

专访吕毅：链家网技术架构的演进之路

吕毅 极客时间 2016-07-25

链家网虽然成立于2010年，但是其技术团队却于2014年正式创立。此前技术开发采用的是传统模式，每个业务都会单独地重新开发，不仅造成各个模块孤立，并且开发人力投入成本巨大。鉴于互联网时代企业业务发展迅速，原有的传统化方式已经不适用，链家网正式建立技术团队，在原有的传统架构基础上开始了优化工作。

团队对已有的业务进行抽象，将各个业务模块中的公共部分综合出来，据此添加了一层公共的服务层，实现了平台服务化，扩展技术基础能力。此外，链家还重新搭建系统监控并完成日志监控，双级监控完善技术运营能力。目前的技术架构充分满足业务发展情况。InfoQ就技术架构变迁对链家网平台服务架构师吕毅进行了采访。

InfoQ：谈谈您为什么会选择加入链家？

吕毅：关于自我成长，相对于前两家我任职过的成熟互联网公司，在我入职时，链家网相对初期。这就预示着链家网这里会有很多挑战，我一直清晰认识到自己是有惰性的，所以我希望找到一家技术挑战多的公司，这样这些挑战会刺激我兴奋、鞭策我成长。在这里的一年，的确验证的当初我的这个想法，刚入职是我与两位伙伴做话务平台项目，不满一年，我们平台架构组已是8人的团队，负责着公司6项平台化服务，挑战与成长并行，大家都感受到架构设计、技术选型与实现、上下游沟通上突飞猛进的成长。

关于职业选择，我是技术角色，顺应之前的职业发展，我在抉择机会时候还是会选择互联网公司。然而，我将互联网公司分为两类，纯互联网公司与传统企业+互联网公司。在纯互联网公司这座大山里，我作为技术角色站在过最好的山头，虽未登顶但也看清了这座山头的风景，转头看，最近些年才被关注的传统企业+互联网公司的这片山林，一片荆棘，而我具备的能力可以开垦这片大山，为了追求开垦的乐趣从而来了这里。至于为什么是链家，当时做选择时链家人为我描绘的链家蓝图吸引了我，当时了解到产品整体负责人闫觅之前是百度的高工，这也让我对链家好感增加不少，我想技术出身的PM在产品规划上应该会更理性些。也恰巧那会做选择时想咨询下前辈鸟哥意见，鸟哥说他下个月也去链家，这让链家的Offer加分不少。

入职前的种种对传统行业的顾虑，在入职后逐渐的全部解除了。这里提供的空间与自由度，管理的扁平程度，都是让我在入职前难以想象的。在链家网，时常在露台碰见闫觅、鸟哥，吐槽吐槽对产品的不快，对公司的发展的不理解，他们都会耐心的解释为何如此，从什么角度考虑做的决策，每次聊完后我都能自我纠正了想法，同时我也乐意将这些解答传递给我们的组员。链家网给予员工充分的空间，在保证自身工作的基础上，

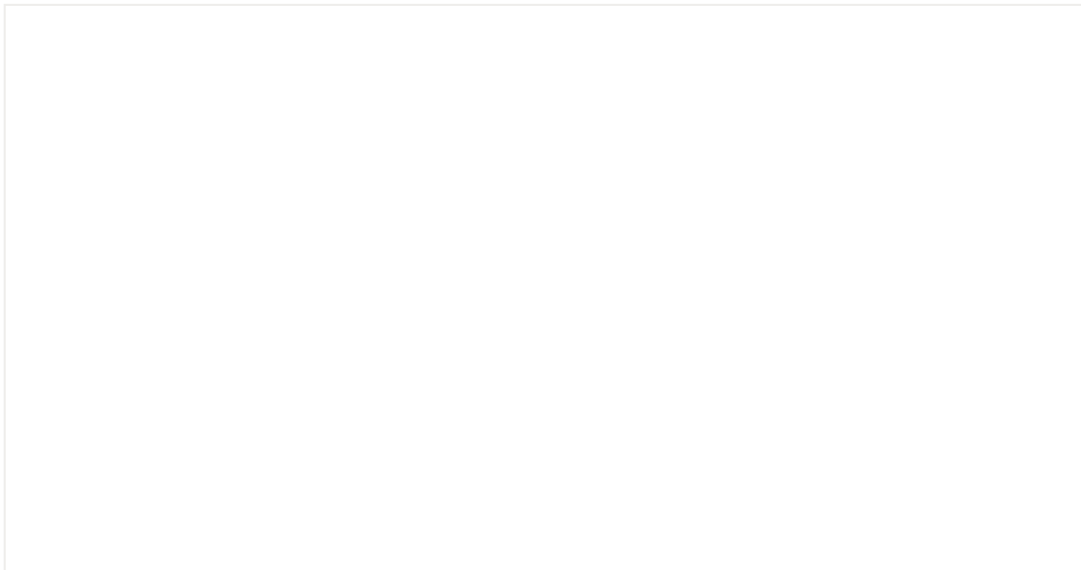
我可以与HRBP组织录制平台化服务系列课程、与员工发展部规划初级工程师培训、与外部技术社团联络到链家来组织技术活动、与一些公司技术团队做技术交流都得到了大家的肯定与支持。这里有足够的空间，只怕你的想象力不够。

InfoQ：谈谈目前链家网的主要业务及团队规模？

吕毅：链家网主要业务在很多公开资料中都有谈到，二手房方向是链家网目前的重心，围绕二手房开展的面向业主与客户的链家网、链家APP，面向链家十多万经纪人的Link项目，都是链家网目前的重点业务。链家网目前员工1000人，产品技术角色占比达一半。

InfoQ：您提到链家网这两年在架构上做了很多的调整，可以谈谈链家网的架构演进历程吗？

吕毅：链家网的架构演化的确从未停歇，从技术团队建立至今的两年时间，架构上大致分为两个阶段，如下图所示：



2014~2015年技术上关注两件事，一块是将链家集团曾交付外包实现的面向经纪人业务改造为自主研发，另一块是从零打造面向用户的业务链家网与掌上链家APP（目前已改名为链家APP），即第一年的技术架构重在业务的建设。

2015年至今，业务的逐渐成熟，引入了新的挑战。业务方向内的子业务细化伴随着链家新业务的开展，此时迫切希望抽离公共技术部分，避免重复造车的同时也希望由公共服务来支撑好业务线发展，让业务线更好的满足产品迭代。从2015年开始，一系列的公共服务的建设，构建成了整体技术架构上的服务层，同时也扩展了公司基础技术能力，并推动着建设起了系统、日志级的监控。

InfoQ：您对目前的架构满意吗？下一步优化的方向是什么？

吕毅：从技术架构的演化可见链家网成立的两年间，技术架构从无到有、从有到完善，一切都在快速的发展着。我们目前的技术架构是充分满足当前的业务发展情况的。要说对架构是否有不满，的确是有的。如今链家网业务各个方向发展迅猛，技术架构上的压力不小，理想状态是希望技术架构在支撑业务时，时刻游刃有余，但这会是条任重道远的路。

下一步的调整与优化方向，还是配合业务上一起做好服务化。业务层面希望将目前打包在一起的功能逐渐服务化，而技术架构上需要提前调整与优化，提供服务维护、服务治理、服务监控、服务通信等一系列围绕产品服务、技术服务的周边技术支持，这块是明确的方向。然而，一些未知的方向，还得与业务线、管理者们常沟通，不断摸索，按需计划着开展。

InfoQ：服务层的优化已经进行一年多，这过程中有什么经验可以和大家分享的？

吕毅：在来链家网前到现在，算起来的确有一年多在做平台化、公共服务这些事情，伴随着自我成长，感慨的确很多，列三点深刻感受的在这里与大家共勉。

第一，服务源自需求。只有业务需要的才是值得做的，只有多个业务线都需要的，才是值得拿出来做成平台服务的。多去关注业务，寻找业务线研发团队共性部分，才能发现需求，通过与业务线的沟通才能发现做平台服务的价值，所以做平台服务的技术角色切忌闭门造车，只有把自己当做半个业务线的RD才能感受到业务线的痛点。

第二，“第三选择”。常见行业内做公共服务的同学与业务线同学有些碰撞，双方各持观点，对一些边界问题拿捏不清。此时大家秉持第三选择，一起寻找更好的解决办法，目标达成一致，问题便可化解。当然第三选择的解决办法用在哪里都是合适的。

第三，服务的解耦。这是平台化服务的基本原则，服务间的解耦、服务内部的功能解耦，都是在日常设计、开发中需要注意的，只有将一块事情拆成多个点，才好做点与点之间的联系，以有限的功能点构建出无限的能力。这是老生常谈的点了，但依然值得再提一遍。

InfoQ：请讲讲你们的日志分析平台架构吧？目前的日志处理流程是怎么样的？处理的日志大概是怎样的数量级？

吕毅：日志分析平台在前面技术架构图中，属于纵向的监控部分中的“日志监控”环节，主要解决业务模块、服务模块日志字段的数据收集、展示、监控。架构设计引用公开资料中的这张截图来说明。



日志通过Kafka收集，根据日志所属RD配置的统计、监控规则通过Apache Storm实时分析日志，并将结果集数据存储数据库，实时分析期间若触发了监控规则阈值，则触发报警。数据库中的数据可以做实时的数据展示，整套方案可以让研发、测试角色实时查看日志情况，避免了大家日常合并日志文件再做shell统计的问题，并提供平台可以持续使用。

目前日志平台每秒处理的30万行日志，处理结果的展示与报警延迟在2s以内，并且这套解决方案有计划在后续开源，让业内同学低成本掌握并构建到生产环境中。

InfoQ：要设计一个高可用、高吞吐的日志平台，您认为需要重点考虑和解决哪几个方面的问题？

吕毅：设计这样一套日志方案，有以下几点需要关注：数据的收集、数据的处理、数据的存储、数据的展示。

其中数据的收集部分行业内部不少用flume采集日志，我们最早期的beta版本也是如此，但在我们方案中缓存日志数据的组建Kafka在年初版本中发布了Kafka Connector功能，实现了类似flume能力，故我们上线前就改用Kafka全套解决日志收集问题。

数据的处理，因为要求实时性，行业内也有两套方案，Spark Streaming与Apache Storm。两者共性很多，但选择Storm的原因主要是Storm的设计与Feature更专注实时运算，而Spark做离线的大并发流式处理是不错的，例如流式批运算、图片处理等等。

数据的存储，选型比较多，大家做DB选型时需关注企业级日志系统需要有大批量数据写入，特别是业务高峰时期，那么选择一个良好支持高并发写入特性的DB即可，我们使用的HBase。

数据的展示，这块就比较灵活，根据自己需求，从DB中选择数据通过组织拼装成格式化数据，配合上前端特效展示即可。

InfoQ：在你们的平台中，日志数据会采集哪些信息？你们是如何统一其他开发人员的日志格式和信息的？

吕毅：如日志平台的架构图所示，数据的采集基本是全量的，日志文件大部分是研发关注的信息，这些被关注的日志文件都会被收集，用作实时分析计数。

在日志格式这块，我们在接入层做了统一日志格式，故这份日志将会是所有业务的请求日志全集。对于业务模块、服务模块自己的日志，我们会给予建议，但没有强制规定日志格式，这部分有差异的日志格式，会在日志平台中研发角色配置统计规则时，通过正则匹配自助扣取想要的日志字段用作统计、监控。

InfoQ：与日志监控相比，系统级监控对业务层、服务层的监控指标及目的是怎样的？

吕毅：系统监控和日志监控比较，有这么几点差异：

面向人群

系统监控主要面向运维角色，RD角色关注较少，RD对于系统监控，更关注报警。日志监控主要面向RD、QA角色，他们关注业务日志某些字段出现的频次、某些值的最大最小值等，同时也可以设置基于日志的业务报警。

数据源

系统监控更多收集syslog、机器设备数据；日志监控主要手机业务线自己打的日志内容。

两者互补，在需要了解业务所属的服务器信息时候，在系统监控上查看；想了解业务数据情况时，在这套日志监控平台上看。

InfoQ：链家也是一家从传统公司转型为互联网公司的代表，你认为这中间，最大的挑战是什么？你有什么经验分享？

吕毅：传统公司转型互联网，我认同链家集团老总发表的观点，这样的转型并不是将原有业务打死，全部线上化，而是传统公司加入互联网属性，是传统业务与互联网业务相

结合的过程。链家网闫觅也曾说过“这是一次人文革命”，互联网更多的是一个工具，一个新型的商机渠道，一个提升效率的系统方案。

那么链家在做线上线下结合的过程中，**最大的挑战应该有两点：传统业务的梳理与互联网化改造，线上业务与传统业务的融合。**

传统业务的梳理与互联网化改造，链家网从创立之初就在践行，至今仍是重点工作之一，可见难度并非一般。在几次与经纪人业务侧的PM交流后，深感房产交易的复杂性，就北京而言交易中的一个环节可能会有30多种角色参与，全国各个城市政策不一，更是难上加难，而链家网在努力优化流程，给业主、用户提供清晰明了的房产业务体验。这方面的挑战折射在技术上，就是复杂的业务流程控制、风控体系、数据安全等技术难点。

线上业务与传统业务的融合，存在于在To C业务与To B业务存在交集的部分。To C业务是我们所长，但如何让还未进入房产交易流程内的普通用户，快速了解链家房产信息、交易流程，让普通用户在有需求时第一时间通过链家网、掌上链家APP发起沟通、交易，这都是挑战。折射在技术上就是如何保证好线上用户如何快速、满足需求的查找到信息、如何与经纪人快速建立有效沟通、如何沉淀意向客户并建立起关系等等的技术问题。

作为链家网平台架构组负责人，与组内7位伙伴一起，对技术追求极致的同时，我们也时刻关注着业务线的难点、痛点，希望通过技术手段提供支持与帮助，我想所有链家网角色的信念是一致的：让房产交易不再难，提供诚信便捷的服务体验！

受访嘉宾介绍

吕毅，链家网平台服务架构师，负责链家网多个公共服务。云计算浪尖时参与建设国内第一家PaaS平台SAE，移动浪尖时作为初期成员与团队打造了国内Top3的SuperAPP手机百度，重度垂直时代加入链家网负责平台化服务建设。

聊聊架构发起的顶级Docker容器技术会议

CNUTCon全球容器技术大会，我们邀请到了广发证券、平安科技、宜信、上交所、京东、阿里巴巴、腾讯、百度、携程、同程旅游、eBay、Docker、CoreOS、Mesosphere、新浪微博、搜狗、华为等公司的技术负责人前来分享他们的容器和微服务实践，极客邦InfoQ出品，目前8折报名中，团购有更多优惠。点击[阅读原文](#)链接，了解详情。



阅读原文