

**Directions (Please read carefully):** Please make sure you write your Student ID and Name. You are required to answer all questions. There is only one correct answer for multiple choice questions, unless you are asked to choose specific number of right answers. Circle your answer for easy readability.

*Multiple Choices (Each carry 2 marks):*

1. A software standard is a:  
a. Certification process  
☒ b. Specification  
c. License  
d. None of the above
2. A software platform is a technology stack that includes:  
a. Interface languages  
b. Communication protocols  
c. Run time components  
☒ d. All of the above
3. IBM Web sphere MQI is an example of:  
a. Communication layer  
b. Security certificate  
☒ c. An application programming interface  
d. None of the above
4. Typical tasks of a data base administrator include:  
a. Creating objects (tables, indexes, views, etc.)  
b. Grant privileges (select, update, delete, etc.)  
c. Maintenance (load, statistics, copy, recover, etc.)  
☒ d. All of the above
5. In RDBMS, the "automatic enforcement" of referential constraints is called:  
a. Data Integrity  
☒ b. Referential Integrity  
c. 2<sup>nd</sup> form of Normalization  
d. 3<sup>rd</sup> form of Normalization
6. If you want to list the employees whose salary is the highest, the outer query's WHERE Clause may be coded:  
a. WHERE SALARY > (SELECT SALARY FROM EMPLOYEE..)  
b. WHERE SALARY = (SELECT MAX (SALARY) FROM EMPLOYEE..)  
c. WHERE SALARY > ALL (SELECT SALARY FROM EMPLOYEE..)  
☒ d. None of the above
7. A data base management system provides:  
a. Logging & Recovery  
b. Query Optimization  
c. Data security & integrity  
☒ d. All of the above

79  
100

8. Which one is not an Inter Process Communication(IPC) mechanism:

- a. Signals
- b. Unnamed Pipes
- c. Message queues
- ☒ d. Net BEUI

9. The basic principle of Application Integration is to(choose 3 answers):

- ☒ a. Let information flow freely across application
- b. Let application reuse data and knowledge
- ☒ c. Let application connect to customers, partners and suppliers(external applications)
- ☒ d. Let developers create a single application out of many

10. The following is not correct about Semaphores:

- a. internally implemented as a set
- b. guarding access to resources
- c. used for coordinating activities of processes
- ☒ d. None of the above

11. Which of the following is not an example of Program to program communication:

- a. Conversational(sockets)
- b. Call-and-return(RPC)
- c. Messaging(MQ)
- ☒ d. Distributed File system (Hadoop)

12. Which of the following statement is true about HDFS NameNode :

- a. Maintains file system tree and meta data for all files/directories in the tree
- b. NameNode will crash if it runs out of RAM
- c. Coordinates access to DataNodes but data never goes on NameNode
- ☒ d. All of the above

13. Match the following Apache Open source products which constitute an Open ESB platform:

a) Apache Active MQ - 3	1) Business Process orchestration
b) Apache Camel - 4	2) Restful Web Services
c) Apache CXF - 1	3) Message Queue
d) Apache ODE - 2	4) Enterprise Integration pattern

14. An Enterprise Service Bus or ESB does not provide

- ☒ a. Security
- b. Routing
- c. Service Level Agreement
- d. Dynamic generation of test

15. In Service Oriented Architecture, a service is:

- a. An architecture style
- b. A specification
- ☒ c. A repeatable business task.
- d. None of the above

16. Which of the following statements about a Linux process is not correct?

- a. A process is an instance of a program running on the computer
- b. A process owns system resources such as child threads, memory, files, data etc.-
- ☒ c. A process normally runs in kernel mode
- d. In order for a process to execute system calls it needs to change to system/kernel mode



17. Consider you have an EMPLOYEE table and a DEPARTMENT table. The Foreign key EMPLOYEE.DEPARTMENTNAME was defined with ON DELETE SET NULL referential integrity rule. If a DEPARTMENT called "HR" is deleted from the DEPARTMENT table what effect it will have on EMPLOYEE table?

- a. All the employee belong to HR dept. will be removed from the EMPLOYEE table
- ☒ b. All the employee belong to HR dept. will have their new dept. set to null values
- c. Nothing will happen as this integrity rule is restricted
- d. The effect will depend on the privileges on the table

18. If multiple threads are using the same shared memory then thread synchronization can be achieved using:

- a. Event
- b. Mutex
- c. Semaphore
- ☒ d. Any of the above

True / False (Each carry 2 marks, please circle either True or False as your answer):

- 1. In relation database table, all column names must be unique. – True / ☒ False
- 2. The following SQL query is syntactically correct – ☒ True / False  
`SELECT WORKDEPT, AVG(SALARY) FROM EMPLOYEE WHERE AVG(SALARY) > 20000 GROUP BY WORKDEPT`
- 3. In Linux operating system, Processes with a higher nice value get less CPU time than processes with a low nice value. – ☒ True / False
- 4. "CREATE PRIMARY KEY" and "CREATE FOREIGN KEY" statements can be used to define a primary key and a foreign key respectively – True / ☒ False
- 5. Signals, the most common IPC mechanism are delivered in user mode and received in system mode. – ☒ True / False
- 6. Message queues are like your phone voice mail in the sense that, messages can be removed from front, middle or anywhere from the queue. – ☒ True / False
- 7. IBM's Deep Q&A system called Watson quickly searched Internet to provide answers to Jeopardy questions – True / ☒ False
- 8. In HDFS, Datanode sends Heartbeat to Namenode periodically. When Namenode does not get heartbeat from a datanode it marks that datanode as dead and stops forwarding any I/O requests to that datanode. – ☒ True / False
- 9. Map Reduce is a **Programming Language** for distributed computing – True / ☒ False
- 10. WHERE clause predicates must be separated from each other by commas. – ☒ True / False
- 11. Data Clustering helps improve query performance by Storing logically related records together hence reducing physical I/O – ☒ True / False
- 12. In a Map Reduce job, tasks are scheduled by Task Tracker and executed by Job tracker. – ☒ True / False

### Short Answer Questions

1. What is Complex Event Processing (CEP)? Give one problem example that CEP can solve. [4 Marks]

- If there are various number of events are processing at a certain point of time, then it is called Complex event processing.

→ Complex event processing means the execution of the events that are sharing the same resources at a time.

- The coordination of various events can perform in a sequential manner.

ex 2

2. Does MapReduce programming model provide a way for reducers to communicate with each other? Justify your answer. [4 Marks]

- No.

- MapReduce programming model conceptualizes the mapping of similar chunks of data and reducing them with the similar set of data, by assigning a key to that same type of data.

- Once, the reduction of chunks of data assigned to a specialized key, then the whole chunk is represented as that defined keyword.

3. Does Linux operating system provide an undelete utility to recover deleted files? Justify your answer with why / why not. Is there any way to recover accidentally deleted files on Linux? [4 Marks]

- No, Linux doesn't provide an undelete utility to recover deleted files.

- Because, the deletion process executes in kernel mode.

- Ways to recover accidentally deleted files:-

↳ We can search the reference of the file, as a file is never deleted, only its physical value is being removed when deletion occurs.

↳ Or else, we can change the permission of the file in prior.



4. In Map Reduce paradigm, define "map" and "Reduce" step with an example [6 Marks]

- "Map" maps i.e. searches the similar types of data which are having similar data value sets, and gathers them into chunks of data.
- "Reduce" allocates a specialized key for a time these similar chunks of data which are pre-mapped.
- For example, there is an online survey on the pros and cons of the recently launched product.
- Public will like it, will dislike it or won't bother to comment.
- So, here, "map" will gather the common likes or dislikes into a chunk of data which are certainly having similar meaning.
- "Reduce" will assign a specialized keyword for the likes, or dislikes in a whole, as in assigning a overall survey that these many people likes or dislikes the product.

5. In Linux operating system boot process [4 Marks]

a. what is the first process to execute when the Linux kernel is loaded?

- ~~First~~ It will mount the root file system.

b. What could happen, if kernel cannot execute this process?

- It will show an error of "missing file".

6. Define a Linux process. Name at least 4 resources owned by a process. [4 Marks]

- Linux process is a sequential execution of program.
- Process owns files, memory, data, threads, etc.

7. All Linux processes are started by another process, what are the two reasons a process can be terminated? [4 Marks]

Process can be terminated by the following two reasons:-

- If it is interrupted by another process.
- If it is terminated by itself.

8. Please study the STUDENT\_COURSE table below in 1NF (1<sup>st</sup> Normalized form) [10 Marks]

STUID	COURSEID	COURSE DESC
000030	141	Database concepts
000250	002	Operating Systems
000250	011	Project Mgmt
000250	067	Operation Research
000270	415	Cloud computing
000270	287	Software Testing
000300	011	Project Mgmt
000300	032	User Design

What are 3 problems with the STUDENT\_COURSE table? List them below:

- There is no certain primary key.
- Redundancy of fields.
- There is no certain foreign key from which we can relate.

Also draw the 2<sup>nd</sup> Normal form version of the above table in the space below:

COURSEID	COURSE DESC	COURSE ID	STUID
141	Database Concepts	141	000030
002	Operating Systems	002	000250, 000300
011	Project Mgmt	011	000250
067	Operation Research	067	000250
415	Cloud Computing	415	000270
287	Software Testing	287	000270
032	User Design	032	000300

Here, the primary key becomes

- Here, the new COURSEID becomes the primary key in table-1 and foreign key in table-2.

...asked to choose specific number of right answers. Circle your answer for multiple  
Multiple Choices (Each carry 2 marks):

1. A software standard is a:
  - a. Certification process
  - ☒ b. Specification
  - c. License
  - d. None of the above
2. A software platform is a technology stack that includes:
  - a. Interface languages
  - b. Communication protocols
  - c. Run time components
  - ☒ d. All of the above
3. IBM Web sphere MQI is an example of:
  - a. Communication layer
  - b. Security certificate
  - ☒ c. An application programming interface
  - d. None of the above
4. Typical tasks of a data base administrator include:
  - a. Creating objects (tables, indexes, views, etc.)
  - b. Grant privileges (select, update, delete, etc.)
  - c. Maintenance (load, statistics, copy, recover, etc.)
  - ☒ d. All of the above
5. In RDBMS, the "automatic enforcement" of referential constraints is called:
  - a. Data Integrity
  - ☒ b. Referential Integrity
  - c. 2<sup>nd</sup> form of Normalization
  - d. 3<sup>rd</sup> form of Normalization
6. If you want to list the employees whose salary is the highest, the outer query clause may be coded:
  - a. WHERE SALARY > (SELECT SALARY FROM EMPLOYEE)
  - ☒ b. WHERE SALARY = (SELECT MAX (SALARY) FROM EMPLOYEE)
  - c. WHERE SALARY > ALL (SELECT SALARY FROM EMPLOYEE)
  - d. None of the above
7. A data base management system provides:
  - a. Logging & Recovery
  - b. Query Optimization
  - c. Data security & integrity
  - ☒ d. All of the above



8. Which one is not an Inter Process Communication(IPC) mechanism:

- a. Signals
- b. Unnamed Pipes
- c. Message queues
- ☒ d. Net BEUI

9. The basic principle of Application Integration is to(choose 3 answers):

- ☒ a. Let information flow freely across application
- ☒ b. Let application reuse data and knowledge
- ☒ c. Let application connect to customers, partners and suppliers(external applications)
- d. Let developers create a single application out of many

10. The following is not correct about Semaphores:

- a. internally implemented as a set
- b. guarding access to resources
- c. used for coordinating activities of processes
- ☒ d. None of the above

11. Which of the following is not an example of Program to program communication:

- a. Conversational(sockets)
- b. Call-and-return(RPC)
- c. Messaging(MQ)
- ☒ d. Distributed File system (Hadoop)

12. Which of the following statement is true about HDFS NameNode :

- a. Maintains file system tree and meta data for all files/directories in the tree
- b. NameNode will crash if it runs out of RAM
- c. Coordinates access to DataNodes but data never goes on NameNode
- ☒ d. All of the above

13. Match the following Apache Open source products which constitute an Open ESB platform:

a) Apache Active MQ (3)	1) Business Process orchestration
b) Apache Camel (4)	2) Restful Web Services
c) Apache CXF (2)	3) Message Queue
d) Apache ODE (1)	4) Enterprise Integration pattern

14. An Enterprise Service Bus or ESB does not provide

- a. Security
- b. Routing
- ☒ c. Service Level Agreement
- d. Dynamic generation of test

15. In Service Oriented Architecture, a service is:

- a. An architecture style
- b. A specification
- ☒ c. A repeatable business task
- d. None of the above

16. Which of the following statements about a Linux process is not correct?

- a. A process is an instance of a program running on the computer
- b. A process owns system resources such as child threads, memory, files, data etc.
- ☒ c. A process normally runs in kernel mode
- d. In order for a process to execute system calls it needs to change to system/kernel mod



4. In Map Reduce paradigm, define "map" and "Reduce" step with an example [6 Marks]

sol: Suppose you are searching for a number of times the word "and" has occurred in a very large file of Tb. The map step what it does is it works on the small chunk of large file that is assigned to slave node and generates key value pairs. Like this there will be many map tasks and each one produces its own key, value.

Now after this the Reduce step what it does it collects the data on the output of map tasks and generates a consolidated output. Suppose there are two tasks and they gave output as ("and", 3), ("and", 7)

("and", 4) respectively. The reduce consolidates and generates ("and", 14) which

5. In Linux operating system boot process [4 Marks]

a. what is the first process to execute when the Linux kernel is loaded?

Init process is the first process to execute when kernel

b. What could happen, if kernel cannot execute this process?

If kernel cannot execute this process, it just freezes and drops into root shell.

6. Define a Linux process. Name at least 4 resources owned by a process. [4 Marks]

→ In linux, a process is a program in execution is the one that owns resources.

→ The four resources owned by a process are

a) memory

d) address space.

8. Please study the STUDENT\_COURSE table below in 1NF (1<sup>st</sup> Normalized form) [10 Marks]

STUID	COURSEID	COURSE DESC
000030	141	Database concepts
000250	002	Operating Systems
000250	011	Project Mgmt
000250	067	Operation Research
000270	415	Cloud computing
000270	287	Software Testing
000300	011	Project Mgmt
000300	032	User Design

What are 3 problems with the STUDENT\_COURSE table? List them below.

- If the course description for a course changes, you need to make changes in all rows.
- This table ~~only~~ only shows those courses taken by a student, and does not show those courses that are offered but still offered.
- ~~If the course~~ If the course is no longer offered, you need to delete all those rows.

draw the 2<sup>nd</sup> Normal form version of the above table in the space below.

→ here the 2<sup>nd</sup> normalized version is:

Student id	course id
------------	-----------



### Short Answer Questions

1. What is Complex Event Processing (CEP)? Give one problem example that CEP can solve. [4 Marks]

→ A set of events that ~~occ~~ occur which triggers ~~an event~~ and which trigger a new event is called CEP.

→ One problem that can be solved by CEP is that of credit card fraud or even that of a false insurance claim where one person has submitted claim report from of two different hospitals within duration and those places are long apart and the person

2. Does MapReduce programming model provide a way for reducers to communicate with each other? Justify your answer. [4 Marks]

4 Yes MapReduce programming model provides a way for reducers to communicate with each other.

The reducers that have worked on output of the map function communicate using a queue which consolidates the output to produce the consolidated output.

3. Does Linux operating system provide an undelete utility to recover deleted files? Justify your answer with why / why not. Is there any way to recover accidentally deleted files on Linux? [4 Marks]

4

17. Consider you have an EMPLOYEE table and a DEPARTMENT table. The Foreign key EMPLOYEE.DEPARTMENTNAME was defined with ON DELETE SET NULL referential integrity rule. If a DEPARTMENT called "HR" is deleted from the DEPARTMENT table what effect it will have on EMPLOYEE table?
- All the employee belong to HR dept. will be removed from the EMPLOYEE table
  - All the employee belong to HR dept. will have their new dept. set to null values
  - Nothing will happen as this integrity rule is restricted
  - The effect will depend on the privileges on the table

18. If multiple threads are using the same shared memory then thread synchronization can be achieved using:
- Event
  - Mutex
  - Semaphore
  - Any of the above

True / False (Each carry 2 marks, please circle either True or False as your answer):

1. In relation database table, all column names must be unique. **True / False**
2. The following SQL query is syntactically correct **True / False**  
`SELECT WORKDEPT, AVG(SALARY) FROM EMPLOYEE WHERE AVG(SALARY) > 20000 GROUP BY WORKDEPT`
3. In Linux operating system, Processes with a higher nice value get less CPU time than processes with a low nice value. **True / False**
4. "CREATE PRIMARY KEY" and "CREATE FOREIGN KEY" statements can be used to define a primary key and a foreign key respectively. **True / False**
5. Signals, the most common IPC mechanism are delivered in user mode and received in system mode. - **True / False**
6. Message queues are like your phone voice mail in the sense that, messages can be removed from front, middle or anywhere from the queue. - **True / False**
7. IBM's Deep Q&A system called Watson quickly searched Internet to provide answers to Jeopardy questions - **True / False**
8. In HDFS, Datanode sends Heartbeat to Namenode periodically. When Namenode does not get heartbeat from a datanode it marks that datanode as dead and stops forwarding any I/O requests that datanode. **True / False**
9. Map Reduce is a **Programming Language** for distributed computing - **True / False**
10. WHERE clause predicates must be separated from each other by commas. - **True / False**
11. Data Clustering helps improve query performance by Storing logically related records together hence reducing physical I/O **True / False**
12. In a Map Reduce job, tasks are scheduled by Task Tracker and executed by Job tracker. - **True / False**



CMPE 172-272 Mid-Term Exam 10/12/2011 6:15-8:15pm

Student ID: 008051797

Student Name: KOTI KRISHNA SUDHAKAR  
TIRUVEEDHOLA

Directions (Please read carefully):  
Please make sure you write your Student ID and Name. You are required to answer all questions. There is only one correct answer for multiple choice questions. Circle your answer for easy readability.

Multiple Choices (Each carries 4 mark):

- 4
1. Data Clustering helps improve query performance by:
    - a. Storing logically related records together hence reducing physical I/O
    - b. Utilizing intra parallelism on a multi-cpu system
    - c. Distributing data over cluster of nodes
    - d. All of the above.
  2. All of the above Typical tasks of a data base administrator include:
    - a. Creating objects (tables, indexes, views, etc.)
    - b. Grant privileges (select, update, delete, etc.)
    - c. Maintenance (load, statistics, copy, recover, etc.)
    - d. All of the above
  3. In RDBMS, the "automatic enforcement" of referential constraints is called:
    - a. Data Integrity
    - b. Referential Integrity
    - c. 2<sup>nd</sup> form of Normalization
    - d. 3<sup>rd</sup> form of Normalization
  4. If you want to list the employees whose salary is the highest, the outer query's WHERE Clause may be coded:
    - a. WHERE SALARY > (SELECT SALARY FROM EMPLOYEE..)
    - b. WHERE SALARY = (SELECT MAX (SALARY) FROM EMPLOYEE..)
    - c. WHERE SALARY > ALL (SELECT SALARY FROM EMPLOYEE..)
    - d. None of the above
  5. Scalar functions supported by RDMS are used to
    - a. Manipulate strings within SQL statements
    - b. Perform basic mathematical functions within SQL statements
    - c. Perform filtering to limit number of rows returned
    - d. All of the above
- A DBMS maintains Referential Integrity by enforcing the following rules:
- a. A foreign key value must match a unique key value or be null
  - b. Primary key values cannot be null
  - c. All non-null values must match a value in the referenced column (column set)
  - d. All of the above
- Which statement below is incorrect?
- a. In 1970 IBM's E.F. Codd first proposed the relational model for data
  - b. In 1996, Ralph Kimball published his famous book "The Data Warehouse Toolkit" challenged the Enterprise Data Model proposed by Bill Inmon
  - c. Multi-dimensional data model was introduced by Microsoft in 1997
  - d. None of the above



**True / False Questions (each carries 2 marks):**

1. Distributed Computing refers to multiple CPUs performing jobs within the same shared memory system. - True / False
2. Thread level parallelism can be achieved by running multiple tasks utilizing separate threads to run on multiple CPUs at the same time. - True / False
3. Basic communication unit in MPI is a message - a piece of data sent from one machine to another. - True / False
4. Access Control List (ACL) uses a set of three-tuples <User, Resource, Privilege> to specify which users are allowed to access which resources with which privileges. - True / False
5. In MPI paradigm, a process finds out its own rank by calling MPI\_Comm\_rank(). - True / False
6. In functional programming world, a higher-order function is a function that takes other functions as arguments or returns a function as result. Map() and fold() are examples of higher order functions. - True / False
7. Erlang and Haskell are two examples of functional programming. Software written using functional programming techniques are typically smaller in size hence tends to be less error prone. - True / False
8. Beowulf cluster is a network of very expensive machines that runs proprietary operating systems. - True / False
9. Distributed computing refers to multiple computers with their own memory communicating and accomplishing intensive work. - True / False
10. Cookies are popular method of HTTP authentication in which Temporary cookies are cached until browser shut down, persistent ones cached until expiry date. - True / False
11. Disabling unused services and accounts on a server is a cheap way of mitigating security attack risk. - True / False
12. "Hidden" parameters in HTML are not really hidden, can be manipulated. - True / False
13. In relation database table, all column names must be unique. - True / False
14. A DBMS accesses data from its physical storage utilizing Disk manager and file manager subsystems. File manager is responsible to request and provide actual data record whereas disk manager is responsible for write and read data from actual disk. - True / False
15. Static analysis tool when run on the code can uncover Buffer overflow vulnerabilities. - True / False
16. "CREATE PRIMARY KEY" and "CREATE FOREIGN KEY" statements can be used to define a primary key and a foreign key respectively. - True / False
17. WHERE clause predicates must be separated from each other by commas. - True / False
18. MD-5 and SHA-1 are examples of hashing algorithm used for checking data and message integrity. - True / False
19. In Map Reduce paradigm, the "map" function reads a set of key/value pairs, processes them any way it deems fit, and produces intermediate (key, data) data structures. - True / False
20. The following SQL query is syntactically correct - True / False  
`SELECT WORKDEPT, AVG(SALARY) FROM EMPLOYEE WHERE AVG(SALARY) > 20000 GROUP BY WORKDEPT`



8. DNS Cache Poisoning is the type of a security attack in which:
- a. An alert user can spot a suspicious URL
  - b. an attacker attempts to insert a fake address record for an Internet domain into the DNS server
  - c. For as long as the fake entry is cached by the server subscriber's browsers or e-mail servers will automatically go to the address provided by the compromised DNS server
  - d. Both (b) and (c) above are correct
9. Which of these native functions is known to cause Buffer overflow?
- a. malloc()
  - b. strcpy()
  - c. strcat()
  - d. both (b) and (c)
10. Which of the following is true about MPI(Message Passing Interface)?
- a. MPI programs are thread safe
  - b. MPI supports network heterogeneity
  - c. MPI supports dynamic process spawn
  - d. All of the above
11. If multiple threads are using the same shared memory then thread synchronization can be achieved using:
- a. Event
  - b. Mutex
  - c. Semaphore
  - d. Any of the above
12. What is not true about Functional Programming?
- a. Focuses on "what" should be done not necessarily "how"
  - b. Functions are treated first class
  - c. Supports lazy evaluation
  - d. Order of function invocation does matter
13. The following is not the method of authentication:
- a. One time password (OTP) cards such as SecureID
  - b. Biometrics such as Finger print scan
  - c. Access Control List (ACL)
  - d. None of the above
14. Risk of SQL Injection can be mitigated by
- a. Validating all user input
  - b. Programmatically review possible injection and take appropriate actions
  - c. Limit privileges on Database objects
  - d. All of the above
15. If you want to list the employees whose salary is the highest, the outer query's WHERE Clause may be coded:
- a. WHERE SALARY > (SELECT SALARY FROM EMPLOYEE..)
  - b. WHERE SALARY = (SELECT MAX (SALARY) FROM EMPLOYEE..)
  - c. WHERE SALARY > ALL (SELECT SALARY FROM EMPLOYEE..)
  - d. None of the above is correct

15) Between ~~any~~ rows of the result table

4 a) Sorted in Ascending

b) Sorted in Descending

c) Not sorted at all