

Московский физико-технический институт
(государственный университет)
Факультет радиотехники и кибернетики
Кафедра информатики и вычислительной техники

Реализация состояния глубокого сна микропроцессора Эльбрус-4С средствами ядра Linux

Выпускная квалификационная работа
(бакалаврская работа)

Выполнил:
студент 313 группы Демидов А. А.
Научный руководитель:
Кравцунов Е. М.

Москва, 2017

Постановка задачи

Цель работы

Снижение энергопотребления вычислительных комплексов на базе микропроцессоров Эльбрус-4С во время простоя (idle)

Задача

- Реализация алгоритма перевода вычислительного ядра в состояние глубокого сна
- Встраивание алгоритма в подсистему `cruidle`

Требования

Возможность управления через интерфейс `sysfs`

Состояния ядер микропроцессора

Выделяют три состояния энергопотребления ядер микропроцессоров семейства Эльбрус

- C0 – оперативный режим
- C1 – отключение конвейера
- C3 – гашения кэш-памятей и отключение синхросигнала (глубокий сон)

В ядре ОС Эльбрус нет реализации перехода в состояние C3

Перевод вычислительного ядра в СЗ

Алгоритм перевода вычислительного ядра согласно спецификации целевых микропроцессоров Эльбрус содержит следующие шаги:

1) Инвалидация регистров подготовки перехода

Позволяет избежать слабоконтролируемых спекулятивных обращений в память от незавершенных операций подготовки перехода

2) Гашение кэш-памятей

Требуется из-за невозможности поддержания когерентности памяти отключенного ядра

3) Занесение кода отключения ядра в IB без исполнения

Позволяет избежать слабоконтролируемых спекулятивных обращений в память от незавершенной предподкачки прямой ветви кода

4) Отключение ядра

Алгоритм реализован в функции `e2k_enter_idle()`

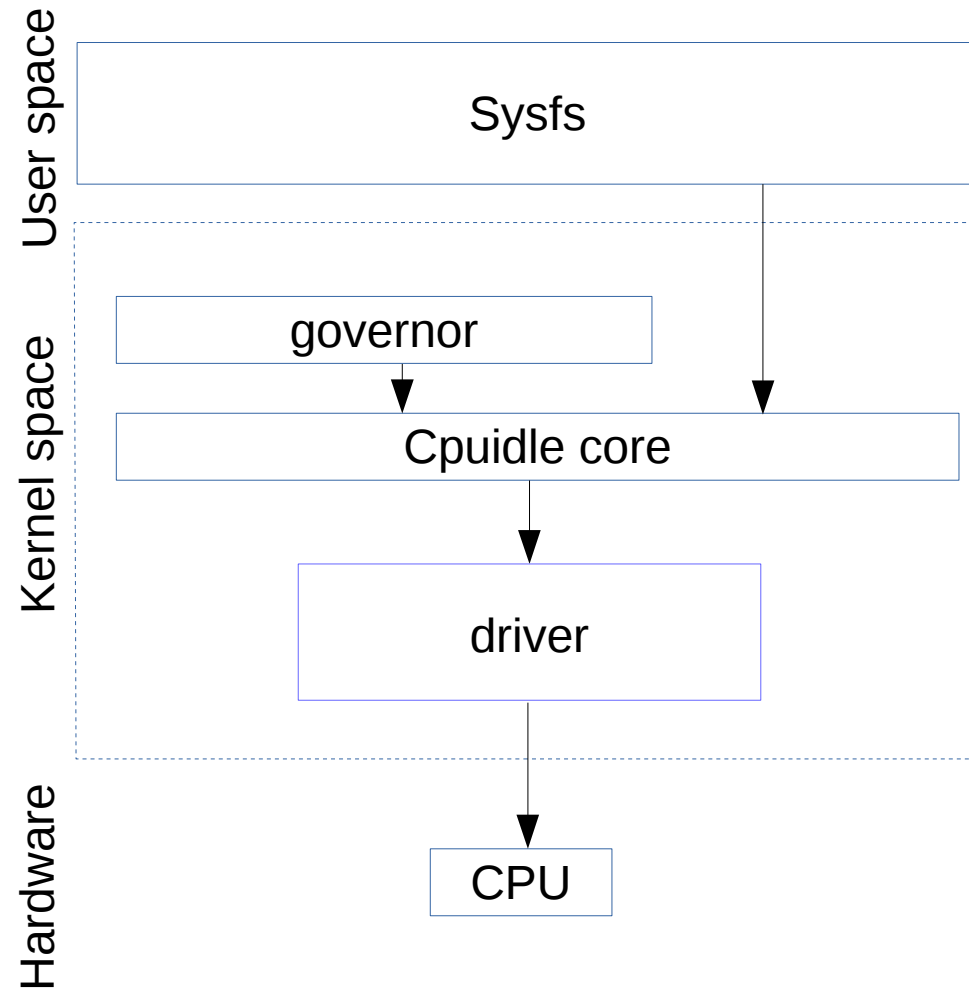
Функция входит в состав интерфейса `e2k_idle_driver`

Встраивание в подсистему cpuidle

Общая схема

Подсистема ядра Linux, переводящая ядра микропроцессора в различные состояния сна в состоянии простоя процессора, которая включает в себя:

- governor — элемент подсистемы cpuidle, определяющий политику перевода вычислительного ядра в оптимальное состояние сна
- driver — элемент подсистемы cpuidle, реализующий перевод вычислительного ядра в состояние сна
- Cpuidle core
- Sysfs интерфейс



Встраивание в подсистему cpuidle

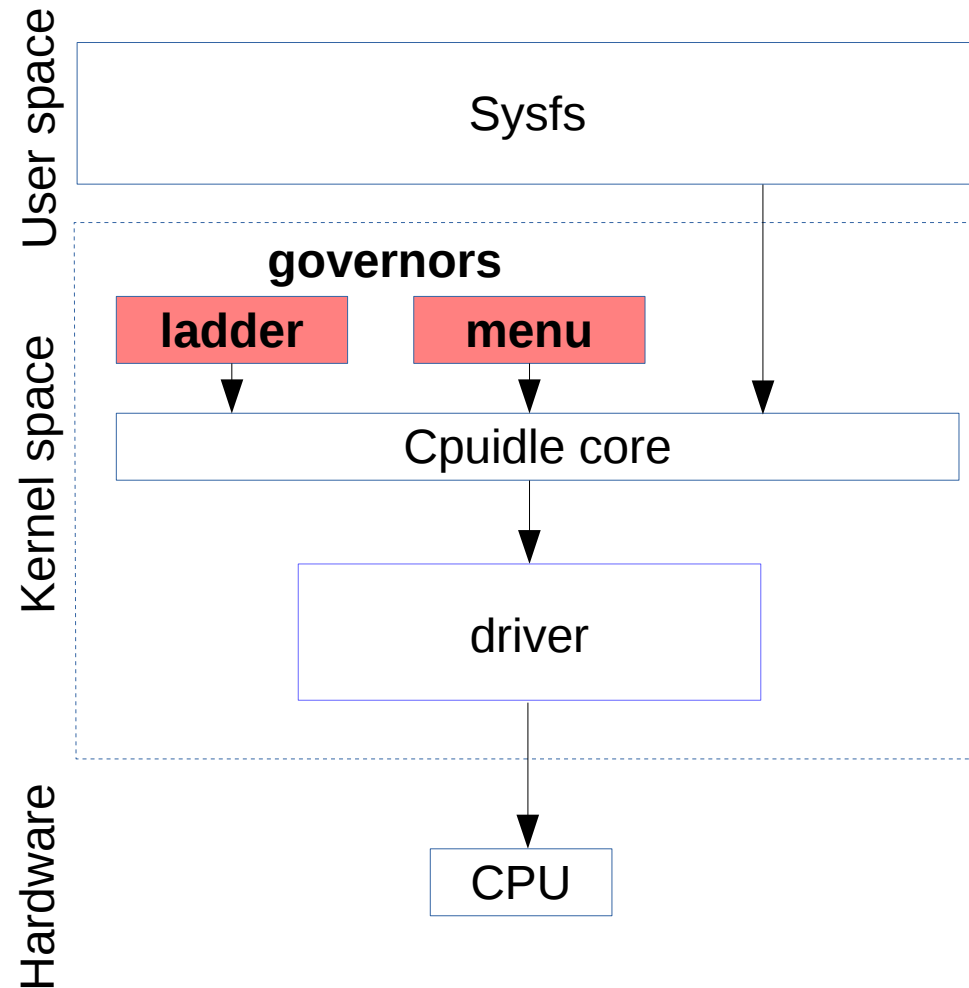
Governor

Содержит функции:

- `select()` — вызывается перед вхождением в состояние сна, возвращает состояние, в которое нужно перевести CPU согласно политике данного governor
- `reflect()` — вызывается после выхода из состояния сна для сбора информации о ядре, находившемся в этом состоянии

`ladder` и `menu` — реализации governor

- `ladder` — перевод в смежное состояние сна
- `menu` — перевод в заданное состояние сна



Встраивание в подсистему cpuidle

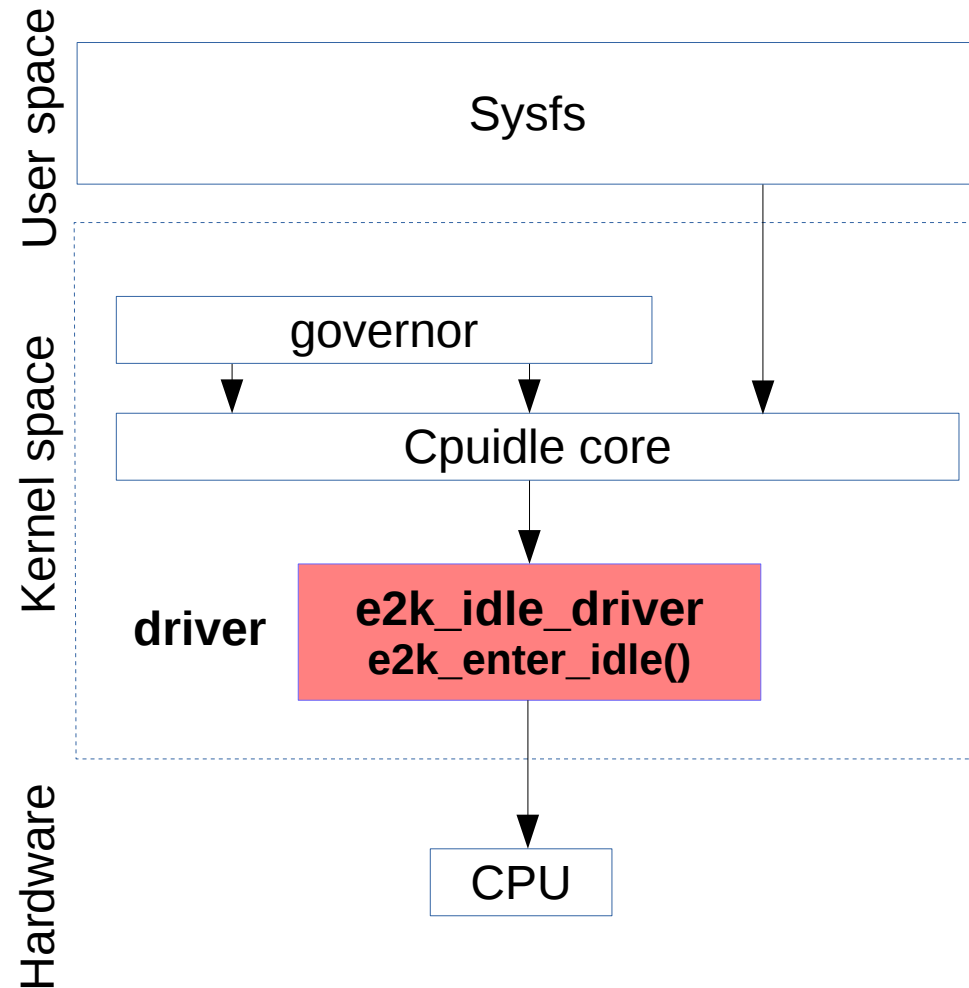
Driver

Переводит ядро в состояние сна.

Параметры перевода описываются структурой `cpuidle_state`:

- `e2k_enter_idle()` — непосредственно переводит ядро в состояние сна
- `name` — имя состояния
- `exit_latency` — задержка пробуждения в этом состоянии
- `target__residency` — предполагаемое время нахождения в этом состоянии

Массив структур передается в `cpuidle core`.



Встраивание в подсистему cpuidle

Пользовательский интерфейс

Интерфейс управления

`/sys/devices/system/cpu/cpuidle/`

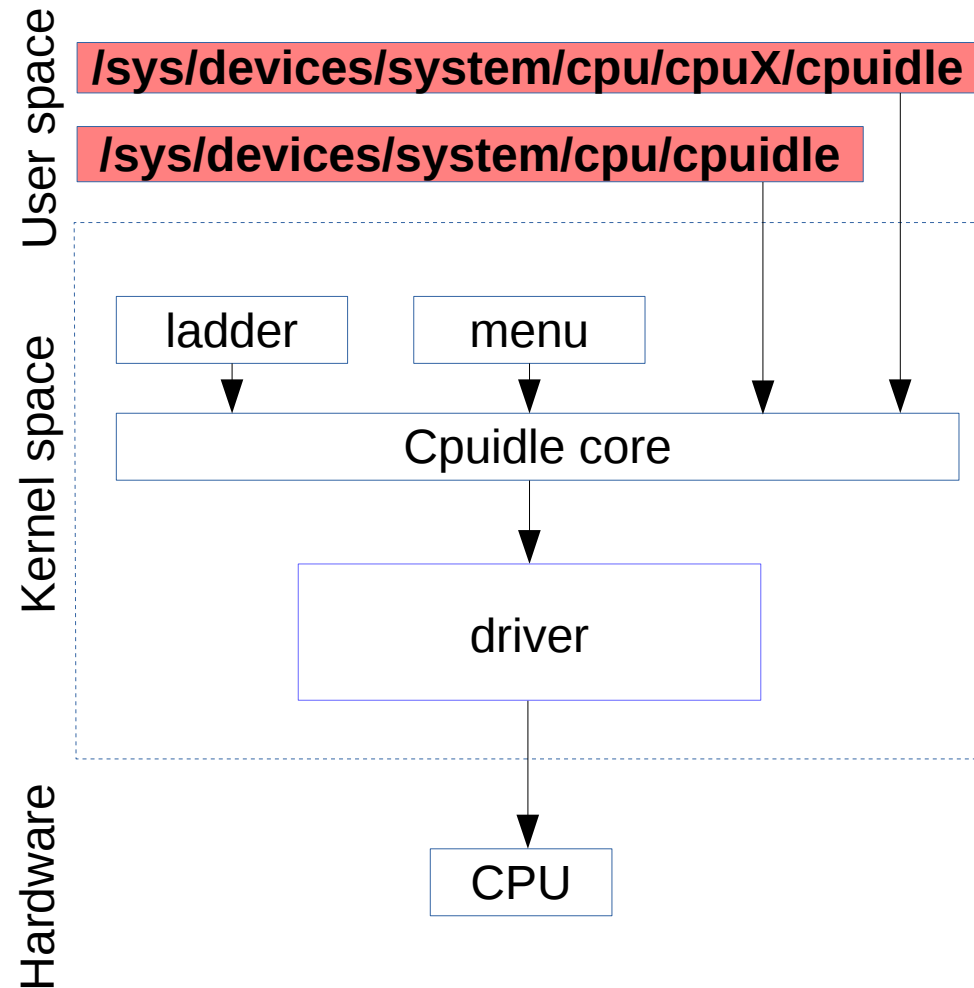
- `current_governor` — перезаписав это значение на `ladder` или `menu`, инициализируется соответствующий `governor`

Интерфейс мониторинга

`/sys/devices/system/cpu/cpuX/cpuidle/stateY/`

где X — номер ядра, Y — номер состояния

- `usage` — число вхождений в это состояние с момента включения модуля
- `time` — время, затраченное на это состояние сна
- `latency` — задержка пробуждения для данного состояния

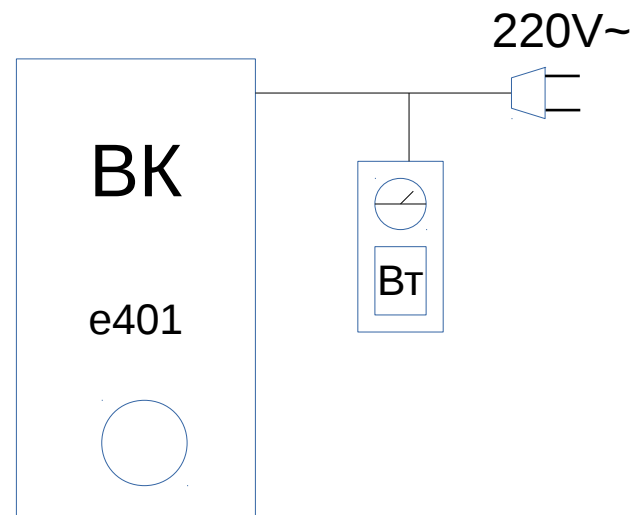


Управление модулем cpuidle

Действия	Команды
Добавление опции в параметры загрузки ядра Linux	<code>cpuidle_sysfs_switch=1</code>
Загрузка модуля	<code>modprobe cpuidle-e2k</code>
Изменение политики выбора состояния сна	<code>echo "ladder" > /sys/devices/system/cpu/cpuidle/current_governor</code> <code>echo "menu" > /sys/devices/system/cpu/cpuidle/current_governor</code>
Отображение количества вхождений в состояния	<code>cat /sys/devices/system/cpu/cpu{X}/cpuidle/state{Y}/usage</code>

Тестирование

Схема стенда



Длительность теста:

30 секунд

Описание аппаратуры:

тест проводился на машине Эльбрус 401

Процесс тестирования: ядра микропроцессора находятся в состоянии простоя, включается подсистема `cruidle` и в качестве `governor` используется `ladder`

Результаты:

Количество вхождений вычислительных ядер в состояния сна:

- C0 ~ 10
- C1 ~ 50
- C3 ~ 1000

БК Эльбрус 401 с отключенной подсистемой `cruidle` потребляет **50.6 Вт**

БК с включенной подсистемой потребляет **41.8 Вт**

Результаты

- Реализован алгоритм перевода вычислительного ядра в состояние глубокого сна
- Алгоритм встроен в подсистему `cruidle`
- Осуществлена возможность управления через интерфейс `sysfs`
- При включении соответствующего модуля наблюдается уменьшение энергопотребления ВК Эльбрус 401 на **17%** в состоянии простоя