Subject: [PATCH 0/3] cgroup: block device i/o bandwidth controller (v3)
Posted by Andrea Righi on Fri, 20 Jun 2008 10:05:32 GMT
View Forum Message <> Reply to Message

The goal of the i/o bandwidth controller is to improve i/o performance predictability and provide better QoS for different cgroups sharing the same block devices.

Respect to other priority/weight-based solutions the approach used by this controller is to explicitly choke applications' requests that directly (or indirectly) generate i/o activity in the system.

The direct bandwidth limiting method has the advantage of improving the performance predictability at the cost of reducing, in general, the overall performance of the system (in terms of throughput).

Detailed informations about design, its goal and usage are described in the documentation.

Tested against latest git (2.6.26-rc6).

The all-in-one patch (and previous versions) can be found at:
http://download.systemimager.org/~arighi/linux/patches/io-throttle/

Changelog: (v2 -> v3)
- scalability improvement: replaced the rbtree structure with a linked list to store multiple per block device I/O limiting rules; this allows to use RCU to protect the whole list structure, since the elements in the list are supposed to change rarely (this also provides zero overhead for cgroups that don't use any I/O limitation)
- improved user interface
  - now it's possible to specify a suffix k, K, m, M, g, G to express bandwidth values in KB/s, MB/s or GB/s
  - current per block device I/O usage is reported in blockio.bandwidth
- renamed cgroup_io_account() in cgroup_io_throttle()
- updated the documentation

TODO:
- implement I/O throttling using a token bucket algorithm, as suggested by Carl Henrik Lunde, in addition to the current leaky bucket approach
- provide a modular interface to switch between different i/o throttling algorithms at run-time

-Andrea