上海科技大学教学实训平台
用户使用手册（V1.1）

图书信息中心

2022 年 5 月 20 日

版本历史

版本	提案人	批准人	日期	描述
1.0	刘祎阳	张迎冬	2020-9-20	初建
1.1	刘祎阳	孙思思	2022-5-20	修改部分说明

目录

1. 教学实训平台主页

https://etraining.shanghaitech.edu.cn

【注意】平台不支持 IE 内核的浏览器，必须使用 Chrome 或 Firefox 登录平台，建议分辨率为 1920x1020，如使用过低分辨率，请尝试将浏览器视图缩小到 80%。

2. 登录方式

平台登录支持两种方式。

方式一：上海科技大学统一认证系统登录，用户自动同步相应权限。

图 1

方式二：如果管理员已经分配了用户名和密码，直接通过平台登录。

图 2

如果没有用户名和密码，请联系管理员。

找回用户密码

当您遗忘自己账户的密码时，需要与系统管理员联系，通过系统管理员帮您进行密码重置。

Linux 账户信息

在创建账户时，系统会默认为您分配一个配套的 Linux 账户，系统会为 linux-user 分配一个默认的 ID，并将这个 ID 进行加密存储，在分配运行实例的时候，系统会将这个 Id 作为运行实例内部默认用户的 ID，从而将物理环境和容器环境的用户打通。

在进入运行实例与 Linux 系统进行交互时系统将以此 Linux 用户进行登入，您不需要记录此账户与密码，通过平台访问 Linux 运行实例是免密登陆的。

在平台的 Linux 运行实例（即容器实例）中，默认是以普通用户身份（非 root 用户）登入系统的，但在平台内进行实例交互时您可以随时通过 su 命令进行安全的提权，通过 exit 命令切换回普通用户，如下所示。

```
sudo su - exit
```

例如，在安装软件时，您可能需要 root 权限，如下命令片段所示：

```
sudo yum install vim
```

root 用户的权限仅在容器实例内部有效，不必担心破坏宿主机器的软件和系统环境，您可以大胆的在容器虚拟实例中进行任何操作。

如果您使用的镜像不是系统内置或以内置镜像衍生的，则可能在镜像内未安装 sudo 和 su 命令，那么可能就无法以免密方式切换至 root 用户了。

3. 基本概念

在进一步使用上海科技大学教学实训平台之前，您需要了解一些平台内的具体概念，上海科技大学教学实训平台的设计理念基于简单易懂的原则，因此您可以很容易理解这些基础概念。

3.1. 容器与镜像

在上海科技大学教学实训平台中您不需要深刻的理解容器技术，平台为您提供了所见即所得的构建容器以及镜像的方式，您可以像使用普通 Linux 主机一

样的方式在上海科技大学教学实训平台内灵活的构建和维护您的应用。

在上海科技大学教学实训平台中最基本的运行实例是容器（Container），无论如何进行业务组合，最小运行单位都是容器。

在容器中封装了应用或服务程序、相关依赖环境和软件包、应用数据以及相关网络配置，您可以将容器当做一个打包好的运行整体看待。而这些封装的静态存储方式是容器的镜像（以下简称镜像）。

镜像的本质是一个打包后的文件，因此它具有移植性，在上海科技大学教学实训平台系统内容器镜像被系统的镜像库所管理。因此您不需要了解更多的细节，只需要知道一个容器运行实例是通过对应的镜像启动的即可。

通过镜像启动的容器运行实例将具备镜像中所存储的应用或服务程序、相关依赖环境和软件包、应用数据以及相关网络配置。同一个镜像可以启动很多容器运行实例，因此只需要构建一次镜像，就可以运行多个相同的应用。

如果您是学习编程的技术人员，那么镜像和容器运行实例可以被比喻为类和对象的概念。

3.2. 模板、模板实例、实例的生命周期、实例性能要求

在上海科技大学教学实训平台中业务对象都是以模板的概念进行组织的。

在系统中，模板可以创建一个或多个该模板的运行实例，有个生动的比喻就是“月饼的模具可以创造很多相同的月饼”。

模板与模板实例是镜像与容器运行实例的更高一层的封装。模板中包含了使用哪些镜像启动实例的描述，同时也约束了这些实例启动的性能配置要求、数据配置要求、运行控制要求和模板实例存活的生命周期描述。

例如模板中会描述以 4 核 CPU、16GB 内存、10GB 本地存储和 1 块 GPU 卡的方式，使用 A 镜像，启动 5 个容器运行实例。

模板实例的生命周期包含如下相关事件：

- 申请或撤销申请
- 审批或驳回
- 排队等待资源分配或取消排队
- 创建运行实例
- 到期销毁或主动销毁、强制销毁

- 释放资源

3.3. 模板 to 实例与实例 to 模板

模板 to 实例就是通过模板创建运行实例的过程，用户在执行这类操作时通常系统会给出申请表单，让用户可以对模板的运行参数进行相关调整，然后依据用户的最终需求执行实例创建的过程。

模板 to 实例强调的是使用，用户可以使用自己的私有模板创建运行实例，也可以通过他人共享或发布的模板创建运行实例。

实例 to 模板是相反的过程，是上海科技大学教学实训平台为用户提供简单直观的固化运行环境的方式，用户可以造作运行实例，并将变化保存至模板对象中，以便下次启动时携带已经变更的环境或程序改动。用户也可以将他人共享的或发布的模板另存为自己的私有模板，并在此基础上进一步修改变成新的应用。

3.4. 集群应用中的镜像、Role、实例的概念

在任何分布式集群应用环境中，一个应用都是由多种实例角色(Role)组成的，每个实例角色都是以相同运行配置、相同应用配置运行的实例组成的，其逻辑关系如下图所示：

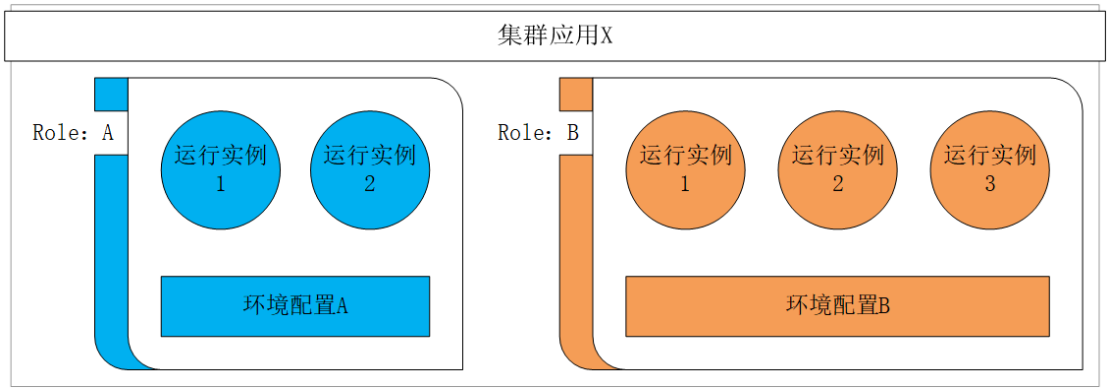

图 3

例如在高性能计算领域通常一个集群调度系统会有两种角色存在，分别是调度节点和计算节点，调度节点可以分为 Master 和 Slave 两个等价实例互为主备备份，计算节点可以使 1~n 个物理服务器组成，调度节点可以使用一致的软件和环境配置，计算节点可以使用一致的软件和环境配置，它们共同组成了一个分布

式集群应用环境。

例如在深度学习模型训练领域，比较重要的角色概念是 PS 与 Worker 以及 Master 主机，其关系图如下所示：

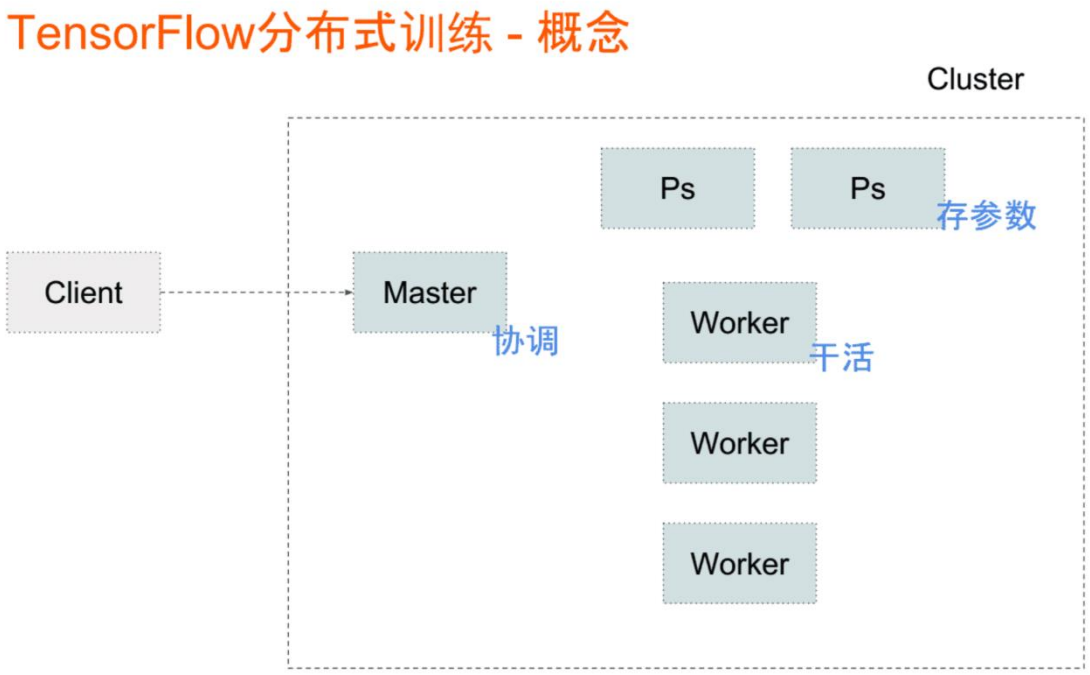

图 4

根据上图所示，首先简单介绍下参数服务器的基本概念，便于读者理解角色划分的概念。

一般而言，机器学习的参数训练过程可以划分为两个类别：

第一类： 根据参数算算梯度。

第二类：根据梯度更新参数。

对于小规模训练，数据量不大，参数数量不多，一个 CPU 就足够了，两类任务都交给一个 CPU 来做。

对于普通的中等规模的训练，数据量比较大，参数数量不多，计算梯度的任务负荷较重，参数更新的任务负荷较轻，所以将第一类任务交给若干个 CPU 或 GPU 去做，第二类任务交给一个 CPU 即可。

对于超大规模的训练，数据量大、参数多，不仅计算梯度的任务要部署到多个 CPU 或 GPU 上，而且更新参数的任务也要部署到多个 CPU。

如果计算量足够大，一台机器能搭载的 CPU 和 GPU 数量有限，就需要多台机器来进行计算能力的扩展了。参数服务器是一套分布式存储，用于保存参数，

并提供参数更新的操作。

简而言之 TensorFlow 的分布式训练集群应用将参数的存储与更新计算放在了 PS 服务器上，因为这部分涉及到的“计算”任务量和存储量较小，将计算梯度的任务放在 Worker 服务器上，Worker 服务器可以进行横向扩容以便提高梯度计算的计算能力。

从这种模式来看，TensorFlow 的分布式计算应用就可以划分为 PS 角色和 Worker 角色，这些角色下面可以根据计算规模的不同设定不同的角色实例，例如对于人脸识别模型，300 万训练集规模的训练任务，可以设定两台 PS 角色实例和 30 台 Worker 角色实例，每个 Worker 实例上都可以配置 GPU 以加速梯度计算，这样综合下来平均每台 Worker 并行的负载约为单机计算的 1/30，使得训练的效率提升了约 30 倍。

在平台中为了支持复杂集群场景，默认所有的应用都是由角色和实例组成的，您可以创建多个角色，配置选择不同的镜像，并给以角色不同的运行性能配置（例如 Worker 服务器都会配置 GPU，PS 服务会配置更大的内存）。

当系统以此配置完成环境启动时，您将拥有一套多角色、多实例的集群应用环境，环境的具体性能可以根据计算量的需求在不同场景下进行调整。不仅如此，由于每个角色实例在平台中都是容器运行实例，因此您在平台上所构建的集群应用具备良好的打包移植性，可以将这些容器镜像部署至性能更高的独立运行主机上，或直接迁移到机构的另一个机房。

3.5. 平台中的业务分类

在平台中为例支撑全场景的生态环境，系统将业务划分为：

- 自助上机业务
- 实训实验业务
- 数据管理业务
- 镜像仓库
- 工作流业务

自助上机业务：主要面向的场景是研发、科研、业务计算，用户可以像使用虚拟机那样申请任意配置的环境。

实训实验业务：面向的场景是企业内的实训，支撑企业的技术传递培训、练习场景，在此场景下内容的创建者可以限制使用的资源数量和时间

数据管理业务：是一项支撑业务，系统内置数据存储管理的能力，并能够将这些数据挂载至其他业务实例中，从而将应用与数据绑定，形成完整的运行环境。

镜像仓库：是系统内的核心支撑业务，平台中所有环境的构建都离不开镜像仓库，此镜像仓库是平台内的私有镜像仓库，您可以放心创建新的镜像。

工作流：workflow 业务是平台面向生产级的业务，在此业务中您可以创建各式各样的自动化工作流，例如：顺序工作流、并发工作流、条件工作流以及有复杂依赖关系的工作流。工作流业务所使用的核心资源是容器镜像，您可以直接从镜像仓库中获得，也可以使用私有的镜像构建工作流。

3.6. 私有网盘与数据集的概念

在数据管理业务中，您可以在平台上申请一块私有的存储空间，即私有网盘，用于代码和数据的持久化存储，默认所有运行的环境实例都会将您的私有网盘挂载上，运行实例的生命周期是有时间限制的，而私有网盘的生命周期则是随用户的注销而消亡的。

数据集是平台数据管理业务的基本管理单位，您可以创建、销毁一个特定的数据集，例如面向目标识别创建一个 ImageNet 数据集存储待训练的图像和标签数据。

私有网盘其实是一个特殊的数据集，有别于数据集的选择挂载特性，私有网盘是默认挂载，该特性使得您在使用任何实例时就像是使用一块虚拟的 USB 存储设备一样方便。

您可以利用平台提供的发布、共享机制为平台生态贡献数据生态的力量，使得数据可以传播、受控传播。

3.7. 业务公有、私有、共享的概念

在平台内所有的业务对象都具备的一个属性即是业务的可见性、可使用性，

也就是公有和私有的概念。

用户自己创建的业务模板对象（自助上机的业务环境、实训实验环境、自动化工作流模板、数据集），默认情况下其业务的可见属性都是私有属性，也就是仅创建者自己可以看到和使用，它们存储于用户的个人业务列表中，如下图所示：

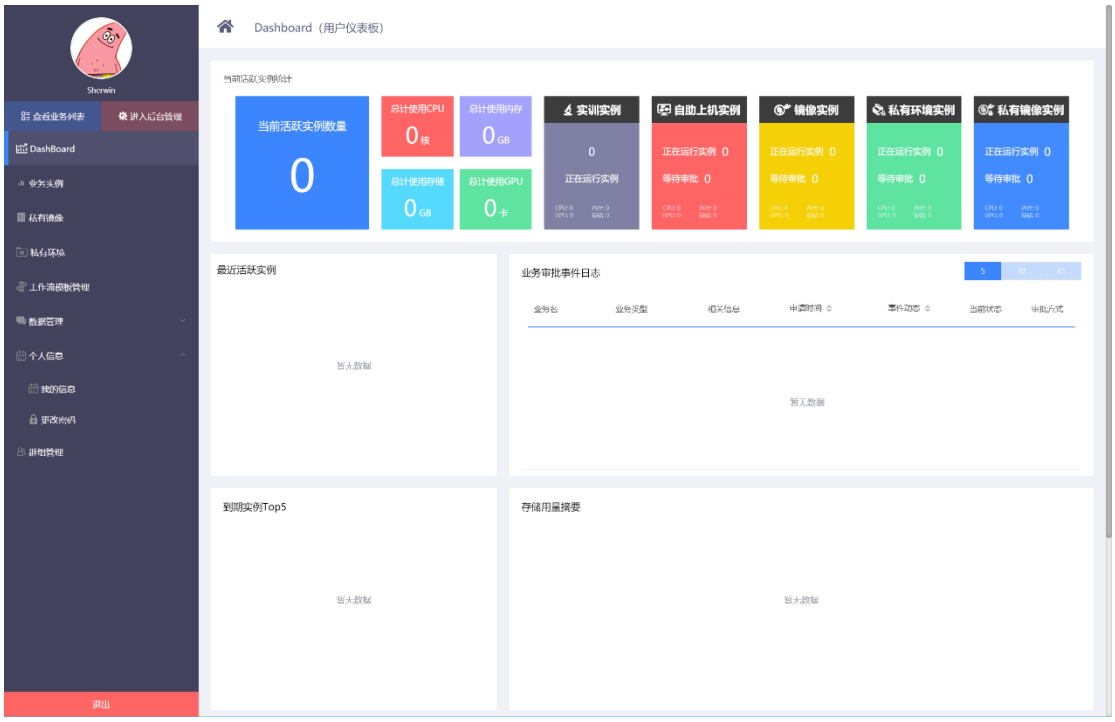

图 5

用户可以通过发布和共享操作调整这些业务模板对象的可见性和可使用性。当您公开发布一个业务模板后， 那么用户便可以从公有业务列表中找到您发布的业务对象，进而可以使用它并在您构建的业务环境基础上进一步优化、扩展，再发布，以此平台便为企业提供了一个可控的开放式平台生态环境，不断的积累、迭代优化各种业务和创新的研发成果。

对于公有、私有之间的转换，用户可以通过将一个公有业务对象另存为私有的方式和将私有业务对象发布为公有方式完成。

当您认为目前的业务尚未成熟或基于业务密保隔离要求不能进行公开发布时，您还可以利用平台提供的群组关系进行局部的共享，以达到在局部范围内发布的目的，此时仅有您允许的用户可以见到并使用这些业务对象。

3.8. 角色与权限的概念

用户账户是平台内唯一的鉴权依据，每一个用户账户都会绑定一个用户角色，而每一个用户角色都会配置一系列的可操作权限，当您使用具体业务和系统功能时，只有权限与功能匹配时系统才允许您使用。

例如，将一个自助上机环境模板另存为私有环境模板时，需要系统对您的权限进行鉴权，只有鉴权通过才能允许您完成另存为的操作，若您所绑定的角色没有这个权限，则鉴权失败，系统不会让您完成此次另存为操作。

系统内大部分鉴权分布在有资源占用的特性，例如数据集的创建权限将消耗系统存储资源，将公有自助上机模板存储为私有自助上机模板时会消耗镜像库的存储资源。

用户角色具备哪些权限由系统管理人员进行定义，也可以有管理人员在线进行调整，当您遇到所需业务但没有权限使用时您可以通过联系系统管理人员帮您提升权限（即为您的账户变更绑定角色）。

4. 功能简介

对于学员角色来说，系统一共有两项主要的业务功能，它们分别满足学员不同的使用场景，它们分别是：

4.1. 实训实验

学员可以通过在线浏览自己感兴趣的课程参与既定课程的实训练习，练习的方式在时间和内容维度上随学员的兴趣爱好，在使用的资源和课程设定上是由课程设计师设计好的，学员需要跟随课程的设定进行实训练习，此业务特性比较适合清晰的知道自己的擅长点和待巩固知识点的场景。

4.2. 自助上机

在日常学习中，有时候学员更需要的是一个灵活的练习环境，自助上机业务是学校传统机房的一种扩展，采用云技术让学员选择自己的上机环境（性能、系统要求可配置的上机环境）。学员可以根据自己的需求进行申请，

申请后管理员会对申请的合理性进行审批，审批通过后学员可以长期使用自己的上机环境，该业务没有课程设计的相关约束，拥有更灵活的使用时间安排。

4.3. 数据集

Linux是一个多用户的操作系统，用户的数据有严格的权限设定。在进行机器学习等计算相关的实验时，用户经常会碰到数据共享的问题，通常只有两种方式，一是将数据另外拷贝一份或改变数据的权限；另一个是将自己的账号给其他人使用。这两种方式都存在一定的问题。数据集是一种逻辑的数据共享方式，用户可以指定数据，设置为可以共享或发布的数据集，其他用户在使用数据时，只是将数据挂载到相应的环境中，无需额外浪费空间。

5. 实训实验

打开系统主页后便可以看到所有的实训实验课程列表，每一个课程项都会展示，课程的标题、难度系数、预计学习课程的时长和访问热度，可以根据这些基本信息来识别和确定某一个实训课程的参与意愿。

5.1. 课程时长

课程的时长通常来讲都是一个估计值，对于学员的学习来说都是有一定的延长的，例如1个小时可以完成的实训，课程设计时都会给3个小时的有效时长设定，主要是考虑到学员还要在学习的过程中对内容进行理解、尝试、解决问题等行为。

5.2. 课程的生命周期

课程的时长设定涉及到实训实例有效期的概念，在进行自助实训的时候是存在有效期限制的，有效期到时，系统将会自动的回收正在使用的实训实例环境，以便供其他人继续使用，因此需要注意实训实例到期时间，在实训即将到期之前系统会进行提醒。

实训实验的延期申请：系统也考虑到了课程有效期不足的情况，当您在潜心研究课程的理论知识和实训案例的时候，有可能会花费大量的时间和精力去查找资料，尝试编写不同的程序，修改不同的参数进行试错或改进课程设计师给出的案例，这都将使您获得更多的知识，并对理论理解的更加透彻。在这种情况下系统会给您提供免审批的延期申请方法，在“实训实验业务”中，您可以在实训进程超过一半有效期的时候提出延期申请，系统会自动帮您延长一倍的有效期。

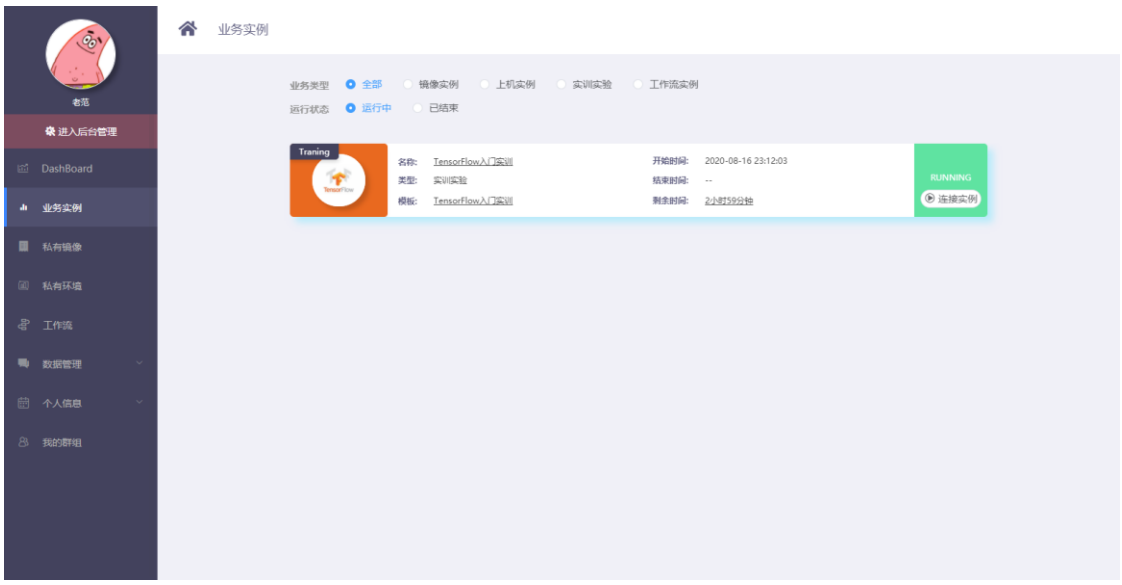

图 6

系统限制了每个用户使用系统资源的数量，您至多可同时参加3个自助实训课程。系统在尽可能的避免系统资源的浪费，也请您在不需要使用时及时释放资源，在必要时通过申请延期增加使用资源的时间。

5. 3. 浏览实训教材

您可以在概要介绍的页面中打开此课程对应的教材列表，这能够让您更加深刻的了解即将要参加的课程，点击阅读教材后，教材列表会从系统的左侧弹出，点击列表以外的位置收起教材列表。当然您在实训操作的过程中也可以随时看到教程与教材。

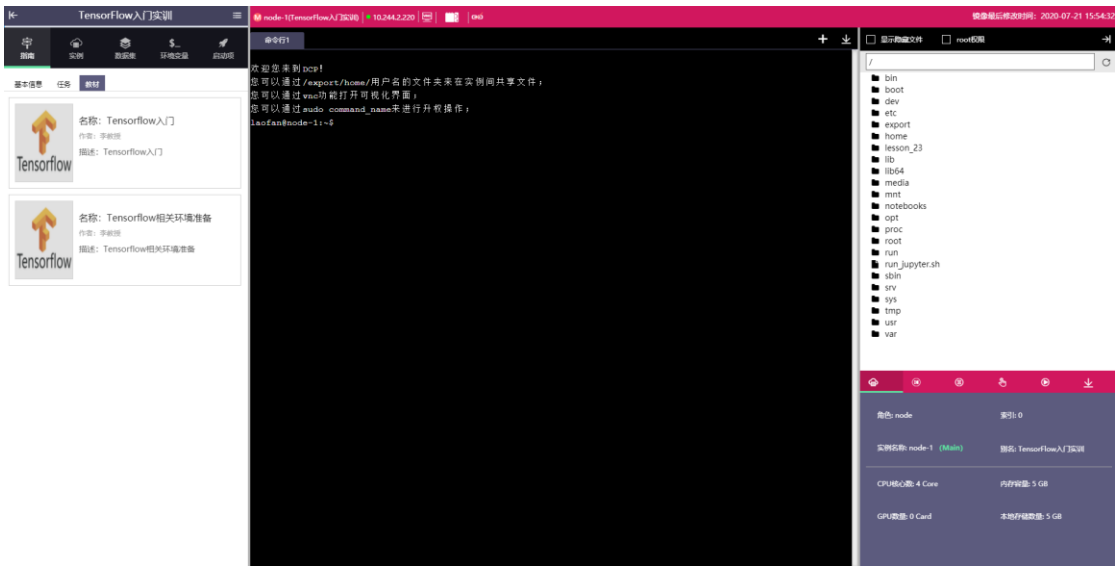

图 7

5. 4. 参加实训

当您确定需要参加该实训的时候，可以点击开始上机，此时系统会出现连接实训实例的画面，稍等片刻就可以打开实训的交互系统了。稍等片刻系统便会帮助您连接到实训环境中，在系统内您不需要记住实训环境的账户和密码，也不需要记录实训的I P地址，这些全部都是自动化分配与连接的。

一切正常的话您将看到实训环境的操控界面：

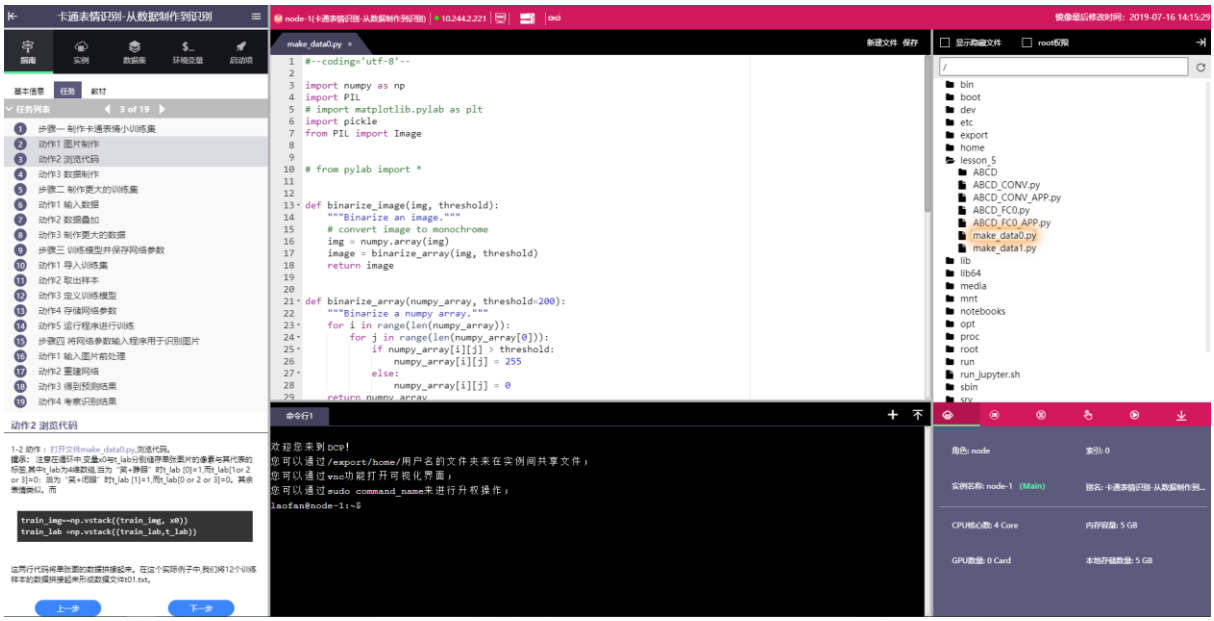

图 7

5. 5. 交互系统

交互系统是上海科技大学教学实训平台的核心功能，它让您能够非常轻松的与后端复杂的计算机集群和系统软件环境进行交互式操作。无论是“实训实验”、或是 “自助上机”都会用到交互系统。

上海科技大学教学实训平台的交互系统被称之为：UXSYS，它不仅能够让您轻松的访问和操控云端的实训环境，还提供大量的配套工具帮助您完成实训课程，这让初入计算机领域学员 避免了很多的麻烦。(AI 、HPC、大数据等都属于交叉学科，大部分行业背景的学员在上手计算机领域时都会遇到一些麻烦，UXSYS 在尽可能地简化计算机的交互使用过程。)

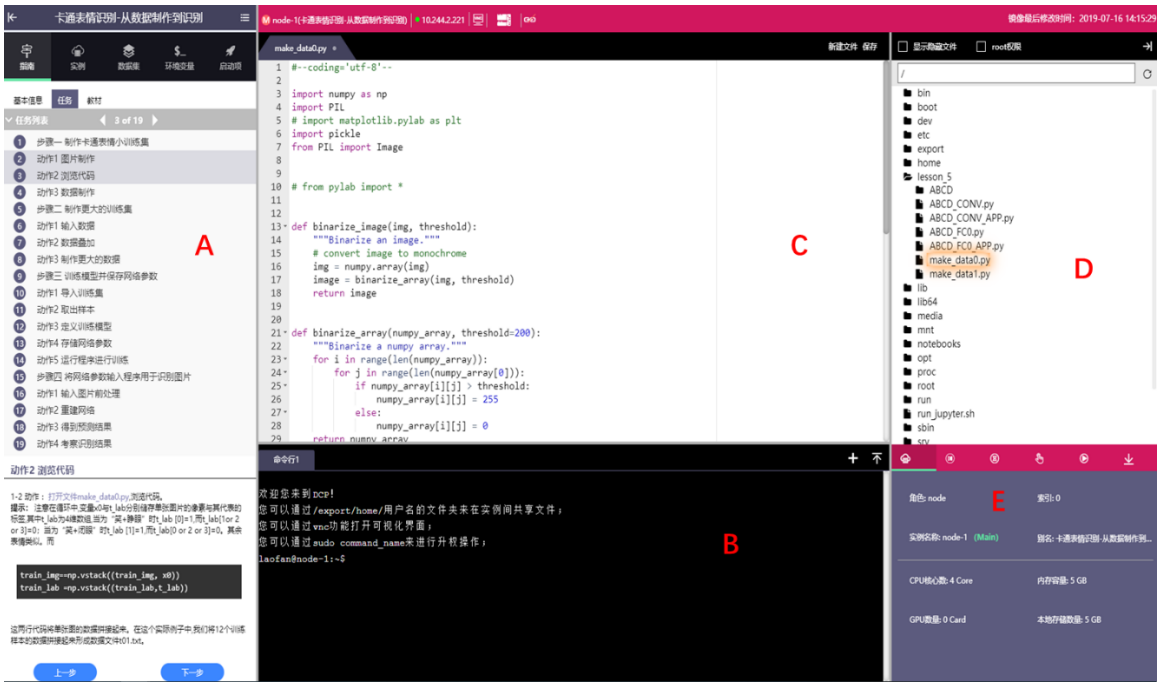

图 8

UXSYS 界面是一个集成界面，一共分为5个区域：

A. 区域A为实训引导系统界面，在该区域内展现的是实训课程的引导指南，该指南可以让您在学习的过程中一步一步地跟随课程设计师的指引完成最终的实训任务，在指南中提供大量的便捷工具（知识点的提醒、快速输入命令行、复制编码、寻找指定目录或文件）帮助您屏蔽知识学习以外的麻烦。

B. 区域B为标准命令行输入输出窗口，它后面连接的是系统为您准备好的实训环境， 并且该实训环境已经为您安装了必要的软件和部署了实训所需要的数据（如深度学习中所需的训练数据）。

C. 区域C为实训环境的文本、代码编辑器，在实训的过程中您不需要单独准备代码 编辑器了，它可以高亮您的代码，让您在集成的环境中完成实训环境的操控和代码编辑，当然您也可以通过命令行窗口使用VIM进行编码。

D. 区域D为便捷式的文件资源管理器，它连接后端的实训环境，并以更方便的 GUI 方式提供给您进行操控。在文件资源管理器中，您可以双击打开文件进行编辑，也可以创建文件、目录、对文件重命名等常规文件操作。

E. 区域E为信息查看区。用户可以查看启动环境的一些基本信息，如：资源利用率。

5.6. 基本概念说明

接下来我们先了解一下上海科技大学教学实训平台系统里的几个名词与概念，这样可以让您更好的使用UXSYS。

- **什么是实训实例？**

实训实例在上海科技大学教学实训平台中所代表的是一次实训过程所使用的所有资源，这些资源包含 环境实例、实训指南、实训教程与教材、实训所使用的ShowApp应用、实训所使 用 的远程桌面连接等。可以这样理解，您参加实训就会产生一条实训记录，这条实训记录就是一个实训实例，在这个实训实例的生命周期内您可以使用这个实训实例中的环境资源、ShowApp、教材等内容。

- **什么是环境实例？**

环境实例是承载实训的主要环境载体，在上海科技大学教学实训平台中是您所能够连接的计算机设备，您在使用过程中看到的是一台标准Linux系统的计算机设备，而实际上，它们是被虚拟化技术和容器技术所包装出来的虚拟计算机（但它们表现的行为和真实的计算机是一样的）。

- **多实例是怎么回事以及实训实例与环境实例的关系？**

在一个实训实例中是可以包含多个环境实例的，即上海科技大学教学实训平台支持集群式应用，您可以在参加一个实训的时候同时控制多个环境实例，环境实例有多个，而他们都属于 一个实训实例。通过环境实例标签，在不同环境实例之间进行切换，在同一个环境实例中可以创 建另一个SSH连接，这样可以方便您运行一个程序的同时，观察这个程序的输出以及资源占用情况。

- **什么是ShowAPP**

ShowApp在上海科技大学教学实训平台中就像微信小程序一样，课程设计者设计了很多解决实际问题的实训案例，ShowApp则以一种非常形象的方式为学员展现了这些实训案例。

您可以在窗口的最上方找到S howApp入口，单击ShowApp启动开关，系统会在新的页面中打开ShowApp，使得您可以对自己改造的程序进行交互式的测试，正如下图所示（下图是与人工智能对弈的实训实例，您可以换成您

的深度网络然后试试您自己创建的人工智能智商如何）。

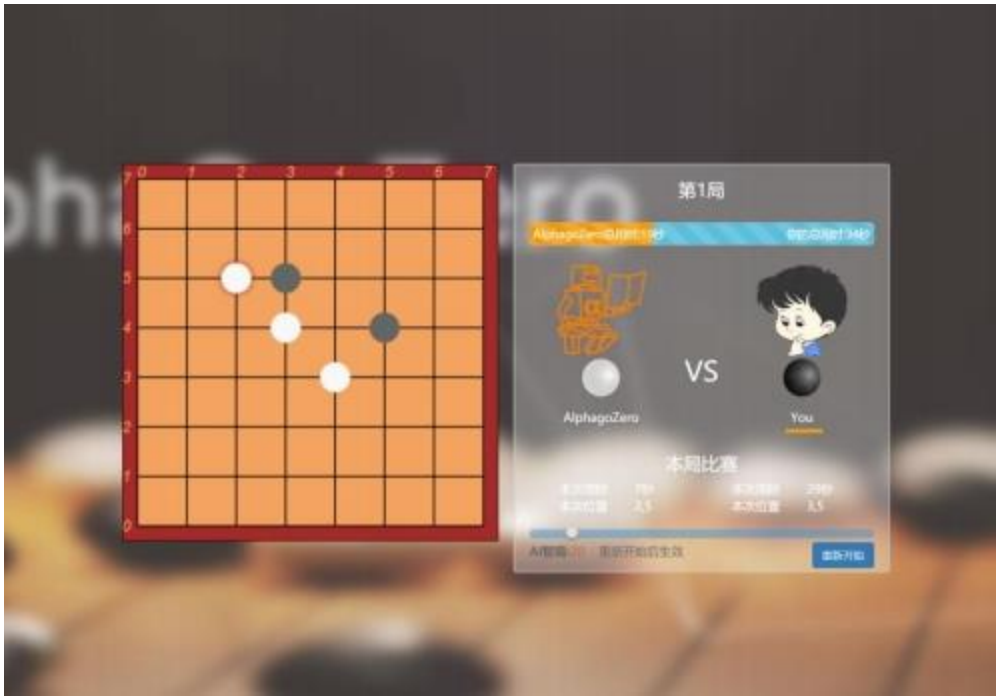

图 9

- **远程桌面功能**

为了满足部分实训是在GUI 系统下进行的，实训平台提供了远程可视化桌面的方案，学员在实训过程中可以随时通过远程桌面连接协议连接到环境实例的桌面上，从而操作一些图形化界面的相关实训。如图所示：

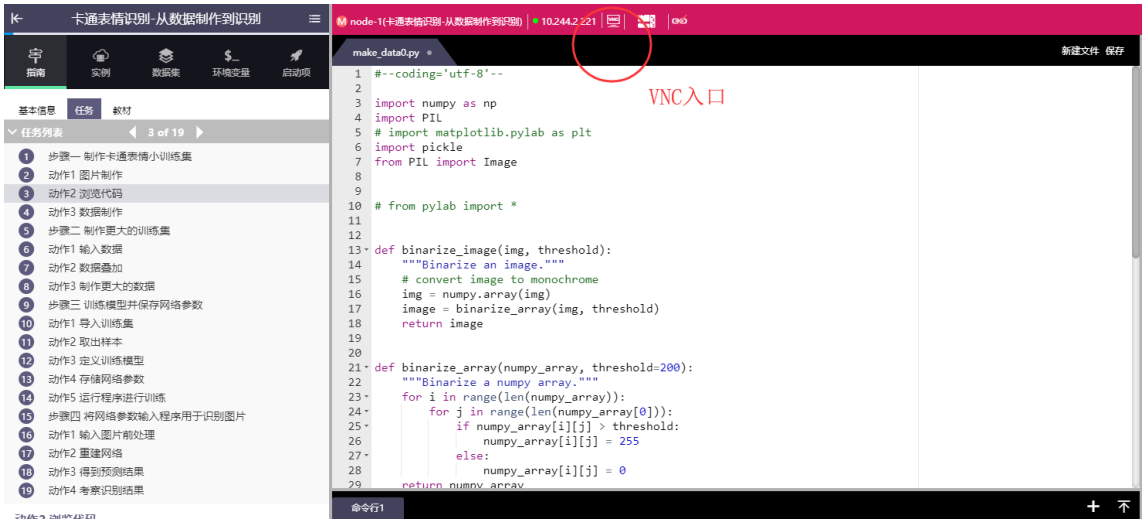

图 10

图 11

- **联动交互:GuideTag系统**

GuideTag系统是上海科技大学教学实训平台非常具有特色的子系统，它为实训课程提供了更多的交互 方式，其主要目标是让实训过程中的理论知识呈现层次化，更加突出实训的重点。GuideTag提供了非常多的交互方式，例如当课程设计师需要您在某一个目录下创 建一个文件时，虽然是一个非常简单的动作，但它取决于课程设计师是否精确清晰地说出了目录的位置，这个时候GuideTag系统提供了一个非常方便的办法，您只需点击一下指南中带有Tag的链接，系统便会联动环境实例中的文件列表高亮要查找的目录。

① 图像、原理图放大展现

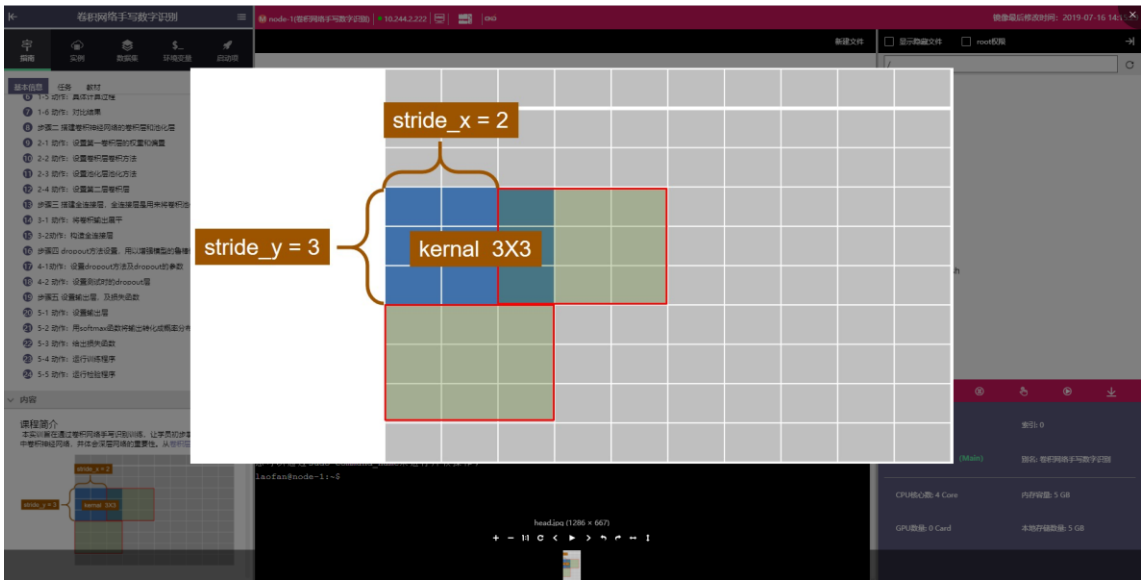

图 12

② 自动命令行填充与执行

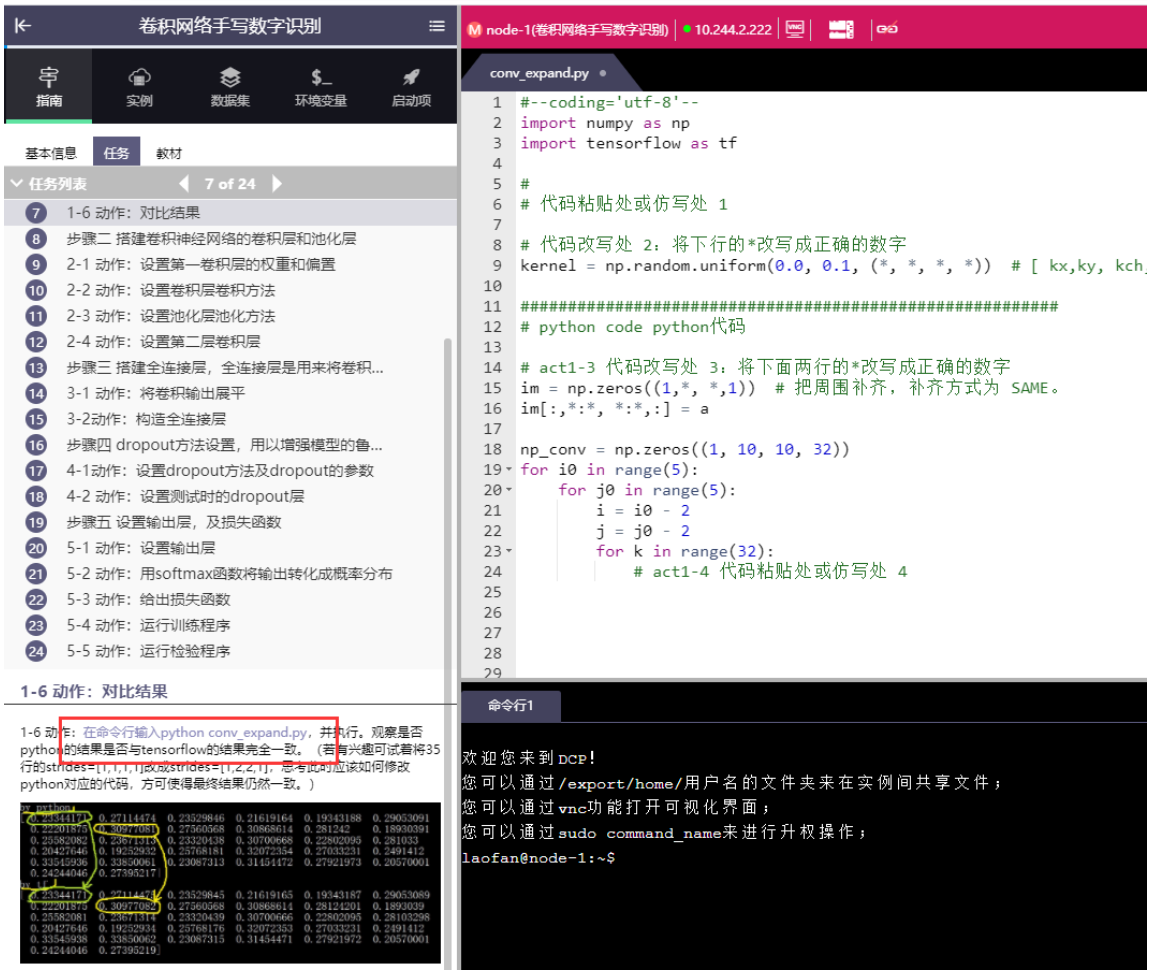

图 13

③ 从指南中自动打开文件

在上海科技大学教学实训平台的实训课程中，您随处可见这种细节的交互，帮助您省去很多时间去查找注解资料、找寻目录或文件位置、手动输入很长的命令行，上海科技大学教学实训平台在所有的细节 上希望您能够将时间用在知识的获取上，而不是这些琐碎的任务上。

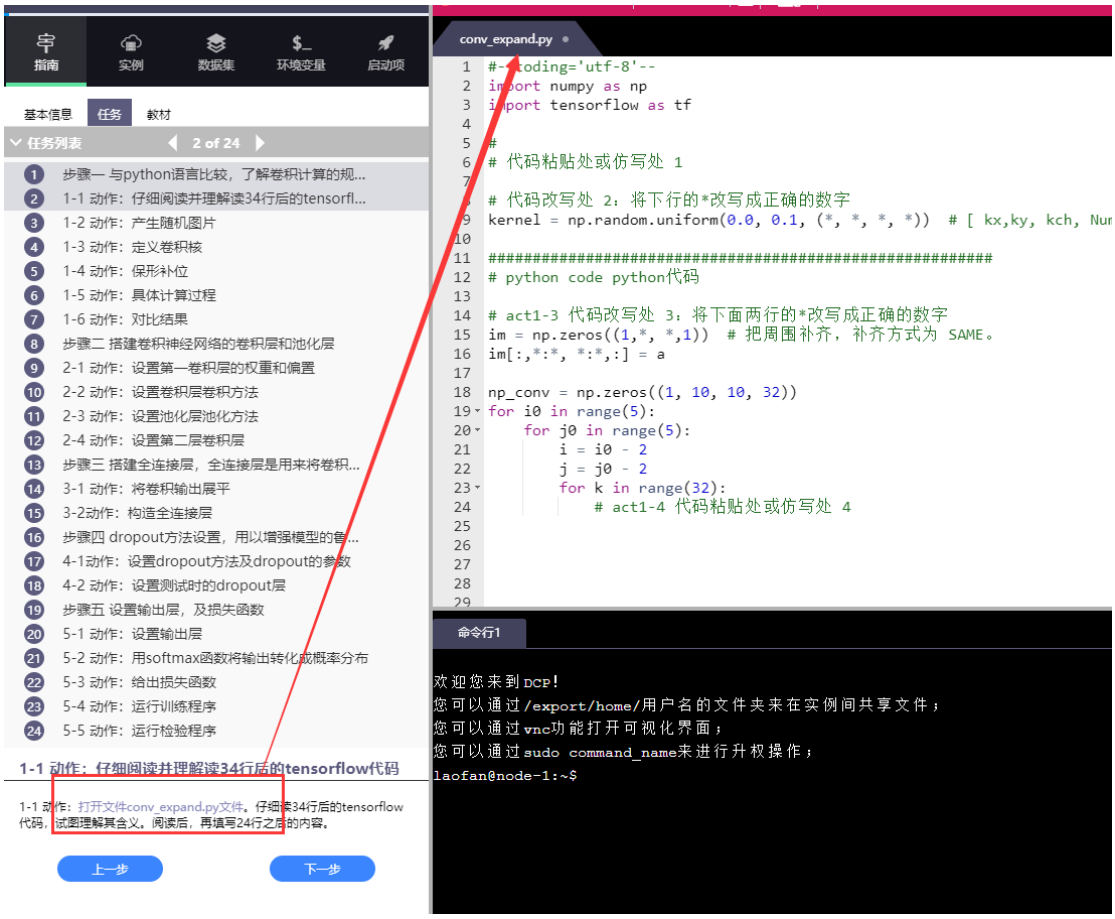

图 14

- 任务系统

在实训的过程中您需要按照课程设计者的设计，一步一步地完成一个实际问题的 解题，课程设计者将一个复杂的问题拆解成了一系列的任务，系统会清晰的将任务呈现给您，不仅如此，系统还会智能地检查任务的完成情况。

完成一个任务时，如果看到了任务标题后面有一个“检查”按钮，您可以点击它，如果任务完成则系统会提醒您成功完成了这个任务，并且整体的实训进展也会向前前进一大步。

- 有关退出交互系统的说明

在上海科技大学教学实训平台中退出交互系统非常简单，只需要您关闭

交互系统的浏览器界面即可，虽然您退出了交互系统，但是实训实例以及实训实例中包含的环境实例都未消失，它们都在云端，当您下次通过系统再次连接到实训实例时您所运行的应用仍然在继续运行，您所存储的数据依然在文件系统中保存着。

6. 自助上机

实训实验的所有场景都是默认资源，在用户做实验的过程中，有可能需要对资源进行自定义，平台的自助上机业务就是为满足这种需求设计。在自助上机业务中，用户可以自定义申请资源的使用时间、节点数、每个节点的配置、加载的数据集等资源。

下面以申请一个包含两个节点 Tensorflow 环境为例说明自助上机的操作方法。

6.1. 登录进入平台后，选择“自助上机”

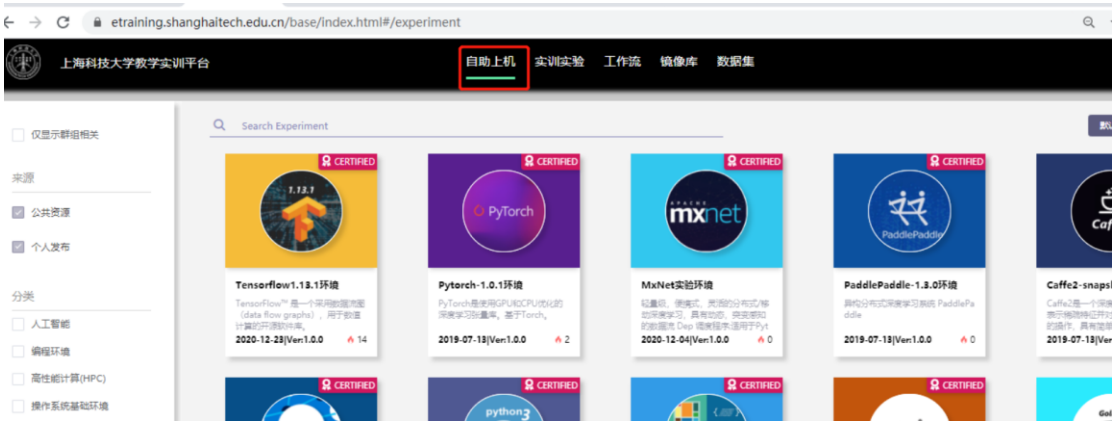

图 15

6.2. 选择其中一个实训环境

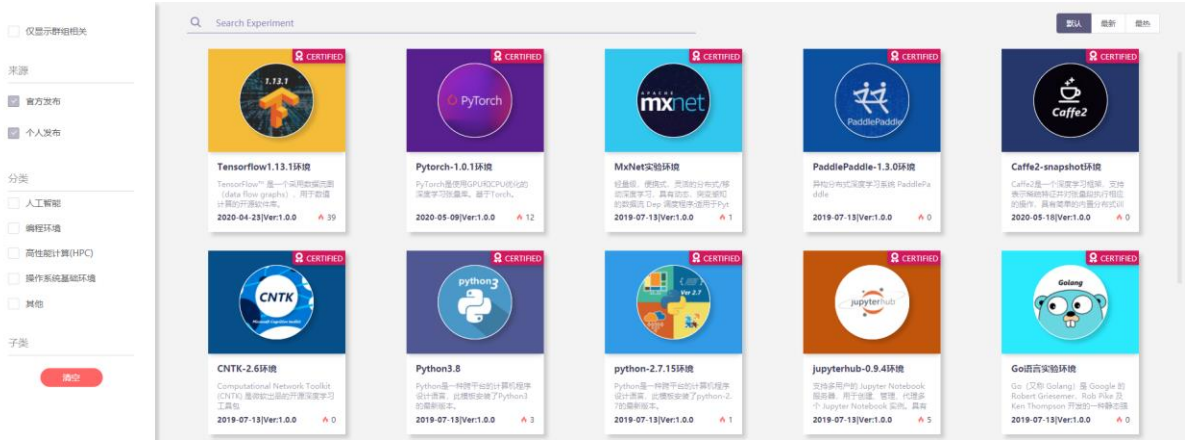

图 16

6.3. 进入申请申请环境实例

如下图，根据需要依次选择填写：

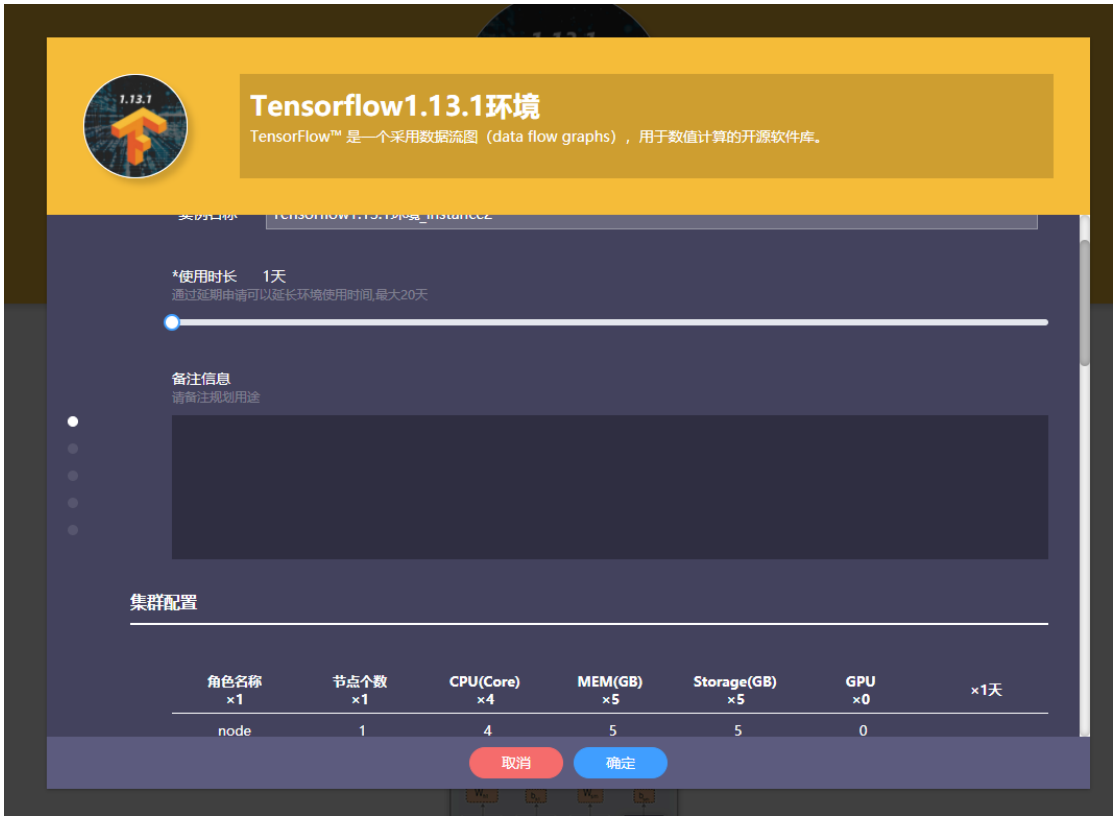

图 17

定义“使用时长”、集群配置（平台暂时没有对存储进行约束），“Storage”

部分可以忽略：

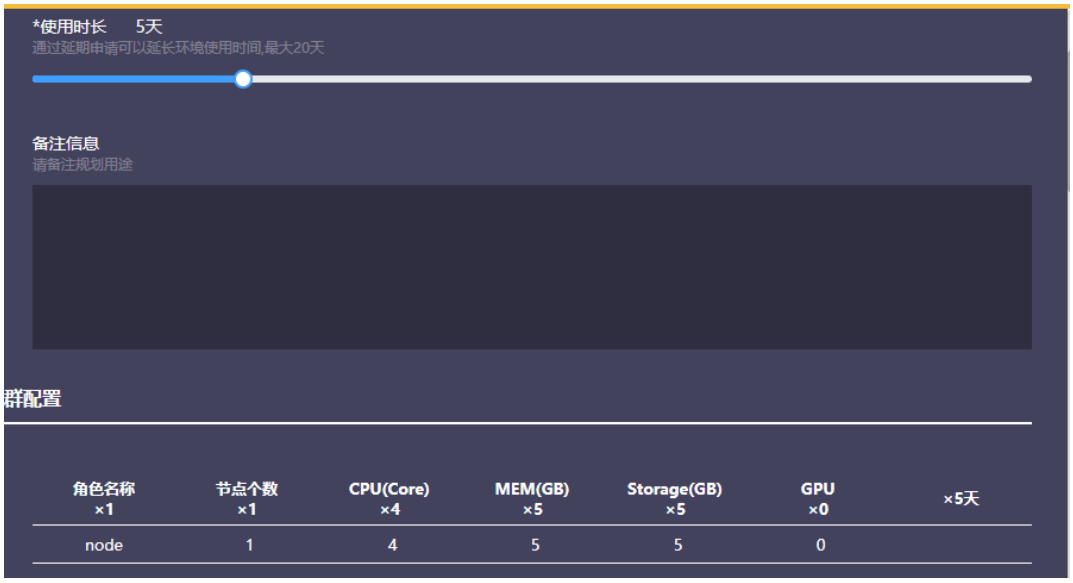

图 18

选择环境中要挂载的数据集：

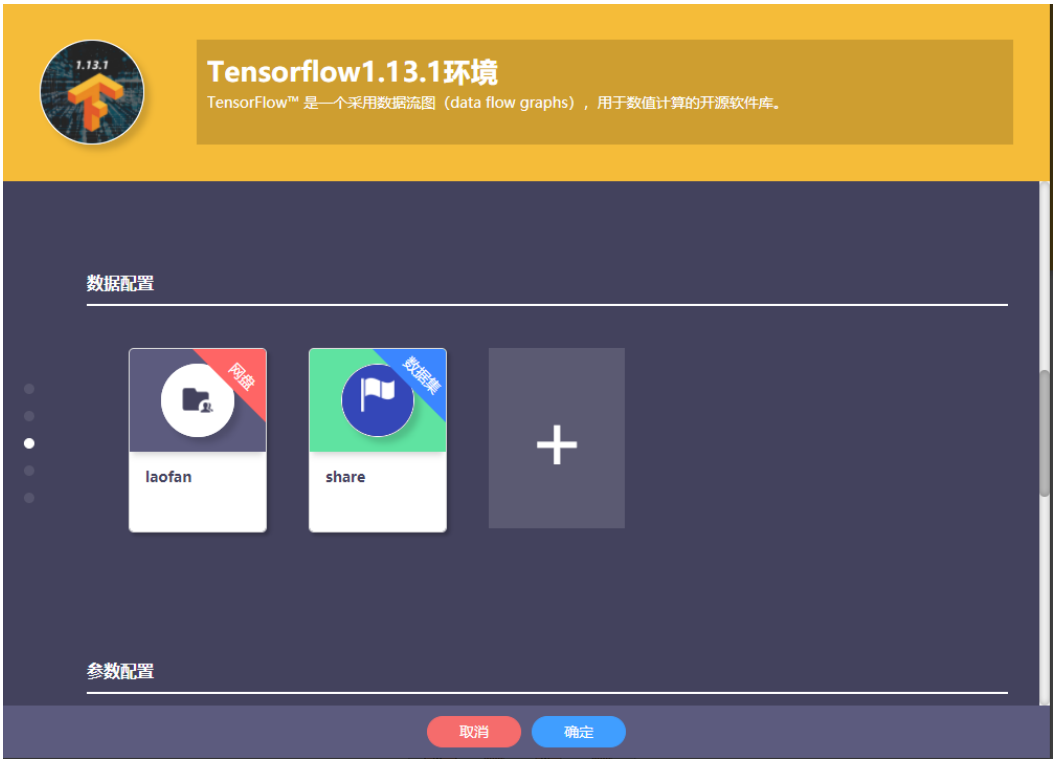

图 19

参数配置和启动配置两个选项，如果环境中没有配置，则无需更改：

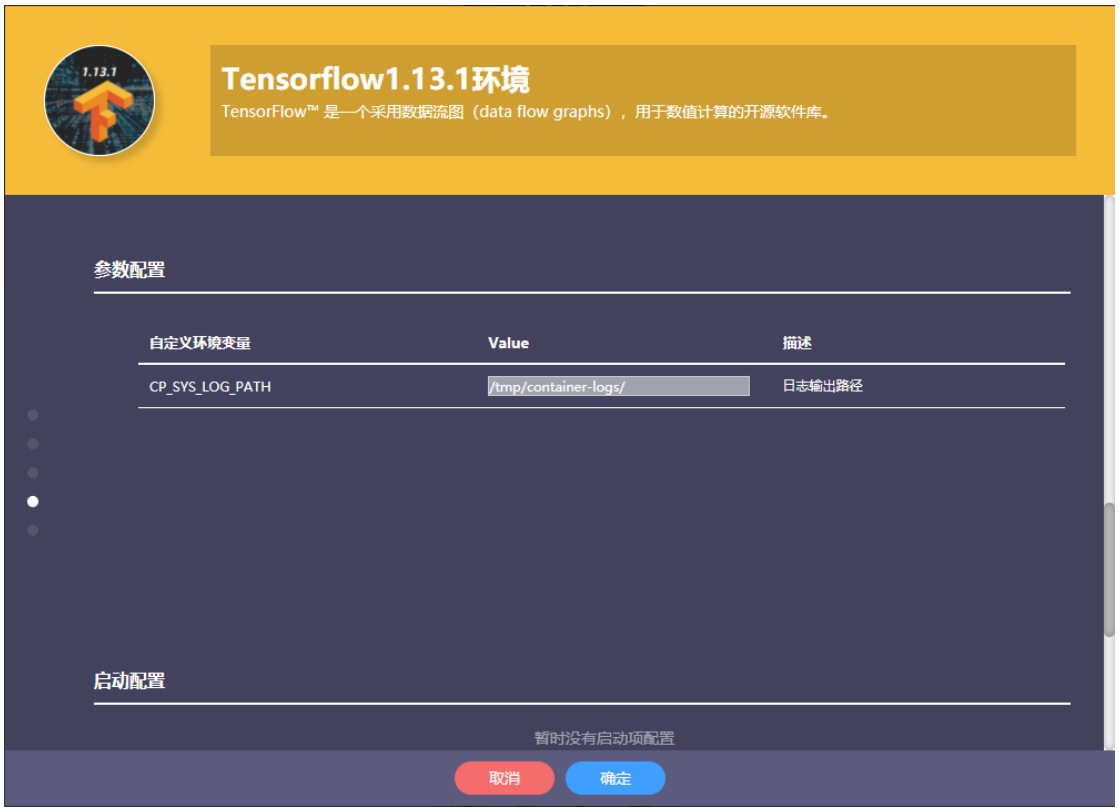

图 20

填写完成后，点击“确定”，选择进入“进入自助上机实例列表查看”：

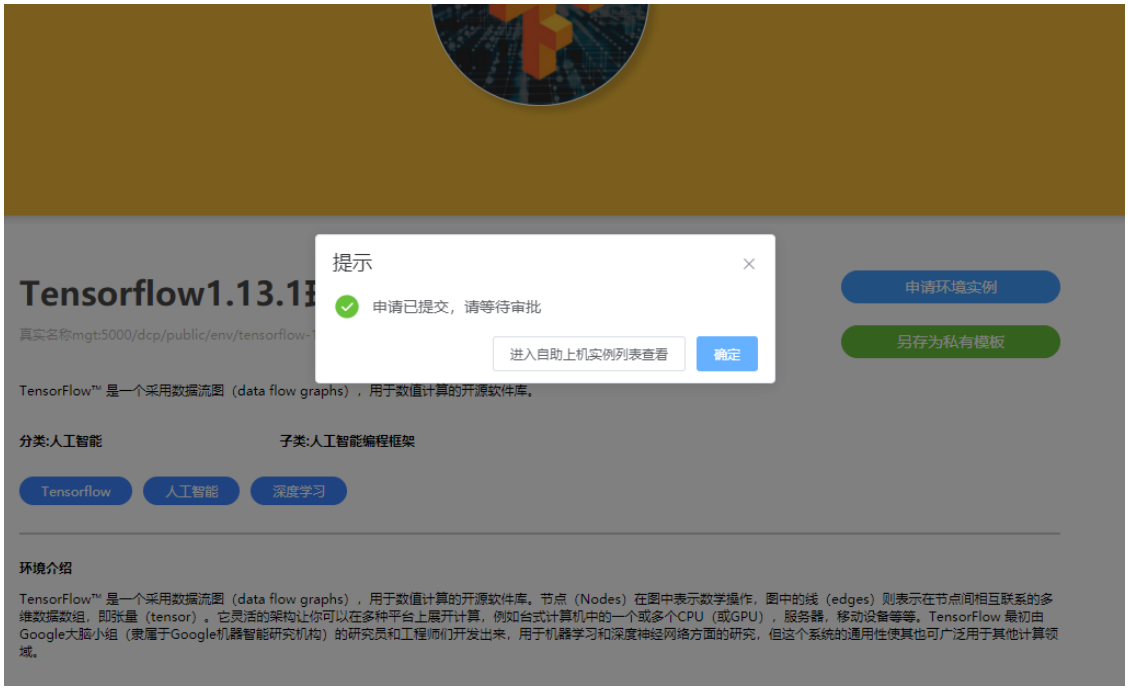

图 21

6.4. 查看申请结果

可以看到已经申请的环境。如果申请的资源满足系统设置的权限，则直接获得资源；如果超出系统设置，则需要通过管理员的审批；当系统资源不足时，该申请需要排队。

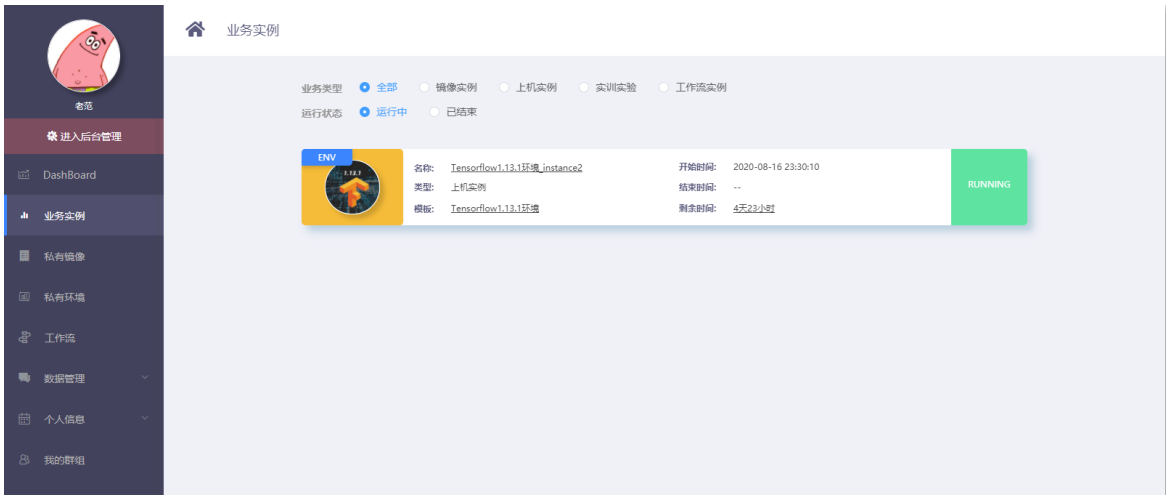

图 22

6.5. 进入实验环境

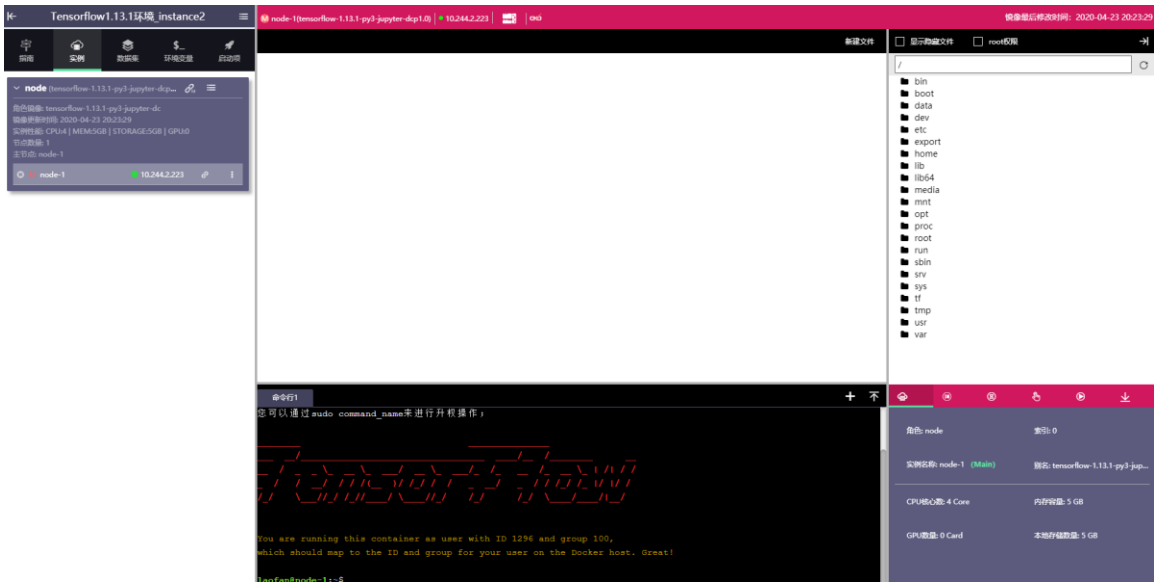

图 23

如下图，操作命令在“命令行”窗口中，默认其上方空白处可以用于编辑程序代码，不使用的话可将“命令行”窗口用鼠标上拉增大操作界面。

右侧 home 目录为“命令行”窗口所在操作系统的目录，存放于此目录内的文件，在实验结束后将会随开启的操作系统一起消失删除。

export 目录下的子目录为实训平台存储的共享目录，此目录下存放的文件在实验结束后仍会保留。

请根据需要选择文件目录，注意节省公用的共享目录空间。

6.6. 创建或上传文件：

在命令行中可使用 mkdir 等命令创建子目录文件夹或者文件。 右侧选择目录后，例如**********之后，右键菜单中也可以新建、上传、下载文件或者文件夹。

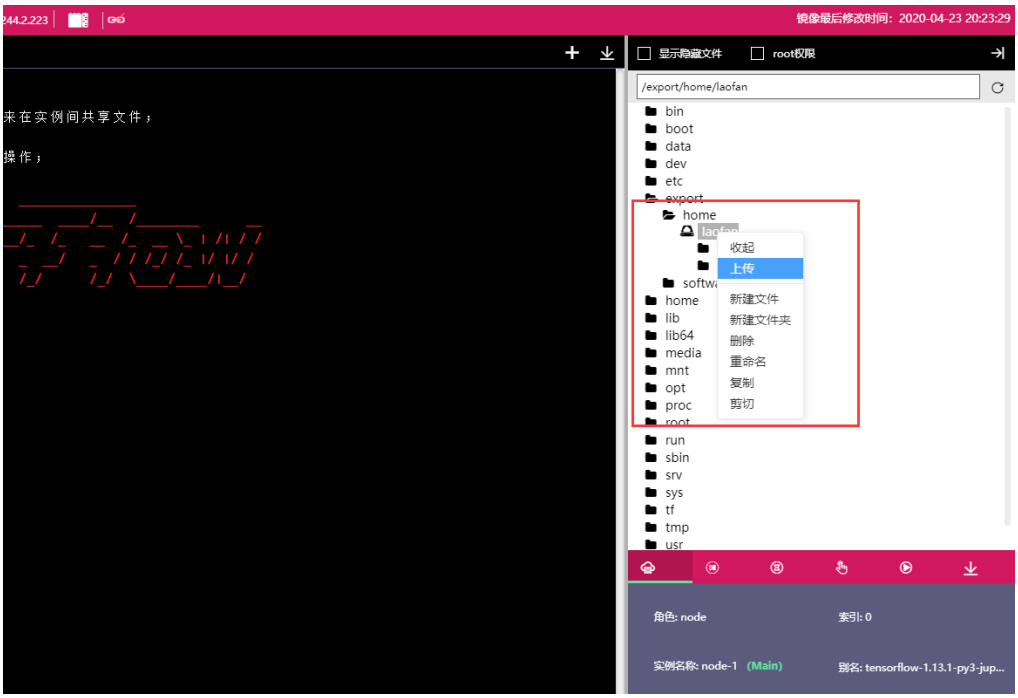

图 24

选择文件，然后浏览自己电脑上的输入文件目录，选择需要上传的文件。

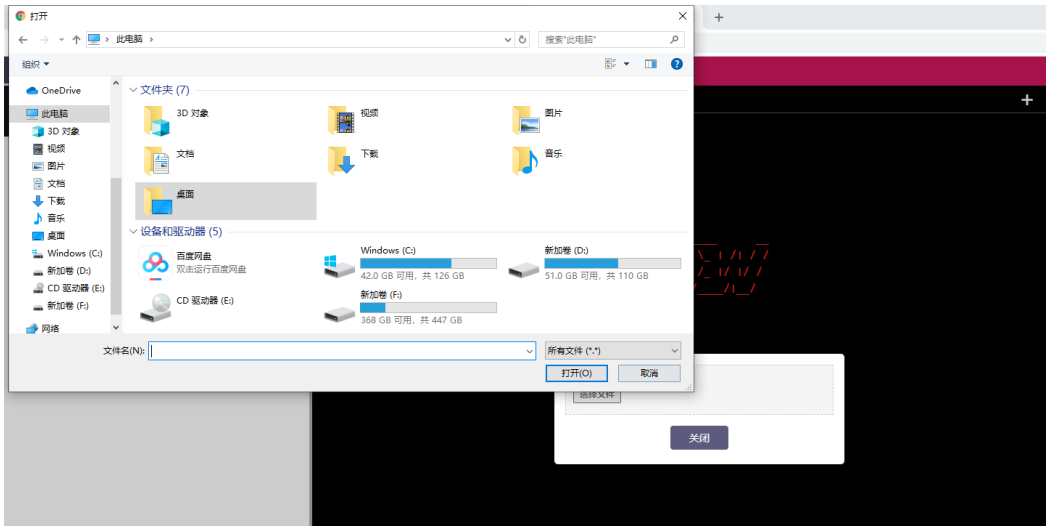

图 25

上传成功后，可以在命令行界面内看到自己的文件了，然后执行命令进行实验了。

6.7. 下载文件

待作业计算完成后，有两种下载文件的方式：

1） 选择右侧目录树位置，选择要下载的文件到自己的本地电脑上进行后期处理。

2） 在命令行界面利用 tar 命令压缩要下载的文件（例如：tar zcvf test.tar.gz test，将 test 文件夹压缩成 test.tar.gz），再单独下载压缩后的文件到自己本地电脑上进行后期处理。对于一次下载多个文件的情况，最好使用这种方式。

6.8. 结束实验

结束计算后，请在个人中心->自助上机位置，选择结束实验，即时释放系统资源，让其他用户得以排队使用。

7. 数据集

用户在进行实验的过程中，会存在对数据进行共享的需求。当数据比较大或者比较敏感时，传统的拷贝方式就显得不太合理，本系统提供了数据集功能，其基于数据得逻辑挂载为用户提供数据得共享途径，既满足了共享的需求，又满足了对数据权限可控的需求。

进入系统后，可以看到上方的“数据集”链接，如图所示，可以按照来源和分类对数据集进行筛选。“官方发布”是指平台统一发布所有人可见的，“个人发布”是指登录用户发布的数据集，“仅显示群组相关”是指与同组成员可见的共享数据集：

选择一个数据集打开，可以进行数据集的订阅：

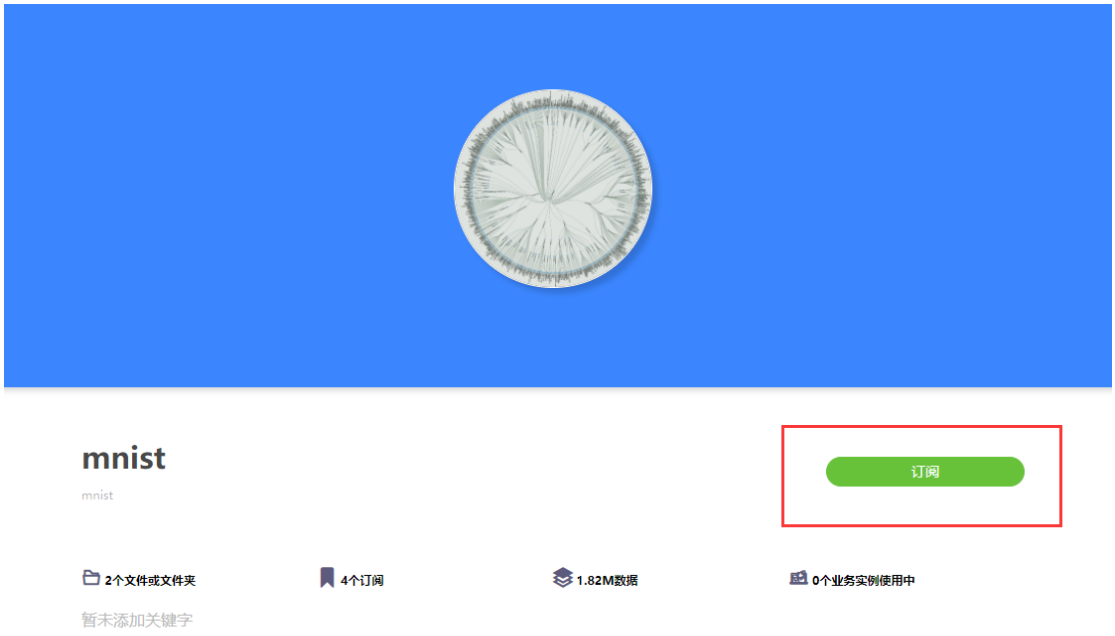

图 26

订阅后的数据集可以启动分析环境，或者在申请环境的过程中选择挂载：

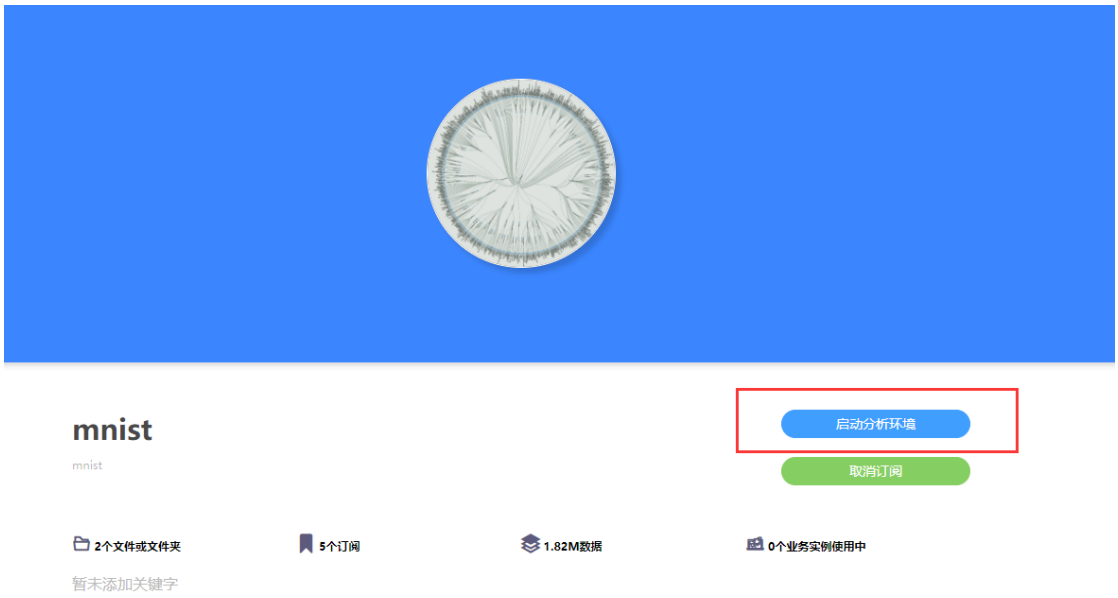

图 27

当个人需要创建一个数据集的时候，首先进入个人中心—数据管理—数据集，点击“新建数据集”，如下图所示，这里可以有几种选择：1）创建全新数据集：选择存储路径、数据集名称和简写，系统会按照设置创建一个新的目录；2）选择网盘子路径：如果数据已经存在，可以在这里直接选择；3）选择其他数据集子路径：如果数据集是在其他数据集的路径下，可以使用这种方式；4)手动输入路径：用户可以自定义数据集路径。

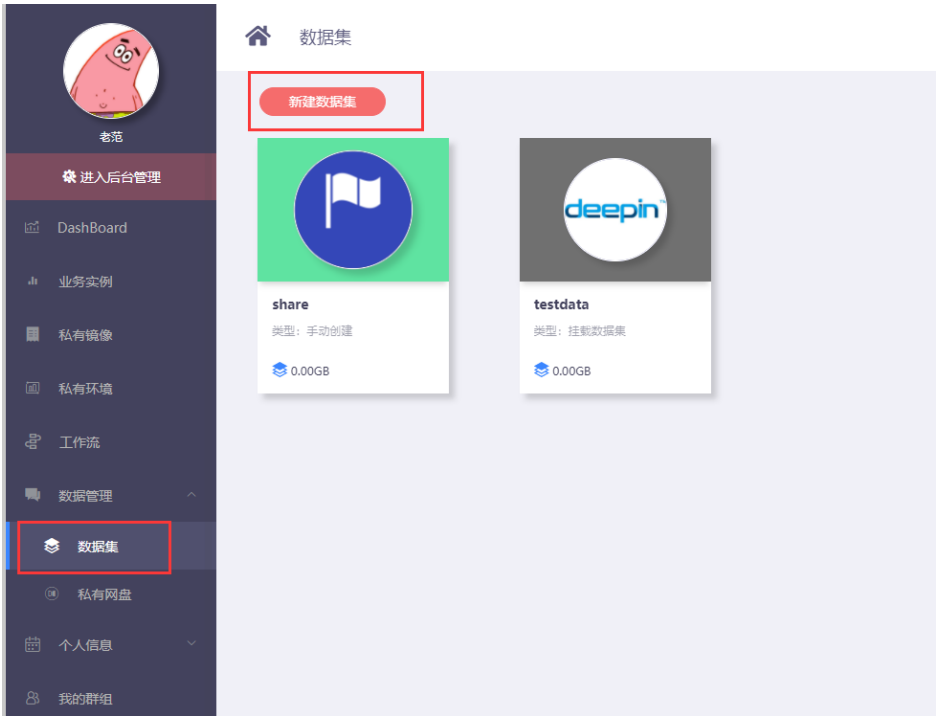

图 28

点击进入新创建的数据集：

图 29

“基本信息”一栏中可以修改数据集的名称、路径、分类等信息（如果需要修改数据集的权限，需要将数据集下架，修改后重新发布）。

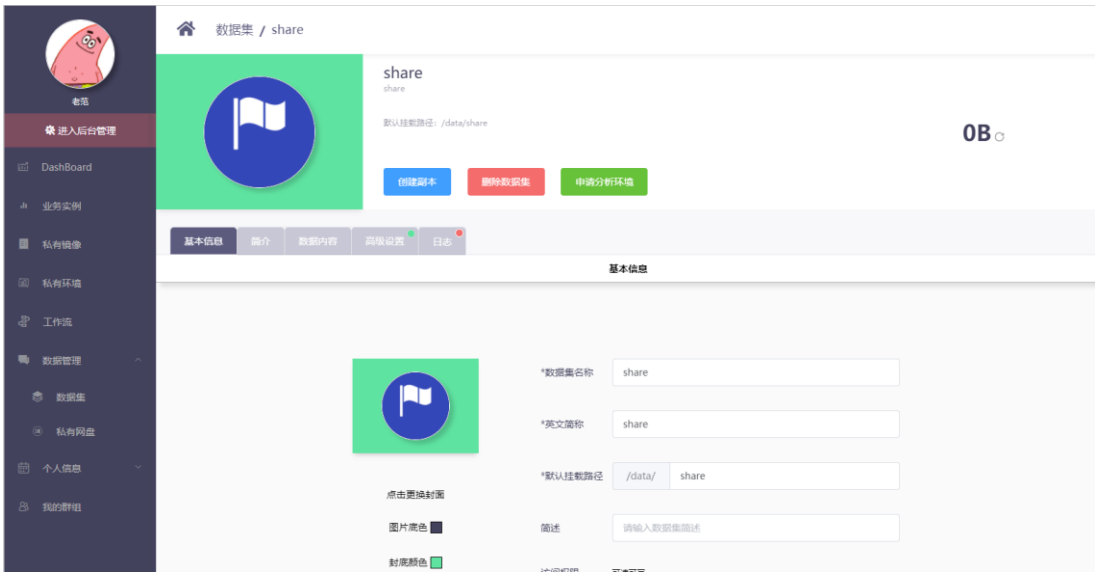

图 30

“数据内容”提供了可视化的数据管理方式，可以进行新建、删除、重命名、复制、剪切等操作：

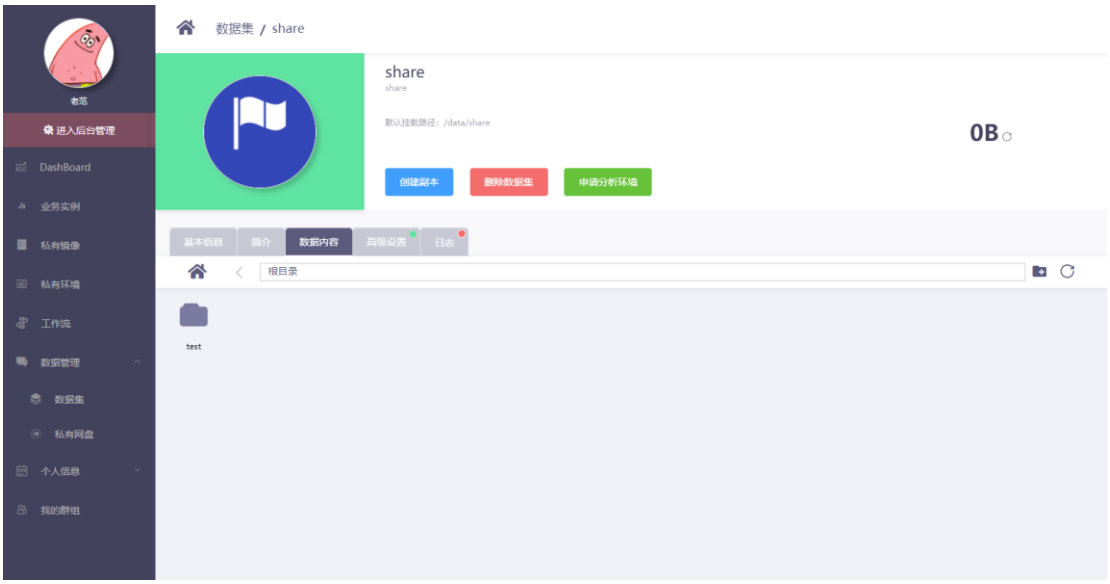

图 31

“高级设置”提供了数据的共享和发布操作。共享是指在指定群组的用户共享数据，发布是指对平台所有用户共享数据：

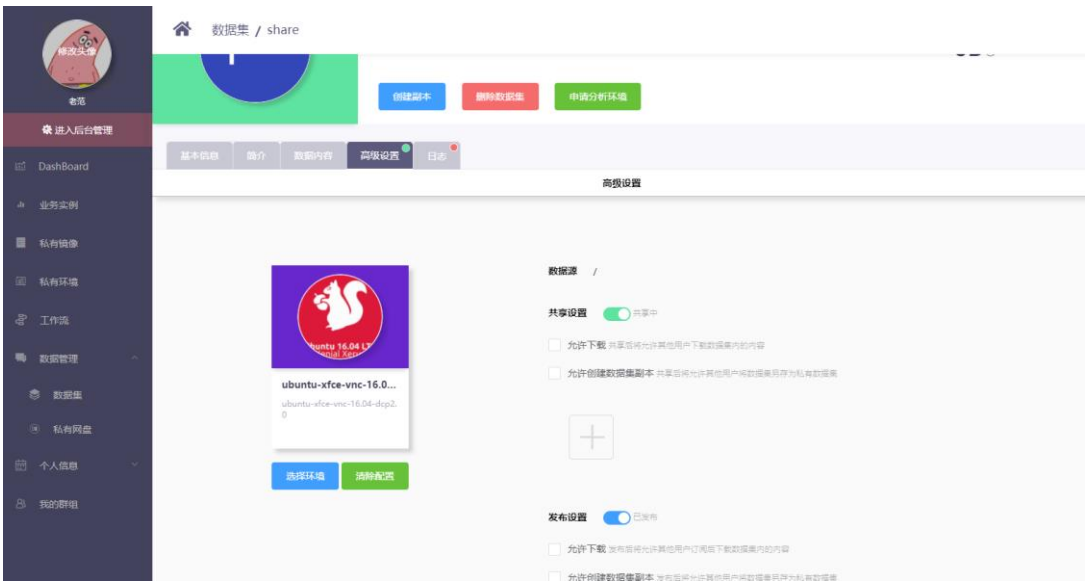

图 32

共享数据集时，需要首先指定群组：

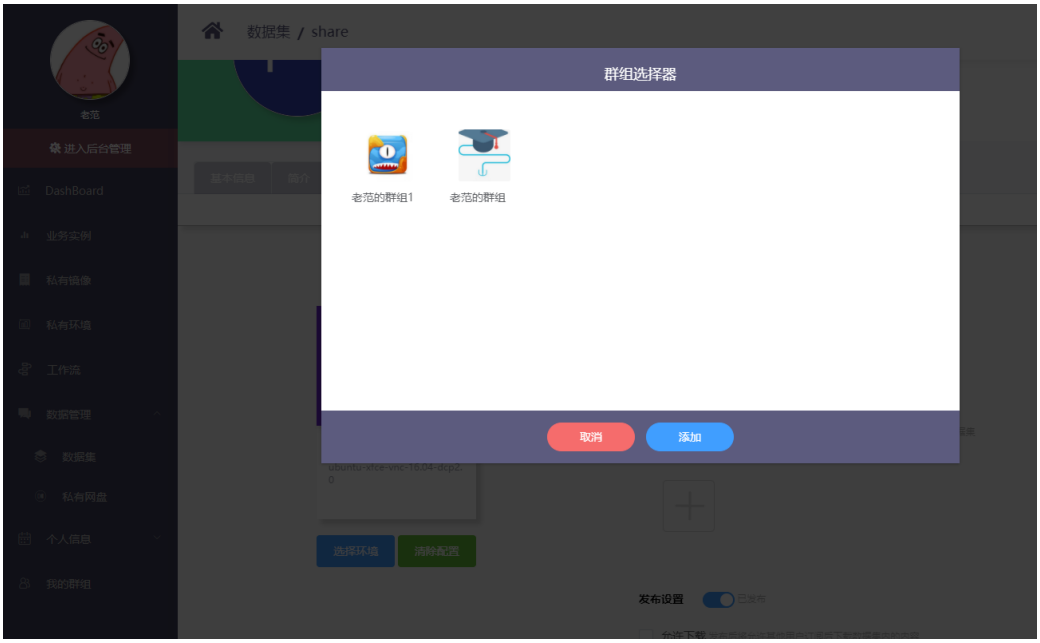

图 33

在个人中心—我的群组中，可以对群组进行设置：

图 34

8. 发布课程

在平台中，自助上机和实训实验业务都时由环境构建的。所谓“环境”，就是由镜像、计算资源、数据集、启动项等集成起来，为用户提供服务的整体内容。

要发布课程，首先需要建立私有环境（在建立私有环境之前，同时也需要将镜像和数据集准备好，镜像可以使用平台公共的，也可以使用私有的）。

创建私有环境有两种方法，一是基于平台已有的自主上机和实训实验，将其另存为私有环境模板，如图所示：

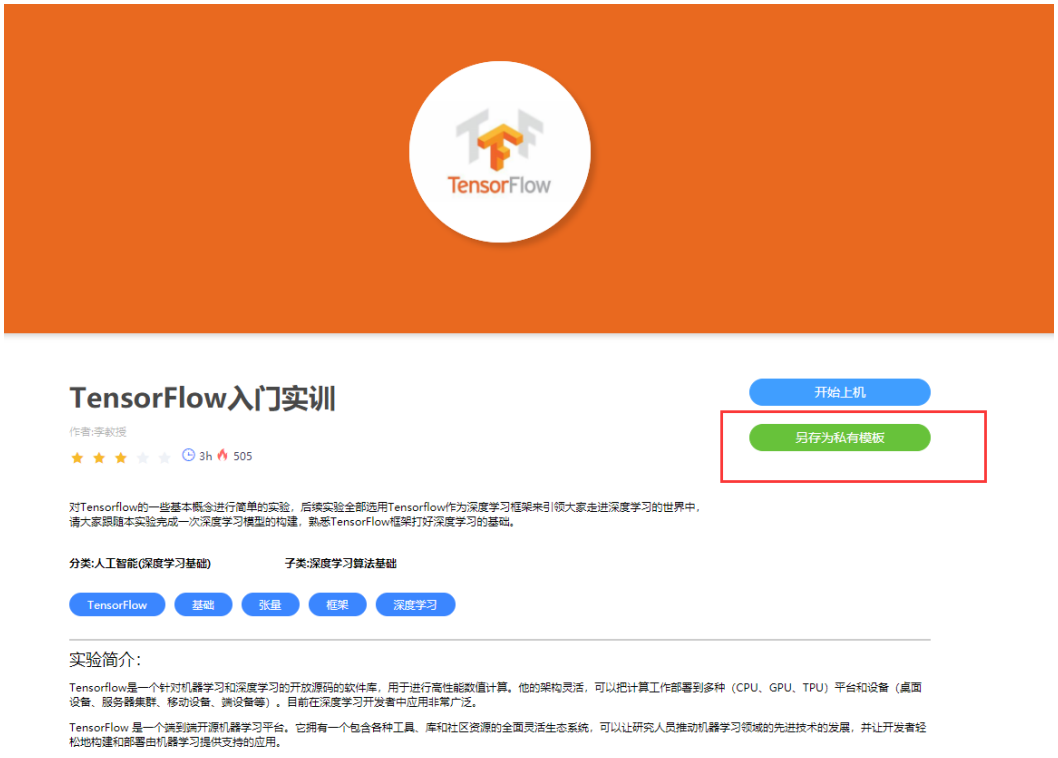

图 35

保存后的私有环境在个人中心—私有环境中可以看到：

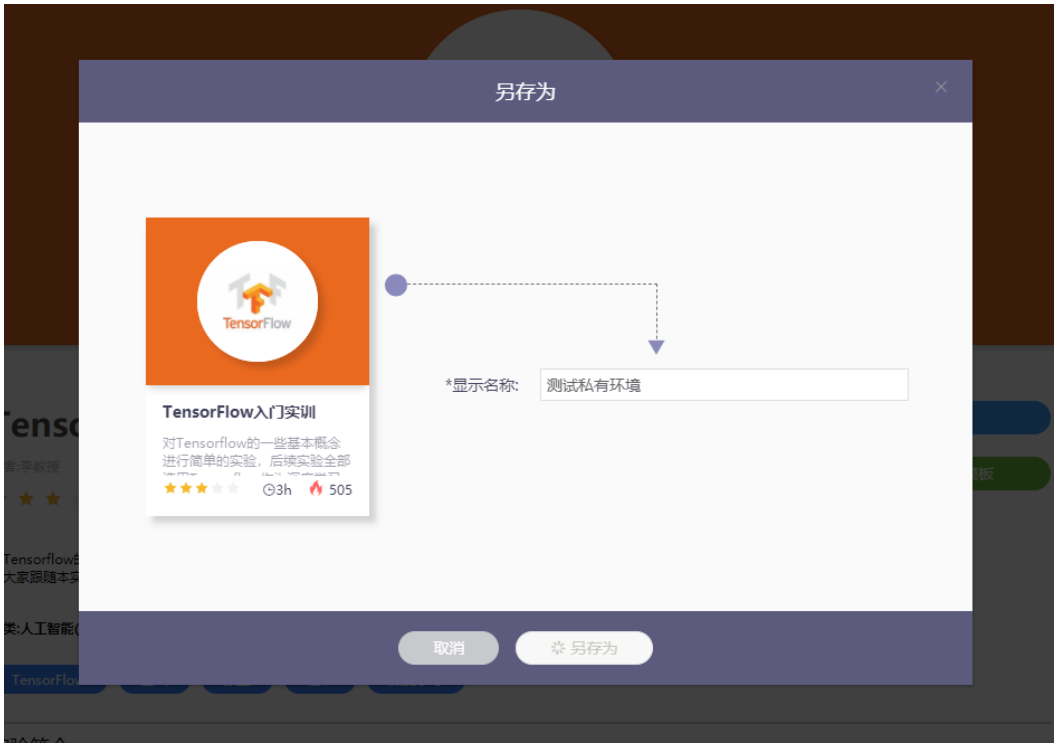

图 36

图 37

另一种方法是新建环境模板，进入个人中心—私有环境，选择新建环境模板：

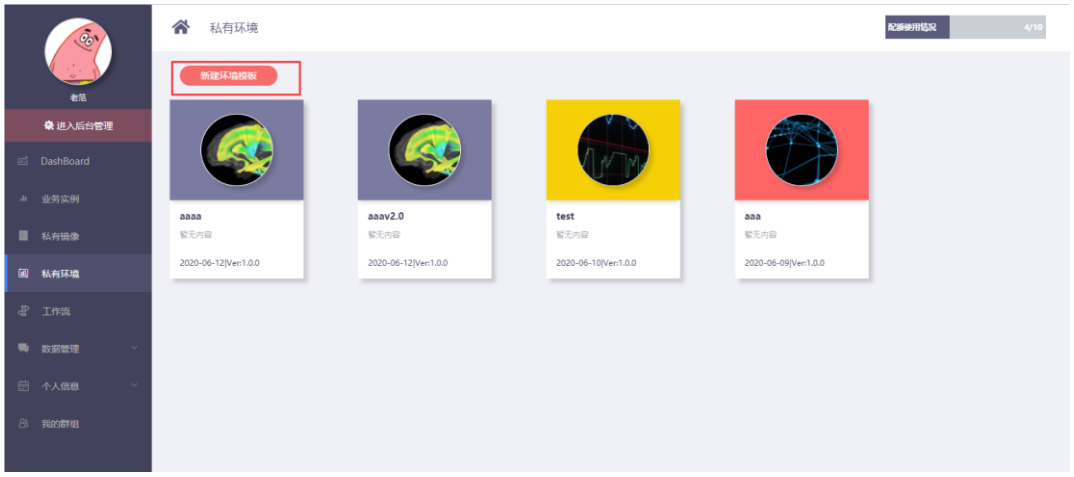

图 38

填写环境名称后，建立一个空模板：

图 39

进入新建的模板，这里可以查看模板的基本信息、简介、系统配置、共享配置、环境变量、启动项、数据集、访问端口等：

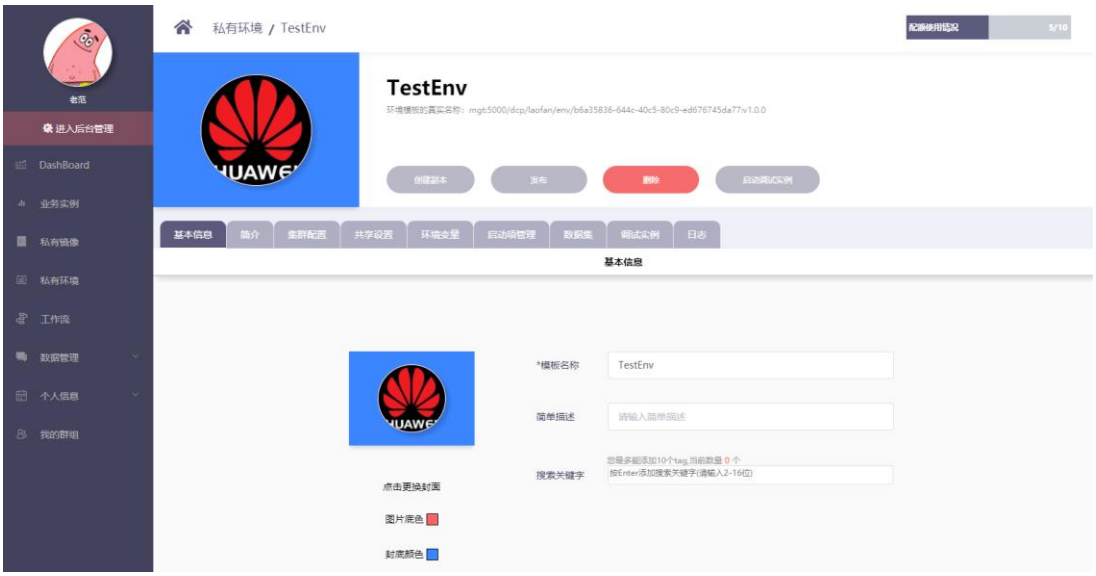

图 40

用户可以首先通过调试实例对空模板进行编辑：

图 41

启动调试实例时，可以忽略对资源的配置，直接点击“确定”，然后点击“连接实例”，进入调试界面。如下图所示，此时可以看到的时一个空白模板，UXSYS 界面中的几个区域都是空白的。左侧的菜单中，可以对实例、指南、数据集、环境变量、启动项和应用端口进行配置：

图 42

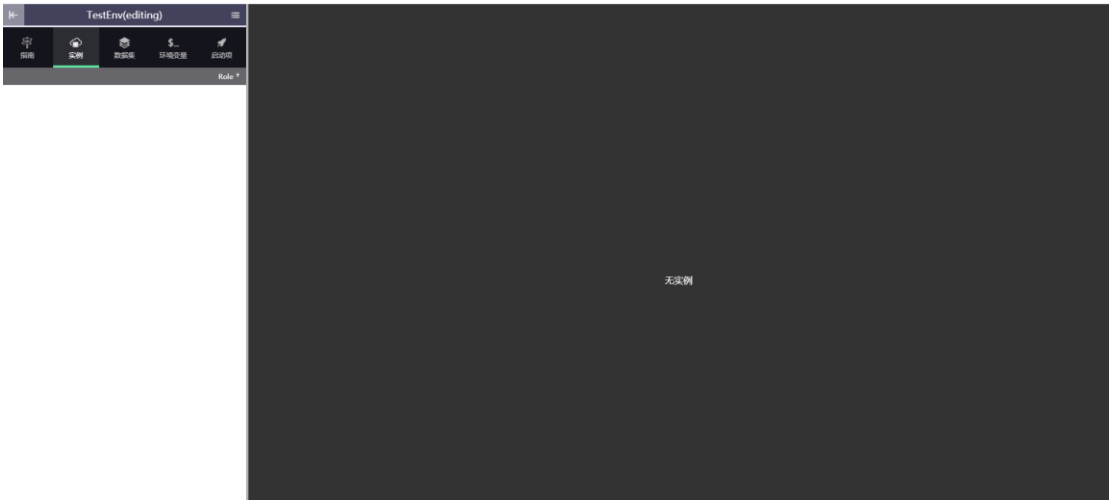

图 43

实例中可以配置环境使用的镜像和资源，每一个角色在环境中对应一个镜像，每个角色可以配置多个实例，其都对应相同的镜像模板（例如，在 Tensorflow 中，ps 是一种角色，worker 是一种角色），下图为配置一个 ps 角色实例和两个 worker 角色实例：

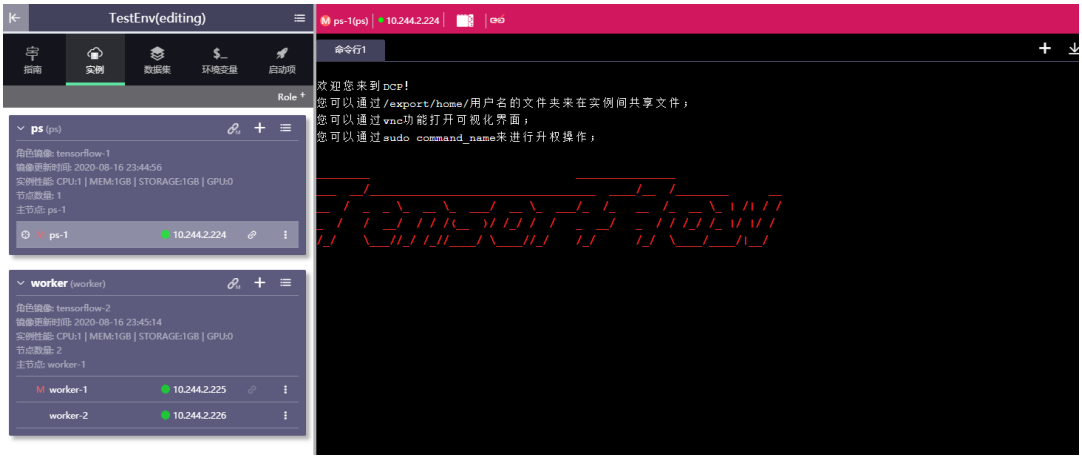

图 44

指南是为实训实验配置说明的入口，可以配置任务和上传教材。点击“+”添加任务，编辑任务内容。系统提供了可视化的编辑器，可以对文字、表格、图片、命令等多种元素进行编辑，如图所示：

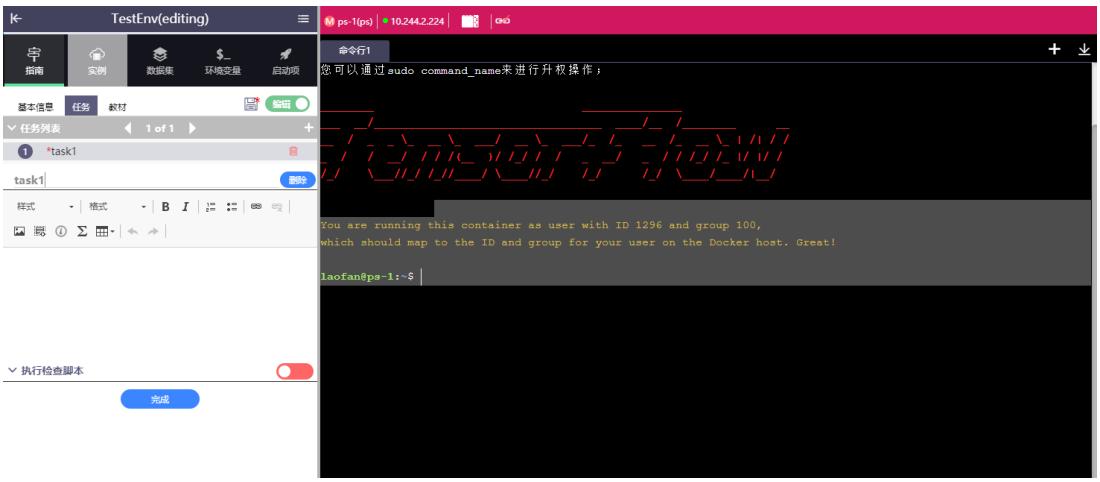

图 45

在环境启动和运行的时候，可能会使用到环境变量，例如，某个角色的实例启动时需要获取本机的主机名、IP 等信息，这些都可以使用环境变量来指定。环境变量分为内置环境变量和自定义环境变量。内置环境变量包括如下内容（可以启动一个实例查看），用户也可以通过自定义环境变量对名称、值等进行配置：

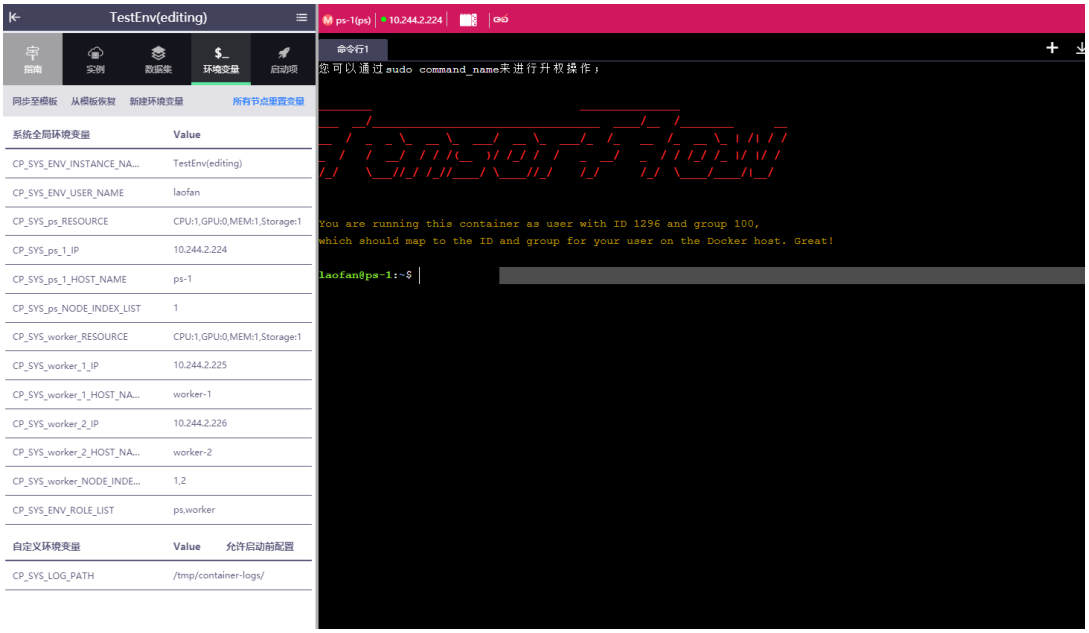

图 46

启动项是指实例在启动的时候需要加载的脚本，可以配合环境变量进行操作，例如：某个实例在启动时需要修改主机名和 IP 的关联文件，否则，由于实例每次启动时 IP 地址会发生变化，应用软件无法正常启动。这时，需要为环境

配置启动项，在实例启动的时候执行相应的指令。下图是配置 Slurm 作业调度系统的启动项：

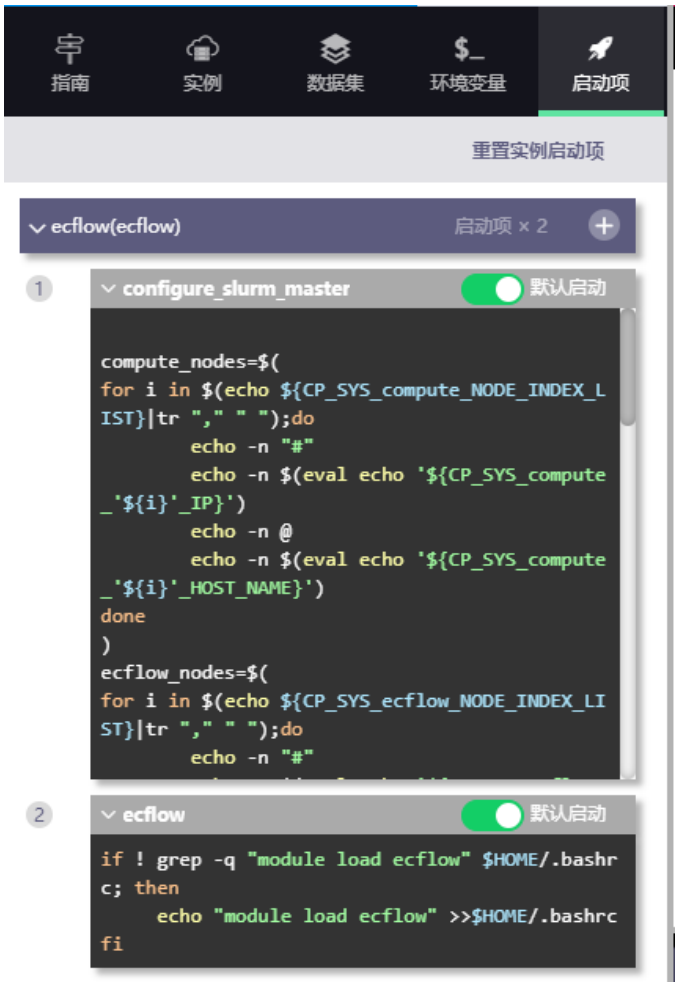

图 47

应用是指对应用端口进行配置。例如，需要为系统的 SSH 服务配置一个端口，让用户从管理节点的外网连接 IP 直接访问实例，此时可以按照下图的配置方法，首先创建应用，指定应用名称、实例、应用端口，保存配置：

图 48

然后打开 SSH 应用，选择激活，此时系统为 SSH 服务绑定一个端口，如下图所示：

图 49

基本信息和简介是环境模板展现给用户的内容介绍，包括模板名称、简述、搜索关键字、图标等：

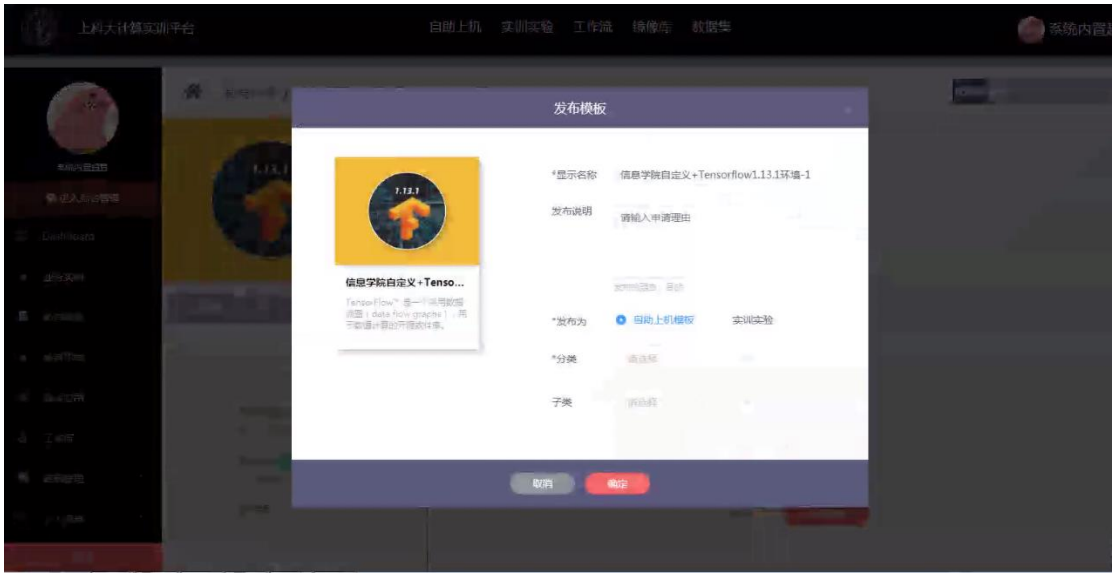

图 50

在共享设置中，可以将环境共享给特定的群组：

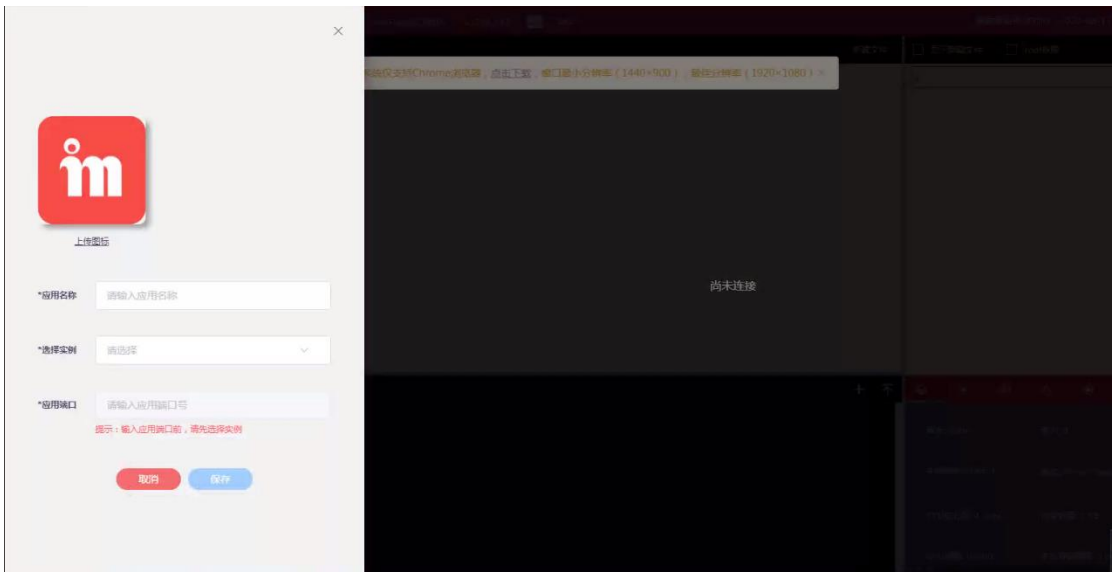

图 51

与数据集的属性一样，发布是指将环境发布到平台所有人可见：

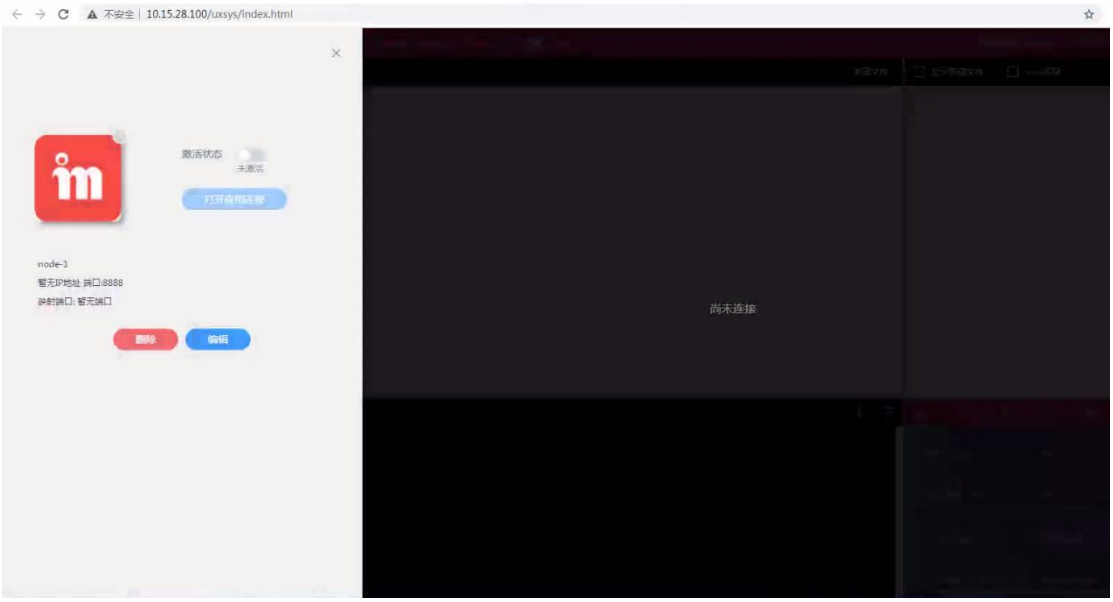
不安全 | 10.15.28.100/uxsys/index.html
激活状态
未激活
node-1
暂无IP地址 端口:8888
映射端口: 暂无端口
删除
编辑
尚未连接

图 52