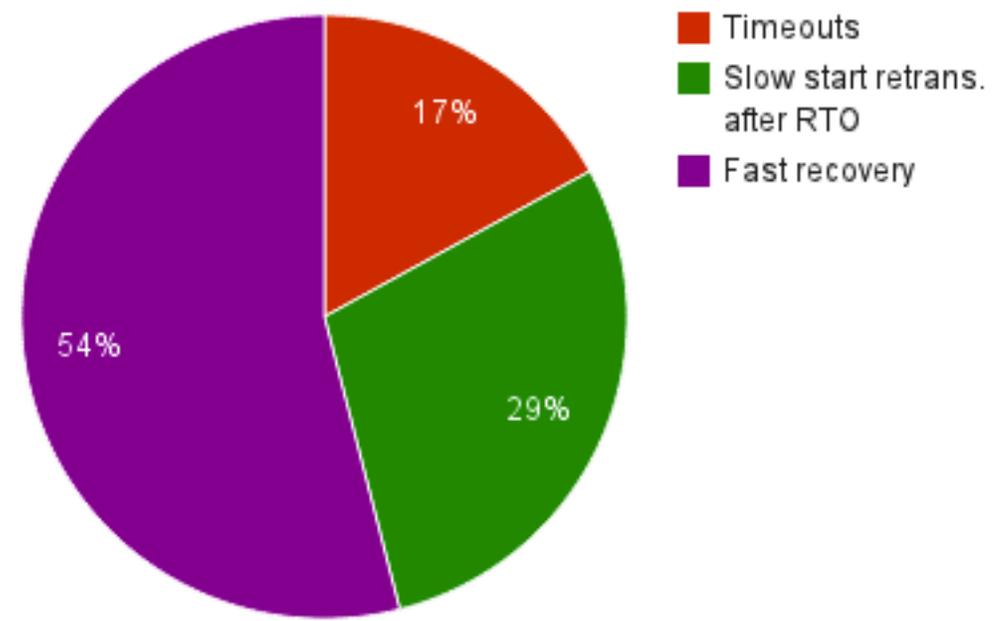
# TCP丢包的恢复方式

Google WEB



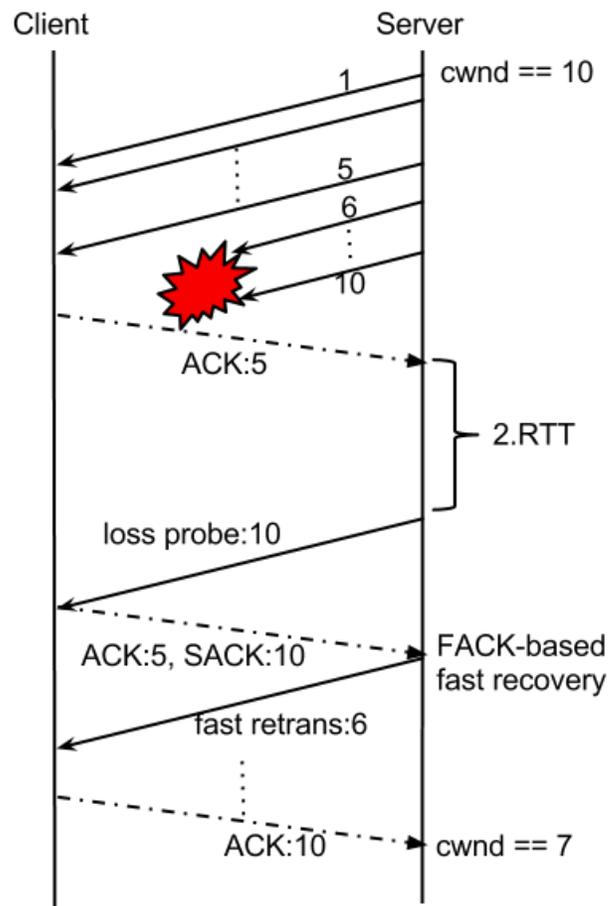
✓ 尾包丢失是首包丢失的两倍

YouTube



✓ 丢包请求耗时是非丢包请求10倍

# TLP原理

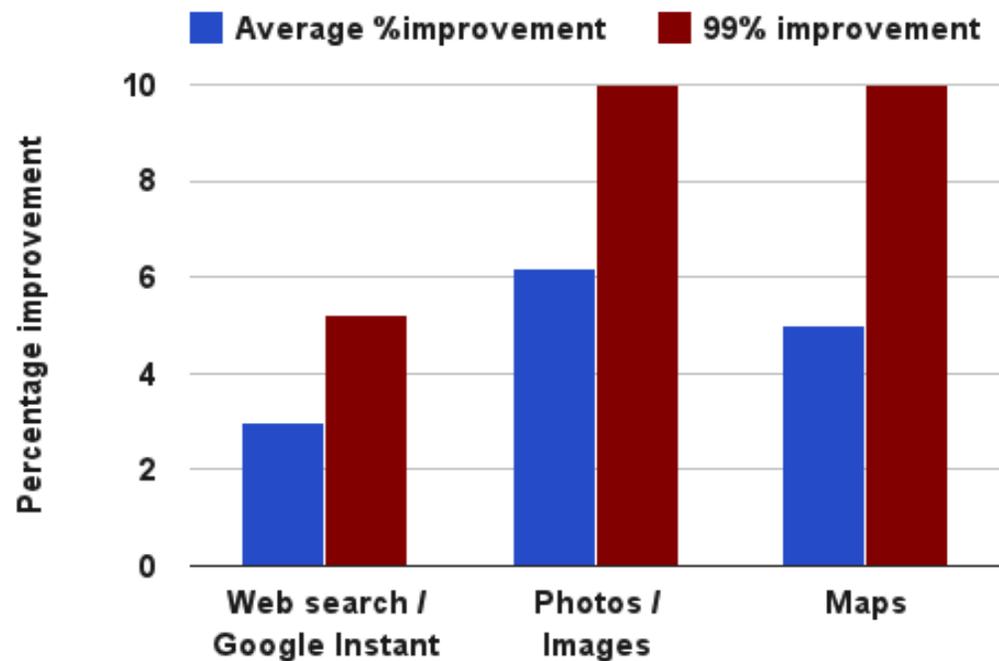


1. PT0触发尾包重传
2. 尾包的ACK带上SACK信息
3. SACK触发FACK快速重传和恢复
4. 避免了RTO导致的慢启动和延迟

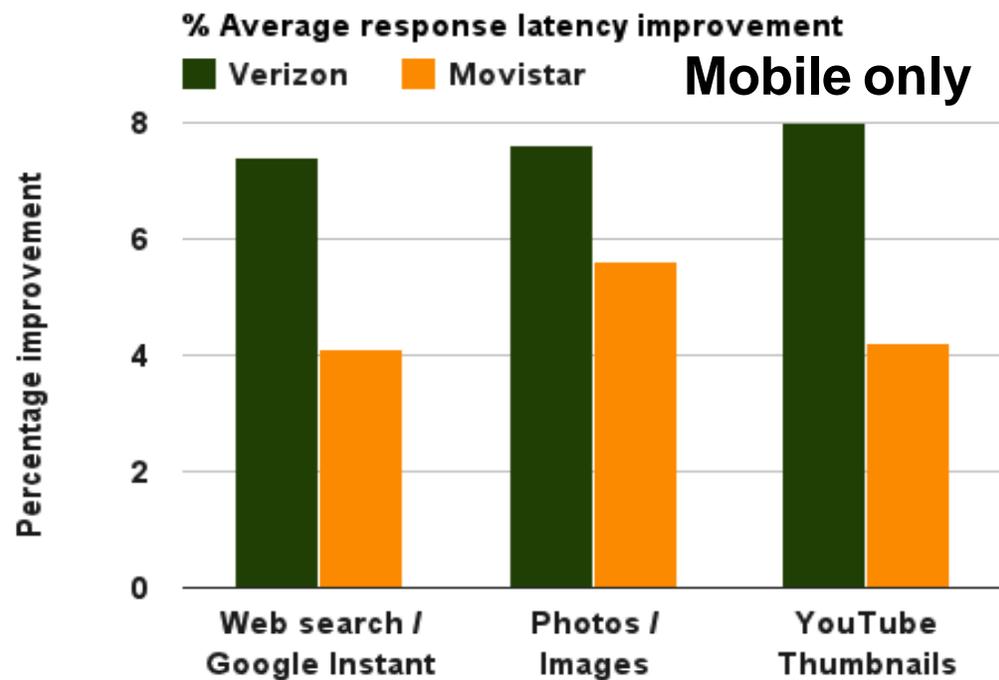
# TLP对重传补充

#losses	scoreboard after TLP ACKed	mechanism	outcome
AAAL	AAAA	TLP loss detection	All repaired
AALL	AALS	Early retransmit	All repaired
ALLL	ALLS	Early retransmit	All repaired
LLLL	LLLS	FAACK fast recovery	All repaired
>=5 L	...LS	FAACK fast recovery	All repaired

# TLP效果

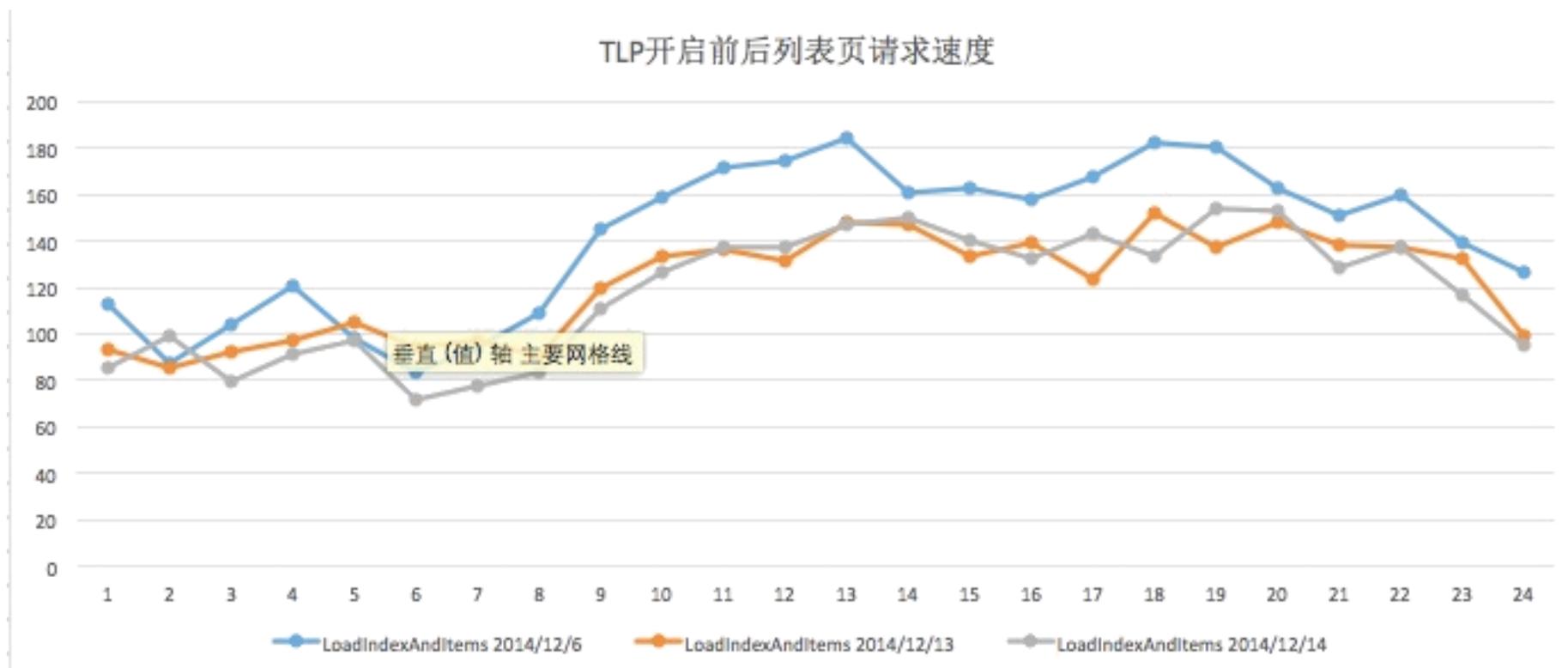


✓ 6%减小的图片搜索延迟.



✓ 10%减小RTO重传.

# TLP在腾讯的实践



# TCP一些思考

虽然TCP不断改进下，越来越适合移动互联网

但从RFC制定到Kernel层实现到用户覆盖很漫长

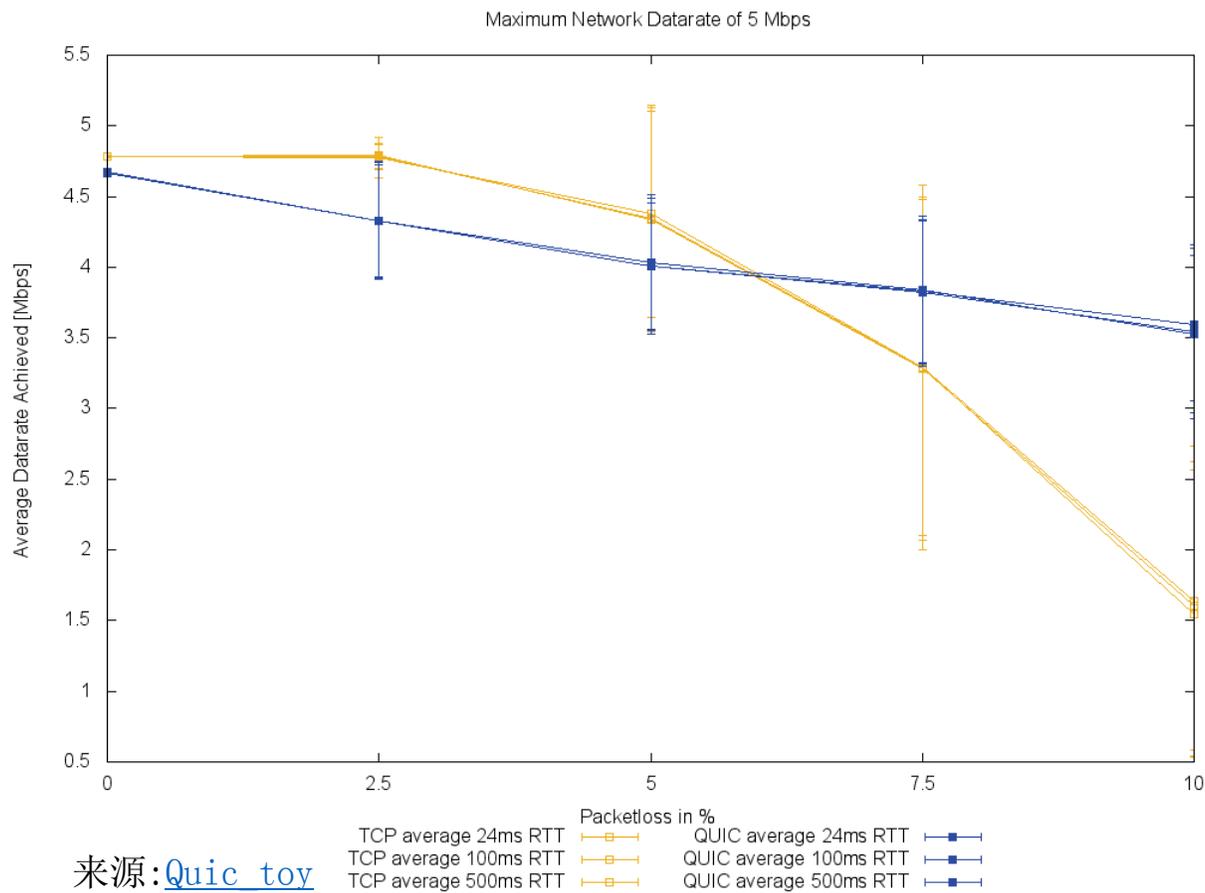
新特性(NACK & FEC)接纳慢

别开蹊径的QUIC

# QUIC Quick UDP Internet Connection

1. RFC 6298 (RTT computation)
2. FACK Loss Recovery (paper)
3. RFC 3782, RFC 6582 (NewReno Fast Recovery)
4. TLP (draft)
5. RFC 5827 (Early Retransmit) with Delay Timer
6. RFC 5827 (F-RTT)
7. RFC 6937 (Proportional Rate Reduction)
8. TCP Cubic (draft) with optional RFC 5681 (Reno)
9. Hybrid Slow Start (paper)
10. FEC & NACK
11. Head of line block
12. 0-RTT Connect
13. Sprout-EWMA (congestion control)

# QUIC



可以说QUIC就是为移动互联网量身定做

- ✓ QUIC十分合适做可复用连接
- ✓ 由于互联网NAT实现问题, 不合适做长连接PUSH通道

# 应用层策略

## 通用策略

- 复合连接
- 合理的超时
- 减少数据量
- 协议合并
- 合包发送
- 业务重试
- 网络敏感重试

## 小数据

- 相对超时(读/写)
- 绝对超时
  - 首包
  - 读写
- 动态超时
- PB Zlib

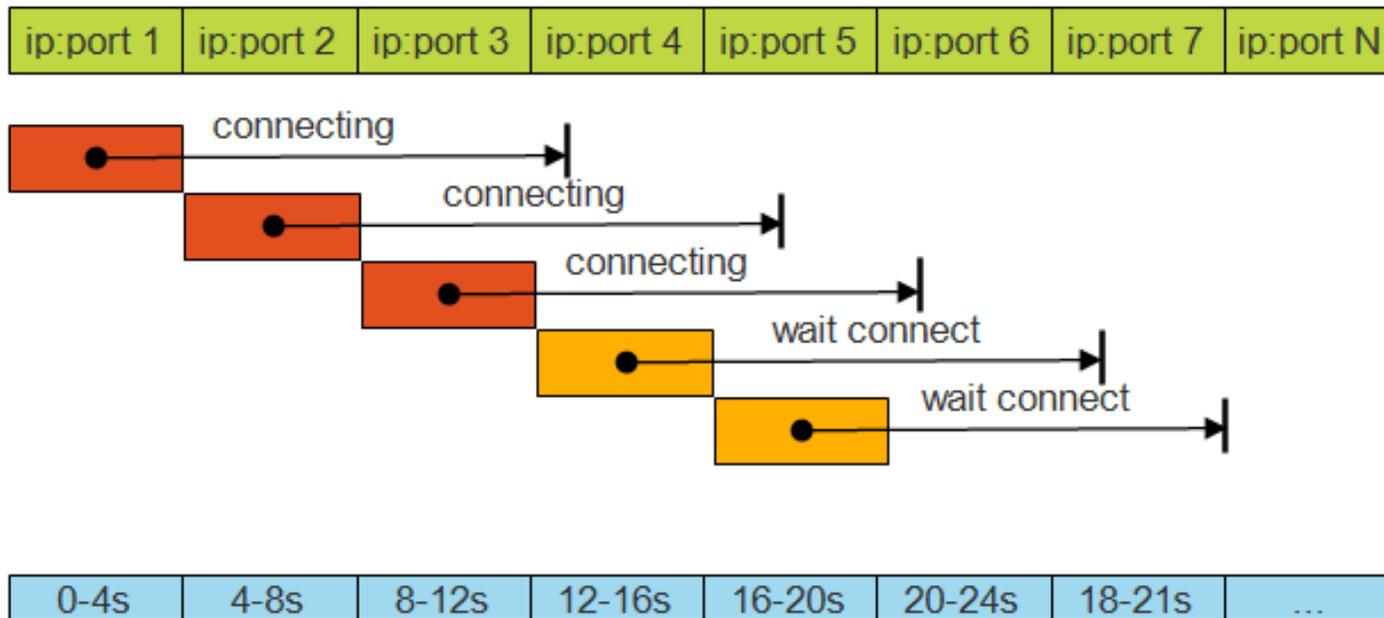
## 大数据

- 图片webp  
hevc
- 低成本重传
- 有损上传
- 有损下载

✓ 发图

✓ 事务型

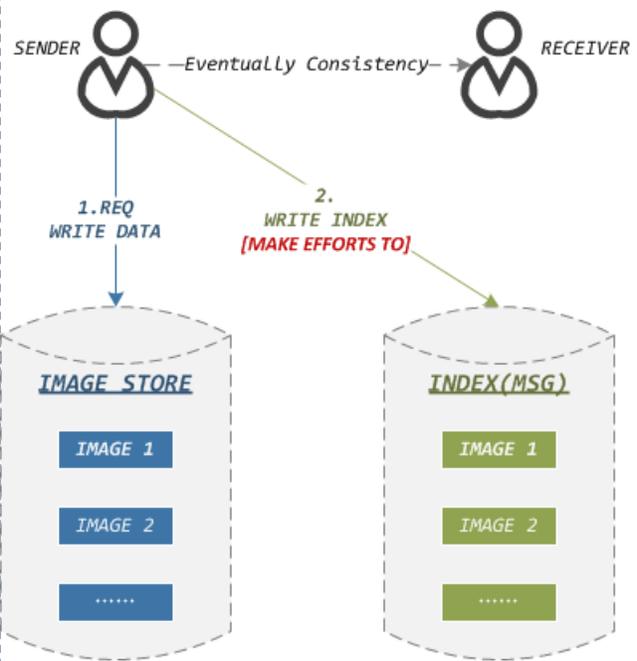
# 发图-复合连接



1. ip1+port1 0s连接, 10s超时
  2. ip2+port2 4s连接, 14s超时
  3. ip3+port3 8s连接, 18s超时
  - ...
- 任一连接成功, 关闭其他连接

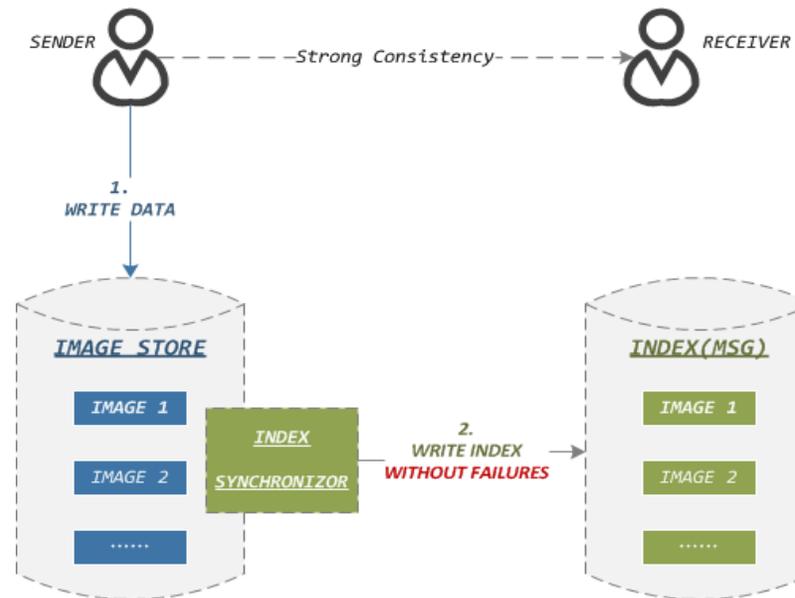
- ✓ 连接成功率提升5%
- ✓ 更快找到可用链路和IP轮转

• 耗时 = T1 + T2

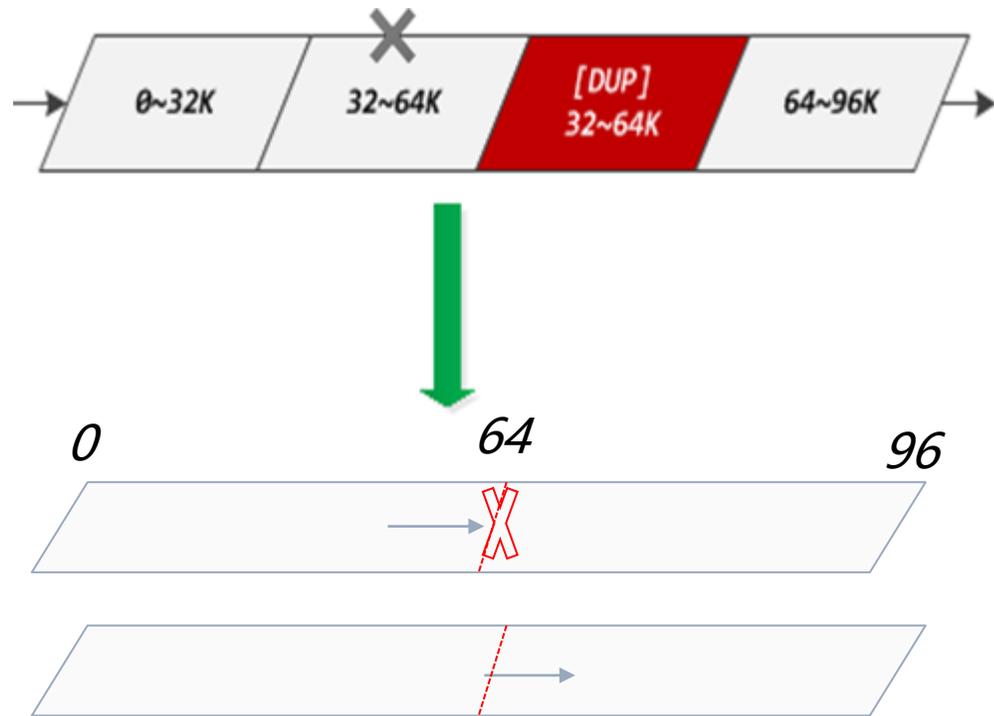


• 耗时 = T1

• (20% FASTER)



# 发图-低成本重传



## 分包

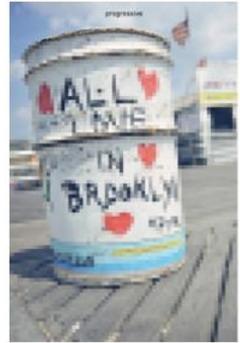
- 降低包大小
- 增加并发
- 包头损耗

## 流式

- 确认粒度策略灵活
- 单线程

# 发图-渐进式图片

- ✓ JPG支持不完整数据的解码
- ✓ 利用这个特性，可以增加弱网络下的可用性



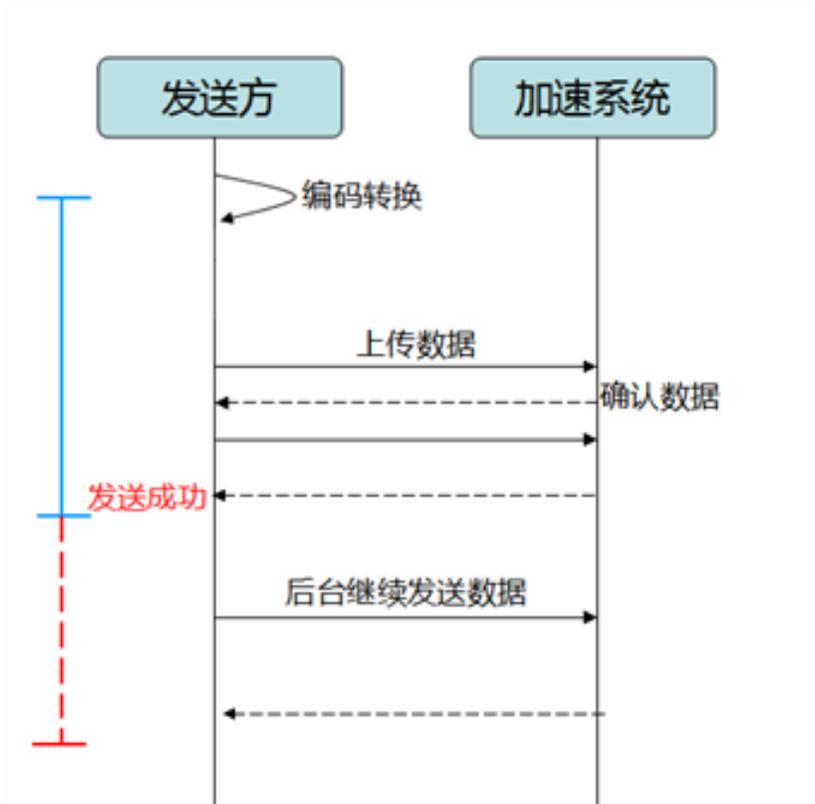
基线式图片

渐进式图片

JPEG渐进式在编码的时候计算开销大约是基线式的3-5倍，两者编码出来的文件大小基本相同

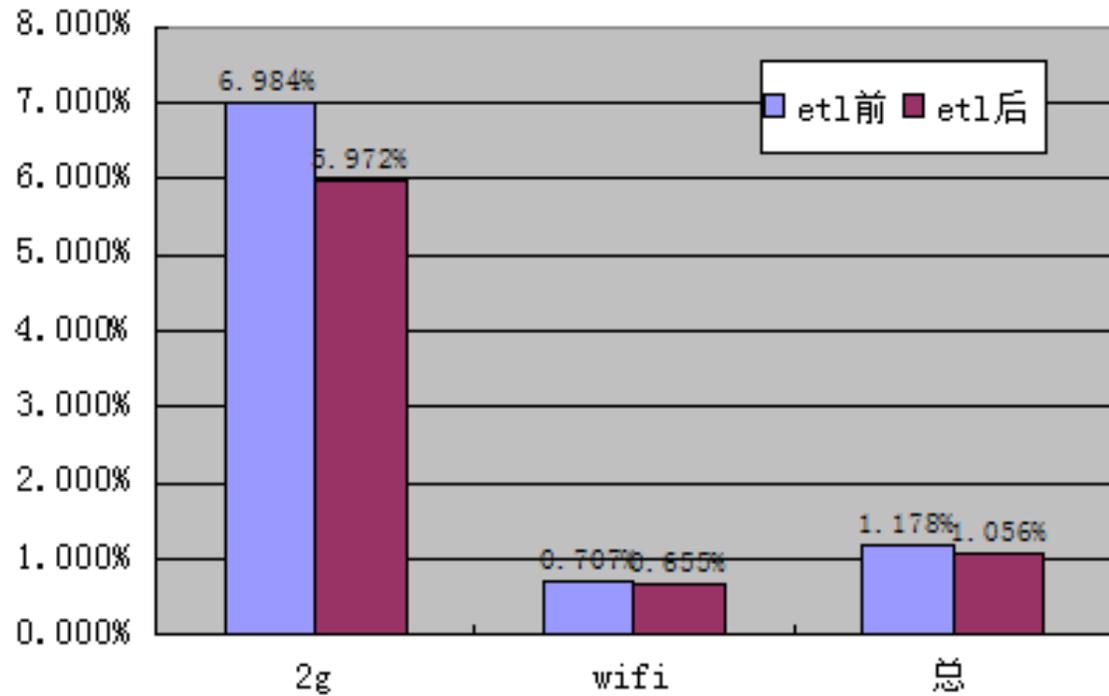
除了JPEG 2000，支持渐进式解码的图片，我们都可以用来做有损服务

# 发图-有损上传



1. 发送渐进式图片
2. 服务器接收数据且回复数据确认包
3. 当数据足够时候(50%)，回复发送成功确认包
4. 发送方继续补充数据
  - 网络正常，数据完整
  - 网络异常，认为已发成功
5. 服务器通知接收者

# 发图-有损上传数据



失败率对比

效果：

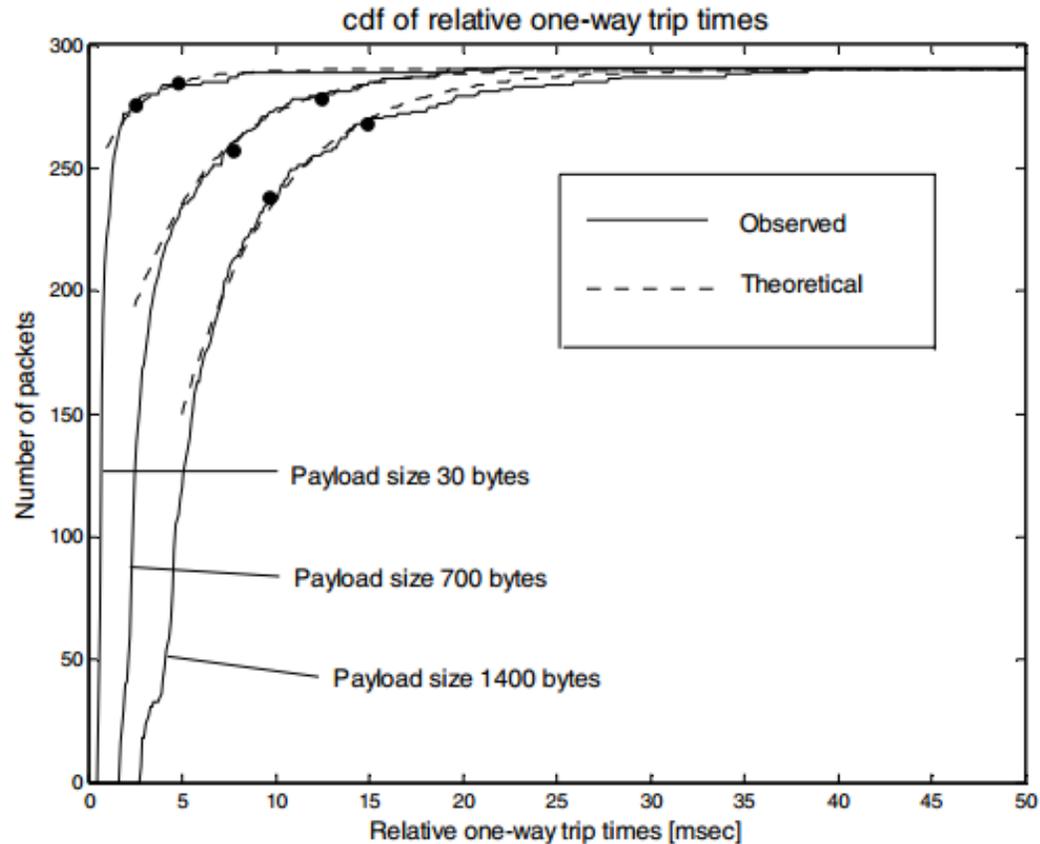
客户端总体失败率降低10.39%

2g环境客户端总体失败率降低达14.49%

# 发图-有损下载



# 小数据-数据量的影响

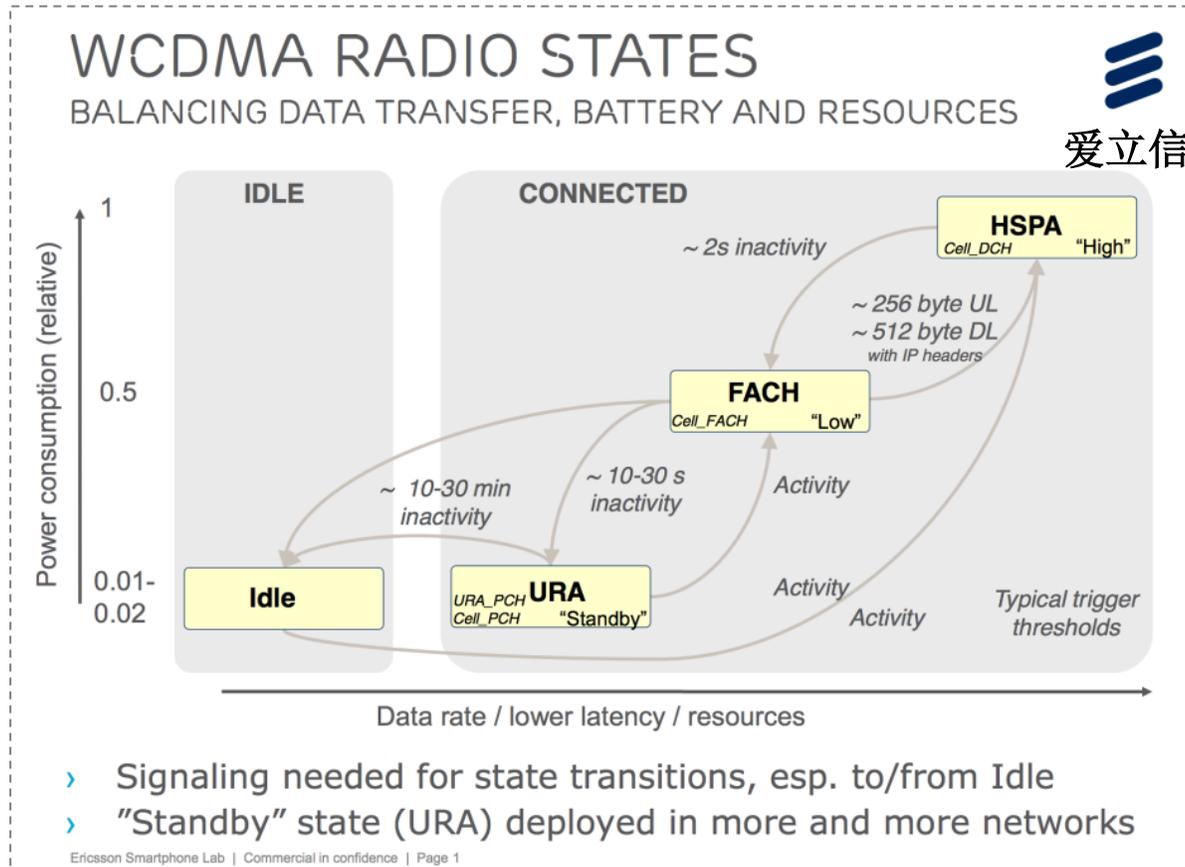


Payload size	$P$
30 bytes	0.1054
700 bytes	0.3203
1400 bytes	0.4860

- ✓ 延迟受数据量影响
- ✓ 丢包受数据量影响 (BER)
- ✓ 丢包会增加延迟 (link layer 重传)

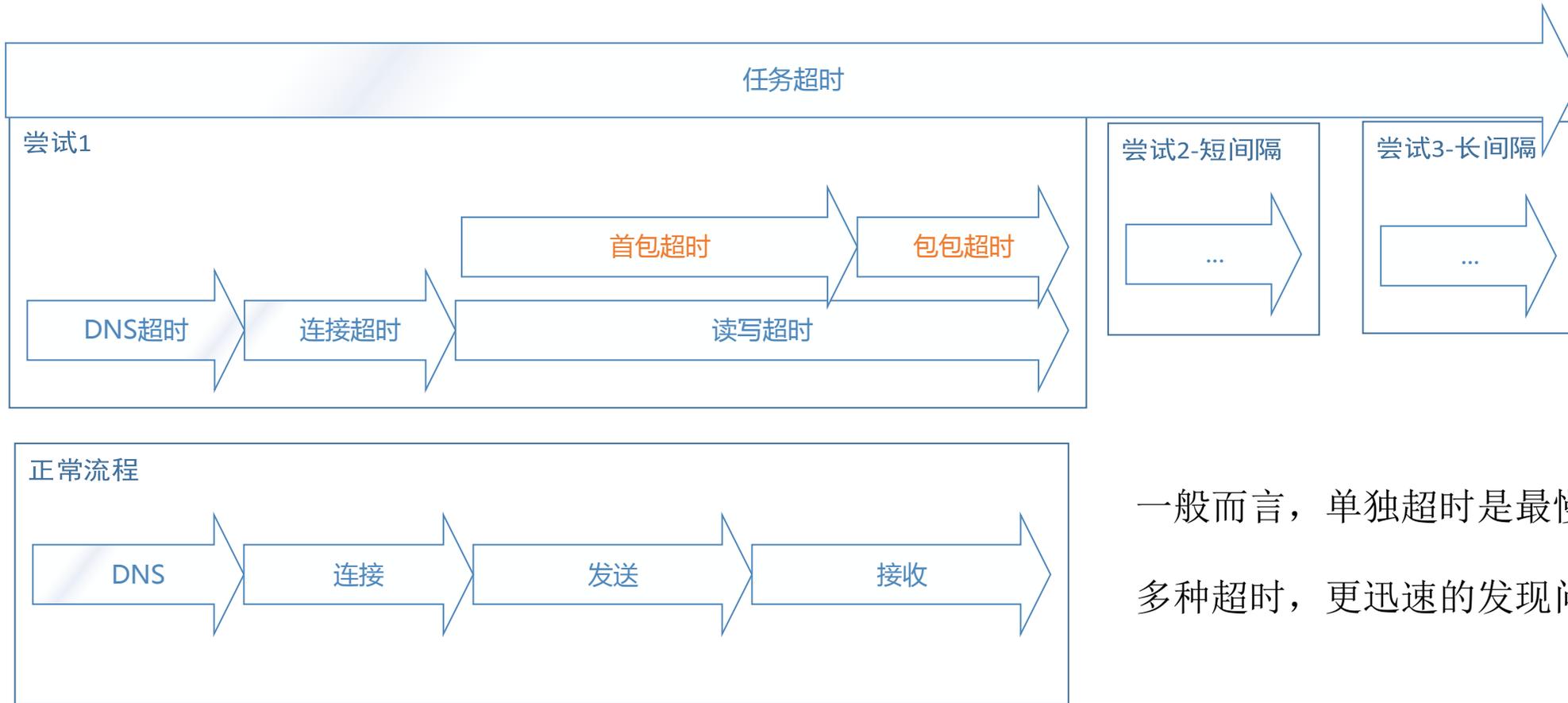
来源:Effect of Packet Size on Loss Rate and Delay in Wireless Links

# 事务-数据量的影响



- ✓ 超时信令回收
- ✓ 信令分配耗时高
- ✓ 上传或者下载超过阈值信道会跳变

# 事务-多种超时



一般而言，单独超时是最慢的(例如RTO)

多种超时，更迅速的发现问题(例如TLP)

# 事务-其他超时

## 包包超时

- 每次读取或发送的间隔

## 动态超时

- 根据网络情况，调整其他超时的系数或绝对值

## 包包超时

获取sock snd buf内未发数据

iOS:

getsockopt 读取SO\_NWRITE

Android:

ioctl 读取 SIOCOUTQ

Balakrishnan, Hari, et al. "Improving TCP/IP performance over wireless networks." Acm Conf on Mobile Computing & Networking 2013:2-11.

"HALA ELAARAG" - "Improving TCP Performance over Mobile Networks" – "ACM Computing Surveys, Vol. 34, NO 3, Sep 2002, pp 357-374"

Mikko V. J. Heikkinen, and Arthur W. Berger. "Comparison of User Traffic Characteristics on Mobile-Access versus Fixed-Access Networks." International Conference on Passive and Active Measurement 2012:32-41.

Cheng, Yuchung, et al. "Tail Loss Probe (TLP): An Algorithm for Fast Recovery of Tail Losses." (2013).

Elaarag, Hala. "Improving TCP performance over mobile networks." *Acm Computing Surveys* 34.3(2013):357-374.

Korhonen, J., and Y. Wang. "Effect of packet size on loss rate and delay in wireless links." *3(2005):1608-1613 Vol. 3.*

Iyengar J, Swett I. QUIC Loss Recovery And Congestion Control[J]. 2016.

# Mars

跨平台开源组件

STN-信令网络，小数据传输

CDN-数据分发网络，大数据传输

XLOG-高性能客户端日志组件

# 联系我们



rayye@tencent.com



微信终端开发公众号

# THANKS

聚焦前沿技术 传递实践经验

主办方 **Geekbang**  **InfoQ**  
极客邦科技