Subject: [RFC][PATCH] memory controller per zone patches take 2 [0/10] introduction
Posted by KAMEZAWA Hiroyuki on Fri, 16 Nov 2007 10:11:07 GMT
View Forum Message <> Reply to Message

Hi, this is updated version of patch set implementing per-zone on memory cgroup.

I still uses x86_64/fake NUMA (my ia64/NUMA box is under maintainance....)
So, RFC again. (I'd like to do 3rd update in the next week.)

Major Changes from previous one.
- per-zone-lru_lock patch is added.
- all per-zone objects of memory cgroup are treated in same way.
- page migration is handled.
- restructured and cleaned up.

Todo:
- do test on "real" NUMA.
- merge YAMAMOTO-san's background page reclaim patch set on this. (If I can)
- performance measurement at some point
- more cleanup and adding meaningful comments
- confirm added logic in vmscan.c is really sane.

Overview:

All per-zone obects are put into
==
```
struct mem_cgroup_per_zone {
        /*
         * spin_lock to protect the per cgroup LRU
         */
        spinlock_t              lru_lock;
        struct list_head        active_list;
        struct list_head        inactive_list;
        unsigned long count[NR_MEM_CGROUP_ZSTAT];
};
```
==
And this per-zone area is accessed by following functions.
==
```
mem_cgroup_zoneinfo(struct mem_cgroup *mem, int nid, int zid)
page_cgroup_zoneinfo(struct page_cgroup *pc)
```
==

Typical usage is following.
==
```
        mz = page_cgroup_zoneinfo(pc);
        spin_lock_irqsave(&mz->lru_lock, flags);
        __mem_cgroup_add_list(pc);
```

spin_unlock_irqrestore(&mz->lru_lock, flags);
==

Thanks,
-Kame

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