

Disk Performance

I/O Schedulers

In Linux, the I/O scheduler is responsible for determining the order in which block I/O operations are submitted to storage devices. The scheduler affects the performance and behavior of disk I/O operations, impacting both throughput and latency. Different I/O schedulers are available, each optimized for specific workloads and scenarios.

By default, I/O tasks are scheduled by fifo - first in first out.

Common Linux I/O Schedulers

1. CFQ (Completely Fair Queuing) - Provides a balanced approach to I/O scheduling, aiming to give each process a fair share of the I/O bandwidth.
2. Deadline - Designed to prevent starvation of I/O operations by imposing deadlines on requests.
3. NOOP - Implements a simple FIFO (First-In, First-Out) queue, essentially a passthrough scheduler.
4. BFQ (Budget Fair Queuing) - Aims to provide predictable I/O performance by distributing I/O bandwidth according to budgets assigned to tasks.
5. MQ Deadline (Multiqueue Deadline) - Similar to the Deadline scheduler but designed for multiqueue block devices.
6. Kyber - A relatively new scheduler designed to work well with modern hardware and to provide low latency.
7. BFQ (Budget Fair Queuing) - Bypasses software I/O scheduling, relying entirely on hardware-level I/O management.

Changing the scheduler

The scheduler is set on a per disk basis, not per filesystem.

(none persistent) method:

The scheduler can be set within the below file:

```
/sys/block/diskname/queue/scheduler
```

Within this file, you'll likely see a number of the potential I/O scheduler types. The one surrounded by [] is the currently selected scheduler, ie:

```
cat /sys/block/sda/queue/scheduler
[mq-deadline] none
```

(persistent) method:

To persistently set a disk scheduler, we'll need to alter the grub configuration

```
/etc/default/grub
```

append the below to the line prefaced with 'GRUB_CMDLINE_LINUX='

```
elevator=schedulertype
```

Save the new GRUB configuration:

```
update-grub2
```

or

```
grub-mkconfig
```

=====

Disk Performance Troubleshooting Tools

sar

lsof

We can use the lsof command to check what parts of a disk a process is accessing.

```
lsof -p pid
```

```
lsof -c command
```

sysstat tools

The `sysstat` package includes a collection of performance monitoring tools for Unix-like systems.

iotop

```
iotop [options]
```

| | |
|----|---|
| -a | aggregate disk IO over time (duration of command) |
|----|---|

iostat

Similar to `iotop`. Provides statistical information about I/O device loading. It reports on CPU utilization, device I/O statistics, and system throughput, making it useful for overall system performance analysis.

```
iostat
```

ioping

`ioping` is basically the `ping` command but for disks. It is used to measure the IO latency of storage devices - basically it measures how long a storage device will take to respond to an I/O request.

```
ioping [options] target
```

| | |
|----|---|
| -c | specify a number of IO requests to make |
| -i | interval between requests |
| -s | request size (default 4kb) |
| -q | suppress regular output, only show statistics |

Revision #3

Created 2024-06-09 23:12:30 UTC by Daniel

Updated 2024-06-23 22:00:47 UTC by Daniel