

INSTRUMENTAL COMPLEX OF PARALLEL SOFTWARE SYSTEM DEVELOPMENT AND OPERATING ENVIRONMENT SUPPORT FOR DISTRIBUTED PROCESSING WITHIN MULTITRANSPUTER SYSTEMS, TRANSSOFT

Boris E. Polyachenko and Filipp I. Andon

Institute of Software Systems,
Ukrainian Academy of Sciences
40, Glushkov prospect
252207, KIEV - 207, UKRAINE
CRIT%SOFTGIC.KIEV.UA@RELAY.USSR.EU.NET

1. System Features

TRANSSOFT presents both complete system to manipulate with Data Bases (DB) and Knowledge Bases (KB) for different level users and also the instrumental tools to develop own information systems and complex applications with efficient multiprocessors (multiprocesses) DB and KB handling in new operational environments, with new hardware tools (new multiprocessors architectures), supporting new possibilities for design and functions extension. TRANSSOFT supports dBASE files structure as a real DB standard. High level of portability is supported, so the considered tools can be easily used for various operating environments. TRANSSOFT ensures the following main functions and features implementation:

- parallel software products synthesis based on sequential modules written by such standard languages as C, Pascal and FORTRAN;
- DB and KB manipulating, access and parallel processing in local and/or distributed multiprocessor systems;
- automatic parallelization of DBs processing;
- supporting Object-Oriented (OO) paradigm;
- implementation the original techniques for the effective parallel DB and KB processing;
- providing source C-functions of complete set to design of information systems and serious applications in any software and hardware environments;
- distributed environment maintenance for application

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

SIGMOD /5/93/Washington, DC, USA

© 1993 ACM 0-89791-592-5/93/0005/0548...\$1.50

programs processing;

- description and initiating the set of sequential programs interacting through shared DBs and KBs with automatic computations parallelization for the multiprocessor (multiprocess) dynamic environments;
- providing the adaptability to multiprocessor computing system architecture and generating the version for concrete hardware and software architecture, operating system;
- increasing the computers performance, reduction the problem processing time and programmers labour expenditures.

2. Object-Oriented Model

We have been using OO paradigm in our system to describe classes of DB objects and their relationships. Class has unique name and maps the set of objects with corresponding sets of attributes and methods or operations on objects. Relationships define interclasses communication. TRANSSOFT gives the possibility to define user's own environment which consists of user's own objects, classes, operations and relationships with corresponding access constraints and possibilities of DB sharing. User's own environment defines DB subset and is locked and used by this user in exclusive mode. TRANSSOFT supports manipulating and processing of DB items of user's own environment in fast (main) computing system memory during the session of user work by means of special buffering and control system of the first level. Once the session is finished the data is transferred to DB store and made available for general usage by other users. Production-based system (KB Processor) is supported with rules description on the level of objects and operations.

3.Objects and Relationships Operations

TRANSSOFT supports the set of functions to manipulate with objects and relationships. This set presented can be extended and/or modified by means of special interactive tools.

4.Parallel Computations

Parallel computations in TRANSSOFT are based on our original techniques, which are used for wide range of problems and provide high (controlled) degree of parallelism. These ones referred as vertical, horizontal and data structure binding paralleling techniques. The vertical paralleling technique (VPT) supports the computations in asynchronous way controlled by DB and KB items accessibility. DB (KB) query is mapped into the query graph with the set of nodes corresponded to the set of DB (KB) processing statements. Each DB (KB) processing statement is the mapping of the set of its input arrays into the set of its output arrays. Communication between statements provides by common arrays. The array notion is logical one and corresponds to real DB (KB). VPT provides parallel processing of depended by DB (KB) statements (query graph) with minimal volume of exchanges with these DB (KB) by means of special control tools. Therefore, VPT provides multilevel hierarchical approach to DB and KB parallel processing, because the results of VPT of i -th level can be used in VPT of $(i+1)$ -th level, etc. Horizontal paralleling technique provides data structure decomposition and its parallel processing in multiprocessor environment by means of equivalent transformation of the DB, KB processing statement into the number of such statements, each of which processes its own substructure of initial structure. These statements are processing in parallel way. Data structure binding supports logical inference on the data structure level.

5.Mathematical Models

For formal DB and KB query representation the algorithm algebra under the memory is used. This algebra defines the set of elementary conditions, the sets of elementary functional and communicating statements. This model is based on OO paradigm supporting objects, their classes, properties and object inheritance, interobject communications and interface with parallel multiprocessor computing system. Discrete dynamic system is defined as the state set and the system of transition rules which are given on this set. By means of such models we have described traditional computation process organization and/or 'rendez-vous' - like principle, as well as asynchronous (vertical and horizontal paralleling) computations, proved their features for different important data processing program classes and examined deadlock conditions.

6.System Architecture

TRANSSOFT has being developed for transputer and/or INTEL/486 based multiprocessor complexes. IBM PC and other computers, their network complexes with or without multiprocessor transputer motherboards INMOS,

TRANSTECH and similar can be used as the host computers for transputer based implementation. TRANSSOFT supports dBASE structure as a real DB standard. High level of portability is supported, and the considered tools can be easily used for various operating environments. TRANSSOFT has hierarchical architecture and consists of the number of logical processors. There are several levels of TRANSSOFT's users and corresponding interfaces. Programmer-user will be supported by C and Assembler source programs and corresponding tools for own DB systems development. Non-programmer-user will use parameterized standard queries and procedures, views, menus and their design possibilities on highest level. SQL - language tools are also supported both in package and interactive modes with tools for automatic query optimization and transformations. Parallel application software products developing by means of TRANSSOFT tools can be provided on the basis of sequential modules on C, Pascal and FORTRAN. No application program modification is required. TRANSSOFT system has an open architecture and the capabilities of functions extensions. Program components distributing and processing are performed with accounting for their requirements on resources and component informational dependencies .

7.Performance Evaluation

7.1.Transputer implementation

Experimental version of TRANSSOFT has been developed for following hardware architecture. Multiprocessor Computing System is based on IBM PC computer with TRANSTECH TMB08 transputer motherboard. Computers are working under MS DOS or UNIX-like operating systems communicated with 3L and HELIOS transputer environments. DB, Optimization and Parallelization Processors are based on our own DB access programs parallelized also by means of vertical and horizontal paralleling techniques. Performance evaluation has been realized in 3L parallel transputer environment for IBM PC/286 (16 MHz) host. dBASE files were stored on 40 Mb Seagate ST 351AX hard disk. Their processing were organized both by means of standard DB tools (t1) and under TRANSSOFT control (t2). The implementation under TRANSSOFT control provides decreasing the execution time of the DB processing problems on 3 transputers (TRANSTECH TTM6-8-F) depending on problem class, if compared with standard dBASE tools.

DB volume, K Bytes	t1, s	t2, s	$K=t1 / t2$
10	4,25	1,52	2,796
100	46,83	8,07	5,803
100	24,12	3,99	6,045
100	16,19	2,17	7,461

7.2.IBM OS/2 implementation

Uniprocessor version provides remote DBs access and processing within LAN, maximal level of parallelism based on multiprocessors (processes) mode of IBM PC/486 based (with or without transputer board) DB server. Communications with transputer-based computer is provided at the DB level. Experimental version has been developed for IBM PC/486 DB server and provides several times decreasing the query graph processing time in comparison with standard IBM OS/2 1.3 and 2.0 tools.