

# [PATCH] [27/34] i386: pte clear optimization

Source: <http://linux.derkeiler.com/Mailing-Lists/Kernel/2007-04/msg12035.html>

- From: Andi Kleen <ak@xxxxxxx>
- Date: Mon, 30 Apr 2007 17:50:00 +0200 (CEST)

From: Zachary Amsden <zach@xxxxxxxxxxx>

When exiting from an address space, no special hypervisor notification of page table updates needs to occur; direct page table hypervisors, such as Xen, switch to another address space first (init\_mm) and unprotects the page tables to avoid the cost of trapping to the hypervisor for each pte\_clear. Shadow mode hypervisors, such as VMI and lhype don't need to do the extra work of calling through paravirt-ops, and can just directly clear the page table entries without notifying the hypervisor, since all the page tables are about to be freed.

So introduce native\_pte\_clear functions which bypass any paravirt-ops notification. This results in a significant performance win for VMI and removes some indirect calls from zap\_pte\_range.

Note the 3-level paging already had a native\_pte\_clear function, thus demanding argument conformance and extra args for the 2-level definition.

Signed-off-by: Zachary Amsden <zach@xxxxxxxxxxx>  
Signed-off-by: Andrew Morton <akpm@xxxxxxxxxxxxxxxxxxxxxx>  
Signed-off-by: Andi Kleen <ak@xxxxxxx>

---

include/asm-i386/pgtable-2level.h | 5 +++++  
include/asm-i386/pgtable.h | 2 +-  
2 files changed, 6 insertions(+), 1 deletion(-)

Index: linux/include/asm-i386/pgtable-2level.h

```

=====
--- linux.orig/include/asm-i386/pgtable-2level.h
+++ linux/include/asm-i386/pgtable-2level.h
@@ -36,6 +36,11 @@ static inline void native_set_pmd(pmd_t
#define pte_clear(mm,addr,xp) do { set_pte_at(mm, addr, xp, __pte(0)); } while (0)
#define pmd_clear(xp) do { set_pmd(xp, __pmd(0)); } while (0)

+static inline void native_pte_clear(struct mm_struct *mm, unsigned long addr, pte_t *xp)
+{

```

[PATCH] [27/34] i386: pte clear optimization

```
+ *xp = __pte(0);
+}
+
static inline pte_t native_ptep_get_and_clear(pte_t *xp)
{
return __pte(xchg(&xp->pte_low, 0));
Index: linux/include/asm-i386/pgtable.h
```

```
=====
--- linux.orig/include/asm-i386/pgtable.h
+++ linux/include/asm-i386/pgtable.h
@@ -344,7 +344,7 @@ static inline pte_t ptep_get_and_clear_f
pte_t pte;
if (full) {
pte = *ptep;
- pte_clear(mm, addr, ptep);
+ native_pte_clear(mm, addr, ptep);
} else {
pte = ptep_get_and_clear(mm, addr, ptep);
}
-
```

To unsubscribe from this list: send the line "unsubscribe linux-kernel" in the body of a message to majordomo@xxxxxxxxxxxxxxxxxxx  
More majordomo info at <http://vger.kernel.org/majordomo-info.html>  
Please read the FAQ at <http://www.tux.org/lkml/>