Lilliput на горизонте

JEP 450: Compact Object Headers

Максим Дегтярёв
HUAWEI

Maccimo 13 августа 2022 в 21:00

Атака на String.hashCode: прообразы и коллизии

Ненормальное программирование, Программирование, Java

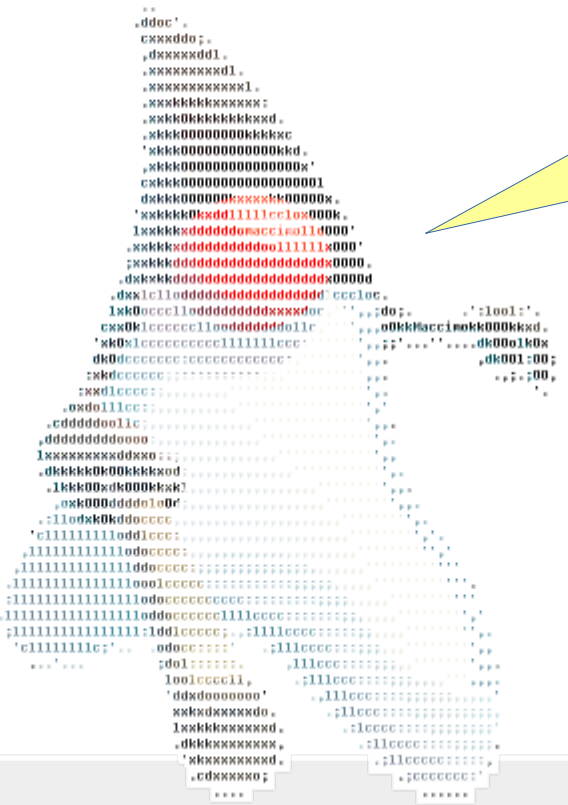
Технотекст 2022 | Recovery Mode | Сезон Java

Maccimo 13 августа 2022 в 21:00

Атака на String.hashCode: прообразы и коллизии

Ненормальное программирование, Программирование, Java

Технотекст 2022 | Recovery Mode | Сезон Java



"Joker 2023 (CHEE=RRM) "

Maccimo 13 августа 2022 в 21:00

Атака на String.hashCode: прообразы и коллизии

Ненормальное программирование, Программирование, Java

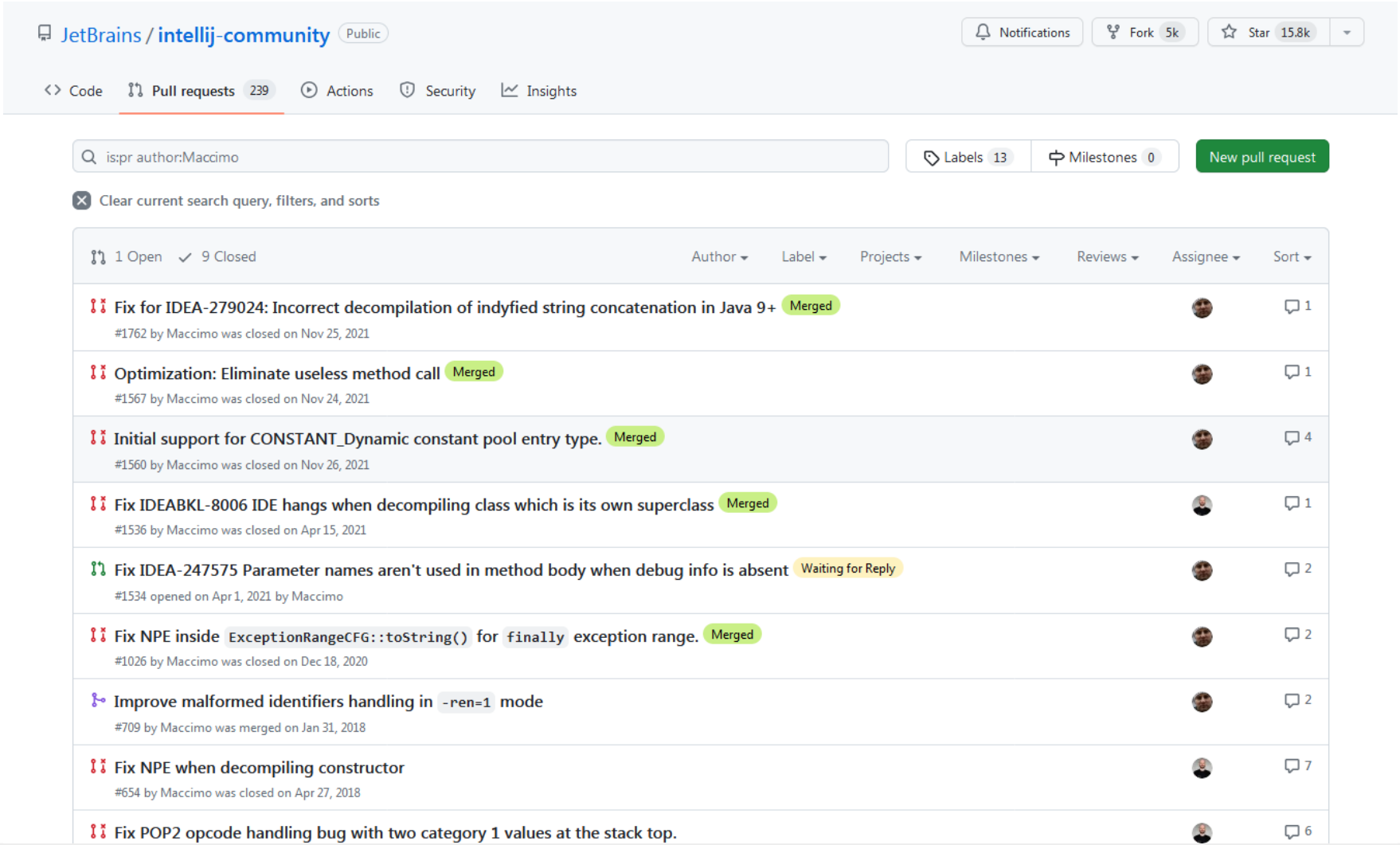
Технотекст 2022 | Recovery Mode | Сезон Java

"Joker 2023 (CHEE=RRM) "

https://habrahabr.ru/post/674816

Decompiled .class file, bytecode version: 65.0 (Java 21) Choose Sources...

```
1   //
2   // Source code recreated from a .class file by IntelliJ IDEA
3   // (powered by FernFlower decompiler)
4   //
5
6   public class Joker2023 {
7       public Joker2023() {
8       }
9
10      public static void main(String... args) {  args: []
11          System.out.println("Конференция для опытных Java-разработчиков".hashCode());
12          System.out.println("Joker 2023 (CHEE=RRM) Конференция для опытных Java-разработчиков".hashCode());
13      }
14  }
15
```


JetBrains / intellij-community Public
Notifications
Fork 5k
Star 15.8k
Code
Pull requests 239
Actions
Security
Insights
is:pr author:Maccimo
Labels 13
Milestones 0
New pull request
Clear current search query, filters, and sorts
1 Open
9 Closed
Author
Label
Projects
Milestones
Reviews
Assignee
Sort
Fix for IDEA-279024: Incorrect decompilation of indyfied string concatenation in Java 9+ Merged
#1762 by Maccimo was closed on Nov 25, 2021
1
Optimization: Eliminate useless method call Merged
#1567 by Maccimo was closed on Nov 24, 2021
1
Initial support for CONSTANT_Dynamic constant pool entry type. Merged
#1560 by Maccimo was closed on Nov 26, 2021
4
Fix IDEABKL-8006 IDE hangs when decompiling class which is its own superclass Merged
#1536 by Maccimo was closed on Apr 15, 2021
1
Fix IDEA-247575 Parameter names aren't used in method body when debug info is absent Waiting for Reply
#1534 opened on Apr 1, 2021 by Maccimo
2
Fix NPE inside ExceptionRangeCFG::toString() for finally exception range. Merged
#1026 by Maccimo was closed on Dec 18, 2020
2
Improve malformed identifiers handling in -ren=1 mode
#709 by Maccimo was merged on Jan 31, 2018
2
Fix NPE when decompiling constructor
#654 by Maccimo was closed on Apr 27, 2018
7
Fix POP2 opcode handling bug with two category 1 values at the stack top.
6

- Экспериментальный статус
- OpenJDK HotSpot JVM
- AMD64 (Intel64) и Aarch64
- Доступны тестовые сборки
- Дата релиза не определена

План доклада

План доклада

План доклада

Объект:

Объект:

- Определённый тип

Объект:

- Определённый тип
- Неизменный Identity Hash Code

Объект:

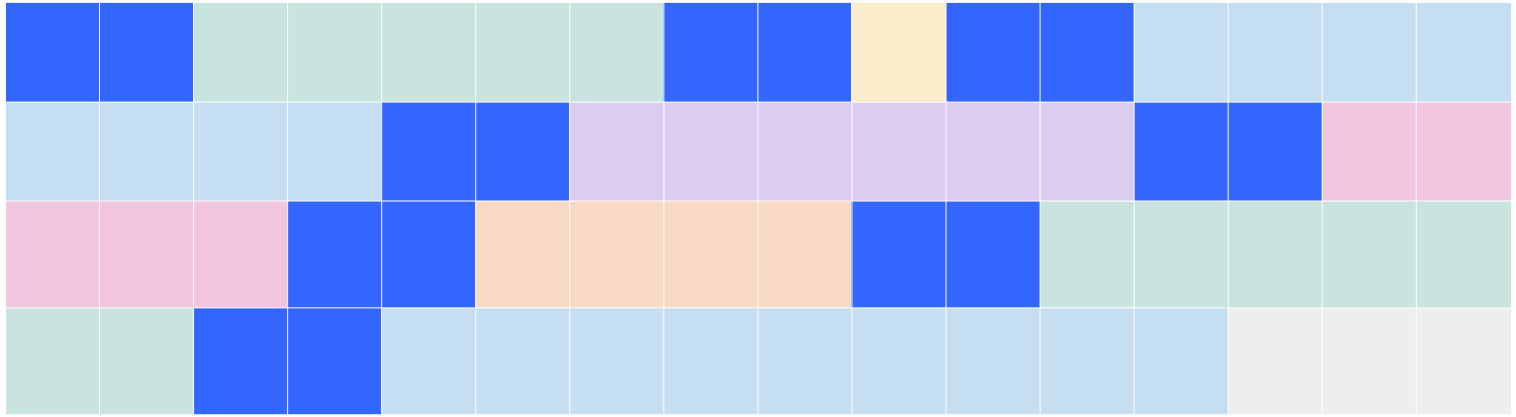
- Определённый тип
- Неизменный Identity Hash Code
- Может быть примитивом синхронизации

Объект:

- Определённый тип
- Неизменный Identity Hash Code
- Может быть примитивом синхронизации
- Автоматическое управление памятью (GC)

Где хранить?

Заголовок!



Заголовок объекта

Заголовок объекта

- Есть у каждого объекта

Заголовок объекта

- Есть у каждого объекта
- Располагается перед данными объекта

Заголовок объекта

- Есть у каждого объекта
- Располагается перед данными объекта
- Размер заголовка — 2 указателя

Заголовок объекта

- Есть у каждого объекта
- Располагается перед данными объекта
- Размер заголовка — 2 указателя
- На 64-битных платформах равен 16 байтам (128 бит)

Сферический объект в вакууме

Lilliput Experiment Results

Created by Roman Kennke, last modified on Apr 20, 2023

```
All measures are averages over the given number of samples.
Smpls:    Number of samples taken
Num-objs: Average number of objects in the heap
Size:     Average size of objects in the heap, in bytes
LDS:      Live-data-set: Average number of bytes occupied by live (i.e. reachable) objects in the heap
IHashed:  Number of objects that have an identity hashcode
Locked:   Number of locked objects
%Ihashed: How many % of live objects have an identity hashcode
%Locked:  How many % of live objects have been locked
```

https://wiki.openjdk.org/display/lilliput/Lilliput+Experiment+Results

Сферический объект в вакууме

- Типичный объём пользовательских данных в объекта — от 32 до 64 байтов

Сферический объект в вакууме

- Типичный объём пользовательских данных в объекта — от 32 до 64 байтов
- До 20% памяти — служебная информация

128

128

64

128

64

32

Зачем?

Зачем?

- Экономия до 10% памяти

Зачем?

- Экономия до 10% памяти
- Улучшение локальности данных

Зачем?

- Экономия до 10% памяти
- Улучшение локальности данных
- Ускорение из-за локальности данных

Что в заголовке?

Что в заголовке?

Что в заголовке?

Mark Word

Что в заголовке?

Mark Word

Class Word

Что в заголовке?

Mark Word

Class Word

Что в заголовке?

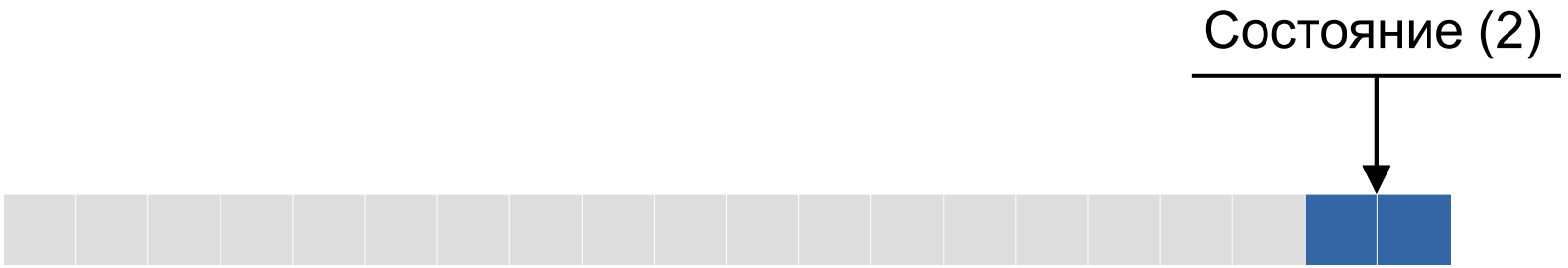
Mark Word

Class Word

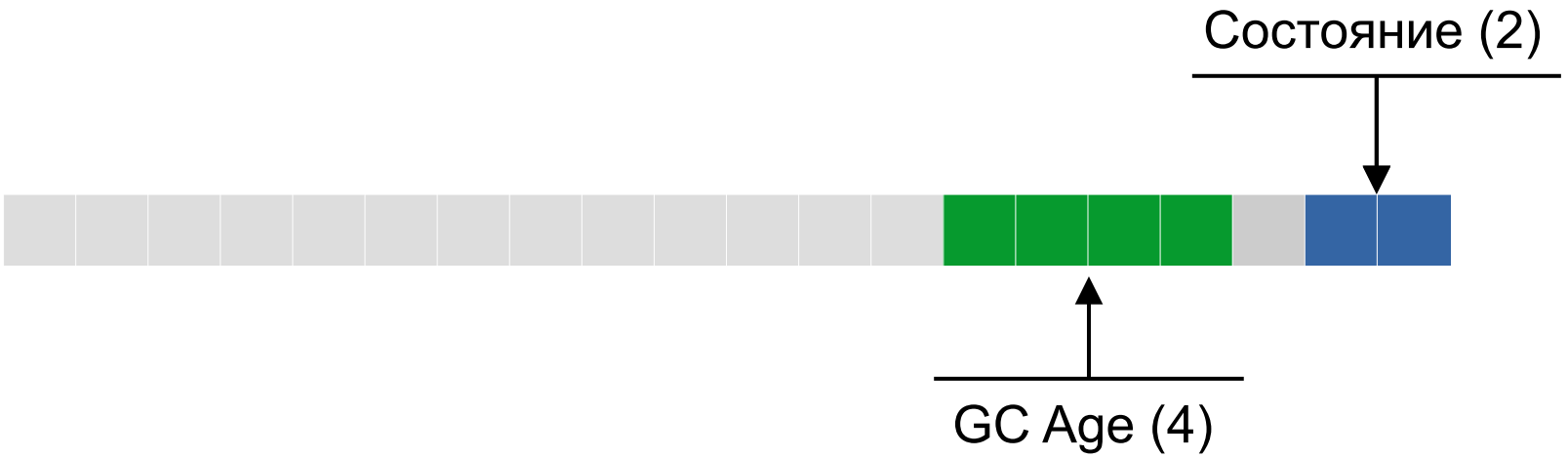
Mark Word

Mark Word

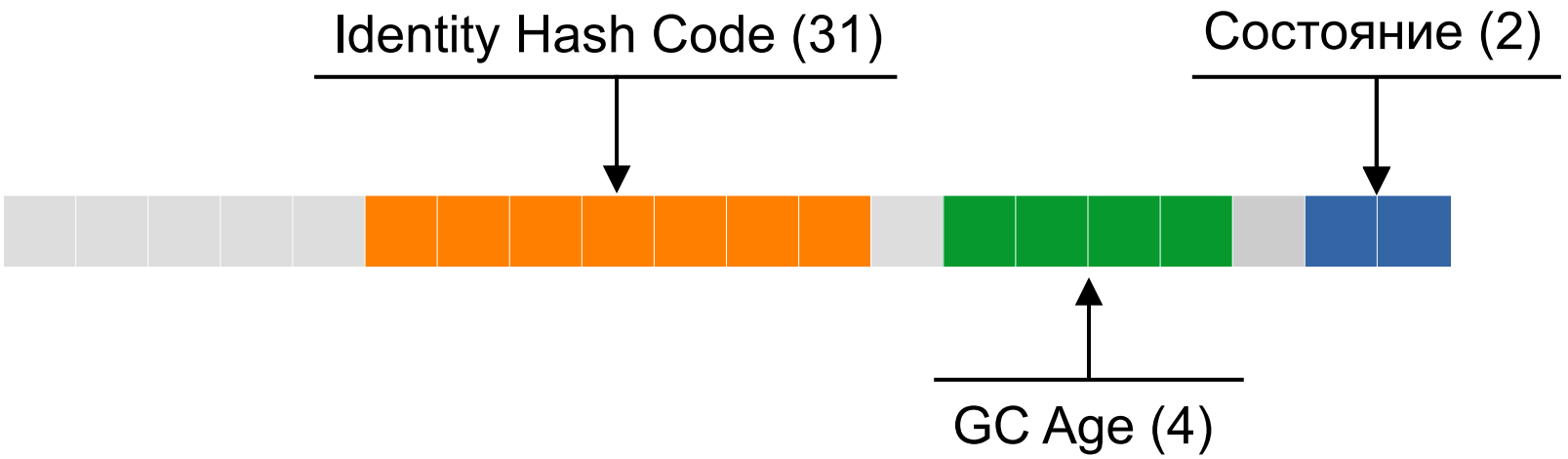
Mark Word — Normal



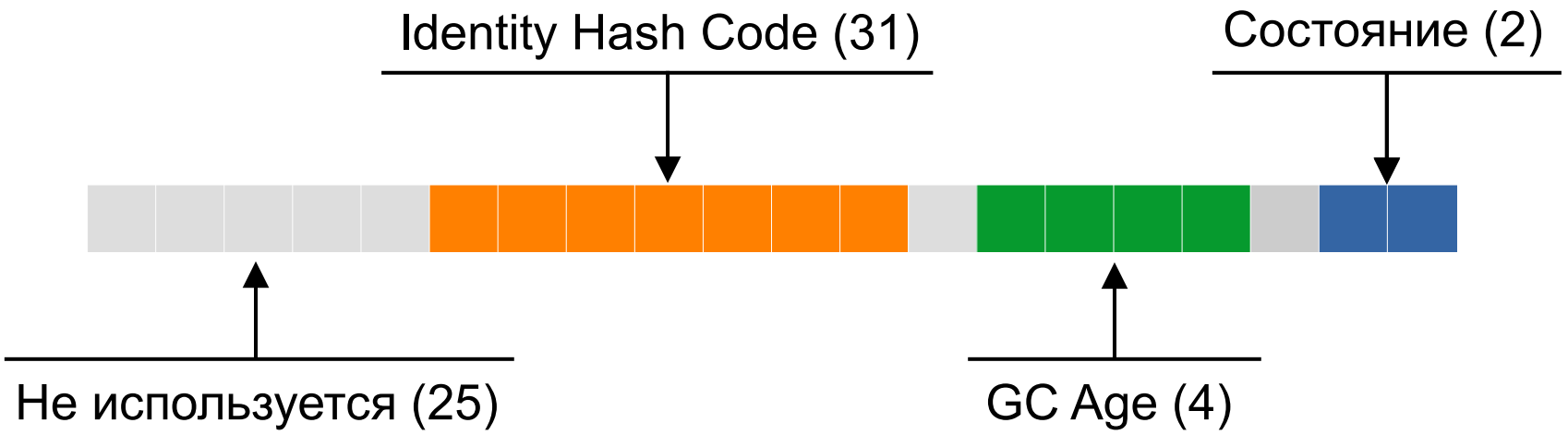
Mark Word — Normal



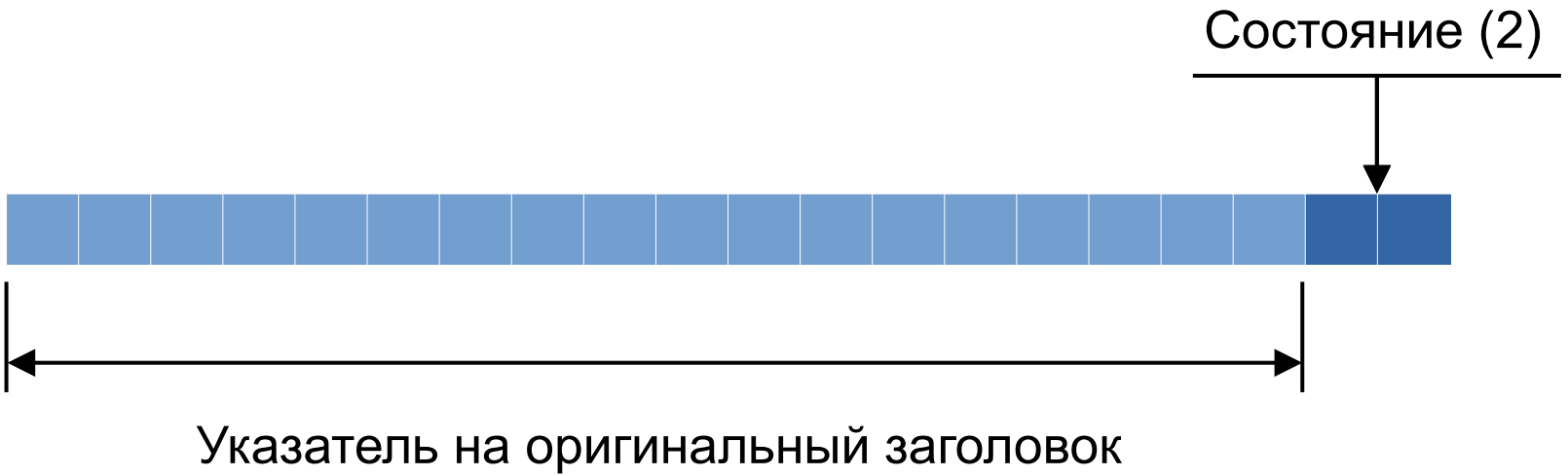
Mark Word — Normal



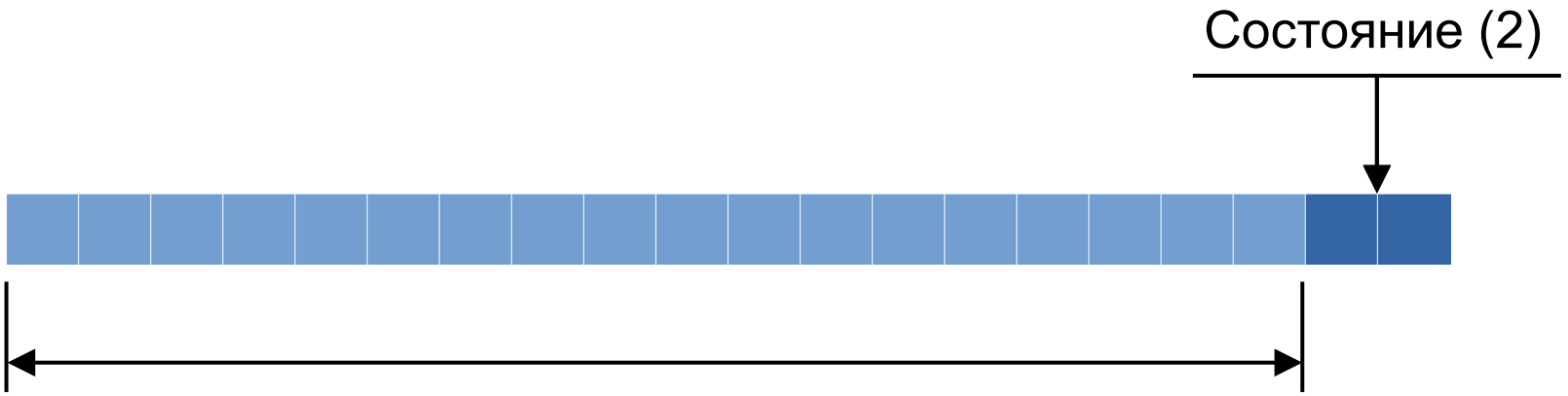
Mark Word — Normal



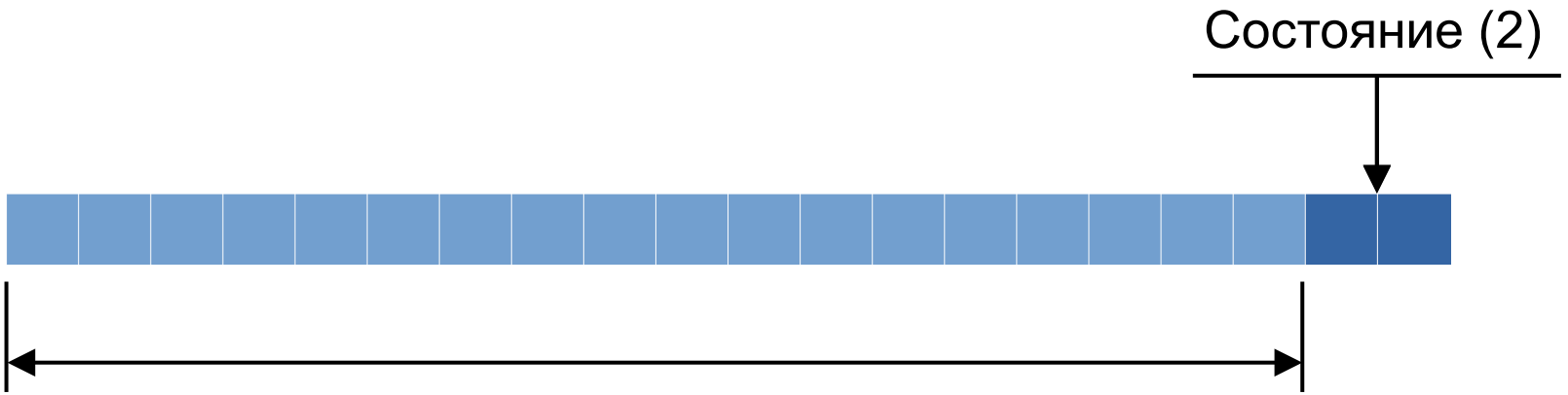
Mark Word — Stack locking



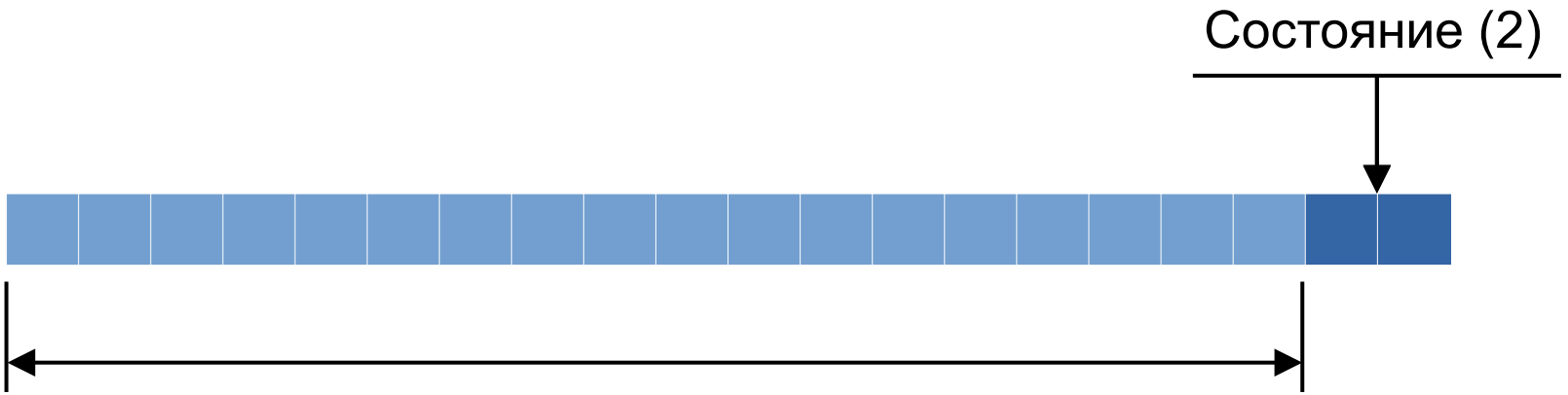
Mark Word — Monitor locking



Mark Word — Marked



Mark Word — Marked



Новый адрес этого объекта
Оригинальный заголовок доступен по ссылке
(или нет)

Сжимаем!

Уменьшаем ClassWord

- 2^{64} = 16 экзабайт, это слишком много

Уменьшаем ClassWord

- 2^{64} = 16 экзабайт, это слишком много
- JDK 6: Сжатие указателей

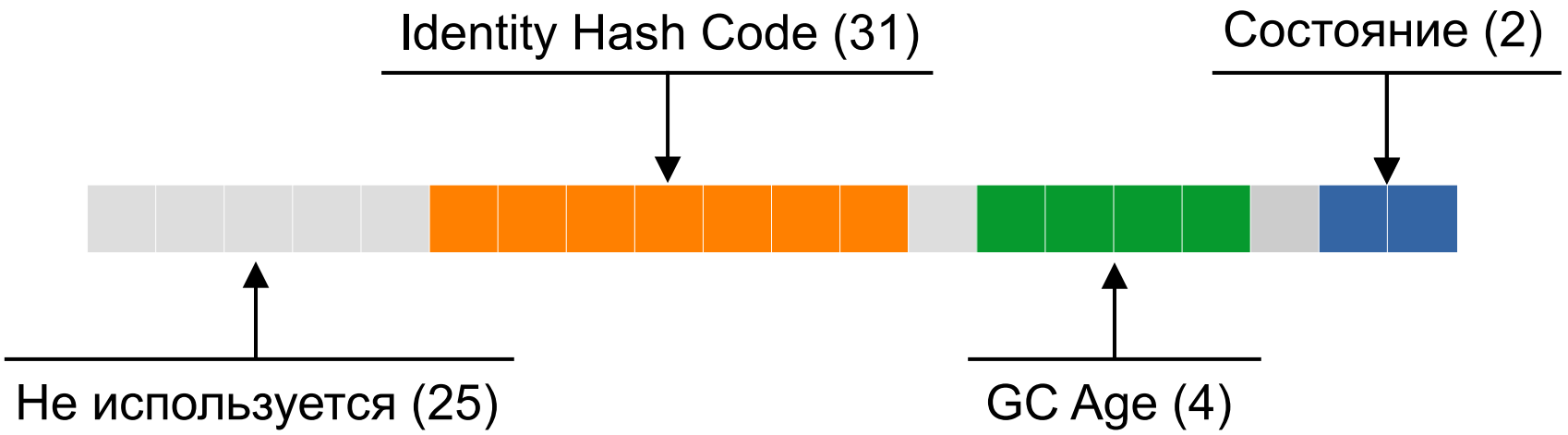
Уменьшаем ClassWord

- 2^{64} = 16 экзабайт, это слишком много
- JDK 6: Сжатие указателей
- JDK 15: Сжатие указателей на класс можно включить независимо

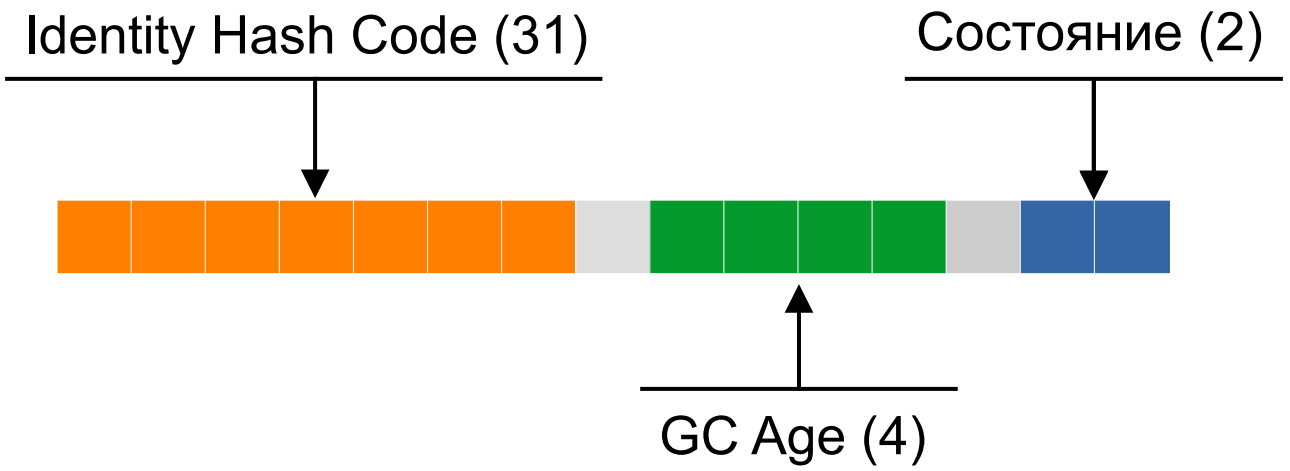
Уменьшаем ClassWord

- 2^{64} = 16 экзабайт, это слишком много
- JDK 6: Сжатие указателей
- JDK 15: Сжатие указателей на класс можно включить независимо
- Lilliput: ClassWord всегда использует сжатие

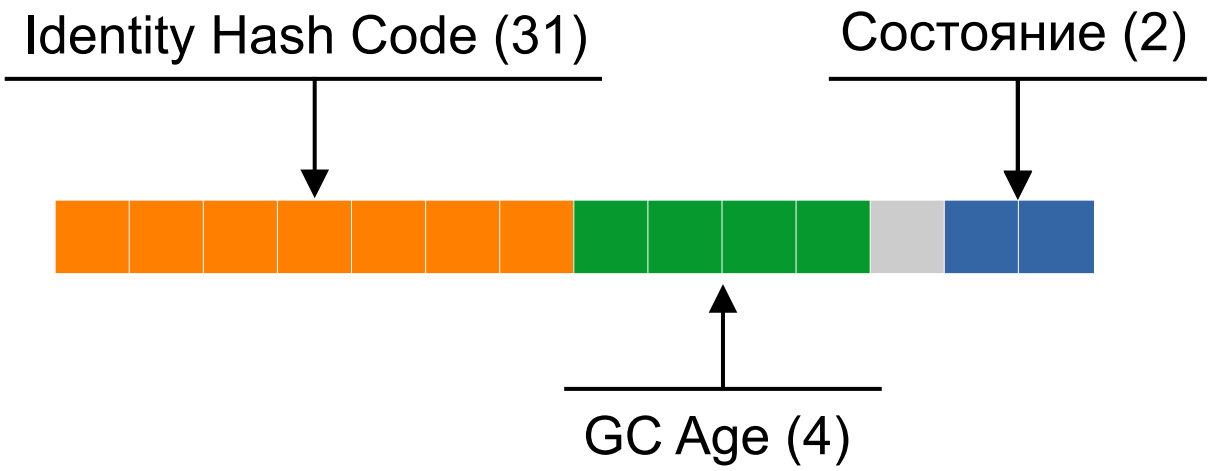
Уменьшение MarkWord — Normal



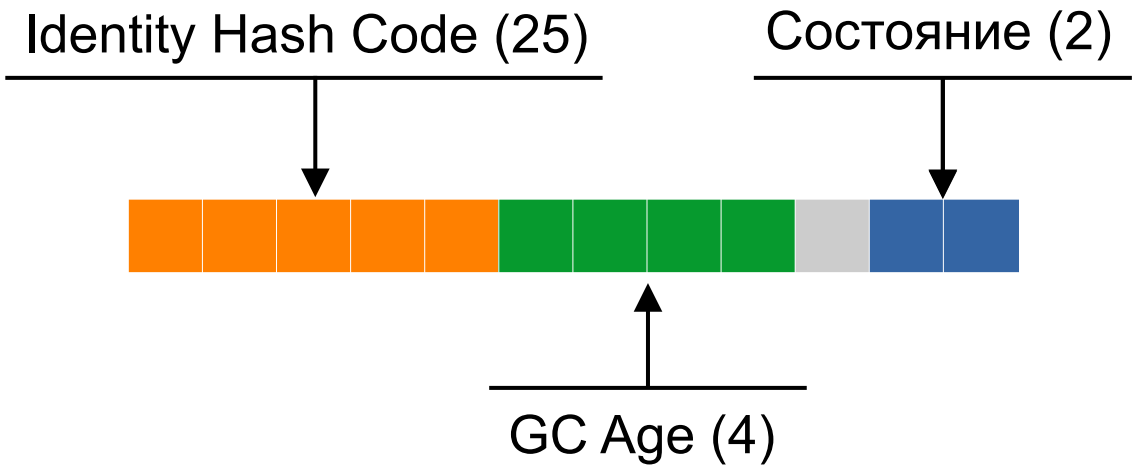
Уменьшение MarkWord — Normal



Уменьшение MarkWord — Normal



Уменьшение MarkWord — Normal



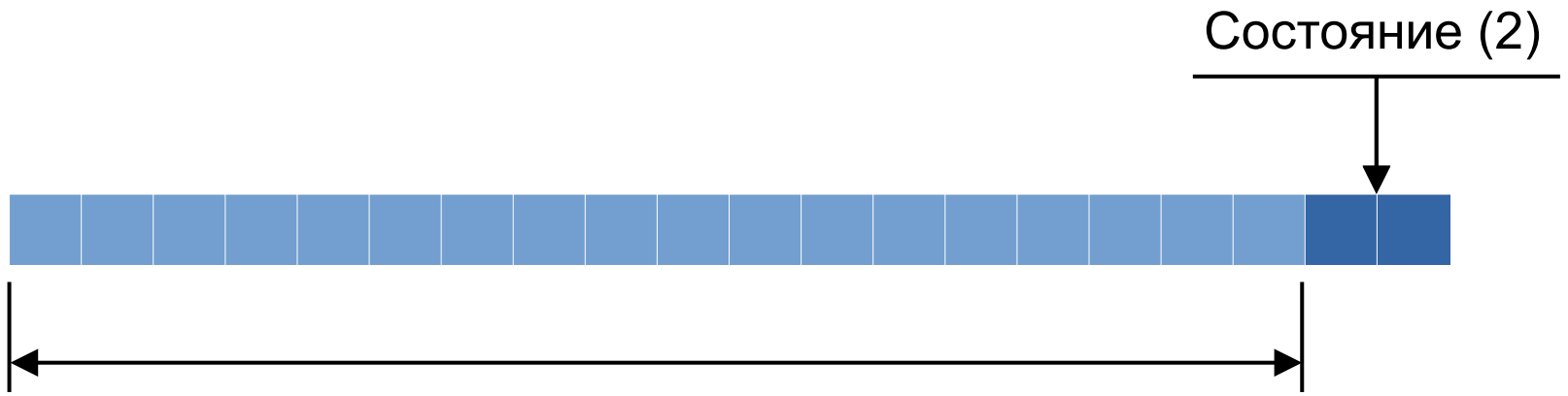
Сложная часть: указатели в MW

- При 64-битном заголовке 64-битный указатель занимает его полностью, и ClassWord и MarkWord
- Заголовок нам важен, особенно — ClassWord

Mark Word — Stack locking

- Stack locking заменён на Alternative fast-locking scheme
- https://bugs.openjdk.org/browse/JDK-8291555
- -XX:+UnlockExperimentalVMOptions -XX:LockingMode=2
- Заголовок не перезаписывается (кроме битов состояния)
- Необходимые данные о lock хранятся в thread-local структуре данных.

Mark Word — Monitor locking



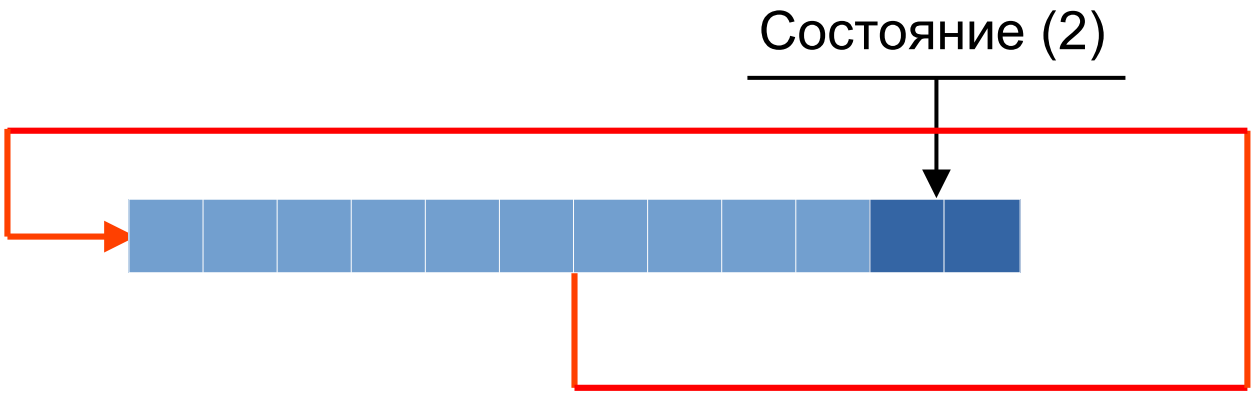
Mark Word — Monitor locking

- В заголовок записывается ссылка на MonitorObject
- Оригинальный заголовок сохраняется внутри MonitorObject
- Всё что нужно — не допускать прямого чтения ClassWord

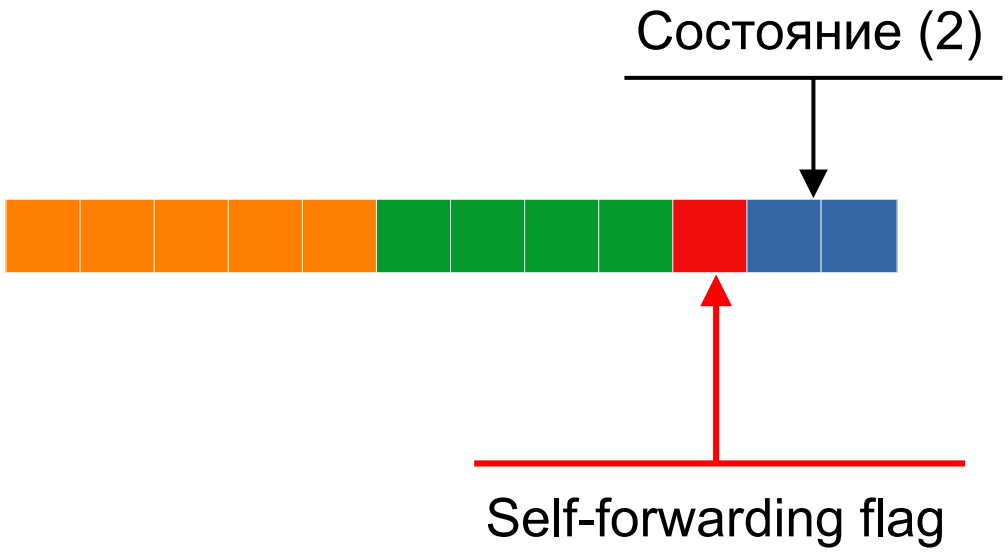
Mark Word — Marked

- В норме GC копирует объект в новый регион и записывает в старый заголовок адрес новой копии
- Никаких специфичных для Lilliput изменений

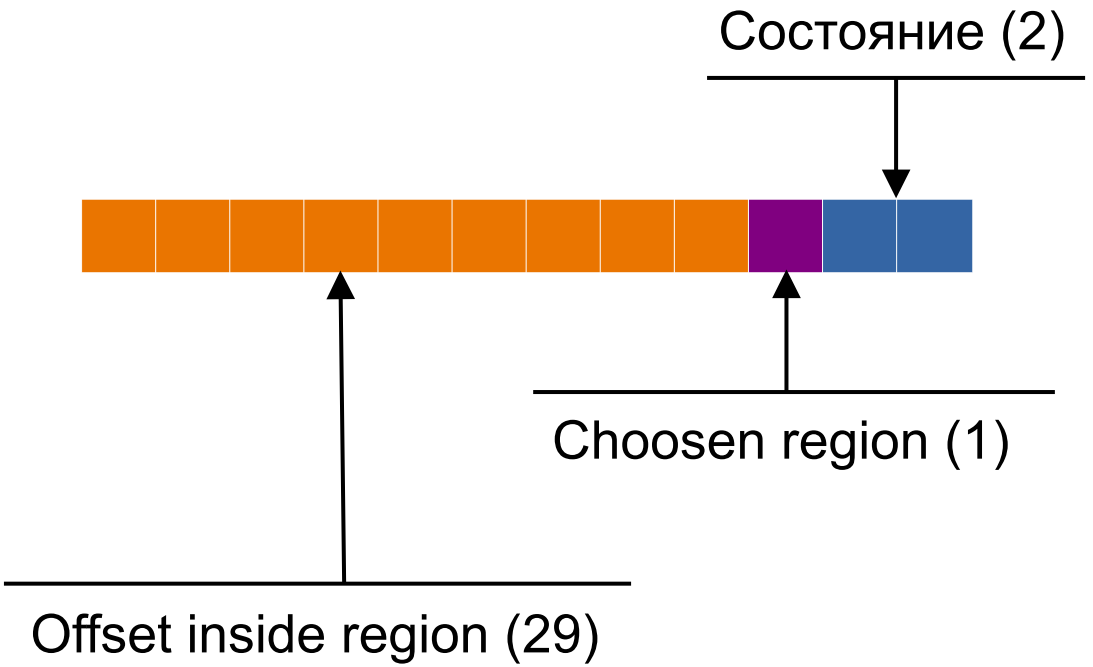
Mark Word — Marked (self-forwarding)



Mark Word — Marked (self-forwarding)



MarkWord — Marked (Sliding GC)



Планы на будущее: 32 бита

- Сократить Identity Hash Code до 2 бит в заголовке
- Сократить указатель на класс до 23 бит
- Хранить forwarding pointer в старом теле объекта

Как попробовать?

Сборки от Amazon Corretto (Windows, Linux, macOS):

https://downloads.corretto.aws/#/downloads?build=nightly&branch=lilliput

Флаги:

-XX:+UnlockExperimentalVMOptions
-XX:+UseCompactObjectHeaders

Полезные ссылки

- JEP 450: https://openjdk.org/jeps/450
- Репозиторий Lilliput на GitHub: https://github.com/openjdk/lilliput
- Список рассылки: https://mail.openjdk.org/mailman/listinfo/lilliput-dev
- Сборки Lilliput от Шипилёва (Linux)
  https://builds.shipilev.net/openjdk-jdk21-lilliput/
- Сборки Lilliput от Amazon Corretto
  https://downloads.corretto.aws/#/downloads?build=nightly&branch=lilliput

Q&A