

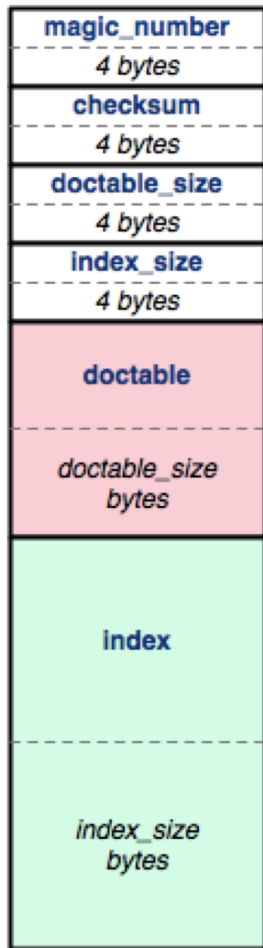
CSE 333 – SECTION 6

- HW3 Hex View

Overview

- HW3 due next Thursday
- Midterms are posted – regrade requests can be submitted through Gradescope
- Questions?
- Reading HW3 index files

Index File



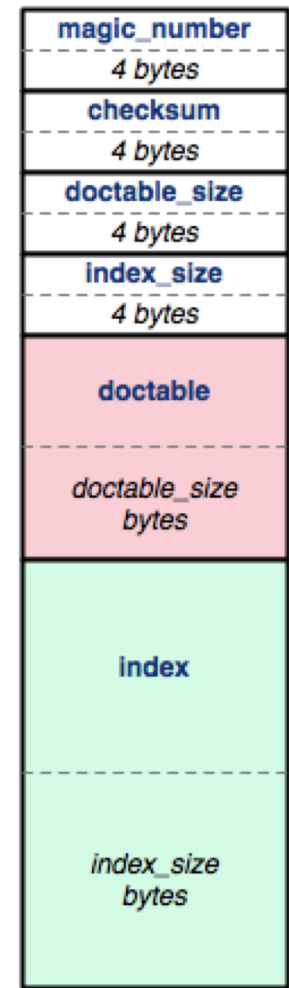
index file

- magic_number: 0xCAFEF00D
- checksum: mathematical signature
- doctable_size: in bytes
- index_size: in bytes

Index File Header - HEX

- Find a hex editor/viewer of your choice
 - `xxd <indexfile>`
 - `hexdump -vC <indexfile>`
 - Pipe the output into a file or `less` to view

```
0000000: cafe f00d 1c42 4620 0000 205b 0000 075d  ....BF .. [...]
0000010: 0000 0400 0000 0000 0000 2014 0000 0001  .....
0000020: 0000 2014 0000 0001 0000 2031 0000 0001  .. ..... 1....
0000030: 0000 204e 0000 0000 0000 206b 0000 0000  .. N..... k....
0000040: 0000 206b 0000 0000 0000 206b 0000 0000  .. k..... k....
0000050: 0000 206b 0000 0000 0000 206b 0000 0000  .. k..... k....
```



index file

The header:

Magic word **Checksum** **Doctable size** **Index size**

man xxd
man hexdump

Hex View

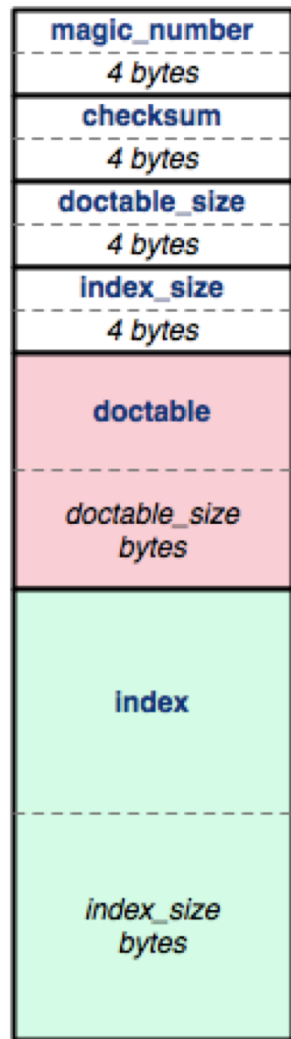
- emacs “M-x hexl-mode”

```
File Edit Options Buffers Tools Hexl Help
87654321 0011 2233 4455 6677 8899 aabb ccdd eeff 0123456789abcdef
00000000: cafe f00d ce52 0578 0000 205e 0000 0944 .....R.x.. ^...D
00000010: 0000 0400 0000 0000 0000 2014 0000 0001 .....
00000020: 0000 2014 0000 0001 0000 2032 0000 0001 .. ..... 2....
00000030: 0000 2050 0000 0000 0000 206e 0000 0000 .. P..... n....
00000040: 0000 206e 0000 0000 0000 206e 0000 0000 .. n..... n....
00000050: 0000 206e 0000 0000 0000 206e 0000 0000 .. n..... n....
```

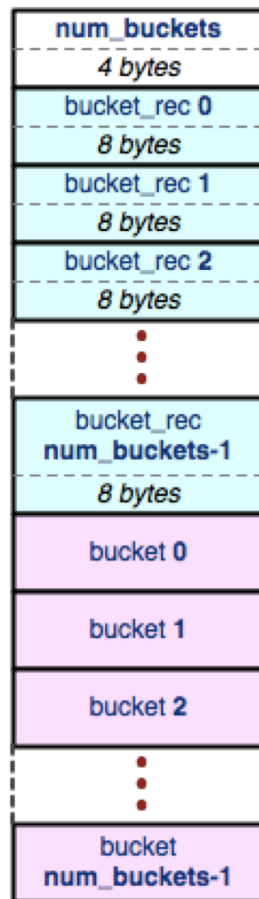
- vim “:%!xxd”

```
0000000: cafe f00d 1c42 4620 0000 205b 0000 075d .....BF .. [...]
0000010: 0000 0400 0000 0000 0000 2014 0000 0001 .....
0000020: 0000 2014 0000 0001 0000 2031 0000 0001 .. ..... 1....
0000030: 0000 204e 0000 0000 0000 206b 0000 0000 .. N..... k....
0000040: 0000 206b 0000 0000 0000 206b 0000 0000 .. k..... k....
0000050: 0000 206b 0000 0000 0000 206b 0000 0000 .. k..... k....
```

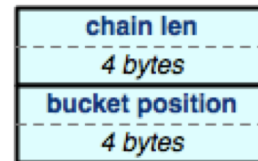
doctable



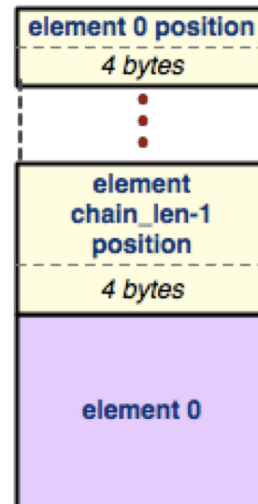
index file



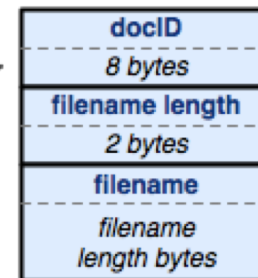
doctable



bucket_rec

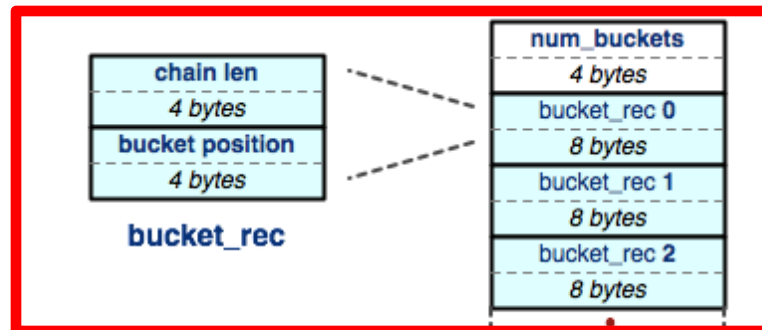


bucket



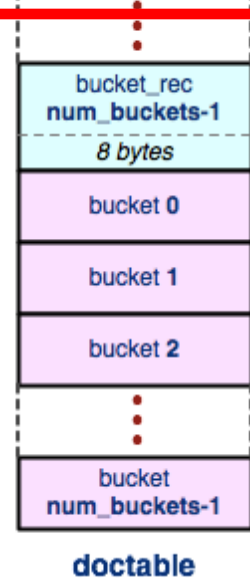
element

doctable



```

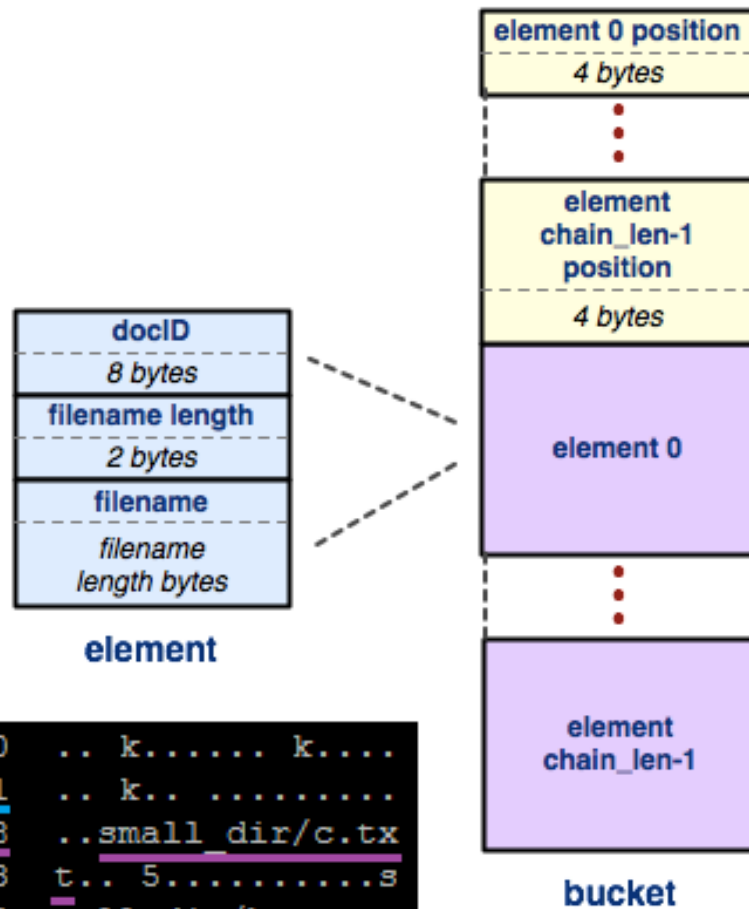
0000000: cafe f00d 1c42 4620 0000 205b 0000 075d  ....BF .. [...]
0000010: 0000 0400 0000 0000 0000 2014 0000 0001  .....
0000020: 0000 2014 0000 0001 0000 2031 0000 0001  .. ..... 1....
0000030: 0000 204e 0000 0000 0000 206b 0000 0000  .. N..... k....
0000040: 0000 206b 0000 0000 0000 206b 0000 0000  .. k..... k....
0000050: 0000 206b 0000 0000 0000 206b 0000 0000  .. k..... k....
0002000: 0000 206b 0000 0000 0000 206b 0000 0000  .. k..... k....
0002010: 0000 206b 0000 2018 0000 0000 0000 0001  .. k.. .....
0002020: 000f 736d 616c 6c5f 6469 722f 632e 7478  ..small_dir/c.tx
0002030: 7400 0020 3500 0000 0000 0000 0200 0f73  t.. 5.....s
0002040: 6d61 6c6c 5f64 6972 2f62 2e74 7874 0000  mall_dir/b.txt..
0002050: 2052 0000 0000 0000 0003 000f 736d 616c  R.....smal
0002060: 6c5f 6469 722f 612e 7478 7400 0000 8000  l_dir/a.txt....
0002070: 0000 0000 0024 6f00 0000 0000 0024 6f00  ....$o.....$o.
  
```



The header

Num buckets (Chain len Bucket offset)*

doctable



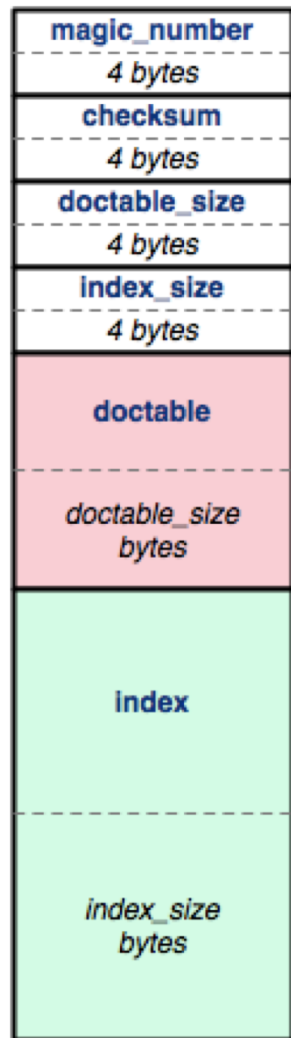
```

0002000: 0000 206b 0000 0000 0000 206b 0000 0000 .. k..... k....
0002010: 0000 206b 0000 2018 0000 0000 0000 0001 .. k.. .....
0002020: 000f 736d 616c 6c5f 6469 722f 632e 7478 ..small_dir/c.tx
0002030: 7400 0020 3500 0000 0000 0000 0200 0f73 t.. 5.....s
0002040: 6d61 6c6c 5f64 6972 2f62 2e74 7874 0000 mall_dir/b.txt..
  
```

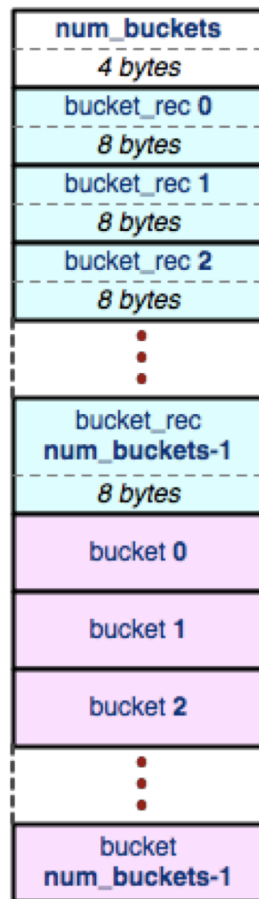
The buckets:

$$\left(\left(\text{Element offset} \right)^n \left(\text{DocID} \quad \text{Filename len} \quad \text{Filename} \right)^n \right)^*$$

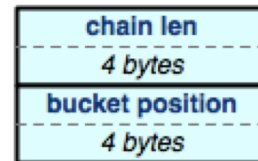
doctable



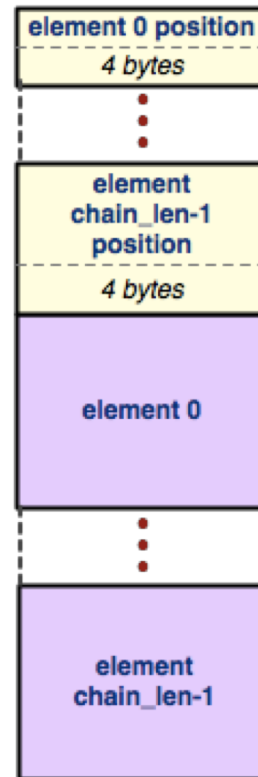
index file



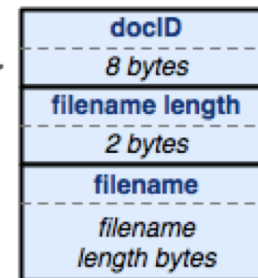
doctable



bucket_rec

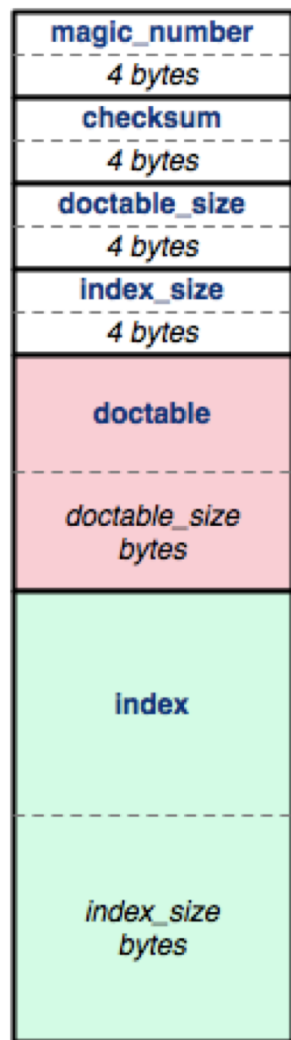


bucket

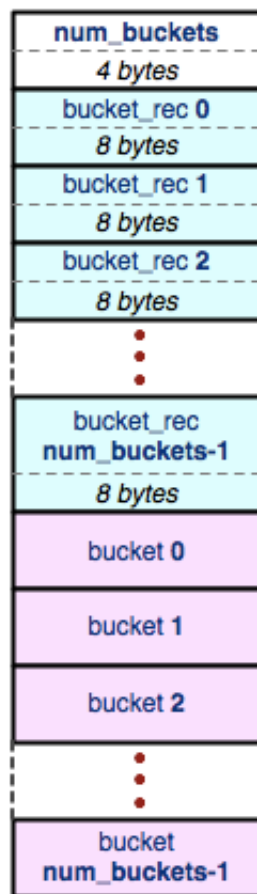


element

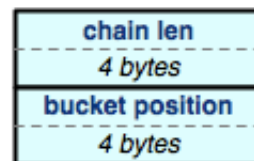
index



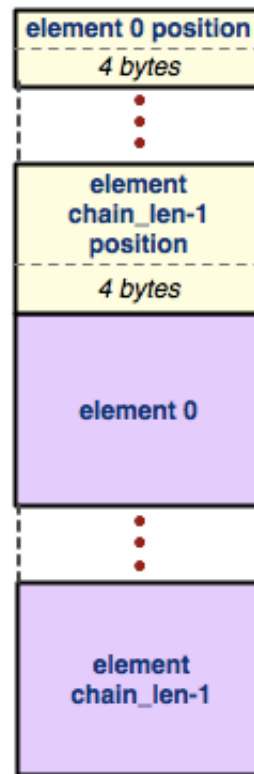
index file



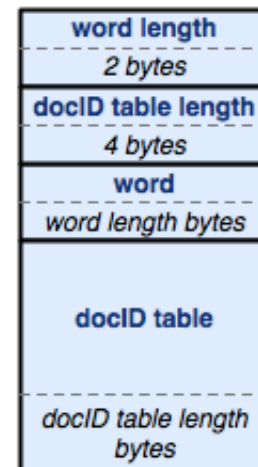
index



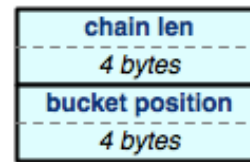
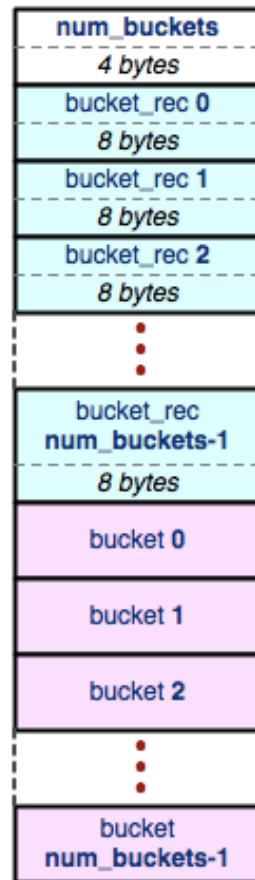
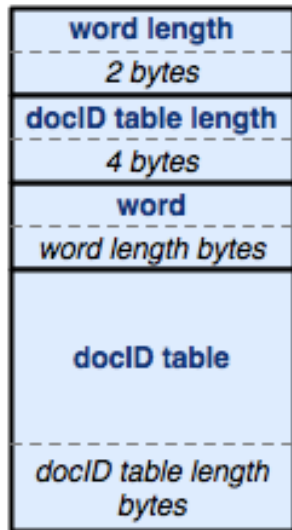
bucket_rec



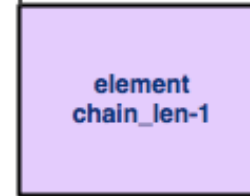
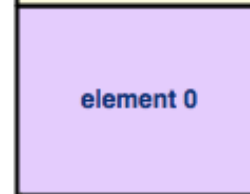
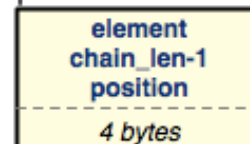
bucket



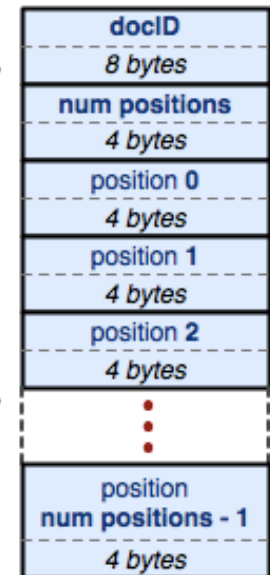
docID table



bucket_rec



bucket



element