

УДК 004.75

АНАЛИЗ ПРОИЗВОДИТЕЛЬНОСТИ И ОПТИМИЗАЦИЯ ВЫСОКОНАГРУЖЕННЫХ ГЕОПОРТАЛЬНЫХ СИСТЕМ

¹Ямашкин С.А., ¹Ямашкин А.А., ²Ямашкина Е.О.

¹ФГБОУ ВО «Национальный исследовательский Мордовский государственный университет им. Н.П. Огарёва», Саранск, e-mail: yamashkinsa@mail.ru;

²ФГБОУ ВО «МИРЭА – Российский технологический университет», Москва

В статье дана характеристика проекту, посвященному решению научной проблемы анализа производительности и оптимизации высоконагруженных геопортальных систем как узлового компонента цифровых инфраструктур пространственных данных, формируемых на основании того, что для оптимизации процессов хранения и эффективного использования пространственных данных ИПД должна содержать связанные с ними программные интерфейсы подсистемы хранения, анализа и синтеза, визуализации и распространения пространственных данных. Для решения задачи анализа производительности и оптимизации высоконагруженных геопортальных систем разработана методика расчета интегральной оценки производительности. Данный показатель является прикладным, количественным и интерпретируемым. Он решает недостатки методики Apdex, заключающиеся в недостаточной гибкости аппарата и ориентации главным образом на показатели отклика системы. На основании определенных правил значения производительности системы могут быть сопоставлены с интерпретируемыми с позиции «хорошо» – «плохо» терминами. Во второй части статьи дан перечень рекомендаций по развертыванию геопортальных систем, позволяющих повысить их надежность и отказоустойчивость, посредством которого достигаются следующие свойства геопортальной системы: устойчивость при падении отдельного сервера, а также при недоступности дата-центра любого региона; автоматическая перестройка репликации и балансировки; возможность автоматического развертывания новых окружений для разработки при создании новых ветвей; непрерывная интеграция и разработка через тестирование; развертывание новых версий без простоя.

Ключевые слова: инфраструктура пространственных данных, геопортал, анализ производительности, оптимизация, высоконагруженные системы

PERFORMANCE ANALYSIS AND OPTIMIZATION OF HIGH-LOADED GEOPORTAL SYSTEMS

¹Yamashkin S.A., ¹Yamashkin A.A., ²Yamashkina E.O.

¹Federal State Budgetary Educational Institution of Higher Education
«National Research Ogarev Mordovian State University», Saransk, e-mail: yam-ashkinsa@mail.ru;

²Federal State Budgetary Educational Institution of Higher Education
«MIREA – Russian Technological University», Moscow

The article describes a project dedicated to solving the scientific problem of performance analysis and optimization of highly loaded geoportals, as a nodal component of digital infrastructures of spatial data, formed on the basis that for optimization of storage processes and effective use of spatial data, SDI should contain related software interfaces of the subsystem of storage, analysis and synthesis, visualization and distribution of spatial data. To solve the problem of performance analysis and optimization of highly loaded geoportals, a methodology for calculating the integral performance assessment has been developed. This indicator is applied, quantitative and interpretable. It solves the disadvantages of the Apdex technique, which are the lack of flexibility of the apparatus and focus mainly on the performance of the system's response. Based on certain rules, the values of system performance can be compared with terms interpreted from the standpoint of «good» – «bad». In the second part of the article, a list of recommendations for the deployment of geoportals is given, which makes it possible to increase their reliability and fault tolerance and through which the following properties of a geoportals system are achieved: stability when a separate server falls, as well as when a data center of any region is unavailable; automatic rebuilding of replication and balancing; the ability to automatically deploy new development environments when creating new branches; continuous integration and test-driven development; deploying new versions without downtime.

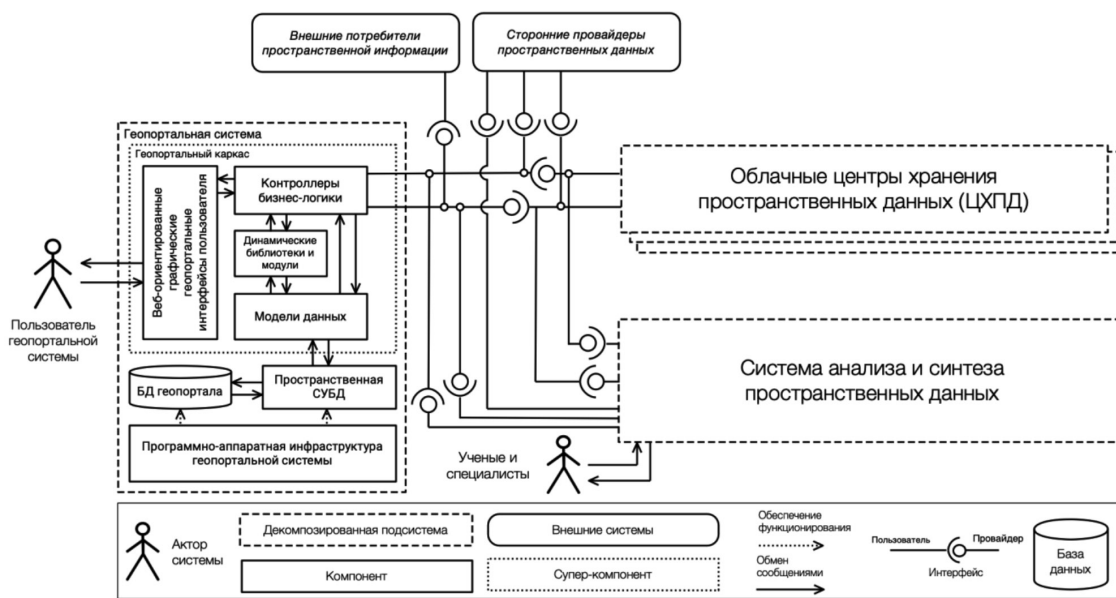
Keywords: spatial data infrastructure, geoportals, performance analysis, optimization, high-load systems

Геопорталы как внешний компонент цифровых инфраструктур пространственных данных (ИПД) относятся к классу распределенных веб-систем и, следовательно, наследуют их характерные особенности. В частности, этот факт накладывает отпечаток на нефункциональные требования, предъявляемые к геопортальным системам [1]. Так, с точки зрения безопасности важно обеспечение свойств конфиден-

циальности (защиты от несанкционированного доступа), доступности (защиты при параллельном доступе к ресурсам) и целостности (сохранности данных) системы. Актуальной задачей при проектировании геопорталов является проблема обеспечения качества обслуживания, заключающаяся в необходимости выполнения определенного набора клиентских требований.

Качественное требование масштабируемости определяет способность геопортальной системы отвечать на вызовы роста функциональности и увеличения рабочей нагрузки [2]. При этом целесообразно, чтобы время поиска по базам геопространственных данных не росло экспоненциально при увеличении объема накапливаемой информации о природно-социально-произ-

пространения пространственных данных (рисунок). Роль компонента, выполняющего две последние функции, традиционно возлагается на геопортальные системы [3]. Внешними компонентами по отношению к ИПД при этом выступают ее ключевые акторы: пользователи геопортальных систем, ученые и специалисты, потребители и провайдеры геопространственных данных [4].



Геопортальная система как узловой компонент цифровой инфраструктуры пространственных данных

водственных системах, а величина использования ресурсов не увеличивалась бесконтрольно с ростом количества пользователей.

Данная публикация посвящена решению научной проблемы анализа производительности и оптимизации высоконагруженных геопортальных систем, которая неизбежно возникает в рамках процессов итерационной разработки, внедрения и технической поддержки геопортальных систем и связана с необходимостью получения интегральной оценки производительности.

Материалы и методы исследования

Анализ опыта построения проектно-ориентированных цифровых ИПД показывает, что процесс формирования архитектуры систем данного класса должен быть основан на положении о том, что для оптимизации процессов хранения и эффективного использования пространственных данных ИПД должна содержать связанные за счет программных интерфейсов подсистемы хранения, анализа и синтеза, визуализации и рас-

При первичном развертывании геопортала необходимо иметь гарантии устойчивости системы к нагрузкам определенного уровня и обеспечения при этом необходимой производительности (задача первого типа). Для достижения обозначенной цели требуются последовательно решить задачи, направленные на: 1) определение целевой производительности P_i геопортальной системы на основе формализованного набора критериев; 2) проведение нагрузочного тестирования, оценку фактической производительности P_f и сравнение ее с целевыми показателями; 3) развертывание геопортала, оценку производительности при условии работы реальных пользователей P_r и сопоставление полученных показателей с целевыми.

При наличии рабочего геопортального решения может потребоваться решение задачи второго типа: достижение повышения текущей производительности посредством оптимизации текущей версии геопортальной системы на основе последовательного решения следующих задач: 1) интегральная оцен-

ка текущей производительности системы P_{r1} ; 2) выявление узких мест геопортальной системы, ее программно-аппаратных компонентов, особенности функционирования которых ограничивают возможность достижения целевых показателей производительности P_j ; 3) оптимизация системы с последующей оценкой производительности новой версии P_{r2} и сравнение полученного результата с целевым показателем P_r .

Для решения обозначенных выше задач необходимо определить методику формирования оценки производительности P . Данный показатель должен быть прикладным (отражающим реальную неабстрактную производительность геопортальной системы), интегральным (учитывающим все ключевые аспекты работы системы и варианты ее использования); количественным (измеряющимся в числах для обеспечения возможности сравнения производительности при разных условиях для различных версий системы); интерпретируемым (определяемым в терминах «хорошо» – «плохо»).

Примем за множество ключевых операций геопортальной системы Op совокупность из конечного числа n операций, являющихся критичными для функционирования инфраструктуры пространственных данных, недостаточная производительность которых ведет к потерям различного типа. Отдельного внимания при этом заслуживают операции, выполняемые параллельно большим количеством пользователей.

$$Op = \{op_i | i \in \mathbb{N} \wedge i \leq N\}.$$

Далее решение задачи анализа производительности геопортальной системы с целью оптимизации ее функционирования в условиях высокой нагрузки будем проводить с позиции исследования особенностей выполнения ключевых операций.

Отметим, что для оценки производительности корпоративных систем в условиях высокой нагрузки существуют известные методики и стандарты. Так, методика Apdex (Application Performance Index, или индекс производительности приложений) представляет собой международный открытый стандарт [5], разработанный для решения задачи объективной оценки производительности информационных систем разного класса и позволяющий свести к простому значению разнородные факторы и статистические данные о производительности информационной системы. К положительным особенностям рассматриваемого аппарата следует отнести ранжирование анализируемых операций по приоритетности, учет объективных фактических данных, интуитивную понят-

ность оценки для конечных пользователей системы и ее владельцев. Однако у данного подхода есть и недостатки, заключающиеся в недостаточной гибкости аппарата и ориентации главным образом на показатели отклика системы.

Результаты исследования и их обсуждение

Приведем последовательность шагов по формированию методики объективной оценки производительности геопортальной системы. Для каждой ключевой операции op_i определим приоритет π_j , выражаемый натуральным числом, обратно пропорциональным важности ключевой операции и позволяющим отсортировать операции в порядке критичности влияния качества их производительности на работу геопортальной системы в целом.

Следующим этапом необходимо собрать данные о производительности для каждой ключевой операции op_j , которые должны быть зафиксированы в виде множества показателей производительности операции T_j , имеющего мощность m :

$$T_i = \{\tau_{ij} | j \in \mathbb{N} \wedge j \leq M\}.$$

Показатель производительности τ_{ij} при этом должен быть выражен численным значением: это могут быть замер времени выполнения ключевой операции, показатель отказов, факт запуска определенного сценария и активации условий, число ошибок и т.д. Также каждому показателю производительности τ_{ij} может быть задан приоритет π_{ij} , указывающий на важность критерия. Для каждого показателя производительности τ_{ij} определим получаемое на основе функций $f_{ij}: \tau_{ij} \rightarrow X_{ij}$ конечное число утверждений, характеризующих кластеризованные по определенному принципу значения набора показателей производительности:

$$X_{ij} = \{x_{ij(k-1)} | k \in \mathbb{N} \wedge k \leq K\}.$$

Каждому утверждению x_{ijk} сопоставим значение фактора влияния риска ω_{ijk} . Значениям каждого показателя производительности должно быть сопоставлено минимум два утверждения: x_{ij0} , для которого $\omega_{ij0} = 0$ (для случая, при котором наступает отрицательное влияние на работу геопортальной системы), и x_{ij1} , для которого $\omega_{ij1} = 1$ (для положительного результата). Дополнительно можно определить конечное число утверждений x_{ijk} с факторами влияния риска ω_{ijk} , характеризующими состояния различной степени умеренности, причем исход можно считать тем более положительным, чем больше принятое для него значение ω_{ijk} .

Обозначим через N_{ijk} количество исходов за серию испытаний, приведших к получению утверждения x_{ijk} . Таким образом, проведя серию из L наблюдений относительно выполнения операций op_i , с точки зрения оценки показателя производительности τ_{ij} , можно дать оценку производительности $P_{\Sigma ij}$ относительно операции i с точки зрения показателя j для анализируемой геопортальной системы:

$$P_{\Sigma ij} = \frac{\sum_{k=0}^{K-1} \omega_{ijk} N_{ijk}}{\sum_{k=0}^{K-1} N_{ijk}} : (\omega_{ij0} = 0) \wedge (\omega_{ij(K-1)} = 1) \wedge (0 < \omega_{ij(k-1)} < 1 : k \in (1, K-1)).$$

Показатель производительности $P_{\Sigma ij}$ при этом будет измеряться в интервале от нуля (худший, неудовлетворительный исход) до единицы (идеальный тест).

Интерес представляет альтернативный показатель, выстраиваемый по принципу гистограмм P_{Hij} и характеризующий распределение количества показателей производительности при сериях из L_{ij} испытаний по A_{ij} группам. Для его расчета определим через $\tau_{ij(L)}$ вектор, сформированный на основе значений параметров τ_{ij} , полученных на основе L_{ij} испытаний.

$$P_{Hij} = \left\{ h_{ija} = \sum_{l=1}^{L_{ij}} \tau_{ijl} \begin{cases} \tau_{ijl} \geq \min(\tau_{ij(L_{ij})}) + (a-1)\sigma \\ \tau_{ijl} < \min(\tau_{ij(L_{ij})}) + a\sigma \\ \wedge (a \in \mathbb{N} \wedge a \leq A_{ij}) \\ \wedge \sigma = \frac{(\max(\tau_{ij(L_{ij})}) - \min(\tau_{ij(L_{ij})}))}{A_{ij}} \end{cases} \right\}.$$

При этом количество групп A может быть подобрано вручную с целью повышения информативности метрики или определено на основе правила. Так, согласно правилу Стёрджеса [6], количество групп зависит от объема данных:

$$A_{ij} = \lfloor 1 + \log_2 L_{ij} \rfloor.$$

Однако это правило не годится для искаженных наборов данных и может быть модифицировано с целью учета особенностей выборок с нестандартным распределением вероятности путем учета коэффициента асимметрии g_{ij} и стандартного отклонения $\sigma_{g_{ij}}$.

$$A_{ij} = 1 + \left\lfloor \log_2 L_{ij} \left(1 + \frac{|g_{ij}|}{\sigma_{g_{ij}}} \right) \right\rfloor.$$

Для интегральной оценки $P_{\Sigma i}$ производительности ключевой операции op_i относительно всего множества показателей производительности T_i можно прибегнуть к использованию агрегатных функций, например взятию минимума.

$$P_{\Sigma i} = \min_{1 \leq j \leq M} P_{\Sigma ij} \left[\pi_{ij} \geq \pi_{ij(\min)} \right].$$

В этом случае значение $P_{\Sigma i}$ будет равно худшей оценке производительности ключевой операции op_i относительно каждого критерия из множества T_i , что позволяет метрике давать характеристику надежности геопортальной системы относительно информации о значении наиболее низкого показателя производительности и характеризует наиболее узкое место в системе. Учитывая значения $\pi_{ij(\min)}$ (пороговый значимый приоритет), можно отсеять при этом операции определено низкой значимости и не учитывать их производительность.

$$P = \langle P_{\Sigma}, P_H \rangle = \left\langle \begin{bmatrix} P_{\Sigma 1} \\ P_{\Sigma 2} \\ \dots \\ P_{\Sigma N} \end{bmatrix}, \begin{bmatrix} P_{H1} \\ P_{H2} \\ \dots \\ P_{HN} \end{bmatrix} \right\rangle.$$

На основании определенных правил значения производительности системы P могут быть сопоставлены с интерпретируемыми с позиции «хорошо» – «плохо» терминами.

Представим далее перечень рекомендаций по развертыванию геопортальных систем, позволяющих повысить их надежность и отказоустойчивость:

1) развертывание и настройка облачных мультимодельных хранилищ данных [7] в нескольких регионах, а также узлов с геопортальным фреймворком в разных зонах доступности с балансировкой нагрузки над ними и оркестрацией узлов;

2) организация инструментария жизненного цикла геопортальной системы, представляющего возможности для управления репозиторием, отслеживания ошибок, непрерывной интеграции;

3) проведение юнит- и функциональных тестов в изолированном окружении, мониторинг и отслеживание статусов геопортальных сервисов и серверов и сетевого оборудования, сбор бизнес-метрик и уведомлений об отказах системы;

4) организация резервного копирования данных в облачное хранилище в зашифрованном end-to-end виде и с дедупликацией.

Таким способом достигаются следующие свойства геопортальной системы: устойчивость при падении отдельного сервера, а также при недоступности дата-центра любого региона; автоматическая перестройка репликации и балансировки; возможность автоматического разворачивания новых окружений для разработки при создании новых ветвей; непрерывная интеграция и разработка через тестирование; развертывание новых версий без простоя.

Выводы

В статье дана характеристика проекту, посвященному решению научной проблемы анализа производительности и оптимизации высоконагруженных геопортальных систем. Данная проблема возникает в рамках итерационной разработки, внедрения и технической поддержки геопортальных систем и связана с необходимостью получения интегральной оценки производительности: при первичном развертывании геопортала необходимо иметь гарантии относительно устойчивости системы к нагрузкам определенного уровня и обеспечения при этом необходимой производительности,

а при наличии рабочего геопортального решения может потребоваться повышение текущей производительности посредством оптимизации текущей версии геопортальной системы на основе последовательного решения следующих задач.

Для решения обозначенных выше задач разработана методика формирования оценки производительности в виде прикладного, интегрального, количественного, интерпретируемого показателя, устраняющего проблему методики Ardex, заключающуюся в недостаточной гибкости аппарата и ориентации, главным образом, на показатели отклика системы. На основании определенных правил значения производительности системы могут быть сопоставлены с интерпретируемыми с позиции «хорошо» – «плохо» терминами.

В заключение статьи представлен перечень рекомендаций по развертыванию геопортальных систем, позволяющих получить следующие свойства: устойчивость при отказе отдельного сервера; автоматическая перестройка репликации и балансировки; возможность автоматического развертывания новых окружений для разработки при создании новых ветвей; непрерывная интеграция и разработка через тестирование; развертывание новых версий без простоя.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-37-70055.

Список литературы

1. Gamez M.R., Perez A.V., Falcones V.A.M., Bazurto J.J.B. The geoportals as strategy for sustainable development. International Journal of Physical Sciences and Engineering. 2019. Vol. 3. № 1. P. 10–21.
2. Ямашкин С.А., Ямашкин А.А., Федосин С.А. Разработка проектно-ориентированной инфраструктуры пространственных данных с применением облачных технологий // Радиопромышленность. 2019. Vol. 29. № 3. С. 79–90.
3. Jiang H., Mazzetti J., Koo H., Chen M. Current status and future directions of geoportals. International Journal of Digital Earth. 2019. Vol. 13. № 7. P. 1–22.
4. Gkonos C., Iosifescu I.E., Hurmi L. Spinning the wheel of design: evaluating geoportals Graphical User Interface adaptations in terms of human-centred design. International Journal of Cartography. 2019. Vol. 5. № 1. P. 23–43.
5. Sevcik P. Defining the application performance index. Business Communications Review. 2005. Vol. 20. P. 8–10.
6. Sturges H.A. The choice of a class interval. Journal of the American statistical association. 1926. Vol. 21. № 153. P. 65–66.
7. Скворцов М.А., Большакова М.В., Ямашкин С.А., Ямашкин А.А. Сравнительный анализ подходов к управлению базами данных для организации хранилища репозитория нейросетевых моделей // Современные наукоемкие технологии. 2021. № 6 (1). С. 108–113.