THE ROLE OF DBMS IN ANALYTICAL PROCESSES OF THE LOGISTIC OF STOCK RESERVES

Julian Vasilev

Abstract: One of main problems of corporate information systems is the precise evaluation of speed of transactions and the speed of making reports. The core of the problem is based on the DBMS that is used. Most DBMS which are oriented for high performance and reliability of transactions do not give fast access to analytical and summarized data and vice versa. It is quite difficult to estimate which class of database to use. The author of the article gives a concise observation of the problem and a possible way to be solved.

Keywords: Database management systems (DBMS), Information technology, Cache, Interbase, post-relational DBMS.

ACM Classification Keywords: H.2.8 Database Applications, H.4 information systems applications.

Introduction

Stock reserves affect all activities in the enterprise. Their management is directly connected to the appliance of different methods. According to Gatorna [1,333] "…logistics includes two main spheres of activity: management of materials entering the production and the management of distribution of final products."

According to the information needs of the operational management we meet requirement of different groups of users who give priorities for fastness of defined actions. For instance, people who get stock from suppliers and store it in the factory warehouse want high speed of their transactions. By analogy with supplies, the sales department needs high speed in making, saving and printing invoices. Having in mind these requirements, the key role of the DBMS (Database management system) is to provide high speed of transactions. As we mentioned these transactions concern registering documents in the information system of the enterprise. In practice certain raw materials are stored and "wait" their participation in production. At the end of the enterprise final products form also stock reserves. This type of production activities is usual for Bulgarian enterprises. Without involving logistics, these activities are organized in accounting software.

Layout

According to managers, production operations are subject to monitoring using certain indicators, such as effectiveness, profit, costs, revenues. A full control of material flows is a precondition for the science logistics to give a solution to several problems in the sphere of creating an order for supply. Supply department needs to know the amount of order, the frequency of sending, stocks included, supplier. Analytical processes refer to a higher level of data aggregation and extracting synthesized data. Some authors [2,100] give account of necessity of expert systems in logistics for solving complex problems. According to other experts [3,182] in building "computers are used for solving complex problems of planning and for techno-economical problems when building plans for material and technical providence". According to other authors [5] "multi-dimensional data structures are the base of the conception of direct analytical process". Top managers are interested in the dynamics of several indicators in order to monitor the state of the organization. The warehouse of the enterprise generates big amounts of data. Data increases throughout the time. Observations in practice show that new technical data storage devices are with bigger volume, than their precedents, so data storage is not a problem for IT specialists. Problems occur when we need to show dynamic indicators. The use of DBMS, accenting to high speed of transactional performance, for example Cache (a product of Intersystems), guarantees quick and reliable registration of data. DBMS of that class usually make reports slower than other classes of DBMS such as Oracle or MS SQL Server. The application of last two databases is combined with a spread application development instruments for OLAP (on-line analytical data processing) data analysis, which make them adjusted to solve managers' problems of high level. Information systems based on Oracle require significant hardware

resources. Another disadvantage is the speed of transactions. Keeping indexed data in relational structures is a requirement for fast speed of OLAP instruments. But DBMS need much time to keep indexed data actual. With increase of data more time is needed for a transaction to be saved. We made an observation. Results are summarized in Table 1.

Table 1: Comparison between Cache, MS SQL Server and MS Access

DBMS/Indictor	Cache	MS SQL Server	MS Access
SQL access	yes	yes	yes
Support of huge data structures	yes	yes	yes
Speed of grouping and summarizing data	low	high	middle
Support of OLAP	no	yes	no
Speed of saving transactions	high	middle	low

We consider that the key problem of the DBMS role is the seeking and finding of an objective compromise between high speed of transactions and high speed of analytical processes in management. Most IT experts prefer popular DBMS such as Oracle, MS Access, Interbase, DB2, MS SQL Server. Advertisements in IT magazines and newspapers, application development environments make them attractive. Managers need to monitor a set of economic indicators concerning activities in the enterprise. The compromise is oriented to high speed of registration accounting documents from one side and to the logistics on the other side. A huge document flow requires high speed of transactional performance. If it is not provided, customer services are slowed down. Moreover the company needs personnel to process these documents. Customer services have to be fast and with high quality. These are main priorities of marketing. Whereas future sales are subject to research of the science “Forecasting”, future supplies are a matter of organization of the procurement department and are subject of discussion by scientists in the sphere of logistics. Forecasting sales and organization of future supplies is deeply connected to analysis of big arrays of data, meaningless methods which can be applied. Backing our opinion other authors [4, 57] think that “Forecasting of material requirements is based on values from historical data. Proving of future needs is helped by statistical methods”.

Each DBMS is optimized either by the processing of transactions (OLTP – on-line transaction processing) or – by analytical processing of data - OLAP (on-line analytical data processing). Their performance is compared by the execution of test done by independent companies, such as TPC – Transactional Processing Council. SQL Server support tools for optimizing queries, such as SQL Query Analyzer, Query Governor – for the determination of priorities of execution of queries. The creators of DBMS Cache have built several tools to fasten processing of queries. One of them is called “Write Demon”. All queries which require operations including adding or modifying data are non-synchronic – the server accepts the query and gives back to the client a flag for successful execution. In the meanwhile, Write Demon adds or modifies data. Another process is called “Garbage Collector”. Its purpose is processing of queries for deleting data. It functions likewise “Write Demon”. More information about comparing DBMS-s can be read in an article, published by Bloor Research [7, 64].

As we said the fastness of a DBMS can be in the sphere of transaction processing or analytical processing. If we study the problem carefully we will find out that the approach for making indexes in the DBMS is the nostrum. If we compare Intersystem’s Cache and MS SQL Server the approaches for making indexes are different. SQL Server uses B-tree technology for indexing. It supports cluster and non-cluster indexes. The common rule is: the more indexes we have in a table, the more time we wait on adding data, the less time we wait on extracting summarized data. Cache uses block structure to save physically records of data. The organization of blocks resembles the file allocation structure in operating systems: UNIX, DOS, Windows. Cache is optimized for multi-user access and processing of high intensity flow of transactions. Grouping and searching of data is executed slower than relational DBMS. For the purpose of retrieving data faster, designers of information systems have to create artificial structures analogical to index data. This approach leads to: firstly, increasing the size of the database because doubled data is stored and secondly, business logic complicates. In contrast to relational DBMS where adding of data takes more time, the same operation in a post-relational DBMS takes less time. For the purpose of our research we made a comparison between Cache and Interbase. Results are summarized in Table 2.

Table 2: Comparison between Cache and Interbase

DBMS/Indicator	Cache	Interbase
Time to insert 100 000 records in the database, each is 255 bytes long	8 seconds	350 seconds
Filtering 100 000 records and retrieving 5 000 records by a given criteria	4 seconds	8 seconds

It is indisputable that there is a wide choice of DBMS. In accordance with the opinion of some authors [6] "For building databases commercial DBMS can be used such as ORACLE, MS-SQL, DB2 and others, as well as free versions." We consider that for building corporate systems IT experts cannot rely on DBMS which does not have support. Experiments were done on Win XP, with 128 MB RAM, Intel Celeron processor 466 MHz. To visualize resulting data we used following technologies: HTML (hyper text markup language), ASP (active server pages) and GUI (graphic user interface) application. The GUI application visualizes data faster than ASP technology which needs an IIS (Internet information server).

In a post-relational DBMS (such as Cache) records are stored as blocks. Each block has: a unique address, Boolean flag (free or occupied), data range, address of the next block. Adding data means rerouting the address of previous block to the next block and changing the flag from "free" to "reserved".

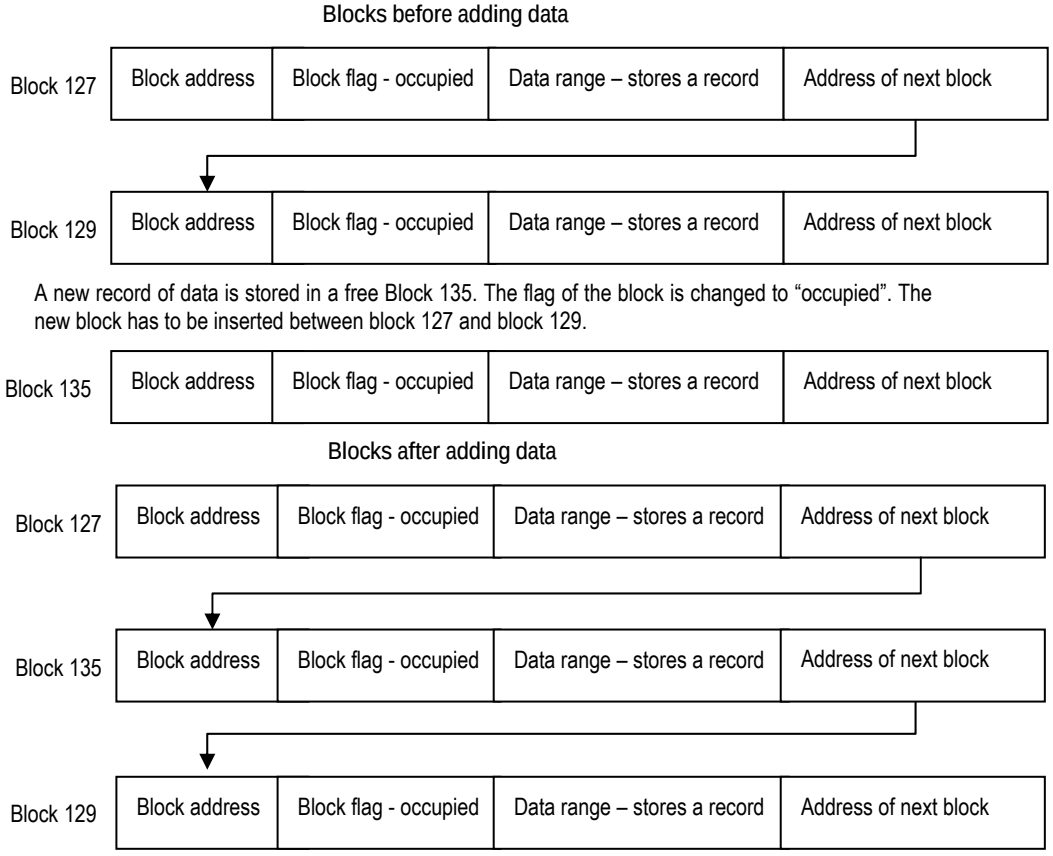

Fig. 1. Adding data in block structures

This organization of adding data resembles the management of RAM (random access memory) and the management of queues and stacks. The difference is that blocks are stored on hard disk drives. Deleting of data is executed in the following order – the flag of the block is changed to "free", bits in data range are set to "0" and the pointer "address to next block" is set to "null". In figure 2 we can see sample deletion of block 135 (which we already added).

Computer programmers don't manipulate directly blocks. They write "Insert" or "Delete" SQL clauses. These clauses are interpreted by Cache and transformed to block operations. In this way programmers use the

language MUMPS (Massachusetts University Multiprogramming System), COSL (Cache Object Script Language) or SQL queries.

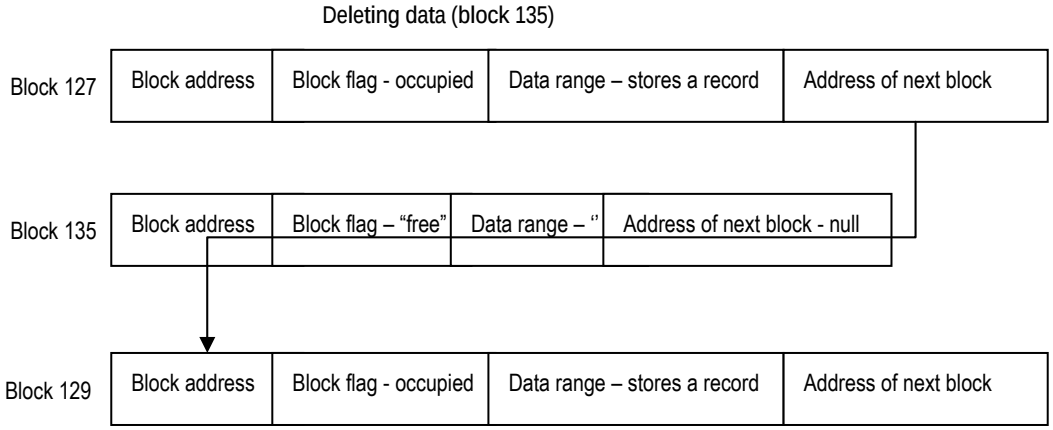

Fig. 2 Deleting data in block structures

For now, there is not a DBMS which is optimized in both directions – transactional processing and analytical processing. That is why, when a company wants to choose a DBMS, it has to estimate which category of DBMS needs. In the sphere of logistics – for the registering a high intensity flow of transactions we need an OLTP system. From another side, for the decision support processes we need OLAP instruments and an analytical DBMS. These facts open another problem – defining the interface between both DBMSs.

Conclusions

Estimating DBMS we need to bare in mind not only its popularity but its orientation to fast transactions or high speed of analytical processes. The right choice is based on logistic processes, hardware resources, personnel and an objective forecast for the increase of data arrays. The last factor is usually ignored and sometimes it is decisive. Its correct evaluation is done after several years of functioning of a store or accounting system. One of the ways to solve this problem is to be built an application server, which acts as a Windows process which transports data from one DMBS to another. For instance we can have an information system based on Cache, and a second one based on MS SQL Server. The first one is used for registering documents such as invoices and stock receipts, the second one for OLAP analysis. The mediator is the upper mentioned application server. In our opinion it is a temporary solution until further development of DBMSs.

Bibliography

[1] М Гаторна, Д. Основи на логистиката и дистрибуцията, Бургас, 1996.

[2] Гаджинский, А. Логистика, Москва, 2000.

[3] Генов, Г. Логистика в строителството, ИУ Варна, 1998.

[4] Коралиев, Я. Логистика на снабдяването, София, 1998.

[5] Несторов, К. Организация на данни в програмни продукти за счетоводно отчитане, сп. Известия, ИУ Варна, бр. 2, 2002г.

[6] Петров, П. Проблеми на изграждането на автоматизирани информационни системи в големи производствени предприятия сп. Известия, ИУ Варна, бр. 1, 2002г.

[7] Howard, P., On-Line Database Technology, Bloor Research Volume 1, 2004

Author's Information

Julian Vasilev – Chief assistant professor, Department of Informatics, Varna University of Economics; 77, Kniaz Boris I str.; Varna; Bulgaria; e-mail: vasilev@ue-varna.bg