



## 全米デジタル・マモグラフィー・アーカイブ： ペンシルバニア大学コンソーシアムとIBMが開発した 乳ガン検診用コンピューティング・グリッド



### 概要

#### 課題

患者ファイルの高速検索および診断評価のため、複数の場所からそのファイルを電子的に取り込み、管理、保存すること。

#### ソリューション

IBM® @server™、DB2®、Universal Database™、および GPFS ファイルシステムを利用した IBM グリッド・コンピューティング

#### メリット

全米デジタル・マモグラフィー・アーカイブは、患者のセキュリティ保護されたスケーラブルな電子診断ファイルを90秒以内で提供できる。

#### 課題

ペンシルバニア州フィラデルフィアにあるペンシルバニア大学では数年にわたり、全米科学財団 (NSF)、米国医療図書館 (NLM)、米国国立衛生研究所 (NIH)、および次世代インターネット (NGI) などの関係機関から補助金を受けて、画期的な電子医療記録 (EMR) データグリッドとリポジトリを開発しています。患者中心の医療画像記録システムを開発し、どこからでも医療ファイル全体の取り込みを可能にすることが課題です。ファイルのデータには、忠実度の高い医療画像 (CT、MRI、乳房 X 線写真)、医療記録、病歴などが含まれます。この課題を達成するためには、患者の記録を電子データとして取り込むためのネットワーク化したシステムの構築、高速検索や、比較、診断補助のための大容量ファイルの管理と保存、患者の記録に関する米国連邦および州の規制を満たすために必要となるセキュリティおよびプライバシー基準の保証が必要になります。

ペンシルバニア大学のロバート・ホルビーク博士の指導のもとに、EMR データグリッドの研究として、まずスケーラブルな全米デジタル・マモグラフィー・アーカイブ (NDMA) の開発から取りかかりました。乳房 X 線撮影 (マモグラフィー) にはすでに規格とプロトコルが提供されており、最大のデータ・ファイルが蓄積されているからです。IBM からの共同大学研究 (SUR) の補助金をもとに、ペンシルバニア大学主導で NDMA 構築のための共同事業が始まりました。現在は開発の第 2 段階に入っており、オークリッジ国立研究所 (BWXT の ACT) のグループの協力によって、ペンシルバニア大学グリッドには、ペンシルバニア大学、シカゴ大学、およびノースカロライナ大学の各病院と、トロントのサニーブルック女子大付属病院が接続されています。システムには IBM @server UNIX と IBM の DB2 Universal Database を実装した Linux® を組み合わせて、今までにない強力なソリューションが構築されています。

患者の健康管理に必要な電子データの取り込み

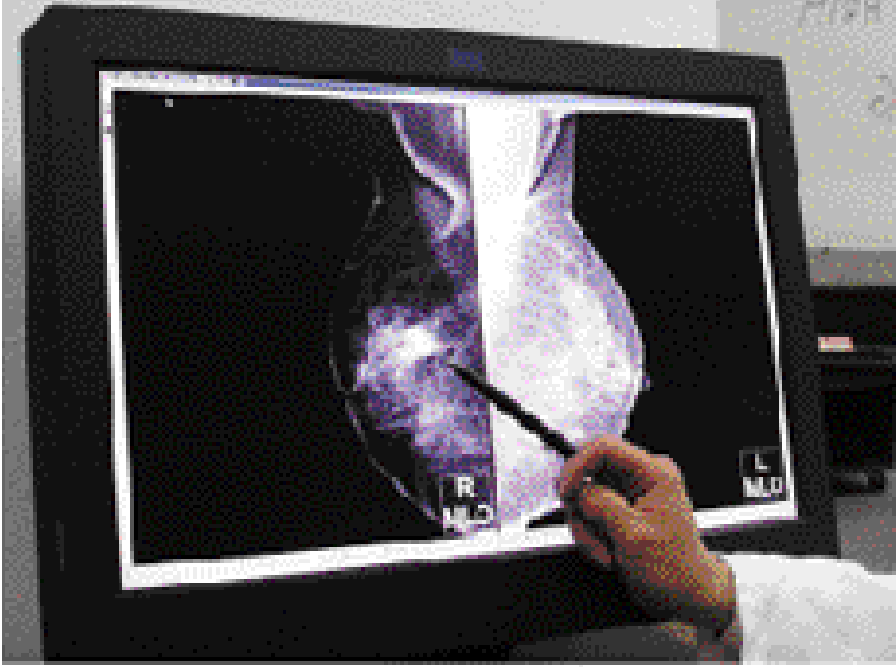
病院によっては、さまざまな理由から過去の乳房X線写真のファイルが利用できず、放射線専門医は比較する記録なしに診断しなければなりません。その結果、生検などの診断をさらに行わなければなりません。患者が別の場所で乳房X線撮影を受けると、やはり記録が途切れてしまいます。NDMA では、現在および過去の患者の記録を2～90秒でアクセスできるようになります。診断に必要な記録が最適化され、患者の健康管理を強化します。それに加えて、複数の病院や診療所で使われていた膨大な「カルテ」やフィルム・ファイル・システムなどのオーバーヘッドと管理コストを削減できます。また、患者の健康の安全性を高め、潜在的な訴訟費用を削減するといったメリットがあります。

オープン・スタンダードに基づいて構築されたペンシルバニア大学グリッドは、安全なインターネット接続を介して、電気や水道のようなユーティリティ・サービスとしてコンピューティングリソースを供給する、大規模な分散型コンピューターです。数千におよぶ病院がデジタル形式で乳房X線写真を保存することが可能で、許可を受けた医療関係者がほとんど同時に患者の記録にアクセスでき、高価なレントゲン・フィルムの必要性が減少します。病院はセキュリティー保護されたインターネット・ポータルを介してグリッドに接続されます。許可を受けた医師はこのポータルで、デジタル化されたレントゲンデータをアップロード、ダウンロード、および解析して、腫瘍の可能性などの問題を特定することができます。高度なアルゴリズムを使って、特定のコミュニティでのガン「集団発生」や、病気の異常集中など、その住民にみ

「患者の乳房X線写真をシステムに取り込めば、異常を検出する強力なツールを使って、現在のレントゲン写真を数年前の写真と比較し、すばやく診断を行うことができます。個々の患者のレントゲン写真はさまざまな医療機関に散らばっていることが多く、従来は必要なときに探し出すことが困難でした。このグリッドを利用すると、許可を受けた医師に、患者の重要なデータをすべて非常にすばやく、効率よく、安全に提供できるようになります」

ロバート・ホルビーク医学博士  
ペンシルバニア大学 National Scalable Cluster (NSCP) 研究所所長

られるパターンを発見することができます。医師が個々の症状を診断し、ガンの特定を支援する解析ツールが導入されています。また、医師、医学生、およびインターンが乳ガンや関連する病気について学ぶのを助けるための一連のツールもグリッドに導入されています。放射線専門医の養成用ツール、コンピューター支援診断 (CAD) ツール、および診断比較のために乳房X線写真を「疑いがある」か「疑いはない」かに分類するフラクタイレ・ベースの解析用ツール (FarGen: Focus Attention Region Generation) なども導入されています。



新しいIBM T221、92メガピクセルの  
フラット・パネル・モニターで科学および医療用の  
超高解像度画像を表示

#### グリッドの利点

高圧送電網を介して電気が各家庭に供給されるように、コンピューティンググリッドを介して、地理的に分散している組織でアプリケーション、データ、およびコンピューティング・リソースを共有することができます。コンピューティングの新しいモデルであるグリッドは、インターネットを

介して接続されたサーバーなどのリソース・クラスタでありGlobus オープンソース・コミュニティ (Globus.org) が提供するプロトコルや、Linux などのオープン技術を利用します。患者のデータのアクセスおよび管理にグリッド・コンピューティングを利用する利点としては、高速のデータ検索、スケーラビリティ、コストの節約などがあります。

## IBM eServer および DB2: NDMAグリッドの仕組み

ペンシルバニア大学はグリッドを、ほとんどすべてIBMハードウェア上で稼働するポータルシステムを使って構築しました。LinuxおよびWindows® 2000 が稼働する16台の分散型IBM Netfinity®サーバーなどが使用されています。NSCP 研究所に設置されたメインのアーカイブシステムもIBM機器に大きく依存しています。このグリッド・コンピューティング・アーカイブの開発に関する問題としては、次の点が挙げられます。

- ・データ量。1回の検査ごとに最大160メガバイトの画像が必要になります。もし、検診に必要な年齢の女性の60%が毎年乳房X線写真の検診に応じるとすると、1回の検診に160メガバイトとして、NDMAのデータ量は年間5.6ペタバイトを超えることになります。
- ・現在の検査の保管と、現在と過去の検査の比較のためのトラフィックのスケジューリング。1日のトラフィックは、最低でも8テラバイトと予測されます。
- ・専門医の診察や予定外の検査に必要な高速アクセスから、低速のインターネット・アクセスに対応する調整可能なネットワーク帯域幅。
- ・安全な公開ネットワーク上での画像転送や個々の患者のデータを含む情報すべての暗号化。

データを取得したデバイスやワークステーションから、地方の病院またはエンド・ノードの安全なポータルを使って、データがポータルにロードされます。ここでIBM eServer Cluster 1600 UNIXシステムを使った外口ポータル・ハブやエリア・アーカイブにデータが転送されます。グリッドが完全に配備されると、外口ポータル・ハブやエリア・アーカイブが、リソースのプールのために1台の大容量の地域ハブにまとめられます。現在、この地域ハブはIBM eServer Linux Clusterで試作されています。分散型アーカイブの容量の総計と同じ容量を持つひとつの巨大なアーカイブを、分散型アーカイブがエミュレートします。地方の病院からのアクセスはトランスペアレントです。

## DB2 Universal Database の 処理能力

ペンシルバニア大学のチームとIBMのチームは共同で、デジタル化したレントゲンデータの安全で使いやすいインデックスとして機能するよう、高速アクセスが可能な大容量のDB2 Universal Databaseを実装しました。このシステムの最高速度での動作はDB2 Universal Databaseの平行化テクノロジーの機能によって強化されています。平行化テクノロジーは、ギガバイトの性能で多重に送受信が可能な平行化ハードウェアシステムと同様に、こうしたネットワークの大きな帯域幅遅延を扱うための主要コンポーネントです。このテクノロジーを使って、複数のプロセッサで並行して実行が可能な実行コンポーネントに処理を

細分することにより、大規模データベースでのスケーラブルな性能を実現します。

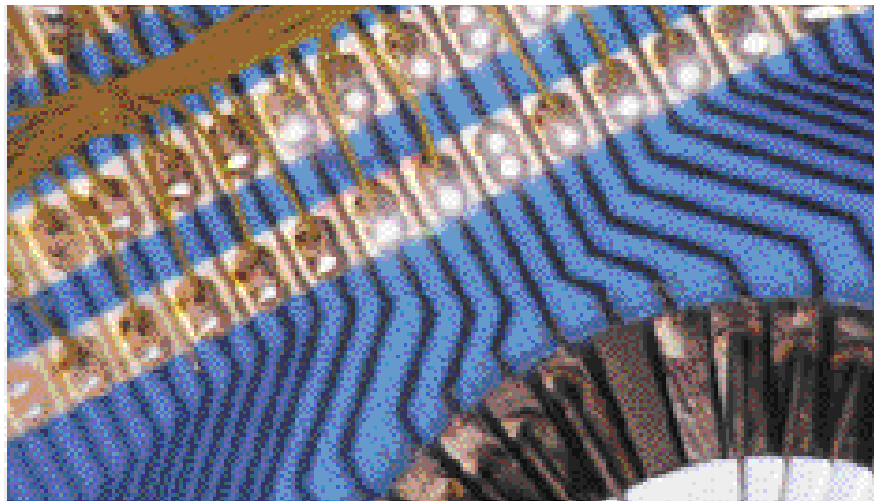
DB2 Universal Database Extended-Enterprise Edition (UDB EEE)の平行化データベースは、平行化クエリー・メカニズムとともに、入力データのインデックス化とカタログ化、および収集されたデータ全体への高速アクセスの供給に非常に高速なデータ・パフォーマンスを提供します。DB2 Universal Database EEEは、データ検索の最も効果的な手段を選択するオプティマイザーのパフォーマンスにおいても、その最先端にあることを示しています。

将来の発展のための  
スケーラブルな統合システム

NDMAグリッドの3層アーキテクチャーは、Globusのオープン・プロトコルを使って、IBM @server の長所を活用しています。ユーザーレベルでは、グリッドに参加している各病院に2台のIBM @server xSeries システムからなるポータルが設置されています。Series 1台は、デジタルデータの一時的なリポジトリとして機能し、もう1台はインターネット、またはインターネット2と呼ばれる次世代インターネットへのリンクとして機能します。3層のシステムでは、AIX、Linux、およびWindowsが稼働し、本質的に異機種混在のコンピューティンググリッドの例証となっています。

IBM クラスタ・システム・マネージメント (CSM) ツールが、Linux クラスタの一元管理として使用される予定です。このツールは、ノードやノード・グループ情報の管理、ハードウェアのモニタリングと制御、複数ノード上でのリモート・コマンドの実行、およびシステムイベントのモニタリングなどを行います。すべてのノードの構成ファイルの管理や同期化、さらにセキュリティ保護や障害診断支援も行います。

NMDAグリッドはIBM ゼネラルパラレルファイルシステム(GPFS)の機能も利用します。これはアーカイブへの読み書きを行うための、非常に高速で使いやすいパラレルファイルシステムです。GPFSは、データ要求のピークの総計が分散型ファイルシステム・サーバー1台の容量を上回るような環境に特に適しています。GPFSは大型の画像ファイルの主要なストレージ・メカニズムになります。



## セキュリティおよびプライバシーの問題

患者の情報の電子データの取り込み全体を通して、臨床試験用か NDMAのような全米医療アーカイブ用かに関わらず、医療用アプリケーションにおけるセキュリティが、世論、米国連邦法、州法、および連邦の規制で義務付けられています。これらのうち主要なのは、1996年の「医療保険の携行性と責任に関する法律 (HIPAA)」です。HIPAAの情報保護要件には次の項目が含まれています。

- ・機密保護を保証するポリシーと要領
- ・コンピューターおよびネットワークの資産、設備、アクセス制御の物理的な保護を保証する物理的安全策
- ・ユーザーのアクセス制御の実施、識別および認証、アイデンティティ・ロールやコンテンツに基づくアクセス制御およびユーザーの動作の監査を保証する技術的セキュリティ・サービス
- ・通信セキュリティのための暗号化、データの保全制御、エンド・エンティティの認証などの技術的セキュリティ・メカニズム
- ・電子署名の組み込みや、ポリシーの実施の保証を行う補助的安全策

NDMAでのセキュリティへのアプローチは、各組織および政府の有効なすべてのポリシーへの対応を可能にする仮想ファイル・ルーム・コンセプトに基づいています。HIPAAなどの規制基準の準拠を保証するためにNDMAグリッドは、物理的セキュリティ、ハードウェア・セキュリティ、ソフトウェア・セキュリティ、および通信セキュリティを組み込んだ多層の統合セキュリティ・サービスを特徴とするセキュリティ・アーキテクチャーを土台としています。このセキュリティ・グリッドに適宜組み込まれる技術には、仮想プライベート・ネットワーク暗号方式、証明書認証、スマート・カード認証、ロールベースのアクセス制御、ローカルレベルおよびアーカイブ・レベルでのポリシーの定義と実施、状況モニタリング、および患者の識別番号の削除などがあります。

IBMのグリッドにおけるリーダーシップ  
ペンシルバニア大学のグリッドは、この分野でのIBMのリーダーシップを示す一連のグリッド・プロジェクトのうち最新のものです。IBMはノースカロライナのバイオインフォマティクス・グリッドの構築にも選ばれました。このシステムは、GlaxoSmithKline Inc.、Biogen、ノースカロライナ大学、デューク大学などの組織と共同で開発されます。IBMは、米国の4つの研究センターのコンソーシアムからも選ばれて、世界で最強のコンピューティング・グリッドを構築し、一連のLinuxクラスターを相互連結して毎秒13.6兆回の演算を処理できるようにします。このグリッド・システムはDistributed Terascale Facility (DTF) と呼ばれ、米国中の数千人の科学者が世界最高速の研究用ネットワークを介してコンピューティング資源を共有し、ライフサイエンス、気候モデリングなどの重要分野の画期的な研究を行います。IBMは英国のナショナル・グリッドのいくつかのセンターとも提携して、プロジェクトのための主要技術とインフラストラクチャーを提供しています。このプロジェクトは、英国全土のコンピューターの大規模ネットワークとリンクし、オランダの大学にも強力なコンピューティング・グリッドを構築しています。

「このグリッド・プロジェクトでIBMと提携したことを誇りに思っています。私たちには、高速なデータ検索、スケーラビリティ、そしてセキュリティを提供してくれる信頼性の高いハードウェアとオープン・システム・ソフトウェアが必要です。こうしたニーズは、IBMの@server クラスターの処理能力と汎用性に、IBM DB2 Universal Database、および IBMの GPFS パラレル・ファイル・システムを組み合わせることで、直接対処できます。デジタル化したレントゲン・データの安全で使いやすいリポジリーの構築にあたり、IBMのハードウェアとDB2は、多層の統合キャパシティーとセキュリティ・サービスを特長とする安全なアーキテクチャーを構築する土台になりました」

ロバート・ホルビーク医学博士  
ペンシルバニア大学 National Scalable Cluster (NSCP) 研究所所長

IBM ライフ・サイエンスソリューション、および  
IBM ライフ・サイエンス・ビジネス・パートナーの詳細については、  
IBM Web サイト、  
**ibm.com/solutions/lifesciences** を参照してください。

または、IBM ライフ・サイエンス・スペシャリスト、  
ls@us.ibm.com にお問い合わせください。

IBM, IBMロゴ, @server, DB2 Universal Database, DB2, Netfinity, xSeries, AIX は、IBM Corporation の商標。

Microsoft, Windows, Windows NT およびWindows ロゴは、Microsoft Corporation の米国およびその他の国における商標。

UNIX は、The Open Group がライセンスしている米国およびその他の国における登録商標。

Linux は Linus Torvalds の登録商標。

他の会社名、製品名、サービス名等は、それぞれ各社の商標または登録商標。

このお客様の話は、ペンシルバニア大学から提供された情報に基づいており、ある組織がIBM 製品をどのように使用しているかを示すものです。記載されている結果とメリットには、多くの要因が関連している可能性があります。M では、常に同じような結果が得られることは保証致しかねます。

---

IBM製品・サービスの詳細情報等については、  
IBM ホームページ <http://www.ibm.com/jp/>  
をご利用ください。

---

お問い合わせは、弊社営業担当員  
または、ダイヤルIBM (☎0120-04-1992)へ。  
受付時間：月～金 9:00～18:00 (祝日、12/30～1/3を除く)  
携帯電話でおかけのお客様は下記の電話番号をご利用ください。  
ダイヤルIBM 044-221-1522 (この場合、通話料金はお客様のご負担となります。)

---

'02-12月版



日本アイ・ビー・エム株式会社

〒106-8711 東京都港区六本木3-2-12  
12-02 Printed in Japan

このカタログの情報は2002年12月現在のものです。製品、サービス等詳細については、弊社もしくはIBMビジネス・パートナーの営業担当員にご相談ください。