

ПАРАЛЛЕЛЬНАЯ СИСТЕМА УПРАВЛЕНИЯ ПОЛНООБЪЕКТНЫМИ ЗАЩИЩЕННЫМИ БАЗАМИ ДАННЫХ КАРТОГРАФИЧЕСКИХ СЦЕН

Р.Ф. Гибадуллин, С.В. Пыстогов

*Казанский национальный исследовательский технический университет
им. А.Н. Туполева*

Анализируются тенденции универсализации баз данных картографических сцен (БД КС) и их защиты. Рассматривается исследуемая архитектура СУБД КС. Приводятся разработки: схемы универсальной БД КС с ассоциативной защитой, метода учета погрешностей векторизации, параллельных алгоритмов и программ синтеза БД КС на вычислительном кластере.

Введение

В настоящее время системы управления картографическими базами данных развиваются на стыке многих научных дисциплин и применяются в различных сферах человеческой деятельности, например при изучении природных ресурсов и управлении ими, при создании и ведении кадастров и управлении муниципальными образованиями и в других сферах.

Картографическая продукция имеет следующие особенности:

- Отличается высокой себестоимостью работ по ее получению. Топографами и геодезистами затрачиваются немалые усилия на формирование этой продукции.
- Для последующей обработки требуются векторизация растровых изображений, устранение дефектов топологической структуры при создании векторных карт и организация хранения их в базе данных картографической сцены и соответствующей системы управления.
- Картографическая сцена содержит контурные изображения объектов разных типов с их привязкой к координатам.
- Картографические базы данных характеризуются большой объемностью. Согласно общим сведениям о геоинформационной системе «Панорама», разработанной с учетом широкого круга пользовательских потребностей, векторная карта может содержать несколько тысяч листов. Всего один лист карты может содержать до 4 миллиардов объектов. Например объем векторной кадастровой карты Российской Федерации без учета семантического описания объектов может достигать нескольких десятков гигабайт [1].

Коммерческая и другая ценность картографической продукции требует организации ее информационной безопасности. Обработка запросов к защищенным базам данных картографических сцен связывается со значительными временными затратами. Это приводит к необходимости организации параллельной обработки: распределения картографических баз данных по узлам вычислительного кластера, применения технологий MPI, NVIDIA CUDA и др. Присутствие на тематических слоях карты различных типов объектов (точечных, линейных и площадных) и применение к ним единого механизма защиты требует разработки полнообъектной структуры хранения таких данных. Все это обуславливает необходимость разработки полнообъектной защищенной системы управления базами данных картографических сцен (СУБД КС).

1. Существующие решения

На сегодняшний день для организации хранения и управления пространственными данными в СУБД широко применяется серверное программное обеспечение ArcSDE в составе ArcGIS Server [2]. ArcSDE тесно интегрирована с пакетами ArcEditor, ArcInfo, ArcView для работы с многопользовательскими картографическими базами данных. Например, ArcView действует как клиентское приложение для просмотра и анализа картографических данных.

ArcSDE использует сжатый бинарный формат для хранения картографических данных. Этот бинарный формат был разработан для нескольких платформ СУБД. Важно отметить, что многопользовательская база данных хранится не в собственной структуре, а в стандартных таблицах, управляемых СУБД. В результате пользователи могут обращаться к многопользовательской базе данных посредством ArcSDE.

Ведущие места среди СУБД, обладающих встроенными механизмами защиты баз данных, занимают: *Oracle*, *Microsoft SQL Server*, *Sybase Server*. Каждая из них имеет свои определенные преимущества. *Oracle* обладает такими возможностями, как защита и ограничения доступа пользователей к важным данным и приложениям; защита данных при передаче по сети. Но при этом она имеет серьезный недостаток – отсутствие шифрования данных непосредственно в базе данных, что делает систему уязвимой при внутренних атаках (в случае кражи жестких дисков). Большинство средств защиты *Microsoft SQL Server* обеспечивается функциональными возможностями операционной системы, а не с помощью инструментов СУБД. Например, файловая система *NTFS* операционной системы *Windows* обеспечивает полное шифрование каталогов и файлов. Эта функциональная возможность *Windows* используется в СУБД для защиты файлов базы данных от прямых атак. Такой подход не обеспечивает высокой производительности при работе с защищенными базами данных. *Sybase Server* предлагает надежную защиту, включая стандарт тройного шифрования данных (*triple DES*) в базе данных. Это хорошо защищает СУБД от внутренних атак на данные. Защита, предлагаемая в СУБД компании *Sybase*, превосходит защиту *Oracle* и является более интегрированной по сравнению с защитой *Microsoft SQL Server*.

Перечисленные СУБД достаточно универсальны, однако построение СУБД КС, ориентированных на работу с защищенными БД картографии, может существенно повысить эффективность управления такими БД по критерию быстродействия при требуемом уровне стойкости защиты.

2. Базовые принципы формирования полнообъектной защищенной БД КС

Базовые принципы формирования базы данных картографических сцен для случая точечных объектов подробно описаны в статье [3].

Соответствующая структура базы данных показана на рисунке (рис.1).

Описание отношений:

- α, β, γ определяют тип слоя: слой точечных, слой линейных и слой площадных объектов соответственно.
- *Name* – название слоя.
- *Code** – код тематического слоя.
- *Obj_name* – имя точечного объекта.
- *Obj_code** – код точечного объекта.
- *Cluster_num** – номер кластера в тематическом слое.
- *Global_X**, *Global_Y** – глобальные координаты нижнего левого угла данного кластера.

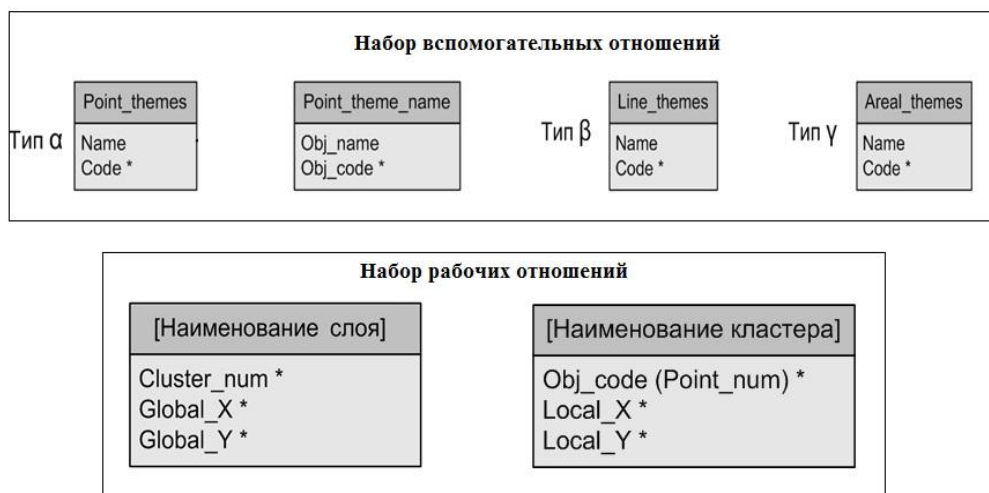


Рис. 1

- $Obj_code (Point_num)^*$ – код точечного объекта (номер узловой точки линейного или площадного объекта).
 - $Local_X^*$, $Local_Y^*$ – локальные координаты точки/узла внутри кластера.
 - Данные по атрибутам, отмеченные звездочкой, хранятся в сокрытом *виде*.
 - Наименование слоя и наименование кластера являются составными компонентами.
 - [Наименование слоя] = <Тип_слоя> <Стегокод_слоя>.
 - [Наименование кластера] = <Наименование слоя> <№ кластера>.
- Различают рабочие и вспомогательные отношения.

Рабочие отношения – набор записей одинакового формата, которые несут в себе основную информацию об объектах картографируемого участка местности в рамках предлагаемого подхода хранения данных в терминах «объект–координаты».

Вспомогательные отношения – это набор записей одинакового формата, несущих дополнительную информацию об объектах, тематических слоях картографируемого участка местности. Данная информация необходима для правильной работы системы с хранимыми данными.

Каждое рабочее отношение с именем типа [Наименование слоя] описывает отдельный тематический слой (фрагменты внутри слоя). Отношения с именем типа [Наименование кластера] описывают по одному фрагменту (точечные объекты и узловые точки внутри него).

В качестве пользовательского приложения для анализа карт в СУБД КС выбрана геоинформационная система (ГИС) MapInfo. Так как она оперирует координатными данными в виде 31-битных целочисленных переменных (знаковый разряд не используется), то минимальная единица деления экстенда карты, полученной посредством картографической проекции Гаусса-Крюгера (рис. 2), составит:

$$(X_{\max} - X_{\min}) / (2^{31} - 1) = 2 \times 10^7 / 2 \times 10^9 = 0,01 \text{ м.}$$

Погрешность векторизации по координатным осям в таком случае составляет

$$\mu = 0,01/2 = 0,005 \text{ м.}$$

Неравенство $\varepsilon > \mu$ неприемлемо из практических соображений, где ε – погрешность представления точечного объекта в разрабатываемой параллельной системе. Равенство погрешностей ($\varepsilon = \mu$) достигается сужением границ защищаемого участка картографической сцены до $10 \times 10 \text{ км}^2$ с шагом локальной координатной сетки, равным 0,01 м.

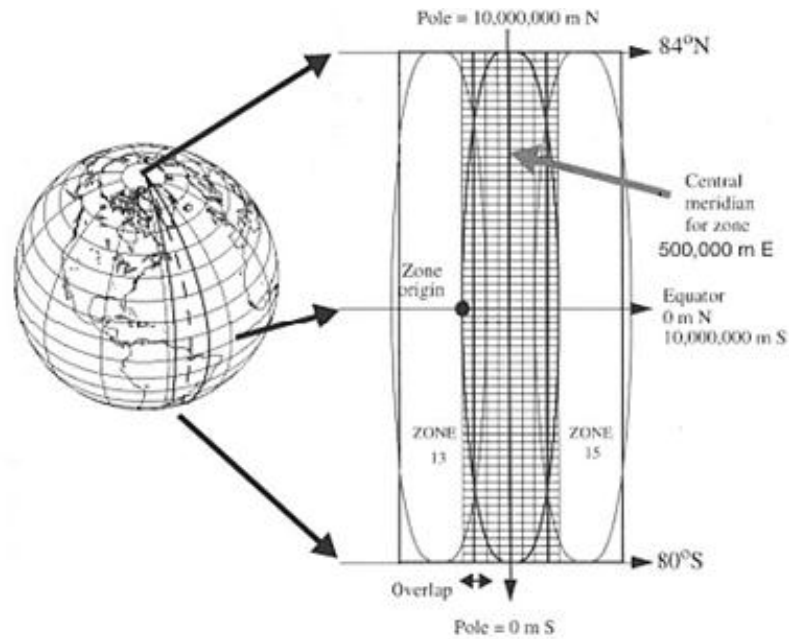


Рис. 2. Геодезическая зона в проекции Гаусса-Крюгера

Таким образом, шаг локальной координатной сетки при формировании баз данных картографических сцен должен определяться минимальной единицей деления экстенда карты пользовательской ГИС.

3. Архитектурные особенности системы

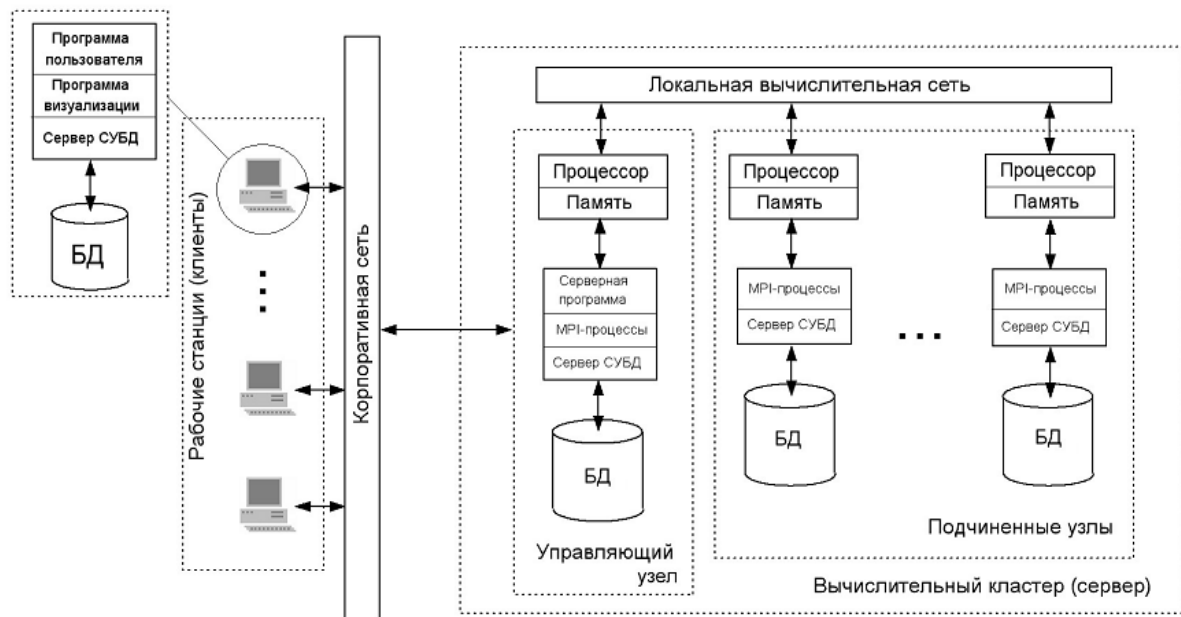


Рис. 3

В системе за основу берется «интеллектуальная» файл-серверная организация взаимодействия. Под интеллектуальностью в данном случае понимается выполнение сервером не только функции хранения, но и функции предварительной обработки поступающих запросов и «интеллектуальный» сбор результатов. Файловый сервер принимает запросы, поступающие по сети от компьютеров-клиентов, и, в соответствии с запросами, передает им требуемые данные. Клиентская часть получает от сервера ответ (ре-

зультат). Данный ответ должен быть предварительно обработан программой клиента для возможности дальнейшей работы пользователя с полученными данными.

Серверная часть состоит из множества однородных узлов, среди которых выделяется один узел, выполняющий функции управления. Именно этот узел является ключевым в задачах получения запросов от клиентов, распределения задач между счетными узлами кластера, сбора результатов и т.д. Остальные узлы кластера являются вычислительными.

К серверу по корпоративной сети подключается множество пользователей. На их рабочих местах должно быть установлено специальное программное обеспечение для работы с БД КС.

4. Синтез БД КС

1. Считывание информации об объектах (точечных, линейных и площадных) картографируемого участка местности из файлов формата MIF/MID в текстовый файл.
2. Создание на управляющем узле кластера структуры будущей базы данных на основе информации, вводимой пользователем или считываемой из входных файлов.
3. Равномерное распределение полученного файла с информацией об объектах по всем узлам кластера (либо по определенным узлам согласно указаниям администратора).
4. Параллельная обработка информации об объектах узлами кластера. Заполнение таблиц базы данных на управляющем узле.
5. После формирования базы данных с открытой информацией на управляющем узле происходит распределение таблиц по узлам кластера для последующих хранения и обработки запросов. Распределению методом горизонтального деления подвергаются рабочие отношения, в то время как вспомогательные отношения полностью дублируются в базе данных каждого узла.
6. После распределения выполняется операция маскирования всеми узлами. Её можно назвать параллельной, т.к. все узлы в совокупности обрабатывают полностью всю БД КС, хотя каждый узел маскирует свой набор записей.

Литература

1. Сообщество специалистов в области ГИС и ДЗЗ [Электронный ресурс]. GIS-Lab © 2012. URL: <http://www.gis-lab.info>.
2. Картографическая платформа для вашей организации [Электронный ресурс]. Esri © 2012. URL: <http://www.arcgis.com>.
3. Райхлин В.А., Вершинин И.С., Гибадуллин Р.Ф. Конструктивное моделирование систем в приложении к защите данных картографии // Методы моделирования: Труды Респ. научн. семинара АН РТ. – Казань: «ФЭН» (Наука), 2010. Вып. 4. С. 68–95.