Coffee & Code To Go

- This presentation: https://github.com/unmrds/cc-tech-days/blob/master/Coffee_and_Code_2_Go.ipynb (https://github.com/unmrds/cc-tech-days/blob/master/Coffee_and_Code_2_Go.ipynb)
- Presentation Repository: https://github.com/unmrds/cc-tech-days (https://github.com/unmrds/cc-tech-days)
- Coffee & Code Platform: phttps://github.com/unmrds/cc-content-platform] (https://github.com/unmrds/cc-content-platform) (https://github.com/unmrds/cc-content-platform](https://github.com/unmrds/cc-content-platform))

Playground: http://cc.unmrds.net:8888 (http://cc.unmrds.net:8888)

Everyone's favorite workshop series has a new, portable architecture!

The only software requirement for playing along (or deploying on any server or operating system) is Docker: https://www.docker.com/ (https://www.docker.com/)

About Coffee & Code: Whys and Wherefores

The series and selection of topics are motivated by factors including:

- Developing and promoting the CSEL DEN
- Outreach programming
- Fun!

The content and structure of each session is likewise informed by various models.

	Carpentries	Data Management Curricula	Librarian Skill & Professional Development
Programs	Software Carpentry	DataONE Education Modules	MANTRA DIY Research Data Management Training Kit for Librarians
	Data Carpentry	NECDMC	
	Library Carpetry	Research Data MANTRA	
Example Concepts and Skills	tabular data management and analysis (OpenRefine, SQL, R)	data management plans and planning	research data management concepts
	version control	metadata and data documentation	data archiving and preservation

More about...

The Carpentries (https://carpentries.org/)

New England Collaborative Data Management Curriculum (https://library.umassmed.edu/necdmc/modules)

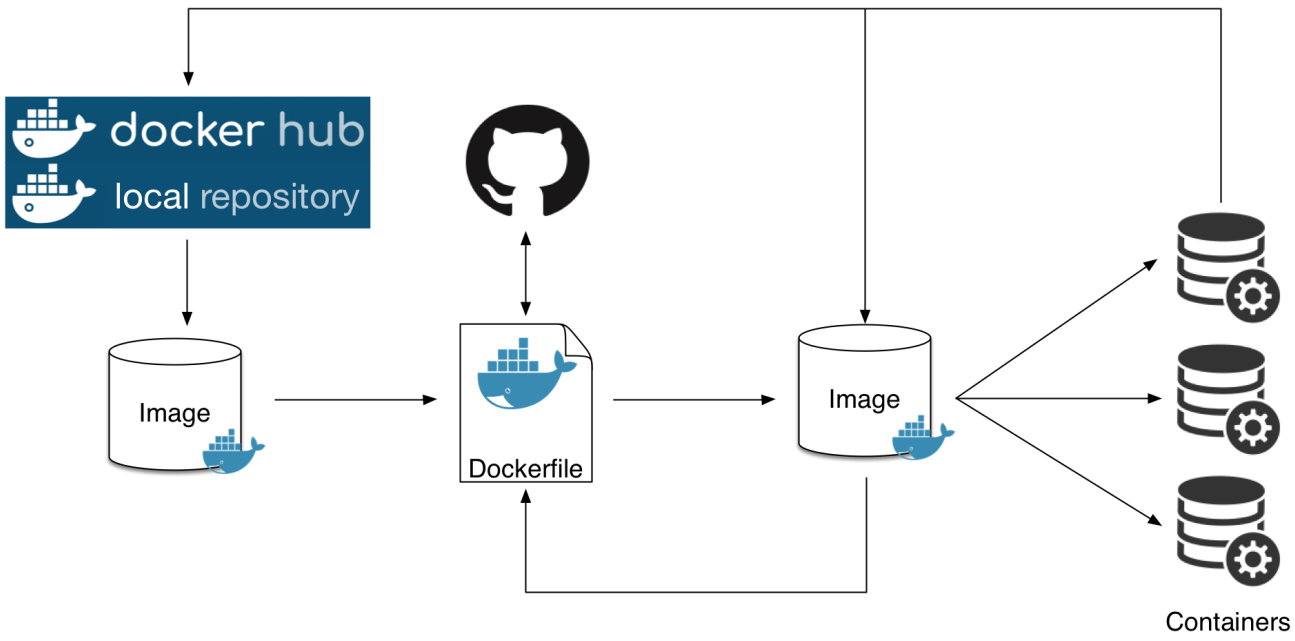
Interlude: Revamped Workshop Architecture

Version 1: Docker Containers for Each Individual Workshop

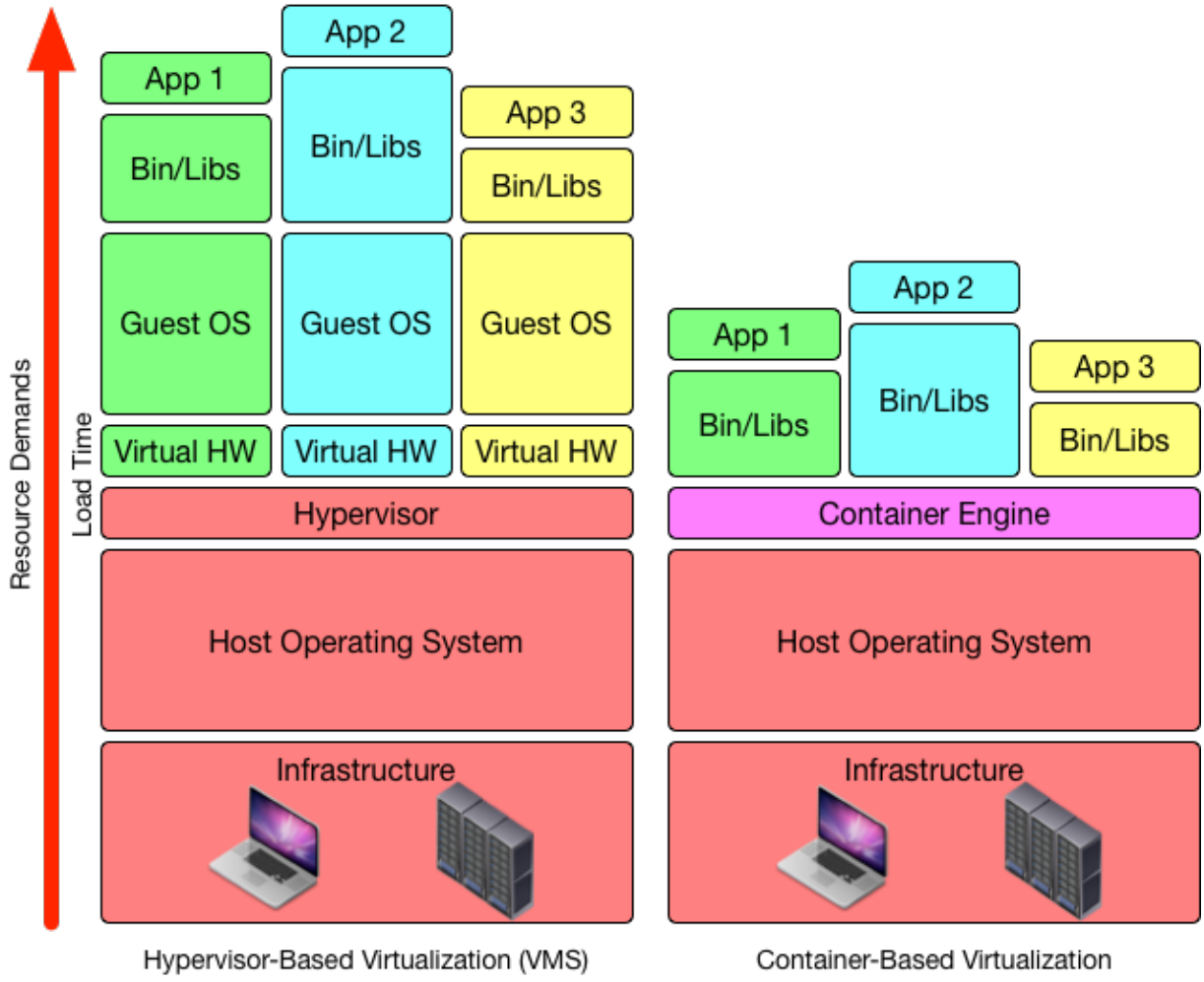
The first iteration of the *Coffee and Code* platform consisted of individual Docker (https://www.docker.com) containers (one from which the instruction is provided, a second *playground* container that workshop participants could experiment with) and associated workshop materials that were maintained in separate repositories (https://github.com/unmrds) within the UNM RDS GitHub organization.

A Docker Primer

Containerization is the process of defining the software components required to run a specific application. In the context of a Docker-based containerization process this definition is written into a Dockerfile. A Dockerfile is then used to create a Docker Image. That Docker Image can then be used as the basis for one or more separate Docker Containers, or as the starting point for defining a new Dockerfile.



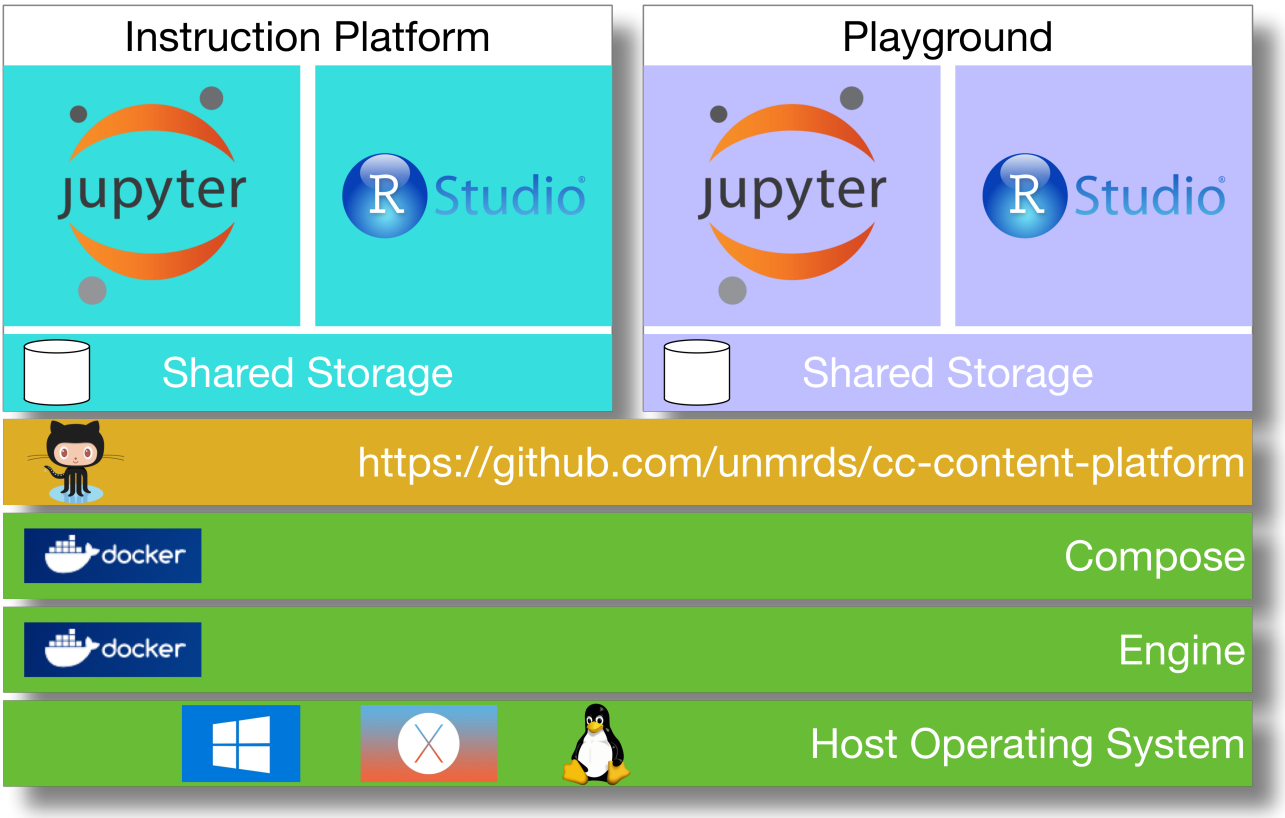
and



Note: I would refer you to a more detailed introduction that was provided as part of our [Coffee and Code workshop (https://github.com/unmrds/cc-containers/blob/master/01-Docker%20Overview.ipynb)](https://github.com/unmrds/cc-containers/blob/master/01-Docker%20Overview.ipynb) focused on containerization for a more detailed discussion of this topic.

Version 2: A Single Orchestrated Docker Environment That Automatically Contains Current Versions of All Workshops

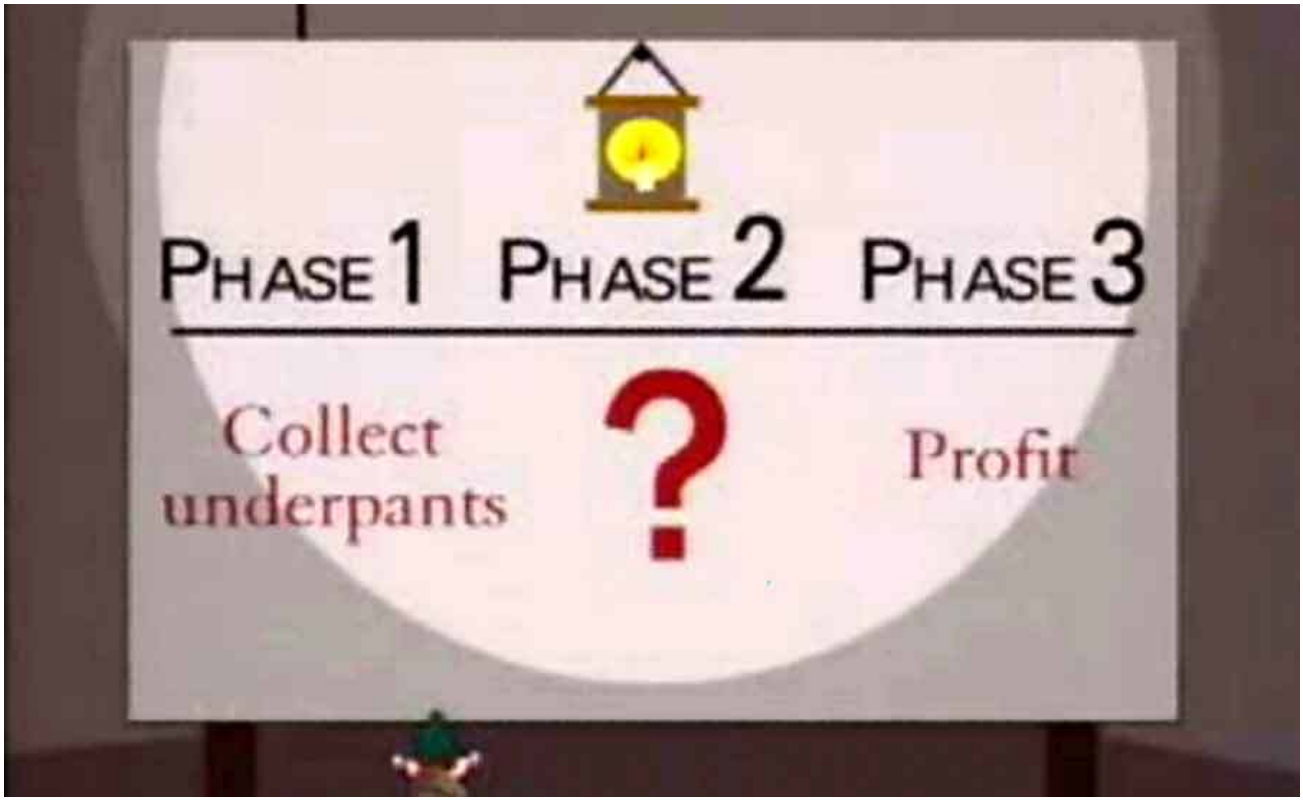
To streamline the preparation and presentation of the workshops we adopted the *[Docker Compose (https://docs.docker.com/compose/)](https://docs.docker.com/compose/)* approach of defining a single configuration file that could be used to run multiple containers from a single command. This significantly streamlined the deployment of updated versions of the workshops and accelerated our development and testing cycle for workshop content.



As part of the execution of the configuration file additional editable setup files are run that refresh the content of the running environment with the current set of workshops that are hosted in GitHub.

Version 2.1: Packaging Coffee and Code To Go in a Free-standing GitHub Repository for Streamlined Access and Execution on Instruction and Student Computers

While not quite as easy as the business plan developed by the South Park *underpants gnomes* ...



the process of replicating our instruction and learning platform is a simple three-step process:

1. Install the Docker Platform (https://docs.docker.com/install/) including *Docker-compose* on your computer.
2. Download the current *Coffee and Code* platform from the UNM RDS GitHub Repository (https://github.com/unmrds/cc-content-platform) by either *cloning* the repository using a local copy of the Git distributed version control system, or by simply downloading the ZIP archive (https://github.com/unmrds/cc-content-platform/archive/master.zip).
3. Run the `docker-compose` command to build (the first time) and spin up the containers. Once they are running the instruction and learning platform can be interacted with through a web browser.

Upcoming Workshops

Subject to change

- August 10: Advanced Version Control with Git, GitHub, and LoboGit
  - Resetting and reverting commits
  - Resolving conflicts
  - Forking, branching, merging
- September 14: Quantitative Analysis Using R and RStudio
  - Big Data Analytics
  - Subsetting and visualizing large datasets
- October 12: Natural Language Processing with Python
  - Word counts, ngrams, stop words
  - NLTK: The Natural Language Toolkit
  - Sentiment analysis
- November 9: Database Fundamentals part 2: MySQL and Postgres
  - Data modeling, primary and foreign keys
  - Optimizing performance with indexing
  - Joins
- December 14: NoSQL and Graph Databases with Elasticsearch, MongoDB and NodeJS
  - Fun stuff
  - More fun stuff
  - Applied fun stuff

In [ ]: