

- 1. recap: fs and file
 - 2. file-mapping structures
 - 3. fs namespace
 - 4. hierarchical fs is dead?
 - 5. egos-fs preview
-

1. Recap: fs and file

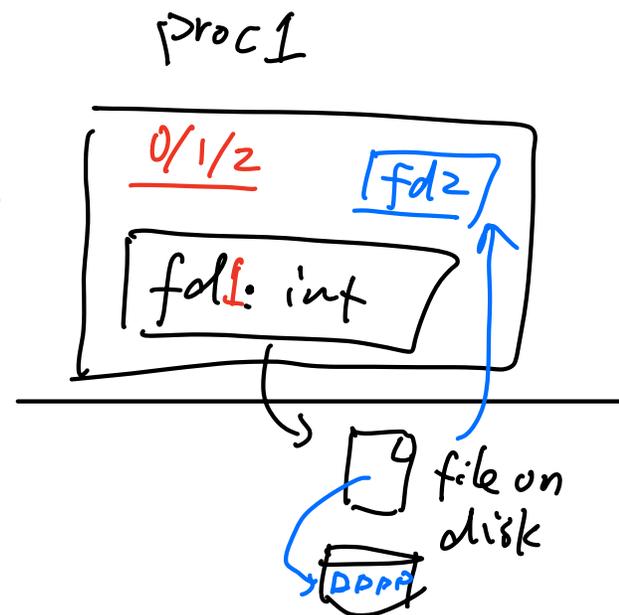
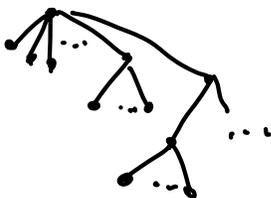
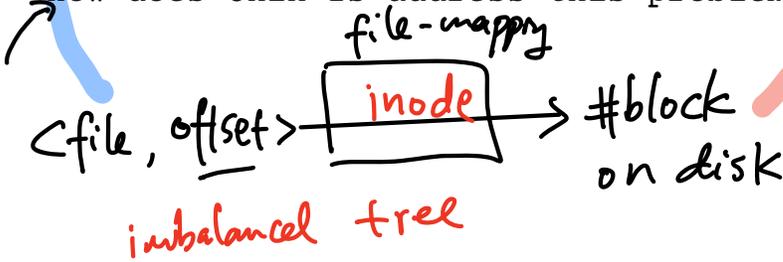
Q: what does a FS do?

1. persistency
2. name a seq of bytes
3. dirs: human-friendly names

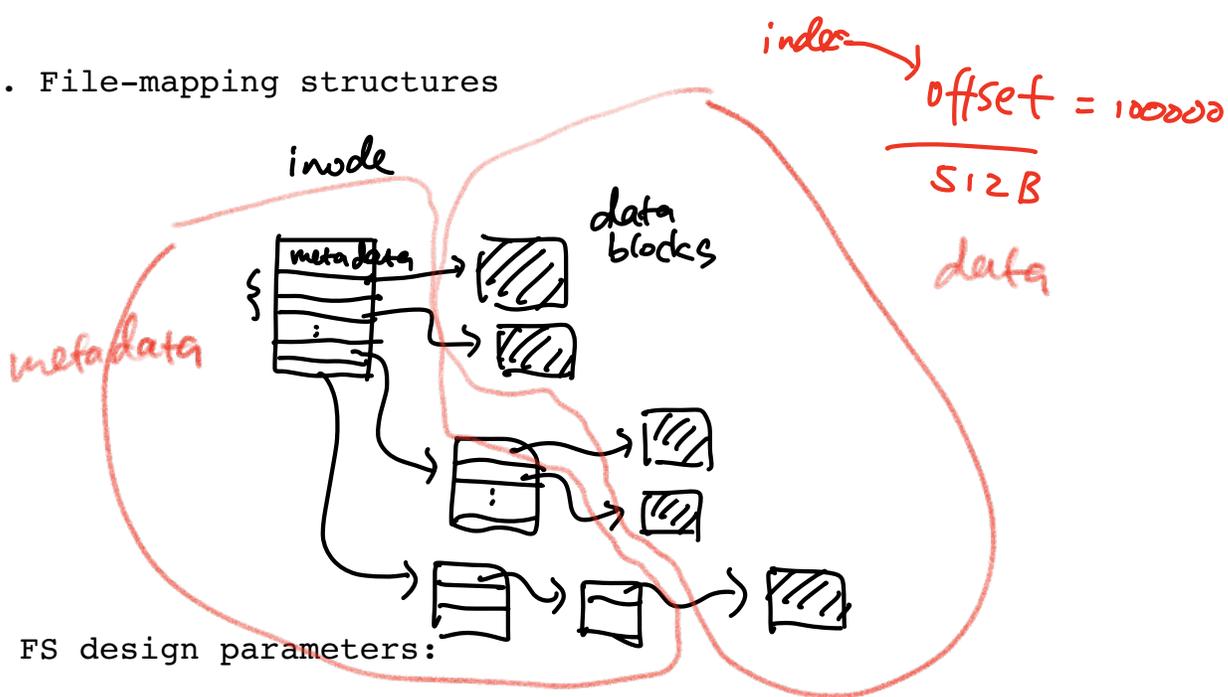
Q: what is a file?

- Users: a named array of bytes
- OS: a set of blocks

Q: What problem does file mapping aim to solve?
How does Unix fs address this problem?



2. File-mapping structures



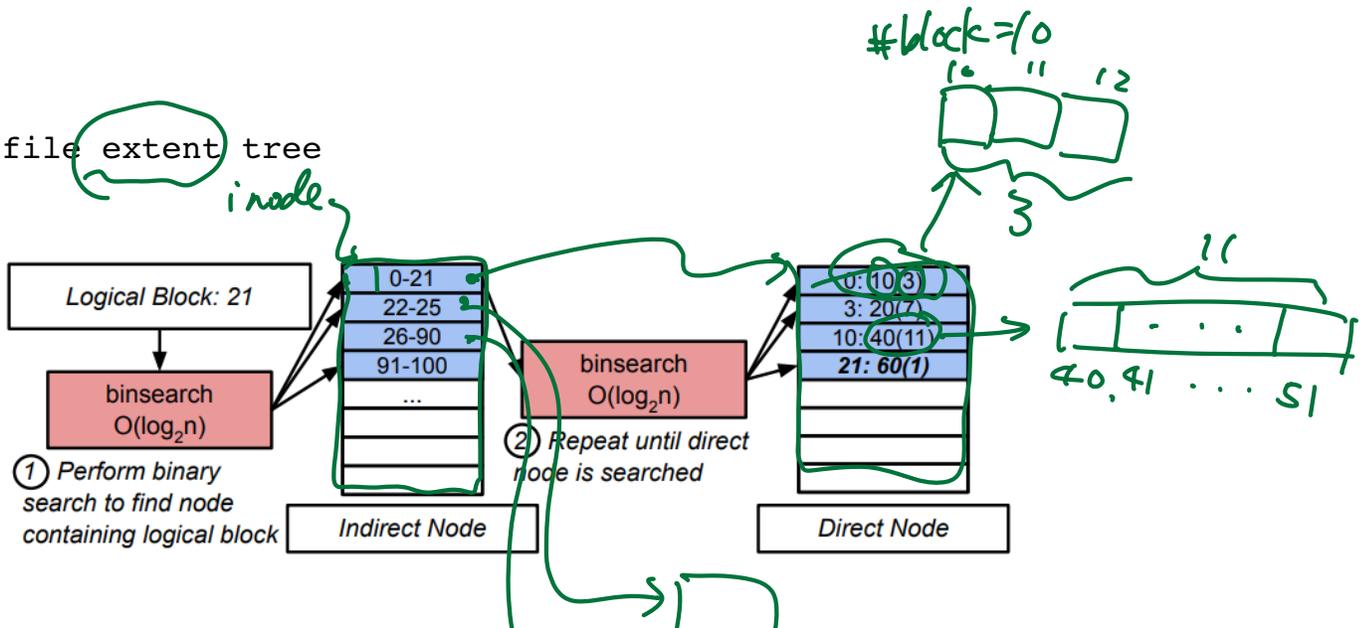
* FS design parameters:

- small files (most files are small) vs. large files (much of the disk is allocated to large files)
- sequential access vs. random accesses
- prefetching
- disk utilization (metadata overhead and fragmentation)

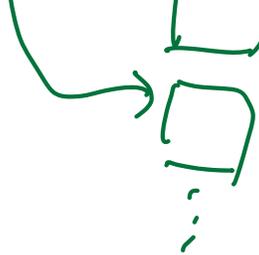
For UNIX fs

- [:)] small files
- [~] large files
- [:)] sequential access
- [~] random accesses
- [:)] prefetching
- [:)] disk utilization

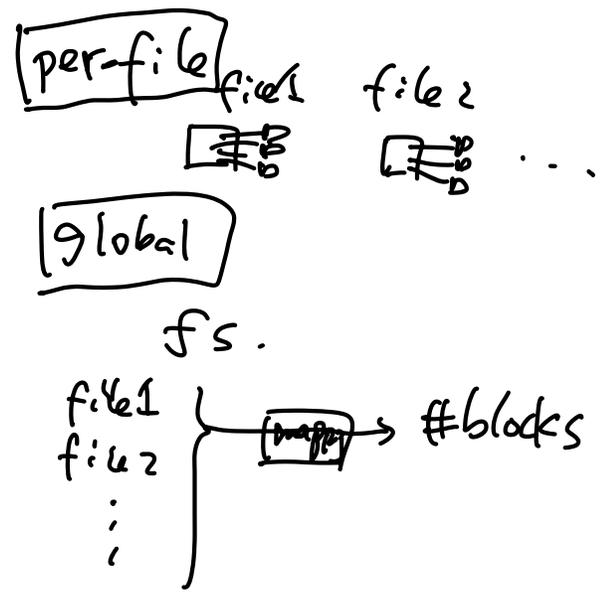
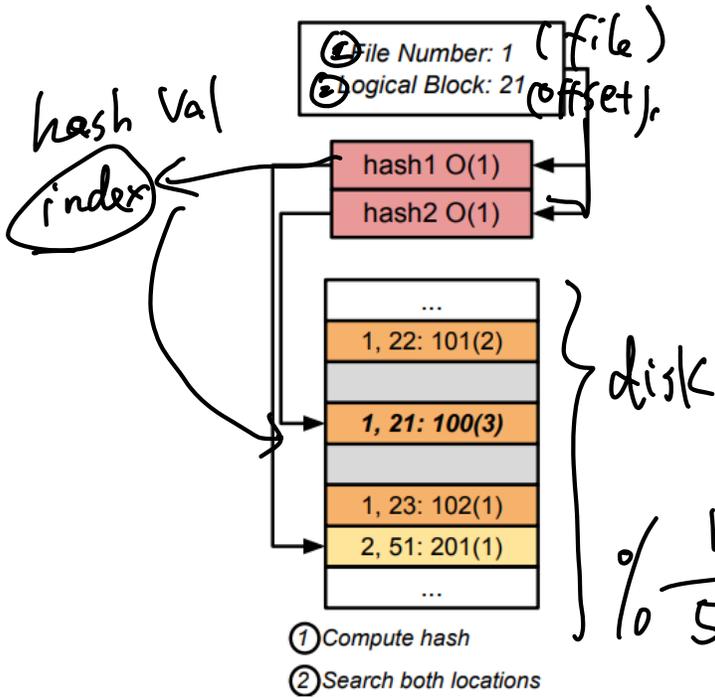
A. per-file extent tree



- [] small files
- [] large files
- [] sequential access
- [] random accesses
- [] prefetching
- [] disk utilization



B. global cuckoo hash table



$$\frac{1TB}{512B}$$

file = 1, offset = 21 →
 file = 1, offset = 22 →

- [~] small files
- [~] large files
- [: (] sequential access
- [] random accesses
- [: (] prefetching
- [:)] disk utilization

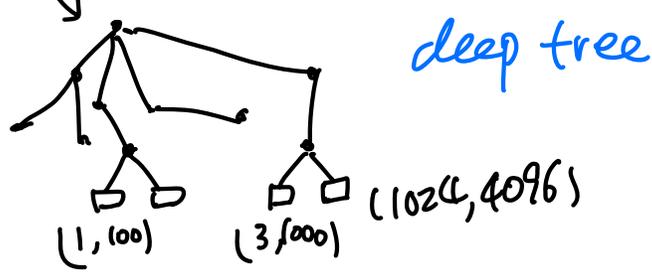
→ Q: Why not per-file hash mapping? What's wrong with it?

• unknown file size

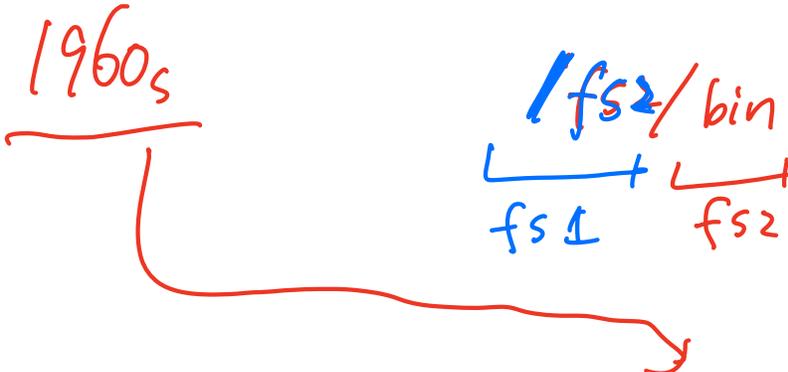
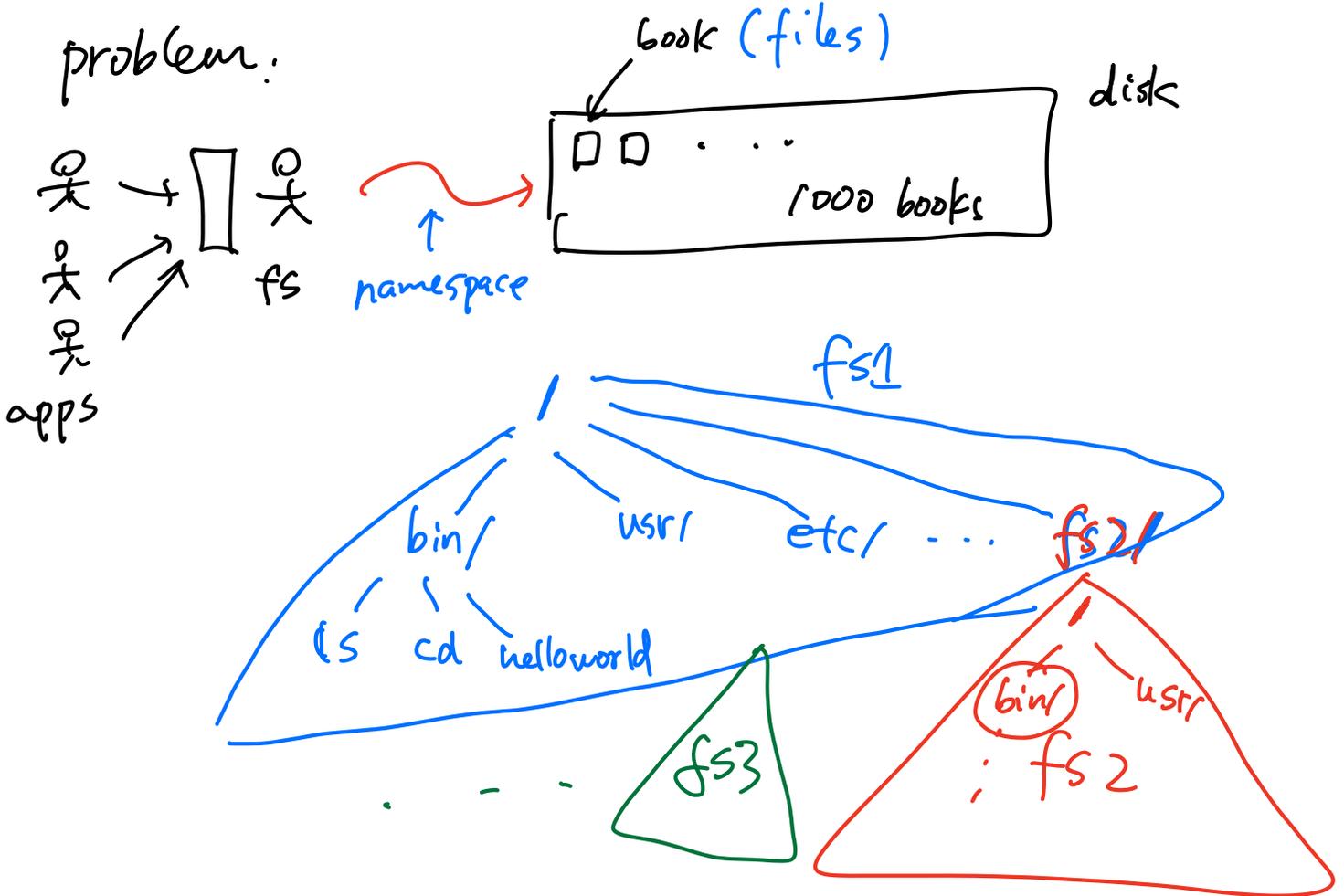
Q: Why not global extent trees? What's wrong with it?



#files is unknown



3. fs namespace



4. "Hierarchical file systems are dead"? (2009) by Margo Seltzer and Nicholas Murphy, HotOS'09

-- **hFSD** argues,
hierarchical namespace works good in the past.
"The situation, however, has evolved"

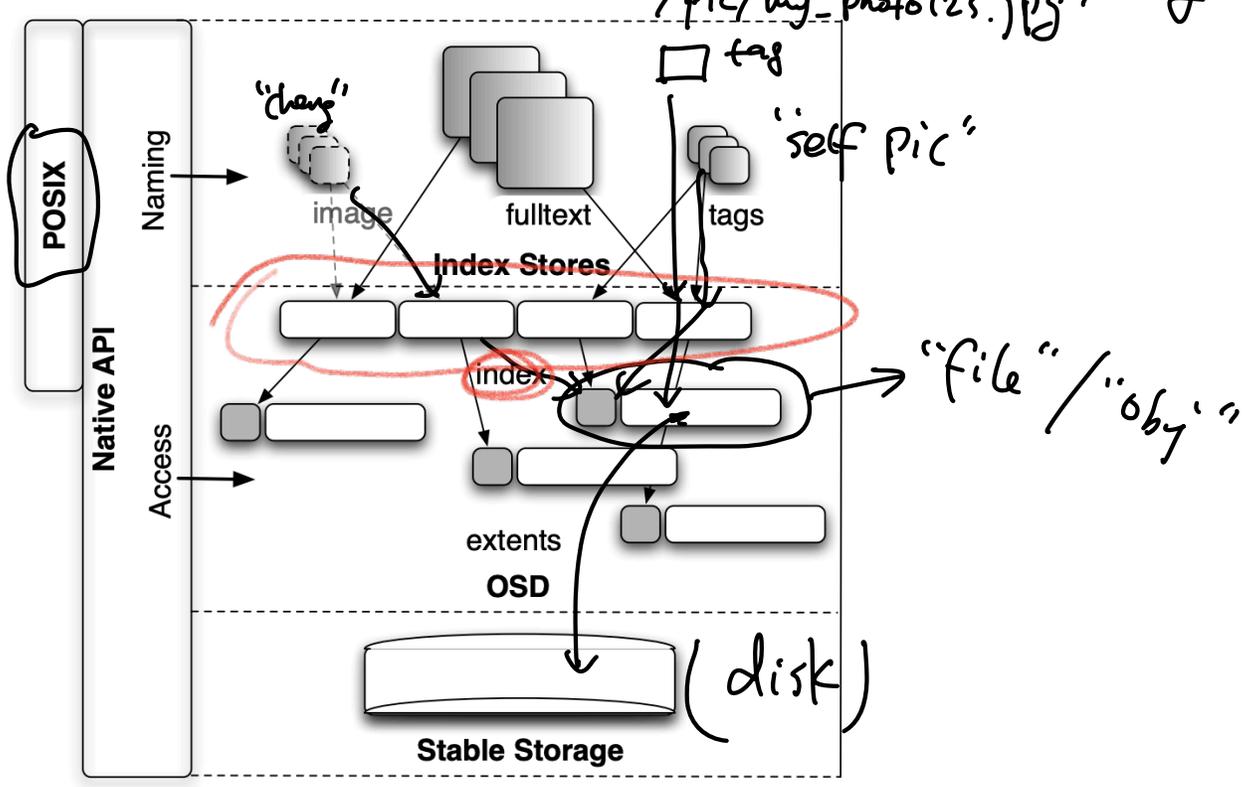
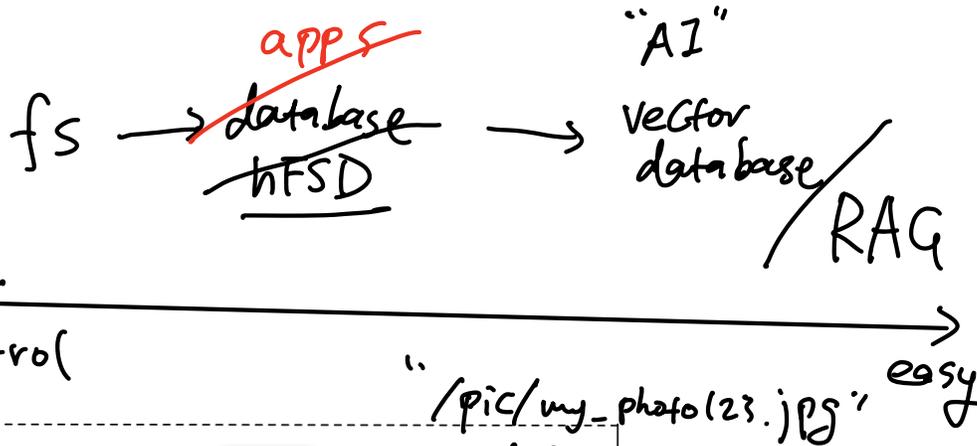
i) storage size grows

← 1000x

- 1992: 300MB disk
- 2009: 300GB disk (the paper's time, 23 years later) ["Amazon-buyable"]
- 2023: 22TB disk & 8TB SSD (14 years later than paper)
- 2025: 36TB disk & 8TB SSD (16 years later than paper)
- 2026: 36TB disk & 8TB SSD (17 years later than paper)
- 2030(?): 1000TB SSD

ii) "...they [file sizes] have not increased by the same margin."
 larger space &
 file size wasn't growing that fast =>
 more files =>
 harder to manage

iii) "Google is a verb" (or "ChatGPT" now?)
 what they want instead of where it lives



engineers
 ⚙️

Users
 👤 ... 👤

name/abc/1's ...