摘要

（目的）在全球经济受到影响的背景下，“降本增效”的意义价值凸显，数据中心的资源浪费问题越来越被关注。混部是目前业界公认缓解资源浪费的解决方案。混部的概念由 Google 于 2015 提出，并已经发展多年，国内大型互联网公司都早已开始采用混部提高集群资源利用率。

（方法）混部提升资源利用率的前提是保证应用服务质量，目前国内多数企业采用 Kubernetes 作为混部的编排调度框架，然而 Kubernetes 调度关注的主要资源是 CPU 与内存，但是对应用的实际资源负载需求没有感知。随之引起的问题包括：(1)集群跨节点资源负载不均衡。(2)部分对应用性能影响较大的共享资源竞争将被忽略，如内存子系统的资源竞争。这些问题限制了集群利用率的进一步提升，若调度团队不解决以上问题，而继续提升资源利用率，应用存在被调度到资源负载过高的节点而服务质量严重下降的风险。

（方法）为了填补 Kubernetes 框架的缺陷，本文在所处的基于 Kubernetes 调度平台之上，设计与实现了一套解决方案，增强集群调度能力。本文的主要工作：(1)硬件指标采集工具，该工具补全了容器的内存子系统压力相关指标与 cycles-per-instructions(CPI)的监控，前者丰富调度系统的资源感知，后者用于集群数据分析。(2)通用的内存子系统压力定义，包括 last level cache(LLC)与内存控制器资源的压力定义。(3)应用资源画像与基于画像的调度策略，本文采用通用且易生产落地的方法构建了应用资源画像，并且基于画像设计了基于高斯估计的调度优化策略，（部分结果）为集群跨节点资源负载(cpu/cache/memory controller)的均衡度提升最高达(38%,28%,32%)。(4)调度仿真工具，实现对 Kubernetes 组件的端到端的调度验证。

（结论）本文的部分实现已投入某互联网公司生产，并且本文的实现都充分考虑通用性，经过本文的实践验证后，将尝试贡献到开源社区，分享本文的解决方案。

关键词： 内存子系统压力，资源感知，集群调度，资源均衡性