
Subject: Re: [PATCH v7 06/10] tcp buffer limitation: per-cgroup limit
Posted by [KAMEZAWA Hiroyuki](#) on Wed, 30 Nov 2011 02:00:29 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Tue, 29 Nov 2011 21:56:57 -0200
Glauber Costa <glommer@parallels.com> wrote:

> This patch uses the "tcp.limit_in_bytes" field of the kmem_cgroup to
> effectively control the amount of kernel memory pinned by a cgroup.
>
> This value is ignored in the root cgroup, and in all others,
> caps the value specified by the admin in the net namespaces'
> view of tcp_sysctl_mem.
>
> If namespaces are being used, the admin is allowed to set a
> value bigger than cgroup's maximum, the same way it is allowed
> to set pretty much unlimited values in a real box.
>
> Signed-off-by: Glauber Costa <glommer@parallels.com>
> CC: David S. Miller <davem@davemloft.net>
> CC: Hiroyouki Kamezawa <kamezawa.hiroyu@jp.fujitsu.com>
> CC: Eric W. Biederman <ebiederm@xmission.com>

you need one more fix.
(please add changelog.)

```
> +static int tcp_update_limit(struct mem_cgroup *memcg, u64 val)
> +{
> + struct net *net = current->nsproxy->net_ns;
> + struct tcp_memcontrol *tcp;
> + struct cg_proto *cg_proto;
> + u64 old_lim;
> + int i;
> + int ret;
> +
> + cg_proto = tcp_prot.proto_cgroup(memcg);
> + if (!cg_proto)
> + return -EINVAL;
> +
> + tcp = tcp_from_cgproto(cg_proto);
> +
> + old_lim = res_counter_read_u64(&tcp->tcp_memory_allocated, RES_LIMIT);
> + ret = res_counter_set_limit(&tcp->tcp_memory_allocated, val);
> + if (ret)
> + return ret;
> +
> + for (i = 0; i < 3; i++)
```

```
> + tcp->tcp_prot_mem[i] = min_t(long, val >> PAGE_SHIFT,  
> +      net->ipv4.sysctl_tcp_mem[i]);  
> +  
> + if (val == RESOURCE_MAX)  
> +      jump_label_dec(&memcg_socket_limit_enabled);
```

```
if (val == RESOURCE_MAX && old_lim != RESOURCE_MAX)
```

Thanks,
-Kame
