

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 第三期 总第40期

本期主题

运维管理

No.3



上海证券交易所
SHANGHAI STOCK EXCHANGE

ITRDC
证券信息技术研发中心 (上海)

内部资料 2020 年第三期（总第 40 期）
准印证号：沪（K）0671

NO.3

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
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



篇首语

在证券期货行业，系统、机房、基础设施的运维管理十分重要，承担了保障市场安全运行的责任。证券行业系统种类多、实时性要求高，从开市前、开市期间到闭市后都要保持紧张的监测状态，随着近期全球证券领域安全运行事件多发，运维效率和应急响应能力得到了越来越多的关注。本期《交易技术前沿》以“运维管理”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，《中金财富证券在智能化异常根因定位方面的探索与实践》通过聚类与贝叶斯方法实现智能化的根因定位，以期提高异常根因定位的效率，解决多对多的机器架构和告警数量爆发式增长带来的两大运维挑战。《网络自动化运维系统自主研发的探索与实践》为满足防火墙策略手工配置重复性高且繁琐的问题，提出防火墙策略自动化配置的解决方案，分享了自主研发的实践经验。《加速行情实现，提升金融交易能效——基于 FPGA 的超低延时行情系统》结合国内外金融环境与技术，介绍了基于 FPGA 研发的超低延时行情系统，包括 FPGA 硬件设计与实现、延时性能等。《构建在知识中台基础上的企业画像》分析了企业画像面临的数据、业务和技术问题，针对数据治理、知识转化与沉淀、算法模型天花板带来的挑战给出了基于知识图谱技术的三种思路和架构。《华泰证券财经快讯自动生成发布实践》基于自然语言处理技术设计了一套包含公告文件解析、财务信息抽取和快讯生成的财经快讯自动生成解决方案，并应用到了自研的金融舆情事件服务系统中。《实时计算系统建设经验分享》利用实时计算技术搭建实时计算系统的过程和经验，使客户在集中交易、两融、OTC、期权、黄金五大系统内的数据资产同步变化。

运维管理经历了手工、脚本到自动化运维等阶段，近年来 Devops 和智能运维也成为新的焦点。随着技术的快速迭代，我们拥有了更丰富的工具，但是立足深入的业务理解和运维管理中的故障分析与决策，仍然是做好安全运行的圭臬。

《交易技术前沿》编辑部

2020 年 9 月 30 日

目录 Contents

本期热点 Hotspot

- | | |
|--------------------------------------|----|
| 1 中金财富证券在智能化异常根因定位方面的探索与实践 / 孟庆江、田忠毅 | 4 |
| 2 网络自动化运维系统自主研发的探索与实践 / 梁德汉、唐新华、武孟军 | 11 |
| 3 关于 IDC 运管平台建设的思考 / 林杭宁、黄勇刚 | 21 |

实践探索 Exploration

- | | |
|--|----|
| 4 加速行情实现，提升金融交易能效——基于 FPGA 的超低延时行情系统 / 赵月、张强、连成浩、赵健榆 | 30 |
| 5 构建在知识中台基础上的企业画像 / 齐伟 | 37 |
| 6 财经快讯自动生成技术探索和实践 / 邱震宇、张弛、刘园、孔亚洲、朱德伟、肖驰 | 42 |
| 7 基于大数据的智能化客户服务平台 / 应原、任荣、王东、王洪涛 | 49 |
| 8 证通云外呼系统智能化实践 / 王晴、何浩 | 56 |

行业观察 Observation

- | | |
|--|----|
| 9 实时计算系统建设经验分享 / 肖钢、李宁、张海园 | 65 |
| 10 数据中心智慧园区思考 / 黄飞、蓝汉明 | 70 |
| 11 海通证券混合金融云平台生态体系建设之云原生迁移服务 / 罗秋清、陆颂华、王朝阳、张真真、崔恒春 | 77 |

信息资讯采撷 Information

- | | |
|----------|----|
| 监管科技全球追踪 | 84 |
|----------|----|

Return on Investment

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus quis erat quam. Maecenas ac nibh sagittis, convallis augue accumsan, laoreet orci.

Ut eget ipsum libero. Duis eleifend tellus nec augue auctor placerat. Donec rutrum laoreet venenatis. Etiam sapien neque, ultricies vitae erat eu, consectetur ornare nunc. Vestibulum mollis neque sit amet fermentum vehicula. Pellentesque semper condimentum odio, quis fermentum lectus venenatis sit amet.

Performance Results by Year

Periods	Profit (M)
2010	11
2011	15
2012	9
2013	14
2014	
2015	
2016	
2017	
2018	

Region



PROFIT



H 本期热点 otspot

- 1 中金财富证券在智能化异常根因定位方面的探索与实践
- 2 网络自动化运维系统自主研发的探索与实践
- 3 关于 IDC 运管平台建设的思考

中金财富证券在智能化异常根因定位方面的探索与实践

孟庆江、田忠毅 / 中金财富证券股份有限公司信息技术部



证券行业是对连续性、稳定性要求最高的行业之一，客户交易体验永远是券商关注的话题。当交易系统出现异常时，快速定位异常根因并实施恢复可减少异常对客户体验造成的影响。但传统的根因定位方案目前存在一些困难：一是目前券商交易系统在各个环节的衔接广泛采取了多对多的机器架构，传统的链路定位在这种情况下效果并不理想；二是监控技术的发展使得告警数量呈爆发式的增长，导致告警与异常间的关联关系变得难以确定，增加了异常根因定位的难度，而智能化运维的出现，为我们解决这些问题提供了新的思路与方法。本文首先概述了中金财富证券近年来在智能化运维方面所作的工作，而后重点介绍了为解决以上两个运维痛点所作的探索与实践：通过聚类与贝叶斯方法实现智能化的根因定位，以期提高异常根因定位的效率。

一、引言

（一）智能化运维的产生与发展

在金融科技的不断发展、金融产品的不断增

加及金融监管的愈发加强等因素的影响下，证券公司的 IT 架构也逐渐复杂化，其表现在机器设备不断增加、业务链条不断增长、系统间的相互关联性逐渐上升，这给券商运维团队提出了新的



图 1.1：中金财富智能运维价值闭环

难题。受限于人力、经验等客观条件的制约，使用传统的、基于规则的运维方式已无法满足券商业务的连续无中断性要求，因此，券商运行维护系统需要紧紧跟上证券公司业务创新发展的潮流，脱离依赖人工和经验为主的传统模式，运用新的技术解决新的问题。

随着大数据、人工智能等先进科技不断给金融行业赋能，智能化运维逐渐成为证券行业运维系统的发展方向。智能化运维顾名思义，是将运维数据和大数据、机器学习技术相结合，针对运维场景开发成一系列的智能策略，融入到运维系统中。智能化运维赋予了 IT 系统从历史中自动获取经验知识的能力，同时借助大数据平台的强大算力实现了横向的拓展，有效的解决了运维团队对于经验和人力的依赖。自 2015 年开始中金财富逐步开始建立基于大数据、人工智能为技术基础，以事前预测、事中检测、事后分析为应用场景的智能化运维体系。该体系使用后，不仅可以有效的解决由于业务发展所带来的运维问题，提升了系统安全性，同时也帮助运维团队摆脱了以往简单重复性的工作，可以将更多精力投入到提升服务质量的需求开发中，增加了运维团队的工作价值。如图 1.1 所示，智能化运维促使券商 IT 系统形成了以业务需求推动监控发展、以监控发展促进智能化运维算法落地、智能化运维提升运维自动化水平、进而提升业务能力的闭环。

（二）中金财富在智能化运维中的探索、实践

中金财富证券从自身业务需求出发，以事前预测、事中检测、事后分析为发展方向，自 2015 年以来不断探索、实践智能化运维算法，并推进落地，如图 1.2 所示，通过对历史数据的学习与挖掘、对实时数据的即时分析及对异常模式的识别，实现了基于动态数据拟合基线的指标异常检测、基于 FT-Tree 算法的日志概率异常检测、基于趋势预测算法的容量预测与评估、智能流量调度决策，基于 XGBoost 算法的机器健康度模型、基于聚类与贝叶斯算法的异常根因定位等等。在本文后续的章节，我们将主要介绍中金财富在智能化根因定位方面所作的一些探索与实践。

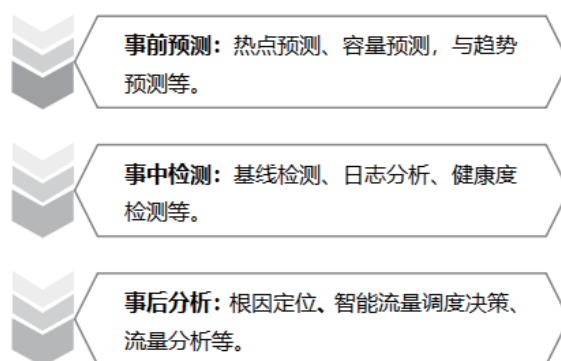


图 1.2 中金财富智能运维实践场景

（三）智能化根因定位的发展

根因定位是智能化运维中一项重要且难于实现的领域。受限于当前的技术水平，即使再

完善的系统也无法避免会出现异常，因此如何将快速定位异常的根因并实施恢复以降低异常造成的影响就显得至关重要。券商实行智能化异常定位的目标就是要尽量降低券商系统出现异常后给客户体验甚至资金安全造成的影响。但随着券商交易系统的规模不断扩大及各系统间的相互影响趋于复杂，让异常定位变得愈发困难。目前行业内外已有多种根因定位方案投入使用，如 Twitter 和大众点评等开源了他们基于调用链的根因定位方案，百度公司采用基于异常范围搜索的 HotSpot 算法进行根因分析，建设银行、光大证券等采用了基于时间序列相关性分析的算法等。但因为系统之间差异性及根因定位的复杂性，目前仍然没有一套能在所有系统行之有效的方案。因此，中金财富证券在学习业内外根因定位思路的基础上，为解决自身运维痛点，提出了两项智能化的根因定位方案，下文会对这两项方案做详细的介绍。

（四）中金财富智能化根因定位方案

在中金财富证券交易系统中，前端机器的响应时间为客户从下单到收到响应的的时间，可以反映客户体验的好坏。一些前端机器的响应时间异常是由后台机器的响应时间异常引起，但前端机器与后台机器、后台机器之间的关联性会随时变动，为异常定位增加了难度，传统的基于链路的异常根因定位在这类情况下效果并不理想。因此，我们设计了一种基于 DBSCAN 聚类算法，在识别到异常发生后，通过将异常的前端机器指标值序列与所有可能与其关联的后台机器指标值序列聚类，从而快速定位到异常根因。同时交易系统有一部分引起异常的根因无法从数值中体现，而在异常发生前后以文字告警的形式出现。但随着监控技术的发展券商运维团队时刻会收到大量的告警，这就需要我们区分哪些告警是与交易系统异常相关，哪些无关。因此，我们基于贝叶斯方法，运用历史数据从概率的角度推断告警与

异常的关联关系，对告警进行排序，从而定位异常根因。在下文中我们会详细介绍这两种方案在中金财富证券的落地实践，供大家参考。

二、基于聚类的异常根因定位方案

在这一方案中，我们主要实现了异常识别、告警压缩、特征处理、聚类实现四项功能。并以中金财富证券的一种移动端交易渠道 A（以下称 A 交易系统）的响应时间为例，对方案进行了实践。本部分主要叙述了异常识别与聚类实现两项功能，以及在 A 交易系统的实践结果。

（一）异常识别

中金财富证券通过基线的方式发现交易系统指标的异常，基线是通过学习历史经验，为业务指标提供判断标准的基准线。在实际应用中，为避免偶发告警造成误告，通常配置将 M 分钟内 N 次超出基线的事件判断为一次异常。为保障异常发现的时效性，M 不宜过大，同时为兼顾异常发现的灵敏性，N/M 也不宜过大。在本方案中我们选择 $M=5$ ， $N=3$ 。这一方案对基线的准确性提出了较高的要求，在中金财富运维实践中，A 交易系统的基线存在两种风险：1) 系统配置发生变更后，与之相关的前端机器响应时间会随之发生变化，但基线的变化存在滞后性，使得与当前响应时间不匹配，导致误告警发出。2) 基线学习了错误的历史经验而偏离，造成误告警。对此，我们利用配置变更校正算法和滑动 T 检验法，降低这两类误告警的概率。

（1）配置变更校正算法

A 交易系统存在经常性的配置变更，在变更后关联机器的响应时间也会发生整体平移，往往会超出基线范围。为应对这种情况，我们设计了配置变更校正算法，通过计算基线中轨序列与响应时间序列间的差值并使用差值校正响应时间，减少此类误告警。算法步骤如下：

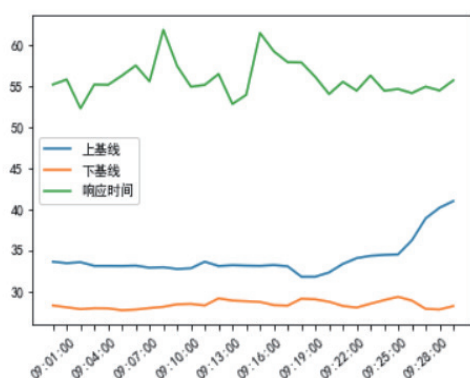


图 2.1 : 变更后的响应时间

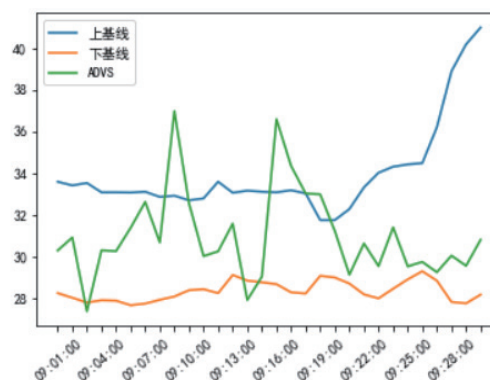


图 2.2 : 经修正后响应时间

1) 读取从判定时刻起前 30 分钟内指标的基线值, 并求出这 30 分钟内指标的基线均值序列 BLVS:

$$BLVS = ((up_1 + low_1)/2, (up_2 + low_2)/2, \dots, (up_{30} + low_{30})/2)$$

其中: up_i 、 low_i 为从判定时刻起前 i 分钟的上、下基线值。

2) 读取从判定时刻起前 30 分钟内指标的真实值, 并与基线均值做出这 30 分钟的差值序列 DSVS:

$$DSVS = ((value_1 - blv_1), (value_2 - blv_2), \dots, (value_{30} - blv_{30}))$$

其中: $value_i$ 为从判定时刻起前 i 分钟的真实值, blv_i 为 BLVS 的第 i 个值。

3) 对 DSVS 中的值进行异常过滤, 设过滤后的序列为 FDSV, 对过滤后的序列求均值, 为变更校正值 BIAS:

$$BIAS = \text{mean}(FDSV)$$

4) 指标校正值序列 $ADVS = \text{value} - \text{BIAS}$, value 为 A 交易系统响应时间序列。

如图 2.1 所示, 中金财富 A 交易系统在一次变更后, 一台前端机器的业务响应时间偏离基线的情况, 图 2.2 为经校正的指标值与基线的比较。利用校正后的指标序列再次与基线进行比较, 可有效减少因配置变更造成的误告警, 该算法在中金财富 A 交易系统的实践中可过滤超过 98% 的告警。

(2) 滑动 T 检验算法

滑动 T 检验算法常用于气象研究中, 用于

检测时间序列是否发生突变。本方案引用滑动 T 检验法, 对异常序列进行突变检测, 若异常序列并未发生突变, 则可认定为并未发生异常。

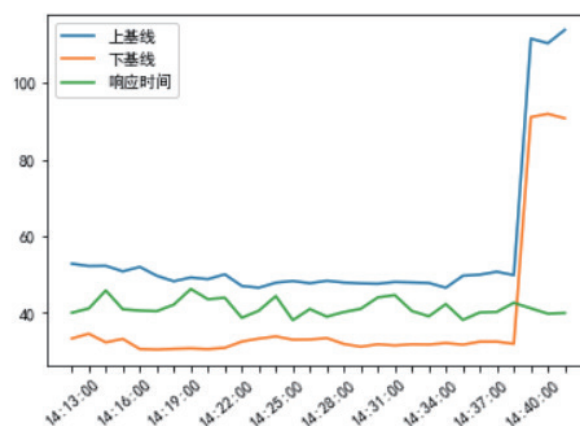


图 2.3 基线偏离事件

图 2.3 为中金财富 A 交易系统的一台前端机器发生的一次基线偏离, 经滑动 T 检验算法, 该次异常被过滤。滑动 T 检验算法可有效降低由于基线偏离而产生误告警的概率。该算法在中金财富 A 交易系统的实践中可过滤超过 10% 的告警, 进一步提高了告警的准确性。

(二) 聚类实现

为实现指标聚类, 首先需定义指标序列间的距离度量, 本方案对比了三种距离度量: 欧式距离、基于 Pearson 相关系数的度量和基于 Spearman 相关系数的度量, 假设 CDS_1 、 CDS_2 为两个时间序列, CD_{ij} 为 CDS_1 的第 j 个分量, n 为

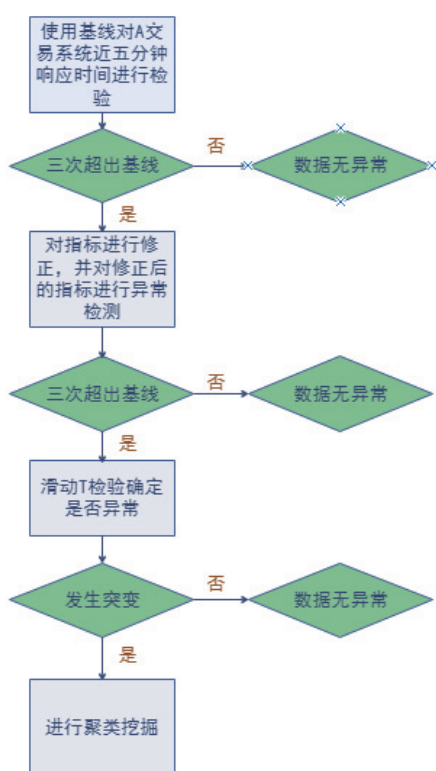


图 2.4：基于聚类的根因挖掘方案流程图

时间序列的长度，他们的距离定义如下：

1) 欧式距离：

$$d(CDS_1, CDS_2) = \sqrt{(CD_{11} - CD_{21})^2 + (CD_{12} - CD_{22})^2 + \dots + (CD_{1n} - CD_{2n})^2}$$

2) 基于 Pearson 相关系数的度量：

$$d(CDS_1, CDS_2) = 1 - \frac{\sum_{i=1}^n (CD_{1i} - \text{mean}(CD_1))(CD_{2i} - \text{mean}(CD_2))}{\sqrt{\sum_{i=1}^n (CD_{1i} - \text{mean}(CD_1))^2 \sum_{i=1}^n (CD_{2i} - \text{mean}(CD_2))^2}}$$

3) 基于 Spearman 相关系数的度量：

$$d(CDS_1, CDS_2) = 1 - \frac{\sum_{i=1}^n (rk_{1i} - \text{mean}(rk_1))(rk_{2i} - \text{mean}(rk_2))}{\sqrt{\sum_{i=1}^n (rk_{1i} - \text{mean}(rk_1))^2 \sum_{i=1}^n (rk_{2i} - \text{mean}(rk_2))^2}}$$

其中， rk_1 、 rk_2 为 CDS_1 、 CDS_2 中的分量转换为降序位置指标后的序列。由于方案主要考察 A 交易系统前端机器响应时间与后台机器响应时间之间有无同增同减的线性关系，因此基于 Pearson 相关系数的度量与基于 Spearman 相关系数的度量更适合本方案，在中金财富 A 交易系统的实践中，基于 Pearson 相关系数的度量对于时间序列的距离度量最为准确。

在异常传导的过程中可能会跨分钟，因此，

对于后台机器相关指标，我们同时取 $((t_i-1) \text{ to } (t_i-1)), (t_i \text{ to } t_i)$ 这两个时间段的值序列，并取这两个序列中与 A 交易系统前端机器异常响应时间序列距离的最小值作为最终距离。

本方案采用 DBSCAN 聚类方法，DBSCAN 基于密度对指标进行聚类，无需事先确定类的个数，与本方案的场景相符。图 2.4 展示了基于聚类的根因挖掘方案流程。

(三) 实践结果

本次实践以 A 交易系统为例：A 交易系统是中金财富某一移动端交易系统。用户在移动端下达交易请求后，首先将请求传送至 A 交易系统的前端机器进行处理，前端机器处理后再按照请求信息，调用 A 交易系统后台机器的各项服务以执行客户请求。因此，前端机器对请求 g 的响应时间包括了前端机器处理请求的时间与各类后台机器执行请求的时间。如以下表达式所示， $AnswerTime(g)$ 为一笔请求 g 的响应时间， $Processtime(g)$ 为请求 g 在前端的处理时间， $AnswerTime_i(g)$ 为后台机器在执行请求 g 时第 i 阶段的响应时间：

$$AnswerTime(g) = Processtime(g) + \sum_{i=1}^k AnswerTime_i(g)$$

当客户请求在前端机器的响应时间发生异常后，我们使用聚类根因定位算法，快速分析用户的请求是在哪一环节出现了问题，以便实施恢复。我们使用 A 交易系统 2020 年 5-6 月份数据进行回溯，共识别出 51 次响应时间异常，成功挖掘出 26 次异常的根因。

图 2.5 为 A 交易系统某前端机器在一时刻响应时间的异常情况，使用本方案进行聚类后，识别出三台后主机的响应时间异常（图 2.6- 图 2.8）为本次异常的根因。

三、基于贝叶斯推断的异常根因定位方案

A 交易系统中，存在一部分引起异常的根因

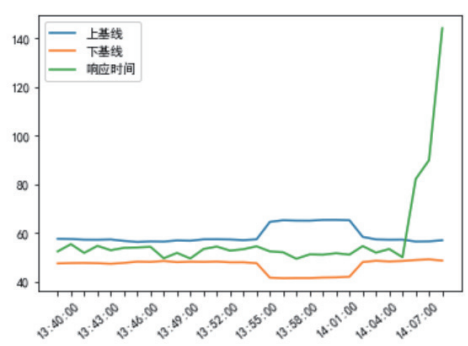


图 2.5 : A 系统某前端机器响应时间

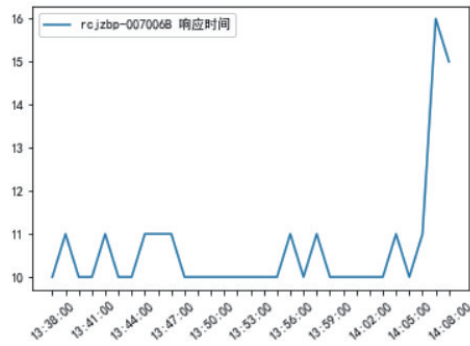


图 2.6 : 相关后台机器甲响应时间

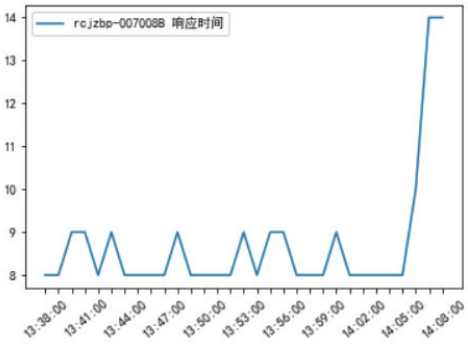


图 2.7 : 相关后台机器乙响应时间

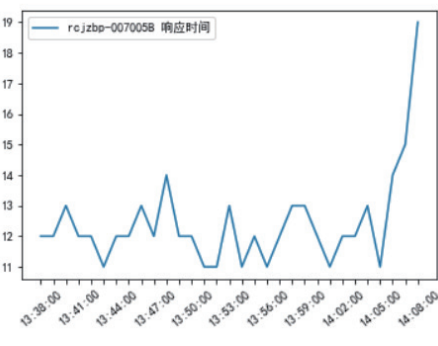


图 2.8 : 相关后台机器丙响应时间

无法从数值中体现，而在异常发生前后以文字告警的形式出现。随着监控技术的发展、监控手段的丰富，中金财富集中告警管理平台每分钟可能收到数以百计的告警，使得区分哪些与异常相关、哪些与异常无关存在一定的困难。因此，我们基于贝叶斯方法，通过挖掘告警与异常之间的关联概率从而推断他们的关联关系，并在中金财富 A 交易系统中进行了实践。下文对方案实现及实践结果进行了叙述。

(一) 方案实现

由贝叶斯公式，告警事件 X 会导致 A 交易系统发生异常事件 Y 的概率 $P(Y|X)$ 可经如下计算得到：

$$P(Y|X) * P(X) = P(X|Y) * P(Y)$$

$$P(Y|X) = P(X|Y) * P(Y) / P(X)$$

因此，条件概率 $P(Y|X)$ 正比于 $P(X|Y)$ /

$P(X)$ ，基于此，我们设计以下步骤实现该推断：

- 1) 统计在历史上，Y 事件发生的前五分钟以内有哪些告警事件发生，并计算每个异常告警 X 出现的条件概率 $P(X|Y)$ 。
- 2) 统计在历史上任意五分钟内 X 事件发生的概率，得到 $P(X)$ 。
- 3) 计算关联概率相对值 $P(X|Y) / P(X)$ ，数值越大，表明告警 X 事件越可能导致 Y 事件。
- 4) 收集 A 交易系统发生异常事件 Y 之前五分钟内的告警，并根据计算所得的关联概率相对值，由高到低将告警排序输出。

图 3.1 展示了基于贝叶斯推断的异常根因挖掘方案模型训练及推断流程。

(二) 实践结果

中金财富在 2015 年建立了全局统一的告警事件平台，用于接受各类监控所产生的告警。由

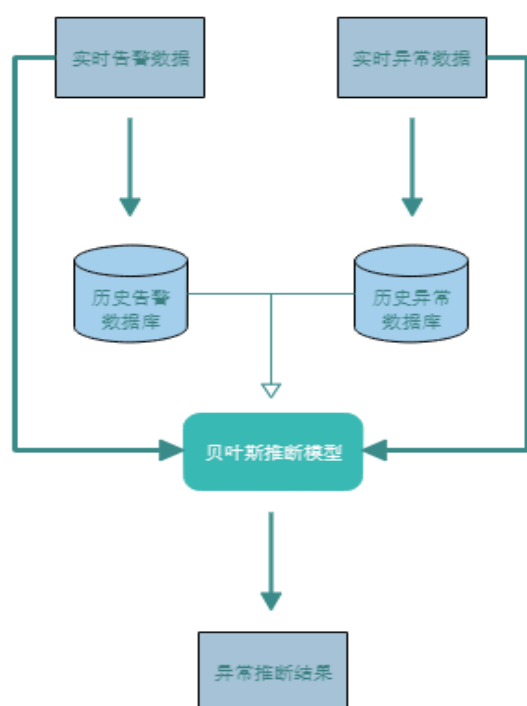


图 3.1：基于贝叶斯推断的异常根因挖掘方案流程图

于监控系统不断完善，告警时间平台在单位时间内收到的告警数量不断增长，其中大部分为轻微、偶发的告警，因此当交易异常发生后，我们通过使用基于贝叶斯推断的根因分析算法确定哪些告警与本次异常相关，从而掌握更多异常细节信息便于实施恢复。我们使用了 2020 年 2-6 月份 A 交易系统的数据进行回测，由于贝叶斯方法的推断需要较大的数据量才能准确，

因此我们选取了一个发生次数较多的异常事件进行挖掘：A 交易系统前端机器 V 交易响应时间异常。经回测发现如下结果：最有可能导致 A 交易系统 V 机器交易响应时间异常的三个告警事件和最不可能导致它的一个告警事件及其推断相对值（见表 3.1）。

贝叶斯推断的结果具有较高的分离度，各类告警与 A 交易系统异常间的关联关系相对值差距明显，作为排序依据有一定的参考价值。

四、总结与展望

为保障客户体验，中金财富证券在行业内较早的实现了基于大数据的自动化运维，并一直致力于智能化运维的推广与落地。本文介绍了中金财富在智能化运维中异常根因定位方面所作的一些探索与实践，所涉及的方案仍存在一定的局限性，在异常根因定位的成功率方面还有一定的提升空间。在近期，智能化根因定位出现了一批新的算法方案，如基于知识图谱与基于自然语言处理的根因定位方案，中金财富会结合新的算法，持续探索、不断改进。随着智能化技术的不断发展，智能化运维必将愈发成为证券企业降本增效的利器，中金财富也会不断地升级技术手段，为客户提供更高质量的服务。

表 3.1：贝叶斯推断结果

事件	相对值
某电子商务网络交换机告警	846.15
甲营业部汇聚路由器告警	52.88
乙营业部汇聚路由器告警	42.31
丙营业部汇聚路由器告警	0.20

网络自动化运维系统自主研发的探索与实践

梁德汉、唐新华、武孟军 / 安信证券股份有限公司



对于网络运维人员来说，日常工作中有一类属于重复性高且繁琐的内容：防火墙策略配置、定期设备配置备份、定期密码修改、配置内外网地址映射、交换机 VLAN 端口划分等等。特别是防火墙策略配置，随着应用系统数量增加，业务访问逻辑的复杂化，导致访问策略开通申请数量剧增。原始的手工配置方式不仅工作量大，而且差错率高，已经不能满足网络运维需求。故，网络自动化运维是 IT 系统运维必然的发展方向，也是大势所趋。但市场上此类成熟产品凤毛麟角，大多要深度定制开发。其中，防火墙策略开通自动化配置，是技术含量最高，也是最难实现的部分。本文重点分享了我司在网络自动化运维工作中为满足防火墙策略自动化配置需求，提出的防火墙策略自动化处理解决方案和自主研发实践经验，希望能“抛砖引玉”，为行业解决此类问题提供一个参考思路。

一、系统简介

网络系统是整个券商 IT 系统的重要基础设施和支撑平台，“牵一发而动全身”，其重要性不言而喻。相较于其他应用系统，网络系统中设备种类繁多，如路由器、交换机、防火墙、负载均衡器等。即使是同一种类设备，由于品牌和型号不同，配置管理方式和命令格式也不尽相同。因

此，在 IT 运维自动化技术发展如火如荼的今天，网络系统自动化运维进程可谓举步维艰，传统的手工网络运维方式依然占据主流。

随着业务的发展，我司“两地三中心”IT 系统架构的成型，网络运维工作量也成倍增加，传统运维模式差错率高、效率低等种种问题凸显出来。为改变这种被动局面，信息技术中心组织技术力量，分析网络日常运维工作中的弊端，结

合目前出现的自动化运维新技术，决定自主研发、搭建一套能够解决网络日常运维工作中重复性高，需求量大，工作量大（需批量处理）的网络自动化运维系统。

1.1 功能需求和系统定位

经过分析，下列运维工作内容全部或部分适用于自动化处理：

- (1) 配置防火墙策略
- (2) 负载均衡设备配置
- (3) 定期设备配置备份
- (4) 定期修改设备密码
- (5) 交换机端口 VLAN 划分
- (6) 网络设备工作状态自动定期巡检
- (7) 应急处理（应急方案自动化）
- (8) 批量设备 IOS 升级

另外，系统定位于实现网络设备的自动化配置功能，不涉及网络设备监控功能。但是系统可以与监控系统建立通信接口，与监控系统建立联动机制，完成某些操作动作。比如，根据监控系统提供的监控数据和预先设置的阈值，触发自动化系统执行某些应急操作命令。

上述功能需求中，防火墙策略配置自动化需求尤为急迫：每天需要处理的防火墙策略开通流程多达 30 ~ 50 个，单个流程中列出的源、目的访问关系平均 5 ~ 10 条。这样，每天运维人员需要配置的防火墙策略条数接近百条，全部需要

手工处理，占用了大量运维人力。同时，这也是所有功能需求中实现起来技术含量最高，最困难的部分。为此，我们决定先实现这部分功能，如果能顺利攻克这个“制高点”，后续部分的功能需求实现起来，也就水到渠成，容易多了。

1.2 系统整体架构

系统整体架构如图 1 所示，包括后台处理功能模块，系统界面和数据库三个主要部分。

其中，数据库保存各种程序数据配置信息，如设备信息，配置文件，包括各类用户输入的流程相关数据。系统界面实现用户接口，管理员配置、操作系统功能，它保存系统信息到数据库，也可以通过通信接口和功能模块通信，调用功能模块进行处理。功能模块实现一个可扩充的处理系统，由实现各种自动化处理功能的子模块组成。软件环境要求如下：

数据库：MYSQL 5.6

操作系统：CentOS 7

Web 服务器：Apache 2.4.41，PHP 7.0

功能模块开发语言：python >= 2.7

经过努力，目前已经完成了系统整体平台的搭建，以及防火墙策略的自动化处理：包括策略信息读取，自动生成配置脚本，自动下发配置等，实现了预期功能。使用本系统处理 50 条左右的策略，全部处理完成耗时在 1 分钟左右，极大提高了工作效率。本文后续内容，重点就防火墙策

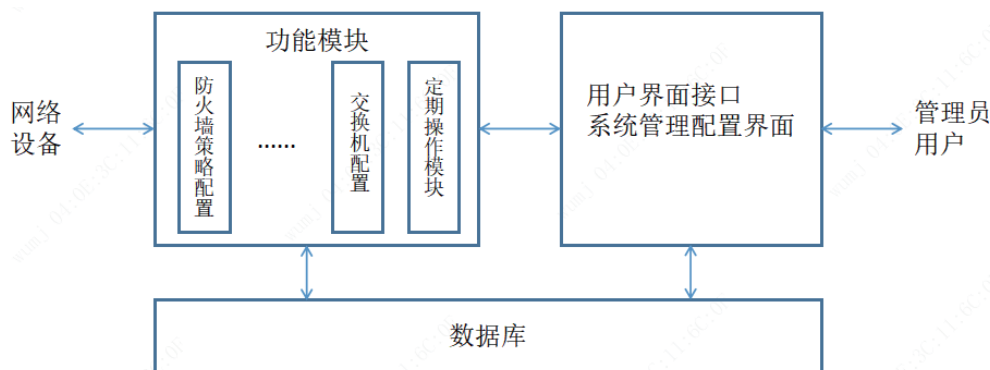


图 1：系统整体架构

略自动化处理功能的实现进行介绍。

二、网络访问策略自动化处理解决方案

2.1 网络访问策略申请及处理流程

默认情况下，网络数据中心安全域之间主机相互访问是全部禁止的，网络访问按照“最小化”策略原则开通。在IT应用系统建设和运维工作中，如涉及到安全域之间主机的访问关系变更，就需要申请开通防火墙策略。通常，系统负责人会通过OA/ITIL系统提一个网络访问权限开通申请流程，经过审批，流转网络运维人员处理后再流转回申请人处测试验证。具体流程如下：

► 申请人

申请人需要清楚应用系统各个组成部分的相互访问关系，即需要开通哪些主机访问哪些主机的哪些端口。这个访问关系随应用系统规模的不同，复杂度也不一样，但终归可以归结为1条或多条以下的访问关系，如表1所示。

在防火墙策略术语中，主机IP地址172.1.1称为源地址，172.10.5.5称为目的地址，TCP端口8080称为服务。源地址，目的地址和服务，可以认为是一条防火墙策略中最关键的三要素。申请人的工作，是要把这些访问关系，以三要素的形式，提交给运维人员处理。

► 网络运维人员

网络运维人员的处理过程如下：

1. 首先根据源和目的地址，确定访问网络路径，路径上部署了哪些防火墙，最终确定需要在哪些防火墙上开通策略。完成这个工作的基础，是运维人员要十分熟悉网络的整体拓扑，以及整个网络中防火墙的分布情况。

2. 假定需要在防火墙F1和F2两台防火墙上

上开通策略，下一步要确定F1和F2是什么品牌的防火墙？因为不同品牌防火墙，配置命令大相径庭。

3. 一般的防火墙设备都提供web页面和命令行方式进行管理。登录防火墙，输入配置一条策略的必需信息，完成配置。如果是web页面，会通过表单提交信息。如果是命令行方式，则需要逐条输入完整的配置命令。

2.2 防火墙策略自动化处理

通过上面的过程分析，可以看到，申请人的工作量主要集中在访问关系的整理上；运维人员的工作重点则是在步骤1-3上，这也是最耗费人力，最费时的工作内容。策略自动化系统，就是要将运维人员的工作步骤1-3全部实现自动化。

也就是说，自动化配置系统需要根据申请人提供的策略信息，完成下面三部分工作：

- 选择要配置的防火墙
- 生成具体的配置脚本
- 下发配置脚本到具体的设备

其中，确定访问路径中的防火墙，首先要将网络中防火墙部署关系，也就是拓扑图用某种数据类型表示。比如将网络安全域表示为图的节点，网络链路表示为图的边。企业网网络拓扑相对稳定，不会频繁发生变化。这样，访问路径的确定实际上转化为有环路无向图的最短路径问题。

确定了访问路径，路径中有哪些防火墙，也就可以确定了。当然，如果网络规模不大，结构相对简单，也可以采用简单的手工确定方式，将网络中所有网段之间的访问关系列出来，每种访问关系经过哪些防火墙，保存到配置文件中直接使用。

在设备上下发配置策略，有两种方式可选，一是模拟手工通过命令行方式进行，即通过远程

表1：访问关系

源地址	目的地址	协议	端口	备注
172.1.1.1	172.10.5.5	TCP	8080	

登录，将配置命令逐行提交设备。二是目前很多设备提供各种 API 接口，调用这些编程接口可以实现配置策略的功能。综合比较，考虑到通用行、执行效率和灵活性，选择使用第一种方式。远程登录可以使用 python 语言提供的 telnetlib，或第三方库 paramiko 编写脚本实现，也可以考虑和目前流行的各种自动化运维平台结合，由第三方运维平台提交。

大部分重要和繁琐的工作，是生成策略配置脚本的过程。可以说，这个功能实现的差异，直接决定自动化配置系统的成败和智能化程度的高低。下面着重予以说明。

2.2.1 防火墙安全策略分析

图 2 是一条典型的防火墙安全策略。

可见，一条安全策略包括源安全域，目的安全域，源地址集合，目的地址集合，端口集合（包括 TCP，UDP），动作（允许，拒绝，转发到 VPN 隧道）等几个重要组成元素。这条策略的控制意义是：如果有一个数据包从源安全域进入，经过防火墙处理后，要从目的安全域出去，数据包的源 IP 地址包含在策略的源地址集合中，目的地址包含在策略的目的地址集合中，且端口包含在端口集合中，则称为该数据包匹配了这条安全策略，防火墙根据策略中指定的动作，允许或拒绝该数据包通过。

注 1：防火墙中的安全域可以简单理解一个命名的若干物理网络接口的集合。服务端口集合也叫服务，除了 TCP，UDP 端口，还可以包括其他协议，比如 ICMP。简单起见，这里只涉及 TCP，UDP 端口。

其中，源地址集合，目的地址集合，端口集合是组成一个策略的三个关键元素。在防火墙策

略中，描述网络地址的形式可以是网络地址 + 掩码，一段起止地址范围，也可以是域名。

地址也可以以地址对象的形式出现。一个命名地址对象是若干个上述地址形式的集合，可以代表一组地址。同样地，一个命名服务对象是若干 TCP，UDP 端口范围的集合，表示一组服务端口。在定义策略时，网络地址可以直接在策略中使用，也可以将一组地址定义为一个地址对象后在策略中引用。服务必须先定义为一个服务对象后才能在策略中引用。

注 2：不同防火墙品牌在定义策略时地址和服务端口使用规定也不尽相同。另外，不同防火墙在单条策略中允许出现的地址条目数量，服务条目数量也不同。使用时请查询使用手册，这里不展开讨论。

如何在防火墙上配置一条这样的策略呢？以山石网科防火墙命令行方式为例（下同），分为以下几个步骤：

- (1) 定义端口对象
- (2) 定义地址对象（如需要）
- (3) 定义策略

注 3：策略中地址条目比较少的话，可以直接使用地址定义，不用单独定义地址对象。如果地址过多且不连续，或地址信息中包含有域名条目，需要定义地址对象。

另外，要考虑的问题如下：

- (1) 防火墙上已经存在了一条这样的策略吗？或已经存在了一条包含了这条新增策略访问关系的策略吗？如是，则这条策略无需再配置了；
- (2) 新开策略，端口对象存在吗？如是，则可以直接在新策略中引用；否则新定义一个命名端口对象；



图 2：防火墙中一条实际的策略

(3)如需定义地址对象,现有配置中存在吗?如是,则可以直接在新策略中应用;否则新定义一个命名地址对象;

(4) 上述步骤 (2) 或 (3) 中, 如果不存在相等的对象, 是否存在 2 个或几个已有对象, 他们的并集等于新增对象呢? 如有也可以直接拿来使用;

注 4: 在配置新策略时, 之所以尽量引用防火墙中已有配置的地址和端口对象, 不新增加配置项, 一是可以减少配置文件大小, 二是出于提高设备运行性能考虑。因为在配置文件中的地址或端口对象, 运行时最终都要加载到内存中的。

2.2.2 配置文件分析

由前面的分析可以知道, 掌握防火墙现有的配置信息, 是高效配置安全策略的基础。我们采用下载设备当前配置文件的方式, 对整个配置文件逐行进行分析, 解析出设备中现有配置的各种对象、策略等重要配置信息保存, 为后续的策略生成决策提供数据分析依据。

配置文件的分析的过程其实就是文本处理, 比如在配置文件中, 一条完整的策略配置信息如下:

```
rule id 2305
  action permit
  src-zone "B_untrust"
  dst-zone "B_trust"
  src-ip 172.97.1.204/30
  src-ip 172.0.242.62/31
  dst-ip 172.0.36.152/31
  service "tcp_1521"
  description "CM004184"
  exit
```

每行的组成由关键字 + 配置参数组成。一般情况下, 关键字都在行的开头。因此, 可以根据判断关键字的结果, 从配置文件中分析出一条完整的策略信息。上面的例子中, 一条策略的开始由“rule id 2305”这一行开始, 后续的行为该条策略的其它配置信息, 直到“exit”结束。

以此类推, 可以从配置文件中解析出已定义的地址对象和服务对象等关键配置信息, 保存备用。

2.3 生成策略配置命令序列

有了设备当前配置信息, 后面的策略配置命令生成就容易多了。这里我们仍以配置一条策略最重要的三个元素为例, 说明如何产生对应命令序列。

在生成配置命令序列之前, 需要做的一个重要步骤是, 在当前已配置策略中进行查找比对, 确定是否已经存在策略, 包含或等于本次新增的策略。另外还可以查找新增策略是否和当前策略中有可以合并的策略, 如有, 则建议合并以减少设备中策略数量。查找当前策略和新增策略进行比较的算法略显复杂, 这里就不再展开叙述了, 下面讲述新增策略的脚本生成步骤。

(1) 源地址和目的地址

在一条策略中, 如果源地址有 10 个, 分别是 172.0.5.1--172.0.5.10, 那么, 在策略配置命令中, 可以是这样的单个地址分列的形式:

```
...
src-ip 172.0.5.1/32
src-ip 172.0.5.2/32
...
src-ip 172.0.5.10/32
...
```

也可以是地址范围的形式:

```
...
src-range 172.0.5.1 172.0.5.10
...
```

还可以是先定义地址对象, 再在策略中引用的形式:

```
address "my_source_address"
range 172.0.5.1 172.0.5.10
...
src-addr "my_source_address"
```

单条策略中定义源地址语句有数量限制，一般为 8、16 或 32 个。而一个地址对象中可定义地址语句则可以多于 256 个。因此，当地址经过合并处理，比如，连续地址用一个网段表示，或用一段地址范围表示，使得在策略中可以用几个语句定义完成的，在策略中直接定义即可；否则，定义一个地址对象，然后在策略中引用。

在定义地址对象时，应充分利用当前配置中的地址对象。如有完全相同的，则直接引用，无需再定义地址对象；如无，则应查找是否有几个当前的地址对象合并后等于现有的地址集合，如有，直接在策略中引用这几个地址对象。更进一步，如果可以找到有限几个地址对象的并集，接近新增地址集合，只要在策略中定义地址语句不超过数量限制，就可以在先引用这几个地址对象，然后再将未被包含的新增地址集合中的地址，使用单独的地址定义语句添加进去即可。如此劳力费神操作，目的还是遵循这样一个重要原则：新增策略配置尽量少在设备中增加新的配置元素。

目的地址的处理和源地址处理过程完全相同，只不过在对应配置语句中，将关键字 "src" 全部替换为 "dst" 即可。

(2) 服务

简单起见，服务只考虑 TCP、UDP 端口。和地址不同的是，在策略配置中，服务必须先配置为一个服务对象，然后才能在策略配置中引用。且策略配置中服务的关键字只有一种，就是 "service"，后跟已经定义的服务对象。

同样定义服务对象也应该采用和定义地址对象相同的算法，尽量利用当前配置中已有配置项，减少新增配置项。例如，当前配置文件中有如下服务对象配置：

```
...
service "tcp_7700"
  tcp dst-port 7700
  exit
```

```
service "tcp_9016"
  tcp dst-port 9016
  exit
...
```

假定本次新增的服务又恰好只有 TCP 7700 和 9016 两个端口，那么在策略定义中就可以直接使用如下定义：

```
...
service "tcp_7700"
service "tcp_9016"
...
```

最后，分别将三部分命令序列组合起来，加上其它固定部分，比如源和目的安全域，按顺序组合到一起，形成一个完成策略配置命令序列。至此，一条防火墙策略的完整配置脚本就生成了。

三、项目实施成果

安信证券网络自动化运维系统于 2019 年 9 月份立项建设，2020 年 1 月份上线试运行。项目一期的主要目标，是实现公司“两地三中心”网络中（包括运营商 IDC 托管机房）近 30 台物理防火墙日常运维中策略开通工作的自动化配置处理。

3.1 系统主要功能

- 根据申请信息中的源，目的地址判断相关主机所处的安全域，进而找出正确的访问路径，找出路径中的防火墙，最终确定需要在哪些防火墙上配置策略。
- 根据防火墙现有配置信息，判断新增策略是否已经存在或被包含，如是则忽略该条策略。
- 根据防火墙现有配置信息，判断新增策略是否可以和已存在策略合并。
- 根据防火墙现有配置信息，判断新增策略是否可以利用已经存在的端口，地址条目等配置信息，减少新增配置信息。

- 判断自否需要做地址转换，如需要，则根据地址转换规律生成地址转换命令序列。
- 每天生成一个配置脚本文件，永久保存。
- 策略下发完毕，生成配置反馈文件，记录实际配置结果，永久保存。

3.2 系统架构及处理流程

系统架构如图 3 所示。

整个系统由策略信息录入服务器、数据库和处理机组成。逻辑功能上分为两部分，一是完成申请人通过 Web 页面提交策略参数信息，二是后台策略数据处理部分。

Web 服务器为用户提供登录认证、策略参数录入页面。录入数据在前端进行合法性检查后，保存到数据库。事务处理机在指定时间内处理录

入数据，产生配置脚本，然后下发到目标设备。

申请流程需要通过 ITIL 系统发起，Web 服务器和 ITIL 系统做了接口对接，跟踪申请流程在 ITIL 系统中的审批状态。审批通过后的流程，相关信息被保存到待执行表中。

另外，系统提供了一些附加功能：策略查询功能，路径查询功能，内外网映射查询等。

3.3 主要功能展示

3.3.1 策略信息录入功能

策略信息输入界面如图 4 所示。申请人通过输入界面，将流程需要开通的策略，逐一输入。同时，为提高输入效率，还提供了已输入策略的复制编辑功能。申请人也可以在历史申请流程中选择一个进行完整复制，然后再编辑提交。

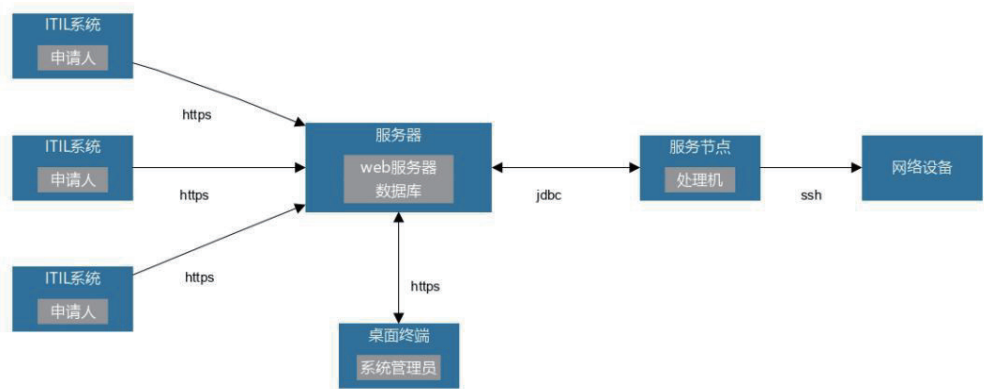


图 3：系统架构

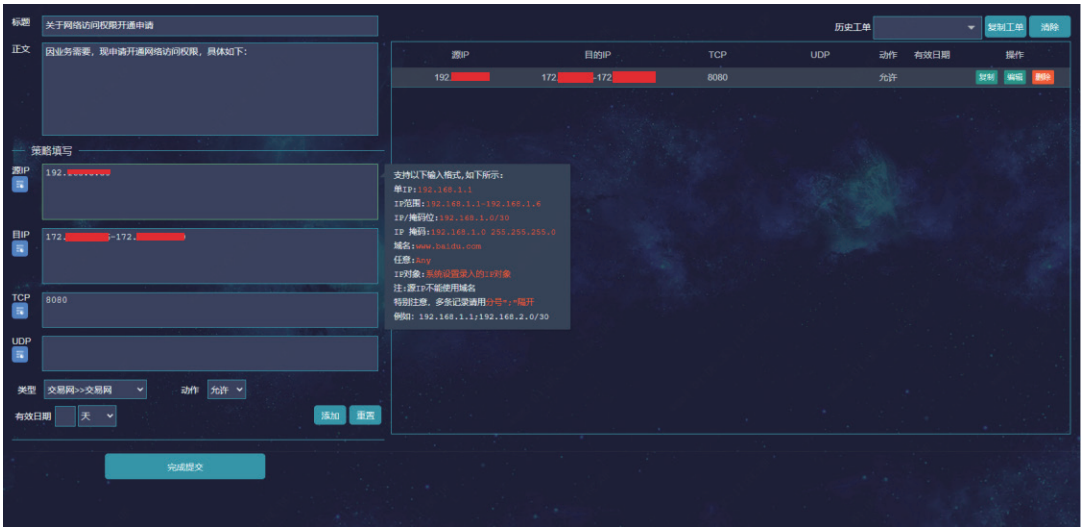


图 4：策略信息录入界面

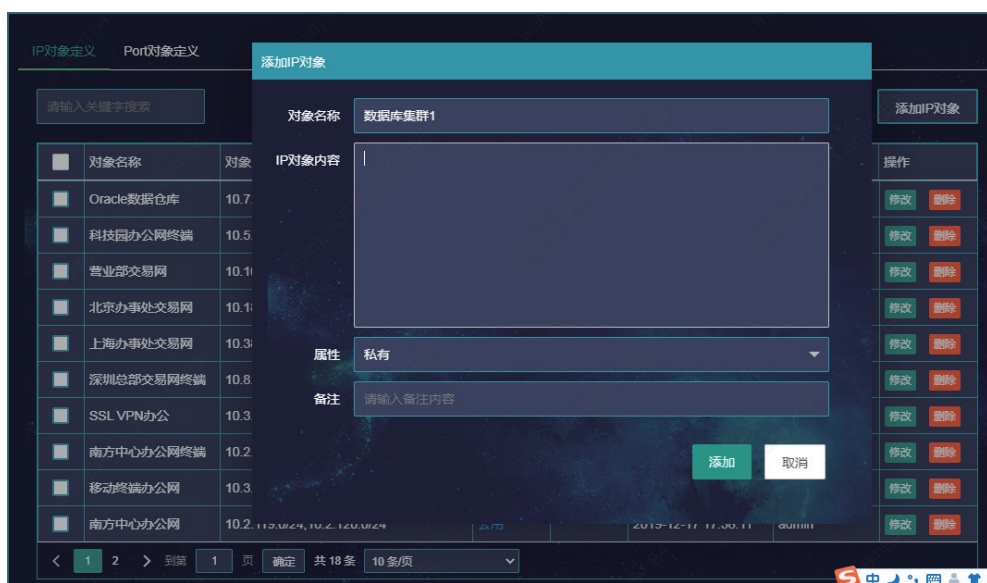


图 5：对象定义

信息录入指定了单个地址，地址段，地址+掩码以及域名等形式限制，否则在添加策略时会给出错误提示。另外，在策略录入完成后提交流程时，会对每条策略进行初步检查。比如不能访问互联网的主机在策略中申请访问互联网，源和目的主机位于同一安全域内（不需要开放防火墙）等。

3.3.2 对象定义功能

对象定义功能见图 5。系统管理员可以预先定义一些面对全体用户的，常用的大地址段，比

如“XX 办公网”“XX 数据库集群”“XX 负载均衡组”等。这些地址可能经常在策略中被引用，且包含的地址多而分散。定义好地址对象之后，就可以直接在策略申请输入信息时使用。用户也可以预先定义一些私有的地址对象，仅供自己使用。

除了地址对象，系统还支持端口对象定义。

3.3.3 拓扑图及查询功能

拓扑图及查询功能见图 6。这里的拓扑并非

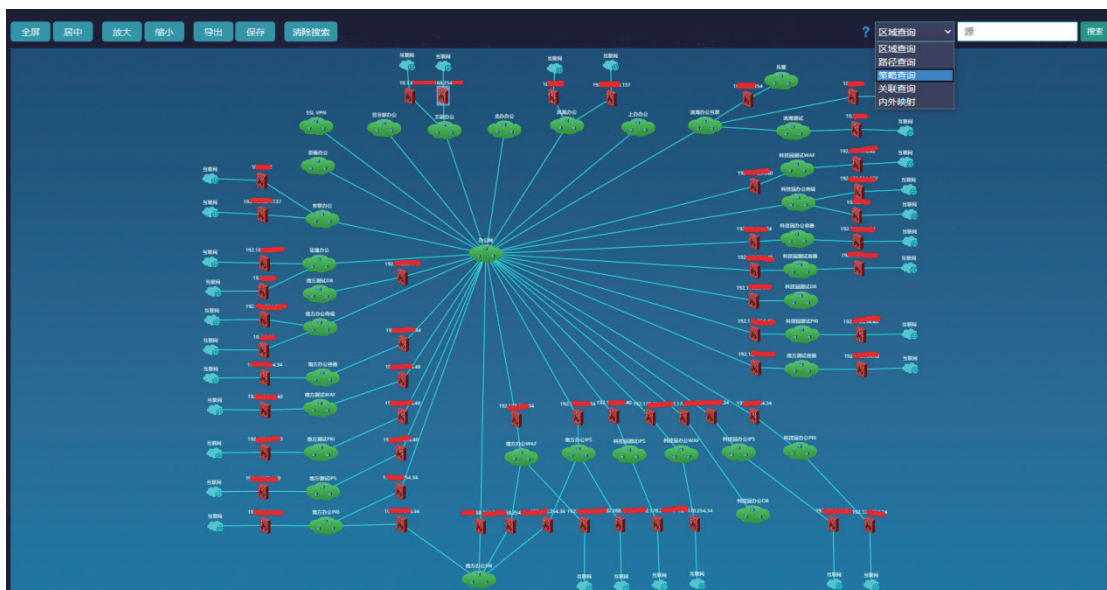


图 6：拓扑图及查询

传统意义上的网络拓扑图，更确切地，应该称其为防火墙部署图。图中每个节点代表一个网络安全域。连接两个节点之间的连线，如果有防火墙，则两个节点之间的主机相互访问，就需要处理防火墙策略。如果连线上没有防火墙，说明两个节点之间的安全等级相同，可以视为同一个安全域。

基于这个基本拓扑，可以实现一些查询功能。比如某台主机所处的安全域？某台主机是否可以访问互联网？某台主机访问另外一台主机，是否经过防火墙？是否存在对应的策略？等等。

3.3.4 流程处理功能

流程处理，暂时采用在指定时间段手动触发方式。点击“开始处理”后，由后台处理机启动处理程序，对数据库待执行表中的策略逐条读出，分析处理，生成对应的脚本文件，并在窗口中展示出来。图7是一个系统实际生成的脚本配置文件的一部分。

可以看到在图中框内，自动化系统搜索到一

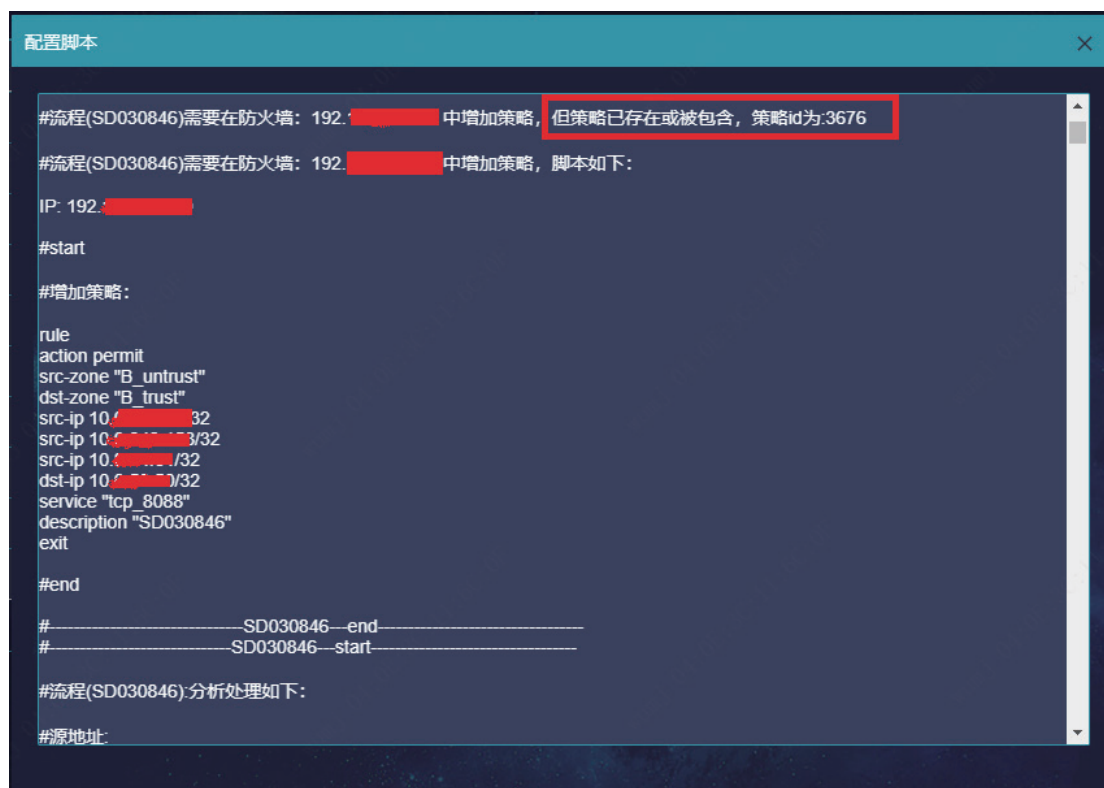
条已经存在的策略，包含了本次要新增的策略，结果将忽略该条策略的后续处理。

运维人员可以检查脚本的正确性，如有错误，在窗口直接进行编辑修改后，保存提交，点击“执行”，则后台处理机启动下发进程，远程连接防火墙设备，逐条提交配置命令。图8是配置下发完成后，记录的结果文件内容中一部分。

下发过程采用多线程处理方式，每个需要配置的防火墙启动一个线程，这样，下发策略需要时间最长的设备，即为本次下发操作耗费的时间。

四、结语

本系统从功能上模拟了人工处理防火墙策略开通流程的整个过程，实现了自动化、智能化处理策略配置。自动化比人工处理优势明显：可自动在成千上万条策略中查找是否有包含新增策略、自动查找需要当前配置中是否已存在



```
配置脚本

#流程(SD030846)需要在防火墙: 192.168.1.1 中增加策略, 但策略已存在或被包含, 策略id为:3676

#流程(SD030846)需要在防火墙: 192.168.1.1 中增加策略, 脚本如下:

IP: 192.168.1.1

#start

#增加策略:

rule
action permit
src-zone "B_untrust"
dst-zone "B_trust"
src-ip 10.0.0.0/32
src-ip 10.0.0.0/32
src-ip 10.0.0.0/32
dst-ip 10.0.0.0/32
service "tcp_8088"
description "SD030846"
exit

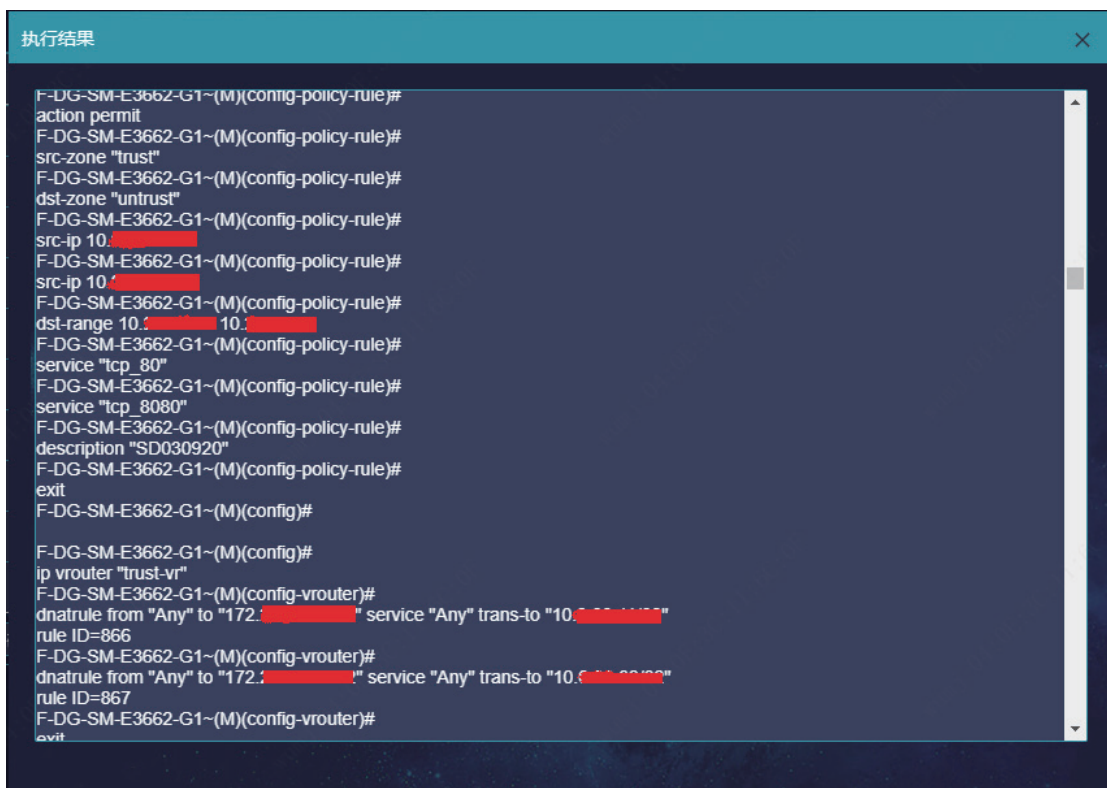
#end

#-----SD030846---end-----
#-----SD030846---start-----

#流程(SD030846)分析处理如下:

#源地址:
```

图7：配置脚本生成



```
执行结果
F-DG-SM-E3662-G1~(M)(config-policy-rule)#
action permit
F-DG-SM-E3662-G1~(M)(config-policy-rule)#
src-zone "trust"
F-DG-SM-E3662-G1~(M)(config-policy-rule)#
dst-zone "untrust"
F-DG-SM-E3662-G1~(M)(config-policy-rule)#
src-ip 10.10.10.1
F-DG-SM-E3662-G1~(M)(config-policy-rule)#
src-ip 10.10.10.1
F-DG-SM-E3662-G1~(M)(config-policy-rule)#
dst-range 10.10.10.1 10.10.10.1
F-DG-SM-E3662-G1~(M)(config-policy-rule)#
service "tcp_80"
F-DG-SM-E3662-G1~(M)(config-policy-rule)#
service "tcp_8080"
F-DG-SM-E3662-G1~(M)(config-policy-rule)#
description "SD030920"
F-DG-SM-E3662-G1~(M)(config-policy-rule)#
exit
F-DG-SM-E3662-G1~(M)(config)#

F-DG-SM-E3662-G1~(M)(config)#
ip vrouter "trust-vr"
F-DG-SM-E3662-G1~(M)(config-vrouter)#
dnatrule from "Any" to "172.16.1.1" service "Any" trans-to "10.10.10.1"
rule ID=866
F-DG-SM-E3662-G1~(M)(config-vrouter)#
dnatrule from "Any" to "172.16.1.1" service "Any" trans-to "10.10.10.1"
rule ID=867
F-DG-SM-E3662-G1~(M)(config-vrouter)#
exit
```

图 8：脚本下发记录文件

地址和服务对象、高效发布变更至防火墙设备、大幅降低人工操作出差概率等。由于这段时间我司数据中心机房处于搬迁阶段，新、旧两个机房共存，不仅策略开通量大增，而且网络拓扑变化频繁，自动化系统很好地应对了这种复杂环境，运行平稳。

网络运维自动化系统自 2020 年 1 月份投入试运行，至今共处理申请流程近 4000 条。其中，每条流程包含访问权限开通申请平均 2.2 条，剔

除已有包含策略，实际产生并下发到防火墙的具体访问策略条目接近 8000 条。自动化系统每天处理策略，下发配置脚本耗费时间控制在 1 分钟左右，差错率控制在 1% 以下。

后期，将不断增补其它网络设备的自动化配置功能，例如负载均衡设备的自动化配置、交换机和路由器等网络基础设备的日常运维自动化配置等，实现大部分日常运维工作内容自动化处理。

关于IDC运管平台建设的思考

林杭宁 / 上交所技术有限责任公司

黄勇刚 / 上海海旭软件有限公司



近年来，各行业核心机构的托管数据中心逐步建成并向行业投放，数据中心的规模越来越大，而数据中心的运营管理手段也需匹配相应规模。本文针对行业数据中心的现状未来发展，调研市场托管数据中心的运营管理经验，提出建设统一 IDC 运管平台的设想。IDC 运管平台作为 IDC 数据中心运营管理的生产指挥系统，承担着 IDC 日常的运营管理和运维管理，涉及 IDC 运营相关的所有职能：客户管理，资源管理，业务开通，资产管理，日常运维，运营分析，管理门户，备品备件，系统管理。通过平台的建设，提升 IDC 运营管理水平，支撑售前 / 售中 / 售后管理，促进精细化运营。

一、前言

随着上交所金桥数据中心的建设完成，IDC 的业务量与日俱增，IDC 的运营与运维管理也愈

发重要。金桥数据中心作为行业技术发展的基础设施，已得到行业内用户的充分认可。为有效支撑 IDC 业务发展，实现对 IDC 客户、合同、订单、业务、资源调度、计费等统一集中运营管理的目

标，做到 IDC 业务全量资源准确掌握、业务产品种类灵活定制、业务上线过程准确监控、业务运营数据精确获取、违规业务及时管控，需要建设一套 IDC 运管平台，以支撑 IDC 的日常运营管理。

二、遇到的挑战

金桥数据中心作为亚洲金融行业最大的数据中心基地，运营管理工作涉及范围非常广泛。除了原有的机房、机柜、楼内链路等管理，更增加了楼内结构化布线、跨楼互联、园区管井等，资源种类繁多且数量巨大，资源管理工作十分复杂，如何保证资源有效管理、提高业务开通速度、提升运维管理质量、保障数据中心安全运行与业务发展，存在诸多挑战，主要包括：

• 资源的有效管理

从整体角度来看，资源管理分散在不同部门进行单独管理，手段主要为业务管理系统中简单的列表式管理与 excel 表格的静态管理，管理粒度主要是机架粒度，没有相应的资源配置流程，没有客户、资源、合同的关联管理，无法实现精细化管理。

• 业务的快速开通

现有市场 IDC 业务的运营机构之间的竞争愈发激烈。金桥数据中心可售资源数量大，要想

在蓬勃发展的 IDC 服务市场上占据一席之地，需要快速响应客户需求、快速开通业务、提升 IDC 服务品质。

• 资源的可视化手段

作为亚洲金融行业最大的数据中心，如何可视化的展示与管理整个数据中心的资源，如何提高资源利用率，如何在故障时进行快速业务分析和倒换，迫切需要建设资源可视化手段，以提升数据中心运营管理能力。

• 运维流程化管理

对于日常的值班管理、人员设备进出管理、巡检管理等，需要有更贴近一线运维需求的电子化、流程化支撑手段，保障日常运维工作的顺利开展，提升运维管理能力。

三、体系架构

平台技术架构如图 3.1。平台按照云化三层的 IT 架构搭建，基础设施由云化资源池提供能力承载，软件架构通过服务总线实现前后端分离，将业务逻辑统一收敛到后端应用，前端系统轻量化，关注使用者交互和体验；随着互联网式营销，业务并发量要求平台具备在线自动扩展能力，通过分布式数据总线实现应用与数据解耦，实现应用随并发量而伸缩，数据随着业务量增长

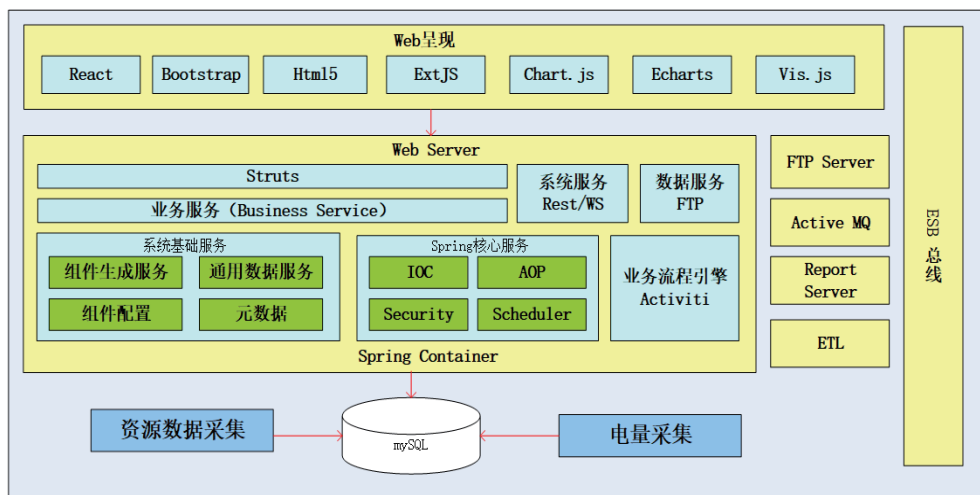


图 3.1：平台技术架构

而自动分片。

平台功能架构如图 3.2。通过建设权威的、全面的、标准的 IDC 管理信息库，将 IDC 业务开展中涉及到的各类资源（物理资源、备品备件、耗材等）全部入库进行信息化管理，对空间的管理力度细化到地板、机架、MCB，对设备的管理细化到插盘、端口。在 IDC 资源的分配和使用上，通过规范化、流程化的流程进行资源配置和调度。支撑值班、巡检、工具管理、人员设备出入等日常运维工作。

• 资源管理

通过自有资源、客户资源、业务的逻辑关系，实现覆盖机房、机架、U 位、IP、链路、带宽、端口等物理和逻辑资源的全量精确管理和灵活调度，并能在统一的资源管理基础上实现对于业务开通、运维支撑、运营分析、用户服务等上层管理功能的支撑。

• 业务开通

提供客户管理、合同管理、产品管理、定单管理、工单管理等功能，支持日常业务开通，提高业务开通的及时率、正确率和完整率。具备灵

活的全生命周期管理功能，具备灵活的流程定制能力。

• 运维管理

对日常运维工作进行电子化、流程化的管控。通过机房出入管理、值班管理、考勤管理、巡检管理等流程化手段，提升日常运维水平。

• 运营分析

通过运营分析，进行业务趋势分析、工单趋势分析。通过客户经理视图，查看客户的业务、资源开通情况。通过容量视图，查看 IDC 数据中心的机架、端口、IP、工位的容量情况。

• 客户 Portal

通过面向客户的安卓 APP 客户端，支持客户自助服务、问题咨询、远程协助、资源查看、人员 / 设备进出申请等。

四、关键问题分析和实践

4.1 链路自动调度

对于 IDC 数据中心而言，如何才能更好地进行光缆资源调度，是其正面临的重要问题。现

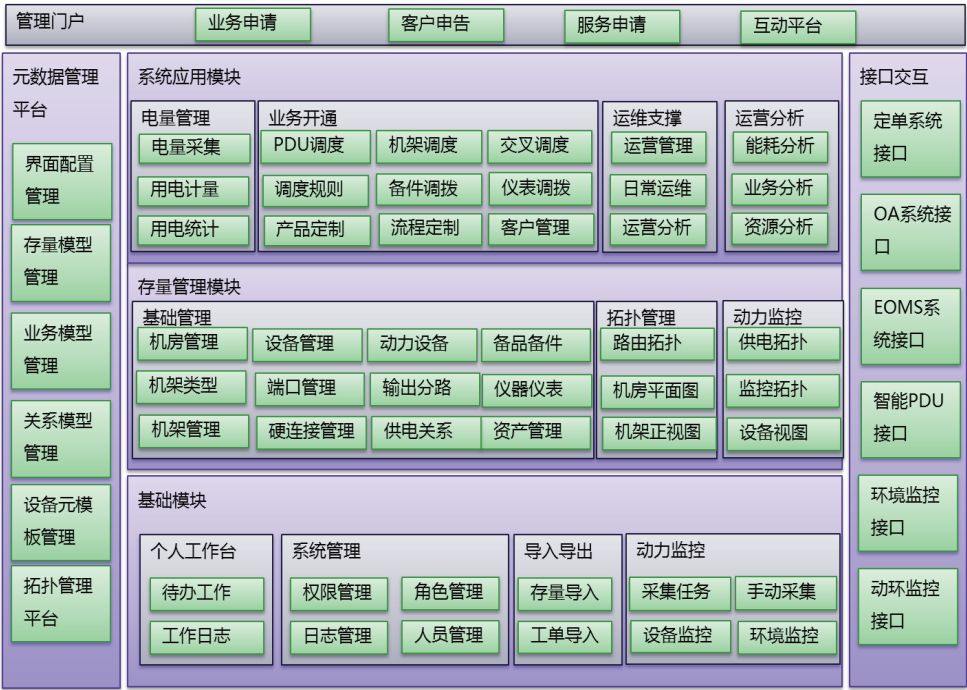


图 3.2：平台功能架构

有技术方案采用人工表格记录的方法，没有统一的全局视图，没有将业务、资源、客户进行关联，没有对光缆调度进行闭环管理，没有对调度规则进行统一建模和抽象，无法满足日常的运营维护要求。日常的业务开通和运行维护，都需要操作人员具备很强的记忆和经验，对现网数据非常熟悉，跨系统 / 表格进行数据维护，利用记忆和经验实现业务开通和光缆运维，给维护人员增加了大量的维护工作量。同时因为无法实现对光缆的精细化管理，导致资源利用率低，开通效率低，客户满意度差。无法实现光缆可管理、可分析、可维护、可视化的管理目标。需要解决以下技术问题：

- ▶ 没有一个统一的视图，不能进行多维度关联展示

- ▶ 没有对规则进行统一的梳理和抽象，没有办法基于规则进行自动化调度

针对以上问题，主要考虑如下：

1) 建立统一视图。通过对 IDC 相关资源统一建模，涵盖机楼、楼层、机房、地板、机架、DDF 架、ODF 架、U 位、光缆、光缆段、光纤、光路、定单、工单、客户等，实现资源、业务、客户的多维度关联，并基于日常运维需求，建立一套基于机房和光缆的统一视图，全面分析展示 IDC 的资源提供能力。

2) 基于已经建立的统一视图进行光缆的自动调度，通过分析日常光缆调度场景，抽象出调度规则模板（路由最短、转接点最少、指定路由节点、光缆冗余选择、光缆缆序选择），将规则进行封装并设定优先级，基于规则算法实现自动化调度。

如图 4.1，选择起始机房 / 设备 / 端子、终止机房 / 设备 / 端子，选择调度规则，系统通过算法自动进行路由计算，生成端到端图形路由，通过图形中的点、线，都可以查看到相应的资源信息。

4.2 资源全生命周期管理

资源的全生命周期管理是平台的核心功能。将资源对象全生命周期状态管理引入到平台中，根据资源对象状态来集中管理网络资源规划、设计、施工、竣工验收、运行维护、退网、报废。规划部门可以利用现有的资源信息数据来完成资源规划，规划后的成果交给设计部门进行设计，设计完成后的成果交给施工部门进行施工，施工完成后，由竣工验收部门对施工数据进行验收，验收完成后交给维护部门使用。这些过程中只有资源对象状态的批量自动转移，没有大量的数据重复录入和倒入。节约了大量的人力和时间投资，保证了数据的正确性和及时性。

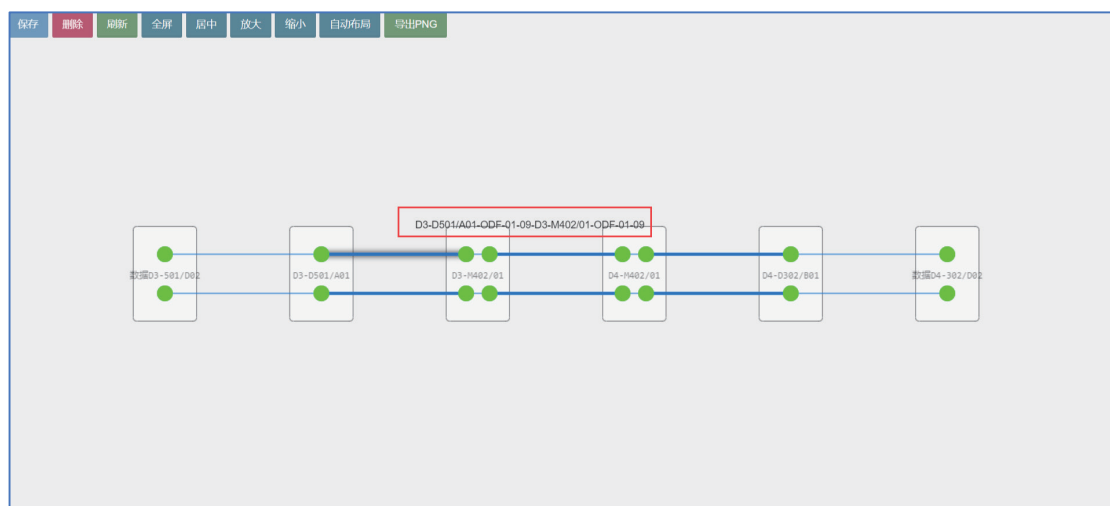


图 4.1：链路自动调度

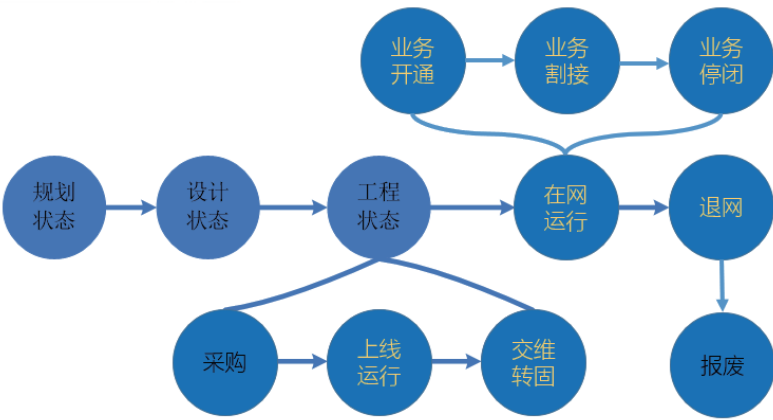


图 4.2：资源全生命周期管理视图

IDC 运管平台建设中，主要围绕支撑资源全生命周期的管理，存在以下流程，包括资源数据入库流程、业务查勘流程，资源开通流程、资源退网流程等，其中业务流程进一步细分为业务预受理（查勘）流程、开通流程、变更流程、撤销流程、预占到期提醒流程、测试到期提醒流程、欠费停机流程、复机流程等。

4.3 数据中心可视化

数据可视化不仅直观的展示数据，提高运维效率，而且通过数据分析，展示容量情况，展示业务发展情况，展示运营和经营结果的关联关系，提高操作的便捷性和系统的易用性。主要从以下几个维度考虑：

- 机楼导览图
平台通过机楼导览图，用图形方式展示数据中心的整体机楼情况，楼层情况，机房分布情况等。通过机楼楼层分布图，展示某机楼所有楼层分布图。
- 综合布线图
平台不仅要管理楼内的光缆、设备资源，还需要管理园区内楼栋间的光缆、管线资源。光缆带宽是数据中心的核​​心资源，也是对外业务的核心竞争力。

平台通过综合布线图，用图形方式呈现各机房之间的光缆连接状况和机架光缆链接情况。平台自动根据光缆两端机架所在机房，动态生成以机楼、机房为节点的光缆全局图，展示所有光缆

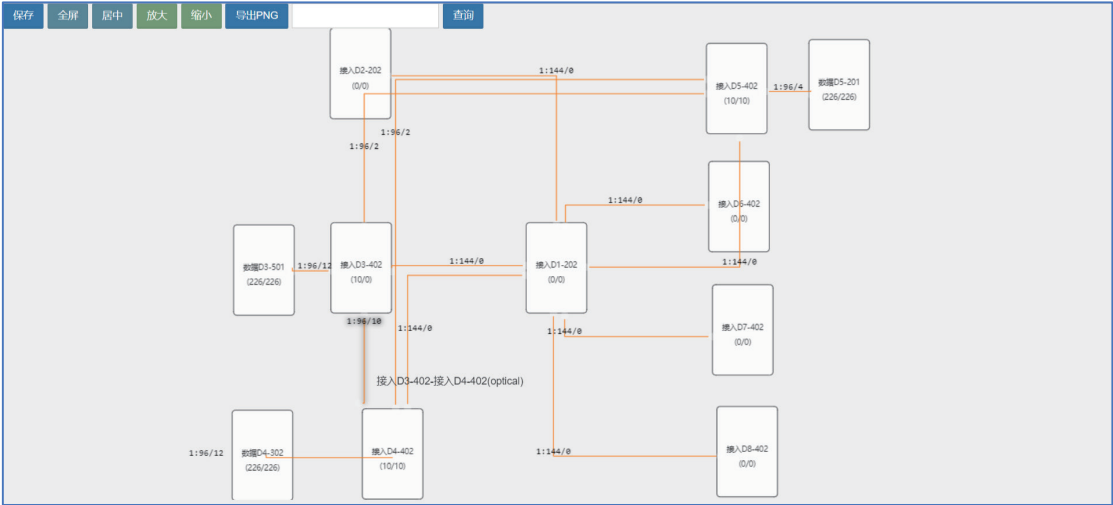


图 4.3.1：综合布线图

的连接走向，宏观了解光缆连接情况。以图形方式展示机楼与机楼间、机房与机房间的光缆连接；通过查看光缆可以查看纤芯使用情况，使用的客户及对应的定单，查看纤芯关联的链路信息；通过查看光缆的铺设信息，查看光缆的客户、业务信息，并进一步查看管道路由图、人井展开图、管道断面图，查看管孔 / 子孔占用信息。从而实现客户、业务、资源的关联分析，实现故障的快速定位和影响分析，便于业务快速开通和故障快速修复。

• 机房平面图

平台通过对机房平面图的管理，能真实展示机房的形状和空间使用情况。通过机房平面图，直观展示机架的使用情况和客户信息，可以进入机架正视图、ODF 正视图，进一步显示地板、机架、交换机、服务器等设备信息 / 客户信息 / 链接信息。实现日常运维的图形化管理。

如图 4.3.2，按照所见即所得的方式，真实展示机房实施情况，包括门柱、房门、地板，也展示相应的温度、湿度情况。对于机架机位，展示机架编号、机架所属客户、机架状态、超电情况等。

• 机架正视图

机架正视图直观展示机架上设备的放置情况

与业务状态信息，以图形化方式显示机架上各设备的排列情况及设备状态信息。显示机架订单号，可查看订单详情。显示选中 U 位的设备基本信息，显示设备名称、设备类别、所属客户、端子数、维护负责人、订单号、合同电量、当前电量、上游供电设备等信息。可关联显示选中设备的端子信息、关联显示选中设备的订单号，可查看订单详情。

如图 4.3.3，展示机架相关的信息，包括基本信息、客户信息、定单信息、供电情况、超电情况（图形展示）、设备安装情况等。

• 光路路由图

光路路由图是客户光纤租用的直观展示，通过路由图，客户可以直观的了解租用的资源信息。路由图显示光路的端到端详细路由，包括设备 / 端口信息、ODF 端子信息、纤芯信息，并可以继续点击钻取，可关联显示选中的资源信息，可关联显示选中链路的订单信息，可关联显示选中的光路的路由表信息。

4.4 运维管理

运维管理是数据中心日常的重要工作内容，工作琐碎却很重要，运维管理水平是衡量 IDC 等级的重要指标。主要从以下几个方面考虑：



图 4.3.2：机房平面图

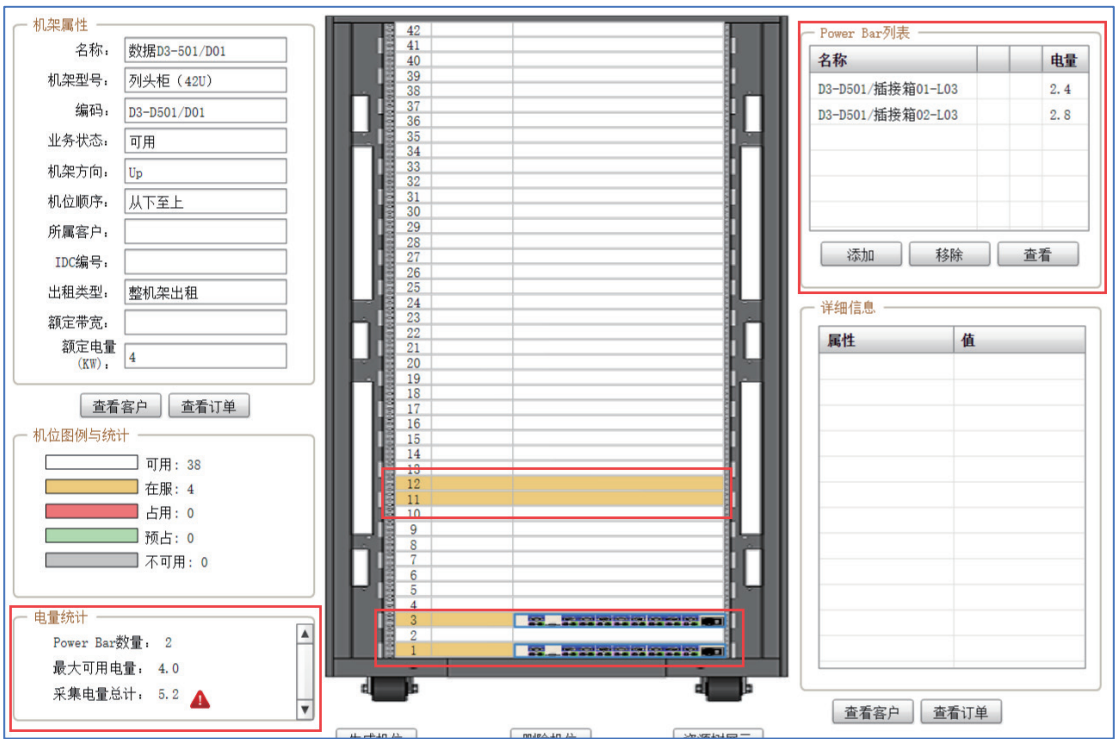


图 4.3.3：机架正视图

• 巡检管理

平台可以自定义巡检内容，设置巡检周期，两者结合生成相应的巡检任务，满足上交所巡检灵活性的要求。平台形成巡检任务后，通过 PAD 对巡检机房 / 设备进行二维码扫描，在 PAD 上即可填写巡检结果。完成后，提交巡检单，最终在平台中完成审批。支持在 PAD 上查看用户本人的超时未执行任务预警。

• 值班管理

值班管理不仅要满足日常的工作考勤，还要支撑日常的运维管理，包括节假日管理、休假管理、排班信息、排班日历、值班以及值班日志等功能。平台记录用户值班、换班、上岗、下岗等交接班的记录，包括未处理事务、已处理事务、新事务、巡检问题等的交接，并能对历史记录进行查询和管理，从而实现考勤与日常工作无缝结合。

• 进出场管理

设备和人员的进出，不仅关系到财产安全，

而且还与数据中心的安全运行密切相关。平台设置了严格的审批流程。在人员进出时，需要审核人员姓名、身份证件信息、所属客户、审批状态、允许入场时间、要求离场时间、入场原因、所属订单 / 工单、联系手机、联系邮箱等信息。在设备进出时，需要提交设备型号、数量、SN 码、所上机架、电源功耗等信息，并且需要经过机房管理人员审批才能进场。

五、回顾和总结

5.1 系统建设是基础，管理提升是关键

IT 支撑系统的使用，必然对传统的管理、作业模式带来变革与冲击，需要有一个逐步适应过程：

- 从原来的自由、随意的操作向规范的管理推进。
- 从原来的邮件模式向流程化、电子化管理迈进。

- 从原来的开放型管理向闭环管理改进。
- 从原来的粗放式管理向精细化管理演进。

随着平台的建设，对于原有的业务流程需要进行优化，对于原有的组织结构要进行完善，对于原有的考核办法要进行量化，从而提升整体的运营效率。没有管理的提升，平台必将成为日常工作的绊脚石。

5.2 流程建设是前提，数据准确是核心

IDC 运管平台中涉及到大量的流程建设，与业务开通相关的资源查勘流程、业务开通流程（包含开通、停闭、复通、测试等）等，按业务类型又可以分成机架、光缆、动力、空间等，细化到流程实现有近二十个。流程要以规范化、电子化、闭环管理为方向，又要以不增加工作人员日常工作量为目标，因此需要在流程的自动化上多做考虑。

IDC 运管平台能够正常运行的核心是数据准确性，必须给予数据三性（准确性、完整性、一致性）足够重视。在初期平台建设时，需要大量的基础数据，平台提供数据导入功能，实现批量数据入库，同时要加强入库数据的校验。在数据入库完成后，只要通过流程驱动才可以对数据进行修改，并定期开展数据的核查工作，确保客户、定单、资源三者关联一致。

5.3 整体规划是目标，分布建设是保障

通过公司整体规划，可以保证平台整体建设方向的正确性，可以保证周边系统功能边界清晰，可以保证数据流、业务流满足生产要求。整体规划，可以极大的避免系统的杂乱建设、重复建设、规划遗漏等情况。

在公司的整体规划中，对 IDC 运管平台的定位、功能和周边系统的关系等进行了明确。而

采用分步建设，一方面可以稳步推进，规避风险；另一方面，可以与流程管理和业务发展相协同，贴近生产管理现状需求。

六、 未来规划

IDC 运管平台已经成为日常生产和运维的重要系统，围绕平台融合、应用推进、效率提升的目标，主要体现在如下方面：

- 实现对业务开通的支撑，提高开通的及时率 / 准确率；
- 实现对资源的精细化管理，提高资源利用率；
- 实现统一的资源视图，支撑售前 / 售中 / 售后；
- 实现日常运维管理，提高运维管理水平；
- 实现外部系统支持，支撑财务、工程管理等。

为了更好的提升数据中心的管理效率，提高业务支撑能力，提升运维服务质量，结合公司的整体规划，IDC 运管平台的未来规划如下：

- 移动化：提供移动终端上的能力，实现移动化运营和运维。
- 位置化：提供基于位置的服务能力，包括资源定位、日常巡检、作业计划等。
- 自动化：提高自动化能力，包括基于 RFID 的数据自动采集、网管数据自动采集、自动巡检、资源自动分配等。

• IDC 门户：建立 IDC 门户，实现企业自服务管理，满足售前 / 售中 / 售后 / 客户自服务的需求。

• 外部应用：通过与 ERP 平台、运维监控平台、IDCISP 平台的对接，实现资产对接（ERP 平台）、效率提升（短信邮件、运维监控平台）、安全管理（IDCISP）。



E 实践探索 xploration

- 4 加速行情实现，提升金融交易能效
——基于 FPGA 的超低延时行情系统
- 5 构建在知识中台基础上的企业画像
- 6 财经快讯自动生成技术探索和实践
- 7 基于大数据的智能化客户服务平台
- 8 证通云外呼系统智能化实践

加速行情实现，提升金融交易能效

——基于FPGA的超低延时行情系统

赵月 / 上海金仕达软件科技有限公司 (yue.zhao2@kingstarfintech.com)

张强 / 上海金仕达软件科技有限公司 (qiang.zhang3@kingstarfintech.com)

连成浩 / 上海金仕达软件科技有限公司 (chenghao.lian@kingstarfintech.com)

赵健榆 / 上海金仕达软件科技有限公司 (jianyu.zhao@kingstarfintech.com)



金融量化人才大批涌现、量化选股与多因子策略的国内落地、股指与两融业务相继推出……进入中国才短短十余年的量化投资发展迅速。当前，国内量化体系正处于从低风险量化转向主动量化的行业变革期：大资管行业变革；个股期权、股指期货等金融工具接连推出；交易所技术系统不断升级等。上述变化的发生离不开对应硬件（系统）与软件（人才）的双重升级。硬件方面，拥有更低延时通道、更灵敏的系统响应、更稳定的交易系统是量化策略快速计算、快速运行的必备条件。

由于摩尔定律陷入瓶颈，以及行业日新月异的发展需求，FPGA（Field Programmable Gate Array，即现场可编程门阵列）因其自身独特的可编程、高并行、高带宽等特性优势进入大众视野，并在人工智能、云计算、通信、医疗、航空航天、自动化控制等领域的各应用场景中得到了广泛应用。

金仕达结合国内外金融环境与技术，以基于 FPGA 实现超低延时的沪深 Level 2 行情为切入点，进而深度思考和挖掘 FPGA 在金融交易领域的研究方向。本文通过介绍基于 FPGA 研发的超低延时行情系统，对 FPGA 硬件设计与实现、延时性能等内容进行深入剖析，希望有助于交易团队、科技公司等相关方进一步探索 FPGA 的研究和应用，进而提升交易服务质量、改善投资环境。

一、概述

1.1 市场的低延时行情需求

作为交易环节的重要组成部分，行情揭示了市场细微实时的变化，更快更精准地获取行情信息则意味着在“起点”快人一步。因此，建设超低延时行情成为提升整体交易速度的必经之路。随着市场基础设施的高速建设与发展，量化交易、算法交易、高频交易等多种自动化投资工具在交易行为中得到了有效的应用。相较于传统的人工盯盘、手动下单交易方式，自动化交易执行拥有更高的处理能力与更快的响应速度，更易在瞬息万变的市场行情中快人一步捕获交易机会并完成交易。在此背景下，低延时的交易数据获取及系统交互广受市场参与者的追捧。低延时通道不仅成为服务商差异化竞争的“法宝”，也成为交易团队在竞争加剧、优胜劣汰的市场中获得更为丰富投资回报的利器。一场围绕构筑超低延时系统的无形军备竞赛正如火如荼地进行。

1.2 FPGA 建设超低延时行情

FPGA（Field Programmable Gate Array，即现场可编程门阵列）是一种可以通过硬件编程语言改变其内部连接结构和逻辑单元，从而可实现自定义功能的芯片，可灵活编程是其与传统 ASIC（专用集成芯片）的显著区别。作为硬件体系结构，FPGA 具备硬件的运行速度和可靠性等优势，尤其是在逻辑/算数运算定制、高并行、高带宽等方面表现优异。由 FPGA 硬件电路固定延时代替系统软件延时抖动，不仅实现了低延时，同时提高金融交易系统的延时稳定性。

国外对冲基金已将 FPGA 硬件加速技术深度植入到交易环节的多个部分。对于国内而言，FPGA 产品在期货行业（如行情、交易柜台、网关风控等多个场景）中的落地为证券行业的低延时建设思路提供了可靠的技术与产品依据。基于在期货行业技术与业务深度积累，我司 FPGA 团

队将 FPGA 低延时特性应用在沪深行情的解析中，开发出具有超低延时特性的行情系统。

二、FPGA 硬件行情加速的设计与实现

2.1 系统整体架构设计及设备选型

FPGA 作为硬件芯片在对数据进行并行计算处理（流水线并行和数据并行）方面具有得天独厚的优势；另一方面，基于 CPU 的软件程序在系统交互实现的便捷性优势亦无需多言。故而，我们采用“FPGA（RTL）+ 软件”的异构架构来满足前述的市场需求——通过 FPGA 来实现 TCP/IP 协议栈，以及以太网 MAC 控制器，将原本属于主 CPU 的这部分工作由额外的硬件电路完成，从而减轻 CPU 的负担，加速网络处理的能力，提升系统的性能。

具体而言由 FPGA 硬件部分实现沪深交易所行情数据解析的核心环节，软件部分实现与周边系统交互的管理功能，包括与交易所行情网关（MDGW/VDE）的登录连接管理、行情用户的登录连接管理，以及系统监控等功能。FPGA 板卡与服务器之间通过 PCIe 接口进行数据交互。为实现行情解析全链路低延时的性能方案，我们采用自带网络接口的 FPGA 板卡，在 FPGA 内部完成行情的网络协议解析/重组，实现行情信息解析的端到端收发以及处理。系统整体架构如图 1 所示。

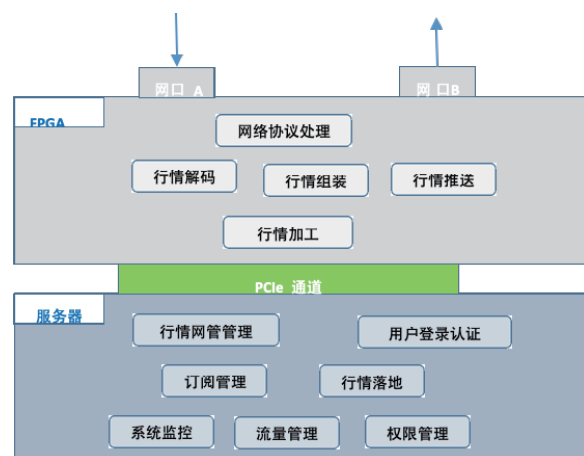


图 1

基于沪深行情业务，并且考虑到业务的实际应用场景，该 FPGA 板卡需要支持如下特性：(1) 万兆以太网接口两个及以上，该接口用于接收交易所实时行情源，组播行情推送等功能。(2) PCIe Gen3x8 或 PCIe Gen3x16，并向下兼容 Gen1/2。(3) 支持部分可重构（partial configuration），实现热加载硬件 FPGA 版（实现 FPGA 版本更新升级与软件版本更新升级相同便捷度）。

2.2 FPGA 硬件行情加速架构

采用自带 FPGA 芯片的硬件单板，对原始行情数据进行独立的纯硬件解析处理，实现行情接收、行情解码、行情组播等全链路功能。

如图 2 所示，交易所的原始 TCP 行情数据从万兆以太网口 A 进入 FPGA，通过 RTL 级设计的行情处理电路在其内部完成行情解析，最终将解码后行情以 UDP 组播方式以万兆以太网口

B 发出。

硬件行情处理按处理流程可划分为以下几个部分，如图 3 所示。

1) 万兆以太网接入层：实现从光纤上的串行比特数据流中识别出有效数据前导码，将串行数据经过 CRC 校验后，转换成并行数据接口的标准以太网包结构。

2) TCP/IP 协议解析 / 过滤层：将从万兆以太网接口接收到的包结构按以太网协议格式解析，根据 IP 字段，TCP 字段信息过滤非行情数据包，完成 TCP 保序并提取行情数据包中有效的行情源信息字段，发往行情解析层。

3) 行情解析层：根据交易所（沪、深分别处理）行情传输协议进行解码，解析出所需的各个行情数据字段，发往 UDP 协议组件层。

4) UDP 协议组建层：将解码后的行情数据，按以太网标准协议组建 UDP 组播包，如图 3，发往万兆以太网接出层。

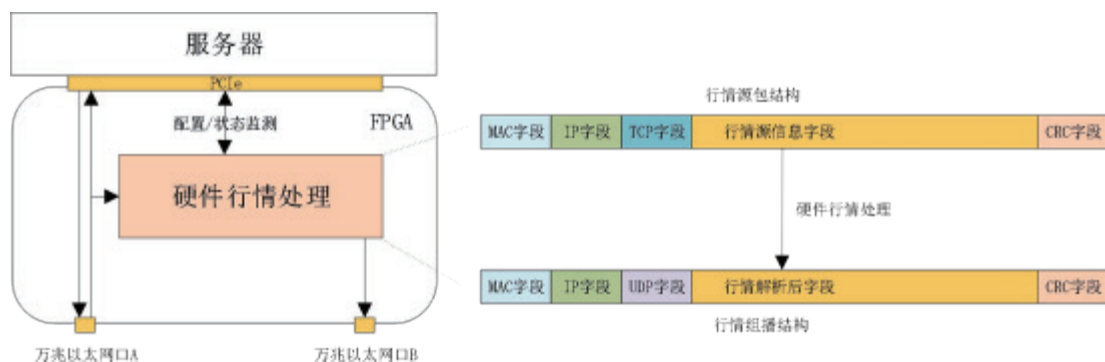


图 2

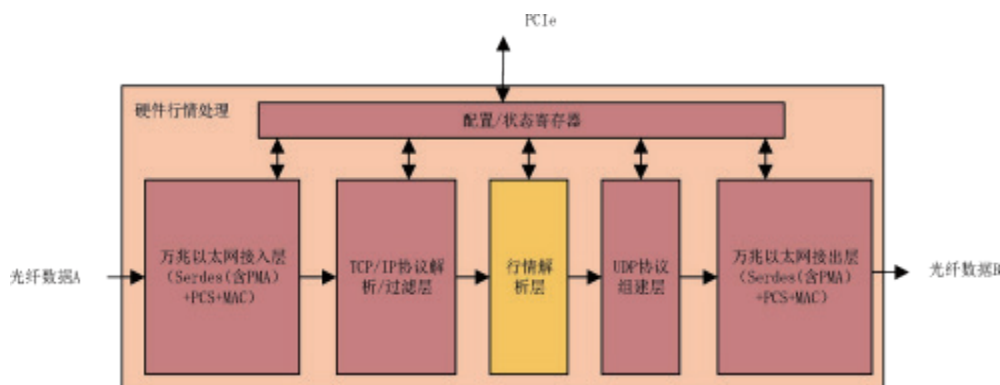


图 3

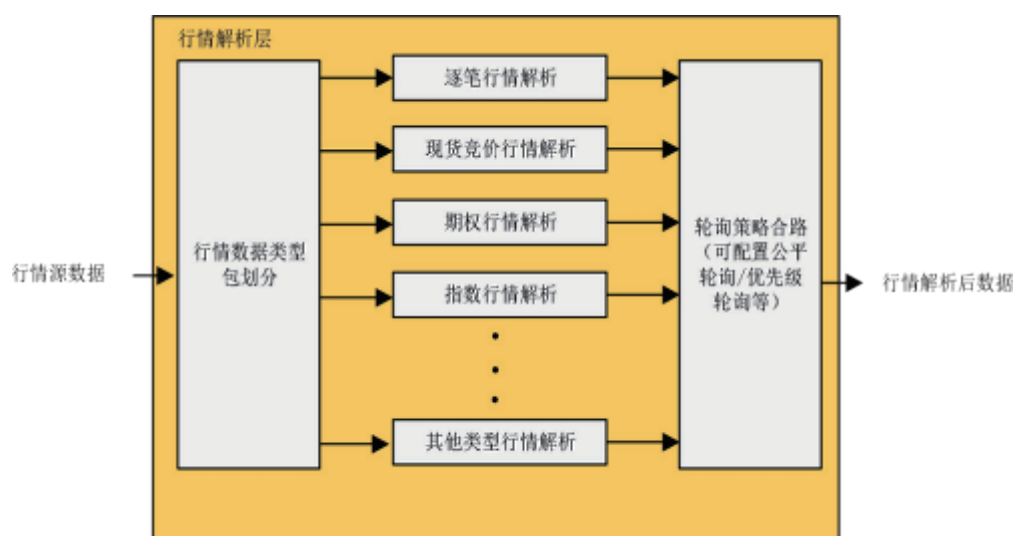


图 4

5) 万兆以太网接出层：实现与接入层相反的功能，将 UDP 组播包运算出 CRC 字段，并实现将并行包结构数据接口转成串行码流发往光纤。

6) 配置寄存器：该模块通过 PCIe 接口与服务器 HOST 交互。例如：HOST 将行情源的 IP/PORT 信息配置到 FPGA 寄存器中，TCP/IP 协议解析/过滤层获取对应寄存器的值，实现对非行情源包的过滤。同时，也会将 UDP 组播包 IP/PORT 信息配置到 FPGA 寄存器中，实现 UDP 协议包组建。

7) 状态寄存器：为了实现服务器对 FPGA 运行状态监控，各层模块开放一系列内部运行状态寄存器供行情系统软件部分读取，实现行情系统软件部分对硬件监控并做微秒级重启复位等异常后处理。

2.3 行情协议解析功能实现

硬件行情加速架构可适用于多种金融应用场景。针对沪深行情而言，行情源中涉及到多种业务（股票/基金/债券/期权/指数等）行情信息。为了实现并行化，提高行情业务数据处理带宽，行情解析层可以采用高并行、业务高扩展结构，

如图 4 所示（以出口为单个以太网口进行多路轮询合为一路为例）。

同时，针对不同业务的重要性以及速度优先级，可以通过软件灵活置寄存器接口，实现对不同业务数据包的发送优先级。

2.4 行情性能优化

2.4.1 数据压缩传输

交易所提供的行情源信息字段中大量数据采用 Unit64 表示，而某些字段实际业务中存在上限值，因此，可以用更短位宽进行表示。例如：交易所规定单笔订单数量最大不超过一百万。数据压缩：（1）可以提高网络报文传输效率，大大减小客户交易端接收行情的延时（通过降低数据长度的方式来调整输出字段格式，将 UDP 发送长度缩减 30%~70%）；（2）可以优化数据存储结构，在上交所全量恢复等大量存储访问场景中，减少存储访问次数能够有效地降低快照行情处理延时。

2.4.2 数据全内部 RAM 存储

行情解析中涉及部分场景需要存储全量或者历史行情信息，如果对位于延时关键路径上的数据处理采用 FPGA 外挂的 DDR 存储器进行存储，

其访问周期远大于 FPGA 内部的 BLOCK RAM (BRAM) 以及分布式 RAM。FPGA 内部的存储器能在单时钟周期完成读写, 而 DDR 存储器则需要数十个时钟周期延时。在上交所快照全量恢复场景中, 延时关键路径上的数据存储采用全 BRAM 的方式, 这样可以极大降低快照解析延时。

2.4.3 多时钟域方案

对于硬件数字电路而言, 时钟频率越高, 相同电路运算延时越低。在解析字段长度不确定等复杂协议 (如上交所 Fast 协议) 时, 除了在设计解析逻辑时通过提高并行度等方法降低延时以外, 还可以对该模块使用更高频率的时钟, 使得其运算速率达到最佳, 减小关键路径延时。

2.5 性能验证

2.5.1 延时性能

• 验证测试方法 (见图 5)

本延时测试方案是测量从行情源发出时刻 T0 到客户接受到解析后行情时刻 T1 的延时 (T0 与 T1 在同一台服务器上获取)。鉴于实时行情源 (MDGW 或 VDE) 连接限制的特殊性, 故采用行情回放的方式可以有效记录行情源发出时刻 T0, 从而进行验证测试。为了有效模拟真实行情系统上下游交互场景, 测试需两台服务器, 如图 5 所示, 服务器二插上 FPGA 板卡, 服务器一通过行情回放程序进行行情播放, 原始行情通过光纤进入服务器二上 FPAG 板卡的万兆以太网口

A, FPGA 解析后行情在万兆以太网口 B, 通过万兆光纤回传到服务器一的接收网口, 行情接收程序接收到该解析后行情, 并落地该行情信息以及时间戳 T1。同步, 服务器一在发出原始行情时, 需要另一种软件行情解析软件程序也对原始行情进行解析处理, 以便得到解析后标准行情数据 (以便与 FPGA 解析后行情进行逐条对应) 以及其发送时间 T0, 以此用一一对应的方式对硬件解析后行情进行延时测量。

具体步骤测试如下:

Step 1——获取回放源: 在生产环境中使用网络抓包工具录制一段原始行情数据, 即从 MDGW/VDE 中发出的网络数据包, 生成行情 pcap 包 (TCP 协议包)。

Step 2——发送行情并获取发送时间: 由行情回放程序对 pcap 包进行读取, 将该包向网口发出, 在发出该 TCP 包的同时打上时间戳, 该时间戳即为该 TCP 包中所有行情数据的起始时间, 记为 T0。

Step 3——获取标准数据源: 为了对硬件解码行情进行延时测量, 行情回放程序同时将对 pcap 源通过另一种软件行情解析程序 (位于服务器一上) 进行解析, 获得逐条解析后行情数据。对每条行情数据对应上发送该 TCP 包的时间戳 T0。

Step 4——获取 FPGA 解析后行情, 并计算延时: FPGA 板卡接收到原始行情 TCP 协议

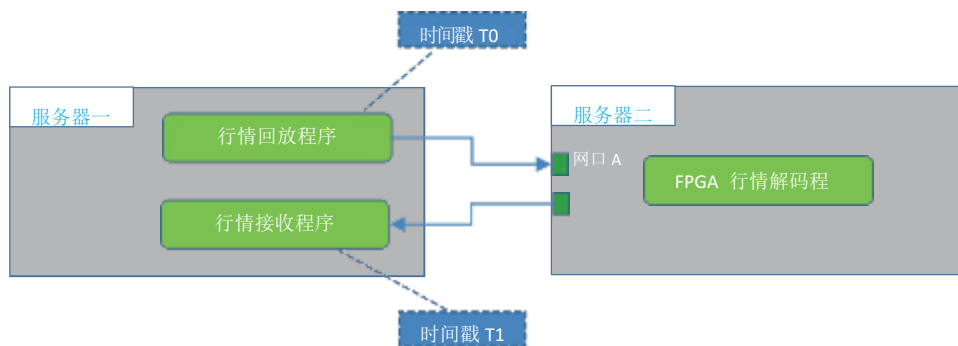


图 5

包进行解析处理，并将解码后的行情数据通过网口 B 发往行情接收分析程序。行情接收分析程序在接收到行情数据后逐条记录行情数据到达时间，记为 T1，时间戳使用 clock_gettime 获取，并形成落地文件。行情解析穿透延时即为 $T=T1-T0$ 。

• 测试结果数据

表 1

交易所	FPGA 仿真延时
上交所-快照	800ns~1.2μs

注：因上交所逐笔委托 Fast 解析累积机制，因包的笔数不同会造成延时数据波动性较大，无法规律统计，这里仅展示快照数据。

2.5.2 吞吐量性能

• 验证测试方法

同延时性能测试方式，通过调整行情回放程序回放速度倍数来实现吞吐量压力测试，测试过程中观察分析行情服务器运行状态以及解析后行情数据，以获取行情服务器的最大吞吐负载。

• 测试结果数据（见表 2）

2.5.3 数据准确性 / 完整性

• 验证测试方法

为验证 FPGA 行情系统解析数据的正确性以及完整性，将 FPGA 行情系统落地的数据与目前在市面上广泛运用的另一套行情系统解析结果进行对比测试，测试部署方案如图 6 所示。

具体测试步骤如下：

Step 1——服务器 A、服务器 B 上分别启动 FPGA 行情系统和另一套行情系统，而且两套系统对应的数据接收落地程序、落地格式保持一致。

Step 2——行情网关发出的数据通过分光器发往两台服务器，以保证两套行情系统解析的数据源一致。

Step 3——编写行情数据对比程序，逐条数据每个数据项一一比对数据正确性，同时判断是否有数据条丢失。输入：两套系统的落地数据，输出：每支股票每条数据的对比结果。

• 测试结果数据

经过长期稳定的测试，根据比对程序输出结

表 2

回放速度	系统状态	行情解析状态
1-10 倍回放速率	抗压正常运行	数据完整且准确
10-40 倍回放速率	抗压运行 10min 左右	抗压运行期间数据完整且准确
>40 倍回放速率	抗压运行 30s 左右	抗压运行期间数据略有丢失

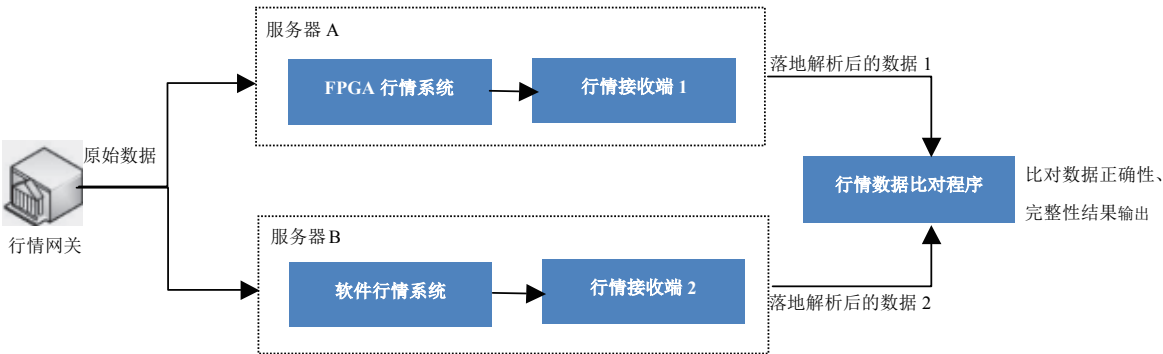


图 6

果，FPGA 行情系统解析数据稳定无丢失，且数据项均正确。

三、结语——未来 FPGA 应用展望

FPGA 硬件加速技术在交易环节的技术突破需要充分发挥其低延时、高并行、可扩展性、分布式等优点。目前，FPGA 硬件行情产品的商业落地及稳定运行不仅得到了客户的高度认可，同时也为其在证券业务其他场景中的应用潜力提供了有力支持。在此趋势下，金仕达结合市场需求在低延时交易柜台领域进行了尝试与实践，现已

实现基于 FPGA 异构的证券低延时交易柜台（柜台稳定穿透延时 $3\mu\text{s}$ ）。

未来量化策略的发展将呈现多元化、复杂化，投资者机构化，以及海外机构加速布局。出于对提升市场有效性的迫切需要，大资管行业将加速升级技术系统，FPGA 作为低延时赛道的核心技术也因此将成为金融生态系统的关键。鉴于此，金仕达将积极推动高频研究创新中心、仿真态 FPGA 实验中心、FPGA 金融行业云中心的创立，并且将联合知名 FPGA 厂商及芯片制造商，打造专用高性能硬件设备，为客户提供长期稳定可靠的硬件设施。

构建在知识中台基础上的企业画像

齐伟 / 恒生电子股份有限公司 咨询事业部 qiwei@hundsun.com



现状

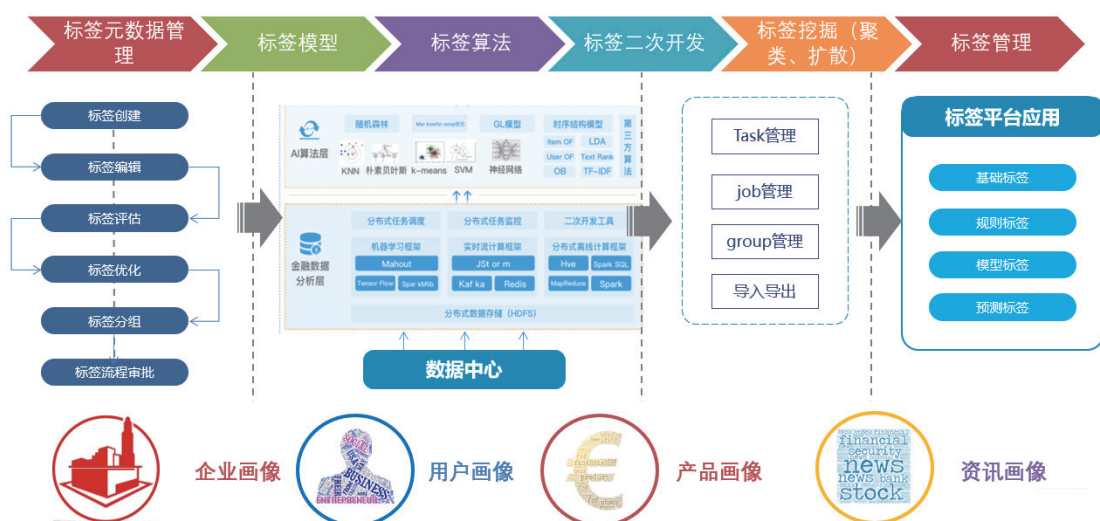
金融行业是信息驱动的行业，金融机构对用户的营销服务、适当性管理催生了用户画像业务；对投资标的的投研风控的分析、资产配置的要求，催生了产品画像业务；对新闻资讯的精准搜索、个性推荐需求，催生了资讯画像业务。无论用户画像、产品画像、资讯画像，其依赖的数据主要是金融机构内部的经营数据，包括用户在金融机构端的行为数据、账户数据、交易数据、行情数据等等。这些数据基本上都是结构化的，数据价值密度高，形成一个有效果的画像一般只会面临工程层面的问题。

当前，金融机构对画像业务的关注点正在转向企业画像，一方面这源于大数据投研风控的需

求，对于股债基等等投资标的基本面研究，需要从投资标的对应的企业的分析入手，甚至于孤立的企业分析也不太够用，还需要从企业所处的产业角度去发现一个公司的价值与风险。另外一方面，企业也是我们金融机构的客户，VC、PE、企业投融资等金融业务的开展，需要企业尽调、合规监管的审查，对于企业客户的营销服务也需要对企业客户做画像。无论是基于产业的投研风控分析还是对企业客户做金融业务的营销服务，都会涉猎到大量非上市公司。企业画像的难点就在于对这些非上市公司的打标以及关联分析上。

传统方法

基于标签的画像方法是当前画像系统的主流



技术，给企业打一组标签，每个标签给一个权重，把这些信息整合在一个工作界面上构成企业全息信息视图，同时提供检索、分析、报表等功能，一个画像系统就算完成了。这类软件或者平台，有个共同技术支撑平台：一个是画像门户，一个是打标平台。

画像门户是数据展示的端，天眼查、企查查就是这类软件。打标平台则是其核心后台技术。打标平台是一个一站式解决方案，方案涉及标签元数据管理、标签模型、算法的支撑、标签的二次开发、标签挖掘以及标签管理等内容。标签作为画像系统的核心服务，又分为基础标签、规则标签、模型标签、预测标签几个维度。基础标签比拼的数据源的多少与质量；规则标签是对基础标签的规则化处理，主要靠专家来制定规则，泛化能力有限；模型标签，是通过机器学习建模，对某类标签进行有监督或者无监督的计算方法，比如通过聚类方法实现用户的自动分类；预测标签是对某种业务属性的一种预测，诸如商机预测、风险预警都属于这类标签。相对于模型标签的模式识别，预测标签则代表对未来某种潜在标签达成的概率预测。模型标签和预测标签是比较有技术含量的标签体系，是引入大数据或者大数据处

理方法的一种泛化。

问题

大数据多源、异构、数据价值密度低、数据质量参差不齐，基于这样的数据环境，企业画像面临诸多问题与挑战，与依托金融机构内部数据的企业画像，有着本质不同。企业画像面临的问题有数据问题、业务问题和技术问题，接下来我们逐一详解。

数据问题——数据治理的问题

从工程实践的结果上反过来分析，金融机构关注的企业画像所需信息包括几个维度：1) 企业基础信息，包括董监高行业地域规模等等；2) 股权层面，质押、担保、对外投资、实控人、受益人、资本系等；3) 产业层面：主营产品、细分行业、供应关系以及上下游等相关信息；3) 新闻舆情：包括新闻资讯、舆情传播、事件影响等；4) 另类数据包括司法诉讼、专利、行政处罚、政策法规、监管问询等等。

上面所列数据从数据来源来说难点主要两个方面，一个难点是产业链数据，申万 28 个行业，

细分子行业是 2+N 的层级，加上国际统计局的 GIS 产品标准和类目，随便一个行业其涉及到的行业细分产品的数目可能就会达到万级，关键这些数据还是动态更新的，需要按需修订。另一个难点是舆情画像，我们知道做舆情的企业动则监控几万的站点，几十台上百台的服务器日夜无休的爬取各类数据。还要随时面对网站改版、安全策略更新等突发情况。

以上这些数据基本来源于工商数据、行业数据、政府公开数据以及其他另类数据，相对于金融机构内部数据，可以说是真正的多源异构，如何将这些信息有效的关联起来本身是一项非常有挑战的工作。拿产业链数据来说，市面上做产业链的公司几十家应该是有的，而且因为其行业分析师的行业出身不同，导致每家只做了 1、2 条实际可用的产品级产业链，每家数据标准、业务标准都不统一，将这类数据集成在一起的时候，如果用数据中心的统一资讯的方法，会非常灾难，那些产品类目的对齐就是一项难以完成的工作。标准化容易工程落地难！

业务问题——知识转化与沉淀的问题

企业画像有待解决的业务问题主要体现在企业大数据与企业金融分析之间的业务鸿沟。以股权数据为例，我们可以轻易的从工商数据中获取到相关股权数据，经过简单加工处理，也可以获取一份清晰的股权链路图，即我们通常所说的股权图谱。很多工商数据的厂商对股权数据加工基本上止步于此。但金融行业对股权的加工是有明确要求和目标的，“人行 235 号文”对基于股权数据的疑似受益人、实际控制人在关联人、股权比例、企业类别方面有明确要求，很显然简单的股权链路无法满足金融行业的要求，我们需要按照金融场景来定制处理。这就带来了面向金融行业的企业画像的一个业务难题，就是什么样的标签和关联分析是符合金融场景需求的？这不仅仅是大数据处理的问题，也涉及到大量的金融专业知

识，所以，金融行业的企业画像是很有必要引入专家经验的，所以人机如何协同？专家如何将金融知识落地沉淀到企业画像系统，也是画像系统要解决的问题。

技术问题——算法模型的天花板

企业画像因为没有金融机构内部行为数据的支撑（至少是冷启动建设开始的事实是这样），所以画像标签面临不完整、不准确的问题。标签不完整是源于很多非上市公司数据来源非常有限，极端情况下我们只能拿到其工商数据，那么我们定义的画像各种数据维度的槽，是无法填充掉的，传统打标平台无法解决这块信息的补全。标签的不准确表现的为缺乏背景知识的语义理解不准确，举个例子，如果某篇资讯通篇在提光缆、通信设备、基站以及车联网相关的话题，那么我们在给与上述实体标签基础上，最应该打上还有个“5G”标签，即便文章中并未提到 5g，但事实上，文章内容是对 5g 产业链的一个综合性描述，而机器（打标软件）如何知道 5g 跟那些关键词的对应关系的呢？这就需要我们赋予“机器”一个 5g 产业链的背景知识。很显然，单纯的机器学习模型哪怕是预测模型也是解决不了这类标签准确性问题的。

对于一份数据稀疏的企业大数据，用传统机器学习方法打标，很快还会遇到一个算法的天花板。当下工程界大家的算法和调参基本处在同一个起跑线，语料是决定算法结果质量的关键，企业画像数据过于稀疏、数据价值密度低，使得我们从显性数据里获取的语料非常有限，好在我们知道在显性数据背后还有很多隐性数据关联，类似实控人这类标签业务，就是有效关联带来的成果。那么问题是，我们如何发现这些隐性关联，以及这些关联后的新知识呢？

应对

要想解决上述问题，我们需要引入知识图谱

技术，并且在工程实践中做出 3 个思维方式的改变：1) 基于过程的分析转向基于领域的分析，这主要应对的是企业画像的业务问题；2) 基于显性 ER 关系转向隐性关联关系的发现，这点主要应对的是数据问题；3) 基于 ER 数据的模型走向基于语义的模型，这点应对的是算法模型的天花板问题。

基于过程的分析转向基于领域的分析

我们现在做画像业务主要是两种模式：一个是数据驱动，意即有什么样的数据画什么样的画像；一个是场景驱动，就是有什么样的场景，找什么样的数据，再画什么样的画像。为了解决前述分析的企业画像系统的各种问题，我们需要引入第三种模式即“领域驱动”的模式。无论是数据驱动还是场景驱动，都是过程驱动，数据落地表现为一个个 ER 关系图，要么是从场景中找到对应的表结构，要么从现有数据 ER 结构推出新的画像 ER 结构。领域驱动要解决的问题则是要把 ER 关系转化为语义网络关系，把表与表之间的关键字连接，变成表中描述的实体与另外表中描述的实体建立语义上的连接。这种转变就是过程转化为领域的过程。我们看个具体的例子，我们可以将企业在产业中的信息用 ER 关系进行表达，于是会形成企业信息表、企业关联行业表、企业主营产品表，如果我们寻找到新的数据源，比如企业供应关系，那么增加一张表就解决问题了。这样一个 ER 关系如果转化成领域图，那么就是一个产业链的知识图谱，这张图里有企业、行业、主营产品、供应关系、上下游等数据结构，所有数据表达采用统一的模式图“三元组”，形成的就是产业链知识图谱。面向过程的 ER 图的使用，需要人来找数据用数据，而面向领域的产业链知识图谱，既可以作为行业分析人员的背景知识，更重要的是，基于三元组的语义网络的数据组织形式，是可以让计算机理解产业链数据，进而可以产生机器认知的某种能力，诸如知识推理、

知识补全等等，这不是面向过程的数据组织方式可以提供的能力。

基于显性 ER 关系转向隐性关联关系的发现

当我们的思考角度能够从基于过程的分析转向为基于领域的分析之后，我们就有能力从显性 ER 里发现隐性关联关系。这种转变，对于企业画像来说，意义重大。1) 我们具有了发现新数据/新知识的一种能力，就像我们在股权关联里隐藏了实控人，产业链里隐含了行研逻辑一样；2) 我们自动有了新的标签。对于标签系统来说，知识图谱的 schema 就是一个丰富的标签库，schema 描述的实体间的关系，以及实体属性、关系属性，本质上都是对某个实体的标签描述，甚至于 1 度、2 度关系也可以是某种标签描述；3) 通过实体在知识图谱关系网络上的漫游，我们可以为标签理解增加了大量的背景知识的同时，也为标签系统提供了语料数据，为机器学习增加了大量的可选择特征和样本数据；4) 通过图分析、推理和挖掘，我们可以让计算机来参与到标签发现、知识发现和数据关联发现的工作中来。

基于 ER 数据的模型走向基于语义的模型

我们将画像业务从基于过程的分析转变为领域分析后，需要借助知识图谱做工程落地，有了知识图谱，我们就可以利用知识推理、挖掘、发现、关联，从显性 ER 关系获取更多的隐性数据关联。也可以将专家逻辑跟 AI 逻辑做统一整合。进而，企业画像系统依赖的机器学习标签、预测标签，可以从基于 ER 数据的模型，走向基于语义的模型。机器学习可以理解为聪明的数理统计模式识别的方法和工具，而知识图谱可以赋予机器学习大量背景知识，那么一个既聪明又有学识的 AI 是否能更加胜任我们的企业画像业务呢？

ER 模型走向语义模型我们可以有几条路径去驱动：1) 特征支持，图谱的 schema 是特征的源，相对于实体属性关系网络特征应该是区分度

最好的特征来源之一；2) 语料样本数据支持，图谱的三元组数据是一个个事实，隐性关联的知识也是这样的一个个事实，从而对于标签理解或者机器学习标签可以提供语料或者样本数据；3) 图分析的一整套方法可以补充进机器学习标签模型里，丰富算法空间。4) 基于知识图谱的表示学习可以和深度学习结合，完成更加有挑战的标签数据的建设。

架构

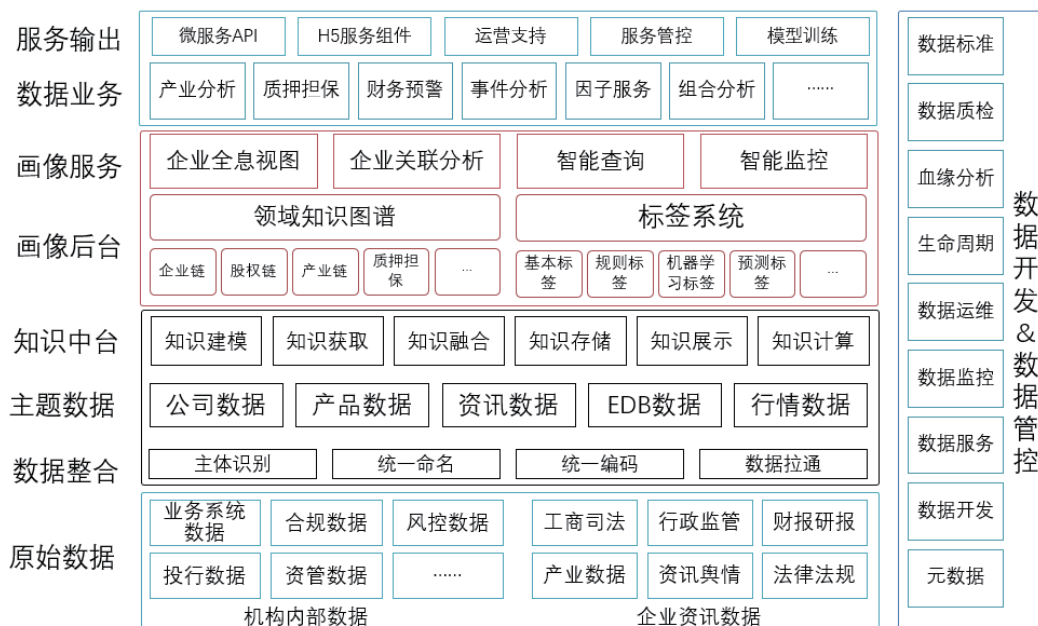
从工程实践的角度，我总结了如下一个企业画像的架构图。

架构要点有两个：

第一点：引入知识中台概念，建立“数据中台+知识中台”双中台驱动。体现在1) 数据->信息->知识->智能，数据中台主要定位在“数据->信息”这个环节，知识中台定位在“信息->知识”这个环节；2) 知识中台是知识相关、事实相关、关系相关的信息的再组织，它无法取代数据中台对应的数据中心，当然，如果你处理数据

的主要目的就是找关系，那么也可以绕开数据中台直接建设知识中台，但大多数情况下，数据中台和知识中台是共生的，他们合在一起是更加广义的数据中台。数据中台为知识中台提供大数据支撑，知识中台为数据中台赋能 AI 认知能力；3) 画像系统以及其他业务系统按需从数据中台、知识中台或者原始数据中获取数据，他们不是多选一，而是多方协作；

第二点：知识图谱结合标签系统共同为画像系统服务。打标系统为知识图谱赋能，体现在1) 指标进入知识图谱，为知识图谱中的实体和关系增加属性特征；2) 指标进入图谱，为知识图谱关联分析提供边际变化，比如事件标签进入图谱后形成事件图谱，也可以产生事件传播影响的分析业务。知识图谱为打标系统赋能，体现在1) 知识图谱为打标系统提供背景知识，增强打标准确性；2) 知识图谱的 schema 实体关系属性本身输出为打标系统的标签；3) 知识图谱为打标系统提供知识与事实，作为语料和标注数据；4) 知识图谱的隐性关联关系又会产生新的标签。



财经快讯自动生成技术探索和实践

邱震宇 / 华泰证券股份有限公司 qiuzhenyu@htsc.com

张弛 / 华泰证券股份有限公司 zhangchi015258@htsc.com

刘园 / 华泰证券股份有限公司 liuyuan@htsc.com

孔亚洲 / 华泰证券股份有限公司 kongyazhou@htsc.com

朱德伟 / 华泰证券股份有限公司 zhudewei@htsc.com

肖驰 / 华泰证券股份有限公司 xiaochi@htsc.com



证券研究业务涉及到大量研究报告的撰写工作，例如针对上市公司定期报告的财经点评快讯。通常，这类财经快讯的撰写需要大量信息整理工作。这些工作较为繁琐且重复性较高，在一定程度上限制了研究人员的研究产出。我们研究发现，这些财经快讯通常包含基础财务指标描述、主营业务相关情况描述和异动财务与经营指标及其原因等基础元素。一方面，这些基础元素均可利用技术手段从上市公司的公告文件中自动解析出来。另一方面，这类财经快讯大多属于客观陈述，一般不涉及复杂的深度观点剖析。基于上述分析，我们借助自然语言处理技术设计了一套包含公告文件解析、财务信息抽取和快讯生成的财经快讯自动生成解决方案，该解决方案已经实际应用到华泰证券自研的金融舆情事件服务系统中。

1 研究背景

在证券研究业务中,研究人员时常要根据上市公司发布的定期报告,撰写公司财经点评快讯。以华泰证券公司发布的部分财经点评报告为例,公司 2019 年对外发布中文研报 5558 篇,其中财经点评类报告 1160 篇,占比达 20.87%。这些财经快讯撰写工作较为机械繁琐,在一定程度上限制了研究人员的研究产出。如何自动地生成财经快讯,减少研究人员的工作量是当前研究的热点。

近年来,随着人工智能和自然语言处理技术的研究进展,自然语言生成技术 (Natural Language Generation, NLG) 技术展现出较强的实用价值和广阔的应用前景。自然语言生成问题要求模型在理解自然语言的基础上,自动地生成一段文本序列。根据模型输入输出的类型,自然语言生成任务可以进一步细分为多个不同的子任务。当输入数据为结构化的表格型数据时,我们称之为 Data-to-Text 任务。譬如,根据比赛数据,实时生成体育比赛快讯。当输入为长文本,输出为简练的短文本时,我们称之为摘要任务。譬如,根据输入的新闻报道,自动地生成新闻标题。

通过研究,我们发现财经快讯报告通常包含基础财务指标描述、主营业务相关情况描述和异动财务与经营指标及其原因等基础元素。这些财务基础元素均可利用自然语言处理技术从上市公司的公告文件中自动解析出来。另一方面,这类财经快讯报告大多属于客观陈述,一般不涉及复杂的深度观点剖析。基于上述分析,我们借助于自然语言处理技术设计了一套财经快讯自动生成解决方案。该方案包含上市公司公告文件解析、财务信息抽取和财经快讯生成等功能。为了解决训练数据集不足的问题,我们设计了一套基于冷启动的数据集构建方案。目前,该财经快讯自动生成解决方案已经实际应用到华泰证券自研的金融舆情事件服务系统中。

财经快讯自动生成可以产生以下价值:(1) 减

少研究人员搜集、整理和加工财经数据的重复性工作;(2) 使研究人员可以专注于深度观点剖析研究;(3) 该技术可以帮助研究人员快速应对政策变化,适应研究标的覆盖面的快速外延。

2 技术方案

在这一节中,我们将详细介绍财经快讯自动生成问题面临的难点和相应的解决方案。首先,我们先给出财经快讯自动生成的问题描述。然后,我们将展示解决方案的整体框架。最后,我们将详细地介绍每一个模块的设计思路和运行细节。

2.1 问题描述

财经快讯自动生成问题希望从上市公司的定期报告中自动识别出财务指标描述、异动财务与经营指标等元素,并根据这些元素生成精炼的财经快讯。假设给定一篇 PDF 文档 D , 财务指标集合 $A=\{a_1, a_2, \dots, a_m\}$ 。模型需要先将文档转化为结构化数据 $S_D=\{T_{\text{text}}, T_{\text{table}}\}$, 其中, T_{text} 代表文本数据, T_{table} 代表表格数据。接着,模型需要根据财务指标集合抽取出关键指标信息 R_{index} 和关键文本信息 R_{text} 。最终,模型根据抽取出来的关键信息 R_{index} 和 R_{text} 生成最终的财经快讯文本 Y 。

2.2 整体框架

根据上述问题描述,我们设计的财经快讯自动生成方案主要分为三大模块:文件解析模块、信息抽取模块和文本生成模块。这三个模块以流水线方式串联执行,整体架构如图 1 所示。

模型的输入数据为上市公司发布的原始定期公告文件,包括年报、半年报、季度报等。这些公告文件大多属于 PDF 格式,它们无法直接应用到下游任务中。因此,模型需要先借助 PDF 解析技术对文本进行结构化处理,将这些公告文件转化为结构化的数据。处理得到的结构化数据除了包含基础财务数据,还存在大量冗余信息。模型

需要通过财务信息抽取模块抽取出有用的财务数据，这些数据包括基本财务指标相关数据、指标变动原因文本、主营业务介绍以及相关财务变动原因文本等。待抽取的数据既包含结构化的表格型数据，也有半结构化、非结构化的文本数据。因此，财务信息抽取模块综合了规则模板、序列标注模型和句法分析等多种方法来共同完成财务信息抽取任务。最后，模型将这些关键财务信息输入到财经快讯生成模块中进行快讯生成任务。在这个模块中，我们首先设计了一种基于冷启动的语料构建方法来生成必备的训练语料。然后，我们借助基于规则模板的生成方法和基于神经网络的序列生成模型来生成最终的财经快讯。接下来，我们将详细介绍上述三个基本模块的构造细节。

2.1 公告文件解析

本文模型处理的对象主要包括上市公司的定期财务报告，如年报、半年报、季报等公告文档。不同行业、不同公司、不同事件或不同时间发布的公告文档结构差异较大。因此，模型需要对这些 PDF 文件进行预处理，将其处理成统一的、方便处理的数据格式。模型借助于自然语言处理技术实现了定期公告文档智能解析，主要包含以下模块：

文档采集模块 该模块主要负责从指定数据源上实时获取上市公司的公告文档。数据源包括互联网、本地文件和数据库等。

文档解析模块 该模块主要负责对 PDF 公告文档进行粗粒度解析。解析得到的数据包括文档的结构、文本、表格和图片等。

数据提取模块 该模块对上一步得到的结果进行更细粒度地信息抽取处理。一方面，解析得到的文本大多以块的形式存在，模型需要把同一段落内的文本块进行合并处理。另一方面，模型设计了专用的卷积神经网络 (Convolutional Neural Network, CNN) 来提取图片中的表格数据。

数据后处理模块 该模块主要是对抽取得到的数据做进一步地格式化处理，包括金额数据归一化、时间单位转换等。

关联分析模块 该模块结合文档的篇章结构等目录信息，进一步将单一指标数据进行关联，包括指标明细关联、表格历史关联和指标历史关联等。

数据标注模块 该模块主要是支持人工对机器自动抽取的数据进行查看、校验和错误修改，帮助提升文档解析的质量。

数据存储模块 模型将解析得到的数据存到数据库中以备后用。

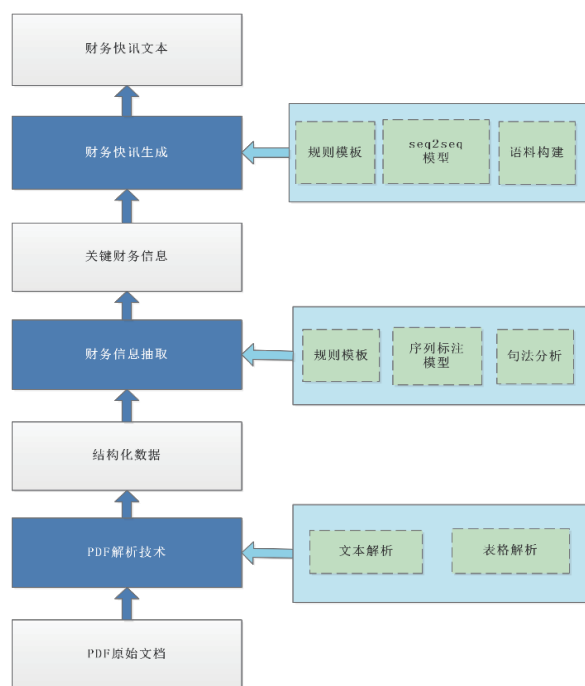


图 1：自动生成方案整体框架图

2.2 财务信息抽取

PDF 解析模块得到的数据不仅包含了基础的财务数据，还存在大量冗余信息。模型需要对这些数据进行信息抽取处理，抽取出关键的财务信息以传递到下游模块中。针对不同特点的财务数据，模型设计了专门的抽取方法进行处理。

2.2.1 基本财务指标数据和主营业务指标数据抽取

基本财务指标数据主要关注财务指标名称及其对应的变动百分比，如营业收入、归母净利润等。主营业务指标数据关注公司在所属的不同行业、输出的不同产品、不同地理区域上的营业状况。这两部分数据的获取较为简单，模型可以直接从解析得到的结构化数据中索引获取。

2.2.2 异动财务指标及原因抽取

异动财务指标数据主要指那些数据变动较为异常的财务指标，包括基本财务数据和主营业务数据。根据研究员提供的专业指导意见，我们针对不同指标设置了对应的异动阈值。若该指标数值变动超过了这个阈值，则该指标将被判定为异动指标。除此之外，我们还需要抽取出这些异动指标产生的原因。通常，指标异动原因描述会以两种形式存在于公告中：表格栏目数据和非结构化文本描述。对于前者，系统可以定位到表格相应的指标异动原因栏目，直接抽取。对于非结构化指标异动原因，系统需要结合机器学习相关技术进行文本抽取，主要包括数据集构建和机器判别两个部分。在项目初期，我们先从文本中抽取出存在异动指标的子句片段，并将其交由标注人员判别包含该指标的子句是否为该指标异动的原因。

接着，我们将指标异动的原因抽取任务抽象为异动文本匹配问题。具体地，我们将先前步骤得到的指标异动情况转化为一个短文本语句。例如：“营业收入”指标项对应的变动值为“50%”，我们将其转化为“营业收入相比同期增长了50%”。然后，我们以 BERT 模型为基础，构建一个二分类问题。模型将前述步骤拼接的异动指标情况描述以及候选原因子句作为输入，来判断该候选原因子句是否为该指标的异动原因描述。通过上述步骤，模型可以从文档中找出异动指标及其对应的异动原因。

2.3 财经快讯文本生成

通过前述步骤，我们已经获取了公司公告

文档中的基本财务指标的结构化数据、异动财务指标的结构化数据以及相关异动原因的非结构化数据。接下来就可以将这些信息整合起来生成一段完整的财经快讯。针对财经快讯生成任务训练数据量不足的问题，本模块首先设计了一种多样化构建语料库的方法。然后我们结合规则模板、Seq2Seq 序列生成技术来完成财经快讯文本自动生成任务。该模块的主要流程如图 2 所示。

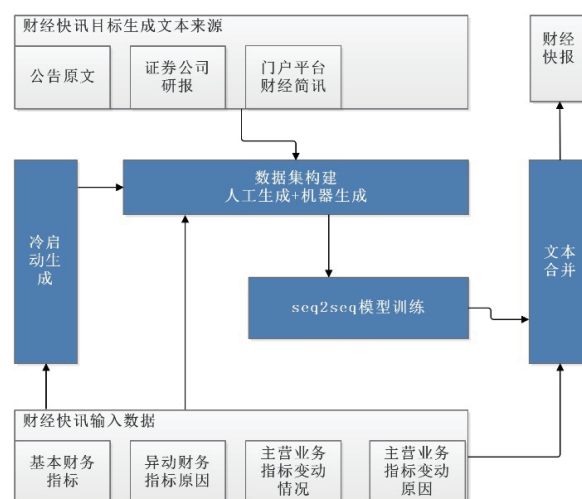


图 2：文本生成流程

2.3.1 数据集构建

数据集构建对深度学习模型落地非常关键。财经快讯生成数据集的构建难度较大。这主要体现在以下两个方面：(1) 目标文本整理难度较大。我们构建的目标文本来源于各家门户网站的人工点评文本以及证券研究员撰写的研究报告。然而，这些文本并非全部符合要求。我们需要专业的标注人员从中筛选出我们需要的内容；(2) 目标文本和输入数据匹配关系较为复杂。在查找到目标文本后，模型还需要将文本与模型的输入数据进行对齐，即需要将具体某个公司的某个公告与该文本进行匹配，保证数据的正确性和有效性。基于上述问题，我们设计了一个人工标注结合机器标注的数据集构建方式。具体流程如下：

收集原始语料 收集公告发布后 3 日内与该公告相关的研报点评，并将这些数据添加到候选语料数据集中；

人工标注语料 由专业人员对这些语料数据进行人工标注；

机器辅助标注 我们设计了一种基于相似句检索的机器标注方式。首先，对已有的指标情况使用简单短句形式进行拼接，对于异动指标原因则可以直接使用其文本。然后，使用开源中文分词工具 LAC 对文本进行分词，并使用中文词向量结合单词的 TFIDF 表示对句子进行编码。接着，使用向量余弦相似度来匹配原始文本中相似度高的文本，并添加到候选目标文本中。最后，我们根据文本在文中的位置将所有文本进行重排序，并由标注人员进行校验。虽然这种方式还是需要人力参与，但是标注工作量大大减小。

2.3.2 冷启动生成

项目运行初期，相关数据积累较少，使用监督训练的难度较大。因此，我们选择使用基于规则模板的方式来进行冷启动生成。根据财务指标情况描述，我们为每个指标设计一个带有槽位的描述模板。其中，槽位类型包含指标名称、指标当期数值和指标变动百分比等。模板实例如下：“{ 指标名称 } 亏损 { 当期数值 } 元，亏损同比增长 { 变动百分比 } %”。在生成时，将检索得到的指标值填入到对应的模板槽位中构造对应的财经快讯子句。另一方面，根据子句推荐度，直接从前述步骤构造的指标异动原因子句集合中，选择一个子句作为最终的原因描述。

通常来说，规则模板生成的财经快讯指标描述准确性较高，但是生成的文本样式和表达风格较为固定，缺少变化。另外，规则模板的维护和扩展也需要较高的人力成本。因此，我们研究了基于深度学习的自然语言生成技术。在实际业务中，基于深度学习的模型通常需要大量的标注数据。除了人工标注的数据外，我们的模型还将冷启动阶段获得的部分快讯作为一部分训练数据加入到数据集中，辅助训练。

2.3.3 基于 Seq2Seq 的文本生成模型

本阶段模型主要根据输入的结构化指标数据

生成客观的指标描述文本。而异动指标变动原因描述文本均直接抽取自公告本身的原文，不参与本阶段的序列生成步骤。

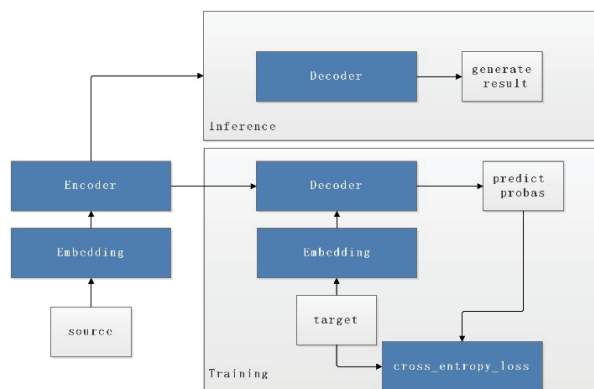


图3：Seq2Seq 整体框架

深度学习文本序列生成模型的骨干网络是Seq2Seq网络，该网络由编码器Encoder和解码器Decoder组成。编码器负责对输入数据进行语义编码，并将编码得到的文本语义向量输入到解码器进行解码。解码器的作用是根据编码器编码得到的输入语义信息，逐步生成目标文本。Seq2seq整体框架如图3所示。

与一般的文本序列生成任务相比，财经快讯指标描述生成任务有以下两点不同。首先，财经快讯生成模型的输入数据为结构化的表格数据，模型需要对编码器进行重新设计。其次，财经快讯生成模型需要对输入的财务指标进行筛选，从中选取重要的关键指标进行描述。除此之外，为了使得生成的描述文本符合常规阅读顺序，模型还需要决定指标的生成顺序。

针对上述问题，我们参考了论文^[1]以及论文^[2]的设计思路。模型在Seq2Seq基础框架中引入内容选择模块和内容编排模块。内容选择模块是让模型对不同财务指标属性内容的重要程度进行学习，内容编排模块根据内容选择模块的学习结果，对财务指标进行筛选和排序。另外，模型还引入了指针网络机制。该机制能够让模型直接从原文中复制数值、百分比等类型的数据，该机制能提高生成文本的数值准确度。该生成模型的整体架

^[1] Puduppully R, Dong L, Lapata M, et al. Data-to-Text Generation with Content Selection and Planning[C]. national conference on artificial intelligence, 2019, 33(01): 6908-6915.

构如图 4 所示。

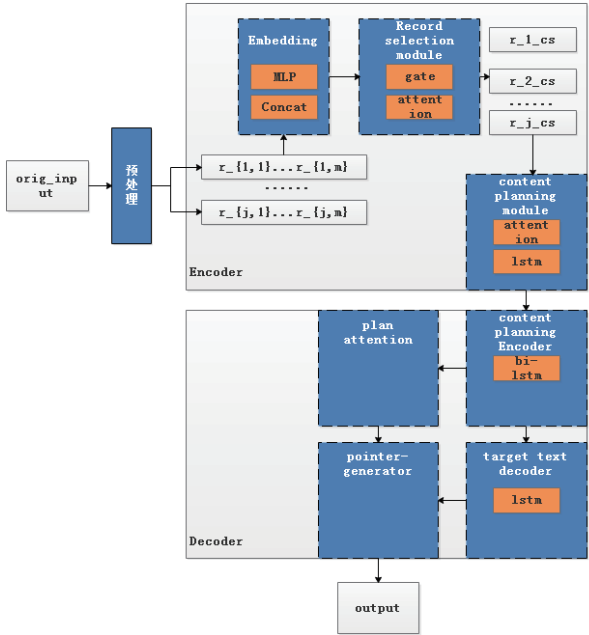


图 4：生成模型整体框架

数据预处理 原始的结构化数据无法直接输入到编码器中进行编码，模型需要对输入的结构化数据做预处理。具体地，模型将数据转化为三元组形式，每一行表示一个条目，如表 1 示：

表 1：条目数据示例

指标名称	属性类型	属性值
营业收入	当期数值	a
营业收入	变动百分比	b%
每股收益	当期数值	c
每股收益	变动百分比	d%

其中，每个条目可以表示为一个三元组 $r_i = \{r_{i,1}, r_{i,2}, r_{i,3}\}$ 。

数据嵌入编码 预处理得到的数据属于离散数据，不能直接输入到神经网络中进行学习。模型需要对输入数据进行 Embedding 编码，将所有属性值编码成浮点数向量表示 E_{r_i} 。

内容选择模块 内容选择模块主要是希望学习出每个条目属性内容的重要性，从而对信息进行筛选。为了实现上述功能，我们首先利用自注意力机制得到条目的新向量表示 r_j^{att} 。然后，模型引入门限机制来控制哪些信息可以流入到下游

的网络，从而让解码器能够聚焦在重要的条目信息上。

内容编排模块 该模块的主要信息是将编码器输出的条目信息进行内容编排，确定这些指标的顺序。模型引入一个隐变量 z 来表示编排后的条目列表，表 3 就是一个可能的 z 列表。为了简化问题，我们将研报文本的指标及其属性抽取出来，并按照其在原文的顺序进行排列，这样就得到每个目标文本的“近似最优编排” $target_z$ 。模型以 $target_z$ 为参照，通过 seq2seq 模型来学习最优的指标编排方式，从而筛选出重要的指标及其出现顺序。模块结构如图 5 所示。

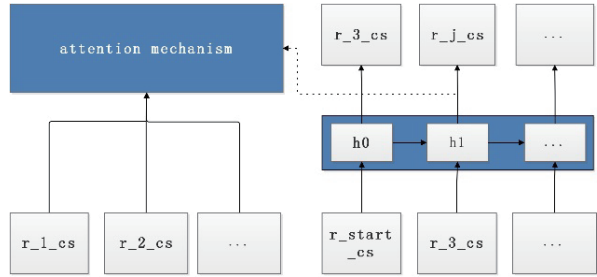


图 5：内容编排模块

解码生成模块 在解码阶段，模型会在解码器之前再添加一个编码网络，用于编码编排后的条目序列 z 。在构建词表时，模型一般会将数值、金额、日期等属性过滤掉。因此，模型引入指针网络来控制模型在某个时刻是使用原文复制的内容还是模型自己生成的内容。

3 效果展示

目前我们在公司自研金融事件舆情系统“泰极”上上线了财经快讯生成的部分功能。首先，图 6 展示了顺络电子 2020 年半年年报结构化解析的结果。从图中我们可以看出，模型已经将公告中的表格数据完整的解析成结构化数据。

关于财务快讯生成的效果，我们主要是从两个方向进行验证测试。首先，我们验证了财务快讯文本生成的时效性。我们对比了当前国内全网较快的财务快讯发布平台格隆汇以及同花顺财经，

[2] Nan Jiang, Jing Chen, Ri-Gui Zhou, Changxing Wu, Honglong Chen, Jiaqi Zheng, Tao Wan. PAN: Pipeline Assisted Neural Networks Model for Data-to-Text Generation in Social Internet of Things. 2020

Sunlord 深圳赛格电子股份有限公司 2020 年半年度报告

四、主要会计数据和财务指标

公司报告期末总资产较期初增加 1.58 亿元

单位：元

	本报告期	上年同期	本报告期比上年同期增减
营业收入（元）	1,442,035,173.06	1,217,417,618.36	18.40%
归属于上市公司股东的净利润（元）	242,301,177.06	189,084,137.10	24.09%
归属于上市公司股东的扣除非经常性损益的净利润（元）	231,204,007.10	181,876,123.88	27.12%
经营活动产生的现金流量净额（元）	223,134,244.26	210,834,300.00	-11.03%
基本每股收益（元/股）	0.30	0.24	25.00%
稀释每股收益（元/股）	0.30	0.24	25.00%
加权平均净资产收益率	5.40%	4.48%	0.92%

五、境内外会计准则下会计数据差异

1、同时按照国际会计准则与按照中国会计准则披露的财务报告中净利润和净资产差异情况

2、同时按照境外会计准则与按照中国会计准则披露的财务报告中净利润和净资产差异情况

图 6：定期公告解析结果实例

公告解读：泰禾集团半年度亏损15.82亿 盈转亏

同花顺财经 08-14 21:05

8月15日，泰禾集团发布2020年半年报，公司2020年1-6月实现营业收入24.63亿元，同比下降83.02%，归属于上市公司股东的净利润为亏损15.82亿元，同比止盈转亏，去年同期净利15.61亿元，每股收益为-0.6355元。

公司所属的房地产开发行业已披露半年报个股的平均营业收入增长率为7.13%，平均净利润增长率为27.06%。

泰禾集团(000732.SZ)：上半年由盈转亏15.82亿元

格隆汇 08-14 20:57

2020年8月15日，泰禾集团发布了2020年半年度报告。财报显示，报告期内泰禾集团实现营业收入24.63亿元，同比下降83.02%，主要原因为报告期内地产集中交付项目，造成收入较上年同期大幅下降，且归属于上市公司股东的净利润亏损15.82亿元，而去年同期盈利15.61亿元。在业绩大幅下滑的同时，在经营费用、管理费用、研发费用、财务费用、利息费用等几项费用中，财务费用的支出较高，达7.36亿元，同比增长了80.5%；此外，财务费用在上述几项费用中较上年同期增幅最大，主要原因是利息支出较上年同期增加；资产减值方面，公司实现资产总计2255.16亿元，同比增长0.54%；归属于上市公司股东的权益合计180.57亿元，同比下降8.34%。

格隆汇8月14日 | 泰禾集团(000732.SZ)发布2020年半年度报告，营业收入24.63亿

图 7：财经快讯生成实例

以泰禾集团相关公告为例，格隆汇在 2020 年 8 月 14 日 20 时 56 分发布了关于泰禾集团 2020 年半年报的快讯发布，我们的泰极舆情系统也在 2020 年 8 月 14 日 20 时 57 分发布了相关快讯，而同花顺财经的半年报快讯则是在 21 时 05 分发布。图 7 展示了上述平台发布的财务生成实例。从图中我们可以发现，泰极舆情系统生成的自动财务快讯能够快速准确地展示出公司相关的财务数据，同时也能展示部分指标变动原因。

4 总结与展望

本文主要介绍了华泰证券在财经公告快讯自动生成任务方面的技术探索和实践。我们的技术方案提供了自动获取财务公告、PDF 文本解析、关键财务信息抽取和财经快讯生成等一套完整的解决方案。从生成的文本内容来看，机器自动生成财经快讯是可行的。但是目前仍然存在一些问题，我们计划在以下几个方面进行继续优化：

(1) 文本数据准确性。财经快讯文本的数据准确性对财经快讯生成任务是最重要的。未来，我们计划采取两阶段的生成方式，一阶段模型生成带数值槽位的快讯模板，二阶段则是抽象为一个槽位填充任务。

(2) 文本内容完整性。我们计划将财务指标异动原因文本也整合到模型生成过程中，而非当前的简单拼接。

(3) 文本风格多样性。我们计划研究基于变分自编码器的深度学习模型，将多样性的指导融入到模型学习中，以此提升文本的多样性。

(4) 数据集构建优化。在数据集构建时，机器标注方式还存在一定问题。在积累充足语料后，我们将尝试构建基于 BERT 模型的文本匹配模型来帮助构建数据集。

基于大数据的智能化客户服务平台

应原、任荣、王东、王洪涛/海通证券股份有限公司 信息技术管理部

随着近年来人民生活水平的不断提高，越来越多的人通过参与金融市场寻求资产的保值增值，证券公司的客户数不断增加，对客户经理精细化服务客户、营销客户的要求不断提高，传统的客户关系管理系统逐渐难以支撑不断发展的业务规模。结合大数据的计算能力和智能化工具的便利，打造的客户服务平台正逐渐取代传统的客户关系管理系统。本文主要介绍了海通证券基于大数据的智能化客户服务平台的设计及其重要组成模块等。



一、背景

在迅猛发展的互联网形势下，监管适当性要求日益趋严，各业务部门和分支机构对客户服务的要求也在不断提高，客户关系管理系统作为证券行业维护客户关系、提高客户服务质量的重要系统，采集、处理并展示了大量的客户信息。按公司客户量不同，每日需要采集及处理的数据从

百万级、千万级发展至今上亿级的规模。

传统的客户关系管理系统一般采用关系型数据库用来处理数据，随着客户数量的增长、交易数据的增长，数据的处理加工的时间越来越长，消耗的机器性能和资源越来越多，成本不断提高。同时，近年来大数据生态圈蓬勃发展，根据国家网信办发布的《数字中国建设发展报告(2017年)》预计2020年中国大数据市场产值将超万亿。在

这样的背景下，海通证券启动了基于大数据的智能化客户服务平台建设，构建以客户信息管理为基础、以客户服务为目标、推进资源共享，覆盖客户全生命周期管理的统一客户服务体系，结合智能化工具以更好的为分支机构和客户经理提供展业支持。

二、总体设计

本平台提供了PC端和APP端服务，PC端提供了营销管理和事务管理等功能模块，实现了任务建立、任务分配、任务执行、执行跟踪、任务评价等一体化的服务流程；APP端涵盖了客户经理版本和管理人员版本，搭建首页、客户、产品、任务、个人中心五个方面功能模块。自推广上线以来，日均登录人数千余人，覆盖了大部分前台营销人员，尤其是APP端，为营销人员外出移动展业提供了有力的支持。

本平台结合大数据搭建了高效的Spark-MOT引擎，为一线客户经理生成名下客户的各类场景任务，帮助客户经理维护及服务客户；提供了灵活的客户综合查询功能，利用其高效的分布式检索能力实现智能化和精准化的客户营销与运营；借助Impala高效的运算能力组建了海量数据批处理模块，对各个源系统提供的客户数据

进行采集和分析挖掘，构建全方位的客户画像，方便业务人员认识并了解自己的客户；通过高可靠的HBase历史数据查询模块，为追溯客户服务记录、信息发送历史等提供便利；结合智能化的RPA工具及全覆盖的智能监控中心，为平台的智慧运维保驾护航；搭建了高吞吐量的消息中心，为快速、有效的触及大批量客户进行服务与营销提供了有力手段。

图1为本平台的组成图，本文第三章对各主要组成模块进行了简要介绍。

三、平台各组成模块

3.1 高效的Spark-MOT引擎

Spark是专为大规模数据处理而设计的快速通用的计算引擎，用来构建大型的、低延迟的数据分析应用程序，Spark的一个主要特点是能够在内存中进行计算，其运算效率相较传统的MapReduce可提高100倍以上。Spark具有高效性、易用性、通用性、兼容性等特点。

本平台使用基于Spark的MOT引擎实现了百余个场景任务的每日生成与分发，每日处理千万级数据，生成与分发的场景任务数达到数十万个，耗时不到20分钟，极大的加强了分支机构对客户服务的支持力度。

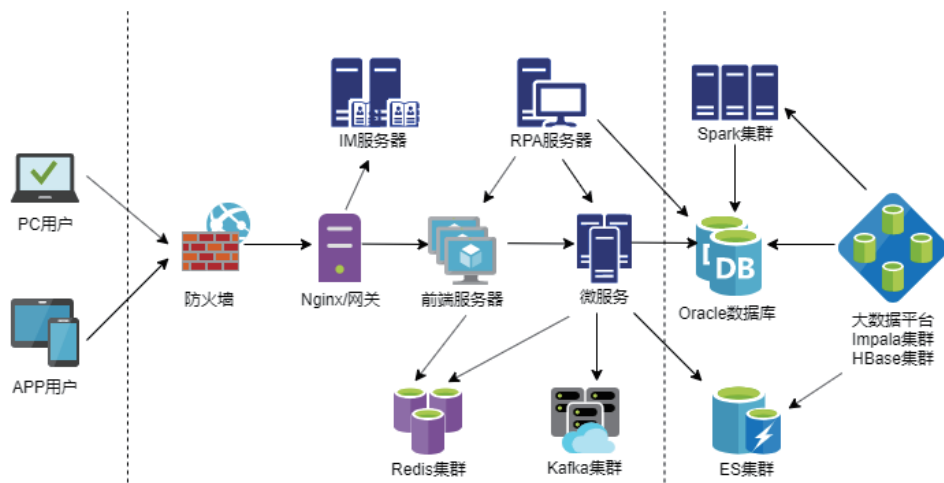


图1：本平台组成图

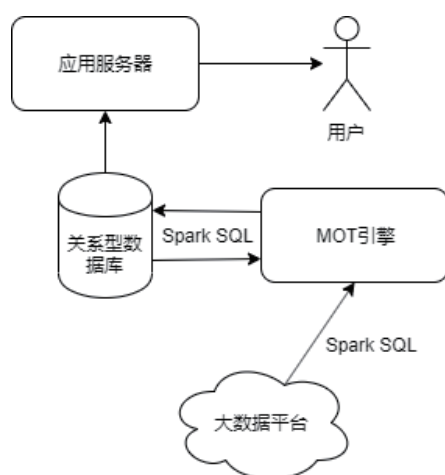


图 2：基于 Spark 的 MOT 引擎逻辑图

基于 Spark 的 MOT 引擎具有如下特点：

(1) 高效率

相较于传统的关系型数据库上的存储过程，基于 Spark 的 MOT 引擎生成效率可达到 5 倍以上（与基于关系型数据的传统系统并通过存储过程生成每日 MOT 事件进行对比）。

(2) 跨数据源

Spark SQL 可从多个数据源抽取数据，注册在内存中形成虚拟表，如从大数据平台抽取计算后的客户相关指标数据，从关系型数据库抽取实时的客户关系信息用于事件的分配，避免了在多个平台上数据的多次导入导出及冗余的问题，同时提高了数据的时效性。

(3) 易扩展

目前 MOT 引擎已集成了预计算的百余个客户相关的指标，新增一个使用预计算指标的 MOT 事件可在前端界面上直接配置并实时生效，无需后台升级等繁琐步骤。

(4) 低成本

除用于调度、管理 MOT 引擎的应用服务器以外，Spark 计算集群采用已有的大数据平台的资源，通过合理的资源配置管理以支持 MOT 引擎运行。Spark SQL 对于资源的调度和利用非常优秀，200G 的内存即可支撑 MOT 引擎千万级数据量的数据处理。

3.2 灵活的客户综合查询

ElasticSearch（简称 ES）是一个基于 Lucene 的搜索服务器。它是一个分布式、高扩展、高实时的搜索与数据分析引擎。它能很方便的使大量数据具有搜索、分析和探索的能力。根据 DB-Engines 的排名显示，ES 是最受欢迎的企业搜索引擎。

通过基于 ES 索引查询技术支持近千余指标字段的客户筛选查询，所有字段均可以灵活组合，为每个用户提供个性化的查询方案，同时在千万级数据量的基础上可实现秒级的结果返回，为智能营销提供有力支撑。

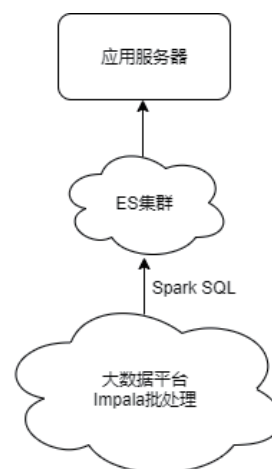


图 3. 基于 ES 的客户综合查询逻辑图

基于 ES 的客户综合查询具有如下技术特点：

(1) 高效率

在千万级数据量上，支持对客户近千余指标字段进行单字段查询、多字段查询、范围查询、聚合查询等，可快速返回结果集。相较于传统的关系型数据库查询效率提升 100 倍以上。

(2) 冷热分离的读写技术

客户指标类的数据需要每日更新，为了提高数据推送至 ES 的效率，采用了冷热分离的读写技术，即推送数据直接写入热节点（SSD 固态硬盘），而数据的备份（replica）同步在冷节点（STAT 盘）上。数据写入热节点可以提高读写速度，减少每日数据准备的时间。

(3) 双模型切换

为保障客户综合查询的功能持续可用，采取了双模型切换的方式，每次向 ES 推送数据时会自动检测正在使用的模型，并向另一个模型推送数据，当日数据推送完成后，进行模型切换，整个过程仅需几秒即可完成，确保功能随时可用。

(4) 低成本

客户综合查询的 ES 模型建立在统一的 ES 查询集群上，由各应用系统共用该查询集群，实现资源共享统一管理。

3.3 快速的海量数据批处理

Impala 是一个基于的 MPP (大规模并行处理) 的 SQL 查询引擎，提供对 HDFS 数据的高性能、低延迟的交互式 SQL 查询功能。它基于内存进行计算，能够对 PB 级数据进行交互式实时查询、分析，Impala 作业无需转换为 MapReduce，直接读取 HDFS 的数据，从而大大降低了延迟，具有数据仓库的特性，可对 hive 数据直接做数据分析，同时支持 JDBC/ODBC 远程访问。

本平台为了兼顾应用系统对数据的实时高效访问及夜间批处理对海量数据的聚合计算等，将客户关系管理系统所需的数据分为了两部分，一部分数据为客户的当前数据及已加工数据，数据量较小，存储在关系型数据库中支持实时查询及业务处理；一部分数据为客户的历史数据及未加工数据，数据量较大，存储在大数据集群上，通过 Impala 数据批处理加工后导入关系型数据库中。

数据仓库采集各源系统的数据包括客户信息、交易信息等等，在数据仓库中加工基础指标，加工完成后推送到大数据平台，在大数据平台上通过 Impala 进一步加工应用系统所需的数据，包括各维度、各时间段汇总及同比环比等数据，最后通过 Sqoop 同步到应用系统数据库。

基于 Impala 的海量数据批处理具有如下技术特点：

(1) 高效率

以本平台为例，Impala 清算逻辑千余个，清算耗时约 2 小时，总体耗时约 3 小时（含数据推送等），相较于基于关系型数据库的传统系统提升约 3 倍（原数据处理约 6-8 小时），同时随着大数据的资源扩充，Impala 的并发度可以不断提升，进一步压缩总体的执行时间。

(2) 低门槛

通过结合 CDH 大数据上的开源调度组件 Oozie 及相应的调度脚本，开发人员只需要编写通用的 SQL 语言脚本（部分函数需要替换）即可实现调用 Impala 进行数据清洗，无需了解大数据集群的授权、任务分发等较为复杂的过程。

(3) 易扩展

Impala 基于大数据平台，扩容仅需要在新增集群节点上增加 Impala 角色即可使该节点成为 Impala 的计算节点并对集群数据进行 rebalance 即可。大数据平台近年来经历了多次扩容，现已达到几十个节点的规模，多次扩容均无大量修改工作。

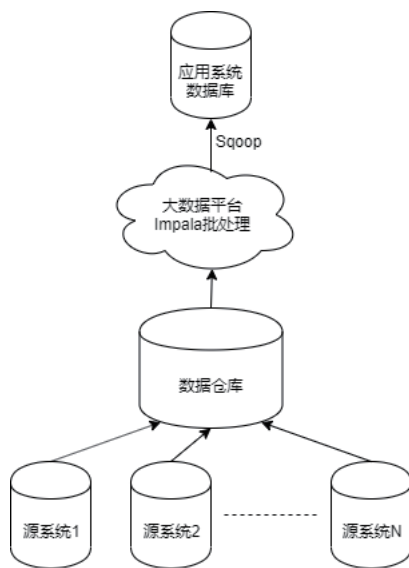


图 4：基于 Impala 的数据批处理逻辑图

3.4 高可靠的历史数据查询

HBase 是一个高可靠性、高性能、面向列、可伸缩的分布式存储系统，利用 HBase 技术可

在廉价 PC Server 上搭建起大规模结构化存储集群。Phoenix 查询引擎支持使用 SQL 进行 HBase 数据的查询，会将 SQL 查询转换为一个或多个 HBase API，协同处理器与自定义过滤器的实现，并编排执行。

为了解决随运行时间不断增加导致历史数据积压在关系型数据库中，拖累整体运行效率的问题，本平台采取保留近 1 年的数据在关系型数据库，历史数据导入大数据平台并提供基于 HBase 的数据查询接口进行历史数据查询以解决该问题。

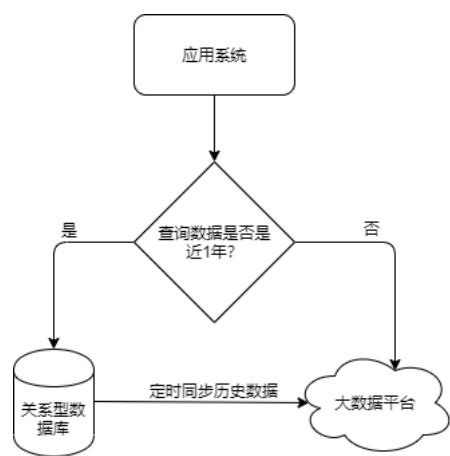


图 5：基于 HBase 的历史数据查询逻辑图

基于 HBase 的历史数据查询具有如下技术特点：

(1) 低成本

传统的关系型数据库随着数据量不断增加需要不断进行纵向扩容，即时使用一体机等高性能的机器也难以避免性能的瓶颈且成本会成倍上升，而将历史数据迁移至大数据平台使应用的关系型数据库的内容增长是平缓可控的，大数据平台本身的扩容是横向的，不会带来大量的成本提升。

(2) 高可靠性

依托于大数据平台的海量数据容量及易扩展性，可以长期储存历史数据，HBase 本身具有 Write-Ahead-Log (WAL) 机制，用于数据的容错和恢复。同时通过将历史数据统一存储于

HBase 并提供查询接口的方式，即使应用系统发生颠覆性的更迭也不用担心历史数据的迁移或者丢失的问题。

(3) 易开发

采用 Phoenix 与 HBase 结合，开发人员只需要编写标准的 SQL 脚本即可查询 HBase 上的数据提供应用系统查询。

3.5 智能化工具模块

本平台提供了多个智能化工具，助力提升营销人员客户服务及技术人员自动化运维。包括通过 OCR (Optical Character Recognition, 光学字符识别) 提供客户名片的自动识别及录入等、ASR (Automatic Speech Recognition, 语音识别技术) 提供服务记录、工作日志的快速录入，知识库的快速检索等及通过 RPA (Robotic Process Automation, 软件流程自动化) 提供智能运维。

RPA 是指用软件自动化方式实现在各个行业中本来是人工操作计算机完成的业务。它让软件机器人自动处理大量重复的、基于规则的工作流程任务。针对上线初期普遍存在不稳定及用户问题较多的情况，本平台结合 RPA，实现了数据自动校验、自动定时巡检、交流群记录收集等智能化运维的功能。

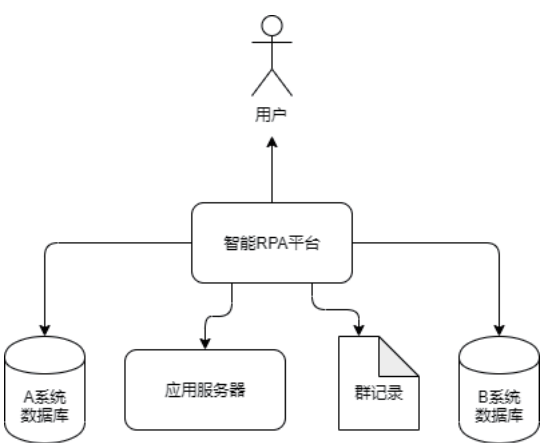


图 6：智能 RPA 逻辑图

其包括如下技术特点：

(1) 数据自动校验

客户的数据及指标纷繁复杂，本平台上线初期存在新老系统并行的情况，需要不断核对数据差异，存在大量的重复性工作，为了确保数据的质量，RPA 每天定时从两个数据库中抽取部分重要的客户指标数据，进行自动化比较，并生成相应的报告发送技术人员，极大的减少了重复性劳动。

(2) 自动定时巡检

为了保障高可用性，及时发现可能的问题，RPA 会定时登录本平台，模拟操作各类功能模块，对于连续多次无法完成操作会发送提醒至相应的技术人员。

(3) 交流群记录收集

本平台上线后通过沟通交流群与用户进行交流，用户对各项新功能总会有各种各样的问题，为了更好的收集归纳问题，形成长期的知识库，RPA 对交流群中的聊天记录进行了定时收集并进行智能化的筛选，对于共通性的问题会纳入知识库中进行分享。

3.6 全覆盖的智能监控中心

Logstash 是一个数据收集引擎，具有实时管道功能。Logstash 可以动态地将来自不同数据源的数据统一起来，并将数据标准化到指定的目标。

通过将应用日志按照规定格式打印输出，日志平台 Logstash 进行采集并推送至 ES 集群，实现了对登录用户数、每日请求数、访问异常、平均请求耗时等关键指标的监控，展示成监控大屏，供技术人员及时了解平台运行情况，同时根据指标的异常情况智能化的判断健康程度。日志平台还提供通过关键字对日志进行搜索的功能，避免登录生产环境查询日志带来的操作风险。

基于 Logstash 的日志监控大屏及日志搜索有如下技术特点：

(1) 高效率

智能监控中心每日采集约千万级条日志，通过 Filebeat 及 Logstash 准实时采集并进行展示，

覆盖各项监控指标，同时提供准实时的日志搜索以供问题排查使用。

(2) 搜索速度快

基于 Lucene 语法的日志搜索可快速搜索 3 个月内的应用系统日志，通过关键字可在几秒内定位到日志，并按照需求展示该条日志前后的相关日志信息，为方便定位问题提供了极大的便利。

(3) 长期存储

针对 3 个月内的日志，存储在 ES 集群上以便快速搜索，超过 3 个月的数据归档至大数据平台，以文件形式长期存储，为排查问题、监管调阅资料等提供支持。

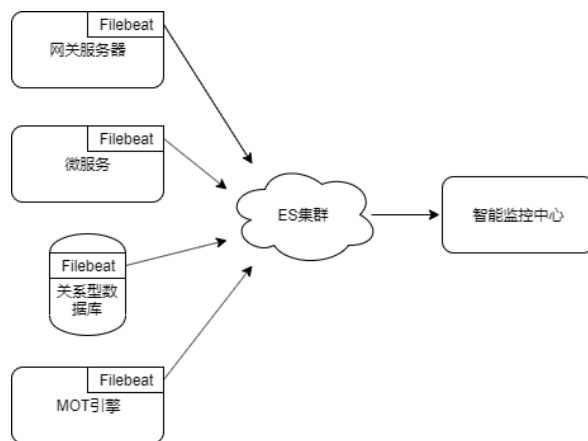


图 7：智能监控中心逻辑图

3.7 高吞吐量的消息中心

Kafka 是一种高吞吐量的分布式发布订阅消息系统，通过 $O(1)$ 的时间复杂度进行顺序写入磁盘数据结构提供消息的持久化，这种结构对于即使数以 TB 的消息存储也能够保持长时间的稳定性能，即使是非常普通的硬件搭建的 Kafka 集群也可以支持大量消息的吞吐。

随着券商业务的不断发展，对于发送短信、微信、邮件等发生触及客户的需求不断提升，本平台采用基于 Kafka 集群的消息中心，每小时可吞吐百万数量级的消息，同时可根据需求推送至客户的微信、邮箱或短信上，为快速触及用户提供了极大的便利。

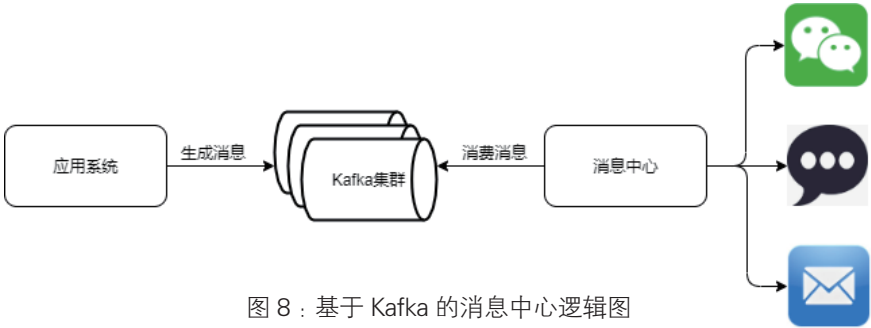


图 8：基于 Kafka 的消息中心逻辑图

基于 **Kafka** 的消息中心具有如下技术特点：

(1) 高并发

本平台使用的 **Kafka** 集群，支持多个应用系统、多个进程同时写入和读取消息，可支持上千个进程同时读写。

(2) 多渠道

根据需求可将消息推送到微信、短信、邮件等传统渠道，同时支持渠道扩展，包括手机端消息推送、微博、公众号等等。

(3) 最多一次的消费模式

消费端采用最多一次（**at most once**）的消费方式，确保每条数据最多会被消费一次，避免产生重复发送消息对客户带来的困扰。

针对客户经理向客户发送的消息，进行关键词、敏感词过滤；针对推荐的产品，根据产品的风险等级与客户的风险等级等进行适当性匹配，对于不适当的产品无法进行推荐，保障信息的合规性。

(3) 持久性

所有的聊天记录等内容都会长期保存在服务器上，确保可追溯可查证，达到服务留痕的目的。

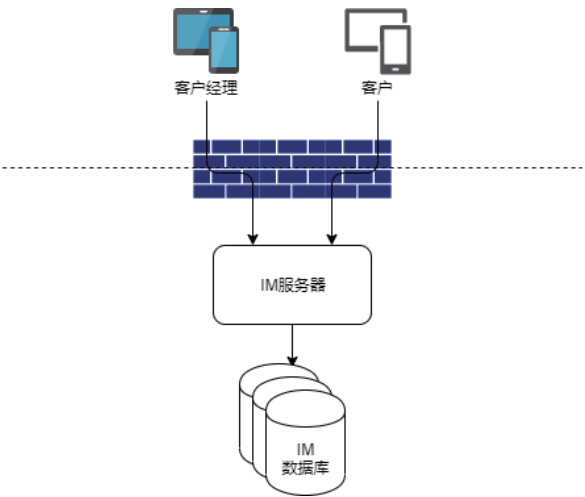


图 9. 安全合规的即时聊天模块逻辑图

3.8 安全合规的即时聊天模块

近年来，随着即时聊天软件的普及，越来越多的沟通通过微信等第三方工具进行，随之暴露出了客户信息泄漏、适当性不匹配等问题。为保障沟通记录的私密性、合规性，本平台提供了私有化的 **IM**（即时聊天）模块，为客户经理与客户经理之间及客户经理与其服务的客户之间的在线聊天、音视频聊天、附件收发、群发消息等提供了安全合规的渠道。

本模块具有以下技术特点：

(1) 私密性

所有的聊天记录，包括文字、音视频、附件等均保存在私有化的内网服务器中，外网无法直接访问，同时通过权限隔离、内容加密等方式限制相关内容在服务器端的访问，确保其私密性。

(2) 合规性

四、展望

未来本平台会进一步发挥背靠大数据的优势，深度对接企业图谱、人工智能、智能外呼、**AI** 平台等模块，加强数字化营销、智能营销、统一客户信息等模块的建设，同时，结合机构业务条线的需求，打造一款统一的客户管理、服务营销和业务管控平台；围绕财富管理转型主线，前瞻性推进 **APP** 建设，提高营销服务和经营管理效率，打造业内员工移动展业支持领先品牌。

证通云外呼系统智能化实践

王晴 / 证通股份有限公司 wangqing@ect888.com

何浩 / 上交所技术有限责任公司 hhe@sse.com.cn



本文主要叙述了证通股份基于上交所技术证通云在智能外呼平台建设过程中的探索与实践，具体阐述了外呼系统的整体设计及实现模式，希望对智能客服+专有云建设的“一站式”云SAAS解决方案有较好的借鉴价值。秉承“降低投入成本，提供持续服务”的机构服务理念，深入研究证券行业外呼对话策略设计方案，实现基于文本精准分类的意图识别，语义深度匹配的智能问答，模块关键计算的流程引导等技术能力。核心智能引擎云部署通过云管理平台统一管理，业务流程与关键数据安全隔离，确保业务运营安全合规。

一、概述

伴随人工智能的不断演进，人机对话技术能力取得了突破性进展，尤其是“语音识别”、“语音合成”以及“语义理解”上述三个关键任务的执行效果获得了大幅提升，识别能力已接近人类水平，这使得“机器人给人类打电话”的科幻场景走进了我们的生活。通过基于人机对话技术能力打造的智能外呼系统，可大幅降

低机构给客户拨打电话的成本，并充分利用机器人的核心计算能力收集客户反馈、通知客户关键事项、挖掘客户行为线索等方式，为机构提供更有效的客户触达手段。目前，智能外呼的应用已受到行业内机构广泛关注。

然而在业界逐渐接受智能外呼作为新触达客户的有效手段的同时，笔者也发现了一些机构通过智能外呼服务客户中暴露出的一些问题。首先，智能外呼作为人工智能技术的应用产品，

需要机构花费高额成本进行智能语音相关产品采购。其次，人工智能技术的应用依托于人力持续的维护，良好的智能外呼服务效果需要机构组织专职人员持续对话术模型进行调优才能达到，而大部分机构在这方面的人员投入难以保障，从而影响了智能外呼的服务效果。除此之外，由于工信部对外呼电话线路资源的管控，使得诸多机构难以获得足够的外呼线路资源为客户提供服务，导致了机构即使购买了相关服务也难以顺利开展外呼业务。

基于此，上交所技术有限责任公司（以下简称“上交所技术”）联合证通股份有限公司（以下简称“证通股份”）为行业通过提供云模式的智能外呼服务，大幅降低机构投入成本，为机构提供话术模型的持续优化服务，节省了机构的话术模型维护成本。同时，通过与运营商达成合作帮助机构获取外呼线路，解决了大量机构外呼线路资源不足的问题。同时，深入研究证券行业外呼对话策略设计方案，将自然语言理解及深度学习等 AI 技术嵌入智能外呼系统，实现基于文本精准分类的意图识别，语义深度匹配的智能问答，模块关键计算的流程引导等技术能力，有效提升智能外呼业务的成功率及服务满意度。同时，将核心智能引擎云化部署，业务流程和关键数据安全隔离，在保证业务运营合规安全的基础上，实现智能能力资源充分复用，有效降低机构用户使用成本。本文将基于证通股份与上交所技术联合建设的证通云智能外呼平台，阐述外呼系统的整体设计及实现模式。

二、价值

证通云智能外呼伴随业务的发展为机构客户带来多重价值。首先，智能外呼技术解决了传统呼叫中心的关键矛盾点“业务发展不断增加的外呼量 VS 呼叫中心人工坐席不足”，通过智能机器人辅助客服中心有效完成外呼任务。

其次，由于智能外呼低成本服务的特性，可满足机构以更低成本为提升客户投资体验进行服务。同时，作为触达客户的有效手段，通过机器人外呼可以通过与客户的交流，挖掘客户的行为背后的意图，并通过人工智能与数据分析对客户行为进行记录并了解其背后的原因，从而更好的帮助机构进行用户运营与数据分析。在带来业务价值的同时，证通云智能外呼通过云模式有效降低了机构采购相关服务的投入成本，并通过证通云的针对话术提供的运营服务降低了机构维护的成本。机构可以通过证通云向运营商申请外呼线路资源，全面降低机构开展智能外呼投入成本。

三、技术应用

智能外呼在证券与期货机构的合规服务回访、服务通知等领域具有广泛的应用场景，例如：开户断点回访、存量回访、新开户回访、新股中签通知等。

以“断点开户回访”为例对技术应用做相应的解释说明：

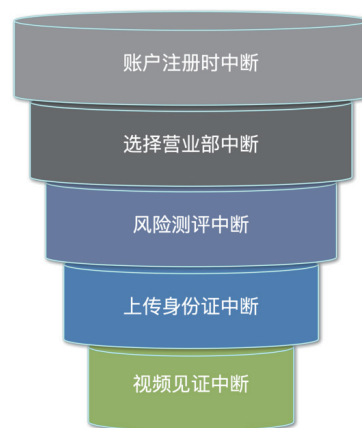


图 3.1. 开户断点漏斗模型

（一）业务需求背景

针对进入开户流程但没有完成开户的客户，进行机器人电话回访，通过解答客户在开户中遇到的问题，引导客户完成开户流程。

（二）机器人外呼模型介绍

开户断点场景主要针对开户过程中的 5 个环节设计外呼流程，通过主动询问客户“是否遇到了什么问题？”，获取用户开户过程中遇到的难点，并针对不同的客户情况智能化解答。

目前，通过话术模型的调优与实际运营，行业经营机构通过开户断点回访可达到 16% 以上的客户开户召回率，计算公式如下：

$$(\text{呼叫后完成开户的客户量}) / \text{总外呼客户量} * 100\%$$

四、建设思路

1、整体方案设计

证通云智能外呼服务采用 IaaS+ 多租户的设计模式，各租户享有独立的业务运营空间和数据存储空间的同时，充分复用底层硬件资源和核心 AI 技术，包括：自动语音识别 / 合成、自然语言理解，以及深度学习网络，以及外呼系统的关键组件 IVR、CTI 等，其总体的方案架构如图 4.1 所示。

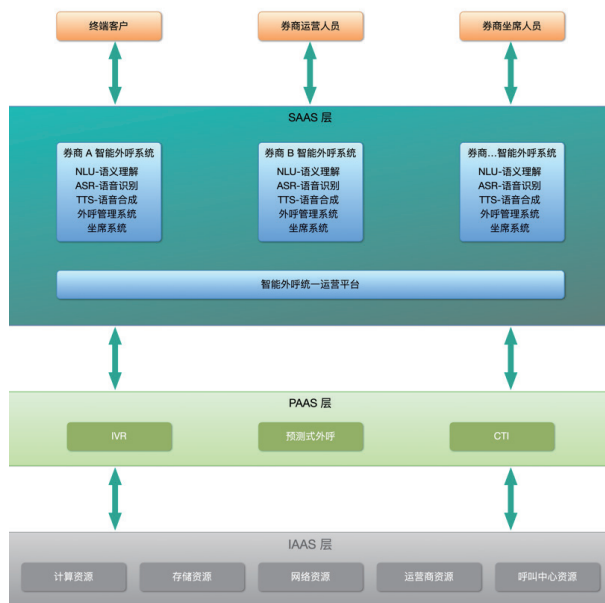


图 4.1：证通云智能外呼服务的整体架构图

整体设计思路：IaaS 的基础资源使用证通云成熟、安全的计算、存储、网络及运营商资源，

从而提供高安全等级、高稳定性的资源服务，支持应用层的即取即用、安全稳定、弹性扩展。同时，将核心的智能模块（语音识别 / 合成、自然语言理解、IVR/CTI 等）封装成平台层的公共服务，实现智能服务的按需即取、动态扩展。系统基于平台层的智能服务输出，结合机构用户的应用场景，构造独立的业务运营空间和数据存储空间，保证机构用户可快速借助云外呼机器人，实现独立运营管理本机构的外呼业务和相关业务数据。

2、逻辑架构

证通云智能外呼服务整体的逻辑架构分为基础资源层、技术层、平台管理层、机构管理层、机构应用层和接入层，如图 4.2 所示：



图 4.2 证通云智能外呼服务的逻辑架构图

上交所技术证通云提供基础资源服务，支撑上层应用的快速构建。证通股份基于自身在软交换、智能语音、自然语言处理等 AI 技术上的积

累构建智能云服务平台，为其上层业务系统输出核心 AI 能力，实现不同场景下的智能化服务应用。平台管理层围绕业务的核心功能，提供辅助性的公共服务，如机构管理、权限管理、应用配置和平台监控等，并围绕话术模型的创建、训练、分配等管理功能，设计数据标注、训练和管理的闭环解决方案。独立于公共的平台管理之上，每个机构形成各自独立运营管理空间，保证功能应用和数据存储的有效隔离。机构应用层则根据具体的业务场景来设计，实现相应的业务流程和业务功能，为机构提供如智能外呼、人工坐席、智能培训、智能质检等智能服务。目前机构仅需与云平台完成系统和线路对接、将业务场景话术提供至证通股份进行模型配置，即可通过上传外呼名单创建任务，实现智能自动化的外呼服务。

3、整体流程设计

证通云智能外呼服务通过使用智能机器人与终端客户进行对话沟通，辅助证券基金经营机构开展外呼业务，其主要流程围绕智能机器人与客户的语音交互进行，通过发起外呼——客户应答——语音路由——语音识别——语义理解——语音合成——响应回复，形成一个完整的语音流转、交互的闭合回路，如图 4.3 所示。

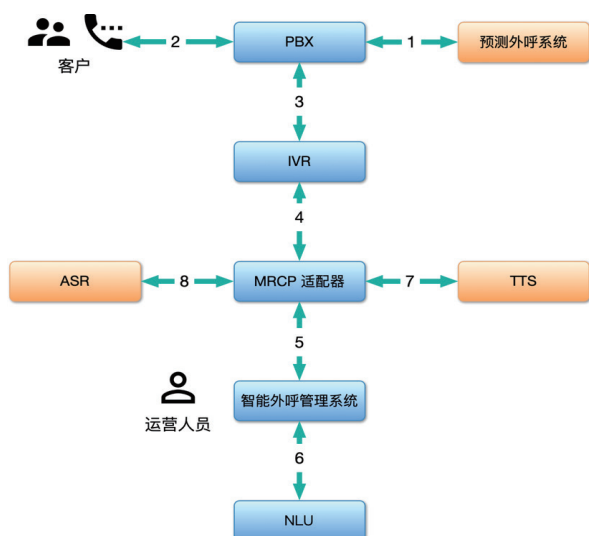


图 4.3 证通云智能外呼服务的交互流程图

业务流程简单归述如下：① 预测外呼系统发起预测式外呼；② PBX 接收外呼指令后拨打客户电话；③ 客户接听后，PBX 将线路转到 IVR；④ 如果首次接通，则向 MRCP 适配器发送首次接通指令；⑤ MRCP 适配器发指令给智能外呼管理系统；⑥ 智能外呼管理系统从 NLU（语义理解）中获取接听后首次需要播报的文本（外呼流程已预先配置好）；⑦ MRCP 将第 6 步获取的文本发送至 TTS 获取语音播报内容，通过 IVR 播报给客户；⑧ 客户根据 TTS 播报内容进行回答，MRCP 从 IVR 获取语音流发送给 ASR 进行识别，并获取识别结果；⑨ 重复第 5、⑥ 步骤，直至预先设置流程的挂机指令。

五、关键技术

证通云智能外呼系统使用多租户技术实现多机构单环境下共用系统组件，并且确保各租户数据的有效隔离。同时作为任务型对话管理系统，智能外呼旨在完成机构的特定任务需求，通常需经过多轮交互为用户提供服务，为此，对话模块整体架构的设计参考了 Steve Young 提出的 POMDP 框架。

1、多租户技术

多租户技术（Multi-Tenancy Technology）是一种软件架构技术，保证租户内独立的业务运营空间和数据存储空间的同时，充分复用底层硬件资源和核心 AI 技术。

• 数据隔离方面

证通云智能外呼使用独立数据库的方式实现机构数据存储空间的隔离，相比于共享数据库的数据隔离方案，证通云智能外呼使用的数据隔离方案具有数据隔离级别最高，安全性佳等优点，相反带来的缺点是部署成本较高。

• 应用程序方面

针对核心模块对话模型，系统使用多进程

的方式进行隔离，各进程之间互相独立，互不影响，可充分保障外呼过程中的通话效果。

同时系统使用微服务架构来保障系统的高可用性和可扩展性。系统按照功能分为统一运营平台、外呼管理平台、人工坐席系统、对话引擎等4大模块。每个模块都由若干微服务组成。根据业务运行情况，系统可便捷扩充相关模块的性能指标，满足支持证券行业上百家机构的外呼的需求。同时，任何微服务都采用双活方式部署，可保障在单点服务出现问题时，系统继续保持可用。

2、对话引擎

证通云智能外呼对话引擎综合利用智能语音交互技术（包括语音识别（ASR）、语音合成（TTS）、自然语言处理技术（包括自然语言理解（NLU）、对话状态跟踪（DST）、回复决策（DPL）及语言生成（NLG）），实现基于电话线路的人机对话。系统可在会话过程中的数据实时识别获取，客户意图解析，需求追踪并判断当前对话状态的功能，并基于用户意图识别及回复策略生成系统对话语句，通过语音合成技术返回给客户，从而实现“一问一答”的多轮人机对话交互。整体实现如下图所示：

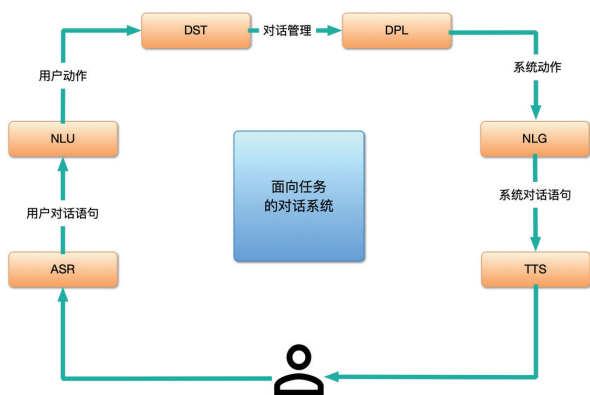


图 5.1

• 智能语音交互

语音识别（ASR）：提供将录音文件或者实时音频流转成文字的服务，能够支持中文、英文、

方言普通话等多个语种，支持噪音环境下和带背景音的语音识别，支持机器自动分离不同人声，支持快速定制识别中的热词以提升准确度。

在智能外呼场景中，语音识别的准确率和响应时间是核心指标，目前证通云智能外呼对语音识别引擎进行了深度定制，已经形成证券行业语言学模型和证券行业声学模型，确保在人机对话过程中，语音识别准确率可达到一个很好的效果。

语音合成（TTS）：通过先进的深度学习技术，将文本转换成自然流畅的语音。使用海量的音频数据训练合成数据，合成音真实饱满、抑扬顿挫、富有表现力，MOS 评分达到业内顶级水准。目前有多种音色可供选择，并可以调节语速、语调、音量、停顿等多项指标。

• NLU 自然语言理解

在面向任务的对话系统中，基于 NLU 技术将识别出来的文本信息转换为机器可以理解的表达方式，包括句子检测，分词，词性标注，句法分析等。自然语言理解参考了框架语义分析的解析模式，针对会话中的语法、语义进行分析，将用户输入的内容包括用户表述及语境映射为用户的意图和相应的槽位值，并采用计算机可读的结构化形式，结合上下文语境完成槽位填充（slot-value）及意图排序，最终输出用户意图（user intention）。即通过语义分析输出所谓的用户需求，为对话系统提供任务执行方向及执行该任务时输出的参数值。

意图识别流程如图 5.2。

针对外呼任务执行时用户输入的内容，基于 NLP 处理实现用户意图的识别功能。首先，系统对输入词法进行基础的分析，包括中文分词及词性标注两部分，即将一句话的每个词标注为名词、动词、形容词等等。词性标注完成后，对句子中的实体进行命名实体识别，此处采用 CRF 条件随机场算法进行提取操作，并根据正则算法判断是否为规则配置的实体，有效得将

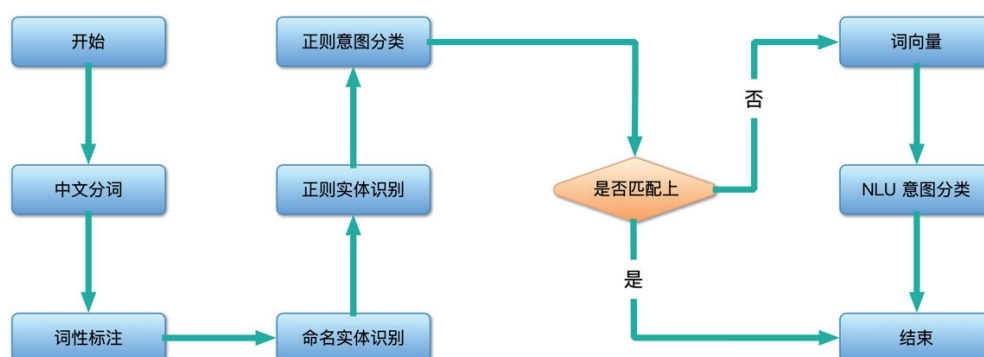


图 5.2

命名实体及正则从输入内容中抽取出来，为后期意图的精准分类奠定基础。其中，系统将正则实体根据配置好的正则意图分类规则进行匹配，如若匹配成功则完成 NLP 处理操作，并将结果返回；如若尚未匹配成功，系统则进入词向量阶段，将文本编码为一个向量，生成意图分类的输入内容，基于模型的深度学习神经网络来预测文本的意图，最终实现将预测结果返回的功能。

• DST 对话状态追踪

DST (Dialogue State Tracking) 对话状态追踪模块是在某一时刻结合当前用户的输入及历史对话状态及相应信息追踪用户需求并推断对话状态和目标用户。该模块针对每一轮次对用户的目标进行预估，以多轮对话历史、当前的用户输入作为数据依据，结合上下文总结、推理和理解，管理每轮会话的输入和历史，输出用户含义。此模块采用上下文替换和主体补全技术，综合考虑用户的多轮动作辅助对话系统精准识别用户的真正需求。目前对话状态表示主要包括三个部分：目前槽位填充情况、本轮对话过程中的用户输入、对话历史。其中槽位填充情况是对话状态追踪的核心指标。

然而，由于在 ASR 和 NLU 这两个环节会存在误差，因此语音识别和语义理解输出的每个句子或者每个意图结果都带有一个相应的置信度值，DST 往往也是输出各个状态的概率分

布。为提升追踪精准度，DST 在判断当前的对话状态采用两种不同的规则方式：1-Best 方式和 N-Best 方式。

(1) 1-Best 方式

1-Best 方式是指 DST 判断当前对话状态时只考虑置信度值最高的情况，输出的状态也是确定型，因此维护对话状态的表示时，只需要等同于槽位数量的空间，在较为复杂的场景下适用性不强。

(2) N-Best 方式

证通云智能外呼使用此方式。N-Best 方式是指 DST 判断当前对话状态时给出多个置信度值，综合考虑所有槽位，以更好的保证结果。因此，DST 模块在上述多个置信度值的基础上，不仅要维护一个对话状态，而是要考虑每个槽位的结果，输出组合后用户意图的整体概率分布。最终，通过整体置信度值判断当前对话状态。

• DPL 对话策略学习

DPL (Dialogue Policy Learning) 即为对话策略，又被称作对话策略优化 (Optimization)，是系统根据对话状态追踪 DST 输出的对话置信状态，设计对话策略决定下一步执行系统动作的过程。比如根据 DST 知道有哪些槽位未填充，用规则的方法选取一个优先级最高的槽位进行提问，DPL 即为一个序列的决策过程，因此常用的方法有规则方法、CRF 方法和强化学习等。

目前证通云智能外呼系统是基于神经网络

及框架管理的方式，预先搭建真实业务场景模型，供机器离线学习对话策略，并通过规则对用户输入的槽位填充情况进行管理，实现对话过程中直接应用学习结果、明确槽位相应执行动作，从而生成当前对话的回复策略。

• NLG 自然语言生成

NLG (Natural Language Generation) 自然语言生成的任务是将对话策略 DPL 模块输出的语义内容表示转化成自然语言，最终反馈给用户。NLG 模块主要基于规则，将各个系统动作映射生成自然语言表达，此过程的每个动作都配有相应的回复话术作为生成文本，包括固定文本和通过函数及 API 调用返回动态结果两种。目前，为了实现机器回复的多样性，团队正在积极尝试使用基于深度学习的神经网络模型技术来生成回复话术，以提升交互友好度。

六、难点与挑战

(一) 难点：

1、创新性

自然语言需要具有强大的创新活力，需实时引入对最新概念、表述和意义的表达能力，常见于新词以及旧词新意的出现。比如：科创板。

2、递归性

递归性 (Recursion) 是人类语言的最重要的特性。比如：政府部门层层转发通知导致的“转发《上海市教育厅关于〈教育部关于做好放假安全工作的紧急通知〉的通知》的通知”的标题，两个通知，读起来就非常费力。

3、多义性

语言中本身存在不确定性，同一句话存在大量同音字和一词多义的现象，即一个字或词往往兼顾多个词义。同样的意思可以使用多种句式表达，且同一句话在不同的上下文语句中，表达的意思不同。一句话，调整一个词或者调整顺序及语调，也可能表达意思相差甚远。比如：

单身狗单身的原因有两个：(1) (单身狗) 谁都看不上；(2) 谁都看不上 (单身狗)。

4、主观性

不同的个人经历和认知水平带来强烈的主观性，且往往出现不流畅、错误、重复等情况，而对于机器人来说，这句话都是一个意思。例如：A 说你这是什么意思？B 回答没什么意思，意思意思。

5、社会性

不同社会地位的人在相互交流时，地位低的人会从语言风格上适应地位高的人，而地位高的人则不会主动调整自己的语言风格适应别人。

(二) 挑战：

1、结构语义表示空间构建

现在的语义表示方案中，符号表示过于粗略，无法考虑语言符号背后反映的丰富语义信息；而分布式表示虽然具有更强大的表示能力和自由度，但目前只能通过特定任务下的数据学习，建立满足特定需求的语义表示，缺少可解释性，鲁棒性差，通用性和迁移性不足；同时，这些与人类丰富的语义表示能力相比，仍差之千里，未来需要探索更强大的结构化语义表示空间。

2、多模态复杂语境理解

人类并非孤立地使用语言，通常会夹杂着复杂的语境。在多义的语言单元情况下，往往需要其外部的复杂语境信息进行消歧：例如字的多义性利用所组成的词来消歧；词的歧义性需要所在的句子来消歧；句子的意思时常要放在语篇或对话语境中，甚至需要复杂的知识来辅助理解。

然而，语境是开放的，也是多模态的，如上下文句子的文本信号、对话者的语气等语音信号、所处环境的视觉信号等往往会存在差异。因此，理解多模态复杂语境是目前技术难以突破的难点问题。

七、总结及展望

证通云智能外呼系统的应用，将自然语言理解和深度学习能力与现有规则判断相融合，致力于为行业机构提供精准文本解析的用户意图识别，深度语义理解的智能化问答。与此同时，力争实现在及时响应用户需求的基础上，根据流程

模型策略，智能化引导用户完成任务，提升成功率，从而实现低成本高效率的一站式外呼服务管理体系。而针对目前遇到的技术实现难点和挑战，将作为云外呼系统的建设目标，深入探究技术方案，优化对话系统策略及技能，助力企业实现高效、精准的云智能外呼服务，最大程度降低企业运营成本。

0 行业观察 bservation

9 实时计算系统建设经验分享

10 数据中心智慧园区思考

11 海通证券混合金融云平台生态体系建设之云原生迁移服务



实时计算系统建设经验分享

肖钢、李宁、张海园 / 中信建投证券股份有限公司信息技术部 zhanghaiyuan@csc.com.cn



实时计算技术已经应用到广告、电商、游戏、文娱等各个领域，比如电商网站实时分析用户属性，基于分析结果给客户推送相关商品；网络游戏实时分析玩家数据，进而对游戏参数和平衡性进行调整。本文重点讲述中信建投证券利用实时计算技术搭建实时计算系统的过程和经验。

1. 概述

1.1 什么是实时计算

实时计算是指对无限数据做无界数据处理，一般都是针对海量数据进行的，延迟要求为秒级。实时计算主要分为两部分：数据的实时入库和数据的实时计算。

用户属性，基于分析结果给客户推送相关商品；网络游戏实时分析玩家数据，进而对游戏参数和平衡性进行调整。

中信建投证券将其应用到金融领域，搭建了一套实时计算系统，实时汇总客户在股债基、理财、期权、两融、贵金属等业务的资产，为客户提供全方位的资产查询和分析服务。

1.2 应用领域

目前，实时计算技术已经应用到广告、电商、游戏、文娱等各个领域，比如电商网站实时分析

1.3 实时计算系统简介

得益于逐渐成熟的实时计算技术，中信建投证券于 2018 年底建设了实时计算系统，为员工

和客户提供实时资产查询服务，2019 年又加入了实时盈亏、预期收益、理财资产可用日和可取日、计息天数等指标。

实时计算系统把客户在集中交易、两融、OTC、期权、黄金五大系统（以下简称“五大系统”）内的数据资产汇总到一起，并与五大系统的数据同步变化。

2. 建设经验

2.1 系统架构

五大系统包含的客户数量是千万级别的，如此多的客户，在开市期间的各种操作会产生海量数据，如何在如此大规模数据的条件下满足实时数据处理的需求，是实时计算系统必须要解决的问题。

开市前，实时计算系统从五大系统采集清算后的 T-1 日数据作为基础，开市后，实时计算系统采集五大系统变化的数据，加工后将其更新到基础数据里。为了便于理解，笔者在系统原有架构的基础上做了简化，详见图 1：

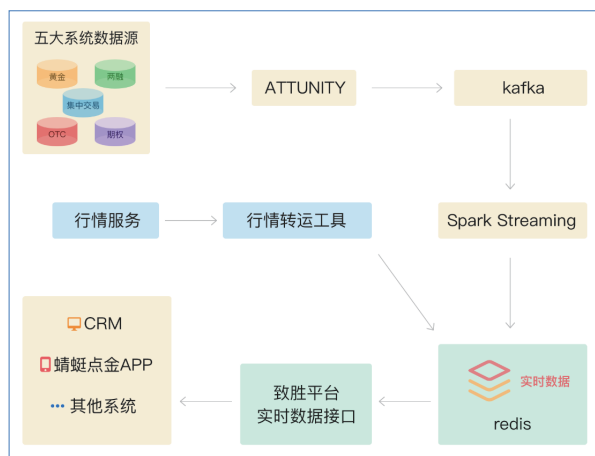


图 1：实时计算系统结构图

图 1 中实时数据流向如下：

1. Attunity 从五大系统数据源采集盘中产生的实时数据推送到 Kafka；同时，行情转运工具从行情服务获取行情数据，推送到 Redis；

2. 实时计算系统集群的 Spark Streaming 按

先进先出的顺序消费 Kafka 内的数据；

3. Spark Streaming 将实时数据加工后更新到 Redis 集群内的基础数据；

4. 其他系统访问致胜平台上的实时数据接口；

5. 实时数据接口从 Redis 集群内获取最新的客户数据，并将其转化为 json 格式的数据返回给其他系统。

系统建设过程中，项目团队还论证过另一种方案，即直接从五大系统数据库采集数据，汇总后提供给访问者。但经过讨论分析后，发现这种方案存在诸多缺陷：

1. 由于安全原因，五大系统不能直接暴露数据库给实时计算系统，这就需要五大系统开发专门的接口供实时计算系统使用，接口的开发和维护工作量很大；

2. 即使五大系统提供了查询接口，接口被调用时，五大系统也要从全部客户的数据里将被查询客户的数据检索出来，而且不止查询一个表，查询效率低下；

3. 此方案需要专门的服务器来运行，必须要投入一定数量服务器和网络资源，并不比其他方案需要投入的成本低。

2.2 关键技术

2.2.1 实时数据采集

系统建设初期，项目团队选择了 Oracle GoldenGate（简称“OGG”）作为实时数据采集工具。

OGG 软件是一种基于日志的结构化数据复制软件，能够实现大量交易数据的实时捕捉、变换和投递，实现源数据库与目标数据库的数据同步。Oracle GoldenGate 通过解析源数据库在线日志或归档日志获得数据的增删改变化，再将这些变化应用到目标数据库，实现源数据库与目标数据库同步、双活。

一期系统建设完成后，项目团队发现 OGG

给运维团队造成了很大的工作量，每次五大系统数据库维护或升级，运维团队都要耗费大量时间来恢复 OGG 服务。因此，项目团队决定用一个更合适的数据同步工具—Attunity Replicate（简称 Attunity）来替换 OGG。

Attunity 是一款基于数据库日志的异构数据库变化量同步工具。它采用 B/S 架构，通过浏览器图形化界面、鼠标点击拖拽的方式进行数据源、目标表的配置和维护，支持包括 Oracle、SQLserver、DB2 等十几种同构和异构数据库之间的跨平台数据同步。此外，它还支持并行工作，可以自定义并发度，这使得它比 OGG 延迟更低。

比于 OGG，Attunity 部署和维护更简便。Attunity 无需在源端或目标端操作系统上安装软件，只需一台独立的服务器安装 service，基于 WEB 的维护页面简单方便，只需要很少的操作即可完成数据同步配置，极大地减少了运维的难度和工作量。

2.2.2 中间数据存储

五大系统在开市期间产生海量的实时数据，即便是运行在集群上的 Spark Streaming 也无法在数据产生的瞬间将其完全消费，因此，就需要一个缓冲池来临时存放实时数据。综合各方面因素考虑，实时计算系统选用 Kafka 作为缓冲池。

Kafka 是一种高吞吐量的分布式发布订阅消息系统。实时计算系统架设了专门的集群来运行 Kafka，该集群有足够大的内存空间作为缓冲池存储实时流数据。Attunity 作为生产者，向 Kafka 实时流数据；Spark Streaming 作为消费者。为了保障客户的每一步操作都能精确地同步到实时计算系统，系统严格按照先进先出的方式消费 Kafka 内的数据。

2.2.3 实时流计算

因为实时计算系统对时效性要求很高，所以必须有一个能快速处理数据的工具。项目团队选用 Spark Streaming 来承担数据处理的任务。

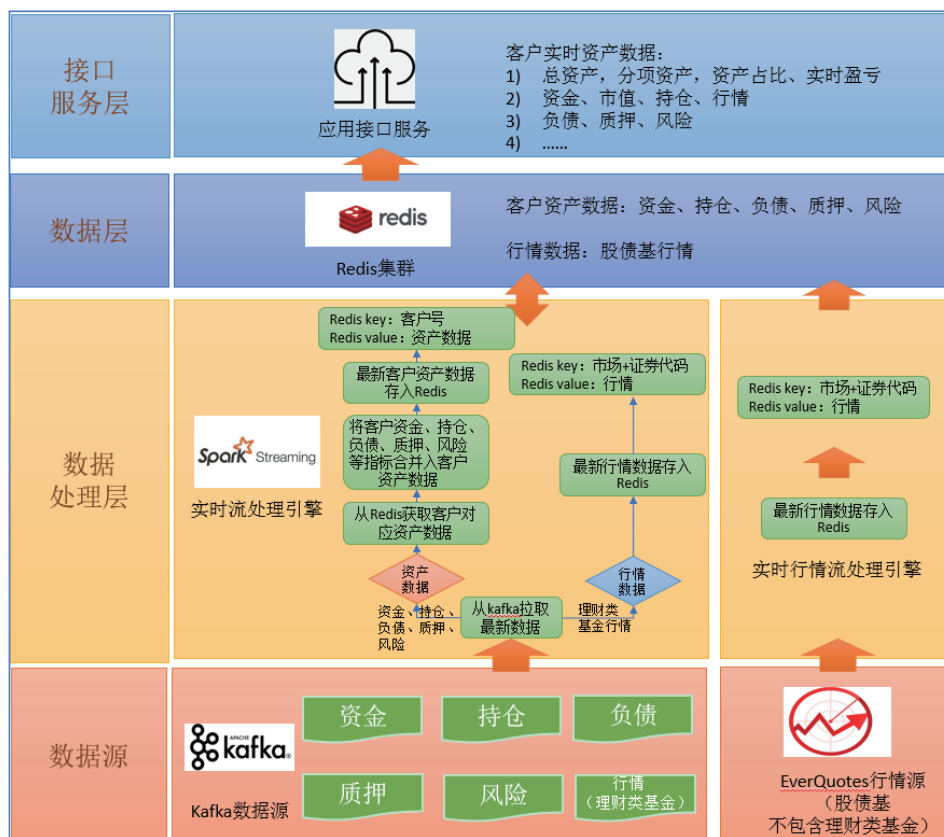


图 2：数据处理逻辑

Spark Streaming 是 Spark 核心 API 的一个扩展，可以实现高吞吐量且具备容错机制的实时流数据处理功能，这种特性能满足实时计算系统低延时要求。

开市前，实时计算系统先将 T-1 资产推送到 Redis 里，开市后，Spark Streaming 持续从 Kafka 里消费数据，然后对其进行处理。Spark Streaming 内部的数据处理逻辑如图 2。

图 2 中的数据处理层 Spark Streaming 与上下游工具配合，承担了实时流处理的工作。以客户开市期间的一笔股票交易为例，Spark Streaming 处理该客户的资产过程简化后如图 3。

接口服务层负责将客户资产按照相关业务规则分类。为了给调用者提供一个大而全的资产汇总数据，系统在接口服务层把客户在五大系统里的资产数据汇总到一起，为每个客户生成了一条以客户号为唯一标识的 json 数据，返还给调用者。

接口返回的 json 数据内包含资产类型按币种划分包含：人民币，美元，港币。其中，有美元和港币资产的客户较少，业务相对简单；有人

民币资产的客户较多，业务种类繁多。当上述各类资产单独存在于各自的系统时，资产和负债情况比较简单，无需繁琐计算即可得出其总资产、总负债、净资产等数据，但是如果要把五大系统的资产和负债汇总到一起，就要从全局的角度考虑各种情况，比如：客户在认购场外开放式基金时如何处理普通账户上的在途资产；客户用现金理财产品购买股票时如何修正其可用资金；客户提前赎回理财产品后如何更新其可用资金。为了解决这些复杂的问题，项目团队做了大量的调研和梳理工作，最终形成了一个包含五大系统资产的客户全景资产模型，图 4 是该模型的中客户人民币资产的内容。

2.2.4 对外服务接口

经过 Spark Streaming 加工汇总的客户数据，需要有一个容量大、响应快的容器来容纳；当需要对外提供数据时，还需要一个高并发的服务平台。中信建投证券自主研发的致胜平台^[1]能够满足上述需求，项目团队基于致胜平台开发了一套接口服务。

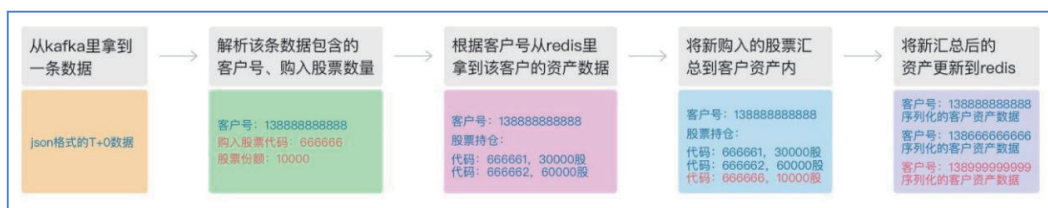


图 3：股票交易同步逻辑

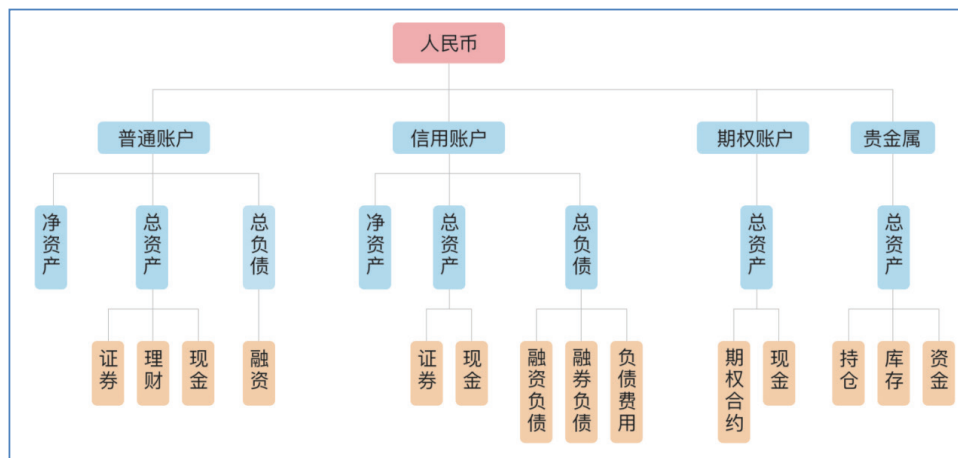


图 4：客户人民币资产概览



图 5：接口调用原理

致胜平台为实时计算系统分配了一组 Redis 集群，专门用来存储汇总后的客户数据。为了最大限度的提高 Redis 集群的吞吐量，数据在 Redis 集群内以序列化的形式存储。当其他系统调用系统接口时，接口从 Redis 里抽取相应数据并加工成 json 格式，返回给调用者。Redis 里的数据保持实时更新，接口服务随时等待被调用，两者互不干涉。接口调用示意图见图 5。

3. 未来展望

随着技术的日益发展和项目团队的不断探索，实时计算系统未来还可以应用到更广泛的领域，比如股票实时策略推荐、两融业务实时预警等。中信建投证券将依托于实时计算系统为客户提供全方位的金融服务，其金融科技水平将也迈上一个新的台阶。

^[1]“致胜平台”是遵循统一的技术架构和规范的一整套微服务体系架构，包括技术中台、业务中台和数据中台三个部分，详见文章。

数据中心智慧园区思考

黄飞 / 深圳证券通信有限公司 huangfei@sscc.com

蓝汉明 / 深圳证券通信有限公司 lanhm@sscc.com



近年为解决证券期货行业基础设施资源不足等问题，行业内各核心机构大力投入及建设行业数据中心，随之而来的数据中心园区管理也成为各数据中心管理者不得不思考的问题。本文从智慧园区概念出发，结合当前园区管理工作，对智慧园区建设进行思考。试图提出一个整体的、全方位的、多层次的平台架构和建设思路，解决园区管理普遍存在的需求不明、系统独立、不够便捷等问题，从而为更好的支持行业数据中心业务发展。

一、概述

随着人工智能、区块链、云计算、大数据、物联网、机器人、智能终端和 5G 等新一轮科学技术的迅猛发展和应用落地，当今世界正经历一场深刻的技术变革。新技术渗透到各行各业和每一个角落，逐步成为日常生产和生活的一部分，也成为未来园区发展的驱动力。

近年来在各行各业智慧化转型的浪潮中，园

区也从传统型园区向智慧型园区演进。近年证券期货行业也在积极部署和建设数据中心，要想打造“世界一流行业信息技术中心”，势必要洞察数据中心智慧园区发展的规律，摸清数据中心智慧园区发展的脉络，思考数据中心智慧园区建设的方向。

到底何为“智慧园区”？如何实现“智慧”？未来走向何处？成为智慧园区建设中不得不思考的问题。

二、智慧园区定义

何为“智慧园区”？

从技术角度讲，智慧园区是运用新一代信息与通信技术，具备迅捷信息采集、高速信息传输、高度集中计算、智能事务处理能力，深度融合人事物的有机生命体和可持续发展空间。

（一）运用新一代信息与通信技术

“数字中国”、“新基建”、“互联网+”等策略部署，在金融科技行业乃至国家信息化发展战略中不断的被提出和强化，其中都包括了“以技术创新为驱动”这一理念。所以，在园区转型和建设智慧园区的过程中，运用新一代信息与通信技术具有极为重要的战略意义。

新一代信息与通信技术包括了人工智能、区块链、云计算、大数据、物联网、机器人、智能终端和 5G 等技术。其中以云计算、大数据、5G 为代表的全新基础设施和人工智能、机器人、智能终端等全新解决方案相结合，为智慧园区建设奠定了技术基础。

（二）具备“四种能力”

智慧园区和传统园区相比较，具备迅捷信息

采集、高速信息传输、高度集中计算、智能事务处理能力。同样，这些需要具备的能力也是传统园区当前发展的瓶颈和未来转型时迎接的挑战。

信息采集要求迅捷、全面、精确，借助各种工具、传感器实现对园区内各类元素和活动的“全面感知”。首先需要用到 IoT、GIS、机器人、智能终端等技术，园区内各元素的物理信息转化为数字化信息。

信息传输要求高速、完整，借助统一网络实现园区各类元素的“泛在联接”。需要用到 5G、Wi-Fi6、Ethernet、Optical、e-LTE、ICT 等网络技术，同时要打通 IT 与 OT 间的通道，使园区各元素及信息可互联互通。

集中计算要求集中、有效，借助一体化的信息与数据平台对数据进行计算，实现“数据增值”。需要用到边云协同、大数据等技术，使各系统数据可融合丰富和实现分析治理。

事务处理要求智能、高效，通过对园区信息的充分掌握和智能处理，高效地解决园区内出现的各类需求与问题。需要用到计算机视觉、语音处理、机器学习、人工智能等技术。

（三）持有“三种特性”

未来的智慧园区将是融合了各种元素，并可



图 1 数据中心智慧园区“四种能力”

以像“生命”一样能有机调动这些元素，做到可持续发展的绿色、高效物理空间。

未来的智慧园区是人事物深度融合的有机整体，是可持续发展且具备良好可扩展性的，既满足当前生产需要又保持一定的发展潜力。通过创新的技术和服务，不断的提升园区的技术能力和空间价值。同时，未来智慧园区也是可自我更新的，具备完整的自我进化动力。通过各元素的协同和促进，不断的迭代更新技术能力和业务产品。例如在有限的空间资源下，园区将会朝着服务和技术方向拓展，通过 SaaS 提升基础设施的利用率和潜在价值，通过创新技术服务提高园区的核心竞争力和品牌价值。

三、智慧园区建设

如何实现“智慧”？

针对大型数据中心智慧园区项目规划，应结合国内外以往经验，提高立足点来进行统一规划，需要从数据中心初期的空间设计、产品选型、网络布局就介入，一开始就要在“安全、稳定”的基础上提出“便捷、有序、高效、绿色、增值、持续”的新理念为建设“智慧型”园区做出良好铺垫。

（一）智慧园区蓝图框架

首先，确定数据中心智慧园区蓝图框架，用

于指导数据中心智慧园区规划与建设。

愿景，结合组织意向以高立足点提出方向性的长远导向，并可根据具体定位、项目、业务进一步细化。

目标，紧密围绕数据中心“人事物”三个要点，从人员体验、事务处理和物资利用出发，实现“便捷 & 有序”、“高效 & 绿色”、“增值 & 持续”的三大核心目标，并进一步在此基础上深化设计子目标。

业务，从数据中心运营角度出发确定“动环”、“IT”和“园区”三个常规的业务系统，并提出在实际的业务和平台中分别从信息、网络、计算、事务等四个层面实现“全面感知”、“高速互联”、“集中计算”和“智能处理”等理念，指导后期智慧园区平台建设方向。

基础，作为数据中心重要“物理”和“数字”元素包括：人员、业务、空间、设施、联接、计算等六类，用于指导信息采集的方向和内容。

（二）智慧园区架构

其次，建立数据中心智慧园区架构，明确智慧园区实现的功能和要点。

智慧园区架构从“感、通、融、用”四个方面出发，按照“纵向解耦、横向融合”的原则，构建“基础终端层”、“网络传输层”、“数字融合层”和“智慧应用层”四层架构。每层着重实现的功能和具备的要求都各不相同。



图 2：数据中心智慧园区“三种特性”



图 3：数据中心智慧园区蓝图框架

基础终端层主要是通过各类终端或传感器对园区各个元素和活动行为进行“全面感知”，实现园区“全覆盖、精细化”的管理。对终端设备要求其尽量做到 IP 化、无线化、集约化，并可简单处理部分信息（智能终端）。

网络传输层主要通过各种网络设备将园区各个元素接入，实现“泛在联接”并做到安全管控。

对网络层要求安全、高速、可扩展，并且在不同网络区域间做到可互通、可控制、可审计。

数字融合层主要通过一体化信息平台对各专业系统进行集成，对数据进行融合和计算。在终端和网络夯实的基础上，深度感知园区内人事物各元素的信息和变化，并实现无缝对接与协同，在园区平台呈现信息化映像和综合态势感知。

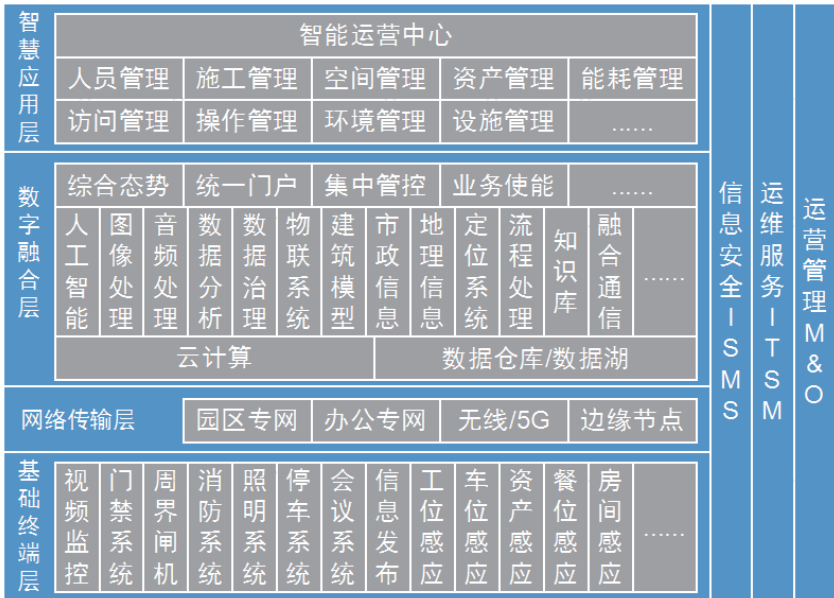


图 4：数据中心智慧园区架构

智慧应用层是根据实际业务需求对园区进行整体的智能管理。在园区数字信息被融合、丰富、深度挖掘的前提下，结合园区日常业务，实现多资源、多系统、多数据、多部门的高效协同工作。从而改变过去单一维度解决问题带来的偏差和资源不足等问题，突破各种壁垒与条件限制，使数据广泛被管理者所用，协助智能管理，提供智慧服务。

（三）智慧园区方案

最后，提出数据中心智慧园区方案，确保智慧园区架构中各功能有相应技术得以实现。

元素终端，实现园区万物数字化、信息化转型，使园区物理世界与数字世界相融合。依靠感应器、智能终端、视频等设备，构建智慧园区信息基础。

联接技术，实现园区系统互联互通，使园区各元素、各系统相连接。依靠 5G、Wi-Fi6、红外、蓝牙、NFC 等技术，搭建智慧园区通信桥梁。

软件平台，实现园区数据融合、丰富，使园区各信息、各数据相支撑。依靠大数据、云计算、人工智能、边云协同提供超强能力，实现智慧园区计算核心。

业务应用，实现园区事务便捷、智能，使园区各事务、各服务相合作。依靠终端开发、流

程梳理、可视化展示等方式，实现智慧园区应用窗口。

四、智慧园区展望

未来走向何处？

从园区的发展趋势来看，智慧化的探索和建设将是一个长期的过程，智慧园区的建设会一直持续下去。但整体来看，未来的智慧园区将会聚焦在空间的利用、用户的体验、数据的价值、业务的创新、运营的精细化和发展的可持续上。未来智慧园区将会让服务更智能、管理更高效、环境更舒适、能耗更绿色。

（一）便捷通行，智能服务

结合计算机视觉、人员定位和红外成像等技术实现便捷通行。在办公、生活区域（财务、监控除外）不再必须持有各类门禁卡和充值卡，实现刷脸通行和支付。门禁系统、视频监控会采集面部信息，对每位员工、客户、来访人员面部进行识别，了解其位置、路径和状态。对闯入未授权区域的人员进行告警；对施工人员进行监控；对来访人员进行参观引导和讲解；对体温异常人员或非正常行为进行识别。通过人脸识别和互联网支付实现餐厅、咖啡厅、住房、娱乐设施、生

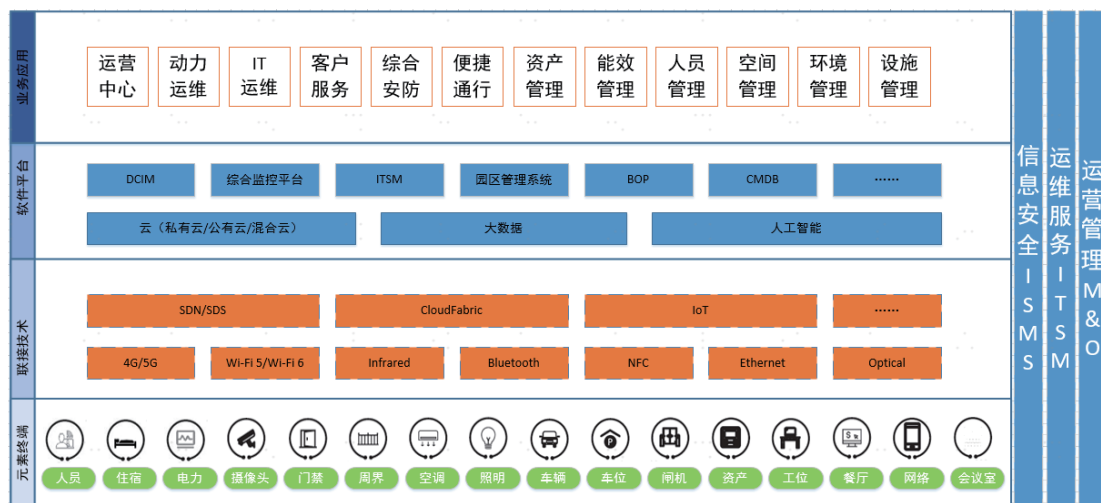


图 5：数据中心智慧园区方案

活缴费（住宿区水电、网络、电视）等快捷支付。通过人脸识别和机器人等技术实现仓储、物流的快速核验及收发。

同时，通过一体化数据与信息平台提供智能服务。对办公、生活服务不在需要繁琐的申请和操作，只用在智慧园区的平台或 APP 上一次性提交申请。实现园区会议、房间、车位、餐位、运动设施、（非固定）工位、订餐、访问参观、物流投递等服务的预定，平台会根据客户需要联动相应的设备、人员和部门进行精细化全方位服务，不需要申请人过度的介入和处理。如预定会议室后，平台会自动根据会议人数、等级和要求自动开启灯光、投影、空调、远程会议等设备，通知会议管理人员准备相应的水牌、茶水、纸笔、椅子等物品，提前传送会议资料和会议信息给参会人员。如预定访问参观后，平台会根据来访人员的行程、级别、要求自动申请临时车辆、车位、会议室、房间、用餐、门禁权限、网络接入等，并自动通知接访和安保人员，自动开启展示厅和信息发布迎宾标语，安保人员可通过 AR 自动识别来访人员身份，不需要刻意拦截、询问。

（二）安全管理，主动维护

传统园区智能化水平低，技防手段弱。主要依靠人防和物防，虽然安装了大量摄像机，但是仅能用于事后追溯，无法提供智能分析，无法在海量视频中，为安保人员提供有价值的报警预判信息。安防子系统分散，无联动，感知体验差，缺乏发布和获取信息的平台，访客“门难进、路难找、车难停、费难缴”，属典型的被动防护。

智慧园区建构的综合安防应用，将变被动为主动，通过数字化手段对园区进行精准运营，打通各个离散的安防子系统，实现数据共享、安防联动。结合 AR、机器人、无人机、人脸识别、人员定位、Wi-Fi6、5G 等技术提升园区安保级别，对园区人员、车辆、环境进行全方位管理。AR 技术协助安保人员快速识别来访人员身份；通过

人脸识别和人员定位，对人员访问区域进行管控；通过无人机和机器人实现自动安防巡检。

同时，通过一体化的数据与信息平台消除各设备缺陷和短板，降低漏报或误报情况。通过园区各传感器、监控、周界、闸机、门禁和智能设备等跨平台智能联动，结合历史数据、人员使用和维护内容对设备设施进行主动维护提醒。

（三）能效管理，绿色节能

通过资源整合，重点打造可视、可管、可优化的层次递进的能效管理平台 and 场景应用。能效管理的中心思想是“能效受控”，在不影响建筑舒适性的前提下，降低能源消耗，提升能源使用效率。

要实现能效受控，首先要全员参与，包括集团领导、用能人员、维护人员，让所有人知道能源运行情况、运行效率、各类指标等，将能耗可视化。再进一步实现能源运行的调配、控制、事件处置等，寻找最优能效控制方案。

能效控制方案，通过采集和监控建筑中各类用能系统（例如：配电、照明、空调、给排水等）的实际运行状态，找出关键耗能点和异常耗能点，制定成熟可靠、行之有效的能效控制方案。例如：根据室内光照强度，自动调节灯光数量；会议室人员感应，长时间无人时，自动关闭灯光、空调；具备远程控制和管理，结合实际采集数据，应用 PDCA 管理循环，进行微调，最终寻找到适应四季变化的、符合实际状况的最优能效控制方案。

（四）协同办公，无碍联接

通过 5G、AR、VR、云计算和移动会议等技术，突破地域限制实现远程办公、远程会议、远程参观、在线交流、远程技术支持等应用，达到与现场面对面同等效果，确保信息传递的完整、真实和便捷。

通过语言处理、语音追踪、电子名牌、云端存储、无线投屏等技术提高会议工作效率，支持

多场地、多终端、多方式会议和办公接入。

通过 5G、机器人和 AR 技术实现远程技术维护和支持，实现虚拟信息与现实物品的充分融合，实现专家“永远在场”的现场感，并通过信息记录联动工单系统进行维护和处理。

（五）设施管理，提升效率

在智慧园区中有大量的设施和设备。不同类型的设备有不同的用途，如能耗类的电表，环境类的烟感设备，安防类的摄像头等。在传统园区中，一般都是由不同的垂直子系统来分别管理设备信息的。因此设备信息要重复在多个系统中录入，难以保证信息一致，难以实现统一管理。各个垂直子系统之间也难以使用设备信息进行交互，导致难以实现跨子系统的业务协同。为拉通不同应用场景和不同垂直子系统的设备信息，更方便高效地实现各自系统与设备交互的功能，有必要通过统一的“设施管理应用”进行集中管理。

园区的运营人员能够管理设施基础的信息，包括设备编码、设备名称、设备分类、对外型号、设备位置、GIS 坐标（经度、纬度、高度）、设备使用状态、设备重要级别、设备描述信息等。

五、结论

智慧园区的发展方兴未艾，未来将会是一个长期而持续的事情。智慧园区体系和知识涉及面广、技术方案多样、实施环节复杂。所以，在执行部署时不应一成不变，应以实际需求为导向，按照“急用先行、成熟先上”的原则，制定合适的智慧园区建设计划，推动重点技术的研制、应用和推广。

文本仅为行业数据中心智慧园区建设及发展提供一个初步的思路，阐明一种想法，构建一个框架，意图展开一个讨论的话题和方向。对于其中实际的场景和详细的措施，还有待后续更多的理论探讨和实践检验。

海通证券混合金融云平台生态体系建设之云原生迁移服务

罗秋清、陆颂华、王朝阳、张真真、崔恒春 / 海通证券股份有限公司

海通证券混合金融云平台以多云的视角，提供给用户全自主化的操作体验，在集团内取得了非常良好的效果，全面助力海通证券数字化转型。云管理平台对私有云、行业云和公有云进行了异构管理和统一编排，不但解决了资源统一管理、分配、编排等控制层面的需求，也解决了数据层面的互联互通问题。本文主要基于混合金融云平台中的云原生迁移服务，介绍用户如何将业务系统在不同的云平台间灵活迁移的技术原理与最佳实践。



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，海通证券建设完成中国证券行业首个应用创新的统一纳管、编排私有云、行业云和公有云的混合金融云平台。

混合云平台以开源自主可控的 OpenStack 框架为基础，集成了国产硬件 SDN 实现云网联动、集成了国产负载均衡实现流量负载智能分摊、集成了集中存储和分布式存储实现性能容量按需分配。此外，还提供了完善的云平台生态体系，如云原生迁移服务、负载均衡即服务、高精度时钟源服务、DNS 服务、YUM 源服务、监控告警服务、补丁服务、病毒库服务和正版激活服务等。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，使用的部门有集团总部、子公司和分支机构。

根据 Gartner 对云管理平台的研究¹，云管理平台²应具备 11 大核心模块（如图 1）：资源自动化交付和编排、服务请求管理、治理与策略、监控和度量、多云管理、成本分析与优化、容量和资源优化、云迁移和灾备、服务等级管理、安全和认证、配置自动化。其中，云迁移和容灾是云管理平台中非常重要的模块之一。但是目前能借鉴的云原生迁移的成功案例较少，究其原因可能有两点：

1、目前对云管理平台的很多需求仅停留在“管”，很少涉及到“通”，所以云迁移的需求被忽略了。但是随着公司内部多云战略的不断深化，对多云管理平台的理解日臻完善，云上的迁移和容灾的需求将变得越来越迫切。

2、目前市面上的迁移产品多为传统容灾服务商开发，更多的考虑是基于物理机迁移场景，对于云原生的支持不够友好。一方面没有针对云平台资源的 API 充分利用，造成无法持久化的“集成”，扩展性差；另一方面没有提供相应的 API 接口，从而无法“被集成”。

基于以上两个方面，海通证券根据自身业务需求，在云管理平台中集成了面向多租户的云原生迁移服务，不但便于用户将自己的业务系统在不同的云平台间灵活迁移，也进一步完善了云管理平台的功能。

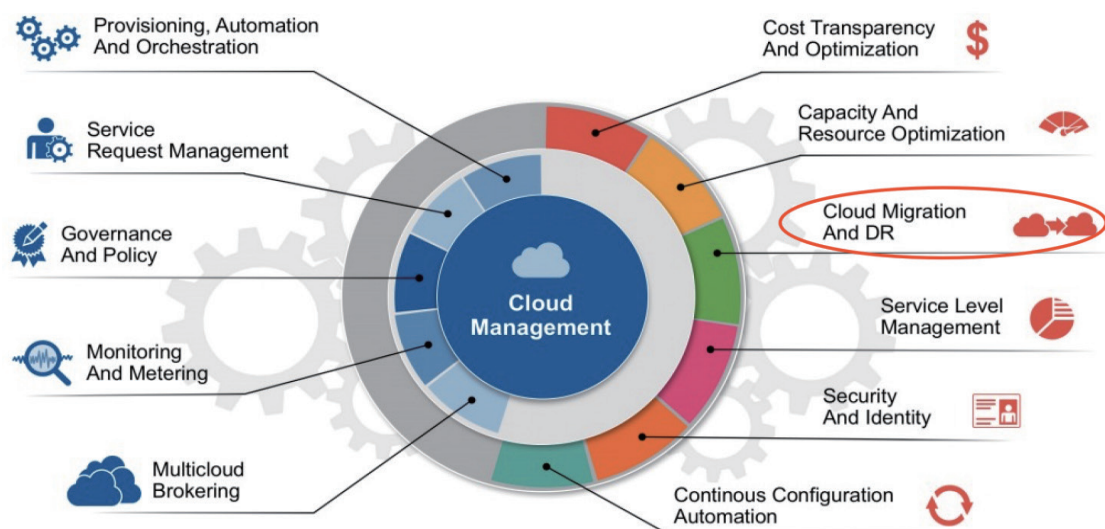


图 1：Gartner 定义的云管理平台核心模块

¹ <https://www.gartner.com/document/3873016>

² 云管理平台：Cloud Management Platform, CMP

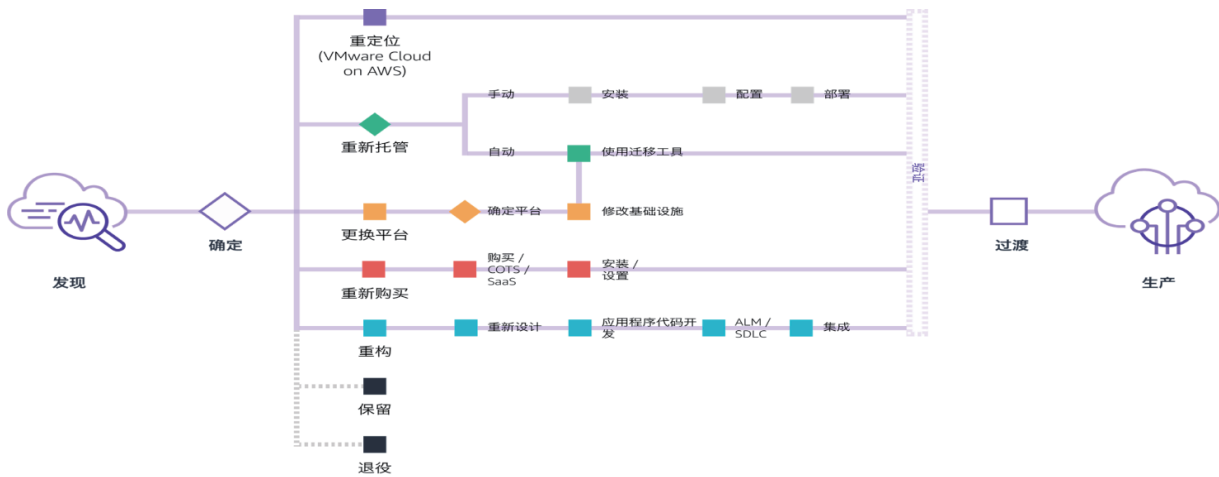


图 2：AWS 定义的 7 种云迁移策略

2 海通证券云原生迁移服务简介

根据 AWS 对云迁移策略的定义（如图 2），常见的七种策略（7R）³有：重新托管（Rehosting）、更换平台（Replatforming）、重新购买（Repurchasing）、重构（Refactoring）、退役（Retire）、保留（Retian）。在混合云环境下，对于用户成本最低、效率最高的方式就是重新托管（Rehosting）方式，即原封不动照搬过来，用户从操作系统到应用系统都不会发生任何改变，可以在线迁移，迁移之后可以立即使用。

海通证券混合金融云平台中提供的云原生迁移服务则具备以下特点：

- 1、实现“搬家（Re-Host）”效果。
- 2、实现“热迁移”效果。在用户迁移时，采用块级差别量复制技术实现准“热迁移”，用户无须关机即可完成在线迁移。迁移过程中，用户业务系统不中断，保证业务连续性。
- 3、云原生“集成”能力。通过云平台接口驱动层完成与多云之间的对接，充分利用云上资源的特性，在不借助任何特殊设备情况下进行迁移。
- 4、“被集成”能力。云迁移服务本身的功能全部以 RESTful API 方式提供给云管理平台或第三方应用，使得用户无须单独操作迁移平台，通

过云管理平台与办公流程联动，在资源配额满足需求的前提下，仅需几步简单配置便可在线完成迁移。

5、面向多租户设计。用户订阅迁移服务，自助完成迁移动作，无须管理员介入。

6、智能驱动适配，实现一键式迁移效果。迁移服务会智能的判断目标云平台的虚拟化类型，选择合适的磁盘、网卡等驱动进行驱动注入，保证迁移后的系统在目标云平台能够正常使用。

3 云原生迁移服务的技术原理与最佳实践

3.1 技术原理

要实现云原生的“热”迁移方式，需要解决三个关键问题：

数据怎么读出来？如果要想实现整体搬迁的效果，仅仅依靠文件级别的同步是无法实现的，所以必须从块级别实现同步，才能满足需求。另外，在同步过程中，还需要支持全量和增量的同步，以实现“热”迁移的效果。

数据怎么传？这里面需要解决两个维度的问题：传输方式和传输后数据的存储。只有把数据直接存放在云原生的资源中，才能实现最终的数据变云主机的过程。所以我们采用云平台的块存

³ <https://xiandin.s3.cn-northwest-1.amazonaws.com.cn/MAttC/Asset+1+eBook+Strategies+for+Accelerating+Migration+to+AWS.pdf?aliId=eyJpJjoicmg1Y3FVZ1pleWNhdFY1RyIsInQiOjPMG5vVm9JOXdzM1wvaUNlcXR2ZHI0dz09In0%253D>

储服务作为数据存放的目标。

数据如何变成云主机？在此过程中需要解决如何将块设备变为云主机，并且保证系统启动正常。这里我们通过云平台原生 API 接口实现高度自动化的调度流程，并实现智能驱动适配手段，解决了迁移最后一公里的问题。

3.1.1 HyperMotion

云原生迁移控制服务，主要提供用户操作的 UI 界面 /API 接口，做源端、目标端连通性检查等工作。一般一个集群 /Region 仅需要安装一个 HyperMotion 服务即可。

3.1.2 HyperGate

云原生迁移代理服务，主要完成目标存储管理与智能驱动适配等工作。一般一个租户需要安装一个或者多个 HyperGate 服务。

存储解耦：解除源系统和对应存储的绑定关系，将任意云平台存储作为迁移存储提供给应用主机使用。

智能驱动适配：解决跨平台场景中普遍存在的驱动转换等问题，提供驱动自动适配、注入功能，且无需人为介入。

云平台调度处理：用户不再受存储绑定，不受异构平台驱动转换限制，可以任意选择云平台。

3.1.3 块级别复制：

复制的对象是操作系统 + 应用程序 + 应用配置 + 数据，迁移时不受文件系统类型约束，是实现整体（业务级）挂载恢复的基础。

3.2 最佳实践

海通证券在提供云原生迁移服务的过程中，遇到了一些问题，也通过解决这些问题积累了一定的经验。

3.2.1 共享迁移租户

HyperGate 是云原生迁移服务中负责一个租户内接收数据和云资源 API 接口调用的核心组件，寓意是通往“任意云”的“任意门”。根据租户数量和网络互通的复杂性来看，有两种可选的实现方案：

方案 1 是云平台部署一个 HyperMotion 服务（如图 3），用户每个需要迁移云主机的租户拥有独立的 HyperGate 组件，HyperGate 只负责该租户内云主机的迁移工作。这种方案的优势是用户在迁移云主机的同时已经形成了数据隔离，不会出现跨租户中转的情况，即云主机从源租户直接到达了目标租户。但是缺点也很明显，首先是资源浪费，每个需要迁移云主机的租户都需要预留部署 HyperGate 组件的资源，这部分资源的配额费用是否要用户承担；其次是管理复杂，源租户

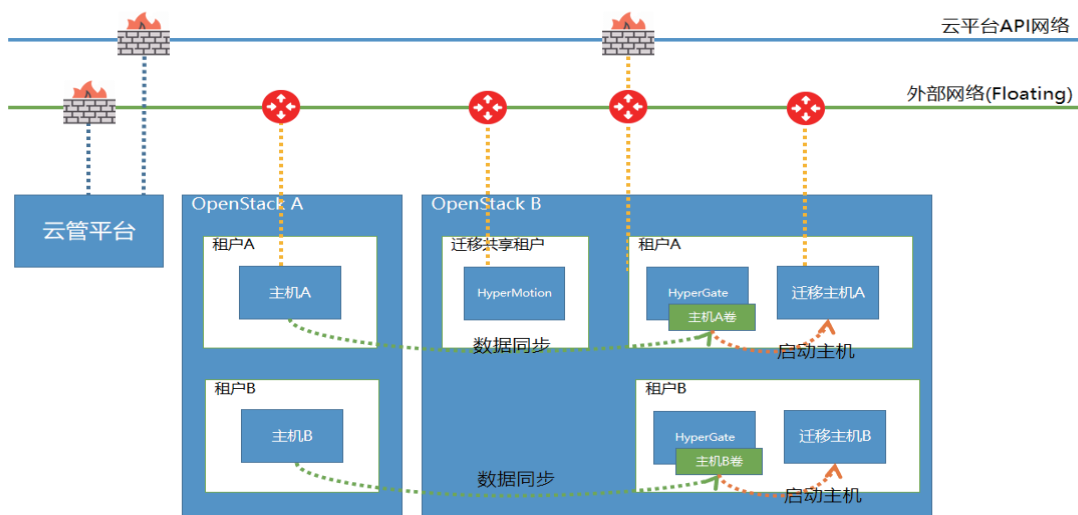


图 3：每个租户分别部署 HyperGate 的方案

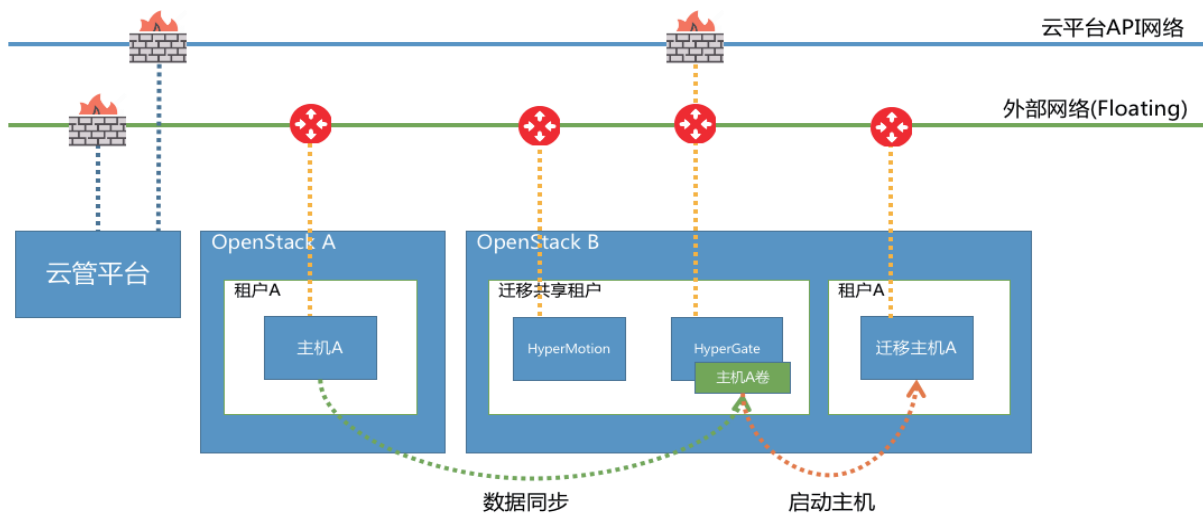


图 4：共享租户部署 HyperGate 的方案

和目标租户都要进行外部网络开通，而且下发的网络策略和云平台资源需要在什么时间进行回收是一个不得不慎重考虑的问题。

方案 2 是由管理员创建一个共享租户，并且部署一个 HyperMotion 服务和一个 HyperGate 服务（如图 4），用户所有需要迁移云主机的租户共享此 HyperGate 组件。这种方案的优势，首先是无需过多的迁移组件资源，而且这部分资源配额的费用可以由平台承担；其次所有待迁移主机的租户与共享租户的网络互通策略比较简单，可以很好利用 SDN 的优势，定制模板并批量下发。缺点是迁移的云主机都需要在共享租户中转；一个 HyperGate 组件无法高效支持多个租户同时迁移；而且每个租户下需要创建一个迁移管理员用户。

海通证券混合金融云平台目前有 800 多个租户，如果使用方案 1 的话将会对平台资源造成极大的浪费，而且网络策略回收将会是一个复杂的问题。方案 2 虽然需要在共享租户中转，但只在启动云主机时，将卷转移到目标租户内进行启动，不会影响迁移的效率；通过在共享租户中安装多个 HyperGate 组件便可支持多个租户同时进行云主机迁移工作，并且结合迁移限速，可以有效控制多个租户同时迁移对平台的影响；租户中的用

户可以通过云管理平台自动创建，仅需修改少量业务逻辑代码即可完成。资源和管理复杂度的考虑，最终选择了方案 2。

同步数据时，由迁移共享租户内的 HyperGate 服务负责接收数据并写入云硬盘中，每次同步后创建云硬盘快照。当用户需要进行迁移测试或业务切换时，利用 OpenStack Volume Transfer 功能将创建好的卷，转移到指定租户内，之后调用 Boot From Volume 接口进行启动。

3.2.2 Ceph 存储的解链

默认情况下，OpenStack 使用 Ceph 链式快照，启动主机后，如果要清理迁移过程中产生的临时资源，由于依赖关系造成资源残留。为了解决这个问题，在启动主机时，提供两种方式：迁移测试和迁移切换。如果是迁移测试，则默认使用快照方式，保证测试的启动速度。如果是迁移切换，则调用 cinder backup 服务，完成解链的过程后再清理资源。

3.2.3 支持限速、停止同步及断点续传

在迁移过程中，如果同步速度过快势必会影响用户业务系统的运行，所以需要提提供限速功能来保证业务系统流量正常。

如果在迁移过程中，用户业务系统有一个突发情况希望终止迁移，则需要使用停止同步



图 5：云原生迁移服务

的功能来结束目前的迁移。在适当的时间点重新开始后，将从上一次结束同步的位置继续同步数据。

3.2.4 与 CMP 集成

按照 Gartner 的定义，云原生迁移服务已经成为了海通证券混合金融云平台的核心模块之一，与云管理平台深度集成（如图 5），在配额内，通过基本信息收集 -> 预检 -> 注册 -> 同步 -> 启动 -> 清理等 6 步自助服务，即可完成虚拟机迁

移的工作。

4 建设总结

通过提供云原生迁移服务，使得混合金融平台的用户获得了很大的价值。如：简化流程，在配额内用户自助服务；提高效率，简单点击几步即可完成操作；节省资源，平台自动完成资源的生成与清理。



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

国际证监会组织发布咨询报告《市场中介与资产管理人对人工智能的应用》

此前，国际证监会组织（IOSCO）曾就如何监管市场中介和资产管理机构使用人工智能（AI）和机器学习（ML）发布了一份拟议指南。2020年6月25日，IOSCO委员会就这一指南发起了公众意见征询，截止日期为10月26日。IOSCO表示，上述两类技术的使用能够有效提高机构运营速率、降低投资成本，但是同时也会给金融市场带来不小的风险。如今，越来越多的监管机构将目光锁定在了这两项新兴技术之上。而IOSCO的拟议指南建议，监管机构应该确保使用AI和ML的市场中介和资产管理拥有明确的治理管控框架、透明的信息披露、稳健的开发流程和足够的知识储备。这份拟议指南并非强制要求，但是IOSCO仍然希望成员机构在法律和监管框架内认真考虑指南中的各项措施。

国际清算银行探讨全球稳定币的建立和运营

在二十国组织（G20）一份题为《加强跨境支付：全球路线图的基础》的新报告中，国际清算银行的支付和市场基础设施委员会（CPMI）指出，当前的跨境支付领域存在着高成

本、低速及透明度的问题，并呼吁采用“全球方法”解决这些问题。国际清算银行表示，（稳定币）这个重点领域比其他领域更具探索性，并且可能会走得更长远。但全球稳定币的建立和运营成本预计会很高，且目前无法量化；由于全球稳定币安排主要基于尚待大规模测试的新技术，因此它们可能面临重大运营风险，需要适当管理和缓解。但是，该行也提供了许多解决方案，据称它们可以帮助实现全球稳定币，如根据国际标准、国内法规和政策设计和实施全球稳定币安排；实施国际协调的监管和监督方法；在所有司法管辖区的跨司法管辖区澄清法律待遇等。

国际清算银行发表关于央行数字货币的工作论文

8月24日，国际清算银行（BIS）发表了题为《中央银行数字货币的兴起：驱动因素，方法和技术》的工作论文。文章指出，中央银行数字货币（CBDC）受到了前所未有的关注。然而，发行CBDC的动机因国家而异，政策方法和技术设计也各有不同。BIS调查研究了发展CBDC的经济和制度驱动因素，评估CBDC的开发设计工作，并建立了一个关于发行技术方法和政策立场的综合数据库。同时，BIS将中国、加拿大和瑞典的数字货币作为典型案例进行了深入分析。

欧美动态

FDIC 发起现代化银行财务报告专题竞赛

2020 年 6 月 30 日，美国联邦存款保险公司 (FDIC) 宣布开展一项银行财务报告建模大赛。共有来自全美的 20 家技术机构应邀参与这一竞赛，他们都是金融服务、数据管理、数据分析、人工智能 / 机器学习领域的行业领军力量。各方将在未来几个月开发相应解决方案并提交 FDIC 审议，帮助银行更加高效地提交财务报告，为 FDIC 提供更及时、更详细的行业健康数据，并促进对单个银行的更有效监管。

CFTC 科技咨询委员会召开电话会议，讨论虚拟货币等议题

美国商品期货交易委员会 (CFTC) 下属科技咨询委员会 (Technology Advisory Committee, 简称 TAC) 于 7 月 16 日举行远程电话会议，就自动和现代交易市场、分布式账本技术和市场基础结构、虚拟货币、网络安全等问题展开讨论。TAC 的成立，是为了就影响美国市场完整性和竞争力的各种监管和市场问题征求意见并向 CFTC 提出建议。TAC 促进了 CFTC 与美国市场、贸易公司、市场参与者和商业最终用户之间的沟通。

美联储分享其在数字货币领域的研究和实验

8 月 13 日，美联储强调了其为增加对央行数字货币相关的机会和 risks 的理解而开展的研究和实验项目。为确保提供安全、可访问和高效的美元支付系统，美联储一直在对最新的支付

技术进行研究和试验。例如，美联储技术实验室 (TechLab) 正在扩大与数字货币和其他支付创新相关技术的试验，以进一步提高美联储对支付技术的理解并为相关政策的制订提供支撑。TechLab 是一个由董事会和联邦储备银行工作人员组成的多学科团队，在支付，经济学，法律，信息技术和计算机科学方面具有专业知识。此外，波士顿联邦储备银行正在与麻省理工学院的研究人员合作，以建立面向中央银行使用的虚拟数字货币。该研究项目旨在支持美联储在评估中央银行数字货币系统的安全性和效率。

纳斯达克推出自动反洗钱调查技术

9 月 16 日，纳斯达克 (Nasdaq) 在其官网宣布推出了基于云的自动反洗钱 (AML) 调查技术。Nasdaq 指出，在极端情况下，AML 事务监视系统每月可能会触发多达 20 万至 30 万条警报。为解决告警过多的情况，许多银行在其 AML 交易监控系统中收紧了参数，或增加了其他评分机制，从而减少了警报。而纳斯达克的自动反洗钱调查技术可从交易监控系统中提取告警数据，整理分析所需的所有必要数据，然后复制负责的人为决策过程，为所有告警提供清晰、可审计的理由，并且无论数量多少，处理速度均以秒为单位。

欧洲央行发布报告，强调警惕大型科技公司在金融和云服务领域的潜在影响

2020 年 8 月欧洲中央银行 (ECB) 发布了一份有关数字金融服务的咨询文件。该文件指出，如今欧盟金融服务提供商对非欧盟关键服务

和技术基础设施（例如云计算）提供商的依赖性成为了欧盟数字金融服务面临的一个重要挑战。尽管总部位于欧盟的技术机构正在努力崛起，但是仍然没有出现一家能够与美国等国家大型技术商相匹敌的机构。未来，会有越来越多的金融机构将其关键信息迁移至 Amazon Web Services、Microsoft Azure 和 Google Cloud，也就是说大型高科技公司在金融服务市场的份额将持续走高，这会导致严重的集中化风险。如果一旦出现网络攻击和各种操作失误，将给欧盟金融系统稳定性带来巨大威胁，因此必须对其加大监管。除此之外，欧洲央行还关注了人工智能、分布式账本技术等联合外包的风险，而在线和数字平台能耗大幅上升带来的环境可持续问题也在欧洲央行本次的讨论范围之内。

欧洲银行业管理局发起监管科技行业调查，预计明年发布结果

8月12日，欧洲银行业管理局（EBA）启动了一项监管科技专项调查，希望包括金融机构、信息通信技术服务第三方提供商在内的利益相关方积极参与，分享其对使用监管科技解决方案的看法和经验。据悉，本次调查将主要关注以下四个方面：1）反洗钱与反恐怖主义融资；2）信用资格评估；3）安全标准与要求（信息安全、网络安全、支付服务等）合规情况；4）监管报告。本次调查截止日期为2020年9月30日，最终结果预计2021年上半年发布。

欧盟委员会与欧洲区块链联盟合作开展区块链领域测试验证

9月24日，欧盟委员会数字创新与区块链团队官网信息显示，欧盟委员会将与欧洲区块链

联盟（EBP）合作测试欧洲区块链服务基础架构（EBSI）中的区块链和数字资产用例，包括数据可移植性、B2B数据空间、智能合约和数字身份在健康、环境、交通、能源和其他关键领域的应用。EBSI是欧盟委员会和EBP的一项联合计划，旨在使用区块链技术在整個欧盟范围内提供跨境数字公共服务。此外，欧盟委员会还计划在2021至2022年间启动一个泛欧洲区块链监管沙箱。

英国政府启动金融科技行业审查，预计明年初结束

7月20日，英国财政部经济大臣 John Glen 宣布，英国政府将对该国的金融科技发展展开专项审查。据悉，本次“金融科技战略审查”由 Worldpay 前首席执行官 Ron Kalifa 领导，希望借此机会明确金融科技发展现状与潜在机遇，为行业机构、政策制定者和监管机构提供帮助。本次审查计划最早于2020年财政预算中提出，最终结果预计明年初发布。据统计，英国金融科技行业总规模约为70亿英镑，累计提供就业机会6万多个，2019年吸引风险投资41亿英镑，自2015年以来市场规模已经增长近70%。

英国金融行为管理局发布第六批监管沙盒机构名单，申请与入选数量双下滑

7月27日，英国金融行为管理局（FCA）发布了第六批监管沙盒入选机构名单。本次共有来自零售银行、支付、零售借贷等领域的68家机构提出了申请，22家机构最终入选，申请和入选机构数量同比均出现下滑。市场分析人士认为，盈利前景不明朗、资金支持不足和成功经验

难以获得等因素是导致这一现象的主要原因。相比于前五次沙盒计划，本次计划更加关注“普惠金融”和“绿色金融”相关的创新服务。

法国央行开展央行数字货币试点

7月20日，法国银行宣布下一步将与埃森哲、欧洲清算系统、汇丰银行、Iznes、LiquidShare、ProsperUS、Seba Bank 和法国兴业银行合作进行央行数字货币银行间结算测试。未来，法国银行将就以下方面展开测试，包括：1) 探索将金融工具（不包括加密资产）交换为央行货币的新方法；2) 测试中央银行数字货币的结算，以改善跨境支付的执行条件；3) 修改提供中央银行资金的安排。法国银行表示，希望从这些试验中汲取的经验教训能够直接促进欧元体系对央行数字货币价值进行更全球化的反思。

法国金融市场管理局设立科学顾问委员会

8月20日，法国金融市场管理局（AMF）宣布建立了由来自学术界和金融界的知名人士组成的科学顾问委员会，以扩大其研究和战略情报能力。科学顾问委员会将有三方面的职责：向 AMF 提供有关金融领域正在进行的学术研究的信息；确定可能对 AMF 的工作产生影响的发展变化；进行与监管机构关注的问题相关的项目。

德国财政部提出基于区块链的证券数字化解决方案

德国联邦财政部（BMF）发布了有关引入电子证券（eWpG）的草案，其中包括一项数字化

证券的区块链解决方案。德国证券法协会在新闻稿中指出，该法案将对德国证券法和相应的监管法进行修订，重点关注区块链战略。此外，该法案草案还提供了更高的监管透明度，即联邦电子监管局将与 eWpG、德国银行法（KWG）和主要证券存管机构达成协议，跟踪“去中心化登记册”的启动和维护，以作为新的金融服务。该新闻稿称，通过采用区块链和其他新技术对法律框架进行的拟议变更旨在巩固德国的商业中心地位，并扩大“透明度，市场诚信度和对投资者的保护”。

西班牙证交所携手巴塞罗那科技城创立金融科技中心

2020年7月20日，西班牙证券交易所（BME）宣布成为巴塞罗那科技城全球合作伙伴。作为合作的一部分，BME 将在巴塞罗那建立一个占地1000平方米的金融科技中心，为初创企业、投资方、风险投资公司、传统银行和保险公司提供合作交流空间和资本输入新渠道，在整个价值链中产生协同效应。

维也纳证券交易所上市比特币和以太坊交易产品

9月2日，维也纳证券交易所上市了21Shares的比特币和以太坊交易所产品交易产品“ABTC”和“AETH”。维也纳证交所发展部主管 Thomas Rainer 强调，上述两项产品上市后，有经验的本地投资者在交易比特币和以太坊时可以享受“受监控、受监管和透明的交易，并通过其常规经纪商获得实时信息和安全的结算”。这次上市意味着加密货币现已可在德国、奥地利和瑞士这三个德语区国家交易。

亚太动态

韩国金融服务委员会推出促进金融大数据应用的新措施

7月1日，韩国金融服务委员会（FSC）宣布推出数项推动金融大数据应用的措施，其中包括：对 CreDB 系统进行改造：增加数据集类别，扩充服务器容量，以及建立一个专门服务融合数据集的数据库、促进基于 CreDB 数据与其他非金融数据集之间的数据融合研究；在韩国金融电信和清算研究所（KFTC）新建开放金融大数据系统，相关方可在 2021 年上半年通过网站、API 以及根据需求使用这一服务。

新加坡区块链支付网络 Ubin 项目完成开发

2020 年 7 月 13 日，新加坡金融管理局（MAS）与淡马锡（Temasek）携手发布了《Ubin 项目第五阶段：实现广泛的生态系统机遇》专题报告，这也标志着 Ubin 项目第五阶段，即最后一阶段的顺利完成。本次报告研究了区块链技术在不同行业商业应用中的使用情况，指明了这些应用与已开发的基于区块链的支付网络原型集成后所带来的各种益处。目前，这些益处已经通过参与 Ubin 项目第五阶段测试的 40 多家金融和非金融公司举办的研讨会得到了验证。据悉，Ubin 项目由 MAS 于 2016 年 11 月引进，专门探索以区块链和分布式账本技术（DLT）为基础的支付和证券交易清算、结算，特别是银行间支付。

农业银行、交通银行先后成立金融科技子公司

7月28日，农业银行发布公告，其附属金融科技子公司农银金融科技有限责任公司正式成立。

农银金科注册地为北京通州，注册资本 6 亿元。主要业务方向是运用金融科技手段，开展技术创新、软件研发、产品运营与技术咨询。8月25日，交通银行附属金融科技子公司交银金融科技有限公司在中国（上海）自由贸易试验区临港新片区正式注册成立，注册资本 6 亿元。交银金科的主营方向为以金融科技为手段，开展软件研发、技术咨询与创新、信息系统集成服务等。至此，国内已有 12 家商业银行成立金融科技子公司，除邮储银行外，其他五大行均已正式成立金融科技子公司。

国家网信办印发《数字中国建设发展进程报告（2019 年）》

9月10日，国家互联网信息办公室印发《数字中国建设发展进程报告（2019 年）》。《报告》梳理总结了各地区、各部门 2019 年信息化发展情况，分析了数字中国建设面临的新形势、新挑战和新机遇，提出了努力方向和发展重点。《报告》包括五部分内容：1）数字中国建设向纵深发展；2）数字中国建设发展水平稳步提升；3）各地区信息化发展成效不断显现；4）携手构建更加美好的数字世界成为全球共识；5）在危机中育新机，于变局中开新局，扎实推进数字中国建设高质量发展。

标普发布《亚太金融科技投资报告》

8月25日，标普发布了 2020 年第二季度的《亚太金融科技投资报告》。报告指出，2020 年第二季度亚太地区金融科技企业融资总量较上一季度上涨 9.1%，达到 14 亿美元。其中，东南亚和澳大利亚表现最为抢眼，分别融资 4.55 亿美元和 3.69 亿美元。东南亚地区中，电子钱包类业务的投资吸引力最大。而在澳大利亚，数字银行获得的融资最多。

2020年四季度《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年四季度征稿主题如下：

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），请通过投稿邮箱 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 12 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）



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>

内部资料 免费交流

本资料仅为内部交流使用，本季度印750册，编印单位为上交所技术有限责任公司，面向证券期货行业发送，印刷日期为2020年11月，印刷单位主人印刷厂。

部分图片或文字来源于互联网等公开渠道，其版权归属原作者所有。如有版权相关事宜，请发送邮件至ftt.editor@sse.com.cn