

DOI: 10.3963/j.issn.1671-4431.2009.18.028

# 基于对象存储系统的负载均衡模型分析

刘 群<sup>1</sup>, 冯 丹<sup>2</sup>, 李 坚<sup>2</sup>

(1. 华中科技大学网络与计算中心, 武汉 430074; 2. 华中科技大学计算机学院, 武汉 430074)

**摘 要:** 基于对象存储系统吸收现有的网络存储系统优点, 提供对象接口, 其属性描述了对对象的特征与信息, 并存于对象存储设备中。随着对象存储设备的数目增多, 在不同的时间、空间和状态等条件下, 基于对象存储系统日益复杂化。元胞自动机模型是一种描述和模拟复杂系统行为有效的方法, 以对象存储设备设为元胞, 每个元胞只取读或写状态, 并局部状态规则的改变, 这样构成基于存储对象的负载均衡模型, 演绎着基于对象存储系统变化过程。

**关键词:** 元胞自动机; 对象存储系统; 负载均衡; 模型分析

中图分类号: TP 333

文献标识码: A

文章编号: 1671-4431(2009)18-0105-03

## A Model Analysis for Load Balancing Based on Object-Based Storage System

LIU Qun<sup>1</sup>, FENG Dan<sup>2</sup>, LI Jian<sup>2</sup>

(1. Network and Computing Center, Huazhong University of Science and Technology, Wuhan 430074, China;

2. School of Computer, Huazhong University of Science and Technology, Wuhan 430074, China)

**Abstract:** The Object Based Storage System (OBSS) brings together the benefits of existing network storage systems. It provides a new command set called object, which includes attributes describing characteristics of the data and storing in the Object Storage Device (OSD). With the number of OSDs increasing, the OBSS is complicating in the different time, space and state. The cellular automaton is an effective method that draws and simulates a complex system. As a cellular, the OSD only has reading or writing state that locally changes. So it forms a load balancing model based on OSDs, and deduces the process of OBSS.

**Key words:** cellular automata; object-based storage system; load balancing; model analysis

随着数字化时代的到来,人们越来越注重数据存储,但不只是简单地将它们存放在物理介质上。理想的存储体系结构应具有安全性、可跨平台数据共享、具有高性能以及可扩展性。传统的网络存储有直连存储(Direct Attached Storage, DAS)、存储区域网(Storage Area Network, SAN)和附网存储(Network Attached Storage, NAS) 3种。

由于 DAS、SAN 和 NAS 均存在不同程度的缺陷, 随着源源不断增加的数据和不断增加的物理设备对这些存储体系架构提出了巨大挑战, 存储设备由相对愚钝、外部管理转变为智能、自主管理, 并知晓它们所服务的存储应用, 这样就需要一个具有表达能力的接口, 催生了基于对象存储(Object Based Storage, OBS)<sup>[1, 2]</sup>的产生, OBS 吸收 NAS 和 SAN 优点, 具有 SAN 的高性能和 NAS 的数据共享和安全性, 采用对象接口, 对象包含任何类型的数据, 如文件、数据库记录、图像以及多媒体视频音频等, 并可动态地扩大和缩小。对象还具有属

收稿日期: 2009-05-07.

基金项目: 国家重点基础研究发展规划 973 项目(2004CB318201)和国家自然科学基金项目(60873028)。

作者简介:刘群(1969),女,博士,工程师, E-mail: liliuqun@mail.bust.edu.cn

性,用于描述对象的特征与信息<sup>[3]</sup>。把保存对象的存储设备称为对象存储设备(Object-based Storage Device, OSD)。OSD是一个智能设备,维护所有与数据空间分配、空闲空间等有关的元数据管理,并有能力承担更多的智能处理,从而提高了存储系统的智能性和可管理性。

纵观目前与网络存储相关研究,主要采用马尔科夫过程(Markov Process)和Petri网数学模型。马尔科夫过程只是时间离散、状态离散的动力学模型<sup>[4]</sup>,各个时刻的状态的转变是由一个预先设定好的状态转移概率矩阵控制;Petri网用于描述和分析并发、分布式系统的数学模型工具<sup>[5]</sup>,没有考虑在空间位置之间相互关系,它们无法完整地阐述OBSS持久运行后的负载均衡现象,因此需要采用复杂系统的理论方法对OBSS的动态行为特征进行探索。元胞自动机模型<sup>[6]</sup>就是这种描述、认识和模拟复杂系统行为强有力的方法。因此,文中提出了基于对象存储系统的元胞自动机模型,对负载均衡进行分析与研究。

## 1 元胞自动机

元胞自动机最早由数学家Stanislaw M. Ulam和Von Neumann于20世纪50年代提出,表现为一个时空离散的动力系统,其散布于规则网格中的每一元胞均取有限的离散状态,各元胞遵循相同的演化规则进行同步演化,且仅和它相邻的元胞发生相互作用,大量元胞通过简单的相互作用而构成动态系统的演化过程。因此,元胞、网格、邻居和规则构成元胞自动机的重要组成部分。

元胞是元胞自动机的最基本的组成部分。元胞分布在离散的一维、二维或多维几何空间的晶格点上。状态可用 $\{0, 1\}$ 的二进制形式,或是 $\{s_0, s_1, \dots, s_i, \dots, s_k\}$ 整数形式的离散集,元胞自动机中元胞有一个或多个状态变量。

网格是指元胞所分布的空间网点集合。在理论上,网格可以是任意维数的几何空间规则划分,但目前研究集中在二维和三维元胞自动机上。

元胞自动机还须定义一定的邻居规则,明确哪些元胞属于该元胞的邻居。在一维中,元胞自动机通常以半径来确定邻居,距离一个元胞内的所有元胞均被认为是该元胞的邻居。

规则是根据元胞当前状态及其邻居状况确定下一时刻该元胞状态的动力学函数。

元胞自动机是由一个元胞空间和定义于该空间的变换函数所组成。

从20世纪80年代后期起,元胞自动机在许多领域都得到广泛的应用,成为研究复杂系统、复杂性的有效途径之一。

## 2 基于OSD的负载模型

假设OBSS中为同构OSD,一个元胞最多只能为一个OSD,由OSD的属性记载OSD的负载情况,当系统中有请求时,则考虑每一个OSD当前负载情况而进行分配最合适的OSD,可用一维元胞自动机对此进行描述。

元胞空间:元胞空间代表系统的存储空间,即存储对象的个数。采用周期型边界条件,即元胞空间的首位相接,构成一个拓扑环。

元胞状态:元胞状态表示OSD在系统某个时刻的状态,每一个元胞有2种状态:0表示无负载未使用,1表示有负载正在使用。

邻居:邻居代表存储环境,假设下一个请求根据OSD当前负载决定是否被使用,定义邻居半径为1,即左右一个元胞作为中心元胞的邻居。

规则:在元胞空间中,OSD为系统所共享(假设分配存储时OSD为独占),当有请求来临时,先分配后转化状态,即

$$\begin{aligned} \text{若 } S_i^t = 1, S_i^{t+1} &= 0 \\ \text{若 } S_i^t = 0, S_i^{t+1} &= \begin{cases} 1 & n \neq 0 \\ 0 & n = 0 \end{cases} \end{aligned}$$

其中,0表示未使用状态,1表示使用状态, $S_i^{t+1}$ 表示 $t+1$ 时刻 $i$ 中心元胞的状态, $n$ 为两个相邻元胞中处于使用状态的个数。由于3邻居有着8种可能组合状态:111、110、101、100、011、010、001、000,根据上述规则,在下一时刻中心元胞的取值方法如下:

$S_{i-1}^t S_i^t S_{i+1}^t$	111	110	101	100	011	010	001	000
$S_i^{t+1}$	0	0	1	1	0	0	1	0

这个模型恰好是 Wolfram 初等元胞自动机的 50( 00110000) 号元胞自动机。这样就构建了一个基本的一维 OBSS, 图 1 为状态变迁图。

图 2 显示这个模型随着时间的演化, 元胞空间为 32, 初始在第 16 个元胞的状态取 1, 其他取 0, 经过规则演化 16 步的一个过程。当此模型在随机初始状态时, 经过随着十几步的演化, 元胞空间趋于一个空间平稳、均衡的构形(如图 3 所示), 可以说明按照此规则进行分配, 系统负载逐步均衡了。

由此可见, 这种基于元胞自动机模型的负载分析是对 OBSS 一个高度概括, 将 OSD 视为元胞, 并赋予了一定意义的实际含义, 模

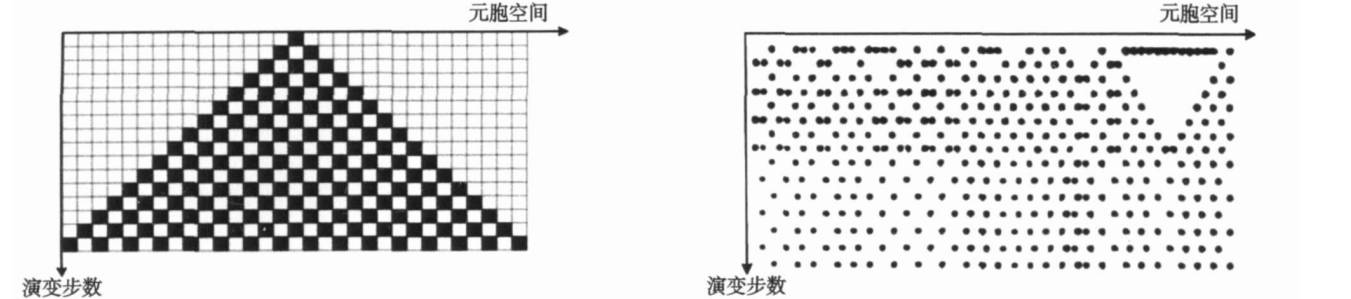


图2 OBSS的负载分配元胞自动机模型的演化

拟出 OBSS 一个简单的负载分配的动态变化。显然, 正是由于 OSD 为整个系统所共享, 并按照某种规则主动、并行地工作, OBSS 逐步达到一个负载平衡的状态。

3 结 语

元胞自动机是一个时间、空间和状态离散的动力系统, 在规则格网中的每一元胞取有限的离散状态, 遵循相同的作用规则, 依据确定的局部规则作同步更新, 大量的元胞通过简单的相互作用而构成了动态系统的演化过程。

基于存储对象的负载分配模型是一个一维元胞自动机模型, 它将 OSD 赋予为元胞, 按照某种规则主动、并行地工作, 模拟了一个简单的负载均衡的动态变化, 高度概括了 OBSS 的演变过程, 采用元胞自动机以期成为存储领域中一种新的分析问题方法, 探寻网络存储系统的演化规律。

参考文献

[ 1 ] Mesnier M, Ganger G R, Riedel E. Object based Storage [ J ] . Communications Magazine, IEEE, 2003, 41: 84 90.

[ 2 ] Mesnier M, Ganger G R, Riedel E. Object based Storage: Pushing More Functionality Into Storage [ J ] . Potentials, IEEE, 2005, 24 ( 2 ) : 31 34.

[ 3 ] Qun Liu, Dan Feng, Qin Ling jun, et al. A Framework for Accessing General Object Storage[ C ] // Proceedings of the 2006 International Workshop on Networking, Architecture, and Storages ( IWNAS2006 ) . Dalian, China: IEEE, 2006. 145 148.

[ 4 ] Ribeiro do Val, J. B. Felix Salles, J. L. Analysis and Algorithm for the Stochastic Control Problem of a Production/ Storage System[ C ] . Proceedings of the 30th IEEE Conference on 11- 13 Dec. 1991 Page( s ) . 477 482.

[ 5 ] Qun Liu, Dan Feng. An Approximate Analytic Performance Model of Object Based Storage[ C ] // Proceedings of the International Conference on Computational Science and Its Applications, Glasgow, UK, May 8 11, 2006. LNCS 3980. 1045 1052.

[ 6 ] 李才伟. 元胞自动机及复杂系统的时空演化模拟[ D ] . 武汉: 华中科技大学, 1997.