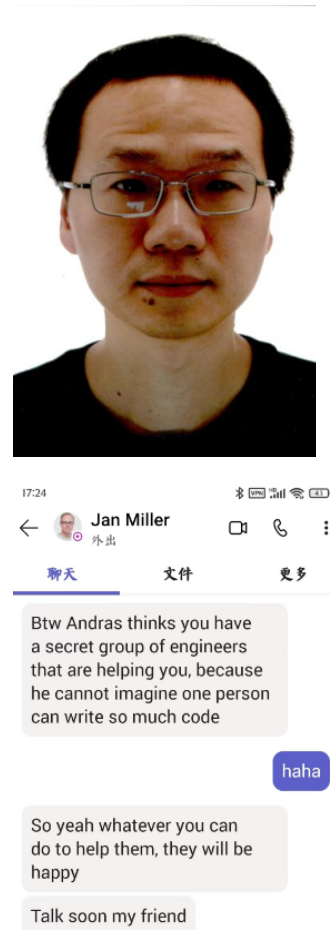


简历

姓名 许明
性别 男
毕业院校 电子科技大学
学历 本科
工作经验 18 年
最近就职企业 FileScan GmbH, OPSWAT
任职岗位 自由职业/技术伙伴
联系方式 186-2812-8964
inshua@gmail.com
blog <http://www.cnblogs.com/inshua>
github <https://github.com/inshua/>
主要能力 Team Leader、需求分析、产品设计、
业务架构设计、全栈开发
适合领域 应用架构(Application Architect)
系统架构
Team Leader
高级程序员
CTO、技术合伙人（创业公司）



家在苏州，择业地点苏州、上海、杭州、南京，也可远程

本人在校起开发小游戏，游戏修改器等，工作后一直是所谓的开发高手。2003 年因有志于从教，在北大青鸟担任讲师约 2 年，这段经历大大提升了表达能力。2005 年回到深圳，加入欣新科技，一家 GPS 初创企业，一干就是 7 年。我从接触软件开发后一直在思考开发方法。2001 年自学 UML 后意识到状态机、静态图、序列图的价值，试图从 UML 把握软件开发。在此期间发明了代码生成器、自绘编辑器等小东西，我曾经在短短的一周内仿制出 Google 地图式的 Web 地图，数据库设计也炉火纯青。这是开发高手阶段。到 2008 年，随着思考的进一步深入，我自学了 SICP 和真正的关系理论，真正实现了随心所欲的创建软件产品。后面的状态类似《黑客与画家》中描述的黑客。

2012 年离开深圳后，有 2 次作为技术主力创业成功，3 次从 0 开始组队。最近的这次创业成功经历是和德国安全专家 Jan Miller 的合作。反恶意软件领域又是一个我没有涉足过的领域，由于我充分掌握软件方法论，很快就如鱼得水，以极高的效率打造出一款轻量级的 Emulator，最终成功被美国安全企业 OPSWAT 收购。上面附图中 Andras 是 OPSWAT 的高级程序员。

研发能力

- 在业务场景中、在编码工作中、在工具使用中琢磨和想象，构思出各种存在的不存在的软件产品、开发框架、优化套路
- 既善于从零开始搭建系统，也善于建设性破坏，找出现有系统面临的困境，为其注入新的生命力，令产品在不断的使用中不断的更新迭代
- 基于系统的理论基础进行开发，切实理解 SICP、关系型运算理论、Erlang 并发理论等等，不依赖于野路子小技巧或流行框架三板斧
- 真正掌握多门语言，像用 JavaScript 一样用 Java，像用 erlang 一样用 Python，使用语言时都符合各自的编码规范，在需要时随时能进入新语言新框架的世界
- 精通大并发，真正理解并发思想，像一个掌握 akka 的人一样用 akka，像一个掌握 asyncio 的人一样用 asyncio，即使在使用 Java 时也保持 no sharing，单机单核也能做到百万日请求
- 擅长面向对象分析，结合新技术，设计灵活经济的架构方案，既能适应风云多变的未来发展，带来更高的开发效率，也能降低频繁返工的成本
- 善于组织 UI 语言表达方便好用、科技感强、轻灵舒展的界面，做出让用户喜出望外的产品
- 思维活跃，认识透彻，数据结构等基本功扎实，真正做过可以商用的编译器
- 产品经历丰富，后端前端桌面端，从事过车联网、小游戏、爬虫、串口、协议、电话语音卡、商城、CMS、ELearning、IM、OA、编译器、安全软件等等
- 和腾讯阿里百度雅虎道富拼多多等等背景同行都有共事经历，有和国外同行共事经历
- 开源项目：
 - **vba-interpreter** 用 Java 实现的 VBA/VB6 解释器: <https://github.com/inshua/vba-interpreter> ; 含编译、解释
 - **d2js** 用 java 实现的运行 js 的数据框架，比 mybatis 好用一百倍
<https://github.com/inshua/d2js>
 - **molecule** 前端框架 <https://github.com/inshua/molecule>，配合 jQuery 用非常爽，没有 vdom，加上 vdom 如虎添翼
 - **js-on-js** chn-js 用 js 实现的 js 解释器以及一个中文化的 js 编程语言，
<https://github.com/inshua/chn-js>

Team Leader

- 眼光好，善于发现优秀的小伙伴，无论他是海归研究生还是三本技校出身，都能看到他身上闪烁的技术光芒，把他拉到队伍中来
- 三次从零开始组队，一次异地组队
- 我认为团队中的人就像神经网络里的神经元，需要兴奋传导激活，形成一个活跃的神经网络，这样常常可以涌现出新的想法
- 把小伙伴当成产品人，破除分工壁垒，不把人当螺丝钉，让大家全方位参与，表达对产品的想法，变成敏捷而活跃的团队
- 真正扁平化合作，设计研发全团队共同参与，不搞设计研发上下游分割
- 严格把握时间节点，激发敬业精神，避免松懈

职业素质

基督徒，职业道德良好

坚持行业规避，十几年来没有从事过相同的行业，所以 3 年内也不涉足版权保护安全播出等领域

多次婉拒客户馈赠

拒过竞争对手高薪挖墙脚

从业经历

FileScan GmbH, OPSWAT **2020.04-2024.01, 4 年** **自由职业/技术伙伴**

<https://www.filescan.io/> 推特号 @filesan_itsec

2020 年 4 月，德国安全专家 Jan Miller 联系我，希望借助我的 VBA 解释器项目(<https://github.com/inshua/vba-interpreter>)实现 Office 文档的恶意软件分析，合作形式为自由职业。2020 年 6 月，Jan Miller 为此创办了 FileScan GmbH。经过一段时间合作，我发现他需要的是一个支持少量 API 及 COM 技术的 Windows 模拟器。

2020 年 9 月起，全职投入该项目。

2022 年底，FileScan 公司成功被 OPSWAT 收购，该公司在中国没有办公室，我继续以自由职业形式合作。2024 年 1 月，由于 OPSWAT 和美国政府有合作不能继续聘用中国人，我也不能接受到越南工作，故合作结束。

在本项目中，本人为该恶意软件分析模拟器的主要设计者和唯一开发者。

FileScan Emulator 的功能有：

1. 进一步完善的 VBA 解释器，VBScript 解释器
2. 模拟 COM 系统，模拟小量 Win32 API、模拟 Process、搭建虚拟文件系统、模拟注册表
3. 根据 MS-DOC、MS-XLS、MS-OFORMS、MS-OLEDS 等等完善 Apache POI 进行 Office 文档解析
4. 支持 Excel Macro Sheet(XLM)，XLM 实际上是一种图灵完备的表格形态的编程语言
5. 改造 rhino、HTMLUnit 等，支持 JScript 恶意软件分析
 - a) 开发 JavaScript 反混淆技术
 - b) 支持 MS JScript 独有的语言特征如宏、ActiveXObject 等
6. 支持 WSF、XSL、HTA 等恶意软件分析
7. 改造 MS Powershell (<https://github.com/PowerShell/PowerShell>)，与本模拟器对接，以支持恶意 Powershell 脚本分析。Powershell 是一门基于 .net core 的面向对象编程语言，开发语言 C#，主要技术是 DLR。
8. 将 ReactOS 中的 CMD 从 C 语言移植到 Java，支持 CMD 及 BAT 的恶意命令分析
9. 手写 RTF 解释器，以支持从混淆的 RTF 文档提取 OLE

本产品还有许多独到之处，不一一列举。截止离开时，本产品对 VBA 类恶意软件的覆盖率约 96%，每日处理约一万件文档。在开发过程中本人亲手处理的文档超过 5000 个，留存测试用例 700 余个。由于样本极多，本人开发了基于 pgvector 的日志分析工具进行相似样本的聚类。

系统使用了 Java 作为开发语言，部分功能使用 Python。Powershell 部分使用 C# 开发，通过 gRPC 与模拟器通讯。

此外，本人还开发了类似 <https://urlscan.io/> 的服务，在使用深度学习技术分析钓鱼网站过程中，发现近似的 phishpedia (<https://github.com/lindsey98/Phishpedia>) 并做了集成。该功能开发语言为 Java、Groovy、Python，使用 vertx 框架。

阜博通（杭州）信息技术有限公司 2017.04-2020.02, 3 年 架构师
（港股 03738.HK）

<http://www.vobilegroup.com/>

阜博通是杭州知名上市企业，由国家“千人”计划入围者王扬斌先生于 2005 年创办，总部位于美国硅谷，全球研发中心落于杭州，同时在北京、东京、新加坡及伦敦设有办事处。其自主研发的核心专利技术“影视基因”（该项目获艾美技术奖）在各项专业测试和市场检验中久经考验，已经成为行业的事实标准。Vobile 阜博通主要从事影视作品版权保护，并在该领域奠定了在业界的领导地位。先后签约 Disney 迪斯尼、Universal 环球、Warner Bros. 华纳兄弟、20th Century-Fox 20 世纪福克斯、Paramount 派拉蒙、Columbia 哥伦比亚等好莱坞各大传媒集团，并与 CBS、NBC、ABC、FOX 美国四大广播电视网达成合作；2008 年，获 IOC 国际奥委会认可，与 CCTV 央视网联合，成功保护 2008 北京奥运新媒体的版权。在国内公司还与华数、广东广播电视台、新华社、迅雷、腾讯、搜狐等知名企业建立合作关系。公司于 2018 年在香港上市，股票代码 03738.HK。

阜博通经多年发展，在本人加入前公司最多时有 200 余名研发人员，开发有数个庞杂的软件产品。公司曾经从阿里、道富、雅虎等等知名企业引进许多优秀人才，然而由于公司创业阶段时缺乏良好规划，架构方面的劣势愈演愈烈已经影响了公司的成长。譬如，公司在 Google Cloud 上年投入当时已达 200 万美元。这些优秀人才创造的成果也难以形成产品。尽管公司采用了各种高大上的技术，如 Google Cloud、BigTable、K8s、ZooKeeper、Kafka、RabbitMQ、akka 等等，业务架构和产品设计的先天缺陷始终无法解决。系统短板令产品缺乏弹性，难以适应新的市场变化，妨碍了利润增长。

本人加入后主要任架构师兼 Team Leader 等岗位。针对业务情景构造了可根本解决问题的软件架构及产品方案，并已形成 DEMO，但后面由于投入不足并未让系统脱胎换骨，幸运的是该方案覆盖的 YouTube 网站因性能特别突出一直在生产环境使用。后由于公司欲拓展版权分发市场，本人受命从零开始远程组建广州研发团队，开发 TVOD 系统一套。在 TVOD 业务推进中，受外部因素影响又将焦点转向短视频运营。对很多系统来说业务场景大调整对系统的扩展性都是极大的挑战。可喜的是本系统架构设计通透，具备良好的扩展性，2 个月后伴随着业务转型即已成功覆盖短视频运营服务并实现盈利。

除了上述两个系统，作为架构师兼 Team Leader，本人还完成有其它一些规模较小的工作。

本人在阜博通的主要成果有：

- 带领队伍采用 Python 语言基于 actor-message 模型、规则贯穿的版权保护新系统
- 基于上述系统开发完成关键词优化功能及 Youtube 爬虫，日爬取次数由 20 万左右提升到每日 200 万以上，同时后台产出量由 80~100 万下降到 30 万左右，无效产出大大下降
- 开发货架识别系统业务部分
- 破解某知名视频网站视频地址
- 带领队伍为广东广播电视台开发广播音频实时安全传输系统，担任需求分析、产品设计、架构设计及 Team Leader
- 从 0 开始在广州异地组建队伍（含前后端工程师、产品经理、UI 设计师），开发完成 TVOD 版权分发系统一套，基于该系统开发完成短视频运营系统一套，在过程中担任需求分析、产品设计、架构设计、Team Leader
- 独立开发视频字幕翻译工具（客户端）
- 独立开发视频拆条小工具（客户端）
- 进行 mysql 优化，将 VTWeb 数据库（MySQL）使用率从 80% 降到 20%，每月节省成本 1k 美元，过程中发明 hotview 技术
- 在爬虫优化过程中发明 python-executor

四川阿凡达信息技术有限公司**2016.05-2017.03****CTO**

成都阿凡达是四川骐骥人力资源公司的软件开发兄弟公司，其主要合作伙伴为四川航空。本人加入时已有项目在运作，另外在本人任职期间公司 CEO 丁华山还得到功夫财经项目的开发委托。本人加入后同样从 0 组队（含后端工程师、APP 工程师），

本人从 0 开始组队（含后端工程师、APP 工程师），完成有：

- 功夫财经 APP（财经类互联网 APP，各大应用市场均可下载）
- 四川航空 E 学院
- 四川骐骥人力资源公司人格测评产品
- 四川航空 HR 系统（部分参与）

因成都雾霾严重离职赴杭。

上海博涛企业管理咨询有限公司**2015.04-2016.02****CTO**

本公司系民航方面的创业公司，由吉祥航空安全总监任国强先生及东航数名机长创办，主要愿景是打造一个飞行员相关的 APP。本人加入后担任 CTO，从 0 开始招人组队（含前后端工程师、APP 工程师、产品经理、UI 设计师），在半年时间内靠 3 名主力完成一个集 IM、eLearning、个人任务、飞行记录、电子书、QA 为一体的民航飞行员使用的行业 APP。

南京智精灵科技**2014.04-2014.11****网页小游戏主程**

南京智精灵公司由心理学海归博士向华东及知名心理学学者王晓怡创办，公司的主要目标是将心理学训练范式以小游戏的形式呈现给用户。本人加入智精灵时公司将目标投向医疗康复市场，然而开发小游戏的人员一直没有就绪，之前一直靠外包进行，质量和工期难以保证。本人加入前开发了一套从 PSD 自动生成游戏的小工具，加入公司后在短短的一个月内即完成小游戏 40 个。

在智精灵完成有：

- 大量认知心理学方面 HTML5 游戏开发（cocos2d-js）
- 游戏代码生成工具
- 北京民航总院认知心理实验室某心理测评项目

后因有一创业机会离职。

依托这些小游戏，该公司资本运作获得成功，目前接近港股上市。

<https://finance.sina.cn/chanjing/gsxw/2023-09-07/detail-imzkwizn3419366.d.html?vt=4>

杭州软景信息技术有限公司**2013.09-2014.04****技术经理**

杭州软景原来主要从事医疗领域的进销存系统开发维护，后升级为医疗行业供应链系统。本人在软景完成有：

- 医疗方面供应链整合 ETL 工具（具有各行业通用价值）
- 药监码信息提取

整理 D2JS 框架（后端）及基于 D2JS 的用户权限框架。

欣新科技（深圳）有限公司**2005.05-2012.05****软件部经理**

2005 年，加入创业公司欣新科技（深圳）有限公司，从事 GIS 产品研发。先后任软件工程师、软件部经理，软件部最多时 15 人。

欣新科技由归国华侨熊锐先生创办，熊锐先生持有 GPS+图像传输发明专利，其邀请了美国软件专家周立一起在深圳二次创业，2005 年收购领航通科技，2006 年与航盛集团成立了航盛移动视讯，与交通部交通科学研究院成立了交科视讯。公司最多时有 200 余人。本人加入欣新科技时正是公司起步阶段，所有 GPS 相关软件系统都一手打造或完善过。除了一般 GPS 系统开发者需要掌握的通讯服务器、Web 展示方面的知识技能外，本人独立研发了 Web 地图组件、独立研发了规则服务器及规则设计器等一系列独有的子系统。在欣新科技本人由一个较为优秀的程序员成长为在高并发领域、软件产品设计领域、软件工程、团队招聘与领导等方面均有较深认识的软件人。

在欣新科技期间主导完成产品：

- GPS 监控管理平台
- 自主研发 Web 地图组件
- Web 地图服务器
- BS 客户端（前端、后端）
- CS 客户端（.net WinForm 技术）
- 规则服务器（通过 Java 生成 JavaScript 脚本运行于 Java 容器）
- 规则设计器
- GPS 通讯服务器
 - 采用 Erlang 编写，支持多种 GPS 厂商协议，含交通部 JT-808 JT-809 协议
- 出租车电召终端及 Erlang 电召服务器
 - WinCE .NET COMPACT FRAMEWORK
- 广告服务器（运用于出租车顶灯）
- 垃圾车系统
- 物流配送系统
- 公交排班系统
- 公务车辆排班系统

并获得相关专利一项（专利号 CN202548878U）。

基于该平台本人领导完成的二次开发项目主要有：

- 中国移动（深圳）物流通运营平台
- 世能达（天津）公司宝洁服务车辆监控平台
- 湖南中烟工业公司 GPS 平台
- 湖南怀化金大地水泥厂 GPS 监控平台
- 湖南省渣土局渣土车监控平台
- 山东某油田 GPS 系统
- 深圳西乡街道办网格化管理政务系统（使用智能手机）

在产品和项目中，开发有 Web 框架一套（使用到 DSL、AJAX 技术，动态语言、RESTful 理念，该框架是 D2JS 框架前身）；可用于 WinCE/Win32 的仿 iOS6 GUI 框架一套（完整的线程池、消息系统、组件树、自绘 UI、布局、窗口系统、动画）。

北大青鸟成都名流培训中心

2003.09-2005.04

.NET 培训讲师、教研组长

教育工作是我的夙愿，在北大青鸟担任讲师约 2 年时间。

深圳阿斯克科技有限公司

2002

工程师

开发交通银行香港分行人事管理系统。

教育经历

电子科技大学

1997.09-2001.07

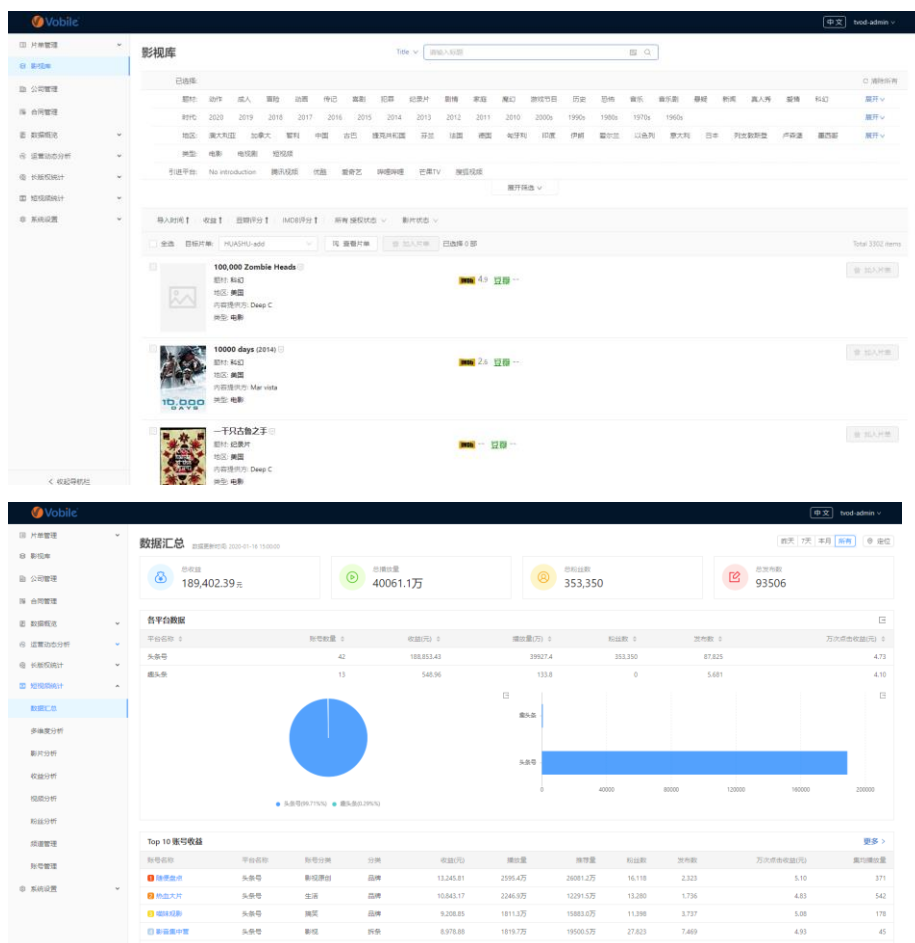
应用物理系 全日制本科

电子科技大学（University of Electronic Science and Technology of China）坐落于四川省会成都市，直属中华人民共和国教育部，由教育部、工业和信息化部、四川省和成都市共建。位列”世界一流大学和一流学科“、“985工程”、“211工程”，入选 2011 计划、111 计划、卓越工程师教育培养计划、国家建设高水平大学公派研究生项目、中国政府奖学金来华留学生接收院校、全国深化创新创业教育改革示范高校，两电一邮成员。学校原名成都电讯工程学院，1956 年由交通大学（现上海交通大学、西安交通大学）、南京工学院（现东南大学）、华南工学院（现华南理工大学）的电讯工程有关专业合并创建而成；1960 年被中共中央列为全国重点高等学校；1961 年被中共中央确定为七所国防工业院校之一。1988 年更名为电子科技大学。2000 年由原信息产业部主管划转为教育部主管。2017 年进入国家建设“双一流”A 类高校行列。

产品与项目成果

版权分发系统暨短视频运营系统（2019）

本系统为实现 TVOD 业务开发，包含 CMS、UAC、TVOD 三大部分，兼顾撮合式交易与总代模式，覆盖了意向片单、片单入库、片单出库、作品库检索、收益分析等等业务场景。后因业务转向短视频，又迅速开发出短视频运营相关产品功能。在以产品为中心的思想推动下，研发和运营高度默契，短视频业务很快实现盈利。



系统后端采用 d2js 技术，前端采用 react 技术。IMDB 豆瓣等全球影视库爬取采用 python3 实现。系统大量运用了 Postgres JSONB 类型，系统数据库结构简单，表数目少表达能力强。本小组以一个产品经理、一个前端工程师、一个后端工程师、半个爬虫工程师的规模完成了产品研发并持续支持业务迭代。本人从零开始在广州异地招聘组建小组，担任管理、架构、技术指导等各项工作。

YouTube 爬虫（2018-2020）

```
vt@search-bankend:log$ grep -c 'fetch succ' vt-2017.log.2020-02-0*
vt-2017.log.2020-02-01:2093347
vt-2017.log.2020-02-02:2083647
vt-2017.log.2020-02-03:2162832
```

YouTube 是全球最大的 UGC 视频网站，日产量很高，在本行业具有代表性。为了验证本人重构方案的有效性，并验证关键词优化的价值，基于上述方案开发了 YouTube 爬虫系统及关键词/作品优化模块。YouTube 爬虫将代理池理解为搜索机会，将需要锁的竞争问题转化为分配问题，充分挖掘代理服务器资源，使单日爬取次数从 20±万次提升到 200±万次。另外，根据实际运行数据的仔细观察和分析，本人编制了关键词/作品优化算法。该算法降低了无效关键词的频次，使每日新发现 URL 量从 100 万左右下降到 30 万左右，发现量下降随之降低了后续下载比对环节的成本。同时算法增加了高产作品的出场频次，使系统能更及时的发现新的侵权内容。本爬虫采用 asyncio 技术，单核执行，采用 leveldb 去重，LRU 内存用量约 2G。由于 python 不适合 CPU 密集型场景，关键词优化模块采用 Java 重写，与爬虫通过 RabbitMQ 交互。

在爬虫开发中本人担任指导、编码、维护等各项工作。

hotview(2019)

在 MySQL 优化过程中，本人参考 PostgreSQL 的 Continuous View 设计发明了 hotview。

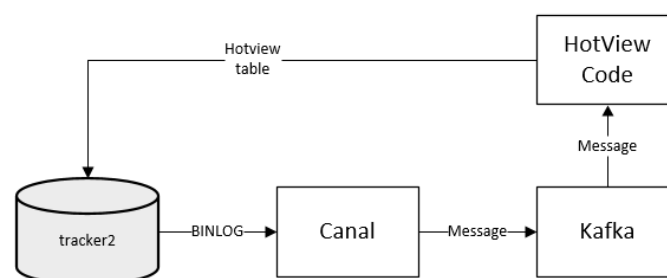
一个典型的 hotview 采用 python 编写，声明式风格，可监听相关若干表的变动，同步 hotview 持有的数据表。

```

        AND b.notice_status = 'sent'
        AND b.company_id = a.company_id
        AND a.matchedVideo_id = mv.id
        AND a.company_id = mv.company_id
        AND mv.takeoff_time = 0
        AND mv.trackingMeta_id = me.meta_id
        AND mv.company_id = me.company_id
    }
}

def __init__(self, view_db: Engine, logger: logging.Logger, cache_db: Engine = None):
    logger.info('start hot_takedownNoticeItemForResend')
    meta = MetaData()
    mapping = [
        TableMapping('takedownNoticeItem', meta,
            ColumnMapping('id', cond_alias='a.id', view_alias='parentId', is_fixed=True, primary_key=True),
            ColumnMapping('notice_send_count', view_alias='noticeSendCount', is_fixed=False),
            ColumnMapping('updated_at', view_alias='updated_at', is_fixed=False),
            trace_insert=True,
            predicate=lambda row: int(row['parent_id']) == 0 and row['is_stop_resending'] in ['2', 'false', 2]),
        TableMapping('takedownNotice', meta,
            ColumnMapping('id', cond_alias='b.id', view_alias='takedown_notice_id', is_fixed=True, primary_key=True),
            trace_insert=True, # 有时先插入 takedownNoticeItem 后插入 takedownNotice
            predicate=lambda row: row['notice_status'] in ['2', 'sent'] and int(row['id']) > 0
        ),
        TableMapping('matchedVideo', meta,
            ColumnMapping('id', view_alias='tracking_relation_id', cond_alias='mv.id', is_fixed=True, primary_key=True),
            ColumnMapping('last_refresh_at', view_alias='lastRefreshAt'),
            ColumnMapping('count_send_notice', view_alias='countSendNotice'),
            ColumnMapping('hide_flag', view_alias='hide_flag'),
            ColumnMapping('last_send_notice_date', view_alias='lastSendNoticeDate', is_fixed=False),
            ColumnMapping('is_poster_whitelist_filtered', view_alias='is_poster_whitelist_filtered'),
            predicate=lambda row: row['takeoff_time'] == '0000-00-00 00:00:00' or row['takeoff_time'] == 0
        ),
        TableMapping('metaExtraInfo', meta,
            ColumnMapping('meta_id', view_alias='metaId', cond_alias='me.meta_id', is_fixed=True, primary_key=True),
            ColumnMapping('trackingSetting_id', view_alias='trackingSetting_id')
        )
    ]
```

hotview 示例



原理图

系统采用阿里 canal 框架同步 mysql binlog，canal 可将 binlog 变动投递到 kafka 消息队列。hotview 从 kafka 收听变动，根据 hotview 定义更新 hotview 数据表。

hotview 可适用于许多被认为需要 Lambda 大数据架构的场景，非常适合 SQLer 提升统计报表、dashboard 等功能的性能。

功夫财经 APP (2016)

“功夫财经”打造了一个“最强网红经济学家”的江湖联盟，由王牧笛、匡澜联合创始，集结了李大霄、马光远、王福重、时寒冰四大掌门，以及琢磨先生、胡润等数位财经界著名人士。

新闻参考：功夫财经获合一集团 1500 万人民币投资，视频网站入局财经内容市场_36 氪
<https://36kr.com/p/5047038.html>

<https://itunes.apple.com/cn/app/gong-fu-cai-jing/id1166858886?mt=8>



成都阿凡达公司承接了该 APP 开发，本人主要负责架构设计和部分后台功能开发、开发人员招聘、技术指导等。系统后台使用 PostgreSQL 数据库，其中大量应用 JSONB 数据类型，抓大放小，动静交织，使数据库架构大为精简，达到关系型数据库与 NoSQL 的优雅融合。由于该公司影响较大，此处不多介绍技术方案。

其余成果包括 GPS 监控系统（Web 地图、规则服务器、WINCE UI 框架）、心理学小游戏、飞行员 APP（IM、阅读、飞行记录本、E Learning）等等，请参考

<https://github.com/hsuming/resume/raw/master/Products.pdf>。