

Архитектура и сценарии распределенной обработки медицинских изображений веб-ориентированной PACS

Двойкин А. А., Привалов М. В., Воронова А. И.

Донецкий национальный технический университет, г. Донецк
кафедра автоматизированных систем управления
94dvoikin@gmail.com

Цель работы: повысить быстродействие веб-ориентированной системы архивирования, хранения, передачи и обработки медицинских изображений за счёт разработки архитектуры системы и сценариев распределённой обработки данных.

Введение.

При установлении диагноза и проведении лечения врачи всё больше полагаются на медицинские изображения, к которым относятся рентгенограмма, УЗИ, магнитно-резонансная томография, компьютерная томография, томография на позитивном излучении и т.д. Использование медицинских изображений непрерывно растёт по мере того, как во многих больницах (сначала в базовых больницах медицинских институтов) устанавливаются системы с высокой пропускной способностью для архивирования изображений. Медицинские изображения дают основной объём информации о пациенте, однако, самих по себе их недостаточно, поскольку требуется анализ и интерпретация в контексте истории болезни пациента (т.е. нужны метаданные, связанные с изображениями). Существует ряд факторов, которые значительно усложняют лечение пациента, основанное на медицинских изображениях: медицинские данные о пациентах собираются в разных лечебных учреждениях; врачи не всегда имеют доступ ко всем историям болезней всех своих пациентов; медицинские изображения представляются очень большими объёмами данных со сложной структурой. Объёмы обусловлены наличием 3D-изображений, показаний, полученных в последовательные моменты времени, многочисленных протоколов описания изображений, а сложность структур обусловлена клинически и эпидемиологически значимыми показателями, такими как возраст, питание, образ жизни и история болезни пациента, параметры получения изображений и анатомические и физиологические изменения. Часто изображений одной модальности оказывается недостаточно, поскольку на получение изображения влияет много

параметров, и дополнительная информация накапливается различными системами сбора физических данных (physical acquisition systems) [1].

Следовательно, информационные системы, которые предоставят возможность доступа к изображению с любого устройства и любого места при наличии Интернет-соединения, являются и будут оставаться актуальными ещё долгое время, ведь компьютерная обработка медицинских изображений постоянно развивается, появляются новые физические методы трёхмерного сканирования и визуализации внутренних органов человека [2].

Обзор существующих систем.

В данный момент существует достаточное количество систем архивации и передачи медицинских изображений (PACS – Picture Archiving and Communication Systems). Они решают следующую проблему: более широкое применение цифровой патологии в лаборатории означает, что размер изображений и метаданных, подлежащих хранению, продолжает увеличиваться, что приводит к увеличению спроса на места хранения. Наряду с увеличением требований к пространству для хранения данных скорость, с которой эти изображения необходимо анализировать и извлекать, также должна соответствовать растущим требованиям этого динамического диагностического поля [3]. Система должна быть масштабируемой и поддаваться регулярным обновлениям, учитывая быстрые изменения, которые разработчики делают с существующими решениями для анализа и хранения [4].

Для решения данной проблемы возможно использование сервис-ориентированных и распределённых Интернет-архитектур [5-9]. Однако, сервис-ориентированная архитектура на основе паттерна REST предлагает такие возможности: веб-ориентированность системы, не требующую больших затрат времени и громоздких задач по установке и настройке, простое и удобное добавление и модификацию новых алгоритмов разработчиками, доступ в

любое время и из любого места за счет веб-доступа [10].

Преимуществами таких систем является: наличие функционального веб-ориентированного клиентского приложения, возможность изменения и добавления новых алгоритмов в серверное приложение, возможность интеграции в системы анализа и управления изображениями и т.д.

Но основными и главными для данного исследования недостатками являются: отсутствие возможности распределения нагрузки между клиентскими терминальными устройствами и центром обработки данных (ЦОД) медицинского учреждения, а также добавления сервисов и расширения клиентского приложения. Так, например, открытие клиентского приложения на мобильном устройстве, которое не сможет произвести фрагментацию огромного изображения целиком, будет проблематичным, но, тем не менее,

обработка и отображение на экране какого-то фрагмента вполне возможны. Или если мы используем достаточно мощный для обработки ноутбук, но при этом кластеры ЦОД медицинского учреждения заняты, возможна ситуация, когда пользователь либо не получит изображение, либо будет ждать своей очереди.

Решением данной проблемы является разработка архитектуры, которая будет обладать высокой гибкостью и масштабируемостью, а также будет способна определять сценарий распределённой обработки изображений в зависимости от текущей ситуации.

Разработка архитектуры подсистемы хранения и обработки медицинских изображений

На рис. 1 изображена общая архитектура подобной подсистемы.

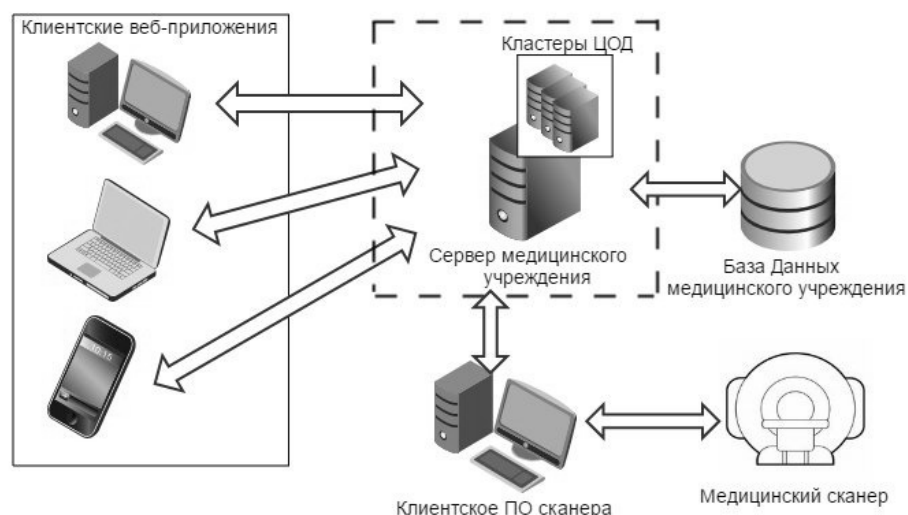


Рисунок 1 – Общая архитектура подсистемы хранения и обработки медицинских изображений

Из рис. 1 видно, что подсистема предоставит доступ к данным, которые хранятся в БД медицинского учреждения, с клиентских устройств благодаря взаимодействию клиентского веб-приложения с подсистемой, являющейся частью сервера медицинского учреждения. Данный сервер имеет прямой доступ к БД медицинского учреждения или удалённой БД, расположенной на кластере ЦОД. Более подробно архитектура подсистемы показана на рис. 2.

Взаимодействие между подсистемами основано на принципах сервис-ориентирования. Такие связи присутствуют между сервисами серверных подсистем и клиентским

приложением, аналогично построено взаимодействие и с DICOM-сервером.

Все сервисы подсистемы можно разделить на 5 основных групп:

- сервисы подготовки данных к передаче клиентскому приложению;
- сервисы обработки запросов клиентского приложения;
- сервисы динамического определения сценария обработки;
- API-сервисы;
- сервисы распределённого анализа изображения.

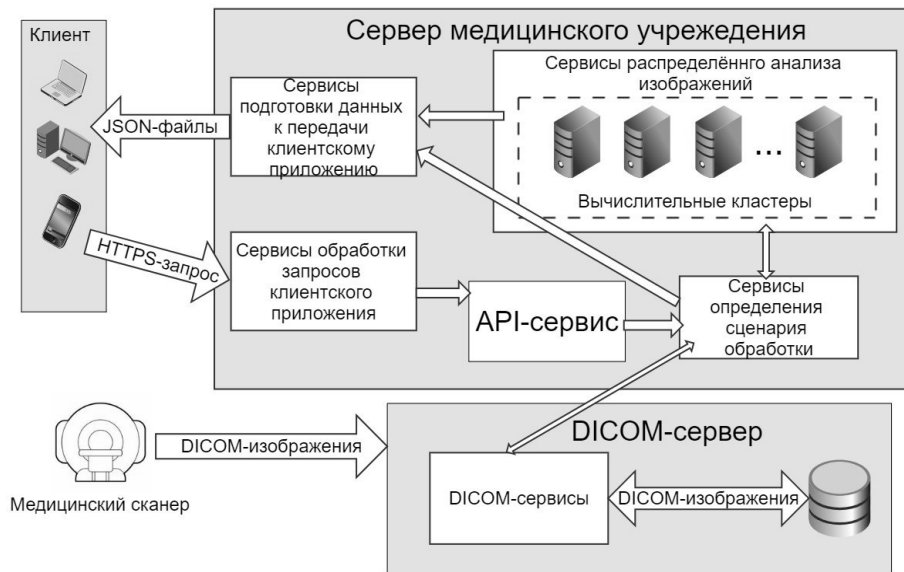


Рисунок 2 – Подробная архитектура подсистемы

Принцип работы подсистемы: с клиентского приложения по защищённому протоколу HTTPS поступает HTTP-запрос, который обрабатывается в сервисах обработки запросов. Сервисы обращаются к API-сервисам, которые подают запрос на определение сценария обработки соответственно клиентскому запросу, вследствие чего сервисы определения сценария получают от DICOM-сервера изображение и данные о нём, а также данные о загруженности ЦОД медицинского учреждения и вычислительной мощности клиентского устройства. После определения сценария задача обработки изображения выполняется в ЦОД, и клиентскому приложению приходит результат обработки, либо изображение в полном виде отправляется клиентскому приложению.

Определение сценария

Главным критерием при выборе сценария является время обработки t_o :

$$t_o = t_{mc} + t_{ck}, \quad (1)$$

где t_{ck} – время передачи данных с сервера на клиентское приложение;

t_{mc} – время обработки изображения.

Время обработки определяется по формуле:

$$t_{mc} = t_c + t_n + t_\phi, \quad (2)$$

где t_c – время считывания изображения;

t_n – время преобразования изображения определёнными методами;

t_ϕ – время формирования данных для отображения.

Время выполнения каждого этапа обработки зависит от следующих факторов:

- производительности (M) устройства, на котором обрабатывается изображение;
- размера (S) изображения;
- сложности (D) алгоритма каждого этапа обработки.

Следовательно, формула (2) имеет вид:

$$t_{mc} = \frac{M}{S} \cdot \left(\frac{1}{D_c} + \frac{1}{D_n} + \frac{1}{D_\phi} \right), \quad (3)$$

где D_c – сложность алгоритма считывания изображения;

D_n – сложность алгоритма преобразования изображения определёнными методами;

D_ϕ – сложность алгоритма формирования данных для отображения.

Главный параметр $t_{ок}$, который учитывается при принятии решения об обработке на клиентском оборудовании, определён формулой:

$$t_{ок} = t_{скк} + t_{мск}, \quad (4)$$

где $t_{скк}$ – время передачи необработанных данных с сервера клиенту;

$t_{мск}$ – время обработки на клиентском терминальном оборудовании.

Время обработки изображения на клиентском терминальном оборудовании ($t_{мск}$) зависит от:

- производительности (M_k) клиентского терминального оборудования;
- возможности (B) выполнения обработки ($B = 0$, если возможности нет, иначе $B = 1$);
- размера (S) изображения;
- сложности (D) алгоритма каждого этапа обработки.

Таким образом, время обработки изображения на клиентском терминальном

оборудовании может быть определено по формуле:

$$t_{мск} = \frac{B \cdot M_{к}}{S} \cdot \left(\frac{1}{D_c} + \frac{1}{D_n} + \frac{1}{D_{\phi}} \right) \quad (5)$$

Главный параметр $t_{оц}$, который определяет условие обработки изображения в ЦОД медицинского учреждения, определяется как:

$$t_{оц} = t_{скц} + t_{мсц}, \quad (6)$$

где $t_{скц}$ – время передачи результатов обработки с сервера клиенту;

$t_{мсц}$ – время обработки изображения в ЦОД медицинского учреждения.

В свою очередь, $t_{мсц}$ определено как:

$$t_{мсц} = \frac{M_{кл} \cdot K}{S} \cdot \left(\frac{1}{D_c} + \frac{1}{D_n} + \frac{1}{D_{\phi}} \right), \quad (7)$$

где K – количество кластеров, между которыми можно распараллелить обработку изображения;

$M_{кл}$ – производительность одного кластера.

Критерий, на основании которого принимается решение об обработке в ЦОД или на клиентском оборудовании, может быть сформулирован следующим образом:

– при $t_{ок} > t_{оц}$ обработку следует выполнять в ЦОД с учётом ограничений: загрузка кластеров ЦОД на момент запроса, возможность обработки изображения на клиентском терминальном оборудовании;

– при $t_{ок} < t_{оц}$ обработку следует выполнять на клиентском терминальном оборудовании.

Выводы

Развитие методов диагностики, основанных на получении и анализе медицинских изображений, а также темпы накопления этих изображений позволяют сделать вывод, что повышение быстродействия, гибкости и масштабируемости систем хранения и обработки медицинских изображений – актуальная задача. Анализ архитектур современных информационных систем показал, что масштабируемость и гибкость могут быть повышены за счёт применения принципов сервис-ориентирования. Показано, что оперативность работы и доступность системы возможно повысить за счёт автоматического определения сценария обработки на основании объёма задачи, характеристик клиентского оборудования и текущей доступности ЦОД. Сформулирован критерий выбора узла для обработки. Направлением дальнейших исследований является экспериментальное исследование применяемых методов обработки изображений, их реализаций и типов медицинских изображений с целью определения характеристик производительности

оборудования, сложности алгоритма, а также уровня загрузки сети и ЦОД, которые позволят эффективно выполнять выбор сценария обработки.

Список литературы

1. Назаренко Г. И., Гулиев Я. И., Ермаков Д. Е. Медицинские информационные системы: теория и практика / под ред. Г. И. Назаренко, Г. С. Осипова. – М.: ФИЗМАТЛИТ, 2005. – 320 с.
2. Гонсалес Р., Вудс Р. Цифровая обработка изображений / М.: Техносфера, 2005. – 1072 с.
3. Попов С. Б. Концепция распределенного хранения и параллельной обработки крупноформатных изображений / С. Б. Попов // Компьютерная оптика – 2007. – т. 31 – №4 – С. 77-85.
4. Таненбаум Э. Распределенные системы. Принципы и парадигмы / Э. Таненбаум, М. Ван Стеен. – СПб.: Питер, 2003. – 877 с.
5. Юдифь Гурвиц и др. Сервис-ориентированная архитектура для чайников. = Judith Hurwitz et al, Service Oriented Architecture for dummies, Wiley, 2007. – 387pp.
6. Томас Эрл, Стивен Беннетт. Управление СОА. Управление разделяемыми службами локально и в облаке. = Thomas Erl, Stephen Bennett, SOA governance. Governing shared services on-premise and in the cloud. // Prentice Hall, 2011. – 704pp.
7. Томас Эрл, Бенджамин Карлайл, СОА с REST. Принципы, шаблоны и ограничения для построения корпоративных решений на основе REST. = Thomas Erl, Benjamin Carlyle, SOA With REST. Principles, patterns and constraints for building enterprise solutions with REST. // Prentice Hall, 2012. – 624pp.
8. Евгений Алексеев, Обзор Windows Workflow Foundation на примере построения системы электронного документооборота [Электронный ресурс]. – <https://habrahabr.ru/company/luxoft/blog/181562/>
9. Амбили К.К., Настройка Windows Workflow Foundation 4.5: разработка унифицированных и масштабируемых приложений. = Ambily K. K., Customizing Windows Workflow Foundation 4.5: Unified and Scalable Application Development, Amazon Digital Services LLC, 2015. – 126 pp.
10. Симоненко А. А., Привалов М. В. Разработка веб-ориентированной подсистемы доступа к медицинским изображениям // Информатика, управляющие системы, математическое и компьютерное моделирование в рамках II форума «Инновационные перспективы Донбасса» (ИУСМКМ-2016): VII Международная научно-техническая

конференция, 26 мая 2016, г. Донецк – Донецк:
ДонНТУ, 2016. – 624 с. – С. 274-279.

Двойкин А. А., Привалов М. В., Воронова А. И. Архитектура и сценарии распределённой обработки медицинских изображений веб-ориентированной PACS. Выполнен анализ и описаны ключевые особенности систем обработки медицинских изображений, описана необходимость распределения нагрузки между клиентским приложением и сервером (центром обработки данных). Показаны основные сценарии обработки данных в такой системе и предложены критерии для их выбора.

Ключевые слова: системы обработки медицинских изображений, распределённая обработка, распределение нагрузки.

A.A. Dvoykin, M.V. Privalov, A.I.Voronova Architecture and distributed medical image processing scenarios of the Web-oriented PACS. Analysis of the medical image processing systems was performed and their key features were described. It is shown that workload balancing between client application and server (data center) is required. Main data processing scenarios and criteria for their selection were proposed.

Keywords: medical image processing systems, distributed processing, workload balancing.

*Статья поступила в редакцию 20.09.2017
Рекомендована к публикации д-ром техн. наук В.Н. Павлышом*