



基于 MapR 的 IPTV 用户收视行为分析的方案与实践

方艾, 张玉忠, 徐雄, 金铎

(中国电信股份有限公司广州研究院, 广东 广州 510630)

摘要: 互动电视业务开放平台是运营商的互联网+IPTV 转型实践, 是用户、终端、内容、商家等之间相互连接的枢纽。平台目前拥有百万级的活跃用户, 每天产生接近 1 TB 的用户收视行为、系统运行日志等数据, 这些数据可用于业务运营分析、用户画像、系统运维分析等。MapR 是适合此类分析的大数据处理框架之一, 结合实际需求介绍 MapR 的主要特点、核心组件, 阐述基于 MapR 的分析系统的架构设计及应用效果。

关键词: 大数据; 互联网+; MapR; Apache Drill; SQL on Hadoop

中图分类号: TP302.1

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2017030

Scheme and practice of IPTV user viewing behavior analysis based on MapR

FANG Ai, ZHANG Yuzhong, XU Xiong, JIN Duo

Guangzhou Research Institute of China Telecom Co., Ltd., Guangzhou 510630, China

Abstract: Interactive TV business open platform is the internet and IPTV transformation practice for operators. It's the interconnected hub between users, terminals, content, business and other. The platform now has millions of active users, generating nearly 1 TB of daily viewing activity and system log data. They can be used for business analysis, user portraits and system analysis. MapR was one of the large data processing frameworks suitable for this kind of analysis. The main features and core components of MapR were introduced combined with actual demand, and the architecture design and application effect of analysis system based on MapR was elaborated.

Key words: big data, Internet+, MapR, Apache Drill, SQL on Hadoop

1 引言

IPTV 面向移动互联网转型, 互动电视应用将成为一个重要的突破口, IPTV 需要建立用户与终端、用户与用户、用户与内容、人与商家之间的连接, 让用户可以充分参与内容、发现服务、社交互动, 使电视真正成为家庭媒体娱乐中心。互动电视业务开放平台是互动电视应用的解决方案在广东 IPTV 的实践, 目前平台用户每日 PV (page view, 页面浏览量) 超过 7 亿, 每天产生近 1 TB 的数据文件, 内容

包括用户收视行为日志、系统运行日志等。通过对数据的实时处理、关键质量指标分析, 通过系统监控, 可指导系统的持续优化, 为用户提供更好的体验; 从用户行为与业务使用上分析, 输出业务运营指标, 可为业务优化提供依据, 并形成反馈、验证的闭环, 从用户的行为中分析挖掘用户需求、分析用户特点, 可为用户提供更有价值的、个性化的服务。

实现上述在运维、运营、用户关怀等方面业务需求的一个难点是系统产生的海量数据, 由于数据产生的速度

快、数据量大,对数据分析系统提出了较高要求:一方面要求高时效性,做到实时采集应用服务器集群产生的数据文件,并实时完成数据清洗、关联、加载等预处理;另一方面要求高效率分析,即能够方便、高效地进行数据分析,快速得到结果。

数据应用的场景可分为两大类:一类是传统的分析型应用,如针对历史数据的统计、分析与建模;另一类是基于数据与事件驱动的运营型应用,将用户的行为或事件产生的实时数据与历史积累的数据结合,实现快速响应的实时应用,如实时用户趋势、实时推荐等。

2 MapR 框架选型

2.1 系统目标

2.1.1 业务运营分析

系统提供从用户、产品、内容等不同角度,从地区、时间等不同维度进行全方位多层次的实时分析能力,快速为业务运营提供决策依据。

分析工作可分为实时与非实时两类:实时分析提供了对当前用户收视行为的实时监控与统计分析,例如针对收视率的实时分析,实时获知当前全省及各个城市最热门的直播频道与节目,实时获知用户流向等。实现实时发现热点频道、实时向用户提醒热门节目信息,或指导后台视频源服务动态调度等。离线分析提供系统日常运营报表,包括平台 UV、PV、TopN 等业务统计。

运营人员通过查询界面,在线查询、分析验证运营效果。

2.1.2 用户分析建模

基于用户行为及业务数据的建模形成用户特征画像,感知用户兴趣,为个性化服务、精准营销提供基础。例如通过用户最常看的频道与节目内容,进行在线时段/时长等分析,指导在合适的时间推送合适的内容,屏蔽不恰当的

内容。通过用户聚类分析,衍生出社交应用,如圈子推荐、活动组织等。

2.1.3 系统运维分析

针对系统运维层面提供系统运作的关键质量指标,为保障系统正常运作提供支撑,结合业务负载、系统资源等分析提供系统优化依据并验证系统优化的效果。实时分析应用于系统异常的实时告警,当设定的指标超过一定阈值时,提醒运维人员处理。通过对系统负载、响应时间、用户体验指标等的分析与跟踪,指导系统的优化,对比优化前后的分析结果并对结果进行确认。

2.2 框架选型

借助大数据处理框架 Hadoop 可实现上述目标,目前 Hadoop 存在多个商业发行版,例如 cloudera 版本(CDH)、Hortonworks 版本(HDP)、MapR 等,但它们各有侧重,需选择适合本系统的框架,通过对比主流大数据框架的特点与差异,结合本系统的实际需求,最终选取 MapR 作为数据分析的基础框架,主要原因有:MapR 建立在适合实时分析的文件系统上,满足本系统的应用要求,支持 Apache Drill、Spark 等工具进行方便高效的数据分析,Drill 可查询文本文件,特别适合本系统当前数据变动频繁的情况。

2.2.1 MapR 主要特点

MapR 提出了融合分析型应用与运营型应用的架构^[1],如图 1 所示,通过 3 个组件——MapR-FS(文件系统)、MapR-DB(数据库)和 MapR streams(流处理)构建了底层数据框架,在此基础上,支持各种开源与商业的分析工具,如 Hadoop 生态、Spark 高性能运算框架、Apache Drill 分析工具等。另外,还提供管理监控的界面或接口,可与第三方监控工具结合,方便进行集群的管理和监控。

2.2.2 MapR-FS 介绍

MapR-FS 采用标准 POSIX 文件系统,与 HDFS 相比^[2],

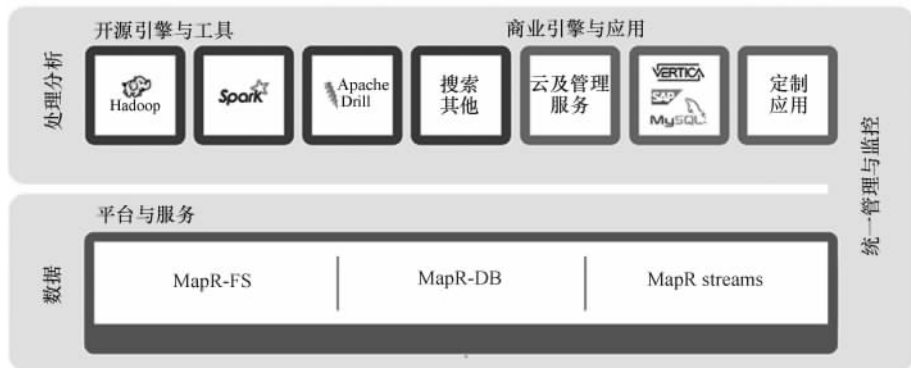


图 1 MapR 架构示意



MapR-FS 采用不同的架构,一是解决 namenode(文件命名空间管理)单点的问题,实现 namenode 角色的分布式;二是实现更灵活高效的文件操作,特别是支持多次读写。由于 HDFS 文件采用一次写入方式,因此无法直接修改,需删除后再写入,这种机制不适合实时数据的写入,例如在每秒更新的场景下,会造成许多文件碎片,这与设计初衷不符。MapR-FS 则支持全面的读写操作,并采用 Linux 方式的 ACE 权限控制。通过 NFS(network file system,网络文件系统)服务器,MapR-FS 可被挂载到 Linux 的本地目录,像操作本地文件一样操作。相比 HDFS 通过客户端或 REST 接口操作数据的方式,更为方便。另外 MapR-FS 实现了对用户透明的压缩机制,实测中,纯文本占用的空间仅是原文件的 1/3,正好弥补了冗余所需的空間。

MapR-FS 由以下几个部分构成^[9]。

- 存储池(storage pool):由一组磁盘构成,一般是 3 块磁盘,用于写数据。
- 容器(container):抽象实体,一类用于存放数据文件、文件块、表块等,另一类用于保存命名空间等信息,相当于 HDFS 的 namenode,但由于分布在多节点中,避免了 namenode 单点的问题。
- 容器位置数据库(CLDB):用于记录容器位置的服务。
- 卷(volume):管理实体,用于存储和组织容器,也用于分布元数据、设置权限。一般由一个命名容器和若干个数据容器组成。图 2 示意一个跨越了 3 个节点的卷。卷可以被快照,发生问题时可用于回滚,也可以镜像到其他集群作为冗余,或在异地集群上实现灾备。

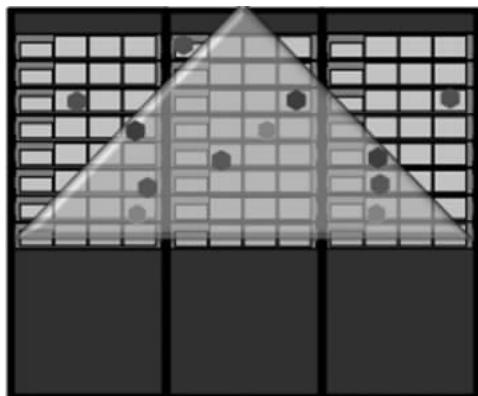


图 2 跨越 3 个节点的卷示意

2.2.3 Apache Drill 介绍

Apache Drill 是 SQL on Hadoop^[4]的方式之一,与 CDH

的 Impala 类似,是对 Google Dremel 的开源实现,旨在提供低时延的在线数据查询与分析,支持 ANSI 标准 SQL。本系统采用 Apache Drill 的另一个原因是其支持对文本文件的操作,不依赖 Hive 表的元数据,通过定义存储格式(storage),主要是文件路径、文件格式(如文本、CSV、JSON 等)、分隔符等,可使用 SQL 语法查询 MapR-FS 或本地文件。由于本系统在发展初期变化较快,数据内容与格式不时发生变动,如采用需预定义表结构的存储方式,如 Hive,则由于数据更新频繁,带来较大的工作量。采用 Apache Drill 时,可将数据解析成 JSON 格式,支持灵活的结构,大大加快了开发流程。当然,由于每个字段都带有键值,存储空间要求更多。

程序或用户可通过 C++API、REST 接口、Drill Shell 或 ODBC/JDBC 与 Drill 交互。在每个节点上部署 Drill-bit 服务,Drill-bit 接受客户端的 SQL 查询请求后,先解析成为逻辑计划,经过优化后转换成物理计划。物理计划进一步被并行器组织成一个执行树^[9](如图 3 所示),执行树的叶子是一批并行运行的分片,负责扫描、过滤数据,并通过树枝一层层地返回结果。

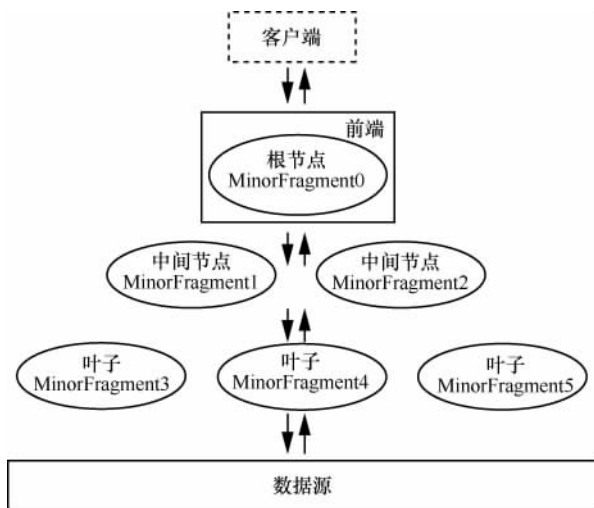


图 3 Drill 执行树示意

3 系统架构与实践

3.1 MapR 部署

本系统将 MapR 基础框架部署在大数据定制型物理机上,每台服务器配备了较大内存和硬盘空间,CPU 则为中等配置。由于使用了 MapR 的社区版,未有高可用(high available,HA)模式。

系统部署示意如图 4 所示。系统部署的主要组件介绍如下。

- YARN(资源调度):每个节点都部署节点管理器,其中一台还作为资源管理器。
- MapR-FS(文件系统):每个节点上部署了文件服务器,除操作系统的磁盘外,所有磁盘都作为存储使用。其中一个节点作为 NFS 网关,另一个节点作为 CLDB(管理集群容器位置的数据库)。
- 分析工具 Apache Drill,每个节点安装了 Drill-bit 服务。
- Web 服务器:选取一个节点安装 Web 服务器,作为集群管理界面的入口。
- Zookeeper (集群协同服务),部署在奇数个节点上。

3.2 架构设计

系统框架设计如图 5 所示。数据主要来源是互动电视平台应用服务器产生的日志。最初,日志通过实时流方式进入平台,随着数据量的增大,实时流采集方式由于受处理能力、网络带宽、网络稳定性等因素的影响,有可能发生丢失,为确保历史数据的完整性,后续加入了准实时同步方式。实时应用则可按需订阅数据,只获取部分指定数据内容或字段,减少网络传输的流量。数据采集使用了自主开发的程序,未使用 Flume 或 Kafaka 等组件。

数据经过 ETL 后存储到 MapR-FS,对历史数据的 ETL 流程主要采用了 Hadoop 流处理,结果以 JSON 格式存放在 MapR-FS。对实时数据的处理则按应用所需提取。



图 4 系统部署示意

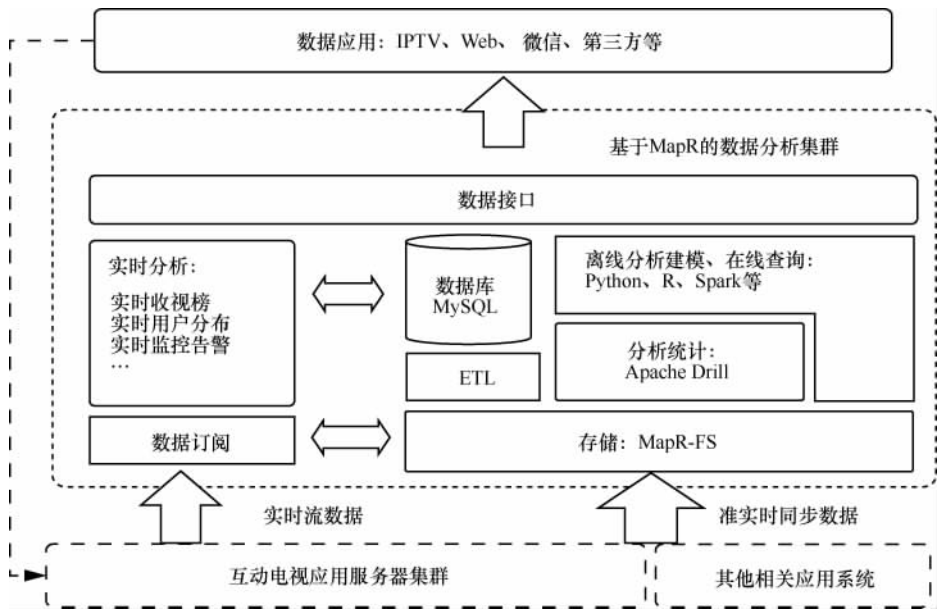


图 5 系统框架设计



3.3 分析模型

系统支持不同类型的分析任务,按数据更新频率可分为实时分析与离线分析:实时分析一般指数据一旦产生即进行处理,输出计算结果的频率一般为秒级;离线分析一般指时效性要求不高的批量分析作业,涉及数据量较大的历史数据。计算方法包括过滤、筛选、分类、汇总、关联、累计等。典型的分析模型如图6所示:从数据源筛选必要的字段,进行一系列的计算,例如对用户账号、所在频道进行频道用户数累计等,每隔一定间隔(如以1 s为时间窗口)输出结果;每次计算的耗时应小于时间窗口,以确保新产生的数据能够被及时处理。

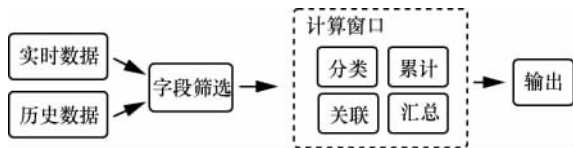


图6 典型的分析模型

3.4 应用实践

(1) 实时分析

直播频道实时收视榜单是本系统实时分析的应用之一,目的是随时了解当前最受欢迎的频道以及用户的流向,收视行为数据中记录了用户的账号、用户进入的频道、进入的时间点等信息,系统每隔4 s计算一次当前各个频道的用户数与占比、进出各个频道的用户数等,之后通过Web socket输出到Web页面,展示成图表。图7所示的界面显示了在某个时刻,某个城市最受欢迎的前几个频道及相应的用户数、占比以及当时进入、离开的人数。该榜单每4 s更新一次。

目前,该应用平均每秒需处理上万条数据记录,总体的计算耗时为毫秒级,满足了实时输出的要求。

(2) 离线分析

业务运营报表是离线分析的一类应用,目前系统每天需完成10多个分析作业,输出包括每小时的用户数、访问量,频道用户分布,城市用户分布,机顶盒开机分布等多个

报表,随着业务发展和数据量的增长,参与计算的数据量从200 GB增长到500 GB,相应的作业处理总时间从25 min增加到41 min(如图8所示),显示出较好的性能,满足了需求。

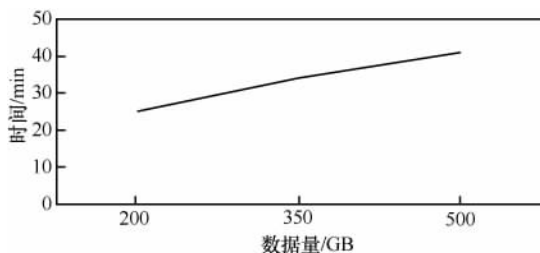


图8 分析作业运算性能

4 结束语

MapR(v5.1)提供了运行稳定、操作方便的基础环境,MapR-FS发挥了其优势作用,系统运行5个月来基本稳定,目前仅使用了数量较少、性能中等的几台服务器,实现了多项应用需求,特别是实时榜单、实时告警等应用。目前使用的Apache Drill(v1.6)仍存在一些不足,需要特别的注意或处理,例如对空值、中文、同一字段含有不同类型数据的处理,文件目录中存在空文件时出错等,需要上、下游的流程配合。

系统建设初期一般采用默认配置,随着数据量的增长与任务的增加,系统产生瓶颈需进行调优,可先在操作系统参数、集群配置、分析工具参数配置等方面挖掘纵向潜力,再考虑水平拓展。借助资源监控与测试分析瓶颈。本系统运行中出现过局部内存不足等问题,已通过调整配置^[6]解决。随着业务需求的发展和数据量的增加,未来也可考虑使用MapR streams进行实时流计算、尝试应用MapR-DB等,结合生态中的相关工具,实现更优化的架构。

参考文献:

- [1] MapR converged data platform[EB/OL]. (2015-10-01)[2016-11-10]. <https://www.mapr.com/products/mapr-converged-data-platform>.

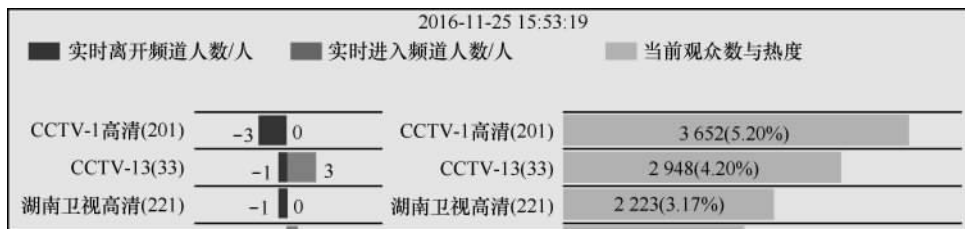


图7 实时收视分析结果

- [2] MapR-FS vs. HDFS: the 5-minute guide to understanding their differences-whiteboard walkthrough [EB/OL]. (2016-01-10)[2016-11-10]. <https://www.mapr.com/blog/mapr-fs-vs-hdfs-5-minute-guide-understanding-their-differences-whiteboard-walkthrough>.
- [3] MapR-FS provides high performance enterprise-grade storage for big data [EB/OL]. (2016-02-10)[2016-11-10]. <https://www.mapr.com/products/mapr-fs>.
- [4] SQL on Hadoop[EB/OL]. (2015-02-10)[2016-11-10]. <https://www.mapr.com/why-hadoop/sql-hadoop>.
- [5] Drill query execution[EB/OL]. (2016-01-10)[2016-11-10]. <https://drill.apache.org/docs/drill-query-execution/>.
- [6] Performancetuningintroduction[EB/OL]. (2016-02-10)[2016-11-10]. <https://drill.apache.org/docs/performance-tuning-introduction/>.

[作者简介]



方艾(1981-),男,中国电信股份有限公司广州研究院工程师,主要研究方向为移动互联网与大数据应用。



张玉忠(1974-),男,中国电信股份有限公司广州研究院工程师,主要研究方向为移动互联网与大数据应用。



徐雄(1974-),男,中国电信股份有限公司广州研究院高级工程师,主要研究方向为移动互联网应用。



金锋(1970-),男,中国电信股份有限公司广州研究院高级工程师,主要研究方向为移动互联网应用。