universität
innsbruck

AutoDAG

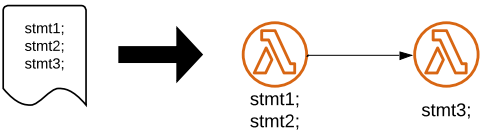
An end-to-end automated platform for the creation and optimization of serverless DAGs in federated clouds

Philipp Gritsch, Simon Triendl, Sashko Ristov

Automating Serverless DAGs

1. **FaaSification:** Assigning tasks to functions

But maybe better configurations exist:

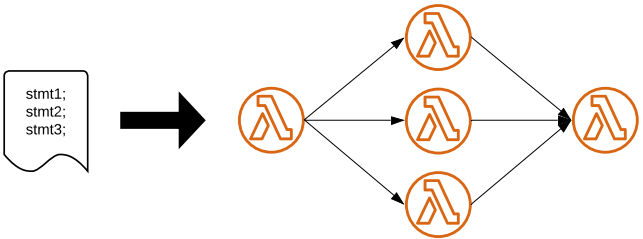


Automating Serverless DAGs

❷ **Optimization:** Assigning optimal memory to each function

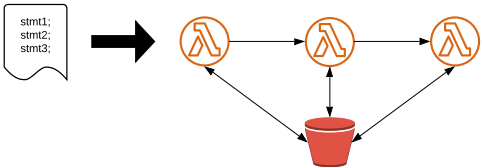
Automating Serverless DAGs

❸ **Profiling:** Recognize opportunities for parallelism

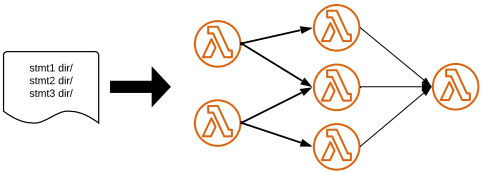


Automating Serverless DAGs

❹ **Profiling:** Derive explicit and implicit dependencies

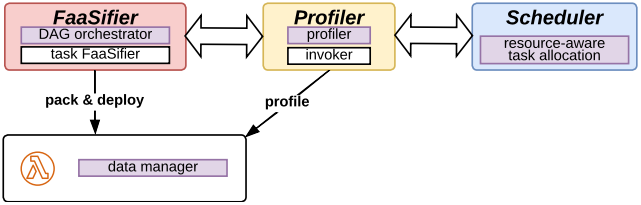


naive: shared storage



better: per function to reduce data transfers

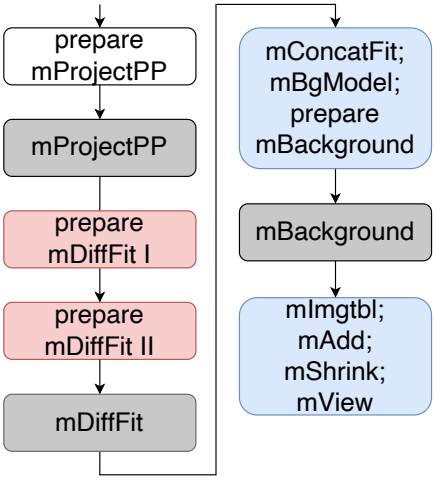
AutoDAG Architecture



FaaSifier: Converts monolithic code into serverless DAGs
Profiler: Profiles file management and data flows
Scheduler: Resource-aware task allocation for optimal cost

Main Innovation
Steps are inter-dependent and require end-to-end support and automation

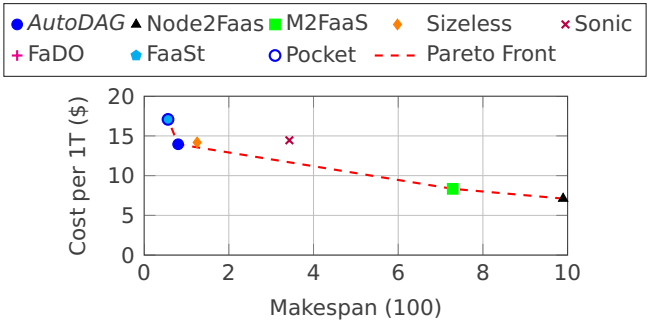
Exemplary Results (Montage on AWS)



Key Insights

- 7 functions (naive allocation: 12)
- `prepareMDiffFit` is not fused because one of its task is very lightweight
- Otherwise, file transfers are reduced
- GCP DAG has one more function, because of GCP's higher bandwidth

Exemplary Results (Montage on AWS)



Key Insights

- *AutoDAG*'s cost: $14.08 (0.93 % error in estimation)
- 12.31× performance improvement over Node2FaaS (no parallelism)
- 1.55× performance improvement over Sizeless (no task allocation)

Thank you for your attention!