

УДК 004.75

Р.А. Родригес Залепинос

МЕТОДЫ РАСПРЕДЕЛЕННОГО ВЫПОЛНЕНИЯ АНАЛИТИЧЕСКИХ ЗАПРОСОВ НАД БОЛЬШИМИ ОБЪЕМАМИ РЕТРОСПЕКТИВНЫХ ГЕОГРАФИЧЕСКИХ ДАННЫХ

Описаны методы и технологии, которые впервые предоставили возможность доступа к большому объему ретроспективных климатических данных в реальном времени для множества одновременных пользователей. С помощью графического 3D-интерфейса пользователь может интуитивно передвигаться во времени, визуализировать данные и интерактивно выполнять запросы вычисления статистических величин. Необходимые данные для визуализации и расчетов не хранятся на машине пользователя, а доставляются ему в реальном времени через Интернет с компьютерного кластера. Приведенные результаты экспериментов для нескольких тысяч клиентов подтверждают эффективность подхода. Представленные в данной статье методы реализованы в клиент-серверной системе Climate Wikience. Она способна передавать и визуализировать в реальном времени данные климатического архива повторного анализа ERA-Interim, спутниковые данные MODIS L3, а также строить изолинии по этим данным. Climate Wikience позволит существенно улучшить понимание ретроспективных географических данных большого объема и предоставит беспрецедентные возможности исследователям наук о Земле.

Компьютерный кластер; распределенная обработка данных; 3D-визуализация; реальное время.

R.A. Rodrigues Zalipynis

METHODS OF DISTRIBUTED PROCESSING OF ANALYTICAL QUERIES OVER BIG VOLUMES OF RETROSPECTIVE GEOGRAPHICAL DATA

The paper describes methods and technologies that for the first time enabled the real-time access to large volumes of retrospective climate data for numerous concurrent clients. With graphical 3D interface, a user is able to navigate intuitively in time, visualize data and interactively issue queries for calculating statistical values. The required data for visualization and computing are not stored on the client machine. They are delivered in real-time over Internet from computer cluster. Provided experimental results for several thousand users evidence the effectiveness of the approach. The methods presented in this paper are implemented in client-server system Climate Wikience. It is capable to transfer and visualize in real-time data from ERA-Interim climate reanalysis archive, MODIS L3 satellite data as well as build isolines based on these data. Climate Wikience will enable significantly enhance the understanding of large volumes of retrospective geographical data and provide unprecedented opportunities for Earth science researchers.

Computer cluster; distributed data processing; 3D-visualization; real time.

Введение. Современный мир испытывает взрывообразный рост количества данных, которые генерируются чрезвычайными темпами. Географические данные, к которым относятся климатические архивы повторного анализа [1, 2] и спутниковые данные дистанционного зондирования Земли [3], предоставляют беспрецедентные возможности лучше понять, а следовательно, подготовиться и адаптироваться к штормам, засухам, суровым погодным условиям и будущему климату.

Климатические архивы повторного анализа являются наиболее популярными в исследованиях климата, начиная от построения простых трендов до отслеживания циклонов [4]. Первый такой архив – NCEP/NCAR – был выпущен в 1996 году. С тех пор более 12 200 публикаций ссылаются на статью с его описанием [5] (25 февраля 2012 г.).

Данные хранятся в особых научных форматах, в частности NetCDF [6] и HDF [7]. Помимо самих данных, файлы насыщены множеством метаданных, особенно в спутниковых данных. Каждый раз исследователи повторяют те же самые рутинные операции по заказу файлов, написанию небольших программ для извлечения данных из них с помощью специальных библиотек, дублируя усилия своих коллег. В то же время они постоянно сталкиваются с теми же трудностями из-за отсутствия визуальных инструментов, которые позволили бы интуитивно и интерактивно исследовать существующие данные. Например, даже их самая простая визуализация очень трудоемкая, требует знаний деталей форматов, отвлекает от первоначальных целей и задерживает результаты.

Данные. Данные повторного анализа содержат более 80 метеорологических показателей с 1979 года по настоящее время с интервалом в 6 часов. К показателям относятся температура, влажность, атмосферное давление на поверхности и среднем уровне моря, скорость и направление ветра. Для каждого момента времени значения показателя размещены в узлах регулярной широтно-долготной решетки с разрешением $1,5 \times 1,5^\circ$. Решетки одного показателя представляются двумерными массивами и разбиты по времени на ряд файлов формата NetCDF. Каждый файл хранит один год данных. Размер файла для одного показателя приблизительно составляет 190 мегабайт.

Спутниковые данные MODIS L3 также расположены на регулярной широтно-долготной решетке, но с разрешением $1 \times 1^\circ$. Отличие состоит в том, что это не точечные значения в узле решетки, а статистическое значение для всей площади ячейки $1 \times 1^\circ$, которые получаются обработкой более детальных сцен. Данные хранятся с 2000 года по настоящее время с интервалом в 1 день. Каждый файл представляет собой набор решеток для более 600 параметров за один день. Размер одного файла приблизительно равен 100 мегабайт.

Организация хранения данных на узлах компьютерного кластера. Одной из главных целей рассматриваемой высокопроизводительной распределенной системы хранения и анализа является предоставление доступа к любой части данных в реальном времени для нескольких тысяч одновременных пользователей. Большой объем имеющихся данных не может быть загружен в оперативную память, поэтому при каждом запросе существует большая вероятность обращения к диску для получения следующей порции данных.

Следовательно, для ускорения выполнения запросов необходимо использовать методы декластеризации. В данной системе используется репликация данных для балансировки нагрузки и отказоустойчивости.

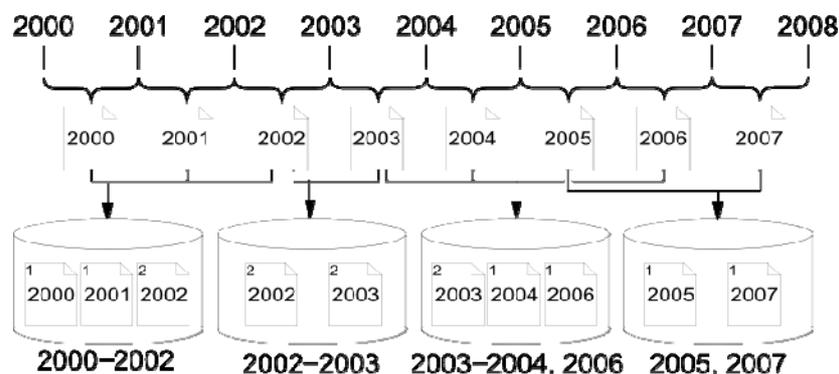


Рис. 1. Вариант размещения данных на узлах компьютерного кластера

Все файлы для одной переменной имеют приблизительно одинаковый размер. Набор этих файлов равномерно распределяется по узлам кластера. Каждый файл дублируется на несколько разных узлов. При этом наиболее часто запрашиваемые файлы могут дублироваться большее количество раз, чем другие файлы (рис. 1).

На рис. 1 сверху изображена шкала времени. Каждый файл хранит данные за равные отрезки времени. Цилиндрами показаны узлы компьютерного кластера. Подписи под ними – промежутки времени, за которые хранятся данные на узле.

Архитектура системы и распределенная схема выполнения аналитических запросов. Система состоит из трех компонент: клиентов, рабочих узлов компьютерного кластера и шлюза. Шлюз и узлы кластера объединены в локальную сеть. Клиенты устанавливают соединение со шлюзом через Интернет и иницируют запросы (рис. 2).



Рис. 2. Схема распределенного выполнения аналитических запросов

Аналитический запрос, выполняемый клиентом, представляет собой вычисление статистических величин для требуемой регулярной широтно-долготной решетки. Клиент задает переменную и момент времени, для которого необходимо получить решетку (1) в виде строки с SQL-подобным синтаксисом. Например, «SELECT DATA FROM r2.pressure.msl WHERE TIME = 01.01.2003 00:00» вернет регулярную широтно-долготную решетку давления на среднем уровне моря за 1 января 2003 года.

Шлюз получает строку, разбирает ее, выбирает рабочий узел с данными и иницирует их чтение с диска (2). Когда данные готовы, рабочий узел асинхронно передает их на шлюз (3), который, в свою очередь, упаковывает их и отправляет клиенту (4). За один запрос могут вернуться данные для одной решетки. Если необходимо несколько решеток, то нужно выполнить несколько запросов. При таком подходе система лучше масштабируется для большого количества клиентов.

Статистические расчеты над решеткой выполняются на клиенте. Это позволяет значительно разгрузить систему и обеспечить ее доступность для большего числа клиентов, а также существенно упростить архитектуру облачного сервиса, задачи балансировки нагрузки и обеспечения безопасности.

Анализ и визуализация данных. Для удобства экспертного анализа больших объемов ретроспективных географических данных разработан графический интерфейс. Он впервые в мире способен визуализировать в 4D (3D+время) в реальном времени данные климатического архива повторного анализа и спутниковых данных MODIS L3. Под визуализацией в реальном времени понимается доставка по Интернету и вывод данных на экран без заметной задержки для пользователя.

Визуальный интерфейс представляет собой интерактивную трехмерную модель планеты «Земля». Пользователь может изменять масштаб отображения и вращать сферу. Для удобного исследования ретроспективных данных разработан специальный инструмент временных линеек для передвижения во времени.

В узлах регулярной широтно-долготной решетки с регулируемым шагом отображаются маркеры, цвета и размеры которых пропорциональны значению переменной исходных данных в узле. Разработан и реализован алгоритм построения и визуализации изолиний для регулярной широтно-долготной решетки. Он отличается высокой скоростью и генерацией изолиний в виде замкнутых полигонов, а не просто набора отрезков.

Пользователь может выделить с помощью визуальных инструментов произвольную область на глобусе. Система определяет узлы решетки, попавшие в выбранную область, и вычисляет описательную статистику для значений в этих узлах.

При изменении позиции камеры либо времени для ретроспективных данных, прозрачно для пользователя формируется строка запроса к распределенной системе на компьютерном кластере и инициируется его передача и выполнение. В результате запроса возвращается новая решетка, которая отображается на глобусе. Время от создания строки запроса до возвращения результата составляет меньше 500 мс.

Экспериментальная оценка производительности системы. Для моделирования большого количества клиентов была разработана многопоточная программа. Каждый поток моделирует одного клиента. Шлюз, рабочие узлы и машины клиентов находились в одной 1 Гбит/с локальной сети. Шлюз содержит 8 GB оперативной памяти, Cent OS 6.0, Core i7 (8 ядер, 3.46 ГГц). Использовалось 6 машин с Windows 7, 4 GB ОП, Intel Core Quad 2 (2.66ГГц) – четыре для клиентов, на двух рабочих узлах виртуальные машины (Ubuntu 10.04, 1GB ОП и одно ядро). Рабочие узлы были заполнены ранее описанными данными. Для обработки запросов от клиентов на шлюзе было запущено 64 потока и по 32 потока на рабочих узлах.

На клиентских машинах запускалось 512 потоков, что в общем составляет 2048 клиентов. Каждый поток случайным образом выбирает тип данных и время, для которого составляет запрос и инициирует его выполнение. Поток генерирует запросы непрерывно, один за другим. Следующий запрос не генерируется, пока не получен результат выполнения предыдущего запроса. Моделирующие программы работали в таком режиме в течение 5 минут. За это время система успела ответить на 66 282 запроса, т.е. 220 запросов в секунду. В реальной ситуации это значение может быть еще выше, поскольку наибольшее количество времени для выполнения запроса расходуется на клиенте из-за задержек, связанных с большим числом потоков. Статистика приведена для всех этапов выполнения запроса, начиная от генерации строки, передачи данных и до его распаковки на стороне клиента.

Выводы. Описанные методы и технологии позволяют в реальном времени организовать распределенное выполнение запросов и доставку данных тысячам одновременных клиентов. Клиентское приложение доступно для свободной загрузки по адресу wikience.donntu.edu.ua. Сейчас система способна в реальном времени визуализировать в 3D климатические архивы повторного анализа и спутниковые данные MODIS L3, хранящиеся на кластере Донецкого национального технического университета. На сегодняшний день в мире нет аналогов этой системе.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Kalnay et al.* The NCEP/NCAR 40-year reanalysis project, Bull. Amer. Met. Soc. (77), 1996. – P. 437-470.
2. *Kanamitsu M. et al.* NCEP-DOE AMIP-II Reanalysis (R-2), Bull. Amer. Met. Soc., 2002. – P. 1631-1643.

3. *Savtchenko A. et al.* Terra and Aqua MODIS products available from NASA GES DAAC // *Advances in Space Research.* – 2004. – № 34. – P. 710-714.
4. *Hannachi A., Awad A., Ammar K.* Climatology and classification of Spring Saharan cyclone tracks // *Clim. Dyn.* – 2011. – № 37. – P. 473-491.
5. Страница поиска Google Scholar, 25 Февраля 2012 г. <http://scholar.google.com.ua/scholar?q=Kalnay+The+NCEP%2FNCAR+40-year+reanalysis+project>.
6. *Rew R. et al.* The NetCDF Users Guide, UNIDATA, 2011.
7. The HDF Group, HDF5 Reference Manual, 2011.

Статью рекомендовал к опубликованию д.т.н., профессор А.В. Боженюк.

Родригес Залепинос Рамон Антонио – Донецкий национальный технический университет; e-mail: rodrigues@csm.donntu.edu.ua; 83001, Украина, г. Донецк, ул. Артема, 58; тел.: +380623010769; кафедра компьютерных систем мониторинга; аспирант.

Rodrigues Zalipynis Ramon Antonio – Donetsk National Technical University; e-mail: rodrigues@csm.donntu.edu.ua; 58, Artema street, Donetsk, 83001, Ukraine; phone: +380623010769; the department of Computer Systems for Monitoring; postgraduate student.

УДК 681.327

Л.С. Берштейн, А.В. Боженюк

ОПРЕДЕЛЕНИЕ ВНУТРЕННЕЙ УСТОЙЧИВОСТИ НЕЧЕТКОГО ТЕМПОРАЛЬНОГО ГРАФА*

Рассматривается понятие нечеткого темпорального графа, который является обобщением с одной стороны нечеткого, а с другой стороны – темпорального графов. В нечетком темпоральном графе степень связности вершин изменяется в дискретном времени. Введены понятия максимального нечеткого внутренне устойчивого подмножества темпорального подграфа с наибольшей степенью внутренней устойчивости, а на его основе – нечеткого множества внутренней устойчивости как инварианта нечеткого темпорального графа. Рассмотрен метод нахождения всех максимальных нечетких внутренне устойчивых множеств с наибольшей степенью нечеткости, что позволяет находить нечеткое множество внутренней устойчивости. Рассмотрен пример нахождения нечеткого множества внутренней устойчивости нечеткого темпорального графа.

Нечеткий темпоральный граф; суграф нечеткого темпорального графа; степень инцидентности; нечеткое множество внутренней устойчивости.

L.S. Bershtein, A.V. Bozhenyuk

INTERNAL STABLE DEFINITION OF FUZZY TEMPORAL GRAPH

In this paper the notion of temporal graph is considered. Which one is a generalization of a fuzzy graph on the one hand, and a temporal graph on the other hand. The incidence of graph vertices is changed in the discrete time in fuzzy temporal graph. The notions of maximum fuzzy internal stable subset and internal stable fuzzy of fuzzy temporal graph are introduced. The method of definition of internal stable fuzzy set is considered. The example of definition of internal stable fuzzy set is considered too.

Fuzzy temporal graph; subgraph of fuzzy temporal graph; incidence degree; internal stable fuzzy set.

* Работа выполнена при поддержке гранта РФФИ, проект № 10-01-00029а.