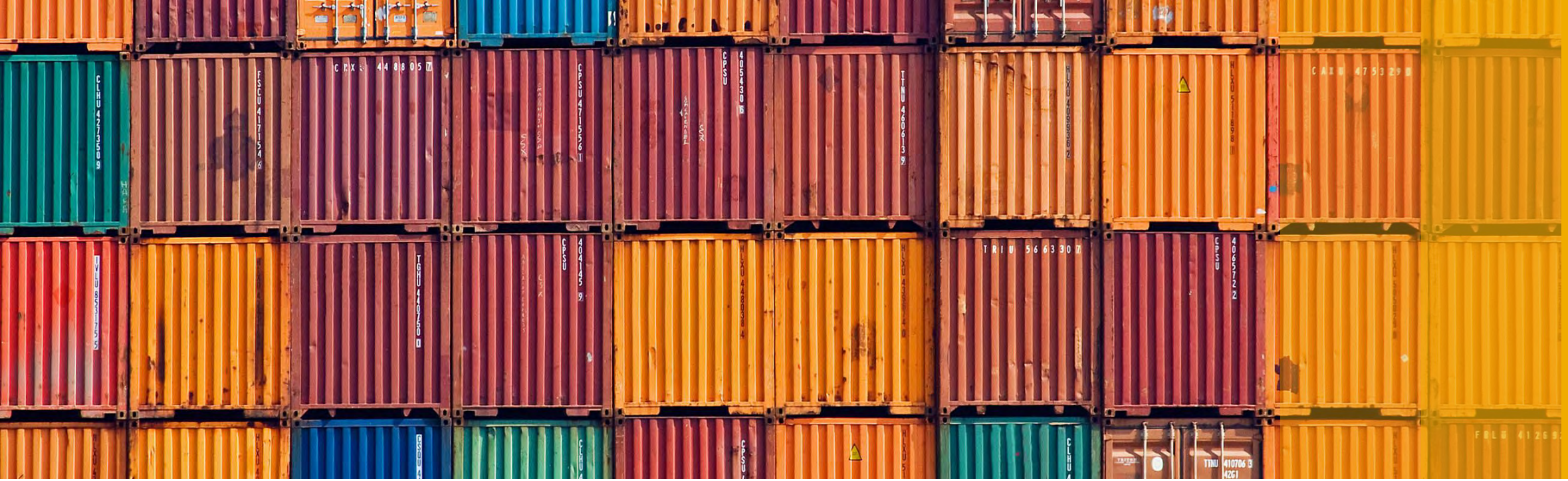


Performance of Containerized Database Management Systems

Kim-Thomas Rehmann, Enno Folkerts, SAP SE

7th International Workshop on Testing Database Systems (DBTEST)

Agenda

OS containers and databases

Comparing container configurations

Container resources

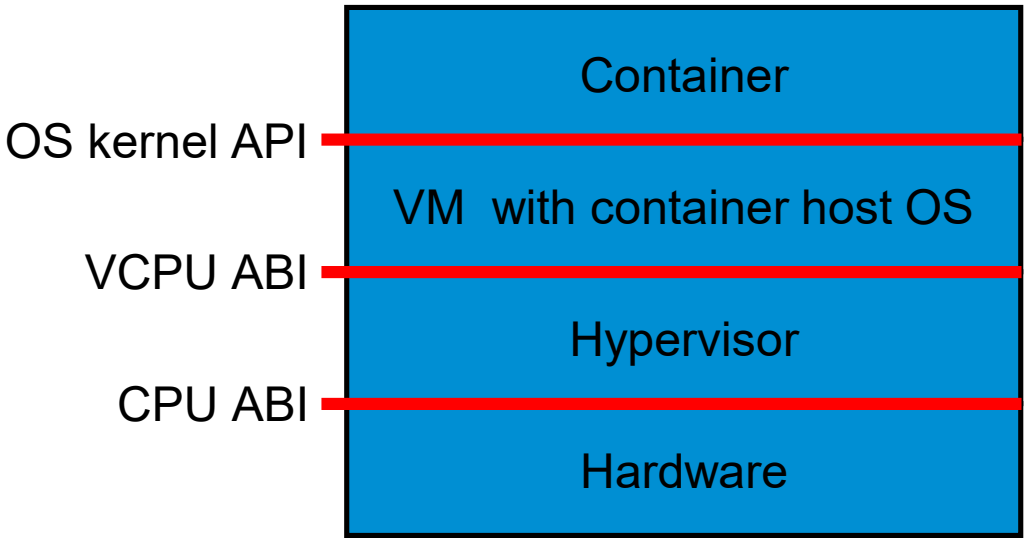
Operations

Evaluation

Containers are the new VMs

Virtualization at OS-kernel level instead of hardware level

- More lightweight because of higher abstraction level
- Software deployed as container image, which includes the userland OS
  - Several layers, the top layer is mutable but volatile
  - Persistent storage must be mapped into the container
- Largely independent of hardware → kernel API and resource assignments are relevant
  - Partitioning or virtualization, to colocate containers
  - Isolation of collocated containers is possible but hard to achieve
- Container orchestration (application execution e.g. by Kubernetes)
  → cross-OS compatibility



Comparing container configurations

Container-based deployment of software in the cloud

IaaS and container orchestration have various configuration options

- E.g. CPU allocation, storage class and QoS, network

How to compare these options?

Contribution

- Testing methodology to compare performance of different configurations under a wide range of workload
- Exemplified with Docker container engine and SAP HANA DB

Container resources

CPU

- Two options to assign hardware threads, dynamically changable
  - By total number (cpus)
  - By selection of threads (cpusets)

Volatile memory

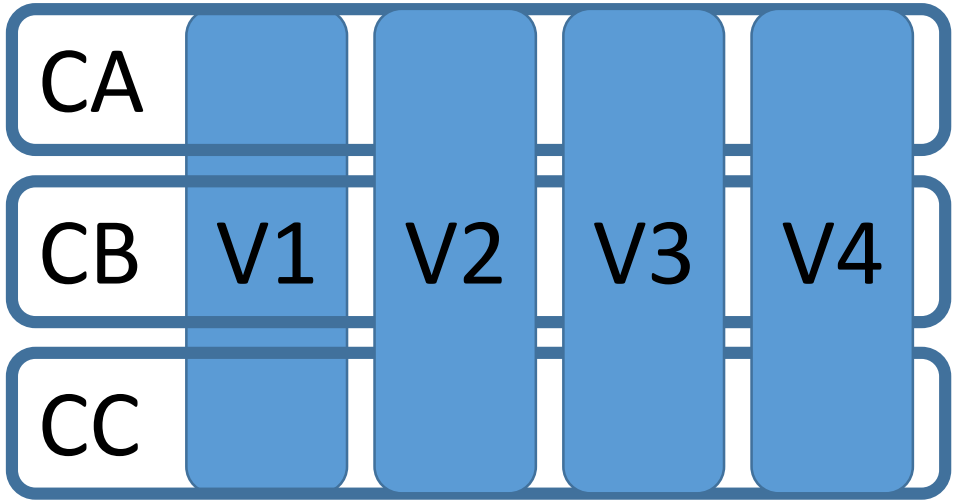
- Overprovisioning not allowed to avoid OoM hazard
- Should take NUMA into account

Persistency

- Persistent filesystem must be mapped from outside

Network

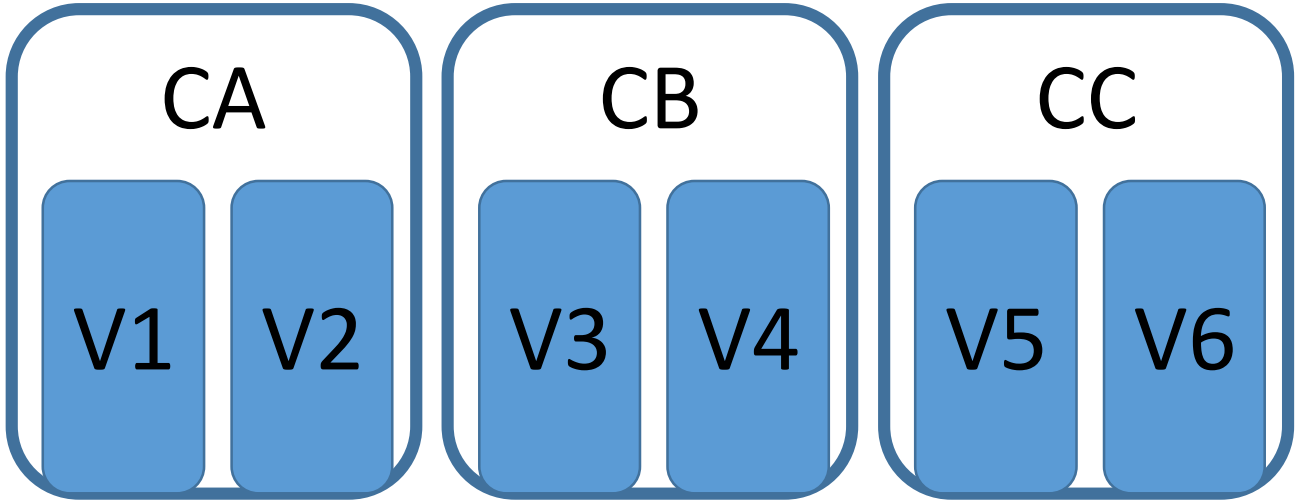
Operations: persistency striping for log volume



Shared volumes

Parallel accesses allow for increased performance

Low fragmentation



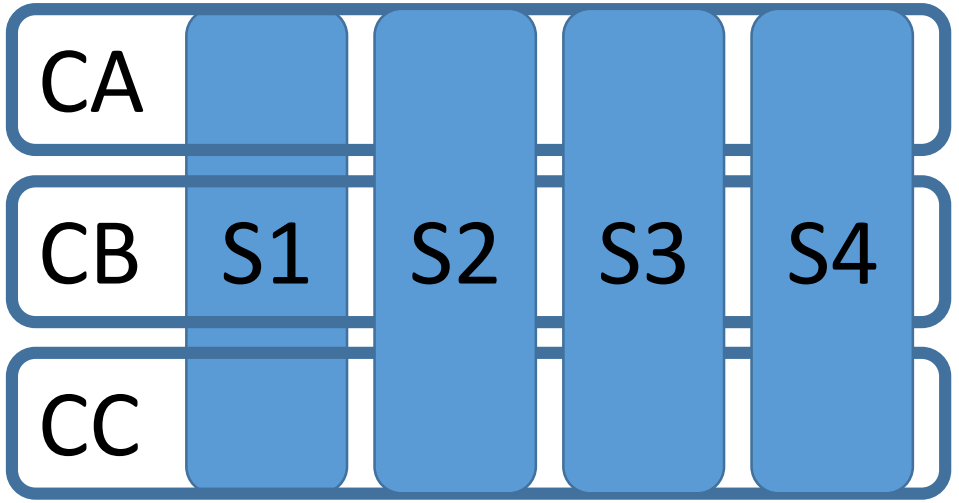
Dedicated volumes

Physical separation of data

Lower performance with reduced number of stripes

Higher storage consumption and fragmentation

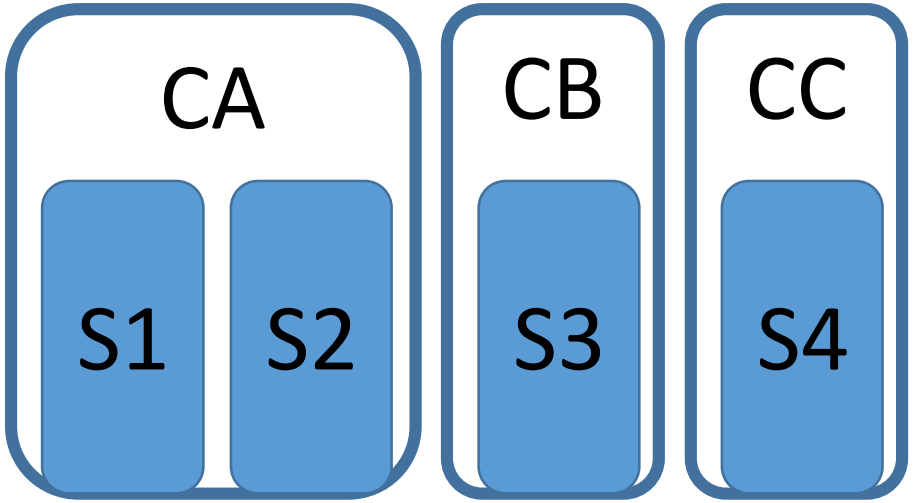
Operations: CPU elasticity



cpus: shared sockets

Easy to overprovision

No control over NUMA placement



cpusets: pinned sockets

Partitioning, no simple overprovisioning

Supports NUMA-aware database

Operations: container migration

Hardware issues, maintenance, changing resource demand → container to move to different host

VMs usually support live migration, but container live-migration is not generally supported

Database can be transferred with near-zero downtime using system replication

- Steps: add secondary, replicate data, takeover, remove secondary (former primary)

If system replication is too expensive, external persistency needs to be switched

Depending on storage configuration

- Dedicated volumes can be re-mounted on target host
- Shared volumes need to be copied

Performance testing methodology

Aim

- Simple comparison of container configuration options
- Based on diverse workload
- In isolation as well as in presence of noisy neighbors

Test suite

- Approx. 700 measurement points reporting either CPU time, elapsed time or cycle time
- OLTP: 5 queries on 1 or two tables with 100 clients
- OLAP: 4 queries on up to 38 tables
- 95% of measurement points have a coefficient of variation less than 5%

How to get an overview of the comparison of 700 measurement points μ?

min(μ):= min(avg(μ,A), avg(μ,B))

delta(μ):= avg(□,B) – avg(μ,A)

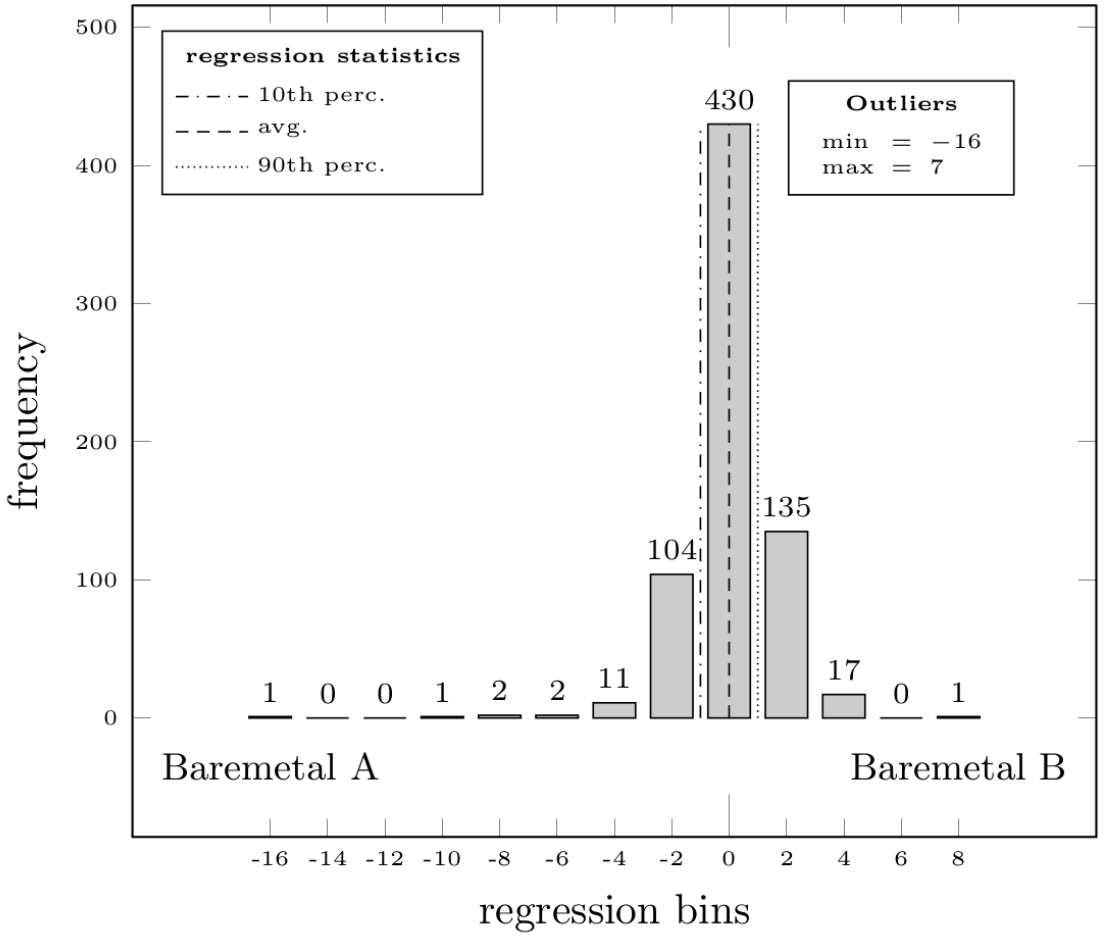
regression(μ):= 100 * *delta*(μ) / *min*(μ)

Procedure

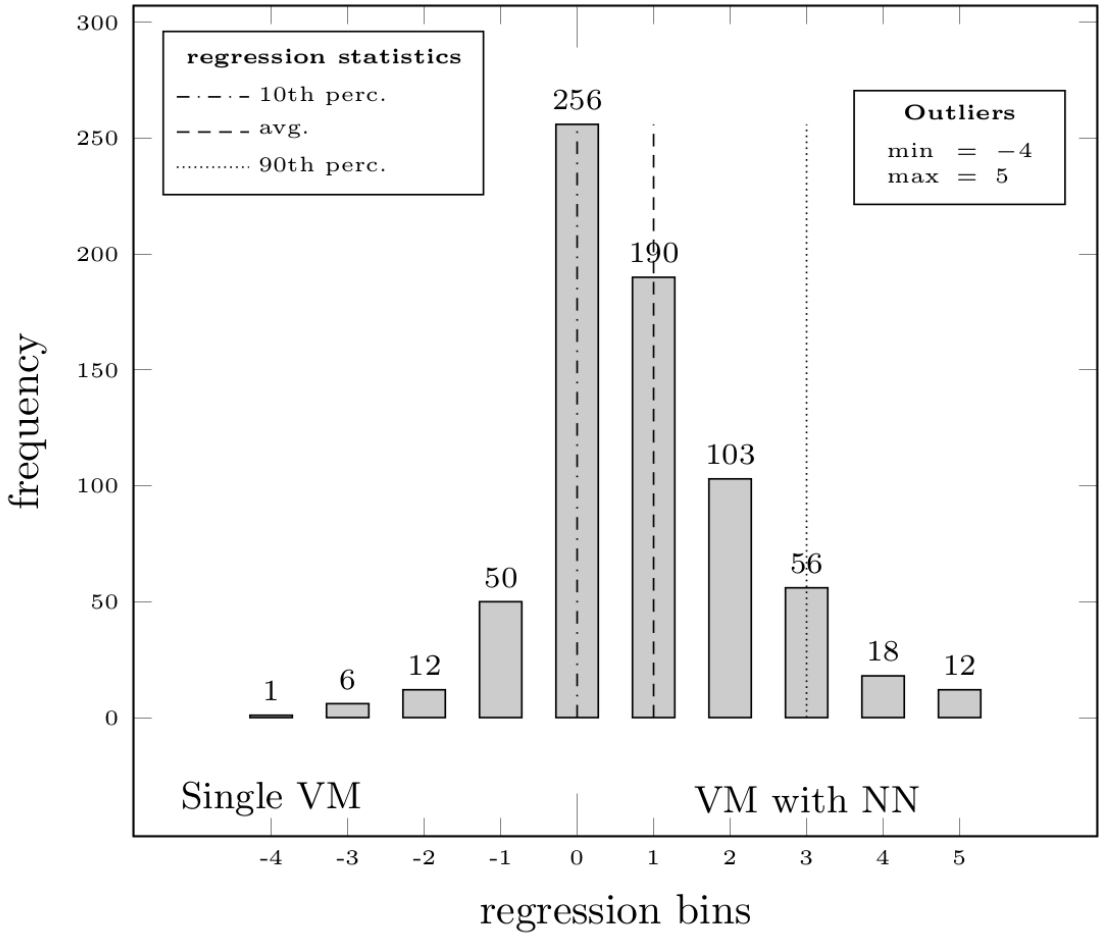
- Create bins of regression values
- Display the bins as histogram, together with mean and percentiles
- Skew shows if and how much the configurations differ

Evaluation: baremetal baseline and VMs with noisy neighbor

baremetal

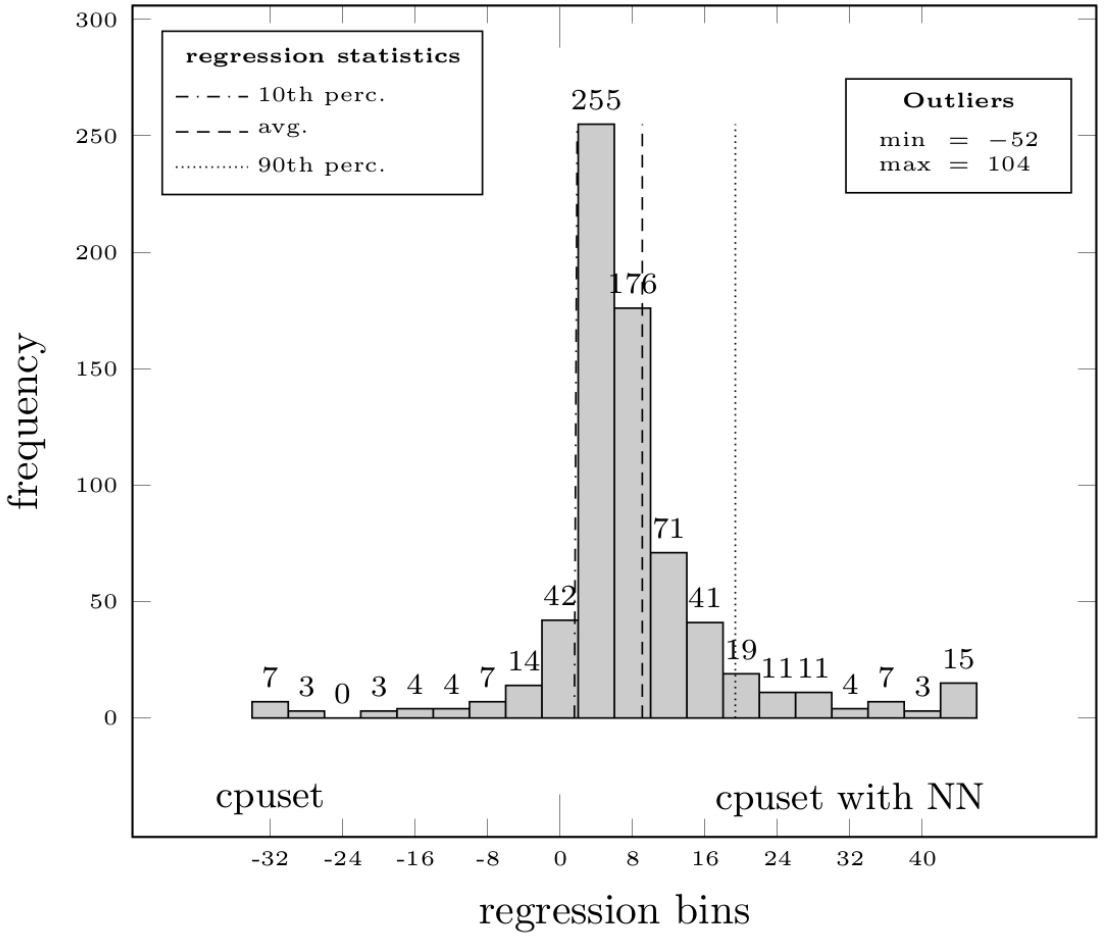


VMs without/with noisy neighbor

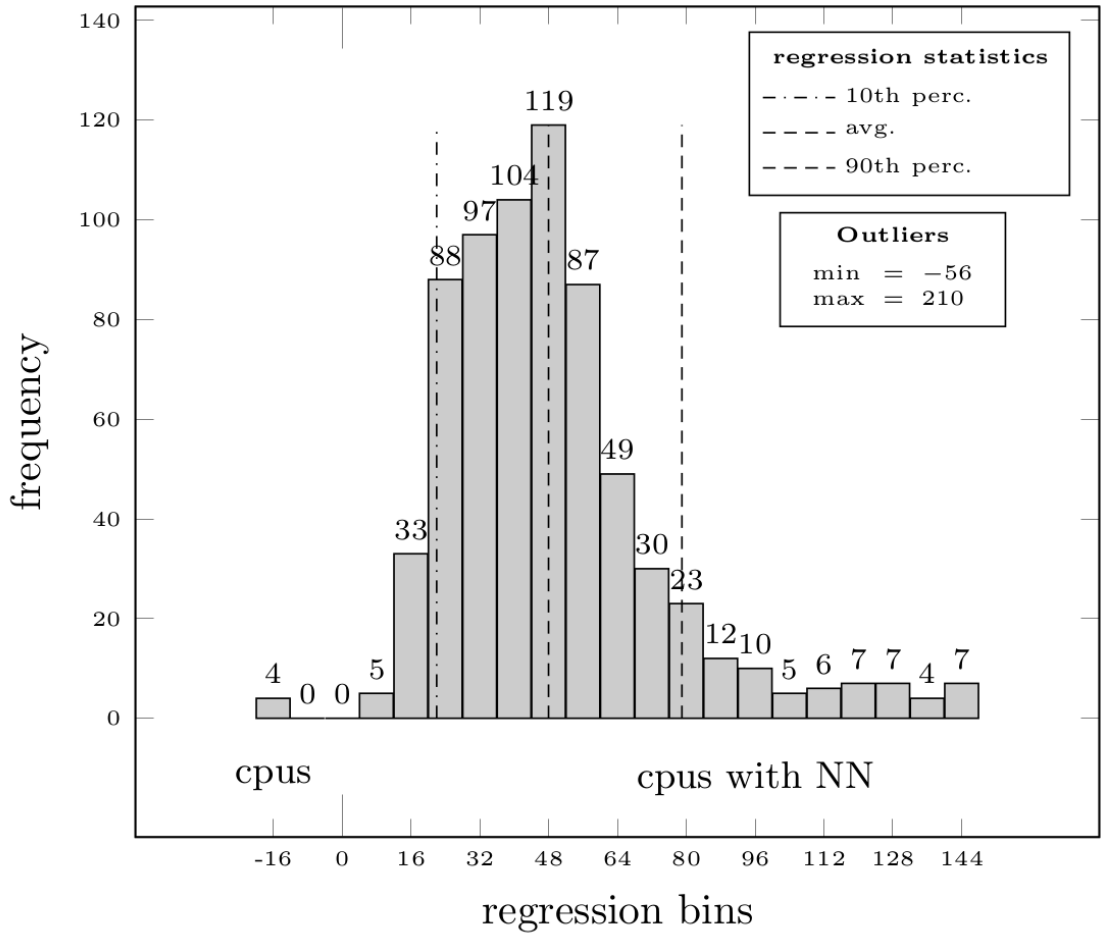


Evaluation: CPU assignment (impact of noisy neighbor)

cpuset (pinned CPUs)

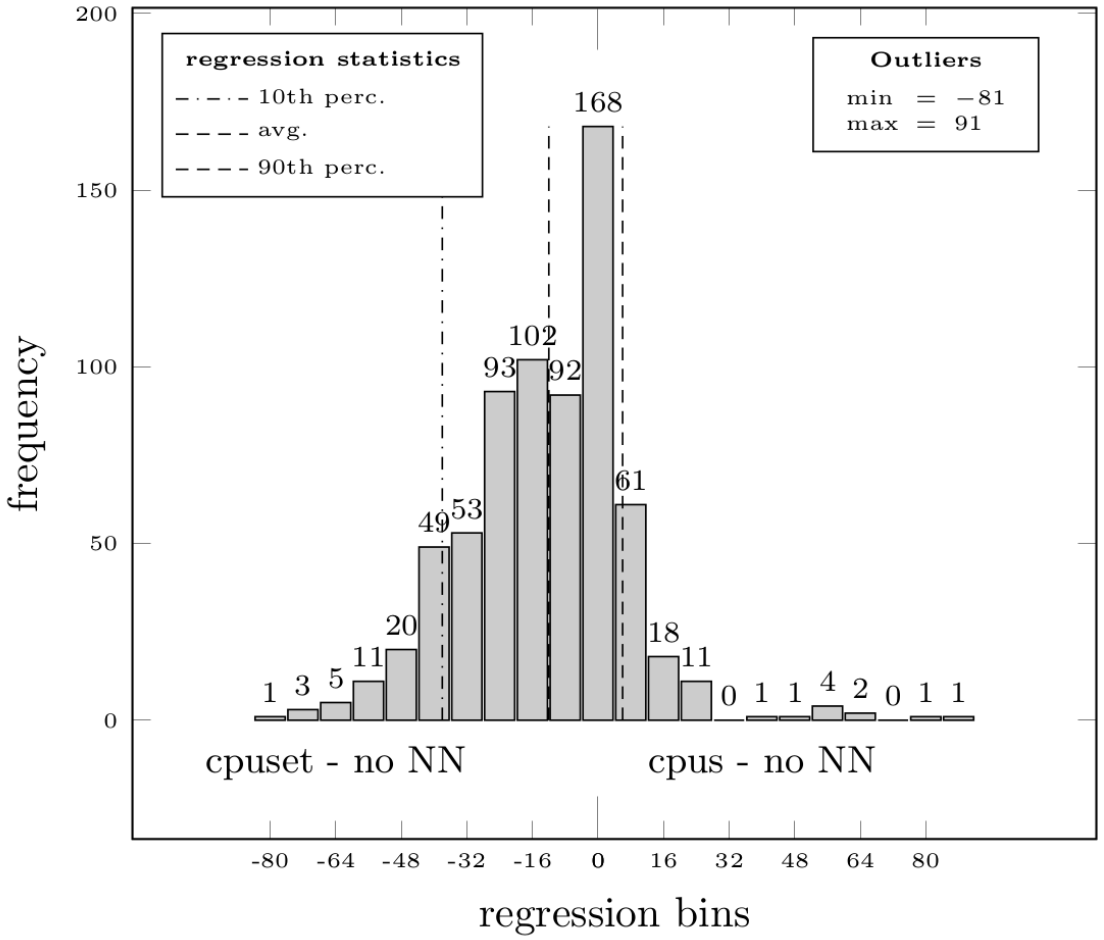


cpus (shared sockets)

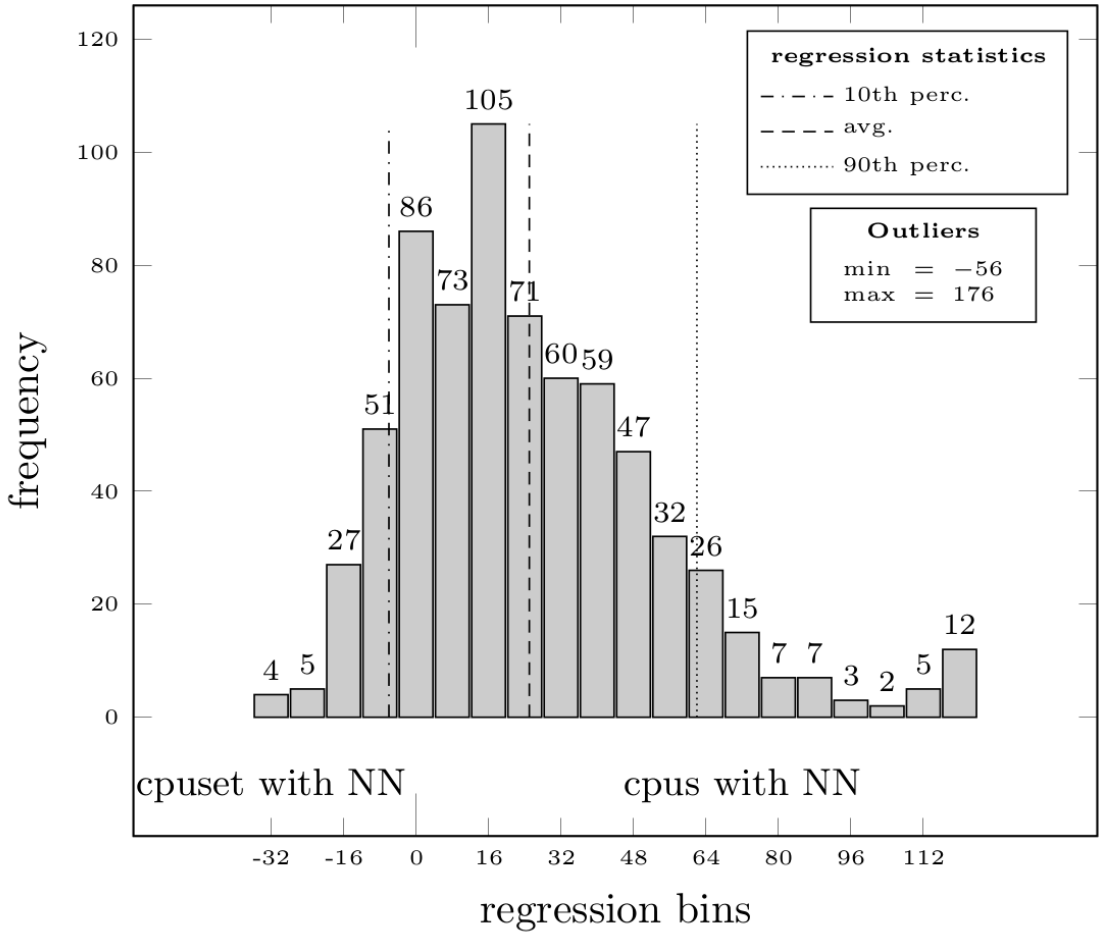


Evaluation: CPU assignment (cpuset vs. cpus)

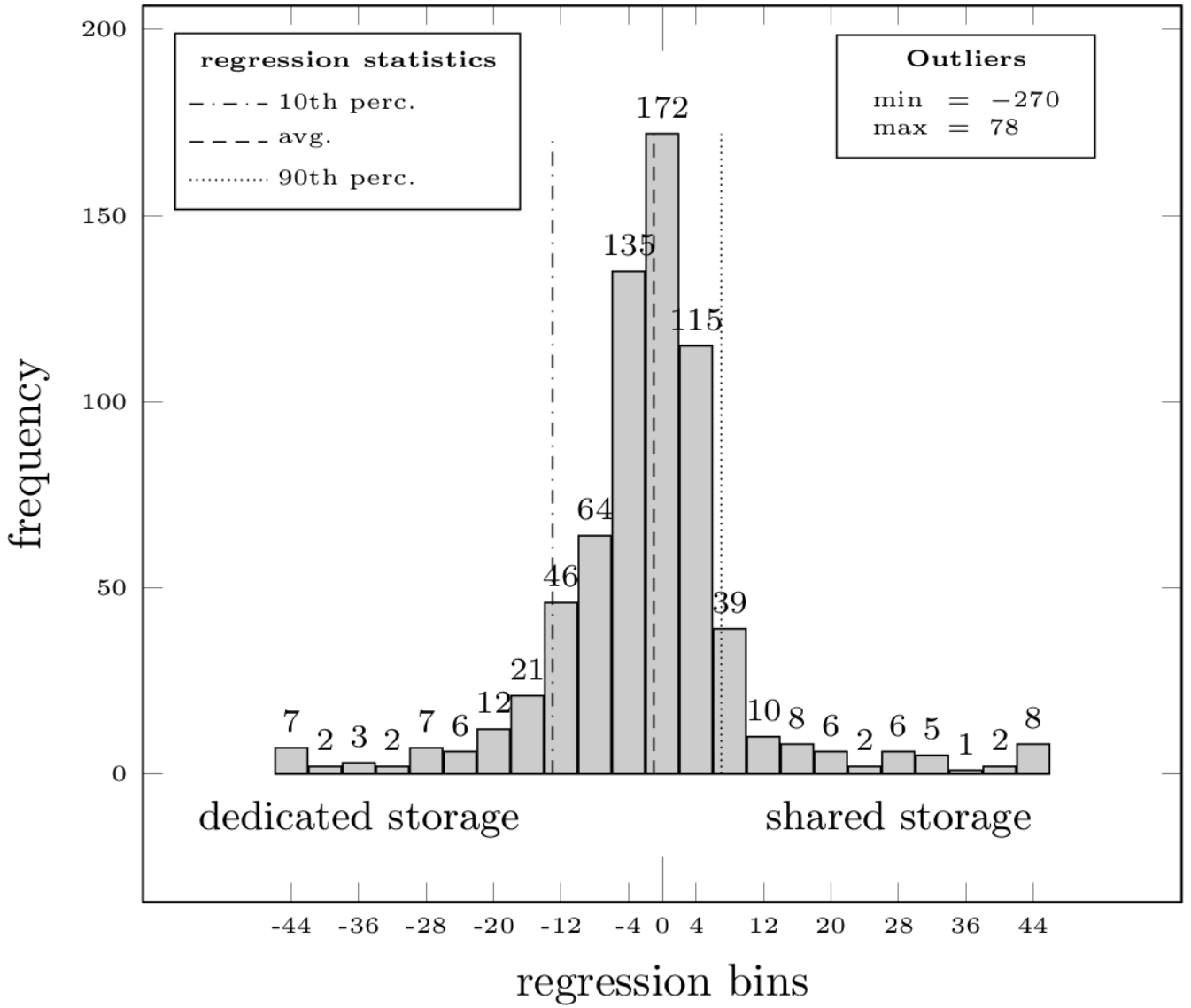
without noisy neighbor



with noisy neighbor



Evaluation: storage striping (dedicated storage vs. shared storage)



Thank you.

Contact information:

Kim-Thomas Rehmann, SAP SE
Senior Developer
Dietmar-Hopp-Allee 16, 69190 Walldorf
+49 6227 749096
Kim-Thomas.Rehmann@sap.com

THE BEST RUN SAP®