SIMULATION STUDIES OF THE IMPLEMENTATION OF CENTRALIZED TWO-PHASE LOCKING IN DDBMS

Alina Culciar and Svetlana Vasileva
College - Dobrich
University of Shumen
9302, Dobrich, Bulgaria
E-mail: taistaif@abv.bg, svetlanaeli@abv.bg

KEYWORDS
Simulation models, concurrency control, distributed transactions, 2PL, distributed database.

ABSTRACT

One of the most important problems in distributed database systems is the concurrency control. This paper considers algorithms simulating the implementation of centralized two-phase locking (2PL) in distributed database systems and simulation results. It describes specifically the simulations of two-version 2PL and 2PL with integrated timestamp ordering mechanism. In concurrency control method 2PL may take place deadlocks of the transactions. Therefore, in the modeling algorithms described here are integrated algorithms for deadlock avoiding: two-version architecture of database and timestamp ordering strategy “wait-die”. There are also presented, the results of the simulations of these two variants of the 2PL method at different scales of the networks for the transmission of data and at different intensities of inflow transactions. Modeling algorithms are developed by means of the system for simulation modeling GPSS World Personal Version.

INTRODUCTION

Many of today's information systems and management systems have a distributed structure, because the distributed computer systems meet the requirements of the relevant application areas much better than centralized ones. The modular and open structures increase the opportunities for expansion and modification of these systems, and their flexibility. (Connolly and Begg 2002).

All this is especially true for the distributed database (DDB) systems and vital component to them - Distributed Database Management System (DDBMS). In DDBMS as a result of fragmentation and replication of data is increasing the reliability and throughput of the system. A major problem in these systems is concurrency control to general data parallel executed transactions. (Bernstein and Goodman 1981; Carey and Livny 1988; etc) Concurrency control techniques are generally divided into: Locking, Timestamp ordering and Optimistic strategies - Validation check up. (Srinisava et al 2001) In the last two methods was obtained a better retention of transactions in the system when it is saturated with conflicts (due the frequent rollbacks of transactions). Therefore, it is desirable to use the method of Two-phase locking (two-phase locking - 2PL). But in servicing the transactions method 2PL may arise deadlocks of transactions. In DDBMS this problem is even more complicated to solve, because the transactions work at different sites in the distributed system. There are two basic methods for avoiding of transactions of deadlocks: using multiversion (two-version) architecture of data in database (Chardin 2005) and embed timestamping algorithm, working in one of two strategies: "wait-die" and "wound-wait" (Connolly and Begg 2002; Garcia-Molina et al. 2001; etc). The problems in the application of 2PL method in DDBMS and avoiding transactions of deadlocks require testing. Methods 2PL in DDBMS are four varieties: Centralized 2PL, Primary Copy 2PL, Distributed 2PL and Voting 2PL. In the paper we consider algorithms, modeling first of these protocols – the protocol for Centralized 2PL (Bernstein and Goodman 1981; Connolly and Begg 2002; Garcia-Molina et al. 2001). For the development environment modeling algorithms Centralized Two-version 2PL (2V2PL) and 2PL with integrated Timestamping method (strategy "wait-die") was chosen general purpose system simulation - GPSS World.

“System GPSS World is … a comprehensive modeling tool, covering the fields of both discrete and continuous computer modeling, possessing a high level of interactivity and visual representation of information. Using GPSS World allows you to assess the impact of engineering solutions in extremely complex systems of real world” (GPSS World). The above, the studied literature on simulation (Tomashevski and Zhdanova 2003) and many other unnamed here, convince us in the benefits of the environment GPSS World for simulation systems for processing transactions, such as database management systems (DBMS). In front of the authors is standing the problem of studying the transaction concurrency control in distributed DBMS, in which the data are subjected to fragmentation and replication on the nodes of the system.

The review of published modeling studies (Agrawal et al. 1987; Al-Jumah, 2000; Madria et al. 2003;

Proceedings 29th European Conference on Modelling and Simulation ©ECMS Valeri M. Mladenov, Petia Georgieva, Grisha Spasov, Galidiya Petrova (Editors)
ISBN: 978-0-9932440-0-1 / ISBN: 978-0-9932440-1-8 (CD)

Srinivasa et al. 2001; etc) showed that most of them are parametric - do not describe detailed transaction concurrency control algorithms. Therefore, it is necessary to synthesize detailed imitation models of Two-phase locking of distributed transactions in DDB systems.

DISTRIBUTED DATABASE AND DISTRIBUTED DBMS MODEL

Consider distributed database (DDB) system for the management of which the scheduler works by the method of two-phase locking. In particular, we consider distributed DBMS with Central Lock manager (LM). Processing the transactions in such systems is discussed in detail in (Garcia-Molina et al. 2001; Connolly and Begg 2002 etc). In the only node, which is the LM, is stored all the information about the locks of data elements in the system only Lock table.

When starting the global transaction from site S_i:

I. The transaction coordinator node TC_i divides it into several subtransactions, using information from the global system catalogue. If the transaction updates replicated data element, TC_i must ensure the renewal of all existing replicas of this data element. Therefore TC_i must require an exclusive (exclusively) locks on all copies of the updates data element, and then release those locks. To read updated data element TC_i it can select any of the existing replicas. Usually the transaction reads the local replica, if it exists.

II. Central Lock Manager verifies the eligibility of incoming requests to block data elements, taking into account the current state of the lock of these elements:

If the lock is admissible, the Lock Manager sends to the site-initiator a message that states that the request for the element block is granted; otherwise the application shall be placed in a queue and is there until the requested lock is not provided.

Fig. 1 shows a structural model of such distributed DBMS (DDBMS). Modules involved in the processing of the transactions, are depicted by circles, and the necessary data - with rectangles. For clarity in the model site S_0, where is located the only Lock manager LM_0 and the only Lock table LT_0, are not shown the remaining components: Transaction manager TM_0, Data manager DM_0, Transaction coordinator TC_0, Local database LDB_0 and the service for data transferring. The remaining indications of the scheme of Figure 1 are:

S_i – Site - source of transactions;

TC_i – Transaction coordinator in S_i (has included a component for transmitting data);

SC_i – System catalogue.

The transaction manager keeps a record of the transactions, decides to abort and commits to the transactions and transmits query locking to the lock manager. The Lock manager receives queries about interlock, checks in the Lock table compatibility and transmits operations read/write of the data manager, respectively returns the results in TM_i. Data manager asynchronously performs the operations derived from the Lock manager.

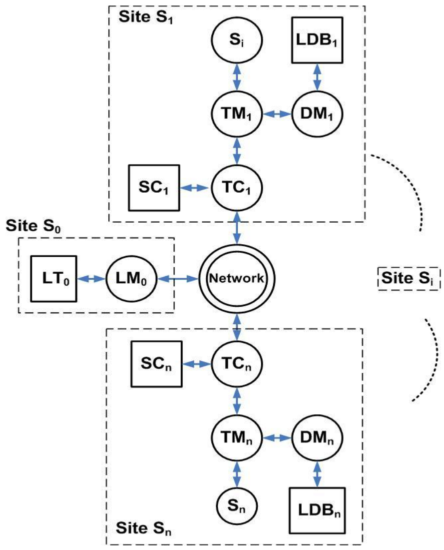


Figure 1: Model of the DDBMS structure at Centralized 2PL

GPSS MODEL OF THE QUEUING SYSTEM

In the construction of probabilistic models of computer systems of different types and especially information systems for public use, the processing of user requests is described as service requests in a closed queuing system (QS). The number of service phases - the number of QS, forming the network, the probability of transmission of requests from one phase to another, the number of parallel channels for service workers in phases, average times and laws of distribution of service requests phases depend on the configuration of the test system (Bernstein et al. 1987; etc). Furthermore, the structure of the model and the topology model depends a lot on the algorithm processing the requests of users in the system. For the network of QS in the model it is assumed that:

- The flow of the receipt of global transactions in sites DDBMS is Poisson using parameter λ;
- The duration of service transactions has exponential distribution with parameter m;
- Serving devices are arranged alternately in parallel channels;
- Discipline of service queues is „First Input – First Serviced";
- Incoming transaction does not leave the system until it is serviced.

In our case transactions go to the sites on an exponential law. They are sent on parallel channels, recorded in certain incoming and outgoing buffers reach sites with the necessary data and transmitted to the recipient (leave the system).

GPSS MODEL OF CENTRALIZED 2PL OF DISTRIBUTED TRANSACTIONS

The main component in Database management system, carrying out the transaction concurrency control is the Scheduler. When working on 2PL protocol, is called the Lock manager. When DDBMS works on Centralized 2PL protocol, the system has only one Lock manager – LM_0 and only one Lock table – LT_0, located on one of the sites of the distributed system. (Connolly and Begg 2002; Garcia-Molina et al. 2001; etc)

In different sites S_1 – S_k of the distributed system enter transactions, which are managed by the relevant transaction managers TM_1 – TM_k. We consider the work of DDBMS on centralized 2PL protocol, therefore all queries about locks (or release locks) are sent as operations transactions to the only lock manager LM_0. Indications for incoming operations: $r_n(x)$ – transaction with number n (T_n) wants to read element x; $w_n(x)$ - T_n recording the value of the element x; c_n - T_n successfully completes its operation and fixes; a_n – T_n interrupts his performance and restarts.

The Two-version 2PL protocol is described in details in (Chardin 2005), and the modeling algorithm of Centralized 2V2PL is presented in (Milev and Vasileva 2009). Fig. 2 represents a structural scheme of the execution simulating algorithm of global (distributed) transaction, initiated from site S_{P2} and refreshing the data elements *El1* having copies in local databases LDB_{P6} and LDB_{P7} and the data element *El2* with copies in local databases LDB_{P8} and LDB_{P9} located in corresponding sites respectively S_{P6} and S_{P7}, and S_{P8} and S_{P9}.

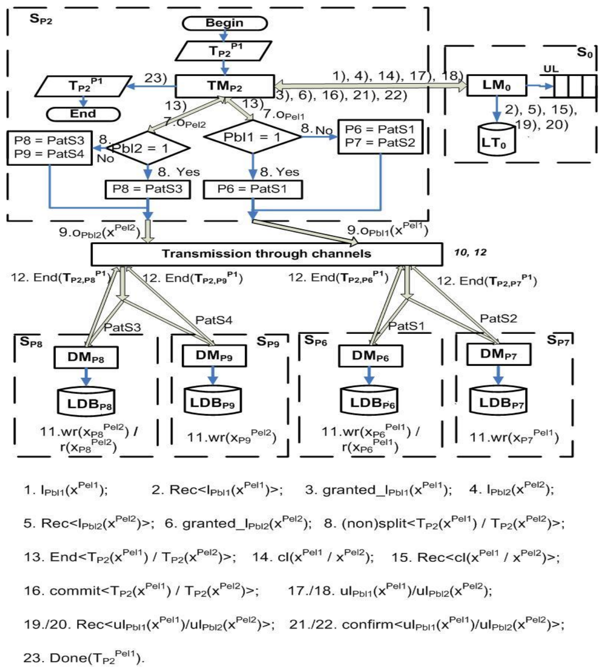


Figure 2: Structural scheme of the algorithm modeling the execution centralized 2V2PL of transactions

The algorithm modeling Centralized 2PL with an integrated mechanism of timestamping (strategy "wait-die") is described in (Vasileva 2011).

The main elements of a simulation model of Centralized 2PL

Based on the analysis of 2PL protocols for modeling DDBMS 2PL algorithms have to use the following initial parameters:

- Cash flows (number of sites in the distributed system) generated transactions;
- Average intensity of each flow generated transactions;
- Law of distribution of each input stream;
- Number of data elements in the modeled DDB (variable *NumEl*), which will compete in parallel performed transactions;
- Number of elements to which transactions will lock states (length transaction);
- Number of replicas of data element - in the presented models each data element has two replicas;
- Functions, modeling the distribution of copies on sites modeled DDBMS.

Parameters of the generated transaction in simulation models

To the generated GPSS transactions are given values of the following parameters:

P1 – Number of the transaction. The value is the sum of System Numeric Attribute (SNA) *MP2* (the difference between the relative time model and the contents of the second parameter of the GPSS transaction) and the number of the site;

P2 – Number of the site, which is generated by the transaction. The value is an integer from 1 to <the number of streams transactions>. In the presented models, <cash flows transactions> = 6;

P$Nel – Length of the modeling of GPSS distributed transaction. Value the synthesized models is 1 or 2, selected with probabilities given by the function *FN$BrEl*, respectively 0.30 and 0.70. The model assumes that the "long" transactions enter the system more often than the shorter;

P$El1 / P$El2 – Number of the first / second element that generates a transaction will "read" or "updated". The value is chosen at random number uniformly distributed in the range [1, NumEl];

P$Bl1 - Type of lock applied to the first element, processing of the generated transaction. The value is taken from one of the numbers: 1 (Reading), 2 (update) or 3 (for recording), selected with probabilities given by the function FN$LockTip1 (Originally are set 0.40, 0.30 and 0.30 respectively);

P$Bl2 – Type of lock applied to the second element. The value is taken from one of the numbers: 1 (reading), 2 (update) or 3 (for recording), selected with probabilities given by the function FN$LockTip2 (Originally are set 0.10, 0.30 and 0.60, respectively);

P$CHTN1, *P$CHTS1* – If P$Bl1 = 1, the transaction will only „read” the element with number P$El1, if the element is not available, but the lock is allowed, in these two parameters are recorded, the number of previous transaction, blocked the element and the numbers generated by the site (by LT);
P$CHTN2, P$CHTS2 – If P$Bl2 = 1, the transaction will only „read” the element with the number P$El2, if the element is not available, but the lock is allowed, in these two parameters are recorded the number of previous transaction, blocked the element and the numbers generated by the site;
P5 – Value 0, if the transaction is in the first phase – seizure of locks and a value of 1 if the transaction completes its work and must release the locks. In simulation of longer transactions, where it can take place deadlocks on transactions can be assigned a value of 1 and P5 of transaction, which would, needs to reboot;
P6 / *P7* – It records the number of the site where the first / second copy of the first data element (*P$El1*) is processed by the transaction;
P8 / P9 - It records the number of the site where the first / second copy of the second data element (*P$El2*) is processed by the transaction;
In the synthesis of models with a greater number of copies of the data element can be used more parameters of the transactions (for example P10 and P11 for recording numbers on the sites where the third copies respectively to the first and the second element).
P$vr – This parameter is used in the model of 2PL with timestamping - *P$vr* = 1 in the case that the lock of the element is not possible and the number of the GPSS subtransaction is smaller than the number of the transaction, which set the lock, then, according to timestamping mechanism strategy "wait-die" (Connolly and Begg 2002) if subtransaction does not continue and does not restart, it waits for the release of the elements in the user chain, whose number is the number of elements.

Auxiliary variables and matrices, used in modeling algorithms

Besides the parameters of transactions in 2PL modeling algorithms shall also apply the following options of GPSS World:
Matrices. In the synthesized simulation models mainly is used the matrix *MX$LTA* for modeling the lock table. Each row of the table corresponds to the data element from DDB. The matrix has the following columns:
- Type of the resulting lock the GPSS transaction blocked free data elements. This column, records the value parameter *P$Bl1* or *P$Bl2* of the same (depending on whether the element is the first or the second processed by transaction);
- Number the GPSS transaction, blocked free element. This column records the value the *P1* parameter transaction, which placed the lock in the first column;
- Number of the site - initiator the GPSS transaction blocked first element (the value of parameter P2 of the transaction, borrowed the element in the lock table);
- Number of the GPSS transaction, received a shared lock element, blocked by another "reading" transaction before. This column records the value the P1 parameter of the transaction, whose request for "reading" is compatible with the shared lock set by the first column;
- Number of the GPSS transaction, received an interlock "record" elements blocked by "reading" transaction before. This column records the value the P1 parameter of the transaction, whose request for "record" is compatible with the shared lock, set by the first column 2V2PL modeling algorithm.
In the synthesized simulation models are used also and two matrices - *MX$RAZST* and *MX$RAZDEV* to set the mean and standard deviation of the retention time of the transactions in the transmission of messages between the nodes of the distributed database system modeled for communication costs.
Variables. In the simulation models primarily are used the following variables:
V$ElemN1 – in it are calculated random selection numbers of the first element that will process the transaction;
V$ElemN2 – in it are calculated the random selection of the numbers of the second element which will process the transaction. In the statements which calculates variable ElemN1 and the ElemN1 is involved and the random number generator RN2;
V$RAZRBL1 – in it is calculated the admissibility of the first lock element processed by the transaction. The value is the product of the parameter *P$Bl1* of the GPSS transaction requesting the lock and the value in the first column of the matrix LTA;
V$RAZRBL2 – in it is estimated eligibility of the second element lock, processed by the current transaction. The value is the product of the parameter *P$Bl2* of the transaction requesting the lock and value in the first column of the LTA matrix;
Cells. In the synthesized simulation models are used the following cells serving as counters:
X$BROITR – total number of generated GPSS transactions during the modeling;
X$ZAVTR – total number of transactions leaving the model served;
X$BROITR1 – total number of generated transactions, processing one data element;
X$BROITR2 – total number of generated GPSS transactions, processing two data elements;
X$ZAVTR1 – total number of transactions with length 1 leaving the modeled served system;
X$ZAVTR2 – total number of transactions with length 2 leaving the modelled served system;

SIMULATION RESULTS

The studies over the simulation models, are enabling in the GPSS World environment to be collected, summarized and analyzed the results of the modeling algorithm of Centralized 2V2PL and modeling algorithm of Centralized 2PL with an integrated mechanism of timestamping in DDB. Calculate the main characteristics of service transactions DDBMS: Throughput (TP), Service Probability (SP) and Response time (RT). Throughput of one system is calculated in the number of requests serviced per unit time (Tomashevski and Zhdanova 2004). This indicator is calculated by the formula (1):

$$X = \frac{N_c}{T_n} \qquad (1)$$

Where N_C is the total number of the committed transactions (in the synthesized models it is the value of the cell *X$ZAVTR*), for the monitoring period of the system - T_n (time modeling at different startups of the modeling algorithm).
Time modeling developed in the GPSS World algorithms is set in milliseconds. All streams transactions are received upon an exponential law with a variable at different studies with an average length of the interval. In all modeling algorithms we consider 6 streams generated by GPSS transactions modeling 6 sites in distributed database system, from which Poisson law shall go global transactions.
The diagram of fig. 3 presents the results of simulations of Throughput Centralized 2PL algorithms at the same intensities of input flows depending on the monitoring period (in seconds): The graph marked with a thin blue dashed line (2V2PL) and the graph indicated by a thick black line and square markers (2PL TSwd) – 6 streams, each with an average intensity 4,17 tr/s (minimum load - intensity cumulative flow 25 tr/s); The graph marked with thin black line and asterisks (2V2PL) and in the graph illustrated by dashed lines with triangular markers (2PL TSwd) - 6 streams, each with an intensity of 8,33 tr/s (average load - intensity of the aggregate stream 50 tr/s); The graph indicated by the thin dotted line (2V2PL) and the graph indicated by a thin blue line (2PL TSwd) - 6 streams of medium intensity 16,67 tr/s (maximum load - intensity of aggregate stream 100 tr/s).
The diagram in fig. 3 shows Throughput for both models at three operating modes of load increased with the increase of the model during and after a certain time (5 minutes after the operational start of the system); the occurrence of stationary mode reaches a maximum value and does not change with time. The diagram also is similare with the Throughput template chart in (TPC Counsil 2010).

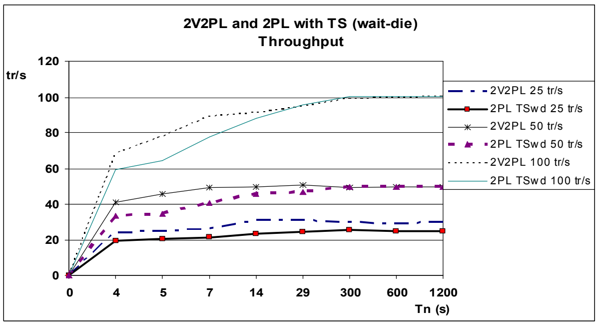


Figure 3: Throughput of the systems

Service probability factor or completeness of service transactions serves to assess the dynamic properties of DDBMS. The probability of service P_s of distributed transactions is calculated by the formula (2) (Tomashevski and Zhdanova 2004):

$$P_s = \frac{N_c}{N_g} \qquad (2)$$

where N_C is the total number of fixed transactions (cell value of *X$ZAVTR* after modeling, and N_g is the total number of transactions generated for the same period of time (cell value of *X$BROITR* after modeling).
Fig. 4 presents the results for the probability of service of distributed transactions at simulation algorithm Centralized 2V2PL and Centralized 2PL with an integrated mechanism of timestamping at the same intensities of inflows (as for Figure 3).

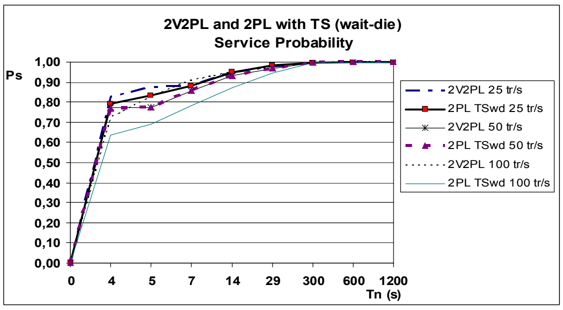


Figure 4: Service probabilities in the 2V2PL model and the 2PL with TS “wait-die” model

Table 1 presents the results obtained for the average Response time for modeling of Centralized 2V2PL and Centralized 2PL with TS "wait-die". The data are taken automatically from the generated reports by GPSS World after the end of each simulation. In the codes of the models is set a table DaTable, which records the "residence time" (SNA of GPSS - *M1*) model of each GPSS transaction. Furthermore, the Response Time can be fixed and from the table of queues in the accounts of the implementation of modeling algorithms. In the report of the queues (realized by GPSS blocks QUEUE <name of a queue> - entry in the respective queue and

DEPART <name of a queue> - exit from the respective queue), for each queue GPSS World model displays the following statistics: Maximum length of the queue; Average length of the queue; Number of entries in the queue; Average time of residence in the queue of all transactions; etc.

Table 1: The Average Response Time in the models of Centralized 2V2PL and Centralized 2PL with built timestamping (in seconds)

Tn	Cntr2	Cntr3	Cntr4	Cntr5	Cntr6	Cntr7
s	100 t/s	50 tr/s	25 tr/s	100 t/s	50 tr/s	25 tr/s
3,6	0,747	0,700	0,671	0,996	0,830	0,773
4,5	0,745	0,695	0,681	1,047	0,854	0,763
7,2	0,759	0,701	0,682	1,178	0,876	0,793
14,4	0,769	0,722	0,690	1,417	0,912	0,789
28,8	0,813	0,721	0,687	1,625	0,913	0,774
300	0,860	0,728	0,701	1,655	0,929	0,787
600	0,867	0,730	0,704	1,679	0,928	0,786
1200	0,872	0,731	0,707	1,674	0,930	0,787

In all models, after the generation of each transaction, are assigned values of its parameters (how many data elements will process; the numbers of the elements that will process; will read or updated elements;, the numbers of sites where are the copies of the elements that will process; etc.). After setting the values of "common" for each transaction parameter, the transaction enters into the queue TOTALTIM. Once the transaction is fixed, exits the TOTALTIM queue (in which the transaction "submit" the information about itself in the automatically collected statistic by GPSS) before leaving the model.
The data in Table. 1 is obtained by the following intensities:
- in Column 2 (2V2PL) and Column 5 (2PL TSwd) - each of the six flows has an average interval of receiving the transactions 60 model units. The intensity of the aggregate flow is the highest of those surveyed - 100 tr/s (in the general case, the transaction ends her performance, long after the generation of the next transaction in the stream);
- in Column 3 (2V2PL) and Column 6 (2PL TSwd) – each of the six fluxes has a mean interval of placing the model units 120 transactions. The intensity of the aggregate flow is the average value in the surveyed - 50 tr/s (in each transaction, the flows have the opportunity to leave the served, shortly after the generation of the next transaction).
- in Column 4 (2V2PL) and Column 7 (2PL TSwd) – each of the six fluxes has a mean interval of placing the model units 240 transactions. The intensity of the aggregate flow is the lowest of those surveyed - 25 tr/s (in each transaction, the flows have the opportunity to leave the served, before the generation of the next transaction).

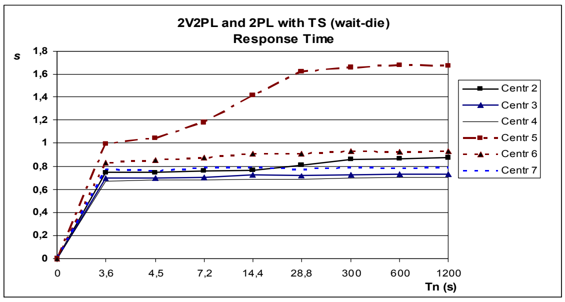


Figure 5: Response time of transactions in model systems at different intensities of inflows

In the diagram of Figure 5 are shown in graphical form the data from Table 1. It can be seen that the graphics of RT measurements have the same kind - a rapid increase in the beginning, slow growth and stationary mode.
From the diagram of fig. 5 it appears that the greatest is the response time in the model of Centralized 2PL with timestamping "wait-die" at the intensity of the incoming stream 100 tr/s. This can be explained by the specificity of the protocol. If the transaction, which wants locking the element is older than the occupied by the element transaction (has a smaller number), it restarts with a reservation number. So, the next race for locking elements will be with the smallest number and will receive the locking of the elements, to accomplish after the read / write operation. This way appear the delays by the lock manager to the site-initiator and after that back to the lock manager, this increases the time of service of the transaction.
On fig. 6 graphically are given the values that are obtained for throughput by Formula (1) by substituting the fixed in the receiving reports values of the cell *X$ZAVTR*. Intensity of inflows transactions are the same as the graphs of fig. 3, was changed only the matrix of distances *MX$RAZST* - the values in it are increased twice compared to models whose results are reported in the graphs of fig. 3.

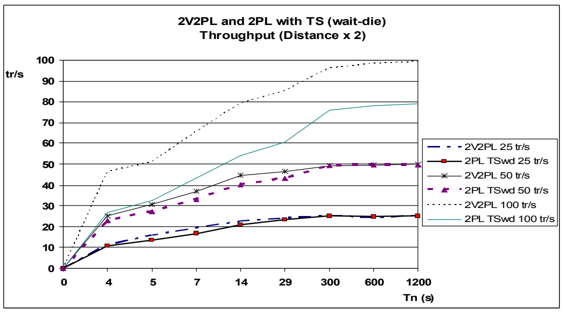


Figure 6: Throughput in the models at doubled distance between the sites in the system

From the graphs of fig. 6 it can be concluded that with the increase of the distance between sites in the system, the throughput graphs are "spaced apart" more. This is very evident in the graphs at maximum load of the system (intensity of inflow 100 tr/s).

By varying the values of cells in the matrix *MX$RAZST* we can conduct research whether and how the gaps in the net, impact on Throughput, Service probability and other performance indicators of concurrency control algorithms.

For example, by five times, increasing the cell values of the matrix *MX$RAZST* are obtained the values of the throughput graphs on fig. 7 from the simulation reports of the Centralized 2V2PL protocol and the Centralized 2PL with timestamping "wait-die" at the same intensities of inflows as Figure 3.

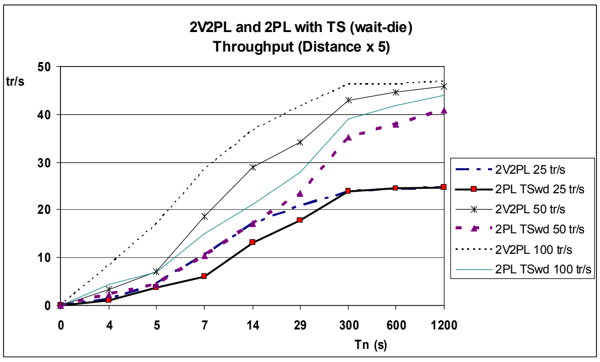


Figure 7: Throughput of the system (distance x 5)

By increasing the distance between sites in the system, except the increasing divergence of the graphs throughput, is observed slower occurrence of stationary mode, and also "instability" of the values in the transition mode. This is shown in the graph for the Centralized 2V2PL maximum intensity of the inflow transactions - graph, marked with dotted lines in fig. 8.

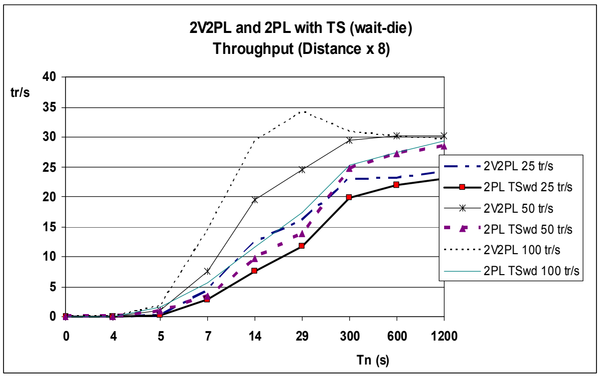


Figure 8: Throughput of the system (distance x 8)

Another indicator of the performance of concurrency control algorithms is Frequency distribution of Response time of transactions. Diagrams of Frequency distribution of RT are built automatically by the formulated in the GPSS model tables. On fig. 9 is demonstrated the histogram of Frequency distribution of RT in modeling Centralized 2V2PL at the total intensity of the input streams 100 tr/s (maximum load on the system) and observation time 28.8 seconds. A histogram is generated automatically by GPSS World on table Response time distribution DATABLE, whose values are brought into the simulation report.

On fig. 10 is demonstrated the histogram of Frequency distribution of RT in modeling Centralized 2PL with timestamping (wait-die) at total intensity of the input streams 100 tr/s (maximum load on the system) and observation time 28,8 seconds (as in Figure 9).

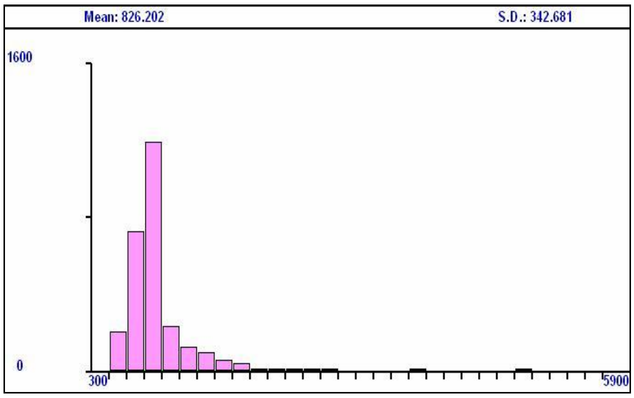


Figure 9: Frequency Distribution of RT in modeling Centralized 2V2PL at λ = 100 tr/s

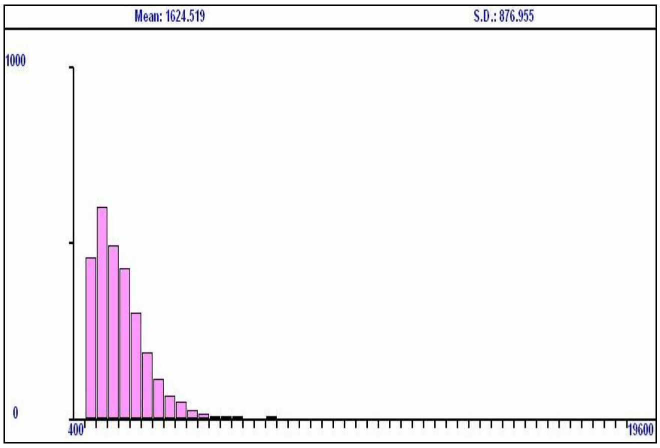


Figure 10: Frequency Distribution of RT in modeling Centralized 2PL with TS (wait-die) at λ = 100 tr/s

The tables of Frequency distribution of RT besides that serve comparative analysis of concurrency control algorithms, serve also to assess the reliability of modeling algorithms by comparing with the template chart of Frequency distribution of RT (TPC Council 2010).

Similarly, can be compared with the template graphics and charts for throughput of fig. 3, fig. 5 and fig. 7.

CONCLUSION

The system of simulation GPSS World permits creation of effective simulation models of transaction concurrency control (in particular models of Two-phase locking of transactions in DDBMS with one-version and two-version architecture of replicated data.

The synthesized simulation models of GPSS World for Centralized 2V2PL and for Centralized 2PL with built-in mechanisms of timestamping (strategy "wait-die") allow experimentation and accumulation of statistical data to conduct comparative analysis of algorithms.

The proposed simulation models of Centralized Two-version 2PL and Centralized 2PL with timestamping in the DDB are convergent and give good results in the experimental research done.

The comparative analysis of the results of modeling of one-version and two-version Centralized 2PL showed no deadlocks, while 2V2PL throughput is high and the response time is significantly lower than that of the one-version 2PL algorithm.

REFERENCES

Agrawal, R., M. Carrey, M. Livny. 1987. "Concurrency Control Performance Modeling: Alternatives and Implications". *ACM Transactions on Database Systems*, Vol. 12, No. 4, (Dec), 609-654.

Al-Jumah, N., H. Hassanein, M. El-Sharkawi. 2000. "Implementation and modeling of two phase locking concurrency control – a performance study". *Information and Software Technology* 42 (2000), 257-273.

Bernstein, Ph. & N. Goodman. 1981. "Concurrency Control in Distributed Database Systems". *Computing Surveys*, Vol. 13, No. 2, 185-221.

Bernstein, Ph., V. Hadzilacos, N. Goodman. 1987. "Concurrency control and recovery in Database Systems". Addison-Wesley Publishing Company.

Carey, M., M. Livny. 1988. "Distributed Concurrency Control Performance: A Study of Algorithms, Distribution and Replication". In *Proceedings of the 14th VLDB Conference*, Los Angeles, 13-25.

Chardin, P. 2005. "Data multiversion and transaction concurrency control". http://citforum.ru/database/articles/multiversion..

Connolly, T., C. Begg. 2002. "Database Systems". Addison – Wesley.

Garcia-Molina, H., J. Ullman, J. Widom. 2001. "Database Systems: The Complete Book". New Jersey 07458, Prentice Hall.

Madria, S., S. Maneshwari, B. Chandra. 2003. "A Concurrency Control Algorithm for an Open and Safe Nested Transaction Model: Formalization and Correctness". *Journal of Information Science and Engineering*, 19, 283-328.

Milev, A. and S. Vasileva. 2009. "Simulation of Algorithms for 2-Phase Locking in Distributed Databases". In *Proceedings of the International Conference Automatics and Informatics'09* (Sofia, BG, Oct.3–4), I-13 – II-16.

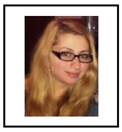
Srinisava, R., C. Williams, P. Reynolds. 2001. "Distributed transaction Processing on an Ordering Network". Technical Report CS-2001-08, (Feb).

Tomashevskii, V. and E. Zhdanova. 2003. "Simulations in GPSS". Bestseller, Moscow.

TPC Council. 2010. "TPC – C Benchmark". Standard Specification Revision 5.11. Transaction Processing Performance Council, (Feb).

Vasileva, S. 2011. "An algorithm for deadlock avoiding in distributed database management systems". In: *Journal of the Technical University Sofia, branch Plovdiv*. "Fundamental Sciences and Applications", Vol. 16(1), International Conference Engineering, Technologies and Systems, Plovdiv, BG (May), 377-382.

AUTHOR BIOGRAPHIES



ALINA CULCIAR was born in Medgidia, Romania and attends the University of Shumen (College - Dobrich), where she is studying Informatics and Information Technologies chemical technology and will obtain her degree in 2015. She is interested in imitation modeling applications in the educational process in disciplines such as Operation systems, Database systems, etc.

Publications: Vasileva, S., A. Culciar. Some approaches to the application of imitation modeling in educating students. // International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE - 2013), December 6 -7th, 2013, UNWE, Sofia, 2013, Bulgaria, pp. 169-178; Vasileva, S. Options of GPSS World for integrated demonstration models in the educational process. // Proceedings of Science and Information Conference 2014, The Science and Information Organization(SAI), ISBN: 978-0-9893193-1-7, 2014, pp. 933-937.

e-mail address is: `taistaif@abv.bg`

SVETLANA Zh. VASILEVA was born in Plodiv, Bulgaria and went to the Saint-Petersburg State Electrotechnical University (LETI), where she studied Engineer Process Systems "Computer Aided Design" and obtained her degree in 1994. She has also a degree as a PhD in Informatics 01/01/12 - acquired on 06.07.2010. She works since 2001 in University of Shumen, College Dobrich as an assistant professor (2001-2010), currently she is a chief assistant professor. Areas of scientific interests: Control systems for distributed databases; Simulation modeling of processes and systems; Application of information technologies in education.

Her e-mail address is : `svetlanaeli@abv.bg`

Webpage:
`http://shubg.net/faculties/pk/prepodavateli?faculty=pk&teacherId=465#biography`