



# 基于大数据技术的网管系统数据管理层

李 玮

(中国移动通信集团设计院有限公司 北京 100080)

**摘 要:**数据管理层是运营商网管系统中数据处理和管理的核心。从网管现状和问题分析入手,提出打破专业网管之间的隔离、推进数据的公有共享,进一步细分数据管理层的数据域和重构数据模型,并使用 Hadoop+MPP 混搭的技术架构建设数据管理层平台,为网管系统的规划和设计指明了方向。最后,还对数据管理层平台和企业级大数据共享服务平台的关系进行了论述。

**关键词:**网管系统;数据管理层;大数据;数据模型;技术架构

**doi:** 10.11959/j.issn.1000-0801.2015265

## Network Management System's Data Management Layer Based on Big Data Technology

Li Wei

(China Mobile Group Design Institute Co., Ltd., Beijing 100080, China)

**Abstract:** Data management layer is the core of data processing and management in network management system. The current situation and problems of network management system was introduced. To break the isolation between the professional network management system and promote data sharing, data domain should further be subdivided and the data model should be reconstructed, and technical framework of Hadoop+MPP was a better choice. The direction for the planning and design of the network management system was pointed out. Finally, the relation of data management layer platform and enterprise wide data sharing service platform was discussed.

**Key words:** network management system, data management layer, big data, data model, technical framework

### 1 引言

网管支撑系统(network management support system,简称网管系统或者网管)是网络和业务运维的信息化管理平台。如果以 1990 年 IETF 在 RFC1157 中正式公布的 SNMP (simple network management protocol, 简单网络管理协议)和 ISO 在 ISO/IEC 7498-4 文档中定义出的五大功能作为网管系统发展的起点,那么网管已经从简单的维护管理工

具,发展成为运营商提升服务质量和支撑业务可持续性发展的关键生产运维系统。在这 20 多年中,新技术的不断涌现,尤其是近年兴起的大数据技术,为运营商网管系统进一步充分利用、分析、挖掘其掌握的规模巨大、种类多样、覆盖全面、真实权威的数据带来崭新契机。

### 2 现状分析

目前,运营商的网管主要还是面向网元、网络以及业

收稿日期:2015-05-14;修回日期:2015-10-10

论文引用格式:李玮. 基于大数据技术的网管系统数据管理层. 电信科学, 2015265

Li W. Network management system's data management layer based on big data technology. Telecommunications Science, 2015265



务的分专业网管,如话务网管、数据网管、传输网管等。这些专业网管虽然在专业网络管理、所属业务简单开通和质量保持上发挥着重要作用,但已经不能很好地满足运营商内部能力提升的要求。

- 不同专业网管之间信息需要长流程传递,无法做到实时交互,导致跨专业的业务开通能力受限,也不利于跨专业故障的快速准确定位。
- 专业网管系统各自孤立,信息分散存储,难以支撑全局意义上的信息共享和关联挖掘分析,无法很好地满足管理层在资源整体规划设计、预测分析上的要求,也无法很好地支撑客户体验度 (quality of experience, QoE) 提高的需求。
- 数据共享粒度弱,数据重复采集和重复处理现象普遍存在;与集中模式相比,建设和运维成本居高不下;数据和应用耦合度紧密,开放性较差。

来自于市场环境、方针政策等外部要求,对网管系统也提出了新的要求。

首先,基于互联网及移动互联网的应用和业务高速发展,尤其是 4G 商用规模的不断扩大,促使网管由传统的面向网络和业务管理的服务模式向面向经营管理和客户感知的服务模式转变。在这一模式要求下,不仅数据的来源、类型、内容和组织结构会持续变化更新,而且数据体量会急剧增加,这些都是现有网管的数据模型和技术架构不能应对的。

其次,虚拟运营商、三网融合等政策导向促使“一家独大、一家专营”向社会资源高效配置和跨产业链合作方式转变。由此,网管不仅可能、而且必须成为连接运营商及其合作伙伴的关键一环。在精确划分客户、发掘利润点、消除业务增长的制约和牵制因素等方面发挥出更大作用。在从后台支撑系统向前台服务系统转变的过程中,当前按照专业划分的网管界限必须被打破,建设一个能够集中处理、支持合作开放、与应用耦合度更松散的数据处理平台正在成为业内共识。

从根本上解决上述问题,关键点是要将网管系统中原有的自有、独享、离散、偏重于当前交易应用(transaction application)和业务管理的专项数据,扩充为公有、共享、聚合、兼顾长期分析应用(analysis application)和客户感知的集合数据。从体系结构的角度分析,就是要重构重建网管系统的数据管理层。顾名思义,数据管理层负责各专业资源、性能和告警数据的全面管理和动态建模,并通过数

据链接将不同类型的数据链接在一起,最终实现全数据的统一管理和分析功能。在网管技术体系结构上,数据管理层位于采集层之上,收集来自采集层的标准化数据;位于应用层之下,为应用层提供统一的接口协议和信息模型服务。虽然不同网管系统略有差别,但一般为减少各类数据与采集以及应用功能的互相过渡依赖,都会引入这一层次。

当然,并不是说数据管理层改造和革新就会解决所有的问题,而是说作为一个承上启下的数据管理和处理组件,如果能够将原有分散在不同专业网管的数据加以汇聚和开放共享,那么就为实现端到端的复杂业务管理、跨专业数据合作、上层经营管理、用户行为感知提供了可行的数据基础和技术环境。

### 3 数据模型的重构

数据模型(data model)是数据特征的抽象,描述了被管理数据的属性信息、关系信息等,是构建网管系统数据管理层的关键,其优劣决定了网管系统数据存储、处理和应用的一致性、扩展性、易用性和快捷性。

传统专业网管以自有资源为主来构建数据模型;但各自的系统没有办法建立跨专业的数据模型,导致无法快速准确实现业务的端到端管理,更无法做到客户感知层面的分析。因此需要在数据管理层重新进行数据组织和模型构建。重构后的数据模型,不仅应该可以有效整合各专业的数据、实现数据的标准化和开放共享、全面管控数据质量,还应该具有缓慢变化维技术对历史分析实体归属变化追溯本源的性能,能够在诸多不确定的数据中发现确定信息,使得数据管理层能够对所有数据可视、可管和可用。

从当前来看,在新一代网管中,数据模型不再简单地仅为本专业上层应用提供数据或者存储数据,还需要提供强大的分析能力。之所以在新的数据管理层中强调 OLAP (on line analytical processing,联机分析处理)的作用,是因为这一技术可以较好地解决用户感知、决策支持等传统专业网管无法应对的问题。

由于数据建模的层次、维度、对象和关系等要素在规模和复杂度上大大增加,因此数据管理层中的数据集合有必要进一步按照来源、类型、目标、应用等进行细分,本文暂将其划分为基础数据域和数据汇总域,并针对每一个数据域建设各自的数据模型,二者的比较见表 1。

表 1 基础数据域和汇总数据域比较

| 序号 | 项目   | 基础数据域                                                        | 汇总数据域                                                                   |
|----|------|--------------------------------------------------------------|-------------------------------------------------------------------------|
| 1  | 数据来源 | 经处理后的采集层数据,如信令数据                                             | 经 ETL(抽取(extract)、转换(transform)、加载(load))处理后的采集层数据和基础层数据,如资源数据          |
| 2  | 数据类型 | 结构化数据和非结构化数据                                                 | 一般是结构化数据                                                                |
| 3  | 模型要求 | 存取方便且速度快,结构清晰,容易理解和扩展                                        | 满足布尔逻辑运算和数学运算规则形成的查询要求;能搜索、组合和比较分析各类数据;增加和删除数据方便;具有更高的数据独立性、更好的安全保密性    |
| 4  | 主要目标 | 形成稳定和统一的基础性明细数据集                                             | 为数据挖掘或多维分析应用提供整合后的、高质量的汇总数据集                                            |
| 5  | 上层调用 | 支持大规模简单写入和读取,无需支持随机修改和写操作;通过计算可以生成各种简单维度的汇总数据和部分直接供上层调用的应用数据 | 支持经过预处理后,一定规模的写入和读取,需支持随机修改和写操作;通过计算可以生成各种复杂维度的统计汇总,计算数据和部分直接供上层调用的应用数据 |
| 6  | 数据体量 | 巨大,一般大于 PB                                                   | 大,一般在数据基础层的几分之一                                                         |
| 7  | 保存时间 | 30~60 天                                                      | 按需存储                                                                    |

4 技术架构的混搭

网管正在逐步从设备和业务的管理迈向面向用户感知、经营管理的新阶段,因此必须对更加广泛的、粒度更细的数据加以采集、处理、分析和挖掘。数据管理层中不仅有传统网管所需资源、配置、告警、用户信息等,还需要引入信息量更丰富、涵盖环节更全面的信令数据,这就导致了新一代网管系统数据管理层处理的数据量远远大于传统网管。这里以 4G 客户对覆盖与上网速率的主观感知分析的应用需求为例。实现这一应用功能,需要采集大量 LTE S1-MME、S6a、S1-U、Uu、X2 等接口信令,还要辅以 Gn、Mc 等信令数据。根据推算每万用户忙时生成的信令(XDR)量大概在 200~400 GB/天,那么按照 1 000 万户 4G 用户考虑,每天需要处理的原始数据量就会达到 195~390 TB,而基础数据域保存 30 天的数据量将达到 5.7~11.4 PB、数据仓库层存储周期内的数据也将达到 1~2 PB。这一数据规模已经超出了运营商一般省份所有专业网管的存储数据之和。这是“小型机+关系型数据库+集中存储”的传统技术架构根本无法实现的。进而言之,只有在数据管理层引入大数据技术,才可能在有效时限内完成如此大规模的数据的处理。

对于大数据技术,如 Hadoop、MPP 的起源、定义、原理、部件组成的研究材料较多,因此本文更多地是从技术架构的角度对大数据技术进行探讨。

首先面临的问题是,在数据管理层的技术架构中,单一的大数据技术是否能够满足要求。下面从起源初衷、擅长领域和工程造价等因素入手进行探讨。

Hadoop 产生的初衷实际上与两个因素有关,一是互

联网数据分析;二是数据的海量性。因此在场景上,Hadoop 非常适用于基于海量非结构化、结构化数据的分布式数据简单快速处理查询;而不太适合做复杂数据模型、复杂关联分析以及按照网元、空间、时间、业务做各维度的数据汇聚和挖掘分析,而这些正是另一个大数据技术——MPP(massively parallel processing)可以大展身手的领域。除此之外,MPP 在 OLAP 方面的优势也非常突出,如有大量成熟完备的第三方工具可以与之集成、完全支持 SQL 等。

而从实现角度来看,MPP 必须预先为访问做优化(分区、索引等),造成其数据装载效率较低;同时在运营商领域,当前 Hadoop 集群节点数已经超过 1 000 台以上;而 MPP 的节点数很少能够扩展到 300 台;此外从采购和建设成本的角度来看,Hadoop 综合成本在 5 000 元/TB;而基于通用 x86 架构的 MPP 成本在 3 万~5 万元/TB,是 Hadoop 成本的数倍。

总之,在网管系统数据管理层的技术架构选择上,Hadoop、MPP 数据库以及传统数据库并非是互相取代的关系,而是取长补短、互为补充的,因此采用混搭架构是当前实现数据管理层方案的可行之选,如图 1 所示。

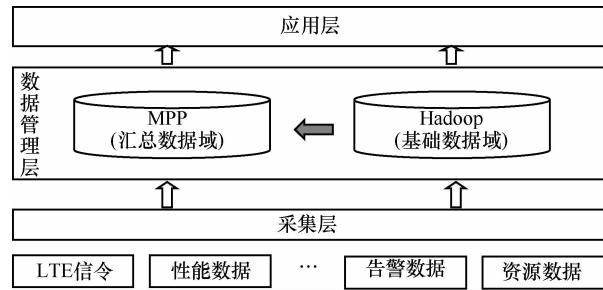


图 1 混搭结构的数据管理层技术架构



## 5 未来演进的设想

网管系统数据管理层在引入大数据技术后,成为管理和加工运营管道中全面、真实、实时、权威、海量数据的超级引擎。不仅可以在 O 域大展身手,而且还可以通过监测、聚合、分析等手段,挖掘更多的信息为 B 域甚至 M 域服务(O 域数据一般占到整体数据规模的 85%以上)。于是这就出现另外一个问题,能否将网管系统的数据管理层建设成为一个为运营商全局服务的企业级大数据共享服务平台,让网管数据逐步成为整个公司而非某个部门的核心资产。大数据共享服务平台既可以推动运营商 IT 一体化和 BOM 三域融合战略落地,又可以利于它开展跨企业、跨领域的合作。这一设想虽然优势明显,但面临的困难也是巨大的。

首先是技术队伍的问题。和互联网公司不同,运营商在网管的自运营与核心自研能力方面还处于一个比较初级的水平,较为严重地依赖于第三方厂商。如果把原有 O 域中的数据管理层拓展为一个面向所有部门、甚至合作伙伴的平台,则各方诉求的快捷准确响应要求对数据管理层的运营者和研发人员都提出极高的要求。其次,运营商部门之间已经形成了较为完整和清晰的分工界面。如果由某个现有部门来筹建和运维大数据共享服务平台,那么必然会造成其他部门在优先级别、服务水平上的担心。另外从技术上看,当前大数据处理中多租户等相关资源分配和权限管理方面的一些问题还在演进和研究的过程中。因此从现状来看,网管中的数据管理层虽然在定位、功能、涵盖范围、技术架构与企业级大数据共享服务平台有高度的吻

合,但是在管理上却有必要成立专职的运营研发部门,在技术上也有诸多困难。

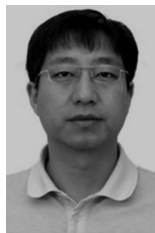
## 6 结束语

网管发展到今天,海量数据的统一建模、统一管理、统一服务等要求正在成为一种必要,而大数据技术的出现使之成为一种可行。基于大数据技术搭建的数据管理层平台,为网管系统数据架构的分布式、标准化、集中化奠定了基础,在数据向高纬度内聚的同时,促进数据与应用的解耦,为网管未来的应用开发模式奠定了坚实基础。

## 参考文献

- 1 刘军. Hadoop 大数据处理. 北京:人民邮电出版社, 2013  
Liu J. Hadoop Big Data Processing. Beijing: Posts and Telecom Press, 2013
- 2 孙博辉. 电信网管系统数据存储及处理方法的研究. <http://www.docin.com/p-318263416.html>, 2014  
Sun B H. Research on data storage and processing method of telecommunication network management system. <http://www.docin.com/p-318263416.html>, 2014

### [作者简介]



李玮,男,中国移动通信集团设计院有限公司网络所高级工程师、高级咨询设计师,主要从事大数据、网管支撑系统、信息安全以及网络方面的规划研究与咨询设计。

2015265-4

## 勘误说明

陈向飞,唐松同志所撰写的稿件《光子集成研究进展》(doi:10.11959/j.issn.1000-0801.2015282)已在我刊 2015 年第 10 期上刊登,因作者失误,2015282-7 页,左栏第 6 节第 3 行:国家自然科学基金委重大项目“高速光电子集成基础研究”(No.61090392),应将项目编号更改为 No.61090390。

特此声明!

《电信科学》编辑部  
2015 年 11 月