一种多层次 Hadoop 平台设计

李兆兴，马自堂
（解放军信息工程大学 密码工程学院，河南 郑州 450000）

摘 要：大数据和大数据中心规模的进一步扩大为 Hadoop 平台的可扩展性提出了新的要求．在系统地分析了制约 Hadoop 平台性能和可扩展性因素的基础上，提出了一种多层次 Hadoop 平台．通过将 Hadoop 平台划分为若干区域，形成全局-区域-节点的多层次结构．由 Master 节点负责系统级的元数据管理和任务分发，区域管理节点负责区域级的元数据管理和任务分发，从而提高 Hadoop 平台可扩展性和容错性．实验表明，该平台可有效提高 Hadoop 的可扩展性，具有一定的可行性．

关键词：Hadoop；NameNode；SeMNode；多层次
中图分类号：TP391 **文献标识码**：A **文章编号**：1000－7180(2016)02－0027－04

A Multi-Layer Hadoop Platform

LI Zhao-xing, MA Zi-tang
(College of Crypgraphy Engineering, PLA Information Engineering University, Zhengzhou 450000, China)

Abstract: With the further expansion of Big Data and Big Data Centre, new extensibility requirements are proposed on the Hadoop Platform. In the base of systematically analysis constraints on performance and extensibility of Hadoop Platform, this paper presents a multi-level Hadoop Platform which has a structure of Global-Region-Node after being divided. In this platform, Master Node is responsible for metadata management and task development on the global level, and regional management node takes charge of metadata management and task development on the regional level, which improve both scalability and fault tolerance of Hadoop Platform. Experiments show that this platform can effectively promote the scalability of Hadoop and has certain feasibility.

Key words: Hadoop; NameNode; SeMNode; multi-layer

作者简介：

李兆兴 男，(1990-)，博士研究生．研究为大数据、数据检索、数据挖掘．E-mail：hoops@whu.edu.cn

马自堂 男，(1962-)，教授．研究方向为大数据、云计算、信息．