**Tutorial 7: File Management**

Q1. (a) List **FOUR**  functions of file managers.

|  |
| --- |
| **Functions**   * Track where each file is stored. * Determine where and how files will be stored. * Allocate each file when a user has been cleared for access to it, then record its use. * Deallocate file when it is returned to storage |

(b) Explain the factors that can affect the efficiency of file managers.

|  |
| --- |
| **Efficiency depends on:**   * how system’s files are organized (sequential, direct, or indexed sequential). * how they’re stored (contiguously, non-contiguously/link, or indexed). * how each file’s records are structured (fixed-length or variable-length). * how access to these files is controlled |

Q2. (a) Explain **FOUR (4)** characteristics which should be considered by a System Analyst when selecting the best organization for a file.

|  |
| --- |
| **Criteria**   * Volatility of data—frequency with which additions & deletions made. * Activity of file—% records processed during a given run. * Size of file. * Response time—amount of time user is willing to wait before requested operation is completed. |

(b) Describe the differences between sequential record organization, indexed sequential record organization and direct file organization

|  |
| --- |
| **Sequential record organization**   * Easiest to implement because records are stored &amp; retrieved serially, one after other. * File search: beginning until record found * To speed process some optimization features may be built into system. - Select key field from record and sort before storage - Original order must be preserved every time records added or deleted. (rewritten) * Typically used in batch applications   **Indexed sequential record organization**   * Combines best of sequential and direct access. * Created and maintained through Indexed Sequential Access Method (ISAM) software package. - Doesn’t create collisions because it doesn’t use result of hashing algorithm to generate a record’s address. Uses info to generate index file through which records are retrieved. * Divides ordered sequential file into blocks of equal size. Size determined by File Manager to take advantage of physical storage devices & to optimize retrieval strategies.   **Direct record organization**   * Uses direct access files which can be implemented only on direct access storage devices. * Give users flexibility of accessing any record in any order without having to begin search from beginning of file. * Records are identified by their relative addresses (their addresses relative to beginning of file). These logical addresses are computed when records are stored &amp; again when records are retrieved. * User identifies a field to be used as the key field. The program used to store the data follows a set of instructions called a hashing algorithm. The hashing algorithm transforms each key into a number: the record’s logical address. * This logical address is then given to the File Manager, which takes the necessary steps to translate the logical address into a physical address. |

Q3. (a) Allocation of disk space to files can be done using one of the following techniques:

Contiguous allocation

Linked allocation

Indexed allocation

Assess each allocation technique with respect to disk fragmentation and disk space usage.

|  |
| --- |
| Contiguous Storage   * Records stored one after another * Any record found once starting address, size known * Easy direct access * External fragmentation will occurs if the largest chunk of free disk space is insufficient for a request. * Wastage of disk space is substantial   Linked Allocation Storage   * Files use any available disk storage space * File records stored in contiguous manner if enough empty space * Eliminated external fragmentation as any block of the disk can be used * However, some percentage of disk space is used to store pointers   Indexed Storage   * Allows direct record access * Brings pointers together * Links every extent file into index block * Eliminated external fragmentation as any block of the disk can be used * However, some space is needed for index block for every file |

(b) Which of the allocation technique is the most efficient for random access of data? Explain your answer.

|  |
| --- |
| Linked Allocation Storage  This is very flexible in terms of file size. File size can be increased easily since the system does not have to look for a contiguous chunk of memory. This method does not suffer from external fragmentation. This makes it relatively better in terms of memory utilization. |

(c) “File allocation methods are essentially influencing the efficiency of disk scheduling algorithms”. Do you agree with this statement? Justify.

|  |
| --- |
| Yes.  Contiguous allocation – consecutive blocks are allocated to a file. Therefore, a program reading a contiguously allocated file will generate several requests that are close together on the disk, resulting in limited head movement.  Under index allocation methods, blocks are widely scattered. Thus access data in indexed allocated files will result in greater head movement. |

Q4. Describe THREE methods that are commonly used in operating systems for managing free disk space. Recommend the most suitable method to be used for a file system that applies the contiguous allocation strategy. Justify your answer.

|  |
| --- |
| **Methods of implementing the free-space list**   * bit vector * linked list * grouping * counting   Bit vector or counting. Either one is suitable method to be used for a file system that applies the contiguous allocation strategy. |

Q5. In multi-user environments, the issues of file sharing and protection are important.

A user wanting to share his file X with others has the following options:

Option A: create links from the other user’s home directories to X.

Option B: copy X to the home directories of the other users.

Compare the benefits and drawbacks of each option.

|  |  |
| --- | --- |
| **Option A** | **Option B** |
| * Any update is reflected to all users sharing data. * Save storage space * When the owner of data deletes it, dangling pointer may arise. | * Copies of data may become inconsistent and stale information may be used accidentally. * Storage waste * No dangling pointer problem |

Self-Review

Q1. What is *file organization* that is handled by File Manager? Explain both *sequential file organization* and *direct file organization* on a magnetic disk.

|  |
| --- |
| * + File organization concerned with how records are arranged and characteristics of medium used to store it.   + Sequential file organization   – store and retrieve data serially   * + - Search data from its beginning until the requested record is found   + Direct file organization   – access any data in any order without having to begin search from beginning of file (regardless of its position in the file)   * + - Data are identified by their relative addresses (their addresses relative to beginning of file) |