(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 110362544 A
(43)申请公布日 2019.10.22

(21)申请号 201910447683.8

(22)申请日 2019.05.27

(71)申请人 中国平安人寿保险股份有限公司
地址 518000 广东省深圳市福田区益田路5033号平安金融中心14、15、16、41、44、45、46层

(72)发明人 石晓龙

(74)专利代理机构 深圳市赛恩倍吉知识产权代理有限公司 44334
代理人 杨毅玲

(51)Int.Cl.
G06F 16/17(2019.01)
G06F 16/18(2019.01)
G06F 16/182(2019.01)
G06F 9/54(2006.01)

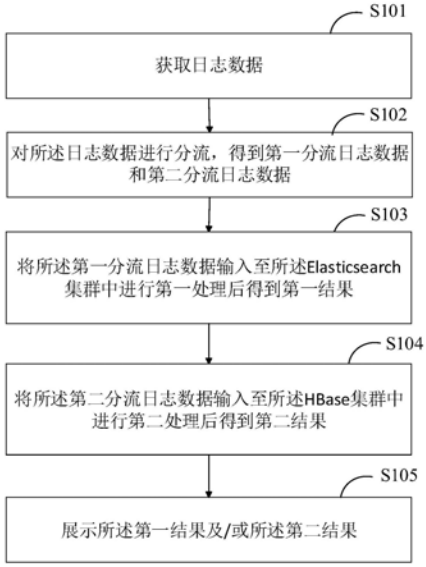
权利要求书2页 说明书9页 附图4页

(54)发明名称

日志处理系统、日志处理方法、终端及存储介质

(57)摘要

本发明实施例提供一种日志处理系统，包括日志采集模块，用于获取日志数据；Kafka日志分发集群，用于对日志数据进行分流，得到第一分流日志数据和第二分流日志数据；Elasticsearch集群，用于对所述第一分流日志数据进行第一处理后得到第一结果；HBase集群，用于对所述第二分流日志数据进行第二处理后得到第二结果；结果展示模块，用于展示所述第一结果及/或所述第二结果。本发明实施例还提供一种日志处理方法、终端及计算机可读存储介质。利用本发明实施例，由Elasticsearch集群存储短期的日志数据，进行日志实时处理，由HBase集群主要负责离线的日志数据处理，从而提高了日志处理效率。



CN 110362544 A

1.一种日志处理系统,其特征在于,所述日志处理系统包括:

日志采集模块,用于获取日志数据;

Kafka日志分发集群,用于对日志数据进行分流,得到第一分流日志数据和第二分流日志数据;

Elasticsearch集群,用于对所述第一分流日志数据进行第一处理后得到第一结果;

HBase集群,用于对所述第二分流日志数据进行第二处理后得到第二结果;

结果展示模块,用于展示所述第一结果及/或所述第二结果。

2.一种利用如权利要求1所述的日志处理系统进行日志处理的日志处理方法,其特征在于,所述日志处理方法包括:

获取日志数据;

对所述日志数据进行分流,得到第一分流日志数据和第二分流日志数据;

将所述第一分流日志数据输入至所述Elasticsearch集群中进行第一处理后得到第一结果;

将所述第二分流日志数据输入至所述HBase集群中进行第二处理后得到第二结果;

展示所述第一结果及/或所述第二结果。

3.根据权利要求2所述的日志处理方法,其特征在于,所述对所述日志数据进行分流,得到第一分流日志数据和第二分流日志数据包括:

通过所述Kafka日志分发集群将获取到的日志数据转换为Kafka消息队列;

对所述Kafka消息队列中缓存的日志数据进行分流处理,分为实时日志数据与非实时日志数据,其中,所述第一分流日志数据为实时日志数据,所述第二分流日志数据为非实时日志数据。

4.根据权利要求3所述的日志处理方法,其特征在于,在所述将所述第一分流日志数据输入至所述Elasticsearch集群中进行第一处理后得到第一结果之前,所述方法还包括:

接收所述Kafka消息队列中缓存的不同topic里的实时日志数据;

通过Logstash日志解析模块按照预设解析规则对所述实时日志数据进行解析操作。

5.根据权利要求4所述的日志处理方法,其特征在于,所述将所述第一分流日志数据输入至所述Elasticsearch集群中进行第一处理后得到第一结果包括:

通过所述Elasticsearch集群保存所述经解析处理后的所述第一分流日志数据;

对所述第一分流日志数据进行实时日志数据处理,得到实时日志数据处理结果,其中,所述实时日志数据处理包括以下中的一种或多种的组合:实时检索处理、实时告警处理与在线统计处理。

6.根据权利要求3所述的日志处理方法,其特征在于,在所述将所述第二分流日志数据输入至所述HBase集群中进行第二处理后得到第二结果之前,所述方法还包括:

读取预定解析规则;

通过Spark集群按照预定解析规则对所述第二分流日志数据进行解析操作。

7.根据权利要求6所述的日志处理方法,其特征在于,所述将所述第二分流日志数据输入至所述HBase集群中进行第二处理后得到第二结果包括:

通过所述HBase集群保存所述经解析处理后的所述第二分流日志数据;

对所述第二分流日志数据进行离线日志数据处理,得到离线日志数据处理结果,其中,

所述离线日志数据处理包括以下中的一种或多种:离线分析处理、日志备份处理与日志还原处理。

8.根据权利要求2所述的日志处理方法,其特征在于,所述展示所述第一结果及/或所述第二结果包括:

获取所述日志处理系统当前正在处理的日志数据信息;

当所述日志处理系统当前正在处理的日志数据信息为第一分流日志数据时,展示所述第一结果;

当所述日志处理系统当前正在处理的日志数据信息为第二分流日志数据时,展示所述第二结果。

9.一种终端,其特征在于,所述终端包括处理器,所述处理器用于执行存储器中存储的计算机程序时实现如权利要求1所述日志处理系统或者实现如权利要求2至8中任意一项所述日志处理方法。

10.一种计算机可读存储介质,所述计算机可读存储介质上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现如权利要求1所述日志处理系统或者实现如权利要求2至8中任意一项所述日志处理方法。

日志处理系统、日志处理方法、终端及存储介质

技术领域

[0001] 本发明涉及日志生成过程优化技术领域，具体涉及一种日志处理系统、日志处理方法、终端以及计算机可读存储介质。

背景技术

[0002] 随着计算机和网络的发展，日志数据的数据处理量越来越大，日志数据的数据量级通常是百万级以上，甚至是百万亿级、千万亿级以上。针对如此庞大的日志数据体系，首先对日志数据的处理提到了较高的要求。现有技术中，日志系统一般采取两种方案，一种是利用ELK的架构，ELK是一种以Elasticsearch(实时全文搜索和分析引擎)、Logstash(用来搜集、分析与过滤日志的工具)以及Kibana(一种基于Web的图形界面，用于搜索、分析和可视化存储在Elasticsearch指标中的日志数据)三者为核心套件的基本架构，这种方式实时性好、查询方便，但是由于Elasticsearch查询是Http协议，所以不适合大批量的对外提供日志；还有一种是基于Hadoop的架构，这种方式可以将日志汇聚起来，然后对外提供日志文件，可是实时性较差，查询也不够方便。

[0003] 因此，现在亟需一种针对日志数据处理的改进方法。

发明内容

[0004] 鉴于以上内容，有必要提供一种日志处理系统、日志处理方法、终端以及计算机可读存储介质，其可以将ELK生态与Hadoop生态结合，由Elasticsearch集群存储短期的日志数据，主要负责日志数据的实时处理，由HBase集群负责离线的日志数据处理，提高了日志处理效率。

[0005] 本发明实施例第一方面提供一种日志处理系统，所述日志处理系统包括：

[0006] 日志采集模块，用于获取日志数据；

[0007] Kafka日志分发集群，用于对日志数据进行分流，得到第一分流日志数据和第二分流日志数据；

[0008] Elasticsearch集群，用于对所述第一分流日志数据进行第一处理后得到第一结果；

[0009] HBase集群，用于对所述第二分流日志数据进行第二处理后得到第二结果；

[0010] 结果展示模块，用于展示所述第一结果及/或所述第二结果。

[0011] 本发明实施例第二方面提供一种利用上述的日志处理系统进行日志处理的日志处理方法，所述日志处理方法包括：

[0012] 获取日志数据；

[0013] 对所述日志数据进行分流，得到第一分流日志数据和第二分流日志数据；

[0014] 将所述第一分流日志数据输入至所述Elasticsearch集群中进行第一处理后得到第一结果；

[0015] 将所述第二分流日志数据输入至所述HBase集群中进行第二处理后得到第二结

果；

[0016] 展示所述第一结果及/或所述第二结果。

[0017] 进一步的，在本发明实施例提供的上述日志处理方法中，所述对所述日志数据进行分流，得到第一分流日志数据和第二分流日志数据包括：

[0018] 通过所述Kafka日志分发集群将获取到的日志数据转换为Kafka消息队列；

[0019] 对所述Kafka消息队列中缓存的日志数据进行分流处理，分为实时日志数据与非实时日志数据，其中，所述第一分流日志数据为实时日志数据，所述第二分流日志数据为非实时日志数据。

[0020] 进一步的，在本发明实施例提供的上述日志处理方法中，在所述将所述第一分流日志数据输入至所述Elasticsearch集群中进行第一处理后得到第一结果之前，所述方法还包括：

[0021] 接收所述Kafka消息队列中缓存的不同topic里的实时日志数据；

[0022] 通过Logstash日志解析模块按照预设解析规则对所述实时日志数据进行解析操作。

[0023] 进一步的，在本发明实施例提供的上述日志处理方法中，所述将所述第一分流日志数据输入至所述Elasticsearch集群中进行第一处理后得到第一结果包括：

[0024] 通过所述Elasticsearch集群保存所述经解析处理后的所述第一分流日志数据；

[0025] 对所述第一分流日志数据进行实时日志数据处理，得到实时日志数据处理结果，其中，所述实时日志数据处理包括以下中的一种或多种的组合：实时检索处理、实时告警处理与在线统计处理。

[0026] 进一步的，在本发明实施例提供的上述日志处理方法中，在所述将所述第二分流日志数据输入至所述HBase集群中进行第二处理后得到第二结果之前，所述方法还包括：

[0027] 读取预定解析规则；

[0028] 通过Spark集群按照预定解析规则对所述第二分流日志数据进行解析操作。

[0029] 进一步的，在本发明实施例提供的上述日志处理方法中，所述将所述第二分流日志数据输入至所述HBase集群中进行第二处理后得到第二结果包括：

[0030] 通过所述HBase集群保存所述经解析处理后的所述第二分流日志数据；

[0031] 对所述第二分流日志数据进行离线日志数据处理，得到离线日志数据处理结果，其中，所述离线日志数据处理包括以下中的一种或多种：离线分析处理、日志备份处理与日志还原处理。

[0032] 进一步的，在本发明实施例提供的上述日志处理方法中，所述展示所述第一结果及/或所述第二结果包括：

[0033] 获取所述日志处理系统当前正在处理的日志数据信息；

[0034] 当所述日志处理系统当前正在处理的日志数据信息为第一分流日志数据时，展示所述第一结果；

[0035] 当所述日志处理系统当前正在处理的日志数据信息为第二分流日志数据时，展示所述第二结果。

[0036] 本发明实施例第三方面还提供一种终端，所述终端包括处理器，所述处理器用于执行存储器中存储的计算机程序时实现上述所述日志处理系统或者实现上述任意一项所

述日志处理方法。

[0037] 本发明实施例第四方面还提供一种计算机可读存储介质，所述计算机可读存储介质上存储有计算机程序，所述计算机程序被处理器执行时实现上述所述日志处理系统或者实现上述任意一项所述日志处理方法。

[0038] 本发明实施例提供一种日志处理系统、日志处理方法、终端以及计算机可读存储介质，所述日志处理系统包括：日志采集模块，用于获取日志数据；Kafka日志分发集群，用于对日志数据进行分流，得到第一分流日志数据和第二分流日志数据；Elasticsearch集群，用于对所述第一分流日志数据进行第一处理后得到第一结果；HBase集群，用于对所述第二分流日志数据进行第二处理后得到第二结果；结果展示模块，用于展示所述第一结果及/或所述第二结果。利用本发明实施例，其可以将ELK生态与Hadoop生态结合，由Elasticsearch集群存储短期的日志数据，主要负责日志数据的实时处理，由HBase集群主要负责离线的日志数据处理，保证了实时处理的性能与实时性，从而提高了日志处理效率。

附图说明

[0039] 为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本发明的实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据提供的附图获得其他的附图。

[0040] 图1是本发明第一实施方式提供的日志处理方法的流程图。

[0041] 图2是本发明第一实施方式提供的日志处理系统的结构示意图。

[0042] 图3是本发明一实施方式的终端的结构示意图。

[0043] 图4是图3所示的终端的示例性的功能模块图。

[0044] 主要元件符号说明

[0045]

终端	1
存储器	10
显示屏	20

[0046]

处理器	30
日志处理系统	100
日志采集模块	101
Kafka 日志分发集群	102
Elasticsearch 集群	103
HBase 集群	104
结果展示模块	105

[0047] 如下具体实施方式将结合上述附图进一步说明本发明实施例。

具体实施方式

[0048] 为了能够更清楚地理解本发明实施例的上述目的、特征和优点，下面结合附图和具体实施方式对本发明进行详细描述。需要说明的是，在不冲突的情况下，本申请的实施方式中的特征可以相互组合。

[0049] 在下面的描述中阐述了很多具体细节以便于充分理解本发明实施例，所描述的实施方式仅仅是本发明一部分实施方式，而不是全部的实施方式。基于本发明中的实施方式，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施方式，都属于本发明实施例保护的范围。

[0050] 除非另有定义，本文所使用的所有的技术和科学术语与属于本发明实施例的技术领域的技术人员通常理解的含义相同。本文中在本发明的说明书中所使用的术语只是为了描述具体的实施方式的目的，不是旨在于限制本发明实施例。

[0051] 图1是本发明第一实施方式的日志处理方法的流程图。所述日志处理方法可以应用于终端1，所述终端1可以是例如智能手机、笔记本电脑、台式/平板电脑、智能手表以及个人数字助理(Personal Digital Assistant，PDA)等智能设备。如图1所示，所述日志处理方法可以包括如下步骤：

[0052] S101：获取日志数据。

[0053] 在本实施方式中，通过日志采集模块从相关应用的运营平台中获取的日志数据，所述日志数据的类型可以包括用户行为数据、应用状态数据或设备状态数据，此处并不对日志数据的内容、来源进行限定。所述日志采集模块可以使用Filebeat进行日志数据采集(称为Filebeat日志采集模块)，所述Filebeat为日志数据采集器。所述Filebeat日志采集模块支持在日志处理系统100中定制各类日志数据的发送方，所述Filebeat日志采集模块用于获取日志数据，并将所述日志数据输出至各类日志数据的接收方。具体的，所述Filebeat日志采集模块启动一个或多个探测器(prospectors)去检测指定的日志目录或文件；对于所述探测器找出的每一个日志文件，所述Filebeat日志采集模块启动收割进程(harvester)；每一个所述收割进行读取一个日志文件的新内容，并发送所述日志文件的新

内容到处理程序(spooler),所述处理程序会集合这些日志数据,最后所述Filebeat日志采集模块会发送集合的日志数据到指定的地点。可以理解的是,在所述获取日志数据之后,所述方法还包括:根据预设结构对所述日志数据进行转化,具体的,所述日志数据的预设结构可以包括日志时间、日志级别、日志输出类以及日志内容等。

[0054] S102:对所述日志数据进行分流,得到第一分流日志数据和第二分流日志数据。

[0055] 在本实施方式中,通过日志采集模块进行日志数据的采集,并将所述日志数据推送给Kafka日志分发集群作为所述日志数据的缓存层。其中,所述Kafka日志分发集群是一种分布式消息缓存中间件,具有高吞吐量的特点(即使是使用非常普通的硬件,Kafka也可以支持每秒数十万的消息),用于海量数据的缓存,通过消息队列的方式,对数据进行分发和控制。所述Kafka日志分发集群可以将接收到的日志数据转换为Kafka消息队列。所述Kafka日志分发集群可以对所述Kafka消息队列中缓存的日志数据进行分流处理,所述Elasticsearch集群与所述HBase集群为所述Kafka日志分发集群的消费者。也即,所述Kafka日志分发集群可以将一份日志数据输出给所述Elasticsearch集群,一份日志数据输出给所述HBase集群。

[0056] 所述对所述日志数据进行分流,得到第一分流日志数据和第二分流日志数据包括:通过所述Kafka日志分发集群对所述日志数据进行分流,将所述日志数据分为实时日志数据与非实时日志数据,其中,所述第一分流日志数据即为所述实时日志数据,所述第二分流数据即为所述非实时日志数据。对于所述第一分流日志数据,将其输出给所述Elasticsearch集群;对于所述第二分流日志数据,将其输出给所述HBase集群。所述通过所述Kafka日志分发集群对所述日志数据进行分流包括采用Strom流式计算框架得对所述Kafka消息队列中缓存的日志数据进行分析处理,得到实时日志数据以及非实时日志数据。在其他实施方式中,还可以通过Zookeeper(ZooKeeper是一个分布式的,开放源码的分布式应用程序协调服务)日志分发集群对所述日志数据进行分类,得到第一分流日志数据和第二分流日志数据。

[0057] S103:将所述第一分流日志数据输入至所述Elasticsearch集群中进行第一处理后得到第一结果。

[0058] 在本实施方式中,所述第一分流日志数据为实时日志数据,在将所述第一分流日志数据输入至所述Elasticsearch集群中进行第一处理后得到第一结果之前,所述方法还包括:接收所述Kafka消息队列中缓存的不同topic里的实时日志数据;通过Logstash日志解析模块按照预设解析规则对所述实时日志数据进行解析操作。所述通过Logstash日志解析模块按照预设解析规则对所述实时日志数据进行解析包括通过Logstash日志解析模块对所述第一分流日志数据进行清洗和加工,并将所述第一分流日志数据结构化成不同的字段。通过Logstash日志解析模块对日志文件进行解析,能够识别出待处理的所述第一分流日志数据中的有用信息,过滤掉垃圾数据。所述Logstash日志解析模块中配置有所有日志来源的解析文件,所述预设解析规则为所述解析文件中设置的规则。

[0059] 将经过所述Logstash日志解析模块解析处理后的所述第一分流日志数据输出给所述Elasticsearch集群。所述将所述第一分流日志数据输入至所述Elasticsearch集群中进行第一处理后得到第一结果包括:通过所述Elasticsearch集群保存所述经解析处理后的所述第一分流日志数据;对所述第一分流日志数据进行实时日志数据处理,得到实时日

志数据处理结果，其中，所述实时日志数据处理包括以下中的一种或多种的组合：实时检索处理、实时告警处理与在线统计处理。所述Elasticsearch集群保存日志数据采取的方式为分布式存储方式，所述第一分流日志数据通过倒序索引的方式将关键文字与日志数据进行映射。其中，所述关键文字包括时间、字段、关键字等。将一个索引进行分片，不同的分片存在不同的集群节点上，能够备份日志数据防止节点损坏导致文件丢失，能够将日志数据信息进行展现，而且可以通过输入所述关键文字(例如，时间、字段、关键字)的方式快速搜索到需要的信息。

[0060] S104：将所述第二分流日志数据输入至所述HBase集群中进行第二处理后得到第二结果。

[0061] 在本实施方式中，所述第二分流日志数据为非实时日志数据，在将所述第二分流日志数据输入至所述HBase集群中进行第二处理后得到第二结果之前，所述方法还包括：读取预定解析规则；通过Spark集群按照预定解析规则对所述第二分流日志数据进行解析操作，将所述第二分流日志数据解析为HBase数据表格式，将解析后的HBase数据表格式存储至所述HBase集群中。其中，所述预定解析规则可以是系统开发人员预先设置的，所述预定解析规则可以包括正则表达式、KeyValue解析、字段值拆分(例如，利用split函数进行拆分)、String类型转换成数值型、Json解析、URL解码、时间戳识别以及UserAgent解析中的一种或多种。

[0062] 将经过解析处理后的所述第二分流日志数据输出给所述HBase集群。所述将所述第二分流日志数据输入至所述HBase集群中进行第二处理后得到第二结果包括：通过所述HBase集群保存所述经解析处理后的所述第二分流日志数据；对所述第二分流日志数据进行离线日志数据处理，得到离线日志数据处理结果，其中，所述离线日志数据处理包括以下中的一种或多种：离线分析处理、日志备份处理与日志还原处理。

[0063] S105：展示所述第一结果及/或所述第二结果。

[0064] 在本实施方式中，通过结果展示模块展示所述第一结果及/或所述第二结果，所述结果展示模块保存至Web客户端中。所述展示所述第一结果及/或所述第二结果包括：获取所述日志处理系统100当前正在处理的日志数据信息；当所述日志处理系统100当前正在处理的日志数据信息为第一分流日志数据时，展示所述第一结果；当所述日志处理系统100当前正在处理的日志数据信息为第二分流日志数据时，展示所述第二结果。

[0065] 本发明实施例还提供Mysql数据库、Mongo数据库与Web应用程序。所述Web应用程序与所述Mysql数据库及Mongo数据库连接。其中，所述Mysql数据库是一种开放源代码的关系型数据库管理系统，所述Mysql数据库中主要存放资源配置相关数据。所述Mongo数据库是一个基于分布式文件存储的数据库，旨在为WEB应用提供可扩展的高性能数据存储解决方案，所述Mongo数据库中主要存放日志数据的统计分析结果。

[0066] 所述Web应用程序还与Web服务器相互连接，所述Web服务器用于接收Web客户端上传的用于与Web应用程序进行数据交互的交互数据，并将所述交互数据通过接口输出给Web应用程序，Web应用程序对交互数据进行处理后，得到处理结果，并将处理结果反馈给Web服务器，通过Web服务器将处理结果反馈至客户端，通过所述客户端中的结果展示模块将结果进行展示。

[0067] 本发明实施例提供一种日志处理方法，获取日志数据；对所述日志数据进行分流，

得到第一分流日志数据和第二分流日志数据;将所述第一分流日志数据输入至所述Elasticsearch集群中进行第一处理后得到第一结果;将所述第二分流日志数据输入至所述HBase集群中进行第二处理后得到第二结果;展示所述第一结果及/或所述第二结果。利用本发明实施例,其可以将ELK生态与Hadoop生态结合,由Elasticsearch集群存储短期的日志数据,主要负责日志数据的实时处理,由HBase集群主要负责离线的日志数据处理,在HBase集群上运行一些较耗时的离线分析任务时,可同时在Elasticsearch集群中进行日志的查询与告警等实时处理,从而提高了日志处理效率。

[0068] 图2是本发明第一实施方式提供的日志处理系统的结构示意图。如图2所示,所述日志处理系统100包括日志采集模块101、Kafka日志分发集群102、Elasticsearch集群103、HBase集群104以及结果展示模块105(其中,所述结果展示模块105图未示出,所述结果展示模块保存至Web客户端中,所述Web客户端在图中也未示出)。其中,所述日志采集模块101可以用于获取日志数据(Log);所述Kafka日志分发集群102可以用于对日志数据进行分流,得到第一分流日志数据和第二分流日志数据;所述第一分流日志数据为所述实时日志数据,所述第二分流数据为所述非实时日志数据。在将所述第一分流日志数据输出至所述Elasticsearch集群之前,还需要通过Logstash日志解析模块按照预设解析规则对所述第一分流日志数据进行解析(也即对所述第一分流日志数据进行清洗和加工,结构化成不同的字段)操作。在将所述第二分流日志数据输出至所述HBase集群之前,还需要通过Spark集群按照预定解析规则对所述第二分流日志数据进行解析操作。所述Elasticsearch集群可以用于对所述第一分流日志数据进行第一处理后得到第一结果;所述HBase集群可以用于对所述第二分流日志数据进行第二处理后得到第二结果;结果展示模块可以用于展示所述第一结果及/或所述第二结果。本发明实施例还提供Mysql数据库、Mongo数据库与Web应用程序。所述Web应用程序与所述Mysql数据库及Mongo数据库连接。其中,所述Mysql数据库中主要存放资源配置相关数据,所述Mongo数据库中主要存放日志数据的统计分析结果。所述Web应用程序还与Web服务器相互连接,所述Web服务器用于接收Web客户端上传的用于与Web应用程序进行数据交互的交互数据,并将所述交互数据通过接口输出给Web应用程序,Web应用程序对交互数据进行处理后,得到处理结果,并将处理结果反馈给Web服务器,通过Web服务器将处理结果反馈至客户端,通过所述客户端中的结果展示模块将结果进行展示。

[0069] 图3是本发明一实施方式的终端1的结构示意图,如图3所示,终端1包括存储器10,存储器10中存储有日志处理系统100。所述的终端1可以是手机、平板电脑、个人数字助理等具有应用显示功能的终端1。所述日志处理系统100可以获取日志数据;对所述日志数据进行分流,得到第一分流日志数据和第二分流日志数据;将所述第一分流日志数据输入至所述Elasticsearch集群中进行第一处理后得到第一结果;将所述第二分流日志数据输入至所述HBase集群中进行第二处理后得到第二结果;展示所述第一结果及/或所述第二结果。利用本发明实施例,其可以将ELK生态与Hadoop生态结合,由Elasticsearch集群存储短期的日志数据,主要负责日志数据的实时处理,由HBase集群主要负责离线的日志数据处理,从而提高了日志处理效率。

[0070] 本实施方式中,终端1还可以包括显示屏20及处理器30。存储器10、显示屏20可以分别与处理器30电连接。

[0071] 所述的存储器10可以是不同类型存储设备,用于存储各类数据。例如,可以是终端

1的存储器、内存，还可以是可外接于该终端装置1的存储卡，如闪存、SM卡(Smart Media Card，智能媒体卡)、SD卡(Secure Digital Card，安全数字卡)等。此外，存储器10可以包括高速随机存取存储器，还可以包括非易失性存储器，例如硬盘、内存、插接式硬盘，智能存储卡(Smart Media Card,SMC)，安全数字(Secure Digital,SD)卡，闪存卡(Flash Card)、至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。存储器10用于存储各类数据，例如，所述终端1中安装的各类应用程序(Applications)、应用上述日志处理方法而设置、获取的数据等信息。

[0072] 显示屏20安装于终端1，用于显示信息。

[0073] 处理器30用于执行所述日志处理方法以及所述终端1内安装的各类软件，例如操作系统及应用显示软件等。处理器30包含但不限于处理器(Central Processing Unit，CPU)、微控制单元(Micro Controller Unit,MCU)等用于解释计算机以及处理计算机软件中的数据的装置。

[0074] 所述的日志处理系统100可以包括一个或多个的模块，所述一个或多个模块被存储在终端1的存储器10中并被配置成由一个或多个处理器(本实施方式为一个处理器30)执行，以完成本发明实施例。例如，参阅图4所示，所述日志处理系统100可以包括日志采集模块101、Kafka日志分发集群102、Elasticsearch集群103、HBase集群104以及结果展示模块105。本发明实施例所称的模块可以是完成一特定功能的程序段，比程序更适合于描述软件在处理器中的执行过程。

[0075] 可以理解的是，对应上述日志处理方法中的各实施方式，终端1可以包括图4中所示的各功能模块中的一部分或全部，各模块的功能将在以下具体介绍。需要说明的是，以上日志处理方法的各实施方式中相同的名词相关名词及其具体的解释说明也可以适用于以下对各模块的功能介绍。为节省篇幅及避免重复起见，在此就不再赘述。

[0076] 日志采集模块101可以用于获取日志数据。

[0077] Kafka日志分发集群102可以用于对日志数据进行分流，得到第一分流日志数据和第二分流日志数据。

[0078] Elasticsearch集群103可以用于对所述第一分流日志数据进行第一处理后得到第一结果。

[0079] HBase集群104可以用于对所述第二分流日志数据进行第二处理后得到第二结果。

[0080] 结果展示模块105可以用于展示所述第一结果及/或所述第二结果。

[0081] 本发明实施例还提供一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时实现上述任一实施方式中的日志处理方法的步骤。

[0082] 所述日志处理系统100/终端1/计算机设备集成的模块/单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。基于这样的理解，本发明实现上述实施方式方法中的全部或部分流程，也可以通过计算机程序来指令相关的硬件来完成，所述的计算机程序可存储于一计算机可读存储介质中，该计算机程序在被处理器执行时，可实现上述各个方法实施例的步骤。其中，所述计算机程序包括计算机程序代码，所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读存储介质可以包括：能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器(ROM，

Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、电载波信号、电信信号以及软件分发介质等。

[0083] 所称处理器30可以是中央处理单元(Central Processing Unit,CPU),还可以是其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现成可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等,所述处理器30是所述日志处理系统100/终端1的控制中心,利用各种接口和线路连接整个日志处理系统100/终端1的各个部分。

[0084] 所述存储器10用于存储所述计算机程序和/或模块,所述处理器30通过运行或执行存储在所述存储器内的计算机程序和/或模块,以及调用存储在存储器10内的数据,实现所述日志处理系统100/终端1的各种功能。所述存储器10可主要包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需的应用程序(比如声音播放功能、图像播放功能等)等;存储数据区可存储根据手机的使用所创建的数据(比如音频数据、电话本等)等。

[0085] 在本发明所提供的几个具体实施方式中,应该理解到,所揭露的终端和方法,可以通过其它的方式实现。例如,以上所描述的系统实施方式仅仅是示意性的,例如,所述模块的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式。

[0086] 对于本领域技术人员而言,显然本发明实施例不限于上述示范性实施例的细节,而且在不背离本发明实施例的精神或基本特征的情况下,能够以其他的具体形式实现本发明实施例。因此,无论从哪一点来看,均应将实施例看作是示范性的,而且是非限制性的,本发明实施例的范围由所附权利要求而不是上述说明限定,因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本发明实施例内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。权利要求中陈述的多个单元、模块或装置也可以由同一个单元、模块或装置通过软件或者硬件来实现。

[0087] 以上实施方式仅用以说明本发明实施例的技术方案而非限制,尽管参照以上较佳实施方式对本发明实施例进行了详细说明,本领域的普通技术人员应当理解,可以对本发明实施例的技术方案进行修改或等同替换都不应脱离本发明实施例的技术方案的精神和范围。

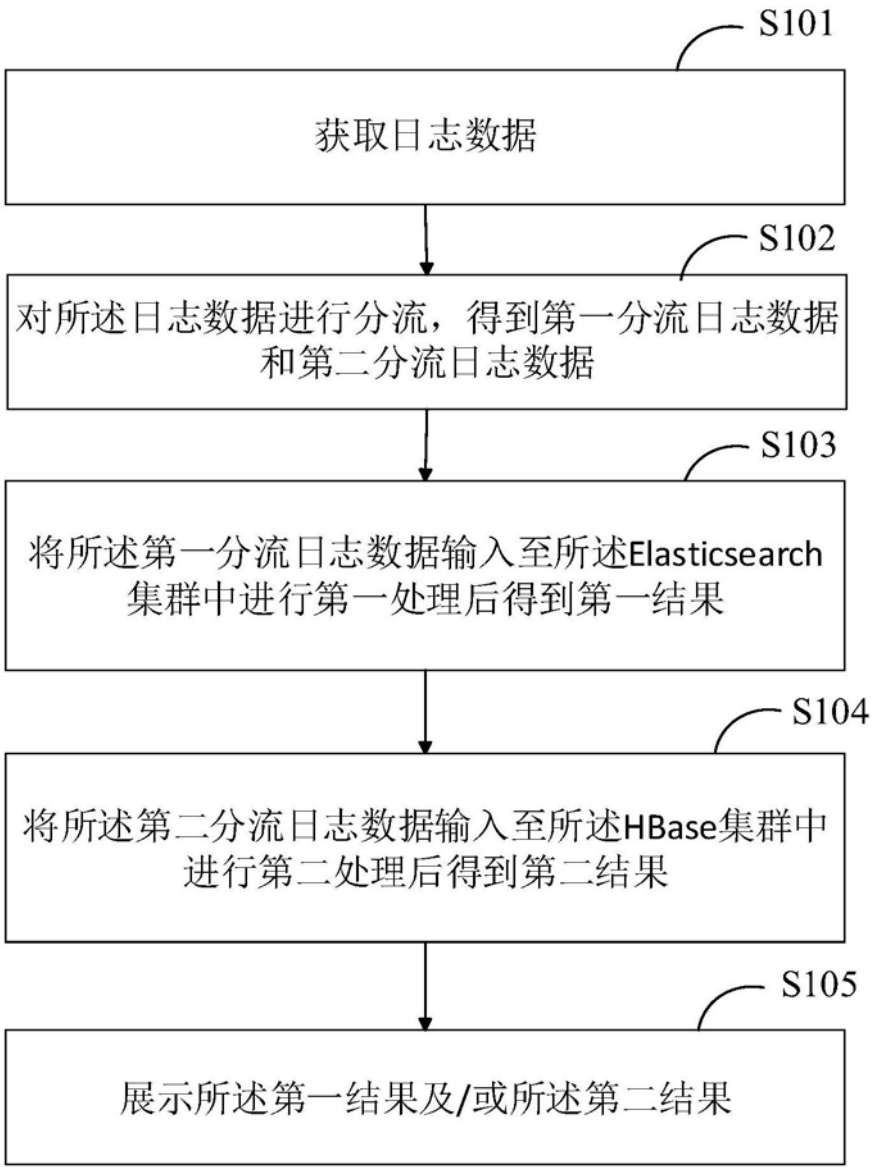

S101
获取日志数据
S102
对所述日志数据进行分流，得到第一分流日志数据和第二分流日志数据
S103
将所述第一分流日志数据输入至所述Elasticsearch集群中进行第一处理后得到第一结果
S104
将所述第二分流日志数据输入至所述HBase集群中进行第二处理后得到第二结果
S105
展示所述第一结果及/或所述第二结果

图1

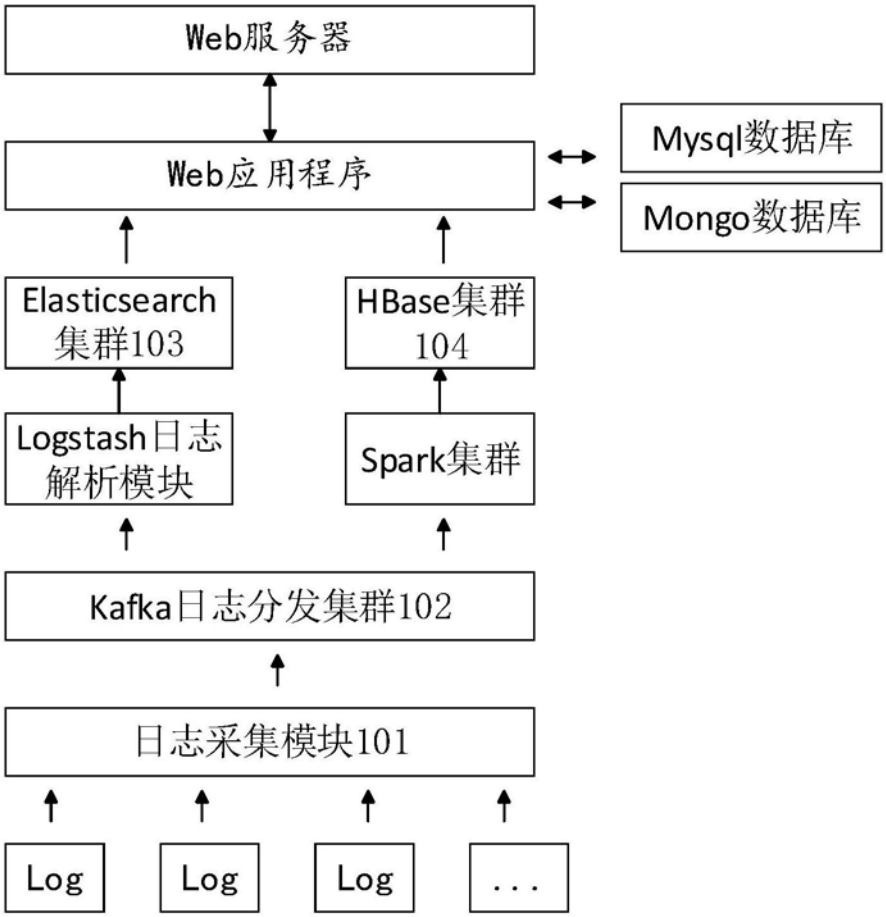

Web服务器
Web应用程序
Mysql数据库
Mongo数据库
Elasticsearch
集群103
HBase集群
104
Logstash日志
解析模块
Spark集群
Kafka日志分发集群102
日志采集模块101
Log
Log
Log
. . .

图2

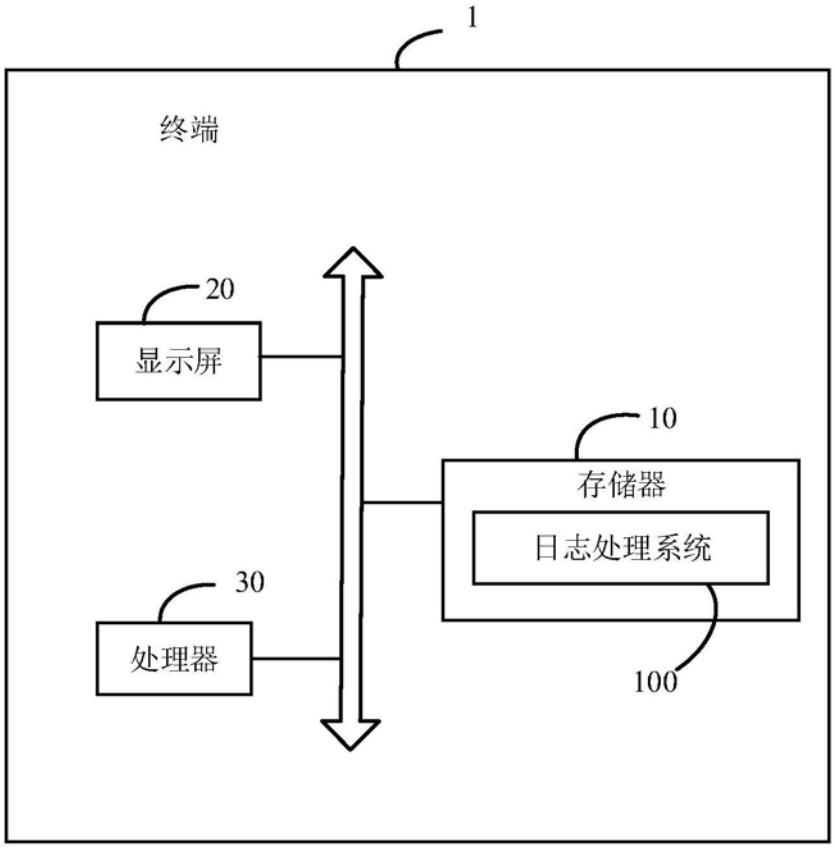


图3

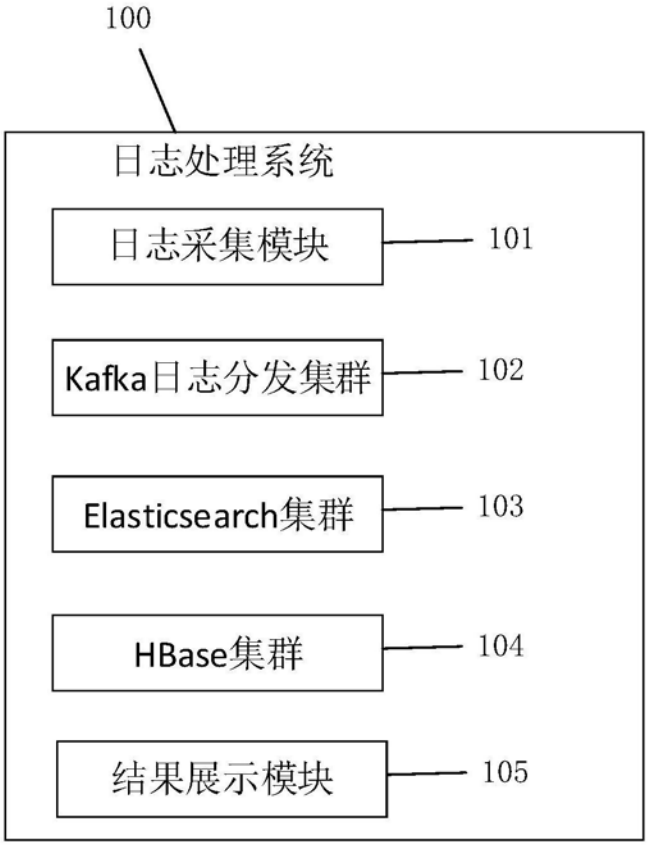

100
日志处理系统
日志采集模块
101
Kafka日志分发集群
102
Elasticsearch集群
103
HBase集群
104
结果展示模块
105

图4