

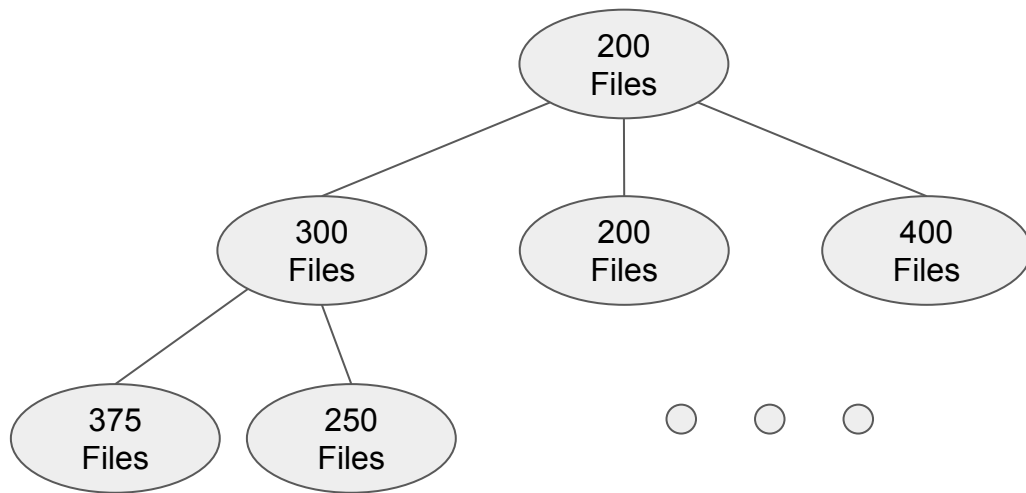
# Tree Walk Algorithm

# Idea 1: Go directory by directory

- Current approach.
- Deploy the ExifTool directory by directory.
- Possibility to work with multiple threads.

## Idea 2: Be more efficient with big directories

- Solutions that expects there to be **only directories with a huge amount of files** per directory.
- **Example:**



- **TODO:** Find out which approach is more efficient:

1) **Split the threads evenly** among the directories

**Example:** 8 directories and 8 threads => Each directory gets 1 thread

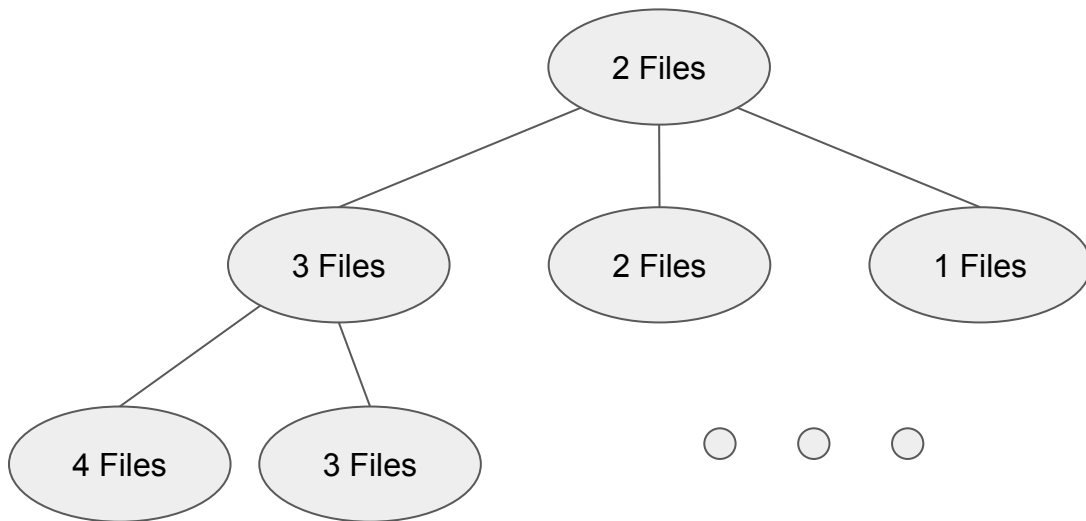
2) Go **directory by directory** and give each the maximum amount of threads.

→ Split the files in X work packages

**Example:** 8 directories and 8 threads => Give the first all 8

## Idea 3: Be more efficient with small directories

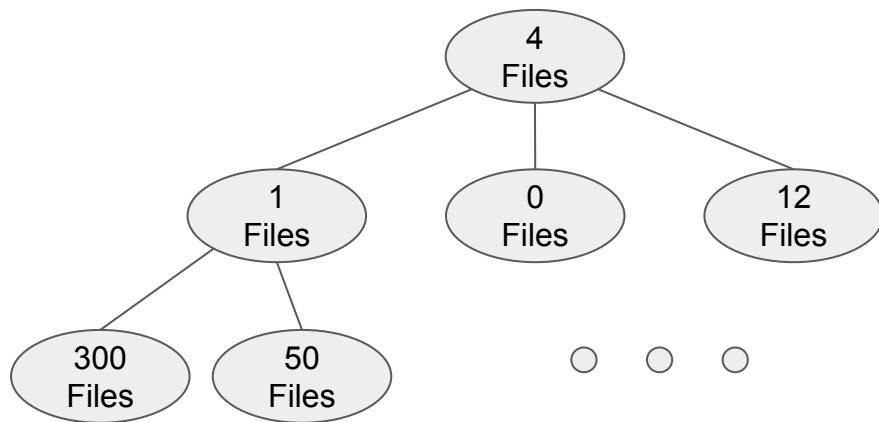
- Solutions that expects there to be only directories with a small amounts of files per directory.
- **Example:**



- **TODO:** Find out which approach is more efficient:
  - 1) Go directory by directory and give each directory one thread
  - 2) Try to combine the directories in some way to save ExifTool startup time

# Idea 4: Allround Approach

- Solution that expects a huge variance in the amount of files per directory.
- Example:



→ Try to combine idea 1 and idea 2

- Create a Hash Map that maps directories to the amount of files in the directory.

**Key:** Number of files

**Entry:** Directory

- Deploy the ExifTool on the directories with large amount of files first.
  - **TODO:** Determine X
- Use idea 2 for this.
- After all large projects are completely scanned, start working on the small ones.
- Use idea 1 for this.