8.6.4 I/O Scheduling

I/O schedulers provide a way to optimize disk access requests (read/write). In many cases they try to do this by merging I/O requests to similar location on the disk. Using this method, the drive doesn't need to seek as often, improving the overall response time for disk operations.

This lesson covers the following topics:

- I/O schedulers
- Viewing I/O schedulers
- Changing I/O schedulers

I/O Schedulers

Linux provides several I/O schedulers to choose from. The following table gives a brief summary of three such schedulers:

Scheduler	Description
Noop	Noop is the simplest scheduler. It places all I/O requests into a First in, First Out (FIFO) queue. In addition, read/write requests of a similar purpose are also combined to reduce the number of disk operations and increase the length of system calls. This scheduler is often used for systems that do not need an I/O scheduler. For example, a virtual machine (VM) is running on a host computer which is already using its own I/O scheduler.
Deadline	The Deadline scheduler creates a read queue and a write queue. Since each I/O request has an associated time stamp (used by the kernel for an expiration time), the Deadline scheduler utilizes this time stamp to push I/O request that have reached their deadline to their highest priority. The default Deadline values are 500 ms for read operations and 5,000 ms for write operations. If needed, these values can be adjusted. Because of these values, the Deadline scheduler is often considered the optimal scheduler for read-heavy workloads.
CFQ	The Complete Fairness Queuing (CFQ) input/output (I/O) scheduler works by creating a per-process I/O queue. The goal of CFQ is to provide a fair I/O priority to each process. This is accomplished by first ordering the queues to reduce disk seeking and then servicing these per-process I/O queues, in a round-robin fashion. The benefits of using the CFQ scheduler is that it tries to provide each process with the same priority for disk access. The disadvantage of using the CFQ scheduler is that it makes this scheduler less optimal for environments that might need to prioritize one request type (such as reads) from a single process.

Each disk device has its own scheduler and can be configured independent of the other devices. You can use the **lsblk -o KNAME,TYPE,SIZE,MODEL** command to view these devices.

Viewing I/O Schedulers

To determine which I/O scheduler your system is using run **cat /sys/block/*[disk device]*/queue/scheduler**. A list of the available schedulers will be shown. The scheduler enclosed in brackets is the scheduler currently being used.

Example:
cat /sys/block/sda/queue/scheduler

Result/Output:
noop [deadline] cfq

In this example, this system is using the scheduler named Deadline. Also notice that the noop and cfq schedulers are available.

Changing I/O schedulers

Changing which I/O scheduler is used can be done at runtime or by modifying the Grub boot loader. When changing the scheduler at runtime, the change is applied immediately but on the next boot, the system will run the default scheduler. Modifying the scheduler via the Grub boot loader causes the system to use the same scheduler each time the system is started.

To modify the scheduler at runtime run **echo *"[scheduler]"* > /sys/block/*[disk device]*/queue/scheduler**.

Example:
echo "cfq" > /sys/block/sda/queue/scheduler