

# 目录

## ■ 工业互联网

海尔：COSMOPlat 海尔工业互联网实践 ..... 2

## ■ 城市大脑

阿里云：城市大脑开源创新平台 ..... 11

## ■ RISC-V 开源指令系统架构 CPU

倪光南：迎接开源芯片新潮流 ..... 17

编者的话<sup>1</sup>：自研 RISC-V DSP 芯片量产，突破美国芯片封锁 ..... 28

## ■ 5G

项立刚：5G 带来商业模式新变革和关于 6G 的思考 ..... 30

## ■ 网约车和自动驾驶

滴滴出行：采用基于开源的深度信息技术推动交通出行变革 ..... 50

## ■ 边缘计算

九州云：在边缘计算中的创新与实践 ..... 63

## ■ 人脸识别

编者的话：美欧点评商汤人脸识别 ..... 80

## ■ “缺芯少魂” 攻关

编者的话：点评“缺芯少魂”攻关 ..... 82

---

<sup>1</sup> COPU 编者：陈伟、鞠东颖

# 工业互联网

# COSMOPlat 工业互联网实践

海尔集团

## 一、工业互联网理解

以基于开源的新一代信息技术为核心的数字基础设施为经济发展释放巨大潜力，已成为全球共识。工业互联网作为新一代 ICT 技术与制造业深度融合的产物，符合产业升级需要，也代表着中国经济高质量发展的方向。

工业互联网已成为工业经济高质量发展的新引擎。工业互联网不仅支撑起工业智能制造的高效率生产，而且建立的大规模定制新模式使产销高精度对接，推动工业经济高效率、高精度发展，工业互联网成为连接工厂与用户的新渠道。

工业互联网是数字新基建的核心。新基建中的 5G 大数据中心以及人工智能将以工业互联网平台为载体，与工业深度融合，发挥出数字新基建的基础支撑作用，共同为中国制造由制造大国向制造强国转变赋能。

## 二、工业互联网实践目标

早在 2012 年海尔已经意识到工业互联网在带动企业研发、生产、运营、营销和管理方式转型中起到的基础支撑作用，制定出工业互联网“三步走”战略目标：

第一步完成自身数字化转型，实现传统制造向智能化生产的转

变；

第二步裂变商业模式，实现从传统大规模生产到以用户为中心的大规模定制模式的转变；

第三步打造全球引领的物联网生态品牌，构建共生、互生的“热带雨林式”生态。

### 三、卡奥斯 COSMOPlat 平台简介

承接上述目标的卡奥斯 COSMOPlat 工业互联网平台被海尔集团董事局主席、首席执行官张瑞敏列为海尔集团的“一把手工程”，举全集团之力推进 COSMOPlat 工业互联网平台建设和应用推广。卡奥斯是在海尔“人单合一”模式指引下产生的不断孕育新物种的工业互联网生态品牌平台，旨在为混沌中寻求新生的企业提供转型升级解决方案，联合各方资源缔造共创共享、面向未来的物联网新生态。

卡奥斯最大的差异化之一是基于 BaaS（全称 Business & Best Practice as a Service）的理论创新，开创了工业互联网的体系创新（企业云、平台云、生态云），创造了工业互联网大规模普及的新载体。

Business（商业逻辑）代表着先进的生产关系，它促进生产力的解放和提高；

Best Practice（最佳实践）代表着先进的生产力，它驱动生产关系的发展和变革；

BaaS 把两者（生产关系和生产力）有机的融合在一起，裂变成一种新的力量即先进的生态力；

BaaS（先进的生态力） = Business（先进的生产关系） + BestPractice（先进的生产力）。

中国管理科学学会“管理科学奖”，对卡奥斯 COSMOPlat 首创的先进生态力理论（融合管理创新和实践创新）的成果进行了充分肯定，称其为“物联网时代最重要的一种企业家战略思维”。



在卡奥斯先进生态力理论指导下，具体从三个层面介绍平台的工业互联网实践之路。

#### 四、工业互联网实践过程以及效果

在智能化生产层面，COSMOPlat 平台首先对海尔三十多年的制造最佳实践经验进行总结提炼，抽象出研发仿真模型、业务流程模型、工业机理模型和数据算法模型四大类工业模型，其次利用数字孪生、虚拟仿真、三维空间多源异构数据、先进传感等软硬件技术，实现工业生产和虚拟模型之间的映射，做到了智能化排产、可视化追溯与预防性维修。海尔先后在全球基于 COSMOPlat 平台打造了多个工业园、制造中心、互联工厂，实现了高精度、高效率的海尔智能

制造。

在大规模定制层面，COSMOPlat 将传统面向流程的大规模制造应用和技术层面解构为基于工业微服务的方式，打通了采购、生产、营销链路，并基于区块链智能合约在关键节点自执行，杜绝了执行主体和交易双方的“道德风险”，实现了用户与制造、网器、全流程互联的柔性化，打造出以用户体验为中心的大规模定制新模式，不入库率同行业遥遥领先。

在生态建设层面，COSMOPlat 把集“技术创新+模式创新+管理创新”于一体的平台体系，解构成了交互、研发、营销、采购、生产、物流、服务等七大模块，进而封装成可复制、可推广的云化解决方案，实现制造能力和工业知识的模块化，支撑起跨行业跨领域推广应用。在此基础上，又沉淀出了针对企业转型的六大赋能能力，为平台上的企业提供软硬件一体化的整体解决方案。这体现为：

- 全要素互联的数据力；
- 革新物理世界数字化表达的仿真力；
- 实施大规模定制模式的定制力；
- 开源开放满足企业个性化需求的开源力；
- 保障企业数据的安全力；
- 支持企业共创共享共赢的生态力。

通过提供软硬件一体化的整体解决方案，利用“六大能力”赋能平台可持续发展，聚合企业、用户、资源，打造“热带雨林”式生态。



COSMOplat 卡奥斯已经构建了包括建材、房车等行业在内的 15 个链群，涌现出海享淘、幸达等新物种，为用户构建了衣、食、住、行、康、养、医、教等全方位的大生态。

## 五、卡奥斯 COSMOplat 平台获奖情况以及获奖理由

2020 年 8 月，卡奥斯以 557.67 亿元的品牌价值，挺进中国品牌价值百强榜，成为唯一入榜的工业互联网平台。

在智能制造领域，2020 年，海尔基于卡奥斯 COSMOplat 率先成为全球唯一一个同时拥有两座“灯塔工厂”的企业，并基于打造“灯塔工厂”的经验，卡奥斯 COSMOplat 赋能的海尔中德智慧园区集合海尔中央空调互联工厂等 4 家互联工厂打造了“灯塔基地”，让“灯塔经验”覆盖更多企业；

在大规模定制领域，卡奥斯 COSMOplat 已在国际三大标准组织 IEEE、ISO、IEC 主导制定了大规模定制的国际标准，具备了领先的标准话语权，探索出大规模定制的新模式；

在双跨领域，卡奥斯平台已连续两年被国家工信部认定为跨行

业、跨领域的工业互联网“双跨”平台之首；

在国际上，平台位居全球知名机构 Forrester 工业互联网领导者第一象限，海尔 COSMOPlat 在 18 项指标拿下满分，在产品提供维度，海尔在应用使能功能、分析和数据等指标得分较高，在战略布局维度，海尔在合作伙伴、商业模式、创新路线、平台差异化、交付模式等指标得到高分，此外，在市场表现维度，海尔在客户群、链接设备等指标获得了多个高分。

## 六、海尔工业互联网实践与开源关系

平台能在这么短的时间内完成海尔工业互联网转型，打造物联网生态，并得到国家和国际的认可，这与平台一开始就坚持开源策略是分不开的。工业互联网是互联网和制造业深度融合所形成的产业和应用生态，其发展不能单靠一、两项技术突破，也不是一两家企业的独角戏，要立足于建设开源的生态系统，构成一个相互作用、依赖和发展的整体，因为与工业经济时代独占、排他思维方式不同，数字经济时代需要共享和生态，开源已成为当前全球的一种创新和协作模式，以开源模式构建工业互联网已经成为一种趋势，西门子的 MindSphere，施耐德的 EcoStruxure，包括 Predix，这些平台如今已经吸引了大量软件开发者加入，布局工业互联网，所以推进工业互联网发展，就要打好开源战。

平台在一开始就积极利用开源资源赋能自身建设，在数据平台和数据服务方面，引入了 Hadoop、Spark、Kafka 等开源技术，应用和中间件服务方面，选用 Tomcat、Hyperledger、Tensorflow 等内

容，在 IaaS 和 PaaS 方面，也选用了 Openstack、Docker 等成熟开源技术，极大的提高了平台建设质量与效率。另外在原有开源技术基础上，结合工业应用场景和生态，深度定制，实现平台技术创新。例如在 Dubbo/Dubbox、Spring Cloud、Istio 等开源组件基础上，深度定制，打造了工业领域微服务及治理平台，完成无侵入服务治理技术攻关与服务灰度发布与服务监控模块建设，同时通过“自研 + 社区 + 技术合作伙伴”的开放协作模式，不断扩大平台开源的参与广度，让包括自身研发队伍、社区聚合的开发资源以及第三方技术合作伙伴等更多的力量加入到开源推进中来。

过去开源大部分发生在消费互联网领域，工业互联网领域开源项目并不多。平台从工业 PaaS 这一角度切入，以低代码+高控制力双引擎平台支撑应用开源演进，目前正在推进面向跨行业、跨领域复制的工业 APP 开发框架和基于开源 OpenFaaS 构建工业函数服务等开源项目，打造开源生态，积极构建“生态开放、共创共建”的平台格局。

## 七、海尔工业互联网远景预期

对于工业互联网领域而言，本行业的开源属性逐渐显现。近几年国内开源的有百度发布 OpenEdge 边缘计算开源项目、腾讯的物联网实时操作系统 TencentOS tiny 以及开源工业互联网联盟推出的 IOT3000 开源方案等。未来需要借鉴其他行业的开源风险，尽量规避开源许可证变更、不合规使用开源等风险问题，另外也提议开源联盟组织推广国内开源项目，搭建开源配套基础设施，进一步推

进标准制定，助力工业互联网领域开源发展。

综上，海尔工业互联网实践取得的成就是行业与国家对于海尔COSMOPlat平台工业互联网建设的认可，未来，海尔将坚定不移地发展工业互联网“新基建”，坚持开源、开放的共创、共享策略。“尽己之心 推己及人”，海尔卡奥斯平台将继续为工业互联网带动中小企业高质量跨越式发展贡献自己的力量，成为引领中国工业互联网发展的巨大平台。

# 城市大脑

# 城市大脑开源创新平台

阿里云

城市大脑国家新一代人工智能开放创新平台，是基于人工智能和大数据处理技术，以城市大脑系统为蓝本，构建一整套计算、数据处理、应用服务平台，为城市安防治理、公共服务及其他各行业的智能应用构建起开放、多元的生态体系。目前，项目实施稳步推进，取得了多项阶段性成果。

## 一、平台建设进展

升级建成基于公共云或私有云的城市大脑平台，基于自研 NPU 芯片、面向多任务场景和海量数据的异构计算集群优化框架，搭建异构集群资源平台，开发大规模视频实时结构化分析流计算系统和超大规模数据检索与回溯系统，实现超大规模视频分布式处理和数据检索与回溯，利用可伸缩式海量交通数据实时汇聚与弹性计算平台，为海量城市数据开发和全城协同治理提供强有力的工具支持。汇聚基础城市数据，形成标准的训练、验证与测试城市数据集，建设人工智能基础城市资源数据库。基于云边协同技术，为城市大脑中的视频等大数据提供近场边云协同的计算资源和计算能力的服务，为城市大脑中跨区域智能应用提供云上离线训练、云边在线优化、边云协同推理的机制。

平台将以机器视觉为主，在成熟的开放服务机制下，继续为产

业链和创新链上下游单位提供开放服务，包括开放基础公共数据服务应用接口，开放基础智能分析算法接口，并开源通用并行化深度学习框架，为高校、科研院所、中小型企业、开发者提供一站式的数据、资源、研发工具及解决方案支撑，降低行业门槛。



图 1 城市大脑整体架构

## 二、技术创新与场景落地情况

城市大脑的持续加大技术研发和创新投入，目前形成了一些阶段性成果。

### 1、建立百亿级别城市视觉数据搜索系统

基于海量的城市视觉数据（交通监控、视频图像等），研发城市视觉数据搜索系统，搜索对象包括人、车、非机动车等城市移动对象，也包括标志建筑、公共场所等固定对象。在多个场景空间（可疑人员、肇事车辆追踪等）进行验证。建立了百亿级别的城市视觉数据库，支持人、车、非机动车、建筑、公共场所的视觉搜索，支持万级并发的查询请求响应，支持一秒以内的搜索响应，支持图片检索

和视频检索，检索 mAP 高于 50%。

## 2、实现高精度城市车流人流预测

基于城市中海量异构数据，深入、全面了解城市交通系统的运行规律，开发对城市车流、人流进行准确预测的算法组件。综合利用城市历史和当前的视频数据，通过跨摄像头联合分析，实时预测出城市各区域一段时间后的车流量、人流量。交通方面，交警能够提前进行交通疏导和信号控制，防止出现交通拥堵和事故；安保方面，安保人员能够提前做好人员疏散，防止出现人员踩踏事件等安全问题；大型商业中心和商铺方面，能够提前做好相应服务的准备工作，合理安排运营资源，更好、更高效地服务顾客。能够综合利用和处理车流、人流速度和密度，天气与气候，路网建模，特殊事件等多维交通数据，有效预测城区 80%以上面积的路网流量。对城市核心区域，1 小时内车流、人流预测的误差不高于 8%。

## 3、构建交通事件事故检测系统

基于城市交通监控实时视频数据，研发交通事件事故检测系统。系统通过对各种车型的机动车、非机动车和行人等交通元素以及摄像机巡航位、道路线圈的检测识别，利用智能规则实现对城市地面道路、快速路及绕城高速路上发生的违停、闯禁、抛锚等交通事件事故的实时捕捉报警，同时保存录像片段作为非现场执法的证据。能够支持上万个路口的交通监控视频流的实时分析，对违停、闯禁、抛锚等交通事件事故的实时捕捉报警，事件事故的检测准确率达到 90%，支持录像片段的单独存储和调取。

## 4、搭建城市智能交通优化与控制系统

基于海量实时计算平台，利用多源数据融合匹配的技术，将交警数据、公交数据、视频数据、地图互联网数据、环境数据整合处理，通过模糊认知反演算法，发现复杂场景背后的超时、超距弱关联，成功应用到道路交通分析治理领域，在交通中实现从单个点、单条线到整个城市的交通优化控制。实现智能交通控制系统降低拥堵指数 20%以上；提升城市主要拥堵区域通行速度 15%，压缩特种车辆到达时间 50%；提升城市公交主干线服务人数 20%；通过导航系统为 5 亿人群提供躲避拥堵服务。

目前，城市大脑已经在杭州、苏州、上海、衢州、澳门、海南等 30 多个国家落地 60 多个智慧城市项目。以杭州城市大脑为例，现已升级至第 3 代，1.0 时代，数据打通开始在交通场景发挥作用；2.0 时代，更多的数据开始跨领域跨部门的共享利用；3.0 时代，城市数据多维多元深度融合，并且智能化全面介入。



图 2 城市大脑 3.0



图 3 杭州城市大脑 3.0

### 三、开放共享及生态建设情况

#### 1、开放共享

城市大脑平台向社会全面开放，城市大脑走向更普惠的生态型平台。面向对象将包括政府管理部门、大学与科研机构、咨询公司、中小微企业和行业开发者。城市大脑平台的全面开放，有力助推我国人工智能行业开放共享及生态建设。对于科研机构和政府部门，更切实的将城市治理从理论走向实践；对于中小微企业和行业开发者，避免了重复造轮子，提升创新创业的成功几率；对于核心零部件厂商，提供了参与新一轮标准制定的机会；对于芯片等相关企业，意味着拥有底层驱动的机会。

2020 年 6 月，阿里云将城市大脑中的核心机器视觉能力形成视觉智能开放平台向社会开放，视觉智能开放平台是基于阿里巴巴视觉智能技术实践经验，面向视觉智能技术的开发与应用用户，为其提供好用、易用、普惠的视觉智能 API 服务，帮助企业、开发者快速建立视觉智能技术的应用算法的综合性视觉 AI 算法平台。近期平台上线了 30+个 AI 算法， AI 算法数量突破了 100 大关。为了更好

的帮助中小企业和独立开发者快速对接视觉 AI 算法，阿里云视觉智能开放平台免费开放平台上现有的 100 余种视觉 AI 算法服务的使用权限，服务调用不收取任何费用。

## 2、生态带动

一方面，融入垂直行业合作伙伴。人工智能、云计算和大数据是阿里的优势，阿里选择开放与投资企业、合作伙伴共建垂直行业，积极发展生态合作伙伴，带动垂直行业技术企业共同发展。例如向生态合作伙伴提供接入服务，也把最新的技术注入到传统企业的解决方案中，构建了包括浙大中控、银江科技、浙江大华、图盟科技、盛迅信息、宝信软件、半云科技、浩鲸科技等企业的行业生态。

另一方面，向中小微企业和行业开发者开放。城市大脑平台开放之后，中小微企业和行业开发者加入其中，成为生态中的一员，他们可以和阿里共享智慧城市领域的相关技术及数据。基于阿里大规模云端计算容量及能力，能够低门槛实现超强的虚拟运行能力，进而形成一个快速迭代的生态闭环。此外，中小微企业和行业开发者还能在这个生态中更迅速、更容易的接触到产业链上下游的生态伙伴，从而有效提升创新创业的成功率。

# RISC-V 开源指令系统架构 CPU

# 迎接开源芯片新潮流

倪光南

## 处理器（CPU）芯片现状

- CPU芯片和操作系统是网信领域最基础的核心技术
- 中国在这方面还受制于人，常被比拟为“**缺芯少魂**”
- 实践表明，每一种CPU在市场竞争中要取得成功在很大程度上取决于其生态系统
- 目前世界上生态最为成熟的二种指令集架构的CPU：

➤ **Intel x86指令集架构 CPU**——历史悠久

垄断桌面和服务器领域，被美国的Intel和AMD两家公司拥有，罕有对外授权



➤ **ARM指令集架构CPU**——历史较长

在移动和嵌入式等领域占垄断地位

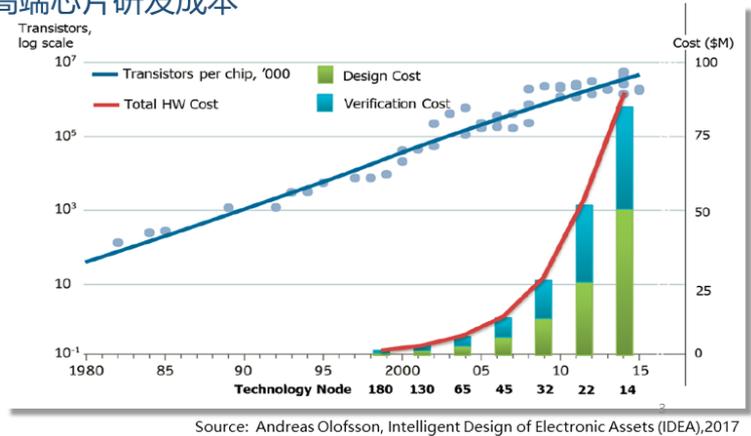
世界上很多公司花费数百万元~数亿元费用，向ARM公司购买CPU许可



2

## 芯片设计门槛极高

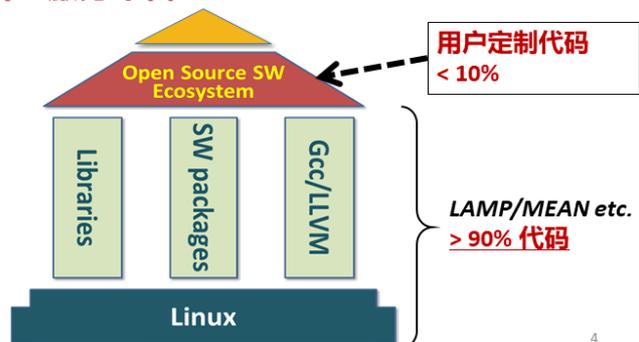
- 14nm工艺为例，**十亿元以上**研发经费，投入**几百**人年
- 只有**少数**企业能承受中高端芯片研发成本
- **制约了**芯片领域的**创新**



## 开源软件的成功经验

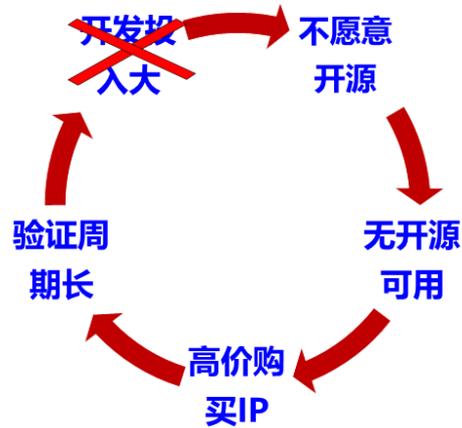
- **降低互联网创新的门槛**
  - **3-5位**开发人员用**几个月**即可快速开发创新业务，滴滴、摩拜等
- **提高互联网企业自主能力**
  - 互联网公司能构建**基于开源软件的自主服务平台**

- “开源”（开放源代码）是一种开发模式和商业模式
- 开源软件（以Linux为代表）已成为当前软件产业的主流
- “开源”能否应用于芯片领域？



4

## 开源芯片“死结”



5

## RISC-V指令集：降低处理器芯片IP成本的新模式

- 由美国加州大学伯克利分校计算机科学部门于2010年发明
- 基于广泛使用且标准十分宽松的BSD许可证 (Berkeley Software Distribution license)
- 用户**可自由免费**地使用RISC-V进行CPU设计、开发并添加自有指令集进行拓展，自主选择是否公开发布、或商业销售、或更换其他许可协议、或完全闭源使用



Field	Standard	Free, Open Impl.	Proprietary Impl.
Networking	Ethernet, TCP/IP	Many	Many
OS	Posix	Linux, FreeBSD	M/S Windows
Compilers	C	gcc, LLVM	Intel icc, ARMcc
Databases	SQL	MySQL, PostgreSQL	Oracle 12C, M/S DB2
Graphics	OpenGL	Mesa3D	M/S DirectX
ISA	??????	-----	x86, ARM, IBM360

## RISC-V的优势

### (1) 设计优势

- 免除授权费用和知识产权风险的完全开源免费，这是RISC-V存在的主要意义。传统的芯片设计有时需要上亿研发费用、投入上百人的团队和花费大量时间，使得中小型科技企业不可能承担芯片研发，被迫只能使用市场上已有产品，既造成了芯片市场的绝对垄断，又丧失发展的主动权。而开源芯片设计可以将芯片设计门槛大大降低，甚至于3-5人的小团队在3-4个月内，只需几万元便能研制出一款有市场竞争力的芯片，十分适应于中小企业创业，市场前景广阔。能够有效促进芯片产业的繁荣

### (2) 技术优势

- RISC-V指令集在最初设计时就明确表示要追求简约，丢弃历史包袱。目前成型的技术代码集小、支持模块化，性能十分优越，能够满足从微控制器到超级计算机等各种复杂程度的处理器需求，支持从FPGA、ASIC乃至未来器件等多种实现方式。同时能够高效地实现各种微结构，支持大量定制与加速功能，并与现有软件及编程语言可良好适配

7

## RISC-V的优势（续）

### (3) 市场优势

- 随着人工智能和物联网等新一代信息技术的发展，各种场景下对于CPU需求日益碎片化，嵌入式应用前景广阔，低功耗、低成本和定制化需求越来越大，这使得RISC-V的精简、低功耗、模块化和可扩展的优势与数字经济未来发展方向十分契合

### (4) 管理优势

- 最初的RISC-V指令集是20世纪50年代，因而大多数专利早已过期，在商业化和开源使用方面不会存在大的专利壁垒和成本问题；另外，RISC-V由2005年成立的基金会管理，这是一个非盈利的中立机构和开放社区，主要负责日常维护RISC-V指令集标准手册与架构文档，参与RISC-V相关规范地制定和软硬件生态系统的发展。基金会采取会员制管理，目前共有上百家会员单位，中科院计算所是发起和审计会员，阿里、华为是白金会员，这种开放式的管理模式能够有效地促进RISC-V社区的交流和创新，形成生态系统，从而降低研发成本，打破芯片市场的垄断

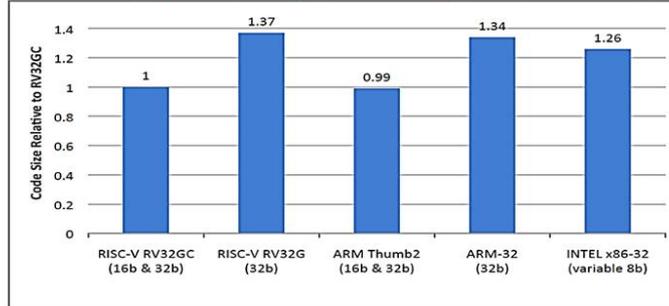
8

## RISC-V技术优势：追求简约

### 以页数和单词数衡量的 RISC-V、ARM-32 和 x86-32 指令集手册比较

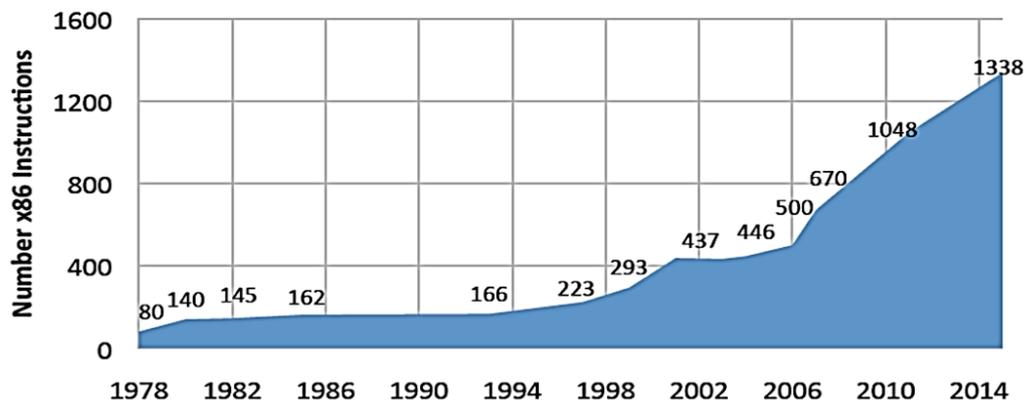
ISA	Pages	Words	Hours to read	Weeks to read
RISC-V	236	76,702	6	0.2
ARM-32	2736	895,032	79	1.9
x86-32	2198	2,186,259	182	4.5

### RV32G, ARM-32, x86-32, RV32C和Thumb-2程序的相对大小



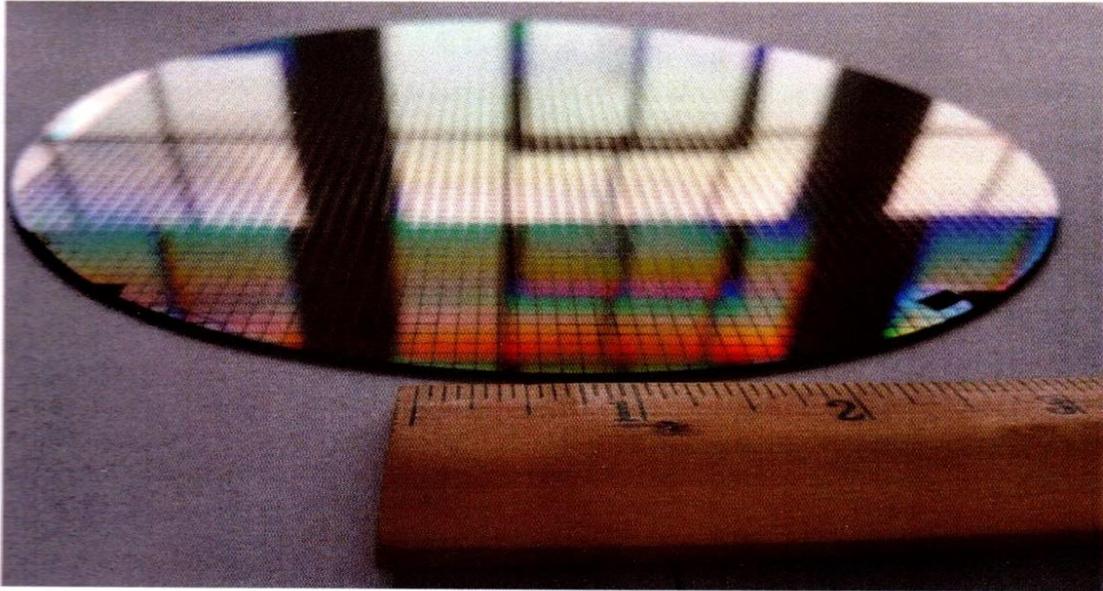
9

## RISC-V技术优势：丢弃历史包袱



x86指令集自诞生以来指令数量的增长  
自1978年诞生时的80条指令，到2015年增长了16倍，  
达1338条指令，且仍在增长，历史包袱沉重

10

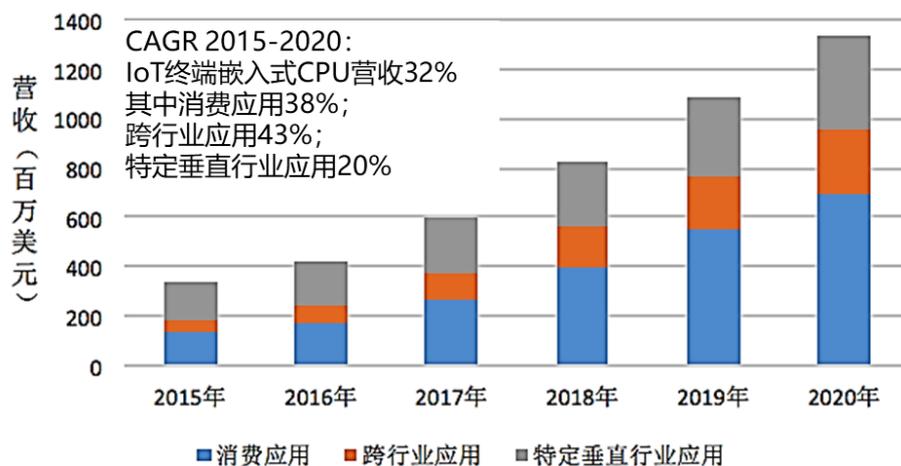


由SiFive设计的直径为8英寸的RISC-V晶圆。它有两种类型的RISC-V芯片，使用较旧的较大加工线。  
FE310芯片为2.65mm×2.72mm，SiFive测试芯片为2.89mm×2.72mm。  
一片晶圆上有1846片FE310和1866片SiFive测试芯片，总共3712个芯片。

11

## RISC-V市场优势：新一代信息技术领域潜在市场

全球应用于IoT终端的嵌入式CPU市场



12

## RISC-V与x86及ARM指令集架构CPU的比较

技术与市场	x86或ARM架构	RISC-V
架构篇幅	数千页	少于三百页
指令数	指令数繁多，且不同分支不兼容	基本指令集40余条
模块化	不支持	支持
可扩展性	不支持	支持
硬件实现	复杂性高	硬件设计与编译实现非常简单
商业运作	X86封闭，ARM架构授权昂贵	开源、免费
生态环境	成熟	快速起步中
应用市场	服务器与桌面PC(x86)，移动和嵌入式(ARM)	从物联网切入，可覆盖从微控制器到超级计算机的全计算领域
应用特点	服务器、PC、移动和嵌入式应用领域居垄断地位	高性能（面积、功耗和性能）、普适（可自由扩展和裁剪）、可控（满足差异化和定制化）
应用风险	不可控、缺乏应用弹性、较高成本	生态不足、碎片化、专利问题

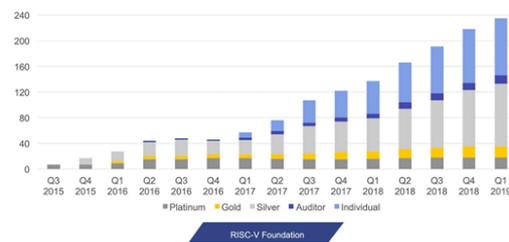
## RISC-V管理优势：RISC-V国际基金会

中国成员：计算所（发起、审计）、清华（审计）、华为（金）、阿里（金）.....

RISC-V基金会2015年成立，目前包括235家会员单位

- 白金会员19家
- 普通会员216家

RISC-V Foundation Growth  
September 2015 – February 2019



来源：RISC-V基金会

### Platinum Members



### Gold, Silver & Auditor Members



基金会规模迅速壮大，进入加速发展期

## RISC-V国际生态现状



注：以上仅为部分代表厂商  
15

## RISC-V国内生态现状



注：以上仅为部分代表厂商  
16

## RISC-V中国组织——目标相同，协作共赢

### 中国开放指令生态 (RISC-V) 联盟 (CRVA)



由网信办、工信部和中科院三个单位指导，于2018/11/8第五届世界互联网大会上宣告成立，挂靠在中国计算所。CRVA旨在召集从事RISC-V指令集、架构、芯片、软件、整机应用等产业链各环节企事业单位及相关社会团体，自愿组成一个全国性、综合性、联合性、非营利性的社团组织。截至2019年7月，联盟会员规模已由创始时的12家增长至58家

### 中国RISC-V产业联盟 (CRVIC)



联盟成立大会于2018/9/20在上海集成电路行业协会成功召开。由上海芯原、芯来科技、上海高性能集成电路设计中心、上海集成电路行业协会、上海物联网行业协会等单位共同发起

17

## 基于RISC-V国产CPU未来展望

CPU架构	基于x86架构CPU	基于ARM 架构授权国产CPU	申威、龙芯等国产CPU	基于RISC-V 国产CPU
自主可控测评	难以满足要求	现阶段能满足要求 未来发展存在变数	能满足要求	可望满足要求
既有生态状况	成熟	成熟	发展中	起步中
新兴市场预测	不详	不详	不详	前景看好
授权费用	缺乏成熟的授权模式	架构授权费用高	无	无

18

## 开源芯片的前景



编者的话：

## 自研 RISC-V DSP 芯片量产，突破美国芯片封锁

2021 年 4 月 12 日讯，中科昊芯公司研制的 RISC-V DSP 芯片即将量产，突破美国芯片封锁。

RISC-V 是与 X86、ARM 并列为三大指令集架构，RISC-V 是开源的，DSP 芯片可对数据进行实时处理，为华为的鸿蒙系统推广提供平台。

# 5G（第五代移动通信）

# 5G 带来商业模式新变革和关于 6G 的思考

项立刚

2020 年中国已经建成 72 万个 5G 基站。5G 带来的业务变化很容易直接看到，而商业模式的变化却是间接的，隐含在业务的后面，不认真研究无法看到，所以对于绝大部分普通人来说，不太会看到业务后面的业务模式和商业模式变化，这就是别人看不到的机会。

4G 网络只能传输 1080P 的视频，如果更高的清晰度，网络就很难承受，尤其是大量用户使用的环境下，用了 5G 网络，就可以支持 4K 甚至是 8K 的高清传输，这一步对于任何一个人都是显而易见的，从这个角度去开发产品，最为容易，一般人都能看到这个机会，而除了一眼可以看出来的视频效果提升，随着视频能力提升，带来建立在视频业务之上的商业模式改变，一般人就很难看清，而能看到这些，就更容易走在别人的前面，找到属于自己的机会。即使是同样的业务，也可以形成不同于别人的商业模式，形成自己的差异化。

## 一、技术能力的提升一定会带来商业模式的变迁

全世界 2020 年建设约 100 万个 5G 基站，中国占比 70%。

我们分析几个移动互联网时代商业模式的改变，作为 5G 时代新机会的参照。

传统的音乐销售模式主要是两种方式，一种是歌唱家的演唱会，另一种是卖唱片。把歌唱家们的歌曲和音乐录制在唱片、光盘这种载体上，向大众去销售。这个市场很长时间是一个非常小的市场，唱片的时代，有唱机的家庭很少，唱片的价格也贵，一般人买不起，所以唱片销售要到百万，那就是一个了不起的数字，而能够靠出唱片获得较大经济利益的音乐家也很少。

在光盘这种新的介质出现时，音乐可以存储在光盘中，播放器量大大增加了，普通家庭的电脑也可以播放音乐，市场要较唱片扩大了很多，但是这时出现了另一个问题，就是复制太容易，复制成本太低，普通人甚至可以用自己的电脑刻录光盘，专业的刻录机价格也不贵，任何一张光盘就可以作为母盘被用翻录。这个时候音乐家们依然没有在光盘的出现获得经济效益，甚至要维护自己权利变得更加困难了，很长时间音乐家都中呼吁大家使用正版，但是在价格巨大差异面前，守得住道德底线的人很少，音乐传播的主要渠道还是盗版市场。音乐家、歌唱家们没有从自己的音乐作品中获得更大的收入。很长一段时间，中国的唱片（光盘）市场只有10亿产值，相对于十几亿中国人，这个数字非常惨淡。

移动通信出现，移动通信加入到音乐的传播中，方便的传播、随时随地的下载，最初也并没有给音乐家们带来福音，反面是盗版传播效率更高，传播更加容易了。但是一种新的业务模式出现，为音乐传播带来新改变。这种业务是彩铃业务。

彩铃不是为手机用户自己提供服务的产品，它是为给机主打电话的用户提供服务，机主设置了彩铃之后，任何拨打机主的用户听

到回声不再是嘟嘟的声音，而是个性化的音乐，这个音乐很大体现了机主的品味、人生态度，对社会人生的看法，通过这个彩铃附载了机主本人的感情。

大多数的彩铃是用歌声，电信运营商就需要取得音乐家们的授权，在音乐家们呼吁了很多年没法解决的正版问题，在这里解决了，因为电信运营商不可能用盗版音乐，它是一个庞大的商业机构，很容易被音乐人起诉，来维护自己的权利。而对电信运营商来说，支付版权费用并不是问题，因为每一首彩铃电信运营商都会向用户收费，一个月 3-5 元的费用，大部分用户都支付得起，和话费一起收取，收费非常方便。移动通信的加入，把一个很多年没有办法解决的正版音乐问题解决了一部分，音乐家们的收入很大一块来自于电信运营商的彩铃分成。很快彩铃市场在中国达到了 80 亿元左右的市场规模。音乐人从彩铃市场分得的收入很快就超过了唱片和光盘的销售。

移动通信的加入，改变了音乐的销售与收入模式，这是大家没有想到的，但是它确实起到了这样的作用。彩铃在 2G 时代就可以支持，随着 3G 和 4G 的到来，中国的音乐销售模式被彻底改变了，今天我们已经基本听不到正版的呼吁了，不是不能盗版，音乐不能方便的免费传输，而是基于 4G 的网络能力，新的音乐销售模式被建立起来。

4G 的到来，随着智能手机的普及，年轻人这个音乐主要的消费群体，已经不需要下载音乐到 MP3 播放器这种音乐播放和存储设备中，智能手机已经代替了 MP3 播放器，最为重要的是，随着 4G

的普及，电信运营商资费也大幅度降低，听歌已经不需要很麻烦地把音乐下载到手机中，“下载”这个词已经开始被年轻人忘记了，需要听音乐，只要在客户端中方便地找到音乐，它有方便的分析，各种推荐，在线就可以听。用户根本不需要担心流量，也不会出现卡顿，播放质量不好的情况，因为 4G 的高带宽，年轻人已经不满足 MP3 的音乐，多声道、Hi-Res 音乐也受到更多追捧，48.0kHz、24bit、PCM\_S24LE 这些音乐也在手机端可以方便地在线播放。

差不多每一个青少年都在自己手机里安装了两个以上的音乐平台，成为音乐的包月会员，可以随时听自己想听的音乐。这些平台提供了包月会员、每首歌收费等各种收费模式，在线随时都可以听到自己想听音乐，可以最方便的交费，年轻人渐渐接受了这样一个模式。花时间去盗版传输，不如交很少的费用，自己不再去浪费时间。用金钱来换时间，用金钱来换服务，今天年轻人都越来越接受这个商业模式。

今天中国的在线音乐市场已经达到 600 亿规模，极大地改变了音乐销售的模式，而这个市场还在以较高速增长。

4G 到来时，对于电信运营商而言，音乐下载似乎并不是 4G 最大的特点，电信运营商介绍 4G 业务时，更愿意说高清视频下载，因为这才能把 4G 高速度的特点体会出来。很少会把音乐作为 4G 有代表性的业务，但是却是 4G 让音乐从下载音乐，转变为在线音乐，音乐的销售模式发生了根本性变化，看到了这样的机会，建立起音乐销售平台，当然就是抓住了 4G 的商业机会。其实离 4G 最近的电信运营商，并没有真正抓住这个机会，对于音乐平台的建设、

推动也不是特别的认真，很大程度上就是没有理解 4G 会带来商业模式的变迁。

4G 的商业模式变迁还体现在游戏业务上。我们都知道最早的电子游戏是单机版的电子游戏，电子游戏靠每一个游戏的销售来获得收入，游戏开发者卖的就是一个个拷贝，和音乐一样，游戏市场也同样面临盗版问题，大量的游戏玩家不去买正版游戏，只要用很便宜的价格买一个盗版光盘就可以，一个盗版 5 元人民币的盗版光盘可能会有 30 甚至上百个游戏。为了防止盗版，游戏企业不得不开发出各种认证办法，而这些认证无一例外都很会被破解，多种解码工具被开发出来，随着游戏一起做到盗版光盘中间去。

很长一段时间，游戏业一直在呼吁大家玩正版游戏，游戏市场的空间也只有 80 亿左右市场，游戏市场并不大。

虽然随着互联网的发展，游戏市场也在不断增长，但是真正改变电子游戏市场的还是 3G 尤其是 4G 的到来，4G 的出现，整个手机游戏市场出现了格局的变化，这个格局不是来自游戏业务的本身，而是游戏市场的商业模式革命性改变。

在智能手机和移动互联网全面发展之后，电子游戏的商业模式最大的改变是不再是单机版游戏，而基本上所有的游戏都变为网络游戏，这些游戏不需要销售拷贝，下载、安装、注册都是免费的，这样大大增加了用户群，用户也可以自由地体验，感受这些游戏。因为游戏的安装使用都是免费的，迅速发扩大了用户群，让更多的普通人成为用户，极大地推进了电子游戏的普及，尤其是通过社交的传播，很多游戏在较短时间，病毒一般地扩散开来。

电子游戏免费下载、安装、注册的模式扩大了用户群，它的收入不再靠销售拷贝，而是在玩游戏过程中，卖道具、卖装备、卖皮肤，这降低了用户抵触和反感，因为方便付费，用户个性化、战胜对手的心态的刺激保证了大量的游戏用户愿意付费，腾讯著名的《王者荣耀》创造一天亿元这样惊人的收入，而一个皮肤就卖了亿元。

中国的游戏产业正是因为移动互联网的出现，智能手机的出现发生了质的变化，不仅是智能手机让用户随时、随地都可能玩游戏，大量的等人、公交车上、短时间休息都可以来玩游戏，而社交帮助游戏可以病毒一样进行传播，移动互联网让游戏的推广模式和收费模式发生了较大变化，用户对收费从抵触转为较易接受。

今天中国的手机游戏产业已经达到 1800 亿产值，这个市场还在增长，如果没有 3G、4G 出现，移动互联网的加入，导致分发、传播、收费这些模式的变化，这个市场是不可能想象的。

无论是音乐还是游戏的市场，从一个较小的市场到爆发性增长，3G、4G 都在其中扮演了重要角色，但是无论是电信运营商还是产业界，在 3G 和 4G 推进时，都不太把音乐和游戏看作是较大的机会，因为这些业务早就存在，它也不一定需要特别大的带宽，似乎没有把大带宽的特点发挥出来，这不是最好的 3G、4G 的例子。但是正是移动互联网带来这个两个产业根本性的变化。这些变化最大特点不是业务本身，而是商业模式。

## 二、5G 时代商业模式可能会比业务冲击更大

5G 的重要机会也会来自商业模式，早可能在商业模式上形成 5G 机会的，是这些业务在 4G 已经有了，也是通过智能手机提供服务，也就是说在 4G 的发展中，早已经做好了足够的准备，仅仅是因为 4G 的带宽还是不够，比如对高清视频支持效果不好，到了 5G 随着效果的提升，资费的下降，很快会导致商业模式上的新变化，我相信这些业务会最早出现了短视频、视频社交、在线教育这些领域。

用视频社交一个例子看 5G 的商业模式变化。4G 时代社交也开始有了视频功能的雏形，但是绝大部分社交是还是文字与语音为主要的载体，虽然微信这种平台也提供社交的功能，但是因为 4G 的能力总体不足，无法支持用户大规模使用高速度网络，所以微信的视频能力都是做了限制，服务器端的支持也不够，我们使得 4G 网络用微信视频经常会卡顿，声音效果也不好，视频质量经常不够好。

随着 5G 的到来，因为速度更快，总体网络的资源大大提升，大量用户都使用高清视频成为可能，5G 时代，很快智能手机支持的视频都会提升为高清视频，今天在华为的旗舰手机上，我们就已经用上了“畅连通话”，这个能力就有强大的视频支持能力，视频都是高清视频，体验效果远远超过微信的视频能力，只要有了这样的视频支撑能力，在它下面建立更多的业务就有了可能与机会。

2020 年春天这场新型冠状病毒的疫情牵动了每一个中国人的心，也是对社会经济产生了巨大影响，如果我们已经有了强大的 5G 通信能力，有了高清视频支持，可以在这个基础上建立一个远

程诊疗系统，用户通过高清视频接入医院端诊疗系统，不需要一定去医院，其实今天我们很多的诊疗也不一定必须面对面的交流，体温这样的数据可以通过用户自行按医生指导来完成，高清视频保证了交流的畅通，也给了病人信任感和亲切感。这个强大的诊疗系统还可以在医生与病人进行交流时，病人的既往病史、过敏史等信息出现在医生的屏幕上，医生在与病人交流，了解基本情况，对照病人的检查数据，决定病人是否需要到医院治疗，或是就可以居家治疗，或是指派社区医生上门辅助治疗，比如打针之类。

我相信这样的一个系统如果逐渐形成，可以大大减少医院的压力，很多小病可以通过远程系统来解决。最近笔者体检血压略高，去医院看医生，第一次去医院经过复杂的挂号、取号、开卡程序，去看医生，医生不过量一次血压，看我自己最近一周量的血压数据，得出结论可能是高血压，开两周的药，吃完再看情况。两周后，再去复查，医生看吃药的效果不错，那就继续吃。每次看病用时不超过3分钟，这种常见的小毛病，如果通过远程诊疗系统，病人不需要跑医院，减少了很多感染的几率，医生也降低了受感染的可能，减轻了压力，甚至一般性的问题也可以远程随时沟通，帮助病人解决日常中遇到的健康问题。

这样的远程诊疗系统不仅可以防止交叉感染，可以大大降低社会成本，降低对医疗资源不足的压力，这也是分级诊疗、家庭门诊的一个部分，甚至会大大降低伤医案件。

当然，这样的远程诊疗系统，似乎在技术上并不一定需要5G，用WiFi、4G也是一样可以建构，一个技术最后能成为产品，

成为强大的服务，经常的情况不是技术能不能做，而是这个技术实现的成本是多少。如果通过 WiFi 就需要在每个家庭建立 WiFi 接入，还需要进行维护，因为这项工作只能由用户自己来做。事实上用户要做好这项工作，还是非常困难的。5G 的价值是整个系统是电信运营商已经做好了网络的维护，不需要用户去干预就保证了网络的畅通，而大家都会拥有智能终端，不但不需要用户去投入，而且支付、注册这些系统也基本都完善。5G 做到的是对网络速度的保证远超过 4G，它不需要用户做更多的投入，学习成本很低。

其实基于传统互联网和 3G、4G 都是可以远程医疗系统，但是要大规模普及，形成商业模式，5G 提供了最大的可能和机会。开发者一定要明白一件事，一个产品或是服务能被广泛接受，是多种因素共同形成的力量，渐渐做到水到渠成，能够实现这种业务，不是产品服务的机会，低成本、不需要学习和维护、能方便提供服务，甚至还能形成强大的传播和推广能力，这才是真正的机会。5G 就不仅是提供了能力，它还会低成本、免学习，这才是发展起来的机会。

### 三、精细化的计费将是电信运营商机会

5G 带来的机会不是业务本身，而是商业模式和业务模式，对于电信运营商更是如此。

很长期以来，电信业一直在讨论，不要沦为管道，电信运营商除了网络建设之外，希望能找到更多的机会，在移动互联网的业务发展过程中，电信运营商也希望在这个领域有所作为，应该说全

球电信运营商都做过努力，成效并不明显，相对于富有竞争力，决策机制简单的互联网公司，电信运营商显然不适合这种即时反应的领域。

5G 除了建设网络之外，会不会还有机会，我相信除了建设网络卖流量之外，5G 会给电信运营商带来更多的机会。

电信业这几十年有了高速度发展，之所以能取得较高的收入，除了网络之外，一个精细化的计费平台是它生存发展的根本。20年前，广电行业和电信行业相比，普通大众看电视的时间远多于打电话时间，电视覆盖的人群也远远超过电信业覆盖的人群，但是广电业获得的收入总体只有电信业的 1/10，一个重要原因，就是广电业从来没有建立起一个较为完善的收费平台，电视只是单向信息传输，对于用户完全没有管理，谁看了电视，不知道；看了什么节目，不知道；用户看了多长时间，不知道。一部电视剧，有没有用户看，只能靠调查。广电的主要收入模式就是广告，虽然央视黄金时间一秒钟广告达到了几十万元，但是总收入的水平很低，今天全中国最顶级的中央电视台一年的收入，不过是一个中型企业的收入水平。

除了电视台之外，广电业的有线电视虽然也是建立了一个收费平台，但是这个平台是极为粗放的，虽然已经有了较大的带宽接入，但是它还是一个单身的模式，为了适应这个模式，计费就是极为粗放的模式，一个机顶盒一户，每个月收一份钱，因为它面向所有用户的收费的标准是一样的，影响这个标准制定的就是全社会的

人，社会上收入水平低的人影响了标准制定，有线电视的收费标准就是面向低收入人群制定的。所以有线电视的收入也并不高。

电信业和广电业最大的区别就是在开始提供服务时，就建立了一个较为复杂的收费系统，谁是自己的用户，知道；用户用了多长时间，知道；用户在什么地方，知道。虽然用户量一开始并不如广电用户多，但是因为计费平台的精细化，可以在这个平台上形成计费，也形成了收入的差别。似乎表面上计费单元是一样的，打一分钟电话，收 0.40 元，但是因为通过了时长进行计费，一个经济水平很一般的用户，平时主要是接电话，并不常打电话，一个月可能只花费 20 元，他基本是可以接受。一个工作非常繁忙的商务人士，他需要经常打电话，一个月可能花费 800 元的电话费，但是对他而言，这通信费是生产成本，公司是可以报销的，对于企业而言，这个支出可以带来收入回报，能够承受的起。这个一个商务人士的收入就相对了 40 个普通用户的收入。

近年随着技术的发展，资费的下降，对于普通大众来说，通信的要求比看听广播看电视的要求更高，而且通信设备也可以看电视、听广播了，通信业的收入增长更是远远超过了广电。

因此，精细化的计费在电信业发展中扮演了极为重要角色，如果没有精细化的计费，也采用包月，一定会定出一个低价，电信业的收入就会腰斩，甚至还会减少更多。

随着电信业的竞争，电信业有放弃精细化计费的趋势，近年来，电信运营商搞不限流量套餐，就是在往这个方向发展，一个用户甚至一个家庭，包月，不限流量，这种情况的结果就是最初 199

元一月套餐，不限流量随使用，那么另一家运营商一定可以推出188元套餐，也可以不限流量随使用，然后就会有169元，129元，甚至88元套餐，这种情况在残酷的竞争中，电信运营商的收入只会绝对下降，而无任何增长可能，因为今天中国已经拥有14亿手机用户，这意味了基本上是每人一部手机，新增用户的空间很小，在现有的用户中，单个用户收入是下降的，这对于电信运营商的发展可谓是灭顶之灾。

不限流量的套餐，不但会让电信运营商计费粗放，收入下降，也会导致社会更加不公平。精细化的计费，就是使用得多更多付费，使用得少，较少付费。大量使用网络资源的商务用户，更多承担网络的成本。而较少使用的用户，较少的承担成本。包月不限流量的模式，是经济水平高，使用网络多的人，资费降下来了。而收入水平低，使用网络不多的人，费用要涨上来，承担更多的费用。这其实是不公平的。

粗放的计费模式也会让服务更加粗化，一些用户的要求无法满足。一个网络直播的播主，他需要在任何地方都是网络稳定的，在大量使用较为拥塞的情况下，他必须是畅通的，对于这种情况，他是可以多付一些费用。但是今天的网络大家都是一致的，收费一样，服务也是一样，要堵所有人都堵住，你想多付费也不行，这就导致了一些特殊的服务得不到满足。

4G的网络主要销售的产品主要是语音、短信、流量，产品较为单一，对于这些产品的计费模式也在变得越来越简单，这很大程度上造成了电信运营商收入无法形成增长空间。

5G 却不同，首先它的产品变得更加丰富，除了传统的语音、短信、流量之外，还会有更多的不同产品。即使是流量产品，也会有高速度流量、一般速度流量，除了速度之外，还有可靠性保证。计费单元从原来的时长、短信的条数和流量的多少，增加了速度、稳定性、可靠性等诸多的单元，这些单元被引到计费体系中，不仅计费更加复杂，同时在精细化的计费过程中，还会形成更多新的服务。这些精细化的服务，形成的能力就是电信运营商的新的收入空间。

仅仅是单一的流量，价格一定是越来越低，这个大势是无法改变的，但是 5G 的到来，运营商提供的不仅是流量，还会有稳定性、安全性、云服务、计费能力，它就不仅是卖流量，而会有更多的收入和发展空间。

我们都知道 5G 除了 NSA 这样一个路径，还有 SA 的路径，SA 走的路就是独立组网，要针对 5G 的业务和要求，不仅把基站改造成 5G 基站，可以支持 5G 的高速度，而且是要把核心网改造成 5G 核心，在用户管理、业务管理、计费体系上形成新的系统。通过这个新的系统，让网络管理、业务管理更加强大，计费更加精细。网络切片会成为 5G 网络的一个重要部分，一个无差别的公共网络，在网络上业务是无差别的，用户是差别的，会过渡为一个复杂的网络，这个网络会对不同业务进行智能化的管理化管理，也可以实现不同的安全级别，还可以针对不同用户不同的要求，提供不同的服务，当然这个网络还会提供不同的计费模式，甚至输出计费能力。这些带来的收入会远远超过卖流量的收入。

分享几个相关的案例：还是前文所述主播的例子，一个女主播，她希望在任何时候都有畅通的网络，而事实上在一些演唱会、体育场，人群聚集较多的情况下，所有人都在用这个网络，无论是4G还是5G都有可能出现拥堵的情况，在一个演唱会上她要做直播，出现了拥堵，她希望这时如果别人拥堵，因为他们只是一般用户，但是她作为主播，这是她的工作，她愿意为此付较高费用，比如每月1千元的费用，因为相对于她一个月5万元的直播打赏收入，这1千元的通信费是生产成本。在4G的时代，这个女主播的要求没法得到满足，即使是她愿意多付钱也不可能。

但是5G的智能网络管理和业务管理就可以把这个用户做成一个网络切片，把这个用户切出来，保证她一定网速，在别人网络出现拥堵时，她的网速可以得到保证，因为为她提供的专门服务，运营商也应该有更多的收入。这样的业务相信不仅是主播而且很多需要网络保证的用户都可能采用。

一个矿山，今天大部分还是需要由人深入地下进行采掘作业，把矿山装车，通过轨道或是传送带把矿山运送出来，因为地下条件非常复杂，每年事故不断，经常会造成伤亡事故，对于产业影响非常大，现在人力损失也需要很大赔付成本。大家都在探索远程进行管理控制的模式。

但是对于矿井下面，以往是不需要建设通信网络的，在井下建设通信网络，在技术上还需要一些需要解决的问题，如安全问题，井下是不允许电火花，需要使用泄漏电缆，还需要建立独立计算、数据处理系统，用边缘计费的能力，来实现低时延。这在很大程度

是构建一个小的局域网，通过个网络系统提供低时延高可靠的网络，这个网络可以远程采掘业务，支持井下运输管理，这样必须要通过 5G 来建立新的系统，当然也会有更高的成本支出，这个系统中流量其实要求并不那高，运营商也不可能向矿山卖流量，销售的是一个基于 5G 的矿井采掘与服务系统，同时还要负责这个系统的管理与运营，它应该获得的收入也应该更高。

除此之外，电信运营商还可能输出自己的计费能力。电信运营商已经建立起一个庞大的、复杂的计费系统，这个计费体系个人或是小的互联网平台没有能力建设，未来的智能互联网业务绝大部分是需要收费的，电信运营商可以把计费、管理的能力提供给合作伙伴成为它的计费、收费、管理平台。一个简单的例子，2019 年，我曾经有一个讲什么是 5G 的视频在互联网上流传非常广泛，在各个平台都有很多传播，全网播放上亿次，有非常大的影响，但是并没有给我带来任何经济收入，因为这都是免费的。如果这个视频电信运营商提供了管理与计费能力，视频到一半时，愿意再看下去的用户需要付 1 元的费用，如果还有一半人愿意再看，方便地通过运营商平台付费，这一个视频就有可能获得 5000 万的收入，我相信这无论对于视频制作者还是电信运营商都会有很好的经济的回报。

总之，一直卖流量，把计费的搞得越来越简单，电信运营商的收入一定没有办法增长，因为用户已经达到极限，再有用户的增长的空间已经很小，要在现在的用户中增长收入，就是需要提供更多的有价值的服务，这种服务就是需要针对用户的个性化来提供，精细化的计费和管理，是实现个性化的重要一步。运营商能不能也参

加去做业务，不是不可能，但是电信运营商要和小的互联网公司竞争，它的灵活性、成本都很难实现高效果。只有通过一个较为庞大的体系，提供更多用户使用，而一个精细化的计费体系，可以根据用户的需要提供服务，这才是未来的一个发展机会。

## 四、关于 6G 的思考

### 1、6G 的定位是 5G 的延伸与完善而不是一次技术革命

每一代通信技术的发展都有其使命，我们以移动通信来看，第一代移动通信是建立起固定通信到移动通信的体系；2G 建立了数字通信能力；3G 建立了数据通信能力。这都是具有革命意义的技术突破，也需要较长的时间，需要在技术做较多的准备。4G 基本上就 3G 的延伸的完善，需要解决的技术问题不多，甚至业务也是 3G 基础上的推进，部署和发展很快，但是因为极大地提升了 3G 的体验，4G 时代是把 3G 要解决的问题圆满地解决了，用户体验有极大提升，用户接受度和认同度非常高，达到了极好的市场效果。

5G 不仅在速度上提升，同时把低时延、低功耗、高可靠的愿景引入移动通信中，打破只是提升速度的旧模式，为工业互联网、无人驾驶、智慧城市的需求提供了可能。这也是技术的革命，5G 要全面使用产生良好效果，可能面对困难远远超过 4G，在 5G 网络部署完成之后，终端、业务、商业模式的完善，要走漫长的道路。

6G 很大程度上和 4G 一样是 5G 延伸和完善，而很难做革命性的技术突破，在愿景上，也很难完全打破 5G 的愿景，上一个更高

的台阶。理解好 6G 的定位，对于我们提出 6G 愿景和进行战略规划具有很重要指导价值。

## 2、速度将不再是 6G 关注的重点

对于 6G 的愿景，很多专家把更高的速度作为一个重要方向，提出了 300M/s,甚至理论网速 1TB/s 的 6G 理论网速。

更高的网速，对于业务体验的提升确实有非常大的价值，人类对于网络速度提升的要求是无止境的，但是在 5G 网速已经达到一个较高的情况，是否有业务应用需要把 5G 网络速度提升 10 倍甚至 100 倍，业务前景并不明朗。

但是随着无人驾驶、工业控制、社会管理等相关业务的发展，这些业务未必追求的是不断提升传输速度，但是对于时延、可靠性提出了非常高的要求，也对网络的智能化提出了更高要求，安全性也是未来的网络的一个极大问题。

从这个意义上，我们对于 6G 技术的思考，从提升速度这一个维度，要转向多个维度，在 6G 这个阶段，可能提升速度这个维度并不是最重要的，追求 10 倍速度的提升并不是最重要的方向，在这个层面更多下功夫，它带来价值对于业务能力提升帮助不大，而付出的成本过高。

6G 我们应该更加关注低时延、高可靠的广域覆盖的可能性。

## 3、6G 频谱有可能不会大规模采用太赫兹

6G 采用太赫兹是较为广泛的声音，这是基于 6G 需要大量带宽来支持高速率为基础的，但是 6G 如果不以大幅度提升速率为目标，是不是要采用太赫兹就存在较大疑问。

目前 5G 采用的频段主要是 Sub6，毫米波有几十 GHz 的频段，在十年左右的时间内，能否利用好毫米波，还是一个较大的问题，太赫兹作为一个研究方向当然是可以研究，但是期望在 10 年内，成为新一代通信技术的频段，有可能存在巨大风险。

我相信太赫兹作为一个研究方向非常有价值，但是作为下一代移动通信的规划，把太赫兹作为重要方向存在较大风险。6G 的主要频段应该以毫米波为主。

十年之内，把太赫兹频段真正变为移动通信技术，基础研发、技术积累还存在太多问题没有解决，在这个时间段实现技术突破，达到可以商用状态，存在巨大不确定性。

#### **4、天地一体化进行广域覆盖是 6G 重要愿景**

5G 基本上解决了速度、低时延、低功耗这些智能互联网的核心问题，但是广域的覆盖问题并没有根本解决，在偏远地区、高山、海洋中无法实现覆盖。而智能互联网发展，却需要在这些地方进行广域覆盖，才能让一些以前无法做到的业务与应用成为可能，如高山上的环境监测、远洋航行的监控、深海养殖的管理，还有可能会在平流层出现的飞行器的管理，这些覆盖目前 5G 基站的模式没法做到。

采用低轨道卫星和基站通信网络整合起来，用低轨道卫星对海洋、高原、偏远地区进行覆盖，用地面基站对城市地区进行覆盖，形成一个体化整合的体系。

加快低轨卫星的建设与发展，将其与移动通信网络进行整合，在 6G 时代形成广域的全面覆盖。

## 5、重新构建互联网底层架构

传统的互联网是为信息传输而建立，追求信息的自由畅达，但是安全性、管理能力、控制能力、多设备接入能力都不够。传统互联网的架构由美国为主导的非政府机构控制，传统互联网的架构已经无法适应智能互联网发展，但是针对智能互联网进行底层架构的建构进展极慢。

互联网的底层架构要面向智能互联网进行重新改造，这应该成为 6G 的一个部分，6G 不仅需要速度，还需要安全，需在适应大接入量，需要有针对性适应不同领域的要求。

通过 6G 标准的建立，从底层构建互联网的架构，将互联网从信息传输到生活服务，向社会管理、工业控制发展，这些应用提供安全、高效、节能的底层。

在 6G 标准的框架里建立起一个新的互联网架构，这对于互联网发展意义重大，也到应该重新建构新体系的时候了。

# 网约车和自动驾驶

# 采用基于开源的深度信息技术推动交通出行变革 ——滴滴出行的应用实践

张博 章文嵩 吴国斌 脱立恒 金曦

我国拥有全世界复杂、庞大、多元的出行市场。我国拥有 8 亿城镇人口，以约 70% 的成年劳动人口计，平均每天大概有 11 亿人次的出行需求，我国的城市人口密度和大城市数量也远远超过其他国家。城市拥堵、汽车利用率低等问题严重影响了出行的质量与效率，庞大的人口规模、差异化的社会经济条件以及多样化的出行需求，让出行预测、规划和实时派单的难度呈几何级增加。如何高效地满足人们的出行需求成为世界级难题。

滴滴出行的使命是让出行更美好。自 2012 年成立以来，滴滴始终坚持技术驱动，不断加大对基于开源的深度信息技术（大数据、云计算、人工智能等）的研发投入，利用科技创新力量推动使命的达成，积极投身到我国交通强国建设中。经过几年的积累，滴滴在交通出行领域的技术、产品、业务创新方面走到了世界前沿。

滴滴的科技战略是“人工智能改变交通”（AI for Transportation）。滴滴认为交通的变革主要体现在三个层面：最下面一层是交通的基础设施，包括红绿灯、道路资源等等，滴滴已经与交管部门紧密合作，用滴滴拥有的交通大数据进行组织优化；中间一层是交通工具本身的变革，未来十年智能化、电动化会越来越

普及；最上面一层则是共享出行的网络，未来的车会越来越的共享而不是被私人拥有。

## 一、共享出行层变革：共享平台的智能预测与派单关键技术达到世界领先水平

滴滴的共享出行平台将司机和乘客通过网络链接起来，为乘客提供全方面服务，满足乘客多样的出行需求和高质量的服务体验，同时提高司机收入、减少空驶、提高交通运行效率。

平台服务包括发单、派单、接驾、送驾四个主要流程。发单时，利用预计时间到达进行价格估算，同时实时预测用户目的地，推荐用户上车点。派单时，利用深度强化学习进行派单。接驾时，通过ETA和路径规划，快速接驾。送驾时，通过多项措施保证送驾安全。同时，平台通过供需预测与调控，预测供需，调节运力，提升平台整体效率。

在平台系统结构中，预估时间到达、供需预测、智能派单等服务位于系统的服务层，组成了滴滴共享出行平台的核心。预估时间到达为供需预测提供时空特征支撑，为智能派单提供决策支撑；供需预测为智能派单提供策略迭代；智能派单为供需预测提供供需优化。

### 1、预计到达时间估计技术

预计到达时间（Estimated Time of Arriva, ETA）是共享出行平台一项必不可少的基础服务，无论是行程前的预估接驾时间、预

估价格显示，还是派单、调度、拼车等系统决策，离不开高精度 ETA 服务。

对行程时间的预估受到路径规划选择、道路状况、行车速度、司机驾驶行为习惯等多重因素影响。在深度学习出现之前，基于道路物理结构的规则模型通过对每一个路段和路口的时间预估累加进行预测，准确率很难提高。

滴滴 2017 年首次提出基于深度学习模型的预计到达时间方法。通过将深度学习模型引入 ETA 计算，模型每日处理规模特征达亿级，准确率显著提升。目前，滴滴每天有数百亿次的 ETA 请求，滴滴 ETA 误差率已经从出行行业 16%降低至 10.5%。

## 2、供需预测技术

供需预测是指根据历史一段时间内的供需数据准确地预测未来一段或者多段时间内的供需变化趋势。其中，供需数据是城市中某区域某时段内网约车供给与订单需求量的量化表征。供需预测本质是时空数据预测问题。时空预测主要有两个挑战，第一，时空数据的表示通常是结构化的，在表示数据的同时需要表示区域间、时段间的复杂关系；第二，基于这种结构化的表示，需要提供相应的机器学习模型进行时空特征抽取，进行精准预测。

滴滴供需预测模型通过挖掘城市区域间的多种隐式关系，提出将区域间关系用图结构表示，并建立多个图卷积神经网络进行空间特征提取。并基于供需问题的时间序列特性，构建循环神经网络进行时间特征提取。最后，对多个模型进行融合。该模型帮助平台预

测任意时段、各区域的订单需求和供给分布情况。目前，滴滴对 15 分钟后的需求预测准确率已达 85%。

### 3、智能派单技术

派单是把乘客发出的订单分配给在线司机的过程。该问题主要面临的挑战有供需和路况存在高不确定性；待匹配司乘规模大，以一城市为例，每秒钟参与计算的司机、订单数量可达到千级别；所有的派单决策需要在秒级做出；并且要考虑司乘体验与司乘长期收益等多个目标。

滴滴提出了融合强化学习和组合优化的智能派单系统，通过智能算法决定司机、乘客最优匹配，撮合司乘交易，实现智能派单。滴滴从抢单演进到派单模式，乘客的应答率有了 20 个点以上的提升。目前，滴滴每天为乘客匹配千万量级的出行订单。在高峰期时，滴滴每分钟要处理超过 6 万的叫车订单，每 2 秒计算就要匹配成百上千的乘客和司机。滴滴对车辆的智能分配模式每天能够帮助平台上的司机与乘客节省共 30 万小时的出行时间。

经过中国人工智能学会组织的科技成果鉴定会鉴定，滴滴共享出行平台智能预测与派单的关键技术在面向交通出行行业和复杂交通环境下的新一代人工智能技术研究达到国际领先水平。相关技术获得了多项国际学术奖项。

## 二、基础设施层变革：基于浮动车数据的智慧信号控制系统在多地落地应用

传统的交通信号灯使用默认时间转换灯色，虽然转换灯色的时间会根据数据每几年更新一次，但是随着交通模式发展，传统系统很快就会过时。而滴滴人工智能驱动的智慧信号控制系统则利用滴滴实时浮动车数据，以动线可追溯和预判手段，结合政务雷达传感器和摄像头，监控交通状况，然后利用先进的滴滴人工智能算法决定灯色转换时间，通过人工智能和交通控制理论融合应用，优化了城市道路网络中的交通流量。

## 1、滴滴浮动车数据介绍

在出行场景中由于车辆导航、付费等多方面应用需求产生了大量的车辆运行过程中的轨迹数据，其数据中包含了 GPS 经纬度、方向、速度、时间等基础信息，同时更新频率不超过 3 秒。为了让数据更加精准可靠，在实际应用中滴滴重点采用载客状态下的轨迹数据，最大程度保障了驾驶目的的单一性（最快将乘客送往目的地）。另外，载客状态下的轨迹数据跟交易价格强关联，因此对轨迹质量的要求也更高。通过着重采用载客状态下的数据，保障高质量的数据输入。

基于轨迹数据，滴滴不仅可以掌握车辆出行的 OD 信息，还能回溯每条轨迹整个行驶过程中的路线、行驶速度、途经的每个信控路口转向级延误信息，停车次数、停车位置等信息。基于每个路口，滴滴可以提取在固定时间段内通过路口的全部轨迹信息，并进行相应集聚，这样便可以得到路口所有转向级延误指标、停车次数，停车位置等信息。根据各流向的基础信息，可计算得到过饱和、失衡、溢

流等报警信息，结合当前时段路口运行信号灯配时方案及实际道路交通组织，可进一步计算信号灯优化分析结果及方案建议。

## 2、滴滴浮动车数据在信号优化服务中的应用

传统的信号优化服务存在信号灯配时问题难以发现、交通流特征难以判断、优化效果难以评估、交通报告素材不够丰富、路段速度难以采集、信号优化工程师水平经验不一致等问题，需要通过更多的人工、更复杂的工作方式来缓解。通过滴滴浮动车数据，可以更加有效的解决上述问题，并且提供更多的赋能。

### 应用一：发现信号灯配时问题

通过滴滴浮动车的轨迹数据，可以获得路口之间的速度、路口各进口道的转向级延误数据、路口各进口道的停车次数和停车位置数据。从而可以进一步精准定位路口和信号灯配时相关的问题，并且及时进行拥堵报警。

### 应用二：判断交通流特征

通过滴滴的浮动车数据，可以有效了解路口各个转向的全天延误情况，并且能够有效的区分工作日和节假日之间的交通特征。并可以根据全天不同转向的拥堵情况，精细化调整路口方案控制的固定时段，还可以追溯到特殊日期的交通流特征。为节日管控、大型活动管控等特殊场景提供有效的数据支撑。

### 应用三：精准评估优化效果

滴滴的浮动车数据可以得到量化的路口指标，不仅能够直接获取到延误、停车次数等优化路口的前后对比数据，更可以通过干线

评价（速度、停车次数、延误）、区域评价（速度、延误）等多维度指标来整体评估优化前后效果以及对周边路口/段的影响。

#### **应用四：报告自动生成，数据维度丰富**

在优化服务的过程中，很重要的一部分工作就是需要产出分析报告。滴滴信控平台提供了多类型报告生产的功能，方便在不同业务场景中使用。同时，报告可以针对不同的需求和侧重点，选取想要分析对比的数据模块，在单点、干线、区域等不同纬度的报告中，可以针对运行状态、运行情况、运行指标等多角度进行分析。同时，在分析中，可以选取日报、周报、月报等不同的时间跨度，满足更复杂的应用场景。

#### **应用五：提供准确的路段行驶速度**

在协调优化的设计过程中，很重要的一个参数就是路段上的行驶速度。行驶速度如果计算不准确的话，会直接影响到协调方案的实施效果。通过滴滴的浮动车数据，可以多样本，全时段计算路段之间的行驶速度，并且可以结合滴滴地图的路段长度数据、信号机背景方案，直接生成不同时段下的绿波方案时距图（在计算数据中，已经充分考虑了起步、停车等减速场景），方便绿波方案设计。

#### **应用六：提供标准化工具**

滴滴通过浮动车数据，自主研发了一套以浮动车数据作为主要数据输入的信号控制平台，平台不仅能够依据不同地市的管控习惯、车流特征去设置自定义参数，也能针对不同业务场景下的需求，自动生成优化方案、时段方案以及专业报告，弥补人员作业流程不规范，实施参差不齐等问题。

### 3、滴滴智慧信号控制系统的落地应用

目前，滴滴智慧交通项目已在北京、济南、深圳、苏州、南京、柳州等国内外 30 多个城市落地，优化信号灯超过 2500 个，平均拥堵缓解 10%-20%。

部分应用效果如下：

济南应用实践：应用并优化了济南 349 套信号灯，经实测效果检验，济南经十路山大路到舜耕路路段，工作日早高峰平均延误时间下降 10.73%，晚高峰平均延误时间下降 10.94%。在停车次数方面，经十路的数据也有明显下降，工作日晚高峰停车次数下降 8.7%，早高峰下降 6.7%。

武汉应用实践：应用于武汉 170 套信号灯，经实测效果检验，武汉江发路园博大道段到发展一路段、京汉大道前进一路段到义和巷段的干线道路延迟指数有大幅下降。江发路早高峰各路口平均通行时间降低 13.3%，各路口平均延误时间降低 29.7%，干线方向整体延误时间降低 23.4%；京汉大道平峰期各路口平均通行时间下降 4.9%，各路口平均延误时间降低 8.5%，干线方向整体延误时间降低 5.8%。

柳州应用实践：应用于柳州 218 个路口的信号灯，通过轨迹数据的实时检测，以分钟级对信号灯进行动态优化。通过该试点，柳州在不依赖渠化及路口改造前提下，仅通过信控优化及软件升级，全城拥堵减缓 7.4%，车均停车次数下降 4%，干线延误及停车次数下降达 5%-25%，过饱和及多次停车比例下降 37.5%。

### 三、交通工具层变革：滴滴自动驾驶进入载人示范应用阶段

滴滴从 2016 年开始布局自动驾驶，致力于打造世界领先的 L4 级及以上自动驾驶技术，希望通过科技让出行更安全、更高效。2019 年 8 月，滴滴宣布将自动驾驶部门升级为独立公司，专注于自动驾驶研发、产品应用及相关业务拓展。

#### 1、滴滴自动驾驶发展现状

滴滴目前已具备感知、预测、规划、控制、定位、高精地图、仿真系统等完备的自动驾驶研发模块，团队规模超 500 人，在全球申请自动驾驶相关专利超过 600 件。滴滴已相继在北京、上海、合肥、苏州、美国加州等地获得路测资格。

2020 年 6 月 27 日，滴滴在上海率先开展自动驾驶载人示范应用项目，用户可以通过滴滴 APP 报名申请在测试区域内进行自动驾驶车辆的试乘体验。这也标志着自动驾驶载人示范将进入规模化阶段。截止 2021 年 4 月，已经累计有 5 万多用户报名体验，滴滴自动驾驶公司的实际路测数据也达到了数百万公里。

#### 2、滴滴自动驾驶的独特 DNA 优势

滴滴自动驾驶具备独特的 DNA 优势，D 即 Data（大数据）、N 即 Network（运营网络）、A 即 AI（人工智能技术）。

在数据方面，滴滴是目前世界上唯一拥有千亿级公里出行场景数据的公司。除了高精度地图以外，真正做自动驾驶需要一套完整的体系，将战术地图和高精度地图整合到一起。基于滴滴大数据，对于自动驾驶产生三张地图：经营地图，精准锁定高经营价值区域；

安全地图，筛查评估道路安全事故；技术地图，圈定技术可行性场景。这三张地图的整合形成战术地图，来判断最合适的驾驶场景。庞大的数据量能帮助滴滴提升运营区域洞察和落地的精准度与效率，加速自动驾驶应用部署。

在运营网络方面，滴滴能为乘客提供自动驾驶和普通网约车“混合派单”服务，一方面弥补现阶段自动驾驶仅能在特定区域使用的局限，另一方面也能让自动驾驶不仅仅停留在新奇体验阶段，更能成为人们日常出行的切实补充。此外，滴滴八年多以来构建了覆盖全国的车队运营网络。滴滴全球领先的网约车运营网络，还能自动驾驶提供全流程车队运维、管理、充电、保险等。

在人工智能技术层面，滴滴自动驾驶技术致力于 L4 级及以上系统解决方案。目前积累了以 AI 技术为核心的 L4 级全栈自动驾驶方案，包括感知、预测、规划、控制、定位、高精地图、仿真系统等完备的研发模块。滴滴自主研发的“桔视”车载摄像设备，覆盖了滴滴 50% 以上的网约车订单和 300 多个城市。“桔视”设备收集的长尾场景数据，将帮助滴滴搭建更丰富的仿真测试场景库，推动自动驾驶算法不断迭代进化。

实现自动驾驶是一个复杂的系统工程，需要精准感知环境，理解交通参与者的意图并能够在广泛丰富的场景下，实现稳定安全的无人驾驶。基于独特的 DNA 优势，滴滴自动驾驶将继续迎接复杂的自动驾驶场景、环境的不确定性的挑战，探索更多落地应用，推动全球汽车与交通产业的变革。

## 四、滴滴国际化

依托于大数据、云计算、人工智能等基于开源的深度信息技术，滴滴为全球超过 5.5 亿用户提供出租车、快车、专车、豪华车、公交、代驾、企业级、共享单车、共享电单车、汽车服务、外卖、支付等多元化的服务，年运送乘客超过 100 亿人次。同时，滴滴与 Grab、Lyft、Ola、Uber、99、Taxify、Careem 等全球七大领先的出行企业构建了一个全球移动出行服务网络，服务全球超过 80% 的人口。

自 2015 年出海以来，滴滴积极响应国家号召，投身参与“一带一路”建设，沿着拉美的旅途，一路向北，相继进军巴西、墨西哥、智利和哥伦比亚等国家。在滴滴国际化战略中，拉美是重点区域。未来，滴滴将在拉美地区聚焦多元化出行、电动车运营服务、智慧交通等方向。其中，在电动车运营服务方面，2020 年 1 月，滴滴在墨西哥全国范围内开始运营电动和混合动力汽车。滴滴成为拉美第一家线上运营新能源车队的网约车平台。

截至目前，滴滴国际化业务已经发展至海外 14 个国家，包括澳大利亚、新西兰、日本、巴西、墨西哥、智利、哥伦比亚、秘鲁、哥斯达黎加、巴拿马、俄罗斯、多米尼加、阿根廷和南非。滴滴同时正全力推进国际化外卖和创新业务发展，为全球更多用户提供本地化服务。

滴滴平台走出国门，为更多的人们提供了极大的便捷出行与生活方式，带动了移动互联网业务的繁荣，实现共享出行等产品的规模化应用，带来了巨大的经济和社会效益。滴滴将持续积极响应“一

带一路”倡议，将构建的移动出行和生活平台解决方案和技术，输出到海外，赋能全球交通出行，并为当地经济社会的长远发展作出积极贡献。

采用基于开源的深度信息技术，滴滴以科技改变交通出行为己任，积极开展智慧交通、自动驾驶、车路协同等布局。滴滴在拥抱开源的同时，也积极回馈开源。滴滴开源项目总数已达到 63 个，国内企业开源贡献排名第 5，全球开源贡献排名进入前 100，使用滴滴技术的组织和企业超过 5000 家。滴滴将持续打造内源外源双循环的技术开放生态。对内，通过激励内部开源、完善内源社区的方式，激发科技创新活力、提升成果转化效率；对外，通过开源项目、与开源基金会合作等，使滴滴内部的优秀开源项目帮助到更多开发者，回馈开源社区并创造更多社会价值，以持续推动交通出行领域的变革。

# 边缘计算

# 在边缘计算中的创新与实践

九州云

## 一、边缘计算发展综述

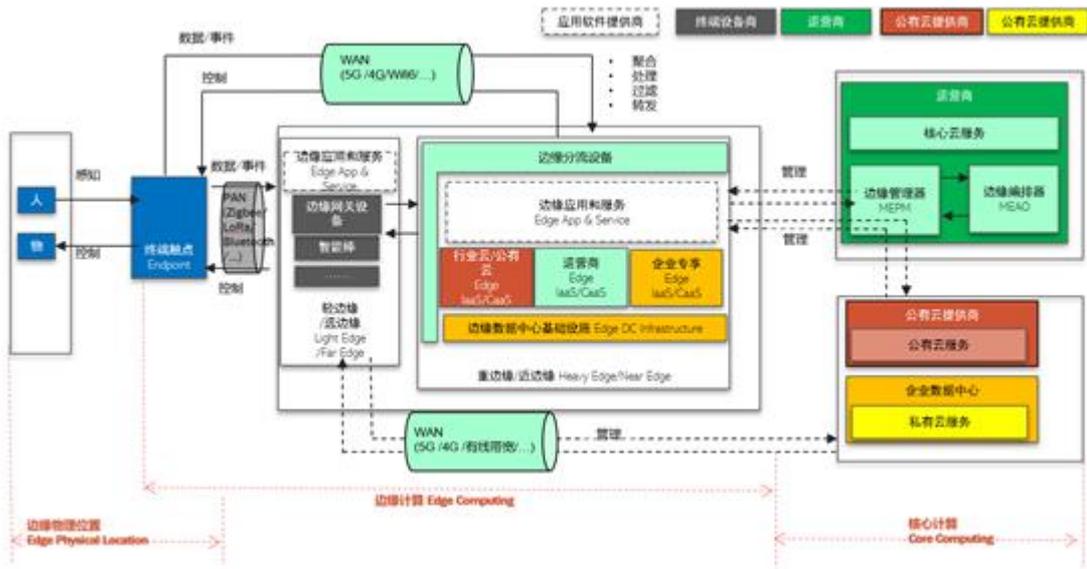
边缘计算随着 5G、IOT、AI 等相关技术演进越来越引人瞩目。边缘计算的第一次提出是在 2014 年,高通工程总裁 Karim Arabi 在 IEEE DAC 2014 Keynote 提出边缘计算为云之外的在网络的边缘侧的所有计算,2015 年 8 月,ETSI 第一次提出了 MEC 这一概念,将边缘计算从 IOE 的视角扩展到 ICT 的视角。ETSI 定义边缘计算为“为应用开发者和内容提供商提供在(运营商)的网络边缘侧的云计算能力和 IT 服务,这一环境的特点是极低的延时和极大的带宽,支持针对应用侧无线网络的实时访问”,2017 年 3 月,Gartner 提升了对 Edge Computing 这一领域的重视,Gartner VP Thomas J. Bittman 发表博客,认为“边缘将吃掉云”。在 2018 年,从云到边缘是被列为十大战略性科技发展趋势之一。2019 年,边缘计算再一次登上 Gartner 的趋势榜单。2018 年麦肯锡《新需求、新市场:边缘计算对硬件公司意味着什么》的报告中,指出边缘计算领域,在智慧交通、交叉市场、先进工业、零售、新媒体、公共服务和能源领域,到 2025 年,将会有潜在 1750 亿到 2150 亿的市场规模。2020 年 KPMG 也在 2020

年发表报告预测：“全球到 2023 年，随着 5G 和边缘计算的更广泛应用，这五个目标行业将在涵盖连接性、硬件、软件和服务的整个生态体系中创造超过 5,000 亿美元的年收入”。

## 二、边缘计算技术体系综述

针对以上不同来源的边缘计算定义，各家对边缘计算的定义虽然有所差异，但基本强调新边缘管道和分布式部署方式产生的对业务形态的重塑。在这里我们对边缘计算的要素进行总结，主要的重点包括以下三点：

- 边缘计算必须包含终端触点（Endpoint）的方式，实现物理世界到数字世界的双向翻译（数据 / 事件 / 控制信令），并通过网络管道实现一定规模的终端和空间的覆盖，实现人物交互 / 物物交互，单一具备计算能力的终端（如具备自处理能力的摄像头）如果未通过网络管道接入一个更广泛的数字世界，不适合归入边缘计算范畴
- 边缘计算的数字计算能力，可以分为轻边缘（远边缘）和重边缘（近边缘）
- 边缘计算的网络管道能力，可以分为 WAN（5G/LTE/WiFi6 等）和 PAN（Zigbee/LoRa/Bluetooth 等）

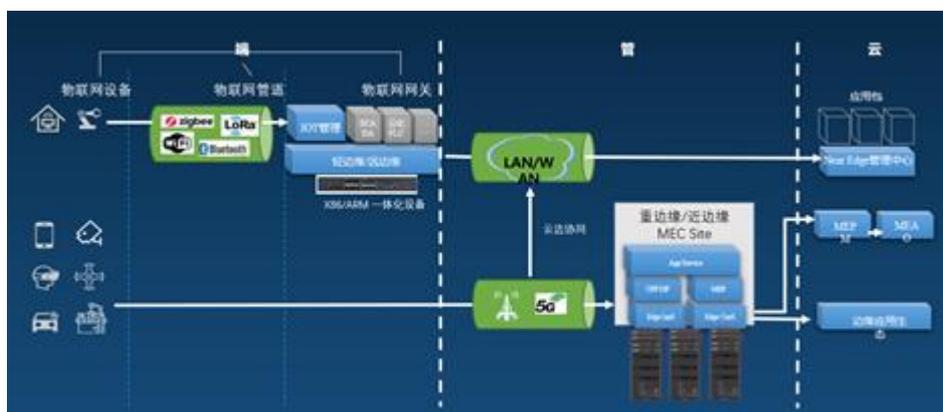


基于 5G 的 MEC 边缘和基于 IoT 设备的边缘一体机是两种比较常见的边缘形态,以 5G 为代表的 WAN 管道的边缘计算,相对以 LoRa、NB-IOT、Zigbee、WiFi6 等 PAN 管道,在大带宽、控制/内容传输兼具、室外 (Outdoor) 空间扩展性、移动性、SLA 保障和计算延展性上具备更大优势。WAN 管道和 PAN 管道的不同,形成了重边缘和轻边缘两种形态。

	重边缘 (5G MEC)	IoT 一体机轻边缘
部署位置	通信基站后面	通信基站前面
放置位置	通信机房	以挂壁等形式放置于产线/设备旁边
通信协议	WAN (5G)	PAN (Zigbee/LoRa/Bluetooth)

通信接口	IP 网络接口	比较丰富的工业物理接口，如 RS232/RS485/RS422/USB/CAN 等
计算能力	以数据中心服务器组件的集群，一般具备数百个 CPU 物理核、数百 TB 级别内存、PB 级别存储	以台式机服务器的单体，一般具备 4-8 个物理核 CPU，32~64GB 内存，1-2TB 存储
GPU 处理能力	配置相应的 GPU 处理集群	受限于一体机插槽，一般不提供
空间延展性	覆盖空间大 (Wide Area)	覆盖空间小 (Local Area)

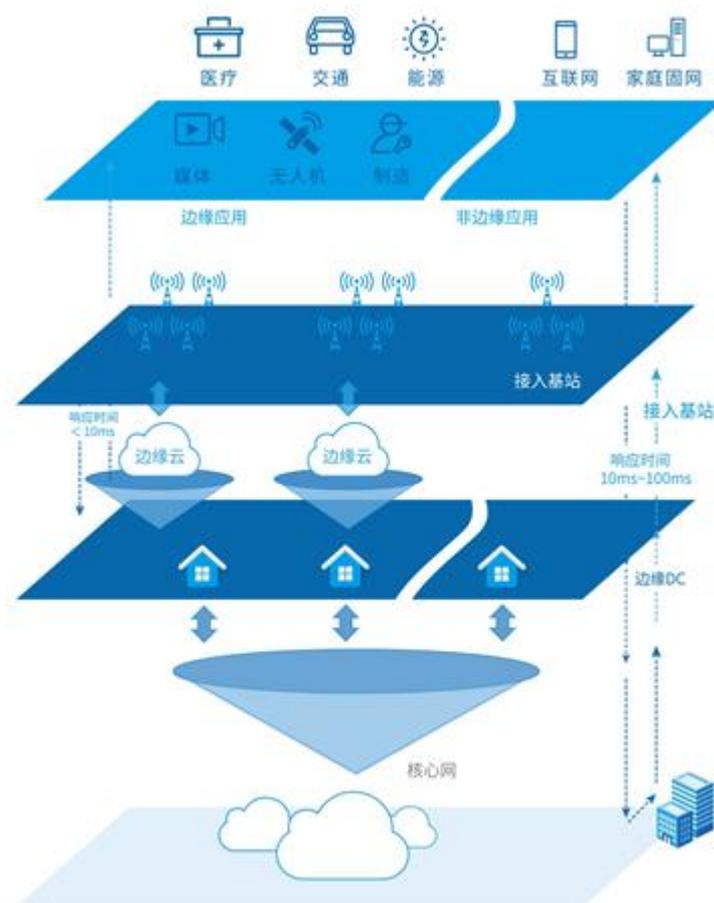
区分重边缘和轻边缘两种形态的另外一个因素是放置位置，重边缘一般和 5G 结合，部署在基站后面，而轻边缘一般部署在基站前面，下沉物联网设备或者工业产线现场。



总的来说,重边缘对 5G 管道等刚性要求比较高,建设成本更大,但是带来更好的空间延展性、计算能力和网络带宽。和轻边缘比较,5G MEC 重边缘计算在大带宽和低延时并存的场景中具备更广阔的前景,比如制造、港口、文娱和园区类场景,5G MEC 重边缘能够提供更好的数据上下行所需的带宽能力、出局处理所需的计算能力、指令控制所需的低延时和安全所需的数据私密性。

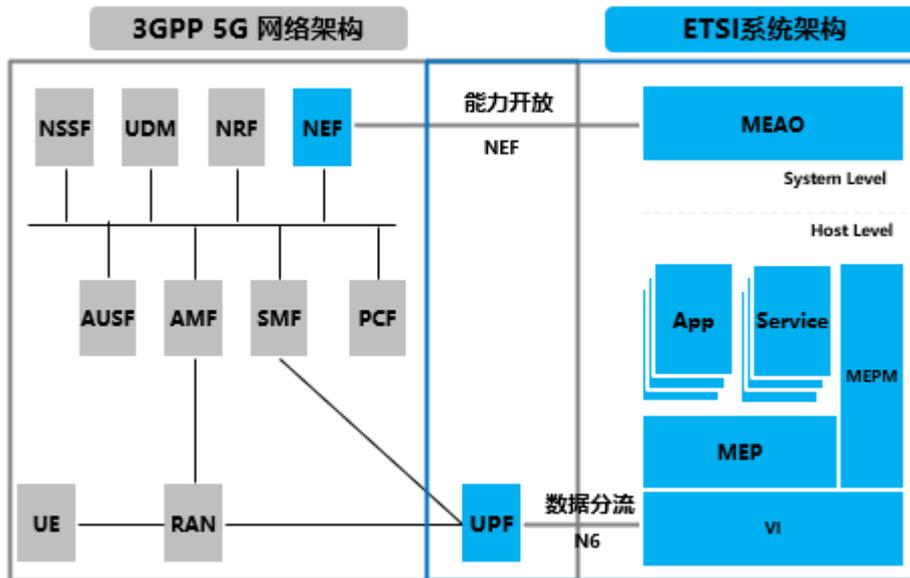
行业	设备类型	典型规模	单路带宽	总带宽	延时
制造	工业相机	1000	5Mbps	5Gbps	<20ms
	工业 PAD	1000	5Mbps	5Gbps	<30ms
	AGV 小车/RGV 小车	50	200Mbps	10Gbps	<20ms
港口	塔吊、龙门吊、轮胎吊	100	60Mbps	6Gbps	<10ms
	高清摄像头	1000	5Mbps	5Gbps	<30ms
文娱	AR	5000	2Mbps	10Gbps	<10ms
	本地互动游戏	400	50Mbps	20Gbps	<20ms

	高清直播视频	100	100Mbps	10Gbps	<20ms
园区	园区办公	100	10Mbps	1Gps	<30ms



### 三、九州云在边缘计算中的创新

九州云的边缘计算发展主要集中在5G 边缘计算领域(重边缘)领域，九州云的设计理念是“源于标准，高于标准”。在5G 边缘计算领域相关的标准主要是来自3GPP 和 ETSI。



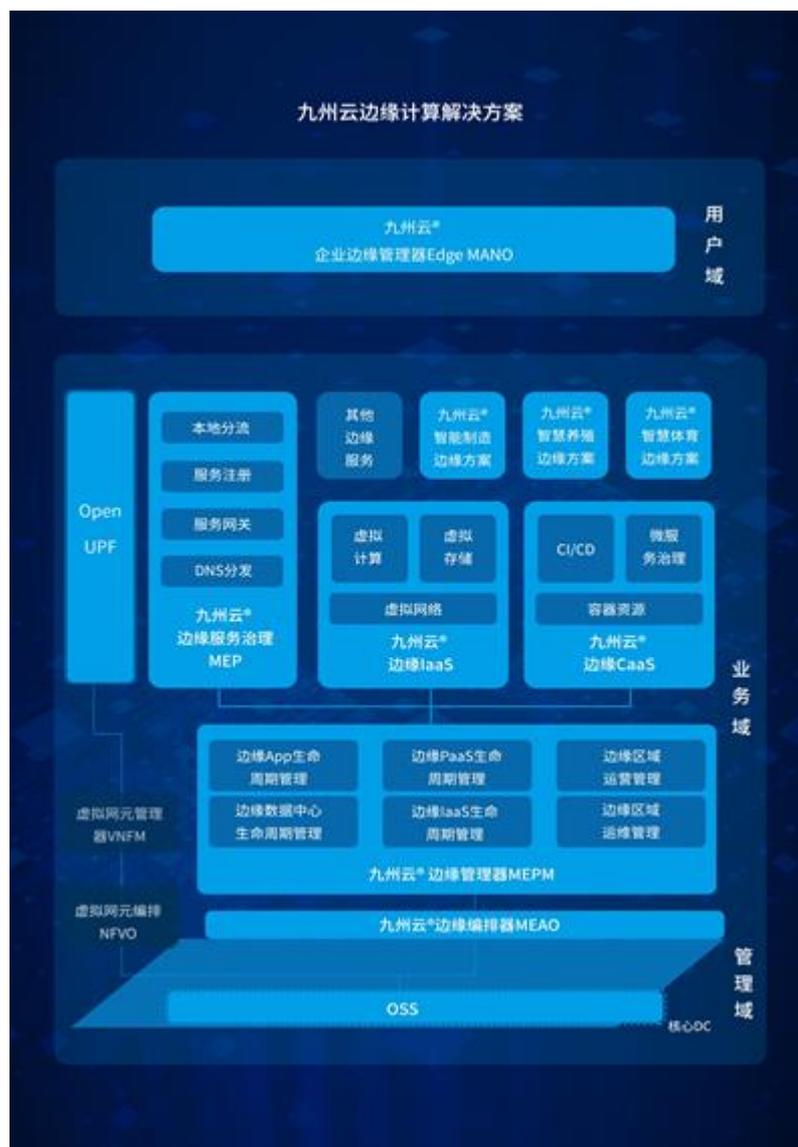
3GPP 定义了 5G 中 C/U 分离的网络架构，UPF 是边缘计算的数据锚点；ETSI 定义了 MEC 的商业框架，包含软件架构、应用场景和 API 接口。5G Re115 标准，网络架构创新：核心网 CU 分离，U 平面可灵活按需部署。分流机制设计：核心网设计三种流量卸载机制，满足 5G 不同场景下分流需求。业务连续性：多种业务连续性保障机制，满足不同业务的连续性需求。动态交互机制：设计应用与网络动态交互机制，支持 UPF 在本地按需增加和删除。5G Re116 将通过网络功能增强支持 5G V2X，URLLC 等低时延高可靠业务。ETSI 阶段一一定义 MEC 基本架构，服务管理 API，无线信息开放北向 API。阶段二完成 MEC 架构与 NFV 架构融合，应用迁移机制研究，定义固网，WiFi 接入 API，分析 MEC 与切片，及 5G 网络的关系。阶段三与垂直行业组织合作(例如 5GAA, ITS 等)，开展面向垂直行业的平台能力设计。

九州云结合自身实践总结，以及面向打造差异化能力的边缘服务要求，在 3GPP/ETSI 标准上做了更多的增强和扩展。九州云推出遵循 ETSI 规范、基于 TOSCA 电信标准和 EdgeGallery、K8S、OpenStack 等开放技术的边缘计算整体解决方案，涵盖的 5G MEC、OpenUPF 六大模块，实现在电信级和企业级的边缘整体解决方案，满足边缘应用的部署和服务的生命周期管理需求。

- **MECM**: 实现边缘服务的统一编排，包括应用引入、上线、部署、状态查询、升级、下线等生命周期管理。实现边缘资源的统一管理，包括边缘应用包、边缘基础架构包等资源的管理，支持边缘机房的注册、上线、启用、下线等生命周期管理，支持边缘基础架构的发现、部署、状态查询、资源监控、扩容、升级、下线等等生命周期管理。根据边缘资源的资源局限，支持不同形态的部署模式
- **ME-IaaS+ME-CaaS 边缘底座**: 提供边缘基础架构服务，提供虚拟计算、虚拟存储、虚拟网络和容器等基础能力
- **MEP 边缘协同器**: 实现 5G UPF 网元、平台服务和应用服务协同，为应用提供符合 ETSI 标准 MP1 接口，暴露平台服务和应用服务
- **Edge-MANO 企业云边管理器**: 为企业用户提供云边融合的自服

务能力平台，实现边缘计算和企业私有云的云边协同统一管理能力

- Open-UPF：基于 3GPP 标准的 5G UPF 网元，支持 5GC 的 N4 解耦标准接口
- ME-Gallery 边缘生态模块：和 EdgeGallery 开放社区应用商店打通，提供智能制造、智慧体育、智能农业等行业垂直生态接入服务



九州云的创新主要集中在以下几个方面：

系统	子模块	增强部分
5G 对接	数据平面接口	针对 5G 的用户在 N6 平面流量实现精细流控和安全增强等能力，实现用户 UE 感知、园区数据不出园等增强能力
	控制平面接口	针对控制平面，实现 IP 分流、域名解析、负载均衡等能力
	Open 5GC/UPF	针对下沉专供类 5G 专网的能力，实现亲量化的 5GC/UPF，支持 N4 接口的开放解耦
边缘业务平台	MECM	对接 OSS 实现对 MEC 全局资源的管理和调度,Edge DC LCM、VAS LCM、区域运维能力的增强
边缘部署平台	MEP	宏服务和微服务增强、UE 流量感知、轻量化 UPF 专供模式分流等
	ICT-IaaS	IaaS/CaaS 融合,CT 能力增强、Edge POD 模式、实时操作系统、加速硬件能力融合、轻量化、无人

		值守和远程运维
	MEC 服务	API 网关在鉴权、计量、性能损耗等方面增强
边缘生态平台	开放生态	和 EdgeGallery 开放社区应用商店打通，提供智能制造、智慧体育、智能农业等行业垂直生态接入服务

在 MEC 中最核心的组件是 MEP 能力，边缘接入协同平台 (MEP) 负责管理边缘业务，承担着 MEC 边缘云中 “云、网、边、端、业” 的核心交互工作。在 ETSI 的 MEP 规范中，MEP 主要的职责是流量控制 (Traffic Rule Control)、DNS 处理 (DNS Handling) 和服务分发 (Service Delivery)，基于以上基本能力，进行了大量的细化和提升。主要包括以下几个方面：

- 在 ETSI 标准的 MEP 流量控制 (Traffic Rule Control) 基础上，增强了 UPF/GW-U/DP N4 流量控制、IT/CT 防火墙 vFW 流量控制、云边防火墙 vFW 流量控制、vNAT 和 UE Devcie ID 感知和流量控制等能力。
- 在 ETSI 标准的 DNS 处理 (DNS Handling) 基础上，整合了 DNS 域名解析服务 (DNS Library) 和 DNS 分流 (DNS Re-direction)

能力

- 在 MEP 服务分发(Service Delivery)的基础上, 增强了兼容宏服务(Macro-Service)和微服务(Micro-Service)的 API 网关(API Gateway)
- 在 MEP 定义的功能之外, 实现了安全漏扫、云边协同、边边协同、服务漫游等能力实现

在 MEC 边缘云的场景中, 边缘节点(ME-DC)承载的资源, 和传统的 IT 或者 CT 数据中心相比更加复杂。在一个典型的边缘数据中心中, 即包括和 CT 相关的 UPF/GW-U/DP 等资源, 也包括和 IT 相关的 SDN 网元、边缘接入协同平台 MEP、APP 应用和 VAS 服务。MEC 边缘云基础底座能力是边缘业务承载的基石, 九州云在边缘云底座上做了优化和提升, 形成边缘 ICT 融合解决方案。

- VM-as-a-service 和 Container-as-a-service 能力融合: 集成 VM-as-a-service 和 Container-as-a-service 的能力, 能够同时提供虚拟机(VM)和容器(Container), 满足 CT 和 IT 的不同需求
- 边缘二次分流能力: 增强了网络二次分流能力, 实现 In-Line、Break-out 和 Tap 等不同层面精细分流的需求

- 看门狗(WatchDog)功能：通过虚拟机内 repair-timeout 配置协同，实现在对关键的 VNF 和 IT 业务的僵死状态感知，自动实现重启或者其他自愈干预操作
- RT 实时操作系统：实时(Real-Time)操作系统支持，提供高敏感的实时虚拟机，满足高灵敏度 CT 和 IT 业务对处理任务性能和延迟抖动的严格要求
- 加速硬件能力融合：针对硬件加速有特殊要求的 VNF 和 IT 业务，提供 DPDK/SR-IOV 加速、GPU 加速透传、NUMA CPU 预留、大页内存分配能力等能力
- 轻量化：针对一些边缘资源紧凑的节点，支持轻量化模式，实现边缘侧分布式存储与控制、计算共享宿主机的集成模式。通过镜像缓存模式，在中心集中放置镜像，边缘按需进行拉取。
- 无人值守和远程运维：安全补丁或新功能更新，通过远程推送实现边缘 ME\_ICT-IaaS 的升级。故障上报能力，根据运营商 OSS 要求向中心上报故障，实现集约化运维。

#### 四、九州云在边缘计算中的实践

随着 5G 新基建的逐步成熟，5G 边缘计算也在 2B 应用和 2C 应用领域探索也逐渐进入了深水区，九州云在过去两年参与了 100 多

个边缘场景的落地，积累了丰富的经验。

在 2B 智能制造领域，九州云和传化智联合作，在某汽车制造厂的车间，双方通过 5G 边缘计算替代传统 WiFi 网络，为 AGV 小车提供低延时、高带宽的 5G 网络，大大提升了 AGV 小车工作效率。和货物堆积如山的物流公司车间里，众多 AGV 小车穿梭来往，紧张有序地装运和卸载货物。借助 5G 网络连接，AGV 小车能够迅速接收和执行指令，行驶过程中可灵活躲避障碍物。物流车间传统上采用 WiFi 网络，面临着覆盖范围有限、信号衰减严重、后期维修复杂等问题；采用 5G 边缘计算后，得益于低延时、高带宽、大连接的特性，网络连接信号更稳定，建设维护更简单，仓储机器人可将物品自动运送到指定地点，实现自动化无人仓储运输。展望未来，5G 网络如同一把钥匙，将开启更为丰富精彩的想象空间。



在 2B 智慧交通领域，华东院联手九州云在本项目研制了一套基于 5G 混合云、边缘云和数据中台为底层框架的轨道交通安全监管平台。平台分为地图总览展示和协同办公模式两块功能，并在移动端、PC 端和大屏端分别实现。2022 年第 19 届亚运会将在杭州举办，为打造“平安地铁”“智慧地铁”、“人文地铁”，通过云边协同能力，为智慧轨交保驾护航。通过移动端实现第三方巡检、施工方、监理、监测单位、设计单位上传实时的风险动态和施工资料，整合各方数据汇聚成轨道交通四张图：基本风险图、动态进站图、问题工点图、核心风险图，最终实现统览杭州轨道交通在建工程的全局态势。该项目可实现风险动态分级，以及隐患排查治理闭合的线上动态管控。同时实现各主体，全事件，全过程安全管控。政府主管部门，建设，勘察，设计，施工，监理，第三方等协同可见。实现重大问题可启动专家会诊中心。

在 2C 领域，九州云和某运营商合作，实现基于 5G+MEC 的云游戏模式，不仅支持自己和运营商实现云游戏的业务拓展，也支持和其他的一些云游戏平台生态接入。云游戏的本质是云端算力的重新分布，具有云端的扩展性，稳定性，灵活性和集中管理等特性。云游戏从“端”到“云”和“边”，解放了用户侧终端的限制，进而对中心算力需求大增，并且有助于实现算力的高效合理利用和分配。

随着 5G 时代的到来，无论是用户需求还是游戏市场前景，云游戏都将成为 5G 在 C 端的第一个杀手级应用，成为 5G 边缘计算产品。对比 5G+MEC、5G+公有云本地机房、300M 固定宽带+公有云本地机房以及 4G+公有云本地机房这四种环境，5G+MEC 的具备最低的延时和网络稳定性是最好的（实测数据见上图）。云游戏是传输视频，每一个抖动对画面影响会很大，MEC 环境下不仅仅是延时和带宽的优势，更体现在大带宽下的稳定性（标准差小），保证流畅不丢帧。

	720P画质 + 2Mbps带宽		标准差
	网络延迟	网络延迟	
5G+MEC	8ms	9ms	2.99
5G+公有云本地机房	27ms	30ms	12.84
300M固定宽带+公有云本地机房	46ms	55ms	104.88
4G+公有云本地机房	42ms	103ms	145.30

未来，九州云的 MEC 云游戏商用方向主要包括这三个方面：通过云游戏提高互联网 APP 的内容供应；将云游戏与游戏直播结合，提供沉浸式直播互动社交体验；将云游戏方案运用在电竞赛事中。除游戏外，其实九州云 MEC 云游戏架构有很多应用的场景，例如，云游戏+直播、云游戏+互动视频、云游戏+VR、云游戏+教育、云游戏+广告等等。

# 人脸识别

## 编者的话： 中国商汤科技等研发的人脸识别在全球领先

德国《商报》在 2021 年 4 月 3 日一篇“中国对全球经济攻势”的文章中谈到商汤科技（SenseTime）的人脸识别一体机，0.3 秒极速验证准确率 99.99%。

2 月下旬，谷歌前 CEO 埃里克-施密特受拜顿总统委托，在评估中美人工智能发展水平时也谈到：“在人脸识别上，中国超过了美国，在全球是顶尖的”。

## “缺芯少魂” 攻关

## 编者的话：点评“缺芯少魂”攻关

CPU 芯片和操作系统是我国网信领域最基础的核心技术，中国曾在这方面受制于人，常被比拟为“缺芯少魂”。

近年来，我国“缺芯少魂”的局面正在解决过程中。本期发表倪光南院士的文章：“迎接开源芯片新潮流”，谈研发开源的 RISC-V 指令集 CPU，目的是为解决“缺芯”的问题。

深度信息技术（精品）专辑（第一期）发表宋可为、陈渝的文章：“自主操作系统和自建生态”，短期内国内自主开发了 66 款操作系统和自建生态（含嵌入式操作系统），这是解决“少魂”问题的，可见国内的巨大潜力！

如果不计嵌入式部分，为各种智能终端、PC 和服务器开发的操作系统为 47 款，绝大部分采用 Linux 宏内核。现在的问题：一个国家如此分散开发这么多的操作系统，确有精简、优化和整顿的必要！宋、陈在第一期文章中说是“阶段成果”，这是正确的。

这 47 款操作系统几乎都是开源的，但其中有 37 款输入开源（采用 Linux 宏内核），整体输出不开源，这是需要详细调查研究是什么原因，其中是否有对开源误解的原因？

从开源的性质来看，人们从互联网上自由、免费下载（操作系统内核）的开源代码，肩负自由传播开源的使命：可自由复制、修改、剪裁、使用或者自由进行二次发布，但这里有一个制约条件，即不能违背开源许可证的规定，中断或破坏被移植、剪裁或被应用的开源软件自由传播的特性，这时你如果将其闭源（不开源），就

是破坏其自由传播，也就是侵犯了开源软件的知识产权！如果出现这种情况是不允许的。

这里也谈一谈第一期中宋、陈提到的华为鸿蒙(Harmony)操作系统问题：

1、鸿蒙不但是一个内核，也是一个操作系统；

2、早期华为自主研发的操作系统采用微内核，主要是为 IOT 服务的；

3、2019 年 9 月 10 日，华为发布鸿蒙 1.0 操作系统，定义多内核（或混合内核），即①Linux 宏内核，②微内核，③LiteOS 的内核。华为并提出运行 1+8+N（硬件终端设备）多场景。其中：1 为手机；8 为智慧屏、音响、眼镜、手表、车载、耳机、平板、PC；N 为外设。

4、2020 年 4 月 18 日，华为发布搭载智能手机的鸿蒙 2.0（操作系统），其内核为 Linux 宏内核（方案一），宣布鸿蒙开发、测试、制作进程：

2021 年 2 月 22 日，完成鸿蒙操作系统 2.0 版本；

2021 年 3 月 31 日，完成鸿蒙操作系统 3.0 版本，同时完成自建 HMS 生态系统。

与此同时，华为发布自主设计的微内核（方案二）的开发、测试、制作进程：

2020 年 11 月，自研设计微内核；

2021 年 2 月，微内核初步测试成功；

2021 年 4 月，微内核完成全面测试。