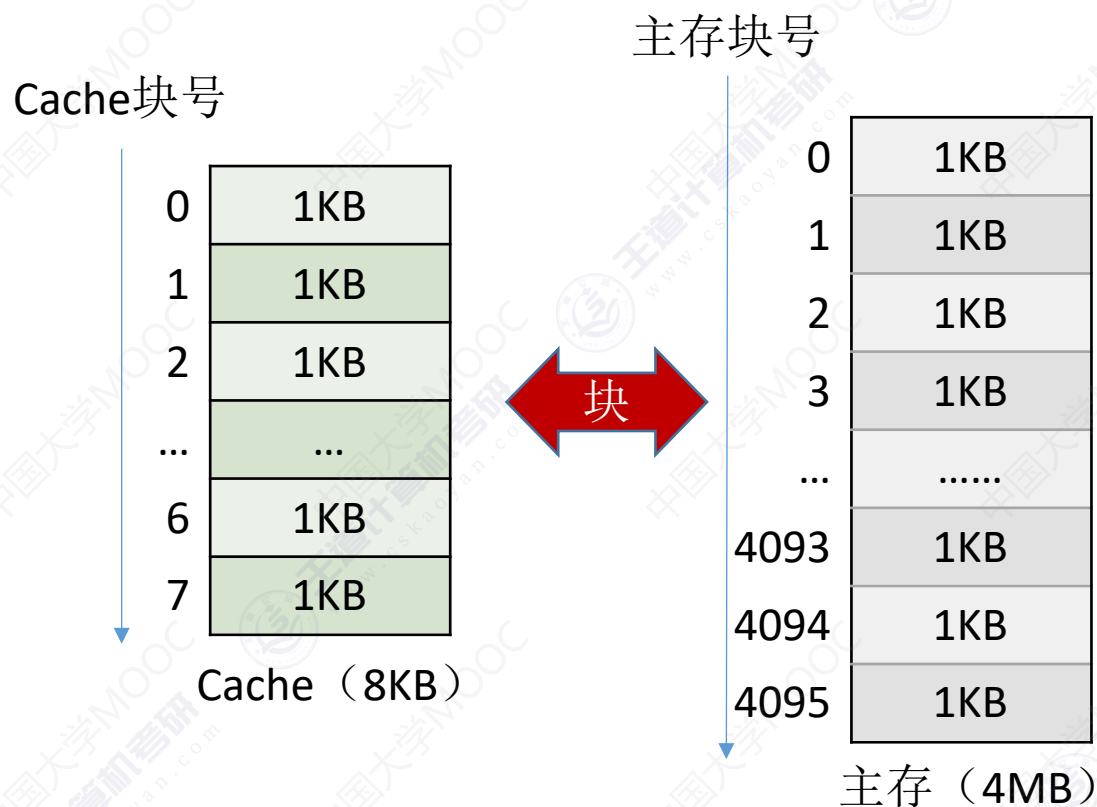


本节内容

Cache-主存 映射方式

有待解决的问题



注意：每次被访问的主存块，一定会被立即调入Cache

主存的地址共22位：

块号	块内地址
12位	10位

$$4M=2^{22}, 1K=2^{10}$$

整个主存被分为 $2^{12} = 4096$ 块

- 如何区分 Cache 与 主存 的数据块对应关系？——Cache和主存的映射方式
- Cache 很小，主存很大。如果Cache满了怎么办？——替换算法
- CPU修改了Cache中的数据副本，如何确保主存中数据母本的一致性？——Cache写策略

有“标记”就够了？

还要增加“有效位”

二进制表示，
初始都为0

有效位

标记

1	9
1	8
1	5
0	0
0	0
0	0
0	0
1	0



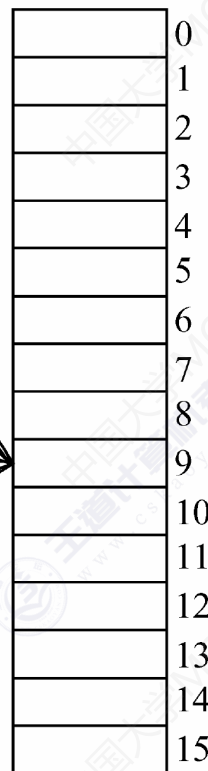
本节总览



如何区分Cache中存放的是哪个主存块？

给每个Cache块增加一个“标记”，记录对应的主存块号？

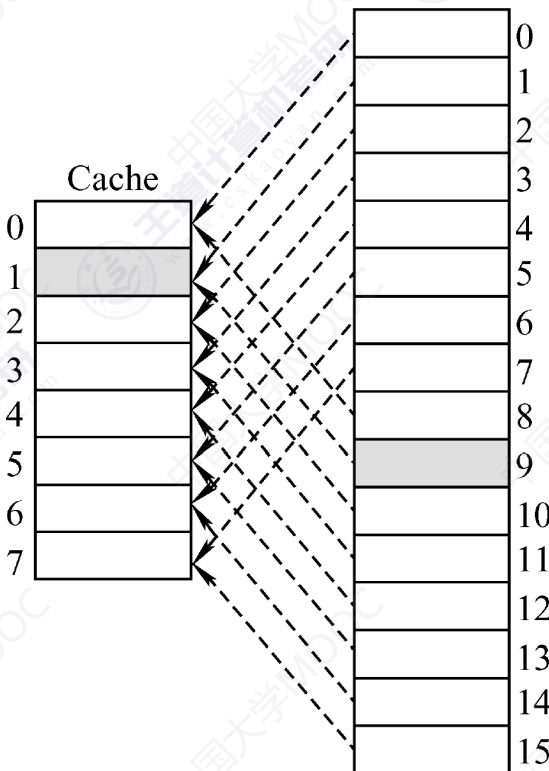
主存



(a) 全相联映射

主存块可以放在
Cache的任意位置

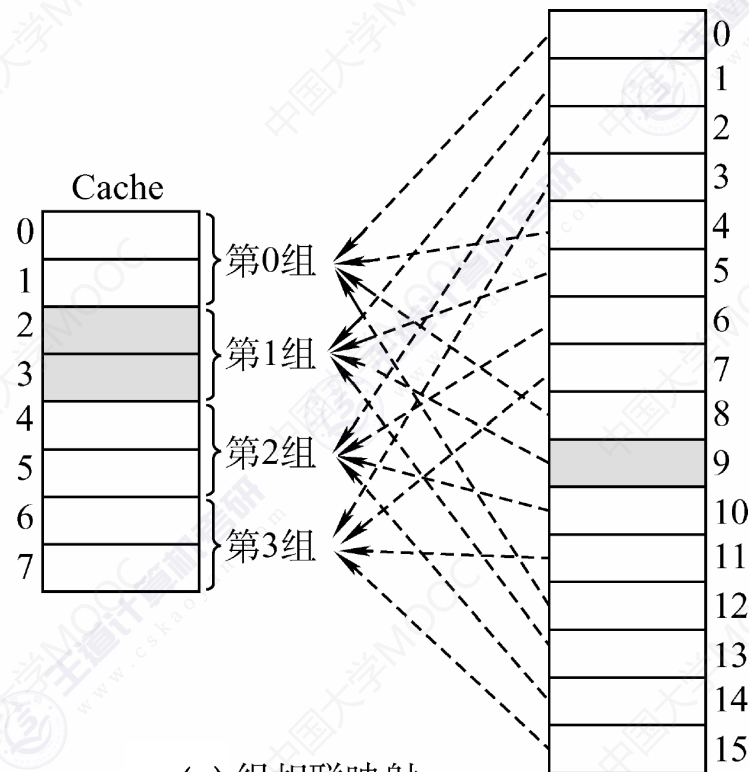
主存



(b) 直接映射

每个主存块只能放到一个特定的位置：
 $\text{Cache块号} = \text{主存块号} \% \text{Cache总块数}$

主存



(c) 组相联映射

Cache块分为若干组，每个主存块可放到特定分组中的任意一个位置
 $\text{组号} = \text{主存块号} \% \text{分组数}$

全相联映射（随意放）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

即Cache块，与主存块的大小相等

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位

Cache	
0	
1	
2	
3	
4	
5	
6	
7	

Cache: $8 \times 64\text{B} = 512\text{B}$

块号	主存	每个主存块的地址范围
0		$0 \dots 000000000000 \sim 0 \dots 0000111111$
1		$0 \dots 0001000000 \sim 0 \dots 0001111111$
2		$0 \dots 0010000000 \sim 0 \dots 0010111111$
	⋮	
$2^{22}-3$		$1 \dots 1101000000 \sim 1 \dots 1101111111$
$2^{22}-2$		$1 \dots 1110000000 \sim 1 \dots 1110111111$
$2^{22}-1$		$1 \dots 1111000000 \sim 1 \dots 1111111111$

主存: 256MB

全相联映射（随意放）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

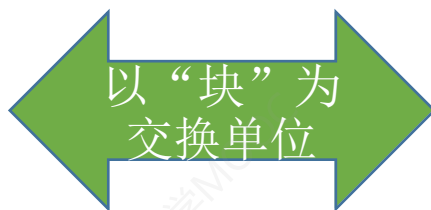
即Cache块，与主存块的大小相等

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位

Cache	
0	64B
1	64B
2	64B
3	64B
4	64B
5	64B
6	64B
7	64B

Cache: $8 \times 64\text{B} = 512\text{B}$



块号	主存	每个主存块的地址范围
0	64B	$0 \dots 000000000000 \sim 0 \dots 0000111111$
1	64B	$0 \dots 0001000000 \sim 0 \dots 0001111111$
2	64B	$0 \dots 0010000000 \sim 0 \dots 0010111111$
⋮		
$2^{22}-3$	64B	$1 \dots 1101000000 \sim 1 \dots 1101111111$
$2^{22}-2$	64B	$1 \dots 1110000000 \sim 1 \dots 1110111111$
$2^{22}-1$	64B	$1 \dots 1111000000 \sim 1 \dots 1111111111$

主存: 256MB

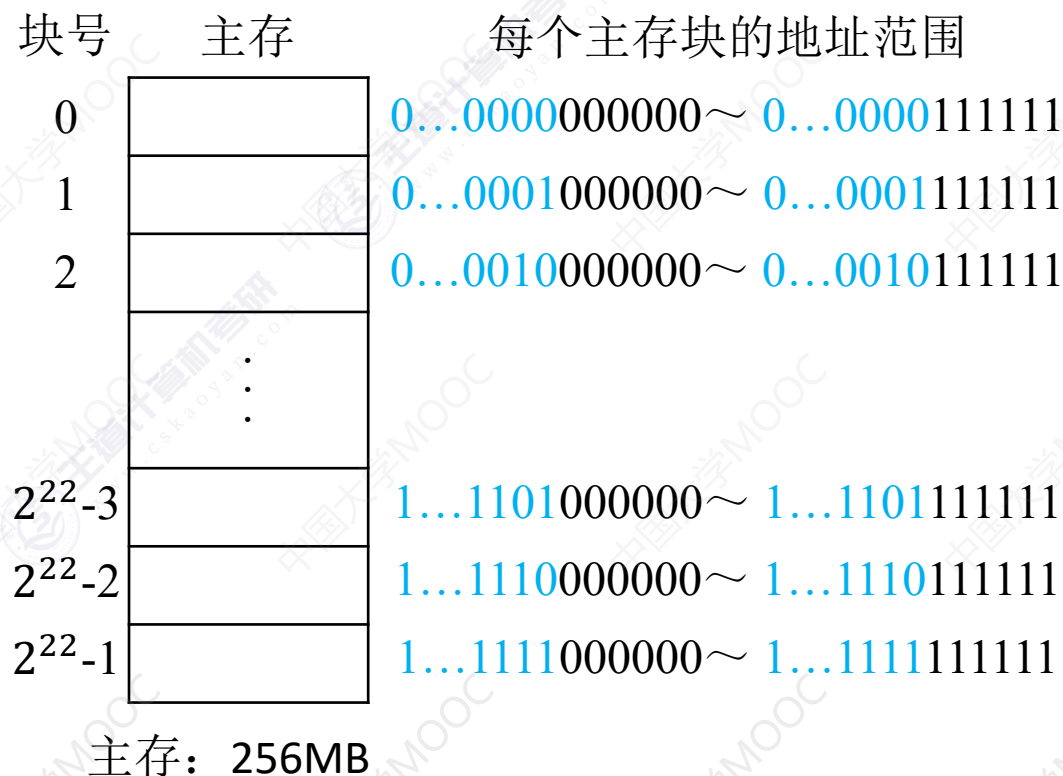
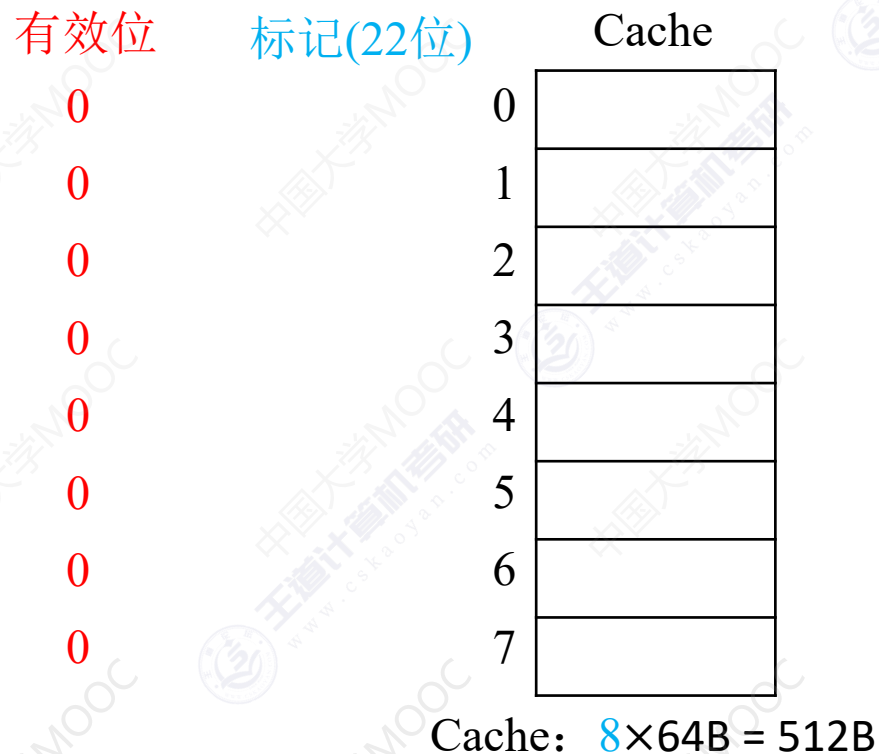
全相联映射（随意放）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

即Cache块，与主存块的大小相等

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位



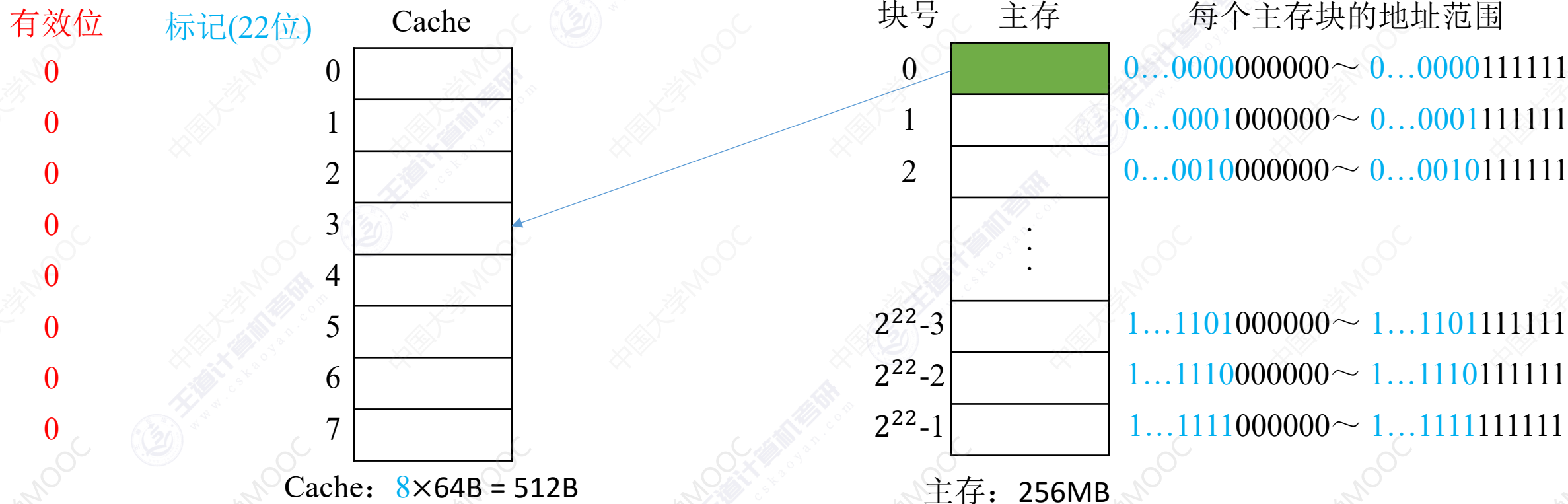
全相联映射（随意放）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

即Cache块，与主存块的大小相等

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位



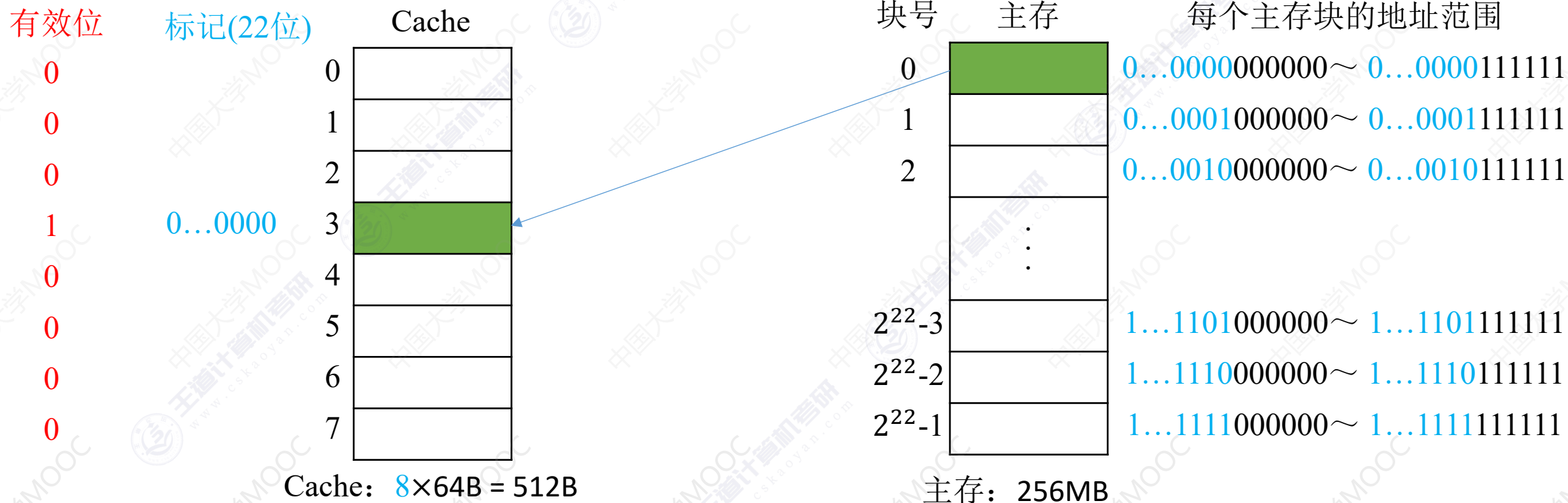
全相联映射（随意放）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

即Cache块，与主存块的大小相等

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位



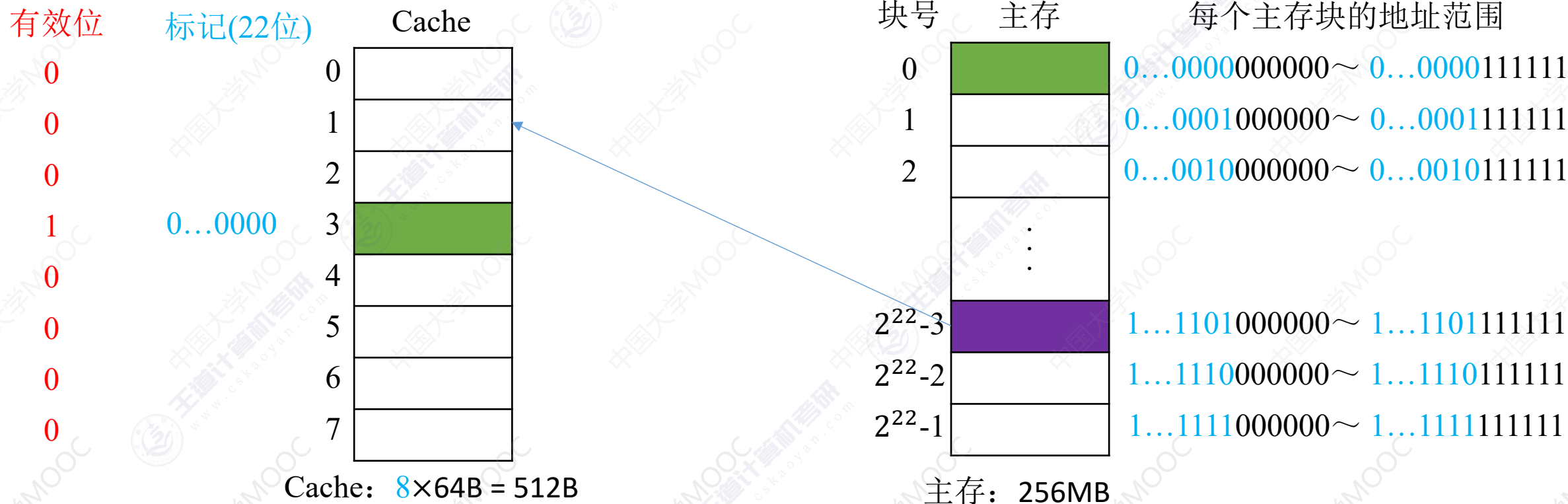
全相联映射（随意放）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

即Cache块，与主存块的大小相等

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位



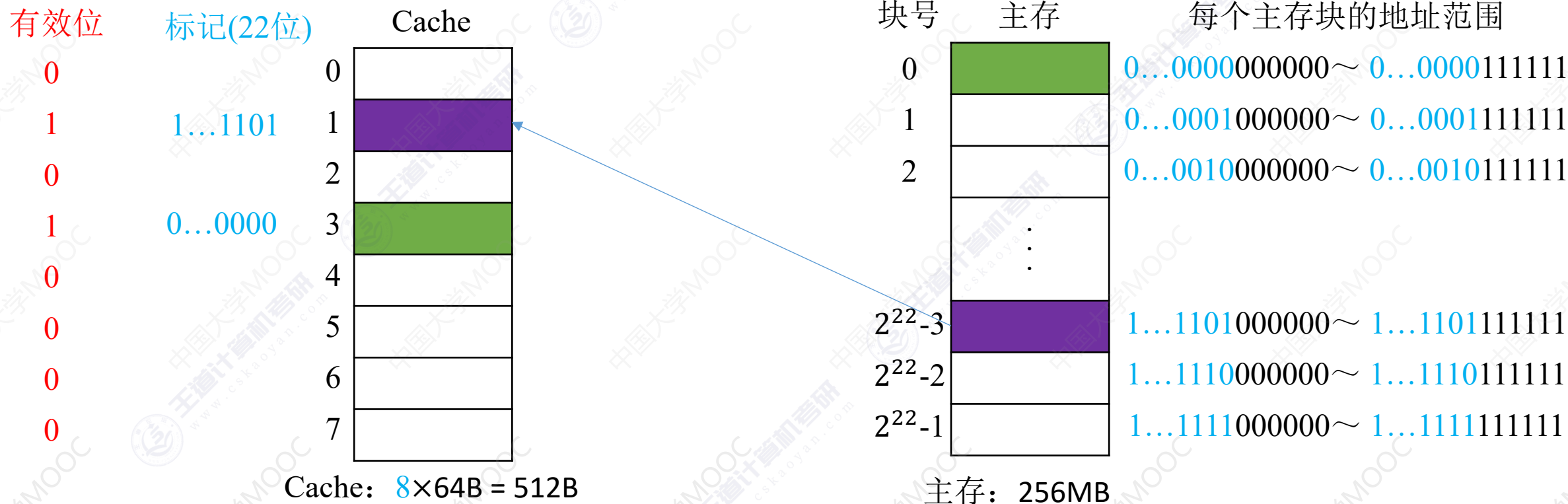
全相联映射（随意放）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

即Cache块，与主存块的大小相等

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位



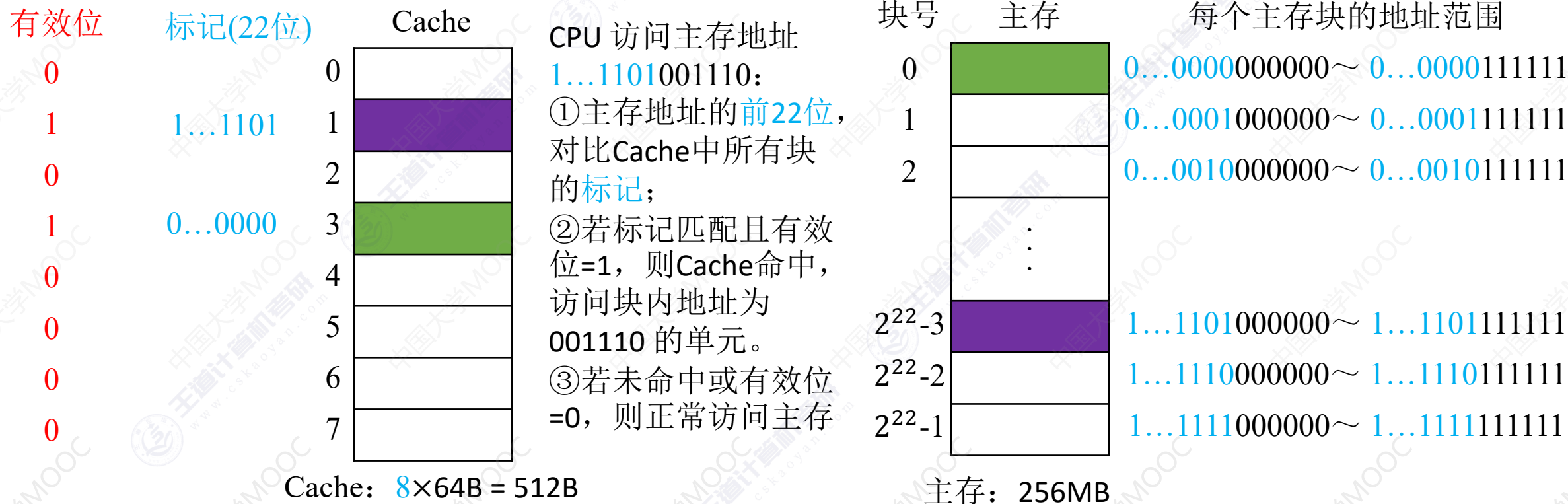
“全相联映射”如何访存？

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

即Cache块，与主存块的大小相等

256M=2²⁸ 主存的地址共28位：

主存块号	块内地址
22位	6位



直接映射（只能放固定位置）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

直接映射，主存块在Cache中的位置=主存块号%Cache总块数

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位

Cache	
0	
1	
2	
3	
4	
5	
6	
7	

Cache: $8 \times 64\text{B} = 512\text{B}$

块号	主存	每个主存块的地址范围
0		$0 \dots 000000000000 \sim 0 \dots 0000111111$
1		$0 \dots 0001000000 \sim 0 \dots 0001111111$
2		$0 \dots 0010000000 \sim 0 \dots 0010111111$
	⋮	
$2^{22}-3$		$1 \dots 1101000000 \sim 1 \dots 1101111111$
$2^{22}-2$		$1 \dots 1110000000 \sim 1 \dots 1110111111$
$2^{22}-1$		$1 \dots 1111000000 \sim 1 \dots 1111111111$

主存: 256MB

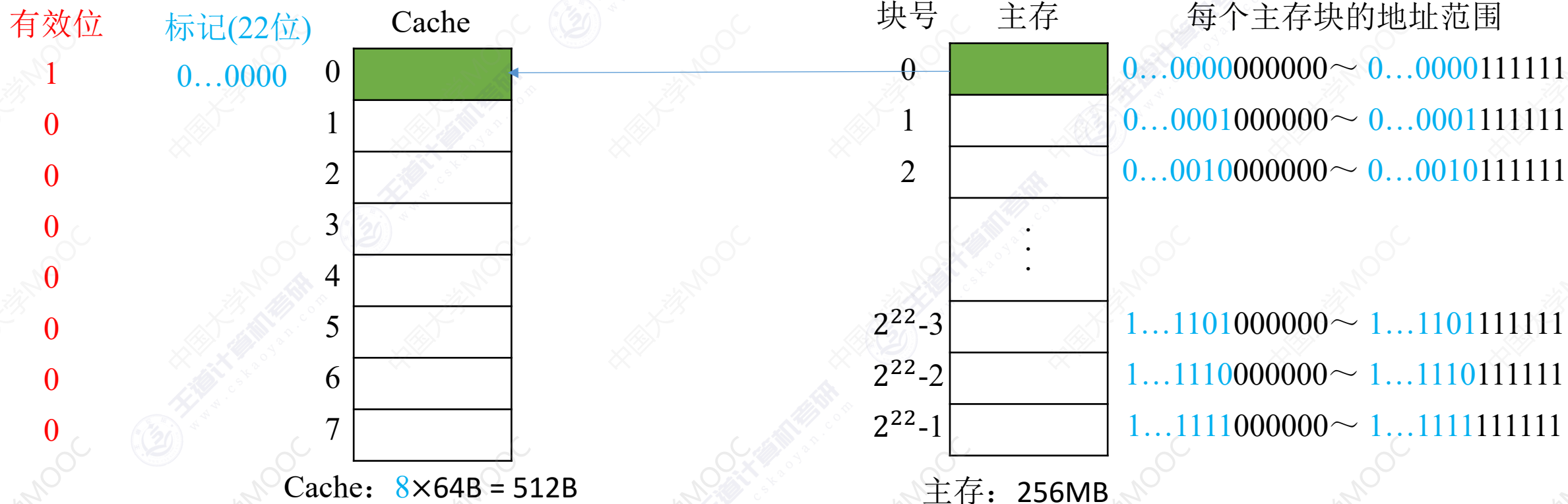
直接映射（只能放固定位置）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

直接映射，主存块在Cache中的位置=主存块号%Cache总块数

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位



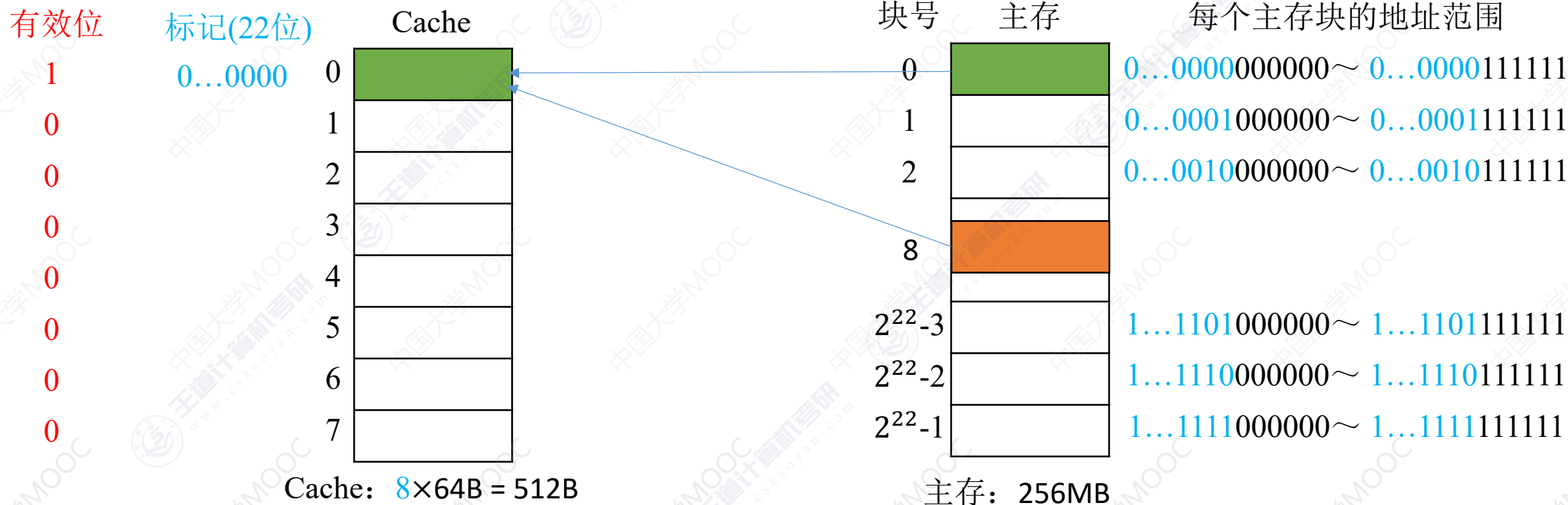
直接映射（只能放固定位置）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

直接映射，主存块在Cache中的位置=主存块号%Cache总块数

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位



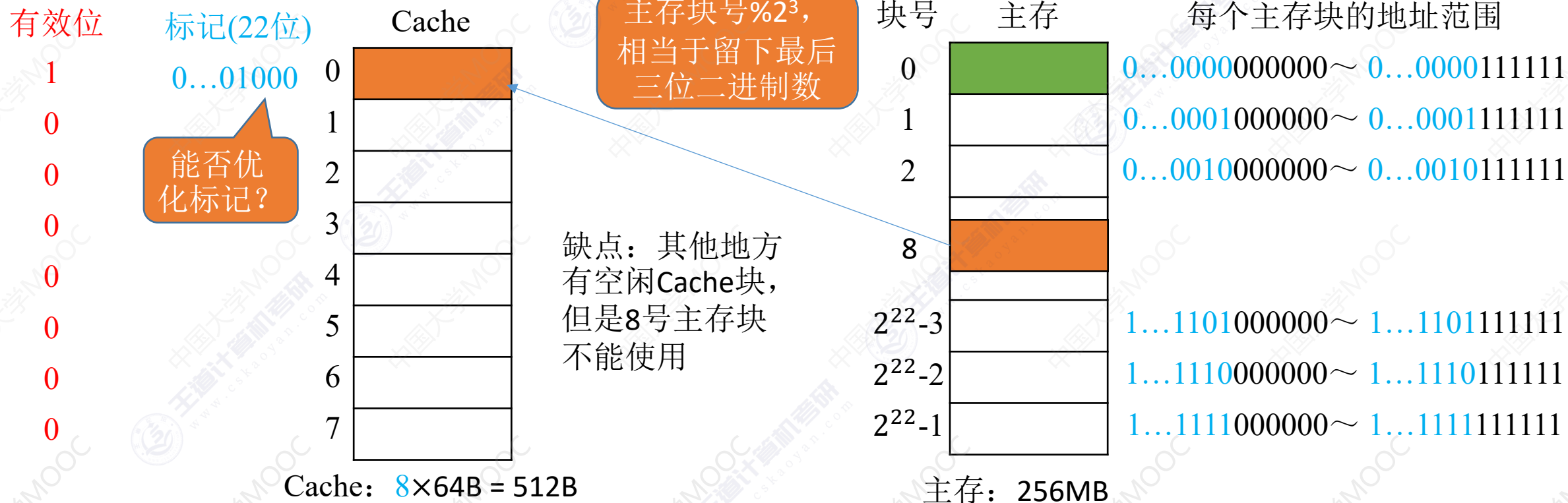
直接映射（只能放固定位置）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

直接映射，主存块在Cache中的位置=主存块号%Cache总块数

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位



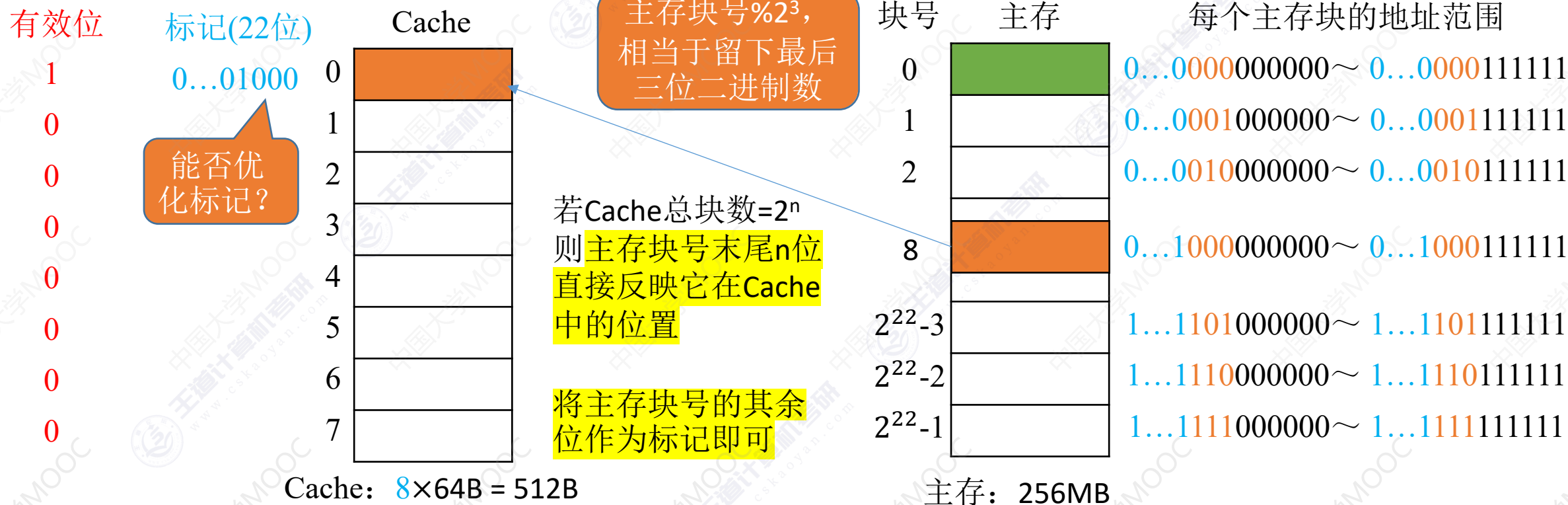
直接映射（只能放固定位置）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

直接映射，主存块在Cache中的位置=主存块号%Cache总块数

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位



直接映射，主存块在Cache中的位置=主存块号%Cache总块数

主存块号	块内地址
22位	6位



“直接映射”如何访存

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

直接映射，主存块在Cache中的位置=主存块号%Cache总块数

256M=2²⁸ 主存的地址共28位：

主存块号		块内地址
22位		6位
19位 标记	3位 行号	6位块内 地址

Cache 共2³ 行

有效位	标记(19位)	Cache
1	0...01	0 000
0		1 001
0		2 010
0		3 011
0		4 100
0		5 101
0		6 110
0		7 111

Cache: 8×64B = 512B

CPU 访问主存地址

0...01000 001110 :

①根据主存块号的后3位确定Cache行

②若主存块号的前19位与Cache标记匹配

且有效位=1，则Cache命中，访问块内地址为001110的单元。

③若未命中或有效位=0，则正常访问主存

块号 主存

0	
1	
2	
8	
2 ²² -3	
2 ²² -2	
2 ²² -1	

主存: 256MB

每个主存块的地址范围

0...000000000000 ~ 0...0000111111
0...000100000000 ~ 0...0001111111
0...001000000000 ~ 0...0010111111
0...100000000000 ~ 0...1000111111
1...110100000000 ~ 1...1101111111
1...111000000000 ~ 1...1110111111
1...111100000000 ~ 1...1111111111

组相联映射（可放到特定分组）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

组相联映射，所属分组=主存块号%分组数

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位

Cache	
0	
1	
2	
3	
4	
5	
6	
7	

Cache: $8 \times 64\text{B} = 512\text{B}$

块号	主存	每个主存块的地址范围
0		$0 \dots 000000000000 \sim 0 \dots 0000111111$
1		$0 \dots 0001000000 \sim 0 \dots 0001111111$
2		$0 \dots 0010000000 \sim 0 \dots 0010111111$
	⋮	
$2^{22}-3$		$1 \dots 1101000000 \sim 1 \dots 1101111111$
$2^{22}-2$		$1 \dots 1110000000 \sim 1 \dots 1110111111$
$2^{22}-1$		$1 \dots 1111000000 \sim 1 \dots 1111111111$

主存: 256MB

2路组相联映射——2块为一组，分四组

组相联映射（可放到特定分组）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

组相联映射，所属分组=主存块号%分组数

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位



Cache: $8 \times 64\text{B} = 512\text{B}$

2路组相联映射——2块为一组，分四组

块号	主存	每个主存块的地址范围
0		$0 \dots 000000000000 \sim 0 \dots 0000111111$
1		$0 \dots 0001000000 \sim 0 \dots 0001111111$
2		$0 \dots 0010000000 \sim 0 \dots 0010111111$
	⋮	
$2^{22}-3$		$1 \dots 1101000000 \sim 1 \dots 1101111111$
$2^{22}-2$		$1 \dots 1110000000 \sim 1 \dots 1110111111$
$2^{22}-1$		$1 \dots 1111000000 \sim 1 \dots 1111111111$

主存: 256MB

组相联映射（可放到特定分组）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

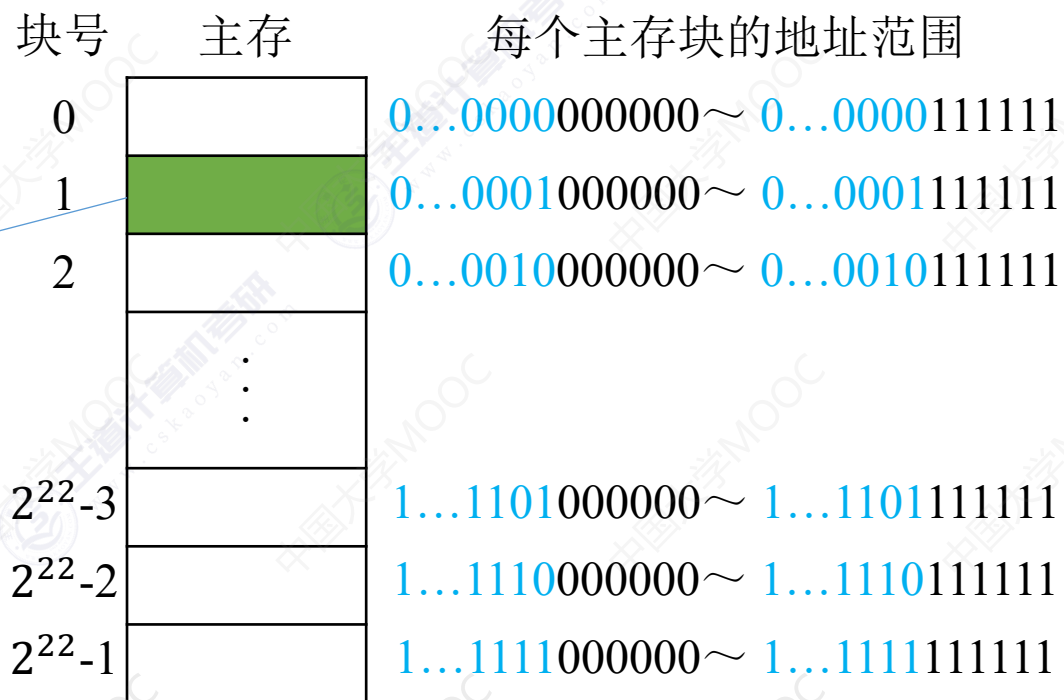
组相联映射，所属分组=主存块号%分组数

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位



Cache: $8 \times 64\text{B} = 512\text{B}$



主存: 256MB

2路组相联映射——2块为一组，分四组

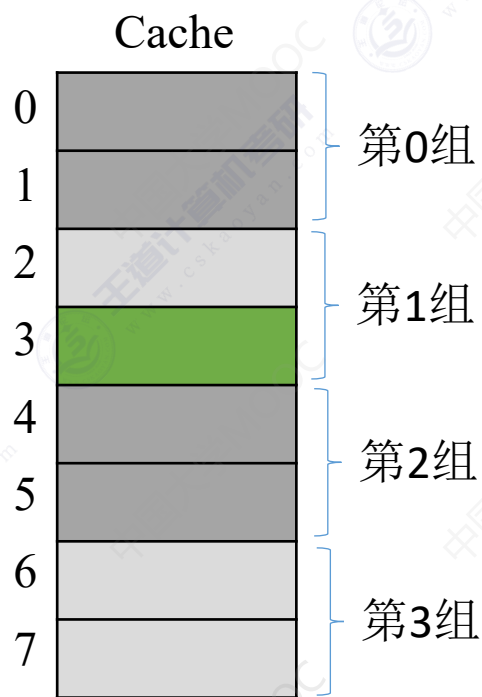
组相联映射（可放到特定分组）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

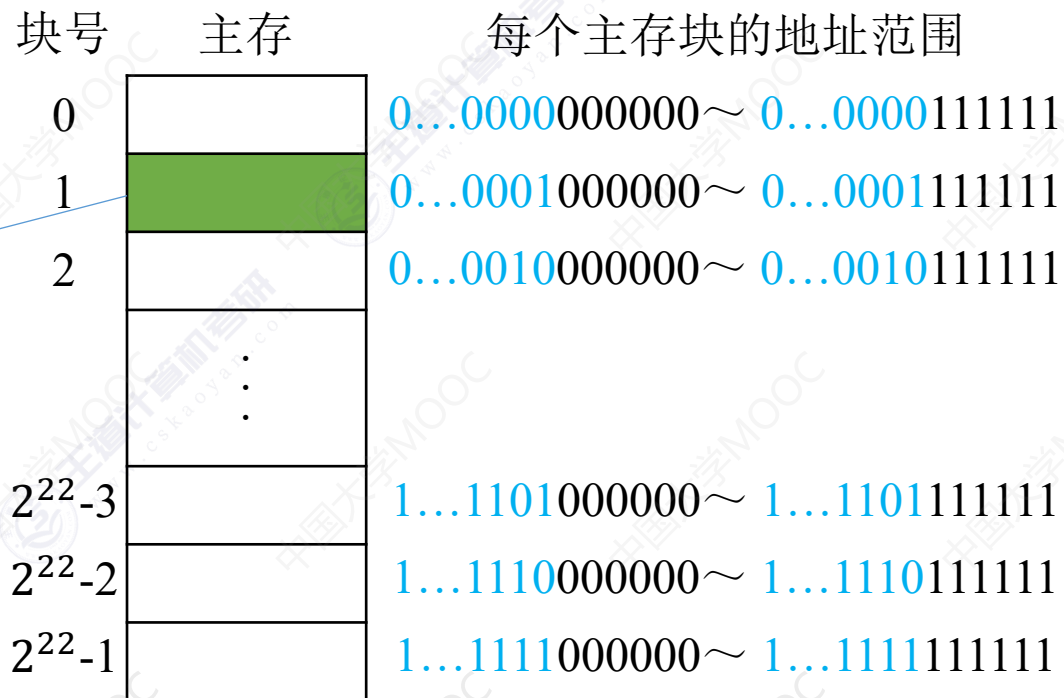
组相联映射，所属分组=主存块号%分组数

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位



Cache: $8 \times 64\text{B} = 512\text{B}$



主存: 256MB

2路组相联映射——2块为一组，分四组

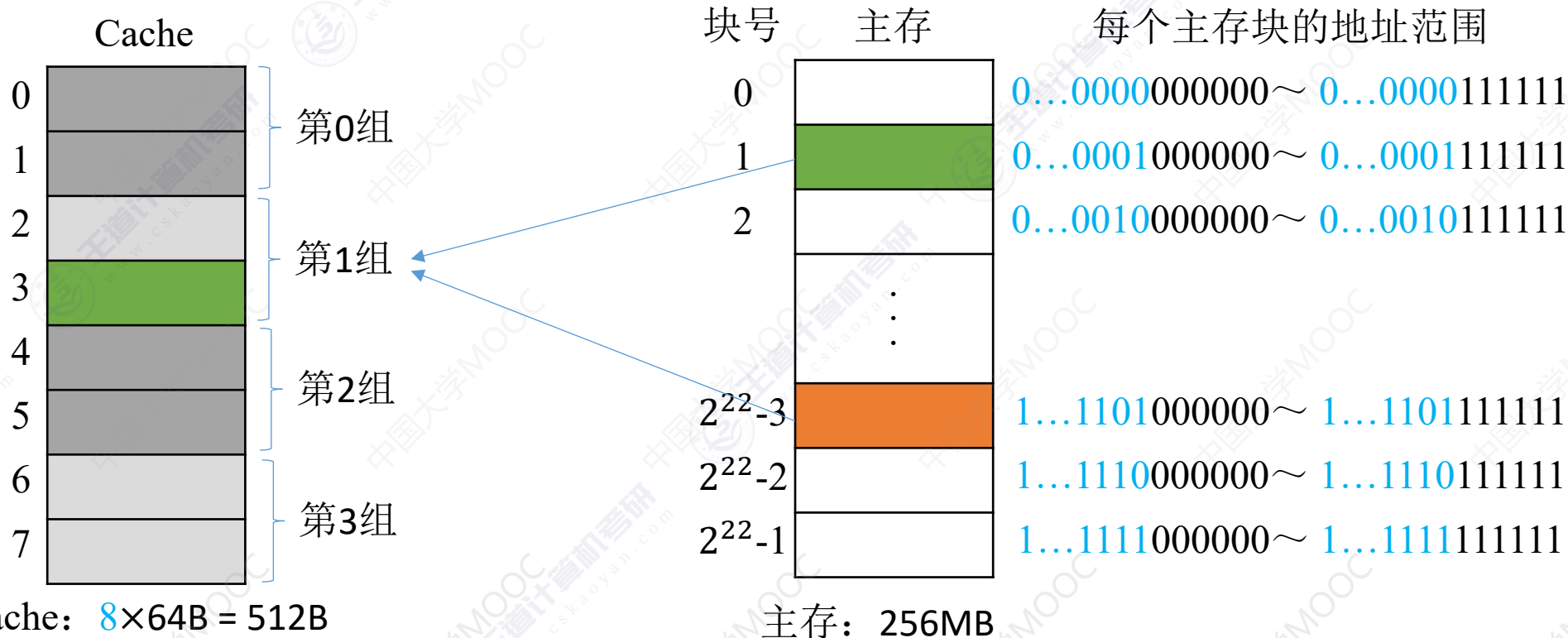
组相联映射（可放到特定分组）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

组相联映射，**所属分组=主存块号%分组数**

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位



2路组相联映射——2块为一组，分四组

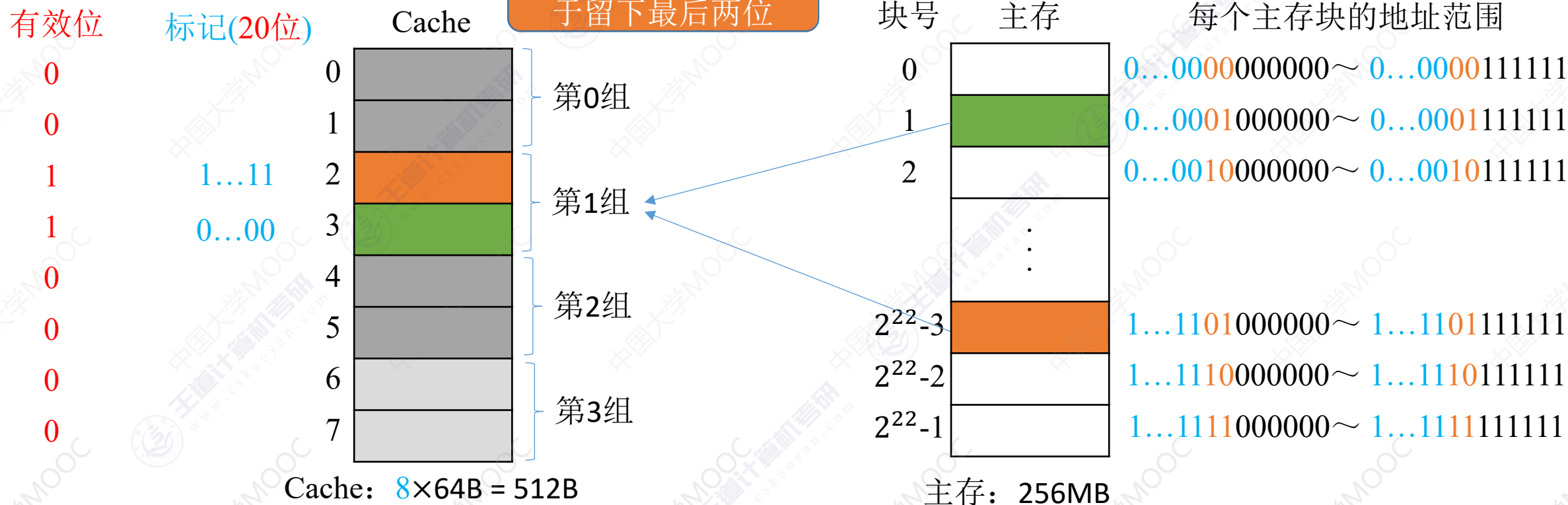
组相联映射（可放到特定分组）

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

组相联映射，所属分组=主存块号%分组数

$256\text{M}=2^{28}$ 主存的地址共28位：

主存块号	块内地址
22位	6位



2路组相联映射——2块为一组，分四组

组相联映射（可放到特定分组）

256M=2²⁸ 主存的地址共28位：

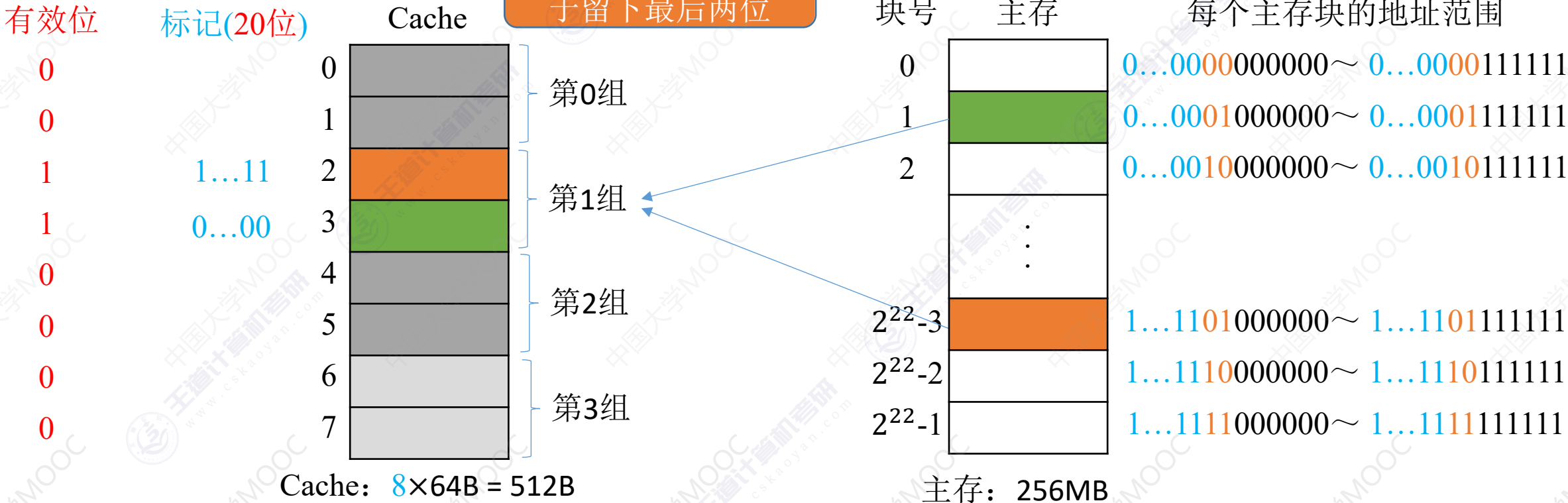
主存块号		块内地址
22位		6位
20位 标记	2位 组号	6位块内 地址

Cache 分为2²组

假设某个计算机的主存地址空间大小为256MB，按字节编址，其数据Cache有8个Cache行，行长为64B。

组相联映射，所属分组=主存块号%分组数

主存块号%2²，相当于留下最后两位



2路组相联映射——2块为一组，分四组

组相联映射, 所属分组=主存块号%分组数

主存块号		块内地址
22位		6位
20位 标记	2位 组号	6位块内 地址

Cache 分为 2^2 组

有效位	标记(20位)		Cache
0		0	00
0		1	00
1	1...11	2	01
1	0...00	3	01
0		4	10
0		5	10
0		6	11
0		7	11

2路组相联映射——2块为一组，分四组

1...1101001110 :

②若主存块号的前20位与分组内的某个标记匹配且有效位=1，

③若未命中或有效位=0，则正常访问主存

块号	主存	每个主存块的地址范围
0		0...0000000000000000 ~ 0...000011111111
1		0...0001000000000000 ~ 0...000111111111
2		0...0010000000000000 ~ 0...001011111111
	⋮	
$2^{22}-3$		1...1101000000000000 ~ 1...110111111111
$2^{22}-2$		1...1110000000000000 ~ 1...111011111111
$2^{22}-1$		1...1111000000000000 ~ 1...111111111111

王道考研/CSKAOYAN.COM

知识回顾

Cache-主存映射方式

Cache 中存储的信息

有效位 (0/1)+ 标记 + 整块数据

其中“标记”用于指明对应的内存块，不同映射方式，“标记”的位数不同

全相联映射

主存块可以放到 Cache 的任意位置

主存地址结构

标记 (整个主存块号)+ 块内地址

优点：Cache 存储空间利用充分，命中率高；缺点：查找“标记”最慢，有可能需要对比所有行的标记

直接映射

主存块只能放到特定的某个 Cache 行，行号=主存块号 % 总行数

主存地址结构

标记 (主存块号前几位)+ 行号 (主存块号末几位) 块内地址

优点：对于任意一个地址，只需对比一个“标记”，速度最快；缺点：Cache 存储空间利用不充分，命中率低

组相联映射

主存块可以放到特定分组中的任意位置，所属组号=主存块号 % 总组数

主存地址结构

标记 (主存块号前几位)+ 组号 (主存块号末几位) 块内地址

优点：另外两种方式的折中，综合效果较好

术语：n 路组相联映射——每 n 个 Cache 行为一组

结合每种地址映射方式的地址结构思考：给定一个主存地址，如何拆分地址，并查找Cache、访存？



公众号：王道在线



b站：王道计算机教育



抖音：王道计算机考研