Exploring the Effects of Hardware Heterogeneity on FaaS Performance Variability

Jonas Heisterberg | 3S | Master Thesis | 19. December 2025
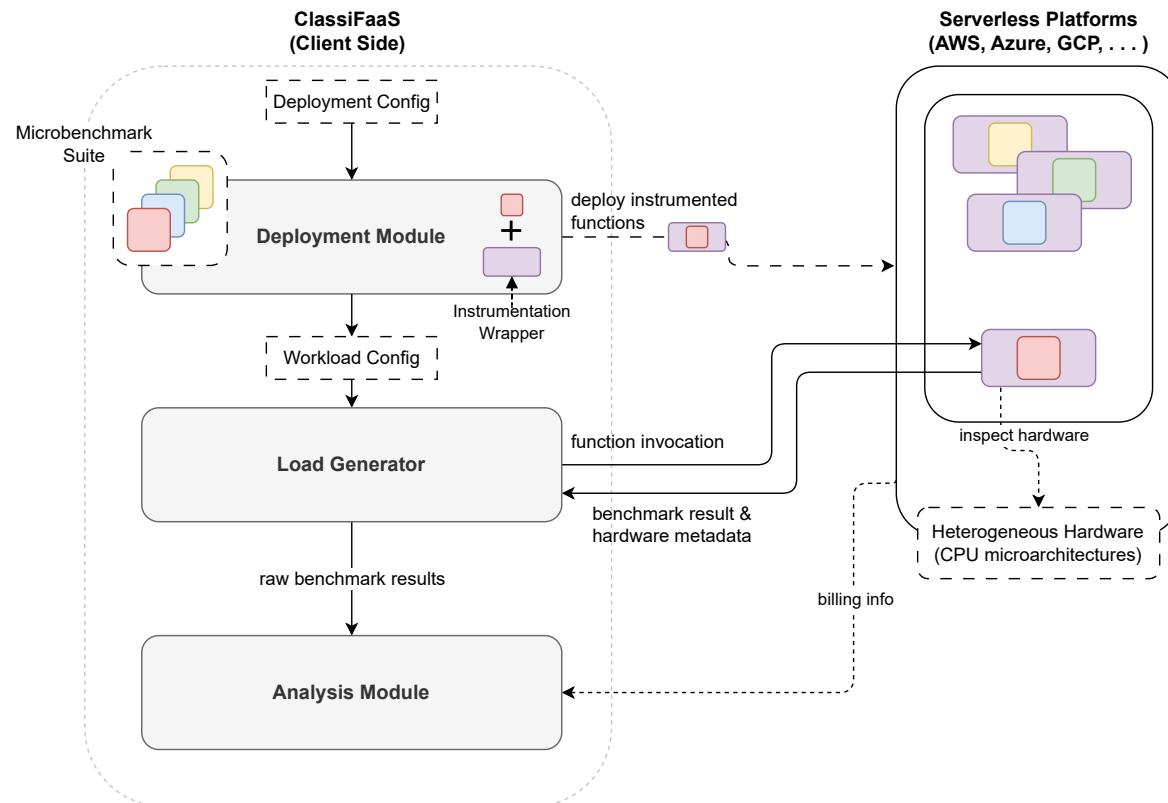
Platzhalter für Sublogo in Weiß

Motivation: Hardware is abstracted, variability is not

− FaaS abstracts infrastructure from developers

− Same function, same config → different performance

− Pay-per-use billing → performance variance = cost variance

− **Hypothesis:** Identical configurations are backed by heterogeneous CPU generations, causing performance variation

− **Goal:** quantify the "hardware lottery" effect on performance

Jonas Heisterberg | 3S | Master Thesis | 19. December 2025

Platzhalter für Sublogo
in Weiß

Benchmark Design

- Workload:
  - Closed model
  - Synthetic microbenchmarks
- Maximize hardware exposure by:
  - Sequential benchmark execution
  - High concurrency
  - Forced instance termination

CPU Identification Across Platforms

Provider	AWS	Google Cloud Platform	Alibaba Cloud	Azure
Service	Lambda	Cloud Functions (gen 1)	Function Compute 3.0	Functions (Flex Consumption)
CPU identification from **/proc/cpuinfo**	**model name** (generic)	**numeric model identifier**	**model name** (sometimes generic)	**model name** (unmodified)
Identifier example	Intel Xeon @ 2.50GHz	85	Intel Xeon 2.50GHz **or** Intel Xeon Platinum 8269CY 2.50GHz	Intel Xeon Platinum 8370C @ 2.80GHz
Distinct CPU types observed	**5**	**6**	**5**	**3**
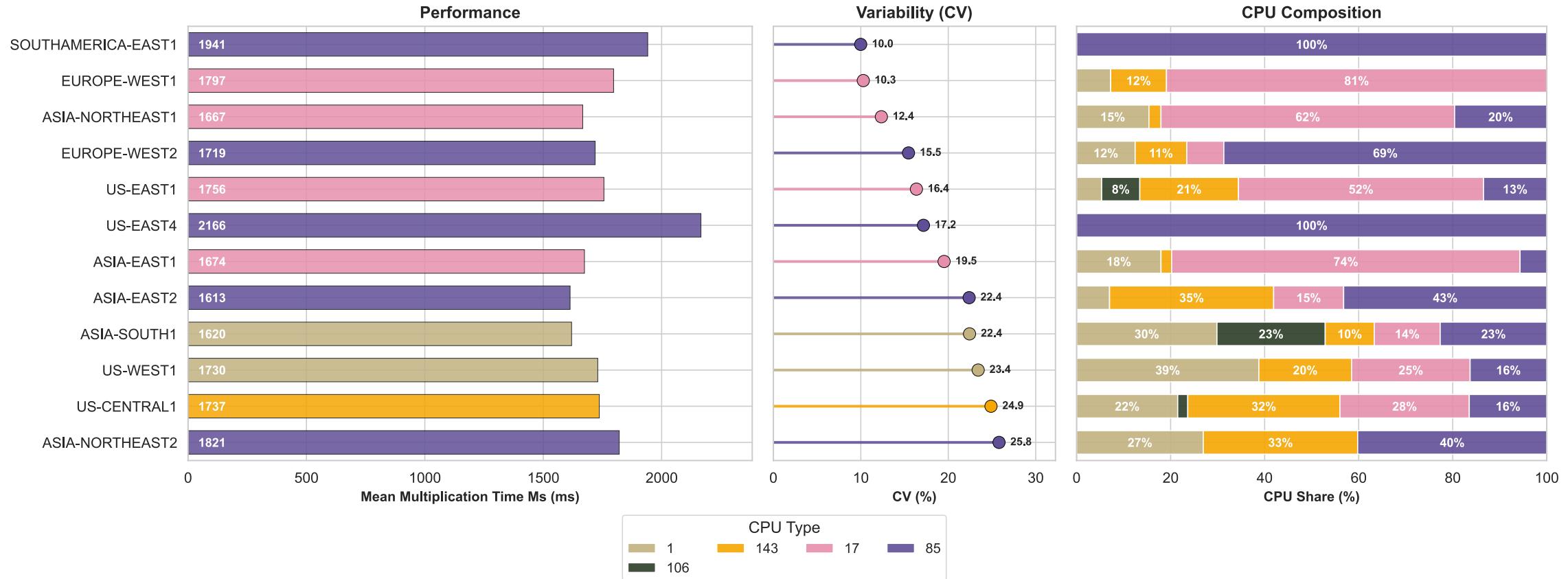
# Experimental Design: Progressive Stages

- Stage A — Regional Heterogeneity Baseline

  - Investigate CPU diversity across regions (12 regions per provider)

- Stage B — Memory Configuration Effects

  - Assess the impact of configuration choice (1 region x 3 configurations)

- Stage C — Temporal & Workload Effects

  - Explore hardware assignment over time (1 region x 4 times a day x 7 days)

  - Quantify impact on different workloads (5 distinct workloads)
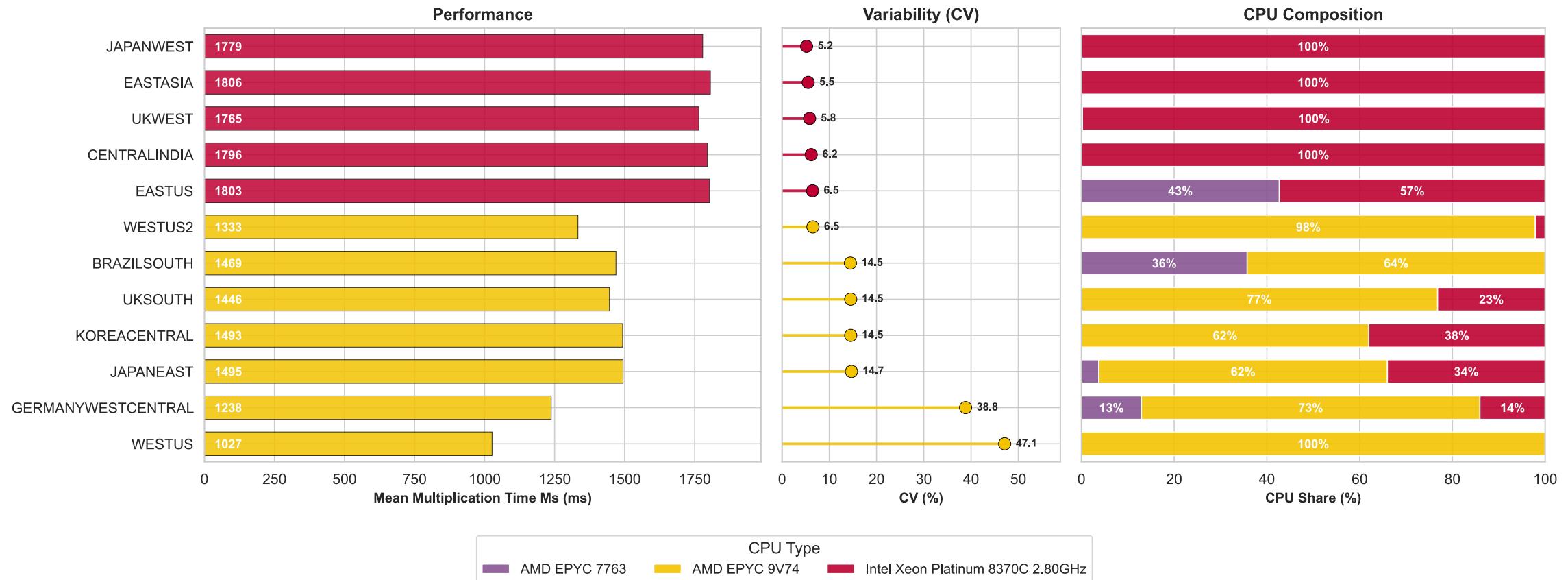
GCP - Regional CPU mix and performance variability (Matrix Multiplication, 512 MB)
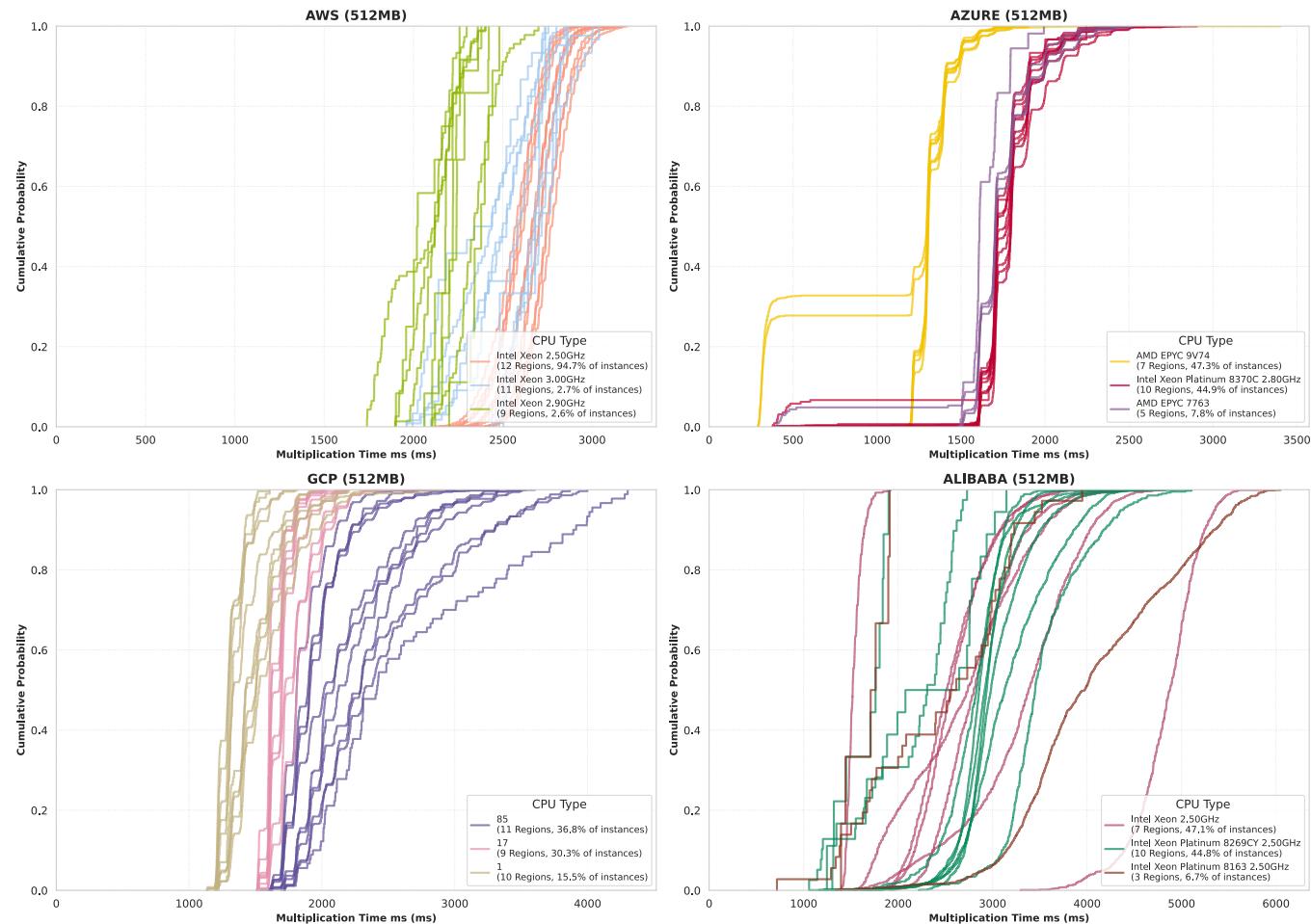
Jonas Heisterberg | 3S | Master Thesis | 19. December 2025

AZURE - Regional CPU mix and performance variability (Matrix Multiplication, 512 MB)

Jonas Heisterberg | 3S | Master Thesis | 19. December 2025
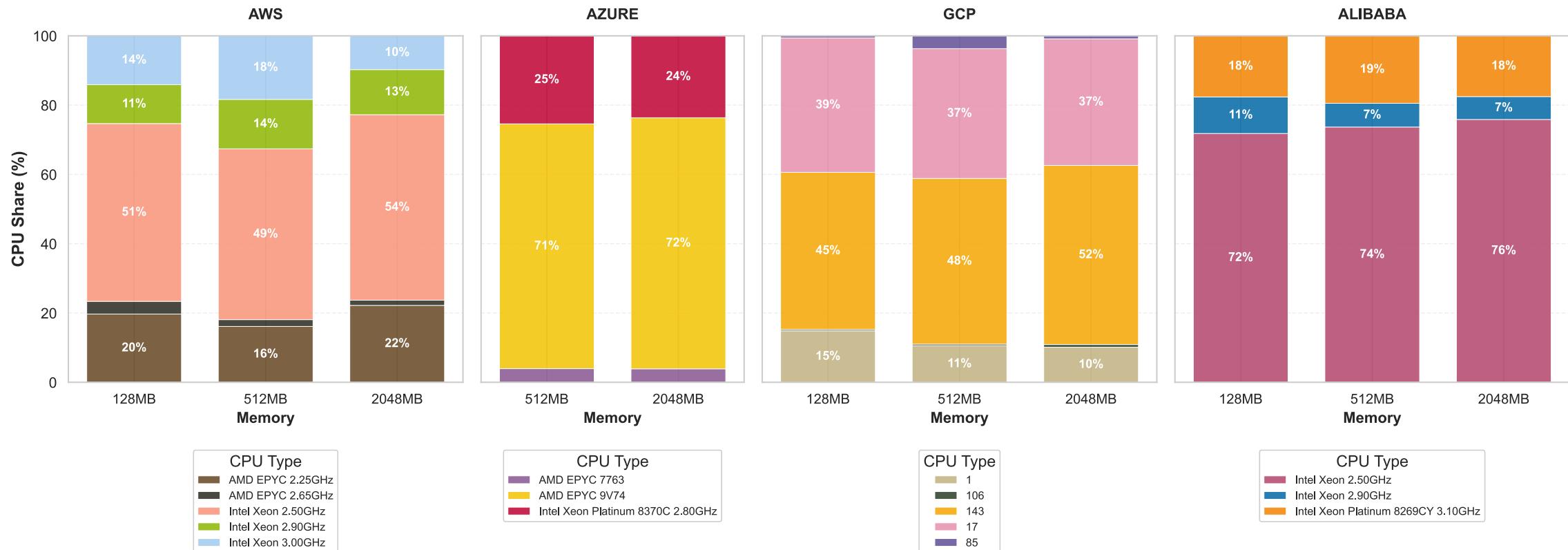
Stage A

ECDF - Regional performance consistency by CPU (Matrix Multiplication)

Jonas Heisterberg | 3S | Master Thesis | 19. December 2025
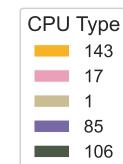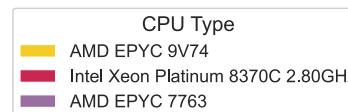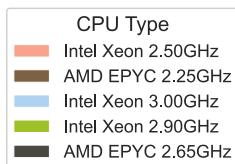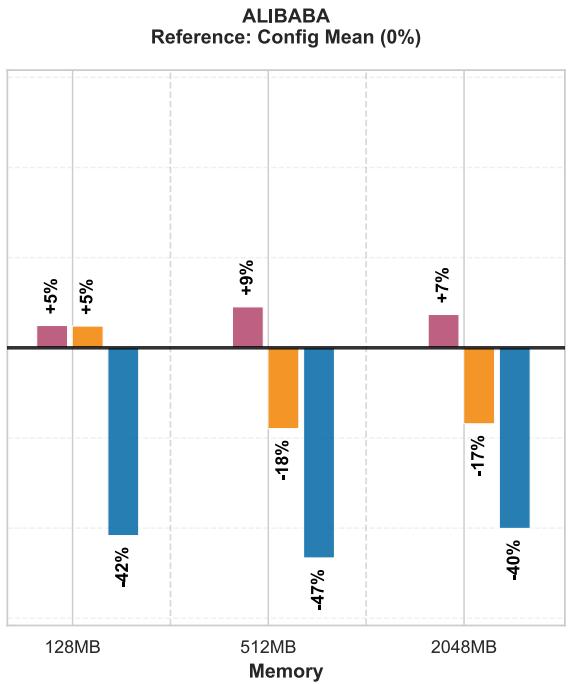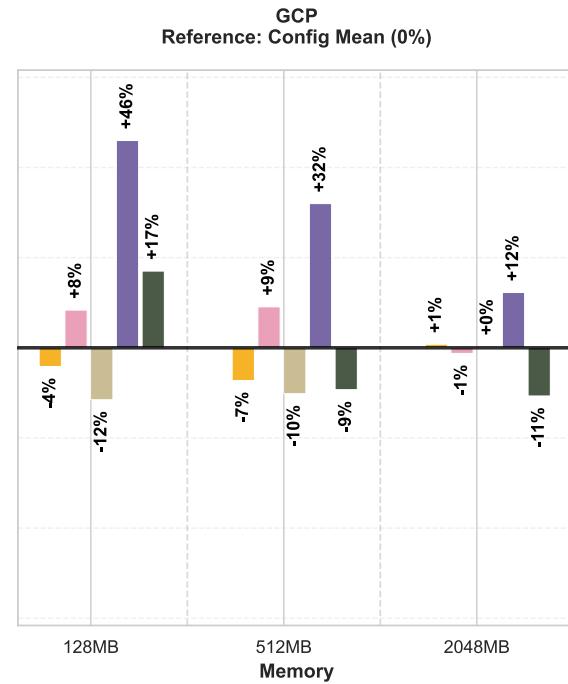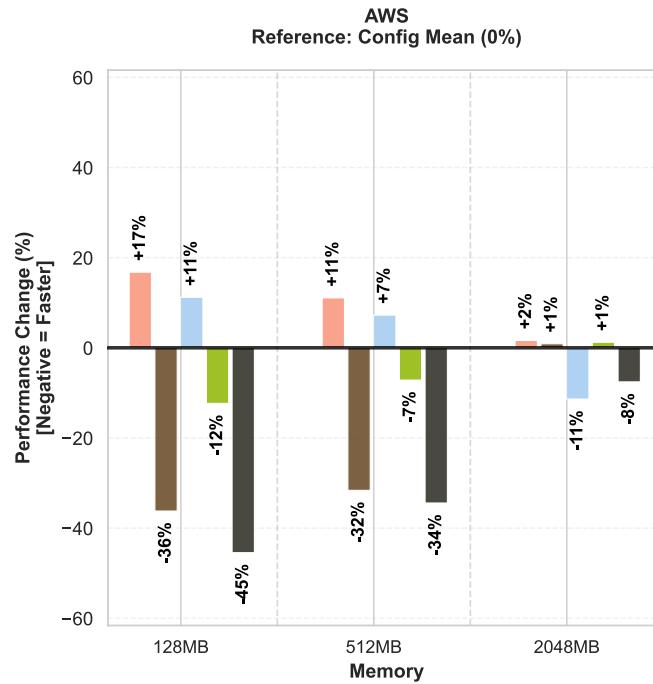
CPU mix by memory size (Matrix Multiplication)

Jonas Heisterberg | 3S | Master Thesis | 19. December 2025
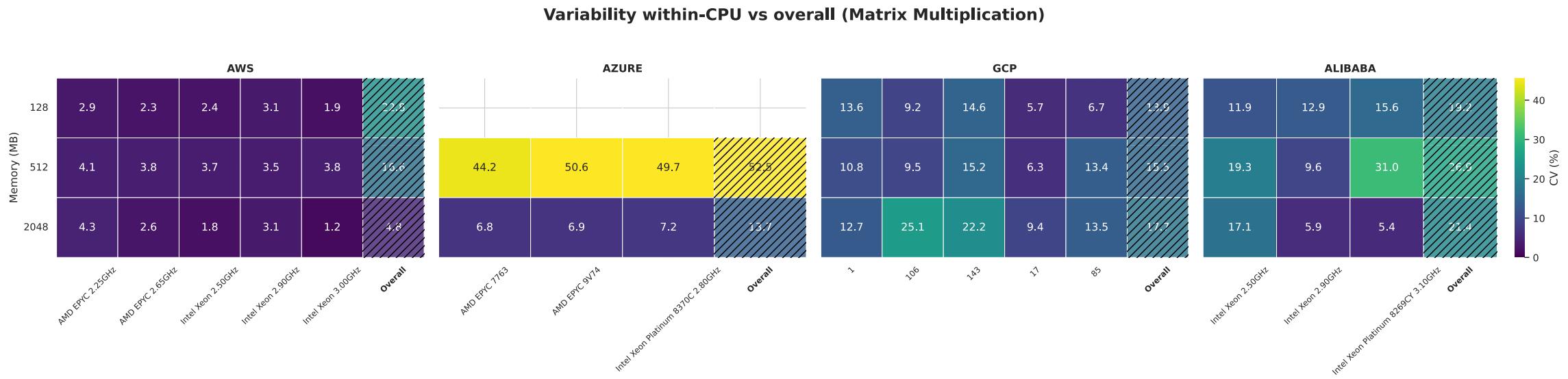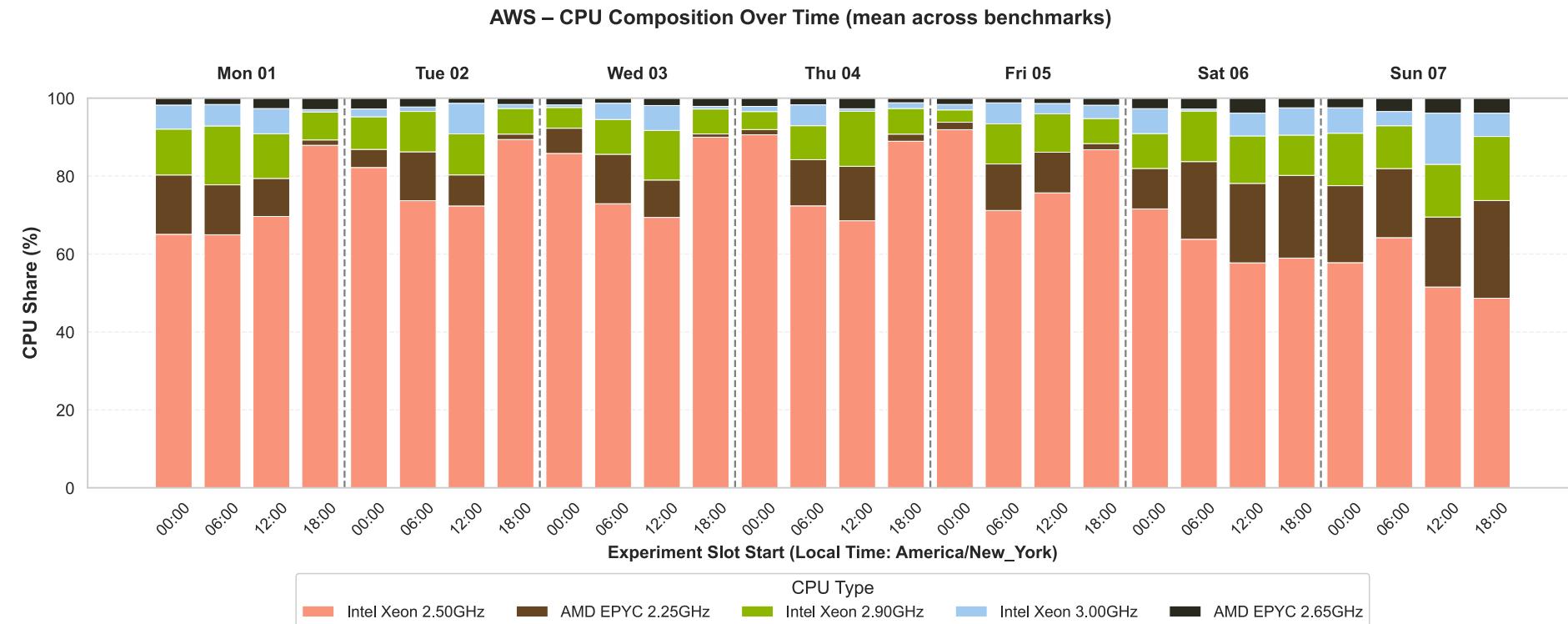
CPU performance deviation from config mean (Matrix Multiplication)

Jonas Heisterberg | 3S | Master Thesis | 19. December 2025

Stage B

Variability within-CPU vs overall (Matrix Multiplication)

Jonas Heisterberg | 3S | Master Thesis | 19. December 2025

Stage C

AWS – CPU Composition Over Time (mean across benchmarks)
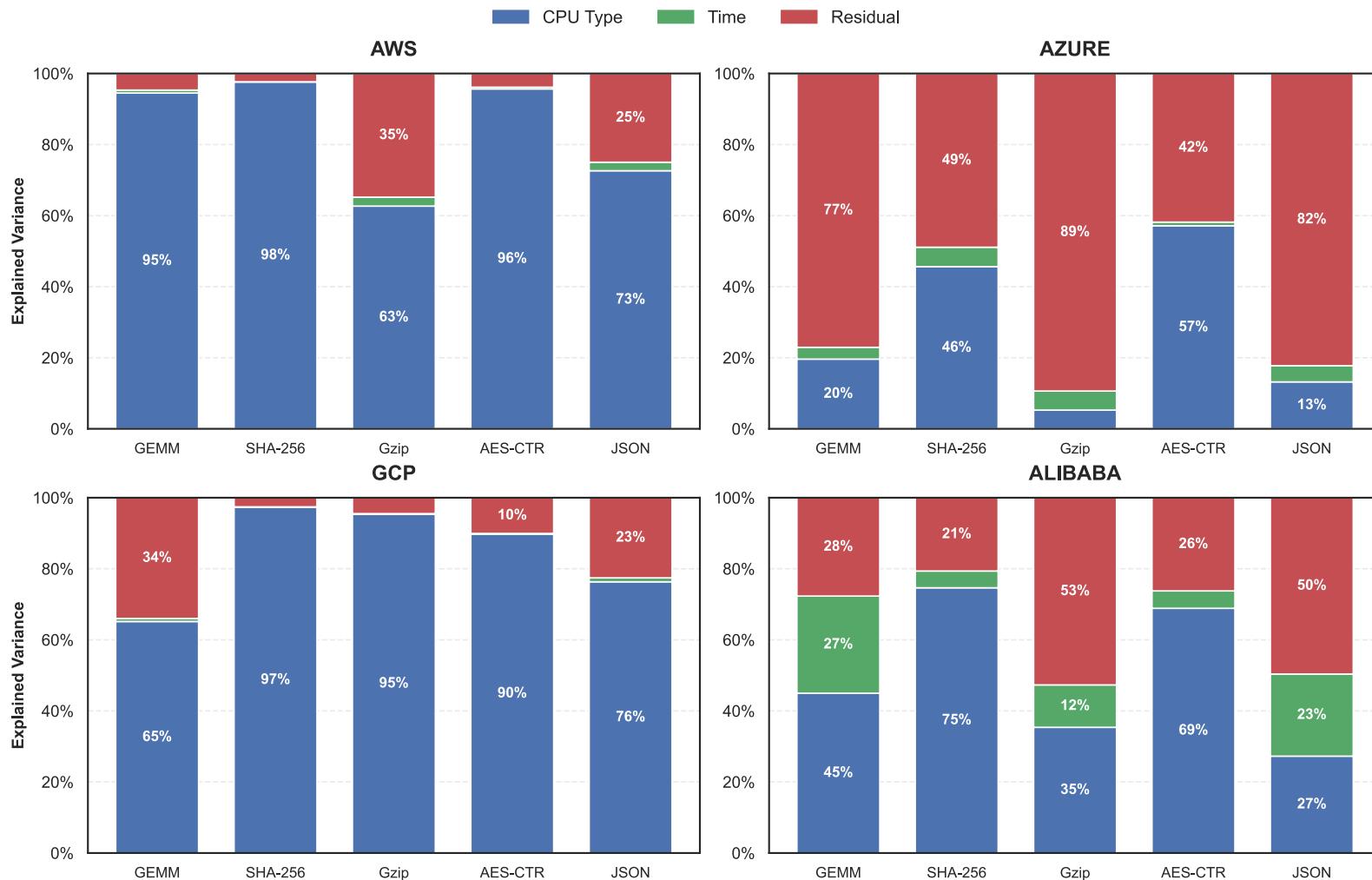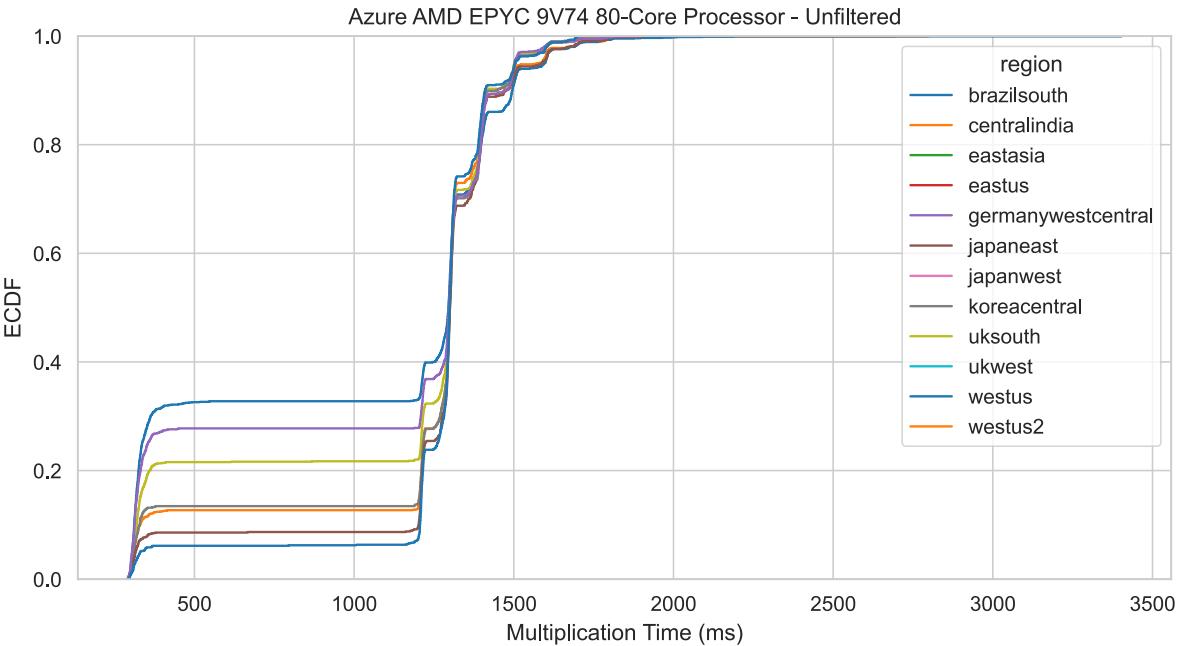
- All Platforms stable except AWS Lambda

Jonas Heisterberg | 3S | Master Thesis | 19. December 2025
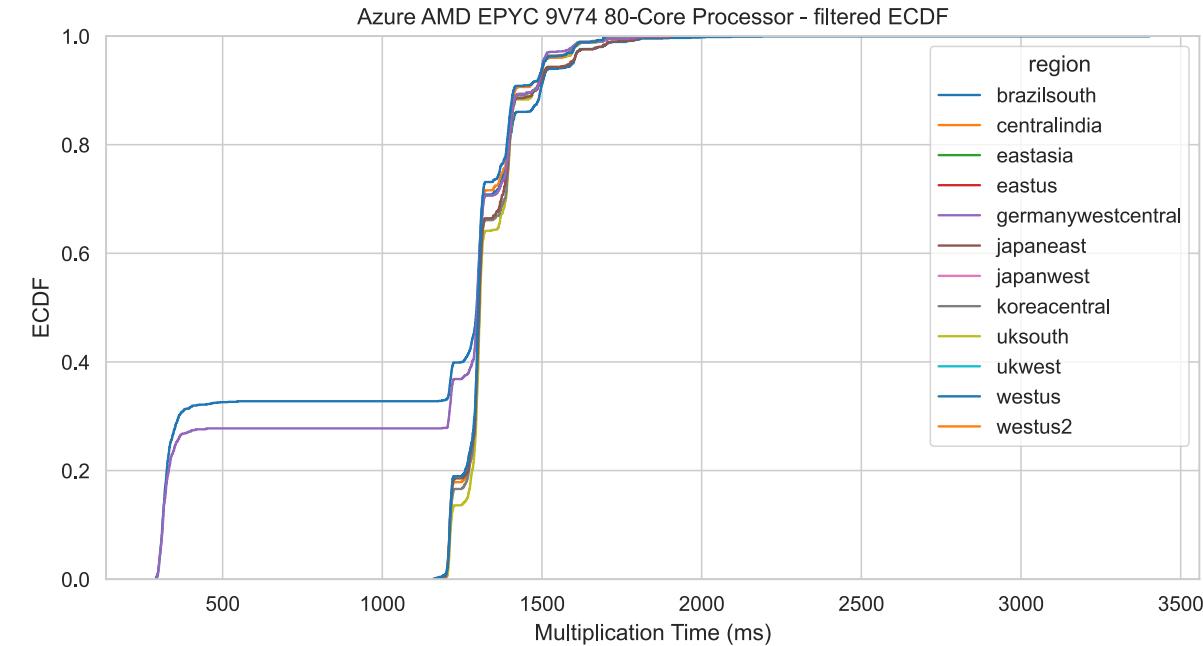
Variance Decomposition: 512MB RAM
(Type-II ANOVA)

Jonas Heisterberg | 3S | Master Thesis | 19. December 2025

Problem: Tukey Outlier Filtering

Azure AMD EPYC 9V74 80-Core Processor - Unfiltered

Azure AMD EPYC 9V74 80-Core Processor - filtered ECDF

Raw Data

After Tukey Filter

Jonas Heisterberg | 3S | Master Thesis | 19. December 2025