東京エレクトロンデバイス

超並列処理技術を実装して費用対性能を最大化した **DWH用データベースエンジン [Greenplum Database]**

基盤はオープンソーステクノロジー

「Greenplum Database」は、オ ープンソーステクノロジーを基盤と した次世代型データウェアハウス (DWH) 用データベースエンジン である。PostgreSQLにシェアー ド・ナッシング・アーキテクチャを 拡張することで、複数の PostgreSQLデータベースインスタ ンス間の通信を可能にし、システム を1つの論理データベースとして動 作させる。そのデータベースに対す る並行クエリ実行や並列データロー ドで、大規模なDWHを構築する。 主な特長は次のとおりである。

・クエリとロードの並列化:既存 DWHと比較して10~100倍のクエ リ、ロード性能が可能。

- ・高い拡張性:ユーザーの20%が 100 テラバイト以上の DWH を構 築/サーバノード1台単位でシステ ム拡張が可能/ソフトウェア特有の 柔軟なシステム構成。
- ・業界標準の汎用的なインタフェー **ス**: SQL92/99/2003OLAP、JDBC、 ODBCの汎用インタフェース。
- ・複雑なSQL・処理に対応:デー タマイニングや機械学習(Perl、 Python、R) に対応/MapReduce 機能によりアプリケーション処理も 分散並列処理が可能。
- ・汎用的なハードウェアを選択可 能:汎用IAサーバ + Linux、UNIX を利用可能/既存のDWHアプライ アンスより数倍~数十倍の高いコス トあたり性能を実現。

メント・ホストを組み合わせて、単 一データベースとして稼働する。 オークションを手がけている米国の eBay社は、Greenplum Databaseに より、6.5ペタバイトにもおよぶ世界 最大規模のDWHを構築した。同 DWHは、eBayサイトの全てのアク セスログを収集しており、その量は 1日あたり18テラバイト(1,300億レ コード)に達している。

お問い合わせ先

ペタバイトクラスまで拡張が可能

図1は、Greenplum Databaseア

ーキテクチャである。マスタ・ホス

トは、コネクションの受付、SQLの

処理、ワークロードの配布、結果の

回収・表示、グローバル・システ

ム・カタログ管理等を行い、そして ユーザーは、マスタ・ホストを介し

てセグメント・ホストにアクセスす

る構成になる。一方セグメント・ホ

ストはユーザーデータの格納を行い、

ユーザー定義のテーブル等セグメン

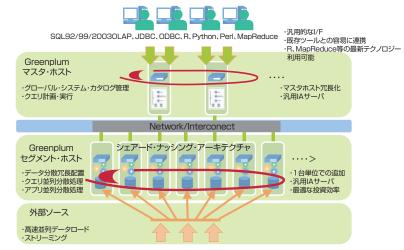
ト数に応じて配備され、複数のセグ

ネット通信販売やインターネット

東京エレクトロン デバイス(株) CN事業統括本部CNプロダクト本部 プロダクト推進部 担当:二宮

TEL: 03-5908-1984

E-mail: ninomiya.j@teldevice.co.jp URL: http://cn.teldevice.co.jp/



Greenplum Database アーキテクチャ

ビジネスコミュニケーション 2010 Vol.47 No.3