CRAFT 磁体性能研究平台数据归档系统

马树良[1,2] 刘华军[2] 刘 方[2] 张舒庆[2,3] 李 童[2] 施 毅[2] 郭 亮[2]

1（安徽大学 物质科学与信息技术研究院 合肥 230601）

2（中国科学院合肥物质科学研究院 合肥 230031）

3（中国科学技术大学 合肥 230026）

摘要 聚变堆主机关键系统的综合研究设施的磁体性能研究平台（Magnet Performance Research Platform，MPRP）是为先进超导磁体实验建立的大型实验平台，其历史数据在海量存储情况下存在检索速度慢的问题。因此，对系统检索速度进行研究并开发了 MPRP 数据归档系统（MPRP Data Archiving System，MPDAS）。MPDAS 设计了 EPICS（Experimental Physics and Industrial Control System）数据归档插件并采用 MongoDB 分片和副本集机制搭建高扩展性数据存储架构。为提高数据检索速度，MPDAS 借鉴最近最少使用（Least Recently Used，LRU）、使用频率最低（Least Frequently Used，LFU）、先进先出（First In First Out，FIFO）三种传统缓存替换算法核心思想，基于牛顿冷却定律建立数据温度模型并提出一种综合访问时间、访问频率以及存储顺序的多维度特征数据划分算法。根据数据划分算法标识冷热历史数据实现数据分层存储。MPDAS 在查询历史数据时优先访问 Redis，根据命中结果和数据完整性选择不同的检索策略。系统测试结果表明：MPDAS 功能特征满足设计要求，其搭载的冷热数据划分算法相比 FIFO、LRU、LFU 在热数据库保存 1% 历史数据量时的 Redis 命中率分别提升了 38.05%、26.91% 和 11.06%。通过提高热数据命中率能够直接减少数据检索平均响应时间，MPDAS 通过量化历史数据热度并进行冷热划分，有效地提升了系统检索响应速度。

关键词 历史数据，MongoDB，热数据，PV，分布式集群

中图分类号 P315.69

DOI: 10.11889/j.0253-3219.2023.hjs.46.020401

CRAFT magnet performance research platform data archiving system

MA Shuliang[1,2] LIU Huajun[2] LIU Fang[2] ZHANG Shuqing[2,3] LI Tong[2] SHI Yi[2] GUO Liang[2]

1(*Institute of Physical Science and Information Technology*, *Anhui University*, *Hefei 230601*, *China*)

2(*Hefei Institutes of Physical Science*, *Chinese Academy of Sciences*, *Hefei 230031*, *China*)

3(*University of Science and Technology of China*, *Hefei 230026*, *China*)

Abstract **[Background]** The comprehensive research facility for fusion technology magnet performance research platform (MPRP) is a large-scale experimental platform established for advanced superconducting magnet experiments. The retrieval speed of MPRP historical data is slow due to massive storage. **[Purpose]** The study aims to develop a MPRP data archiving system (MPDAS) and increase its retrieval speed. **[Methods]** First of all, the experimental physics and industrial control system (EPICS) data archiving plug-in was designed for MPDAS. Both MongoDB Sharding and Replica Set mechanism were employed to build a highly scalable data storage architecture. Then, the core ideas of three traditional cache replacement algorithms, LRU (least recently used), LFU (least

中国科学院战略性先导科技专项(No.XDB250020200)资助

第一作者：马树良，男，1998年出生，2020年毕业于山东农业大学，现为硕士研究生，研究领域为计算机技术

通信作者：刘华军，E-mail：liuhj@ipp.ac.cn

收稿日期：2022-09-08，修回日期：2022-11-23

Supported by Strategic Priority Research Program of Chinese Academy of Sciences (No.XDB250020200)

First author: MA Shuliang, male, born in 1998, graduated from Shandong Agricultural University in 2020, master student, focusing on computer technology

Corresponding author: LIU Huajun, E-mail: liuhj@ipp.ac.cn

Received date: 2022-09-08, revised date: 2022-11-23

frequently used) and FIFO (first in first out) were drawn by MPDAS to establish a data temperature model based on Newton's law of cooling. A multi-dimensional feature data partitioning algorithm was implemented to integrate access time, access frequency and storage order, hence the hot and cold historical data were identified to realize data tiered storage. Finally, the retrieval speed of MPDAS was improved by preferentially accessing Redis when querying historical data, and selecting different retrieval strategies based on hit results and data integrity. **[Results]** The system test results show that the functional characteristics of MPDAS meet the design requirements. Compared with FIFO, LRU, and LFU, the Redis hit rate of the MPDAS when the hot database stores 1% of the historical data is increased by 38.05%, 26.91%, and 11.06% respectively. **[Conclusions]** By increasing the hit rate of hot data, the average response time of data retrieval can be directly reduced. The retrieval response speed of MPDAS is effectively improved by quantifying the heat of historical data and dividing the heat and cold.

Key words Historical data, MongoDB, Hot data, PV, Distributed cluster

聚变堆主机关键系统综合研究设施(Comprehensive Research Facility for Fusion Technology,CRAFT)要建设具有国际领先水平的超导磁体研究系统和偏滤器研究系统,以解决中国聚变工程试验堆一些关键研发问题[1−2]。磁体性能研究平台(Magnet Performance Research Platform,MPRP)作为CRAFT超导磁体研究系统的重要组成部分,建成后将用于开展未来聚变堆超导磁体在强磁耦合、高磁能及强电磁干扰工况下的性能研究[3−4]。MPRP规模庞大,构成复杂,拥有2 000多路信号,数据生成的峰值速度将近100 MB·s^{-1},预计20年内其历史数据存储规模将达到PB级。在长期运行工况下,其历史数据规模不断增大,将导致数据访问的响应时间延长,影响用户体验。对MPRP来说,开发一个吞吐率高、可用性强的数据归档系统至关重要。

MPRP采用基于EPICS(Experimental Physics and Industrial Control System)的分布式控制系统。EPICS是一个广泛应用于大科学装置的大型分布式控制系统的软件运行环境和开发平台,主要由输入输出控制器IOC、操作接口OPI以及其特定的网络通讯协议Channel Access(CA)组成[5]。为存储历史数据,EPICS曾发布以文件为基础的数据存档工具Channel Archiver,被广泛应用于早期的SSRF(Shanghai Synchrotron Radiation Facility)、BEPC-II(Beijing Elec-tron Positron Collider II)等大科学装置[6−7]。后来Channel Archiver由于在数据管理和稳定性方面存在不足,被基于关系型数据库的数据归档系统所取代。一些专家学者采用关系型数据库对科学装置数据归档系统进行了升级。王春红等[8]将表空间和表分区等技术引入Oracle重建BEPC-II历史数据库解决了数据获取软件稳定性难题。罗江波等[9]基于MySQL数据库,并通过双机热备、读写分离、分区优化及建立索引等技术提高了加速器驱动次临界系统历史数据检索效率。随着历史数据累积量不断增加,关系型数据库查询性能在PB级数据存储条件下达到瓶颈状态,非关系型数据库逐渐得到广泛重视。EPICS社区发布的Archiver Appliance是一种基于文档型数据库的存档工具,文件存储为Protocol Buffers序列化后的二进制数据,I/O性能优异,已在HSL-II[10]、钍基熔盐堆[11]等装置上展开应用。为提升数据查询效率,一些研究人员开始采用非关系型数据库建设新型数据归档系统。乔予思等[12]采用MongoDB研发了BEPC-II新型数据存档软件,通过无序批量写入及覆盖索引提高了系统吞吐量。Kikuzawa等[13−14]采用HBase构建日本质子加速器分布式数据库,满足了系统存储扩展的高性能和高可用性需求。

在科学装置广泛使用的数据库中,Oracle、MySQL等关系型数据库存在海量数据存储情况下检索速度慢的问题,Archiver Appliance、HBase等查询方式灵活性不足。MongoDB[15]作为一种可扩展的高性能分布式存储数据库,具备优秀的海量数据存储和查询性能,能够对数据建立索引并且支持强大的查询语言,更符合MPDAS的使用需求。MPDAS面向EPICS IOC并采用多线程、MongoDB批量写入等技术并行数据同步流盘,基于分片和副本集机制搭建MongoDB分片复制集群。为提升数据检索响应速度,我们基于牛顿冷却定律并结合访问时间、访问频率以及存储顺序三个数据属性特征提出一种冷热数据划分算法,通过历史数据温度衡量其未来被频繁访问程度,划分历史数据为冷数据或热数据并将热数据抽取到Redis数据库。基于此设计,开发了基于MongoDB和Redis的MPRP数据归档系统。

1 系统架构

MPDAS主要功能是为MPRP提供稳定可靠的

历史数据存储服务以及为科研人员提供数据检索和可视化分析的人机交互界面。为实现分布性强、可用性高的数据归档系统，MPDAS采用B/S架构进行系统设计。系统总体架构图如图1所示，系统由数据归档模块（Data Archiving Module）、数据划分模块（Data Partition Module）、数据源（Data Source）和Web服务模块（Web Service Module）组成。

数据归档模块采用CA通讯协议与EPICS控制系统建立通道链接，通过JCA接口与IOC服务器交互，调用数据采集引擎获取PV（Process Variable）信息，能够同时监控所有物理信号。数据归档模块具备实时解析PV数据并进行BSON格式转换功能，为避免过度消耗数据库连接池资源，建立多个缓存区暂存BSON数据，当缓存区使用率达到阈值，通过Batch Write Engine将批量缓存数据并发写入MongoDB分布式集群。

数据划分模块提出一种基于牛顿冷却定律（Newton's Law of Cooling）的数据热度分析策略，综合访问时间、存储顺序和访问频率等多个数据属性构建数据温度计算模型。MPDAS调用数据划分引擎计算历史数据温度，创建数据划分表实现温度数据持久化存储，在服务器访问低峰期集中进行冷热数据划分，根据数据冷热标识将热数据抽取到高速存储介质。

MPDAS数据源包括MongoDB和Redis两种数据库，用于存储数据划分算法标识的冷热历史数据。MPDAS基于Sharding和Replica Set两种MongoDB集群模式建立分布式数据库并保存历史数据副本，实现数据存储模块高鲁棒性和高扩展性。建立Redis分布式集群分别存储各个MongoDB分片热数据，提高系统吞吐率。

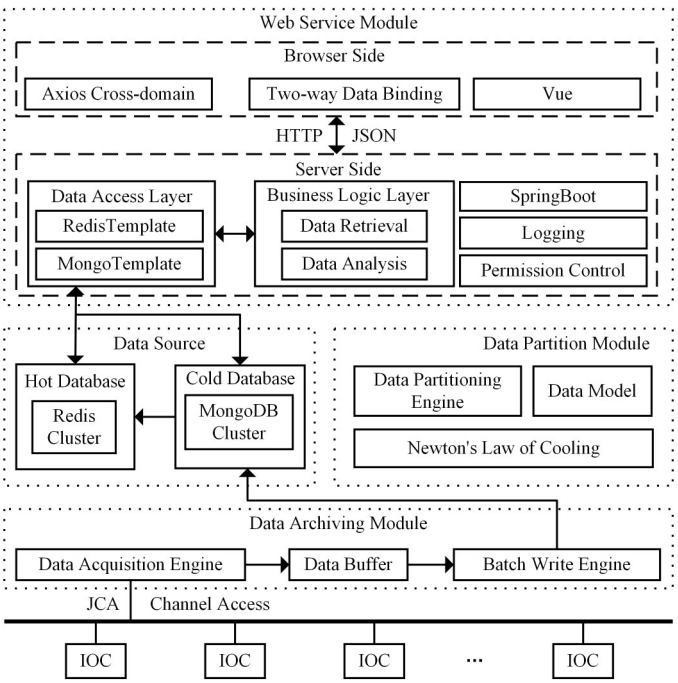

图1 MPDAS系统总体架构图
Fig.1 Overall architecture diagram of MPDAS system

Web服务模块采用前后端分离的系统架构，通过超文本传输协议（Hyper Text Transfer Protocol，HTTP）进行资源请求和数据响应，采用轻量级文本格式JSON进行数据交互。浏览器端采用基于MVVM模型的Vue框架开发单页面应用，通过双向数据绑定技术实现自动更新实时数据，运用Axios技术并配置代理服务器实现跨服访问服务器资源。服务器端采用MVC设计模式，将后端程序划分为业务逻辑层和数据访问层两个层次并实现访问权限控制和记录系统日志功能。业务逻辑层主要实现数据检索及分析、用户及系统管理和平台故障监控等业务逻辑。数据访问层屏蔽数据源驱动底层实现，提供抽象的统一访问接口。

2 系统实现

2.1 数据归档

MPRP物理信号数量繁多，部分脉冲数据产生频率较高，在磁体失超瞬间生成大量实时数据，预计20年内其数据规模将达到PB级。历史数据分为工

程数据和实验数据两种类型。工程数据自平台投入使用后持续产生，数据生成的频率范围为1~100 Hz。实验数据仅在科研人员进行实验时产生，采集频率从1 Hz~10 kHz不等。为存储MPRP生成的海量历史数据，首先，实现了EPICS数据归档插件获取IOC数据；其次，应用MongoDB建立实验数据库和工程数据库分别存储生成周期、采集频率差异较大的两种历史数据；最后，建立MPRP分布式数据库集群应对持续增长的海量数据存储需求。

MPDAS数据归档插件基于JCA API并采用多线程技术实现并行数据归档功能，各线程之间彼此独立、互不干扰，能够同时获取多路EPICS PV。MPDAS数据归档流程图如图2所示。通过分析所有信号采集频率，平衡划分PV集，加载properties格式配置文件生成CA上下文，创建多个数据采集线程与IOC服务器建立通道链接，为子PV集添加侦听器和监视器，监控PV变化，获取数据并存入缓存区。应用MongoDB批量写入功能，循环读取数据缓存区，按照PV名对数据分组，调用多线程数据存储任务将每组数据批量插入到MongoDB集合，提高数据存储速度。

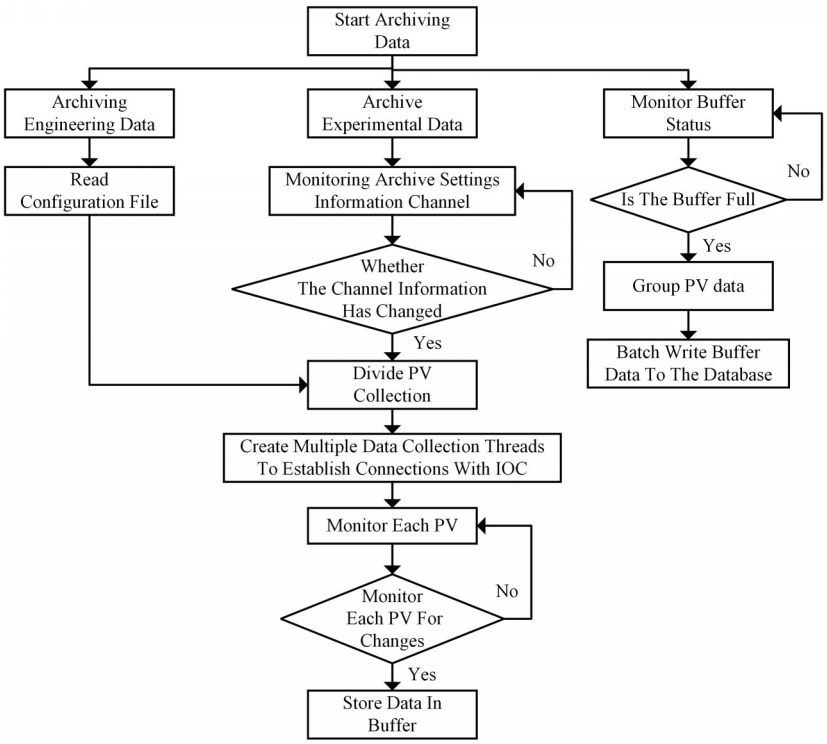

图2　MPDAS数据归档流程图
Fig.2　Flowchart of MPDAS data archiving

MongoDB是MPDAS主数据源，其工程数据库和实验数据库具有相似的文档逻辑结构。实验数据库设计两种集合模型，分别为ArchiveInfo集合和PVData集合。ArchiveInfo和PVData集合文档模型如图3所示。ArchiveInfo集合存储了实验名称、PV列表、数据单位及类型等实验总体性信息。PVData集合记录了具体的通道数据、时间戳、数值状态、报警等级等PV细粒度内容，通过实体数据模型外键属性建立对ArchiveInfo集合的依赖关系。工程数据库在保持PVData文档结构基础上对上一层次的集合模型进行更细粒度划分。由于工程数据的信号种类和PV集合是确定的，其数据生成具备可预见性，我们按照PV名和数据生成时间创建工程数据库PVData类型集合，并重构ArchiveInfo存储PVData元数据建立集合间对应关系。

MPRP分布式数据库集群采用Sharding和Replica Set两种模式，具备高扩展性和强可靠性。MongoDB分布式集群架构如图4所示，采用三台物理机搭建，每台物理机部署5个实例，包括1个Router Server、1个Config Server和3个分片实例。MPDAS为每个分片节点配置一个副本集，分别将分片和其副本部署于不同物理机，保证数据库集群在任意一台物理机宕机情况下仍能够正常提供服务。合适的分片键能够充分发挥auto-sharding机制，使数据读写分布更为均衡[16]。MPDAS写操作需求远大于读操作，片键的设计应首先考虑数据写分布性能。PVData集合的PV名字段基数较高、易于拆分且是常用的查询参数，以PV名作为片键能够在提高

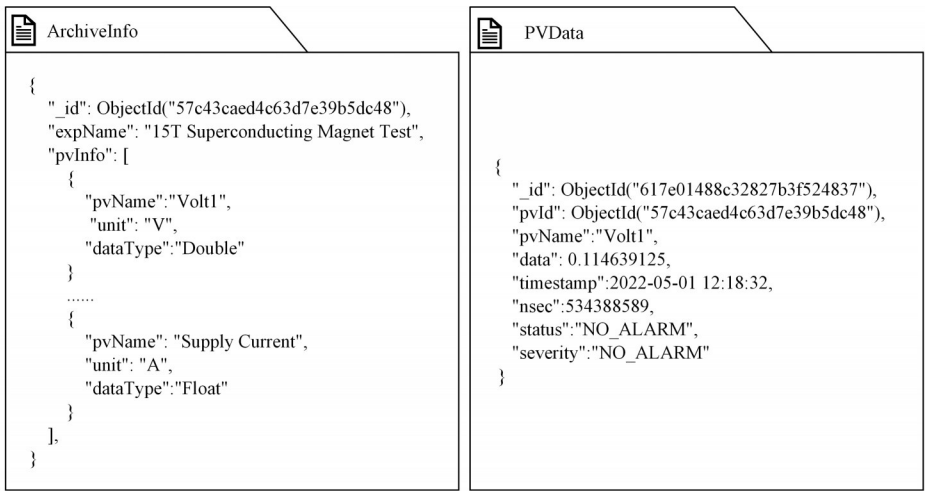

图3　ArchiveInfo和PVData集合文档模型
Fig.3　Collection document models of ArchiveInfo and PVData

数据写分布的同时获得出色的查询性能。在数据归档过程中，所有PV数据同时生成，PV名片键可以将实时数据均匀分发到各个分片节点，能够避免升序片键数据写热点问题。MongoDB片键在保证分布性的前提下需要具备良好的数据局部性，否则数据将分散存储到多个分片的不同数据块，导致查询操作频繁进行磁盘IO，消耗更多数据整合时间。PVData集合中的时间戳字段是典型的升序片键，其自增特性能够将连续时间数据持续写入同一数据块，减少查询操作时间消耗，提升数据读取性能。因此，我们基于PV名和时间戳的复合片键策略实现MongoDB分布式存储数据库。

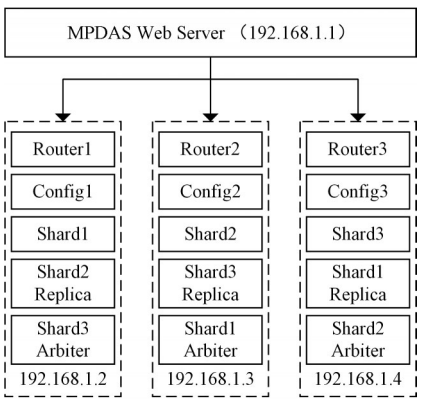

图4　MongoDB分布式集群架构
Fig.4　Distributed cluster architecture of MongoDB

2.2　数据划分

历史数据访问具有局部性特征，即少量数据被频繁访问，大部分数据很少或几乎不被访问[17]。历史访问频繁的数据称为热数据，反之，就认为是冷数据，将热数据存储到高速存储介质可以明显提高数据检索的响应速度[17]。数据冷热划分算法可以类比操作系统的缓存替换算法，传统的缓存替换算法包括FIFO（First In First Out）、LRU（Least Recently Used）、LFU（Least-Frequently Used）等[18–19]。FIFO、LRU、LFU分别根据数据存储先后顺序、访问时间、访问频率的单一特征区分冷热数据，存在局限性。本文克服以上三种传统算法的单一化不足，提出了一种综合数据访问时间、访问频率和存储顺序的多维度特征冷热数据划分算法（Division of Hot and Cold Data，DHCD）。

DHCD采用牛顿冷却定律构建对数据访问时间敏感的温度模型。牛顿冷却定律描述了高温物体在低温环境中其温度随时间呈指数衰减规律的变化过程，指数衰减数学模型被应用于多个自然科学领域，例如放射性衰变、RC电路电流减小、大气压力随海拔高度减少等[20–22]。现实世界数据的冷热程度同样是随时间衰减的，其降温过程与物体冷却过程类似，同牛顿冷却定律基本含义相一致[20,23]。数据温度在被访问后上升一定高度，随后在不被访问的时间段内快速“冷却”并趋向于0，符合指数衰减模型，因此可基于牛顿冷却定律建立历史数据随访问时间衰减的温度模型[23]。牛顿冷却定律[24–25]认为物体的冷却速率与其和环境的温差成正比关系，其公式如式（1）所示：

$$\frac{\mathrm{d}T(t)}{\mathrm{d}t} = -k\left(T(t) - T_v\right) \tag{1}$$

式中：T表示物体温度；T_v表示环境温度；t为时间；k为物体温度冷却速率与其和环境的温差的比例系数。

对式（1）求解可得物体温度式（2）：

$$T(t) = T_0 \mathrm{e}^{-kt} + T_v \tag{2}$$

对现实数据而言，环境温度没有意义。在仅考

虑数据访问时间的条件下，忽略环境温度，数据在t_n时刻的温度可由式(3)计算得到：

$$T(t_n) = T(t_{n-1})\mathrm{e}^{-k(t_n - t_{n-1})} \tag{3}$$

式中：t_{n-1}为数据上次受到访问的时间。

式(3)仅考虑了访问时间对数据温度的影响，然而，数据库中的数据冷热程度同样依赖于访问频率。根据LFU算法思想，同一时间段内数据访问次数越多的数据温度越高。当数据被访问时，其温度将得到一定增幅，将温度增幅定义为W，对式(3)变形可得兼顾访问时间和访问频率的温度计算式(4)：

$$T(t_n) = T(t_{n-1})\mathrm{e}^{-k(t_n - t_{n-1})} + W \tag{4}$$

数据个体的差异性决定了其重要性不尽相同，故W的取值不能简单地定义为常量，应根据数据属性单独计算。FIFO算法认为最先进入存储队列的数据未来被访问的概率最小，即数据存储时间越长的数据价值越低。因此，不同存储时间的数据在受到访问后，其增温幅度应有所不同，数据温度增幅随存储时间增加而减少。根据FIFO算法思想，采用历史数据存储时间计算不同数据被访问后的温度增幅，基于物质科学领域普遍用来描述物理量随时间变化过程的“e指数规律”构建温度增幅关于存储时间的指数衰减模型[21]。引入MPRP运行时间以限定温度增幅取值范围，W定义如式(5)所示：

$$W = c \times \mathrm{e}^{-\frac{d}{D}} \tag{5}$$

式中：c为一天内数据的访问次数；d表示数据存储时间；D表示MPRP运行时间。根据e指数函数性质，在单次访问(c=1)条件下，当数据存储时间等于MPRP运行时间时温度增幅取最小值e^{-1}，当数据存储时间趋向于0时温度增幅接近1，故单次访问的W取值范围为e^{-1}~1。

将式(5)代入式(4)得到数据温度的最终计算式(6)：

$$T(t_n) = T(t_{n-1})\mathrm{e}^{-k(t_n - t_{n-1})} + c \times \mathrm{e}^{-\frac{d}{D}} \tag{6}$$

式中：$T(t_{n-1})$表示数据在t_{n-1}时刻的温度；$T(t_n)$为数据在t_n时刻的温度。比例系数k能够调整数据冷却速率，可以根据数据划分需求进行选择，MPDAS设置k值为0.005。

DHCD通过多维度数据特征量化并标识数据冷热程度，利用式(6)可以轻松计算得到历史数据温度。为提高数据读取速度，MPDAS采用基于内存和SSD(固态硬盘)的混合存储架构存储历史数据。由于内存存在易失性，本文创建数据划分表对最近访问时间、访问次数、存储时间、数据温度等信息进行持久化存储。通过分段计算并持久化存储所有数据的温度值，MPDAS可以在任意时刻读取数据划分表获取历史数据温度信息。MPDAS数据划分流程图如图5所示。根据数据温度对历史数据排序，从MongoDB数据库中抽取热数据并冗余存储到Redis数据库。调用热数据存储监控线程对Redis存储空间使用率进行实时监控，如果Redis中的数据量达到设定阈值，则批量淘汰温度较低的历史数据。当系统发生数据访问时，能够自动调用数据访问监控模块，重新计算一天内数据的访问次数。如果频繁进行温度计算、热数据抽取等操作，将导致处理器、内存等资源占用率过高，本文采用集中处理策略在服务器负载较轻的时间段对历史数据统一进行冷热划分，该时间段一般设置为3:00 AM到5:00 AM。

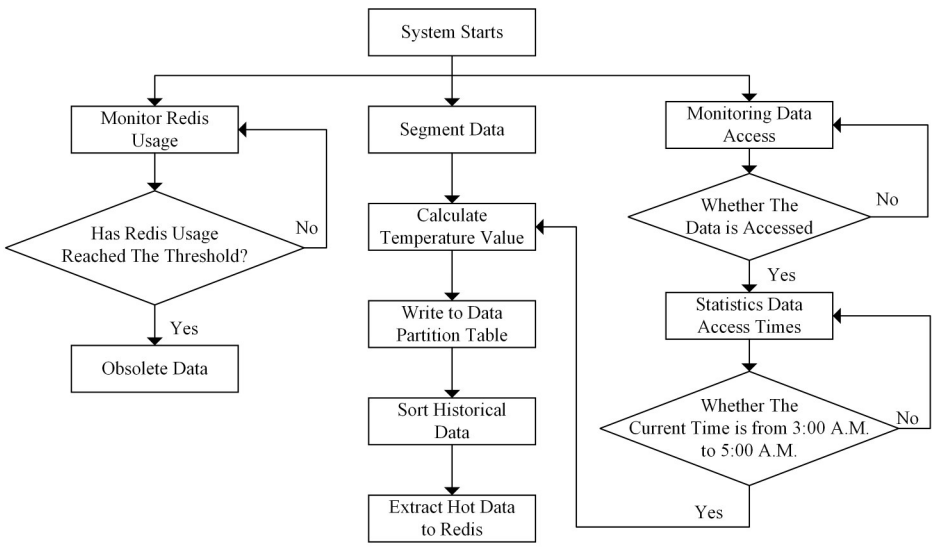

图5　MPDAS数据划分流程图
Fig.5　Flowchart of MPDAS data division

MPDAS通过基于Redis内存数据库的热数据冗余存储机制提高数据检索的响应速度。当用户访问历史数据时，系统首先到Redis中查找。如果用户请求在Redis命中，根据查询条件中时间范围和PV名称判断所命中数据是否完整。若Redis命中数据完整，则返回数据并更新数据划分表。如果Redis未命中或命中数据不完整，系统将根据查询条件调用MongoDB数据检索服务。MongoDB数据库存储了完整的历史数据，能够命中所有合法数据检索请求。

3 系统测试

MPRP正在筹建中，各子系统工程建设工作推进扎实有序。为评估软件质量并验证数据划分算法的功效，MPDAS分别开展系统功能测试和热数据命中率测试。系统服务器搭载8核16线程Intel Xeon E5-2630@2.4 GHz CPU、128 G Samsung DDR4 1 866 MHz内存、2 TB西部数据固态硬盘。

3.1 系统功能测试

在中国科学院等离子体物理研究所超导电工实验室低温测试平台对MPDAS进行黑盒测试，在现实低温超导实验环境中使用真实物理数据评价系统功能特征，在不考虑程序内部结构和特性的情况下检测系统功能是否正常。低温测试平台如图6所示，平台中包括一台电源机柜（包含2台1 200 A/10 V规格的AMETEK电源）、3台研华工控机、6台Keithley 2182A、2台LakeShore224、2台DAQ6510等仪器设备。测试系统搭建流程如下：

1）采用5台Keithley 2182A、一台LakeShore224以及电源、工控机等设备搭建采集系统。

2）使用基于Labview的数据采集软件通过CA_Lab接口向局域网发送数据。

3）建立IOC服务器并定义运行数据库接收数据采集软件的上传数据。

4）更改MPDAS配置文件连接IOC并启动应用服务器和数据库服务器。

图6 低温测试平台
Fig.6 Snapshot of cryogenic test platform

通过15T超导磁体测试实验测试MPDAS面向EPICS的数据归档功能，并在浏览器检索实验数据。MPDAS数据检索界面截图如图7所示，图7中展示了15T超导磁体测试实验中5个PV的检索结果。经过功能测试，历史数据归档、数据搜索及处理、曲线属性设置、清除数据、图表缩放及平移、导出数据等功能满足系统设计要求。

3.2 热数据命中率测试

为提升数据检索速度，本文在FIFO、LRU、LFU三种传统算法的基础上提出了一种基于牛顿冷却定律的冷热数据划分算法，划分历史数据为冷数据和热数据并抽取热数据到内存型数据库Redis，通过内存的高I/O处理性能和DHCD的高命中率减少平均访问时延。在内存固定的条件下，MPDAS的热数据

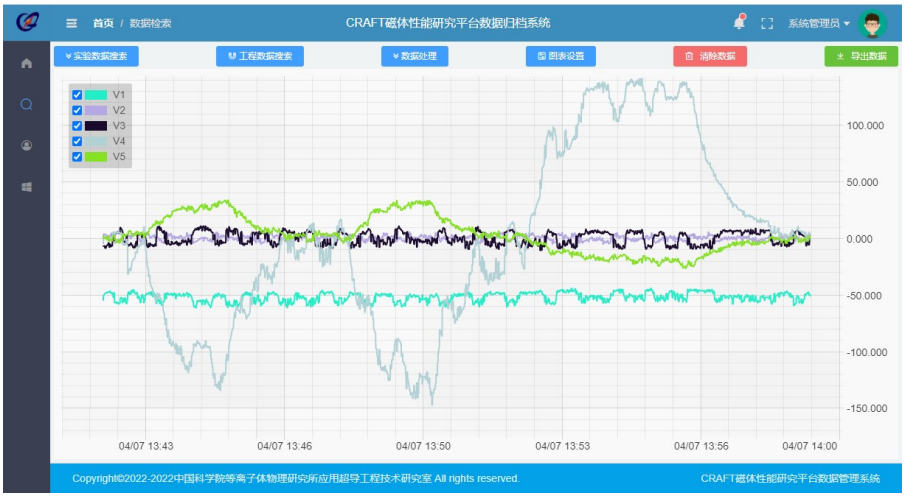

图7 MPDAS数据检索界面截图
Fig.7 Screenshot of MPDAS data retrieval interface

库命中率直接决定了系统的数据检索响应速度。本文通过对比DHCD与FIFO、LRU、LFU的热数据库命中率来验证冷热数据划分算法带来的数据检索性能提升。由于历史数据的用户访问特性在统计学规律上服从Zipf分布，我们采用Python语言的NumPy工具生成10 000条满足Zipf分布的测试数据模拟用户访问行为，分别从用户访问次数和热数据库容量两个角度对比热数据命中率[23,26-27]。

热数据命中率随访问次数变化趋势如图8所示。其中，横坐标为用户访问次数，纵坐标为热数据命中率。本实验在固定热数据库容量前提下进行4种算法在不同访问次数条件下热数据命中率的实验验证，热数据库容量为总数据量的5%。从图8可知，4种算法热数据命中率在用户访问测试开始后迅速上升并且在开始阶段命中率变化基本一致，原因是热数据库在实验测试初期阶段存储了所有用户访问的热点数据。在实验开始前Redis为空，若发生用户访问未命中事件，则将目标数据直接调入热数据库直到数据量达到设定阈值。4种算法命中率在未发生缓存替换前随着热数据持续调入而快速上升，在同一用户访问测试模式下热数据库存储状态一致，故呈现出相同的命中率变化。随着用户访问次数增加，Redis逐渐被填满，开始出现热数据置换现象。FIFO算法在进行缓存替换时优先丢弃最早进入Redis的历史数据，由于过早换出高热点数据，其热数据命中率快速下降并趋于平缓。LRU、LFU和DHCD开始进行热数据替换时由于频繁换入换出行为导致命中率波动较为明显，随着访问次数增加，命中率变化逐渐稳定。从图8可以看出，LRU命中率随访问次数递减，LFU命中率缓慢增长后逐渐稳定，而DHCD命中率明显上升，且相对其他算法的命中率提升随访问次数不断增加。总体而言，本文提出的DHCD算法拥有更高的热数据命中率，更加适用于高访问率的数据归档系统。

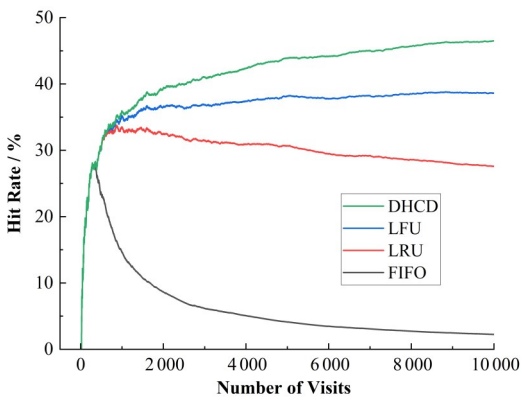

图8 热数据命中率随访问次数变化趋势
Fig.8 Trend of hot data hit rate with number of accesses

不同热数据库容量下命中率对比如图9所示。其中，横坐标为热数据库数据量占总数据量百分比，纵坐标为热数据命中率。从图9可知，4种算法的命中率均随热数据库容量递减，DHCD算法在不同层级热数据库容量的命中率对比中均拥有最高的热数据命中率。与此同时，DHCD算法相比其他算法的命中率提升随热数据库容量减少而不断增加。DHCD在热数据库存储30%的数据量时，命中率最高，但相比FIFO、LRU、LFU的命中率提升仅分别为35.54%、3.21%、1.03%，以上数值在热数据存储容量为1%时分别增加到了38.05%、26.91%和11.06%。可以看出，DHCD算法相比其他算法的命中率提升在热数据占总数据量比例越小时越有优势。由于服务器内存容量有限，而历史数据总量却随运行时-间不断增长，使得这一特性在大科学装置数据库系统中拥有更高的应用价值。因此，本文提出的冷热数据划分算法相比FIFO、LRU、LFU等算法命中率更高且更加稳定，通过提升热数据命中率能够有效减少数据查询平均响应时间，提升数据检索速度。

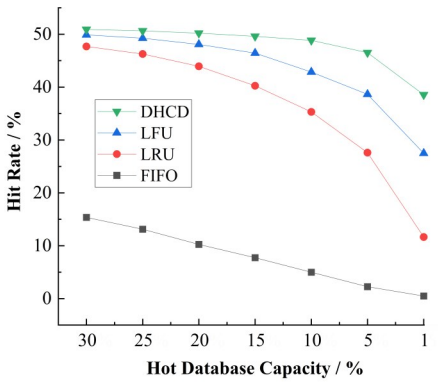

图9 不同热数据库容量下命中率对比
Fig.9 Comparison of hit rate under different hot database capacities

4 结语

为实现CRAFT磁体性能研究平台历史数据统一、长期、稳定地存储，MPDAS基于MongoDB和Redis数据库并应用企业级Web技术实现了交互友好、吞吐量高、扩展性强的数据归档系统。本文应用EPICS JCA接口并基于多线程技术设计数据归档插件，采用MongoDB分片和副本集策略将数据写入压力均匀地分发给各个数据库节点，同时实现了扩展灵活、配置简单、安全性高的分布式数据存储架构。受到FIFO、LRU、LFU三种缓存替换算法启发，本文基于牛顿冷却定律提出一种综合数据存储顺序、最近访问时间和数据访问频率的三个维度特征的温度模型，量化历史数据未来被频繁访问程度，设计数据

划分算法划分冷热数据，将热数据抽取到Redis数据库，降低热数据库数据检索时间复杂度，提高数据检索响应速度。经过系统测试，MPDAS系统功能能够满足科研人员使用需求，其冷热数据划分算法相比传统算法拥有更高的热数据命中率。通过提高热数据命中率减少数据查询平均响应时间，提升了历史数据检索速度。下一步我们将进一步对数据实时监控进行研究，以方便科研人员实时观测数据变化，及时掌握平台运行状态。

作者贡献声明 马树良：酝酿、设计并实现系统，起草论文初稿；刘华军：指导，对文章知识性内容作批评性审阅；刘方：指导系统测试，获取研究经费，行政及技术支持；张舒庆：实施研究；李童：指导系统设计，材料支持；施毅：指导，支持性贡献；郭亮：指导，支持性贡献。

参考文献

1 Hu L, Zhang Q, Zhu Z, *et al*. Conceptual design of cryogenic system for comprehensive research facility for key fusion reactor core systems[J]. IOP Conference Series: Materials Science and Engineering, 2019, **502**: 012113. DOI: 10.1088/1757-899x/502/1/012113.

2 Zhou H S, Yuan X G, Li B, *et al*. A new high flux plasma source testing platform for the CRAFT project[J]. Journal of Fusion Energy, 2020, **39**(6): 355 – 360. DOI: 10.1007/s10894-020-00277-y.

3 Li T, Liu H J, Wu Y, *et al*. Control and diagnostic system for CFETR CSMC testing platform[J]. IEEE Transactions on Plasma Science, 2020, **48**(6): 1789 – 1792. DOI: 10.1109/TPS.2019.2946193.

4 李童. CSMC磁体测试平台诊断控制系统设计与研究[D]. 合肥: 中国科学技术大学, 2020.
LI Tong. Design and research of diagnostic control system for CSMC magnet test platform[D]. Hefei: University of Science and Technology of China, 2020.

5 About EPICS [EB/OL]. 2022-05-26. https://epics-controls.org/about-epics/.

6 祝晴, 蒋舸扬, 李林, 等. 上海光源数据存档引擎的设计与实现[J]. 核电子学与探测技术, 2007, **27**(3): 521 – 525. DOI: 10.3969/j.issn.0258-0934.2007.03.024.
ZHU Qing, JIANG Geyang, LI Lin, *et al*. Design and implementation of archiver engine based on XML technology for SSRF[J]. Nuclear Electronics & Detection Technology, 2007, **27**(3): 521 – 525. DOI: 10.3969/j.issn.0258-0934.2007.03.024.

7 李洛峰, 王春红, 王金灿, 等. EPICS数据获取与查询系统的研究与实现[J]. 核电子学与探测技术, 2013, **33**(9): 1043 – 1046, 1061. DOI: 10.3969/j.issn.0258-0934.2013.09.001.
LI Luofeng, WANG Chunhong, WANG Jincan, *et al*. Research and implementation of data acquiring and data query system for EPICS[J]. Nuclear Electronics & Detection Technology, 2013, **33**(9): 1043 – 1046, 1061. DOI: 10.3969/j.issn.0258-0934.2013.09.001.

8 Wang C H, Li L F. A new data acquiring and query system with oracle and EPICS in the BEPCII[C]//Proceedings of ICALEPCS2015. Melbourne, Australia, 2015: 865−868.

9 罗江波, 郭玉辉, 刘海涛, 等. 加速器驱动次临界系统注入器II数据归档系统[J]. 强激光与粒子束, 2016, **28**(10): 134 – 140. DOI: 10.11884/HPLPB201628.160034.
LUO Jiangbo, GUO Yuhui, LIU Haitao, *et al*. Data archiving system for injector II of accelerator driven sub-critical system[J]. High Power Laser and Particle Beams, 2016, **28**(10): 134 – 140. DOI: 10.11884/HPLPB201628.160034.

10 Song Y F, Li C, Xuan K, *et al*. Automatic data archiving and visualization at HLS-II[J]. Nuclear Science and Techniques, 2018, **29**(9): 129. DOI: 10.1007/s41365-018-0461-6.

11 李嘉曾, 韩利峰, 李丹清, 等. 基于大数据平台的EPICS历史数据归档系统[J]. 核技术, 2019, **42**(11): 110603. DOI: 10.11889/j.0253-3219.2019.hjs.42.110603.
LI Jiazeng, HAN Lifeng, LI Danqing, *et al*. EPICS historical data archiving system based on big data platform[J]. Nuclear Techniques, 2019, **42**(11): 110603. DOI: 10.11889/j.0253-3219.2019.hjs.42.110603.

12 乔予思, 雷革. BEPC-II新型数据存档软件系统设计[J]. 核电子学与探测技术, 2017, **37**(2): 152 – 155. DOI: 10.3969/j.issn.0258-0934.2017.02.010.
QIAO Yusi, LEI Ge. A novel data archiving and retrieving software system for BEPC-II[J]. Nuclear Electronics & Detection Technology, 2017, **37**(2): 152 – 155. DOI: 10.3969/j.issn.0258-0934.2017.02.010.

13 Kikuzawa N, Ikeda H, Kato Y, *et al*. Status of operation data archiving system using Hadoop/HBase for J-PARC [C]//Proceedings of PCaPAC2014. Karlsruhe, Germany, 2014: 193−195.

14 Kikuzawa N, Yoshii A, Ikeda H, *et al*. Development of J-PARC time series data archiver using distributed database system[C]//Proceedings of ICALEPCS2013. San Francisco, USA, 2013: 584−586.

15 MongoDB tutorial[EB/OL]. 2022-05-26. https://www.

mongodb.org.cn/tutorial/.

16 熊峰, 刘宇. 基于MongoDB的数据分片与分配策略研究[J]. 计算机与数字工程, 2019, **47**(4): 892 – 897. DOI: 10.3969/j.issn.1672-9722.2019.04.029.
XIONG Feng, LIU Yu. Research of the strategy of data fragmentation and allocation based on MongoDB[J]. Computer and Digital Engineering, 2019, **47**(4): 892 – 897. DOI: 10.3969/j.issn.1672-9722.2019.04.029.

17 梁懿, 陈又咏, 李森, 等. 自适应业务场景的数据库冷热数据识别算法[J]. 现代电子技术, 2022, **45**(7): 107 – 111. DOI: 10.16652/j.issn.1004-373x.2022.07.020.
LIANG Yi, CHEN Youyong, LI Sen, *et al*. Database cold and hot data identifying algorithm for transaction scenery adaption[J]. Modern Electronics Technique, 2022, **45**(7): 107 – 111. DOI: 10.16652/j.issn.1004-373x.2022.07.020.

18 Dan A, Towsley D. An approximate analysis of the LRU and FIFO buffer replacement schemes[C]//Proceedings of the 1990 ACM SIGMETRICS Conference on Measurement and Modeling of Computer Systems - SIGMETRICS '90. University of Colorado, Boulder, Colorado, USA. New York: ACM Press, 1990: 143 – 152. DOI: 10.1145/98457.98525.

19 Robinson J T, Devarakonda M V. Data cache management using frequency-based replacement[C]//Proceedings of the 1990 ACM SIGMETRICS Conference on Measurement and Modeling of Computer Systems - SIGMETRICS '90. University of Colorado, Boulder, Colorado, USA. New York: ACM Press, 1990: 134 – 142. DOI: 10.1145/98457.98523.

20 许佳欣. 面向冷热数据的存储机制研究与实现[D]. 成都: 电子科技大学, 2021.
XU Jiaxin. The research and implementation of storage mechanism for hot and cold data[D]. Chengdu: University of Electronic Science and Technology of China, 2021.

21 陈学文, 江秦, 张家伟, 等. 对大学物理课程中“e指数规律”的探讨[J]. 物理与工程, 2020, **30**(2): 49 – 53. DOI: 10.3969/j.issn.1009-7104.2020.02.010.
CHEN Xuewen, JIANG Qin, ZHANG Jiawei, *et al*. Discussion on “e-exponential law” in college physics[J]. Physics and Engineering, 2020, **30**(2): 49 – 53. DOI: 10.3969/j.issn.1009-7104.2020.02.010.

22 陈昌兆. 一阶RC电路的深度透视: 概念、思想和方法[J]. 物理与工程, 2018, **28**(5): 44 – 49. DOI: 10.3969/j.issn.1009-7104.2018.05.008.
CHEN Changzhao. Depth perspective on first-order rccircuits: concept, principle and method[J]. Physics and Engineering, 2018, **28**(5): 44 – 49. DOI: 10.3969/j. issn. 1009-7104.2018.05.008.

23 解玉琳. 基于数据温度的冷热数据识别机制研究[D]. 杭州: 浙江大学, 2019.
XIE Yulin. Research on recognition mechanism of hot and cold data based on data temperature[D]. Hangzhou: Zhejiang University, 2019.

24 刘素美, 马红章, 尹玉芳, 等. 变质量冷却法测定油品比热容[J]. 大学物理, 2017, **36**(3): 25 – 27. DOI: 10.16854/j.cnki.1000-0712.2017.03.007.
LIU Sumei, MA Hongzhang, YIN Yufang, *et al*. The changing mass cooling method for measuring the specific heat of oil[J]. College Physics, 2017, **36**(3): 25 – 27. DOI: 10.16854/j.cnki.1000-0712.2017.03.007.

25 熊欢欢, 秦政, 秦诗童, 等. 基于牛顿冷却定律的炉温曲线优化模型[J]. 实验科学与技术, 2022, **20**(2): 9 – 14. DOI: 10.12179/1672-4550.20210278.
XIONG Huanhuan, QIN Zheng, QIN Shitong, *et al*. Optimal model of furnace temperature curve based on Newton's law of cooling[J]. Experiment Science and Technology, 2022, **20**(2): 9 – 14. DOI: 10.12179/1672-4550.20210278.

26 方宇哲. 无线融合网络中基于强化学习的分布式缓存技术研究[D]. 上海: 上海交通大学, 2020.
FANG Yuzhe. Research on distributed cache technology based on reinforcement learning in wireless converged networks[D]. Shanghai: Shanghai Jiao Tong University, 2020.

27 张志伟. 基站环境下媒体流行度预测与缓存策略研究[D]. 合肥: 中国科学技术大学, 2016.
ZHANG Zhiwei. Research on media popularity prediction and caching under radio access network[D]. Hefei: University of Science and Technology of China, 2016.