

Published: 12/18/67

Identification

Load Control Overview
C. Marceau

Purpose

The load control module has 2 responsibilities:

- 1) to control system load by enforcing a limit on the number of user process groups which may be logged in concurrently;
- 2) to decide, if demand on the system exceeds the maximum set on system load, which process-groups have to go.

Controlling System Load

To fulfill its first function, load control will ultimately include modules which measure system performance (throughput, response time) and dynamically compute a limit on the number of user process groups. (In initial versions of Multics this limit is a system constant which may be reset by patching.)

Note on Terminology

Since a user may log in more than once, he may have several process-groups logged in concurrently. In this section, as elsewhere, it is convenient to speak of users rather than of user process-groups or "instances" of a user. This should cause no confusion if the reader remembers that the system machinery sees process-groups, not personalities. Thus to say "n users can log in" is only a more convenient (although less precise) way of saying that n user process-groups may be logged in concurrently. They could conceivably all be instances of a single user.

Some Users Lose

If the maximum number of user process-groups allowed is n , and m ($m > n$) users wish to log in, then $m-n$ users lose. The load control module is responsible for determining which process-groups must be automatically logged out (if necessary). It does this by computing a ranking for each user process-group before the group is created. Every user process-group is created by either user control or the absentee monitor. Before creating a user group

these modules ask load control to compute a rank for the prospective group. If the computed rank indicates that the user group should not be created, load control informs its caller of this fact. A negative response from load control causes user control to forbid the user to log in. It causes the absentee monitor to postpone unshelving the absentee job.

(Note: some MSPM sections refer to this aspect of the load control module as the process-group ranker.)

Load Control Process

The load control process contains those modules of load control which measure system performance and dynamically compute the limit on user groups. If the load control process notices that system supply exceeds demand it notifies the absentee monitor that more absentee process groups may be allowed to compete for system resources.

Process-Group Ranking

Load Control maintains a ranking of all process groups currently logged in, together with the current limit n on user process groups, and the limit m ($m \leq n$) on interactive process groups. Ranking policy is described in BQ.5.02. When a user tries to log in, the projected rank of his process-group is computed. If the system is saturated, and he outranks the lowest ranking logged-in user, load control automatically logs out that low ranking user, and allows the new user to log in.