

本节内容

# 置换-选择 排序

# 上上节知识回顾

若要进行k路归并排序，则需要在内存中分配k个输入缓冲区和1个输出缓冲区

## 外部排序

### 步骤

- ①生成  $r$  个初始归并段（对  $L$  个记录进行内部排序，组成一个有序的初始归并段）
- ②进行  $S$  趟k路归并， $S = \lceil \log_k r \rceil$

### 如何进行k路归并

- 把  $k$  个归并段的块读入  $k$  个输入缓冲区
- 用“归并排序”的方法从  $k$  个归并段中选出几个最小记录暂存到输出缓冲区中
- 当输出缓冲区满时，写出外存

### 外部排序时间开销

读写外存的时间 + 内部排序所需时间 + 内部归并所需时间

### 优化

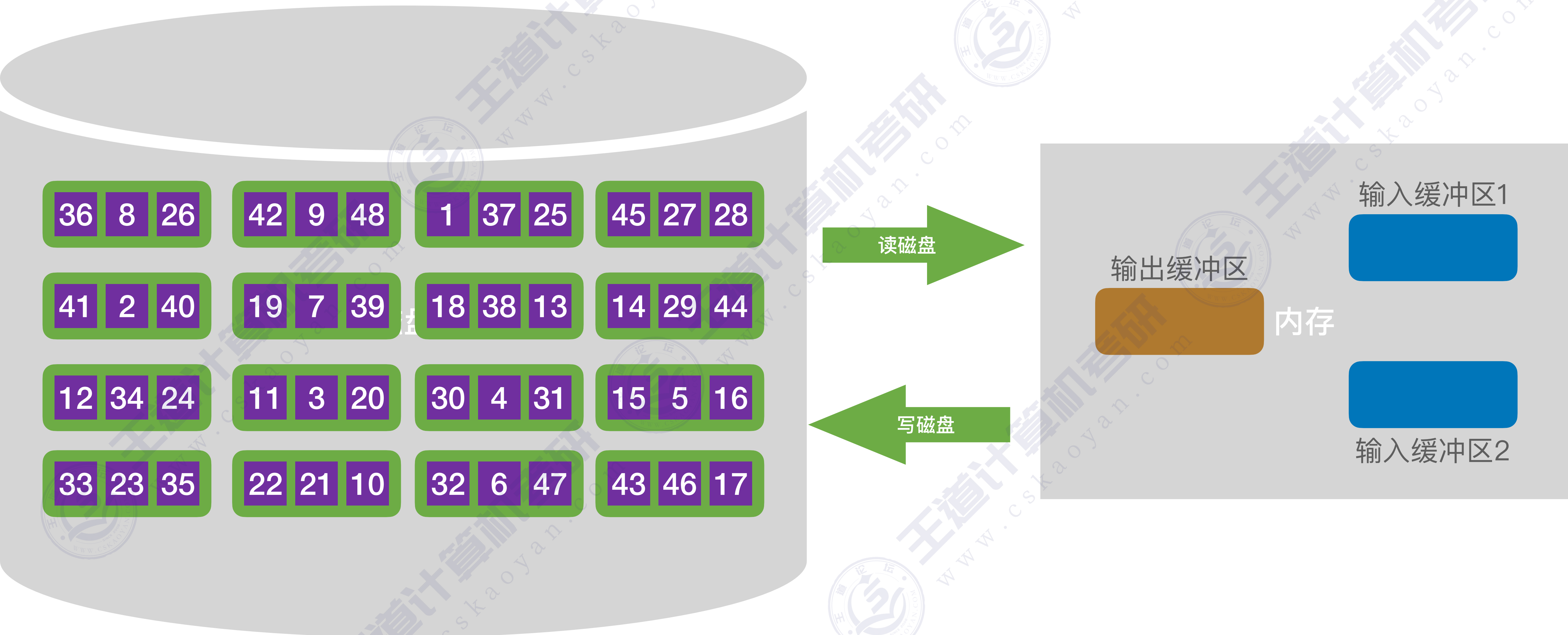
- 增加归并路数  $k$ ，进行多路平衡归并
  - 代价 1：需要增加相应的输入缓冲区
  - 代价 2：每次从  $k$  个归并段中选一个最小元素需要  $(k-1)$  次关键字对比
- 减少初始归并段数量  $r$

可用“败者树”减少关键字对比次数

可用“置换-选择排序”进一步减少初始归并段数量

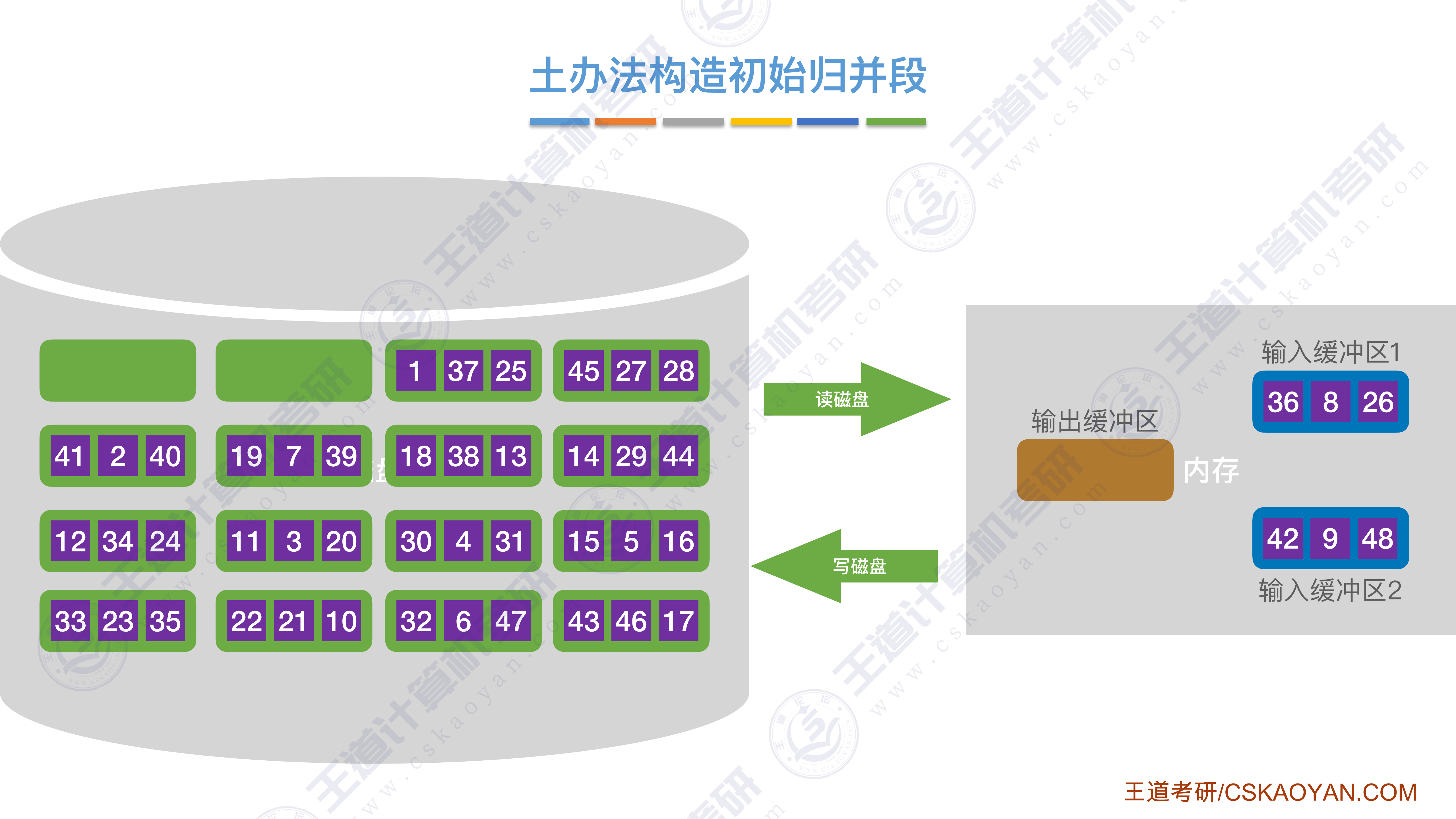
注：按照本节介绍的方法生成的初始归并段，若共  $N$  个记录，内存工作区可以容纳  $L$  个记录，则初始归并段数量  $r = \lceil N/L \rceil$

# 土办法构造初始归并段

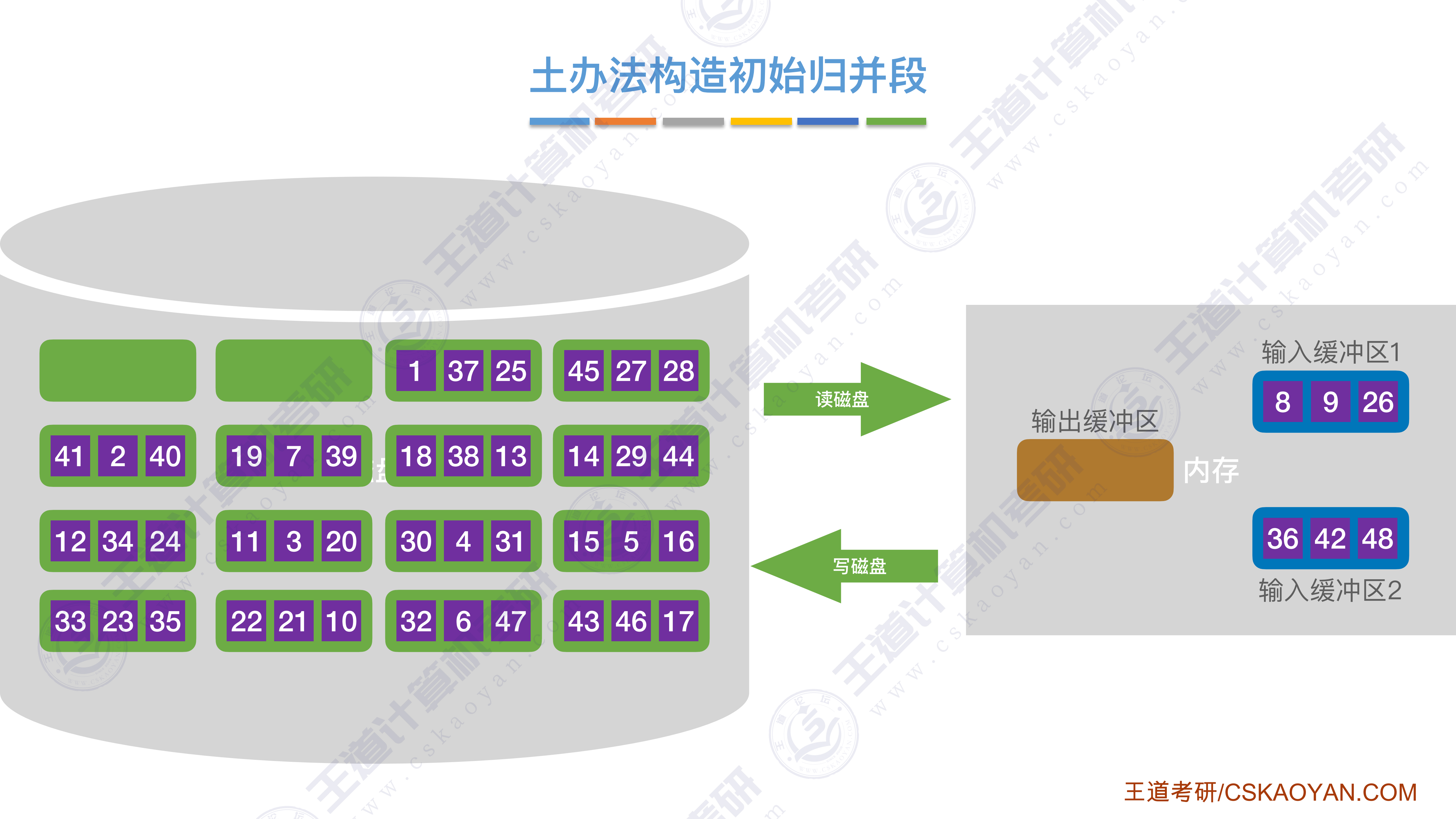




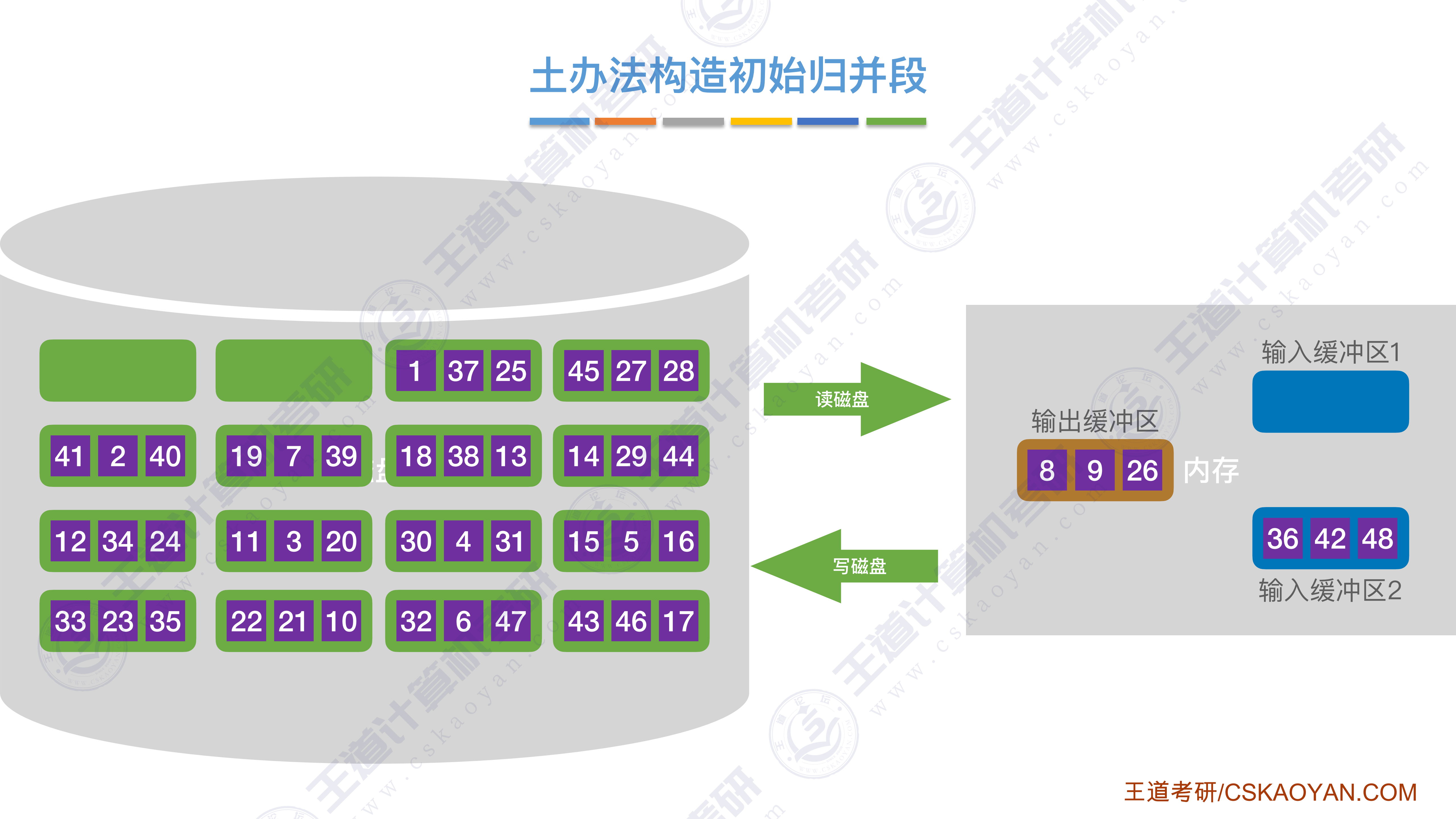
# 土办法构造初始归并段



# 土办法构造初始归并段

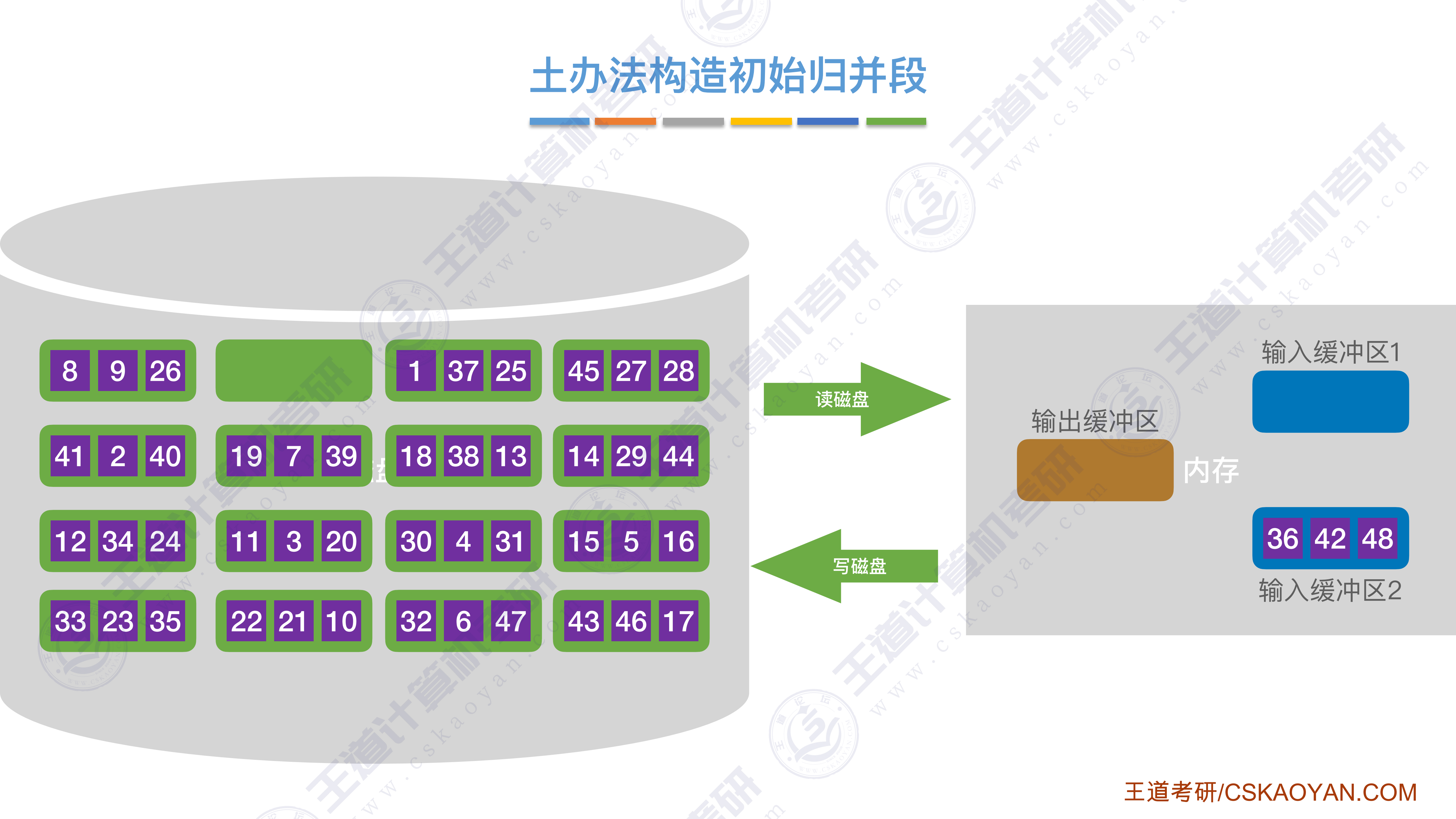


# 土办法构造初始归并段

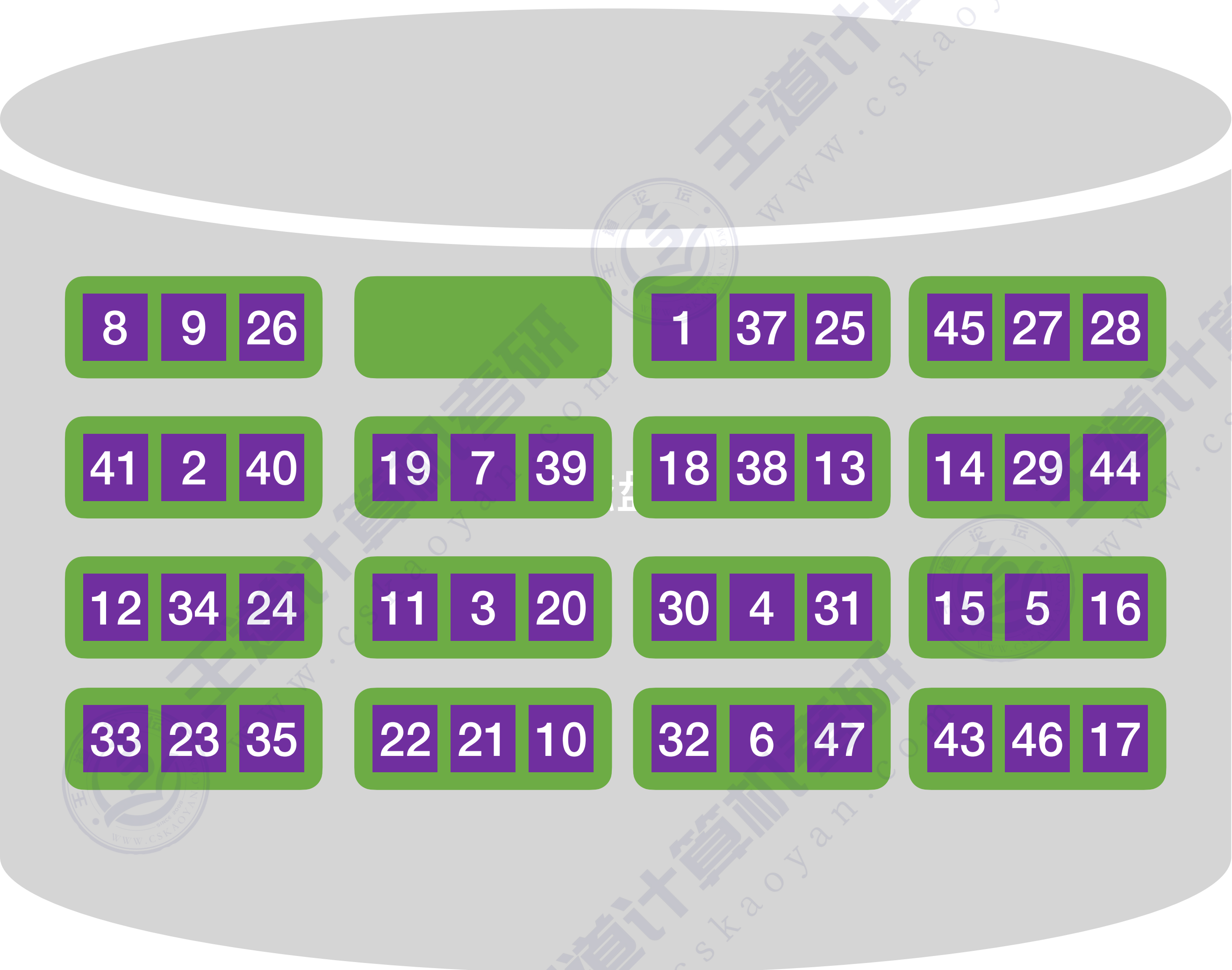




# 土办法构造初始归并段

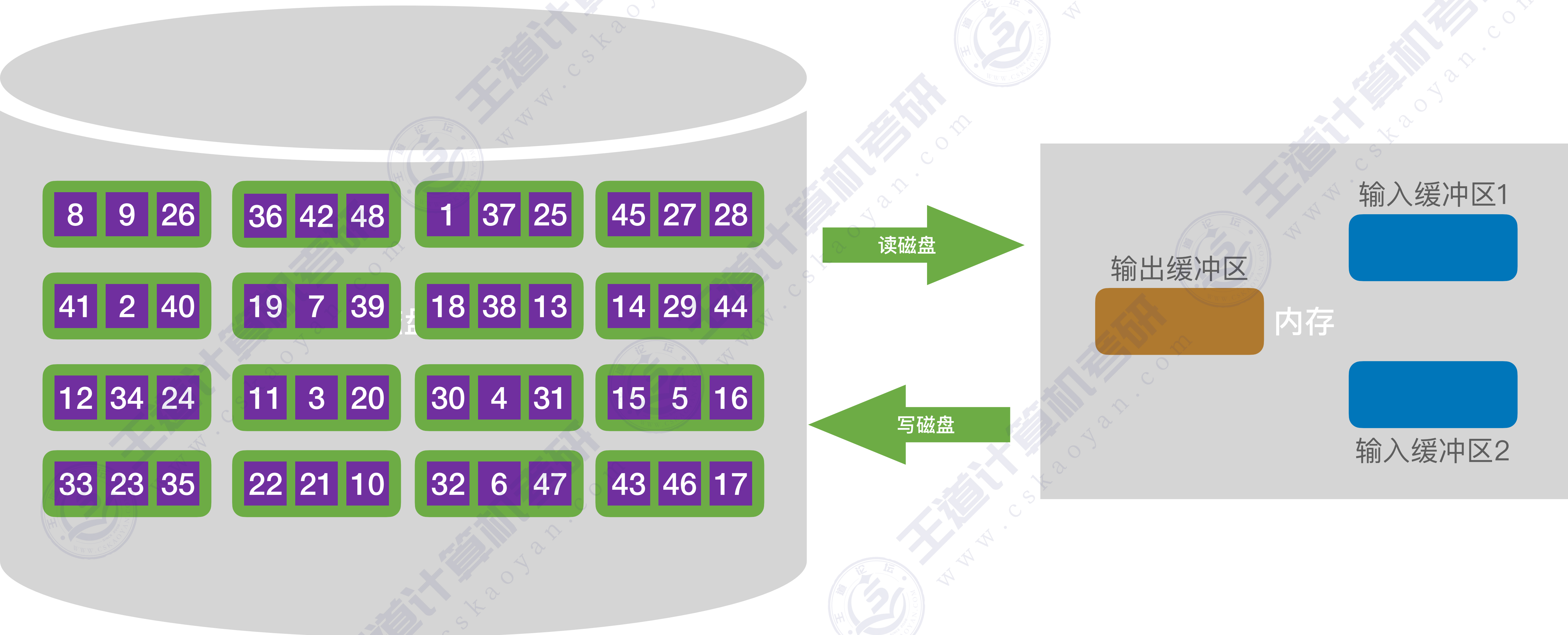


# 土办法构造初始归并段





# 土办法构造初始归并段



# 土办法构造初始归并段



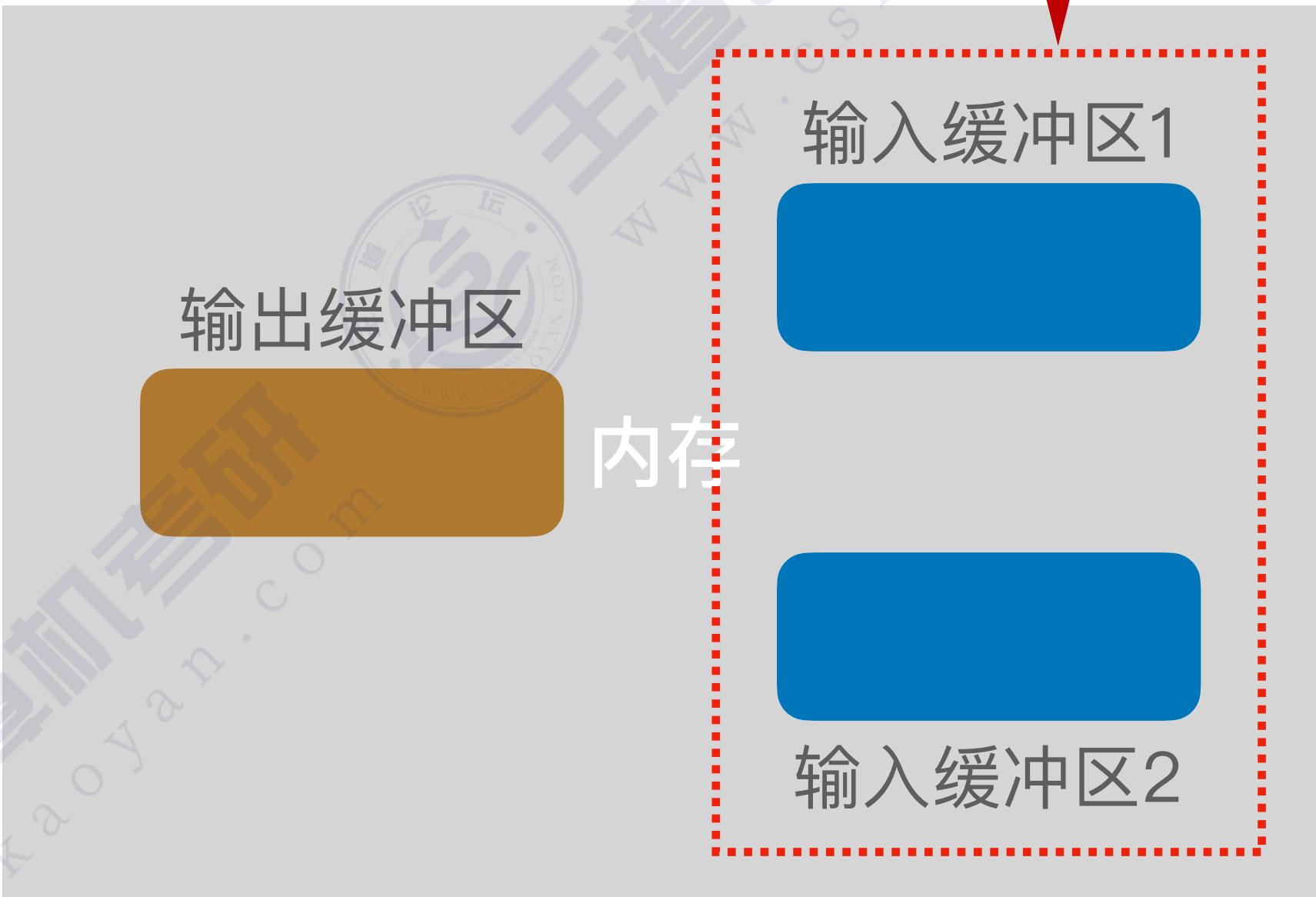
“初始归并段”中包含6个记录

8	9	26	36	42	48	1	37	25	45	27	28
41	2	40	19	7	39	18	38	13	14	29	44
12	34	24	11	3	20	30	4	31	15	5	16
33	23	35	22	21	10	32	6	47	43	46	17

读磁盘

写磁盘

用于内部排序的内存工作区只能容纳6个记录



# 土办法构造初始归并段

每个“初始归并段”  
可包含18个记录

可以用一片更大的内存区域来进行内部排序（如：可容纳18个记录）

36	8	26	42	9	48	1	37	25	45	27	28
41	2	40	19	7	39	18	38	13	14	29	44
12	34	24	11	3	20	30	4	31	15	5	16
33	23	35	22	21	10	32	6	47	43	46	17

读磁盘

写磁盘

内存  
输出缓冲区

输入缓冲区1  
输入缓冲区2

用于内部排序的内存工作区WA 可容纳  $l$  个记录，则每个初始归并段也只能包含  $l$  个记录，若文件共有  $n$  个记录，则初始归并段的数量  $r = n/l$



# 置换-选择排序



初始归并段输出文件 FO:



内存工作区WA

初始待排序文件 FI:

4	6	9	7	13	11
16	14	10	22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序

初始归并段输出文件 FO:

把最小的元素  
“置换”出去



内存工作区WA

初始待排序文件 FI:

			7	13	11
16	14	10	22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1: 4

MINIMAX= 4



内存工作区WA

初始待排序文件 FI:

			7	13	11
16	14	10	22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序



初始归并段输出文件 FO:

归并段1: 4

MINIMAX= 4



内存工作区WA

初始待排序文件 FI:

				13	11
16	14	10	22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

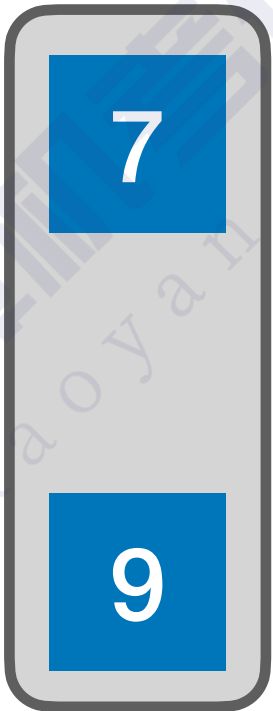
# 置换-选择排序

初始归并段输出文件 FO:

归并段1: 

4	6
---	---

MINIMAX= 6



内存工作区WA

初始待排序文件 FI:

				13	11
16	14	10	22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序

初始归并段输出文件 FO:

归并段1: 

4	6
---	---

MINIMAX= 6

7
13
9

内存工作区WA

初始待排序文件 FI:

					11
16	14	10	22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序

初始归并段输出文件 FO:

归并段1: 

4	6	7
---	---	---

MINIMAX= 7



内存工作区WA

初始待排序文件 FI:

					11
16	14	10	22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:

4	6	7
---	---	---

MINIMAX= 7

11
13
9

内存工作区WA

初始待排序文件 FI:

16	14	10	22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序

初始归并段输出文件 FO:

归并段1:

4	6	7	9
---	---	---	---

MINIMAX= 9

11
13

内存工作区WA

初始待排序文件 FI:

16	14	10	22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序

初始归并段输出文件 FO:

归并段1: 

4	6	7	9
---	---	---	---

MINIMAX= 9

11
13
16

内存工作区WA

初始待排序文件 FI:

	14	10	22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序

初始归并段输出文件 FO:

归并段1: 

4	6	7	9	11
---	---	---	---	----

MINIMAX=11



内存工作区WA

初始待排序文件 FI:

	14	10	22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:

4	6	7	9	11
---	---	---	---	----

MINIMAX=11

14
13
16

内存工作区WA

初始待排序文件 FI:

		10	22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序

初始归并段输出文件 FO:

归并段1: 4 6 7 9 11 13

MINIMAX=13



内存工作区WA

初始待排序文件 FI:

		10	22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序

初始归并段输出文件 FO:

归并段1: 

4	6	7	9	11	13
---	---	---	---	----	----

MINIMAX=13

14
10
16

内存工作区WA

初始待排序文件 FI:

			22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:

4	6	7	9	11	13
---	---	---	---	----	----

MINIMAX=13

不可能放到归并段1的末尾

14
10
16

内存工作区WA

初始待排序文件 FI:

			22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序



初始归并段输出文件 FO:

归并段1:

4	6	7	9	11	13	14
---	---	---	---	----	----	----

MINIMAX=14



内存工作区WA

初始待排序文件 FI:

			22	30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:

4	6	7	9	11	13	14
---	---	---	---	----	----	----

MINIMAX=14

22
10
16

内存工作区WA

初始待排序文件 FI:

				30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:

4	6	7	9	11	13	14	16
---	---	---	---	----	----	----	----

MINIMAX=16



内存工作区WA

初始待排序文件 FI:

				30	2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序



初始归并段输出文件 FO:

归并段1:

4	6	7	9	11	13	14	16
---	---	---	---	----	----	----	----

MINIMAX=16

22
10
30

内存工作区WA

初始待排序文件 FI:

					2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序

初始归并段输出文件 FO:

归并段1: 

4	6	7	9	11	13	14	16	22
---	---	---	---	----	----	----	----	----

MINIMAX=22



内存工作区WA

初始待排序文件 FI:

					2
3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1: 

4	6	7	9	11	13	14	16	22
---	---	---	---	----	----	----	----	----

MINIMAX=22



内存工作区WA

初始待排序文件 FI:

3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序



初始归并段输出文件 FO:

归并段1:

4	6	7	9	11	13	14	16	22
---	---	---	---	----	----	----	----	----

MINIMAX=22

2
10
30

内存工作区WA

初始待排序文件 FI:

3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:

4	6	7	9	11	13	14	16	22	30
---	---	---	---	----	----	----	----	----	----

MINIMAX=30



内存工作区WA

初始待排序文件 FI:

3	19	20	17	1	23
5	36	12	18	21	39

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:



MINIMAX=30



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序

初始归并段输出文件 FO:

归并段1:



若WA内的关键字都比 MINIMAX 更小, 则该归并段在此截止

MINIMAX=30



内存工作区WA

初始待排序文件 FI:



注: 假设用于内部排序的内存工作区只能容纳3个记录

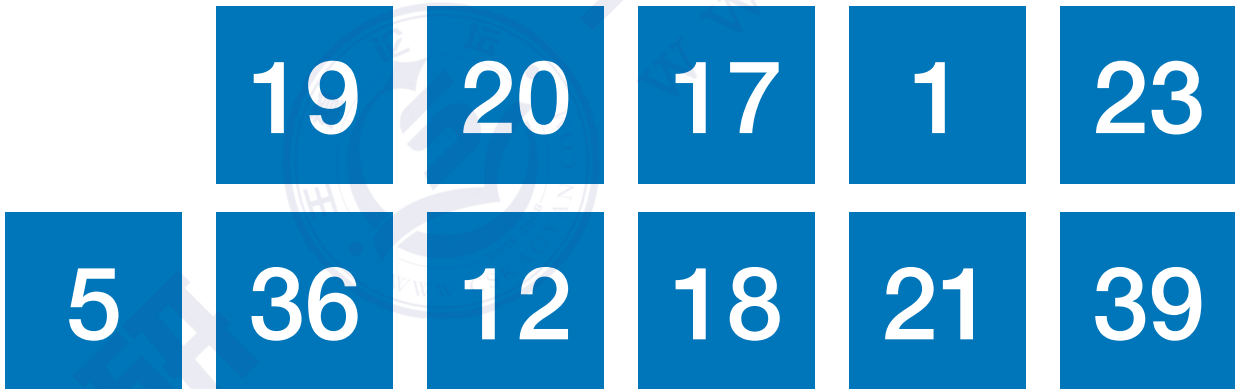
# 置换-选择排序

初始归并段输出文件 FO:



归并段2:

初始待排序文件 FI:



内存工作区WA

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序

初始归并段输出文件 FO:

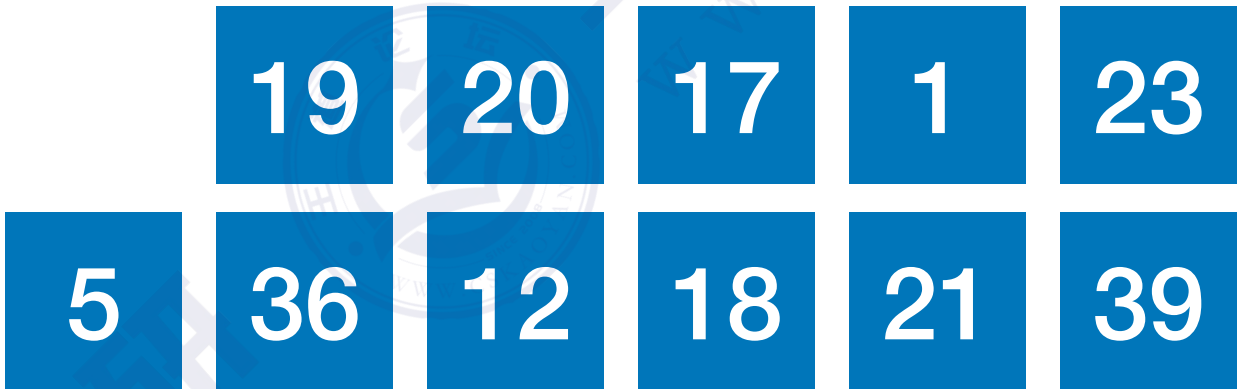


MINIMAX= 2



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序



初始归并段输出文件 FO:

归并段1:



归并段2:



MINIMAX= 2



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:



归并段2:



MINIMAX= 3



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序

初始归并段输出文件 FO:

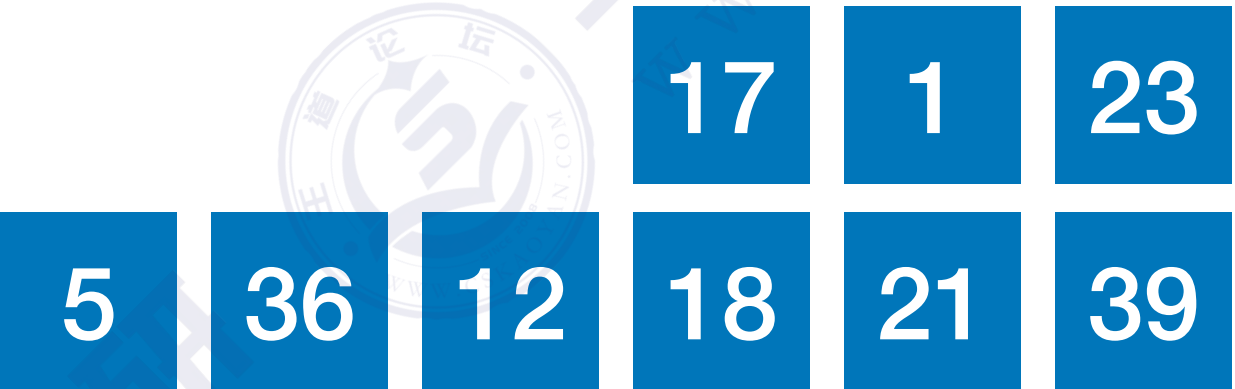


MINIMAX= 3



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序

初始归并段输出文件 FO:



MINIMAX=10



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:



归并段2:



MINIMAX=10



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:



归并段2:



MINIMAX=17



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序



初始归并段输出文件 FO:

归并段1:



归并段2:



MINIMAX=17



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:



归并段2:



MINIMAX=17



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:



归并段2:



MINIMAX=19



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序

初始归并段输出文件 FO:



MINIMAX=19



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:



归并段2:



MINIMAX=20



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:



归并段2:



MINIMAX=20

初始待排序文件 FI:



内存工作区WA

注：假设用于内部排序的内存工作区只能容纳3个记录

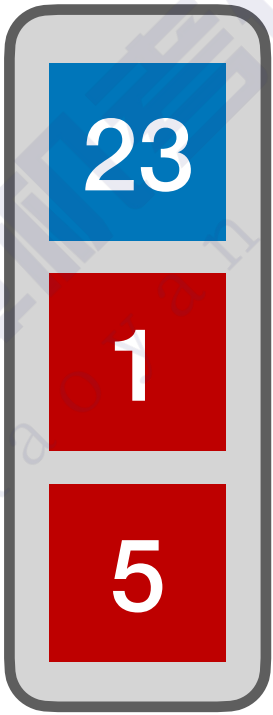


# 置换-选择排序

初始归并段输出文件 FO:



MINIMAX=20



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:



归并段2:



MINIMAX=23

初始待排序文件 FI:



内存工作区WA

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:



归并段2:



MINIMAX=23



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序



初始归并段输出文件 FO:

初始待排序文件 FI:

归并段1:



归并段2:



MINIMAX=36



内存工作区WA



注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

初始待排序文件 FI:

归并段1:



归并段2:



MINIMAX=36



内存工作区WA



注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1:



归并段2:



MINIMAX=36



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序



初始归并段输出文件 FO:

初始待排序文件 FI:

归并段1:



归并段2:



归并段3:



内存工作区WA



注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序

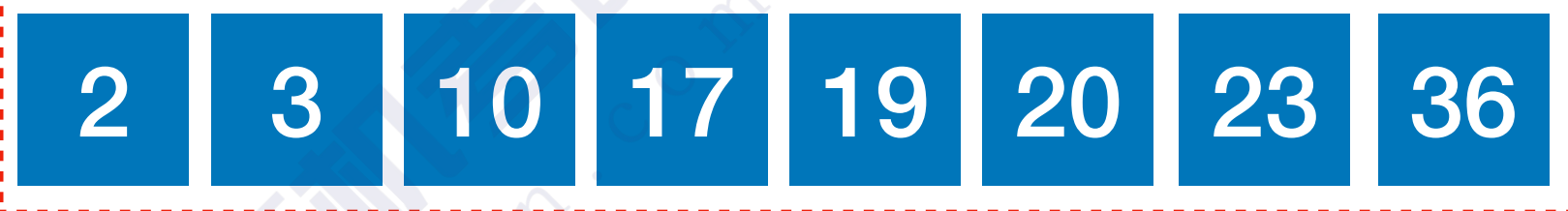


初始归并段输出文件 FO:

归并段1:



归并段2:



归并段3:



MINIMAX= 1

初始待排序文件 FI:



内存工作区WA

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序

初始归并段输出文件 FO:

归并段1: 4 6 7 9 11 13 14 16 22 30

归并段2: 2 3 10 17 19 20 23 36

归并段3: 1

初始待排序文件 FI:

21 39

MINIMAX= 1

12  
18  
5

内存工作区WA

注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序

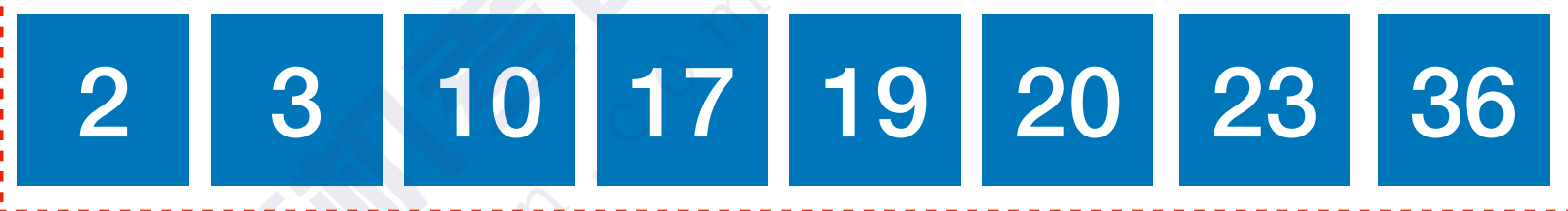


初始归并段输出文件 FO:

归并段1:



归并段2:



归并段3:



MINIMAX= 5



内存工作区WA

初始待排序文件 FI:



注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

归并段1: 

4	6	7	9	11	13	14	16	22	30
---	---	---	---	----	----	----	----	----	----

归并段2: 

2	3	10	17	19	20	23	36
---	---	----	----	----	----	----	----

归并段3: 

1	5
---	---

MINIMAX= 5

12
18
21

内存工作区WA

初始待排序文件 FI:

39
----

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

初始待排序文件 FI:

归并段1:



归并段2:



归并段3:



MINIMAX=12



内存工作区WA

39

注：假设用于内部排序的内存工作区只能容纳3个记录



# 置换-选择排序

初始归并段输出文件 FO:

归并段1: 

4	6	7	9	11	13	14	16	22	30
---	---	---	---	----	----	----	----	----	----

归并段2: 

2	3	10	17	19	20	23	36
---	---	----	----	----	----	----	----

归并段3: 

1	5	12
---	---	----

MINIMAX=12

39
18
21

内存工作区WA

初始待排序文件 FI:

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

初始待排序文件 FI:

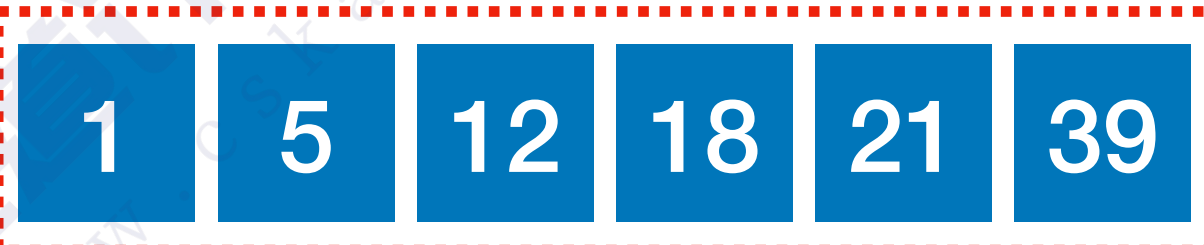
归并段1:



归并段2:



归并段3:



内存工作区WA

注：假设用于内部排序的内存工作区只能容纳3个记录

# 置换-选择排序



初始归并段输出文件 FO:

初始待排序文件 FI:

归并段1:



归并段2:



归并段3:



内存工作区WA

使用置换-选择排序，可以让每个初始归并段的长度超越内存工作区大小的限制

注：假设用于内部排序的内存工作区只能容纳3个记录



## 知识回顾与重要考点

设初始待排文件为FI，初始归并段输出文件为FO，内存工作区为WA，FO和WA的初始状态为空，WA可容纳 $w$ 个记录。置换-选择算法的步骤如下：

- 1) 从FI输入 $w$ 个记录到工作区WA。
- 2) 从WA中选出其中关键字取最小值的记录，记为MINIMAX记录。
- 3) 将MINIMAX记录输出到FO中去。
- 4) 若FI不空，则从FI输入下一个记录到WA中。
- 5) 从WA中所有关键字比MINIMAX记录的关键字大的记录中选出最小关键字记录，作为新的MINIMAX记录。
- 6) 重复3) ~5)，直至在WA中选不出新的MINIMAX记录为止，由此得到一个初始归并段，输出一个归并段的结束标志到FO中去。
- 7) 重复2) ~6)，直至WA为空。由此得到全部初始归并段。