

VOCs 监测预警云计算平台的设计与实现

张涛¹, 蒋晓濛², 焦正²

(1. 晋城职业技术学院, 山西 晋城 048026; 2. 上海大学 环境与化学工程学院, 上海 200444)

摘要: 挥发性有机污染物(volatile organic compounds, VOCs)的监测预警工作是一项复杂的系统工程, 也是当下环境科学领域研究的热点与挑战。在分析现有监测预警机制不足的基础上, 设计了基于云计算的 VOCs 监测预警云平台, 介绍了该平台的总体架构、构成要素以及主要模块, 分析了该平台的关键技术以及优势所在, 从而进一步证明了该平台的应用价值。

关键词: 挥发性有机污染物; 云计算; 监测预警; 负载均衡

中图分类号: X 84

文献标志码: A

文章编号: 1007-2861(2018)05-0791-09

Design and realization of the cloud computing platform for VOCs monitoring and early warning

ZHANG Tao¹, JIANG Xiaomeng², JIAO Zheng²

(1. Jincheng Institute of Technology, Jincheng 048026, Shanxi, China;

2. School of Environmental and Chemical Engineering, Shanghai University, Shanghai 200444, China)

Abstract: Monitoring and early warning of volatile organic compounds (VOCs) belongs to complicated system engineering. It is a research hotspot and a challenge in environmental science. In this paper, by analyzing shortcomings of existing mechanisms for monitoring and early warning, a VOCs monitoring and early warning cloud platform based on cloud computing is designed. The overall architecture of the system is introduced. The system components and main modules of the platform construction are illustrated. On this basis, key technologies and advantages of the platform are analyzed, showing its value in applications.

Key words: volatile organic compound; cloud computing; monitoring and early warning; load balancing

由于环境污染的日益加重, 国内外学者们已在大气扩散模式、污染气象学以及大气污染预报等领域陆续开展了多方位的研究^[1-2]。王海鲲等^[3]通过国际机动车排放模型的运用, 估算了上海市机动车的污染物排放量。陆涛^[4]、刘娟^[5]、李莉等^[6]在长三角区域也开展了多方位的研究。曹骝等^[7]创造性地将环保与云计算有机结合, 并在此基础上研究了基于大数据平台的PM_{2.5}监测预警系统。虽然我国的环境监测预警机制已逐步完善, 但仍然存在一些急需解决的问题。比如网络建设缺乏系统规划, 监测因子缺乏代表性, 监测质量和规范化有待完善, 信息系统有待统一建立和管理以及经费需求量庞大等^[8-9]。

收稿日期: 2016-12-07

基金项目: 国家自然科学基金资助项目(61174011)

通信作者: 焦正(1972—), 男, 教授, 博士生导师, 博士, 研究方向为环境功能材料。E-mail: zjiao@shu.edu.cn



借鉴欧美以及中国香港等众多发达国家和地区在建设空气质量监测网络方面的发展历程和先进经验^[10-16],在已有的监测站点基础上,结合大气污染的特征和现状,本工作设计了针对挥发性有机污染物(volatile organic compounds, VOCs)的监测与预警平台。本平台是基于云计算的大气质量监测预警云平台,通过前端大规模部署的特征因子监测设备和空气质量监测设备来监测空气质量,包括 $PM_{2.5}$, H_2S , NH_3 , NO_2 , SO_2 等无机气体、有机溶剂气体、可燃气体等存在于大气中的污染因子。本平台可通过空气质量云监测与预警平台(数据立方)实时感知破坏环境的行为,并通知执法人员第一时间进行阻止。本平台还能够通过海量的科学数据分析对比,得到反映局部区域(特别是重度污染区域重点监测)的相关参考值,从而反映特定区域的空气质量,也为该区域的环保工作提供一个科学的信息化支持。

1 平台设计和构成

考虑到 VOCs 监测预警云计算平台建设的发展需求,以实现统一管理、高效应用、平滑扩展为目标,以“先进、安全、成熟、开放、经济”为总体设计原则,设计了本监控平台。

1.1 总体架构

本平台的总体构架如图 1 所示。首先,通过前端特定的采集设备采集监测区域内空气的特征参数,经由 GPRS 等无线传输方式将参数传送到有公网 IP 地址的服务器上,对其进行初步处理。然后,将数据汇入数据立方并通过本平台进行存储和相应的计算。最后,将实时的预警信息以邮件或者短信等方式发送给执法人员。此外,本平台还可以结合前期建设的视频平台来监控各个监测点的污染情况,视频平台中的摄像终端也可以记录整个污染过程,再由视频服务器和 Web 服务器对数据进行最终的后处理,并向公众公布。如此就可以通过移动终端或者 Web 页面实时地了解监测区域内所有监测点的污染情况,并且对破坏环境的违法行为做到执法有据。此外,所有数据的云存储运用了数据流与控制流相互分离的技术,对数据的读取和存储在各个节点上同时进行,且随着存储节点的增加,体系的性能和容量也会呈现线性增长趋势。同时,由 JobKeeper 调度平台对不同需求的数据进行统一处理。

1.2 主要模块

本平台主要由传感器智能连接系统、风速风向仪、声光报警器,以及空气传感器系统通信等模块构成。

1.2.1 传感器智能连接系统

智能连接系统负责将传感器单元探测到的数据经由无线网络传送至本平台的 Internet 网络,是整个无线网络的基本组成单元。构成探测器系统的传感器单元通过简单的接口(如 RS485, CAN 等,支持多种连接技术)与数据集中器(包含了无线传输模块)相连,从而组成智能连接系统(见图 2)。传感器主要负责探测数据。本平台选用的传感器种类众多,包括 H_2S 气体传感器、 NH_3 气体传感器、 NO_2 气体传感器、 SO_2 气体传感器以及可燃气体传感器。数据集中器承担着询问并上传传感器探测到的数据的任务。这个过程的实现依靠两个关键点:①通过 RS485, CAN 等接口,可以实现数据集中器与传感器单元之间的通信;②数据集中器中包含无线模块,可以实现数据的无线传输。此外,数据集中器可以挂载几十个不同类型的传感器,可同时自动识别不同的传感器类型。由于各个传感器单元之间是相互独立的,因此若单个传感器单元出现了故障,不会对平台的正常运行造成影响,并且平台会检测出有故障的设备,并向上汇报。



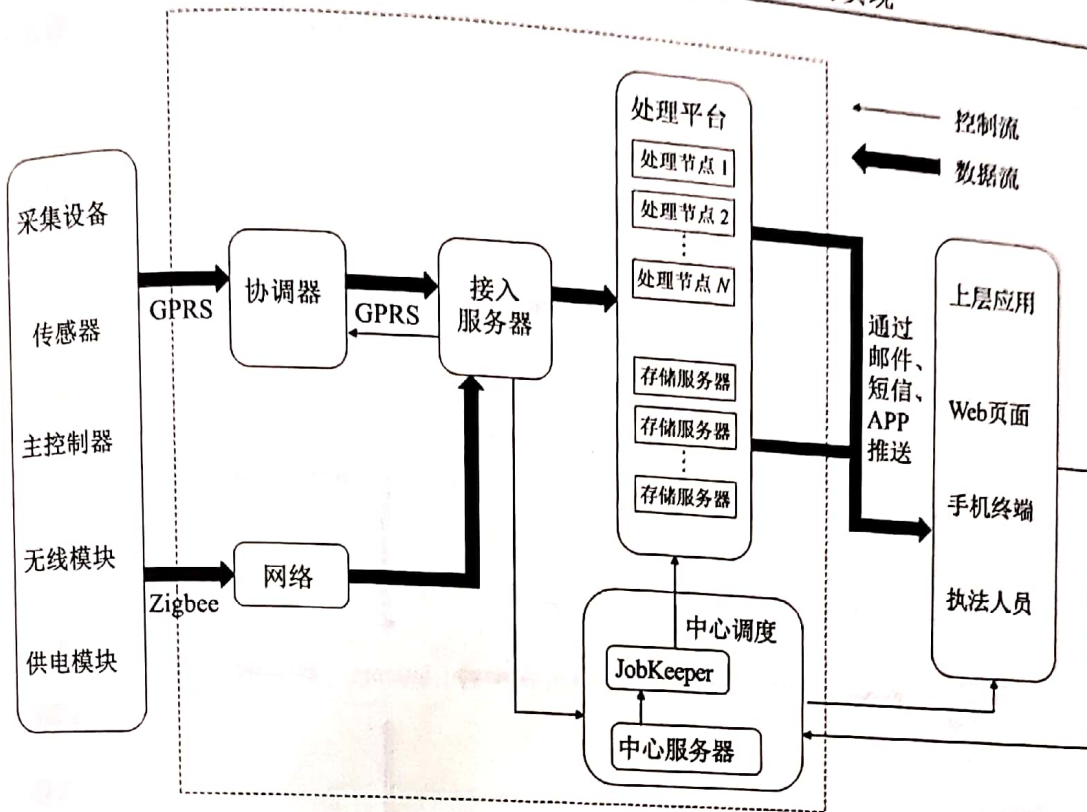


图1 总体架构

Fig. 1 Scheme of the overall structure

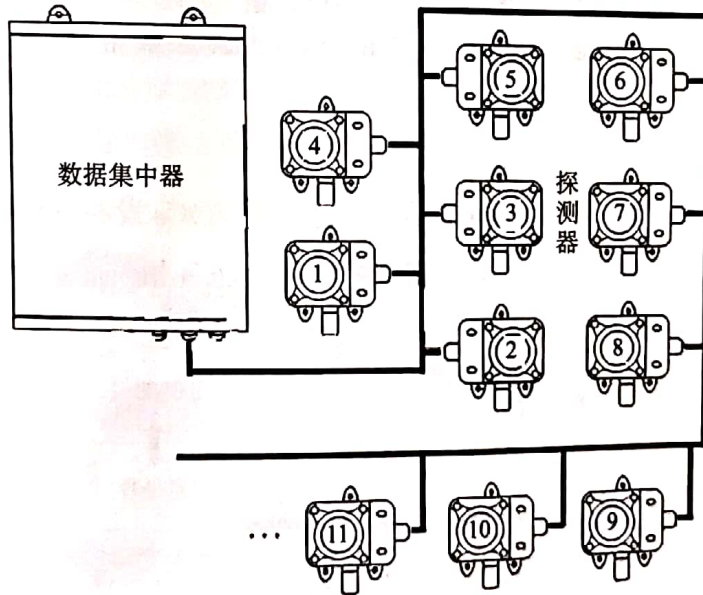


图2 智能连接系统

Fig. 2 Scheme of the connection system

1.2.2 风速风向仪和声光报警器

风速风向仪用于测量风向、风速的瞬间值和平均值,可为空气污染传播过程的演化与污染源的追踪提供实时气象数据。风速风向仪结构合理、自动化程度高,能连续对瞬时风向、风速进行监测和报警。此外,每个监测点又布署了一套声光报警器。当VOCs的质量浓度超标时,系统能够根据监测到的不同空气质量浓度发出不同的报警声音,也可以进行自定义的语音报警。

需要特别说明的是,声光报警器是隔爆型声音报警和光提示型报警的结合体,需要与隔爆

型气体探测器搭配使用。当目标气体的质量浓度大于气体探测器的设定值时, 探测器在本身报警的同时, 也会激发声光报警器, 发出高强度的声光报警信号。

1.2.3 空气传感器系统通信

空气传感器系统通信可以通过 GPRS 进行数据的无线传输, 也可以直接接入现有的专网网络。GPRS 数据传输示意图如图 3 所示。可见, 数据由前端发送设备依次经由 GPRS、移动网络、Internet 网络进行传输, 再被公共网络的服务器接受, 待数据入库并处理后, 通过 Web 服务器得以展现。前端设备也可以直接接入现有的电信专网。

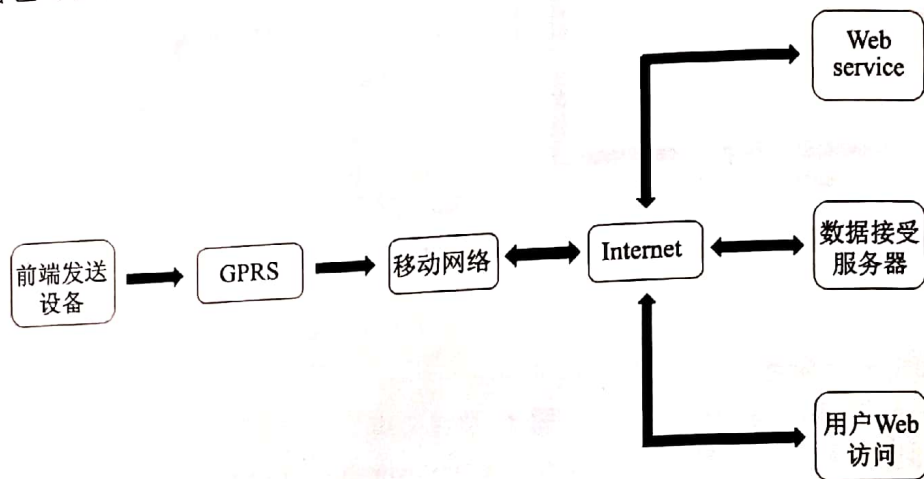


图 3 GPRS 数据传输

Fig. 3 Scheme of GPRS data transmission

2 平台组成架构

本平台主要由 VOCs 监测平台, VOCs 预警平台以及 VOCs 发布平台 3 部分组成(见图 4)。VOCs 监测平台的设计基于 B/S 结构, 可在实现对 VOCs 相关信息高效监测的同时, 将数

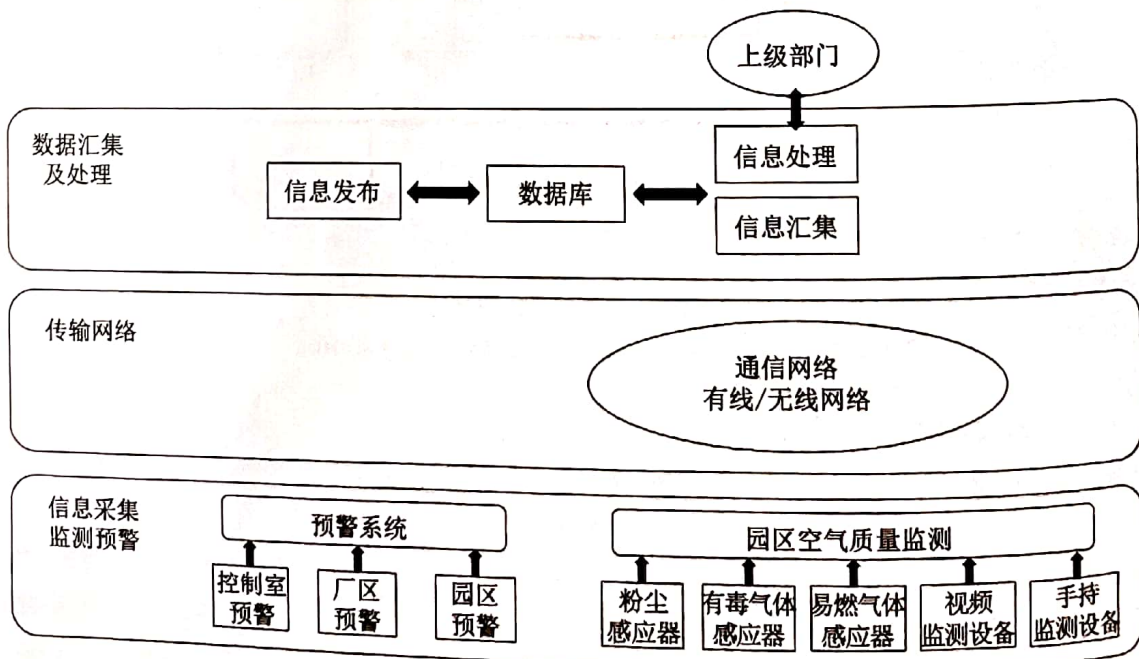


图 4 VOCs 监测与预警平台

Fig. 4 VOCs monitoring and early warning platform



据以 GPRS、移动网络等方式传送到系统内部并整编入库,并做出相应的分析处理. 监测设备采集到的信息将作为空气质量数据进行存储,为今后的环保工作提供真实准确的数据基础.

本平台充分利用了云计算的技术优势,解决了单台服务器不能建立大规模数据模型的问题,并且采用了国际上先进的 AERMOD 大气扩散模型以及高斯扩散算法,建立了科学的环境监控数据模型,使得任何区域点的空气质量数据值都可以按照该模型进行推算,有利于及时发现污染最严重的区域作为重点治理对象. 此外,本平台还可以查找出各种污染源相互影响的结果.

作为 VOCs 监测平台的有效补充, VOCs 预警平台承担了更多及时而有效的信息化支持. 预警平台是通过传感器体系和数据处理平台的分析结果启动的. 当 VOCs 监测平台的分析结果超过设定值时,便会自动开启预警平台,并且明确预警范围和级别(根据生产车间监测数据决定),同时启动相应的报警程序,并以 APP 推送、短信以及邮件等方式将其传递给工作人员.

最后,根据公众对于空气质量状况的需求,特别是雾霾非常严重的冬季, VOCs 发布平台可以向公众提供所在区域的实时空气质量状况.

3 关键技术

负载均衡技术是本平台进行云存储的一项关键技术. 该技术通过自动均衡各服务器的负载,提高数据处理集群和计算与存储集群的可靠性,进而将各存储节点的性能调整到最优.

3.1 数据处理集群的可靠性与负载均衡技术

本平台中不可避免地会出现一些单一故障点问题,此时负载均衡技术可以通过单点失效容错处理来解决这一问题,从而提高平台的可靠性.

3.1.1 负载均衡处理机的单点失效容错处理

负载均衡处理机的分布如图 5 所示. 首先,将全部的原始数据都导入交换机中,并将数据

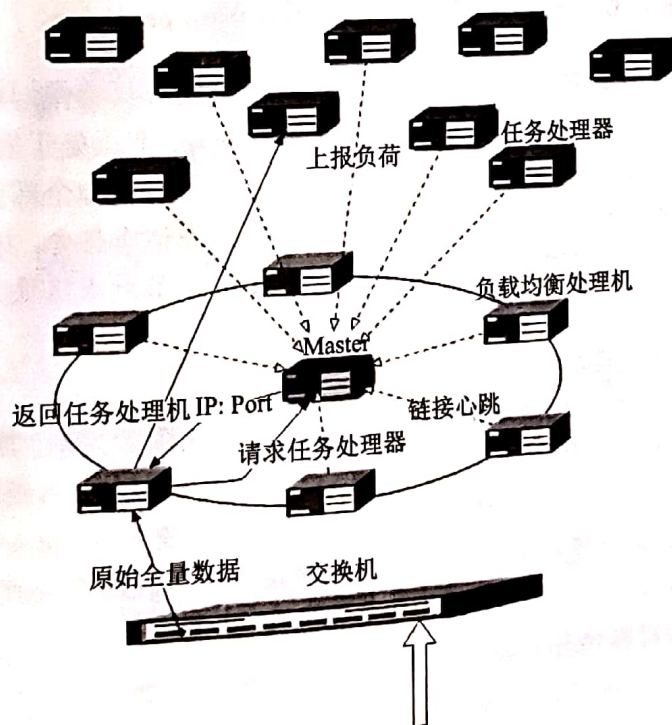


图 5 负载均衡处理机的分布图

Fig. 5 Distributions of the load balancing processor



发送给指向返回任务处理机 IP 地址的负载均衡处理机。然后, 负载均衡处理机通过 Master 节点向任务处理机发出请求指令, Master 节点返回任务处理机的 IP 端口信息。最后, 负载均衡处理机依据任务处理机的 IP 端口信息将数据发向任务处理机。

当某一台负载均衡处理机发生宕机时, 根据如图 6 所示的预案执行负载均衡操作, 具体步骤如下: ① Master 节点可以通过链接心跳获知负载相对均衡处理机处于宕机状态; ② Master 节点会挑选出负载较小的均衡处理机, 然后执行虚拟 IP 及其他所有的相关程序; ③ 通过虚拟 IP 的工作, 对提交的原始数据进行接收处理, 并在此基础上进行负载均衡处理。

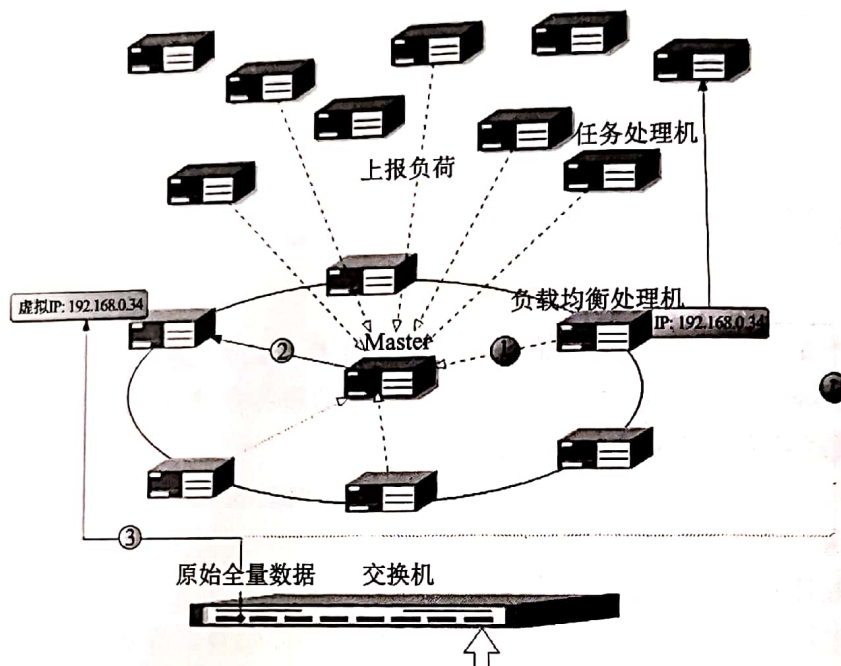


图 6 负载均衡处理机的宕机预案

Fig. 6 Crash plan of the load balancing processor

当 Master 节点宕机后, 按照如图 7 所示的预案执行负载均衡操作, 具体步骤如下: ① 存储计算节点以及负载均衡节点可以通过链接心跳获知 Master 节点处于宕机状态; ② 全部负载均衡节点会挑选出新的 Master 节点, 并且把新的 Master 节点告知全部存储计算节点, 然后以负载均衡 IP 为依据, 存储计算节点向新的 Master 节点发出请求指令, 并在此基础上向新的 Master 节点反映负载状态; ③ 全部工作任务都由新的 Master 节点来接受并处理, 同时对之前处于宕机状态的 Master 节点进行失效容错处理。

3.1.2 查询处理机的单点失效容错处理

当任务处理机宕机时, 查询处理机的单点失效容错处理如图 8 所示, 具体步骤如下: ① 负载均衡处理机得知任务处理机宕机后, 缓存接管数据, 同时 Master 节点得知任务处理宕机后, 将该任务处理机信息从待分配列表中删除; ② 负载均衡处理机向 Master 节点请求任务处理机, Master 节点为负载均衡处理机重新分配任务处理机; ③ 负载均衡处理机将原始数据发向新的任务处理机, 进行查询。

3.2 计算与存储集群的可靠性与负载均衡技术

3.2.1 Master 节点的单点失效容错处理

AvatarNode0 以 Primary 方式启动, 作为 NameNode 节点与用户交互。AvatarNode1 以 Standby 方式启动, 是一个处于 SafeMode 的 NameNode, 可以通过定期读取 AvatarNode0 的



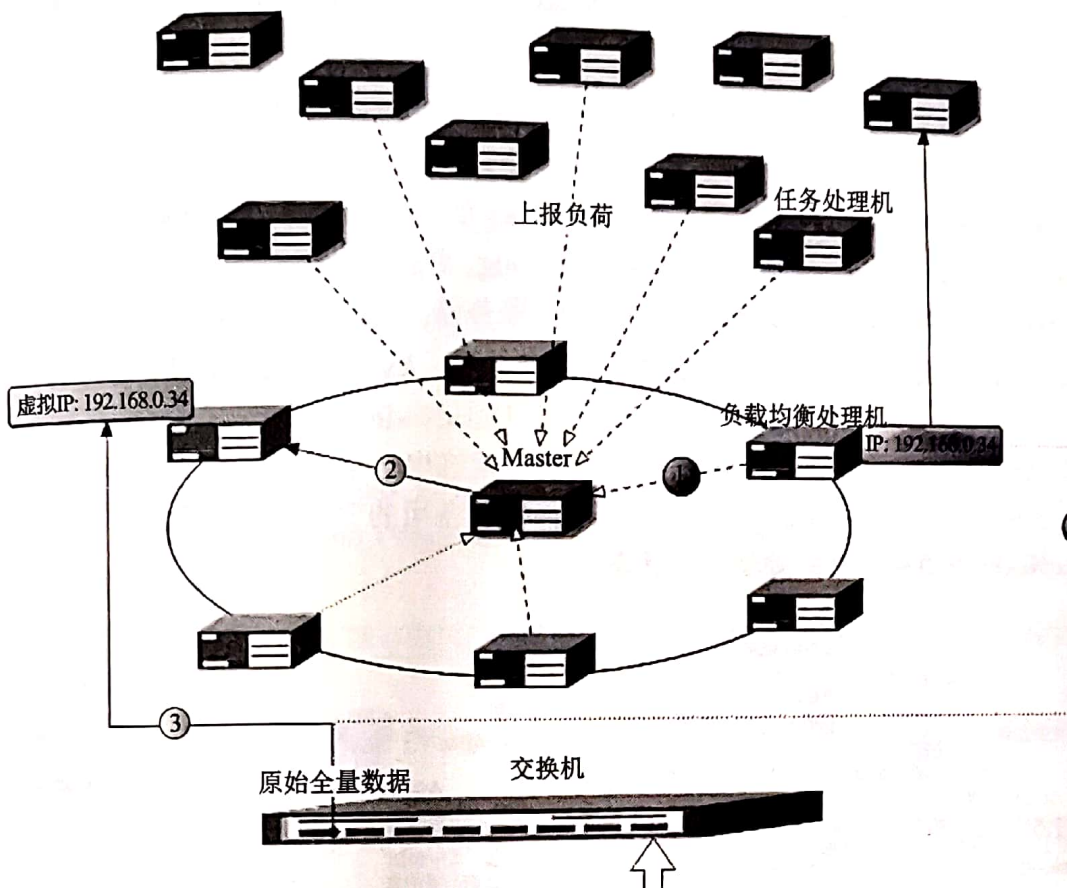


图7 Master节点的宕机预案
Fig.7 Crash plan of the master nodes

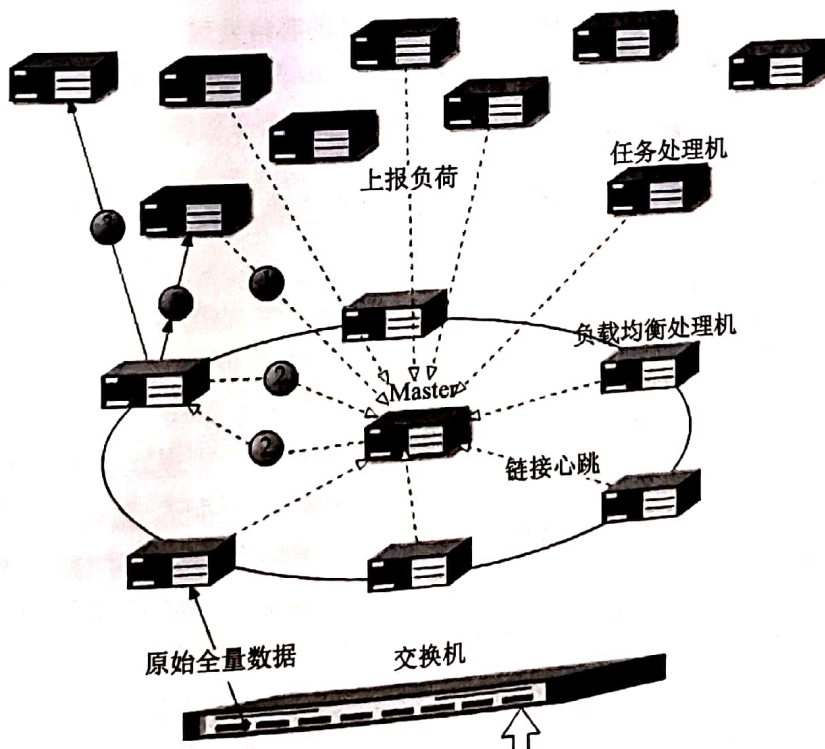


图8 查询处理机单点失效的容错处理
Fig.8 Fault tolerant of the processor query failure

日志来更新自己的元数据和内存, 并执行 CheckPoint, 同时更新 AvatarNode0 上的 Fsimage 以及 Editlog.

DataNode 向 AvatarNode0 和 AvatarNode1 同时发送心跳信息和 BlockReport, 其中包括 Block 的位置信息.

NFS 服务器存储 AvatarNode0 和 AvatarNode1 的 Fsimage 和 Editlog. AvatarNode 的 Primary 节点写入 Editlog, Standby 节点读取 Editlog, 更新内存中的元数据信息, 并且定期执行 CheckPoint, 将 Fsimage 及 Editlog 回写到 NFS 服务器.

Master 节点单点失效的容错处理如图 9 所示. 可见, AvatarNode0 以 Primary 方式启动, AvatarNode1 以 Standby 方式启动, 然后启动各个 DataNode; 当确定 AvatarNode0 发生故障后, 可以将 AvatarNode1 手动切换到 Primary, 同时挑选出一个备用的节点, 并以 Standby 方式启动, 如此即便 AvatarNode1 发生了故障, 也能够把备用的节点切换为 Primary 状态, 从而保证 NameNode 节点可以始终处于工作状态.

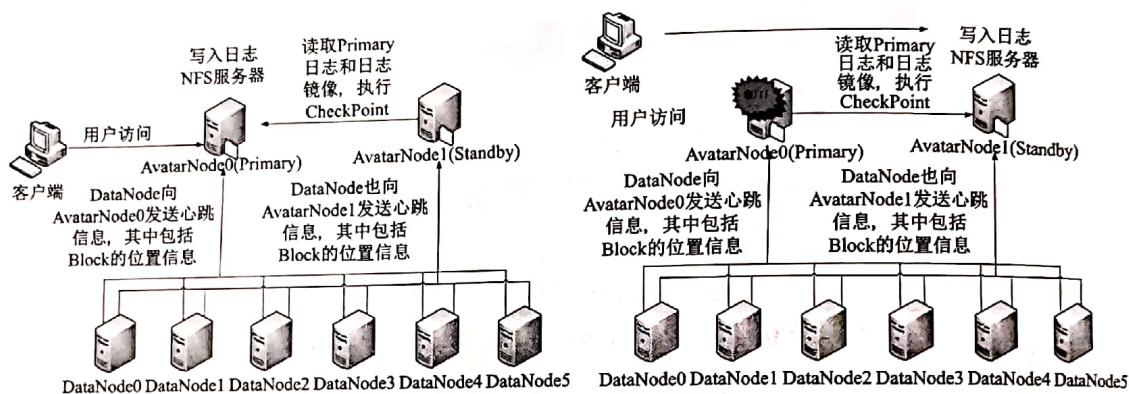


图 9 Master 节点单点失效的容错处理

Fig. 9 Fault tolerant of the master nodes query failure

3.2.2 计算与存储集群的负载均衡处理

本节所提到的“负载均衡”主要包括合理分配初始任务, 以及事后时刻监督、及时调整两个阶段的事务.

HDFS 分为 Replication Target Chooser 和 Balancer 两大类. 前者起着为数据块寻找“婆家”的作用. 总体上, 数据块的分配、备份数量、数据的服务器位置以及客户端地址之间有着紧密的联系, 而计算时只需考虑静态位置的信息, 以期将多重备份分配到不同的机器上(具体算法, 自行参见源码). 后者是对前者的补充. 因为后者是为了实现动态的负载调整而存在的, 其运作过程遵循 Client Protocol 和 NameNode Protocol 两个协议, 通过与主控服务器之间的信息传递, 获知单个数据服务器的负载情况, 从而实现及时而有效的调整. 此外, Balancer 会向相关的目标数据服务器发出一个 DataTransferProtocol.OP_REPLACE_BLOCK 消息. 当数据服务器接收到这个消息时, 就会写入新的数据块并且向主控服务器发出通知, 使其删除同一个数据服务器上最初写入的那个数据块. MapReduce 使用公平调度法进行负载平衡, 可以对多个用户指派任务, 其优势在于可以使全部任务在时间的推移下, 得到与其权值相应的共享资源. 如果只有一个任务在执行, 那么整个集群都会工作; 而一旦上传了其他任务, 平台就会给空闲的任务计算槽提供新任务, 从而使所有任务基本都能得到与权值相应的计算时间. 这个特性无疑带来一个好处, 即可以在合理时间内完成短任务的同时, 又可确保对长任务的完成质量.



4 结束语

作为基于云计算的 VOCs 监测预警云平台,本平台不需要昂贵的设备,只采用了包括特征因子监测器、各类传感器、数据集中器、风速风向仪以及声光报警器等在内的传统设备,在对分析海量科学数据的基础上,实现了在宏观和微观上对特定区域 VOCs 的监测预警功能。本平台使用的采集设备以及分析测试方法具有自动化程度高、测定速度快以及数据处理量大而准确的优点,在真正意义上实现了对 VOCs 污染事件的秒级响应,同时还可以与现有的监测点形成互补优势,从而保证了对所发布空气质量数据的参考价值。

参考文献:

- [1] ROBERTO S J, JUAN L P, JOSE L M. Simulations with MM5-CMAQ and WRF/CHEM models of a high elevated PM₁₀ and PM_{2.5} episode in Germany during winter (2003) [M]. NC: Chapel Hill, 2009: 1-6.
- [2] 韩志伟,杜世勇,雷孝恩,等. 城市空气污染数值预报模式系统及其应用[J]. 中国环境科学, 2002(22): 202-206.
- [3] 王海鲲,陈长虹,黄成,等. 应用 IVE 模型计算上海市机动车污染扬排放[J]. 环境科学学报, 2006, 26(1): 1-9.
- [4] 陆涛. 上海世博会长三角区域空气质量预警联动系统开发与应用[J]. 中国环境监测, 2013(1): 141-146.
- [5] 刘娟. 长三角区域环境空气质量预测预警体系建设的思考[J]. 中国环境监测, 2012(4): 135-140.
- [6] 李莉,陈长虹. 长江三角洲地区大气 O₃和 PM₁₀ 的区域污染特征模拟[J]. 环境科学, 2008, 29(1): 237-245.
- [7] 曹骝,吴修. 基于大数据平台的 PM_{2.5} 监测预警系统研究[J]. 互联网天地, 2015(4): 74-79.
- [8] 刘莹莹,乔丽英,刘振新. 构建先进的环境监测预警体系的探讨[J]. 资源节约与环保, 2014(4): 75.
- [9] WU Q Z, WANG Z F, CHEN H S, et al. An evaluation of air quality modeling over the Pearl River Delta during November 2006 [J]. Meteorology and Atmospheric Physics, 2012, 116: 113-132.
- [10] 王少毅,曾燕君,琚鸿,等. 广州市环境空气监测与预警体系在亚运会中的应用[J]. 广州环境科学, 2011(2): 1-3.
- [11] 黄爱平. 化工园区构建大气污染事故预防预警体系的思考[J]. 化工管理, 2016, 29: 302.
- [12] 曾定章. 城市环境 VOCs 连续监测技术探讨[J]. 仪器仪表用户, 2016(7): 13-14.
- [13] 刘俐,李继罡,叶智,等. 德阳市全域环境监测网络现状及建设思考[J]. 四川环境, 2012(5): 30-35.
- [14] 张宇波. 上海空气质量自动监测数据管理平台设计与开发[D]. 上海: 复旦大学, 2012.
- [15] 许亚宣,李小敏,于华通,等. 邯郸市大气复合污染特征的监测研究[J]. 环境科学学报, 2015, 35(9): 2710-2722.
- [16] JE C H, STONE R, OBERG S G. Development and application of a multi-channel monitoring system for near real-time VOC measurement in a hazardous waste management facility [J]. Science of the Total Environment, 2007, 382(2/3): 364-374.

本文彩色版可登陆本刊网站查询: <http://www.journal.shu.edu.cn>

