

智能变电站海量在线监测数据处理方法

王德文,肖磊,肖凯

(华北电力大学 控制与计算机工程学院,河北 保定 071003)

摘要: 通过对智能变电站状态监测数据的特点以及 Hadoop 云计算技术的分析,研究数据检测机制以及索引构建方法,提出一种基于 Hadoop 的智能变电站在线监测数据可靠存储及快速查询方法。该方法将变电站在线监测所收集的海量变电设备在线监测数据冗余存储在 Hadoop 分布式文件系统(HDFS)中,对在线监测数据的索引表结构进行优化,并存储在分布式结构化数据库(HBase)中,实现海量在线监测数据的快速查询。搭建基于 Hadoop 的在线监测数据处理实验平台,进行了基准测试、排序测试以及在线监测数据读写性能测试等。测试结果表明,上述方法及平台在查询时间、读写效率及吞吐量等方面具有良好的性能,能够满足智能变电站大规模在线监测数据的可靠存储及高效处理的需要。

关键词: 智能变电站; 监测; 云计算; Hadoop; 快速查询; 数据处理

中图分类号: TM 73

文献标识码: A

DOI: 10.3969/j.issn.1006-6047.2013.08.024

0 引言

随着智能变电站信息一体化的发展^[1-3],电力状态监测数据呈现以下特点:数据量呈几何性增长,数据规模不断扩大;数据类型结构复杂多样,不仅包括各种实时在线监测数据,还包括设备台账信息、试验数据、缺陷数据等离线信息^[4];数据广域分布、种类繁多,包括各类结构化和半结构化数据,且各类数据查询与处理的频度及性能要求也不尽相同;平台间数据通信困难、交互性差等。现有电力系统的处理方案^[5-6],大都采用常规的数据存储与管理方法,存储硬件采用磁盘阵列,数据库管理软件采用关系数据库系统,导致系统可伸缩性与可靠性较差、数据处理能力不足,难以适应智能变电站对状态监测数据的可靠存储和快速查询的更高要求。

云计算技术具有成本低、可靠性高、易扩展等优势,为智能变电站信息平台的建设提供了全新的解决思路。云计算利用分布式计算数据存储技术,为海量数据提供了不怕故障的存储;其基于列的数据管理技术能够高效地管理各类多元数据;其强大的并行处理能力,能够将电力系统业务逻辑与并行计算的复杂细节划分开来,为快速查询与计算提供便利^[7]。云计算目前主要应用于互联网、商业和科学计算等领域^[8-10],在电力领域已经展开初步研究^[11-12],将云计算技术引入智能变电站状态监测系统,既是一种有益的实践,又具有重要的研究价值与实际意义。

本文结合智能变电站状态监测系统的特点与功能需求,研究基于 Hadoop 云计算技术的变电设备状态监测数据存储方式、数据检测机制以及索引构建方法等关键问题,提出基于 Hadoop 云计算技术构建智能变电站信息平台,使用面向列的分布式结构化数据库 HBase 在开源的 Hadoop 集群上实现智能变电站中在线监测数据的可靠存储及快速查询。

1 变电设备在线监测系统

智能变电站是指采用智能电子设备,自动完成信息采集、测量和监测等基本功能,并支持智能调节、分析决策、协同互动等高级功能的变电站^[13-14]。根据相关技术导则以及智能变电站技术规范的要求^①,变电设备在线监测系统结构如图 1 所示。

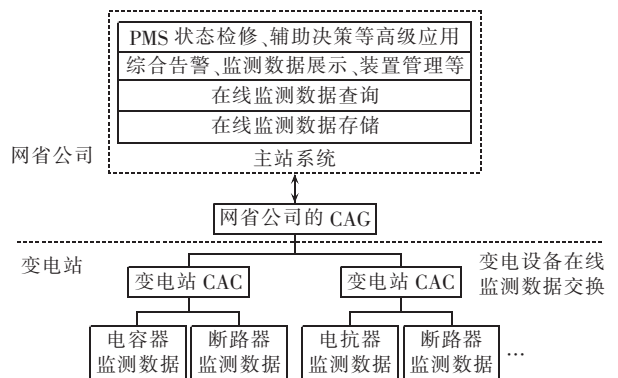


图 1 变电设备在线监测系统结构

Fig.1 Structure of substation equipment online monitoring system

a. 各级地市变电站中都有各自的在线监测接入控制器(CAC),负责收集各种在线监测数据,包括电容器监测数据、电抗器监测数据、断路器监测数据

收稿日期:2012-07-30;修回日期:2013-06-19

基金项目:国家自然科学基金资助项目(61074078);中央高校基本科研业务费专项资金资助项目(12MS113)

Project supported by the National Natural Science Foundation of China(61074078) and the Fundamental Research Funds for the Central Universities(12MS113)

① 国家电网公司. 输变电设备状态监测系统总体框架设计. 2010.

等,然后将这些数据合成处理,生成在线监测数据记录,经过变电设备在线监测数据交换传送到网省公司的在线监测接入网关机(CAG)中。

b. 网省公司的变电站 CAG 对传入的在线监测数据进行处理汇总,并且与 CAC 进行通信,实现对 CAC 的异地维护、升级、控制等工作,同时将收集合成的在线监测数据转发到主站系统。

c. 变电站主站系统通过对变电站内变压器、断路器、避雷器等电气设备状态进行监测,将设备状态监测数据汇总,实现综合分析、故障诊断和状态评估的功能,并为数据挖掘与辅助决策等高级应用提供技术支撑。主站系统的在线监测数据的存储与查询性能很大程度上决定着整个主站系统在线监测数据处理平台的安全、高效和稳定运行。因此,对于在线监测数据的可靠存储和高效查询的方法的研究就显得尤为重要。

2 基于 Hadoop 的在线监测数据处理平台

2.1 基于 Hadoop 的在线监测数据处理框架

Hadoop 是 Apache 开源组织的一个分布式计算框架,支持在大量廉价的硬件设备组成的集群上运行“数据密集型”应用,具有高可靠性和良好的可扩展性。引用 Hadoop 技术,搭建在线监测数据处理平台,可以解决海量在线监测数据存储和查询问题,并且为更高级应用提供高性能的分布式计算环境。

将电力系统总部以及各个网省公司的数据中心服务器、存储设备等硬件资源进行虚拟化,使其升级为云节点,以虚拟机为单位进行统一管理。虚拟化基础设施能够有效地简化数据中心的管理与维护工作,提升服务的可用性与可扩展性^[15]。在虚拟化平台基础上,搭建 Hadoop 在线监测数据处理集群,利用 Hadoop 分布式文件系统(HDFS)以及面向列的 HBase 分布式数据库系统,为海量的在线监测数据提供可靠的存储和查询服务,并为更高级的数据挖掘与辅助决策提供便利。具体过程如图 2 所示。

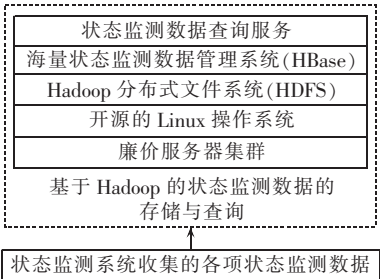


图 2 基于 Hadoop 的在线监测数据处理平台
Fig.2 Online monitoring data processing platform based on Hadoop

在线监测系统收集的各項在线监测数据存储到在线监测数据存储系统中,利用 HDFS 对海量的在

线监测数据进行存储。HDFS 采用主从(Master/Slave)架构,具备较为完善的冗余备份和故障恢复机制,能保证数据读写操作的安全性、可靠性等各方面的要求^[16]。基于 Hadoop 的在线监测存储系统由一个 Namenode 和一定数目的 Datanodes 组成。Namenode 是一个中心服务器,管理在线监测文件系统的名字空间以及客户对文件的访问,其余节点为 Datanodes 负责管理所在节点对在线监测数据的存储。

HBase 建立在 HDFS 之上,是具有高可靠性、高性能、列存储、可伸缩、实时读写特点的数据库系统,通过行键(row key)和主键的范围来检索数据,用来存储非结构化和半结构化的松散数据。HBase 的数据表结构非常松散,表与表之间没有任何关联,其查询效率比较高,能够为海量的在线监测数据查询服务提供支持。

2.2 基于 HBase 的在线监测数据索引建立方法

HBase 表的结构如表 1 所示,其由行键、时间戳以及列族组成。行键为表的主键,表中的记录按照行键排序;时间戳表示每次数据操作对应的时间;列族可以由任意多个列组成,并且支持动态扩展。

表 1 基于 HBase 的变电设备在线监测数据概念视图
Tab.1 Conceptual view of substation equipment online monitoring data based on HBase

行键	时间戳	列族		
		被监测设备码	监测时间	顶层油温/℃
021001	T1	23M00000005482691		
021005	T2		2012-06-30 12:20:20	80
021005	T3	23M00000005482672		78

行键可设置为监测类型编码,例如 021001 代表监测类型为变压器及电抗器的局部放电,021005 则代表监测类型为变压器及电抗器的顶层油温。列族包括相应的在线监测数据内容,如被监测设备码、监测装置标识、被监测设备相别组成和监测时间等。设备编码采用国家统一编码方式,由 3 段共 17 位字符组成,格式如下:xxMxxxxxxxxxxxxxx,其中第 1 段 xx 为网省公司标识;第 2 段采用固定标识符 M;第 3 段采用 14 位流水号标识。

HBase 需要保证用户在数据的低冗余及一致性的基础上,对海量的在线监测数据进行高性能的范围检索。随着数据量的大规模增加,这种方法给用户查询带来了巨大的不便。利用 HBase 创建索引表,以行键来存储索引,列族用来存储被监测设备标识、监测时间、监测装置标识等其他数据。

HBase 索引表往往是 1 个行键对应 1 个索引,然后依靠表的行键来实现范围检索。这种方法的特点是简单、扫描效率相对较高,但在查询上还是不够

简洁方便,并且在查询处理以及数据的合并汇总上都会出现瓶颈。为了改善查询效率,采用组合索引的方式,将建立索引的字段拼接组合,提高查询效率,如表 2 所示。

表 2 基于 HBase 在线监测数据的组合索引表
Tab.2 Comprehensive index table based on HBase for online monitoring data

行键	列族		
	被监测设备码	监测装置标识	相别
021005_yyyy-MM-dd	17 位设备码	17 位设备码	XX
021005_yyyy-MM-dd	17 位设备码	17 位设备码	XX
021005_yyyy-MM-dd	17 位设备码	17 位设备码	XX
021005_yyyy-MM-dd	17 位设备码	17 位设备码	XX

将监测时间 yyyy-MM-dd 与监测类型编码 021005(变压器顶层油温)组合拼接到行键中,建立索引表。通过这种组合索引的方式,就可得到该天所有变压器顶层油温的数据,并分析数据完成在线监测数据的查询。另外,可以根据不同的场景和需要创建不同的查询索引,例如将被监测设备码放置到行键中创建查询索引(如 021005_23M00000005482691),所得到的结果就是针对 23M00000005482691 这一被监测设备所收集到的所有在线监测数据。组合索引方案的优点是查询速度非常快、效率高,用户可以根据查询条件,创建查询索引,通过对查询索引表的检索即可得到数据结果集。

2.3 基于 Hadoop 的在线监测数据快速查询方法

在线监测数据文件存储 HDFS 中,为了提高快速查询处理能力,针对在线监测数据建立查询索引,并将这些查询索引存放到 HBase 中。基于分布式文件存储系统以及查询索引,完成大规模在线监测数据的查询。查询方法分为以下 3 个步骤,具体如图 3 所示。

a. CAC 以及 CAG 负责在线监测数据的收集,然后传送到 HDFS 集群中,状态数据以文件形式组织,现阶段的处理方式是直接将数据以二进制的形式存放到文件里,不包含任何冗余数据。后续查询将通过文件名和偏移量来获取数据。

b. HDFS 设置一个主节点的数据监测接口用于检测 HDFS 中新的状态数据文件,然后该主节点把检测到的数据文件分发给不同的子节点完成查询索引

创建。索引创建过程如下:读取文件,对在线监测数据进行扫描,将需要建立索引的常用字段提取出来加入到在线监测数据记录列表中。

c. 根据这些常用字段在 HBase 中建立索引表,方便数据的查询。基于 HBase 的查询索引,接受并处理用户的状态数据查询请求,到 HBase 中完成具体的查询处理,在获得结果在线监测数据记录列表后,再到存储在 HDFS 的状态数据文件中读出详细的在线监测记录,并将得到的数据逐一返回给客户端。

该方法中将会实时监测是否有新的文件生成,而且查询索引处理过程是分布式的;且通过使用 Zookeeper,管理者不必关心系统的负载均衡和单点失效问题。

3 智能变电站云实验平台及性能测试

3.1 智能变电站云实验平台的搭建

根据上述方法,搭建一个基于 Hadoop 的智能变电站在线监测实验平台。整个实验平台由 10 台 PC 机(作为 10 个节点)和 1 个交换机构成,PC₁ 作为 Name-node 和 jobtracker,PC₂ 至 PC₁₀ 作为 Datanode 和 task-tracker。每个节点安装 VirtualBox 虚拟机,在虚拟机上安装 Ubuntu Linux 操作系统,并安装 Apache Hadoop 云计算平台。HDFS 管理界面显示 Datanode 各个节点使用空间、使用率、大小等内容。

将模拟产生的在线监测的数据,导入到基于 Hadoop 的在线监测数据处理平台中,利用 TeraSort、Sort、TestDFSIO、YCSB 等测试工具对所搭建的在线监测实验平台的性能进行测试。

3.2 在线监测数据排序能力测试

Sort 是一种主要测试 Hadoop MapReduce 性能的工具,用于衡量分布式数据处理框架的数据处理能力。Sort 自带一个部分排序的程序,实验过程通过生成随机的数据,执行排序,然后验证结果。

首先通过使用 RandomWriter 生成一些随机的在线监测数据。它以每个节点 10 个 map 的方式运行一个 MapReduce 作业,并且每一个 map 生成近似 10 GB 的随机二进制在线监测数据,带有不同长度的键和值。

执行排序,对产生的在线监测数据文件进行排序测试,实验结果表明,实验平台对所传送的在线监测数据文件能够进行 sort 排列。取 5 次排序时间的平均值,分别为 324 s、298 s、374 s、341 s 和 382 s,平均时间为 342.8 s。实验表明,在线监测实验平台具有良好的数据计算能力。

3.3 大规模在线监测数据块基准测试

智能变电站监测的设备众多,数据量庞大,数据结构复杂。对所搭建平台进行大数据块基准测试,

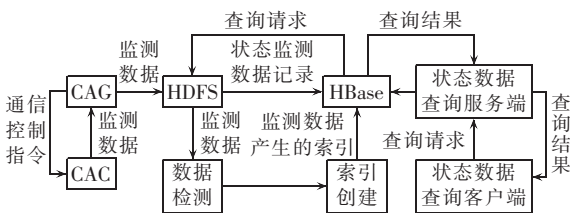


图 3 基于 Hadoop 的在线监测数据查询方法
Fig.3 Online monitoring data query based on Hadoop

常用测试工具为 TestDFSIO。通过生成数据,提交作业,统计作业完成时间对实验平台的整体 I/O 性能进行了测试。测试程序用一个 MapReduce 作业对 HDFS 进行高强度的 I/O 操作,来测试集群整体的并行写入以及读取数据的性能。

实验通过改变数据规模、文件大小、文件数量,来对获取的运行时间进行比较,测试其对集群 I/O 性能的影响,如表 3 所示。

表 3 基准测试读写运行时间

Tab.3 Read/write speed by benchmark test

测试	文件数量	文件大小/MB	运行时间/s
Write	10	1 000	330.861
Write	10	2 000	554.465
Write	10	5 000	679.542
Write	20	1 000	386.252
Write	20	2 000	412.975
Write	20	5 000	598.346
Read	10	1 000	189.415
Read	10	2 000	220.855
Read	10	5 000	268.358
Read	20	1 000	200.568
Read	20	2 000	254.622
Read	20	5 000	315.639

实验说明,逐渐增大数据规模进行读、写操作,系统执行时间增长趋缓,平均访问时间有效降低,相比传统数据管理系统平台,随着数据量的增加,基于 Hadoop 的在线监测数据管理系统对文件的读写效率更好。这是由于传统数据库在随着数据量的增加的同时,负荷逐渐增加,导致性能越来越慢,然而基于 Hadoop 在线监测数据处理平台在大规模文件处理方面显示了很强的性能,实验显示 Hadoop 集群适合处理大数据量的读写。

在实验过程中,开始取 10 个 1 000 MB 的文件,进行读写测试,随着测试数据和文件的不断增大,系统性能也逐渐提升。再次用 10 个 1 000 MB 的文件进行测试时,运行时间仅仅用了 162.673 s,接近开始时的 330.861 s 的一半。实验结果表明,该实验平台的系统性能在处理过程中逐渐提升,如果继续增加服务器节点的数量,其读写速度还会有更大的提高。

3.4 在线监测数据读写查询性能测试

在线监测要求对所采集的数据进行分析汇总,有效提取故障信息,并对海量的在线监测数据进行快速查询。YCSB 是 Yahoo 公司用来对云服务进行基础测试的工具,具有很好的可扩展性和灵活性^[17]。通过 YCSB 进行写入数据、读取数据的性能测试,验证实验平台的在线监测数据查询性能。

3.4.1 数据导入测试

根据所设计的 HBase 存储模式,通过模拟 CAG 接口传入的 10 万条变压器顶层油温数据记录(例如

监测类型编码 021005,被检测设备标识 23M000000-05482691,顶层油温 80 ℃,监测时间为 2012-06-30 12:30:25),累计约为 10 GB,导入 HBase 数据库。总体运行时间为 1 004.155 s,系统吞吐量达到每秒操作数 142.33,单条记录导入平均时延 39.471 ms。

3.4.2 组合索引对比测试

以变压器顶层油温在线监测数据为例,建立组合索引表,将冗余在线监测数据的监测时间设置到行键中进行检索(021005_2012-06-30),见表 4。

表 4 变压器顶层油温在线监测数据的组合索引表

Tab.4 Comprehensive index table for online monitoring data of top transformer oil temperature

行键	列族	
	被监测设备码	监测装置标识
021005_2012-06-30	23M00000005482691	23M00000005482691
021005_2012-06-30	23M00000005482691	23M00000005482691
021005_2012-06-30	23M00000005482691	23M00000005482691
021005_2012-06-30	23M00000005482691	23M00000005482691

未采用查询索引优化时,即行键为 021005,监测时间为列族时,扫描所用时间平均为 1 823 ms,吞吐量为每秒操作数 162.09,经过组合索引优化后的数据扫描时间为 1 412 ms,吞吐量为每秒操作数 228.23。实验结果显示,经过组合索引优化,扫描操作时间更短,采用组合索引查询的效率和吞吐率方面的性能均得到提升。

3.4.3 读、写延迟对比测试

在变电站在线监测系统中,大多数应用场景都是对数据的多次读取操作,数据的读取操作往往比写入操作更加频繁,写入操作通常只进行一次,更新数据也很少进行,因此平台的并行读取性能就显得格外重要。

对已存储的 10 万条变压器顶层油温在线监测数据,执行 10 000 次读写操作,包括 80% 的读指令和 20% 的写指令。取 3 次测试结果的平均值,图 4 描述了单条指令执行的不同的读写平均延迟。3 次测试的吞吐量变化平缓,系统运行较稳定,数据读取操作的平均延迟比写操作的平均延迟要小,该实验平台具有良好的数据读取性能,可以满足大规模数据的读取和查询的要求。

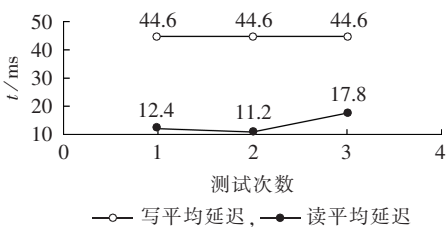


图 4 3 次试验中读、写平均延迟

Fig.4 Average read/write latency of three tests

4 结语

本文通过研究智能变电站状态监测数据特点,结合 Hadoop 云计算技术,提出一种在线监测数据存储和查询的解决方法,即基于 Hadoop 的平台利用查询索引对海量的在线监测数据进行快速查询。实验结果显示在存储容量、查询延迟、数据可靠性和可扩展性等方面,所提方法均显示了较高的性能,在一定程度上验证了该方法的可行性与有效性,能够满足智能变电站大规模在线监测数据的存储与处理的要求。

参考文献:

- [1] 周邨飞,徐石明. 智能变电站数据中心初探[J]. 电力系统自动化,2011,35(18):89-92.
ZHOU Yefei,XU Shiming. A preliminary investigation on data center in smart substation[J]. Automation of Electric Power Systems,2011,35(18):89-92.
- [2] 薛晨,黎灿兵,黄小庆,等. 智能变电站信息一体化应用[J]. 电力自动化设备,2011,31(7):110-114.
XUE Chen,LI Canbing,HUANG Xiaoqing,et al. Application of intelligent substation information integration[J]. Electric Power Automation Equipment,2011,31(7):110-114.
- [3] 蒋宏图,袁越,杨昕霖. 智能变电站一体化信息平台的设计[J]. 电力自动化设备,2011,31(8):131-134.
JIANG Hongtu,YUAN Yue,YANG Xinlin. Design of integrated information platform of smart substation[J]. Electric Power Automation Equipment,2011,31(8):131-134.
- [4] 张怀宇,朱松林,张杨,等. 输变电设备状态检修技术体系研究与实施[J]. 电网技术,2009,33(13):70-73.
ZHANG Huaiyu,ZHU Songlin,ZHANG Yang,et al. Transmission and distribution equipment condition based maintenance technology systems research and practice[J]. Power System Technology,2009,33(13):70-73.
- [5] 王晓波,樊纪元. 电力调度中心统一数据平台的设计[J]. 电力系统自动化,2006,30(22):89-92.
WANG Xiaobo,FAN Jiyuan. Design of power dispatch center unified data platform[J]. Automation of Electric Power Systems,2006,30(22):89-92.
- [6] 林峰,胡牧,蒋元晨,等. 电力调度综合数据平台体系结构及相关技术[J]. 电力系统自动化,2007,31(1):61-64.
LIN Feng,HU Mu,JIANG Yuanchen,et al. Architecture and related techniques of a power dispatching data platform[J]. Automation of Electric Power Systems,2007,31(1):61-64.
- [7] 陈康,郑纬民. 云计算:系统实例与研究现状[J]. 软件学报,2009,20(5):1337-1348.
CHEN Kang,ZHENG Weimin. Cloud computing:system instances and current research[J]. Journal of Software,2009,20(5):1337-1348.
- [8] di COSTANZO A,de ASSUNCAO M D,BUYIA R. Harnessing cloud technologies for a virtualized distributed computing infrastructure[J]. IEEE Internet Computing,2009,13(5):24-33.
- [9] 郑湃,崔立珍,王海洋,等. 云计算环境下面向数据密集型应用的数据布局策略与方法[J]. 计算机学报,2010,33(8):1472-1480.
ZHENG Pai,CUI Lizhen,WANG Haiyang,et al. A data placement strategy for data-intensive applications in cloud[J]. Chinese Journal of Computers,2010,33(8):1472-1480.
- [10] IOSUP A,OSTERMANN S,YIGITBASI M N,et al. Performance analysis of cloud computing services for many-tasks scientific computing[J]. IEEE Trans on Parallel and Distributed Systems,2011,22(6):931-945.
- [11] 赵俊华,文福拴,薛禹胜,等. 云计算:构建未来电力系统的核心计算平台[J]. 电力系统自动化,2010,34(15):1-8.
ZHAO Junhua,WEN Fushuan,XUE Yusheng,et al. Cloud computing:building a power system of the future core computing platform[J]. Automation of Electric Power Systems,2010,34(15):1-8.
- [12] 王德文,宋亚奇,朱永利. 基于云计算的智能电网信息平台[J]. 电力系统自动化,2010,34(22):7-12.
WANG Dewen,SONG Yaqi,ZHU Yongli. Information platform for smart grid based on cloud computing[J]. Automation of Electric Power Systems,2010,34(22):7-12.
- [13] 鲁东海,孙晓军,王晓虎. 智能变电站中在线监测系统的设计[J]. 电力自动化设备,2011,31(1):134-137.
LU Donghai,SUN Xiaojun,WANG Xiaohu. Online monitoring system for smart substation[J]. Electric Power Automation Equipment,2011,31(1):134-137.
- [14] 王孟超,王允平,李献伟,等. 智能变电站及技术特点分析[J]. 电力系统保护与控制,2010,38(18):59-63.
LI Mengchao,WANG Yunping,LI Xianwei,et al. Smart substation and technical characteristics analysis[J]. Power System Protection and Control,2010,38(18):59-63.
- [15] 王德文. 基于云计算的新一代电力数据中心基础架构及其关键技术[J]. 电力系统自动化,2012,36(11):67-71.
WANG Dewen. Based framework and key technology for a new generation of data center in electric power corporation based on cloud computing[J]. Automation of Electric Power Systems,2012,36(11):67-71.
- [16] WBITE T. Hadoop 权威指南[M]. 2版. 北京:清华大学出版社,2011:44-49.
- [17] COOPER B F,SILBERSTEIN A,TAM E,et al. Benchmarking cloud serving systems with YCSB[C]//ACM Symposium on Cloud Computing. Indianapolis,Indiana,USA:[s.n.]:143-154.
- [18] 王德文,阎春雨,毕建刚,等. 输变电状态监测系统的分布式数据交换方法[J]. 电力系统自动化,2012,36(22):83-88.
WANG Dewen,YAN Chunyu,BI Jiangang,et al. Distributed data exchange method in condition monitoring system for power transmission and transformation equipment[J]. Automation of Electric Power Systems,2012,36(22):83-88.
- [19] BALIGA J,ROBERT W A A,HINTON K,et al. Green cloud computing:balancing energy in processing,storage and transport[J]. Proceedings of the IEEE,2010,99(1):149-167.
- [20] ISOM P K,MILLER-SYLVA S L,VAIDYA S. Intelligent enterprise architecture[J]. Research and Development,2010,54(4):1-2.

作者简介:

王德文(1973-),男,黑龙江克山人,副教授,博士,研究方向为电力系统自动化和计算机网络(E-mail:wdewen@gmail.com);

肖磊(1987-),男,河北保定人,硕士研究生,研究方向为电力信息化与云计算数据处理;

肖凯(1987-),男,河北邯郸人,硕士研究生,研究方向为电力信息化与云计算数据挖掘。

(下转第156页 continued on page 156)

- [J]. 中国电机工程学报,2006,26(6):14-18.
- ZHANG Dongxia,TANG Yong,ZHU Fang,et al. Equivalent model for induction-motor load connected to high voltage bus [J]. Proceedings of the CSEE,2006,26(6):14-18.
- [18] 王庆平,陈超英,陈礼义. 暂态全过程仿真中面向对象的负荷自适应模型[J]. 电网技术,2002,26(8):29-33.
- WANG Qingping,CHEN Chaoying,CHEN Liyi. Object-oriented self-adaptive load model in whole transient simulation [J]. Power System Technology,2002,26(8):29-33.
- [19] 侯俊贤,汤涌,张红斌,等. 感应电动机的综合方法研究[J]. 电网技术,2007,31(4):36-41.

HOU Junxian,TANG Yong,ZHANG Hongbin,et al. Study on integration method for induction motor [J]. Power System Technology,2007,31(4):36-41.

作者简介:

陶 华(1983-),男,湖南衡阳人,硕士,主要从事电力系统动态仿真、智能电网方面的研究工作(E-mail:miaoyan0702@163.com);

许津津(1987-),女,福建泉州人,硕士,主要从事电力系统动态仿真、状态检修方面的研究工作。

Model conversion from BPA to PSCAD

TAO Hua,XU Jinjin,ZOU Wencong

(Quanzhou Electric Power Bureau,Quanzhou 362000,China)

Abstract: The electromagnetic transient simulation of electric power system is an important test means of relay protection performance,as well as a main reliability analysis method of AC/DC hybrid transmission system. The present power system transient simulation in China is mostly for the electromechanical transient process and the grid data are consequently stored in the data format of electromechanical transient simulation software,resulting in the shortage of simulative electromagnetic transient data. The difference between BPA(a kind of electromechanical transient simulation software) and PSCAD(a kind of electromagnetic transient simulation software) is analyzed in detail and an implementation method of model conversion from BPA to PSCAD is presented. BPA and PSCAD are applied respectively to cases and its electric topology and transient characteristics are compared between two kinds of software,which proves the effectiveness and accuracy of the proposed conversion method.

Key words: electromagnetic transient; BPA; PSCAD; model conversion; electric power systems; computer simulation

(上接第146页 continued from page 146)

Processing of massive online monitoring data in smart substation

WANG Dewen,XIAO Lei,XIAO Kai

(Department of Computer Science,North China Electric Power University,Baoding 071003,China)

Abstract: The characteristics of smart substation monitoring data and the technology of Hadoop cloud computing are analyzed and the mechanism of data examination and the method of index establishment are studied. A method based on Hadoop is proposed for the reliable storage and fast query of smart substation online monitoring data,which stores the collected massive online monitoring data of substation equipments in HDFS(Hadoop Distributed File System) redundantly,optimizes the structure of index table for online monitoring data and stores it in HBase(Hadoop column-oriented dataBases) to achieve the fast query of massive online monitoring data. An online monitoring platform based on Hadoop is built,with which,the benchmark test,sort test and read/write efficiency test are carried out. Test results indicate that,the proposed method and platform have high performance in query speed,read/write efficiency,latency and throughput,meeting the requirements of smart substation for reliable storage and efficient processing of massive online monitoring data.

Key words: smart substation; monitoring; cloud computing; Hadoop; fast query; data processing