



基于 Hadoop 的电信 BSS 大数据平台建设研究

陈 娜¹, 徐歆壹¹, 宋红兵², 何 毅³

(1. 中国电信股份有限公司广东研究院 广州 510630; 2. 中国电信集团公司 北京 100032;

3. 亚信联创集团股份有限公司 南京 210013)

摘 要: 当前电信运营商从传统的“语音+短信+增值业务”的业务模式转变为“语音+应用+流量”的业务模式, 电信的核心战略转向流量经营。在该背景下, 大数据是电信 IT 支撑面临的首要技术课题。本文主要研究使用 Hadoop 平台搭建流量经营大数据管理和大数据服务的一种分布式平台, 同时通过实验数据, 论证该平台与传统 IT 架构对比的优势。

关键词: 大数据; 平台; Hadoop; Hbase; 高可用

doi: 10.3969/j.issn.1000-0801.2013.03.008

Research on Big Data System of Telecom Based on the Hadoop

Chen Na¹, Xu Xinyi¹, Song Hongbing², He Yi³

(1. Guangdong Research Institute of China Telecom Co., Ltd., Guangzhou 510630, China;

2. China Telecom Corporation, Beijing 100032, China; 3. AsiaInfo Linkage Group Limited by Share Ltd., Nanjing 210013, China)

Abstract: Telecom operators are transforming their business model from “voice+SMS+VAS” to “voice+application+traffic”, and they are choosing the mobile traffic operation as their key development strategy. In this scenario, big data is becoming the next frontier for the telecom IT support system. A distributed architecture based on Hadoop to provide big data management and services for traffic operation was described, and with the experimental data, demonstration of the outstanding performance of the platform was introduced.

Key words: big data, platform, Hadoop, Hbase, high availability

1 引言

随着智能手机的普及以及 3G+网络的建设, 移动互联网的出现带来了电信行业新的革命, 电信运营商的角色也发生了转变。CSP 从传统的电信产业链的组织和领导者, 转变为管道和功能更加突出的产业的重要参与者。同时客户的行为模式也发生了转变, 从传统的“语音+短信+增值业务”的模式转变为“语音+应用+流量”的模式。

在以上大背景下, 当前的电信运营商系统中存在

大量的数据, 这些数据总体包括: BSS 域的客户/产品/渠道等基本数据、CDR 数据、业务订购数据、工单数据、计费数据、客服数据、客户访问日志; OSS 域的信令数据、网元数据; MSS 域的 ERP 数据、内部办公数据; 公共领域的网页数据、物联网数据等。其中的很多数据已经在智能分析系统中进行分析, 并对当前的运营起到了重要作用。但总体而言, 当前电信运营商单一的纵向 IT 架构已经难以处理近年来体量猛增的非结构化和半结构化数据, 如 DPI 数据等, 亟需在大数据领域开辟一条新的道路。

2 大数据解决方案分类

近年来,各行业业务模式的转变和业务量的增加导致数据的体量和结构发生了质的变化,基于大数据的处理和分析需求也越来越多,越来越苛刻。基于这些复杂而苛刻的大数据需求,目前主流的大数据解决方案主要分为两大类:一体机解决方案和纯软件解决方案。

一体机解决方案提供完全黑盒的硬件和软件集成产品,能很好地支撑已定制的业务,在流计算方面也有很好的表现,但面临高昂的成本和较低的需求响应速度。一体机解决方案的典型产品,如 SAP HANA、Oracle Exadata 等。

纯软件解决方案提供开放的大数据平台软件,软件可部署于 x86 等多种硬件环境下,部署模式可根据具体的业务场景进行设计,且用户可基于大数据的软件平台设计自己的上层应用。其拥有较低的成本和较快的需求响应速度,但纯软件平台将面临更多的软件开发工作。基于大数据处理的技术架构近年来涌现出众多的新技术,当前业界流行的纯软件解决方案典型产品,如 Cassandra、IBM BigInsights、Cloudera Hadoop、Intel Hadoop 发行版、EMC Greenplum 等^[1]。其中,由 Apache 基金会开发的 Hadoop 开

源架构已被众多互联网公司采纳,在一定范围内可以借鉴相关经验。

3 适合电信 BSS 的大数据平台研究

移动互联网的普及导致电信 BSS 需要处理的数据量大幅增加,然而单位容量的数据所创造的利润已经远远小于 2G 时代,因而电信 BSS 大数据平台必须具备较低的成本;移动互联时代终端设备、应用和用户喜好的变化周期大幅缩短,因而电信 BSS 大数据平台必须具备快速多变的需求响应速度。基于以上两点,电信 BSS 大数据平台的建设应选择纯软件的解决方案,设计合适的架构部署于廉价的 x86 服务器上。本文以海量 DPI 数据的入库和查询场景为依据,提出一种基于 Cloudera 4dh Hadoop 平台的分布式大数据平台,通过模拟大数据的实验,搭建电信 BSS 大数据平台,并论证该平台与传统 IT 架构相比在性能、容灾、成本方面的优势。

3.1 平台功能架构设计

在功能架构上,本平台采用“三层结构”的设计思想,在逻辑上按“数据服务层、业务处理层和接入层”3 层结构设计,如图 1 所示。将接入层独立出来,使得平台的访问和使用更灵活方便,易于实现个性化和客户化;将业务处理

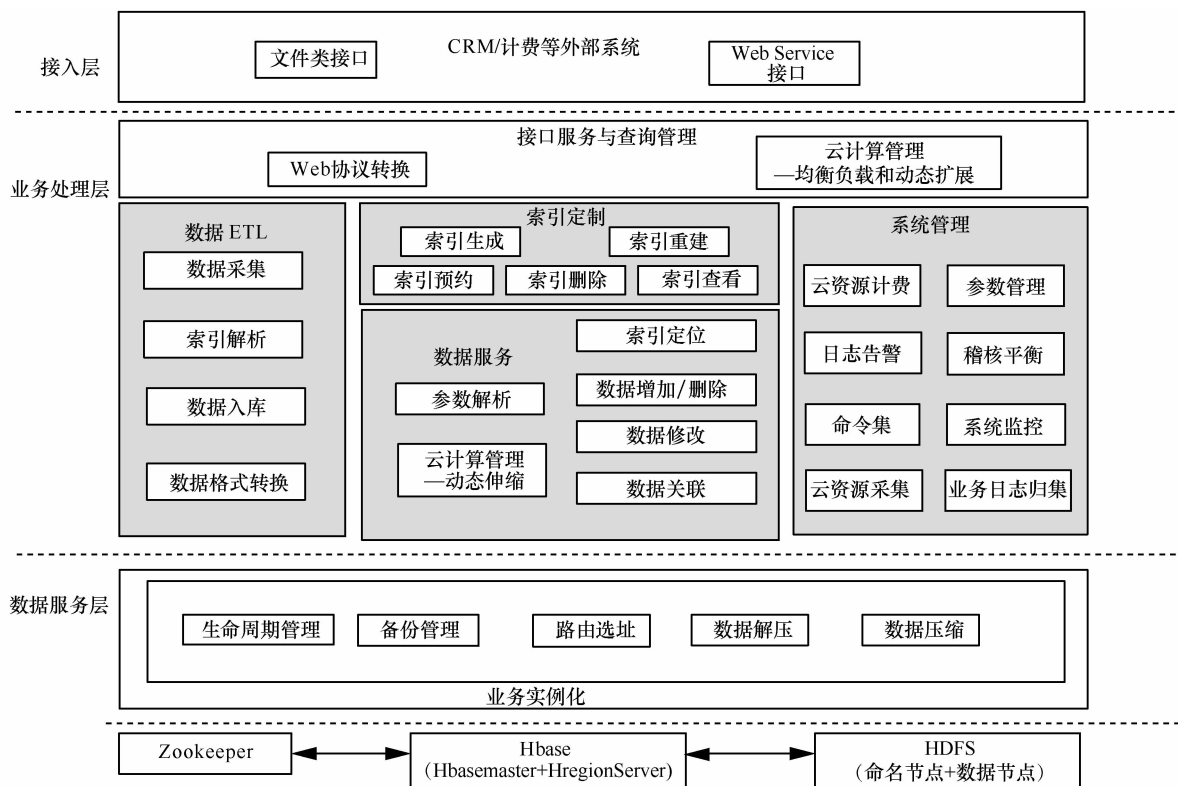


图 1 系统功能架构



层和数据服务层分开,可以屏蔽业务数据的存储、组织和访问的细节,实现业务数据的充分共享,从而实现横向组合,具体介绍如下。

(1)数据服务层

数据服务层为整个平台提供分布式的数据服务,包括分布式文件系统(HDFS)、分布式数据库(Hbase)、分布式协调器(Zookeeper)和业务实例化部分^[1]。本层是大数据平台的底层,主要用于完成 Hadoop 自服务管理以及相关的 Hadoop 参数调优、监控、管理等工作。大数据平台的系统安全、数据安全、容灾等功能也在本层实现。

(2)业务处理层

业务处理层构建于数据服务层之上,封装了整个平台的核心业务处理逻辑,包括接口服务与查询管理、数据 ETL、索引定制、数据服务、系统管理功能。

接口服务功能主要是响应外围系统的查询请求,大数据平台作为大数据中心,为外围系统提供数据,需要一个统一、简单、不断增加的接口,支持 API、Socket、Web Service、Tuxedo 中间件等接口方式。同时对外围系统的数据导入和查询请求,有一个优化的均衡负载策略,包括数据在各节点的均衡分布以及查询请求的均衡负载、动态伸缩和扩展功能,保证大数据平台面对大数据冲击时的性能优化和系统安全。

数据 ETL 一般以文件方式为主,因为大数据平台的数据源广泛且复杂,有结构化、半结构化和非结构化数据,必须支持文件、消息、数据库等接口方式,且具备一定的数据转换和清洗功能以及合并、分拆、统计、压缩、解压等辅助功能。

索引定制和数据服务主要根据不同数据格式的数据特点和业务特性,制定有利于存储和提取的索引方式,追

求存储和查询性能的最优化。

系统管理功能需具备一般数据平台的数据稽核、盘点、安保、生命周期管理等功能,支持数据可查询和可跟踪,保证数据平台的完整性以及相关 Hadoop 平台的监控、维护等工作。

(3)接入层

接入层提供本平台和外部系统的接口,包括输入接口和输出接口。如果大数据平台的外围系统较多且数据源格式复杂,一般会为大数据平台建设一个辅助的数据传输管理平台,一点集中和传输,完成大数据平台与外围系统的数据交换工作。

3.2 大数据平台物理部署架构设计

本平台采用分布式的部署,在部署设计上需要充分发挥各主机的性能。因而部署设计主要分为 Hadoop 平台的部署设计和应用层的部署设计。其中,应用层的部署设计又分为导入应用的部署设计和查询应用的部署设计。

3.2.1 Hadoop 平台的部署设计

Hadoop 平台层的部署设计架构如图 2 所示,具体介绍如下。

(1)主从节点规划

主机 1 和主机 2 作为 Hadoop 平台层的主节点,工作模式为主备模式;主机 3~主机 6 作为 Hadoop 平台层的从节点,工作模式为负载均衡^[2]。

(2)HDFS 部署规划

主机 1 和主机 2 部署 HDFS 的命名节点,配置主机 1 上的命名节点为主用,主机 2 上的命名节点为备用。主机 1~主机 6 上分别部署 HDFS 的数据节点^[3]。

(3)MapReduce 部署规划

主机 1 和主机 2 部署 MapReduce 的 Jobtracker,配置主

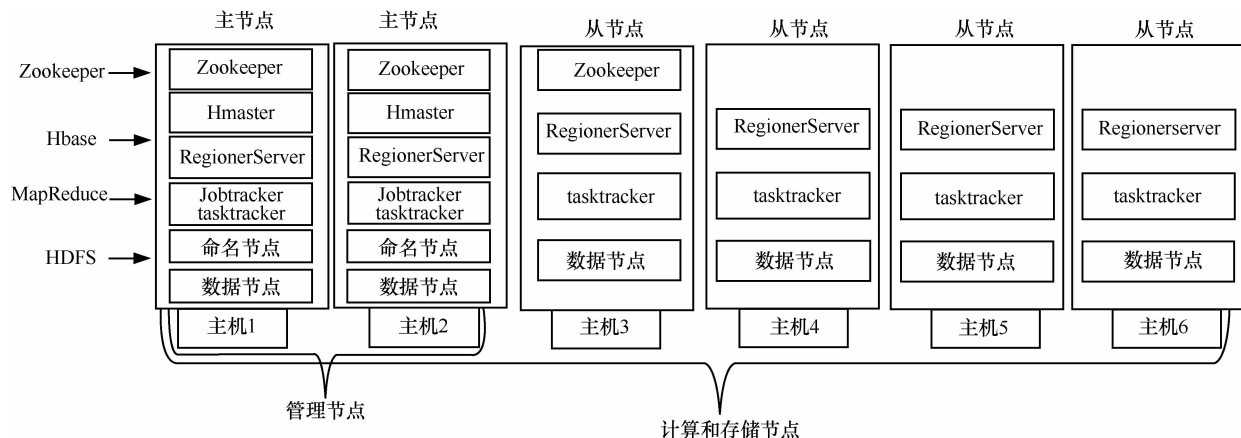


图 2 Hadoop 平台层的部署设计架构

主机 1 上的 Jobtracker 为主用, 主机 2 上的 Jobtracker 为备用。主机 1~主机 6 上分别部署 MapReduce 的 tasktracker^[4]。

(4)Hbase 部署规划

主机 1 和主机 2 部署 Hbase 的 Hmaster,配置主机 1 上的 Hmaster 为主用,主机 2 上的 Hmaster 为备用。主机 1~主机 6 上分别部署 Hbase 的 RegionServer^[5]。

(5)Zookeeper 部署规划

主机 1、主机 2、主机 3 上分别部署 Zookeeper。

3.2.2 应用层的部署设计

3.2.2.1 查询应用的部署设计

查询应用的部署设计架构如图 3 所示, 具体介绍如下。

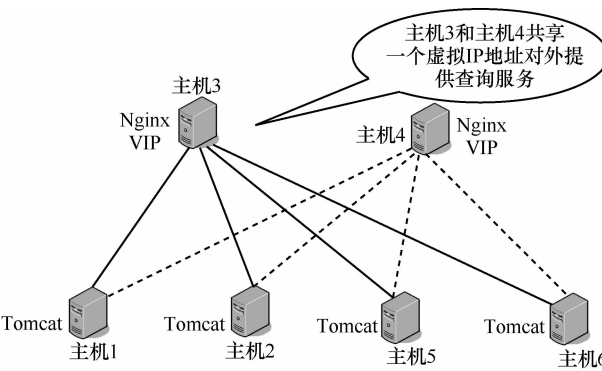


图 3 查询应用的部署设计架构

(1)主从节点规划

主机 3 和主机 4 作为查询应用的主节点,工作模式为主备模式;主机 1、主机 2、主机 5、主机 6 作为查询应用的从节点,工作模式为负载均衡^[6]。

(2)Tomcat 部署规划

主机 1、主机 2、主机 5、主机 6 上分别部署 Tomcat 应用服务器。

(3)Nginx 部署规划

主机 3、主机 4 上分别部署 Nginx 消息分发器,并分别配置 Nginx 将前端业务消息负载均衡转发到主机 1、主机 2、主机 5、主机 6 的 Tomcat 上^[7]。

(4)VIP 部署规划

主机 3 和主机 4 上分别部署虚拟 IP 地址管理软件,并配置两主机的共用虚拟 IP 地址,设置主机 3 为主节点,主机 4 为备节点。配置虚拟 IP 地址切换脚本,当主机 3 发生硬件故障或查询应用故障时,将该虚拟 IP 地址切换到主机 4 上^[8]。

3.2.2.2 入库应用的部署设计

入库应用的部署设计如图 4 所示。主机 1、主机 4、

主机 5、主机 6 上分别部署入库程序,其中主机 1 规划入库 20 亿条记录,主机 4、主机 5、主机 6 规划各入库 10 亿条记录。

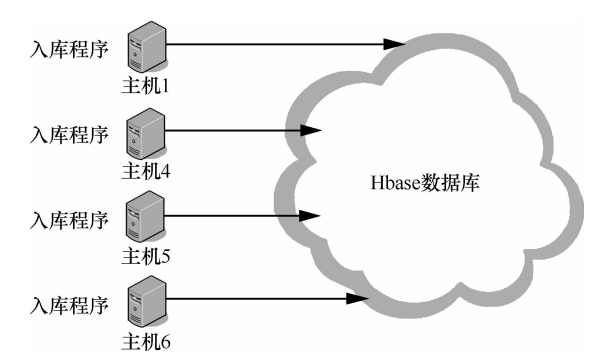


图 4 入库应用的部署设计

4 平台功能和性能验证对比

4.1 验证数据和硬件环境

实验选择的验证数据是流量经营 DPI 数据,DPI 数据是基于 DPI 技术获取的 TCP/IP 应用层数据,基于此能够对数据内容进行分析,因此 DPI 数据在电信流量经营业务背景下具有很重要的意义。实验选择的硬件环境为基于 x86 架构的廉价 PC 服务器。

4.2 数据特性

本实验选择的流量经营 DPI 数据的数据特性见表 1。

表 1 DPI 数据的数据特性

业务类型	文件编码方式	原始记录字段数(条)	需入库的记录字段数(条)	需入库记录数(条)
HTTP 业务	ASCII	65	14	25 亿
WAP 业务	ASCII	40	14	25 亿

4.3 验证方法和评估指标

本实验主要验证 50 亿条 DPI 数据的入库、查询性能和整个系统的高可用性,具体的验证方法和评估指标见表 2。

表 2 具体的验证方法和评估指标

评估点	评估指标
50 亿条数据的入库性能	入库总耗时 (LTC)、每秒入库记录数 (RNPS)
50 亿条数据的查询性能	平均查询响应时长 (ATRT)、平均事务处理能力 (TPS)
高可用性验证	是否有单点问题、RTO 与 RPO
初次投资成本	成本



4.4 硬件配置

本实验的主机环境为 6 台 DL380 Gen8, CPU 为 E5-2630×1 (12 核), 内存 16 GB, 联机磁盘各为 300 GB, 构建分布式文件系统的存储容量各为 2 TB。实验网络带宽采用吉比特以太网网络。

为体现本平台和传统 IT 架构的功能和性能差别, 下文中用于对比测试的传统 IT 架构基于 2 台 HP rx8640, 16CPU/128 GB 硬件, 数据库为 Oracle10g。

4.5 入库性能对比

入库 50 亿条数据, 本平台与传统 IT 架构(小型机+关系数据库)的测试结果对比见表 3。

表 3 入库性能测试结果对比

	LTC	RNPS(万条/s)
传统IT架构	11时46分	11.8
本平台	5时20秒	28.1

4.6 查询性能对比

查询 50 亿条数据, 本平台与传统 IT 架构(小型机+关系数据库)的测试结果见表 4。

表 4 查询性能测试结果对比

查询方法	虚拟用户数(个)	传统 IT 架构 ATRT(s)	传统 IT 架构 TPS	本平台 ATRT(s)	本平台 TPS
查 50 亿条	300	0.457	1 829	0.038	3 769
边入边查 50 亿条	300	0.461	1 684	0.045	3 132
边删边查 50 亿条	300	0.459	1 775	0.039	3 669

4.7 高可用性对比

本平台在高可用性方面的测试结果见表 5。

表 5 本平台高可用性测试结果

验证点	是否有单点问题	RTO	RPO
停查询应用的非主节点	否	0	0
停查询应用的主节点	否	3 s	0
停 Hadoop 的非主节点	否	0	0
停 Hadoop 的主节点	否	0	0

在传统 IT 架构(小型机+关系数据库)下, 数据库服务器的高可用性测试结果见表 6。

表 6 传统架构下的高可用性测试结果

验证点	是否有单点问题	RTO	RPO
停查询应用的主节点	是	N/A	N/A
停数据库服务器的主节点	否	169 s	0

4.8 成本对比

本平台的初次投资成本设置见表 7。

表 7 本平台的初次投资成本设置

成本项目	数量	配置	总价(元)
DL380 Gen8 服务器	6	E5-2630×1 (12 核) 16 GB 内存、2.3 TB 存储空间	48 万
操作系统	6	Redhat	6 万
系统软件	1	Hadoop+Hbase+Zookeeper	0
App 中间件	1	Tomcat	0
合计			54 万

传统 IT 架构的初次投资成本设置见表 8。

表 8 传统 IT 架构的初次投资成本设置

成本项目	数量	配置	总价(元)
HP rx8640 服务器	2	16CPU/128 GB 内存	200 万
存储	1	12 TB (HP EVA8400)	10 万
数据库软件	1	Oracle10g	80 万
App 中间件	1	Tomcat	0
合计			290 万

5 结束语

本文论述了移动互联网背景下一种适合电信 BSS 领域的大数据平台, 该平台在技术上基于 x86 架构的 PC 服务器和 Hadoop 开源框架。同时通过真实数据的测试对比, 可以看出该平台与传统 IT 架构相比, 可节约 80% 的成本, 且具有更高的入库、查询性能和高可用性。

参考文献

- 1 Borthakur D. Hadoop distributed file system. <http://hadoop.apache.org/>, 2007
- 2 White T. Hadoop the Definitive Guide. United States of America: O'Reilly, 2009
- 3 Owen O'Malley. Programming with Hadoop's MapReduce. ApacheCon EU, 2008
- 4 Apache. WelcometoHadoop. <http://hadoop.apache.org/>, 2007
- 5 Borthakur D. The Hadoop distributed file system: architecture and design. http://hadoop.apache.org/common/docs/r0.20.0/hdfs_design.pdf, 2011
- 6 David E Culler, Jaswinder Pal, Singh Anoop 等著. 李晓明等译. 并行计算机体系结构: 硬件/软件结合的设计与分析 (第 2 版). 北京: 机械工业出版社, 2003
- 7 曹羽中. 用 Hadoop 进行分布式并行编程 (第 1 部分). <http://www.ibm.com/developerworks/cn/opensource/os-cn-hadoop1/index.html>, 2008
- 8 朱珠. 基于 Hadoop 的海量数据处理模型研究和应用. 北京邮电大学硕士学位论文, 2008

(下转第 47 页)

全威胁的基础上,建立云计算模式下移动互联网的安全保护体系和体系架构,研究多层次的云计算模式下移动互联网安全保障方法,提出相关安全技术标准与测评体系,解决移动终端安全、客户数据安全和隐私保护、虚拟化运行环境安全和风险评估及安全监管等问题,增强了云计算模式下移动互联网的机密性、完整性、可验证性和可用性,为基于云计算模式下的移动互联网业务发展提供了参考。

参考文献

- 1 中国通信标准化协会. 移动环境下云安全技术研究, 2012
- 2 张云勇, 陈清金, 潘松柏等. 云计算安全关键技术. 电信科学, 2010(8)

[作者简介]



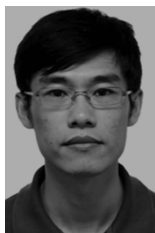
房秉毅,男,博士,中国联通集团研究院高级工程师,主要研究方向为云计算、核心网新技术,已发表论文 20 余篇。



张云勇,男,博士后,中国联通集团研究院平台与云计算研究中心主任、高级工程师,中国通信学会、电子学会、计算机学会高级会员,中国人工智能学会会员,主要研究方向为下一代开放网络、固定移动融合核心网、移动互联网及业务、公共运算,已出版中文论著 14 部、英文论著 1 部,已发表论文 58 篇。



吴俊,男,博士,北京邮电大学副教授,主要研究方向为移动互联网应用、云计算商业价值评价等。



徐雷,男,博士,中国联通集团研究院高级工程师,主要研究方向为云计算、移动互联网新技术。

(收稿日期:2013-02-20)

(上接第 40 页)

[作者简介]



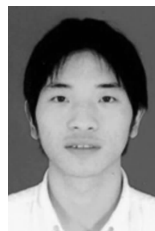
陈娜,女,硕士,中国电信股份有限公司广东研究院高级工程师,主要从事电信 IT 支撑系统的相关研究工作。



宋红兵,女,硕士,中国电信集团公司工程师,长期从事中国电信业务支撑运营管理和研究工作。



徐歆壹,男,学士,现就职于中国电信股份有限公司广东研究院,主要从事 BSS 领域的研究工作。



何毅,男,硕士,亚信联创集团股份有限公司工程师,主要从事电信 IT 支撑系统的设计和开发工作。

(收稿日期:2013-03-05)