Database Systems: 1st Mid

Date: 7th February 2018 **Duration: 1.5 hrs**

1. **No clarifications during the exam.**
2. **Make *reasonable assumptions* and *clearly state* them to answer *ambiguous* questions.**
3. **Show your steps. Be concise and organized.**
4. **Calculators allowed. Sharing of calculators *not* allowed.**

1) The Megatron 777 disk has the following characteristics:

1. There are 10 surfaces, with 10,000 tracks each.
2. Tracks hold an average of 1000 sectors of 512 bytes each.
3. 20% of each track is used for gaps.
4. The disk rotates at 10,000 rpm.
5. The time it takes the head to move n tracks is $1+0.001n$ milliseconds.

(a) What is the capacity of the disk?
(b) If all tracks hold the same number of sectors, what is the density of bits in the sectors of a track?
(c) What is the maximum seek time?
(d) What is the maximum rotational latency?
(e) If a block is 16,384 bytes (i.e. 32 sectors), what is the transfer time of a block? **[10]**

2) Suppose we are scheduling I/O requests for a Megatron 777 disk, and the requests are made as below, with the head initially at track 4000. At what time is each request serviced fully if:
(a) We use elevator algorithm (it is okay to start moving in either direction at first).
(b) We use first come first served scheduling. **[10]**

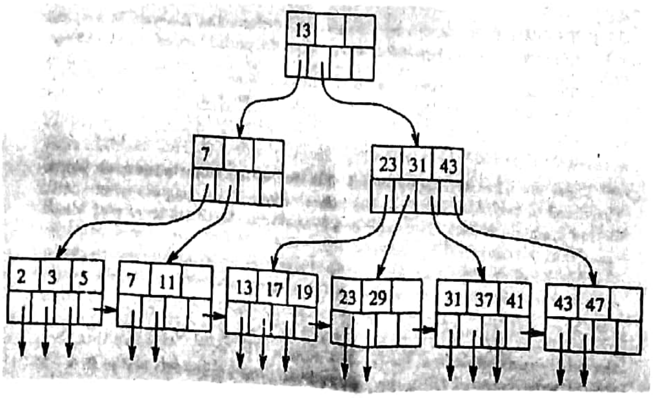
Cylinder of Request:	**1000**	**6000**	**500**	**5000**
First time available:	**0**	**1**	**10**	**20**

3) Suppose we swizzle all pointers automatically, we can perform the swizzling in half the time it would take to swizzle each one separately. If the probability that a pointer in main memory will be followed at least once is p, for what values of p is it more efficient to swizzle automatically than on demand? **[10]**

4) Suppose blocks hold either 3 records, or 10 key-pointer pairs. As a function of n, the number of records, how many blocks do we need to hold a data file and:
(a) A dense index? (b) A sparse index? **[10]**

5) Describe steps, using diagrams if necessary, to execute the following operations on the shown B+ tree:

a. Lookup all records in range 20 to 30
b. Lookup all records less than 30
c. Insert a record with key 1
d. Insert records with keys 14 through 16
e. Delete record with key 23 **[10]**



System and Network Security (CS 5470)

Mid Semester Examination - 1 (Spring 2018)

International Institute of Information Technology, Hyderabad

Time: 1 Hour and 30 Minutes Total Marks: 40

Instructions: Answer ALL questions.

This is a closed books and notes examination.

Write your answers sequentially as given in the question paper and also all the parts of a question at the same place.

No query is allowed in the examination hall.

Use of Regular Calculator is allowed.

1. **(a)** Consider the following two encryption methods.

Encryption scheme 1: Have only the encryption and decryption functions and these are kept secret to the sender and receiver only. No key is used in this method.

Encryption scheme 2: Key is being used. However, the encryption and decryption functions are made public.

Explain why the keys are necessary. Why not just choose one encryption function and its corresponding decryption function?

(b) Define the notions behind the "unconditionally secure" and "computationally secure" cryptographic schemes.

(c) Explain the chosen ciphertext attack (CCA). How is it practical for an adversary?

[2 + 2 + 2 = 6]

2. **(a)** Show theoretically that the Hill cipher is insecure against plaintext attack.

(b) Suppose Alice wishes to send a plaintext M to Bob using the RSA algorithm. Let the public key for Bob be the pair $(n, e) = (187, 7)$. Note that $187 = 17 \times 11$ and $7 \times 23 \equiv 1 \pmod{160}$. Alice uses an alphabetic set of only 11 letters and encodes them as: $A = 0$, $C = 1$, $D = 2$, $E = 3$, $I = 4$, $N = 5$, $O = 6$, $R = 7$, $T = 8$, $U = 9$, and the blank space is encoded with 10. Alice transmits the messages in blocks, where each block corresponds to two letters which are encoded into their numerical equivalent. For example, "NO" becomes "56" and then it is encrypted using the RSA algorithm.

(i) If Alice wants to send the plaintext "NO", what will be the ciphertext received by Bob?

(ii) If Bob receives the ciphertext "19", what was the original message transmitted by Alice?

[4 + (4 + 4) = 12]

3. **(a)** Give two examples where the message authentication code and one-way hash function can be used to provide confidentiality and authentication.

(b) Explain the end-to-end encryption (EEE) method for encrypting a communication channel. What are the drawbacks behind this approach?

(c) Discuss the internal error control mechanism to provide the datalink layer security. Justify the following statement–"External error control is better than internal error control".

[(2+2) + 3 + 3 = 10]

4. **(a)** To withdraw the offline password guessing attack in the password management scheme used in UNIX, the biometric-based fuzzy extractor technique can be useful. Explain the biometric-based approach to secure the password management scheme in this case.

(b) Assume that two participants, say A and B agree on the following digital signature scheme. The participant A signs a binary message m of any arbitrary length. The participant B can verify that signature by using the public key of A.

Participant A executes the following steps in the key generation part:
(i) Select two primes p and q such that $q \mid (p-1)$.
(ii) Select a random integer g with $1 < g < p-1$, such that $\alpha = g^{(p-1)/q}(\text{mod} p)$ and $\alpha > 1$.
(iii) Select a private key a, $1 \leq a \leq q-1$.
(iv) Compute $y = \alpha^a (\text{mod} p)$.
The public key of A is (p, q, α, y).

After key generation, A signs the message m as follows:
(i) Select a random secret integer k, $1 \leq k \leq q-1$.
(ii) Compute $r = \alpha^k \pmod p$, $e = H(m||r)$, and $s = (ae + k) \bmod q$, where $H()$ is a one-way cryptographic hash function.
(iii) Select two random secret integers u and v, $0 < u < q$ and $0 < v < q$, and compute $r' = r\alpha^{-u}y^{v} \bmod p$.
(iv) Compute $e' = H(m||r')$ such that $e' = e - v$ and $s' = s - u$.
A then sends the signed message $(m, (e', s'))$ to the verifier B.

Design a verification algorithm for the verifier B. Also, provide the correctness proof of the verification algorithm.

[5 + (5 + 2) = 12]

****************** End of Question Paper **********************

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

MID SEMESTER EXAMINATION-1

IIIT, Hyderabad, Spring 2018
Subject: Intro to Parallel Scientific Computing
(Code: CSE504)
Instructor: Dr. P. Kumar

08.02.2017
12:00 - 1:30

Duration: 90 minutes

Maximum Marks: 30

Instructions

There are 2 pages and 9 questions. All questions are compulsory. This paper is divided into three sections: *Section A*: Basic Matrix/Vector Computations, *Section B*: Fast Fourier Transform, and *Section C*: Rotation and Reflection Matrices. *Note:* You may first want to tick those questions that you are most confident of answering, then try the remaining ones later. Calculator is allowed.

Basic Matrix/Vector Computations (10 Marks)

1 1. Given $A \in \mathbb{R}^{n\times n}, u, v \in \mathbb{R}^{n}$. Write an efficient algorithm to compute $A + uv^T$. Determine the flop count. [1]

2. Write an efficient algorithm to compute $(xy^T)^k$, $x, y \in \mathbb{R}^n$. Determine flop count. [1]

2 3. Given a lower triangular matrix $L \in \mathbb{R}^{n\times n}$, a upper triangular matrix $U \in \mathbb{R}^{n\times n}$, and a diagonal matrix $D \in \mathbb{R}^{n\times n}$. Write an efficient algorithm to compute x from $UDLx = b$. Determine flop count. [2]

4. Given a block matrix

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1p} \\ A_{21} & A_{22} & A_{23} & \cdots & A_{2p} \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ A_{p1} & A_{n2} & A_{n3} & \cdots & A_{pp} \end{bmatrix},$$

where each block A_{ij} is a $n \times n$ matrix. Write an efficient algorithm to compute $A^T A$. Determine flop count. [2]

5. Given $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{p\times q}$. Write an efficient algorithm to compute the tensor (or Kronecker) product $A \otimes B$. [2]

2 6. Storage of structured matrices. [2=1+1]

   (a) For a given symmetric matrix, storing the matrix entries requires memory for n^2 floating point number. Show a compact storage scheme that requires storing roughly $n^2/2$ floating point numbers only.

   (b) For a banded matrix of upper bandwidth p and lower bandwidth q, write an optimal storage scheme by showing the mapping of $(i, j)th$ entry to the $(i, j)th$ entry of banded matrix.

Fast Fourier Transform (4 Marks)

3 7. If $x \in \mathbb{C}^n$ and $n = 2^t, t > 2$ then write an algorithm that computes the discrete Fourier transform $y = F_n x$. Discuss the steps involved by doing a Fourier transform for [4]

$$x = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 5 & 7 & 8 \end{bmatrix}\begin{bmatrix} 1 & 3 & 5 \\ 3 & 4 & 7 \\ 3 & 8 & 1 \end{bmatrix} = \begin{bmatrix} \cdot & & \\ & \cdot & \\ & & \cdot \end{bmatrix}$

No need to explicitely evaluate ω in your calculations.

LU Factorization (10 Marks)

8. Answer the following. [1+1+1+2+2+2+1=10]

   (a) Given a lower triangular matrix $L \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$, write an algorithm to solve $Lx = b$, which is a forward substitution.

   (b) Similarly, given an upper triangular matrix $U \in \mathbb{R}^{n \times n}$, write an algorithm to solve $Ux = b$, which is a backward substitution.

   1 (c) Write an algorithm to solve $LUx = b$ by first solving $Lt = b$ then by solving $Ux = t$. Determine flop counts.

   1 (d) Write an algorithm to compute the LU factorization for a given matrix $A \in \mathbb{R}^{n \times n}$. The L and U factors must be stored (overwritten) in A. Describe the steps of LU factorization for the following matrix

$$A = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

   (e) Describe a variant of Gaussian elimination that introduces zeros into the columns of A in the order $n : -1 : 1$, and which produces the factorization $A = UL$, where U is unit upper triangular and L is lower triangular.

   2 (f) For a rectangular matrix $A \in \mathbb{R}^{m \times n}$, write the modified LU algorithm. Show major steps to obtain LU for the following rectangular matrix:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}.$$

   1 (g) The LU algorithm for a given square matrix $A \in \mathbb{R}^{n \times n}$ fails when a zero pivot in encountered. Show that LU algorithm is guaranteed to not fail if the determinant of $A(1:k, 1:k)$ for all $1 \leq k \leq n$ is non-zero.

Rotation/Reflection Matrices (6 Marks)

9. Answer the following. [2+2+1+1=6]

   (a) Given vectors $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$. Find a rotation matrix Q such that $y = Qx$.

   (b) Given a vector $v \in \mathbb{R}^n$. Write an algorithm to compute the Householder matrix that rotates the vector

$$v = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

   to a new vector u, where

$$u = \begin{bmatrix} \|v\|_2 \\ 0 \\ 0 \end{bmatrix}.$$

   (c) Let

$$u = \begin{bmatrix} 10^{-49} \\ 10^{-50} \\ 10^{-50} \\ 10^{-10} \end{bmatrix}$$

   Given that computer precision is such that 10^{-100} or less than this value is rounded to 0, what precutations one has to take to prevent roundoff error during construction of Householder matrix for rotating u?

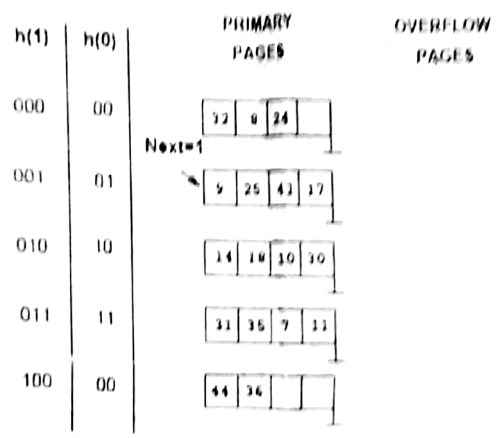
   (d) Reflect the vector $[1, 1]^T$ across the line $[\sin(30), \cos(30)]^T$ using reflection matrix (not Householder).

Database Systems: 2nd Mid

Date: 16th March 2018 Duration: 1.5 hrs

1. **No clarifications during the exam.**
2. Make *reasonable assumptions* and *clearly state* them to answer *ambiguous questions*
3. Show your steps. Be concise and organized
4. Calculators allowed. Sharing of calculators *not* allowed.

1) Consider the Linear Hashing index shown below. Assume *we split whenever an overflow page is created*. Answer the following questions about this index.



(a) What can you say about the last entry that was inserted into the index if you know that there have been no deletions from this index so far?
(b) Suppose you know that there have been no deletions from this index so far. What can you say about the last entry whose insertion into the index caused a split?
(c) Show the index after inserting an entry with hash value 4.
(d) Show the index after inserting an entry with hash value 15.
(e) Show the index after *fully* deleting the entries with hash values 36 and 44.

[10]

2) Suppose we store a relation *R(x,y)* in a grid file. Both attributes have a range of values from 0 to 1000. The partitions of this grid file happen to be uniformly spaced; for x there are partitions every 20 units, at 20, 40, 60, and so on, while for y the partitions are every 50 units.

(a) How many buckets do we have to examine to answer the range query
SELECT *
FROM *R*
WHERE 310 < *x* AND *x* < 400 AND 520 < *y* AND *y* < 730;

(b) How many disk accesses are needed to answer the above query using the grid file?

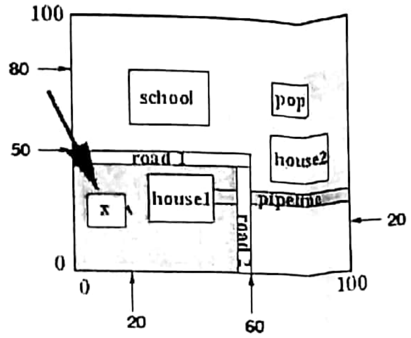
3) Suppose a database has the following schema: [10]
TRIP(fromAddrId: INTEGER, toAddrId: INTEGER, date: DATE)
ADDRESS(id: INTEGER, street: STRING, townState: STRING)
(a) Write an SQL query that returns the street of all addresses in 'Stony Brook NY' that are destination of a trip on '5/14/02'.
(b) Translate the SQL query in (a) into the corresponding "naive" relational algebra expression.
(c) Draw a query tree for the expression in (b)

(d) Translate the relational algebra expression in (b) into an equivalent expression using pushing of selections and projections.

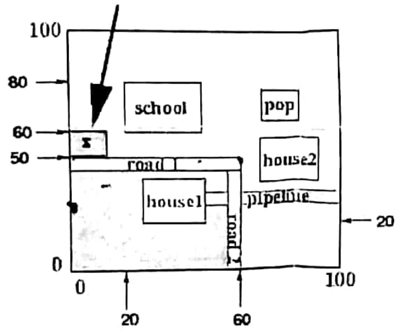
(e) Translate the relational algebra expression in (c) into a most directly corresponding SQL query.

[10]

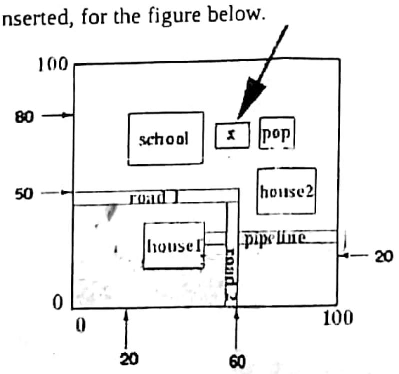
4) (a) Show the structure of an R-tree to store the regions in the figure, and show the steps of insertion of region x into that R-tree, assuming region x is inserted last.



(b) Show how region x is inserted, for the figure below.



(c) Show how region x is inserted, for the figure below.



[10]

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

END SEMESTER EXAMINATION
IIIT, Hyderabad, Spring 2018
Subject: Introduction to Parallel Scientific Computing
(Code: CSE504)

23.04.2017
03:00 - 06:00

Time: 3 hours

Instructor: Dr. P. Kumar

Maximum Marks: 100

Instructions

There is 1 page and 8 questions. All questions are compulsory. Calculators are allowed. When using calculators, keep 3 digits after decimal point. For example, if a calculation gives result 3.12345, then use the approximation 3.123. In other words, your results will be accurate upto 3 decimal places.

1. Write Grahm-Schmidt algorithm. Orthonormalize the given set of vectors using Grahm-Schmidt algorithm. [10]

$$\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\\0\end{bmatrix}.$$

2. Write an algorithm to compute the eigenvalues of a matrix. Show your steps for the following matrix. [20]

$$\begin{bmatrix}1 & 2 & 1\\2 & 1 & 1\\1 & 1 & 1\end{bmatrix}.$$

3. Write an algorithm to compute SVD of a matrix. Show you steps by computing SVD of the following matrix. [10]

$$A = \begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix}.$$

4. Write algorithm to compute QR of a given matrix A. Compute QR of the following matrix [10]

$$A = \begin{bmatrix}1 & 2 & 3\\0 & 1 & 1\\1 & 0 & 1\end{bmatrix}.$$

5. Describe an algorithm to solve least squares problem [10]

$$\min_x \|b - Ax\|$$

using QR factorization. Solve the following least squares problem.

$$\begin{bmatrix}1 & 2\\2 & 0\\1 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}1\\1\\1\end{bmatrix}.$$

6. What is the basic difference between quantum and classical parallelism? Describe Deutsche's quantum algorithm, and draw quantum circuit. Where is parallelism in this algorithm, and how it is different from classical parallel algorithm to solve the same problem as Deutsche's algorithm solves? [10]

7. Describe Grover's algorithm. Draw the circuit. Write a classical algorithm for database search. What is the difference between classical parallel search, and Grover's database search algorithm? [20]

8. Define entangled dates. Write an algorithm for quantum teleportation. Draw the circuit. [10]

Student's name:

End of exam

Database Systems

End Semester Exam

25th April 2018 Duration: 3 hrs, Marks: 90

1. **No clarifications during the exam.**
2. **Make *reasonable assumptions* and *clearly state* them to answer *ambiguous* questions.**
3. **Show your steps. Be concise and organized.**
4. **Calculators allowed. Sharing of calculators *not* allowed.**

1) The Megatron 747 disk has the following characteristics:

1. There are 4 platters providing 8 surfaces, with 8192 tracks per surface.
2. Tracks hold an average of 256 sectors of 512 bytes each.
3. 10% of each track is used for gaps.
4. The disk rotates at 3840 rpm.
5. The time it takes the head to move n tracks is 1+0.002n milliseconds.

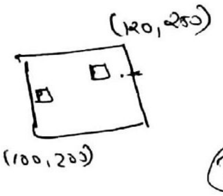
Suppose the Megatron 747 disk head is at track 1024, i.e. 1/8 of the way across the tracks. Suppose that the next request is for a block on a random track. Calculate the average time to read this block.

[10]

2) Suppose that blocks can hold either ten records or 99 keys and 100 pointers. Also assume that the average B-tree node is 70% full; i.e., it will have 69 keys and 70 pointers. The data file is a sequential file, sorted on the search key, with 10 records per block. The B-tree is a dense index. Determine: (i) the total number of blocks needed for a 1,000,000-record file, and (ii) the average number of disk I/O's to retrieve a record given its search key. You may assume nothing is in memory initially, and the search key is the primary key for the records.

[10]

3) Suppose we store a relation R(x,y) in a grid file. Both attributes have a range of values from 0 to 1000. The partitions of this grid file happen to be uniformly spaced; for x there are partitions every 20 units, at 20, 40, 60, and so on, while for y the partitions are every 50 units.



We wish to perform a nearest-neighbour query for the point (110, 205). We begin by searching the bucket with lower-left corner at (100, 200) and upper-right corner at (120, 250), and we find that the closest point in this bucket is (115, 220). What other buckets must be searched to verify that this point is the closest?

[10]

$5^2 + 15^2$

225

$\sqrt{250}$

$5 \times 5 \times 10$

$5\sqrt{10}$

4) Suppose B(R) = B(S) = 10,000, and M = 1000. What value of M would we need to compute R ⋈ S using the nested-loop algorithm with no more than 1,00,000 disk I/O's?

[10]

5) Below are the vital statistics for four relations, W, X, Y, and Z:

W(a, b): T(W) = 100; V(W, a) = 20; V(W, b) = 60
X(b, c): T(X) = 200; V(X, b) = 50; V(X, c) = 100
Y(c, d): T(Y) = 300; V(Y, c) = 50; V(Y, d) = 50
Z(d, e): T(Z) = 400; V(Z, d) = 40; V(Z, e) = 100

5, 3

3 2 2 1

4 4 5

4, 2

10 9 7

4, 4

Estimate the sizes of relations that are the results of the following expression:
(a) W ⋈ X⋈ Y ⋈ Z
(b) $\sigma_{a=10}$ (W)

3 + 3 + 3

9

[10]

1 4 2 2 4 5

④ ④ ④

3 4 5 1

9x4 = 36

4 3 5 5

105

$250 = (X - 205)^2$

$200 = x^2 - 400x + 40000$

$= x^2 - 400x - 39 =$

~~225~~

220

6) With respect to transaction serializability, consider these protocols:

(A) 2PL

(B) Validation test

(C) Timestamp ordering

Is the following schedule valid in each of the above protocols? Describe why or why not, for each one.

$R_1(a), R_2(b), R_3(c), R_1(d), R_2(c), R_3(d), W_3(c), W_1(d), W_3(d)$

[15]

7) The following is a sequence of undo-log records written by two transactions T and U: <START T>; <T, A, 10>; <START U>; <U, B, 20>; <T, C, 30>; <U, D, 40>; <COMMIT U>; <T, E, 50>; <COMMIT T>. If there is a crash and the last log record to appear on disk is:

a. <START U>
b. <COMMIT U>
c. <T, E, 50>
d. <COMMIT T>

For each of these situations, what values written by T and U *must* appear on disk?

[10]

8) Consider the following sequence of undo/redo log records: <START S>; <S, A, 60, 61>; <COMMIT S>; <START T>; <T, A, 61, 62>; <START U>; <U, B, 20, 21>; <T, C, 30, 31>; <START V>; <U, D, 40, 41>; <V, F, 70, 71>; <COMMIT U>; <T, E, 50, 51>; <COMMIT T>; <V, B, 21, 22>; <COMMIT V>. Suppose that we begin a non-quiescent checkpoint immediately after one of the following log records has been written (in memory):

a. <S, A, 60, 61>
b. <T, A, 61, 62>
c. <U, B, 20, 21>
d. <T, E, 50, 51>

For each, tell:

i. At what points could the <END CKPT> record be written, and

ii. For each possible point at which a crash could occur, how far back in the log we must look to *find all possible incomplete* transactions. Consider both the case that the <END CKPT> record was or was not written prior to the crash.

[15]