|  |  |  |  |
| --- | --- | --- | --- |
| 需求名称 |  | 需求版本 | Vnn.1.6.4 |

**运维信息分析系统二期建设项目**

**需求说明书**

**需求申请部门：系统运维中心**

|  |  |  |  |
| --- | --- | --- | --- |
| 拟制 |  | 日期 | 2018*-8-10* |
| 评审人 |  | 日期 |  |
| 批准 |  | 日期 |  |
| 签发 |  | 日期 |  |

**修订记录**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 编号 | 版本号 | 修订内容简述 | 修订日期 | 修订人 | 审核 |
| 1 | 1.0 | 初订 |  | 陈俊 |  |
| 2 |  |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  |  |
| 5 |  |  |  |  |  |
| 6 |  |  |  |  |  |
| 7 |  |  |  |  |  |
| 8 |  |  |  |  |  |
| 9 |  |  |  |  |  |
| 10 |  |  |  |  |  |

1. **业务需求概述**

**1.1需求背景**

运维信息分析系统建设项目是实现我行“全行十三五发展规划131号任务：推进运维大数据的分析和应用”的主要着力点。项目从2017年开始建设，至2018年底完成一期建设，应用大数据和机器学习技术，打造了准实时的运维大数据存储、计算和检索环境，基于其上研发了系统画像、智能监控、日志分析、运维内控这四大应用场景。

当前功能、使用情况和发挥作用简述如下：

1、准实时的运维大数据存储、计算和检索环境，实现各运维工具、流程系统整合运维数据，支持前端自助式类SQL语法配置分析、监控视图或预警邮件任务。目前系统运维中心各技术条线已自助开发了200多项报表、仪表盘，应用不同的应用场景，主要是用于历史分析、实时监控。

2、通过构建全面、微观、多维度的系统画像，满足运维人员、业务人员多维度、全视角的服务需求，夯实数字化经营基础，推动运维数据集约化共享应用。此场景支持科技十三五规划98号任务“提升系统运行处理能力”，通过健康智能评分、运维指标数据查看、运维指标数据下钻分析，支撑 “SEE”、“容量管理”等工作开展，同时辅助管理员发现或预测系统运行中的容量瓶颈。目前运维指标类型包括事实、模型和预测，涵盖tps、交易量等指标，系统指标覆盖大部分系统。

3、整合系统的历史性能指标数据和实时性能指标数据，根据对历史性能指标数据分析，总结系统运行指标规律，分辨出周期性趋势特征，动态建立一段时间窗口内的历史基线，以此监控系统的实时性能指标数据。此场景提升了系统监控预警能力，解决了传统监控工具固定阈值模式监控的灵活度、敏感性不足的问题，现已覆盖绝大部分系统的CPU、内存、IO、TCP链接等多个指标实时监测及文件系统使用率、数据库空间使用率等二十余项容量预警。

4、整合系统运行监测数据，通过统计分析、关联分析以及机器学习算法，对系统运行数据进行规整、计算，能够将系统的运行情况和知识库中的系统运行数据进行聚合和分类，支持运维管理。此外，通过与其它监控、自动化工具跨界联动，针对部分场景探索故障自愈。此场景目前实现自动出具全辖系统性能容量自动化评估报告，辅助晚间值班人员技术值班告警分析和决策。

5、辅助运维内控管理工作，及时防范生产运行风险隐患，提升运维内控管理水平。通过对接现有的系统，获取运维人员在系统的上得具体操作信息，对运维人员、登录主机、账号等进行聚类分析，发现管理员行为异常变化，对运维人员的生产操作命令、操作时间等进行统计分析和合规性识别。通过对各类内控数据的关联分析，及时发现运维风险隐患。此场景目前实现全辖运维人员画像，辅助内控管理员识别高危操作命令。

综上，运维信息分析系统已成为系统运维中心数据分析的最重要技术手段，在运维工具蓝图体系建设中，目前也作为唯一的建设系统承担了数据分析决策域的各块职能，为持续推进系统功能优化，按照“全行十三五发展规划131号任务：推进运维大数据的分析和应用”的计划，拟定本项目的建设需求目标。

**1.1.1业务需求分类（单选）**

|  |  |
| --- | --- |
| **需求大类** | **需求细类** |
| 业务类 | □新的产品及业务  □现有产品及业务升级：  □现有产品及业务流程优化  □现有功能升级  □业务产品推广支持 |
| 管理类 | □新管理要求（附管理办法、行业批文，参照标准等）  □已有管理的优化（简述已有流程，优化后的流程、优化思路） |
| 运营操作类 | □新增业务流程  □已有流程优化 |
| 监管要求类 | □新监管要求  □原要求补充或调整 |
| 核算规则类 | □新监管要求  □原要求补充或调整 |
| 系统运维管理类 | 🗹系统效能提升  □基础设施建设  □软硬件升级 |
| 数据运用类（固定报表、查询、多维分析（OLAP）、统计分析和数据挖掘） | □新增数据运用需求 |
| □现有数据运用需求的优化 |

**1.1.2业务需求重要程度**

|  |  |
| --- | --- |
| 编号 | 重要程度 |
| 1 | □非常重要 |
| 2 | 🗹重要 |
| 3 | □一般 |

**1.2需求的业务可行性和必要性**

**1.2.1可行性分析**

项目一期的建设实现了大部分运维数据的整合与多种应用分析，主要内容概括为三方面：

1. 建立实时、通用的海量数据分析和科学家探索平台，消除运维数据分散各处形成的数据孤岛、统一数据保存周期不一的问题，降低查询分析比对的难度，为全体运维人员提供灵活友好的数据分析平台。
2. 适用于运营场景的算法模型体系并工程化落地，系统建立了一套面向交易分析、面向IT基础软硬件的运维辅助，面向人员的内控管理三大板块的算法模型体系,包括22余项自研或改进算法、500余组模型，并将之工程化落地。
3. 试点监控、自动化和大数据联动场景，实践以配置为纽带，将系统的硬件、软件、网络、存储和备份等十余种监控和自动化工具串接建立实时监控感知，结合算法模型体系处理和计算，实现从监控感知到数据分析乃至自动化处置的线上闭环处理。

同时，同业也在积极探索此领域研究：

1. 银联从12年开始hadoop生态圈组件研究和建设，近三年其云平台万台服务器的监控从原有传统监控产品被大数据开源组件解决方案所替代，实现秒级上百监控指标的准实时采集、聚合计算与可视化展现，目前考虑进一步拓展应用场景。
2. 工商银行建设运维大数据平台和数据指标池，实现运维数据的统一供给和挖掘分析，推动数据中心智能化转型，包括流数据平台+应用监控、应用系统性能容量平台、网络数据包解析等方面建设，在网络应急智能化、磁盘性能画像预警、主机交易率预估模型、智能问题定位方面有所应用。
3. 交通银行于16年底至今开展了运维大数据的建设，将部分运维数据进行了采集清洗，利用开源的hadoop生态圈技术进行存储和计算，通过神经网络等算法进行异常监测应用，弥补传统固定阈值监控短板，于18年三季度开始二期建设，期望持续扩大数据规模、深入算法模型在运维领域的应用。
4. 农业银行通过生产运行自动化三年规划建设形成48个运维工具，存在数据孤岛，导致数据无法整合全局分析和共享，18年发起IT数据驱动运营分析及流程优化整合，希望引入大数据采集、预处理、存储、分析工具，在此基础上逐步建立运维数据关联性分析工作功能模型，通过在线和离线分析和计算，为运维人员提供各类运维决策支持提供参考，达到通过高效的数据分析辅助生产运维决策的目的。
5. 华夏银行提出基于运维大数据分析的生产系统风险智能预警模型，通过采集28套重要信息系统的历史监控数据对生产系统风险智能预警模型进行验证，证实可以发现故障隐患，为系统管理员主动采取事前应对措施，避免风险和损失提供了决策信息等。

综上，我行及同业都积极开展运维大数据方面研究与应用，部分场景上已取得了良好效果，且此领域是后面几年会逐步加大投入的热点，技术实现上更加成熟、案例将更为丰富。

**1.2.2必要性分析**

必要性主要包括以下三方面的内容：

1. 支持内部重点工作

实现或支持2项全行十三五发展规划任务，包括98号任务-重要生产系统运行能力持续提升和131号任务-推进运维大数据的分析和应用；实现或支持5项系统运维中心19年工作计划，包括运维大数据建设、加强运维数据治理、加强运维管理和流程优化、分行运维管理、推动运维人员岗位调整；满足4项监管要求，包括商业银行十三五发展规划白皮书、网络安全法、监管数据报送、监管审计；直接满足或间接支持系统运维中心三处16个小组的58项数据分析领域需求。

2、满足外部监管指引

十三五期内，内外部对数据中心的运维管理工作提出了新的要求。《中国银行业信息科技“十三五”发展规划白皮书》中提到，加强运维大数据分析，利用运维大数据加强业务风险防控，探索利用大数据、智能化技术推动业务流程优化并支持业务创新；相应地，我行提出了“以客户为中心，科技引领，打造一流数字生态银行”为十三五规划目标，积极拥抱开源新技术，推进大数据的分析和应用，着力提升运维智能化水平。

3、系统运维中心转型发展的内在需求。

新形势下，数据中心未来将持续推进以“稳定、可靠、安全”为主题的IT运维管理体系建设，同时又要逐步向以“体验、效率、效益”为主题的IT运营分析阶段发展。IT运营分析的关键在于数据分析，本项目通过应用大数据、智能化等技术，建立PB级全行运维大数据平台，基于其上实现交易分析、智能监控、数据探索、运维内控和操作控制台功能，着力提升监控感知能力、大数据分析能力、运维自动化能力、数据可视化能力，努力实现科技从支持保障到与业务融合发展乃至引领业务的重要转变。

4、解决项目推广或持续建设过程中暴露的资源不足和数据质量的问题

一期已完成系统运维中心大部分生产运维系统的数据集中和汇总，开发出面向系统管理员的运维信息数据分析平台，以类SQL的方式提供查询接口，极大降低了数据分析的难度，能快速以数据可视化报表的方式反馈系统管理员系统状态。随着数据接入种类的不断丰富和数据接入量的不断增长，在实际应用中也出现了一些问题：

1. 存储资源不足，目前重要数据设置存储周期为半年，非重要数据存储一个月，超出生命周期则采用清除策略，导致数据分析训练的数据量不足，且随着设备数增加、新数据接入也支撑不久。
2. 算力资源不足，随着接入数据规模增大、实时和离线批量计算的需求增长，二期需要提升平台的算力，增强数据处理能力。
3. 数据质量不一，从各运维辅助工具采集来的数据因为数据格式、标准不一导致对同一事物描述不一，实时采集、解析过程中对数据的清洗能力有限，需要通过ETL调度、数仓建模来进一步规整数据，提升数据质量。

4） 为了支持今后向AIOPS转型，需要持续建设和提升运维大数据平台能力，提供稳定可靠的数据服务、充足的数据处理能力、安全灵活的数据API接口、算法打造的智能引擎等基础功能服务。

**1.3需求目标**

二期项目总体目标如下：

1、夯基础，开展数据治理、建立运维数仓、建设运维指标体系、强化数据安全控制，让数据完整、准确，为运维和业务分析提供基础数据支撑；

2、优框架，拓展平台框架功能、建立数据分析API接口，实现基础服务能力与应用场景开发解耦，提供强大、便捷、灵活的大数据分析处理能力；

3、增效益，持续探索应用机器学习算法，拓展对预警预测、故障诊断、画像建模等应用场景的支持，试点故障发现、故障定位、故障规避和故障处置；

4、扩容量，随着数据持续接入，容量满足至2020年6月的大数据存储和分布式计算资源，并对至2020年6月前到期的大数据产品组件版本升级。

**1.4需求范围**

**1.4.1行内机构适用范围**

|  |  |
| --- | --- |
| □全行性需求 | 🗹总行需求 |
| □分行个性化需求 | □海外机构 |
| □自贸机构 | □村镇机构 |
| □外包合作公司 | □其他 |

**1.4.2用户适用范围**

|  |  |
| --- | --- |
| □柜员 | □总行业务管理人员 |
| □对公客户经理 | □总分行高管 |
| □对私客户经理 | □运营操作员 |
| □交易员 | 🗹行内科技人员 |
| □分行业务管理人员 | □其他 |

**1.4.3客户适用范围**

|  |  |
| --- | --- |
| □对公客户 | □非本行代理他行客户 |
| □对私客户 | □行外监管部门 |
| □金融机构 | 🗹行内客户 |

**1.5预期时间要求**

|  |  |  |
| --- | --- | --- |
| **需求阶段** | **预期时间** | **是否必填** |
| 需求正式提交时间 | 2018年10月 | 必填 |
| 需求分析完成时间 |  | 非必填 |
| 提交业务测试时间 |  | 非必填 |
| 上线投产时间 | 2019年8月基础软硬件及数仓功能投产  2019年7月至2020年6月其他功能迭代上线 | 必填 |

**1.6与年度需求计划的对应关系**

|  |  |  |  |
| --- | --- | --- | --- |
| **是否列入年度需求计划(是/否)** | **需求编号** | **需求计划名称** | **本项需求的范围与需求计划的对应关系（全部/部分）** |
| 是 |  | 运维信息分析系统二期建设项目 | 全部 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**1.7质量保证要求**

|  |  |
| --- | --- |
| 需求实施会议 | □每周 □每两周 □每月 |
| 需求建设报告周期 | □每周 □每两周 □每月 |
| 是否参加科技开发部需求评审 | 🗹是 □否 |

1. **业务概述**

**2.1名词解释**

1）智能化：是指由现代通信与信息技术、计算机网络技术、行业技术、智能控制技术汇集而成的针对某一个方面的应用。从感觉到记忆再到思维这一过程称为“智慧”，智慧的结果产生了行为和语言，将行为和语言的表达过程称为“能力”，两者合称“智能”。智能一般具有这样一些特点：一是具有感知能力，即具有能够感知外部世界、获取外部信息的能力，这是产生智能活动的前提条件和必要条件；二是具有记忆和思维能力，即能够存储感知到的外部信息及由思维产生的知识，同时能够利用已有的知识对信息进行分析、计算、比较、判断、联想、决策；三是具有学习能力和自适应能力，即通过与环境的相互作用，不断学习积累知识，使自己能够适应环境变化；四是具有行为决策能力，即对外界的刺激作出反应，形成决策并传达相应的信息。具有上述特点的系统则为智能系统或智能化系统。

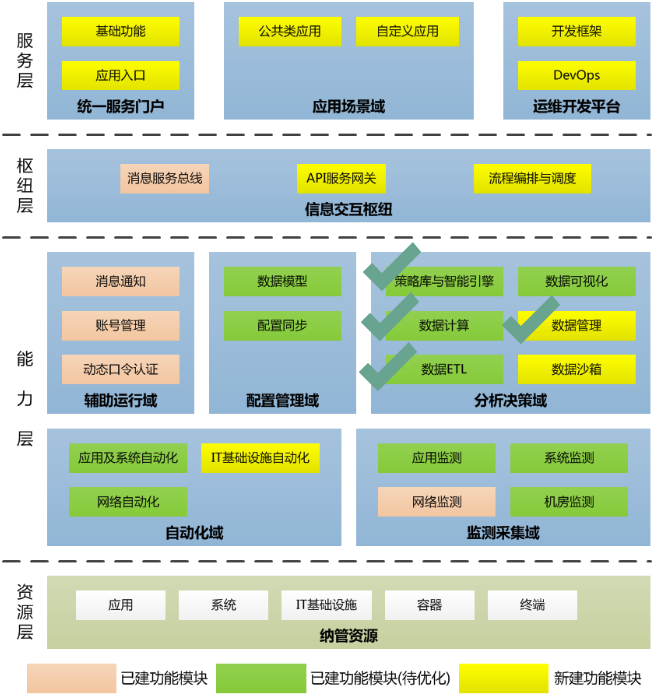
2）机器学习：机器学习是一门多领域交叉学科，专门研究计算机怎样模拟或实现人类的学习行为，以获取新的知识或技能，重新组织已有的知识结构使之不断改善自身的性能。它是人工智能的核心，是使计算机具有智能的根本途径，其应用遍及人工智能的各个领域，它主要使用归纳、综合而不是演绎。

3）学件：学件的概念由南京大学周志华教授提出，于2016年在FCS的Perspectives专栏，发表文章Learnware: on the future of machine learning，是一种性能良好的预训练机器学习模型，其具有一套解释模型意图和/或特性的规约，即学件=模型+规约，模型具有可重用、可演进、可了解的特性，规约能够给出模型的合适刻画，可以是基于逻辑的描述，和/或揭示模型所针对的目标的统计数据，和/或甚至一些简化了的、能够描述模型应用场景的训练样本。

**2.2业务架构**

系统运维中心对后续运维工具建设进行了详尽梳理和统一规划，如下为

规划蓝图：



本项目涉及其中打勾的四个模块功能的优化与新建，包括“策略库与智能引擎”、“数据计算”、“数据ETL”，“数据管理”。

在此大的框架下，设计本项目的业务总体架构图如下所示：

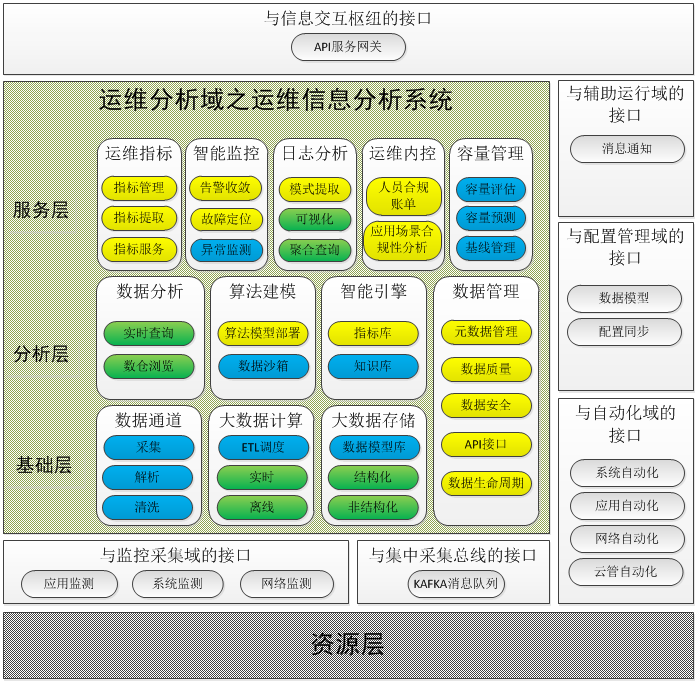


图 2.1业务架构图

1） 基础层

基础层主要功能是大数据基础功能实现，主要分为三个子模块：数据通道、大数据计算和数据存储。

（1）数据通道：一期项目完成了数据通道的建设，实现数据流的分发采集、数据流式处理和解析等功能，无缝对接集中采集总线KAFKA。随着数据接入种类和总量的不断提升，二期项目全面提升数据通道的承载能力、提升采集负载均衡能力、流式处理能力。

（2）大数据计算：一期开发了大量ETL的数据流，二期需增加ETL统一调度的功能，使得管理更加清晰，并提高任务的并发效率。

（3）大数据存储：二期项目计划开展运维数仓建设，将处理过的高质量的运维数据长久保存在数据模型库之中，为智能运维算法、数据分析提供扎实的数据基础。

2） 分析层

分析层的主要功能是将基础层所采集的数据做加工，合并，归一化等工作，将数据加工成各服务层所需要输入的多维数据。在一期项目的建设过程中，已完成了通过HUE对于底层数据进行实时的浏览和查询，并通过数据流的join等功能做了数据的加工等工作。

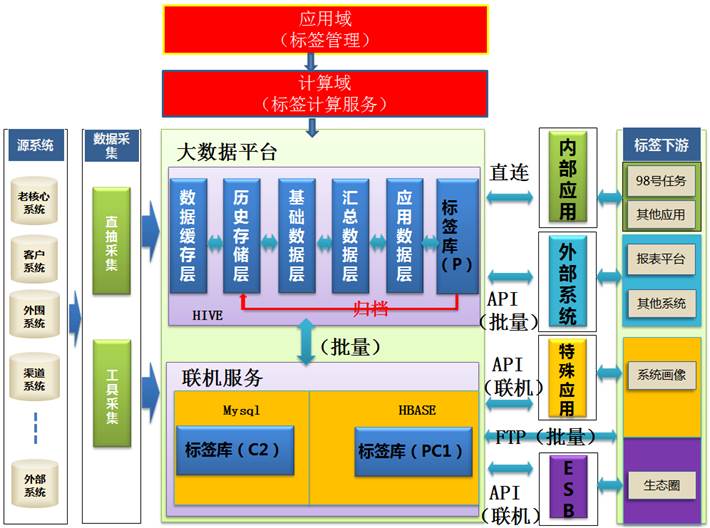
3） 服务层

服务层涉及分析层、基础层各功能模块的原子化封装、管理类页面配置等，各块内容通过持续开发丰富数据管理的API接口，注册到外部的API服务网关，联动其他域开放的API功能，进一步实现各应用场景的开发。

4） 外部接口设计

外部接口设计需要对接不同的外部运维工具，如自动化运维工具，配置域工具等。同时为了规范化数据管理，大数据平台在二期项目中进行API服务网管建设，统一对外进行数据交互及权限管理。

**2.3系统画像业务架构**



标签体系配置信息： Mysql

标签体系跑批信息（标签库）：Hive

系统画像场景所需标签信息：HBASE或kylin，部分静态标签数据后续也可考虑放入。

1. **功能性需求**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 功能说明 | | | | |
| 模块一级分类 | 模块二级分类 | 功能名称 | 功能说明 | 新增/修改 |
| 服务层 | 运维指标体系 | 运维指标管理 | 实现对运维指标的生命周期管理、要素管理，并建立多维度的运维指标视图，方便用户多维度、全景式查阅。 | 新增 |
| 运维指标提取 | 实现支持用户自助式简单配置，动态生成运维指标提作业，可视化对作业进行管理和监控功能。 | 新增 |
| 运维指标服务 | 运维标签数据对外提供联机、批量和互动查询服务，实现运维数据标签对外数据服务的监控管理。 | 修改 |
|  | 系统画像 |  | 基于API接口管理功能，调整原系统画像底层获取运维指标的接口、支持自定义图表展示、系统画像运维指标钻取、建立系统间关联图谱、优化原智能化评分算法、支持运维指标在CALL浦端的展示和使用。 | 修改。 |

**3.4 标签体系建设**

**3.4.1功能说明**

本期需求主要是基于已搭建运维信息分析系统中海量运维数据的历史沉淀和系统画像前期数据运维指标成果积累，进一步通过对不同口径、不同来源、不同结构的信息进行统一的加工整合和分类管理，利用大数据挖掘技术（例如统计算法、机器学习等）构建覆盖生产系统、生产资源的运维数据运维指标体系，规范数据运维指标设计，提高数据运维指标质量，实现数据运维指标持续建设、迭代优化、定期淘汰的闭环模式，持续发掘提升数据资产业务价值，为系统画像、容量管理、SEE、98号任务等多场景、跨板块运维指标应用提供联机和批量运维数据运维指标服务，满足运维数据分析人员多维度、全视角的数字服务需求，夯实数字化经营基础，推动运维数据集约化共享应用。

**3.4.2功能描述**

1. **运维指标体系**



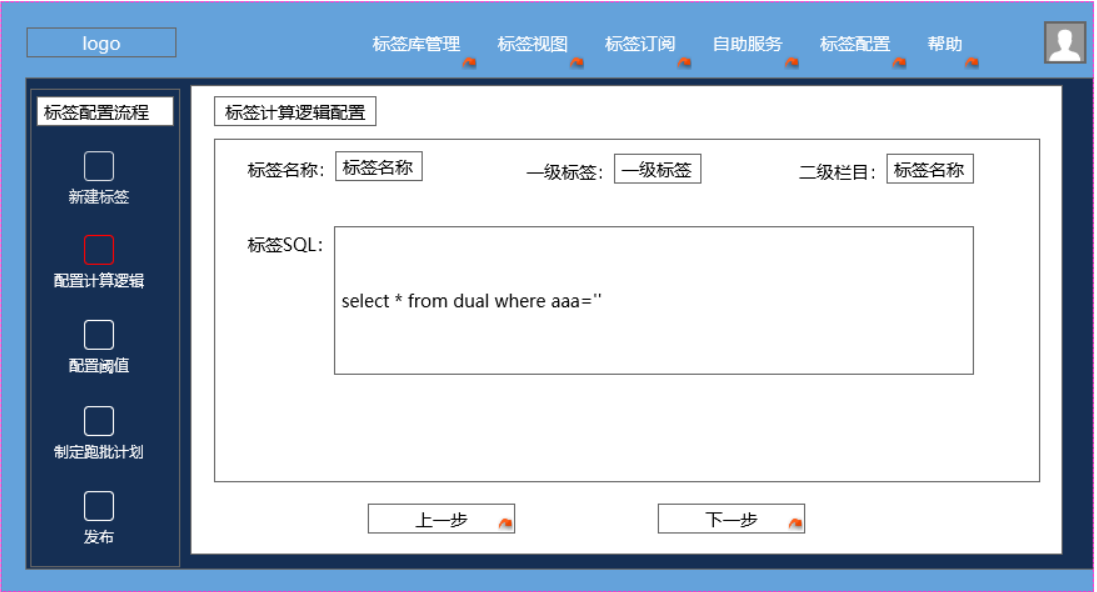
* + **运维指标管理**
  1. 生命周期管理：(对应标签配置)

支持运维数据运维指标新增、修改、发布、应用、评估、冻结、下线、归档等。 允许对数据运维指标进行维护，支持对运维指标数据进行发布，根据运维指标数据的应用进行效能评估，支持对数据运维指标进行冻结、下线处理，定期将历史运维指标数据插入到各自的历史运维指标表中，实现运维指标数据的归档。

新增标签：



配置计算逻辑：



配置阈值：



制定跑批计划：



发布：





审核：



审核通过后，试运行：



应用（应用后，可在行内数据之家上展现）：



冻结：



下线：



归档：



* 1. 运维指标要素管理：基于行内数据管理要求，配套建立运维数据运维指标相关管理要素。要素包括但不限于：中文名称、英文名称、主题类目、父级运维指标、运维指标属性（事实类/模型类/预测类）、基础/衍生运维指标标志、适用场景、数据来源、更新频度（日/月/季/年等）、安全分级（低敏感/中敏感/高敏感）、历史数据保存周期等；

标签库管理主页面（超级管理员）：



标签查看



标签修改：



标签库管理首页面（标签管理员）：



* 1. 建立多维度的运维数据运维指标视图，使数据运维指标使用者可以从多个维度对数据运维指标状况进行全景式浏览，维度包括运维指标目录（按框架体系、主题类目、系统类目等，结合运维指标层级结构，帮助用户快速清晰的定义各个类目）、运维指标血缘（根据运维指标的父级运维指标或数据来源，实现运维指标的血缘关系分析，辅助定位数据问题）、访问热度（根据用户查询、调用等情况进行统计，对运维指标访问进行统计排序）、访问推荐（根据用户查询、调用情况，结合大数据算法分析，实现运维指标访问推荐建议）等；

标签目录主页面：



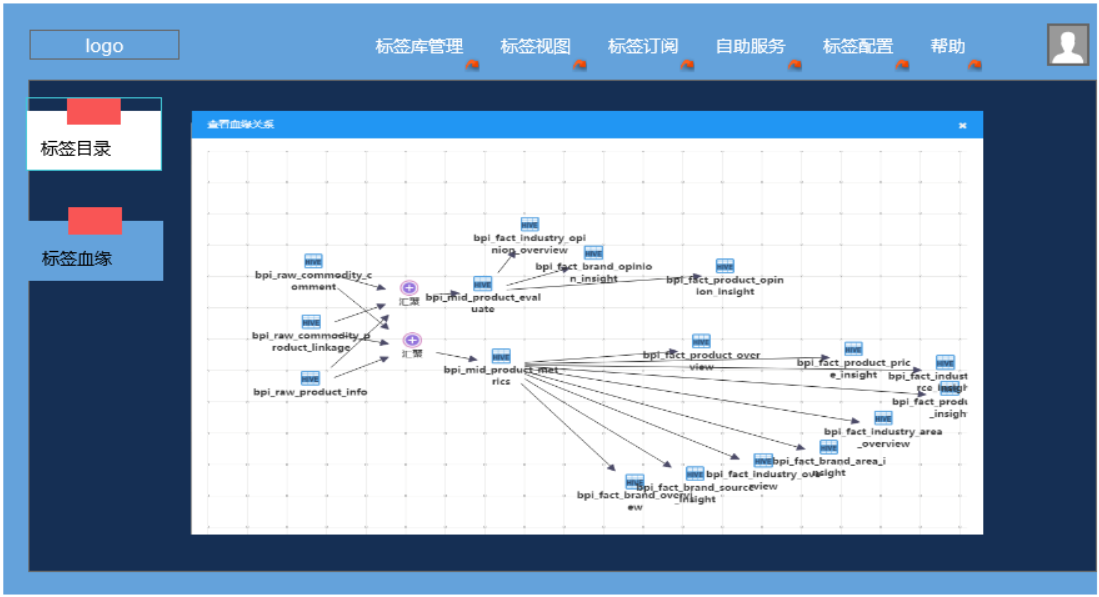
标签目录详细页面：



标签关系钻取页面：



标签血缘：



标签关系详细页面



自助服务：



群体分析:



* + **运维指标提取**
  1. 运维指标作业管理，基于运维信息分析系统的数据模型库数据模型库，实现运维数据运维指标的数据存储、批量加载、批量计算等，建立运维数据运维指标集市；支持运维指标重新打标服务，实现运维指标回流至运维指标集市；支持按数据运维指标元素中的更新频度配置批量提取作业，支持作业的定时或手工发起。



* 1. 运维指标作业监控，监控每日计划运行作业数的情况，按照不同数据源进行分组监控，对失败的作业提供明确信息标识以及错误信息提示，统计跑批成功/失败的记录，支持邮件、Call浦等渠道告警通知。



* + **运维指标服务**

标签API管理：



标签订阅：



**3.5 系统画像**

**3.5.1功能说明**

* 1. 基于运维信息分析系统的大数据模型库和运维数据运维指标体系设计，完成系统画像的数据获取接口等适应性调整。

需求目标：完成数据获取接口适应性调整。其中标签结果数据,接口考虑Hbase或者kylin ;标签钻取数据，接口考虑Hive或者ES。

* 1. 基于运维信息分析系统的大数据模型库API接口，支持自定义图表展示，实现系统画像运维指标的钻取。

需求目标：实现类似于夏洛克平台中 仪表盘的图形展示。

* 1. 基于已有数据运维指标成果，引入大数据智能分析，实现评分指标智能选择、评分权重智能设置。

需求目标：暂定。

* 1. 基于数据运维指标实现系统间的关联图谱。

需求目标：基于系统->标签的访问形式，实现标签->系统的关联访问。

* 1. 实现系统画像配套优化。

需求目标：实现对当前系统画像存在的一些问题进行优化，包括页面和功能等。

* 1. 实现运维指标在移动端（CALL浦端）的展示和使用。

需求目标：将目前系统画像上的标签信息在CALL端展示。

1. **非功能性需求**

## 4.1集群性能指标

从数据搜索能力、海量数据吞吐能力及场景应用三个方面对现有集群进行测试，并根据二期数据使用规划提出集群性能的新需求。

搜索性能需求：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 一期测试 | | 二期需求 | |
| 数据量(TB) | 耗时（s） | 数据量(TB) | 耗时（s） |
| 固化关键字搜索 | 1T | 2.33 | <=2 | 2 |
| >2 | 5 |
| 动态聚合搜索 | 1T | <8 | <=2 | 5 |
| >2 | <15 |

海量数据吞吐能力需求：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 一期测试 | | 二期需求 | |
| 均值 | 峰值 | 均值 | 峰值 |
| 日增数据 | 3TB | 无 | 11.71TB | 无 |
| 单节点采集速率 | 3万/s | 15万/s | 3万/s | 15万/s |
| 整体处理能力 | 3万/s | 60万/s | 10万/s | >60万/s |
| 入库延迟 | 0.3s | <1分钟 | <1s | <1分钟 |

场景应用需求：

|  |  |  |
| --- | --- | --- |
|  | 一期测试 | 二期需求 |
| 应用场景app | 26 | 50 |
| 可视化报表 | 214 | 500 |
| 配置数据流 | 71 | 150 |

**4.2业务量指标**

|  |  |  |  |
| --- | --- | --- | --- |
|  | 年内初始值 | 年增量（%） | 3年目标值 |
| 机构数 | 无 | 无 | 无 |
| 用户数 | 100 | 20% | 140 |
| 签约客户数 | 无 | 无 | 无 |
| 产品协议签约数（协议包括账户、合同, 具体名称根据需求而定， 如活期账户、国债账户、贷款合约等） | 无 | 无 | 无 |
| 并发访问（接入）量 | 460项任务包括：  1、接入任务95项  2、处理任务115项  3、存储任务230项  4、查询任务20项 | 20% | 660项任务 |
| 针对BS架构系统，补充在线用户数需求 | 20 | 50% | 30 |
| 日均联机交易量 | 无 | 无 | 无 |
| 高峰日联机交易量 | 无 | 无 | 无 |
| 批量业务日处理量（如批量代发代扣、清算） | 无 | 无 | 无 |
| 批量业务高峰日处理量 | 无 | 无 | 无 |

**4.3数据存储要求**

|  |  |
| --- | --- |
| 历史数据导入年限： | 2017年9月至本项目投产前 |
| 历史数据保存年限（如涉及交易类数据、账户类数据、日志数据、报表数据等分别说明保存年限）： | 永久 |
| 在线数据查询年限： | 秒级KPI数据1个月  其他3个月 |
| 近线数据查询年限： | 永久 |

针对二期的数据需求和一期实际占用存储的数据总量（不计算副本），日增数据匡算如下：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 分类 | 数据源 | 当前每日增量统计 | 二期每日增量预估 | 估算依据 |
| 性能容量数据 | Oracle、DB2、SQLServer数据库 | 10G | 72G | 目前供数频率为30分钟，二期调整为5分钟以内；业务增长估算20%，故总体增长约为7.2倍 |
| 中间件 | 5G | 6G | 业务增长估算20%，故总体增长约为1.2倍 |
| 服务器-监控系统采数 | 63G | 378G | 目前供数频率为5分钟，二期调整为1分钟，部分乃至秒级；业务增长估算20%，故总体增长至少5.2倍 |
| 服务器-AIX脚本秒级采数 | 2T | 3.5T | 合肥数据接入，按照与生产80%规模估算，业务增长估算20%，故总体增长约为1.2倍 |
| Docker性能容量 | 10G | 50G | Docker业务增长较传统为指数倍，增长估算为当前5倍 |
| 存储 | 10G | 12G | 业务增长估算20%，故总体增长约为1.2倍 |
| 网络 | 10G | 24G | 根据二期网络组提出需求需要接入网络线路信息、网络配置等信息，日增数据可以达到20G,再考虑业务20%的增长 |
| Informix、Mysql、Sybase数据库 | 无 | 500G | 试点采集Informix数据库秒级性能和sysbase推算 |
| 试点三家分行日志 | 无 | 300G | 根据一家分行采集数据100G预估 |
| 华为DDOS、华为存储、BPC、Zabbix、VC等十种产品数据 | 无 | 500G | 相关数据提供方预估 |
| 日志数据 | 告警 | 1G | 5G | 目前接入criticl级别告警，后续其他级别告警也接入，并考虑业务增长估算，预计5G |
| 应用性能（splunk） | 300G（200G秒级聚合后的KPI，1.3的膨胀,还有一些维度聚合表） | 450G | 业务增长估算50%，故总体增长约为1.2倍 |
| 应用性能  （应用监控） | 140G | 140G | 不会增长 |
| 应用日志（splunk替代方案试点） | 20G（单个系统，全量日志明细，非秒级聚合的KPI） | 240G | 一期试点手机银行单个系统日接数据量为20G，二期试点范围至10个重要系统，并考虑20%的业务增长，故总体约为20\*10\*1.2=240G |
| SSL设备日志 | 500G | 600G | 业务增长估算20%，故总体增长约为1.2倍 |
| 系统日志 | / | / | 容量估算包含在《信息科技部系统日志集中管理平台迁移项目》中 |
| 指控系统数据 | 无 | 1T | 原始报文约3-4T，解析完毕有效文本内容约占1T |
| 运维管理数据 | 人员合规类数据  （变更、事件、操作命令） | 20G | 24G | 业务增长估算20%，故总体增长约为1.2倍 |
| 总计 |  |  | 7.76T |  |

**4.4服务等级指标**

|  |  |  |
| --- | --- | --- |
| 服务等级指标 |  | 备注 |
| 是否为监管类应用 | 否 |  |
| 面向客户服务应用功能 | 否 |  |
| 外联机构 | 无 |  |
| 应用服务时段要求 | 7\*24 |  |
| 交易量占比分析 | 无 |  |
| 联机高峰交易量发生时段 | 9：00-17：00 |  |
| 联机高峰交易量发生日期 | 无 |  |
| 单笔实时交易响应时间要求 | <3s |  |
| 高峰批量业务时段 | 0时至8时 |  |
| 高峰批量业务日期 | 无 |  |
| 应用等级重要程度 | 中断时间不大于2小时 | 系统发生灾难性问题时系统恢复能力的要求（如允许中断的时间要求） |
| 系统运行级别（7\*24，5\*8等 | 7\*24 |  |
| 页面要求响应时间 | 在线数据查询<3s  近线数据查询<15s |  |
| 报表生成响应时间 | 在线数据查询<3s  近线数据查询<15s |  |

1. **风险预估及应急方案**

无

1. **需求解释及验收负责人**

信息科技部系统运维中心