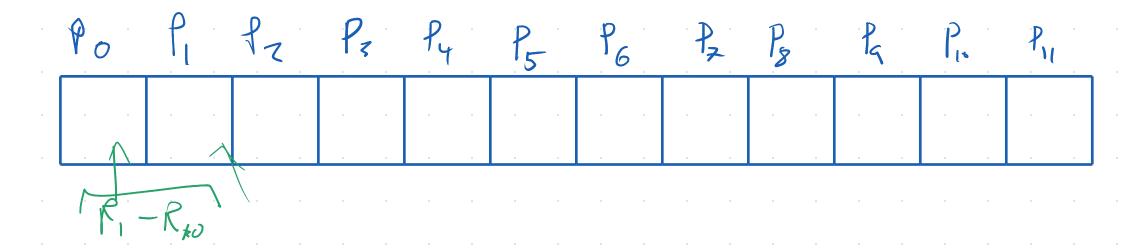
# CSE 350

Advanced Data Structures

Topic 9: Storage in Trees 📦 🌲



### File Layout



Goals

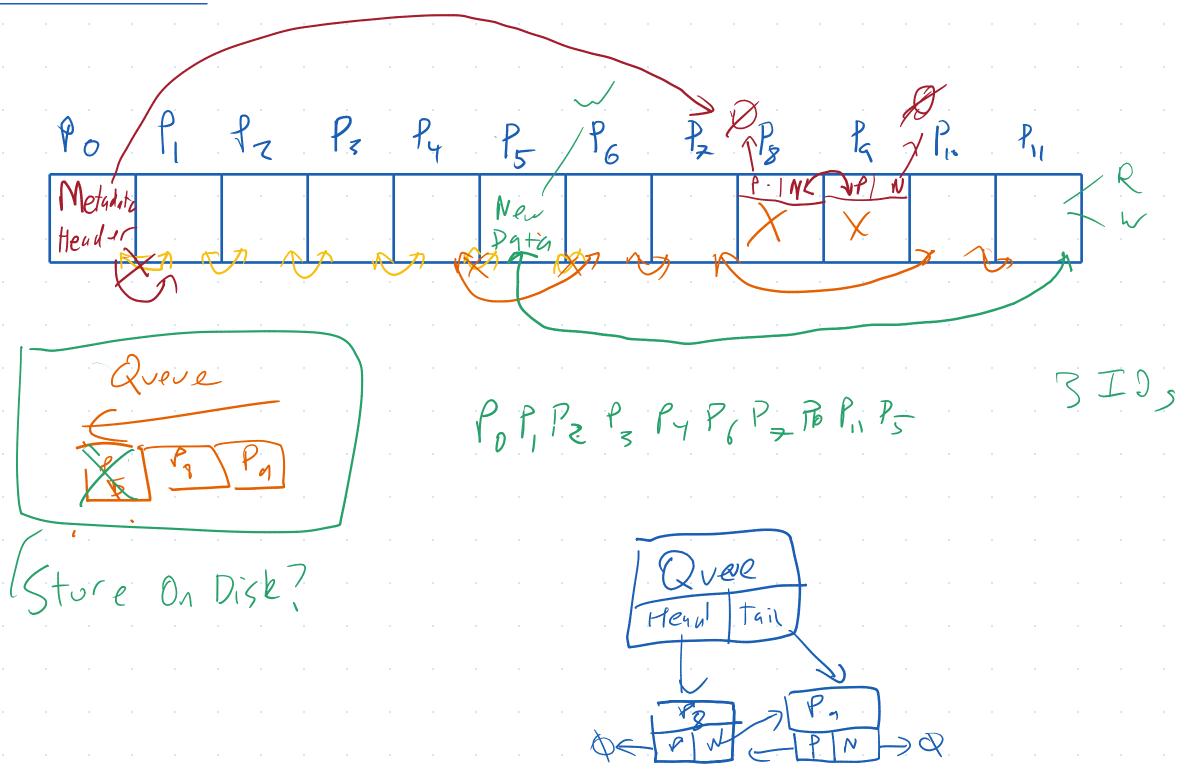
- Low wasted space > Fast 5 cans
- Data mutability

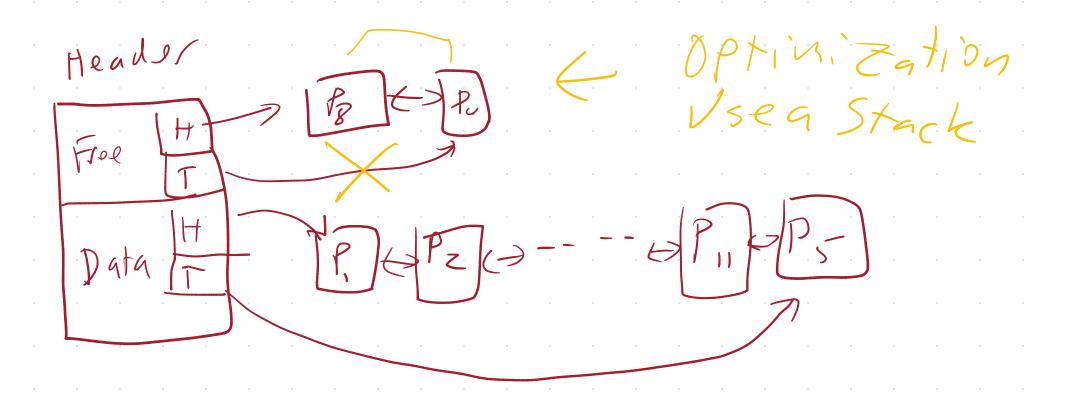
Flat File New Idea for nam? Type Address Bittalo Electrical -Delew Electrical --Buttul Load Page 5

Find Ps Ry

J GFINA RY

#### <u>Linked List</u>





### Finding Records

How is a record identified?

La Pointer (Page Recordindex)

L> Key

(Any attribute to id a collection of records)

#### What are the tradeoffs of each?

Key, ( Mable may not have one ( =) Need to map key to Location (Mapping function)

Point

 $\oplus$ 

OSQLONY talks in

8 Fast look Ups

How can we change the file layout to make key lookups faster?

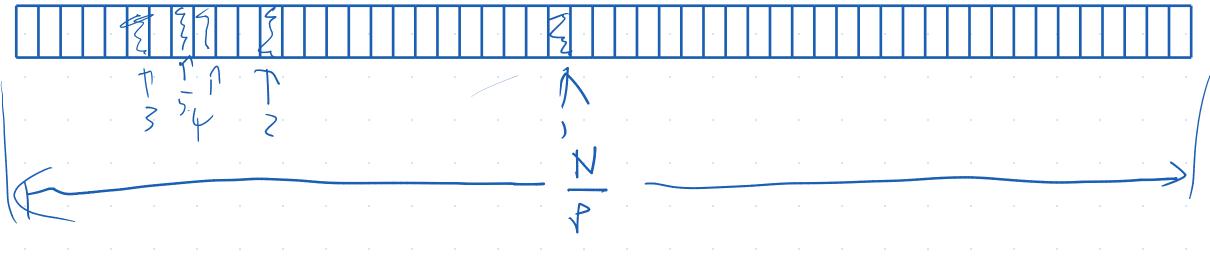
Problem
La No information relating key to location
by detailt

Hash Table

Sort it

#### Idea 1: Sort the Records

How do we find a record?



How Much Id?

Nirecords

Precords/Page

Search Space

No lanes

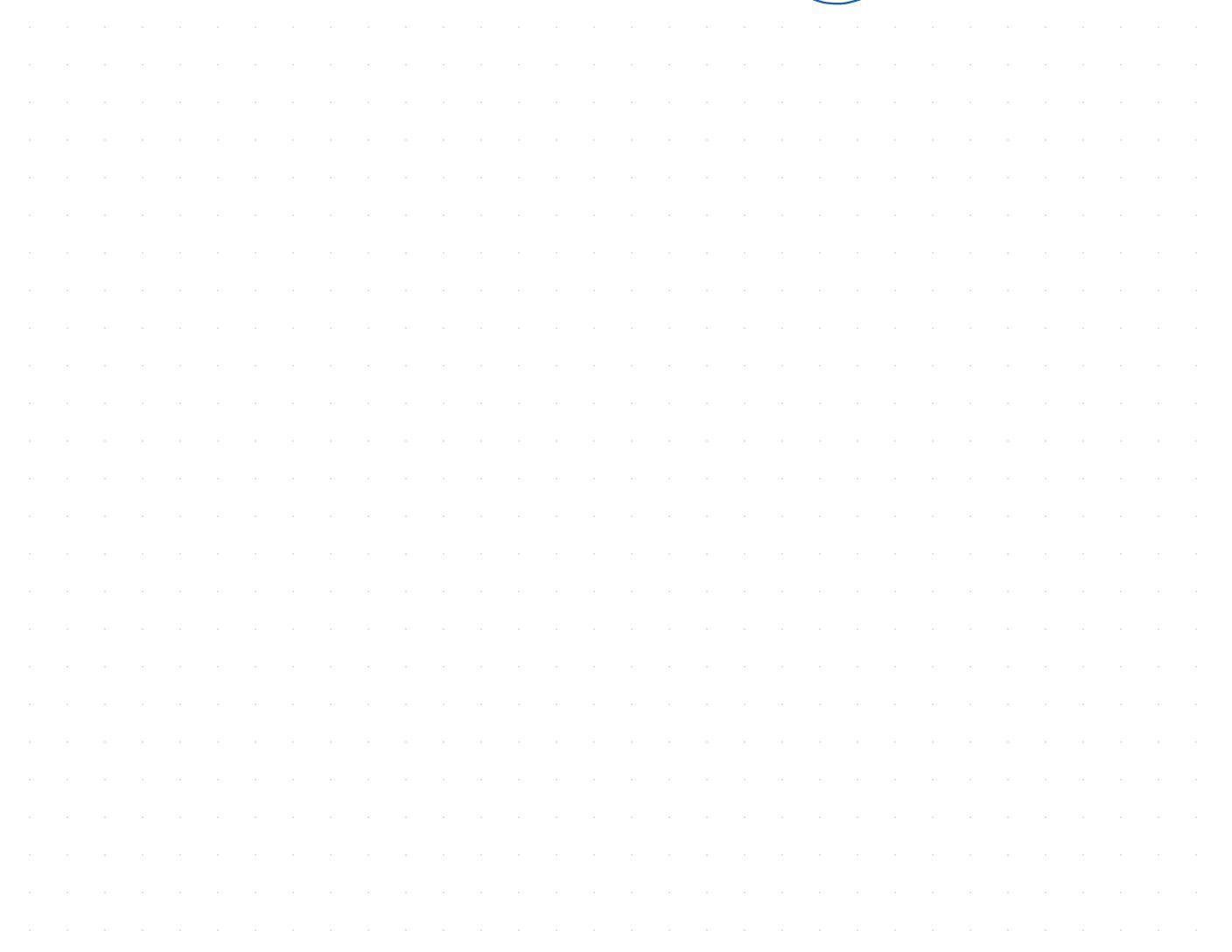
No lanes

No lanes

No lanes

No lanes

tog\_f = leg\_N-leg\_f,



What is the IO complexity of binary search?

On a sorted file

Log N-log Pages read

 $\mathcal{O}((092(N)))^{2}$ 

Not great

# What is the Memory complexity of binary search?

P+(
Pasebeing

analyzed

### **Relative Sizes**

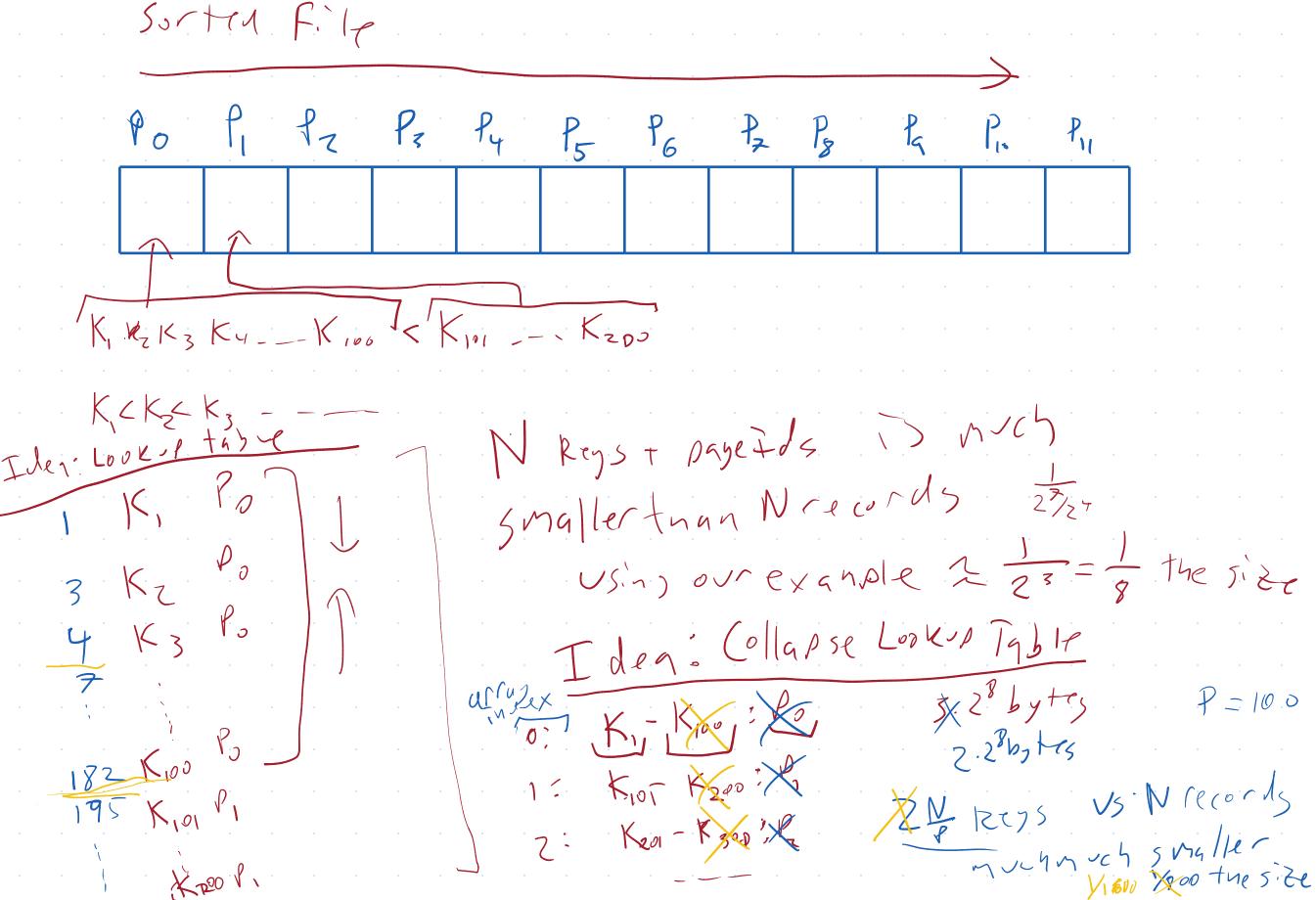
```
128B record

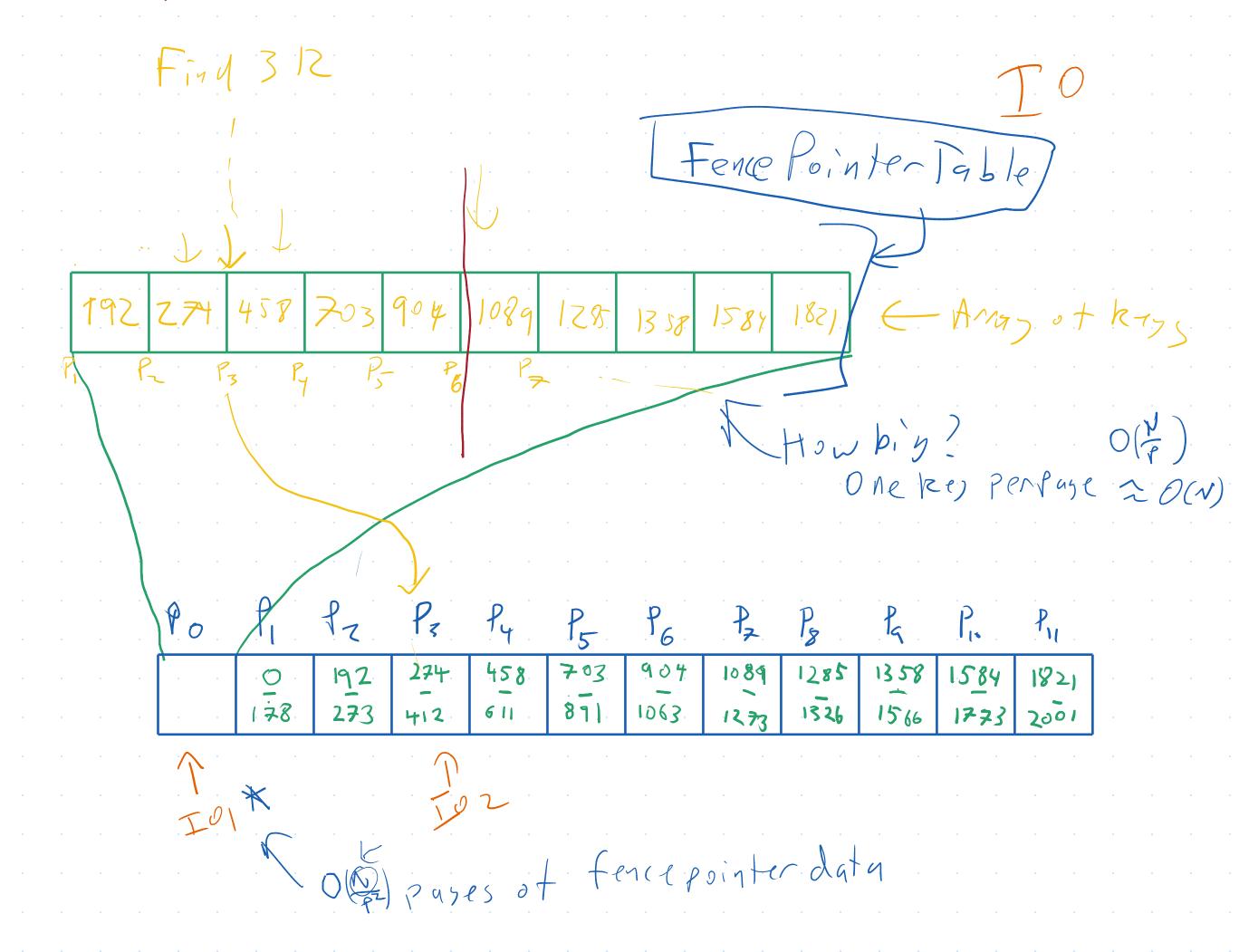
8B key - Porm + ID (in + 1)

How big is 20 million records? 2^{31} \approx 2GB

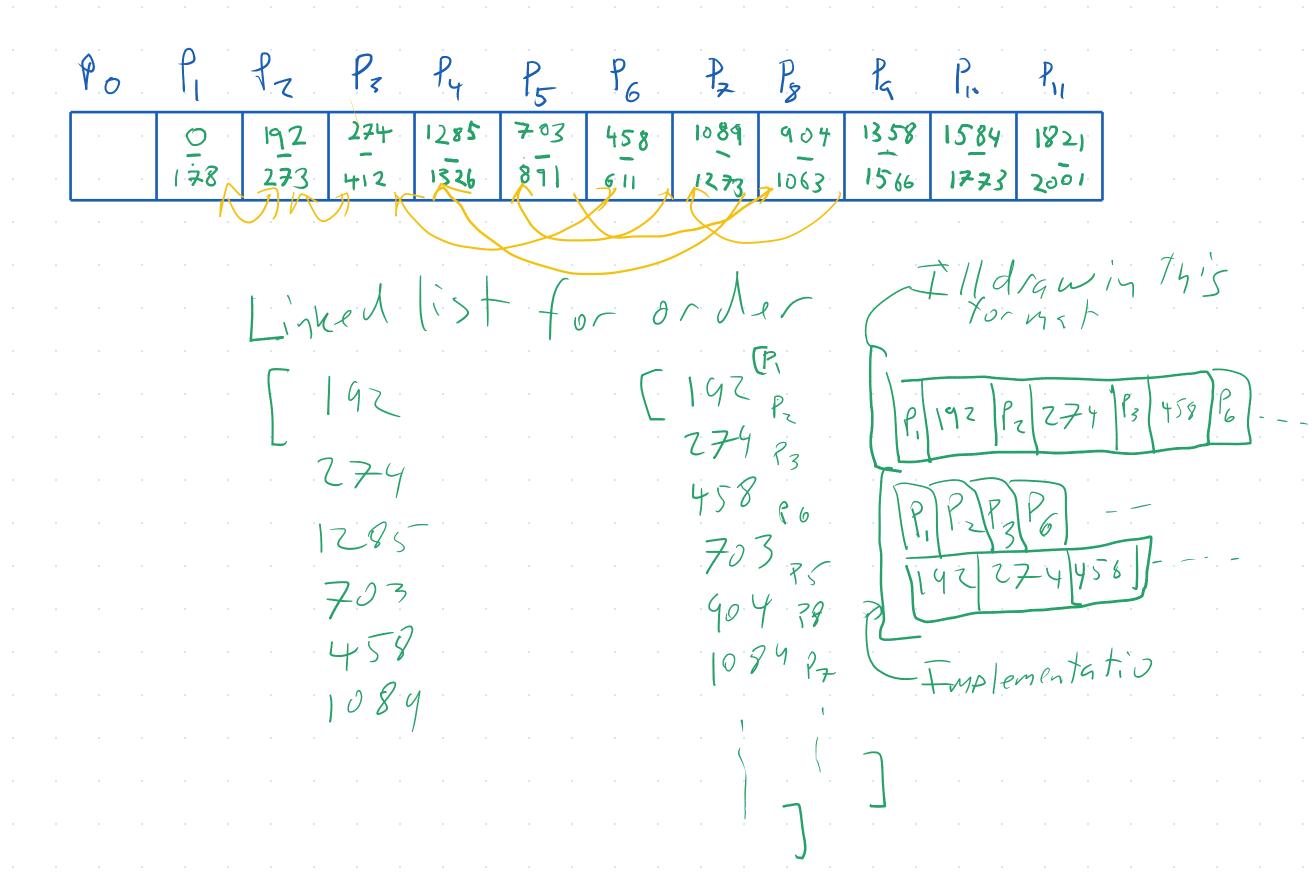
How big is 20 million keys? 2^{27} \approx 129 \text{ MD}
```

## Idea 2: Track key information for each page





#### Variation: Unordered Data Pages



What is	the IO	compl	exity c	of binar	y sea	rch?	Tab Lp
	30+ rea						G (045/47)
						5 g & Bi	G (onstant

Variation: Keer FP Table In memory
b) ITO Exactly

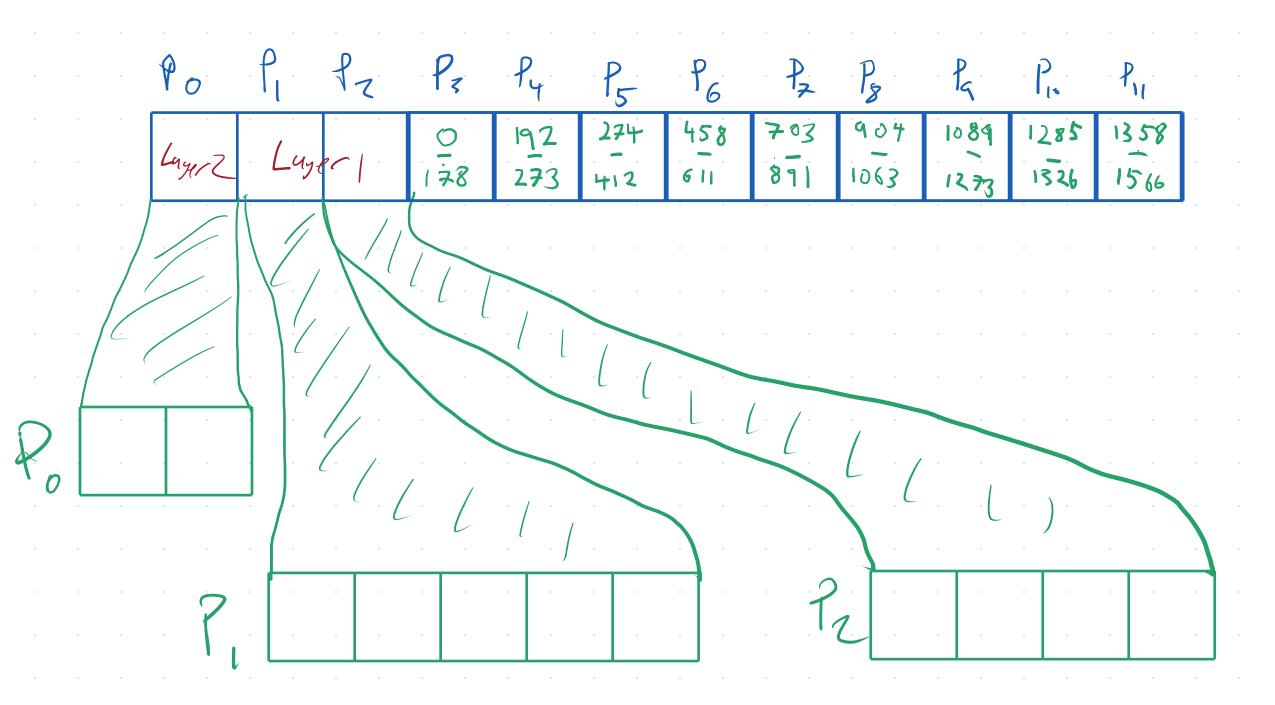
What is the Memory complexity of binary search?

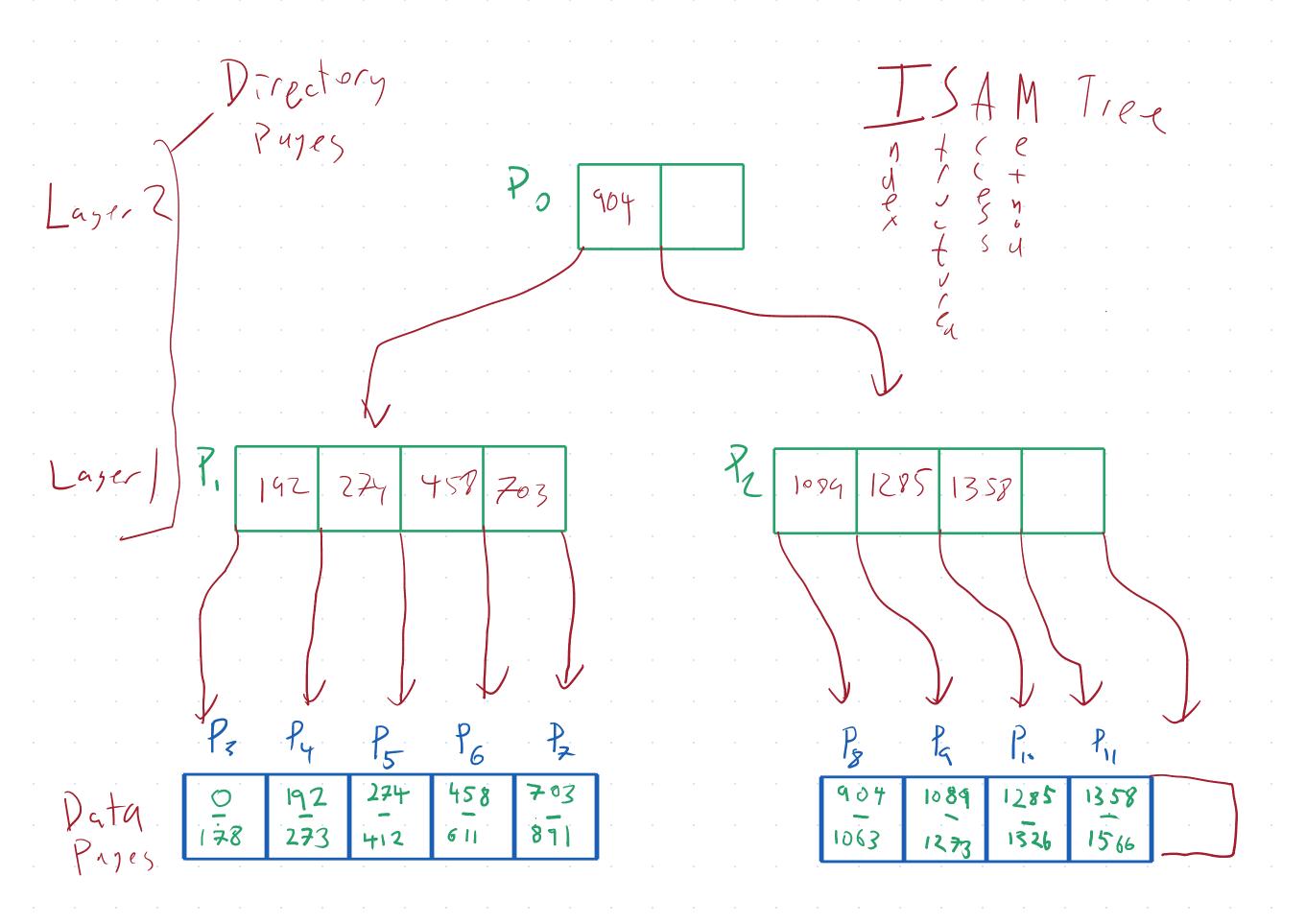
1. Page + Small constant

Variation: Keep FPTable In Momon



## Idea 3: Layers





How deep is the tree?

M, A of C(0-d)P # 2 + data pages Page average, expected K: Key Panz > Pages Key level);  $\frac{1}{P_{l}}$  Kegs Level Z: P. Keys -> N P.KZ3 Pages Level 3 - N Keys DIV PoK3y Pages level i : Non  $\frac{1}{\sqrt{\frac{1}{2}}} \int_{\mathbb{R}^{3}} q_{3} q_{3}$ 

ve to-\frac{1}{2} \frac{1}{2} \left( \frac{1}{2} \left( \frac{1}{2} \right) \left

will have I page for Noverords 109c ( = Loyk

What is the IO complexity of binary search? Log ( P) + 1 Reads N-100021024 3/2006 VS 10 levels

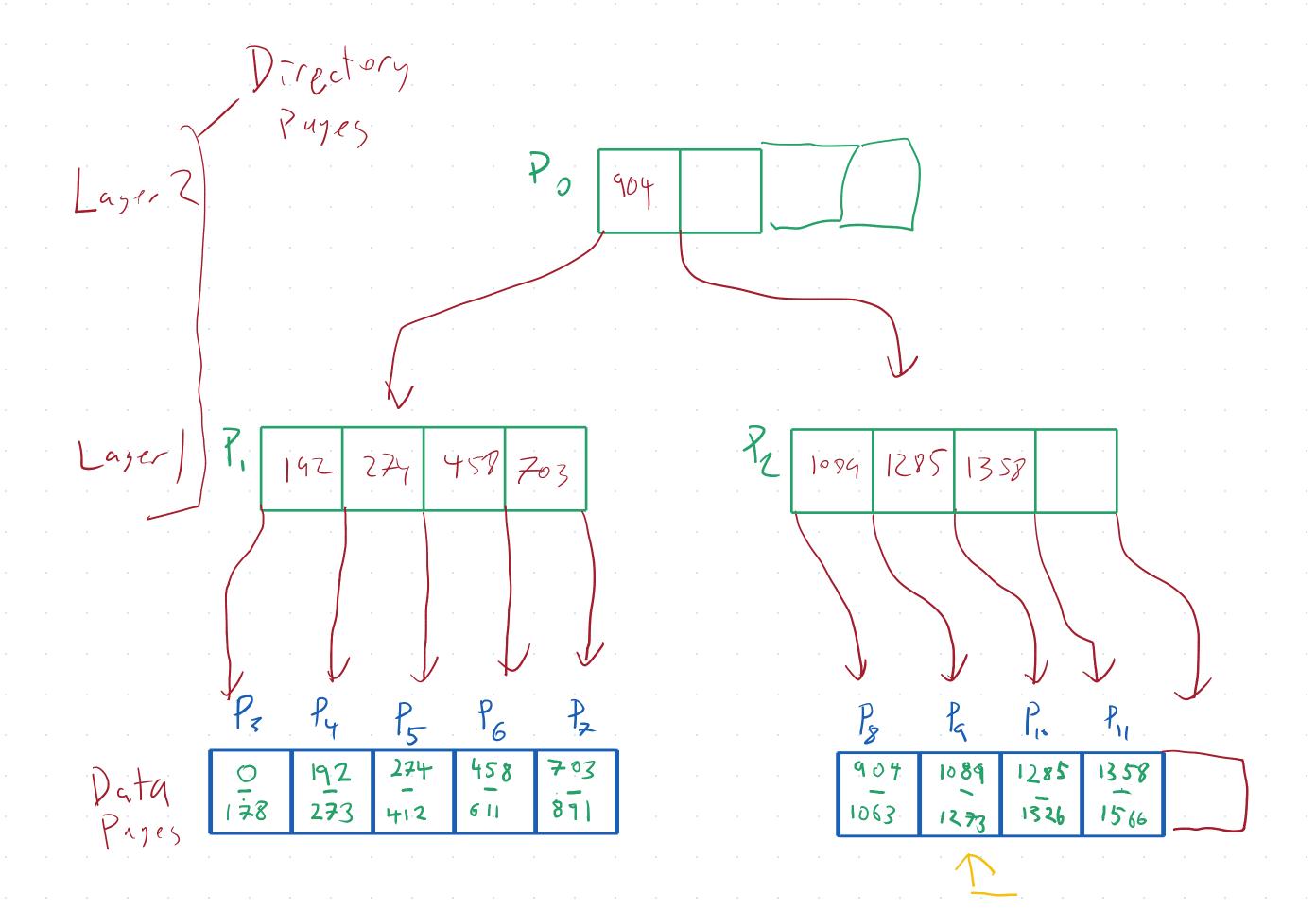
24 binas

search Big Diff. ( q 15 ) V5 log(N) for simple binary search

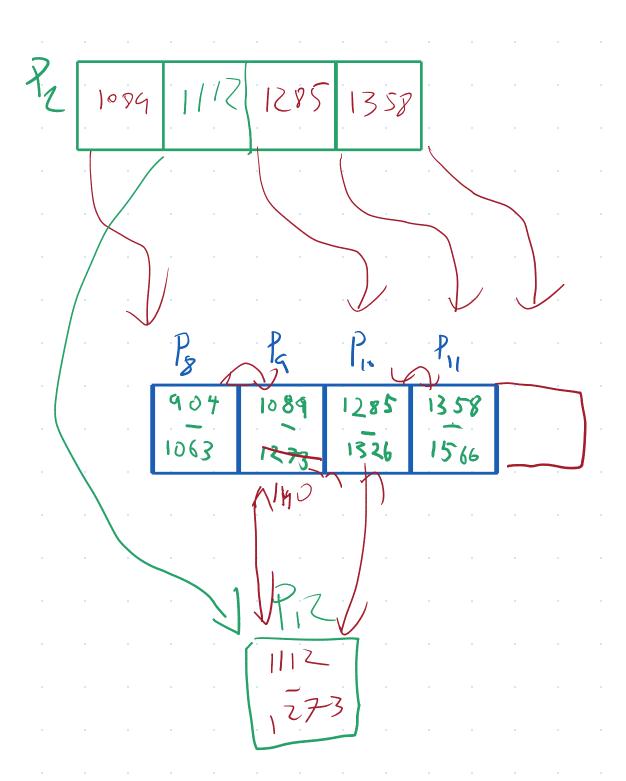
What is the Memory complexity of binary search?

I paye t constant

### What needs to change to allow changes?



Problem! NOSPACE to insert on agta pages 11 directory
Pages LProblem Z: No space 11 delete dounto Necords of X Problem 3: No way to reclaim tree layers
Problem 4: No way to grow the tree
Wilmore data Iden Pages as linkedtist



Instr GIF Space on Page / Aunt WSplit the page vsing LL DINSINT NEW Split, into Parint directory page b) It space in dir page Vlong W Split dir paye Drepent w/parent (If no parent (at root) Screate new layer u/newrot

### Idea 4: B+ Trees

Resspordir page

Resspordir page

Precords per datapage

tree

Pa Pro Pri

ad Paye has

Preconds

Steal;

Legeral

Merge

Lizz

P

Delete Record b) > Precords remain? La Cards remain b) Try to Steal from alt Paye L) Modity Purent dir paga DMerge w/either.ads Page L) Remove geparator transacent Glarenthas 2 Kz ergs Vdone Weart has LK Roys 4 5+eal US Merse Exception Rotcanhave

Litrothas only one child. Remove Footlevel Like an ISAM Tree Detapuses are a linked list 4) Dr- 2 Data Pages MUST Be half full DException: Root

( ) depth tree