(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2025 年 9 月 4 日 (04.09.2025)

WIPO | PCT

(10) 国际公布号
WO 2025/181562 A1

(51) 国际专利分类号:
G06F 16/22 (2019.01) *G06F 16/901* (2019.01)

(21) 国际申请号: PCT/IB2025/050170

(22) 国际申请日: 2025 年 1 月 8 日 (08.01.2025)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202410232858.4 2024年2月29日 (29.02.2024) CN

(71) 申请人: 云智能资产控股（新加坡）私人股份有限公司 **(CLOUD INTELLIGENCE ASSETS HOLDING (SINGAPORE) PRIVATE LIMITED)** [SG/SG]; 新加坡勿拉士峇沙路 51 号来赞达一号大厦#03-06 189554 (SG)。

(72) 发明人: 徐中 **(XU, Zhong)**; 中国浙江省杭州市西湖区三墩镇灯彩街1008号云谷园区 310030 (CN)。黄贵 **(HUANG, Gui)**; 中国浙江省杭州市西湖区三墩镇灯彩街1008号云谷园区 310030 (CN)。楼江航 **(LOU, Jianghang)**; 中国浙江省杭州市西湖区三墩镇灯彩街1008号云谷园区 310030 (CN)。方孝健 **(FANG, Xiaojian)**; 中国浙江省杭州市西湖区三墩镇灯彩街1008号云谷园区 310030 (CN)。

(81) 指定国(除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UY, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明，要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE,

(54) **Title:** COLUMN-BASED STORAGE METHOD, DEVICE, COMPUTER PROGRAM PRODUCT, AND STORAGE MEDIUM

(54) 发明名称: 一种列式存储方法、设备、计算机程序产品及存储介质

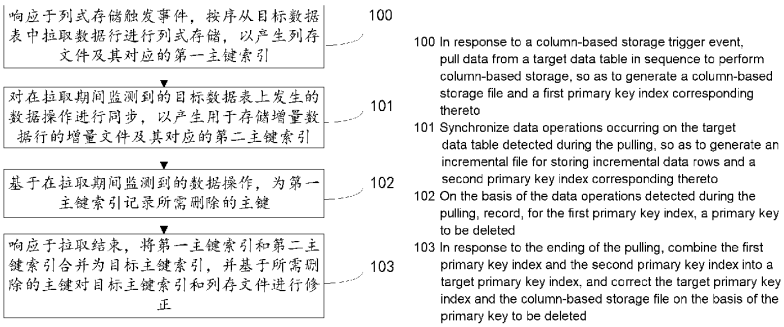


图 1

(57) **Abstract:** Embodiments of the present disclosure provide a column-based storage method, a device, a computer program product, and a storage medium. Batch streams and incremental streams can be pulled in parallel while normal on-line service provision with row-based storage is maintained. A column-based storage file and a first primary key index can be generated by pulling batch streams; by synchronizing data operations occurring on a target data table detected during the pulling to pull incremental streams, an incremental file and a second primary key index can be generated, so that it can be ensured that column-based storage can keep consistency with row-based storage in terms of incremental processing; further, a primary key to be deleted is recorded for the first primary key index on the basis of the data operations detected during the pulling; and in response to the ending of the pulling, the first primary key index and the second primary key index are combined into a target primary key index, and the target primary key index and the column-based storage file are corrected on the basis of the primary key to be deleted, so that it can be ensured that column-based storage can also achieve consistency with row-based storage in terms of batch processing.

WO 2025/181562 A1

[见续页]

BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

根据细则4.17的声明:
— 关于申请人有权要求在先申请的优先权(细则4.17(iii))

本国际公布:
— 包括国际检索报告(条约第21条(3))。

(57) 摘要: 本公开实施例提供一种列式存储方法、设备、计算机程序产品及存储介质。在保持按行式存储正常提供在线服务的前提下，可并行拉取存量流和增量流。通过拉取存量流可产生列存文件及第一主键索引。通过对在拉取期间监测到的目标数据表上发生的数据操作进行同步，以拉取增量流，可产生增量文件及第二主键索引，并可保证列式存储在增量方面能够与行式存储保持一致性。以及还可基于拉取期间监测到的数据操作，为第一主键索引记录所需删除的主键；响应于拉取结束，将第一主键索引和第二主键索引合并为目标主键索引，并基于所需删除的主键对目标主键索引和列存文件进行修正，这可保证列式存储在存量方面也能够与行式存储达到一致性。