

# Intel Xeon D-1500: неочевидные факты

27 Мар 2016

[Илья Манусов](#)



Реализация центрального процессора с полнофункциональным набором системной логики в одной микросхеме давно под силу современной микроэлектронике. Но разработчики High-End платформ в силу ряда объективных факторов, таких как рассеиваемая мощность, а также с учетом принципов декомпозиции сложных устройств, предпочитали конструктивно разделять центральный процессор и подсистему ввода-вывода. Эволюция линейки CPU Intel Xeon нарушила эту традицию.

Процессор состоит из десятков подсистем, простое перечисление которых превысит размер статьи, поэтому остановимся только на нескольких моментах, обычно не упоминаемых в презентациях, но показавшихся нам важными с точки зрения исследования архитектуры платформ.

## Холодный расчет

Посмотрев на значения **TDP** или *Thermal Design Power* (Рис.1) нетрудно понять, что в серверном сегменте ожидается (вернее уже идет) очередной этап миниатюризации.

**Table 4-2. VCCIN Supply Current (Iccin\_max and Iccin\_tdc) Specification**

Segment/SKU	TDP	ICC_MAX @ VCCIN	ICCIN_TDC @VCCIN	Pmax <sup>2</sup> (SoC)	Notes <sup>1</sup>
D-1520, D-1521, D-1531, D-1540, D-1541 D-1548, D-1557, D-1559, D-1577, D-1571	45W	55A	26A	90W	
D-1518, D-1527, D-1528, D-1537, D-1539	35W	43A	21A	70W	
D1508, D-1517, D1591	25W	30A	15A	50W	
D1507	20W	24A	12A	40W	

**Рис. 1.** Характеристики потребляемой мощности для основного источника питания процессора

Напомним, один из принципов, позволяющих строить компактные и эффективные системы, состоит в определении потребляемой мощности в виде двух параметров:

1. Максимальная потребляемая мощность (Pmax) важна при учете предельных электрических характеристик системы электропитания.
2. Усредненная мощность TDP (Thermal Design Power), используется для расчета характеристик теплоотвода.

Второй параметр существенно меньше первого, именно поэтому радиатор может быть компактным, а иногда и просто отсутствовать. Предполагается, что потребление мощности, превышающей TDP в течение длительного времени, достаточного для перегрева радиатора, крайне маловероятно. Но если это и произойдет, а теплоотвод, спроектированный на основании TDP (но не Pmax!) не справится со своими обязанностями, в действие вступит *thermal throttling*. Понижение тактовой

частоты, а точнее, частичный пропуск тактов, позволит избежать аварийной ситуации не прерывая работу сервера.



Рис. 2. Одноплатный компьютер на базе Intel Xeon-D от [компании X-es](#), анонсированный 18 марта с.г., используется [без активного охлаждения](#)

## Интегрированный ввод-вывод

Как известно, *северный мост* системной логики, включающий контроллер оперативной памяти и «широкие» линки шины PCI Express, стал компонентом процессора уже достаточно давно, поэтому акцентируем внимание на ресурсах *южного моста* или подсистеме ввода-вывода, также именуемой *PCH (Platform Controller Hub)*, перемещение которого в процессор для серверов, является событием сравнительно недавним.

### PCI Devices and Functions

Bus:Device:Function	Function Description
Bus 0:Device 31:Function 0	LPC Controller <sup>1</sup>
<b>Bus 0:Device 31:Function 2</b>	<b>SATA Controller #1</b>
Bus 0:Device 31:Function 3	SMBus Controller
Bus 0:Device 31:Function 5	SATA Controller #2 <sup>2</sup>
Bus 0:Device 31:Function 6	Thermal Subsystem
Bus 0:Device 29:Function 0 <sup>3</sup>	USB EHCI Controller #1
Bus 0:Device 28:Function 0	PCI Express* Port 1
Bus 0:Device 28:Function 1	PCI Express Port 2
Bus 0:Device 28:Function 2	PCI Express Port 3
Bus 0:Device 28:Function 3	PCI Express Port 4
Bus 0:Device 28:Function 4	PCI Express Port 5
Bus 0:Device 28:Function 5	PCI Express Port 6
Bus 0:Device 28:Function 6	PCI Express Port 7
Bus 0:Device 28:Function 7	PCI Express Port 8
Bus 0:Device 25:Function 0	Gigabit Ethernet Controller
Bus 0:Device 22:Function 0	Intel® Management Engine Interface #1
Bus 0:Device 22:Function 1	Intel Management Engine Interface #2
Bus 0:Device 22:Function 2	IDE-R
Bus 0:Device 22:Function 3	KT
Bus 0:Device 20:Function 0	xHCI Controller

Рис. 3. Список ресурсов PCH, перемещенных в состав микросхемы CPU

Реализация подсистемы ввода-вывода, ставшей частью процессора, достаточно типична для современных платформ Intel и включает в себя такие компоненты:

1. Дисктовую подсистему, конфигурируемую для работы в режимах: эмуляции *IDE*, контроллера *AHCI* или *RAID*. Производительность и совместимость на выбор. Наличие узла *DMA Engine* и технология *IntelQuickData* позволяют создавать фактически аппаратную реализацию RAID-контроллера средствами центрального процессора.
2. Интерфейс шины *Low Pin Count*, применяемой для связи с медленной периферией.
3. Интерфейс шины *System Management Bus*, обеспечивающей дополнительный канал взаимодействия между ресурсами платформы.
4. Подсистема термоконтроля.
5. Подсистема USB, содержащая как контроллер USB2 EHCI, необходимый для совместимости с устаревшими ОС, так и современный контроллер USB3 xHCI, полнофункциональный с точки зрения обеспечения всех скоростных режимов универсальной последовательной шины.
6. Гигабитный контроллер Ethernet.
7. Автономный интеллект платформы в виде подсистемы *Intel Management Engine Interface*.
8. Подсистема, обеспечивающая дистанционное управление платформой, включая функции *IDE Redirection* и *Keyboard and Text Redirection* соответственно, для удаленной загрузки и создания консоли. Для этой подсистемы используется последовательный интерфейс в виде *memory-mapped UART*.
9. Многочисленные интерфейсы *PCI Express* с реконфигурируемыми топологией и скоростными режимами.

**KT IO/Memory Mapped Device Register Address Map**

Address Offset	Register Symbol	Register Name	Default Value	Attribute
0h	KTRxBR	KT Receive Buffer Register	00h	RO
0h	KTTHR	KT Transmit Holding Register	00h	WO
0h	KTDLLR	KT Divisor Latch LSB Register	00h	R/W
1h	KTIER	KT Interrupt Enable register	00h	R/W, RO
1h	KTDLMR	KT Divisor Latch MSB Register	00h	R/W
2h	KTIIR	KT Interrupt Identification register	01h	RO
2h	KTFCR	KT FIFO Control register	00h	WO
3h	KTLCR	KT Line Control register	03h	R/W
4h	KTMCR	KT Modem Control register	00h	RO, R/W
5h	KTLSR	KT Line Status register	00h	RO
6h	KTMSR	KT Modem Status register	00h	RO

**Рис. 4.** Регистровая модель 82х50 совместимого последовательного интерфейса не утратила своей актуальности и успешно применяется в составе подсистемы *KT (Keyboard and Text Redirection)* современных платформ

## POST-диагностика процессора Xeon-D

Идея [оптической передачи POST-кодов](#) с использованием модуляции яркости свечения *Power Led*, на первый взгляд кажется слишком необычной. Но если объективно взвесить соотношение минимальных аппаратных затрат и максимальной получаемой функциональности, то можно понять, почему полупроводниковый гигант продвигает именно ее уже в нескольких поколениях системной логики как рекомендуемый метод отладки и диагностики платформ.

### Serial Post over GPIO Reference Circuit

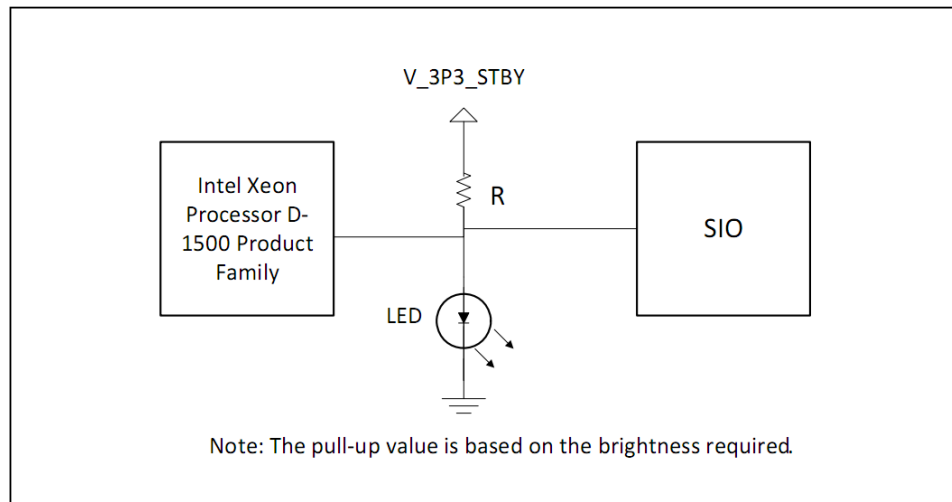


Рис. 5. Подключение светодиода Power Led: при выводе диагностических кодов, сигнал формируется системной логикой, входящей в состав процессора

Как видим, преемственность коснулась не только преимуществ такой схемы, но и ее недостатков. Пробитый выход микросхемы SIO (ведь мы диагностируем *неисправную* платформу!) не позволит выполнить вывод [POST-кодов](#) в диагностический порт светодиодного индикатора.

## Подсистема LAN

Гигабитный LAN-контроллер, скромно упомянутый выше, является всего лишь дополнением к значительно более мощной подсистеме, образованной двумя 10-гигабитными высокоинтеллектуальными контроллерами, поддерживающими технологию аппаратной [виртуализации ввода-вывода SR-IOV](#). Одно из объяснений такой функциональной насыщенности, а именно, сосуществования двух видов контроллеров, может состоять в планируемой опциональности наличия производительных NIC.

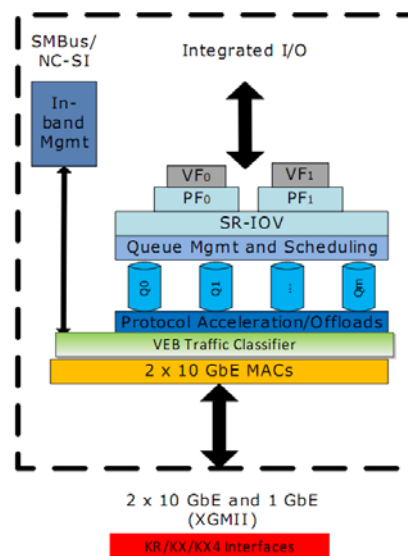


Рис. 6. Два 10-гигабитных сетевых контроллера поддерживают аппаратную виртуализацию SR-IOV: для взаимодействия со средой передачи используется интерфейс XGMII (10 Gigabit Media Independent Interface)

В соответствии с принципами **SR-IOV**, контроллеры представляются в адресном пространстве как набор виртуальных функций, отдельно передаваемых в распоряжение гостевых ОС. А семейство технологий Protocol Acceleration / Offloads позволяет освободить центральный процессор от ряда рутинных операций сетевого ввода-вывода.



Рис. 7. Одна из первых реализаций платформы на Intel Xeon-D от [компании MEN Micro](#): кроме разъемов для двух 10G-подключений на плате [G25A](#) в наличии имеется также гигабитный интерфейс управления

## Вместо послесловия: о первоисточниках

---

Информация подготовлена на основе открытой документации *Intel Xeon Processor D-1500 Product Family Datasheet*. Документ состоит из четырех томов:

- *Volume 1 Integrated Platform Controller Hub* описывает периферийные контроллеры, схемы системной поддержки и другие ресурсы, традиционно расположенные в составе *южного моста* системной логики, являющиеся частью микросхемы процессора в системах SoC.
- *Volume 2 Registers* содержит описание ресурсов *северного моста* системной логики, также входящего в состав процессора. В центре внимания IMC (Integrated Memory Controller), принципы формирования конфигурационного пространства, аппаратная виртуализация VT-d, а также DMA Engine, пришедший на смену Legacy-контроллеру Intel 8237.
- *Volume 3 Electrical* определяет процессор как электрофизическую систему, нормируя набор питающих напряжений, потребляемые токи и множество сигнальных цепей, обеспечивающих интерфейс микросхемы CPU с внешним миром. Для каждого сигнала описаны электрические параметры и функциональное назначение.

- *Volume 4 LAN Controller* определяет архитектуру контроллера локальной сети, реализованного в составе процессора. Показательно, что одной этой подсистеме посвящен отдельный том и он самый большой (758 страниц). Здесь рассмотрено достаточно много связанных тем: интерфейс с Flash-памятью конфигурации сетевого адаптера, поддержка технологии виртуализации SR-IOV для NIC.

Четырехтомник ориентирован на разработчиков аппаратных платформ и системного программного обеспечения, в частности *UEFI firmware* и низкоуровневых драйверов. Отметим, что набор инструкций x86 и его многочисленные функциональные расширения описываются другим многотомным документом — *Software Developers Manual*, ориентированным на прикладных программистов. Эта информация в описания модельных рядов процессоров не включается.