

基于预留缓存的并发访问及多任务并行支撑技术研究

党宏杰¹, 秦留洋², 韩惠婕²

(1. 北京跟踪与通信技术研究所, 北京 100094; 2. 北京航天测控技术有限公司, 北京 100041)

摘要: 在航天装备在役考核试验执行中, 对于存在时空重叠的多个任务以及子任务, 需要应用并行执行支撑技术, 从而完善任务分配的合理性、提高任务处理的效率、数据处理的实时性等约束性质; 以典型航天装备在役考核综合评估任务为对象, 从任务流层面和数据流层面, 对并行控制与处理算法进行研究, 采用基于预留缓存的多线程并行并发技术, 对多任务进行并行处理和执行; 实现任务的合理自主规划、数据处理与读写的多线程触发与执行; 依托自主研发的航天装备在役考核综合评估系统, 对方法进行验证; 系统运行结果表明, 采用基于预留缓存的并发访问及多任务并行支撑技术可以高效、快速地对大纲设计、数据采集等多个过程进行并发执行控制, 对多源数据进行并行处理计算, 有效提高了在役考核试验执行和综合评估的效率。

关键词: 航天装备; 在役考核评估; 多任务并发执行; 多线程并行处理; 预留缓存

Research on Concurrent Access and Multi-Task Support Technology Based on Reserved Cache

DANG Hongjie¹, QIN Liuyang², HAN Huijie²

(1. Beijing Institute of Tracking and Communication Technology, Beijing 100094, China;

2. Beijing Aerospace Measurement and Control Technology Co., Ltd., Beijing 100041, China)

Abstract: Multi-Task Concurrent Execution Technology (MTCET) is applied to support time-space-overlapped tasks and sub-tasks of analysis and evaluation for In-Service Inspection (ISI) of Space Equipment, which is able to improve the rationality of task assignment, efficiency of task execution and instantaneity of data process. Concurrent technology based on Multi-thread and reserving cache is adapted to execute multi-task in the aspects of both task flow and data flow of overall assessment of Space Equipment. The method is verified based on the self-developed Overall Evaluating System of Space Equipment, whose result shows that several processes such as outline design and data collection can be concurrently executed and controlled effectively by applying presented multi-task concurrent execution technology, and multi-source data can be processed and calculated concurrently, which lifts the efficiency of in-service inspection test and overall evaluation.

Keywords: Space Equipment; multi-task concurrent execution; multi-thread; reserving cache

0 引言

航天装备体系的建设为各类型作战应用提供信息支援和支撑保障^[1]。涵盖多类、多型装备的组成形式, 决定了其体系的复杂性和任务的多元性。由于航天装备对作战保障具有全面性、全局性的特点, 大量资源被投入用于航天系统的建设, 导致太空中各类航天器呈爆发式增长^[2]。目前多元作战力量、多维作战任务、高实时性是航天装备的突出特点^[3], 大量的计算密集和时间敏感型应用在通信资源和计算资源方面对航天装备的通信基础设施和数据计算服务等设施的建设和使用提出了更高的要求^[4-7]。同时航天装备体系本身也面临着支持卫星数目多、多卫星同时过境、卫星相继过境间隔时间缩短、卫星任务需求多、卫星任务种类多、卫星任务突发等难题^[8], 单线程串行任务执行模

式难以满足任务数量、形式、种类、要求等多维需求, 严重影响卫星应急任务服务需求的应急能力、应变能力和处理能力。

目前, 从任务调度层面进行任务约束规划开展的研究较多, 有了一些成熟的算法模型。文献 [9] 对任务调度问题进行了研究, 在可视时间窗口内考虑任务权重, 建立了整数规划模型。文献 [10] 对任务调度问题的约束规划模型进行研究, 设置了问题中特有的约束, 最终建立了相应的约束规划模型。文献 [11] 通过建立时间约束着色 Petri 网模型, 建立了统一的设备库所, 按类型和服务要求对设备着色。但是对并发任务执行的底层数据层面支撑研究较少, 如何将任务流层面的合理规划与数据流层面的并行计算进行融合, 从而实现真正的“并行处理”, 成为航天装备

收稿日期: 2022-12-28; 修回日期: 2023-03-20。

作者简介: 党宏杰(1983-), 女, 山东枣庄人, 硕士, 助理研究员, 主要从事航天装备仿真评估方向的研究。

引用格式: 党宏杰, 秦留洋, 韩惠婕. 基于预留缓存的并发访问及多任务并行支撑技术研究[J]. 计算机测量与控制, 2023, 31(5): 255-259.

领域多任务控制研究的重点。

本文结合航天装备体系在役考核综合评估的任务执行实际情况,设计实现了一种基于预留缓存的兼顾任务层面和数据层面的多任务并行执行支撑方法,应用数据的并行处理和线程并行缓存等方法,对任务分类、并行组织、数据读写、数据处理等过程进行并行迭代,实现航天装备在役考核综合评估的多任务并发支撑。

通过自主研发的航天装备在役考核综合评估系统,采用基于预留缓存的数据处理方法,实现对多个评估任务、多层级评估任务的有效、灵活的并行控制和运行,实现多任务并发访问及并行运行支撑。

1 数据多线程并行处理技术

航天装备在役考核综合评估涉及复杂的多维数据,包括任务训练模拟数据、任务执行过程数据、装备在役环境数据、装备实例主体数据、装备作用对象数据、装备使用维护数据等^[12],数据内容涵盖航天装备研制、试验、生产、使用和任务执行等多个周期和节点,数据种类包含航天装备遥测数据、遥控数据、测定轨数据、性能参数数据、地面仿真数据等多维数据。为了提高繁杂、庞大、多元数据的数据处理效率,设计并实现了基于多线程的数据并行处理技术,并在通用技术的基础上适配应用于航天装备在役考核这一特殊领域。

数据处理过程被划分为数据切分、数据读取、数据处理、数据存储和结果反馈等多个步骤。图 1 为数据多线程并行处理技术的原理示意图。

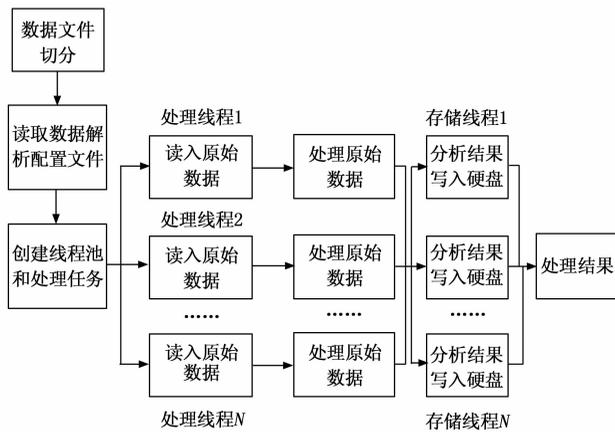


图 1 数据多线程并行处理技术的原理示意图

1.1 数据切分

首先,对引接的航天装备原始数据进行切分,数据切片的切分准则和标准由数据类型和数据的应用场景决定。数据切分的参数由数据应用方进行初始设定,根据数据业务和数据逻辑将原始数据切分为可被独立处理的碎片单元。

航天装备数据需要依据数据的分布特征、来源和时间序列进行细致的切分分片处理。数据在采集端就会设置相应的采集节点标志和时间标志,通过智能匹配,生成节点、时间、数据三者的映射关系,生成对应的数据标签,从而

为数据切片提供基础和依据。

从上述分析中可以看出,数据标签的准确程度和合理性的最主要表征是数据的时空合理性,因此对航天装备而言,数据的时间标签和任务执行中数据产生时的装备姿态具有最显著的关系,这两个维度上的数据标签也就成为数据切片过程中最重要的校标依据。

在获取数据过程中,为实现数据的传输效率,采用连续记录的整条数据的传输和下载方式,将其下载到处理单元之前的节点。节点的内部时钟标准校时设备的时间进行比对,进行时间的校准并获取实际数据的记录时间和标准校时设备记录的时间之间的校准偏差值。

对于装备姿态,以方位偏移角度进行表征,航天装备的相应传感器在记录数据的过程中会重点记录当前的装备姿态参数,从而为后续的数据切片提供空间维度的依据。

数据切分是获取合理的数据切片的重要手段,主流的数据切分包括垂直切分、水平切分、混合切分三种方式^[13]。由于混合切分交替使用水平和垂直方式,融合了两种切分方式下的方法优势,具有切分方法多样、过程灵活可控的特点,因此被广泛应用。

在进行数据切分时,通常根据各个数据标志节点实现数据连贯、全面、合理、合规的切分。通过获取并应用不同数据节点的时间漂移值,在数据提取时,就计算出标志节点在此刻的时钟漂移,完成校正数据的提取和预处理。对于如航天装备的姿态数据、载荷数据等数据内容,会在数据提取时用节点在此刻的姿态参数或载荷参数对采集数据进行处理和校正^[14]。

在航天装备体系的数据库系统中,切分后获取的多个数据片段被存储到不同的相应数据库里。

切分完毕之后根据各个单元数据加载对应的数据解析协议,并通知处理单元创建对应的线程池和处理任务,进入数据的并行处理阶段。

该阶段针对原始数据进行处理,采用单线程模式,保证数据处理的完整性、准确性和流程的连续性。

1.2 数据读取

在完成完整数据的切分之后,进入数据处理的流程。数据处理过程主要包括“①数据读取(输入)—②数据处理(计算)—③数据写入(存储)”。根据业务特点,数据的读取和计算是紧密连接的过程,拆分处理不易于数据处理的接续性,为数据的组织带来额外负担,因此单片的数据读取和数据计算应设计为串行的单线程过程,即①和②组成完整的处理线程。

数据读取作为数据计算处理的输入和基础,分多个线程,通过数据接口对分片数据进行引接。

1.3 数据处理

数据处理主要对读取的单片数据进行解析、转换、计算、统计和分析。每个处理线程的数据处理源均为该线程读取的单片数据,保证了数据处理的连续性和结果的一致性。按照数据切片的规范,单片数据之间存在一定的数据

重叠,以保障数据抽取、插值等处理的完整性。

为提高数据解析与处理的伸缩性、延展性、通用性和易用性,满足多任务、多型号、多装备、多场景的数据信息的差异化、分级化、高并发、多剖面等应用需求,对航天装备的多维数据处理过程和方法进行形式化与标准化描述,通过模型化语言,构建多层解耦的数据协议模型^[15]。通过深化深入地分析和梳理航天装备多维数据的数据处理流程,航天装备在役考核数据的协议模型可以按照数据层次分为源层、准备层、逻辑层和加工层等4个不同的层面,依据数据处理的粗细粒度,实现由底向上层级递进,每层的处理资源由下一层提供^[16]。其中:

源层按遥测信息、遥控信息、测控信息、地勤信息、环境信息、通信信息等不同方向和数据内容要求对接数据源,引接数据,为数据处理做数据资源准备。

准备层将源层引接的多维数据按照预设的帧结构、包结构和数据组织规则等进行解析和数据域的预处理,为类型匹配、方法匹配等后续的匹配操作提供依据。

逻辑层根据航天装备在役考核的特殊业务特点和数据帧组织形式,将数据以数据流的形式进行分类,同时针对不同类型数据构建标准规则库,以库的形式,实现通用逻辑标准的确立。

加工层实现对解析后数据按需进行格式转换、关联分析、统计分析、数据主题生成、数据存储与转发等数据层面的应用加工,并对数据进行精细化、多级化、通用化、统一化的管理。

1.4 数据存储

数据存储通过数据的写操作将数据备份到相应的介质,为后续的数据应用提供基础。常见的数据存储包括两种存储模式,一种是实时存储,即为每个数据处理线程配置一个数据存储线程,即时存储处理过的数据。该模式下,数据存储和数据处理为串行流程,每个单片数据的存储过程必须在本数据完成处理之后进行,在保证实时性的情况下,大幅降低了存储效率,存在存储计算资源的浪费。

在实际应用中,单片分数据的应用对实时性的要求并不高,因此采用串行存储会造成大数量下的数据积压。为了尽可能快速、高效地将大量通过处理线程完成数据处理的结果工程量数据(即已经具备清晰的数据含义和表达性与可读性的规范化数据)进行存储,采用了基于预留缓存的多线程并行存储技术。

整个数据流的后处理过程可以理想化地区分为数据处理过程和数据存储过程,分别安排相应线程进行处理。数据存储线程和处理线程独立运行,每个存储线程按自然顺序对任一个处理线程处理结束的单片进行存储,并对数据片进行标记,实现数据的按序存储。多线程并行存储可以最大化利用存储资源,提高数据的计算分析效率。

1.5 结果反馈

在分片数据完成存储之后,根据数据片的序列关系对数据进行重组和整合,得到处理结果,并按照完整数据形式进

行数据结果的校验和反馈,为数据的后续应用提供基础。

2 基于预留缓存的数据读写

航天装备体系承载的有效载荷正朝着数量众多、功能多元、性能严苛的方向发展,航天装备在执行相应的任务和服役考核等试验鉴定过程中承载的数据流的数据量、准确度、实时性等要求越来越高。大量、高速的数据输入和处理需求,数据的存储器和处理器过长的启动时间和反应过程会增加装备数据和载荷数据丢失或者出错的风险^[17]。而数据文件读写过程是数据处理、存储和应用的关键过程,是依照特定的数据协议和规范对数据进行提取、保存的过程描述。因此如何减少航天装备数据存储与处理系统的数据读取、写入等关键过程时间、提高快速访问速度成为解决上述问题的重要方面。

为实现快速高效的数据读写,通常采用缓存技术对数据进行缓存加载,如前所述,数据的读取和存储在大数据量的约束下,需要进行并行处理。而预留缓存和多线程并发是数据并行处理的重要手段,通过在任务调度层面和数据计算与处理层面融合采用基于预留缓存的多线程并发技术,可以有效提高航天装备的任务执行效率,提高作战保障能力。

所谓“线程”,指的是计算机系统程序中的一个执行流,是程序执行的最小单位。每个线程都有独立的寄存器、指针、程序计数器。多线程就是一个进程中同时包含多个执行流,且每个执行流互不影响的独立完成各自任务。多线程方法作为一种并发型、多任务的工作机制,具有提高程序响应、改善程序结构、使多核处理和计算资源可以并行工作的优点^[18],从而大大提高处理效率。

数据读取按照数据切片的结果,依照预定分片数量参数设计缓存,处理线程1到处理线程N为一组并行线程,进行同步处理分析,有效提高数据处理的效率。而数据写入则有高占用率高实时性串行写入和低占用率低实时性的并行写入两种模式。

针对数据的处理结果,在以结果工程量为存储模式的前提下,系统对数据的实时性要求降低,而对写入效率要求明显提高。因此串行写入虽然可以提供更高的实时性,却会占用和浪费存储资源,造成效率的下降。并行写入通过拆分文件写操作,实现了有限资源的高效利用,是支撑任务并行化处理的有效手段,是航天装备试验鉴定和任务执行过程中的数据处理的支撑。

数据在正式应用之前需要进行处理和存储,传统数据处理流程中基于并行时序理念,将处理线程作为存储线程的严格前序线程,导致一个数据片段的数据存储只对应一个特定数据处理线程的处理结果。数据处理线程组和存储线程组按原始数据的时间递增序列进行同步分析。不难看出,这样的处理方式数据一一对应,数据处理的实时性更强,但要付出存储线程轮空等待的代价,对计算资源的要求高,导致系统占用率较高,在数据量过大或数据分片过

多的情况下，数据处理的效率会大幅降低。

另外可以通过开辟缓存区，将碎片化的数据处理结果进行缓存，并进行再拼接，然后统一进行存储。此种方法实时性和并行性较差，但是系统负担较小，出错率低。

传统的基于数据流时序和数据处理线程的数据文件读写技术，不能在大数据量、多构型数据分别处理的北京下雪，满足多源数据处理的处理和并行存储要求。

为提高数据存储的效率，在多线程并行存储中采用了内存映射技术和基于数据分片的数据并行处理技术，通过缓存预留加载，设计了基于预留缓存的数据并行读写技术。该技术有效的解决了常规数据存储技术只能逐字节地写入数据存储文件的问题，可以实现结果数据按文件地址快速、并行地写入工程量文件。如图 2 所示，通过预留缓存实现了数据读取速度的有效提高，缓存中保留了规则话存储的数据片段，依据数据切分的结果和片段数量，系统自适应判别并划分相应的数据处理线程。

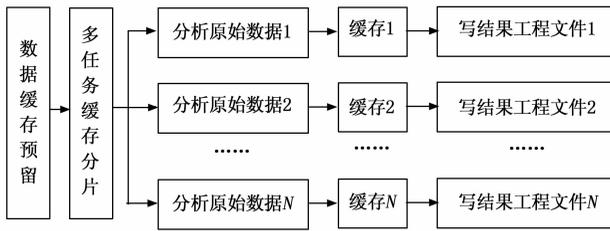


图 2 基于预留缓存的数据并行处理示意图

应用现阶段的数据存储技术，“结果工程量文件写 1”线程的执行必须被赋予首要的优先级，即该线程处理的文件必须是第一个被写入硬盘的数据。与后续的结果工程量数据存在严格的串行关系，每个线程必须要等到上一个线程结束才可以进行。此时，即使其余线程（如“结果工程量文件写 2”或者“结果工程量文件写 N”）先完成了数据分析，具备数据存储和写入的条件，也必须在“结果工程量文件写 1”线程执行结束后进行写操作。串行的模式极大的增加了线程的等待时间，致使整个进程的效率降低。

而采用预留缓存技术，数据存储的线程之间为，并行关系，没有必然的联系，各个线程的执行不需要等待其余线程的执行结束，消除了线程的等待时间，提高了整个数据处理效率。

多线程并行缓存实现如下：

1) 设定处理参数总数为 $ParaSum$ 、采样率为 $ParaCyl$ 、处理总时间为 $TimeSum$ ，数据处理信息字节数为 $ParaInfoSize$ ，则需创建的工程量文件的大小为：

$$ResFiSize = ParaSum \times TimeSum \times ParaCyl + ParaInfoSize$$

2) 进行内存映，创建结果工程量文件，其字节数为 $ResFiSize$ ；

3) 执行该次数据处理信息的数据存储写操作；

4) 数据进入处理线程池，进行相应的数据处理，将单

元时间内处理好的结果数据交给存储线程组；

5) 数据进入存储线程池，将结果数据写入工程量文件，实现数据的并行处理和快速存储。

单位时间内处理好的结果数据按指定地址存储。

设定起始时刻为 $TimeBeg$ ，起始地址为 $AddrBeg$ ，则数据在任意的 $TimeN$ 时刻的存储地址为：

$$TimeNaddr = AddrBeg + (TimeN - TimeBeg) \times ParaSum \times ParaCyl$$

当存储线程接收到单元时间的结果数据以及该数据的相关信息，按照 $TimeN$ 就可以知道结果数据的存放首地址，实现将该数据存储写入工程量文件。

3 航天装备在役考核多任务并行执行支撑技术验证

航天装备体系由一系列的卫星和相应配套的地面站构成，共同完成包括信息获取、信息传输和时空基准等支援保障任务^[19-20]。

规模庞大、组成结构复杂、多星多任务背景等复杂性是航天装备体系的重要特点，装备体系在役考核阶段难于快速、高效开展定性、定量评估。为了开展对航天装备的评估分析工作，必须要从体系的视角出发进行研究，需要对多星多任务进行合理且全面的剖析。

为实现航天装备体系在役考核的综合评估，自主研发了综合评估系统。根据航天装备在役考核的综合评估需求，系统涵盖了任务规划、试验设计、数据采集、分析评估和结果展示等多个模块，实现了多层次、多装备、多指标的综合评估。图 3 为系统的功能架构。

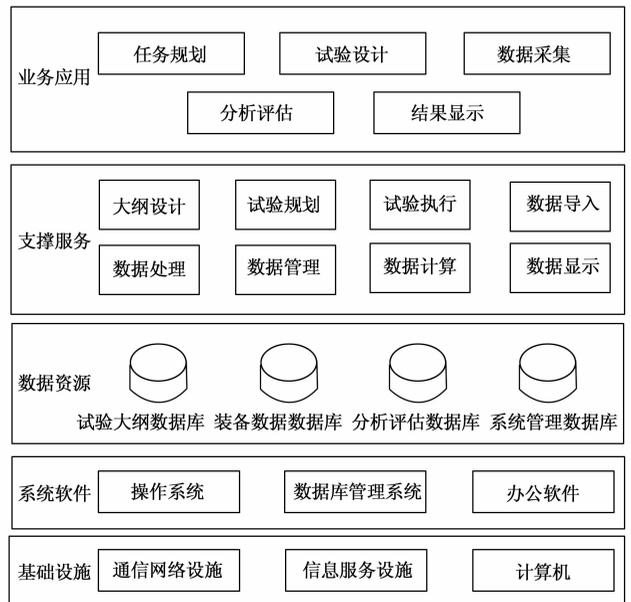


图 3 航天装备在役考核评估系统

通过科学合理的设计综合评估流程，划分试验科目和数据采集顺序，从任务层面实现了试验科目设计、试验执行、数据采集、特定指标计算的并行流程。图 4 航天装备在役考核评估系统任务流层面并行控制与执行软件界面示

意图。在整体流程的执行流程中, 各个环节均可以根据实际任务要求进行并行规划。任务规划可以对不受时序约束的不同任务进行分别规划, 试验设计(与执行)可以根据试验的实际执行过程进行不同试验科目的并行执行, 从而实现不同数据的并行采集和计算等过程。

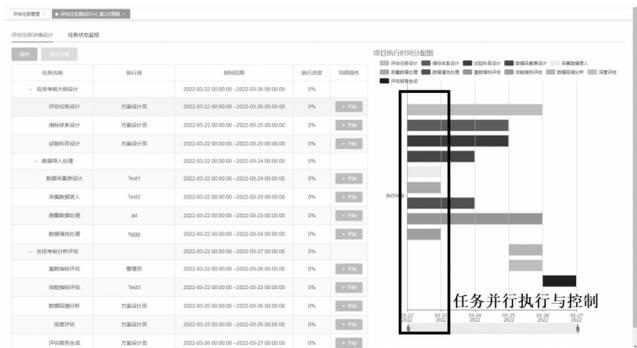


图 4 任务层面并行控制与执行软件界面示意图

其中, 数据采集和处理过程采用本文基于预留缓存的数据并行分析处理方法, 对采集的原始数据进行分组、分割、处理、存储和融合。对采集数据导入、测量数据处理和数据清洗变换等预处理实现高效、快速的执行, 从而在任务层面的并行执行之外, 实现数据底层计算逻辑和计算资源的并行执行。图 5 为航天装备在役考核评估系统数据流层面并行控制与执行软件界面示意图。



图 5 数据流层面并行控制与执行软件界面示意图

系统的应用和测试结果显示, 与传统的数据串行处理和存储相比, 数据的并行处理、存储和应用效率大大提高。

4 结束语

本文结合航天装备体系评估的实际需求, 针对装备类型多、执行任务多、数据量大、数据结构复杂等问题, 从任务执行层面和数据底层计算逻辑与资源分配层面等多层面分析了数据的串行处理分析存在的效率低的问题, 从而通过任务规划、科目划分、数据分片、预留缓存、合理规划多线程等多种方式实现了任务的并行执行和数据的并行处理, 提高了航天装备体系在役考核评估的处理效率。

参考文献:

[1] 美军参联会, 美空军, 美陆军. 太空作战 [M]. 李智, 侯

妍, 张占月, 译. 北京: 国防工业出版社, 2014.

[2] 赵辉, 严晓芳, 曹晨. 美军天基信息系统与网络中心战 [J]. 航天制造技术, 2010 (3): 11-16.

[3] 罗小明, 朱延雷, 何榕. 天基信息支援指挥控制体系运行机制及复杂网络建模 [J]. 指挥控制与仿真, 2016, 38 (5): 1-7.

[4] HE B, MAO Y, GUO Z, et al. Comet: batched stream processing for data intensive distributed computing [C] // Acm Symposium on Cloud Computing. ACM, 2010.

[5] LIU Y Q, et al. Towards Edge Intelligence: Multi-Access Edge Computing for 5G and Internet of Things [J]. IEEE Internet of Things Journal, 2020, 1.

[6] 施巍松, 孙辉, 曹杰, 等. 边缘计算: 万物互联时代新型计算模型 [J]. 计算机研究与发展, 2017, 54 (5): 907-924.

[7] 吕华章, 陈丹, 范斌, 等. 边缘计算标准化进展与案例分析 [J]. 计算机研究与发展, 2018, 55 (3): 487-511.

[8] MARCUCCIO S, ULLO S, CARMINATI M, et al. Smaller Satellites, Larger Constellations: Trends and Design Issues for Earth Observation Systems [J]. IEEE Aerospace and Electronic Systems Magazine, 2019, 34 (10): 50-59.

[9] 程思微, 张辉, 沈林成. 利用混合整数规划的卫星操作规划问题研究 [J]. 计算机工程与应用, 2011, 47 (3): 229-232.

[10] 陈英武, 方炎申, 李菊芳, 等. 卫星任务调度问题的约束规划模型 [J]. 国防科技大学学报, 2006 (5): 126-132.

[11] 张汉锋, 卢显, 陈立云. 天基信息支援下装备战场抢修有色 Petri 网模型研究 [J]. 指挥控制与仿真, 2011, 33 (1): 108-112.

[12] 魏启, 王克, 任文静, 等. 一种海洋地震节点数据切分方法 [C] // 中国地球物理学会地球物理技术委员会第九届学术会议—全域地球物理探测与智能感知学术研讨会会议摘要集, 2021: 211-215.

[13] 李阳, 王华, 张得华. 测控系统指显数据处理方法 [J]. 计算机系统应用, 2019, 28 (2): 270-273.

[14] 董献伦. 基于关系型数据库的数据切分问题研究 [D]. 济南: 山东大学, 2016.

[15] 周淦, 刘贵聪, 申展瑜, 等. 航天发射通用化指显数据处理方法及应用 [J]. 电子技术应用, 2022, 48 (11): 121-126.

[16] 李欣, 禹霁阳, 牛跃华, 等. 基于分级缓存加速的高可靠高速星载固件设计 [J]. 计算机测量与控制, 2019, 27 (10): 186-191.

[17] 刘兵, 郑怀洲, 肖东响. 装备指挥训练模拟数据收集系统研究 [J]. 系统仿真学报, 2006 (S2): 246-249.

[18] 周文刚, 赵宇, 朱海. 基于多线程的 AES-ECB 改进设计与性能分析 [J]. 计算机测量与控制, 2015, 23 (1): 207-210.

[19] 王兆耀. 军事航天技术 [M]. 北京: 中国大百科全书出版社, 2009.

[20] 孙盛智, 侯妍. 天基信息支援层次化保障模式及其应用 [J]. 火力与指挥控制, 2018, 43 (10): 99-104.