

---

# ТЕХНИЧЕСКИЕ НАУКИ

## TECHNICAL SCIENCE

УВК 004.41/.42  
ББК 32.973.202  
К 43

### **Кiryushkin M.A.**

*Аспирант кафедры теоретической физики и компьютерных технологий Кубанского государственного университета, тел. 89183509695, e-mail: shadekma@mail.ru*

### **Chizhikov V.I.**

*Доктор физико-математических наук, профессор, заведующий кафедрой теоретической физики и компьютерных технологий Кубанского государственного университета, тел. 89183351736, e-mail: cvi@phys.kubsu.ru*

## **Анализ этапов масштабирования систем** (Рецензирована)

### **Аннотация**

*Исследованы основные этапы масштабирования систем и представлены основания для реализации системы до определенного этапа.*

**Ключевые слова:** информационные системы, масштабирование, шкалирование, реализация.

### **Kiryushkin M.A.**

*Post-graduate student of Theoretical Physics and Computer Technologies Department, Kuban State University, ph. 89183509695, e-mail: shadekma@mail.ru*

### **Chizhikov V.I.**

*Doctor of Physics and Mathematics, Professor, Head of Theoretical Physics and Computer Technologies Department, Kuban State University, ph. 89183351736, e-mail: cvi@phys.kubsu.ru*

## **Analysis of stages of system scaling**

### **Abstract**

*This work examines the main stages of system scaling and presents reasons for realization of the system to a certain stage.*

**Keywords:** information systems, scaling, realization.

## **Введение**

Масштабирование системы [1-3] очень дорогой и трудоемкий процесс. Многие успешные системы могут столкнуться с проблемой невозможности выйти за определенные границы масштабирования. Это может быть ограничение количества пользователей, невозможность удаленных пользователей работать с системой из-за скорости или невозможности доступа в принципе.

Данные проблемы могут иметь критические последствия – от потери нескольких клиентов до потери места на рынке (поскольку конкурирующие фирмы, видя успех такой системы, могут предложить свою реализацию, но с возможностью работать с большим количеством людей и на большем расстоянии).

*Цель работы.* Провести анализ этапов масштабирования и определить цепочку развития системы от небольшой системы до распределенной системы мирового уровня. Вывести основные требования к системе на каждом этапе. Анализ позволит определить требования к системе на этапе планирования, что позволит снизить затраты средств и

---

времени в процессе будущего масштабирования.

## Этапы масштабирования

Любая система может находиться на разных этапах масштабирования. На каждом из них она способна поддерживать различное количество пользователей, независимых административных единиц и удаленных клиентов.

Системы на каждом этапе различаются по трем различным показателям:

1. **По размеру** – показатель системы, который характеризует, насколько легко можно подключить к системе новых пользователей и новые ресурсы.

2. **По местоположению** – показатель системы, характеризующий возможность распределения и расширения в пространстве, то есть – географически.

3. **По административному признаку** – возможность распределения системы между совершенно разными независимыми административными единицами (организациями, группами).

На каждом этапе показатели разные и могут варьировать в зависимости от используемого оборудования, местоположения сервера и так далее. Однако, чем выше уровень масштабирования, тем выше показатели.

Рассмотрим основные этапы масштабирования, начиная с самого низкого уровня и заканчивая самым высоким:

### 1. Масштабирование системы вглубь

На данном этапе основную роль в масштабировании играет оборудование (когда вместо небольшого сервера приобретается более крупный многопроцессорный сервер, далее, возможно, понадобится мэйнфрейм или даже суперкомпьютер). Несмотря на то, что это самый низкий уровень масштабирования, он так же является одним из самых дорогих, покупка сервера может обойтись компании в очень значительную сумму.

На саму систему не накладывается практически никаких ограничений. Крайне редко может оказаться, что система откажется работать на более мощном сервере.

### 2. Масштабирование системы вширь (локально)

Повышение производительности на этом этапе практически идентично масштабированию вглубь. Но имеется два существенных отличия – это цена (стоимость нескольких небольших серверов может оказаться существенно ниже одного большого сервера) и требования к системе.

Основное требование системы – необходимость работать рассредоточено. Должна быть возможность переноса части функционала с единого сервера на несколько других и/или возможность распараллелить одну задачу на несколько машин.

Многие старые системы не могут удовлетворить этим требованиям. Такие системы строились как единое целое – неделимое. Часто даже внесение незначительных изменений требовало больших усилий, а уж выделить определенный функционал и вовсе не представлялось возможным.

### 3. Метод сокрытия времени ожидания связи

Рано или поздно, при масштабировании, любая система столкнется с проблемой ограничения линии связи.

Первое, но не всегда эффективное решение данной проблемы – метод сокрытия времени ожидания. Он заключается в том, чтобы позволить пользователю отправить запрос и не ждать ответа, а совершать какие-либо другие действия. Такие запросы называют асинхронными и их часто используют в глобальных сетях. Сам метод не решает проблему связи, а только помогает ее скрыть.

---

Для реализации данного метода требуется, чтобы система была способна работать с асинхронными запросами. Эту способность лучше заложить в систему на этапе разработки, иначе могут возникнуть большие сложности при изменении системы в будущем.

На этапе планирования стоит проанализировать выгоду от этого метода – иначе может получиться, что асинхронные запросы будут бесполезны, например, если пользователю особо нечем будет заняться во время выполнения запроса и/или время ожидания будет настолько велико, что целесообразнее будет перейти к следующему этапу масштабирования.

#### 4. Репликация

Самый мощный и самый сложный вариант масштабирования системы.

Репликация способна ускорить работу системы на несколько порядков, но вносит в систему проблему непротиворечивости. При решении проблемы непротиворечивости приходится искать компромисс между потерей производительности и вероятностью получить неодинаковые копии.

При поддержке системы такого уровня требуется учитывать множество параметров каждой реплики. И добавление новых реплик напрямую зависит от состояния существующих (например, от скорости работы, скорости доступа к ним, стабильности работы и так далее).

### Цепочка развития системы

На основании описанных выше этапов масштабирования можно построить цепочку развития системы (рис. 1).

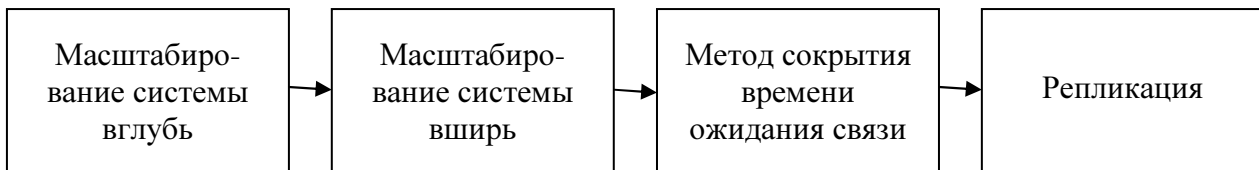


Рис. 1. Цепочка развития системы.  
Этапы развития системы от небольшой до распределенной

Далеко не каждая система пройдет все этапы масштабирования. Какие-то системы могут остановиться на одном из первых этапов, другие – наоборот, проскочить один или несколько первых этапов.

Системы следует проектировать и разрабатывать таким образом, чтобы сразу учесть возможные переходы с одного этапа на другой (конечно, с начала разработки невозможно узнать, насколько успешным окажется проект, и сколько клиентов он должен будет поддерживать, однако всегда можно предположить возможные варианты работы и масштабы разработки). Требования для успешного перехода с одного этапа на другой даны в описании каждого этапа. Если система изначально не была спроектирована для перехода на следующий этап, переход может потребовать значительных затрат как временных, так и денежных.

Например, если система строилась как одно целое и каждая часть функционала тесно связана с набором других функций, то вряд ли системе удастся малыми затратами уйти дальше первого этапа. Так же, если система критично относится к актуальности содержащимся в ней данным и малейшее несоответствие может привести к катастрофе, то репликация станет для системы не методом повышения, а методом понижения производительности из-за постоянных синхронизаций реплик, при изменении данных на одной из них.

---

Третий этап масштабирования является обособленным, а возможность и целесообразность использования его полностью зависит от конкретной системы (какие действия сможет совершать пользователь во время запроса, нужно ли пользователю ждать результат запроса).

### **Заключение**

В статье представлены основные этапы масштабирования системы. Даны их описания и требования при реализации. Построена цепочка развития системы, с помощью которой можно определить до какого этапа система должна поддерживать возможность перехода. И, благодаря правильному выбору поддерживаемых этапов, можно производить масштабирование системы до соответствующего этапа, обходясь малыми затратами.

#### **Примечания:**

1. Танненбаум Э., ван Стеен М. Распределенные системы. Принципы и парадигмы. СПб.: Питер, 2003. 877 с.
2. Александер Стив. Масштабируемость // Computerworld Россия: еженедельник. 2000. № 26.
3. Алдунин Ю.А. Синхронизация времени в распределенных системах в зависимости от выбранной модели противоречивости // Вестник РГРТУ. Рязань, 2007. Вып. 22. С. 18-21.

#### **References:**

1. Tannenbaum E., van Steen M. Distributed systems. Principles and paradigms. SPb.: Peter, 2003. 877 p.
2. Alexander Steve. Scalability // Computerworld of Russia: a weekly. 2000. No. 26.
3. Aldunin Yu.A. Time synchronization in the distributed systems depending on the chosen model of inconsistency // The Bulletin of the RGRTU. Ryazan, 2007. Iss. 22. P. 18-21.