

# 信息化运维系统的图模型存储<sup>①</sup>

周振煜, 陈望旭, 韩笑, 林加国

(南京南瑞信息通信科技有限公司, 南京 210003)

通讯作者: 周振煜, E-mail: [zhouzhenyu@sgepri.sgcc.com.cn](mailto:zhouzhenyu@sgepri.sgcc.com.cn)



**摘要:** 基于新型 IT 基础设施, 如云计算环境, 大数据平台等, 信息化应用获得了灵活可靠的底层服务. 同时通过新型服务形式, 如物联网应用, 移动端服务等, 增强了业务服务能力. 但这些新技术使得信息运维面临运维对象频繁变化, 对象关系不断调整, 运维数据格式多样的难题. 本文基于电网企业当前运维现状, 给出一种基于图模型的运维数据存储设计, 提高了对动态, 非结构化运维数据的处理能力. 经实验验证, 模型处理普适性更广, 数据读写性能稳定, 方案有效解决了新运维形式下的电网企业信息化运维系统的业务难点.

**关键词:** 图模型; 数据存储; 信息化运维系统

引用格式: 周振煜, 陈望旭, 韩笑, 林加国. 信息化运维系统的图模型存储. 计算机系统应用, 2021, 30(5): 128-133. <http://www.c-s-a.org.cn/1003-3254/7951.html>

## Graph-Model Storage of Information Operation and Maintenance System

ZHOU Zhen-Yu, CHEN Wang-Xu, HAN Xiao, LIN Jia-Guo

(NARI Group Corporation Information & Communication Technology Co. Ltd., Nanjing 210003, China)

**Abstract:** On the basis of new IT infrastructure, such as cloud computing environment and big data platform, information application has obtained flexible and reliable underlying services. At the same time, through new service forms, such as Internet of Things application and mobile service, the business service ability is enhanced. However, these new techniques make information operation and maintenance face a lot of problems, such as frequent changes of operation and maintenance objects, constant adjustment of object relationships, diversity of operation and maintenance data formats. According to the current operation and maintenance status of power grid enterprises, we present a design of operation and maintenance data storage based on a graph model, which improves the ability of processing dynamic and unstructured operation and maintenance data. The experimental results show that the model has wider universality and stable data reading and writing performance. Therefore, the proposed scheme in this study effectively solves the business difficulties in the information operation and maintenance system of power grid enterprises in the new form of operation and maintenance.

**Key words:** graph model; data storage; information operation and maintenance system

## 1 引言

随着电网企业生产经营服务对信息和信息系统的需求不断增加, 电网企业 IT 服务和环境的复杂性也在不断增加, 特别在近些年电网企业内部依据不同业务

或服务需要, 构建起的云计算环境、大数据平台、物联网应用、移动端服务等各类复杂的 IT 基础设施和业务服务系统给电网企业信息运维带来巨大的挑战<sup>[1]</sup>.

早期的电网企业信息化运维活动中, 需要处理的

① 基金项目: 南京南瑞信息通信科技有限公司科技项目 (5246DR200016)

Foundation item: Technical Project of NARI Group Corporation Information & Communication Technology Co. Ltd. (5246DR200016)

收稿时间: 2020-09-25; 修改时间: 2020-10-21; 采用时间: 2020-11-17; csa 在线出版时间: 2021-04-28

是有限且较为稳定的运维对象和业务流程,如主机、网络设备、中间件、数据库、运维人员、业务审批流程、设备检修流程等,但在当前信息化运维活动中电网企业面临的是不断变化的运维对象和不断调整的业务流程,如虚拟机、容器服务、POD节点、临检流程、设备申请流程等。信息化运维系统作为运维工作的重要平台,在日益复杂的运维活动中,被要求除了要监管上述或稳定或不断变化的运维对象外,还需要对不断增加的系统日志数据、系统服务间调用情况等诸多海量非结构化、相互关联的数据实现采集和监管。

面对复杂的运维对象和数据,电网企业当前基于关系型数据库的信息化运维系统越来越难以满足监控要求。容器服务、POD节点等运维对象,在云环境中不断生成和删除,系统日志数据、服务调用关系链,在系统运行中不断增加和调整,关系型数据库在应对大量动态、非结构化数据时显得捉襟见肘<sup>[2]</sup>。

综合上述基本情况,本文基于电网企业当前信息化环境的实际运维对象特点及运维要求提出一种基于图模型的运维数据存储设计,以此优化电网企业信息化运维系统对动态、非结构化运维数据的处理能力。

## 2 运维现状

电网企业面对信息化环境的变化针对诸多运维需求和角度进行了探索实践,实践中运维系统的各类定制化拓展丰富了运维对象,延伸了运维内容,优化了运维流程。归纳来讲,运维系统的拓展落脚在运维对象上,即分为对传统运维对象的管理优化,对云环境等新运维对象的监控探索,以及结合两类运维对象的综合性运维管理实现。

### 2.1 数据分类

传统运维对象主要涉及设备硬件、软件、虚拟资源、基础支撑资源。

其中设备硬件包括如PC服务器、刀片机、小型机等主机设备,磁盘阵列、磁带库、光盘库等存储设备,路由器、交换机、负载均衡器等网络设备,防火墙、防病毒网关、入侵监测等安全设备,手持终端、平板电脑、云终端等终端设备,UPS、KVM、打印机等外部辅助设备,内存条、电源模块、网卡等备品备件。

软件包括操作系统、数据库、中间件以及ERP系统、财务管理系统、人力资源系统等各类业务系统。

虚拟资源为由VMware等传统虚拟化工具生成的

虚拟服务器、资源池、虚拟桌面等对象。

基础支撑包括建筑场地、IP、账号权限等其他不在软、硬件和虚拟化对象中的内容。

传统运维对象在运维过程中涉及了资产、采购、服务、维护、运行、流程管理、关联关系等内容。

资产内容包含产权信息、ERP信息等,采购内容涉及采购时间、供应商等信息,服务内容涉及售后服务时间、服务商、序列号等信息,维护内容涉及领用信息及检修信息,运行内容涉及内存大小、CPU数量等,流程管理内容涉及流程状态、审批内容等,关联关系涉及关联、依赖、连接、备份、包含等。

云环境等新运维对象主要涉及基础云平台、商业云平台、虚拟资源以及其他在运维实践中产生的对象内容等。

其中基础云平台包括OpenStack、Kubernetes,商业云平台包括阿里云、华为云、京东物联平台等,虚拟资源涉及容器、POD、容器控制器、虚拟交换机、虚拟化存储、虚拟路由器等由云平台产生的虚拟对象,其他新的运维对象还涉及服务调用链路、拓扑视图等。

### 2.2 存储方案

在电网企业现有的运维实践中,针对传统运维对象的存储依旧主要采用关系型数据库,而对云环境等新运维对象则尝试采用了文档数据库、列式数据库、图数据库等多种数据库。

传统运维对象的生命周期长,对象属性变化频率低,运维对象间关系稳定。以关系型数据库存储相应的对象时,对象模型以表结构的形式存在,模型的属性即为表的列,模型及其属性的变更映射为表结构的变更,模型间的关系通过表的关联关系表达,既有强约束的外键,也存在弱约束的引用字段。

云环境等新运维对象生命周期短,运维对象多数随着环境访问、运行状态自动触发生成或删除;对象属性及运维对象间的关系随着运行不断变化。以非关系型数据库存储相应的对象时,主要依据数据特点选用不同的数据库,如文档数据库多用于存储容器、POD等频繁生成删除的对象,利用文档数据库高读写性能,可以有效应对短生命周期对象的采集存储及监控需要;列式数据库的存储中将模型结构存于代码逻辑之中,数据库只负责存储,将所有模型扁平化,一条数据即为一个实例,模型属性变更便捷;图数据库实现模型及资源间的关系存储,高性能的图形数据读写能力对业务

关系的管理提供有力支撑<sup>[3]</sup>。

### 2.3 运维要求

由前述分析可见电网企业当前信息化环境的实际运维对象具备如下特点:

- (1) 传统运维对象和云环境等新运维对象共存;
- (2) 对象属性同时具备运维管理及资产维护的双重要求;
- (3) 对象间存在多重继承;
- (4) 对象间具有可自定义的关联关系。

尽管针对某些特定运维需要,在具体的运维实践中对特定存储有了较好的设计及应用,但电网企业在整合上述实践过程中也提出如下运维要求:

- (1) 保持对已有运维工作的支持;
- (2) 能快速且自动化反应云环境及其之上运行系统的运行时状态;
- (3) 提供可自定义的各类拓扑视图展示。

## 3 图结构模型设计

基于上述运维对象特点及运维要求,本文对电网企业信息化运维系统涉及的对象存储进行重新梳理设计,提出一种基于图模型的运维数据存储设计。

### 3.1 设计条件

首先,根据对象模型的继承特点及建模规则:

- (1) 模型的继承关系是单向性的,方向为由于子孙模型指向父祖模型;
- (2) 模型在创建时需要确定继承关系;
- (3) 模型只能继承于已经存在的模型。

可以推出,模型无法通过多次继承关系实现对自身的继承,即模型间的继承关系组成有向无环图。

其次,根据对象模型的关联特点:

- (1) 模型的关联关系是单向性的,方向为由某一模型指向另一模型;
- (2) 依赖关系存在的模型必然直接或间接关联一个不依赖关系存在的独立模型;
- (3) 两个依赖关系存在的模型不相互依赖。

还可以推出,模型间关联关系可以形成环路,但环路中的关系必存在可空。

即模型设计应基于下述必要而可行的条件:

- (1) 模型间继承关系组成一张有向无环图;
- (2) 模型间关联关系如果形成环路,则必然有一条关联关系为可空关系。

### 3.2 存储结构设计

可以直观地看出,上述条件下的模型最适合的存储形式为图形数据,即以模型为节点,继承、关联关系为边,模型的增删改查业务逻辑转化为对图数据节点及边的相关操作。

此外根据资源实例同模型的对应关系,亦可轻易地将资源实例纳入图模型中,即以资源实例为节点,将模型同资源实例的关系、资源实例间的关系作为边。

综上设计出基于图模型的运维数据存储结构,见图1。

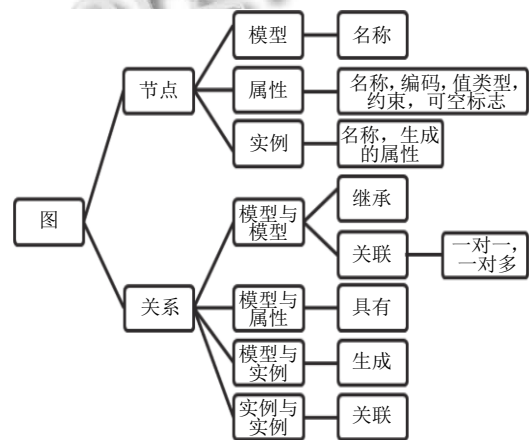


图1 运维数据存储结构

在设计中,模型和实例实现为图结构的节点,模型间关系、模型和实例关系、实例间关系,实现为图结构的边。由于模型属性管理的灵活性需要,模型属性作为节点实现,模型和模型属性间的关系作为边。但并未相应地将资源属性作为节点,而是仍然作为资源实例节点的属性处理。这样的设计在一定程度上破坏了模型和资源实例的存储结构一致性,使得两者在管理操作的实现上需要区别处理,即模型和属性间以关系进行操作,实例和属性间以节点属性进行操作,但是在资源实例查询上简化了操作提升了性能。

## 4 存储管理实现及实践

基于上述图结构信息化运维模型,本文对该模型的存储管理进行了实现,同时结合具体的信息化运维管理平台对该模型及其存储管理功能进行实际验证。

### 4.1 数据库选型

结合当前电网公司在信息化运维领域的探索实践,

选取其中运用较多的关系型数据库、列式数据库、文档数据库、图数据库4类数据库作为备选数据库。

关系型数据库的存储中,如前现状分析,模型的变更将引起数据库结构的变动,且无法较好地处理云资源等新运维对象快速存取的需要。此外由于本文设计的模型需要实现大量复杂关系运算,通过表结构实现的关联关系表达会在多表联查等操作中不可避免地存在性能低下、关系处理代码逻辑复杂等问题,因此不适合作为本文设计模型的底层存储数据库。

列式数据库在存储中可将所有模型扁平化,属性表现为列族的划分,列族中的列来自相应的父类模型。由于数据库只负责存储,数据无结构化,模型与实例的变更非常方便,但更新实例尤其是按条件批量更新难以实现,查询受限于设计的key,难以实现灵活的多维度查询。此外模型结构的处理无法基于数据库本身实现,需要通过代码逻辑进行处理,模型关系需另行存储,关系处理会比较复杂,亦不适合作为本文设计模型的底层存储数据库。

文档数据库通过将模型文档化,能够支持对模型及其实例高性能的变更操作。模式自由的特性也意味着无需预定义模型结构,在系统生命周期中,模型结构的变化也不影响实例数据的存储。通过索引和动态查询能力的支撑,文档数据库可以满足单模型查询的绝大部分需求,但无法支持跨模型级联查询等复杂关系查询。故若以文档数据库为存储数据库,则需要基于该数据库实现额外的关系处理功能。

图数据库存储和处理的数据对象是节点及其之间的关系,该数据对象同本文设计的模型有较高的契合度。在图数据库存储中,模型是节点,属性也是节点,无需依赖代码逻辑或者其他外部因素,模型及属性的变更简单且不产生数据结构的变化。此外模型间关系、模型与属性的关联、实例与模型关联、实例间关系,均转化为节点之间的关系,通过统一的图查询方法实现不同的查询需要,能满足高效的复杂多级关系查询。

综上所述,文档数据库和图数据库在模型存储及管理上均有良好支撑,满足动态的模型调整且对数据结构影响较小,但在关系处理上图数据库相较文档数据库更具优势,故本文采用图数据库作为底层存储数据库,对上述基于图模型的运维数据存储设计进行实现。在开源图数据库中,Neo4j相较DGraph等其他图数据库在社交网络、智能推荐等领域得到了广泛应

用<sup>[4]</sup>,支持文档丰富、便于生产改造。通过Cypher语句,Neo4j提供了对节点、关系及其属性等半结构化数据的创建和维护,高性能的检索、遍历图形化数据的能力支撑了基于图形数据的业务应用实现。

#### 4.2 总体结构

尽管Neo4j可以有效地管理图形化后的运维对象,但通过Cypher语句操作的数据组织形式同结构化数据存在较大差异<sup>[5]</sup>。此外存储结构不代表业务结构,业务结构可以在业务层进行重构。为了减少对上层业务的影响,降低开发、部署的兼容改造工作,故需要在Neo4j存储之上实现模型适配接口。同时为提供个性化的业务需求支持,在模型适配接口之上提供内容转换和查询引擎模块。据此基于图模型的运维数据存储功能总体结构如图2。

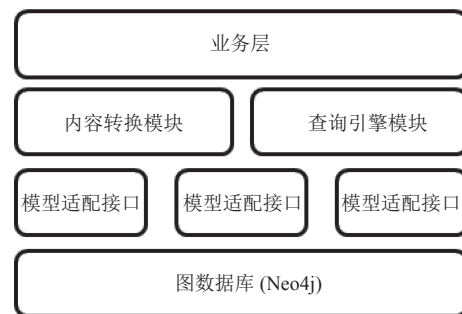


图2 存储功能总体结构

模型适配接口模块实现了资源模型、模型属性、模型继承和关联关系、资源实例、资源实例间关系、实例所属模型等数据的增加、删除、修改、查询的功能,以REST接口对外提供服务。服务基于Cypher语句对Neo4j数据库进行操作,对应了资源模型、模型属性、资源实例3类节点和继承、关联、具有属性、生成实例4类边的增加、删除、修改、查询操作。

内容转换模块实现了Neo4j数据库中存储结构同业务结构间的转换功能,即将从Neo4j数据库中查询到的,如模型继承关系、属性分组标签、机房拓扑结构、地域组织关系等信息,转换为模型树、属性分组、机房管理、地域管理等业务可视结构,部分高频转换以数据词典形式固化,提升了效率。

查询引擎模块实现定制查询及模糊查询等优化查询功能。通过对常用的复杂查询需要进行定制,提升查询效率的同时减少了业务端的开发工作。同时基于

ElasticSearch 搜索引擎,实现模型、资源节点的快速定位,实现全文模糊查询能力.

### 4.3 集群服务

由于 Neo4j 社区版不提供集群服务功能<sup>[6]</sup>,为保证 Neo4j 在生产环境中的可靠性和稳定性,需要实现 Neo4j 的集群能力.集群的方案首推多读多写形式,由于 Neo4j 社区版是适用于单实例部署的版本,无法较好地处理数据分片,考虑高性能分布式图数据处理也是业界难点,亦非本文研究重点内容<sup>[7]</sup>,因此使用 Neo4j 社区版构建数据库集群服务采用了一写多读形式,其功能结构如图 3.

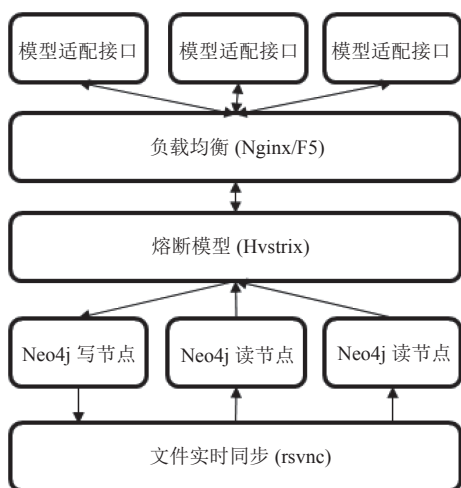


图 3 数据库集群服务

通过多组 Neo4j 社区版单实例组成集群节点,选取某一节点作为写入节点,其他节点作为读取节点.为保障读取节点同写入节点的数据一致性,节点之间采用文件实时同步工具 rsync 实现数据同步.模型适配接口模块对 Neo4j 集群的访问无需关注具体的 Neo4j 单实例节点,各读写节点前端通过负载 (Nginx 或 F5) 和熔断 (SpringCloud-Hystrix) 模块实现访问控制,在保障读写节点分离的同时,实现节点之间的主备切换及读写节点的功能转换,提高了集群服务整体可靠性.

### 4.4 实践验证

为验证本文方案,在电网企业某省公司信息化运维管理平台测试环境中进行部署和评估.

验证环境搭建在 6 台 32 核 64 GB 内存的虚拟机上,其中 3 台搭建 Kubernetes 容器云平台用于部署存储管理,另外 3 台搭建 Neo4j 数据库集群服务.

存储管理的模型适配接口、内容转换模块、查询引擎模块基于 Spring Cloud 实现,均采用容器化部署,内容转换模块、查询引擎模块的容器副本数均设置为 2,模型适配接口的容器副本数设置为 4.

上层业务应用选取该省公司信息化运维管理平台的资源配置、采集监控和资源管理 3 个模块,并根据接口需要,在存储管理业务层定制开发相应接口.测试环境中修改信息化运维管理平台对应模块的接口配置,指向存储管理相关接口.同时根据业务需要,对原运维管理平台中硬件资源、软件实例、云平台及虚拟化、网络资源等四大类 65 个模型及其模型属性等信息进行初始化,同时导入对应模型的资源数据.经过运维人员日常使用,本文方案对运维管理平台功能支撑完整,模型管理灵活度高,可根据实际运维需要由运维人员进行模型定制处理.

方案性能测试评估从前端应用页面调用接口及后端采集数据批量入库调用接口 2 方面进行验证.接口调用情况通过信息化运维管理平台接口自监控模块进行记录统计,前端调用情况如图 4,后端调用情况如图 5.

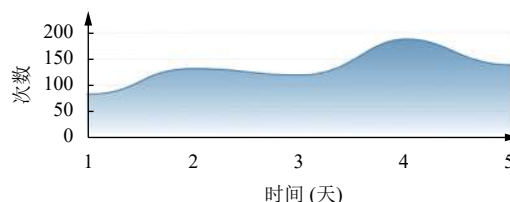


图 4 页面接口调用统计

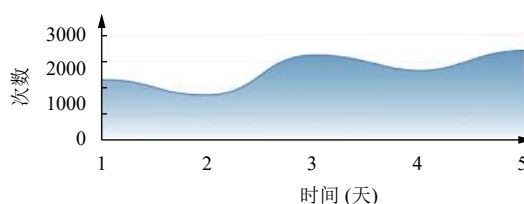


图 5 外部系统调用接口统计

通过反复压力测试,运行中的存储管理功能稳定,Neo4j 数据库集群服务稳定,数据库集群服务单节点写入及数据读取性能满足当前数据处理需要.新运维模型在数据处理及展示中相较关系模型更为直观,在资源监控模块的监控拓扑图的实时生成和变更操作中更具便利性,模型设计及应用符合预期目标.

## 5 结论与展望

本文分析了电网企业信息化运维的数据现状和运维要求,在电网企业大量运维探索实践的基础上,设计了基于图模型的运维数据存储方案,给出了模型继承特点和模型关联特点下的图形管理依据以及模型、属性和模型实例等全部运维对象的基础数据存储结构等创新点,阐述了该方案的存储管理设计及 Neo4j 数据库集群设计。最后,以电网企业某省公司信息化运维管理平台为现实应用场景,给出了基于图模型的运维数据存储方案的应用验证评估,生产试运行实践表明,该存储方案能够纳管所有运维对象、模型管理自由程度高、新运维对象支撑效果好,关系查询便捷,能够为电网企业的信息化运维管理平台提供良好的基础数据管理支撑。后续将针对自动化运维、统计分析、告警管理等运维数据进一步拓展并优化该解决方案。

## 参考文献

- 1 高赞. 基于 ITIL 的运维管理系统设计与实现 [硕士学位论文]. 成都: 电子科技大学, 2020.
- 2 戴声, 王波. 基于图数据库的大规模配置管理数据库关联查询. 计算机与现代化, 2017, (1): 51-56.
- 3 陈喜春, 牛晓莉. 基于图数据库的装备信息联想式查询. 计算机系统应用, 2019, 28(5): 244-247. [doi: 10.15888/j.cnki.csa.006895]
- 4 郭蕴颖. 基于知识图谱的电网信息搜索引擎的设计与实现 [硕士学位论文]. 沈阳: 中国科学院大学 (中国科学院沈阳计算技术研究所), 2020.
- 5 The Neo4j cypher manual v4.2. <https://neo4j.com/docs/cypher-manual/current/>. [2020-04-16].
- 6 The Neo4j operations manual v4.2. <https://neo4j.com/docs/operations-manual/current/introduction/>. [2020-04-16].
- 7 李宗沅. 分布式图数据库关键技术的设计与实现 [硕士学位论文]. 成都: 电子科技大学, 2020.