



(12) 发明专利

(10) 授权公告号 CN 109407977 B

(45) 授权公告日 2021.08.31

(21) 申请号 201811119218.3

G06F 16/182 (2019.01)

(22) 申请日 2018.09.25

G06F 16/176 (2019.01)

(65) 同一申请的已公布的文献号
申请公布号 CN 109407977 A

(56) 对比文件

(43) 申请公布日 2019.03.01

CN 104320401 A, 2015.01.28

CN 106599308 A, 2017.04.26

(73) 专利权人 佛山科学技术学院
地址 528000 广东省佛山市南海区狮山镇
仙溪水库西路佛山科学技术学院

CN 102387179 A, 2012.03.21

CN 104657497 A, 2015.05.27

CN 105608155 A, 2016.05.25

CN 106156359 A, 2016.11.23

(72) 发明人 张彩霞 郭静 王向东 王新东
胡绍林

CN 103905537 A, 2014.07.02

CN 108268208 A, 2018.07.10

CN 103729436 A, 2014.04.16

(74) 专利代理机构 广州嘉权专利商标事务所有
限公司 44205

审查员 戴自立

代理人 王国标

(51) Int. Cl.

G06F 3/06 (2006.01)

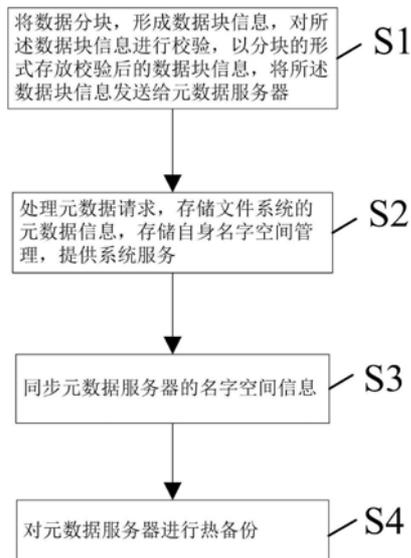
权利要求书2页 说明书4页 附图1页

(54) 发明名称

一种大数据分布式存储管理方法及系统

(57) 摘要

本发明涉及大数据存储技术领域,具体涉及一种大数据分布式存储管理方法及系统,通过将数据分块和校验,以分块的形式存放校验后的数据块信息,将所述数据块信息发送给元数据服务器;通过元数据服务器处理元数据请求,存储文件系统的元数据信息,存储自身名字空间管理,提供系统服务;将元数据信息以镜像文件和日志文件的方式持久化到共享存储池中;通过同步元数据服务器的名字空间信息从而使得集群中每个元数据服务器利用备用节点进行热备份,从而实现对元数据服务器的热备份,本发明能保证大数据存储访问过程中数据的可靠性。



1. 一种大数据分布式存储管理系统,其特征在於,所述系统包括数据节点、元数据服务器、共享存储池、备用节点;

所述数据节点,用于将数据分块,形成数据块信息,对所述数据块信息进行校验,以分块的形式存放校验后的数据块信息,将所述数据块信息发送给元数据服务器;

所述元数据服务器,用于处理元数据请求,存储文件系统的元数据信息,存储自身名字空间管理,提供系统服务;

所述共享存储池,用于同步元数据服务器的名字空间信息;

所述备用节点,用于对元数据服务器进行热备份;

其中,所述数据节点,用于定期对元数据服务器进行数据采集,将内存元数据信息和日志内容合并成一个新的镜像文件,并将所述镜像文件发送到共享存储池中进行存储;

所述元数据服务器还用于根据各数据节点信息及列表项配置信息生成数据节点列表,将数据存放到元数据服务器对应的数据节点,每台数据节点的项数按以下公式计算:

$$U_i = \left[(C - \sum_1^{i-1} u_i) / (n - i) \right]$$

其中, U_i 表示第*i*台数据节点在列表内所出现的次数, C 表示列表的项数, n 表示数据节点总数, U_i 即为每台数据节点的项数;

所述系统服务包括:通过将多个元数据服务器构成一个全局的名字空间,作为共享存储池中共享存储节点,保存元数据服务器镜像和日志文件的副本信息;

所述系统还包括切换模块,用于当元数据服务器处于离线状态时,将元数据服务器服务进程切换至备用节点。

2. 一种大数据分布式存储管理方法,其特征在於,所述方法基于权利要求1所述的一种大数据分布式存储管理系统,包括以下步骤:

步骤S1、将数据分块,形成数据块信息,对所述数据块信息进行校验,以分块的形式存放校验后的数据块信息,将所述数据块信息发送给元数据服务器;

步骤S2、处理元数据请求,存储文件系统的元数据信息,存储自身名字空间管理,提供系统服务;

步骤S3、同步元数据服务器的名字空间信息;

步骤S4、对元数据服务器进行热备份;

其中,所述步骤S1还包括:定期对元数据服务器进行数据采集,将内存元数据信息和日志内容合并成一个新的镜像文件,并将所述镜像文件发送到共享存储池中进行存储;

所述步骤S2还包括:根据各数据节点信息及列表项配置信息生成数据节点列表,将数据存放到元数据服务器对应的数据节点,每台数据节点的项数按以下公式计算:

$$U_i = \left[(C - \sum_1^{i-1} u_i) / (n - i) \right]$$

其中, U_i 表示第*i*台数据节点在列表内所出现的次数, C 表示列表的项数, n 表示数据节点总数, U_i 即为每台数据节点的项数;

所述系统服务包括:通过将多个元数据服务器构成一个全局的名字空间,作为共享存储池中共享存储节点,保存元数据服务器镜像和日志文件的副本信息;

所述方法还包括：当元数据服务器处于离线状态时，将元数据服务器服务进程切换至备用节点。

一种大数据分布式存储管理方法及系统

技术领域

[0001] 本发明涉及大数据存储技术领域,具体涉及一种大数据分布式存储管理方法及系统。

背景技术

[0002] 随着集群规模的扩大,在大数据分析和处理过程中,网络故障和节点失效等错误发生的概率增加,需要底层分布式数据系统提供高可用保证,保持数据系统服务的连续性。基于共享存储池,集群中每个元数据服务器利用影子节点SN进行热备,无需额外的数据系统或设备支持就能快速恢复故障。通常情况下,影子节点不提供元数据服务,处于备用状态,但始终维护与元数据服务器一致的名字空间。

[0003] 因此,如何提供一种高可用,保持数据系统服务的连续性的大数据分布式存储管理途径成为值得解决的问题。

发明内容

[0004] 本发明提供一种大数据分布式存储管理方法及系统,能够保证大数据存储访问过程中数据的可靠性。

[0005] 本发明提供的一种大数据分布式存储管理系统,所述系统包括数据节点、元数据服务器、共享存储池、备用节点;

[0006] 所述数据节点,用于将数据分块,形成数据块信息,对所述数据块信息进行校验,以分块的形式存放校验后的数据块信息,将所述数据块信息发送给元数据服务器;

[0007] 所述元数据服务器,用于处理元数据请求,存储文件系统的元数据信息,存储自身名字空间管理,提供系统服务;

[0008] 所述共享存储池,用于同步元数据服务器的名字空间信息;

[0009] 所述备用节点,用于对元数据服务器进行热备份。

[0010] 进一步,所述数据节点,用于定期对元数据服务器进行数据采集,将内存元数据信息和日志内容合并成一个新的镜像文件,并将所述镜像文件发送到共享存储池中进行存储。

[0011] 进一步,所述元数据服务器还用于根据各数据节点信息及列表项配置信息生成数据节点列表,将数据存放到元数据服务器对应的数据节点,每台数据节点的项数按以下公式计算:

$$[0012] \quad U_i = \left\lceil \frac{C - \sum_1^{i-1} u_i}{n - i} \right\rceil$$

[0013] 其中, U_i 表示第*i*台数据节点在列表内所出现的次数, C 表示列表的项数, n 表示数据节点总数, U_i 即为每台数据节点的项数。

[0014] 进一步,所述系统服务包括:通过将多个元数据服务器构成一个全局的名字空间,作为共享存储池中共享存储节点,保存元数据服务器镜像和日志文件的副本信息。

[0015] 进一步,所述系统还包括切换模块,用于当元数据服务器处于离线状态时,将元数据服务器服务进程切换至备用节点。

[0016] 本发明提供了一种大数据分布式存储管理方法,所述方法基于上述任一所述的一种大数据分布式存储管理系统,包括以下步骤:

[0017] 步骤S1、将数据分块,形成数据块信息,对所述数据块信息进行校验,以分块的形式存放校验后的数据块信息,将所述数据块信息发送给元数据服务器;

[0018] 步骤S2、处理元数据请求,存储文件系统的元数据信息,存储自身名字空间管理,提供系统服务;

[0019] 步骤S3、同步元数据服务器的名字空间信息;

[0020] 步骤S4、对元数据服务器进行热备份。

[0021] 进一步,所述步骤S1还包括:定期对元数据服务器进行数据采集,将内存元数据信息和日志内容合并成一个新的镜像文件,并将所述镜像文件发送到共享存储池中进行存储。

[0022] 进一步,所述步骤S2还包括:根据各数据节点信息及列表项配置信息生成数据节点列表,将数据存放到元数据服务器对应的数据节点,每台数据节点的项数按以下公式计算:

$$[0023] \quad U_i = \left\lfloor \left(C - \sum_1^{i-1} u_i \right) / (n - i) \right\rfloor$$

[0024] 其中, U_i 表示第*i*台数据节点在列表内所出现的次数, C 表示列表的项数, n 表示数据节点总数, U_1 即为每台数据节点的项数。

[0025] 进一步,所述系统服务包括:通过将多个元数据服务器构成一个全局的名字空间,作为共享存储池中共享存储节点,保存元数据服务器镜像和日志文件的副本信息。

[0026] 进一步,所述方法还包括:当元数据服务器处于离线状态时,将元数据服务器服务进程切换至备用节点。

[0027] 本发明的有益效果是:本发明公开一种大数据分布式存储管理方法及系统,通过将数据分块和校验,以分块的形式存放校验后的数据块信息,同步元数据服务器的名字空间信息;对元数据服务器进行热备份,从而保证大数据存储访问过程中数据的可靠性。

附图说明

[0028] 下面结合附图和实例对本发明作进一步说明。

[0029] 图1是本发明一种大数据分布式存储管理方法的流程示意图。

具体实施方式

[0030] 本发明提供了一种大数据分布式存储管理系统,所述系统包括数据节点、元数据服务器、共享存储池、备用节点;

[0031] 所述数据节点,用于将数据分块,形成数据块信息,对所述数据块信息进行校验,以分块的形式存放校验后的数据块信息,将所述数据块信息发送给元数据服务器;

[0032] 所述元数据服务器,用于处理元数据请求,存储文件系统的元数据信息,存储自身名字空间管理,提供系统服务;所述元数据信息包括数据索引节点号、数据属性、创建时间,

在逻辑上以树形结构组织,并以镜像文件和日志文件的方式保持到共享存储池中。

[0033] 所述共享存储池,用于同步元数据服务器的名字空间信息,从而使得集群中每个元数据服务器利用备用节点进行热备份。

[0034] 所述备用节点,用于对元数据服务器进行热备份。

[0035] 进一步,所述数据节点,用于定期对元数据服务器进行数据采集,将内存元数据信息和日志内容合并成一个新的镜像文件,并将所述镜像文件发送到共享存储池中进行存储。

[0036] 进一步,所述元数据服务器还用于根据各数据节点信息及列表项配置信息生成数据节点列表,将数据存放到元数据服务器对应的数据节点,以减少元数据的内存占用开销,每台数据节点的项数按以下公式计算:

$$[0037] \quad U_i = \left\lfloor (C - \sum_1^{i-1} u_i) / (n - i) \right\rfloor$$

[0038] 其中, U_i 表示第*i*台数据节点在列表内所出现的次数, C 表示列表的项数, n 表示数据节点总数, U_i 即为每台数据节点的项数。

[0039] 进一步,所述系统服务包括:通过将多个元数据服务器构成一个全局的名字空间,作为共享存储池中共享存储节点,保存其他元数据服务器镜像和日志文件的副本信息。

[0040] 进一步,所述系统还包括切换模块,用于当元数据服务器处于离线状态时,将元数据服务器服务进程切换至备用节点。

[0041] 参考图1,本发明提供一种大数据分布式存储管理方法,所述方法基于上述任一所述的一种大数据分布式存储管理系统,包括以下步骤:

[0042] 步骤S1、将数据分块,形成数据块信息,对所述数据块信息进行校验,以分块的形式存放校验后的数据块信息,将所述数据块信息发送给元数据服务器;

[0043] 步骤S2、处理元数据请求,存储文件系统的元数据信息,存储自身名字空间管理,提供系统服务;所述元数据信息包括数据索引节点号、数据属性、创建时间等,逻辑上以树形结构组织,并以镜像文件和日志文件的方式持久化到共享存储池中;

[0044] 步骤S3、同步元数据服务器的名字空间信息从而使得集群中每个元数据服务器利用备用节点进行热备份;

[0045] 步骤S4、对元数据服务器进行热备份。

[0046] 进一步,所述步骤S1还包括:定期对元数据服务器进行数据采集,将内存元数据信息和日志内容合并成一个新的镜像文件,并将所述镜像文件发送到共享存储池中进行存储。

[0047] 进一步,所述步骤S2还包括:根据各数据节点信息及列表项配置信息生成数据节点列表,将数据存放到元数据服务器对应的数据节点,以减少元数据的内存占用开销,每台数据节点的项数按以下公式计算:

$$[0048] \quad U_i = \left\lfloor (C - \sum_1^{i-1} u_i) / (n - i) \right\rfloor$$

[0049] 其中, U_i 表示第*i*台数据节点在列表内所出现的次数, C 表示列表的项数, n 表示数据节点总数, U_i 即为每台数据节点的项数。

[0050] 进一步,所述系统服务包括:通过将多个元数据服务器构成一个全局的名字空间,作为共享存储池中共享存储节点,保存其他元数据服务器镜像和日志文件的副本信息。

[0051] 进一步,所述方法还包括:当元数据服务器处于离线状态时,将元数据服务器服务进程切换至备用节点。

[0052] 以上所述,只是本发明的较佳实施例而已,本发明并不局限于上述实施方式,只要其以相同的手段达到本发明的技术效果,都应属于本发明的保护范围。

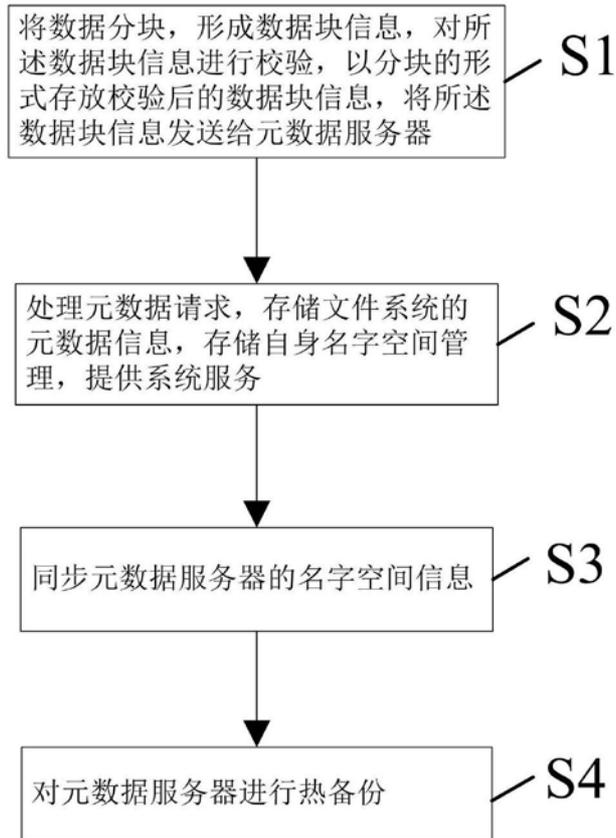


图1