processing until all portions of all log records for the transaction have been written to the physical log.

SUMMARY OF THE INVENTION

It is an object of the invention to ameliorate one or more disadvantages of the prior art.

According to one aspect of the invention there is disclosed in a data processing system a method executed by a data processor for fetching a selected key record in a group of record keys by utilizing a portion of a key record through an index tree having a modifiable structure during a transaction in said data processing system wherein other transactions may concurrently modify the structure of said index tree, said index tree having at least a root node, each root node having a key record reference to one or more nodes in a next lower ordered level and having bottom nodes that provide access to said key record data in an ordered sequence of key records, said method comprising the steps performed within said data processing system of: traversing across said nodes within said data processing system from said root node by using key record reference until an appropriate bottom node is reached; identifying said selected key record in said bottom node; requesting a conditional access restriction on said selected key record; fetching said selected key record if said conditional access restriction is granted; requesting an unconditional access restriction of said selected key record if said conditional access restriction is not granted; examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or note said appropriate bottom node has been substantially altered; fetching said selected key record is said appropriate bottom node has not been substantially altered; and traversing across said nodes from said root node by using said key record reference until a second appropriate bottom node is reached if said appropriate bottom node has been substantially altered.

According to another aspect of the invention there is disclosed in a data processing system a method executed by a data processor for inserting a single key record in a group of record keys according to a key record through an index tree having a

modifiable structure during a transaction in said data processing system wherein other transactions may concurrently modify the structure of said index tree, said index tree having at least a root node, each root node having a key record reference to one or more nodes in a next lower ordered level and having bottom nodes that provide access to an ordered sequence of key records, said method comprising the steps performed within said data processing system of: traversing across said nodes within said data processing system from said root node by using said key record reference until an appropriate bottom node is reached; identifying a next higher key record than the single key record to be inserted; requesting a conditional access restriction on said next higher key record; inserting said single key record into said appropriate bottom node if said conditional access restriction is granted; requesting an unconditional access restriction on said next higher key record if said conditional access restriction is not granted; examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or not said appropriate bottom node has been substantially altered; inserting said single key record into said appropriate bottom node if said appropriate bottom node has not been substantially altered; and traversing across said node from said root node by using said key record reference until a second appropriate bottom node is reached if said appropriate bottom node has been substantially altered.

According to another aspect of the invention there is disclosed in a data processing system a method executed by a data processor for deleting a single key record in a group of record keys according to a key record through an index tree having a modifiable structure during a transaction in said data processing system wherein other transactions may concurrently modify the structure of said index tree, said index tree having at least a root node, each root node having a key record reference to one or more nodes in a next lower ordered level and having bottom nodes that provide access to an ordered sequence of record keys, said method comprising the steps performed within said data processing system of: traversing said nodes within said data processing system from said root node by using said key record reference until an appropriate bottom node is reached;

requesting a conditional access restriction on a next higher key record than the single key record to be deleted; deleting said single key record from said appropriate bottom node if said conditional access restriction is granted; requesting an unconditional access restriction on said next higher key record if said conditional access restriction is not granted; examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or not said appropriate bottom node has been substantially altered; deleting said single key record from said appropriate bottom node if said appropriate bottom node has not been substantially altered; and traversing across said nodes from said root node by using said key record reference until a second appropriate bottom node is reached if said appropriate bottom node has been substantially altered.

According to another aspect of the invention there is disclosed a data processing system for fetching a selected key record in a group of record keys by utilizing a portion of a key record through an index tree having a modifiable structure during a transaction in said data processing system wherein other transactions may concurrently modify the structure of said index tree, said index tree having at least a root node, each root node having a key record reference to one or more nodes in a next lower ordered level and having bottom nodes that provide access to said key record data in an ordered sequence of key records, said data processing system comprising: means for traversing across said nodes within said data processing system from said root node by using said key record reference until an appropriate bottom node is reached; means for identifying said selected key record in said bottom node; means for requesting a conditional access restriction on said selected key record; means for fetching said selected key record if said conditional access restriction is granted; means for requesting an unconditional access restriction of said selected key record if said conditional access restriction is not granted; means for examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or not said appropriate bottom node has been substantially altered; means for fetching said selected key record if said appropriate bottom node has

not been substantially altered; and means for traversing across said nodes from said root node by using said key record reference until a second appropriate bottom node is reached if said appropriate bottom node has been substantially altered.

According to another aspect of the invention there is disclosed a data processing system for inserting a single key record in a group of record keys according to a key record through an index tree having a modifiable structure during a transaction in said data processing system wherein other transactions may concurrently modify the structure of said index tree, said index tree having at least a root node, each root node having a key record reference to one or more nodes in a next lowered ordered level and having bottom nodes that provide access to an ordered sequence of key records, said data processing system comprising: means for traversing across said nodes within said data processing system from said root node by using said key record reference until an appropriate bottom node is reached; means for identifying a next higher key record than the single key record to be inserted; means for requesting a conditional access restriction on said next higher key record; means for inserting said single key record into said appropriate bottom node if said conditional access restriction is granted; means for requesting an unconditional access restriction on said next higher key record if said conditional access restriction is not granted; means for examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or not said appropriate bottom node has been substantially altered; means for inserting said single key record into said appropriate bottom node if said appropriate bottom node has not been substantially altered; and means for traversing across said node from said root node by using said key record reference until a second appropriate bottom node is reached if said appropriate bottom node has been substantially altered.

According to another aspect of the invention there is disclosed a data processing system for deleting a single key record in a group of record keys according to a key record through an index tree having a modifiable structure during a transaction in said data processing system wherein other transactions may concurrently modify the structure

of said index tree, said index tree having at least a root node, each root node having a key record reference to one or more nodes in a next lower ordered level and having bottom nodes that provide access to an ordered sequence of record keys, said data processing system comprising: means for traversing said nodes within said data processing system from said root node by using said key record reference until an appropriate bottom node is reached; means for requesting a conditional access restriction on a next higher key record than the single key record to be deleted; means for deleting said single key record from said appropriate bottom node if said conditional access restriction is granted; means for requesting an unconditional access restriction on said next higher key record if said conditional access restriction is not granted; means for examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or not said appropriate bottom node has been substantially altered; means for deleting said single key record from said appropriate bottom node if said appropriate bottom node has not been substantially altered; and means for traversing across said nodes from said root node by using said key record reference until a second appropriate bottom node is reached if said appropriate bottom node has been substantially altered.

According to still anther aspect of the invention there is disclosed a data processing system comprising a concurrent access database system for providing access to data in the database by a plurality of users concurrently by means of a tree structure, said access comprising first operations in which the tree structure remains unchanged and second operations in which the tree structure is modified, said data processing system comprising means for performing a transaction comprising a plurality of said operations, and recovery means for storing the sequential positions of operations from which a transaction can be rolled back to allow recovery of operations in the event of a system failure, characterised in that the recovery means comprises means for storing an indication of the sequential position of a second operation in a sequence of operations modifying the tree structure immediately prior to commencement of said sequence;

means for storing a compensation log record pointing said indication on completion of said sequence, and means for rolling back each of said plurality of said operations in the transaction, by reference to said stored indication, when the system fails prior to completion of said sequence, and for rolling back, on system failure, only operations not in said sequence when said sequence has completed, by reference to said log compensation record.

According to still another aspect of the invention there is disclosed a method of operating a data processing system comprising a concurrent access database system for providing access to data in the database by as plurality of user concurrently by means of a tree structure, said access including first operations in which eh tree structure remains unchanged and second operations in which the tree structure is modified, the method comprising the steps of: performing a transaction comprising a plurality of said operations; storing an indication of the sequential position of a second operation in a sequence of operations modifying the tree structure immediately prior to commencement of said sequence; storing a compensation log record pointing said indication on completion of said sequence; rolling back each of aid plurality of said operations in the transaction, by reference to said stored indication, on system failure prior to completion of said sequence; and rolling back, on system failure only operations not in sequence when said sequence has completed, by reference to said log compensation record.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The preferred embodiment of the invention will now be described, by way of example, with reference to the accompanying drawings, in which:

FIGURE 1 is a block diagram of a concurrent access database system in accordance with the present invention;

FIGURE 2 is a logic flow chart illustrating an initial search operation through a database;

FIGURE 3 is a logic flow diagram illustrating a fetch operation through a database;

FIGURE 4 is a logic flow chart illustrating an insert operation through a database;

FIGURE 5 is a logic flow chart illustrating a delete operation through a database;

CLAIMS

1. In a data processing system a method executed by a data processor for fetching a selected key record in a group of record keys by utilizing a portion of a key record through an index tree having a modifiable structure during a transaction in said data processing system wherein other transactions may concurrently modify the structure of said index tree, said index tree having at least a root node, each root node having a key record reference to one or more nodes in a next lower ordered level and having bottom nodes that provide access to said key record data in an ordered sequence of key records, said method comprising the steps performed within said data processing system of:

traversing across said nodes within said data processing system from said root node by using key record reference until an appropriate bottom node is reached;

identifying said selected key record in said bottom node;

requesting a conditional access restriction on said selected key record;

fetching said selected key record if said conditional access restriction is granted;

requesting an unconditional access restriction of said selected key record if said conditional access restriction is not granted;

examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or note said appropriate bottom node has been substantially altered;

fetching said selected key record is said appropriate bottom node has not been substantially altered; and

traversing across said nodes from said root node by using said key record reference until a second appropriate bottom node is reached if said appropriate bottom node has been substantially altered.

2. The method according to claim 1, wherein said step of examining said appropriate bottom node after said unconditional access restriction has been granted to determine whether or not said appropriate bottom node has been substantially altered further

includes the step of determining whether or not said selected key record is a first key record in said ordered sequence of key records within said appropriate bottom node.

3. The method of according to claim 1, wherein said step of examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or not said appropriate bottom node has been substantially altered further includes the step of determining whether or not a next lower key record in said ordered sequence of key records within said appropriate bottom node is lower than said selected key record.

4. The method according to claim 1, wherein said step of examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or not said appropriate bottom node has been substantially altered further includes the step of determining whether or not a key record exists in said appropriate bottom node which is lower than said selected key record.

5. In a data processing system a method executed by a data processor for inserting a single key record in a group of record keys according to a key record through an index tree having a modifiable structure during a transaction in said data processing system wherein other transactions may concurrently modify the structure of said index tree, said index tree having at least a root node, each root node having a key record reference to one or more nodes in a next lower ordered level and having bottom nodes that provide access to an ordered sequence of key records, said method comprising the steps performed within said data processing system of:

traversing across said nodes within said data processing system from said root node by using said key record reference until an appropriate bottom node is reached;

identifying a next higher key record than the single key record to be inserted;

requesting a conditional access restriction on said next higher key record;

inserting said single key record into said appropriate bottom node if said conditional access restriction is granted;

requesting an unconditional access restriction on said next higher key record if said conditional access restriction is not granted;

examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or not said appropriate bottom node has been substantially altered;

inserting said single key record into said appropriate bottom node if said appropriate bottom node has not been substantially altered; and

traversing across said node from said root node by using said key record reference until a second appropriate bottom node is reached if said appropriate bottom node has been substantially altered.

6. The method according to claim 5, wherein said step of examining said appropriate bottom node after said unconditional access restriction is granted further includes the step of determining whether said single key record is bounded on said appropriate bottom node.

7. The method according to claim 5, wherein said step of examining said appropriate bottom node after said unconditional access restriction is granted further includes the step of determining whether said appropriate bottom node is participating in a structure modification operation which is not yet complete.

8. In a data processing system a method executed by a data processor for deleting a single key record in a group of record keys according to a key record through an index tree having a modifiable structure during a transaction in said data processing system wherein other transactions may concurrently modify the structure of said index tree, said index tree having at least a root node, each root node having a key record reference to one or more nodes in a next lower ordered level and having bottom nodes that provide access to an ordered sequence of record keys, said method comprising the steps performed within said data processing system of:

traversing said nodes within said data processing system from said root node by using said key record reference until an appropriate bottom node is reached;

requesting a conditional access restriction on a next higher key record than the single key record to be deleted;

deleting said single key record from said appropriate bottom node if said conditional access restriction is granted;

requesting an unconditional access restriction on said next higher key record if said conditional access restriction is not granted;

examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or not said appropriate bottom node has been substantially altered;

deleting said single key record from said appropriate bottom node if said appropriate bottom node has not been substantially altered; and

traversing across said nodes from said root node by using said key record reference until a second appropriate bottom node is reached if said appropriate bottom node has been substantially altered.

9. The method according to claim 8, wherein said step of examining said appropriate bottom node after said unconditional access restriction is granted further includes the step of determining whether said appropriate bottom node is participating in a structure modification operation which is not yet complete.

10. A data processing system for fetching a selected key record in a group of record keys by utilizing a portion of a key record through an index tree having a modifiable structure during a transaction in said data processing system wherein other transactions may concurrently modify the structure of said index tree, said index tree having at least a root node, each root node having a key record reference to one or more nodes in a next lower ordered level and having bottom nodes that provide access to said key record data in an ordered sequence of key records, said data processing system comprising:

means for traversing across said nodes within said data processing system from said root node by using said key record reference until an appropriate bottom node is reached;

means for identifying said selected key record in said bottom node;

means for requesting a conditional access restriction on said selected key record;

means for fetching said selected key record if said conditional access restriction is granted;

means for requesting an unconditional access restriction of said selected key record if said conditional access restriction is not granted;

means for examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or not said appropriate bottom node has been substantially altered;

means for fetching said selected key record if said appropriate bottom node has not been substantially altered; and

means for traversing across said nodes from said root node by using said key record reference until a second appropriate bottom node is reached if said appropriate bottom node has been substantially altered.

11. The data processing system according to claim 10, wherein said means for examining said appropriate bottom node after said unconditional access restriction has been granted to determine whether or not said appropriate bottom node has been substantially altered further includes means for determining whether or not said selected key record is a first key record in said ordered sequence of key records within said appropriate bottom node.

12. The data processing system according to claim 10, wherein said means for examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or not said appropriate bottom node has been substantially altered further includes means for determining whether or not a next lower key record in said ordered sequence of key records within said appropriate bottom node is lower than said selected key record.

13. The data processing system according to claim 10, wherein said means for examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or not said appropriate bottom node has been substantially

altered further includes means for determining whether or not a key record exists in said appropriate bottom node which is lower than said selected key record.

14. A data processing system for inserting a single key record in a group of record keys according to a key record through an index tree having a modifiable structure during a transaction in said data processing system wherein other transactions may concurrently modify the structure of said index tree, said index tree having at least a root node, each root node having a key record reference to one or more nodes in a next lowered ordered level and having bottom nodes that provide access to an ordered sequence of key records, said data processing system comprising:

means for traversing across said nodes within said data processing system from said root node by using said key record reference until an appropriate bottom node is reached;

means for identifying a next higher key record than the single key record to be inserted;

means for requesting a conditional access restriction on said next higher key record;

means for inserting said single key record into said appropriate bottom node if said conditional access restriction is granted;

means for requesting an unconditional access restriction on said next higher key record if said conditional access restriction is not granted;

means for examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or not said appropriate bottom node has been substantially altered;

means for inserting said single key record into said appropriate bottom node if said appropriate bottom node has not been substantially altered; and

means for traversing across said node from said root node by using said key record reference until a second appropriate bottom node is reached if said appropriate bottom node has been substantially altered.

15. The data processing system according to claim 14, wherein said means for examining said appropriate bottom node after said unconditional access restriction is granted further includes means for determining whether said single key record is bounded on said appropriate bottom node.

16. The data processing system according to claim 14, wherein said means for examining said appropriate bottom node after said unconditional access restriction is granted further includes means for determining whether said appropriate bottom node is participating in a structure modification operation which is not yet complete.

17. A data processing system for deleting a single key record in a group of record keys according to a key record through an index tree having a modifiable structure during a transaction in said data processing system wherein other transactions may concurrently modify the structure of said index tree, said index tree having at least a root node, each root node having a key record reference to one or more nodes in a next lower ordered level and having bottom nodes that provide access to an ordered sequence of record keys, said data processing system comprising:

means for traversing said nodes within said data processing system from said root node by using said key record reference until an appropriate bottom node is reached;

means for requesting a conditional access restriction on a next higher key record than the single key record to be deleted;

means for deleting said single key record from said appropriate bottom node if said conditional access restriction is granted;

means for requesting an unconditional access restriction on said next higher key record if said conditional access restriction is not granted;

means for examining said appropriate bottom node after said unconditional access restriction is granted to determine whether or not said appropriate bottom node has been substantially altered;

means for deleting said single key record from said appropriate bottom node if said appropriate bottom node has not been substantially altered; and

means for traversing across said nodes from said root node by using said key record reference until a second appropriate bottom node is reached if said appropriate bottom node has been substantially altered.

18. A data processing system according to claim 17, wherein said means for examining said appropriate bottom node after said unconditional access restriction is granted further includes means for determining whether said appropriate bottom node is participating in a structure modification operation which is not yet complete.

19. A data processing system comprising a concurrent access database system for providing access to data in the database by a plurality of users concurrently by means of a tree structure, said access comprising first operations in which the tree structure remains unchanged and second operations in which the tree structure is modified, said data processing system comprising means for performing a transaction comprising a plurality of said operations, and recovery means for storing the sequential positions of operations from which a transaction can be rolled back to allow recovery of operations in the event of a system failure,

characterised in that the recovery means comprises means for storing an indication of the sequential position of a second operation in a sequence of operations modifying the tree structure immediately prior to commencement of said sequence, means for storing a compensation log record pointing said indication on completion of said sequence, and means for rolling back each of said plurality of said operations in the transaction, by reference to said stored indication, when the system fails prior to completion of said sequence, and for rolling back, on system failure, only operations not in said sequence when said sequence has completed, by reference to said log compensation record.

20. A data processing system according to claim 19, in which said tree structure is an index tree structure having a root node including key record references to nodes in a next lower ordered level node, and respectively ordered lower level nodes, each referenced by keys in the next higher ordered node, down to bottom nodes having keys referencing the

data in the data base, each said operation comprising traversing the nodes, using said key record references to an existing bottom node, and including means for requesting conditional access restriction on reaching a key record in the bottom node, means for accessing referenced data in the data base on grant of the access restriction, means for requesting an unconditional access restriction if said conditional access restriction is not granted, and means for traversing said nodes from the root node to a new bottom node if said existing bottom node is substantially altered.

21. A data processing system according to claim 20, including means for examining said bottom node on grant of said unconditional access restriction to determine whether a key record identified in the bottom node is the first key record in the node.

22. A data processing system according to claim 21, including means for determining, in said bottom node, whether a key record therein is lower than a requested key record.

23. A method of operating a data processing system comprising a concurrent access database system for providing access to data in the database by as plurality of user concurrently by means of a tree structure, said access including first operations in which eh tree structure remains unchanged and second operations in which the tree structure is modified, the method comprising the steps of:

performing a transaction comprising a plurality of said operations;

storing an indication of the sequential position of a second operation in a sequence of operations modifying the tree structure immediately prior to commencement of said sequence;

storing a compensation log record pointing said indication on completion of said sequence;

rolling back each of aid plurality of said operations in the transaction, by reference to said stored indication, on system failure prior to completion of said sequence; and

rolling back, on system failure only operations not in sequence when said sequence has completed, by reference to said log compensation record.

24. A method according to claim 23, in which said tree structure is an index tree structure having a root node including key record references to nodes in a next lower ordered level node, and respectively ordered lower level nodes, each referenced by keys in the next higher ordered node, down to bottom nodes having keys referencing the data in the data base, each said operation comprising traversing the nodes, using said key record references, to an existing bottom node, further including the steps of:

requesting conditional access restriction on reaching a key record in the bottom node;

accessing referenced data in the data base on grant of the access restriction;

requesting an unconditional access restriction if said conditional access restriction is not granted; and

traversing said nodes from the root node to a new bottom node if said existing bottom node is substantially altered.

25. A method according to claim 24, including the step of: examining said bottom node on grant of said unconditional access restriction to determine whether a key record identified in the bottom node is the first key record in the node.

26. A method according to claim 25, including the step of: determining, in said bottom node, whether a key record therein is lower than a requested key record.

TREE STRUCTURE DATABASE SYSTEM

ABSTRACT

A database system in a data processing system has an index tree structure with a root node (102) having reference keys to nodes in the next lower level and so on until bottom nodes (116) having keys referencing the database records. Multiple users may access the database concurrently. Transactions in which the tree structure is to be modified (for example removal of a record) are performed by traversing the tree to the appropriate bottom node and setting an indication of pendency of the modification. At the same time other accesses to the database, through nodes at which no tree structure modifications are to take place, can continue. When the structure modification has been performed, a log record of it is made to prevent undoing the modification in the event of a system failure, and database access through all the bottom nodes then continues.

Fig. 2