

# 分散データベース

Hiroshima Institute of Technology

1

## 授業計画

- 第1回 ガイダンス・データベースの基本概念
- 第2回 データモデル
- 第3回 関係代数
- 第4回 データベース設計
- 第5回 リレーションの正規化
- 第6回 中間まとめ
- 第7回 関係データベース言語(SQL1)
- 第8回 関係データベース言語(SQL2)
- 第9回 計算機実習
- 第10回 データの検索機構 MySQL実習
- 第11回 トランザクション管理 MySQL実習
- 第12回 障害回復 MySQL実習
- 第13回 分散データベース MySQL実習
- 第14回 期末まとめ
- 第15回 応用技術と将来動向 MySQL実習

## 11. 分散データベース

### 講義内容

- 分散データベースの概要
- 透過性
- テーブルの結合方法
- 分散入れ子ループ法
- 分散ソートマージ法
- セミジョイン法
- コミットメントプロトコル

## 11. 分散データベース

### 1. 分散データベースの概要

- クライアントサーバ方式の分散処理形態が普及し、**データベースもネットワーク上に分散配置**して利用する方式がとられている。
- 分散データベースは、**1つのDBMSが複数のCPUに接続**されている記憶装置群を制御する形態のデータベースである。
- 物理的には同じ場所の複数台のコンピュータで構成される場合や、コンピュータネットワークで相互接続されたコンピュータ群に分散されている場合などがある。

## 11. 分散データベース

### 1. 分散データベースの概要

- 分散システムとは、**地理的または論理的に分散**した複数のシステムがコンピュータネットワークなどの通信手段を介して、個々のシステムが協調して何らかのまとまった処理を行うシステムである。
- 分散データベースシステムは、地理的または論理的に分散した複数のDBMSが通信手段を介して結合し、ユーザにはあたかも**1つのDBMSのように見せる**システムのことである。
- **分散データベース**とは、クライアントサーバ方式の分散ネットワーク上にデータベースサーバを分散配置し、**データ処理の垂直分散や水平分散**を実現するデータベースシステムのことである。

Hiroshima Institute of Technology

5

## 11. 分散データベース

### 1. 分散データベースの概要

#### 分散データベースの利点

<b>危険分散</b>	<ul style="list-style-type: none"><li>・災害やデータベースシステムの障害が発生しても、その影響を問題が発生したサイトに局所化することが可能となり、他のサイトのデータベースの利用を継続することが可能となる。</li><li>・データを複数のサイトに重複して格納しておくことによって、災害やデータベースシステムの障害が発生したサイトを除いてアクセスすることにより<b>サービスを継続</b>できる。</li></ul>
<b>負荷分散</b>	<ul style="list-style-type: none"><li>・データベースを複数のコンピュータで管理することになるので、<b>負荷を分散</b>することができる。</li><li>・負荷に偏りが発生した場合には、サイト間でデータの移動などによって負荷を均等化する余地がある。</li><li>・システムの拡張は、新たなデータベースシステムを分散データベースを構成するネットワークに追加すればよいので比較的容易である。</li></ul>

Hiroshima Institute of Technology

6

## 11. 分散データベース

### 1. 分散データベースの概要

- **垂直分散**(vertical distribution)は、**主従関係**のある分散データベース形態である。
- 垂直分散では、データベースに関して主サイトと従サイトの関係がある。
- 主サイトのデータベースは、クライアント機能をもたず従サイトのデータへのアクセスはできない。
- 従サイトのデータベースは、クライアントからのアクセスに対応するサーバ機能と、主サイトのデータベースをアクセスするクライアント機能から構成される。ただし、従サイト間のアクセスは行わない。
- 実現が容易で管理しやすいという特徴があるが、**主サイトのサーバにアクセスが集中**するので通信負荷が高く障害に比較的弱いという欠点がある。

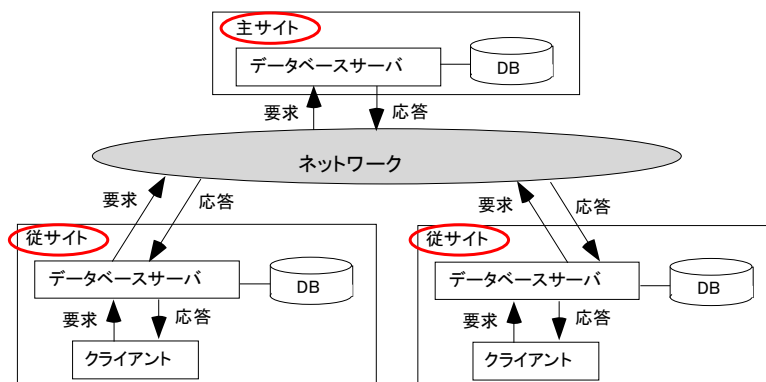
Hiroshima Institute of Technology

7

## 11. 分散データベース

### 1. 分散データベースの概要

#### 垂直分散の例



Hiroshima Institute of Technology

8

## 11. 分散データベース

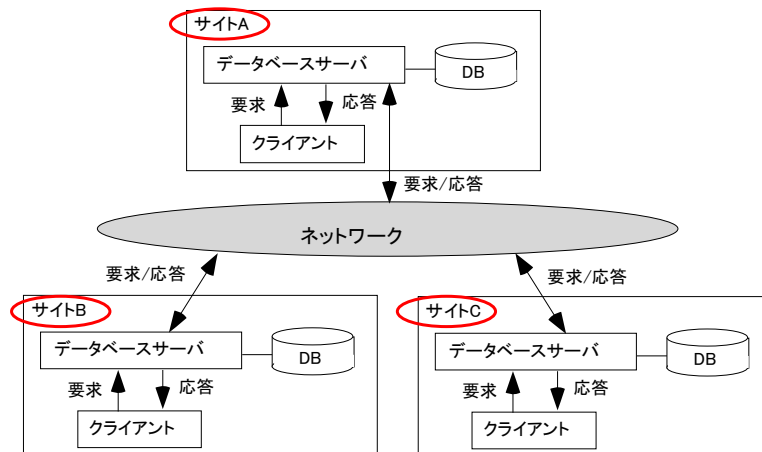
### 1. 分散データベースの概要

- **水平分散 (horizontal distribution)**は、**各サイトがサーバ機能とクライアント機能**をもち、各サイト間のデータベースが相互に均等な関係でアクセスしあう分散データベース形態である。
- 水平分散のデータベースシステムは、拡張性や耐障害性が高く、1つのサーバにトラフィックが集中しないという特徴があるが、設計面で考慮すべき課題が多く**管理がかなり複雑**になるという欠点がある。

## 11. 分散データベース

### 1. 分散データベースの概要

#### 水平分散の例



## 1 1. 分散データベース

### 2. 透過性

- 分散データベースにおいて重要なことは、ユーザに対しては分散されて結合されているシステム全体を1つのシステムとして見せることである。
- 分散データベースをあたかも集中データベースのようにユーザに見せることを**分散透過性**(distributed transparency)という。

分散透過性

アクセスに対する透過性

位置に対する透過性

移動に対する透過性

分割に対する透過性

重複に対する透過性

障害に対する透過性

Hiroshima Institute of Technology

11

## 1 1. 分散データベース

### 2. 透過性

#### 分散データベースの透過性

透 過 性	内 容
アクセスに対する透過性	ローカルサイトにあるデータでも、リモートサイトにあるデータでも、ユーザが同じ方法でアクセスできることである。
位置に対する透過性	データの存在する場所またはサイトをユーザが知らなくてもアクセスできることである。
移動に対する透過性	運用の都合や性能上の目的で、表を格納しているサイトを変更しても、ユーザにはその表の格納先の移動を意識させないことである。
分割に対する透過性	1つの表が複数のサイトに分割して格納されていても、ユーザにその表の分割を意識させないことである。
重複に対する透過性	1つの表が複数のサイトに重複して格納されていたり、サイト間でデータの複製が存在していても、ユーザにそれを意識させないことである。
障害に対する透過性	分散データベースを構成するいずれかの要素において障害が発生しても、それを隠ぺいすることである。

Hiroshima Institute of Technology

12

## 1 1. 分散データベース

### 3. テーブルの結合方法

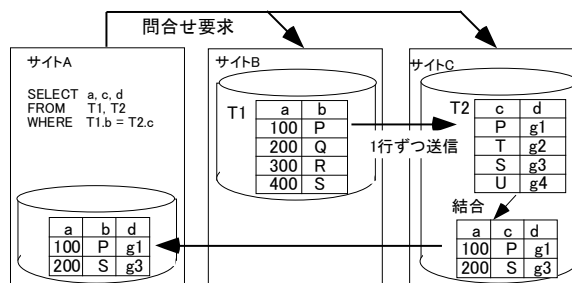
- 分散データベースでは、表が複数のサイトに存在する場合、複数のサイト間にまたがる結合演算が必要な場合がある。
- 複数サイト間の表を結合する場合には、通常の結合処理に加えて、**通信負荷を考慮**して結合方法を決定する必要がある。
- 分散データベースシステムにおける**結合演算方法**には、分散入れ子ループ法、分散ソートマージ法、セミジョイン(準結合)法などがある。

## 1 1. 分散データベース

### 3. テーブルの結合方法

#### ■ 分散入れ子ループ法

- 分散入れ子ループ法は、2つの表の結合において、一方の行を外側のループとして取り出したものに対して、他方の表の全ての行を内側のループとして照合して結合演算する方法である。

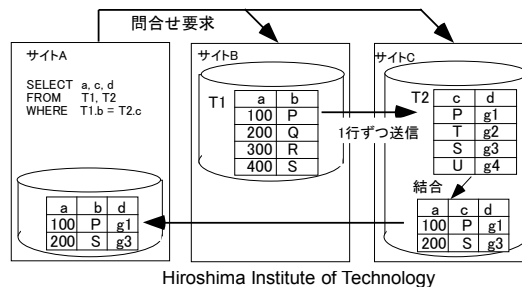


## 1 1. 分散データベース

### 3. テーブルの結合方法

#### ■ 分散入れ子ループ法

- 外側ループの表をもつサイトから行をネットワークを介して1行ずつ転送し、内側ループの表をもつサイトで結合演算を行う。
- 分散データベースシステムで入れ子ループ法を採用した場合、**外側ループの表の全行を内側ループの表をもつサイトに転送するため、外側ループの表の大きさがそのまま通信負荷となり、分散問合せの処理効率に大きな影響を与えることに注意が必要である。**



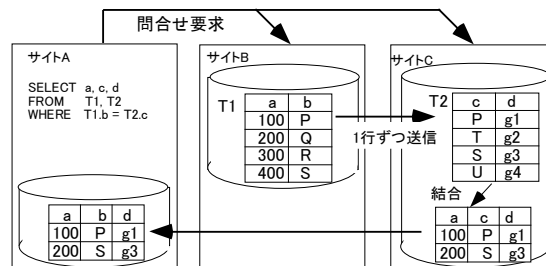
15

## 1 1. 分散データベース

### 3. テーブルの結合方法

#### ■ 分散入れ子ループ法

- ・サイトBの表を外側ループとして処理し、サイトCの表を内側ループとして処理している。
- ・サイトBは**行単位でサイトCに転送**し、サイトCでその行の結合演算が終わると、次の行を転送する処理を繰り返す。
- ・サイトCでは、サイトBから転送された行とサイトCの表を全行照合し、結合条件が成立する行との結合を行う。
- ・サイトBの表の全行の転送が終了すると、**サイトCでの全ての結合演算が完了**することになる。
- ・最後に、**この結果を分散問合せ処理を要求したサイトAに転送**するとトランザクションが完了する。



16



## 1 1. 分散データベース

### 3. テーブルの結合方法

#### ■ 分散ソートマージ法

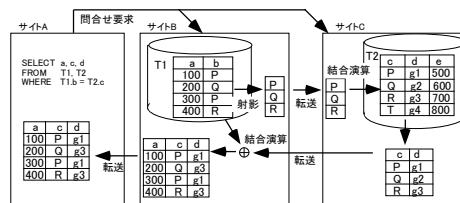
- 分散ソートマージ法は、結合演算対象となる表をもつ各サイトでそれぞれの表を結合対象の列でソートし、一方のサイトから他方のサイトへ表全体を転送後、マージ処理で結合演算する方法である。
- ソートマージ法は、結合対象の表が入出力バッファに収まらない場合に用いられる。
- 分散ソートマージ法は、入出力バッファに収まらないような大きな表全体を転送することになり、効率のよい方法ではない。

## 1 1. 分散データベース

### 3. テーブルの結合方法

#### ■ セミジョイン法

- セミジョイン法は、分散データベースシステムの結合演算において、その通信量を減らすために考案された結合処理の方法である。
- 結合対象の2つの表において、一方のサイトの表の結合対象列のみ他のサイトへ転送し結合演算をし、その結合結果を返信し、再度結合演算を行う。
- 分散データベースシステムでセミジョイン法を採用した場合、結合演算は2段階になるが、通信量が大幅に減少し、全体の処理効率が向上する。

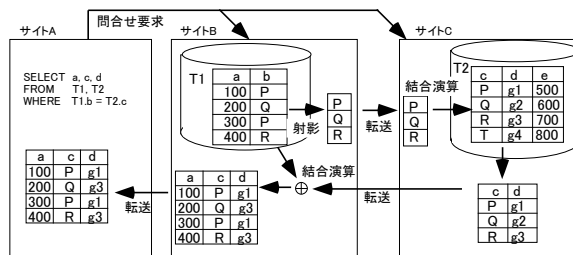


## 1 1. 分散データベース

### 3. テーブルの結合方法

#### ■ セミジョイン法

- ・サイトBの結合対象列を射影演算で抽出し、サイトCへ転送する。
- ・サイトCでは、受信した結合対象列と結合演算を行い、その結合結果をサイトBに返送する。
- ・サイトCから受け取った結合結果と再度結合演算を行い、最終的な結合結果を導き出して、分散問合せ要求したサイトAに返送する。



Hiroshima Institute of Technology

19

## 1 1. 分散データベース

### 4. コミットメントプロトコル

- 分散データベースにおいて、複数のデータベースを同時に更新する場合、一方のデータベースに障害が発生してもデータベース間の整合性を保つための制御方法に、**2相コミットメント**や**3相コミットメント**がある。

#### ■ 2相コミットメント

- 2相コミットメント(**2PC**)は、主サイトから複数の従サイトに直接コミット命令を発行するのではなく、2つのフェーズに分ける方法である。
- まず、主サイトが対象となる複数の従サイトに対し、“**コミット可否問合せ**”を行う。
- 更新準備が整っているサーバは“**準備完了**”の応答を返し、すべての従サイトが準備を終えたことを確認した上で、主サイトは“**コミット命令**”を通知し、データベースが一斉に書き換えられる。

Hiroshima Institute of Technology

20

## 11. 分散データベース

### 4. コミットメントプロトコル

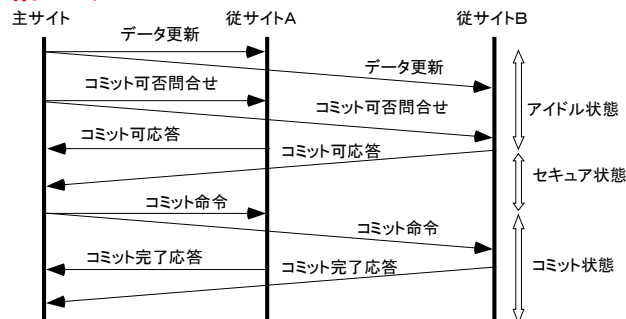
#### ■ 2相コミットメント

- 書き換え中にいずれかのデータベースで異常が発生した場合、異常が生じたサーバは失敗を伝え、主サイトはすべての従サイトに処理撤回を通知して、ロールバック処理を行なうように指示する。
- 一連の手順が**コミット準備**と**コミット実行**の2段階に分かれていることから、**2相コミットメント**と呼ばれる。
- 従サイトにおいて、コミット可否応答を返してからコミット命令を受けるまでの状態を**セキュア状態**という。
- セキュア状態では、更新処理の対象となるデータベースは全てコミットもロールバックも可能な状態である。

## 11. 分散データベース

### 4. コミットメントプロトコル

#### ■ 2相コミットメント



- 従サイトが**セキュア状態**になっている時点で、主サイトの障害や通信回線上の障害が発生してコミット命令やロールバック命令が通知できなくなる可能性がある。

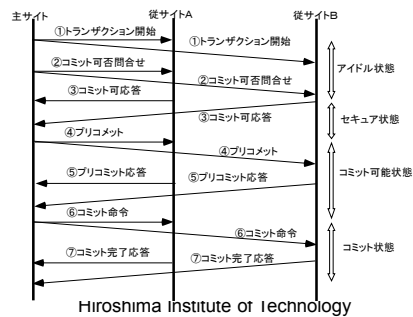
→ この場合には、従サイトはコミットもロールバックも行えない事態となる。

## 11. 分散データベース

### 4. コミットメントプロトコル

#### ■ 3相コミットメント

- 3相コミットメント(3PC)は、2相コミットメントの問題点を克服するために、**セキュア状態確認**後に、更に**プリコミット状態**で確認を取る方式である。
- 3相コミットメント制御は、データの一貫性を高度に高めるが、トランザクションの応答性能を悪化しネットワーク負荷上昇の問題もあり実用性には問題がある。



23

## 11. 分散データベース

### 講義内容

- 分散データベースの概要
- 透過性
- テーブルの結合方法
- 分散入れ子ループ法
- 分散ソートマージ法
- セミジョイン法
- コミットメントプロトコル

Hiroshima Institute of Technology

24