

МОДЕЛИРОВАНИЕ И СОЗДАНИЕ РАСПРЕДЕЛЕННЫХ СИСТЕМ ВИЗУАЛИЗАЦИИ ГЕОПРОСТРАНСТВЕННЫХ РЕЗУЛЬТАТОВ ОБРАБОТКИ СПУТНИКОВЫХ ДАННЫХ

Н.И. Ильин, А.Н. Кравченко

Институт космических исследований НАН и НКА Украины

В даній роботі запропоновано підхід до моделювання Web-орієнтованих розподілених систем візуалізації геопросторових даних [1], наведено модель продуктивності та розглянуто реалізацію запропонованого підходу в сервісах екологічного моніторингу Інституту космічних досліджень НАНУ-НКАУ.

1. Предусловия возникновения распределенных систем визуализации

Прогресс в области использования спутниковых данных для ряда задач, в частности мониторинга экологических параметров окружающей среды, наблюдения и прогнозирования развития природных и техногенных катастроф, привел к возникновению необходимости в эффективных и недорогих специализированных системах визуализации геопространственных данных – информации о некоторых пространственных характеристиках земной поверхности и околоземного пространства.

На текущем этапе развития было создано ряд специализированных сервисов визуализации и предоставления данных, преимущественно web-ориентированных [2]. Уточнение существующих и создание новых моделей на основе интеграции результатов этих сервисов привели к возникновению проблем стандартизации существующих систем, часто настолько серьезным, что сервисы мониторинга приходится разрабатывать заново, выбрасывая предыдущие разработки.

Сложность обработки и большие объемы обрабатываемых данных привели к возникновению распределенных систем визуализации и распределенных каталогов данных, близких по своей структуре к Grid-системам. Очевидность стандартизации интерфейсов в таких системах не вызывает сомнений, что оправдывает деятельность в области выработки стандартов предоставления и визуализации. Наиболее успешными на сегодня являются стандарты WMS [3], WFS [4] и WCS [5] консорциума OGC, принятые в международной системе GEOSS.

Несмотря на прогресс в области реализации стандартов и построения конечных систем визуализации, оценка и предсказание параметров производительности таких систем все еще остается в большинстве случаев предметом экспериментального исследования уже реализованной системы. Принимая во внимание стоимость аппаратных и программных средств, ошибки проектирования как результат отсутствия методов оценки рабочих характеристик разрабатываемой системы визуализации приводит к значительным финансовым потерям. Поэтому исследования в области моделирования и оценки производительности систем визуализации имеют, кроме научного, большое прикладное значение.

2. Подход к моделированию и модель производительности Web-ориентированной распределенной системы визуализации

Основными характеристиками (метриками) производительности являются: 1. Время отклика (response time), 2. Пропускная способность (throughput), 3. Готовность (availability), 4. Стоимость (cost).

Предлагаемый подход заключается в построении сети массового обслуживания на основе представления компонентов визуализации как систем массового обслуживания с последующим применением методов операционного анализа для получения оценок граничных значений времени отклика и пропускной способности всей системы на основе экспериментальных усредненных характеристик представляемых компонент.

Рассмотрим систему из $k = \overline{1, n}$ ресурсов. Каждый ресурс характеризуется следующими параметрами:

X_k – средняя пропускная способность (производительность) k ;

S_k – среднее время обработки запросов на ресурсе k за одно посещение;

V_k – среднее число посещений ресурса k ;

D_k – среднее время обработки запроса ресурсом k ;

U_k – коэффициент использования ресурса k ;

R_k – среднее время отклика на запрос ресурса k ;

N – среднее число запросов в сети массового обслуживания;

λ – средняя частота поступления запросов к ресурсу сети массового обслуживания.

Z – среднее время ожидания пользователя перед отправкой запроса;

В дальнейших выкладках слово «среднее» будет пропущено. В случае, когда система открыта (число запросов неограничено) в состоянии равновесия ($\lambda = X$) имеет место зависимость:

$$U_k = X_k \times S_k, \quad X_k = \lambda \times V_k \Rightarrow U_k = \lambda \times D_k \leq 1$$

для всех k . Поэтому

$$U_{\max} = \lambda \times D_{\max} \leq 1,$$

где D_{\max} – необходимость обслуживания на ресурсе с максимальным временем обслуживания. Введем параметр частоты насыщения запросов:

$$\lambda_{\text{sat}} = \frac{1}{D_{\max}}.$$

Время отклика для открытой системы:

- Задержки при поступлении запросов отсутствуют:

$$D = \sum_k D_k \Rightarrow R = D;$$

- n запросов поступают одновременно каждые $\frac{n}{\lambda}$ единиц времени.

В случае закрытых систем (количество запросов в системе ограничено) предельные значения метрик производительности:

для больших нагрузок:

$$U_k(N) = X(N) \times D_k \leq 1, \quad \forall k \Rightarrow X(N) \leq \frac{1}{D_{\max}};$$

для малых нагрузок:

$$X = \frac{1}{D + Z}.$$

Асимптоты в случае отсутствия задержек:

- для каждого запроса $X = \frac{1}{D + Z}$;
- для N запросов $X = \frac{N}{D + Z}$ (верхняя граница).

Асимптоты в случае, когда каждый запрос задерживается другими $N-1$ запросами:

- суммарное время пребывания в системе $(N-1)D + D = ND$;
- суммарное время между обработкой каждого запроса $ND + Z$;
- пропускная способность для одного запроса $\frac{1}{ND + Z}$;
- нижний предел $X(N) = \frac{N}{ND + Z}$.

Точка пересечения верхней и нижней асимптот называется точкой насыщения. После точки насыщения значение производительности остается постоянным и не зависит от числа запросов:

$$\frac{1}{D_{\max}} = \frac{N^*}{D+Z} \Rightarrow N^* = \frac{D+Z}{D_{\max}}$$

В общем случае, не детализируя классы запросов, для пропускной способности X справедливо

$$\frac{N}{ND+Z} \leq X(N) \leq \min\left(\frac{1}{D_{\max}}, \frac{N}{D+Z}\right)$$

Для времени отклика R :

$$\max\left(D_{\max}, \frac{D+Z}{N}\right) \leq \frac{R(N)+Z}{N} \leq \frac{ND+Z}{N}, \max(ND_{\max} - Z, D) \leq R(N) \leq ND$$

Таким образом можно получить граничные значения производительности компонент системы визуализации, а также оценить производительность всей инфраструктуры в целом.

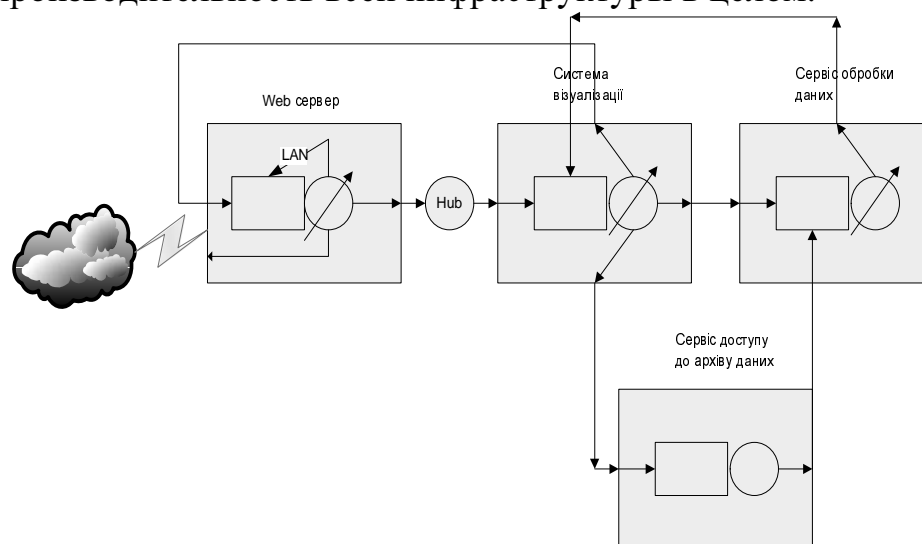


Рис. 1. Система визуализации как сеть массового обслуживания.

Рассмотрим применение предлагаемого подхода на примере системы визуализации, разрабатываемой Институтом космических исследований НАНУ-НКАУ. Компоненты системы визуализации могут быть представлены как системы массового обслуживания с задержками, не зависящими от нагрузки (рис.1).

На основе построенной модели была реализована система визуализации для ряда сервисов экологического мониторинга. Было проведено тестирование общей производительности системы, полученные результаты совпадают с модельными для более чем 80% случаев. Полученная точность является высокой по сравнению с

результатами исследований методами, применяемыми для оценки производительности Web-систем.

3. Примеры реализации

Рассмотренная в предыдущем разделе система визуализации геопространственных данных нашла применение в сервисах оценки биоразнообразия в Причерноморском регионе [6] и мониторинга водных ресурсов в Днепровском лимане [7].

При реализации сервисов использовано бесплатное программное обеспечение с открытым кодом UMN MapServer 4.10 (визуализация геопространственных данных, поддержка стандартов предоставления WMS и WCS), CartoWeb 3.3 (интерфейс пользователя, балансировка нагрузки, разграничение доступа, интернационализация), Apache 1.34+mod_ssl (web-сервер, разграничение доступа, поддержка защищенного канала передачи). В качестве базовой операционной системы при разработке сервисов, реализации серверов доступа и обработки данных использована Gentoo Linux. Реализованная система не требует дополнительных затрат на лицензирование или закупку специализированного ПО.

Выводы. В данной работе предложен подход к моделированию распределенных Web-ориентированных систем визуализации геопространственных данных. Построенная на основе предлагаемого подхода модель для системы визуализации ИКИ НАНУ-НКАУ при верификации на реализованной системе показывает точность более 80%, что является очень хорошим результатом по сравнению с существующими подходами для Web-сайтов. Полученная модель использована при построении системы визуализации для систем мониторинга экологических параметров окружающей среды на основе спутниковых данных.

Список литературы

1. Shelestov A., Kravchenko O., Ilin M. Geospatial data visualisation in Grid system on Ukrainian segment of GEOS/GMES// Proc. of the V-th International Conference "Information Research&Applications". — Varna (Bulgaria). — June 26-30, 2007. — Vol. 2. — P. 422-428
2. Кравченко О.М., Шелестов А.Ю. Застосування реалізацій стандартів ОГС для створення розподілених систем візуалізації та надання геопросторових даних, Проблеми програмування. – 2006. - №2-3. – С. 135-139. ISSN1727-4907
3. WMS 1.1.1 specification, <http://www.opengeospatial.org/standards/wms>.
4. WFS specification, <http://www.opengeospatial.org/standards/wfs>.
5. WCS specification, <http://www.opengeospatial.org/standards/wcs>.
6. Сервис оценки биоразнообразия в Южно-Причерноморском регионе, www.dos.ikd.kiev.ua.
7. Сервис мониторинга водных ресурсов Днепровского лимана, www.geo-ukraine.org.ua.

Получено 14.05.07