



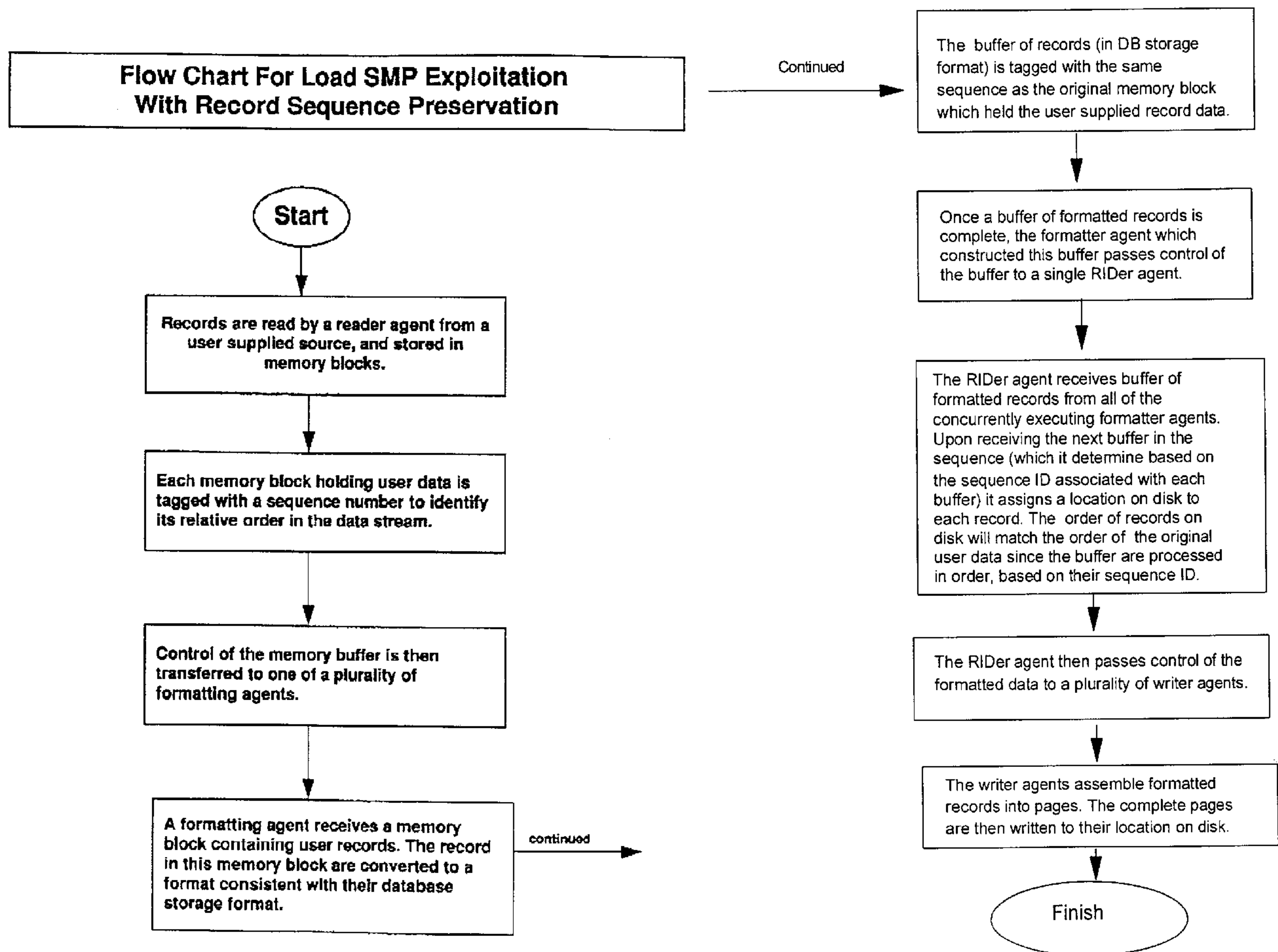
(11) (21) (C) **2,209,549**
(22) 1997/07/02
(43) 1999/01/02
(45) 2000/05/02

(72) LIGHTSTONE, Sam, CA
(72) COMEAU, Alain C., CA
(73) IBM CANADA LTD. - IBM CANADA LIMITÉE, CA

(51) Int.Cl.⁶ G06F 17/30, G06F 15/16, G11B 20/12

(54) **METHODE ET DISPOSITIF POUR CHARGER DES DONNEES
DANS UNE BASE DE DONNEES DANS UN ENVIRONNEMENT
MULTIPROCESSEUR**

(54) **METHOD AND APPARATUS FOR LOADING DATA INTO A
DATABASE IN A MULTIPROCESSOR ENVIRONMENT**





(11) (21) (C) **2,209,549**
(22) 1997/07/02
(43) 1999/01/02
(45) 2000/05/02

(57) Méthode et appareil utilisant le traitement parallèle pour charger, dans une base de données, des données d'une source groupées en enregistrements selon un ordre prédéterminé. Au moyen d'un ou de plusieurs agents de lecture, les enregistrements de données sont lus dans la source où ils sont stockés. Ces enregistrements de données sont stockés dans des groupes d'enregistrements. Chaque groupe d'enregistrements stockés est étiqueté au moyen d'un identificateur de séquence correspondant à l'ordre prédéterminé des données. Un dispositif permet de formater les enregistrements de données des groupes en utilisant un certain nombre d'agents de formatage en parallèle. Les enregistrements formatés sont stockés dans des groupes d'enregistrements formatés. Des identificateurs de séquence sont indiqués pour chacun des groupes dans le groupe d'enregistrements formatés correspondant. Un dispositif d'identification d'enregistrement associe des emplacements de page aux enregistrements au moyen des identificateurs de séquence afin de maintenir l'ordre prédéterminé.

(57) This invention provides a method and apparatus for loading data having a predetermined order of data records from a source into a database using parallel processing. Using one or more reading agents, data records are read from the source in which they are stored. These data records are stored in groups of records. Each group of stored records are tagged with a sequence identifier corresponding to the predetermined order of data. Apparatus is provided for formatting the data records in the groups using a plurality of formatting agents in parallel. The formatted records are stored in formatted record groups. Sequence identifiers are written for each of the groups to the respective formatted record group. A record identification apparatus is provided for assigning page locations for the records using the sequence identifiers to maintain the predetermined order.

CA9-97-022

ABSTRACT**METHOD AND APPARATUS FOR LOADING DATA INTO A DATABASE
IN A MULTI PROCESSOR ENVIRONMENT**

5

This invention provides a method and apparatus for loading data having a predetermined order of data records from a source into a database using parallel processing. Using one or more reading agents, data records are read from the source in which they are stored. These data records are stored in groups of records. Each group of stored records are tagged with a sequence identifier corresponding to the predetermined order of data. Apparatus is provided for formatting the data records in the groups using a plurality of formatting agents in parallel. The formatted records are stored in formatted record groups. Sequence identifiers are written for each of the groups to the respective formatted record group. A record identification apparatus is provided for assigning page locations for the records using the sequence identifiers to maintain the predetermined order.

10

15

METHOD AND APPARATUS FOR LOADING DATA INTO A DATABASE IN A MULTI PROCESSOR ENVIRONMENT

FIELD OF THE INVENTION

This invention relates to the loading of data into a database operating in a data processing system comprising a network of multiple Central Processing Units (CPUs), typically known as
5 Symmetric Multi processors (SMPs).

BACKGROUND OF THE INVENTION

With the advent of databases it is incumbent on the user or operator of a database to INSERT
10 user supplied data into the database in a form that is consistent with the internal storage form
of data for the database. A Database Management System (DBMS) or LOAD utility is used
to convert the user supplied data into the internal database form. The actions of parsing,
conversion, and formatting data into this internal form is a CPU intensive operation. LOAD
utilities reading from user supplied source media, such as disks, or tapes containing user
15 supplied data incur the overhead of these conversions, as do INSERTs issued to the DBMS.
Because of these CPU intensive operations, insertion of user data may become a CPU bound
activity.

It is well known in the art that database records are grouped into pages of predetermined size.
A number of pages are grouped together into a table, and each page in the table contains a
20 number that denotes its sequence in the table. Pages of a table are typically written to
persistent storage (such as disk).

Symmetric Multi processors (SMPs) comprise a class of computers containing multiple CPUs.
Operating systems running on SMPs dispatch processes and threads to different CPUs in order
25 to distribute the workload assigned across the available processors. For a given program to
exploit the power of an SMP, it is advantageous that it is designed so that it performs portions

CA9-97-022

2

of its workload in separate dispatchable units of work which the operating system of the SMP can distribute to the various CPUs in the SMP.

5 It would be a definite advantage in many cases that when data is loaded into the database of an SMP system, the ability to perform parsing, conversion and formatting processes in concurrent dispatchable operating system units in order to exploit the advantages of the SMP system hardware and operating system.

10 The difficulty is that the design of a method and apparatus in order to load data for processing in concurrent dispatchable operating system units is nontrivial because of the numerous items of state information which must be maintained as part of the database table meta-data (stored data that describes the database table concerned), such as free space control records, a table descriptor record, etc. as will be appreciated by those skilled in the database art. Despite the complexity of the problem, solutions to it have been attempted in the past.

15 However, in the normal case, adding true parallel processing (i.e.. True decoupled concurrent processing) the data is processed by each CPU in the system independently, resulting in the data being loaded into the database table in an arbitrary sequence. This means that the data is stored in an arbitrary sequence in the table, and as well, in an arbitrary physical sequence
20 on the database storage device used by the data processing system.

25 As will be appreciated by those skilled in the art reviewing this application, the arbitrary sequence of data both logically in the database table and physically on the storage device being used by the data processing system to store database information can pose a problem if the source data was intentionally supplied by the user in a significant order (such as cluster order). It can also result in poor exploitation of the buffer pool and I/O prefetchers used by the data processing system to accelerate or optimize data retrieval. The net result of which is ultimately that corruption of the sequence of the records in the source data may negatively impact subsequent query performance. As will be appreciated, query performance is one of

CA9-97-022

3

the primary judgement criteria for which database products compete.

5 The very requirement for data order and the advantages of parallel processing appear to be opposed to each other. In order to take advantage of parallel processing it appears that the prior art would require sacrifice of the required data order, while corrupting the data order appears to negatively impact processing performance including query performance.

The performance results are major indicators of product performance, and are heavily used by customers in deciding which Database products to buy.

10

The Transaction Processing Performance Council (TPC) regularly establishes guidelines for transaction processing and database benchmarks against which Database vendors regularly compete. Database vendors regularly publish their TPC compliant performance results. The official TPC benchmarks include both query performance, as well as the database creation time (of which LOADING data is a major component). So, both the creation of the database and the subsequent query performance are major factors which customers consider, and for which Database product producers aim.

15

The term transaction is often applied to a wide variety of business and computer functions. From the point of view of a computer function, a transaction could refer to a set of operations including disk read/writes, operating system calls, or a type of data transfer from one system or subsystem to another.

20

While TPC benchmarks involve the measurement and evaluation of computer functions and operations, the TPC regards a transaction as it is commonly understood in the business world: a commercial exchange of goods, services, or money. A typical transaction, as defined by the TPC, would include the updating to a database system for such things as inventory control (goods), airline reservations (services), or banking (money).

25

CA9-97-022

4

SUMMARY OF THE INVENTION

The current invention relates to methods and apparatus for data loading into a database table which exploits true SMP parallelism. The preferred embodiment disclosed below has been laboratory tested and shown to yield excellent parallel processing, scalability permitting good performance on SMP systems to at least 12 parallel CPUs, and which preserves the user provided order of the table data during the loading process.

In one aspect of the invention which is designed for use in a MULTI PROCESSOR data processing system capable of parallel processing, a method is provided for loading data having a predetermined order of data records into a database using parallel processing,

using one or more reading tasks, reading each data record from the location in which it is stored; storing each said record in a buffer memory block of predetermined size;

tagging each buffer with a sequence identifier; transferring control of each said buffer to the control of one of a plurality of formatting tasks;

converting said records of data into a suitable database storage format;

storing said formatted data records in buffers;

writing said sequence identifier to said buffers of formatted data;

assigning a database page location for each record using said formatted data buffer order and said sequence identifier to maintain said predetermined order;

assembling said records into pages;

writing said pages to storage at the locations specified.

In another aspect of the invention a record is read from data records having a pre selected order from a source location of the data. Each record is stored in buffer blocks of pre selected size, and each buffer block is tagged with a sequence number corresponding to said pre selected order of said record in its source location. Control of the buffers is transferred to a plurality of formatter tasks or agents for processing in parallel. The records of data are converted to a suitable database storage format, and stored in buffers. A sequence identification number of each buffer is written to the buffers of formatted record data.

CA9-97-022

5

Control of the buffers of formatted data is transferred to a RIDer agent. For the data records of each buffer a database page location is assigned maintaining the pre selected order by means of the formatted data record buffer order and sequence identification number. Control of the formatted data is transferred to a plurality of writer agents. The formatted records are assembled into pages which are written to system storage in the assigned storage locations.

In still another aspect of the invention apparatus is provided for loading data records having a pre selected order using a parallel data processing system to load the data into the data storage of a database while maintaining the pre selected data order. The apparatus includes at least one reader agent to read the data, preferably a number of formatter agents for converting the data records read by the reader agent to a format suitable for the database. Apparatus is provided for storing the formatted records in a buffer, and assigning with sequence number to the buffer of the data records. A Record Identification (RIDer) agent is provided for assigning data processing system storage location for each read data record. Apparatus is provided for assembling the formatted records into pages, and a number of writer agents are provided for writing each page to the appropriate assigned storage location.

Yet another aspect of the invention provides a method of loading data having a predetermined order of data records from a source into a database using parallel processing comprising:

- 20 using one or more reading agents, reading data records from the source in which they are stored;
- storing said records in groups of records;
- tagging each group with a sequence identifier corresponding to said predetermined order of data;
- 25 formatting the data records in said groups using a plurality of formatting agents in parallel;
- storing said formatted records in formatted record groups;
- writing said sequence identifier of each said group to the respective formatted record group;

CA9-97-022

6

assigning page locations for said records using said sequence identifiers to maintain said predetermined order.

Still another aspect of the invention provides a computer program product comprising:

- 5 a computer usable medium having computer readable program code means embodied therein for causing a computer to load user data, the computer program product comprising :
- computer readable program code means for causing a computer to effect apparatus for loading data records having a pre selected order using a parallel data processing system to load said data into the data storage of a database using said data processing system while
- 10 maintaining said pre selected data order, comprising;
- computer readable program code means for causing a computer to effect at least one reader agent to read said data;
- computer readable program code means for causing a computer to effect buffer storage means for storing said data in buffers;
- 15 computer readable program code means for causing a computer to effect tagging means for tagging each buffer containing said data with a sequence number corresponding to the pre selected order of the data;
- computer readable program code means for causing a computer to effect a plurality of formatter agents for converting said data records read by said at least one reader agent to
- 20 a format suitable for said database;
- computer readable program code means for causing a computer to effect means for storing said formatted records in formatted record buffers, with the sequence numbers of the buffers containing said data;
- computer readable program code means for causing a computer to effect a (RIDer)
- 25 agent for assigning page locations for each said formatted data records;
- computer readable program code means for causing a computer to effect means for assembling said formatted records into pages, and;
- computer readable program code means for causing a computer to effect a plurality of writer agents for writing each said pages to said assigned storage locations.

CA9-97-022

7

Another aspect of this invention provides a program storage device readable by a machine, tangibly embodying a program of instructions executable by the machine to perform method steps for loading data having a predetermined order of data records from a source into a database using parallel processing said method steps comprising:

- 5 using one or more reading agents, reading data records from the source in which they are stored;
- storing said records in groups of records;
- tagging each group with a sequence identifier corresponding to said predetermined order of data;
- 10 formatting the data records in said groups using a plurality of formatting agents in parallel;
- storing said formatted records in formatted record groups;
- writing said sequence identifier of each said group to the respective formatted record group;
- 15 assigning page locations for said records using said sequence identifiers to maintain said predetermined order.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a flow diagram of a preferred method of the invention.

- 20 Figure 2 is a block diagram of a process model for a preferred embodiment of the invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS OF THE INVENTION

- Referring to Figure 2 which illustrates a process model for a preferred specific embodiment of the invention it can be seen that a single Reader Agent 1 is provided to read user data records. This information is transferred to the control of a number of Formatter Agents 2
- 25 which convert the user data records to records in a suitable database format corresponding to that used by the database of the data processing system for which the invention can be used. The Record Identification (RIDer) Agent 3 assigns each record a location within a page

CA9-97-022

8

of the database table.

5 The Writer Agents 4 construct pages (i.e., assemble formatted records into pages) based on the location assigned to each record by the RIDer agent, in conjunction with the formatted record data. The Writer Agents write fully constructed pages to persistent storage (e.g., disk). It should be noted that page construction could be done by other agents.

The agents shown in Figure 2 combine to form a data LOADER, i.e., a collection of software agents that together load user data into a database table.

10

Figure 1, which represents a flow chart for the load SMP exploitation with record sequence preservation system of the present invention is described below:

15

A data processing system user having sequential data records to be entered into a database on an SMP system would initiate loading of the data by the data processing system.

20

In order that the data records can be quickly processed using the capabilities of the SMP system parallel processing the invention provides the following process which may be conveniently implemented using software agents or tasks running on the data processing system

25

At the start of record entry the user supplied records are read by a reader agent from the source supplied by the user, e.g., diskette or tape, and stored in memory block buffers in the memory area of said data processing system.

In order to assist in record order preservation during the currency of the operation each memory block buffer holding user data is tagged with a sequence number to identify its relative order in the data stream resulting from the data loading process.

CA9-97-022

9

Once this tagging has been accomplished in a buffer, control of that memory buffer is then transferred to one of a number of formatting agents. The Formatter agent reads the user data records from the buffer received from the Reader, converts each record into internal database storage format, and stores the converted records in a second memory block or buffer in the same order that these records appeared in the memory buffer containing the user data (received from the Reader agent). The buffer containing the records, now in the database storage format is tagged with the same sequence as the original memory buffer which held the user supplied record data. The operation of the formatter agents is carried on concurrently taking advantage of the parallel processing capabilities of the SMP system on which they are operating.

Once the formatted record data in a buffer is complete the formatter agent which constructed the formatted data records passes control of the buffer to a single RIDer agent.

The RIDer agent (The name RIDer was chosen as the RIDer agent assigns a Record Identifier (RID) to each record. The RID represents the location of a record on a page (including identification of the page)) receives control of buffers of formatted records from all of the concurrently executing formatter agents (i.e. parallelly executing formatter agents). Upon receiving the next buffer in the sequence (which it determines based on the sequence ID associated with each buffer of formatted records) it assigns a location on a page in the database table to each record.

The RIDer agent uses the sequence number on each formatted record buffer to ensure that it processes them in the originating sequence (that is, the sequence of the user data)

Once its work is done the RIDer agent then passes control of the formatted data to a number of writer agents implemented by the invention and operating on the data processing system for parallel or concurrent application. The writer agents then assemble the formatted records into pages, i.e.. the Writer agents construct pages based on the location assigned to each

CA9-97-022

10

record by the RIDer agent, in conjunction with the formatted record data. The fully constructed pages are then written to their location on the data processing system's database storage. This can be done by the Writer agents or some other agent.

5 As result of this, the order of records on the (e.g. disk) storage of the data processing system will match the order of the original user data since the lists are processed in order based on their sequence Ids.

PSEUDO CODE

10

The preferred embodiment above as illustrated in Figures 1 and 2 may be further understood by more particular information present in the following pseudo code which illustrates a software implementation of the invention which achieves load SMP exploitation with record sequence preservation when suitably formulated for use on a SMP data processing system.

15

PSEUDO CODE FOR READER AGENT:

{

Open user data source for reading;

Read nK bytes of user data into the bottom half of a memory buffer of 2nK bytes;

20 Strip any existing partial record from the bottom of this nK bytes of data just read;

Store a copy of the partial record in a temporary storage area;

If there was a partial record in the previous read operation, then prefix that partial record to the start of the nK memory area;

25

Tag the memory buffer with a sequence identifier;

Transfer control of the memory buffer to one of the Formatter agents;

CA9-97-022

11

```
    Continue reading the next nK bytes of user data until all user data is read;  
}
```

5 Comments: There can be more than one Reader agent used, as long as a relationship is provided e.g.. in the form of a software algorithm to correctly tag the user data records in the buffers with suitable sequence Ids. In the specific embodiment illustrated only a single Reader agent is illustrated so the assignment of sequence Ids is straightforward e.g.. the first block of user data has sequence ID 0, the second block has ID 1, etc.

10 While the pseudo code of the Reader agent has been generalized for a buffer storage area of 2nK bytes, the implementation tested as described below is a specific implementation for an arbitrary record length limit. Assuming that no user record could exceed 32K, a 64K buffer was allocated so that a whole number of records could be read into a block of memory. In
15 the 64K buffer 32K of user data was read into the bottom half of the buffer. (While the 32K size is up to the system designer the important consideration is that the bottom half of a buffer large enough to hold at least one record is used. The pseudo code reflects that the buffer is twice the size of the expected data block.) The data was examined by the Reader, to determine if a partial record existed at the end of the block. If a partial record existed (and it frequently did) the partial record was copied to a temporary storage space, and the end of
20 the memory block was redefined to be the byte prior to the start of the partial record. When the next 32K bytes of user data was read (again into the bottom half of the 64K buffer) the partial record from the end of the previous block was prefixed to this buffer. The pseudo code illustrated enables the particular embodiment described to handle complete records rather than partial records. This is a useful feature for the situation in which records have
25 varying lengths.

Sequence Ids can be in any form sufficient to identify the order of the data, such as : alphabetic identifiers, integer identifiers, the modulo of the formatter count, etc.

CA9-97-022

12

In another implementation of the invention user data can be alternatively read by any of the other agents in the process model.

PSEUDO CODE FOR FORMATTER AGENT:

```

5  {
    Receive control of a memory buffer containing user record data in user format;

    Convert the next user record into a formatted record consistent with the form of
records in database storage;
10
    The converted record, now in database form, is added to a buffer of formatted
records;
    If this is the last record in the buffer:
    {
15      Tag the buffer containing formatted records with a sequence ID identical to the
          sequence ID of the user data area;
      Transfer control of the buffer containing database-formatted record to the
RIDer agent;
      Obtain control of a new buffer of user data from the Reader agent;
20    }
    else, there are more user records to process;
    {
      continue formatting the next user record;
    }
25 }

```

Comments:

There can be any number of formatter agents, but the performance of this invention can be optimized by correlating (but not necessarily equating) the number of formatting agents with

CA9-97-022

13

the number of system CPUs online.

PSEUDO CODE FOR WRITER AGENT:

{

5 Receive control of a buffer of formatted records, and their associated location information from the RIDer agent;

 Construct database data pages;

10 Write pages to storage devices;

 Obtain next buffer from RIDer agent, until complete;

}

Comments:

15 There can be any number of Writer agents. Typically, the number of Writer agents should be related to the number of database storage devices used.

Database data pages could be written by agents other than the Writer agent, and could be constructed by agents other than the Write agent.

20

PSEUDO CODE FOR RIDer AGENT:

{

 Obtain a buffer of formatted database records from one of the Formatter agents;

25 if this buffer is the next expected buffer in the sequence of buffers, based on the sequence ID:

 {

 For each record in the buffer of the formatted records , assign a location within a page of the database table.

CA9-97-022

14

After processing the entire buffer, process any cached buffers of formatted records that are in sequence;

```

    After all cached buffers of formatted records in perfect sequence have been
5 processed, return to the first step "Obtain a buffer of formatted data records...";
    }
    else this buffer will need to be processed later:
    {
        Cache this buffer, until it can be processed in sequence;
10    }
    }

```

Comments:

In the implementation illustrated "preserving record sequence" is an option on the loader.
15 By specifying that the LOADER not exploit this capability, the RIDer is free to process data in any sequence, which can result in greater scalability on SMP systems.

Once written to disk the records can be conveniently and quickly accessed achieving a high degree of read efficiency.

20

Normally, to preserve the data sequence, the RIDer has to wait until it gets the next buffer of formatted data in the sequence. For example, while waiting for formatted data buffer #6, it may receive formatted data buffers #8 and #9, then #7. It waits until #6 arrives. Once #6 arrives from the formatter that generated it the RIDer can then process buffers #6,#7,#8,
25 and #9 even though they arrived in order 8,9,7,6.

But, if preserving order is not important, the RIDer can just process the buffers of formatted data as they arrive, regardless of their sequence ID's. (i.e.. in this example the RIDer would process buffers in the order # 8, #9, #7, and then #6) So the RIDer doesn't need to wait

CA9-97-022

15

for the correct buffer to show up, and wastes less time waiting.

This variation of the invention takes advantage of the parallel processing capabilities of the LOADER of the invention providing significant speed advantages but does not preserve the entire original record sequence.

EXAMPLE

The following example should assist in clarifying one embodiment of the invention. In this embodiment user data is stored in a file on disk, for example, in a format compatible with the LOAD utility of the database management system in use on the data processing system in use.

Shown in this example is a 3 column table stored in delimited ASCII. The 3 columns are integer, 18 byte character, and floating point, respectively. Note that a new line character, 0x0A, is implied at the end of each line. Columns are delimited by the “,” character:

5, My dog, 23.4

7, Your dog, 69.7

12, Bob's dog, 30.0

14, Sue's dog, 10.4

11, Ken's dog, 11.2

9, John's dog, 26.1

The user data can be represented as follows in memory:

The data, in HEX representation looks like:

352C4D79 20646F67 2C32332E 340A372C

596F7572 20646F67 2C36392E 370A3132

2C426F62 27732064 6F672C33 302E300A

31342C53 75652773 20646F67 2C31302E

The same data in buffered ASCII:

5,My dog,23.4,7,

Your dog, 69.7,7,12

,Bob's dog,30.0,

14,Sue's dog,10.

CA9-97-022

16

340A3131	2C4B656E	27732064	6F672C31	4,11, Ken's dog, 1
312E320A	392C4A6F	686E2773	20646F67	1.2,9, John's dog
2C32362E	31			,26.1

- 5 The Reader agent stores data in a buffer, where a few bytes of storage are reserved for the sequence tag. For example, if the sequence tag was a 4 byte integer, and the memory set-aside for this tag was the first four bytes in the memory buffer, then the tagged memory (tagged with sequence 0x00000000), would have the following HEX representation:

10 00000000

352C4D79	20646F67	2C32332E	340A372C
596F7572	20646F67	2C36392E	370A3132
2C426F62	27732064	6F672C33	302E300A
31342C53	75652773	20646F67	2C31302E
15 340A3131	2C4B656E	27732064	6F672C31
312E320A	392C4A6F	686E2773	20646F67
2C32362E	310A0A		

- 20 If the same data happened to appear in the second buffer of user data, it would be tagged with sequence ID 1 (instead of 0). In HEX form, this is 0x00000001. The memory buffer would then look like:

00000001

352C4D79	20646F67	2C32332E	340A372C
596F7572	20646F67	2C36392E	370A3132
25 2C426F62	27732064	6F672C33	302E300A
31342C53	75652773	20646F67	2C31302E
340A3131	2C4B656E	27732064	6F672C31
312E320A	392C4A6F	686E2773	20646F67
2C32362E	310A0A		

CA9-97-022

17

When the formatter agent receives this buffer, it immediately can identify the sequence ID of the data, and thus it's relative position in the stream of data by reading, and storing this sequence ID.

- 5 the formatter converts each of the user records to a form consistent with the storage form of records in the DBMS. The formatted DBMS type records are stored in a buffer. The buffer is tagged with the same sequence ID as the buffer holding the corresponding user data.

10 For example: Consider a 64K buffer holding a list of DBMS style records, where the first 4 bytes of the buffer was reserved for a sequence ID.

4 byte seq. ID

N1 bytes representing DBMS form of user record 1

N2 bytes representing DBMS form of user record 2

N3 bytes representing DBMS form of user record 3

15 N4 bytes representing DBMS form of user record 4

N5 bytes representing DBMS form of user record 5

N6 bytes representing DBMS form of user record 6

remaining unused bytes in 64K buffer.

20 Note that the number of bytes required to represent the data record in DBMS (database management system) form may vary from record to record. N1,N6 may or may not be the same.

CA9-97-022

18

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

- 1 1. A method of loading data having a predetermined order of data records from a source
2 into a database using parallel processing comprising:
3 using one or more reading agents, reading data records from the source in which they
4 are stored;
5 storing said records in groups of records;
6 tagging each group with a sequence identifier corresponding to said predetermined
7 order of data;
8 formatting the data records in said groups using a plurality of formatting agents in
9 parallel;
10 storing said formatted records in formatted record groups;
11 writing said sequence identifier of each said group to the respective formatted record
12 group;
13 assigning page locations for said records using said sequence identifiers to maintain said
14 predetermined order;

- 1 2. The method of claim 1 including:
2 assembling said records into pages.

- 1 3. The method of claim 2 including:
2 writing said pages to storage at the locations specified.

- 1 4. The method of claim 1 in which said reader agents read nK bytes of user data into the
2 bottom half of a memory buffer area of $2nK$ bytes;
3 strips any existing partial record from the bottom of said nK bytes of data just read;
4 stores a copy of said partial record in a temporary storage area;
5 if a partial record of a previous read operation exists, then prefixes that partial record

CA9-97-022

19

6 to the start of said nK memory area;
7 tags said memory storage area with a sequence identifier;
8 transfers control of said memory buffer to one of said formatter agents;
9 continues reading the next nK bytes of user data until all user data is read.

1 5. For a MULTI PROCESSOR data processing system capable of parallel processing, a
2 method of loading data having a predetermined order of data records into a database using
3 parallel processing comprising:

4 using one or more reading agents, reading data records from the source in which they
5 are stored;

6 storing said records in memory buffers;

7 tagging each buffer with a sequence identifier corresponding to said predetermined
8 order of data;

9 transferring control of each said buffer to the control of one of a plurality of formatting
10 agents;

11 converting said records of data into a suitable database storage format;

12 storing said formatted records in formatted record buffers ;

13 writing said sequence identifier of each said buffer to the respective said formatted
14 record buffer;

15 forwarding said formatted record buffers to a record identifier agent;

16 assigning a page location for each record in said formatted record buffers using said
17 sequence identifiers to maintain said predetermined order.

1 6. The method of claim 5 in which said formatted record buffers are processed by a
2 record identifier agent as they are received.

1 7. The method of claim 6 in which said formatted record buffers are processed by a
2 record identifier in the order of their sequence identifiers.

CA9-97-022

20

1 8. The method of claim 5 including assembling said records into pages and writing said
2 pages to storage at the location specified.

1 9. For a parallel processing system, a method of loading data records having a pre selected
2 order from a source location, in parallel, into a database, while maintaining said order,
3 comprising:

4 reading a record from said source location;

5 storing said record in memory buffers of predetermined size;

6 tagging each buffer with a sequence number corresponding to said pre selected order
7 of said record in its source location;

8 transferring control of said buffers to a plurality of formatter agents which operate in
9 parallel;

10 converting each record to a suitable database storage format;

11 storing said formatted records in formatted record buffers;

12 writing a sequence identification number of each said buffer to said corresponding
13 formatted record buffer;

14 transferring control of said formatted record buffers to a RIDer agent;

15 for each formatted record buffer being received in sequence assigning a page location
16 to each record;

17 transferring control of said formatted data buffer to a plurality of writer agents;

18 assembling formatted records into pages

19 writing said pages to said assigned storage locations.

1 10. The method of claim 9 for a parallel processing system, a method of loading data
2 records having a pre selected order from a source location, in parallel, into a database, while
3 maintaining said order, comprising:

4 reading a record from said source location;

5 storing said record in memory buffers of predetermined size;

6 tagging each buffer with a sequence number corresponding to said pre selected order

CA9-97-022

21

7 of said record in its source location;
8 transferring control of said buffers to a plurality of formatter agents;
9 converting each record to a suitable database storage format;
10 storing said formatted records in formatted record buffers;
11 writing a sequence identification number of each said buffer to said corresponding
12 formatted record buffer;
13 transferring control of said formatted record buffers to a RIDer agent;
14 for each formatted record buffer being received processing said formatted record
15 buffers in order of their respective sequence identification numbers assigning page locations
16 to said records maintaining said pre selected order by means of said sequence number order
17 and sequence identification number;
18 transferring control of said formatted data buffer to a plurality of writer agents;
19 assembling formatted records into pages writing said pages to said assigned storage
20 locations.

1 11. Apparatus for loading data records having a pre selected order using a parallel data
2 processing system to load said data into the data storage of a database using said data
3 processing system while maintaining said pre selected data order, comprising;
4 at least one reader agent to read said data;
5 buffer storage means for storing said data in buffers;
6 tagging means for tagging each buffer containing said data with a sequence number
7 corresponding to the pre selected order of the data;
8 a plurality of formatter agents for converting said data records read by said at least one
9 reader agent to a format suitable for said database;
10 means for storing said formatted records in formatted record buffers, with the
11 sequence numbers of the buffers containing said data;
12 a RIDer agent for assigning page locations for each said formatted data records;
13 means for assembling said formatted records into pages;
14 a plurality of writer agents for writing each said pages to said assigned storage locations .

1 12. A computer program product comprising:

2 a computer usable medium having computer readable program code means embodied
3 therein for causing a computer to load user data, the computer program product comprising :

4 computer readable program code means for causing a computer to effect apparatus
5 for loading data records having a pre selected order using a parallel data processing system to
6 load said data into the data storage of a database using said data processing system while
7 maintaining said pre selected data order, comprising;

8 computer readable program code means for causing a computer to effect at least one
9 reader agent to read said data;

10 computer readable program code means for causing a computer to effect buffer
11 storage means for storing said data in buffers;

12 computer readable program code means for causing a computer to effect tagging
13 means for tagging each buffer containing said data with a sequence number corresponding to
14 the pre selected order of the data;

15 computer readable program code means for causing a computer to effect a plurality
16 of formatter agents for converting said data records read by said at least one reader agent to
17 a format suitable for said database;

18 computer readable program code means for causing a computer to effect means for
19 storing said formatted records in formatted record buffers, with the sequence numbers of the
20 buffers containing said data;

21 computer readable program code means for causing a computer to effect a (RIDer)
22 agent for assigning page locations for each said formatted data records;

23 computer readable program code means for causing a computer to effect means for
24 assembling said formatted records into pages, and;

25 computer readable program code means for causing a computer to effect a plurality
26 of writer agents for writing each said pages to said assigned storage locations.

1 13. A program storage device readable by a machine, tangibly embodying a program of

CA9-97-022

23

2 instructions executable by the machine to perform method steps for loading data having a
3 predetermined order of data records from a source into a database using parallel processing
4 said method steps comprising:

5 using one or more reading agents, reading data records from the source in which they
6 are stored;

7 storing said records in groups of records;

8 tagging each group with a sequence identifier corresponding to said predetermined
9 order of data;

10 formatting the data records in said groups using a plurality of formatting agents in
11 parallel;

12 storing said formatted records in formatted record groups;

13 writing said sequence identifier of each said group to the respective formatted record
14 group;

15 assigning page locations for said records using said sequence identifiers to maintain said
16 predetermined order.

FIGURE 1

**Flow Chart For Load SMP Exploitation
With Record Sequence Preservation**

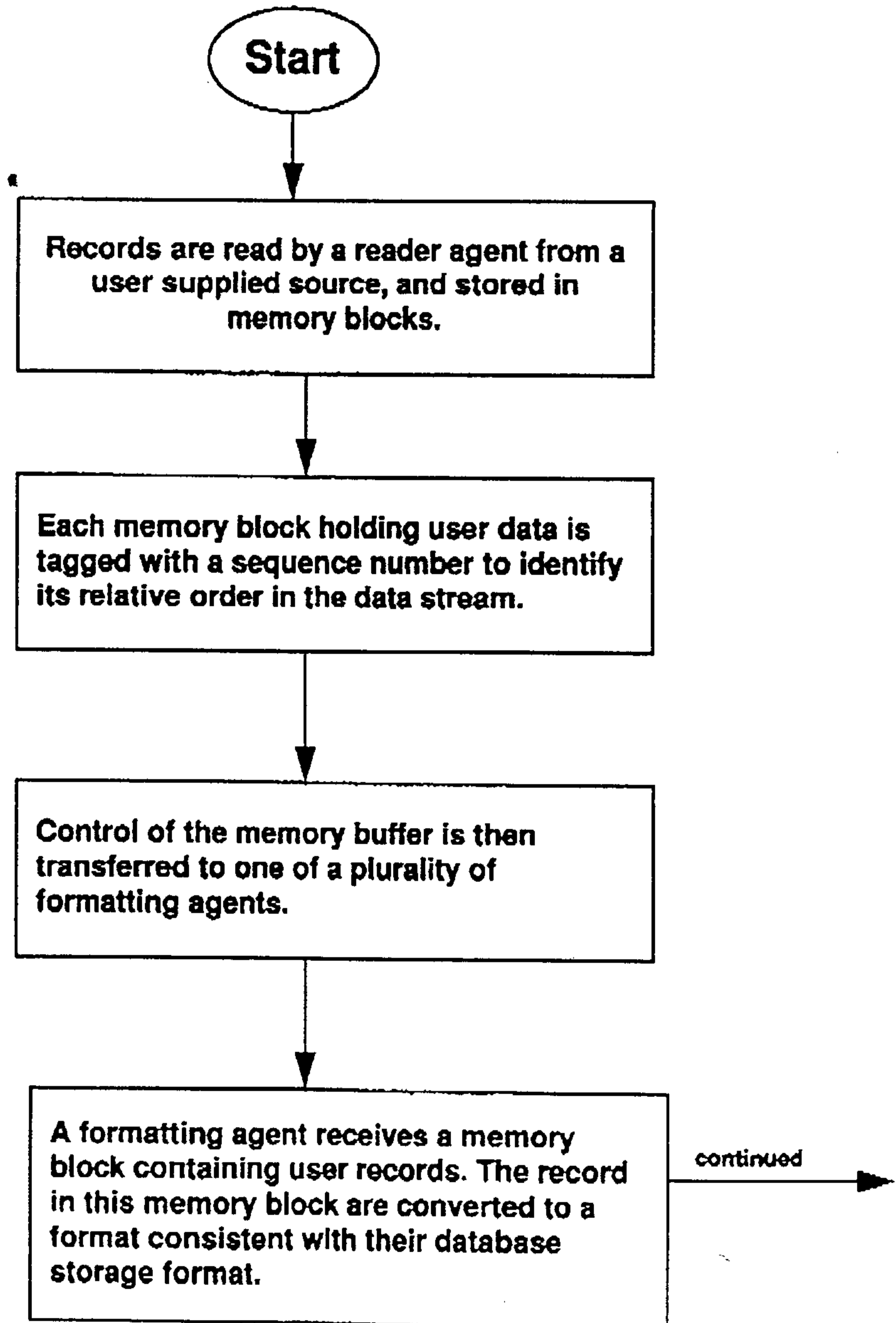


Figure 1 (Con't)

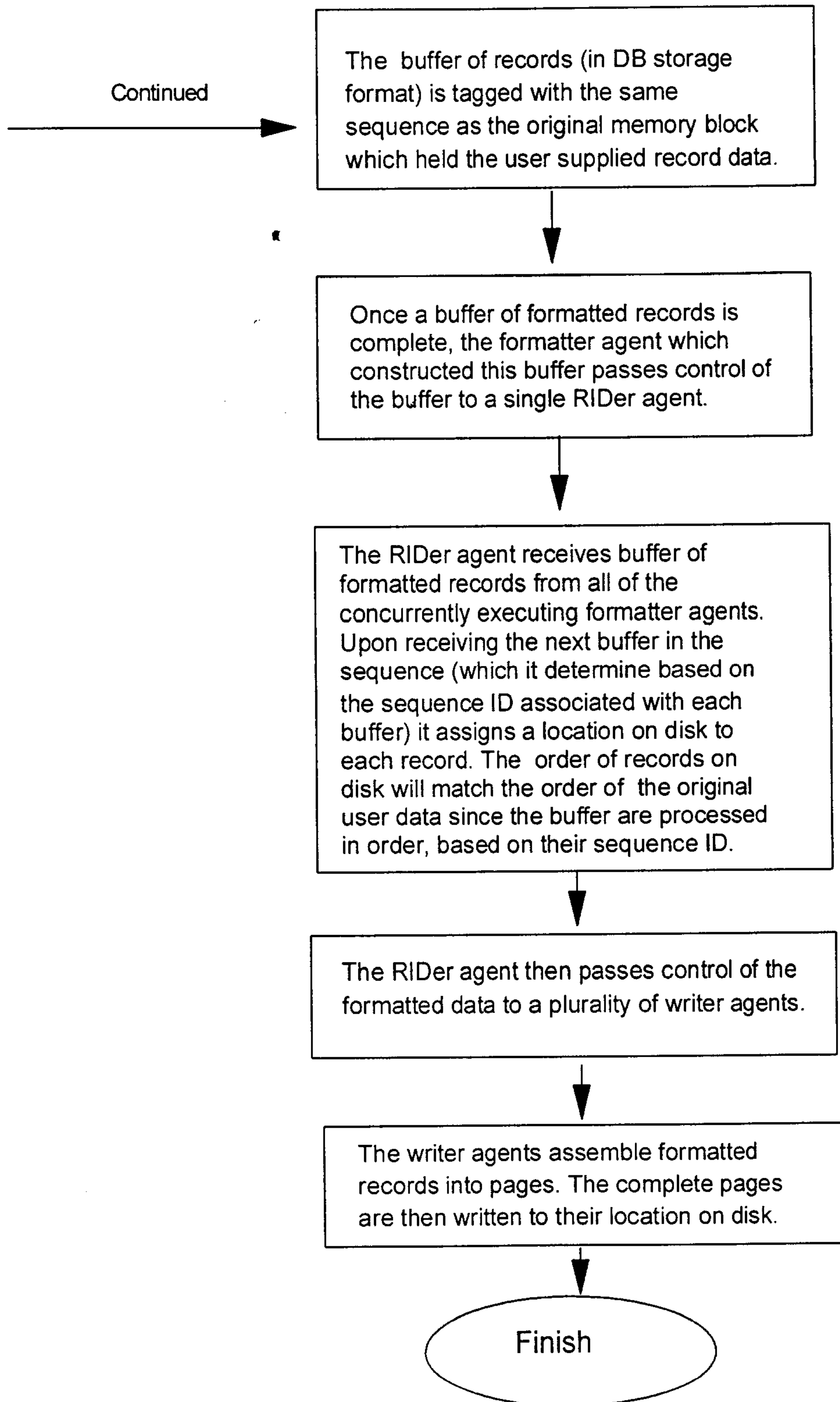


FIGURE 2

Process Model For the Preferred Embodiment of the Invention

