

Научно-исследовательский журнал «Экономический вестник / Economic Bulletin»
<https://eb-journal.ru>
2025, Том 4 № 6 2025, Vol. 4, Iss. 6 <https://eb-journal.ru/archives/category/publications>
Научная статья / Original article
УДК 338.24



¹ Рахматуллин Т.Г.,
¹ Mytona, г. Санкт-Петербург

**Модель оценки экономической эффективности гибридной потоковой обработки данных
в высоконагруженных системах: методология снижения операционных задержек**

Аннотация: целью исследования является раскрытие авторской методологии снижения операционных задержек и повышения экономической эффективности при применении гибридной потоковой обработки данных в высоконагруженных системах.

Методы: в качестве методов исследования использованы анализ и синтез научных источников в области потоковой обработки данных и экономической оценки ИТ-инфраструктуры, сравнительный анализ архитектурных подходов (Batch Processing, Streaming-only, гибридные архитектуры), экономическое моделирование предотвращенных убытков и оптимизации выручки, а также кейс-стади, основанный на моделировании типовых инцидентов эксплуатации высоконагруженных цифровых продуктов.

Результаты (Findings): в исследовании предложена авторская модель оценки экономического эффекта гибридной потоковой обработки данных, которая учитывает предотвращенные убытки, оптимизацию принимаемых решений и совокупную стоимость владения инфраструктурой. Доказано, что применение гибридной архитектуры в соответствии с Lambda Architecture обеспечивает линейный характер повышения инфраструктурных затрат при увеличении нагрузки, устраняет эффект информационной слепоты и позволяет существенно снизить прямые и косвенные экономические потери по сравнению с пакетными и проприетарными SaaS-решениями аналитики.

Выводы: полученные результаты подтверждают, что экономическая эффективность потоковой обработки данных определяется не столько максимальной технической производительностью, сколько соответствием архитектуры профилю бизнес-рисков и управленческих задач. Разработанная модель может использоваться в качестве инструмента обоснования инвестиций в аналитические системы и оценки ROI цифровых трансформаций в высоконагруженных системах.

Ключевые слова: гибридная потоковая обработка данных, экономическая эффективность, высоконагруженные системы, операционные задержки, информационная слепота бизнеса (Data Blindness), совокупная стоимость владения, Lambda Architecture, предотвращенные убытки

Для цитирования: Рахматуллин Т.Г. Модель оценки экономической эффективности гибридной потоковой обработки данных в высоконагруженных системах: методология снижения операционных задержек // Экономический вестник. 2025. Том 4. № 6. С. 55 – 64.

Поступила в редакцию: 15 сентября 2025 г.; Одобрена после рецензирования: 11 ноября 2025 г.; Принята к публикации: 28 декабря 2025 г.

¹ Rakhmatullin T.G.,
¹ Mytona, St. Petersburg

**A model for assessing the economic efficiency of hybrid stream data processing
in high-load systems: a methodology for reducing operational latency**

Abstract: the purpose of this study is to present an author's methodology for reducing operational latency and increasing economic efficiency through the application of hybrid stream data processing in high-load systems.

Methods: the research employs analysis and synthesis of scientific literature in the fields of stream data processing and economic assessment of IT infrastructure, comparative analysis of architectural approaches (Batch Pro-

cessing, streaming-only, and hybrid architectures), economic modeling of prevented losses and revenue optimization, as well as a case study based on modeling typical operational incidents in high-load digital products.

Findings: the study proposes an original model for assessing the economic effect of hybrid stream data processing that accounts for prevented losses, optimization of managerial decision-making, and the total cost of ownership of infrastructure. It is demonstrated that the application of a hybrid architecture based on the Lambda Architecture concept ensures a linear growth of infrastructure costs as workload increases, eliminates the effect of business data blindness, and significantly reduces both direct and indirect economic losses compared to batch processing and proprietary SaaS analytics solutions.

Conclusions: the findings confirm that the economic efficiency of stream data processing is determined not by maximum technical performance alone, but by the alignment of architectural solutions with business risk profiles and managerial objectives. The proposed model can be used as a tool for justifying investments in analytical systems and for evaluating the ROI of digital transformation initiatives in high-load environments.

Keywords: hybrid stream data processing, economic efficiency, high-load systems, operational latency, business data blindness (Data Blindness), total cost of ownership, Lambda Architecture, prevented losses

For citation: Rakhmatullin T.G. A model for assessing the economic efficiency of hybrid stream data processing in high-load systems: a methodology for reducing operational latency. Economic Bulletin. 2025. 4 (6). P. 55 – 64.

The article was submitted: September 15, 2025; Approved after reviewing: November 11, 2025; Accepted for publication: December 28, 2025.

Введение

В реалиях современной цифровой экономики, характеризующейся повышенным объемом обмена данными, активной цифровой трансформацией и развитием платформ, важным фактором обеспечения конкурентоспособности становится скорость принятия решений и реагирования на события. Информационные системы функционируют в режиме реального времени и высокой интенсивности входящих потоков данных; закономерно, задержки их обработки становятся обстоятельством, которое сказывается не только на общей производительности системы, но и на ее экономической эффективности.

В частности, стоит подчеркнуть факт того, что традиционные архитектуры потоковой обработки данных, основанные на масштабируемых распределенных вычислительных кластерах, в целом позволяют достичь требуемого уровня пропускной способности. Однако при их масштабировании нередко проявления эффектов роста затрат, связанных с избыточным резервированием ресурсов и неэффективным использованием вычислительной инфраструктуры [6; 7]. Кроме того, традиционные подходы к аналитике, основанные на суточной пакетной обработке (Batch Processing), создают выраженный временной лаг между событием (например, сбоем в транзакции или запуском рекламной кампании) и реакцией менеджмента. В среднем подобный временной лаг длится в диапазоне от 24 до 48 часов и характеризуется высокой латентностью принятия решений (Decision Latency). В индустриальной практике данный феномен также описывается термином «информационная слепота» (Data Blindness), при которой принимаемые в компании решения основываются на фактически устаревших данных, что, в случае принятия неверных или менее оптимальных решений, способствует возникновению в лучшем случае упущенной выгоды, а при худшем сценарии сводится к прямым убыткам. В ответ на данные ограничения на протяжении последних лет активное развитие получили гибридные архитектурные подходы, под началом которых объединяются облачные и локальные вычислительные ресурсы, а также гетерогенные вычислительные компоненты (CPU, FPGA и специализированные ускорители) [2; 8].

Как отмечают в своем исследовании Е.А. Сычев и А.О. Мусатов, попытка решить проблемы масштабируемости на устаревших монолитных архитектурах нередко приводит к накоплению технического долга, который блокирует инновации и экспоненциально увеличивает затраты на поддержку систем. Следовательно, бизнесу требуется формировать новый архитектурный подход, который позволит снизить задержки без критического повышения инфраструктурных затрат [11]. Одновременно с этим возрастает и потребность в применении экономических моделей оценки потоковой обработки, ориентированных не только на показатели задержки и пропускной способности, но и на совокупную стоимость владения, стоимость масштабирования и чувствительность архитектуры к изменяющимся нагрузкам [6; 9; 10].

В связи с рядом вышеприведенных обстоятельств, все более актуальной и научно обоснованной задачей становится разработка модели оценки экономической эффективности гибридной потоковой обработки данных, направленной на снижение операционных задержек при контролируемых издержках.

Материалы и методы исследований

Исследование основано на открытых материалах авторов научных статей, посвященных вопросам потоковой обработки данных в высоконагруженных системах и затрагивающих вопросы повышения их эффективности. Работа ориентирована на конкретизацию как информационно-технологических, инженерных, так и экономических данных при применении гибридной системы.

В частности, основываясь на современных исследованиях в области потоковой обработки данных, стоит подчеркнуть, что постепенно происходит смещение фокуса с сугубо технических характеристик систем к экономическим последствиям их применения в деятельности бизнеса, прежде всего выраженных во влиянии задержек обработки данных на принимаемые внутри системы управления решения и финансовые результаты организаций.

Так, например, в исследовании О.С. Дробковой и Д.М. Мирохиной справедливо указывается, что развитие архитектур хранения и обработки больших данных, в частности Data Lake, обусловлено не только повышением объемов информации, но и потребностью бизнеса в снижении издержек доступа к данным при расширении возможностей их бизнес-аналитики. Авторы считают, что Data Lake экономически эффективнее классических хранилищ ввиду эффектов снижения стоимости хранения и инвариантности аналитических сценариев; вместе с тем авторами косвенно фиксируется ограничение подобных решений, поскольку данные при отсутствии потоковой обработки (несмотря на доступность) остаются временно устаревшими, что делает их менее ценными для применения под нужды управления [1]. Данный вывод соотносится с понятием «стоимости информации во времени» и упомянутым ранее феноменом Data Blindness, используемым в настоящем исследовании.

Экономика потоковой обработки детально раскрыта в исследовании С. Hochreiner, М. Vögler, S. Schulte и S. Dustdar, в котором доказывается, что традиционные модели масштабирования распределенных потоковых систем, ориентированные исключительно на показатели загрузки ресурсов, приводят к повышению операционных затрат и нередко становятся основанием возникновения скрытых издержек. В связи с этим авторы вводят понятие стоимости масштабирования как самостоятельного экономического фактора, сводящегося к издержкам миграции, перераспределению ресурсов и кратковременным потерям производительности [7]; основываясь на тезисах авторов, стоит подчеркнуть, что низкая задержка обработки не всегда эквивалентна высокой экономической эффективности системы.

В частности, развитие данной идеи представлено в исследовании М. Borkowski, С. Hochreiner и S. Schulte, которыми выявляется, что чрезмерная реактивность систем потоковой обработки данных перед краткосрочными всплесками нагрузки увеличивает совокупную стоимость владения (или TCO) и может негативно влиять на соблюдение SLA. В результате авторы приходят к выводу, что снижение количества операций масштабирования позволяет минимизировать операционные издержки без потери качества сервиса [6], что с экономической точки зрения подтверждает тезис о необходимости оценки не только технической производительности, но и предельной полезности ускорения обработки данных.

С позиции архитектурных решений особое внимание необходимо уделить исследованию именно гибридных моделей обработки данных; так, в работе Р. Pérez-Arteaga и соавторов на примере Lambda-архитектуры показано, что комбинирование потокового и пакетного слоев обработки позволяет снизить TCO аналитической инфраструктуры за счет разграничения данных по уровню их экономической значимости. Авторы подмечают, что реальное время обработки должно применяться исключительно к тем данным, задержка по которым приводит к прямым или косвенным финансовым потерям, а историческая аналитика может обрабатываться в более экономичном пакетном режиме [10]. Схожие доводы приводит в своей работе и Р. Kodakandla, который рассматривает гибридные системы в качестве инструмента балансировки между производительностью и экономической эффективностью; автор указывает, что облачные ресурсы целесообразны для обработки пиковых и критичных по времени потоков, а стабильные нагрузки экономически выгоднее обслуживать на предсказуемой инфраструктуре [8]. Как показывает сравнительный анализ облачно-нативных и гибридных потоковых конвейеров, представленный М. Lee, эмпирически подтверждается эффективность гибридных архитектур при длительных высоконагруженных сценариях, поскольку они обеспечивают более низкую стоимость обработки одного события при сопоставимых показателях задержки. Иными словами, делается вывод о том, что экономическая эффективность потоковой обработки определяется не максимальной производительностью, а соответствием архитектуры профилю бизнес-нагрузки [9], что усиливает аргументацию в пользу моделей оценки, ориентированных на TCO и предотвращенные

убытки. В контексте издержек на управление подобное сводится к возможности контролируемого повышения затрат, что критично для расчета ROI цифровых трансформаций. При этом, учитывая результаты представленного литературного обзора, отметим, что вопросы оптимизации обработки данных раскрывались автором в предыдущих исследованиях [3-5], на основании результатов которых можно подтвердить, что автоматизация ETL-процессов с использованием Apache Airflow, оптимизация шардирования и индексирования в MongoDB, а также разработка масштабируемых решений с Apache Flink и Kafka позволяют в целом снизить накладные расходы на обработку данных и повысить предсказуемость времени доставки аналитики. Таким образом, экономическая эффективность гибридной системы формируется не только на уровне вычислительного ядра, но и за счет оптимизации сопутствующих процессов обработки и хранения данных.

В совокупности по итогам литературного обзора формируется выраженное противоречие, представленное тем, что в современных исследованиях накоплен существенный объем знаний о технических способах снижения задержек и оптимизации потоковой обработки. Однако большинство работ рассматривает экономический эффект фрагментарно через снижение затрат на инфраструктуру или соблюдение SLA. В то же время недостаточно разработан интегральный подход, который бы учитывал предотвращенные убытки, эффекты повышения эффективности управления и устранения Data Blindness как самостоятельного источника экономического эффекта. Данный пробел и определяет цель, границы, а также научную новизну настоящего исследования.

Цель исследования – раскрыть авторскую методологию снижения операционных задержек и повышения экономической эффективности при применении гибридной потоковой обработки данных в высоконагруженных системах.

Основанием для оценки экономической эффективности гибридной модели обработки данных послужил Lambda Architecture подход, предполагающий разделение потоков данных на два уровня, с разной экономической стоимостью. Первым является уровень Speed Layer (или Speed-слой; скоростной слой), который направлен на обработку данных в реальном времени и используется в целях оперативного реагирования (LiveOps); соответственно, на техническом уровне данный слой представляет собой технологии потоковой обработки, экономическая эффективность которых доказательно подтверждается по результатам предыдущих исследований [4]. Вторым является пакетный слой (или Batch-слой; от англ. Batch Layer), который обеспечивает эталонную точность и хранение исторических данных, отличается низкой удельной стоимостью хранения, и реализуется на основании применения инструментов оркестрации (по типу Apache Airflow). Комбинированный подход, основанный на данных слоях, предполагал их совместное применение для разделения потоков данных на основании экономической стоимости.

Тем самым в основу исследования положена гипотеза, предполагающая, во-первых, что совокупная стоимость владения гибридной системой (или от англ. TCO Lambda Architecture) окажется ниже, чем TCO у потоковой системы, при условии сброса «холодных» данных в дешевый Batch-слой; во-вторых, что экономический эффект гибридной системы будет выше, чем у пакетной, поскольку будут предотвращаться убытки в Speed-слое.

В целях оценки представленной гипотезы применяется формула расчета экономического эффекта (далее – E) от внедрения гибридной системы аналитики: $E = (L_{prevented} + R_{opt}) - (C_{infra} + C_{dev})$; где: $L_{prevented}$ (Loss Prevented) – предотвращенные убытки за счет сокращения времени реакции на инциденты; R_{opt} (Revenue Optimization) – дополнительная выручка от оптимизации маркетинга (UA) в реальном времени; C_{infra} – стоимость инфраструктуры (серверы, облачные ресурсы); C_{dev} – затраты на разработку и поддержку. При расчете параметра C_{dev} в модели учитывается специфика Lambda Architecture, которая предполагает увеличение трудозатрат на поддержку двух кодовых баз (пакетной и потоковой). Однако в высоконагруженных системах (High-Load) данная статья расходов становится фиксированной величиной, которая при масштабировании нагрузки размывается. В то же время инфраструктурные расходы (C_{infra}) в полностью потоковых архитектурах (Streaming-only) возрастают экспоненциально из-за требований к отказоустойчивости stateful-операций. За счет использования данной формулы удалось доказать, что при определенном пороге нагрузки экономия на инфраструктуре многократно перекрывает повышенные затраты на штат инженеров.

Результаты и обсуждения

Итак, отдельного внимания заслуживают вопросы влияния технологических факторов на снижение TCO, в чем сокращение издержек достигается путем выбора технологий, которые позволяют масштабировать систему линейно, а не экспоненциально. В этой связи отдельного внимания заслуживают следующие аспекты:

1. Использование Redis для кэширования «горячих» данных, которое позволяет снизить нагрузку на дорогие основные СУБД, что закономерно сокращает потребность в использовании вычислительных мощностей.

2. Применение стратегий шардирования в MongoDB, которое позволяет использовать горизонтальное масштабирование (commodity hardware) вместо покупки дорогостоящих вертикальных серверов.

На практике использование данных решений обеспечивает предсказуемость бюджета; как показывают результаты расчетов, при увеличении нагрузки в 6,5 раз (с 1000 до 6500 событий в секунду) стоимость инфраструктуры возрастает именно линейно, без скрытых комиссий, характерных для SaaS-решений. Для сравнения, проприетарные решения (например, Firebase или Amplitude) при превышении лимитов вводят экспоненциальные тарифы или сэмплирование данных, что снижает их экономическую эффективность для бизнеса (при возникновении потребности в обработке большего количества событий).

Тем самым проведенное исследование позволило перейти от теоретического обоснования экономической целесообразности гибридной потоковой обработки данных к практической верификации выдвинутой гипотезы и анализу полученных эффектов в контексте высоконагруженных цифровых систем (рис. 1).

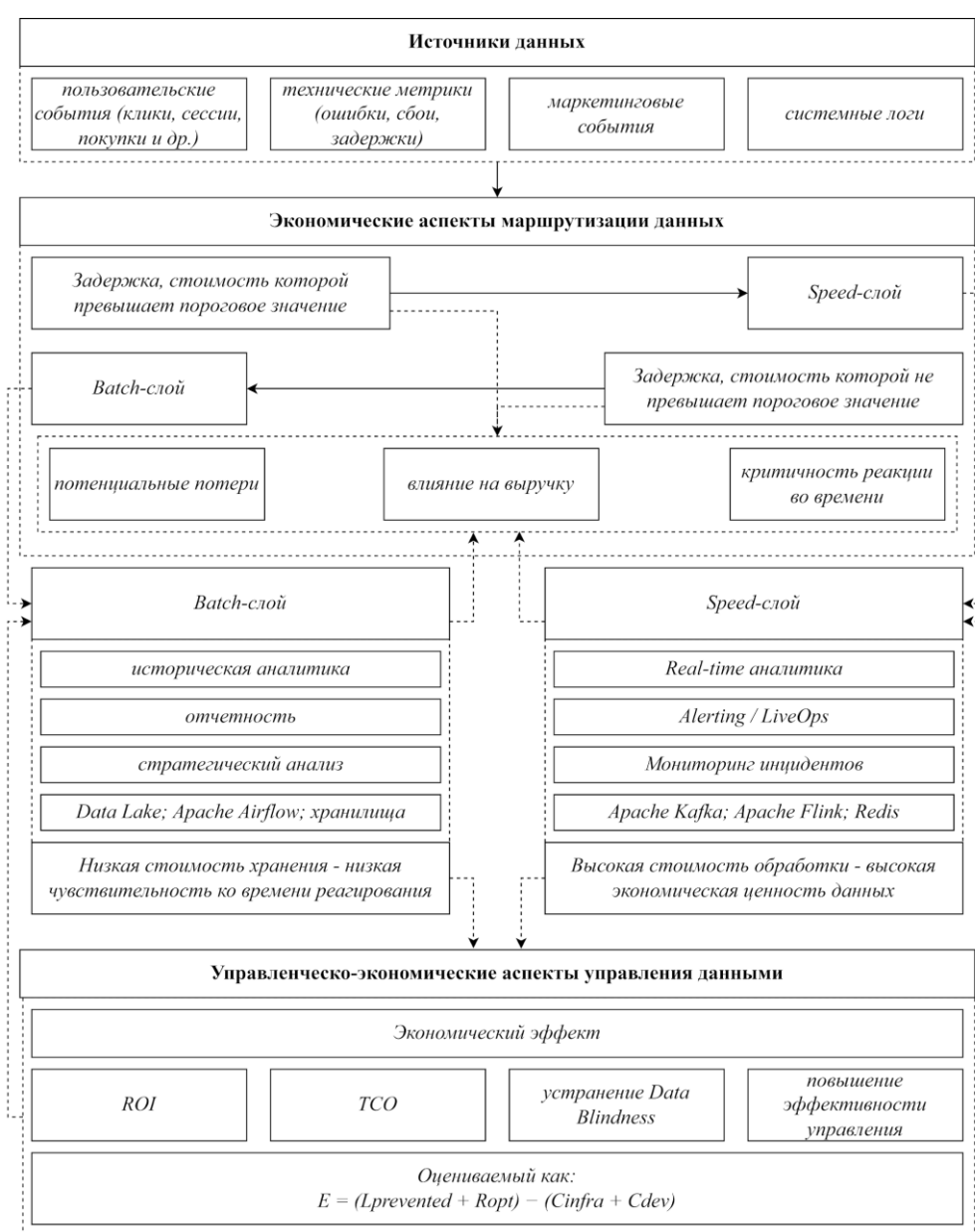


Рис. 1. Экономически ориентированная схема гибридной потоковой обработки данных.
Fig. 1. Economically oriented architecture of hybrid stream data processing.

Апробация предложенной модели была проведена в крупной международной IT-компании, оперирующей цифровыми продуктами с аудиторией более 100 млн. пользователей. Внедрение системы позволило сократить время доставки данных (Data Freshness) с 24 часов до 2-5 минут, что привело к некоторым выраженным экономическим эффектам.

В первую очередь, подтверждаются метрики эффективности – за счет сокращения времени доставки данных удалось полностью устранить эффект информационной слепоты, в связи с чем количество расхождений в отчетах достигло значения в менее 1% (при 10% до изменений). Одновременно с этим увеличилось использование аналитического портала сотрудниками, и достигло 40% за счет доверия к системе.

С другой стороны, проявился выраженный экономический эффект, полученный на основании моделирования инцидентов. Приведем один из таких инцидентов. Критический сбой (Crash Rate – показатель сбоев свыше 2%) в продукте с активной аудиторией. На основе данных компании, расчетные потери выручки при таком сбое составляют около 3200 долл. США в час. При обнаружении данных в течение 24 часов совокупный прямой убыток составляет (при текущей стоимости часа потерь) не менее 76800 долл. США. В случае гибридной системы и обнаружения за 20 минут (при эффективности системы в 5 минут), убыток составляет 1066 долл. США, т.е. гибридная архитектура позволяет предотвратить около 75000 долл. США убытков лишь в одном инциденте (рис. 2).

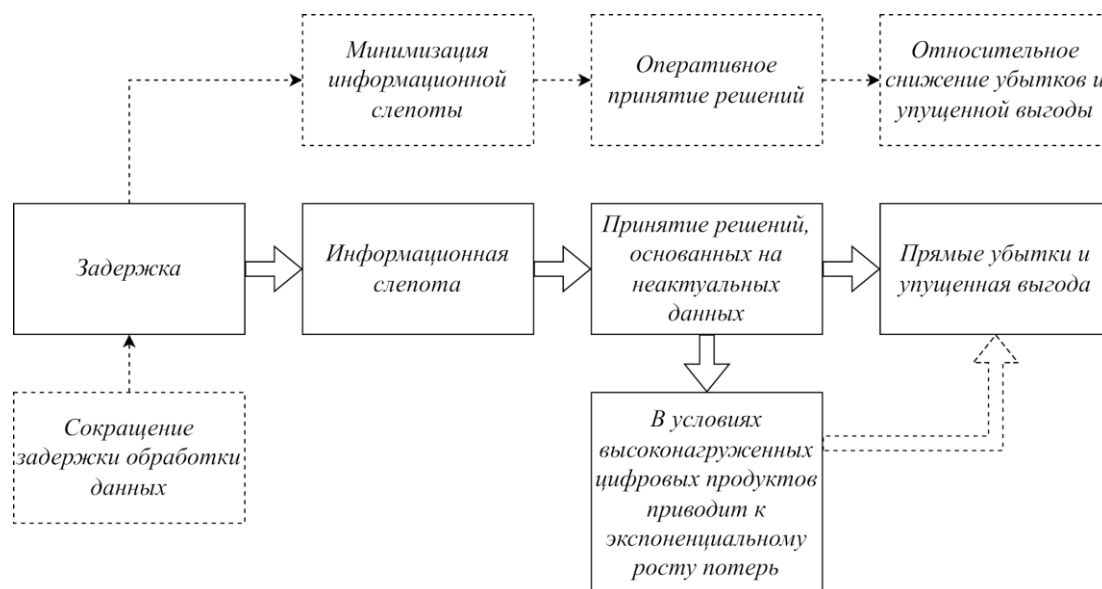


Рис. 2. Влияние задержек обработки данных на экономические потери бизнеса.

Fig. 2. Impact of data processing latency on business economic losses.

Схожие экономические эффекты выражаются в контексте расчета оптимизации маркетинга – за счет оперативного отключения убыточного канала трафика в первые часы запуска (благодаря LTV-прогнозу первого дня), отказ от ожидания недельного отчета экономит до 15-30% маркетингового бюджета на тестовый период.

Отметим, что в целом в основе интерпретации полученных результатов лежит сопоставление трех принципиально различных архитектурных подходов к обработке данных:

- 1) традиционной пакетной модели (Batch Processing);
- 2) полностью потоковой модели (Streaming-only);
- 3) гибридной модели, реализованной в виде системы Lambda Architecture с разграничением потоков по экономической значимости.

Отдельно необходимо обосновать выбор Lambda Architecture в противовес набирающей популярность Карпа Architecture, которая предполагает использование единого потокового лога (например, Kafka) как единственного источника истины. Сравнительный анализ показал, что для исследуемых высоконагруженных систем Карпа Architecture обладает существенным экономическим недостатком при сценариях «глубокого пересчета» данных (Reprocessing). В Карпа Architecture исторические данные хранятся в системе, оптимизированной под последовательный потоковый доступ, а не под экономичное долгосрочное хранение,

соответственно полная перечитка (Replay) петабайтов данных через потоковый движок требует использованиякратно больших вычислительных мощностей и времени, чем пакетная обработка тех же данных из специализированных «холодных» хранилищ (S3/HDFS), что реализовано в Batch-слое Lambda Architecture.

При этом основное внимание уделяется не абсолютному снижению задержек как технического показателя, а их трансформации в измеримый экономический эффект, выраженный через предотвращенные убытки, оптимизацию выручки и изменение структуры совокупной стоимости владения аналитической инфраструктурой.

Прежде всего, результаты подтверждают справедливость выдвинутого положения о том, что задержка обработки данных является не нейтральным техническим параметром, а экономически нагруженной переменной, которая напрямую влияет на финансовые показатели бизнеса. Сокращение времени доставки аналитических данных с суточного интервала до диапазона 2-5 минут принципиально трансформирует процесс принятия решений, поскольку в пакетной модели аналитика носит преимущественно ретроспективный характер, а в гибридной архитектуре данные становятся инструментом оперативного управления (то есть решения принимаются быстро). Таким образом, достигаются качественные изменения, с переходом от описательной аналитики к аналитике превентивной, что принципиально важно для современных цифровых продуктов ввиду их высокой чувствительности к сбоям и поведению пользователей.

Экономический эффект от внедрения гибридной модели был рассчитан с использованием предложенной формулы. Так, по проведенной оценке, наибольший вклад в суммарный эффект вносит именно компонент предотвращенных убытков, возникающих вследствие сокращения времени реакции на инциденты. Моделирование типового технического сбоя с повышением crash rate позволило установить, что при пакетной аналитике обнаружение проблемы происходило спустя сутки, что приводило к накоплению значительных финансовых потерь. В условиях гибридной архитектуры инцидент фиксировался в течение первых десятков минут, что позволяло локализовать проблему до достижения критического масштаба. Таким образом, экономическая ценность гибридной обработки формируется ввиду снижения масштабов потенциального ущерба в связи с влиянием фактора времени, что подтверждает применимость концепции стоимости информации во времени.

При этом, в целях обоснования конкурентных преимуществ гибридной модели был произведен сравнительный анализ с распространенными SaaS-решениями аналитики. Результаты представлены в табл. 1:

Таблица 1
Сравнительная характеристика аналитических решений.

Table 1

Comparative characteristics of analytical solutions.

Показатель	Гибридная система	Firestore Analytics	Amplitude Enterprise
Максимальная нагрузка, RPS	Свыше 15000	Около 1000	5000
Задержка аналитики	Менее 5 мин	Менее 30 мин	Менее 10 мин
Задержка алертов	Менее 1 мин	Не поддерживается	Менее 5 мин
Хранение данных	Неограниченное	13 месяцев	Свыше 5 лет
Масштабирование	Линейное	ограниченное	условно-линейное

Отдельного обсуждения заслуживают результаты анализа совокупной стоимости владения инфраструктурой. Сравнение гибридной архитектуры с полностью потоковой моделью позволяет утверждать, что попытка обрабатывать весь массив данных исключительно в режиме реального времени приводит к нерациональному повышению затрат. Высокочастотная обработка «холодных» данных, которые не обладают экономической ценностью в текущий момент времени, формирует избыточную нагрузку на вычислительные ресурсы и увеличивает стоимость масштабирования. Гибридный подход, напротив, позволяет сегментировать потоки данных по их влиянию на финансовый результат, направлять критичные события в скоростной слой, а исторические и агрегированные данные в более экономичный Batch-слой. В результате достигается снижение ТСО за счет перераспределения нагрузки и использования дешевых хранилищ и оркестрации пакетных процессов.

В рамках обсуждения архитектурных эффектов целесообразно выделить следующие механизмы, обеспечивающие экономическую эффективность гибридной модели:

– во-первых, линейный характер масштабирования, при котором рост нагрузки сопровождается пропорциональным увеличением затрат без скачкообразных эффектов (рис. 3);

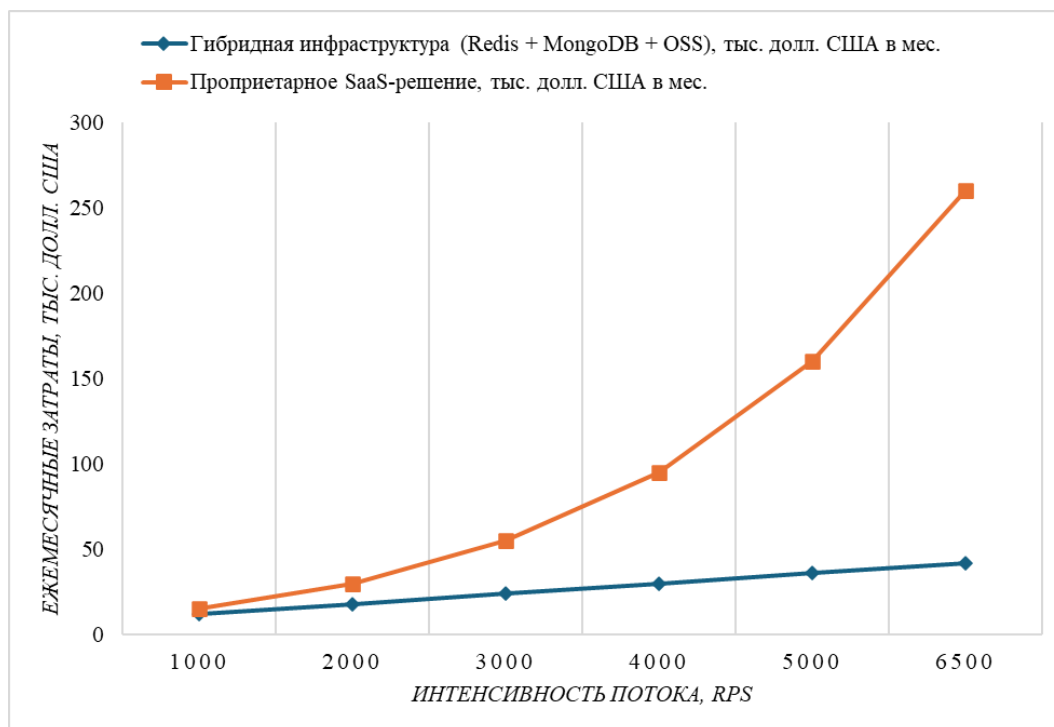


Рис. 3. Зависимость инфраструктурных затрат от интенсивности потока событий.
Fig. 3. Relationship between infrastructure costs and event stream intensity.

Так, на рис. 3 представлена зависимость инфраструктурных затрат от интенсивности потока событий для гибридной архитектуры и проприетарного SaaS-решения. Полученные данные доказывают, что при увеличении нагрузки гибридная инфраструктура обеспечивает практически линейное повышение затрат, в то время как SaaS-решения характеризуются экспоненциальным удорожанием после превышения тарифных лимитов. Тем самым использование открытых технологий (Redis, MongoDB с шардированием, OSS-стек потоковой обработки) формирует предсказуемую модель TCO и снижает инфраструктурные риски при эксплуатации высоконагруженных систем.

- во-вторых, отсутствие скрытых комиссий и лимитов, характерных для проприетарных SaaS-решений;
- в-третьих, снижение зависимости от внешних поставщиков и повышение предсказуемости бюджета.

Приведенные факторы в совокупности формируют условия для управляемой цифровой трансформации, при которой инвестиции в аналитику поддаются экономическому планированию и оценке ROI.

Обобщая полученные результаты, можно констатировать, что предложенная модель оценки экономической эффективности гибридной потоковой обработки данных подтверждает выдвинутую гипотезу исследования. С одной стороны, совокупная стоимость владения гибридной системой оказывается действительно ниже, чем у полностью потоковой архитектуры, что достигается путем исключения нецелесообразной обработки данных с низкой экономической значимостью. С другой стороны, экономический эффект гибридной модели существенно превосходит показатели пакетной аналитики благодаря предотвращению убытков и повышению эффективности принимаемых решений. Тем самым гибридная архитектура выступает оптимизационным решением, которое позволяет одновременно снижать задержки и контролировать издержки.

В контексте научного обсуждения полученные результаты позволяют сформулировать более обширный вывод, выраженный в том, что экономическая эффективность систем потоковой обработки данных определяется не максимальной технической производительностью, а степенью соответствия архитектурных решений структуре бизнес-рисков и профилю задач, решаемых системой управления. Гибридная модель в данном случае выступает инструментом институционализации экономической сущности аналитики, что позволяет переводить данные из вспомогательного ресурса в актив, который обладает измеримой стоимостью во времени.

Выводы

Таким образом, в рамках проведенного исследования была разработана и апробирована модель оценки экономической эффективности гибридной потоковой обработки данных в высоконагруженных системах, ориентированная на снижение операционных задержек при контролируемых инфраструктурных издержках. В отличие от преобладающих в научной литературе подходов, которые фокусируются преимущественно на технических показателях производительности и масштабируемости, предложенная модель учитывает экономическую значимость времени обработки данных, предотвращенные убытки, эффекты оптимизации принятия решений и совокупную стоимость владения инфраструктурой. Полученные результаты подтверждают выдвинутую гипотезу о том, что гибридная архитектура, реализованная в Lambda Architecture, обеспечивает более высокий экономический эффект по сравнению с традиционными пакетными и полностью потоковыми решениями. Практическая апробация модели в условиях реальной эксплуатации высоконагруженного цифрового продукта показала существенное сокращение времени доставки аналитики, устранение эффекта информационной слепоты, снижение прямых финансовых потерь и повышение эффективности использования маркетинговых бюджетов. Установлено, что основополагающим фактором эффективности гибридной модели является линейный характер масштабирования инфраструктурных затрат, который обеспечивает предсказуемость ТСО и управляемость инвестиций в аналитику. В совокупности изложенное обстоятельство позволяет рассматривать гибридную потоковую обработку данных не только как технологическое, но и как экономически обоснованное управленческое решение, обеспечивающее повышение конкурентоспособности организаций в условиях цифровой экономики.

Список источников

1. Дробкова О.С., Мирохина Д.М. Применение технологии Data Lake как способ повышения эффективности деятельности промышленных предприятий // Вопросы инновационной экономики. 2024. Т. 14. № 4. С. 1381 – 1400. DOI: 10.18334/vinec.14.4.122269
2. Радионова Е.А. Гибридная вычислительная архитектура CPU/FPGA для задач потоковой обработки данных // Актуальные научные исследования: сборник статей XXIV Международной научно-практической конференции. Пенза: МЦНС «Наука и Просвещение», 2025. С. 50 – 53.
3. Рахматуллин Т.Г. Автоматизация ETL-процессов с использованием Apache Airflow // Актуальные исследования. 2024. № 8 (190). Ч. I.
4. Рахматуллин Т.Г. Оптимизация потоковой обработки игровых метрик с использованием Apache Flink и Kafka: опыт разработки масштабируемых решений // Актуальные исследования. 2024. № 45 (227). Ч. I. С. 33 – 39.
5. Рахматуллин Т.Г. Оптимизация работы с большими данными в MongoDB: стратегии шардирования и индексирования // Актуальные исследования. 2024. № 50 (232). Ч. I. С. 41 – 46.
6. Borkowski M., Hochreiner C., Schulte S. Minimizing cost by reducing scaling operations in distributed stream processing // Proceedings of the VLDB Endowment. 2019. Vol. 12. No. 7. P. 724 – 737. DOI: 10.14778/3317315.3317316
7. Hochreiner C., Vögler M., Schulte S., Dustdar S. Cost-efficient enactment of stream processing topologies // PeerJ Computer Science. 2017. Vol. 3. Article e141. DOI: 10.7717/peerj-cs.141.
8. Kodakandla P. Balancing performance and economics in hybrid cloud data architectures // International Journal for Research Trends and Innovation. 2022. Vol. 7, no. 2. P. 135 – 140.
9. Lee M. Cost-efficient stream processing architectures: comparative analysis of cloud-native and hybrid Kafka-Spark-BigQuery pipelines // 2024. August. (Научная статья).
10. Pérez-Arteaga P., Castellanos C., Castro H., Correal D., Guzmán L., Denneulin Y. Cost comparison of Lambda architecture implementations for transportation analytics using public cloud software as a service // Proceedings of the 13th International Conference on Software Technologies (ICSOFTE 2018). 2018. P. 855 – 862. DOI: 10.5220/0006869308550862.
11. Sychev Y.A., Musatov A.O. Incremental refactoring to reduce technical debt: migrating from an MVC monolith to microservice APIs // International Scientific Journal of Innovative Science. 2025. No. 4 (1.00). P. 28 – 30.

References

1. Drobkova O.S., Mirokhina D.M. Application of Data Lake Technology as a Way to Improve the Efficiency of Industrial Enterprises. *Issues of Innovative Economics*. 2024. Vol. 14. No. 4. P. 1381 – 1400. DOI: 10.18334/vinec.14.4.122269
2. Radionova E.A. Hybrid CPU/FPGA Computing Architecture for Streaming Data Processing Tasks. *Current Scientific Research: Collection of Articles from the XXIV International Scientific and Practical Conference*. Penza: MCNS "Science and Education", 2025. P. 50 – 53.
3. Rakhmatullin T.G. Automation of ETL Processes Using Apache Airflow. *Current Research*. 2024. No. 8 (190). Part I.
4. Rakhmatullin T.G. Optimization of Stream Processing of Game Metrics Using Apache Flink and Kafka: Experience in Developing Scalable Solutions. *Current Research*. 2024. No. 45 (227). Part I. P. 33 – 39.
5. Rakhmatullin T.G. Optimization of Working with Big Data in MongoDB: Sharding and Indexing Strategies. *Current Research*. 2024. No. 50 (232). Part I. P. 41 – 46.
6. Borkowski M., Hochreiner C., Schulte S. Minimizing Cost by Reducing Scaling Operations in Distributed Stream Processing. *Proceedings of the VLDB Endowment*. 2019. Vol. 12. No. 7. P. 724 – 737. DOI: 10.14778/3317315.3317316
7. Hochreiner C., Vögler M., Schulte S., Dustdar S. Cost-efficient enactment of stream processing topologies. *PeerJ Computer Science*. 2017. Vol. 3. Article e141. DOI: 10.7717/peerj-cs.141.
8. Kodakandla P. Balancing performance and economics in hybrid cloud data architectures. *International Journal for Research Trends and Innovation*. 2022. Vol. 7, no. 2. P. 135 – 140.
9. Lee M. Cost-efficient stream processing architectures: comparative analysis of cloud-native and hybrid Kafka-Spark-BigQuery pipelines. 2024. August. (Scientific article).
10. Pérez-Arteaga P., Castellanos C., Castro H., Correal D., Guzmán L., Denneulin Y. Cost comparison of Lambda architecture implementations for transportation analytics using public cloud software as a service. *Proceedings of the 13th International Conference on Software Technologies (ICSOF 2018)*. 2018. P. 855 – 862. DOI: 10.5220/0006869308550862.
11. Sychev Y.A., Musatov A.O. Incremental refactoring to reduce technical debt: migrating from an MVC monolith to microservice APIs. *International Scientific Journal of Innovative Science*. 2025. No. 4 (1.00). P. 28 – 30.

Информация об авторе

Рахматуллин Т.Г., старший инженер-программист, «Mytona», г. Санкт-Петербург, timur860@mail.ru

© Рахматуллин Т.Г., 2025