

浙 江 大 学

研 究 生 课 程 论 文



题目： 离线混部技术与数据中心能耗优化：
为元宇宙的可持续发展提供支持

学号： **22451040**

姓名： 刘京宗

专业： 软件工程

所在学院： 软件学院

课程名称： 工程前沿技术研究讲座

混部技术与数据中心能耗优化：为元宇宙的可持续发展提供支持

摘要

在本文中，深入探讨了离线混合部署技术及数据中心能耗优化对元宇宙可持续发展的重要性。随着元宇宙概念的逐渐普及，其对计算资源的需求预计将以指数级增长，这对数据中心的设计和管理提出了新挑战。数据中心不仅面临快速扩张的问题，还需有效应对能耗优化，以实现绿色可持续发展。

本文指出，现有数据中心的资源利用率普遍较低，导致大量能源浪费，增加了企业运营成本和环境负担。为应对这些挑战，各国和企业开始推行绿色计算，探索提升能源效率的方法。在中国，国家提出的“东数西算”战略旨在优化数据中心的布局和资源分配，以减少电力消耗和能耗压力。此外，绿色云数据中心的理念逐渐受到重视，通过可再生能源来降低运营成本，提升能效。

离线混合部署技术通过结合在线任务和离线作业，能够有效提升资源的整体利用率和降低运营成本。在线服务通常要求高可用性和低延迟，而离线任务则具有较高的容错能力和资源利用率。这种混合部署的方式可以在高峰期平衡资源使用，减少闲置，从而提升数据中心的经济效益和环境友好性。然而，技术应用中依然存在许多挑战，例如性能干扰、资源隔离和动态资源分配等问题，这些问题可能影响服务质量并阻碍元宇宙的可持续发展。

在回顾国内外在离线混合部署及数据中心能耗优化方面的研究现状时，本文提及了多种业界解决方案，如美团的 LAR 系统、阿里巴巴的 Fuxi 和 Google 的 Borg 系统，这些系统通过创新的资源管理和调度策略，成功提升了资源利用率并降低了能耗。

最后，本文展望了未来研究的方向，强调了几个重要方面：建立更精确的性能干扰模型、设计基于微服务的调度算法、探索绿色能源的整合与应用，以及发展智能化资源管理系统。此外，政策与标准的制定将对数据中心能耗优化产生深远影响，推动绿色数据中心的建设，为可持续发展提供制度保障。

关键词：离线混合部署，数据中心，能耗优化，元宇宙，绿色计算

一、研究背景

1.1 研究背景

随着元宇宙的概念逐渐渗透到各行各业，数据中心作为支撑其发展的核心基础设施，其重要性愈发凸显。元宇宙的实现依赖于虚拟现实、增强现实以及大规模在线交互，这些技术均需要强大的计算能力和高效的数据处理能力。根据相关研究，预计未来数年内，全球对计算资源的需求将以指数级增长，这对数据中心的设计和管理提出了新的挑战^[1]。

在这一背景下，数据中心不仅面临数量和规模的快速扩张，还需解决能耗优化的问题。现有的数据中心在资源利用方面存在普遍低效现象，许多服务器在负载不满时仍需运行，导致能源浪费。这种情况不仅增加了企业运营成本，也加重了环境负担，导致碳排放显著上升。为了实现可持续发展，许多国家和企业开始倡导绿色计算，力求通过提高能源利用效率，减少碳足迹，来应对全球变暖的挑战。

在中国，国家提出“东数西算”战略，以促进数据中心的合理布局和高效利用。通过将计算资源从东部发达地区向西部地区转移，不仅可以优化能源资源的配置，还可以减少电力消耗和能耗压力。此外，绿色云数据中心的理念逐渐受到重视，采用风能、太阳能等可再生能源，为数据中心的运营提供清洁能源解决方案，成为提升能效的重要途径^[2]。

在此背景下，离线混部技术的研究应运而生。这一技术通过将在线任务与离线作业有效结合，旨在提高资源的整体利用率和降低运营成本。在线服务通常需要高可用性和低延迟，而离线作业则具备较高的容错能力和资源利用率。将这两者混合部署，可以在高峰期平衡资源使用，有效减少闲置，从而提升数据中心的经济效益和环境友好性^[3]。

然而，尽管离线混部技术的前景广阔，实际应用中仍面临诸多挑战，例如性能干扰、资源隔离、动态资源分配等问题。这些问题不仅影响了系统的整体性能，也为实现元宇宙的可持续发展设置了障碍。因此，深入研究数据中心能耗优化和离线混部技术，探索其在元宇宙背景下的应用，将为推动技术进步和行业发展提供重要支持。

本文的组织结构如下：

第一节介绍了在离线混部和数据中心能耗优化的背景和问题挑战。

第二节介绍了在离线混部和数据中心能耗优化的研究现状。

第三节介绍了业界为提高系统资源利用率，提出的各种方案。

第四节为文章的总结内容。

1.2 问题与挑战

随着世界云计算的繁荣，数据中心无论在数量上，还是在规模上，均获得空前发展，而且它所带来的大量能耗已成为工业界亟待解决的难题，已有数据中心资源利用率普遍较低，导致能耗浪费较大，它关系到整个社会所面临的能源枯竭的问题、环境污染问题现状反差明显；数据中心会造成很大的电费开支和能源消耗，而一些企业难以支撑如此大的开支进而限制了数据中心的发展。

许多企业达到一定的规模后，实际上，在线机器的发展是好的，因为线上和生意基本上是成比例的，是 QPS，就拿现在的 100 来说吧，来年想做 200，然后我把机器加进去；同时，由于每年机器都比前一年好，因此，在预算的视角下，公司认为这很有道理，生意增长了 20%，你们的机器增长如 15%，这肯定是可以的，起码不需要太多的钱，也是可以接受的。

主要总结有如下几大问题与挑战：数据中心的能耗优化问题、混部作业性能干扰问题、混部作业调度问题、作业的共享资源隔离问题和资源动态分配问题。

二、研究现状

2.1 在离线混部概念

一般来说，企业的 IT 环境通常运行两大类进程，一类是在线服务，一类是离线作业。

在线任务就是指需要长时间运行在服务器上，服务流量和资源利用率具有潮汐特征和时延敏感的特点，对于服务 SLA 的需求极高，例如消息流的 Feed 服务、电商交易服务、互联网公司提供的网站服务、类似 QQ 微信的通信软件和广大论坛等等。

而离线任务就是指一次性任务，或者运行时间分区间，在运行过程中资源利用率高，时延不灵敏，容错率较高，中断通常使重运行成为可能，例如人工智能中模型训练任务，以及 Hadoop 生态中 MapReduce、Spark 操作等。这些工作分布在天文信息处理，互联网服务，数据挖掘，等多个领域^[4]。

所以在离线混部技术就是把在线任务和离线任务混合在一起部署。在线服务资源利用率有更明显的起伏特征，所以混部的主要场景是通过填充离线作业把在线服务各个时段的空闲资源利用起来，减少企业与日俱增的成本开支。

2.2 在离线混部国内外现状

在离线混部一直是学术界和工业界为提高数据中心的系统资源利用率在不断发展的一种部署思路。

混部目前的问题在于性能干扰，性能干扰由多作业竞争 CPU、内存等资源而产生。而在离线混部场景下的性能干扰对集群的调度和管理会带来很大的影响。

另一个问题是资源隔离问题，目前主流方案还是采用 Linux 系统的 CGroup（control groups 的缩写），这是一种 Linux 内核功能，用于限制、说明和隔离一组进程的资源使用。

此外是资源动态分配问题。现有的操作系统和虚拟化技术提供了一系列的资源隔离技术，可以在作业运行时动态调整该作业的可用资源。例如 Linux 提供的 CGroup 提供了包括 CPU、内存、网络、磁盘等资源的隔离，可以在作业运行时动态调整该作业的可用资源。三级缓存方面，Intel 提供了 Intel Resource Director Technology (RDT)，可以为作业分配私有的 L3 缓存。以现有的资源隔离技术为基础，研究人员研究了资源动态分配算法，在运行时动态调整各个作业对于共享资源的使用量，提升作业的运行效率^[5]。

2.3 数据中心能耗优化国内外现状

近年来，各大企业一直在设法降低数据中心集群的能耗，如利用更好的能耗均增计算技术，更优秀的调度算法系统等等^[6]。由于大规模的二氧化碳（CO₂）排放，在电力生产过程中，ICT 设施间接造成大量温室气体排放。这些人口密集的数据中心产生的热量需要大型冷却装置才能将温度保持在运行范围内。显然，这些冷却装置增加了总能源消耗并有自己的碳足迹^[7]。

KEIDS 论文^[3]的作者考虑了由边缘节点和云节点组成的 Kubernetes 集群，这些节点可能具有绿色清洁能源，例如水能或太阳能。该论文考虑了 Pod 中使用的容器类型（例如 CPU 密集型或网络密集型）以及与作业截止日期、所需资源，将调度问题表述为一个多目标优化问题，其目标是 최소화能源消耗和减少干扰。该论文将整体能耗降低了 14%，此外，集群性能干扰和数据中心碳足迹水平分别降低了 47% 和 31%^[8]。

Heats 提出^[9]针对具有 GPU、FPGA、x86 以外的架构（如：ARM、IBM Power8）或具有某些特定拓展（例如英特尔 SGX）的异构云实例，使用机器学习的方式，根据节点性能与能耗特征，对节点建立相应模型，从而将应用部署到所需能耗最低的节点，经过评估，Heats 最多可以节省 8.5% 的能源。

KCSS^[10]引入了节点的多标准选择，考虑的标准包括节点上 CPU、内存和磁盘的利用率、节点的功耗、运行容器的数量以及传输容器镜像所需的时间，使用了优先级排序（TOPSIS）算法寻找理想解，相较于其他容器调度算法，在不同场景下性能均有提高，功耗均有降低。

三、业界解决方案

3.1 美团 LAR

LAR（Load Auto-Regulator）系统，旨在保证服务质量的同时提高系统的资源利用率。

LAR 提出了两个核心创新：资源池化升级，动态负载和静态资源映射。架构分为了资源分级管理模块、资源池配置管理模块、服务质量保障模块、资源隔离管理模块和策略配置模块通过分级池化资源模型。

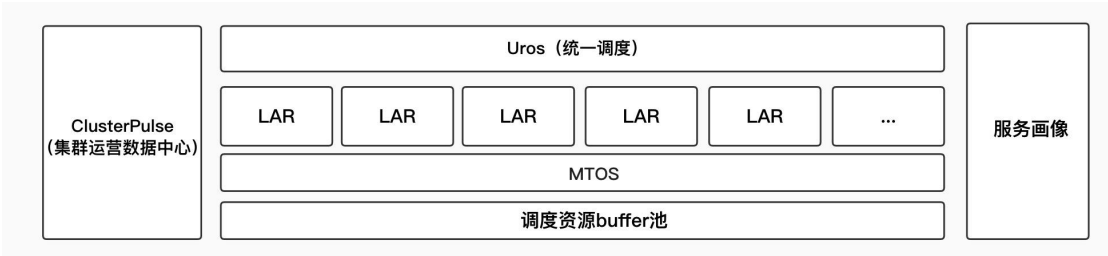


图 1 LAR 架构图

混部主要就是通过将延时和稳定性容错性更高的离线服务和在线服务混合部署，实现在线服务和离线服务在资源使用时空上的削峰填谷。

3.2 阿里 Fuxi

阿里巴巴的 Fuxi 支持数据分析工作负载；它从 2009 年开始运行。与 Borgmaster 一样，中央 FuxiMaster（为容错而复制）从节点收集资源可用性信息，接受来自应用程序的请求，并相互匹配。Fuxi 增量调度策略与 Borg 的等价类相反：Fuxi 不是将每个任务与一组合适的机器匹配，而是将新可用的资源

与待处理的积压工作进行匹配。同时，Fuxi 允许定义“虚拟资源”类型。能达到的 CPU 利用率为 38%。

3.3 Google Borg

来自 Google 的 Borg 系统是一个集群管理器。它结合了准入控制，高效任务打包，机器共享等技术，在进程级进行性能隔离，从而达到高利用率^[11]。

它支持具有可最大限度减少故障恢复时间的运行时特性和可降低相关故障概率的调度策略的高可用性应用程序。Borg 通过提供声明性作业规范语言、名称服务集成、实时作业监控以及分析和模拟系统行为的工具来简化用户使用。采用的资源隔离技术有容器技术、CAT 以及 CPU 强制调度技术^[12]。

四、总结与展望

在云计算与元宇宙交汇的时代，数据中心的能耗优化与离线混部技术的发展变得尤为重要。本文探讨了在元宇宙背景下，如何通过有效的资源管理和能耗优化来支持大规模计算需求。研究表明，提升资源利用率不仅可以降低运营成本，还能显著减少环境影响，为实现可持续发展目标奠定基础。

总结来看，当前数据中心在能耗管理和资源调度方面仍面临诸多挑战。离线混部技术虽然能够提高资源利用率，但其在性能干扰、资源隔离及动态资源分配等问题上亟待解决。这些问题的存在可能导致在线和离线任务之间的干扰，从而影响服务质量。因此，针对这些技术难题的深入研究，显得至关重要。

未来的研究方向值得关注，主要包括以下几个方面^[13]：

性能干扰模型研究：针对离线混部技术中多作业竞争资源造成的性能干扰，需建立更为精确的模型，以量化和预测不同作业之间的相互影响。这将有助于优化调度算法，提升系统的整体性能。

基于微服务的作业调度算法：随着微服务架构的广泛应用，研究如何在离线混部中有效调度微服务任务，将是一个重要课题。设计能够适应动态变化的调度策略，将提高资源利用率和系统的灵活性。

绿色能源的整合与应用：探索如何将绿色能源（如风能、太阳能）与数据中心运营相结合，以实现能耗的持续优化。这不仅能够减少碳排放，还可以降低对传统能源的依赖，推动行业的可持续发展。

智能化资源管理系统：随着人工智能技术的发展，利用 AI 算法进行数据中心资源管理将成为一个重要趋势。通过机器学习等智能技术，动态优化资源配置，提高数据中心的运营效率。

政策与标准制定：国家政策和行业标准的制定将对数据中心的能耗优化产生深远影响。推动政策引导和行业标准的建立，有助于促进绿色数据中心的建设，为可持续发展提供制度保障。

综上所述，离线混部技术和数据中心能耗优化在元宇宙的可持续发展中具有重要意义。通过持续的技术创新和政策支持，未来的数据中心将更加高效、环保，为数字经济的高质量发展提供坚实基础。

参考文献

- [1] 王康瑾, 贾统, 李影. 在离线混部作业调度与资源管理技术研究综述[J/OL]. 软件学报, 2020, 31(10): 3100-3119.
- [2] 周人杰. 实施“东数西算”工程 打造算力一张网[J]. 经营管理者, 2022(03): 1.
- [3] KAUR K, GARG S, KADDOUM G, 等. KEIDS: Kubernetes-Based Energy and Interference Driven Scheduler for Industrial IoT in Edge-Cloud Ecosystem[J/OL]. IEEE Internet of Things Journal, 2020, 7(5): 4228-4237. DOI:10.1109/JIOT.2019.2939534.
- [4] 乔佳彬. 面向混部云的性能评测方法的研究[D/OL]. 大连理工大学, 2018[2022-10-12].
- [5] HUMMAIDA A R, PATON N W, SAKELLARIOU R. Adaptation in cloud resource configuration: a survey[J/OL]. Journal of Cloud Computing, 2016, 5(1): 7. DOI:10.1186/s13677-016-0057-9.
- [6] 邓维, 刘方明, 金海, 等. 云计算数据中心的新能源应用:研究现状与趋势[J]. 计算机学报, 2013, 36(03): 582-598.
- [7] SHUJA J, MADANI S A, BILAL K, et al. Energy-efficient data centers[J/OL]. Computing, 2012, 94(12): 973-994. DOI:10.1007/s00607-012-0211-2.
- [8] MITTAL S, VETTER J S. A Survey of Methods for Analyzing and Improving GPU Energy Efficiency[J/OL]. ACM Computing Surveys, 2015, 47(2): 1-23.
- [9] ROCHA I, GÖTTEL C, FELBER P, et al. Heats: Heterogeneity-and Energy-Aware Task-Based Scheduling[C/OL]//2019 27th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP). 2019: 400-405. DOI:10.1109/EMPDP.2019.8671554.
- [10] MENOUEUR T. KCSS: Kubernetes container scheduling strategy[J/OL]. The Journal of Supercomputing, 2021, 77(5): 4267-4293.
- [11] TOKA L, DOBREFF G, FODOR B, et al. Adaptive AI-based auto-scaling for Kubernetes[C/OL]//2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID). 2020: 599-608.
- [12] VERMA A, PEDROSA L, KORUPOLU M, et al. Large-scale cluster management at Google with Borg[C/OL]//Proceedings of the Tenth European Conference on Computer Systems. New York, NY, USA: Association for Computing Machinery, 2015: 1-17[2022-11-19].