

# NADC 科研云平台使用指南

## 1. NADC 科研云平台简介

NADC 科研云平台是面向大数据时代的在线科研环境。平台通过互联网连接跨地域的资源节点，形成统一的科学计算系统。科学平台的资源节点以天文科学数据为中心，配备相应的计算资源及软件环境。平台支持科学用户通过浏览器访问平台内可公开访问的数据、存储及计算资源。科学用户可根据科研任务需求，包括计算、数据规模等，选择合适的资源节点，创建交互式应用或提交批量计算任务，实现面向大规模科学数据的在线科研，有助提高科研效率。

NADC 科研云平台现有国台 VEGA 集群(VEGA)、阿里云 ECS 集群(ALIYUN)、紫金山天文台云资源(PMOVIM)三个资源节点，其中 VEGA 集群包含 38 台 CPU 服务器(2432 核)，一台 GPU 服务器(8 卡 A40)以及一台胖节点服务器(内存 2TB)，是平台现有的主要计算资源。平台主界面如图 1 所示。

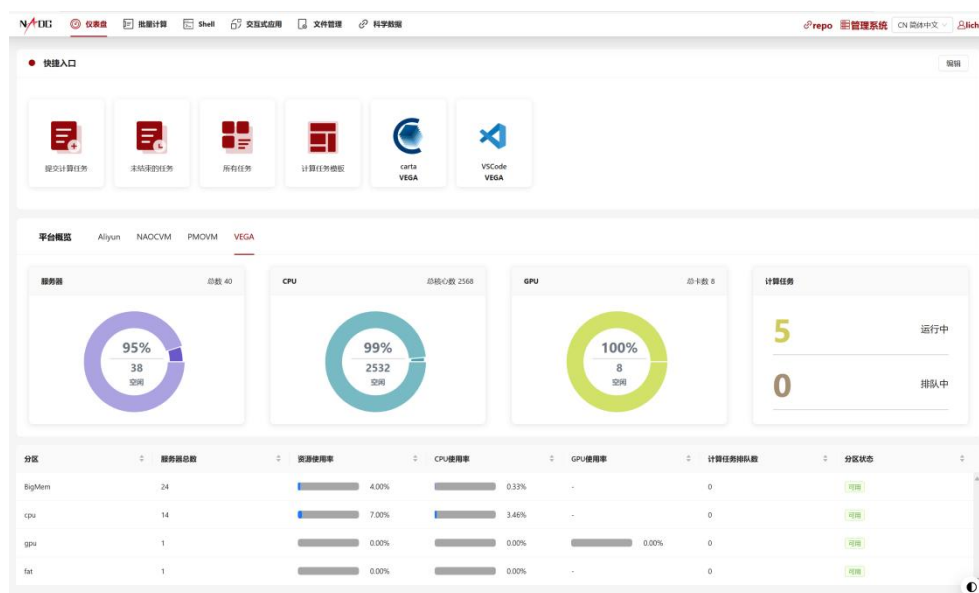


图 1：科学平台主界面

## 2. 平台帐号申请

(1) 科研云平台用户采用与 NADC 统一认证系统同样的用户信息,但是由于计算平台涉及后台 LINUX 系统的操作,用户名称需要符合 LINUX 系统用户名的命名规范,因此,NADC 统一认证系统的用户需要设置短用户名后,方可登录计算平台。设置方法:登录 NADC 统一认证平台

(<https://oauth.china-vo.org/>),进入我的工作台,点击左侧短用户名按钮,然后在右侧的短用户名输入框中输入短用户名,点击提交。如图 2 所示:

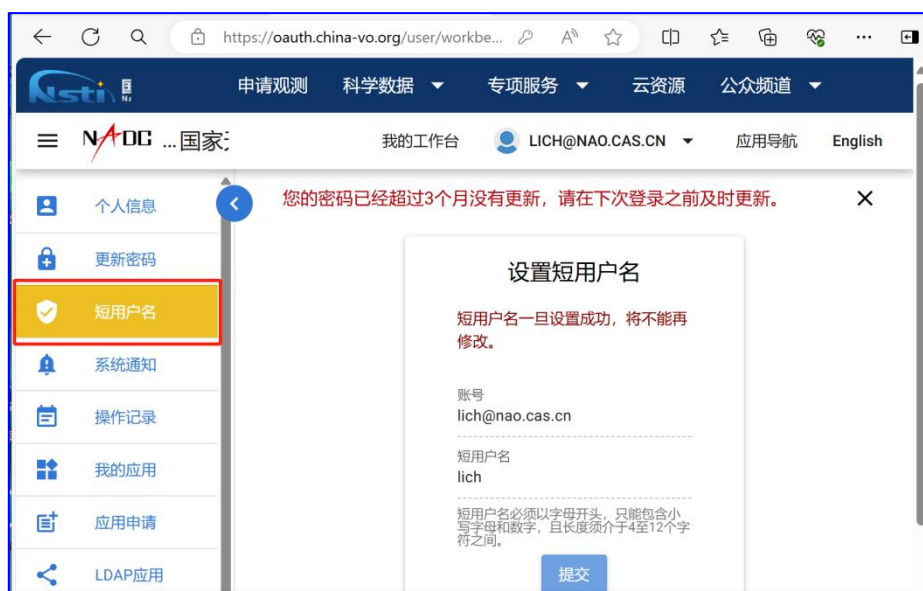


图 2: 短用户名设置

(2) 对于没有设置过密码的用户,同时也需要重新设置一个登录密码。设置方法如图 3 所示:

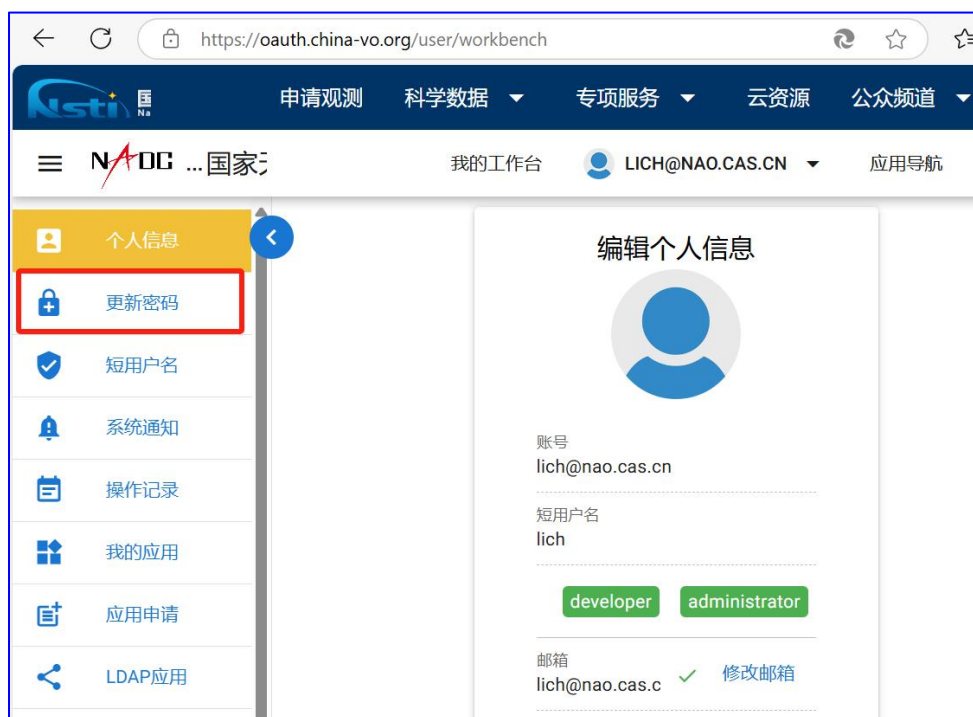


图 3：设置或更新密码

### 3. 资源申请

为了保证计算平台的安全性,NADC 平台用户需要提交资源使用申请表并得到平台管理员的审核通过后,方可登录计算平台。当登录提示出现如图 4 所示界面时,点击 application 链接,便可在线提交申请。

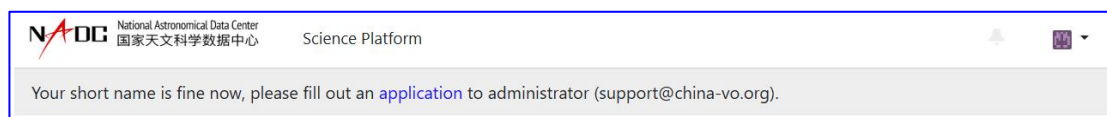


图 4： 提交帐号开通提示

### 4. 平台登录

平台登录 URL: <https://science.china-vo.org/>

输入此地址后,默认将进入到 NADC 的统一认证登录界面,输入 NADC 的帐号及密码即可完成登录,并自动进入到科学平台主页。

同时,平台保留了单独的登录界面,如果上述登录不成功,将进入到平台独立的认证界面,如图 5 所示。



图 5：平台自定义登录界面

此时，用户名为 NADC 统一认证系统中设置的短用户名，密码跟统一认证系统的登录密码一致。

此外，为了方便习惯于使用 SSH 工具进行操作的用户，部分节点也支持用户从内网直接通过 SSH 进行登录，登录用户名及密码与 WEB 端一致。。

VEGA 节点的登录服务器 IP： 10.3.15.142， 采用 MobaXterm 工具示例如图 6 所示，登录后可以提交作业及文件传输。

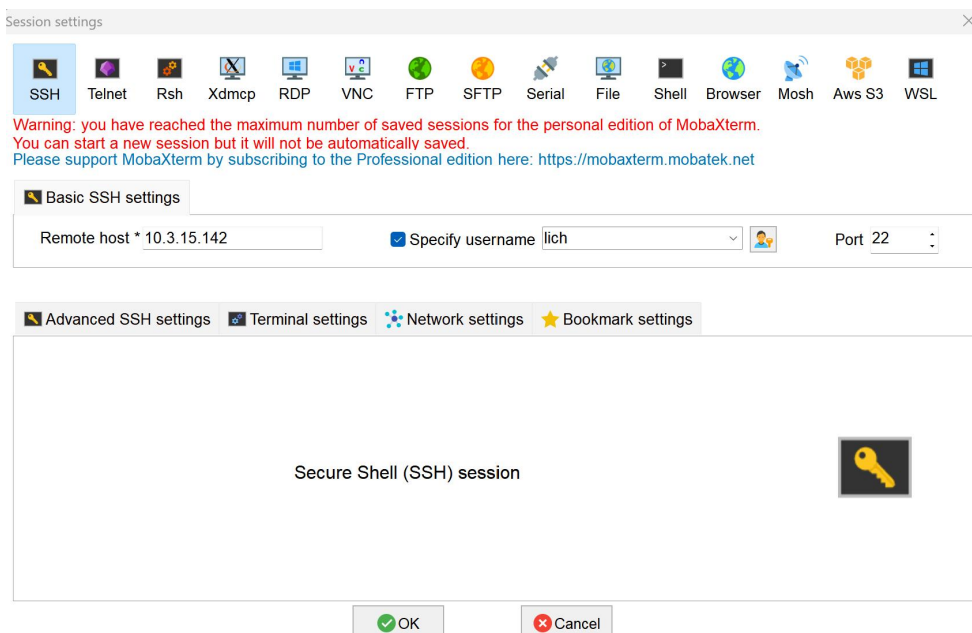


图 6： MobaXterm 登录到资源节点的示例图

## 5. 个人环境设置

完成帐号注册后，用户在平台各资源节点上就拥有了计算帐号，可以通过交互式应用、批量计算提交计算任务，或通过 Shell 直接 SSH 到资源节点的登录服务器上。

在资源节点上，已经部署了 conda 包管理工具，并安装了 python 3.7，用户通过 webserv 登录后执行 `conda activate python3.7` 即可激活 python3.7 的 python 环境，此环境下已安装了大部分数据处理与机器学习相关的库文件，如果用到的库不存在，用户也可自行安装。

## 6. 创建在线应用

登录后，点击交互式应用，在下拉菜单上选择相应节点后将显示二级菜单。二级菜单包含已创建应用和创建应用两个选项。点击已创建应用时，将显示在此计算集群上已经创建的应用列表。如图 7 所示：

点击创建应用时，将显示此计算集群上已部署的应用，如图所示：



图 7： 资源节点上部署的科研应用

点击某一应用图标后，即可进入应用配置界面。

配置项依赖于具体集群的配置，请参考集群配置信息。

填完相关信息后，点击提交，平台将在指定节点上部署应用，并自动跳转到已创建应用列表。部署完成后，应用列表项将出现连接按钮，点击连接即可进入到新创建的应用界面。

## 7. 提交计算任务

平台资源节点以 slurm 作为资源调度的基础软件，支持 MPI 跨节点并行计算。提交计算任务可以采用 Shell 命令，也可以通过 WEB 功能界面来提交。命令行界面下可以直接运行 slurm 的相关命令，如查看队列情况（squeue），查看资源总体情况（sinfo），提交计算任务命令（sbatch）。关于 slurm 软件的详细使用，请参考文件 slurm.pdf

如果对 slurm 命令不熟悉，也可以直接通过批量计算功能界面来提交计算任务。提交计算任务如图 8 所示：

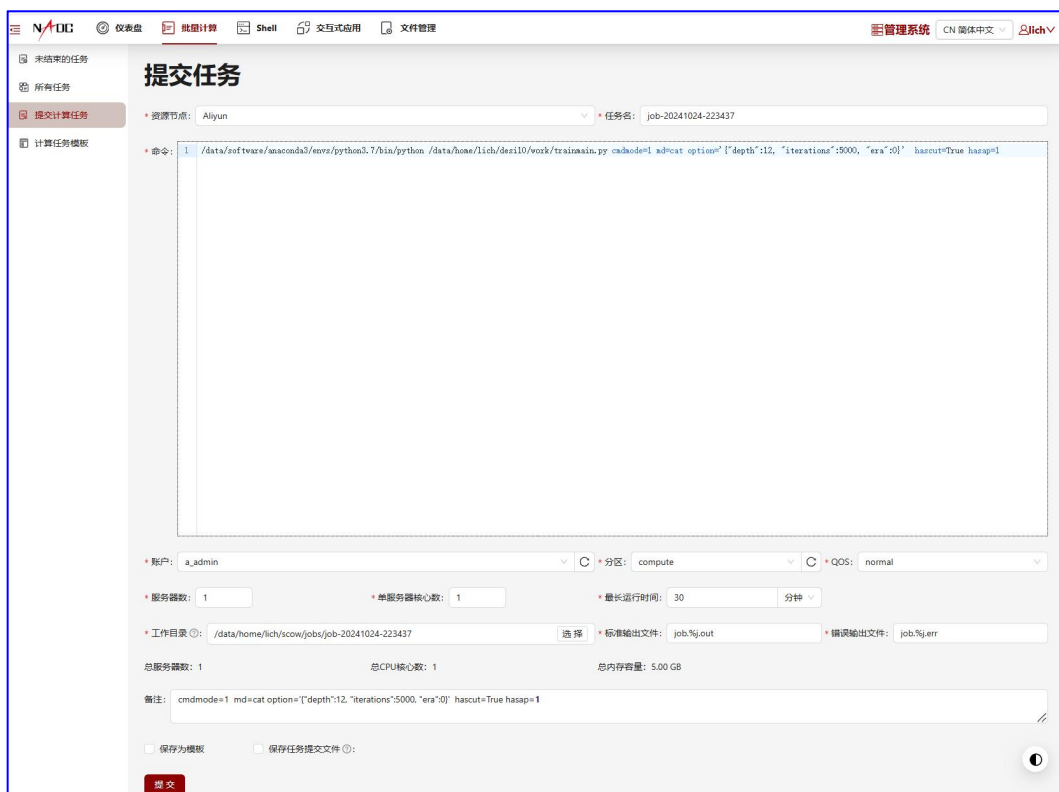


图 8：提交计算任务主界面

提交任务可以定义任务的名称，指定执行任务的资源节点。执行的命令写在命令输入框中，一条命令一行。分区及 QOS 指资源节点的队列及优先级，在 slurm 服务端配置时定义。服务器数为本计算申请执行的服务器数量，默认为 1，若非 MPI 作业，一般无需修改。单服务器核心数为单台服务器上执行本任务的 CPU 核数。**最长运行时间**指申请执行本任务的最长时间，这个需要用户根据任务情况适当设置，如果不太了解，可以适当设置长一些，以免因时间到期导致作业中断。全部输入完成后，点击提交即可将计算任务提交到指点计算节点上进行计算。

## 8. 天文数据集挂载

平台集成了天文数据中心已汇交的大量已公开天文科学数据集。在科学数据/数据列表栏目浏览并选择自己需要使用的数据，如图 9 所示。

