A Note on Dynamic File Organization Models

Choice of file organization may be made from the unorganized 'pile' organization though a limited number of fundamental organizations to those more sophisticated forms found in database systems. This note is aimed at discussing the dynamic performance and cost of this spread of organizations and refers to a recently published dynamic analytical model.

The choice of file organization selected for analysis in the recently developed dynamic model[1] was largely dictated by the work of Weiderhold.[2] Measures of performance used and the formulae derived are generated from this basic approach and as a result the format adopted in discussing alternative model structures relies on Weiderhold's proposals.

The range of complexity reflected by these organizations varies from the unorganized 'pile' file organization through a number of fundamental file organizations, such as sequential, indexed-sequential, direct and inverted, to the multi-ring organization, which would be better considered to be the first level of a simple database system. It is not really appropriate to compare the multi-ring with the other simpler organizations which are often used as building blocks in database systems. Indeed, once this observation has been made, a more complex area is entered in which evaluation of generalized database management systems is under consideration as opposed to the performance of individual file organizations.

Furthermore, the choice of organization referred to as the 'pile' is rather unrealistic, and a more suitable replacement at this simple level would be a single record serial file.

Table 1. Performance of single record type serial file

$T_f = \frac{1}{2}(n + \overline{I + Mt_i})R/t', R = aV$
$T_n = T_f$
$T_i = s + 3r + B/t$
$T_u = T_f + 2r + T_i$
$T_d = T_f + \underline{2r}$
$T_x = (n + \overline{I + Mt_i})R/t'$
$T_{xs} = T_x + 2n(R/t')\log_2 n_i$
$T_y = T_x + n_i' R/t'$

The simple single record type serial file has performance times quoted in Table 1, from which costs on a dynamic basis may be calculated using the notation developed earlier.[1] Costs emerge at approximately 90% those of the pile organization.

A restricted set, therefore, of fundamental file organizations is proposed, namely the sequential, indexed-sequential, direct and inverted organizations. This excludes the pile or serial file as lacking sufficient organization and the multi-ring file as offering excessive organization.

R. H. DAVIS
Department of Computer Science,
Heriot-Watt University,
Edinburgh, UK

References

1. R. H. Davis and P. Coumpas, A dynamic file organisation model, *The Computer Journal* in press.
2. G. Weiderhold, *Database Design*, McGraw-Hill Kogashuka (1977).

Received July 1983