

Пути реализации стандарта ACPI 4.0 для многопроцессорных вычислительных комплексов на базе процессоров Эльбрус-2S.

Кравцунов Е.М.
Москва, ЗАО "МЦСТ"

Стандарт ACPI (Advanced Configuration and Power Interface Specification)

Последняя актуальная ревизия стандарта ACPI (Advanced Configuration and Power Interface Specification) Revision 4.0a, была выпущена 5-го апреля 2010 года [1]. Целью стандарта является унификация интерфейсов для управления энергопотреблением устройств, находящихся на борту вычислительного комплекса. Стандарт вводит определение глобальных состояний системы, переходов между ними, а также состояний устройств, и центральных процессоров. Определены интерфейсные регистры управления и мониторинга вычислительным комплексом и общая логика, которая должна быть реализована аппаратно для поддержки ACPI.

Для процессоров (ядер процессоров) стандарт определяет 3 набора состояний: Sx, Px, T. C0 – состояние working. В состоянии C0 на процессор (ядро процессора) подается напряжение и синхроимпульс определенной частоты, происходит выполнение инструкций задач пользователя. Состояние C0 подразделяется на несколько состояний Px (performance state), каждое из которых задается парой значений частоты синхроимпульса и напряжения, подаваемое на ядро процессора (w, U). Управление энергопотреблением в состоянии C0 реализуется за счет переключений между состояниями Px. Переключением управляет ядро операционной системы. Критерием для понижения частоты и напряжения питания ядра процессора, то есть, например переключения с состояния (C0; P0) на состояние (C0; P1), является малая загрузка процессора задачами пользователя за определенный интервал времени.

C1 – Cn – это состояния сна процессора, в которых инструкции задач пользователя не выполняются, и синхроимпульс на процессоре может быть отключен.

T (throttling) – состояние, при котором синхроимпульс с процессора снят, а напряжение осталось без изменений по сравнению с последним состоянием (C0; Px). Процессоры переводятся в состояние T для решения задачи пассивного охлаждения в случае перегрева. Переход (C0; P0) -> T не приводит к существенному энергосбережению, так как напряжение питания не изменяется. Однако количество выделяемого тепла в состоянии T уменьшается, так как снижаются потери на активных сопротивлениях, и энергия уходит не в тепло, а тратится на реактивных сопротивлениях, например зарядку – разрядку конденсаторов при утечке.

Процессор Эльбрус-2S и проект вычислительного комплекса

Основные технические характеристики ВК «Эльбрус-2S» таковы:

- количество процессорных ядер внутри процессора (кристалла) - 4;
- количество процессоров в составе комплекса – 4;
- диапазон тактовой частоты процессора – от 800 МГц до 1000 МГц;

Межпроцессорное взаимодействие реализовано с помощью устройств Link, входящих в состав северного моста. Северный мост является частью кристалла. В состав северного моста входят 4 устройства LAPIC (Local APIC), по одному на каждое ядро. Устройства LAPIC обеспечивают доставку внешних, внутренних и межпроцессорных прерываний процессорным ядрам.

Внешние прерывания устройства LAPIC получают от контроллера внешних прерываний IOAPIC, являющегося частью южного моста. Для взаимодействия с южным мостом используется устройство I/O-LINK, работающее на одном из четырех процессоров ВК. Прерывания поступают в IOAPIC от внешних устройств, находящихся на одной из шин, поддерживаемых южным мостом. Так прерывания от устройств мониторинга и регулировки питания (микросхема LTC2970IUFD), регулировки частоты (микросхема MPC92432FA) и датчиков температуры (LM95231) поступают в IOAPIC из контроллера шины I2C, клиентами которой являются перечисленные устройства.

Маршрутизация внешних прерываний на ядра процессоров осуществляется устройством IOAPIC с помощью таблицы Redirection Table устройства APIC_TABLE. Таблица является программно доступной для ядра операционной системы.

На борту каждого кристалла находится один встроенный датчик температуры, позволяющий отслеживать температуру кристалла в точке расположения датчика через контроллер LM95231. Наличие датчика позволяет отследить среднюю температуру кристалла, которая может использоваться в качестве критерия для того чтобы стартовать средствами ОС пассивное охлаждение кристалла.

Для решения этой задачи на процессоре предусмотрена возможность полного отключения синхроимпульса на процессорных ядрах. Отключение производится с помощью регистра PWR_MGR0. При этом частота снимается только с ядер, а на северный мост и коммутатор взаимодействия с памятью частота подается. То есть фактически осуществляется перевод процессора из состояния C0 (working) в состояние T (throttling). Такой переход описан в стандарте.

Пути реализации стандарта

Реализация полной совместимости со стандартом ACPI на вычислительном комплексе Эльбрус-2S невозможна, так как в процессорах Эльбрус-2S отсутствует аппаратная поддержка отключения питания отдельных устройств, интерфейсных регистров и генерации прерываний по событиям простоя и перегрева процессора.

Однако на вычислительном комплексе Эльбрус-2S возможна экспериментальная реализация части состояний и методов управления энергопотреблением ACPI путем доработки только ядра операционной системы linux-2.6.33.1, перенесенного на архитектуру E2S. Благодаря присутствию на борту вычислительного комплекса устройств управления частотой синхроимпульса и питания, заложенной в процессор возможности полного отключения синхроимпульса и программной доступности устройства IOAPIC реализуемы наборы состояний Sx, Px, T процессора, переходы между этими состояниями. Переходы реализуются с помощью стандартных методов _PTS, _GTS, _BFS, которые в свою очередь реализованы путем введения ряда архитектурно-зависимых конструкций в архитектурно-независимую часть реализации ACPI в linux-2.6.33.1 [2], [3].

Литература

- [1] Advanced Configuration and Power Interface Specification, Hewlett-Packard Corporation, Intel Corporation, Microsoft Corporation, Phoenix Technologies Ltd, Toshiba Corporation, Revision 4.0a, April 5, 2010
- [2] ACPI in Linux: Architecture, Advances, and Challenges, Len Brown, Anil Keshavamurthy, David Shaohua Li, Robert Moore, Venkatesh Pallipadi, Luming Yu, Proceedings of the Linux Symposium 2005, p. 51
- [3] The Ondemand Governor: Past, Present, and Future, Venkatesh Pallipadi, Alexey Starikovskiy, Proceedings of the Linux Symposium 2006, p. 657