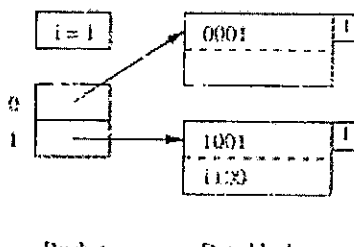## OCTOBER 2020: IN SEMESTER ASSESSMENT
B Tech 5th SEMESTER
TEST – 1

UE18CS315 (4 credits) - Database Technologies

Time: 2 Hrs	Answer All Questions	Max Marks: 60

1.	a)	In physical database design, list the actions taken to improve performance of the applications	3
	b)	How does the data block size impact performance?	3

	c)	Estimate the secondary storage required for storing 1,000,000 tuples in the customer table. Consider fixed length record format, Block size = 8kb and block header = 12 bytes. Also, estimate the number of bytes wasted in each block.
		
		create table Customer (CustId int not null primary key,
		            CustName varchar(50) not null,
		            Street varchar(50) not null,
		            City varchar(20) not null,
		            State varchar(20) not null,
		            PostalCode char(6) not null,
		            Phone char(10) not null)
			4

2.	a)	Insert 8, 5, 1, 7, 3, 12, 9, 6 into a BTree with the number of keys in each node = 3	3
	b)	Insert a record whose key hashes to the sequence 1000 into the below Extensible hash table Buckets Data blocks	3

	c)	Recommend suitable secondary indexes to be created for the following query
		
		select C.CustName,
		        C.City,
		        CO.OrderDate,
		        OD.Quantity,
		        OD.UnitPrice,
		        I.Description as ItemDescription,
		        IC.Description as ItemCategory
		from Customer C,
		        CustomerOrder CO,
		        OrderDetail OD,
		        Item I,
		        ItemCategory IC
		where C.CustId between 200000 and 201000 and
		        CO.CustId = C.CustId and
			4

OD.OrderNumber = CO.OrderNumber and
OD.ItemID = I.ItemID and
IC.CategoryID = I.CategoryID
order by 3 desc

ItemCategory (CategoryID, Description)
Item (ItemID, CategoryID, Description, UnitPrice, Discount)
Customer (CustId, CustName, Street, City, State, PostalCode, Phone)
CustomerOrder (OrderNumber, OrderDate, CustId, OrdAmt)
OrderDetail(OrderNumber, OrderLineNumber, ItemID, Quantity, UnitPrice)

3.	a)	Write the logical query for the below SQL statement Select E.Name, E.Salary, D.Dept_Name from Employee E, Department D where E.ID between 100 and 200 and D.Dept_Name in ('CSE', 'ME');	3
	b)	Consider B(R) = 100 and B(S) = 50. In one pass algorithm for binary intersection operation R ∩ S, what is the number of disk I/O's performed and the buffers needed?	3
	c)	Consider B(R) = 200 and B(S) = 100. What is the minimum number of buffers required to do a nested loop join? If M = 51, Estimate the number of disk I/O's performed in a nested loop join.	4

4.	a)	Illustrate Two-Pass duplicate elimination operation using sorting with an Example	3
	b)	Explain two pass hash based algorithm for joining relations R(X,Y) and S(Y,Z)	3
	c)	Consider B(R) = 1,000, T(R) = 20,000 and V (R, a) = 100. Estimate the number of disk I/O's for the operation $\sigma_{a=c}$ (R) for the following cases 1. There is no index on R.a 2. There is a clustered index on R.a 3. There is a non clustered index on R.a 4. R.a is the primary key	4

5.	a)	List the operations which can be executed in parallel	3
	b)	Explain the Map-Reduce Parallelism Framework	3
	c)	Suppose that a disk I/O takes 100 milliseconds. Let B(R) = 100. What is the speedup if this selection σ_c (R) is executed on a parallel machine with 5 processors assuming that the blocks are uniformly distributed?	4

6	a)	Consider R(X, Y) ⋈ S(Y, Z) where R and S reside at different nodes of a network. What is the optimum way to handle this operation?	3
	b)	What are the various types of data replication?	3
	c)	What is the role of a transaction coordinator and a transaction manager in a distributed transaction?	4