# 目录

目录	1
全局管理简介	2
资源管理	2
用户管理	2
活动日志	2
个人中心	2
用户管理	2
用户列表	2
查看配额	2
设置算力型号	3
算力分配	4
设置权限	4
用户资源监控	5
算力消耗周报表	5
加速卡使用信息	5
CPU使用信息	5
内存使用信息	6

用户指南 2021-12-14

## 全局管理简介

在KingAI全局管理模块中,支持管理员对KingAI平台资源进行管理与监控,同时支持对子账号进行算力的分配与管理,全局管理一共分为资源管理、用户管理、活动日志、个人中心、镜像管理、AI商店六个功能模块。

#### 资源管理

在KingAI资源管理模块中,支持主账号对所有子账号消耗的资源进行监控与管理,包含算力监控、集群监控和集群节点管理三部分。仅有主账可以显示该功能模块。

- 1. 算力监控:支持监控平台资源的使用情况,监控指标如下:
  - · 集群算力: 以节点为监控粒度进行监控,显示当前空闲资源和总资源(可调度部分)。
  - 用户容器使用情况:支持查看用户的容器使用情况,显示信息包含用户名称、已使用算力时、CPU资源使用率、内存资源使用率、GPU资源使用率和资源闲置率等,支持按时间进行筛选,查看 历史的使用情况
  - 。 集群资源使用率:集群资源使用率为集群CPU资源使用情况。
  - 系统事件:显示任务的事件信息,包含事件名称、时间等字段。
  - 。 系统资源占用分析:显示各模块占用系统资源的占比情况,支持按不同CPU、GPU计算资源进行筛选展示。
  - 。用户资源使用排名:显示主账号下所有子账号的资源使用排名情况,支持按时间、不同CPU、GPU计算资源维度进行筛选展示。
- 2. 集群节点管理: 支持用户对集群节点进行管理, 支持对不同节点进行购买、启用、停用、删除等操作。
- 3. 资源监控: 支持监控平台集群的资源使用情况,监控集群、服务器、容器等多多维度,监控多项指标的动态变化。如下为一些示例指标及介绍:
  - 。 CPU使用情况:该监控项展示了集群、各节点、各Pod的CPU使用率百分比,当CPU使用率比较高时会影响KingAI服务。
  - 。 CPU平均负载: 平均负载是指固定分钟同时处于就绪状态的平均进程数,监控维度分为1分钟、5分钟和10分钟三个维度。
  - o GPU使用情况:该监控项展示了集群、各节点、各Pod的GPU使用率百分比。
  - 。 内存使用情况:监控项展示了集群、各节点、各Pod的HeapMemory使用率百分比。
  - 。 磁盘使用量:该监控项展示了集群、各节点、各Pod的磁盘使用率百分比。
  - 。 磁盘吞吐: 该监控项展示了集群、各节点、各Pod在一秒内,磁盘进行多少次 I/0 读写。
  - 。 IOPS: 指单位时间内系统能处理的I/O请求数量,分为读、写两个维度。
  - 。 网络接收/传输: 指网络接收和传输的速度; 分为接受、传输两个维度。

#### 用户管理

在KingAI用户管理模块中,支持管理员将平台中的算力资源进行统一管理分配,给用户进行统一算力上限管理,算力分配等,提高算力资源利用效率。仅有主账可以显示该功能模块。设置算力型号包

- CPU使用上限,单位核心。
- 内存使用上限,单位GB
- GPU型号(包含P40、tesla t4、titan V等多种GPU型号)和使用上限。

注意:子账号初次进入KingAI平台,算力资源上限和算力时均为0,需要管理员对子账号进行设置算力型号和算力分配的操作。

## 活动日志

在KingAI活动日志模块中,支持显示用户任务的调用事件(包含新建、暂停、编辑、删除等),详情如下:

- 显示事件名称、事件开始类型、事件结束事件、状态等基本信息。
- 支持根据时间跨度、事件类型进行筛选。

### 个人中心

在KingAI个人中心模块中,支持显示个人中心的算力信息,详情如下:

- 支持显示用户算力信息,包含当前可用算力时、累计获取算力时、累计消耗算力时等。
- 支持显示用户资源使用上限,包含CPU 使用上限、内存使用上限、GPU型号和使用上限。
- 支持显示用户的算力消耗记录,包含时间、算力用量、占用信息,其中占用信息显示任务名称,支持任务名称跳转到该条任务中。
- 支持显示用户的算力获取记录,包含时间、分配算力等信息。
- 以正在运行和历史运行两个维度显示用户的任务统计,包含信息如下: 任务名称、任务类型、算力用量、开始和结束时间。

## 用户管理

在用户管理模块,支持主账号查看子账号的资源配额信息,对子账号进行设置算力型号、分配算力等操作。

### 用户列表

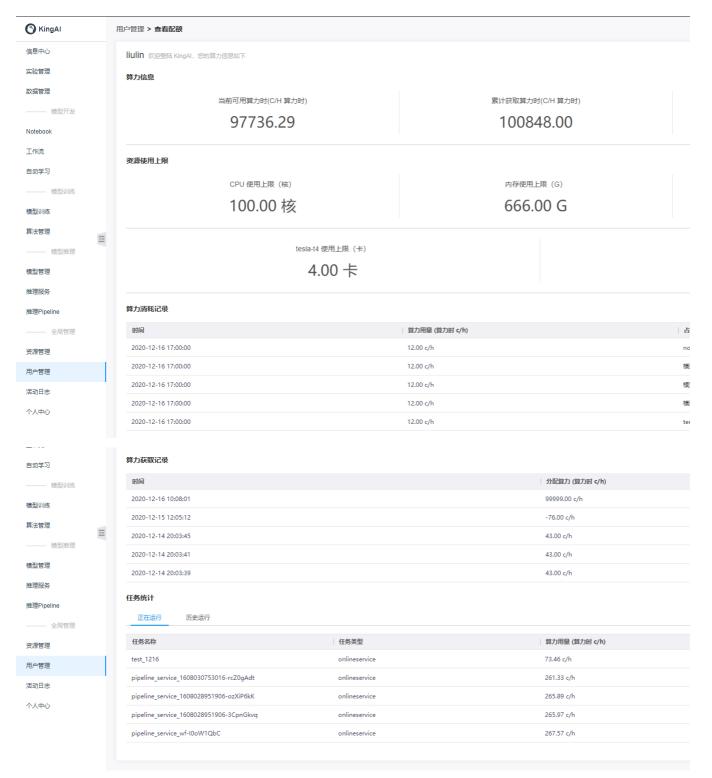
### 查看配额

支持主账号查看子账号的资源配额信息

### 操作步骤

- 1. 登录人工智能KingAI控制台。
- 2. 在左侧导航栏中,单击全局管理>用户管理。
- 3. 在用户管理页面,选择任意一个子账号,单击该列表操作列中的查看配额,可查看的配额信息如下:
  - 。 **算力信息**:显示用户当前可用算力时、累计获取算力时和累计消耗算力时。
  - 。 资源使用上限:显示用户CPU使用上限、内存使用上限和GPU使用上限(显示全部GPU型号和使用上限)。
  - 。 算力消耗记录: 支持的详细信息和功能如下:
    - 显示用户的任务算力消耗记录,包含时间、算力用量和占用信息等基本信息,以时间为维度,默认展示最近的5条任务的算力消耗记录。
    - 单击**查看更多**,查看全部任务的算力消耗记录。 单击**任务名称**,跳转到该条任务页面。
  - 算力获取记录:显示用户的所有算力获取记录,包含时间和分配算力,以时间为维度,默认展示最近的5条任务的算力获取记录,单击查看更多,支持查看全部任务的算力获取记录。
  - 。 任务统计:以正在运行和历史运行两个维度展示任务算力用量信息,统计的任务名称、任务类型、算力用量、时间等字段,默认展示最近的5条任务统计记录,支持单击查看更多,支持全部

金山云 2/6



## 设置算力型号

支持主账号设置子账号的算力额度上限,可以限制子账号CPU、GPU、内存等资源量的大小,子账号的资源型号和配置信息不能超过主账号设置资源上限。

## 操作步骤:

- 1. 登录人工智能KingAI控制台。
- 2. 在左侧导航栏中,单击**全局管理>用户管理**。
- 3. 在**用户管理**页面,选择任意一个子账号,单击该列表操作列中的**设置算力型号**。
- 4. 弹出设置算力型号弹框,可以设置用户的GPU、CPU、内存的使用上限。上限额度不可超过总算力 (即可被调度的部分),上限仅能设置正整数。

金山云 3/6



### 算力分配

支持主账号给子账号分配算力信息,子账号使用时,将根据自己使用的资源类型设置算力系数,进行算力计算,当算力消耗到小于0时,需要主账号进行算力充值。

## 操作步骤

- 1. 登录人工智能KingAI控制台。
- 2. 在左侧导航栏中,单击**全局管理>用户管理**。
- 3. 在**用户管理**页面,选择任意一个子账号,单击该列表**操作**列中的更多**算力分配**。
- 4. 可以设置用户的算力信息,在用户现有算力时(C/H)上增加新的算力时,可设置负算力时,分配算力仅支持一位小数。

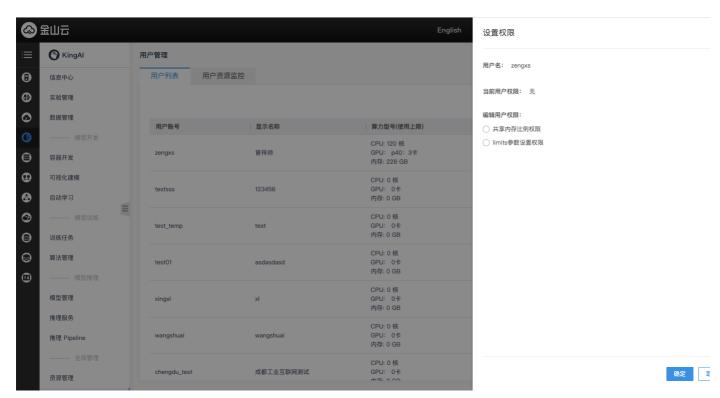


## 设置权限

支持主账号对其子账号设置内存比例权限和limits参数权限。 操作步骤

- 1. 登录人工智能KingAI控制台。
- 2. 在左侧导航栏中,单击全局管理>用户管理。
- 3. 在用户管理页面,选择任意一个子账号,单击该列表操作列中的**更多\*\***设置权限\*\*。 4. 主账号可以对子账号设置内存比例权限 、1 imits参数设置权限。

金山云 4/6



## 用户资源监控

支持主账号查看其子账号加速卡、CPU、内存等算力资源消耗情况。

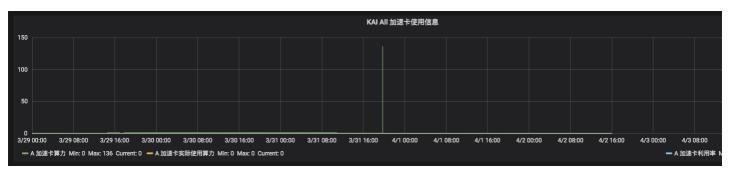
- 1. 登录人工智能KingAI控制台。
- 2. 在左侧导航栏中,单击全局管理>用户管理>用户资源监控。
- 3. 在用户资源监控页面,会显示该主账号用户维度的算力、内存消耗信息。

#### 算力消耗周报表

显示各个子账号以周维度的各算力消耗,主要包括:加速卡、CPU和内存三种算力的算力消耗和利用率情况。

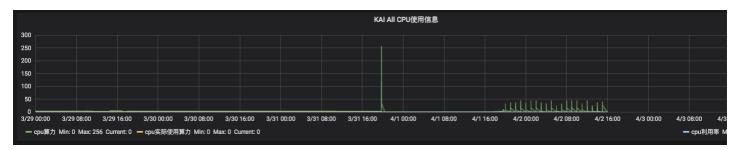


## 加速卡使用信息

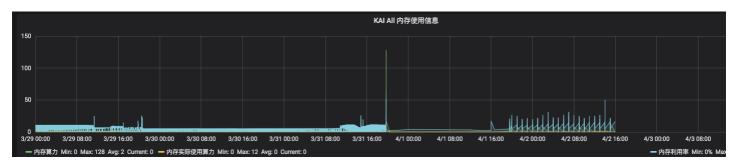


CPU使用信息

金山云 5/6



## 内存使用信息



金山云 6/6