本节内容

外部排序

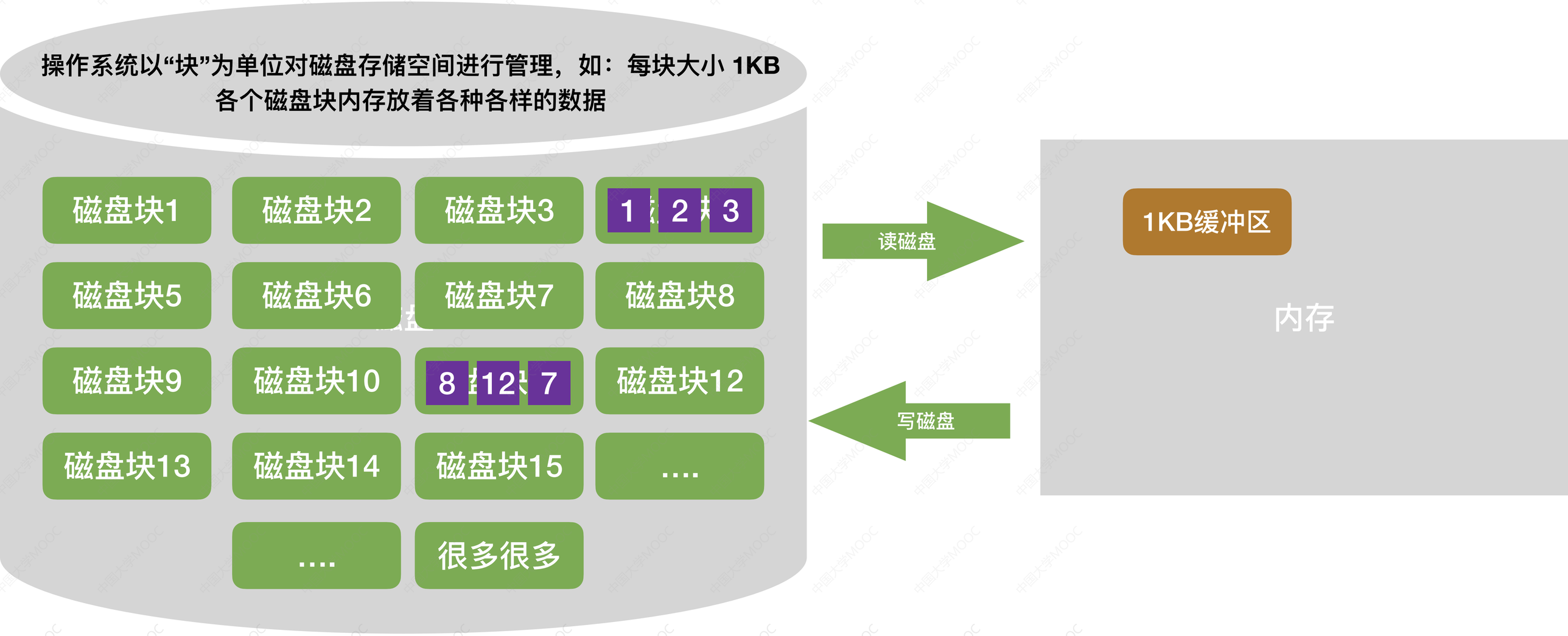
王道考研/CSKAOYAN.COM

知识总览

- 外部排序
  - 外存与内存之间的数据交换
  - 外部排序的原理
  - 影响外部排序效率的因素
  - 优化思路

王道考研/CSKAOYAN.COM

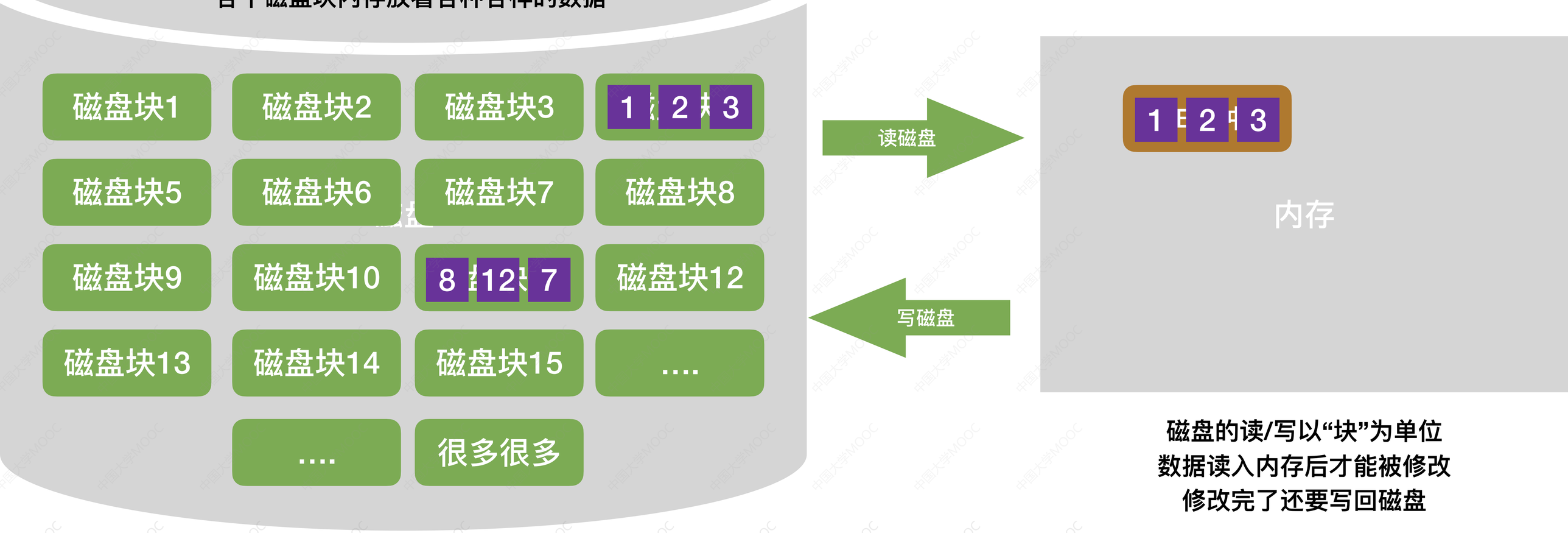
外存、内存之间的数据交换



王道考研/CSKAOYAN.COM

外存、内存之间的数据交换

操作系统以“块”为单位对磁盘存储空间进行管理，如：每块大小 1KB
各个磁盘块内存放着各种各样的数据

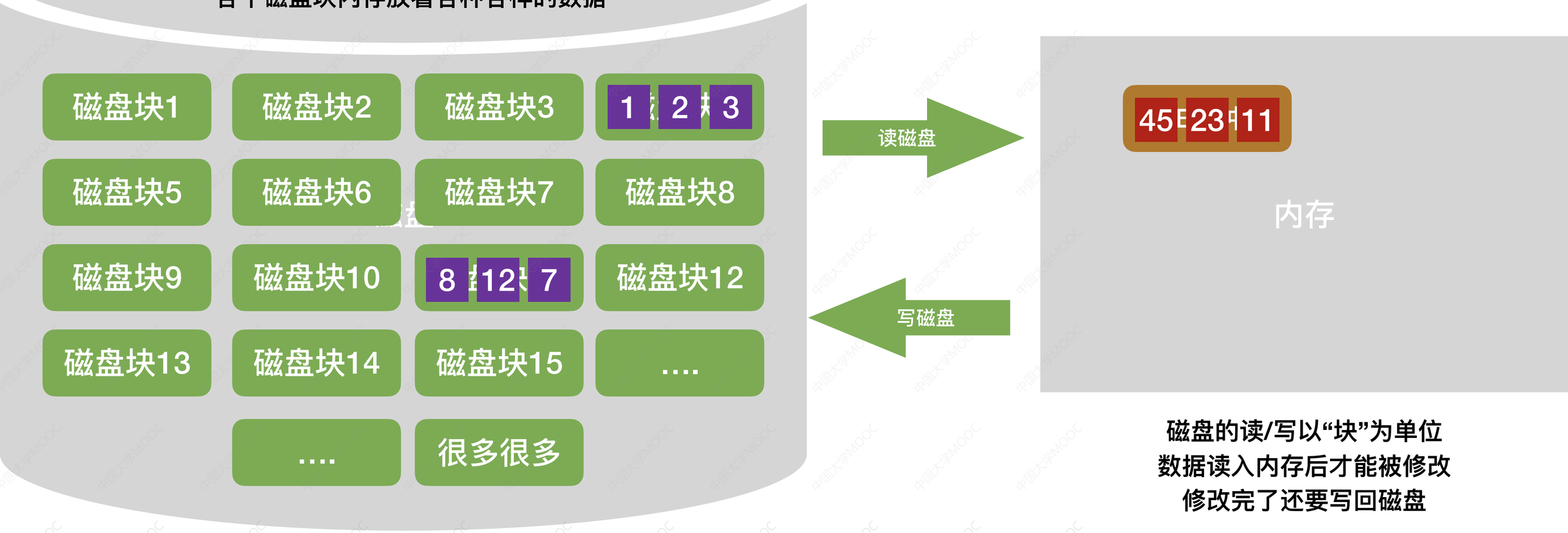


磁盘的读/写以“块”为单位
数据读入内存后才能被修改
修改完了还要写回磁盘

王道考研/CSKAOYAN.COM

外存、内存之间的数据交换

操作系统以“块”为单位对磁盘存储空间进行管理，如：每块大小 1KB
各个磁盘块内存放着各种各样的数据

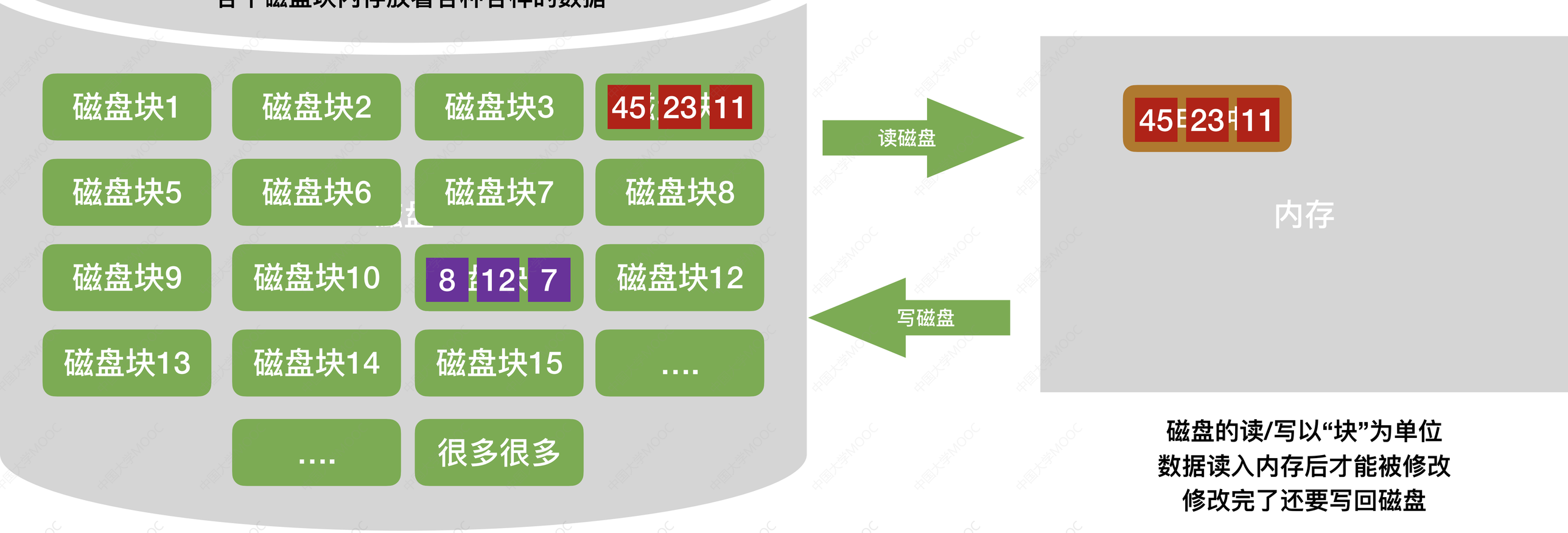


磁盘的读/写以“块”为单位
数据读入内存后才能被修改
修改完了还要写回磁盘

王道考研/CSKAOYAN.COM

外存、内存之间的数据交换

操作系统以“块”为单位对磁盘存储空间进行管理，如：每块大小 1KB
各个磁盘块内存放着各种各样的数据

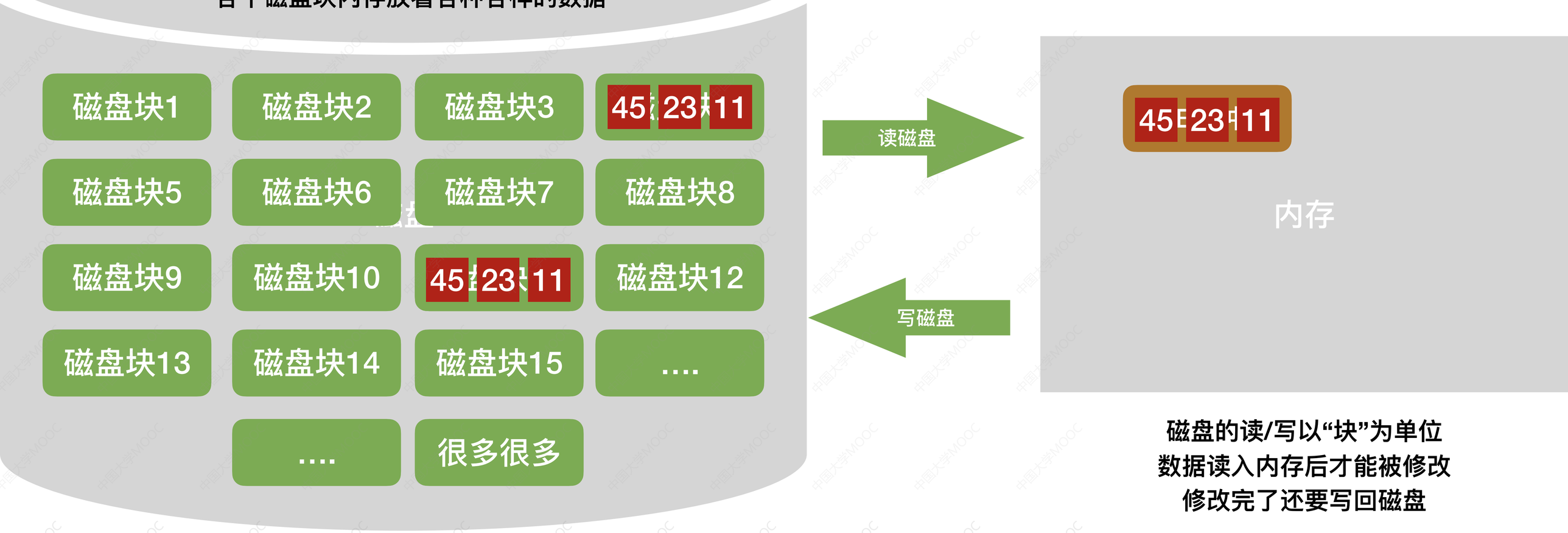


磁盘的读/写以“块”为单位
数据读入内存后才能被修改
修改完了还要写回磁盘

王道考研/CSKAOYAN.COM

外存、内存之间的数据交换

操作系统以“块”为单位对磁盘存储空间进行管理，如：每块大小 1KB
各个磁盘块内存放着各种各样的数据

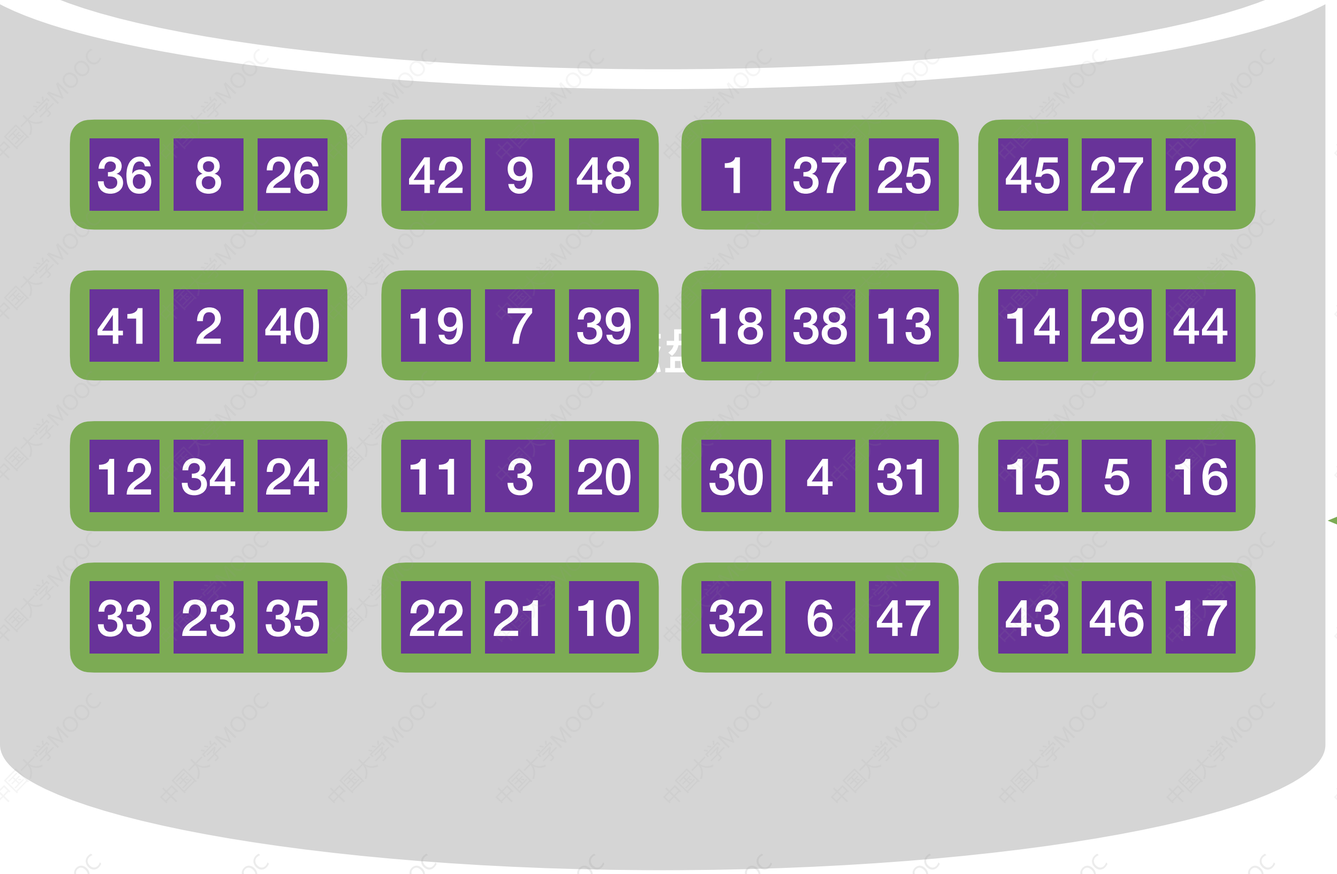


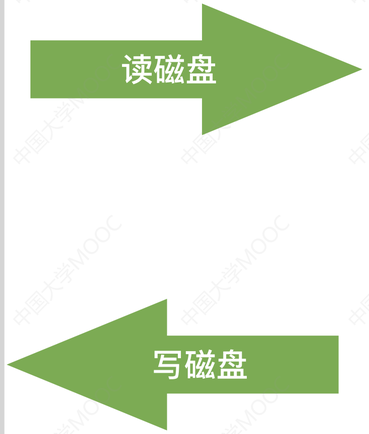
磁盘的读/写以“块”为单位
数据读入内存后才能被修改
修改完了还要写回磁盘

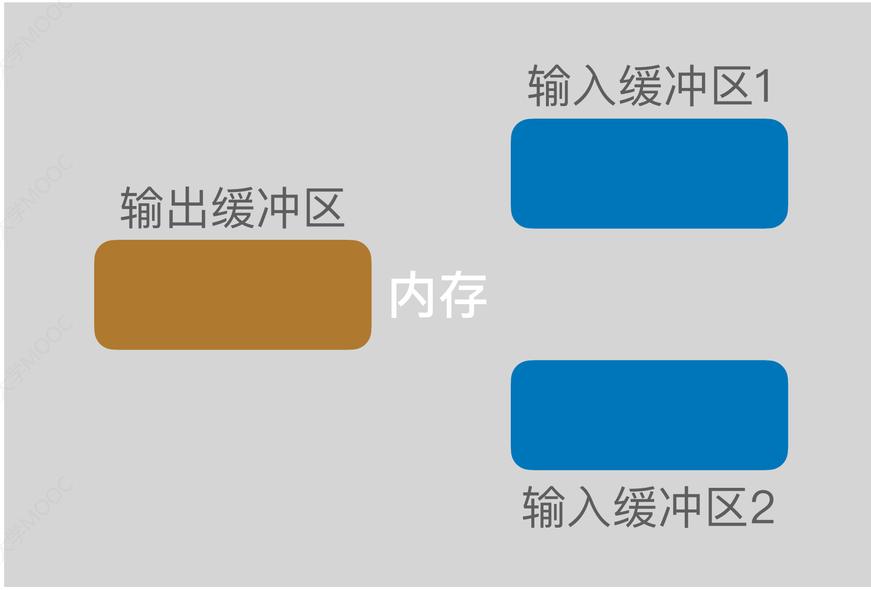
王道考研/CSKAOYAN.COM

外部排序原理

外部排序：数据元素太多，无法一次全部读入内存进行排序





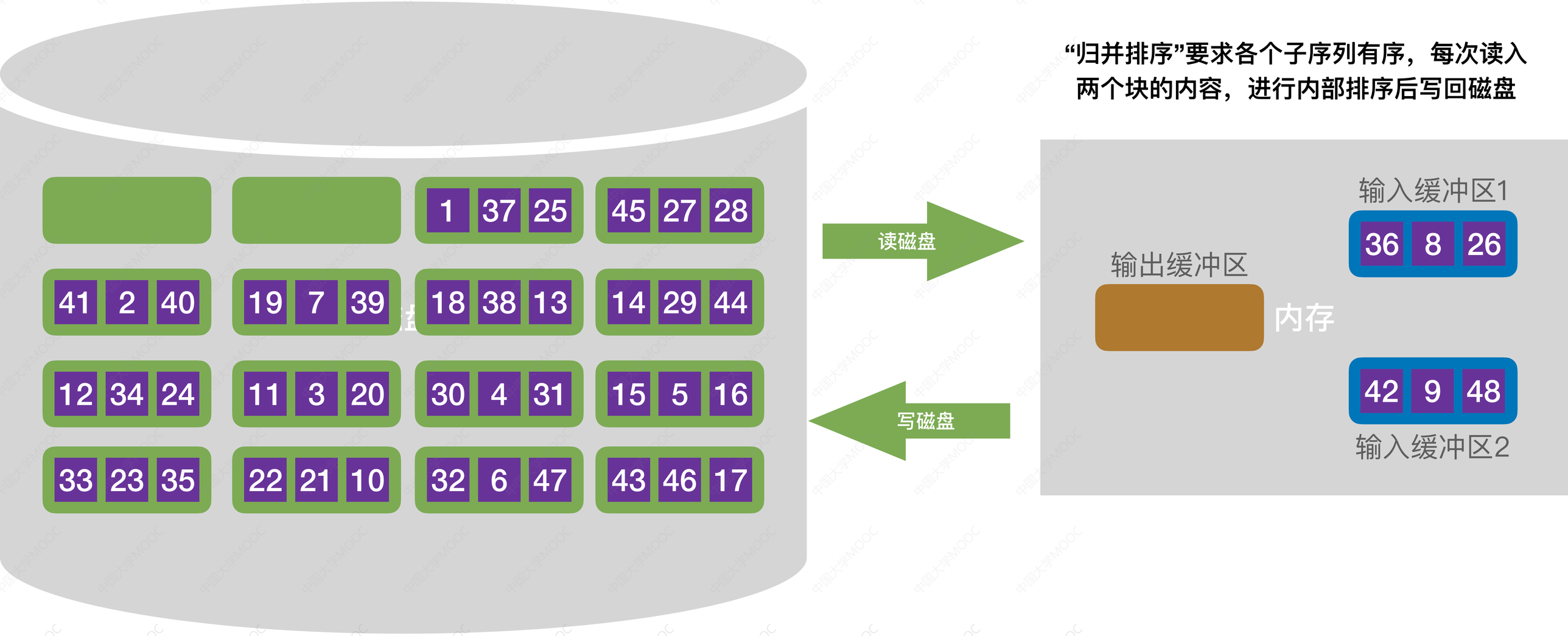


使用“归并排序”的方法，最少只需在内存中分配3块大小的缓冲区即可对任意一个大文件进行排序

王道考研/CSKAOYAN.COM

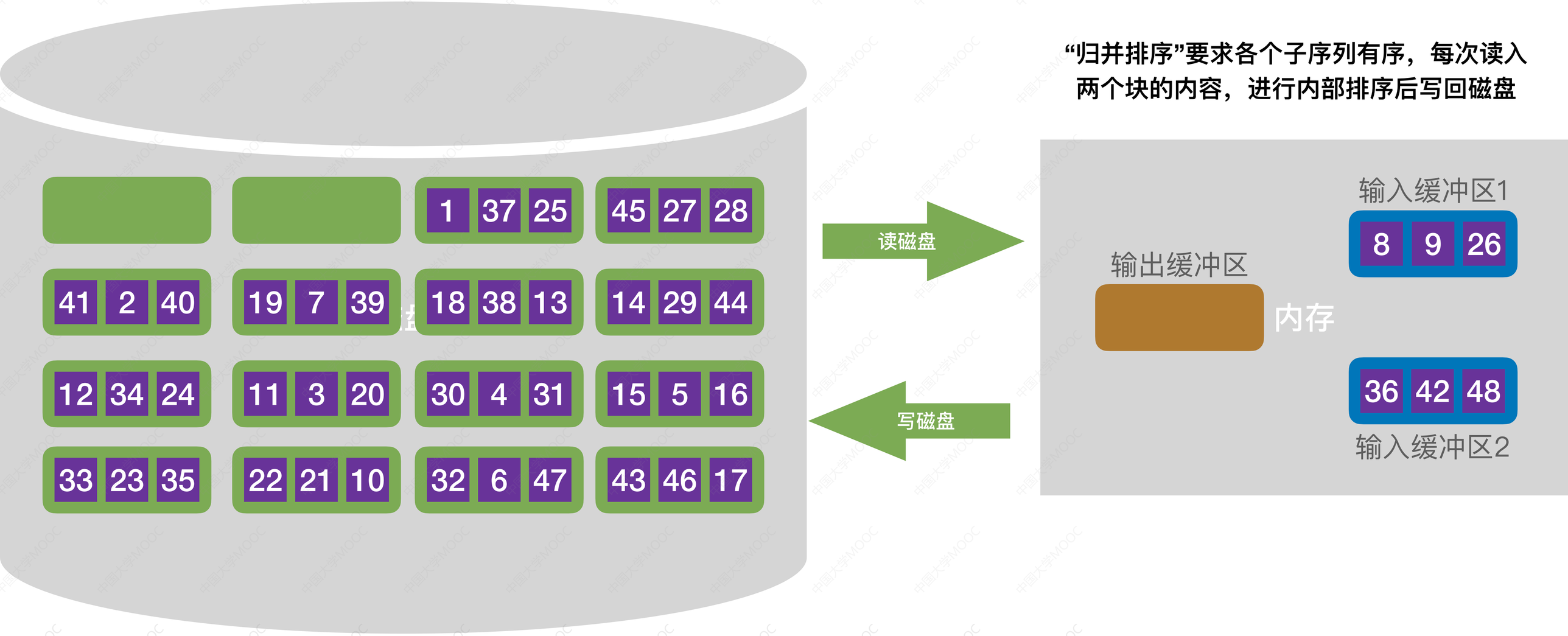
构造初始“归并段”

“归并排序”要求各个子序列有序，每次读入两个块的内容，进行内部排序后写回磁盘



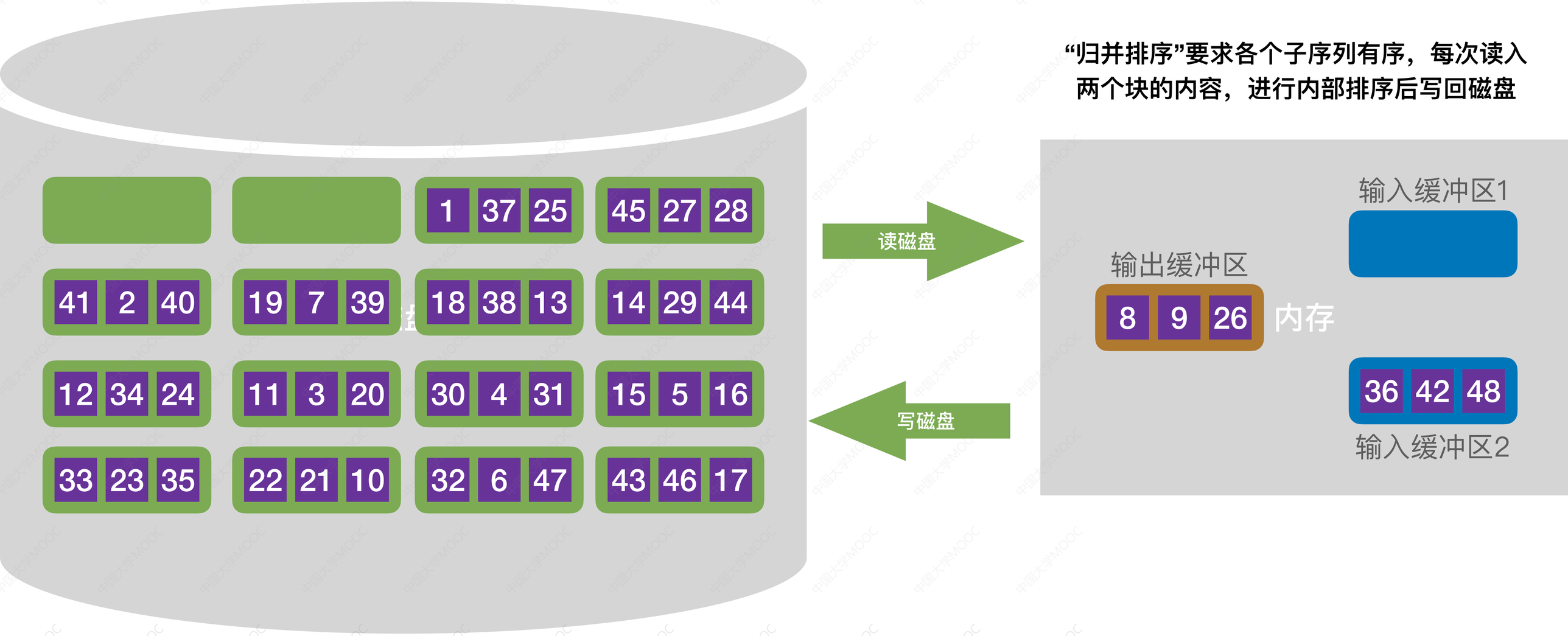
构造初始“归并段”

“归并排序”要求各个子序列有序，每次读入两个块的内容，进行内部排序后写回磁盘



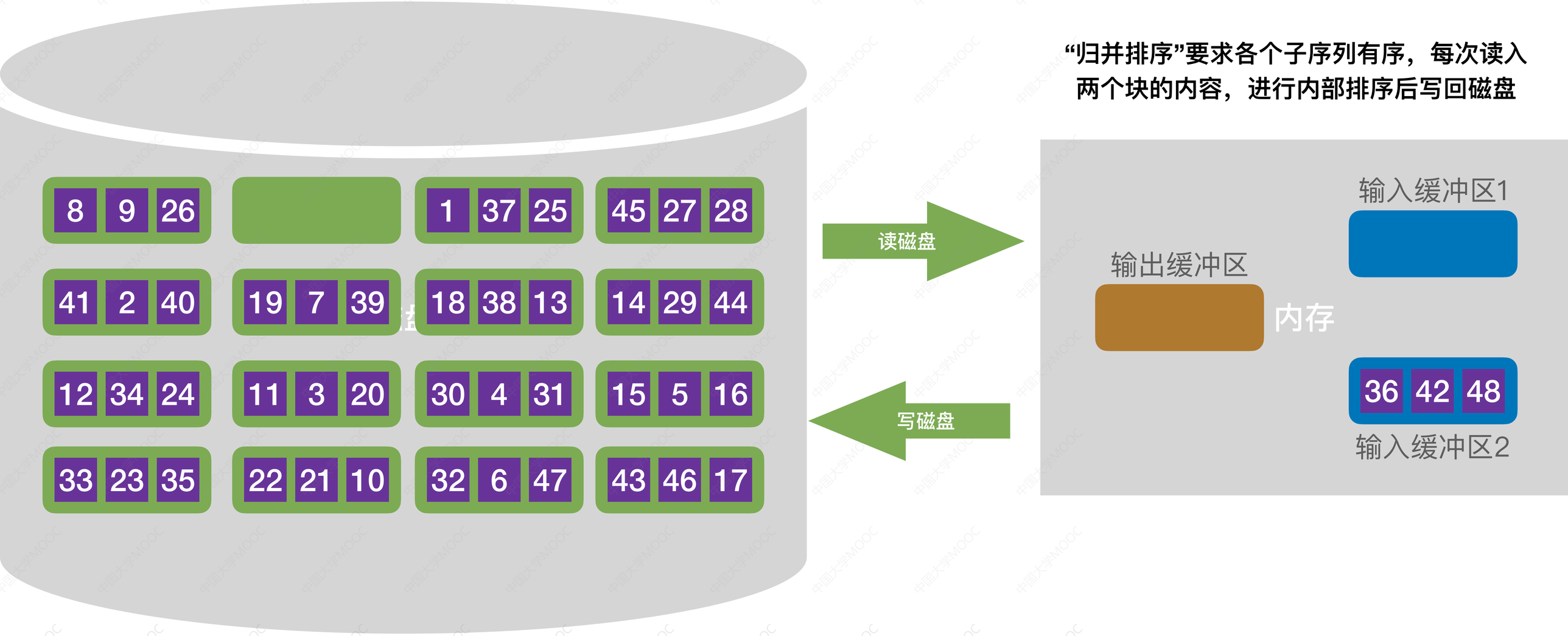
构造初始“归并段”

“归并排序”要求各个子序列有序，每次读入两个块的内容，进行内部排序后写回磁盘



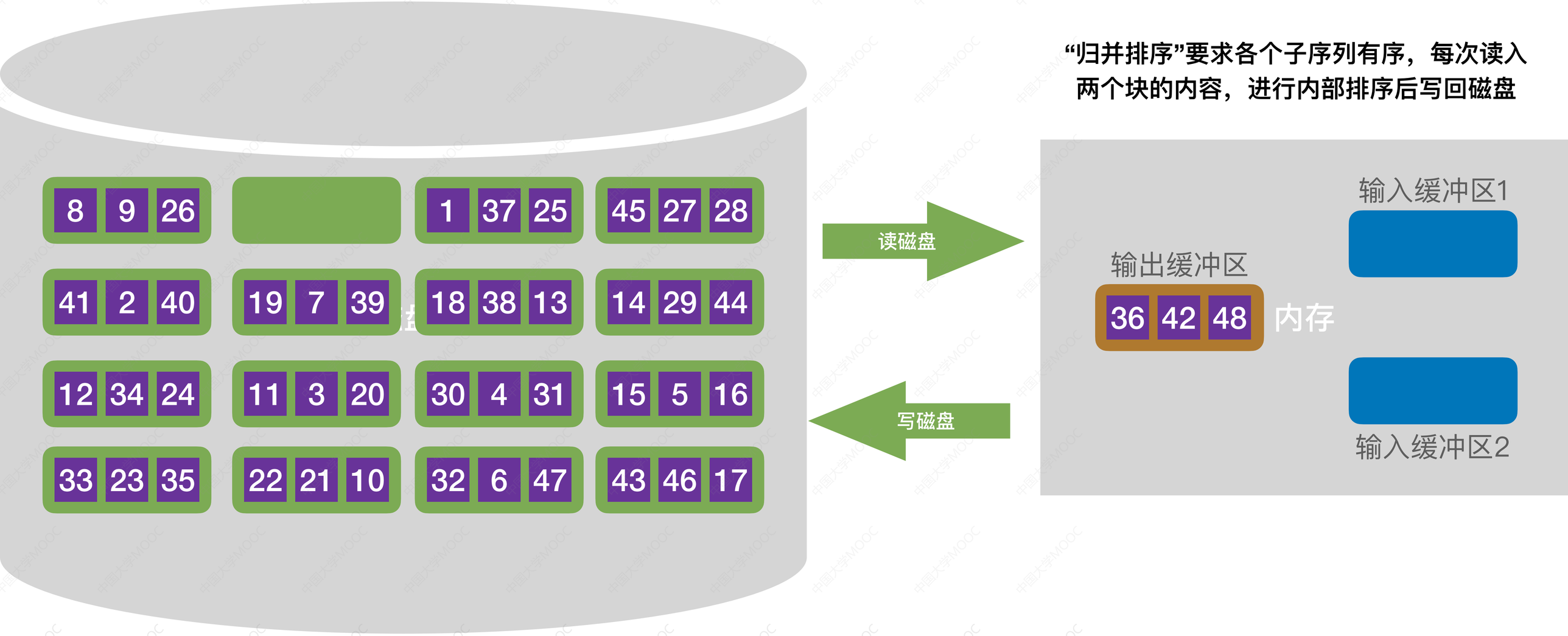
构造初始“归并段”

“归并排序”要求各个子序列有序，每次读入两个块的内容，进行内部排序后写回磁盘



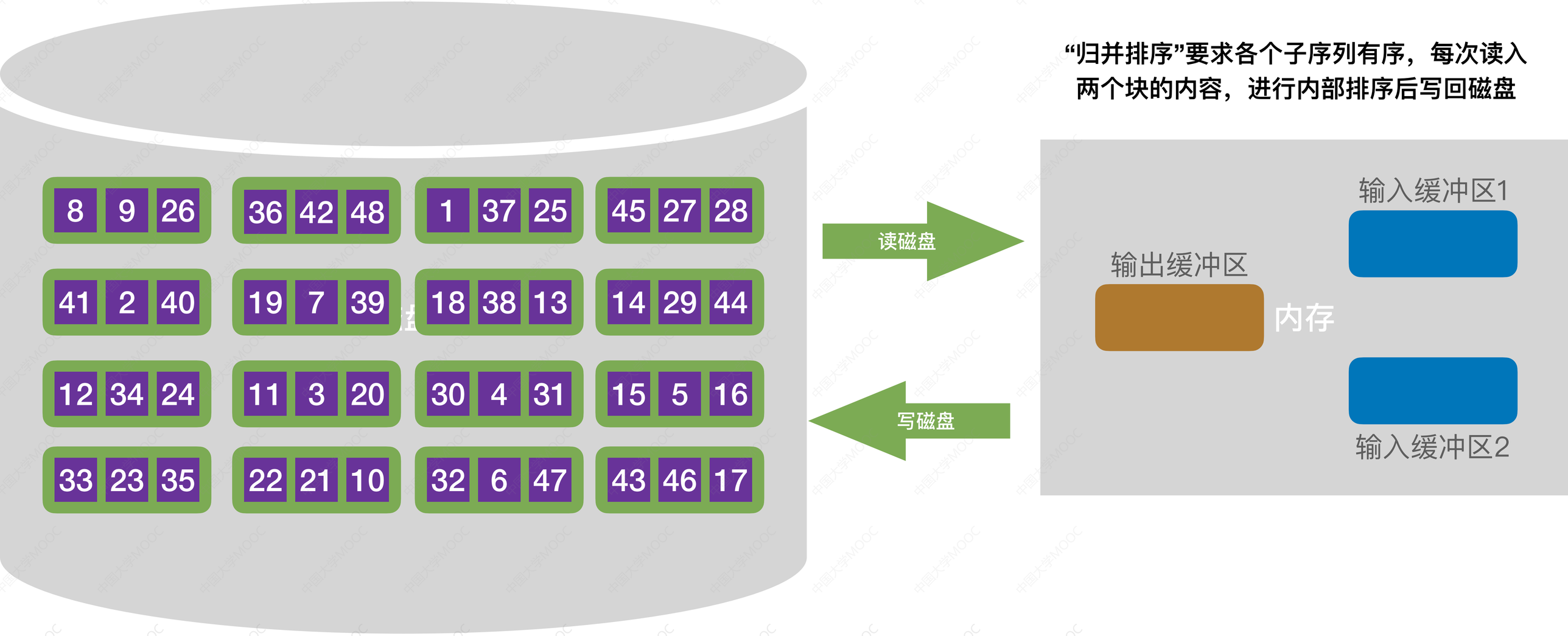
构造初始“归并段”

“归并排序”要求各个子序列有序，每次读入两个块的内容，进行内部排序后写回磁盘



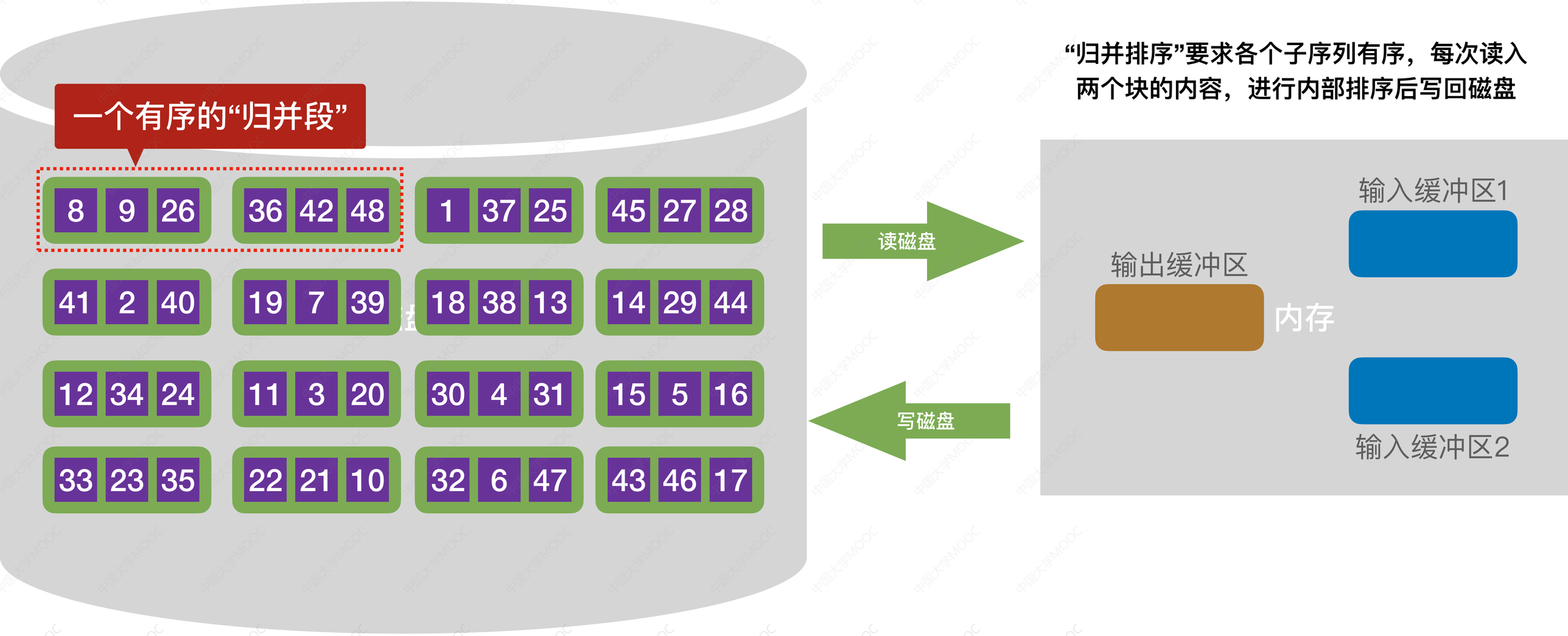
构造初始“归并段”

“归并排序”要求各个子序列有序，每次读入两个块的内容，进行内部排序后写回磁盘



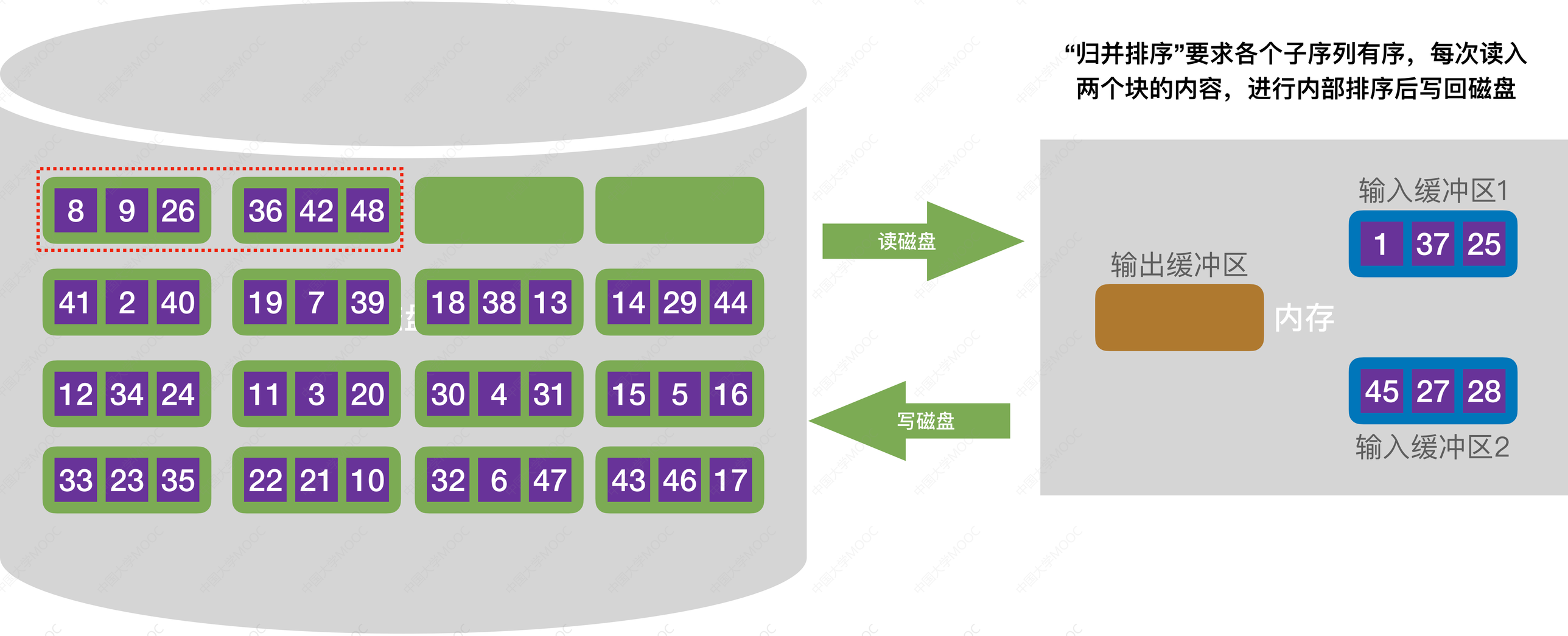
构造初始“归并段”

“归并排序”要求各个子序列有序，每次读入两个块的内容，进行内部排序后写回磁盘



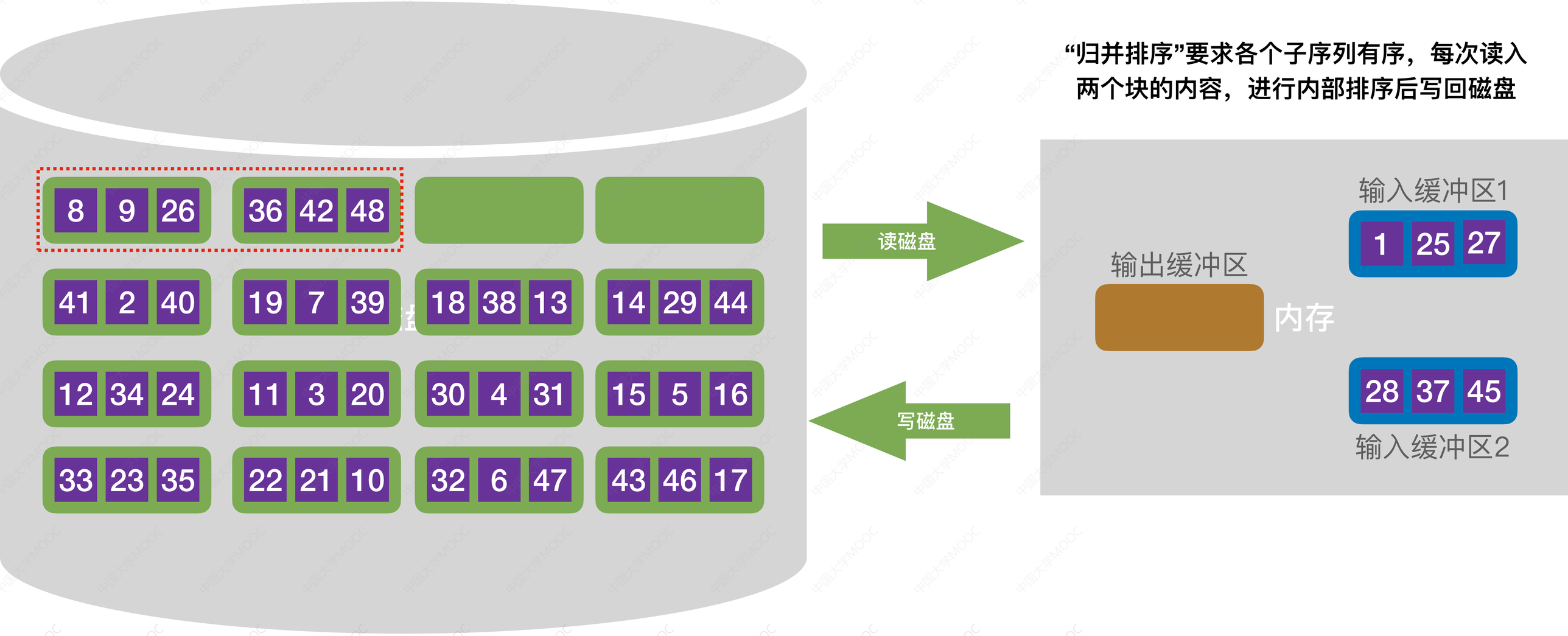
构造初始“归并段”

“归并排序”要求各个子序列有序，每次读入两个块的内容，进行内部排序后写回磁盘



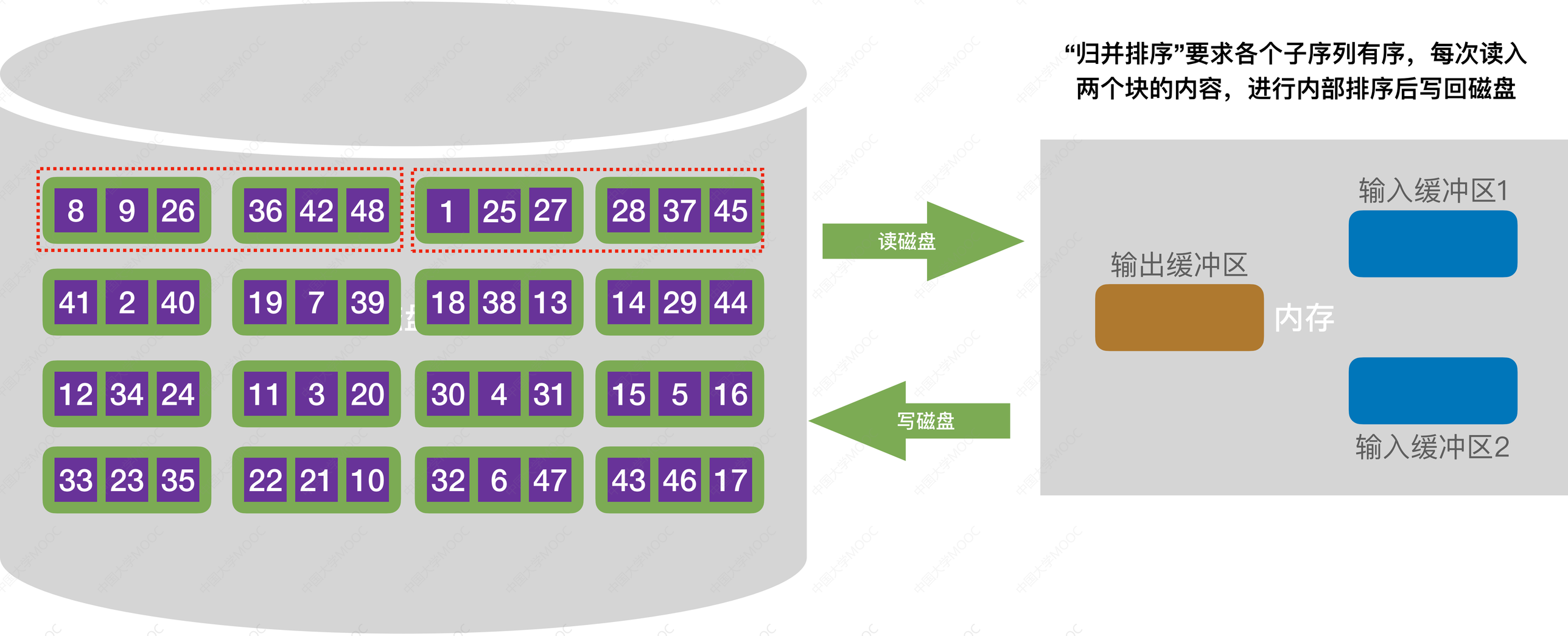
构造初始“归并段”

“归并排序”要求各个子序列有序，每次读入两个块的内容，进行内部排序后写回磁盘



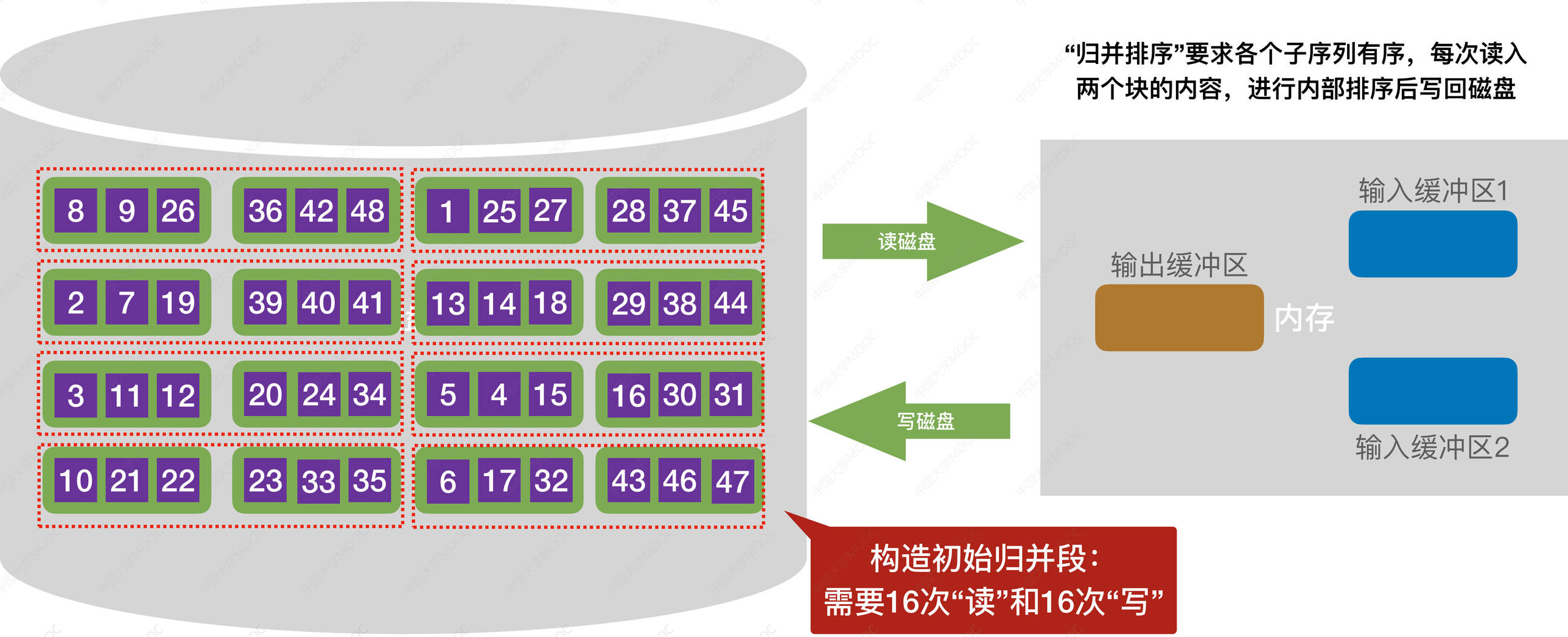
构造初始“归并段”

“归并排序”要求各个子序列有序，每次读入两个块的内容，进行内部排序后写回磁盘



构造初始“归并段”

“归并排序”要求各个子序列有序，每次读入两个块的内容，进行内部排序后写回磁盘

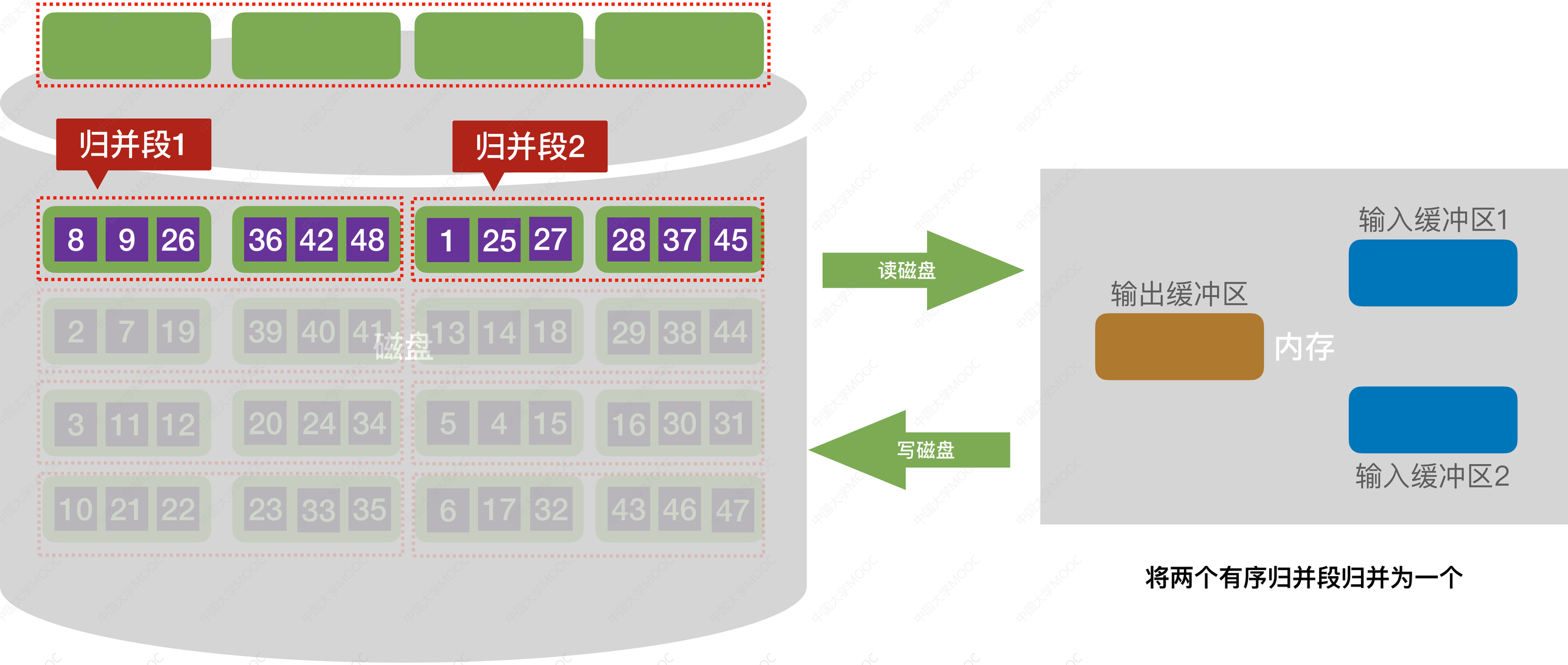


构造初始归并段：
需要16次“读”和16次“写”

王道考研/CSKAOYAN.COM

第一趟归并

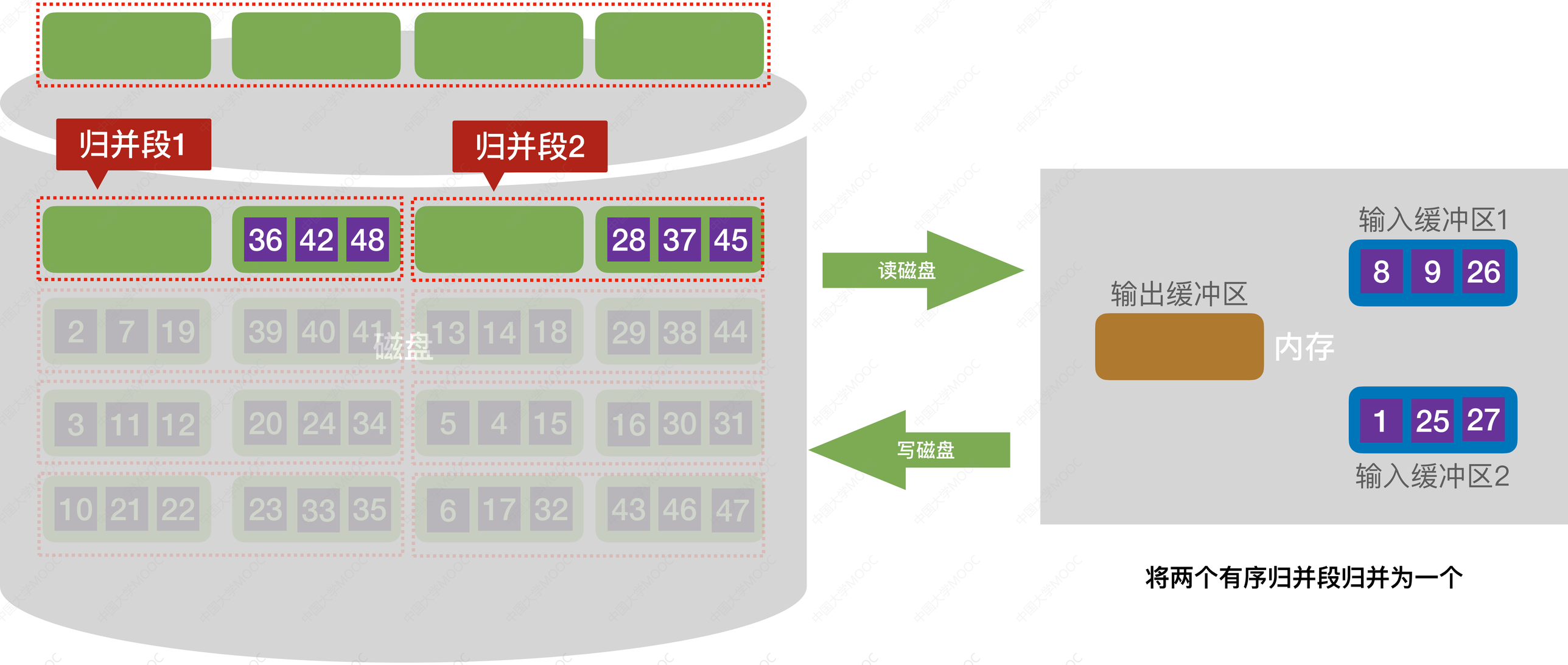
把8个有序子序列（初始归并段）两两归并



将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM

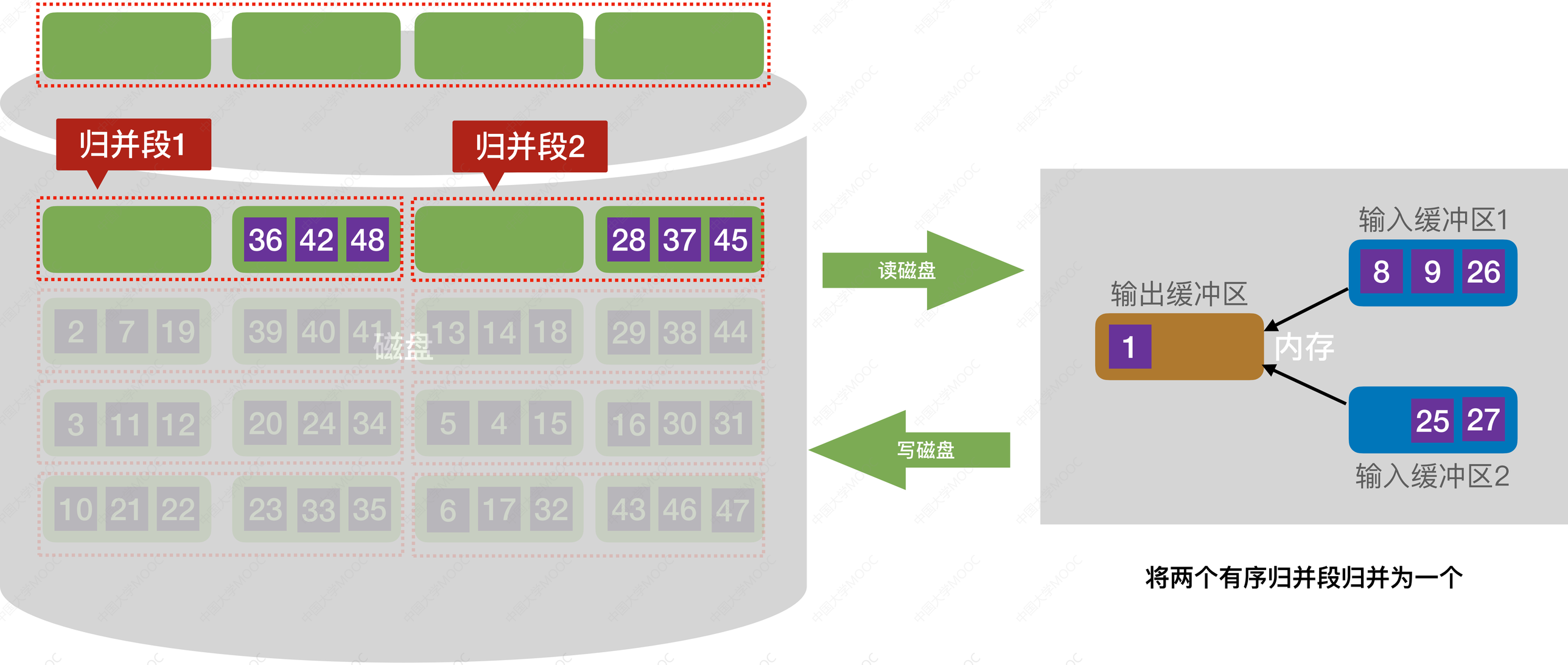
第一趟归并



将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM

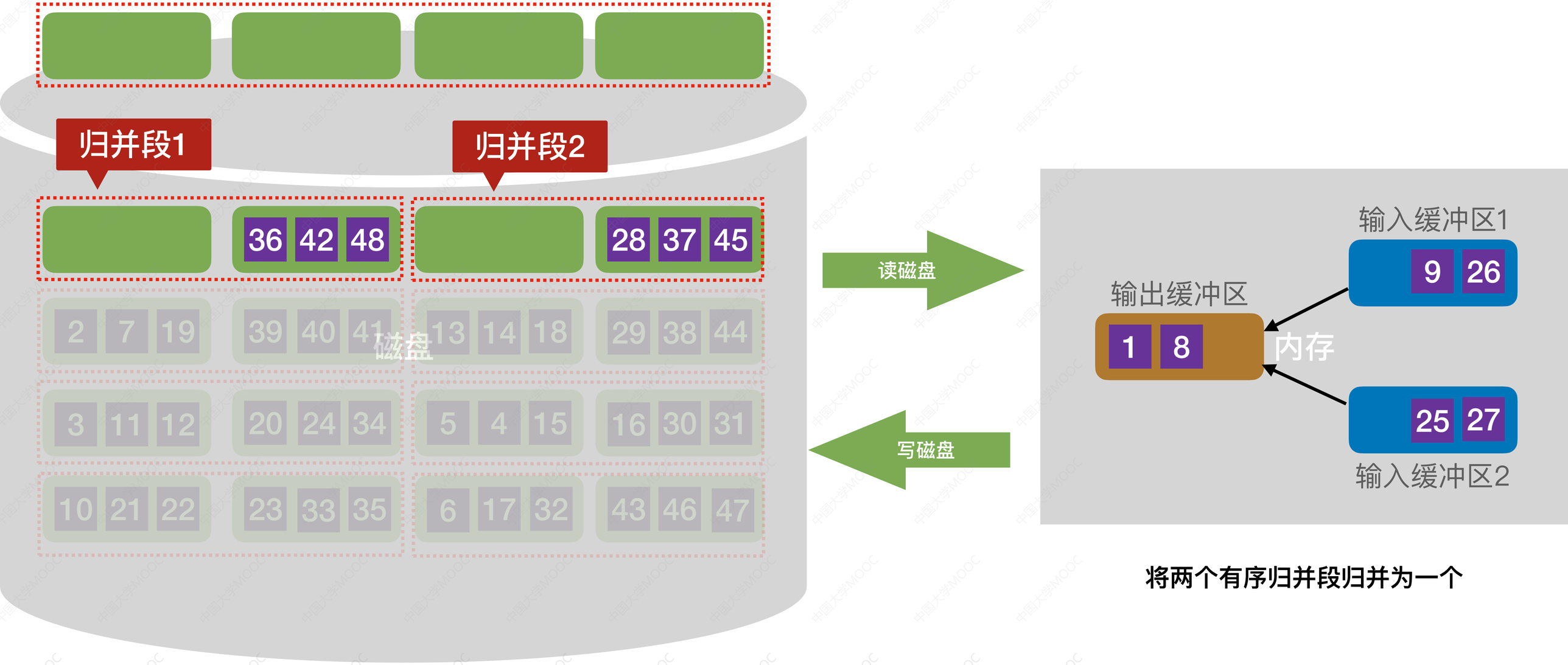
第一趟归并



将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM

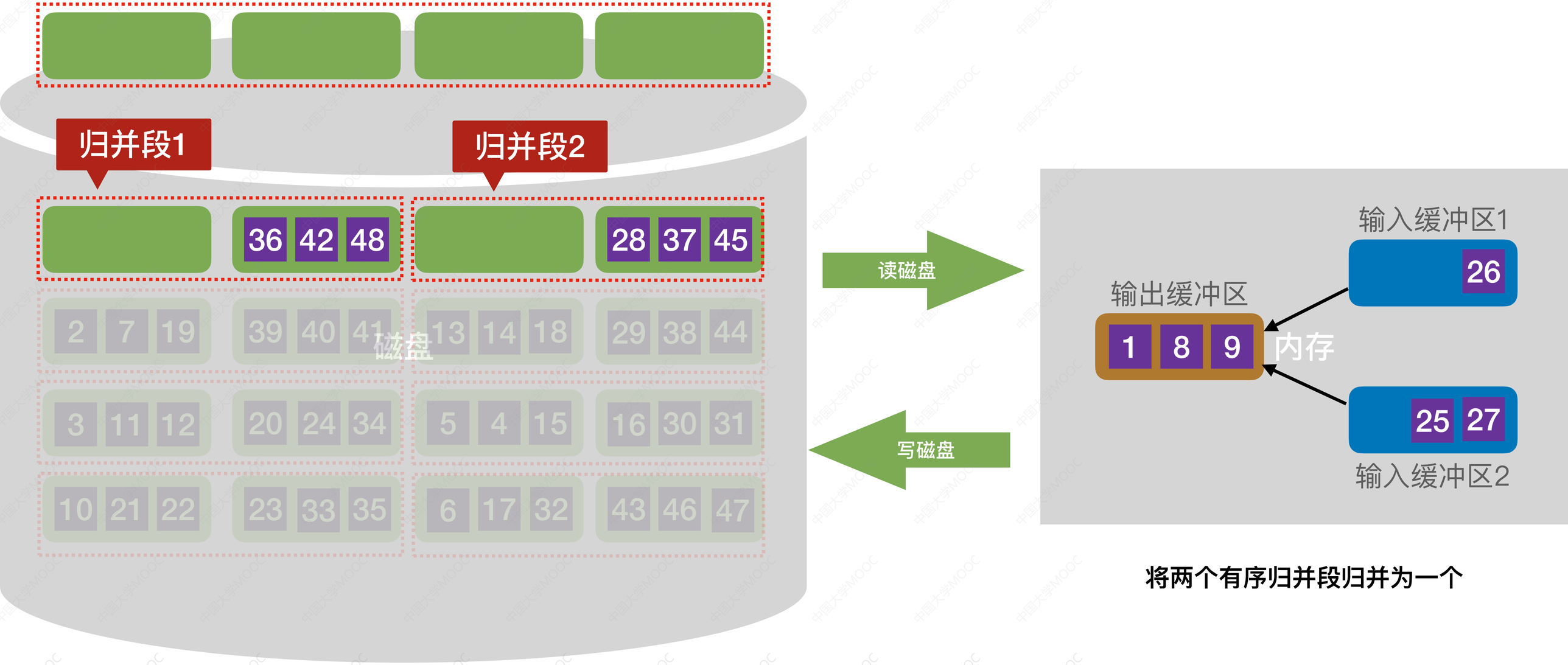
第一趟归并



将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM

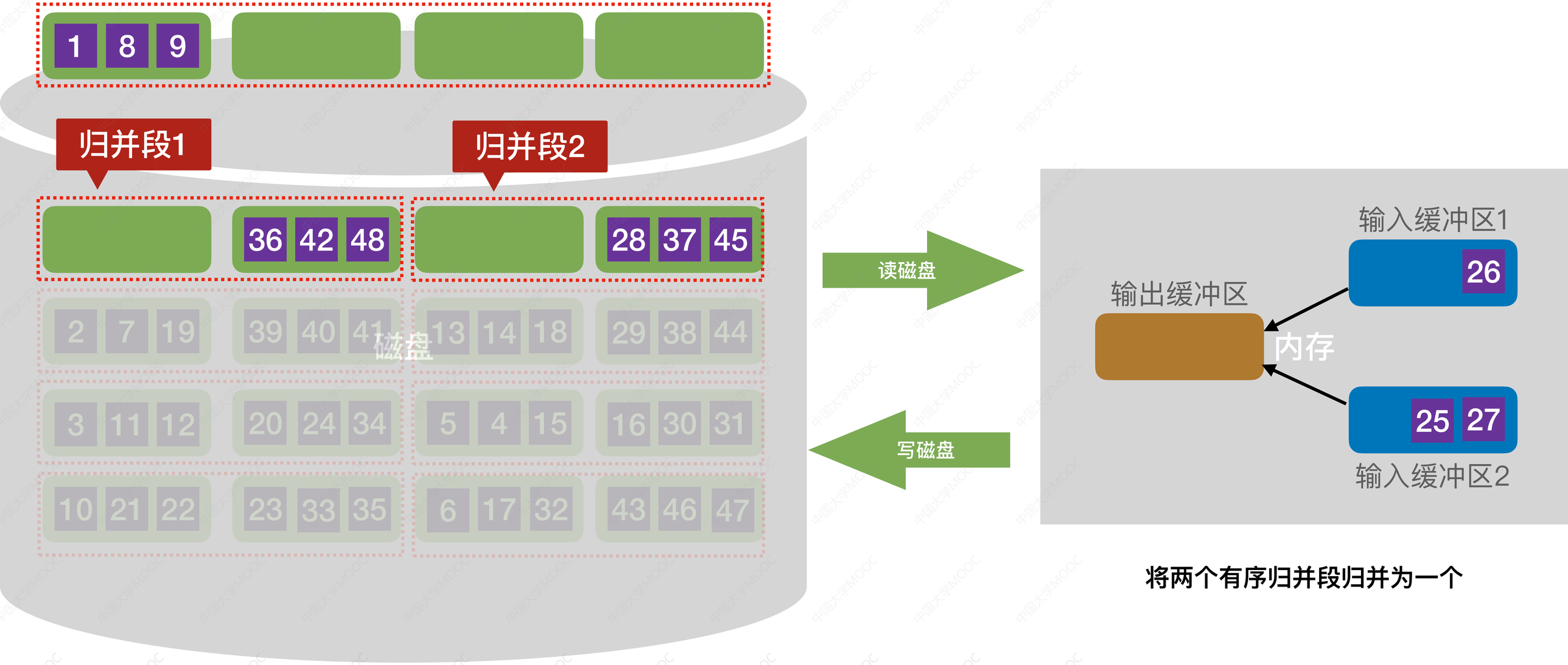
第一趟归并



将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM

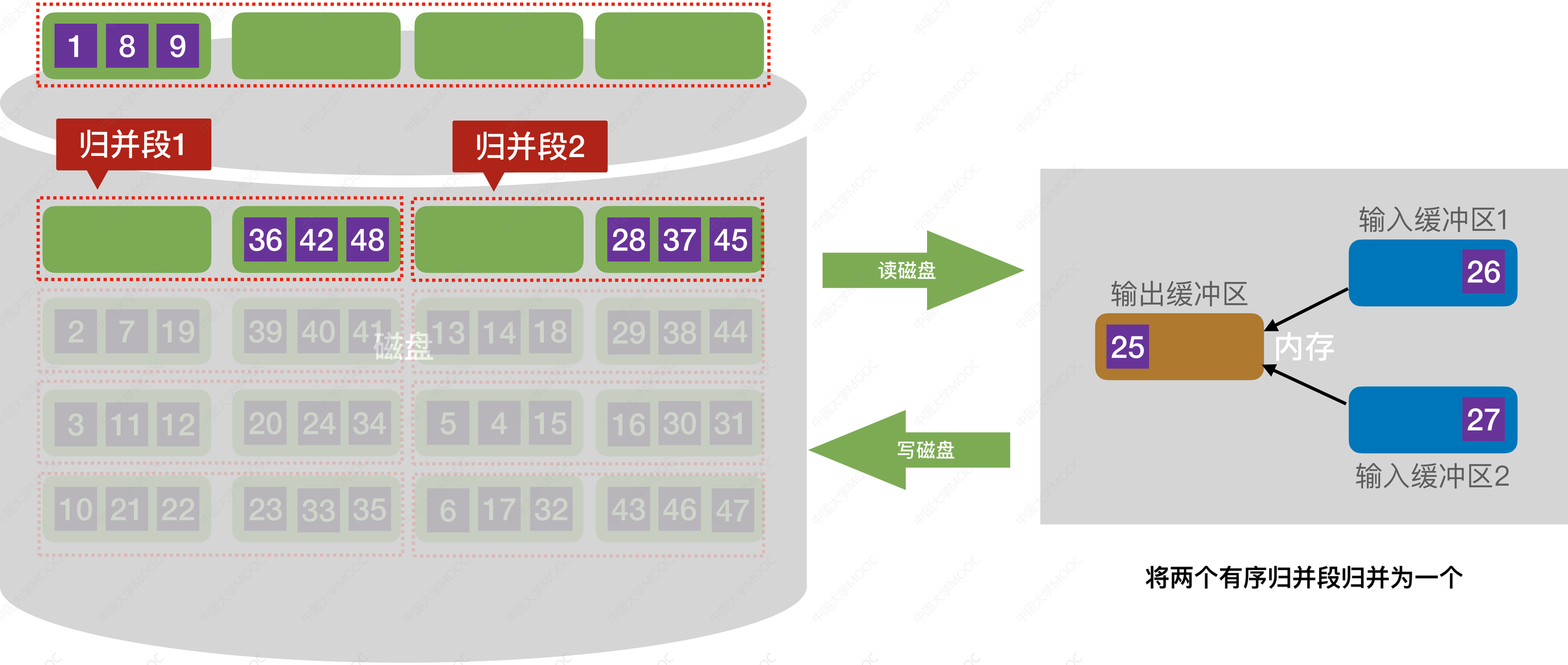
第一趟归并



将两个有序归并段归并为一个

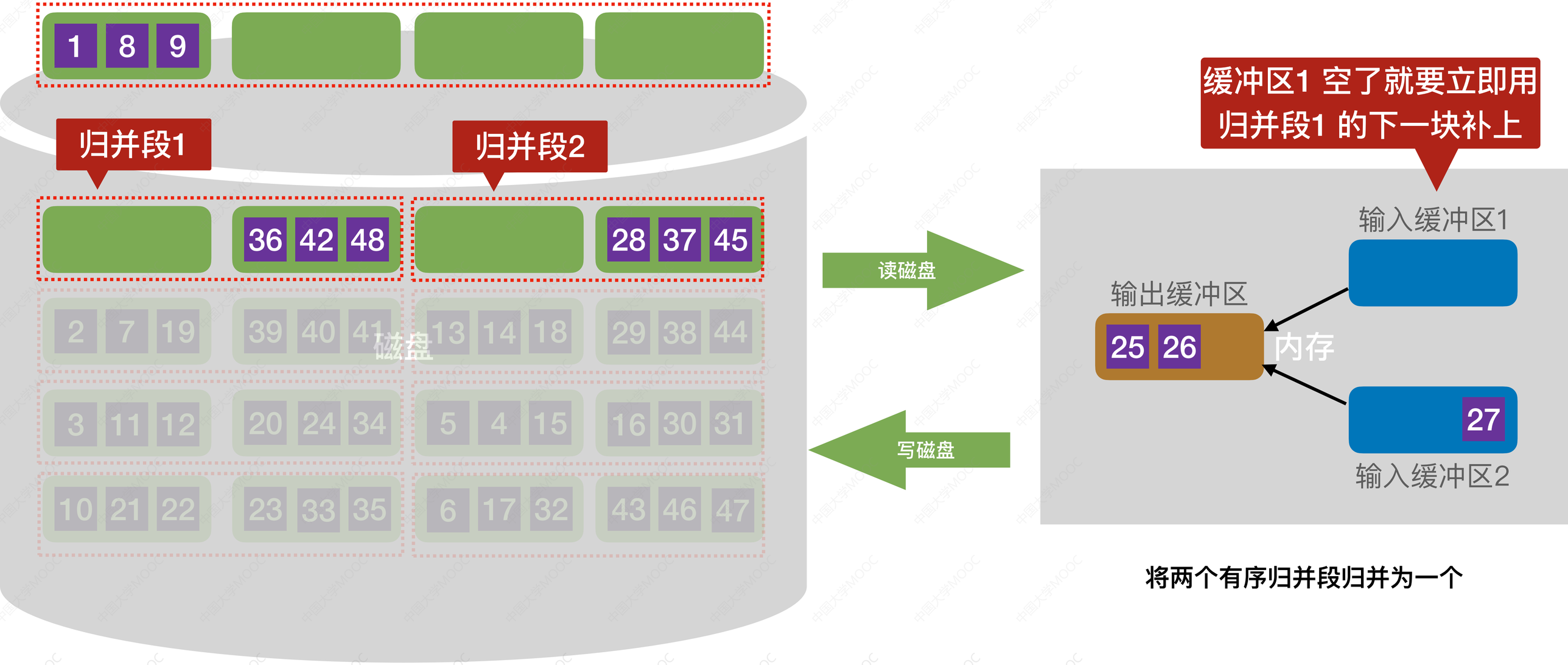
王道考研/CSKAOYAN.COM

第一趟归并



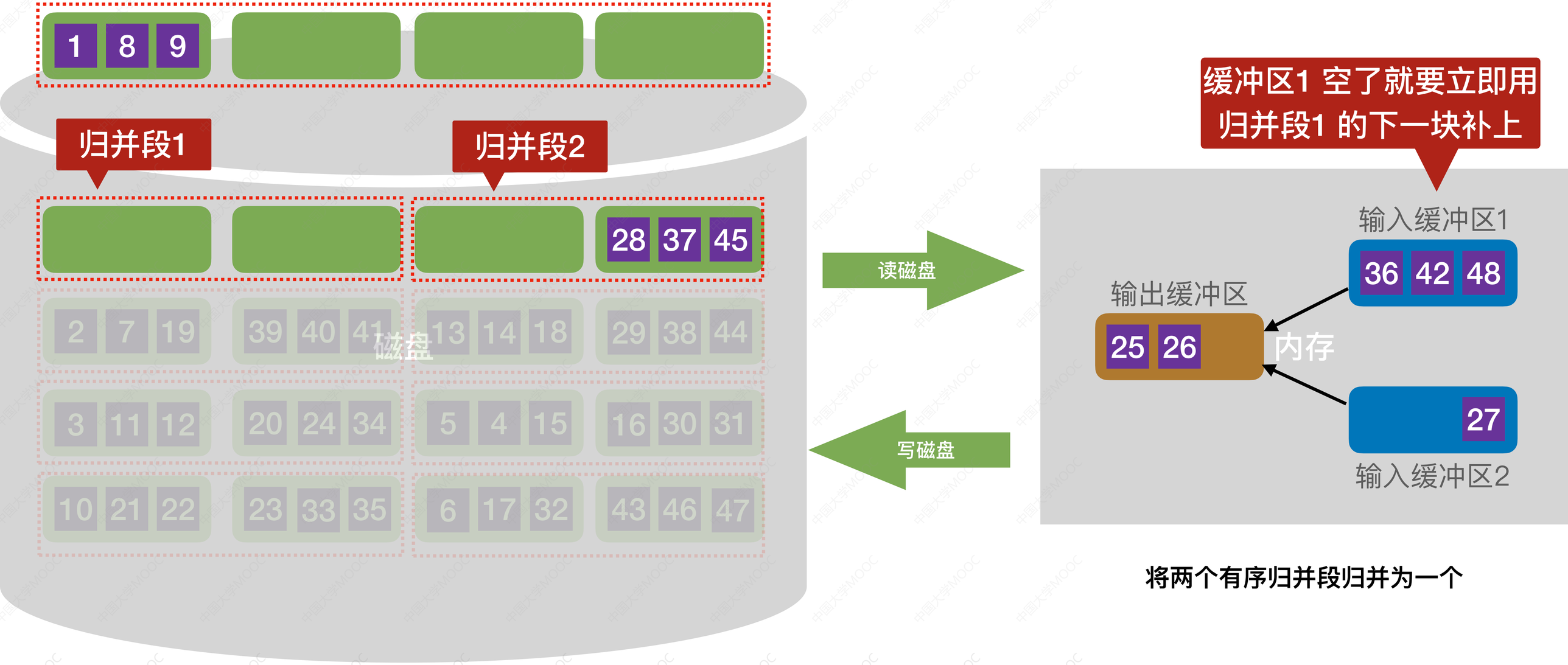
将两个有序归并段归并为一个

第一趟归并



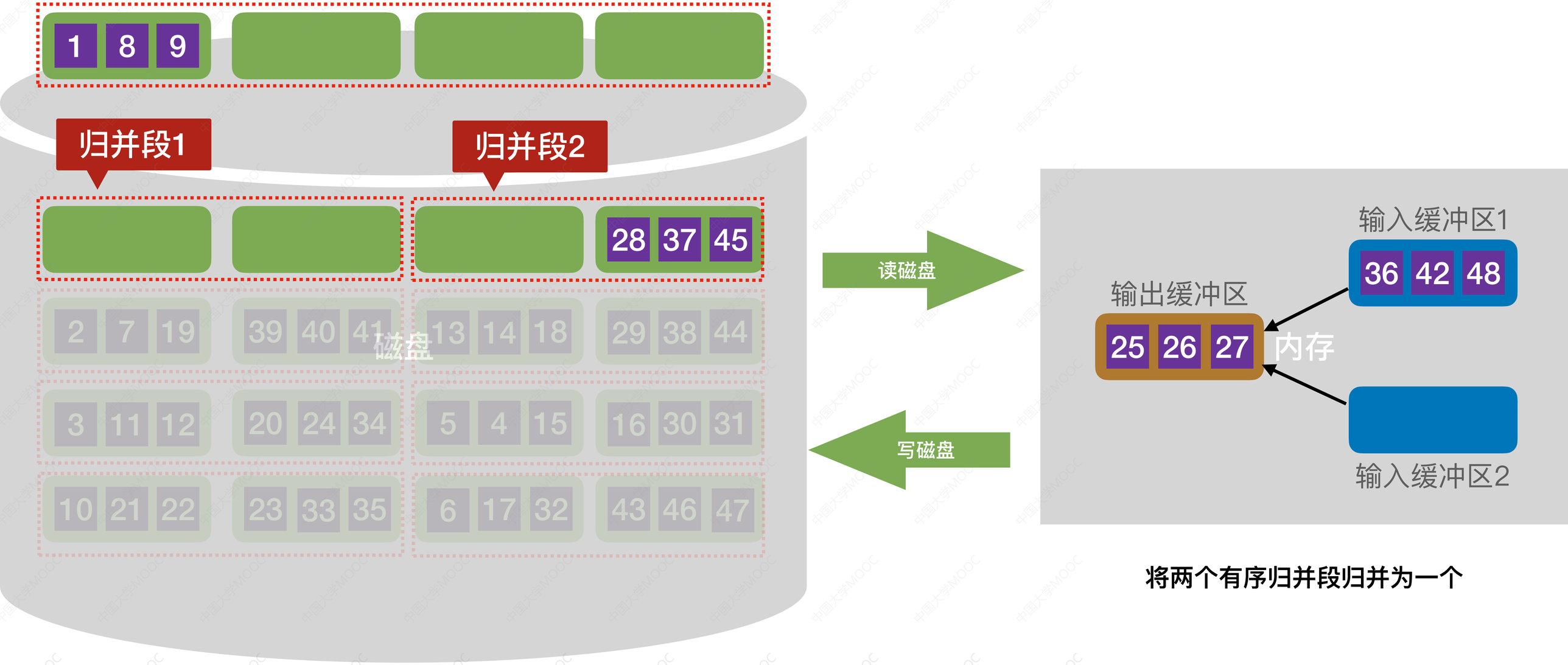
将两个有序归并段归并为一个

第一趟归并



将两个有序归并段归并为一个

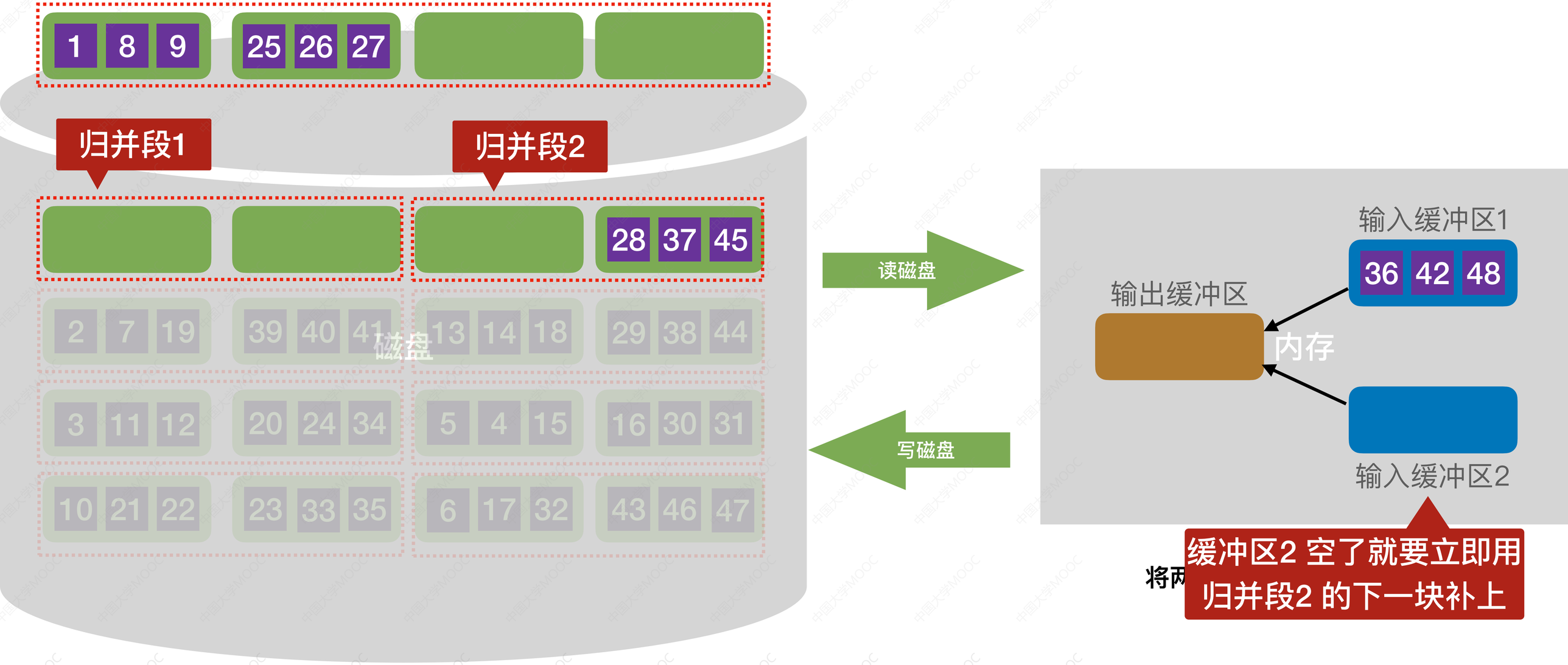
第一趟归并



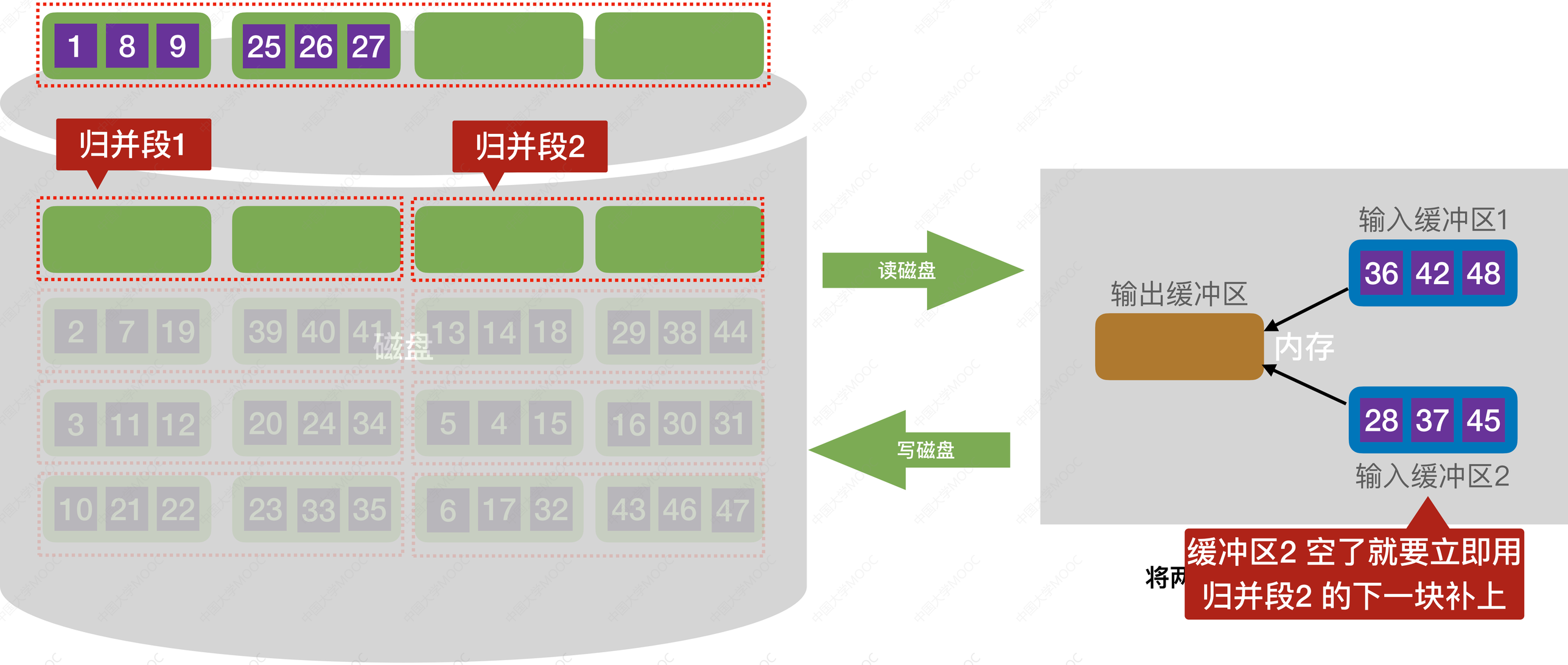
将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM

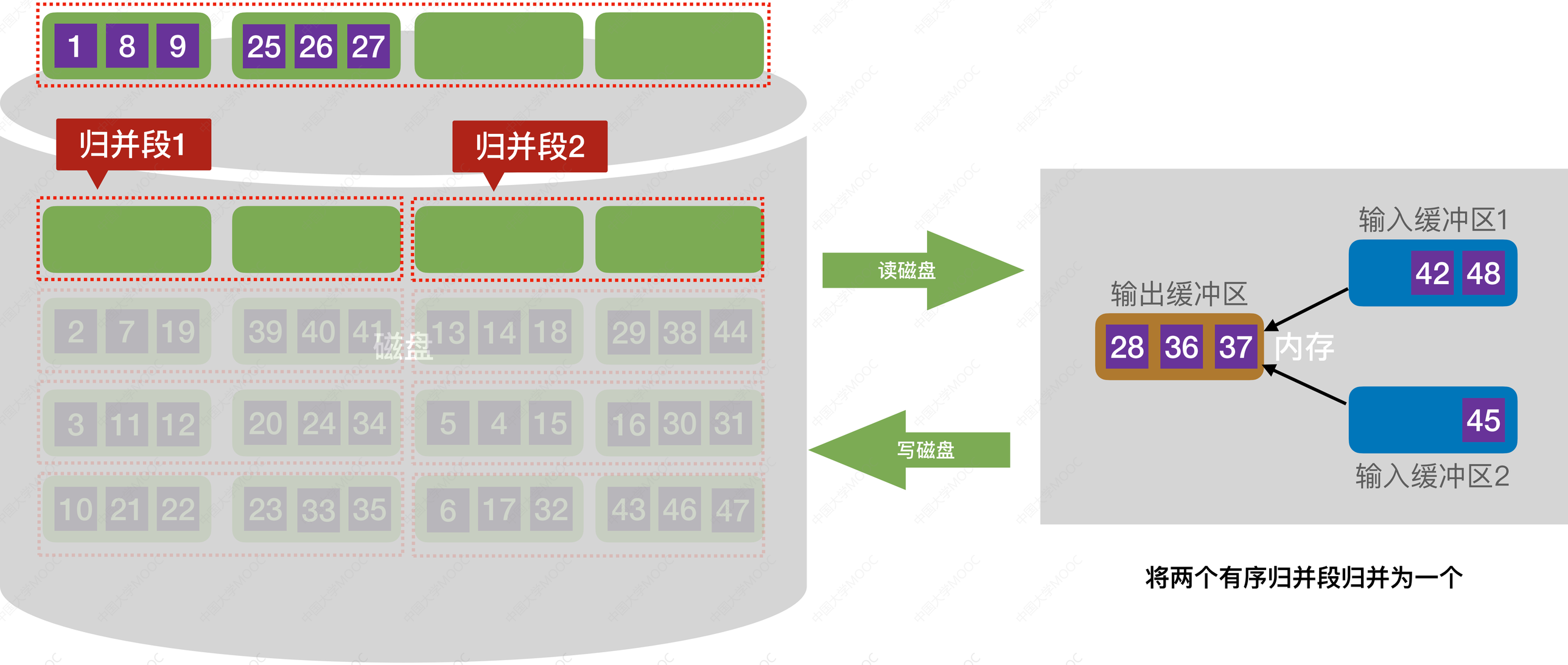
第一趟归并


1 8 9
25 26 27
归并段1
归并段2
28 37 45
2 7 19
39 40 41
13 14 18
29 38 44
磁盘
3 11 12
20 24 34
5 4 15
16 30 31
10 21 22
23 33 35
6 17 32
43 46 47
读磁盘
写磁盘
输入缓冲区1
36 42 48
输出缓冲区
内存
输入缓冲区2
缓冲区2 空了就要立即用
归并段2 的下一块补上

第一趟归并


1 8 9
25 26 27
归并段1
归并段2
2 7 19
39 40 41
13 14 18
29 38 44
磁盘
3 11 12
20 24 34
5 4 15
16 30 31
10 21 22
23 33 35
6 17 32
43 46 47
读磁盘
写磁盘
输入缓冲区1
36 42 48
输出缓冲区
内存
28 37 45
输入缓冲区2
缓冲区2 空了就要立即用
归并段2 的下一块补上

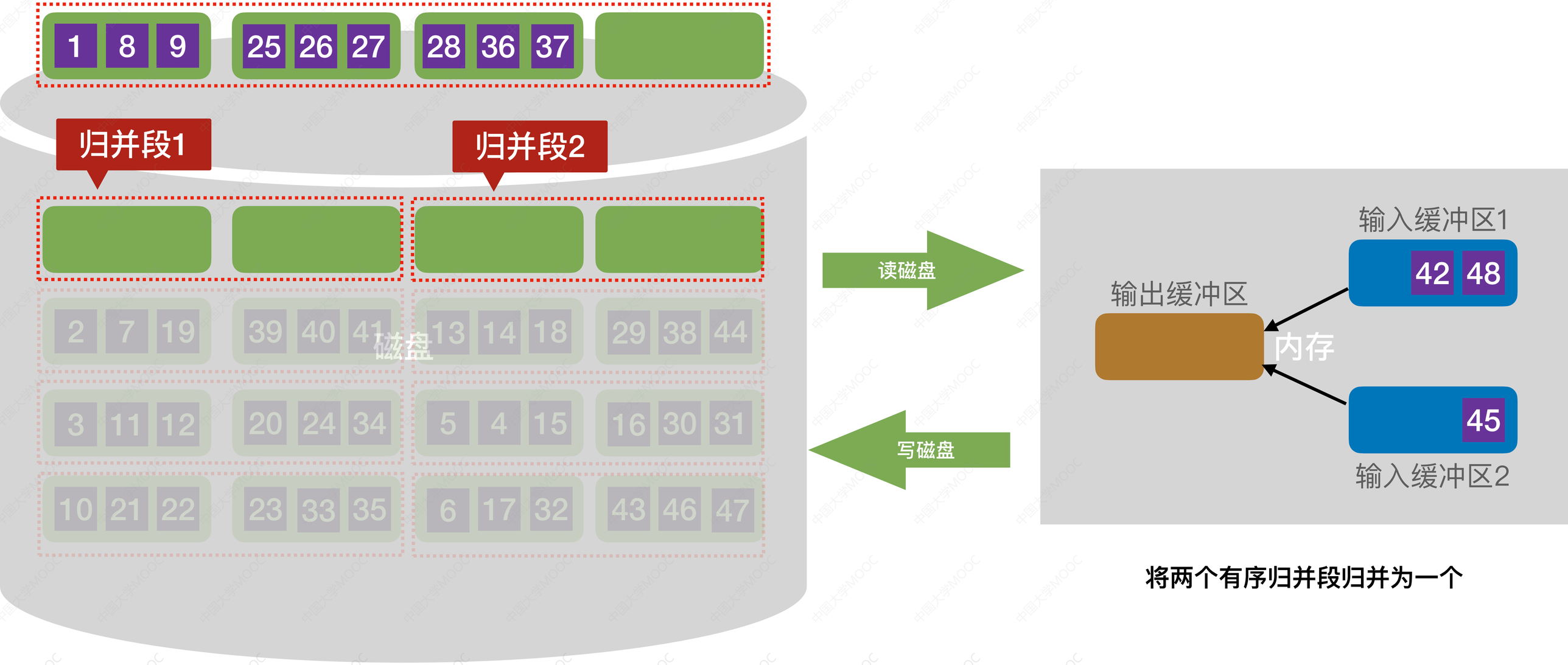
第一趟归并



将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM

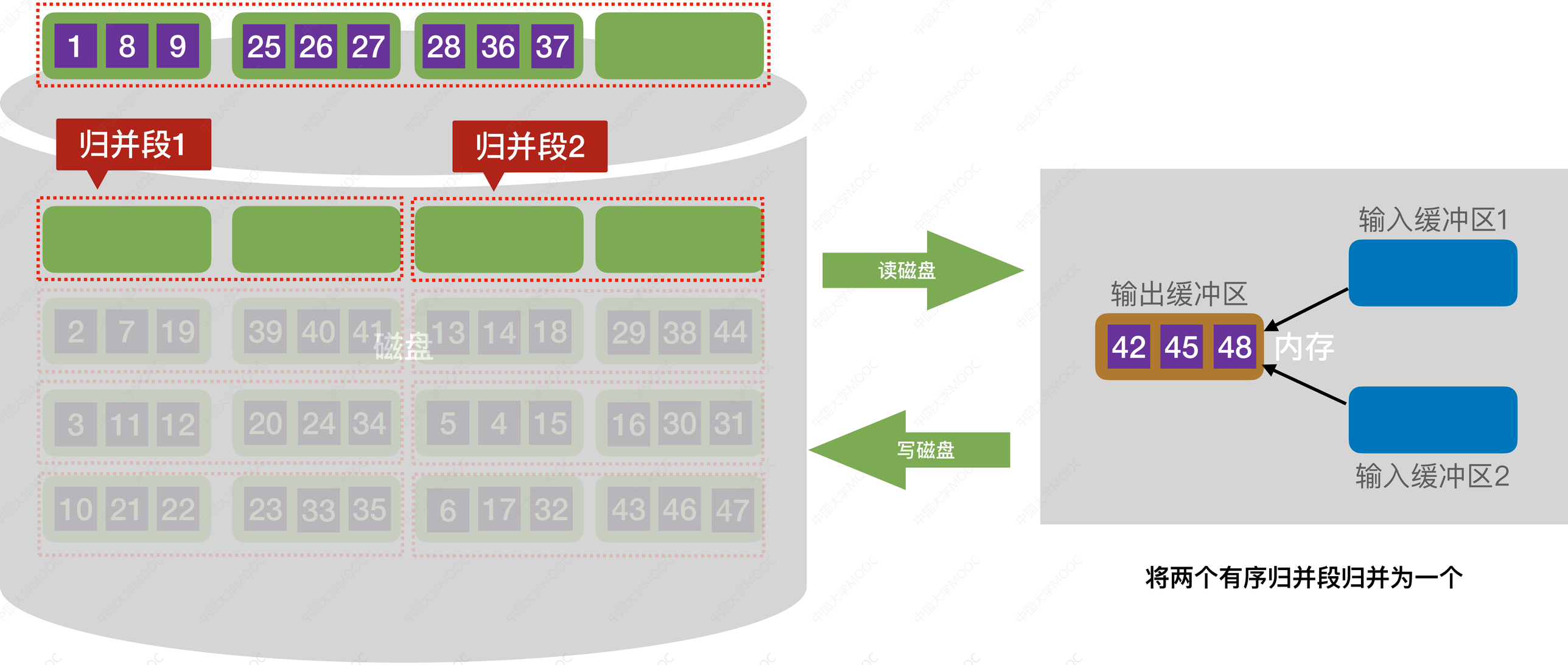
第一趟归并



将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM

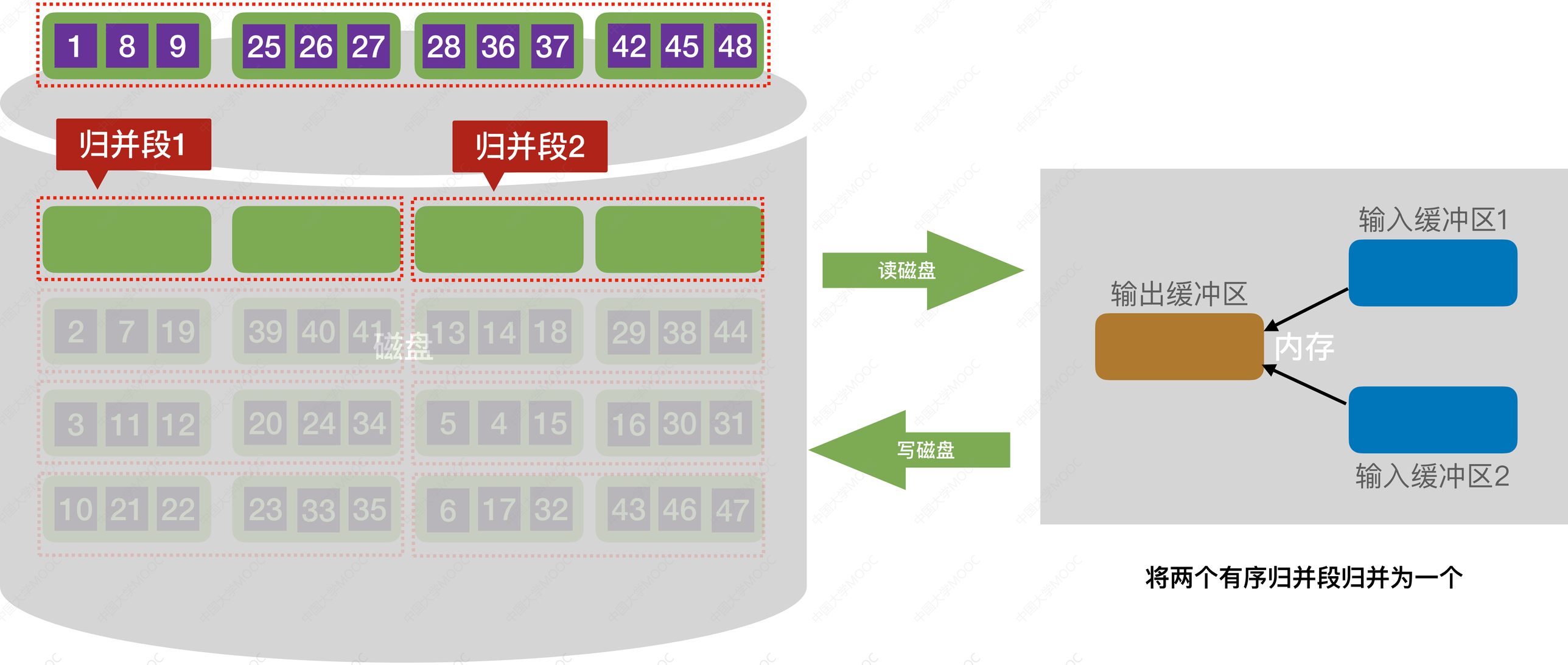
第一趟归并



将两个有序归并段归并为一个

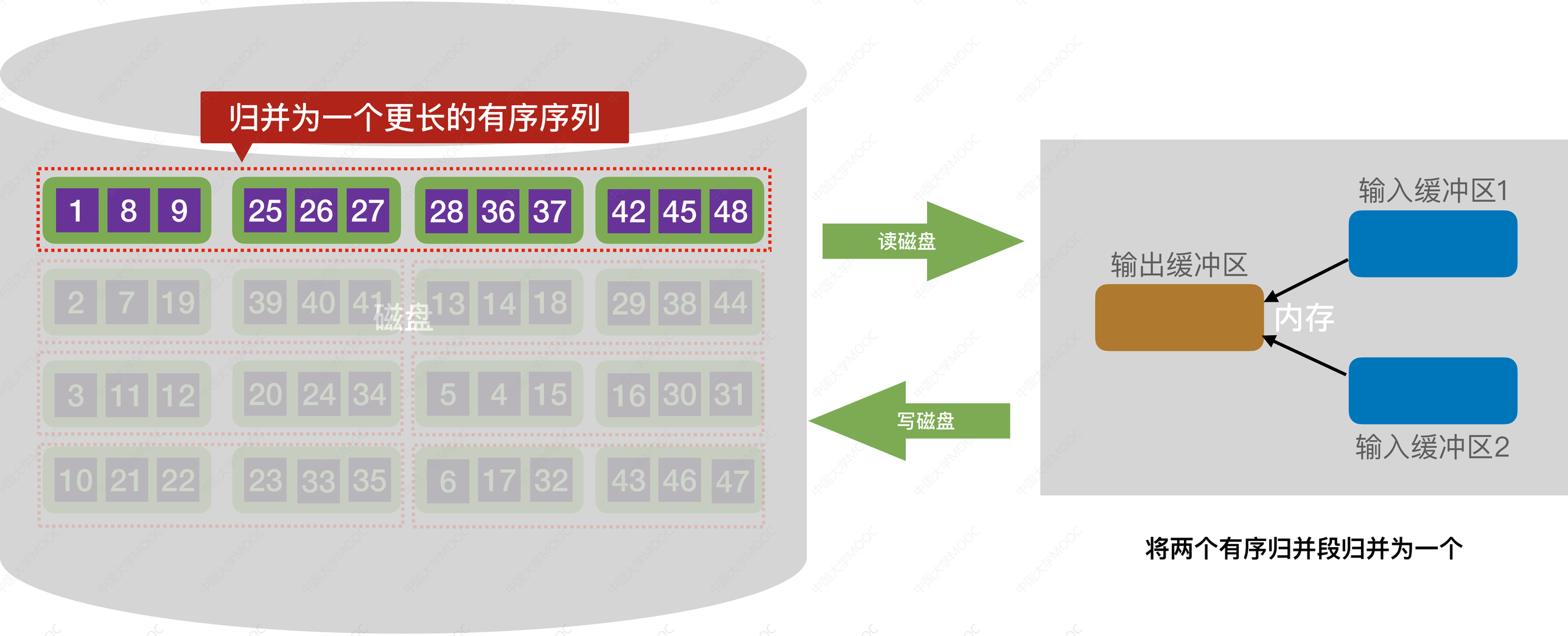
王道考研/CSKAOYAN.COM

第一趟归并



将两个有序归并段归并为一个

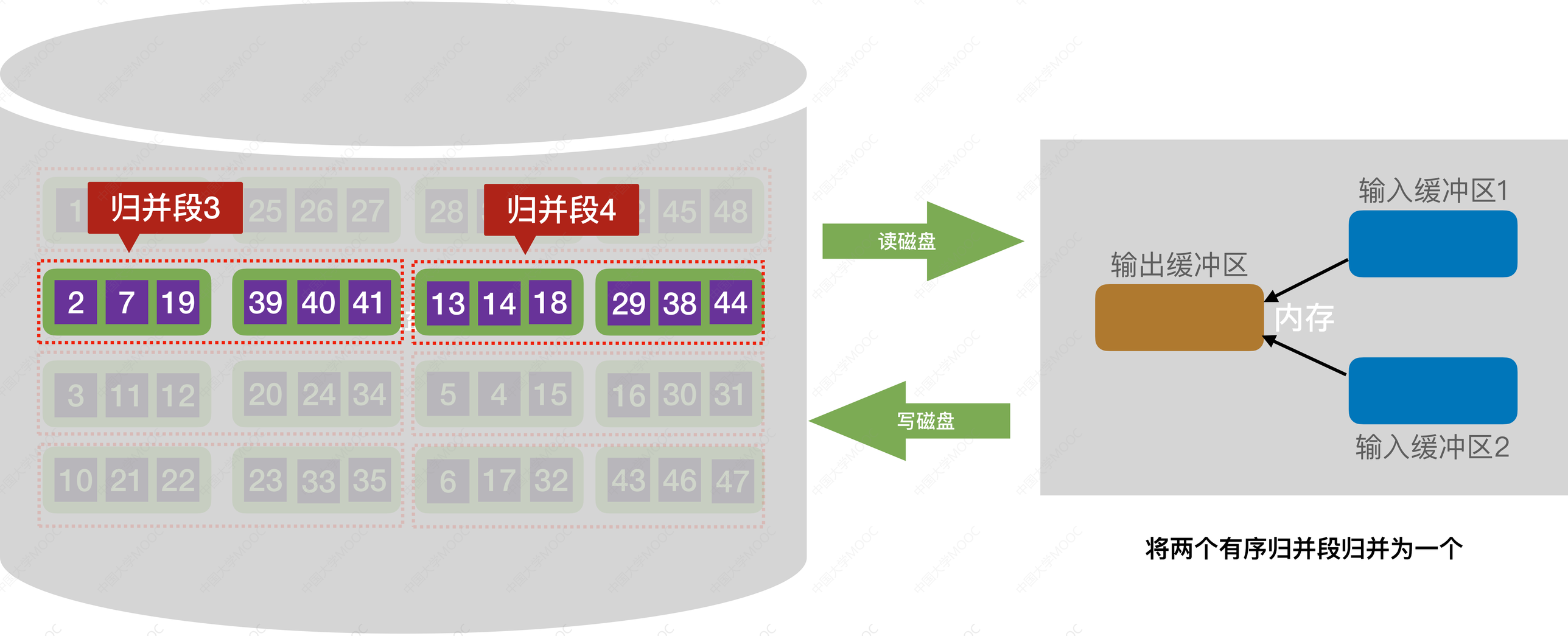
第一趟归并



将两个有序归并段归并为一个

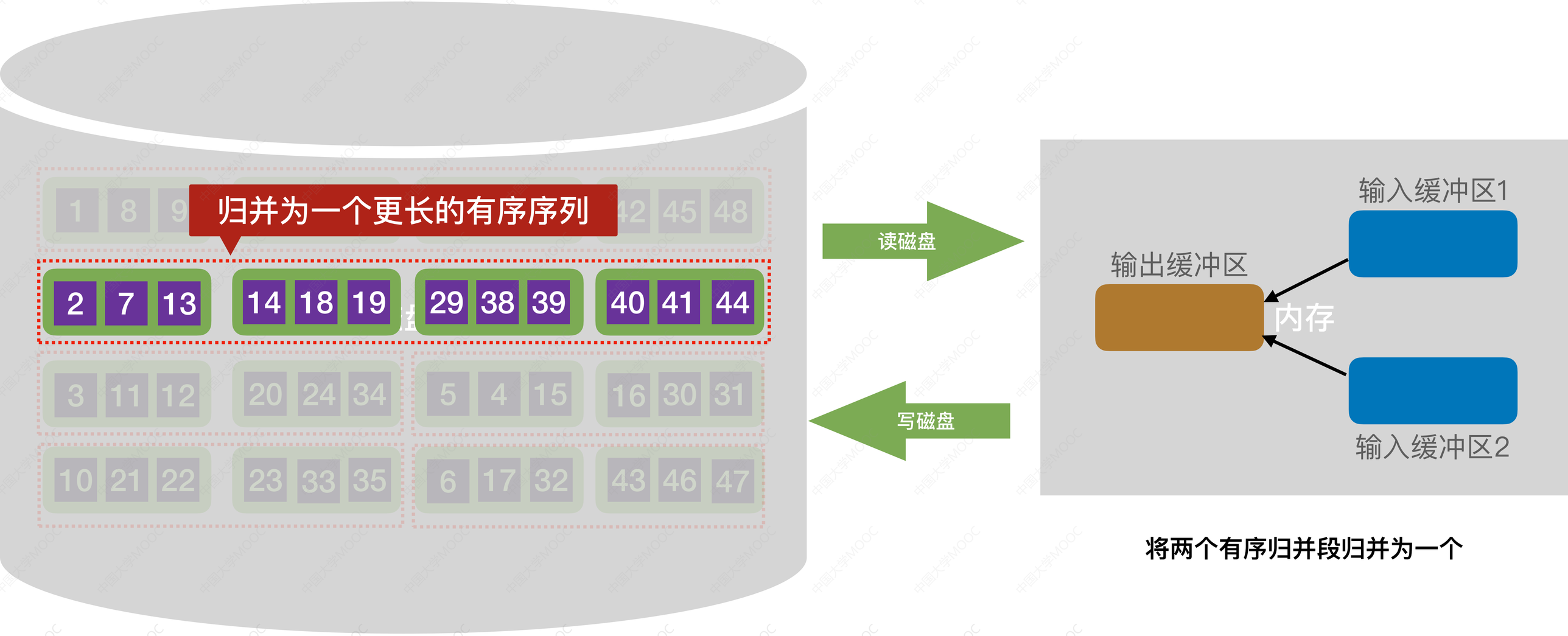
王道考研/CSKAOYAN.COM

第一趟归并



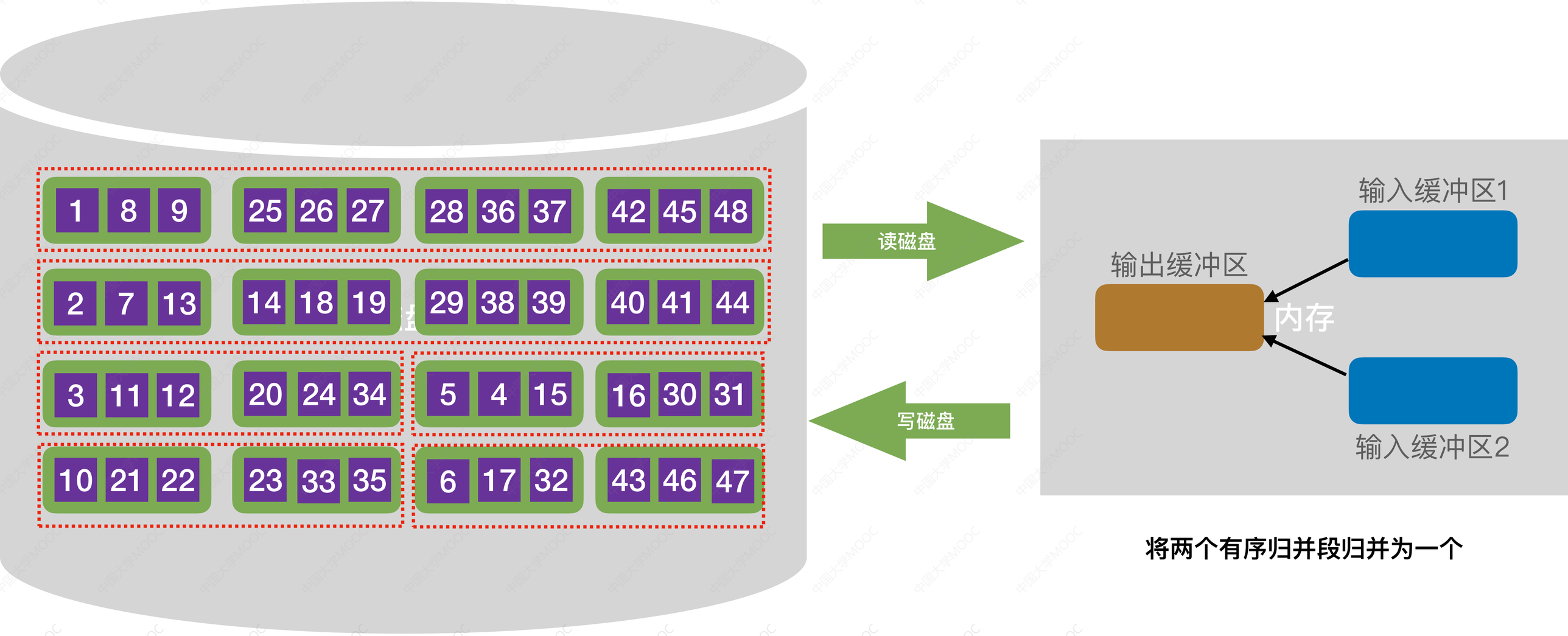
将两个有序归并段归并为一个

第一趟归并



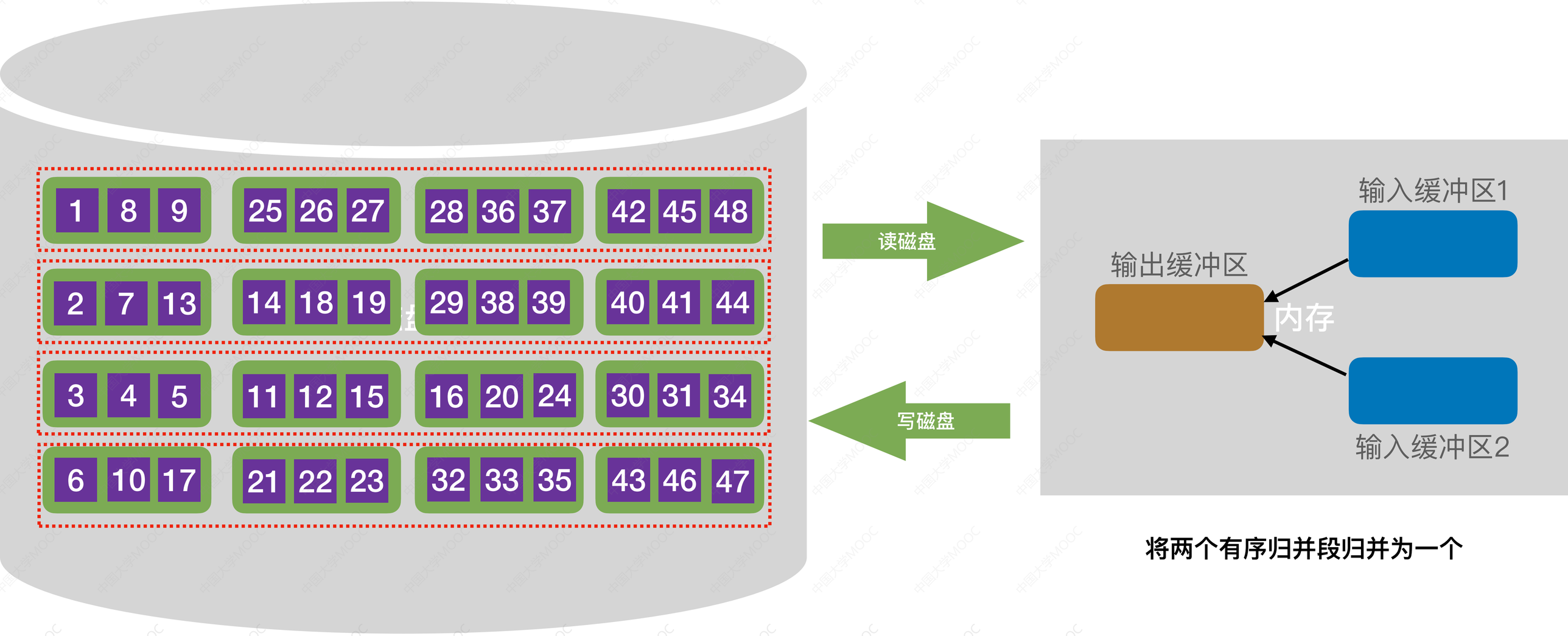
将两个有序归并段归并为一个

第一趟归并



将两个有序归并段归并为一个

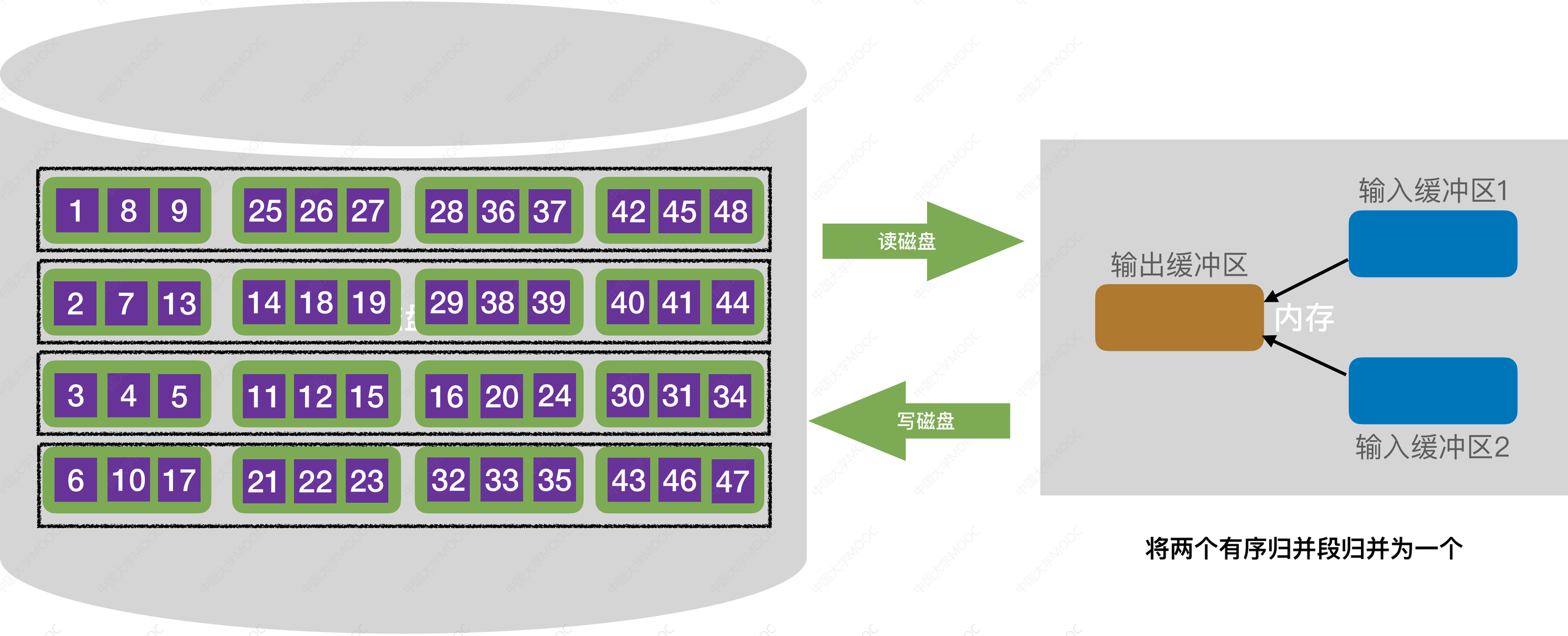
第一趟归并



将两个有序归并段归并为一个

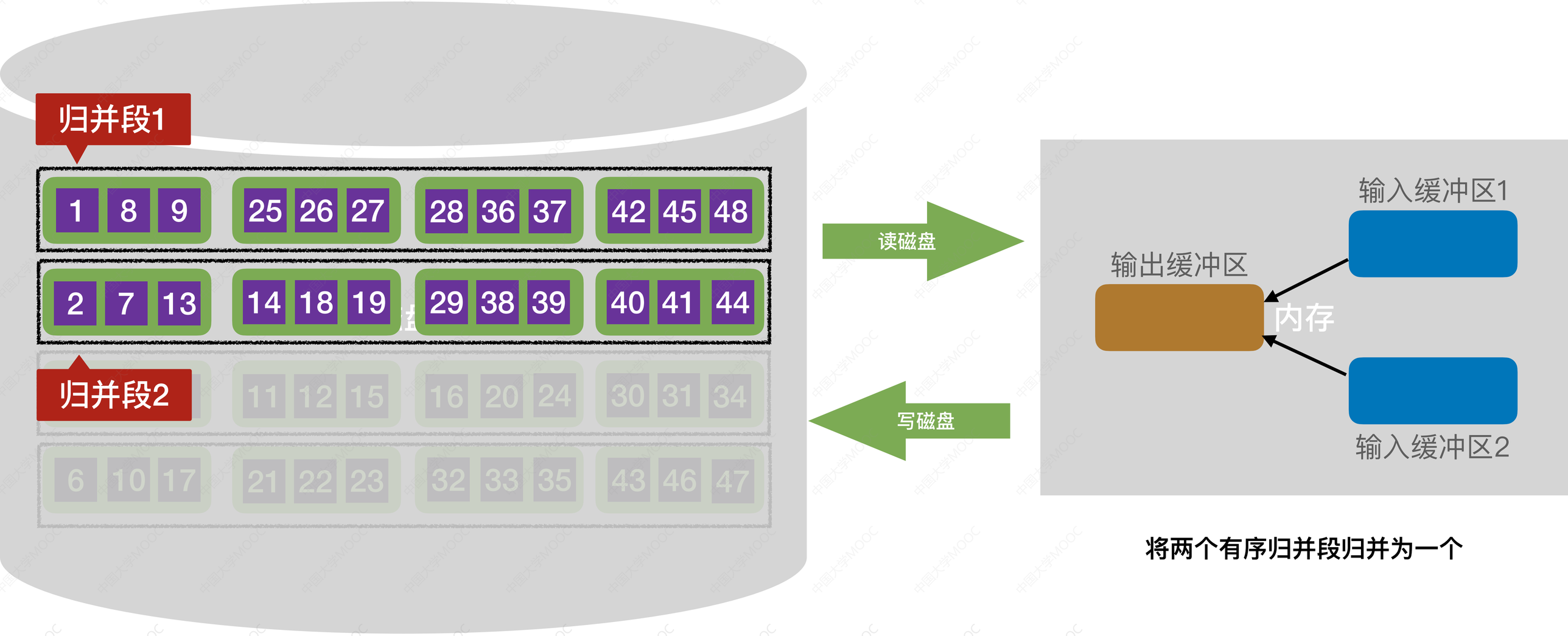
第二趟归并

把4个有序子序列（归并段）两两归并



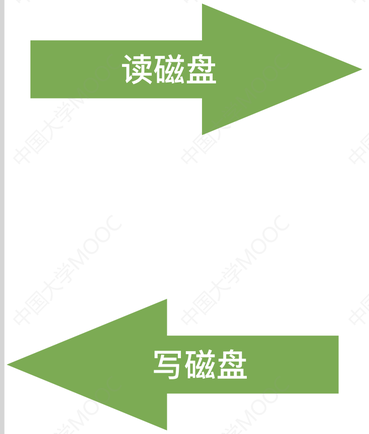
将两个有序归并段归并为一个

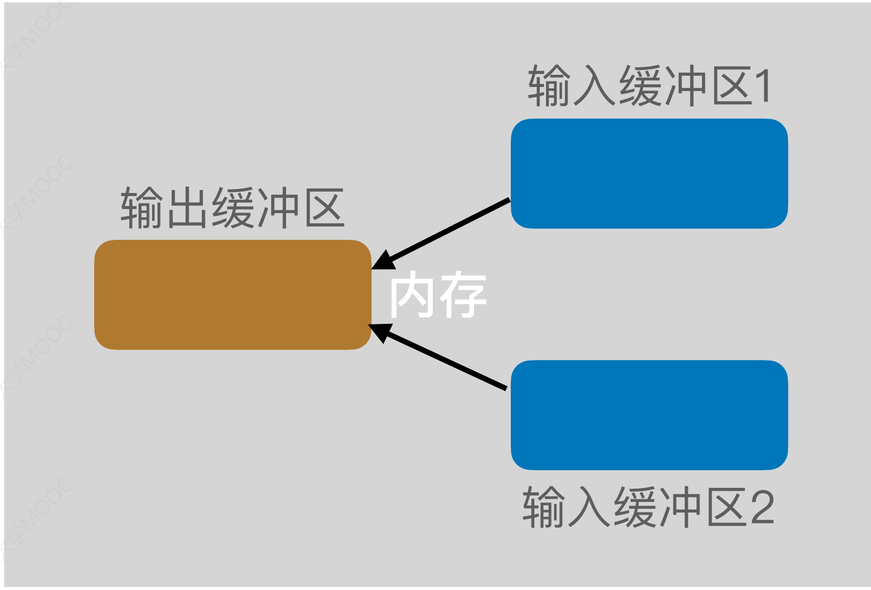
第二趟归并



将两个有序归并段归并为一个

第二趟归并

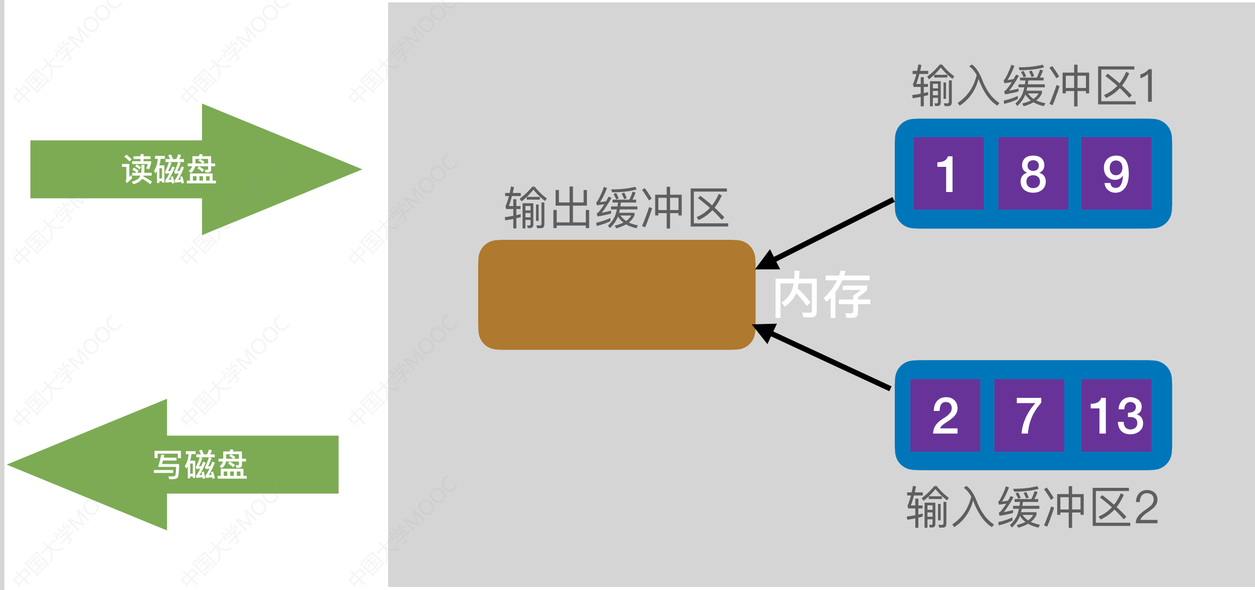




将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM

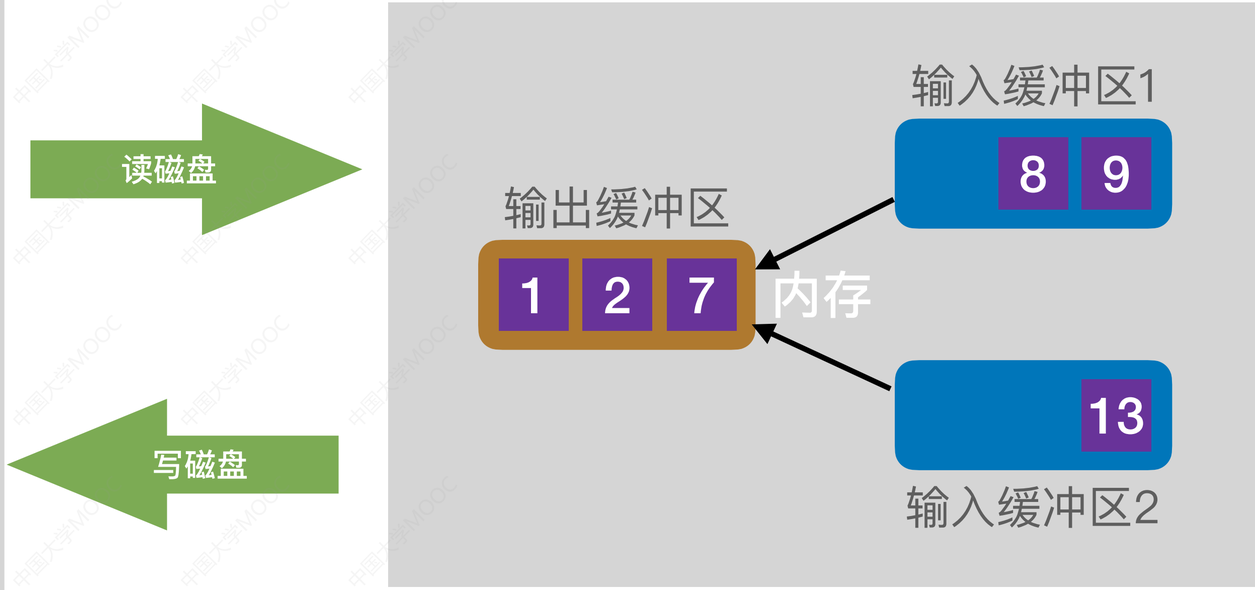
第二趟归并



将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM

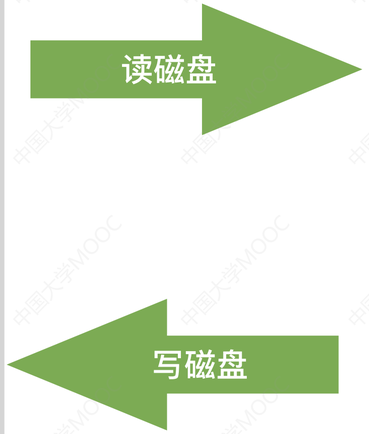
第二趟归并

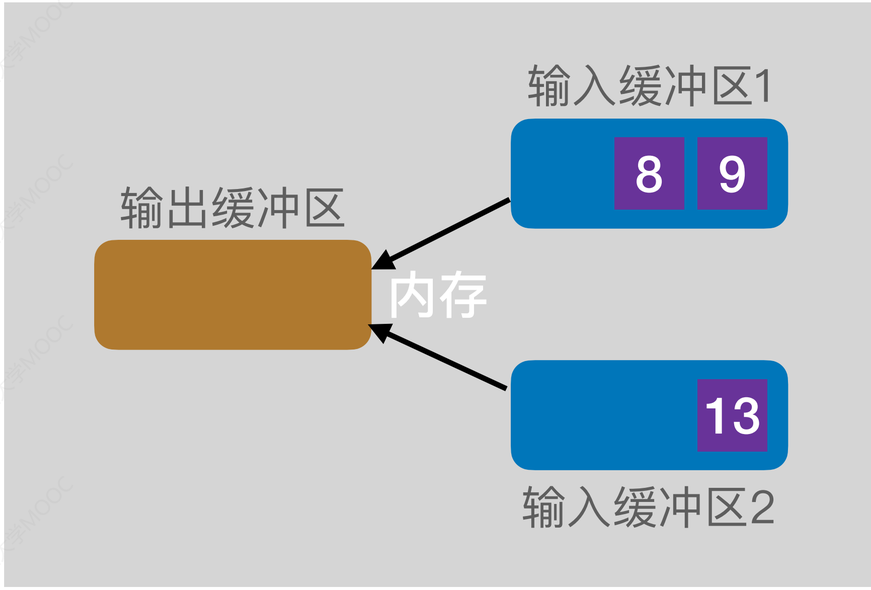


将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM

第二趟归并

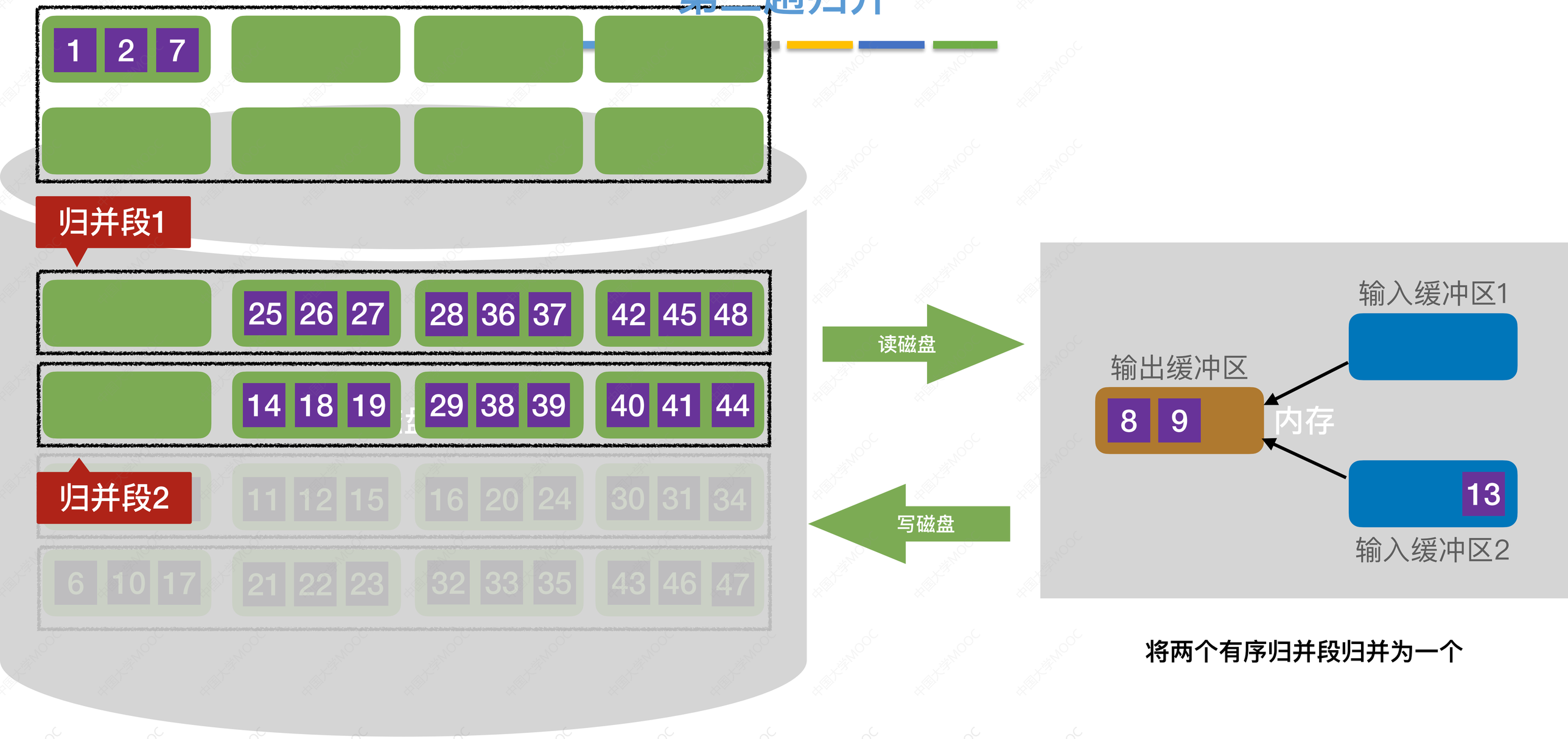




将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM

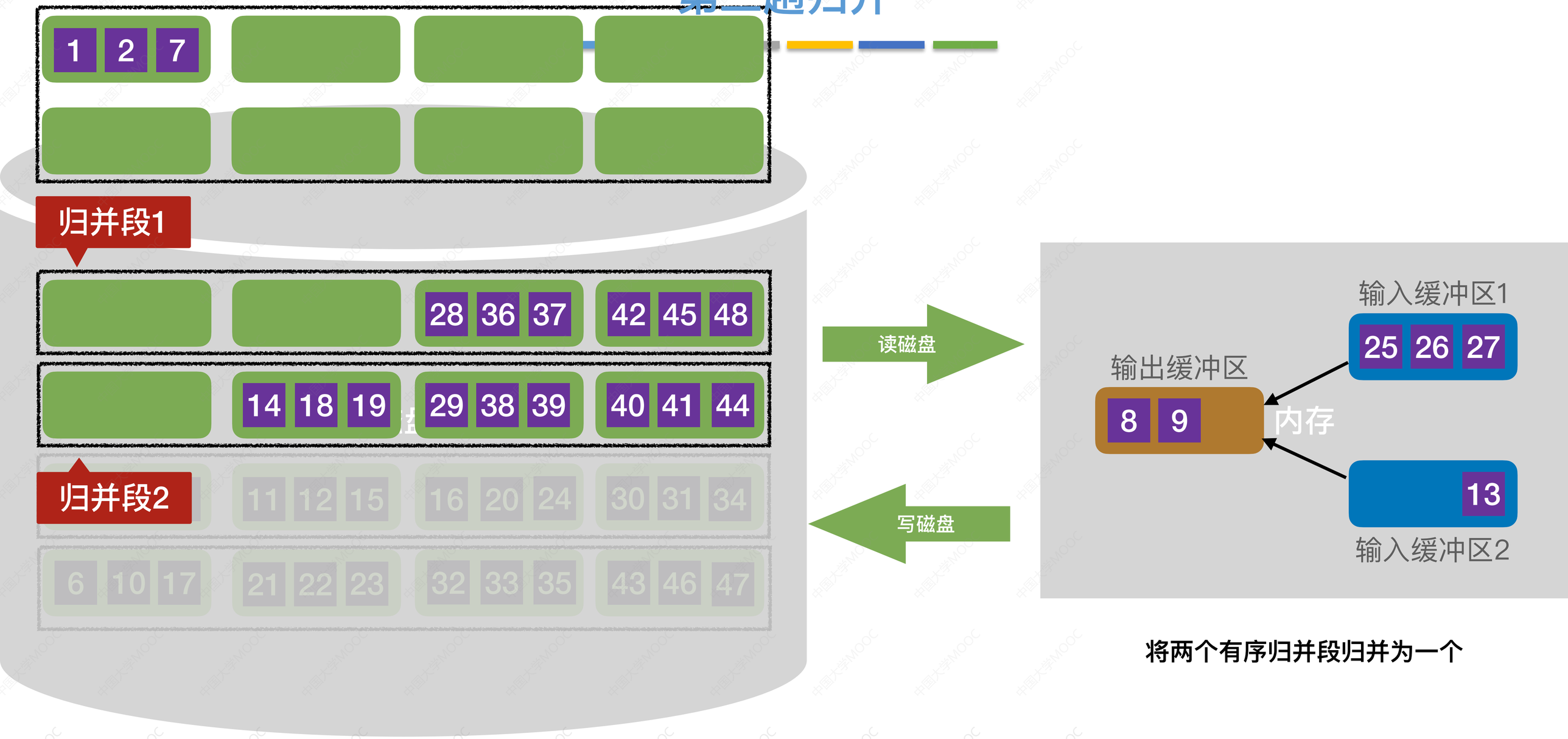
第二趟归并



将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM

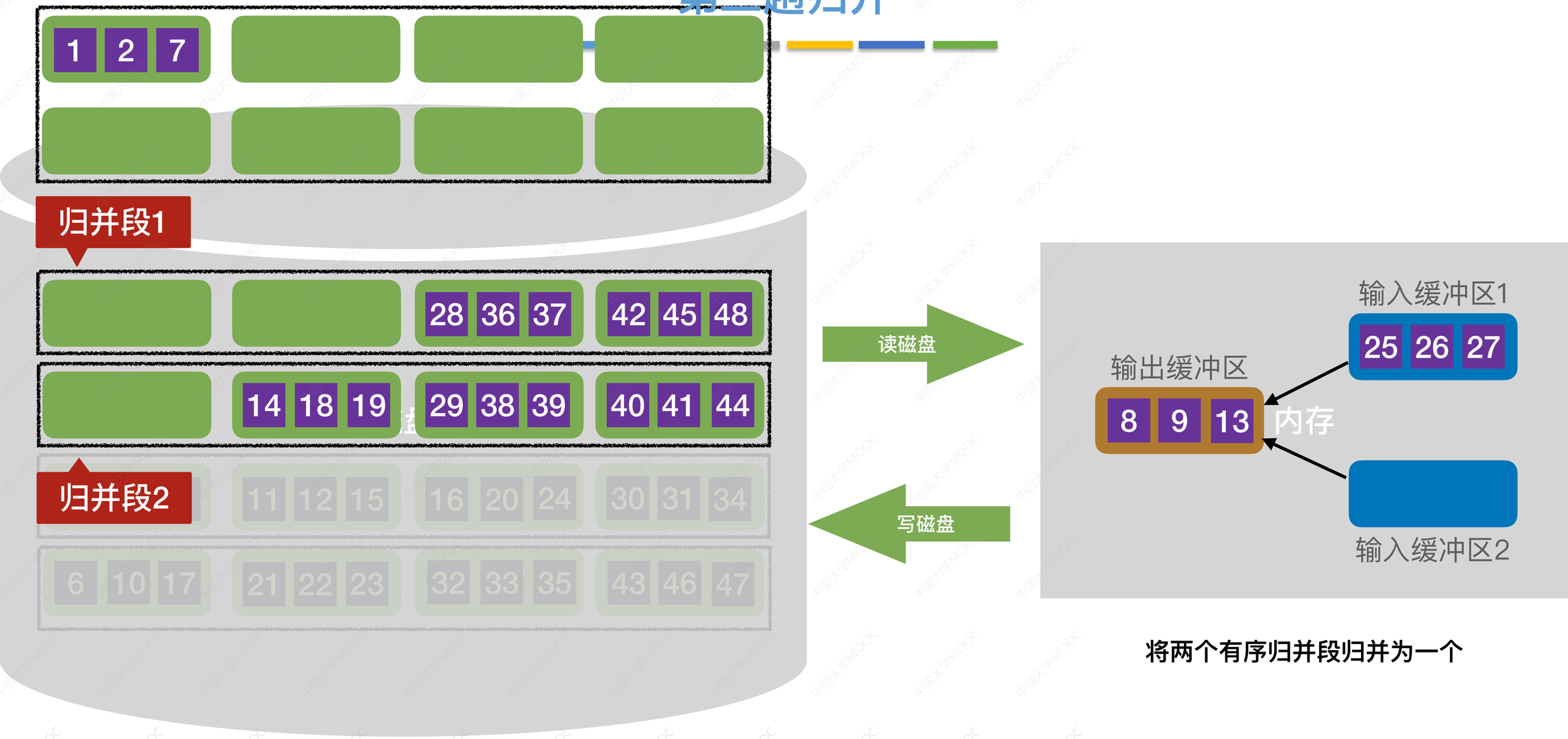
第二趟归并



将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM

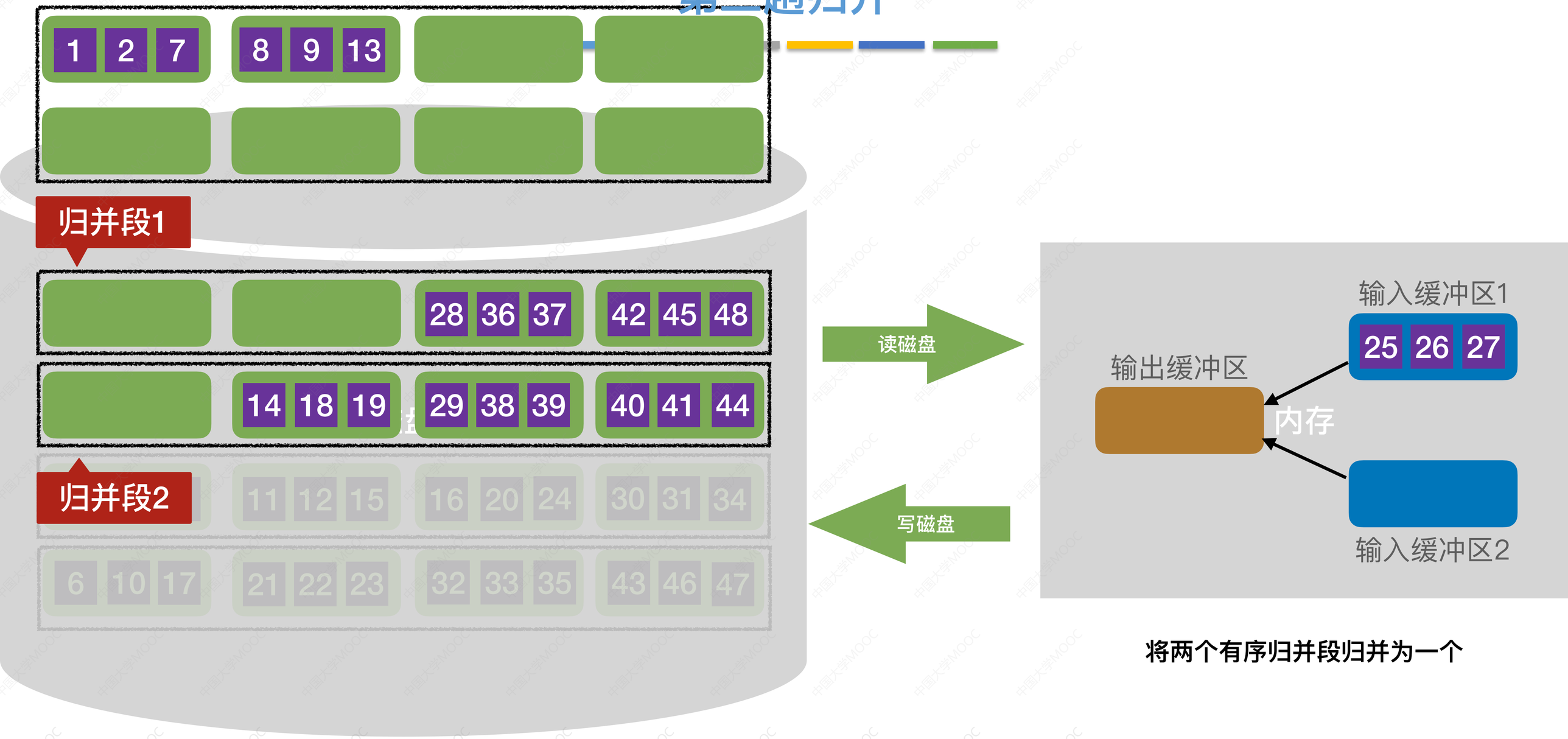
第二趟归并



将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM

第二趟归并



将两个有序归并段归并为一个

王道考研/CSKAOYAN.COM