COLLEGE OF ENGINEERING, PUNE

(An Autonomous Institute of Govt. of Maharashtra)

END-SEM EXAM

Distributed Operating System

Program: F.Y.M. Tech (Computer Engineering)

Year: 2011-12 Semester I Date: 29/11/2011

Duration: 3 hrs Max. Marks: 100

Instructions:

1. Answer all questions.
2. Figures to right indicate full marks.
3. Draw neat figures wherever required.
4. Assume suitable data if necessary.

			Marks
Q.1	a	What are the commonly used semantics for ordered delivery of multicast messages?	10
	b	Suggest whether at least one or exactly one semantic should be used for each of the following applications with appropriate reasoning i) for making a request to a compilation server to compile a file ii) for making a request to a database server to update a bank account iii) for making a request to a booking server to cancel an already booked seat iv) for making a request to a file server to read a file	6
Q.2	a	i) Explain why RPC semantics is normally different from the conventional procedure call semantics.	4
		ii) What are the communication protocols for RPCs?	6
	b	Discuss the access matrix model for access control in security system.	6
Q.3	a	What are the various location policies to select destination node in load balancing?	8
	b	What are the issues involved in freezing & restarting the process in process migration?	10
Q.4	a	How does file replication differ from file caching?	4
	b	What are the commonly used approaches for handling the issue of maintaining consistency among multiple copies of the same file on different servers?	14

P.T.O.

Q.5	a	What is false sharing in DSM?	**6**
	b	Explain the following Weak consistency model and Release consistency model	**10**
	OR		
	b	Explain the distributed approach for implementing mutual exclusion in distributed system.	**10**
Q.6		What is replicated migrating block strategy for implementing sequentially consistent DSM system? What are the data-locating mechanisms used for this strategy?	**16**
	OR		
Q.6		Explain the following approaches for recording file updates in reversible manner i) file version approach ii) write ahead log approach	**16**

--------xxxxxxx----------

College of Engineering, Pune

(An Autonomous Institute of Government of Maharashtra)

End Semester Examination (2011-12)

Class: M. Tech. Semester: I Branch: Computer Engineering

Course Code: CT 521 Course Name: Machine Learning

Date: 27th Nov., 2011 Max Marks: 50 Duration: 3 hrs

Note:

1. Figures to the right indicates Marks
2. Solve any **five** questions
3. Each question carries equal marks
4. Assume data wherever necessary

Q1 A: Define
1. Hypothesis
2. Version Space
3. Semi-Supervised Learning

3

B: Comment on the futility of Bias Free learning 2

C: State and prove version space representation theorem. 5

Q2 A: Derive the Bayes' Estimator. State use for mean, variance, etc. 5

B: Write the log likelihood for a multinomial sample and derive equation 5

$$\hat{p} = \frac{\sum_t x_i^t}{N}$$

Q3 A: Use ID3 algorithm for generation of decision tree from following data. 8

Day	*Outlook*	*Temperature*	*Humidity*	*Wind*	*PlayTennis*
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

	B:	What is multivariate Regression (two line definition). Write a formula for multivariate linear model.	2
Q4	A:	Explain k-means clustering algorithm. Where it is useful, give two application examples.	5
	B:	What are smoothening parameters? How to choose smoothening parameters?	3
	C:	What is Mahalonobis Distance? Explain its utility in multivariate regression.	2
Q5	A:	Show the perceptron that calculates the parity of its three inputs.	5
	B:	Discover a simple SVM from following data that accurately discriminates the two classes. Positively labeled data points=$\left\{\binom{3}{1}, \binom{3}{-1}, \binom{6}{1}, \binom{6}{-1}\right\}$ Negatively labeled data points=$\left\{\binom{1}{0}, \binom{0}{1}, \binom{0}{-1}, \binom{-1}{1}\right\}$	5
Q6	A:	Define: i. Vapnik-Chervonenkis dimension ii. Noise iii. Covariance iv. Entropy v.	2
	B:	Explain Hidden Markov model. Give suitable example.	5
	C:	What types of problems are addressed by Genetic algorithms? Give the details of the Genetic operators.	3

COLLEGE OF ENGINEERING ,PUNE.

An Autonomous Institute of Government of Maharashtra, Pune - 411005)

End Semester Examination

Semester – I

Sub: Topics in Databases

Year : F .Y.M.Tech **Branch : Computer Engg.**
Academic Year : 2011-12 **Date :**
Duration: 3 hrs **Max. Marks :** 50

Instructions: 1. All questions are compulsory.
2. Figures to right indicate full marks.

Q1 For each of the following statements, indicate if it is **TRUE** or **FALSE**. If the statement is false, give a simple counter-example. To construct counter-examples, use schedules that use the following two transactions **10**

T1: w1(A) w1(B) r1(C) c1
T2: w2(A) r2(B) w2(C) c2

a) All serial schedules are conflict-serializable.

b) All avoids-cascading-rollback schedules are recoverable.

c) All strict schedules are serial.

d) All recoverable schedules are conflict-serializable.

e) All strict schedules are conflict-serializable.

f) Any schedule produced by a lock scheduler using shared and exclusive locks is conflict serializable.

g) Multi-key indexes are better than partitioned hash tables for answering range queries on multiple key attributes.

h) Undo logging is suitable for use with archiving, though it's less efficient than redo and undo/redo logging.

i) No deadlocks can occur under 2PL.

j) Sequence pointers are used in leaf nodes of B+ trees to speed up range queries.

Q 2

A. Device a timestamp based protocol which avoids phantom phenomenon. 3

B. ARIES assumes space for LSN in each page. Suggest a technique(supporting physical redos) to handle the situation where an entire page may be used by an object leaving no space for the LSN. 4

OR

Discuss advantages and disadvantages of Map Reduce

C. Relation R has 4000 records. Suppose a clustered B+ tree index is built on R's key attribute. A block can hold a B+ tree node (interior or leaf node) with 20 keys and 21 pointers. Calculate minimum number of blocks needed to hold the entire B+ tree index. 3

Q 3

A. Consider relations r1(A,B,C), re(C,D,E), and r3(E,F), with primary keys A,C,E respectively. Assume r1, r2 and r3 have 1000, 1500 and 750 tuples respectively. Estimate the size of r1 ⋈ r2 ⋈ r3, and give an efficient strategy for computing the join. 4

OR

What is the effect on cost of merging runs if the number of buffer blocks per run is increased, keeping overall memory available for buffer runs fixed.

B. Examine the schedule given below. There are four transactions, T1, T2, T3, and T4. 6

	T1	T2	T3	T4
1				
2	READ salary			READ tax
3				
4		READ tax		WRITE tax
5		WRITE tax		
6	READ tax			
7	WRITE salary			
8			READ salary	
9	WRITE tax			
10			WRITE salary	
11				READ salary
12				WRITE salary

a) Draw the precedence graph for this schedule.
b) What is the equivalent serialization order for this schedule? If no order is possible, then state 'none'.
c) Assume that transaction T4 did not run at all. What is the precedence graph in this case?
d) What is the equivalent serialization order for this second schedule? If no order is possible, then state 'none'.

Q 4

A. Distinguish serial schedule from serializable schedule. 2

B. Consider a query optimizer that uses histograms. In particular, the following information is known about the attribute A of relation R where R has the attributes R(A,B,C). Attribute A is of type integer. 8

- There are 100 tuples with A values between 1 and 10.
- There are 300 tuples with A values between 11 and 20.
- There are 200 tuples with A values between 21 and 30.

(a) Suppose that for each range, the possible A values of tuples are uniformly distributed. For example, a tuple with an A value within the range of 21 and 30 has one of the 10 possible values with equal probability. How many tuples are expected in the result of the query $\sigma A=7(R)$? Expected number of tuples:

(b) Now suppose for the same R we have the following information.

- There are 3 unique A values (uniformly distributed) within the A values between 1-10.
- There are 5 unique A values (uniformly distributed) within the A values between 11-20
- There are 10 unique A values(uniformly distributed)within the A values between 21-30

How many tuples are expected in the result of the query $\sigma A=17(R)$ when 17 is one of the 5 unique values for the second range? Expected number of tuples:

(c) Using the same R from (b), now consider the query $R \bowtie A=D\ S$, where S has attributes S(D,E, F). Assume that the following information is known about the attribute D of relation S. Attribute D is of type integer.

- There are 50 tuples with D values between 1 and 10. In this range, there are 5 unique D values uniformly distributed.
- There are 100 tuples with D values between 11 and 20. In this range, there are 5 unique D values uniformly distributed.
- There are 200 tuples with D values between 21 and 30. In this range, there are 5 unique D values uniformly distributed.

Assuming that the containment of value sets assumption holds when joining R and S, how many tuples are expected in the answer? Expected number of tuples:

Q 5

A. Compare deferred- and immediate-modification versions of the log-based recovery scheme. 3

B. Consider a database that uses a validation mechanism for concurrency control. There are only 4 kinds of transactions in the system. The read, write actions of the 4 kinds of transactions are given below. 7

(i): r(A) r(B) w(A) w(C) w(E)
(ii): r(C) w(F) w(B)
(iii): r(D) r(B) w(B)
(iv): r(c) r(D)

Note that (i) - (iv) are transaction *types* and not transactions. In particular we can have two different transactions of the same type. Two transactions are said to execute *concurrently* if at some point of time, both the transactions have started, but neither has finished.

a) Determine for each pair of transaction types *(m, n)* (including the case $m = n$), if the transaction of type *n* can be aborted due to a transaction of type *m* when executing concurrently (in the absence of any other transaction). Please indicate your answer in the table below by placing a **YES** in column *n* and row *m* if a transaction of type *n* can be aborted due to a transaction of type *m*, and **NO** otherwise.

	i	ii	iii	iv
i				
ii				
iii				
iv				

b) What is the condition under which two transactions can *never* execute concurrently in the validation scheme (without one of them being rolled back)? Express your answer in terms of read-sets ($RS(T_1)$, $RS(T_2)$) and write-sets ($WS(T_1)$, $WS(T_2)$) of the two transactions T_1 and T_2.

OR

Write detailed note on HDFS architecture

COLLEGE OF ENGINEERING, PUNE

(An Autonomous Institute of Govt. of Maharashtra)

END SEM EXAM

F.Y.M.Tech - Computer Engineering

CT-503 – Advanced Computer Networks

Academic Year: 2011- 12
Semester -I

Timing: 3 hrs
Max. Marks: 50

Instructions:

1. Answer any Five Questions.
2. Draw neat diagrams wherever necessary.
3. Make appropriate assumptions if required.
4. Figures to the right indicate full marks.

			Marks
Q. 1	A)	How do you intend to solve the problem of bottleneck at hierarchical level "i" in the processor hierarchy? Will adding another level in hierarchy help why or why not?	05
	B)	Discuss the limitations due to scaling by number of processors as against scaling by amount of memory.	05
Q. 2	A)	Discuss applications that can be designed using NICs in promiscuous mode.	03
	B)	Suppose you are network Designer for a dedicated network protocol device with 4 ports and 4 network cores. Which one of the below approach you will adopt and why? Keep in mind in future we may have to support multiple protocols. 1. One thread for each layer 2. One thread for each protocol 3. Multiple threads for each protocol 4. Threads for protocols plus timer management thread 5. One thread for each packet 6. One thread per port	03
	C)	What is Data Fragmentation?	01
	D)	What is data line with reference to hardware architecture?	01
	E)	Differentiate between dynamic and static packet classification.	02

Q. 3	A)	Differentiate MPLS Network technology with ATM Networks and pure IP Networks.	03
	B)	Discuss various applications in security management space that influence network management.	03
	C)	Summarize the set of operations associated with egress processing.	02
	D)	What is CAM and how it supports rapid data storage?	02
Q. 4	A)	When multiple equal cost routes to a destination exist OSPF may distribute traffic equally among routes. This is load balancing. What effect does such load balancing have on transport layer protocol such as TCP? Discuss.	05
	B)	Name any one distributed network management architecture.	01
	C)	What is tunneling?	01
	D)	State three important differences in characteristics between authentication and encryption.	03
Q5	A)	Illustrate network security requirements in a wireless environment.	05
	B)	Messages arrive at a switching centre for a particular outgoing communication line in a Poisson manner with a mean arrival rate of 180 messages /hr. Message length in distributed exponentially with a mean length of 14400 characters line speed in 9600 bps 1] What is the mean waiting time in the switching activity? 2] How many messages will be waiting in the switching centre for transmission on the average?	03
	C)	What is ATM cell size?	02
Q6	A)	Describe drawbacks of adaptive routing.	02
	B)	Illustrate requirements for Network design when applied by storage protocols.	04
	C)	Enumerate different strategies to resolve the below problems in flow control: 1. Protection against overrun errors in a reliable way. 2. Protection against message loss.	04

College of Engineering, Pune
(An Autonomous Institute of Government of Maharashtra, Pune – 411 005)
M. Tech (Computer Engg)
Sub : **Advanced Computer Architecture (ACA)**

21st November 2011

Year: **F. Y. M.Tech** Branch: **Comp. Engg.**
Academic Year: **2011-12** Max. Marks: **100**
Duration: **Three Hours.**
Instructions: **Solve any five questions**

Q.1. (i) Explain the characteristic differences between GPU & CPU architectures.
(ii) Draw the architectural block diagram of Nvidia GeForce 6800.
(iii) Explain the salient features of Nvidia's Tesla GeForce 8800 architecture & CUDA Programming Environment.
(iv) Enlist the High Lights of 38th List of "TOP 500 Supercomputers" released recently.
(20)

Q.2. (i) Compare Message Passing Systems : MPI & PVM .
(ii) Explain the various Collective Communication calls in MPI.
(iii) What are MPI–2 extensions?
(iv) Explain the salient features of POSIX Threads Model and OPEN MP Standard.
(20)

Q.3. (i) Explain in details IBM SP2 System Architecture and its salient features.
(ii) Explain in details the operation of High Performance Switch (HPS) and Communication Adapter deployed in IBM SP2 System.
(20)

Q.4. (i) Draw the block diagram of Intel's 'Nehalem' Architecture
(ii) Explain Intel's Turbo Boost technology.
(iii) List down the various X86-64 Operating Modes in Tabular Form.
(iv) Explain the various versions of Hardware Multithreading.
(20)

Q.5. (i) What is SCI?
Explain SCI Interconnect topologies, link and packet format and Node Interface structure in details.
(ii) With suitable block diagram, explain :
SEQUENT's CC-NUMA Q-2000 Architecture in details.
(20)

Q.6. Explain in details the Case Study of C-DAC's "PARAM Yuva" System & its Applications.
(20)

****