



Удвоение пропускной способности шины PCIe 6.0: переход на 4-уровневую частотно-импульсную модуляцию сигнала PAM4

Юлия Гарсия

В связи с высоким спросом на производительные серверные решения в центрах обработки данных (ЦОД) и в системах на базе алгоритмов искусственного интеллекта и машинного обучения большое значение приобретает увеличение скорости обмена данными по локальным шинам. Удвоение пропускной способности шины PCI Express до 64 ГТ/с увеличением частоты дискретизации цифрового сигнала ведёт к возрастанию влияния частотно-зависимых помех, при которых полезный сигнал становится неотличим от шума. В статье рассмотрен метод решения проблемы рабочей группой альянса PCI-Special Interest Group (PCI-SIG) для будущей спецификации 6.0 интерфейса PCIe.

Технология обмена данными PCI Express (PCIe) повсеместно используется в устройствах, требующих высокопроизводительных, ускоренных вычислений, подключения сетевых карт и графических ускорителей, но не исчерпывается только этим. Обладая преимуществом в высокой скорости передачи

данных перед параллельными шинами PCI, PCI-X и AGP, а также низкой задержкой (временем использования шины периферийным устройством), интерфейс PCIe получил широкое распространение в построении иерархических структур хранения данных (например, подключения ОЗУ и DRAM через сло-

ты PCIe с поддержкой протокола внутрисистемных соединений CXL (Compute Express Link) на базе PCIe 5.0.

Каждые три-семь лет скорость передачи данных PCI Express при сохранении обратной совместимости с предыдущими поколениями PCIe удваивалась (см. табл. 1).

Таблица 1
Эволюция интерфейса PCIe

PCIe Spec	Скорость передачи данных на линию, ГТ/с	Кодирование	Пропускная способность, x16, Гбит/с	Год
1,0	2,5	8b/10b	32	2003
2,0	5	8b/10b	64	2007
3,0	8	128b/130b	126	2010
4,0	16	128b/130b	252	2017
5,0	32	128b/130b	504	2019
6,0	64,0 (PAM4)	FLIT	1024	2021*

*прогнозируемая.

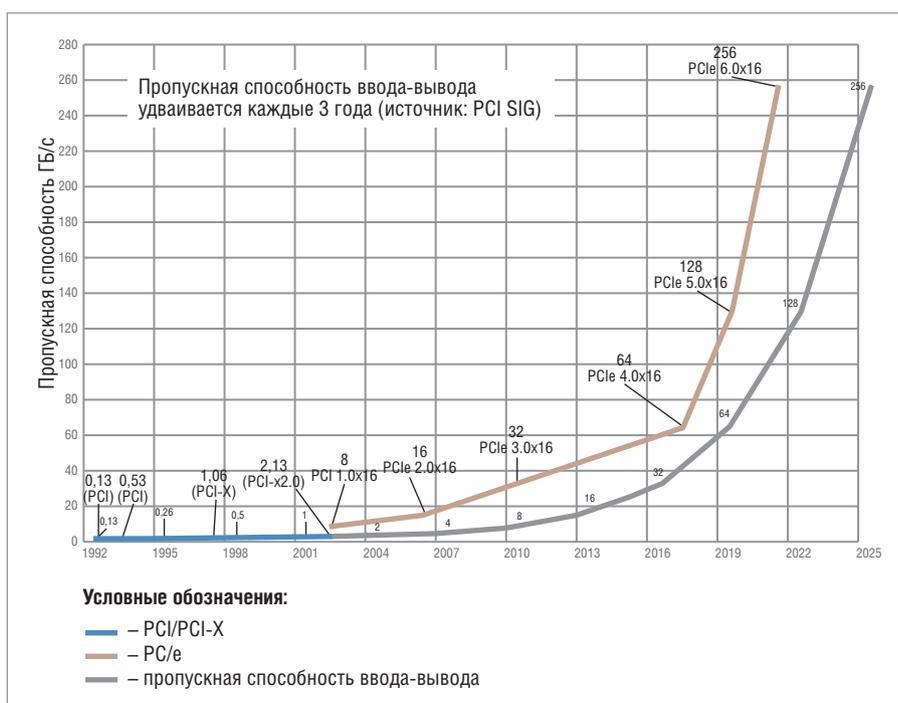


Рис. 1. Пропускная способность интерфейса PCI Express (с перспективой дальнейшего увеличения)

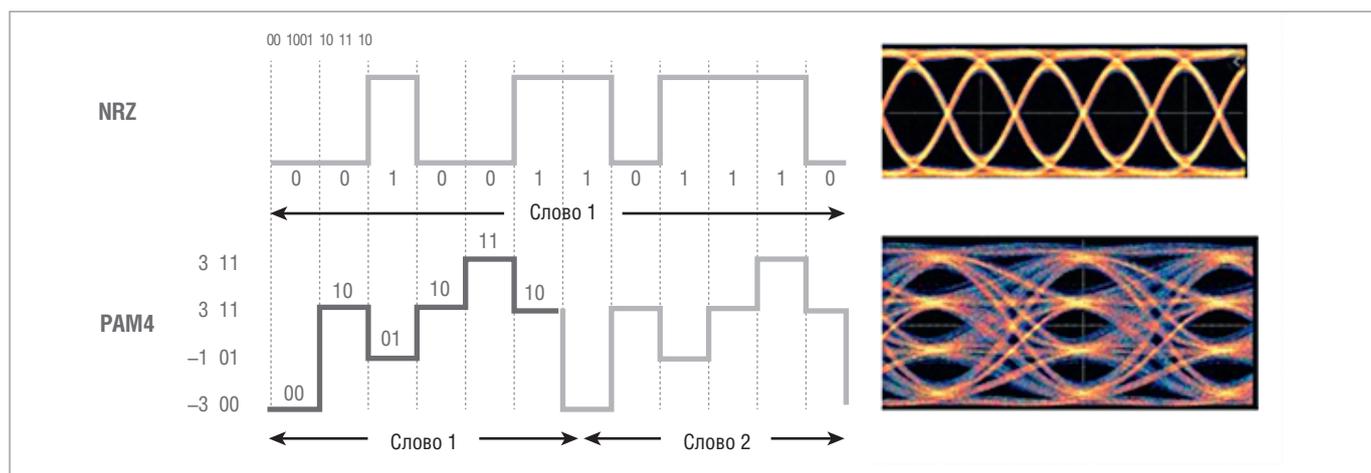


Рис. 2. Отличие амплитудно-импульсной модуляции PAM4 от NRZ-кодирования

Разработка приложений искусственного интеллекта (AI – Artificial intelligence) и машинного обучения (ML – Machine Learning), построение высокопроизводительных серверных систем и организация облачных вычислений продолжают способствовать генерации трафика гигантских объёмов феноменальными темпами. Стараясь соответствовать непрерывно растущим запросам на увеличение скорости обмена данными между процессором и компонентами, установленными на плате (GPU, FPGA, память), консорциум PCI-SIG (PCI Special Interest Group) представил предварительную версию спецификации PCIe 6.0 со скоростью передачи данных до 64 ГТ/с, окончательное утверждение которой ожидается к концу 2021 года (рис. 1).

Рассмотрим подробнее новую версию интерфейса и его реализацию.

Главная проблема PCIe 6.0

Чтобы избежать дорогостоящей модернизации инфраструктуры, новый интерфейс должен соответствовать требованиям обратной совместимости с предыдущими спецификациями, например, электрические параметры устройств нового поколения PCIe 6.0 должны соответствовать более старым версиям на объединительной плате (например, поддерживать скорость передачи данных до 28 Гбит/с). Требование механической совместимости разъёмов PCIe не позволяет увеличивать количество линков для увеличения пропускной способности.

В спецификации PCIe 5.0 используется моделирование цифровых сигналов кодированием без возврата к нулю (NRZ, no-return-to-zero). При этом искажение сигнала для каналов может достигать 36 дБ (частота Найквиста, рав-

ная половине частоты дискретизации, при частоте тактирования 32 ГГц составляет 16 ГГц). При удвоении скорости передачи данных с 32 ГТ/с (гигатранзакций в секунду) до 64 ГТ/с кодированием без возврата к нулю частота Найквиста равна 32 ГГц, при этом частотно-зависимые потери канала увеличиваются до 70 дБ [1].

Таким образом, если пропускная способность будет увеличиваться за счёт увеличения частоты, например, до 56 ГГц, вносимые потери IL (Insertion loss) на частоте Найквиста (28 ГГц) составляли бы ~60 дБ, а отношение вносимых потерь сигнала IL к перекрёстным помехам (ICR – insertion-loss-to-crosstalk ratio) при этом стремилось бы к нулю. Это обстоятельство делает невозможным увеличение скорости передачи данных до 56 Гбит/с традиционным методом повышения частоты квантования (дискретизации). Удвоение скорости передачи данных вносит в сигнал существенные искажения, даже если речь идет о небольшом расстоянии, на которое передаётся сигнал.

Частота передачи сигнала свыше 32 ГГц делает его более нестабильным, практически неотличимым от шума.

Переход на PAM4-кодирование

Новая ревизия интерфейса PCIe использует вместо NRZ-кодирования 4-уровневую амплитудно-импульсную модуляцию (PAM4), основанную на использовании не 2, а 4 значений напряжения, и передаёт 2 бита за минимальный интервал времени между изменениями состояния сигнала (Unit Interval), в отличие от кодирования без возврата к нулю, которое передаёт только 1 бит за тот же интервал (рис. 2). Этот метод передачи сигнала позволяет увеличить пропускную способность PCIe 6.0 в 2 раза, поддерживая искажение сигнала на том же приемлемом уровне, что и в предыдущей версии стандарта PCIe 5.0.

На рис. 3 хорошо заметно, что для канала PAM4 вносимые потери IL составляют ~31 дБ, перекрёстные ICR – ~30 дБ на частоте 14 ГГц. Также можно

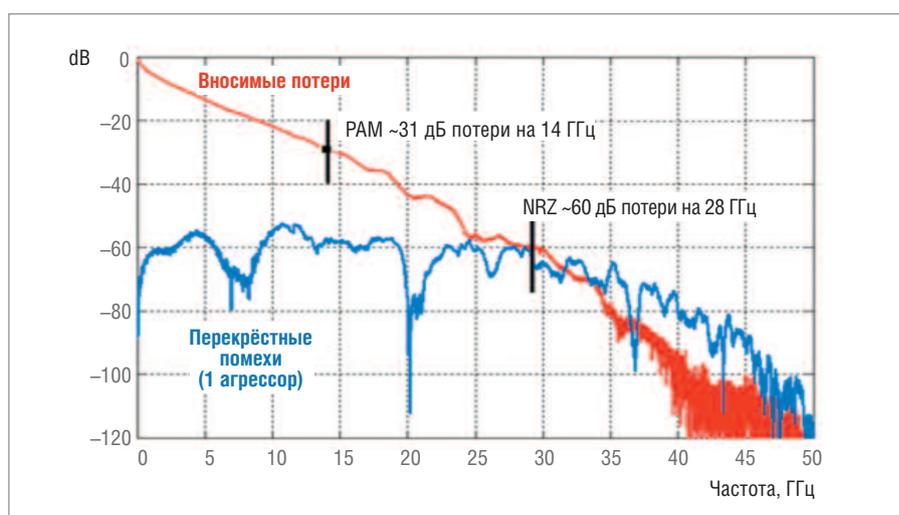


Рис. 3. Сравнительная характеристика частотно-зависимых потерь каналов PCIe 6.0 и PCIe 5.0

заметить, что значения вносимых ИЛ и перекрёстных ICR помех при моделировании сигналов NRZ (канал на объединительной плате предыдущих версий стандарта) на частоте 28 ГГц составляют 60 дБ.

Однако преимущество новой сигнальной структуры PAM4 обходится дорого: запас помехоустойчивости для PAM4-кодирования снижается на 9,5 дБ (33%). Это усугубляет неблагоприятное воздействие отражённого сигнала и шумов от источника питания.

FEС и код Грея

Несмотря на то что в интерфейсе PCIe 6.0 удвоена скорость передачи данных за счёт использования PAM4-кодирования, пониженное соотношение сигнал/шум (SNR – signal-to-noise-ratio) делает его более восприимчивым к помехам по сравнению с кодированием NRZ, способствует высокой частоте битовых ошибок и может привести к сбоям в работе системы или снижению производительности. Метод восстановления целостности сигнала в стандарте

PCIe 6.0 – упреждающая коррекция ошибок FEC (forward-error-correction) – предусматривает отправление избыточных данных вместе с полезными при условии, что частота ошибок ниже определённого порогового значения. Циклическая проверка избыточности (CRC – cyclic redundancy check) выполняется для обнаружения и исправления битовых ошибок, если CRC обнаруживает ошибки после FEC, запускается механизм повторной проверки.

Также в качестве повышения помехоустойчивости сигнала в PCIe 6.0 применяется двоичный циклический код (код Грея). Код Грея оперирует самым старшим битом (MSB – most significant bit) и самым младшим битом (LSB – least significant bit) таким образом, чтобы ошибка, вызванная электрическими помехами, приводила максимум к ошибке в одном разряде (рис. 4).

Предыдущие поколения PCIe поддерживали режим экономичного энергопотребления за счёт динамического изменения ширины канала. PCIe 6.0 вводит режим экономии энергопо-

требления L0p (Low Power State), который позволяет изменять потребляемую мощность пропорционально пропускной способности без прерывания трафика.

В спецификации PCIe 6.0 предусмотрено кодирование на основе блока управления потоком FLIT (Flow Control Unit) для обеспечения меньшей величины задержки, связанной с применением алгоритмов FEC и CRC. Таким образом, добавление вышеуказанных механизмов самокоррекции в PCIe 6.0 не должно существенно увеличивать задержку (латентность) по сравнению с версией PCIe 5.0.

Рабочая группа PCI-SIG доказала, что для PCIe 6.0 её уровень не превышает 10 нс (рис. 5).

Линейность

Для метода PAM4 характерен так называемый эффект нелинейности, хорошо видимый на глаз-диаграмме (рис. 6). В левой части приведена идеальная линейность, для которой высоты разделения уровней одинаковы. Интервал

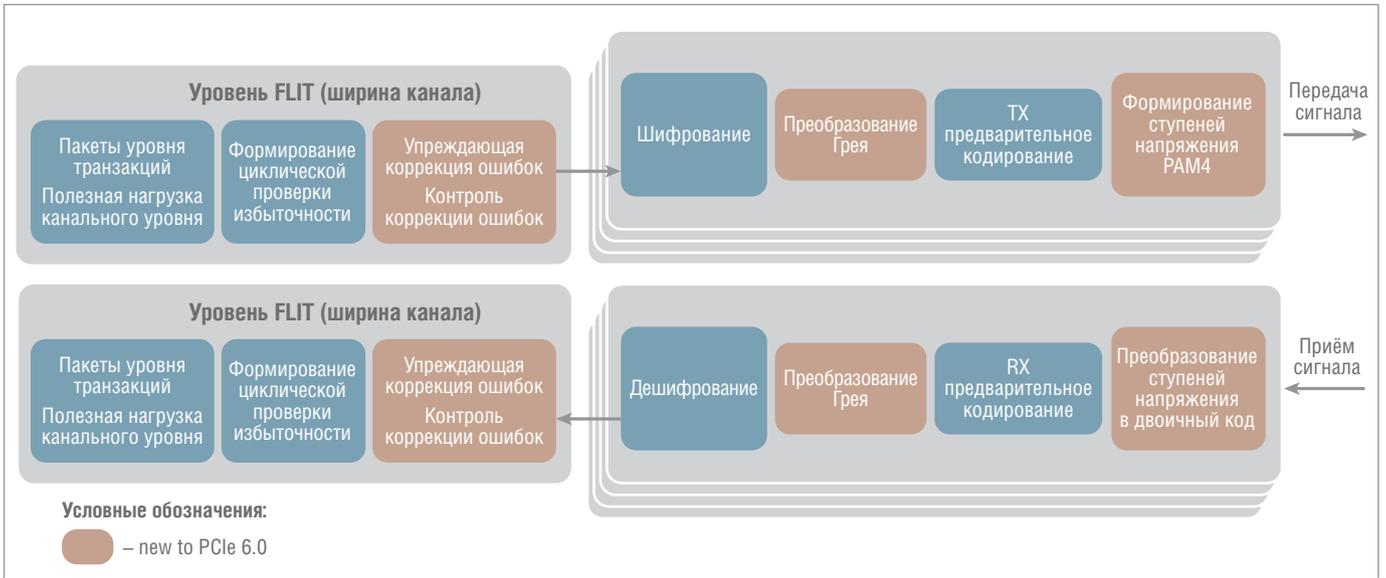


Рис. 4. Добавление FEС и кода Грея в PAM4-кодирование

Объём данных	Объём пакета	(X1 Link)			(X16 Link)				
		Задержка (нс) при кодировании 128b/130b и скорости передачи данных 32 ГТ/с	Задержка (нс) в режиме FLIT и скорости передачи данных 64 ГТ/с	Накопление задержки (нс)	Задержка (нс) при кодировании 128b/130b и скорости передачи данных 32 ГТ/с	Задержка (нс) в режиме FLIT и скорости передачи данных 64 ГТ/с	Накопление задержки (нс)		
0	4	6.09375	18	11.90625	0	4	0.380859375	1.125	0.744140625
4	8	10.15625	20	9.84375	4	8	0.634765625	1.25	0.615234375
8	12	14.21875	22	7.78125	8	12	0.888671875	1.375	0.486328125
16	20	22.34375	26	3.65625	16	20	1.396484375	1.625	0.228515625
32	36	38.59375	34	-4.59375	32	36	2.412109375	2.125	-0.287109375
64	68	71.09375	50	-21.09375	64	68	4.443359375	3.125	-1.318359375
128	132	136.09375	82	-54.09375	128	132	8.505859375	5.125	-3.380859375
256	260	266.09375	146	-120.09375	256	260	16.63085938	9.125	-7.505859375
512	516	526.09375	274	-252.09375	512	516	32.88085938	17.125	-15.75585938
1024	1028	1046.09375	530	-516.09375	1024	1028	65.38085938	33.125	-32.25585938

Рис. 5. Сравнение задержки для 1 линии контакта и 16 линий

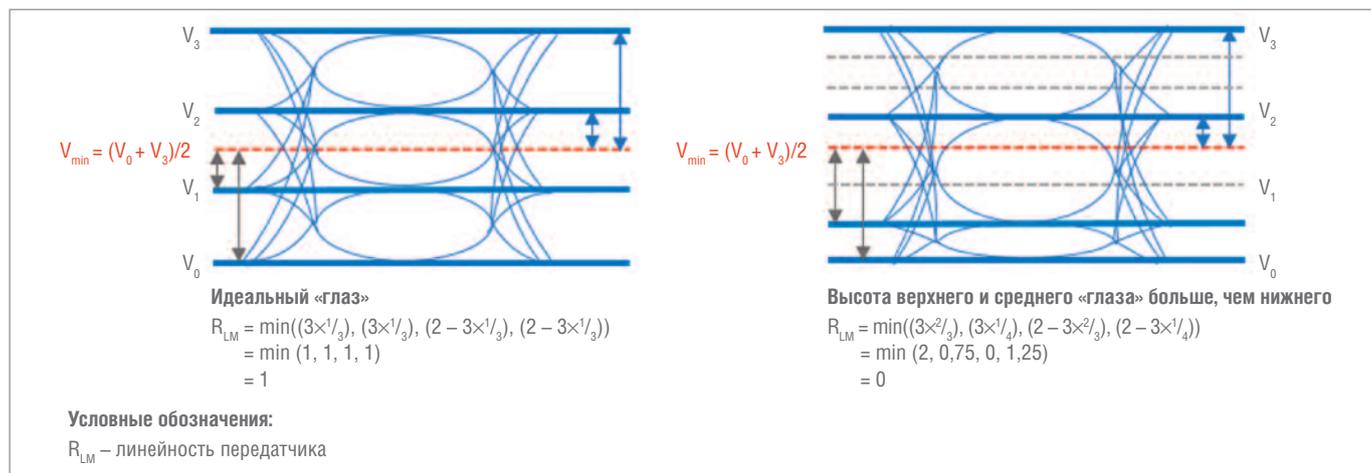


Рис. 6. Графическое представление цифрового сигнала, позволяющее оценить качество передачи (глаз-диаграмма)

между V_1 и V_{min} составляет одну треть интервала между V_0 и V_{min} . Аналогичным образом расстояние между V_2 и V_{min} составляет одну треть от расстояния между V_3 и V_{min} . В идеальном случае коэффициент рассогласования расстояния между уровнями RLM равен 1. Чем ближе реальный показатель RLM к 1, тем лучше линейность.

Для примера на правой части рис. 6 показан «глаз» с плохой линейностью – уровень сигнала V_1 и V_2 настолько низок, что интервал между V_1 и V_{min} составляет две трети от V_0 до V_{min} , а интервал от V_2 до V_{min} равен всего одной четверти интервала между V_3 и V_{min} . В этом случае расчётный показатель RLM равен 0.

Выравнивание амплитудно-частотной характеристики и усиление сигнала могут внести такие искажения, что два уровня напряжения из четырёх будут зафиксированы приёмником сигнала

как один. В целом недостаточная линейность, описанная правой частью глаз-диаграммы (рис. 6), приводит к неустраняемым битовым ошибкам [2].

ЗАКЛЮЧЕНИЕ

Переход на 4-уровневую импульсно-амплитудную модуляцию (PAM4) сигнала в настоящий момент является оптимальным средством увеличения пропускной способности при условии сохранения совместимости линков и приемлемого соотношения сигнал/шум.

Несмотря на то что целый ряд задач ещё ждёт своих решений от разработчиков PCI-SIG, параметры помехоустойчивости новой спецификации интерфейса соответствуют возможностям современных приёмопередающих устройств PCIe 6.0, что обеспечит центрам обработки данных, телекоммуника-

ционным и другим системам с повышенными требованиями к пропускной способности скорость передачи данных до 64 Гбит/с на линию. ●

ЛИТЕРАТУРА

1. Whitepaper: Pushing the Envelope with PCIe 6.0: Bringing PAM4 to PCIe, Tony Chen, Cadence. [Электронный ресурс] // URL: https://www.cadence.com/content/dam/cadence-www/global/en_US/documents/tools/ip/design-ip/pushing-the-envelope-with-pcie-6-wp.pdf.
2. AN 835: PAM4 Signaling Fundamentals Intel. [Электронный ресурс] // URL: <https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/an/an835.pdf>.

**Автор – сотрудник
 фирмы ПРОСОФТ
 Телефон: (495) 234-0636
 E-mail: info@prosoft.ru**

НОВОСТИ НОВОСТИ НОВОСТИ НОВОСТИ НОВОСТИ НОВОСТИ

Компании ААЕОН вручили почётную награду «Образцовые стандарты коммерческой деятельности»

Объявлен победитель 26-й Тайваньской Национальной премии качества, присуждаемой предприятиям с наивысшим уровнем качества управления. В этом году награда заслуженно досталась компании ААЕОН – тайваньскому производителю встраиваемых систем и компонентов для создания приложений искусственного интеллекта, граничных вычислений, инфраструктуры Умных городов и промышленного Интернета вещей.

В течение двух лет оргкомитет государственной премии проводил тщательный анализ, оценивая претендентов по следующим критериям:

- руководящий состав и идеология бизнеса;
- стратегическое управление;

- исследования, разработки и новаторская деятельность;
- расширение границ рынка и клиентской базы;
- кадровый потенциал и управление знаниями;
- стратегия управления информацией;
- управление технологическими процессами;
- эффективность коммерческой деятельности.

Благодаря многолетним усилиям сотрудников всех подразделений и служб контроля качества компания ААЕОН смогла обойти других участников и получить высокую государственную награду. С самого дня основания руководство компании ААЕОН осу-

ществляло коммерческую деятельность, опираясь на потребности своих заказчиков и добросовестный, высококачественный сервис. В компании ААЕОН функционирует система контроля качества, соответствующая международным стандартам. Приоритетным принципом разработки и производства в компании является принцип сохранения окружающей среды, который помогает выполнить социальные обязательства перед будущими поколениями. Компания ААЕОН наследует ценные знания о встраиваемых технологиях прошлого и изучает современные, анализирует современную компонентную базу и направления развития вертикальных рынков для того, чтобы стать в авангарде фирм-изготовителей вычислительных решений для приложений искусственного интеллекта. ●