ADAPTIVE REPLICATION ALGORITHM FOR MOBILE COMPUTERS

Karol MATIAŠKO, Michal ZÁBOVSKÝ
Department of Informatics, Faculty of Management Science and Informatics,
University of Žilina, Univerzitná 8215/1, 010 26 Žilina, tel. 041/513 4050, 041/5134 181,
E-mail: karol.matiasko@fri.uniza.sk, michal.zabovsky@fri.uniza.sk

ABSTRACT

The need for mobile access to large databases leads to problems with mobile computers connectivity. Mobile computers often suffer from limited connectivity or lack of network access. Existing replicated databases are not well suited for mobile scenarios as well as algorithms used for data replication. Distributed database systems and distributed database management systems have been developed in response to trend of distributed computation. In distributed models of computations several sites are interconnected via a communications network. The advantages of data distribution involve increased availability, distributed access to data and improved performance of parallel processing. Disadvantages, on the other hand, are given by increased overhead and complexity in the system design and implementation. In the following paper we introduce algorithm for adaptive data replication including mobile computers.

Keywords: *distributed database systems, distributed database management systems, wireless networks, data replication.*

1. INTRODUCTION

The growth of distributed models of computation requires the ability to provide services for both, wire and wireless clients. Mobile, nomadic, computers often suffer from limited connectivity. Even the complete lack of network access can be observed for TCP/IP based protocols [1]. Nomadic computers mostly use a wireless access and, because of limited bandwidth, this type of connection is more expensive than wire communication. Therefore is important to access mobile databases in minimal communication cost. We will show how mobile clients can be involved in the replication schema and propose solutions for problems we observed during our research.

2. DISTRIBUTED DATABASES

The term distributed databases is used collectively for distributed database systems and distributed database management systems. Components of distributed database consist of database portions spread out over multiple *sites – nodes*. The sites are connected by a communication network with a given topology. Local site may have its own local database which behaves as a common database management system. Each site contains *fragments* – parts of global database distributed over the set of nodes. Fragmentation of global database is managed by application and communication processing software.

Advantages of data distribution include local autonomy, performance, reliability, availability and scalability. One of the most important features is availability which refers to the probability that a system is continuously available during some time interval. For centralized system means the failure of a single site failure of the entire system. In distributed database failure of a single site affects only data located at that site. Another advantage is the performance improvement since local access to data is much faster than remote query. The performance depends on *replication schema* which defines data distribution over the sites involved in distributed database system. This is mostly an advantage when the database is distributed geographically over different locations.

Disadvantages of data distribution are given by a significant cost in terms of performance (when replication schema is not optimal), software complexity and administration difficulty. The response time for transactions across different sites must be kept in acceptable range. In the distributed database systems is the design and replica placement a critical task. The objective of *fragmentation* is to achieve necessary level of data distribution. Consequently, optimal placement of fragments must be found over the available sites. This process is called *allocation* of fragments. Allocation must fulfill needs for consistency and efficient data access.

Formal description of a basic distributed database allocation problem is shown in following text. The traditional approach is based on methods of operational research. The criterion is based on *access pattern* defined by operations read and write. The result is decision, how to distribute the fragments over the sites.

Consider the database *D* consists of a collection *S* of *m* sites, where each site *i* has its capacity c_i,

$$S = \{c_1, c_2, \dots, c_i, \dots, c_m\} \tag{1}$$

and a set *F* of *n* fragments, where each fragment *j* is characterized by its size s_j,

$$F = \{s_1, s_2, \dots, s_j, \dots, s_n\} \tag{2}$$

Each fragment is required by at least one of the sites. The requirements for each fragment are indicated by the requirements matrix,

$$R = \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,n} \\ r_{2,1} & r_{22} & \cdots & r_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m,1} & r_{m,2} & \cdots & r_{m,n} \end{bmatrix} \tag{3}$$

where the element of matrix indicates the requirement by site i for fragment j. The value is represented by a real number or by zero-one value that indicates requirement for given fragment by defined site. Communication costs are defined by the transmission cost matrix,

$$T = \begin{bmatrix} t_{1,1} & t_{1,2} & \cdots & t_{1,m} \\ t_{2,1} & t_{2,2} & \cdots & t_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ t_{m,1} & t_{m,2} & \cdots & t_{m,m} \end{bmatrix} \quad (4)$$

where element t indicates the cost for site i to access a fragment located on site j.

Given the above definitions, the distributed database allocation problem is one of finding the optimal placement of the fragments at the sites. Hence we wish to find the placement,

$$P = \{p_1, p_2, \ldots, p_j, \ldots, p_n\} \quad (5)$$

where $p_j = i$ indicates that fragment j is located at site i. By adding the capacity constraint

$$\sum_{j=1}^{n} r_{ij} s_j \le c_i \qquad \forall i \;\; 1 \le i \le m \quad (6)$$

is specified that the capacity of any site is not exceeded ant the total transmission cost

$$\sum_{i=1}^{m} \sum_{j=1}^{n} r_{ij} t_i p_j \quad (7)$$

is minimized [3].

This kind of allocation problem can be transformed to the bin packing problem by adding restrictions for use of requirements matrix and zero transmission costs. The bin packing problem is known to be NP-complete. The DDB allocation problem is even more difficult so it is NP-complete too [3]. Thus we must use heuristic methods to find near optimal solution.

3. DYNAMIC DATA ALLOCATION

By the static fragment allocation are fragments located at the sites from which they are most frequently accessed. Since the distributed database system is rather dynamical, the main problem with previous, static, allocation of fragments is changing workload. This occurs when the access frequencies to various portions of database from a particular site vary with time. Even very simple methods for dynamic data allocations are able to improve system throughput by 30 percent. Experimental evaluation of dynamic data allocation strategies can be found in [4]. To determine when a re-allocation is needed, algorithms proposed in [4] maintain weighted counters of the number of access from each site to each block. For effective estimation the aging factor is used to update counters. The main problem can be divided into the two problems: how to detect changes in workload and how to dynamically re-allocate fragments of database result in improved throughput.

Considering mobile clients, communication network characteristics must be evaluated carefully since they form transmission cost matrix. LAN characteristics for wire clients, such a throughput and latency (or round-trip time), can be estimated reliably by tools such a *tstat* [1] or *pathchar* [5]. For wireless client is estimation more difficult than for wire clients. Technology for wireless communication based on the IEEE 802.11 standards provides connection with variable transmission characteristics. Rather than throughput and latency a signal-to-noise ratio (SNR) is a characteristic that has to be involved in replication model. SNR affects both characteristics throughput and latency in significant way [6] (see Fig. 1).

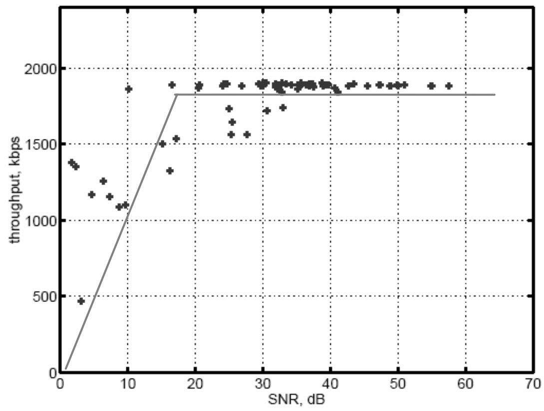

Fig. 1 Throughput dependency on SNR with linear prediction

The linear mathematical model for throughput prediction based on previous observations looks as follow [6]:

$$T = T_{\max} \qquad SNR > SNR_C \quad (8)$$

$$T = A \times (SNR - T_0) \qquad SNR \le SNR_C \quad (9)$$

where T_{max} is a saturation throughput, A defines slope, SNR_0 is a cutoff SNR and SNR_C defines critical threshold. Respective exponential model is also described in [6], for proposed solution linear algorithm is sufficient enough to describe communication network characteristics. Now we can define a set SNR_C of m elements, where each value snr_{ci} represents critical threshold for site i.

$$SNR_C = \{snr_{c1}, snr_{c2}, \ldots, snr_{ci}, \ldots, snr_{cm}\} \quad (10)$$

When the site i is a wire client, critical threshold is zero. Finally, we need to define function which returns current *SNR* value. Such a function is necessary to implement on each site from replication schema.

4. HEURISTIC ALGORITHM FOR ADAPTIVE REPLICATION

Now we can formulate the algorithm for adaptive replication based on simple algorithm described in [4].

The algorithm consists of two tests. The *test of expansion* is executed after the specified number of transactions and is responsible for replication schema expansion when such change improves solution. The *test of expansions* is defined by following steps:

1. The control process examines read counters for each fragment.
2. The site with the highest counter value is marked as a candidate for fragment re-allocation.
3. If the candidate is the site on which fragment is currently located, go to step 6.
4. For the candidate site get *SNR* and *SNRc* values. For wired nodes return *SNR=100* and *SNRc=0*.
5. If *SNR>SNRc* then move fragment from original site to the candidate site. Otherwise, choose site with the highest counter value from the set of unmarked sites and mark it as a candidate for fragment re-allocation – then go to step 3.
6. Wait for specified number of transactions to be completed and then go to step 1.

The first step can be modified by using aging factor to improve efficiency. The *test of contraction* solves the problem with wire nodes included in the replication schema. The motivation is given by assumption that is easier to prevent site failure due to the communication network problem than to solve its failure. Since the test is performed repeatedly, is possible to release a site from the replication schema when requests for replica may cause failure. The *test of contraction* is specified as follows:

1. The control process monitors *SNR* value for each site included in the replication schema.
2. If the site does not response then the site is released from replication schema and its counters are reset.
3. If *SNR≤SNRc* then the site is release from replication schema.
4. Wait specified time and then go to step 1.

The performance of the system under dynamic re-allocation scheme with and without (load sensitive algorithm [4] was used) proposed algorithm was compared during the experimental phase. We used communication network statistics produced by *tstat* tool [1] and a set of tools we created for wireless connection evaluation. Network characteristics were acquired from the real network. Requests for data and replication process were simulated.

The results show that for the *SNR greater than SNRc* our solution is from 1 to 4 percent more expensive than solution that is not using proposed algorithm. The percentage change represents the overhead caused by the algorithm. In the case that *SNR oscillates around the SNRc* is the cost of solution with proposed algorithm for about 20 percent worse than for solution without the adaptive algorithm. Hence the *SNRc* value must be evaluated carefully since the algorithm reacts to possible critical threshold but the communication network parameters are near to regular values. Finally, if the *SNR is less than SNRc* our solution improves overall performance in a significant way. The observed results show about 60 percent better overall response time than for load sensitive algorithm. This is given by the fact that proposed algorithm is avoiding replica transmission to the site with communication problems.

5. CONCLUSION

Performance in distributed database systems is heavily dependent on allocation of data among the sites of the database. The static allocation provides only limited response to workload changes. The situation is even worse when mobile wireless computers are included in replication schema. We presented algorithm for dynamic re-allocation of data with a mobile computers included in replication schema. Proposed algorithm offers significantly increased performance for nomadic nodes with limited connection. Our experiments make a practical case for future development of algorithms for changing environment such as intelligent transportation systems, location aware application and information systems for mobile user.

REFERENCES

[1] Franceschinis, M., Melia, M., Meo, M., Munafo, M.: Measuring TCP over WiFi: A Real Case, http://www.tlc-networks.polito.it/ mellia/papers/winmee.pdf

[2] Huang, Y., Sistla, P., Wolfson, O.: Data Replication for Mobile Computers, ACM 0-89791-639-5/94/0005, 1994

[3] Corcoran, A. L., Hale, J.: A Genetic Algorithm for Fragment Allocation in a Distributed Database System, ACM 089791-647-6/94/0003, 1994

[4] Brunstrom, A., Leutenegger, S. T., Simha, R.: Experimental Evaluation of Dynamic Data Allocation Strategies in a Distributed Database with Changing Workloads, ACM 0-89791-812-6/95/11, 1995

[5] Downey, A.: Using pathchar to estimate Internet link characteristics. - SIGCOMM 1999, Cambridge, MA, 1999, pp. 241-250

[6] Na, C., Chen, J., Rappaport, T. S.: Measured Traffic Statistics and Throughput of IEEE 802.11b Public WLAN Hotspots with Three Different Applications, IEEE Transactions on Wireless Communications, 2006, pp. 3296-3305

[7] Huang, Y-F., Chen, J-H.: Fragment Allocation in Distributed Database Design, Journal of Information Science and Engineering, 2001, pp. 491-506

Received February 6, 2008, accepted February 22, 2008

BIOGRAPHIES

Karol Matiaško was born in 1952. In 1975 he graduated (MSc.) in the study field of Cybernetics at the Faculty of Electrical and Mechanical engineering of the University of Transport and Communications in Žilina. He received PhD. in Technical Cybernetic in 1988 at University of Transport and Communication in Žilina like employee of the Research Institute of Transport. Since 1990 he was joined with the Faculty of Management Science and Informatics at University of Žilina and his research and educational activities have been oriented in the area of database system, distributed processing and data processing.

Michal Zábovský was born on October 26, 1975. In 1999 he graduated (MSc.) at the Faculty of Management Science and Informatics at University of Žilina. He defended his PhD. in the field of applied informatics in 2006; his thesis title was "An Adaptive Database Allocation in Distributed Database System". He is working as a lector with the Department of Informatics since 1999. From 2003 he is visiting lector at University of Applied Sciences in Jyväskylä, Finland. His scientific research is focusing on database systems, distributed systems and data processing.