

Example “Real” Scheduler

The priority of each thread is the sum of three priority values,

- Base priority
- Dynamic priority
- Compensation priority

When the scheduler chooses a process to run, it chooses the process with the highest sum.

Each thread’s base priority level is initialized to 0 and is intended to be adjusted externally (such as via a `setpriority` syscall or “renice” operation.) The values for base priority must lie in the range $-256 \dots 256$.

Each thread’s dynamic priority level is adjusted to prioritize IO-bound threads and penalize CPU-bound threads. A thread that blocks or yields voluntarily has its dynamic priority raised. A thread that is stopped by the clock interrupt has its dynamic priority lowered. The change made to dynamic priority is `SCHED_DYNPRI_DELTA`, which is fixed at 16.

The dynamic priority is clamped to the range `SCHED_DYNPRI_MIN` to `SCHED_DYNPRI_MAX` ($-256 \dots 256$). The lower bound prevents totally cpu-bound threads from being abused too badly. The upper bound keeps threads from getting too much credit for doing tons of little I/Os.

Finally, each thread has a compensation priority level. This is cleared to zero when a thread is chosen to run, and incremented each time a thread is *not* chosen to run. This means that even extremely low-priority threads will eventually be chosen to run.

The change made to compensation priority level, the compensation priority delta, is a variable. The normal value of this variable is 1. However, it could potentially be set to 0, which would, in combination with a low base priority, cause a thread to be run more or less only when nothing else is runnable. It could also be potentially set higher, which could help to ensure a fast response time on a busy system. One could conceivably adjust this value dynamically as well.

The compensation priority level is not clamped, but is limited in practice by the fact that a thread with a high compensation priority will eventually be run and lose the compensation priority.

The values picked for all these constants are somewhat arbitrary and could be tuned.