

1.假如我有 100亿条数据，但是我们的内存只有1M，但是我们磁盘很大，

1.假如我有 100亿条数据，但是我们的内存只有 1M，但是我们磁盘很大

我们现在要对这100亿条数据进行排序，是没法把所有的数据一次性的load进行内存进行排序的，这就涉及到一个外部排序的问题，我们的1M内存只能装进1亿条数据，每次都只能对这 1亿条数据进行排序，排好序后输出到磁盘，总共输出100个文件，最后怎么把这100个文件进行merge成一个全局有序的大文件。

我们可以每个文件（有序的）都取一部分头部数据最为一个 buffer，并且把这 100个 buffer放在一个堆里面，进行堆排序，比较方式就是对所有堆元素（buffer）的head元素进行比较大小，然后不断的把每个堆顶的 buffer 的head 元素 pop 出来输出到最终文件中，（想象链表的有序合并）然后继续堆排序，继续输出。如果哪个 buffer 空了，就去对应的文件中继续补充一部分数据。最终就得到一个全局有序的大文件。