Methods of Using Business Intelligence Technologies for Dynamic Database Performance Administration

Veska Mihova*

Summary:

Today productivity and performance are the main problems of business applications. Business applications become increasingly slow while their stored and managed data are growing.

The problem of monitoring and productivity optimization of a business application's database is a traditional and well-known problem. Therefore, there are many various concepts and software products developed for its solution. But all solutions perform monitoring on the current state of the information system, which makes the process of optimization slow and ineffective.

The current paper proposes a concept for monitoring the future state of a database. This solution gives the business enough time for reaction to an occurring problem with the application's performance and respectively the database's performance. The concept presents a metadata model for generating a specific data warehouse that is generated in dependence of specific RDBMS. The proposed architecture offers an additional layer of web services which provides the necessary data for monitoring the future state of the database so that any business organization can integrate, visualize and implement the data in a different and appropriate way. Also the current paper proposes methods for dynamic database performance administration through which the overall concept is created.

Key words: database, database monitoring, database administration, database performance, performance forecast.

JEL Classification: C8, L86.

1. Introduction

The main problems of business applications are relevant to their productivity and performance.

Business information systems' productivity is an indicator determining the data processes' execution time and users' response time.

Together with the complex information systems' architecture, a very important indication for the development of the information system is the permanent growth of stored and managed data. This growth worsens the system's productivity and decreases the system's performance.

The solution of the problem with the systems' productivity was created, but in some aspects there are serious weaknesses provoking the search for better solutions.

* PhD, University of National and World Economy, Department of Information technologies and communications
e-mail: vmihova@unwe.bg

The problem of business information systems' monitoring and optimization is a current topic in recent years because the need for a solution is recognized in more and more business areas.

1. State of the art

With the development of new technologies in the modern world, the work of a business organization is monitored by the so-called monitoring systems. The monitoring of an information system's operation and database monitoring in particular is a task which has been defined and implemented for a long time. The problem of monitoring and productivity optimization of the database of a business application is a traditional problem and for its solution there are many various concepts and software products developed.

The basic concepts are developed in two main directions: monitoring of the condition and providing opportunities for change. Most of the modern database management systems (DBMS) integrate various implementations of these concepts in the form of software components, system elements and integrates external ones.

1.1. Monitoring

Monitoring is a continuous process of collecting information to monitor the implementation of a process. (Bartle, 2011) It is an ongoing and systematic collection and analysis of information which is used in the process of further management, control and decision making.

Monitoring is literally following a certain process - monitoring, analysis and evaluation. Monitoring can be realized on many and different levels - on a particular activity, on a project, on an information system or on the overall performance and effectiveness of a unit, on the operation and development of a team or on individual members, etc.

2.2. Main principles of database monitoring

The modern means of database monitoring and proposal for database optimization techniques are based on the available data on the current state of the database, described by qualitative indicators such as volume data, CPU activity, number of concurrent queries, execution time of queries, frequency of execution of queries, etc. Providing such information for the database status is a task which is solved by different means and instruments in various database management systems. (Lane, 2013) (ManageEngine, 2015) (ManageEngine, 2012)

Database monitoring systems perform present time monitoring, systematic and periodic collection and use of information for management control and decision-making. The monitoring systems are designed to monitor the work, the results in a system and the resources which are available, whether they are sufficient and whether they are well-used by the system. The purpose of database monitoring systems is to provide information about the system and to provide effective and efficient database system operation.

Monitoring systems provide a consolidated summary that exposes the current status of the monitored system. These reports help database administrators and system administrators to stay informed of current problems. (Yuhanna, et al., 2007) (Proctor, et al., 2009)

The existing systems in this area include monitoring the system operation, the database productivity, the critical points of systems, and the detection of errors and system-related problems. The problem is that they follow the current state, i.e. when the problem is ongoing. A stable decision would be one that allows predicting upcoming problems. However, the stable decision would one that

suggests appropriate behavior in case an exceptional situation occurs (Murjeva a, et al., 2011) (Murjeva b, et al., 2010). Data forecast and data correlation extraction is well-known activities in other scientific areas and they require data history on the basis of which the forecast values can be calculated. Searching for a solution in this direction should be preceded by the monitoring of the information system and the collection of sufficient information in terms of both in quantity and quality about it.

2.3. Monitoring of the future state of the database

The mentioned above clarifies what monitoring of a database is and what the information is available for monitoring. It is also clear how to collect information for a database as well as the fact that monitoring is performed on the state of the information system that makes optimization slow and ineffective. On the basis of the collected information for the work of the database it is clear that the future state of the database may be forecasted, as science provides a powerful statistical forecasting unit.

Given all that, the next interesting problem appears - is it possible to combine all to create monitoring for the future or, in other words, is prediction monitoring possible. (Murjeva a, et al., 2011) (Murjeva b, et al., 2010)

Applying any technique for optimizing the performance and respectively to remove the problem in most cases requires a long period of time for taking an adequate decision. The reason is that in some cases it may be necessary to apply more drastic techniques in order to resolve the problem, e.g. redesign of database which requires enough time for planning the activities, distribution of the problems, organization and planning for solving the problem.

2.4. Data provision

Monitoring the performance of information systems is related to the collection of data for occurring events (applications, processes, etc.). Modern technologies allow toperform monitoring at each layer of the system's architecture but the highest level of data provision is the layer for data management.

This data provision is better due to the database management systems supply of large number of data and statistics for the functioning of the system.

For performing qualitative monitoring, the information that should be collected is for: (Lane, 2013) (Oracle a, 2015) (Oracle b, 2014) (Oracle c, 2013)

- Data volume;
- The execution of queries;
- The execution time of queries;
- Execution plans;
- Database users;
- Loading of the system, i.e. simultaneous operation of multiple users, number of user connections to databases at different times;
- Time to connect to the database;
- Number of queries to the database;
- Regularity and frequency of executions of queries;
- The database structure (tables, columns, relations, etc.);
- Number of created indexes;
- The operation of the cache buffer;
- The execution of transactions;
- The number and duration of execution of current operations;
- The number and duration of sessions;
- Number and duration of the current lock resources;
- And others.

The information for the performance of a system may be provided from the database management system itself. Each system provides system tables and

views which contain information about the structure of the database, its work and the work of the system. (Bartle, 2011) (Oracle a, 2015) This information may be used for monitoring, as it is available at different levels of detail – individual values, cumulative values. This information is available in each database management system and it represents the current status of database including the collected statistics which are a picture of the database. The larger DBMS (database management system) provide functionality for collecting historical information, but with significant restrictions. Therefore the layer of data management in the whole architecture of the business application is well provided with information. This is a good opportunity for collecting historical information, which could be used for forecasting the future state.

2.5. Architecture of business intelligent system for monitoring the current and the future state of a database

The next interesting task is whether it is possible to simulate the state of the database at a later point in time on the basis of the forecasted values of its state. As a solution the paper proposes creating forecasted reality, a database that contains only the data necessary for the database optimizer without really having data. (Murjeva a, et al., 2011)

Since the basic optimization techniques and the monitoring systems are based on the notion of statistics, it is interesting whether these statistics can be replaced with the forecasted ones. The idea is to get the future database so that the optimizing systems can offer how to optimize the forecast database (figure 1).

The proposed concept would give the business enough time for reaction and timely organization of problem-solving.

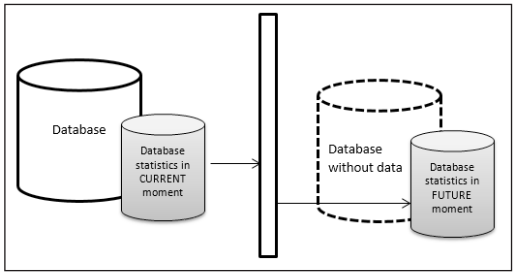

Fig. 1. Replacement of the database statistics with forecast database statistics

The presented idea covers tasks and concepts that are very similar to those of business intelligent systems - data collection, storage history, producing new values based on the accumulated history, forecast. (Eckerson, 2005) (Brackett, 1999) (Quarles van Ufford, 2002) For this reason, this paper takes the idea of the architecture of business intelligent systems and develops it further to serve the idea of future forecasted environment, monitoring the expected state of the database.

This idea can be realized with the help of a business intelligent system which would have the following architecture as shown in figure 2:

The data source in the proposed architecture is the database of the business information system and the database that stores metadata for the management process of generating the structure of the data warehouse. This metadata is stored with a set purpose for the automatic generation of data warehouse depending on the specifics of the database. This metadata can indicate what the corresponding database's name is, what DBMS it works with, and the path to the database, according to how long it takes to make a forecast and other relevant data. The idea is to generate a data warehouse with a structure that is appropriate to a different DBMS. The data warehouse is designed to store historical information that is extracted from the system tables and tables of statistics of a particular DBMS so its data

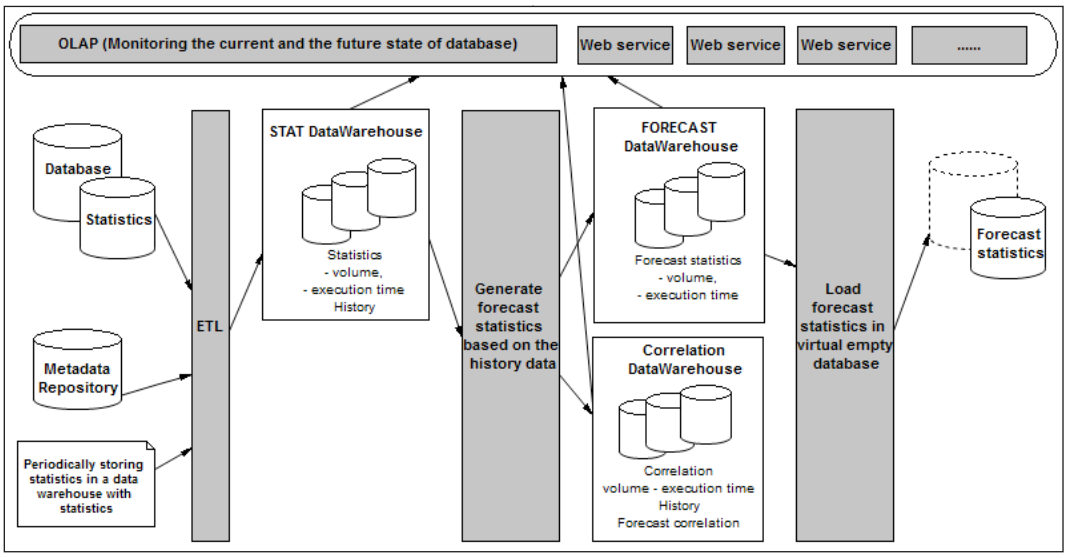

Fig. 2. General BI architecture for monitoring the current and the future state of a database

source is potentially one and homogeneous. Therefore it's reasonable to propose storage with an analogous structure to the system tables, equipped with an indicator of time. In such a structure it is possible to provide procedures for generating these structure-based meta descriptions.

It follows a model of metadata for managing the generating process and the generation of specific data warehouse, shown in figure 3 (Murjeva, A. et al., 2011) (Microsoft, 2015):

The data warehouse is loaded by ETL processes. The data warehouse is the structure of tables that stores statistics for the execution time of queries, for data volumes in different database objects and other necessary relevant statistics for generating forecast values.

ETL tools, stored procedures, offline processing or other appropriate tools provide generation capability of statistical values at a precise time, depending on how dynamic the process of system operation is.

After collecting enough historical information about the system's operation, periodically predicted values are generated based on the gathered statistics. The generation of these forecasted statistics can be done using statistical packages, Data Mining, stored procedures, MS Excel and many other instruments.

In the proposed architecture a data warehouse component with forecasted data is provided. This is still storage structure, since this is historical data, accumulated values for different periods, with the only specific that these periods are in the future. Apart from data warehouse with forecasted information about data values that are stored in the system and the execution time of queries, the architecture allows collecting correlation statistics between data volumes and the execution time of queries for a certain period of time possible.

The correlation coefficient between the two variables is calculated and stored within the data warehouse.

As the survey shows, at different moments of time one query is affected by different resources. The idea is that BIS can accumulate historical information about these dependencies and, within a particular time, can record their correlation value.

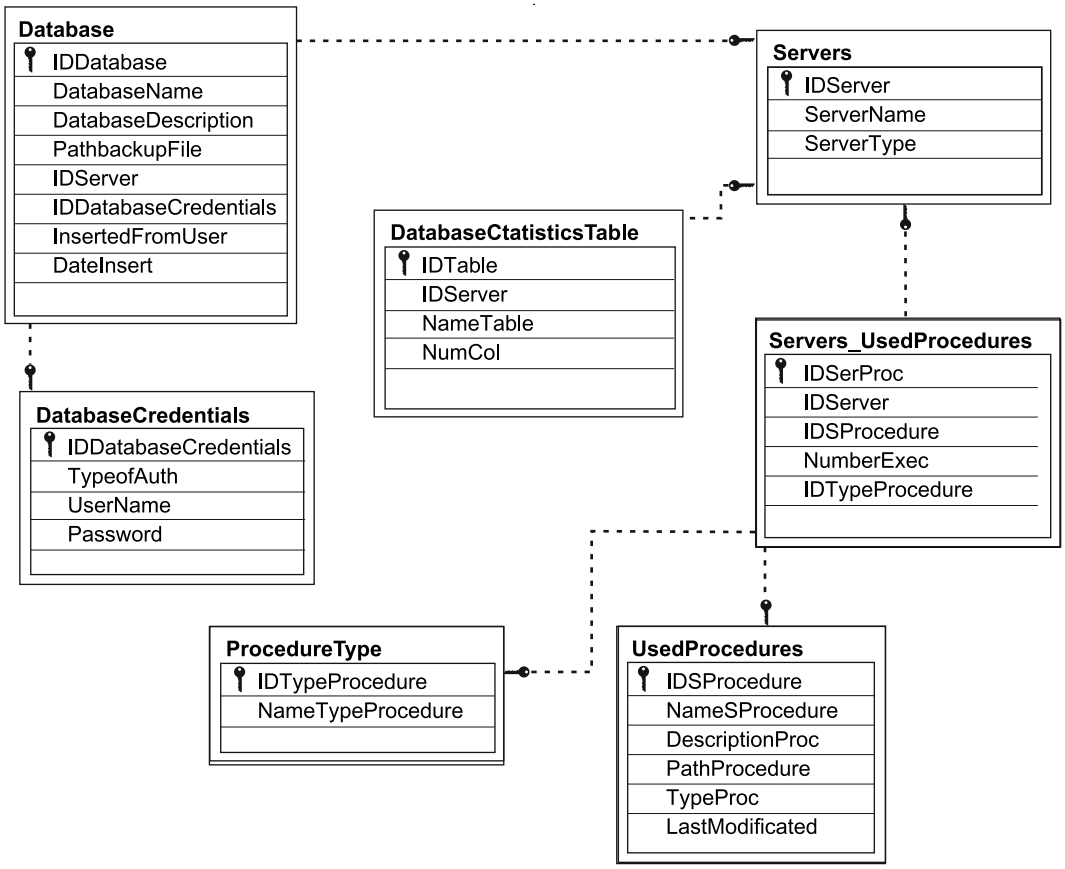

Fig. 3. Metadata model for generation management of a specific Data Warehouse

After enough forecasted database statistics are generated, these forecasted values will be loaded in a new empty virtual database. That means that actually there will be no data in the database, it will be loaded only with forecasted database statistics.

Considering that monitoring systems are based on statistical data, monitoring of the future state of a database can be performed.

Of course, the monitoring can be realized using different instruments or systems, for example using OLAP tools.

Monitoring is a process that may occur irregularly and remotely, which requires consideration of technological capabilities to create a suitable user interface for presenting the available information. This interface should be appropriate for isolating the specifics of the overall architecture to enable a choice for different technical realizations, depending on the specifics, needs and policies of the organization. Creating a layer of services to be used as interface for accessing data in the overall architecture is a good approach.

Furthermore, the architecture offers an additional layer of web services which provides the necessary data for monitoring the future state of the database so that any business organization can integrate, visualize and implement the data in a different and appropriate way.

2. Methods for dynamic database performance administration

2.1. The Registration of queries for monitoring, forecasting and optimization method:

This part of the system is designed to work with the common properties of the database that will be monitored. This component system should also allow the user – database administrator to specify business problems, represented in the system using as large queries, which the organization needs to optimize and monitor. The user should also be able to set critical thresholds for execution time of the queries.

- Creating stored procedures for generating storage structures for statistical data about the work of the database depending on different servers.

Analysis of tables and views that hold or access statistical data for the different database management systems respectively. Creating stored procedures corresponding to a selected server or database management system.

- Collection and management of data for generating a specific data warehouse.

This step is designed to include collecting of data necessary for creating a specific data warehouse in a prepared in advance universal storage structure.

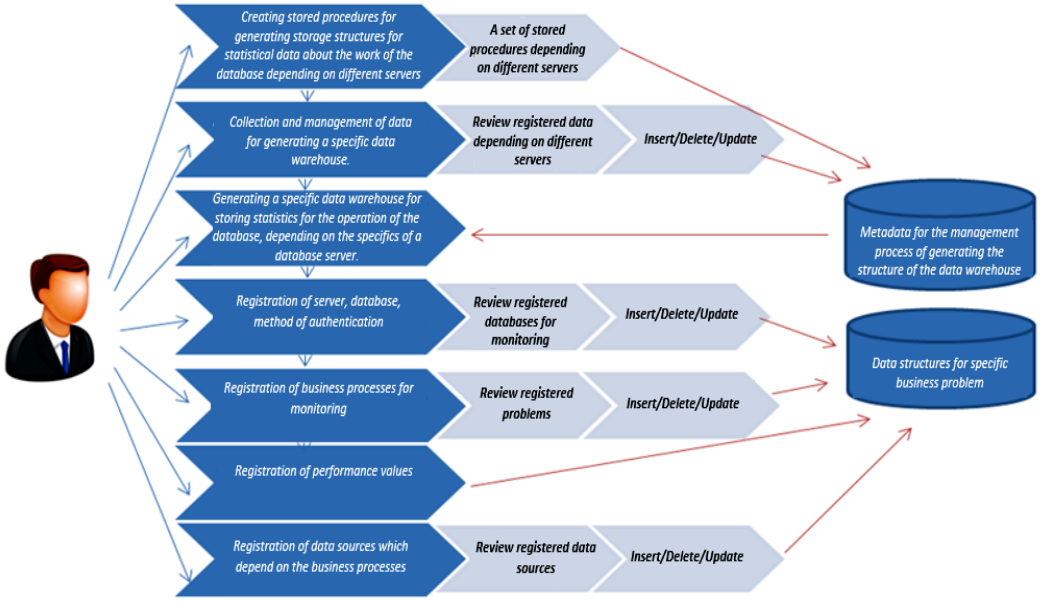

Fig. 4. Sequence of steps in Method 1

The so-called execution steps briefly describe the main functionality of the Registration of queries component system. The sequence of steps involved in this method is shown in figure 4. These are seven steps that require intervention of database administrator:

The data warehouse, which will be generated, is defined as specific, because it is created according to some specific properties such as the database server, which is used for the database and some others.

Each database management system stores statistical data for the work of each database in various structures. These

structures are usually tables that differ not based on the databases, but based on the server and the database management system, on which they are stored and used.

- Generating a specific data warehouse for storing statistics for the operation of the database, depending on the specifics of a database server;
- On this step the data warehouse is generated according to the data for the database, collected during the first two steps.
- Registration of server, database, method of authentication;
- This step includes collecting information for the name of the server of the database which will be monitored as well as the name of the database. Another important thing is to choose a way of identification for the server– windows or sql user.
- Registration of business processes for monitoring;
- The suggested method equalizes the registration of a business problem with registration of a critical for the business in terms of query execution time.
- Registration of sources which depend on the business processes;
- For research purposes there should be registered different data sources, which participate in the studied business process, in order to monitor, forecast and detect potential problems with the database performance depending on the given business process.
- Registration of performance values.

The database administrator is responsible for determining critical thresholds for the execution time of a given query and in this way the process will be monitored according to its optimal values.

2.2. The Forecasting database performance method

This component system is the heart of the system. This basic aim is to collect data for the working of the system based on the previous methods' results for specific business problems. This method requires using statistical methods to forecast values and to measure the relation between specific variables.

The sequence of steps involved in this method is shown in figure 5. These are seven steps that are managed from inputted data in method 1. The intervention of database administrator in this method is minimal. In this part, he is required to set automatic procedures for periodically collecting and generating statistics and forecasting statistics for execution time of query, database objects and also for correlation coefficients and their forecasting values. After that the database administrator only starts the procedures.

- Collection of statistics for execution time of queries;
- Generate forecast values for execution time of queries;
- Collection of statistics for database objects;
- Generate forecast values for the database objects;
- Calculation of correlation coefficients based on statistical data;
- Generate forecast values for correlation coefficients.
- Periodic collection and generation of statistics and their forecast values, calculating the correlation coefficients and generating forecast values for correlation coefficients.

2.3. The Monitoring the current and the future state of the registered queries method

This part of the system allows the database administrator to monitor the system and its indicators through a dynamic dashboard. In this way the administrator will have a view containing all registered

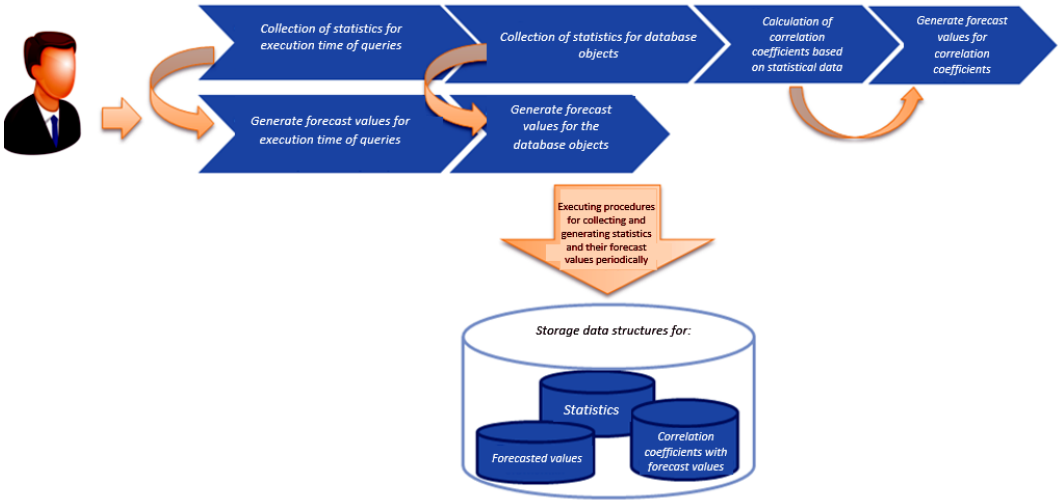

Fig. 5. Sequence of steps in Method 2

business processes, their current and forecasted state, the volumes of the data sources participating in these processes, as well as the execution time for each registered query in separate moments. All these current and forecasted values can be compared to critical thresholds, set for each of the business processes. What is more, in this dynamic dashboards can be monitored and detected dependences between the different data sources participating in a given business process. Monitoring of the calculated forecasted values for detecting correlations between the different sources is a feature of this method too. In this way an accurate and timely decision can be made for detecting and preventing problems with the database performance.

The following steps cover the main functionality of this component. The sequence of steps involved in this method is shown on figure 6. There are six steps in this method that visualize the results of previous two methods with an appropriate tool. Thereby the database administrators receive information about the system's working in the current and the forecasted moment. In this method it is not an expected sequence of steps. The database administrator can perform the steps in order which is needed.

- Monitoring the execution time of registered queries for a specified period compared with critical thresholds;
- Monitoring forecast values for execution time of queries registered for a specified period;
- Monitoring the values of database objects associated with the registered queries;
- Monitoring forecast values of database objects associated with the registered queries;
- Monitoring of the calculated correlation coefficients based on statistical data;
- Monitoring forecast values for correlation coefficients based on statistical data.

2.4. The Database performance optimization method

As all systems for monitoring and optimization of database performance follow the idea of analysing diverse statistics about the objects in the database, this method is aimed to organize the process of entering all the statistics collected by the rest of the

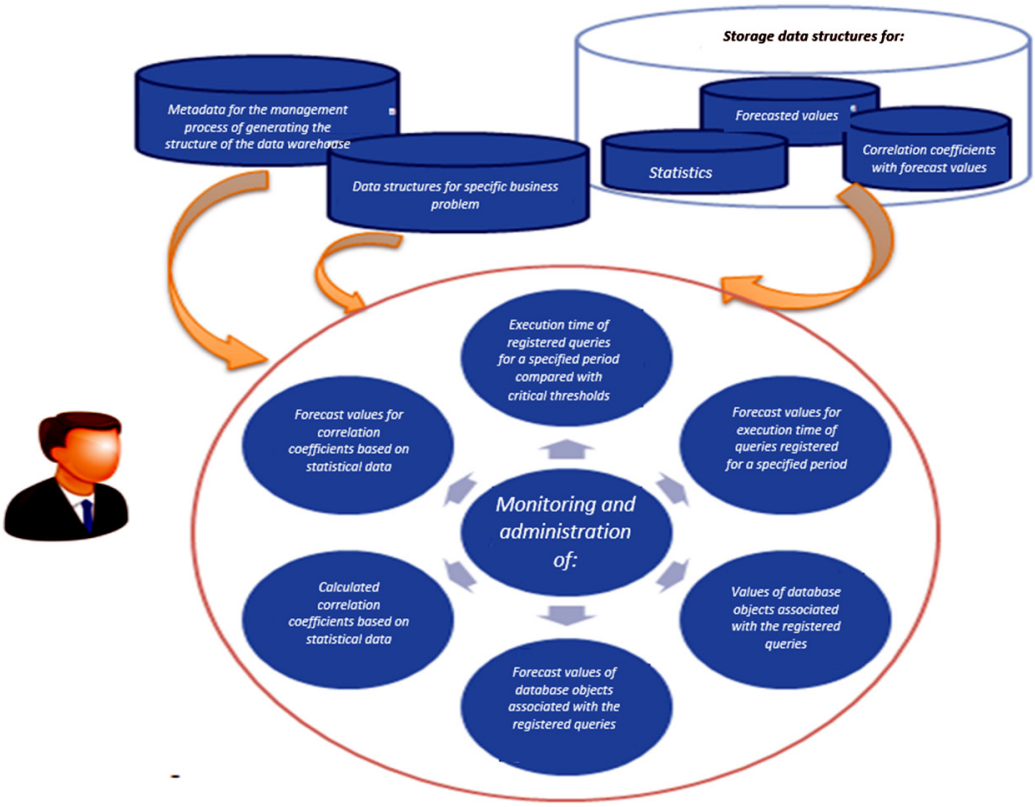

Fig. 6. Sequence of steps in Method 3

methods into an empty virtual database. This means that this database will contain no data, just the statistics, in order to allow the query optimizers to propose accurate solution for optimization of the already forecasted values for volumes of data in the different objects in the database. (ManageEngine, 2012) (DellSoftware, 2013) (SolarWinds, 2013) In this way the administrators will be able to look one step ahead in the development of the database volumes, which they are responsible for.

Based on all previous methods, this method provides time for reaction and flexibility in decisions.

The sequence of steps involved in this method is shown in figure 7. These are two basic steps that are an extension of the previous three methods. The intervention of database administrators is basically on the first step in implementing the current method, as in the second step, the generation of an execution plan is done by the relational database management system. The database administrator only needs to start the process of generation. The relational database management system provides optimizers that offer plan for optimization. The only thing that has to be done is to show the optimizer the forecasted statistics.

- Load forecast statistics in database;
- Generating Execution plan based on forecast statistics.

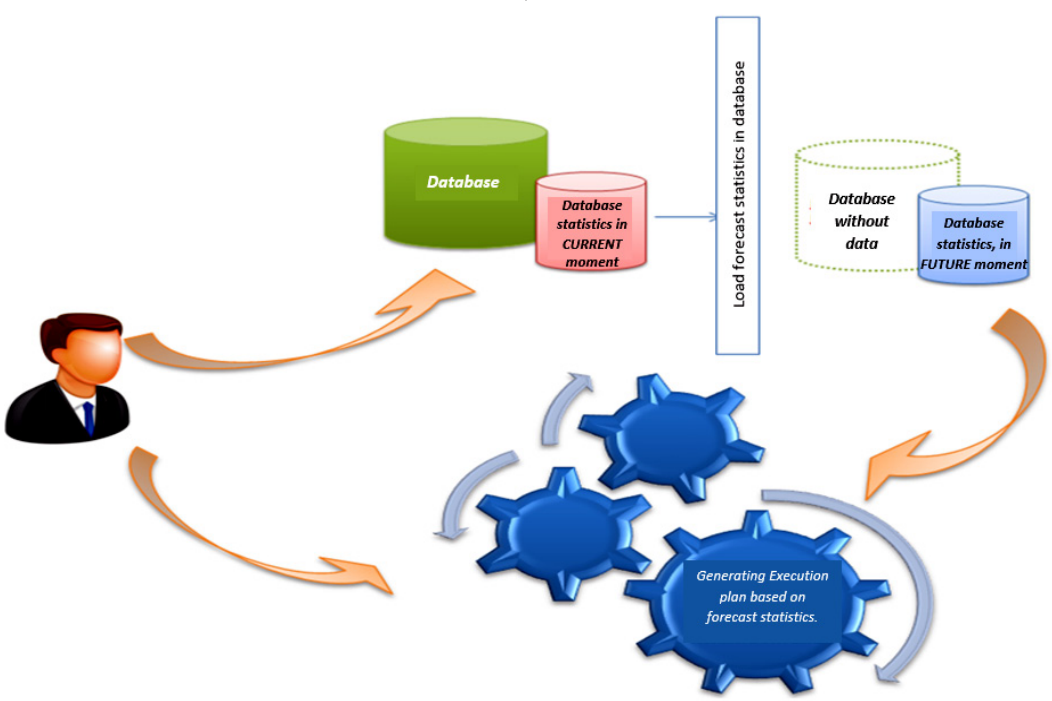

Fig. 7. Sequence of steps in Method 4

Conclusion

In conclusion, this paper presents a concept for monitoring the future state of a database. In the beginning the main problems with business applications are outlined. Then monitoring and database monitoring are reviewed. The architecture of a business intelligent system for monitoring the current and the future state of a database, and methods for dynamic management of database performance are offered. The implementation of the proposed methods and architecture would have a positive effect on a business application and a database, respectively will yield a positive economic effect. There are objective opportunities for the realization of the proposed methods and architecture.

It is important to have a tool for database performance management in order to identify potential risk areas for improvement. The author has developed a prototype system for database performance management based on the proposed architecture. The author intends to make a number of improvements for further development of the proposed prototype.

The proposed solution gives the business enough time for reaction to an occurring problem with the application's performance and respectively the database's performance.

References

Bartle, Phil. 2011. The nature of monitoring and evaluation. *http://cec.vcn.bc.ca.* 2011. [Online] Available at: http://cec.vcn.bc.ca/cmp/modules/mon-wht.htm [Accessed: June 1, 2015].

Brackett, M. 1999. Business Intelligence Value Chain. [Online] Available at: http://www.dmreview.com/editorial/ [Accessed: June 1, 2015].

DellSoftware. 2013. Performance monitoring solutions | Dell Software. *http://www.quest.com.* [Online] Available at: http://www.quest.com/performance-monitoring/ [Accessed: June 1, 2015].

Eckerson, W., Howson, C. 2005. Enterprise Business Intelligence: Strategies and Technologies for Deploying BI on an Enterprise Scale (Report Excerpt). [Online] Available at: http://www.tdwi.org/Publications/WhatWorks/ [Accessed: June 1, 2015].

Lane, Adrian. 2013. Database activity monitoring keeps watch over your data. *http://searchsecurity.techtarget.com.* [Online] Available at: http://searchsecurity.techtarget.com/article/Database-activity-monitoring-keeps-watch-over-your-data[Accessed: May 1, 2015].

ManageEngine. 2012. Compare Database Monitoring Tools for SQL Server Monitoring. *http://www.manageengine.com.* [Online] Available at: http://www.manageengine.com/products/applications_manager/sql/compare-database-tools.html [Accessed: June 10, 2015].

Microsoft. 2015. SQL Server, SQL Statistics Object. *http://msdn.microsoft.com.* [Online] Available at: http://msdn.microsoft.com/en-us/library/ms190911.aspx [Accessed: June 10, 2015].

Murjeva A, Mihova, V. 2011. *Monitoring and analysis the future state of business applications' performance.* Sofia : Anniversary conference "Statistics, Information technologies and Communications" 24-25, 11, 2011.

Murjeva A, Mihova, V.. 2010. *Business intelligent systems for optimization and management of application systems.* Sofia : s.n., 2010. International conference "Communications, information technologies and statistics. Current issues of theory and practice".

Oracle a. 2015. Gathering Optimizer Statistics. *http://docs.oracle.com/.* [Online] Available at: http://docs.oracle.com/cd/B10501_01/server.920/a96533/stats.htm [Accessed: June 1, 2015].

Oracle b. 2014. Oracle JD Edwards World Applications. *http://www.oracle.com.* [Online] Available at: http://www.oracle.com/us/products/applications/jd-edwards-world/overview/index.html [Accessed: June 1, 2015].

Oracle c. 2013. Oracle E-Business Suite. *http://www.oracle.com.* [Online] Available at: http://www.oracle.com/us/products/applications/ebusiness/overview/index.html [Accessed: May 10, 2015].

Proctor, Paul E., Litan, Avivah and Nicolett, Mark. 2009. Pattern Discovery With Security Monitoring and Fraud Detection Technologies. [Online] Available at: http://www.gartner.com/DisplayDocument?id=1160514 [Accessed: June 1, 2015.].

Quarles van Ufford, Deborah. 2002. Business Intelligence: The Umbrella Term, VU Amsterdam. [Online] Available at: http://www.few.vu.nl/stagebureau/ [Accessed: June 1, 2015].

SolarWinds. 2013. Performance Monitoring & Management Solutions. *http://www.solarwinds.com.* [Online] Available at: http://www.solarwinds.com/solutions/sql-database-monitor.aspx [Accessed: June 1, 2015].

Yuhanna, Noel, Penn, Jonathan and Smillie, Katie. 2007. Forrester Research:Research: The Forrester Wave™: Enterprise Database Auditing And Real-Time Protection, Q4 2007. [Online] Available at: http://www.forrester.com/The+Forrester+Wave+Enterprise+Database+Auditing+And+RealTime+Protection+Q4+2007/fulltext/-/E-RES41970?objectid=RES41970 [Accessed: June 1, 2015].