WPPCC-基因赛道

使用说明文档

**队长**

韩培义

**队员**

段少明、郭蕴哲、靳小鹏

（按姓氏拼音排序）

2022年1月

# Docker 镜像导入

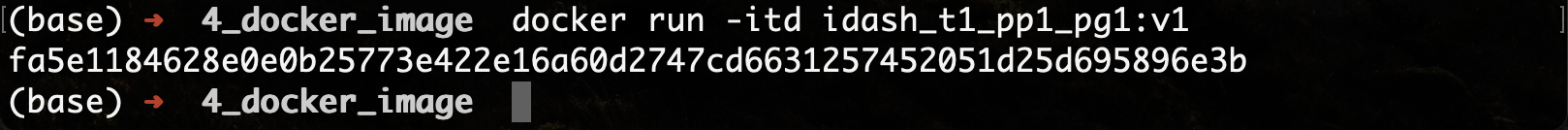
镜像文件名idash\_t1\_pp1\_pg1\_v1.1.tar.gz，(md5sum: 6ef0865066c66cefc07b1f76e807db74)，使用命令 docker load < idash\_t1\_pp1\_pg1\_v1.1.tar.gz 导入 docker中，如下图所示（更新版本中，docker的版本号变为了v1.1）：

一張含有 文字 的圖片

自動產生的描述

# 运行Docker 镜像

使用命令docker run -itd idash\_t1\_pp1\_pg1:v1.1运行docker镜像，如下图所示：



根据返回的ID，使用docker exec -it {容器ID} bash进入容器，直接进入了工作目录，可以看到数据以及程序文件，如下图所示：

一張含有 文字, 螢幕, 靠近 的圖片

自動產生的描述

# 运行第一个程序

使用命令 python start.py 运行程序，程序第一次运行时会自动解压数据集到 data/ 目录，运行过程中程序会提示时间、运行步骤、操作执行方等信息，如下图所示：

一張含有 文字 的圖片

自動產生的描述

程序执行的具体提示信息，如下图所示：

一張含有 文字 的圖片

自動產生的描述

程序执行结束后，会给出输出的NJ树的相对路径，如下图所示（后续可通过docker cp 命令拷贝到docker外部）：

一張含有 文字 的圖片

自動產生的描述

# 切换数据集

Start.py 支持命令行传入参数，可以使用 python start.py L1 使用L1数据集，将上述命令的L1替换为L2或者L3，可以切换L2或L3数据集，如下图所示：

一張含有 文字, 裝置 的圖片

自動產生的描述

在上一章节中，L0数据集是我们为了快速验证程序整个流程可用，使用L1数据集剪裁的，当Start.py没有命令行输入时，默认执行L0数据集

# 目录说明

整个docker镜像中的文件和目录如下图所示：

一張含有 文字, 螢幕, 螢幕擷取畫面 的圖片

自動產生的描述

对目录结构进行说明：

* **202111-data.zip**：为大赛下发的数据文件
* **A\_local**：甲方本地存储的变量，包括同态加密所用的公钥和解密私钥；
* **A\_to\_B**：在整个计算过程中，甲需要发给乙的变量；
* **B\_to\_A**：在整个计算过程中，以需要发给甲的变量；
* 可以通过衡量以上两个文件的大小，计算整个计算过程中的通信量，执行新任务时候，会自动删除上次任务中，在以上三个目录中遗留的文件；
* **clean\_up.sh：**执行该脚本可以删除程序执行中产生的文件，但不会删除生成的NJ树；
* **config.py**：包含了本程序可配置的变量信息；
* **data**：存放解压之后的数据；
* **output**：存放程序输出的NJ树文件；
* **start.py**：程序入口；
* **utills**：存放除了start.py和 config.py 之外的程序源码；

# 改变其他运行参数

由于需要适配多个数据集，所以我们使用比例的方式控制M和N变量，需要在config.py 中进行修改，如下图所示：

一張含有 文字 的圖片

自動產生的描述

我们在 docker中安装了vi和emacs文本编辑器，便于大赛评委编辑；如需自定义基因编码或each\_cnt变量，相关变量也在config.py定义。