

Calpont InfiniDB® 概要

Release 3.5.1
Document Version 3.5.1-1
December 2012





Copyright © 2012 Calpont Corporation. All rights reserved.

InfiniDB および Calpont 製品名は、Calpont の登録商標です。他社およびその製品への参照は、各社が所有する登録商標を使用しており、参照のみを目的としています。

この文書の情報は予告なしに変更される場合があります。この文書にいかなる誤りがある場合も、Calpont に責任はないものとします。

ソフトウェアライセンス

「GNU Free Documentation License」の条項のもと、この文書をコピー、配布、修正する権限が付与されます。Free Software Foundation によって発行されたバージョン 1.3 以降には、変更のない項、表紙、裏表紙が含まれていません。ライセンスのコピーについては「GNU Free Documentation License」の項に記載されています。

目次

	はじめに	i
	対象読者	i
	表記規則	ii
	マニュアルリスト	iii
	マニュアルの入手	iii
	マニュアルへのフィードバック	iii
	追加リソース	iii
第 1 章	分析用データベース InfiniDB の概要	1
	分析用データベースの概要	1
	InfiniDB の主な機能	4
	サポートするハードウェアとソフトウェア	7
第 2 章	InfiniDB のアーキテクチャ	9
	InfiniDB の概要	9
	InfiniDB のコンポーネント	12
	ユーザーモジュール	12
	パフォーマンスモジュール	14
	InfiniDB のストレージ概念	17
	InfiniDB データの一般的なストレージ概念	17
	圧縮とリアルタイム解凍	18
	バージョンバッファおよびファイル	19
	トランザクションログ	19
	エクステンタマップ	20
	エクステンタマップの機能	20
	I/O 処理およびワークロードの分散	22
第 3 章	データベースの管理	25
	SQL インタフェース	27
	InfiniDB インスタンスの起動および停止	27

セキュリティの概念	27
オペレーティングシステムのセキュリティ	27
データベースのログインアクセス権	28
オブジェクトのセキュリティ	28
InfiniDB 構成ファイル	28
新しいデータベースの作成	29
管理ユーティリティ	30
システムログ	31
バックアップおよびリカバリ	31
バックアップ手順の概要	31
リカバリ手順の概要	34
GUI ツールのサポート	35

第 4 章	オブジェクト管理	37
--------------	-----------------	-----------

表	37
表の変更および削除	38
ストアドプロシージャおよび SQL 関数	38
サポートされていないデータベースオブジェクト	40
InfiniDB システムカタログ	40

第 5 章	データのロードおよび操作	43
--------------	---------------------	-----------

InfiniDB のバルクローダー	43
シンプルバルクロードの例	44
複数の表のバルクロード	46
従来のバルクロードの例	46
データロード用の MySQL ユーティリティ	48
DML のサポート	49
トランザクション管理	50
同時実行の制御	51

第 6 章	アプリケーションおよびビジネスインテリジェンスの開発	53
--------------	-----------------------------------	-----------

アプリケーションツールおよび BI ツールを使用した InfiniDB への 接続	53
アプリケーション開発ツールおよびビジネスインテリジェンス ツール	54

第 7 章	パフォーマンスのチューニング	55
	監視および診断ユーティリティ	55
	アラームおよび通知	60
	SQL の監視および診断ユーティリティ	62
第 8 章	InfiniDB への移行	67
	一般的な移行の概念	67
	MySQL のデータの移行方法	69
	Oracle のデータの移行方法	69
	Microsoft SQL Server のデータの移行方法	70
付録 A	GNU Free Documentation License	71



はじめに

『InfiniDB 概要』へようこそ。本書では、分析用データベース InfiniDB の概要と、InfiniDB のアーキテクチャおよびこのデータベースの機能について説明します。このマニュアルを読み終えると、InfiniDB の理解を深め、管理および開発の両方の観点から、このデータベースを有効に利用する準備ができるようになります。

本書には、Calpont InfiniDB Enterprise にのみ有効な情報が含まれています。

対象読者

本書は、以下に示す様々な役割の読者を対象としています。

- ◆ データベース管理者
- ◆ アプリケーションおよび Web の開発者
- ◆ データの設計者
- ◆ システムおよび Web の管理者

表記規則

本書では、次の表記規則およびユーザーへの警告を使用しています。

表 1: 表記規則

項目	記述
太字	表示されたとおりに入力する文字。 例： getLogInfo と入力します この場合、 getLogInfo と入力します。
斜体	変数またはプレースホルダ。文字列を適切に置き換えて入力します。複数の単語で構成される変数はアンダースコア () で連結して表示されています。 例： <i>ID</i> を入力します。 ID 番号 34878 を入力します。 <i>IP_address</i> を入力します。 IP アドレス 110.68.52.01 を入力します。

表 2: ユーザーへの警告

項目	記述
	注意：役立つ情報であることを示します。
	警告：データの損失または破損の原因となるハードウェアやソフトウェアのエラーを発生させる可能性があることを示します。

マニュアルリスト

Calpont InfiniDB のマニュアルは、様々な読者を対象とした複数のガイドで構成されています。次の表を参照してください。

表 3: マニュアル

マニュアル	記述
『Calpont InfiniDB 管理者ガイド』	Calpont InfiniDB を管理するための詳細な手順について説明します。
『Calpont InfiniDB 最小推奨仕様ガイド』	Calpont InfiniDB の実装に必要なハードウェアおよびソフトウェアの最小の推奨仕様を示します。
『Calpont InfiniDB インストールレーションガイド』	分散構成に Calpont InfiniDB をインストールするために必要な手順の概要について説明します。
『Calpont InfiniDB SQL 構文ガイド』	Calpont InfiniDB に固有の構文について説明します。
『Calpont InfiniDB パフォーマンスチューニングガイド』	分析用データベース InfiniDB をパラレル化および拡張するためのチューニングに役立つ情報について説明します。
『Calpont InfiniDB マルチ UM 同期ガイド』	Calpont InfiniDB で 2 つ以上のユーザーモジュールの同期を保持するために使用するオプションの概要について説明します。

マニュアルの入手

英語版のマニュアルは、(<http://www.infinidb.org/> および <http://www.calpont.com>) で入手することができます。追加の支援が必要な場合は infinidb_doc@ashisuto.co.jp にご連絡ください。

マニュアルへのフィードバック

マニュアルの改善に向けて、フィードバック、コメントおよび提案をいただけますようお願いいたします。マニュアル名、バージョンおよびページ番号を添えてコメントを infinidb_doc@ashisuto.co.jp にご送付ください。

追加リソース

Calpont InfiniDB のインストールおよびチューニング、または Calpont InfiniDB を使用したデータの問合せに関して支援が必要な場合は infinidb_doc@ashisuto.co.jp までご連絡ください。

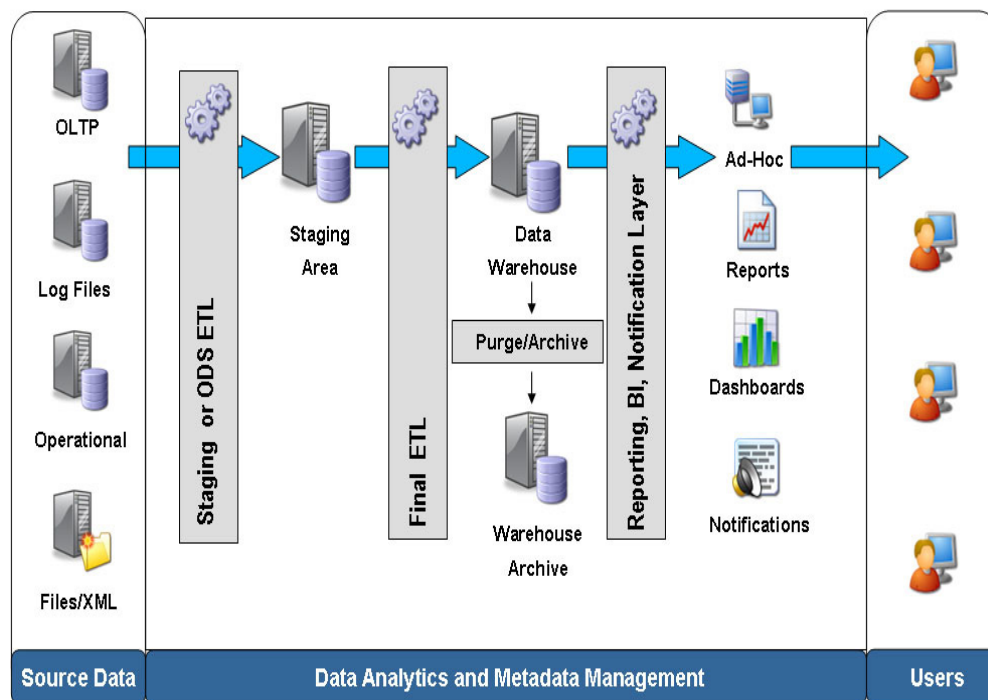
分析用データベース InfiniDB の概要

分析用データベース InfiniDB は、大規模な収集情報の管理および取得に特化して設計されたソフトウェアソリューションで、通常、データウェアハウス、データマート、分析用データベースまたは分析用データストアと呼ばれているものです。従来のリレーショナルデータベースは分析操作またはビジネスインテリジェンス操作ではなくトランザクション処理操作に設計されていますが、InfiniDB は、この従来のリレーショナルデータベースに見られる制約をなくすように設計されています。InfiniDB のマルチスレッドの列指向アーキテクチャによって、安価な汎用ハードウェアを利用し、マシンに CPU および RAM を追加した拡張構成でテラバイトレベルのデータを効率的に管理することができます。また、従来のリレーショナルデータベース管理システム (RDBMS) では完了するのに時間がかかっていた問合せの応答時間を大幅に短縮することもできます。さらに、エンタープライズ版の InfiniDB では、大規模パラレル処理 (MPP) 方法で展開するために、スケールアウト設計を実現する有償の製品版オプションも用意しています。この構成では、すべての参加マシン間で処理がパラレル化されるため、実際よりも多いノードがシステムに追加されたかのようにパフォーマンスをリニアに向上させることができます。

ビジネスインテリジェンスおよびデータ分析は、日々の重要な意思決定および競争力の維持を目的として使用するビジネスの主要なメカニズムとなっているため、Calpont (InfiniDB のプロバイダ) は多くの方にこの強力な分析用データベース機能をご利用いただけるものと確信しています。そのため、Calpont は分析用データベース InfiniDB の無償オープンソース版 (InfiniDB コミュニティ版) を提供します。誰でも GPL ライセンスのもと無償でダウンロードして使用できます。Calpont は高度な機能を提供し、技術サポートを週 7 日、24 時間体制で行う分析用データベースの製品版 (エンタープライズ版) も提供しています (日本でのサポートは 9:00-17:00 です (2013 年 4 月現在))。

分析用データベースの概要

ビジネスインテリジェンス (BI) およびデータ分析は、ビジネスの重要な意思決定を行うための情報の円滑な利用に関係します。通常、BI およびデータ分析のライフサイクルは次に示すとおりです。



この図では、業務データおよびトランザクションデータが抽出 - 変換 - ロード（ETL）フェーズに入り、最終的に分析用データベースまたはデータウェアハウスに移入される仕組みを示しています。その後、分析用データベースは様々なビジネスインテリジェンス処理で使用され、非定型の問合せ、事前にビルドされたダッシュボードとレポート、または主要なデータ変更が発生したときに自動送信される通知を介してデータにアクセスする意思決定者をサポートします。

現在提供されているほぼすべてのリレーショナルデータベースは、オンライントランザクション処理（OLTP）のワークロードに対処することを主な目的として設計されています。通常、トランザクション（Amazon など、Web ベースで本を販売する業者から本をオンラインで注文する行為など）は、リレーショナルデータベースの 1 つ以上の行に対応し、大部分の RDBMS の設計は行単位のパラダイムに基づいています。トランザクションベースのシステムの場合、このアーキテクチャは受信データの入力を処理するのに適しています。

ただし、読取りが非常に集中していて、リクエストされる情報に対して選択問合せを実行するアプリケーションの場合、OLTP データベースの設計は一般的に適していません。トランザクションは行ベースですが、ほとんどの場合、データベースの問合せは列ベースです。トランザクションデータの挿入と削除は行ベースのシステムで適切に処理できますが、表の少数の列のみを対象とする選択問合せは列指向アーキテクチャの方がはるかに適切に処理できます。平均的に、行ベースのシステムは、列ベースのデータベースで同じ情

報を取得する場合に比べて 5 倍から 10 倍の物理 I/O を実行します。問合せで最も時間がかかる処理は通常は物理 I/O であること、また分析問合せでは一般的なトランザクションでのデータベース操作よりはるかに多い行数のデータを処理することを考えると、行指向アーキテクチャと列指向アーキテクチャのパフォーマンスの差は、データベースが拡大するにつれて広がることがよくあります。

選択問合せの効率を低下させないため、行ベースの RDBMS は索引付け、水平パーティション化、マテリアライズドビュー、サマリー表、パラレル処理を活用しています。これらの技法はすべて、集中的な問合せには有益ですが、それぞれに欠点もあります。たとえば、索引付けは確かに問合せをより迅速に完了させるために役立つ場合がありますが、それにはより多くの格納領域を必要とし、挿入、更新、削除、バルクロードの操作の妨げになり（元になる表だけでなく索引も保持する必要があるため）、断片化が進むとパフォーマンスが低下する可能性があります。さらに、ビジネスインテリジェンスと経営分析の場合は、非定型の問合せが実行されるため、どの列に索引を付ける必要があるか予測することはほぼ不可能です。そのため表に索引を付けすぎるか（そして負荷と保持の問題を引き起こすか）、または適切に索引を付けられず、多くの問合せでは期待する時間よりも終了がかなり遅くなります。

特に分析用に設計された列指向データベースは、行ではなく列に基づいてデータを保存、管理、照会することで、従来型の RDBMS の制約をなくします。行全体ではなく、問合せで指定された必要な列だけにアクセスするため、I/O アクティビティも全体的な問合せ応答時間も大幅に削減されます。その結果、中程度の情報量（数十または数百 GB）でも大量のデータ（数 TB）に対しても、標準的な RDBMS に必要な時間の何分の 1 かの時間で照会し問合せ結果を戻せます。

従来の行ベースの RDBMS ではなく分析用に設計されたデータベースを選択することで得られる明らかなメリットがいくつかあります。列指向データベースについてのレポートで、Philip Howard 氏（Bloor Research のリサーチディレクタ）は次のように述べています。

「列のほうが高コストであり、かつ少ない占有領域で、優れたパフォーマンスを実現する。問合せのパフォーマンスに深く関心を持っている企業が、列ベースのソリューションをなぜ検討しないのか理解し難い。」¹

1. Philip Howard、『What's Cool About Columns』、Bloor Research、2008 年 3 月。

InfiniDB の主な機能

分析用データベース InfiniDB は、読取りのみの Web アプリケーション、データマート、データウェアハウス、その他のビジネスインテリジェンス、分析、または問合せに重点を置いた使用事例など、集中的な読取りが発生する環境でパフォーマンスを向上させることを目的とした、様々な機能を提供します。

InfiniDB は無償のオープンソースソフトウェア（コミュニティ版）のデータベースを提供しています。これは、単一サーバー上で動作し、マルチ CPU やマルチコアを搭載する今日のマシンのメリットを活用して、問合せと書き込みを並行してマルチスレッド方式で処理できるように設計されています。オープンソース InfiniDB サーバーの主な機能は次のとおりです。

- ◆ **列指向アーキテクチャ**：InfiniDB では、従来の行ベースの形式ではなく、列ベースのアーキテクチャを採用しています。そのため、InfiniDB は、200-300GB から数 TB もの情報を、行指向データベースより高速に任意の場所から読み取る必要がある問合せに適しています。
- ◆ **マルチスレッド設計**：InfiniDB はマルチスレッド設計であるため、今日の新しいマルチ CPU やマルチコアベースのハードウェアを活用できます。CPU やコアの数が増えるほど InfiniDB のパフォーマンスは向上し、アプリケーションを修正する必要はありません。
- ◆ **垂直および水平の自動パーティション化**：列指向の InfiniDB は、垂直パーティション化を透過的に使用してデータを格納し、問合せに必要な列だけを読み取ります。ただし、ストレージの特別な配置や設計を必要とせずに、論理的な水平レンジパーティション化も使用できます。垂直パーティション化と論理的な水平レンジパーティション化の両方を使用することで、InfiniDB はどちらの方向（列と行）でも I/O を削減できます。垂直および水平のパーティション化は、いずれも InfiniDB データベースによって自動的に処理されるため、ユーザーの介入は不要です。
- ◆ **高レベルの同時実行機能**：同時実行に関しては、InfiniDB はサーバーマシンの処理能力によってのみ制限され、理論上の制限はありません。
- ◆ **高速データローダー**：InfiniDB は従来型の挿入、更新、削除処理をサポートしていますが、大量のデータをロードするには高速ロードユーティリティを利用できます。InfiniDB のローダーを使用すると、TB 単位のデータを従来型の方法よりはるかに高速に InfiniDB データベースにロードできます。
- ◆ **DML のサポート**：データの高速バルクロードのサポートに加えて、InfiniDB は DML（挿入、更新、削除）操作もすべてサポートします。
- ◆ **トランザクションのサポート**：InfiniDB データベースは ACID 準拠のトランザクションをサポートします。トランザクションは簡単にコミットまたはロールバックすることができ、競合の解決手段としてデッドロック検出もサポートします。
- ◆ **クラッシュからのリカバリ**：InfiniDB にはクラッシュからのリカバリ機能があります。システムクラッシュが発生すると、InfiniDB は自動的にデータの整合性を保持し、システム再起動時にロールフォワードとロールバックの操作によってデータベースを一貫した状態に戻すようサポートします。

- ◆ **MVCC 設計** : InfiniDB は多版型同時実行制御 (MVCC) つまり「スナップショット読み取り」機能をサポートしているため、問合せ操作が表でブロックされることはありません。表は常に、その問合せを発行したときに存在していたデータを参照します。
- ◆ **索引付けが不要** : InfiniDB は垂直パーティション化と論理的な水平パーティション化の両方を透過的に使用するため、索引付けは不要です。本質的に、列指向データベースのデータが索引になります。また、InfiniDB はエクステントマップという小さい重要な構造を自動的に保持して、I/O を削減するために利用します。エクステントマップによって手動によるデータのパーティション化も不要になります。
- ◆ **保守の軽減** : 表の索引付けが不要になるほか、InfiniDB には、マテリアライズドビュー、サマリー表などの、パフォーマンスを向上させるためのオブジェクトも不要です。これは、一般的なデータベースで必要とされる管理と保守の作業を不要にし、システム全体の複雑さを軽減するためにも役立ちます。
- ◆ **ディクショナリの文字列の圧縮** : InfiniDB はトークンディクショナリ方式によって 9 バイト以上の固定長文字列と 8 バイト以上の可変長文字列を管理します。列ファイルには、長い文字列を含む個別のディクショナリファイルへのトークン (ポインタ) が格納されます。表内で重複した文字列値はディクショナリの 1 つの場所に、共有文字列エントリを指す重複したトークンとともに格納されます。
- ◆ **表定義変更のサポート** : ALTER TABLE コマンドを使用して表に対して列を追加および削除することもでき、オンライン状態で列を追加する機能も特別にサポートされています。
- ◆ **パフォーマンス診断** : パフォーマンスチューニングを支援するため、InfiniDB には監視と診断のユーティリティが付属しています。これは、ユーザーがデータベースを監視し、実行が遅い SQL のトラブルシューティングを行うために役立ちます。
- ◆ **MySQL フロントエンド** : InfiniDB はフロントエンドに MySQL を利用します。そのため、MySQL に精通したユーザーであれば、InfiniDB で、すぐに生産性を向上させることができます。MySQL をよく知らないユーザーも、MySQL は ANSI が定める標準 SQL 操作のほぼすべてをサポートしているため、容易に習得できます。さらに、MySQL と他のベンダーによって無償の GUI ツールが多数提供されており、InfiniDB データベースの開発と管理に使用できます。

InfiniDB サーバーのオープンソース版の他に、Calpont は製品版のデータベースも提供しています。InfiniDB Enterprise では、次に示す拡張機能を提供する、有償のスケールアウトオプションが追加されます。

- ◆ **圧縮とリアルタイム解凍** : InfiniDB は優れた圧縮機能を提供するとともに、ディスクからの読み取りのパフォーマンスを最適化する圧縮方法を採用しています。これによって、I/O バウンドのシステムで、ディスクからの読み取りのパフォーマンスを向上させることができます。圧縮はデフォルトで有効になっていますが、インスタンスレベル、表レベル、列レベルで有効と無効を切り換え

ることができます。列指向のストレージでは、類似したデータが各列ファイルに格納されるため、優れた圧縮性が実現されます。

- ◆ **パーティションの無効化とドロップ**：InfiniDB は、DBA がパーティションにあらかじめ名前を付けたり定義したりする必要がないように、自動パーティション化を採用しています。パーティションをドロップするときは、`show partitions` コマンドによって、データの各パーティションに存在する値の範囲を表示します。業務に関連する行がなくなったパーティションは、無効にしてデータを論理的に削除するか、またはドロップして関連するファイルとともに削除します。
- ◆ **ユーザー定義関数 (UDF)**：InfiniDB では、InfiniDB エンジン内に統合された関数として実行できる、完全な並行分散型の UDF を作成できます。これによって、InfiniDB の統合マッピング機能をもっと最大限に活用したカスタム関数を作成し、分散レイヤー内の利用可能なすべてのコアにカスタマイズした機能を分散できます。
- ◆ **大規模パラレル処理 (MPP) 対応**：InfiniDB は複数の汎用ハードウェアマシンを使用して、総合的なパフォーマンスをリニアに向上させることができます。InfiniDB 構成に安価なハードウェアを追加してデータベースの処理能力を向上させることが可能で、新しいノードを追加するだけで応答時間を半分に短縮できる場合もあります。これによって、増大するデータ量への対応、増加するユーザーアクティビティへの対応、または総合的なパフォーマンス目標達成のための調整が容易になります。さらに、InfiniDB システムを停止またはオフラインにする必要がない、動的な方法で新しいノードを追加できます。
- ◆ **シェアードナッシング分散データキャッシュ**：複数ノードによる InfiniDB 構成では、データは様々なノードとそのデータキャッシュに分散されます。ノード間でデータを共有することはありませんが、問合せのためにデータを読み取るときは、InfiniDB MPP アーキテクチャ内のすべてのノードがアクセスされます。つまり、InfiniDB は、すべての参加ノードが並行して分散形式でアクセスする、1 つの大きな論理データキャッシュを作成すると考えられます。これによって、十分な数のノードと豊富なメモリーがあれば、InfiniDB は文字通り、大規模なデータベースをキャッシュできます。
- ◆ **自動フェイルオーバー**：InfiniDB のアーキテクチャには、ユーザー問合せの管理と実際のデータベース I/O 操作を処理するソフトウェアモジュール用に、組み込みの自動フェイルオーバー機能が含まれています。
- ◆ **同時実行の自動スケールアウト**：構成にノードを追加すると、自動的に 1 台以上のマシンに同時実行をスケールアウトできます。
- ◆ **ソフトウェアパッチの自動管理**：複数の InfiniDB サーバーにパッチまたはアップグレードを適用する必要がある場合、自動処理がアップグレードが適用された最初のノードからソフトウェアを取得し、他のすべての参加ノードを自動的にアップグレードします。

ここで説明したすべての機能が、分析用データベースやデータウェアハウスを設計し実装するための専門知識を持たない初心者でも、高速に動作するデータマート、データウェアハウス、その他の分析用データベースを作成し管理できるように支援します。

サポートするハードウェアとソフトウェア

InfiniDB データベースは次のハードウェアとソフトウェアによる構成をサポートします。

- ◆ Intel および AMD の汎用ハードウェア、64 ビット
- ◆ 次の Linux
 - Red Hat Enterprise
 - CentOS
 - Debian
 - Ubuntu
 - AWS (Amazon) EC2

InfiniDB のアーキテクチャ

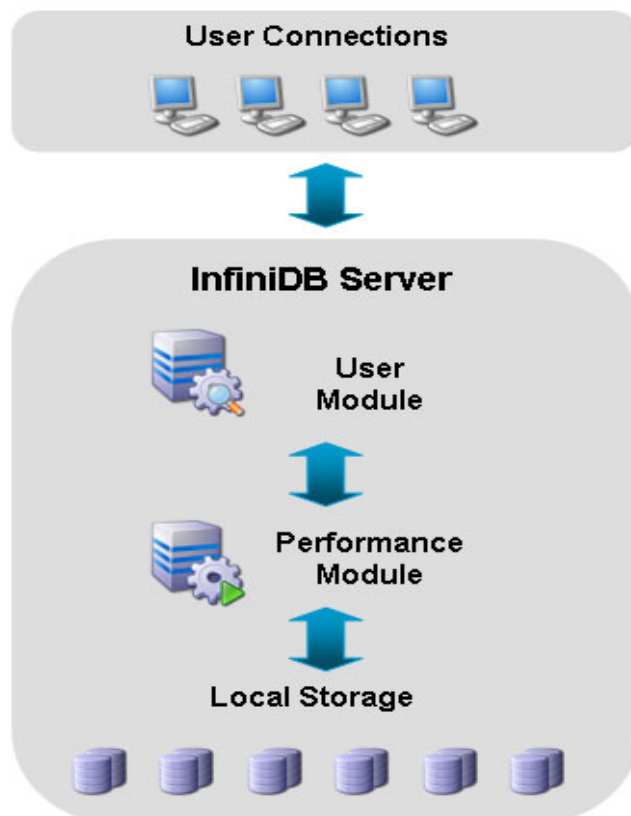
この項では、InfiniDB のアーキテクチャおよびデータベースを構成する様々なコンポーネントについて説明します。

InfiniDB の概要

InfiniDB のアーキテクチャは様々なコンポーネントで構成されており、すべてのコンポーネントが連携して InfiniDB データベースを構成しています。これらのコンポーネントには次のものがあります。

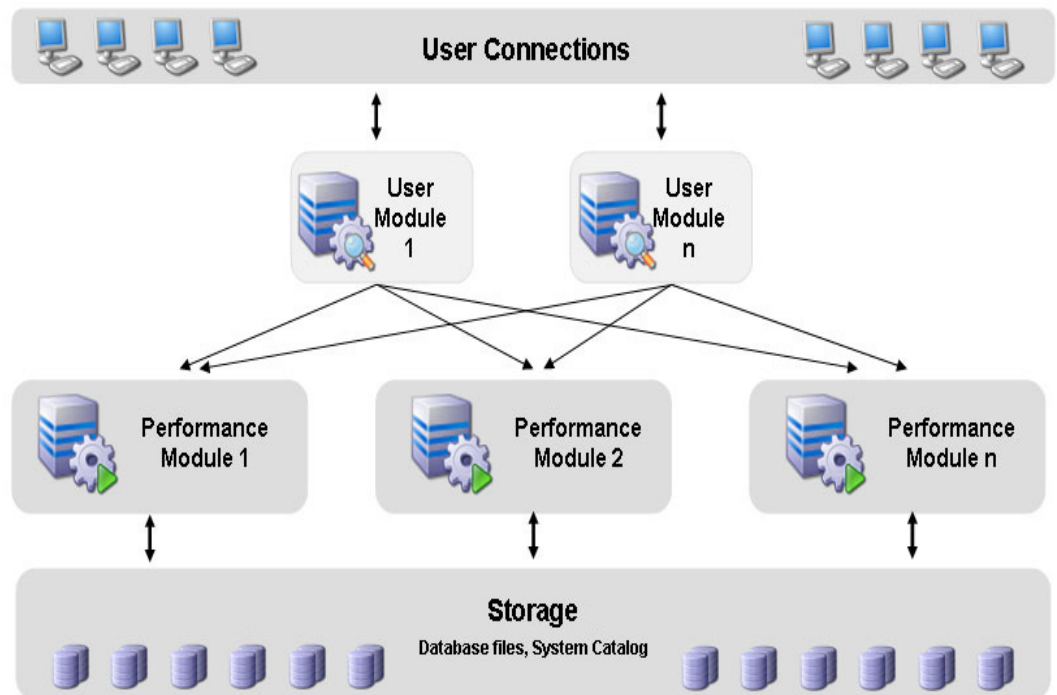
- ◆ **ユーザーモジュール**：ユーザーモジュールは、小さい MySQL インスタンス（個別にディレクタモジュールと呼ばれることもある）、および同時実行スケールリングを処理する多くの InfiniDB プロセスで構成されています。また、ユーザーモジュールは、SQL リクエストを分割して 1 つ以上のパフォーマンスモジュールに分散します。パフォーマンスモジュールによって、リクエストされたデータがメモリーキャッシュまたはディスクのいずれかから実際に取得されます。最後に、ユーザーモジュールは、関連するパフォーマンスモジュールからのすべての問合せ結果をまとめ、包括的な問合せ結果セットを形成してユーザーに返します。InfiniDB Enterprise の構成には複数のユーザーモジュールが存在する可能性があることに注意してください。この構成では、高可用性に対応した構成および問合せのワークロードバランシング用の方法を実現することができます。
- ◆ **パフォーマンスモジュール**：パフォーマンスモジュールは、データの格納、取得および管理、問合せ操作に対するブロックリクエストの処理、および問合せリクエストをファイナライズするためにユーザーモジュールへのデータの引渡しを行います。パフォーマンスモジュールは、ディスクからデータを選択して、パフォーマンスモジュールが存在するサーバーの一部であるシェアードナッシング方式のデータキャッシュにキャッシュします。InfiniDB Enterprise は、ユーザーが必要な数のパフォーマンスモジュールを構成できるようにすることで MPP のメリットを実現しています。追加された各パフォーマンスモジュールによって、データベース全体にキャッシュが追加され、処理能力が高まります。
- ◆ **ストレージ**：InfiniDB では、ローカルストレージまたは共有ストレージ（SAN など）のいずれかを使用してデータを格納できます。ユーザーは、1 つのみのサーバーを InfiniDB サーバーとして動作させ、すべてを構成してそのサーバー上で実行するか、または複数のサーバーにスケールアウトして、シェアードナッシングのアーキテクチャを構成できます。

InfiniDB データベースには、設計可能な多くの構成があります。オープンソースの InfiniDB コミュニティ版のデータベースは次のように示すことができます。



前述の構成では、すべてのモジュールが同じサーバーで動作し、マシン上のすべての CPU 処理能力を活用してスケールアップした処理能力を利用できます。

MPP スケールアウトオプションを使用する製品版の InfiniDB Enterprise には、設計可能な多くの構成があります。その 1 つが共有ディスクアーキテクチャです。



前述の構成では、同時実行スケールアウトはユーザーモジュールノードを追加することで実現されていますが、I/O スケールアウトはパフォーマンスモジュールノードを追加することで実現されています。

ユーザーモジュールとパフォーマンスモジュールは互いに近接した場所にある場合もあれば、異なるデータセンターおよび地理的な場所に分けられている場合もあることに注意してください。

InfiniDB のコンポーネント

この項では、InfiniDB データベースの主要な各コンポーネントおよびデータベースの動作全体に影響する関連構造について詳しく説明します。

ユーザーモジュール

簡単に説明すると、InfiniDB ユーザーモジュールは、エンドユーザーの問合せ操作の管理および制御を行います。各問合せの状態を維持し、問合せの代わりに1つ以上のパフォーマンスモジュールにリクエストを発行して作業を行い、すべての関連パフォーマンスモジュールの結果セットを集計して最終的にエンドユーザーに返す結果セットにする問合せを実行します。

ディレクタモジュールおよび MySQL インタフェース

InfiniDB ユーザーモジュールの個別のコンポーネントはディレクタモジュールと呼ばれる場合があります。基本的に、これは InfiniDB が次の基本機能の実行に利用する MySQL インスタンスです。

- ◆ 接続の検証
- ◆ SQL 文の解析
- ◆ 一般的な SQL 計画の生成
- ◆ 最終結果セットの配付

このコンポーネント内で、InfiniDB はサーバーを管理するためのいくつかのプロセスを提供します。プロセスマネージャ (ProcMgr) はすべての InfiniDB プロセスの起動、監視および再起動を行います。また、各マシン上でプロセスモニター (ProcMon) と呼ばれる別のプロセスを使用して InfiniDB プロセスを追跡します。

InfiniDB では、MySQL 形式の問合せ計画を利用する多くの接続ライブラリを使用して問合せ計画を InfiniDB 形式に変換します。InfiniDB 形式も基本的には解析ツリーですが、ユーザーモジュールで解析ツリーを InfiniDB ジョブリストに変換するときに役立つようにオプティマイザから実行ヒントが追加されます。また、InfiniDB 用の MySQL への MySQL ストレージエンジン API (ハンドラ) インタフェースを提供する MySQL コネクタライブラリ (libcalmysql.so) もあります。

主要なユーザーモジュール操作

ユーザーモジュールは、InfiniDB データベースに対して次の主要な機能を実行します。

- ◆ InfiniDB ジョブリストへの MySQL 計画の変換
- ◆ InfiniDB システムカタログからの InfiniDB OID (オブジェクト ID) の検索の実行
- ◆ 不要なエクステンツの除外によって実現される I/O の削減のためのエクステンツマップ (詳細はこのマニュアルの後半を参照) の検査
- ◆ パフォーマンスモジュールへの指示 (「プリミティブ操作」とも呼ばれる) の発行
- ◆ 必要に応じたハッシュ結合の実行 (結合での小さい方の表のサイズに基づく)。処理が必要なハッシュマップをパフォーマンスモジュールに送信することによる、分散されたハッシュ結合の管理の支援
- ◆ ハッシュ結合の後に発生する表をまたいだ関数および式の実行
- ◆ パフォーマンスモジュールからのデータの受信および必要に応じたパフォーマンスモジュールへの返信
- ◆ すべての集計処理および重複行の削除処理に対する補足ステップの実行
- ◆ MySQL インタフェースへのデータの返信

ユーザーモジュールの主要なジョブは同時実行スケーリングの処理です。データベースファイルに対して直接処理が行われることはないため、データベースファイルの可視性も必要とされません。マシンの RAM を一時的に使用して部分的な問合せ結果を集計し、最終的にユーザーに返される包括的な応答にします。この時点で、問合せは 1 つのユーザーモジュールにのみ存在しますが、利用可能なすべてのパフォーマンスモジュールに問合せをブロードキャストして実際の I/O 動作をスケールアウトできます。

ユーザーモジュールのプロセス

ユーザーモジュールには、実行マネージャ (ExeMgr)、DML マネージャ (DMLProc)、DDL マネージャ (DDLProc) およびインポート分散マネージャ (cpimport) という、いくつかのプロセスが含まれています。ExeMgr は、TCP/IP ポートでディレクトタモジュール (MySQL インスタンス) の問合せ解析ツリーをリスニングします。ExeMgr は、問合せ解析ツリーを InfiniDB ジョブリスト (問合せへの応答に必要な一連の指示) に変換します。ExeMgr は、問合せ解析ツリーを移動しながら、ジョブステップを繰返し生成してジョブリストの最適化および再最適化を行います。ジョブステップの主要なカテゴリは、列フィルタの適用、表結合の処理および返された列の投影です。DMLProc および DDLProc は、適切なパフォーマンスモジュールに DML および DDL を分散します。cpimport は、ユーザーモジュールで実行されている場合に、パフォーマンスモジュールにソースファイルを分散します。

計画内の各ノードは、ジョブリスト自体によって並行して実行されます。ノードは、ユーザーモジュール全体、パフォーマンスモジュール全体またはこれらの組合せに対

して動作できます。各ノードでは、エクステントマップを使用して、作業指示の送信先となるパフォーマンスモジュールを決定します（この使用方法の詳細は、後述のエクステントマップに関する説明を参照してください）。これらのすべてのノードはユーザーレベルのスレッドです。各スレッドで、データおよび利用可能なシステムリソースの配分に基づいて追加のスレッドを生成できます。

ユーザーモジュールのスケールアウト

エンタープライズ版の InfiniDB では、透過的なユーザーモジュールのスケールアウトを実現できます。管理者は、システムをスケールアップしてより多くのユーザー接続をサポートできるように、複数のユーザーモジュールが含まれるように InfiniDB システムを構成できます。

パフォーマンスモジュール

この項では、Calpont のパフォーマンスモジュールの様々な側面について説明します。

一般的な操作

パフォーマンスモジュールは、問合せ処理および書込み処理において I/O 操作を実行します。また、実行する作業に関してユーザーモジュールから指示を受信します。パフォーマンスモジュールは、問合せ自体は確認せず、ユーザーモジュールによって与えられた一連の指示のみを確認します。この点において、パフォーマンスモジュールでは、標準的な意味での問合せの実行ではなく、ブロック指向の I/O 操作の実行に重点が置かれています。

InfiniDB Enterprise の大規模パラレル処理またはスケールアウト機能は、InfiniDB 構成にパフォーマンスモジュールを追加することで実現されます。特定の設定にノードを追加すると、異なるパフォーマンスモジュール間で問合せがパラレル化されるため、パフォーマンスのリニアな向上が実現されます。パフォーマンスモジュールには、データベースの動作のスケールアウトに必要となる 3 つの重要な動作（スキャンの分散、ハッシュ結合の分散および集計の分散）があります。これら 3 つの動作を組み合わせることによって、問合せが集中する環境での MPP の動作をチューニングできます。

パフォーマンスモジュールで実行されるプロセスには、PrimProc（問合せ実行を処理するプロセス）、WriteEngineServer（パラレル書込みを調整するプロセス）および cpimport（実際のデータベースファイルの更新を実行するプロセス）の 3 つがあります。

ロード処理および書込み処理

パフォーマンスモジュールのノードは、基礎となる永続ストレージへのロードおよび書込みを実行するタスクを割り当てられています。WriteEngineServer は、各パフォーマンスモジュールで、DML、DDL およびインポートを調整します。cpimport は、どのモジュールで実行されているかを認識し、パフォーマンスモジュールで実行されている場合には、データベースディスクファイルの実際の更新を処理します。このように、InfiniDB は完全に並行なロード機能をサポートします。DDL の変更は、すべての InfiniDB メタデータを追跡する InfiniDB システムカタログ内で維持されます。InfiniDB Enterprise では、高可用性構

成をサポートするために書込みノードに対してフェイルオーバー機能が提供されています。

シェアードナッシング方式のデータキャッシュ

エンタープライズ版の InfiniDB では、複数のパフォーマンスモジュールでシェアードナッシング方式のデータキャッシュが使用されています。初めてデータへのアクセスが行われたときに、パフォーマンスモジュールはユーザーモジュールによって指示されたデータの量に基づいて動作し、後続のアクセス用に LRU ベースのキャッシュにデータをキャッシュします。InfiniDB 専用マシンでは、コンピュータの RAM の最大 85% をパフォーマンスモジュールのデータキャッシュ専用に使えます。

パフォーマンスモジュールのキャッシュはシェアードナッシング方式で設計されているため、パフォーマンス上の主要なメリットが少なくとも 2 つあります。1 つは、関連パフォーマンスモジュールノード間でデータブロックの ping が発生しないことです。これは複数インスタンスまたは共有ディスクのデータベースシステムでは発生することがあります。もう 1 つは、多くのパフォーマンスモジュールノードがシステムに追加されているため、データベースの全体的なキャッシュサイズが大幅に大きくなることです。これによって、(場合によっては) データベース全体を RAM にキャッシュして非常に高速なアクセスを行うことができます。

InfiniDB がスタンドアロンサーバーで使用される場合、データキャッシュは他の標準 LRU ベースのデータキャッシュと同様に動作することに注意してください。単一サーバーの実装では、存在するパフォーマンスモジュールは 1 つのみのため、データベースからリクエストされたすべてのデータがキャッシュを介して渡されます。

パフォーマンス向上のための機能

InfiniDB Enterprise 構成へのパフォーマンスモジュールの追加は、追加サーバー用のユーザーモジュールで一連の基本構成手順を実行すると実現されます。この処理は InfiniDB システム全体で透過的であり、通常新しいパフォーマンスモジュールノードは 60 秒以内に追加されます。

パフォーマンスモジュールのフェイルオーバー

複数ノードの InfiniDB Enterprise のパフォーマンスモジュール構成では、すべてのノードがオンラインであり、特定のパフォーマンスモジュールでエラーが発生した場合に透過的なフェイルオーバーが行われることがハートビートメカニズムによって保証されます。パフォーマンスモジュールが異常終了した場合、パフォーマンスモジュールを使用して処理中の問合せではエラーが返されます。パフォーマンスモジュールを使用していない問合せでは、動作している残りのパフォーマンスモジュールを使用して処理が続行されます。パフォーマンスモジュールでエラーが発生したためエラーを受信したユーザーは、問合せを再発行して、残りのパフォーマンスモジュールを使用して作業を実行できます。

フェイルオーバーが発生した場合、冗長データが構成されているか、基礎となるストレージアーキテクチャがシェアードエブリシング方式（EC2 EBS、SAN など）であれば、動作しているパフォーマンスモジュール間でパフォーマンスモジュールのデータブロックのマッピングが再編成され、ユーザーモジュールのエクステントマップが再評価されることにより、残りの適切なパフォーマンスモジュールに問合せが送信されます。この処理はユーザーに対して透過的であり、手動操作は必要ありません。

エラーが発生したパフォーマンスモジュールがオンラインに戻ると、InfiniDB はそのパフォーマンスモジュールを自動的に構成に戻し、作業に使用し始めます。

ユーザーモジュールおよびパフォーマンスモジュールを使用した問合せ処理の概要

InfiniDB によるエンドユーザーの問合せの処理方法のトップダウンの見解は次のとおりです。

1. MySQL インタフェースを介してリクエストを受信します。MySQL によって、リクエストの処理に必要なすべての表に対する表操作が実行され、最初の問合せ実行計画が取得されます。
2. InfiniDB で、MySQL ストレージエンジンインタフェースを使用して MySQL の表オブジェクトが InfiniDB オブジェクトに変換されます。その後、これらのオブジェクトはユーザーモジュールに送信されます。
3. ユーザーモジュールで、MySQL 実行計画が変換され、これらのオブジェクトが InfiniDB 実行計画に最適化されます。ユーザーモジュールで、問合せの実行に必要な手順およびそれらを実行するタイミングが決定されます。
4. ユーザーモジュールで、問合せに必要なデータの場所に関してエクステントマップが参照され、エクステントマップ内に含まれている情報に基づいてエクステントの除外が実行されます。
5. ユーザーモジュールで、1 つ以上のパフォーマンスモジュールにコマンドが送信されてブロックの I/O 操作が実行されます。
6. パフォーマンスモジュールで、条件のフィルタ処理、結合処理、データの初期集計が実行され、最終結果セットの処理のためにユーザーモジュールにデータが送信されます。
7. ユーザーモジュールで、最終結果セットの集計が実行され、問合せの最終結果セットが作成されます。
8. ユーザーモジュールから、ユーザーに結果セットが返されます。

InfiniDB のストレージ概念

この項では、InfiniDB データベースの基本的なストレージ概念について説明します。

InfiniDB データの一般的なストレージ概念

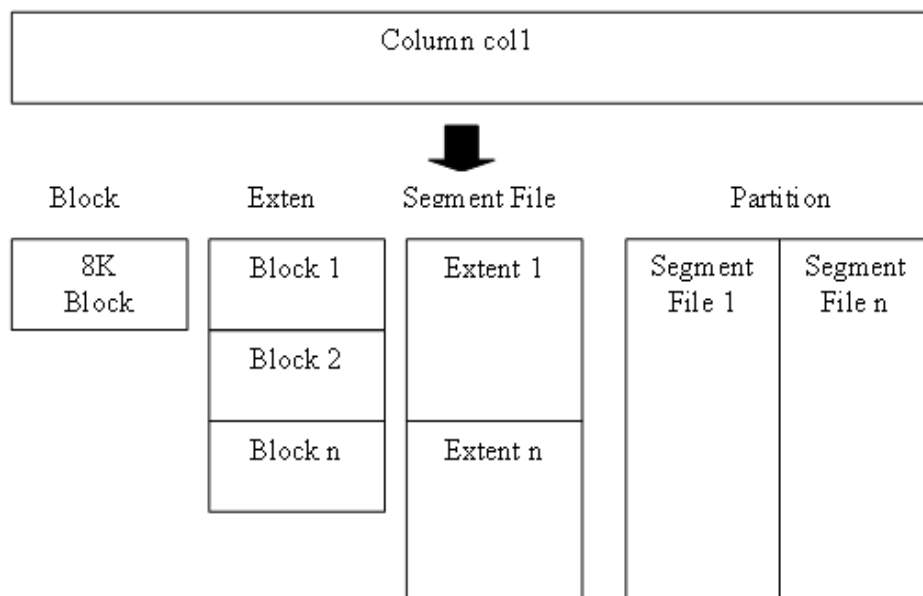
InfiniDB は列指向データベースであるため、表が作成されると、システムによって表の列ごとに 1 つ以上のファイルが作成されます。したがって、たとえば、3 つの列を含む表が作成されると、個別にアドレス指定可能な論理オブジェクトが少なくとも 3 つ、SAN またはサーバーのディスクドライブに作成されます。

次に、InfiniDB データベースのストレージビルディングブロックを下位から上位の順に示します。

- ◆ **ブロック**：ブロック（一部の RDBMS ではページとも呼ばれる）は、サイズが 8KB の物理オブジェクトです。データベース内のすべてのデータベースブロックは論理ブロック識別子 (LBID) で一意に識別されます。InfiniDB のブロックサイズは設定変更できませんが、InfiniDB の先読みメカニズムで読み取られるブロックの数はカスタマイズできます。
- ◆ **エクステンツ**：エクステンツは、1 つの物理的なセグメントファイル内に存在する 800 万行の論理的測定単位です。エクステンツは、1 バイトのデータ型では 8MB、2 バイトのデータ型では 16MB、4 バイトのデータ型では 32MB、8 バイトのデータ型では 64MB、可変サイズのデータ型では 64MB が消費されます。エクステンツが一杯になると、新しいエクステンツが自動的に作成されます。
- ◆ **セグメントファイル**：セグメントファイルは、列のデータを保持するディスク上の物理ファイルです。セグメントファイルがエクステンツの最大数に到達すると、新しいセグメントファイルが自動的に作成されます。
- ◆ **パーティション**：パーティションは、表または索引に存在するデータを水平に分割するためにパーティション化を使用する従来の行ベースのデータベースの概念とは異なる概念です。InfiniDB では、パーティションは論理オブジェクトであり、1 つ以上のセグメントファイルで構成されます。1 つのパーティションに含めることができるセグメントファイルの数は `FilesPerColumnPartition` パラメータによって制御されます。列に含めることができるパーティションの数に制限はありません。

InfiniDB のストレージ概念は、次のような図で表すことができます。

```
create table t1 (col1 int, col2 varchar(20), col3 date)
engine=infinidb;
```



ディスクのサブシステムへのセグメントファイルの配置に関して、InfiniDB ストレージに使用するディスクリソースを InfiniDB 構成ファイル内にディレクトリで指定できます。InfiniDB では、データの物理的な読取り率が最大化されるように、指定したすべてのディスクリソースにすべてのセグメントファイルが自動的に分散されます。

圧縮とリアルタイム解凍

InfiniDB は、優れた圧縮機能を提供するとともに、ディスクからの読取りのパフォーマンスを最適化する圧縮方法を採用しています。これによって、I/O バウンドのシステムで、ディスクからの読取りのパフォーマンスを向上させることができます。

圧縮はデフォルトで有効になっていますが、表レベルまたは列レベルで有効と無効を切り替えることができます。また、変数 `infinidb_compression_type` を設定してセッションレベルで制御することもできます。デフォルト設定では、リリース 2.0 より後のリリースで作成されたすべての表に対して圧縮が有効になります。以前の既存の表は変更されないことに注意してください。圧縮されていない以前の表を圧縮するには、新しい表を圧縮して作成した後、データを選択し、高速インポートメソッド (cpimport) を使用して

圧縮した新しい表にロードします。その後、表の名前を変更し、圧縮されていない表を削除できます。

列指向のストレージでは、類似したデータが各列ファイルに格納されるため、優れた圧縮性が実現されます。ほとんどのデータセットで 65-95% の領域を節約する優れた圧縮率が示されますが、実際に節約される領域は、データのランダムさおよび存在する個別値の数によって異なります。

InfiniDB の圧縮方法は、解凍速度を高速化してディスクからの読取りパフォーマンスが最大化されるようにチューニングされています。これは、通常 WORM (Write Once Read Many) であるシステムでの標準的なデータウェアハウスの使用に理想的です。ディスクからの読取りの高速化は、ベストプラクティスに従い、既存のデータ操作を最小化するシステムに適しています。初期ロード後に頻繁に更新される列については、その列が未圧縮と定義されている場合のほうが、InfiniDB では更新のパフォーマンスが向上することがあります。列の追加または削除に、列データベースの表を再構築する必要はありません。したがって、複数のオプションについて並行して評価されます。

バージョンバッファおよびファイル

InfiniDB では、データベースの問合せで一貫したビューを提供できるように、トランザクションのロールバックアクティビティを管理し、データベースの MVCC (多版型同時実行制御) または「スナップショット読取り」機能を提供するバージョンバッファという構造が使用されます。InfiniDB のすべての文は、システムによってシステム変更番号 (SCN) として参照されるデータベースの特定のバージョン (またはスナップショット) で実行されます。これはバージョンバッファと呼ばれますが、メモリ構造とディスク構造の両方で構成されています。

バージョンバッファでは、メモリー内のハッシュ表を使用して処理中のトランザクションの情報へのメモリーアクセスが提供されます。スタートアップ時の初期サイズは 4MB です。メモリー領域は、トランザクションによって変更されたブロックを処理するためにこのサイズから増えていきます。ハッシュ表の各エントリは、変更された 8K のブロックに対する 40 バイトの参照情報です。

バージョンバッファファイルは、各 DBRoot でサイズが 1GB のファイルにデフォルト設定されます。これは、VersionBufferFileSize パラメータを使用して構成できます。バージョンバッファファイルは、システムの各 DBRoot に分散されます。

トランザクションログ

トランザクションを管理し、クラッシュリカバリの状況が発生した場合にデータベースに対してロールフォワード操作を実行するために、InfiniDB ではトランザクションログが使用されます。トランザクションログには、DDL 文と DML 文のエントリおよびバルクロード操作のマーカーが含まれています。Linux での InfiniDB の現行の実装では、トランザクションログは data_mods.log という名前で /var ファイルシステムにインストールされています。

バックアップおよびリカバリの目的で、トランザクションログのアーカイブが時間ベースの間隔で実行されます。これは、InfiniDB 構成ファイル (TransactionArchivePeriod) で構成できます。各トランザクションログダンプの末尾には、マシン上で一意のファイルになるようにタイムスタンプが追加されます。トランザクションログがダンプされた後、アーカイブログは切り捨てられ、新しいエントリに対応できるようになります。

トランザクションログの情報を表示するために、ユーザーは実行された DDL 文および DML 文についてレポートする ReplayTransactionLog ユーティリティを使用することができます。バルクロード操作はトランザクションログには含まれないことに注意してください。

エクステントマップ

問合せの処理速度を高めるために、InfiniDB ではエクステントマップと呼ばれるスマート構造が使用されます。これによって、行ベースのデータベースでそのアーキテクチャによる制約をなくすために実装が必要な索引付け、手動での表のパーティション化、マテリアライズドビュー、サマリー表、およびその他の構造とオブジェクトが必要なくなります。

前述のとおり、エクステントは物理的なセグメントファイル内に存在する領域の論理ブロックで、サイズは 8-64MB の間です。各エクステントはディスク上のより少ないスペースを使用して、より小さいデータ型で、同じ数の行をサポートしています。エクステントマップには、すべてのエクステントおよびそれらに対応するブロック (LBID) がカタログ化されます。また、エクステントマップには、エクステント内の列のデータの最小値および最大値が保持されます。

InfiniDB では、問合せリクエストを処理するためにエクステントマップおよびバージョンバッファの両方が使用されることに注意してください。これら 2 つの構造の組合せはブロック解決マネージャ (BRM) と呼ばれます。

InfiniDB Enterprise 構成では、エクステントマップのマスターコピーはプライマリのパフォーマンスモジュールに存在します。システムのスタートアップ時、ファイルはメモリーに読み取られた後で、ディザスタリカバリおよびフェイルオーバーに備えて関連するその他すべてのユーザーモジュールおよびパフォーマンスモジュールに物理的にコピーされます。すばやくアクセスできるように、すべてのノードでエクステントマップがメモリーに保持されます。エクステントが変更されると、すべての参加ノードにも更新がブロードキャストされます。

エクステントマップの機能

エクステントマップを使用すると、InfiniDB で問合せへの対応に必要なブロックのみを取得できますが、それとは別に論理的レンジパーティション化というメリットもあります。これは、エクステントマップ内に含まれている各エクステントの最小値および最大値によって実現されます。InfiniDB では、まず (必要な列のみがスキャンされる) 列指向の

アーキテクチャによってエクステントの除外が実行されますが、エクステントマップで実装される論理的水平パーティション化のため、この処理は高速化されます。

この自動エクステント除外の動作は、データが頻繁にロードされ、時間に基づいて参照されることが多い一連のデータ、順序付けられたデータ、パターン化されたデータまたは時間ベースのデータに適しています。データの先頭に対する問合せではぼりアルタイムでロードすることによって、すべての日時列および昇順のキー値に適したエクステント除外を簡単に示すことができます。クラスタ化された値を含む列はエクステント除外に適しています。

列スキャンにフィルタが設定されているときにエクステントを除外する場合、そのフィルタの値は列の各エクステントに格納されている最小値および最大値と比較されます。

Coll
Ext 1 Min 1 Max 100
Ext 2 Min 101 Max 200
Ext 3 Min 201 Max 300
Ext 4 Min 301 Max 400

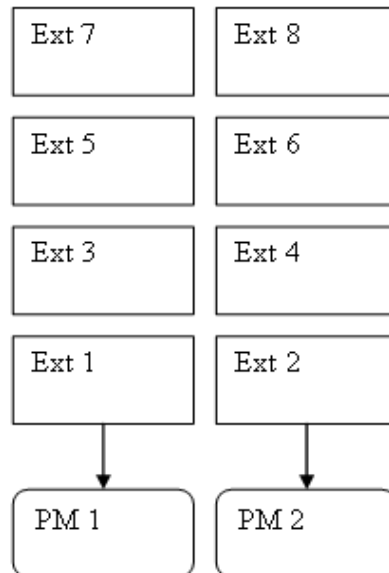
前述の図では、「COL1 BETWEEN 220 AND 250」という WHERE 列フィルタが指定されている場合、InfiniDB ではエクステント 1、2 および 4 がスキャンから除外されるため、I/O の 4 分の 3 が節約され、多くの比較操作が省略されます。エクステントにフィルタと一致する行が含まれている可能性がない場合、そのエクステント全体が無視されます。また、1 つの列に対して除外された各エクステントによって表内の他のすべての列の同じエクステントも除外されるため、使用できないフィルタチェーンを I/O の必要なく簡単に識別できます。たとえば、次の 2 つの列およびそれらのエクステントマップ情報について考えてみます。

Col1	Col2
Ext 1 Min 1 Max 100	Ext 1 Min 100 Max 10000
Ext 2 Min 101 Max 200	Ext 2 Min 10100 Max 20000
Ext 3 Min 201 Max 300	Ext 3 Min 20100 Max 30000
Ext 4 Min 301 Max 400	Ext 4 Min 30100 Max 40000

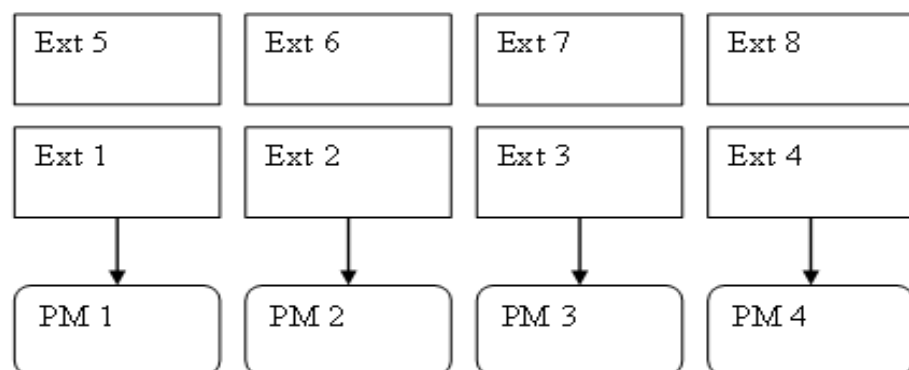
「COL1 BETWEEN 220 AND 250 AND COL2 < 10000」という WHERE フィルタが指定されている場合、InfiniDB では、エクステント 1、2 および 4 が最初の列フィルタから除外されます。その後、COL2（つまりエクステント 3 のみ）に対して一致するエクステントのみが検索され、一致するエクステントが存在しないことが確認されると I/O を実行することなく 0 行が返されます。

I/O 処理およびワークロードの分散

InfiniDB Enterprise 構成では、ブロックはエクステント内に割り当てられており、エクステントはディスク上の連続領域（セグメントファイル）にマップされるため、任意の数のパフォーマンスモジュールに作業を均等に分散できます。たとえば、2つのパフォーマンスモジュールシステムでは、8つのエクステントが含まれている列のスキャンは次のように分散されます。



4つのパフォーマンスモジュールシステムでは次のようになります。



このアーキテクチャでは、InfiniDBによって作業がすべての関連パフォーマンスモジュールに適切に分散されるため、多くの問合せにおいてパフォーマンスをリニアに向上させることができます。

大規模なスキャン操作のために、InfiniDBでは、ColScanReadAheadBlocksパラメータによって構成される先読みメカニズムを採用しています。このメカニズムは現

在 512 にデフォルト設定されますが、問合せへの対応に必要でない可能性があるデータの読取りを回避するために、InfiniDB ではプリフェッチしきい値パラメータ (PrefetchThreshold、デフォルト設定は 5) も使用されます。これによって、問合せに必要なブロックの数がしきい値を下回っていることがシステムで確認された場合、必要なブロックのみが読み取られます。

もう 1 つの注意点は、InfiniDB ではハッシュ結合がサポートされているという点です。このマニュアルの作成時点では、ハッシュ結合は MySQL ではサポートされていません。また、InfiniDB のハッシュ結合は、複数ノードの MPP 構成が使用されている場合、分散して並行で処理できます。

データベースの管理

製品版の InfiniDB には、InfiniDB データベースに管理コマンドを発行するのに使用される InfiniDB 管理（コマンド）コンソールが付属します。管理者は、これを使用して 1 つ以上の InfiniDB サーバーの構成、監視および管理を行うことができます。ユーティリティ名は `cmconsole` で、InfiniDB バイナリに対する実行権限を持つ任意のオペレーティングシステムユーザーが起動できます。

サポートされているコマンドに関するヘルプを参照するには、**help** と入力してコマンドのリストを確認します。特定のコマンドのヘルプは、**help** に続けて目的のコマンドを入力すると参照できます。

```

InfiniDB> help
help    Fri Oct 26 14:11:12 2012

List of commands:
Note: the command must be the first entry entered on the command line

Command      Description
-----
?            Get help on the Console Commands
addDbroot    Add DBRoot Disk storage to the Calpont InfiniDB System
addExternalDevice  Add External Device to Configuration file
addModule    Add a Module within the Calpont InfiniDB System
alterSystem-disableModule  Disable a Module and Alter the Calpont InfiniDB System
alterSystem-enableModule  Enable a Module and Alter the Calpont InfiniDB System
assignDbrootPmConfig  Assign unassigned DBRoots to Performance Module
assignElasticIPAddress  Assign Amazon Elastic IP Address to a module
disableLog   Disable the levels of process and debug logging
enableLog    Enable the levels of process and debug logging
exit         Exit from the Console tool
getActiveAlarms  Get Active Alarm list
getActiveSQLStatements  Get List Active SQL Statements within the System
getAlarmConfig  Get Alarm Configuration Information
getAlarmHistory  Get system alarms
getAlarmSummary  Get Summary counts of Active Alarm
getCalpontSoftwareInfo  Get the Calpont InfiniDB RPM detailed information
getExternalDeviceConfig  Get External Device Configuration Information
getLogConfig   Get the System log file configuration
getModuleConfig  Get Module Name Configuration Information
getModuleCpu   Get a Module CPU usage
getModuleCpuUsers  Get a Module Top Processes utilizing CPU
getModuleDisk  Get a Module Disk usage
getModuleMemory  Get a Module Memory usage
getModuleMemoryUsers  Get a Module Top Processes utilizing Memory
getModuleResourceUsage  Get a Module Resource usage
getModuleTypeConfig  Get Module Type Configuration Information
getProcessConfig  Get Process Configuration Information
getProcessStatus  Get Calpont InfiniDB Process Statuses
getStorageConfig  Get System Storage Configuration Information
getStorageStatus  Get System Storage Status
getSystemConfig  Get System Configuration Information
getSystemCpu   Get System CPU usage on all modules
getSystemCpuUsers  Get System Top Processes utilizing CPU
getSystemDisk  Get System Disk usage on all modules
getSystemInfo  Get the Over-all System Statuses
getSystemMemory  Get System Memory usage on all modules
getSystemMemoryUsers  Get System Top Processes utilizing Memory
getSystemNetworkConfig  Get System Network Configuration Information
getSystemResourceUsage  Get System Resource usage on all modules
getSystemStatus  Get System and Modules Status
help          Get help on the Console Commands
monitorAlarms  Monitor alarms in realtime mode
movePmDbrootConfig  Move DBRoots from one Performance Module to another
quit         Exit from the Console tool

```

InfiniDB コマンドの全リストについては、『Calpont InfiniDB 管理者ガイド』を参照してください。

SQL インタフェース

InfiniDB データベースに対する SQL コマンドは、`idbmysql` ユーティリティを使用して入力できます。このコマンドラインツールは、用途と機能の点で、標準 MySQL の `mysql` コマンドラインユーティリティに類似しています。すべての標準 SQL 文 (SELECT、INSERT、UPDATE、DELETE、GRANT など) は、このユーティリティを介して発行できます。

InfiniDB インスタンスの起動および停止

関連するすべてのユーザーモジュールおよびパフォーマンスモジュールの起動および停止は、`cmconsole` ユーティリティを使用して非常に簡単に実行できます。主な 3 つのコマンドを次に示します。

- ◆ `startsystem`
- ◆ `stopsystem`
- ◆ `restartsystem`

管理者は、基本的なオペレーティングシステムのサービス起動コマンドを活用することもできます。たとえば、管理者は、個々の Linux マシンで `service infinidb start` コマンドを発行して InfiniDB インスタンスを起動できます。`service infinidb stop` を実行した場合は InfiniDB インスタンスが停止されます。

セキュリティの概念

この項では、Calpont のセキュリティの概念およびプラクティスについて簡単に説明します。

オペレーティングシステムのセキュリティ

InfiniDB は、実行可能ファイルおよび基礎となるデータベースファイルの保護に関して、オペレーティングシステムの標準的なセキュリティを使用します。このレベルでは、特殊なものは使用されません。InfiniDB データベースをインストールして管理する管理者は、InfiniDB バイナリまたはデータベースファイルを含むファイルシステムやディレクトリへのアクセス権をどのオペレーティングシステムアカウントに付与するかという点に関して、標準のセキュリティ手順に従う必要があります。

データベースのログインアクセス権

InfiniDB では、標準 MySQL のセキュリティコマンドを使用して、InfiniDB データベースに対するユーザーの追加および削除を行います。ユーザーアカウントには、システム全体および作業する特定のデータベースへのアクセス権が付与されます。

たとえば、「rms」という名前の新しいユーザーに、システムへのアクセス権および「db1」という名前のデータベースに対する完全な権限を付与する場合、管理者は次のコマンドを入力します。

```
mysql> grant all on db1.* to rms identified by 'password' ;  
Query OK, 0 rows affected (0.00 sec)
```

InfiniDB データベースからユーザーアカウントを削除する場合、管理者は `drop user` コマンドを使用します。

ユーザーアカウントの追加および削除の詳細は、最新オンライン版の MySQL のリファレンスマニュアルを参照してください。

オブジェクトのセキュリティ

データベースやシステムのアクセスセキュリティと同様に、InfiniDB では標準 MySQL のセキュリティコマンドを使用します。これには、選択されたユーザーアカウントに対してオブジェクトレベルの権限を付与および削除する、ANSI 標準 SQL の `GRANT` および `REVOKE` コマンドが含まれます。たとえば、「db1」データベースの「t1」という名前の表に対するすべての権限を「rms」という名前のユーザーに付与する場合、管理者は次のコマンドを発行します。

```
mysql> grant all on db1.t1 to rms;  
Query OK, 0 rows affected (0.00 sec)
```

`GRANT` および `REVOKE` の機能の全リストについては、最新オンライン版の MySQL のリファレンスマニュアルを参照してください。

InfiniDB 構成ファイル

InfiniDB 構成ファイルには、InfiniDB データベースの操作を制御するパラメータが含まれます。これは XML ベースのファイルで、システムのスタートアップおよびシャットダウン時に、InfiniDB システム内の関連するすべてのノードで自動的にレプリケートされます（または明示的なコマンドを発行してファイルをレプリケートします）。構成ファイルのマスターコピーは、プライマリのパフォーマンスモジュールに保持されます。システムは、スタートアップ時に構成ファイルから情報を読み取って、システムのアクティビティに必要なリソース（データキャッシュなど）を割り当てます。

構成ファイルをテキストエディタで直接編集することはお薦めしません。かわりに、ファイルを変更するためのコマンドラインツール configxml が提供されています。スクリプトに対して各種パラメータを使用して、構成情報を表示および設定します。

たとえば、アクティブなトランザクションログのアーカイブログがディスクに書き込まれる時間間隔の値（分単位で設定）を表示する場合、DBA は次を入力します。

```
[root@server1 bin]$ ./configxml.sh getconfig SystemConfig
TransactionArchivePeriod
Current value of SystemConfig / TransactionArchivePeriod is
10
```

アーカイブログの時間間隔を 60 分に変更するには、setconfig オプションを使用して次を入力します。

```
[root@server1 bin]$ ./configxml.sh setconfig SystemConfig
TransactionArchivePeriod 60
Old value of SystemConfig / TransactionArchivePeriod is 10

/usr/local/Calpont/etc/Calpont.xml backed up to /usr/
local/ Calpont/etc/Calpont.xml.1255027253

Old value of SystemConfig / TransactionArchivePeriod is 10
SystemConfig / TransactionArchivePeriod now set to 60
```

Calpont で使用される構成変数の全リストについては、『Calpont InfiniDB 管理者ガイド』を参照してください。

新しいデータベースの作成

InfiniDB のパラダイムは、サーバーのインスタンスが実行されて、データベースがそのインスタンスの下で管理されるという点で、MySQL および Microsoft SQL Server に類似しています。インスタンスとデータベースには 1 対多の関係があります。これは、インスタンスとデータベースが通常 1 対 1 である Oracle と異なります。

InfiniDB では、新しいデータベース（MySQL ではスキーマと呼ばれることもある）を非常に簡単に作成できます。

```
mysql> create database db1;
Query OK, 1 row affected (0.00 sec)
```

作成後、次の例のように、データベースに表を作成できます。

```
mysql> use db1
Database changed
mysql> create table t1 (c1 int, c2 date) engine=infinidb;
Query OK, 0 rows affected (0.03 sec)

mysql> desc t1;
+-----+-----+-----+-----+-----+-----+
| Field | Type   | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
| c1    | int(11)| YES  |     | NULL    |       |
| c2    | date   | YES  |     | NULL    |       |
+-----+-----+-----+-----+-----+-----+
2 rows in set (0.00 sec)
```

管理ユーティリティ

InfiniDB では、様々なコマンドライン管理ユーティリティが提供されています。よく使用されるユーティリティを以下に示します。

- ◆ colxml : バルクロードユーティリティ cpimport によって使用されるジョブファイルを作成します
- ◆ configxml.sh : 構成ファイルのパラメータの読取りおよび更新に使用されます
- ◆ cpimport : データベースへのデータのバルクロードに使用されます
- ◆ dumpcol : 列データを出力します
- ◆ edititem : エクステンタマップのレポートおよび編集に使用されます
- ◆ edittxn : システム変更番号 (SCN) を出力し、複数ノード構成で一貫性の検証に役立ちます
- ◆ file2oid.sh : 列のいずれかのデータファイルが指定されている場合に、その列のオブジェクト ID を出力します
- ◆ oid2file : オブジェクトの最初のファイルのディレクトリ情報を表示します
- ◆ replaytransactionlog : アーカイブされたトランザクションログおよび現在のトランザクションログのすべてのトランザクション情報を含むファイルを作成します

システムログ

InfiniDB は、次のようないくつかのログにシステム情報を書き込みます。

- ◆ サーバーの操作に関する一般情報を含む情報ログ (info.log)
- ◆ サーバーで発生したエラーに関するデータを含むエラーログ (err.log)
- ◆ 将来的なシステムエラー（サーバーのディスク領域の不足など）を示す可能性があるサーバーの状態を通知する警告ログ (warning.log)

InfiniDB では、これらのファイルは一般的なロギングディレクトリに格納されます（たとえば、Linux の場合、この場所は /var/log/Calpont です）。

バックアップおよびリカバリ

この項では、InfiniDB のバックアップおよびリカバリの操作について説明します。

バックアップ手順の概要

InfiniDB システムの全体バックアップを作成する場合の基本的な手順は次のとおりです。

- ◆ システムでの書込みアクティビティを一時的に停止します（製品版の InfiniDB でのみ利用可能な cmconsole ユーティリティを使用して実行）
- ◆ MySQL インスタンスをバックアップします
- ◆ InfiniDB データベースファイルをバックアップします
- ◆ InfiniDB のトランザクションログをバックアップします
- ◆ システムでの書込みアクティビティをリストアします（製品版の InfiniDB でのみ利用可能な cmconsole ユーティリティを使用して実行）

データベースの書込みの一時停止

InfiniDB システムでデータベースの書込みを一時停止するには、cmconsole 内で次のコマンドを発行します。

```
calpont> suspendDatabaseWrites
suspenddatabasewrites    Wed Sep 30 08:58:06 2009
```

```
This command suspends the DDL/DML writes to the InfiniDB
Database
```

```
Do you want to proceed: (y or n) [n]: y
```

```
Suspend InfiniDB Database Writes Request successfully
completed
```

MySQL インスタンスのバックアップ

MySQL データベースは、いくつかの方法を使用してバックアップできます。

方法 1 :

MySQL が提供するユーティリティ (`mysqldump` や `mysqlhotcopy` など) を使用します。InfiniDB では MySQL 内のデータベースデータは格納されないため、MySQL インスタンスのバックアップは迅速に実行する必要があることに注意してください。

MySQL のバックアップおよびリカバリの詳細な処理については、最新オンライン版の MySQL のリファレンスマニュアルを参照してください。

方法 2 :

MySQL フロントエンドには Calpont InfiniDB 固有の実装はされないため、通常の手順のかわりに、次のディレクトリをバックアップします。

```
/usr/local/Calpont/mysql/db
```

例 :

```
cp -rp /usr/local/Calpont/mysql/db /mnt/InfiniDB/backup/
frontend
```

注意 : `-rp` オプションは、ディレクトリを再帰的にコピーし、所有権情報を保存するためのものです。

選択したデータベースのみをバックアップする場合は、かわりに前述のディレクトリ内のデータベースディレクトリをバックアップします。

InfiniDB データファイルのバックアップ

InfiniDB データファイルをバックアップするには、そのファイルをバックアップ場所にコピーするか、システムまたはベンダーが提供するバックアップまたはスナップショットのユーティリティを使用して InfiniDB データが存在するディレクトリのスナップショットを作成します。これらの場所の次のファイルをバックアップする必要があります。

- ◆ DBRM ファイル。たとえば、Linux の標準インストールの場合、このファイルは /usr/local/Calpont/data1/systemfiles/dbrm/ ディレクトリにあります。
- ◆ すべてのデータファイル。たとえば、Linux の標準インストールの場合、すべてのファイルは /usr/local/Calpont/dataN (N は data1、data2 などの一意のディレクトリを表す) にあります。

トランザクションログのバックアップ

InfiniDB トランザクションログを適切にバックアップするには、管理者は次の手順を実行する必要があります。

- ◆ 現在のトランザクションログをコピーするか、このファイルのスナップショットを作成します。たとえば、Linux の標準インストールの場合、トランザクションログは /usr/local/Calpont/data1/systemFiles/dataTransaction にあります。
- ◆ ポイントインタイムリカバリに使用可能なアーカイブされたトランザクションログファイルをコピーするか、このファイルのスナップショットを作成します。Linux の標準インストールの場合、このファイルは /usr/local/Calpont/data1/systemFiles/dataTransaction/archive にあります。

データベースの書き込みの再開

バックアップの完了後にデータベースの書き込みアクティビティを再開するには、cmconsole 内で次のコマンドを発行します。

```
calpont> resumeDatabaseWrites
resumedatabasewrites    Wed Sep 30 08:58:23 2009

This command resumes the DDL/DML writes to the InfiniDB
Database

Do you want to proceed: (y or n) [n]: y

Resume InfiniDB Database Writes Request successfully
completed
```

リカバリ手順の概要

InfiniDB データベースの完全リストアを行うには、次の手順を実行します。

- ◆ MySQL インスタンスをリストアします
- ◆ InfiniDB データベースをリストアします
- ◆ トランザクションログを再生します（手動のポイントインタイムリカバリが必要な場合）

MySQL インスタンスのリストア

MySQL データベースは様々な方法でリストアできますが、最も一般的な方法は、mysqldump によって作成したバックアップファイルを取得し、mysql コマンドラインユーティリティからそのファイルを実行することです。これによって、そのファイル内に含まれるすべての SQL 文が実行されます。

MySQL のバックアップおよびリカバリの詳細な処理については、最新オンライン版の MySQL のリファレンスマニュアルを参照してください。

InfiniDB データファイルのリストア

InfiniDB データファイルをリストアするには、そのファイルをバックアップ場所からコピーするか、システムまたはベンダーが提供するスナップショットユーティリティを使用して、InfiniDB データが存在したディレクトリのスナップショットをリストアします。リストアするファイルには次のものがあります。

- ◆ DBRM ファイル。たとえば、Linux の標準インストールの場合、このファイルは /usr/local/Calpont/data1/systemfiles/dbrm/ ディレクトリに戻されます。
- ◆ すべてのデータファイル。たとえば、Linux の標準インストールの場合、すべてのファイルは /usr/local/Calpont/dataN (N は data1、data2 などの一意のディレクトリを表す) に戻されます。

InfiniDB トランザクションログの再生

InfiniDB では、アーカイブされたトランザクションログおよびアクティブなトランザクションログを読み取るユーティリティ (ReplayTransactionLog と呼ばれる) が提供されています。このユーティリティは、アーカイブされたトランザクションログファイルおよびアクティブなトランザクションログファイルを調べて SQL 文 (DML および DDL) を生成します。この SQL 文は、InfiniDB SQL インタフェースからの実行およびデータベース内での再実行が可能なファイルに書き込まれます。

DBA またはシステム管理者は、データベースの最後の全体バックアップ以降に発生してデータベースへの適用が必要なトランザクションがある場合に、`ReplayTransactionLog` を使用して手動のポイントインタイムリカバリ操作を実行できます。現行バージョンの InfiniDB では、`ReplayTransactionLog` ユーティリティは、利用可能なすべてのアーカイブされたログファイルと現在のトランザクションログを読み取って出力ファイルを生成することに注意してください。DBA は、InfiniDB SQL インタフェースで実行されるファイルに各種の文が含まれないようにする場合、ファイルを手動で編集してこれらの文を削除する必要があります。

最後のバックアップ以降に発生したバルクロード操作も再実行する必要があることに注意してください。

GUI ツールのサポート

現在、Calpont は、InfiniDB データベースで動作する GUI ツールを提供していません。一般的なデータモデリング、問合せの開発、一般的な MySQL のセキュリティおよびオブジェクト管理用に、開発者または DBA は MySQL から無償の GUI ツール (MySQL Workbench および古い GUI ツールのバンドル) をダウンロードできます。

また、開発者または DBA が InfiniDB データベースオブジェクトに対して問合せや管理を実行できるツールが多数のサードパーティのソフトウェアベンダーによって作成されています。SQLYog、Quest Software (TOAD for MySQL)、Mentat Technologies (DreamCoder for MySQL) など、多数のベンダーが、無償版の GUI ツールおよび高度な機能を含む有償エディションの両方を提供しています。

オブジェクト管理

この項では、InfiniDB データベースオブジェクトの作成および一般的な管理について説明します。

表

リレーショナルデータベースである InfiniDB は論理表構造でデータを格納します。InfiniDB と他の多くの RDBMS の主な違いは、InfiniDB では列指向の構造が使用されることです。これは、従来の行ベースの実装より、分析環境またはデータウェアハウス環境に適しています。ただし、データベース表の作成に関しては、InfiniDB と他のリレーショナルデータベースに表面上の違いはありません。表内のデータの作成、移入、操作および表示には、ANSI 標準の SQL 文、DDL 文または DML 文が使用されます。

たとえば、列が 3 つの「db1」というデータベース内に新しい表を作成する場合、ユーザーは次のように入力できます。

```
mysql> use db1
Database changed
mysql> create table t1 (c1 int, c2 date, c3 varchar(20))
engine=infinidb;
Query OK, 0 rows affected (3.03 sec)
```

InfiniDB によって各表に連続した領域が事前に割り当てられるため、create table コマンドの実行では、表が空で作成されている場合でも表に領域が事前に割り当てられない他のデータベースシステムより少し時間がかかることに注意してください。

作成後、ユーザーは Calpont のバルクロードユーティリティを使用して大量のデータを迅速に表にロードしたり、MySQL ユーティリティを使用してデータを表にロードしたりできます。また、1 つの DML 文を使用してデータの挿入、更新および削除を行うこともできます。表への領域割当てに関してユーザーによる操作は必要ありません。InfiniDB による領域の割当て方法の詳細は、InfiniDB のストレージ概念に関する前述の項を参照してください。

InfiniDB の現行の実装では、主キー、外部キーまたは列レベルの制約はサポートされていません。このサポートは今後のリリースのサーバーで追加されます。

CREATE TABLE コマンドの処理の詳細は、『InfiniDB SQL 構文ガイド』または MySQL のリファレンスマニュアルを参照してください。

表の変更および削除

InfiniDB では、表の列の追加および削除に対して標準の ALTER TABLE コマンドがサポートされています。表の変更は、標準の MySQL（または他のデータベース）より InfiniDB での方が短時間で実行されます。これは、InfiniDB では索引付けが使用されないため操作時に索引の保守を実行する必要がないためです。

また、InfiniDB では、特別な方法を使用してオンラインで表に新しい列を追加することもできます。操作時、表の読取り操作は中断なく続行されます。オンラインで表に新しい列を追加するために calonlinealter 関数コールが使用された後、操作が完了すると MySQL データディクショナリが同期化されます。たとえば、新しい整数列（c7）を表（t1）にオンラインで追加するには、ユーザーは次のコマンドを発行します。

```
select calonlinealter('alter table t1 add column c7 int;');
alter table t1 add column c7 int comment 'schema sync only';
```

表の削除は、単純な DROP TABLE コマンドを使用して実行します。InfiniDB では、表の削除時に、基礎となる物理データファイルもすべてサーバーから削除されるため、後で管理者が手動クリーンアップを実行する必要はありません。

ALTER TABLE および DROP TABLE コマンドの処理の詳細は、『InfiniDB SQL 構文ガイド』または MySQL のリファレンスマニュアルを参照してください。

ストアドプロシージャおよび SQL 関数

InfiniDB では、MySQL のストアドプロシージャおよび SQL 関数がサポートされています。ストアドプロシージャは、サーバー自体に格納されているプロシージャコードオブジェクトであり、データのプログラムによる処理の実行および機密データに対するセキュリティの強化のために使用される場合があります。DBA は、ストアドプロシージャに実行権限を付与してアクセスする表内の列およびそのアクセス方法を制限できますが、基礎となる表自体は制限できません。

たとえば、ユーザーが顧客番号を入力したときに「customer」という表から 2 つの列を取得するストアドプロシージャを作成する必要がある場合、DBA は次のようなプロシージャを作成できます。

```
delimiter //
create procedure p_cust (p1 int)
begin
select c_name,c_address from customer
```



```

where c_custkey = p1;
end
//
delimiter ;

```

ストアドプロシージャのコールは次のように実行します。

```

mysql> call p_cust(1);
+-----+-----+
| c_name                | c_address      |
+-----+-----+
| Customer#000000001    | IVhzIApeRb     |
+-----+-----+
1 row in set (0.26 sec)

```

ストアドプロシージャおよび InfiniDB を使用する現在の制限では、一度に使用できるオープンカーソルは1つだけです。

また、SQL 関数も作成して使用できます。簡単な関数の例を次に示します。

```

mysql> CREATE FUNCTION hello (s CHAR(20))
-> RETURNS CHAR(50) DETERMINISTIC
-> RETURN CONCAT('Hello, ',s,'!');
Query OK, 0 rows affected (0.00 sec)

mysql> select hello('world');
+-----+
| hello('world') |
+-----+
| Hello, world!  |
+-----+
1 row in set (0.03 sec)

```

ストアドプロシージャおよび関数の詳細は、最新版の MySQL のリファレンスマニュアルを参照してください。

サポートされていないデータベースオブジェクト

現行バージョンの InfiniDB では次のオブジェクトはサポートされていません。

- ◆ トリガー
- ◆ 主キー
- ◆ 外部キー
- ◆ NOT NULL とデフォルト値以外の列制約

InfiniDB システムカタログ

InfiniDB をインストールすると、InfiniDB オブジェクトに関連付けられているメタデータの追跡に使用される小さいシステムカタログが構成されます。このシステムカタログは、MySQL の mysql および information_schema データベース内にある標準オブジェクトとともに使用されます。データベース名またはスキーマ名は calpontsys です。

InfiniDB システムカタログデータベース内に含まれているオブジェクトは次のとおりです。

```
mysql> desc systable;
+-----+-----+-----+-----+-----+-----+
| Field          | Type          | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
| tablename      | varchar(128)  | YES  |     | NULL    |       |
| schema         | varchar(128)  | YES  |     | NULL    |       |
| objectid       | int(11)       | YES  |     | NULL    |       |
| createdate     | date          | YES  |     | NULL    |       |
| lastupdate     | date          | YES  |     | NULL    |       |
| init           | int(11)       | YES  |     | NULL    |       |
| next           | int(11)       | YES  |     | NULL    |       |
| numofrows      | int(11)       | YES  |     | NULL    |       |
| avgrowlen      | int(11)       | YES  |     | NULL    |       |
| numofblocks    | int(11)       | YES  |     | NULL    |       |
| autoincrement  | int(11)       | YES  |     | NULL    |       |
+-----+-----+-----+-----+-----+-----+
11 rows in set (0.00 sec)
```

```
mysql> desc syscolumn;
```

Field	Type	Null	Key	Default	Extra
schema	varchar(128)	YES		NULL	
tablename	varchar(128)	YES		NULL	
columnname	varchar(128)	YES		NULL	
objectid	int(11)	YES		NULL	
dictobjectid	int(11)	YES		NULL	
listobjectid	int(11)	YES		NULL	
treeobjectid	int(11)	YES		NULL	
datatype	int(11)	YES		NULL	
columnlength	int(11)	YES		NULL	
columnposition	int(11)	YES		NULL	
lastupdate	date	YES		NULL	
defaultvalue	varchar(64)	YES		NULL	
nullable	int(11)	YES		NULL	
scale	int(11)	YES		NULL	
prec	int(11)	YES		NULL	
autoincrement	char(1)	YES		NULL	
distcount	int(11)	YES		NULL	
nullcount	int(11)	YES		NULL	
minvalue	varchar(64)	YES		NULL	
maxvalue	varchar(64)	YES		NULL	
compressiontype	int(11)	YES		NULL	
nextvalue	bigint(20)	YES		NULL	

```
22 rows in set (0.00 sec)
```


データのロードおよび操作

この項では、InfiniDB データベース内でのデータのロードおよび操作について説明します。

InfiniDB のバルクローダー

InfiniDB では、InfiniDB の表にデータを迅速に効率よくインポートする高速のバルクロードユーティリティが提供されています。このユーティリティは、データのフィールド（表内の列など）がデリミタで区切られたデータを含む任意のフラットファイルを入力することができます。デフォルトのデリミタはパイプ (|) 文字ですが、カンマなど、その他のデリミタも使用できます。

InfiniDB のバルクロードユーティリティ (cpimport と呼ばれる) は、InfiniDB データベースへのデータのインポート時に次の操作を実行します。

- ◆ 指定されたフラットファイルからデータを読み取ります
- ◆ InfiniDB の列指向のストレージ設計に合うようにデータを変換します
- ◆ 冗長データをトークン化し、論理的に圧縮します
- ◆ データをディスクに書き込みます

cpimport を使用する最も一般的な方法には次の 2 つがあります。

1. UM から : cpimport がすべてのパフォーマンスモジュールに行を分散します
2. PM から : cpimport が起動元の PM にのみインポートされた行をロードします

これらほど一般的ではありませんが、cpimport には他にも多くのデータのロード方法があります。cpimport の使用方法の詳細については、『Calpont InfiniDB 管理者ガイド』を参照してください。

cpimport ユーティリティの 2 つの主な使用手順は次のとおりです。

1. フラットファイルから複数の表にデータをロードするのに使用するジョブファイルを作成します (オプション)
2. cpimport ユーティリティを実行して、データのインポートを実行します

バルクロードは表への追加操作であるため、処理中も既存のデータの読取りが可能であり、既存のデータは影響を受けないことに注意してください。また、バルクロードでは、データ操作がトランザクションログに書き込まれません。これらの操作にトランザクションの特性はなく、この時点ではアトミック操作と見なされます。ただし、情報マーカがトランザクションログに配置されるため、DBA はバルク操作が発生したことを確認できます。

ロード操作が完了すると、各列ファイルの最高水位標がアトミック操作で移動されます。このアトミック操作により、新たにロードされたデータを後続の間合せて読み取ることができるようになります。この追加操作では、読取り一貫性を実現できますが、データのロギングのオーバーヘッドは発生しません。

シンプルバルクロードの例

DBA が「db1」データベースの「t1」表をロードし、「t1.tbl」という名前のロードファイルを使用することを想定します。このファイルには次のデータが含まれ、パイプデリミタを使用してデータが区切られています。

```
1|Smith|2000-01-01
2|Jones|2000-02-01
3|Fields|2000-03-01
4|Johnson|2000-04-01
5|Simms|2000-05-01
```

Calpont の現在のバージョンでは、`-f` パス名オプションが使用されていない場合、このファイルの場所は、デフォルトで現在のディレクトリに設定されます。現在、ユーティリティでは `<table name>.tbl` のネーミング形式がデフォルトで設定されます。

表の構造は次のとおりです。

```
mysql> desc t1;
+-----+-----+-----+-----+-----+-----+
| Field | Type          | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
| c1    | int(11)       | YES  |     | NULL    |       |
| c2    | varchar(20)   | YES  |     | NULL    |       |
| c3    | date          | YES  |     | NULL    |       |
+-----+-----+-----+-----+-----+-----+
3 rows in set (0.03 sec)
```

1 つの表のみをロードする場合は、`cpimport` ユーティリティを直接コールできます。DBA は、`cpimport` ユーティリティ（InfiniDB の `bin` ディレクトリにある）を実行して、実際にデータを表にロードします。

```
[root@server1 test]# pwd
/home/test

[root@server1 bin]# /usr/local/Calpont/bin/cpimport db1 t1 t1.db1

Bulkload root directory : /home/test
job description file : Job_501.xml
2009-10-06 15:08:16 (17995) INFO : successfully load job file /
usr/local/Calpont/data/bulk/job/Job_501.xml
2009-10-06 15:08:16 (17995) INFO : PreProcessing check starts
2009-10-06 15:08:16 (17995) INFO : PreProcessing check completed
2009-10-06 15:08:16 (17995) INFO : preProcess completed, total run
time : 0 seconds
2009-10-06 15:08:16 (17995) INFO : No of Read Threads Spawned = 1
2009-10-06 15:08:16 (17995) INFO : No of Parse Threads Spawned = 3
2009-10-06 15:08:17 (17995) INFO : For table db1.t1: 5 rows
processed and 5 rows inserted.
2009-10-06 15:08:17 (17995) INFO : Bulk load completed, total run
time : 1 seconds
```

複数の表をロードする場合は、後述する複数の表のバルクロード例を参照してください。

STDIN のバルクロードの例

InfiniDB では、cpimport を簡単に使用した、標準入力からのデータのバルクロードがサポートされています。前述の同じロードファイルを使用して、ファイルを cpimport プロセスに送信できます。たとえば、最初の 100,000,000 行のみをインポートするように t1.tbl ファイルを制限でき、このファイルを cpimport に送信します。

```
head -100000000 t1.tbl | /usr/local/Calpont/bin/cpimport db1 t1
-s '\t'
```

表から選択した内容のバルク挿入

標準入力、任意の SELECT 文からの出力を高速のバルクローダーに直接送信するために使用することもできます。ここでは、データ以外の書式設定を削除するために -N フラグを使用して db2.source_table から選択しています。

```
idbmysql -e 'select * from source_table;' -N db2 | /usr/local/
Calpont/bin/cpimport db1 t1 -s '\t'
```

複数の表のバルクロード

インポートユーティリティを使用して複数の表を同時にインポートできます。シンプルバルクロードでは、インポートごとに表が一意であるかぎり、複数のジョブを同時に発行できます。従来のバルクロードでは、スキーマ内のすべての表をインポートするか、または colxml で -t オプションを使用して特定の表をリストすることで複数の表をインポートできます。表およびジョブ ID が一意であるかぎり、従来のインタフェースを使用して複数のジョブを発行できます。

従来のバルクロードの例

DBA は、従来のバルクロードを使用して、スキーマ全体または複数の表を 1 つのバルクロードジョブでロードすることもできます。たとえば、DBA が、「tpch2」という名前のデータベースにある次の表にデータをロードするとします。

```
mysql> show tables;
+-----+
| Tables_in_rms |
+-----+
| customer      |
| lineitem      |
| nation        |
| orders        |
| part          |
| partsupp      |
| region        |
| supplier      |
+-----+
8 rows in set (0.00 sec)
```

ロードファイルでデフォルトのパイプ (|) フィールドデリミタを使用し、インポートディレクトリにすべての表のロードファイルを配置してから、「tpch2」データベース自体に対するロードジョブ用に colxml ユーティリティを実行できます。

```
[user1@server1 bin]$ ./colxml tpch2 -j500
Running colxml with the following parameters:
2009-10-07 15:14:20 (9481) INFO :
    Schema: tpch2
    Tables:
    Load Files:
    -b      0
    -c     1048576
    -d      |
```



```
-e      10
-f      CSV
-j      500
-n
-p      /usr/local/Calpont/data/bulk/job/
-r      5
-s
-u
-w      10485760
-x      tbl
```

File completed for tables:

```
tpch2.customer
tpch2.lineitem
tpch2.nation
tpch2.orders
tpch2.part
tpch2.partsupp
tpch2.region
tpch2.supplier
```

Normal exit.

その後、実際のバルクロード操作を実行できます。

```
[user1@server1 bin]$ ./cpimport -j 500
```

```
Bulkload root directory : /usr/local/Calpont/data/bulk
job description file : Job_500.xml
2009-10-07 15:14:59 (9952) INFO : successfully load job file /usr/
local/Calpont  data/bulk/job/Job_500.xml
2009-10-07 15:14:59 (9952) INFO : PreProcessing check starts
2009-10-07 15:15:04 (9952) INFO : PreProcessing check completed
2009-10-07 15:15:04 (9952) INFO : preProcess completed, total run
time : 5 seconds
2009-10-07 15:15:04 (9952) INFO : No of Read Threads Spawned = 1
2009-10-07 15:15:04 (9952) INFO : No of Parse Threads Spawned = 3
2009-10-07 15:15:06 (9952) INFO : For table tpch2.customer: 150000
rows processed and 150000 rows inserted.
2009-10-07 15:16:12 (9952) INFO : For table tpch2.nation: 25 rows
processed and 25 rows inserted.
2009-10-07 15:16:12 (9952) INFO : For table tpch2.lineitem:
6001215 rows processed and 6001215 rows inserted.
```

```
2009-10-07 15:16:31 (9952) INFO : For table tpch2.orders: 1500000
rows processed and 1500000 rows inserted.
2009-10-07 15:16:33 (9952) INFO : For table tpch2.part: 200000 rows
processed and 200000 rows inserted.
2009-10-07 15:16:44 (9952) INFO : For table tpch2.partsupp: 800000
rows processed and 800000 rows inserted.
2009-10-07 15:16:44 (9952) INFO : For table tpch2.region: 5 rows
processed and 5 rows inserted.
2009-10-07 15:16:45 (9952) INFO : For table tpch2.supplier: 10000
rows processed and 10000 rows inserted.
2009-10-07 15:16:45 (9952) INFO : Bulk load completed, total run time
: 106 seconds
```

STDIN のバルクロードの例

InfiniDB では、標準入力からのデータのバルクロードがサポートされています。この際、複数表の cpimport で -fSTDIN フラグを指定します。前述の同じロードファイルを -j501 コマンドで使用して、ファイルを cpimport プロセスに送信できます。たとえば、最初の 100,000,000 行のみをインポートするように t1.tbl ファイルを制限でき、このファイルを cpimport に送信します。

```
head -100000000 t1.tbl | /usr/local/Calpont/bin/cpimport -j501 -s
'\t' -fSTDIN
```

表から選択した内容のバルク挿入

標準入力は、任意の SELECT 文からの出力を高速のバルクローダーに直接送信するために使用することもできます。ここでは、データ以外の書式設定を削除するために -N フラグを使用して db2.source_table から選択しています。

```
idbmysql -e 'select * from source_table;' -N db2 | /usr/local/
Calpont/bin/cpimport -j501 -s '\t' -fSTDIN
```

データロード用の MySQL ユーティリティ

MySQL では、データベースの表にデータをロードする load data infile コマンドが提供されています。このコマンド（ユーティリティ）は InfiniDB でサポートされており、使用可能ですが、InfiniDB のバルクロードユーティリティ cpimport を使用したほうが高速です。

次の表を例とします。

```
mysql> desc t3;
+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+
```

Field	Type	Null	Key	Default	Extra
c1	int(11)	YES		NULL	
c2	int(11)	YES		NULL	

次のファイルをロードします。

```
[root@server1 local]# more t3.tst
1,2
2,3
3,4
```

次のように実行してロードできます。

```
mysql> load data infile '/usr/local/t3.tst' into table t3
fields terminated by ',';
Query OK, 3 rows affected (0.96 sec)
```

DML のサポート

InfiniDB では標準 DML がサポートされているため、ユーザーは、InfiniDB の表に対して SQL の INSERT、UPDATE および DELETE 文を発行できます。InfiniDB は列指向データベースであるため、InfiniDB で INSERT 文および DELETE 文を実行した場合、行指向データベースで実行した場合よりも、通常は低速になることに注意してください（より多くのブロックにアクセスする必要があります）。ただし、UPDATE 文はかなり高速に実行されます。

INSERT 文の場合、InfiniDB では MySQL の複数挿入操作がサポートされることに注意してください。次の表を例とします。

```
mysql> desc t3;
+-----+-----+-----+-----+-----+-----+
| Field | Type   | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
| c1    | int(11)| YES  |     | NULL    |       |
| c2    | int(11)| YES  |     | NULL    |       |
+-----+-----+-----+-----+-----+-----+
```

ユーザーは次の INSERT 文を発行できます。

```
mysql> insert into t3 values (1,2),(3,4),(5,6);
```

```
Query OK, 3 rows affected (0.34 sec)
Records: 3  Duplicates: 0  Warnings: 0
```

DML 構文の詳細は、『InfiniDB SQL 構文ガイド』を参照してください。

トランザクション管理

InfiniDB では、完全なコミット、ロールバックおよびデッドロック検出機能とともに、ACID 準拠のトランザクションがサポートされています。InfiniDB のデフォルトモードは、すべての DML 文が自動的に確定される自動コミットです。この設定は、InfiniDB のセッションごとに簡単に変更できます。たとえば、次の例では自動コミットモードをオフに設定し、表を更新して変更をロールバックし、再度表を更新して、データベースに対しトランザクションをコミットします。

```
mysql> set autocommit=0;
Query OK, 0 rows affected (0.00 sec)

mysql> select c1 from t3;
+-----+
| c1    |
+-----+
|      1 |
|      2 |
|      3 |
|      1 |
|      1 |
|      3 |
|      5 |
+-----+
7 rows in set (0.26 sec)

mysql> update t3 set c1=4;
Query OK, 7 rows affected (0.26 sec)
Rows matched: 0  Changed: 0  Warnings: 0

mysql> rollback;
Query OK, 0 rows affected (0.23 sec)

mysql> select c1 from t3;
+-----+
```

```

| c1 |
+-----+
| 1 |
| 2 |
| 3 |
| 1 |
| 1 |
| 3 |
| 5 |
+-----+
7 rows in set (0.26 sec)

```

```

mysql> update t3 set c1=5;
Query OK, 7 rows affected (0.14 sec)
Rows matched: 0  Changed: 0  Warnings: 0

```

```

mysql> commit;
Query OK, 0 rows affected (0.21 sec)

```

```

mysql> select c1 from t3;
+-----+
| c1 |
+-----+
| 5 |
| 5 |
| 5 |
| 5 |
| 5 |
| 5 |
| 5 |
+-----+
7 rows in set (0.14 sec)

```

同時実行の制御

InfiniDB では、ページレベルのロックが使用され、分離レベルでのコミット読取りロックが採用されており、コミットされたデータのみが読み取られるようにしています。

InfiniDB では、InfiniDB にバージョンングインフラストラクチャを提供する多版型同時実行制御 (MVCC) も使用されます。これは、他のデータベースでは「スナップショット読取り」と呼ばれることもあります。スナップショット読取りは、問合せが常に、問合せが発行された時点のデータを参照することを意味します。これによって、問合せに関係するすべての表のすべての列で読取り一貫性が実現されます。

InfiniDB のバージョンングの最も重要なメリットは、読取りが遮断されないことです。つまり、バージョンングされた (古い) ブロックを、あるセッションが更新しているときに、同時に別のセッションが読み取ることができます。InfiniDB のすべての文は、InfiniDB によってシステム変更番号 (SCN) として参照されるデータベースの特定のバージョン (またはスナップショット) で実行されます。循環バージョンバッファ、バージョン置換構造 (VSS) およびバージョンバッファブロックマネージャ (VBBM) 構造を使用して、リクエストされたブロックの適切なバージョンが検索されます。これらのオブジェクトのサイズは、予測負荷でのバッファオーバーランを回避できるように構成可能です。

バージョンングがトランザクションの DML 文のみに適用され、追加操作を使用して行われるバルクロードには適用されないことに注意してください。

アプリケーションおよびビジネスインテリジェンスの開発

この項では、アプリケーション開発およびビジネスインテリジェンス環境での InfiniDB との連携動作について簡単に説明します。

アプリケーションツールおよび BI ツールを使用した InfiniDB への接続

InfiniDB では MySQL フロントエンドが使用されるため、InfiniDB への接続には任意の標準 MySQL コネクタを使用できます。無償で入手できるコネクタには、次のものがあります。

- ◆ ADO.NET
- ◆ ODBC
- ◆ JDBC
- ◆ C++
- ◆ C
- ◆ PHP
- ◆ Perl
- ◆ Python
- ◆ Ruby

MySQL コネクタは、MySQL Web サイトから自由にダウンロードして使用できます。現在の URL の場所は、<http://www.mysql.com/products/connector/> です。

アプリケーション開発ツールおよびビジネスインテリジェンスツール

InfiniDB の MySQL フロントエンドにより、InfiniDB では、現在 MySQL をサポートするアプリケーション開発ツールまたはビジネスインテリジェンスツールが動作します。Eclipse、Microsoft Visual Studio など、多くのアプリケーション開発ツールが MySQL をサポートしており、InfiniDB で正常に動作します。

現在、MySQL に接続するビジネスインテリジェンスツールには、Business Objects、Cognos、Microsoft Analysis Services など、主要ベンダー独自仕様のものがすべて含まれます。Pentaho、Jaspersoft、Talend などのオープンソースの BI ツールも Calpont で動作します。

パフォーマンスのチューニング

この項では、InfiniDB データベースの監視およびチューニングについて説明します。

監視および診断ユーティリティ

製品版の InfiniDB に含まれる `cmconsole` ユーティリティを使用すると、DBA または `sysadmin` は、単一または複数ノードの InfiniDB 構成で様々な監視操作を実行できます。DBA は、`cmconsole` ユーティリティを使用して、ユーザーモジュールおよびパフォーマンスモジュールが動作するデータベースとサーバーの両方を監視できます。

各種の監視および診断コマンドのリストを次に示します。これらのコマンドの多くは、その名前から内容がわかるようになっています。

コマンド	説明
?	コンソールコマンドに関するヘルプを表示します
<code>addDbroot</code>	DBRoot ディスクストレージを Calpont InfiniDB システムに追加します
<code>addExternalDevice</code>	外部デバイスを構成ファイルに追加します
<code>addModule</code>	Calpont InfiniDB システム内でモジュールを追加します
<code>alterSystem-disableModule</code>	モジュールを無効にして Calpont InfiniDB システムを変更します
<code>alterSystem-enableModule</code>	モジュールを有効にして Calpont InfiniDB システムを変更します
<code>assignDbrootPmConfig</code>	割り当てられていない DBroot をパフォーマンスモジュールに割り当てます
<code>assignElasticIPAddress</code>	Amazon Elastic IP アドレスをモジュールに割り当てます

コマンド	説明
disableLog	プロセスおよびデバッグのログギングのレベルを無効にします
enableLog	プロセスおよびデバッグのログギングのレベルを有効にします
exit	コンソールツールを終了します
getActiveAlarms	アクティブなアラームリストを取得します
getActiveSQLStatements	システム内でアクティブな SQL 文のリストを取得します
getAlarmConfig	アラームの構成情報を取得します
getAlarmHistory	システムアラームを取得します
getAlarmSummary	アクティブなアラームの集計数を取得します
getCalpontSoftwareInfo	Calpont InfiniDB RPM の詳細情報を取得します
getExternalDeviceConfig	外部デバイスの構成情報を取得します
getLogConfig	システムログファイルの構成を取得します
getModuleConfig	モジュール名の構成情報を取得します
getModuleCpu	モジュールの CPU 使用率を取得します
getModuleCpuUsers	CPU を利用している上位のモジュールプロセスを取得します
getModuleDisk	モジュールのディスク使用率を取得します
getModuleMemory	モジュールのメモリー使用率を取得します
getModuleMemoryUsers	メモリーを利用している上位のモジュールプロセスを取得します
getModuleResourceUsage	モジュールのリソース使用率を取得します
getModuleTypeConfig	モジュールタイプの構成情報を取得します
getProcessConfig	プロセスの構成情報を取得します
getProcessStatus	Calpont InfiniDB プロセスの状態を取得します
getStorageConfig	システムのストレージ構成情報を取得します
getStorageStatus	システムのストレージステータスを取得します
getSystemConfig	システムの構成情報を取得します

コマンド	説明
getSystemCpu	すべてのモジュールにおけるシステムの CPU 使用率を取得します
getSystemCpuUsers	CPU を利用している上位のシステムプロセスを取得します
getSystemDisk	すべてのモジュールにおけるシステムのディスク使用率を取得します
getSystemInfo	システム全体の状態を取得します
getSystemMemory	すべてのモジュールにおけるシステムのメモリー使用率を取得します
getSystemMemoryUsers	メモリーを利用している上位のシステムプロセスを取得します
getSystemNetworkConfig	システムのネットワーク構成情報を取得します
getSystemResourceUsage	すべてのモジュールにおけるシステムのリソース使用率を取得します
getSystemStatus	システムおよびモジュールの状態を取得します
help	コンソールコマンドに関するヘルプを表示します
monitorAlarms	リアルタイムモードでアラームを監視します
movePmDbrootConfig	DBRoot をパフォーマンスモジュールから別のモジュールに移動します
quit	コンソールツールを終了します
removeDbroot	DBRoot ディスクストレージを Calpont InfiniDB システムから削除します
removeExternalDevice	外部デバイスを構成ファイルから削除します
removeModule	モジュールを Calpont InfiniDB システムから削除します
resetAlarm	アクティブなアラームをリセットします
restartSystem	停止またはシャットダウンされた Calpont InfiniDB システムを再起動します
resumeDatabaseWrites	Calpont InfiniDB データベースへの書込みの実行を再開します

コマンド	説明
setAlarmConfig	アラームの構成パラメータを設定します
setExternalDeviceConfig	外部デバイスの構成パラメータを設定します
setModuleTypeConfig	モジュールタイプの構成パラメータを設定します
setProcessConfig	プロセスの構成パラメータを設定します
setSystemConfig	システムの構成パラメータを設定します
shutdownSystem	Calpont InfiniDB システムをシャットダウンします
startSystem	停止またはシャットダウンされた Calpont InfiniDB システムを起動します
stopSystem	Calpont InfiniDB システムの処理を停止します
suspendDatabaseWrites	Calpont InfiniDB データベースへの書き込みの実行を一時停止します
switchParentOAMModule	アクティブな親 OAM モジュールを別のパフォーマンスモジュールに切り替えます
system	システムシェルコマンドを実行します
unassignElasticIPAddress	Amazon Elastic IP アドレスの割り当てを解除します

コマンドのヘルプを参照するには、「**help**」に続けてコマンド名を入力します

cmconsole を使用して、多数のシステム診断を監視および分析できます。問合せアクティビティ、システム全体の状態、ディスク使用率、CPU とメモリーのリソース消費などの問合せおよびレポートが可能です。たとえば、DBA または sysadmin は、次のコマンドを実行して、構成全体の状態をすばやく取得できます。

```
calpont> getSystemStatus
getsystemstatus    Wed Oct 26 12:34:28 2012

System qperfd01

System and Module statuses

Component          Status              Last Status Change
-----
```

```

System      ACTIVE      Wed Oct  7 15:10:47 2009
Module um1   ACTIVE      Tue Oct  6 15:50:09 2009
Module pm1   ACTIVE      Tue Oct  6 15:50:19 2009
Module pm2   ACTIVE      Wed Oct  7 15:10:45 2009
Module pm3   MAN_DISABLED Tue Oct  6 13:25:09 2009
Module pm4   MAN_DISABLED Tue Oct  6 13:25:39 2009
Module pm5   MAN_DISABLED Tue Oct  6 11:19:38 2009
Module pm6   MAN_DISABLED Tue Oct  6 11:20:10 2009
Module pm7   MAN_DISABLED Tue Oct  6 11:20:44 2009
Module pm8   MAN_DISABLED Tue Oct  6 11:21:18 2009

```

Active Parent OAM Performance Module is 'pm1'

この出力は、現在、1つのユーザーモジュールおよび2つのパフォーマンスモジュールがオンラインのため、問合せに利用可能であるのに対し、パフォーマンスモジュール3-8が現在オフラインであることを示しています。DBAまたはsysadminは、構成内にあるいずれかのノードに関するCPU、ディスクおよびメモリーの消費を表示する場合、次のようなコマンドを入力できます。

```

calpont> getModuleCpu pm1
getmodulecpu    Wed Oct 26 12:35:36 2012

```

Module 'pm1' CPU Usage % = 47

```

calpont> getModuleDisk pm1
getmoduledisk   Wed Oct 26 12:36:02 2012
Module 'pm1' Disk Usage (in 1K blocks)

```

Mount Point	Total Blocks	Used Blocks	Usage %
/	66719320	57662764	87
/boot	101086	19987	20
/usr/local/Calpont/data1	1124868224	830641000	74
/usr/local/Calpont/data2	1124868224	835241732	75
/usr/local/Calpont/data3	1124868224	834809896	75
/usr/local/Calpont/data4	1124868224	833940516	75
/usr/local/Calpont/data5	1124868224	835826024	75
/usr/local/Calpont/data6	1124868224	833620112	75
/usr/local/Calpont/data7	1124868224	832053424	74
/usr/local/Calpont/data8	1124868224	832324408	74
/home/qa/srv	235332176	88043896	38

```

calpont> getModuleMemory pm1
getmodulememory Wed Oct 26 12:37:20 2012

```

Module Memory Usage (in K bytes)

Module	Mem Total	Mem Used	cache	Mem Usage %	Swap Total	Swap Used	Swap Usage%
pm1	16440436	15472352	136236	94	2031608	0	0

cmconsole を使用してローカルマシンおよびリモートマシンを監視することで時間を節約することができます。さらに、これによって、DBA および sysadmin は、1 つのツールでデータベースと物理サーバーの両方の診断を実行できます。

アラームおよび通知

InfiniDB では、DBA または sysadmin は、データベースまたはシステムで障害が発生する前に問題の通知を受けるために使用可能な各種のシステムアラームを構成することもできます。管理者は、ディスク、CPU、メモリー、システムのスワップ領域、非互換のソフトウェアバージョンなどに対してしきい値を設定できます。

アラームおよび通知は、cmconsole ユーティリティを使用して管理します。管理者は、次のコマンドを入力すると、構成可能なアラームのリストを表示できます。

```
calpont> getAlarmConfig
getalarmconfig   Wed Oct  7 15:45:50 2009

Alarm Configuration

Alarm ID #1 Configuration information
BriefDesc = CPU_USAGE_HIGH
DetailedDesc = The usage on the indicated CPU has exceeded its
high threshold
Severity = CRITICAL
Threshold = 100

Alarm ID #2 Configuration information
BriefDesc = CPU_USAGE_MED
DetailedDesc = The usage on the indicated CPU has exceeded its
medium threshold
Severity = MAJOR
Threshold = 70

Alarm ID #3 Configuration information
BriefDesc = CPU_USAGE_LOW
DetailedDesc = The usage on the indicated CPU has exceeded its
low threshold
```

```
Severity = MINOR
```

```
Threshold = 50
```

```
Alarm ID #4 Configuration information
```

```
BriefDesc = DISK_USAGE_HIGH
```

```
DetailedDesc = The usage on the indicated Disk Drive has  
exceeded its high threshold
```

```
Severity = CRITICAL
```

```
Threshold = 100
```

```
.
```

```
.
```

```
.
```

管理者は、setAlarmConfig コマンドを使用して、アラーム番号とそのしきい値を指定することによって、アラームの現在の設定を変更できます。たとえば、約 95% に達した CPU に関して通知する CPU_USAGE_HIGH アラームを設定する場合、管理者は次のコマンドを実行します。

```
calpont> setAlarmConfig 1 Threshold 95
```

```
setalarmconfig    Wed Oct   7 15:55:47 2009
```

```
Successfully set Threshold = 95
```

トリガーされたアラームは、getActiveAlarms コマンドを使用して表示できます。

```
calpont> getActiveAlarms
```

```
AlarmID           = 1
```

```
Brief Description = CPU_USAGE_HIGH
```

```
Alarm Severity    = CRITICAL
```

```
Time Issued       = Tue Oct   6 14:50:06 2009
```

```
Reporting Module  = pm2
```

```
Reporting Process = ServerMonitor
```

```
Reported Device   = CPU
```

```
AlarmID           = 30
```

```
Brief Description = CONN_FAILURE
```

```
Alarm Severity    = CRITICAL
```

```
Time Issued       = Tue Oct   6 15:48:08 2009
```

```
Reporting Module  = um1
```

```
Reporting Process = ExeMgr
```

```
Reported Device   = 10.100.7.5 PrimProc
```

Nagios、HP OpenView または Tivoli などのグローバルシステム監視ソフトウェアに SNMP トラップを送信するようにアラーム通知を構成することもできます。

SQL の監視および診断ユーティリティ

適切に記述されていない SQL は、多くの場合、データベースでパフォーマンスの問題が発生する 1 番の原因となります。適切に実行されていない SQL 文を特定して修正できるように、InfiniDB では多数のコマンドおよびユーティリティが提供されています。

まず、一般的な MySQL サポートの観点では、ユーザーは、ログオンしているユーザーとそのユーザーが現在実行している SQL に関する簡単な情報を `show processlist` コマンドを使用して表示できます。

```
mysql> show processlist;
+----+-----+-----+-----+-----+-----+-----+-----+
| Id | User | Host       | db    | Command | Time | State | Info                |
+----+-----+-----+-----+-----+-----+-----+-----+
| 73 | root | localhost | ssb10 | Query   | 0    | NULL  | show processlist    |
+----+-----+-----+-----+-----+-----+-----+-----+
1 row in set (0.01 sec)
```

InfiniDB の `cmconsole` ユーティリティでも、データベースで現在実行されている SQL を表示する `getActiveSQLStatements` コマンドが提供されています。

```
calpont> getActiveSQLStatements
getactivesqlstatements   Wed Oct  7 08:38:32 2009
```

```
Get List of Active SQL Statements
=====
```

Start Time	Time (hh:mm:ss)	Session ID	SQL Statement
Oct 7 08:38:30	00:00:03	73	select c_name, sum(lo_revenue) from customer, lineorder where lo_custkey = c_custkey and c_custkey = 6 group by c_name

個々の SQL 文のパフォーマンスに対してトラブルシューティングを行う場合、使用可能なユーティリティは 3 つあります。1 つ目の標準 MySQL の `EXPLAIN` ユーティリティは、InfiniDB では索引や MySQL I/O 機能が使用されないため、若干有用性に欠けます。ただし、次の例に示すように、`EXPLAIN` でいくつかの情報を取得できます。

```
mysql> explain select select count(*) from dateinfo, customer, part,  
supplier, lineorder where s_suppkey = lo_suppkey and d_datekey = lo_or and
```



```
p_partkey = lo_partkey and c_custkey = lo_custkey and s_nation =
'BRAZIL';
```

id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	dateinfo	ALL	NULL	NULL	NULL	NULL	2000	
1	SIMPLE	customer	ALL	NULL	NULL	NULL	NULL	2000	Using join buffer
1	SIMPLE	part	ALL	NULL	NULL	NULL	NULL	2000	Using join buffer
1	SIMPLE	supplier	ALL	NULL	NULL	NULL	NULL	2000	Using where with pushed condition; Using join buffer
1	SIMPLE	lineorder	ALL	NULL	NULL	NULL	NULL	2000	Using where; Using join buffer

InfiniDB が提供する SQL 統計コマンド (calgetstats) およびトレースユーティリティ (caltrace) を使用すると、SQL 文のパフォーマンスに関して、より高度な診断情報を取得できます。

DBA または開発者は、calgetstats をコールするだけで、SQL インタフェースを介して実行された問合せの SQL パフォーマンスの統計を表示できます。次に例を示します。

```
mysql> select c_name, sum(l_quantity) from customer, orders,
lineitem where c_custkey = o_custkey and l_orderkey = o_orderkey
and l_shipdate = '1992-01-02' group by c_name;
+-----+-----+
| c_name          | sum(l_quantity) |
+-----+-----+
| Customer#000094199 |          35.00 |
| Customer#00009263  |          19.00 |
.
.
17 rows in set (1.21 sec)

mysql> select calgetstats()\G
***** 1. row *****
calgetstats(): Query Stats: MaxMemPct-0; NumTempFiles-0;
TempFileSpace-0MB; ApproxPhyI/O-0; CacheI/O-7977; BlocksTouched-
7977; PartitionBlocksEliminated-0; MsgBytesIn-5MB; MsgBytesOut-
0MB; Mode-Distributed| 1255540221 457489
1 row in set (0.00 sec)
```

統計出力には、次の情報が含まれます。

- ◆ MaxMemPct：大規模なハッシュ結合操作のサポート時など、ユーザーモジュールのピーク時のメモリー利用率。
- ◆ NumTempFiles：一時ファイルに関するレポート。一時ファイルは、利用可能なメモリーを超える大きさの問合せ操作（通常は、比較的小さい表結合カーディナリティが、構成可能な何らかのしきい値を超える特殊な結合操作）をサポートするために作成されます。
- ◆ TempFileSpace：一時ファイルが使用する領域に関するレポート。一時ファイルは、利用可能なメモリーを超える大きさの問合せ操作（通常は、比較的小さい表結合カーディナリティが、構成可能な何らかのしきい値を超える特殊な結合操作）をサポートするために作成されます。
- ◆ PhyI/O：ディスク、SSD、その他の永続ストレージから読み取られる 8K ブロックの数。ほとんどの場合、個々の I/O 操作の数はブロックの数に比べて非常に少ないため、Calpont InfiniDB では 1 回の I/O 操作で 512 個のブロックが同時に読み取られます。
- ◆ CacheI/O：必要な個別の物理 I/O コールの数により下方調整されている、メモリー内で処理された 8K ブロックの概数。
- ◆ BlocksTouched：メモリー内で処理された 8K ブロックの概数。
- ◆ PartitionBlocksEliminated：InfiniDB がエクステンツマップによって除外したブロックアクセス数。
- ◆ MsgBytesIn、MsgByteOut：問合せをサポートするためにノード間で送信されるメッセージのバイト数（MB 単位）。

InfiniDB の SQL トレースユーティリティ caltrace の基本的な使用手順は次のとおりです。

1. 特定のセッションのトレースを開始します
2. 問題の SQL 文を実行します
3. この文に対して収集されたトレースを確認します

例として、次のセッションでトレースを開始して、6,000,000 行のファクト表および 300,000 行のディメンション表に対して問合せを発行し、トレースからの出力を確認します。

```
mysql> select calsettrace(1);
+-----+
| calsettrace(1) |
+-----+
|                1 |
```

```

+-----+
1 row in set (0.04 sec)

mysql> select c_name, sum(lo_revenue)
      -> from customer, lineorder
      -> where lo_custkey = c_custkey and
      -> c_custkey = 5
      -> group by c_name;
+-----+
| c_name          | sum(lo_revenue) |
+-----+
| Customer#000000005 | 552483078.00 |
+-----+
1 row in set, 0 warning (1.23 sec)

```

```

mysql> select calgettrace();
+-----+
-----
-----
-----
-----
-----
-----
-----+
| calgettrace()
|
+-----+
-----
-----
-----
-----
-----
-----
-----+
|
Desc Mode Table      TableOID ReferencedOIDs PIO  LIO  PBE Elapsed Rows
BPS  PM  customer  3327      (3328,3336)    0   150   0   0.017   1
BPS  PM  lineorder 3278      (3281,3291)  4096 29808 0   0.766  63
HJS  PM  lineorder 3278      -          -    -    -    0.000   -
ADS  UM  -          -          -          -    -    -    0.724   -
|
+-----+
-----
-----
-----
-----
-----
-----
-----+
1 row in set (0.02 sec)

```

各ステップの物理および論理 I/O、経過時間などに関する情報が表示されます。

InfiniDB への移行

この項では、他のデータベースから InfiniDB データベースに移行する方法について説明します。

一般的な移行の概念

通常、InfiniDB への移行では、別のデータベースからのスキーマまたはデータの移動、および場合によってはストアドプロシージャなどのストアドコードオブジェクトの移動を実行します。通常、次のロードマップに従います。

1. ソースデータベースオブジェクトのドキュメント化
2. InfiniDB オブジェクトの設計
3. ソースシステムから InfiniDB へのデータの移動
4. InfiniDB へのコードオブジェクトの変換（オプション）

既存のデータベース構造のドキュメント化は、通常は手動のデータディクショナリ検索などを実行する標準の手動処理で試行すると困難な場合があります。ほとんどのデータベースには操作に役立つメタデータディクショナリが含まれていますが、手動ですべてのメタデータ（表、列など）を取得する処理は、大規模なデータベースでは非常に時間がかかるイベントとなる場合があります。

ほとんどの場合の最良の方法は、変換に必要で適用可能なすべてのメタデータを自動的にカタログ化する、コンピュータによるリバースエンジニアリング処理を使用する方法です。これには、PowerDesigner、ERWin および ER/Studio などの適切なサードパーティのデータモデリングツールを使用できます。これらのすべてのツールで複数のデータベースタイプのリバースエンジニアリングがサポートされています。

InfiniDB ではほとんどの重要なデータ型がサポートされているため、通常、InfiniDB へのスキーマ構造の移動は困難なタスクではありません。InfiniDB でのデータ型のサポートの詳細は、『InfiniDB SQL 構文ガイド』を参照してください。

ソースデータベースのメタデータを取得して理解したら、次の手順は InfiniDB ターゲットデータベースの設計です。これには、ソースオブジェクトおよびそのプロパティ（列のデータ型など）を InfiniDB の対応するものに変換することが含まれます。一部のモデリングツールでは、ユーザーは基礎となるデータベースを物理的な設計で変更するオプションを選択することで、あるデータベーススキーマを別のデータベー

スキーマに自動的に変換できます。すべてのデータ型が自動的に変更されますが、データ型が 1 対 1 で一致しない場合、一部のデータ型に手動での調整が必要なことがあります。

ストアドプロシージャなどのコードオブジェクトの変換は別問題です。Oracle、Sybase、SQL Server などのストアドプロシージャから MySQL のストアドプロシージャに変換する、Ispirer の SQLWays 製品などのいくつかのサードパーティツールがあります。

スキーマオブジェクトの変換の場合は、可能性がある 1 つの方法として、無償の MySQL Migration Toolkit の使用があります。MySQL Migration Toolkit を使用すると、ツールによってリバースエンジニアリングされたソースデータベースが MySQL ベースの対応するものに自動的に変換され、すべてのデータ型変換が行われます。ユーザーは、ソースデータベースから移行するオブジェクトおよび MySQL の特定の指定方法に関して包括的に制御できます。

Calpont 用の無償の MySQL Migration Toolkit を使用する場合、通常、ユーザーは次の手順に従います。

1. Migration Toolkit をダウンロードしてインストールします。MySQL の Web サイトの GUI ツールのダウンロードセクション（現在、<http://dev.mysql.com/downloads/gui-tools/5.0.html>）にあります。
2. ツールを使用してソースデータベースに接続します。
3. ツールを使用して InfiniDB ターゲットデータベースに接続します。
4. InfiniDB に移行するスキーマおよびオブジェクトを選択します。
5. オブジェクトマッピングの画面で、ユーザー定義のラジオボタンを選択して、作成する SQL スクリプトに「**engine=infinidb**」と入力します。
6. オブジェクトをオンラインで自動的に作成するのではなく、SQL の作成をスクリプトに送信することを選択します。
7. エディタで作成スクリプトを確認し、InfiniDB で必要なデータ型またはその他の変更を行います。
8. InfiniDB SQL エディタでスクリプトを実行して InfiniDB 表オブジェクトを作成します。

MySQL のデータの移行方法

汎用的な MySQL データベースのデータを InfiniDB に移動するには、いくつかの異なる方法を使用できます。1 つ目では、ユーザーは MySQL の `SELECT INTO OUTFILE` コマンドを使用して一連の表出力をフラットファイルにスプールできます。たとえば、ユーザーは、2 つの列がある「t1」という表から、パイプ (|) 文字で区切られたファイルにデータを書き出す必要がある場合、次のように行うことができます。

```
mysql> select c1, c2
-> from t1
-> into outfile 't1.tbl' fields terminated by '|';
```

MySQL ファイルへの書込みが行われたら、そのファイルを InfiniDB インポートディレクトリに移動し、`cpimport` ユーティリティを使用して InfiniDB にインポートできます。

もう 1 つの方法では、無償のオープンソース ETL（抽出 - 変換 - ロード）ツールを使用して、ある MySQL サーバーから、ある InfiniDB サーバーにデータを移動します。ツールは Pentaho、Jaspersoft、Talend などから入手できます。

短時間で表全体からデータを移動して InfiniDB にインポートする 3 つ目の方法では、MySQL の CSV ストレージエンジン（MySQL 5.1 以上で使用可能）を使用します。カンマ区切り値ストレージエンジンによって、ファイル内の列がカンマで区切られているフラットファイルに表のデータが格納されます。ユーザーは、InfiniDB 以外のインストール環境内の既存の MySQL 表を使用し、現行のエンジンではなく CSV エンジンを使用するようにその表を変更して表ファイルを InfiniDB サーバーにコピーした後で、表を元のエンジンに戻すことができます。唯一の注意事項としては、CSV ストレージエンジンでは NULLABLE 列定義はサポートされていないため、列に NULL 制約が関連付けられていないようにする必要があります。

ユーザーは、InfiniDB サーバーにカンマ区切り形式の表ファイルを含めた後、そのファイルを InfiniDB インポートディレクトリに移動し、`cpimport` ユーティリティを使用して InfiniDB にロードできます。

Oracle のデータの移行方法

Oracle のデータを InfiniDB に移動する方法は多くあります。MySQL に接続可能なすべてのデータ移動ツールが Calpont へのデータの移動に適しています。たとえば、DBA は、必要に応じてデータを移動し、変換を実行するのに、Talend または Pentaho の Kettle などのオープンソース ETL ツールを自由にダウンロードして使用できます。

Oracle ではデータベース用のアンロードツールは提供されていないため、DBA は Oracle データをアンロードするためのサードパーティのアンロードツールを所有していないかぎり、SQL*Plus セッションからフラットファイルをスプールアウトする必要

がある場合があります。たとえば、DBA は、3 つの列がある「workorder」という表を含む InfiniDB にロードするために Oracle からフラット表を作成し、パイプ (|) 文字でデータを区切る場合、次のようなスクリプトを使用できます。

```
SET TERMOUT OFF
set newpage 0
set space 0
set linesize 500
set pagesize 0
set echo off
set feedback off
set heading off
set trimspool on
col wonum      format a10
col leadcraft format a8
col location   format a50
spool c:\workorder.txt
SELECT
WORKNUM || '|' ||
CONTACT || '|' ||
PRIMARY_LOCATION
FROM SCHEMA1.WORKORDER;
spool off
exit
```

Microsoft SQL Server のデータの移行方法

Microsoft では、SQL Server 製品との統合サービスが提供されています。これは、バンドルされたデータの移動および BI のプラットフォームです。SQL Server の DBA は、これを使用して InfiniDB に接続し、InfiniDB にデータを移動できます。

Oracle とは異なり、SQL Server にはデータベース用の bcp という高速アンロードツールがあります。bcp ユーティリティを使用して SQL Server 表データのフラットファイルを作成し、cpimport ユーティリティを使用して InfiniDB にロードできます。大量のデータに対しては、SQL Server データを bcp で出力し、cpimport で InfiniDB にインポートする方法が最良の方法であると考えられます。

GNU Free Documentation License

GNU Free Documentation License
Version 1.3, 3 November 2008

Copyright © 2000, 2001, 2002, 2007, 2008 Free Software Foundation, Inc. <<http://fsf.org/>>

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondly, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

The "publisher" means any person or entity that distributes copies of the Document to the public.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each

Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

- ◆ A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.
- ◆ B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.
- ◆ C. State on the Title page the name of the publisher of the Modified Version, as the publisher.
- ◆ D. Preserve all the copyright notices of the Document.
- ◆ E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
- ◆ F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
- ◆ G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
- ◆ H. Include an unaltered copy of this License.
- ◆ I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document

as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

- ◆ J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.
- ◆ K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.
- ◆ L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
- ◆ M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.
- ◆ N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.
- ◆ O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties—for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations

requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See <http://www.gnu.org/copyleft/>.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free

Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

11. RELICENSING

"Massive Multiauthor Collaboration Site" (or "MMC Site") means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A "Massive Multiauthor Collaboration" (or "MMC") contained in the site means any set of copyrightable works thus published on the MMC site.

"CC-BY-SA" means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

"Incorporate" means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is "eligible for relicensing" if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.