



(12)发明专利

(10)授权公告号 CN 103873501 B

(45)授权公告日 2017.07.18

(21)申请号 201210533741.7

(22)申请日 2012.12.12

(65)同一申请的已公布的文献号  
申请公布号 CN 103873501 A

(43)申请公布日 2014.06.18

(73)专利权人 华中科技大学  
地址 430074 湖北省武汉市洪山区珞瑜路  
1037号

(72)发明人 周可 王桦 王春燕 赵雨虹

(74)专利代理机构 华中科技大学专利中心  
42201

代理人 方放

(51) Int. Cl.

H04L 29/08(2006.01)

H04L 29/12(2006.01)

(56)对比文件

CN 102693168 A, 2012.09.26, 全文.

CN 101902498 A, 2010.12.01, 全文.

CN 102414673 A, 2012.04.11, 说明书第  
[0044]-[0055]段及图5-6.

审查员 楚彩虹

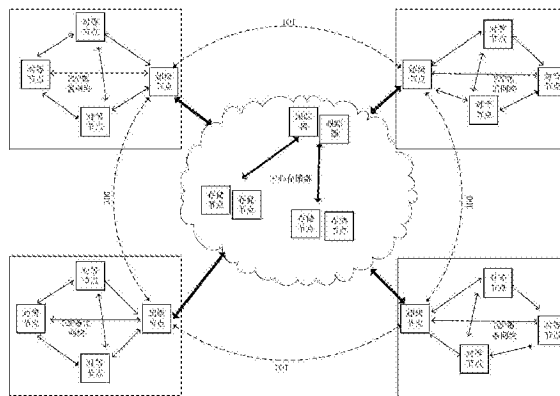
权利要求书4页 说明书7页 附图4页

(54)发明名称

一种云备份系统及其数据备份方法

(57)摘要

一种云备份系统及其数据备份方法,属于计算机数据备份系统,解决现有备份系统及数据备份方法所存在的存储层次单一,扩展性、容错性、效率不高的问题。本发明云备份系统,包括中心存储器、M个超级节点和M个P2P覆盖网络,中心存储器包括跟踪器和多组存储节点,每个跟踪器分别连接各存储节点;跟踪器和各存储节点分别与各超级节点通过网络连接,M个超级节点分别连接自身的P2P覆盖网络,各P2P覆盖网络包括多个对等节点。本发明的数据备份方法,包括数据备份作业步骤和数据恢复作业步骤。本发明针对开放式系统,数据在本系统中形成P2P覆盖网络存储、超级节点缓存、中心存储器数据管理三级存储,为企业或用户提供高扩展性、高容错性的云备份服务。



1. 一种云备份系统数据备份方法,包括数据备份作业步骤和数据恢复作业步骤,所述云备份系统,包括中心存储器、M个超级节点和M个P2P覆盖网络, $M \geq 1$ ;所述中心存储器包括1~2个跟踪器和1~8组存储节点,每组存储节点由2~3个存储节点构成,每个跟踪器分别连接各存储节点;

所述中心存储器的跟踪器和各存储节点分别与各超级节点通过网络连接,M个超级节点分别连接自身的P2P覆盖网络,各P2P覆盖网络由多个对等节点按照kademilia协议自组织而成;

每个对等节点都有一个节点标识,由SHA1哈希算法或者MD5哈希算法计算由该节点IP地址、端口和主机名串联起来的信息而获得该节点标识,每个对等节点具有内存和本地磁盘,在自身内存中维护160个K桶,在本地磁盘上维护一个数据库、一个配置文件、一个引导文件、一个块文件目录和一个元数据文件目录;所述数据库用于保存数据块和服务节点的对应关系;所述配置文件用来记录超级节点的IP地址和端口,该超级节点对外服务的端口,块文件目录路径和元数据文件目录路径;所述引导文件用来在磁盘上动态记录K桶中对等节点的网络信息,最初只有超级节点的网络信息,网络信息包括IP地址和端口;所述块文件目录用于该对等节点作为主服务节点或副本服务节点时存放所接收的数据块文件,文件名为数据块标识加后缀“.dat”;所述元数据文件目录用于该对等节点作为备份节点时存放每次备份作业产生的元数据文件;

各超级节点之间通过网络连接,各超级节点具有内存和本地磁盘,在本地磁盘维护一个数据库、一个块文件目录和一个作业信息文件目录,所述数据库用于保存超级节点在各对等节点上收集的数据块标识和将其发送到中心存储器时由中心存储器返回的新数据块标识的对应关系;所述块文件目录用于该超级节点存放从相关节点所接收的数据块;所述作业信息文件目录用于存放各个备份作业相关的作业信息文件,作业信息文件用来记录各个对等节点的各个备份作业号和其相对应的元数据文件各文件块标识的对应关系;

所述对等节点能够作为备份节点、主服务节点或副本服务节点,所述对等节点也能够作为恢复节点;其特征在于:

A. 所述数据备份作业步骤:

某个P2P覆盖网络中的某个对等节点向本P2P覆盖网络中其它对等节点发出备份命令,发起备份命令的对等节点称为备份节点,备份命令由类型和参数构成,类型为备份,参数为备份目标路径,备份目标路径由备份目标及其在备份节点上的存储位置构成;包括下述子步骤:

A1. 备份节点在自身开辟一块字节数为L的内存作为缓冲池,用于存储待发送数据块,然后同时启动子步骤A2和子步骤A3;

A2. 备份节点从备份目标按定长字节数 $L1$ 分块读取数据,依次存入缓冲池,直到整个备份目标读取完毕,备份目标中某个文件最后一块数据字节数不足 $L1$ 也作为一个数据块读取, $2KB \leq L1 \leq 4MB, L1 \leq L \leq 6L1$ ;

A3. 备份节点判断缓冲池里是否有数据块,是则转于步骤A5,否则进行子步骤A4;

A4. 判断子步骤A2是否结束,是则转于步骤A7,否则等待子步骤A2读取数据块到缓冲池后,进行子步骤A5;

A5. 备份节点按存入次序,从缓冲池里依次读取并清除一个数据块,通过SHA1哈希算法

或者MD5哈希算法对该数据块产生一个摘要,作为该数据块标识,备份节点根据数据块标识和备份节点标识,选择主服务节点,对其传输该数据块及其数据块标识,同时向主服务节点发送副本传递指示;主服务节点将接收到的数据块作为一个文件存储到自身的块文件目录下,文件名为数据块标识加后缀“.dat”,并按照副本传递指示,根据数据块标识和主服务节点标识,选择副本服务节点,对其传输该数据块及其数据块标识,并向备份节点返回副本服务节点的标识;副本服务节点也将接收到的数据块作为一个文件存储到自身的块文件目录下,文件名为数据块标识加后缀“.dat”,进行子步骤A6;

A6. 备份节点将该数据块标识以及该数据块对应的主服务节点标识和副本服务节点标识,按照对应关系放入备份节点的数据库中,返回子步骤A3;

所述主服务节点和副本服务节点均定期将接受到的数据块及其数据块标识发送给所在P2P覆盖网络的超级节点,超级节点将接收到的数据块作为一个文件存储到自身的块文件目录下,文件名为数据块标识加后缀“.dat”;超级节点定期将收集到的数据块发送到中心存储器;

A7. 本次备份作业的所有数据块备份完毕,确定该次备份作业号,并产生元数据文件,将其存储在本地磁盘元数据文件目录下,再进行子步骤A8;元数据文件的文件名为备份作业号加后缀“.dat”,元数据文件的内容包括本次备份作业的总文件数以及每个文件的文件名、文件路径以及各个文件的每个数据块的数据块信息,数据块信息包括:块号、偏移、块长和数据块标识;块号指该块属于原文件的第几块,偏移指该块的第一个字节相对于文件第一个字节的距离;

A8. 备份节点将所述元数据文件按照定长字节数L1分块,针对每一个数据块,生成数据块标识,根据各数据块标识和备份节点标识,分别选择主服务节点,对其传输数据块,同时向主服务节点发送副本传递指示;所述主服务节点根据副本传递指示,同时根据各数据块标识和主服务节点标识,分别选择副本服务节点,向副本服务节点传递相应数据块,并向备份节点返回副本服务节点标识,备份节点将各数据块标识以及该数据块对应的主服务节点标识和副本服务节点标识,按照对应关系放入备份节点的数据库中;

A9. 所述元数据文件的所有数据块备份完毕,备份节点将该次备份作业号和元数据文件的各数据块标识的对应关系按次序发送给超级节点,超级节点将这些信息写入自身磁盘作业信息文件目录下的一个作业信息文件中,该作业信息文件名为备份作业号+后缀“.lst”;至此备份作业步骤完成;

B. 所述数据恢复作业步骤:

某个P2P覆盖网络中的某个对等节点向本P2P覆盖网络中其它对等节点发出恢复命令,发起恢复命令的对等节点称为恢复节点,恢复命令由类型和参数构成,类型为恢复,参数为恢复节点先前作为备份节点的某次备份作业号;包括下述子步骤:

B1. 恢复节点根据恢复命令中的备份作业号,在本地磁盘的元数据文件目录下查找与此次恢复作业相关的元数据文件是否存在,是则转于步骤B2,否则转于步骤B11;

B2. 解析元数据文件,根据文件名和文件路径依次创建各个待恢复文件,然后对每个待恢复文件进行子步骤B3,直至元数据文件中记录的每个文件都恢复完毕,本次恢复作业步骤完成;

B3. 依次读取元数据文件中各个文件的每个数据块的数据块信息,每次读出数据块信

息后,进行子步骤B4,直至该文件的所有数据块信息均读出完毕,该文件恢复完毕,转于步骤B2;

B4. 恢复节点根据数据块信息中的数据块标识,到本地数据库中查找到主服务节点标识,判断该主服务节点是否在线,是则进行子步骤B5,否则转于步骤B6;

B5. 连接该主服务节点,并向其发送恢复数据块请求,请求参数为所述数据块标识,主服务节点根据该数据块标识在本地块文件目录下读取相应的数据块传递给恢复节点,恢复节点将该数据块写入待恢复文件,转于步骤B3;

B6. 根据数据块信息中的数据块标识,从本地数据库中找到该数据块的副本服务节点标识,判断该副本服务节点是否在线,是则进行子步骤B7,否则转于步骤B8;

B7. 连接该副本服务节点,并向其发送恢复数据块请求,请求参数为所述数据块标识,副本服务节点根据该数据块标识在本地块文件目录下读取相应的数据块文件传递给恢复节点,恢复节点将该数据块写入待恢复文件,转于步骤B3;

B8. 连接超级节点,向超级节点发送恢复数据块请求,请求参数为所述数据块标识,超级节点判断是否缓存有对应的数据块,是则进行子步骤B9,否则转于步骤B10;

B9. 超级节点根据该数据块标识在本地块文件目录下读取相应的数据块文件传递给恢复节点,恢复节点将该数据块写入待恢复文件,转于步骤B3;

B10. 超级节点根据该数据块标识,在本地数据库中查找到该数据块标识在中心存储器中的对应新数据块标识,向中心存储器发送恢复数据块请求,请求参数为新数据块标识,中心存储器根据该新数据块标识向超级节点传递所述数据块,之后超级节点再将所获取的数据块传输给恢复节点,恢复节点将该数据块写入待恢复文件,转于步骤B3;

B11. 根据恢复命令中的备份作业号,恢复节点在本地元数据文件目录下创建该作业号对应的元数据文件作为待恢复元数据文件,并从超级节点获取本次恢复作业相关的元数据文件的各数据块标识,进行子步骤B12;

B12. 对于每个数据块标识,进行子步骤B13,直至该元数据文件的所有数据块均写入待恢复元数据文件,元数据文件恢复完成,转于步骤B2;

B13. 恢复节点根据该数据块标识,到本地数据库中查找到主服务节点标识,判断该主服务节点是否在线,是则进行子步骤B14,否则转于步骤B15;

B14. 连接该主服务节点,并向其发送恢复数据块请求,请求参数为所述数据块标识,主服务节点根据该数据块标识在本地块文件目录下读取相应的数据块文件传递给恢复节点,恢复节点将该数据块写入待恢复元数据文件,转于步骤B12;

B15. 恢复节点根据数据块信息中的数据块标识,从本地数据库中找到该数据块的副本服务节点标识,判断该副本服务节点是否在线,是则进行子步骤B16,否则转于步骤B17;

B16. 连接该副本服务节点,并向其发送恢复数据块请求,请求参数为所述数据块标识,副本服务节点根据该数据块标识在本地块文件目录下读取相应的数据块文件传递给恢复节点,恢复节点将该数据块写入待恢复元数据文件,转于步骤B12;

B17. 连接超级节点,向超级节点发送恢复数据块请求,请求参数为所述数据块标识,超级节点判断是否缓存有对应的数据块,是则进行子步骤B18,否则转于步骤B19;

B18. 超级节点根据该数据块标识在本地块文件目录下读取相应的数据块文件传递给恢复节点,恢复节点将该数据块写入待恢复元数据文件,转于步骤B12;

B19. 超级节点在本地数据库中查找到该数据块标识在中心存储器中的对应新数据块标识,向中心存储器发送恢复数据块请求,请求参数为新数据块标识,中心存储器根据该新数据块标识向超级节点传递所述数据块,之后超级节点再将所获取的数据块传输给恢复节点,恢复节点将该数据块写入待恢复元数据文件,转至步骤B12。

2. 如权利要求1所述的云备份系统数据备份方法,其特征在于:

所述数据备份作业步骤的子步骤A5中,备份节点根据数据块标识和备份节点标识,选择主服务节点的过程为:对数据块标识和备份节点标识进行异或二进制运算,得到二者距离,根据二者距离,定位到备份节点的某个K桶,再从该K桶中选择第一个对等节点作为主服务节点;

主服务节点根据数据块标识和主服务节点标识,选择副本服务节点的过程为:对数据块标识和主服务节点标识进行异或二进制运算,得到二者距离,根据二者距离,定位到主服务节点的某个K桶,再从该K桶中选择第一个对等节点作为副本服务节点。

3. 如权利要求1或2所述的云备份系统数据备份方法,其特征在于:

所述数据备份作业步骤的子步骤A6中,超级节点将收集到的数据块发送到中心存储器的过程为:

首先由中心存储器的跟踪器之一向超级节点返回某组存储节点内任一个存储节点信息;然后超级节点向该存储节点发送数据块,该存储节点存储该数据块后,计算产生新数据块标识,并将该新数据块标识返回给超级节点,该新数据块标识用作超级节点需要向中心存储器获取该数据块时提供索引;所述存储节点再将该数据块复制到本组内其他各存储节点。

## 一种云备份系统及其数据备份方法

### 技术领域

[0001] 本发明属于计算机数据备份系统,具体涉及一种云备份系统及其数据备份方法。

### 背景技术

[0002] 现有备份系统一般由客户端、元数据服务器和存储节点组成,其中客户端为备份、恢复作业的发起端,元数据服务器为控制中心,负责用户注册、作业调度、存储节点管理等管理沟通工作,存储节点则负责数据的最终存储,三者通过传输控制协议(TCP)进行网络通信。这样一种简单的三方架构其扩展性、容错性和效率都不甚理想。

[0003] 数据的爆炸式增长对数据归档与备份提出越来越高的要求,高扩展性、高容错性、高效率成为数据备份系统越来越看重的特点,把P2P网络引入存储备份系统被认为是一种极富前景的应用,然而P2P网络的高抖动性也对数据的可用性构成威胁。

[0004] 按照其所适用的环境,P2P网络分为两大类:封闭式系统和开放式系统;封闭式系统是指系统有较为严格的中心认证、审计及管理功能,保证系统中节点的持续运行,封闭系统中的节点相对较为稳定,虽然可能出现暂时错误,但是会及时修复并重新回到系统中,并且不会随意退出系统,节点之间非常协作。封闭式系统仅适用于非常理想的情况,比如仅用于一个局域网少数几个机器,实际应用并不多见。

[0005] 开放式系统是指每个节点可以随意加入或退出系统,节点不保证持续在线提供服务的系统,开放式系统中的临时错误(节点暂不在线)和永久错误(节点退出系统)相对于封闭系统更加频繁,节点之间的合作度很低,甚至有不少理性用户的存在。虽然开放式系统存在更多问题和挑战,但实际应用中大多为开放式系统。

[0006] 由于P2P网络的抖动性,必须提供一定的数据冗余方案。在P2P存储系统中,目前主要有两种提供数据冗余的方式:多副本和纠删码。多副本顾名思义就是保存多个存储数据的完整副本,要占用更多存储空间,但实现简单。纠删码是指将要存储的数据先切分成 $m$ 个部分,然后通过编码算法将 $m$ 个部分变换为 $n$ 个部分, $n > m$ , $n$ 个部分中任意 $t$ 个部分可以用来恢复数据, $t \geq m$ ;相对于多副本方式,纠删码造成一定的计算量,也增加系统设计和实现的复杂度。

[0007] 为了充分理解本发明,以下对涉及到的概念进行说明:

[0008] kademia协议:简称Kad协议,属于一种典型的结构化P2P覆盖网络协议,以分布式的应用层全网方式进行信息的存储和检索。

[0009] OSI参考模型:开放式系统互联参考模型,是国际标准化组织(ISO)在1985年研究的网络互联模型,该参考模型定义了网络互连的七层框架,分别是:物理层,数据链路层,网络层,传输层,会话层,表示层,应用层。

[0010] UDP:用户数据报协议,为User Datagram Protocol的简称,是OSI参考模型中一种无连接的传输层通信协议,提供面向事务的简单不可靠信息传送服务。

[0011] TCP:传输控制协议,为Transmission Control Protocol的简称,是OSI参考模型中一种面向连接的、基于字节流的传输层通信协议。

[0012] 端口:传输层使用一个16位的端口号来标识一个端口,它的作用是标志本计算机应用层中各个进程在和传输层交互时的接口。

[0013] 距离:在使用Kad协议的P2P覆盖网络中,每个节点都有一个160bit的标识,两个节点的距离由对两个节点的标识进行异或二进制运算获得。

[0014] K桶:在使用Kad协议的P2P覆盖网络中,每一个对等节点均维护了160个链表(list),其中每个链表均被称之为一个k桶(k-bucket),在当前对等节点的第i个链表中,用k个条目,分别记录了与自身距离为 $2^i \sim 2^{i+1}$ 的其他k个对等节点的网络信息,每个条目包括节点标识,IP地址,UDP端口, $0 \leq i \leq 159, k \leq 10$ 。

[0015] 备份目标:一次备份作业的对象,为对等节点本地磁盘的一个目录或者一个文件,一个目录可以包含多个子目录和多个文件。

[0016] 主服务节点:针对一个数据块,备份节点按照一致性哈希方法从K桶中选择一个对等节点就叫这个数据块的主服务节点。

[0017] 副本服务节点:针对一个数据块,主服务节点按照备份节点指示,在需要传递副本时按照一致性哈希方法在自身的K桶中选择一个对等节点叫做这个数据块的副本服务节点。

[0018] SHA1哈希算法:是美国国家安全局(NSA)设计,美国国家标准与技术研究院(NIST)发布的一系列密码散列函数,SHA1是其中一个版本,对于长度小于 $2^{64}$ 位的数据,进行一系列位运算,产生一个160位的摘要,用来作为所述数据的标识。

[0019] MD5哈希算法:一种著名的哈希算法,通过一系列计算,可以将任意长度的数据的二进制值映射为长度为128位的二进制值,这个小的二进制值称为哈希值,可以用来作为所述数据的标识。

[0020] 备份作业号:用来标识一次备份作业,与元数据文件具有一一对应关系,采用备份目标的路径名。

[0021] FastDFS:是一个轻量级的开源分布式文件系统,主要解决了大容量的文件存储和高并发访问的问题,文件存取时实现了负载均衡。

## 发明内容

[0022] 本发明提供一种云备份系统,同时提供其数据备份方法,解决现有备份系统及数据备份方法所存在的存储层次单一,扩展性、容错性、效率不高的问题。

[0023] 本发明所提供的一种云备份系统,包括中心存储器、M个超级节点和M个P2P覆盖网络, $M \geq 1$ ;其特征在于:

[0024] 所述中心存储器包括1~2个跟踪器和1~8组存储节点,每组存储节点由2~3个存储节点构成,每个跟踪器分别连接各存储节点;

[0025] 所述中心存储器的跟踪器和各存储节点分别与各超级节点通过网络连接,M个超级节点分别连接自身的P2P覆盖网络,各P2P覆盖网络由多个对等节点按照kademilia协议自组织而成;

[0026] 每个对等节点都有一个节点标识,该节点标识由SHA1哈希算法或者MD5哈希算法计算由该节点IP地址、端口、主机名串联起来的信息而获得,每个对等节点具有内存和本地磁盘,在自身内存中维护160个K桶,在本地磁盘上维护一个数据库、一个配置文件、一个引

导文件、一个块文件目录、一个元数据文件目录；所述数据库用于保存数据块和服务节点的对应关系；所述配置文件用来记录超级节点的IP地址和端口，自己对外服务的端口，块文件目录路径和元数据文件目录路径；所述引导文件用来在磁盘上动态记录K桶中对等节点的网络信息，最初只有超级节点的网络信息，网络信息包括IP地址和端口；所述块文件目录用于该对等节点作为主服务节点或副本服务节点时存放所接收的数据块文件，文件名为数据块标识加后缀“.dat”；所述元数据文件目录用于该对等节点作为备份节点时存放每次备份作业产生的元数据文件；

[0027] 各超级节点之间通过网络连接，各超级节点具有内存和本地磁盘，在本地磁盘维护一个数据库、一个块文件目录和一个作业信息文件目录，所述数据库用于保存超级节点在各对等节点上收集的数据块标识和将其发送到中心存储器时由中心存储器返回的新数据块标识的对应关系；所述块文件目录用于该超级节点存放从相关节点所接收的数据块；所述作业信息文件目录用于存放各个备份作业相关的作业信息文件，作业信息文件用来记录各个对等节点的各个备份作业号和其相对应的元数据文件各文件块标识的对应关系；

[0028] 所述对等节点能够作为备份节点、主服务节点或副本服务节点，所述对等节点也能够作为恢复节点。

[0029] 所述P2P覆盖网络作用在于：各对等节点启动时首先向该组超级节点注册信息，并获取组内其他对等节点信息，根据此信息形成一个动态的P2P自组织覆盖网络，之后组内各对等节点可自由发起备份或恢复操作，数据在组内各对等节点中采取多副本形式存储，以满足抖动P2P网络的数据可用性。所述超级节点负责维护该P2P覆盖网络组成员相关信息，同时负责缓存各对等节点向其发送的数据块，收集的所有数据块最终也要存储到中心存储器上。中心存储器采用分组方式存储，组内数据互备。

[0030] 本发明所述云备份系统的数据备份方法，包括数据备份作业步骤和数据恢复作业步骤，其特征在于：

[0031] A.所述数据备份作业步骤：

[0032] 某个P2P覆盖网络中的某个对等节点向本P2P覆盖网络中其它对等节点发出备份命令，发起备份命令的对等节点称为备份节点，备份命令由类型和参数构成，类型为备份，参数为备份目标路径，备份目标路径由备份目标及其在备份节点上的存储位置构成；包括下述子步骤：

[0033] A1. 备份节点在自身开辟一块字节数为L的内存作为缓冲池，用于存储待发送数据块，然后同时启动子步骤A2和子步骤A3；

[0034] A2. 备份节点从备份目标按定长字节数L1分块读取数据，依次存入缓冲池，直到整个备份目标读取完毕，备份目标中某个文件最后一块数据字节数不足L1也作为一个数据块读取， $2KB \leq L1 \leq 4MB$ ， $L1 \leq L \leq 6L1$ ；

[0035] A3. 备份节点判断缓冲池里是否有数据块，是则转于步骤A5，否则进行子步骤A4；

[0036] A4. 判断子步骤A2是否结束，是则转于步骤A7，否则等待子步骤A2读取数据块到缓冲池后，进行子步骤A5；

[0037] A5. 备份节点按存入次序，从缓冲池里依次读取并清除一个数据块，通过SHA1哈希算法或者MD5哈希算法对该数据块产生一个摘要，作为该数据块标识，备份节点根据数据块标识和备份节点标识，选择主服务节点，对其传输该数据块及其数据块标识，同时向主服务



节点发送副本传递指示;主服务节点将接收到的数据块作为一个文件存储到自身的块文件目录下,文件名为数据块标识加后缀“.dat”,并按照副本传递指示,根据数据块标识和主服务节点标识,选择副本服务节点,对其传输该数据块及其数据块标识,并向备份节点返回副本服务节点的标识;副本服务节点也将接收到的数据块作为一个文件存储到自身的块文件目录下,文件名为数据块标识加后缀“.dat”,进行子步骤A6;

[0038] A6.备份节点将该数据块标识以及该数据块对应的主服务节点标识和副本服务节点标识,按照对应关系放入备份节点的数据库中,返回子步骤A3;

[0039] 所述主服务节点和副本服务节点均定期将接受到的数据块及其数据块标识发送给所在P2P覆盖网络的超级节点,超级节点将接收到的数据块作为一个文件存储到自身的块文件目录下,文件名为数据块标识加后缀“.dat”;超级节点定期将收集到的数据块发送到中心存储器;

[0040] A7.本次备份作业的所有数据块备份完毕,确定该次备份作业号,并产生元数据文件,将其存储在本地磁盘元数据文件目录下,再进行子步骤A8;元数据文件的文件名为备份作业号加后缀“.dat”,元数据文件的内容包括本次备份作业的总文件数以及每个文件的文件名、文件路径以及各个文件的每个数据块的数据块信息,数据块信息包括:块号、偏移、块长和数据块标识;块号指该块属于原文件的第几块,偏移指该块的第一个字节相对于文件第一个字节的距离;

[0041] A8.备份节点将所述元数据文件按照定长字节数L1分块,针对每一个数据块,生成数据块标识,根据各数据块标识和备份节点标识,分别选择主服务节点,对其传输数据块,同时向主服务节点发送副本传递指示;所述主服务节点根据副本传递指示,同时根据各数据块标识和主服务节点标识,分别选择副本服务节点,向副本服务节点传递相应数据块,并向备份节点返回副本服务节点标识,备份节点将各数据块标识以及该数据块对应的主服务节点标识和副本服务节点标识,按照对应关系放入备份节点的数据库中;

[0042] A9.所述元数据文件的所有数据块备份完毕,备份节点将该次备份作业号和元数据文件的各数据块标识的对应关系按次序发送给超级节点,超级节点将这些信息写入自身磁盘作业信息文件目录下的一个作业信息文件中,该作业信息文件名为备份作业号+后缀“.lst”;至此备份作业步骤完成;

[0043] B.所述数据恢复作业步骤:

[0044] 某个P2P覆盖网络中的某个对等节点向本P2P覆盖网络中其它对等节点发出恢复命令,发起恢复命令的对等节点称为恢复节点,恢复命令由类型和参数构成,类型为恢复,参数为恢复节点先前作为备份节点的某次备份作业号;包括下述子步骤:

[0045] B1.恢复节点根据恢复命令中的备份作业号,在本地磁盘的元数据文件目录下查找与此次恢复作业相关的元数据文件是否存在,是则转子步骤B2,否则转子步骤B11;

[0046] B2.解析元数据文件,根据文件名和文件路径依次创建各个待恢复文件,然后对每个待恢复文件进行子步骤B3,直至元数据文件中记录的每个文件都恢复完毕,本次恢复作业步骤完成;

[0047] B3.依次读取元数据文件中各个文件的每个数据块的数据块信息,每次读出数据块信息后,进行子步骤B4,直至该文件的所有数据块信息均读出完毕,该文件恢复完毕,转子步骤B2;

- [0048] B4. 恢复节点根据数据块信息中的数据块标识,到本地数据库中查找到主服务节点标识,判断该主服务节点是否在线,是则进行子步骤B5,否则转子步骤B6;
- [0049] B5. 连接该主服务节点,并向其发送恢复数据块请求,请求参数为所述数据块标识,主服务节点根据该数据块标识在本地块文件目录下读取相应的数据块传递给恢复节点,恢复节点将该数据块写入待恢复文件,转子步骤B3;
- [0050] B6. 根据数据块信息中的数据块标识,从本地数据库中找到该数据块的副本服务节点标识,判断该副本服务节点是否在线,是则进行子步骤B7,否则转子步骤B8;
- [0051] B7. 连接该副本服务节点,并向其发送恢复数据块请求,请求参数为所述数据块标识,副本服务节点根据该数据块标识在本地块文件目录下读取相应的数据块文件传递给恢复节点,恢复节点将该数据块写入待恢复文件,转子步骤B3;
- [0052] B8. 连接超级节点,向超级节点发送恢复数据块请求,请求参数为所述数据块标识,超级节点判断是否缓存有对应的数据块,是则进行子步骤B9,否则转子步骤B10;
- [0053] B9. 超级节点根据该数据块标识在本地块文件目录下读取相应的数据块文件传递给恢复节点,恢复节点将该数据块写入待恢复文件,转子步骤B3;
- [0054] B10. 超级节点根据该数据块标识,在本地数据库中查找到该数据块标识在中心存储器中的对应新数据块标识,向中心存储器发送恢复数据块请求,请求参数为新数据块标识,中心存储器根据该新数据块标识向超级节点传递所述数据块,之后超级节点再将所获取的数据块传输给恢复节点,恢复节点将该数据块写入待恢复文件,转子步骤B3;
- [0055] B11. 根据恢复命令中的备份作业号,恢复节点在本地元数据文件目录下创建该作业号对应的元数据文件作为待恢复元数据文件,并从超级节点获取本次恢复作业相关的元数据文件的各数据块标识;
- [0056] B12. 对于每个数据块标识,进行子步骤B13,直至该元数据文件的所有数据块均写入待恢复元数据文件,元数据文件恢复完成,转子步骤B2;
- [0057] B13. 恢复节点根据该数据块标识,到本地数据库中查找到主服务节点标识,判断该主服务节点是否在线,是则进行子步骤B14,否则转子步骤B15;
- [0058] B14. 连接该主服务节点,并向其发送恢复数据块请求,请求参数为所述数据块标识,主服务节点根据该数据块标识在本地块文件目录下读取相应的数据块文件传递给恢复节点,恢复节点将该数据块写入待恢复元数据文件,转子步骤B12;
- [0059] B15. 恢复节点根据数据块信息中的数据块标识,从本地数据库中找到该数据块的副本服务节点标识,判断该副本服务节点是否在线,是则进行子步骤B16,否则转子步骤B17;
- [0060] B16. 连接该副本服务节点,并向其发送恢复数据块请求,请求参数为所述数据块标识,副本服务节点根据该数据块标识在本地块文件目录下读取相应的数据块文件传递给恢复节点,恢复节点将该数据块写入待恢复元数据文件,转子步骤B12;
- [0061] B17. 连接超级节点,向超级节点发送恢复数据块请求,请求参数为所述数据块标识,超级节点判断是否缓存有对应的数据块,是则进行子步骤B18,否则转子步骤B19;
- [0062] B18. 超级节点根据该数据块标识在本地块文件目录下读取相应的数据块文件传递给恢复节点,恢复节点将该数据块写入待恢复元数据文件,转子步骤B12;
- [0063] B19. 超级节点在本地数据库中查找到该数据块标识在中心存储器中的对应新数

据块标识,向中心存储器发送恢复数据块请求,请求参数为新数据块标识,中心存储器根据该新数据块标识向超级节点传递所述数据块,之后超级节点再将所获取的数据块传输给恢复节点,恢复节点将该数据块写入待恢复元数据文件,转至步骤B12。

[0064] 所述的数据备份方法,其特征在于:

[0065] 所述数据备份作业步骤的子步骤A5中,备份节点根据数据块标识和备份节点标识,选择主服务节点的过程为:对数据块标识和备份节点标识进行异或二进制运算,得到二者距离,根据二者距离,定位到备份节点的某个K桶,再从该K桶中选择第一个对等节点作为主服务节点;

[0066] 主服务节点根据数据块标识和主服务节点标识,选择副本服务节点的过程为:对数据块标识和主服务节点标识进行异或二进制运算,得到二者距离,根据二者距离,定位到主服务节点的某个K桶,再从该K桶中选择第一个对等节点作为副本服务节点。

[0067] 所述的数据备份方法,其特征在于:

[0068] 所述数据备份作业步骤的子步骤A6中,超级节点将收集到的数据块发送到中心存储器的过程为:

[0069] 首先由中心存储器的跟踪器之一向超级节点返回某组存储节点内任一个存储节点信息;然后超级节点向该存储节点发送数据块,该存储节点存储该数据块后,计算产生新数据块标识,并将该新数据块标识返回给超级节点,该新数据块标识用作超级节点需要向中心存储器获取该数据块时提供索引;所述存储节点再将该数据块复制到本组内其他各存储节点。

[0070] 本发明针对开放式系统,为了降低复杂度,本发明采用多副本冗余方式。所述主服务节点和副本服务节点均定期将接收到的数据块传递给所在P2P覆盖网络的超级节点,超级节点定期将收集到的数据发送到中心存储器,这样数据在本系统中就形成P2P覆盖网络存储、超级节点缓存、中心存储器数据管理三级存储,P2P存储主要是提高备份恢复性能;超级节点缓存是为了收集本组所有数据并最终发送到中心存储器;中心存储器则做数据的最终保障。这样一种三层存储架构,提供高扩展性,同时兼顾用户的备份速度,并保证了数据的高可用性,为企业或用户提供高扩展性、高容错性的云备份服务。

## 附图说明

[0071] 图1为本发明的系统结构示意图;

[0072] 图2为中心存储器结构示意图;

[0073] 图3为数据备份作业步骤流程框图;

[0074] 图4为数据恢复作业步骤流程框图。

## 具体实施方式

[0075] 下面结合附图和实施例对本发明进一步说明。

[0076] 如图1所示,本发明云备份系统的一个实施例,包括中心存储器、4个超级节点和4个P2P覆盖网络;所述中心存储器包括2个跟踪器和2组存储节点,每组存储节点由2个存储节点构成,每个跟踪器分别连接各存储节点;

[0077] 所述中心存储器的跟踪器和各存储节点分别与各超级节点通过网络连接,4个超

级节点分别连接自身的P2P覆盖网络,各P2P覆盖网络由3个对等节点按照kademilia协议自组织而成;

[0078] 如图2所示,中心存储器采用FastDFS架构,本实施例包括两个跟踪器(跟踪器1和跟踪器2)和两组存储节点,每组存储节点包括两个存储节点,第1组存储节点包括存储节点11和包括存储节点12,第2组存储节点包括存储节点21和包括存储节点22。跟踪器作为管理节点,其主要作用是负载均衡和调度,并在内存中记录存储节点分组和各个存储节点的状态信息.存储节点存储文件,完成文件管理的所有功能。同组内的多个存储节点之间是互备关系,即其上的数据文件是完全一致的。

[0079] 本发明方法的实施例,包括数据备份作业步骤和数据恢复作业步骤;

[0080] 如图3所示,数据备份作业步骤流程如下:某个P2P覆盖网络中的某个对等节点向本P2P覆盖网络中其它对等节点发出备份命令,启动备份作业步骤:该备份节点在内存中开辟一个缓冲池,用于存放待发送数据块,同时启动生产者线程和消费者线程。生产者线程负责从备份目标中按定长读取数据块到缓冲池,消费者线程负责从缓冲池中读取数据块发送到网络。针对每个数据块,消费者线程选择主服务节点传输数据块,所述主服务节点根据副本传递指示选择副本服务节点传输数据块。备份完一个数据块后,备份节点将数据块标识和主服务节点标识、副本服务节点标识存入本地数据库中。当所有数据块均备份完毕时,生成这次备份作业的元数据文件。备份节点将元数据文件也分块存储到P2P网络中。之后,备份节点将该次备份作业号和元数据文件的各数据块标识对应关系发送给超级节点。主服务节点和副本服务节点会定期将接收的数据块发送给超级节点,超级节点也会定期将收集的数据块发送给中心存储器,这样形成了数据在P2P覆盖网络、超级节点、中心存储器的三级存储。

[0081] 如图4所示,数据恢复作业步骤流程如下:某个P2P覆盖网络中的某个对等节点向本P2P覆盖网络中其它对等节点发出恢复命令,启动恢复作业步骤:首先判断此次恢复作业相关的元数据文件是否存在本地存在,若存在则直接解析,否则从超级节点获取元数据文件数据块信息并从P2P网络获取数据块,在本地恢复出元数据文件。根据元数据文件记录的信息,依次创建各个待恢复文件,针对每个待恢复文件,依次恢复各个数据块,针对每个数据块,首先判断主服务节点是否在线,如果主服务在线,则从主服务节点获取数据块,否则判断副本服务节点是否在线,若副本服务节点在线,则从副本服务节点获取数据块,否则询问超级节点是否缓存该数据块,若有则从超级节点获取数据块,否则超级节点从中心存储器获取数据块并传递给恢复节点。每个待恢复文件的每个数据块按照上述方法恢复完成后,此次恢复作业完成。

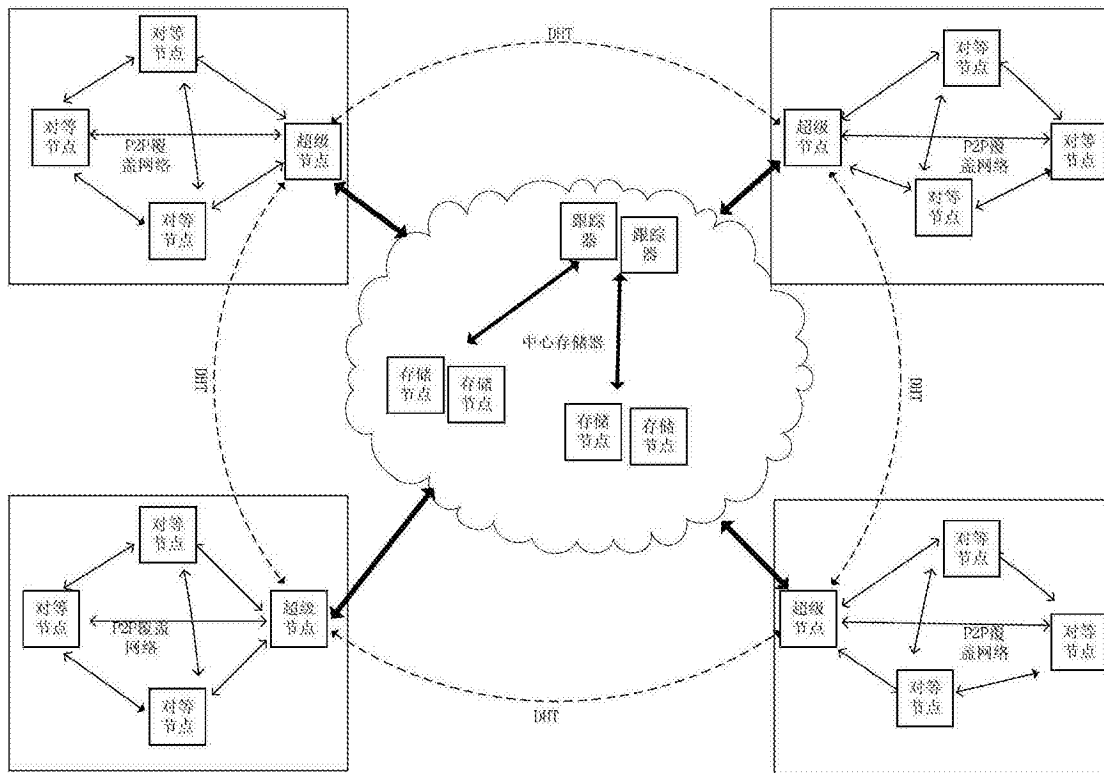


图1

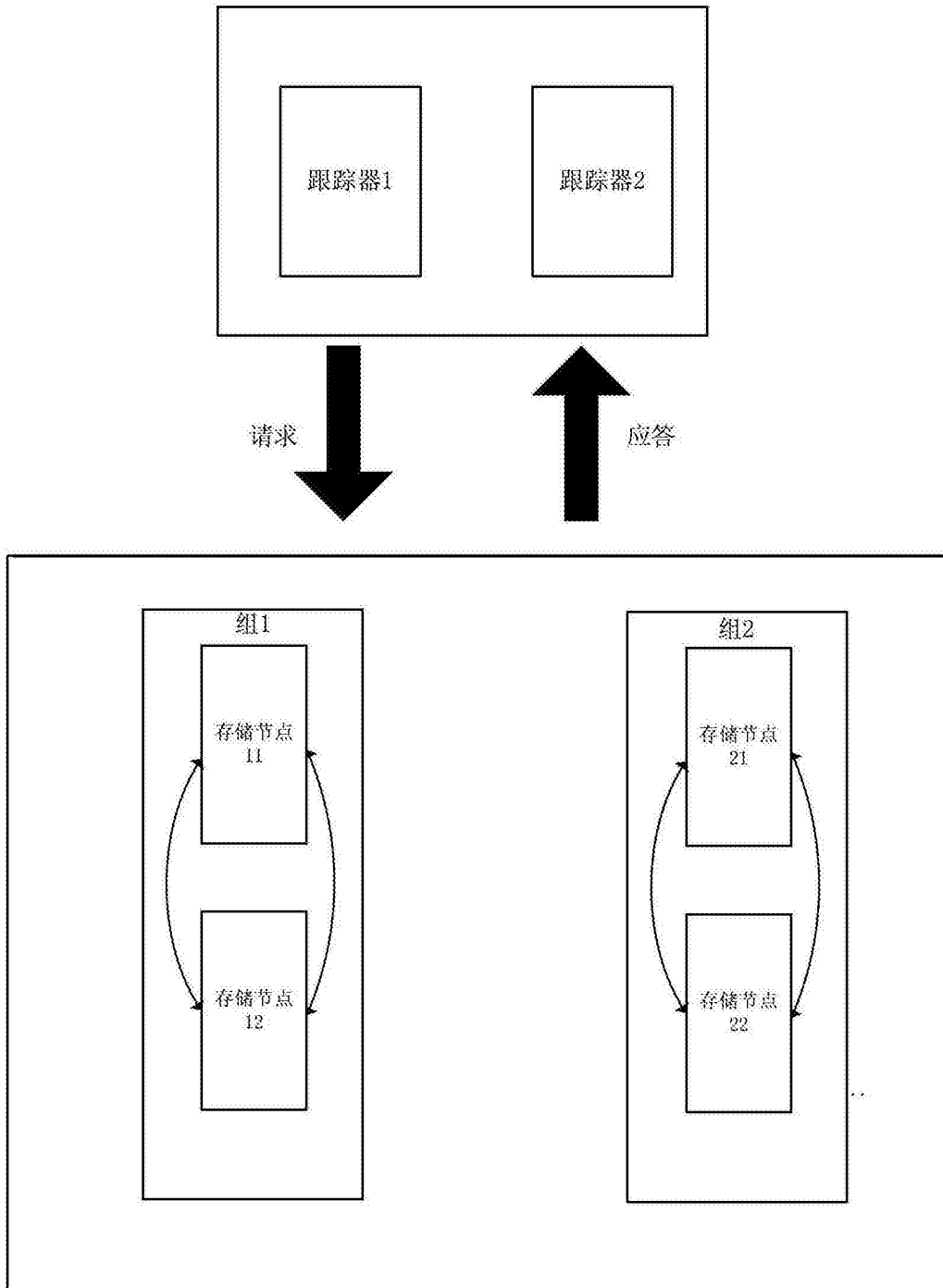


图2

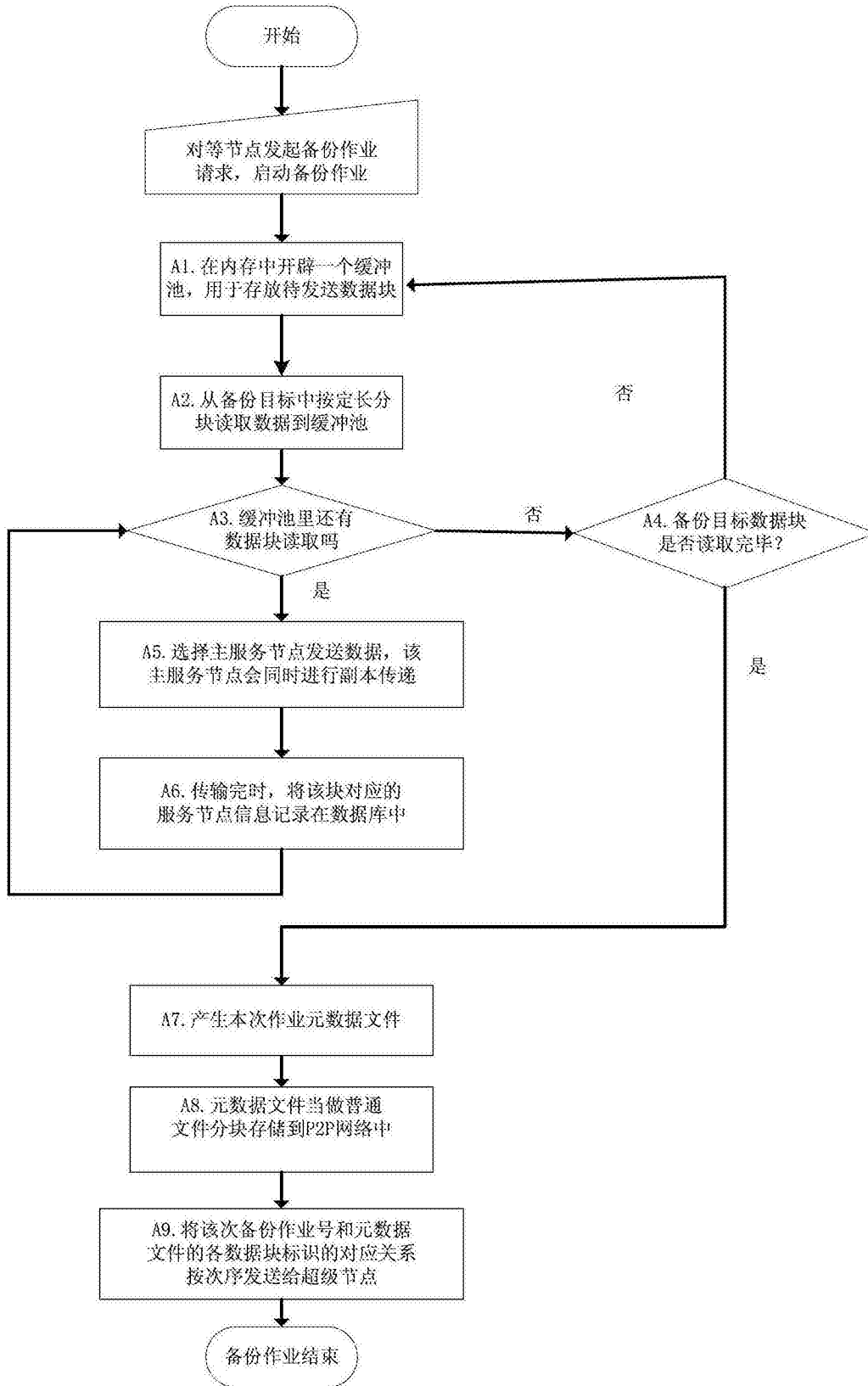


图3

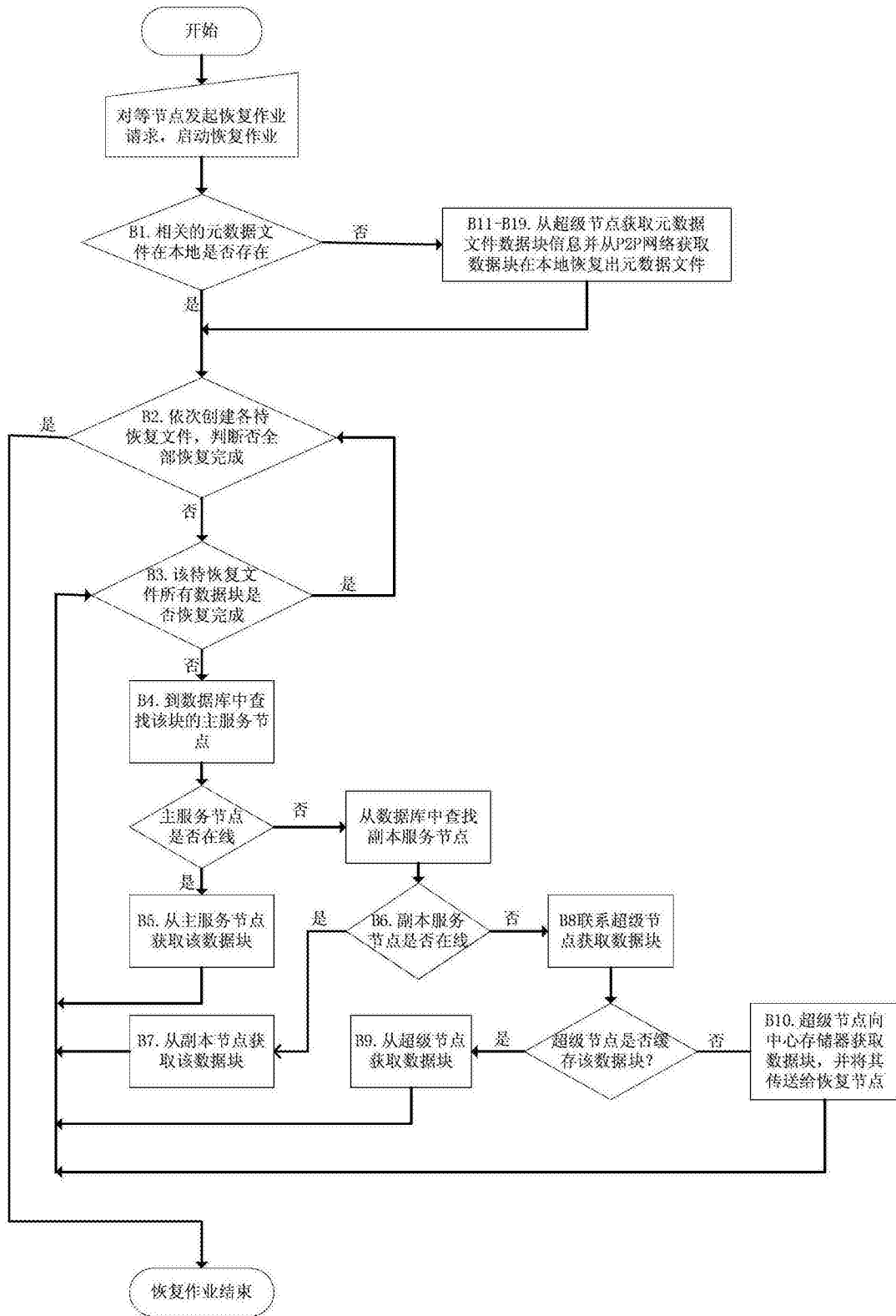


图4