

Для цитирования: Билялетдинов И. Е., Ометов А. Е., Тимин Л. С. Оптимизация параметров высокоскоростных каналов процессора с целью повышения отказоустойчивости вычислительного комплекса // Вопросы радиоэлектроники. 2018. № 2. С. 87–92. УДК 004.052.3

И. Е. Билялетдинов^{1, 2}, А. Е. Ометов³, Л. С. Тимин^{1, 3}

¹ ПАО «ИНЭУМ им И. С. Брука», ² МТУ (МИРЭА), ³ АО «МЦСТ»

ОПТИМИЗАЦИЯ ПАРАМЕТРОВ ВЫСОКОСКОРОСТНЫХ КАНАЛОВ ПРОЦЕССОРА С ЦЕЛЬЮ ПОВЫШЕНИЯ ОТКАЗОУСТОЙЧИВОСТИ ВЫЧИСЛИТЕЛЬНОГО КОМПЛЕКСА

В современных вычислительных средствах семейства «Эльбрус» используются высокоскоростные каналы передачи данных – каналы памяти DDR3, каналы межпроцессорного взаимодействия и ввода-вывода, построенные на базе PCI Express. Их физические уровни имеют большое число настроек, таких как номинал терминирующего резистора, preemphasis и equalization, амплитуда выходного сигнала и т.д. Все они влияют на работоспособность канала и целостность сигналов. Для повышения отказоустойчивости вычислительного комплекса требуется найти наиболее оптимальный вариант настройки для физического уровня высокоскоростных каналов. Ввиду большого количества возможных настроек необходимо иметь возможность произвести оценку за наименьшее время и с минимальным количеством дополнительного оборудования. Авторы данной статьи разработали методику определения качества настроек, использующую внутренние механизмы физического уровня каналов.

Ключевые слова: DDR3, PCI Express, внутреннее самотестирование, глазковые диаграммы.

Введение

В современных вычислительных средствах семейства «Эльбрус» используются высокоскоростные каналы передачи данных – каналы памяти DDR3, а также каналы межпроцессорных связей и ввода-вывода, реализованные на базе физического уровня PCI Express.

Физический уровень каналов памяти имеет большое число различных настроек и параметров – это и сопротивления (например, выходные сопротивления сигналов), и настройки временных характеристик, и параметры внутренних блоков и механизмов.

Один из наиболее применяемых способов, позволяющих выяснить качество выбранных настроек, заключается в том, чтобы запустить длительный прогон на нескольких вычислительных комплексах. Другой – в том, чтобы, используя очень точный осциллограф, следить за изменениями сигналов и их параметрами. Однако при реализации в оптимальном варианте оба они требуют большого количества дополнительного оборудования и временных затрат. Если же проводить настройки неоптимальным образом, то с большой вероятностью может снизиться надежность вычислительного комплекса (возрастает вероятность сбоя) или скорость его работы (чтобы обеспечить надежность, снижается частота).

Для решения задачи необходимо получить численную характеристику, отражающую работоспособность канала настройки. Такой характеристикой, которую можно получить быстро, является ширина области работоспособности на глазковой диаграмме.

Глазковые диаграммы сигналов

Глазковая диаграмма – это суммарный вид всех битовых периодов измеряемого сигнала, наложенных друг на друга [1]. В теории она должна выглядеть как правильный многоугольник (рис. 1а), однако в реальности его границы размываются (рис. 1б).

Размытие возникает из-за нестабильности задающего генератора, изменений параметров линий

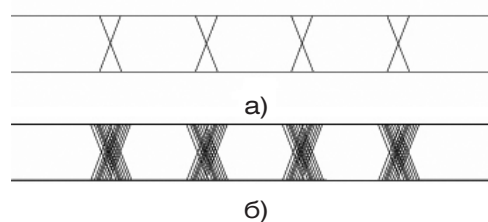


Рисунок 1. Глазковая диаграмма: а – идеальное представление; б – представление с учетом фазового дрожания сигналов

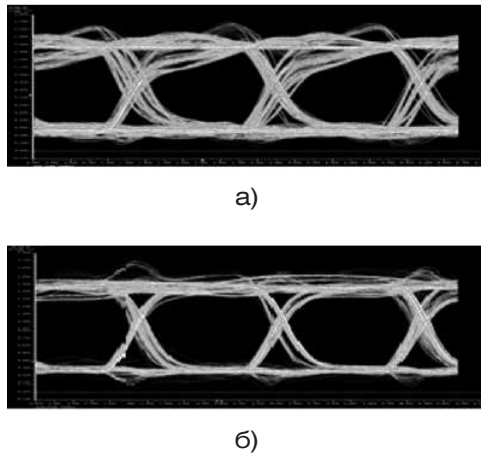


Рисунок 2. Глазковые диаграммы для неоптимального (а) и оптимального (б) вариантов настройки

(например, температуры), различной скорости распространения частотных составляющих сигнала, отражений и помех на линии. Все эти эффекты характеризуются как джиттер, или фазовое дрожание цифрового сигнала.

Глазковая диаграмма позволяет быстро и наглядно оценить качество сигнала и выявить ошибки, связанные с помехами и искажениями сигналов. Как показано на рис. 2, при правильном подборе параметров выходных сопротивлений и терминирующих резисторов величина фазового дрожания сигнала уменьшается.

Одной из характеристик глазковой диаграммы является ширина «глаза данных» – области диаграммы с наибольшим отношением величин сигнала и шума. Пример, приведенный в [1], показан на рис. 3.

Использование глазковой диаграммы для определения работоспособности системы памяти

Для определения ширины глаза данных применительно к системе памяти DDR3 без помощи осциллографа авторами был применен метод,

основанный на перемещении точки захвата данных. Найти эту границу позволяет тот факт, что при сдвиге за границы глаза система становится неработоспособной. Захват данных в системе памяти DDR3 происходит по фронту строба DQS. Соответственно, границы глаза можно найти, сдвигая строб относительно данных. На физическом уровне канала памяти функционируют управляемые блоки задержек сигналов строба DQS, обозначаемые как BDL и LCDL. Начальные значения задержек выставляются во время процедуры инициализации физического уровня. Это выполняется механизмами подстройки [2], однако сами значения также доступны для чтения и записи пользователем. Управлять задержками можно отдельно в следующих режимах: для каждого байта, для режима чтения и режима записи (при процессорном чтении данными и стробом управляет модуль памяти, при записи – процессор). Алгоритм определения окна работоспособности предполагает следующую последовательность действий:

1. Установка начальной задержки.
2. Увеличение задержки.
3. Запуск теста, контролирующего работоспособность.
4. Увеличение задержки и повторный запуск при успешном выполнении теста.
5. Завершение теста с ошибкой означает, что строб вышел из зоны работоспособности. Последняя рабочая задержка фиксируется (запоминается) как правая граница окна работоспособности.
6. Установка начальной задержки.
7. Уменьшение задержки с запуском теста работоспособности.
8. Если тест прошел успешно, то – дальнейшее уменьшение задержки и запуск теста.
9. Завершение теста с ошибкой означает, что строб вышел из зоны работоспособности. Последняя рабочая задержка фиксируется (запоминается) как левая граница окна работоспособности.
10. Установка начальной задержки.

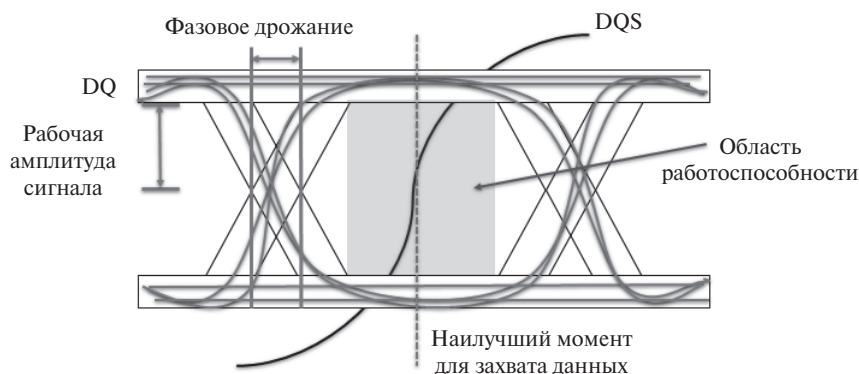


Рисунок 3. Область работоспособности на глазковой диаграмме

Отдельное управление задержкой позволяет определять окна работоспособности для каждого байта и каждого режима отдельно, что, в свою очередь, дает оценить возможный технологический разброс характеристик каналов памяти.

Тест на работоспособность

В качестве теста на работоспособность может выступать широкий набор тестов, однако для уменьшения времени исполнения разумно использовать внутренние механизмы физического уровня памяти. Таких механизмов два: встроенное самотестирование (BIST) и встроенный модуль управления (DCU) [2].

Встроенное самотестирование позволяет записывать псевдслучайный код в память по программируемым адресам. Определенный недостаток состоит в том, что одновременно можно тестировать только один байт памяти. С одной стороны, это не так плохо, т.к. позволяет снимать характеристики для каждого байта, с другой стороны, исключается влияние соседних байтов друг на друга и уменьшаются шумы на питании, которые могут повлиять на работоспособность вычислительного комплекса в дальнейшем.

Встроенный модуль управления позволяет провести тестирование в обход контроллера памяти. Он может исполнять все внутренние команды памяти, включая команды управления и настройки модуля памяти, но есть ограничения, обусловленные

размерами кэшей управления: всего можно исполнить 16 команд. К недостаткам следует отнести относительно длительное заполнение кэшей и трудность программирования на языке памяти (например, простая запись одного значения занимает 4 команды из 16 доступных). Преимущество командного модуля относительно встроенного механизма самотестирования только одно – возможность оперировать со всеми байтами одновременно.

Результаты исследования канала памяти

Границы области работоспособности канала памяти оценивались с помощью механизма встроенного самотестирования для каждого байта отдельно. Были сняты характеристики окон работоспособности применительно к нескольким вычислительным комплексам на базе различных микропроцессоров семейства «Эльбрус». По итогам исследований было отмечено большее влияние настроек памяти на комплексы с несколькими слотами памяти на один канал, а также большая зависимость относительных значений задержек (в единицах отсчетов LCDL) от величины напряжения питания физического уровня памяти и температуры. В связи с этим в физическом уровне памяти был включен механизм внутренней подстройки по напряжению и температуре VT_drift. Результаты исследований для одного из каналов вычислительного комплекса «Эльбрус 4.4» представлены на рис. 4.

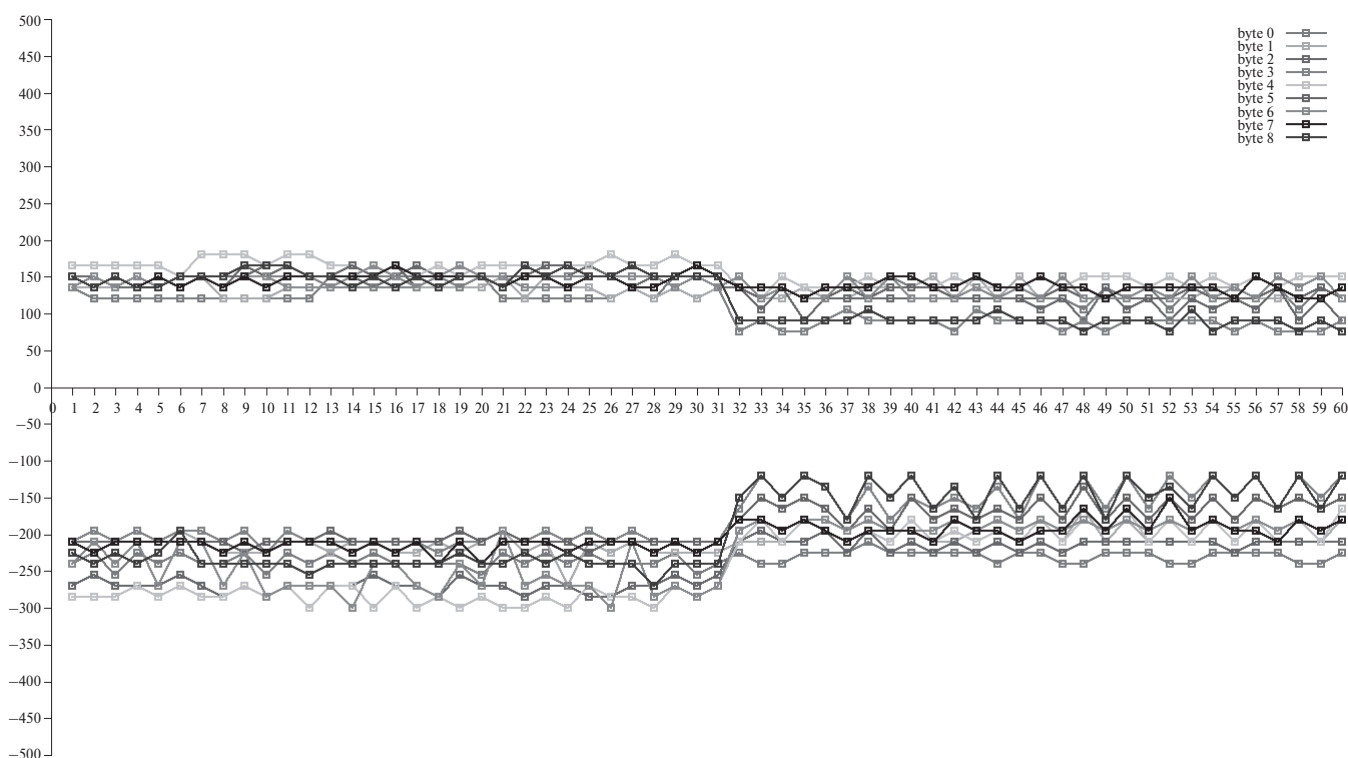


Рисунок 4. Результаты тестирования для различных вариантов настройки DDR3: по оси абсцисс – номера в списке настроек; по оси ординат – граница окна работоспособности в пикосекундах

Точки на графике в отрицательной области соответствуют левой границе области работоспособности, в положительной области – правой границе. Точки построены для каждого из девяти байтов данных отдельно. При оптимальных вариантах настройки ширина области работоспособности увеличивается до 50%.

Настройка канала PCI Express

Связи в вычислительных комплексах на базе микропроцессоров «Эльбрус» реализованы на основе физического уровня PCI Express. Микропроцессоры связаны по принципу «каждый с каждым» четырьмя соединениями, которые состоят из четырех двусторонних линий. Пропускная способность этих линий достигает 6 Гб/с, поэтому очень важно правильно подобрать настройки физического уровня так, чтобы сигнал не затухал (линии на серверной плате достаточно длинные) и не искажался [2, 3].

Межпроцессорные связи имеют следующий набор параметров, правильная настройка которых необходима для повышения отказоустойчивости комплекса:

- Выходная амплитуда передатчика.
- Величина pre-emphasis и equalization.
- Пропускная способность канала.

Чтобы проверить работоспособность набора параметров, необходимо провести тестирование линий. Одним из известных решений является стандарт IEEE1149.6, который позволяет тестировать высокоскоростные интерфейсы с помощью JTAG-интерфейса [4]. Однако у данного стандарта есть два существенных недостатка: во-первых, его необходимо внедрить и поддерживать в дополнение к распространенному стандарту IEEE1149.1, во-вторых, представленный в стандарте метод не позволяет тестировать интерфейс на рабочей частоте, т.к. у JTAG она не превосходит 10–30 МГц [4].

В силу этих причин для проверки набора параметров следует использовать встроенный модуль самотестирования PCIe (рис. 5). Он состоит из двух частей: генератора тестовых данных и блока сравнения. Первый встроен в передатчик и может посылать кодовые последовательности типа PRBS7, PRBS15, PRBS23 и PRBS31, а также пользовательские данные. Блок сравнения размещен в приемнике и может синхронизоваться с генератором, принимать значения и подсчитывать количество ошибок. Управление этим модулем, изменение параметров связей, их инициализация и повторная инициализация производятся с помощью управляющих регистров, которые доступны через JTAG-интерфейс.

Установлен следующий порядок тестирования совокупности параметров:

1. Выбор набора настроек, инициализация физического уровня передатчика и приемника с их применением.
2. Выбор режима работы генератора на передатчике, включение компаратора на приемнике в аналогичном режиме.
3. Синхронизация механизма сравнения с генератором, после чего – сброс счетчика ошибок и начало тестирования.
4. Чтение значения счетчика ошибок каждой линии.
5. Выключение генератора и механизма сравнения.
6. Анализ полученных данных, вывод о работоспособности настроек.

Подобным образом проверяется каждый набор. В результате на основании данных счетчика ошибок можно определить окно работоспособности межпроцессорной связи, проанализировать работу каждой линии и выбрать наилучшую точку для повышения отказоустойчивости системы.

На рис. 6 представлены результаты подбора настроек для вычислительного модуля на базе

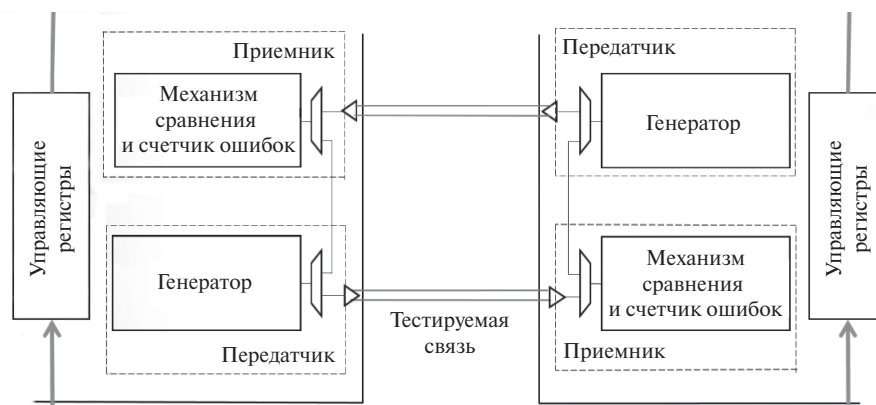


Рисунок 5. Принципиальная схема модуля самотестирования PCIe

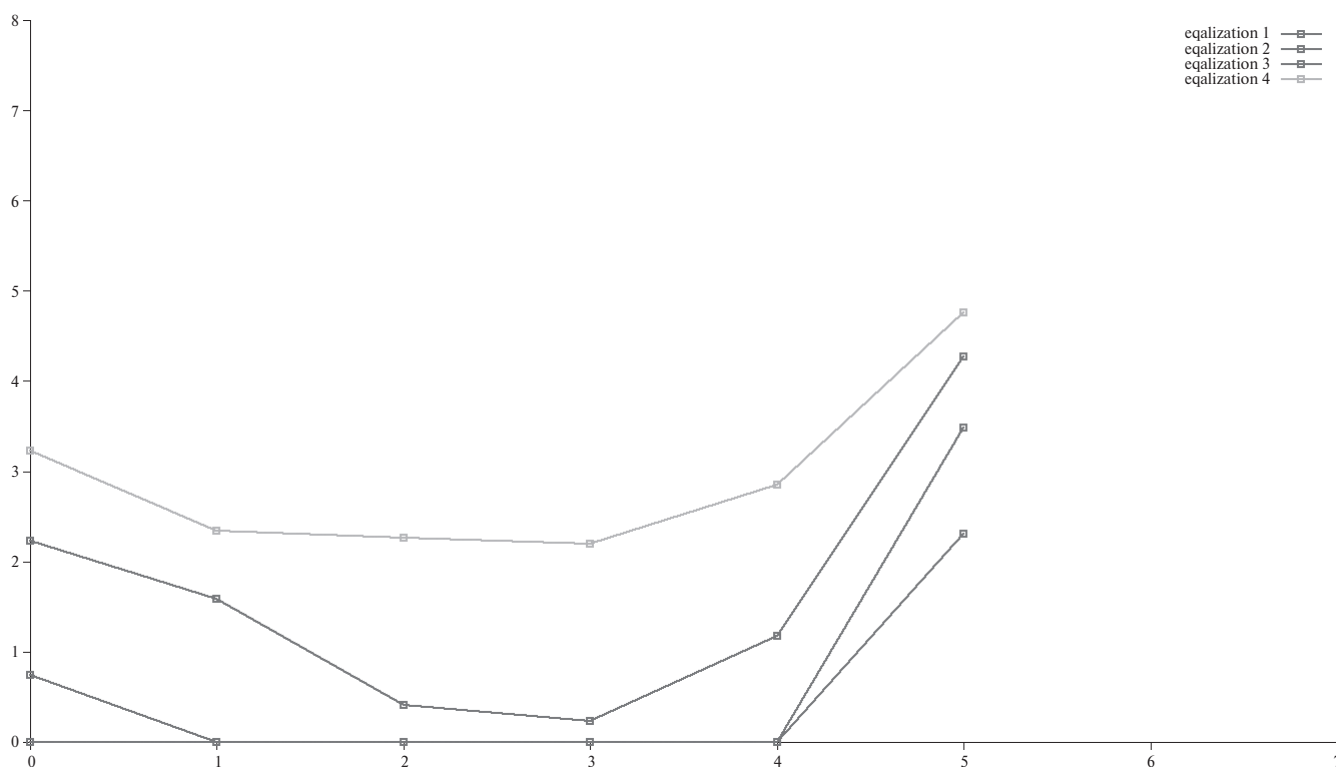


Рисунок 6. Результаты тестирования для различных вариантов настройки PCIe: по оси абсцисс – номера в списке настроек; по оси ординат – десятичный логарифм числа ошибок при данной настройке

микропроцессора «Эльбрус-8С». С целью большей наглядности для различных настроек equalization построены отдельные графики.

Заключение

В данной работе представлены разработанные и исследованные авторами методы, позволяющие оценить качество настроек и за короткое время с использованием минимального количества дополнительного оборудования подобрать оптимальные. Было проведено большое число экспериментов

с различными вариантами настроек высокоскоростных каналов процессора для различных вычислительных комплексов, в ходе которых выяснилось, что для повышения отказоустойчивости подбор параметров следует проводить для каждого вычислительного комплекса. Это связано с разбросом характеристик плат при производстве и монтаже, а также с условиями эксплуатации. В связи с этим быстрый и простой в использовании способ оценки настроек чрезвычайно важен и имеет реальные перспективы применения.

СПИСОК ЛИТЕРАТУРЫ

1. Foster G. Anatomy of an Eye Diagram – a Primer. SyntheSys Research, Inc., 2004, 9 p.
2. Wang L.-T., Stroud C. E., Toubia N. A. System-on-Chip Test Architectures: nanometer design for testability. Burlington, Morgan Kaufmann Publishers, 2008, 856 p.
3. Wang L.-T., Wu C.-W., Wen X. VLSI Test Principles and Architectures: Design for Testability. San Francisco, Morgan Kaufmann Publishers, 2006, 777 p.
4. IEEE Std 1149.1–2001: IEEE Standard Test Access Port and Boundary-Scan Architecture. New York, Institute of Electrical and Electronics Engineers, 2001, 208 p.

ИНФОРМАЦИЯ ОБ АВТОРАХ

Билялетдинов Илья Евгеньевич, аспирант, МТУ (МИРЭА); инженер-программист, ПАО «ИНЭУМ им И.С. Брука», 119334, Москва, ул. Вавилова, д. 24, тел.: 8 (495) 363-95-03, e-mail: ilia.e.bilyaletdinov@mcst.ru.

Ометов Александр Евгеньевич, инженер-программист, АО «МЦСТ», 119334, Москва, ул. Вавилова, д. 24, тел.: 8 (495) 363-95-03, e-mail: alexander.e.ometov@mcst.ru.

Тимин Леонид Сергеевич, начальник сектора, ПАО «ИНЭУМ им И.С. Брука», АО «МЦСТ», 119334, Москва, ул. Вавилова, д. 24, тел.: 8 (495) 363-95-03, e-mail: le0@mcst.ru.

For citation: Bilyaletdinov I.E., Ometov A.E., Timin L.S. Optimization of parameters of high-speed channels of the processor in order to increase the fault tolerance of the computer complex. Voprosy radioelektroniki, 2018, no. 2, pp. 87–92.

I.E. Bilyaletdinov, A.E. Ometov, L.S. Timin

OPTIMIZATION OF PARAMETERS OF HIGH-SPEED CHANNELS OF THE PROCESSOR IN ORDER TO INCREASE THE FAULT TOLERANCE OF THE COMPUTER COMPLEX

The modern computer facilities of the Elbrus family use high-speed data channels – DDR3 memory channels, interprocessor communication channels and I/O channels, built on the basis of PCI Express. Their physical level have a large number of settings, such as the rating of the termination resistor, preemphasis and equalization, the amplitude of the output signal, etc. All of these affect the integrity of the signals and the way channel operate. To increase the fault tolerance of the computer complex, the most optimal option is required for the physical level of high-speed channels. Due to the large number of possible settings, it is necessary to be able to evaluate in the shortest time and with a minimum amount of additional equipment. In this paper we developed a technique for determining the quality of settings using external mechanisms of the physical layer of channels.

Keywords: DDR3, PCI Express, build-in-self-test, eye diagrams.

REFERENCES

1. Foster G. Anatomy of an Eye Diagram – a Primer. SyntheSys Research, Inc., 2004, 9 p.
2. Wang L.-T., Stroud C. E., Toubia N. A. System-on-Chip Test Architectures: nanometer design for testability. Burlington, Morgan Kaufmann Publishers, 2008, 856 p.
3. Wang L.-T., Wu C.-W., Wen X. VLSI Test Principles and Architectures: Design for Testability. San Francisco, Morgan Kaufmann Publishers, 2006, 777 p.
4. IEEE Std 1149.1–2001: IEEE Standard Test Access Port and Boundary-Scan Architecture. New York, Institute of Electrical and Electronics Engineers, 2001, 208 p.

AUTHORS

Bilyaletdinov Ilya, graduate student, MTU (MIREA); engineer-programmer, PJSC Brook INEUM, 24, ulitsa Vavilova, Moscow, 119334, Russian Federation, tel.: +7 (495) 363-95-03, e-mail: ilia.e.bilyaletdinov@mcst.ru.

Ometov Aleksandr, engineer-programmer, JSC MCST, 24, ulitsa Vavilova, Moscow, 119334, Russian Federation, tel.: +7 (495) 363-95-03, e-mail: alexander.e.ometov@mcst.ru.

Timin Leonid, head of Sector, PJSC Brook INEUM, JSC MCST, 24, ulitsa Vavilova, Moscow, 119334, Russian Federation, tel.: +7 (495) 363-95-03, e-mail: le0@mcst.ru.