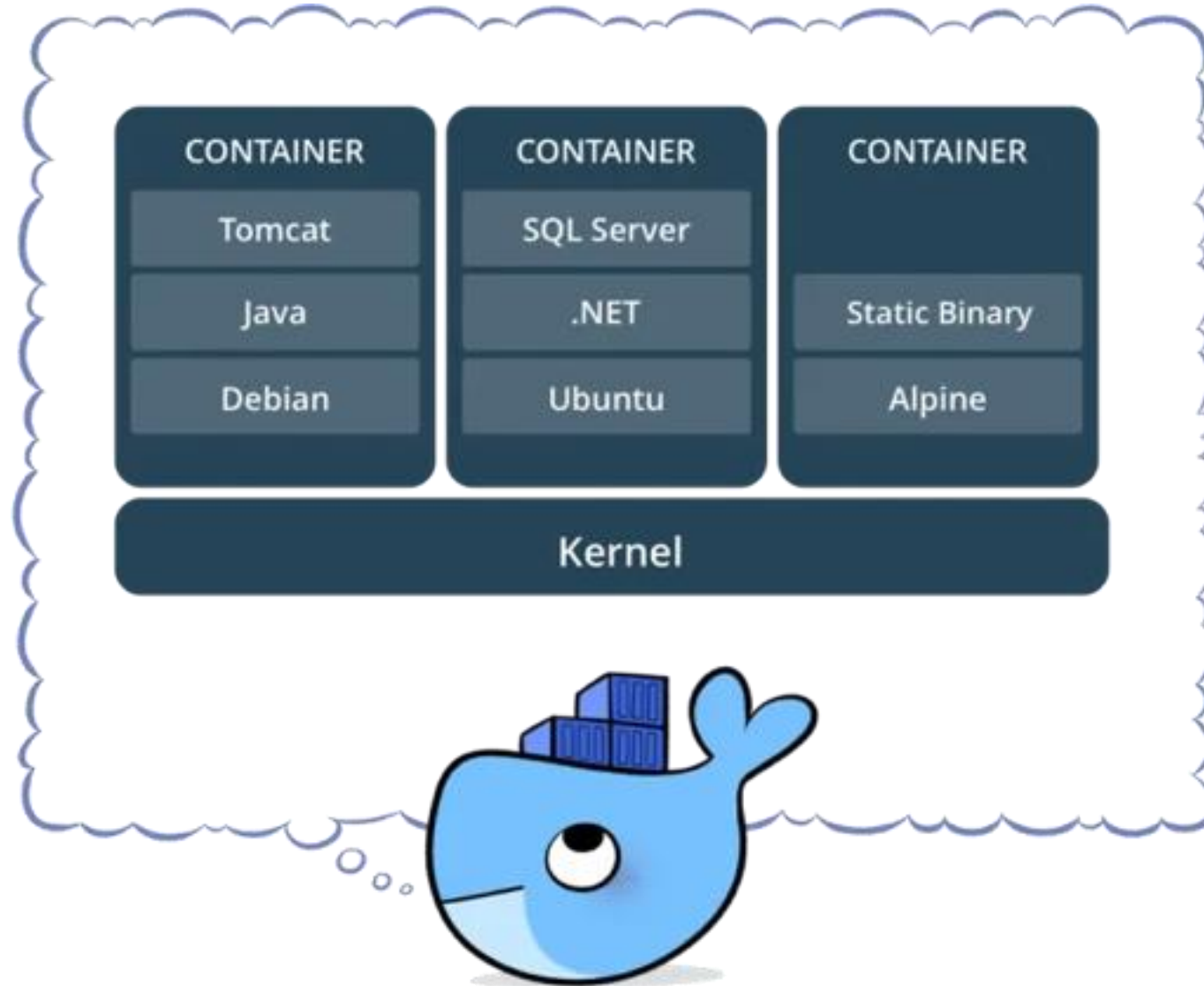# Intro to Docker

Computing Environments

- Laptop
    - You administer
    - You control completely

- Shared compute cluster
    - A sysadmin or group administers it
    - You control very little

- Docker (containers)
    - Sysadmins handle the hardware
    - You control the software almost completely

Docker containers

Docker Exercise

`https://gist.github.com/chrisamiller/fab79aa56b259e01744330e1d245085e`

Anatomy of a docker command

```
docker run -v /workspace:/data -it mgibio/gatk-cwl:3.6.0 /bin/bash
```

Anatomy of a docker command

`docker run` `-v /workspace:/data -it mgibio/gatk-cwl:3.6.0 /bin/bash`

tells the **docker** program that you want to **run** an image

Anatomy of a docker command

```
docker run -v /workspace:/data -it mgibio/gatk-cwl:3.6.0 /bin/bash
```

mounting directories:
the `/workspace` directory on your actual computer will be linked to
the `/data` directory inside your image

Any part of your computer that isn't linked explicitly
will be invisible inside the docker container!

Sometimes we'll use bash substitution to link the current directory:
`-v $(pwd -P):/data`

Anatomy of a docker command

`docker run -v /workspace:/data -it mgibio/gatk-cwl:3.6.0 /bin/bash`

run docker interactively, that is, drops you into a terminal

Anatomy of a docker command

```
docker run -v /workspace:/data -it mgibio/gatk-cwl:3.6.0 /bin/bash
```

the docker image that you want to run

this one is the `gatk-cwl` image,
hosted in the `mgibio` organization on
dockerhub
and we're pulling the image that is tagged `3.6.0`

If you don't provide a tag, it looks for a tag
called `latest`

(Why might using tags be important for your
analyses?)

Anatomy of a docker command

```
docker run -v /workspace:/data -it mgibio/gatk-cwl:3.6.0 /bin/bash
```

The command that you want to run inside of docker.

In this case, we're starting a shell so that we can do command line stuff

Anatomy of a docker command

```
docker run -v /workspace:/data mgibio/gatk-cwl:3.6.0 \
gatk --java-options '-Xmx7g' HaplotypeCaller \
-R /workspace/inputs/references/genome/ref_genome.fa \
-I align/WGS_Norm_merged_sorted_mrkdup_bqsr.bam \
-O /workspace/germline/WGS_Norm_HC_calls.vcf \
--bam-output /workspace/germline/WGS_Norm_HC_out.bam
```

In this example, we removed **-it** and just added the command we want to run. Docker will automatically exit after running this command.