File Systems - Introduction

- What is a File
- File System: Interface, Naming, Persistence
- File Access
- File Organization

What is a File?

File: A collection of data elements, grouped together for purpose of access control, retrieval, and modification.

Persistence: Often, files are mapped onto physical storage devices, usually nonvolatile.

Some modern systems define a file simply as a sequence, or stream, of data units.

What is a File System?

File: A collection of data elements, grouped together for purpose of access control, retrieval, and modification.

File System: Software responsible for

- creating, destroying, reading, modifying, moving files
- naming files
- controlling access to files
- management of resources used by files.

Logical Organization of a File

A file is perceived as an ordered collection of **records**,

$R_0, R_1, R_2, R_3, R_4, \ldots, R_n$.

Record: Contiguous block of information transferred during a logical read/write operation.

Can be of fixed or variable length.

File Access

A file is perceived as an ordered collection of **records**,

$R_0, R_1, R_2, R_3, R_4, ..., R_n$.

Sequential Access: Access records in sequence.

Example:
```
File f("foo");
while (!f.Eof()) { Record r = f.Read(); ...}
```

Random Access: Access records in arbitrary order.

Example:
```
File f("bar");
for(;;) { Record r = f.Read(rand()%N); ...}
```

Logical Organization of a File

A file is perceived as an ordered collection of **records**,

$R_0, R_1, R_2, R_3, R_4, ..., R_n$.

Record: Contiguous block of information transferred during a logical read/write operation.

Can be of fixed or variable length.

File Organization: How are the records organized in file?

- Pile, Sequential File, Indexed Sequential File, Indexed File, etc.

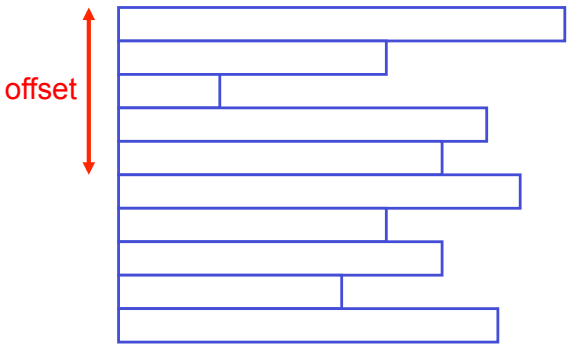
Pile

- Records stored in chronological order

Pile File

Pile

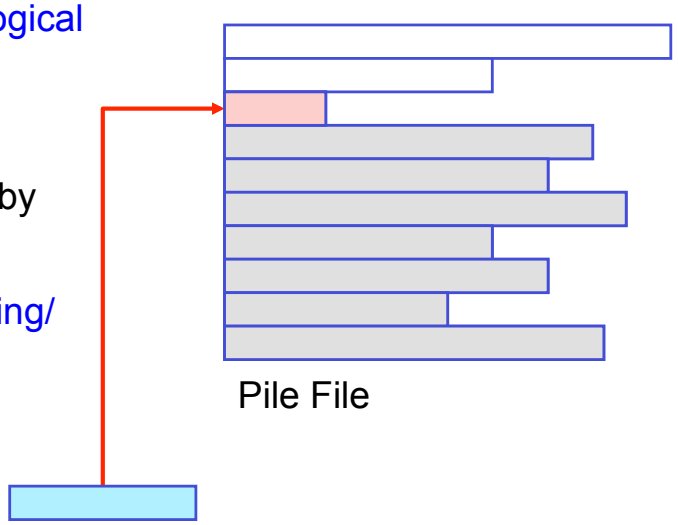
- Records stored in chronological order
- Variable-length records
- Random access to record by search of whole file.



Pile File

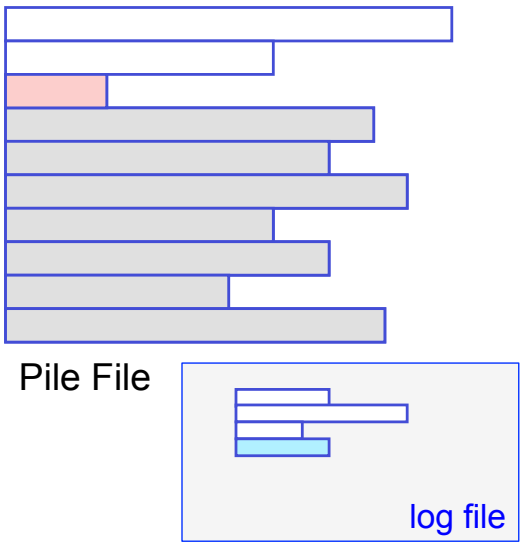
Pile

- Records stored in chronological order
- Variable-length records
- Random access to record by search of whole file.
- What about inserting/deleting/ modifying records?



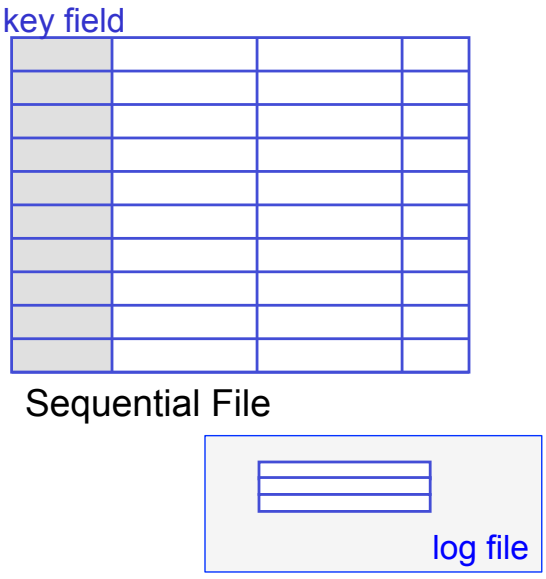
Pile

- Records stored in chronological order
- Variable-length records
- Random access to record by search of whole file.
- What about inserting/deleting/ modifying records?



Sequential File

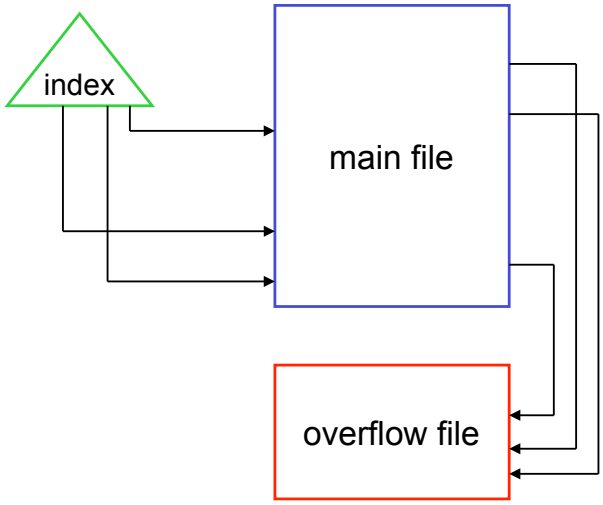
- Fixed-format records
- Records often stored in order of *key field*.
- Good for applications that process all records.
- No adequate support for random access.
- Q: What about adding new record?
- A: Separate pile file keeps *log file* or *transaction file*.



Sequential File

Indexed Sequential File

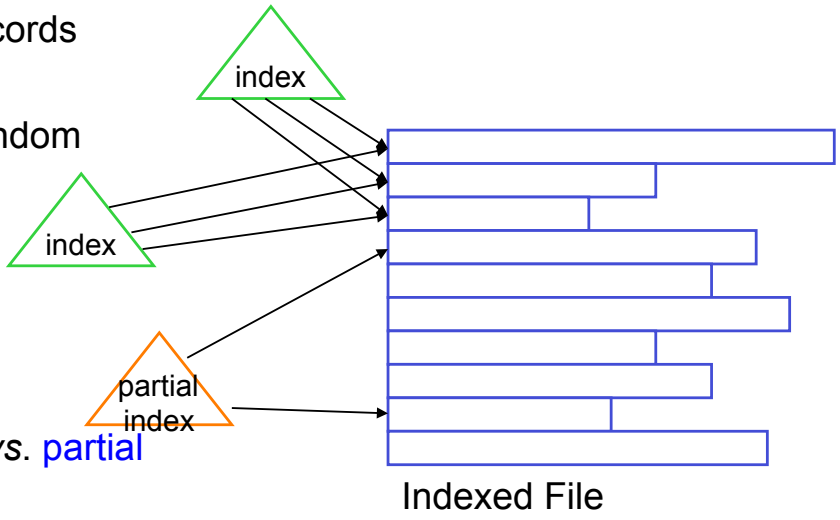
- Similar to sequential file, with two additions.
  - Index to file supports random access.
  - Overflow file indexed from main file.
- Record is added by appending it to overflow file and providing link from predecessor.



Indexed Sequential File

Indexed File

- Variable-length records
- Indices support random access
- Multiple Indices
- Exhaustive index *vs.* partial index



File Systems - Introduction

- What is a File
- File System: Interface, Naming, Persistence
- File Access
- File Organization