GAP Docker containers for reproducible computational experiments (and more ...)

Alexander Konovalov

Centre of Interdisciplinary Research in Computational Algebra

University of St Andrews

Computational Mathematics with Jupyter ICMS, Edinburgh, 16-20 January 2017

We acknowledge financial support from:

- OpenDreamKit Horizon 2020 European Research Infrastructures project (#676541)
- Collaborative Computational Project CoDiMa (CCP in the area of Discrete Computational Mathematics (EP/M022641/1).







Computational Discrete Mathematics

What if ...

- You need to quickly get fully functional installation of GAP and packages, but ...
 - Some packages are not compiling on your computer
 - Some packages have complicated dependencies
 - Some may be only installed on a particular machine so you need network connection to access it
- But you want to:
 - Be able to take this with you to the journey
 - Share your experimental setup with collaborator
 - Provide environment for reproducing your experiment
 - Automate deployment on the cluster / in the cloud
 - Save time in a large-scale tests or experiments

Ideas?

- Virtual machines
- VMWare, VirtualBox, ...
- One more computer to maintain?
- Eventually becomes outdated
- Could automate installation scripts
- Wouldn't it be nice to have something more lightweight?

What is Docker?

- Solution for virtualisation
- Lightweight "containers" instead
- <u>https://www.docker.com</u>
- Documentation: <u>https://docs.docker.com/</u>
- For Linux, OS X, Windows, major cloud platforms, ...
- To get it, go to <u>https://www.docker.com/products/overview</u>
- Docker toolbox for older versions of OS X and Windows: <u>https://www.docker.com/products/docker-toolbox</u>

A simple example

- run Docker container for GAP and packages:
 - First, pull the container from the Docker Hub
 - Final Formation Formation The Territor of territor of territor of territors and territors are territor of territors and territors are territor
 - GAP options can be added after, e.g. `gap -A`
- See this container in
 - Docker Hub: <u>https://hub.docker.com/r/gapsystem/gap-docker/</u>
 - GitHub: <u>https://github.com/gap-system/gap-docker</u>
- Explore container's filesystem and GAP installation

```
docker pull gapsystem/gap-docker
docker run --rm -i -t gapsystem/gap-docker gap
```

More use cases in README.md

Using network

- Mounting local folders
- Let's see the demo

docker run -v /Users/alexk/dockermount:/work --rm -i -t gapsystem/gap-docker
gap /work/homophony.g

History of Docker support in GAP

- Hackday project at Collaborations Workshop 2015, with Sarah Mount (University of Wolverhampton), Devasena Inupakutika (Software Sustainability Institute), and myself
- Shortly after that joined by Sebastian Gutsche
- We have since established a pipeline of containers
- As alternative GAP distributions, they are usually updated shortly after the GAP release

	gapsystem/gap-docker public automated build	3 STARS	734 PULLS	DETAILS
<image/>	gapsystem/gap-container public automated build	1 STARS	158 PULLS	DETAILS
	gapsystem/gap-docker-base public	1 STARS	119 PULLS	DETAILS
Sustainability Institute	gapsystem/gap-docker-jupyter public automated build	1 STARS	70 PULLS	DETAILS

GAP Containers pipeline

- Sontainer for the core GAP system and only GAPDoc package
 - https://hub.docker.com/r/gapsystem/gap-container/
 - <u>https://github.com/gap-system/gap-container</u>
- Intermediate container for external software used by GAP packages.
- Needed to build gap-docker
 - https://hub.docker.com/r/gapsystem/gap-docker-base/
 - https://github.com/gap-system/gap-docker-base
- Container for GAP and all packages redistributed with GAP
 - <u>https://hub.docker.com/r/gapsystem/gap-docker/</u>
 - <u>https://github.com/gap-system/gap-docker</u>
- Container for the GAP Jupyter interface
 - https://hub.docker.com/r/gapsystem/gap-docker-jupyter/
 - https://github.com/gap-system/gap-docker-jupyter

Use case: testing GAP packages using Travis CI

<u>https://github.com/gap-system/gap-docker-pkg-tests</u>
<u>https://travis-ci.org/gap-system/gap-docker-pkg-tests</u>

✓ # 23.1		PKG_NAME=ace	() 3 min 1 sec
× # 23.2	🖧 > C	PKG_NAME=alnuth	() 2 min 56 sec
× # 23.3	🖧 > C	PKG_NAME=anupq	() 13 min 8 sec
× # 23.4	🖧 > C	PKG_NAME=atlasrep	() 9 min 43 sec
✓ # 23.5	🖓 > C	PKG_NAME=automgrp	() 3 min 28 sec
✓ # 23.6	🖓 > C	PKG_NAME=browse	() 4 min 55 sec
✓ # 23.7	🗞 > C	PKG_NAME=circle	() 2 min 54 sec
✓ # 23.8	🖓 > C	PKG_NAME=cohomolo	() 3 min 27 sec

Reproducibility

- Code inside new container
- Code outside existing container
- Recommend to read this Bill Mills' blog post:
 - "Full Stack Science"
 - http://billmills.github.io/blog/full-stack

What's in the future?

- Reliable to "use as you go"
- Reasonable to expect to run not too old containers for a while
- Unclear (to me at least) how much it is reliable for long term preservation
- Docker's first release made in March 2013
- Popular and active project
- Docker Hub (<u>https://hub.docker.com/</u>) free for public repositories
- If situation will change, the mindset will still stay

Coding sprint at this workshop?

- Installing and trying GAP Jupyter interface
- Update for GAP 4.8.6
- Help with issues on GitHub
- Frying and testing