

TOWARDS IMPROVED CLOUD FUNCTION SCHEDULING IN FUNCTION-AS-A-SERVICE PLATFORMS

Student: Edwin F. Boza

<eboza@espol.edu.ec>

Advisor: Prof. Cristina L. Abad

<cabad@fiec.espol.edu.ec>

Affiliation: Escuela Superior Politécnica del Litoral, ESPOL

Guayaquil - Ecuador

Research: Distributed systems, key-value stores as cache, workload modeling,
scheduling in serverless...

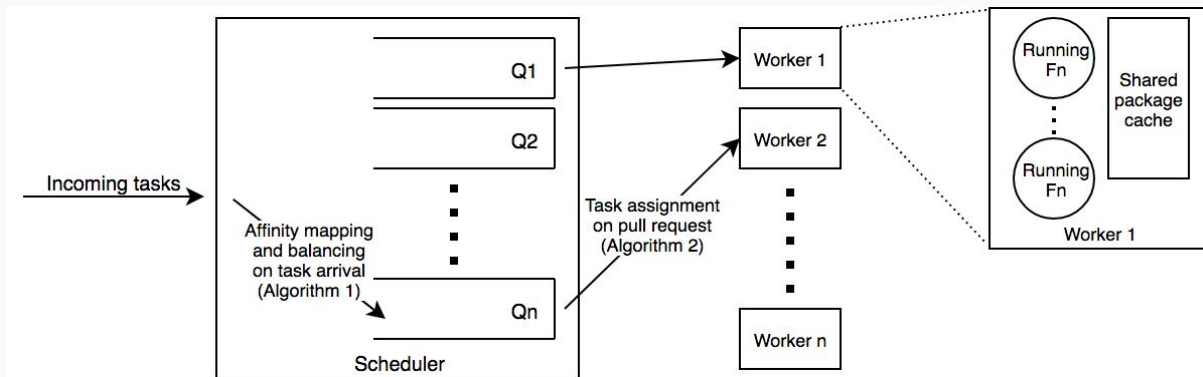


ESPOL
"Impulsando la sociedad del conocimiento"

Towards Improved Cloud Function Scheduling in Function-as-a-Service Platforms

- FaaS architectures are increasingly popular, but present several performance challenges time-sensitive services.
- Potential high latencies, derived from the initialization phase, make this architecture unsuitable for some use cases.
- ***Smart scheduling cloud functions, can improve performance in FaaS architectures***
- By working on the scheduler I could:
 - Reduce function launch time by minimizing control plane and data plane communications?
 - Reconcile possibly conflicting goals of the function scheduling process? (e.g., balance worker load, maximize data locality, maximize code locality, meet QoS guarantees)
 - Allow the provider to offer differentiated services in terms of launch time latency?

Preliminary results: Improve function latency by increasing code locality



Results:

- Moderate improvements in **hit rates of the package caches** at the worker nodes. Our algorithm improves the average hit rate by actively seeking to improve package affinity.
- Median latency improves by 66.25%
- Tail latency improves by 38.6%
- Adds (manageable) node imbalance.

Questions?

eboza@fiec.espol.edu.ec

Thanks!



ESPOL
"Impulsando la sociedad del conocimiento"