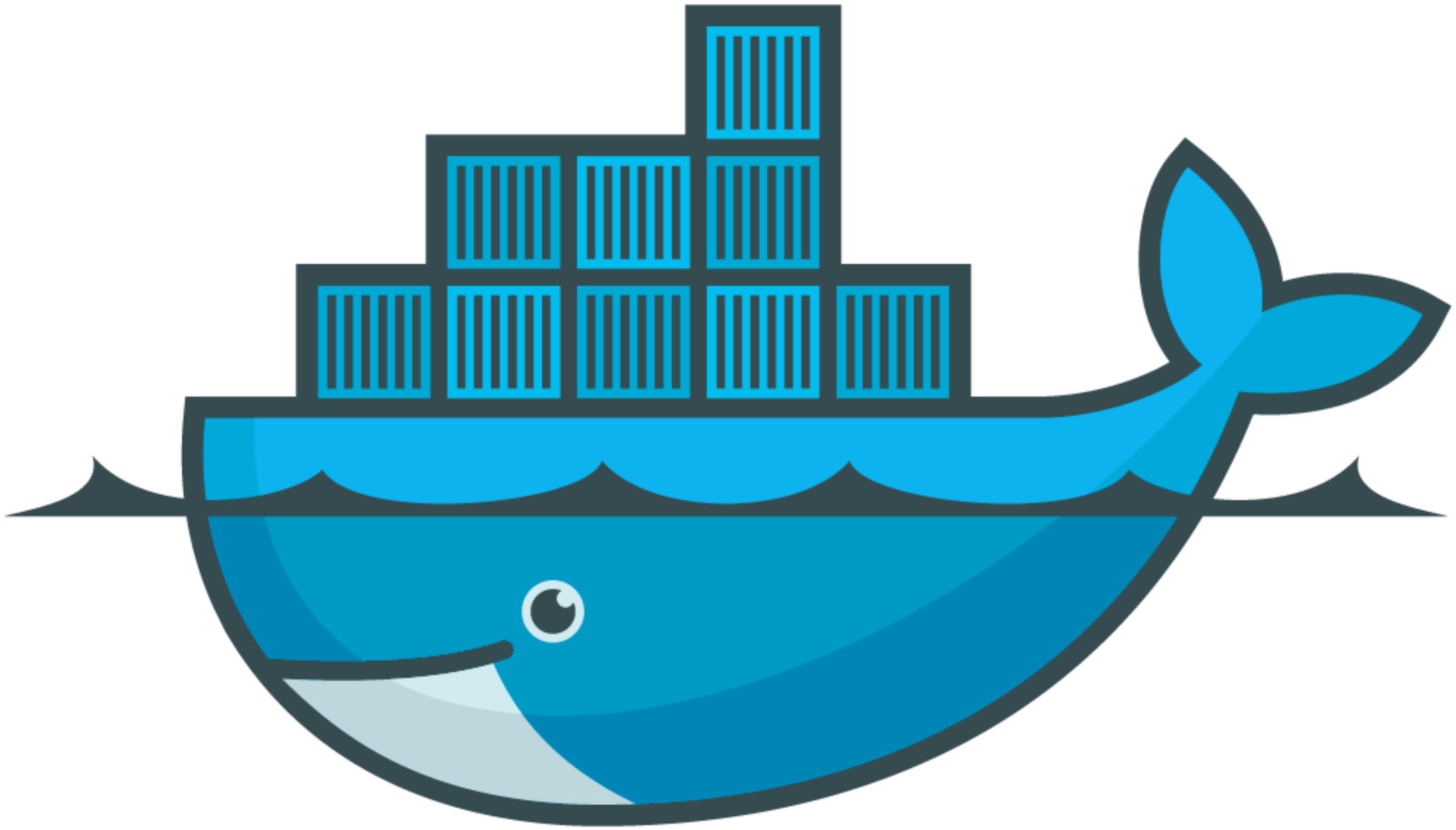
Docker for teaching



Naupaka Zimmerman, June 2019

Goal

Students will be able to log on to shared remote infrastructure to perform analyses using both command-line (bash) and graphical interfaces (RStudio, Jupyter Notebook)

Problem:

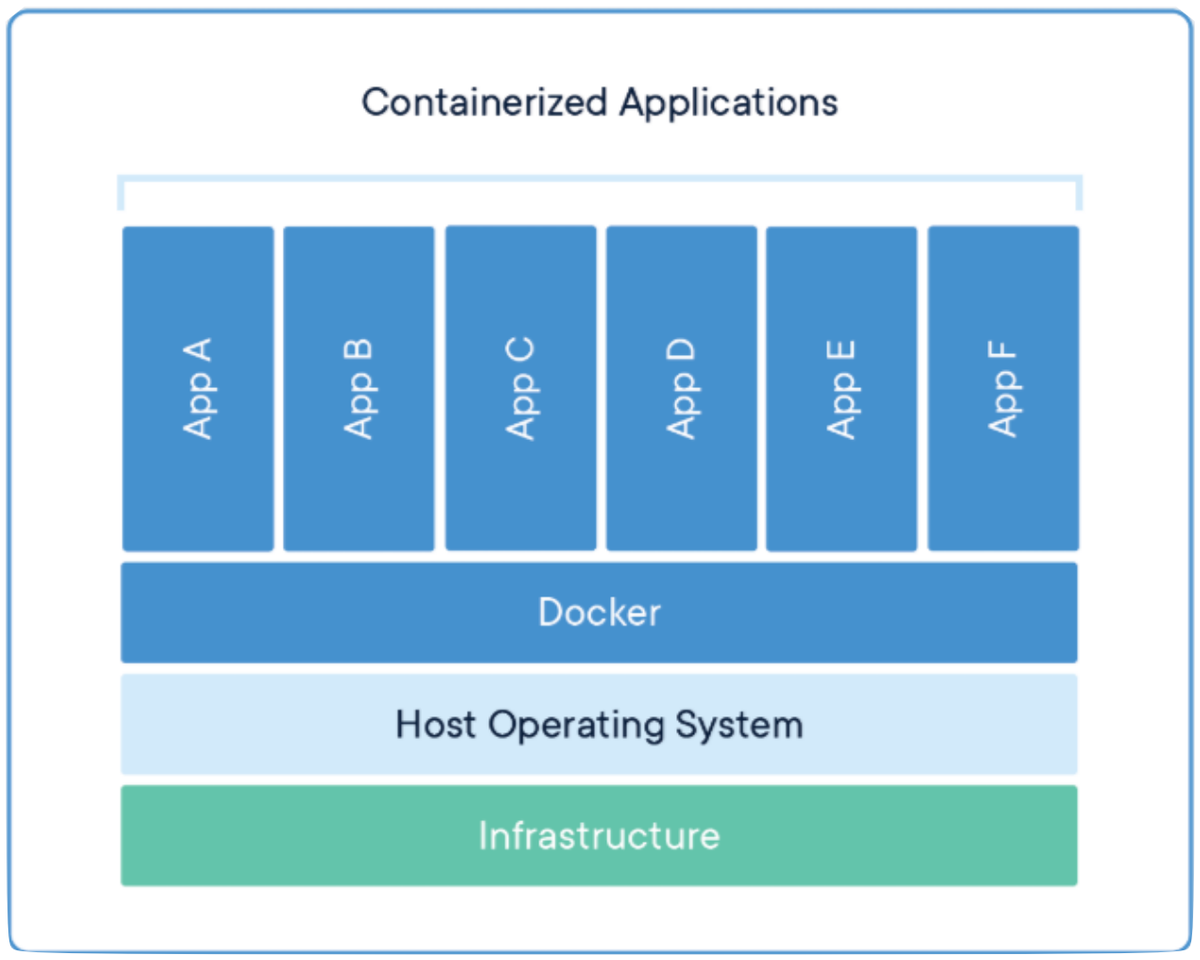
It is difficult to manage installation, updates, and maintenance of computational infrastructure for classes involving novice learners

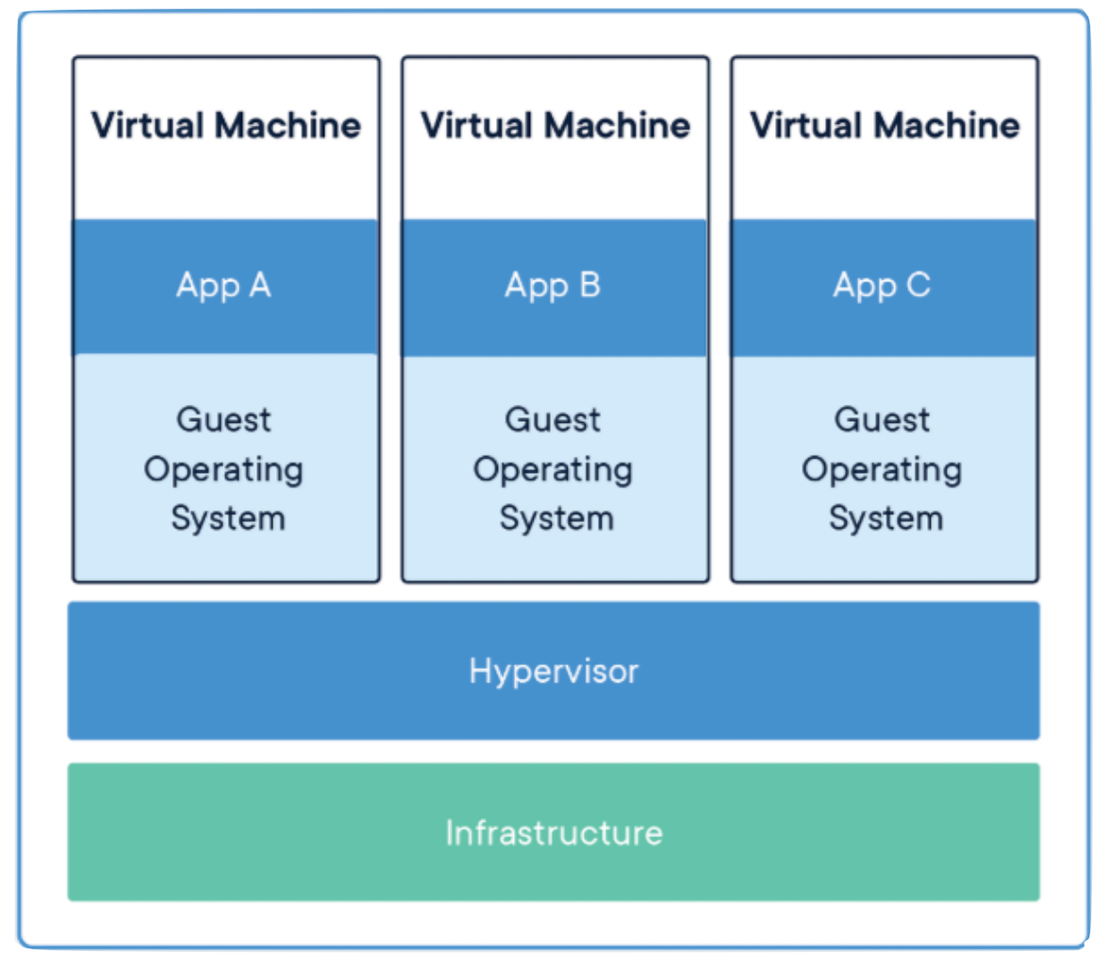
Problem:

It is difficult to manage installation, updates, and maintenance of computational infrastructure for classes involving novice learners

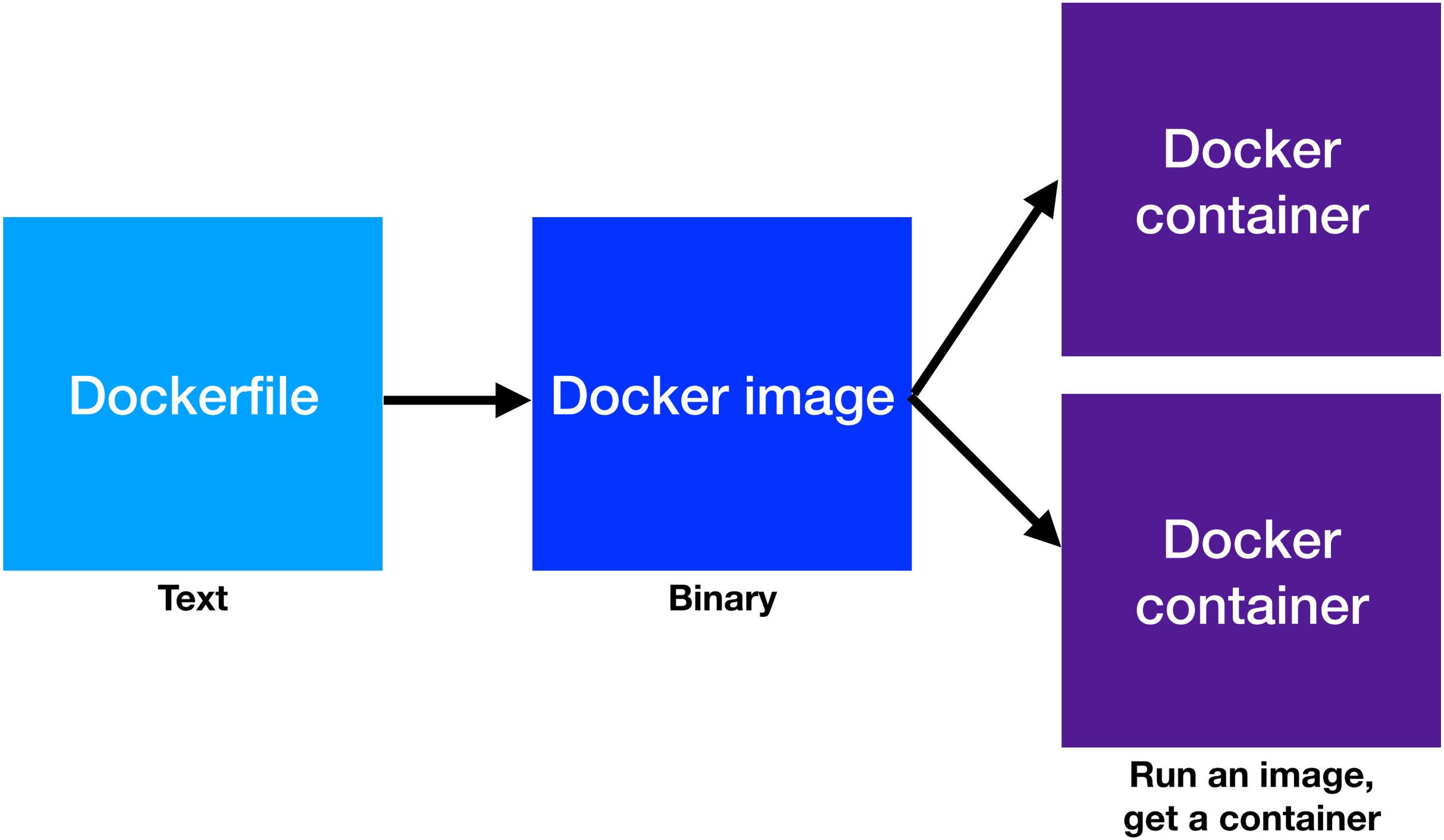
computational resources are finite
installation of necessary libraries and applications
different environments for different assignments
instructor access to help troubleshooting
want isolated filesystems (e.g. chroot)
scriptability

What is Docker?

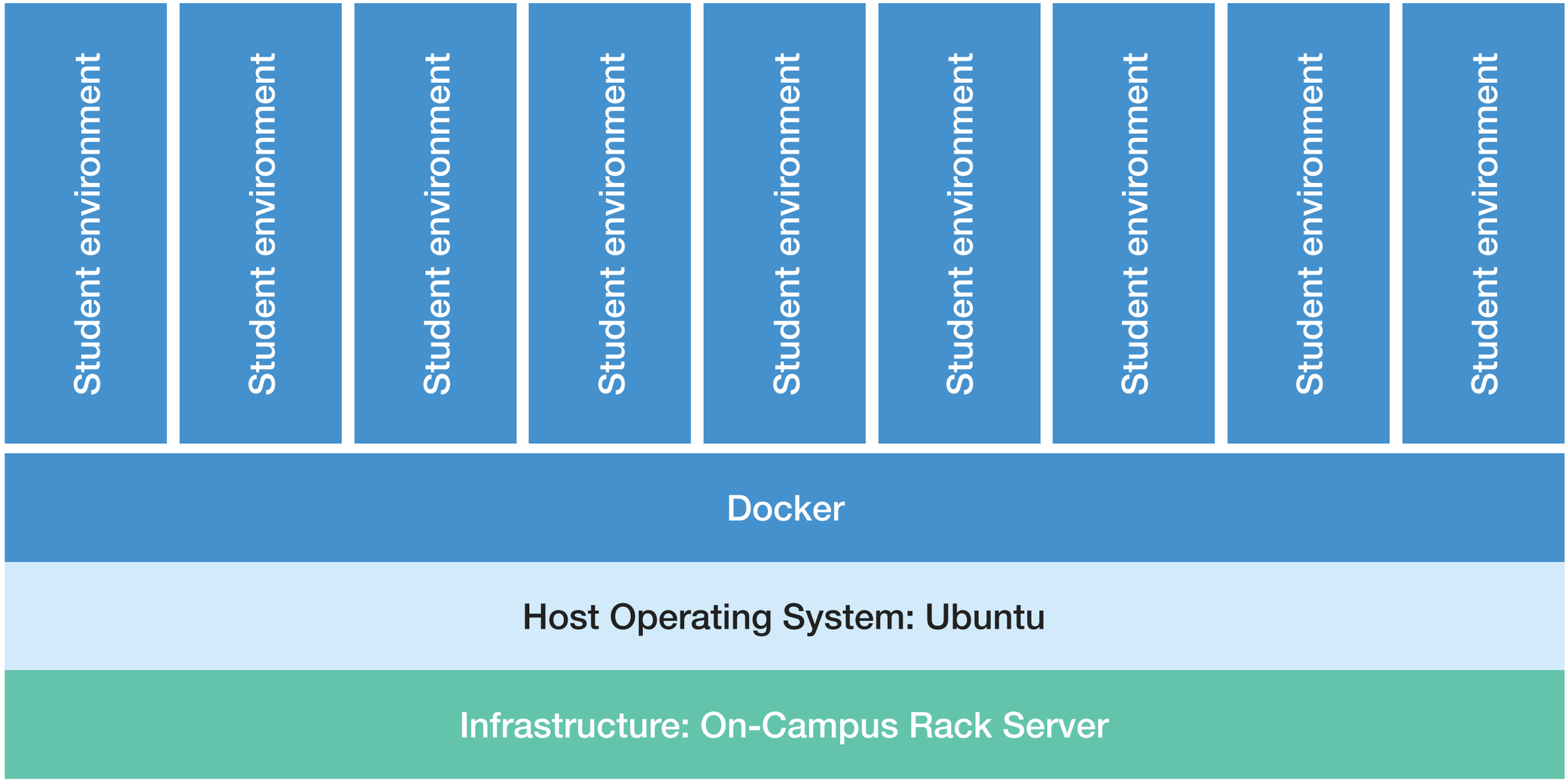




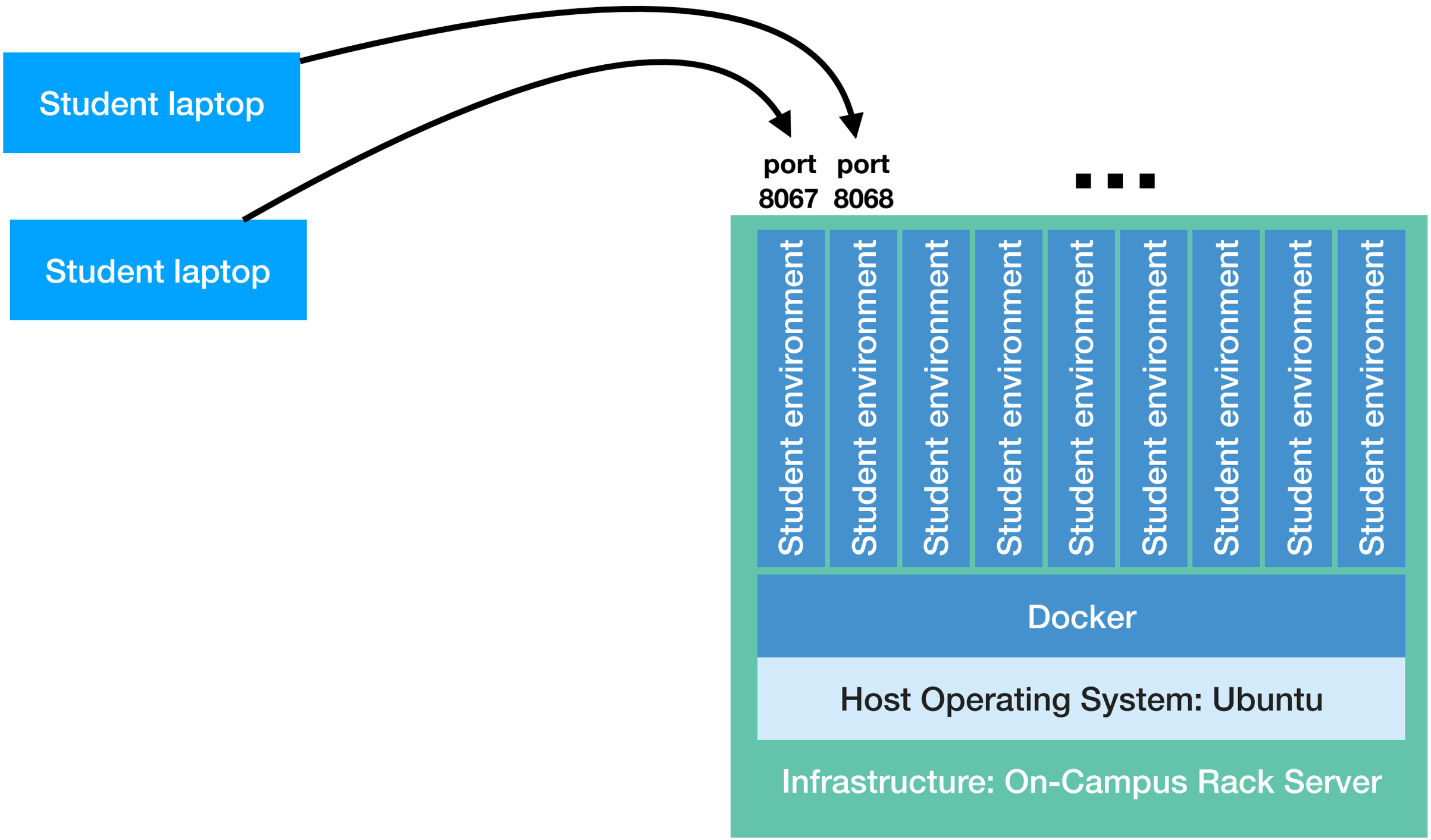
What is Docker?



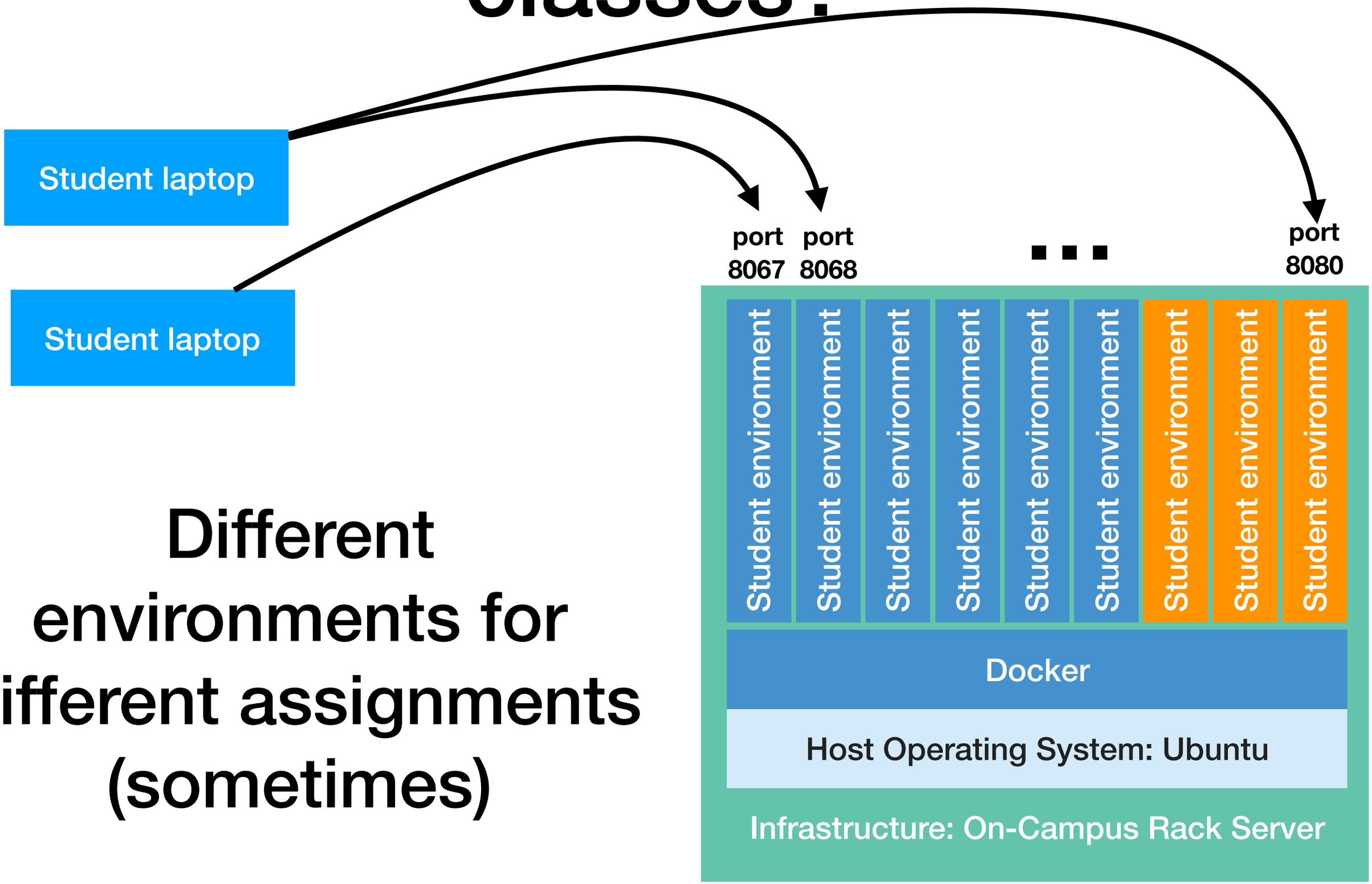
How do I use Docker with classes?



How do I use Docker with classes?

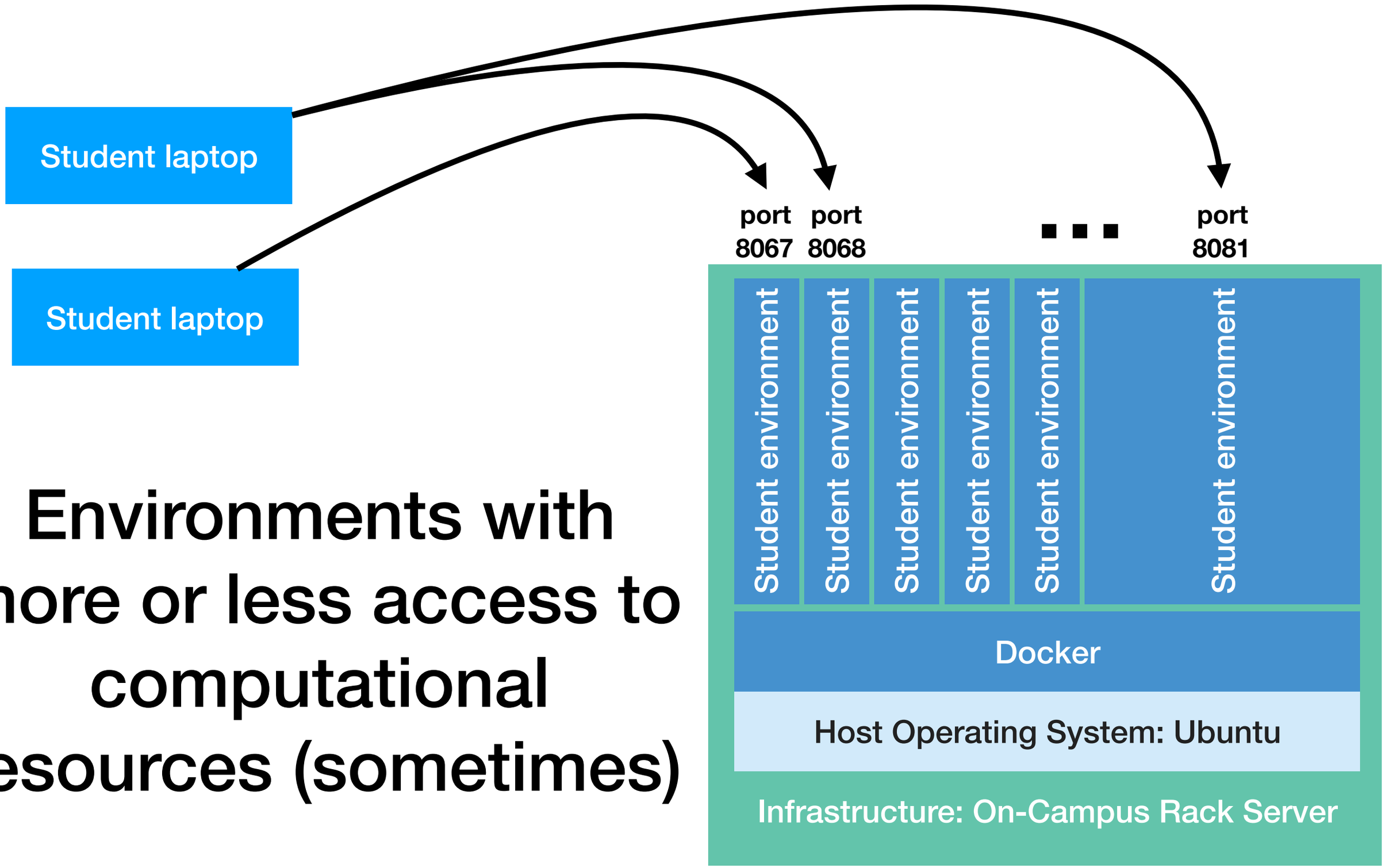


How do I use Docker with classes?



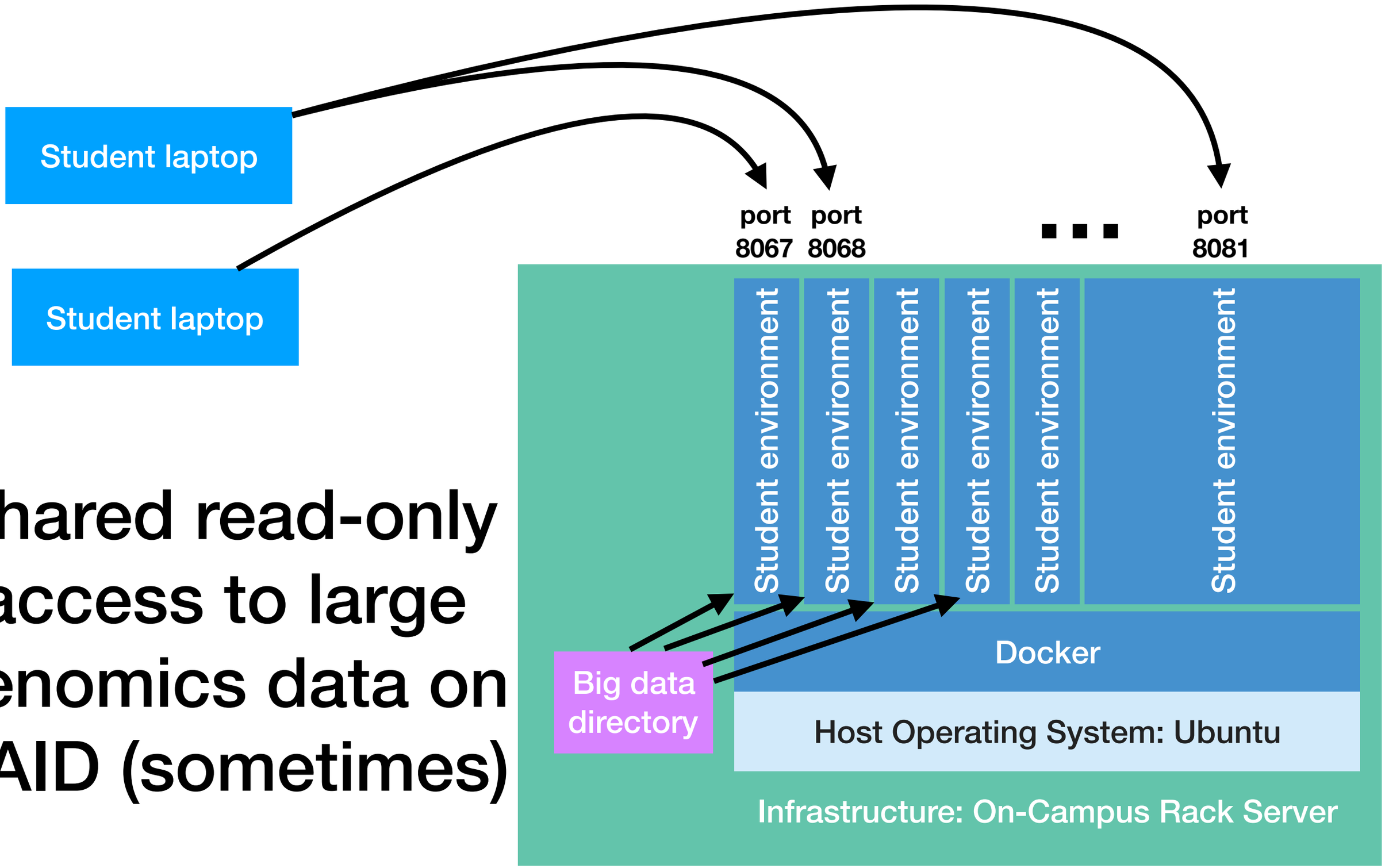
Different environments for different assignments (sometimes)

How do I use Docker with classes?



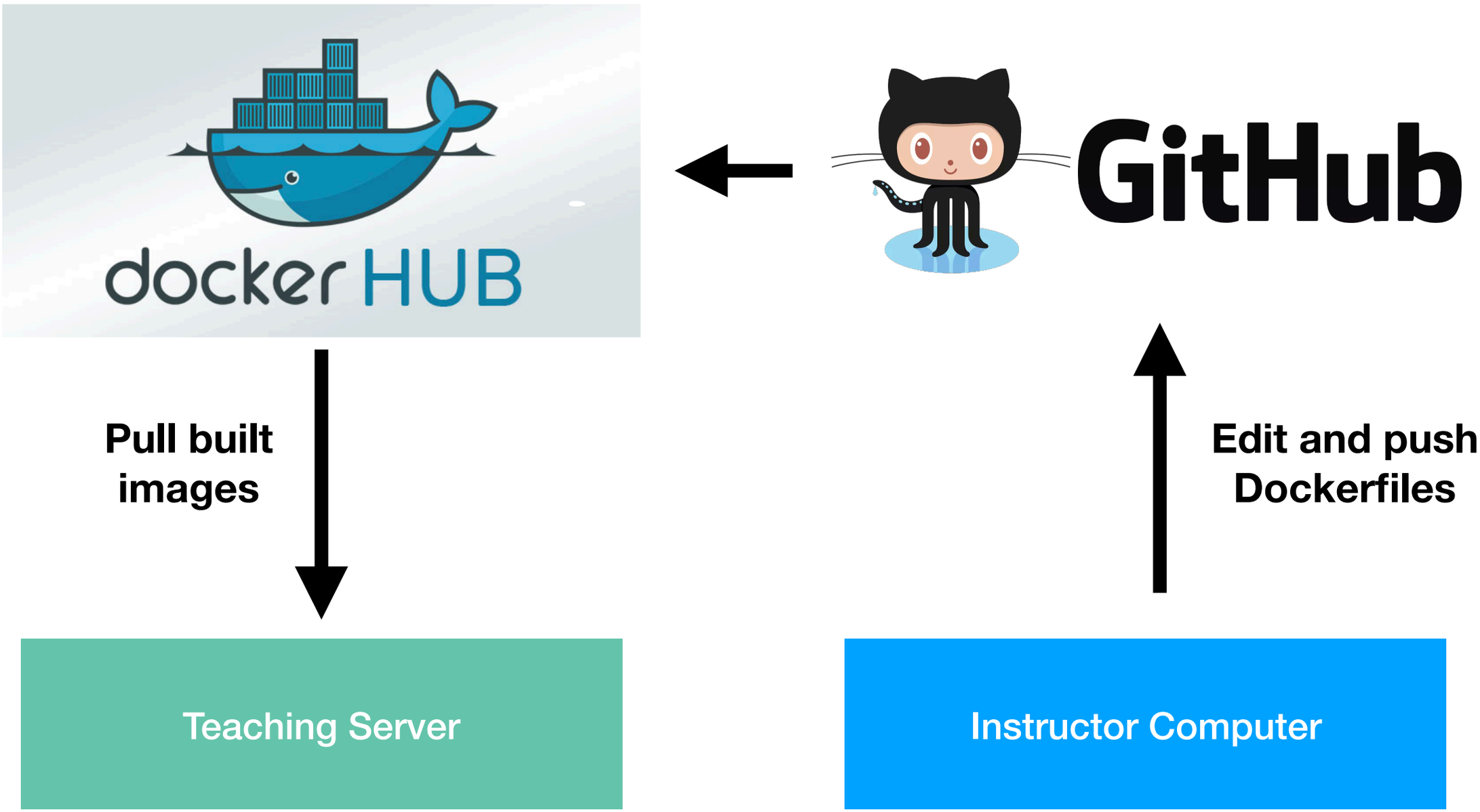
Environments with more or less access to computational resources (sometimes)

How do I use Docker with classes?



Shared read-only access to large genomics data on RAID (sometimes)

Keeping track of it all



Dockerfiles

Branch: ubuntu-sshd-bl... ubuntu-sshd / Dockerfile

naupaka Add markdown syntax highlighting to nano

2 contributors

67 lines (47 sloc) | 2.26 KB

```
FROM       ubuntu:16.04
MAINTAINER Naupaka Zimmerman "https://github.com/naupaka"

RUN apt-get update
RUN apt-get install -y openssh-server tmux nano git unzip \
    trimmomatic fastqc bison byacc ncbi-blast+ curl wget tar \
    make gcc libz-dev shellcheck

# there is a problem with normal fastqc installation
# have to fix by downloading config files from source
# http://www.bioinformatics.babraham.ac.uk/projects/download.html#fastqc
# and then putting the three files from the Configuration folder
# into /etc/fastq/Configuration

# Download the source and extract to get out config files
RUN curl http://www.bioinformatics.babraham.ac.uk/projects/fastqc/fastqc_v0.11.5_source.zip ·
RUN unzip /home/fastqc_v0.11.5_source.zip -d /home
```

Dockerfiles

Branch: ubuntu-sshd-bl... ubuntu-sshd / Dockerfile

naupaka Add markdown syntax highlighting to nano

2 contributors

67 lines (47 sloc) | 2.26 KB

```
FROM       ubuntu:16.04
MAINTAINER Naupaka Zimmerman "https://github.com/naupaka"

RUN apt-get update
RUN apt-get install -y openssh-server tmux nano git unzip \
    trimmomatic fastqc bison byacc ncbi-blast+ curl wget tar \
    make gcc libz-dev shellcheck

# there is a problem with normal fastqc installation
# have to fix by downloading config files from source
# http://www.bioinformatics.babraham.ac.uk/projects/download.html#fastqc
# and then putting the three files from the Configuration folder
# into /etc/fastq/Configuration

# Download the source and extract to get out config files
RUN curl http://www.bioinformatics.babraham.ac.uk/projects/fastqc/fastqc_v0.11.5_source.zip ·
RUN unzip /home/fastqc_v0.11.5_source.zip -d /home
```

Dockerfiles

Branch: ubuntu-sshd-bl... ▾ ubuntu-sshd / Dockerfile

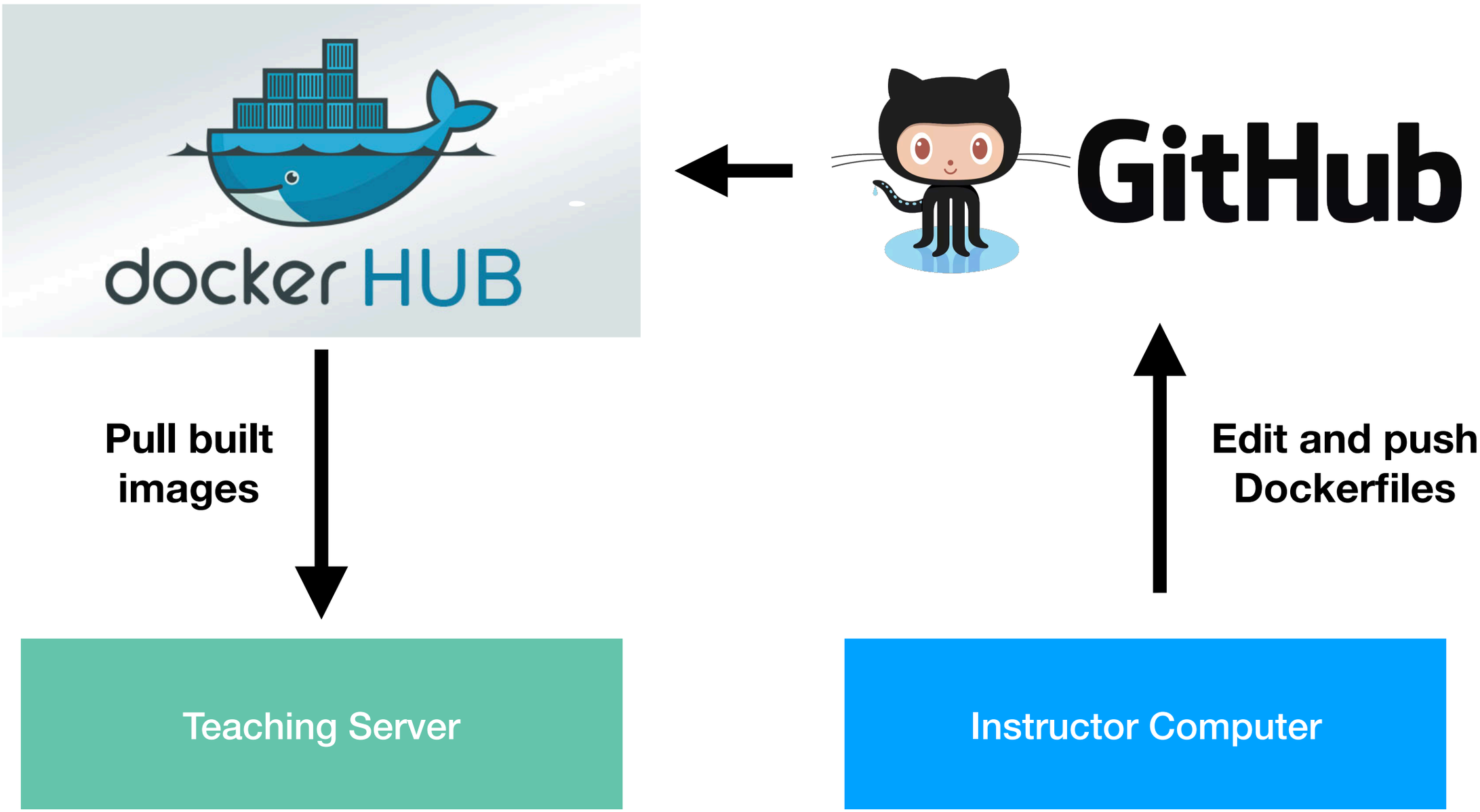
naupaka Add markdown syntax highlighting to nano

2 contributors

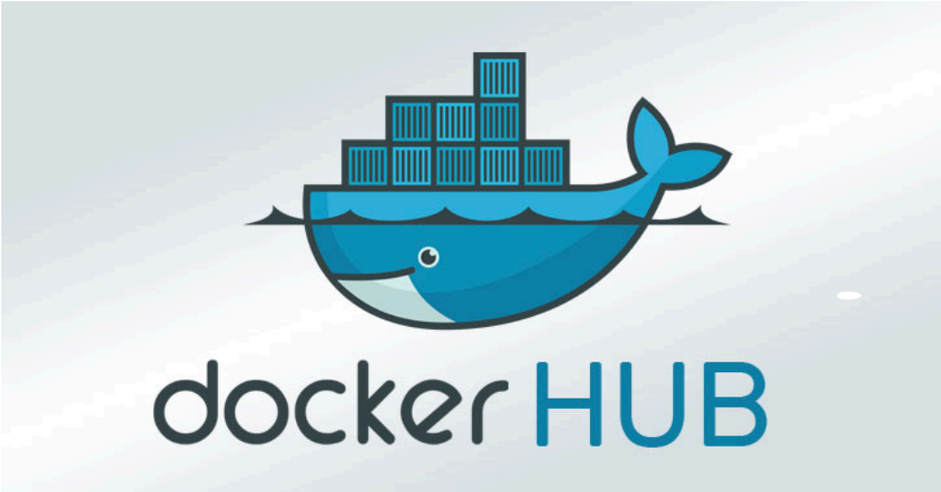
67 lines (47 sloc) | 2.26 KB

```
FROM       ubuntu:16.04
MAINTAINER Naupaka Zimmerman "https://github.com/naupaka"

RUN apt-get update
RUN apt-get install -y openssh-server tmux nano git unzip \
    trimmomatic fastqc bison byacc ncbi-blast+ curl wget tar \
    make gcc libz-dev shellcheck

# there is a problem with normal fastqc installation
# have to fix by downloading config files from source
# http://www.bioinformatics.babraham.ac.uk/projects/download.html#fastqc
# and then putting the three files from the Configuration folder
# into /etc/fastq/Configuration

# Download the source and extract to get out config files
RUN curl http://www.bioinformatics.babraham.ac.uk/projects/fastqc/fastqc_v0.11.5_source.zip ·
RUN unzip /home/fastqc_v0.11.5_source.zip -d /home
```

Keeping track of it all

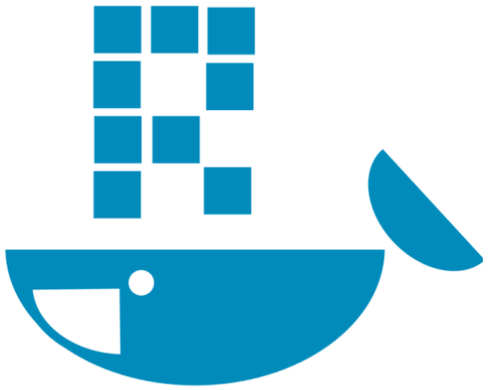


Keeping track of it all



Pull built images

Teaching Server



Rocker project

Pre-built options

For R, Rocker is Awesome

Versioned stack (builds on r-ver)

These images build on `rocker/r-ver`. Each of these include tags to specify the desired version of R, from `3.1.0` - `3.4.0`, `:latest` and `:devel`. See rocker-versioned repo for details.

image	description	size	metrics	build status
rstudio	Adds rstudio	355.1MB 21 layers	docker pulls 3M	docker build automated
tidyverse	Adds tidyverse & devtools	708.6MB 24 layers		
verse	Adds tex & publishing-related packages	1.1GB 26 layers		docker build automated
geospatial	Adds geospatial libraries	1.5GB 28 layers		

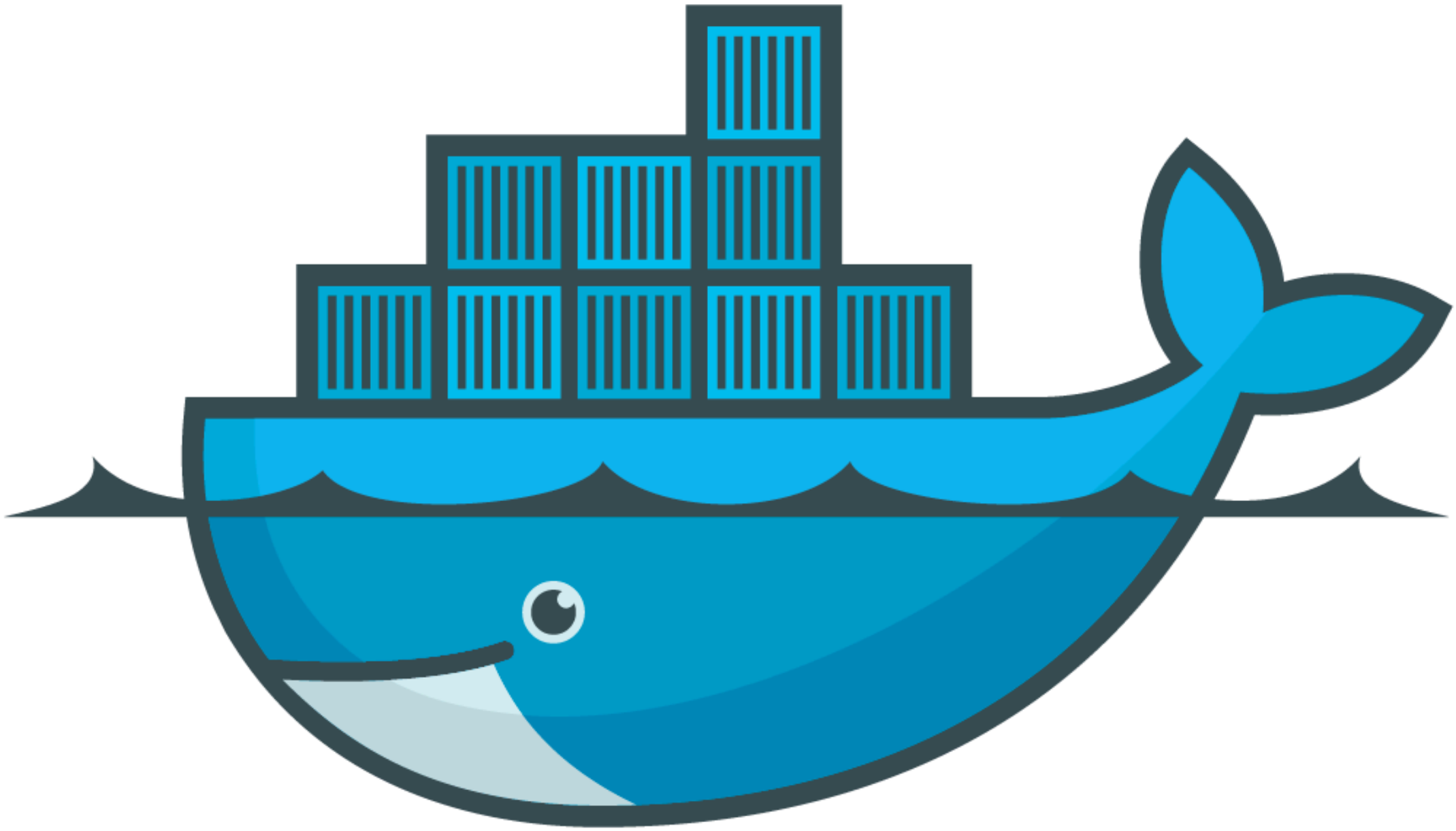
Anyone interested in proposing or collaborating on additional use cases should read our guide to contributing and get in touch

For R, Rocker is Awesome

```
sudo docker run -d -p 8787:8787 -e \
   PASSWORD=<password> --name rstudio \
   rocker/rstudio
```

How do I explain it all to students?

Questions?



Naupaka Zimmerman — @naupakaz — June 2019