

美团即时物流的分布式系统架构设计

宋斌 美团技术团队 2018-11-22

点击关注“美团技术团队”，阅读更多技术干货



总第308篇

2018年 第100篇

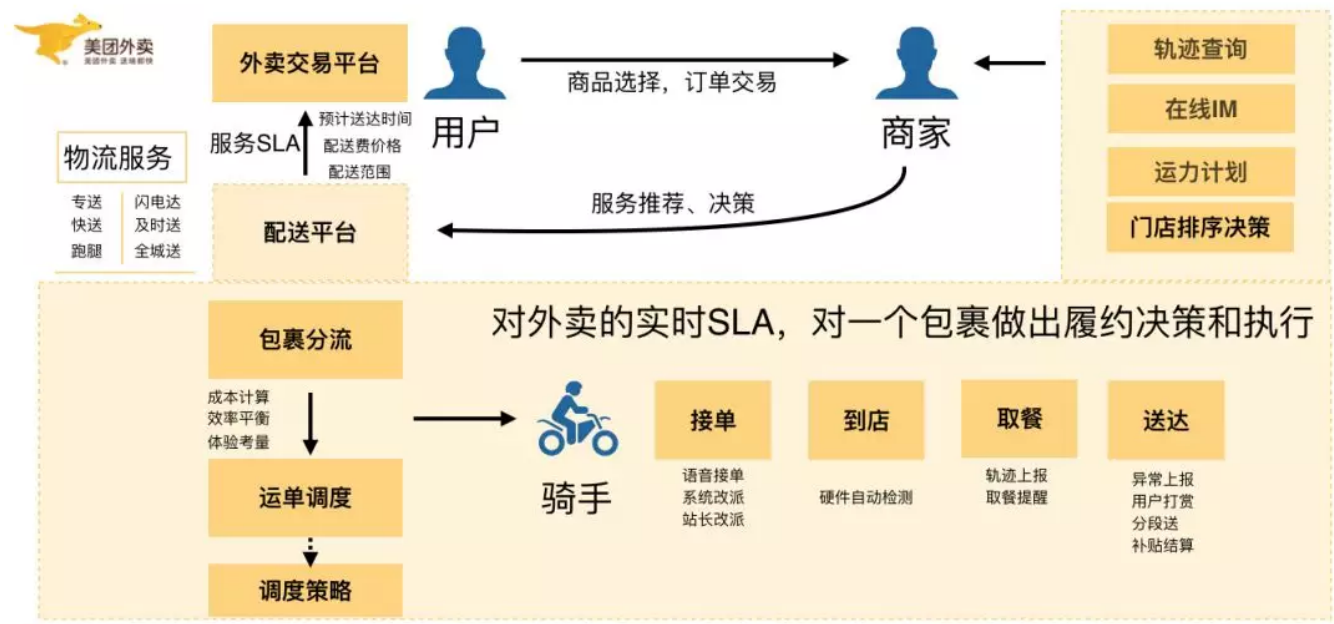


本文根据美团资深技术专家宋斌在ArchSummit架构师峰会上的演讲整理而成，主要介绍在美团即时物流分布式系统架构逐层演变的进程中，遇到的技术障碍和挑战，还有我们的解决思路。

背景

美团外卖已经发展了五年，即时物流探索也经历了3年多的时间，业务从零孵化到初具规模，在整个过程中积累了一些分布式高并发系统的建设经验。最主要的收获包括两点：

1. 即时物流业务对故障和高延迟的容忍度极低，在业务复杂度提升的同时也要求系统具备分布式、可扩展、可容灾的能力。即时物流系统阶段性的逐步实施分布式系统的架构升级，最终解决了系统宕机的风险。
2. 围绕成本、效率、体验核心三要素，即时物流体系大量结合AI技术，从定价、ETA、调度、运力规划、运力干预、补贴、核算、语音交互、LBS挖掘、业务运维、指标监控等方面，业务突破结合架构升级，达到促规模、保体验、降成本的效果。

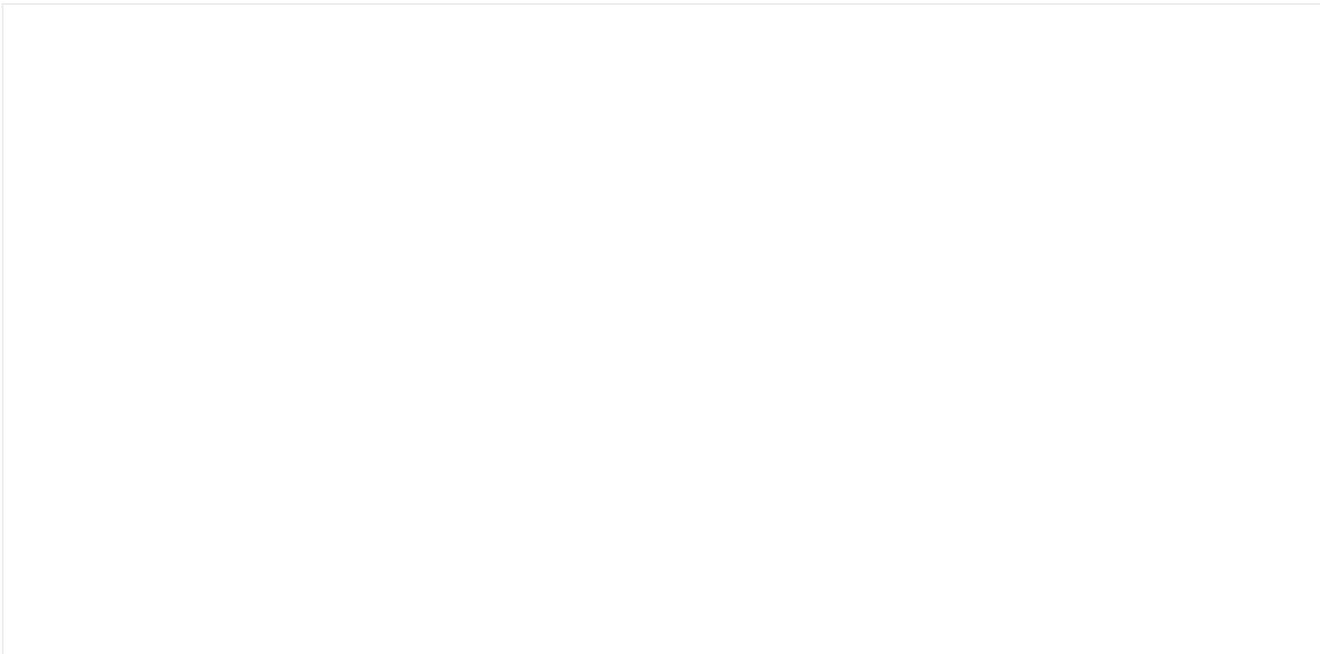


本文主要介绍在美团即时物流分布式系统架构逐层演变的进程中，遇到的技术障碍和挑战：

- 订单、骑手规模大，供需匹配过程的超大规模计算问题。
- 遇到节假日或者恶劣天气，订单聚集效应，流量高峰是平常的十几倍。
- 物流履约是线上连接线下的关键环节，故障容忍度极低，不能宕机，不能丢单，可用性要求极高。
- 数据实时性、准确性要求高，对延迟、异常非常敏感。

美团即时物流架构

美团即时物流配送平台主要围绕三件事展开：一是面向用户提供履约的SLA，包括计算送达时间ETA、配送费定价等；二是在多目标（成本、效率、体验）优化的背景下，匹配最合适的骑手；三是提供骑手完整履约过程中的辅助决策，包括智能语音、路径推荐、到店提醒等。



在一系列服务背后，是美团强大的技术体系的支持，并由此沉淀出的配送业务架构体系，基于架构构建的平台、算法、系统和服务。庞大的物流系统背后离不开分布式系统架构的支撑，而且这个架构更要保证高可用和高并发。

分布式架构，是相对于集中式架构而言的一种架构体系。分布式架构适用CAP理论（Consistency 一致性，Availability 可用性，Partition Tolerance 分区容忍性）。在分布式架构中，一个服务部署在多个对等节点中，节点之间通过网络进行通信，多个节点共同组成服务集群来提供高可用、一致性的服务。

早期，美团按照业务领域划分成多个垂直服务架构；随着业务的发展，从可用性的角度考虑做了分层服务架构。后来，业务发展越发复杂，从运维、质量等多个角度考量后，逐步演进到微服务架构。这里主要遵循了两个原则：不宜过早的进入到微服务架构的设计中，好的架构是演进出来的不是提前设计出来的。

分布式系统实践



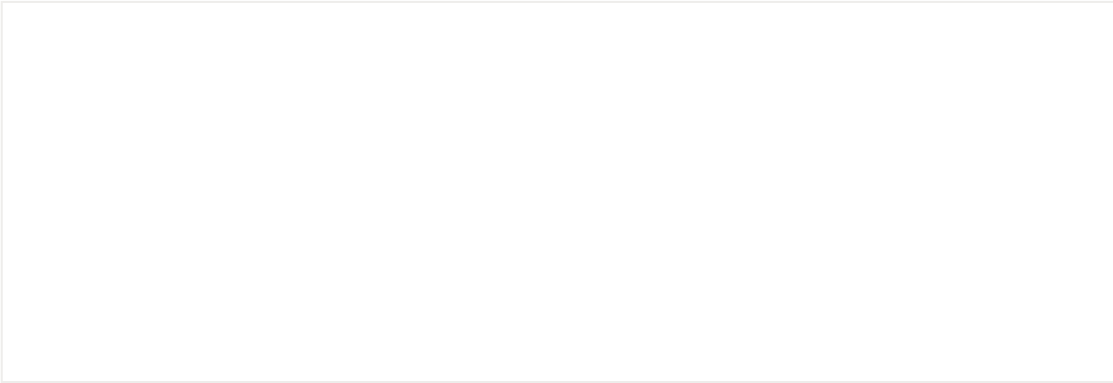
上图是比较典型的美团技术体系下的分布式系统结构：依托了美团公共组件和服务，完成了分区扩容、容灾和监控的能力。前端流量会通过HLB来分发和负载均衡；在分区内，服务与服务会通过OCTO进行通信，提供服务注册、自动发现、负载均衡、容错、灰度发布等等服务。当然也可以通过消息队列进行通信，例如Kafka、RabbitMQ。在存储层使用Zebra来访问分布式数据库进行读写操作。利用CAT（美团开源的分布式监控系统）进行分布式业务及系统日志的采集、上报和监控。分布式缓存使用Squirrel+Cellar的组合。分布式任务调度则是通过Crane。

在实践过程还要解决几个问题，比较典型的是集群的扩展性，有状态的集群可扩展性相对较差，无法快速扩容机器，无法缓解流量压力。同时，也会出现节点热点的问题，包括资源不均匀、CPU使用不均匀等等。

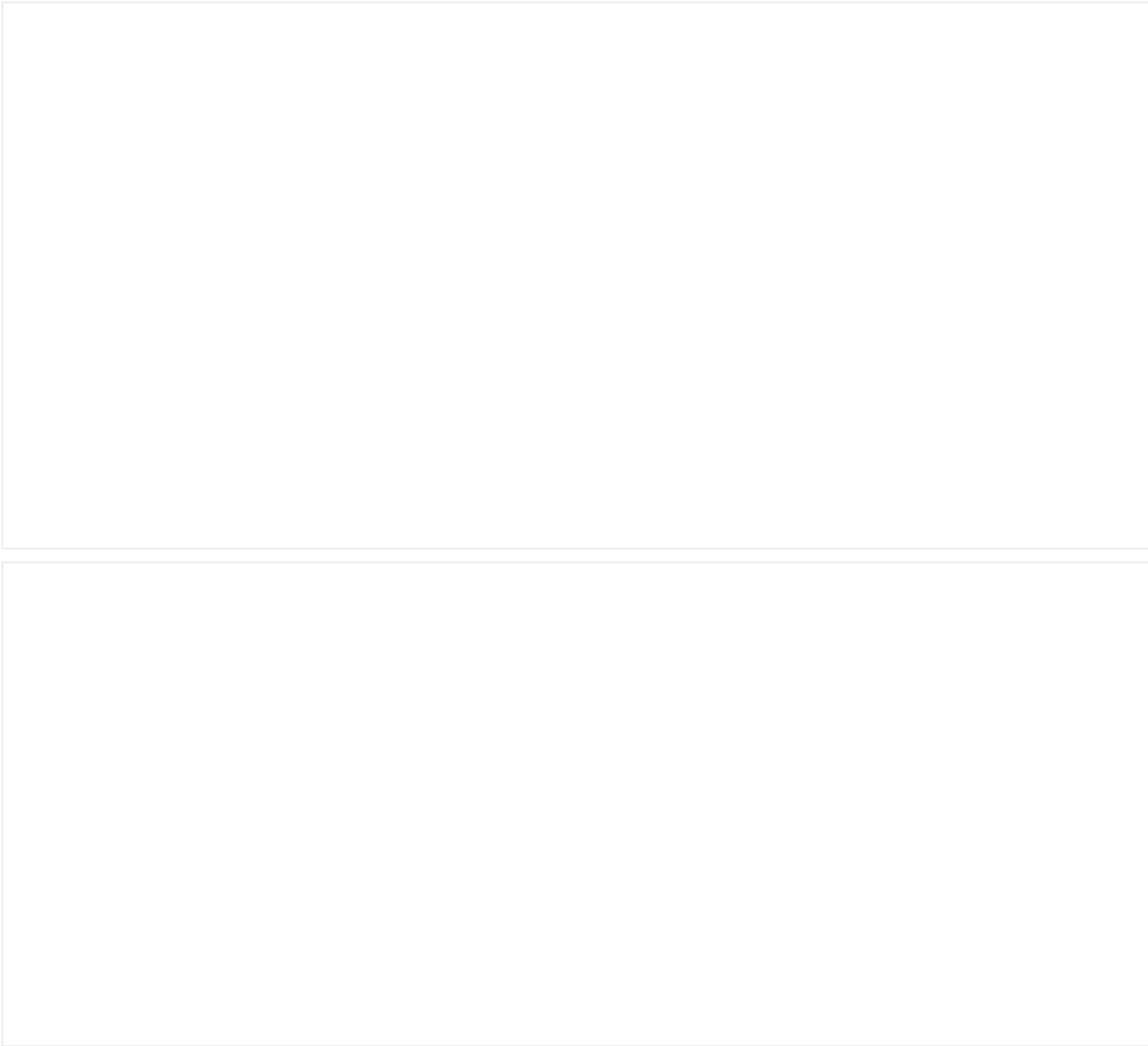


首先，配送后台技术团队通过架构升级，将有状态节点变成无状态节点，通过并行计算的能力，让小的业务节点去分担计算压力，以此实现快速扩容。

第二是要解决一致性的问题，对于既要写DB也要写缓存的场景，业务写缓存无法保障数据一致性，美团内部主要通过Databus来解决，Databus是一个高可用、低延时、高并发、保证数据一致性的数据库变更实时传输系统。通过Databus上游可以监控业务Binlog变更，通过管道将变更信息传递给ES和其他DB，或者是其他KV系统，利用Databus的高可用特性来保证数据最终是可以同步到其他系统中。



第三是我们一直在花精力解决的事情，就是保障集群高可用，主要从三个方面来入手，事前较多的是做全链路压测评，估峰值容量；周期性的集群健康性检查；随机故障演练（服务、机器、组件）。事中做异常报警（性能、业务指标、可用性）；快速的故障定位（单机故障、集群故障、IDC故障、组件异常、服务异常）；故障前后的系统变更收集。事后重点做系统回滚；扩容、限流、熔断、降级；核武器兜底。



单IDC的快速部署&容灾

单IDC故障之后，入口服务做到故障识别，自动流量切换；单IDC的快速扩容，数据提前同步，服务提前部署，Ready之后打开入口流量；要求所有做数据同步、流量分发的服务，都具备自动故障检测、故障服务自动摘除；按照IDC为单位扩缩容的能力。



多中心尝试

美团IDC以分区为单位，存在资源满排，分区无法扩容。美团的方案是多个IDC组成虚拟中心，以中心为分区的单位；服务无差别的部署在中心内；中心容量不够，直接增加新的IDC来扩容容量。



单元化尝试

相比多中心来说，单元化是进行分区容灾和扩容的更优方案。关于流量路由，美团主要是根据业务特点，采用区域或城市进行路由。数据同步上，异地会出现延迟状况。SET容灾上要保证同本地或异地

SET出现问题时，可以快速把SET切换到其他SET上来承担流量。

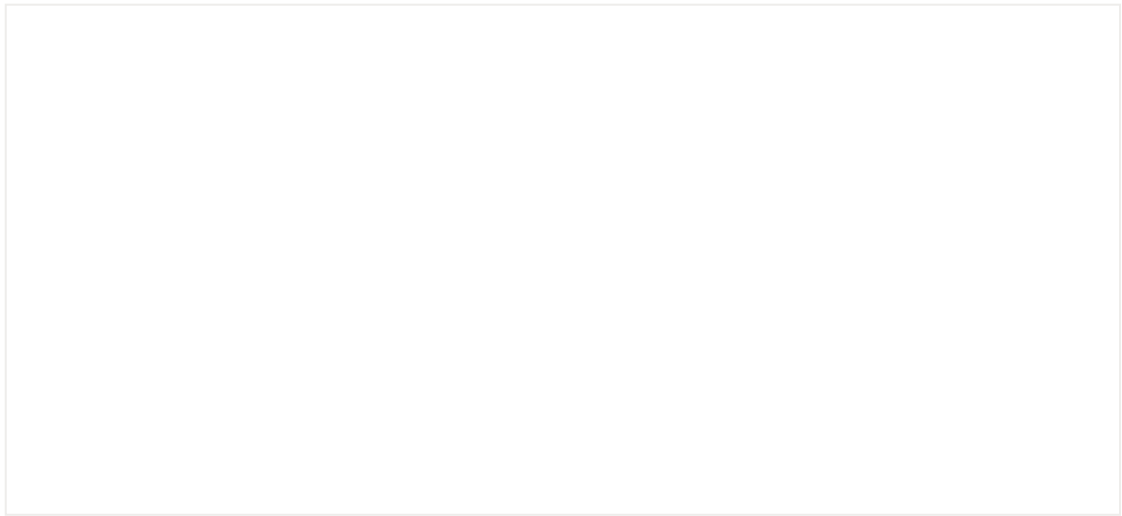


智能物流的核心技术能力和平台沉淀

机器学习平台，是一站式线下到线上的模型训练和算法应用平台。之所以构建这个平台，目的是要解决算法应用场景多，重复造轮子的矛盾问题，以及线上、线下数据质量不一致。如果流程不明确不连贯，会出现迭代效率低，特征、模型的应用上线部署出现数据质量等障碍问题。



JARVIS是一个以稳定性保障为目标的智能化业务运维AIOps平台。主要用于处理系统故障时报警源很多，会有大量的重复报警，有效信息很容易被淹没等各种问题。此外，过往小规模分布式集群的运维故障主要靠人和经验来分析和定位，效率低下，处理速度慢，每次故障处理得到的预期不稳定，在有效性和及时性方面无法保证。所以需要AIOps平台来解决这些问题。



未来的挑战

经过复盘和Review之后，我们发现未来的挑战很大，微服务不再“微”了，业务复杂度提升之后，服务就会变得膨胀。其次，网状结构的服务集群，任何轻微的延迟，都可能导致的网络放大效应。另外复杂的服务拓扑，如何做到故障的快速定位和处理，这也是AIOps需要重点解决的难题。最后，就是单元化之后，从集群为单位的运维到以单元为单位的运维，也给美团业务部署能力带来很大的挑战。

作者简介

宋斌，美团资深技术专家，长期参与分布式系统架构、高并发系统稳定性保障相关工作。目前担任即时物流团队后台技术负责人。2013年加入美团，参与过美团外卖C端、即时物流体系从零搭建。现在带领团队负责调度、清结算、LBS、定价等业务系统、算法数据平台、稳定性保障平台等技术平台的研发和运维。最近重点关注AIOps方向，探索在高并发、分布式系统架构下，如何更好的做好系统稳定性保障。

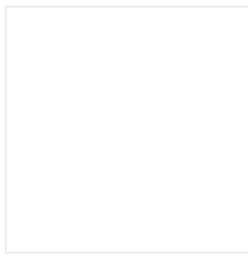
招聘信息

美团配送技术团队诚招 LBS领域、调度履约平台、结算平台、AIOps方向、机器学习平台、算法工程方向的资深技术专家和架构师。共建全行业最大的单一即时配送网络 and 平台，共同面对复杂业务和高并发流量的挑战，迎接配送业务全面智能化的时代。

欢迎有兴趣的同学投送简历到 songbin@meituan.com，chencheng13@meituan.com。

加个小广告

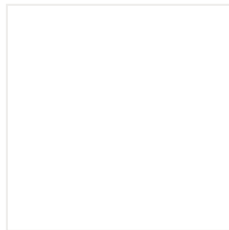
12月7日-8日，在即将召开的北京ArchSummit全球架构师峰会上，美团技术专家将分享知识图谱建设、外卖系统异常检测系统建设、打造Native体验Hybrid App实践、客户端自动化测试实践等方面的宝贵经验，欢迎大家报名参加。



前端技术专题出品人 | 冯湧

现就职于美团金服担任美团支付、保险平台业务研发总监，美团金服技术委员会执行主席。拥有16年大型互联网公司研发经验，9年研发管理经验。

挑战 Hybrid 体验技术瓶颈 —— 打造 Native 体验 Hybrid App 实践



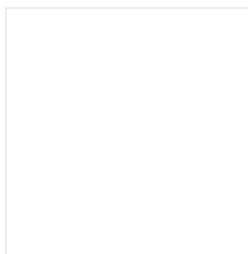
陈禹霖 | 美团 Web 前端负责人&技术专家

2015年加入美团，目前在金服负责美团支付和支付平台的前端团队，包括收银台、钱包、理财等产品。有丰富的前端架构及Hybrid 领域经验，拥有多项Hybrid领域专利。从0到1创建美团金服前端团队，组织参与美团支付收银台建设、钱包体系建设。主导了美团支付前端技术体系的构建。

议题简介

在移动端研发领域中，Native研发速度慢体验好，Hybrid研发速度快体验差。EH是完全自主研发的多端多平台Hybrid 体验技术。创新的使用一系列黑科技解决了行业级的Hybrid体验问题，达到H5的速度，Native的体验，可以在缩短研发周期30%，减小研发投入成本 50%的情况下，保持Native的界面体验。EH具备六大核心功能，包括去白屏，SPA使用 Native导航等等。

基于形变分析模型的美团外卖业务异常检测系统建设与实践



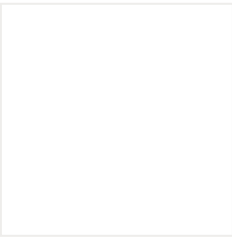
刘宏伟 | 美团资深技术专家

2016年加入美团点评，美团外卖技术保障组负责人，现在负责业务稳定性评估、实时监控、异常检测与故障诊断等方向的建设。

议题简介

美团外卖业务持续高速增长，业务迭代快，逻辑复杂，关联服务多。如何快速准确识别系统各项指标的异常，发现问题根因，并快速解决显得尤为重要。在常规业务指标监控工作中，需要手动维护上万业务指标报警阈值，不仅成本高，效果也不佳。我们尝试使用“形变分析模型”对业务指标自动进行异常检测，无需人工设置阈值。在实践过程中与外卖全链路压测，服务保护等稳定性保障系统进行内联，目前已覆盖绝大部分美团外卖C端核心业务指标，效果显著。

美团外卖2000万订单下的知识图谱建设挑战



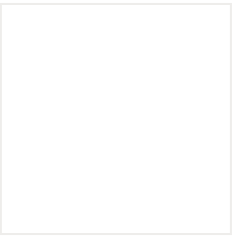
李滔 | 美团点评高级技术专家

2016年加入美团，从零打造美团外卖的智能增长和数据挖掘团队，伴随美团外卖订单从300万增长到超过2000万。

议题简介

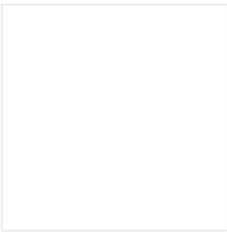
美团外卖平台目前覆盖了上百万商家和上亿餐饮类商品。平台的用户体验、推荐搜索以及商品管控等都强依赖于一个覆盖全面、设计合理、信息精准的知识图谱的支撑。而餐饮类商品具有非标准化、非结构化等特点，给知识图谱的建设带来巨大的挑战。本次演讲介绍我们通过文本、图像等多模态建模，打造了国内最大规模的美食垂类知识图谱，广泛应用在外卖的搜索、推荐以及商品管控等多个业务层面。

美团酒旅客户端自动化测试实践



张杰 | 美团高级技术专家

参与过美团、大众点评、美团旅行等App的开发工作。超过10年的开发经验，7年多的客户端开发经验，前后任职于土豆、百度等。现为美团高级技术专家，专注于客户端架构、自动化测试、基础工具等方面的工作。



胡朝旭 | 美团高级技术专家

2015年加入美团，先后在智能餐厅、酒旅住宿等业务从事Android客户端开发工作。现今在终端基础服务团队，专注自动化测试、业务开发架构、性能监控及优化等相关方向。自然科学爱好者。

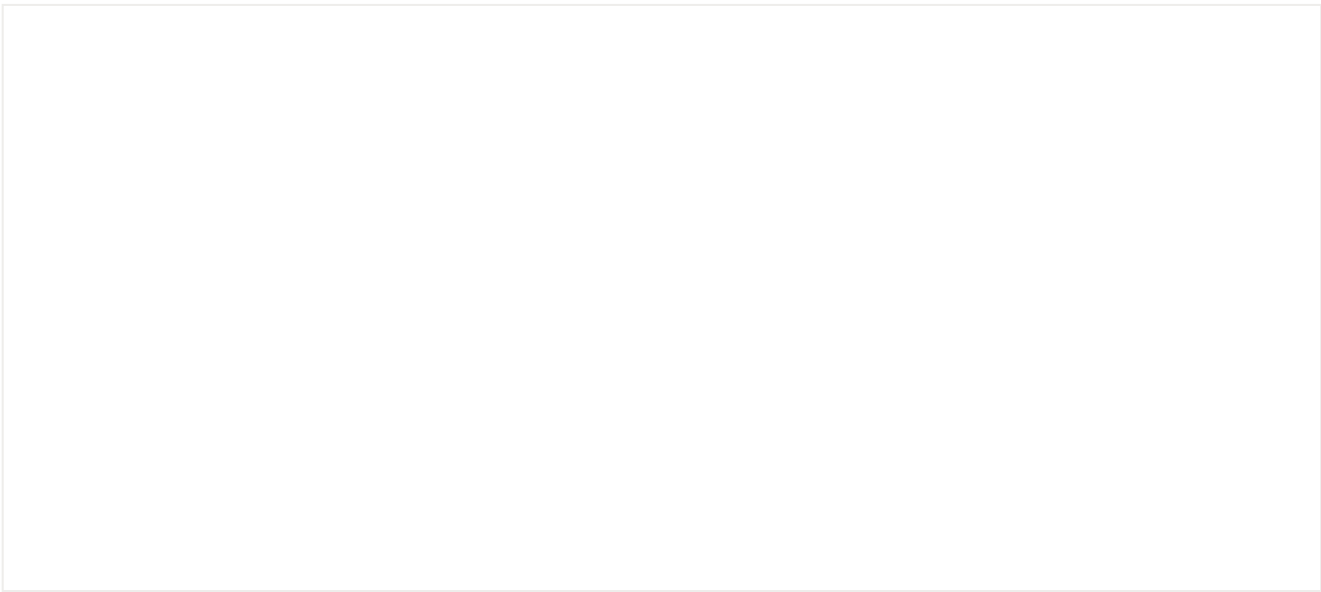
议题简介

随着美团酒旅客户端业务功能的日益繁复，如何保证功能质量、降低测试成本已经成为客户端开发流程中的一个重要课题。美团酒旅客户端团队经过一年多时间的探索，在自动化测试方面积累了较多的经验。

我们选用了功能逻辑测试为主，UI截屏测试为辅的方案，采用白盒技术栈、黑盒方式操作UI，兼具了较好的开发效率和抗重构性。目前酒旅几大业务线已经有2500多条自动化测试Case，可以在半小时内完成一次运行，为我们的质量保证和测试成本控制贡献了重要的力量。

点击“[阅读原文](#)”，即可报名参加ArchSummit全球架构师峰会。

----- END -----



阅读原文