Shortest remaining time

From Wikipedia, the free encyclopedia

Shortest remaining time, also known as **shortest remaining time first (SRTF)**, is a scheduling method that is a preemptive version of shortest job next scheduling. In this scheduling algorithm, the process with the smallest amount of time remaining until completion is selected to execute. Since the currently executing process is the one with the shortest amount of time remaining by definition, and since that time should only reduce as execution progresses, processes will always run until they complete or a new process is added that requires a smaller amount of time.

Shortest remaining time is advantageous because short processes are handled very quickly. The system also requires very little overhead since it only makes a decision when a process completes or a new process is added, and when a new process is added the algorithm only needs to compare the currently executing process with the new process, ignoring all other processes currently waiting to execute.

Like shortest job first, it has the potential for process starvation; long processes may be held off indefinitely if short processes are continually added. This threat can be minimal when process times follow a heavy-tailed distribution.[1]

Like shortest job next scheduling, shortest remaining time scheduling is rarely used outside of specialized environments because it requires accurate estimations of the runtime of all processes that are waiting to execute.

References [edit]

1. ^ Harchol-Balter, Mor; Schroeder, Bianca; Bansal, Nikhil; Agrawal, Mukesh (2003). "Size-Based Scheduling to Improve Web Performance". *ACM Transactions on Computer Systems* **21** (2): 207–233. doi:10.1145/762483.762486.

v · t · e	**Queueing theory**	[hide]
Single queueing nodes	D/M/1 queue · M/D/1 queue · M/D/c queue · M/M/1 queue (Burke's theorem) · M/M/c queue · M/M/∞ queue · M/G/1 queue (Pollaczek–Khinchine formula · Matrix analytic method) · M/G/k queue · G/M/1 queue · G/G/1 queue (Kingman's formula · Lindley equation) · Fork–join queue · Bulk queue
Arrival processes	Poisson process · Markovian arrival process · Rational arrival process
Queueing networks	Jackson network (Traffic equations) · Gordon–Newell theorem (Mean value analysis · Buzen's algorithm) · Kelly network · G-network · BCMP network
Service policies	FIFO · LIFO · Processor sharing · Shortest job first · **Shortest remaining time**
Key concepts	Continuous-time Markov chain · Kendall's notation · Little's law · Product-form solution (Balance equation · Quasireversibility · Flow-equivalent server method) · Arrival theorem · Decomposition method · Beneš method
Limit theorems	Fluid limit · Mean field theory · Heavy traffic approximation (Reflected Brownian motion)
Extensions	Fluid queue · Layered queueing network · Polling system · Adversarial queueing network · Loss network · Retrial queue
Information systems	Data buffer · Erlang (unit) · Erlang distribution · Flow control (data) · Message queue · Network congestion · Network scheduler · Pipeline (software) · Quality of service · Scheduling (computing) · Teletraffic engineering
Category		

Categories: Processor scheduling algorithms