

基于 Hadoop 的智能家居信息处理平台

徐源吾^{1,2}, 王 珣^{1,3}

(1. 南京邮电大学 通信与信息工程学院, 江苏 南京 210003;

2. 南京邮电大学 通信技术研究, 江苏 南京 210003;

3. 南京邮电大学 宽带无线通信与传感器技术教育部重点实验室, 江苏 南京 210003)

摘 要:智能家居包含门禁系统、温湿度系统、室内滴灌系统、照明系统等。每一个系统包含了不同的终端, 每一个终端安装有不同的传感器, 传感器感知周围环境产生大量数据。用户数的不断增加将产生海量数据。文中针对传统的海量数据处理方法会出现硬件成本高、计算瓶颈等问题, 在云计算理论的基础上针对智能家居中的门禁系统、温湿度系统、滴灌系统和照明系统搭建了基于 Hadoop 的数据处理平台。该平台对大量传感信息进行分布式并行处理, 并根据处理结果向底层设备发送指令以实现智慧化服务, 有效解决了成本和计算瓶颈问题。

关键词:智能家居; 云计算; Hadoop; 分布式并行处理

中图分类号: TP302.1

文献标识码: A

文章编号: 1673-629X(2014)09-0183-04

doi: 10.3969/j.issn.1673-629X.2014.09.042

Information Processing Platform of Smart Home Based on Hadoop

XU Yuan-wu^{1,2}, WANG Xun^{1,3}

(1. College of Telecommunications & Information Engineering, NUPT, Nanjing 210003, China;

2. Institute of Communication Technology, NUPT, Nanjing 210003, China;

3. Key Lab on Wideband Wireless Communications and Sensor Network Technology of Ministry of Education, NUPT, Nanjing 210003, China)

Abstract: Smart home includes access control system, temperature and humidity system, indoor drip irrigation system and lighting system etc. Each system includes different terminals and each terminal is equipped with different sensors. Sensor perceives its surroundings and generates large amounts of data. With the increasing number of users, it will generate vast amounts of data. As the traditional method of massive data processing has shortcomings of high cost in hardware and has developed to a bottleneck of calculation, a data processing platform based on cloud computing is designed for the access control system, temperature and humidity system, indoor drip irrigation system and lighting system. This platform can realize the distributed parallel processing of large amounts of sensor information, and then send commands to the equipment according to the process result to realize the smart service, which effectively solves the problems in cost and bottleneck of calculation.

Key words: smart home; cloud computing; Hadoop; distributed parallel processing

0 引 言

近年来,在物联网等具有海量数据需求的应用变得越来越普遍时,无论是从科学研究还是从应用开发的角度来看,掌握云计算技术已成为一种趋势。智能家居作为物联网应用的主要领域之一^[1],自然离不开云计算。云计算的发展使用户无需关注复杂的内部工作机制、无需具备丰富的分布式系统知识及开发经验,即可实现大规模分布式系统规定部署,以及海量数据

的并行处理^[2]。

Hadoop 是目前较为成熟和应用较为广泛的云计算架构之一。Hadoop 的核心思想是分布式并行处理。Hadoop 技术中的关键技术是 HDFS (分布式文件系统) 和 Map/Reduce (映射/规约)^[3]。Hadoop 利用 HDFS 存储海量数据,利用 Map/Reduce 进行海量数据的计算^[4]。

文中首先介绍了智能家居与云计算的概念,接着

收稿日期: 2013-11-11

修回日期: 2014-02-16

网络出版时间: 2014-07-17

基金项目: 国家自然科学基金资助项目(61271237)

作者简介: 徐源吾(1989-),男,硕士研究生,研究方向为无线通信与电磁兼容、云计算在物联网中的应用。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20140717.1229.029.html>

针对传统的海量数据处理方法硬件成本太高,存在计算瓶颈的问题,在云计算理论的基础上针对智能家居中的门禁系统、温湿度系统、滴灌系统和照明系统搭建了基于 Hadoop 的数据处理平台。可以对大量传感信息进行分布式并行处理,并根据处理结果向底层设备发送指令以实现智慧化服务。

1 智能家居与云计算

1.1 智能家居

目前通常把智能家居定义为利用电脑、网络和综合布线技术,通过家庭信息管理平台将与家居生活有关的各种子系统有机结合的一个系统^[5]。首先,需要在家居中建立一个通信网络,为家庭信息交互和传输提供必要的通路,在家庭网络的操作系统控制下,通过相应的硬件和执行机构,实现对所有家庭网络上的家电和设备的控制和监测。其次,智能家居系统需要通过一定的媒介平台,提供与外界的通信通道,以实现与家庭以外的世界沟通信息,满足远程控制/监测和交换信息的需求^[6]。

1.2 云计算架构 Hadoop

Hadoop 是一个分布式计算框架,可以在大量廉价的硬件设备组成的集群上运行应用程序,为应用程序提供了一组稳定可靠的接口,旨在构建一个具有高可靠性和良好扩展性的分布式系统^[7]。Hadoop 的主要优点有:扩容能力强、成本低廉、效率高、可靠性高、免费开源及良好的可移植性。

Hadoop 项目包括多个子项目,但主要是由 Hadoop 分布式文件系统 (Hadoop Distributed File System, HDFS) 和映射/规约引擎 (Map/Reduce Engine) 两个主要的子项目构成^[8]。

1.3 分布式文件系统 (HDFS)

HDFS 由一个名叫 NameNode 的主节点和多个名叫 DataNode 的子节点组成,是一种典型的主从式 (Master/Slave) 架构,这种架构方法可以通过主节点屏蔽底层的复杂结构,并向 Client 提供方便的文件目录映射^[9]。NameNode 存储着文件系统的元数据,这些元数据包括文件系统的名字空间等,并负责管理文件的存储等服务,但实际的数据并不存放在 NameNode,而是由 HDFS 中的 DataNode 来存放数据,然后由 Client 直接与 DataNode 建立数据通信。

此外,HDFS 为了可靠地海量存储文件,各个文件以块序列的形式存储。为了保证故障容错,文件的块被复制。块的大小和副本的个数都可以配置。NameNode 控制所有的块复制操作,它周期性地接收来自集群中 DataNode 的“心跳”回应和块报告^[10]。收到一个节点的“心跳”回应表示这个 DataNode 是正常的。

1.4 MapReduce 分布式计算

MapReduce 分布式计算主要是用来在数千部服务器上同时处理保存于 GFS Clusters 中高达几十 TB 甚至更多的数据^[11]。

MapReduce 在温度值统计过程中的运作如下所述。

MapReduce 的作用是对计算机集群上的大型数据集运行分布式计算,替大量数据做平行计算处理。MapReduce 的整个架构是由 Map 和 Reduce 函数组成的^[12],当程序输入一大组 Key/Value 键值对时,Map 函数自动将原本的 Key/Value 分拆成多组中介的键值对,然后 Reduce 函数再合并具有相同 Key 的中介值配对,化简成最后的输出结果。

一个智能家居系统中温度值统计程序所运行的 MapReduce 计算过程概念,如图 1 所示。

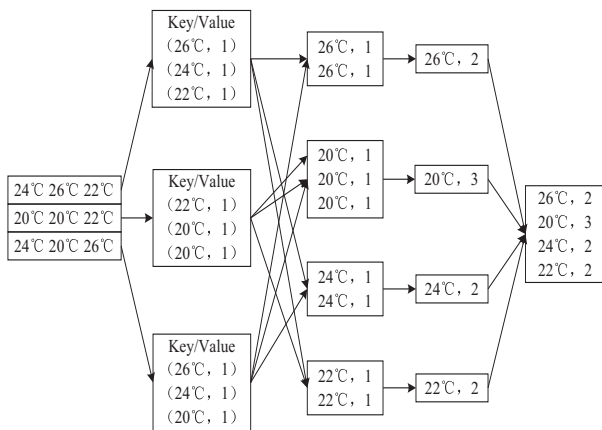


图 1 智能家居温度值统计 MapReduce 计算过程

1.5 智能家居云模型

智能家居拥有 7 大子系统,每一个系统包含了不同的终端,每一个终端都安装有传感器,传感器感知周围环境并产生大量数据^[13]。千千万万的用户会产生海量的数据,而云计算可靠的数据存储,强大的计算能力,低廉的运营成本使得它非常适合处理海量数据。于是每家用户都将通过网络连接到云服务中心,产生的数据均由云计算中心存储和计算处理。

2 开发程序设计分析

本平台针对智能家居中的几个典型应用:门禁、温度、湿度、光照等进行信息采集及分布式并行处理。启动程序后可以看到当前房间内的温度、湿度、土壤湿度、光强度,还可以通过对室内温度、湿度、土壤湿度进行设定来改变当前的状态。

2.1 功能模块设计

根据需求分析可知,系统功能模块共分为四大部分,即门禁系统、温湿度系统、滴灌系统、照明系统。

门禁系统实现的功能是持卡人刷卡,产生持卡人

的卡号,从事先设定好的数据库中查询是否有该卡号,如果有说明该用户为合法用户,门禁打开,并显示持卡人卡号、姓名、卡的类型,否则门禁继续保持关闭。

温湿度系统显示室内当前温度和湿度。当设定一个温度值后,空调工作,根据设定温度调整室内温度,直到室内当前温度达到设定值;当设定湿度值后,室内湿度就会根据所设定的湿度进行调整,直到室内当前湿度达到设定值。

滴灌系统显示当前的土壤湿度。当设定土壤湿度后,滴灌系统工作,调整湿度直到当前湿度为设定的土壤湿度。

在照明系统中,显示室内当前光强度。窗帘和电灯的状态受控于光照强度的大小。

智能家居云计算平台实现的功能主要是能够将产生的传感信息集中交给服务器处理,用户只需将传感器采集来的传感信息传输给服务器即可。服务器为计算机集群,这里使用三台电脑进行模拟。三台电脑对信息进行分布式并行处理,并根据处理结果对底层设备发出指令,以达到智能控制的效果。同时还要将各种传感数据和底层设备的状态写入数据库,如图 2 所示。

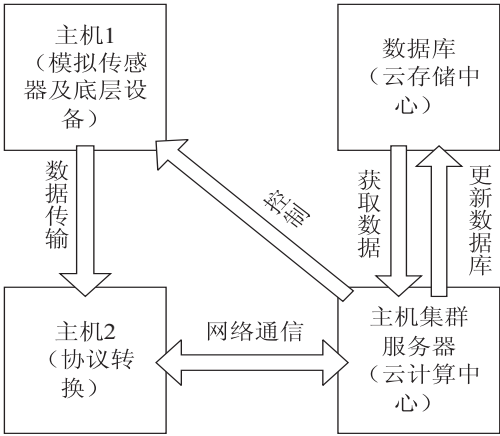


图 2 智能家居云计算平台

2.2 逻辑结构设计

(1) 门禁系统。

门禁状态一开始为关闭。当持卡人来刷卡时,获得持卡人卡号,查询数据库,判断卡的 ID 是否合法,如果 ID 不合法,则门禁保持关闭状态。如果合法则门禁打开,并且显示持卡人的姓名和性别。

(2) 温湿度系统。

系统启动时首先要对温度进行判断。当前温度低于下限值,空调启动制热;当前温度高于上限值,空调启动制冷;当室内当前温度调整到舒适范围时空调自动关闭。系统同时显示室内当前湿度,通过“设定湿度”按钮可以设置室内湿度,当前湿度会慢慢调整到所设置的湿度。

(3) 室内滴灌系统。

滴灌系统的状态主要受控于土壤湿度。这里土壤湿度依然采用相对湿度。

系统开始运行时滴灌处于关闭状态,然后判断土壤湿度是否小于 0.15,如果小于 0.15,滴灌打开,喷水,土壤湿度开始上升,当土壤湿度达到预先设定的上限值时(0.4),滴灌自动关闭。滴灌关闭后由于水分的蒸发,土壤湿度会逐渐减少。系统每隔 1 s 对土壤湿度进行一次判断,如果土壤湿度少于 0.15,则滴灌再次打开。如此反复。当然,土壤湿度的上限值也可以通过“设定土湿”按钮进行设定。具体过程见图 3。

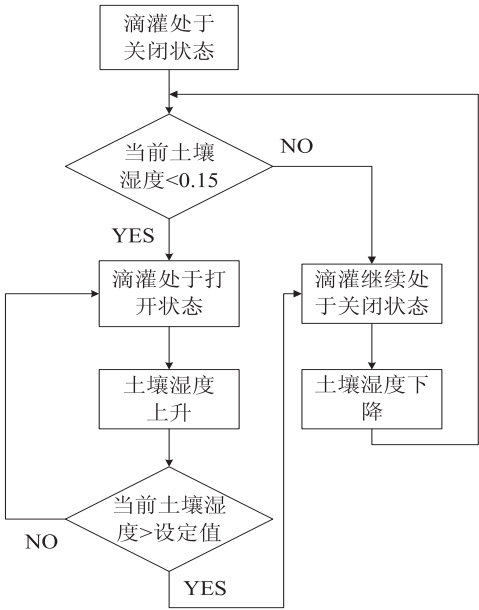


图 3 滴灌系统工作流程图

(4) 照明系统。

照明系统主要由窗帘和电灯两个部分组成。窗帘和电灯的状态会根据光照强度的大小以及当前时间自动改变。根据文献[1]将客厅照明的最低照明度定为临界值,即 150 Lux。

(5) 智能家居云平台。

整个云平台采用五台电脑进行模拟,第一台电脑作为传感器和底层设备的模拟,用于产生传感信息。第二台电脑作为连接底层设备和服务器的网关,剩下三台电脑构成服务器集群,作为云处理中心,并与数据库相连,对传感信息进行分布式并行处理。服务器与客户端之间的通信使用 Socket。

3 平台仿真

3.1 门禁系统测试

运行程序后,门禁处于关闭状态。单击“生成 ID”按钮,会产生当前持卡人卡的 ID 号。然后单击“刷卡”。搜索数据库中已存在的合法 ID 号,合法来访者,门禁打开,并显示持卡人的姓名和性别。门禁打开 5 s

后再次关上。如果 ID 号不存在,则门禁继续保持关闭状态,如图 4 所示。

3.2 温湿度系统测试

运行程序后显示当前温度、湿度,以及默认的下限温度(18 ℃)和上限温度(23 ℃)。系统将当前温度与下限温度和上限温度进行比较,如果当前温度在人体舒适度范围 18 ℃ ~ 23 ℃ 内,则空调继续保持关闭状态。当前室内温度低于下限温度时,空调打开,处于制热状态,直到当前室内温度达到舒适度的下限温度。当前室内温度高于上限温度时,空调打开,处于制冷状态,直到当前室内温度达到舒适度的上限温度。温度的上下限值可以通过设置按钮进行设定。

在“设定湿度”旁的文本框中输入所要设置的室内湿度,然后按“设定湿度”按钮,这时室内的温度就会慢慢改变,上升或下降,直到当前湿度达到所设定的湿度,如图 4 所示。



图 4 门禁系统及温湿度系统

3.3 滴灌系统测试

运行程序后显示当前土壤湿度,滴灌系统一开始处于关闭状态。由于土壤中水分的蒸发,土壤湿度不断下降。当土壤湿度降到 0.15 时滴灌自动打开,开始喷水。滴灌喷水后土壤湿度增加,直到当前土壤湿度达到设定值(默认值为 0.4),这时滴灌再次关闭,如图 5 所示。

3.4 照明系统测试

运行程序后显示当前的室内光强度,如果是早晨七点到晚上六点,则窗帘打开,其他时间段窗帘关闭。电灯的状态不仅受时间段控制,也受室内光强度控制。如果是晚上十点到第二天早晨七点,电灯处于关闭状态。在早晨七点到晚上六点之间,电灯一般是关闭状态,除非室内自然光强度低于 150 Lux,晚上六点到晚上十点,电灯是打开状态的。上午九点四十五的系统运行结果如图 5 所示,此时窗帘是打开状态,电灯是关闭状态。

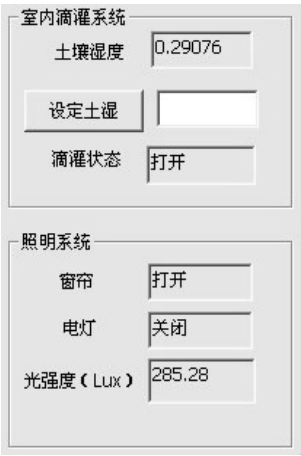


图 5 滴灌系统及照明系统

3.5 云平台测试

将五台电脑建立连接并启动,第一台电脑用于模拟传感器和底层设备,第二台电脑作为连接底层网络和上层网络的网关,剩下三台电脑作为服务器集群(云计算中心)。第一台电脑产生传感数据后传给第二台电脑,第二台电脑作为网关将数据转发给服务器集群,服务器集群采用 Hadoop 计算架构对传感数据进行分布式并行处理,判断当前智能家居各系统的状态,并发出指令给底层设备,达到智能控制的效果。运行结果见图 6。



图 6 云平台运行结果

4 结束语

文中在云计算理论的基础上针对智能家居中的门禁系统、温湿度系统、滴灌系统和照明系统搭建了基于 Hadoop 的数据处理平台。可以对大量传感信息进行分布式并行处理,并根据处理结果向底层设备发送指令以实现智慧化服务。此云计算平台是在 Hadoop 分布式计算框架的基础上采用三台电脑模拟服务器集群实现对海量数据的并行处理,有效解决了成本和计算瓶颈问题。

参考文献:

[1] Singer P. The internet of things[J]. Solid State Technology, (下转第 190 页)

完成 VxWorks 下的三维图形开发,同时可以方便地应用在实际环境中。

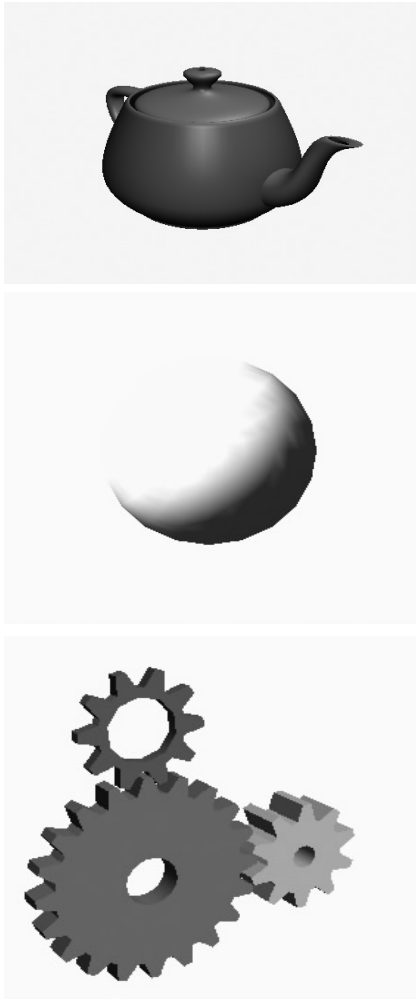


图 4 测试结果

参考文献:

- [1] 陈智宇,温彦君,陈 琪. VxWorks 程序开发实践[M]. 北京:人民邮电出版社,2004.
 - [2] 王学龙. 嵌入式 VxWorks 系统开发与应用[M]. 北京:人民邮电出版社,2003.
 - [3] WindML DDK programmer's guide[M]. USA: Wind River Systems, Inc., 2002.
 - [4] 欧训勇. OpenGL 程序移植技术浅析[J]. 计算机应用研究, 2002, 19(9): 158-160.
 - [5] Shreiner D, The Khronos OpenGL ARB Working Group. OpenGL programming guide[M]. 李 军,徐 波,译. 7th ed. 北京:机械工业出版社,2010.
 - [6] Segal M, Akeley K, Leech J. The OpenGL graphics system: a specification(version 1. 2. 1) [M]. [s. l.]; The Khronos Group Inc., 1999.
 - [7] WindML sdk programmer's guide[M]. USA: Wind River Systems, Inc., 2002.
 - [8] 曹桂平. VxWorks 设备驱动开发详解[M]. 北京:电子工业出版社,2011.
 - [9] 孔祥营. 嵌入式实时操作系统 VxWorks 及其开发环境 Tornado[M]. 北京:中国电力出版社,2002.
 - [10] 张宇坤,袁冬莉,黄鑫鑫,等. 基于 VxWorks 的网卡驱动程序开发[J]. 计算机技术与发展, 2013, 23(1): 18-20.
 - [11] Wind River, 王金刚,苏 琪,等. VxWorks BSP 开发人员指南[M]. 北京:清华大学出版社,2003.
 - [12] 程敬原. VxWorks 软件开发项目实例完全解析[M]. 北京:中国电力出版社,2005.
 - [13] Wind River. VxWorks programmer's guide[M]. USA: Wind River System Inc, 1997.
 - [14] Tornado user's guide(Windows Version 2. 2) [M]. USA: Wind River Systems Inc., 2002.
-
- (上接第 186 页)
 - 2012, 55(3): 4-4.
 - [2] Pearson S. Taking account of privacy when designing cloud computing services[C]//Proc of the 2009 ICSE workshop on software engineering challenges of cloud computing. Vancouver: IEEE, 2009: 44-52.
 - [3] 林清滢. 基于 Hadoop 的云计算模型[J]. 现代计算机: 上下旬, 2010(7): 114-116.
 - [4] 田秀霞,周耀君,毕忠勤,等. 基于 Hadoop 架构的分布式计算和存储技术及其应用[J]. 上海电力学院学报, 2011, 27(1): 70-74.
 - [5] 施 军,黄卫东. 物联网打造智能家居[J]. 中国电信业, 2010(12): 70-71.
 - [6] 俞文俊,凌志浩. 一种物联网智能家居系统的研究[J]. 自动化仪表, 2011, 32(8): 56-59.
 - [7] Porter G, Jolla L. Decoupling storage and computation in Hadoop with SuperDataNodes[J]. ACM SIGOPS Operating System Review, 2010, 44(2): 41-46.
 - [8] 施 游,张智勇. 云计算体系架构[J]. 电脑知识与技术: 学术交流, 2011, 7(1): 83-84.
 - [9] 宋 均,祝 林. 基于云计算的海量数据处理平台设计与实现[J]. 电讯技术, 2012, 52(4): 566-570.
 - [10] Lenk A, Klems M, Nimis J, et al. What's inside the cloud? an architectural map of the cloud landscape[C]//Proceedings of the 2009 ICSE workshop on software engineering challenges of cloud computing. Vancouver, BC: IEEE, 2009: 23-31.
 - [11] Dean J, Ghemawat S. MapReduce: simplified data processing on large clusters[J]. Communications of the ACM, 2008, 51(1): 107-113.
 - [12] Abouzeid A, Bajda-Pawlikowski K, Abadi D, et al. Hadoop-DB: an architectural hybrid of Map/Reduce and DBMS technologies for analytical workloads[C]//Proceedings of VLDB Endowment, 2009, 2(1): 922-933.
 - [13] Oldewurtel F, Mahonen P. Analysis of enhanced deployment models for sensor networks[C]//Proceedings of vehicular technology conference. [s. l.]: [s. n.], 2010: 23-24.


作者：

徐源吾，王珣，[XU Yuan-wu](#)，[WANG Xun](#)

作者单位：

[徐源吾, XU Yuan-wu\(南京邮电大学 通信与信息工程学院, 江苏 南京 210003; 南京邮电大学 通信技术研究所, 江苏 南京 210003\)](#)，[王珣, WANG Xun\(南京邮电大学 通信与信息工程学院, 江苏 南京 210003; 南京邮电大学 宽带无线通信与传感器技术教育部重点实验室, 江苏 南京 210003\)](#)

刊名：

[计算机技术与发展](#)

英文刊名：

[Computer Technology and Development](#)

年，卷(期)：

2014 (9)

本文链接：http://d.g.wanfangdata.com.cn/Periodical_wj fz201409042.aspx