What the Racoo

SUPERFAST JVM STARTUP





HUNDREDSOF

THOUSANDS OF FOSS PROJECTS...



JAVAC COMPILER

BYTE CODE

azul

SOURCE CODE

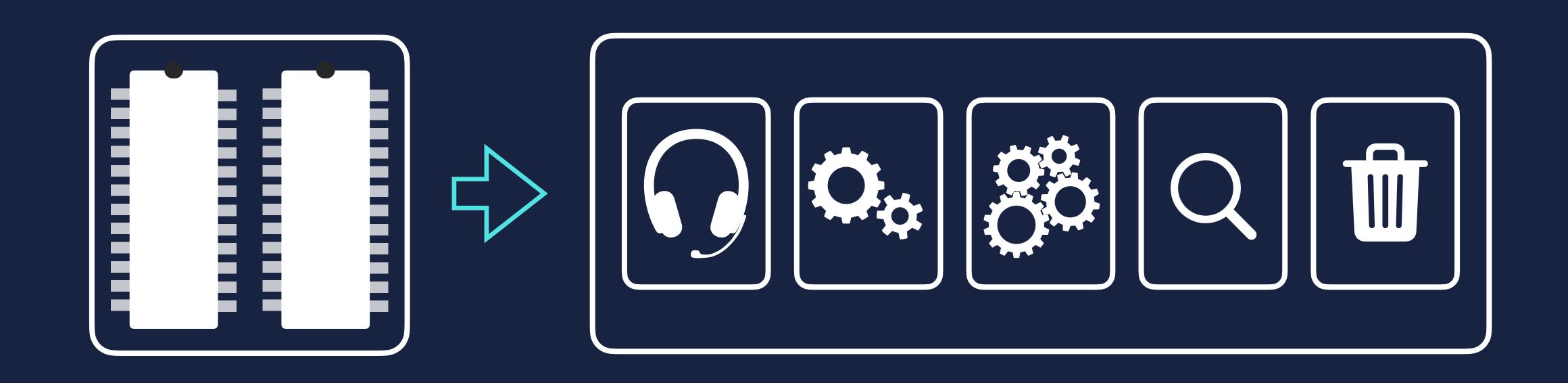


CLASS LOADERS

JVM MEMORY

azul

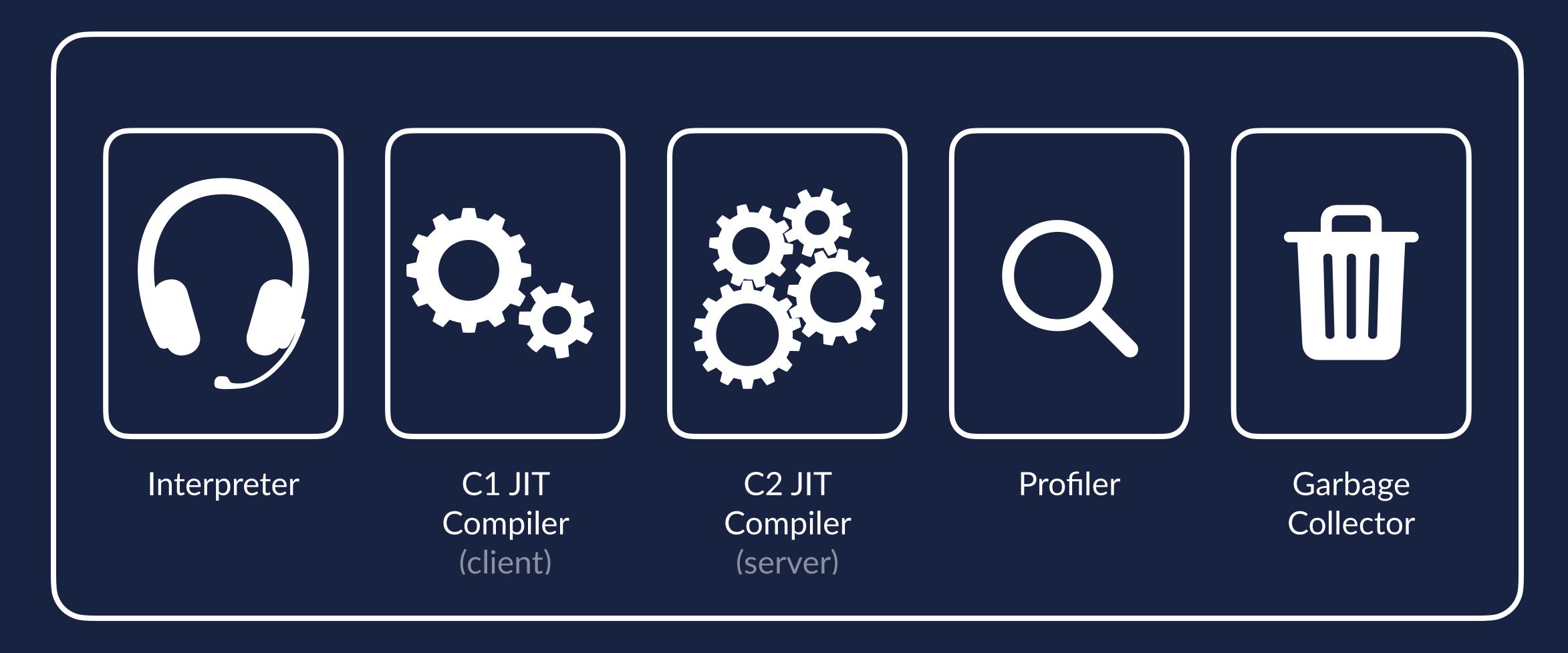
BYTE CODE



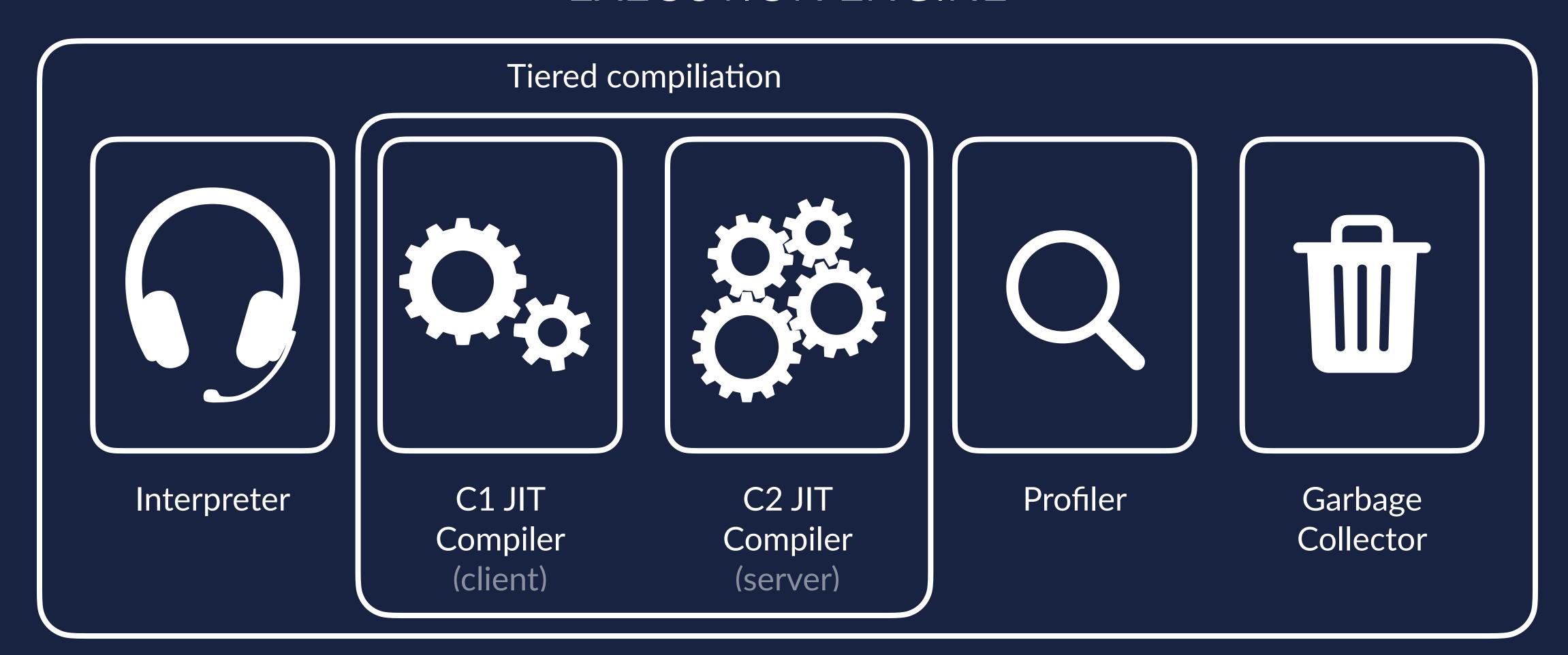
JVM MEMORY

EXECUTION ENGINE

EXECUTION ENGINE

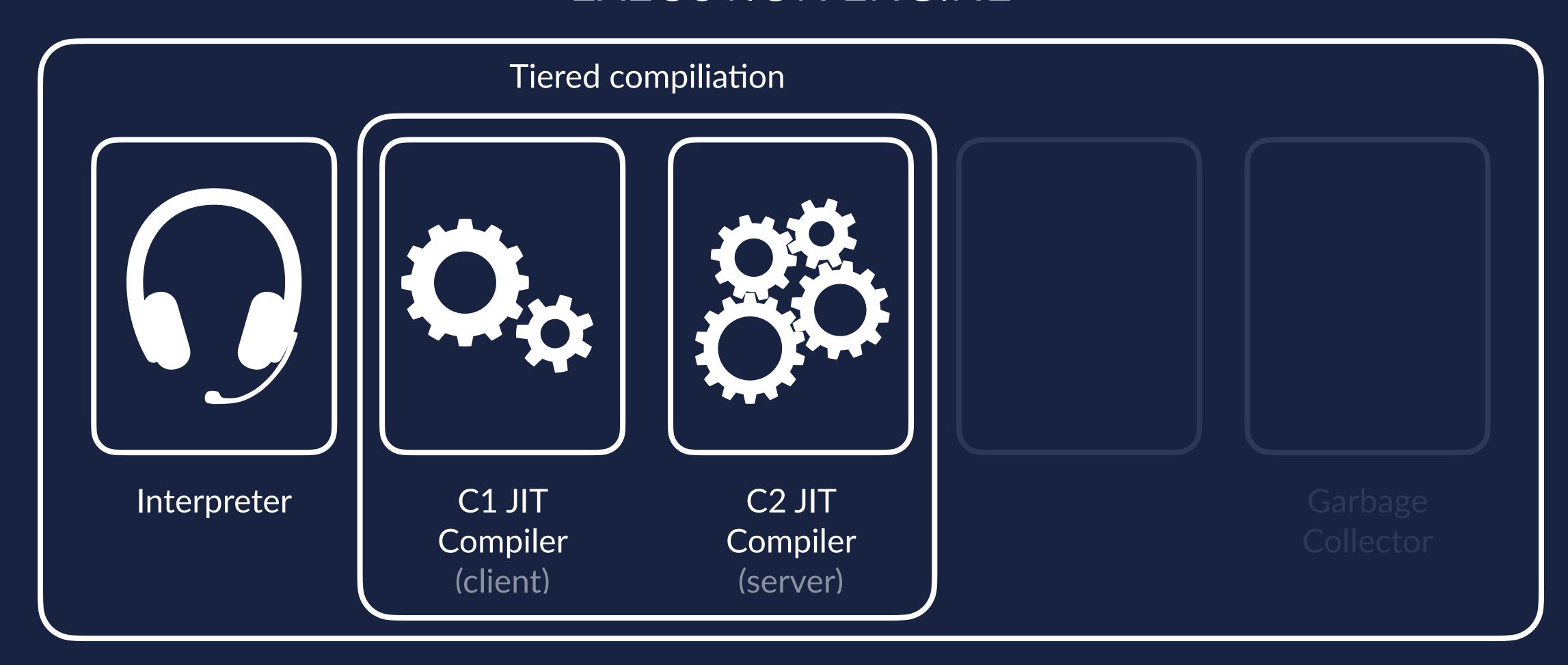


EXECUTION ENGINE



DEFAULT SINCE JDK 8

EXECUTION ENGINE



DEFAULT SINCE JDK 8

Converts ByteCode into instruction set of CPU



INTERPRETER

THRESHOLD
REACHED
(1000 in JDK 17)

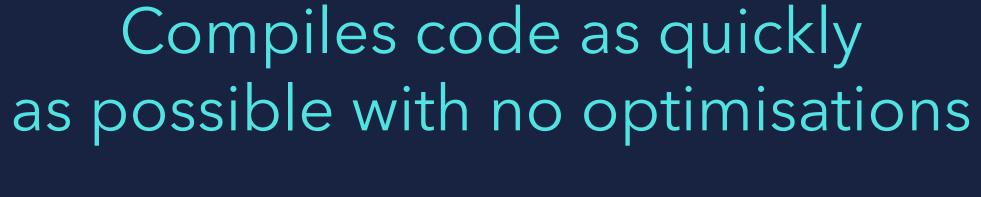
Detects hot spots by counting method calls and loop back edges

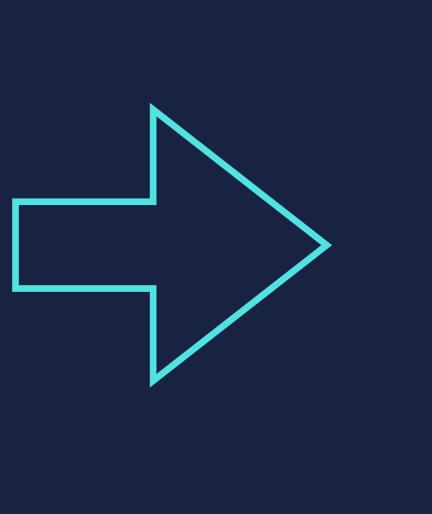


JVV

Pass the "hot" code to C1 JIT Compiler









COMPILER

Compiles code as quickly as possible with no optimisations



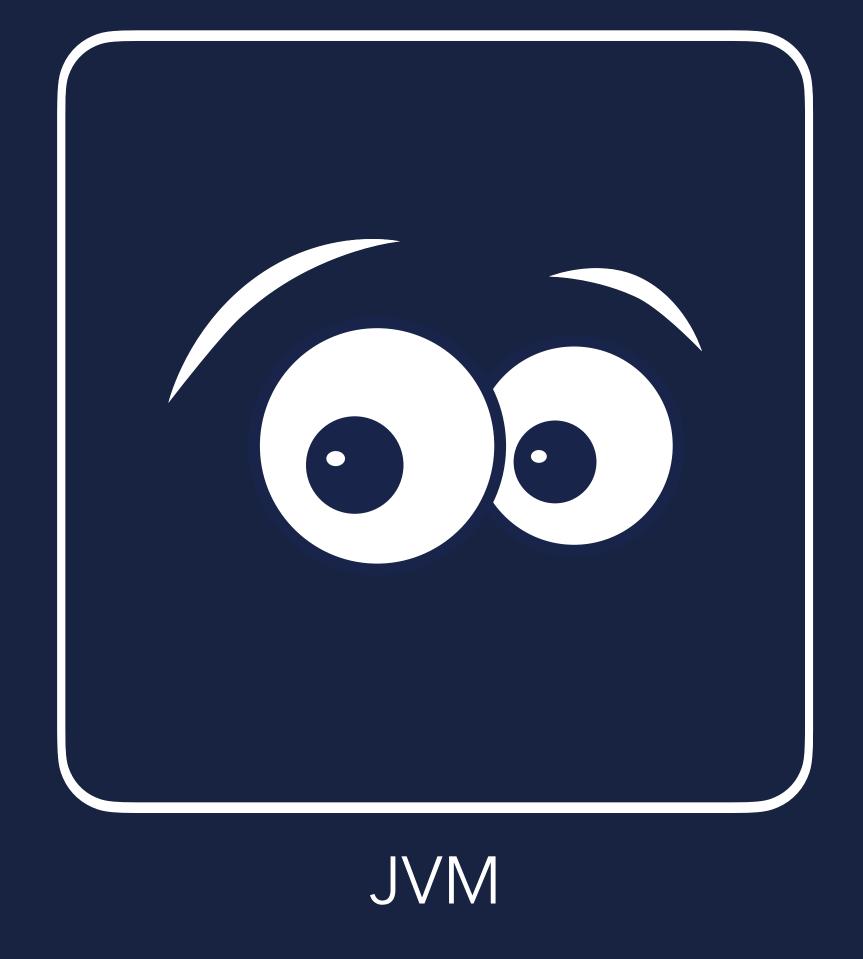
C1 JIT COMPILER

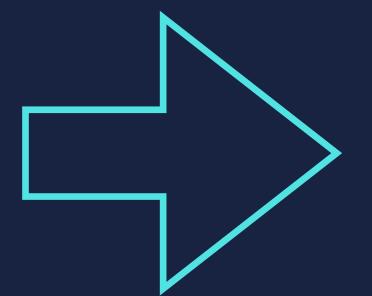
THRESHOLD
REACHED
(5000 in JDK 17)

Profiles the running code (detecting hot code)



Pass the "hot" code to C2 JIT Compiler

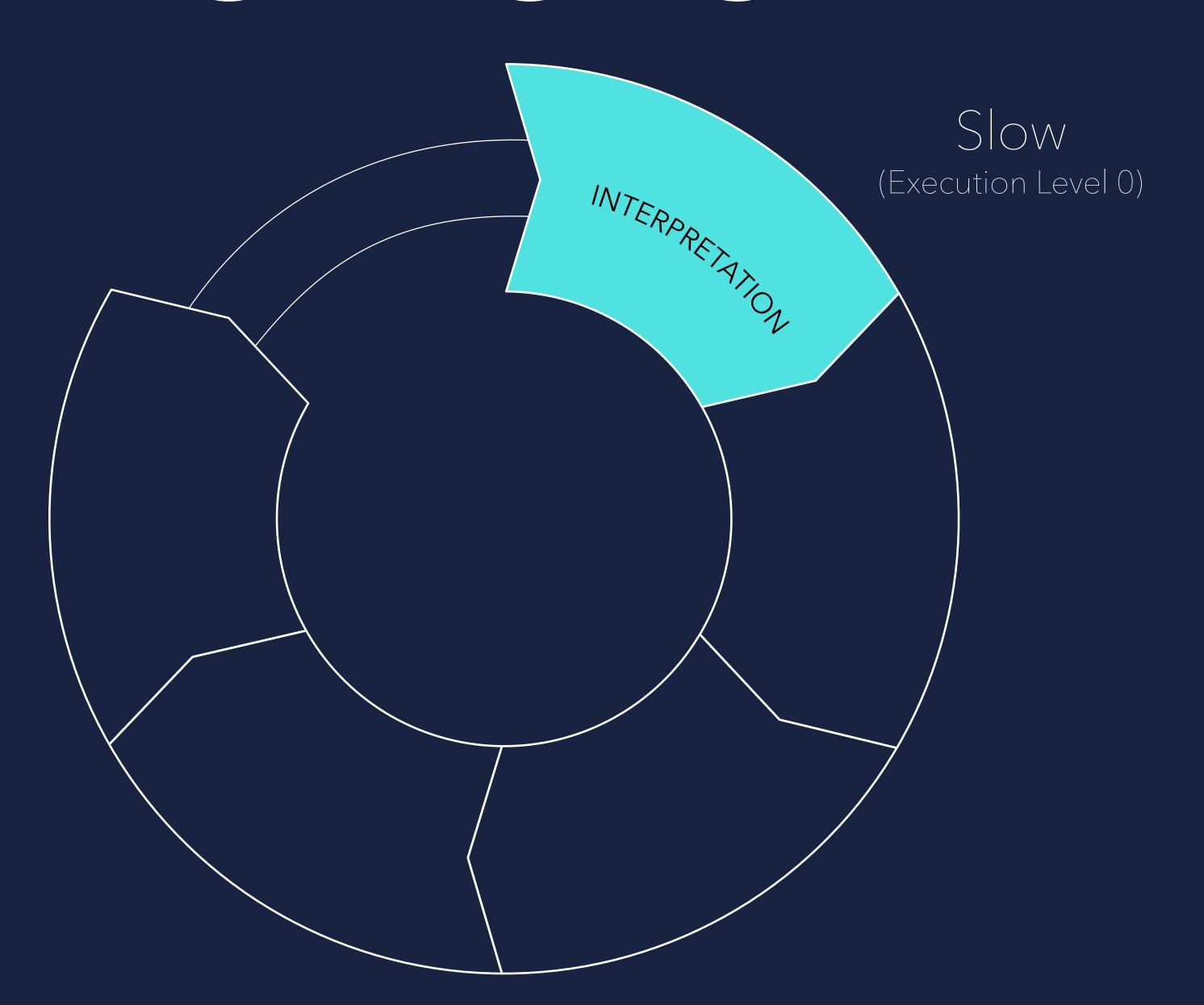




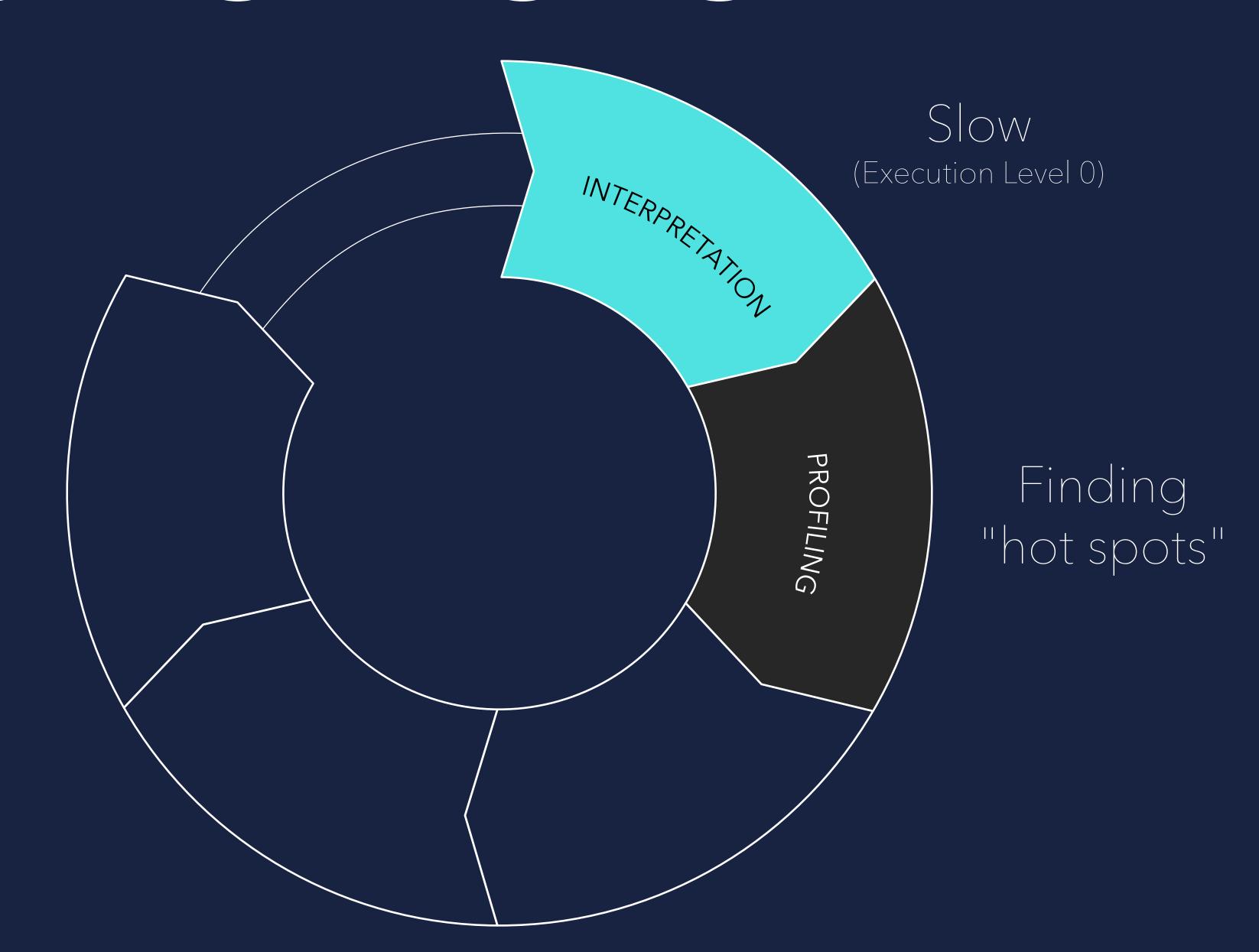
Compiles code with best optimisation possible (slower)



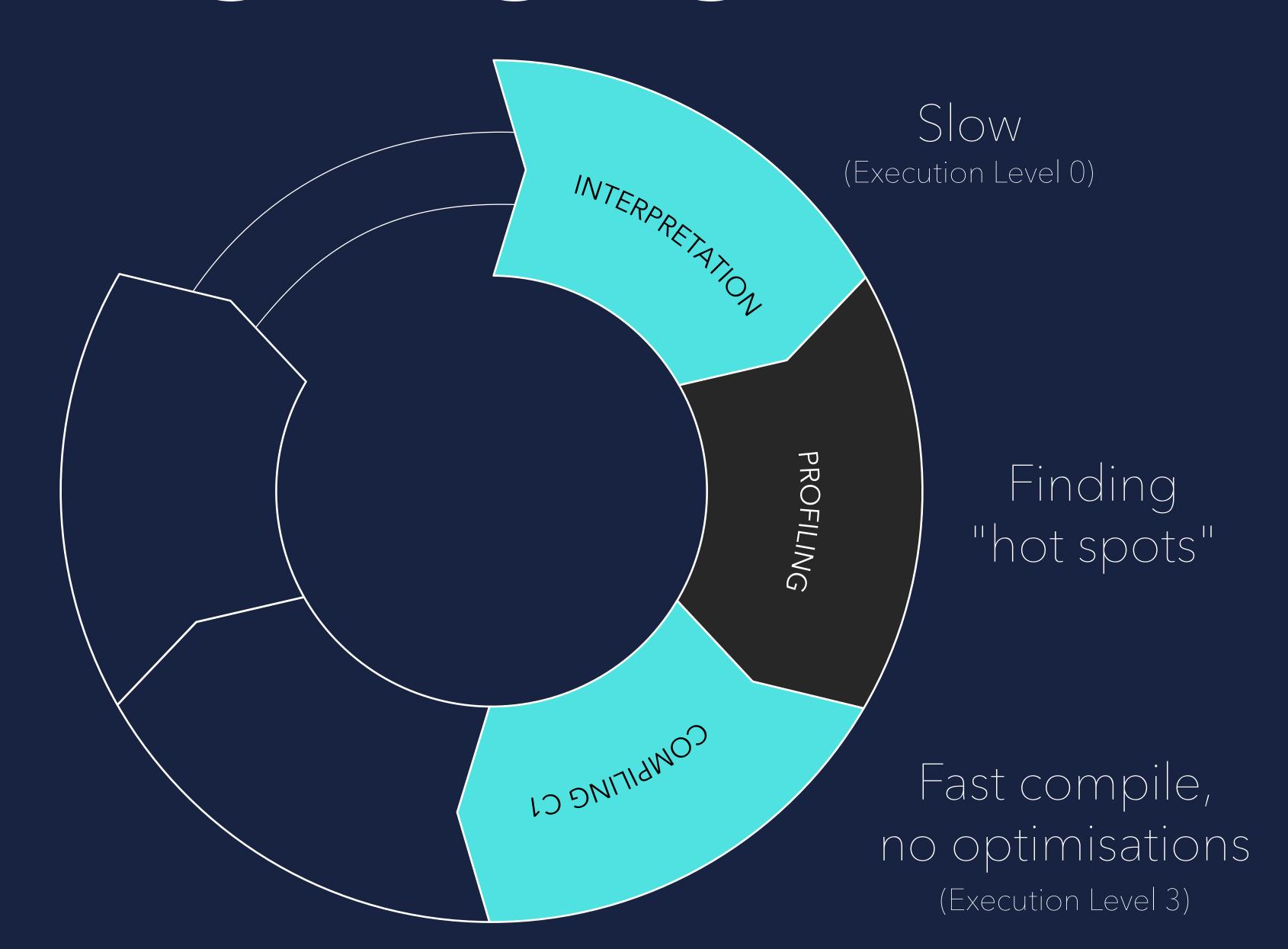
C2 JIT COMPILER

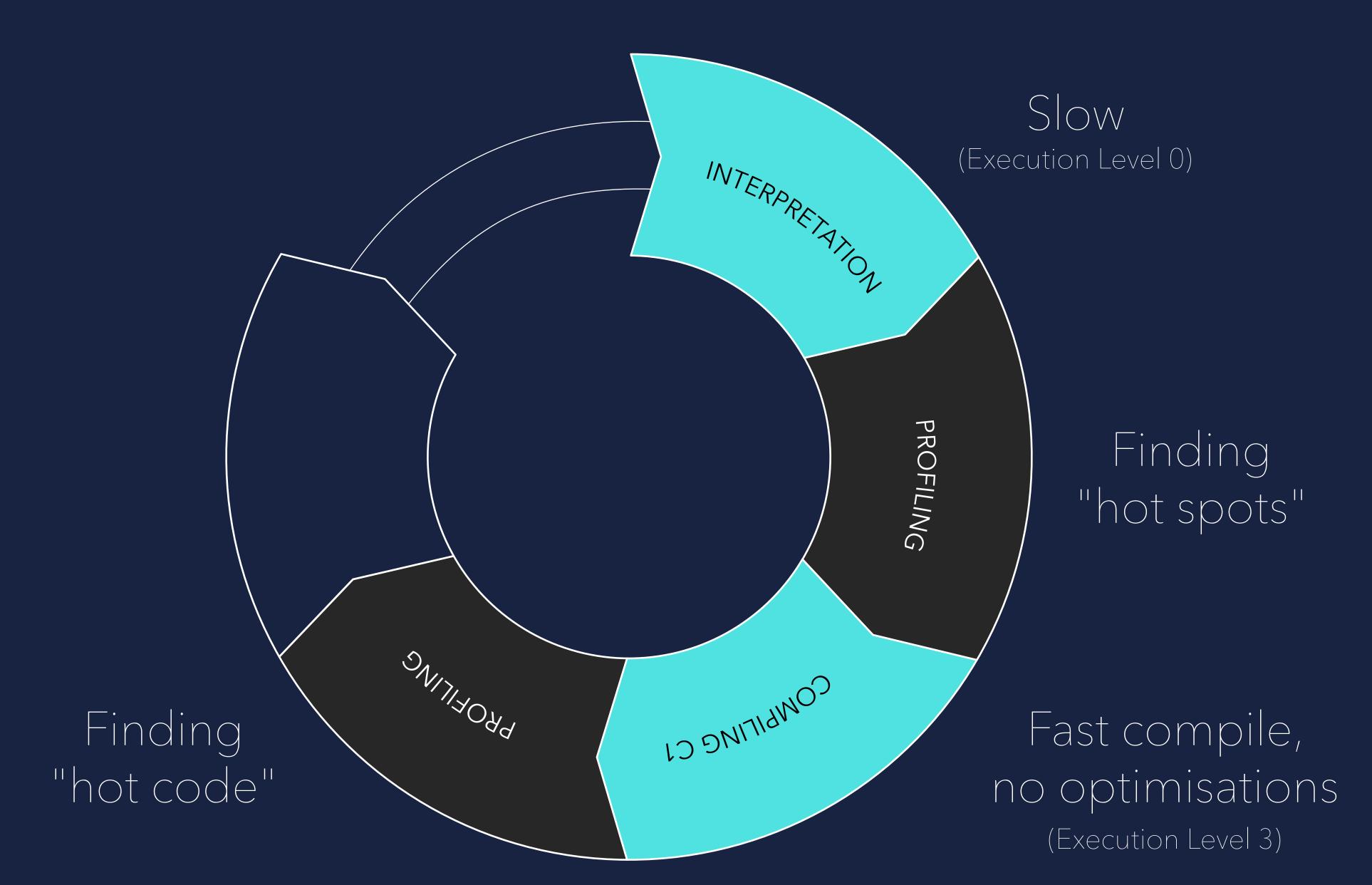


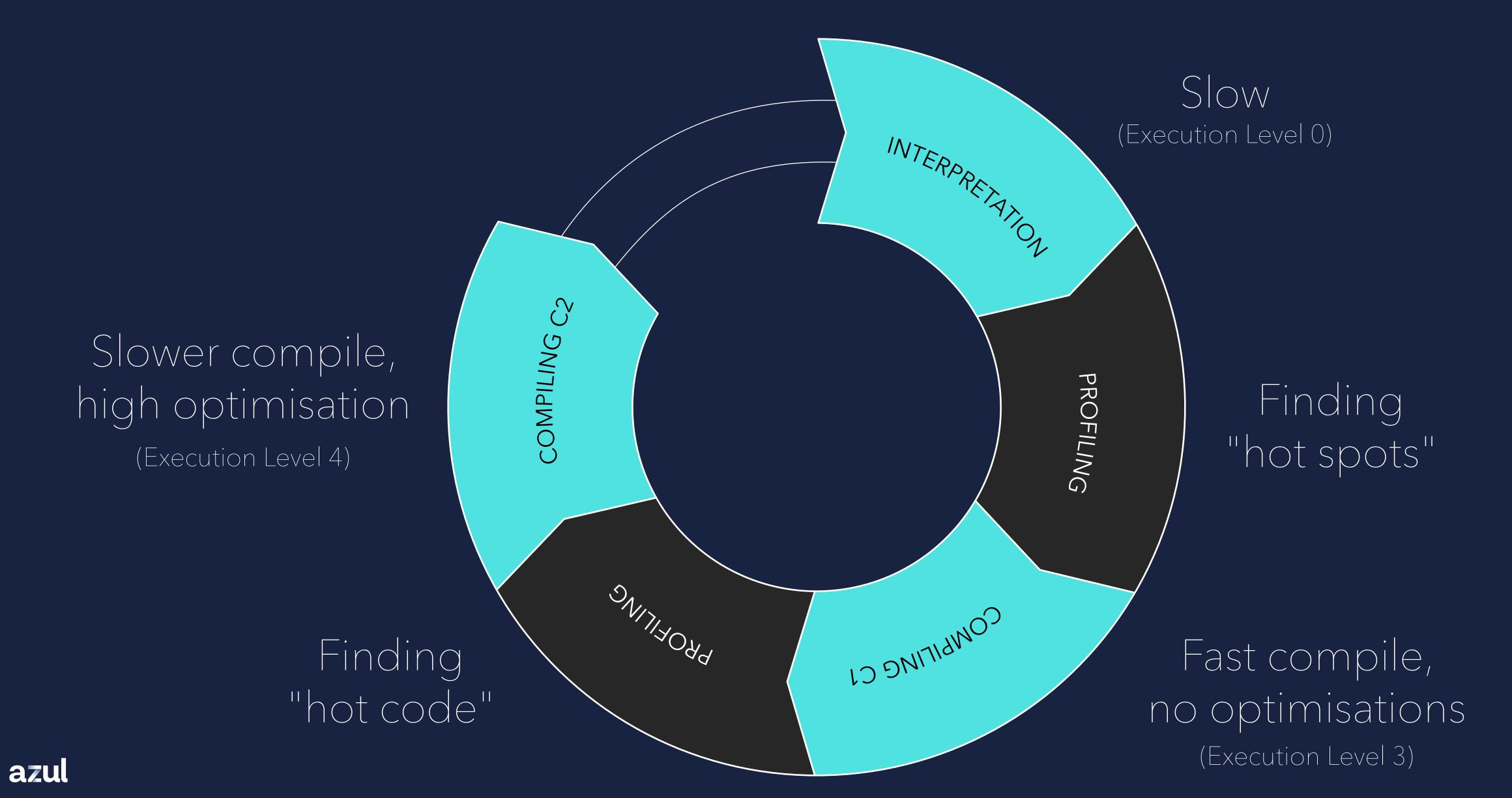
EXECUTION CYCLE

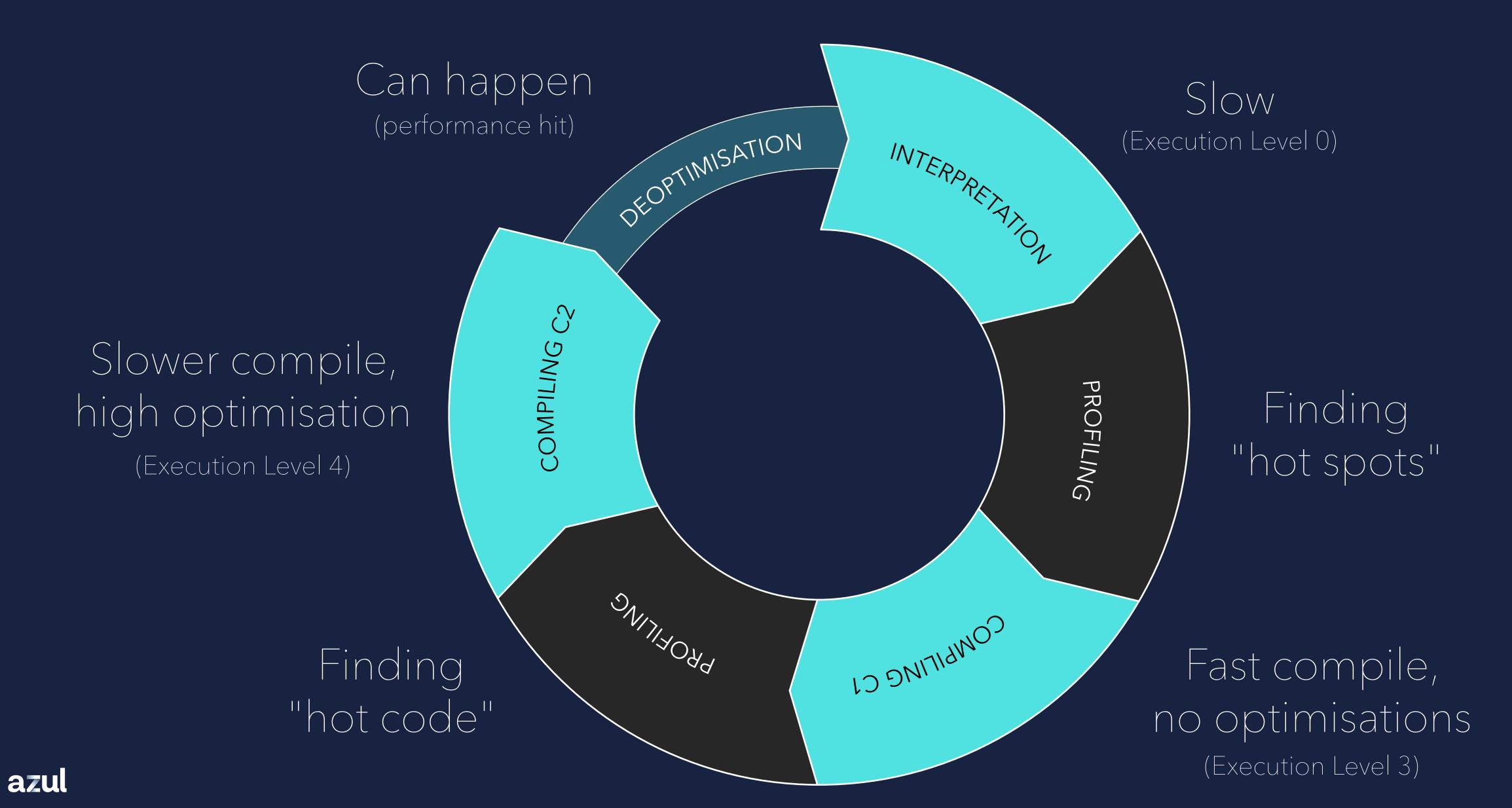


EXECUTION CYCLE

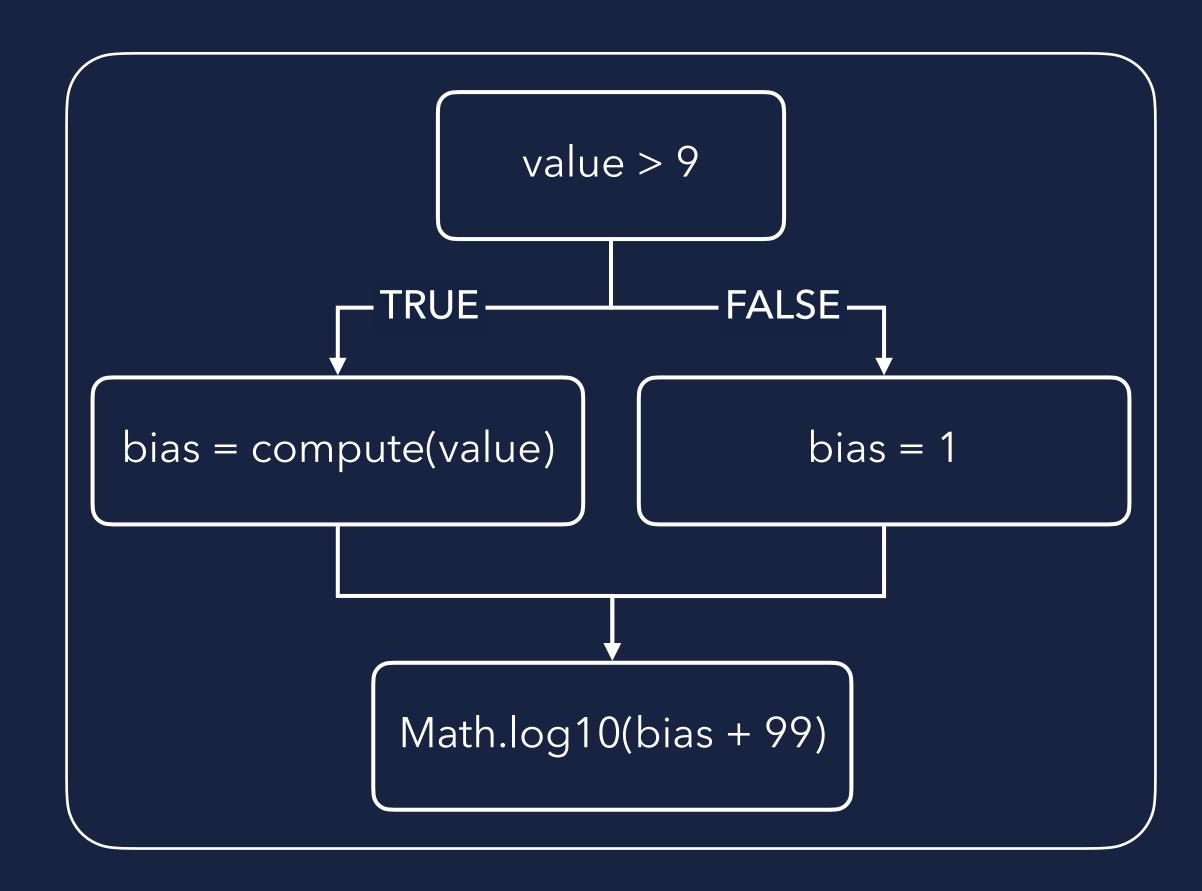




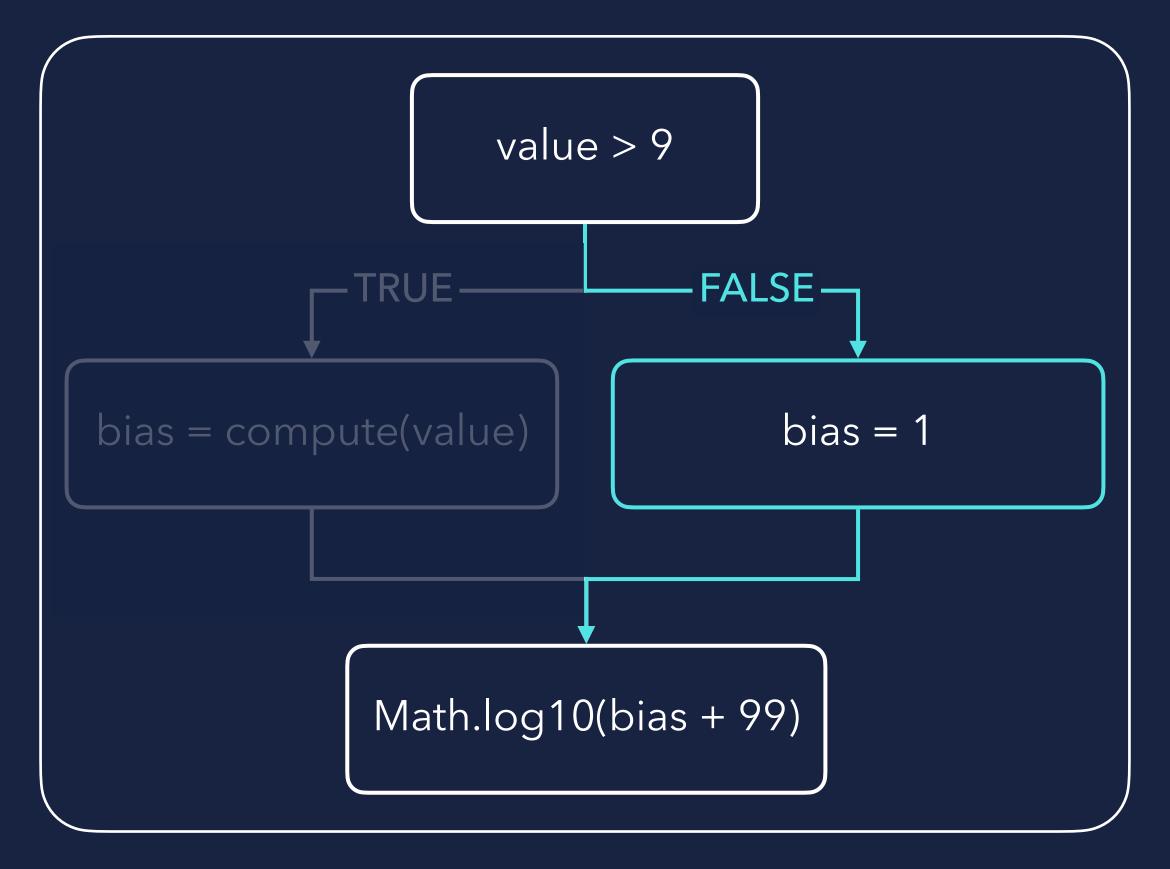




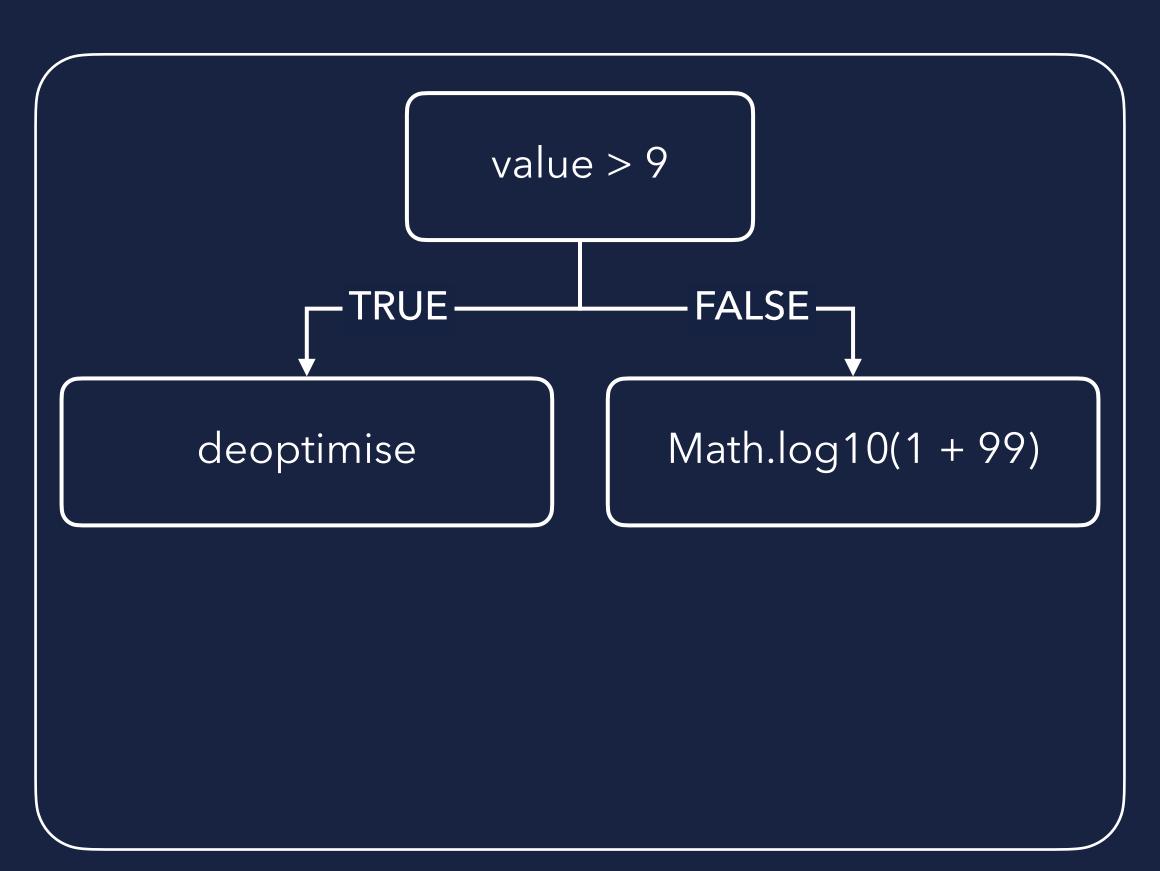
```
int computeMagnitude(int value) {
    int bias;
    if (value > 9) {
        bias = compute(value);
    } else {
        bias = 1:
    return Math.log10(bias + 99);
```



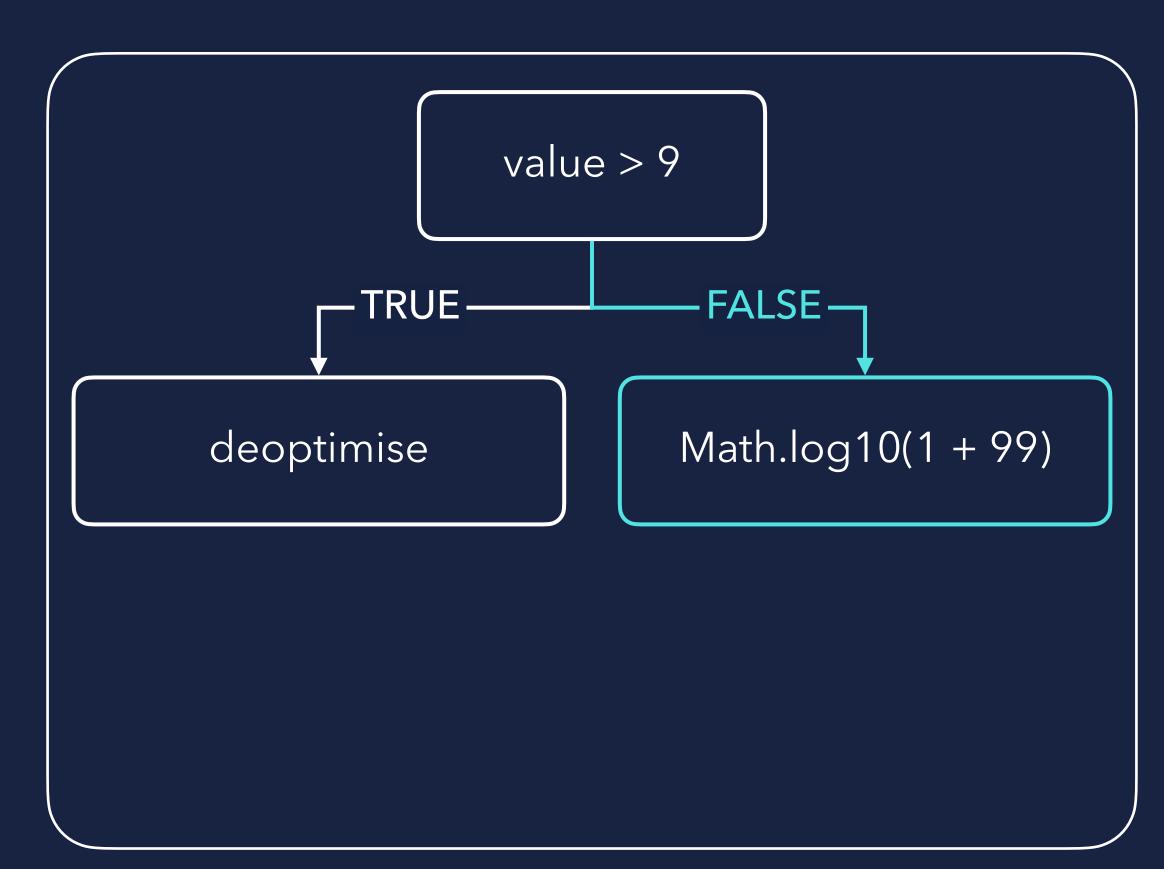
```
int computeMagnitude(int value) {
    int bias;
    if (value > 9) {
        bias = compute(value);
    } else {
        bias = 1:
    return Math.log10(bias + 99);
```



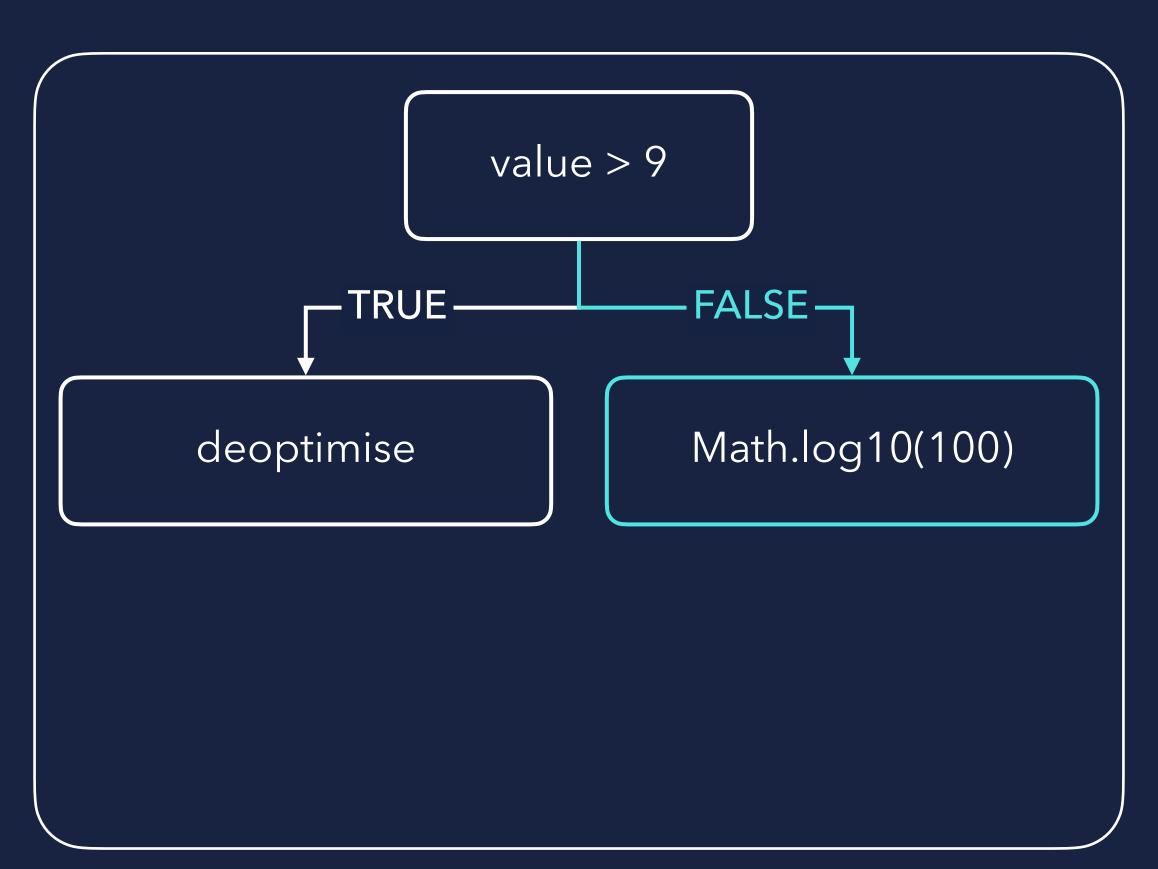
```
int computeMagnitude(int value) {
    if (value > 9) {
        uncommonTrap();
    }
    int bias = 1;
    return Math.log10(bias + 99);
}
```



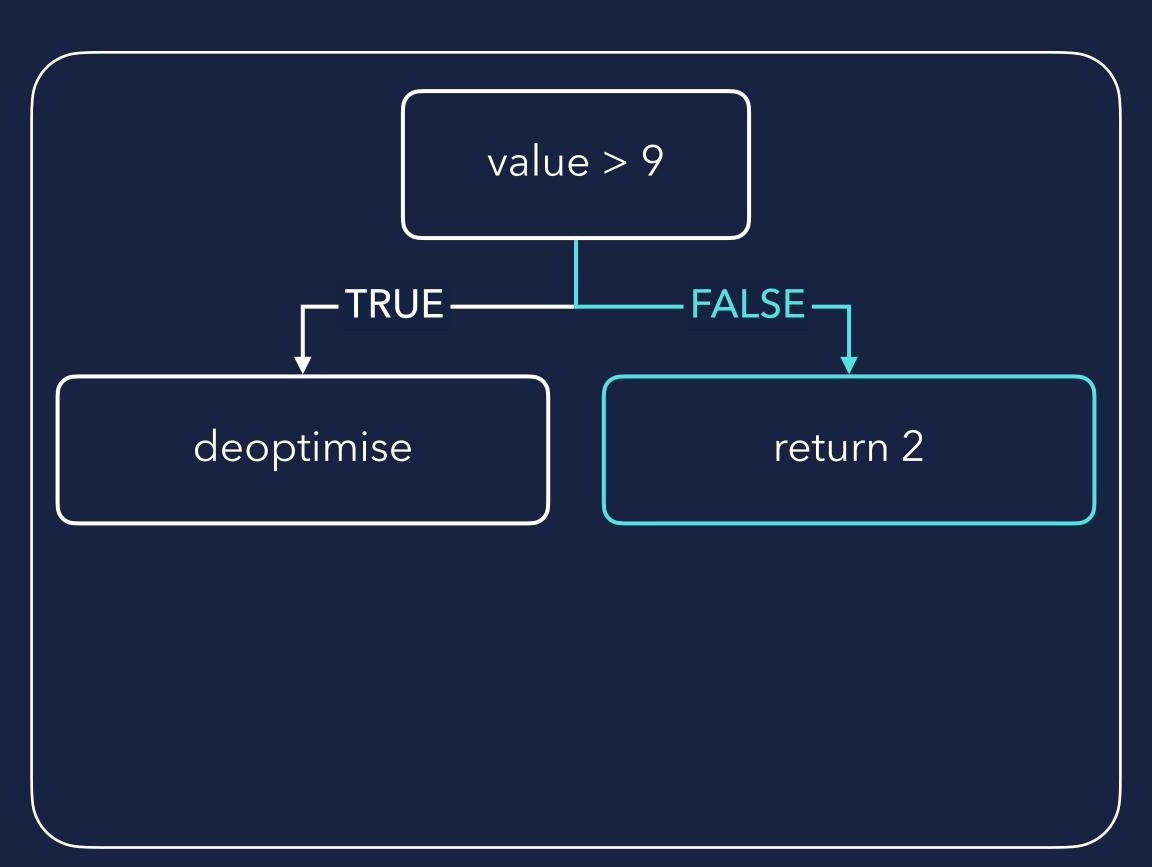
```
int computeMagnitude(int value) {
   if (value > 9) {
      uncommonTrap();
   }
   int bias = 1;
   return Math.log10(bias + 99);
}
```



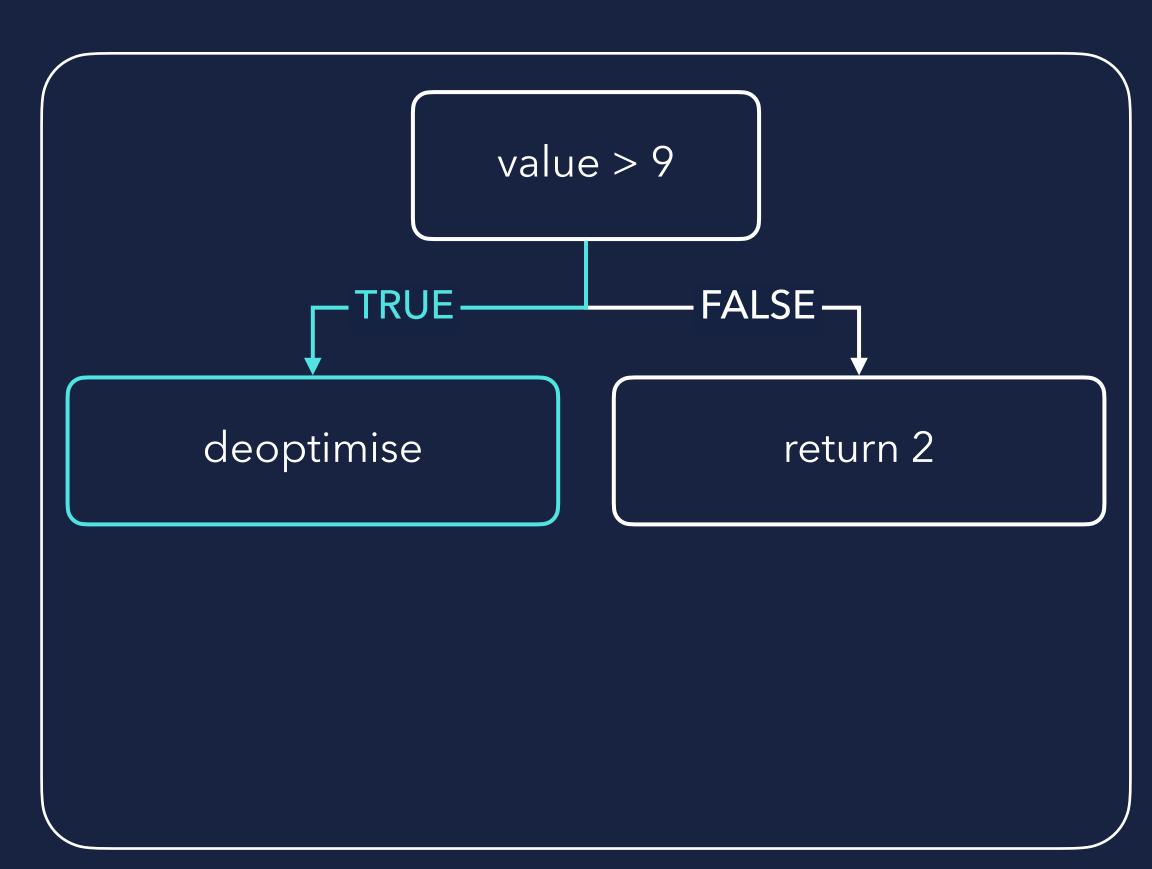
```
int computeMagnitude(int value) {
    if (value > 9) {
        uncommonTrap();
    }
    return Math.log10(100);
}
```



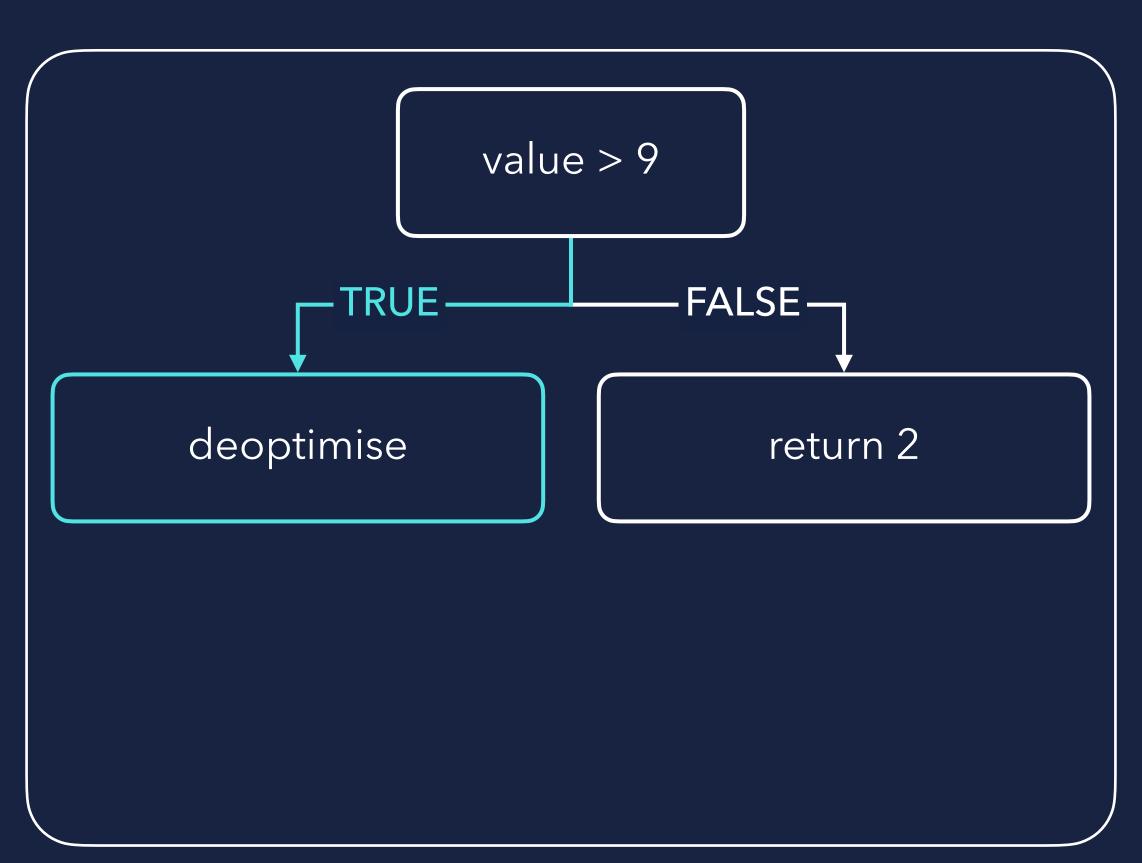
```
int computeMagnitude(int value) {
    if (value > 9) {
        uncommonTrap();
    }
    return 2;
}
```



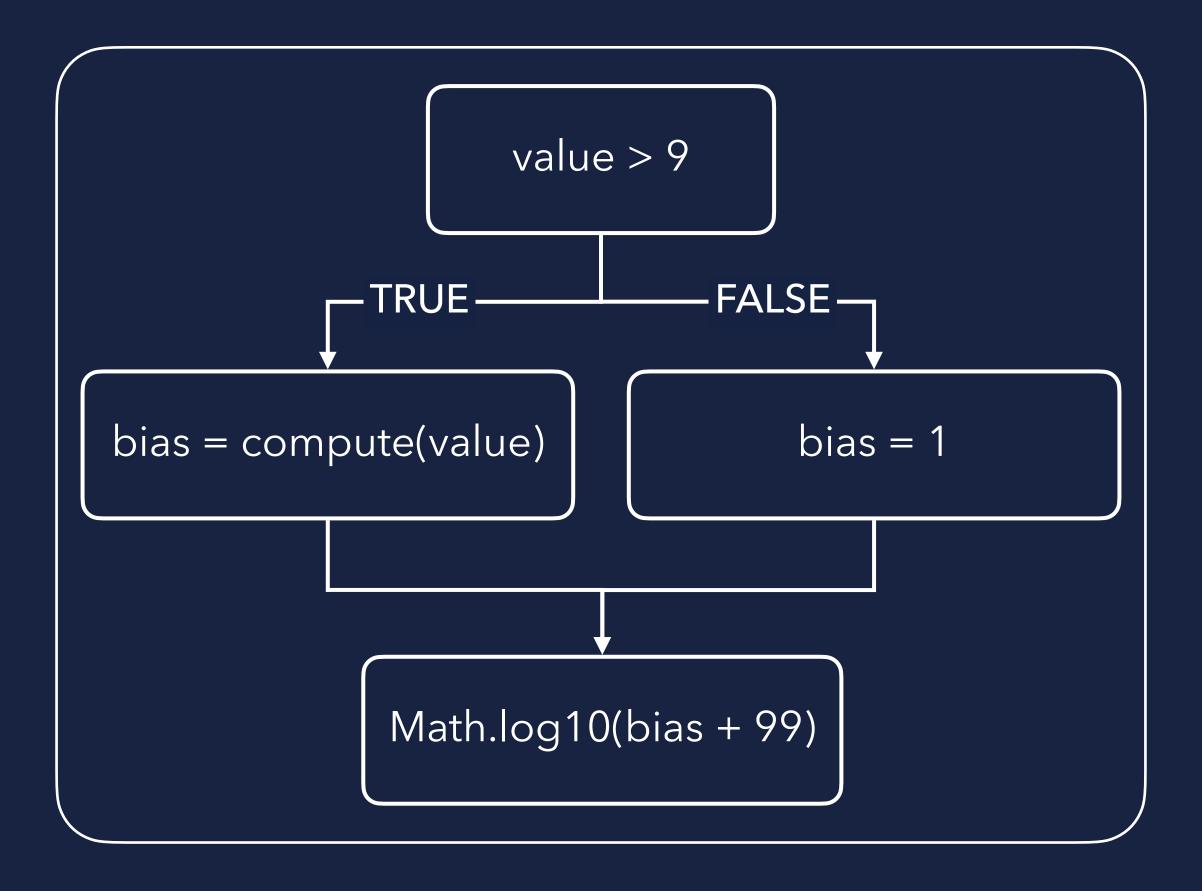
```
int computeMagnitude(int value) {
   if (value > 9) {
      uncommonTrap();
   }
   return 2;
}
```



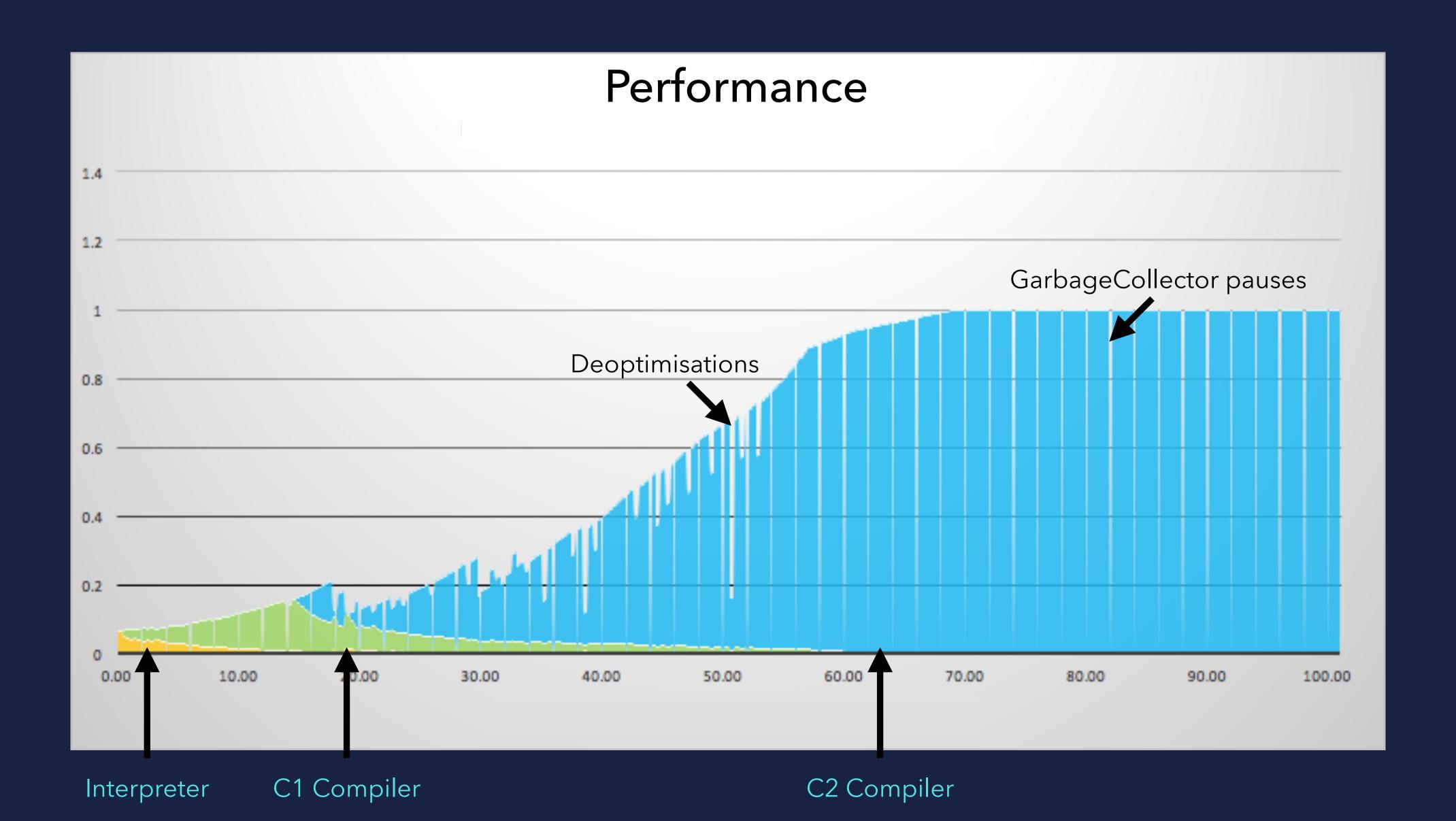
```
int computeMagnitude(int value) {
    if (value > 9) {
        uncommonTrap();
    return 2;
```



```
int computeMagnitude(int value) {
    int bias;
    if (value > 9) {
        bias = compute(value);
    } else {
        bias = 1:
    return Math.log10(bias + 99);
```



JVM PERFORMANCE GRAPH



azul

FAST

JVM START

JVM

- Load & Initialise
- Optimisation

FAST TAKES A BIT JVM START APPLICATION START JVM JVM Load application classes Load & Initialise Initialise all resources Optimisation Kick off application specific logic Optimisation

FAST TAKES A BIT

JVM START

APPLICATION START

JVM

- Load & Initialise
- Optimisation

JVM

- Load application classes
- Initialise all resources
- Kick off application specific logic
- Optimisation

Generally referred to as JVM Startup (Time to first response)

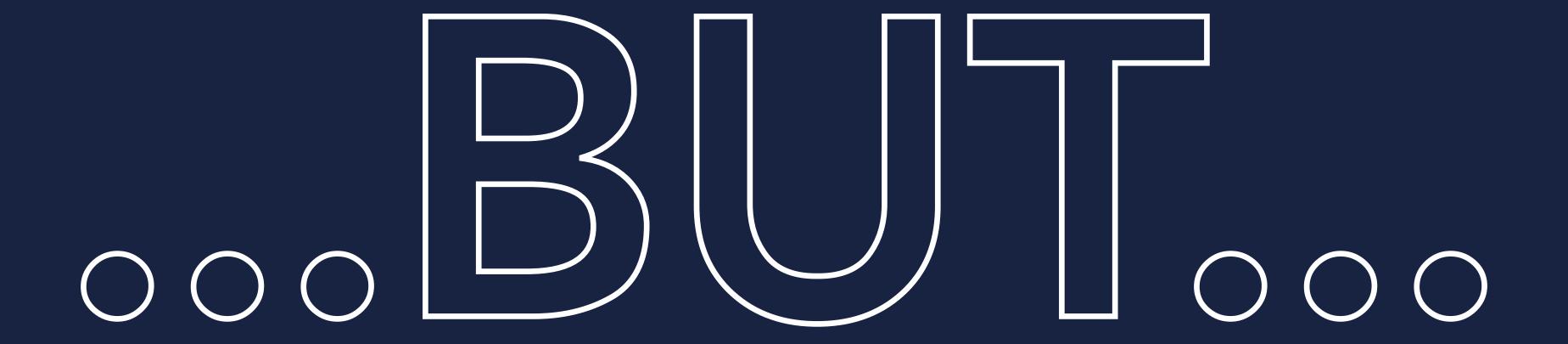
FAST TAKES A BIT TAKES SOME TIME JVM START APPLICATION START APPLICATION WARMUP JVM JVM JVM Load & Initialise Load application classes Optimising (Compile/Decompile) Initialise all resources Optimisation Kick off application specific logic App Optimisation Apply application specific workloads

Generally referred to as JVM Startup (Time to first response)

FAST TAKES A BIT TAKES SOME TIME JVM START APPLICATION START APPLICATION WARMUP JVM JVM JVM Load & Initialise Optimising (Compile/Decompile) Load application classes Optimisation Initialise all resources Kick off application specific logic App Optimisation Apply application specific workloads

Generally referred to as JVM Startup (Time to first response)

Generally referred to as JVM Warmup (Time to n operations)



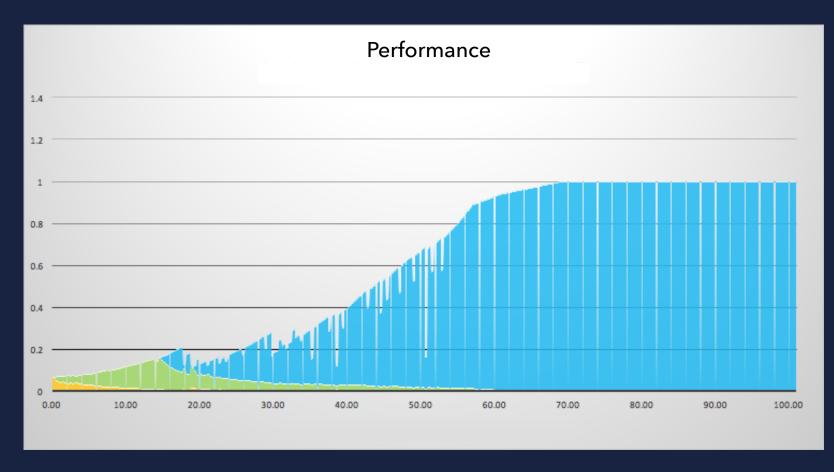
MICROSERVICE ENVIRONMENT

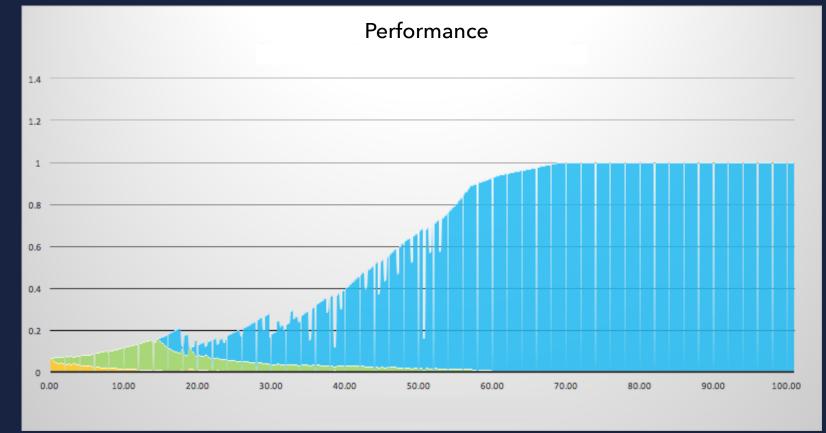
MICROSERVICE ENVIRONMENT

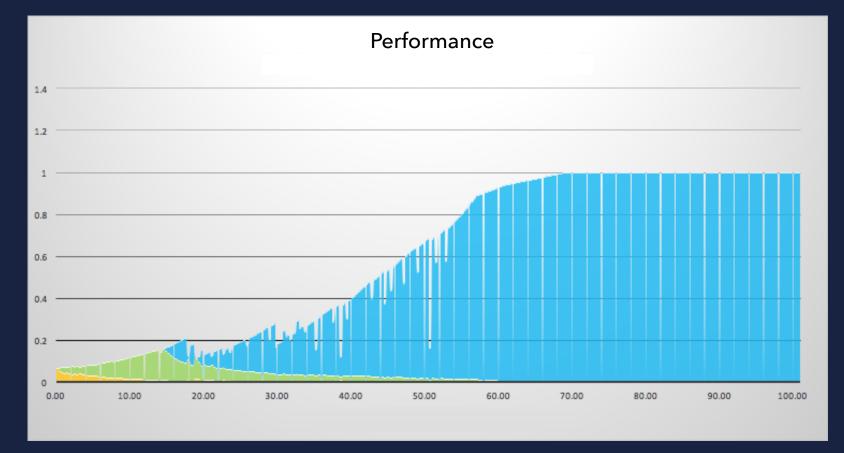
FIRST RUN

SECOND RUN

THIRD RUN







JVM STARTUP

JVM STARTUP

JVM STARTUP

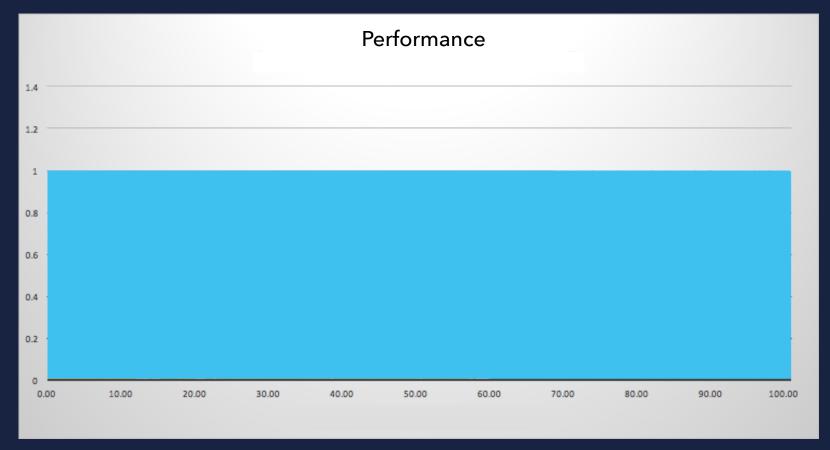


WOULDNIT BE GREAT..?

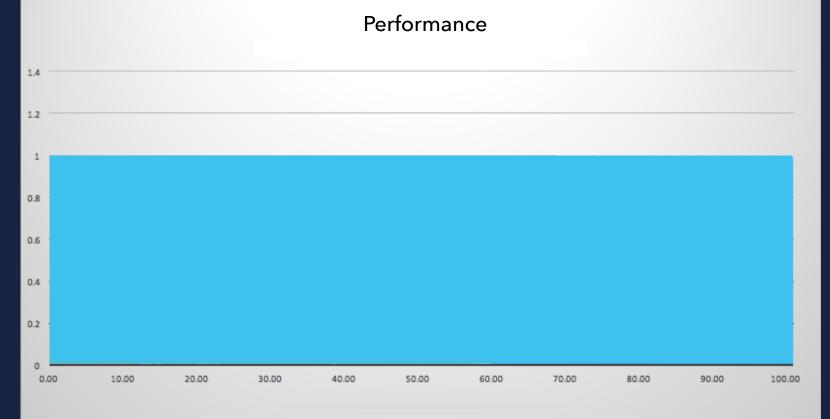
FIRST RUN

Performance 1.4 1.2 1.0 0.8 0.6 0.4 0.2 0.0 0.0 1.0.00 1.0.

SECOND RUN



THIRD RUN



JVM STARTUP

NO STARTUP OVERHEAD NO STARTUP OVERHEAD

Dump internal class representations into file

WHAT ABOUT CDS?

- Dump internal class representations into file
- Shared on each JVM start (CDS)

MHAT ABOUT CDS?

- Dump internal class representations into file
- Shared on each JVM start (CDS)
- No optimization or hotspot detection

- Dump internal class representations into file
- Shared on each JVM start (CDS)
- No optimization or hotspot detection
- Only reduces class loading time

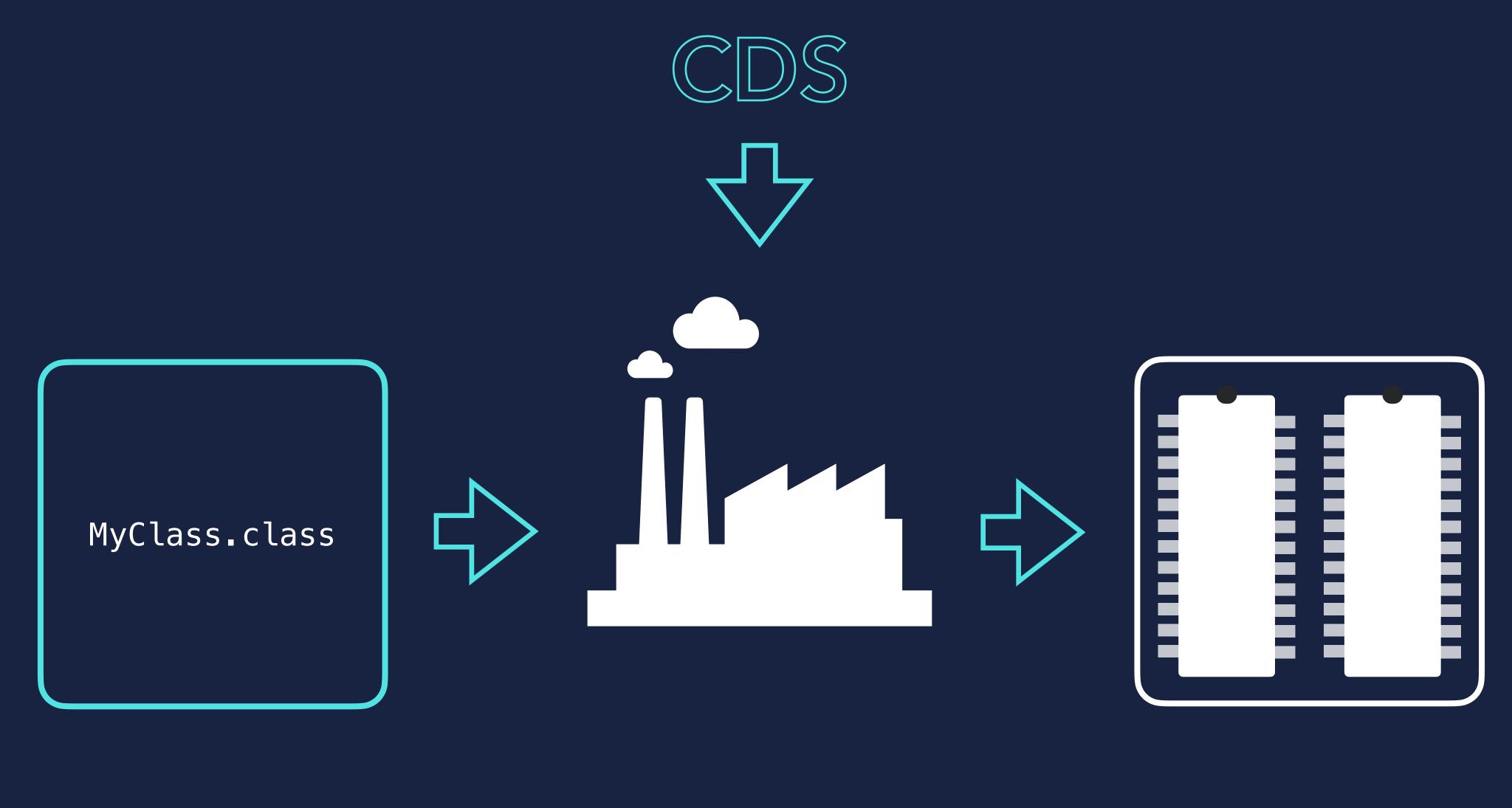
MHAT ABOUT CDS?

- Dump internal class representations into file
- Shared on each JVM start (CDS)
- No optimization or hotspot detection
- Only reduces class loading time
- Startup up to 2 seconds faster

WHAT ABOUT CDS?

- Dump internal class representations into file
- Shared on each JVM start (CDS)
- No optimization or hotspot detection
- Only reduces class loading time
- Startup up to 2 seconds faster
- Good info from lonut Balosin





BYTE CODE

CLASS LOADERS

JVM MEMORY

AHEAD OF TIME COMPILATION

No interpreting bytecodes

- No interpreting bytecodes
- No analysis of hotspots

- No interpreting bytecodes
- No analysis of hotspots
- No runtime compilation of code

- No interpreting bytecodes
- No analysis of hotspots
- No runtime compilation of code
- Start at 'full speed', straight away

- No interpreting bytecodes
- No analysis of hotspots
- No runtime compilation of code
- Start at 'full speed', straight away
- GraalVM native image does that

PROBLEM SOLVED...?

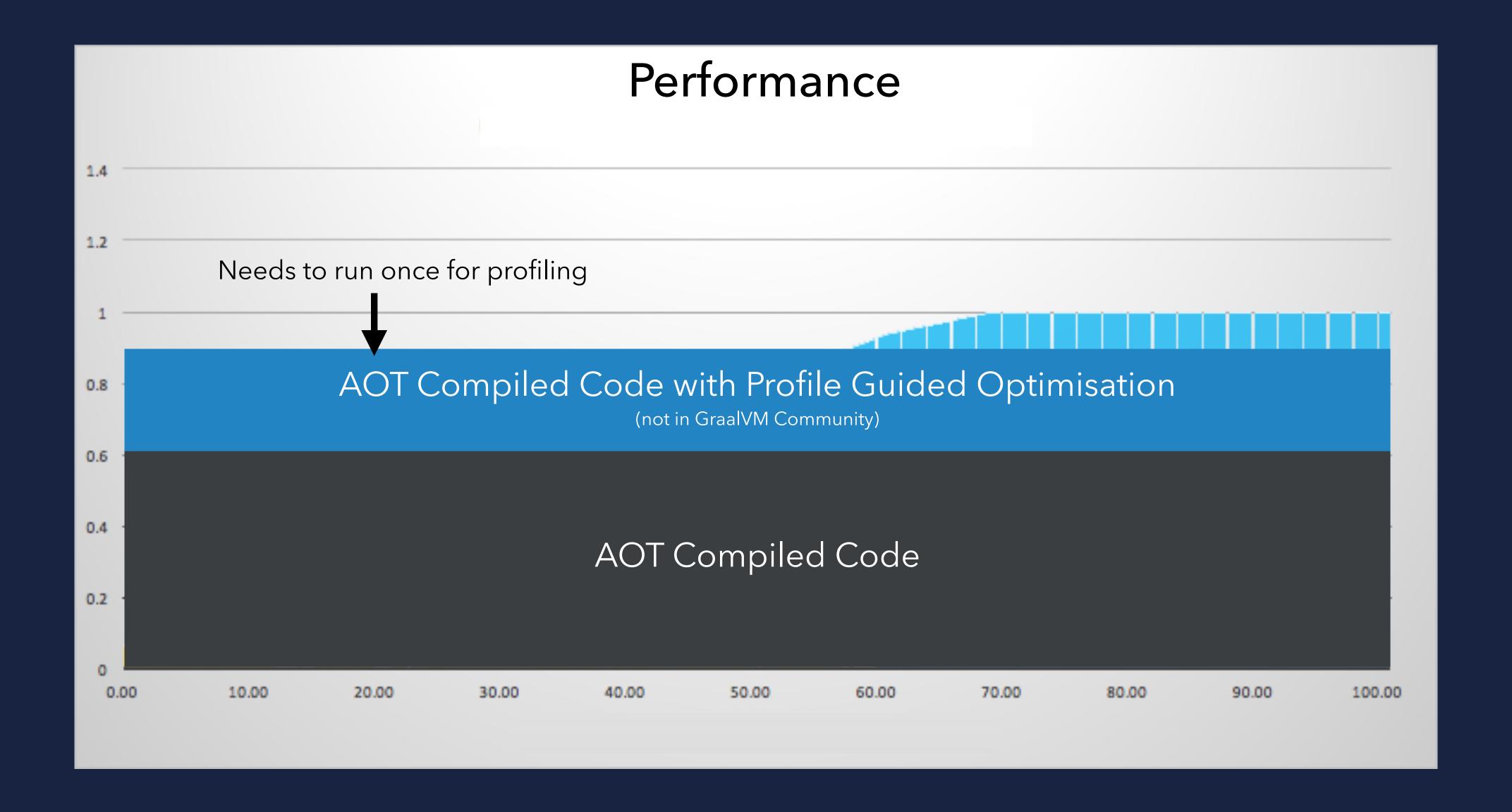
AOT is, by definition, static

- AOT is, by definition, static
- Code is compiled before it is run

- AOT is, by definition, static
- Code is compiled before it is run
- Compiler has no knowledge of how the code will actually run

- AOT is, by definition, static
- Code is compiled before it is run
- Compiler has no knowledge of how the code will actually run
- Profile Guided Optimisation (PGO) can partially help

JVM PERFORMANCE GRAPH







CHECKPOINT RESTORE IN USERSPACE



Linux project





- Linux project
- Part of kernel >= 3.11 (2013)





- Linux project
- Part of kernel >= 3.11 (2013)
- Freeze a running container/application





- Linux project
- Part of kernel >= 3.11 (2013)
- Freeze a running container/application
- Checkpoint its state to disk





- Linux project
- Part of kernel >= 3.11 (2013)
- Freeze a running container/application
- Checkpoint its state to disk
- Restore the container/application from the saved data.





- Linux project
- Part of kernel >= 3.11 (2013)
- Freeze a running container/application
- Checkpoint its state to disk
- Restore the container/application from the saved data.
- Used by/integrated in OpenVZ, LXC/LXD, Docker, Podman and others





Heavily relies on /proc file system





- Heavily relies on /proc file system
- It can checkpoint:
 - Processes and threads
 - Application memory, memory mapped files and shared memory
 - Open files, pipes and FIFOs
 - Sockets
 - Interprocess communication channels
 - Timers and signals
- Can rebuild TCP connection from one side only



CRIU CHALLENGES

Restore from saved state on another machine (open files, shared memory etc.)



CRIU CHALLENGES





Restore multiple instances of same state on same machine (PID will be restored which will lead to problems)

CRIU CHALLENGES

- Restore from saved state on another machine (open files, shared memory etc.)
 - Restore multiple instances of same state on same machine (PID will be restored which will lead to problems)
 - A Java Virtual Machine would assume it was continuing its tasks (very difficult to use effectively, e.g. running applications might have open files etc.)

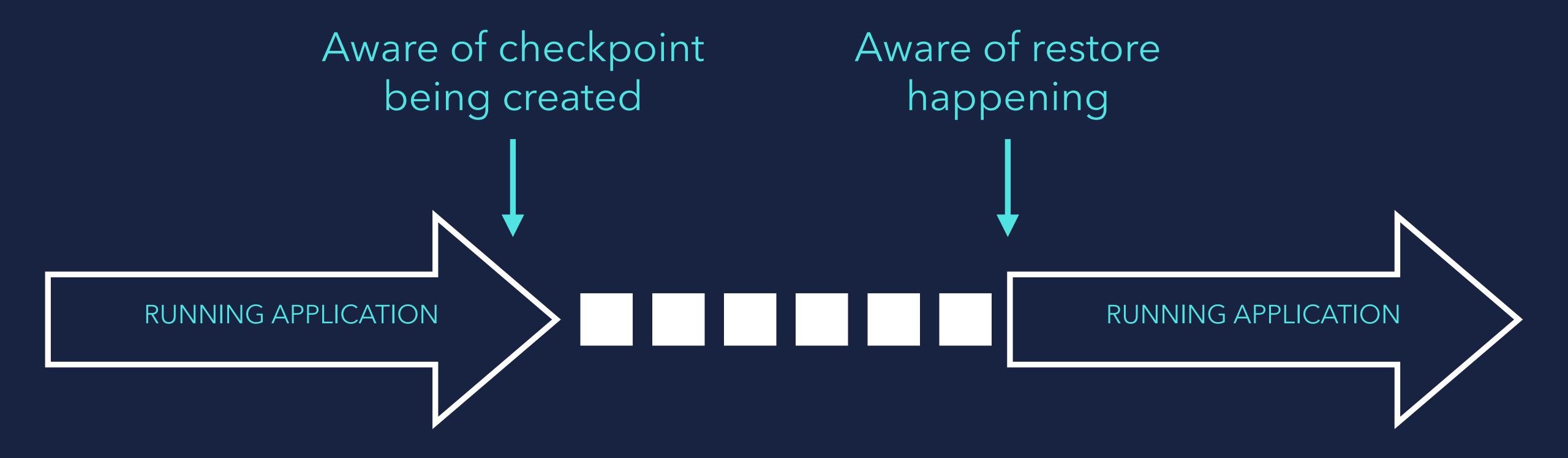




Coordinated Restore at Checkpoint



A way to solve the problems when checkpointing a JVM (e.g. no open files, sockets etc.)





Comes with a simple API



- Comes with a simple API
- Creates checkpoints using code or jcmd



- Comes with a simple API
- Creates checkpoints using code or jcmd
- Throws CheckpointException

(in case of open files/sockets)



- Comes with a simple API
- Creates checkpoints using code or jcmd
- Throws CheckpointException

(in case of open files/sockets)

Heap is cleaned, compacted

(using JVM safepoint mechanism -> JVM is in a safe state)



Additional command line parameters

START

>java -XX:CRaCCheckpointTo=PATH -jar app.jar

RESTORE

>java -XX:CRaCRestoreFrom=PATH



Resource interface (can be notified about a Checkpoint and Restore)

<<interface>>

Resource

- Resource interface (can be notified about a Checkpoint and Restore)
- Classes in application code implement the Resource interface

<<interface>>

Resource

- Resource interface (can be notified about a Checkpoint and Restore)
- Classes in application code implement the Resource interface
- Application receives callbacks during checkpointing and restoring

<<interface>>

Resource

- Resource interface (can be notified about a Checkpoint and Restore)
- Classes in application code implement the Resource interface
- Application receives callbacks during checkpointing and restoring
- Makes it possible to close/restore resources (e.g. open files, sockets)

<<interface>>

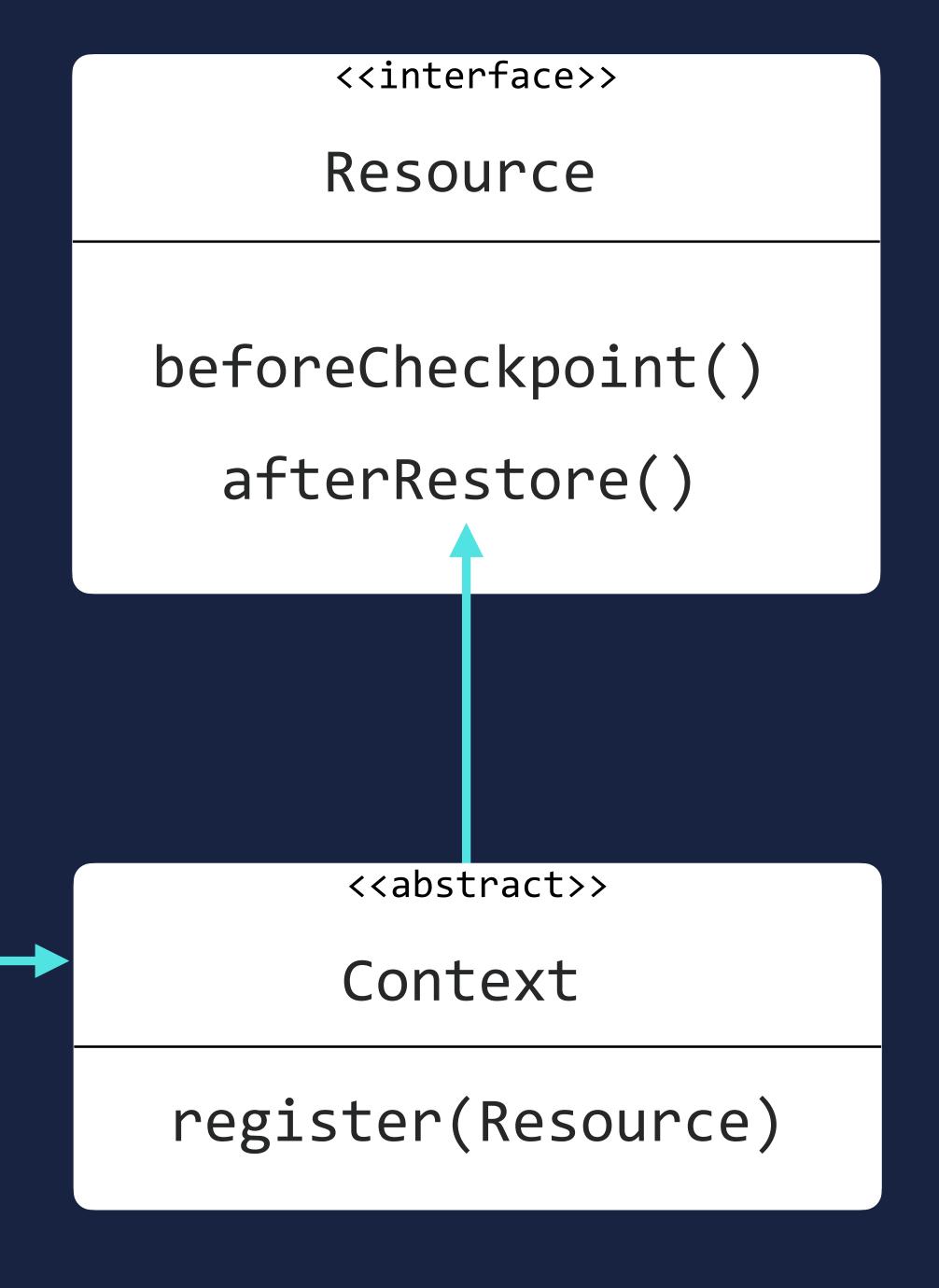
Resource

Resource objects need to be registered with a Context so that they can receive notifications

- Resource objects need to be registered with a Context so that they can receive notifications
- There is a global Context accessible via the static method Core.getGlobalContext()

Core

getGlobalContext()



CREATINGACHECKPOINT

FROMTHE COMMAND LINE:

>jcmd YOUR_AWESOME.jar JDK.checkpoint

>jcmd PID JDK.checkpoint

CREATING A CHECKPOINT

FROM CODE:

Core.checkpointRestore();



Start your app with -XX:+PrintCompilation

- Start your app with -XX:+PrintCompilation
- Apply typical workload to your app

- Start your app with -XX:+PrintCompilation
- Apply typical workload to your app
- Observe the moment the compilations are ramped down

- Start your app with -XX:+PrintCompilation
- Apply typical workload to your app
- Observe the moment the compilations are ramped down
- Create the checkpoint







APPLICATION

JVM startup...

JVM

APPLICATION

RESOURCE 1

beforeCheckpoint()

afterRestore()

RESOURCE 2

beforeCheckpoint()

afterRestore()



APPLICATION

RESOURCE 1

beforeCheckpoint()

afterRestore()

RESOURCE 2

beforeCheckpoint()

afterRestore()

Register resources in global context

JVM

APPLICATION

RESOURCE 1

beforeCheckpoint()

afterRestore()

RESOURCE 2

beforeCheckpoint()

afterRestore()



Warmup the application

JVM

jcmd JDK.checkpoint

APPLICATION

RESOURCE 1

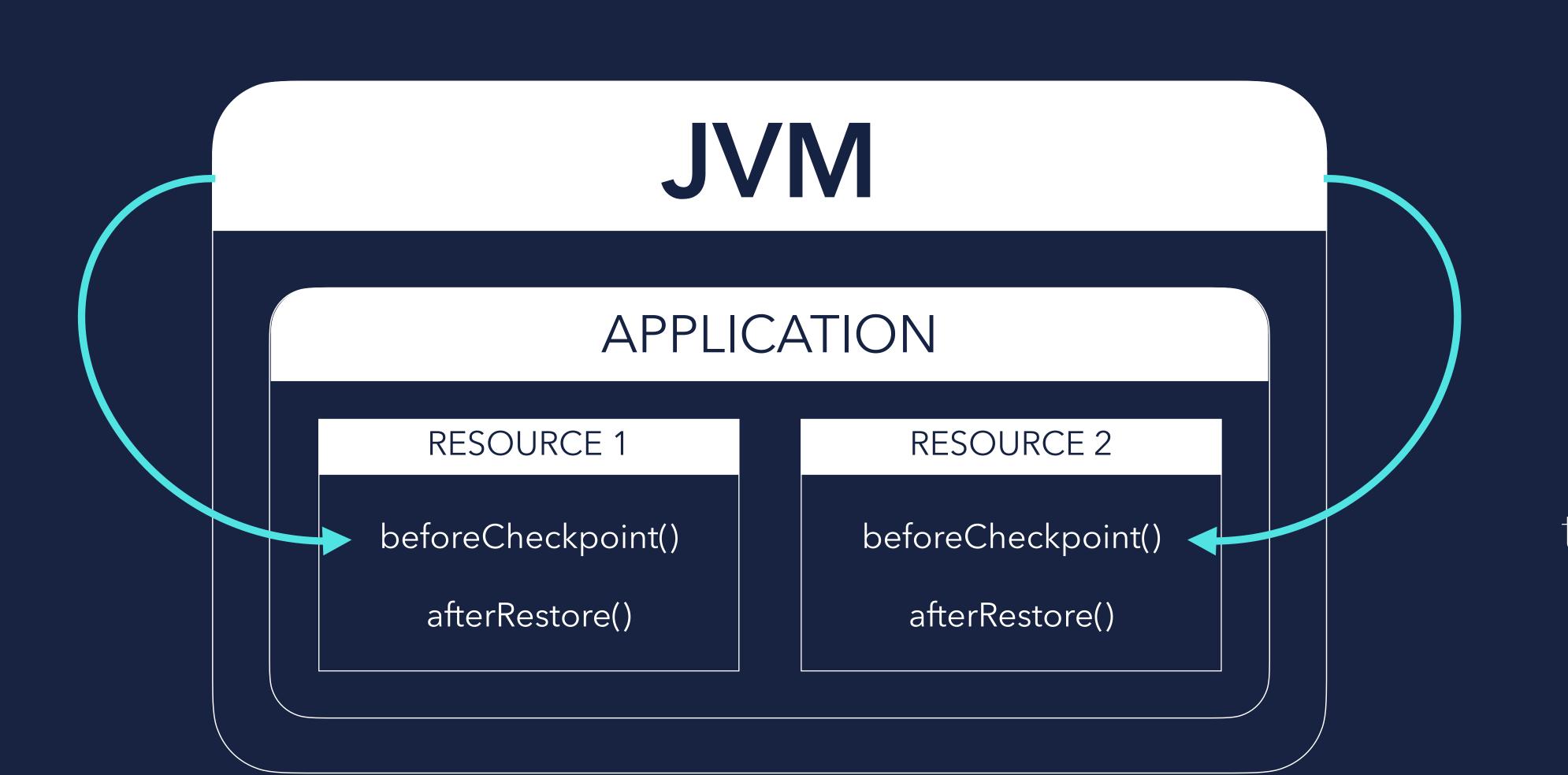
beforeCheckpoint()

afterRestore()

RESOURCE 2

beforeCheckpoint()

afterRestore()



JVIVI notifies the resources

JVM

APPLICATION

RESOURCE 1

beforeCheckpoint()

afterRestore()

RESOURCE 2

beforeCheckpoint()

afterRestore()

Application closes open resources



APPLICATION

RESOURCE 1

beforeCheckpoint()

afterRestore()

RESOURCE 2

beforeCheckpoint()

afterRestore()



JVM stores checkpoint to disc

JVM

APPLICATION

RESOURCE 1

beforeCheckpoint()

afterRestore()

RESOURCE 2

beforeCheckpoint()

afterRestore()



Time...

JVM

APPLICATION

RESOURCE 1

beforeCheckpoint()

afterRestore()

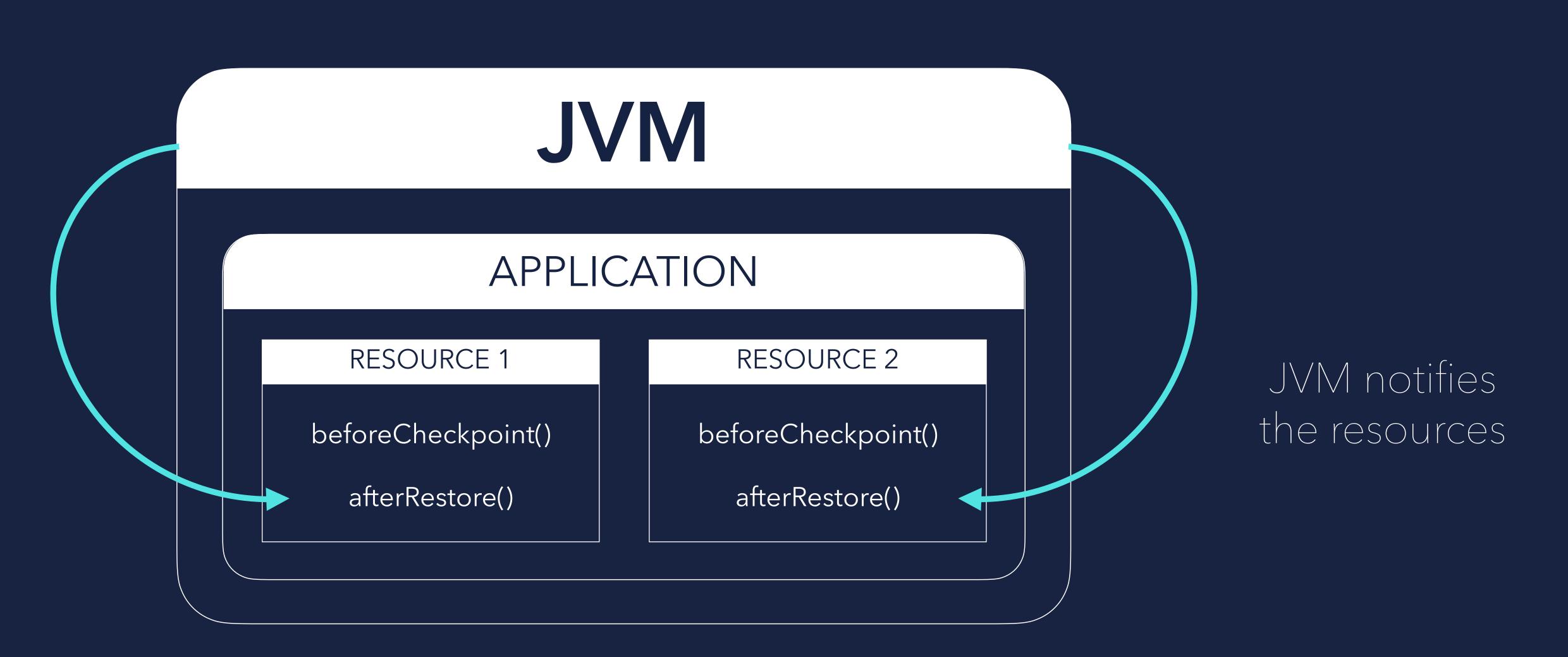
RESOURCE 2

beforeCheckpoint()

afterRestore()

java -XX:CRaCRestoreFrom

Restore from checkpoint



JVM

APPLICATION

RESOURCE 1

beforeCheckpoint()

afterRestore()

RESOURCE 2

beforeCheckpoint()

afterRestore()

Application re-open resources

JVM

APPLICATION

RESOURCE 1

beforeCheckpoint()

afterRestore()

RESOURCE 2

beforeCheckpoint()

afterRestore()

No JVM startup and no application warmup!!!

Run app in a docker container

- Run app in a docker container
- Create checkpoint (store in container or external volume)

- Run app in a docker container
- Create checkpoint (store in container or external volume)
- Commit the state of container (only if checkpoint in container)

- Run app in a docker container
- Create checkpoint (store in container or external volume)
- Commit the state of container (only if checkpoint in container)
- Start the container (point jvm to container or external volume)



Designed to provide smooth CRaC adoption

- Designed to provide smooth CRaC adoption
- Total mirror of jdk.crac api at compile-time

- Designed to provide smooth CRaC adoption
- Total mirror of jdk.crac api at compile-time
- Can be used with any OpenJDK implementation

- Designed to provide smooth CRaC adoption
- Total mirror of jdk.crac api at compile-time
- Can be used with any OpenJDK implementation
- Detects CRaC implementation at runtime

- Designed to provide smooth CRaC adoption
- Total mirror of jdk.crac api at compile-time
- Can be used with any OpenJDK implementation
- Detects CRaC implementation at runtime
- No CRaC support -> won't call CRaC specific code

- Designed to provide smooth CRaC adoption
- Total mirror of jdk.crac api at compile-time
- Can be used with any OpenJDK implementation
- Detects CRaC implementation at runtime
- No CRaC support -> won't call CRaC specific code
- CRaC support -> will forward all CRaC specific calls to jdk.crac



implementation 'org.crac:crac:1.5.0'

Malen

github.com/CRaC/org.crac





Upgrade (Haswell -> restore: Ice Lake, no problem)

- Upgrade (Haswell -> restore: Ice Lake, no problem)
- ODOWNGrade (Ice Lake -> restore: Haswell, problematic)

- Upgrade (Haswell -> restore: Ice Lake, no problem)
- Downgrade (Ice Lake -> restore: Haswell, problematic)
- Solved in CRaC by specific flag (little drop in performance)

- Upgrade (Haswell -> restore: Ice Lake, no problem)
- Oowngrade (Ice Lake -> restore: Haswell, problematic)
- Solved in CRaC by specific flag (little drop in performance)
- Node groups stick to same cpu architecture

- Upgrade (Haswell -> restore: Ice Lake, no problem)
- Downgrade (Ice Lake -> restore: Haswell, problematic)
- Solved in CRaC by specific flag (little drop in performance)
- Node groups stick to same cpu architecture
- Virtualized Linux environments work on all OS's (as long as cpu architecture is x64/aarch64)

Helidon (experimental support)

- Helidon (experimental support)
- Ouarkus (rudimentary support)

- Helidon (experimental support)
- Ouarkus (rudimentary support)
- Micronaut (fair support)

- Helidon (experimental support)
- Ouarkus (rudimentary support)
- Micronaut (fair support)
- Spring 6.1+/SpringBoot 3.2+ (good support)



SPRINGBOOT3.2 PETCLINIC

> java -jar spring-petclinic-3.2.0.jar

START APPLICATION

```
> java -jar spring-petclinic-3.2.0.jar
:: Built with Spring Boot :: 3.2.0
2023-11-29T11:57:27.579+01:00 INFO 3839 --- [
                                                        main] o.s.d.j.r.query.QueryEnhancerFactory
                                                                                                       : Hibernate is in classpath; If
applicable, HQL parser will be used.
                                                        main] o.s.b.a.e.web.EndpointLinksResolver
                                                                                                       : Exposing 13 endpoint(s) beneath
2023-11-29T11:57:28.549+01:00 INFO 3839 --- [
base path '/actuator'
2023-11-29T11:57:28.625+01:00 INFO 3839 --- [
                                                        main] o.s.b.w.embedded.tomcat.TomcatWebServer : Tomcat started on port 8080 (http)
with context path ''
                                                        main] o.s.s.petclinic.PetClinicApplication
                                                                                                       : Started PetClinicApplication in
2023-11-29T11:57:28.639+01:00 INFO 3839 ---
4.619 seconds (process running for 5.051)
Started up in 4099ms with PID: 3839
```

START FROM

AUTO CHECKPOINT

Feature in SpringBoot 3.2+

AUTO CHECKPOINT

- Feature in SpringBoot 3.2+
- Start with -Dspring.context.checkpoint=onRefresh

- Feature in SpringBoot 3.2+
- Start with -Dspring.context.checkpoint=onRefresh
- Creates automatic checkpoint after start of SpringBoot framework

- Feature in SpringBoot 3.2+
- Start with -Dspring.context.checkpoint=onRefresh
- Creates automatic checkpoint after start of SpringBoot framework
- Right before the application will be started

> java -Dspring.context.checkpoint=onRefresh -XX:CRaCCheckpointTo=./tmp_auto_checkpoint -jar spring-petclinic-3.2.0.jar

START APPLICATION AND CREATE CHECKPOINT

```
> java -Dspring.context.checkpoint=onRefresh -XX:CRaCCheckpointTo=./tmp auto checkpoint -jar spring-petclinic-3.2.0.jar
> java -XX:CRaCRestoreFrom=./tmp auto checkpoint
2023-11-29T12:01:37.698+01:00 WARN 15261 --- [1-1 housekeeper] com.zaxxer.hikari.pool.HikariPool
                                                                                                        : HikariPool-1 - Thread starvation
or clock leap detected (housekeeper delta=1h26m17s198ms377µs333ns).
                                                         main] o.s.c.support.DefaultLifecycleProcessor : Restarting Spring-managed
2023-11-29T12:01:37.790+01:00 INFO 15261 --- [
lifecycle beans after JVM restore
2023-11-29T12:01:37.811+01:00 INFO 15261 --- [
                                                         main] o.s.b.w.embedded.tomcat.TomcatWebServer : Tomcat started on port 8080 (http)
with context path ''
2023-11-29T12:01:37.834+01:00 INFO 15261 --- [
                                                         main] o.s.s.petclinic.PetClinicApplication
                                                                                                        : Restored PetClinicApplication in
0.956 seconds (process running for 0.958)
Started up in 697ms with PID: 15261
```

RESTORE FROM CHECKPOINT

START FROM

Start application with -xx:cracCheckpointTo=Path

- Start application with -xx:cracCheckpointTo=Path
- Warm up your application

- Start application with -xx:cracCheckpointTo=Path
- Warm up your application
- Create checkpoint using jcmd

- Start application with -xx:cracCheckpointTo=Path
- Warm up your application
- Create checkpoint using jcmd
- Checkpoint now also contains application

> java -XX:CRaCCheckpointTo=./tmp_manual_checkpoint -jar spring-petclinic-3.2.0.jar

START APPLICATION

```
> java -XX:CRaCCheckpointTo=./tmp_manual_checkpoint -jar spring-petclinic-3.2.0.jar

...

2023-11-29T11:57:28.625+01:00 INFO 3839 --- [ main] o.s.b.w.embedded.tomcat.TomcatWebServer : Tomcat started on port 8080 (http) with context path ''

2023-11-29T11:57:28.639+01:00 INFO 3839 --- [ main] o.s.s.petclinic.PetClinicApplication : Started PetClinicApplication in 4.619 seconds (process running for 5.051)

Started up in 4099ms with PID: 3839
```

> jcmd 3839 JDK.checkpoint

CREATE CHECKPOINT

> java -XX:CRaCRestoreFrom=./tmp_manual_checkpoint

RESTORE FROM CHECKPOINT

```
> java -XX:CRaCRestoreFrom=./tmp_manual_checkpoint

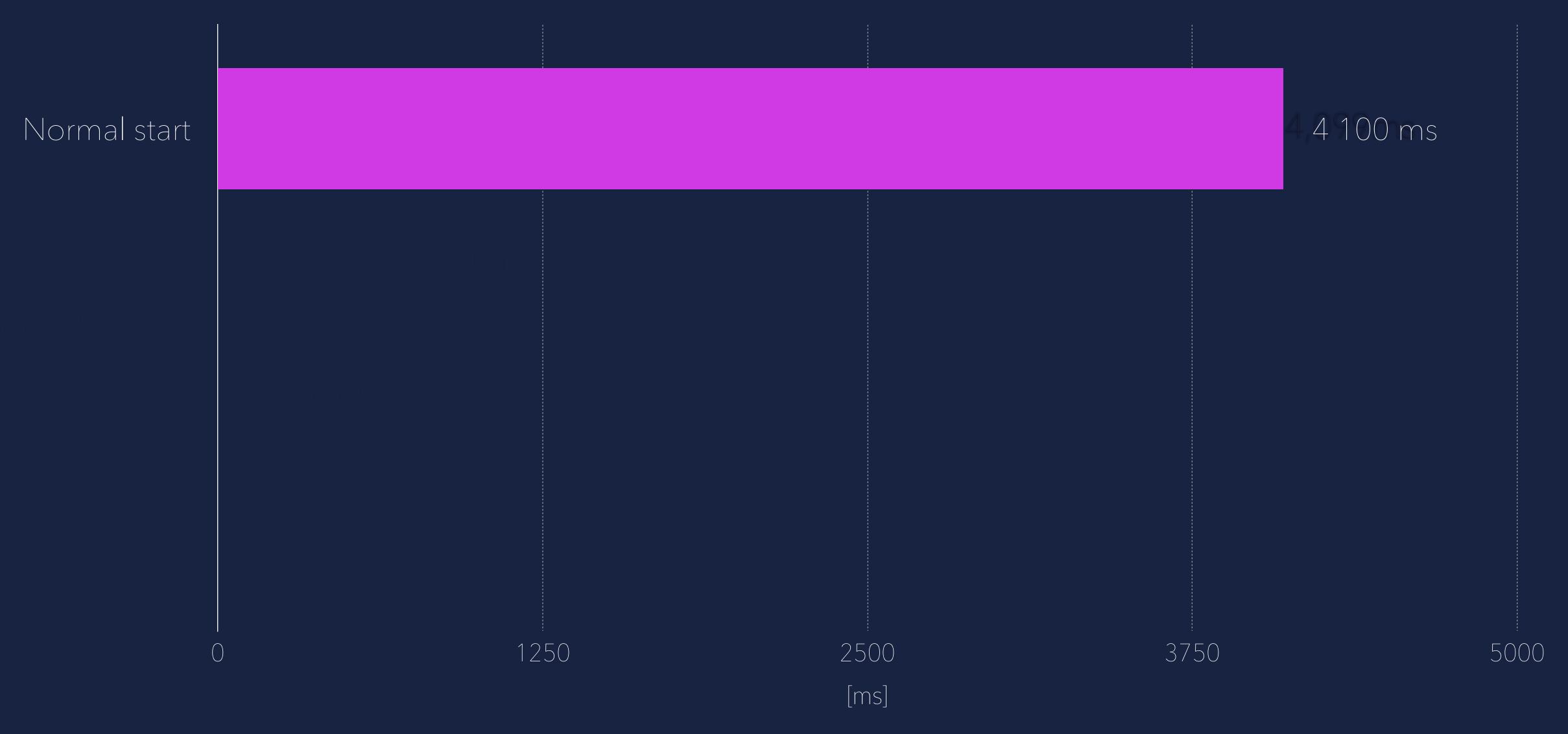
2023-11-29T12:04:32.626+01:00 WARN 15512 --- [l-1 housekeeper] com.zaxxer.hikari.pool.HikariPool : HikariPool-1 - Thread starvation or clock leap detected (housekeeper delta=1h28m32s17ms487µs256ns).

2023-11-29T12:04:32.634+01:00 INFO 15512 --- [Attach Listener] o.s.c.support.DefaultLifecycleProcessor : Restarting Spring-managed lifecycle beans after JVM restore

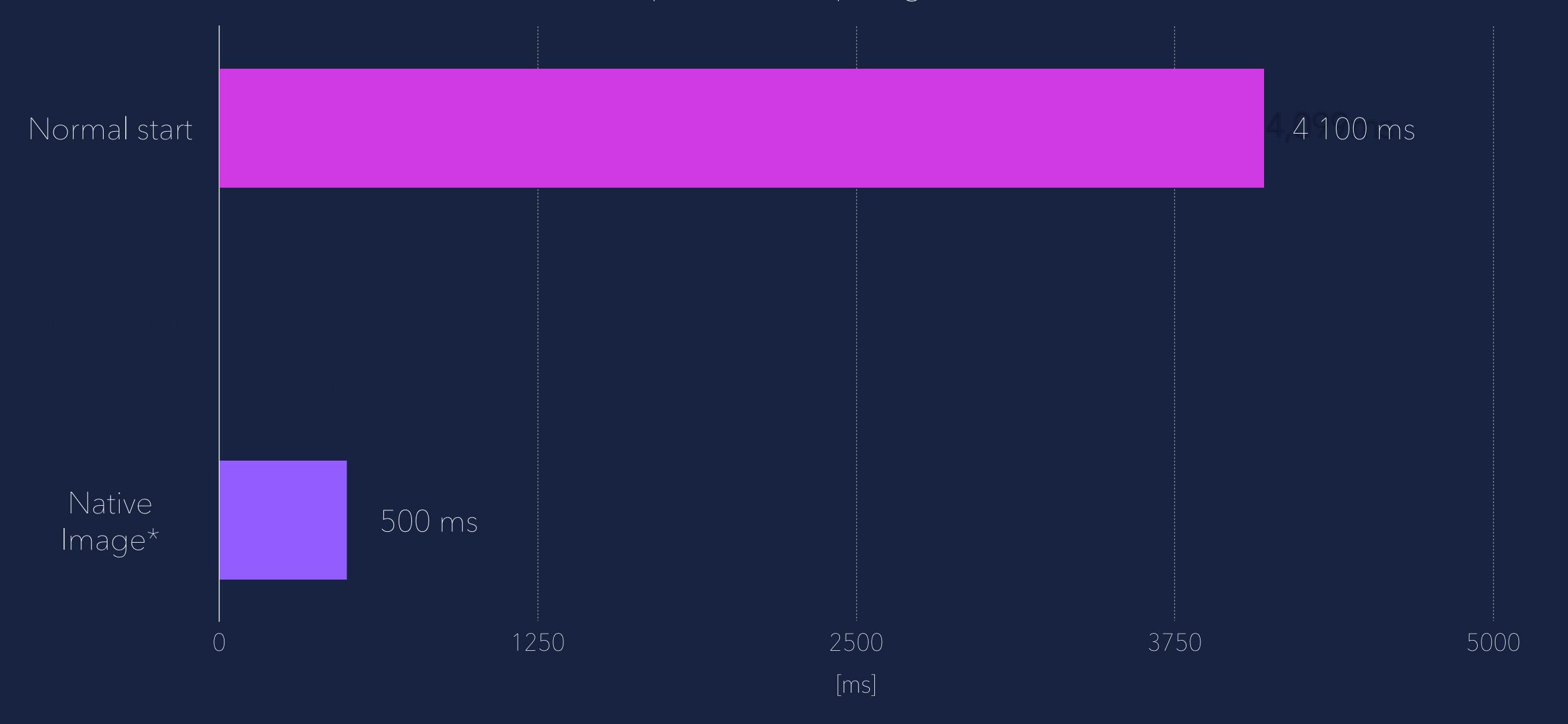
2023-11-29T12:04:32.642+01:00 INFO 15512 --- [Attach Listener] o.s.b.w.embedded.tomcat.TomcatWebServer : Tomcat started on port 8080 (http) with context path ''

2023-11-29T12:04:32.644+01:00 INFO 15512 --- [Attach Listener] o.s.c.support.DefaultLifecycleProcessor : Spring-managed lifecycle restart completed (restored JVM running for 301 ms)
```



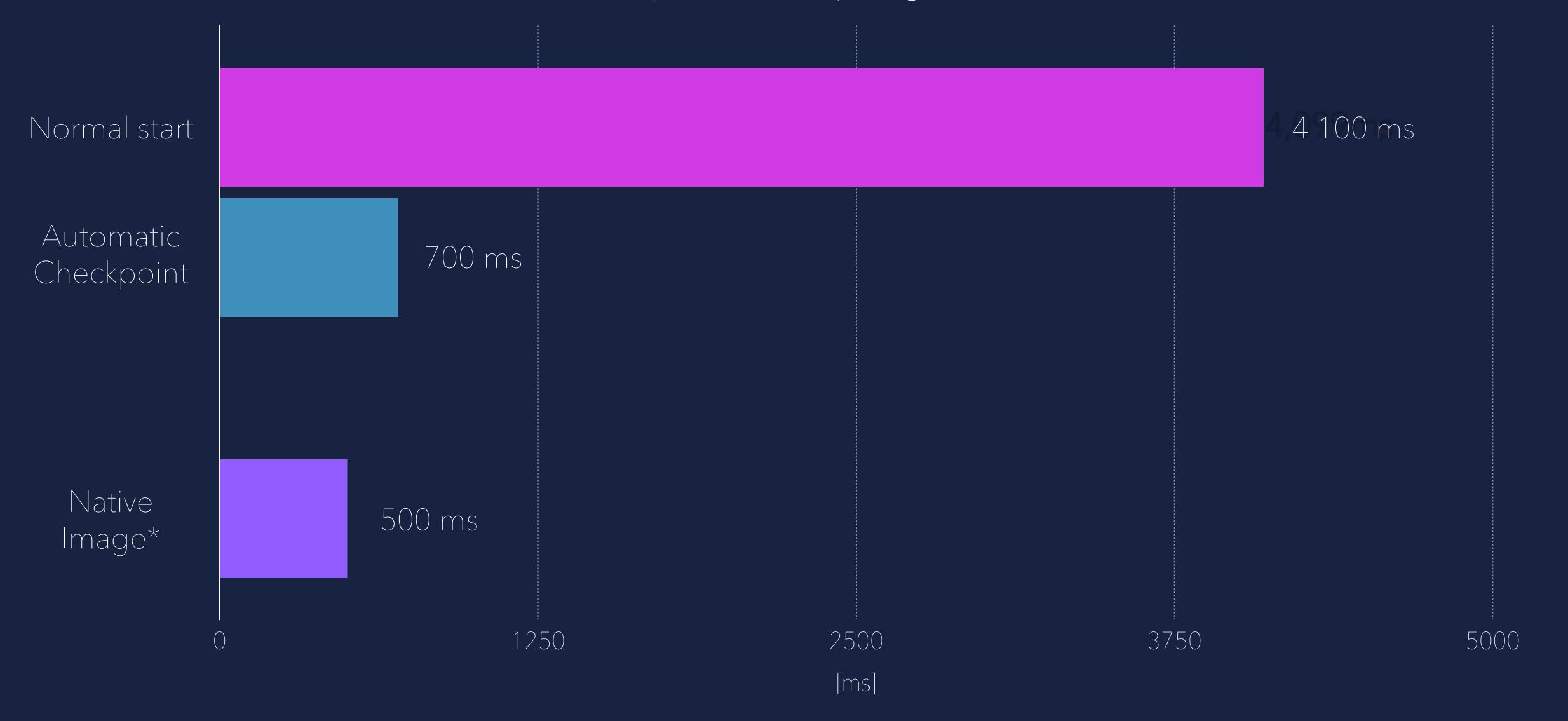


Startup time of SpringBoot PetClinic



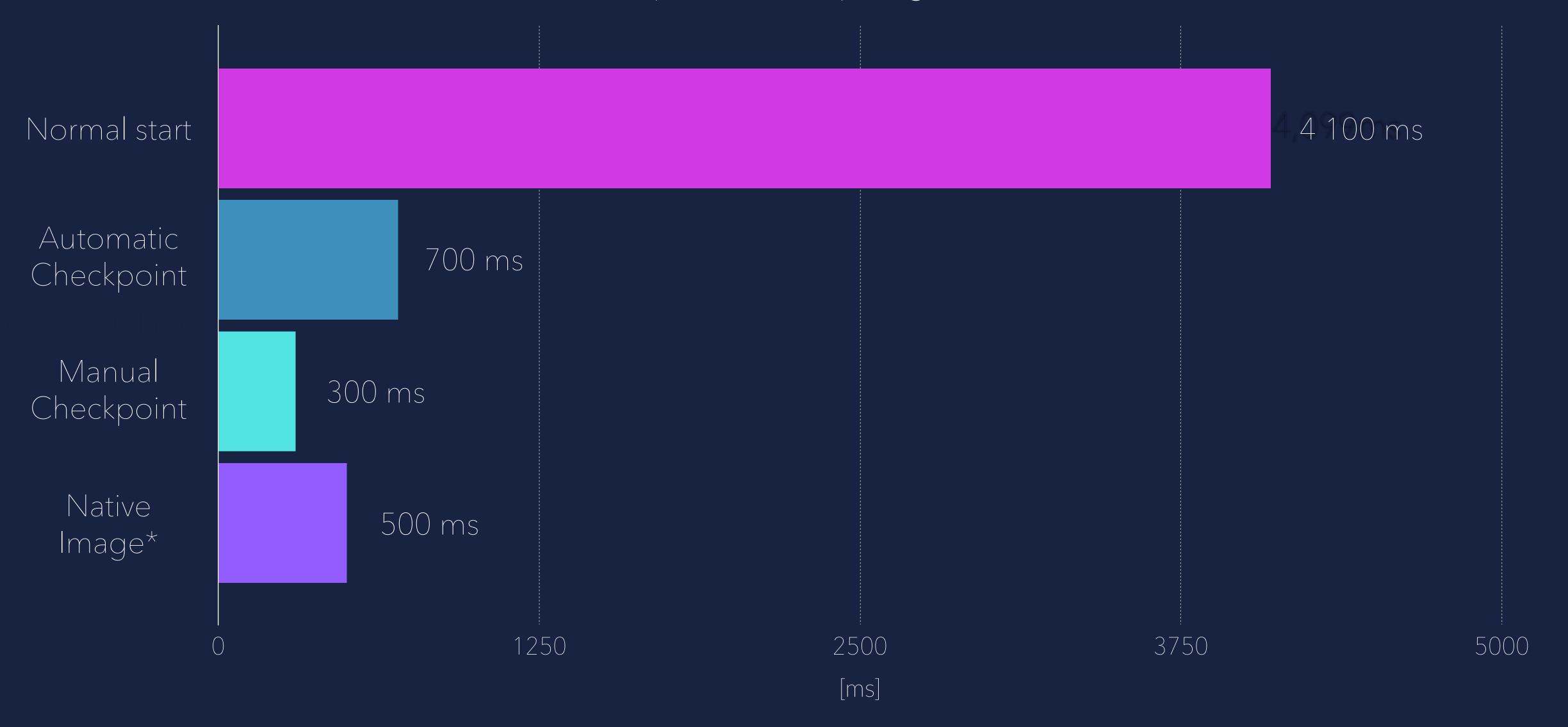


Startup time of SpringBoot PetClinic





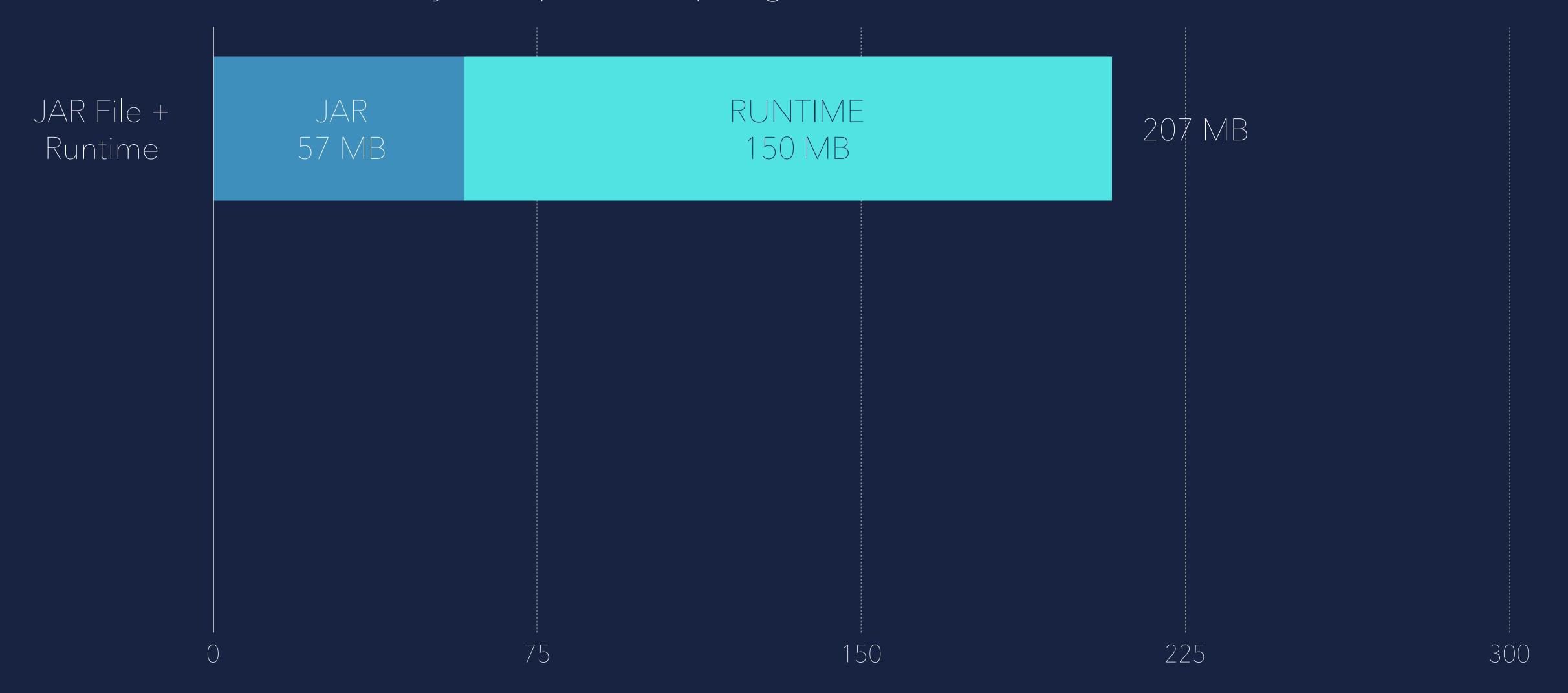
Startup time of SpringBoot PetClinic





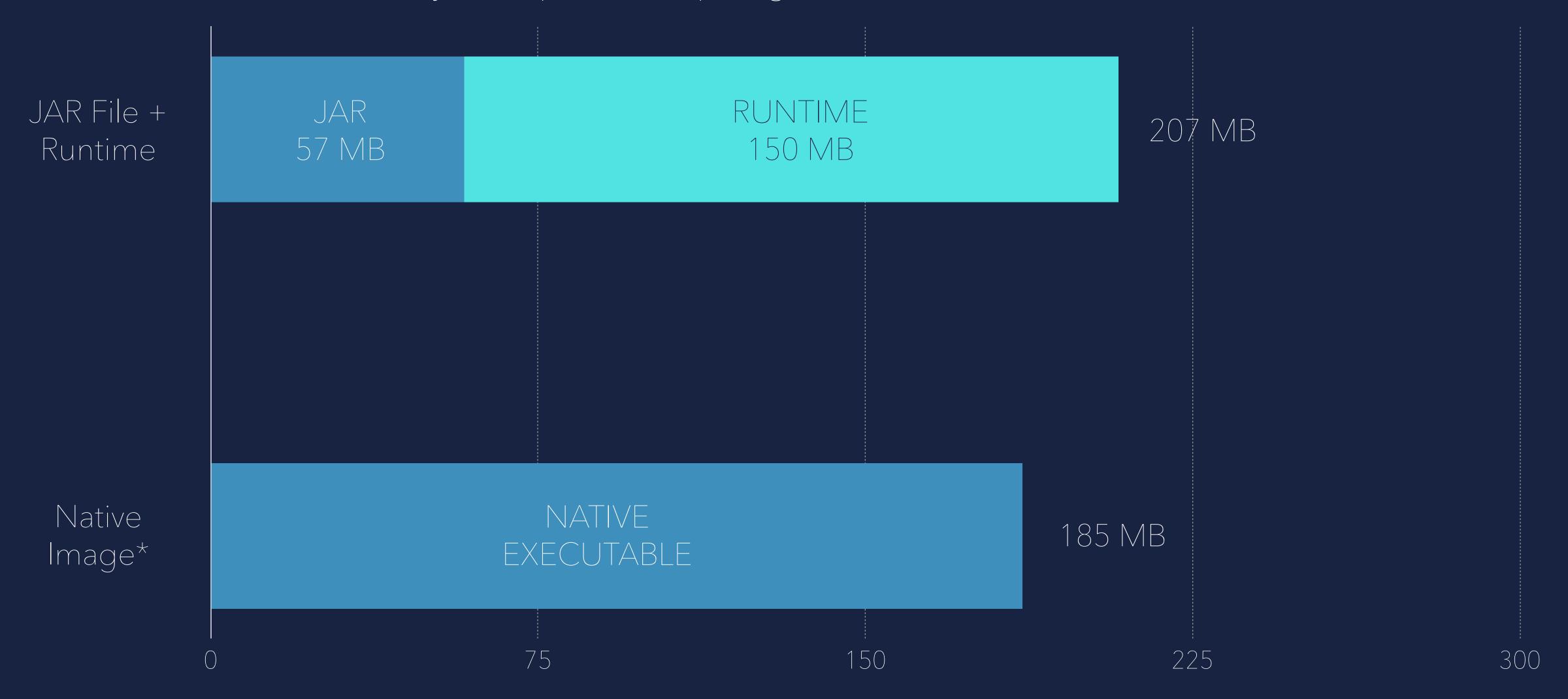
FOOTPRINT.

Memory Footprint of SpringBoot PetClinic (deployment artifact)



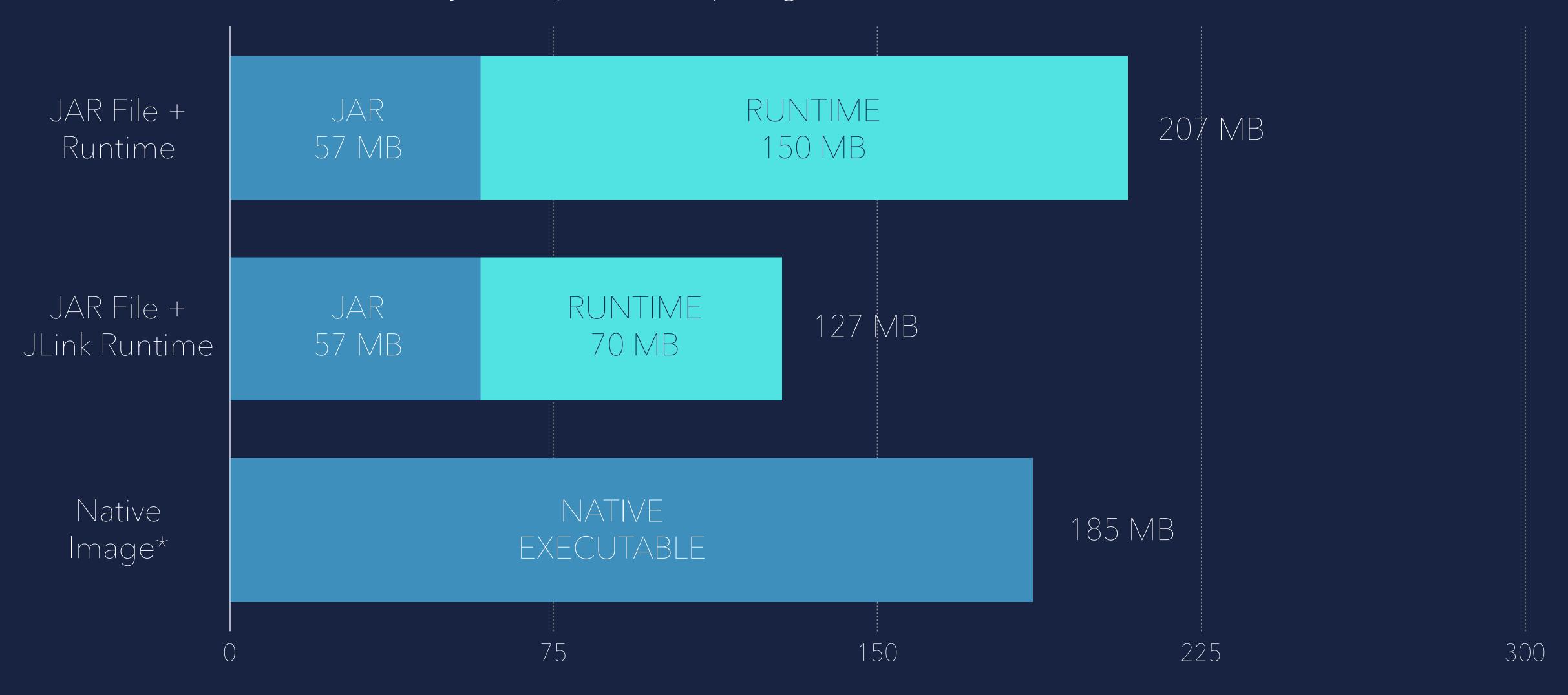


Memory Footprint of SpringBoot PetClinic (deployment artifact)



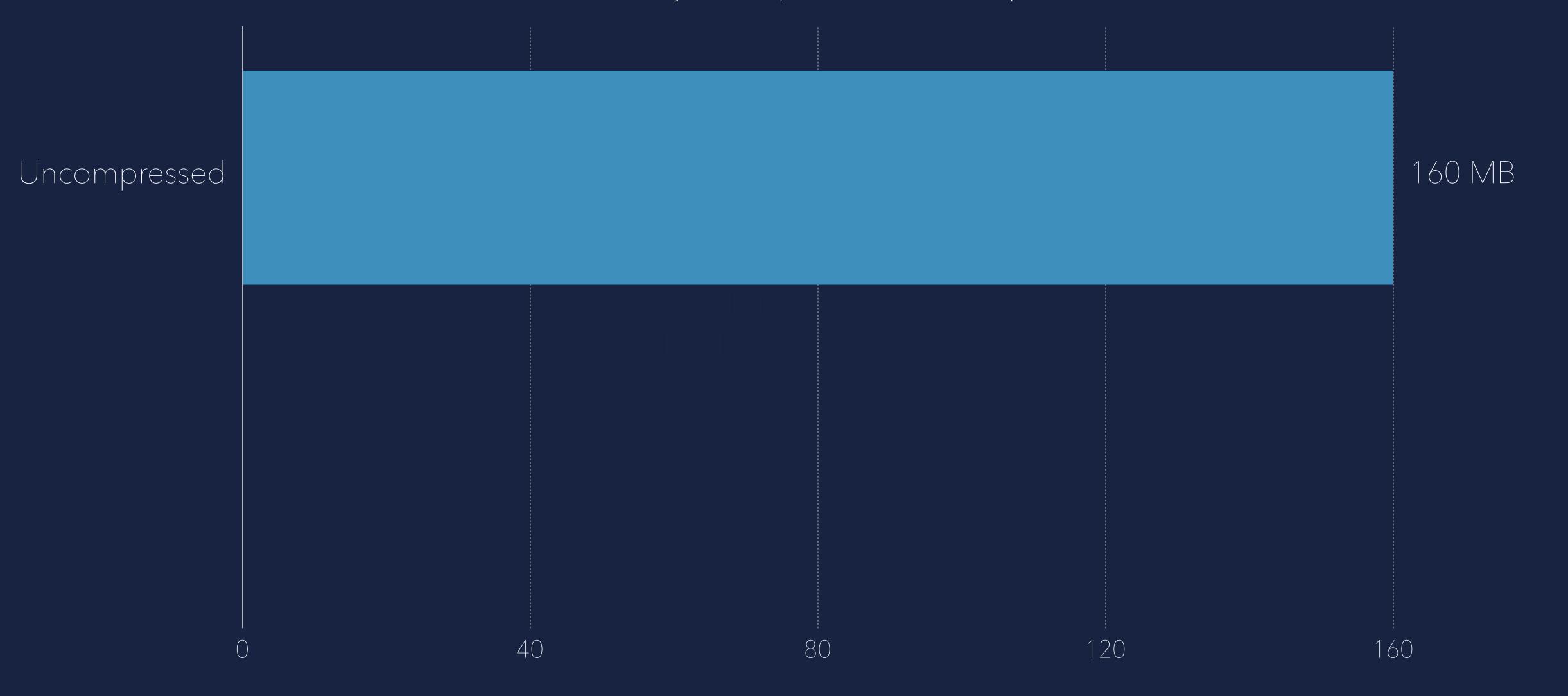


Memory Footprint of SpringBoot PetClinic (deployment artifact)

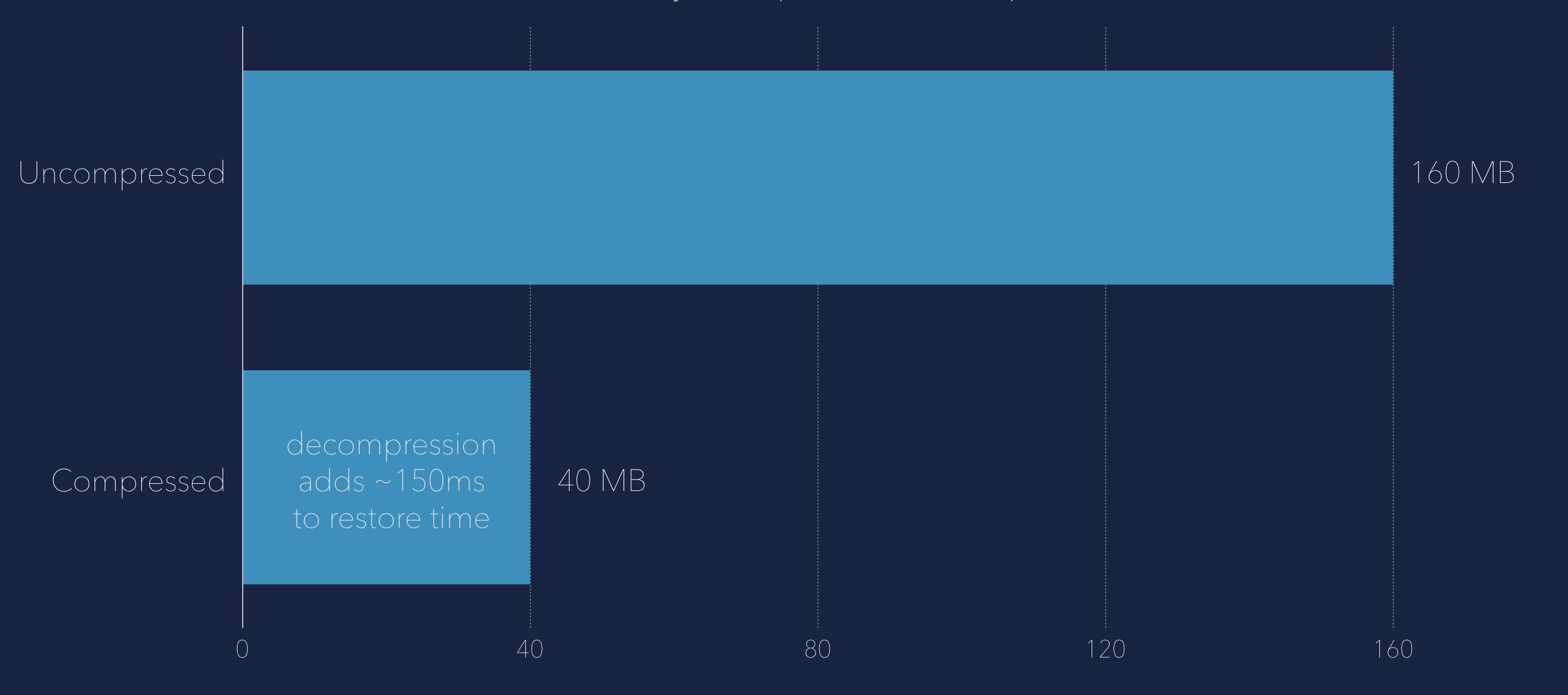




Memory Footprint of checkpoint

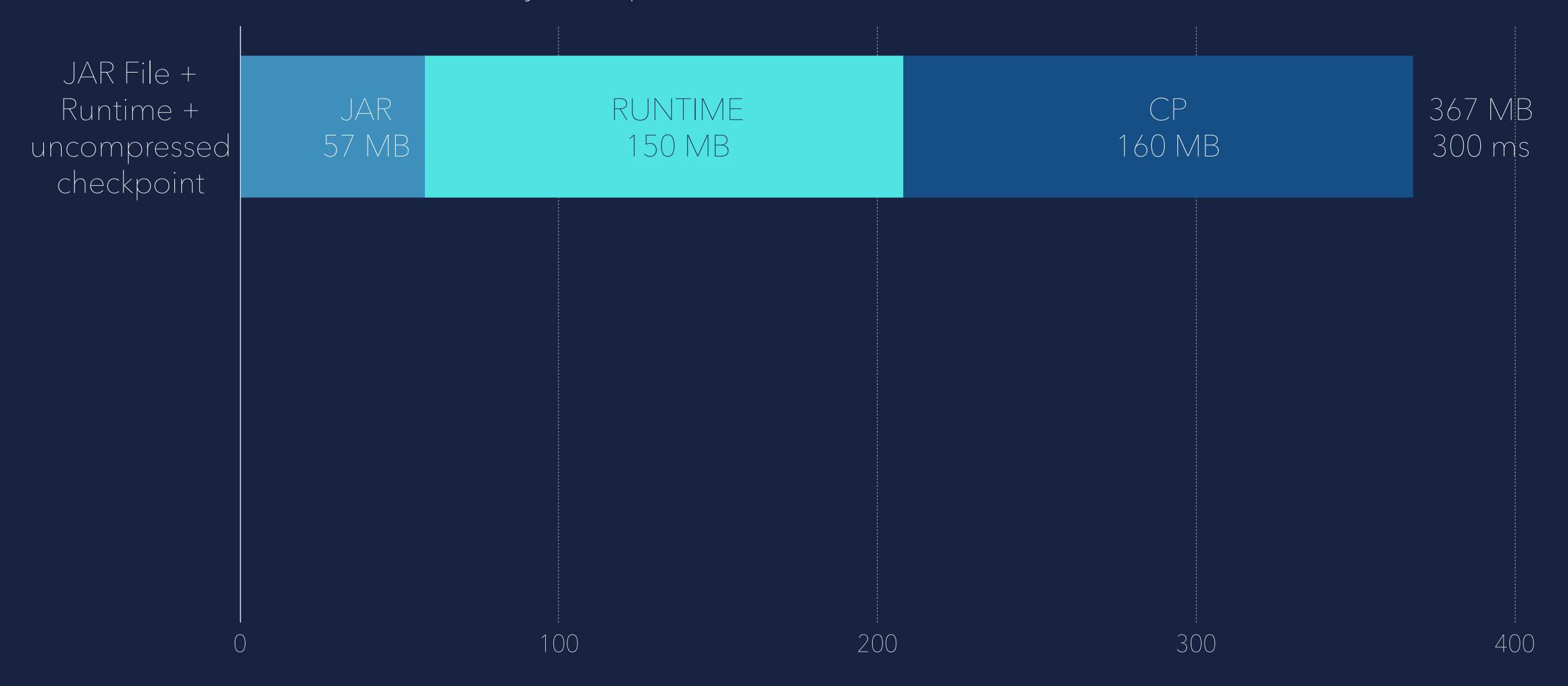


Memory Footprint of checkpoint



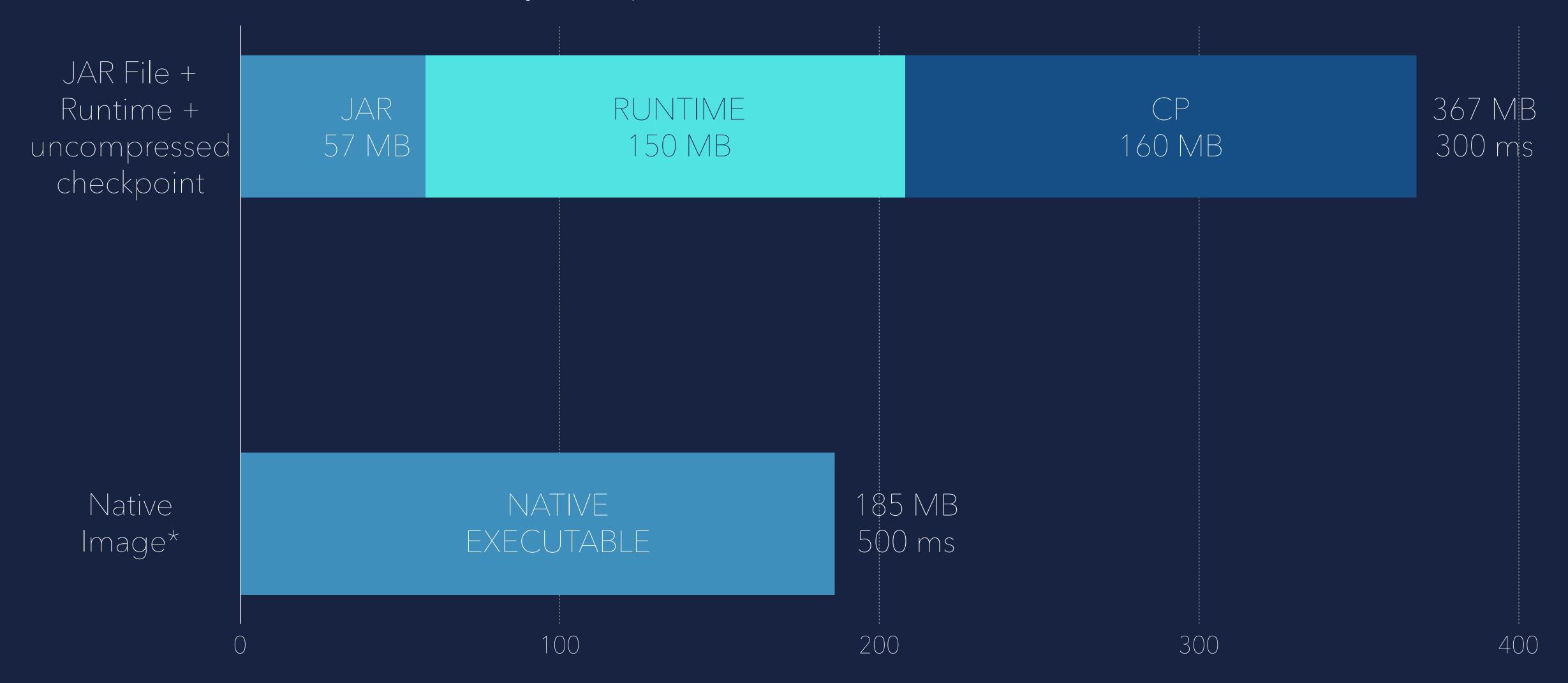


Memory Footprint (SpringBoot PetClinic artifact + checkpoint)



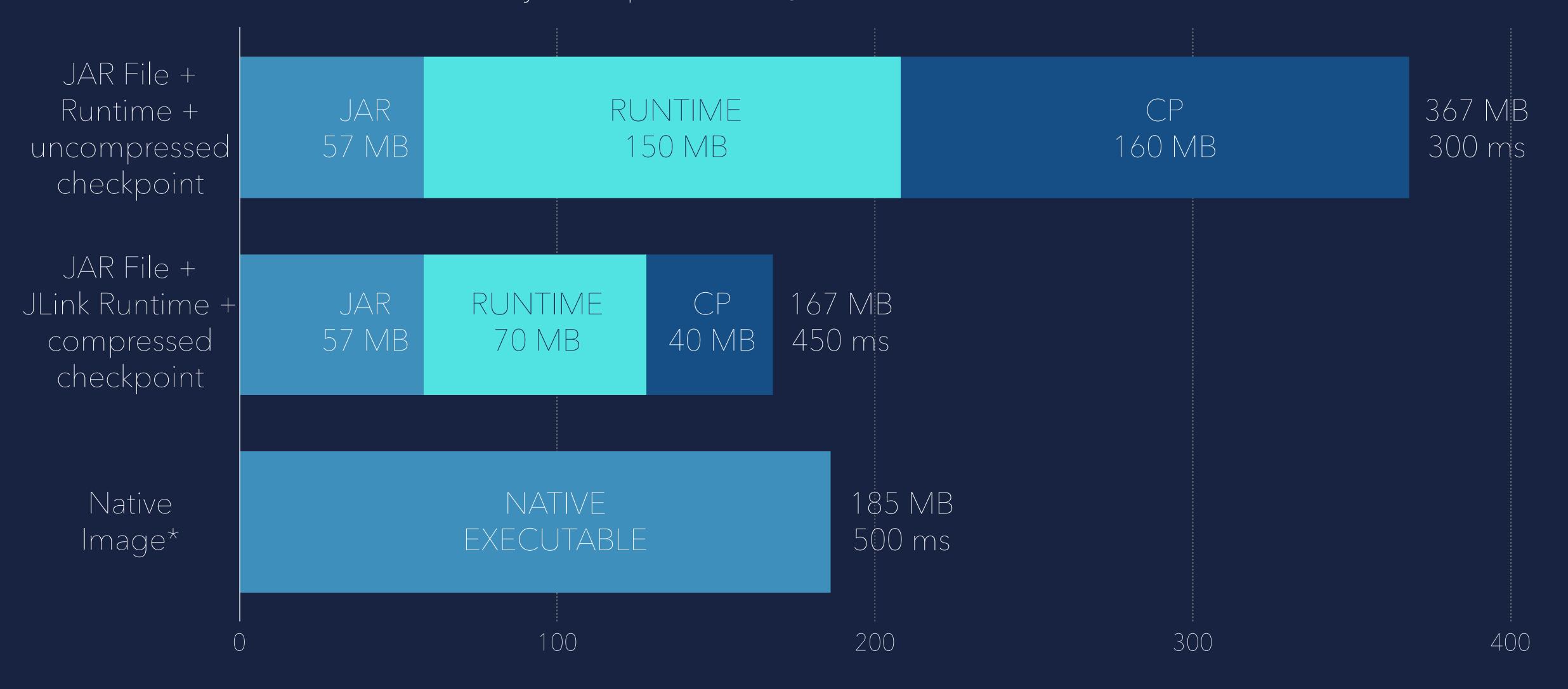


Memory Footprint (SpringBoot PetClinic artifact + checkpoint)





Memory Footprint (SpringBoot PetClinic artifact + checkpoint)





• CRaC engines (already released)

CRIU

- Default engine
- Must be run as root user
- -XX:CRaCEngine=criu (default)
- Runs the same ID as the original process
- Needs privileged access in containers for checkpointing
- Needs CHECKPOINT_RESTORE and SYS_PTRACE for restoring

CRAC ENGINES

WARP

- CRIU independent engine
- Only available in Azul Zulu for x64 (aarch64 will follow)
- -XX:CRaCEngine=warp
- Doesn't run with the same PID as original process
- No privileged access needed
- Optimizations related to compressed images

- CRaC engines (already released)
- Encryption

openjdk.org/projects/crac

Lead by Anton Kozlov (Azul)



CRaC is a way to pause and restore a JVM based application

- CRaC is a way to pause and restore a JVM based application
- It doesn't require a closed world as with a native image

- CRaC is a way to pause and restore a JVM based application
- It doesn't require a closed world as with a native image
- Extremely fast time to full performance level

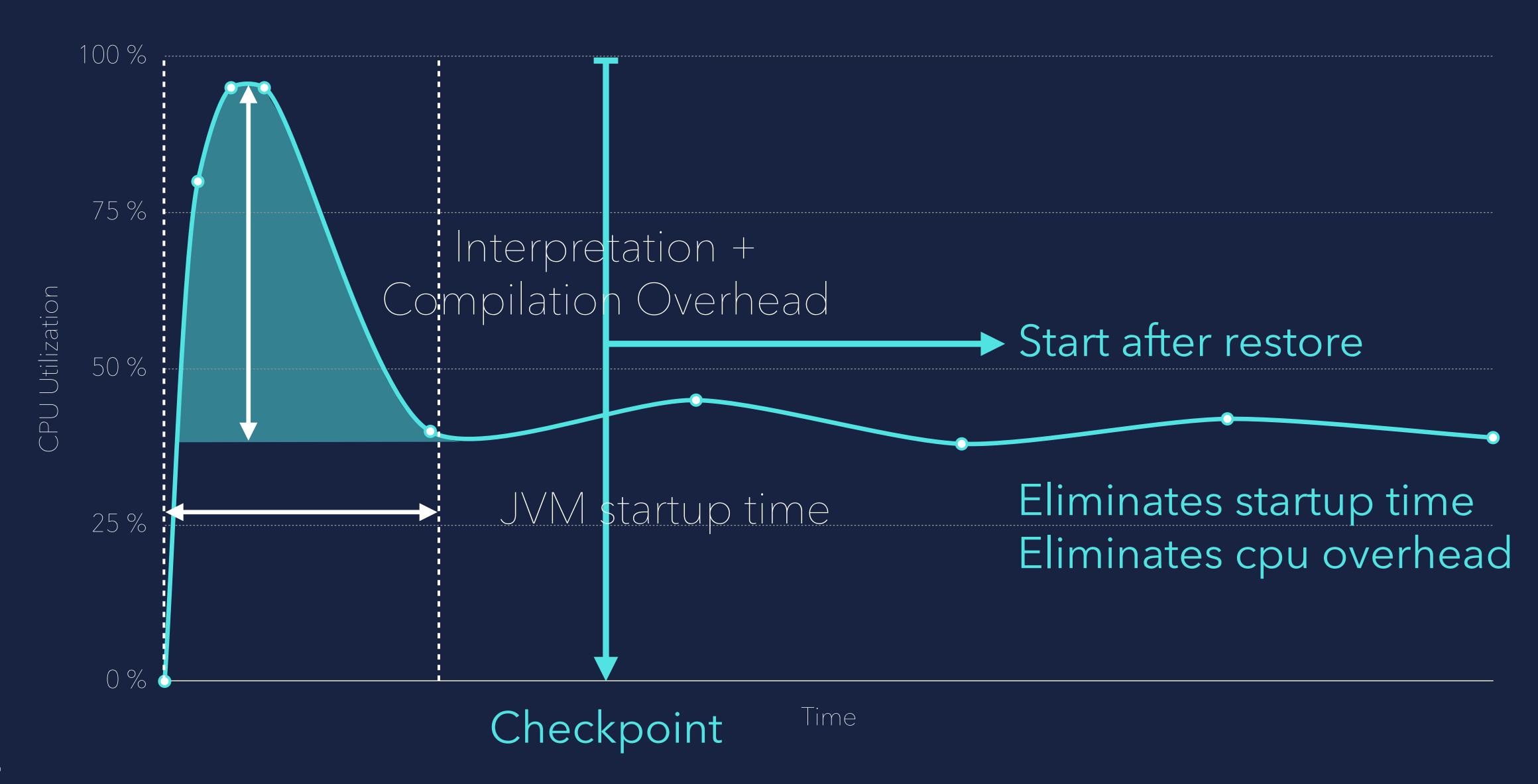
- CRaC is a way to pause and restore a JVM based application
- It doesn't require a closed world as with a native image
- Extremely fast time to full performance level
- No need for hotspot identification, method compiles, recompiles and deoptimisations

- CRaC is a way to pause and restore a JVM based application
- It doesn't require a closed world as with a native image
- Extremely fast time to full performance level
- No need for hotspot identification, method compiles, recompiles and deoptimisations
- Improved throughput from start

- CRaC is a way to pause and restore a JVM based application
- It doesn't require a closed world as with a native image
- Extremely fast time to full performance level
- No need for hotspot identification, method compiles, recompiles and deoptimisations
- Improved throughput from start
- CRaC is an OpenJDK project

- CRaC is a way to pause and restore a JVM based application
- It doesn't require a closed world as with a native image
- Extremely fast time to full performance level
- No need for hotspot identification, method compiles, recompiles and deoptimisations
- Improved throughput from start
- CRaC is an OpenJDK project
- CRaC can save infrastructure cost

INFRASTRUCTURE COST



RIVOW MORE?

github.com/CRaC



JDK 17 / 21 / 23 LINUX X64 / AARCH64



