グリッド技術を利用したセキュアデータベース

分散ファイルシステム(Gfarm)で情報漏洩を防止する

産総研が開発してきた分散ファイルシステム Gfarm を活用し、NTT ネオメイトと共同で「セキュアデータベース」を開発した。これは、データベースのレコードを数バイト単位に細かく分割して分散ファイル上で保管しておき、業務時にのみメモリ上でデータベースを動的に再構築する。アプリケーションを変更せずにデータベースを利用できると同時に、ディスクだけを持ち出しても情報の再現が不可能なシステムである。

A secure database was developed through a collaboration of NTT Neomeit Corporation using Gfarm file systems developed by AIST. Main features of the secure database are 1)Each record of the database is chopped into very small (several bytes) pieces and they are distributed over the Gfarm files systems in different places. 2)When an operator uses the database it is reconstructed temporarily on a RAM disk from the Gfarm file systems, so the data automatically disappear at an instance of system shutdowns. 3)Meta-data which contains locations of the divided pieces are encoded and decoded solely by authorized secret keys.

伊藤 智 Satoshi Itoh
satoshi.itoh@aist.go.jp
グリッド研究センター
ビジネス応用チーム 研究チーム長

グリッドポータルの開発など、グリッド技術のプロジェクトにおける GridASP の開発は主要な成果の一つである。その他、本件のようにグリッド研究センターにおける開発成果をもとにした企業との共同研究開発も積極的に推進している。グリッド技術がさまざまな分野で当たり前のように活用される時代を目指し、広範な分野へ研究を展開していきたい。

Gfarmによるセキュリティ強化

グリッド研究センターでは、大規模なデータを複数の拠点で協調して解析することができる分散ファイルシステム「Gfarm」を研究開発している。Gfarmでは、1つのファイルを任意の大きさに分割して管理できるので、広い地域に分散しているPCのディスクを多数まとめて仮想的に巨大なディスク装置のように扱える。また、データの複製管理機能も持つなど、データの信頼性と並列処理の性能が著しく向上しており、科学技術分野、ビジネス分野における大規模データ処理を実現するシステムとして有効である。

今回、共同開発を行ったNTTネオメイトでは、コールセンターのアウトソーシング業務として、セキュリティ

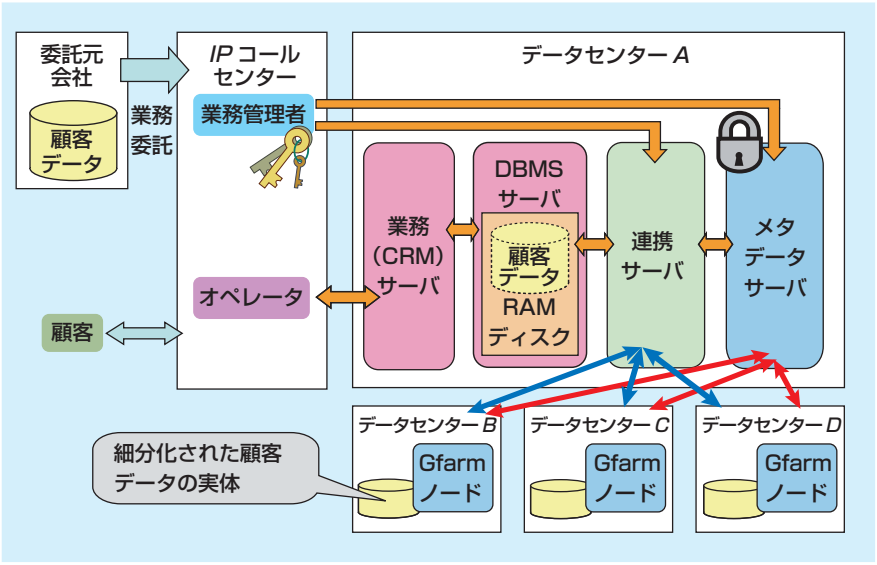

図 1　セキュアデータベースのシステム構成

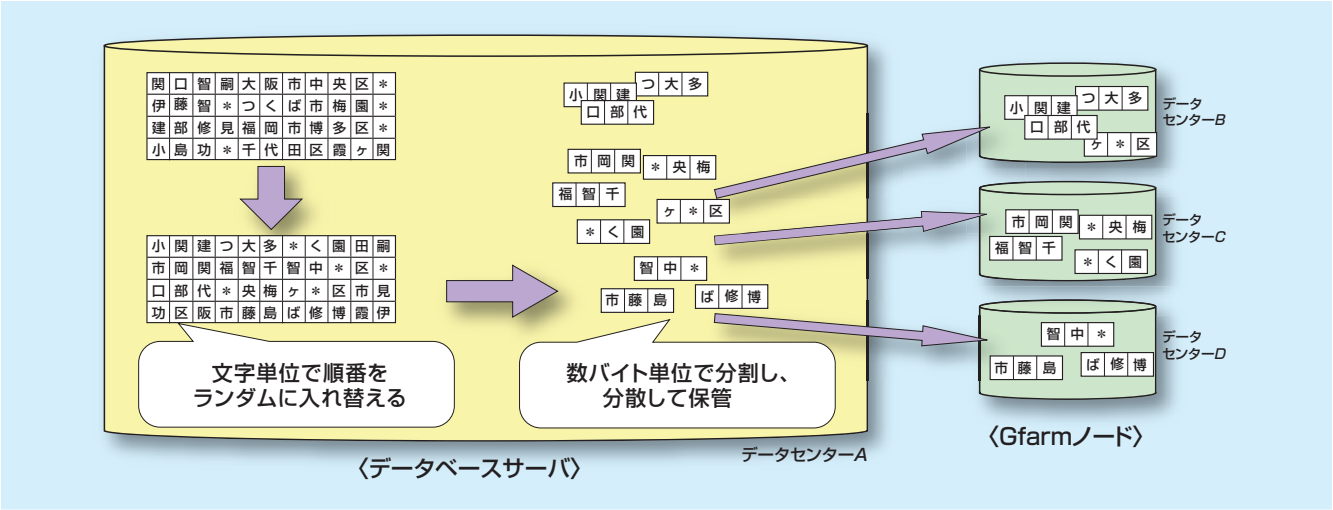

図２　情報の細分化・分散配置のイメージ

機能を強化したIPコールセンターサービスを提供している。ヘルプデスクなどのコールセンター業務では、顧客情報の漏洩防止がきわめて重要な課題となっており、顧客情報を利害関係のない第三者のデータセンターで管理することによってセキュリティを高めていた。この共同開発では、Gfarmを利用した分散ファイルシステム上で顧客情報を断片化して管理することにより、さらなるセキュリティ強化の実現を目指した。開発したシステムの構成を図に示す。

システムの詳細

コールセンターには、顧客情報を納めたデータベースがあり、その上でCRM（Customer Relationship Management）など業務アプリケーションが稼動している。従来は、アウトソーシング契約の終了後にも顧客情報がディスクやそのバックアップなどに残ることがあり得るため、情報の漏洩の可能性があった。しかし、今回の開発では、顧客情報などの重要情報のデータベースをDBMS（データベース管理システム）サーバ内のメモリ上で一時的に構成することで、電源断などで確実に消去される環境を実現した。実際の顧客情報はDBMSサーバから物理的に離れたデータセンター内に保持し、業務の開始前などに動的にデータベースの再構築を行うものである。

データセンターでは、顧客情報のレコードは、文字単位で順番をランダムに入れ替えられ、個人情報として意味のない情報に加工される。その上で、数バイトの長さに分割され、分散したPCの各ディスクなど複数のストレージ（Gfamノード）のバラバラな場所に記録される。ストレージのどこにどのデータが保存されるという位置情報（メタデータ）は、Gfarmメタサーバという別のサーバに保存されているので、データが格納されているハードディスクを持ち出せても、元の顧客情報を再現することは不可能である。さらに、この位置情報を暗号化することで、正規の暗号鍵を持つ業務管理オペレータでないと位置情報それ自身も解読できず、高度なデータ保護が実現できる。また、業務処理中に発生したデータベースの更新は、連携サーバがGfarmファイルノードに随時反映するので、故障などによってメモリデータが消去されても情報の紛失は起こらない。

その他、業務管理オペレータの認証におけるGSI（Grid Security Infrastructure）の利用、連携サーバアクセスにおけるWebサービスプロトコルの利用など、グリッドの標準技術を採用しており、将来の拡張や相互運用性にも配慮している。

システムの高性能化に向けて

今回の開発によって、顧客情報など重要情報の漏洩を二重、三重に防止することが可能となり、NTTネオメイトのIPコールセンターサービスのセキュリティ強化が実現できた。すでにこれを導入する企業も決まり、運用システムの構築が始まっている。

今後、取り扱い可能なレコード数の増加、データベース再構築に要する時間の短縮、メタデータサーバの二重化など、システムの性能向上を目指した共同開発を計画している。

関連情報：
- Gfarm:http://datafarm.apgrid.org/
- プレス発表 2005 年 4 月 8 日：http://www.aist.go.jp/aist_j/press_release/pr2004/pr20040408_2/pr20040408_2.html