

Coordinated and Efficient Huge Page Management with Ingens

Summary:

The objective of this paper is to present the Ingens framework. Having its name originated from the Latin word that literally means "huge", Ingens supports well-coordinated massive-scale paging via just a few basic primitive instructions. It aims to cope with overheads in heavy interactions between CPU and RAM, especially in terms of translating hardware addresses with TLBs, without the need for convoluted combinations of conventional low-level operations. Ingens monitors contiguity as a first-class resource and keeps track of utilization and access frequency of memory pages, and from this, it can deal with issues on overall performance and fairness between processes.

Strengths:

Experiments clearly show that Ingens enhances both fairness and performance in massive-scale paging, and they also show that Ingens greatly reduces tail-latency and memory bloat of core applications within a system. From this, it is clearly proved that Ingens provides better environment for dealing with massive-scale paging which is expected to continuously increasing as hardware technologies progress.