

Кравцов Роман Николаевич

аспирант кафедры информатики и информационных технологий, Московский политехнический университет, Москва. ORCID: 0009-0003-8961-6310. SPIN-код: 7847-1993. AuthorID: 1293068.
Электронный адрес: kravtsov81@gmail.com

Roman N. Kravtsov

Postgraduate at the Department of Informatics and Information Technology, Moscow Polytechnic University, Moscow. ORCID: 0009-0003-8961-6310. SPIN code: 7847-1993. AuthorID: 1293068.
E-mail address: kravtsov81@gmail.com

Винокур Алексей Иосифович

доктор технических наук, профессор, профессор кафедры информатики и информационных технологий, Московский политехнический университет, Москва. ORCID: 0000-0002-6914-2520. SPIN-код: 9976-2104. AuthorID: 690024.
Электронный адрес: alexvinokour@gmail.com

Aleksey I. Vinokur

Doctor of Engineering Sciences, Full Professor, Professor at the Department of Informatics and Information Technology, Moscow Polytechnic University, Moscow. ORCID: 0000-0002-6914-2520. SPIN code: 9976-2104. AuthorID: 690024.
E-mail address: alexvinokour@gmail.com

МЕТОД АППАРАТНОГО УСКОРЕНИЯ ПОСТКВАНТОВЫХ КРИПТОГРАФИЧЕСКИХ ПРИМИТИВОВ НА БАЗЕ НЕЙРОННЫХ ПРОЦЕССОРОВ В IOT

Аннотация. Работа посвящена проблеме эффективной реализации постквантовой криптографии (PQC) в ресурсоограниченных системах интернета вещей (IoT). Программное исполнение криптографии на решетках (LBC) на классических микроконтроллерах приводит к критическим задержкам и энергопотреблению. В статье предложен метод адаптации вычислительных ядер LBC для исполнения на гетерогенных SoC, оснащенных нейронными сопроцессорами (NPU). Математически обоснована эквивалентность операции полиномиального умножения и тензорной свертки, что позволяет делегировать нагрузку NPU-акселераторам. Валидация метода проведена путем моделирования. Расчетная оценка демонстрирует, что предложенная гибридная архитектура обеспечивает значительное ускорение криптографических примитивов с ядрами Cortex-M4. Также проанализированы векторы атак на NPU и предложены архитектурные механизмы защиты памяти.

Ключевые слова: постквантовая криптография, интернет вещей, криптография на решетках, нейронный процессор (NPU), гетерогенные SoC, аппаратное ускорение.

Для цитирования: Кравцов Р.Н., Винокур А.И. Метод аппаратного ускорения постквантовых криптографических примитивов на базе нейронных процессоров в IoT // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. 2026. № 3. С. 151–163. DOI: 10.18137/RNU.V9I87.26.03.P.151

A METHOD FOR HARDWARE ACCELERATION OF POST-QUANTUM CRYPTOGRAPHIC PRIMITIVES USING NEURAL PROCESSING UNITS IN IOT

Abstract. This paper addresses the problem of efficient implementation of post-quantum cryptography (PQC) in resource-constrained Internet of Things (IoT) systems. Software implementation of lattice-based cryptography (LBC) standards on classical microcontrollers results in critical latency and high energy consumption. The article proposes a method for adapting LBC computational kernels for execution on heterogeneous SoCs equipped with Neural Processing Units (NPUs). The equivalence between polynomial multiplication and tensor convolution is mathematically substantiated, allowing the computational load to be offloaded to NPU accelerators. The method was validated through simulation. Performance estimation demonstrates that the proposed hybrid architecture achieves significant speedup for cryptographic primitives compared to Cortex-M4 cores. Additionally, attack vectors targeting NPUs are analyzed, and architectural memory protection mechanisms are proposed.

Keywords: post-quantum cryptography, internet of things (IoT), lattice-based cryptography (LBC), neural processing unit (NPU), heterogeneous SoC, hardware acceleration.

For citation: Kravtsov R.N., Vinokur A.I. (2026) A method for hardware acceleration of post-quantum cryptographic primitives using neural processing units in IoT. *Vestnik of Russian New University. Series: Complex Systems: Models, analysis, management*. No. 3. Pp. 151–163. DOI: 10.18137/RNUV9187.26.03.P.151 (In Russian).

Введение

Криптографическая стойкость классических ассиметричных стандартов (RSA, ECC) оказывается под угрозой ввиду потенциальной возможности реализации алгоритма Шора [1] на квантовых вычислителях, что в совокупности с прогрессом в области наращивания кубитов [2] актуализирует задачу миграции на постквантовую криптографию (PQC).

Критическим фактором для систем с длительным жизненным циклом (государственные архивы, ключи шифрования) являются атаки типа Harvest Now, Decrypt Later (HNDL), формирующие окно уязвимости еще до физического появления квантового компьютера [3].

Теоретические и практические ограничения классической криптографии в IoT

Алгоритм Шора снижает время подбора 2048-битного ключа RSA до нескольких часов при наличии квантового компьютера [4]. Узлы IoT особенно уязвимы к данной угрозе. Причины кроются в длительном сроке эксплуатации оборудования и строгих лимитах энергопотребления. Миграция на PQC требует оценки не только вычислительной сложности операций, но и их энергетической стоимости.

Согласно стандартам Национального института стандартов и технологий США (NIST) эквивалент AES-128 достигается при длине ключа RSA от 3072 бит [5]. Программная реализация такого алгоритма на малоресурсных микроконтроллерах нецелесо-

Метод аппаратного ускорения постквантовых криптографических
примитивов на базе нейронных процессоров в IoT

образна. Вычисления могут вызывать критические задержки работы процессора. Дополнительно возникает проблема блокировки каналов передачи данных в энергоэффективных сетях (LoRaWAN, Zigbee) из-за превышения размера передаваемых пакетов.

По этой причине алгоритм RSA рассматривается в данном исследовании как справочный ориентир. Текущим стандартом индустрии и конкурентом PQC выступает криптография на эллиптических кривых (ECC).

Сравнительный анализ этапов установления защищенного соединения на базе платформы Cortex-M4 представлен в Таблице 1. В качестве основы были выбраны функционально эквивалентные операции протокола рукопожатия со стороны устройств IoT. Для классической криптографии оценивается процедура выработки общего секрета (ECDH). Для постквантовых алгоритмов оценивался механизм инкапсуляции ключа (KEM) и подписи.

Таблица 1

**Сравнительные характеристики операций установления защищенного соединения
на типовых IoT-устройствах**

Алгоритм	Операция	Тактов CPU	Время, мс*	Энергия, мкДж**
X25519 (ECC)	ECDH (Shared Secret)	428 570	2,55	≈255
RSA-2048	Key Transport (Public)	≈2 100 000	12,44	≈1 244
Kyber-768	Инкапсуляция	1 146 556	6,82	≈682
Dilithium3	Подпись	9 289 499	55,30	≈5 530

* Время приведено для частоты 168 МГц (Cortex-M4).

** Энергия рассчитана пропорционально времени исполнения (при полной загрузке CPU).

Источник: таблица составлена авторами на основе [6–8].

Данные Таблицы 1 для алгоритма RSA отражают только операции с открытым ключом (12,44 мс). При этом операции с закрытым ключом (например, генерация цифровой подписи) занимают на микроконтроллерах более 450 мс [8]. Это затрудняет использование RSA в системах реального времени.

Современный стандарт ECC демонстрирует высокую скорость выполнения базовых операций (2,55 мс). Однако миграция на постквантовые алгоритмы (Kyber, Dilithium) увеличивает вычислительные задержки в 2,7–21 раз. Энергопотребление при этом возрастает до пиковых 5530 мкДж. Следовательно, программная реализация PQC-примитивов аналогично затрудняет использование PQC.

Специфика уязвимостей

Специфика инфраструктуры IoT характеризуется длительным сроком эксплуатации (10–15 лет) и сложностью обновлений парка устройств. Согласно отчетам [9–11], промышленный сектор лидирует по числу инцидентов, при этом средняя стоимость утечки достигла 4,88 млн долларов. Пересечение сроков жизни устройств с моментом появления криптографически релевантного квантового компьютера (CRQC) классифицируется NIST как критический риск управления активами, требующий превентивной защиты от HNDL-атак уже сегодня [12].

Собираемые данные могут обладать длительной ценностью, превышающей срок службы самого оборудования. Это создает риск асимметричного ущерба, когда компрометация бюджетного IoT-узла приводит к системным потерям критической инфраструктуры [13], что делает защиту от атак HNDL приоритетной задачей.

Проблема усугубляется невозможностью программной модернизации унаследованного парка. Catuogno и Galdi [14] указывают, что отсутствие доверенного загрузчика и аппаратных криптографических примитивов в низкоресурсных системах делает реализацию защищенных механизмов обновления технически невыполнимой.

Уязвимости к атакам по сторонним каналам

Ресурсные ограничения IoT-устройств создают предпосылки для атак по сторонним каналам при реализации постквантовых алгоритмов. Исследования на платформах Cortex-M4 выявили высокую корреляцию энергопотребления с секретными параметрами при выполнении криптографических операций [15]. Для устройств с ограниченными ресурсами это является критическим вектором атаки, требующим специальных методов защиты, таких как маскирование секретных значений [15].

Также существенны и временные атаки. Вариативность времени выполнения операций в зависимости от загрузки системы позволяет злоумышленнику применять статистический анализ для восстановления ключа. Кроме того, использование единой памяти (HyperRAM) для кода и данных создает возможности для атак по времени доступа [16].

Классификация IoT-устройств

Для систематизации рисков и выработки стратегии защиты предлагается классификация IoT-устройств, основанная на модели жизненного цикла и критичности обрабатываемых данных (Таблица 2) [17–19]. Данная матрица позволяет дифференцировать подходы к миграции, опираясь на методологию NIST по переходу на постквантовую криптографию [17].

Таблица 2

Классификация IoT-устройств по параметрам уязвимости к квантовым угрозам

Класс устройства	Типичные примеры	Срок ввода в эксплуатацию*	Критичность данных (HNDL риск)	Рекомендуемый подход к миграции
Критический	Медицинские импланты, контроллеры SCADA	2018–2022	Высокая (угроза жизни / инфраструктуре)	Немедленная аппаратная замена или гибридная схема (Crypto-Agility)
Высокий риск	Умные замки, автокомпоненты, камеры наблюдения	2020–2024	Средняя / Высокая (персональные данные, видео)	Приоритетное обновление прошивки (OTA) до 2027 г.
Средний риск	Носимая электроника, датчики умного дома	2022–2025	Средняя (поведенческие данные)	Плановая миграция на PQS до 2030 г.
Низкий риск	Одноразовые метки, простые сенсоры	2023+	Низкая (агрегированные данные)	Замена устройства по исчерпанию ресурса

* Срок ввода указывает на устройства, чей жизненный цикл (10–15 лет) пересечется с моментом появления CRQC.

Источник: таблица составлена авторами на основе [17–19].

Метод аппаратного ускорения постквантовых криптографических
примитивов на базе нейронных процессоров в IoT

Для массового сегмента (Consumer IoT, см. Таблицу 2) решения на базе FPGA или ASIC экономически нецелесообразны из-за ограничений управления спецификациями (BOM). Оптимальной стратегией может стать использование уже имеющихся на кристалле NPU-блоков, что не потребует изменения аппаратной конфигурации устройства.

Классификация PQС и критерии выбора

Из финалистов конкурса NIST (Таблица 3) наиболее перспективными для IoT являются алгоритмы на решетках (LBC). Они обладают сбалансированной сложностью $O(n \log n)$ и потенциалом для распараллеливания, в отличие от кодовых схем (большой размер ключей) или хеш-функций (высокая задержка).

Таблица 3

Сравнительный анализ алгоритмов PQС в контексте ограничений IoT

Алгоритм (тип)	Размер ключа (секретный / открытый)	Базовая операция	Потенциал аппаратного ускорения	Вердикт для IoT
Kyber (LBC)	Kyber-512: 1.5/0.8 КБ; Kyber-768: 2.3/1.2 КБ	NTT + полиноми- альное умножение	Высокий (для SIMD, FPGA, потенциально NPU)	Основа для гибридных схем
Dilithium (LBC)	Dilithium3: 3.0/1.9 КБ; Dilithium5: 3.5/2.5 КБ	NTT + полиноми- альное умножение	Высокий (хорошо распараллеливается)	Стандарт ЭЦП для IoT
BIKE / HQC (Code)	BIKE: ~10-20 КБ; HQC: ~50-100 КБ	Битовые операции в $GF(2^m)$ XOR, сдвиги	Средний (требует специализированных FPGA)	Резервный вариант для энергоемких IoT
SPHINCS+ (Hash)	32-64 КБ / 1-4 КБ	Хеш-функции (SHA3/SHAKE)	Низкий (требует SHA- акселератор)	Только для offline-сценариев
McEliece	255-1326 КБ / 255-1326 КБ	Линейная алгебра над $GF(2)$	Ограниченный (плохо параллелится)	Невозможно для части IoT

Источник: здесь и далее таблицы составлены авторами по материалам исследования.

Обоснование выбора LBC

Решетчатые алгоритмы (LBC) являются безальтернативным выбором для массового IoT по совокупности факторов. Алгоритмы на кодах (BIKE, HQC) требуют передачикратно большего объема данных, что критично для узкополосных сетей (NB-IoT, LoRaWAN).

Основным ограничением производительности LBC остается вычислительная сложность полиномиального умножения. В частности, для алгоритма NTRU-HRSS701 ($n = 701$, $q = 8192$) даже при использовании оптимизированных методов (Toom-4/Karatsuba) одно полиномиальное умножение в кольце R_q требует более 170,000 тактов процессора Cortex-M4 [20].

Эволюция аппаратных решений

Сравнительный анализ аппаратных архитектур (Таблица 4) выделяет гибридный подход (CPU-NPU) как компромиссное решение, обеспечивающее баланс между производительностью и стоимостью внедрения в существующую инфраструктуру.

Таблица 4

Сравнительный анализ аппаратных решений

Тип решения	Примеры	Преимущества	Недостатки
ASIC	PQShield	Производительность, энергоэффективность	Высокая стоимость разработки, нет гибкости
FPGA	Xilinx, Intel	Перепрограммируемость	Сложность разработки, цена, энергопотребление
Гибридные CPU-NPU	Hailo, Google TPU	Экономическая эффективность, доступность	Ограниченная точность, отсутствие mod q

Математическая основа синергии

Архитектурная синергия между LBC и NPU имеет строгое математическое обоснование. Ключевой операцией криптографии на решетках является полиномиальное умножение в кольце

$$R_q = \frac{\mathbb{Z}_q[X]}{X^N \pm 1},$$

допускающее представление в виде операций, базовых для нейронных сетей.

Для алгоритмов типа NTRU полиномиальное умножение

$$c(x) = a(x)b(x)(\text{mod}(x^N - 1))$$

эквивалентно циклической свертке векторов коэффициентов:

$$c_k = \sum_{i+j \equiv k \pmod{N}} a_i b_j.$$

Эта операция идентична слою 1D-свертки (Convolution) в глубоком обучении. Альтернативно, она может быть представлена как умножение циркулянтной матрицы, составленной из коэффициентов $a(x)$, на вектор $b(x)$ (GEMV).

На основе выявленной математической двойственности предлагается концептуальная схема разделения вычислительной нагрузки (см. Рисунок). В данной архитектуре наиболее ресурсоемкая тензорная часть (свертка или матричное умножение) делегируется ускорителю NPU, в то время как операции, требующие сложной логики управления или специфической арифметики (генерация полиномов, модульная редукция), выполняются на центральном процессоре.

Метод аппаратного ускорения постквантовых криптографических примитивов на базе нейронных процессоров в IoT

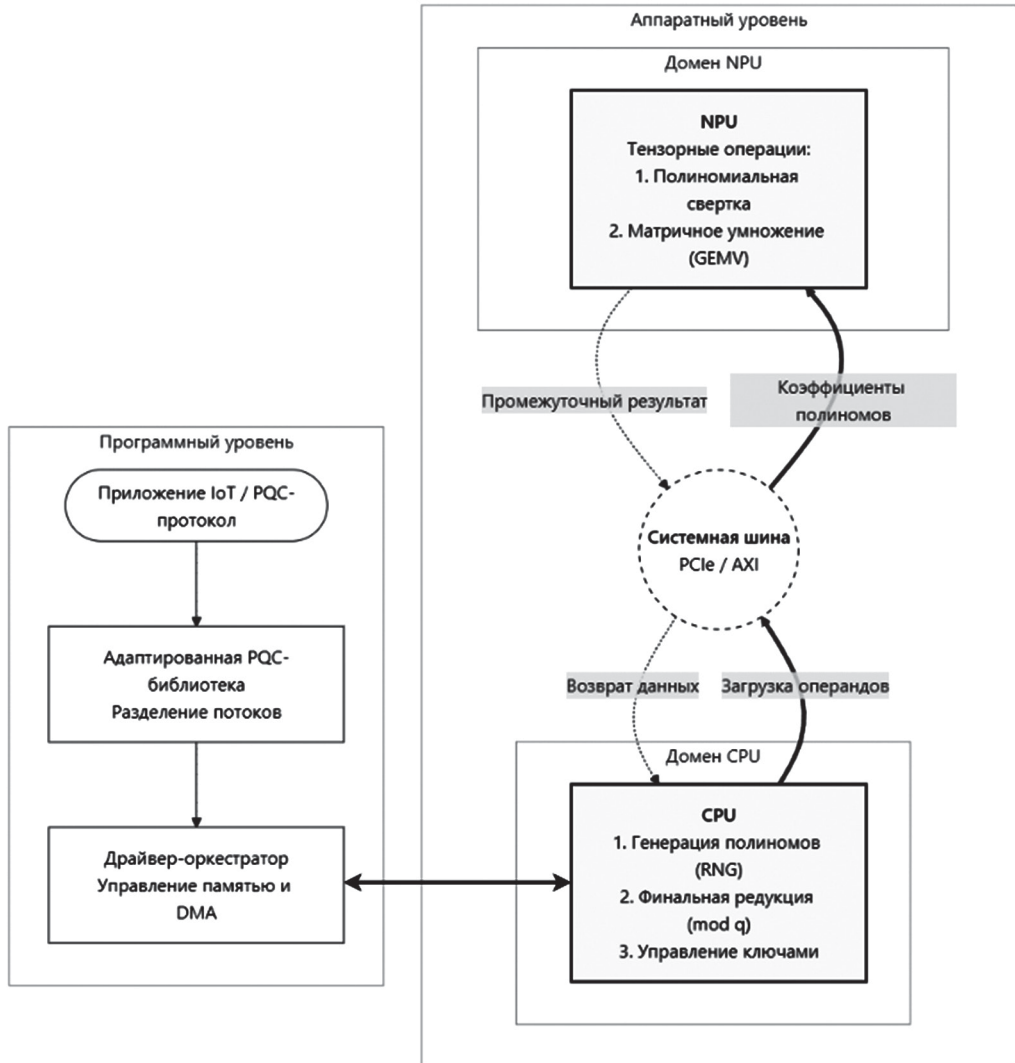


Рисунок. Концептуальная схема гибридной вычислительной архитектуры: разделение функциональных обязанностей между CPU и NPU

Источник: рисунок выполнен авторами

Модель угроз и архитектурная реализация

Внедрение NPU создает прецедент использования архитектурно закрытых компонентов для критических операций. В отличие от документированных CPU микроархитектура NPU зачастую является проприетарной, что затрудняет независимый аудит безопасности. Можно выделить три специфических вектора угроз для таких систем:

1. В архитектуре SoC центральный процессор и NPU, как правило, разделяют общую системную шину и контроллер памяти. В условиях многозадачной ОС без стро-

гой изоляции ресурсов возникает потенциальный вектор атаки по сторонним каналам.

2. Аппаратный пропуск операций умножения на ноль, эффективный для нейросетей, в криптографии создает уязвимость по времени исполнения, раскрывая количество нулевых коэффициентов в секретном полиноме.

3. Использование ускорителей через PCIe или USB создает риск перехвата данных на физическом уровне, а также атак через прямой доступ к памяти.

Для нейтрализации рисков перехвата данных на системной шине предполагается использование механизмов аппаратной изоляции памяти (IOMMU для шлюзов или MPU/TrustZone для микроконтроллеров). В рамках предлагаемой модели драйвер-оркестратор обеспечивает монопольный доступ к NPU на время выполнения криптографической операции, блокируя попытки обращения к памяти ускорителя со стороны сторонних процессов, что нивелирует риски атак по времени доступа.

Теоретическая оценка производительности

Валидация модели в среде PyTorch [21] через побитовое сравнение со стандартной реализацией (SymPy) подтвердила полную математическую эквивалентность тензорной свертки и полиномиального умножения, что гарантирует корректность криптографических преобразований при переносе на NPU.

Для верификации гипотезы проведен сравнительный расчет эффективности выполнения полиномиального умножения (на параметрах NTRU-HRSS-701). В качестве базовых вычислителей, не имеющих аппаратного ускорения, рассмотрены два представителя инфраструктуры IoT:

1. Уровень Gateway – процессор Cortex-A76, использующий только программные векторные инструкции NEON.

2. Уровень Endpoint – микроконтроллер Cortex-M4, опирающийся на DSP-расширения.

Данным программным подходам противопоставляются две предлагаемые гибридные топологии. В качестве референсного сопроцессора выбран тензорный чип Nailo-8L. Для Gateway сценария он выступает моделью внешнего ускорителя, обмен данными с которым идет по системной шине PCIe. Для Endpoint устройств параметры Nailo-8L служат оценкой перспективной системы на кристалле (SoC), где NPU интегрирован рядом с ядром Cortex-M4 и делит с ним высокоскоростную внутреннюю шину AHB/AXI.

Поскольку NTRU и Kyber требует работы с коэффициентами разрядностью 12–13 бит, прямая обработка в INT8 недопустима из-за потери точности. Расчет теоретической производительности NPU выполнен для консервативного режима INT16:

$$P_{\text{eff}} = \frac{13 \text{ TOPS}}{4} = 3,25 \text{ TOPS} \approx 3250 \text{ GOPS}.$$

Параметр P_{eff} отражает асимптотический предел пропускной способности архитектуры. Достижение такой производительности на практике потребует максимальной загрузкой вычислительных ядер, за счет внедрения механизма пакетной обработки входных данных.

Метод аппаратного ускорения постквантовых криптографических
примитивов на базе нейронных процессоров в IoT

Критическим фактором является защита от переполнения разрядной сетки в момент накопления суммы свертки. Современные NPU имеют 32-битные аккумуляторы сложения. Моделирование худшего случая для NTRU-HRSS-701 ($N = 701$, модуль $q = 8192$) показывает, что даже при максимальных значениях коэффициентов итоговая сумма не превышает порог $S_{\max} \approx N \cdot q \approx 5,7 \cdot 10^6$. Выявленный максимум гарантированно укладывается в диапазон знакового целого $\text{INT32} \approx (2,14 \cdot 10^9)$.

Размерность полиномов в стандартах NIST ($N = 701$, $N = 768$) структурно не кратны размерностям аппаратных тензорных ядер. Для запуска вычислений матрицы дополняются нулями. Эта избыточность является незначительной и учтена в итоговых расчетах.

Для Cortex-A76 расчет ведется с учетом задержек шины PCIe и драйвера. Для Cortex-M4 моделируется взаимодействие через высокоскоростную внутреннюю системную шину (AHB/AXI), характерную для интегрированных SoC-решений.

Результаты сопоставления представлены в Таблице 5. Характеристики NPU выделены в отдельную колонку как универсальная математическая модель сопроцессора. Чистое время тензорных вычислений идентично для обоих сценариев. Формирование итоговых задержек гибридной архитектуры зависит исключительно от накладных расходов на транспортировку матриц по локальной шине AHB/AXI для Cortex-M4 или по внешнему интерфейсу PCIe для Cortex-A76.

Таблица 5

**Теоретическая оценка производительности и энергоэффективности выполнения
полиномиального умножения**

Метрика	Endpoint-IoT (Cortex-M4)	Gateway-IoT (Cortex-A76)	NPU-ускоритель (Hailo-8L)	Анализ эффективности (NPU vs CPU)
Режим работы	DSP-расширения	NEON-инструкции	Тензорные вычисления (INT16)	–
Пиковая производительность	$\approx 0,1$ GOPS	$\approx 38,4$ GOPS	≈ 3250 GOPS	Архитектурное преимущество NPU над M4 > 3000 раз, над A76 ≈ 85 раз
Время вычислений (ядро)	$\approx 1\,700$ мкс	≈ 200 мкс	< 1 мкс	Временем вычислений на NPU можно пренебречь
Накладные расходы (интерфейс)	≈ 0 мкс (SRAM/Flash)	≈ 0 мкс (L1 Cache)	$\sim 60\text{--}100$ мкс (зависит от типа шины: AHB/AXI или PCIe)	Узкое место смещается с вычислений на интерфейс передачи данных
Итоговое время операции	$\approx 1\,700$ мкс	≈ 200 мкс	$\approx 60\text{--}100$ мкс	Ускорение: 17х–28х относительно узла, 2х–3х относительно шлюза
Энергия (на 1 операцию)	≈ 170 мкДж	≈ 400 мкДж	$\approx 20\text{--}50$ мкДж	Выигрыш в 3–8 раз по сравнению с CPU
Узкое место	Блокировка ALU	Ширина вектора	Шина данных	Делегирование задач на NPU полностью освобождает CPU

На Gateway-уровне интеграция сопроцессора дает эффект преимущественно за счет асинхронного режима работы. Итоговое время операции здесь упирается не в скорость

вычислений, а в физические задержки при пересылке матриц по внешней шине PCIe. Однако перенос ресурсоемкой математики на тензорные блоки значительно разгружает центральный процессор. Высвобожденные мощности направляются на обслуживание сетевого стека и маршрутизацию пакетов, что кратно увеличивает общую пропускную способность шлюза при массовых подключениях IoT-клиентов.

На Endpoint-уровне применение NPU внутри SoC дает прирост скорости в 17–28 раз. Столь высокий показатель обусловлен изначальной слабостью АЛУ микроконтроллера и низким уровнем интерфейсных ограничений внутри SoC.

Помимо сокращения задержек гетерогенная архитектура решает проблему энергоэффективности. Во время выполнения тензорным ядром криптографических расчетов основной процессор микроконтроллера может перейти в режим сна либо параллельно выполнять другую работу. Для устройств IoT с питанием от батарей такой паттерн работы даст значительное преимущество.

Математическая база алгоритмов Kyber и Dilithium строится на идентичных вычислениях в полиномиальных кольцах. Благодаря этому разработанная модель носит универсальный характер. Она аппаратно ускоряет оба типа криптографических примитивов, комплексно решая проблему ресурсоемкой постквантовой аутентификации в сетях интернета вещей.

Выводы

Проведенное исследование показало, что миграция на программно реализованные алгоритмы LBC на устройства IoT неэффективна и ведет к значительному падению производительности и росту энергопотребления до 8 раз. Для систем интернета вещей, работающих в реальном времени, такие показатели критичны.

В работе предложен способ преодоления этого барьера через использование тензорных ядер нейронных процессоров (NPU). В основу метода легла доказанная математическая эквивалентность полиномиального умножения в кольце R_q и операции одномерной свертки. Это позволяет внедрять постквантовую криптографию (PQC) в современную инфраструктуру IoT без радикальной замены существующей элементной базы.

Несмотря на высокую эффективность, модель имеет ряд ограничений. Требуются дальнейшие исследования в области натурной верификации предложенного подхода, а также разработки специализированных механизмов защиты NPU от атак по побочным каналам, обусловленных архитектурными особенностями тензорных ускорителей.

Литература

1. Shor P.W. (1994) Algorithms for quantum computation: discrete logarithms and factoring; Proceedings 35th annual symposium on foundations of computer science. IEEE. Pp. 124–134.
2. Boixo S. [et al.]. (2018) Characterizing quantum supremacy in near-term devices. Nature Physics. Vol. 14. No. 6. Pp. 595–600.
3. Mosca M., Piani M. (2022) Quantum threat timeline report (2021). Global Risk Institute. 2022.
4. Gidney C., Ekerå M. (2021) How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits. Quantum. Vol. 5. P. 433.

Метод аппаратного ускорения постквантовых криптографических
примитивов на базе нейронных процессоров в IoT

5. *Barker E.* (2020) Recommendation for Key Management: Part 1. (General (Revision 5. NIST Special Publication 800-57. Part 1. Rev. 5. National Institute of Standards and Technology, 2020. Pp. 54–56.
6. *Kannwischer M. J.* [et al.]. (2019) pqm4: Testing and Benchmarking NIST PQC on ARM Cortex-M4. 2019.
7. *Lenngren E.* (2025) Assembly optimised Curve25519 and Curve448 implementations for ARM Cortex-M4 and Cortex-M33. Cryptology ePrint Archive, 2025.
8. Post-Quantum Kyber benchmarks on ARM Cortex-M4: wolfSSL, 2024. URL: <https://www.wolfssl.com/post-quantum-kyber-benchmarks-arm-cortex-m4/> (дата обращения: 26.01.2026).
9. Актуальные киберугрозы: I-II кварталы 2025 года: Positive Technologies, 2025. URL: <https://ptsecurity.com/research/analytics/aktual-nye-kiberugrozy-i-ii-kvartaly-2025-goda/> (дата обращения: 26.01.2026).
10. Threat landscape for industrial automation systems. Q3 2025: Kaspersky ICS CERT, 2025. URL: <https://ics-cert.kaspersky.com/publications/reports/2025/12/11/threat-landscape-for-industrial-automation-systems-q3-2025/> (дата обращения: 26.01.2026).
11. Cost of a Data Breach Report 2024: Financial industry: IBM, 2024. URL: <https://www.ibm.com/think/insights/cost-of-a-data-breach-2024-financial-industry> (дата обращения: 26.01.2026).
12. *Newhouse W.* [et al.]. (2023) Migration to Post-Quantum Cryptography: Preparation for Considering the Implementation and Adoption of Quantum Safe Cryptography: NIST Special Publication 1800-38C: Initial Public Draft. Gaithersburg: NIST, 2023. URL: [https://csrc.nist.gov/pubs/sp/1800/38/iprd-\(1\)](https://csrc.nist.gov/pubs/sp/1800/38/iprd-(1)) (дата обращения: 26.01.2026).
13. *Radanliev P.* [et al.]. (2018) Economic impact of IoT cyber risk-analysing past and present to predict the future developments in IoT risk analysis and IoT cyber insurance. Living in the Internet of Things: Cybersecurity of the IoT-2018, 2018. P. 3.
14. *Catuogno L., Galdi C.* (2023) Secure firmware update: Challenges and solutions. Cryptography. Vol. 7. No. 2. P. 30.
15. *Teague T.* (2022) Side-channel analysis on post-quantum cryptography algorithms. 2022.
16. *Kocher P.* [et al.]. (2020) Spectre attacks: Exploiting speculative execution. Communications of the ACM. Vol. 63. No. 7. Pp. 93–101.
17. *Barker W., Souppaya M., Newhouse W.* (2021) Migration to post-quantum cryptography. NIST National Institute of Standards and Technology and National Cybersecurity Center of Excellence. Pp. 1–15.
18. Post-Quantum Cryptography: Integration Study: European Union Agency for Cybersecurity (ENISA). 2022. URL: <https://www.enisa.europa.eu/publications/post-quantum-cryptography-integration-study> (дата обращения: 26.01.2026).
19. Practical Preparations for the Post-Quantum World: Cloud Security Alliance. 2022. URL: <https://cloudsecurityalliance.org/artifacts/practical-preparations-for-the-post-quantum-world> (дата обращения: 26.01.2026).
20. *Kannwischer M.J., Rijneveld J., Schwabe P.* (2019) Faster multiplication in on Cortex-M4 to speed up NIST PQC candidates. International Conference on Applied Cryptography and Network Security. Cham: Springer International Publishing, 2019. Pp. 281–301.
21. *Кравцов Р.Н.* (2025) Реализация и ускорение NTRU с помощью Pytorch. ИТНОУ: Информационные технологии в науке, образовании и управлении. № 1-2(24-25). С. 20–27. DOI: 10.47501/ITNOU.2025.1.20-27

References

1. Shor P.W. (1994) Algorithms for quantum computation: discrete logarithms and factoring: *Proceedings 35th annual symposium on foundations of computer science*. IEEE. Pp. 124–134.
2. Boixo S. [et al.]. (2018) Characterizing quantum supremacy in near-term devices. *Nature Physics*. Vol. 14. No. 6. Pp. 595–600.
3. Mosca M., Piani M. (2022) Quantum threat timeline report (2021). Global Risk Institute, 2022.
4. Gidney C., Ekerå M. (2021) How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits. *Quantum*. Vol. 5. P. 433.
5. Barker E. (2020) Recommendation for Key Management. Part 1. General (Revision 5). National Institute of Standards and Technology, 2020. Pp. 54–56.
6. Kannwischer M. J. [et al.]. (2019) pqm4: Testing and Benchmarking NIST PQC on ARM Cortex-M4. 2019.
7. Lenngren E. (2025) Assembly optimised Curve25519 and Curve448 implementations for ARM Cortex-M4 and Cortex-M33. Cryptology ePrint Archive, 2025.
8. Post-Quantum Kyber benchmarks on ARM Cortex-M4: wolfSSL, 2024. Available at: <https://www.wolfssl.com/post-quantum-kyber-benchmarks-arm-cortex-m4/> (accessed: 26.01.2026).
9. Positive Technologies (2025). Current cyber threats: Q1-Q2 2025. Available at: <https://ptsecurity.com/research/analytics/aktual-nye-kiberugrozy-i-ii-kvartaly-2025-goda/> (accessed: 26.01.2026). (In Russian).
10. Threat landscape for industrial automation systems. Q3 2025: Kaspersky ICS CERT, 2025. Available at: <https://ics-cert.kaspersky.com/publications/reports/2025/12/11/threat-landscape-for-industrial-automation-systems-q3-2025/> (accessed: 26.01.2026).
11. Cost of a Data Breach Report 2024: Financial industry. IBM, 2024. Available at: <https://www.ibm.com/think/insights/cost-of-a-data-breach-2024-financial-industry> (accessed: 26.01.2026).
12. Newhouse W. [et al.]. (2023) Migration to Post-Quantum Cryptography: Preparation for Considering the Implementation and Adoption of Quantum Safe Cryptography: NIST Special Publication 1800-38C: Initial Public Draft. Gaithersburg. NIST, 2023. Available at: [https://csrc.nist.gov/pubs/sp/1800/38/iprd-\(1\)](https://csrc.nist.gov/pubs/sp/1800/38/iprd-(1)) (accessed: 26.01.2026).
13. Radanliev P. [et al.]. (2018) Economic impact of IoT cyber risk-analysing past and present to predict the future developments in IoT risk analysis and IoT cyber insurance. *Living in the Internet of Things: Cybersecurity of the IoT*. 2018. P. 3.
14. Catuogno L., Galdi C. (2023) Secure firmware update: Challenges and solutions. *Cryptography*. Vol. 7. No. 2. P. 30.
15. Teague T. (2022) Side-channel analysis on post-quantum cryptography algorithms. 2022.
16. Kocher P. [et al.]. (2020) Spectre attacks: Exploiting speculative execution. *Communications of the ACM*. Vol. 63. No. 7. Pp. 93–101.
17. Barker W., Souppaya M., Newhouse W. (2021) Migration to post-quantum cryptography. NIST National Institute of Standards and Technology and National Cybersecurity Center of Excellence, 2022. Pp. 1–15.
18. Post-Quantum Cryptography: Integration Study. European Union Agency for Cybersecurity (ENISA), 2022. Available at: <https://www.enisa.europa.eu/publications/post-quantum-cryptography-integration-study> (accessed: 26.01.2026).
19. Practical Preparations for the Post-Quantum World. Cloud Security Alliance, 2022. Available at: <https://cloudsecurityalliance.org/artifacts/practical-preparations-for-the-post-quantum-world> (accessed: 26.01.2026).

Метод аппаратного ускорения постквантовых криптографических
примитивов на базе нейронных процессоров в IoT

20. Kannwischer M.J., Rijneveld J., Schwabe P. (2019) Faster multiplication in on Cortex-M4 to speed up NIST PQC candidates: International Conference on Applied Cryptography and Network Security. Cham: Springer International Publishing. Pp. 281–301.
21. Kravtsov R.N. (2025). Implementation and acceleration of NTRU using Pytorch. *ITNOU: Information Technologies in Science, Education and Management*. No. 1-2(24-25). Pp. 20–27. DOI: 10.47501/ITNOU.2025.1.20-27 (In Russian).

Поступила в редакцию: 07.07.2026

Received: 07.07.2026

Поступила после рецензирования: 27.08.2026

Revised: 27.08.2026

Принята к публикации: 09.09.2026

Accepted: 09.09.2026