

Реализация состояния глубокого сна микропроцессора Эльбрус-4С при помощи средств ядра Linux

Демидов А.А.

Московский физико-технический институт (государственный университет)
«АО» МЦСТ

Общепринятый подход к снижению энергопотребления вычислительных комплексов (ВК) сформулирован в стандарте ACPI [1], согласно которому для данного ВК определяется несколько наборов состояния энергосбережения, где C0-Cn — состояния сна для процессорного ядра, одним из которых является состояние глубокого сна. Алгоритм перевода вычислительного ядра в состояние глубокого сна согласно спецификации целевых микропроцессоров Эльбрус содержит следующие шаги:

1. Инвалидацию регистров подготовки перехода (позволяет избежать слабоконтролируемых спекулятивных обращений в память)
2. Гашение кеш-памятей (требуется из-за невозможности поддержания когерентности памяти отключенного ядра)
3. Занесение кода отключения ядра в IB (Instruction Buffer) без его исполнения (позволяет избежать слабоконтролируемых спекулятивных обращений в память от незавершенной предподкачки прямой ветви кода)
4. Отключение ядра (запись в регистр `st_core`)

Этот алгоритм реализован автором в составе энергосберегающей подсистемы ядра Linux — `cpuidle` [2], переводящей ядра микропроцессора в различные состояния сна в период простоя процессора (рис.1). Подсистема включает в себя архитектурнозависимый компонент `driver`, содержащий функцию `e2k_enter_idle()`, которая непосредственно переводит ядра микропроцессора в различные состояния сна. При решении поставленной задачи эта функция была модифицирована с целью добавления возможности перевода в (ранее не предусматриваемое) глубокое состояние сна. Политика, позволяющая переводить ядро в оптимальное при данном состоянии процессора состояние сна, определяется компонентом `governor`, управляя которым через `sysfs` интерфейс, можно собрать статистику вхождений ядер в различные состояния сна или изменить политику перевода.

Эти возможности были введены в модифицированный модуль ядра Linux `cpuidle-e2k`. Тестирование проводилось на вычислительном комплексе «Эльбрус-401» в течении 30 сек, состоянии простоя ВК (`idle`). При включении модуля наблюдалось снижение энергопотребления вычислительного комплекса на базе микропроцессора Эльбрус-4С с 50.6 Вт до 41.8 Вт (на 17%).

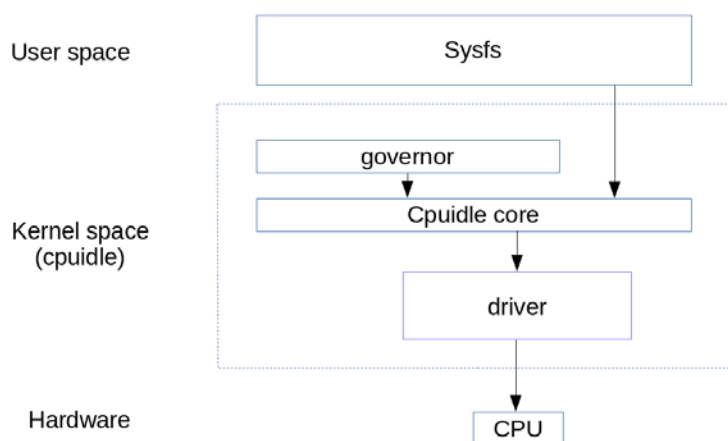


Рис. 1. Структурная схема подсистемы `cpuidle`

Литература

1. Hewlett-Packard Corporation, Intel Corporation, Microsoft Corporation, Phoenix Technologies Ltd, Toshiba Corporation. Advanced Configuration and Power Interface Specification, Revision 4.0a. 2010.
2. *Venkatesh P., Shaohua L., Belay A.*. «cpuidle — Do nothing, efficiently ... » // Proceedings of the Linux Symposium, 2007. V. 2. P. 119.