Analysing Software Containers from a Compilation point of view

Laure Gonnord & Sébastien Mosser

16 octobre 2017

General Context

The adoption of the container approach is tremendously increasing [DC16]. Containers ensure that a given microservice will run the same regardless of its environment, easing the repeatability of build, test, deployment and runtime executions [Boe15, Mer14]. Containers are faster at runtime and boot-time, lighter than virtual machines, and scale better [XNR+13, PHP16, MKK15, FFRR15]. In the container field, the Docker engine quickly became the reference platform for builds and deployments [Rus16].

Building a non-trivial container image is a difficult process. Describing an image is an imperative process, where a *service deployment descriptor* is written (*e.g.*, a *dockerfile* in the Docker ecosystem) to describe as shell commands how the microservice is installed and configured inside the container. Following an *off-the-shelf* approach, a container is defined on top of others (reused as black boxes). However, this implementation is not compliant with the open/closed principle, as it is open for extensions (a descriptor extends another one), but not closed for modifications (a descriptor does not provide a clear interface about its contents, making reuse hazardous). By hiding the contents of an image as a blackbox, deployment instruction can conflict with the hidden one, *e.g.*, overriding executables, duplicating installation of the same piece of software in alternative versions, or shadowing executables. It leads to erroneous deployments, detected at runtime. Moreover, the technologies supporting microservice deployment evolve constantly, to make it more efficient or scalable. This evolution can drastically change the way the deployment engine is implemented, and abstraction leaks can occur (*i.e.*, an internal technological choice inside the deployment engine the final user must take into account when writing a service descriptor). It is up to the service developer to stay up to date with ever-changing guidelines that implements fixes to abstraction leaks.

Internship Objective

In this internship, we propose to address the challenges of deployment descriptor definition from a compilation point of view. We propose to first build an empirical datasets of available dockerfiles (*e.g.*, by crawling GitHub) and identify in these dockerfiles classical flaws encountered by developers (for example by using classification mechanisms). Then, we will work at the descriptor language level to *(i)* provide a sound grammar to safely define and extend descriptors, *(ii)* define formal detection rules to identify in existing descriptors the identified design flaws and *(iii)* provide a bidirectional transformation from our representation to deployable descriptors.

Expected skills

I3S – UMR CNRS / Université Nice Sophia Antipolis & LIP – UMR CNRS / ENS Lyon / UCB Lyon 1 / INRIA
Contact E-mail: Laure.Gonnord@ens-lyon.fr and mosser@i3s.unice.fr

— A taste for empirical approaches ;
— The ability to identify trade-off among multiple solutions ;
— Good communication skills and will to share ideas and discuss results.

This work will be supervised mostly by Sebastien Mosser in Sophia Antipolis and remotely by Laure Gonnord in Lyon.

Références

[Boe15] Carl Boettiger. An introduction to docker for reproducible research. *SIGOPS Oper. Syst. Rev.*, 49(1) :71–79, January 2015.

[DC16] DevOps.com and ClusterHQ. Container market adoption - survey 2016, jun 2016.

[FFRR15] W. Felter, A. Ferreira, R. Rajamony, and J. Rubio. An updated performance comparison of virtual machines and linux containers. In *Performance Analysis of Systems and Software (ISPASS), 2015 IEEE International Symposium on*, pages 171–172, March 2015.

[Mer14] Dirk Merkel. Docker : Lightweight linux containers for consistent development and deployment. *Linux J.*, 2014(239), March 2014.

[MKK15] R. Morabito, J. Kjällman, and M. Komu. Hypervisors vs. lightweight virtualization : A performance comparison. In *Cloud Engineering (IC2E), 2015 IEEE International Conference on*, pages 386–393, March 2015.

[PHP16] René Peinl, Florian Holzschuher, and Florian Pfitzer. Docker cluster management for the cloud - survey results and own solution. *Journal of Grid Computing*, 14(2) :265–282, 2016.

[Rus16] Gareth Rushgrove. Dockercon16 - the dockerfile explosion and the need for higher level tools by gareth rushgrove, jun 2016.

[XNR+13] M. G. Xavier, M. V. Neves, F. D. Rossi, T. C. Ferreto, T. Lange, and C. A. F. De Rose. Performance evaluation of container-based virtualization for high performance computing environments. In *2013 21st Euromicro International Conference on Parallel, Distributed, and Network-Based Processing*, pages 233–240, Feb 2013.