

基于 Solr 的农田数据索引方法与大数据平台构建

苑严伟 冀福华 赵 博 姜含露 王 猛 樊学谦
(中国农业机械化科学研究院土壤植物机器系统技术国家重点实验室, 北京 100083)

摘要: 针对农田数据在高吞吐量、高并发、多条件处理过程中易产生运算负载大、响应速度慢等难题,研究了负载均衡大规模集群数据处理技术,优化了多条件检索时 Hbase 农田数据库,提出了基于 Solr 的二级非主键索引方法,搭建了基于 Hadoop 的农田大数据平台,采用农机深松、植保、保护性耕作等 8 种作业生成的 100 TB 数据对平台进行了检索实验和压力测试实验。实验结果表明,多条件检索时,优化后的技术模型在数据规模达到 5×10^7 条时,系统的响应时间小于 1 s,优化的性能与原生 Hbase 相比提高了 3 倍;在模拟用户达到 5×10^3 次时,系统的 QPS 及 TPS 提高了 1 倍左右、RT 提高了 2.5 倍,系统的平均响应时间为 183 ms。本研究解决了高吞吐量、高并发导致农田数据检索效率低的问题,提高了海量农田数据实时处理的计算能力。

关键词: 农田; 大数据; 索引; Solr; Hbase; Hadoop

中图分类号: S24; TP391 **文献标识码:** A **文章编号:** 1000-1298(2019)11-0186-07

Index Method of Farmland Data Based on Solr and Construction of Big Data Platform

YUAN Yanwei JI Fuhua ZHAO Bo JIANG Hanlu WANG Meng FAN Xueqian
(The State Key Laboratory of Soil, Plant and Machine System Technology,
Chinese Academy of Agricultural Mechanization Sciences, Beijing 100083, China)

Abstract: Aiming at the problems of high throughput, high concurrency and slow response of farmland data in the process of multi-condition processing, such as high computational load and slow response speed. The data processing technology of load balancing large-scale cluster was studied, the Hbase farmland database in multi-condition retrieval was optimized, a two-level non-primary key index method was put forward based on Solr, and a large farmland data platform was built based on Hadoop. The 100TB data was generated by eight operations, such as subsoiling, plant protection and conservation tillage, and those data was retrieved and tested on the platform based on Hadoop. The experimental results showed that the response time of the optimized technical model was less than 1 s when the concurrent volume of farm data was 50 million, and the performance of the optimized model was improved by about four times compared with the original Hbase. When the concurrent volume of simulated users were 500 000, the query per second (QPS) and transaction per second (TPS) of the system were increased by about one time, the response time (RT) of the system was increased by 2.5 times, and the average response time was 183 ms. To a certain extent, this system solved the problem of low efficiency of farmland data retrieval caused by high throughput and concurrency, and improved the computing ability of real-time processing of massive farmland data.

Key words: farmland; big data; index; Solr; Hbase; Hadoop

0 引言

农田是农业生产的载体^[1-4],我国农田呈现数量多、田块小、形态各异、信息多样化等特点,农田信

息的多元异构、海量等特性已成为现代农业信息化发展的难题,为提高我国农业数字化与信息化程度,有必要构建面向精准农业的农田大数据平台,以实时、动态、高效地管理农田信息。

随着物联网、云计算、大数据等信息技术的发展,农田数字化已成为现代农业发展的主要方向之一。国外关于农田数据的研究较早,到目前为止,国外已有成熟的农田信息管理系统^[5-7],如:美国凯斯公司的 AFS 系统、约翰迪尔公司的 GreenStar 系统以及 AgLeader 的 PF(Precision Farming),这些都是相对规范且成熟的农田数据模型。另外,国外研发了与农田数据相关的智能农业决策系统,如 Farmeron 公司^[8]在 2011 年推出了基于 Web 端的农场管理平台,提供相对可靠的生产分析报告,以指导农民进行生产规划;VitalFields 公司^[9]以气象预测、病虫害、成本投入为研究对象,提供农作物种植阶段的相关服务,使得管理农场更加高效;2017 年 OneSoil 公司^[10]通过欧洲空间局的卫星影像和 Mapbox 的数据,利用人工智能的算法,推出了包括欧洲和美国超过 5 700 万块农田的交互式数字地图,用于自动检测地块、识别作物类型和监测作物发育。国外的地块大多为标准农场,而且在作物生产管理的方法和习惯方面与国内存在较大差异,国外的农田田块数据模型不适合我国复杂农情与多元经营模式。因此,国内仍需开发符合我国国情的农田数据模型与大数据平台。

国内关于农田数据的理论研究起步较晚。近年来,在国家十三五规划的带动下,与农田数据相关的技术发展较快,主要表现为多平台、多用途、多方法面向个性化的农田信息存储技术与实现。文献[11]利用 1957—2007 年河南省 30 个县、市的气象、土壤、作物等数据信息,构建了基于元数据的农田信息的存储、管理和共享模型;文献[12]以土地所有者、种植管理者、施肥管理、杀虫喷药等信息作为属性信息,构建了基于图的农田数据模型;文献[13]以土壤信息、环境信息、田间信息、GPS 定位信息等作为属性信息,构建了基于 GDAL 的农田信息数据模型;文献[14]以河南省巩义市吴沟村为例,以卫星遥感图像为建库数字化参照依据,采用目视解译方法,利用 ERDAS Imagine 8.6 和 MapInfo Professional 7.0 SCP 为数据库平台,依次提取旱地梯田、旱坡地、菜地、退耕还林地等不同田块图层,设计了基于 RS、GIS 的村域农田数据库。文献[15]以无线传感器采集平台的空气温湿度、风速、风向、光照为依据,建立了基于无线传感网的农田数据管理系统。国内研究大都基于传统的数据模型进行数据的储存和查询^[16]。而随着数据量的增大,大数据的海量、非结构化的特性导致了传统数据库难以满足大数据的存储与高并发的访问需求,海量异构农田数据多样性、庞杂导致信息无法有效重构,农田田块

大数据处理过程中易产生云端运算负担大、响应时间慢等问题。另外,传统的数据单机处理模式很难提供相对高效的计算资源整合,导致大量的硬件资源浪费。

针对上述问题,本文在前期农机深松监测系统的基础上,对农田数据在解析、存储与处理过程中的技术细节进行研究,建立基于 Hadoop 的农田大数据平台,对农田田块在作物的耕、种、管、收等过程中产生的海量农田数据进行动态管理。

1 大数据平台的设计与构建

1.1 系统总体架构

基于 Hadoop 的农田大数据平台可用于分布在全国各地的农业合作社农田田块信息管理。该平台将大数据、物联网、WebGIS 等技术进行深度融合,向用户提供农田田块数据的快速检索、处理和动态展示。其总体架构如图 1 所示,主要包括感知层、网络层、中间层和应用层。

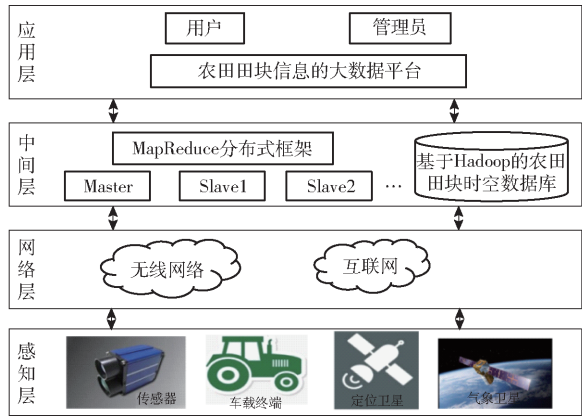


图 1 系统总体架构
Fig. 1 Overall architecture of system

1.2 系统功能设计

基于 Hadoop 的农田大数据平台的功能设计如图 2 所示,主要包括角色管理、地块管理、设备管理、农资管理、作业管理、统计管理,其主要面向对象为各种用户。

(1) 角色管理

平台的用户主要分为管理员和普通用户两种类型,不同类型的用户会分配不同的使用权限,如政府机关具备所管辖合作社的浏览权限、任务分配权限等;合作社具备本合作社业务的日常维护权限。另外为了便于系统的维护,设置了管理员角色作为最高权限。

(2) 地块管理

地块管理是整个大数据平台最重要的环节,它负责管理地块的位置信息、作业信息、气象信息、附着物信息,为地块的精细化管理提供了丰富的数据

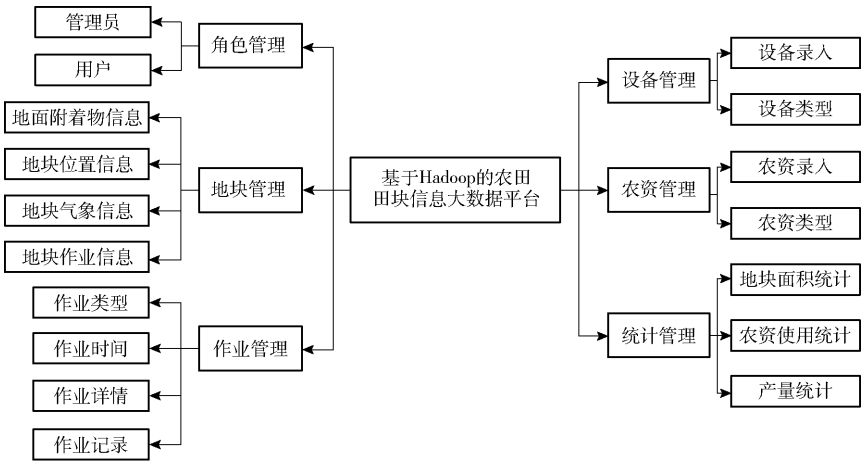


图 2 系统功能设计
Fig. 2 Function module of system

支持,具备地块信息的录入与检索,并通过 WebGIS 技术,将部分数据可视化。

(3) 设备管理

设备管理主要有两个作用:获取田块信息和进行田块作业。设备管理的内容主要为设备的基本信息、工作详情。其中包括录入设备基本信息和设备工作详情,设备工作详情主要指用户在客户端可实时查询设备的位置及机手信息,其中设备位置信息可调用地地图 API 实时获取。

(4) 作业管理

作业管理主要包括作业类型、作业详情、作业记录、作业时间 4 部分。其中作业类型为农机作业的基本类型,如:深松、打捆、籽粒直收、植保等;作业详情为实时展示作业现场并抓拍作业场景;作业完成情况是指作业轨迹、作业的完成率、已作业面积;作业记录是指作业的历史信息,这些信息包括作业轨迹、作业面积、机手信息、作业时间、作业时随机抓拍的图像。

(5) 农资管理

农资管理是对地块的使用资源(化肥、农药等)的统一整合,可对地块的农资使用情况提供较为详尽的记录。另外,可记录地块的产出数据,这些数据都能为农作物科学管理及分析决策提供数据源支持。

(6) 统计管理

统计管理是在农资管理、设备管理、地块管理的基础上,对农田数据进行深层次的加工,比如:统计造成不同地块产量差异性的原因,探究农机作业、农资使用与农作物产量之间关系。

2 农田数据关键技术设计

2.1 数据结构

本文以农田田块为研究对象进行农田数据的分

析,农田田块数据包括田块的位置信息、气象信息、作业信息及附着物信息等。位置信息包含田块经度、纬度、海拔及坡度等地理信息,气象信息包括历年的温度、湿度、光照、风速、风向等;作业信息主要是不同农业机械的历年与实时作业信息,如深松机作业深度、播种机播种施肥量、植保机喷药量、收获机产量分布信息等;附着物信息是指对农业生产影响较大的信息,如水井、电线杆、农舍的经纬度和局部轮廓等。数据结构如图 3 所示。由于数据量大,为了便于检索和存储,现有的农田田块大数据平台将数据进行分表保存处理。

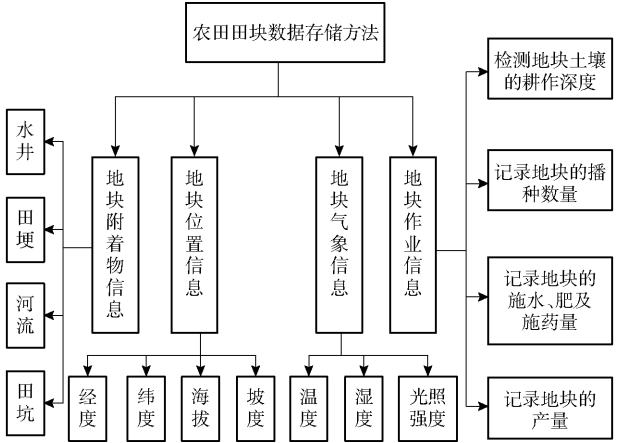


图 3 农田数据结构
Fig. 3 Data structure of farmland

2.2 数据库选取与部署

农田田块数据类型复杂,不仅有常见的结构化数据,而且还有半结构化数据。HBase^[17]作为高性能、列存储、可伸缩、实时读写的分布式数据库,可支持集群存储海量数据。目前由于对农田数据的存储多选用传统的数据库,如 Mysql、PostgreSql 等,这些数据库在解决事务性问题上有其独特的优势,然而随着数据量的增加,数据存储的因素主要从数据检索性能来考量。

数据库 Hbase 的检索过程如下：

(1)客户端从 ZooKeeper 找到存放在 Master 上的 Region 服务器的位置信息、Region 的元数据信息以及 Region 与 Region 服务器的映射关系。

(2)通过 Region 服务器检索存储在 Region 的数据。

(3)通过 HDFS(Hadoop distributed files system)读取存储在磁盘上的数据。

Hbase^[18]的部署服从主从原则,依附于 Hadoop 集群。本文部署的 Hbase 的所有协调工作均由 ZooKeeper 来执行,来实现 HMater 的选举与主备的切换、系统容错、任务管理。而且 Hbase 与 ZooKeeper 为独立部署,这样可避免二者发生强耦合,便于后期集群的升级与维护,由于 ZooKeeper 具备轻量化的特性,部署的 ZooKeeper 仍构建于 HDFS 之上,此外,它还提供了整个大数据集群的负载均衡。

2.3 行键

Hbase 表主要由行键(Rowkey)、列族、列限定符和时间戳等组成。行键是其主键,在满足行键设计长度原则、散列原则、唯一原则的前提下,为提高数据对内存的使用率,对于 64 位操作系统的行键设计应为 8 字节的倍数,本设计中选择 16 字节行键设计。将行键的高位作为散列字段,低位作为时间值,可以使得数据相对均衡地分布在 Region 服务器中。由于 Hbase 的行键按字典顺序增加,为了避免写热点,设计 Rowkey 需使不同行的数据放在同一个 Region 中,如果没有散列字段,直接存放时间值将可能产生所有新数据都在一个 Region 服务器上堆积的热点现象,导致数据检索时负载将会集中在个别 Region 服务器,降低查询效率。在农田大数据平台中,由于地块编号、设备号、作业人及作业时间等字段使用较为频繁,所以在设计行键时包含了上述信息。同时,由于地块编号采用字母+数字的格式,将地块编号写入开端可以避免写入数据时行键的顺序递增的问题。因此,本系统行键设计为:Rowkey=“地块编号+归属人+作业人+时间”。

2.4 列族

农田田块数据主要字段包括:地块编号、地块位置、地块面积、地块归属人、地块管理人、作业类型、设备号、车牌号、作业时间、温度、湿度、光照强度、风速等 128 个字段。对于作业时间要求精确到分钟,采用 8 位占位符进行存储。对于位置信息采用独立的几何属性列族存储;其他字符均采用字符型。各字段长度不一,由于列族少时,底层的 I/O 开销就会

降低,因此所有非几何字段构成一个唯一的列族,其具体设计如表 1 所示。

表 1 农田数据表结构
Tab.1 Structure of land data

Rowkey	列族 1	列族 2	时间戳
Rowkey	landNum	point	t1
	landArea	line	t2
	landPerson	polygon	t3
	machineId		t4
	temp		t5
	⋮	⋮	⋮

2.5 基于 Solr 二级非主键索引的设计

Solr^[19]是 Apache 基金会下的一款基于 Lucene 的全文搜索服务器。用户可通过 Http 协议获得该系统对外提供的类似于 java-web 的 API,并向该系统提交生成索引的 XML 文件,最后可通过 Get 请求获取索引结果。另外,Solr 的搜索引擎采用 Java5 开发,而本文的大数据平台也采用了 Java 开发,可以使得大数据平台很好地兼容 Solr。因此,本文选取了 Solr 进行系统二级索引。

根据表 1 添加了 4 个索引字段,分别为:地块类型(landType)、地块面积(landArea)、地块产量(landOutput)和地块作业类型(workType)。其配置信息中 name 表示字段名;type 表示定义的字段的属性,一般情况下 integer 字段的索引效果更好;indexed 表示是否索引,作为索引字段必需设置为 true;stored 表示是否被存储,为了改进性能,对于不需要作为结果存储的均设为 false;multiValued 表示是否包含多个值,对于可能存在多值的字段全部设为 true,从而避免构建索引时报错,索引字段的类型与配置信息具体如下

```
<fields>
  <field name = " Rowkey" type = " string"
indexed = " true" stored = " true" multiValued =
" false" / >
<field name = " landType" type = " integer" indexed
= " true" stored = " false" multiValued = " false" / >
<field name = " landArea" type = " string" indexed
= " true" stored = " false" multiValued = " false" / >
<field name = " landOutput" type = " integer"
indexed = " true" stored = " false" multiValued =
" false" / >
<field name = " workType" type = " string" indexed
= " true" stored = " false" multiValued = " false" / >
</fields>
```

构建二级非主键索引后,数据传输过程为:采集

到的农田数据分别存储在采集终端和大数据集群服务器上。采集终端可以存储近期内一部分数据,主要是为防止采集终端野外作业时由于信号丢失等引起短时间内数据无法正常传输到服务器端的问题。采集终端采集到的数据(生产者)会通过 kafka 将数据打包发送到数据中心(消费者),数据中心解析之后存储在构建于 HDFS 之上的 Hbase 数据库中。系统的元数据信息存储在 Master 结点的内存中,元数据信息主要包括维护文件系统的 FsImage 和记录操作日志的 EditLog。当客户端发出数据请求时,数据处理相关模块首先判断是否存在主键索引,若存在则直接通过 Hbase 的主键索引检索数据,若检索条件涉及二级非主键索引字段,则通过多条件查询到相关主键值,最后通过主键在 Hbase 里进行检索。该索引模块设计在一定程度上可提高检索速度。

3 实验结果分析

3.1 集群部署

系统部署在阿里云服务器,软件环境为:Hadoop 2.7.1、Hbase 2.0.X、JDK 1.8、ZooKeeper 3.4.6、操作系统为 CentOS7.0;硬件配置信息:CPU 型号为 Intel 酷睿 i7 8700K, 3.7 GHz 主频,32 GB 内存,20 TB 硬盘,千兆以太网,各个节点配置信息如表 2 所示。

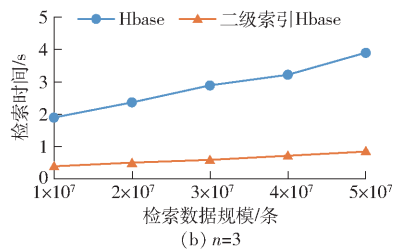
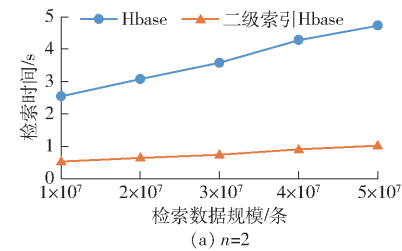


表 2 服务器各节点信息

Tab.2 Server nodes information

IP 地址	主机名	角色名
172.26.199.81	Master	NameNode/JobTracker
172.26.199.82	Slave1	SecondaryNameNode
172.26.199.83	Slave2	DataNode/TaskTracker
172.26.199.84	Slave3	DataNode/TaskTracker
172.26.199.85	Slave4	DataNode/TaskTracker
172.26.199.86	Slave5	DataNode/TaskTracker
172.26.199.87	Slave6	DataNode/TaskTracker
172.26.199.88	Slave7	DataNode/TaskTracker
172.26.199.89	Slave8	DataNode/TaskTracker
172.26.199.90	Slave9	DataNode/TaskTracker
172.26.199.91	Slave10	DataNode/TaskTracker

3.2 数据来源

实验数据选取了某省农机深松、植保、保护性耕作等 8 种作业类型产生的 100 TB 作业数据,该数据集记录了某省 2015—2018 年的作业情况,每条记录不等长,数据信息的类型包括文字、图像、视频等,数据存储结构如图 3 所示。本文用于对比实验的数据是通过 Sqoop^[20] 工具导入,硬件环境与上述 Hadoop 集群保持一致。

3.3 实验结果与分析

3.3.1 多条件检索、性能分析及其优化

检索数据仍采用上述数据,检索条件依次增加,本文设定的检索条件的个数(n)分别为 2、3、4;检索次数为 10 次,检索时间取其平均值。其检索性能对比如图 4 所示。

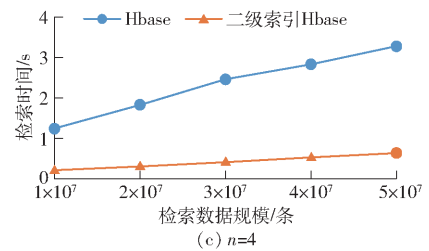


图 4 n 为 2、3、4 多条件检索性能对比

Fig.4 Contrast graphs of retrieval performance

由图 4 可知,随着数据量的增加,Hbase 的响应速度降低较为明显,当数据规模达到千万级时,响应时间大于 3 s。从理论层面上分析 Hbase 是一种面向列的数据库,其底层在行键上建立了基于行键的类 B+ 树索引,可以相对高效地支持基于行键的数据检索;对于多条件查询的非行键数据的检索,系统需要调用全表扫描的功能,数据的整体检索效率低下。随着数据量的不断增加,优化后的索引方法的优势越来越明显,当数据量达到 5×10^7 条时,响应时间小于 1 s,优化的性能与原生 Hbase 相比提高了 3 倍,能满足用户的需求。

为了更好地说明问题,本文对检索数据做了纵向对比,选取检索规模(C)为 3×10^7 条时,检索条件

个数 n 取 2、3、4, Hbase 检索时间分别为 3.576、2.887、2.458 s,二级索引 Hbase 检索时间分别为 0.742、0.582、0.421 s。

可知,二级索引的性能均优于原生 Hbase,而检索时间随着检索条件的增多呈下降的趋势。从理论上分析,由于检索条件的增多,从数据库中检索出符合条件的数据在减少,相当于减少了数据的并发量,因此,检索时间持续下降。

3.3.2 系统压力测试

为了验证系统架构的合理性与安全性,本文选取了请求次数较多的大数据平台首页进行了压力测试,首页界面如图 5 所示。选用 YCSB^[21] 作为系统的压力测试工具,选取每秒查询率(QPS)、吞吐量

(TPS)和响应时间(RT)等作为系统性能量的主要参数。本文通过多线程并发的方式模拟不同角色的用户,表 3 为 5×10^5 次并发用户压力测试结果,测试时间为 5 min。实验结果取 5 次实验的平均值。

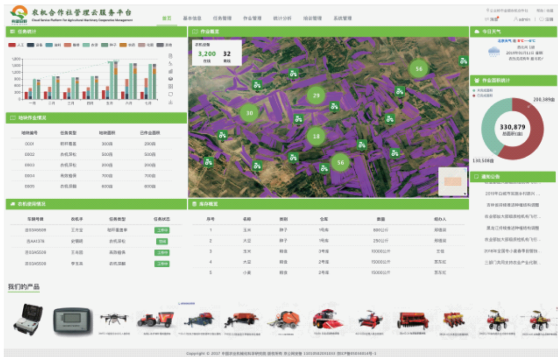


图 5 大数据平台首页

Fig. 5 Home page of big data platform

表 3 并发用户压力测试结果(5×10^5 次)

Tab. 3 Pressure test results(5×10^5)

参数	Hbase	二级索引 Hbase
每秒查询率(QPS)	233 821	424 211
每秒吞吐量(TPS)	4 658	8 657
响应时间(RT)/ms	643	183

通过测试结果,可以得出在高并发量(5×10^5 次)时,优化后系统的 RT 提高了 2.5 倍。系统平均响应时间 183 ms,系统的 QPS、TPS 提高了 1 倍左右,说明该系统在高并发时可高效稳定运行。

此外,对系统的读写能力进行了压力测试,表 4 描述了不同读写操作情况下(分别插入 1×10^5 、 1×10^6 、 1×10^7 条数据)系统插入数据的结果,每组实验进行 5 次,结果取 5 次插入的平均值。

由表 4 可知,在二级索引情况下,插入数据需要消耗的时间更多一些,这是由于二级索引需要触发

表 4 插入响应时间测试结果

Tab. 4 Results of testing for insert performance ms

检索类型	插入条数		
	1×10^5	1×10^6	1×10^7
Hbase	5 201. 8	13 826. 3	103 212. 1
二级索引 Hbase	5 829. 5	15 382. 6	121 131. 5

协处理器将索引数据不断存入 Solr 中,随着数据量增多,与原生 Hbase 相比二级索引插入性能不断降低。从实验结果来看,3 组实验的插入性能平均降低了 13.3%。降低的插入性能与提高的检索性能相比,可忽略。

4 结论

(1)研究了负载均衡大规模集群数据处理技术,实现了农田田间耕、种、管、收及农田田块自身产生的多源异构海量数据的存储与检索,保证了海量农田田块数据存储的安全性及检索的高效性。

(2)针对 Hbase 数据库在农田田间数据进行多条件检索时性能不佳的问题,研究了面向列数据库高效索引技术,提出了基于 Solr 二级非主键索引的检索方法,在数据规模达到 5×10^7 条时,响应时间小于 1 s,农田田块数据多条件关联检索时,相对原生 Hbase 数据库优化性能提高了 3 倍,大大提高了数据库的检索性能。

(3)研发了基于 Hadoop 的农田大数据平台,并对该平台进行了压力测试,测试结果表明,该平台在高并发量(5×10^5 次)时,优化后系统的 RT 提高了 2.5 倍,平均响应时间为 183 ms、系统的 QPS 及 TPS 提高了 1 倍左右,可保持系统的安全与稳定。目前该平台已具备农田田间管理功能,能对农田田块的大数据进行高效整合与处理,对指导大规模农田的安全生产与分析决策具有重大意义。

参 考 文 献

[1] 杨建宇,赵龙,徐凡,等. 基于耕地连片度的高标准基本农田建设划区[J/OL]. 农业机械学报,2017,48(4):142-148. YANG Jianyu, ZHAO Long, XU Fan, et al. High-standard basic farmland construction zoning based on the continuous degree of cultivated land [J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2017,48(4):142-148. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20170418&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2017.04.018. (in Chinese)

[2] 赵冬玲,何珊珊,林尚纬,等. 基于 TOPSIS 热点分析的高标准农田建设优先区选择[J/OL]. 农业机械学报,2017,48(7):153-158. ZHAO Dongling, HE Shanshan, LIN Shangwei, et al. Selection of high-standard priority areas for farmland construction based on TOPSIS and hot spot analysis [J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2017,48(7):153-158. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20170719&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2017.07.019. (in Chinese)

[3] 李道亮,杨昊. 农业物联网技术研究进展与发展趋势分析[J/OL]. 农业机械学报,2018,49(1):1-20. LI Daoliang, YANG Hao. Research progress and development trend analysis of agricultural Internet of Things [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(1): 1-20. <http://www.j-csam.org/jcsam/ch/>

- reader/view_abstract.aspx? flag = 1&file_no = 20180101&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2018.01.001. (in Chinese)
- [4] 陈杰,陈铁桥,梅小明,等. 基于最优尺度选择的高分辨率遥感影像丘陵农田提取[J]. 农业工程学报,2014,30(5):99-107.
CHEN Jie, CHEN Tieqiao, MEI Xiaoming, et al. Extraction of hilly farmland from high resolution remote sensing images based on optimal scale selection[J]. Transactions of the CSAE,2014,30(5):99-107. (in Chinese)
- [5] GEMTOS T, FOUNTAS S, BLACKMORE B S, 等. 欧洲精细农业的发展[J]. 世界农业,2006(10):21-24.
GEMTOS T, FOUNTAS S, BLACKMORE B S, et al. Development of precision agriculture in Europe [J]. World Agriculture, 2006(10):21-24. (in Chinese)
- [6] FORSSMAN N, ROOT-BERNSTEIN M. Landscapes of anticipation of the other: ethno-ethology in a deer hunting landscape [J]. Journal of Ethnobiology, 2018, 38(1):71-87.
- [7] FERRÁNDEZ-PASTOR F, GARCÍA-CHAMIZO J, MARIO N H, et al. Precision agriculture design method using a distributed computing architecture on internet of things context[J]. Sensors, 2018, 18(6):1731.
- [8] WEST J. A prediction model framework for cyber-Attacks to precision agriculture technologies[J]. Journal of Agricultural & Food Information, 2018,4:1-24.
- [9] FOUGHALI K, FATHALLAH K, FRIHIDA A. Using cloud IOT for disease prevention in precision agriculture[J]. Procedia Computer Science, 2018, 130:575-582.
- [10] MOROTA G, VENTURA R V, SILVA F F, et al. Machine learning and data mining advance predictive big data analysis in precision animal agriculture[J]. Journal of Animal Science, 2018,128:231-241.
- [11] 马新民,许鑫,席磊,等. 基于元数据的农田信息存储、管理和共享[J]. 农业工程学报,2010,26(11):209-214.
MA Xinmin, XU Xin, XI Lei, et al. Storage, management and sharing of farmland information based on metadata [J]. Transactions of the CSAE, 2010,26(11):209-214. (in Chinese)
- [12] 杨博宁,杨林楠. 基于图的时空数据模型在田块信息上的应用 [J]. 农机化研究, 2008,30(5):209-214.
YANG Boning, YANG Linnan. Application of graph-based spatio-temporal data model in field information [J]. Journal of Agricultural Mechanization Research, 2008,30(5):209-214. (in Chinese)
- [13] 赵辉辉. 基于 GDAL 的农田信息系统研究[D]. 哈尔滨:东北农业大学,2011.
ZHAO Huihui. Research on farmland information system based on GDAL [D]. Harbin: Northeast Agricultural University, 2011. (in Chinese)
- [14] 乔家君,毛磊. 基于 RS、GIS 村域农田数据库设计研究——以河南省吴沟村为例[J]. 农业系统科学与综合研究,2009, 25(3):312-316,321.
QIAO Jiajun, MAO Lei. Research on the design of village farmland database based on RS and GIS: a case study of Wugou village in Henan Province[J]. Agricultural System Science and Comprehensive Research, 2009, 25(3): 312-316,321. (in Chinese)
- [15] 邹金秋,周清波,杨鹏,等. 无线传感网获取的农田数据管理系统集成与实例分析[J]. 农业工程学报,2012,28(2):142-147.
ZOU Jinqiu, ZHOU Qingbo, YANG Peng, et al. Farmland data management system based on wireless sensor network [J]. Transactions of the CSAE, 2012,28(2):142-147. (in Chinese)
- [16] 孟超英,张雪彬,陈红茜,等. 基于 Hadoop 的蛋鸡设施养殖智能监测管理系统研究[J/OL]. 农业机械学报,2018,49(9): 166-175.
MENG Chaoying, ZHANG Xuebin, CHEN Hongxi, et al. Research on intelligent monitoring and management system of egg facility breeding based on Hadoop [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018,49(9): 166-175. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1&file_no = 20180920&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2018.09.020. (in Chinese)
- [17] JING Weipeng, TIAN Dongxue. An improved distributed storage and query for remote sensing data[J]. Procedia Computer Science,2018,129:238-247.
- [18] WANG Yong, LI Chengjun, LI Meng, et al. HBase storage schemas for massive spatial vector data[J]. Cluster Computing, 2017,20(4):3657-3666.
- [19] WANG Hongman, WANG Hewei, LIU Yuzhang, et al. Design and implementation of SOLR-based information retrieval system for value-added service[J]. The Journal of China Universities of Posts and Telecommunications,2008,19(4):592-602.
- [20] 于金良,朱志祥,梁小江. 一种基于 Sqoop 的数据交换系统[J]. 物联网技术,2016,6(3):35-37.
YU Jinliang, ZHU Zhixiang, LIANG Xiaojiang. A data exchange system based on Sqoop[J]. Internet of Things Technology, 2016,6(3):35-37. (in Chinese)
- [21] SHA H M, JIANG W, DONG H, et al. Towards the design of efficient and consistent index structure with minimal write activities for Non-volatile memory[J]. IEEE Transactions on Computers, 2018, 67(3):432-448.