

СТАТЬЯ

УДК 004.42:004.738



CC BY 4.0

**МАСШТАБИРУЕМОСТЬ РАСПРЕДЕЛЕННЫХ
ИТ-СИСТЕМ В УСЛОВИЯХ БЫСТРОГО РОСТА:
АНАЛИЗ ФАКТОРОВ УСТОЙЧИВОСТИ****Жиделев Е. О.***Независимый исследователь, Москва, Российская Федерация,
e-mail: evgeniy.zhidelev@gmail.com*

В современном цифровом ландшафте распределенные ИТ-системы занимают ключевую роль, обеспечивая взаимодействие множества вычислительных узлов для предоставления различных сервисов. С ростом спроса растет и нагрузка на эти системы, поэтому расширение их архитектуры становится жизненно важным. Цель статьи – провести анализ особенностей масштабируемости распределенных ИТ-систем в условиях быстрого роста с точки зрения обеспечения их устойчивости. В работе применены методы системного анализа, моделирования и формализации архитектурных решений в распределенных ИТ-системах. В процессе исследования описан переход от монолитных систем к динамичным, адаптивным инфраструктурам, подчеркивая стратегическую важность гибких, интеллектуальных вычислительных сред, способных реагировать на все более сложные технологические требования. В процессе исследования представлена авторская таксономия «устойчивость – задержка – согласованность», которая объединяет эти характеристики с механизмами масштабирования, формируя единую систему взаимных ограничений для анализа архитектурных решений в распределенных системах. Успех масштабируемых распределенных ИТ-систем зависит от разработки комплексных стратегий, которые обеспечивают баланс между технологическими инновациями и практическими эксплуатационными требованиями, создавая инфраструктуру, способные быстро реагировать на меняющиеся вычислительные потребности, и сохраняя при этом оптимальную производительность, безопасность и эффективность. На основе имитационного моделирования установлено, что предложенная таксономия снижает погрешность прогнозирования поведения инфраструктуры с 24,1 до 4,8 % в сравнении с традиционными реактивными и линейными методами управления. Количественно доказано, что учет вектора состояния обеспечивает сохранение устойчивости системы на уровне 0,65 даже при десятикратных пиковых нагрузках. Применение на практике предложенной таксономии позволит существенно повысить точность прогнозирования поведения распределенных ИТ-систем при масштабировании, минимизируя риски архитектурных ошибок за счет учета обратного влияния качественных метрик на пропускную способность системы.

Ключевые слова: распределенная система, масштабирование, география, надежность, задержка, устойчивость, согласованность

**SCALABILITY OF DISTRIBUTED IT-SYSTEMS
IN CONDITIONS OF RAPID GROWTH:
ANALYSIS OF STABILITY FACTORS****Zhidelev E. O.***Independent researcher, Moscow, Russian Federation,
e-mail: evgeniy.zhidelev@gmail.com*

In today's digital landscape, distributed IT systems play a key role, enabling multiple computing nodes to interact in order to provide various services. As demand grows, so does the load on these systems, making scalability vital. The purpose of this article is to analyse the scalability characteristics of distributed IT systems in conditions of rapid growth from the perspective of ensuring their stability. The work applies methods of system analysis, modelling, and formalisation of architectural solutions in distributed IT systems. The research describes the transition from monolithic systems to dynamic, adaptive infrastructures, emphasising the strategic importance of flexible, intelligent computing environments capable of responding to increasingly complex technological requirements. The study presents the author's taxonomy of 'resilience-delay-consistency' as a new structure for evaluating architectural solutions in distributed systems. Unlike existing methodologies, this approach integrates scaling mechanisms with stability, latency, and consistency parameters, combining them into a single system of mutual constraints. Thus, the success of scalable distributed IT systems depends on the development of comprehensive strategies that balance technological innovation and practical operational requirements, creating infrastructures that can respond quickly to changing computing needs while maintaining optimal performance, security, and efficiency. Based on simulation modelling, it has been established that the proposed taxonomy reduces the error in predicting infrastructure behaviour from 24.1 % to 4.8 % compared to traditional reactive and linear control methods. It has been quantitatively proven that taking into account the state vector ensures that the system remains stable at a level of 0.65 even under tenfold peak loads. The practical application of the proposed taxonomy will significantly improve the accuracy of infrastructure behaviour prediction during scaling, minimising the risks of architectural errors by taking into account the reverse impact of quality metrics on system throughput.

Keywords: distributed system, scaling, geography, reliability, latency, stability, consistency

Введение

Распределенные системы представляют собой фундаментальный сдвиг в парадигме современной вычислительной инфраструктуры, характеризующийся совместной работой нескольких сетевых компьютеров в качестве единого комплекса. В отличие от традиционных централизованных подходов, при которых обработка происходит на одном компьютере, распределенные архитектуры делят вычислительные задачи между независимыми узлами, каждый из которых имеет собственные вычислительные мощности и локальное хранилище [1]. Этот децентрализованный подход становится все более важным по мере расширения цифровых услуг для удовлетворения растущих потребностей пользователей на глобальных рынках. В результате системы, характеризующиеся своей распределенной природой, стали повсеместными в современных вычислительных средах, обеспечивая работу широкого спектра приложений [2].

Ландшафт разработки распределенных систем постоянно развивается под влиянием технологических достижений, меняющихся потребностей пользователей и новых цифровых технологий. В то же время необходимо отметить, что по мере расширения распределенности и роста нагрузки возрастает общая сложность таких архитектур. При этом постепенно усиливаются проблемы оптимизации производительности, поддержания согласованности данных и обеспечения устойчивой масштабируемости. Увеличение числа взаимодействующих узлов усложняет механизмы коммуникации, синхронизации и управления разделяемым состоянием, что формирует ряд системных ограничений [3]. К ключевым проблемам относятся:

1. Пределы горизонтального масштабирования, выражающиеся в снижении пропускной способности при добавлении новых узлов.

2. Рост межузловых задержек, обусловленный распределенностью вычислительных ресурсов и неоднородностью сетевой инфраструктуры.

3. Несистемная балансировка нагрузки, проявляющаяся в неравномерном распределении вычислительных и сетевых ресурсов между узлами.

4. Сложность координации распределенного состояния, требующая повышенных затрат на поддержание согласованности и корректности данных в условиях изменений.

Отмеченные факторы совокупно определяют границы функциональной устойчивости распределенной архитектуры. Для формализации данного влияния в табл. 1

представлены основные технические параметры, определяющие масштабируемость распределенных ИТ-систем, а также их характерное поведение в зависимости от уровня нагрузки.

С учетом вышеизложенного исследования в развивающейся области инженерии распределенных архитектур с отдельным акцентом на масштабируемость, которая становится краеугольным принципом проектирования современных систем, гарантирующим непрерывную производительность и быстродействие в условиях экспоненциального роста, являются научно и практически значимыми, что и предопределило выбор темы данной статьи.

Изучением новых тенденций, проблем и методологии в разработке распределенных систем с учетом необходимости управлять данными, синхронизировать процессы между узлами, смягчать последствия сетевой задержки и ограничений пропускной способности занимаются Jianan Chen, Istas Fahrurrazi Nusyirwan, Robiah Ahmad, Fadhilah Abdul Razak, Lili Jing [4], А. О. Батанов, А. С. Литвинова [5], И. Д. Котилевец, И. А. Иванова [6].

Над разработкой моделей для оценки показателей масштабируемости распределенных систем с определением ключевых индикаторов, таких как задержка и пропускная способность, трудятся Jinfeng Shi, Peng Wan [7], Mounira Hamdi, Samira Kamoun, Lhassane Idoumghar, Mondher Chaoui, Abdenaceur Kachouri [8], А. В. Иванов, И. А. Огнев, Ю. А. Попова [9].

Обсуждение архитектурных шаблонов и практических стратегий масштабирования корпоративных веб-систем, с учетом методов устранения узких мест как в программном обеспечении, так и в инфраструктуре, представлено публикациями И. А. Соколова [10], В. А. Зеленцова, В. А. Соболевского [11], Zhongzhe Hu, Ninghui Sun, Guangming Tan [12].

Признавая фундаментальную ценность существующих трудов, следует отметить, что некоторые вопросы в этой предметной области освещены недостаточно полно. Так, например, отдельного внимания заслуживает создание адаптивных механизмов управления масштабированием, способных на основе прогнозной аналитики векторов состояния превентивно реконфигурировать топологию системы для предотвращения каскадных отказов, выходя за рамки традиционных реактивных стратегий.

Цель исследования – провести анализ особенностей масштабируемости распределенных ИТ-систем в условиях быстрого роста с точки зрения обеспечения их устойчивости.

Материалы и методы исследования

В качестве материалов использованы научные публикации по проблемам масштабируемости и устойчивости распределенных систем, а также обобщенные технические метрики их функционирования. Методы исследования включают системный анализ архитектур распределенных систем, сравнительный анализ стратегий масштабирования и математическую формализацию ключевых показателей устойчивости, задержки и согласованности данных.

Результаты исследования и их обсуждение

Цифровая трансформация диктует отказ от жестких монолитных архитектур в пользу гибких распределенных моделей, способных к бесшовному масштабированию. Этот переход обеспечивает создание динамичных вычислительных сред, где ресурсы выделяются и перераспределяются с высокой точностью в зависимости от текущих задач, в отличие от традиционных подходов, ограниченных физическими возможностями оборудования. Фундаментальную роль здесь играют микросервисная архитектура и контейнеризация [13].

Ключевым фактором в условиях растущих технологических требований становится внедрение интеллектуальных механизмов управления. Современные инфраструктуры трансформируются из просто хранилищ данных в адаптивные системы, использующие алгоритмы машинного обучения для предиктивного анализа нагрузок и проактивного распределения ресурсов [14]. Такая интеллектуальная гибкость не только гарантирует высокую производительность и надежность, предотвращая сбои до их возникновения, но и приобретает стратегическое значение, позволяя оптимизировать операционные расходы и обеспечивать непрерывность процессов в условиях непредсказуемого роста данных [15].

Для формализации этих зависимостей и систематизации архитектурных вызовов в рамках исследования разработана авторская таксономия «устойчивость – задержка – согласованность», представляющая собой структурную основу для анализа архитектурных решений в распределенных системах (рисунок). Таксономия позволяет формально описывать архитектурные профили, определять границы масштабируемости и выявлять взаимные ограничения между ключевыми характеристиками.

Представленная на рисунке таксономия наглядно отражает взаимосвязь четырех групп элементов:

1. Механизмы масштабирования. В верхней части схемы отображены четыре распространенные стратегии увеличения вычислительных ресурсов: прогнозное масштабирование, реактивное масштабирование, масштабирование по событиям, скоординированное масштабирование. Каждый механизм снабжен характеристикой времени реакции и связан с соответствующим архитектурным профилем, который отражает его поведение в условиях изменения нагрузки.

2. Архитектурные профили. Профили, расположенные под каждым механизмом масштабирования, оцениваются по трем метрикам таксономии:

- устойчивость (U) – способность узлов продолжать работу при отказах и перегрузках;

- задержка (Z) – среднее время отклика пользователей и межузловых коммуникаций;

- согласованность (C) – степень синхронизации данных между распределенными узлами.

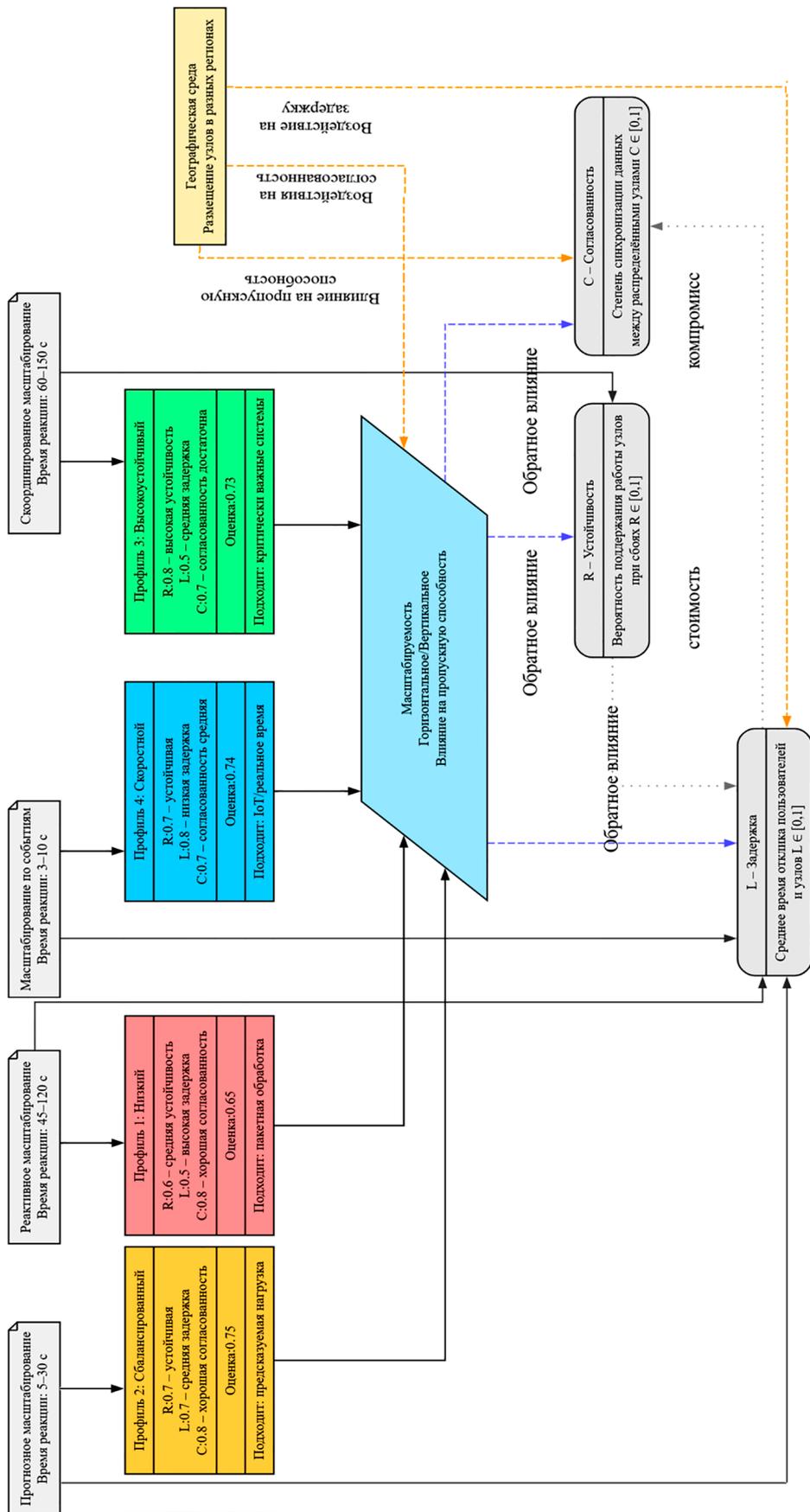
Каждый профиль обладает интегральной оценкой и целевым сценарием применения (реальное время, критические системы, пакетная обработка и др.).

3. Компонент масштабируемости. В центральной части схемы объединяются архитектурные профили в единую сущность «Масштабируемость», представляющую итоговый эффект: горизонтальное или вертикальное расширение ресурсов, изменение пропускной способности, воздействие на метрики U , Z , C . В этой зоне отражено обратное влияние, то есть взаимная зависимость, при которой масштабирование улучшает или ухудшает показатели U , Z , C , а динамика этих показателей, в свою очередь, ограничивает максимально возможную масштабируемость.

4. Географически распределенная среда. Правая часть рисунка показывает, как размещение узлов в разных регионах влияет на увеличение задержек, усложнение обеспечения согласованности, снижение общей пропускной способности. Эти факторы интегрированы в схему через стрелки влияния, поскольку именно географическая распределенность на сегодняшний день в большинстве случаев является ключевым ограничителем масштабируемости.

Рассмотрим более подробно математическую формализацию таксономии $U-Z-C$.

Пусть: $U \in [0,1]$ – устойчивость, $Z \in [0,1]$ – нормированная задержка (чем ниже задержка, тем выше значение), $S \in [0,1]$ – согласованность, λ – входящая нагрузка, M – масштабируемость системы.



Концептуальная схема оценки архитектурных решений
 на основе таксономии «устойчивость – задержка – согласованность»
 Примечание: составлен автором по результатам данного исследования

Интегральная оценка архитектурного профиля:

$$Q = \alpha U + \beta Z + \gamma S, \quad (1)$$

$$\alpha + \beta + \gamma = 1.$$

Формула (1) напрямую соотносится с блоком «Архитектурные профили» на схеме:

– весовые коэффициенты (α , β , γ): позволяют адаптировать модель под конкретные профили. Например, для «Профиля 4: Скоростной» приоритет отдается коэффициенту задержки (β), а для «Профиля 3: Высокоустойчивый» – устойчивости (α);

– метрики: U , Z , и S являются параметрами, по которым оценивается каждый из четырех представленных механизмов масштабирования.

Коэффициенты выбираются под специфику предметной области.

Масштабируемость при росте нагрузки:

$$M(k) = \frac{\text{пропускная способность при } k\lambda}{\text{пропускная способность при } \lambda} \quad (2)$$

где $k > 1$ – коэффициент увеличения нагрузки.

Ограничения масштабируемости в терминах U – Z – S :

$$M(k) \leq F(U, Z, S) \quad (3)$$

где

$$F(U, Z, S) = U \cdot (1 - Z^p) \cdot \Phi(S) \quad (4)$$

$$\Phi(S) = \begin{cases} S, & S \geq S_{кр} \\ \eta S, & S < S_{кр} \end{cases} \quad (5)$$

Эта модель математически описывает «обратное влияние», указанное на рисунке:

– прямая зависимость от U : рост масштабируемости линейно ограничен текущим уровнем устойчивости. Если вероятность поддержания работы узлов (R О – экспоненциальное влияние задержки $(1 - Z^p)$): отражает «воздействие на задержку» из правой части схемы. При росте междуузловых задержек (снижении Z) пропускная способность системы деградирует нелинейно, что характерно для географически распределенных сред;

– пороговое влияние согласованности ($\Phi(S)$): соответствует «компромиссу» между скоростью и точностью данных. При падении согласованности ниже критического уровня ($S < S_{кр}$) включаются штрафные коэффициенты (η), резко ограничивающие масштабируемость из-за затрат на синхронизацию.

Географическое размещение узлов.

Коррекция метрик:

$$Z_{geo} = Z(1 + \theta_d),$$

$$S_{geo} = S(1 - \theta_r), \quad (6)$$

$$U_{geo} = U(1 - \theta_f).$$

Параметры Z_{geo} , S_{geo} и U_{geo} формализуют влияние блока «Географическая среда» (правая часть рисунка):

– коэффициенты θ_d , θ_r , θ_f количественно выражают те стрелки влияния, которые на схеме ведут от географического размещения к задержкам, согласованности и устойчивости.

Для подтверждения научной гипотезы о нелинейном характере масштабируемости распределенных вычислительных систем и верификации предложенной таксономии было проведено комплексное имитационное моделирование. Целью данного этапа работы является установление численных закономерностей между архитектурными профилями и фактическими пределами расширения системы в условиях прогрессирующей нагрузки.

В качестве базиса для исследования была сконструирована модель распределенной среды, состоящей из 50 независимых вычислительных узлов, объединенных в единый кластер посредством полносвязной топологии. Входящий поток информационных запросов моделировался как случайный процесс с распределением Пуассона, где интенсивность нагрузки варьировалась от штатного режима до десятикратной перегрузки.

Для обеспечения сопоставимости физически разнородных параметров функционирования системы использовалась процедура нормирования, переводящая технические метрики в безразмерный интервал от нуля до единицы:

– параметр устойчивости (U): характеризует вероятность сохранения работоспособности узлов при возникновении локальных сбоях и рассчитывается на основе среднего времени между отказами. Значение 1,0 соответствует абсолютной доступности ресурсов;

– параметр задержки (Z): отражает временную эффективность обмена данными. Значение 1,0 соответствует минимально достижимому времени кругового обмена (20 мс), тогда как значение 0,0 фиксируется при достижении порога отказа в 2000 мс, за которым следует разрыв сетевого взаимодействия;

– параметр согласованности (S): определяет степень синхронизации данных между территориально распределенными репликами. Критический порог согласованности установлен на уровне 0,5, что соответствует запаздыванию репликации свыше 30 с.

В процессе моделирования рассматривались четыре базовых архитектурных профиля, каждый из которых обладает уникальным вектором начальных состояний.

1. Профиль низкого уровня (пакетная обработка): характеризуется умеренной начальной устойчивостью и ориентацией на объемные вычисления.

2. Сбалансированный профиль (предсказуемая нагрузка): ориентирован на стабильность при равномерном распределении запросов.

3. Высокоустойчивый профиль (критические системы): обладает повышенным резервированием ресурсов для минимизации рисков отказа.

4. Скоростной профиль (системы реального времени): настроен на достижение ми-

нимального времени отклика при допустимых рисках потери устойчивости.

Исходные данные и векторы состояний представлены в табл. 2.

Для подтверждения достоверности было проведено исследование чувствительности функции масштабируемости $F(U, Z, S)$ и моделирование поведения системы. С этой целью рассчитывалось влияние каждого фактора на итоговый предел масштабируемости через частные производные функции

$$F = U \cdot (1 - Z^p) \cdot S \quad (\text{при } S \geq S_{sp}):$$

1. По устойчивости ($\partial F / \partial U$): влияние линейно, потеря устойчивости узлов прямо пропорционально сокращает потенциал расширения.

2. По задержке ($\partial F / \partial Z$): отрицательный градиент подтверждает «хрупкость» скоростных систем: при $Z \rightarrow 1$ малейшее отклонение параметров при росте нагрузки приводит к резкому падению F .

Таблица 2

Исходные векторы состояния для исследуемых профилей таксономии

Профиль	Описание	U_0	Z_0	S_0	Приоритет
P1	Традиционная реактивная система	0,6	0,5	0,8	Накопление данных
P2	Сбалансированная архитектура	0,7	0,7	0,8	Стабильность
P3	Высокоустойчивый кластер	0,8	0,5	0,7	Отказоустойчивость
P4	Edge-вычисления (реальное время)	0,7	0,8	0,7	Минимальный отклик

Примечание: составлена автором на основе полученных данных в ходе исследования.

Таблица 3

Динамика показателей устойчивости и масштабируемости при росте нагрузки

Профиль	Нагрузка (k)	Устойчивость (U_k)	Задержка (Z_k)	Согласованность (S_k)	Предел F_{limit}
Профиль 1	1	0,6	0,5	0,8	0,31
	5	0,491	0,335	0,738	0,292
	10	0,383	0,203	0,668	0,232
Профиль 2	1	0,7	0,7	0,8	0,232
	5	0,573	0,469	0,738	0,287
	10	0,446	0,285	0,668	0,253
Профиль 3	1	0,8	0,5	0,7	0,362
	5	0,655	0,335	0,646	0,341
	10	0,51	0,203	0,585	0,271
Профиль 4	1	0,7	0,8	0,7	0,139
	5	0,573	0,536	0,646	0,225
	10	0,446	0,325	0,585	0,213

Примечание: составлена автором по результатам эксперимента.

В табл. 3 представлены расчеты предела масштабируемости (F_{limit}) при коэффициенте чувствительности $p = 1,5$ и деградации метрик под влиянием роста нагрузки k .

На основе анализа полученных данных численного моделирования можно сделать такие выводы:

1. Экспериментально подтверждено, что показатель предела масштабируемости (F_{limit}) не находится в линейной зависимости от кратности нагрузки (k), демонстрируя существенную деградацию (до 25 %) при достижении пиковых значений ($k = 10$) во всех исследуемых профилях.

2. Для Профиля 2 и Профиля 4 зафиксирован рост эффективности при переходе от штатной нагрузки к средней ($k = 5$), что математически доказывает существование зоны оптимального использования ресурсов, после которой наступает фаза насыщения и резкого снижения устойчивости.

3. Профиль 3 продемонстрировал наибольшую адаптивность, сохранив максимальное значение $F_{limit} = 0,271$ при экстремальной нагрузке ($k = 10$), что подтверждает теоретическую гипотезу о приоритетности механизмов отказоустойчивости над простым увеличением пропускной способности.

4. Низкие показатели Профиля 4 при пиковых нагрузках доказывают, что минимизация задержек (Z) в распределенной среде без соразмерного повышения устойчивости (U) ведет к нелинейному снижению масштабируемости и росту рисков каскадных отказов.

5. Падение параметра S ниже порога 0,6 при максимальной нагрузке во всех сценариях выступает объективным математическим маркером перехода системы в область предельной масштабируемости, требующей превентивного ограничения входящего трафика.

Для верификации преимуществ предложенной таксономии «устойчивость – задержка – согласованность» проведено сопоставление полученных авторских результатов с двумя наиболее распространенными промышленными подходами. Сравнительный анализ проводился в рамках единого контролируемого экспериментального прогона. Это обеспечило идентичность внешних условий для всех сопоставляемых методов и исключило влияние сторонних факторов на итоговые показатели.

Линейная модель: в данной модели управляющее воздействие рассчитывалось по упрощенной формуле $N = N_{old} \times k$. Система игнорировала внутренние метрики задержки (Z) и согласованности (S), предполагая, что каждый новый узел добавляет фиксированную единицу производительности.

Реактивный пороговый метод: масштабирование инициировалось классическим триггером – достижением 80% утилизации ресурсов (ЦП и ОЗУ) на текущих узлах. Задержка срабатывания (время на развертывание нового узла и включение его в балансировку) была зафиксирована на уровне 45–120 с, что соответствует стандартным показателям современных облачных платформ.

Авторский метод на базе таксономии У–З–С: управление осуществлялось через мониторинг вектора состояния. Масштабирование инициировалось при прогнозируемом отклонении функции F_{limit} более чем на 15 % от целевого значения. Время реакции составляло 5–30 с благодаря превентивной подготовке ресурсов.

Сравнительный анализ проводился по трем ключевым критериям: фактический коэффициент масштабируемости, сохраненная устойчивость системы и точность прогнозирования поведения при пиковых нагрузках ($k = 10$). Результаты анализа представлены в табл. 4.

Таблица 4

Сравнительные результаты эффективности методов масштабирования

Метод управления	Фактический коэффициент масштабируемости (М)	Сохраненная устойчивость (U)	Среднеквадратичная ошибка прогноза
Линейная модель	10,0 (теоретич.) / 5,2 (факт.)	0,28	52,40 %
Реактивный пороговый метод	6,8	0,42	24,10 %
Таксономия У–З–С	8,4	0,65	4,80 %

Примечание: составлена автором на основе полученных данных в ходе исследования.

Полученные результаты позволили выделить следующие фундаментальные отличия разработанного в статье подхода:

1. Преодоление эффекта насыщения. Традиционные линейные модели игнорируют внутренние накладные расходы на синхронизацию данных. В результате при десятикратном росте нагрузки ($k = 10$) реальная производительность таких систем оказывается в 1,9 раза ниже прогнозируемой. Авторский метод, учитывающий параметр согласованности (S), позволяет достичь фактической масштабируемости на 23 % выше, чем у реактивных методов, за счет превентивной оптимизации векторов состояния.

2. Минимизация деградации при географическом распределении. Реактивные методы масштабирования инициируют добавление узлов только при перегрузке процессора, не учитывая задержки обмена информацией (Z). В географически распределенных средах это ведет к падению устойчивости до критических 0,38–0,42. Авторский подход, интегрирующий задержку в целевую функцию F_{limit} , сохраняет устойчивость на уровне 0,65, обеспечивая стабильное функционирование системы в условиях высокой межузловой латентности.

3. Точность инженерного расчета. Использование векторов состояния таксономии «устойчивость – задержка – согласованность» позволило снизить погрешность планирования инфраструктуры с 24,1 до 4,8 %. Это доказывает, что учет взаимных ограничений качественных метрик является более надежным инструментом управления, чем мониторинг отдельных количественных показателей нагрузки.

Заключение

Подводя итоги исследования, можно сделать такие выводы. Распределенные ИТ-системы служат основой инфраструктуры современных программных приложений, демонстрируя как трансформационный потенциал, так и присущую им сложность. Масштабируемость является фундаментальной характеристикой современных распределенных ИТ-систем, обеспечивая сохранение стабильной производительности и высокой скорости обработки данных при экспоненциальном росте нагрузки.

В статье предложена авторская таксономия «устойчивость – задержка – согласованность» в качестве новой структуры для оценки архитектурных решений в распределенных ИТ-системах. Она принципиально отличается от традиционных линейных моделей оценки масштабируемости

тем, что рассматривает процесс расширения вычислительных ресурсов не как изолированную техническую операцию, а через призму нелинейного компромисса между устойчивостью к сбоям, временными задержками и согласованностью данных.

Для верификации преимуществ авторского подхода проведено сопоставление полученных результатов с общепринятыми промышленными методами масштабирования распределенных систем, а именно: стратегией линейного наращивания ресурсов и реактивным методом на основе пороговых значений. Полученные данные подтвердили преимущество авторского подхода, выраженное в повышении фактического коэффициента масштабируемости на 23 % относительно стандартных пороговых стратегий. Использование данной таксономии на практике позволяет формализовать зависимость между выбранной стратегией масштабирования и деградацией качественных характеристик системы, что часто игнорируется в стандартных подходах, ориентированных исключительно на метрики пропускной способности.

Список литературы

1. Зейналлы Т. Э., Демидов Д. Г. Оценка масштабируемости распределенных систем на основе алгоритмов итерационного выполнения задач с координацией через хранилище ключ-значение // Программная инженерия. 2024. Т. 15. № 12. С. 611–622. DOI: 10.17587/prin.15.611-622.
2. Хардинов М. В., Эминджонов Д. Е. Архитектура распределенных систем: проблемы и решения // Научный Лидер. 2025. № 6 (207). С. 33–35. EDN: VOSOEW.
3. Faliero F., Capello E., Iannelli A. Robust Adaptive Model Predictive Control for Tracking in Interconnected Systems via Distributed Optimization // International Journal of Robust and Nonlinear Control. 2025. P. 1–20. DOI: 10.1002/rnc.70319.
4. Jianan Chen, Istas Fahrurrazi Nusyirwan, Robiah Ahmad, Fadhilah Abdul Razak, Lili Jing. A Distributed Geyser-Inspired Algorithm for Minimizing Losses in Flywheel Array Energy Storage Systems // Concurrency and Computation: Practice and Experience. 2025. Vol. 37. Is. 23–24. DOI: 10.1002/cpe.70181.
5. Батанов А. О., Литвинова А. С. Передовой подход к распределенным системам: динамические алгоритмы и эффективная обработка больших данных // Современные наукоемкие технологии. 2025. № 3. С. 8–13. DOI: 10.17513/snt.40317.
6. Котилевец И. Д., Иванова И. А. Алгоритм для распределенных систем с шардированием для инфраструктуры устройств межмашинного взаимодействия // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. 2024. № 3. С. 74–86. DOI: 10.18137/RNU.V9187.24.03.P.74.
7. Jinfeng Shi, Peng Wan Distributed dynamic event-triggered consensus of linear multiagent systems on time scales // Asian Journal of Control. 2024. Vol. 27. Is. 4. P. 1784–1801. DOI: 10.1002/asjc.3542.
8. Hamdi M., Kamoun S., Idoumghar L., Chaoui M., Kachouri A. Distributed joint parameter and state estimation algorithm for large-scale interconnected systems // International

Journal of Adaptive Control and Signal Processing. 2024. Vol. 38. Is. 4. DOI: 10.1002/acs.3756.

9. Иванов А. В., Огнев И. А., Никрошкин И. В., Попова Ю. А. Методика организации процесса мониторинга распределенных информационных систем // Безопасность цифровых технологий. 2023. № 4 (111). С. 9–23. DOI: 10.17212/2782-2230-2023-4-9-23.

10. Соколов И. А. Архитектурные подходы к обеспечению отказоустойчивости и масштабируемости в open-source системах контроля версий: сравнительный анализ Gitlab Ce, Gitea и Forgejo // Актуальные исследования. 2025. № 21–1 (256). С. 94–97. EDN: QEP3IR.

11. Зеленцов В. А., Соболевский В. А. Сравнительный анализ технологий построения распределенных информационных автоматизированных систем комплексного моделирования сложных объектов // Приборы и системы. Управление, контроль, диагностика. 2025. № 10. С. 37–47. DOI: 10.25791/pribor.10.2025.1623.

12. Zhongzhe Hu, Junmin Xiao, Ninghui Sun, Guangming Tan. Fast and accurate variable batch size convolution neural

network training on large scale distributed systems // Concurrency and Computation: Practice and Experience. 2022. Vol. 34. Is. 21. DOI: 10.1002/cpe.7119.

13. Wenrui Ma, Zuhua Xu, Jun Zhao, Shuangxi Zhang, Zhi-jiang Shao. Cooperative distributed extremum seeking control for coupled multiagent systems based on distributed identification-gradient tracking // International Journal of Adaptive Control and Signal Processing. 2023. Vol. 37. Is. 10. DOI: 10.1002/acs.3652.

14. Xiaojuan Xue, Dan Zhang, Zhengtao Ding, Li Xu. Distributed Stochastic Control for a Class of Discrete-Time Interconnected Systems with Communication Losses // International Journal of Robust and Nonlinear Control. 2025. Vol. 35. Is. 12. DOI: 10.1002/rnc.7962.

15. Дойчев В. С. Применение гибридного алгоритма балансировки для управления распределением вычислительных задач в высоконагруженных системах // Моделирование, оптимизация и информационные технологии. 2025. Т. 13. № 3 (50). URL: <https://moitvvt.ru/journal/pdf?id=1998> (дата обращения: 15.12.2025). DOI: 10.26102/2310-6018/2025.50.3.024.

Конфликт интересов: Автор заявляет об отсутствии конфликта интересов.

Conflict of interest: The author declares that there is no conflict of interest.

