

Listed are abstracts from recent papers by IBM authors. Inquiries should be directed to the publications cited.

A clustering algorithm for hierarchical structures, Schkolnik (RES San Jose, CA), *ACM Transactions on Database Systems* 2, No. 1, 27-44 (March 1977). With a paging environment assumed, the problem of determining how to store a hierarchical structure in order to minimize the expected access time is examined. The solution space considered is the set of partitions of the hierarchical structure, each partition being stored in hierarchical order. A very fast algorithm which determines the optimal partition of the tree is described. The algorithm has been used to determine the best partition of an IMS type tree into data set groups as well as to evaluate the cost of different alternatives. Actual measurements against the restructured data base have demonstrated the validity of this technique. Measurements have also shown that selecting the wrong choice of clustering may substantially increase the expected access time.

Communications systems analysis and design using nomographs, T. Thananitayaudom (SCD Kingston, NY), *Computer Networks* 1, No. 3, 147-154 (January 1977). A method is introduced for developing nomographs as tools for obtaining solutions of queuing equations. The approach is found to be particularly useful when repetitive solutions are required as, for example, in the early stages of planning a communication system. The nomograph approach makes use of graphical representations based on the geometrical relationship of scales representing variables of an equation. The accuracy of the solution depends primarily on the assumptions and limitations of the queuing model. Only the scale length and the range of variables chosen affects the accuracy of solutions from the nomograph.

Improving the access time for random access files, P. Clapson (IBM UK Ltd.), *Communications of the ACM* 20, No. 3, 127-135 (March 1977). In a file of records with no inherent order, access to any random record is fastest when a random access method is used. In this method, each record has a key which is transformed into an address location within the file. If the access method is to be used for any file size or key set, the transformation method must be able to cope with general kinds of key properties occurring within key sets that might tend to distort the transformation distribution. A transformation method designed to cope with common key properties is described in this paper.

Physical integrity in a large segmented database, R. A. Lorie (RES San Jose, CA), *ACM Transactions on Database Systems* 2, No. 1, 91-104 (March 1977). A recovery scheme based on maintaining a dual mapping between pages and their locations on disk is proposed for instances of system failure (hardware or software error which causes the contents of main storage to be lost). One mapping represents the current state of a segment being modified; the other represents a previous backup state. At any time the backup state can be replaced by the current state without any data merging. A facility for protection against damage to the auxiliary storage itself is also proposed. It is shown how such protection may be obtained by copying on a tape only those pages that have been modified since the last checkpoint.

Abstracts