

PROCESSAMENTO DISTRIBUÍDO DE IMAGENS MÉDICAS BASEADAS EM CLUSTERING

TAMAE, Rodrigo Yoshio
Faculdade de Ciências Jurídicas e Gerenciais de Garça – FAEG/Garça
MUCHERONI, Marcos Luiz
Faculdade de Informática, Fundação Eurípides de Marília, Marília – São Paulo, Brasil

RESUMO

Existe uma enorme atenção voltada para o desenvolvimento tecnológico baseado em imagens médicas como, por exemplo, aquisição, gerenciamento e distribuição. Tal evolução, no entanto, tem gerado um número cada vez maior de informações textuais e, principalmente, de imagens médicas. Conseqüentemente é preciso criar novos meios para gerenciamento destas informações, bem como, minimizar os tempos de resposta aos clientes em relação a estes dados. Este trabalho tem como objetivo apresentar uma arquitetura que propõe otimizar o impacto desta interação homem-máquina através da tecnologia de clustering.

Palavras-chave: Imagens médicas e clustering.

ABSTRACT

There is an enormous attention directed to technology advances based on medical image, for example, acquisition, management and distribution. This evolution had been generating a growth number of textual information and medical image. Thus, it's necessary create new manner to manage these information and minimize the reply time to the clients. This work shows the architecture that purpose optimize impact the interaction between man and machine by clustering technology.

Keywords: Medical image and clustering.

1 INTRODUÇÃO

Em conseqüência do rápido avanço tecnológico, novos equipamentos de aquisição e processamento de imagens médicas estão sendo desenvolvidos, tornando-se importantes ferramentas na prática da medicina moderna, pois, oferecem não apenas uma forma de visualização não-invasiva a órgãos, tecidos, e demais estruturas do corpo de um paciente, como também um meio para monitorar os efeitos de tratamentos e planejar cirurgias. No entanto, a utilização de equipamentos digitais de diagnóstico geram uma grande quantidade de informações em formato digital, o que torna crítica a realização das tarefas de armazenamento, manipulação, gerenciamento e, principalmente, de distribuição destes dados criando o desafio de otimizar as respostas solicitadas pelos usuários destes sistemas.

Segundo relatório da Sociedade para Programas de Computador em Radiologia (*SCAR - Society for Computer Applications in Radiology*), é necessário uma mudança de paradigma no processo de interpretação radiológica, para ser possível lidar com a crescente quantidade de informação. O relatório aponta que as tarefas de visualização e processamento de imagens, combinadas com sistemas de assistência computacional, como sistemas de apoio à decisão e acesso a bibliotecas de referências, estão entre os seis principais desafios para o futuro da pesquisa em imagens médicas (ANDRIOLI, 2003).

Tais informações médicas são geradas, principalmente, em ambientes PACS (*Picture Archiving and Communication System*) que são sistemas de computadores utilizados para a aquisição, armazenamento, processamento, exibição e distribuição de imagens médicas em formato digital (HUANG, 1999). Esses ambientes disponibilizam os diagnósticos por imagem e permitem o acesso a imagens médicas via rede (SIEGEL, 1999).

Partindo do modelo PACS, esta em desenvolvimento o SISPRODIMEX (Sistema de Processamento Distribuído de Imagens Médicas usando XML) que é o nome atribuído a uma plataforma cuja proposta é de disponibilizar acesso a um banco de dados de informações digitais oriundas de sistemas PACS, geradas no padrão DICOM 3.0, através de tecnologia de ponta e de relativo baixo custo, como as linguagens Java e XML.

O desafio reside na necessidade de obter maior poder de processamento e agilizar a resposta ao usuário no acesso a tais informações em formato digital com o auxílio de um pequeno cluster experimental de computadores.

2 METODOLOGIA

O SISPRODIMEX foi concebido para ser uma aplicação Web e no processo de implementação estão sendo utilizadas as tecnologias baseadas na plataforma Java e na linguagem XML. No atual estágio de desenvolvimento, optamos por construir serviços mais sofisticados usando a tecnologia dos Servlets Java e JSP.

O SISPRODIMEX não tem por objetivo definir qual Sistema Gerenciador de Banco de Dados (SGBD) será utilizado e os testes iniciais com banco de dados estão sendo feitos com o Firebird 1.5 SQL Server.

A segurança de acesso aos dados dos pacientes também foi uma preocupação na concepção do SISPRODIMEX e para isso rotinas de tráfego de dados criptografados baseados no padrão PKCS (*Public Key Cryptography Standards*) da RSA (Rivest, Shamir e Adleman) estão sendo estudados e testados. Para gerenciamento de conteúdo utilizamos o Apache Tomcat 5.0.

No que se refere ao transporte de imagens, parte deste trabalho, tem sido financiado pela FAPESP¹. Os testes iniciais na camada de transporte consistiram de implementações utilizando tanto o protocolo orientado à conexão (TCP) quanto ao protocolo não-orientado a conexão (UDP).

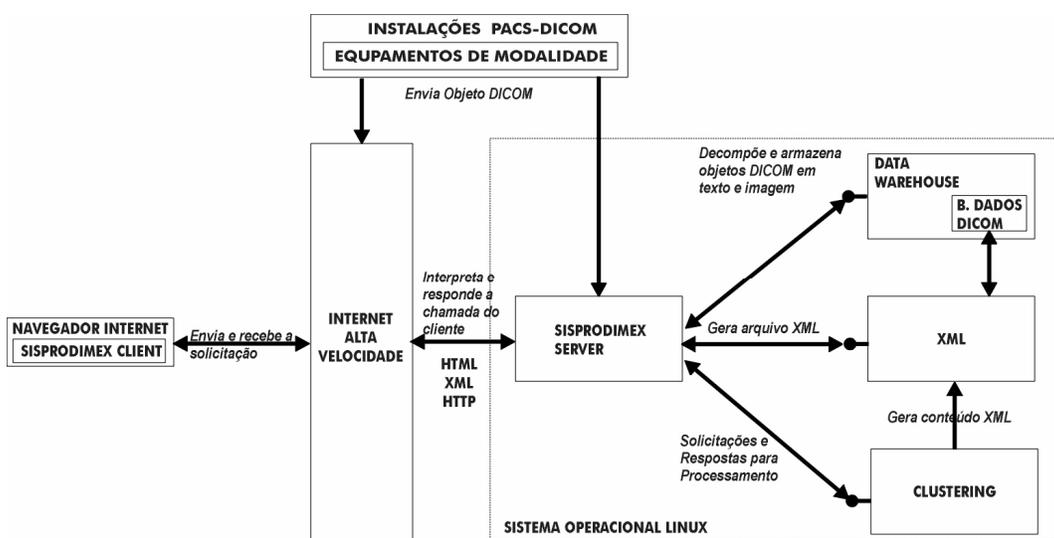


Figura 1 – A arquitetura do SISPRODIMEX com a conexão com o cluster

Como pode ser observado na figura 1, o ambiente PACS envia o objeto DICOM através da Web ou através das conexões de rede ao servidor SISPRODIMEX que decompõe o objeto DICOM, separando as informações textuais das imagens, gravando-as em um banco de dados. A partir de uma solicitação feita através do SISPRODIMEX-Cliente, por meio de um navegador Internet, o SISPRODIMEX-Servidor recebe a solicitação, executa o processamento no cluster e gera resposta ao cliente.

As informações são convertidas em formato XML, com o claro intuito de dar suporte a metadados. Adicionalmente, recursos de Web Semântica e Ontologias (MUCHERONI, 2003), também serão implementados, pois, passam a ter um papel fundamental no que diz respeito a processos de pesquisa e recuperação de dados, sendo este o ponto de partida para a implementação do processo de gerenciamento contextual previsto no SISPRODIMEX. Para isso, serão criadas ontologias e um mecanismo de busca baseado na linguagem XML.

É enfatizado no modelo o uso de Internet de alta velocidade devido ao grande volume de dados previsto a cada solicitação correspondendo. A arquitetura paralela baseada em cluster

visa gerar o poder de processamento necessário para minimizar o impacto na resposta ao usuário final.

3 EXPERIMENTOS COM O CLUSTER

Para incrementar velocidades possíveis em ambientes baseados em clustering, foi retomada a conhecida Lei de Amdahl, onde, supondo que um programa levará um tempo t_s para realizar uma determinada tarefa, de desempenho de máquinas com p processos possíveis de serem paralelizados em r partições paralelizáveis, então, se tem:

$$S(p) = \frac{t_s}{(1-r)t_s + t_s / p} \quad (1)$$

Simplificando e retirando t_s :

$$S(p) = \frac{1}{(1-r) + r / p} \quad (2)$$

Se S é diferenciado com relação a p , tem:

$$\frac{dS}{dp} = \frac{r}{[(1-r)p+r]^2} \geq 0 \quad (3)$$

Incrementando a função de p , com $p \rightarrow \infty$, obtém-se que $S(p)$ é limitado a $1/(1-r)$, ou seja:

$$S = \frac{1}{(1-r)} \quad (4)$$

Em um experimento de extrapolação no tamanho de dados, particularmente importantes em sistemas PACS, assim feito para avaliar o sistema SISPRODIMEX.

O número total de mudanças para certo tamanho igual a R , para um número de elementos processadores igual a NPE , pode ser descrito como:

$$R * N^2 / (N-1) * NPE / (N-1) \quad (5)$$

Ou seja, é possível, no SISPRODIMEX, obter um excelente desempenho, conforme o número de máquinas e o tamanho das imagens que forem tratadas. Assim, esta poderia ser tratada por partições diferentes do cluster.

Para efeito de testes, foi utilizado um arquivo pequeno, de tamanho de 256 kbytes, e executando análises para N processos simulados sob o sistema operacional Linux, em tamanho máximos, de 2 em 2 até 20 (como um número prático em N de processos) para simular o ambiente paralelo, dado a seguir.

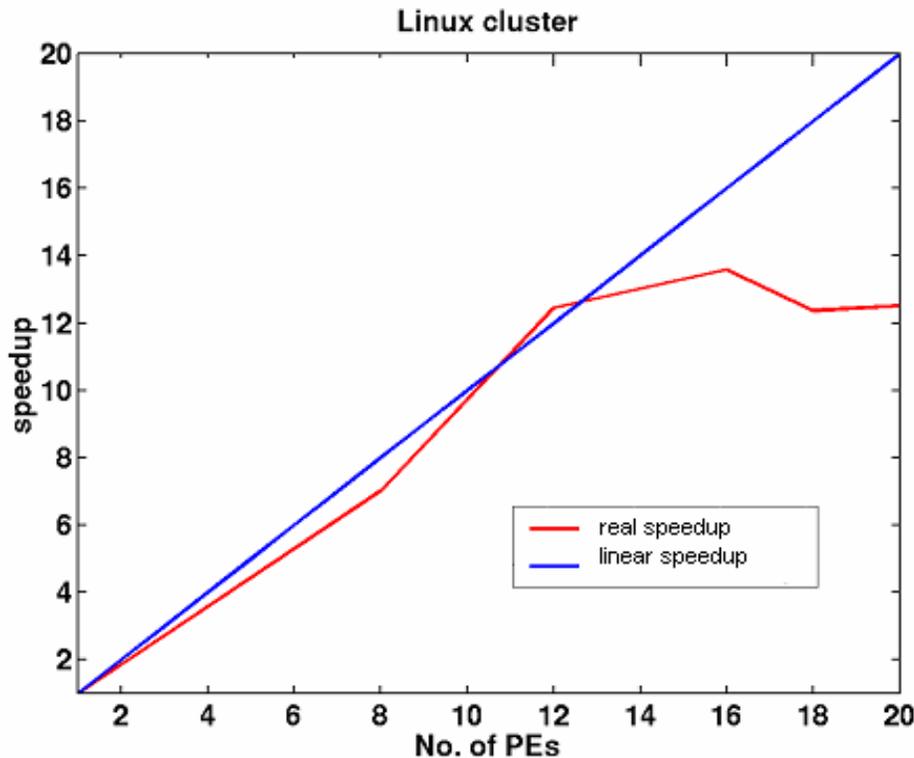


Figura 2 – Resultados práticos com partições uniformes (NPE=2,4,6,8,12,14,16,18,20), onde: (a) caso ideal (b) simulação usando N PÉ s.

4 DISCUSSÃO E CONCLUSÕES

A implementação completa da plataforma SISPRODIMEX possibilitará a pequenos centros hospitalares contar com um banco de dados centralizado de informações textuais e de imagens médicas com independência de localização geográfica através de recursos da Internet e World Wide Web.

Além de gerar o poder de processamento e minimizar o impacto nas respostas solicitadas pelos clientes, o ambiente de processamento paralelo baseado em clusters evita o uso isolado de caríssimos sistemas de Workstations. Além disso, o ambiente tem capacidade de adaptar-se aos diversos formatos de imagens e poderá oferecer um desempenho otimizado conforme o tamanho ou conjunto de imagens a serem tratados.

O SISPRODIMEX poderá disponibilizar recursos em rede a um custo relativamente mais baixo que sistemas semelhantes na área da saúde.

REFERÊNCIAS BIBLIOGRÁFICAS

ANDRIOLI, K.P. **Addressing the Coming Radiology Crisis: The Society for Computer Applications in Radiology Transforming the Radiological Interpretation Process (TRIP™) Initiative**, acessível em: <<http://www.scarnet.org/pdf/TRIPwhitepaper1103.pdf>>. Acesso em: Novembro 2003.

HUANG, H.K. **PACS Basic Principles and Application**, Wiley-Liss, New York, 1999.

SIEGEL, E.L. **Current State of Art and Future Trends, in Filmless Radiology**. Siegel, E.L., Kolodner, R.M. New York City, NY. Springer Verlag. p. 3-20, 1999.

MUCHERONI, M.L.; TAMAE, R.Y.; BRACCIALLI, T. **Advances in Intelligent Systems and Robotics: Auto Organizing Agents in Ontologies with Intentions**, IOS Press, Amsterdam, Netherlands V.101, p.171-179, 2003.