

## Smartly containing the cloud increases computing efficiency, says first-of-its-kind study

February 6 2018

---



Ali Butt, (left) professor of computer science, and Ali Anwar (right), a Ph.D. candidate in the Department of Computer Science, will present their research in Oakland, California, at FAST '18. Credit: Virginia Tech

Not too long ago booting up a computer meant there was time for a lengthy coffee break even before the workday started. For a decade now though, thanks to the cloud, computers have accessed information from virtual machines that exist in the ether, allowing software to launch quickly on demand.

Now, in a first-of-its kind study funded by IBM and the National Science Foundation, Virginia Tech researchers have discovered ways to further improve computing efficiency using management tools for cloud-based light-weight virtual machine replacements called containers—frameworks that allow the microservices that power data retrieval from the ether—to deploy in a more agile manner.

The research team will present their findings in Oakland, California, at FAST'18, the [16th USENIX Conference on File and Storage Technologies](#) in February.

Unlike the software-heavy [virtual machines](#), containers share the core of the underlying operating system, which enables faster deployment of software programs without diminishing performance.

"Containers are just now being studied as part of the cloud infrastructure, but our research indicates that how they function in the cloud is critical to developing and distributing future computer systems that maximize efficiency," said Ali Anwar, lead author on the paper that details the research and a Ph.D. candidate in Virginia Tech's Department of Computer Science in the College of Engineering.

The study was a collaboration with IBM and offers a large-scale survey of the commonly used container management framework known as Docker, a platform that facilitates the deployment of microservices by providing a registry service that acts as a central repository for software components focusing on specific functionalities called images. When

users want to publish their images, the registry makes them accessible to others.

The team analyzed an unprecedented amount of data from five geographically distributed data centers over 75 days spanning 38 million requests and 181.3 TB of traces, or timestamped logs that document a program's execution. The customer set of the study ran the gamut from individuals to small- and medium-sized businesses to large government institutions.

The research uncovered an important aspect of container technology that utilized caching and prefetching of information. The team found that these were important in reducing latency. "This study is crucial to understanding whether containers are amenable to prefetching and how such techniques can improve cloud efficiency," said Ali Butt, co-author and a professor of computer science. "Prefetching data to setup containers even before they are requested by the users allows applications to run far more quickly."

Butt explains the advantage of prefetching in a modern-day metaphor as the difference between setting up a meeting time at 10 a.m. and being ready :30 seconds before with coffee in hand, as opposed to showing up at the designated time.

Existing research of containers indicated that performance issues became apparent in the lifecycle of a container when the number of stored images and concurrent user requests for data increased.

The container registry Docker grows by about 1,500 new daily public repositories, and retrieving images from such a growing repository can account for 76 percent of the container start time. This means that using the cloud for the email you're trying to send or the purchase you are trying to make online takes that much longer to process.

Another key finding includes that younger nonproduction registries experience lower loads compared to longer-running production systems, which can inform how the registry load is affected.

"Our collaboration with Virginia Tech really allowed us to see how data in the wild was performing and how the current microservices were working to achieve tasks of retrieving and posting data," said Mohamed Mohamed, collaborator on the study and member of IBM's container storage research group, Ubiquity. "Without the ability to use such a large and varied data set from IBM, we couldn't have come to the conclusions we did."

In performing this large-scale analysis, the team developed a valuable tool to analyze registry data for further research, and also open source the data and tool for the benefit of the broader cloud computing community.

Ultimately advances in [container](#) technology have the potential for widespread improvement of cloud computing performance. "Container transparency allows a cloud provider to gain insight into applications security, compliance, and performance, enabling new kinds of user-facing application-centric services," said Mohamed.

Fetching your coffee won't be one of them, however.

Provided by Virginia Tech

Citation: Smartly containing the cloud increases computing efficiency, says first-of-its-kind study (2018, February 6) retrieved 18 April 2024 from <https://techxplore.com/news/2018-02-smartly-cloud-efficiency-first-of-its-kind.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private

study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.