

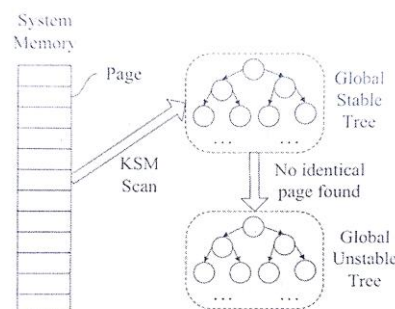
연구 노트

2014314920 박수진

2018.04.27

KSM (Kernel Samepage Merging)

- 리눅스 커널 2.6.32 버전부터 지원하는 메모리 중복제거 기법
- 주기적으로 시스템에 할당된 페이지를 스캔하여 내용이 일치하는 페이지를 하나의 페이지로 병합한다.
- 병합 대상이 되는 페이지는 '레드 블랙 트리 (red black tree)' 를 사용하여 관리
- 공유 페이지는 안정트리(stable tree), 비공유 페이지는 불안정트리(unstable tree)로 관리한다.
- 공유 가능한 페이지는 우선적으로 안정 트리와 비교된 후 안정트리에서 발견되지 않으면 불안정트리에서 검색한다. → 빠른 검색 시간
- `madvise` 로 지정된 익명 페이지(anonymous page)만을 대상으로 한다.
- `int madvise(addr, length, MADV_MERGEABLE)` 시스템콜을 사용해서 스캔대상영역으로 지정한다.



[KSM에서 안정트리와 불안정 트리]

- KSM은 두 개의 global 레드 블랙 트리를 유지한다. (호스트 서버의 전체 메모리 페이지 내부)
- 안정트리와 불안정트리의 각 노드는 페이지의 내용을 직접적으로 index 하고, 이는 레드블랙트리의 key가 된다.
- 과정
 - 1 candidate page는 안정트리 내에 있는지 검색된다.
 - 2 안정트리에서 발견되었다면, 해당 페이지와 병합(merged) 및 공유(shared) 된다.
 - 2.1 발견되지 못했다면, 불안정트리에서 검색한다.
 - 2.1.1 불안정트리에 있다면, 병합 후 불안정트리에서 삭제 → 안정트리로 옮겨서 병합한다.
 - 2.1.2 불안정트리에도 없다면, 불안정트리에 leaf node로 삽입한다.
- 불안정트리의 페이지는 write-protect가 아니기 때문에, 매 스캔 라운드마다 재구성이 필요하다. 스캔 라운드 중에 페이지가 변경될 수 있기 때문이다.

KSM files

- `/sys/kernel/mm/ksm` 디렉토리 내부에 존재하는 파일들로, KSM 모듈을 세팅하거나, 실시간으로 정보를 알 수 있다.
- **pages_shared** : 공유 페이지(=안정 트리에 저장된 페이지)의 개수
- **pages_sharing** : 페이지 병합으로 인해 해제된 페이지의 개수. 즉, 절약된 페이지의 개수
- **pages_unshared** : 공유되지 않은 페이지(=불안정 트리에 저장된 페이지)의 개수.
- **pages_volatile** : 트리에 넣기에 너무 빠르게 변하는 페이지의 개수
- **pages_to_scan** : KSM 이 sleep 되기 전까지 스캔할 페이지의 개수
- **sleep_millisecs** : 다음 스캔 전까지 KSM 이 sleep 할 시간
- **full_scans** : 모든 병합 가능한 영역이 스캔된 총 횟수
- **run** : 0 일 때 KSM 이 중지되고, 병합된 페이지는 유지.
1 일 때 KSM 이 실행되고 페이지 병합
2 일 때 KSM 이 중지되고 병합된 페이지를 모두 병합 해제
- `pages_sharing/pages_shared` 의 비율이 높으면 good sharing
- `pages_unshared/pages_sharing` 은 wasted effort 를 의미한다.