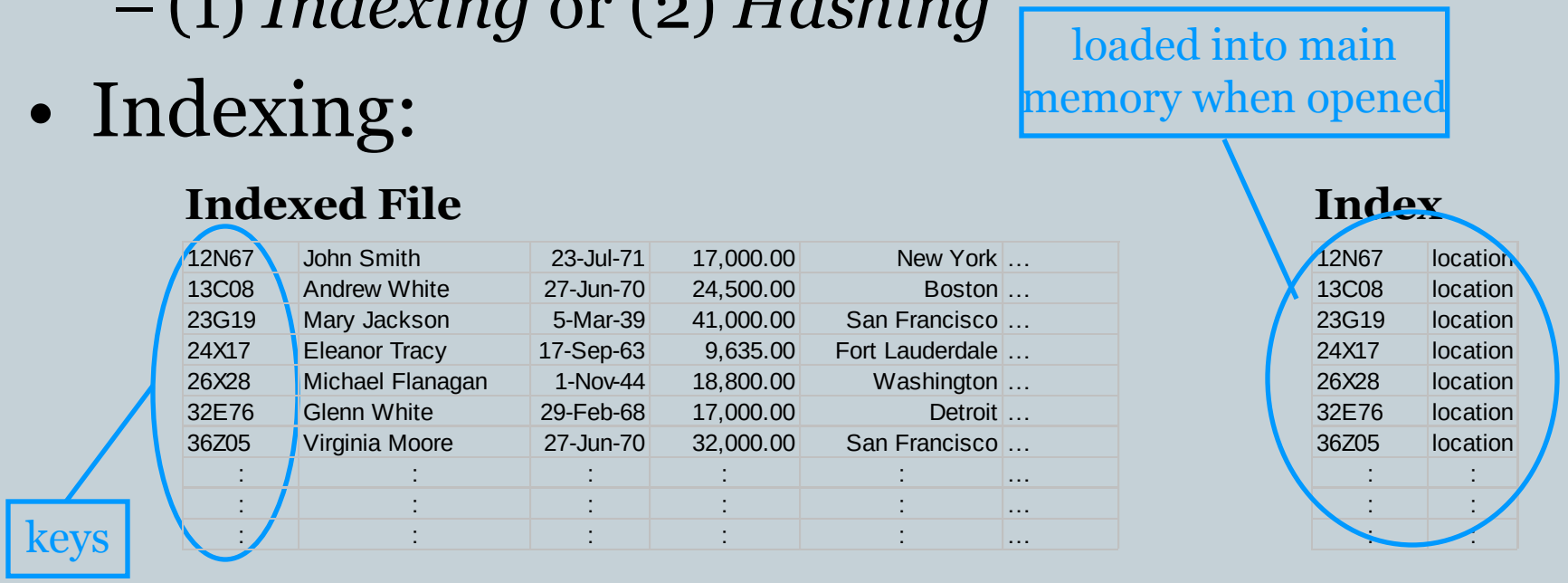
Indexing and Hashing

CHAPTER 8: FILE STRUCTURES
COMPUTER SCIENCE: AN OVERVIEW

LECTURE PREPARED AND DELIVERED BY
DR SYED KHALDOON KHURSHID

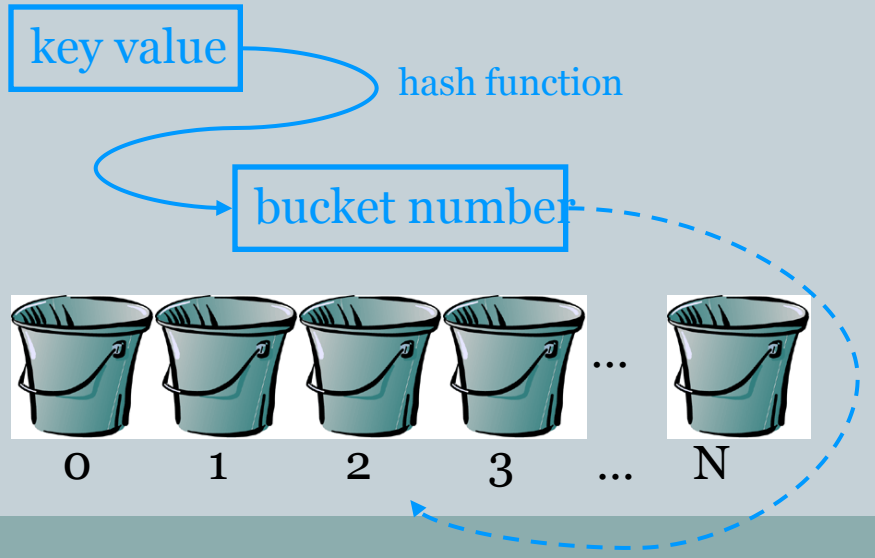
8.3: Quick File Access

- Disadvantage of sequential files:
  - no quick access to particular file data
- Two techniques to overcome this problem:
  - (1) *Indexing* or (2) *Hashing*
- Indexing:



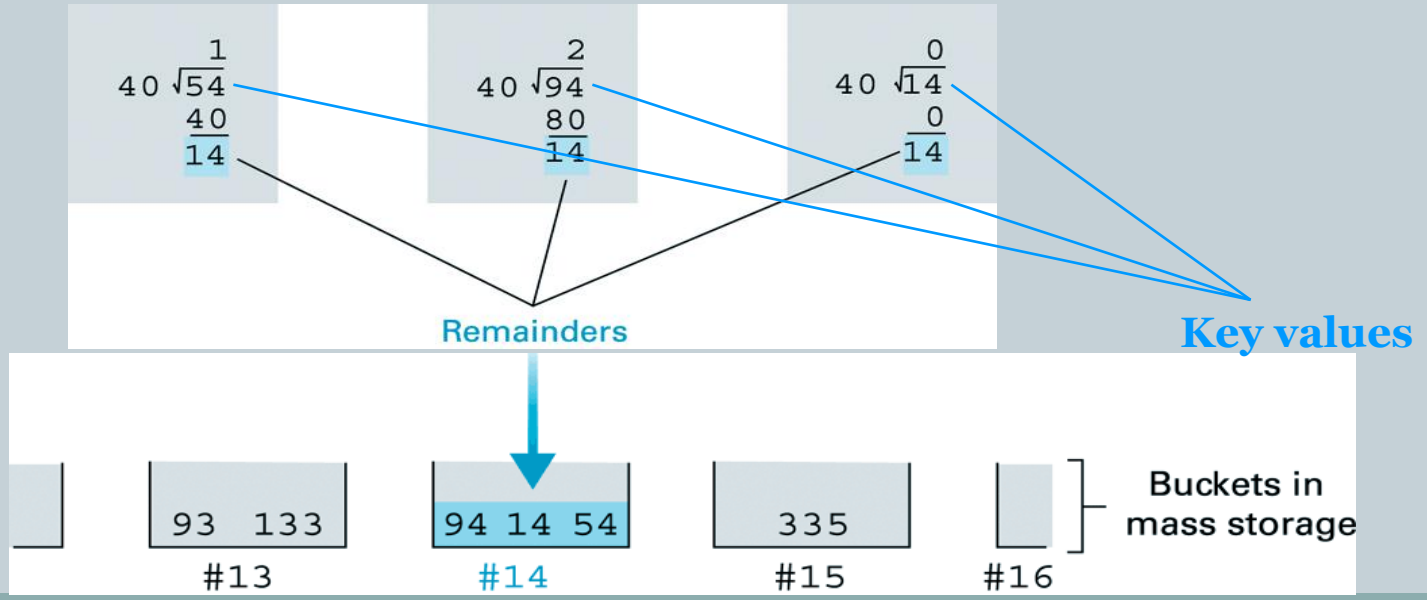
8.4: Hashing

- Solution: '*hashing*'
  - finds position in file using a key value (as in indexing)…
  - … simply by identifying location *directly from the key*
- Disadvantage of indexing is… the index
  - requires extra space
- How?
  - define set of '*buckets*' & '*hash function*' that converts keys to bucket numbers



8.4: Hash Function: Example

- If storage space divided into *40 buckets* and hash function is *division*:
  - key values 14, 54, & 94 all map onto same bucket (collision)




Any
Questions