

- 一、nice
- 二、weight
  - 2.1 cgroup v1
  - 2.2 cgroup v2
- 三、setpriority/getpriority
- 四、参考

Linux系统是抢占式的，对普通进程采用的是完全公平调度算法（CFS）。

Linux的进程调度并未使用直接均分时间片的方式，而是对优先级进行了改进，采用了两种不同的优先级范围，一种是nice值，范围是-20到+19，越大的nice值意味着更低的优先级，低nice值的进程会获得更多的处理器时间（按比例获得），第二种范围是实时优先级，其值是可配置的，默认情况下它的变化范围是从0到99，与nice值意义相反，越高的实时优先级数值意味着进程优先级越高，任何实时进程的优先级都高于普通进程。

## 一、nice

nice 的范围为：[-20, 19] 的整数值，值越小进程优先级越高。

## 二、weight

### 2.1 cgroup v1

在cgroup v1中，nice 没有对应的用户接口。

weight 对应的读写接口为 `cpu.shares`，范围为：[2, 262144] 的整数值。

nice与weight的换算关系（注意：这是一个近似换算关系，Linux对换算结果进行了微调）：

$$weight = 1024 * 1.25^{-nice}$$

### 2.2 cgroup v2

在cgroup v2中，nice 对应读写接口为 `cpu.weight.nice`。

weight 对应的读写接口为 `cpu.weight`，范围为：[1, 10000] 的整数值。

## 三、setpriority/getpriority

```
int setpriority(int which, int who, int prio);
```

setpriority/getpriority 可以设置/获取进程优先级。只有超级用户（root）允许降低此值。

以下设置不会报错：

```
setpriority(PRIO_PROCESS, 0, 20);
```

但实际生效的nice值为19。

## 四、参考

- [setpriority\(3p\) — Linux manual page](#)
- [为什么Greenplum 的CPU有大量是%ni的占用](#)