04 – Database Storage 2

1. Log-structured File Organization
    A. DBMS only stores log records :
    fast, because it uses sequential I/O
    easy to rollback
    re-read log to read records
2. Data Representation
    A. Records are just sequence of bytes
    use catalog to specify data
    B. Real number can be represented as floating point or fixed point
    C. Large values are not allowed in most DBMS,
    if it is allowed, DBMS uses overflow page to store it
    or, external file can be used to store large value, but it can't be manipulated
    (it is used to store data like videos)
3. System Catalogs
    A. "Meta-data of data"

4. Storage Models
    A. Workloads
        i. OLTP : TRX that updates only small amount of data
        ii. OLAP : TRX that reads large amount of data
    B. N-ary Storage Model
        i. All attributes of records are stored contiguously. (row-based)
        ii. Ideal for OLTP (data for a single record are stored in a same page)
        iii. Not good for OLAP (might read useless data)
    C. Decomposition Storage Model
        i. Same Attributes are stored contiguously. (column-based)
        ii. Ideal for OLAP (can read data needed only)
        iii. Tuple identification issue
            1. Fixed length offsets (limits of data length)
            2. Embedded tuple ids (stores more data)
        iv. Not good for OLTP
        (attributes of single data are separated in several pages)
5. What is it?
    A. HTAP (I think it is a kind of workloads… but what does it mean?)
6. Introduced papers
    A. To BLOB or Not to BLOB