# 24.电源管理

DPDK电源管理功能允许用户空间应用程序通过动态调整CPU频率或进入不同的C-State来节省功耗。

* 根据RX队列的利用率动态调整CPU频率。
* 根据自适应算法进入不同层次的C-State，以推测在没有收到数据包的情况下暂停应用的短暂时间段。

调整CPU频率的接口位于电源管理库中。C-State控制是根据不同用例实现的。

## 24.1CPU频率缩放

Linux内核提供了一个用于每个lcore的CPU频率缩放的cpufreq模块。例如，对于cpuX, /sys/devices/system/cpu/cpuX/cpufreq/具有以下用于频率缩放的sys文件：

* fected\_cpus
* os\_limit
* uinfo\_cur\_freq
* uinfo\_max\_freq
* uinfo\_min\_freq
* uinfo\_transition\_latency
* lated\_cpus
* aling\_available\_frequencies
* aling\_available\_governors
* aling\_cur\_freq
* aling\_driver
* aling\_governor
* aling\_max\_freq
* aling\_min\_freq
* aling\_setspeed

在DPDK中，scaling\_governor在用户空间中配置。然后，用户空间应用程序可以通过写入scaling\_setspeed来提示内核以根据用户空间应用程序定义的策略来调整CPU频率。

## 24.2. 通过C-States调节Core负载

只要指定的lcore无任务执行，可以通过设置睡眠来改变Core状态。在DPDK中，如果在轮询后没有接收到分组，则可以根据用户空间应用定义的策略来触发睡眠。

## 24.3.电源管理库API概述

电源管理库导出的主要方法是CPU频率缩放，包括：

* 频率上升：提示内核扩大特定lcore的频率。
* 频率下降：提示内核缩小特定lcore的频率。
* 频率最大：提示内核将特定lcore的频率最大化。
* 频率最小：提示内核将特定lcore的频率降至最低。
* 获取有效的频率：从sys文件中读取特定lcore的可用频率。
* Freq获取：获取当前的特定lcore的频率。
* 频率设置：提示内核为特定的lcore设置频率。

## 24.3.示例

电源管理机制可用于在进行L3转发时节省功耗。

## 24.4.参考

* l3fwd-power: DPDK提供的示例应用程序，实现功耗管理下的L3转发。
* “功耗管理下的L3转发”章节请参阅《DPDK Sample Application’s User Guide》。