

发明专利申请

(初稿·内部评审用)

发明名称：一种多资源动态调度方法、系统及存储介质

申请人：(待填)

发明人：(待填)

编制日期：2026-02-14

摘要

发明名称

一种多资源动态调度方法、系统及存储介质。

摘要正文

本发明公开了一种多资源动态调度方法、系统及存储介质。该方法周期采集处理器、内存和图形处理器资源状态，采用原始快照与平滑快照双视图判定模式；在非紧急模式下执行同一调度周期累计资源投影准入，在紧急模式下执行阻断与定向抢占。所述定向抢占基于瓶颈识别、热点图形处理器识别和归一化回收评分，并在内存回收目标与图形处理器回收目标同时满足后停止。该方案可降低批量任务超发风险，提升紧急回收效率和调度过程可追溯性。

摘要附图

图1。

权 利 要 求 书

1. 一种面向计算设备的多资源动态调度方法，其特征在于，包括以下步骤：
 - S1：周期采集处理器资源、内存资源和图形处理器资源的使用状态，形成同一调度周期的原始资源快照；
 - S2：基于所述原始资源快照和对应平滑资源快照执行双视图模式判定，其中紧急模式判定仅基于原始资源快照，稳态模式判定基于平滑资源快照并结合迟滞退出机制和紧急冷却机制；
 - S3：在非紧急模式下执行同一调度周期内的候选任务累计资源投影准入判定，并在每次准入成功后更新后续候选任务投影基线，所述图形处理器预算至少包括按卡分桶预算和未绑定任务保守预算；
 - S4：在紧急模式下阻断新任务准入，识别瓶颈资源维度和热点图形处理器，并对可回收候选任务计算归一化资源回收评分；
 - S5：基于任务优先级、所述归一化资源回收评分、图形处理器亲和系数和任务启动时序执行定向抢占，并在内存回收目标与图形处理器回收目标同时满足后停止本轮抢占。
2. 根据权利要求1所述的方法，其特征在于，对声明目标图形处理器卡的任务仅在目标图形处理器卡执行显存投影，对未声明目标图形处理器卡的任务将显存需求以保守方式纳入各图形处理器卡风险评估。
3. 根据权利要求1所述的方法，其特征在于，在内存资源和图形处理器资源同时处于紧急状态时，所述归一化资源回收评分引入协同增强项。
4. 根据权利要求1所述的方法，其特征在于，所述图形处理器亲和系数满足同卡绑定任务大于未绑定任务，未绑定任务大于跨卡绑定任务。
5. 根据权利要求1所述的方法，其特征在于，所述方法包括阻断审计事件记录步骤：在任务被阻断时记录任务标识、阻断原因类别和阻断时间戳，并维护阻断事件总数与唯一阻断任务总数。
6. 根据权利要求1所述的方法，其特征在于，所述方法包括终止失败保护步骤：在向目标任务发出终止请求后，若所述目标任务在预设终止超时时间内未终止，则执行强制移除，并记录终止失败事件和超时移除事件。

7. 根据权利要求1所述的方法，其特征在于，所述方法包括自适应参数调优步骤：根据运行统计信号识别重试原因类别，并据此调整后续调度窗口的准入阈值偏置量、调度窗口时长和任务负载参数。
8. 一种多资源动态调度系统，其特征在于，包括资源监控模块、模式判定模块、准入控制模块、定向抢占模块、任务执行模块和事件审计模块；其中：
所述资源监控模块用于采集并输出原始资源快照；
所述模式判定模块用于执行双视图模式判定并输出当前模式；
所述准入控制模块用于在非紧急模式下执行同周期累计资源投影判定并维护按卡分桶预算和未绑定任务保守预算；
所述定向抢占模块用于在紧急模式下执行瓶颈识别、热点图形处理器识别、归一化评分排序和双目标停止控制；
所述任务执行模块用于执行任务启动、任务运行跟踪和任务终止控制；
所述事件审计模块用于记录模式切换、任务阻断、任务启动、任务停止和回收目标达成事件。
9. 根据权利要求8所述的系统，其特征在于，所述定向抢占模块包括归一化评分单元、图形处理器亲和加权单元和双目标停止单元。
10. 一种计算机可读存储介质，其上存储有计算机程序，其特征在于，所述计算机程序被处理器执行时实现权利要求1至7任一项所述的方法。

说 明 书

一、发明名称

一种多资源动态调度方法、系统及存储介质。

二、技术领域

本发明涉及计算机任务调度与系统资源治理技术领域。

更具体地说，本发明涉及一种面向处理器资源、内存资源和图形处理器资源的联合准入控制与紧急回收方法。

本发明同时涉及对应的模块化系统实现方案。

本发明还涉及用于执行所述方法的计算机可读存储介质。

本发明重点面向以下应用环境。

一是本地开发工作站的高并发任务执行环境。

二是边缘计算节点的多作业共存环境。

三是图形处理器密集型计算任务与常规内存密集型任务混合执行环境。

三、背景技术

2.1 背景问题概述

现有调度体系通常在单一资源维度上形成局部优化。

例如有的系统重点处理任务优先级，有的系统重点处理内存压力，有的系统重点处理图形处理器切片与隔离。

在本地多任务环境中，资源风险往往具有跨维度耦合特征。

内存压力与图形处理器显存压力可能同时出现。

单资源策略在高并发提交时容易出现响应滞后、误阻断或回收过度。

此外，批量任务在同一调度周期中连续提交时，如果只做单任务静态判定，可能出现前几个任务都“看似安全”但累计后越线的情况。

因此需要一种既可快速响应突发风险，又能在稳态阶段降低噪声干扰，同时具备跨资源联合约束能力的调度方案。

2.2 现有技术引证与不足分析

以下引证用于说明现有技术基础与不足。
每项引证均给出可核验来源。

2.2.1 引证一：美国专利申请 US20200167197A1

例如，美国专利申请 US20200167197A1 公开了一种通过终止工作负载释放资源并重新调度的方案。

可核验来源为：

<https://patents.google.com/patent/US20200167197A1/en>

该类方案在“过载检测-任务终止-资源释放”链条上具有参考意义。

然而，该方案主要关注抢占式终止逻辑。

该方案未明确给出同一调度周期内的累计准入投影机制。

该方案未明确给出按图形处理器卡分桶的显存投影预算机制。

该方案未明确给出紧急判定与稳态判定分离的双视图机制。

该方案也未公开“内存缺口归一化 + 图形处理器缺口归一化”的联合评分结构。

因此其不足在于缺少多资源联合准入与定向回收的闭环耦合设计。

2.2.2 引证二：Kubernetes Pod Priority and Preemption

又如，Kubernetes 的 Pod 优先级与抢占机制公开了基于优先级驱动的抢占调度能力。

可核验来源为：

<https://kubernetes.io/docs/concepts/scheduling-eviction/pod-priority-preemption/>

该机制可以在资源不足时通过高优先级任务替换低优先级任务。

然而，该机制重点是优先级与可调度性。

该机制未提供按图形处理器卡分桶的显存投影预算。

该机制未提供同一调度周期内的逐任务累计投影更新逻辑。

该机制未提供面向本地图形处理器亲和任务的未绑定保守预算策略。

因此其不足在于缺少细粒度设备级投影和批量同周期准入防超发能力。

2.2.3 引证三：Kubernetes Node-pressure Eviction

又如，Kubernetes 的节点压力驱逐机制给出了内存压力和磁盘压力触发驱逐的工程实现路径。

可核验来源为：

<https://kubernetes.io/docs/concepts/scheduling-eviction/node-pressure-eviction/>

该机制在节点级稳定性方面具有成熟实践意义。

然而，该机制偏向节点级驱逐策略。

该机制通常通过阈值和驱逐顺序控制稳定性。

该机制未提供“瓶颈维度识别 + 热点图形处理器识别 + 归一化联合评分”的定向抢占评分体系。

该机制未给出多资源回收目标同时达成的停止条件定义。

因此其不足在于紧急回收粒度和跨资源评分精度不足。

2.2.4 引证四：中国专利 CN111736987B

又如，中国专利 CN111736987B 涉及图形处理器资源共享与任务调度方向。

可核验来源为：

<https://patents.google.com/patent/CN111736987B/zh>

该类方案在图形处理器资源管理方面具有参考价值。

然而，该类方案主要集中在图形处理器资源分配或共享机制层面。

该类方案未覆盖处理器、内存和图形处理器三资源联合准入。

该类方案未覆盖紧急模式下的多资源归一化定向回收。

该类方案未公开同周期累计预算更新与批量防超发机制。

因此其不足在于系统级多资源联动能力不足。

2.2.5 引证五：Linux OOM Killer 机制

又如，Linux 内核的 OOM Killer 机制通过进程评分选择终止对象。

可核验来源为：

<https://docs.kernel.org/mm/oom.html>

该机制在内存极限条件下可快速回收资源，工程上非常成熟。

然而，OOM 机制以内存压力为主要触发与评分依据。

该机制并不以图形处理器显存瓶颈为核心决策对象。

该机制不提供按图形处理器亲和关系加权的回收评分机制。

该机制不涉及非紧急阶段的同周期累计准入投影控制。

因此其不足在于无法直接解决多资源联合准入与多目标回收的调度问题。

2.2.6 补充引证六：Slurm Preemption 文档

进一步地，Slurm 提供了可配置的抢占策略。

可核验来源为：

<https://slurm.schedmd.com/preempt.html>

Slurm 的策略化抢占为高性能计算调度提供重要经验。

但其公开资料重点在队列策略与调度规则配置。

其未公开本发明所述“双视图模式判定 + 同周期累计投影 + 按卡分桶预算 + 归一化联合评分”的不可分拆组合机制。

2.2.7 补充引证七：GPU 集群调度代表性论文

再如，Gandiva、Tiresias、AntMan、Pollux 分别代表了不同阶段的图形处理器集群调度研究路径。

可核验来源为：

Gandiva (OSDI 2018) : <https://www.usenix.org/conference/osdi18/presentation/xiao>

Tiresias (NSDI 2019) : <https://www.usenix.org/conference/nsdi19/presentation/gu>

AntMan (OSDI 2020) : <https://www.usenix.org/conference/osdi20/presentation/xiao>

Pollux (OSDI 2021) : <https://www.usenix.org/conference/osdi21/presentation/qiao>

上述工作在作业公平性、资源复用和集群级效率方面具有重要价值。

然而，所述工作主要针对集群级任务排队与资源分配优化。

其公开重点不在本地节点级同周期累计准入投影。

其公开重点也不在按图形处理器卡分桶并结合未绑定任务保守预算的准入判定。

其公开重点亦不在“紧急原始值判定 + 稳态平滑值判定”的双视图模式控制。

因此，现有文献虽提供重要背景，但不能直接替代本发明所述组合闭环。

2.3 现有技术综合不足

综合上述现有技术，可以归纳出以下不足。

第一，紧急风险判定与稳态判定缺乏视图分离，导致“要么响应慢，要么误报多”。

第二，批量任务在同周期提交时，缺乏累计预算更新，容易出现准入超发。

第三，图形处理器场景中缺乏按卡预算和绑定关系约束，跨卡误阻断与误放行并存。

第四，抢占回收常以单资源单指标排序，难以同时兼顾内存和图形处理器瓶颈。

第五，调度器参数通常固定，难以根据运行反馈进行结构化调优。

四、发明内容

3.1 要解决的技术问题

本发明至少解决以下技术问题。

问题一，如何在突发资源尖峰时实现低延迟紧急响应，同时在稳态阶段抑制噪声抖动。

问题二，如何在同一调度周期内避免批量任务造成累计超发。

问题三，如何在多图形处理器环境中降低跨卡误阻断，提升设备级准入精度。

问题四，如何在紧急状态下实现瓶颈导向的定向抢占，而非盲目回收。

问题五，如何在保证安全的同时维持可接受吞吐，并形成可审计证据链。

3.2 技术方案

为解决上述问题，本发明提出一种多资源动态调度方案。

该方案包括以下核心技术特征。

特征一，周期采样与双视图判定。

系统周期采集处理器、内存、交换分区和图形处理器状态。

紧急判定使用原始采样值，以捕获突发风险。

稳态判定使用平滑采样值，以降低抖动误判。

模式切换中引入迟滞退出和紧急冷却。

特征二，同周期累计准入投影。

在非紧急状态下，系统按任务优先级执行准入判定。

准入判定不仅参考当前运行基线，也参考本周期已准入任务的累计投影。

每次准入成功后立即更新后续候选任务投影基线。

该机制避免“逐个都安全，累计后越线”的批量超发问题。

特征三，按图形处理器卡分桶预算与未绑定保守预算。

对于绑定目标图形处理器卡的任务，仅在目标卡上投影显存占用。

对于未绑定任务，采用保守预算参与各卡风险评估。

该机制在安全约束下兼顾精度与鲁棒性。

特征四，紧急阻断与定向抢占。

系统在紧急状态下先阻断新任务准入。

随后识别瓶颈维度和热点图形处理器。

再基于归一化资源回收评分执行抢占排序。

评分至少包含内存回收贡献归一化项和图形处理器回收贡献归一化项。

在双瓶颈同时存在时，评分引入协同项，优先回收可同时缓解双瓶颈的任务。

回收停止条件要求达到内存回收目标与图形处理器回收目标。

特征五，参数调优与资源画像自校准。

系统可根据连续运行统计信号识别重试原因类别。

系统据此调整下一调度窗口阈值偏置和负载参数。

系统可根据任务观测峰值更新任务资源画像，用于修正后续资源估算。

3.3 有益效果

本发明至少具有以下有益效果。

效果一，紧急判定响应更快且稳态误报更低。

效果二，同周期累计投影可显著降低批量准入超发风险。

效果三，按卡预算与亲和约束可降低跨卡误阻断。

效果四，归一化联合评分可提高紧急回收效率，减少无关任务被回收概率。

效果五，参数调优可在连续运行中提升“安全与吞吐”双目标达成率。

效果六，事件审计链路完整，便于复核、追责与专利证据固化。

五、附图说明

5.1 附图简述

图1为本发明系统模块结构图。

图2为本发明调度周期流程图。

图3为真机三模式对比图。

图4为自适应调优轨迹图。

5.2 图面规范化说明

当前附图（图1至图4）为实验阶段产物，用于技术方案沟通与审查答复准备。

在正式向 CNIPA 提交前，应执行以下图面规范化动作：

1. 将图面全部标注统一为中文术语，并与权利要求和说明书术语保持一致。
2. 为图1和图2补齐模块、步骤和数据流的参考编号。
3. 将图3和图4中的坐标轴、图例、注释统一中文化。
4. 使图面标号与附图标记说明表逐项一致，避免“图内标号-文字说明”不一致。

六、具体实施方式

6.1 符号与术语说明

为便于说明，以下术语在本文中具有明确含义。

资源快照指某一时刻采样得到的多资源状态集合。

原始快照指未经平滑处理的直接采样结果。

平滑快照指基于指数移动平均处理后的结果。

调度周期指一次“采样-判定-准入或回收-记录”的完整执行窗口。

同周期累计投影指在同一调度周期内，已准入任务对后续任务准入判定基线的累积影响。

绑定任务指声明目标图形处理器卡的任务。

未绑定任务指未声明目标图形处理器卡的任务。

瓶颈维度指当前最需要回收的资源维度。

热点图形处理器指显存占用风险最高的图形处理器卡。

归一化回收评分指将不同资源维度的回收贡献按各自缺口进行无量纲化后得到的综合评分。

双目标停止条件指内存回收目标与图形处理器回收目标同时满足时停止回收。

重试原因类别指触发参数调优的运行反馈类别。

阈值偏置量指在基础阈值上施加的动态调节量。

6.2 系统总体结构与数据流闭环

6.2.1 模块组成

本发明系统优选包括以下模块。

资源监控模块。

模式判定模块。

准入控制模块。

定向抢占模块。

任务执行模块。

事件审计模块。

参数调优模块。

资源画像模块。

6.2.2 模块数据流

步骤一，资源监控模块输出原始资源快照至模式判定模块。

步骤二，模式判定模块输出当前模式至准入控制模块和定向抢占模块。

步骤三，在非紧急模式下，准入控制模块输出任务准入结果至任务执行模块。

步骤四，在紧急模式下，定向抢占模块输出抢占列表至任务执行模块。

步骤五，任务执行模块输出运行状态至事件审计模块和资源画像模块。

步骤六，事件审计模块输出统计信号至参数调优模块。

步骤七，参数调优模块输出下一窗口参数至模式判定模块、准入控制模块和定向抢占模块。

步骤八，资源画像模块输出资源估算修正值至准入控制模块。

上述步骤形成持续闭环。

6.2.3 模块边界说明

资源监控模块不负责决策，只负责数据采集。

模式判定模块不负责任务执行，只负责模式输出。

准入控制模块不负责终止运行任务，只负责非紧急准入。

定向抢占模块不负责创建任务，只负责紧急回收对象选择。

任务执行模块不直接计算策略，只执行启动与停止动作。

事件审计模块不修改调度策略，只记录事件并聚合统计。

参数调优模块不直接管理任务生命周期，只更新后续参数。

资源画像模块不直接做模式判定，只提供估算校准能力。

6.3 核心方法流程

7.1 流程总览

本发明方法流程可概括为以下阶段。

阶段一，采样阶段。

阶段二，模式判定阶段。

阶段三，非紧急准入阶段或紧急回收阶段。

阶段四，状态刷新与事件记录阶段。

阶段五，参数调优与画像更新阶段。

7.2 采样阶段

采样周期可配置。

每个周期采样处理器利用率。

每个周期采样内存占用比例、已用内存、可用内存。

每个周期采样交换分区占用比例。

在启用图形处理器防护时采样图形处理器利用率、显存占用比例、显存已用值、显存总量。

在多图形处理器环境中采样每张卡的独立显存状态。

采样结果统一形成原始资源快照。

7.3 模式判定阶段

模式优选包括普通模式、高压模式和紧急模式。

紧急模式判定优先使用原始快照。

紧急判定条件可包括以下至少一项。

内存占用比例达到紧急阈值。

交换分区占用比例达到紧急阈值。

可用内存低于保留内存阈值。

图形处理器显存占用比例达到紧急阈值。

稳态判定使用平滑快照。

高压模式判定条件可包括以下至少一项。

内存占用比例达到高压阈值。

处理器占用比例达到高压阈值。

图形处理器显存占用比例达到高压阈值。

模式回落时应用迟滞阈值。

紧急模式退出后应用冷却周期。

7.4 非紧急准入阶段

仅在普通模式和高压模式下执行准入。

普通模式的每周期启动预算可高于高压模式。

高压模式可限制每周期启动数量以降低风险。

对每个候选任务执行投影判定。

投影判定至少包含内存投影和处理器投影。

启用图形处理器防护时还包含显存投影。

在同一调度周期内对已准入任务进行累计投影。

每次准入成功后更新投影基线。

若投影结果触及紧急阈值则阻断该任务。

阻断信息记录到事件审计模块。

7.5 紧急回收阶段

在紧急模式下，系统先阻断新任务准入。

随后识别当前瓶颈维度。

瓶颈维度可为内存瓶颈、图形处理器瓶颈或双瓶颈。

在图形处理器瓶颈场景下识别热点图形处理器。

仅从可抢占任务集合中选择候选。

按优先级、归一化评分和时间序进行排序。

执行定向抢占并回收资源。

每次回收后更新回收进度。

当达到双目标停止条件时终止本轮回收。

7.6 状态刷新与事件记录阶段

周期末刷新运行任务状态。

记录任务完成、失败、超时和强制移除事件。

记录模式切换事件。

记录阻断事件；每条阻断事件至少包括任务标识、阻断原因类别和阻断时间戳。

统计阻断事件总数和唯一阻断任务总数，并写入当轮审计汇总。

记录启动和停止事件。

当向目标任务发出终止请求后，若在预设终止超时时间内未完成终止，则执行强制移除，并同时记录终止失败事件与超时移除事件。

记录参数调优事件。

记录资源画像更新事件。

事件日志支持容量上限控制。

日志超上限时按先进先出策略截断。

7.7 参数调优与画像更新阶段

参数调优模块读取聚合统计信号。

根据重试原因类别执行分支化调优。

吞吐不足时优选放宽阈值并延长窗口。

紧急信号不足时优选收紧阈值并增强压力参数。

所述调优参数至少包括准入阈值偏置量、调度窗口时长和任务负载参数。

所述任务负载参数至少包括任务数量、任务持续时长和任务基础内存占用量。

每轮调优结果写入审计轨迹。

资源画像模块根据已完成任务观测峰值更新画像。

当画像样本满足最小样本要求时，修正后续同类任务估算值。

6.4 关键判定条件与计算关系

8.1 平滑关系

平滑快照可采用指数移动平均关系。

平滑值可表示为：

$$S(t) = \alpha \cdot X(t) + (1-\alpha) \cdot S(t-1)$$

其中， $X(t)$ 表示原始采样值。

$S(t-1)$ 表示上一周期平滑值。

α 表示平滑系数。

α 取值范围为0至1。

α 越大，响应越快。

α 越小，平滑越强。

8.2 迟滞关系

高压退出阈值可由以下关系确定。

高压退出阈值 = 高压进入阈值 - 迟滞带宽。

迟滞带宽可大于等于0。

迟滞带宽用于抑制阈值附近抖动。

8.3 同周期累计投影关系

候选任务判定基线可表示为：

判定基线 = 当前运行基线 + 同周期累计投影量。

同周期累计投影量在每次准入成功后立即更新。

因此后续候选任务判定基线单调不减。

该关系保证批量提交场景下的风险可控。

8.4 归一化回收评分关系

在内存瓶颈场景下，内存归一化贡献可表示为：

内存归一化贡献 = 任务内存可回收量 / 内存回收缺口。

在图形处理器瓶颈场景下，图形处理器归一化贡献可表示为：

图形处理器归一化贡献 = 任务显存可回收量 / 显存回收缺口。

在双瓶颈场景下，可采用如下协同组合：

综合评分 = 内存归一化贡献 + 图形处理器归一化贡献 + 二者较小值。

该协同项用于提升双维度同时缓解任务的优先级。

8.5 双目标停止条件关系

回收停止条件至少满足以下关系。

累计内存回收量大于等于内存回收目标。

并且累计图形处理器回收量大于等于图形处理器回收目标。

若当前仅存在单瓶颈，则另一维回收目标可视为已满足。

6.5 实施方式详细描述

9.1 实施方式一：保守安全优先配置

本实施方式适用于内存风险较高的本地工作站。

优选参数关系如下。

内存高压阈值为80%至85%。

内存紧急阈值为85%至92%。

处理器高压阈值为75%至85%。

处理器紧急上限为90%至96%。

图形处理器高压阈值为80%至88%。

图形处理器紧急阈值为92%至97%。

平滑系数为0.5至0.7。

每周期抢占上限为1至2个任务。

调度窗口时长为12秒至30秒。

该配置目标是优先降低系统失稳概率。

在本实施方式中，紧急模式触发后先阻断新任务。

再按照瓶颈导向和归一化评分执行回收。

回收达到双目标后停止。

其优点是稳定性高。

其代价是峰值阶段吞吐可能下降。

9.2 实施方式二：平衡型配置

本实施方式适用于开发与训练混合负载场景。

优选参数关系如下。

内存高压阈值为70%至78%。

内存紧急阈值为80%至88%。

处理器高压阈值为70%至82%。

处理器紧急上限为88%至95%。

图形处理器高压阈值为78%至85%。

图形处理器紧急阈值为90%至95%。

平滑系数为0.4至0.6。

每周期抢占上限为1至3个任务。

调度窗口时长为18秒至60秒。

该配置目标是在安全与吞吐间求平衡。

在本实施方式中，普通模式允许较高每周期启动预算。

高压模式降低每周期启动预算。

紧急模式执行阻断与回收。

参数调优模块根据运行反馈动态调整阈值偏置和负载参数。

其优点是可在多轮运行中逐步靠近双目标。

9.3 实施方式三：图形处理器密集型配置

本实施方式适用于多图形处理器并发作业环境。

优选参数关系如下。

图形处理器高压阈值设置较低，以提前预警。

图形处理器紧急阈值保留安全边界。

内存阈值与处理器阈值保持中等水平。

平滑系数可取0.6至0.8以提高显存风险响应速度。

按卡分桶预算单元必须启用。

未绑定任务保守预算单元必须启用。

亲和权重机制必须启用。

双目标停止条件必须启用。

在该实施方式中，若热点图形处理器显存风险高，优先回收同卡绑定任务。

未绑定任务以中等权重参与评分。

跨卡绑定任务以较低权重参与评分。

该策略降低对非瓶颈卡任务的误伤。

9.4 实施方式四：调优分支策略实施例

本实施方式给出参数调优模块的分支行为。

当统计信号表明吞吐不足时：

调优模块提高准入阈值偏置量。

调优模块延长调度窗口时长。

调优模块保持或轻微下调负载压力参数。

该分支目标是提高完成任务数量。

当统计信号表明紧急信号不足时：

调优模块降低准入阈值偏置量。

调优模块提高任务数量参数。

调优模块提高任务基础内存占用参数。

调优模块适当延长任务持续时长参数。

该分支目标是提升风险触发可观测性并验证回收链路。

当统计信号表明目标满足时：

调优模块停止进一步调节。

保留本轮参数作为候选稳定参数集。

9.5 实施方式五：资源画像自校准实施例

本实施方式用于处理估算误差。

调度器在任务运行期间采样任务资源峰值。

采样值至少包括内存峰值和处理器峰值。

在启用图形处理器采样时还包括显存峰值。

任务结束后将观测峰值写入任务画像。

任务画像采用指数移动平均更新。

当样本数达到最小样本阈值后启用画像校准。

后续同类任务估算值取原估算与画像建议值中的较大值。

该方式可降低估算偏低导致的风险失控。

6.6 数值流程演算

10.1 场景设定

给定一台内存总量为8192MB的设备。
保留内存阈值设为512MB。
内存高压阈值设为85%。
内存紧急阈值设为92%。
图形处理器总显存设为8000MB。
图形处理器高压阈值设为85%。
图形处理器紧急阈值设为95%。
平滑系数设为0.6。
迟滞带宽设为3%。
冷却周期设为2个调度周期。

10.2 第一个调度周期

原始采样显示内存占用为80%。
原始采样显示处理器占用为62%。
原始采样显示图形处理器显存占用为70%。
紧急判定未触发。
稳态判定进入高压模式。
系统开始执行非紧急准入。
候选任务A预计内存占用增量为400MB。
准入后投影内存占用比例约为84.88%。
未触及内存紧急阈值。
任务A准入成功。
同周期累计投影内存增量更新为400MB。

10.3 同一周期后续准入

候选任务B预计内存占用增量为350MB。
若不做累计投影，仅按原始基线判定，可能错误认为可准入。
采用本发明累计投影后，判定基线包含任务A已占用增量。
投影内存占用比例上升至89.15%。
仍低于紧急阈值，但接近高压上沿。
任务B准入成功并更新累计投影。

候选任务C预计内存占用增量为500MB。
本轮判定基线已包含任务A和任务B累计增量。

投影内存占用比例达到95.25%。
超过内存紧急阈值92%。
任务C被阻断并记录阻断事件。
该演算体现同周期累计投影的防超发效果。

10.4 紧急周期演算

下一周期原始采样出现突发尖峰。
原始采样内存占用为93%。
原始采样图形处理器显存占用为96%。
即使平滑值尚未完全上升，紧急判定仍立即触发。
系统进入紧急模式并阻断新任务。
随后执行瓶颈识别。
识别结果为内存和图形处理器双瓶颈。
系统识别热点图形处理器为第1号卡。

10.5 回收评分演算

设任务D内存可回收量为1000MB。
设任务D图形处理器可回收量为1200MB。
设任务E内存可回收量为1300MB。
设任务E图形处理器可回收量为100MB。
设当前内存回收缺口为900MB。
设当前图形处理器回收缺口为1100MB。

则任务D内存归一化贡献约为1.11。
任务D图形处理器归一化贡献约为1.09。
任务D协同评分约为 $1.11 + 1.09 + 1.09 = 3.29$ 。

任务E内存归一化贡献约为1.44。
任务E图形处理器归一化贡献约为0.09。
任务E协同评分约为 $1.44 + 0.09 + 0.09 = 1.62$ 。

因此任务D优先于任务E被回收。
该结果说明协同评分可优先选择双瓶颈同时缓解任务。

10.6 回收停止演算

若回收任务D后累计内存回收达到1000MB。

且累计图形处理器回收达到1200MB。

二者均超过对应缺口。

系统满足双目标停止条件。

本轮回收停止。

避免进一步过度回收造成吞吐损失。

10.7 模式回落演算

紧急风险解除后系统不立即回到普通模式。

先进入冷却窗口。

在冷却窗口中维持紧急模式行为边界。

冷却结束后根据平滑快照与迟滞阈值判定回落。

若平滑值仍高于退出阈值，系统继续保持高压模式。

若平滑值低于退出阈值，系统回落至普通模式。

6.7 与现有技术对比

11.1 对比对象说明

为证明本发明技术效果，本节选择以下典型方案作为对比对象。

对比对象A：Kubernetes Pod 优先级与抢占机制。

对比对象B：Kubernetes 节点压力驱逐机制。

对比对象C：Slurm 抢占机制。

对比对象D：Linux OOM Killer 机制。

对比对象E：US20200167197A1 方案。

11.2 逐特征对比表

对比特征	本发明	对比对象A	对比对象B	对比对象C	对比对象D	对比对象E
紧急判定与稳态判定分离	有，双视图	无明确分离	无明确分离	无明确分离	无	无
同周期累计准入投影	有	无	无	无	无	无明确公开
按图形处理器卡分桶预算	有	无	无	依实现而异	无	无明确公开

未绑定任务保守预算	有	无	无	无	无	无
紧急模式先阻断后回收	有	部分	部分	部分	部分	部分
瓶颈维度识别	有	有限	有限	有限	以内存为主	有限
热点图形处理器识别	有	无	无	依插件	无	无明确公开
归一化双资源评分	有	无	无	无	以内存评分为主	无明确公开
双目标回收停止	有	无	无	无	无	无明确公开
调优分支控制	有	无	无	依策略脚本	无	无明确公开
任务画像自校准	有	无	无	依外部系统	无	无明确公开
逐任务审计闭环	有	有限	有限	有限	有限	有限

11.3 对比结论

与上述对比对象相比，本发明并非仅替换一个评分函数或一个阈值。本发明形成了由双视图判定、同周期累计投影、按卡分桶预算、瓶颈导向归一化回收组成的闭环组合方案。该组合在系统层面同时解决响应速度、批量超发、设备亲和和回收效率问题。因此，本发明具有明确的工程实用价值与组合创新价值。

6.8 附图详细描述

12.1 图1详细描述

图1中资源监控模块位于系统输入端。

该模块与模式判定模块之间通过资源快照数据流连接。

图1中模式判定模块向准入控制模块输出非紧急判定结果。

图1中模式判定模块向定向抢占模块输出紧急判定结果。

图1中准入控制模块将准入成功列表发送至任务执行模块。

图1中定向抢占模块将回收列表发送至任务执行模块。

图1中任务执行模块将状态变更发送至事件审计模块。

图1中事件审计模块将聚合统计发送至参数调优模块。

图1中参数调优模块将更新参数反馈至模式判定模块与准入控制模块。

图1中资源画像模块从任务执行模块读取观测峰值并向准入控制模块输出估算修正值。

12.2 图2详细描述

图2的第一步为采样。

图2的第二步为平滑更新。

图2的第三步为模式判定。

若判定紧急，图2进入阻断分支。

阻断分支先记录待执行队列逐任务阻断事件。

阻断分支再执行瓶颈识别和热点卡识别。

随后执行归一化评分排序与定向回收。

若判定非紧急，图2进入准入分支。

准入分支对候选任务逐个执行累计投影判定。

周期末统一执行状态刷新和日志记录。

12.3 图3详细描述

图3横轴可为实验时间或调度周期编号。

图3纵轴可包括资源占用比例和任务完成数。

图3中无调度曲线用于展示风险上限。

图3中固定并发曲线用于展示基线控制效果。

图3中动态调度曲线用于展示本发明组合机制效果。

图3中应可观测紧急阶段触发。

图3中应可观测抢占事件点。

图3中应可观测回落阶段吞吐恢复。

图3可同时展示安全与吞吐双目标。

图3支撑权利要求中的联合技术效果。

12.4 图4详细描述

图4以轮次为横轴。

图4以阈值偏置量、负载参数和结果指标为纵向观测对象。

图4展示每轮输入统计信号。

图4展示每轮重试原因类别。

图4展示每轮调优动作。

图4展示每轮调优后关键参数。

图4展示每轮安全与吞吐结果。

图4用于说明分支调优的可解释性。

图4用于说明调优终止条件。

图4用于支撑从属权利要求中的参数调优模块。

6.9 权利要求支撑关系

13.1 支撑原则

每条权利要求均需在说明书中清晰支撑。

支撑应包含步骤定义、判定条件、数据流或参数关系中的至少一项。

本节给出权利要求与说明书段落的映射关系。

13.2 支撑映射表

权利要求	支撑章节
方法独立权利要求	第七章、第八章、第九章、第十章
从属2(迟滞与冷却)	第7.3节、第8.2节、第10.7节
从属3(平滑系数)	第8.1节、第9.1节、第9.2节
从属4(同周期累计)	第7.4节、第8.3节、第10.3节
从属5(按卡分桶与未绑定保守)	第7.4节、第9.3节
从属6(目标卡不可用阻断)	第7.4节、第12.2节
从属7(亲和权重)	第7.5节、第9.3节
从属8(协同评分)	第8.4节、第10.5节
从属9(双目标停止)	第8.5节、第10.6节

从属10 (阻断统计)	第7.6节、第12.2节
从属11 (终止失败逃逸)	第7.6节、第9.4节
从属12 (参数调优)	第7.7节、第9.4节
从属13 (分支调优规则)	第9.4节、第10章
从属14 (画像自校准)	第7.7节、第9.5节
系统独立权利要求	第六章、第十二章
系统从属权利要求	第6.1节、第6.2节、第9.3节
介质独立权利要求	第七章、第九章

6.10 可选实现与变体

14.1 采样变体

采样源可来自操作系统资源接口。

也可来自容器运行时接口。

也可来自外部监控代理。

只要能形成统一资源快照即可。

14.2 判定变体

平滑方法除指数移动平均外，也可采用其他低延迟平滑方法。

但紧急判定优选保留原始采样路径。

迟滞和冷却可以独立配置。

普通模式与高压模式分界阈值可按场景调整。

14.3 准入变体

累计投影可支持内存与处理器之外的扩展资源。

图形处理器预算可按卡、按分区或按逻辑设备分桶。

未绑定任务保守预算可采用多种保守分配策略。

只要保持风险可控均属于本发明保护范围。

14.4 回收变体

归一化评分可引入任务业务权重。

归一化评分可引入历史失败代价项。

回收停止条件可扩展为多维目标。

但应保持至少双维瓶颈可解释。

14.5 调优变体

重试原因类别可扩展。

调优动作可扩展为离散动作集或连续调节函数。

调优目标可在安全优先与吞吐优先间切换。

只要调优过程可审计可追溯，均可属于本发明变体。

6.11 可实施性说明

本发明方法可由通用计算设备执行。

本发明不依赖特定品牌硬件。

当图形处理器不可用时，本发明可退化为处理器和内存双资源调度模式。

当图形处理器可用时，本发明自动启用按卡预算与热点卡识别能力。

本发明适用于单机、边缘节点和小规模集群的节点级调度控制场景。

本发明能够在现有软件栈上通过模块化方式部署。

6.12 结论

本发明提出了一种面向多资源约束场景的动态调度方案。

该方案以双视图判定保障紧急响应与稳态稳定。

该方案以同周期累计投影抑制批量准入超发。

该方案以按卡分桶预算和未绑定保守预算提升图形处理器场景精度。

该方案以归一化联合评分和双目标停止条件提升紧急回收效率。

该方案以参数调优和画像校准提升连续运行可用性。

该方案具备完整审计闭环，便于工程落地与专利审查支撑。

6.13 现有技术引证清单

1. US20200167197A1 , Preemptive termination of workloads to free resources and improve efficiency.

链接：<https://patents.google.com/patent/US20200167197A1/en>

2. Kubernetes Pod Priority and Preemption.

链接：<https://kubernetes.io/docs/concepts/scheduling-eviction/pod-priority-preemption/>

preemption/

3. Kubernetes Node-pressure Eviction。

链接：<https://kubernetes.io/docs/concepts/scheduling-eviction/node-pressure-eviction/>

4. CN111736987B。

链接：<https://patents.google.com/patent/CN111736987B/zh>

5. Linux Out Of Memory Handling (含 OOM Killer 机制说明)。

链接：<https://docs.kernel.org/mm/oom.html>

6. Slurm Preemption。

链接：<https://slurm.schedmd.com/preempt.html>

7. Gandiva (OSDI 2018)。

链接：<https://www.usenix.org/conference/osdi18/presentation/xiao>

8. Tiresias (NSDI 2019)。

链接：<https://www.usenix.org/conference/nsdi19/presentation/gu>

9. AntMan (OSDI 2020)。

链接：<https://www.usenix.org/conference/osdi20/presentation/xiao>

10. Pollux (OSDI 2021)。

链接：<https://www.usenix.org/conference/osdi21/presentation/qiao>

6.14 参数配置实施例清单

18.1 参数实施例A (安全优先)

参数项	实施值
内存高压阈值	85%
内存紧急阈值	92%
处理器高压阈值	80%
处理器紧急上限	95%
图形处理器高压阈值	85%

图形处理器紧急阈值	95%
平滑系数	0.6
迟滞带宽	3%
冷却周期	2
每周期抢占上限	1

18.2 参数实施例B (平衡优先)

参数项	实施值
内存高压阈值	72%
内存紧急阈值	82%
处理器高压阈值	78%
处理器紧急上限	95%
图形处理器高压阈值	82%
图形处理器紧急阈值	95%
平滑系数	0.5
迟滞带宽	3%
冷却周期	2
每周期抢占上限	2

18.3 参数实施例C (图形处理器优先)

参数项	实施值
内存高压阈值	75%
内存紧急阈值	88%
处理器高压阈值	75%

处理器紧急上限	93%
图形处理器高压阈值	80%
图形处理器紧急阈值	92%
平滑系数	0.7
迟滞带宽	2%
冷却周期	2
每周期抢占上限	3

6.15 审计字段建议

为保证实现可复核，建议审计日志至少包含以下字段。

事件时间戳。

事件类型。

任务标识。

阻断原因。

模式状态。

抢占原因。

回收进度。

调优动作。

重试原因类别。

关键阈值快照。

6.16 补充说明

本说明书中的“优选”“可选”“至少”不构成对保护范围的单一限定。

本说明书中的阈值范围为工程可行范围示例。

在不脱离权利要求实质精神的情况下，参数可根据设备能力和业务目标调整。

本说明书中的数值演算用于说明流程逻辑，不构成唯一实施条件。

附图说明

0. 图面状态说明（提交前必看）

1. 当前图1至图4属于实验阶段附图，图面文字和标号体系尚未完全提交态。
2. 在正式提交 CNIPA 前，需完成图面中文化、参考编号补齐以及图文一致性复核。
3. 图内编号必须与附图标记说明表一一对应，避免出现“图面标号缺失”或“图文标号不一致”。
4. 图3和图4如存在英文图例或英文坐标轴，需替换为中文并与说明书术语统一。

一、正式附图说明句（可直接放入说明书）

图1为本发明实施例中多资源动态调度系统模块结构示意图。

图2为本发明实施例中单个调度周期流程示意图。

图3为本发明实施例中三种调度模式真机对比示意图。

图4为本发明实施例中自适应调优轨迹示意图。

二、绘图文字描述（给绘图人员）

图1 系统模块结构图（建议横向闭环布局）

1. 画布中心标注“100 多资源动态调度系统”。
2. 左侧输入区域放置“110 资源监控模块”，下方连接“200 资源快照”，并分出“210 原始资源快照”和“220 平滑资源快照”。
3. 中上区域放置“120 模式判定模块”，由 210、220 双输入连接。
4. 中左区域放置“130 准入控制模块”，其输入来自模式判定模块和“300 待执行任务队列”。
5. 中右区域放置“140 定向抢占模块”，其输入来自模式判定模块和“500 抢占候选任务集合”。
6. 下中区域放置“150 任务执行模块”，接受准入控制模块与定向抢占模块输出。
7. 右侧放置“160 事件审计模块”，接收 120、130、140、150 的事件输出。

8. 下右侧放置“170 参数调优模块”，输入来自 160，输出反馈到 120、130、140。
9. 若版面允许，在底部放置“180 资源画像模块”，输入来自 150，输出反馈到 130。
10. 用闭环箭头体现“监控 -> 判定 -> 准入/回收 -> 执行 -> 审计 -> 调优 -> 判定”的反馈路径。

图2 调度周期流程图（建议纵向流程 + 分支）

1. 起点框：S101 采集资源状态并形成原始资源快照。
2. 第二框：S102 更新平滑资源快照并执行双视图模式判定。
3. 判定菱形：是否紧急模式。
4. 非紧急分支：
5. S103 同周期累计资源投影准入判定；
6. S104 准入成功后更新后续候选任务投影基线。
7. 紧急分支：
8. S105 阻断新任务准入；
9. S106 识别瓶颈资源维度与热点图形处理器；
10. S107 计算归一化回收评分并执行定向抢占；
11. S108 判断双目标停止条件是否满足。
12. 汇合框：S109 刷新任务状态并记录审计事件。
13. 结束框：S110 输出本周期调度结果并进入下一周期。
14. 在 S103 附注“按卡分桶预算 + 未绑定保守预算”。
15. 在 S107 附注“可启用协同增强项与亲和加权”。
16. 在 S108 附注“内存目标与图形处理器目标同时满足才停止”。

图3 真机三模式对比图（建议柱状+折线组合）

1. 横轴设置三组模式：无调度、固定并发、动态调度。
2. 柱状指标至少包含：内存峰值占比、图形处理器显存峰值占比。
3. 折线指标至少包含：任务完成数、阻断次数。
4. 对动态调度组加注释“触发紧急模式并执行定向回收”。

5. 对无调度组加注释“高峰风险基线”。
6. 对固定并发组加注释“无资源感知回收”。
7. 标注实验环境说明框（可放右上角）。
8. 图例需清楚区分内存、显存、完成数、阻断数。
9. 数据来源指向 `figures/` 中对应 CSV/JSON。
10. 图注建议说明“本图用于展示安全与吞吐的平衡效果”。

图4 自适应调优轨迹图（建议阶段流程图）

1. 横向按轮次画出 Attempt-1、Attempt-2、Attempt-3（如有）。
2. 每轮分四个节点：输入信号、重试原因、调优动作、输出结果。
3. 输入信号框示例：紧急触发次数、任务完成数、阻断次数。
4. 重试原因框示例：紧急信号不足、吞吐不足。
5. 调优动作框示例：阈值收紧、阈值放宽、负载参数调整。
6. 输出结果框示例：已启动任务数、已完成任务数、抢占次数。
7. 使用箭头表示“按失败原因分支调节”的路径。
8. 在图注中强调“调优是反馈闭环，不是固定脚本”。
9. 与图1中的 170 参数调优模块保持术语一致。
10. 图注建议说明“用于支撑参数自适应实施方式”。

附图标记说明表

说明：本表按“当前 SVG 实际文本”与“建议 CNIPA 参考标号”双列给出，便于提交前重绘统一。

当前 SVG 实际文本	建议 CNIPA 参考标号	建议名称
资源调度系统模块图	100	多资源动态调度系统
资源监控模块	110	资源监控模块
模式判定模块	120	模式判定模块
准入控制模块	130	准入控制模块
抢占回收模块	140	定向抢占模块
任务执行模块	150	任务执行模块
审计日志模块	160	事件审计模块
开始一个 tick	S101	周期开始步骤
采集 raw snapshot 并更新 EMA	S102	采样与平滑步骤
是否 EMERGENCY?	S103	模式判定步骤
阻断准入 + 定向抢占	S104	紧急分支步骤
逐个尝试准入并累计预算	S105	非紧急准入步骤
输出事件与指标并结束 tick	S106	周期结束步骤
A/B/C (图3)	G3-A/B/C	三模式实验组标记
A1-A5 (图4)	G4-A1...A5	调优轮次标记

备注

1. 图3和图4为实验图，当前主要使用实验组标记而非模块标号。

2. 正式提交前建议统一中文图题并在图注中声明实验组含义。

附 图

附图 1：系统模块结构

(见附件 SVG 文件：附图_资源调度_图1_系统模块.svg)

附图 2：调度周期流程

(见附件 SVG 文件：附图_资源调度_图2_调度周期流程.svg)

附图 3：三模式对比

(见附件 SVG 文件：附图_资源调度_图3_真机三模式对比.svg)

附图 4：自适应调优轨迹

(见附件 SVG 文件：附图_资源调度_图4_自适应重试轨迹.svg)