universität
innsbruck

Forget the Cold Start! Experience the "Spawn" Start in Serverless Computing

Sashko Ristov, Christian Hollaus, and Mika Hautz
University of Innsbruck, Austria

Outline

- » Introduction
- » Evaluation Methodology
- » Results
- » Conclusion

Spawn Time

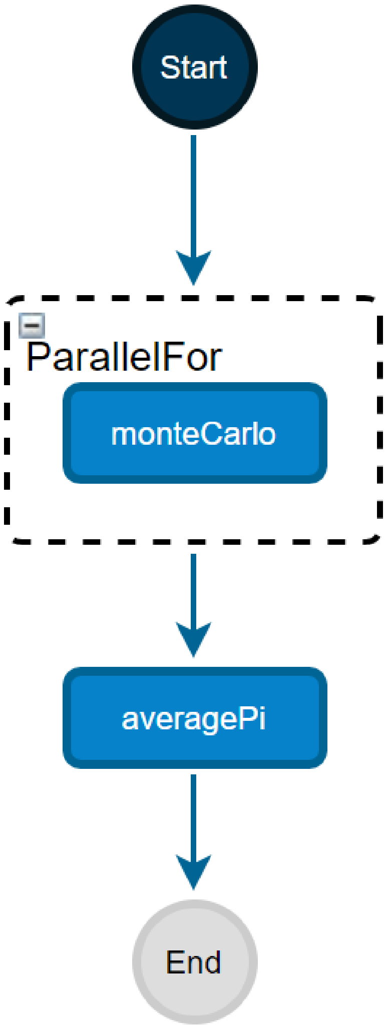
- Considerable *spawn time*

- Spawn **1000 functions** on IBM within **8s**
  - ❖ Josep Sampé, Gil Vernik, Marc Sánchez-Artigas, and Pedro García-López. 2018. Serverless Data Analytics in the IBM Cloud. *Middleware '18. ACM DOI:https://doi.org/10.1145/3284028.3284029*

- Spawn **1000 functions** on AWS within **0.6s**
  - ❖ S. Ristov, S. Pedratscher and T. Fahringer, "xAFCL: Run Scalable Function Choreographies Across Multiple FaaS Systems," in *IEEE Transactions on Services Computing*, doi: 10.1109/TSC.2021.3128137

Evaluation Methodology

- Monte Carlo Serverless Workflow (Function Choreography)
  - Approximates π
  - 100 functions monteCarlo
  - With concurrency 30
  - Measure execution time with SAAF (https://github.com/wlloyduw/SAAF)



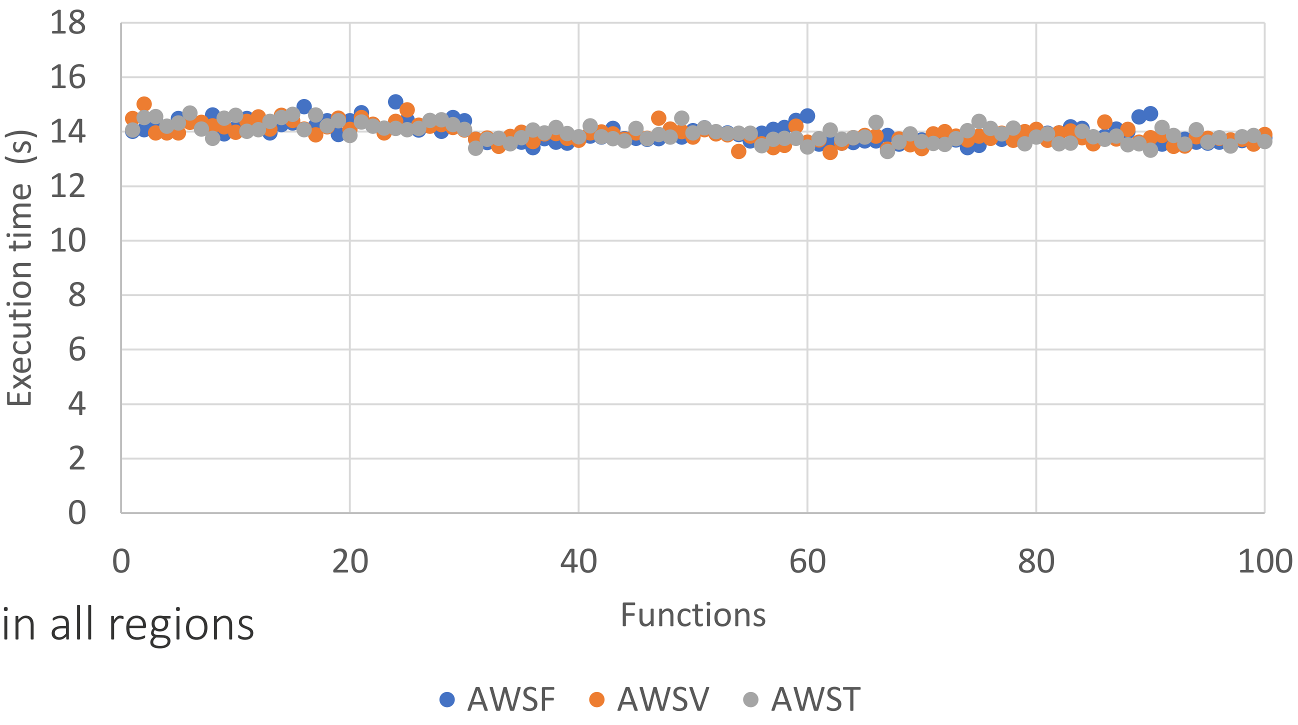
Experiments

- Deployed on 9 regions
  - AWS, IBM, Google
  - EU (Frankfurt), US (Virginia, Washington), Asia (Tokyo)

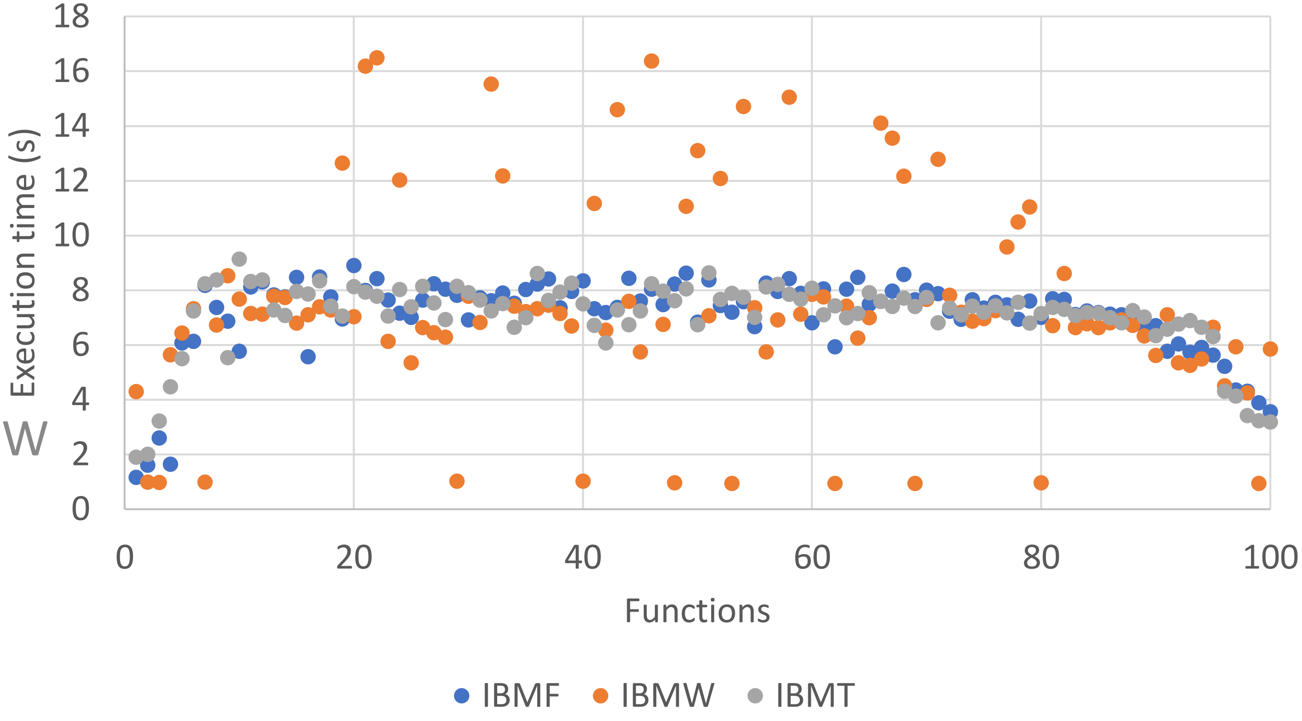
- Run with xAFCL enactment engine
  - S. Ristov, S. Pedratscher and T. Fahringer, "xAFCL: Run Scalable Function Choreographies Across Multiple FaaS Systems," in IEEE Transactions on Services Computing, doi: 10.1109/TSC.2021.3128137
  - https://github.com/sashkoristov/enactmentengine

Results (AWS)

- No spawn start
- Cold start effect
  - Up to 3.5%
  - 480 ms
  - The first 30 functions
- The same pattern in all regions



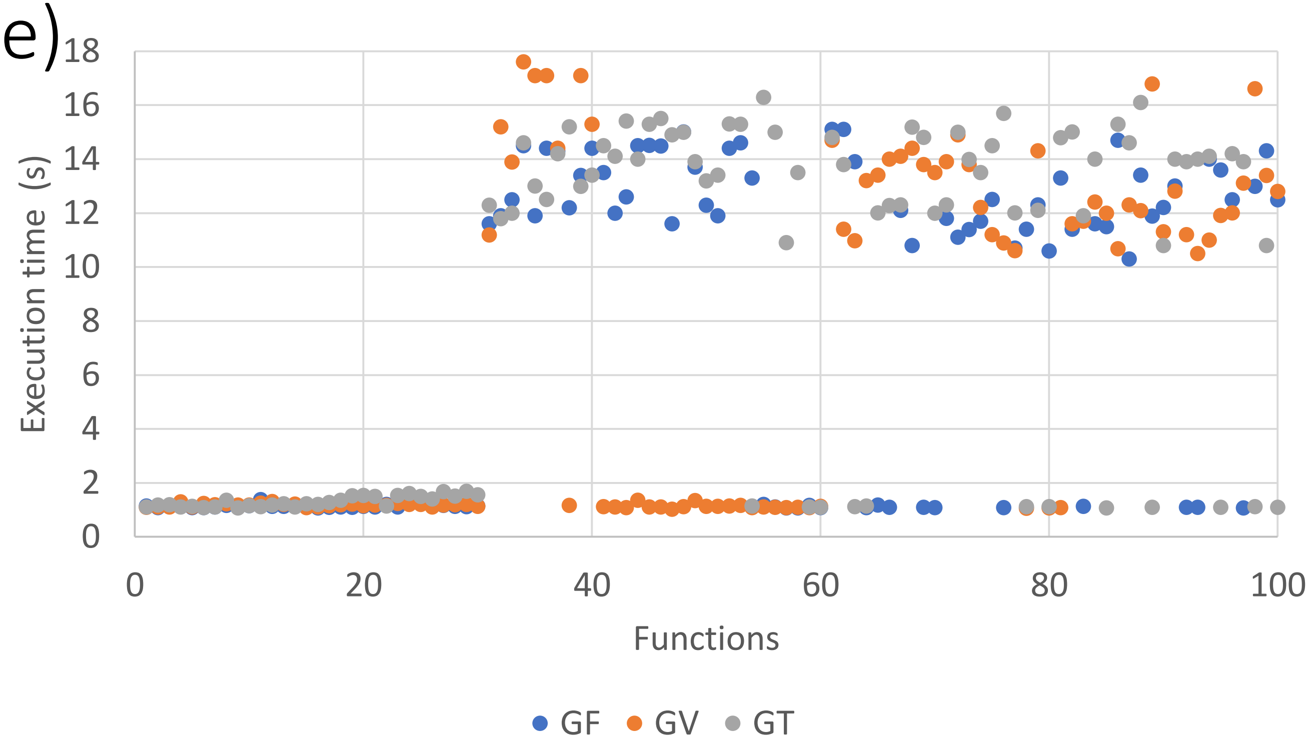
Results (IBM)



- Spawn start effect
  - 6.7x in IBMF
  - 3.8x in IBMT
  - 16.7x in IBMW
  - 7.2 – 15.5s
- The same pattern in all regions
- IBMW provided various CPUs
  - The fastest points in IBMW are due to Xeon(R) Gold 6140
  - But, it does not provide such performance in IBMT!

Results (Google)

- Cold start is HOT!
  - New container = 0
- Spawn start effect
  - 10.5x
  - or 12.6 s

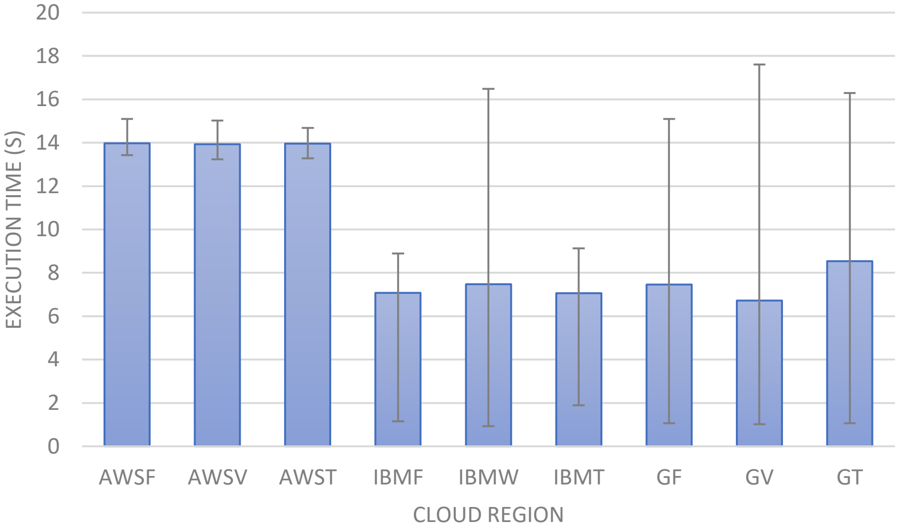


- The same pattern in all regions

Summary Results

- » Google and IBM are fastest for low concurrency

- » But are often slower than AWS for high concurrency



- » IBMF and IBMT are the fastest when spawn 100 functions

Conclusion

- » Classical cold start effects on AWS only
- » Google’s cold start is HOT compared to its spawn start!

- » AWS is not affected by the spawn start
  - But, IBM (15.5s) and Google (12.6s) are affected

- » Be aware, all functions of a parallel loop lie on the critical path
  - Delaying even a single function affects makespan

www.uibk.ac.at