

[PATCH 01/05] mm fix __alloc_pages cpuset ALLOC_* flags

Source: <http://linux.derkeiler.com/Mailing-Lists/Kernel/2005-11/4691.html>

From: Paul Jackson (pj_at_sgi.com)

Date: 11/14/05

Date: Sun, 13 Nov 2005 20:03:29 -0800 (PST)
To: akpm@osdl.org, linux-kernel@vger.kernel.org

Two changes to the setting of the ALLOC_CPUSET flag in mm/page_alloc.c: __alloc_pages()

- 1) A bug fix – the "ignoring mins" case should not be honoring ALLOC_CPUSET. This case of all cases, since it is handling a request that will free up more memory than is asked for (exiting tasks, e.g.) should be allowed to escape cpuset constraints when memory is tight.
- 2) A logic change to make it simpler. Honor cpubsets even on GFP_ATOMIC (!wait) requests. With this, cpuset confinement applies to all requests except ALLOC_NO_WATERMARKS, so that in a subsequent cleanup patch, I can remove the ALLOC_CPUSET flag entirely. Since I don't know any real reason this logic has to be either way, I am choosing the path of the simplest code.

Signed-off-by: Paul Jackson <pj@sgi.com>

```
---
mm/page_alloc.c |    5 +----
1 files changed, 2 insertions(+), 3 deletions(-)
--- 2.6.14-mm2.orig/mm/page_alloc.c      2005-11-12 22:25:03.305301135 -0800
+++ 2.6.14-mm2/mm/page_alloc.c      2005-11-12 22:27:30.519813285 -0800
@@ -933,8 +933,7 @@ restart:
         alloc_flags |= ALLOC_HARDER;
         if (gfp_mask & __GFP_HIGH)
             alloc_flags |= ALLOC_HIGH;
-         if (wait)
-             alloc_flags |= ALLOC_CPUSET;
+         alloc_flags |= ALLOC_CPUSET;

         /*
          * Go through the zonelist again. Let __GFP_HIGH and allocations
@@ -956,7 +955,7 @@ restart:
         nofail_alloc:

                 /* go through the zonelist yet again, ignoring mins */
                 page = get_page_from_freelist(gfp_mask, order,
-                 zonelist, ALLOC_NO_WATERMARKS|ALLOC_CPUSET);
```

Linux-Kernel: [PATCH 01/05] mm fix __alloc_pages cpuset ALLOC_* flags

```
+         zonelist, ALLOC_NO_WATERMARKS);
        if (page)
            goto got_pg;
        if (gfp_mask & __GFP_NOFAIL) {
--
            I won't rest till it's the best ...
            Programmer, Linux Scalability
            Paul Jackson <pj@sgi.com> 1.650.933.1373
-
```

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