Queue sysfs files

This text file will detail the queue files that are located in the sysfs tree for each block device. Note that stacked devices typically do not export any settings, since their queue merely functions are a remapping target. These files are the ones found in the /sys/block/xxx/queue/ directory.

Files denoted with a RO postfix are readonly and the RW postfix means read-write.

hw_sector_size (RO)

This is the hardware sector size of the device, in bytes.

max_hw_sectors_kb (RO)

This is the maximum number of kilobytes supported in a single data transfer.

max_sectors_kb (RW)

This is the maximum number of kilobytes that the block layer will allow for a filesystem request. Must be smaller than or equal to the maximum size allowed by the hardware.

nomerges (RW)

This enables the user to disable the lookup logic involved with IO merging requests in the block layer. By default (0) all merges are enabled. When set to 1 only simple one-hit merges will be tried. When set to 2 no merge algorithms will be tried (including one-hit or more complex tree/hash lookups).

nr_requests (RW)

This controls how many requests may be allocated in the block layer for read or write requests. Note that the total allocated number may be twice this amount, since it applies only to reads or writes (not the accumulated sum).

read_ahead_kb (RW)

Maximum number of kilobytes to read-ahead for filesystems on this block device.

rq_affinity (RW)

If this option is enabled, the block layer will migrate request completions to the CPU that originally submitted the request. For some workloads this provides a significant reduction in CPU cycles due to caching effects.

scheduler (RW)

When read, this file will display the current and available IO schedulers for this block device. The currently active IO scheduler will be enclosed in [] brackets. Writing an IO scheduler name to this file will switch

control of this block device to that new IO scheduler. Note that writing an IO scheduler name to this file will attempt to load that IO scheduler module, if it isn't already present in the system.

Jens Axboe <jens.axboe@oracle.com>, February 2009