

一种成本相关的云提供商数据可靠性保证算法*

钟睿明^{1,3}, 刘川意^{2,3}, 王春露^{1,3}, 项菲^{1,3}

¹(北京邮电大学 计算机学院, 北京 100876)

²(北京邮电大学 软件学院, 北京 100876)

³(可信分布式计算与服务教育部重点实验室(北京邮电大学), 北京 100876)

通讯作者: 刘川意, E-mail: cy-liu04@mails.tsinghua.edu.cn

摘要: 数据可靠性保证和容灾成本控制对云提供商而言是一个相互矛盾的问题. 在分析已有数据保障机制的基础上, 设计了一个基于多个云平台的分布式富云容灾模型, 利用富云容灾系统, 私有云提供商可以借用其他云平台的虚拟资源对自身数据进行冗余备份. 为了减少数据传输的响应时间, 富云容灾模型通过设置多个地理位置隔离的富云代理实现云平台用户任务的分配, 减少私有云平台的工作负荷. 针对富云容灾系统的成本优化和数据可靠性保证问题, 提出了一种成本相关的云计算服务数据可靠性保证算法 CAHRPA. 该算法根据数据传输带宽和容灾费用在多个云平台中动态选择数据副本的存放位置, 从而以一种成本优化的方式为云提供商提供数据容灾方案. 实验结果表明, CAHRPA 能够在保证数据可靠性的同时, 实现一种成本更低的数据容灾策略.

关键词: 云计算; 数据可靠性; 容灾; 差分进化; 云存储

中图法分类号: TP306

中文引用格式: 钟睿明, 刘川意, 王春露, 项菲. 一种成本相关的云提供商数据可靠性保证算法. 软件学报, 2014, 25(8): 1874-1886. <http://www.jos.org.cn/1000-9825/4498.htm>

英文引用格式: Zhong RM, Liu CY, Wang CL, Xiang F. Cost-Aware data reliability provision algorithm for the cloud providers. Ruan Jian Xue Bao/Journal of Software, 2014, 25(8): 1874-1886 (in Chinese). <http://www.jos.org.cn/1000-9825/4498.htm>

Cost-Aware Data Reliability Provision Algorithm for the Cloud Providers

ZHONG Rui-Ming^{1,3}, LIU Chuan-Yi^{2,3}, WANG Chun-Lu^{1,3}, XIANG Fei^{1,3}

¹(School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China)

²(School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China)

³(Key Laboratory of Trustworthy Distributed Computing and Service of Ministry of Education (Beijing University of Posts and Telecommunications), Beijing 100876, China)

Corresponding author: LIU Chuan-Yi, E-mail: cy-liu04@mails.tsinghua.edu.cn

Abstract: Data availability is a desirable feature for all cloud providers. However, adoption of disaster recovery will bring incremental cost due to the continued investment of hardware. This paper illustrates the existing assurance mechanisms for data reliability and points out that the data availability and disaster recovery cost are always two incompatible goals for cloud providers. A rich cloud based disaster recovery (DR) model is designed for cloud providers. In this model, multiple rich cloud proxies are employed in order to reduce the response time and improve the quality of service. With help of this system, a cloud provider is able to replicate data with lower cost by utilizing the virtual resource from other clouds. In order to minimize the DR cost while guaranteeing the data availability, a cost aware high data reliability provision algorithm (CAHRPA) is proposed to tackle such an optimization problem. In the end, the feasibility and efficiency of this CAHRPA is verified by the comparison with some other replication strategies.

Key words: cloud computing; data reliability; disaster recovery; differential evolution; cloud storage

* 基金项目: 国家自然科学基金(60973009); 国家重大科技专项(2012ZX03005010)

收稿时间: 2012-07-07; 修改时间: 2013-03-06; 定稿时间: 2013-09-02

随着信息技术和网络存储技术的不断发展,云计算作为一种新型的网络服务模式,已经成为产业界和学术界的关注的焦点^[1]。但是,任何以互联网为基础的应用都存在一定的危险性,一旦云提供商的存储节点发生硬件故障(人为或自然灾害),存储在云端的大量用户信息就将面临数据丢失的威胁,这会给云提供商带来无法弥补的损失。2009年2月和7月,亚马逊的“简单存储服务(simple storage service,简称S3)”两次中断,导致依赖于网络单一存储服务的网站被迫瘫痪^[2]。2009年10月,微软数据中心宕机,导致了大量T-Mobile用户的数据丢失。数据可靠性保证,已经受到云提供商越来越多的重视。如何为数据提供容灾(disaster recovery,简称DR)保护,是一个云提供商关心的热点问题。

为了提高对灾难事件的抵御能力,目前的云提供商如亚马逊、google都默认采用3-Replicas的数据冗余机制确保数据的可用性^[3]。通过在异地建立自身的数据容灾中心,存储系统可以利用数据的冗余性和地理分散性实现对1份或者多份数据备份的实时维护^[4]。但是,大量的数据冗余存储会给企业带来大量的费用开销。随着云计算技术的不断发展,利用公有云资源为自身私有云平台提供更低成本的数据容灾服务(即,容灾云),已经成为云提供商的一个优先选择。相对于“build your own”方式进行数据容灾,Wood等人^[5]以Data Warehouse的实际容灾费用为例,证明了利用云资源进行数据容灾可以节省11%的容灾费用。但是,利用单一云提供商的资源进行数据容灾会带来数据丢失^[2]和云提供商锁定的问题^[6]。因此,如何利用多个云平台的资源提供更可靠、成本更低的存储服务,已经成为是容灾云面临的一个新问题。

利用公有云资源为自身私有云平台数据进行容灾,已经成为一种新的商业模式^[5],其关键在于如何降低数据容灾的使用成本。影响容灾费用的最主要因素是数据的存储量、传输带宽和存储时间。当前的云提供商,如亚马逊、谷歌、微软等,根据CPU、内存、传输带宽和存储容量等参数的不同,提供了多种可选模式及收费标准^[3],采用不同的云提供商资源执行任务会带来不同的费用开销。当接到大量用户任务请求时,为了保证多备份的要求,私有云平台需要将用户的数据副本存放到多个云存储平台,这种基于多任务多平台的资源分配问题是一个典型的NP困难问题^[7]。为了降低容灾费用,云提供商目前亟需一种成本相关的数据容灾策略。

本文从云提供商的角度出发,设计了一个同时包含私有云和公有云的数据容灾模型。私有云可以以一种pay-as-you-go的方式租用其他云平台的资源为自身数据进行容灾。我们将这种基于多个云平台的数据容灾方式称为富云容灾模式。在富云容灾系统中,云平台的数据容灾服务不再局限于某个云提供商自身的数据中心,私有云可以根据自身资源状况和任务类型的不同,将数据备份存放到自身的容灾中心或是其他云平台的存储资源。为了减少容灾系统的响应时间,避免大规模数据存储的性能瓶颈,本文设计了一个基于多代理机制的分布式富云容灾系统,并以成本优化为目标,提出了一种基于富云模型的数据容灾方案:一种成本相关的云提供商数据可靠性保证算法CAHRPA(cost aware high reliability provision algorithm)。CAHRPA通过分析用户任务、云平台的存储费用及传输带宽等条件,利用离散化差分进化算法的自动寻优能力找到一种成本优化的数据容灾方案,实现了多个云平台资源的优化配置,可以在保证传输带宽和数据可靠性的同时降低云提供商的数据容灾费用。

本文的主要贡献在于:

- (1) 针对云提供商的数据可靠性保证问题,设计了一种基于多代理机制的分布式富云容灾模型,该模型可以实现多个云平台之间的资源互用和相互容灾。针对富云容灾系统的成本优化问题构建了的数学模型,并对云提供商的容灾费用进行了量化分析。
- (2) 针对云容灾模型下的数据备份策略选择这一离散约束优化问题,我们基于离散化差分进化算法,提出了一种成本相关的云平台数据容灾策略,并依据约束条件设计了约束检验操作。该方法可以在保证节点平均传输带宽的前提下实现对数据容灾成本的控制,使得云平台可以在保证数据恢复时间的同时,实现一种成本优化的数据容灾服务。

本文第1节整理分析主要相关工作的情况。第2节介绍利用云资源进行数据容灾的系统模型,分析数据容灾策略需要考虑的各种因素,并给出相关的符号定义。第3节详细阐述基本流程和关键步骤,并给出其形式化算法。第4节通过实验将本文策略与其他数据备份策略进行比较,并对实验结果进行详细分析。第5节对全文工作进行总结。

1 相关工作

1.1 容灾云

云存储技术的不断发展,为容灾服务带来了崭新的解决思路 and 模式.基于已有的云服务模式,Wood 等人^[4]利用云计算的虚拟平台为企业及个人提供数据容灾服务,提出了容灾即服务(DR as a service)的服务模式,明确了云环境下服务可靠性保证的 5 个指标,指出数据恢复时间 RPO 主要与网络节点的传输带宽相关,并以实例证明了采用云资源进行数据备份可以大量节省企业容灾的开销.文献[8]进一步指出:当私有云资源被完全占用时,利用公有云资源执行用户任务可以明显减少任务的执行成本.

1.2 资源分配

随着云计算技术的不断发展,云平台的工作能耗和使用成本问题已经受到了云提供商越来越多的关注.目前,针对云环境的资源分配和任务调度问题主要有 3 类解决方法,分别是基于规则(policy based)的调度策略^[7,9]、基于最优化理论的调度策略^[10-12]和基于进化算法的调度策略^[13-15].

(1) 基于规则的调度策略

针对云环境下的资源分配问题,文献[7]设计了一种基于规则的调度策略 LSTR.LSTR 以最大化云平台的资源利用率为目标,对所有任务进行资源分配.LSTR 虽然可以简单、高效地实现云平台资源的合理分配,但是 LSTR 不是一个实时分配策略,无法满足云平台实时性的要求.

文献[9]从经济学原理的角度提出了云计算经济学架构,设计了基于 SLA 的云资源管理经济模型.该策略为云消费者和供应商提供了有关经济激励的反馈,提高了资源利用率,有助于实现云环境下资源的高效管理,优化配置,可最大限度地满足用户服务质量需求.

(2) 基于最优化理论的调度策略

文献[10]考虑利用多个云平台的虚拟机资源完成用户任务,以虚拟机使用的总成本为优化目标构建了基于多多云平台的虚拟机布局问题模型,采用随机整数规划对问题进行了分析.文献[11]针对混合云环境中计算型任务的分配及费用优化问题进行了分析,文中考虑所有计算任务都带有 deadline 限制,并以总成本最低为目标进行建模和计算,采用整数规划的方法对问题进行求解,得到了一种成本最优的任务分配方案.但是当用户任务数量较大时,计算最优方案需要极大的时间消耗,难以应用于实际的云计算系统.

文献[12]利用博弈论的方法来解决云计算资源分配问题.文中设计了一种基于博弈论的资源分配算法,利用整数规划方法处理单个参与者的独立优化问题.这种基于博弈论的资源分配方法重点是针对一些可分任务调度问题的优化,主要适用于处理一些非常复杂和动态的、应用能分成多个协作任务的资源调度问题.

(3) 基于进化算法的调度策略

针对整数规划计算时间消耗过大的问题^[11],很多学者利用进化算法来解决云环境下的资源分配与任务调度问题研究.其中,文献[13]研究了云计算环境下优先顺序受限的并行应用的调度问题,分别对优先顺序受限并行应用的完成时间和能耗进行了建模;然后提出了一种平行双目标混合方法,即把该调度问题建模为最小完成时间和能耗双目标优化问题,并利用遗传算法求解多目标优化问题.文中提出的并行双目标混合遗传算法兼顾到了能耗和应用完成时间的两个方面.文献[15]提出了一种基于蚁群优化的计算资源分配算法.该算法在分配计算资源时,首先预测潜在在可用节点的计算质量,然后根据云计算环境的特点,通过分析诸如带宽占用、线路质量和响应时间等因素对分配的影响,利用蚁群优化算法得到一组最优的计算资源.

综上所述,基于规则的调度方法具有计算速度快、操作简单的特点^[7],可以满足云平台即时运算即时分配的要求,但是并不能保证云平台资源的最合理分配,无法保证生成的调度策略是最优结果.基于最优化理论的调度方法,具有优化效果好、运算精确的特点,但是该方法运算速度较慢,当运算量较大时,需要非常大的时间消耗,难以应用于实际的云计算系统.基于启发式算法的调度方法可以大量缩短算法运行的时间^[11],可以满足云平台的资源分配的实时性要求.此类方法虽然无法完全获得全局优化问题的最优解,但是可以获得一个误差很小的近似解,能够满足实际工程的任务需求.因此,本文针对云环境下的容灾费用优化这样一个 NP 困难问题,设计了一

种基于进化算法的数据容灾策略,并开展了相关的研究工作.

2 问题模型及数学描述

本节首先介绍了分布式富云容灾系统的组成模块,然后对云存储环境下数据备份问题的相关概念进行阐述,并利用数学建模的方法抽象容灾问题模型.

2.1 系统模型

由于容灾系统具有用户任务量大、随机性强的特点,在面临海量的用户任务请求时,大量远程异地备份可能会存在时间性能问题.因此,本文设计了一个基于多代理机制的分布式富云容灾模型.容灾系统中设置了多个地理位置隔离的富云代理(rich cloud proxy).富云代理可以作为用户和云平台间的中介平台,帮助私有云平台完成用户任务的分配.当发生用户请求时,用户任务会优先发送给距离较近的富云代理,然后由富云代理进行调度分配.在海量任务到来时,多个代理可以平均分配调度任务,减少数据传输的响应时间,减少私有云平台的工作负荷.

分布式富云容灾模型如图 1 所示,其中,Cloud Users 是使用云计算平台完成相关工作的用户,用户可以是个人、企业或者其他的云提供商;Cloud Providers(CP)表示云提供商,其中,CP₁ 是接受用户任务请求的私有云提供商,CP₂~CP_L 是公有云提供商,可以为 CP₁ 提供容灾服务.当富云代理收到用户请求后,会根据容灾成本和带宽限制合理地选择数据副本的存放位置,决定是否需要租用其他云提供商的存储资源,为自身平台提供数据容灾.

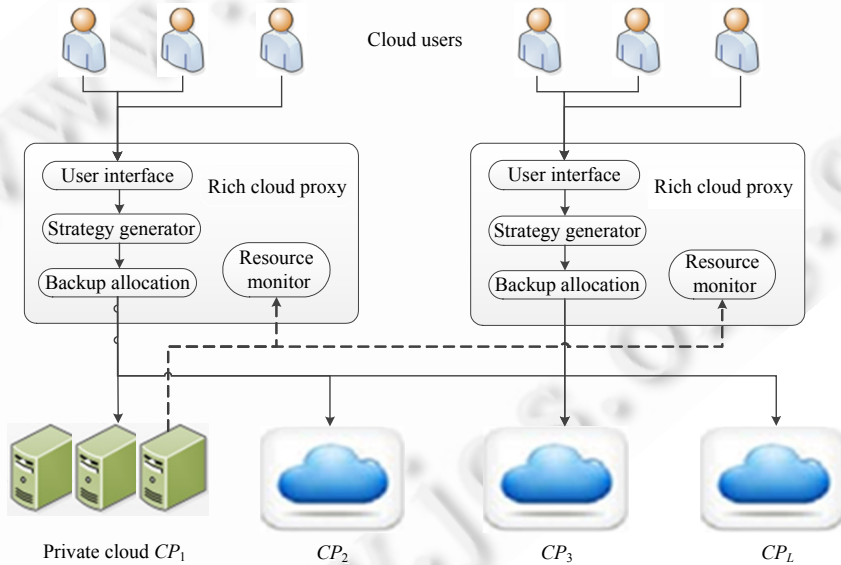


Fig.1 Multiple rich cloud proxy-based disaster recovery system

图 1 基于代理机制的富云容灾系统

在富云代理内部:

- User Interface 是用户与云计算平台信息交互的接口,用户可以通过 User Interface 向云提交使用的申请,查询任务执行的状态,检验所存储数据的完整性.
- Resource Monitor 负责监视云平台资源的动态变化情况,收集私有云云提供商 CP₁ 的相关资源信息,为容灾策略生成器(strategy generator)提供私有云的当前资源状态.
- 容灾策略生成器根据用户的任务请求信息和当前私有云的资源信息,以容灾费用为优化目标生成容灾策略.当有新的用户任务到达时,容灾策略生成器会根据存储费用、通信费用和数据传输带宽动态的将所有的数据备份进行重新分配,以实现较低成本的容灾保护.

- 副本分配器(backup allocation)是富云代理与各云平台的数据传输接口,负责发送和接收容灾数据。

2.2 数学描述

为了保证数据的可靠性,本文采用目前常用的多备份方式实现数据冗余^[3],即,对每个用户任务一次储存多个副本以实现对数据的容灾保护.在系统评价中,我们通过存储费用、通信费用和恢复时间目标(RTO)来衡量容灾系统的费用开销和数据可靠性.针对富云容灾系统的容灾费用优化问题,本文做出如下假定:

- (1) 私有云提供商的存储空间是有限的,为了满足更多用户的存储请求,私有云提供商需要借用其他云提供商的资源进行数据容灾.
- (2) 为了减少通信成本,当接到私有云 CP_1 发送的数据副本后,外部云 CP_k 会存放该备份直到用户申请的存储时间到达,而存放在私有云内部的数据副本可以根据容灾费用动态地存放在外部云.
- (3) 每一个数据副本不可以再分块存储到不同的云平台,即用户任务的 1 个数据副本同时只存储在 1 个云平台中.

容灾系统的数学描述如下:假设云提供商当前有 W 个用户 U_1, \dots, U_W , 每个用户有若干个子任务 N_1, \dots, N_W , 每个子任务的开始执行时间可以不同.为了清楚地表示用户的任务,可以用一个四元组 $\{t_{ij}, TS_{ij}, T_{ij}, D_{ij}\}$ 来表示每个任务的资源请求信息,其中, t_{ij} 表示第 i 个用户的第 j 个子任务, $i=1, 2, \dots, W, j=1, 2, \dots, N_i$; TS_{ij}, T_{ij} 和 D_{ij} 分别表示任务 t_{ij} 的开始时间、存储时间和存储容量.所有的存储备份会依据成本费用和恢复时间存放至不同的云提供商 CP_1, \dots, CP_L .不同的云提供商对服务的收费标准不同,本文重点考虑了数据备份服务的存储费用和通信费用.假定各云提供商数据存储的收费价格和带宽分别为 $Pc_1/GB \times h, \dots, Pc_L/GB \times h$ 和 BW_1, \dots, BW_L , 各云提供商的单位通信价格假定为 $Pt_1/GB, \dots, Pt_L/GB$.云平台 CP_1 出于可靠性保证的考虑,为了保证数据恢复时间会对平均传输带宽有所要求,假设系统对平均传输带宽的最低要求为 $minbw$ MB/s, 则数据容灾备份的费用开销及相应的约束条件可表示如下:

优化目标:

$$Cost = \sum_{i=1}^W \sum_{j=1}^{N_i} \sum_{k=1}^L (Pc_k \times T_{ijk} \times D_{ij} \times O_{ijk} + D_{ij} \times Pt_k \times x_{ijk}) \quad (1)$$

约束条件:

$$\sum_{i=1}^W \sum_{j=1}^{N_i} pc_1 \times T_{ij1} \times D_{ij} \times O_{ij1} \leq maxStorage \quad (2)$$

$$\sum_{k=1}^L O_{ijk} = m \quad \text{对于任意的用户任务,即 } i=1, 2, \dots, W; j=1, 2, \dots, N_i \quad (3)$$

$$\frac{1}{m \times \sum_{i=1}^W N_i} \sum_{i=1}^W \sum_{j=1}^{N_i} \sum_{k=1}^L bw_j \times O_{ijk} \geq minbw \quad (4)$$

其中,

- T_{ijk} 表示用户 i 的任务 j 的一个数据副本存储在第 k 个云平台的时间;
- D_{ij} 表示用户 i 的第 j 个任务的数据存储量;
- O_{ijk} 表示第 k 个云提供商存放的用户 i 的第 j 个任务的数据副本个数;
- x_{ijk} 是一个布尔型变量,当云提供商 CP_k 中存有用户 i 的任务 j 的备份时(即当 O_{ijk} 大于等于 1 时), x_{ijk} 等于 1; 否则, x_{ijk} 为 0;
- BW_k 表示云提供商 CP_k 的带宽;
- L_k 表示云提供商 CP_k 提供的链路条数;
- Tr_{ij} 表示故障发生后的数据恢复启动延迟;
- $maxStorage$ 是云提供商 CP_1 的最大存储空间.

上述各式中,富云容灾系统的费用开销主要由数据存储费用和数据通信费用两部分组成,见公式(1).公式

(2)说明云提供商 CP_1 存储能力的约束, CP_1 所存放的数据容量应小于等于其最大的存储能力.公式(3)表示云提供商会对每一个用户任务都进行 m 次备份以实现容灾服务.公式(4)表示容灾方案应确保一定的平均带宽,以保证数据恢复时间.

3 成本相关的云提供商数据可靠性保证算法(CAHRPA)

在容灾云环境中,用户任务具有工作量大、随机性强等特点,容灾系统的成本优化问题是一个典型的 NP 困难问题,利用传统的整数规划方法求解此类优化问题需要极大的时间消耗^[11],因此,本文利用一种简单、高效的进化算法差分进化算法^[16],设计并实现了一种成本相关的云提供商数据可靠性保证算法.算法的执行流程如图 2 所示.

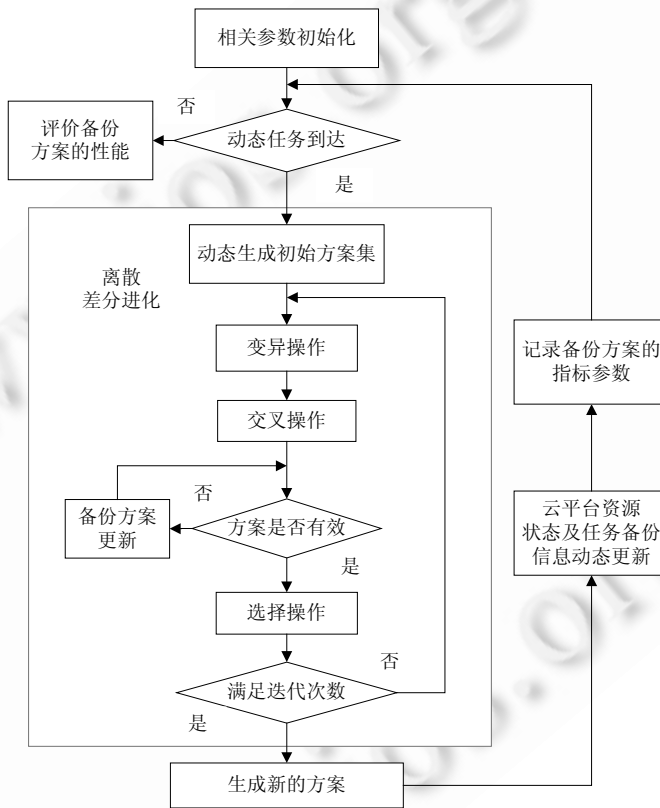


Fig.2 Procedure of CAHRPA

图 2 CAHRPA 的执行流程

本文的容灾问题是一个离散时间的动态调度问题,每单位时刻触发一次策略调度算法完成数据容灾备份.在图 2 中,富云代理会根据接到的用户任务选取容灾策略.CAHRPA 算法首先分析自身平台的存储资源使用情况,对相关参数进行初始化操作,具体包括差分进化算法的变异率、交叉率、种群个数.然后,CAHRPA 根据用户任务数量进行种群初始化、变异、交叉、选择及约束检验等操作,直至满足算法运行的结束条件.当有新任务动态到达时,系统会重新启动优化算法,动态地获取新的容灾方案,并记录云平台的资源使用状态和任务分配信息.在进行优化计算时,我们根据约束条件设计了一个检验更新操作以确保方案的合理性.当完成所有任务后,系统会对容灾方案进行评价分析,综合衡量容灾方案的数据恢复时间和使用费用开销.CAHRPA 算法的关键步骤如下:

步骤 1. 初始化生产方案集.

假设云提供商会为每个用户任务进行 m 次数据备份,因此在个体编码中,用 m 个基因位可以表示一个用户任务的备份方案.假设某一时刻共有云平台共接到 n 个存储任务请求,则在此时共有 $m \times n$ 个数据副本需要分配调度,即,一个个体可以由 $m \times n$ 个基因位组成.为了说明本文采用的编码方式,我们给出一个包含两个用户任务、6 个云提供商、采用 3-replicas 规则备份的编码示例.具体的编码方式见表 1.

Table 1 Encoding example

表 1 编码示例

任务编号	Task 1			Task 2		
备份编号	1	2	3	1	2	3
云提供商编号	0	4	2	3	5	2

编码示例中共有两个存储任务,分别是 Task 1 和 Task 2,每个存储任务有 3 个数据副本,分别编号为 1,2,3.该示例表明了不同任务的每个数据副本的存放位置.在该示例中,任务 Task 1 的 3 个数据副本分别存放到了编号为 1,4,2 的 3 个云存储平台.在算法运行过程中,每个个体可以由云提供商编号组成的元组表示,即,个体可以由 6 个基因位表示,如 {0,4,2,3,5,2}.云提供商编号的取值为 $[0, L-1]$ 间的整数,其中,0 表示私有云提供商.示例中, L 的取值为 6.

假定某一时刻初始种群的规模设定为 NP ,个体长度为 $D(D=m \times n)$,则初始种群可以定义为 $\{x_{ij}|i=1,2,\dots, NP; j=1,2,\dots, D\}$,其中, x_{ij} 为随机初始生成,每个个体可以由 x_i 表示.

步骤 2. 变异操作.

针对数据备份策略选择这一离散优化问题,本文采用了一种离散化的变异操作,通过取模操作对原有实数变异规则改进.具体的变异操作如下:

$$v_{ij} = \begin{cases} \text{mod}((x_{aj} + x_{bj} + x_{cj} + L), L), & \text{rand} < F \\ \text{mod}((x_{aj} + L), L), & \text{rand} \geq F \end{cases} \quad (5)$$

其中, $a, b, c, i \in (1, 2, \dots, NP)$ 是互不相同的随机整数; rand 是 $[0, 1]$ 之间均匀分布的随机数; L 表示可选的云提供商的个数; F 为缩放因子,取值范围通常选为 $[0, 1]$.

步骤 3. 交叉操作.

DE 的交叉操作是为了通过变异向量和目标向量的结合来提高变异向量的多样性.种群通过交叉操作生成新的向量 $u_i = [u_{i1}, u_{i2}, \dots, u_{im}]$.

$$u_{ij}^G = \begin{cases} v_{ij}^{G+1}, & \text{rand} \leq CR \text{ or } j = \text{randr} \\ x_{ij}^G, & \text{rand} > CR \text{ and } j \neq \text{randr} \end{cases} \quad (6)$$

其中, rand 是 $[0, 1]$ 之间均匀分布的随机数. CR 是 $[0, 1]$ 间的常数,称为交叉概率因子. CR 值越大,发生交叉的可能性越大. $CR=0$ 表示没有交叉. randr 是在 $[1, NP]$ 之间随机选择的整数,它保证 u_i 至少要从 v_i 中获得 1 个元素,否则就不会有新的向量产生,群体就不会发生变化. G 表示当前的进化迭代数.

步骤 4. 选择操作.

选择操作通过比较适应度函数的数值来确定个体的优劣,差分进化通过比较父代个体和目标个体的适应度函数值来确定进入子代的个体,从而完成选择操作.

$$x_i^{G+1} = \begin{cases} u_i^{G+1}, & f(u_i^{G+1}) < f(x_i^{G+1}) \\ x_i^G, & \text{otherwise} \end{cases} \quad (7)$$

步骤 5. 约束检验及备份更新操作.

在富云容灾模型中,容灾策略需要综合考虑私有云存储资源限制、用户任务备份次数限制和平均带宽限制,因此,本文的容灾策略选择问题是一个约束优化问题.差分进化算法是一种随机操作算法,其变异交叉后产生的个体可能不满足资源约束的限制.为了确保每个个体都能代表合理的数据备份策略,本文设计了约束检验

操作,对进入选择操作的种群个体进行检验.

在进行选择操作前,为了确保每个种群个体所代表的容灾方案能够满足约束条件,需要对个体进行检验及更新操作.某一时刻,为了判断个体 x_i 是否满足约束条件,CAHRPA 从某一任意基因位开始遍历个体 x_i ,计算存放在私有云平台的用户任务数据量.当某一用户任务 k 的数据副本存放在私有云平台,且私有云平台的存储资源已经达到饱和时,则需要原有编码进行更新操作,即任务 k 的数据副本存放位置需要在 $(1, L-1)$ 范围内重新随机赋值,直至遍历个体所有的基因位.

步骤 6. 生成备份策略.

当差分进化算法达到迭代次数完成进化计算后,CAHRPA 会对新生种群中所有个体(代表一个备份策略)按照容灾费用进行排序,从中选取符合平均带宽限制且容灾费用最少的备份策略作为输出.

4 实验及结果分析

本文针对数据密集型任务设计并实现了一个基于 Java 的富云容灾系统仿真平台,模拟云计算平台从用户提交任务请求、备份策略选择到数据副本存储的全部执行过程.为了进一步分析本文提出的数据备份选择策略 CAHRPA 的实际效果,我们设计了另外 3 种基于备份的数据冗余策略,并通过实验具体分析了在不同用户任务数目时各种容灾策略的优缺点.本文实验的主要评价内容包括:

- (1) 数据容灾费用,包括数据存储费用和数据传输费用.
- (2) 私有云存储空间利用率,计算方式为:私有云占用存储空间/私有云全部存储空间 $\times 100\%$.
- (3) 通过分析仿真实验的相关数据结果,可以为本文提出的数据容灾策略提供有效的评判依据.

4.1 实验环境与配置

系统运行平台的基本配置为 Intel Dual 1.6G Hz, 1 GB 主存, 120 GB 硬盘. 仿真平台由 4 大模块组成:

- (1) 模拟条件生成模块:负责为备份选择策略生成各种模拟条件,包括多个云提供商的收费价格和数据传输带宽,用户任务的开始时间、执行时间和任务存储量.
- (2) 资源及任务分析模块:负责接收云所提供的数据调度请求,收集并提供各云提供商相关资源的动态信息.
- (3) 容灾策略选择模块:根据收集的相关云计算平台的资源信息,结合本文所提出的备份选择算法模拟实际数据中心的备份策略选择过程.
- (4) 结果显示模块:将备份选择策略及相关数据分析的最终结果进行输出.

实验中假定存在 4 个云提供商:私有云提供商 CP_1 和 3 个外部云提供商($CP_2 \sim CP_4$),其中,私有云 CP_1 的最大存储空间为 5 000G, 3 个外部云提供商($CP_2 \sim CP_4$)没有存储空间限制,可以为私有云提供商 CP_1 提供数据容灾服务.实验中,私有云提供商 CP_1 会对每个用户任务进行 3 次备份.我们依据好服务高价格的原则设定了各云提供商的存储价格、通信价格及带宽,具体信息见表 2.

Table 2 Prices and bandwidths of each cloud provider

表 2 云提供商的收费价格及带宽

云提供商	存储价格(\$/G×h)	通信价格(\$/G)	带宽(MB/s)
CP_1	0.6	0.2	100
CP_2	0.2	0.5	10
CP_3	1.0	0.6	30
CP_4	1.5	0.4	15

为了减少发生灾难后的数据恢复时间,保证数据的可靠性,容灾系统会对用户任务的最低平均带宽 $minbw$ 进行约束.若 $minbw$ 设置较小则无法保证数据的恢复时间,若 $minbw$ 设置较大则可能导致无法得出满足可靠性要求的数据备份分配策略.因此,本文取 $minbw$ 为外部云的最大带宽,即 30MB/s.

实验中,所有的用户任务均由模拟条件生成模块随机生成,用户提交的任务信息由开始时间、运行时间和

数据存储量这 3 个部分组成.为了更好地模拟实际的用户任务请求,我们对实验室服务器中 15 个用户的实际数据备份信息进行了分析,发现各用户的存储数据所占空间在 60G~200G 空间不等,平均值在 125G 左右.因此在实验参数设置中,我们设置用户任务的数据存储量服从均值为 125、标准差为 70 的正态分布.任务的开始时间服从 0~10 的随机分布,任务运行时间服从均值 12、标准差为 5 的正态分布.具体的用户任务参数见表 3.

Table 3 Task parameters

表 3 任务参数

参数	取值
任务开始时间	Uniform (0,10)
任务运行时间	Normal (25,15)
数据存储量	Normal (125,75)

为了分析本文数据备份策略 CAHRPA 的实际效果,我们设计了另外 3 种数据备份策略:Random 策略、RTO-Cost 均衡策略、内部资源优先策略.3 种容灾策略具体描述如下:

- Random 策略

以随机的方式为每个用户任务的数据副本选择不同的云提供商进行数据备份.由于私有云提供商 CP_1 的内部存储资源有限,因此,如果随机选取了 CP_1 作为新备份的存放位置,但 CP_1 的存储资源无法满足新的存储任务,Random 策略会在 $CP_2 \sim CP_4$ 中重新选择一个云提供商为用户任务进行数据备份.

- RTO-Cost 均衡策略

首先将一个存储任务的数据副本存放在 CP_1 内部,然后将任务的另外两个数据副本分别存放到传输带宽最大和存储费用最低的两个云提供商.当私有云提供商 CP_1 内部存储资源不足时,选择存储费用较小的云提供商存放数据备份.此调度策略的目的是平衡数据恢复时间和数据容灾费用,实现一种既能保证数据恢复时间又能减少费用开销的容灾策略.

- 内部资源优先

首先对同一时刻的开始执行的用户任务进行排序,将数据量较大的任务备份优先存放在私有云 CP_1 内部,直至完成全部数据容灾任务或私有云 CP_1 的内部存储资源达到饱和.当私有云 CP_1 内部存储资源不足时,此策略先将一个任务备份存放至传输带宽最大的云提供商,再将其他备份存放到存储费用最低的云提供商.此调度策略的目的是尽量利用私有云内部资源进行数据备份,以减少容灾费用开销(适用任务较少情况).

在这 3 种方法中,

- 内部资源优先方法是一种寻求低成本数据容灾的方式,当用户任务较少时,可以更好地为私有云平台提供低成本且恢复速度较快的容灾方式.但是当用户任务较多时,云平台的平均数据恢复速度无法得到保证.
- RTO-Cost 方法是一种寻求容灾成本与恢复速度的折中方法.同时,由于该方法将所有数据副本存放在不同的云平台,可以较好地避免数据丢失给云提供商造成的损失.

4.2 实验结果及分析

本节通过模拟不同任务数量下各备份策略的任务执行情况,对各备份策略下云计算平台的费用开销、平均带宽以及私有云存储空间利用率进行比较.在实验中,CAHRPA 策略的种群数量为个体编码长度的 5 倍,变异率 F 为 0.8、交叉率 CR 为 0.6,最大迭代次数 $G_{\max}=200$.文中所有实验结果是进行 30 次实验的平均值.

图 3 中,随着用户个数的增加,各容灾策略的平均带宽均呈现下降的趋势.这是因为当用户数量较少时,各容灾策略会将更多的数据副本存放至带宽较大的私有云平台,随着存储任务数量的增加,私有云平台内部的存储资源逐渐达到饱和,新提交的用户任务副本会存放至较低带宽的外部云平台.因此,各容灾策略的平均带宽会随着用户个数的增加而下降.

图 4 展示了各容灾策略在不同用户个数及任务数量时的费用开销.随着用户个数及任务数量的增大,各容灾策略的费用开销也在不断增加.与其他几种容灾策略相比,CAHRPA 策略在用户个数不同时都可以获得较小

的容灾费用,而且在用户任务数量较少时,CAHRPA 策略在费用开销方面的优势更为明显.以 20 个用户为例,CAHRPA 策略的容灾费用开销分别是内部资源优先策略的 89.6%、RTO-Cost 均衡策略的 84.5%、Random 策略的 68.9%.当用户个数为 200 时,CAHRPA 策略的容灾费用开销已经大于内部资源优先策略.但是通过比较图 4 中各容灾策略的平均带宽可以发现:当用户个数为 200 时,只有 RTO-Cost 均衡策略和 CAHRPA 策略的平均带宽达到了 30MB/s 平均带宽的约束条件,而 CAHRPA 策略的容灾费用明显小于 RTO-Cost 均衡策略.因此,本文提出的 CAHRPA 策略可以在保证平均带宽限制的同时,为云提供商提供一种成本优化的、高可靠的数据容灾服务,极大地减小了云提供商的容灾费用开销.

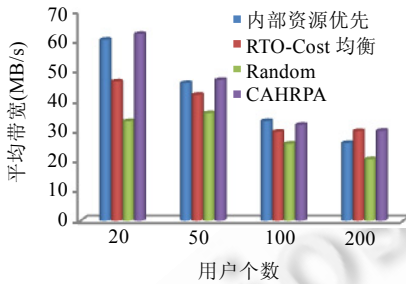


Fig.3 Average bandwidth

图 3 平均带宽

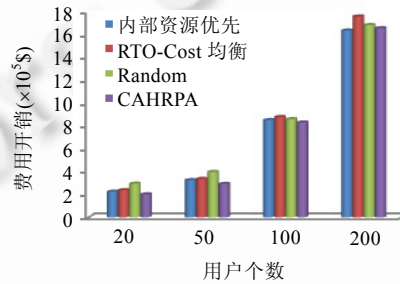


Fig.4 DR cost

图 4 容灾费用

图 5 给出了采用不同容灾策略时私有云提供商 CP_1 的内部存储空间利用率的变化情况.在用户数目不同的情况下,各容灾策略在用户数目不同时的私有云存储空间利用率会出现不同的变化趋势,而本文提出的 CAHRPA 策略都可以较长时间保持较高的私有云存储空间利用率.

当用户个数较少时(如图 5(a)所示),各容灾策略的私有云存储空间利用率均呈现先上升后下降的变化趋势.由于用户存储任务较少,在容灾策略开始执行数据副本分配调度时,用户任务的存储空间占有量远小于私有云的可用存储空间.随着用户任务的不断提交,私有云的存储空间利用率也得以不断提高,达到最高值,并在 10 时刻后开始逐渐下降.

当用户个数为 50 时(如图 5(b)所示),随着时间的推移,不同的容灾策略的云提供商 CP_1 的内部资源利用率呈现了截然不同的变化趋势.其中,Random 策略和 RTO-Cost 均衡策略的私有云 CP_1 存储空间利用率曲线出现了先上升后下降的抛物线形式,并在 7 时刻达到了存储空间利用率的最大值;而采用内部资源优先策略和 CAHRPA 策略时, CP_1 内部资源利用率则呈现由最高点缓慢下降的趋势.在任务开始执行时,由于用户任务数量较少,采用 Random 策略和 RTO-Cost 均衡策略时数据副本更多地存放至外部云平台,无法更好地利用 CP_1 的内部存储资源,而采用内部资源优先策略和 CAHRPA 策略可以更多地使用 CP_1 的内部资源进行数据备份,而且 CAHRPA 策略可以更长时间地保有较高的私有云存储空间利用率,从而实现更低成本的数据容灾.

当用户个数较多时(如图 5(c)、图 5(d)所示),各容灾策略的私有云存储空间利用率都呈现由高到低缓慢下降的趋势.此现象说明:0 时刻云提供商接到的用户任务较多,私有云 CP_1 的内部存储资源几乎被完全占用,有很多数据副本被分配给费用较高的外部云提供商进行容灾备份.10 时刻后,云平台没有收到新的任务请求,私有云存储空间利用率开始逐渐下降,直到完成所有任务.

从图 6 中可以清晰地看出:在不同用户个数时,本文的提出的 CAHRPA 策略都可以达到 70%左右的平均存储空间利用率,而其他策略的平均存储空间利用率只能达到 50%~60%.以 200 个用户为例,CAHRPA 的平均存储空间利用率可以达到 71%,远高于内部资源优先策略的 55%、RTO-Cost 均衡策略的 56%和 Random 策略的 40%.通过比较各容灾策略的平均存储空间利用率可以明显看出:本文提出的 CAHRPA 策略可以更好地利用云提供商 CP_1 自身的内部资源进行数据备份.

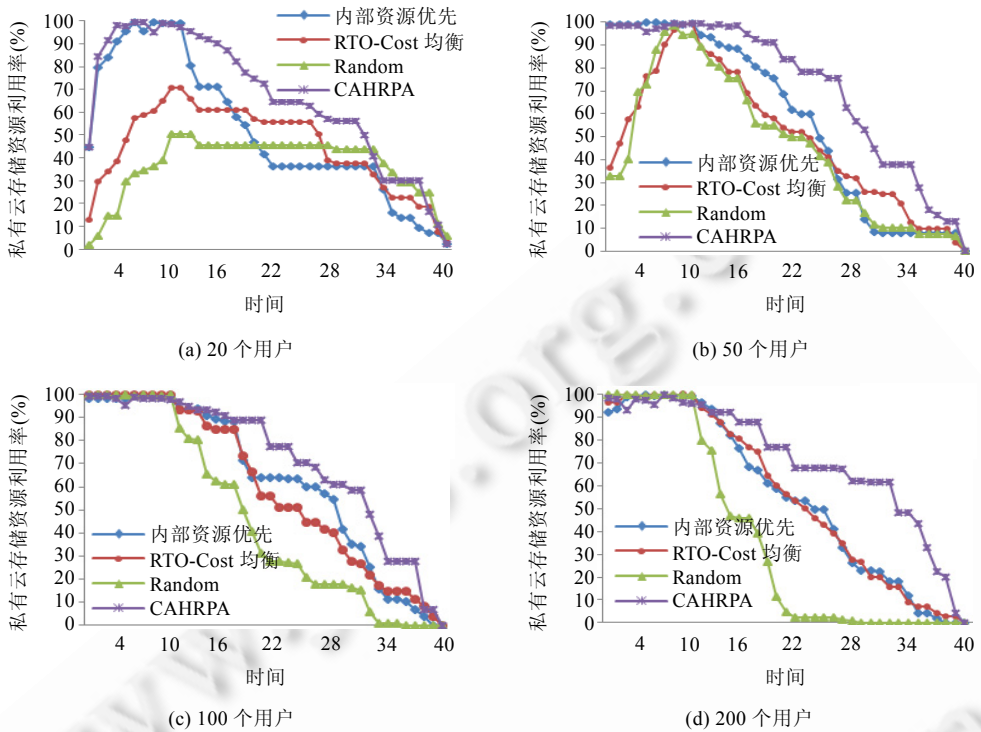


Fig.5 Real time storage space utilization of private cloud

图 5 私有云实时存储空间利用率

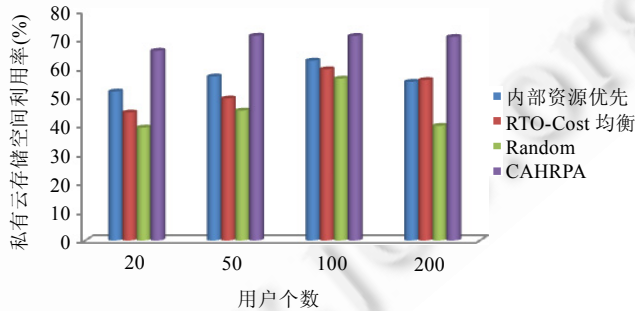


Fig.6 Average storage space utilization of private cloud

图 6 私有云平均存储空间利用率

综上所述,对于不同的用户任务数量,CAHRPA 策略都可以根据恢复时间和容灾费用动态地调整任务备份的存放位置,更好地利用云提供商 CP₁ 自身的内部资源,并在保证传输带宽和数据可靠性的同时,实现一种成本优化的容灾方案,明显优于其他几种容灾策略。

5 结论

数据容灾保护已经成为关乎云提供商利益的重要因素,同时,数据可靠性保证和容灾成本控制对云服务商而言又是一个矛盾的问题.为了减少云提供商的容灾费用开销,本文设计了一种可以借用外部云提供商存储资源进行数据冗余的分布式富云容灾模型,并采用多副本方式实现数据可靠性保护.通过设置多个地理位置隔离

的富云代理,富云容灾模型可以实现多个云平台海量用户任务的分配,从而减少用户任务的响应时间,减少私有云平台的工作负荷.针对富云容灾模型中的容灾费用优化这一 NP 难问题,本文提出了一种成本相关的云提供商数据可靠性保证算法 CAHRPA.CAHRPA 以数据容灾成本为优化目标,并将私有云存储空间和存储数据的平均带宽作为约束条件,对所有用户任务进行动态分配.实验中,为了对 CAHRPA 方法的效果进行评价,我们设计了另外 3 种数据容灾策略(内部资源优先、RTO-COST 均衡和随机策略)作为对比策略.实验结果表明:本文提出的 CAHRPA 既可以满足多备份及平均传输带宽要求,又可以有效降低云提供商的数据容灾费用.本文的方法为云提供商的数据容灾保护提供了一种新的思路.

References:

- [1] Hong C, Zhang M, Feng DG. Achieving efficient dynamic cryptographic access control in cloud storage. *Journal on Communications*, 2011,32(7):125–132 (in Chinese with English abstract).
- [2] Feng DG, Zhang M, Zhang Y, Xu Z. Study on cloud computing security. *Ruan Jian Xue Bao/Journal of Software*, 2011,22(1): 71–83 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3958.htm> [doi: 10.3724/SP.J.1001.2011.03958]
- [3] Wang HY, Jing QF, Chen R, He BS, Qian ZP, Zhou LD. Distributed systems meet economics: Pricing in the cloud. In: *Proc. of the 2nd USENIX Conf. on Hot Topics in Cloud Computing (HotCloud 2010)*. Usenix, 2010. 6–11.
- [4] IBM Corporation. A disaster recovery very solution selection methodology. 2004. <http://www.redbooks.ibm.com/redpapers/pdfs/redp3847.pdf>
- [5] Wood T, Cecchet E, Ramakrishnan KK. Disaster recovery as a cloud service: Economic benefits & deployment challenges. In: *Proc. of the 2nd USENIX Conf. on Hot Topics in Cloud Computing (HotCloud 2010)*. Usenix, 2010. 8–15.
- [6] Abu-Libdeh H, Princehouse L, Weatherspoon H. RACS: A case for cloud storage diversity. In: *Proc. of the ACM SoCC*. ACM, 2010. 229–239.
- [7] Abirami SP, Shalini R. Linear scheduling strategy for resource allocation in cloud environment. *Int'l Journal on Cloud Computing: Services and Architecture*, 2012,2(1):9–17.
- [8] Weinman J. Mathematical proof of the inevitability of cloudcomputing. 2011. <http://www.joeweinman.com>
- [9] Gao HQ, Xing Y. Research on cloud resource management model based on economics. *Computing Engineering and Design*, 2010, 31(19):4139–4212 (in Chinese with English abstract).
- [10] Chaisiri S, Lee BS, Niyato D. Optimal virtual machine placement across multiple cloud providers. In: *Proc. of the IEEE Asia Pacific Services Computing Conf. IEEE*, 2009. 103–110. [doi: 10.1109/APSCC.2009.5394134]
- [11] Bossche RV, Vanmechelen K, Broeckhove J. Cost-Optimal scheduling in hybrid IaaS clouds for deadline constrained workload. In: *Proc. of the IEEE 3rd Int'l Conf. on Cloud Computing. IEEE*, 2010. 228–235. [doi: 10.1109/CLOUD.2010.58]
- [12] Wei GY, Vasilakos A, Zheng Y, Xiong NX. A game theoretic method of fair resource allocation for cloud computing services. *The Journal of Supercomputing*, 2010,54(2):252–269. [doi: 10.1007/s11227-009-0318-1]
- [13] Mezmaz M, Melab N, Kessaci Y, Lee YC, Taibi EG, Zomaya AY, Tuytens D. A parallel bi-objective hybrid metaheuristic for energy-aware scheduling for cloud computing systes. *Journal of Parallel and Distributed Computing*, 2011,71(11):1497–1508. [doi: 10.1016/j.jpdc.2011.04.007]
- [14] Damay J, Quilliot A, Sanlaville E. Linear programming based algorithms for preemptive and non-preemptive rcpsp. *European Journal of Operational Research*, 2007,182(3):1012–1022. [doi: 10.1016/j.ejor.2006.09.052]
- [15] Hua XY, Zheng J, Hu WX. Ant colony optimization algorithm for computing resource allocation based on cloud computing environment. *Journal of East China Normal University*, 2010,(1):127–134 (in Chinese with English abstract).
- [16] Storn R, Price K. Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 1997,11(4):341–359. [doi: 10.1023/A:1008202821328]

附中文参考文献:

- [1] 洪澄,张敏,冯登国.面向云存储的高效动态密文访问控制方法.通信学报,2011,32(7):125–132.
- [2] 冯登国,张敏,张妍.云计算安全研究.软件学报,2011,22(1):71–83. <http://www.jos.org.cn/1000-9825/3958.htm> [doi: 10.3724/SP.J.1001.2011.03958]

- [9] 高宏卿,邢颖.基于经济学的资源管理模型研究.计算机工程与设计,2010,31(19):4139-4212.
- [15] 华夏渝,郑骏,胡文心.基于云计算环境的蚁群优化计算资源分配算法.华东师范大学学报:自然科学版,2010,1:127-134.



钟睿明(1985-),男,辽宁鞍山人,博士,主要研究领域为数据存储,云计算,进化计算.

E-mail: ruiming_zhong@163.com



刘川意(1982-),男,博士,讲师,CCF 会员,主要研究领域为大规模数据存储,云计算.

E-mail: cy-liu04@mails.tsinghua.edu.cn



王春露(1969-),女,副教授,主要研究领域为计算机网络,信息安全,智能交通.

E-mail: wangcl@bupt.edu.cn



项菲(1982-),女,博士,主要研究领域为云存储,多目标进化计算.

E-mail: frey207@sina.com