
Subject: [PATCH v3 28/28] Documentation: add documentation for slab tracker for memcg

Posted by [Glauber Costa](#) on Fri, 25 May 2012 13:03:48 GMT

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

In a separate patch, to aid reviewers.

Signed-off-by: Glauber Costa <glommer@parallels.com>

CC: Christoph Lameter <cl@linux.com>

CC: Pekka Enberg <penberg@cs.helsinki.fi>

CC: Michal Hocko <mhocko@suse.cz>

CC: Kamezawa Hiroyuki <kamezawa.hiroyu@jp.fujitsu.com>

CC: Johannes Weiner <hannes@cmpxchg.org>

CC: Suleiman Souhlal <suleiman@google.com>

CC: Randy Dunlap <rdunlap@xenotime.net>

Documentation/cgroups/memory.txt | 33 ++++++
1 files changed, 33 insertions(+), 0 deletions(-)

diff --git a/Documentation/cgroups/memory.txt b/Documentation/cgroups/memory.txt
index 4c95c00..9accaa1 100644

--- a/Documentation/cgroups/memory.txt

+++ b/Documentation/cgroups/memory.txt

@@ -75,6 +75,12 @@ Brief summary of control files.

memory.kmem.tcp.limit_in_bytes # set/show hard limit for tcp buf memory

memory.kmem.tcp.usage_in_bytes # show current tcp buf memory allocation

+ memory.kmem.limit_in_bytes # set/show hard limit for general kmem memory

+ memory.kmem.usage_in_bytes # show current general kmem memory allocation

+ memory.kmem.failcnt # show current number of kmem limit hits

+ memory.kmem.max_usage_in_bytes # show max kmem usage

+ memory.kmem.slabinfo # show cgroup-specific slab usage information

+

1. History

The memory controller has a long history. A request for comments for the memory

@@ -271,6 +277,14 @@ cgroup may or may not be accounted.

Currently no soft limit is implemented for kernel memory. It is future work to trigger slab reclaim when those limits are reached.

+Kernel memory is not accounted until it is limited. Users that want to just

+track kernel memory usage can set the limit value to a big enough value so

+the limit is guaranteed to never hit. A kernel memory limit bigger than the

+current memory limit will have this effect as well.

+

+This guarantes that this extension is backwards compatible to any previous

+memory cgroup version.

+

2.7.1 Current Kernel Memory resources accounted

* sockets memory pressure: some sockets protocols have memory pressure @@ -279,6 +293,24 @@ per cgroup, instead of globally.

* tcp memory pressure: sockets memory pressure for the tcp protocol.

+* slab/kmalloc:

+

+When slab memory is tracked (memory.kmem.limit_in_bytes != -1ULL), both +memory.kmem.usage_in_bytes and memory.usage_in_bytes are updated. When +memory.kmem.limit_in_bytes is left alone, no tracking of slab caches takes +place.

+

+Because a slab page is shared among many tasks, it is not possible to take +any meaningful action upon task migration. Slabs created in a cgroup stay +around until the cgroup is destructed. Information about the slabs used +by the cgroup is displayed in the cgroup file memory.kmem.slabinfo. The format +of this file is and should remain compatible with /proc/slabinfo.

+

+Upon cgroup destruction, slabs that holds no live references are destructed. +Workers are fired to destroy the remaining caches as they objects are freed.

+

+Memory used by dead caches are shown in the proc file /proc/dead_slabinfo

+

3. User Interface

0. Configuration

@@ -287,6 +319,7 @@ a. Enable CONFIG_CGROUPS

b. Enable CONFIG_RESOURCE_COUNTERS

c. Enable CONFIG_CGROUP_MEM_RES_CTLR

d. Enable CONFIG_CGROUP_MEM_RES_CTLR_SWAP (to use swap extension)

+d. Enable CONFIG_CGROUP_MEM_RES_CTLR_KMEM (to use experimental kmem extension)

1. Prepare the cgroups (see cgroups.txt, Why are cgroups needed?)

mount -t tmpfs none /sys/fs/cgroup

--

1.7.7.6
