

There are a lot of things we want from files

create a new file

- create a new file
- delete a file

- create a new file
- delete a file
- open a file to access it

- create a new file
- delete a file
- open a file to access it
- read data from a file

- create a new file
- delete a file
- open a file to access it
- read data from a file
- write data to a file

- create a new file
- delete a file
- open a file to access it
- read data from a file
- write data to a file
- close a file when we are done

- create a new file
- delete a file
- open a file to access it
- read data from a file
- write data to a file
- close a file when we are done
- rename a file

There are a lot of things we want from files

- create a new file
- delete a file
- open a file to access it
- read data from a file
- write data to a file
- close a file when we are done
- (rename a file)

That last one is actually a directory operation as we shall see in a moment



And directories

create a new directory

- create a new directory
- delete a directory



- create a new directory
- delete a directory
- scan a directory for a filename or directory name

- create a new directory
- delete a directory
- scan a directory for a filename or directory name
- add a file to a directory

- create a new directory
- delete a directory
- scan a directory for a filename or directory name
- add a file to a directory
- remove a file from a directory

- create a new directory
- delete a directory
- scan a directory for a filename or directory name
- add a file to a directory
- remove a file from a directory
- rename a file

And directories

- create a new directory
- delete a directory
- scan a directory for a filename or directory name
- add a file to a directory
- remove a file from a directory
- rename a file

The last three are intertwined





This all is before we come to things like

speed of access

- speed of access
- speed of update

- speed of access
- speed of update
- scalability to large numbers of files

- speed of access
- speed of update
- scalability to large numbers of files
- efficient use of disk space

- speed of access
- speed of update
- scalability to large numbers of files
- efficient use of disk space
- reliability

- speed of access
- speed of update
- scalability to large numbers of files
- efficient use of disk space
- reliability
- protection/security

- speed of access
- speed of update
- scalability to large numbers of files
- efficient use of disk space
- reliability
- protection/security
- simple backup and recovery



We shall be looking at the classical Unix filesystem as an example



We shall be looking at the classical Unix filesystem as an example

Other filesystems are similar in their principles, though modern filesystems are immensely tweaked and tuned



We shall be looking at the classical Unix filesystem as an example

Other filesystems are similar in their principles, though modern filesystems are immensely tweaked and tuned

They vary in their choice of datastructures and algorithms to implement the hierarchy for efficiency or other reasons





Early files had structure, namely records



Early files had structure, namely records

This was a hangover from early systems using things like punched cards



Early files had structure, namely records

This was a hangover from early systems using things like punched cards

A record is a fixed-size block of data, say 80 bytes



Early files had structure, namely *records*

This was a hangover from early systems using things like punched cards

A record is a fixed-size block of data, say 80 bytes

Records could only be read or written as a whole: this meant implementation on the hardware of the time was easy



It also aligned with the way data was regarded at the time: records of peoples names, job classification, salary and so on (*fields*)

 \square



It also aligned with the way data was regarded at the time: records of peoples names, job classification, salary and so on (*fields*)

They would expect an entire record to be read or written at once

 \square



It also aligned with the way data was regarded at the time: records of peoples names, job classification, salary and so on (*fields*)

They would expect an entire record to be read or written at once

Modern filesystems are *byte oriented* and you can access them however you please

 \square