

修士論文

階層的なメタデータサーバを有する
広域分散ファイルシステムを利用した
医用画像保存通信システム

Distributed PACS using Distributed File System

with Hierarchical Meta data Servers

同志社大学大学院 工学研究科 情報工学専攻

博士前期課程 2011年度 735番

南谷 祥之

指導教授 三木 光範教授

2013年1月24日

Abstract

In recent years, storing and sharing medical information is becoming increasingly important. Among medical information, DICOM is the standard and network protocol for medical images. In DICOM standard, metadata such as patient information and examination information is stored with image data. Medical institutions have systems which control DICOM file and this system is called Picture archiving and communication systems (PACS).

PACS is now designed to use medical images stored at each hospital or facility. However, in the near future, several PACSs which exist in different hospitals or institutions will be worked together. There several ways to make collaboration among several PACSs. For example, using cloud platform is one of the solutions and there is an example of constructing the system using Microsoft Windows Azure. When several PACS are trying to be worked together, most of the cases, all the files are stored into one site. However, this strategy does not have scalability to the number of files and hospitals. Not only size of a medical image is large, but also huge number of images is added to the system every day. The amount of medical image information is increasing tremendously with the future development of medical equipment and systems. Based on current trends, it is estimated that over one billion diagnostic imaging procedures will be performed in the USA in 2014, which will generate about 100 petabytes of data.

In the proposed system, there is a meta data server, which manages the information of the files, and DICOM files are stored in locally. This mechanism is very suitable for DICOM files, because DICOM itself consists of metadata and images itself. Using this mechanism, PACS whose medical images are managed in several hospitals and whose information can be shared from outside of hospitals can be constructed. At the same time, when users try to find the certain image, user access to the information which is restored in the metadata. Like this way, the retrieving files can be performed quickly. Moreover, since the proposed system is using distributed file system, the system has the scalability of accessing file speed and the capacities.

However, when the file information is managing on a single server, system has the low fault tolerance. To make it higher, metadata servers should be constructed hierarchically. Thus, in this paper, we propose a system that each medical institution builds a distributed file system individually and coordinates them using hierarchal metadata servers. To implement the proposed system, Gfarm is utilized as a network shared file system. Gfarm have mechanism to coordinate multiple distributed file system and store the metadata of the file as XML data. Gfarm can retrieve the file by using XML in the multiple distributed file system. Therefore, the proposed system can be constructed easily with Gfarm. Using the implemented system, the proposed system is described and the evaluation experiments are performed.

目次

1	序論	1
2	医用画像	2
2.1	DICOM：医用画像規格	2
2.1.1	画像規格	2
2.1.2	通信規格	2
2.2	PACS：医用画像保存通信システム	3
2.3	遠隔医療ネットワーク構築の先行研究	3
3	提案システム	5
3.1	概要	5
3.2	システム要件	5
4	システムの実装	7
4.1	広域分散ファイルシステム Gfarm	7
4.1.1	システム構成	7
4.1.2	XML 拡張属性	8
4.1.3	階層的メタデータサーバ	8
4.2	DCMTK：DICOM Tool Kit	8
4.3	システム構築	8
4.3.1	保存登録処理	9
4.3.2	検索処理	9
4.3.3	取得処理	9
4.3.4	外部の医療機関に対する処理	9
4.4	GUI 実装	9
5	評価実験	10
5.1	実験システム	10
5.2	基盤システムにおける速度評価実験	10
5.2.1	実験概要	10
5.2.2	実験結果	11
5.2.3	議論	11
5.3	DICOM 通信における速度評価実験	12
5.3.1	実験概要	12
5.3.2	実験結果	12
5.3.3	議論	12
6	結論	14

1 序論

近年、医療現場において医療情報の蓄積と共有が重要視されてきている。医用画像に関しては、様々な機器で利用可能な専用規格が策定されており、現在医用画像規格は唯一 DICOM (Digital Imaging and Communications in Medicine) のみが利用されている。医用画像規格 DICOM では、検査情報や患者情報などの多数のメタデータが撮影された画像データと共に同一ファイル内に保存される¹⁾。このような医用画像を管理するため、医用画像保存通信システム (PACS : Picture Archiving and Communication Systems) と呼ばれるシステムが医療機関に広く普及している。現在では、病院内でのみ扱われる PACS については導入が進んでおり、電子患者情報と合わせた管理が行われている²⁾³⁾。一方、近年、遠隔医療への対応の必要性から外部の病院との連携を可能にする PACS の連携システムの構築が求められている。これを実現する手法として、Microsoft Windows Azure を用いたクラウド上で医用画像を扱うような PACS が構築されている⁴⁾。また、国内においては、医療機関が連携して患者情報を共有するシステムとして、「K-MIX : かがわ遠隔医療ネットワーク」が稼動している⁵⁾。しかし、これらのシステムはデータセンターを構築しており、既に各病院の PACS 上で管理されている医用画像をデータセンターのような場所に一元的に集約管理する手法では、連携する病院をさらに広域に拡大させた時、限界があると考えられる。医用画像は画像サイズが非常に大きいだけでなく、毎日膨大な枚数が追加されており、今後も医療機器やシステムの発展にとともに、その情報量は爆発的に増加すると考えられるためである。現在、アメリカにおいて 2014 年の一年間に 10 億以上の画像診断が行われ、それに伴い約 100 ペタバイトのデータが生成されると考えられている⁶⁾。そのため、医用画像の共有に関しては各病院が管理を行い、その上で外部の病院でも利用出来るシステムの構築が必要である。それに伴い、ローカルで管理している医用画像をグローバルに共有するためのシステム開発が進んでいる⁷⁾⁸⁾。

通常 DICOM を利用する場合、検索が行われる。その際、画像全体にアクセスを行うことは非効率であり、メタデータのみを取り出し、検索をメタデータサーバ、画像保存を画像ストレージが行う手法が考えられる。この手法により、メタデータのみを対象とした検索ができ、メタデータの高速度抽出結果を得ることが可能である。またメタデータ、画像データのみの柔軟なアクセスについても提供可能である。

そこで、本研究では各医療機関が個別に分散ファイルシステムを構築した上で、メタデータサーバを階層的に構築するシステムを提案する。大規模かつ高速な環境を提供する分散ファイルシステムとして広域分散システムファイル Gfarm が挙げられる⁹⁾¹⁰⁾。Gfarm は複数の分散ファイルシステムを連携させる機能および、XML をファイルのメタデータとして格納するメカニズムを有しており、複数の分散ファイルシステム内の XML からファイル検索を行う事が可能である。本研究では、この Gfarm を応用し DICOM 規格の医用画像を格納するファイルシステムを提案する。提案システムでは、広域に点在している病院の PACS を Gfarm を用いて連携することで、データセンターのような一元管理ではなく、分散管理による医用画像の共有利用を目指す。

本論文では、第 2 章では医用画像と医用画像保存通信システムについて、第 3 章では提案システムについて、第 4 章では構築したシステムの実装方法について、第 5 章では実験システムによる評価実験について述べる。

2 医用画像

2.1 DICOM：医用画像規格

現在、医療機関でMRI (Magnetic Resonance Imaging) やCT (Computed Tomography)などで撮影された画像ファイルは全てDICOM (Digital Imaging and Communications in Medicine) という形式で保存されている。DICOMは、医用画像のデータ形式と通信仕様に関する世界標準規格である。米国放射線学会 (ACR：American College of Radiology) と北米電子機器工業会 (NEMA：National Electrical Manufacturers Association) によって制定された。DICOM規格の前身はACR-NEMA規格というものであり、DICOMは1992年に正式に制定された。DICOMの規格書は、現在は16のパートから成り総ページ数は1000枚を超える膨大な規格書である。DICOM規格は追加、拡張、修正が継続的に行われており、DICOMを医療情報システムの中でどのように標準的に使用するかのガイドラインが作成され、その標準化が進んでいる¹¹⁾。DICOMはCTやMRIの撮影画像だけでなく、内視鏡や超音波などで作成された動画データや音声データなど、医療診療に関わるデータを包括することができる。DICOM規格の目的は、病院内外で異なった製造業者（マルチベンダー）の異なった種類（マルチモダリティ）のデジタル画像機器を、ネットワークあるいは画像保存媒体で、相互に接続して患者の画像検査情報のやり取りや画像データの伝送を可能とすることである。DICOM規格では、画像規格と通信規格を定義している¹²⁾。

2.1.1 画像規格

DICOM画像規格は画像を主とする様々なデータを内包できるコンテナフォーマットであり、内包物のフォーマットやデータ長を記載するタグ情報を保持している。医用画像規格としては、ISO TC-215では国際標準化が議論されている現在唯一の規格である。医用画像フォーマットとしてのDICOMでは、画像そのものに加えて、データサイズや画像サイズなどの画像情報、患者の氏名や住所などの個人情報、撮影機器や造影剤などの検査情報がメタデータとして付加される。メタデータは3000項目以上が規定されており、患者の検索や診断に大きく寄与している¹⁾。Fig. 1にDICOMの概要を示す。

2.1.2 通信規格

通信規格としてのDICOMでは、MRIやDICOMビューア等のDICOM準拠製品間の通信プロトコルを規定する。現在のDICOMでは、TCP/IP上にDICOM通信コネクションを確立している。これにより、ビューアやMRIなどのDICOM準拠製品間のファイル送受信や問い合わせを行う¹³⁾。

ネットワークにおいては、様々な機器が接続されているため、ネットワークに接続されている機器の機能範囲を調べ、要求するサービスが実現可能かどうかを事前に確認する必要がある。ネットワークにおける機器間の事前確認を折衝と呼ぶが、DICOM通信においても、折衝としてサービスの種類（例：Storage Service Class）と情報内容（画像種別）を事前に伝えている。DICOMでは、サービスの種類をサービスクラス、伝えられる情報を情報オブジェクト定義と呼ぶ。また、この二つを組合わせて「サービスオブジェクト対クラス=SOPクラス」と呼び、実際にはこのSOPクラスを用いて送信する機器側の回答を得ている。Fig. 2にDICOM通信における折衝の流れ図を示す。

DICOMでは画像規格、通信規格が定義されている。医療機器開発においては、このDICOM規格の下開発が進められる。医療機関内の画像管理においてはDICOM規格に準拠したシステムとして、医用画像保存通信システムが稼働している。

2.2 PACS：医用画像保存通信システム

現在、DICOMを管理するシステムとして、医用画像保存通信システム（PACS：Picture Archiving and Communication Systems）と呼ばれるシステムが医療機関で稼働している。PACSは主に医療現場の画像診断に使用されており、医療機関で発生する全ての検査画像をデジタルデータで保存し、配信から診断までの一連のワークフローを総合的に管理するシステムである。PACSは医用画像に関するシステムであるが、医療機関では他に、病院情報システム（HIS：Hospital Information System）および放射線部門システム（RIS：Radiology Information System）が動作している。

- HIS

HIS (Hospital Information System) とは、病院情報システムの略称である。主な業務としては電子カルテシステムが挙げられるが、その他、自動受付、入退院管理、医事会計および薬局管理などの広範囲なシステムによって、病院管理が効率的に行われる。

- RIS

RIS (Radiology Information System) とは、放射線部門情報システムの略称である。主に放射線機器による検査と治療の予約から検査結果までの管理を行うシステムのことであり、患者情報や予約情報などの内容を HIS から取得する。

医用画像の撮影においては、HIS、PACS および RIS の連携が必要不可欠である¹⁴⁾。Fig. 3 のように、HIS は患者情報の管理、RIS は検査予約の管理、PACS は撮影画像の管理を行っている。これらのシステムの連携によって、画像診断の環境が整う。Fig. 3 に各システムの連携図を示す。

2.3 遠隔医療ネットワーク構築の先行研究

現在、病院内でのみ扱われる PACS については導入が進んでおり、電子患者情報と合わせた管理が行われている²⁾³⁾。一方、近年外部の病院との連携を可能にする PACS の連携システムの構築が求められている。撮影された画像を診断する際には、画像診断の専門知識を持つ読影医が必要である。日本においては、読影医が不足する事態が起きており¹⁵⁾、撮影された画像を遠隔地に送信し、診断結果を得る遠隔医療が必要である。また、日本では米国に比べ、セカンドオピニオンの受診率が非常に低い。癌などの画像診断においては、初期状態での早期発見が重要であるが、単独の医師の診断のみでは見落としによる誤診もあり、セカンドオピニオンを受けることが望ましいとされる¹⁶⁾。そのために、PACS の連携システムを実現する手法として、Microsoft Windows Azure を用いたクラウド上で医用画像を扱うような PACS が構築されている⁴⁾。また、国内においては、医療機関が連携して患者情報を共有するシステムとして、「K-MIX：かがわ遠隔医療ネットワーク」が稼働している⁵⁾¹⁷⁾。

しかし、既に各病院の PACS 上で管理されている医用画像をデータセンターのような場所に一元的に集約管理する手法では、連携する病院をさらに広域に拡大させた時限界があると考えられる。医用

画像は画像サイズが非常に大きいだけでなく、毎日膨大な枚数が追加されており、今後も医療機器やシステムの発展にともない、その情報量は爆発的に増加すると考えられるためである。現在、アメリカにおいて2014年の一年間に10億以上の画像診断が行われ、それに伴い約100ペタバイトのデータが生成されると考えられている⁶⁾。そのため、医用画像の共有に関しては各病院が管理を行い、その上で外部の病院でも利用出来るシステムの構築が必要である。それに伴い、ローカルで管理している医用画像をグローバルに共有するためのシステム開発が進んでいる⁷⁾⁸⁾。

3 提案システム

3.1 概要

本研究では、データセンターのように一元的に医用画像の集約管理を行わず、各医療機関が分散的に PACS を管理し、それら PACS を統合するようなシステムの構築を目的とする。Fig. 4 に提案システムの概要図を示す。

第1章で述べたように、現在医用画像は広域に点在している各病院の PACS によって管理されている。膨大なデータ量を保持する医用画像は分散環境にて保存される必要がある。非常に大きなデータサイズの DICOM が広域環境において共有されるためには、各病院が PACS を管理した状態で、それら PACS を統合するようなシステムの構築が必要である。

そこで、本研究では、各医療機関が分散ファイルシステムを構築し医用画像の管理を行った上で、複数の分散ファイルシステムを連携する PACS を提案する。DICOM を対象とする場合に効果が上がる画像格納方法として、DICOM からメタデータを取り出し、検索をメタデータサーバ、画像保存を画像ストレージが行う手法が考えられる。

通常、DICOM を利用する際、DICOM のメタデータを用いて検索が行われる。メタデータを検索するために DICOM 全体にアクセスを行うことは非常に効率が悪い。DICOM からメタデータ情報の切り離しが必要である。そこで、各病院が DICOM のメタデータを格納するサーバと実画像データである DICOM を格納する複数サーバの 2 層のサーバを構築する。さらに、広域環境で医療機関が連携を行うため、医療機関外に新たなサーバを設置し、このサーバを用いて各医療機関で管理している分散ファイルシステムを連携するサーバを構築する。連携サーバでは画像に関する検索機能を持たず、医療機関の場所情報のみを保持する。これにより外部の医療機関で管理されている DICOM を検索する際は、分散処理の観点から連携サーバ内では行わず各病院のメタデータサーバ内にて検索が行われる。連携サーバを構築することにより、新規に病院をシステムに追加する際には、医療機関にシステムを導入後、連携サーバに場所情報を追加するのみで本システムに追加する事が可能となる。また、検索の後、DICOM を送受信する際にはメタデータサーバ、連携サーバを通すことなく画像格納サーバより直接通信により送受信が行われる必要がある。分散ファイルシステムを連携可能であることから、医療機関内においてもデータ量がキャパシティを超える場合には、複数の分散ファイルシステムにスケールアウトすることが可能である。

3.2 システム要件

3.1 節を実現するためには、以下の要件を満たす必要がある。

- (1) 広域の複数拠点にデータを分散保存し、それらを共有利用できる基盤

医用画像は各医療機関に保持される。一方で、他病院で管理されている医用画像に関してもアクセス可能な基盤が必要である。

- (2) 保存されたデータから検索項目を抽出し、ファイルと関連づけて検索する機能

取り扱うデータである DICOM はメタデータを保持している。DICOM メタデータを検索項

目として、ファイルを検索可能な機能が必要である。その際、他病院の DICOM に関しても検索可能でなければならない。

(3) 検索を各拠点にて行う機能

医用画像検索が行われる最も頻度の高い利用場面として、自病院内での検索が予想される。そのために、全病院の画像データに関して検索を行う事は非効率である。また、分散処理の観点からも検索は各医療機関で行われることが望ましい。そのため、各医療機関のメタデータサーバ内にて検索が行われ、必要な際には外部からの検索を受け付けることが可能な機能が必要である。

4 システムの実装

本章ではシステムの実装について述べる。第3章で提案したシステムを実現するため、広域分散ファイルシステム Gfarm に着目した。Gfarm は、システム要件である以下の3点を満たす。

- (1) 広域の拠点にストレージノードを配置しメタデータサーバを用いてストレージノードをまとめることが可能な分散ファイルシステムである。
- (2) 他の分散ファイルシステムにはない独自機能として、ファイルに関連づけて任意の項目から検索可能な機構を有している。
- (3) 複数の Gfarm メタデータサーバを組み合わせ階層的に分散ファイルシステムを構築することが可能となっている。

以上より、Gfarm を用いてシステムの構築を行った。また、既存の DICOM 製品との通信を図るため、DCMTK (DICOM Tool Kit) を利用した。次節より、Gfarm と DCMTK について述べる。

4.1 広域分散ファイルシステム Gfarm

広域分散ファイルシステム Gfarm は、ネットワークで接続された PC のローカルストレージを束ねて構成されるオープンソースのファイルシステムである。LAN 上の PC や PC クラスタ、広域のクラスタ群など、様々な規模での利用が可能であり、NFS (Network File System)、AFS (Andrew File System) に代わる大規模かつ高性能なファイル共有システムとして研究開発が続けられている。Gfarm 広域分散ファイルシステムは、オープンソースで研究開発が進められており、ネットワークで接続された PC のローカルストレージを束ねて構成するファイルシステムである。LAN に接続された複数の PC、単独の PC クラスタ、広域の複数の PC クラスタなどで構成され、高性能な共有ファイルシステムとして利用可能である¹⁰⁾¹⁸⁾。

4.1.1 システム構成

Gfarm ファイルシステムは、3種に分類できるノード群で構成される。ファイルデータ本体が格納される I/O サーバ、各ファイルの保存位置が格納されるメタデータサーバ、ファイルにアクセスするクライアントである。メタデータサーバは、仮想的なディレクトリ階層、実際のファイルの所在などのファイルシステムデータをメタデータとして管理することからメタデータサーバと呼ばれる。Gfarm に格納されたファイルにアクセスするには、まずメタデータサーバに位置を問い合わせ、得られた I/O サーバ中にアクセスする¹⁹⁾。

Gfarm では、ファイルを保存する際、実際のファイル保存先は I/O サーバが担うのに対し、メタデータサーバがファイルの保存先をメタデータとして保持する。これにより、クライアントがファイルを利用する際はメタデータサーバにアクセスすることで、実際にはどの I/O サーバにファイルが保存されているかを意識することなく利用可能である。実際のファイルの送受信に関してはクライアントと I/O サーバが直接行う。Fig. 5 に Gfarm アーキテクチャを示す。

4.1.2 XML 拡張属性

Gfarm メタデータサーバは、ファイルの保存先の管理だけでなく“XML 拡張属性”と呼ばれる XML をデータとして格納する機構を持つ。これは、保存ファイルごとに関連付けられる XML ファイルを指定し、XPath¹によるファイルの検索を可能とする機能である。これにより、保存されたファイルに対して直接アクセスすることなく検索が可能となる。

4.1.3 階層的メタデータサーバ

Gfarm はメタデータサーバの階層的な構築を可能にする機能を有している。メタデータサーバの機能を分散処理することにより、負荷分散を行うだけでなく、各拠点毎に分散ファイルシステムを構築し、そのメタデータ情報を含めた管理を行う事が可能である。

4.2 DCMTK : DICOM Tool Kit

既存の DICOM 製品との接続を可能にするためには、DICOM 通信規格への準拠が必須である。今回、DICOM 通信機能を実装するにあたり、C++実装の DICOM ライブラリである DCMTK (DICOM Tool Kit)²⁰ を利用した。DCMTK はドイツのオルデンプルグ大学に開設された OFFIS (Oldenburg Research and Development Institute for Information Technology) によって BSD ライセンスのもとオープンソースでの開発提供がされている。本稿では、DICOM 通信によりファイル受信機能を実現するため DCMTK のサーバ機能を利用しシステムを構築した。

4.3 システム構築

本提案システムでは、Fig. 4 に示された画像サーバを Gfarm I/O サーバ、メタデータサーバを Gfarm メタデータサーバとしてシステム構築を行った。なお、連携サーバの実態は Gfarm メタデータサーバである。

各医療機関は撮影された DICOM 画像を I/O サーバに保持する。一方で、保存する DICOM のメタデータ部分から XML を作成し、メタデータサーバに登録を行う。各医療機関ではこのような 2 層のサーバを構築することにより、ファイルサイズの大きな DICOM の画像情報にアクセスすることなくメタデータ XML のみを利用することが可能である。これにより、ファイルシステム自体がファイルの検索を行うことが可能である。さらに、各医療機関で管理されている DICOM は外部の複数の医療機関で利用される必要がある。外部の医療機関で管理されている DICOM を検索する際には、まず外部の連携サーバにアクセスされる。この際、この連携サーバ自体は検索を行わず、検索は各医療機関のメタデータサーバで行われる。Fig. 4 に示す連携サーバは Gfarm メタデータサーバであるため、連携サーバをマウントする事によりクライアントは連携された医療機関の Gfarm を単一のファイルシステムとして認識可能である。

また、各 I/O サーバに DICOM 画像が保存される際には、一般的なファイル転送に加えて、既存の DICOM ビューアからのファイル転送を可能にする DICOM 通信機能を実装した。

¹XML 文書の中の特定の要素を指し示す記述方法を定めた規格。

4.3.1 保存登録処理

DICOM を保存する際には、DICOM のメタデータから XML の作成を行い、Gfarm の XML 拡張属性に、作成したメタデータ XML を登録する。XML は各医療機関のメタデータサーバに保存される。ファイルサイズの大きな DICOM の画像情報にアクセスすることなく、ファイルシステム自体がファイルの検索を行うことから高速な検索を期待できる。一方、格納される DICOM は I/O サーバに保存される。保存先 I/O サーバは各サーバの負荷情報から Gfarm が自動的に選択を行う。

DICOM 通信を用いて画像を保存する際にも、受信と同時に DICOM からメタデータ XML が作成され登録が行われる。

4.3.2 検索処理

Gfarm メタデータサーバに格納された XML から検索が行われる。検索項目は複数設定することが可能であり、XPath により指定する。

4.3.3 取得処理

DICOM を取得する際には、メタデータサーバにアクセスを行い目的ファイルの保存先情報を得る。クライアントは I/O サーバとの直接の通信によりファイルを取得する。

4.3.4 外部の医療機関に対する処理

外部の医療機関に対しては、DICOM ファイルの書き込みは必要はなく、検索と読み込みが必要である。検索を行う際には、Fig. 4 に示す連携サーバが各医療機関のメタデータサーバにアクセスを行う。検索は各医療機関で実行され、目的のファイルが保存された I/O サーバより取得を行う。

4.4 GUI 実装

4.3.1～4.3.4 節の機能を実現する GUI の構築を行った。今回、ユーザの利便性を考慮しブラウザから操作が可能な WebGUI を実装する。以下の環境での実装を行った。

構築した GUI の内、DICOM 送信画面を Fig. 6 に、DICOM 検索画面を Fig. 7 に、DICOM 取得画面を Fig. 8 にそれぞれ示す。

5 評価実験

本章では実装した提案システムを用いて、構築システムが十分な性能を示すことを確認すると共に、構築システムが適切に動作を行っている点を確認するため速度評価実験を行った。実験は以下の2種類を行った。

(1) 基盤システムにおける速度比較実験

医用画像向けのファイルシステムとして Gfarm が十分な性能を示すことを確認するため、比較対象として小規模なファイル共有システムとしてよく知られている NFS (Network File System)²¹⁾ を用いて速度比較実験を行った。

(2) DICOM 通信における速度評価実験

通常、既存の DICOM ビューアからは DICOM 通信によって保存、検索、取得が行われる。比較対象として、簡易な PACS としても機能する DCMTK を用いて、Gfarm、構築システムとの速度比較実験を行った。

5.1 実験システム

今回、実験システムとして、メタデータサーバ 3 台 (内 1 台は連携サーバ)、I/O サーバ 4 台を用いて同一 LAN 内に Gfarm を 2 つ構築した。また、メタデータサーバでは、XML 拡張属性を有効にしてインストールを行い、その他はデフォルトで設定、認証方式は共通鍵認証で行った。Fig. 9 に実験システム図を、Table 2 に各サーバの仕様を示す。

5.2 基盤システムにおける速度評価実験

5.2.1 実験概要

本節では、実装した実験システムを用いて、医用画像向けのファイルシステムとして Gfarm が十分な性能を示すことを確認するため評価実験を行った。今回、評価を行うにあたり Gfarm の構築を行い、小規模なファイル共有システムとしてよく知られている NFS を比較対象として用いた。また DICOM のデータサイズに関して、最新の MRI は 2048 × 2048 ピクセル (約 420 万画素) の撮影が可能である。実験にはそのサイズの DICOM 画像を用いた。Table 3 に用いた DICOM と XML のファイルサイズを示す。100 ファイルの DICOM および XML を対象に、書き込みと削除処理に要する時間を計測し評価した。処理に要する時間は time コマンドによる計測を 5 回行った。

実験では以下の 5 点について速度計測を行った。

- (1) DICOM と XML 保存
- (2) DICOM 保存
- (3) XML 保存
- (4) DICOM から XML の作成
- (5) XML 拡張属性を用いての検索性能

5.2.2 実験結果

以下に実験結果を示す。なお、real は実経過時間、user はユーザ CPU 時間、system はシステム CPU 時間を表す。

5.2.2.1 DICOM と XML 保存

Gfarm と NFS に DICOM と XML を同時に保存した場合の平均処理速度の比較を行った。NFS では 1 台に DICOM を保存し、もう 1 台に XML の保存を行った。Gfarm では I/O サーバ 2 台に DICOM を保存し、メタデータサーバで XML 拡張属性を付加した。Fig. 10 に測定結果を示す。

5.2.2.2 DICOM 保存

Gfarm と NFS に DICOM のみを保存した場合の平均処理速度の比較を行った。Fig. 11 に測定結果を示す。

5.2.2.3 XML 保存

Gfarm と NFS に XML のみを保存した場合の平均処理速度の比較を行った。NFS では通常のファイル保存の際の時間を計測し、Gfarm では既に I/O サーバに保存してある DICOM に対してメタデータサーバで XML 拡張属性を付加する時間を計測した。Fig. 12 に測定結果を示す。

5.2.2.4 DICOM から XML の作成

DICOM から XML を作成する時の平均処理速度計測を行った。なお、実装には Python を用いた。Fig. 13 に DICOM100 ファイルから XML を作成した時の測定結果を示す。

5.2.2.5 XML 拡張属性を用いての検索性能

Gfarm の XML 拡張属性を用いて DICOM ファイルを検索した際の時間を計測した。具体的には、DICOM1 万ファイル (126GB) を Gfarm 上に保存し、XML 拡張属性を用いて画像 ID 番号を検索した。なお、画像 ID 番号は全ファイルで異なるものである。検索については外部の Gfarm に対して検索を行った場合の時間についても測定した。Fig. 14 に測定結果を示す。

5.2.3 議論

今回の評価実験において、Fig. 10 に示したとおり Gfarm は NFS より書き込みに要する時間が短いという結果であった。Fig. 11 より、DICOM (1.26GB) 保存の時間が測定時間の大半を占めており、Gfarm が NFS より高速な点はこの処理の結果である。この理由としては、Gfarm はオーバーヘッドを小さく抑える機構が実装されていることが挙げられる。

Fig. 12 より、XML を個別のサーバに書きこむ処理では Gfarm は NFS よりも長い処理時間を要した。これは Gfarm の XML 書き込み処理には、XML 拡張属性の付加処理があるものと考えられる。しかし、処理時間は 0.67 秒と非常に小さな時間である。

Fig. 13 より、DICOM 100 ファイルから XML を作成するために 5.8 秒の時間を要した。XML 作成に 5 秒程度であれば Gfarm において Fig. 10 の結果と合わせても 27.5 秒であり 1.26GB のデータの処理時間として妥当であるといえる。

Fig. 14 より XML 拡張属性を用いた検索では、1 万件 (126GB) の検索に 6.07 秒という結果を得られた。一方、外部の Gfarm に対する検索の場合は 6.08 秒であり内部検索の場合と同時間程度と言え

る。また、今回は同一 LAN 内での評価であるため、今後遠隔地間での性能評価を行う必要があると考えられる。

5.3 DICOM 通信における速度評価実験

5.3.1 実験概要

本節では、DICOM 通信機能において、既存 PACS と Gfarm の性能との比較を行うため速度評価実験を行った。今回、評価を行うにあたり、簡易な PACS としても機能する DCMTK を比較対象とした。また、Gfarm 上に DCMTK を構築し、Gfarm 上で DICOM 通信を実現する際の速度評価についても行った（以下、構築システムとする）。構築システムの概要図を Fig. 15 に示す。

DCMTK では、ファイルを受信と同時に検索インデックスを作成する。構築システムでも同様に、検索インデックスが作成される。一方、保存される DICOM は複数の I/O サーバ上に分散して保存される。

評価実験では、DCMTK、構築システム、Gfarm のそれぞれに対して、DICOM の保存、送信、検索の比較を行う。DCMTK、構築システムでは DICOM プロトコルを用い、Gfarm では通常の通信を用いている。DICOM のデータサイズについては、一般的に用いられているサイズとして 256×256 ピクセル、199KB の DICOM を用いた。評価については 100 回試行の平均値を用いた。

5.3.2 実験結果

5.3.2.1 DICOM 保存

DCMTK、構築システム、Gfarm のそれぞれに DICOM 100 件を保存およびデータベースへ登録した際の処理速度の比較を行った。Fig. 16 に測定結果を示す。Real はネットワーク遅延を含む実経過時間、User はプログラムの処理時間、Sys は OS の実行時間を示す。

5.3.2.2 DICOM 送信

DCMTK、構築システム、Gfarm のそれぞれが DICOM 1 件を送信した際の処理速度の比較を行った。Fig. 17 に測定結果を示す。

5.3.2.3 DICOM 検索

DCMTK、構築システム、Gfarm のそれぞれにおいて ID が全て異なる DICOM 5,000 件から 1 件を検索した際の処理速度の比較を行った。Fig. 18 に測定結果を示す。

5.3.3 議論

今回の評価実験において、Fig. 16 より構築システム (Real) が最も時間を要している事が確認出来る。構築システムでは DCMTK を用いての検索インデックスの作成および Gfarm での分散保存が行われていたため時間を要した。また、Gfarm の User 時間が DCMTK、構築システムの User 時間より時間を要している。Gfarm では検索データベースとして XML が用いられており、受信した DICOM から XML に変換の後データベースへの格納を行っている。構築システムの user 時間は低いことから、データベースへの登録に関しては Gfarm が DCMTK に比べ時間を要していると考えられる。

Fig. 17 より、DICOM の送信に関しては Gfarm が高速であった。DCMTK、構築システムでは DICOM 通信が用いられている。DICOM 通信では DICOM を取得する際でも検索を行った後、目的

の DICOM が送信される。そのため、単純に比較することは出来ない。構築システムと DCMTK の差分より、分散保存した場合の処理時間を確認することができる。

Fig. 18 より、Gfarm の検索性能は件数が増加するにつれ線形に推移していることが確認出来る。一方で、DCMTK は件数が増加しても処理時間が非常に短い。これについて、DCMTK では検索の際よく用いられる項目を優先的に検索出来るアルゴリズムが組まれている。一方の Gfarm では、DICOM のメタデータ情報全てからの検索が行われている。そのため、Gfarm の処理結果が低速であると考えられる。また、Gfarm の処理結果が線形に推移していることから、XML 検索において適切に検索インデックスが作成されていない事が考慮される。構築システムについては、DCMTK の検索インデックスをレプリカしているため DCMTK と同様の結果となった。

6 結論

本論文では、医療機関における医用画像の共有および検索のためのプラットフォームとして、広域分散ファイルシステム Gfarm 上に医用画像規格 DICOM を保存する医用画像保存通信システムを提案した。提案システムでは医療機関が個別に分散ファイルシステムを構築し、それらを連携することで医用画像の共有基盤を構築する。基盤システム上に DICOM を格納する際には、DICOM メタデータをメタデータサーバに登録し、DICOM ファイルを I/O サーバに格納する。DICOM の検索を行う際には各医療機関が個別に管理を行うメタデータ上検索が行われる。外部の病院からの検索リクエストに対しても、連携サーバを通して検索が行われる。

本稿では、以上のようなシステムの実装を行い、その性能について評価を行った。NFS との比較実験を通して、提案システムが高いアクセス性と広域分散システム向けの Gfarm の特徴を利用すると同時に、DICOM が持つメタデータファイルを XML 拡張属性として Gfarm のメタデータサーバに保存する方法で検索性の向上を図れることが確認された。一方で、既存の PACS としての DCMTK との比較においては、DICOM 保存登録に関して Gfarm が既存の PACS よりも検索項目の登録の際に時間を要している事が確認出来た。さらに、検索の処理時間においては、DCMTK に比べ圧倒的に低速である事が確認できた。検索時間も適切な検索インデックスが作成されているならば対数的に推移するはずであるが、検索ファイル数に対して線形に推移している。これらより、Gfarm の検索機能に関して今後検索インデックスの作成および優先的に検索項目を設定する必要性が確認出来た。ファイル送信処理より、一般的な通信方法に比べ DICOM 通信が時間を要することが確認出来た。DICOM 通信では TCP/IP 上で、さらに DICOM プロトコルでの通信を行っているため時間を要したと考えられる。

今回のシステム実装では、研究室内の LAN を用いて小規模なシステム構築を行った。今後、遠隔地間でシステムを連携させた際の遅延や速度評価、および本システムを実際に運用を行った上での評価が期待される。

謝辞

私が本研究室に配属されてからの3年間、同志社大学生命医学部の廣安知之教授には多大なる御指導、そしてご協力を頂き、心より御礼を申し上げます。また、様々な指摘、助言をして下さいました、同志社大学理工学部の本木光範教授に心より感謝致します。また、私がこの研究を進めるにあたり様々な助言を下さり、論文提出の際は励まし面倒を見て頂いた電気通信大学大学院情報システム学研究科の吉見真聡助教に心より感謝致します。

そして、日頃から研究ミーティングにて多く議論して頂いた医療システム班の藤井亮助さん、西村祐二さん、医用画像班の山口浩明さん、布川将来人さんに御礼申し上げます。山口浩明さんには私のコーチとして1年間、多めに議論して頂きました。また、本論文を校正してくださいました藤井亮助さん、星野雄地さん、西村祐二さん、中村友香さん、大西夏子さんに御礼申し上げます。お忙しい中、ご無理をお願いしたにも関わらず丁寧な校正をして頂き感謝しております。

最後に、知的システムデザイン研究室の皆様、医療情報システム研究室の皆様には私の研究に関して数多くの議論や助言をして頂きました。皆様のおかげで、3年間すばらしい研究生活を送ることができました。この場を借りて厚く御礼申し上げます。

参考文献

- 1) H. Oosterwojk. *DICOM Basics Third Edition edition*. OTech Inc, 2005.
- 2) H. Munch, U.Engelmann, A. Schroeter, and et al. The integration of medical images with the electronic patient record and their web-based distribution. *Acad Radiol* 2004.
- 3) O. Ratib, Y. Ligier, D. Bandon, and et al. Update on digital image management and pacs. *Abdom Imaging* 2000, Vol. 25, pp. 333–340.
- 4) C. Teng, J. Mitchell, C. Walker, A. Swan, C. Davila, D. Howard, and T. Needham. A medical image archive solution in the cloud. *Software Engineering and Service Sciences (ICSESS), 2010 IEEE International Conference on*, 2010.
- 5) 原量宏, 横位英人. 病院情報システムと遠隔医療かがわ遠隔医療ネットワークから日本版 EHR の実現へ. 医療機器システム白書 2008~2009, pp. 358–360, 2008.
- 6) Frost & Sullivan. Prepare for disaster & tackle terabytes when evaluating medical image archiving, 2008. <http://www.frost.com>.
- 7) R. Zheng, H. Jin, Q. Zhang, and P. Chu. Heterogeneous medical data share and integration on grid. Proceedings of 2008 International Conference on Biomedical Engineering and Informatics (BMEI '08), IEEE Computer Society Press, 2008.
- 8) S.G. Erberich, J.C. Silverstein, A. Chervenak, R. Schuler, M.D. Nelson, and C. Kesselman. Globus medicus — federation of dicom medical imaging devices into healthcare grids. *Stud Health Technol Inform*, Vol. 126, pp. 269–278.
- 9) O. Tatebe, K. Hiraga, and N. Soda. Gfarm grid file system. *New Generation Computing*, Vol. 28, pp. 257–275, 2010.
- 10) 建部修見, 曾田哲之. 広域分散ファイルシステム gfarm v2 の実装と評価 (グリッド i). 情報処理学会研究報告, Vol. 2007, No. 122, pp. 7–12, 2007-12-07.
- 11) 近藤博史. 電子カルテとの統合 画像システムの組み込みと SBC の導入. 医療機器システム白書 2008~2009, pp. 358–360, 2008.
- 12) National Electrical Manufacturers Association (NEMA), 社団法人日本画像医療システム工業会 (JIRA). DICOM 規格書 (日本語訳). 2000.
- 13) H. Oosterwojk. DICOM 入門. 篠原出版新社, 2008.
- 14) 櫛橋民生, 武中泰樹. PACS の構築と運用のすべて 画像検査部門の IT 化. (株) 医療科学社, 2010.
- 15) 小掠真貴. 癌診断支援のための病理画像解析システム. 病理と臨床, Vol. 24, No. 4, pp. 411–415, 2006.

- 16) 昭彦鈴木, 聰彦伊藤, 憲明大内. 乳癌検診システムの精度向上に向けて. 日本乳癌検診学会誌 = Journal of Japan Association of Breast Cancer Screening, Vol. 18, No. 1, pp. 13–19, mar 2009.
- 17) 横井英人, 福井敏樹, 寅野貴史, 尾藤茂, 原量宏. 遠隔医療ネットワークシステムにおける画像通信機能の強化. <http://www.kcgh.gr.jp/department/d48/files/25th/jcmi25/paper/p10273/p10273.html>.
- 18) 建部修見, 曾田哲之. 広域分散ファイルシステム gfarm v2 の実装と設計と実装. 情報処理学会研究報告, Vol. 2004, No. 81, pp. 145–150, 2004-07-30.
- 19) S. Mikami, K. Ohta, and O. Tatebe. Data intensive distributed computing using mapreduce on gfarm file system. *Information Processing Society of Japan*, Vol. 2010, No. 4, 2010.
- 20) M. Eichelberg, J. Riesmeier, T. Wilkens, A. J. Hewett, A. Barth, and P. Jensch. Ten years of medical imaging standardization and prototypical implementation: the DICOM standard and the OFFIS DICOM toolkit (DCMTK). In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 5371, pp. 57–68, April 2004.
- 21) B. Callaghan, Pawlowski, and Staubach. NFS Version 3 Protocol Specification. *RFC 1813*, 1995.

付 図

1	Outline of DICOM	1
2	Flow of DICOM Negotiation	1
3	Constitution of Hospital systems	1
4	Outline of proposal system	2
5	Gfarm architecture	2
6	DICOM Upload GUI	2
7	DICOM Retrieved GUI	3
8	DICOM Download GUI	3
9	Experimental System	4
10	Results for saving DICOM and XML Files	4
11	Results for saving DICOM Files	5
12	Results for saving XML Files	5
13	Results when Creating XML from DICOM Files	6
14	Results for Retrieval by using XML Extended attributes	6
15	Outline of Construction System	7
16	Results for Upload and Register DICOM Files	7
17	Results for Download DICOM Files	8
18	Results for Retrieval DICOM Files	8

付 表

1	Environment of System	9
2	Specification of a Server	9
3	DICOM and XML File Size	9

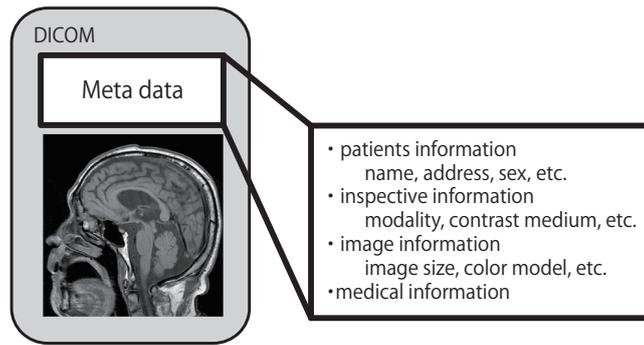


Fig. 1 Outline of DICOM

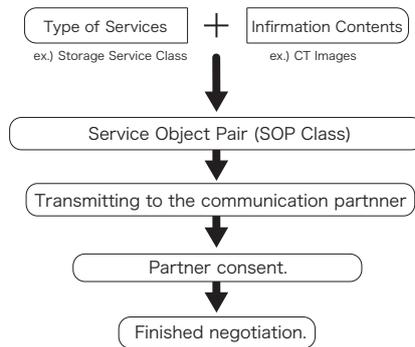


Fig. 2 Flow of DICOM Negotiation

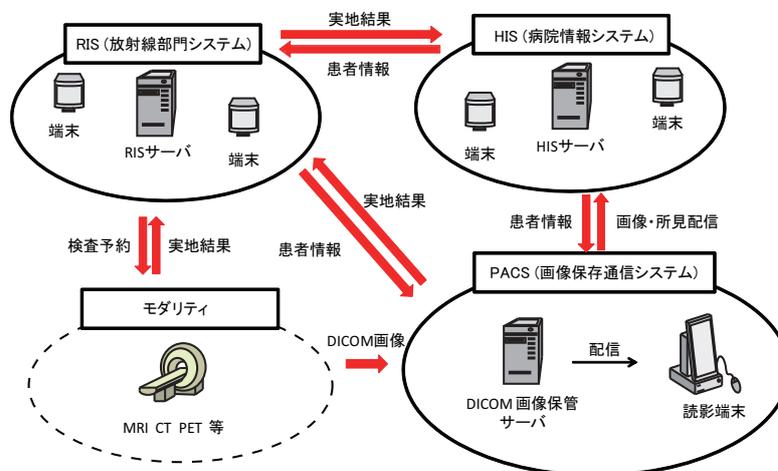


Fig. 3 Constitution of Hospital systems

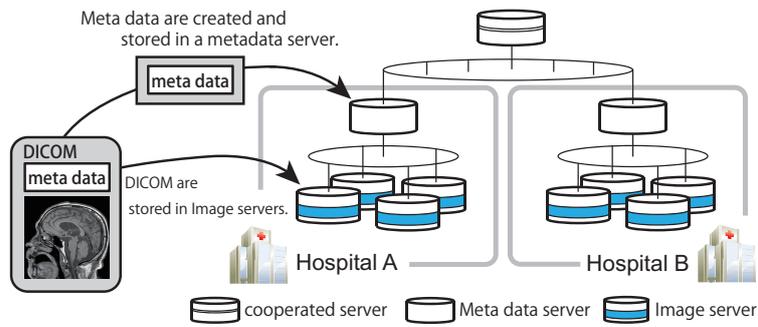


Fig. 4 Outline of proposal system

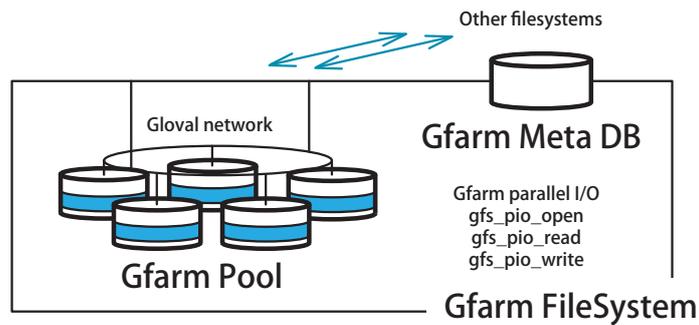


Fig. 5 Gfarm architecture



Fig. 6 DICOM Upload GUI

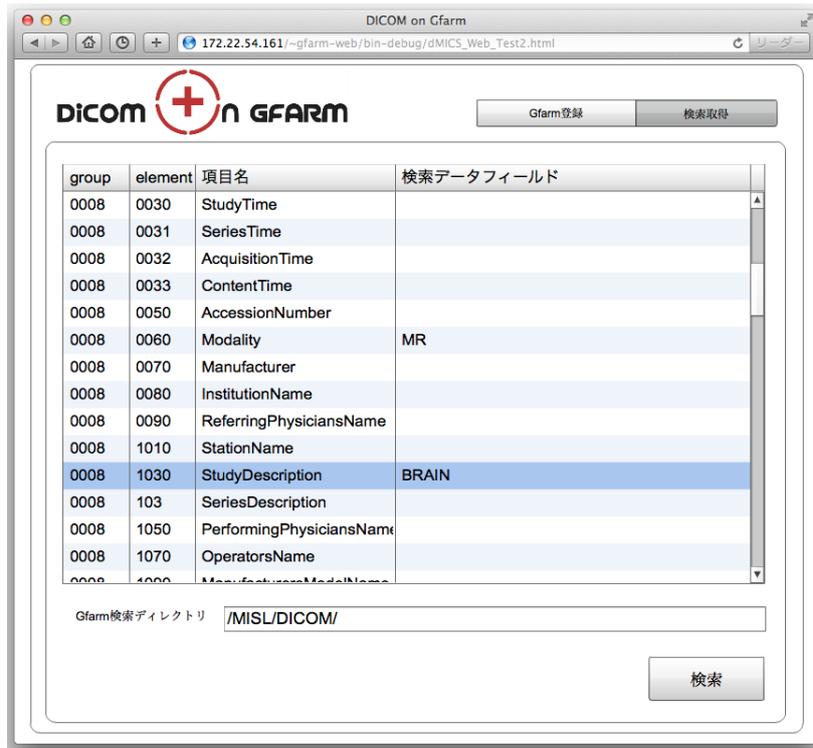


Fig. 7 DICOM Retrieved GUI

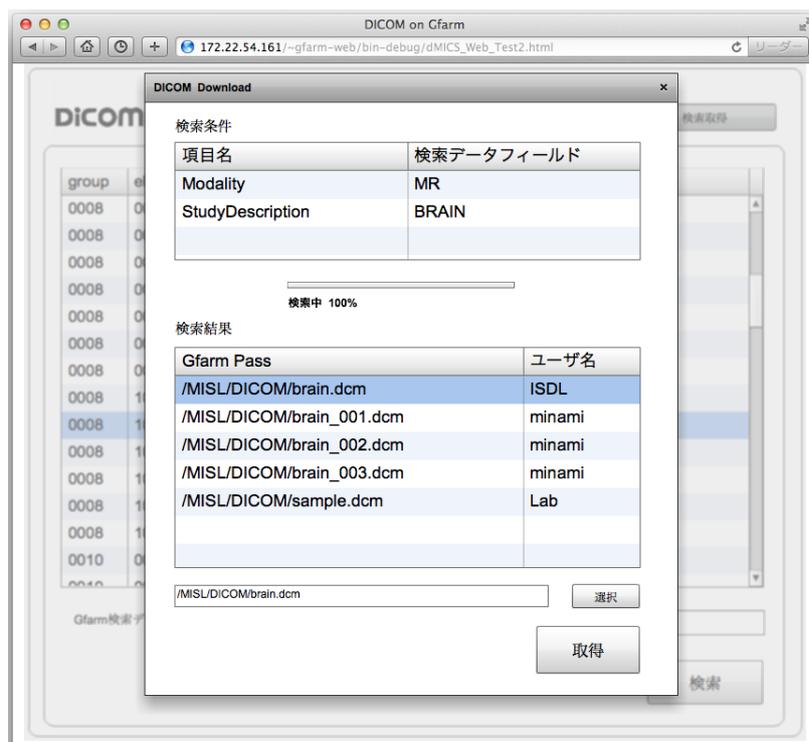


Fig. 8 DICOM Download GUI

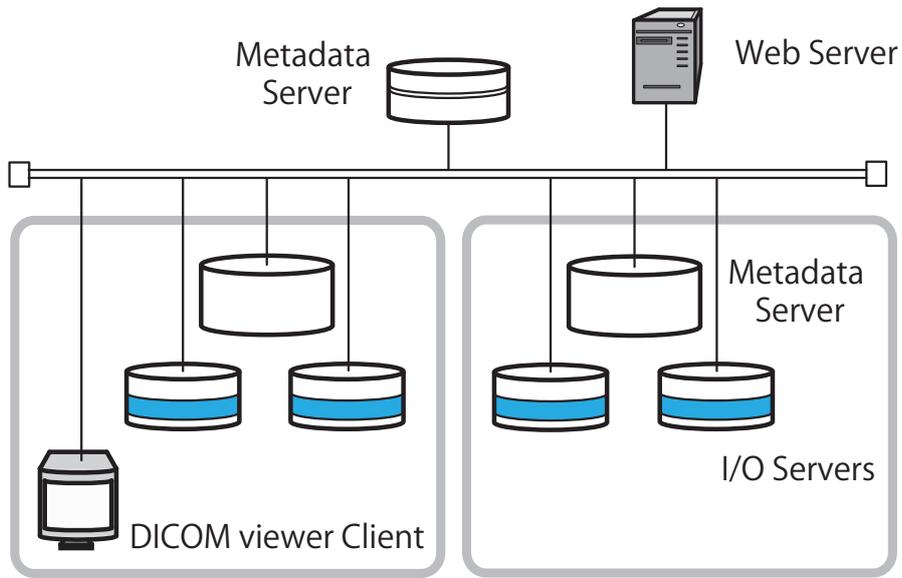


Fig. 9 Experimental System

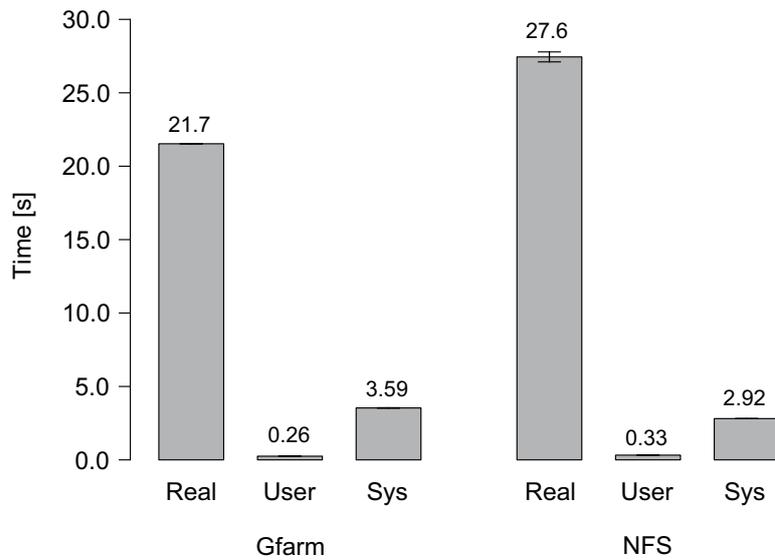


Fig. 10 Results for saving DICOM and XML Files

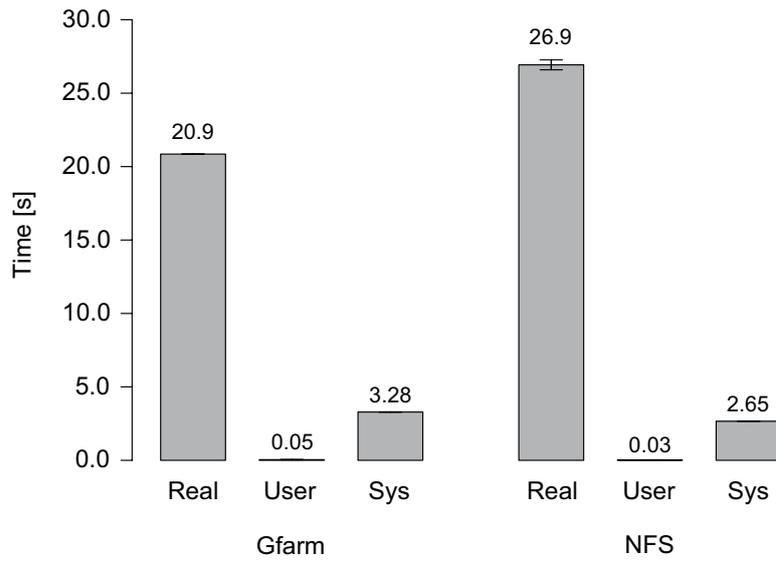


Fig. 11 Results for saving DICOM Files

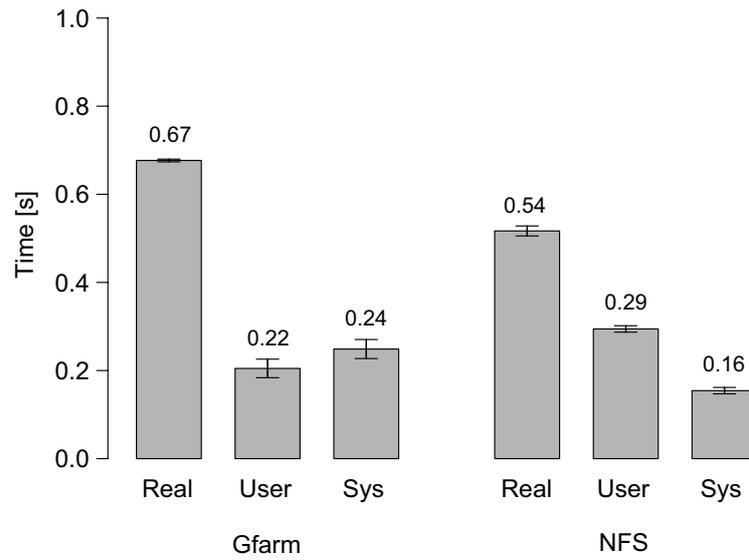


Fig. 12 Results for saving XML Files

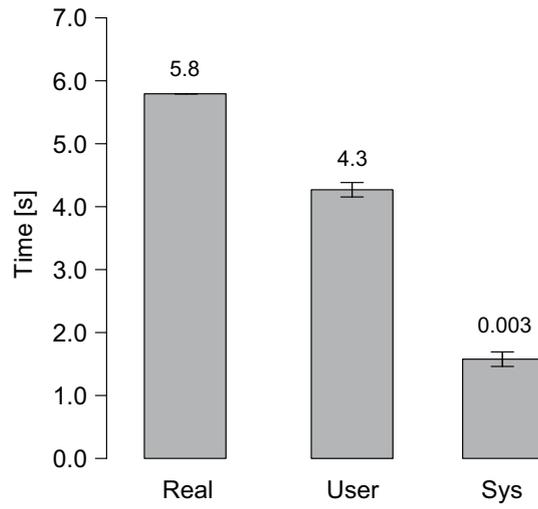


Fig. 13 Results when Creating XML from DICOM Files

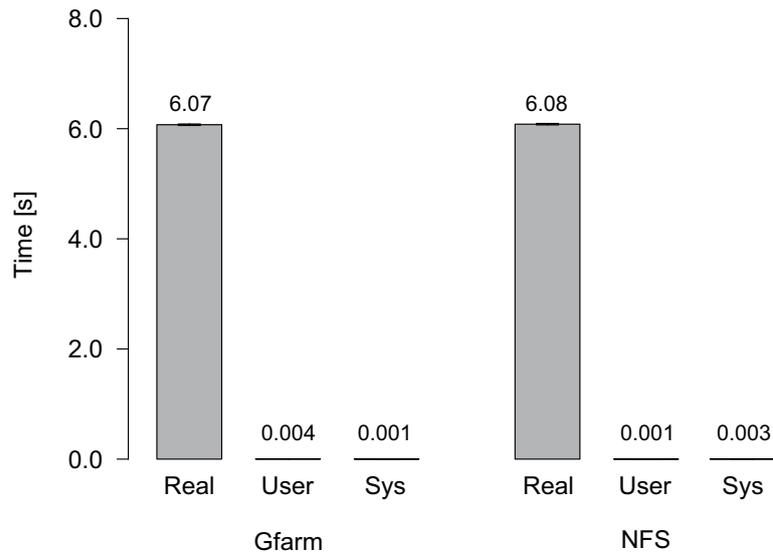


Fig. 14 Results for Retrieval by using XML Extended attributes

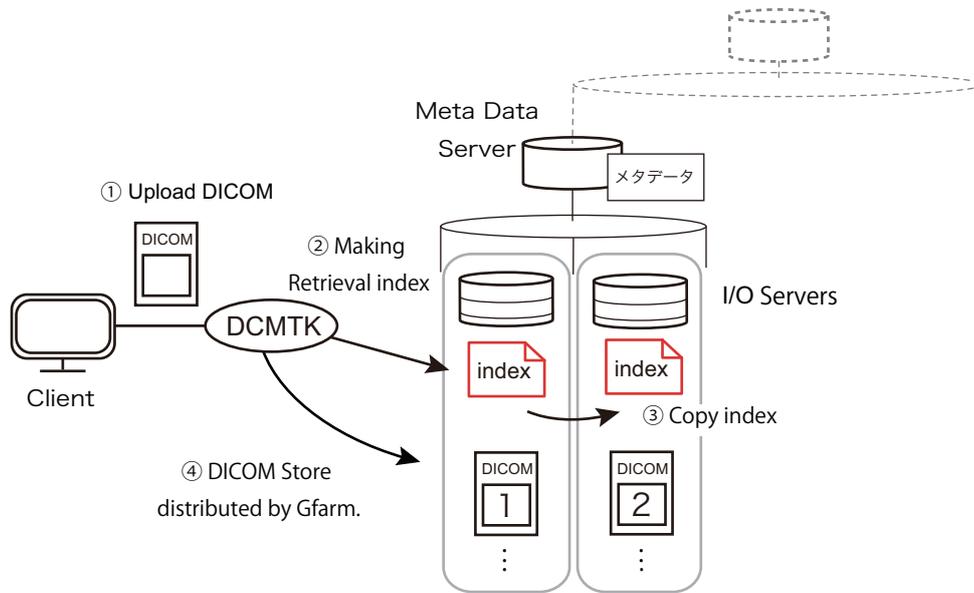


Fig. 15 Outline of Construction System

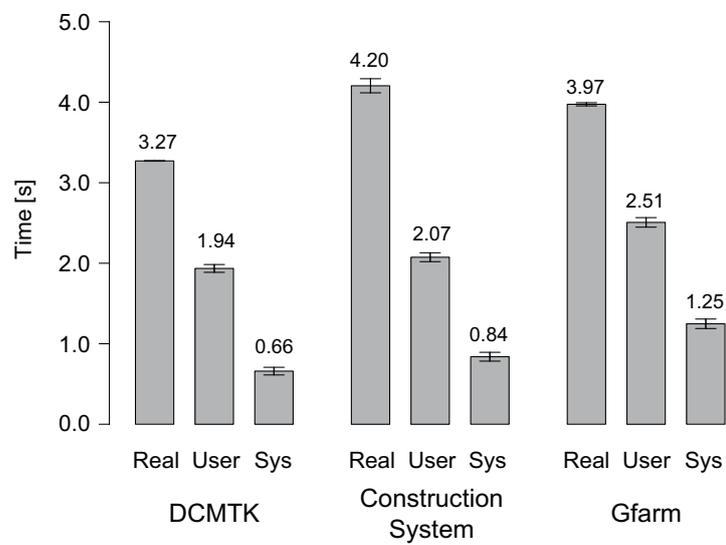


Fig. 16 Results for Upload and Register DICOM Files

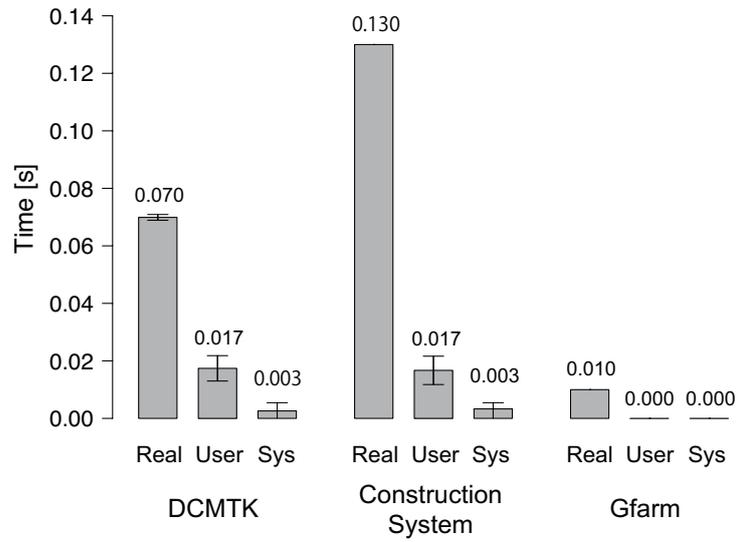


Fig. 17 Results for Download DICOM Files

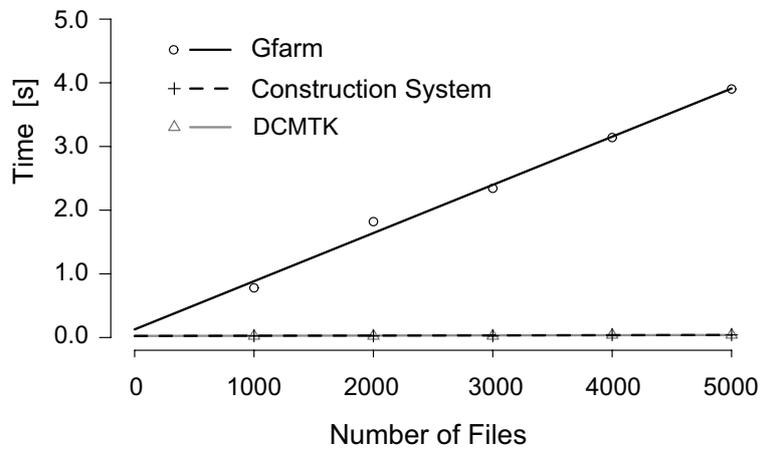


Fig. 18 Results for Retrieval DICOM Files

Table 1 Environment of System

Development	Flash 4.5
Language	Action Script 3.0 Python 2.7.3
Web Server	Apache 2.4.3

Table 2 Specification of a Server

CPU	Quad-Core AMD Opteron 2.3GHz \times 2
Chipset	AMD 8131+8111
Memory	DDR2 667 MHz 8GB
NIC driver	Broadcom Tigon3
OS	Debian 6.0.3
gfarm	2.5.3
gfarm2fs	1.2.6

Table 3 DICOM and XML File Size

File	Size	Number	Sum size
DICOM	12.6MB	100	1.26GB
XML	7KB	100	0.7MB