Newsletter

CloudSkin

March, 2025

GRANNY: Granular Management of Compute Intensive Applications in the Cloud

We're excited to announce that our latest work, "**GRANNY: Granular Management of Compute Intensive Applications in the Cloud**" will be presented at the **USENIX NSDI** (https://www.usenix.org/conference/nsdi25) conference in Philadelphia this coming April 2025. This project is part of the broader CloudSkin initiative, which aims to transform how cloud resources are utilized by enabling efficient and elastic execution of compute-intensive applications.

GRANNY tackles a fundamental limitation of cloud schedulers when executing complex multi-threaded and multi-process applications: they cannot change the allocation of threads and processes to cloud VMs once applications have started executing. This prevents cloud schedulers from adapting to changes in the availability of cloud resources, they cannot elastically add threads to a multi-threaded application if more resources in the same VM become available, and they cannot migrate processes across VMs to improve locality of execution.

by Peter Pietzuch and Carlos Segarra (Imperial College London)

GRANNY enables fine-granular management of multi-thread and multi-process applications by executing threads and processes as "Granules", a new WebAssembly-based execution abstraction that supports efficient snapshotting. GRANNY uses Granule snapshots to vertically scale-up multi-threaded applications, by adding more Granules, or horizontally migrate Granules in a multi-process application, balancing load between VMs. GRANNY can execute unmodifiedMPI and OpenMP applications, and we implement different Granule-aware scheduling policies that considerably improve the resource utilization of cloud VMs, and the performance of individual applications.

We look forward to sharing our research with the community and engaging in discussions about the future of cloud resource management both at NSDI and as part of the CLoudSkin project.

Funded by the European Union