



观点

电信运营商大数据管理网元化的思考

赵 钧, 石屹嵘

(中国电信股份有限公司上海研究院 上海 200122)

摘 要: 大数据对电信运营商的未来发展影响深远,甚至可能是业务模式的颠覆式改变。分析了电信运营商大数据的应用现状和存在的问题,对比了数据集中和分散两种大数据服务模式的优劣,提出了基于分散的数据构建运营商大数据网元化的架构和实施建议。

关键词: 大数据;网元;分布式

doi: 10.11959/j.issn.1000-0801.2015167

Reflection on Network Element Evolution of Big Data Management for Telecom Operators

Zhao Jun, Shi Yirong

(Shanghai Research Institute of China Telecom Co., Ltd., Shanghai 200122, China)

Abstract: Big data has had a profound impact on the future of telecommunications operators, and perhaps even subversive type business model change. The present situation of the application of big data and existing problems of telecom operators were analyzed, and the centralized and decentralized big data service models were compared. The big data network element architecture based on scattered data and implementation of the recommendations were proposed.

Key words: big data, network element, distributed

1 引言

随着移动互联网、智能手机的发展,亿万人参与到了数据生产过程之中,每日全球互联网流量达到EB级,而物联网的兴起,预示着未来5~10年,将会有更多超百亿级的传感设备接入互联网,每日产生的数字信息也将呈现出爆发式增长。大数据时代的来临,不仅仅对个人有影响,对电信运营商的影响更加深远,很可能是业务模式的颠覆式改变。

传统电信行业的收费模式是前向收费,随着4G、百兆宽带的到来,电信运营商未来在网络上因投入的成本增长带来的收益增长将逐步降低,用户仅仅为高带宽付费的欲望不高,而更愿意在免费的互联网/移动互联网上为好的应用和内容付费。这就迫使电信运营商要学习互联网企业,增加后向收费模式的收入占比,要实现这一目标,大数据将起到决定性作用。从目前全球数据的增长态势和大数据技术的发展来看,未来将进入数据为王的时代,谁掌握数据、谁会利用数据,谁就能脱颖而出,成为产业链的

收稿日期:2015-04-22;修回日期:2015-07-07

论文引用格式:赵钧,石屹嵘. 电信运营商大数据管理网元化的思考. 电信科学, 2015167

Zhao J, Shi Y R. Reflection on network element evolution of big data management for telecom operators. Telecommunications Science, 2015167

核心,因此电信运营商能否应对大数据的挑战,抓住大数据带来的机遇,很可能会影响到整个产业的未来发展和格局。

2 电信运营商大数据现状和挑战

经过多年的发展,电信运营商已经积累了大量的数据资产,主要体现在:IT支撑系统中的客户资料、订单、账单、计费话单、客户服务、企业管理等数据;网络中的网元网管、管道流量、位置等数据;业务平台中的用户注册、业务订购、业务使用、业务计费等数据。电信运营商已经认识到了大数据的重要性,利用这些大数据资产在网络优化、流量经营、市场营销、客户服务、企业管理、数据对外经营等领域开展了一些应用实践。但不管是和内部网络建设的投入相比,还是和外部互联网行业在大数据上的投入相比,电信运营商在大数据领域都还处于探索阶段,未形成规模效应^[1]。

影响电信运营商大数据应用的一个重要的制约因素是数据分散在各省、各地市、各部门,分散管理、分散运营,数据之间缺乏共享,区域性、不完整的数据对于客户来说价值大打折扣,难以发挥大数据应有的价值。对比互联网企业全网集中的海量数据,从数据覆盖的完整性角度来看自然欠缺优势,这种状况如果没有改观,假以时日,随着其他各行业在大数据领域的持续投入,电信运营商分散的“大数据”的价值将会急速削弱,未来甚至有可能失去与外部合作伙伴共享大数据资源、开展合作经营的机会。

3 大数据集中与分散的博弈

3.1 数据集中策略

要解决电信运营商大数据分散管理、存储和应用中存在的问题,一个很自然的做法是向互联网企业学习,将电信运营商的系统和数据互联网化,全网集中,一套系统、一套数据,从根本上解决数据分散、不一致的问题。这种思路看似很好,但不一定适合电信行业,是否可行还有待验证。从电信运营商的业务特征分析,有以下几点值得思考。

(1) 大数据采集

互联网企业的业务产品大多基于互联网,其大数据主要来自于和客户接触的门户网站,数据采集集中,没有地域性;而电信运营商最重要的移动用户位置数据、固网管道中的DPI数据、用户计费话单等都需要在属地化的本地

网络上采集。

(2) 大数据存储

电信运营商每日产生的数据量巨大,在没有考虑好如何应用数据的前提下,电信运营商将本地采集存储的大数据集中汇聚存储,不仅要高网络带宽保证,还需要新增集中存储设备,将付出巨额数据迁移成本,而且集中存储的扩展性差、安全性差。

(3) 大数据计算

由于是一点采集数据,对于实时性数据分析需求,互联网企业完全可以通过流计算等技术轻松支持,而电信运营商将大量数据上传到集中存储后再计算,数据应用的实时性需求很难满足。

3.2 数据分散策略

回顾电信运营商大数据遇到的问题,可以发现其实数据分散并不是问题,问题是分散的、各自管理的数据之间缺乏共享,难以构成跨区域、跨部门、跨业务的全网大数据。而解决这个问题除了数据物理集中外,逻辑上集中也是可以考虑的,而这种逻辑集中的模式可能也更符合电信运营商的现状和业务特征。

以谷歌为例,其数据中心分布在全球,数据存储和处理采用分布式技术,用户根本感知不到也无需感知数据的具体位置,其分布式架构完全是物理分散的,但并不影响其业务,相反具有更好的横向扩展性,理论上能够支持数据的无限增长。电信运营商网络地域特性决定了数据采集的分散,但如果架构合理,应用分布式技术,一样可以实现数据物理集中的效果,甚至做得更好,如利用数据分布式特征,将实时任务分解到数据采集节点,经过本地化计算,将少量的结果数据上传、组装、处理后提供给应用使用,既利用了现有分散资源的计算能力,又减轻了网络传输和数据集中存储的压力。

从电信运营商目前分散的数据状态过渡到分布式数据存储和计算架构,还需要对现有分散的数据源做分布式改造,并且未来数据应用也需要支持并行分布式处理架构,技术上有一定难度。但从国内外成功案例分析,大数据分布式存储和计算技术已经逐步成熟商用,且这一技术也是未来应对爆发式数据增长的核心技术。

当然电信运营商和谷歌等互联网企业的实际情况还不同,“去电信化”并不是完全互联网化,完全照搬互联网企业,是“以己之短攻彼之长”。互联网企业的分布式技术引入是为了解决数据从集中到分散的问题,而电信运营商



引入分布式技术是为了解决数据从分散到集中的问题,因此电信运营商部署实施分布式架构的策略和方式应有所不同。长期以来电信运营商的核心业务和产品是基于网络的,培养了大批分散在各地的网络运维人员,具备对这些分散的网元进行电信级服务管理能力,电信运营商完全可以利用自身的这一优势,将 IT 和网络融合,对分布式改造的数据节点进行网元化管理和运营。因此,采用这种分布式化的数据分散策略,既符合电信网络运营优势和网络数据本地化采集特征,实现了分散数据的全网逻辑整合,又符合大数据分布式并行采集、存储和计算架构,针对未来电信运营商整合互联网等外部大数据可能导致的数据爆发式增长,具备良好的扩展性。

4 电信运营商大数据管理网元化架构设计

从上文的分析可以看出,电信运营商利用自身网络管理经验,对分散的数据进行网元化管理,实现分布式特征,是互联网行业与电信行业结合的一种大数据架构设计思路。由于是采用数据逻辑集中、借鉴网元管理的设计理念,所以电信运营商的大数据资产梳理、全网数据服务化、服务分层分权、实时数据采集计算、现有数据节点的网元化和分布式改造等都是架构设计的关键点。这里提出电信运营商的大数据网元化架构的参考设计,如图 1 所示。

电信运营商大数据管理网元化架构参考设计逻辑上分为 3 层:数据网元层、数据汇聚网元层、数据服务对外开放节点层。

- 数据网元层:由数据采集、存储、计算类节点构成,

是提供数据服务的基础。建议改造分散在各地各部门的数据节点实现,充分利用现有资源,在利旧基础上根据需要新增。

- 数据汇聚网元层:可根据需要进行配置,数据汇聚网元的功能是对数据网元的数据进行采集、聚合、加工,提供跨数据网元的数据服务。数据汇聚网元层内部还可以进行分层,上级的数据汇聚网元可采集下级数据汇聚网元/数据网元的数据。
- 数据服务对外开放节点层:面向内外部数据服务应用需求,将需求转换为数据服务请求任务,分发给数据汇聚网元/数据网元执行,并将结果聚合反馈给数据应用。

电信运营商大数据管理网元化参考功能架构包括以下模块。

- 数据资产管理模块:对全网数据资产进行统一配置和管理,包括数据项、数据源地址、数据访问权限等内容,形成全网数据全视图。数据资产管理模块建议全网统一部署、分级分权限使用。
- 数据服务管理模块:对全网数据服务进行统一配置和管理,包括服务名称、输入参数、输出参数等内容,形成全网数据服务全视图,并通过和数据资产管理模块的接口,将数据应用对数据的需求转换为服务调用,分发任务到数据网元,同时汇聚服务调用结果给数据应用。数据服务管理模块建议全网统一部署、分级分权限使用。
- 数据网元管理模块:实现类似传统网管的数据网元

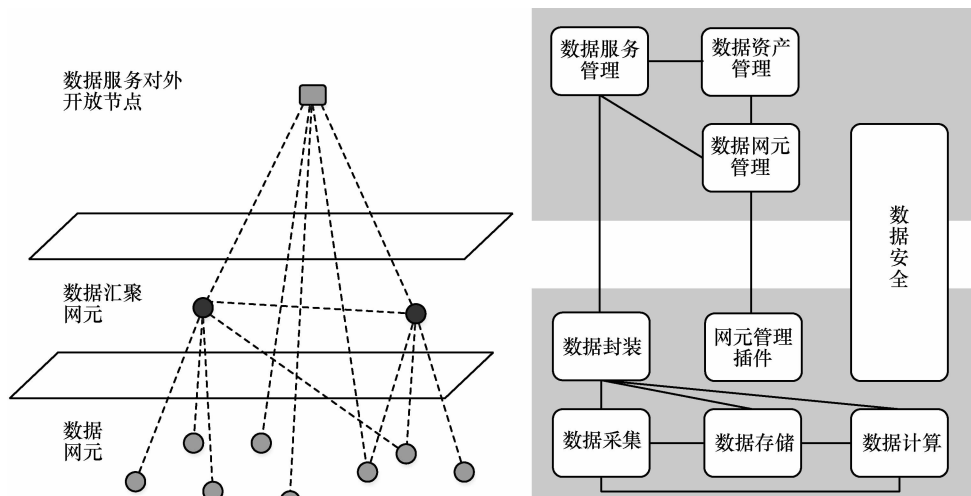


图 1 电信运营商大数据管理网元化架构参考设计

管理功能,通过与数据服务管理模块和数据资产管理模块接口,形成全网数据服务层级拓扑图;可基于拓扑图,与数据汇聚网元/数据网元上部署的网元管理插件接口,对网元进行监控告警,并反映到数据服务管理模块,如禁止无响应数据服务的调用请求。数据网元管理模块建议全网统一部署,分级分权限使用。

- 数据计算模块:实现在原始采集数据基础上进行数据的加工处理功能,分为实时和非实时两类。该模块建议分布式部署,对于有需要提供数据二次加工的数据网元,进行本地化部署或在上级数据汇聚网元进行部署。
- 数据封装模块:在各数据汇聚网元/数据网元上部署,将数据封装成服务。
- 数据存储模块:本地化数据存储,在现有各数据网元上已经具备,可以将采集到的数据存储到本地。对于新建的数据汇聚网元,应具备分布式数据存储能力,如将多个数据网元的数据整合后存储在数据汇聚网元或某几个数据网元上。
- 数据采集模块:现有各数据网元已经具备生成数据或采集数据的能力。对于新建的数据汇聚网元,应具备分布式数据采集能力,能够将数据网元的数据采集到数据汇聚网元上。
- 网元管理插件:部署在数据网元/数据汇聚网元上,支持数据网元管理模块对该网元的监控管理。
- 数据安全模块:实现数据访问控制、数据脱敏、数据加密等安全功能,可根据情况在各数据网元上部署不同级别的安全模块。

电信运营商大数据管理网元化架构参考设计主要引入了服务封装和分布式部署^[2-7]、大数据采集存储和计算^[8-10]、网络管理^[11]、安全管理等技术,这些技术在很多领域,特别是互联网、IT和电信行业都有应用,技术实现上是可行的。对于电信运营商来说,实现大数据管理网元化的难点不在技术,也不在网络和系统建设上,关键是要梳理清楚企业数据资产,通过企业数据资产目录建立全网数据—服务—应用之间的关联和关系,然后再引入技术实现分布式服务调用和数据整合,这样才能将技术真正转化为生产力。

在数据分散基础上,电信运营商大数据管理网元化架构参考设计兼顾了数据集中,可根据实际情况汇聚数据实

现特定场景应用,同时也充分利用了分布式的优点,基于本地化采集和计算,减少了大数据量传输和存储成本,能够更快地支持全网实时、准实时数据应用。

5 电信运营商大数据管理网元化实施建议

电信运营商将大数据网元化架构付诸实施,不是简单的系统改造、技术升级,最重要的是要从挖掘数据价值的角度重新审视企业数据资源,统一规划、组织和管理,将大数据管理网元化工作纳入企业数据发展战略,组建相关各级各部门参与的实施团队,有序推进实施。在具体执行层面,建议重点关注以下几个方面。

(1) 梳理全网数据资产

电信运营商的数据分布在不同层级的不同部门,没有人真正了解企业到底有哪些数据。要实现数据共享和分布式管理,首先要从数据应用需求入手,梳理全网的数据资产目录,其中用于数据网元化管理的包括数据名称、数据类型、数据层级、数据解释、数据采集地址、数据存储地址、数据计算地址、数据服务调用地址、数据有效期等。

(2) 制定数据服务标准

对于各数据网元,需要进行升级改造,具备封装数据、开放服务的能力,同时开放数据服务访问、监控权限给上级网元。数据服务构建的标准应全网统一,各级网元构建的数据服务上报集团统一管理,包括数据服务名称、数据服务描述、输入参数、输出参数、数据服务和数据资产目录中的数据项之间的对应关系、服务有效期等。

(3) 构建全网数据服务地图

基于数据资产目录和各网元数据服务,构建类似网络拓扑的全网数据地图,建立数据网元、数据服务、数据项之间的层级关联关系,让数据管理者、数据使用者能够对企业数据一目了然,同时还应具备传统网管对数据网元的配置管理、性能管理、故障管理、安全管理、计费管理等功能,真正将现有数据节点纳入网元化管理。

(4) 引入大数据技术

利用分布式存储技术,可实现对存储在各数据汇聚网元/数据网元上的数据进行统一调度和管理,存储到空闲网元,实现数据冗余备份或数据汇聚;数据汇聚网元/数据网元不仅可开放基础数据访问服务,还可根据各自系统计算能力,开放计算服务,进行本地化数据加工处理。数据应用利用分布式并行计算技术,将任务拆分到各地数据汇聚



网元/数据网元进行计算处理;对于实时性应用需求高的网元,可以引入流计算技术对采集的数据进行流式处理,生成服务开放所需要的实时数据。考虑到部署的节点较多且需要根据电信需求定制,建议大数据技术基于开源软件,自主研发。

(5) 做好大数据安全保障

由于要开放全网数据,大数据的安全就变得至关重要。因为数据分布在电信运营商内网,所以对于外部数据应用的访问,应先通过电信运营商的安全认证鉴权。各数据网元/数据汇聚网元封装数据服务时,要先对数据进行脱敏,并设置不同的安全访问级别,不同级别的数据服务只能开放给相应权限的服务调用。同时还应 7×24 h 对全网数据服务地图进行监控,及时发现数据安全异常。

6 结束语

本文从大数据对电信运营商未来发展的重要性入手,针对影响电信运营商大数据价值的数据分散、缺乏共享的问题,结合电信运营商网络管理属地化、数据采集点分散的特征,提出了利用电信传统网络管理优势,引入大数据技术,将分散的数据采集和存储节点的管理网元化、数据服务化的思路。通过梳理全网数据资产、构建数据服务地图,集中监控管理,形成分布在全网,逻辑上集中、可随时随地安全地使用的全量数据,开放给外部合作伙伴,让电信运营商大数据发挥真正的价值。

参考文献

- 傅志华. 大数据在电信行业的应用. <http://www.leiphone.com/news/201411/eAYu5wZWfAKAYvrG.html>, 2014
- Fu Z H. Big data in the application of the telecom industry. <http://www.leiphone.com/news/201411/eAYu5wZWfAKAYvrG.html>, 2014
- Josuttis N M. SOA 实践指南: 分布式系统设计的艺术. 程桦译. 北京: 电子工业出版社, 2008
- Josuttis N M. SOA in Practice: the Art of Distributed System Design. Translated by Cheng H. Beijing: Publishing House of Electronics Industry, 2008
- 韩丁, 沈建京, 万芳等. 基于 SOA 的服务构件封装技术研究. 计算机工程与设计, 2009, 30(7)
- Han D, Shen J J, Wan F, et al. Research on encapsulation of service component based on service-oriented architecture. Computer Engineering and Design, 2009, 30(7)

- Hamilton J. On designing and deploying internet-scale services. Proceedings of the 21st Large Installation System Administration Conference (LISA '07), Dallas, USA, 2007
- 刘向, 王伟军, 李延晖. 云计算环境下信息资源集成与服务系统的体系架构. 情报科学, 2014, 32(6)
- Liu X, Wang W J, Li Y H. Architecture of cloud-computing-based integration & service system for information resource. Information Science, 2014, 32(6)
- Coulouris G, Dollimore J, Kindberg T, et al. 分布式系统: 概念与设计. 金蓓弘, 马应龙等译. 北京: 机械工业出版社, 2013
- Coulouris G, Dollimore J, Kindberg T, et al. Distributed Systems: Concepts and Design. Translated by Jin B H, Ma Y L, et al. Beijing: China Machine Press, 2013
- 韩晶. 大数据服务若干关键技术研究(博士学位论文). 北京邮电大学, 2013
- Han J. Research on some key technologies of big data-as-a-service (doctor dissertation). Beijing University of Posts and Telecommunications, 2013
- 涂新莉, 刘波, 林伟伟. 大数据研究综述. 计算机应用研究, 2014, 31(6)
- Tu X L, Liu B, Lin W W. Survey of big data. Application Research of Computers, 2014, 31(6)
- 孙大为, 张广艳, 郑纬. 大数据流式计算: 关键技术及系统实例. 软件学报, 2014, 25(4)
- Sun D W, Zhang G Y, Zheng W. Big data stream computing: technologies and instances. Journal of Software, 2014, 25(4)
- 覃雄派, 王会举, 李芙蓉等. 数据管理技术的新格局. 软件学报, 2013, 24(2)
- Qin X P, Wang H J, Li F R, et al. New landscape of data management technologies. Journal of Software, 2013, 24(2)
- 吴强, 符涛, 周娜. 移动网络网元分布式部署研究. 电信网络技术, 2010(5)
- Wu Q, Fu T, Zhou N. Study on distributed deployment of network element for mobile network. Telecommunications Network Technology, 2010(5)

[作者简介]



赵钧,男,现就职于中国电信股份有限公司上海研究院,主要研究方向为大数据和云计算。

石屹嵘,男,现就职于中国电信股份有限公司上海研究院,主要研究方向为大数据和云计算。