



华中科技大学

# 数据结构

## 第11章 外排序

主讲教师：祝建华



详见：[网学天地](http://www.e-studysky.com) ([www.e-studysky.com](http://www.e-studysky.com))；咨询QQ：2696670126

## 外部排序思想

设磁盘有4500个记录的文件，磁盘的读 / 写单位是250个记录的数据块，内存只能提供1500个记录的空间，试对此文件进行排序。

解： i (1) 从磁盘文件输入三个数据块，计750个记录至内存，排序后再写入磁盘；

(2) 按(1)同样的方法得如下归并段：

R1

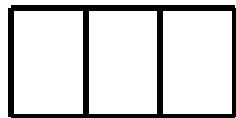
R2

R3

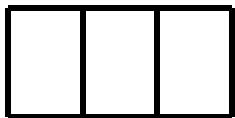
R4

R5

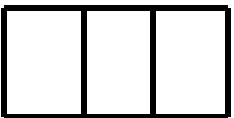
R6



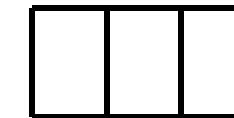
1~750



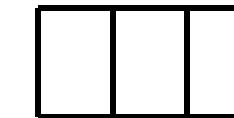
751~1500



1501~2250



2251~3000



3001~3750

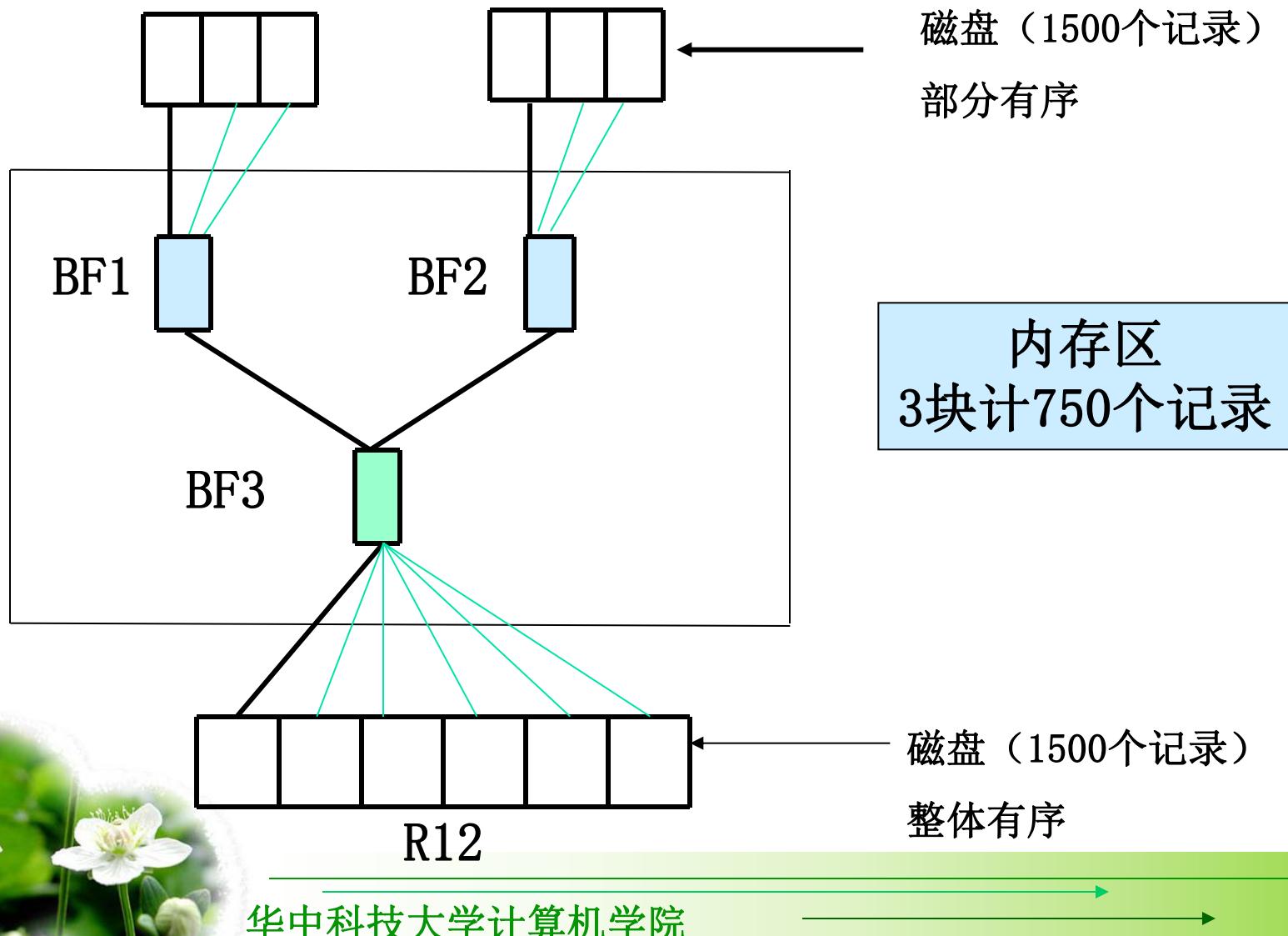


3751 ~ 4500



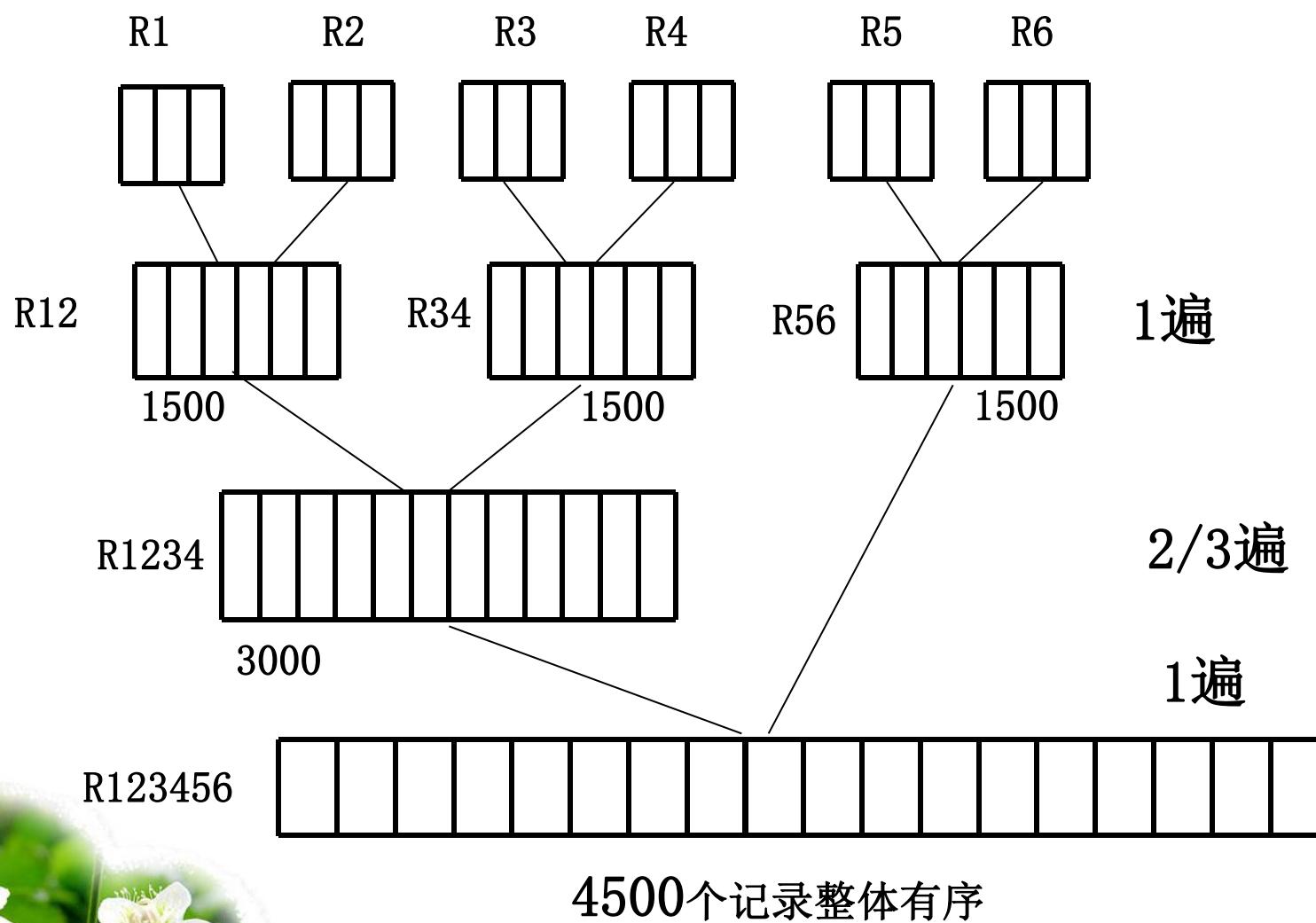


见(1)网学天地([www.e-studysky.com](http://www.e-studysky.com))；咨询QQ：2696670126  
R1 R2



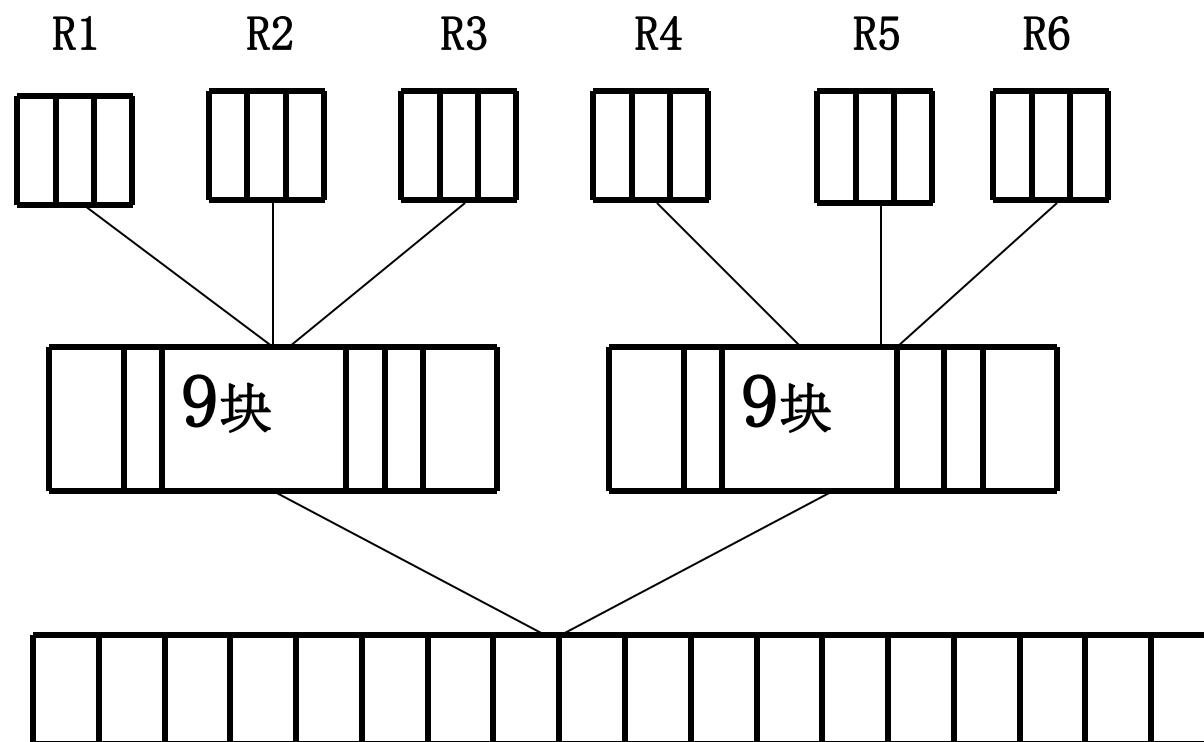


ii 详(2) 按(1)同样得方法得: R34, R56, R1234, R123456





详见：[网学天地](http://www.e-studysky.com) ([www.e-studysky.com](http://www.e-studysky.com))；咨询QQ：2696670126 建立初始归并段要把全部记录（4500）读写一遍，完成全部记录的归并要把（4500）个记录读写8/3遍。下面分析三路归并的情况：



完成全部记录的归并要把（4500）个记录读写2遍。



详见：网学天地（[www.e-studysky.com](http://www.e-studysky.com)）；咨询QQ：2696670126

结论：M个初始归并段进行K路归并时，归并的趟数为：

$$s = \log_K M$$

证明：设M为K的整幂次，即  $M = K^s$  ( $s$ 为整数)

第一趟归并得  $k^{s-1}$ 个为归并段；

第二趟归并得  $k^{s-2}$ 个为归并段；

...

第s趟归并得  $k^{s-s}=1$ 个归并段；

由  $M = K^s$ ，有  $\log_K M = s$

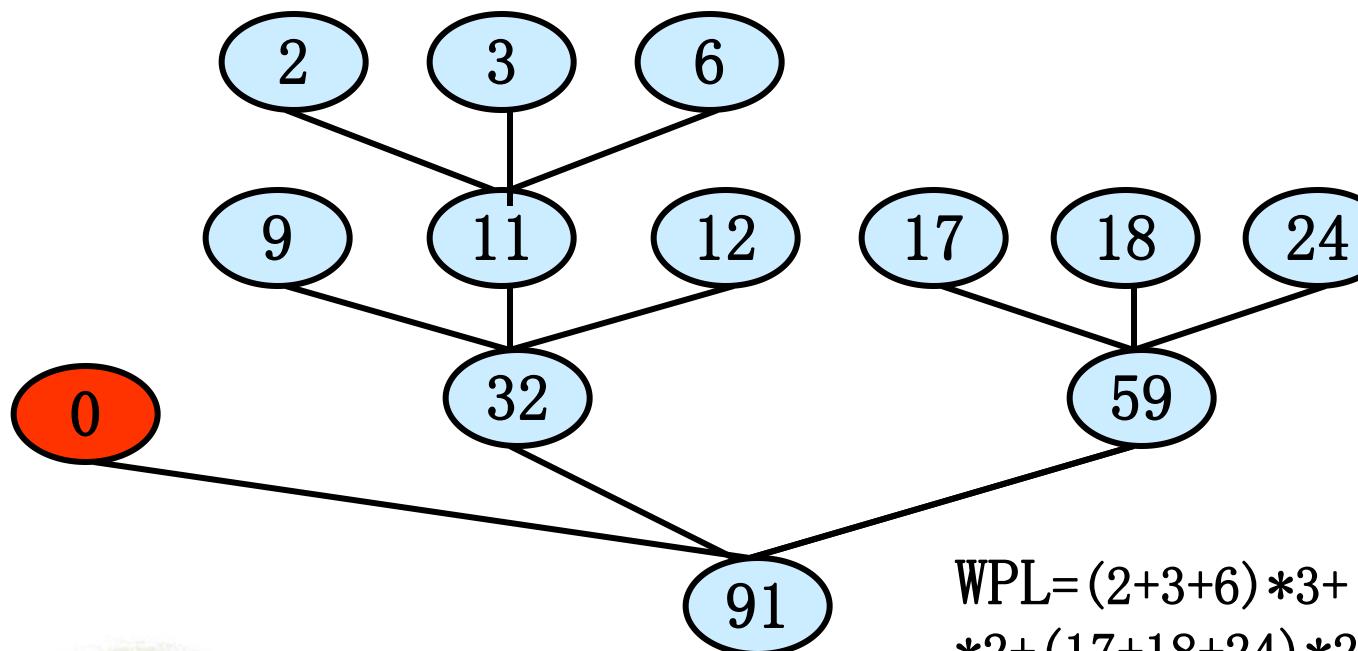
当m不是K的整幂次时， $K^{s-1} < m \leq K^s$  归并趟数是  
 $\log_K m$  的上界函数  $\log_K M$ 。证毕

## 最佳归并树

详见：[网学天地](http://www.e-studysky.com) ([www.e-studysky.com](http://www.e-studysky.com))；咨询QQ：2696670126

设有8个初始归并段，其长度（以外存物理块为单位）分别为2, 3, 6, 9, 12, 17, 18, 24，求其3路最佳归并树，并求其WPL和对外存的访问次数。

解1：



$$\begin{aligned} WPL = & (2+3+6)*3 + (9+12) \\ & *2 + (17+18+24)*2 = 193 \end{aligned}$$

$$\text{读写次数} = WPL * 2 = 386$$

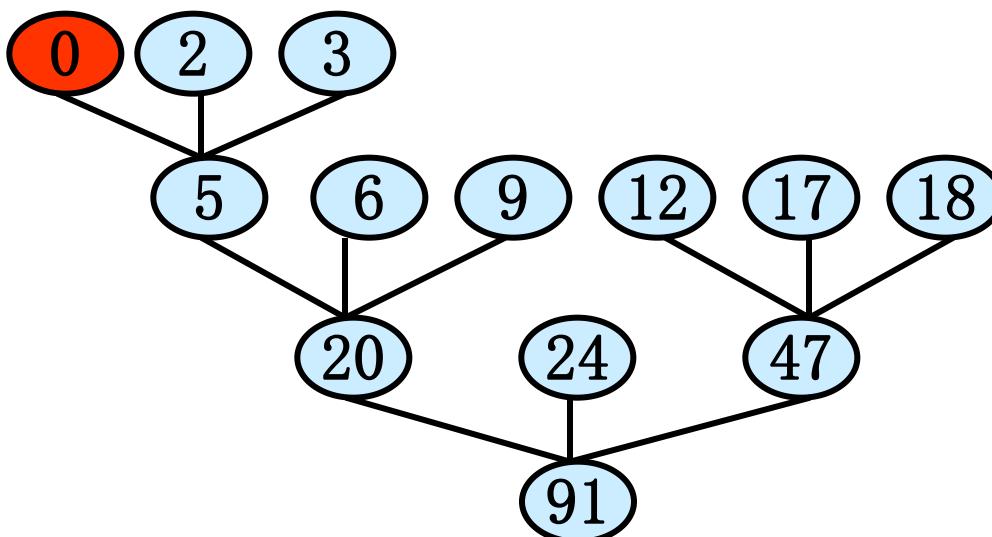
权值愈小的结点，离根愈近？



解

解2：网学天地 ([www.e-study.sky.com](http://www.e-study.sky.com)) 咨询QQ:3-1686610126  
 $(m-1) \bmod (k-1)$   $\equiv 3-1 \bmod (3-1)-1=1$   
 $\therefore$  虚设一个长度为“0”的归并段。

权值愈小的结点，离根愈远！



$$WPL = (0+2+3)*3 + (6+9+12+17+18)*2 + 24*1 = 163$$

$$\text{读写次数} = WPL * 2 = 326$$



详见：网学天地（[www.e-studysky.com](http://www.e-studysky.com)）；咨询QQ：2696670126

## 如何判定虚设段的个数？

令  $M$  —— 初始归并段的个数

$K$  —— 归并的路数

若  $(M-1) \bmod (K-1) = 0$ ，则不加虚设段；  
否则需加  $P$  个虚设段。

$$P = K-1 - (M-1) \bmod (K-1)$$

