
Subject: [RFC PATCH 1/5] memcg VM overcommit documentation

Posted by [Andrea Righi](#) on Mon, 09 Jun 2008 23:32:59 GMT

[View Forum Message](#) <> [Reply to Message](#)

Documentation of the VM overcommit memory controller included in the generic memory controller documentation: basic description and usage.

Signed-off-by: Andrea Righi <righi.andrea@gmail.com>

Documentation/controllers/memory.txt | 29 ++++++
1 files changed, 29 insertions(+), 0 deletions(-)

diff --git a/Documentation/controllers/memory.txt b/Documentation/controllers/memory.txt

index 866b9cd..e984bfb 100644

--- a/Documentation/controllers/memory.txt

+++ b/Documentation/controllers/memory.txt

@@ -12,6 +12,7 @@ c. Provides *zero overhead* for non memory controller users

d. Provides a double LRU: global memory pressure causes reclaim from the global LRU; a cgroup on hitting a limit, reclaims from the per cgroup LRU

+e. Provide distinct cgroup VM overcommit accounting and handling

NOTE: Swap Cache (unmapped) is not accounted now.

@@ -142,6 +143,31 @@ The reclaim algorithm has not been modified for cgroups, except that pages that are selected for reclaiming come from the per cgroup LRU list.

+2.5 VM overcommit accounting and handling

+

+The concept of committed VM is replicated within each cgroup as well as global committed memory concept. Each cgroup can set its own overcommit policy using the files:

+

+memory.overcommit_memory

+memory.overcommit_ratio

+

+These settings override the system sysctl settings (`vm.overcommit_memory` and `vm.overcommit_ratio`) and they apply locally to the cgroup they refer.

+

+Global sysctl settings are initially used by the root level cgroups. Child cgroups initially inherit the parent's settings. Each cgroup can change its own overcommit parameters at any time simply modifying the files `memory.overcommit_memory` and/or `memory.overcommit_ratio`.

+

+Statistics about the current committed space and limit are reported in `memory.overcommit_as` for each cgroup.

+

+Per-cgroup overcommit limit depends of the local cgroup overcommit settings and
+memory limit (RSS + cache) imposed by the memory controller.
+
+See "Documentation/vm/overcommit-accounting" for additional details.

2. Locking

The memory controller uses the following hierarchy
@@ -230,6 +256,9 @@ carried forward. The pages allocated from the original cgroup still
remain charged to it, the charge is dropped when the page is freed or
reclaimed.

+The amount of the task's committed VM, instead, is uncharged from the old
+cgroup and accounted to the newer.

4.3 Removing a cgroup

A cgroup can be removed by rmdir, but as discussed in sections 4.1 and 4.2, a

--
1.5.4.3

Containers mailing list
Containers@lists.linux-foundation.org
<https://lists.linux-foundation.org/mailman/listinfo/containers>
