

Сравнительный анализ систем управления базами данных при реализации распределенного хранилища геоданных

А.А. Саенко, В.С. Габов, А.А. Журавлев

Тихоокеанский государственный университет, Хабаровск

Аннотация В статье представлен сравнительный анализ современных систем управления базами данных (PostgreSQL/PostGIS, Oracle Database, Microsoft SQL Server и MongoDB) в контексте реализации распределённого хранилища геопространственной информации. Рассмотрены особенности работы с векторными, растровыми данными и облаками точек, а также проведено тестирование производительности, масштабируемости и отказоустойчивости в условиях многонодовой архитектуры. На основе результатов экспериментов показано, что PostgreSQL/PostGIS является наиболее универсальной системой, обеспечивающей сбалансированное сочетание функциональности, открытости и масштабируемости. Oracle Database продемонстрировала высокую эффективность при работе с растровыми данными, Microsoft SQL Server подтвердил надёжность при работе с векторной информацией, а MongoDB показала целесообразность использования в задачах гибкого хранения метаданных. Полученные результаты могут быть использованы при выборе технологической платформы для построения геоинформационных систем различного назначения.

Ключевые слова: геоинформационная система, база данных, postgresql, postgis, oracle database, microsoft sql server, mongodb, вектор, растр, облако точек, масштабируемость, производительность, отказоустойчивость.

Введение

Геоинформационная система (ГИС) — информационные системы, обеспечивающие сбор, хранение, обработку, отображение и распространение данных, а также получение на их основе новой информации и знаний о пространственно-координированных явлениях [1]. Такие системы не ограничиваются визуализацией пространственных данных на карте — они позволяют выявлять пространственные закономерности, устанавливать взаимосвязи между различными факторами и формировать обоснованные управленческие решения на основе комплексного анализа геоданных.

Эффективность функционирования ГИС напрямую зависит от качества организации хранения и управления данными [2]. В таких системах используются такие типы данных как пространственные (географические сведения, описывающие местоположение объекта в пространстве относительно других объектов) и атрибутивные (описательные параметры),

различающиеся по форме представления, способу хранения и семантике [3]. Объемы обрабатываемых данных, необходимость обеспечения высокой доступности, масштабируемости и согласованности при работе с распределенными источниками информации ставят перед разработчиками ГИС задачу выбора оптимальной системы управления базами данных (СУБД).

В условиях растущих требований к производительности и отказоустойчивости ГИС, особенно при построении распределенных хранилищ геопространственной информации, актуальной становится задача сравнительного анализа различных СУБД [4, 5]. Особый интерес представляет исследование их функциональных возможностей, степени интеграции с современными ГИС-платформами, поддержки отраслевых стандартов и масштабируемости в распределенной среде.

Целью исследования является проведение сравнительного анализа современных систем управления базами данных в контексте реализации распределенного хранилища геоданных. Анализ проводится с учетом их архитектурных особенностей, поддержки пространственных расширений, соответствия отраслевым стандартам и совместимости с популярными ГИС-инструментами. Результаты работы позволят обосновать выбор наиболее эффективных решений для построения масштабируемых и отказоустойчивых геоинформационных систем.

Виды геоданных в ГИС

Пространственные данные — это информация, описывающая местоположение, форму и взаимное расположение объектов на поверхности Земли. Они являются основой функционирования геоинформационных систем и обеспечивают связь между данными и их географическим контекстом. Эти данные могут быть представлены в двух основных форматах: векторные и растровые.

Векторные данные — используются для представления дискретных объектов. Точки обозначают отдельные объекты, такие как здания или населённые пункты, линии — дороги или реки, а полигоны — административные границы или участки земли.

Растровые данные — представляют информацию в виде регулярной сетки ячеек, где каждая имеет своё значение (например, температура, высота или цвет пикселя). Этот формат применяется для отображения непрерывных явлений, таких как спутниковые снимки или цифровые модели рельефа.

Среди разнообразия растровых данных особое место занимают цифровые модели рельефа (ЦМР) — одно из ключевых представлений высотной составляющей земной поверхности в геоинформационных системах [6, 7]. ЦМР позволяют описывать форму местности в цифровом виде и используются для решения широкого круга задач: от анализа уклонов и построения водосборных бассейнов до трёхмерного моделирования ландшафта. Хотя наиболее распространённым форматом ЦМР является растровый, они также могут быть представлены в векторном виде (горизонтали), в форме триангулированной сети (TIN) или как облака точек.

TIN (Triangulated Irregular Network) – модель рельефа, построенная на основе триангуляции множества точек с заданными высотами, что позволяет точно отражать сложный рельеф местности [8].

Облако точек — это совокупность множества отдельных точек, каждая из которых имеет трёхмерные координаты X , Y , Z в пространстве, а также может содержать дополнительную информацию — например, значения интенсивности сигнала, цветовые параметры (RGB) или классификацию объекта.

Атрибутивные данные — это не пространственные характеристики объектов, например, название улицы, площадь участка, население города или тип дороги, связанные с географическими объектами.

Помимо пространственных и атрибутивных данных, важным элементом информационной структуры геоинформационных систем являются метаданные — информация, описывающая характеристики самих данных. Они содержат сведения о происхождении, точности, формате, актуальности и других параметрах наборов информации, что обеспечивает их достоверность и корректное применение на практике.

Выбор СУБД

Для хранения и обработки разнородной информации, используемой в геоинформационных системах, применяются различные типы систем управления базами данных. Учитывая разнообразие типов информации — от точечных объектов до трёхмерных моделей — необходимо учитывать не только функциональные возможности СУБД, но и их совместимость с ГИС-системами, поддержку стандартов Open Geospatial Consortium и ISO/TC 211, и эффективность обработки конкретного вида данных [9, 10].

Среди существующих решений можно выделить несколько СУБД, каждая из которых имеет свои особенности реализации пространственных расширений и уровень соответствия международным стандартам.

PostgreSQL/PostGIS — открытая объектно-реляционная СУБД, поддерживающая расширение PostGIS, реализующее стандарты OGC и полноценную работу с геоданными.

Oracle Spatial and Graph — коммерческая СУБД с мощной поддержкой пространственных операций, топологии и анализа графов.

Microsoft SQL Server (с модулем Spatial) — корпоративное решение, включающее средства работы с географической и геометрической информацией.

SQLite/Spatialite — легковесная встраиваемая СУБД, ориентированная на мобильные и автономные решения, расширяемая посредством модуля Spatialite для поддержки пространственных данных.

Сравнение всех систем управления базы данных представлено в таблице №1.

Таблица № 1

Сравнение СУБД

СУБД	Поддержка стандартов	Типы данных	Поддержка распределённых БД
PostgreSQL / PostGIS	Полная реализация OGC, поддержка ISO/TC 211	Векторные, растры, 3D	Нет нативной поддержки, но может использоваться с распределёнными решениями (например, Citus)
Microsoft SQL Server	Базовая поддержка OGC	Векторные, частично 3D	Поддерживает горизонтальное масштабирование через Always On Availability Groups
MongoDB	Частичная поддержка OGC	Растровые, метаданные	Полная поддержка распределённых БД через sharding
Oracle Spatial and Graph	Полная реализация OGC, поддержка ISO/TC 211	Векторные, растры, 3D	Поддерживает распределённые БД через Oracle RAC (Real Application Clusters)
SQLite / SpatiaLite	Базовая поддержка OGC	Векторные, растры	Не поддерживает распределённые БД (автономное решение)

Проведённый анализ показал, что PostgreSQL с расширением PostGIS является универсальным решением, поддерживающим все типы геоданных и стандарты OGC. Её открытость и гибкость делают её предпочтительной для общедоступных проектов и научных исследований.

Oracle Spatial and Graph — оптимальный выбор для корпоративных и государственных задач, где важны масштабируемость, отказоустойчивость и сложные топологические операции.

Microsoft SQL Server демонстрирует тесную интеграцию с ArcGIS и удобство использования в Windows-среде, однако ограничен в поддержке 3D-данных по сравнению с PostGIS.

MongoDB эффективна для хранения метаданных и растровых файлов в микросервисных архитектурах, но не подходит для сложных пространственных операций.

SQLite/Spatialite актуальна для локального использования, но не применима в масштабных проектах, поэтому в дальнейшем не рассматривалась в работе.

Тестирование СУБД

Тестирование проводилось в тестовой среде, максимально приближенной к реальным условиям эксплуатации. Все компоненты системы развертывались на физических устройствах с одинаковыми техническими характеристиками.

Проведённое тестирование охватывало функциональное, нагрузочное и тестирование масштабируемости и отказоустойчивости четырёх СУБД — MongoDB, Microsoft SQL Server, PostgreSQL/PostGIS (включая PostGIS Raster и расширение Point Cloud/PDAL) и Oracle Database (с GeoRaster) — для трёх классов геоданных: растры, векторы, облака точек; функциональные метрики усреднялись по 100 последовательным запросам на одной ноде, а нагрузочные профили оценивались при росте числа одновременных клиентов и масштабировании кластера с 1 до 3 нод.

PostgreSQL/PostGIS продемонстрировала стабильную работу с векторными данными при использовании типа GEOMETRY (средние значения составили 6 487 мс для записи, 1 123 мс для чтения и 498 мс для удаления). С растровыми данными (PostGIS Raster) система показала хорошие результаты при распределённой архитектуре, хотя запись оказалась наиболее затратной операцией (28 764 мс против 4 892 мс для чтения и 3 987 мс для удаления). При работе с облаками точек (PCPATCH/PDAL) отмечалась асимметрия операций (18 764 мс на запись, 15 987 мс на чтение и 3 456 мс на удаление), но PostgreSQL/PostGIS показала хорошие результаты при параллельной обработке. PostgreSQL продемонстрировала отличную масштабируемость: в трёхнодовой конфигурации время отклика на 12 клиентах снижалось примерно до 26 000 мс для векторов, 85 000 мс для

растров (против 200 000 мс на одной ноде) и 60 000 мс для облаков точек (против 130 000 мс). Это подтверждает практическую пригодность PostgreSQL как универсального ядра ГИС с максимальной выгодой от горизонтального масштабирования.

Oracle Database хорошо справляется с растровыми данными через модуль GeoRaster, обеспечивая высокую производительность (10 567 мс на чтение и 2 345 мс на удаление). Для векторных данных используется тип SDO_GEOMETRY, который обеспечивает стабильную работу (8 456 мс на запись, 1 876 мс на чтение, 456 мс на удаление), но не оптимизирован для обработки больших массивов точек. Для облаков точек было принято решение хранить данные как сериализованные бинарные массивы (21 432 мс на запись, 6 897 мс на чтение и 2 345 мс на удаление), что повысило скорость доступа, хотя усложнило выполнение пространственных запросов. Oracle показал отличные результаты масштабируемости: при росте числа клиентов с 1 до 3 нод время отклика на 12 клиентах снижалось примерно с 230 000 мс до 150 000 мс и затем до 90 000 мс, а для векторов и облаков точек время отклика сокращалась до 30 000 мс и ещё заметнее при добавлении второй ноды.

Microsoft SQL Server показал надёжную работу с векторными данными при использовании типа GEOMETRY (7 895 мс на запись, 467 мс на чтение и 321 мс на удаление). Однако при скачивании файлов наблюдался значительный рост времени выполнения, что потребовало дополнительной оптимизации — хранения файлов как бинарных объектов. В двух- и трёхнодовых конфигурациях система демонстрировала улучшенную производительность, особенно при средней нагрузке: задержки на 12 клиентах снижались до 26 000 мс. MSSQL Server является хорошим решением для ГИС-приложений с акцентом на векторные данные и распределённую архитектуру.

MongoDB может быть использована для хранения растровых данных через механизм GridFS (4 215 мс на запись и 2 486 мс на чтение), однако она проигрывает традиционным СУБД по скорости обработки и наличию встроенной поддержки пространственных запросов. База данных показала удовлетворительные результаты при малой нагрузке, но быстро достигала предела своих возможностей при увеличении числа клиентов: на 12 клиентах время отклика составляло 117 000 мс на одной ноде против 78 459 мс при двух и 40 000 мс при трёх. MongoDB рекомендуется использовать только в случае необходимости гибкого хранения метаданных вместе с растровым контентом.

Заключение

Таким образом, при разработке геоинформационных систем выбор платформы зависит от приоритетов: для универсального решения и работы с разнородными геоданными оптимальна PostgreSQL/PostGIS, для высоконагруженных корпоративных приложений – Oracle Database, для векторных сценариев — Microsoft SQL Server, а для гибкого хранения метаданных вместе с растровым контентом – MongoDB.

Литература

1. Миронова Ю. Н. Геоинформационные системы // Актуальные проблемы гуманитарных и естественных наук. 2014. №. 3-1. С. 63-65.
2. Филипенков А. В., Кузьмин Е. Л. Сравнение существующих систем управления базами данных в целях выбора наилучшей при реализации требований по сокращению затрат и импортозамещению // Газовая промышленность. 2018. №. 4 (767). С. 24-29.
3. Христодуло О. И. Совместное описание пространственных и атрибутивных данных на основе многомерных информационных объектов // Программные продукты и системы. 2011. №. 3. С. 48-54.

4. Zhang J., Ma B., Huang J. Deploying GIS services into the edge: A study from performance evaluation and optimization viewpoint // Security and Communication Networks. 2020. V. 2020. №. 1. pp. 8822990.
5. Sadaf A., Bruck J., Tahmasebi M. A comparative evaluation of GIS spatial analysis tools for prioritizing natural and hybrid shoreline solutions // Annals of GIS. 2025. V. 31. №. 1. pp. 67-89.
6. Мальцев К. А., Голосов В. Н., Гафуров А. М. Цифровые модели рельефа и их использование в расчётах темпов смыва почв на пахотных землях // Ученые записки Казанского университета. Серия Естественные науки. 2018. Т. 160. №. 3. С. 514-530.
7. Данилова Л. О., Рашевский Н. М., Рекунов С. С., Трудов Я. А., Гуртяков А. С. Классификация плотного облака точек при моделировании рельефа // Academia. Архитектура и строительство. 2024. №. 3. С. 76-81.
8. Выстрчил М. Г., Балтыжакова Т. И., Савина А. В. Алгоритм оценки точности полигональных TIN-поверхностей, получаемых из разреженных облаков точек // Вестник СГУГиТ (Сибирского государственного университета геосистем и технологий). 2024. Т. 29. №. 3. С. 5-19.
9. Di L. The development of remote-sensing related standards at FGDC, OGC, and ISO TC 211 // Proc. IGARSS. 2003. pp. 643–647. DOI: 10.1109/IGARSS.2003.1293868.
10. Giannecchini S., Spina F., Nordgren B., Desruisseaux M. Supporting Interoperable Geospatial Data Fusion by adopting OGC and ISO TC 211 standards // Proc. ICIF. 2006. pp. 1–8. DOI: 10.1109/ICIF.2006.301751.

References

1. Mironova Yu. N. Aktual'nye problemy gumanitarnykh i estestvennykh nauk. 2014. № 3-1. Pp. 63–65.
2. Filipenkov A. V., Kuz'min E. L. Gazovaya promyshlennost'. 2018. № 4 (767). Pp. 24–29.

3. Khristodulo O. I. Programmnye produkty i sistemy. 2011. № 3. Pp. 48–54.
4. Zhang J., Ma B., Huang J. Security and Communication Networks. 2020. V. 2020. № 1. P. 8822990.
5. Sadaf A., Bruck J., Tahmasebi M. Annals of GIS. 2025. V. 31. № 1. Pp. 67–89.
6. Maltsev K. A., Golosov V. N., Gafurov A. M. Uchenye zapiski Kazanskogo universiteta. Seriya Estestvennyye nauki. 2018. T. 160. № 3. Pp. 514–530.
7. Danilova L. O., Rashevskiy N. M., Rekunov S. S., Trudov Y. A., Gurtyakov A. S. Academia. Architecture and Construction. 2024. № 3. Pp. 76–81.
8. Vystrechil M. G., Baltyzhakova T. I., Savina A. V. Vestnik SGUGiT. 2024. T. 29. № 3. Pp. 5–19.
9. Di L. Proc. IGARSS. 2003. Pp. 643–647. DOI: 10.1109/IGARSS.2003.1293868.
10. Giannecchini S., Spina F., Nordgren B., Desruisseaux M. Proc. ICIF. 2006. Pp. 1–8. DOI: 10.1109/ICIF.2006.301751.

Дата поступления: 15.09.2025

Дата публикации: 25.10.2025