

counting Bloom filter

[0000 0000]

add(foo)

$h_1(x) \rightarrow 1$
 $h_2(x) \rightarrow 2$

[0110 0000]

add(bar)

$h_1(bar) \rightarrow 1$
 $h_2(bar) \rightarrow 4$

[0210 1000]

is member(foo) $\begin{matrix} 1 \\ 2 \end{matrix}$

is member(bar) $\begin{matrix} 1 \\ 4 \end{matrix}$

delete(foo) $\begin{matrix} 1 \\ 2 \end{matrix}$

[0100 1000]

is member(foo)? $\begin{matrix} 1 \\ 2 \end{matrix}$

is member(bar)? $\begin{matrix} 1 \\ 4 \end{matrix}$

✓ deleted

✓ still there

Google — web search engine

① obtain
a DB
of "free"
web
(server)

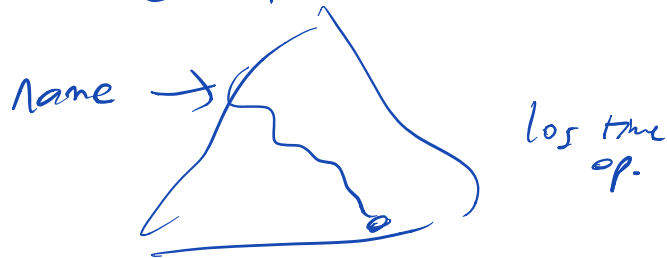
② performs
searches
(client)



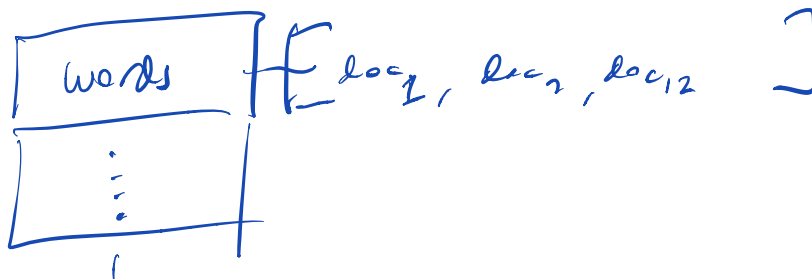
flat data

Searching
everything ~~X~~
INDICES

DB indices — B+ tree



Inverted index



cute	3	15	199						
cat	1	2	3	4	5	6			
dog	4	7	19						
monkey	111	119	199						
pleonasm	145								

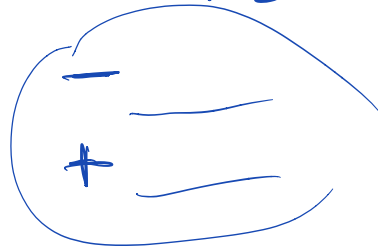
"Cute cat"

Conjunctive query

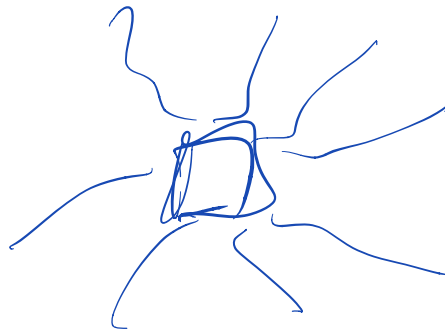
inverted indices

set { 3, 15, 199 }
 (100) (95) (5)

ranked version set



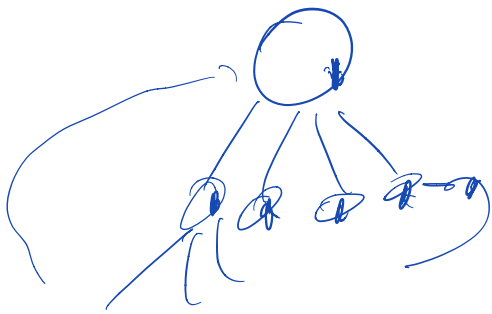
PageRank
random walk



"link spam"

DB of web pages

web spider "crawls" the web



"GC marking"

recursive
search thru links

spider continuously

- Fast

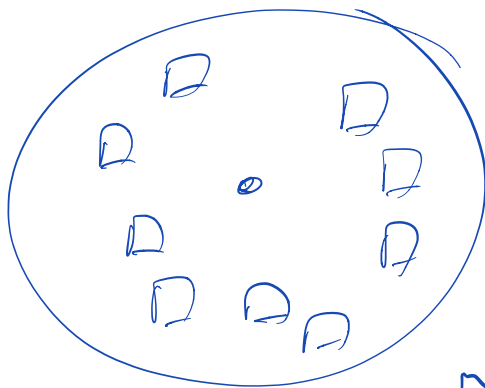
- write

(no updates)

- new versions ?

archive.org

SUPER EFFICIENT

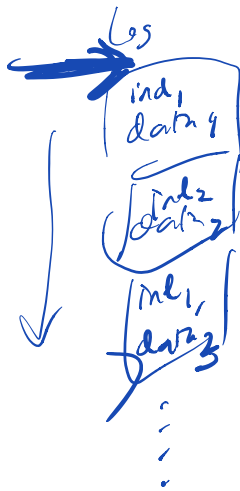
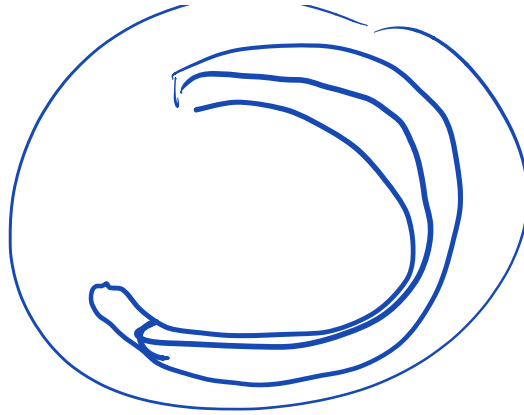


GFS

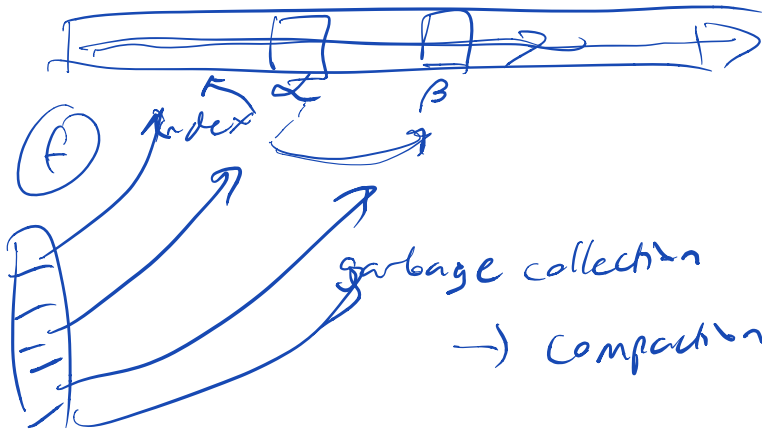
Log-structured File Systems

Rosenblum & Ousterhout

LFS



Replay log \Rightarrow fault tolerance



(Kafka - distributed log)

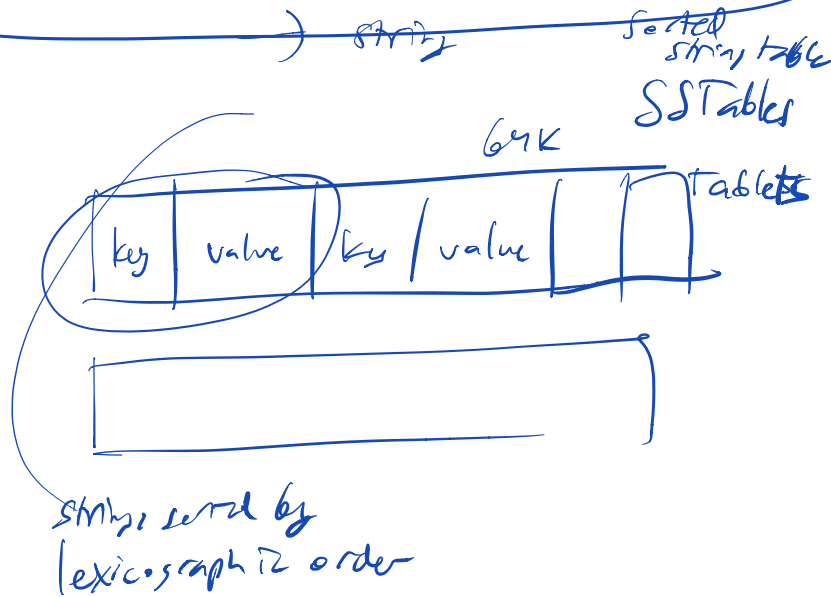
spidering - writes
versions

BigTable HBase Cassandra (FS)

get
(row: string, column: string, time: int64)
→ string

Index

key ₁	offset ₁
key ₂	offset ₂
key ₃	offset ₃

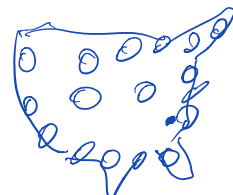


immutable

CAN.com



CDNs content delivery network



can
rom port

can
page₁

can
page₂

explicit control of
Locality

http://www.cnn.com/a
ab
a
z

~~com.cnn~~

com.cnn.www /

edu.umass.

edu.umass.CICS

○

local key access \Rightarrow
local tablet access



sparse — doesn't store empty columns

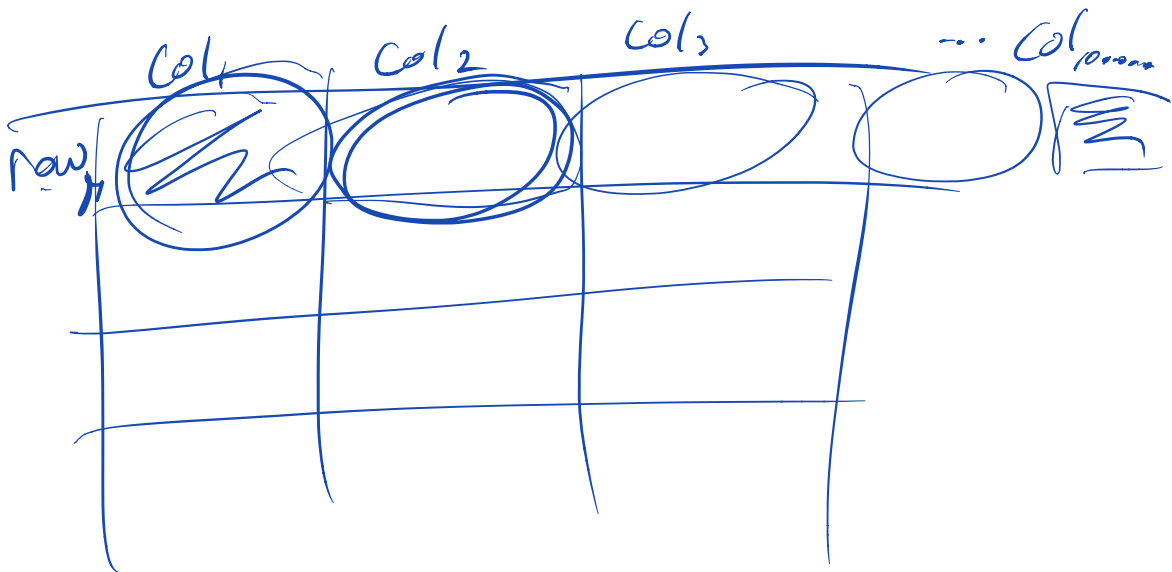
timestamp

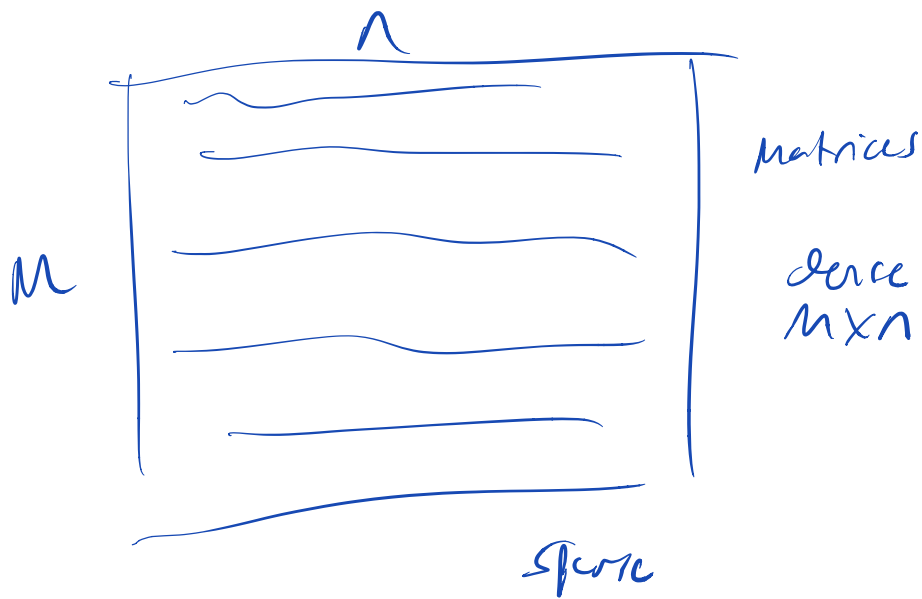
"GC policy"

timestamp older than X days

last 3 versions

⇒ versioning





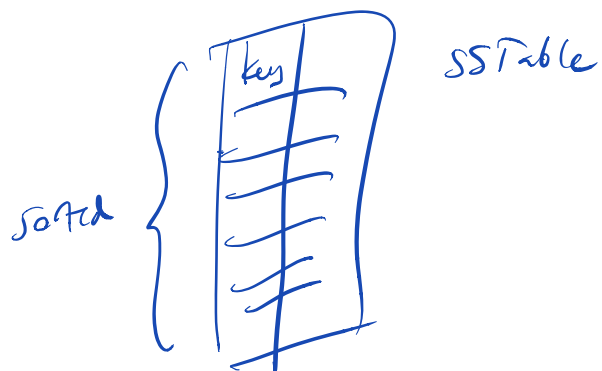
$(r, c, -)$ $(r, c, -)$

"key" — (row, col, times/zip)
 ~ ~ ~ RT69

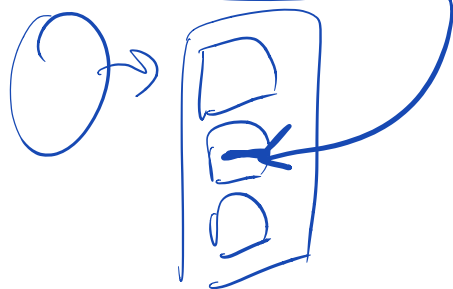
value — web page

— Google Earth / GMaps

MemTables



~~com~~ com / Cate_Cat_Lo1.html



MemTable
local
Cache
for writes

