

(12) 发明专利申请

(10) 申请公布号 CN 102902605 A

(43) 申请公布日 2013. 01. 30

(21) 申请号 201210372575. 7

(22) 申请日 2012. 09. 29

(71) 申请人 山东浪潮齐鲁软件产业股份有限公司

地址 250100 山东省济南市高新区舜雅路 1036 号

(72) 发明人 张华 王刚 张立勇 胡顺杰 陈峰

(51) Int. Cl.

G06F 11/14 (2006. 01)

G06F 17/30 (2006. 01)

H04L 29/08 (2006. 01)

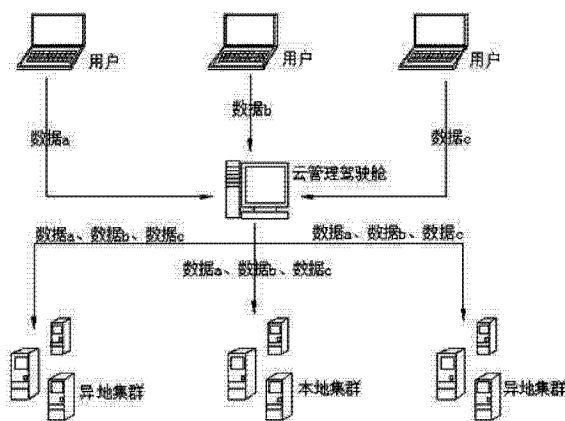
权利要求书 1 页 说明书 3 页 附图 2 页

(54) 发明名称

一种分布式云计算集群组海量数据备份和还原的方法

(57) 摘要

本发明公开一种分布式云计算集群组海量数据备份和还原的方法,该集群组以分布式文件系统为载体,在集群之上建立集群组,在同一区域和不同区域本别部署分布式文件集群,通过云管理驾驶舱对各集群进行统一管理和操作,进行文件的存储和数据的分布式计算,同一份文件在各集群至少保存一份,部分集群异常不影响整个系统运行,云管理驾驶舱自动记录异常集群和丢失文件,并在集群恢复后将丢失文件保存到该集群,系统中文件的备份和还原过程无需操作,软件系统的使用不受影响。该分布式云计算集群组以变通方式实现了海量数据异地备份与还原,提高了分布式计算的效率,保证了软件系统的高可用性,同时为企业和软件使用单位节省了成本。



1. 一种分布式云计算集群组海量数据备份和还原的方法,其特征在于以分布式文件系统为载体,在分布式集群之上建立集群组,在同一区域和不同区域分别部署分布式文件集群,实现海量数据本地、异地的备份和还原,内容包括:集群组分布式计算、云管理驾驶舱和海量数据异地备份与还原三部分,具体步骤如下:

(1) 集群组分布式计算

分散在同一区域和不同区域的集群分别对数据进行计算和抽取,并将结果进行汇总;

(2) 云管理驾驶舱

云计算集群组的CPU,担负着整个集群组的管理与分发,异常的处理与恢复,当云管理驾驶舱收到的是一个文件,则将文件根据不同区域保存到不同的集群组,保证每个区域至少保存一份文件;当某个区域出现异常无法保存时,管理驾驶舱记录该区域的信息以及未保存到该区域的文件,并时时监控异常区域的情况;当异常区域正常连接时,云管理驾驶舱自动从其它集群读取相关数据,将未保存到该区域的文件或数据重新保存到该区域集群中,实现数据的恢复;

当云管理驾驶舱收到的是一组数据和指令时,管理驾驶舱首先检查目前可以连接的集群,并将数据分发到各集群按照指令进行运算,并最终将结果汇总输出;

(3) 海量数据异地备份与还原

在不同区域部署不同的集群,保证数据在不同区域至少保存一份,实现数据的异地备份;当本地集群出现异常时,云管理驾驶舱自动访问异地集群,并保存异常相关信息,等本地集群恢复后,自动将丢失文件存入到本地集群;故障系统的还原无需切换,没有时间消耗。

一种分布式云计算集群组海量数据备份和还原的方法

技术领域

[0001] 本发明涉及到云计算操作系统领域,具体的说是一种分布式云计算集群组海量数据备份和还原的方法。

背景技术

[0002] 随着 web2.0 的诞生,IT 技术的飞跃发展,人们对信息化需求越来越大,传统的结构化数据库面对日益增长的海量数据已明显力不从心。云计算开始登上 IT 的舞台, NoSql 数据库被广泛使用,分布式文件存储已成为主角。

[0003] 对于一些行业,比如卫生、电信,每天的数据增长量非常庞大,海量数据的存储问题可以使用分布式文件系统来存储,但面对庞大的数据,本地和异地的备份与还原的确让人头疼,也是很多 IT 企业一直研究的课题。目前市场上的备份软件大部分收费且很多是单纯的拷贝移动文件,并简单的进行管理。一旦本地出现问题,受带宽和其他因素影响,根本无法在短时期内将异地备份数据还原,无法保证系统的高可用。

[0004] 因此需要一种海量数据的本地、异地备份方案来解决日益增长的数据,来保证系统的高可用性。

发明内容

[0005] 本发明针对目前海量数据本地、异地处理中存在的问题,提出一种分布式云计算集群组,以解决海量数据的本地、异地备份和还原的问题。

[0006] 本发明的目的时按一下方式实现的,以分布式文件系统为载体,在分布式集群之上建立集群组,在同一区域和不同区域分别部署分布式文件集群,实现海量数据本地、异地的备份和还原,内容包括:集群组分布式计算、云管理驾驶舱和海量数据异地备份与还原三部分,具体步骤如下:

1) 集群组分布式计算

分散在同一区域和不同区域的集群分别对数据进行计算和抽取,并将结果进行汇总;

2) 云管理驾驶舱

云计算集群组的 CPU,担负着整个集群组的管理与分发,异常的处理与恢复,当云管理驾驶舱收到的是一个文件,则将文件根据不同区域保存到不同的集群组,保证每个区域至少保存一份文件;当某个区域出现异常无法保存时,管理驾驶舱记录该区域的信息以及未保存到该区域的文件,并时时监控异常区域的情况;当异常区域正常连接时,云管理驾驶舱自动从其它集群读取相关数据,将未保存到该区域的文件或数据重新保存到该区域集群中,实现数据的恢复;

当云管理驾驶舱收到的是一组数据和指令时,管理驾驶舱首先检查目前可以连接的集群,并将数据分发到各集群按照指令进行运算,并最终将结果汇总输出;

3) 海量数据异地备份与还原

在不同区域部署不同的集群,保证数据在不同区域至少保存一份,实现数据的异地备

份；当本地集群出现异常时，云管理驾驶舱自动访问异地集群，并保存异常相关信息，等本地集群恢复后，自动将丢失文件存入到本地集群；故障系统的还原无需切换，没有时间消耗。

[0007] 本发明的有益效果是：本发明所述分布式云计算集群组，在分布式集群之上建立集群组，该集群组在同一区域和不同区域分别部署分布文件集群，通过云管理驾驶舱对各集群进行统一管理和操作，同一份文件在不同区域至少保存一份，只要有一个集群的文件存储成功，则文件保存成功，文件的存储对用户透明，部分集群异常不会影响整个系统运行，云管理驾驶舱会记录未保存成功的集群与文件，在异常集群恢复后自动从其它集群读取该文件数据保存到该集群，变通的实现海量数据异地备份或还原，大大提高系统的可用性和安全性。不同的集群组可以再次组成更高层次的集群组，由更高层次的云管理驾驶舱进行管理。

[0008] 该集群以云计算为理念，充分利用现有资源，在分布式集群之上建立集群组，不受区域限制，不受宽带限制，存储和计算完全对用户透明，以变通方式实现了海量数据异地备份与还原，提高了分布式计算的效率，保证了软件系统的高可用性，同时为企业和软件使用单位节省了成本。

附图说明

[0009] 附图 1 本发明所述集群组的宏观数据走向示意图；

附图 2 本发明所述集群组的微观数据走向示意图。

具体实施方式

[0010] 下面结合附图 1 和附图 2 对本发明的方法作进一步详细说明。

[0011] 本发明一种分布式云计算集群组海量数据备份和还原的方法，是以分布式文件系统为载体，构建对用户透明的集群组式文件存储与计算，其解决所述技术问题采用的技术方案为：该集群组由集群组分布式计算、云管理驾驶舱和海量数据异地备份与还原三部分组成，其中：

集群组分布式计算是指分散在同一区域和不同区域的集群，分别对数据进行计算和抽取，并将结果进行汇总；

云管理驾驶舱是云计算集群组的 CPU，担负着整个集群组的管理与分发，异常情况的处理与恢复；

海量数据异地备份与还原是指在不同区域部署不同的集群，保证数据在不同区域至少保存一份，实现数据的异地备份，当本地集群出现异常时，云管理驾驶舱自动访问异地集群，并保存异常相关信息，等本地集群恢复后，自动将丢失文件存入到本地集群。

[0012] 实施例：

本发明所述集群组由集群组分布式计算、云管理驾驶舱和海量数据异地备份与还原三部分组成，下面分别详细说明：

1. 集群组分布式计算

分散在同一区域和不同区域的集群，分别对数据进行计算和抽取，并将结果进行汇总。相对于分布式计算，集群组分布式计算使运算效率更高；

2. 云管理驾驶舱

云管理驾驶舱,云计算集群组的 CPU,担负着整个集群组的管理与分发,异常的处理与恢复;

当云管理驾驶舱收到的是一个文件,则将文件根据不同区域保存到不同的集群组,保证每个区域至少保存一份文件。当某个区域出现异常无法保存时,管理驾驶舱记录该区域的信息以及未保存到该区域的文件,并时时监控异常区域的情况。当异常区域正常连接时,云管理驾驶舱自动从其它集群读取相关数据,将未保存到该区域的文件或数据重新保存到该区域集群中,实现数据的恢复;

当云管理驾驶舱收到的是一组数据和指令时,管理驾驶舱首先检查目前可以连接的集群,并将数据分发到各集群按照指令进行运算,并最终将结果汇总输出;

3. 海量数据异地备份与还原

在不同区域部署不同的集群,保证数据在不同区域至少保存一份,实现数据的异地备份。当本地集群出现异常时,云管理驾驶舱自动访问异地集群,并保存异常相关信息,等本地集群恢复后,自动将丢失文件存入到本地集群。故障系统的还原无需切换,没有时间消耗,保证了系统的高可用,充分体现了集群组的优势。

[0013] 本发明所述分布式云计算集群组,以现有的分布式文件系统为载体,在集群之上建立集群组,在同一区域和不同区域本别部署分布式文件集群,通过云管理驾驶舱对各集群进行统一管理和操作,进行文件的存储和数据的分布式计算,同一份文件在各集群至少保存一份,部分集群异常时,云管理驾驶舱自动记录异常集群和丢失文件,并在集群恢复后将丢失文件保存到该集群,系统中文件的备份和还原过程无需操作,软件系统的使用不受影响。

[0014] 除说明书所述的技术特征外,均为本专业技术人员的已知技术。

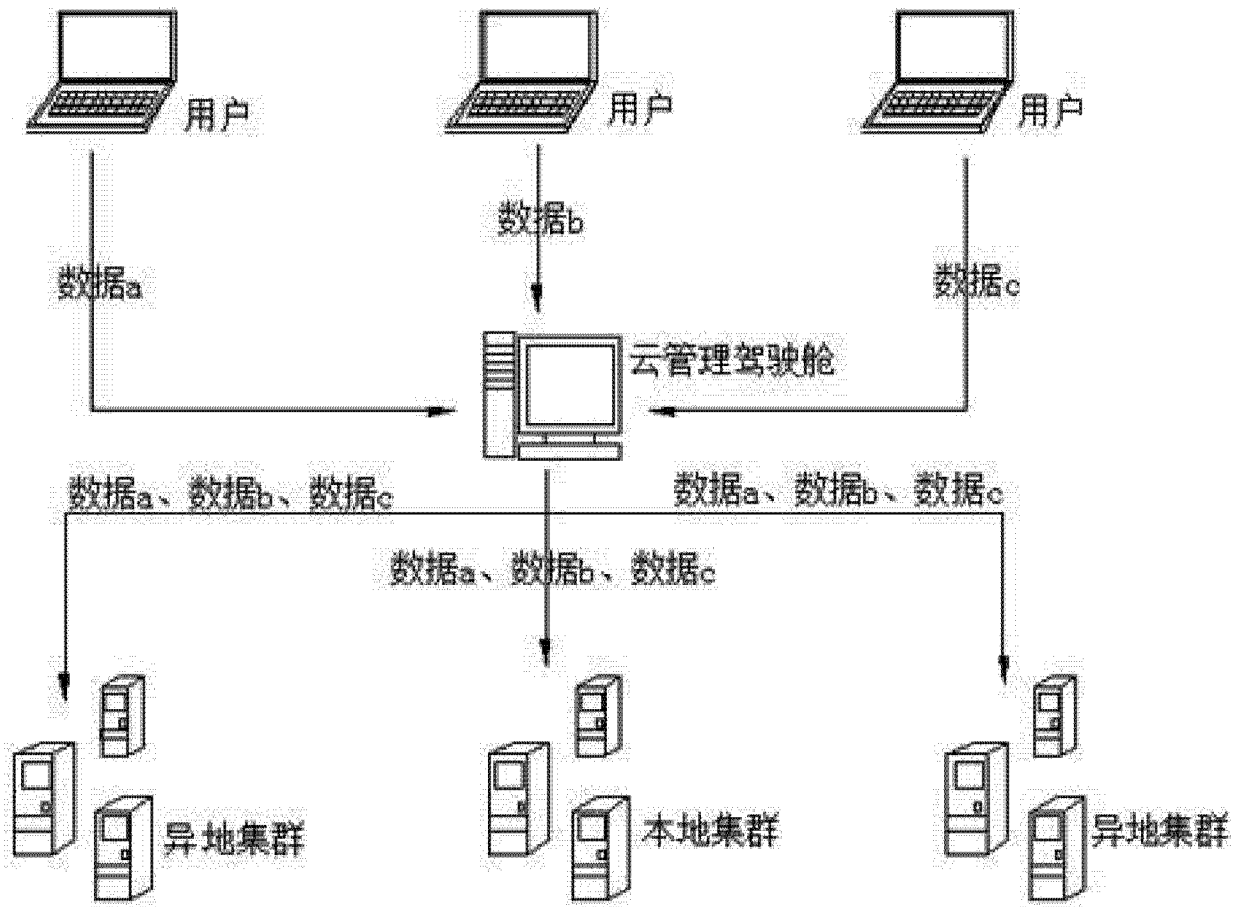


图 1

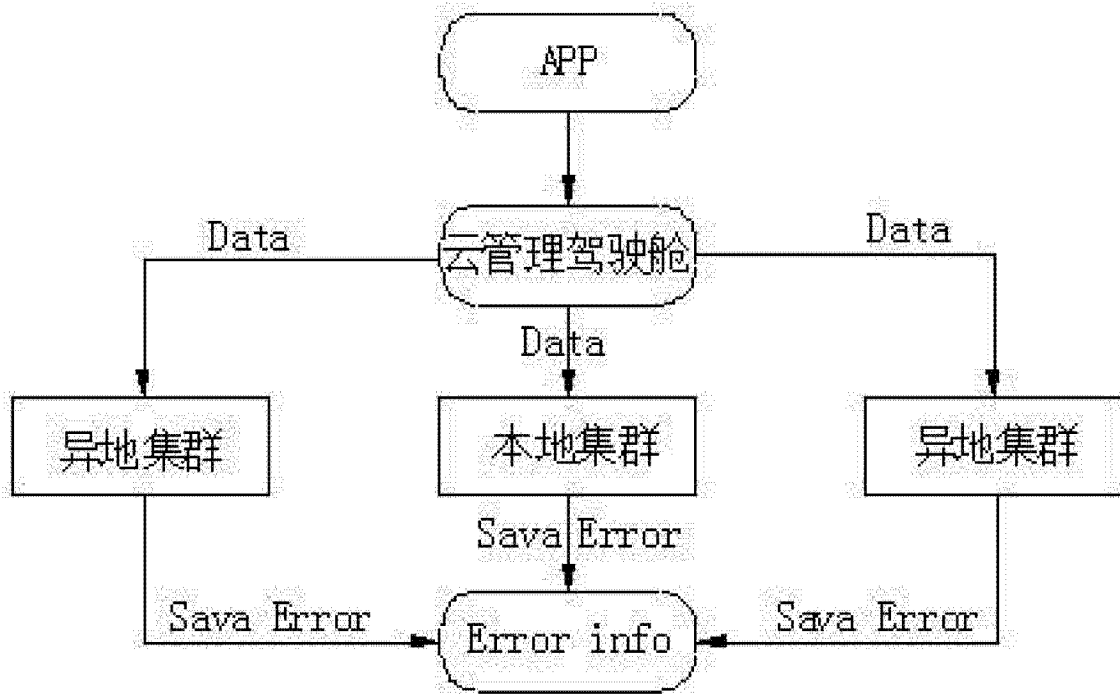


图 2