



(12) 发明专利申请

(10) 申请公布号 CN 112530512 A

(43) 申请公布日 2021.03.19

(21) 申请号 202011608228.0

(22) 申请日 2020.12.29

(66) 本国优先权数据

202011257824.9 2020.11.11 CN

(71) 申请人 北京泽石科技有限公司

地址 100085 北京市海淀区上地东路1号院
1号楼2层203-1室

申请人 泽石科技(武汉)有限公司

(72) 发明人 张坤

(74) 专利代理机构 北京康信知识产权代理有限
责任公司 11240

代理人 董文倩

(51) Int.Cl.

G11C 29/56 (2006.01)

G06K 9/62 (2006.01)

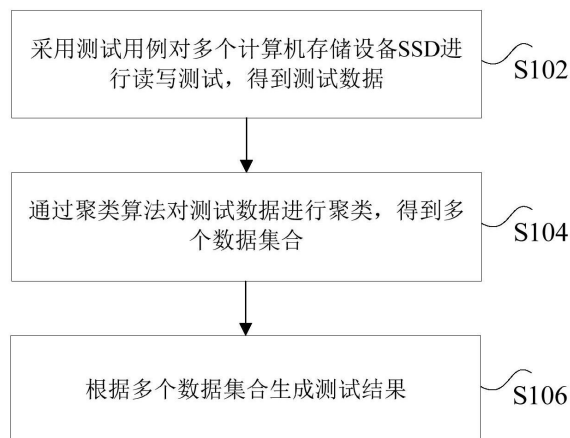
权利要求书2页 说明书9页 附图3页

(54) 发明名称

闪存存储设备的测试方法及装置

(57) 摘要

本发明公开了一种闪存存储设备的测试方法及装置。其中,该方法包括:采用测试用例对多个计算机闪存存储设备进行读写测试,得到测试数据;通过聚类算法对测试数据进行聚类,得到多个数据集合;根据多个数据集合生成测试结果。本发明解决了相关技术中通过专用的读写工具生成测试日志,人工分析测试日志确定闪存存储设备的读写速度,误差较大,效率较低的技术问题。



1. 一种闪存存储设备的测试方法,其特征在于,包括:
采用测试用例对多个计算机闪存存储设备进行读写测试,得到测试数据;
通过聚类算法对所述测试数据进行聚类,得到多个数据集合;
根据所述多个数据集合生成测试结果。
2. 根据权利要求1所述的方法,其特征在于,采用测试用例对多个计算机闪存存储设备进行读写测试,得到测试数据包括:
控制所述多个闪存存储设备运行所述测试用例进行读写测试;
确定所述多个闪存存储设备中成功完成所述测试用例的测试通过闪存存储设备;
获取所述测试通过闪存存储设备进行读写测试得到的数据,并将得到的所述数据作为所述测试数据。
3. 根据权利要求1所述的方法,其特征在于,通过聚类算法对所述测试数据进行聚类,得到多个数据集合包括:
选取多个质心,其中,所述质心与所述数据集合一一对应;
通过聚类算法根据多个所述质心,对所述测试数据进行聚类,得到多个数据集合。
4. 根据权利要求3所述的方法,其特征在于,选取多个质心包括:
选取多个初始质心;
对多个所述初始质心进行优化,得到多个更新质心;
在多个所述更新质心分别与对应的所述初始质心的距离,均小于预设距离的情况下,确定多个更新质心为所述多个质心;
在多个所述更新质心分别与对应的所述初始质心的距离,不小于预设距离的情况下,将多个真实质心作为新的初始质心再次进行优化,直至优化后的多个所述更新质心分别与对应的所述初始质心的距离,均小于预设距离。
5. 根据权利要求4所述的方法,其特征在于,选取多个初始质心包括:
将需要得到的数据集合的数量为所述质心的数量;
确定每个闪存存储设备的测试数据的数据向量,其中,所述测试数据包括读写数据,所述读写数据包括读取速度和写入速度,所述数据向量为根据所述读取速度和写入速度对应的坐标计算得到的;
随机选取一个读写数据作为第一初始质心;
根据测试数据中的读写数据的数据向量和所述第一初始质心,确定第二初始质心;
根据测试数据中的读写数据的数据向量和所述第二初始质心,确定第三初始质心,直至确定出所述数量的初始质心。
6. 根据权利要求5所述的方法,其特征在于,根据测试数据中的读写数据的数据向量和所述第一初始质心,确定第二初始质心包括:
确定所述测试数据中除所述第一初始质心之外的读写数据,与所述第一初始质心的距离;
根据所述距离,确定所述测试数据中除所述第一初始质心之外的读写数据为第二初始质心的概率;
选取所述概率最大的读写数据作为第二初始质心;
根据测试数据中的读写数据的数据向量和所述第二初始质心,确定第三初始质心包

括：

确定所述测试数据中除所述第一初始质心和所述第二初始质心之外的读写数据，与所述第二初始质心的距离；

根据所述距离，确定所述测试数据中除所述第一初始质心和所述第二初始质心之外的读写数据为第三初始质心的概率；

选取所述概率最大的读写数据作为第三初始质心。

7. 根据权利要求4所述的方法，其特征在于，对多个所述初始质心进行优化，得到多个更新质心包括：

根据所述测试数据的读写数据与所述多个初始质心的距离，对所述测试数据的读写数据进行分类，得到多个所述初始质心对应的多个初始集合，其中，所述读写数据属于与所述读写数据距离最近的初始质心对应的初始集合；

确定多个初始集合的实际质心，作为所述初始集合的初始质心对应的更新质心，得到多个所述更新质心。

8. 一种闪存存储设备的测试装置，其特征在于，包括：

测试模块，用于采用测试用例对多个闪存存储设备进行读写测试，得到测试数据；

聚类模块，用于通过聚类算法对所述测试数据进行聚类，得到多个数据集合；

生成模块，用于根据所述多个数据集合生成测试结果。

9. 一种闪存存储介质，其特征在于，所述计算机存储介质包括存储的程序，其中，在所述程序运行时控制所述计算机存储介质所在设备执行权利要求1至7中任意一项所述的闪存存储设备的测试方法。

10. 一种处理器，其特征在于，所述处理器用于运行程序，其中，所述程序运行时执行权利要求1至7中任意一项所述的闪存存储设备的测试方法。

闪存存储设备的测试方法及装置

技术领域

[0001] 本发明涉及硬件测试领域,具体而言,涉及一种闪存存储设备的测试方法及装置。

背景技术

[0002] 相关技术中的闪存存储设备性能测试方法主要是使用专用的硬盘读写工具比如说FIO、BIT等进行速度测试,在程序中设置一个读写速度的阈值作为性能判定标准,对速度在不同的阈值区间进行分类,速度小于该阈值则为性能较差,反之则读写性能优。或者测试结果只是输出到测试工具的日志中,人工分析日志的读写速度数据进行主观判断闪存存储设备读写性能优良。存在误差较大,效率较低的问题。

[0003] 相关技术中通过阈值作为判定标准会使得判定标准僵硬化,不具有过渡性,处在阈值附近的数值会因为较小的差别被错判或者误判(比如说判定标准阈值是300M/S,若闪存存储设备的实测速率是299M/S,由于小于阈值则会错判为速度低);而打印到日志中分析的方法,对于小批量的闪存存储设备还可以考虑,对于大批量的闪存存储设备来说就会效率极低。

[0004] 针对上述的问题,目前尚未提出有效的解决方案。

发明内容

[0005] 本发明实施例提供了一种闪存存储设备的测试方法及装置,以至少解决相关技术中通过专用的读写工具生成测试日志,人工分析测试日志确定闪存存储设备的读写速度,误差较大,效率较低的技术问题。

[0006] 根据本发明实施例的一个方面,提供了一种闪存存储设备的测试方法,包括:采用测试用例对多个闪存存储设备进行读写测试,得到测试数据;通过聚类算法对所述测试数据进行聚类,得到多个数据集合;根据所述多个数据集合生成测试结果。

[0007] 可选的,采用测试用例对多个闪存存储设备进行读写测试,得到测试数据包括:控制所述多个闪存存储设备运行所述测试用例进行读写测试;确定所述多个闪存存储设备中成功完成所述测试用例的测试通过闪存存储设备;获取所述测试通过闪存存储设备进行读写测试得到的数据,并将得到的所述数据作为所述测试数据。

[0008] 可选的,通过聚类算法对所述测试数据进行聚类,得到多个数据集合包括:选取多个质心,其中,所述质心与所述数据集合一一对应;通过聚类算法根据多个所述质心,对所述测试数据进行聚类,得到多个数据集合。

[0009] 可选的,选取多个质心包括:选取多个初始质心;对多个所述初始质心进行优化,得到多个更新质心;在多个所述更新质心分别与对应的所述初始质心的距离,均小于预设距离的情况下,确定多个更新质心为所述多个质心;在多个所述更新质心分别与对应的所述初始质心的距离,不小于预设距离的情况下,将多个真实质心作为新的初始质心再次进行优化,直至优化后的多个所述更新质心分别与对应的所述初始质心的距离,均小于预设距离。

[0010] 可选的,选取多个初始质心包括:将需要得到的数据集合的数量为所述质心的数量;确定每个闪存存储设备的测试数据的数据向量,其中,所述测试数据包括读写数据,所述读写数据包括读取速度和写入速度,所述数据向量为根据所述读取速度和写入速度对应的坐标计算得到的;随机选取一个读写数据作为第一初始质心;根据测试数据中的读写数据的数据向量和所述第一初始质心,确定第二初始质心;根据测试数据中的读写数据的数据向量和所述第二初始质心,确定第三初始质心,直至确定出所述数量的初始质心。

[0011] 可选的,根据测试数据中的读写数据的数据向量和所述第一初始质心,确定第二初始质心包括:确定所述测试数据中除所述第一初始质心之外的读写数据,与所述第一初始质心的距离;根据所述距离,确定所述测试数据中除所述第一初始质心之外的读写数据为第二初始质心的概率;选取所述概率最大的读写数据作为第二初始质心;根据测试数据中的读写数据的数据向量和所述第二初始质心,确定第三初始质心包括:确定所述测试数据中除所述第一初始质心和所述第二初始质心之外的读写数据,与所述第二初始质心的距离;根据所述距离,确定所述测试数据中除所述第一初始质心和所述第二初始质心之外的读写数据为第三初始质心的概率;选取所述概率最大的读写数据作为第三初始质心。

[0012] 可选的,对多个所述初始质心进行优化,得到多个更新质心包括:根据所述测试数据的读写数据与所述多个初始质心的距离,对所述测试数据的读写数据进行分类,得到多个所述初始质心对应的多个初始集合,其中,所述读写数据属于与所述读写数据距离最近的初始质心对应的初始集合;确定多个初始集合的实际质心,作为所述初始集合的初始质心对应的更新质心,得到多个所述更新质心。

[0013] 根据本发明实施例的另一方面,还提供了一种闪存存储设备的测试装置,包括:测试模块,用于采用测试用例对多个闪存存储设备进行读写测试,得到测试数据;聚类模块,用于通过聚类算法对所述测试数据进行聚类,得到多个数据集合;生成模块,用于根据所述多个数据集合生成测试结果。

[0014] 根据本发明实施例的另一方面,还提供了一种计算机存储介质,所述计算机存储介质包括存储的程序,其中,在所述程序运行时控制所述计算机存储介质所在设备执行上述中任意一项所述的闪存存储设备的测试方法。

[0015] 根据本发明实施例的另一方面,还提供了一种处理器,所述处理器用于运行程序,其中,所述程序运行时执行上述中任意一项所述的闪存存储设备的测试方法。

[0016] 在本发明实施例中,采用测试用例对多个闪存存储设备进行读写测试,得到测试数据;通过聚类算法对测试数据进行聚类,得到多个数据集合;根据多个数据集合生成测试结果的方式,通过对多个闪存存储设备进行读写测试,通过聚类算法对测试数据进行聚类,将多个闪存存储设备进行分类,从而快速对测试数据块进行筛选,达到了对海量闪存存储设备同时测试,并对测试数据进行分类的目的,从而实现了提高对海量闪存存储设备的测试效率,以及对测试数据进行查询的效率和准确度的技术效果,进而解决了相关技术中通过专用的读写工具生成测试日志,人工分析测试日志确定闪存存储设备的读写速度,误差较大,效率较低的技术问题。

附图说明

[0017] 此处所说明的附图用来提供对本发明的进一步理解,构成本申请的一部分,本发

明的示意性实施例及其说明用于解释本发明,并不构成对本发明的不当限定。在附图中:

[0018] 图1是根据本发明实施例的一种闪存存储设备的测试方法的流程图;

[0019] 图2是根据本发明实施方式的测试模块的示意图;

[0020] 图3是根据本发明实施方式的测试系统的硬件连接的示意图;

[0021] 图4是根据本发明实施方式的测试方法的流程图;

[0022] 图5是根据本发明实施例的一种闪存存储设备的测试装置的示意图。

具体实施方式

[0023] 为了使本技术领域的人员更好地理解本发明方案,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分的实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都应当属于本发明保护的范围。

[0024] 需要说明的是,本发明的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本发明的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排除他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0025] 根据本发明实施例,提供了一种闪存存储设备的测试方法的方法实施例,需要说明的是,在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行,并且,虽然在流程图中示出了逻辑顺序,但是在某些情况下,可以以不同于此处的顺序执行所示出或描述的步骤。

[0026] 图1是根据本发明实施例的一种闪存存储设备的测试方法的流程图,如图1所示,该方法包括如下步骤:

[0027] 步骤S102,采用测试用例对多个闪存存储设备进行读写测试,得到测试数据;

[0028] 步骤S104,通过聚类算法对测试数据进行聚类,得到多个数据集合;

[0029] 步骤S106,根据多个数据集合生成测试结果。

[0030] 通过上述步骤,采用测试用例对多个闪存存储设备进行读写测试,得到测试数据;通过聚类算法对测试数据进行聚类,得到多个数据集合;根据多个数据集合生成测试结果的方式,通过对多个闪存存储设备进行读写测试,通过聚类算法对测试数据进行聚类,将多个闪存存储设备进行分类,从而快速对测试数据块进行筛选,达到了对海量闪存存储设备同时测试,并对测试数据进行分类的目的,从而实现了提高对海量闪存存储设备的测试效率,以及对测试数据进行查询的效率和准确度的技术效果,进而解决了相关技术中通过专用的读写工具生成测试日志,人工分析测试日志确定闪存存储设备的读写速度,误差较大,效率较低的技术问题。

[0031] 上述闪存存储设备(Solid State Disk)也称固态硬盘,其读写性能是一个重要参数,本实施例通过测试用例对多个闪存存储设备同时进行读写测试,并对得到的测试数据

进行聚类,得到多个数据集合,每个数据集合对应一个质心,该质心对应的读写速度,就可以反映该数据集合的读写性能,根据多个数据集合可以快速有效的生成测试结果。在闪存存储设备的数量较多时,相比于传统的测试方式,其效率提升越明显。

[0032] 在上述采用测试用例对多个闪存存储设备进行读写测试,得到测试数据之前,先对多个闪存存储设备通过交换机和主机相连,其中该主机用于接收上述测试用例。将多个闪存存储设备通过交换机与主机相连后,检测网络连接是否完好,可以保证数据传输,在网络连接完好的情况下,将测试用例传输给多个闪存存储设备进行读写测试。

[0033] 可选的,采用测试用例对多个闪存存储设备进行读写测试,得到测试数据包括:控制多个闪存存储设备运行测试用例进行读写测试;确定多个闪存存储设备中成功完成测试用例的测试通过闪存存储设备;获取测试通过闪存存储设备进行读写测试得到的数据,并将得到的数据作为测试数据。

[0034] 在闪存存储设备未成功完成测试用例的情况下,说明该闪存存储设备测试失败,测试失败的闪存存储设备不会产生有效的测试数据,因此,只对测试通过的闪存存储设备的测试数进行获取,提高测试数据的有效性和准确性。

[0035] 可选的,通过聚类算法对测试数据进行聚类,得到多个数据集合包括:选取多个质心,其中,质心与数据集合一一对应;通过聚类算法根据多个质心,对测试数据进行聚类,得到多个数据集合。

[0036] 上述质心也即是进行聚类的中心,可以根据在坐标系中各个测试数据与质心之间的距离确定该测试数据是否属于该质心的数据集合。上述测试数据包括读写数据,读写数据具体包括读取速度和写入速度,在本实施例中上述坐标系可以分别以读取速度和写入速度为坐标轴建立坐标系。

[0037] 可选的,选取多个质心包括:选取多个初始质心;对多个初始质心进行优化,得到多个更新质心;在多个更新质心分别与对应的初始质心的距离,均小于预设距离的情况下,确定多个更新质心为多个质心;在多个更新质心分别与对应的初始质心的距离,不小于预设距离的情况下,将多个真实质心作为新的初始质心再次进行优化,直至优化后的多个更新质心分别与对应的初始质心的距离,均小于预设距离。

[0038] 在选取初始质心后,初始质心可能并不能保证对测试数据进行有效科学的分类,因此,通过对初始质心进行优化,得到更新质心,在多个更新质心分别与对应的初始质心的距离,均小于预设距离的情况下,确定多个更新质心为多个质心,说明初始质心可以有效的对测试数据进行分类,但是,初始质心相比于初始集合实际的质心,也即是更新质心,更新质心更准确,因此,将更新质心作为上述多个质心。

[0039] 另外,在多个更新质心分别与对应的初始质心的距离,不小于预设距离的情况下,说明初始质心进行聚类的到的数据集合不够合理,将多个真实质心作为新的初始质心再次进行优化,直至优化后的多个更新质心分别与对应的初始质心的距离,均小于预设距离。从而提高质心的合理性,提高多个数据集合的分类准确性和科学性。

[0040] 可选的,选取多个初始质心包括:将需要得到的数据集合的数量为质心的数量;确定每个闪存存储设备的测试数据的数据向量,其中,测试数据包括读写数据,读写数据包括读取速度和写入速度,数据向量为根据读取速度和写入速度对应的坐标计算得到的;随机选取一个读写数据作为第一初始质心;根据测试数据中的读写数据的数据向量和第一初始

质心,确定第二初始质心;根据测试数据中的读写数据的数据向量和第二初始质心,确定第三初始质心,直至确定出数量的初始质心。

[0041] 先随机选择一个第一初始质心,然后根据测试数据中的读写数据的数据向量确定第二初始质心,根据第二初始质心和测试数据中的读写数据的数据向量,以此类推,直至确定出所需数量的初始质心,并记录该初始质心。

[0042] 可选的,根据测试数据中的读写数据的数据向量和第一初始质心,确定第二初始质心包括:确定测试数据中除第一初始质心之外的读写数据,与第一初始质心的距离;根据距离,确定测试数据中除第一初始质心之外的读写数据为第二初始质心的概率;选取概率最大的读写数据作为第二初始质心;根据测试数据中的读写数据的数据向量和第二初始质心,确定第三初始质心包括:确定测试数据中除第一初始质心和第二初始质心之外的读写数据,与第二初始质心的距离;根据距离,确定测试数据中除第一初始质心和第二初始质心之外的读写数据为第三初始质心的概率;选取概率最大的读写数据作为第三初始质心。

[0043] 根据第一质心和测试数据中除第一初始质心之外的读写数据,确定测试数据中除第一初始质心之外的读写数据为第二初始质心的概率,具体可以通过以下概率公式确定:

$$[0044] \quad \tau = \frac{d(rw_i)^2}{\sum_i d(rw_i)^2}$$

[0045] 式中, τ 为第*i*个读写数据为第二质心(下一个质心)的概率, $d(rw_i)$ 为第*i*个读写数据的数据向量 rw_i 与第一质心(当前质心)的距离。

[0046] 通过上述公式,可以根据第一初始质心,选择第二初始质心,然后逐个选择上述数量的初始质心。

[0047] 上述概率公式的含义在于,确定与当前质心距离最远的读写数据,为下一个初始质心。

[0048] 可选的,对多个初始质心进行优化,得到多个更新质心包括:根据测试数据的读写数据与多个初始质心的距离,对测试数据的读写数据进行分类,得到多个初始质心对应的多个初始集合,其中,读写数据属于与读写数据距离最近的初始质心对应的初始集合;确定多个初始集合的实际质心,作为初始集合的初始质心对应的更新质心,得到多个更新质心。

[0049] 在对多个初始质心进行优化,得到多个更新质心之后,通过聚类算法根据更新质心进行聚类,得到多个数据集合。

[0050] 根据该数据集合对应的更新质心的读取速度和写入速度可以确定多个数据集合之间的读写性能差距。也可以确定出读写性能较好的数据集合,完成对多个闪存存储设备的测试。

[0051] 需要说明的是,本申请实施例还提供了一种可选的实施方式,下面对该实施方式进行详细说明。

[0052] 本实施方式应用在闪存存储设备读写性能分析,尤其是对大批量的闪存存储设备读写性能结果分析呈现,避免了划分标准阈值过渡带的误判问题,通过导入不同的测试用例,对不同的测试用例下的读写性能分别分析并生成测试分析报告,方便分析查看不同参数下的闪存存储设备性能。自动获取测试用例,多台测试机同时测试,主机自动采集数据并分析。

[0053] 图2是根据本发明实施方式的测试模块的示意图,如图2所示,本实施方式一共分为五大模块,待测试闪存存储设备、闪存存储设备读写测试模块、主机数据监控模块、主机读写参数(测试用例)模块、主机大数据分析模块。

[0054] 图3是根据本发明实施方式的测试系统的硬件连接的示意图,图4是根据本发明实施方式的测试方法的流程图,如图3和图4所示,本实施方式提供了一种基于K-Means聚类算法的批量闪存存储设备读写性能自动化分析方法。

[0055] 该方法使用现有的FIO读写测试工具,该方法的实现主要执行以下步骤:

[0056] 1、测试机检测网络连接完好,若检测网络连接完好执行步骤2,若监测网络连接不完好,则退出本次测试,对监测网络进行重连;

[0057] 2、各个测试机分别从主机获取测试用例,根据用例进行相应的读写测试,一个样例测试完毕之后要将测试未完成的闪存存储设备剔除;

[0058] 3、重复步骤2直至所有的测试用例完成;

[0059] 4、所有闪存存储设备的测试完成,主机获取所有测试机的读写测试数据;

[0060] 5、主机基于K-Means算法对读写数据处理完成性能分析分类。

[0061] 测试用例IO引擎全部使用libaio,读写分别设置4K顺序读写、128K随机读、32K顺序读写。

[0062] 现对基于K-Means聚类算法将闪存存储设备的读写速率进行分类的实现过程做详细介绍:

[0063] 假设,分类集合数为P(即为聚簇数目),最大迭代次数为M。

[0064] 一、构建读写数据对原始数据集:

[0065] ①设置每块磁盘的读写速率作为一个数据对,测试过得N块盘对于每个测试脚本即有N组读写速率数据对, (r_1, w_1) 、 (r_2, w_2) 、 $(r_3, w_3) \cdots (r_N, w_N)$, 对应的原始数据对向量集合为 $D\{rw_1, rw_2, rw_3 \cdots rw_N\}$, 期望的输出数据对集合为 $O\{\{O_1\}, \{O_2\}, \{O_3\} \cdots \{O_P\}\}$ 。

[0066] 二、选取合适初始质心:

[0067] ②从N组给出的数据对中随机选取一个数据对作为一个质心 c_1 , 对于N组数据对集合中的每一个向量 rw_i ($i=1, 2, 3 \cdots N$), 计算每个向量与当前已选择的质心的最小距离(即离自己距离近的质心的距离) $d(rw_i)$, 计算每个样本被选中成为下一个质心的概率

$\frac{d(rw_i)^2}{\sum d(rw_i)^2}$, 使用轮盘算法选择下一个数据对作为另一个质心。

[0068] ③迭代步骤②直至选择出P个较优的数据对作为质心,使用集合 $C\{c_1, c_2, c_3 \cdots c_P\}$ 表示。

[0069] 三、优化质心位置:

[0070] ④每个质心向量记为 c_j ($j=1, 2, 3 \cdots P$) 根据公式 $d_{ij} = \|rw_i - c_j\|_2^2$ 计算N个数据对分别到P个质心数据对的距离,距离各个质心较小的划分为该质心所属的集合中,比如数据对 rw_i 到质心 c_j 的距离较小,则 rw_i 归为质心 c_j 所在的集合即 $O'_{c_j} = O_{c_j} \cup \{rw_i\}$;

[0071] ⑤所有集合计算完毕,此时共有P个数据对集合,对新的P个数据对集合重新计算P

个新的质心 $c_j = \frac{1}{|O_j|} \sum_{rw_i \in O_j} rw_i$;

[0072] 四、收敛判断与结果输出:

[0073] ⑥计算新的P个质心与原先P个质心的距离,如果距离小于某个收敛值则认为算法收敛,此时可以输出统计结果,如果距离大于既定的收敛值在迭代次数小于最大迭代次数的前提下对④-⑤步骤迭代计算,否则结束计算输出统计结果。

[0074] 在整个算法执行过程中每个读写数据对会携带当前读写速率对应闪存存储设备的SN信息,这样在最后生成的统计结果中可以通过SN方便的查看某块闪存存储设备的读写速率所处的层次。

[0075] 根据以上算法实现方案,具体的实施细节如下:

[0076] 如表1所示,通过10台测试机进行n轮测试之后,在128K顺序读写测试脚本下得出100组的读写速度对。

[0077] 表1 10台测试机进行n轮测试的测试数据

序号	Read(M/S)	Write(M/S)	SN
1	530	180	*****0001
2	532	172	*****0020
3	529	185	*****0015
4	525	175	*****0003
5	300	150	*****0008
6	250	186	*****0011
7	510	160	*****0012
8	511	180	*****0013
9	520	170	*****0014
10	505	175	*****0016
...
100	200	130	*****0099

[0080] 选取这100组读写速率数据对作为原始集合标记为 $D\{rw_1, rw_2, rw_3 \cdots rw_{100}\}$,为了更好的标记以及显示闪存存储设备的性能,选取分类集合数P取值为3。对所有的读写速率数据对分为读写速率低、中、高。

[0081] 根据步骤2所述,从原始集合中随机选取一个点作为第一个初始的质心记为 c_1 ,计算集合D所有向量与 c_1 的最小距离记为 $d(rw_1)$ 、 $d(rw_2) \cdots d(rw_{100})$,计算每个点成为下一个质心的概率,套用轮盘算法计算得出第二个质心;以此类推找出三个最优的质心记为 c_1 、 c_2 、 c_3 。

[0082] 计算集合D所有向量分别到三个质心的距离 d_{ij} ($i=1,2,3\cdots 100$) ($j=1,2,3$), 根据距离三个质心的距离远近归类到相近的集合 O_1 、 O_2 、 O_3 。

[0083] 对新的3个数据对集合 O_1 、 O_2 、 O_3 重新计算3个新的质心, 计算新的3个质心与原先3个质心的距离, 如果距离小于某个收敛值则认为算法收敛, 此时可以输出统计结果, 否则迭代以上步骤直至收敛。这样到程序收敛之后会生成三个数据集合分别用来保存读写速率对, 通过这个算法就可实现快速读写速率性能的识别。

[0084] 本实施方式的基于K-Means聚类算法将闪存存储设备的读写速率进行分类, 实现很好的闪存存储设备速率分布判断, 能比较清楚的描述读写速率处在过渡带以及边缘的闪存存储设备, 而非一刀切。

[0085] 图5是根据本发明实施例的一种闪存存储设备的测试装置的示意图, 如图5所示, 根据本发明实施例的另一方面, 还提供了一种闪存存储设备的测试装置, 包括: 测试模块52, 聚类模块54和生成模块56, 下面对该装置进行详细说明。

[0086] 测试模块52, 用于采用测试用例对多个闪存存储设备进行读写测试, 得到测试数据; 聚类模块54, 与上述测试模块52相连, 用于通过聚类算法对测试数据进行聚类, 得到多个数据集合; 生成模块56, 与上述聚类模块54相连, 用于根据多个数据集合生成测试结果。

[0087] 通过上述装置, 采用测试模块52测试用例对多个闪存存储设备进行读写测试, 得到测试数据; 聚类模块54通过聚类算法对测试数据进行聚类, 得到多个数据集合; 生成模块56根据多个数据集合生成测试结果的方式, 通过对多个闪存存储设备进行读写测试, 通过聚类算法对测试数据进行聚类, 将多个闪存存储设备进行分类, 从而快速对测试数据块进行筛选, 达到了对海量闪存存储设备同时测试, 并对测试数据进行分类的目的, 从而实现了提高对海量闪存存储设备的测试效率, 以及对测试数据进行查询的效率和准确度的技术效果, 进而解决了相关技术中通过专用的读写工具生成测试日志, 人工分析测试日志确定闪存存储设备的读写速度, 误差较大, 效率较低的技术问题。

[0088] 根据本发明实施例的另一方面, 还提供了一种计算机存储介质, 计算机存储介质包括存储的程序, 其中, 在程序运行时控制计算机存储介质所在设备执行上述中任意一项的闪存存储设备的测试方法。

[0089] 根据本发明实施例的另一方面, 还提供了一种处理器, 处理器用于运行程序, 其中, 程序运行时执行上述中任意一项的闪存存储设备的测试方法。

[0090] 上述本发明实施例序号仅仅为了描述, 不代表实施例的优劣。

[0091] 在本发明的上述实施例中, 对各个实施例的描述都各有侧重, 某个实施例中沒有详述的部分, 可以参见其他实施例的相关描述。

[0092] 在本申请所提供的几个实施例中, 应该理解到, 所揭露的技术内容, 可通过其它的方式实现。其中, 以上所描述的装置实施例仅仅是示意性的, 例如所述单元的划分, 可以为一种逻辑功能划分, 实际实现时可以有另外的划分方式, 例如多个单元或组件可以结合或者可以集成到另一个系统, 或一些特征可以忽略, 或不执行。另一点, 所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口, 单元或模块的间接耦合或通信连接, 可以是电性或其它的形式。

[0093] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的, 作为单元显示的部件可以是或者也可以不是物理单元, 即可以位于一个地方, 或者也可以分布到多个

单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0094] 另外,在本发明各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0095] 所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时,可以存储在一个计算机可读取存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可为个人计算机、服务器或者网络设备等)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、移动硬盘、磁碟或者光盘等各种可以存储程序代码的介质。

[0096] 以上所述仅是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

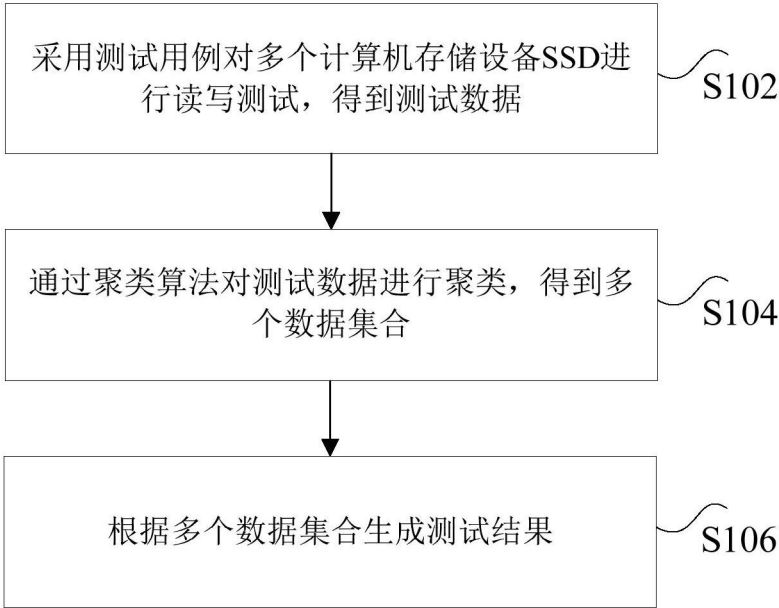


图1

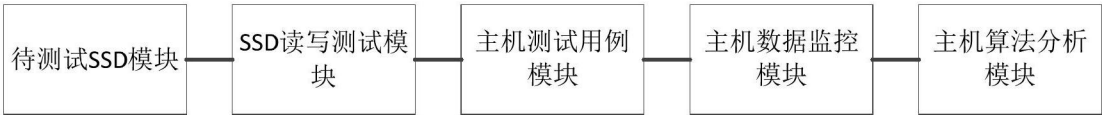


图2

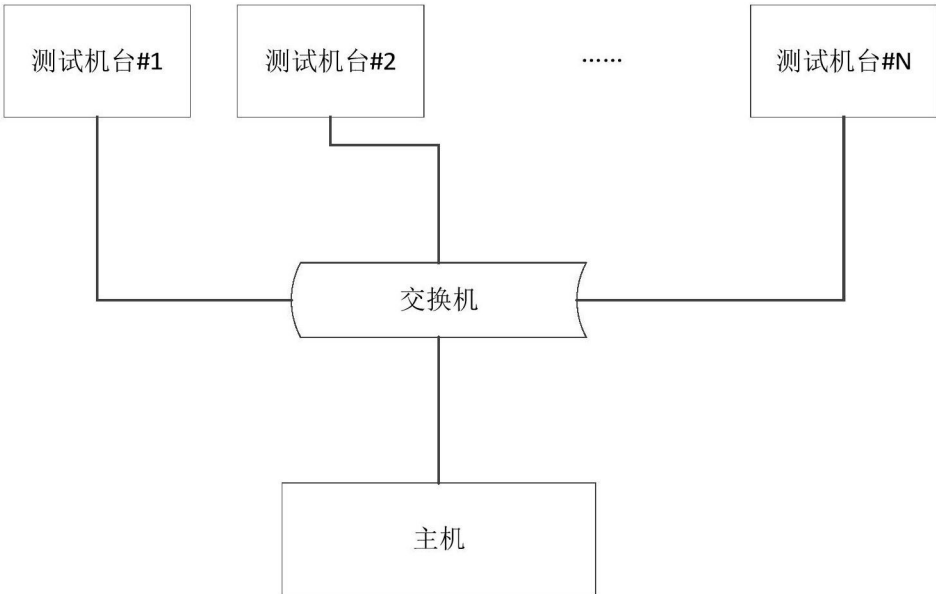


图3

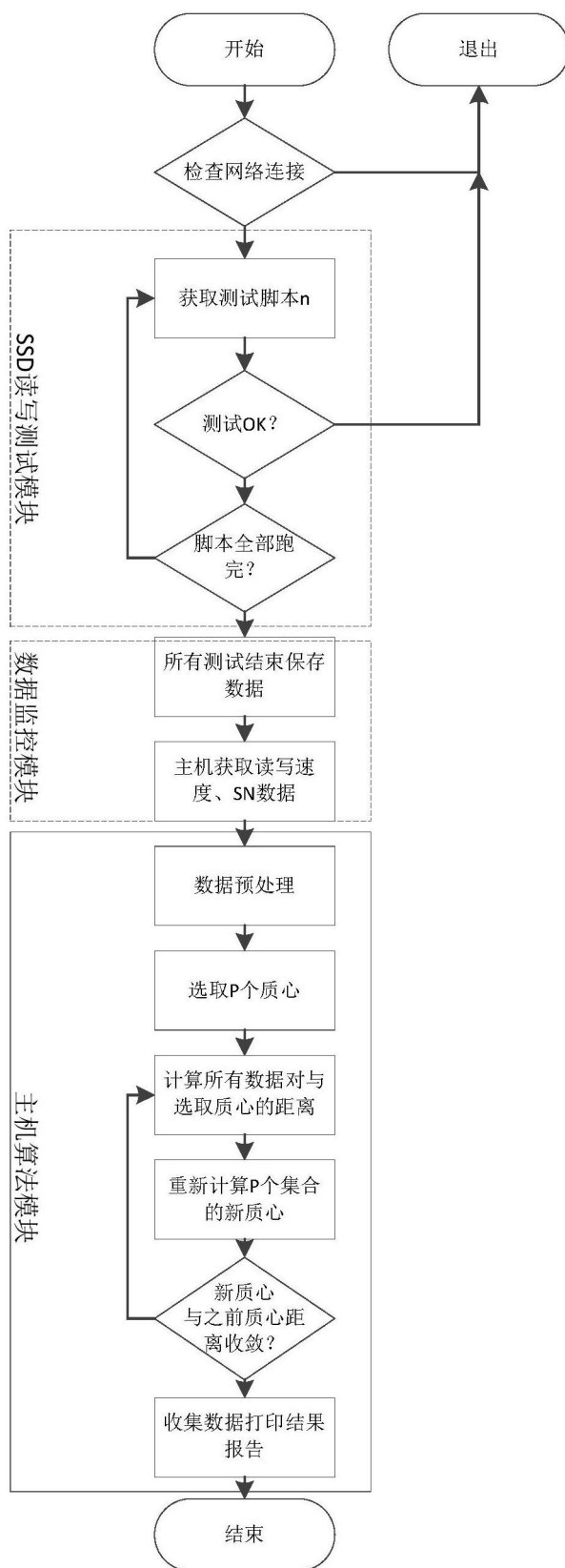


图4



图5