

# 一种基于 MongoDB 和 Hadoop 的海量非结构

## 化物联网数据处理方案

杨鹏 1,2, 林俊晖 1,3

(1 重庆邮电大学 通信与信息工程学院, 重庆 400065; 2 中国信息通信研究院, 北京 100191; 3 重庆高校市级光通信与网络重点实验室, 重庆 400065)

摘要: 随着物联网数据种类的增多和数据规模的增大, 对物联网数据的存储和计算提出了新的挑战, 为了应对海量非结构化物联网数据的存储和计算要求, 提出了一种 NoSQL 数据库技术与 MapReduce 编程框架相结合的方案. 使用典型的 NoSQL 数据库 MongoDB 作为主数据库来存储海量非结构化的物联网数据, 使用 Hadoop MapReduce 作为对物联网数据分析处理的计算框架. 通过对 MongoDB 集群和 Hadoop 集群的重叠部署, 降低了计算时数据传输的开销, 构建了一套高可用、高性能的物联网大数据处理平台. 通过使用该方案对海量非结构化物联网数据的处理分析实验验证了该方案的高可用性及其高效性.

关键词: 物联网; 大数据; NoSQL; Hadoop 平台; MongoDB 数据库

## A Scheme for Massive Unstructured Iot Data Processing Based on MongoDB and Hadoop

YANG Peng<sup>1,2</sup>, LIN Jun-hui<sup>1,3</sup>

(1 School of Communication and Information Engineering, Chongqing University of Post and Telecommunications, Chongqing 400065, China; 2 China Academy of Information and Communications Technology, Beijing 100191, China; 3 Key Laboratory of Optical Communication and Networks, Chongqing 400065, China)

Abstract: As the IoT data type and data volume increase, a new challenge is proposed to the storage and calculation of IoT data. To meet the storage and calculation demands of the massive unstructured IoT data, a scheme that NoSQL database technology combined with MapReduce programming framework has been presented in this paper. Using the MongoDB, a typical database of NoSQL, as the main database to store massive unstructured IoT data. At the same time, the Hadoop MapReduce was adopted as the calculation framework for IoT data analysis and processing. Through the overlapping deployment of MongoDB cluster and Hadoop cluster, the cost of data transmission is reduced, and a set of high availability and high performance large data processing platform is established. The analysis of the experiment that using this scheme to deal with massive IoT data has verified the high efficiency and high efficiency of this method.

Key words: IoT; Big Data; NoSQL; Hadoop platform; MongoDB database

作者简介:

杨鹏男, (1980-), 博士, 高级工程师. 研究方向为下一代移动通信技术、无线泛在网络、下一代互联网理论与技术.

林俊晖 (通讯作者) 男, (1992-), 硕士研究生. 研究方向为无线泛在网络、下一代互联网理论与技术、物联网大数据处理. E-mail: 297876621@qq.com.