



从创新到落地

——以互联网数据传输优化的工业实践为例

李彤

中国人民大学 信息学院/DEKE重点实验室

2022.05.30

提示：
需要一定的技术背景

目录



1 引言



2 创新故事：TACK



3 创新故事：Contact Shield



4 从创新到落地



中國人民大學
RENMIN UNIVERSITY OF CHINA

什么是创新？



引言



“创新” vs “落地”

什么是创新



网页 图片 视频 学术 词典 地图

51,000,000 条结果 时间不限

人的创造性 实践行为

创新从哲学上说是一种人的创造性实践行为，这种实践为的是增加利益总量，需要对事物和发现的利用和再创造，特别是对物质世界矛盾的利用和再创造。人类通过对物质世界的利用和再创造，制造新的矛盾关系，形成新的物质形态。

中文名: 创新
含义: 创造新事物、更新、创造、改变
外文名: innovation
拼音: chuàng xīn

创新

落地

什么才是落地



网页 图片 视频 学术 词典 地图

492,000 条结果 时间不限

让企业从知道到做到，从做到到做好

所谓落地，就是让企业从知道到做到，从做到到做好。

创新

落地



中國人民大學
RENMIN UNIVERSITY OF CHINA

TACK : 互联网传输优化



创新故事一

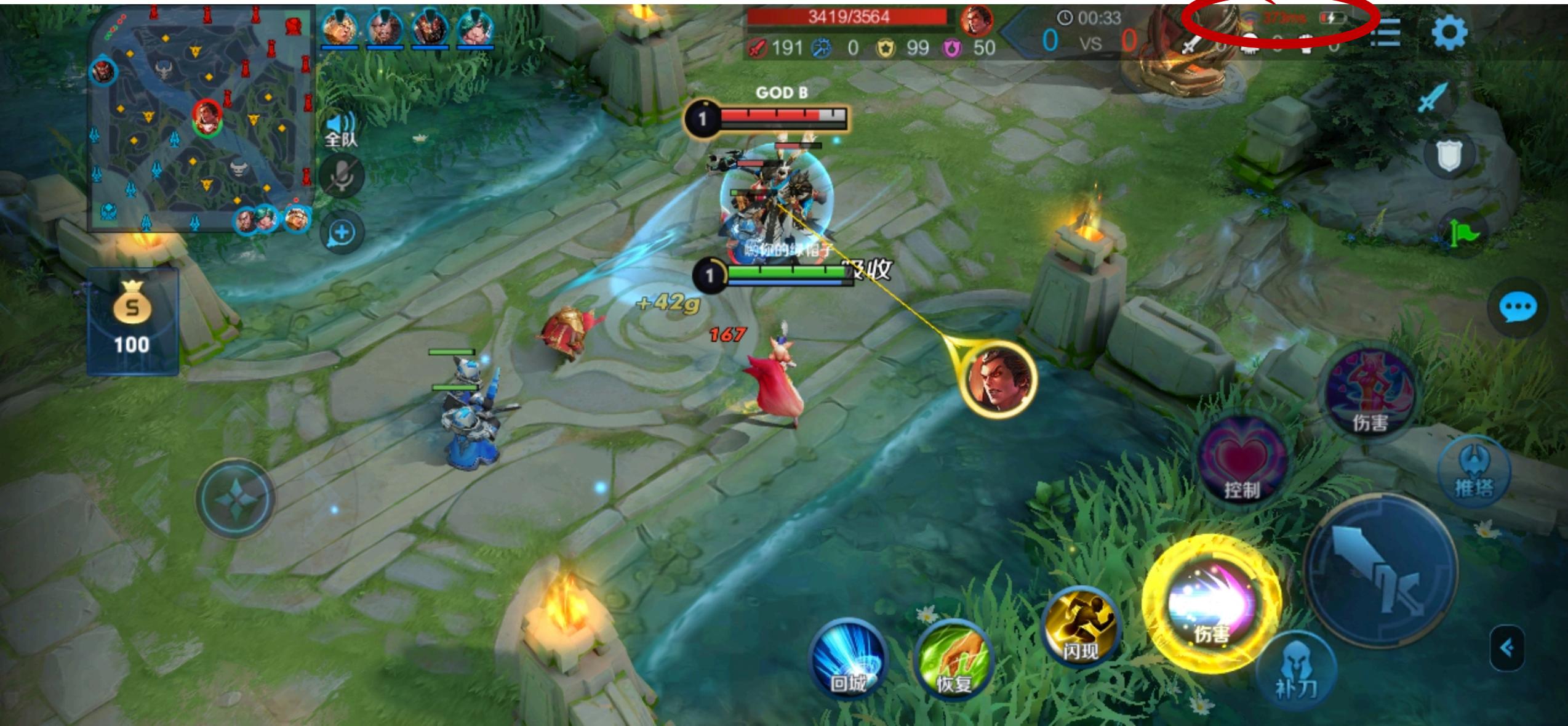


互联网传输优化做什么？

构建高吞吐、低时延的数据传输网络通道，
使能极致用户体验

时延直接影响用户体验

端到端时延
373 ms



未来：基于视频的应用需要高吞吐



极高清(UHD) 无线投屏



VR 交互游戏



AR 交互游戏



Application Requirements	UHD Cameras (Security)	UHD Streaming	VR Streaming	Self Driving Vehicle Diagnostics	Cloud Gaming	UHD IP Video	8K Wall TV	HD VR	UHD VR
Average (Mbps)	16	16	17	20	30	51	100	167	500

Source: Cisco Annual Internet Report, 2018-2023

8K Videos from Youtube	Asteroid Discovery	NORWAY	The Las Vegas Strip	TOKYO HDR Time Lapse	Angel Falls, Venezuela	Tigers Go For A Swim	Henosis(8K Short Film)
Average (Mbps)	81.0	81.1	129.1	62.8	57.7	72.0	40.6
Peek (Mbps)	206.9	196.9	134.2	83.2	81.3	72.1	71.1

Source: Huawei iLab, 2017

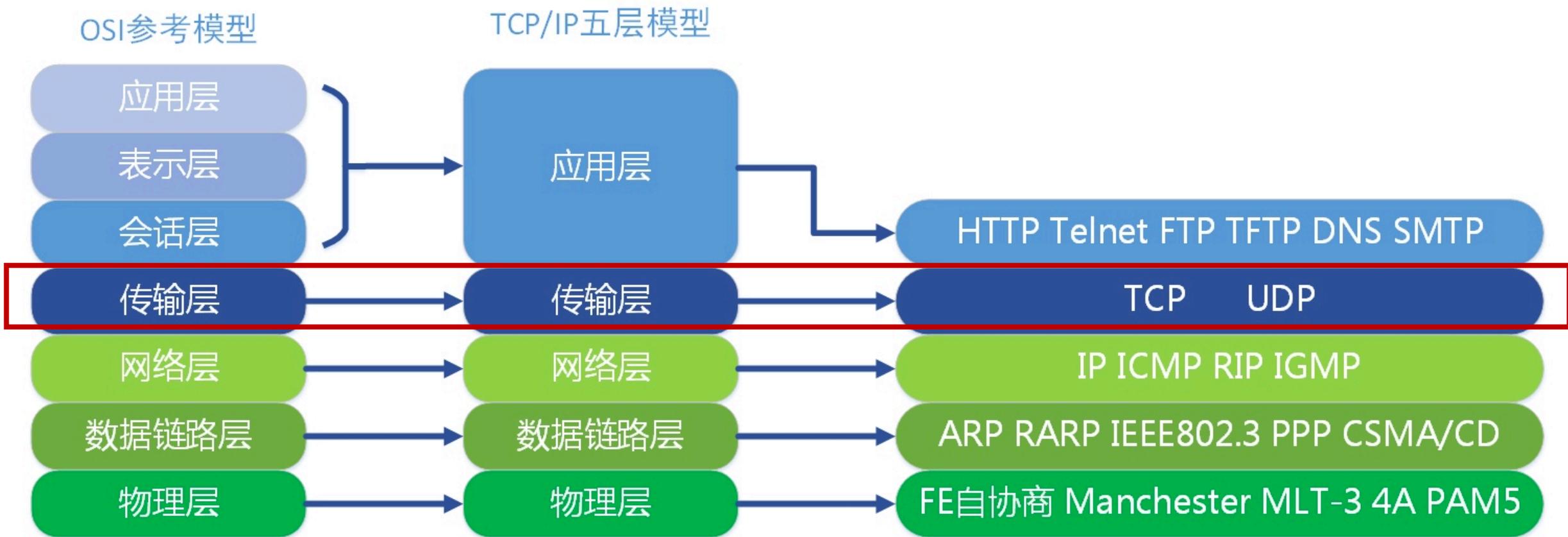
未来：元宇宙八大要素



互联网传输技术创新：
6G、无线局域网、广域网、数据中心网络、算力网络……

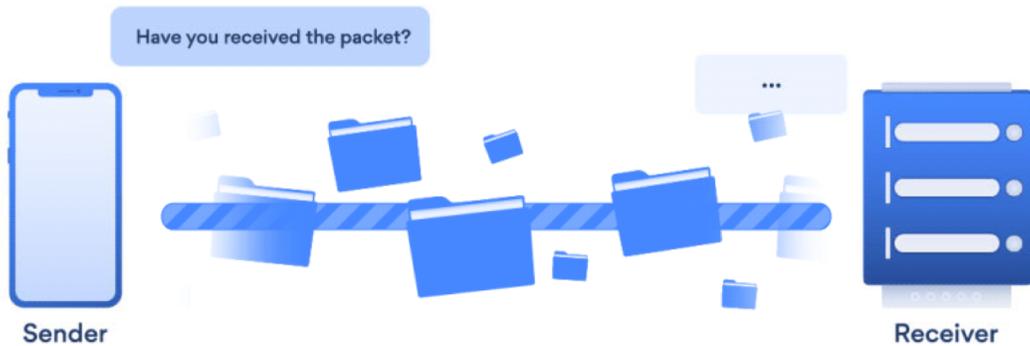
一些背景和网络的基础知识

互联网协议栈



传输层协议：UDP/TCP

How UDP works

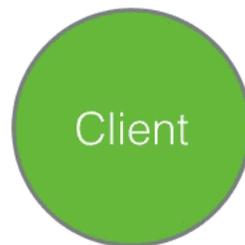


UDP: User Datagram Protocol
TCP: Transmission Control Protocol

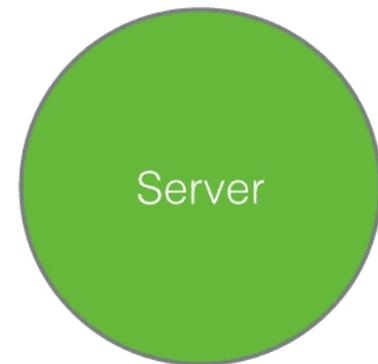
How TCP works



TCP需要接收端确认



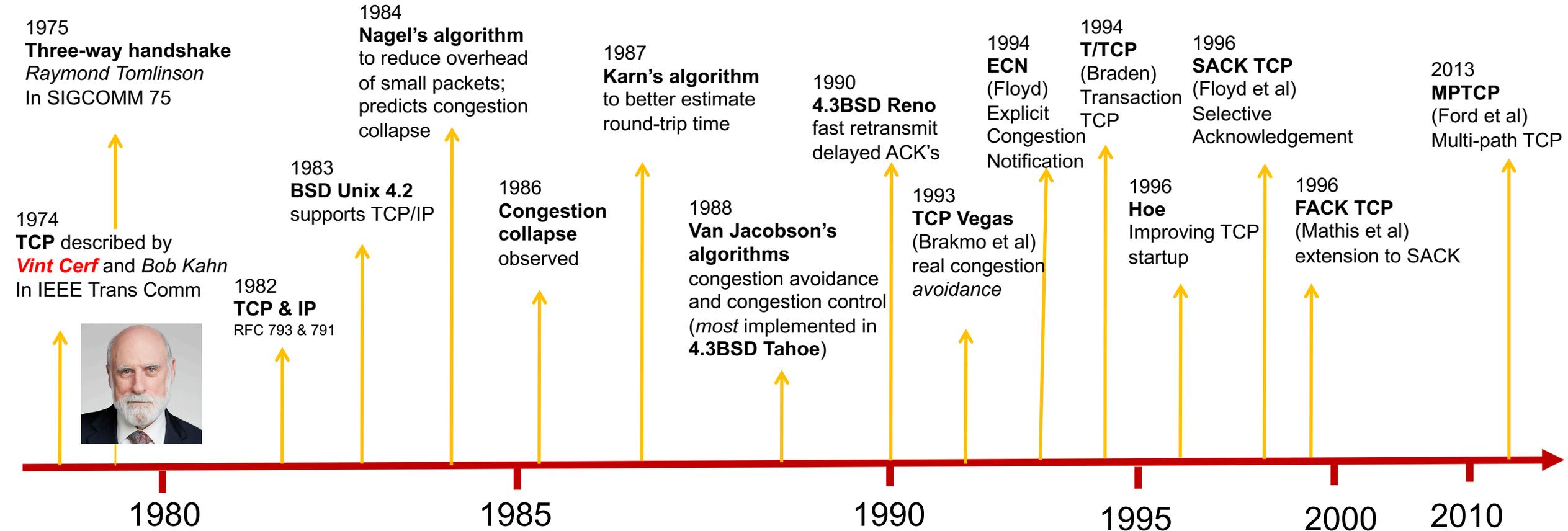
Client



Server

互联网上，超过90%的流量都是TCP!

TCP协议演进历程



追星少年



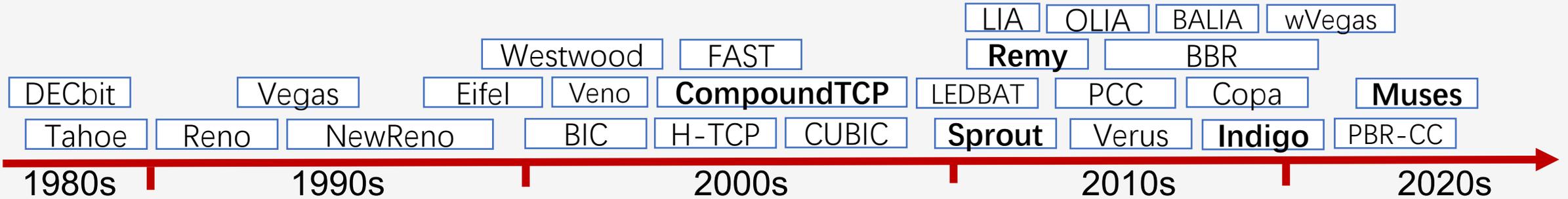
Latency Optimization on Computation Offloading of Mobile Terminals

Dissertation Submitted to
Tsinghua University
in partial fulfillment of the requirement
for the degree of
Doctor of Philosophy
in
Computer Science and Technology
by
Tong Li

Dissertation Supervisor : Professor Ke Xu

April, 2017

TCP关键算法演进历程

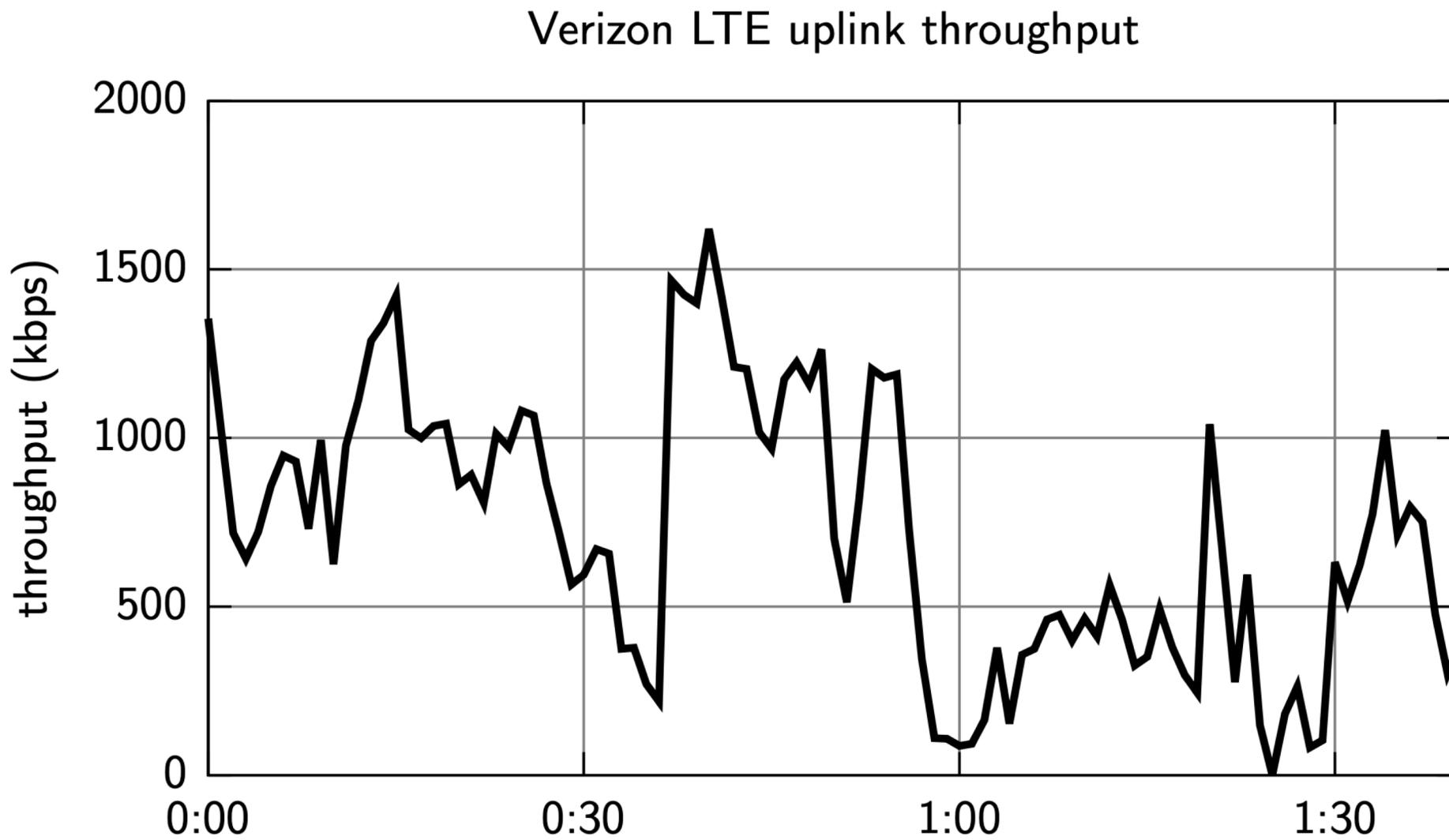


拥塞控制算法 (Congestion Control) 演进



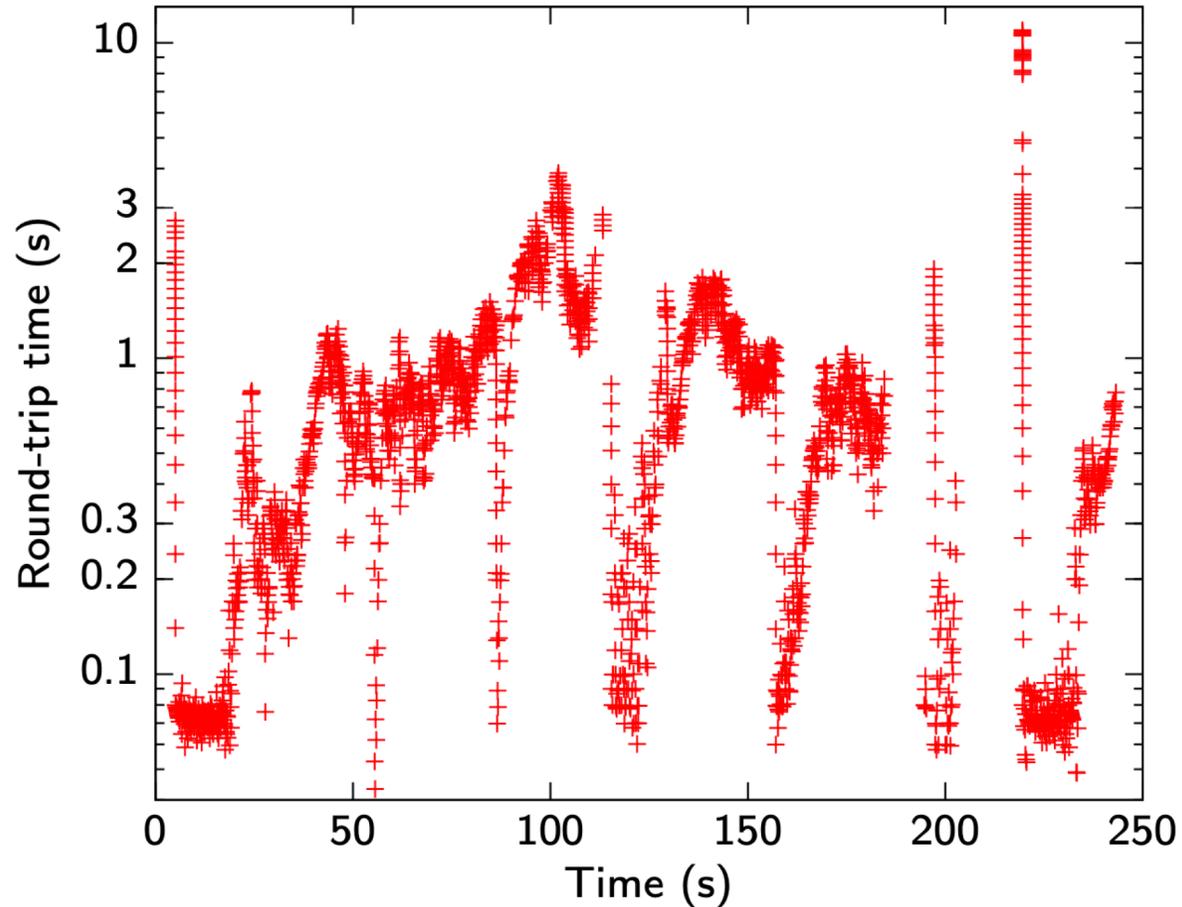
确认机制 (Acknowledgment Mechanism) 演进

TCP在无线网络中，可用带宽抖动剧烈



TCP在无线网络中，时延抖动也很剧烈

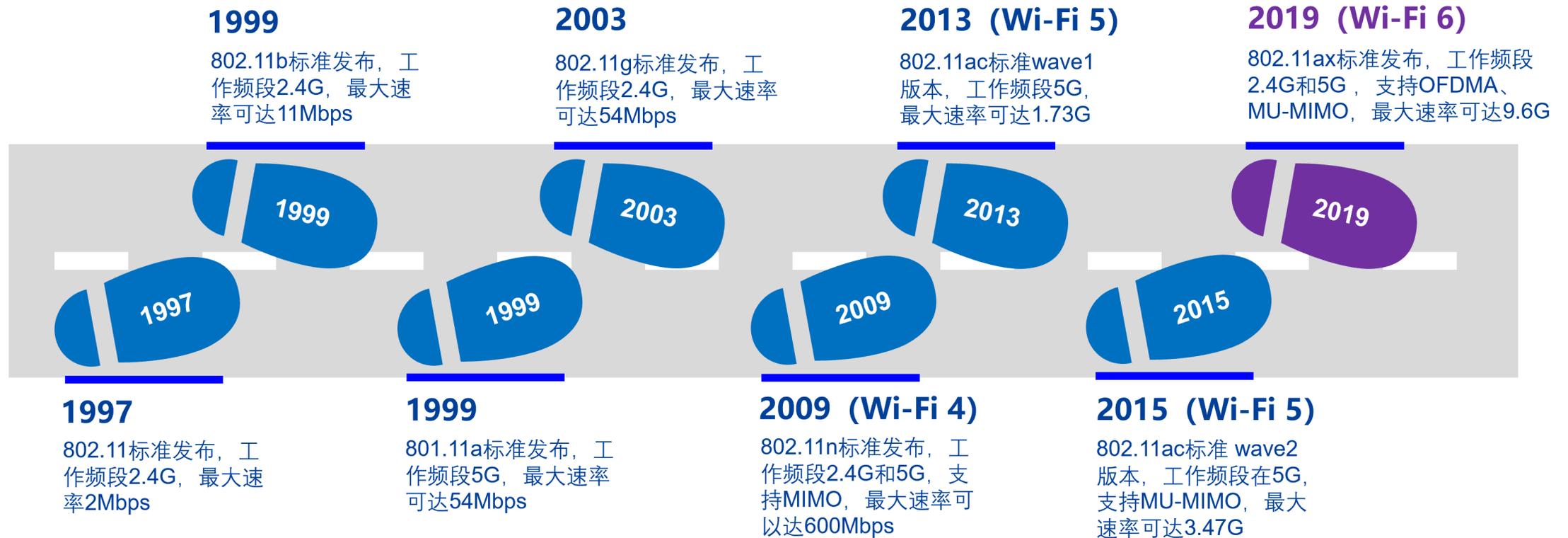
无线网络信道“太不可靠”导致可用带宽动态变化，其链路层却“太可靠”



(Verizon LTE, one TCP download.)

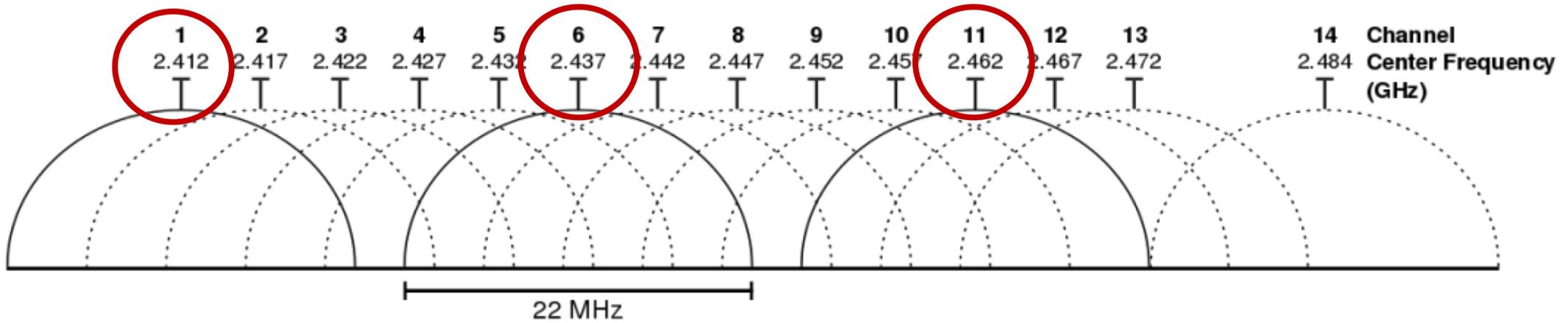
Wi-Fi是最常见的无线网络技术

■ 基于IEEE 802.11标准的Wi-Fi技术发展历程



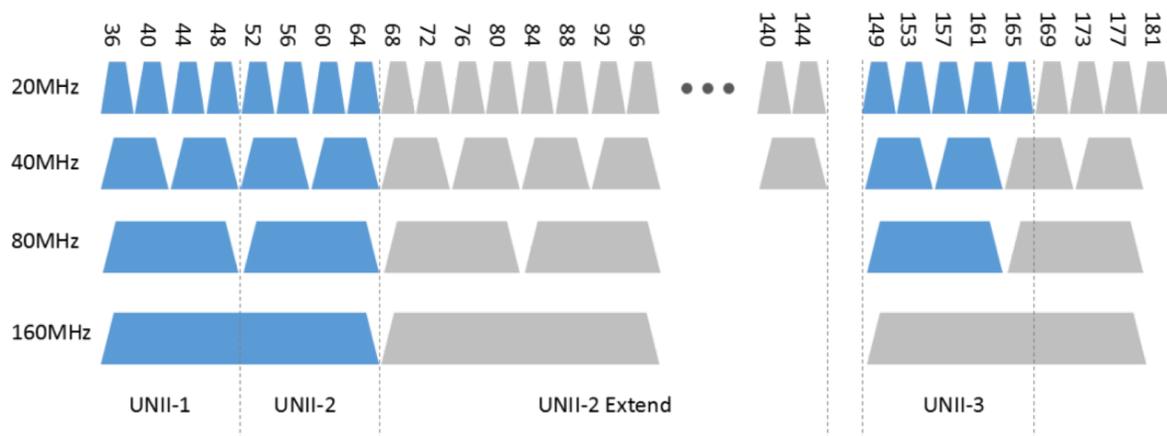
Wi-Fi工作频段

IEEE Wi-Fi 2.4GHz信道分布图



IEEE WiFi 5GHz信道分布图

蓝色的信道在中国可用



“此5G” 非“ 彼5G”



1G



2G



3G



4G



5G



2.4 Kbps



64 Kbps



2,000 Kbps



100,000 Kbps



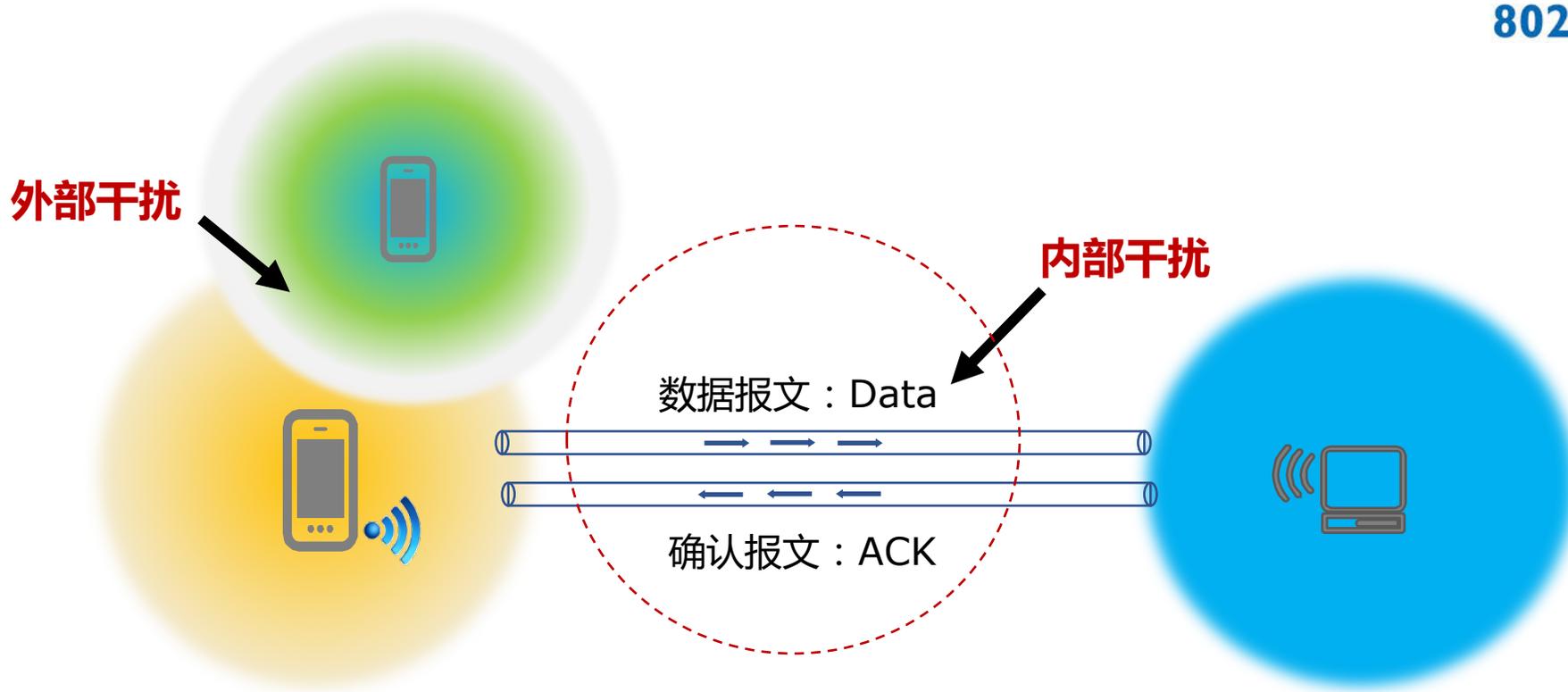
10,000,000 Kbps

Wi-Fi数据传输：同时受到“外部干扰”和“内部干扰”

外部干扰：外部因素引起的干扰

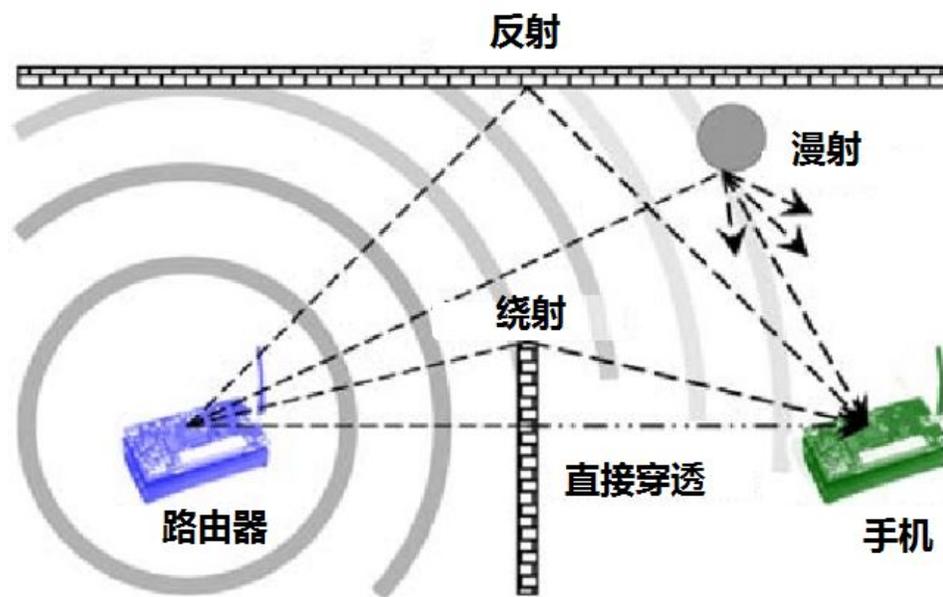
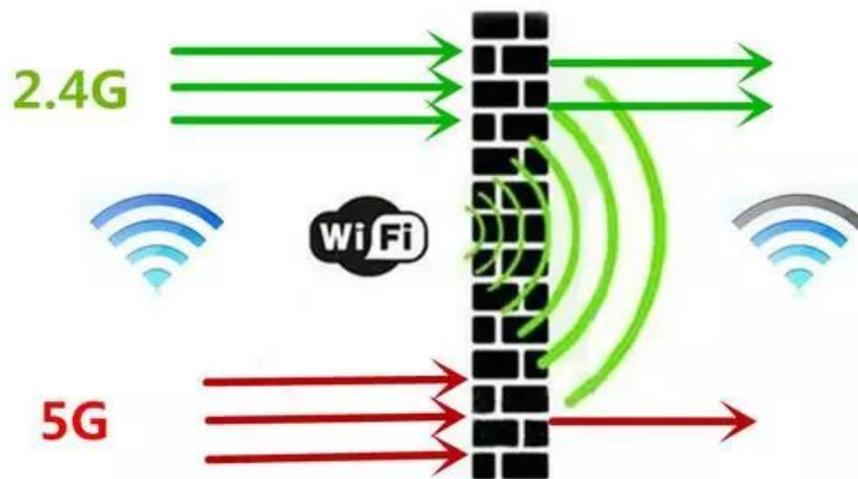
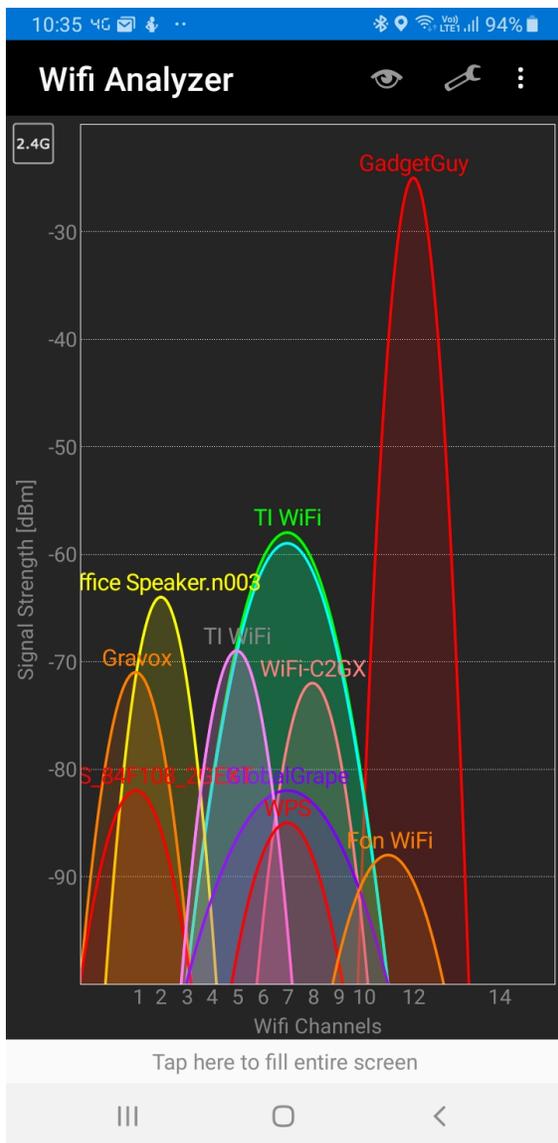
内部干扰：存在于同一个连接中的数据报文和确认报文之间

ACK: 确认报文 (Acknowledgment)

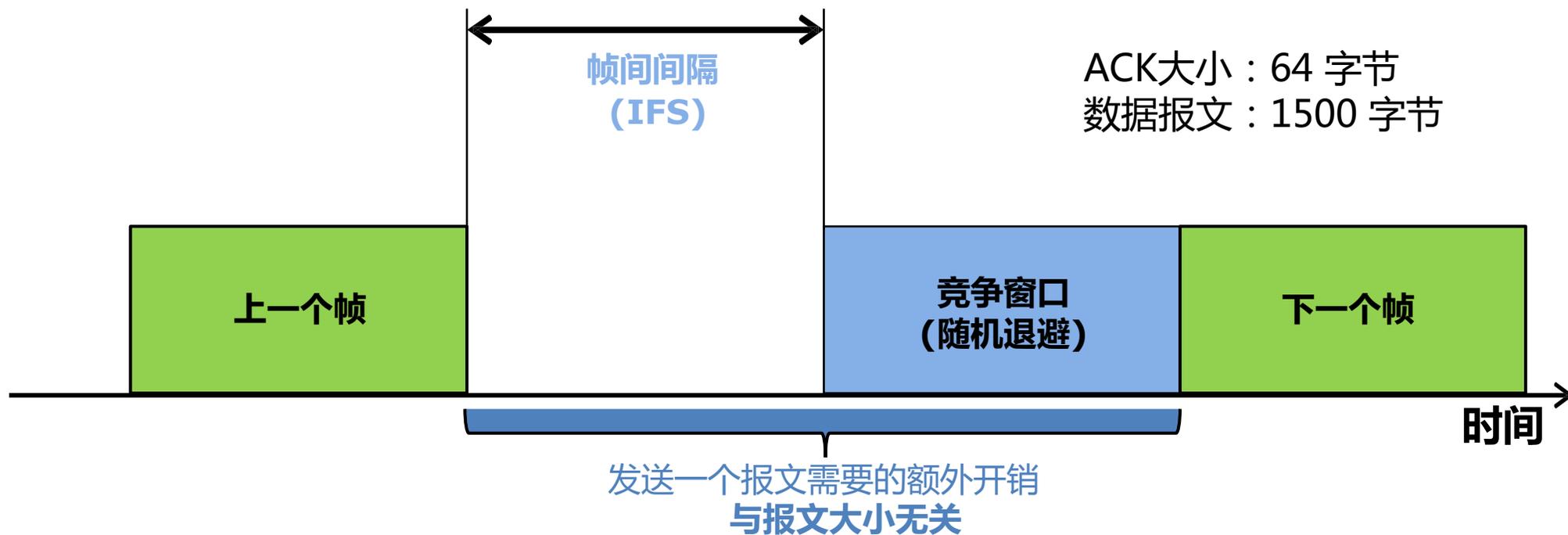


ACK: 确认报文 (Acknowledgment)

外部干扰：同频/邻频干扰、障碍物

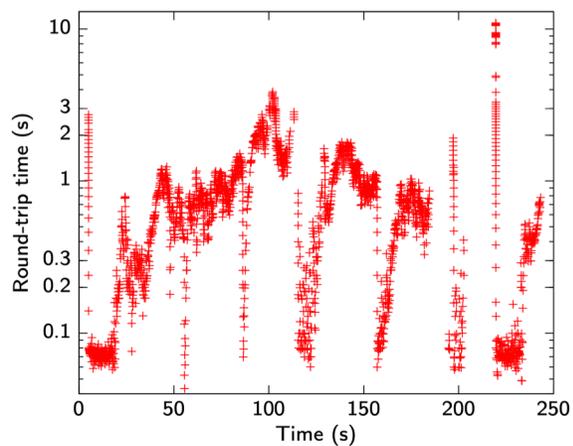
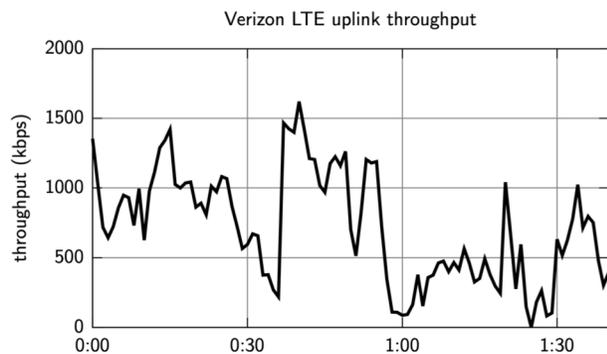


内部干扰：ACK 消耗与数据报文几乎相等的频谱资源



基于 IEEE 802.11 medium access control (MAC) 协议
频谱资源消耗示意图

外部干扰和内部干扰，导致无线网络性能低下

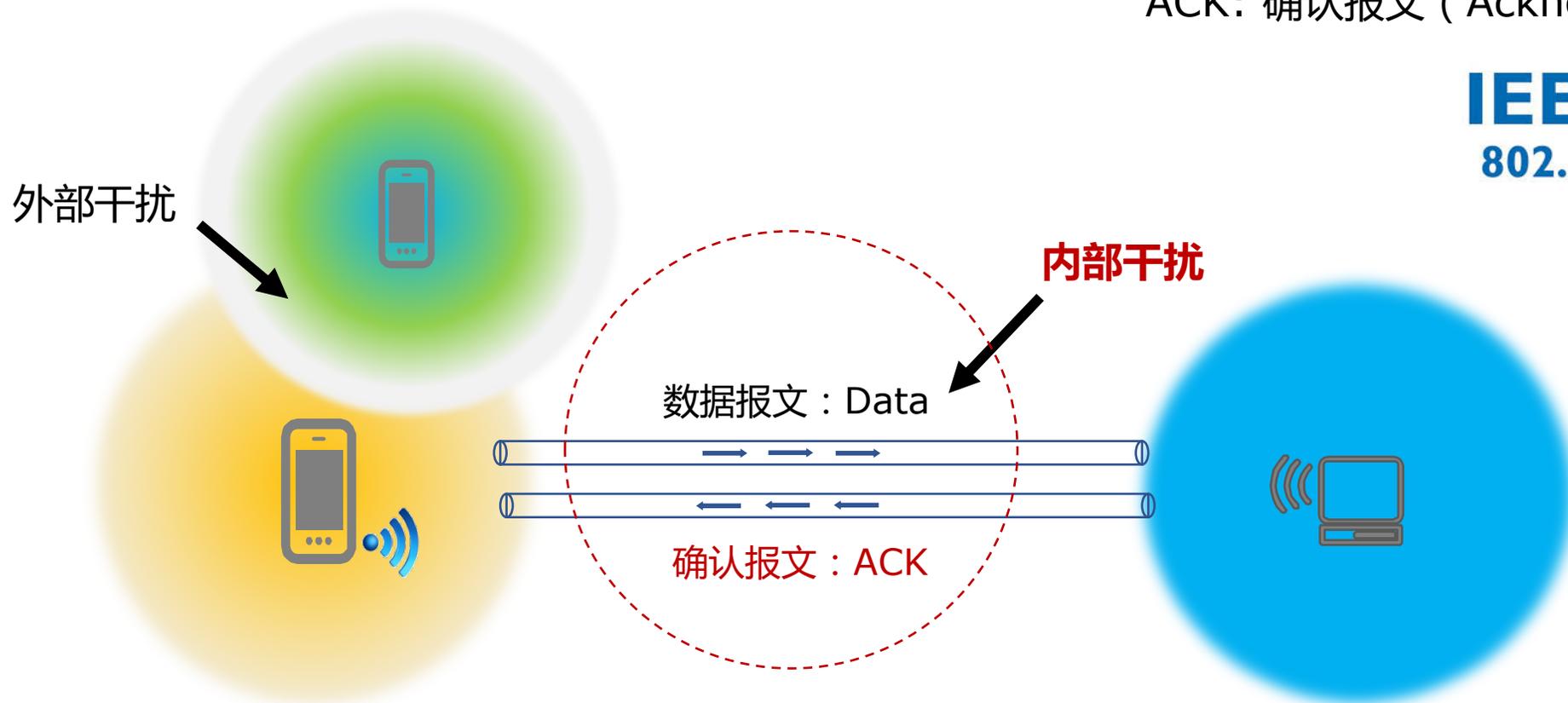


(Verizon LTE, one TCP download.)

我们的创新

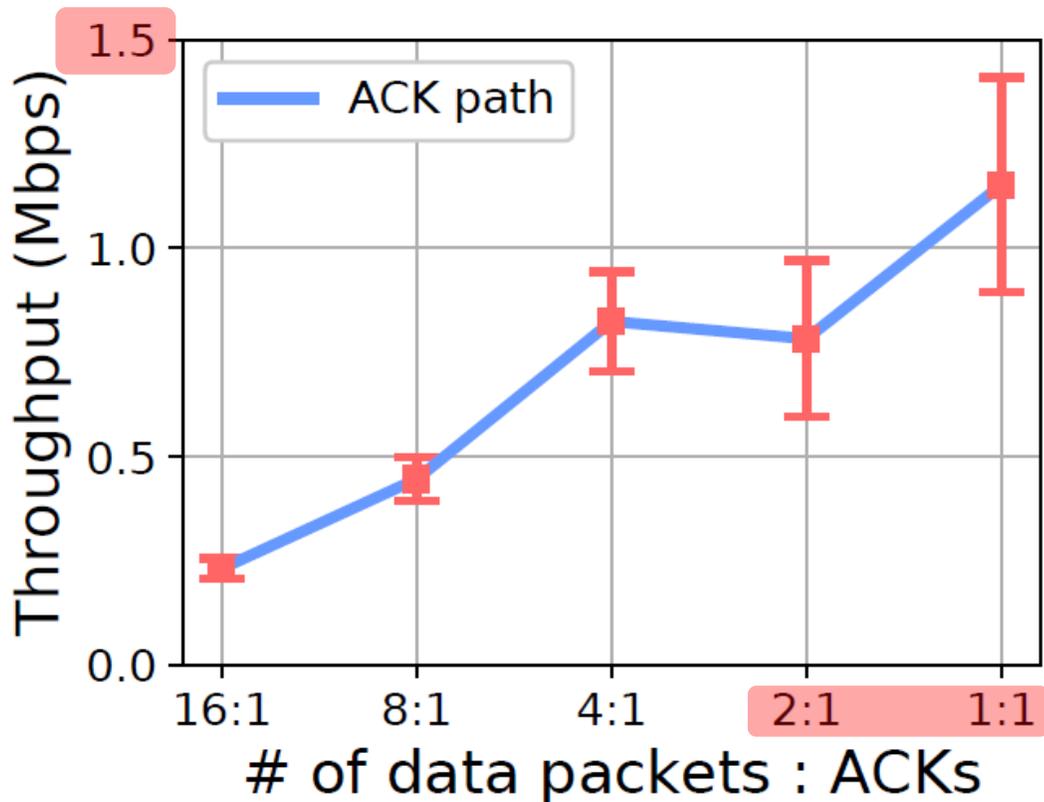
TCP ACK 引入了“内部干扰”

ACK: 确认报文 (Acknowledgment)

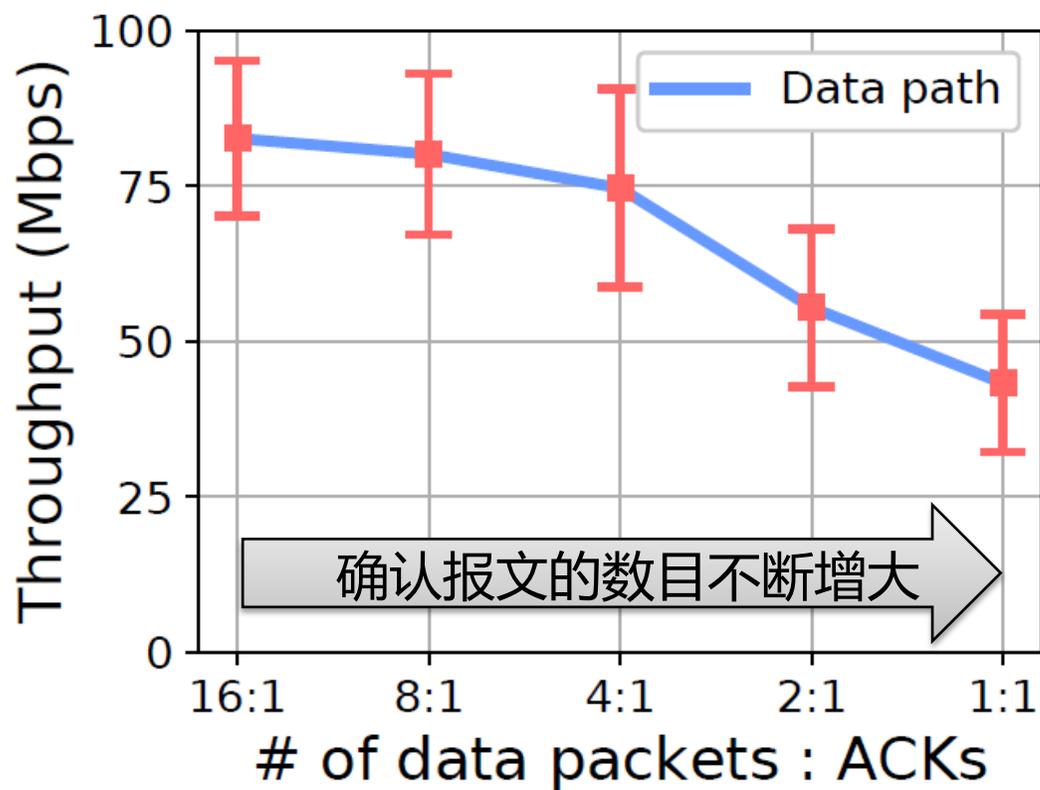


外部干扰：外部因素引起的干扰
内部干扰：存在于同一个连接中的数据报文和确认报文之间

减少ACK的数量可能提升可用带宽，从而可能实现低时延



(a) 确认报文的吞吐量



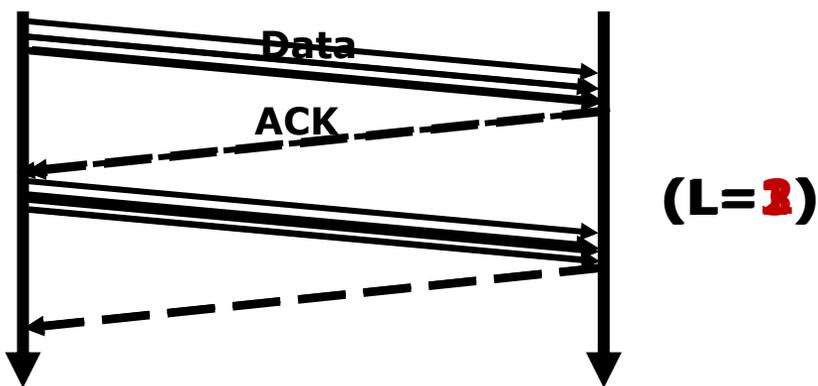
(b) 数据报文的吞吐量

IEEE 802.11n 无线链路上的内部干扰实现

两种减少ACK频率的方式

Byte-counting ACK

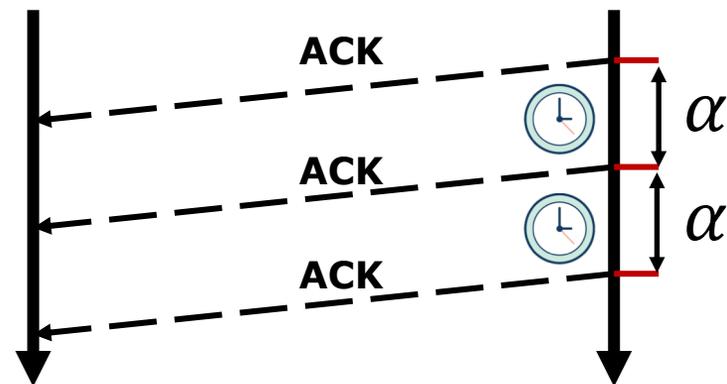
$$f = \frac{bw}{L \cdot MSS}$$



- **f**: ACK频率，单位为Hz, i.e., 每秒发送的ACK数目
- **L**: 发送一个ACK之前收到的数据报文的数目

Periodic ACK

$$f = \frac{1}{\alpha}$$



- **MSS**: 最大报文长度
- **bw**: 带宽
- α : ACK发送周期

前方高能

传统TCP的确认机制无法最小化ACK频率

ACK频率建模：

$$f_{tcp_delayed} = \max \left\{ \frac{bw}{L \cdot MSS}, \frac{1}{\alpha} \right\} = \max \left\{ \frac{bw}{2 \cdot MSS}, \frac{1}{500} \right\}$$

RFC 1122 & RFC 5681

$$\text{if } bw \geq \frac{L \cdot MSS}{\alpha}$$

(Byte-counting ACK)

- **f**: ACK频率，单位为Hz, i.e., 每秒发送的ACK数目
- **L**: 发送一个ACK之前收到的数据报文的数目

$$\text{if } bw < \frac{L \cdot MSS}{\alpha}$$

(Periodic ACK)

- **MSS**: 最大报文长度
- **bw**: 带宽
- **α** : ACK发送周期

传统 TCP 的确认机制并不是最优的

我们的方案：Tame ACK (TACK)

$$f_{tack} = \min \left\{ \frac{bw}{L \cdot MSS}, \frac{1}{\alpha} \right\} = \min \left\{ \frac{bw_{max}}{L \cdot MSS}, \frac{\beta}{RTT_{min}} \right\}$$

if bdp < L · β · MSS

(Byte-counting ACK)

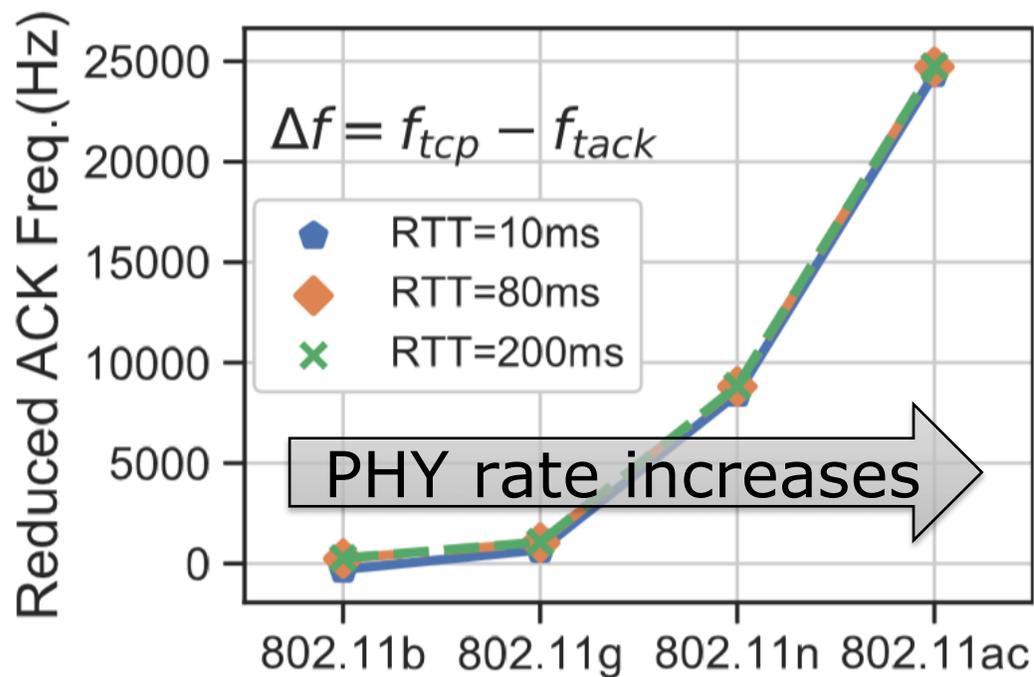
if bdp ≥ L · β · MSS

(Periodic ACK)

- bw_{max} : 评估的最大带宽
- RTT_{min} : 评估的最小往返时延

- β : 每个 RTT_{min} 时间内发送的ACK数目
- bdp : 带宽时延积 ($bw_{max} \times RTT_{min}$)

TACK 机制可以显著减少ACK的数目



(a) ACK frequency reduction

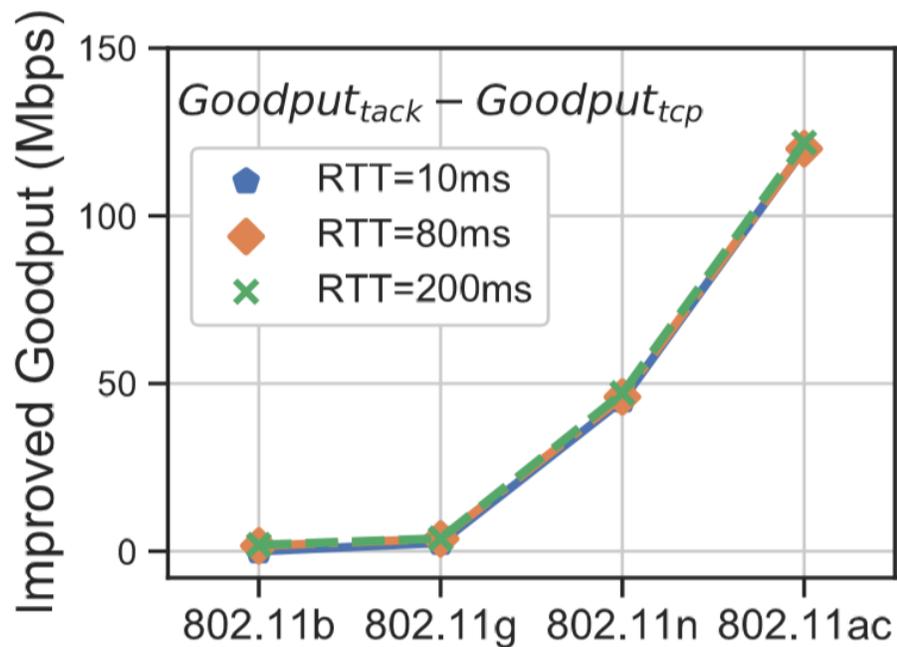
	802.11b		802.11ac	
RTT_{min}	TCP (L=2)	TACK (L=2)	TCP (L=2)	TACK (L=2)
10 ms	294	294	24777	400
80 ms	294	50	24777	50
200 ms	294	20	24777	20

(b) ACK frequency (Hz)

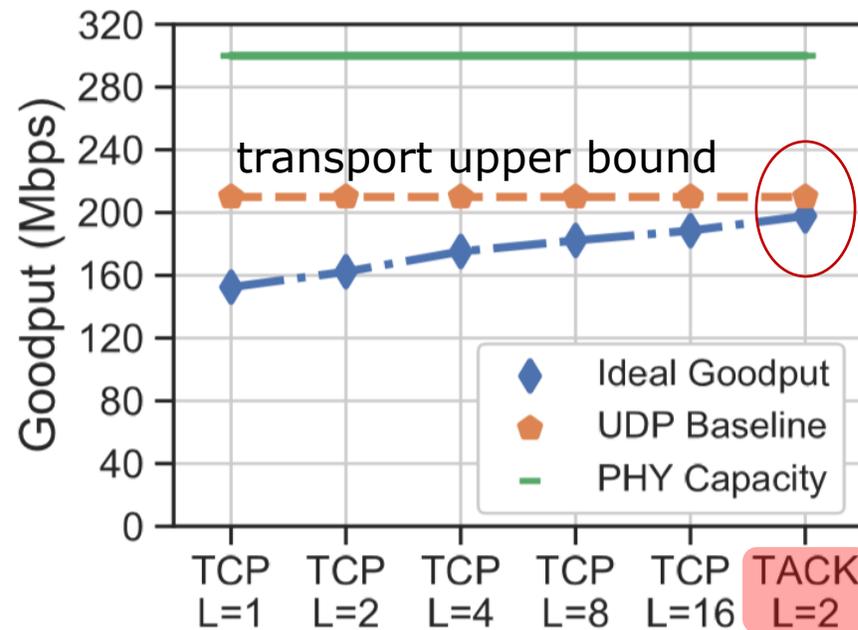
IEEE 802.11 b/g/n/ac 无线链路上的ACK频率定量分析

创新 ≠ 实用的技术

理想情况下，TACK机制可以使用最少的ACK数目，实现最优性能



(a) Goodput improvement

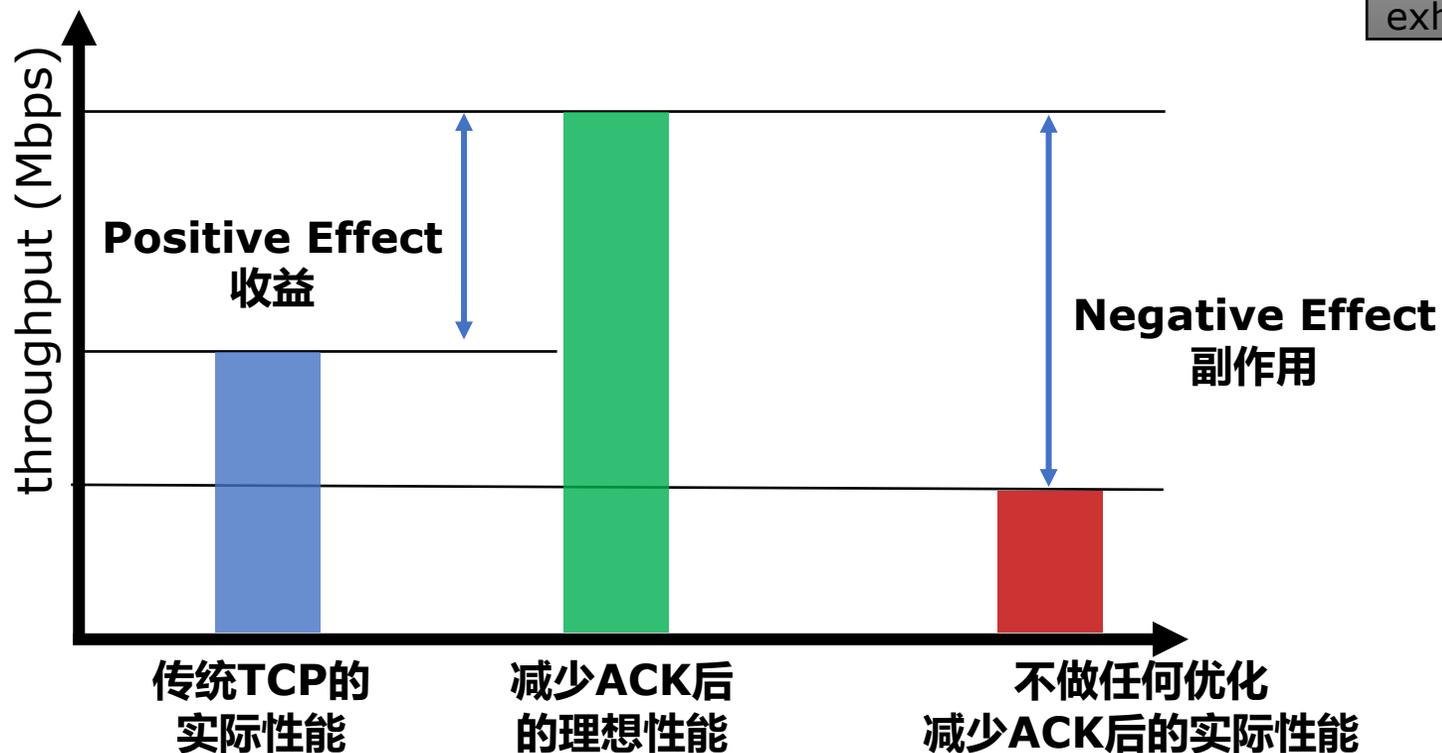


(b) Ideal goodput trend (RTT=80 ms, IEEE 802.11n)

传输上界: 传输控制不会因为减少ACK数目而受到影响

实际情况下，减少ACK的数目，损害TCP的性能

减少ACK的数目不仅有“收益”，还有“副作用”

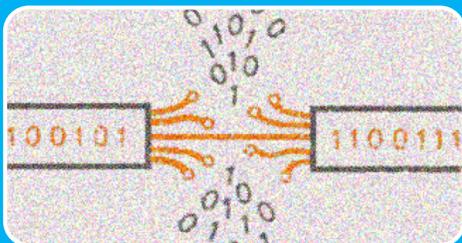


"A TCP receiver which does not generate an ACK for every second full-sized segment exhibits a "Stretch ACK Violation".

--- 1999, RFC 2525, Vern Paxson et.al

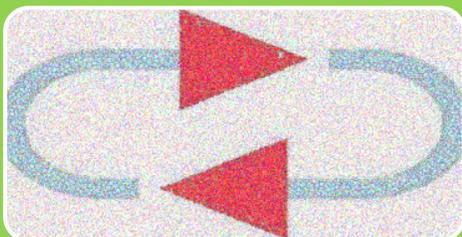
当前的TCP的设计
依赖于频繁地发送
ACK

避免 TACK 的“副作用”面临的三大挑战



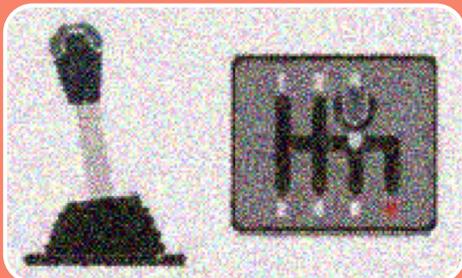
丢包恢复问题

- 减少 ACK 的数目会增大对丢包事件的反馈延迟。ACK 丢失后问题更严重。



时延探测问题

- 在多个数据报文中，仅采样一个 RTT 样本，可能引入测量偏差。



速率控制问题

- 拥塞控制：减少 ACK 的数目将产生突发流量。
- 流量控制：延迟反馈接收缓存，可能导致带宽利用率低下。

基于TACK的协议确认机制

■更多种类的ACK

- 除了TACK, 同时引入一种即时ACK, 称为IACK (“Instant ACK”), 用于响应即时事件
- IACK 和 TACK 是互补的, IACK保证及时性, TACK保证反馈的鲁棒性

■ACK携带更多的必要信息

- 当ACK路径上的丢包率超过阈值时, TACK携带更多的历史反馈过的丢包信息
- 在收发双方之间, 同步ACK延迟、时间戳、丢包率、速率等信息

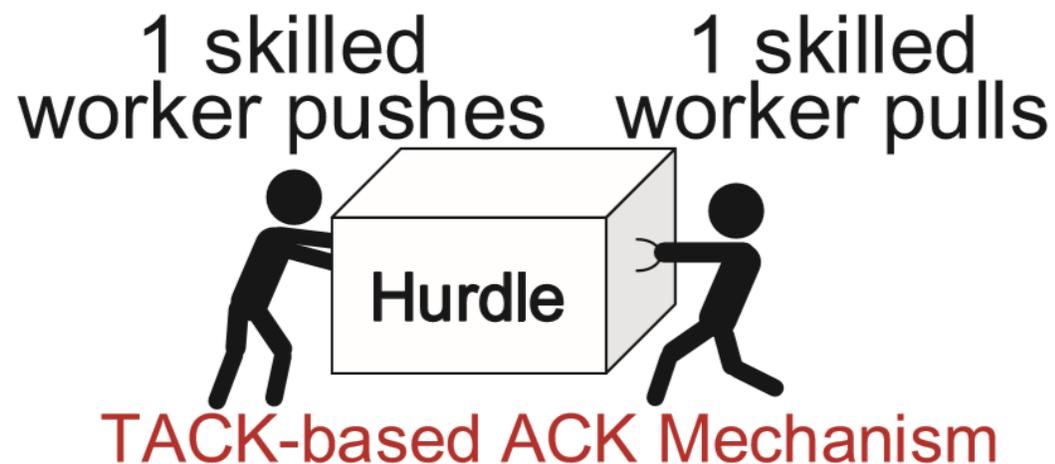
■更少的ACK数目, 但正好是传输需要的

- 新定义的 IACK 只响应即时事件, 所以额外增加的ACK数目通常很低, 可忽略不计

基于TACK的协议确认机制



=



设计原理：类比

基于TACK的协议设计

■ 丢包恢复的改进

- Packet number, TACK+IACK

■ 时延探测的改进

- One-way delay as auxiliary

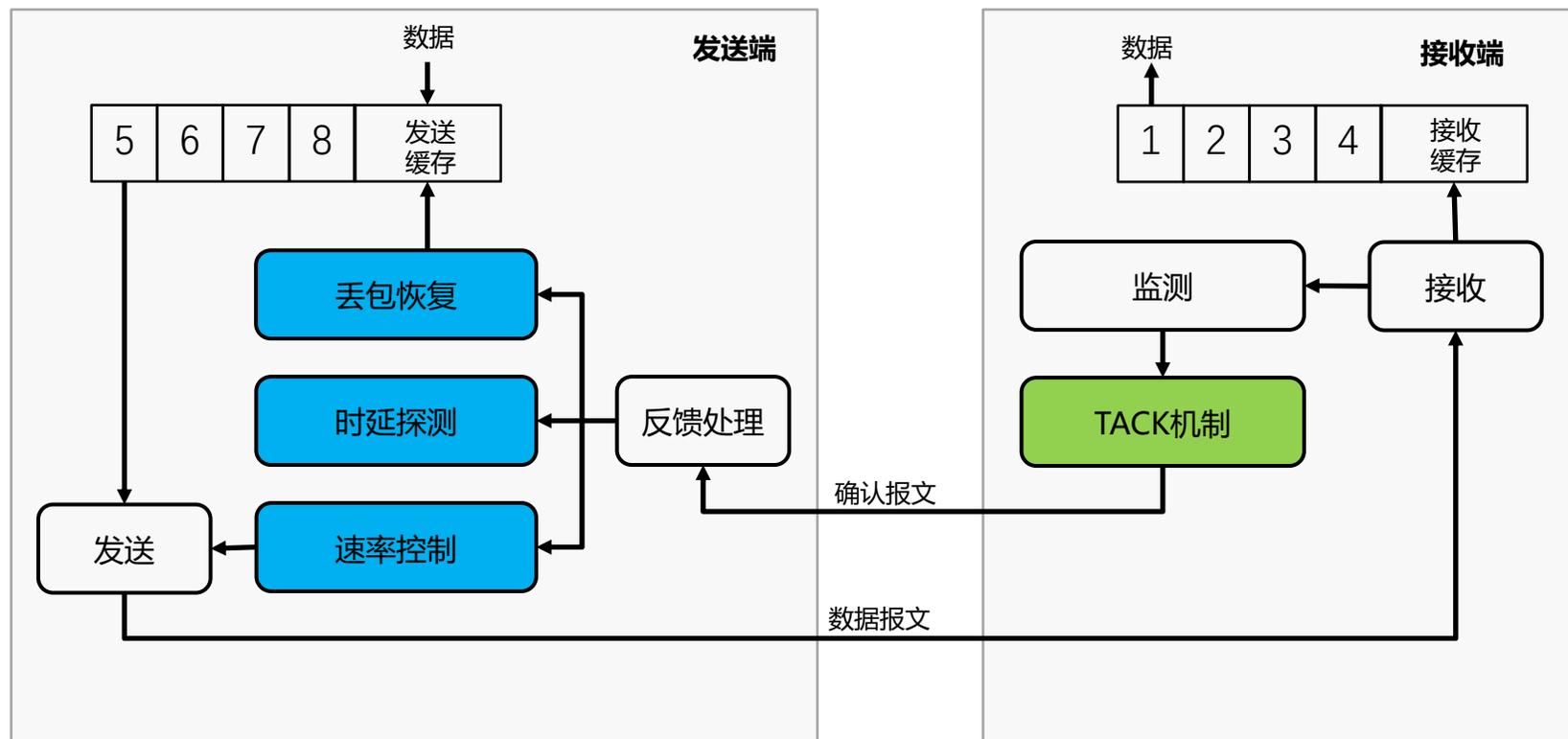
■ 速率控制的改进

- Congestion control: Pacing
- Flow control: IACK

目标: 减少传输控制对频繁的ACK的依赖性

原理: 减少ACK数目后, 接收端比发送端拥有的信息更完全

理念: 发送端驱动 → 接收端驱动



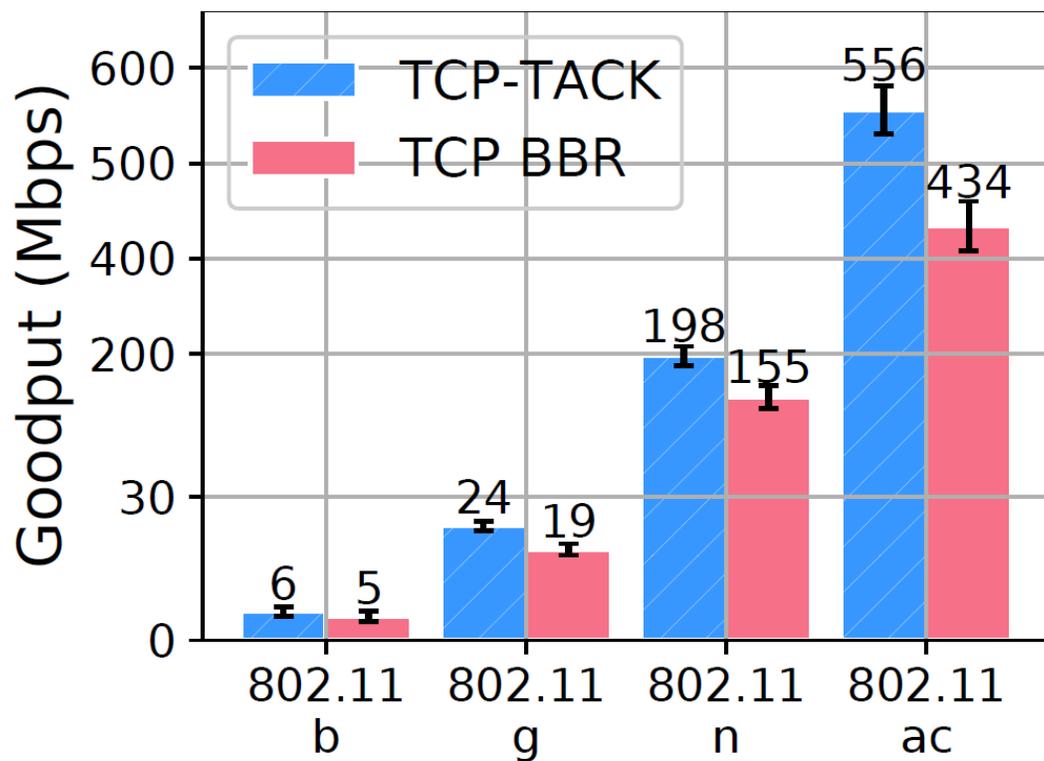
感兴趣可以阅读论文：

Tong Li et al. TACK: Improving Wireless Transport Performance by Taming Acknowledgments. ACM SIGCOMM, 2020.

中文博客：

<https://blog.csdn.net/sbmye/article/details/110175789>

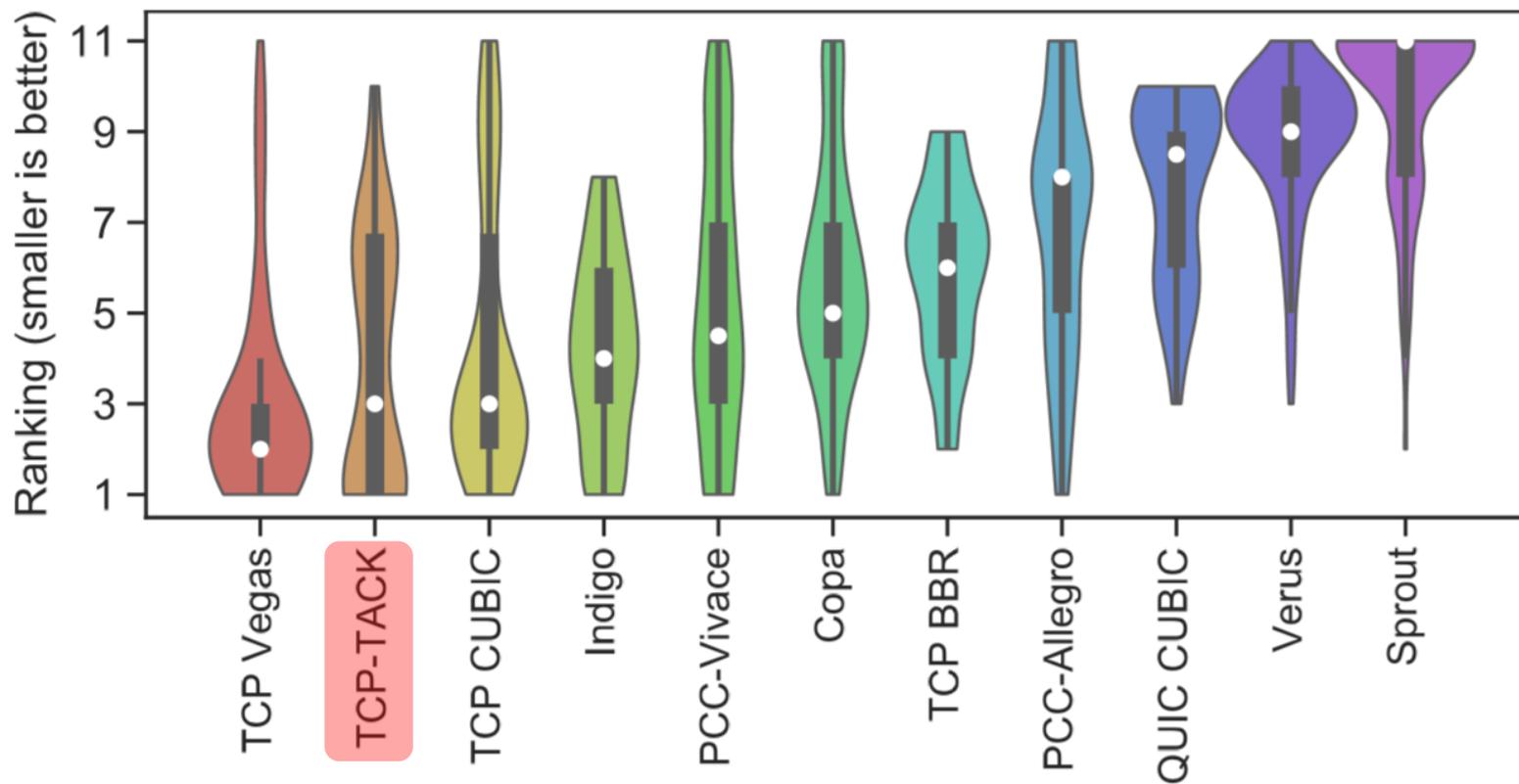
改进后的TACK在无线局域网中可以提升28%的可用带宽



	802.11b	802.11g	802.11n	802.11ac
Number of ACKs reduced	90.5%	95.4%	99.4%	99.8%
Goodput improved	20.0%	26.3%	27.7%	28.1%

在WLAN中，TCP-TACK 相比 TCP-BBR 提升的可用带宽比例

改进后的TACK在广域网中的性能表现良好



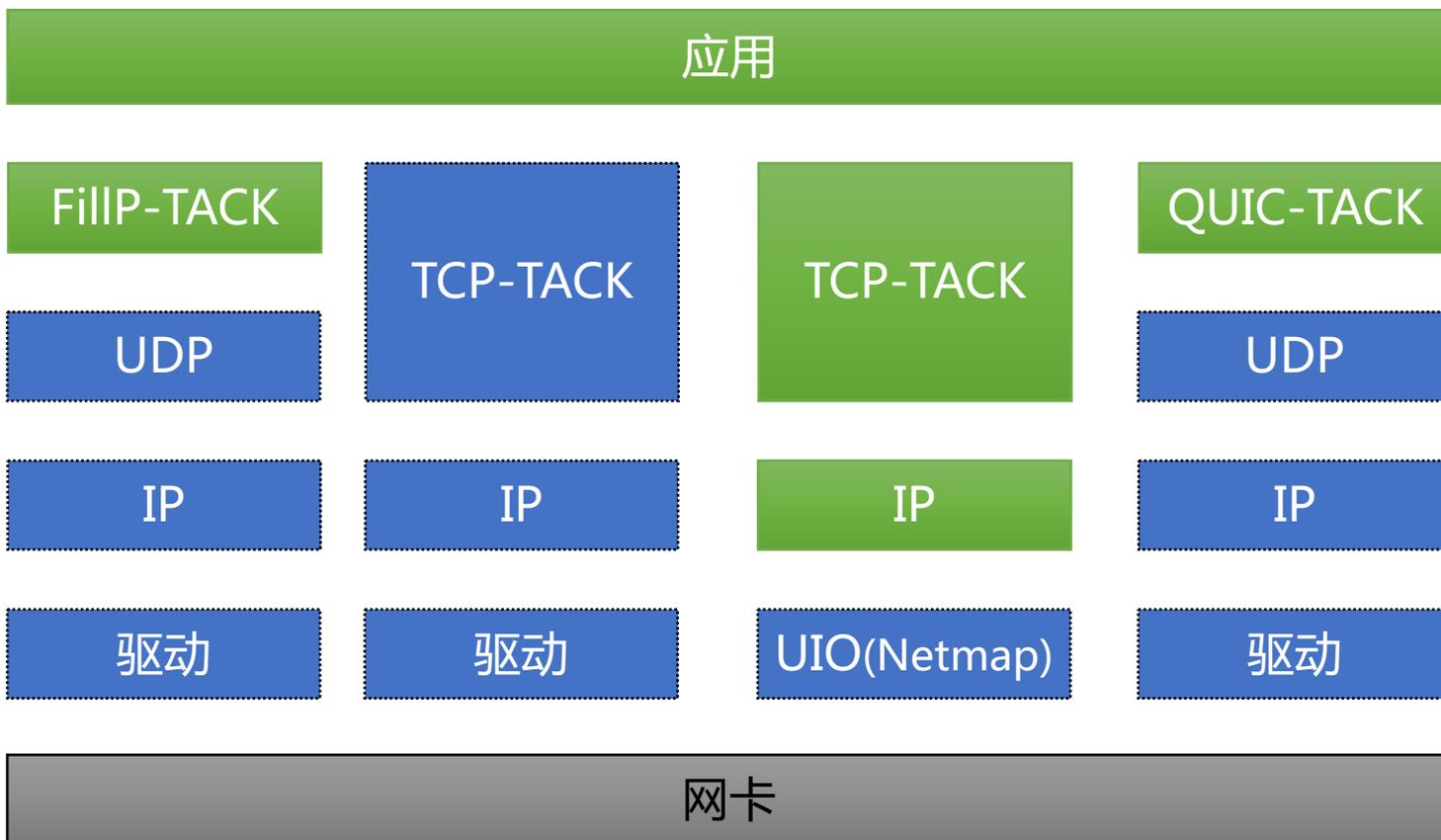
性能排名（200天的数据）

(数据来源：斯坦福大学协议测试比拼平台:<https://pantheon.stanford.edu/summary>)

创新变成了实用的技术

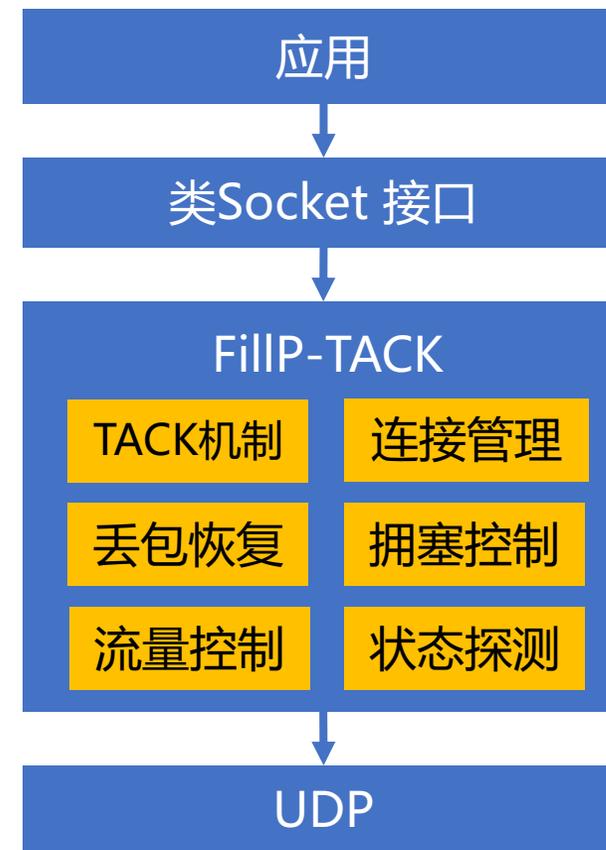
实用的技术 \neq 价值

工业实践：基于TACK的协议实现——FiIP-TACK



■ 内核态 ■ 用户态

QUIC : Quick UDP Internet Connection



FiIP实现示例

创新技术应用用于**华为、荣耀系列终端产品**，成为华为终端操作系统默认的投屏传输技术



超低时延 60帧投屏
游戏玩的爽 观战更爽

50-100ms 平均时延

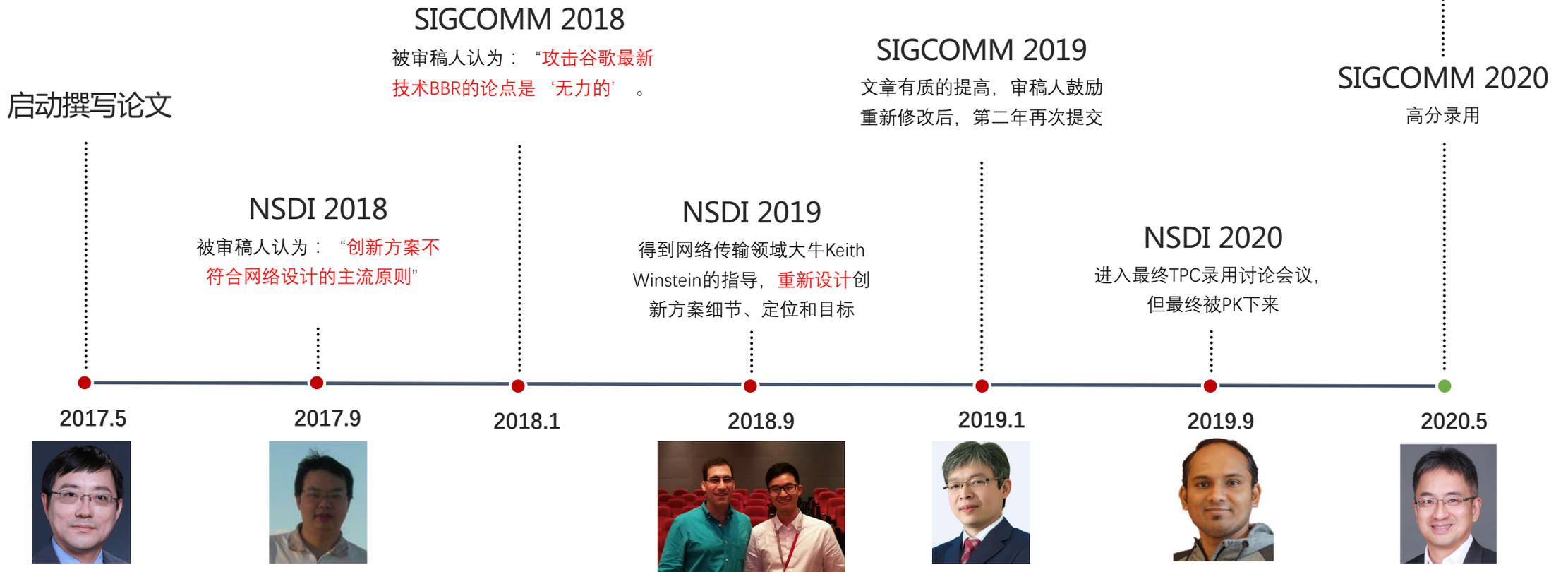
HONOR

创新实现了商业价值变现

复盘：从创新到落地

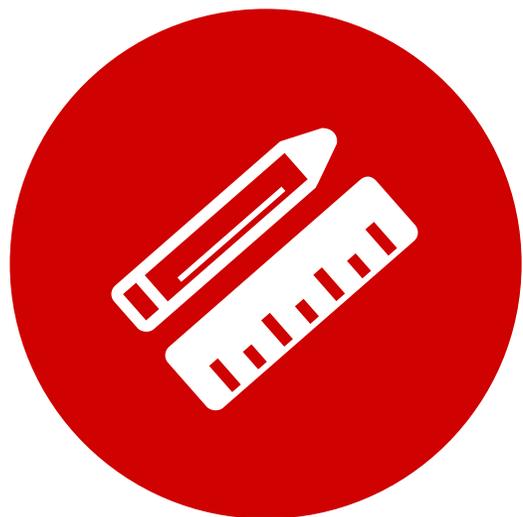


复盘：SIGCOMM之路





中國人民大學
RENMIN UNIVERSITY OF CHINA

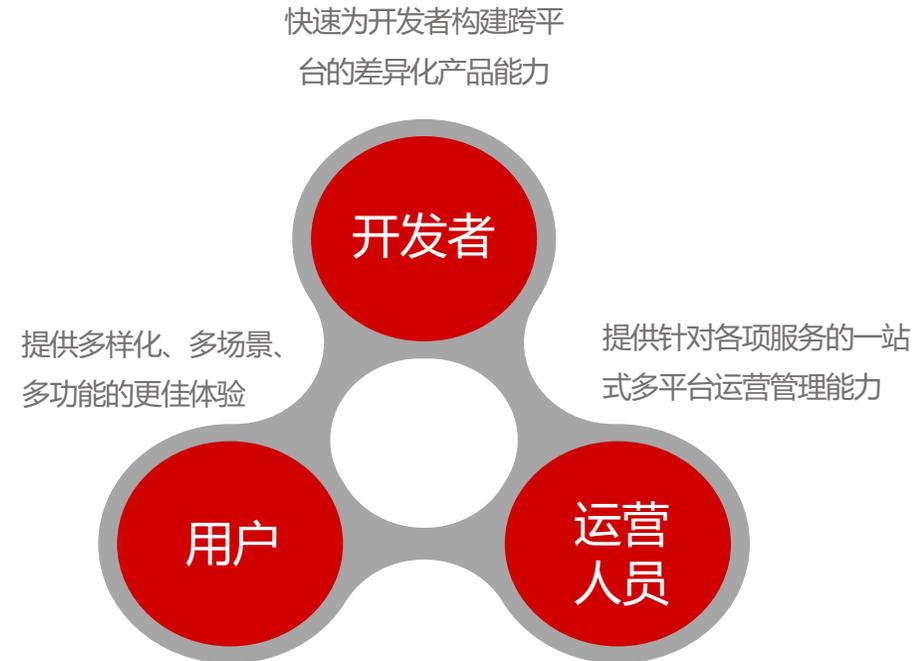
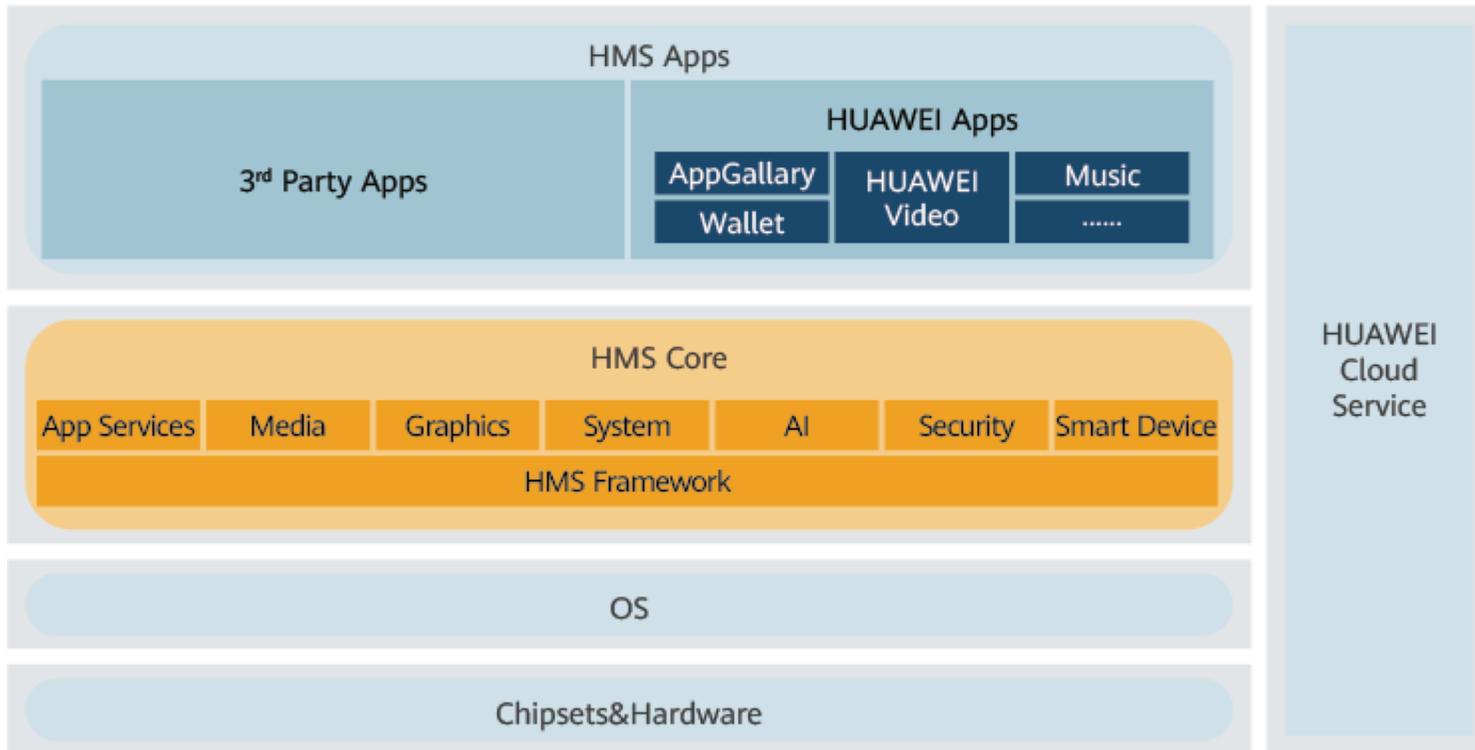


创新故事二

Contact Shield
新冠病毒接触卫士

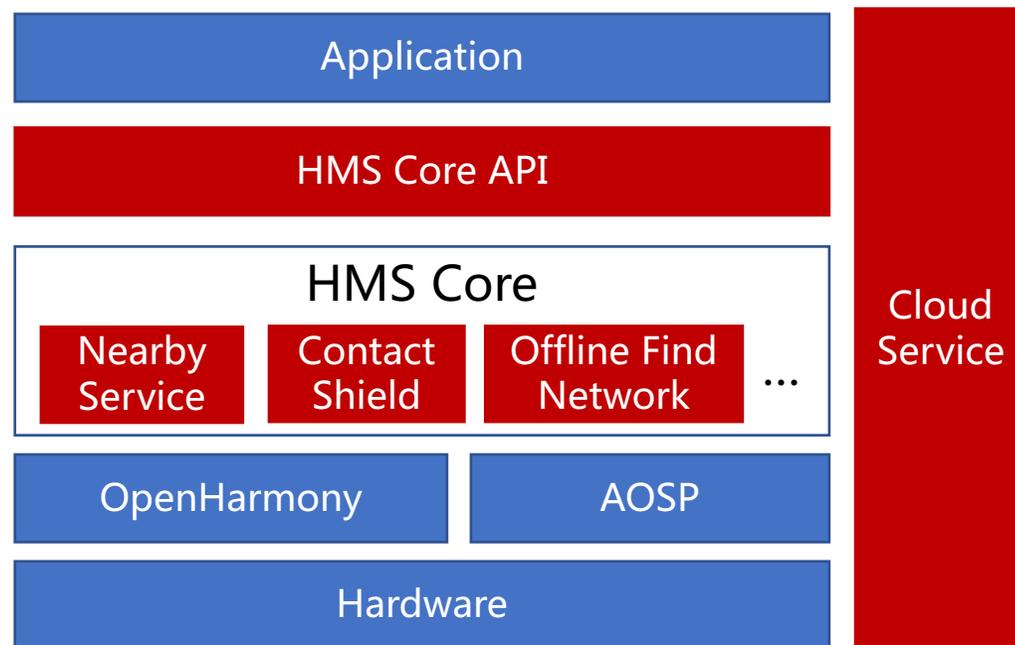
华为移动服务 (Huawei Mobile Servie, HMS)

- HMS Core是华为终端云服务 (HUAWEI Mobile Services) 开放能力合集，**位于开发者应用与操作系统之间**，是为应用开发提供基础服务的平台。同时，依托华为云服务，HMS Core也为这些服务提供云端能力，用于各服务的开通、业务实现及运营



通信相关的三大HMS服务

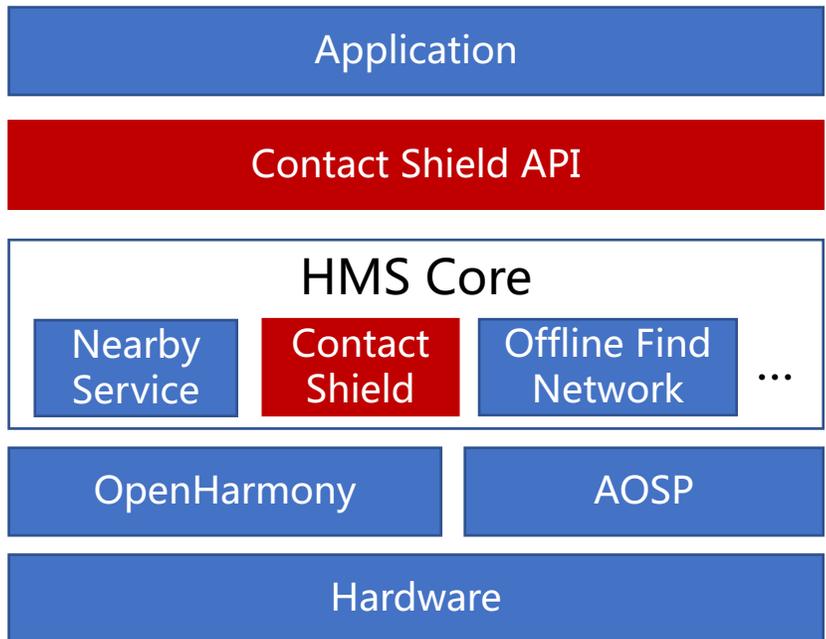
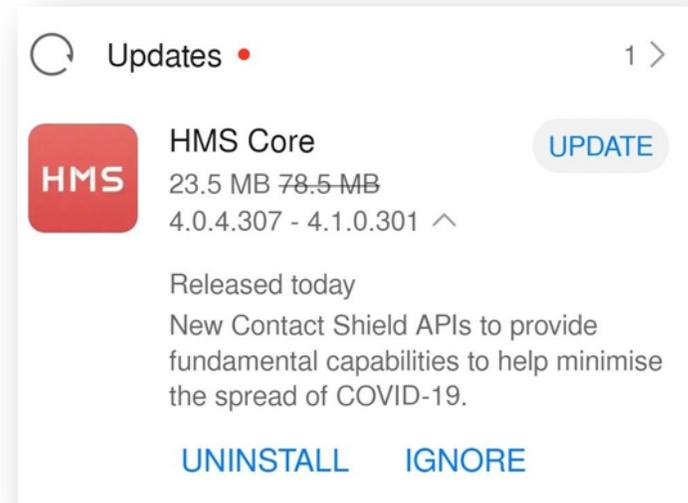
- Nearby Service: 近距离通信服务
- Contact Shield: 新冠病毒接触卫士
- Offline Find Network: 离线查找网络



以新冠病毒接触卫士 Contact Shield为例

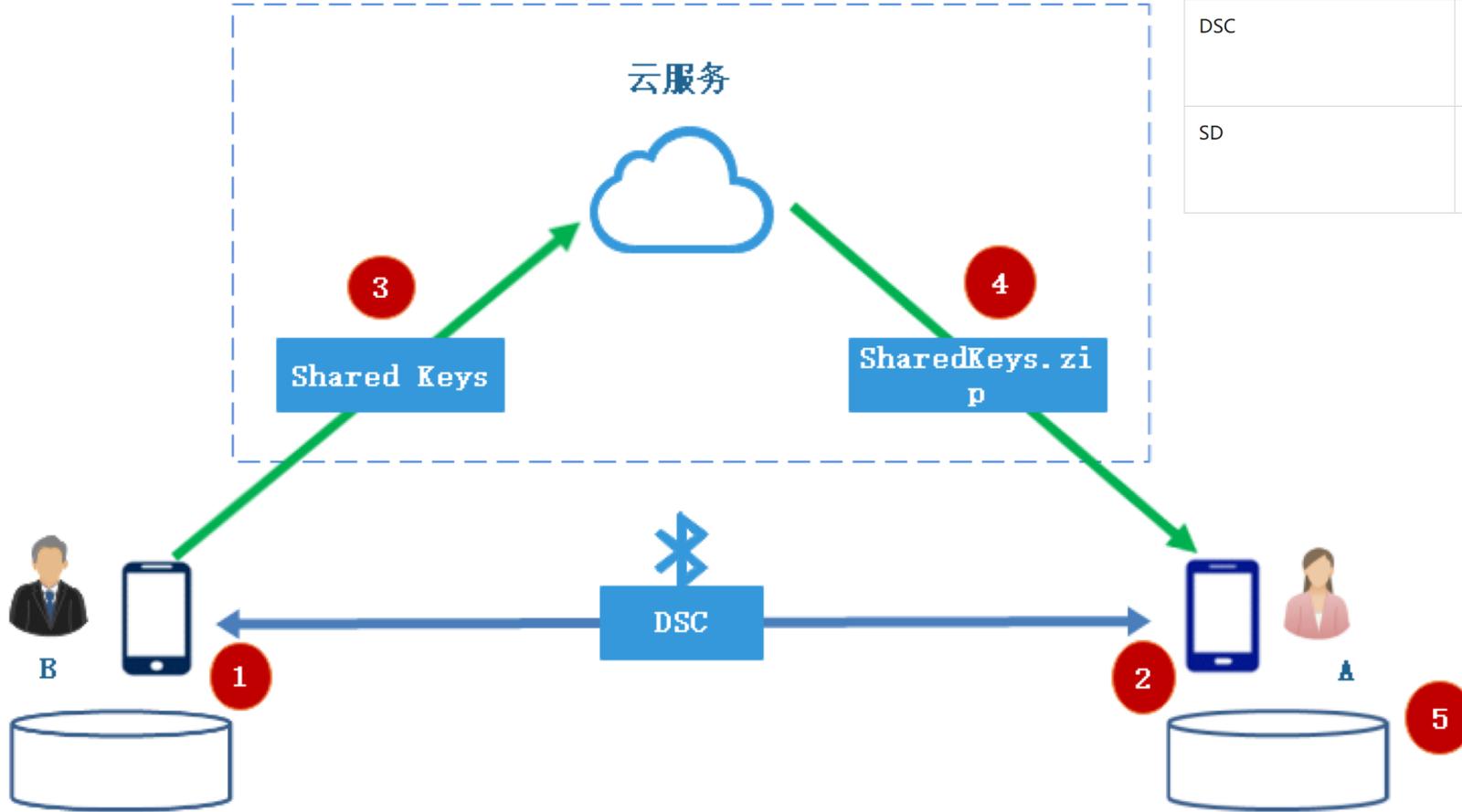
■ 华为HMS Core提供的以保护用户隐私为前提的、基于BLE的接触跟踪基础服务

- 以保护用户隐私为前提，提供设备间接触追踪基础能力
- 华为手机用户与其它安卓、苹果手机用户互联互通
- 各国公共卫生机构可以使用接触卫士API开发新冠病毒接触跟踪的应用



组织/来源	中文	英文
学术名称	接触跟踪	Contact Tracing
学术名称	临近跟踪	Proximity Tracing
Apple Google	暴露通知	Exposure Notification
Huawei	接触卫士	Contact Shield

Contact Shield流程



用户A知道自己是密接，但不知道是因为接触了谁成为了密接，并且用户B和云服务器什么都不知道.....

名称	描述
Periodic Key	周期性密钥，为了保护用户的隐私使用的一种密钥，该密钥每天更新一次。
Shared Key	共享密钥，周期性密钥的子集。当安装了应用的设备所有者（“最终用户”）被确诊为COVID-19感染者后，在征得最终用户许可的前提下，开发者基于接触卫士而开发的应用将会上传有效期为一定天数的周期性密钥，所上传的周期性密钥叫做共享密钥。
DSC	动态共享码（Dynamic Sharing Code，简称DSC），从周期性密钥派生的用于保护隐私的共享码，作为蓝牙广播的载荷数据进行广播，共享码每10分钟会生成一个，BLE广播内容的间隔为15分钟。
SD	辅助元数据（Supplementary Data，简称SD），保护隐私的元数据，该元数据将用于承载协议版本控制和传输功率（Tx_power），以实现更好的接触距离计算。元数据大约每15分钟更改一次，其频率与动态共享码相同，都是为了防止被跟踪。

- ① 通过BLE向周围广播DSC。
- ② 将扫描到的DSC、RSSI存储至本地数据库中。
- ③ 确诊患者将14天内的周期性密钥上传至服务器。
- ④ 定期从服务器下载确诊者的周期性密钥的压缩文件。
- ⑤ 调用Contact Shield进行接触诊断,生成密接诊断报告。

Contact Shield全球落地



通信行程卡 (中国大陆)



安心出行 (中国香港)



Immuni (Italy)



Covid Alert South Africa



Tabaud (Saudi Arabia)



Health Passport Scan (Ireland)



Stop Corona (Austria)



Koronavilkku (Finland)



Exposure Logging Active

Increased Risk

Exposures on 7 days with increased risk

Most recently on 19 May 2022

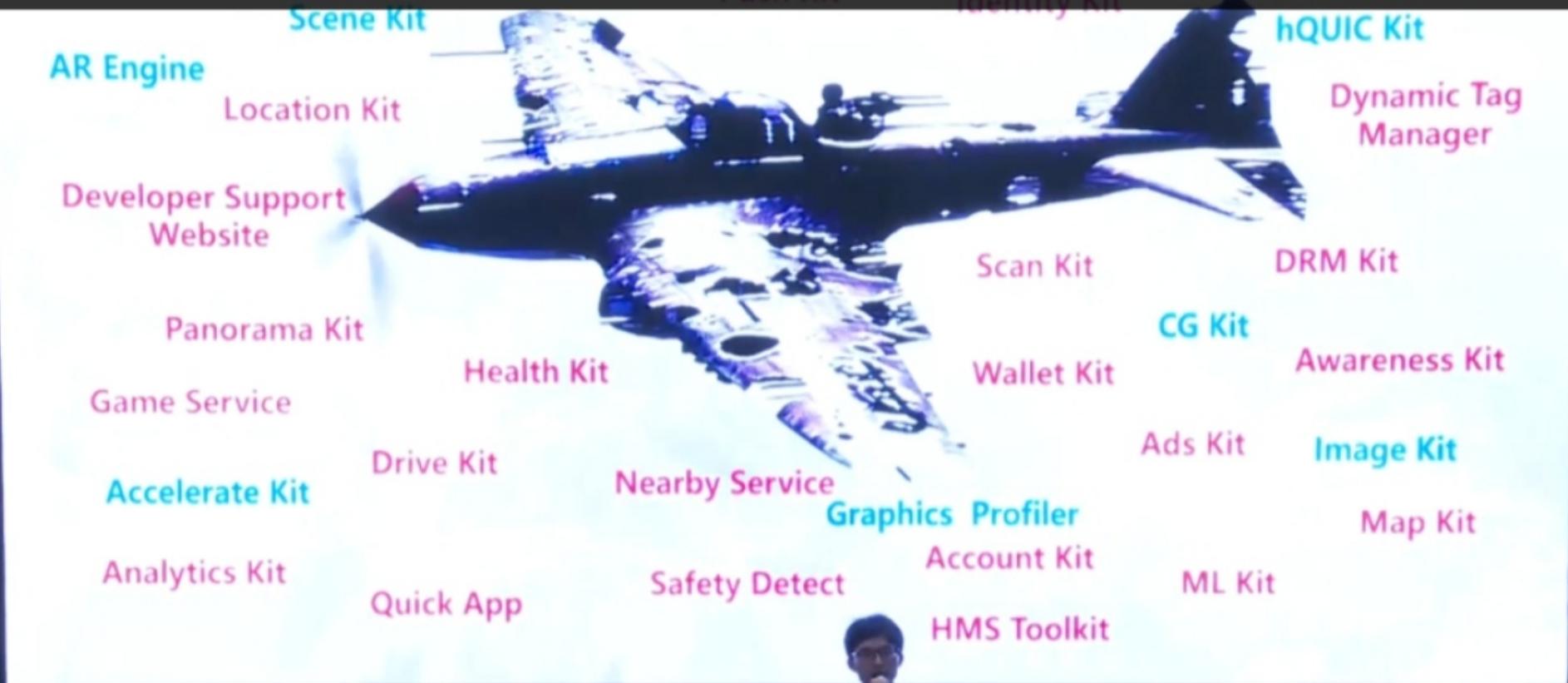
Updated: Today, 14:53

Manage Your Tests

Find testing centers near you and register a test, so you can warn others more quickly, or to request your digital COVID test certificate.

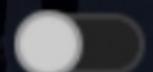
Continue

Corona Warn-App (Germany)



HMS Core

必胜!



落地 \neq 价值

通信行程卡

4G 5G 27 B/s 15:22

深圳卫健委 >

直击深圳“流调”！ 与病毒赛跑，忙起来几天几夜不合眼

1月12日，深圳市报告1例新增新冠肺炎确诊病例。与此同时，该病例14天的活动轨迹、人员接触情况已经清清楚楚，2098名密接者、2625名次密接者全部找到，并按照要求管控起来。这是阻断疫情传播的重要环节。

这么多的人员、场所信息从哪里来？背后，有一支队伍在发挥重要作用——流行病学调查处置团队。



视频来源：南方+

他们是在疫情中跑在最前方的人，是“尖刀连”“侦察兵”。“我们的任务是赶在病毒感染下一个人之前，把它拦住，不让疫情继续扩大。”罗湖流调处置组成员、罗湖区疾控中心健康教育与促进科主任郑庆鸣说。

4G 5G 27 B/s 15:22

行程指南

通信大数据行程卡 使用说明 隐私政策

通信行程卡

"通信行程卡"是由工信部指导，中国信通院、中国电信、中国移动、中国联通共同推出的行程查询服务，可以免费为用户提供本人过往14天内到过的国家(地区)和国内城市证明。

9:41

通信大数据行程卡
疫情防控，人人有责

手机号 获取验证码

请输入验证码

同意并授权运营商查询本人在疫情期间14天内到过的地区信息

授权系统获取新冠肺炎确诊用户蓝牙接触信息

登录

行程卡使用说明

CAICT 中国信通院 中国电信 中国移动 中国联通

4G 5G 0 K/s 15:22

行程指南

行程卡+蓝牙

通信行程卡手机APP版本中加入了蓝牙近距离接触提醒功能，如果您近距离接触过的陌生人诊断为新冠肺炎确诊/疑似患者，使用我们的服务将更及时地提醒您注意健康观察。

行程指南

行程卡2.0的工作原理

01. 下载并安装通信行程卡APP
02. 当进入人员密集或流动性较大的场所时登录通信行程卡APP并按提示开启蓝牙
03. 如果近距离接触的双方都在使用行程卡服务，您与对方的手机将通过蓝牙交互记录下接触事件，接触的时间和距离将加密保存在各自手机中
04. 接触记录将在各自的手机中加密存储，如14天之内接触双方并无感染新冠肺炎，数据将自动清除
05. 如果用户被诊断为新冠肺炎确诊/疑似患者，在经患者同意后，他的加密ID将上传至服务端，您的手机会定期从服务端下载确诊/疑似患者的加密ID，并与手机中存储的接触记录进行比对，如果您的接触列表中有确诊/疑似患者，将在第一时间收到APP和短信提醒



中國人民大學
RENMIN UNIVERSITY OF CHINA



从创新到落地

天时
地利
人和

广告时间

研究方向：应用最新的大数据和机器学习技术提升网络系统的性能和可靠性

无线局域网(WLAN)



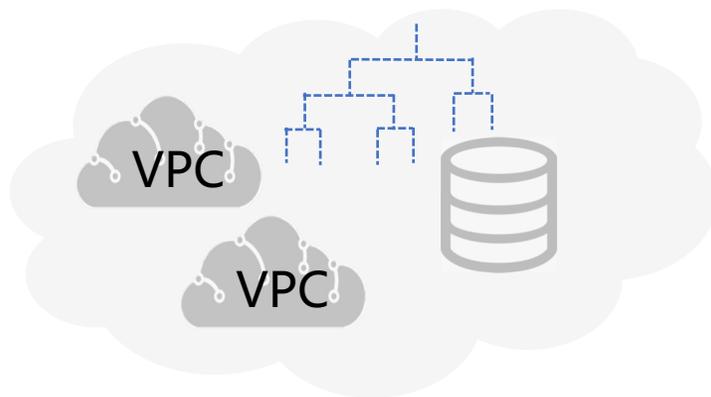
近场通信、物联网(IoT)

技术趋势：单设备智能
→ 多设备协同

关键难题：弱网稳固传输、数据
传输极简化、无线RDMA

抗干扰

数据中心网络(DCN)



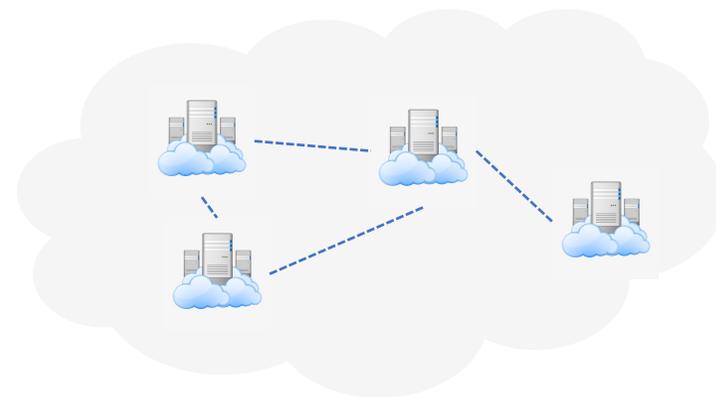
高性能数据库、存储、大
数据、HPC、分布式AI

技术趋势：以计算为中心
→ 以数据为中心

关键难题：解决通信瓶颈、
协议卸载(RDMA)

高性能

广域网(WAN)



算力网络、元宇宙、实时音
视频等互联网应用和服务

技术趋势：移动互联网 →
沉浸式互联网&产业互联网

关键难题：算力接入网、确定
性网络、瓶颈定位(因果分析)

确定性

Feel free to contact me

Email: tong.li@ruc.edu.cn

