THE FORELAND OF
TRADING TECHNOLOGY

内部资料 免费交流
《准印证》编号沪（K）0671

交易技术前沿

2020年 第二期 总第39期

本期主题

信息安全

No.2

上海证券交易所 SHANGHAI STOCK EXCHANGE
ITRDC 证券信息技术研究发展中心（上海）

内部资料 2020 年第二期（总第 39 期）
准印证号：沪（K）0671

NO.2

主管：上海证券交易所
主办：上交所技术有限责任公司
总编：黄红元
副总编：徐毅林
执行主编：王泊、谢毅
责任编辑：黄俊杰、徐丹、郭望

上海市浦东南路 528 号
邮编：200120
电话：021-68828590, 021-68813289
传真：021-68813188
投稿邮箱：ftt.editor@sse.com.cn

扫码阅览历期杂志

篇首语

随着金融科技快速发展，业务与科技不断融合，在技术赋能行业机构的同时，也带来了潜在金融风险。为加强安全风险监管防控，信息技术治理、合规、风控、审计构建了行业机构的四张安全网，共同维护金融科技稳定发展。本期《交易技术前沿》以“信息安全”为主题，收录来自行业十一篇优秀文章，探讨行业技术前沿。

其中，《证券期货机构信息化监管中日志文件使用的初步探索》阐述了证券期货行业监管中对信息化监管和日志文件使用的初步探索，提出一种信息化监管的方法，以提高证券期货行业监管的有效性。《关于＜证券基金经营机构信息技术管理办法＞中“重要信息系统”的信息技术审计实践》为证券基金行业的重要信息系统审计提供新的实践思路，从环境管控、SDLC 管理、权限管理及日志分析四方面展开分析讨论。《“多对多池化”高可用集群技术的实践》叙述了券商单点应用（如银行等外部机构提供的应用）实践“多对多池化”高可用集群技术的探索，有效解决竖井式 IT 架构的服务器单点故障可能引发的诸多问题。《东方证券服务治理建设实践》介绍了东方证券通过构建统一的服务治理框架，支持企业架构向以微服务为核心的现代架构转型。《北金所数据中心迁移》介绍了北金所数据中心的迁移工作，包括总体思路、前期准备、数据库迁移、物理搬迁、模拟搬迁与实际搬迁、风险控制、故障处理及回退方案。

金融科技提高了金融服务效率，但其中隐含的安全风险一旦发生，可能为经营机构和整个行业带来很大损失。近年信息安全事件有所增多，更需补齐信息技术管控短板，同时兼顾发展，保持技术活力。

《交易技术前沿》编辑部

2020 年 6 月 30 日

目录 Contents

本期热点 Hotspot

实践探索 Exploration

行业观察 Observation

信息资讯采撷 Information

本期热点 Hotspot

1 证券期货机构信息化监管中日志文件使用的初步探索

2 关于《证券基金经营机构信息技术管理办法》中“重要信息系统”的信息技术审计实践

证券期货机构信息化监管中日志文件使用的初步探索

谭振宁 / 大连证监局机构处 / tanzhn@csrc.gov.cn

证券期货行业是国民经济发展中不可或缺的组成部分，因此，对于证券期货市场的监督和管理必须要重视起来，以便应对多元化的风险，发挥证券期货行业在经济中的功能和作用。本文主要阐述了证券期货行业监管中对信息化监管和在信息化监管中对日志文件使用的初步探索，提出来一种信息化监管的方法，来提高证券期货行业监管的有效性。

近年来，我国证券期货市场发展迅速，证券期货市场规模不断扩大，社会影响力不断增强。但在证券期货行业欣欣向荣的局面下，信息量爆炸式增长带来的问题也呈现出来。随着《中华人民共和国网络安全法》的颁发和实施，更表明国家已经将全信息安全问题上升到国家高度。证券期货行业的稳定健康发展，关系着亿万投资者的切身利益，也关系着证券期货市场甚至社会稳定

大局，那么证券期货行业的监管能否起到相应作用，则显得尤为重要，证券期货市场能否有效运行则是证券期货市场监管能否发挥其重要作用的鉴证。因此，监管机构应力求突破和完美，在健全制度体系的同时，加强证券期货行业在信息化领域的监管，不断提高监管效能，保障市场安全有效运行。

一、证券期货机构监管的意义和内容

所谓证券期货市场监管，是指证券期货管理机关依照法律、法规和国务院授权，统一监督管理全国证券期货市场，证券期货市场监管是一国宏观经济监督体系中不可缺少的组成部分，对证券期货市场的健康发展意义重大。一方面证券期货业具有特殊性。证券期货业作为经营货币这种特殊商品的特种行业，其业务经营涉及各行各业和千家万户。因此证券期货业都是公众性企业，证券期货业健全与否，已不仅仅是证券期货业本身的事情，还关系到国民经济健康发展，社会政治稳定。另一方面证券期货业具有高风险性。众所周知，证券期货业本身是一个经营风险和社会风险都很大的行业。因此，从保护投资者利益出发，严格对证券期货机构监管，使证券期货风险降低到最低程度，是证券期货管理机关不可推卸的社会责任。

二、证券期货行业信息化监管现状和问题分析

我国已经构建起了严密、高效的网式证券期货监管体系。即以政府为主导、以自律组织为补充，包括中国证监会及其派出机构、证券期货行业协会，证券期货交易所以及其他监控机构在内的，上下一致，纵横相连，职责分明，相互配合，严密、高效的网式监管体系。体系对证券期货市场的各个方面、各个环节都实行有效监管，在这个体系中，政府的主导作用，行业协会的自律作用以及证券期货交易所的一线监管作用缺一不可，形成合力，共同规范市场运作，确保市场的有序运行和健康发展。

但是国内外的证券期货市场监管实践表明，仅仅依靠证监会及其派出机构行政管理和日常检查方式，很难实现对证券期货市场的全面监管。尤其是在证券期货业信息化高速发展的背景下，大多数的监管作用仅体现在安全事件发生后的事件报告和处理，无法及时规避风险和挽救经济损失。因此，在证券期货市场运行过程中监管机构应该重视信息化领域的监管，努力使我国的证券期货市场信息安全程度适应我国现阶段证券期货市场的发展。目前看来，我国证券期货市场的监管人员在数量和信息化能力上，较难满足当前监管工作的需要。提升监管人员的信息化能力固然重要，但在信息时代，最实用的还是完善能够补足监管机构在信息化监管领域短板的技术措施，通过信息化的手段，对证券期货市场进行有效监管。

证监会一直高度重视监管科技工作。2018 年 8 月，证监会正式印发《中国证监会监管科技总体建设方案》（以下简称《方案》），明确监管科技战略规划与愿景。《方案》明确监管科技 1.0、2.0、3.0 各类信息化建设工作需求和内容，提出大数据分析中心建设原则、数据资源管理工作思路和监管科技运行管理“十二大机制”，着力实现三个目标：一是完善各类基础设施及中央监管信息平台建设，实现业务流程的互联互通和数据的全面共享；二是应用大数据、云计算等科技手段进行实时数据采集、实时数据计算、实时数据分析，实现对市场运行状态的实时监测，强化市场风险的监测和异常交易行为的识别能力；三是探索运用人工智能技术，优化事前审核、事中监测、事后稽查处罚等各类监管工作模式，提高主动发现问题能力和监管智能化水平，促进监管模式创新。

三、信息化监管中对日志文件的使用初探

1. 日志信息对于信息化监管的作用

证券期货市场信息化建设较快，信息系统规模普遍较大，信息资产类别和数量较多，包括应用系统、服务器、数据库、安全设备和网络设备等。这些信息资产在系统运行过程中，产生了大量的日志信息。日志的存储和分析是信息安全事件溯源和查找系统内隐藏风险的关键，日志文件的保护和查询对于信息系统的重要程度不言而喻。通过查阅、审计这些信息，管理人员不仅可以了解系统的状态，还可以在系统出现问题时，确定系统当前状态、回溯入侵者踪迹、寻找某特定程序或事件相关的数据。有助于更好地监控和保障信息系统运行，及时识别针对信息系统的入侵攻击、内部违规等信息，同时能够为安全事件的事后分析、调查取证提供必要的信息。

监管机构通过对证券期货市场信息系统操作日志的监管，能够及时发现问题，加强责任追究和追踪，并进行合规性审计等，定期对发现的问题进行反馈和要求整改，确保行业信息系统安全稳定运行。

明确日志在信息化监管中的作用后，如何通过技术措施提升监管机构监管效能和证券期货公司信息安全水平的作用，就是需要解决的重点。

2. 信息化监管科技的初步探索

信息化监管科技主要用来分析和审计系统及事件日志，能够对信息资产产生的日志进行全面收集和细致分析，通过统一的控制台进行实时可视化的呈现。通过定义日志筛选规则和策略，帮助监管人员从海量日志数据中精确查找关键有用的事件数据，准确定位网络故障并提前识别安全威胁。

信息化监管科技主要应分为采集、分析、展现三个层次。通过通用数据接口，与各种安全设备、网络设备、操作系统、数据库、应用系统对接，并采集日志，利用自身算法对日志进行联动分析，并根据自定义的安全策略和阈值，发现异常和报警。对于服务器操作系统、数据库、安全设备和网络设备等产生的日志信息，采用集中收集统计方式比较实用，但对于应用系统信息的收集和分析，则相对比较困难。但由于证券期货行业内的应用系统，仅有几家知名软件公司进行开发，大部分证券期货公司部署的应用系统比较统一，这也为证券期货行业进行统一监管，提供了较大便利。

证券期货市场的信息化监管科技应满足对 Windows/Linux/Unix 等操作系统的服务器的事件日志、路由器及交换机等支持 Syslog 的设备日志、MS SQL 和 Oracle 等主流数据库日志、IDS 入侵和网络审计等安全产品审计日志等进行收集，对采集的所有日志、事件和告警信息统一完整存储，建立一个统一的证券期货市场日志存储平台，为故障排除和信息取证提供可靠的来源和依据。同时根据事件类型、严重程度等设置日志过滤规则，有选择地保存关键的事件日志，便于搜索特定事件和优化数据库容量。应用大数据和人工智能技术实现海量安全日志、告警的汇聚、统计、分析、挖掘。

可以实现：从事件报表深入钻取，查看有关某台主机或设备的事件具体信息；预定义丰富的告警标准，同时支持灵活自定义告警的标准，便于监控特定的对象；当生成的事件符合标准时，系统产生告警并自动通过短信、Email、等方式通知用户，并能够自动执行预定义命令行程序；建立丰富的合规性和用户活动报表，并支持创建自定义报表，便于通过报表和图表数据直观查看和分析事件日志的分析结果，自动生成分析报告；加强防篡改措施，保证输出信息和报表的准确性和可靠性，无法进行恶意篡改，使监管机构看到的问题情况，一定是证券期货公司发生过的，不会发生隐瞒和

误报等情况；提供远程接入监控的方式，方便监管机构实时对信息系统的运行状况进行查看，并对界面的可视化进行一定的侧重研究，通过内置问题库和程序化文件，对发生的问题进行详细描述和确认，达到监管机构能清晰判断问题发生所造成的影响和严重程度的目的。

3. 信息化监管科技的应用

日常运行中，监管机构可以通过远程 VPN 接入的方式进行日志报表的查看和判断，也可以定期要求证券期货公司上报近期的系统运行情况，并现场进行安全检查。发现问题后，及时通知证券期货公司进行事件调查和处理，并从监管的角度，对证券期货公司的日常运行维护提出相关的整改和建议。证券期货公司在系统运行过程中，也可以根据信息化监管科技提供的信息，对系统的安全状况进行了解，及时改进信息系统，达到一定的安全防护能力，避免遭到监管部门的处罚。

监管机构在应用信息化监管科技的过程中，还应采取措施确保内部工作人员按照规定的业务流程和操作规程开展日志采集和管理工作，防止和避免内部人员利用信息系统监守自盗等非授权的行为操作，确保信息系统的安全。

四、结语

随着金融科技的进步，信息安全问题越来越多，监管机构只有依托信息化监管科技的辅助，监管科技、金融科技协调发展、互相促进，才能使证券期货公司的信息系统在全面支持业务发展需求的前提下，向着安全、高效、经济的方向不断完善和发展，才能基本满足资本市场不断创新、持续规范、稳定运行的需要，为资本市场的快速、稳定发展提供坚实的保障。

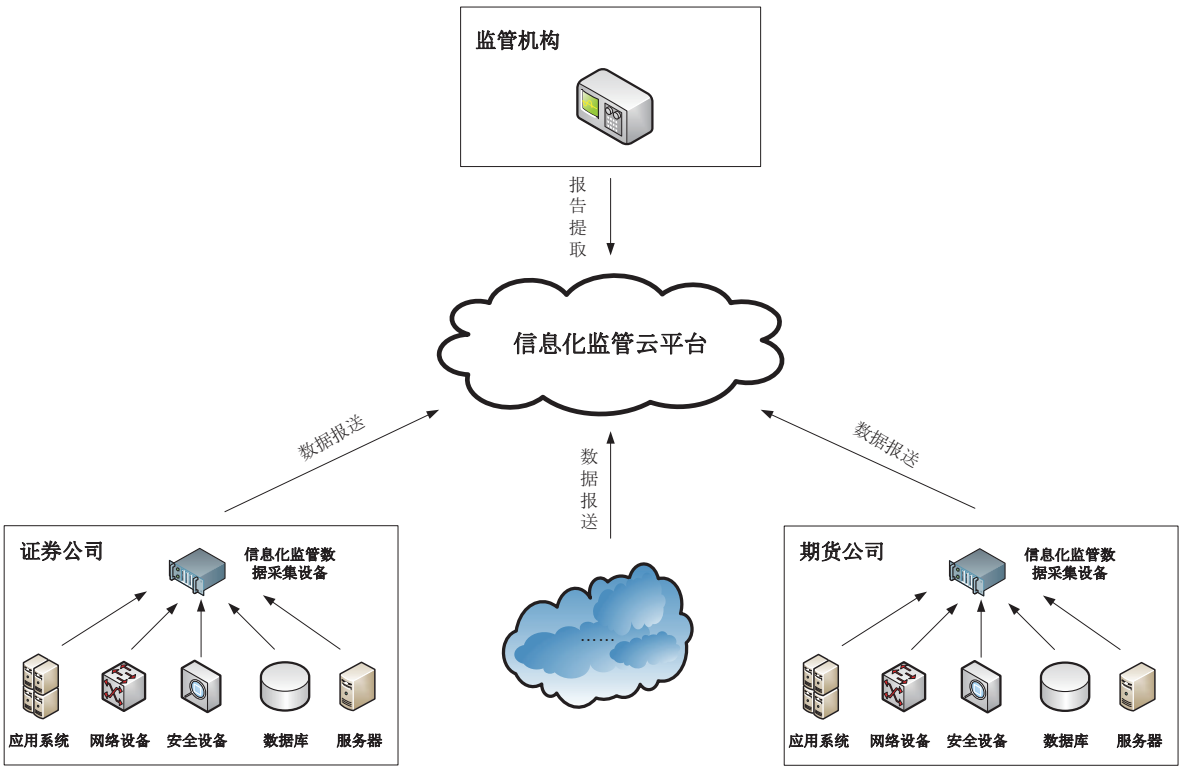


图 1：信息化监管科技的应用模式

关于《证券基金经营机构信息技术管理办法》中“重要信息系统”的信息技术审计实践

胡益民 / 平安证券股份有限公司
顾宏倩 殷昊南 张帆 / 上海艾芒信息科技有限公司 邮箱：fzhang@c-ross.com.cn

一、引言

2018 年 12 月 19 日，证监会发布《证券基金经营机构信息技术管理办法》（以下简称“《办法》”），并于 2019 年 6 月 1 日起正式实施。《办法》作为证监会部门规章级别、包含近十年信息技术监管条线经验的制度设计，结合行业信息技术建设现状，围绕治理、合规及安全三条主线，提出了全面信息技术治理要求及管理目标。随着证券基金经营机构逐步落实《办法》要求，从成立信息技术治理委员会、招聘首席信息官，到搭建灾备数据中心，证券基金经营机构的硬件建设日趋完善，而我们也发现，作为贯穿证券基金经营机构重要信息系统建设全生命周期，体现机构信息技术建设软实力的信息技术审计，在实践中仍缺乏有效的落地方式。

《办法》第六十三条对行业的重要信息系统范围作出界定，共包括 16 类，即集中交易系统、投资交易系统、金融产品销售系统、估值核算系统等；同时《办法》在第十六条也指出，证券基金经营机构应当定期开展信息技术专项审计，频率不低于每年一次，确保三年内完成信息技术管理全部事项的审计工作，包括但不限于信息技术治理、信息技术合规与风险管理、信息技术安全管理、应急管理。

《办法》始终从证券基金经营机构规范开展信息技术建设层面出发，将长期以来“重建设、轻合规”的错误概念，以内部审计和外部审计作为抓手，并与信息技术治理委员会日常管理职能相呼应，从公司层面进一步推进各类信息系统建设，真正发挥对前台业务的支撑。本文旨在结合行业各证券基金经营机构在信息技术管理审计方面的管理实践，讨论关于机构重要信息系统审计的重点关注领域及审计思路。

二、信息系统审计现状与发展趋势

（一）证券基金经营机构信息系统审计现状

一是证券基金经营机构主观意识不足。从目前行业信息技术审计开展情况来看，由于系统的复杂度与专业度，证券基金经营机构长期以来普遍仅重视信息系统建设及业务发展，而对信息系统的安全与合规淡视甚至忽视，缺乏对重要信息系统审计工作重要性的意识。

二是缺乏较好的信息系统审计工具。从具体的信息系统审计实践来看，各类经营机构信息系统建设程度不一，缺乏实施信息技术审计的有效工具；前台业务的复杂带来信息系统的复杂，导致审计较难有统一的工具，而停留在通过 Word、Excel 等简易工具的审计方式必然会导致审计效率的低下，进而导致监管、行业自律组织及各证券基金经营机构对长期的外部审计工作及经营机构内部追踪整改缺乏有效抓手。

三是行业对于信息系统审计缺乏深度及广度。从信息技术外部审计的发展来看，外部机构在信息技术审计方面欠缺专业性，审计人员通常难以同时具备信息技术知识、业务知识及相关法律法规的深刻认识，主流的审计项目仍停留在公司日常业务经营及各项财务指标，缺乏对信息技术治理中合规风控方面的考核。

（二）发展趋势

信息系统审计是监管实施监督管理、证券基金经营机构内部采取控制管理的重要手段，经营机构管理层的重视程度、审计人员的专业胜任能力、可自动化迭代升级并覆盖行业各类审计要点的审计工具是支撑重要信息系统审计工作开展的三个因素，缺一不可。

三、重要信息系统审计实践示例

本段旨在就重要信息系统中最为重要的四部分内容作出审计实践思路的分析及讨论，具体包括环境管控、SDLC 管理、权限管理及日志分析。

（一）环境管控

广义的环境管理涵盖文化环境、物理环境、网络环境等各个方面，本文以重要信息系统为主线讨论其从物理环境到系统环境的各重点审计评价项（如表 1 环境管理审计评价项）。

表 1：环境管理审计评价项

环境管理			
物理环境	系统环境		
数据中心环境	主机环境	数据库环境	终端环境

1、数据中心环境

数据中心的选择需要基于重要信息系统具体的 BCP（业务连续性计划）、DPR（灾难恢复计划）等能力要求，结合《办法》第四十一条关于重要信息系统数据备份能力、故障应对能力、灾难应对能力和重大灾难应对能力的要求，具体指标可参考《证券期货业信息系统备份能力标准》。

在实际关于系统数据中心的审计评价过程中，审计师需要考虑其环境信息、电力支撑、安全防护、设备备份、通信稳定性等各个因素。如结合《数据中心设计规范》（GB 50174-2017）11.1.2 和《证券期货业信息系统审计规范》（以下简称《审计规范》）A.4.2.1.7 电力供应部分所述要求，承载等保三级系统的数据中心应当采用双路市电，故审计师应当基于该项要求了解抽样信息系统所在主机房及备份机房的电源接入方式，以确认系统的单点故障风险是得到控制的。

2、系统环境

本文所指系统环境包括主机环境、数据库环境及终端环境。证券基金经营机构在实际的业务活动中所涉及的主机主要分为 Windows 和 Linux，涉及的数据库主要包括 MS SQL Server、MySql 和 Oracle，终端所涉及架构主要分为 C/S 和 B/S 架构。

机构应当考虑在执行审计的过程中不同的主机、数据库、终端环境对于审计师的专业胜任能力要求，并为审计活动配备相应的技术及人力资源支撑。审计师需要围绕《办法》的要求，结合《审计规范》等标准或指引，对系统环境的身份鉴别、访问控制、安全管理、资源控制等各项内容执行控制测试，具体的测试内容可以包括密码策略、屏幕保护、登入限制、路径共享、完整性校验等。如图 1 所示，在确认数据库连接和穿行测试工作底稿正确后，程序可以自动查看日志信息如（用户信息表 dba_users、检查权限信息表 dba_profiles、检查审计痕迹 dba_audit_trail），将检查结果写入工作底稿中进行保存。程序依据预先设定好的审计标准对相应结果进行自动评分（可辅以人工评分），在测试程序运行结束后系统将自动实现自我评分（可以辅以人工修改）。

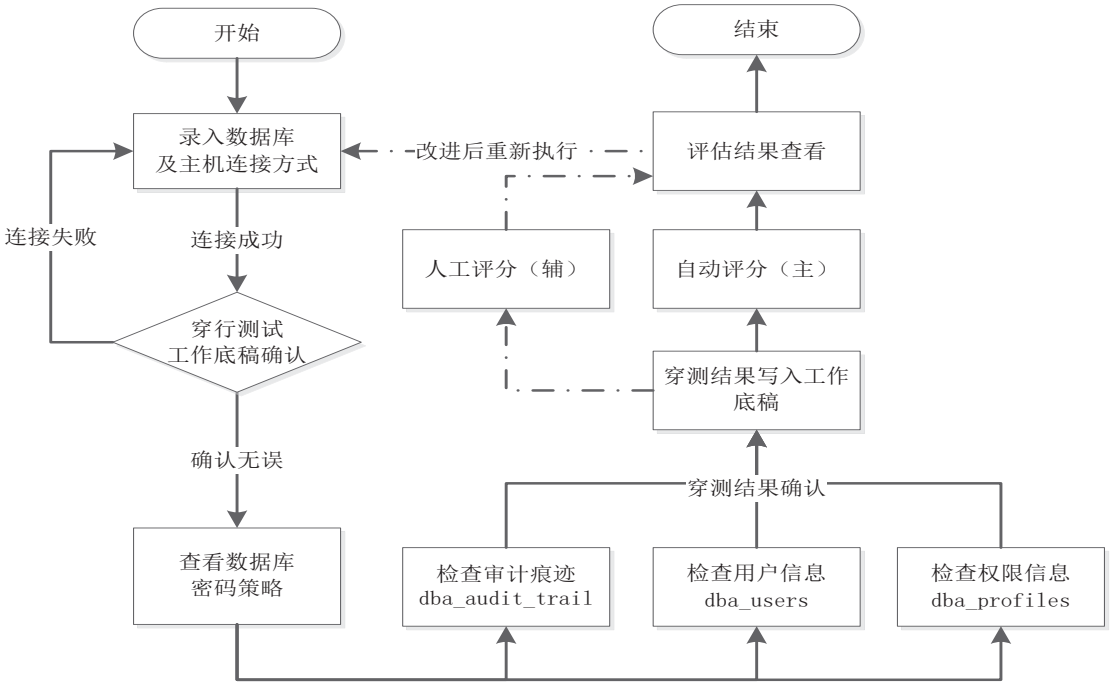


图 1：数据库密码策略穿行测试

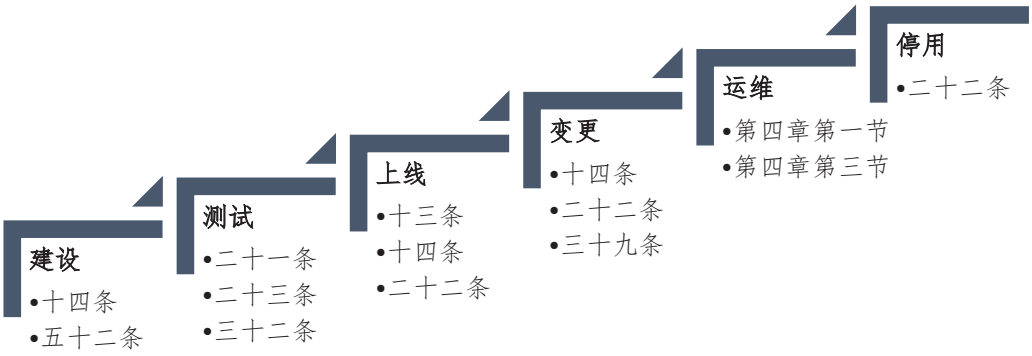


图 2：SDLC 管理审计评价项

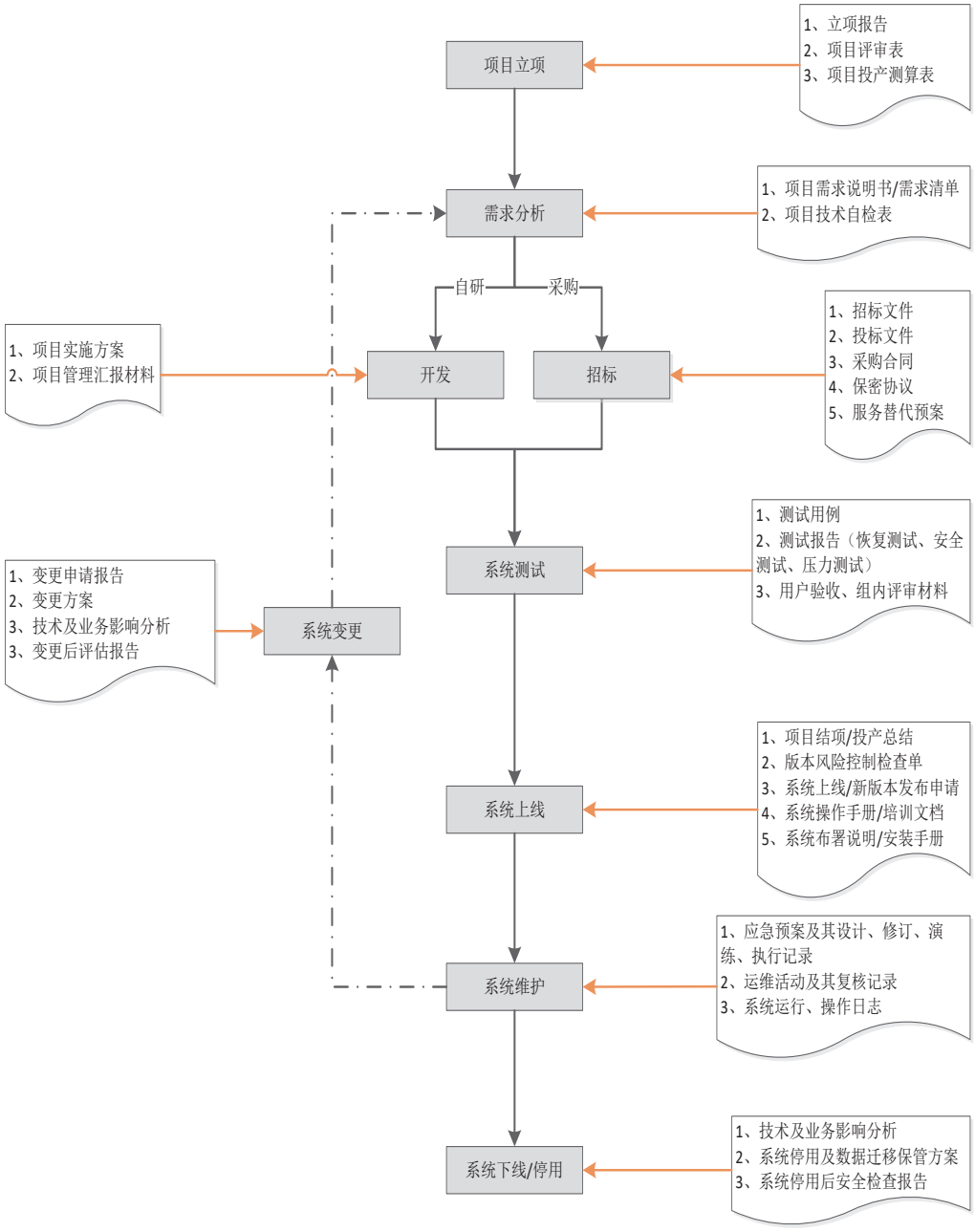


图 3：SDLC 文档留痕

（二） SDLC 管理

SDLC 管理（Software Development Life Cycle）是指系统生命周期，是系统或软件从产生到销毁整个过程的管理。结合《办法》中关于重要信息系统的管理要求，本文以及本文所述 SDLC 涵盖 6 大模块：建设、测试、上线、变更、运维和停用（如图 2 SDLC 管理审计评价项）。

在具体审计评价方面，除《办法》、《审计规范》和《证券期货行业信息系统运维管理规范》（以下简称《运维规范》）相关合规要求，审计师也需要结合机构内部的文档管理要求考虑重要信息系统各阶段的留痕管理要求。结合实际的业务实践，以某机构的系统管理各子流程对应的留痕文档要求为例（图 3 SDLC 文档留痕），审计师可以对留痕文档的授权要素、时间要素等内容进行合规评价。

（三）权限管理

特权用户具备对机构关键信息资产的不受限访问、操作权限，具体包括创建 / 配置文件、定义其他用户权限等，对于非法机构来说是极具吸引力的目标。当特权账户遭遇非法获权或滥用时，机构的信息安全将面临巨大威胁。基于《办法》第三十二条和机构的内部控制的要求，本文认为针对内部特权用户的审计分为 3 个部分：权限管理机制、权限闭环管理和特权用户活动的监控及分析。

1、权限管理机制

权限管理应当考虑两种情况，机构系统权限管理是分散在各系统中的和机构内部具备统一的全权限管理系统。如内部的重要信息系统权限管理是零散、分配于各系统属主 / 管理员的，机构首先应当结合内部数据治理项目对现有重要信息系统权限进行梳理形成重要信息系统权限清单，并定期对该份权限清单进行维护，以确保重要信息系统运行始终处于机构自身控制范围。如机构本身具备身份统一认证的系统管理重要信息系统权限账户的权限，机构应当具备确保该系统能够及时获取到内部各系统的权限变更信息的机制，并定期对该机制的实施有效性进行控制测试、定期对重要信息系统权限清单作出检查及复核。权限清单应当至少包含用户、角色、部门关系、行为权限、域访问权限、权限变更记录等信息。

评价权限管理机制时，审计师应当自上而下地考虑权限管理的管理架构、制度规范、流程及风险 - 控制节点，以评价目前机构权限管理机制设计的合规性及有效性。控制机制的有效落实是机构内部治理及管理的生命力，在机构权限管理机制设计有效的前提下，审计师可以基于权限清单及数据资产清单实施控制测试，以评价机构内部人员及相关外部服务人员对于内部权限管理机制的遵循度。

2、权限闭环管理

在权限闭环管理审计方面（如图 4），审计师可以通过专业判断、抽样等方式针对重要信息系统的权限账户的“权限分配”、“权限审批”和“检查与核对”三阶段形成的管理留痕进行审计，管理留痕信息包括但不限于 OA、邮件、日志等信息。审计师首先应当对机构内部的权限分配机制作出合理性评价，如认为现有的权限管理机制是合理的，则基于现有的机制去考察机制的具体落实情况；如认为不合理，则应当提出相应的合规建议。

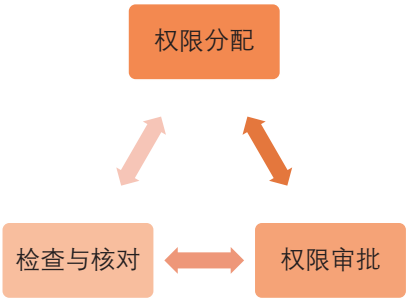


图 4：权限闭环管理节点

在具体的权限落实评价方面，审计师应当关注业务系统及管理系统之间的内在联动、逻辑关系（如图 5 系统权限关系），通过访谈关键人员、

测试管理系统等审计程序获取机构的人员变动信息，后以人员变动信息为切入点考察权限在业务系统、管理系统中的更新及适配情况，具体的权限审查应当重点关注用户 - 角色 - 权限的权责关系、业务流关键风险 - 控制点、权限修改 / 更新时间戳等信息。如存在账户未及时删除或权限不匹配等异常情况，应当了解异常情况发生的原因并评价其是否合理、是否符合机构内部的权限管理策略。

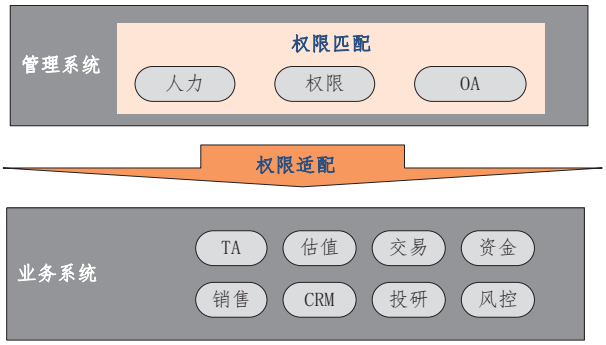


图 5：系统权限关系

3、权限监控指标

在机构的信息技术治理及管理进程中，可以考虑在管理系统中设计监控类指标对机构的重要信息系统账户行为进行监控和分析，管理系统具体可以包括内部审计系统、日志分析系统、内部控制系统等。

在监控指标设计方面，机构可以在内部的合规与风险管理框架内，基于对重要信息系统权限资源和数据资源的梳理，定性和定量地设立权限监控指标集。从机构内部治理及管理的角度，业务部门、运营部门等往往是内部风险管理第一道防线，合规部门、风险管理部门是内部风险管理的第二道防线，监察审计部门则是内部风险管理的第三道防线。考虑到机构内部关于风险第三道防线职责的分工，我们建议将该道防线的监控指标集分为三类：访问类、操作类和控制类（具体如图 6 权限监控指标），以实现在信息系统层面对审计的履职支撑。

（四）日志分析

在《审计规范》和《运维规范》中，全国金融标准化技术委员会对信息系统相关日志记录的最低保管时间提出了标准参考。另外，为保护经营数据和客户信息安全、防范信息泄露与损毁，《办法》中要求需要有日志记录等安全保障措施。合理的日志管理及分析机制可以协助有效管理证券基金经营机构内部的 IT 系统、设备以及应用的日常数据，提升日志管理水平，满足运营、安全和稽核部门要求。

1、日志范围

审计师在审计日志数据时应当注意日志数据范围需要覆盖网络设备、安全设备、操作系统、数据库、中间件、业务系统等（图 7 日志管理范围），在日志数据中需要记录系统、应用用户名、时间、所执行的命令操作及操作结果等详细信息。

图 6：权限监控指标

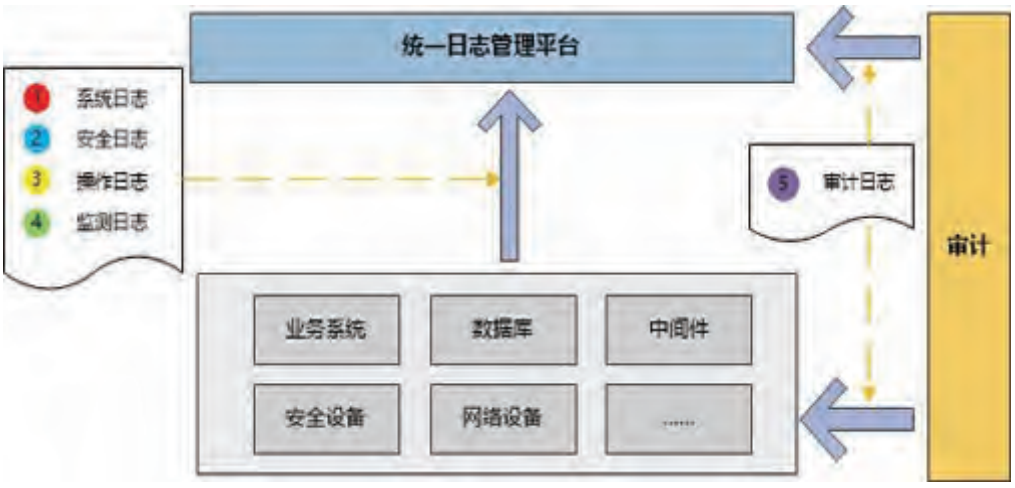


图 7：日志管理范围

2、日志管理

根据用途、来源和记录范围的不同，日志信息可分类为操作日志、系统日志、应用日志、安全日志等，在《审计规范》和《运维规范》中分别对相应日志的保管时间和复核频率加以规定，详见表 2。

系统属主或统一日志分析平台在信息系统进行基础环境变更或调整时需同步完善日志采集及相关的日志事件告警等配置，同时需要对日志中体现的潜在风险点或异常用户行为进行分析预警，根据重要级别制定 / 更新相应的处理方案和应急响应制度。

表 2：日志保管时间及复核规定依据

日志类型	法规依据	保管时限（法规）	复核频率
系统日志	业务系统：《证券期货业信息系统审计规范》A.3.4.3.3	15 年	/
安全日志	/	/	/
审计日志	主机：《证券期货业信息系统审计规范》A.2.4.1.3	6 个月	每季度一次
	业务系统：《证券期货业信息系统审计规范》A.2.4.2.4		
	数据库：《证券期货业信息系统审计规范》A.4.4.2.4		
操作日志	业务系统：《证券期货业信息系统审计规范》5.2.3-5.2.4	1 年	/
监测日志	《证券期货行业信息系统运维管理规范》5.3.8	1 年	/

3、日志检查

审计师需要定期检查和分析日志数据，包括异常日志信息统计、潜在风险或异常行为分析、日志告警趋势及其他相关信息。

机构应将核心交易业务系统监控日志及巡检记录，应纳入机构主动预防体系。相关系统属主应跟踪处理日志分析中发现的异常事件，定期总结分析，全面评估监控日志和操作记录，分析异常情况。

四、结语

行业及各机构信息系统具备体系架构复杂、连续及可用性压力大、一致及准确性要求高、海量数据集中化的特点，这些特点是重要信息系统审计实施困难点的重要影响因素之一。由于审计能力及审计资源的限制，目前的信息技术审计活动更多的是作为财务审计的能力支撑，而缺乏以业务流和风险控制点为导向的信息技术审计。

本文旨在为证券基金行业的重要信息系统审计提供新的实践思路，将原先浮于形式，一味应付公司及监管的信息技术审计乱象扫除，切实为推进证券基金经营机构信息技术建设及数字化转型等工作提前做好风险的预言预判及处置，为证券基金经营机构信息技术的长远发展保驾护航。

E 实践探索 xploration

“多对多池化”高可用集群技术的实践

梁德汉 周金成 王涛 / 安信证券股份有限公司
刘小亮 / 上海英方软件股份有限公司

本文叙述了券商单点应用（如银行等外部机构提供的应用）实践“多对多池化”高可用集群技术的探索，介绍了单点应用从传统的“一对一”冗余技术，到“多对多池化”高可用集群技术的创新与落地，生产上实现了在不改变现有系统架构、应用程序并满足固定 IP 的要求下，以较低的运维人力及软硬件投入成本，有效解决竖井式 IT 架构的服务器单点故障可能引发的诸多问题，切换方式从手动变为自动，故障转移时间从分钟级缩短至秒级。

一、实践背景

券商的业务应用系统，特别是对接银行、交易所、登记公司等的某些应用，在技术架构设计之初，没有充分考虑到冗余架构的需求，导致了架构单点的出现，而且这些单点往往呈现数量较多、不易改造的特点，券商运维团队需投入大量的人力物力来保障应用的高可用性。

本文将重点讨论解决此困境的新方案。下面，我们将从服务器的高可用集群技术展开叙述。

服务器高可用集群是将多个服务器集中在

一起同时进行同一种服务，在应用层看起来就像是一台服务器，是一种以减少服务中断时间为目的的高可用技术。近年来，随着一系列法律法规的颁布，监管单位对机构的业务高可用保护提出了新的要求，例如，根据 GBT 22239-2019《信息安全技术网络安全等级保护基本要求》的要求，需要满足提供重要数据处理系统的热冗余，保证系统的高可用性。

目前，券商机构在确保关键业务系统的高可用方面，采用的主要技术是通过服务器高可用集群，将单点故障对业务连续性的影响降到最低，并随着应用场景的发展和监管需求的提高，衍生了主备架构、双活架构等多种高可用模型。

高可用模型的选择，取决于券商机构的 IT 系统架构。在证券行业，多数对外机构业务的 IT 架构为竖井式业务模型，从上到下分为：上层应用（负责业务交互）、中层系统（负责处理应用数据）、底层存储（负责落地及存储数据）。针对此类业务模型，传统的集群高可用保护方式包括：

1.1. 基于共享存储的高可用架构

此架构由三部分组成：活动主节点，不活动备节点，共享存储。其中两台计算资源节点提供主备角色服务，通过 SAN 网络附加型存储作为数据存储的介质。代表方案有：Windows 故障转移集群、Linux 的 RHCS 集群。

架构优势：(1) 数据强一致性保障，只有数据落盘，数据丢失为零；(2) 共享存储，不需要数据同步的机制，数据不延迟；(3) RTO 时间为秒级；(4) 同机房、同机柜实现，易管理；(5) 支持自动化切换方式，无需过多的人工干预。

架构劣势：(1) 共享存储的同构成本高，远距离高可用接管成本高；(2) 数据存储介质唯一，存储故障风险大；(3) 仅支持一对一架构，不支持级联或一对多高可用。

1.2. 基于双机双柜的高可用架构

双机双柜是一种不依赖共享存储而实现的高可用保护架构，采用主备的高可用保护模式。在双机架构中，生产主机和备机具有物理层的完全独立性，应用、系统、网络和数据都是一

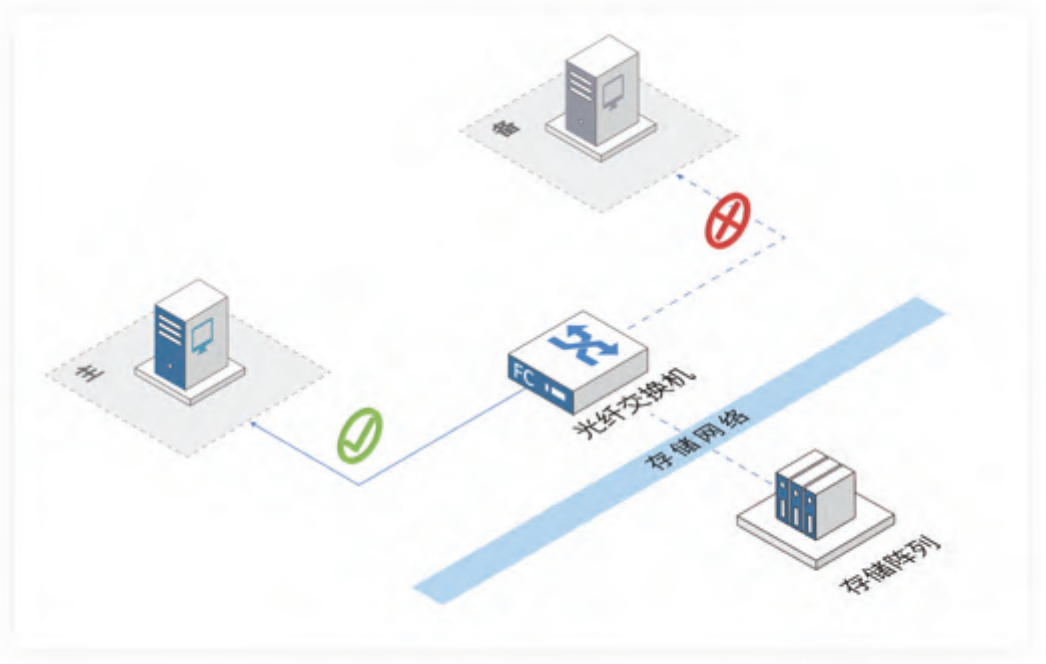


图 1：共享存储架构

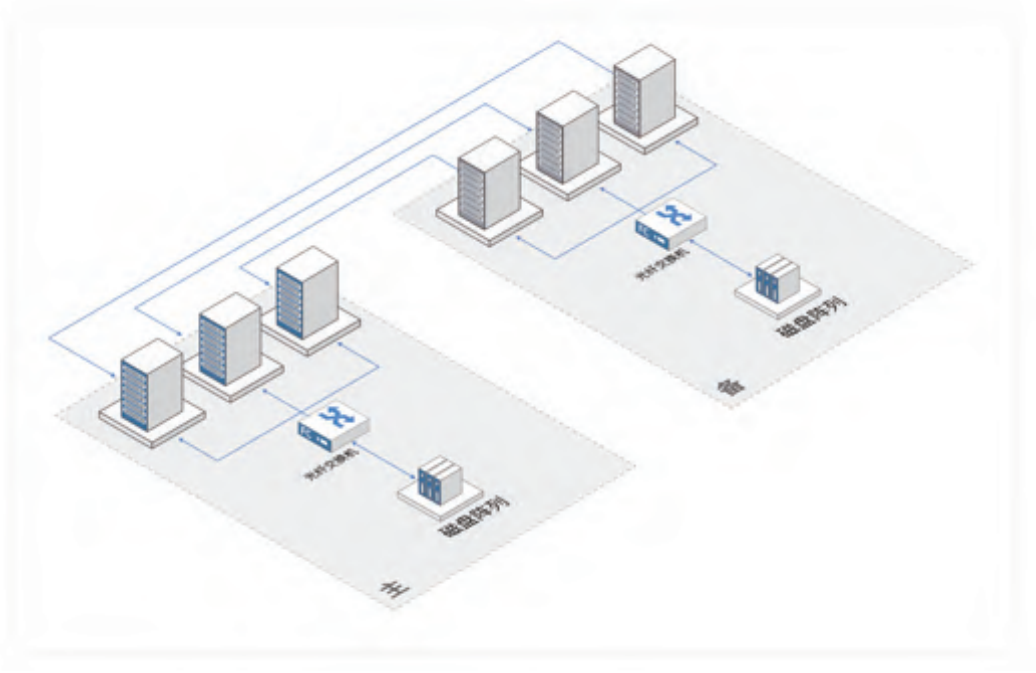


图 2：双机双柜的架构

式两份，彼此之间仅通过网络进行通信和数据传输，不需要 SAN 存储的介入。

架构优势：(1) 不需要共享存储支持，投入成本低；(2) 支持远距离高可用；(3) 双份数据，业务恢复快；(4) 支持自动化和定义组切换，支持深度进程、端口、日志等全方位监控；(5) 支持级联及一对多高可用架构。

架构劣势：(1) 采用异步数据复制，存在数据延迟；(2) 采用同步数据复制，对生产压力大；(3) 需要第三方软件支持，有较高的软硬件成本和运维人力成本。

综上，安信证券团队研究发现，这两种基于竖井式的业务模型的高可用架构，服务器高可用集群采用的是一对一的接管方式，即当生产端的一台服务器出现故障，备端的对应服务器可以启动接管，但通常会采用手动切换，或通过脚本方式启动接管，耗时耗力，且容易出错，不可控的因素多且成本较高。

下文将从安信证券单点应用的服务器“多对多池化”高可用集群技术实践、场景应用和总结展望等方面进行介绍。

二、“多对多池化”高可用集群技术实践

“多对多池化”高可用集群技术立项始于 2019 年初，主要针对单点应用服务器做高可用保护。该单点应用具有三大特点：一是种类及数量较多；二是无数据或少数据的存储需求；三是应用服务的 IP 和端口大多要求固定。根据这三大特点，安信证券做了双机双柜 (一主一备) 的集群高可用的技术方案研究，但是基于一对一的高可用方案并不理想，主要的原因有四个：

一是传统的业务架构，在设计初期没有考虑到后期冗余的需求，难以在短期内进行高可用保护架构的升级，或者新改造方案的成本太高，且存在新的风险；

二是证券系统对业务应用 IP 如何一直保持不变，特别是像银行、交易所、登记公司等的业务应用，IP 和端口是固定的，在这种要求下用传统的方式做集群高可用保护，难度很大。

三是故障完成切换后，只能单点运行，系统的冗余性无法继续保持。

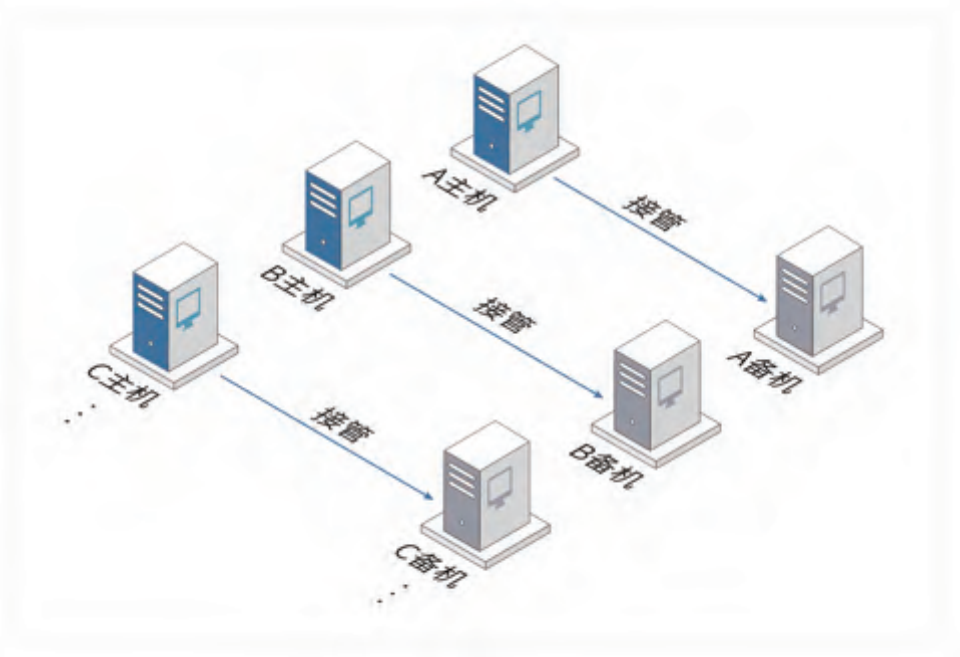


图 3：“一对一”接管架构

四是资源投入成本较大，采用传统的一体机高可用保护方式，硬件数量大，投资回报率低。

如何摆脱高可用性集群建立在“一对一”关系模型上，往“多对多池化”高可用集群技术的方面发展，是突破传统的竖井式 IT 架构集群高可用的关键。由于行业并没有可供参考的合适技术方案，项目团队经过对“一对一”、“级联”、“一对多”等高可用集群模型的研究，开创性地提出了“多对多池化”高可用集群技术保护的模型。

如图 3 所示：传统的“一对一”的关系模型，A 主机与 A 备机的高可用保护形态，数据同步、监控、切换执行等，只发生在 A 主机和 A 备机的关系中。

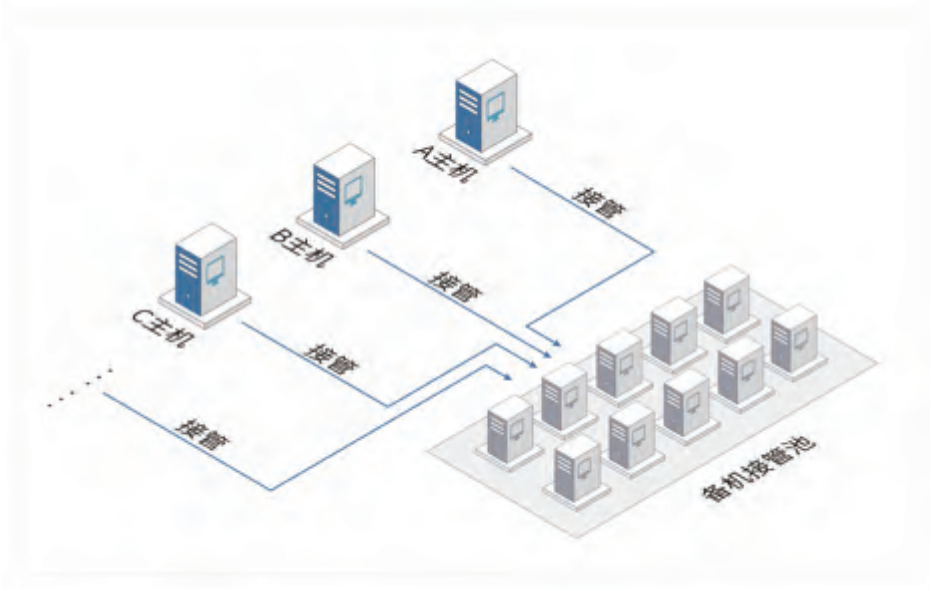


图 4：“多对多”池化接管架构

如图 4 所示：“多对多”的关系模型，则通过构建池化集群，将备端的多台主机组成资源池。当 A 主机发生故障后，由资源池中具备优先接管条件的备机提供接管，备机接管池的高可用保护不再是一对一的固定模式，它们具备了多级接管的能力。即当故障发生时，第 1 台备机如果不具备接管条件，将由第 2 台备机接管，**以保证故障完成切换后，系统的冗余性继续保持。**

2.1 “多对多池化”高可用的组成架构

架构是模型的基本组成，高可用的基本功能定义需要具备如下几个条件，来帮助高可用塑型：

角色：定义活动节点即主节点，不活动节点即备节点；

资源：服务器集群；

数据：一致性保障，确保主备数据一致性；

监控：识别故障及异常的技术方式；

仲裁：第三方防脑裂的判断基准；

VIP ：虚拟 IP，统一用户访问入口；

心跳：高可用切换的依据；

脚本：调用应用启动或实现应用的自动拉起；

告警：邮件或短信或其它通信方式告警；

组别：应用逻辑组或模块的顺序管理，主要应用在组切换场景中；

回切：恢复系统原状。

在服务器“多对多池化高可用集群技术”架构组成方面，由以下组件构成：

切换组件：手动切换、自动切换、不固定目标端角色切换及固定目标端角色，均在切换模块中定义；

资源池组件：由生产服务节点和备用资源共同组成；

数据同步组件：负责将数据从指定服务器同步至目标服务器；

VIP 组件：VIP 自动漂移至目标端，不需要人工干预；

应用联动组件：提供组联动切换功能，一键拉起业务组服务；

监控组件：识别进程、服务、负载、特定监控对象的运行状态等，为切换提供判断依据。

这些组件构成了控制中心组件，负责服务状态的监控，提供资源池配置与管理，并提供 WEB 人机交互的窗口功能。

图 5：“多对多池化”高可用组件架构

图 6：VIP 功能架构图

2.2“多对多池化”高可用的核心功能

在已生产部署的三方存管系统中，单点应用服务器“多对多池化”高可用的核心功能包括：

VIP 功能：在高可用的保护场景中，发生主备切换时，服务器角色会随着切换节点的变更发生物理主机的角色转换，由两台 IP 不一样的主机提供活动节点与不活动节点功能。团队在“多对多池化”高可用架构设计之初，考虑到了主机切换后 IP 地址访问异动的问题，故融入了 VIP 机制的功能，与通过独立负载均衡设备实现的 VIP 功能不同，“多对多池化”高可用采用的是一种软件定义 VIP 的功能。

当备端的资源池高可用构建完成时，所有主节点的主网卡会自动增加一个虚拟 IP 地址，并在当前网络环境下进行正常通信，即一张物理网卡中会有两个地址：第一级主地址和第二

图 7：仲裁服务器架构图

级辅地址。主地址负责集群管理通信，辅地址负责业务数据通信，彼此分工明确。当主节点发生故障，触发了 VIP 切换功能之后，集群将 VIP 辅地址在当前主节点中删除，在备节点进行 IP 地址的附加，从而实现虚拟 IP 地址的漂移。VIP 功能确保服务器切换时，业务的 IP 和端口保持不变。

仲裁机制功能：在高可用保护切换流程中，可能会出现主备同时为主，或者主备同时为备的现象，这个情况称之为脑裂，会造成双主或双备的情况发生，导致业务系统异常，高可用保护或者切换失效。为了防止此类现象发生，可通过独立的第三方主机构建仲裁服务器，站在中间立场来执行唯一性切换判断，解决脑裂、网络闪断造成误切换等问题。

例如，正常的运行模式下，主备双方具备同样的票数确保各角色固定。当心跳网络中断，但主机服务正常的特殊情况下，独立于心跳网络之外的仲裁端开始发挥作用。主备机高可用节点服务器会实时维持与仲裁端的监控同步，识别应用节点正常或异常状态及确定主备角色都各自固定，且唯一存在，当出现如下两种状况时，仲裁开始发挥作用：

1. 主节点正常，但主备心跳中断，依照优先级设置，主备不会发生切换，防止误切换发生，同时会下达指令，控制备端不进行角色接管，并通知运维人员介入处理。

2. 主节点异常，主备心跳中断，能与仲裁连接的节点会被判断为健康节点，不健康的节点会退出高可用性应用。当主备完成切换后，核对切换后的主机和接管后的备机，彼此主备角色不重复，具备唯一性存在。

需要说明的是，当心跳线正常的情况下，仲裁不会起作用。而故障切换的判断条件包括某个服务或进程是否正常，CPU 、内存、磁盘是否正常，网络是否正常，也可以使用脚本自定义其他的判断条件。以上条件多次确认发现问题后，再判断心跳连接和仲裁的情况决定是

图 8：服务器资源池架构图

否切换。仲裁节点节自身出故障时，高可用规则会有告警功能，如果不修复，当主节点故障需要仲裁时可选择两种策略中的一种切换或不切换。

仲裁节点还可以使用高可用方式部署，避免单点故障。部署方式如下：

仲裁机是 1:1 的高可用集群，仲裁 IP 为此集群的对外服务 VIP。集群高可用仲裁有三个场景：

1）如果仲裁 IP 不存在：整个高可用服务器池不能自动切换，可通过手动进行切换（告警发现，需及时修复）。

2）仲裁节点误切换无风险：仲裁 IP 在任何一个安装有代理的机器上，只要有网络权限，就可以起到仲裁节点的作用。

3）如果出现脑裂，当两个都为备节点，则 IP 丢失，运行模式如场景 1；当两个都为主时，两节点会通过优先级去比较，发现初始主节点的优先级比较高，就会回切为主节点，另外一个主节点则会降为备节点。

资源池化功能：服务器资源的池化，采用了云平台的主机池化管理架构，将所有主机和备机统一在一个服务器资源池中进行管理，通过标签定义的功能来识别应用类型。不同应用的主机和备机，构成了资源池的基础，不区分物理机、虚拟机等物理平台类型，所有的切换、接管、监控、策略均发生在可控的资源池中，通过逻辑组，将不同业务划分为不同组别，并进行分类标明，所有组均统一至池中进行管理，不同业务系统、不同组别、不同主机平台，基于节点服务统一进行高可用的切换管控。

在服务器资源池中进行弹性服务器节点的新增或删除时，不需要停机，不影响业务正常运转。例如，可以进行故障服务器的在线更换，即通过资源池的服务器调用，添加完新服务器后，再撤掉旧的服务器，如此可轻松应对设备维护、系统升级和补丁修复等。

迭代切换功能：常规的主备切换模式中，主节点切换至备节点就算一个主备流程完结。在“多对多”高可用的保护设计中，支持迭代切换，即 A 故障 B 接管，B 故障 C 接管，直到资源池最后一个空闲资源节点耗尽为止。迭代切换功能保障了系统的冗余能力最大化。

图 9：迭代切换架构图

图 10：不同接管率架构图

不同接管率支持功能：“多对多池化”高可用保护在面向多台主机保护时，可以根据应用服务保障级别进行服务器的弹性扩容，即通过设计接管率的配比来控制接管主机的投入数量，调配的合理范围一般设置在 10%-100% 内。

举例说明：在 10% 接管率的场景下，100 台生产服务器仅需要 10 台接管主机即可满足要求，即允许同时故障的最大主机数量为 10，**相较于传统“一对一”主备模式，可降低 90% 的服务器资金投入，承担额定故障率的高可用接管。**

代理服务器功能：为了减轻对生产系统的影响，“多对多池化”高可用集群技术方案只在服务器上装一个轻量化的代理程序，然后通过 Web 界面对服务器集群的高可用状态进行监控、展示和管理。

2.3 “多对多池化”高可用技术优势

与传统的“一对一”高可用保护技术相比，“多对多池化”高可用集群技术的优势包括：

化解发生切换后单点运行风险

一对一高可用保护技术，当发生主机切换后，会存在单点运行的风险，即接管设备再发生故障时，业务将中断。“多对多池化”高可用保护技术，通过服务器资源的池化管理，使得备用资源池的服务器大于等于 2 时，整个业务系统的高可用构成方式变成了迭代接管。

对应用系统完全透明

对于证券、银行等传统的系统架构，很多程序短期内无法进行迭代，而架构的修改风险大，成本高。“多对多池化”的高可用保护技术无需对现有的系统架构和程序做任何修改，只需在服务器上安装一个轻量化的代理程序，整个过程对应用系统完全透明。

高性价比

如果采用自动化的高可用接管方式，“一对一”高可用保护技术的方案，可以通过购买一体机的方式进行部署，但投入成本大，一台一体机就可高达二三十万，设备老旧更换的费用高，运维人力资源的投入较大。采用“多对多池化”高可用集群技术，在原有服务器设备基础上，通过高性价比的软硬件投入方式（备机集群资源池甚至可以使用新老设备混用的方

案），实现资源池内服务器资源的灵活调用，**在投入更小的情况下，整个系统的冗余能力反而得到提升。**

维护更可靠

一对一的高可用保护技术，当生产服务器出现故障时，运维人员会手动切换到另外一台服务器，然后再去维修故障服务器，整个切换过程风险大,不可控因素多。采用“多对多池化”高可用集群技术，服务器出现故障时会自动切换，整个系统保持高可用状态，且不会存在切换后单点运行的状态。同时，系统后台会推送告警信息，运维人员无需马上去维护，可以等到收市后的非繁忙时段再去处理，整个过程可视化、更可靠。

2.4“多对多池化”架构的可靠性测试

在上线部署运行之后，本项目组对“多对多池化”架构进行了测试，并于 2020 年 06 月 05 日得出了最新的《安信证券 7.2.21.20060520 版本软件测试报告》的成果，本次测试是根据 Testlink 上的测试用例，对“多对多池化”架构的高可用进行基本功能测试和异常测试验证。

测试场景是通过 2 台控制机组成一个高可用，主备共 109 个节点，根据同网段同类应用形成一个集群的方法，测试共组成约 15 个集群，具体的测试控制机和节点及环境如下：

控制机	Windows Server 2016 R2 Datacenter（2 台） 处理器 8 核 内存 8G
节点	Windows Server 2012 R2 Standard（109 台） 处理器 2 核 内存 4G

测试结果如下：

测试需求项	测试功能描述	测试结论
高可用规则	添加/修改高可用规则，修改生效，新的规则依照修改后的参数运行	pass
	自动切换：达到切换条件时，备节点接管并重新分配备节点	pass
	仲裁节点功能验证	pass
	数据同步(同步对象与同步策略)，定时将主节点的指定数据同步至备节点	pass
	监控对象：监控失败时，触发切换，备节点接管并重新分配备节点	pass
	脚本资源切换：主备切换时，主节点切换为备节点，执行释放资源脚本；备节点切换为主节点，执行获取资源脚本	Pass
	虚 IP 资源切换：主备切换时，主节点切换为备节点，释放业务虚 IP；备节点切换为主节点，添加业务虚 IP	pass
	主节点故障接管(断网、宕机、监控失败、挂起)，备节点接管并重新分配备节点	pass
	备节点故障接管(断网、宕机、监控失败、挂起)，主节点不变，重新分配备节点	pass
	本机接管启用与关闭：启用时，触发切换条件，备节点直接接管；关闭时，触发切换条件后，由备节点从资源池中选举一台机器接管成为主节点	pass
	固定中心节点规则，产生中心节点，添加集群和业务 IP	pass
	强制切换：强制将备节点接管为主节点，主节点切换为备节点	Pass（耗时 2 秒）
	日志收集：节点日志正常上传到控制机	pass
	控制机切换：集群虚 IP 可以正常切换到另一台控制机，使用虚 IP 登录的页面，规则节点信息不受影响	pass

测试需求项	测试功能描述	测试结论
异常测试	就绪状态，重启主节点机器，备节点接管并重新分配备节点	Pass（耗时 9 秒）
	就绪状态，重启备节点机器，重新分配备节点	Pass（耗时 2 秒）
	就绪状态，主节点宕机，备节点接管并重新分配备节点	pass（耗时 9 秒）
	就绪状态，备节点宕机，重新分配备节点	pass
	就绪状态，主节点断网发生切换后，网络恢复并将节点属性转为空闲节点	pass（切换 9 秒，节点状态转换 2 秒）
	就绪状态，备节点断网重新分配备节点后，网络恢复并将节点属性转为空闲节点	Pass（重新分配 2 秒，节点状态转换 2 秒）
	就绪状态，代理程序进程重启	pass
	就绪状态，仲裁节点故障，切换至备节点	pass（耗时 7 秒）
	就绪状态，控制机关机再开机	pass
	多条高可用规则运行，中心节点性能测试	pass

本测试报告还涵盖“服务器池集群、高可用模板、服务标签、权限分离、容量压测、安装卸载、代理程序管理界面”等，测试结论均为“Pass”，限于篇幅及重要性，不再此叙述。

2.5“多对多池化”场景应用实践

在安信证券落地实施的“多对多池化”高可用集群技术场景中，目前已完成三方存管业务、银衍业务的高可用集群部署，图 11 非常清

图 11：场景落地应用架构图

晰地描述了业务逻辑：在银行中间件集群中，生产部署了 9 台主用服务器和 4 台备用服务器；在三方存管业务的三方交易网集群中，部署 2 主 2 备；在银衍业务的三方交易网关集群中，部署 1 主 1 备。

在对单点应用的“多对多池化”高可用保护项目建设过程中，团队实现了以下的生产场景应用：

一是同网段同类应用形成一个集群，即以产品属性自由分配不同的应用集群。

二是多资源池的支持。根据业务应用特点，可划分多个服务器集群的资源池，降低彼此之间的影响。

三是 VIP 功能管理。硬件资源的管理可通过物理 IP 实现，对于有 IP 绑定需求的服务器，系统会增加虚拟 IP 来承载业务入口；如果没有就不会增加，可以节省更多的 IP 资源。

四是服务器资源的维护。在无需停机的情况下，可将系统应用切换至资源池的服务器集群中，待新机替换旧机，或系统升级、补丁修复后，再切换回来，整个过程无需停机，对业务不影响。

三、总结与展望

项目落地实施上线半年多，生产上发生了一次物理机故障，应用程序自动切换到备端服务器资源池的物理设备上恢复业务连续性，整个处理过程秒级切换，无人为干预，客户无感知，整体效果符合预期。

证券是一个对业务连续性要求非常严苛的行业，一分钟的停机事件都会带来巨大的经济损失和影响。“多对多池化”高可用集群技术在成本、可行性、风险管控、创新性方面，对整个行业突破“一对一”冗余保护技术的枷锁，起到了抛砖引玉的良好作用。未来，安信证券更多的单点应用系统将采用此方案进行多对多的高可用保护。

东方证券服务治理建设实践

樊建 / 东方证券股份有限公司 fanjian@orientsec.com.cn
杨子江 / 东方证券股份有限公司 yangzijiang@orientsec.com.cn
胡长春 / 东方证券股份有限公司 huzhangchun@orientsec.com.cn
舒逸 / 东方证券股份有限公司 shuyi@orientsec.com.cn

微服务架构是近几年受到各行业广泛追捧的技术之一，微服务架构具有轻型化、便捷化、敏捷化等特点，不仅能够适应业务创新和变化的需要，而且易于维护、变更、升级，契合当前证券业务发展的需要。然而向微服务架构转型也面临不少挑战，东方证券通过构建统一的服务治理框架，打造了一个多语言、多协议、可视化、灵活配置的服务管理平台，支持东方证券企业架构向以微服务为核心的现代架构转型。文本将介绍东方证券 gRPC-Nebula 服务治理框架与星辰服务治理平台的建设成果与实践经验。

一、概述

近年来，随着市场客户和业务量的不断攀升，以及互联网金融的兴起和金融科技的发展，各证券公司都制定了数字化转型的战略目标。为了把握新一轮数字化技术革命浪潮，企业信息系统架构正在不断升级变迁，很多企业内部的传统软件系统都开始向微服务架构转型，通过服务拆分、降低系统耦合性，提供更加灵活的服务支撑。

随着研发人员对系统进行解耦和拆分，对大

量微服务实例进行有效管控、提升系统运行时的服务质量变得非常困难。在此背景下，东方证券为了顺应互联网＋时代的潮流，响应快速更新的业务需求，迫切需要以统一、服务化的思路来进行系统建设，建设服务治理平台，通过分析服务调用关系及拓扑结构、优化服务质量、制定服务协议规范，达到新建系统与已有系统统一服务治理，实现轻应用（业务为导向，实现业务应用敏捷构建，及时响应市场需求）、重平台（将数据和核心应用转化成平台服务，成为整个架构的核心）、服务化（构建核心服务网络，简化应用开发与部署）的整体企业架构转型目标，实现应用全生命周期管理。

二、微服务架构

（一）单体架构

传统信息系统多采用单体架构，单体架构应用把所有的功能都打包在一个独立单元中，并当做一个整体来开发、测试和部署。Java Web 应用就是典型的单体架构应用，项目被打包成一个 WAR 包部署在同一个 WEB 容器中，其中囊括了数据访问层的 DAO 对象、业务逻辑层的各模块、表示层呈现的 UI 等功能。单体架构的优势是开发、调试、部署简单方便，在业务发展初期，信息系统的规模较小，使用传统的单体架构可以有效地支撑业务的发展。然而，随着业务的爆炸性增长，应用系统规模不断增大，单体架构将给业务系统的开发、维护、部署带来巨大的问题。

第一，开发效率持续下降，庞大的代码规模和错综复杂的业务耦合大大增加了研发新功能的难度，开发者不仅要掌握自己负责的模块，还需要了解整个应用系统的逻辑，否则修改代码后可能会引发冲突；第二，持续迭代存在障碍，任何一个非核心功能的小修改都需要重新部署整个项目，使得系统运维中与发布相关的风险显著增加；第三，系统可靠性变差，传统的单体架构将所有的应用都部署在同一个进程中，如果应用中某个接口发生故障，将会影响整个系统正常提供服务的能力；第四，扩展性先天不足，单体架构的应用只能在一个维度上进行扩展，但是不同的模块可能有不同的资源需求属性，例如有的功能是计算密集型，有的则是 IO 密集型，由于它们运行在一个实例中，无法对特定模块进行扩展；第五，技术僵化无法重构，各个业务使用的技术栈不得不与整个应用的技术栈捆绑在一起，很难更新 SDK 版本或使用新的技术框架。

（二）微服务架构

由于单体架构已不能适应现代企业信息系统的需要，近年来微服务架构被广为推崇，并在越来越多的证券公司中得以实践和落地。微服务架构是由传统的单体架构逐渐演化而来，将大型单体应用按照业务功能设计拆分成多个能独立运行、职能单一的服务，与其他服务之间通过统一协议进行通讯。

微服务架构可以很好地解决单体架构下的诸多问题：第一，将巨大的单体应用拆分成颗粒度更小的服务，服务内逻辑简单、高度内聚，易于开发和维护；第二，各个微服务独立部署，功能修改后可以针对特定部分进行发布，使得各个微服务系统能够持续化部署，加快了迭代的速度；第三，当单个服务系统出现故障时，只需要将出现故障的服务下线修复即可，不会导致整个系统的级联故障；第四，可根据不同微服务系统的访问量和资源需求，动态的实现横向扩展和纵向扩展，这大大的提高系统的利用率；第五，各个研发团队可以根据自己的需求选择编程语言和技术栈，具有更大的灵活性。

虽然微服务架构有着明显的优越性，但是证券公司普遍存在的系统异构化问题也给微服务架构的落地带来了巨大挑战。

首先，业务接口标准不统一，管控风险大。券商行业的核心系统由传统供应商组成，以东方

证券经纪业务核心系统为例，分别由金仕达、新意、恒生、顶点、同花顺等厂商架构组成，SPX、T2、REST 等多种类型服务接口存在于东方证券企业内部，多业务协同适配问题突出，服务多样性对同步、异步、流式数据等都提出了技术需求，统一化难度大；缺乏有效的关键业务流量控制技术手段；全局化平台协同与调度困难重重，缺乏全局视角对内部服务进行统一化管理。

其次，自研系统上线面临诸多困难。随着金融科技的深入发展，证券行业纷纷开始进行自研核心系统，但因为缺乏统一的开发框架，各业务研发团队在具体开发过程中除了业务分析之外，还需同时会关注非常多的技术细节，如依赖服务接口对接，开发语言技能，灵活可扩展架构支撑，客户服务治理保障，对外服务协议选型，服务故障定位，请求流量控制，服务安全配置，配置管理，流量管控等。

理想情况下，业务人员关心业务梳理和场景定义，开发人员把相关业务转换成服务定义，借助代码生成工具自动化生成接口代码，最后根据业务实现接口内部逻辑。由开发框架和外部工具负责架构扩展性、服务治理、配置管理等一系列非业务相关的功能实现，实现业务和框架的解耦，提高开发效率。

三、东方证券服务治理方案

完善的服务治理方案是微服务架构应用稳定运行的基石，东方证券凭借在服务治理领域的技术沉淀和实践经验，在 gRPC 框架基础上新增服务治理特性，建设了 gRPC-Nebula 服务治理框架和星辰服务治理平台，从而实现企业内部及外部服务的统一化管理，构建服务调用关系及拓扑结构，优化改进服务质量，图 1 展示了东方证券服务治理项目的总体架构。

东方证券服务治理方案主要包括如下几个模块：

（1）注册中心

注册中心是一个分布式、高可用的配置维护系统，用于服务的注册和订阅，它存放着所有的

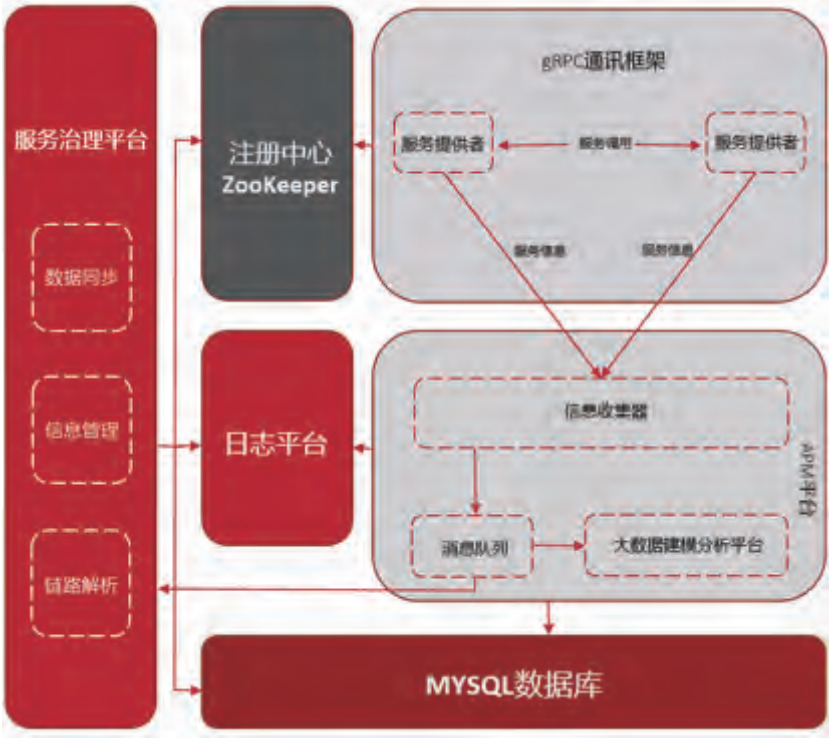


图 1：东方证券服务治理项目总体架构

服务描述信息以及服务接口信息。在微服务框架系统中，服务和接口的数量非常庞大，同时由于系统的动态调整，服务运行的实例数量也是动态变化的，注册中心通过将服务统一管理起来，可以有效地优化服务消费者对服务提供者的感知和管理，避免硬编码地址信息。

（2）服务消费者（客户端）

服务的消费者，通过注册中心交互获取服务注册信息，基于服务注册信息发起对服务端的调用；同时，采集调用端信息发送到数据处理引擎中进行分析处理，为调用链分析提供客户端数据。

（3）服务提供者（服务端）

服务的提供者，通过注册中心对外发布服务信息，响应消费者的服务调用请求；同时，响应控制台等发起的配置管理操作，对服务质量、安全策略、数据收集等进行配置管理。

（4）信息收集器

独立的部署的服务，收集服务调用过程中在服务提供端和服务消费端产生的服务调用、服务响应、服务异常、服务时间、调用链路、内部队列长度、安全事件等信息，收集后统一发送到数据处理引擎进行处理。

（5）数据处理引擎

数据处理引擎，对信息采集器发送过来的信息事件流进行实时分析处理，处理操作包括性能统计、依赖分析、阈值告警、相关聚类、状态跟踪、可用性分析等；同时，数据被存储到性能管理数据库，用于进一步的分析操作。

（6）服务治理平台

服务治理平台汇总了服务治理领域的运行数据和管理系统，它是全公司服务治理的综合门户。在服务治理平台，可以查询每个微服务的实例信息、接口信息、服务状态、依赖和被依赖关系、数据统计、服务调用追踪记录等数据。同时平台支持黑白名单、流量控制、权重配置、主备配置、上下线状态的管理功能，支持调用量、服务质量、故障事件等对象的监控告警功能，是管理人员对微服务进行集中式管理的中控界面。

四、gRPC-Nebula 服务治理框架

（一）技术方案

东方证券调研了目前比较流行的开源微服务框架，包括阿里巴巴的 Dubbo、Facebook 的 Thrift、Google 的 gRPC 以及从 Spring Boot 框架发展而来的 Spring Cloud 项目，它们都具有较好的连通性、健壮性、伸缩性和拓展性，但 Dubbo 和 Spring Cloud 框架不支持多语言，Dubbo 在开源社区最近才重新启动更新

最终我们选择以 gRPC 框架为基础，研发 gRPC-Nebula 服务治理框架。相比其他几种框架，gRPC 有以下优势：(1) 全面的多语言支持，gRPC 支持多种语言，包括 C、C++、Java、Python、PHP、Node.js、C#、Objective-C、Go、Ruby 等。目前券商网上交易和核心交易系统均是 C++ 架构，而其他自研系统大多是 Java 和 Python 架构，gRPC 能有效解决服务的跨语言调用问题；(2) gRPC 在 Google 和广大开源爱好者的大力支持下，目前社区活跃、更新频繁，已在全世界多家大型科技公司内投入生产；(3) gRPC 使用 Google 开源的 Protobuf 3.0 协议定义接口服务，Protobuf 是一种平台无关、语言无关、可扩展且轻便高效的序列化数据结构的协议，广泛应用于网络通信和数据存储，技术人员对 Protobuf 的熟悉有助于 gRPC 技术在企业内的推广；(4) gRPC 的传输使用 HTTP/2 标准，支持同步、异步、双向流，支持 SSL 和自定义鉴权，支持 iOS、Android、Windows、Linux 等平台，可以简单地实现客户端到后台的多路复用与 RPC 调用。

相对于原生 gRPC 框架，gRPC-Nebula 服务治理框架引入了 ZooKeeper 作为注册中心，如图 2 所示，融合了服务注册发现、负载均衡、黑白名单、动态分组、集群容错、流量控制等服务治理机制，本章节后面的部分将详细介绍这些服务治理机制

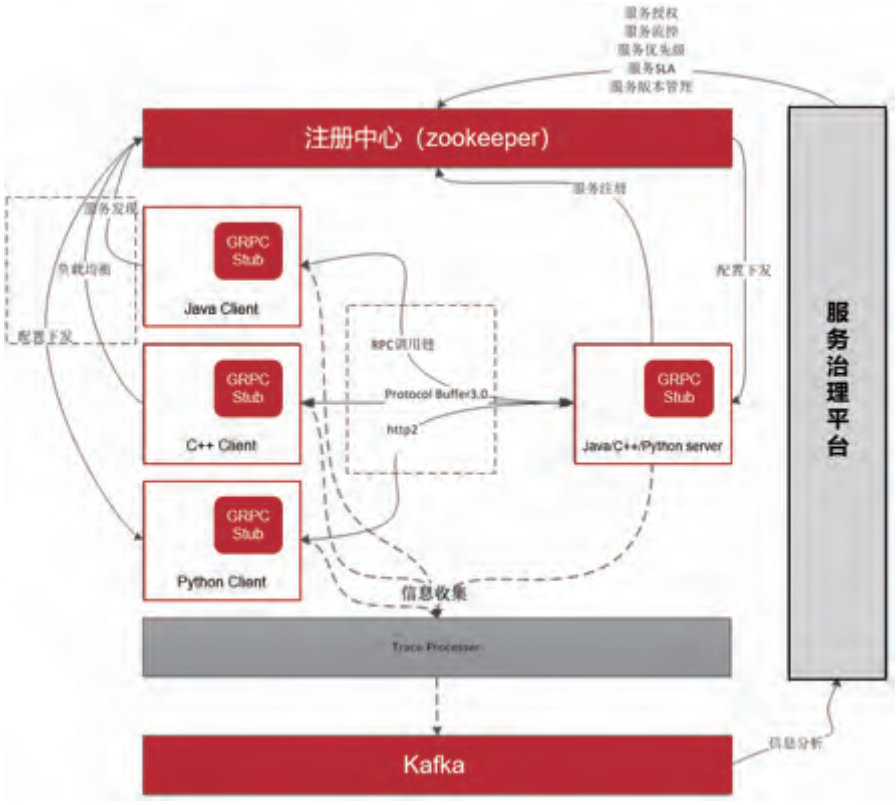


图 2：东方证券 gRPC–Nebula 服务治理框架架构图

表 1：多框架性能测试比较

线程数	Dubbo	原生 gRPC	gRPC-Nebula
1	5867	6059	5452
10	47402	48525	48389
20	87701	86522	86044
50	152902	151536	150091
100	156172	155811	154721

的技术方案。

我们分别对 Dubbo、原生 gRPC、gRPC-Nebula 框架进行了性能测试，如表 1 所示，gRPC-Nebula 框架的性能仅比 Dubbo 和原生 gRPC 框架低 1% 左右，满足高性能服务框架的需求。

2019 年 6 月中旬，东方证券宣布开源 gRPC-Nebula 服务治理框架，目前社区已建设了社区决策委员会，初期拟设 7 名委员，含 1 名委员会主席，设有专人进行 GitHub 代码的跟踪、维护、解决。同时，委员会会定期组织研讨和常态化沟通、社区技术交流、协调开发力量进行社区开发、社区筹款、审议版本 maintainer 的版本路标和功能集、社区委员会选举等工作。社区将秉持金融科技创新，对外技术输出的原则，致力于成为行业内首家基于 gRPC 可治理 RPC 框架下的开源社区。

（二）服务注册发现

服务注册发现是服务治理领域最核心的机制，服务提供者在启动时向注册中心注册它提供的服务信息，服务消费者向注册中心获取服务提供者的地址列表，如图 3 所示。gRPC-Nebula 服务治理框架使用 ZooKeeper 作为注册中心，具有以下特性：

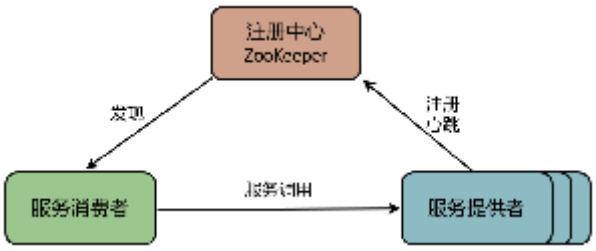


图 3：服务注册发现机制

（1）具备高可用性。当注册中心任意一个节点宕机时，服务能够自动切换连接到其它正常的节点上；当注册中心全部宕机时，只影响新服务的发布与已发布服务的下线，不影响服务的正常运行，服务消费者会使用本地缓存的服务地址列表继续调用。

（2）保证数据一致性。所有服务消费者同一时刻从注册中心不同节点获取到的服务地址列表是同一份数据，不能出现读或写数据的不一致。ZooKeeper 使用 ZAB 协议作为其数据一致性的核心算法，是具有严格的访问控制的能力的分布式协调服务。

（3）服务变更主动推送。服务消费者只需要在启动时向注册中心拉取一次全量服务地址列表，其后向注册中心订阅相关服务的变更事件，一旦服务地址列表发生变更，注册中心会主动将变更的内容推送给服务消费者，服务消费者即时调整调用的服务地址。

（4）实时感知服务状态。注册中心与服务建立长连接，通过心跳检测机制，能够周期性地检测服务的健康状态，当服务进程意外终止或服务器宕机时，注册中心能够立刻向服务消费者推送服务下线的通知，实现故障隔离。

（三）服务路由

在生产环境上，微服务都是多实例部署，服务路由决定了服务消费者如何从服务地址列表中选择服务提供者进行调用。gRPC-Nebula 服务治理框架的服务路由以下三大机制构成：

（1）负载均衡机制

gRPC-Nebula 服务治理框架提供了四种负载均衡算法可供选择：随机策略、轮询策略、权重配置优先策略、一致性哈希策略。

随机策略即随机地选择服务提供者进行调用；轮询策略即遍历服务地址列表，每次调用时依次选择一个服务提供方进行调用；权重配置优先策略可根据配置文件或管理门户对每个服务节点配置的权重比来选择服务提供者；一致性哈希策略中，相同参数的网络请求总由同一个服务提供者处理，当某个服务提供者的节点宕机时，系统基于一致性哈希算法来选择其他的节点。

（2）黑白名单机制

通过设置服务端实例的黑名单、白名单，可以动态实现请求流程的转移以及服务端实例的访问控制。如果将某 IP 加入一个服务的黑名单，部署在这个 IP 上的服务消费者无法从注册中心获取到这个服务的地址列表。

（3）动态分组机制

每个微服务实例都有一个分组属性存放在注册中心，分组属性既可以通过配置文件预先设定，也可以通过管理平台动态配置。通过分组一个微服务的集群可以被划分为多个集合，服务消费者可以按优先级调用某几个特定分组的服务，动态分组机制可以灵活实现同机房调用和业务隔离等场景。

以同机房调用场景为例，在数据机房安全性越来越得到重视的今天，多机房灾备方案被各类企业广泛使用，但是跨机房调用的高耗时可能造成系统的容量降低。如图 4 所示，假设所有服务实例均部署在 A、B 两个异地机房，服务消费者希望优先调用属于同机房的服务提供者，使用 IP 段定义机房的策略灵活性和扩展性不足，服务分组策略可以有效满足这一需求。例如将机房 A 的服务提供者定义为 a 分组，将机房 A 的服务消费者配置成优先调用 a 分组的节点，同时机房 B 的服务也进行类似配置。这样，机房 A 的服务消费者会优先调用机房 A 的服务提供者，避免高耗时

的跨机房调用，当 Server1 和 Server2 全部宕机时，机房 A 的服务消费者会把请求自动切换到机房 B 的 Server3 和 Server4 上。

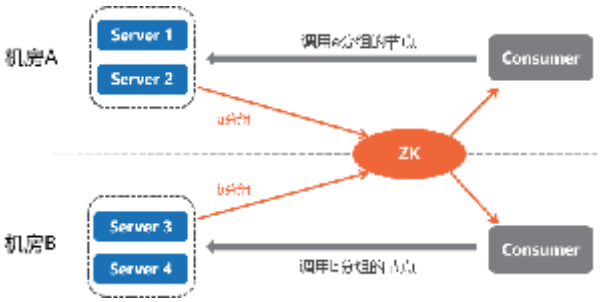


图 4：同机房调用场景

（四）集群容错

当服务提供者无法正常为消费者提供服务时，如连接被拒绝、请求超时、后台服务异常等，服务框架需要进行集群容错，重新进行路由选择和调用，gRPC-Nebula 服务治理框架支持快速失败（Failfast）和失效转移（Failover）两种策略：

快速失败是指如果服务提供者返回异常，消费者不用重试直接报错。这种策略适合一些非核心的服务，可以为重要的核心服务节约宝贵的资源。

失效转移是指当服务调用异常时，重新进行路由选择，查找下一个可用的服务提供者来发起新的调用请求。当调用一个节点连续多次失败或在一段时间内失败率超过限制时，框架认为这个节点当前已不适合再对外提供服务，服务消费者会将其从服务地址列表中剔除，保证一段时间内都不会调用到这个异常节点。这个机制的目的是降低系统对网络抖动的敏感性，不会因为一次偶然的调用失败而调整流量分配，保持服务器负载的稳定性。

（五）流量控制

历史上券商核心系统事故都是由流量冲击引起的，当网络流量瞬间爆发性增长时，对服务器 CPU 和 IO 资源的抢占会造成系统出现瓶颈，服务错误率迅速上升，此时上游或用户的重试行为又进一步加大了网络流量，最终使得服务彻底崩溃且长时间难以恢复，这即是雪崩效应。为了防止雪崩，需要对服务调用过程进行控制，通过一些策略削减流量，保证后台服务接收的请求在可承受的范围内。

gRPC-Nebula 服务治理框架通过设置请求数和连接数限制，动态实现对各服务接口的流控管理。请求数限制即当单位时间内请求数过多时，丢弃多余的请求；连接数限制即控制每个 IP 连接到服务提供者的连接数，当连接数达到阈值时，服务提供者会拒绝建立新连接的请求。

五、星辰服务治理平台

（一）建设目标

由于业务的实际需求和技术发展，开发部门和供应商通常会根据需要选择不同的微服务框架，呈现多样化选型。如何管理好这些服务，成为研发和运维部门需要面对的问题。如果可以将这些框架和服务对接到统一的服务治理平台，将可以大大降低协作开发的成本，并提升整体的版本迭代效率，因此东方证券星辰服务治理平台的建设目标包括：

（1）通用治理能力：引入中间层设计，兼容多框架的通用治理能力，采用分发器和治理组件协调工作统一多框架通用治理能力，由分发器下发任务至不同的治理组件，由理组件完成平台纳管多框架，完成分发器下发治理任务。

（2）平台自服务：平台本身采用微服务架构及容器平台集成，提供更深度的治理功能，提供平台应用生命周期、组件部署管理、灰度、弹性和统一配置支持。

（3）多框架兼容的应用管理：兼容基于 gRPC、Spring Cloud、Dubbo 三种微服务框架，帮助客户快速部署或迁移微服务应用。

（4）业务服务架构防腐化：通过服务注册中心对服务强弱依赖进行分析，结合运行时服务调

用链关系分析，梳理不合理的依赖和调用路径，优化服务化架构，防止代码腐化。

（5）快速故障定界定位：通过服务调用链日志、服务性能 KPI 数据、服务接口日志、运行日志等，实时汇总和分析，实现故障的自动发现、自动分析和在线条件检索，方便运维人员进行故障诊断。

（6）服务微管控：运行态服务治理，包括限流降级、服务迁入迁出、服务超时控制、智能路由、统一配置等，通过一系列细粒度的治理策略，在故障发生时可以多管齐下，在线调整，快速恢复业务。

（7）服务生命周期管理：包括服务的上线审批、下线通知，服务的在线升级，以及上线、下线，自动弹性伸缩，资源扩容。

（二）功能模块

星辰服务治理平台包含以下功能模块：

（1）服务治理

星辰服务治理平台支持对 gRPC、Spring Cloud、Dubbo 三种微服务框架进行管理，如图 5 所示，支持查询注册中心维护的服务实例信息，支持通过控制台配置注册中心、访问控制、主备、分组、黑白名单、流量控制、熔断等信息。

（2）服务地图

服务地图将项目与项目、服务与服务之间的调用关系和调用量通过拓扑图的形式进行展示，如图 6 所示。系统架构师可以从服务地图提炼出全公司的核心系统拓扑图，找出不合理的环形调用链；运维人员可以从服务地图掌握核心系统所依赖的上游系统，并给予核心系统同级别的重点保障。当预期面临流量猛增时，服务地图还可用于流量预估，因为从客户端等入口预估流量最准确，进而可以沿着服务地图下沉计算出各个微服务分摊到的流量，协助后台系统制定扩容预案。

（3）链路跟踪

在微服务架构中，一个用户操作涉及到多个微服务的协同才能完成，在业务调用链路上任何一个微服务出现异常或者网络超时，都会导致失败。通过链路跟踪，我们可以很方便的看到每个请求各个环节的耗时以及异常，帮助我们对系统进行优化。星辰的链路跟踪功能基于 Google 的 Dapper 论文实现，在系统入口接收用户的请求后，会为用户的请求分配一个 TraceID 用来唯一标识调用链。TraceID 会跟随远程调用消息传递到下游服务，直到整个链路的节点都拥有了 TraceID，通

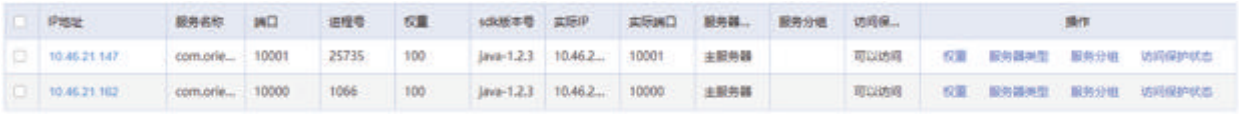


图 5：星辰服务治理平台实例列表

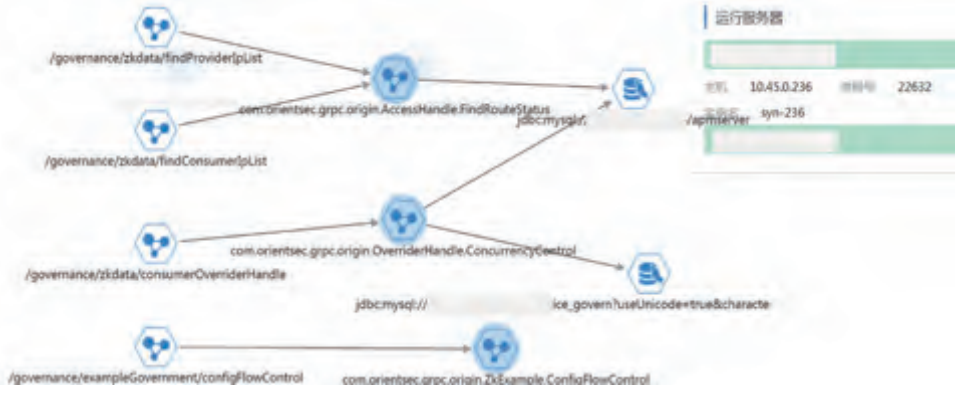


图 6：星辰服务治理平台服务地图

过 TraceID 可以串起这个请求的完整调用链路。

（4）文档中心

文档中心对 ProtoBuf 格式接口定义文件进行自动解析，提供技术人员查询各服务注释信息与接口定义的功能。未来我们还将强化文档中心的交互沟通功能，增加问答与评论功能，打通各服务上下游的交流渠道。

（5）统计分析

统计分析模块支持对服务、实例、端点的性能监控，监控指标包括响应时间、可用性、吞吐量等；支持数据大屏，全景展示当前所有服务的运行状态；记录服务响应时间，并展示响应时间最长的数个服务，即慢服务列表。

（6）告警中心

告警中心支持基于监控数据的告警规则设置，并以自定义的方式发出告警通知。

六、总结

本文探讨了微服务架构领域的关键技术，并详细介绍了东方证券服务治理平台的建设成果与实践经验。东方证券在企业架构层面制定了大中台战略，旨在通过架构转型为公司科技工作的长远发展打下坚实基础。作为大中台战略的核心基础设施，服务治理平台的建设，是公司提高金融科技核心竞争力的重要突破。gRPC-Nebula 框架和星辰服务治理平台已在财富中心、交易中心、账户中心、产品中心、行情中心、东方赢家 APP 和自研机构交易产品等数十个项目中得到应用，随着平台生态的不断优化和发展，未来将在内部全面推广，服务于更多用户和产品线，为公司服务治理规范和体系化架构建设做出更多贡献。

SD-WAN网络在证券行业的探索与实践

陆颂华 杨亚斌 乐剑平 / 海通证券

SD-WAN 的出现，有助于以较低成本拥有更加可靠和更高带宽的广域网线路，为企业在分支互联区域替代广域网专线提供可能。海通证券以“高速互联，品质体验，简易运维”为目标，携手华为共同打造的新一代 SD-WAN 网络，一方面能够为海通的业务发展夯实网络基础设施，促进营业网点智能化等战略的落地；另一方面也可以为科技创新，如海通混合金融云等应用的深化提供技术保障。本文以 SD-WAN 技术的特点和海通证券业务、技术发展需要为切入点，详细描述了 SD-WAN 在海通证券的应用实践，作为行业内率先使用 SD-WAN 技术的探索者，希望能为证券行业网络新技术应用提供一定的借鉴。

一、引言

网络作为联接从云端到边缘的基础设施，在企业数字化转型中扮演着重要角色。

SD-WAN（软件定义广域网）的出现，以低成本的互联网宽带在一定程度上代替了较低流量、价格昂贵的专线来完成企业站点互联，加上安装运维管理的简易性、全局流量调度和可视分析等特性，极大地降低了企业 IT 投入开支。

SD-WAN 是将 SDN 技术应用到广域网场景中所形成的一种服务，这种服务用于连接广阔地理范围的企业网络、数据中心、互联网应用及云服务，可以帮助用户降低广域网（WAN）的成本开支并提高网络连接的灵活性。

近两年，SD-WAN 已经成为网络领域的新风向。根据近期 IDC 针对 SD-WAN 的相关调研，

95% 的企业已经或将在两年内使用 SD-WAN 技术，而 42% 的企业已经完成部署。其中，中国 SD-WAN 应用始于 2017 年，在 2018 年快速增长，2019 年 SD-WAN 市场增速超过 130%，市场规模接近 7000 万美元，使用涉及金融、零售、制造、互联网、媒体、政府、医疗、能源、电力、教育、交通和服务等多个行业。[1]

SD-WAN 市场的快速发展得益于它具有的如下优势：

1、安全可靠。SD-WAN 可在广域网流量传输的过程中对流量进行加密，并通过对网络进行分片来提高网络安全性，确保数据安全。

2、高性价比。SD-WAN 可以让企业有效地利用互联网、4G、MPLS 专线等多种方式构建高性价比的广域网来满足业务需求，而不用担心维护空闲的备份链路。

3、组网灵活。SD-WAN 路由器可以组合多个广域网连接的带宽，并根据企业不同分支机构的地域分布、规模大小以及实际网络需求等，提供灵活的组网方式。

4、部署敏捷。SD-WAN 可以使 WAN 服务快速部署到远程站点，而不需要 IT 人员去部署，真正实现终端设备的零接触部署、零接触维护和策略自动管理。

5、智能选路。为了避免链路故障带来的网络风险，企业往往会订购多个互联网链路，SD-WAN 通过控制器监管链路、网点、应用和设备情况，可基于应用动态智能选择最优路径。

6、云网融合。SD-WAN 让企业对网络的管理更加便捷，通过集中监测、分析网络性能和当前状态，促进公有云、数据中心、分支机构和物联网之间的任意互联。

二、SD-WAN 网络应用背景

海通证券是国内成立最早、综合实力最强的证券公司之一，其经营网点遍及全球 14 个国家和地区，在境内拥有近 340 家证券及期货营业部，服务近 1500 万客户。面对如此复杂的经营网络和庞大的客户群，加上行业业务办理线上化、交易渠道互联网化趋势明显，使得公司的业务开展更加依赖高可靠、大带宽、高性价比和低延时的广域网络。然而，采用传统广域网络支撑公司未来发展面临如下挑战：

1）流量需求快速增加，专线扩容性价比低：智慧营业部引入智能身份识别、高清视频、语音交互等创新金融服务，使营业网点业务流量激增；而传统网络大多采用专线构建营业部到总部的多级广域网络，不仅带宽小（仅 2~10Mbps）且价格昂贵，采用传统专线带宽扩容方式无法既经济又有效的满足智慧营业部的业务流量需求。

2）单链路适配性差，差异化需求无法满足：随着智慧营业部和金融云战略的推进，网点对网络流量的需求不断增加，而不同应用对链路的带宽、时延、抖动、丢包等需求并不一样，证券交易业务需要低时延；高清视频业务更依赖大带宽，而语音交互则对丢包更敏感。传统单链路专线无法针对特殊应用设置专门的可靠性参数，无法针对不同应用的要求提供相应的链路保障。

3）运营模式不灵活，部署过程不可见：海通证券拥有大量轻型营业网点（即 C 类营业部），经营定位与运营模式灵活，需要顺应周边商业环境变化快速设立或撤并；传统专线及 IPSec 的连接模式，往往需要分派专业工程师进行现场调试、安装、部署，这一过程动辄需要数周甚至一个月，一方面无法满足轻型营业网点快速变迁的需求，另一方面，部署过程对需求方的透明，也增加了需求方在等待过程中的焦虑。

4）运维复杂度大，故障排查效率不高：传统网点必须配备专业的网络工程师，在现场进行网络调试和故障定位，一个小的故障往往就涉及数千行命令行的配置，运维过程复杂、效率低下。

为此，海通证券携手华为，在行业内率先启动 SD-WAN 新一代网络建设，以“高速互联，

[1] IDC 咨询：SD-WAN 江湖——原有网络市场格局将被彻底打破，2020 年 3 月 24 日

品质体验，简易运维”为目标，为公司混合金融云和智慧营业网点构建高速互联的通道。

三、海通证券 SD-WAN 应用实践

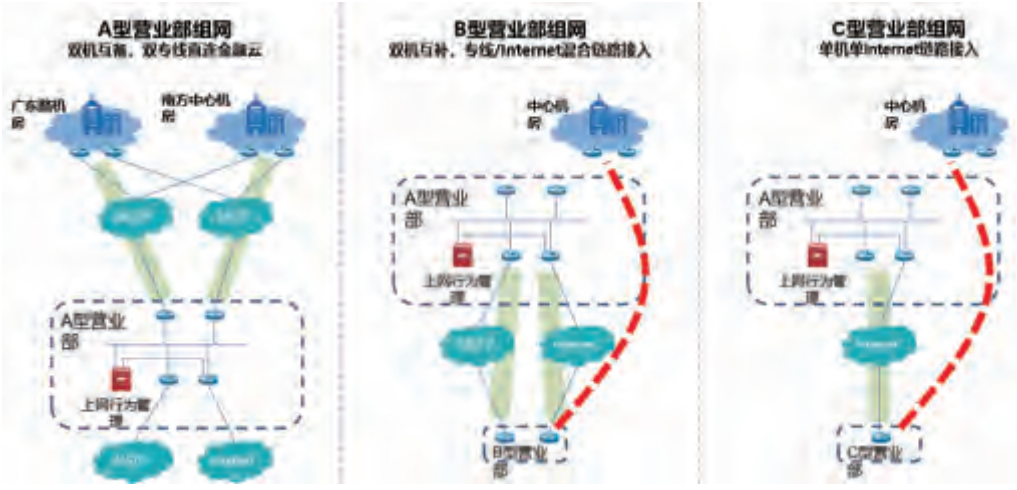
海通证券从 2019 年开始建设 SD-WAN 网络，目前，主要探索实践以下工作：1、灵活引入 MSTP、MPLS VPN、Internet、LTE 等多种链路专线，针对不同营业部的规模和业务特点，构建高性价比的 SD-WAN 高速互联通道；2、多链路互备，应用级智能选路，保障关键证券业务体验；3、营业部无需专业人员上门，设备即插即用，网络分钟级开通；4、全网集中可视管理，全流程自动化管理，提升运维效率；5、通过异地容灾，控制集群的双重冗余设计，与海通证券“两地三中心”的数据中心整体布局相融合，确保 SD-WAN 控制器高可靠，实现全网集中可视管理，运行状态一目了然。

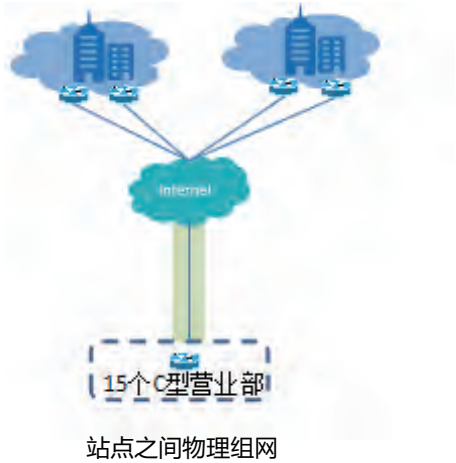
3.1 混合链路接入，高性价比 SD–WAN 互联通道

根据 SD-WAN 网络特点以及实际业务需求，海通证券构建了全新的广域网络架构。

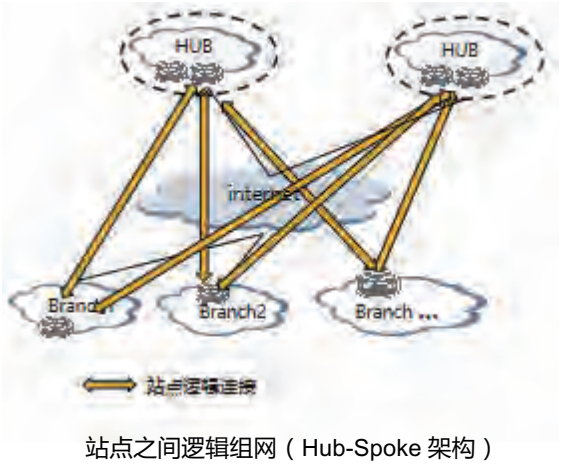
过去，海通证券的广域网络是树形结构，采用 SD-WAN 方案后，借助其丰富的组网模型，除 A 型营业部全部采用专线直连机房的扁平化组网模式外，B 型和 C 型营业部均可根据业务诉求，灵活选择与中心机房建立扁平化或者层次化组网。

基于证券行业业务特点，流量主要以分支





站点之间物理组网



站点之间逻辑组网（Hub-Spoke 架构）

网点到中心机房的上下行流量为主，很少涉及分支网点互访，所以组网架构推荐采用 Hub-Spoke 模型，分支 Spoke 站点只与 Hub 通信，无法直接互访。出于可靠性考虑，Hub 点（即中心机房）选择双设备，Spoke 站点按需可单可双，当前海通已部署的 C 型 Spoke 站点为单设备单出口部署。Hub 侧每台 CPE 各连一条运营商链路，Spoke 侧可根据网点规模，采用单链路或者双链路 Internet 接入。

如上图是海通证券 C 型营业网点采用单条 Internet 接入数据中心的组网模型。

3.2 应用级智能选路，关键证券业务体验可保障

传统网络中，设备无法根据线路质量进行实时切换调整，但 SD-WAN 则不同。海通应用级智能选路包含三种选路方式：

1、基于链路质量的智能选路：当链路质量低于 SLA 阈值（包括丢包、时延、抖动），网络会自动进行流量切换，转移到质量好的网络上去；

2、基于链路带宽流量负载分担：通过应用识别，确定不同业务流带来的链路带宽流量负载，为不同的业务流选择适合它的链路带宽；

3、基于链路带宽利用率的智能选路：如当主链路带宽利用率高于阈值上限 70% 时，切换低优先级应用到备链路，当主链路带宽利用率低于阈值下限 50% 时，回切低优先级应用到主链路。

3.3 零接触开局，网点分钟级网络开通

在部署 SD-WAN 过程中，海通证券充分体验到了 SD-WAN 部署的 ZTP（Zero Touch Provisioning，零接触部署），利用 USB、邮件、DHCP 等多种方式灵活开局，30 分钟内快速上线网点业务，大大加速了证券业务的快速扩张。

以通过邮件方式开局为例。在部署 SD-WAN 时，总部的 IT 工程师只需要提前做好配置数据，然后将配置通过邮件的方式，发给站点开局人员，该员工即可通过邮件内加密链接，完成设备配置部署，不再需要专业 IT 人士到场进行配置安装，实现网点设备“即插即用”，网点应用“随叫随到”。

3.4 多维度可视运维，降低运维复杂度

海通证券 SD-WAN 在组网完成后，建立了全方位的监控网络，以确保网络正常运转。具体包括站点性能监控、站点间性能监控、应用性能监控等，同时借助设备日志管理和故障定位，及时发现问题、解决问题。

站点性能监控：通过站点维度的监控，管理员可以监控全网站点的健康度情况，根据指定站点的日、周、月不同时间纬度查看设备性能、链路吞吐率及带宽利用率、链路质量、应用的流量和质量、访客的流量和应用等；

站点间性能监控：通过站点间维度的监控，管理员可以监控两个站点之间日、周、月不同时间纬度的链路质量，流量，带宽使用率及应用的流量和质量等；

应用性能监控：通过应用维度的监控，管理员可以监控应用质量的分布情况，应用使用流量排名，应用的客户端数排名，指定应用的日，周，月流量和质量趋势等；

设备日志管理：管理员通过 SD-WAN 控制器部署 CPE 的日志策略后，CPE 上报日志给日志服务器，日志服务器对 CPE 日志进行采集，存储和分析；管理员可分别通过 SD-WAN 控制器或日志服务器对控制器及 CPE 的日志进行查询；

故障定位：管理员通过告警和监控及时了解到网络任一位置异常后，通过简易的故障定位手段即可快速发现问题根因，并由此制定相应的修改策略和解决措施。

3.5 初步成效

在海通证券建设智慧营业网点过程中，SD-WAN 网络的部署，优化了海通证券现有网络架构，提升网络基础设施的运行和保障能力；积极探索了证券行业控制器异地容灾网络；通过双重冗余方案设计、控制器异地容灾和站内集群技术保障网络业务的高可靠性。SD-WAN 在海通证券的部署，其成效具体表现为以下三个方面：

1、加速营业部部署速度，支持业务快速扩张。传统营业网点的开通和部署需要 0.5 天，而即插即用的 SD-WAN 只需要 30 分钟，就可以实现网点新建、搬迁、扩容过程中业务快速上线，支撑业务战略；同时，多层 SD-WAN 网络架构，可以对不同层级划分不同租户进行网络管理，实现更加灵活的营业部组网模型，满足不同规模营业部的组网需求。

2、保障证券关键业务高品质体验。SD-WAN 可以快速识别知名应用和证券私有应用，合理规划证券交易，通过多链路负载均衡，让证券交易数据始终运行在质量最优的链路上；充分利用 MSTP，Internet 主备链路，避免主链路拥塞，备链路空闲。不仅提高了应用体验，也提升了带宽的利用率。

3、实现了证券可视化运维，应用、链路、网点、设备状态信息一目了然。其中，基于 GIS 地图的网络监控，应用、链路、网点、设备状态可视；全网业务时延、抖动、丢包率等关键指标实时呈现；网络自动巡检，精准告警信息邮件通知。

四、SD-WAN 实践中的思考

通过 SD-WAN 网络在海通证券的部署实施，我们有如下几点思考：

一是要看全局，重规划。要充分了解广域网技术尤其是企业现有网络体系结构的发展路径，

并站在未来公司业务发展方向来谋划公司网络基础设施的布局。同时，随着服务器虚拟化、云平台的应用、移动端流量激增、智慧网点建设等新趋势，企业广域网络流量需求增加迅速，企业的网络架构需要持续优化。提前做好规划，通过对公司当前网络状况以及使用情况进行深入分析，针对不同分支机构的网络需求进行评估、设计和规划。

二是要分阶段、重实效。在实施和运维过程中，要选择适合企业自身特点的网络解决方案，SD-WAN 组网绝非一蹴而就，而是要根据不同网点，不同区域的业务特点，分步骤，有节奏的逐步推进；如何实现跨地域的大规模组网，如何实现与非 SD-WAN 站点的互联互通，如何确保业务安全，都是实施过程中关注的重点和难点。

三是要重合作、提能力。SD-WAN 是一种复杂、强大且不断发展的技术，企业需要选择合适的合作伙伴来共同规划、实施、运营和完善，在选择 SD-WAN 的合作伙伴时，应重点关注其服务和支持能力以及生态圈的广度和深度。同时，加强自身团队建设，做好人员储备，以应对新型组网对网络架构和人员专业能力的新要求。

五、未来展望

SD-WAN 逐渐成为企业网络迭代的必然选择，海通证券将持续完善网络体系机构，优化网络运营流程，加快 SD-WAN 覆盖；同时，结合云计算，5G 和 AI 等技术，为智慧海通构建高速互联通道。

5G、AI 等技术与 SD-WAN 的融合将带来如下收益：一是 5G、SD-WAN 接入超大带宽、超低时延、海量连接的能力，可极大提升证券网点的接入灵活性和联接带宽，也为轻型证券网点向更多的城市和社区延伸提供了可能；二是 5G 网络的切片功能与 SD-WAN 结合后，能利用其应用识别能力，对业务流进行识别、分类，实现更精准的智能选路，以应对低时延的证券业务需要和智能在线客服的业务量激增；三是融合 5G、SD-WAN 的一体化设备将重塑广域网边缘基础设施，灵活集成路由、安全、广域优化等丰富的边缘特性，降低分支业务部署的复杂度；四是与 AI 技术结合的 SD-WAN 自动化运维技术，将极大简化网络运维复杂度、减少故障定位时间，为分布式、跨地域的证券公司分支机构的网络无人化运维提供可能。

探索个性化TTS技术在券商智能外呼的应用

柯善超 晏强 周朝阳 陈妍 / 光大证券信息技术总部

一、背景

近年来，智能外呼在金融行业应用日益广泛，比如银行账款催收、保险信息核实，金融产品推荐等场景都有成熟的案例，极大缓解人工回访的业务压力，提升回访效率和覆盖率。智能外呼综合应用人工智能技术，包括语音识别 ASR，文字转语音 TTS 以及自然语言处理 NLP，其中回访语音是客户体验的第一印象，热情专业的客服声音能够给客户带来很好的体验，提升客户的兴趣促进回访完成。在智能外呼的场景下，语音播报包括三种方式，全录音播放，全 TTS 合成，以及录音 +TTS 混合模式。我们对上述 3 种方式都深入实践，总结如下：

1、全录音播放模式。录音与人声效果一致，客户体验最佳，但是证券回访场景的话术中含有许多变量的词语，比如客户的姓名，资金账号，股票代码，产品代码，风险等级等，无法使用录音实现全部变量的覆盖。

2、全 TTS 合成模式。TTS 合成技术随着深度学习的应用，已经在资讯播报，语音导航等场景中广泛使用，能够达到媲美人声的效果；但是在外呼场景中仅能使用 8KHz 的采样率，声音效果与真实录音存在一定差距；同时商用

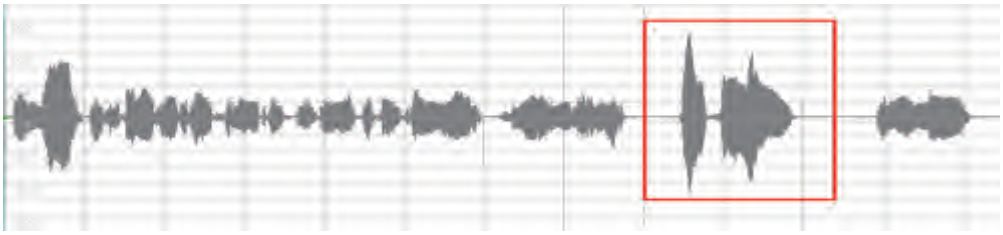


图 1：录音 +TTS 音频波形示例图

TTS 声音缺少个性化，没有针对客服场景进行定制化训练。

3、录音 +TTS 混合模式。为使得录音和 TTS 衔接自然，我们选取了与 TTS 合成声音色最为接近的客服进行音频录制。然而实际使用中仍存在明显差异。图 1 音频波形图可以看到，合成部分前后停顿过长，加上机器感音色，实际听感不佳，容引起客户的反感。

针对以上问题，最成熟的解决方案是通过 TTS 供应商专门为证券外呼 8KHz 场景定制开发，但是成本较高；为了给客户带来更好的外呼体验，实现低成本且个性化的外呼语音，我们在研究当前主流的 TTS 合成模型基础上，结合实际情况，通过一系列优化方法克服了样本数量少、合成音质差等困难，有效解决券商回访场景下变量合成的问题，打造特色的 TTS 语音效果，推动 AI 自主研发的能力提升。

二、TTS 技术概述

原始音频通常具有很高的时间分辨率，以 16k 采样率为例，1 秒音频包括 16,000 个采样点。音频数据变量维度过高且前后存在依赖关系，如果直接将文字编码与音频进行建模，难度太大。目前主流的语音合成系统主要分为两部分：前端和后端。如图 2 所示，前端系统先解析输入的文本，提取语言学特征，映射为音频的中间形态——声谱图；后端即声码器再实现声谱图到声音的转换。其中，声谱图是音频波形数据在低维特征空间的表示，即音频特征，比如频谱、梅尔频谱等，能够保留音频主要的信息。

深度学习自出现以来，在语音合成各个环节得到了广泛应用，与传统的统计参数建模方法相比，极大幅度地提高了模型的精度。

从文本到音频特征这个阶段，2017 年谷歌发布了端到端的模型 Tacotron，利用序列到序列（Sequence-to-Sequence，Seq2Seq）模型实现了文本到原始频谱的转化；2018 年新版本 Tacotron2 发布，引入注意力机制（attention）的基于循环 seq2seq 的特征预测网络，并增加 post-net 来精调 mel-spectrogram。2018 年百度发布了 DeepVoice3，提出了一个全卷积的特征到频谱的架构，能对一个序列的所有元素完全并行计算，并且使用了递归单元使其训练速度比类似的架构极大地加快；可以同时学习数千种不同人的语音，快速实现个性化的声音。

音频特征转换为语音阶段，Griffin-Lim 声码器是早期基于信号处理提出的，根据原始频谱预测相位，进而重构时域信号，无需训练，对于音频的质量有很高的要求，真实环境下的噪声叠

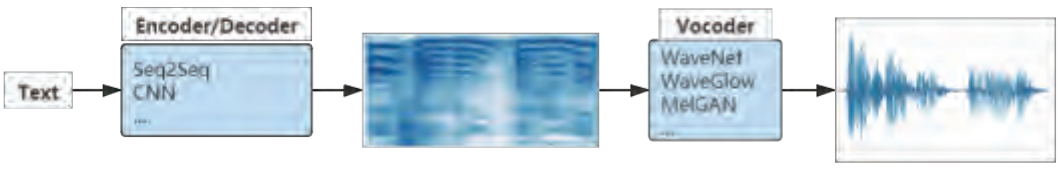


图 2：TTS 合成流程

加对于语音生成效果有很大影响。2016 年，谷歌引用了图像领域广泛应用的自回归模型思想，提出了 WaveNet。这是一个语音后端生成模型，能生成原始音频信号。但 WaveNet 采用了自回归结构，需要依据之前采样点来生成下一个采样点，因而生成速度较慢。同样采用自回归生成模型结构的，还有同年发布的 SampleRNN，该模型是在帧的层面生成语音，提高了生成速度。2018 年 10 月，英伟达提出了一种基于非自回归的神经网络模型 WaveGlow，用于语音合成时能使用 GPU 进行加速，极大提升了预测速度。但是，WaveGlow 的训练成本过高，需要耗费很长的时间。2019 年，MelGAN 模型的提出，首次将 GAN 用于原始音频的生成，在没有额外的蒸馏和感知损失的引入下仍能产生高质量的语音合成模型，而且参数量只有 WaveGlow 的 5%，大大提高了训练和预测速度。

三、TTS 探索实践

我们对上述主流的深度学习方法进行了大量文献调研与实践比较。经过分析，百度 Deep Voice3 具备从文本到语音输出完整的结构，同时与其他端到端模型相比，具备独有的多说话者训练的优势。多说话者训练，能学习不同人之间共同的发声特点，实现参数共享，对于小样本数据集训练有很大优势。因此，我们将 Deep Voice3 模型作为基础研究方向，进行本地化的改造实践。

首先我们在互联网上收集了部分开源语音数据集，进行 Deep Voice3 模型的全流程测试，在约 25 小时的 22KHz 训练音频数据下，单颗 GPU Tesla P40 训练 72 小时后，合成效果能够实现发音清晰、流畅的效果，部分语句发音抖动厉害，与音频原声存在差距，整体效果基本满足需求。接下来我们将数据集替换为客服人员的 8KHz 录音数据，时长约 3 小时，同样环境下训练 72 小时，合成效果较差，声音模糊，部分变调，发音颤抖，较多尖锐的爆破音，基本不可用，针对上述问题，我们对 Deep Voice3 模型做了深入剖析。

模型主体包括编码器、解码器以及转换器三个部分：

编码器将文本转换为模型内部学习表示，首先将文字发音转化为内部向量表示 h_e。嵌入向量 h_e 通过全连接层后，利用卷积网络提取文本中的上下文关联信息，创建注意力键值向量 h_k。最终，按照 $h_v=\sqrt{0.5}(h_k+h_v)$ 计算出注意力向量 h_v。

解码器是利用注意力机制将编码器结果以自回归方式解码成低维特征（梅尔频谱）。输入

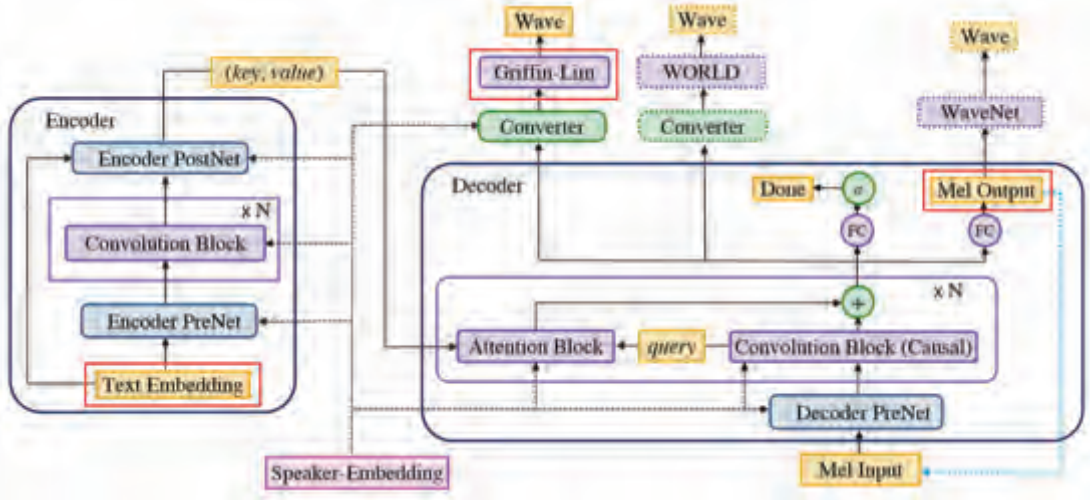


图 3：Deepvoice3 原始模型架构及关键点

一组音频真实梅尔频谱样本，利用多重带有整流线性单元（ReLU）的全连接层进行预处理。然后利用因果卷积与注意力块获取编码器隐藏状态。这里，注意力机制使用查询向量（解码器的隐藏状态）和编码器单位时间键值向量计算注意力权重，并结合权重计算值向量的加权平均值，得到上下文向量。最后通过全连接层，输出下一组梅尔频谱预测结果以及一个二值的终值预测。

转换器负责将解码器最后一层隐藏层输出转换为声码器的参数。对于不同的声码器，转换器使用的损失函数不同。默认使用的是 Griffin-Lim 声码器。

最终模型的训练目标 Loss=Liner_loss + Mel_loss + Done_loss，分别是线性频谱损失，梅尔频谱损失以及结束符预测损失。通过对模型主体结构每个环节的分析，我们定位到造成预测效果差的 3 个主要原因。

1、音频数据的差异。标准的语音训练数据采集环境为专业录音棚及录音软件，录音环境的信噪比高，人声音色、音量、语速一致稳定；而客服自主录音环境较差，在办公室采用手机、录音笔等录音工具，录音周期长语速前后有差异，录音采集时长较短，样本数量不足。

2、声码器 Vocoder 的差异。百度 Deep Voice3 默认采用 Griffin-Lim 作为声码器，利用 frame 之间相位的约束来实现迭代收敛，可以在缺乏原始相位信息的基础上利用频谱重构出语音信号。通过实验，直接输入原始音频的频谱数据生成音频可以发现，Griffin-Lim 声码器合成的声音效果与原声差距明显，用预测的频谱则声音进一步变差，需要用选择更好的声码器。

3、Mel 频谱差异。通过对比 Decoder 的 Mel 频谱输出，我们发现预测的频谱与真实频谱的分布差异明显，真实频谱清晰细腻，而预测的频谱轮廓基本一致但是模糊不清，对应的声音同样也有沙哑不明亮的感觉。

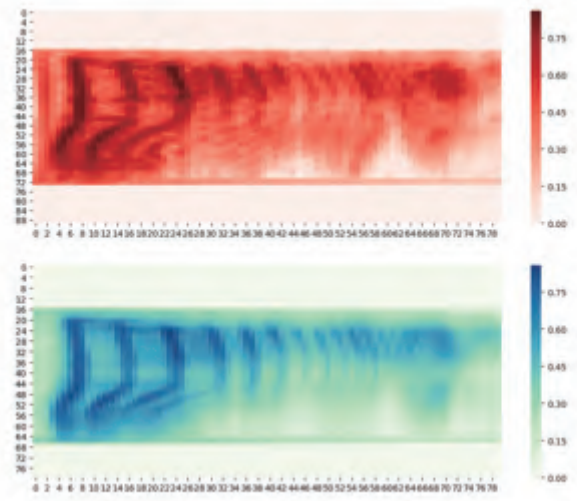


图 4：Mel 频谱对比图

四、TTS 优化方案

4.1 训练数据优化

因不具备录音棚的环境，我们尝试了多种简易录音方式，手机，录音笔，会议拾音设备等，最终采取笔记本 + 外置话筒 + 专业录音软件形式，44KHz 采样率，录音音质较好；同时为了保证录音的前后一致性，我们通过脚本对音频进行初步筛选，将语速明显偏慢和偏快的片段剔除；每个音频前后留白保存一致，去除发音错误的数据，保证训练数据的质量。

经过上述步骤处理后，训练数据大约 9000 条，约 3 个小时，样本量较少，因客服录音时间成本较高，我们借助商用 TTS 合成了数倍的语音数量，作为不同的发音人员联合进行训练。为保障样本的多样性和覆盖率，我们从真实的变量数据中筛选出来数十万条数据，数据筛选规则主要为两点，第一为覆盖面，确保覆盖所有的发音，第二为典型性，发音重复次数较高的变量优先选择。

经过训练数据的优化处理和外部数据补充，模型训练过程更为平稳，loss 稳步下降。

4.2 Vocoder 声码器选择

声码器的训练较为独立，训练数据直接

Vocoder 模型	原理	参数量（百万）	效果对比
WaveNet	基于概率和自回归的原始音频生成模型，可以直接学习音频序列，根据序列前面的采样点预测下一个采样点。	24.7	5.6
WaveGlow	基于流的生成模型，通过学习输入样本的条件概率分布，通过概率密度变化生成音频。	87.9	6.3
MelGAN	基于 GAN 的音频生成模型，模型由生成器和判别器组成，生成器以对抗训练的方式拟合音频。	4.26	7.3

通过音频就可以得到，我们实践对比了几款主流的基于深度学习的 Vocoder，包括 WaveNet, WaveGlow 和 MelGAN 三种。

传统声音效果评估通过平均意见评分（MOS）进行测量，让听众评估待测试音频的听觉质量，要并按照标准评分方案评分。券商在实际评测中较难让听众熟悉一系列规则，我们采取更简单的测试方法，将 3 个模型各生成 10 个语音片段，与 10 个真实录音混合在一起，每名听众对 20 个语音片段判断是否真实录音。最终取 10 个合成语音中判断为真实的数量作为评判依据，可以看到 MelGAN 的合成效果最佳，平均 7.3 个语音被识别成真实语音，效果逼真。

Vocoder 训练最大的难点就是 GPU 算力资源要求较大，在单颗 GPU 的模式下需要数周乃至数月的训练周期，给模型的验证和比较带来较大挑战。如下图展示了 MelGan 在 GPU Tesla V100 下，需要 14 天跑批可以产生清晰明亮的声音效果，28 天 500 万次迭代训练后声音更加细腻逼真。

4.3 模型结构优化

4.3.1 文本特征的选择

DeepVoice3 提出的架构可以将各种文本特

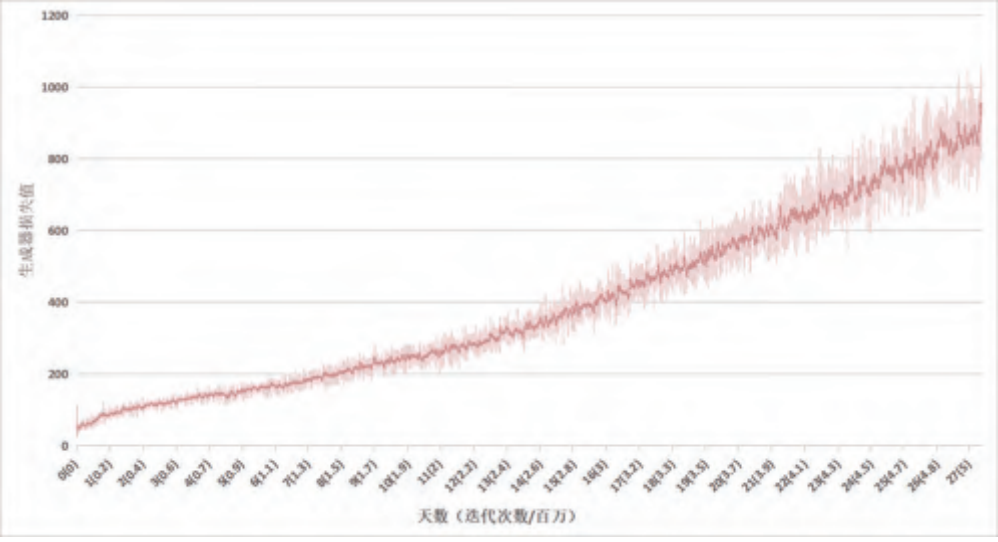


图 5：MelGan 的 Gloss 训练过程

文本特征方案	输入特征	输入范围数量级
字方案	[‘你’,’好’]	10000
完整拼音	[‘ni3’,’hao3’]	1000
拼音声母+韵母	[‘n’,’i3’,’h’,’ao3’]	100
拼音单个字母+音调	[‘n’,’i’,’3’,’h’,’a’,’o’,’3’]	30

征（字、音素、重音）转换为各种声学特征，英文数据集训练时一般通过因素作为文本特征进行编码。中文语音可以转化为拼音，缺少英文更加细致的音素表，我们研究了字、完整拼音、拼音声母＋韵母、拼音单个字母＋音调四种方案，比如输入文本“你好”，各类方案如上表。

通过上述几种方案的验证，字方案输入范围数量过大训练集无法覆盖全，预测时遇到新的词无法转换，方案不可行。另外三种方案测试下来，拼音声母＋韵母以及单个拼音的测试效果较差，预测合成声音极不稳定，部分字完全偏离正常发音；只有拼音的方案能够输出稳定的发音，能明确识别说的内容，最终我们选取此方案作为文本输入特征。

4.3.2 Decoder 结构优化

原始模型结构输出包括线性频谱和 Mel 频谱。首先我们只需要 Mel 频谱的预测，去除了线性频谱生成的部分。另外针对 Mel Output 的结构进行优化，生成更加细致的 Mel 频谱。

1、Mel input 输入时进行了降采样，默认 DownSampling=4。但是在外呼业务的变量文本场景下，训练文本偏短，降采样后实际输入长度过小；将模型降采样过程去除，直接用原 Mel input，最终 loss 收敛值更小，与真实 Mel 频谱更接近。

2、Mel Output 原模型结构处理较为简单，

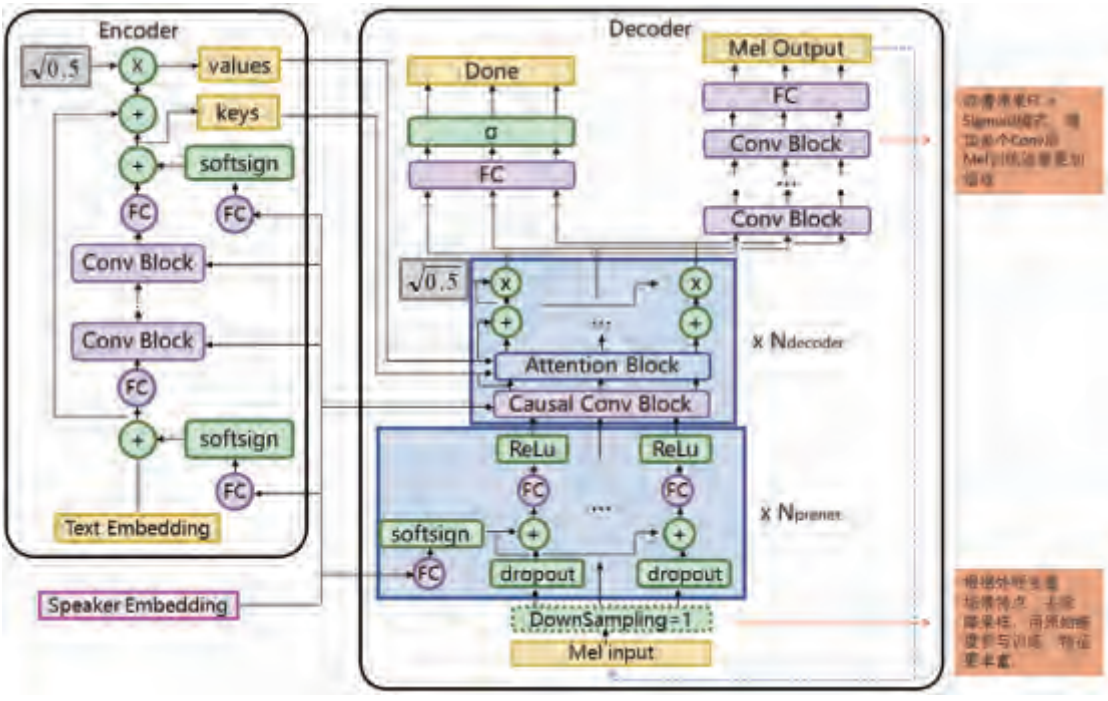


图 6：Decoder 结构优化说明

经过 FC 层后直接做 Sigmoid 转换得到 Mel 频谱，经过多次测试，我们加入了多层 CONV 进行微调，Mel Loss 下降更加快速同时收敛值更低。

4.3.3 Mel 频谱增强

与 NLP 模型不同，文本数据训练目标是明确一致的，而语音数据无可避免的存在差异，同一个字每次发音都有区别，最终预测的 Mel 频谱与真实数据会稍有不同。在 Tacotron2 的训练过程中，作者直接用预测的 Mel 频谱作为输入用于训练声码器，取得更好效果。我们对此方式也进行了尝试，合成效果稍有增强。但是训练过程不稳定，同时需要等待前模型的训练完成，受图像增强相关的技术启发，我们特别提出了 Mel 频谱增强的方式，来进一步提升声音合成的细腻度和真实感。

2017 年 CVPR 超分辨率论文 SRGAN 中备受瞩目的，把图片高清的效果带到了一个新的高度。SRGAN 是基于 GAN 方法进行训练的，有一个生成器和一个判别器，判别器的主体使用 VGG19，生成器是一连串的 Residual block 连接，同时在模型后部加入了 subpixel 模块，提升分辨率的同时减少计算资源消耗。在 SRGAN 模型的基础上，我们进行了部分改造，专门用于 Mel 频谱的增强。

（一）损失函数修改

原论文中给出的损失函数：

$$l^{SR} = \underbrace{l_X^{SR}}_{\text{content loss}} + \underbrace{10^{-3} l_{Gen}^{SR}}_{\text{adversarial loss}} + \underbrace{2 \cdot 10^{-8} l_{TV}^{SR}}_{\text{regularization loss}}$$

perceptual loss (for VGG based content losses)

此损失函数包括两部分，第一部分是感知损失，第二部分是正则化损失。其中感知损失是由李飞飞团队提出的一种损失函数。将生成器生成的假高分辨率图像和真实的高分辨率图像送入 VGG19 网络中进行特征提取，在提取的特征图上再使用均方根误差，能够提升图像的超分辨率效果。但针对 Mel 频谱场景，并不适用图像的感知范畴，content loss 仍沿用 MSE Loss，其余 loss 保持一致。

$$l_X^{SR} = \sum_x \sum_y \left(I^{HR}_{x,y} - \left(G_{\theta_G}(I^{LR}) \right)_{x,y} \right)^2$$

（二）模型结构调整

原模型中输入数据为彩色图像，包括 RGB

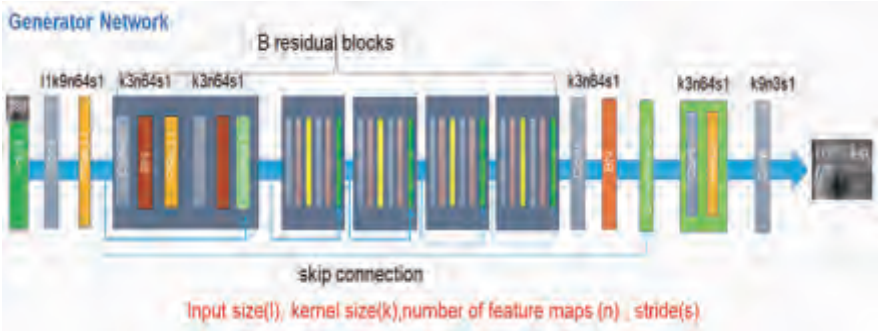


图 7：Generator Network 架构图

预测的 Mel 频谱　　真实的 Mel 频谱　　增强后的 Mel 频谱

图 8：Mel 频谱图增强效果对比

三通道，数据维度为（3,*,*），同时模型训练输入是缩小且模糊的图片，目标是放大高清图片，训练过程包括 UpSampling 步骤；针对 Mel 频谱的数据特征，对模型 Generator Netwok 进行调整如图 7。

通过数万对预测的 Mel 频谱和真实 Mel 频谱进行训练，能够得到较好的 Mel 频谱增强效果，同样增强后的频谱输入到 Vocoder 后产生的声音更加细腻明亮。

4.4 最终模型结构

最终我们完整的模型优化如下构建了如图 9 所示的语音合成模型，该模型由三部分构成。第一部分可将文本特征转换为梅尔频谱，第二部分对梅尔频谱进行增强处理，第三部分通过声码器将梅尔频谱转换为原始波形输出。最终单个音频生成的速度小于 200ms，经过增强的 Mel 频谱产生更加细致、真实的频谱，弥补初始生成的 Mel 频谱过于平滑的缺陷，对应的声音更加细腻清晰。

五、智能外呼效果提升

我们广泛体验了各类场景下的智能外呼服务，包括金融类、互联网、营销类等多种外呼场景，普遍采用的是全 TTS 模式，声音流畅甜美，但是仍然存在机器合成感，能够感觉到是机器人在对话，影响客户的服务体验。我们采用的录音 +TTS 混合模式，在录音播放部分客户是无感知的，和客服直接回答声音效果一致；变量部分通过优化后的 TTS 进行合成，TTS 的训练声音和播放录音都来自同一人，最终合成的效果无论在音色、音调以及语气各方面都与录音原声较为接近；在持续的问答交互中，用户对录音和 TTS 之间的切换基本无感知，通话的整个流程用户实际体验都是和同一个人在交互，在语音方面提升了客户的体验。

在外呼的实际结果对比中，之前通过录音 + 商用 TTS 合成的混合模式下，客户在交互中明确抱怨机器人声音大约占比 1‰，而且抱怨后配合完成回访意愿明显下降，大概率情况下会

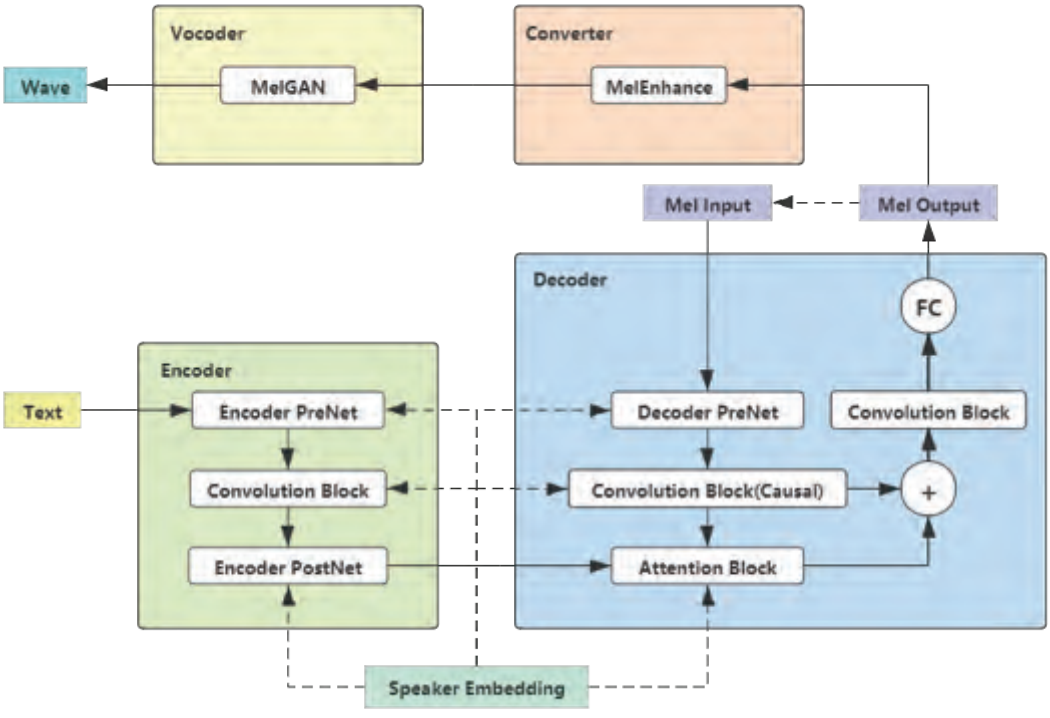


图 9：优化后模型结构

挂断电话或提出转人工回访。在切换到自主优化的 TTS 模型后，外呼一个月内未收集到抱怨机器人声音的记录，给客户带来更好的语音体验，提升整体的回访成功率，初步统计在切换后回访成功率提升 1.5‰，全年可节省数万通外呼电话，节省外呼成本。同时通过掌控个性化的 TTS 模型完整优化流程，后续我们可以定制更多种类的客服声音，根据客户的年龄、地域等属性进行客服声音的匹配，进一步带来个性化的客户体验，打造公司特色 AI 服务。

六、总结展望

本文通过开源的算法模型，结合券商的实际应用场景特点对模型做了大量的研究和优化工作，在简易的录音设备和较少的录音训练数据下，有效解决智能外呼应用中出现的人名合成的问题，对于提高用户体验、减少建设成本以及后续推广等有重要意义。后续在其他语音类智能服务场景下，可以通过自主训练该 TTS 模型，实现不同角色的人声合成，打造更有温度的客户服务体系。

同时，TTS 优化研发有助于推动公司迈向自主探索 AI 应用发展的道路。目前以深度学习为代表的 AI 技术，在 NLP、图像处理、语音数据处理等方面的基础理论和算法流程基本一致，通过 TTS 的自主优化开发积累了丰富的论文学习、模型调参、算法结构调整等经验，对于推动我司 AI 个性化应用场景建设有积极意义。特别是在金融数据安全性等方面因素的制约下，加强公司 AI 自主研发能力，能够有效实现数据更新、模型迭代和产品维护等全流程的自主掌控，持续推动金融科技在证券行业的落地。

海通证券云管理平台微服务化改造实践与思考

陆颂华 王朝阳 张真真 胡晶玉 / 海通证券股份有限公司 wzy10505@htsec.com

海通证券混合金融云平台在集团内部广泛推广和使用，承载的应用超过 500 个，集团总部、子公司和分支机构也对日常使用的云管理平台提出了更高的要求。本文主要介绍为了能够满足更多复杂多变的使用场景，海通证券对云管理平台进行微服务化、容器化改造的最佳实践，以及改造过程中的一些总结与思考。

1 项目背景

以赋能业务创新为基础，提升资源效能为动力，践行合规风控为准则，通过使用国产化的软硬件技术，实践组织与制度的创新架构，并结合证券行业大量真实应用场景，海通证券建设完成了由云管理平台统一纳管、编排私有云（桌面云、研发测试云、生产云、灾备云和托管云）和公有云的混合金融云平台。此平台在集团内部广泛推广和使用，承载的应用超过 500 个，使用的部门有集团总部、子公司和分支机构。

1.1 微服务化改造的必要性

云管理平台对外是面向最终客户和业务运营

侧的管理软件，相对于对内面向基础设施层的多云平台，其面临场景更复杂多变，尤其是随着托管云战略的深入推进，越来越多的子公司和分支机构租户将通过该平台使用和运营混合金融云平台资源，而每一个租户对于云资源使用和运营都可能会有不同的要求，这就要求云管理平台具有更好的“柔性”，能够快速适应复杂多变的租户需求。为此，对云管理平台的功能模块进行更细粒度的微服务化拆分并通过微服务架构进行整合成为必要。

随着容器技术的逐步成熟和在金融行业的落地使用，海通证券也将建设容器云平台。作为新一代的基础设施管理平台，容器云平台提供了一致的环境部署能力、强大的调度和自愈能力、简单便捷的应用快速发布能力。作为混合金融云面向最终用户和业务运营的唯一入口，云管理平台也希望能利用到这些能力，降低平台日常运维工作量，提升开发运维效率，为未来混合金融云大规模对外服务提前做好准备。

1.2 微服务化改造的目标

基于以上原因，海通证券从 2019 年开始对云管理平台微服务化改造并明确了以下三点关键目标。

1、实现云管理平台从面向功能场景设计向面向服务设计转变。通过面向服务的设计方式，将云管理平台内的核心服务进行抽象、复用并独立演进，进一步在面向客户场景支持时实现核心服务的灵活组合。通过这一转变来应对云管理平台在面向越来越复杂场景需求（尤其是托管云场景）时的现实挑战。

2、实现整个云管理平台的容器化部署并在容器云平台内实现自动化调度。云管理平台内所有微服务组件需要遵循主流微服务框架开发，并能够完整在容器云平台内部署运行，充分利用容器云内的调度能力实现平台的分布式高可用部署以及常见故障的自愈处理，从而提升云管理平台的服务水平，降低平台维护成本。

3、构建适应容器环境、基于核心框架的二次研发迭代流程。在完成微服务化改造以及容器化部署后，平台在公司内的日常演进和迭代开发工作也需要构建在相应的容器化开发工具链，这样才能更好发挥云管理平台微服务化的价值，并持续深化。

2 建设思路

在确定好建设目标后，我们将工作的重点放在整个平台提升建设的落地思路上，重点从产品框架模块化、业务逻辑微服务化、部署方式容器化以及交付流程自动化这四个方面着手考虑，具体如下。

2.1 产品框架模块化

对典型业务系统进行微服务化改造是一个众所周知的难题，寄希望于一步到位的改造路径很多时候都会事倍功半。为此，我们本次改造选择了先模块化再微服务化的渐进式模式。传统典型应用的产品框架一般都是分层架构，业务功能及业务逻辑的复用多是通过公共代码库来实现。显然，这些实现方式和框架模块化并不匹配。为此，我们为产品模块化改造引入两条基本原则，即“面向场景划分模块”和“每个模块独立运行”。

所谓“面向场景划分”就是从平台使用场景的角度把相应的功能放到一个模块。比如，用户需要利用云管理平台管理主机资源，则所有和主机相关的功能划分到一个模块。这个场景包括多个不同角色，所以这个模块也需要支持该场景下所有角色的功能。为此，我们将云管理平台按场景分成了十多个模块。并且，每个模块的实现是一个完整的技术堆栈（包括独立数据库或者表、独立的业务逻辑，独立的 UI 界面）。

所谓“每个模块独立运行”是指每个模块有自己独立的运行时，每个模块可以独立启停。系

统层面会有独立的模块注册、加载和管理的能力。需要注意的是，这个阶段，模块之间的数据交互还是通过底层数据库直接完成，并没有通过模块化之间的 API 调用进行。多个模块还是共同访问一个完整的“大”数据库且数据库表结构基本不做太大变化。

基于以上两条原则，整个产品框架模块化改造按以下三步有序推进：

第一步：设计模块框架。这部分工作包括描述模块的基本数据结构，相应元数据的数据库存储结构，以及模块集中注册、加载和管理等生命周期操作。平台需要为每个运行的模块启动一个独立的运行时，并按照生命周期操作来启动或停止相应模块的运行时。另外，任何一个模块运行过程中都可以读取当前模块元数据和当前运行情况，并根据以上信息决定是否在 UI 上显示相关模块功能。

第二步：设计模块细粒度访问权限体系。这部分工作包括对于模块内部功能点的描述方式定义，并能够实现不同角色对于该模块内功能点的细粒度授权管理。目前，平台使用 JSON 格式的描述方式定义每个模块的功能点，并在模块内部界面加载过程中按照实际授权情况动态加载相应的界面 UI 元素。

第三步：将现有平台的功能按“面向场景划分”的原则迁移到相应的模块，并按模块功能需要定义出相应需要进行细粒度管理的功能点。

最后补充说明一下，我们之所以先做模块化改造然后再做业务逻辑微服务化而不是直接就进行公共服务抽取并直接微服务化，是基于以下两点具体考虑：

1、未完成面向场景的模块化设计前，我们很难准确认识到应该如何抽取公共服务。通过面向场景的模块化设计，我们很容易发现重复的业务逻辑和数据结构，不同模块之间的依赖关系，而且可以梳理清楚需要抽象的公共服务具体接口。

2、如果直接简单进行公共服务抽象，我们担心最后的结果是公共服务式抽出来了，剩下的各个不同场景的功能仍然是“一团浆糊”存在于一个遗留工程中。而我们云管理平台的功能迭代和暴露还是要在这个遗留工程中，接下来的开发和迭代仍然很难。与其这样，还不如最开始就从模块化入手来改造项目。

2.2 业务逻辑微服务化

在解决了产品模块化之后，我们需要重点解决云管理平台业务逻辑的微服务化。有了平台模块化建设的基础，这部分工作就会变得容易很多。一方面，需要抽象的公共服务和其需要暴露的接口已经显现出来，另外一方面，通过面向场景的模块化改造，功能模块和公共服务层次关系也变得很清晰。具体来说，这个过程包括以下三个方面工作：

1、引入微服务开发、运行和管理框架，为所有微服务模块的开发、运行、分布式事务和治理提供基础性支撑。在本项目由于之前开发基于 Spring Boot 框架，所以自然延续下来使用 Spring Cloud 微服务治理框架。

2、建设公共微服务。具体来说，我们建设了包括“权限管理中心”、“流程管理中心”、“订单管理中心”、“工单管理中心”、“模块管理中心”、“用户管理中心”、“租户管理中心”和“IP 池管理中心”等公共服务。

3、改造面向用户的模块，使其依赖于公共微服务并将自己也变成一个微服务组件。这次过程中我们改造了“平台管理”、“资源管理”、“主机交付”、“运营分析”、“监控告警”、“容器管理”和“网络管理”等十几个功能模块。

改造过程有一个非常大的工作量就是数据库的拆分和跨服务之间调用的 API 化。我们需要让每一个微服务依赖自己独立的数据库，并通过 API 进行数据交换。所以，通过这个过程后，我们平台后台数据库将从一个“大而全”的实例变

成很多“小而专”的数据库实例。

2.3 部署方式容器化

在微服务化改造前，整个云管理平台的部署方式如图 1。

从以下部署架构图可以看出，平台部署采用的是典型主 - 备部署方案，分成业务接入层、业务逻辑层和数据库存储层三个部分，并部署在两台独立的物理设备上。其中，业务接入层采用的是典型“VIP + keepalive”模式，实现系统故障时候的自动切换。业务逻辑层则在两台独立物理机上分别部署运行，并同时连接到主数据库。而数据库层采用最典型的 MySQL 主从部署方案实现数据库的高可用。

在完成云管理平台微服务化改造后，整个平台是要在容器云环境内运行的，并将系统的高可用交给容器云平台自行编排，而不是现有的传统主从模式配置和切换。具体来说包括以下三个方面。

1、业务接入层不会再是主从切换模式，而应该换成负载均衡模式。用户访问平台的流量会被容器云平台内的负载均衡设备分发，所有运行的云管理业务实例都会同时对外提供服务，这需要进一步细化云管理业务模块中的状态管理。同时，需要增强每个云管理服务的状态监测能力，让负载均衡设备能更精确决策流量分发。

2、所有云管理平台业务都将在容器里面运行，需要改进现在的业务运行监控模式，能够适配在容器环境中采集监控数据并对外展示。这和基于操作系统部署业务的传统模式有非常大的不同，有较多的适配工作。

3、在数据库层面，采用了保守的迁移策略，先将数据库保留在传统操作系统上面并依然按照主从模式部署运行。

经过改造后，整个平台的部署运行图 2。

2.4 交付流程自动化

如前所述，整个云管理平台微服务改造的一个重要目标就是改进软件交付流程。在这一次软件交付流程的改进过程中，我们从传统的基于物理主机的构建、打包和部署的流程切换为基于容器的 CI/CD 流程，并依赖容器云的部署调度能力实现平台的高可用部署。具体的流程如下。

1、首先定义一个流水线（Pipleline）。

图 1：云管理平台改造前部署架构图

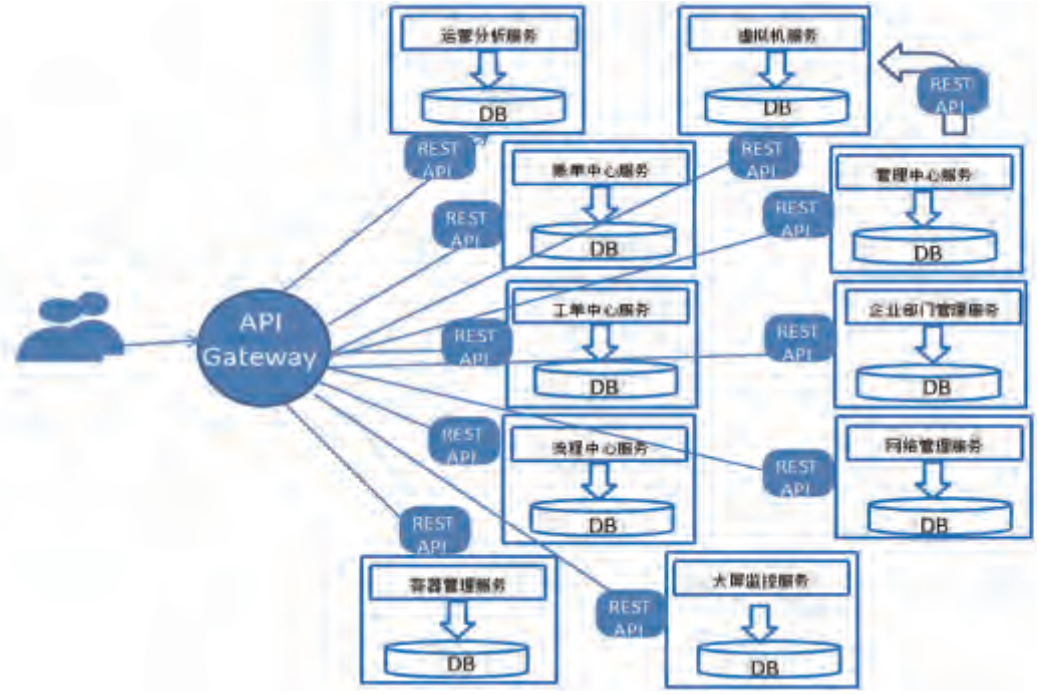


图 2：云管理平台容器化部署架构图

2、源代码在 GitLab 中管理，当新特性的代码合并到主分支后，自动触发 Jenkins 基于 dockerfile 进行构建打包。并把产生的容器镜像存放在容器云平台的制品仓库中。

3、利用容器云平台的发布及编排功能，将生成的容器镜像部署在指定的容器云应用集群并完成相关配置。类似，我们仍然是编制相应的部署描述文件并上传到容器云镜像以实现整个发布过程的自动化。

3 实践总结

回顾整个项目微服务化、容器化改造的过程，有几点实践值得总结和交流。这些实践经验贯彻整个项目改造过程并一直指导项目在正确的道路上前进。

3.1 按产品实际功能需求和定位定义推进方式

微服务化改造对于典型应用是有着极其巨大的挑战，最容易犯的一个错误就是参照经典理念，选择一个微服务框架，然后就进行业务模块的拆分和相应的开发。这样激进的操作非常容易造成整个改造过程的失控。一定要从产品实际定位和真实痛点出发。在本次微服务化改造过程，一个重要的出发点是使平台在面对未来越来越复杂的需求时，具有面向不同场景的“柔性”。所以，在微服务改造过程中，第一个想法不是去抽取公共服务，而是做面向场景的模块化改造。这样一方面方便将不同场景进行切割，另一方面也为接下来做微服务改造打下基础，降低改造风险和项目失控风险。

3.2 微服务模块的粒度和边界控制非常重要

当进行业务逻辑微服务化改造时，如何划分微服务模块（包括公共服务）并控制好其调用关系变的尤为关键。粒度划分过细导致管理难度增加并且会引入不必要的分布式复杂度，划分过粗会导致相互调用和依赖关系混乱，无法合理管理和跟踪。在整个改造过程，我们采用渐进式拆分思路，遵循仅垂直依赖而不是平行乃至反向依赖

的模式。具体来说，只会在有多个模块有共同依赖的情况下才会主动把共同依赖部分拆分成为公共服务，而公共服务部分也采取先大颗粒度建设，直到上层模块依赖复杂到必须进行细化时候再做拆分。

除了模块之间的调用外，另外一个非常困难的地方在于数据库的切分。传统应用集中访问数据库，并利用数据库软件的事务功能解决数据一致性问题，而微服务化后，首先每个微服务模块都将拥有自己的独立数据库，跨微服务的操作会转换成为对多个微服务接口调用，这样就无法利用数据库管理软件内的事务能力，而需要在 API 调用层引入分布式事务管理框架，从而保障数据的一致性。

在拆分过程中，首先定位相应微服务模块的核心能力以及对应的核心数据模型是什么，先将其纳入到微服务模块自己的数据库内，并基于此做相应功能开发和接口暴露，遇到对外部数据库表的依赖时候，首先仍然以对外部数据库的直接方式访问实现。通过这个过程让我们重新梳理模块之间的数据依赖情况，然后再将这些数据库直接依赖转换成为模块之间的 API 调用。

3.3 保持业务在线运行并不断在真实环境中验证

在整个改造过程，我们遵循一个原则，就是业务需要持续在线运行不中断。我们每做一步改造和服务抽象，都会发布一个完整的服务版本，并经过测试环境、仿真环境和生产环境的路径反复验证并试运行，通过这个策略保证整个平台改造质量的把控，同时又能实现无缝升级。当然，这会给整个过程的业务保障带来很大挑战，需要我们对于每一步升级都有完整的升级方案和回滚方案。这个需要在工程管理上做好大量细致性的工作。

4 规划与展望

在完成以上微服务化改造后，整个项目建设已经达到最初设想的整个目标并在真实的容器云环境运行起来，对外提供服务。随着投产运行，我们也对微服务化体系架构及其日常运维管理有了更深一步的理解和认识。基于这些认识，下一步会重点筹备分库后的 DBaaS 建设和持续提升整个微服务架构应用的服务治理能力。

基于OpenCL开发的深交所Binary协议行情解码

邹经纬 马辉 / 国泰君安证券股份有限公司
钟浪辉 陈敏 / 上交所技术有限责任公司

行业现状

证券交易系统中，行情一直是最重要的环节之一。一般情况下，交易所向会员单位提供行情网关程序，会员单位使用行情解码程序，与交易所行情网关程序连接，然后解码收到的行情数据，进行行情处理，最后分发给行情使用者。行情系统包含如下几种类型：传统的集中式行情系统，低延时的组播行情系统和基于 FPGA 的硬件行情系统等等。

传统的集中式行情分发系统，如下图所示。采用软件实现的行情解码程序，连接交易所行情网关 MDGW(Market Data GateWay)，建立 TCP 连接。行情解码程序接收 MDGW 发送的原始行情数据，进行解码，然后将解码后的数据发送到集中的行情中心。行情中心负责行情数据的落地以及分发，将数据落地到历史行情库中，然后将实时数据分发给订阅行情的客户程序。整个链路都是采用软件 TCP 传输。在这种集中式行情系统中，从交易所 MDGW 发送原始数据到客户收到解码行情，时延消耗为几百微秒至几毫秒。

低延时的组播(UDP)行情系统，如下图所示。采用软件实现行情解码，连接交易所 MDGW，

建立 TCP 连接。行情解码程序将解码后的行情采用组播（UDP）的方式，发送到行情组播私网中，在行情组播私网中的客户将接收到组播行情。与集中式行情系统不同，这种低延时组播行情系统不适用于远距离传输，一般在交易所托管机房内部署。这种行情组播系统多采用低延时网卡、搭配低延时交换机和高性能服务器。行情传输时延能达到几微秒到十几微秒，接近于纯网络通信的时延。

基于 FPGA 的硬件行情系统，如下图所示。一般采用具备网络通信功能的 FPGA 板卡和 x86 服务器构建的异构平台，进行行情解码和行情分发。FPGA 一般采用硬件描述语言开发。通过 CPU 负责行情解码控制面，FPGA 负责数据面。FPGA 与 MDGW 建立 TCP 连接之后，接收原始行情数据，在 FPGA 卡内进行行情解码和组播分发，不经过 CPU 处理。最后将行情数据发送到行情组播私网中，加入行情组播私网中的客户将接收到硬件组播行情。由于网络数据不经过 CPU 侧 DDR 处理，因此行情传输时延可以到几百纳秒到 1 微秒。

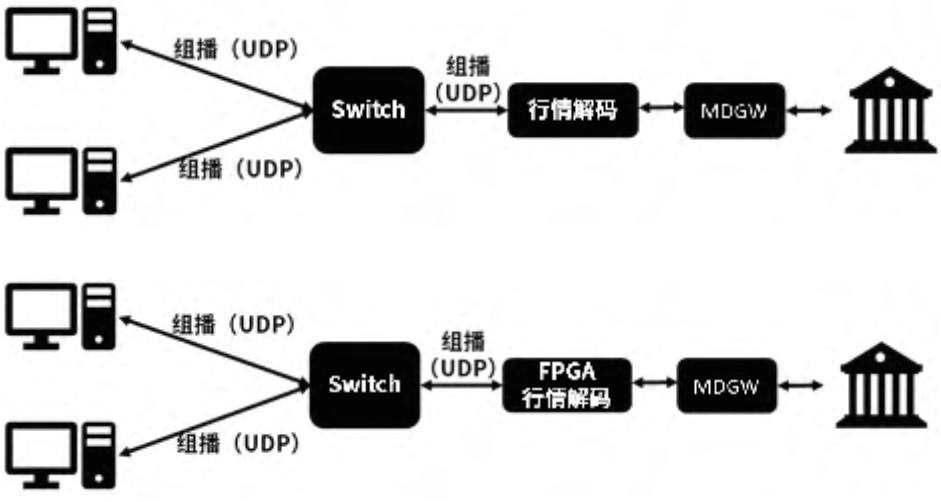
软件开发人员单纯依靠 CPU、低延时网卡和低延时开发套件等，越来越接近时延瓶颈，再考虑到操作系统的调度和时延抖动，很难进入（几）微秒甚至纳秒级的竞争。FPGA（可编程逻辑阵列），作为可编程硬件，可以处理网络数据，内部的大量逻辑资源可以重新“编程”，实现业务逻辑，而且时延抖动小。

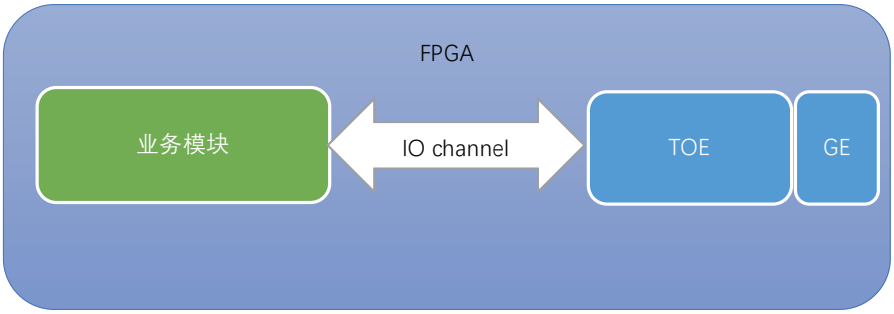
异构加速平台

FPGA 开发大多采用硬件描述语言，开发周期长、难度大，软件开发者很难进入到这个领域。但是随着 FPGA 加速应用的普及，Intel（英特尔）和 Xilinx（赛灵思）也相继推出了 OpenCL（Open Computing Language）和 HLS（high-level synthesis）开发套件，支持采用类 C 语言进行 FPGA 开发。

本方案采用的是 Intel 的 PAC（Programmable Acceleration Card）卡和 x86 CPU 构建的异构加速平台。PAC 卡内置的是 Arria 10GX FPGA，逻辑资源非常多。该异构平台支持使用 OpenCL 进行开发和运行。

CPU 主机侧采用 C++ 开发行情解码控制面程序，FPGA 设备侧采用 OpenCL C（基于 C99）进行数据面开发。将 TOE（TCP over Ethernet）IP（Intellectual Property）核嵌入到 FPGA 中，FPGA 就具备了网络通信的功能，可以支持到传输层协议解析，然后将数据交给应用层处理。这里，应用层就是采用 OpenCL C 开发的业务



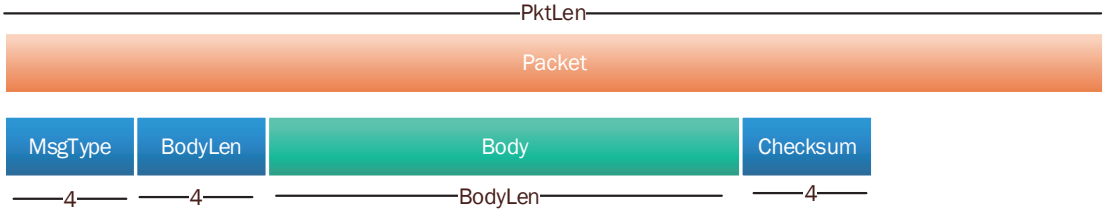


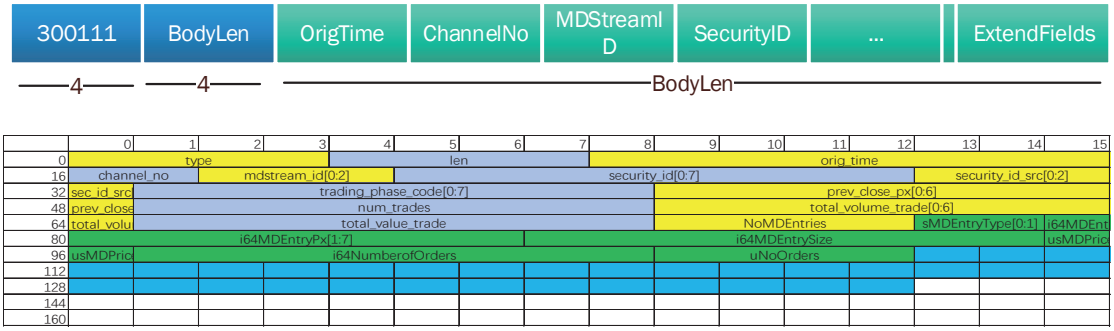
程序，即行情解码。TOE 与业务模块通信，采用 IO-channel 的方式，数据位宽 128bit，也就是 16Bytes。

协议分析

深交所的行情，从 MDGW 发送给下游的解码程序，采用 TCP 连接，因此，在应用层采用分包协议。发送给下游的字节流中，分包协议如下所示。每个应用层行情包，是一个完整的业务报文，可以解码出一类行情信息。一个应用层行情包分成三个部分：消息头，消息体和消息尾。消息头包含 4 字节消息类型 MsgType（如 300111 快照，300192 逐笔委托等），4 字节消息体长度 BodyLength。消息体就是消息类型对应的行情数据。消息尾定义了消息的校验和，计算范围从消息头开始一直到消息体结束。每解完一个应用层行情包，后面紧跟的就是另一个行情包，根据消息类型解析后续的消息体。**行情中所有整数字段均为大端字节序。**

以 300111 行情快照为例，如果解析到 MsgType 为 300111，则后续的消息体使用 300111 对应的 Binary 协议进行解析。依次从消息体中解析 OrigTime（数据生成时间），ChannelNo（频道





	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	type				len				orig_time							
16	channel_no			mdstream_id[0:2]				security_id[0:7]						security_id_src[0:2]		
32	sec_id_src		trading_phase_code[0:7]						prev_close_px[0:6]							
48	prev_close		num_trades						total_volume_trade[0:6]							
64	total_volu		total_value_trade						NoMDEntries				sMDEntryType[0:1]		i64MDEnt	
80	i64MDEntryPx[1:7]							i64MDEntrySize							usMDPric	
96	usMDPric		i64NumberofOrders						uNoOrders							
112																
128																
144																
160																

代码），MDStreamID（行情类别），SecurityID（证券代码），对应的扩展字段等等。

目前 PAC 卡中的 TOE 模块，输出数据的位宽为 128bit（16 字节），为了充分利用 FPGA 的流水线处理特性，业务模块从 IO-channel 中每取到 16 字节进行一次处理，边接收，边解码，而不是收到一个完成行情包再解析（这样会引入较多的时延）。当然，也可以将 16 字节拆分成 8 字节或是 1 字节进行流水线处理。下图为深交所行情快照按照 16 字节摆放，每一个时钟周期数据对应的内容。

OpenCL 行情解码

异构平台程序分主机程序和设备程序。主机程序指运行于 CPU 上的软件程序，设备程序指运行于 FPGA 内部的业务功能模块（OpenCL 编写的业务模块，非 TOE）。主机和设备通过 OpenCL 运行框架通信，进行任务分配。网络功能则通过 TOE 对应的驱动程序，由主机侧调用，通过 TOE 与交易所的 MDGW 建立连接。

设计上分控制面和数据面。CPU 相对灵活，编程容易，适用于逻辑控制。FPGA 虽然可编程，但是功能固定，适合高速数据处理。因此 CPU 负责控制面，FPGA 负责数据面。

控制面

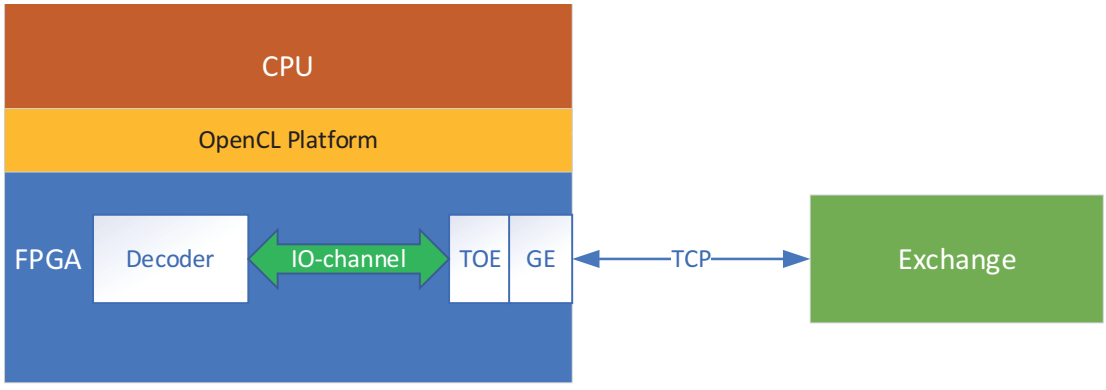
1. 管理上下游链路。与上游交易所 MDGW 建立 TCP 连接，向下游组播私网发送组播行情。

2. 管理 FPGA 上的任务执行。基于 OpenCL Platform，下发 Kernel 任务给 FPGA，包括 FPGA 初始化和释放、MDGW 登录、行情解码等等。

3. 监控行情解码。监控上下游连接状态，监控行情解码的收发包个数。

数据面

1. 对接 TOE。利用 IO-channel 与 TOE Input/



Output 通信，解析和封装 TCP/UDP 报文。

2. 行情解码。进行 TCP 报文的处理、分包，Binary 协议（行情快照、指数快照、逐笔委托、逐笔成交）解析。

软件行情解码，一般是收到一个完整的业务报文，再进行协议解析。由于 CPU 本身有高速缓存 cache，而且主频比 FPGA 高一个数量级，解包效率非常高。但是，数据是从网卡到操作系统协议栈（若使用 kernel bypass 网卡，则进入用户态协议栈），再搬移到 DDR 中，进行行情解析，存在多次数据搬移和拷贝。

FPGA 行情解码，整个行情解码过程中，数据从交易所 MDGW 到 FPGA 板卡 GE 口，进入 TOE，TOE 将数据传入 Decoder 业务模块，Decoder 将解码后的数据又通过 TOE、GE 口发送到对应的行情组播私网中。数据流只在 FPGA 板卡内部，不会进入 FPGA 板卡的片外 DDR，更不需要通过 PCIe 搬移到 CPU 侧 DDR，因此省去很多数据搬移的开销。

Kernel 解码逻辑设计时需要考虑到流水线的合理编排，充分利用 FPGA 硬件资源。TOE 和业务 Kernel 之间的 IO-channel 通道位宽 128bit（16Bytes），即 TOE 每个周期向 IO-channel 输入和读取 16Bytes 数据。为了获得最低延时，业务模块需要同时进行数据接收、行情解码和组播发送。快照解码的伪码如下。

在一个 while 循环体中，不停的从 toe_in 这个 IO-channel 中读取数据，然后进行数据解析。如果这个时钟周期的数据为应用层行情头，即 index 等于 0，则可以从中解析出 msgtype、msglen 和 orig_time 字段；收下一时钟周期的数据，可以解析出证券代码 SecurityId；收到最后一个时钟周期的数据时，解析完。封装解码后的行情数据，发送到中间的缓存 channel（data_out）中。并行运行的发送 Kernel 就一直从 data_out 中取数据，取到数据就通过 toe_out 发送出去。

```
void decoder() {
    int index = 0;
    while(1) {
        uchar16 frame = read_channel(toe_in);
        if (0 == index) {
            // get type, orig_time
        }
        else if (1 == index) {
            // get security_id
        }
        else if (MD_END == index) {
            // 完成一个解包
            finish = true;
        }
        if (finish)  {
            finish = false;
            // 将需要发送的行情数据，组成报文 md 发送出去
            write_channel(data_out, md);
        }
        index = (MD_END == index) ? 0 : (index + 1);
    }
}

void toe_out () {
    while(1) {
        char md[MD_SIZE];
        md = read_channel(data_out);
        // 每个时钟周期 16 字节
        for (int i = 0; i < MD_SIZE; i+=16)) {
            write_channel(toe_out, md[i]);
        }
    }
}
```

TOE 协议的报文分析。TOE 输出的每个应用层报文会包上一个 TOE 头，用于分包，并提供会话、长度信息。紧接着的数据就是应用报文。TOE 头中长度字段在 0 和 1 字节，长度的高字节在 TOE 头的第 0 字节，长度的低字节在 TOE 头的第 1 字节。会话 ID（cid）在第二字节，TOE 一般支持的会话在 128 以下。其余 13 字节都是 0x55。TOE 输入的数据是大端摆放，如果用 uchar16 数据缓存 TOE IO-channel 输入的数据，则 uchar16 的 0~15 字节对应的是 TOE 中的 15~0 字节，相当于数据翻转了。重新翻转之后的 TOE 数据如下：

```
0000:55555555555555555555555555110045
0010:640500000004c2944f000005650047c2
0020:bb56c2a2595a7803c3b4303130303032
0030:333231202031303220543020202020202
0040:2000000000000109c2a0000000000000
0050:01020304050000008dea9c5fb8000000
```

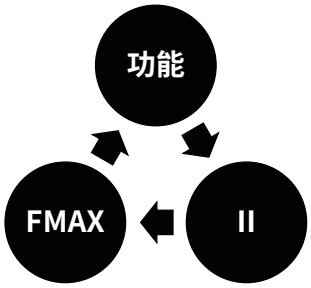
性能调优

在使用 OpenCL 开发时，第一步是功能实现，

第二部是性能调优。因为是使用 OpenCL 开发，gdb 调试，所以功能开发所需要的时间相较传统 RTL 开发少很多。性能调优占据整个开发流程中的一半时间。

II 值和 FMAX

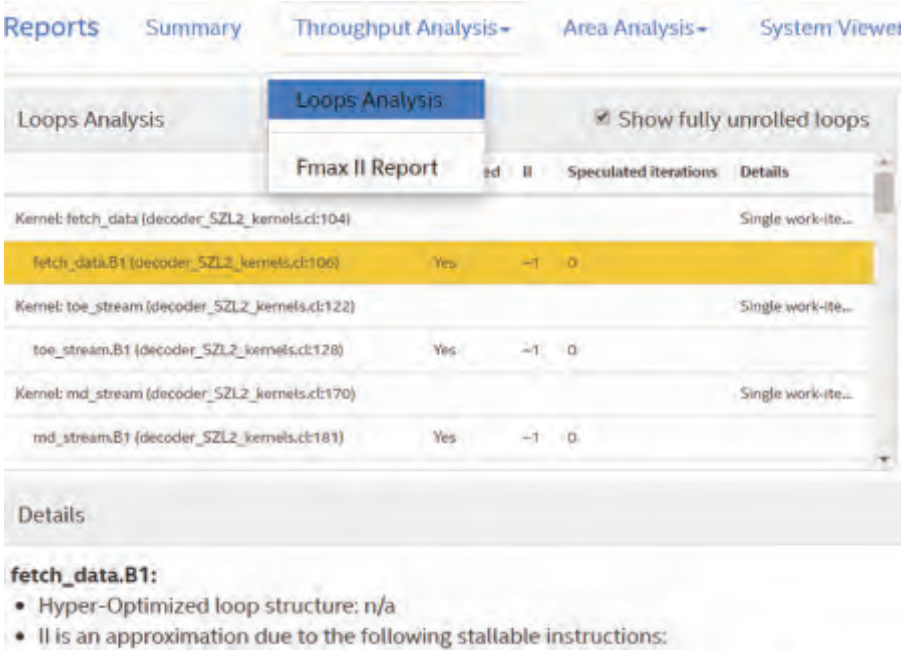
性能调优时，除功能外，需要关注的指标主要有两个：II 值（Initiation Interval）和 FMAX（maximum operating frequency）。



II 值表示连续循环迭代之间的时钟周期。FPGA 是采用流水线执行任务，一条流水线的吞吐取决于最慢的一个环节，在优化时可以认为 II 值就是这个流水线中最慢环节所需的时钟周期。如果流水线设计中，某个环节业务过于复杂，无法在一个时钟周期内处理完，则流水线的 II 值会比较大。II 值越大，表明吞吐越低；II 值越低，则吞吐越大，时延也越低。一般在低延时应用中，II 值目标是等于 1，也就是在每个时钟周期都能“吐出”一个任务结果。

FMAX 值简单理解就是 FPGA 运行的时钟频率。FMAX 越高，则任务处理越快；FMAX 越低，则任务处理越慢。业务模块过于复杂，FPGA 在设计和编译时，FMAX 就“跑”不上去，导致上板后 FPGA 时钟频率较低，业务处理慢。

如果 FMAX 足够低，即时钟周期较长，足够流水线中最慢的环节处理业务，II 值总能为 1；如果 FMAX 过高，即时钟周期短，而流水线中最慢的环节无法在一个时钟周期内处理完，则 II 值大于 1，流水线无法在一个时钟周期内输出一个有效数据。如果 II 值足够大，即流水线中可以容忍在多个时钟周期内执行一次循环迭代（经过多个时钟周期才能输出一个处理结果），FMAX 就可以很高；如果要求 II 值为 1，则要求流水线每一个时钟周期都输出一个结果，那意味着 FMAX 需要降下来，保证流水线的循环迭代能在一个时钟周期内处理完。



编译器总会在 II 值和 FMAX 之间折中。OpenCL 程序设计完后，可以在通过 Quartus 编译生成的报告中查看 II 值和 FMAX。如下图所示，reports 中会给出所有模块的 II 值和 FMAX，如果性能不高或者有优化空间，也会指出问题所在。

调优手段

详细的调优手段可以参考 Intel 官方的《aocl-best-practices-guide》。以下列举在行情解码中用到的一些调优改进方法。

循环展开

使用 while 进行不定次数的循环，使用 for 全展开（unroll）进行固定次数的循环。一般低延时应用中，只允许一层不定次数循环，若有嵌套的不定次数循环，II 值就不等于 1。

例如使用 for 循环拷贝不定长度的数据。拷贝或者处理 num 个数据，num<=16，下述写法 II 值不等于 1，因为存在不定次数循环。

```
for (int i = 0; i < num; i++) {
    frame[i] = data[i];
}
```

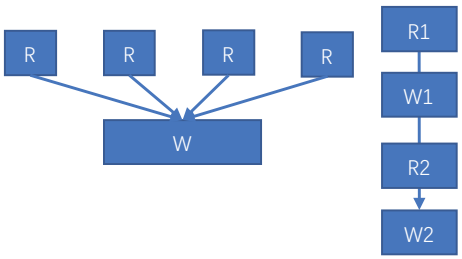
改写成如下写法。就能将 for 循环全展开，且并行执行 16 个数据的处理。由于逻辑固定，可以在一个时钟周期内完成。

```
#pragma unroll 16
for (int i = 0; i < 16; i++) {
    frame[i]=(i<num)?data[i]:frame[i];
}
```

同一变量的读写操作

首先要避免在一个循环体内复用同一个变量（逻辑资源充足，可以定义多个变量）影响编译器优化逻辑依赖关系。

一个循环体内，建议对同一个变量进行多读单写，且写在所有读之前或之后，避免多读多写相互穿插。因为多读单写时，编译器会复制多份变量，这样依赖关系清晰，且多个读代码会并行执行。而多读多写，可能会产生较为复杂的依赖关系，不利于编译器优化。



降低 IF 判断的复杂度

减少 if 嵌套，循环体内尽量只有 1~3 层 if 嵌套，过多的 if 嵌套不利于编译器理解代码逻辑进行优化。

```
if (cond1) {
    if (cond2) {
        if (cond3) {
            // ...
        }
    }
    else {
        // ...
    }
}
```

改写为下面的形式，if 中的判断条件会在一个时钟周期内执行完。这样的逻辑更为清晰，便于编译器优化。

```
if (cond1 && cond2 && cond3) {
    // ...
}
else if (cond1 && !cond2) {
    // ...
}
else if (cond1 && cond2 && cond3) {
    // ...
}
```

结论和展望

基于 OpenCL 开发的深交所 Binary 行情解码，是证券领域利用 OpenCL 进行 FPGA 低延时应用开发的一次探索，配合高性能低延时的 TOE，可以获得接近硬件描述语言开发的性能。使用 OpenCL 开发 FPGA 具备很多好处：开发和迭代速度快，可以使用 gdb 在普通 Linux 环境进行功能调试；OpenCL 编译器也提供详细性能分析报告，便于优化代码性能；还可以调用高性能的 HDL 或者其他语言开发的库。目前已经完成了深交所 Level1/Level2 的 Binary 协议行情解码，包括行情快照（快照和委托队列）、指数快照、逐笔委托和逐笔成交。未来可以在板卡内进行一些指标计算，或者利用 Intel 的 OpenCL FinLib，为客户提供更丰富的行情信息。

0 行业观察 bservation

北金所数据中心迁移

杨硕 / 北京金融资产交易所 系统运行部

北金所是在人民银行、财政部指导下，经北京市人民政府批准，于 2010 年 5 月 30 日正式揭牌的专业化金融资产交易机构。是人行批准的中国银行间市场交易商协会指定交易平台，在交易商协会党委的领导下，北金所为市场提供债券发行与交易、债权融资计划、企业股权交易、市场化债转股转股资产、债权和抵债资产交易，以及债券回购违约处置、到期违约债券转让等服务，为各类金融资产提供从备案、挂牌、信息披露、信息记载、交易到结算的高效服务。

北金所经过 10 年的高速发展，北金所累计交易量已突破 25 万亿元，已成为全球最大的信用类债券集中发行平台，以及国内交易规模最大、交投最活跃的固收类、权益类金融资产交易平台。同时，公司的 IT 系统已发展到拥有双技术平台，百余个信息系统，千余台虚拟化服务器的 IT 架构体量。原有机房无论是从面积、功能还是配电上均不能满足后续公司乃至协会系统的整体 IT 承载需求，为此北金所完成了新数据中心的建设，随之而来的任务就是安全稳妥的开展数据中心迁移的工作。

一、拟搬迁信息系统相关情况

本次数据中心搬迁包括北金所及中债资信（中债资信评估有限责任公司，协会系统下属公司），搬迁信息系统现状：

（一）北金所拟搬迁系统情况

为了支撑北金所“一体两翼”业务架构，落实“双平台”的发展战略，目前北金所信息系统硬件基础平台主要分为 NAFMII 硬件基础平台

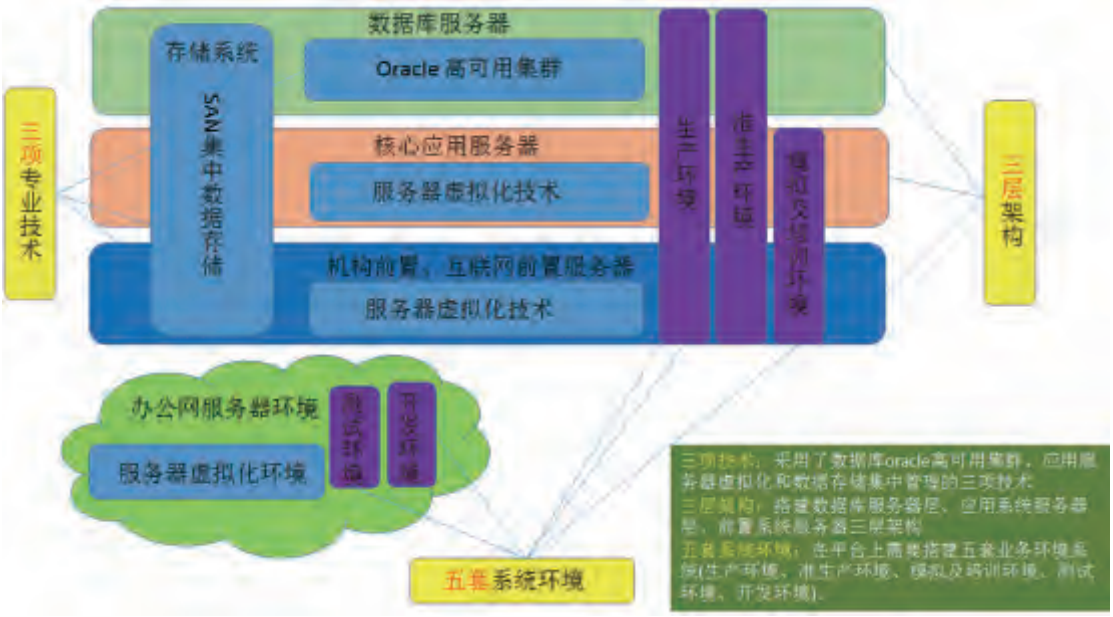


图 1：3+3+5 硬件架构图

（简称‘NAFMII 平台’）和 CFAE 硬件基础平台（简称‘CFAE 平台’）。NAFMII 平台及 CFAE 平台服务器、存储硬件架构均按照“3+3+5”架构进行建设：

三项技术：采用了数据库 oracle 高可用集群、应用服务器虚拟化和数据存储集中管理的三项技术

三层架构：搭建数据库服务器层、应用系统服务器层、前置系统服务器三层架构

五套系统环境：在平台上需要搭建五套业务环境系统 (生产环境、准生产环境、模拟及培训环境、测试环境、开发环境)。

“3+3+5”硬件架构图如图 1。

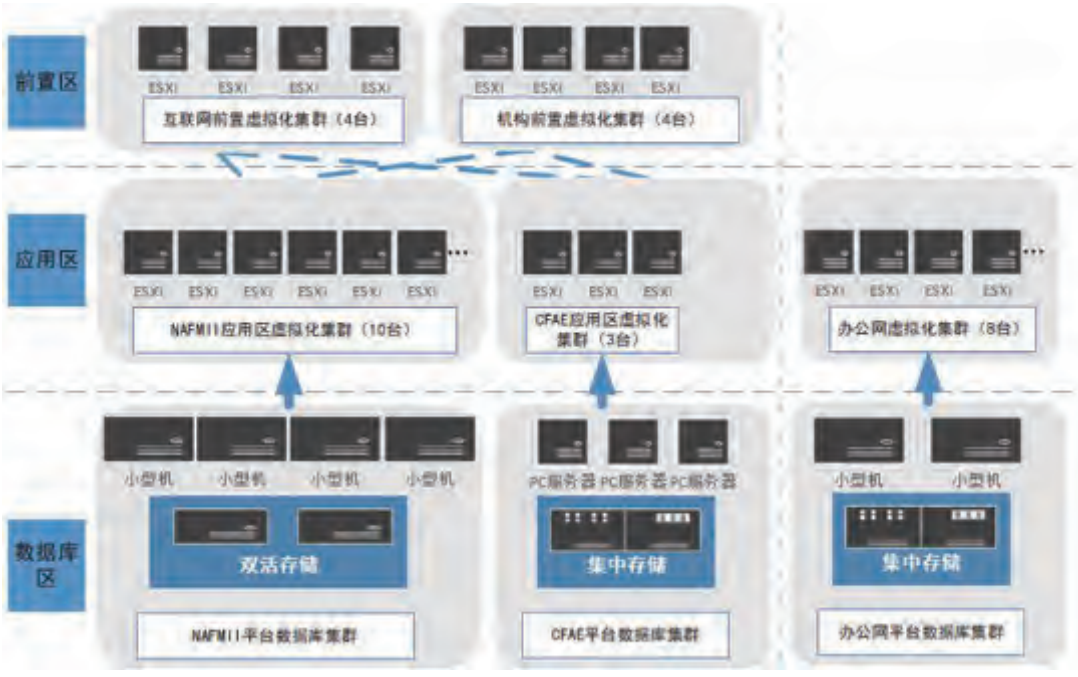


图 2：服务器整体硬件架构图

NAFMII 平台主要分为数据库区、核心应用区、互联网前置区以及机构前置区。CFAE 平台分为数据库区、核心应用区，互联网前置区及机构前置区与 NAFMII 平台共用。另外，CFAE 平台与 NAFMII 平台共享使用备份、监控等管理系统。服务器整体硬件架构图如图 2 所示。

各平台主要设备类型及配置如表 1。

其中 NAFMII 平台承载的信息系统主要包含了集中簿记建档、信息披露一期、信息披露二期、公募孔雀开屏、私募孔雀开屏、远程教育系统、统计分析系统、信用风险缓释工具、信息服务、一号通、统一客户端、信息服务功能迁移、PPN

表 1

区域	功能区	设备类型	配置	设备数量
NAFMII 平台	数据库区	小型机	2 台 IBM E850，2 台 IBM P740	4
		光纤交换机	6 台 48 口 SAN 交换机	6
		存储	EMC Vplex 双活存储，FC 协议，容量 60TB	2
	应用区	PC 服务器	32 核/256GB 内存	10
		光纤交换机	/	共用数据库区交换机
		存储	/	共用数据库区存储
	互联网前置区	PC 服务器	2 台 40 核/512GB 内存 2 台 16 核/128GB 内存	4
		网络交换机	4 台 48 口网络交换机	4
		存储	Netapp 集中存储，iscsi 协议，容量 30TB	1
	机构前置区	PC 服务器	2 台 16 核/256GB 内存 2 台 16 核/128GB 内存	4
		网络交换机	4 台 48 口网络交换机	4
		存储	Netapp 集中存储，iscsi 协议，容量 30TB	1
CFAE 平台	数据库区	PC 服务器	40 核/256GB 内存	3
		光纤交换机	2 台 48 口 SAN 交换机	2
		存储	EMC VNX5600 集中存储，FC 协议，容量 17TB	1
	应用区	PC 服务器	40 核/256GB 内存	3
		光纤交换机	/	共用数据库区交换机
		存储	/	共用数据库区存储
办公网平台	数据库区	小型机	2 台 IBM P720	2
		光纤交换机	2 台 48 口 SAN 交换机	2
		存储	IBM V7000 及 Netapp E5600 存储各 1 台，FC 协议，容量共 79TB	2
	应用区	PC 服务器	6 台 40 核/256GB 内存 2 台 32 核/256GB 内存	8
		光纤交换机	/	共用数据库区交换机
		存储	/	共用数据库区存储

专项投资人、统计数据标准化、做市商、数据运维、统一底层数据等系统的生产、准生产、模拟及测试环境。CFAE 平台承载了债权融资计划、委托债权、竞价系统、国金系统、金融国资信息采集、网上路演、信息服务工作平台、统一参与人、智能客服、债权融资计划区块链技术应用等系统的生产、准生产、模拟及测试环境。另外，部分系统独立部署在非集群 PC 服务器上，如发行前置服务器、交易前置服务器、新门户网站、开发环境统一配置平台等硬件环境。综上本次拟搬迁系统总共包含了 5 套虚拟化集群（32 台 PC 服务器、虚拟机共计 650 余台）、6 台小型机服务器、43 台非集群 PC 服务器、7 套集中存储系统等环境。具体设备类型及数量如表 2。

（二）中债资信拟搬迁系统情况

根据前期的梳理，中债资信共有 11 套信息系统需要进行搬迁，包含 PC Server、磁盘存储及各种网络设备，设备数量总计 55 台。根据计划，搬迁过程首先将在新机房搭建服务器、存储等硬件系统平台并完成数据库环境及虚拟化环境的部署，将拟搬迁信息系统迁移至新机房，经测试迁移成功后，再下线原有业务系统，并对原有的服务器、存储等硬件设备进行利旧。

拟迁移应用系统主机、数据库、存储、网络连接拓扑图如下：

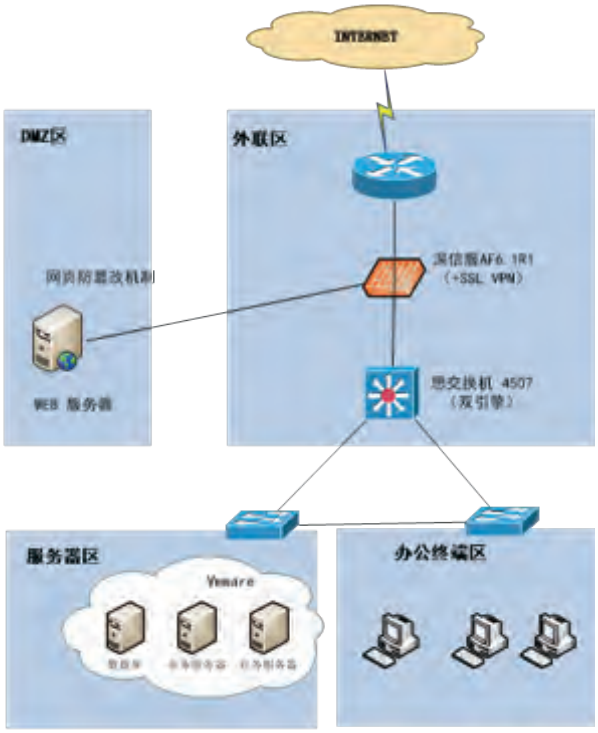


图 3：系统拓扑图

表 3 为中债资信现有设备情况。

表 2：拟搬迁设备类型及数量

设备类型	数量
小型机服务器	6 台
虚拟化集群 PC 服务器	32 台
非集群 PC 服务器	43 台
存储	7 套
磁带库	2 套
SAN 交换机	10 台

表 3

设备类型	数量
机架式服务器	11
存储	4
光纤交换机	2
路由器	2
防火墙	1
交换机	35

二、搬迁总体思路

（一）使用裸光纤实现新、旧机房二层网络连通。本次涉及大量的虚拟机迁移工作，所需的通讯线路数量多，需求的带宽大，拟使用裸光纤进行两数据中心的互联，实现二层网络连通。

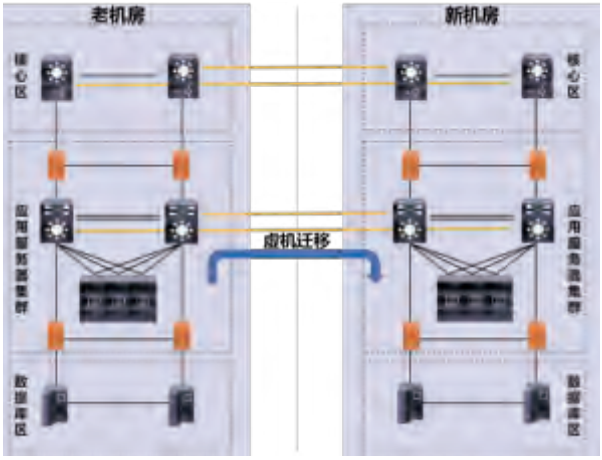


图 4：二层网络连通示意图

（二）分批次、分区域开展 IT 系统搬迁。根据北金所及中债资信各 IT 系统所属不同硬件平台及网络区域的划分，优先开展同硬件平台、同网络区域的 IT 系统的迁移。同时根据各硬件平台及网络区域内 IT 系统的重要性，分批次开展系统迁移，先进行非核心业务系统的迁移，再进行核心业务系统的迁移。搬迁过程中做好联调及快速恢复处理准备，搬迁完毕后进行完整测试、验证及调整工作。

（三）在线迁移、离线迁移、物理搬迁相结合。对于虚拟化平台，计划采用在线迁移的方式，在迁移过程中原系统保持不变，选择某个时间窗口进行最终系统的切换；对于无法在线迁移的虚拟机，将采用停机离线迁移方式。对于独立部署在物理机上的核心生产系统，计划采用在新机房新搭建环境的方式进行迁移；对于运行非核心系统的物理机，计划采用物理搬迁的方式直接迁移至新机房继续提供服务。

（四）新建与利旧相结合。为降低搬迁风险，在搬迁过程中将保留原业务系统便于系统故障回退，在系统切换完成并稳定运行后再下线原系统，确保搬迁过程中数据的安全。与此同时，本次搬迁过程将尽可能的科学、合理规划硬件设备的使用，补充少量新设备用于启动系统迁移同时对原物理设备进行充分利旧，节约了整体的搬迁成本。

北金所信息系统迁移服务器、存储等硬件系统整体建设及设备利旧示意图如图 5。

三、搬迁工作步骤

（一）搬迁准备

1、北金所及中债资信公司信息系统环境详细调研

为保证北金所及中债资信公司信息系统搬迁项目顺利进行，首先要对北金所及中债资信公司机房现有物理设备、虚拟化环境和数据库进行梳理，详细了解相互关系，形成详细的《信息系统调研报告》。原则上所有系统搬迁至新机房后网络配置、操作系统版本、数据库版本、应用架构等应保持不变，如需变化需要各方沟通一致并进行验证。

信息系统调研过程中需要北金所及中债资信公司各业务及技术负责人进行协调及配合。具体内容如下：

机房物理设备及应用系统的详细信息，包括如 IP 地址、CPU、内存、硬盘容量、操作系统类型及版本、是否外挂存储等；

应用系统整体架构及各应用系统间的访问关系；

应用系统重要性、日常工作时间段、关键时间点、允许停机时间窗口等；

数据库基础信息，包含数据库类型及版本、数据库容量、网络端口及参数配置等；

中间件版本及配置信息；

机房网络环境及相关配置信息等。

2、新机房环境服务器及网络环境配置

依据对北金所及中债资信公司现有环境的调

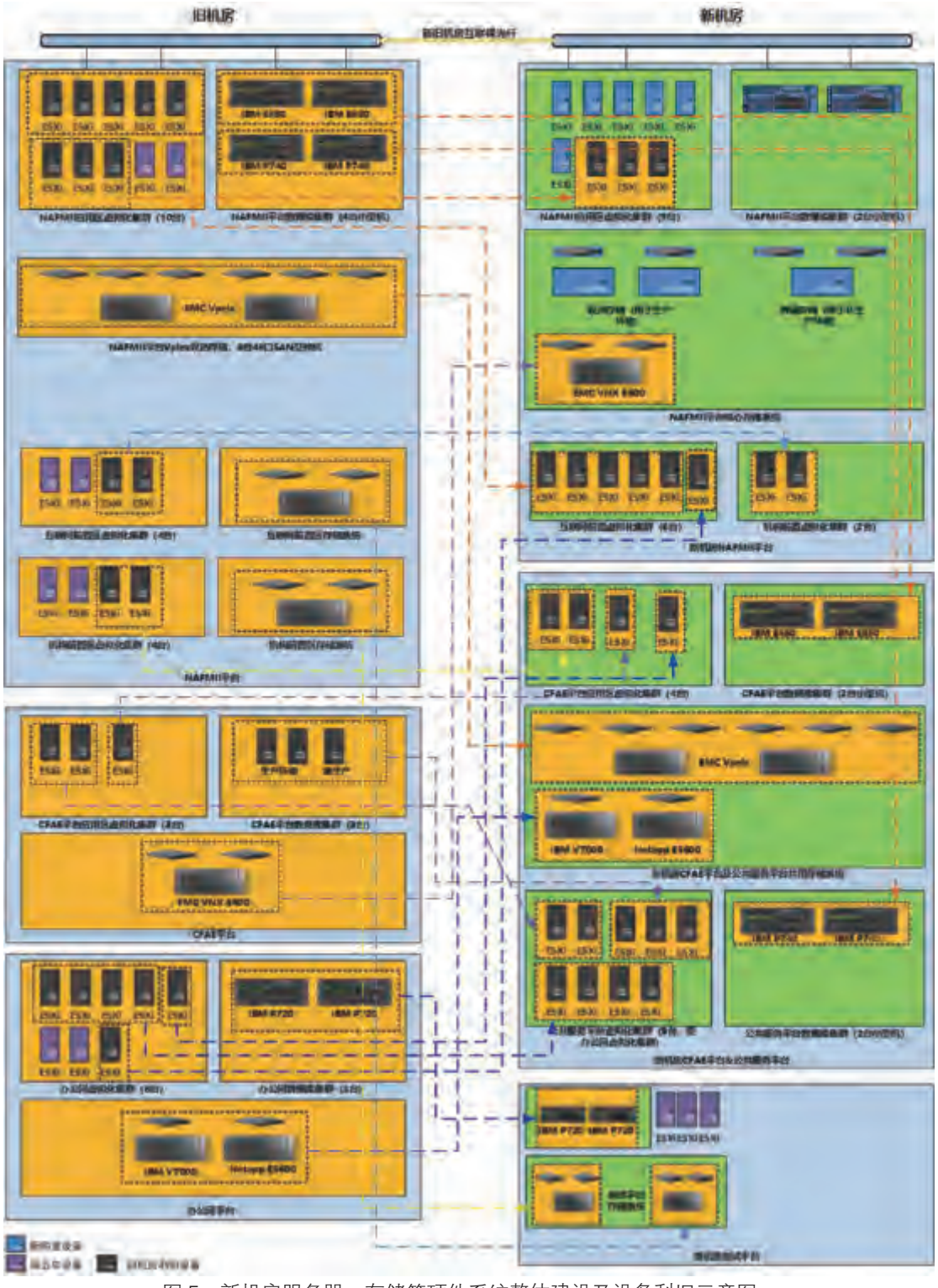


图 5：新机房服务器、存储等硬件系统整体建设及设备利旧示意图

研，在正式实施搬迁工作前，需提前在新机房规划相应的机柜，部署相应版本的小型机操作系统、PC 服务器操作系统、数据库环境以及虚拟化环境，同时新机房网络环境需满足北金所及中债资信公司各业务需求，形成《新机房服务器及网络环境配置说明》，保证迁移后业务可正常切换。

3、备品备件准备

由于应用系统的搬迁对于实时性的要求比较高，需要在实施前按照设备型号准备搬迁设备的备品、备件，在突发硬件故障情况下使用备件进行故障排除。准备的备品、备件详细记录在**《备件清单》**中，需覆盖搬迁系统使用的主要设备型号，如对服务器硬盘、网卡、光纤卡、电源等易损件配备一定数量的备件，存储设备需配备一定比例的硬盘、控制器等。

4、搬迁前的系统健康检查

搬迁前，组织对北金所及中债资信公司拟搬迁的设备及系统进行完整的健康检查，并确定设备及系统运行正常的状态边界，提前通过检查及时发现问题并予以处理，避免搬迁设备及系统本身存在的故障隐患，提高整个搬迁过程的安全性。同时**需要检查各业务系统的数据备份机制，检查各应用系统的软件许可授权，以及检查是否有对硬件设备的特殊需求等相关情况。**在搬迁过程中由各单位的系统维护人员负责进行搬迁前的数据备份及搬迁后的数据恢复等相关操作。所有检查信息需要记录到**《系统健康检查表》**中。

5、迁移方案可行性验证

为确保业务系统迁移后各业务系统正常运行，需要在迁移前对迁移方案可行性进行相关测试工作，制定相应的测试内容、测试目标、实施方法等，并形成**《迁移方案可行性验证报告》**，为后续的搬迁技术方案的选型提供指引。

表 4 为搬迁准备阶段主要工作。

（二）搬迁技术方案选型

对于北金所信息系统搬迁，计划对现有环境中的部分 PC 服务器进行 P2V 迁移，再采用 V2V 方式将虚拟机迁移至新机房；部分仍需使用物理设备的系统以及利旧设备将采用物理搬迁方式迁移至新机房。对于中债资信信息系统搬迁，由于现有生产环境已经全部虚拟化完成，所有信息系统将采用 V2V 方式迁移至新机房。整体搬迁工作主要涉及虚拟机迁移、数据库迁移和物理搬迁三大项，根据技术难度、风险、搬迁速度、搬迁成本的不同，这三项工作均有不同的搬迁技术方案。

在技术方案选型过程中，应根据拟搬迁的业务系统情况和新机房情况，按照测试演练目标，对各类技术方案进行测试对比，以选择最合适的技术方案。以下为主要搬迁技术方案相关介绍：

1、虚拟机迁移方案

虚拟机搬迁主要受到迁移停机窗口、源及目标虚拟化平台软件版本、需要迁移的虚机量、迁移的数据量、迁移方式等因素的影响。常见的服务器虚拟化迁移技术有：VMware VMotion、虚拟机克隆、虚拟机模板导出导入等。不同技术方案的对比情况如表 5。

表 5 的几种跨数据中心虚拟机迁移方案中，针对网络不可达和迁移虚拟机数量较少的情况可采用导出“OVF 模板”的方式；针对 vMotion 网络可满足带宽需求且同时满足业务迁移时间窗口时，可采用“克隆虚拟机之 1 或 2”的方式。通过“克隆虚拟机之 3”方式临时增加 ESXi 主机将虚拟机迁移到该宿主机，然后将该宿主机物理搬迁到新机房，再将其上的虚拟机迁移到新机房的 ESXi 主机上。由于本次拟搬迁系统数据量比较大，采用在两数据中心搭建高速网络进行 VMotion 的迁移方式较为适合本次拟搬迁系统环境，通过反复论证最终也选取了这样的搬迁方式。

首先将所有 vCenter 各个集群中的 DRS 自动化级别改为手动。对即将迁移的虚拟机做克隆备份，作为迁移错误或失败的回退方案。准备好新机房环境中的主机 IP 地址并加入到相应的 vCenter 中。各 vCenter 中的标准交换机和分布式交换机不做变更。各主机在迁移后，上行链路在绑定网卡时，要本着顺序一致的原则。本着先非生产后生产的顺序进行迁移，其中生产虚拟机名称后缀为 pro（生产），非生产虚拟机后缀为 uat（测试）、pre（准生产）、emu（模拟）。所有生产

表 4

工作任务	任务描述	输出文档
项目启动	召开机房搬迁工作启动会，成立搬迁项目组，明确各方职责和分工。	《搬迁工作职责说明书》
项目周期	项目周期确立后，则严格按照计划执行。	
搬迁规划	确定新机房虚拟化环境，数据库环境，SAN 网络，IP 网络	
	搬迁的业务系统做测试演练，提出搬迁要求及改造方案。	
	制定备份与应急恢复方案，论证其可行性。	
	讨论并确定搬迁的顺序、估算搬迁需要的总时间、确定设备搬迁的操作步骤、制定详细搬迁任务清单并确认每个环节的责任人。	
搬迁准备	按照要求准备新机房网络环境，完成新旧机房裸光纤连通性测试	
	对现有的业务环境进行梳理	
	完成新机房数据库 RAC 环境搭建: ● 新购存储和光纤交换机上架调试; ● 新购小型机上架; ● 小型机分区操作系统安装与生产一致; ● 小型机分区操作系统参数与生产一致; ● RAC 安装部署和调整;	《搬迁方案》
	完成新机房数据库虚拟化环境搭建: ● 新购 PC 服务器上架; ● 完成 PC 服务器 ESXi 安装，安装 ESXi 版本与生产一致; ● 完成新机房虚拟化环境配置; ● 确保搬迁前新机房虚拟化环境具备迁移虚拟机条件;	
	完成 P2V 操作: ● 对于旧机房部分物理做 P2V 前调研; ● 割接窗口停止应用保持数据静置状态，备份应用数据; ● 完成 P2V 操作，迁移物理机到虚拟化平台;	

类虚拟机迁移至双活存储，非生产类虚拟机迁移至普通存储。非生产类虚拟机可在工作时间迁移，生产类虚拟机在非工作时间迁移。对于 2 台共享 vmdk 的虚拟机，迁移时需提前关机。首先迁移第一台虚拟机，待迁移完成后，移除第二台虚拟机挂载的共享 vmdk 虚拟磁盘，然后迁移第二台虚拟机，迁移完成后，重新挂载共享的 vmdk 虚拟磁盘。开机验证状态是否正常。其中办公网在

表 5：常见服务器虚拟化迁移技术方案对比

迁移方式	导出 OVF 模板	虚拟机克隆之 1	虚拟机克隆之 2	虚拟机克隆之 3
使用场景	两个数据中心分别使用两个 vCenter，且使用嵌入式方式部署 Platform Services Controller (PSC)和 vCenter。	两个数据中心分别使用两个 vCenter，且使用外部独立部署 Platform Services Controller (PSC)和 vCenter	现有数据中心 vCenter 中的群集和 ESXi 主机可临时接管到新数据中心，原 ESXi 和新数据中心的 ESXi vMotion 网络可通讯	现有数据中心和新数据中心距离较近，可先在现有数据中心临时增加 ESXi 宿主机将虚拟机迁移到该宿主机，然后将该宿主机物理搬迁到新机房，再将其上的虚拟机迁移到新机房的 ESXi 宿主机上
优势	对数据中心 vCenter 群集等配置不需要更改	可实现跨 vCenter 虚拟机迁移	可实现 vCenter 内迁移	适用于无法通过 vMotion 迁移的环境，不需要对现有 vMotion 网络扩容
缺点	要导出的虚拟机需要关机。需要导出再导入多次中转，比较耗时间	若不是使用外部独立部署 Platform Services Controller (PSC)和 vCenter，则对现有环境改动较大。对 vMotion 网络带宽要求高	对 vMotion 网络带宽要求高	需多次中转来完成虚拟机迁移，且需要进行 ESXi 宿主机搬迁
资源需求	群集资源需满足虚拟机使用和冗余资源需求	群集资源需满足虚拟机使用和冗余资源需求	群集资源需满足虚拟机使用和冗余资源需求	需多次中转来完成虚拟机迁移，且需要进行 ESXi 宿主机搬迁。
网络	vMotion 网络	vMotion 网络 ，千兆及以上，最好万兆	vMotion 网络 ，千兆及以上，最好万兆	vMotion 网络 ，千兆及以上，最好万兆
存储	OVF 中转的 vClient ，需要同时能访问两个 vCenter，且有足够空间临时放置 OVF 模板	满足虚拟机迁移后的置备的空间需求	满足虚拟机迁移后的置备的空间需求	满足虚拟机迁移后的置备的空间需求
带宽	千兆及以上	千兆及以上	千兆及以上	千兆及以上

迁移时，先将非生产虚拟机迁移到双活存储上，待普通存储搬迁完成后再做一次迁移。

2、数据库迁移方案

数据库搬迁主要受到迁移停机窗口、源及目标数据库的版本、源及目标数据库的操作系统版本、需要迁移的数据库量、迁移范围（全库或者

是个别用户)、迁移的数据量等因素的影响。常见的数据库搬迁技术有：存储底层同步技术、操作系统复制技术、数据库逻辑备份恢复、数据库使用 rman 恢复、数据库使用 dataguard 同步、数据库利用 ASM Rebalance 特性同步、表空间传输、使用 Gloden Gate 数据同步等。不同技术方案的对比情况如表 6。

表 6 的几种数据库迁移方案中，针对数据量较少并且停机窗口长的情况可采 RMAN+ 升级脚本、EXPDP/IMPDP 方式进行数据迁移；针对需要停机窗口长数据量较小并且跨平台的情况可以采取 TTS 跨平台传输表空间方式；而针对本次数据量适中且具备一定的停机窗口的情况，最终还是选择了采用操作复杂度较低的数据泵（EXPDP/IMPDP）方式进行迁移。数据库迁移的关键在于数据一致性的验证，本次迁移中编制了特定的数据对比脚本，分别在迁移前的原库，和迁移后的新库中各运行一次，通过对字符集、用户数、用户名、无效对象、job、USER PROFILE、DBLink、触发器、关键数据表数据处理比对等信息的抓取对比新旧两个库否一致性，确认一致后，再将数据库交付应用系统使用。

3、物理搬迁方案

对于物理搬迁，主要的技术工作包含：第一，确定设备的数量、型号、配置等相关信息，联系相关设备的供货商或者厂家提供技术支持或备件支持服务，并且在搬迁设备之前必需有详细的表格记录。第二，确定设备的搬迁负责人以及联系方式，保证在搬迁过程中进行统一管理。第三，确定相关的辅助设备，如：配线架、尾纤、跳线、插排等。第四，对关键的数据包括程序、数据库、服务器配置参数等进行备份。第五，确定各服务器关键工作时间点及停机时间窗口。第六，确定各个服务器搬迁后的 IP 及路由变动情况，并做好详细的记录。第七，各种设备提前做好技术检测工作，并登记每个设备的检测情况。

（三）搬迁实施要点

由于本次搬迁实施过程中涉及的系统比较多，使用年限较久，面临很多不确定性的风险，为保障成功完成搬迁计划，在本次搬迁实施方案中有如下要点：

信息系统分批迁移：根据搬迁前对业务系统的梳理，按系统重要性分为 A、B、C 三个级别，结合各业务系统的停机时间窗口，采用不同的数据保障级别，分批对信息系统进行迁移，降低整体搬迁风险。

数据库版本及应用系统兼容性：新机房服务器将使用目前主流的小型机及 PC 服务器，对于安装的操作系统版本支持会产生变化，在迁移前需要充分验证其与原数据库及应用系统的兼容性。

搬迁顺序：由于搬迁信息系统较多，相互之间可能存在数据依存关系，在搬迁实施前，需要详细梳理各系统间的相互关系，确保合理的搬迁顺序，保障系统能顺利切换。

表 6：常见数据库迁移技术方案对比

迁移方案	申请停机时间	对迁移的数据量大小是否敏感	操作复杂度	方案成熟度
RMAN+升级脚本	长	很敏感，停机后做全量数据的迁移	中	高
EXPDP/IMPDP	长	很敏感，停机后做全量数据的迁移	低	高
TTS 跨平台传输表空间	长	很敏感，停机后做全量部分数据的迁移	中	中
XTTS	中	不敏感，停机后做增量部分数据的迁移	高	中
GoldenGate	短	基本不敏感，源和目标环境一直处于同步状态	中	中
Dataguard	短	不敏感，切换前基本做完同步	低	高

数据备份：需要对当天计划搬迁的业务系统进行数据备份操作。对于数据量较小的业务系统，可以在计划搬迁当天进行数据的完全备份；对于数据量较大或者搬迁时间要求比较紧的系统，可以在计划搬迁的前一天做数据的完全备份，计划搬迁的当天做数据的增量备份。

设备标识与名称：在信息系统调研及搬迁过程中涉及设备种类繁多，为了使整个搬迁过程设备更加清晰，所有物品、物件包括信息系统需做到名称统一、标识唯一。

机房环境确认：在机房环境建设中，应加强供电、空调、空间、设备承重、综合布线、服务器及网络环境等方面的检查，确保新机房环境能满足迁移实施的实际需求。

搬迁后验证：对于迁移后的系统，先由技术人员进行技术验证，验证没有问题，再通知相关业务部门进行业务验证。

业务系统状态检查：在搬迁实施前需要检查当天计划搬迁系统的日志状态，确认系统功能是否正常，搬迁完成后需进行相同的状态检查操作，所有检查需各负责人员签字确认。

业务系统关闭：对搬迁系统进行停机操作，要遵循如下原则，先关闭外围的应用，其次关闭中间层的应用，其次关闭数据层的数据库。

表 7

工作任务	任务描述	负责单位/部门	输出文档
准备工作确认	确认搬迁准备工作是否完成。	北金所系统运行部	《搬迁步骤确认表》
业务系统检查	确认是否是当天计划搬迁的业务系统；业务系统功能检查，确保迁移前业务正常。	北金所、中债资信业务和技术部门	
数据备份	对当天计划搬迁的业务系统进行数据备份操作。对于数据量较小的业务系统，可以在计划搬迁当天进行数据的完全备份。对于数据量较大或者搬迁时间要求比较紧的系统，可以在计划搬迁的前一天做数据的完全备份，计划搬迁的当天做数据的增量备份。	北金所、中债资信技术部门	
业务系统关闭	对搬迁系统进行停机操作，要遵循如下原则：先关闭外围的应用，其次关闭中间层的应用，其次关闭数据层的数据库。	北金所、中债资信技术部门	
搬迁	按既定方案进行硬件设备、系统和数据库的迁移工作。	北金所系统运行部	
系统启动	对搬迁后的系统进行启动操作，若系统无法正常启动，立即着手解决故障或做回退处理。	北金所系统运行部	
系统状态检查	在物理设备及虚拟机迁移成功之后，需检查系统的日志状态，确认系统是否正常。如果有故障问题，立即着手解决故障。	北金所、中债资信业务和技术部门	
启动应用	启动应用之后，查看相关软件的日志，确认应用启动是否正常。如果不正常，立即着手解决故障；如果正常，由相应业务人员进行业务测试，如果测试通过，系统正式对外提供服务。	北金所、中债资信技术部门	
恢复业务运行	在确认业务应用没有问题的情况下，系统开始对外提供服务，稳定运行后进行原系统设备利旧。	北金所、中债资信技术部门	

业务系统启动：启动应用之后，查看相关软件的日志，确认应用启动是否正常。如果不正常，立即着手解决故障；如果正常，由相应业务人员进行业务测试。

表 7 为搬迁实施主要操作步骤。

（四）实施模拟搬迁与实际搬迁

模拟搬迁阶段主要根据准备阶段所制定的模拟搬迁目标开展搬迁，计划共进行两轮模拟搬迁演练，并对模拟搬迁的结果进行评估，总结经验，最后调整并优化实际搬迁方案。

实际搬迁阶段根据优化后的搬迁方案进行实施，主要步骤包括：准备工作确认、业务系统检查、数据备份、业务系统关闭、物理机虚拟设备迁移、系统启动、系统状态检查、启动应用、恢复业务运行等。

表 8

编号	风险	几率/影响	应对措施
1	搬迁需采购设备到货延期	低/大	密切关注并推动采购流程 调整系统技术方案/加快到货后调试进度
2	新机房设备出现故障	低/大	尽早接入线路，提前测试 准备相应备机、备件进行替代 安排设备商在搬迁时现场配合
3	业务系统单机运行环境	中/大	提前准备好安装相同业务系统的备机 做好系统数据备份
4	业务停止窗口超出计划	低/中	提前做好业务系统健康检查和测试 及时通报项目组调整计划
5	搬迁中硬件发生损坏	低/大	安排专业硬件平台专家 安排可能损坏硬件的相关现场备件或备机 设备运输保险
6	搬迁中软件故障	低/大	搬迁前健康检查 搬迁前进行备份 相关产品技术支持人员现场待命
7	备份介质损坏	低/大	使用可靠性高的介质备份 核心业务多拷贝备份 搬迁前进行备份恢复演练
8	搬迁过程中设备丢失	低/大	设备运输保险 设备清单多次多方核查 搬迁过程专人监控
9	搬迁过程中数据泄露	低/大	项目组成员签订保密协议 设备清单多次多方核查 搬迁过程专人监控
10	备机、备件损坏	低/小	备机到场后进行测试 多个相同备件 专业技术支持
11	备机、备件数量不够	低/中	优先满足核心业务系统 备件库有足够的备件
12	搬迁后设备无法正常开机	中/大	在搬迁前进行系统健康检查及开关机测试 使用备机、备件进行替代
13	搬迁后应用系统异常	中/大	在搬迁前进行系统健康检查及开关机测试 使用备机、备件进行替代
14	备份时间不足	小/中	根据业务情况及数据量大小，评估大致时间， 预留足够备份的时间
15	应用系统由于原开发厂商服务到期并	中/大	做好充分的应用系统梳理 做好系统数据备份

编号	风险	几率/影响	应对措施
	无法获得原厂支持，影响应用系统搬迁		搬迁前进行备份恢复演练 搬迁完成后做好充分的业务测试
16	虚拟机 vMotion 迁移失败	低/大	虚拟机不会从现网环境的 ESXi 主机删除，不会影响现网业务，可重新启动迁移
17	虚拟机迁移完成后无法启动，迁移回旧环境仍无法使用	低/大	启动现网环境克隆备份，恢复业务
18	互联网线路割接后，无法完成公网连通性测试	低/大	将互联网线路恢复至原机房，恢复业务
19	新、旧机房二层网络线路中断	低/大	保障二层网络线路冗余，搬迁前做好充分的连通测试工作

（五）搬迁验证

在每项业务系统搬迁结束之后，需要对相关硬件、信息系统及业务数据进行一次全面检查及验证，保证业务系统正常切换。检查内容主要包含以下几个方面：

物理设备检查：主要检查设备外观、电源状态、前面板指示灯、磁盘指示灯、系统风扇运转等。

虚拟化环境检查：主要检查虚拟化平台的运行状态，包括虚拟机总体配置、整体运行负载、日志、事件等。

操作系统状态检查：主要检查各系统的 CPU 使用率、内存使用率、磁盘空间、系统日志等。

数据库检查：在数据库迁移完成后业务组需要对数据库整体情况进行检查，确保数据库数据的一致性。

业务验证：业务组负责相关业务模拟操作、检查搬迁后的业务功能是否正常及业务数据是否完整等。

在搬迁验证完成后，对原系统需利旧设备进行搬迁。

（六）风险控制方案

本次搬迁主要为硬件设备、虚拟化及数据库迁移，针对搬迁中潜在的风险，制定服务器迁移风险控制预案和数据库搬迁风险控制预案。针对设备无法启动、切换不成功、复制不成功、同步失败、网络中断、拷贝中断、备份时间不足、虚拟机或服务无法启动、业务数据丢失等情况，提前做好风险应对计划，降低风险造成的影响。表 8 为主要的风险控制方案内容。

（七）故障处理及回退方案

针对本次北金所及中债资信公司信息系统搬迁，原则上采用保留原有硬件系统和网络系统的方式对信息系统进行搬迁，当迁移验证成功后，再下线原有系统，确保搬迁出现问题时能够回退到原系统；对于仍需要运行在物理服务器上的系统将采用物理搬迁方式进行搬迁。

搬迁前对设备进行系统健康性检查，如果发现存在故障情况，必须先处理故障，故障解除后才能实施搬迁工作。

搬迁过程中由于人为因素导致设备硬件故障，由搬迁服务公司负责处理和修复；对于自然损坏的设备由设备服务商负责处理和修复，并由搬迁项目组成员共同确认。

当物理搬迁过程中出现硬件整体损坏而无法修复的情况时，应及时启用备机进行相应的数据恢复。

当虚拟机迁移过程中出现迁移时间超时、虚拟机启动异常、数据库异常、业务验证失败等情况时，则启动原先的业务系统，同时对网络进行回切，避免业务中断。

DevOps在证券互联网研发中的应用与实践

张永启 向元武 于娜娜 / 中泰证券股份有限公司 科技研发部

近些年金融科技在证券行业发挥的作用越来越重要，运用金融科技赋能业务发展，通过个性化服务构建护城河，将金融科技与业务创收和降本增效相结合开始成为证券从业人员所关注的问题，如何提升研发交付效率、小步快跑、快速迭代是所有证券行业科技研发团队共同关心的话题。敏捷为快速迭代提供了理论思想和方法指导，DevOps 为敏捷落地提供了补充和工具支持。

中泰证券股份有限公司科技研发部互联网研发团队通过对 DevOps 相关理论和技术的研究、分析，设计并实现了蜂鸟效能管理平台。通过蜂鸟效能管理平台实现了编码后续研发环节的降本增效。DevOps 是一套创新且有效的文化和思想，本平台借鉴其中的持续集成、持续交付和持续运营的关键思想，并结合互联网研发过程遇到的实际情况，解决了研发、测试和运维等角色沟通协作中遇到的一系列问题，实现了产品多环境交付、流程可视化、测试自动化、运维智能化、流程规范化和效能指标可视化等功能。蜂鸟效能平台上线后的应用实践结果表明，通过运用 DevOps 相关理论和技术能够提升互联网研发在市场快速变化的过程中实现产品应用的快速迭代，从而达到减少产品试错与迭代过程中的时间成本和技术人力成本，为公司业务创收提供技术保障的目的。

一、背景及意义：

DevOps 因其先进性和全面性，已被认为是软件工程的第三次革命；由 PUPPET 和 DORA 联合发布的《2017 State of DevOps Report》报告中，故障恢复时间缩短了 96 倍，业务需求从提出到投产的周期从 3 个月到 6 个月缩短到 3 周甚至更短，使得企业更好适应市场变化。DevOps 已经被证实能在 IT 和商业两方面提升效率。

DevOps 定义：DevOps（Development 和 Operations 的组合词）是一组过程、方法与系统的统称，用于促进开发（应用程序 / 软件工程）、技术运营和质量保障（QA）部门之间的沟通、协作与整合。它是一种重视“软件开发人员（Dev）”和“IT 运维技术人员（Ops）”之间沟通合作的文化、运动或惯例。透过自动化“软件交付”和“架构变更”的流程，来使得构建、测试、发布软件能够更加地快捷、频繁和可靠。它的出现是由于软件行业日益清晰地认识到：为了按时交付软件产品和服务，开发和运维工作必须紧密合作。

特别是敏捷迭代已经成为金融行业研发团队的主流研发模式，这对开发、测试、运维提出了更高效的要求。

中泰证券互联网研发团队采用敏捷研发模式进行团队间的协作，敏捷的实施需要通过小迭代形式不断的交付应用产品。敏捷开发驱动开发人员更快的交付代码，新的代码需要被更快的测试，并需要频繁的被部署到开发、测试和生产，由于运维和测试不能尽快的参与到软件开发生命周期，导致交付流水线阻塞的情况，而通过 DevOps 的运用很好解决了这些问题。

在 DevOps 实施的过程中，涉及的角色主要包括开发、测试（质量）、运维三个角色，见下图：

图 2：开发、测试（质量保证）、运维

其中研发主要关注产品研发的高效、稳定、快速的实现，以及对应的产品开发完成后，交付制品对应上线时间点能够可预期；运维则更

图 1：敏捷迭代流程

多的关注如何通过自动化运维和持续监控等工具降低产品上线后的维护成本；测试（质量保障）角色则关注研发提交过来的产品能够尽快的得到测试，因此在提高质量保障效率的过程中，该角色更多的关注产品的持续自动化测试，以及产品交付质量的提升。而 DevOps 实现了将研发、运维和质量三个角色统一起来，实现了研发、运维和质量的一体化，同时通过持续集成和持续交付的能力，使运维人员更早的参与到产品的交付过程中区，减少了不同角色之间的交付壁垒。

图 3：DevOps 过程

综上所述，尽管新工具思想的推进在一定程度上能够提升产品的交付效率，但由于企业自动化程度低、软件开发流程的不规范导致的交付效率慢、交付流程不规范、线上故障反映不及时、运营数据获取困难等问题的存在，使企业在实际的产品交付过程中依然不能实现快速交付有价值的产品给用户。这就需要一个平台解决以上问题，但是目前市面上已有的相关产品存在不能和流程结合以及不支持混合制品（容器和非容器）的持续集成与持续交付，且不能获取实际场景的业务数据。因此，一个能够解决当前困境的统一自研 DevOps 平台变得尤为重要。

二、中泰证券 Devops 蜂鸟效能平台整体功能规划

蜂鸟效能平台是一个以 DevOps 相关理念

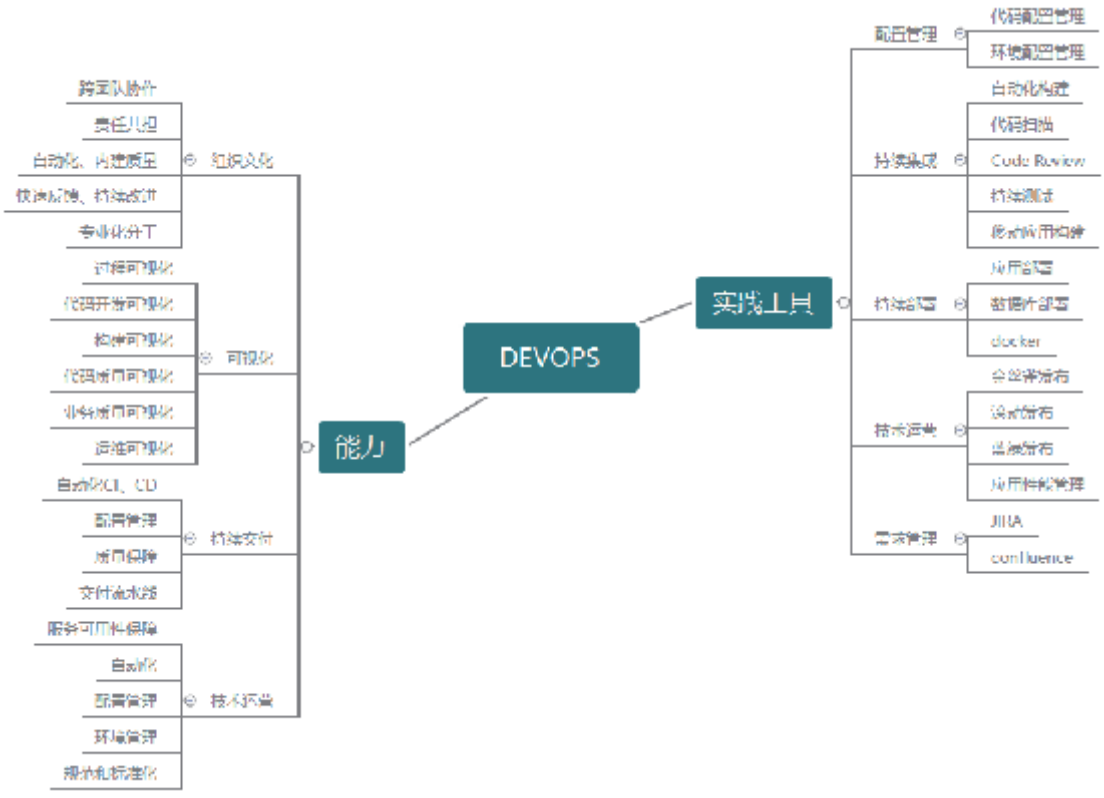


图 4：DevOps 能力地图与实践落地

为指导思想，结合证券行业安全、合规等需求特性实现的一个集多环境（开发、测试、预发布、生产）持续集成 (CI)/ 持续发布 (CD)、代码质量检测、自动化测试、上线流程审批、研发效能数据跟踪及报表统计的综合效能管理平台。通过蜂鸟效能平台在互联网研发过程中的应用，提升了互联网研发在市场快速变化的过程中实现产品应用的快速迭代，从而达到减少产品试错与迭代过程中的时间成本和技术人力成本，并为公司业务创收提供技术了保障的目标。

依托 DevOps 相关理念，结合当前互联网技术中最前沿的容器化技术、容器编排管理 kubernetes、微服务架构、配置中心、分布式消息存储队列、静态代码扫描和自动化测试等技术，蜂鸟效能平台实现了具有持续集成、持续交付和持续运营能力的统一综合效能管理平台，各阶段详细技术如图 5。

蜂鸟效能平台整体技术架构划分为三层，最底层为基础设施层，该层主要为各混合云环境下的基础环境，如 私有云、华为云和行业云等环境下的开发、测试和生产环境，建立在基础层之上搭建了支撑平台的工具，形成了平台的工具层，如需求管理 jira、代码管理 Git、容器管理 k8s 和监控工具等，通过工具层提供的能力，建立并实现了价值流层，价值流层对应的功能直接为对应的职能化人员赋能，主要为持续集成、持续交付和持续运营。

三、蜂鸟效能平台相关技术节点简介

3.1 与项目管理平台 JIRA 集成

为了将具体需求与迭代上线进行对应，从而达到系统上线需求可追踪，同时也为后续对需求进行价值分析提供基础数据，因此需要将项目管理平台 JIRA 与蜂鸟效能平台持续交付进行打通。研发人员在蜂鸟效能平台进行提测和发布上线时可以根据提示选择对应的 STORY，从而完成提测、上线与 JIRA 项目的关联打通。

3.2 CI/CD 流水线整体设计：

蜂鸟效能平台 CI/CD 功能主要包括静态代码扫描、开发环境 CI/CD、测试环境 CI/CD 和自动化测试、上线流程审批、生产环境 CI/CD、交付制品的环境流转等功能。其中应用产品的制品在各环境中详细流转逻辑图如图 6。

从图 6 可知，制品在多环境 CI/CD 流转的

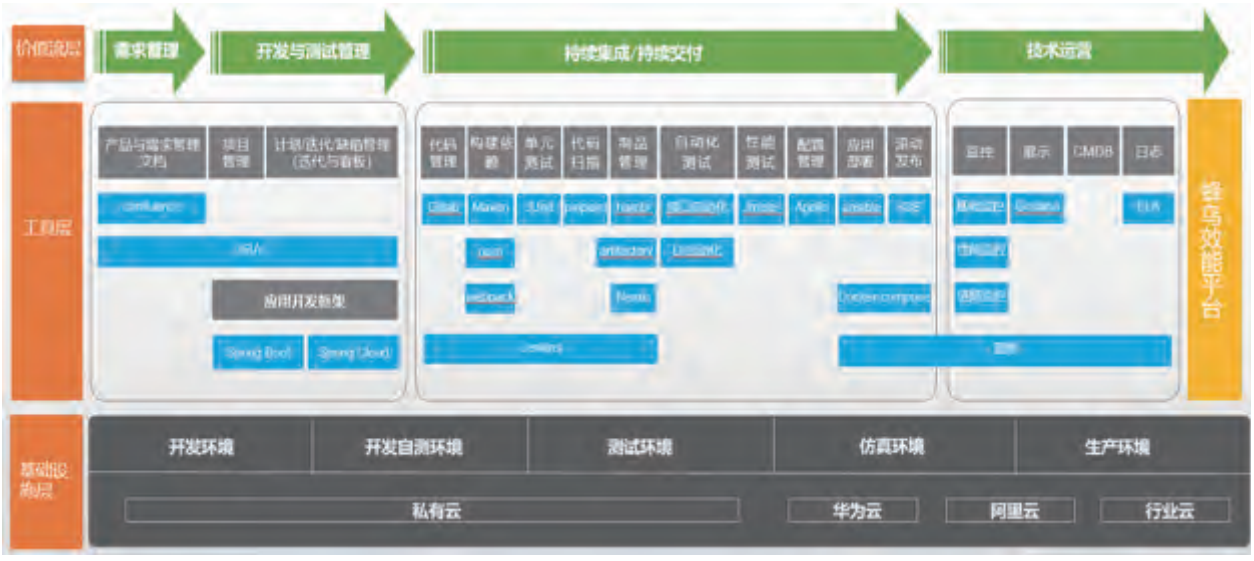


图 5：蜂鸟效能平台规划

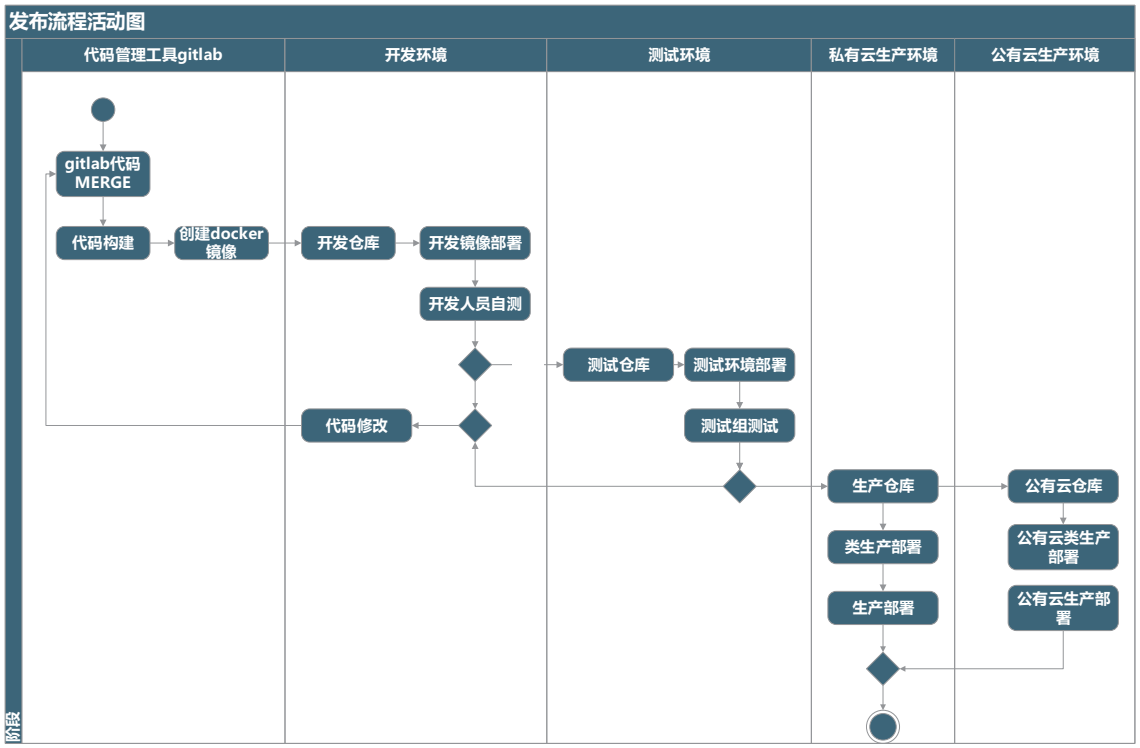


图 6：CI/CD 流程

过程中，主要包括三个环境的流转，分别为开发环境、测试环境和生产环境，其中开发环境采用自动构建、自动集成和自动部署的方案，测试环境的流转与发布根据技术经理在蜂鸟效能平台上的提测，测试组人员可以根据自身需要进行按需自动化发布部署，发布部署完成后，可以对相应的功能模块进行自动化测试。制品在生产环境的流转与发布方式和测试环境的发布部署方式类似，在蜂鸟效能平台上走完审批流程后根据实际需要，运维人员按需自动化发布部署。蜂鸟效能平台实现了线上审批流程和测试、发布操作的关联控制，保证每次提测和发布上线都有严格的流程把控。

其中对于容器化应用各环境流转发布部署细节如下：

1）开发环境镜像生成与发布：在 Gitlab 上创建工程后，研发人员可以自助在平台上对该 Gitlab 工程绑定自动构建和部署模块，当该工程主分支发生 branch 合并时触发自动构建，镜像创建后，会将对应的镜像推送到镜像仓库，然后再触发自动化部署脚本将该镜像自动部署到开发环境。

2）测试环境下的镜像流转与发布：在开发人员将对应的需求开发完后，在蜂鸟效能平台上进行应用产品提测，提测后测试组能够在蜂鸟效能平台上的测试模块看到提测的具体镜像内容，根据提测详情，可以实现一键自动部署，然后再对测试环境的镜像进行自动化功能、接口、性能和 UI 测试。

3）生产环境下的镜像流转与发布：在走完产品上线流程审批后，运维人员能够在蜂鸟效能平台上看到具体的上线流程信息，根据实际情况进行自动化发布部署、回滚和复核。

通过 CI/CD 流水线，使产品、UED、研发、测试、运维和合规等职能化角色可以在其对应职责的权限下完成产品上线过程中对应的具体操作，如测试人员完成自动化测试操作、运维人员完成自动化部署操作和合规人员完成合规审核操作等，最终达到产品交付的目的。同时蜂鸟效能平台隔离了开发、测试、生产环境，

对应的角色只能在对应的环境进行操作，操作环境的隔离符合《证券基金经营机构信息技术管理办法》的相关要求。

3.3 代码管理及构建

代码的科学管理对团队高效协作以及流程规范具有特别重要的作用，蜂鸟效能平台采用 Git 作为代码管理工具。GitFlow 模式是若干模式的集大成者，包含一个主干分支、一个开发分支、许多的特性分支、许多的发布发布分支和 Hotfix 分支，以及许多的合并规则，通过 GitFlow 模式的运用，能够解决开发过程中大部分代码协作的问题。同时通过分支的管理，也为后续开发环境的 CI/CD 奠定了基础。

3.4 制品库管理

蜂鸟效能平台的制品仓库在持续交付的过程中扮演着中转站的作用，如何结合制品仓库实现制品在不同环境中的流转对保持制品交付前后的一致性与可追溯性有着重要的作用。对于当前的制品主要分为 Docker 镜像类和非镜像类制品，为了能够对 Docker 镜像类进行管理，镜像仓库采用开源 Harbor 进行管理，对于非镜像类制品采用 Artifactory 进行管理，为了能够达到环境流转的效果，采用基于仓库的不同账号之间的权限管理方案，具体见关键模块方案。对于唯一性问题，容器类制品采用镜像 id 非容器类制品采用 MD5 码保证制品的唯一性。

3.5 接口管理

蜂鸟效能平台接口管理功能对于不同的研发角色作用不同，开发人员能够利用接口管理功能进行前后端接口调用、多项目接口统一管理、接口调试和多团队协同开发；测试人员能够基于接口管理功能中登记的接口进行简单接口测试、场景化接口测试；运维人员可以基于接口管理功能中登记的接口实现业务监控；产品人员可以快速进行数据统计。

3.6 配置中心

蜂鸟效能平台的配置中心能够实现对不同环境、多云环境的系统参数配置进行管理，同时配置中心也是多环境（开发、测试、仿真、生产）CI/CD 的关键。蜂鸟效能平台的配置中

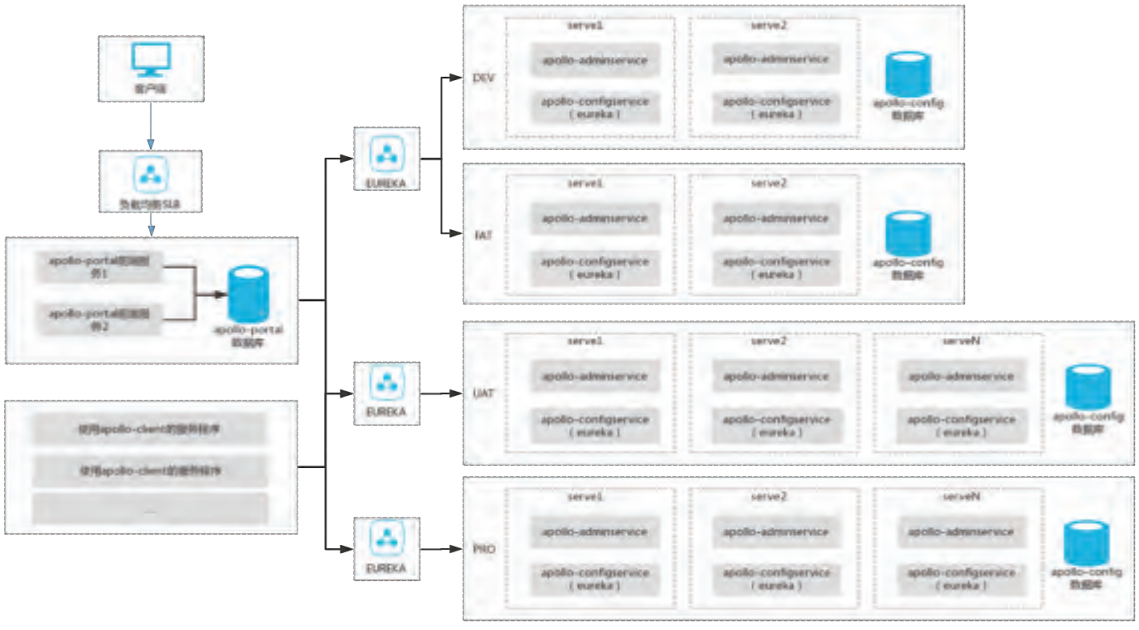


图 7：分布式配置中心

心基于 Apollo 实现，Apollo 是一个开源的分布式配置中心，能够集中化管理应用不同环境、不同集群的配置，配置修改后能够实时推送到应用端，并且具备规范的权限、流程治理等特性，适用于微服务配置管理场景。配置中心实现了开发、测试、生产多环境的系统参数配置功能，同时支持多云架构下的分布式系统配置管理。配置中心为CI、CD的平滑交付提供了技术保障。

3.7 单元测试、静态代码扫描及开发环境 CI/CD

单元测试能够让开发人员在提测前发现新增变动对系统可能造成的不利影响，并通过将单元测试与开发环境 CI/CD 进行结合，当开发人员进行代码提交或者进行代码 merge 时将触发自动单元测试并将结果反馈给对应研发人员。

为了提高研发过程中的代码质量并尽快发现已有系统代码中存在的漏洞缺陷，蜂鸟效能平台提供了静态代码扫描功能，静态代码扫描功能可以对研发人员的代码进行分析并进一步提升编码规范。静态代码扫描功能需要能够识别代码中一些常见的漏洞，如资源类问题（资源释放、无效指针等）、安全性要求（数据污染、注入等）、潜在的缺陷（数组越界、初始化、除零错误、空指针引用等）、多线程和同步性（双重锁定、未释放的锁等）和异常处理（NullPointerException）等。静态代码扫描平台采用增量扫描和全面扫描相结合的方式，日常开发对于不断的代码提交采用自动增量扫描，便于快速发现新增代码中的缺陷，同时结合定时全量扫描和提测前全量扫描的方式，发现代码中所有的缺陷，只有当高危、中危、低危等级的缺陷全部修复完后才能由开发人员在蜂鸟效能平台上提测版本给测试人员，提升了开发人员提测版本的质量和安全性。

3.8 自动化测试及测试环境 CI/CD

蜂鸟效能平台的自动化测试功能实现了互联网研发团队测试人员的效率飞跃，通过将自动化测试平台与持续交付的流程进行结合大大提升了测试效率。在蜂鸟效能平台上，测试人员能够看到研发人员提测的具体内容，并识别出具体的制品版本及唯一码，测试人员可以在蜂鸟效能平台上对开发人员提测的版本进行一键部署和一键自动化测试，最后测试的结果将以报告的方式反馈给研发人员。

通过蜂鸟效能平台的接口管理功能与自动化测试功能的集成。目前已经支持 UI 及接口自动化测试。UI 自动化基于 Appium 实现，Appium 要能真正自动化手机上的应用必须依赖

图 8：UI 自动化流程

于各个移动平台所带的自动化框架；IOS 平台目前依赖于 XCUITest 实现，安卓目前主要依赖于 Uiautomator。框架提供的是运行库，运行库运行在移动设备上。

AppiumServer 服务起来后会在移动设备上安装一个帮助自动化的应用，可理解为“控制许可”或者“代理”应用，通过这样应用可以编译我们自动化给出的指令，然后按指令测试移动设备上的应用。

蜂鸟效能平台的接口自动化功能基于接口分层测试设计的思想，采用 python+unittest+ddt 框架自研实现。把测试数据与测试代码完全分离，将数据操作、用例配置、日志记录、接口请求等公用方法封装成单独类，使用 DDT 数据驱动工具管理每个接口的多种测试场景，使用 unittest 组织、执行多个接口的测试用例集合，通过添加多种断言形式，如接口的状态码、返回值、差异化（diff）对比等对接口测试结果进行判断，最后通过 HtmlTestRunner 生成测试报告，把返回的测试结果用图形和文本形式形象的展现出来。

接口自动化测试功能是 Devops 实践中不可或缺的一部分，具备持续测试能力，大大提升测试效率，使测试人员快速适应敏捷开发工作模式，从而减少了产品迭代过程中的时间成本和技术人力成本，为产品快速迭代和发布提供了质量保障。

3.9 混合云管理及生产发布

在真实的部署环境中，经常涉及多云环境下的发布管理，为了实现多云环境下的制品流转及发布部署，蜂鸟效能平台实现了一套混合云环境的发布管理功能。

中泰证券互联网应用系统的部署环境为一个混合云场景，蜂鸟效能平台通过对混合云环境资源的整合，解决了混合云环境的 CI/CD 发布部署和系统监控问题，具体方案如图 11。

通过该方案与 CI/CD 制品流转设计相结合，使研发人员交付的应用制品可以通过蜂鸟效能平台实现多环境流转，最终发布部署到混合云的环境中去。

3.10 自动化监控及技术运营

为提升线上问题和故障的发现、反馈效率，蜂鸟效能平台集成了自动化监控功能，自动化监控能够提高运维的效率，并能够满足频繁发布部署过程中的应用监控问题，通过对应用服务的自动化监控与故障自愈相结合能够在用户无感知的情况下修复线上故障，同时在蜂鸟效能平台上对线上故障数据进行跟踪收集分析，

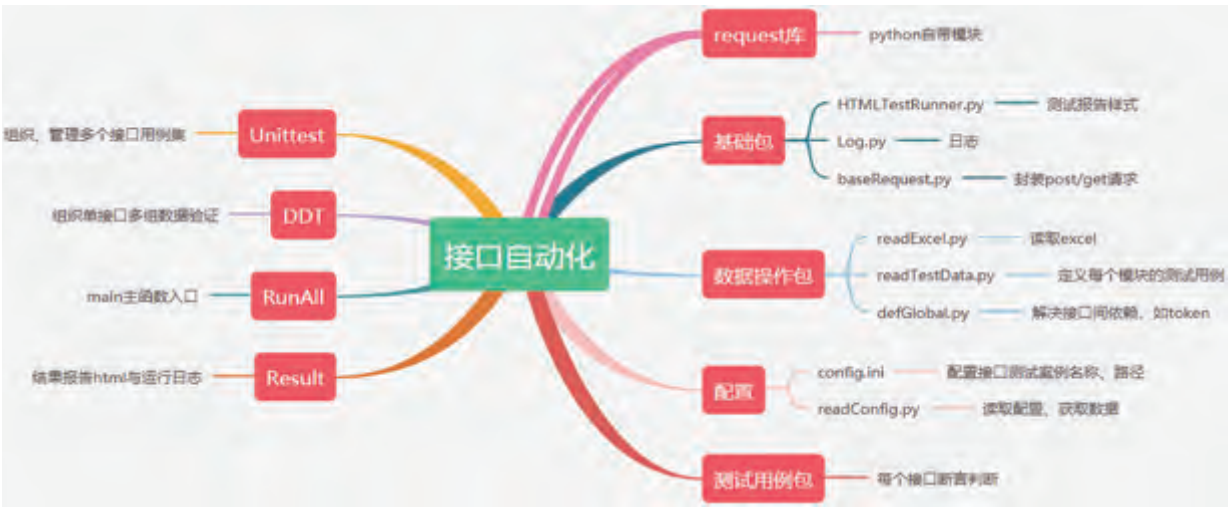


图 9：接口自动化体系

图 10：接口自动化流程

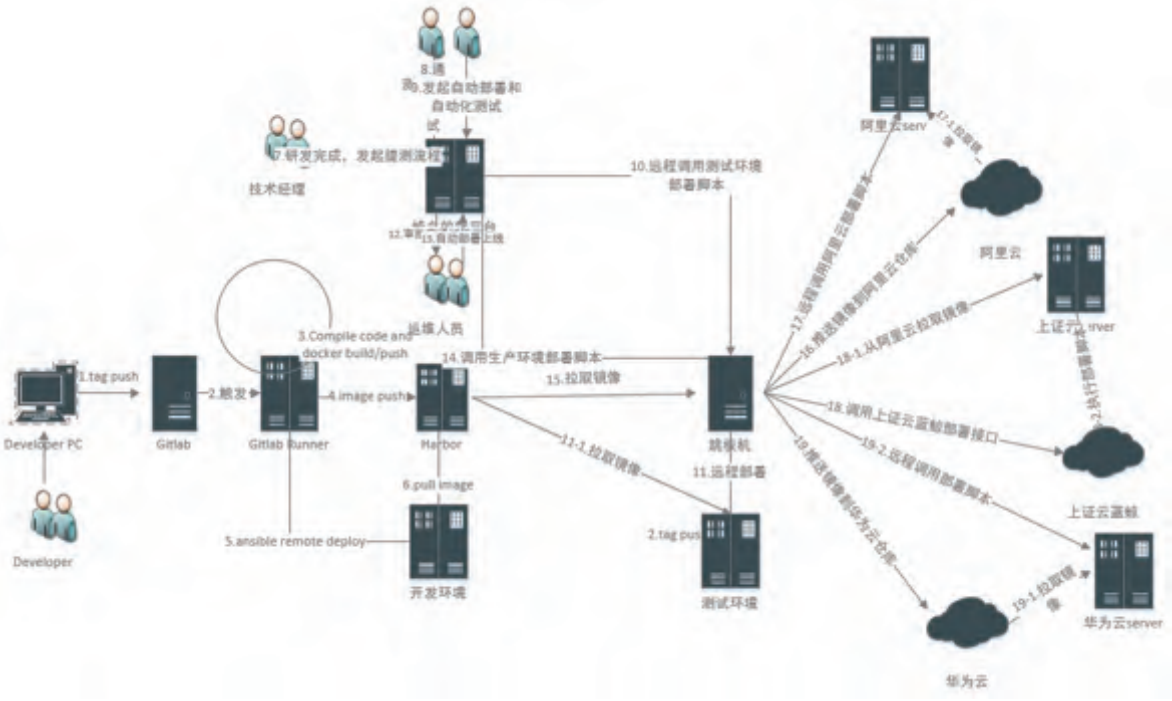


图 11：混合云多云环境部署

达到尽快反馈给研发人员并能够对故障回溯具有参考作用。

3.11 效能指标跟踪及改善

效能指标能够体现研发过程的实施情况，客观的效能数据可以对我们研发改进起到指导效果，没有客观数据和成功标准就无法做到持续反馈和持续改进。蜂鸟效能平台中关键效能指标见图 12。

研发效能数据多维度展示产品交付各阶段的数据，如开发阶段的代码数据统计、测试阶段的自动化测试数据统计、运维上线后的故障数据统计等；其中流水线指标从交付效率和吞吐率两方面主要反映研发过程整体效能情况；项目 / 需求指标能够从需求的成本和上线后的价值两个角度对需求进行价值数据衡量；研发指

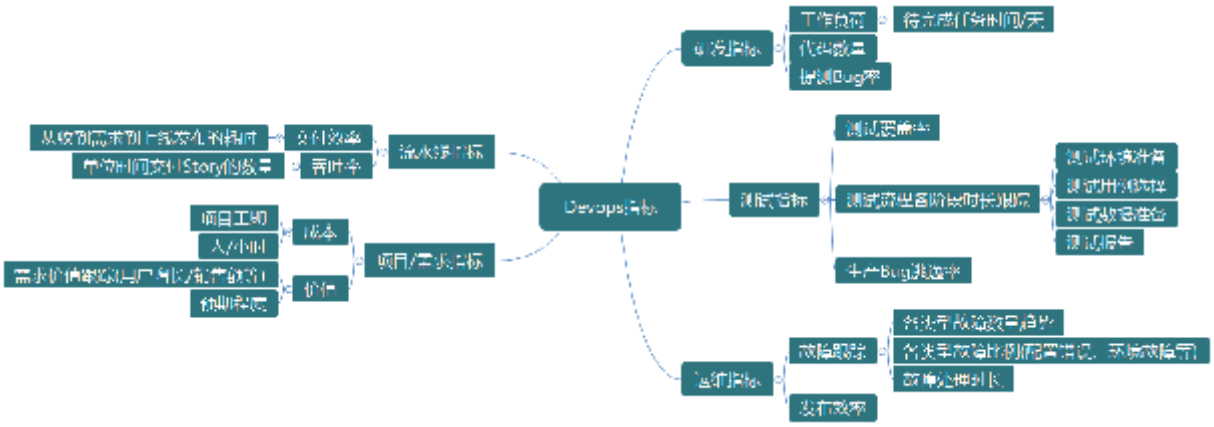


图 12：效能数据

标反映了当前研发人员的工作负荷以及产出及质量情况；测试指标对产品质量保证具有关键指导的作用；运维指标能够反映运维工作效率的情况，如线上故障情况及功能上线发布效率。

四、DevOps 蜂鸟效能平台关键功能截图

目前蜂鸟效能平台已经覆盖中泰证券互联网金融全部业务，接入项目模块 180 多个，支持容器及非容器应用，累计 CI/CD(持续集成 / 持续交付) 次数 6 万 +。

1）产品关键截图一：提测及上线流程跟踪如图 13。

该图为一个完整提测、上线发布流程图，各职能化角色都有参与到该流程中，开发人员提测、测试人员执行测试、技术经理确认上线发布、产品经理发布前验收、产品及技术负责人确认上线、合规确认发布流程、运维人员 A 执行发布、运维人员 B 复核发布、业务部门生

图 13

产验收。

2）产品关键截图二：提测详情页关键截图如图 14。

该图为提测详情图，技术经理可以通过该提测详情页完成容器或者非容器化的提测。

3）产品关键截图三：蜂鸟效能数据截图如下图 15。

该图为效能数据对研发阶段的某一场景进行展示。蜂鸟效能平台通过多维度对不同交付阶段进行数据跟踪并绘制全面的图表对全过程进行效能分析。

4）产品关键截图四：自动化测试结果关键截图如图 16。

通过自动化测试报告能够详细完整的查看具体的功能模块测试详情。

5）产品关键截图五：静态代码扫描关键截图如图 17。

通过静态代码扫描的报告详情，能够查看

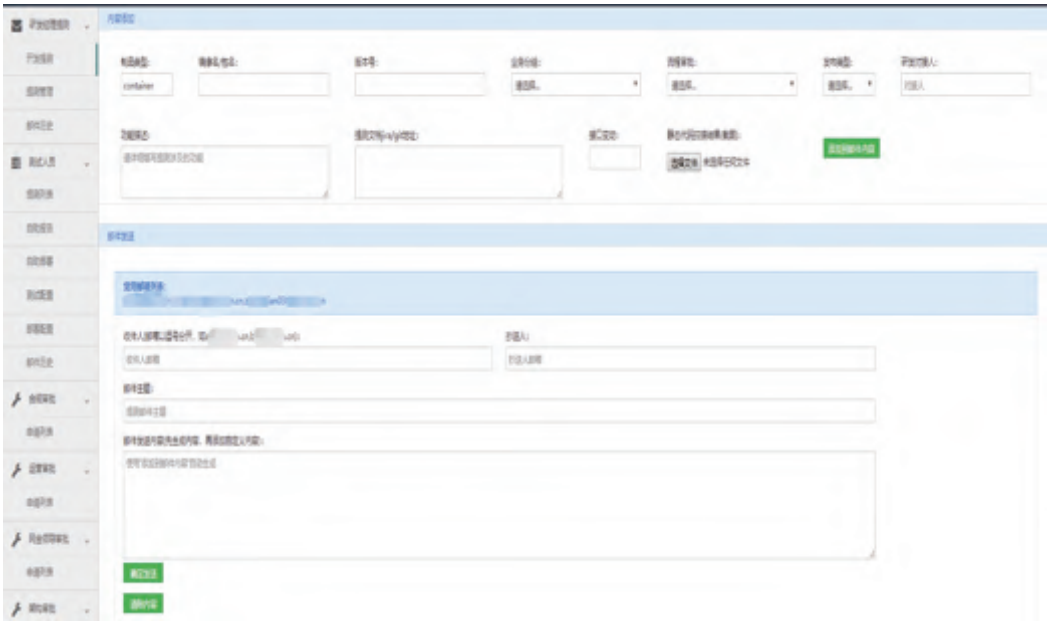


图 14

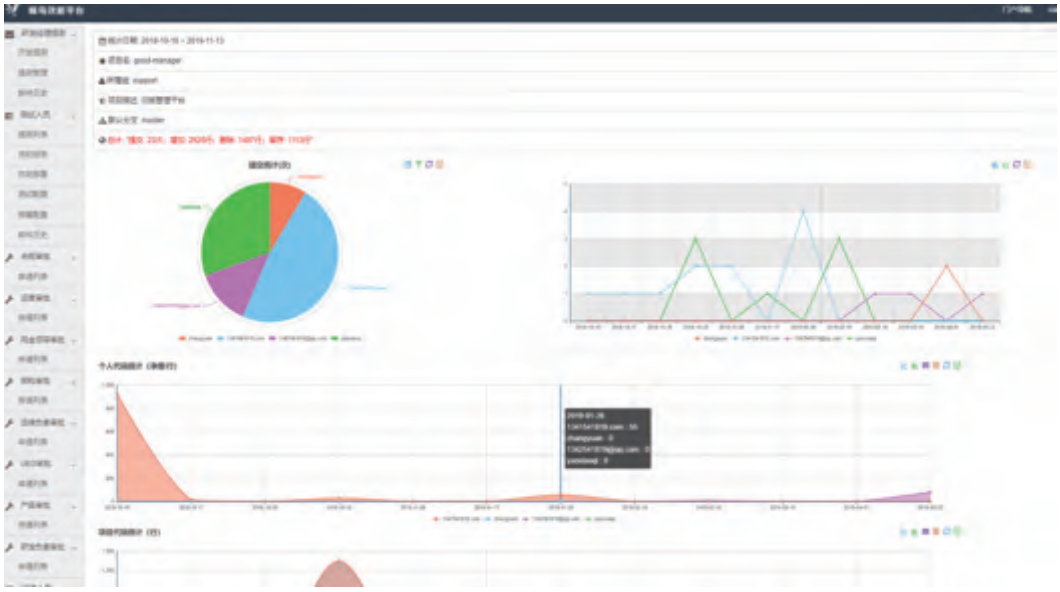


图 15

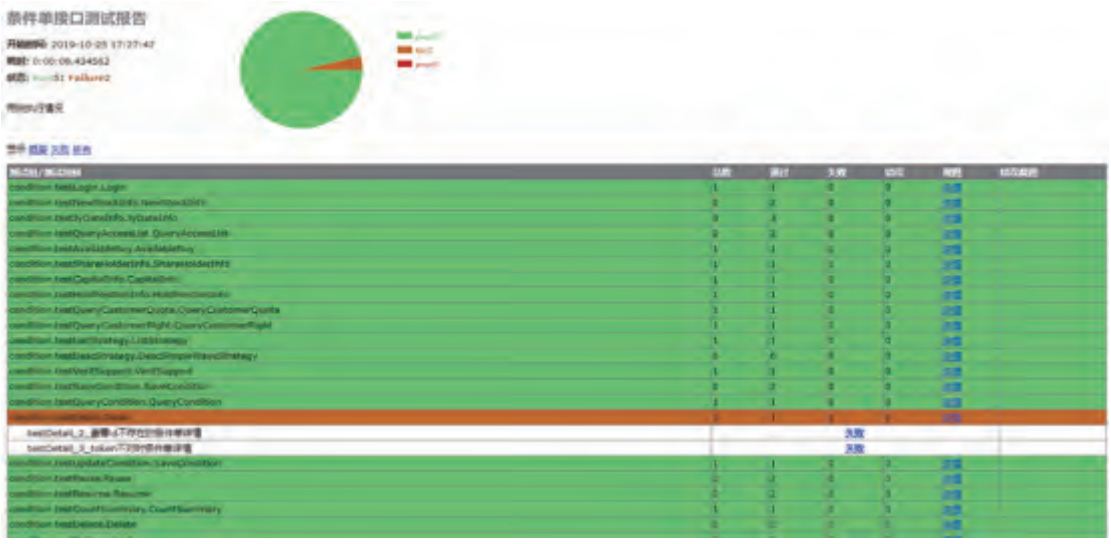


图 16

图 17

当前项目的缺陷情况，同时通过缺陷展示快速定位缺陷位置。

五、总结

中泰证券互联网研发团队基于 DevOps 思想构建的蜂鸟效能平台上线以来，经过累计 6 万 + 次的 CI/CD(持续集成 / 持续交付)，已经成为目前互联网研发过程中不可或缺的核心研发工具平台。带来的价值如下：

1）规范研发过程，实现安全可控

蜂鸟效能平台将开发、测试、生产环境进行了有效隔离，实现了各环节操作的安全可控及合规。所有的发布操作均通过蜂鸟效能平台留痕并且操作绑定流程制度，做到了问题可追溯、减少了误操作、操作符合公司 IT 管理规定。

2）提升交付效率，缩短交付时间

蜂鸟效能平台旨在缩短开发人员完成功能代码编写到生产发布过程中的时间消耗，提升研发交付效率。

图 18

缩短了开发时间周期，通过静态代码扫描可以在 10~20 分钟内实现对一个应用系统的代码检查，提升了开发人员 CodeReview 的效率。每个迭代的时间缩短 0.5~1 天。

缩短了测试时间周期，全功能回归测试从 2~3 天人工测试降低到自动化测试 6~8 小时。每个迭代的时间缩短 2 天左右。

缩短了各环节流转的时间，蜂鸟效能平台自动化 CI/CD 次数自上线以来达到几万次，已经实现了从开发人员提交代码到生产发布过程中的自动化，节约了大量的人力成本。其中各环境下的发布部署改造前后具体参考图如图 18。

从图 18 可知，通蜂鸟效能平台的实现与应用，使应用制品在开发、测试和运维阶段的发布部署过程中，对应的发布部署效率提升 10 倍以上。

3）提升交付质量，减少线上故障

静态代码扫描可以识别一定的代码逻辑设计、编码缺陷及安全漏洞，减少了程序问题导致的生产问题。

应用发布过程中，应用包在开发、测试和生产的环境流转过程中，应用包不用重新编译打包，实现了同一应用包多环境流转，系统、业务配置参数通过配置中心读取，减少了因为应用包变更和参数配置问题导致的生产发布问题。

通过容器化技术实现了各环境对应用的隔离性和封装性，减少了因为开发、测试、生产环境的差异性导致的生产发布问题。

FPGA技术在极速交易场景的应用示范

金乐人 郑宇飞 黎云芃 / 华泰证券股份有限公司

现场可编程门阵列（简称“FPGA”）技术是近年来证券期货市场发展得比较快、研究得比较热的技术之一。其强大的并行处理能力可以在维持大吞吐量的前提下提供稳定的极低处理延迟，非常适用于流式的计算密集型任务和通信密集型任务。

华泰证券正在积极开展 FPGA 应用研究，目标搭建一个高稳定低延迟的金融加速计算基础架构，形成友好易用的开发框架。用户无需关心底层实现，只需专注于开发自己的业务逻辑，更高效地实现客户的个性化业务需求。

本次研究计划对上证行情进行实时解析，以上证指数为基础指标，经过简单策略判断后将计算结果发送到下游系统，同时接收下游系统的反馈信息构建记录表，以这整条链路作为落地应用场景。通过搭建 Demo 环境完善架构设计并验证实际计算加速效果，展现 FPGA 技术在低延时交易领域的应用价值。

一、概述

1.1 低延时交易的挑战

近年来，随着资本市场的快速发展和算法交易技术（尤其是高频交易）在全世界范围内的应用，证券行业在交易低延时领域面临着巨大的技术挑战。高精度、低延时的交易能力是券商的核心竞争力之一。高速的行情数据获取和策略计算能力是各种高频交易算法的核心；

随着证券实时行情发布时延降低到了微秒单位，行业对行情服务和高频交易时延的要求也达到了微秒级。传统的基于通用软硬件建设的交易系统在并行计算和网络协议处理等环节存在明显的性能瓶颈，有必要引入更高速的交易处理解决方案。

1.2 FPGA 在低延时交易场景的应用优势

FPGA 具备低抖动与运行稳定的特点。CPU 模式下存在总线仲裁、内存抢占、进程 / 线程锁等机制，处理业务容易受行情巨幅波动、活跃用户上升、成交 / 委托笔数急剧放大等影响，大型数据库、网络吞吐性能下降严重，系统稳定性下降；

FPGA 模式下使用硬件门电路，所有资源都是预先分配布线，原理上与 CPU 模式有本质的不同，不存在调度问题、仲裁问题。FPGA 技术提供了强大的并行处理能力，可同时进行数据并行和流水线并行计算，在单个时钟周期内可以完成的运算远大于一个 CPU 指令的可完成的运算量，在处理计算密集型应用时效率提升明显；

在数据获取分析过程中，FPGA 板卡上提供的高速光模块可以直接接收处理数据包，同时通过硬件层面的 IP（Intellectual Property）核编程使得 FPGA 在数据传输延迟和网络数据解包能力上都大大优于传统的网络协议解析。

FPGA 高速、稳定、灵活、低成本的特点完美契合金融领域的加速计算需求，是一个理想的低延迟交易解决方案。目前该技术在交易实时风控、行情解析、金融指标计算等方面有较多应用，国内外不少公司已利用 FPGA 技术将交易处理延迟提升微秒级甚至纳秒级。长远来看自主研发和自主掌控 FPGA 相关产品和技术能力可以为券商带来明显的竞争优势。

二、技术和场景选型

2.1 OpenCL for FPGA 技术优势

目前 FPGA 技术在证券期货领域应用有如下困难：

1、开发难度大：传统的硬件开发编码对软件工程师不友好，难度大，需要专业技术背景，金融行业此类人才稀缺。FPGA 技术的开发需要极强的硬件背景，只有通讯、电子设备等少数公司拥有大批 RTL（硬件编程语言）工程师，而且每个合格 RTL 工程师的培养需要 5 年左右的时间。

2、开发周期长：硬件编码整个开发流程包括编写代码、编译、仿真、综合、时序分析、上

表 1

FPGA 特点	
低延迟、低抖动	FPGA 无抢占/调度/仲裁，低延迟，低抖动。
可编程、灵活	可编程与重配置；内嵌丰富的可编程器件；
高性能、低功耗	算力强；功耗小

板测试等过程，开发周期较长。迭代不灵活，跟不上金融行业的上线速度。

3、缺少生态圈：国内金融领域成熟落地的 FPGA 加速方案较少，缺少包括金融机构、FPGA 厂商以及解决方案提供商的完整生态圈。

以上三个难题导致应用 FPGA 进行业务开发的门槛高、成本高、周期长，出问题后定位错误困难。因此证券期货行业急需高性价比、低门槛的开发解决方案。

针对 FPGA 设计的 C-based OpenCL 开发环境允许开发人员用 C 语言描述完整的计算过程，无需学习底层 HDL 编码工作，即可编写内核程序发送至 FPGA。

OpenCL 是为满足异构计算的需求而出现的程序框架标准，可在异构平台（包括 CPU、GPU、FPGA 和其他类型的处理器）上执行，具有良好的开放性和可移植性。

OpenCL 解决方案适合没有硬件背景的 c/c++ 工程师。它解决了用 C 语言如何来描述并行计算结构的问题；它有一套通用的编程规范。程序员可以通过编写 kernel 程序和 host 程序实现功能，不需要去关注太多底层原理，直接按这个 OpenCL 模型与规范写代码就得到一个可以高效运行的系统；最开始接触它的 c 工程师不到 2 个月的时间可以进入开发。非常适合作为软件开发背景行业的 FPGA 开发框架选择。

不过，开发人员对于底层几乎没有什么干预能力，只能使用有限的优化手段和优化方式来影响最终结果，这就需要高效率 IP 库进行弥补；另外，OpenCL 的运行效率一般情况下会低于 RTL（10-20% 左右的性能损失），但是它的开发效率以天计算，显著高于以周计算的 RTL 方式。这就类似 CPU 模式下，汇编语言与 c/c++ 语言的区别。具体的取舍需要综合考虑业务场景特点、性价比、开发周期、人才储备等因素。

2.2 场景说明

本次研究环境在 Intel Arria 10 PAC 卡的基础上搭建，采用 OpenCL for FPGA 技术实现行情数据解析、指标计算、实时股票行情表构建以及计算结果低延时分发等业务逻辑功能，同时结合 Intel TOE IP 实现 TCP/UDP 层网络数据的接收和发送，整体网络处理架构如图 1 所示。

对于计算部分，我们采用实时上证指数指标作为基础指标，在此基础上根据一定策略判断，将计算结果发送到下游系统。

目前上交所每 5 秒推送一次指数行情，时间间隔很长。而股票成交数据是每 3 秒推送一次快照数据，未来会改成逐笔实时成交，本次场景通过获取股票逐笔成交数据来实现更实时的指数计算，使得客户能够更快地根据指数指标进行交易决策。

三、整体架构设计

3.1 环境说明

本场景使用了一台 x86 服务器安装一块 Intel

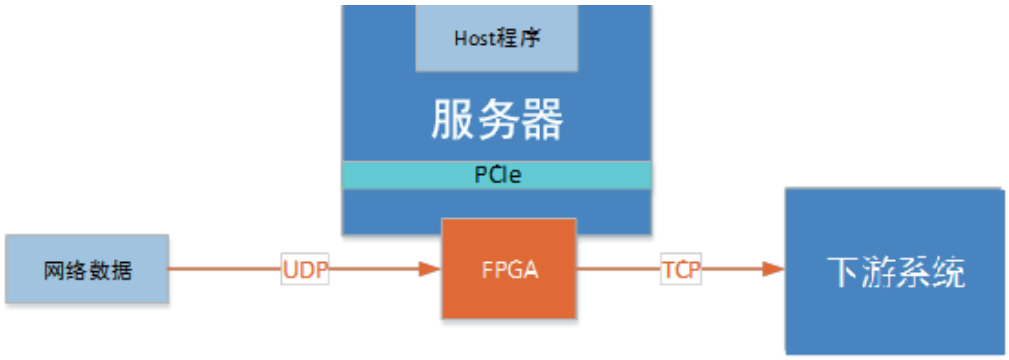


图 1

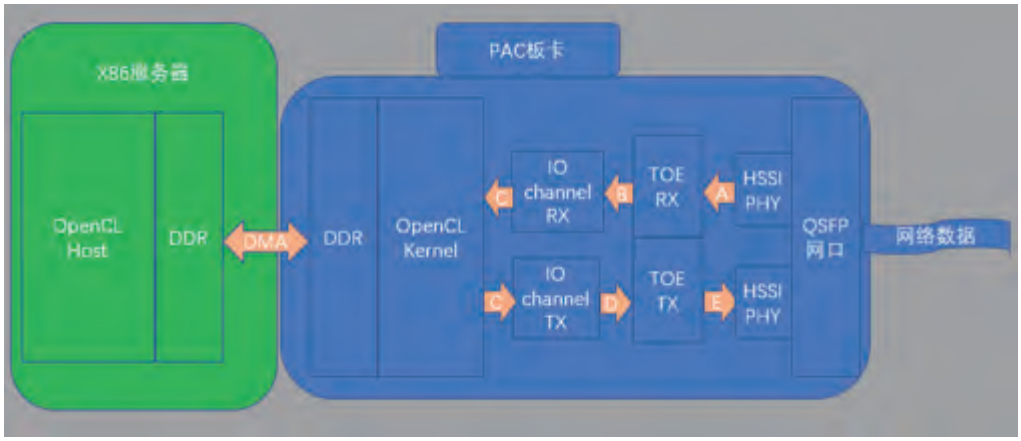


图 2

Arria 10 PAC 板卡。服务器上安装 OpenCL 开发环境、包含 TOE IP 的 OpenCL BSP（板卡驱动）。这台服务器上运行 host 程序，用于初始化 FPGA 环境并启动执行板卡上的 kernel 程序。

FPGA 板卡上启用两个 TOE IP，一个物理上对接行情数据网络传输，另一个对接交易柜台进行下单和接收反馈信息。

整个处理链路从行情接收到下单指令发出都通过网络，所有策略逻辑都在 FPGA 上进行，不经过 CPU。

3.2 网络架构

通过 Intel 的 TOE IP 核，可以实现纳秒级的 MAC 层到 TCP/UDP 层协议数据的卸载处理。通过 API 可以进行网络层面的控制并实现 TOE 状态数据监控和数据处理时钟周期监控。技术架构如图 2。

FPGA 的网口用来接收网络数据和发送处理结果。接收的数据经过 TOE IP 的处理，通过 IO channel 传递给 FPGA kernel 程序进行处理计算，计算结果通过 IO channel 发送给 TOE IP，通过网

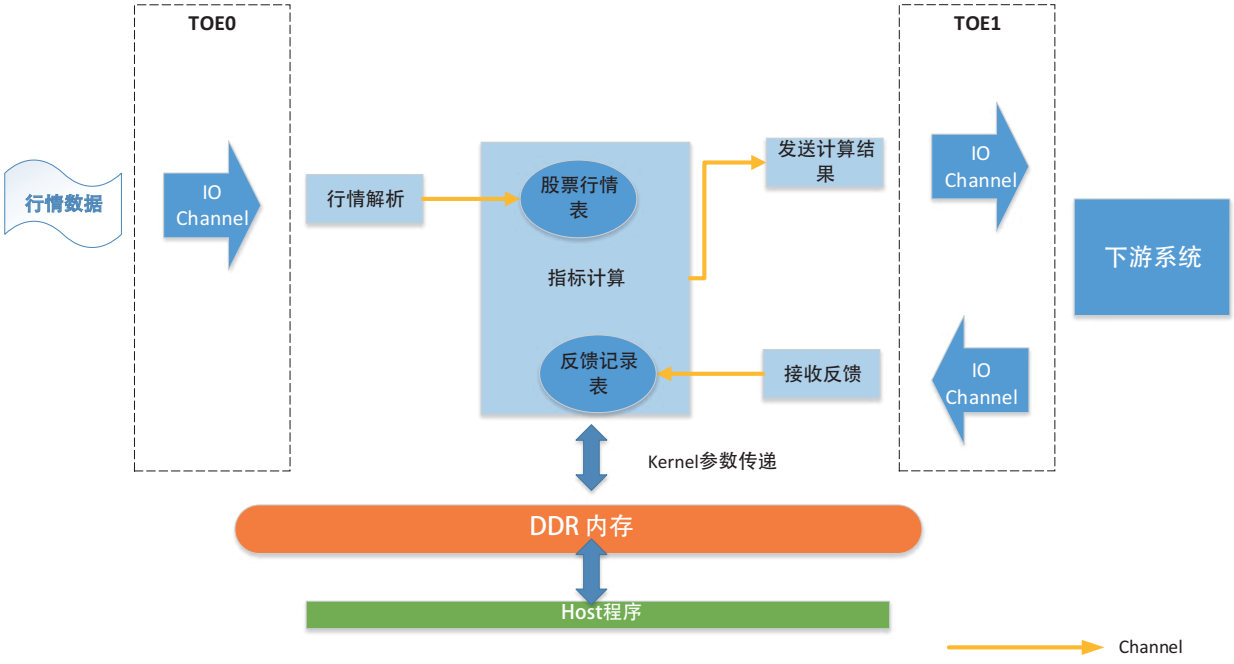


图 3

络将数据发送出去；

基于该技术架构，可以实现并设计网络流数据处理时延敏感类应用场景的通用框架。

3.3 kernel 架构

图 3 方案中，行情解析，实时股票行情表构建，上证指数计算，计算结果发送，反馈信息接收及反馈信息记录表构建等功能都在 FPGA 板卡上实现，分拆为四个 kernel 实现对应的功能。Host 端通过 DDR 内存与 kernel 交互，通过参数传递对策略进行个性化配置，同时通过 DDR 内存获取在 FPGA 本地内存中构建的实时股票行情表和反馈信息记录表，形成一份数据备份。

channel 是 FPGA kernel 程序之间传递数据的机制，是一种 FIFO buffer。启用 channel 可以保障并发运行的 kernel 间的数据传递不经过 host 程序和 DDR 内存，十分高效。通过 channel 可以方便地实现 1 对 N，N 对 1 的并发 kernel 处理架构，如图 4 所示。

具体模块和流程如下：

1、TOE0：用于接收行情数据的 TOE IP，通过 IO Channel 将接收到的数据传递给 FPGA kernel；

2、行情解析 kernel：解码行情数据，截取所需的实时股票成交信息，通过 channel 传递给策略下单 kernel；

3、指标计算 kernel：指标计算和计算结果发送的主体，本次场景中该 kernel 实现以下功能：

1）接收实时股票成交信息，在本地内存中构建实时股票行情表；

2）根据实时股票成交信息，计算实时上证指数；

3）接收下游系统对于计算结果的反馈信息，在本地内存中构建实时反馈信息记录表；

4）基于以上实时表和指数指标，构建简单策略逻辑（比如指数大于一定阀值、在某个时间窗口内结果发送次数、指数在时间窗口的涨幅等）触发计算结果发送，通过 channel 传递给发送计算结果 kernel；

在本地内存中保存的实时数据可以保证最快的读写速度。根据容量测算，全市场数据约在 8~10MB 左右，目前市面上中端或高端的 FPGA 卡都能满足这个需求；

对于策略逻辑实时更新的需求，我们可以将灵活更新的策略判断要素作为 kernel 参数（比如指数变化量、指数阀值等），通过 host 程序传递给指标计算 kernel。同时通过 OpenCL 的 hostpipe

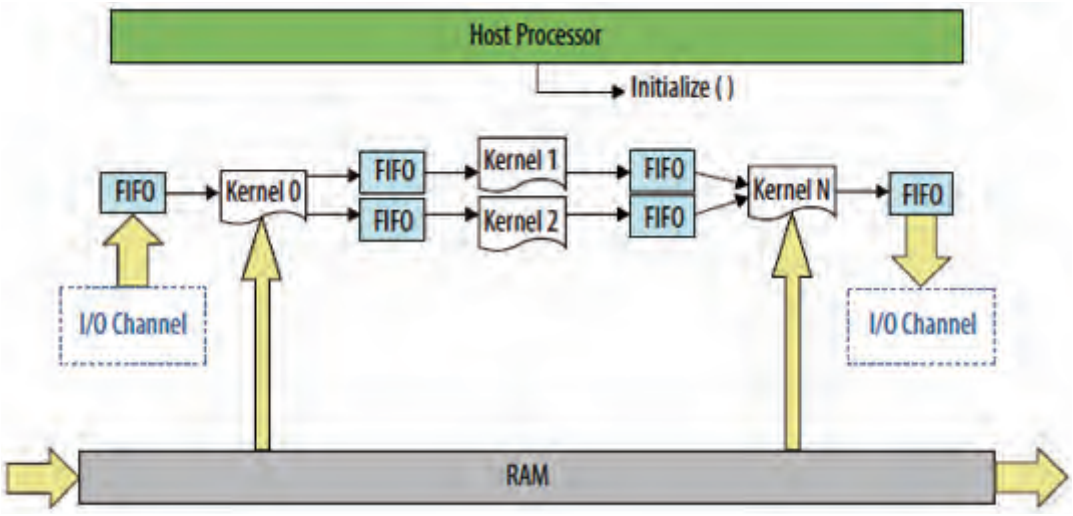


图 4：Channel 数据传输示意图（图中的 FIFO 就是 channel）

功能为 kernel 传入一个信号标志。hostpipe 类似于 host 和 kernel 之间通讯的 channel，通过 hostpipe 传递的数据不经过 DDR 内存，速度很快。kernel 程序每次循环时，都会尝试读取 hostpipe 传递过来的信号标志，如果读取到了，则重新进行参数读取等初始化动作，如果未读取到，则执行常规业务处理逻辑。

4、发送计算结果 kernel：接收计算结果，封装网络协议并发送到 TOE1；

5、接收反馈 kernel：接收下游系统的反馈信息，截取所需的信息，通过 channel 传递给指标计算 kernel；

6、TOE1：用于与交易柜台通讯的 TOE IP，通过 IO Channel 接收并发出计算结果。同时接收下游系统的反馈信息，传递给接收反馈 kernel。

四、测试评估

4.1 性能测试结果

本次研究场景中也编写了模拟 C 程序进行行情解析、指数计算和计算结果分发，使用传统千兆网络进行数据传输，通过比对 FPGA 和 C 程序处理的速度，考察 FPGA 加速计算的效果。

选取单条行情处理延时作为测试指标，计算 FPGA 收到每一条行情数据，到计算完成通过网络发送出去的时间周期。这个时延包括了 TOE 网络处理的时延。

对比发现，FPGA 处理延时平均在 3 微秒内，C 模拟程序会达到几十微秒，有 10~30 倍的差距。同时值得注意的是，FPGA 处理延时稳定在一个非常小的区间，几乎没有波动，而 C 模拟程序的处理延时波动非常大，这充分体现了 FPGA 低延时、低抖动的特性。

延时测量对比如表 2。

4.2 开发模式评估

OpenCL for FPGA 的开发涉及 host 程序和 kernel 程序的编写。这个开发模式的优缺点总结如下：

优点：

host 和 kernel 程序都是基于 C 语言语法来编写。Kernel 程序的语法、变量类型、数据结构都相对简单，没有指针、引用、面向对象以及其他一些高级应用的概念；host 程序与 FPGA 板卡的交互也有固定的步骤和框架，API 使用简单。因此整个开发过程对于软件工程师来说上手较快；

提供模拟编译的功能，可以快速编译 kernel（1 分钟内），使用 CPU 来模拟运行，这样可以快速检查程序内的逻辑错误；

缺点：

实际编译时间较长，随着 kernel 程序的复杂度可能从 2 小时到 5 小时不等。并且即使模拟编译通过了，在线路综合和布局布线阶段仍然会遇到各种错误。而这个阶段的错误需要专业硬件工程师协助才可排查，并且耗时耗力，对于软件工程师来说很难排查处理；

Kernel 程序虽然是用 C 语言来编码，其中的底层逻辑和软件架构是不同的，尤其是一些速度和资源优化的思路，并非传统软件架构的思路。因此软件工程师虽然能快速上手开发程

表 2

	平均延时(us)	最高延时(us)	标准差
FPGA	1.7	1.7	3.68
软件程序	14	53	5082

序，但是要编写出高效的程序需要一个较长的学习理解过程。

五、后续架构改进思路

对于整个 FPGA 极速下单策略架构，如果要在生产环境使用，仍会有一些重要的考量因素，这也是本次研究场景架构后续改进的方向：

1、故障恢复：由于 FPGA 板卡的单点性，以及板卡上内存数据非持久化的特性，建议维护一份软件程序接收行情并策略下单的路径，在 FPGA 板卡或整个链路故障时快速切换；

2、网络协议优化：对于行情接收端，可以采用 UDP 协议，FPGA 端被动接收即可。在这种场景下，甚至不需要 TOE IP，FPGA 在 MAC 层接收网络数据后直接去掉 UDP 协议头即可，进一步缩短网络处理延时；

3、策略复杂度：FPGA 更擅长处理数据并行计算以及简单的触发策略，比如本次研究场景中的指数指标的计算，而不擅长处理过多的控制语句。同时复杂且灵活的控制语句也无法完全通过 kernel 参数传递来自定义，FPGA kernel 程序的开发、编译、测试又是一个以天计算的时间周期，无法灵活变更。

因此在实际生产应用中，建议对策略进行分离，一些基础指标在 FPGA 上进行计算，计算结果通过 PCIe 接口传递给 host 程序，由 host 程序实现复杂的策略判断，结果发送指令再通过 PCIe 接口回传给 FPGA kernel，发出计算结果。该模式能充分利用 FPGA 和 CPU 的不同特性，由 FPGA 处理固定的网络处理、网络协议封装以及行情解析和指标计算动作，由 CPU 处理灵活复杂的策略控制。

六、总结

本文研究了 OpenCL for FPGA 技术在极速交易场景的应用示范。设计并落地了从行情接收到指令下单的整体架构，并与软件模拟程序进行性能比对，论证了 OpenCL for FPGA 技术在高时延敏感场景的落地可行性。本次研究也形成了以下研究结论：

1、OpenCL for FPGA 具有易上手、开发难度低的特点，并且可以通过 host 和 kernel 程序的异构架构，通过 kernel 参数传递有效解决一部分程序变更灵活性的问题。但是本身编译时间较长，编译问题可排查性较差，对开发测试仍有较多不便。本项目也研究落地了基于网络流数据处理场景的通用开发架构，通过该架构能有效降低遇到底层编译问题的概率，解决了并发计算处理、kernel 灵活刷新、kernel 与 host 数据交互等方面的问题，大大提升了 OpenCL for FPGA 开发效率；

2、基于 TOE IP 核，FPGA 在基础架构层面可以作为一个节点直接接入网络，实现诸如网络数据接收传送、多 session 并发、扩展性、高可用性等生产级基础架构的特性，使得 FPGA 技术具备解决网络时延敏感场景的基本能力，具备接入生产环境的基础条件。

3、FPGA 相比传统软件程序在性能和稳定性方面都有了量级的提升，在对时延敏感的场景有非常高的应用价值。

综上所述，FPGA 技术具备在证券行业高时延敏感以及流数据并发计算等场景有较高的应用可行性和应用价值。后续我们也将持续研究 FPGA 技术在其他场景如数据压缩、加密解密、数据批量处理的应用，不断拓展 FPGA 技术在金融计算加速领域的应用。

七、鸣谢

感谢上交所技术有限责任公司的老师对本次研究的大力支持，包括 FPGA 硬件厂商资源的联络协调以及相关技术问题的意见和建议。

I 信息资讯采撷 nformation

监管科技全球追踪

监管科技全球追踪

国际动态

FSB 发布关于网络事故恢复与响应的咨询报告

金融稳定理事会（FSB）于 4 月 20 日发布了一份关于网络事故恢复与响应的咨询报告，旨在为金融机构提供网络事故事前、事中和事后的工具包（toolkit），以遵循 20 国集团（G20）最初在 2018 年提出的要求。该报告指出，网络风险可能来自许多方面，包括多个金融机构之间或金融机构与第三方服务提供商之间相互连接的信息技术系统，主要金融机构或金融机构集团的信息泄露，或事故造成的影响。 该工具包由七个部分组成，包括了 46 种有效做法，同时考虑了司法管辖区的立法，司法和监管框架，受网络事故影响的组织规模以及受影响的组织类型，其中也包括了和大流行病（如新冠肺炎）有关的注意事项。

G20 与 BIS 联合发起 TechSprint 监管科技竞赛

4 月 29 日，G20 轮值主席国沙特与国际清算银行（BIS）创新中心联合启动了 G20 TechSprint 监管科技竞赛，希望通过新技术解决新时期合规和监管领域的种种挑战。新加坡金融管理局（MAS）、金融稳定委员会（FSB）、API Exchange（APIX）和 RegTech for Regulators Accelerator（R2A）也对本次竞赛提供支持。本次竞赛将重点关注以下三个方面：1）针对危机的监管者动态信息共享；2）监控与监视洗钱与恐怖主义融资风险；3）监管与合规报告报送。最终获胜者将于 2020 年 10 月选出。

全球各交易所和清算所积极采取行动应对新冠肺炎疫情影响

随着新冠肺炎疫情（COVID-19）被世卫组织认定为大流行病（pandemic），各国的交易所和清结算所也纷纷采取行动，以应对疫情对日常工作的影响并保持资本市场平稳有序运行。世界交易所联合会（WFE）作为交易所和清结算所的国际性组织，归纳整理了全球八十多家会员和非会员交易所、清结算所面对疫情的应对方案和行动，具体信息可通过 WFE 网站查阅。

Technavio 发布保险科技市场（2019-2023）专题报告

4 月 8 日，Technavio 针对保险科技市场发展前景发布了专题报告。Technavio 指出，2019 年至 2023 年保险科技市场规模将增长 156.3 亿美元。目前的保险科技市场还处于分裂状态，主要的市场参与者包括 Oscar、Quantemplate、Shift Technology 和众安保险。随着市场不断发展，这种支离状态会不断降低。未来，保险科技服务提供商应该更多关注快速增长领域的发展前景，同时保持在原有缓慢增长领域的优势和地位。毫无疑问，数字化将对推动保险科技市场发展起到关键性作用。此外，本次报告还对新冠疫情对保险科技行业的影响分别进行了积极、消极和可能三种情况假设。

欧美动态

欧盟就未来数字金融与金融科技发展发起意见征询

2020 年 4 月，欧盟委员会就数字金融和金融科技产业的未来发展展开了一项意见征询，时间为 2020 年 4 月 3 日至 6 月 26 日。欧盟委员会表示，将根据征询结果在 2020 年第三季度颁布一份全新的数字金融战略提议，充分考虑期间的各类技术进步，为未来五年的金融产业创新奠定政策基础 。

欧盟发起黑客马拉松，鼓励通过科技创新应对新冠疫情

2020 年 4 月 24 日至 26 日，欧盟委员会将举行一次泛欧黑客马拉松活动。据悉，本次活动共吸引了 1.2 万名参赛者，竞赛范围共覆盖 7 个领域（包括金融科技），希望借此加快开发和应用创新方案、应对新冠疫情造成的影响。胜出项目将于 4 月 27 日加入欧洲创新委员会 COVID 平台，投入终端用户（如医院）的实际使用之中。此前，全球各地也相继举行了类似主题的黑客马拉松活动。例如，全球黑客倡议（Global Hack initiative）以及麻省理工学院于 4 月发起的“MIT 新冠肺炎挑战赛：打败大流行病”。

欧洲证券和市场管理局就云外包指南征求意见

欧洲证券和市场管理局于 6 月 3 日发布了一份征求意见文件，内容为关于云服务供应商的外包指南。该指南为金融机构外包给云服务提供商时的外包要求给出指导，帮助企业和主管部门识别、应对和监控云外包活动所带来的风险和挑战，尤其是在数据保护和信息安全方面。ESMA 主席 Steven Maijoor 认为，金融市场参与者应该小心，不应过度依赖云服务供应商，需密切监控云服务供应商的性能和安全措施，并确保能够在必要时取消云外包。ESMA 预计于 2021 年一季度发布该指南的最终版。

欧洲金融科技联盟正式成立

2020 年 6 月 16 日，多家欧洲金融科技企业联合宣布成立“欧洲金融科技联盟”（European FinTech Association，简称 EFA）。这家非盈利组织总部位于布鲁塞尔，集合了曾经的 EFAlliance 和 Fintechs4Europe 等机构的成员，希望为欧洲金融科技企业提供一个统一的信息共享和对话平台。EFA 的成员覆盖支付、贷款、银行、智能投顾、身份验证、软件即服务等多个金融服务领域。

英国金融行为管理局发布 2020-2021 年度工作重点事项

4 月 7 日，英国金融行为管理局（FCA）发布了 2020-2021 年度工作规划（Business Plan 2020/21）。其中，FCA 将应对新冠疫情作为当前的头等大事，希望通过自身监管作用帮助社会各界获得所需金融支持、避免金融诈骗、保持金融市场稳定。未来，FCA 将重点关注 4 个领域，确保消费者获得安全便捷的薪金 / 收益支付、不陷入过度债务之中、对其储蓄进行有效且风险可控的投资、获得公平合理的数字金融服务及产品。此外，FCA 还计划转变运营模式，涉及数据收集、数据分析、信息管理和共享等多个方面。为了完成上述任务，FCA 可能会在今年加大对技术、系统和相关人员的投入。

苏格兰皇家银行宣布关停旗下数字银行平台 Bó

5 月 1 日，苏格兰皇家银行（RBS）发布了 2020 年第一季度报告。报告指出，将逐步关停消费者数字银行平台 Bó，并将其技术融合到中小企业数字银行 Mettle 之中。Bó 于 2019 年 11 月 29 日正式推出，从面世到关停历时不到半年。起初，Bó 的目标是通过数字化技术平台，帮助年轻一代用户进行实时、合理的财务管理。但是在过去的 5 个多月里，Bó 的发展并不顺利，用户总量仅有 1.1 万人，并在今年 2 月因为用户监督认证问题而不得不对 6000 张 1 月 3 日之前发行的银行卡进行了更换。

普华永道携手 Credit Kudos 打造数字银行沙盒平台

总部位于英国的普华永道正在整合多个云端平台，希望创建一个由各类有所专长的金融科技企业组成的全新数字银行，以更快更多地满足用户需求，同时降低 IT 等运营成本。5 月 21 日，普华永道宣布 Credit Kudos 成为这一数字银行生态系统沙盒平台的入选机构之一。Credit Kudos 提供了一种使用公开银行数据而非历史借贷信息来衡量信用度的替代方法，为贷方提供更加精准的风险评估标准，帮助其更快地做出决策，而某些此前被传统金融服务忽视的借款人也可能因此而获得更快的贷款批复。

葡萄牙政府计划成立“科技自由区”，全面推动技术创新与试验

4 月 21 日，葡萄牙政府发布决议，计划出台设立和监管“科技自由区”的总则条文，全面推动新兴技术创新与试验。据悉，这个自由区也是葡萄牙政府“数字转型行动规划”的一部分。葡萄牙政府表示，本次自由区计划将面向金融科技、5G 网络、大数据、物联网、区块链、生物和纳米技术、人工智能等多个方面，通过全面的监管沙盒框架推动葡萄牙的可持续经济转型。

美国 MEMX 证券交易所获批成立

2020 年 5 月 5 日，由摩根士丹利、美银美林、瑞银 UBS 等九家华尔街主流券商和银行联合成立的 Members Exchange（MEMX）证券交易所获得美国证券交易委员会（SEC）批准，并计划在 2020 年第三季度正式开始营业。该交易所总部将设于纽约，与纽约证券交易所、纳斯达克等现有交易所竞争。

纳斯达克推出云数据服务

近日，纳斯达克宣布推出一项云数据服务，允许用户使用一套 API 访问实时交易数据、市场指数和资金数据。纳斯达克表示，这组 API 使用开源交付标准和 SDK 来快速跟踪工程工作，从而消除了对硬件采购、专有协议、文件格式和线路租用的需求。这样就可以轻松集成来自不同来源的数据，大大缩短客户设计应用程序的上市时间，非常适合企业家、金融科技公司和传统应用程序提供方。据悉，这项服务目前将在 AWS 上运行，但仍然可以与多个云供应商兼容。

美国金融业监管局发布《证券业中的人工智能》白皮书

6 月 10 日，美国金融业监管局（FINRA）发布了《证券业中的人工智能》白皮书。在过去的两年中，FINRA 的金融创新办公室就人工智能对证券行业的影响进行了深入研究。FINRA

发现，基于人工智能的应用程序正在证券行业中激增。该本白皮书指出了企业和政策制定者在考虑人工智能在证券行业的应用时应考虑的五个关键要素：1、人工智能不是未来才会出现，而是已经影响着经纪自营商和金融服务；2、人工智能已经成为一个包罗万象的术语，涵盖各种不同技术和应用程序；3、经纪自营商主要在与客户沟通、投资流程和运营中使用人工智能；4、企业在部署人工智能应用前需考虑人工智能的独特挑战，例如可解释性、数据偏差、客户隐私等；5、如果监管得当，人工智能可能为投资者和企业带来巨大利益。

美国证券交易委员会将举行系列在线讨论会，首期主题为监管科技

6 月 11 日，美国证券交易委员会（SEC）创新与金融科技战略中心（FinHub）宣布，将在未来一段时间举行一系列在线讨论会。2019 年，FinHub 曾举办了多场面对面的讨论会，为业界与监管机构沟通交流提供了特定渠道。鉴于目前的疫情发展情况，FinHub 决定将讨论会转为线上。据悉，首个在线讨论会将于 2020 年 7 月 6 日这一周举行，主题为“监管科技”。

在线券商罗宾汉平台再现技术故障

6 月 18 日，在线券商平台罗宾汉（Robinhood）的用户再度遭遇了账户无法登陆的问题，这也是继今年 3 月该交易系统因交易量过大两次宕机以来的又一次“事故”。对此 Robinhood 给出的解释是，“第三方服务导致的交易和转账功能故障”，目前该问题已经得到解决。值得注意的是，这家以免佣金为主要卖点的在线券商平台今年的交易一直非常活跃，并吸引了不少年轻一代的投资者，新增账户超过 300 万个，其中有约一半为首次投资。

亚太动态

Global Ventures 发布西亚与北非地区金融科技报告

近日，总部位于迪拜的风险投资公司 Global Ventures 发布了一份西亚与北非地区金融科技专题报告。在这篇报告中，Global Ventures 详细介绍了该地区金融科技不同领域的发展情况。近几年来，西亚北非地区国家一直在探索经济多样化道路，减少对自然资源的依赖性，这在当前石油价格跌至历史地位的情况下显得尤其重要。为了推动这项转型，该地区各国纷纷推出了监管创新措施，创建经济自由区、启动金融技术加速器和孵化器项目以及设立监管沙盒机制。

阿联酋颁布迪拜国际金融中心数据保护法

6 月 1 日，阿联酋发布了迪拜国际金融中心（DIFC）第五号数据保护法，新法律将于 7 月 1 日生效。新版数据保护法结合了当前各种世界级保护法的最佳做法，如《通用数据保护条例》（GDPR）、《加州消费者隐私发》等，将支持 DIFC 争取获得欧盟委员会、英国和其他司法管辖区的充分认可，从而减轻 DIFC 的业务对数据传输的合规性要求。此外，DIFC 管理局董事会还发布了新的《数据保护条例》，其中规定了在数据保护、问责、记录保存、罚款和跨境转移个人数据方面，

向适当司法辖区的负责人发出通知的流程。

韩国央行宣布开始进行央行数字货币测试

此前，韩国银行（即韩国的央行）一直对央行数字货币保持谨慎观望态度，并认为盲目发行数字韩元可能会影响央行权威和金融系统稳定。然而就在 4 月 6 日，韩国央行却一改以往态度，宣布将采取类似日本和美国的央行数字货币策略。为此，韩国银行将开始展开为期 22 个月的内部测试，从 2020 年 3 月持续到 2021 年底，检测数字货币替代纸币的可能性以及可能遇到的技术问题。

韩国开放公共金融大数据，总量可达千万级

韩国金融服务委员会（FSC）宣布，从 2020 年 6 月 9 日起，将通过 www.data.go.kr 上的开放 API 提供约 4450 万例公共金融数据。其中，有约 58 万例非上市公司的数据是首次免费向公众开放。向公众开放的金融大数据是存储在 FSC 和其他 9 家韩国公共金融机构中大量公共金融数据的链接、融合和标准化的副产品。FSC 表示，随着公共金融大数据的推出，有望在该领域创造更多的工作机会以及更多以数据为驱动的创新服务，促进韩国数据经济和金融科技行业的发展。未来，FSC 还将努力提供更多类型的金融数据，并增加开放金融大数据系统参与机构的数量。

基于区块链技术的广东省中小企业融资平台正式上线发布

近日，以区块链作为底层技术的广东省中小企业融资平台正式上线发布。据了解，区块链技术的应用，使得中小融平台建立了信息共享、隐私保护和互信机制，不仅让企业信息真实不可篡改，也使金融机构得以依法依规实现企业信息查询，此前在供应链金融领域广受诟病的应收账款重复质押融资的现象在区块链技术下也将无所遁形。中国平安旗下的金融壹账通是中小融平台的技术支持方。目前，中小融平台已接入来自 26 个政府部门的 213 类政府数据，对全省 1100 多万家企业信息全面采集，进行企业风险评级和画像，另外对接了工商银行、建设银行、平安银行等省内 129 家金融机构，上线 319 款金融产品。

北京金融科技研究院发布《2020 中国保险科技洞察报告》

4 月，北京金融科技研究院与清华大学互联网产业研究院、北京金融科技与专业服务创新示范区、爱保科技、毕马威、英诺天使基金、中科院资本等科研机构发布了《2020 中国保险科技洞察报告》。该报告作为北京金融科技研究院成立后发布的第一本研究报告，共分保险行业新态势、保险科技生态透视、保险科技行业洞察、保险科技行业展望四个章节。

上海金融科技产业联盟正式成立

5 月 17 日，上海金融科技产业联盟正式成立。目前，联盟形成了“8+44”结构，即：由人民银行上海总部、上海银保监局、上海证监局、市金融局等 8 家管理部门联合出任指导单位，由金融要素市场、金融机构、新金融和金科子公司、科技企业、高校及科研单位、功能性机构等 44 家单位共同构成联合发起单位。联盟定位为联合金融科技相关机构的跨行业、开放性、非营利性的非法人组织，旨在服务上海市及长三角地区金融科技产业的高质量发展，搭建具有国际视野的技术合作与产业促进平台，主要职能包括政策研究、产学研投用合作、组织联盟活动、推动创新应用、提供智力支持、培养专业人才、开展国际交流等。

2020年三季度《交易技术前沿》征稿启事

《交易技术前沿》由上海证券交易所主管，上交所技术公司主办，以季度为单位发刊，主要面向全国证券、期货等相关金融行业的信息技术管理、开发、运维以及科研人员。

2020 年三季度征稿主题如下：

一、云计算

（一）云计算架构

主要包含但不限于：云架构剖析探索，云平台建设经验分享，云计算性能优化研究。

（二）云计算应用

主要包含但不限于：云行业格局与市场发展趋势分析，国内外云应用热点探析，金融行业云应用场景与实践案例。

（三）云计算安全

主要包含但不限于：云系统下的用户隐私、数据安全探索，云安全防护规划、云安全实践，云标准的建设、思考与研究。

二、人工智能

（一）应用技术研究

主要包含但不限于：语音识别与自然语言处理，计算机视觉与生物特征识别，机器学习与神经网络，知识图谱，服务机器人技术。

（二）应用场景研究

主要包含但不限于：智能客服、语音数据挖掘、柜员业务辅助等。

主要包含但不限于：监控预警、员工违规监控、交易安全等。

主要包含但不限于：金融预测、反欺诈、授信、辅助决策、金融产品定价、智能投资顾问等。

主要包含但不限于：金融知识库、风险控制等。

主要包含但不限于：机房巡检机器人、金融网点服务机器人等。

三、数据中心

（一）数据中心的迁移

主要包含但不限于：展示数据中心的接入模式和网络规划方案；评估数据中心技术合规性认证的必要性；分析数据中心迁移过程中的影响和业务连续性；探讨数据中心迁移的实施策略和步骤。

（二）数据中心的运营

主要包含但不限于：注重服务，实行垂直拓展模式；注重客户流量，实行水平整合模式；探寻数据中心运营过程中降低成本和提高服务质量的途径。

四、分布式账本技术（DLT）

（一）主流分布式账本技术的对比

主要包含但不限于：技术架构、数据架构、应用架构和业务架构等。

（二）技术实现方式

主要包含但不限于：云计算 + 分布式账本技术、大数据 + 分布式账本技术、人工智能 + 分布式账本技术、物联网 + 分布式账本技术等。

（三）应用场景和案例

主要包含但不限于：结算区块链、信用证区块链、票据区块链等。

（四）安全要求和性能提升

主要探索国密码算法在分布式账本中的应用，以及定制化的硬件对分布式账本技术性能提升的作用等。

五、信息安全与 IT 治理

（一）网络安全

主要包括但不限于：网络边界安全的防护、APT 攻击的检测防护、云安全生态的构建、云平台的架构及网络安全管理等。

（二）移动安全

主要包括但不限于：移动安全管理、移动互联网接入的安全风险、防护措施等。

（三）数据安全

主要包括但不限于：数据的分类分级建议、敏感数据的管控、数据共享的风险把控、数据访问授权的思考等。

（四）IT 治理与风险管理

主要包括但不限于：安全技术联动机制、自主的风险管理体系、贯穿开发全生命周期的安全管控、安全审计的流程优化等。

六、交易与结算相关

（一）交易和结算机制

主要包含但不限于：交易公平机制、交易撮合机制、量化交易、高频交易、高效结算、国外典型交易机制等。

（二）交易和结算系统

主要包含但不限于：撮合交易算法、内存撮合、双活系统、内存状态机、系统架构、基于新技术的结算系统等。

投稿说明

1、本刊采用电子投稿方式，投稿采用 word 文件格式（格式详见附件），请通过投稿信箱 ftt.editor@sse.com.cn 进行投稿，收到稿件后我们将邮箱回复确认函。

2、稿件字数以 4000-6000 字左右为宜，务求论点明确、数据可靠、图表标注清晰。

3、本期投稿截止日期：2020 年 9 月 30 日。

4、投稿联系方式 021-68828590, 021-68813289 欢迎金融行业的监管人员、科研人员及技术工作者投稿。稿件一经录用发表，将酌致稿酬。

《交易技术前沿》编辑部

证券信息技术研究发展中心（上海）

附件：投稿格式

标题：（黑体 二号 加粗）
作者信息：（姓名、工作单位、邮箱）（仿宋 GB2312 小四）
摘要：（仿宋 GB2312 小三 加粗）

一、概述（仿宋 GB2312 小三 加粗）
二、一级标题（仿宋 GB2312 小三 加粗）
（一）二级标题（仿宋 GB2312 四号 加粗）
1、三级标题（仿宋 GB2312 小四 加粗）
（1）四级标题（仿宋 GB2312 小四）
正文内容（仿宋 GB2312 小四）
图：（标注图 X. 仿宋 GB2312 小四）
正文内容（仿宋 GB2312 小四）
表：（标注表 X. 仿宋 GB2312 小四）
正文内容（仿宋 GB2312 小四）
三、结论 / 总结（仿宋 GB2312 小三 加粗）

杂志订阅与反馈

各位读者，如您想订阅《交易技术前沿》纸质版，欢迎扫描右侧二维码填写问卷进行订阅，同时可以向我们提出关于《交易技术前沿》的建议与意见反馈。如您希望赏阅电子版，欢迎访问我们的电子平台 http://www.sse.com.cn/services/tradingservice/tradingtech/sh/transaction/（或扫描封面尾页二维码）。我们的电子平台不仅同步更新当期的文章，同时还提供往期所有历史发表文章的浏览与查阅，欢迎关注！

扫描在线浏览

联系电话：021-68828590
021-68813289
投稿邮箱：ftt.editor@see.com.cn

ITRDC
证券信息技术研究发展中心（上海）
上海证券交易所 SHANGHAI STOCK EXCHANGE
TECH 上交所技术

中国上海浦东南路528号
邮编：200120
公众咨询服务热线：4008888400
网址：http://www.sse.com.cn