

障害回復

データベース論 I 第11回

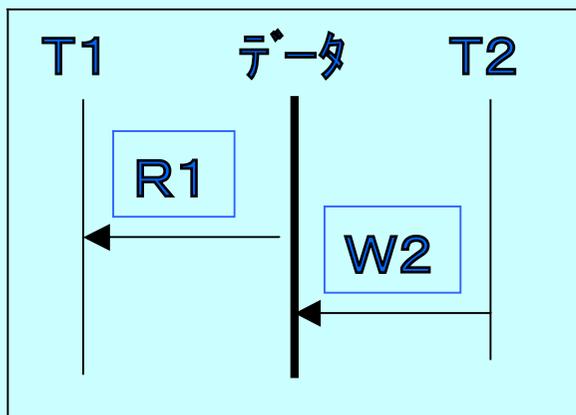
URL <http://homepage3.nifty.com/suetsuguf/>

作成者 末次文雄 ©

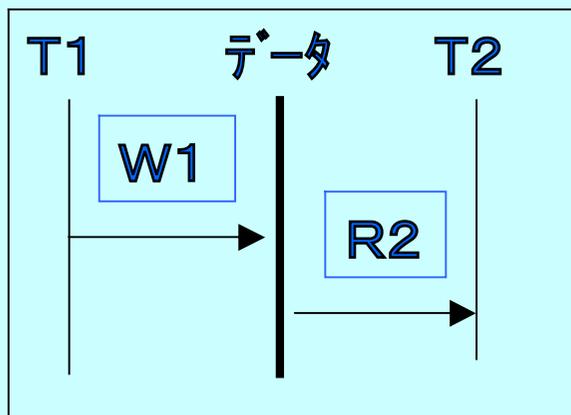
前回の復習（直列可能性、先行グラフ）

アクセス履歴を分析して、**競合する操作パターン**を見出し、**先行グラフ**を作成することにより、**直列可能性の判断**をする方法である。

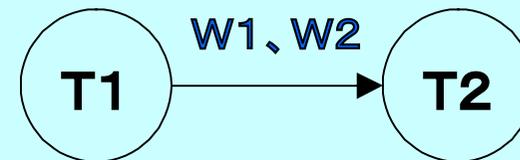
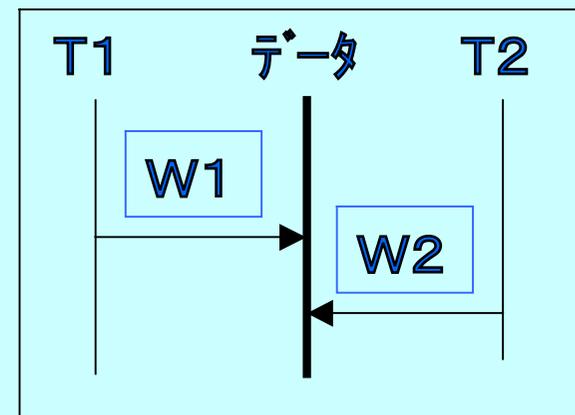
競合操作パターン1



競合操作パターン2



競合操作パターン3



前回の復習 (2相ロック)

2相の例示

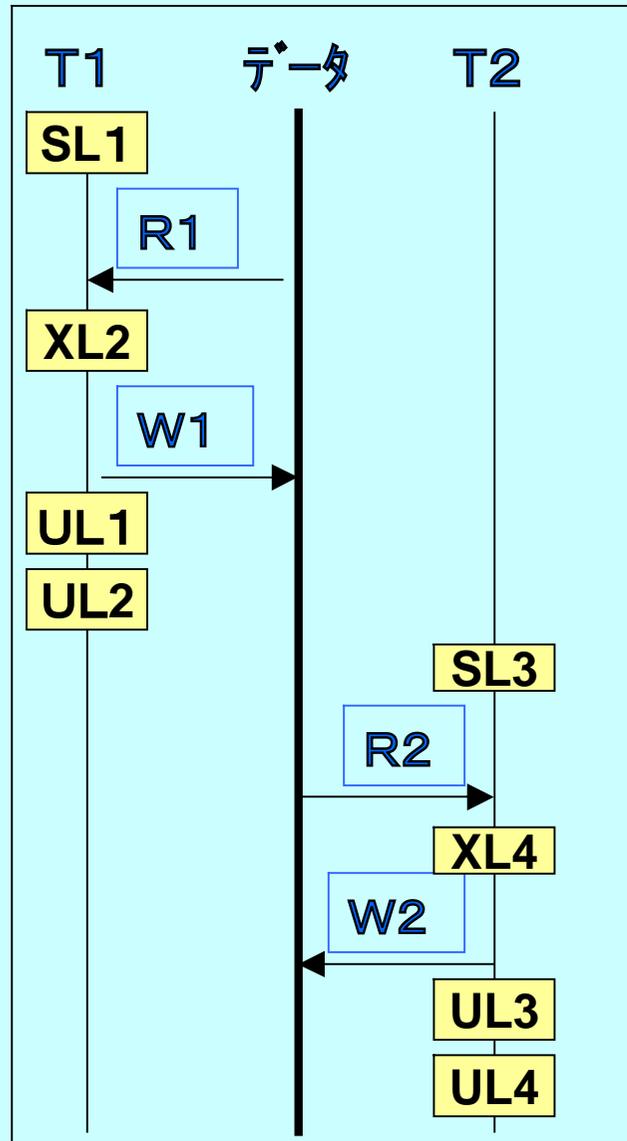
ロック相

解除相

SL: shared Lock

XL: exclusive Lock

UL: unlock



連鎖的アボートを防止するためには、UL前に、commitを取っておくことが望ましい。

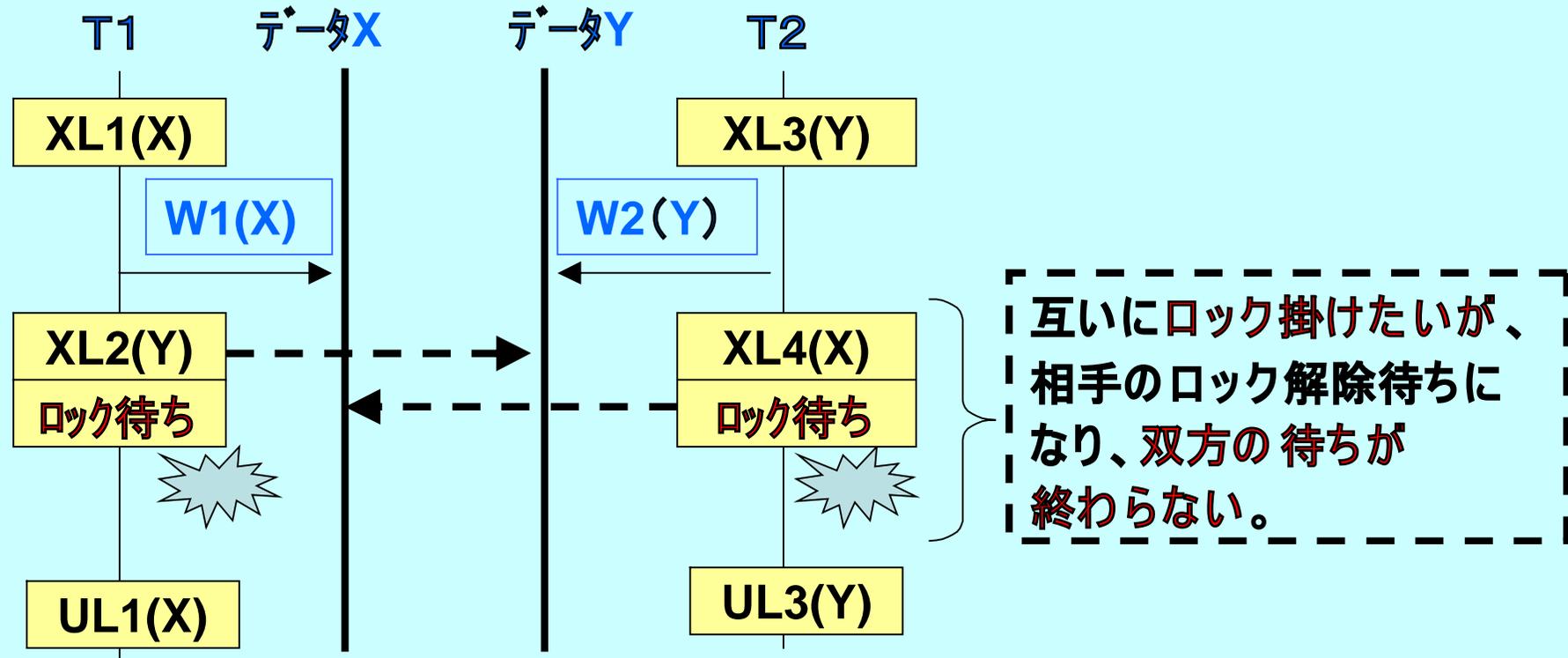
ロック相

解除相

前回の復習 (デッドロックの発生)

① デッドロックの発生原因

相互に複数のデータに対して**ロックを掛けたい時に**、**お互いが相手のロック解除を待つ状態になるため発生。**

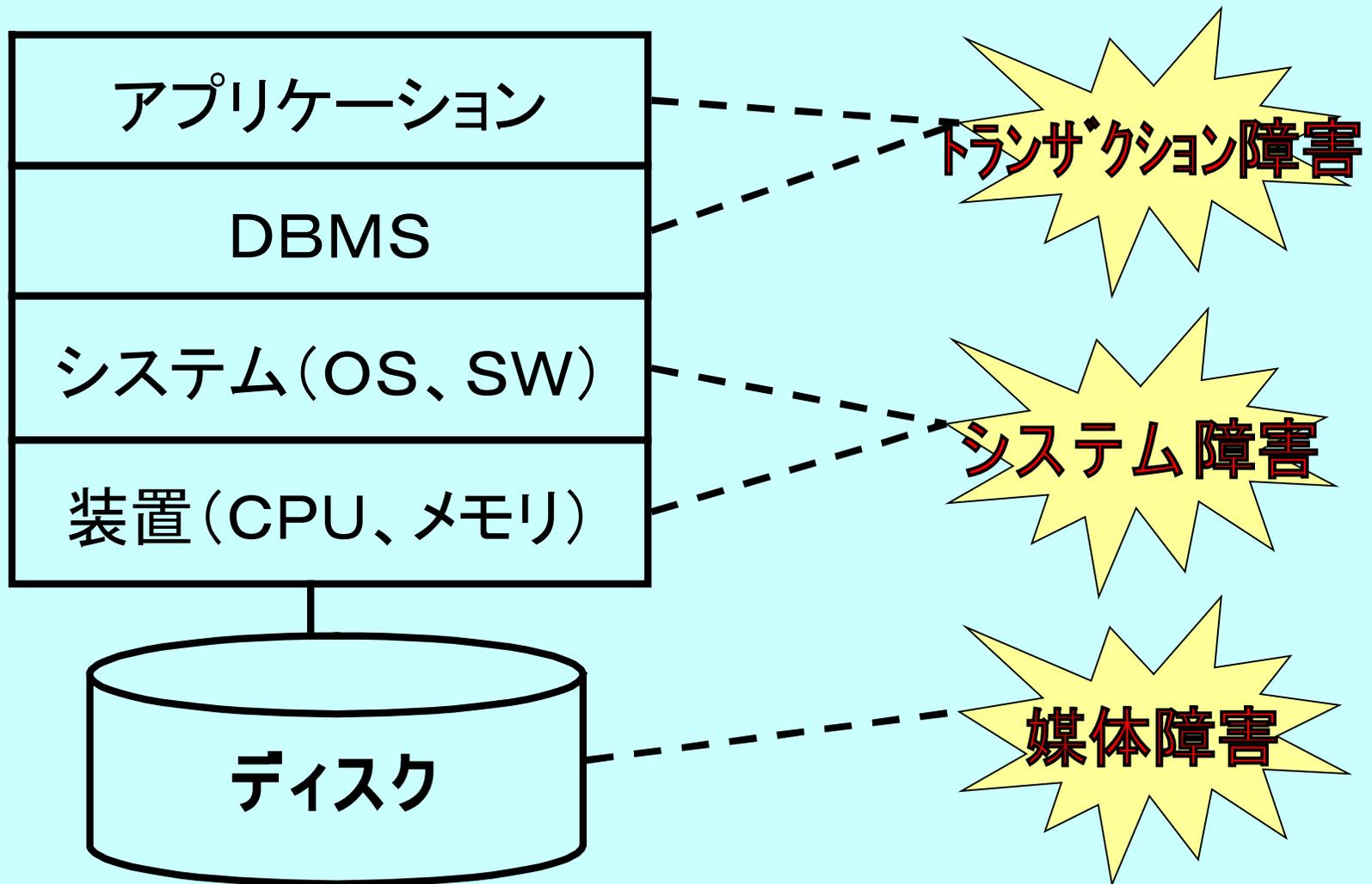


目次

1. 障害の分類
2. 障害回復の準備
3. ログファイルの利用法
4. 障害回復方法
5. DB保守管理
6. レポート課題
7. 参考書ほか

1. 障害の分類

1.1 障害の発生箇所



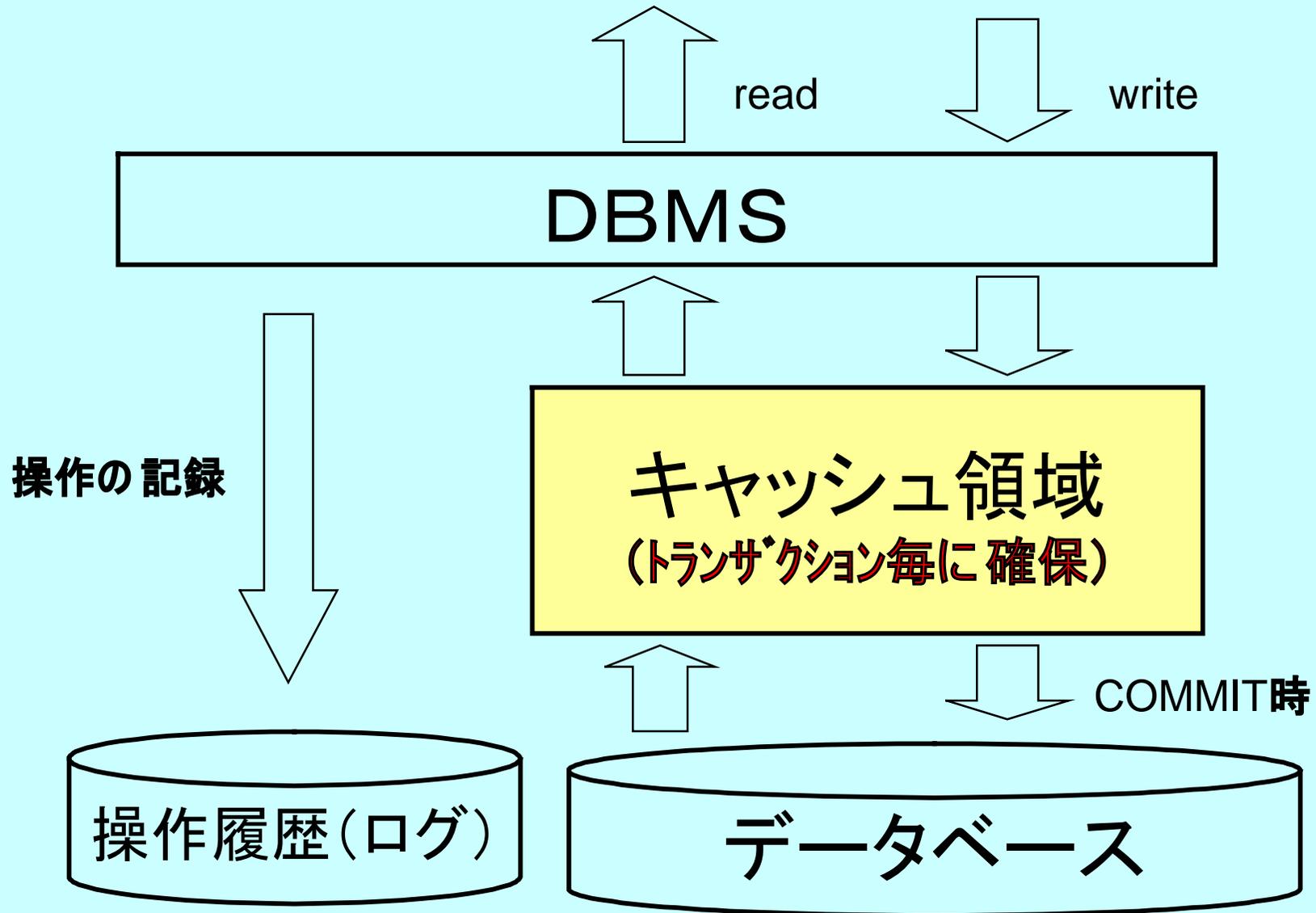
1. 2 障害の種類

種類	原因	対策
トランザクション障害 (transaction failure)	<ul style="list-style-type: none">・プログラム・エラー・デッドロック・・・	<ul style="list-style-type: none">・ロールバック処理
システム障害 (system failure)	<ul style="list-style-type: none">・OSダウン・装置障害・・・	<ul style="list-style-type: none">・リカバリー処理(復旧) (チェックポイント使用)
媒体障害 (media failure)	<ul style="list-style-type: none">・ディスクトラブル	<ul style="list-style-type: none">・バックアップファイルからのリストア(復元)・リカバリー処理(復旧)

1. 3 補足(キャッシュ領域)

- トランザクションのread、write操作時では、
 - ①直ちにデータベースに物理的に反映しない
 - ②**キャッシュ領域**に置かれ、
 - ③コミット命令時にデータベースに反映
(リカバリー時には、**コミット後か否かが重要**)
- キャッシュ領域(バッファ)により、ディスクアクセスが減り、**アクセス効率**が向上する。

1.3 続き (キャッシュ領域)



2. 障害回復の準備

- 障害発生時の状況
 - ー キャッシュ領域の更新記録は消滅
 - ー コミット後では、既にデータベース反映済であり、簡単には元に戻せない
 - ー 媒体障害時は、データベースが消滅
- 準備項目としては、
 - ー DBバックアップ取得
 - ー ログファイル取得
 - ー チェックポイントの設定

2. 1 バックアップファイルの取得

①バックアップの単位、タイミング

- ・DB単位 (INDEX含む)
- ・タイミング (定期的、大量更新直後)

②バックアップの種類

- ・full バックアップ (データベース全体)
- ・差分バックアップ (前回fullバックアップの差)

③媒体

- ・磁気テープ (推奨。複数世代の異地保管)
- ・ディスク (データ、ログとは装置を分ける)

④操作指示 (コマンド、ユーティリティ、ミラーリング)

2. 2 ログファイルの取得

① タイミング

- ・DB更新時 (INDEX含む)

② 方式

- ・先書出し方式 (バッファを持たない)
- ・ログバッファ方式 (処理効率向上)

③ 媒体

- ・通常はディスク (データとは別系列のディスク)
- ・MTとの併用可能

④ 内容

- ・更新前後のレコード (物理ブロック(ページ))
- ・記録のレコード(開始、終了、コミット、チェックポイント)

2.3 チェックポイントの設定

①意味

- ・同時並行処理している全てのトランザクションの更新状況を逐一管理するのは、困難であり、
- ・全てのトランザクションのキャッシュ領域上の更新結果を、強制的にデータベースに反映する。
- ・その結果、この確認点において、キャッシュ領域とデータベースの内容を一致させることが出来る。
- ・障害時に、どこまで回復するかを示す確認点
- ・回復のための処理時間を短縮できる

③タイミング

- ・決められた一定の時間間隔(15分、30分・・・)

④媒体

- ・CP記録はログファイルに書き込まれる。

3. ログファイルの利用法

3.1 ログファイルの形式

・データベースに対する実行操作の履歴

・形式

LSN	操作名	TRS-No	データレコード
-----	-----	--------	---------

ログレコードの
一連番号

トランザクション番号

・操作名の種類

BOTR	トランザクションの開始
WR-BEFR	更新前の物理ブロック
WR-AFTR	更新後の物理ブロック
COMMIT	コミット実行
CHK-PONT	チェックポイント採取
EOTR	トランザクションの終了

3. 2 ログへの書込み順序

- ログ先書込み方式の場合
 1. トランザクション開始レコードの書出し
 2. 更新前の物理ブロックレコードの書出し
 3. 更新後の物理ブロックレコードの書出し
 4. COMMITレコードの書出し
 5. トランザクション終了レコードの書出し

3. 2 (続き) 書き込み事例

事例:

入金、出金の事例

0001	CHK-PONT	T1
------	----------	----

0002	BOTR	T1
------	------	----

0003	BOTR	T2
------	------	----

口座の残金

0004	WR-BEFR	T1	B10	(¥60,000)
------	---------	----	-----	-----------

0005	WR-AFTR	T1	B10	(¥80,000)
------	---------	----	-----	-----------

0006	COMMIT	T1
------	--------	----

0007	WR-BEFR	T2	B10	(¥80,000)
------	---------	----	-----	-----------

0008	WR-AFTR	T2	B10	(¥60,000)
------	---------	----	-----	-----------

0009	COMMIT	T2
------	--------	----

0010	EOTR	T1
------	------	----

0011	EOTR	T2
------	------	----

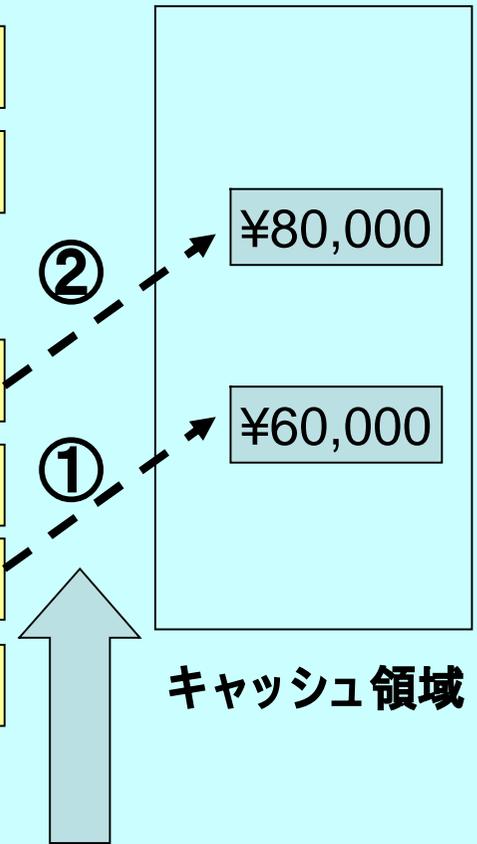
3. 3 使用法1 (ロールバック)

事例:

0001	CHK-PONT	001		
0002	BOTR	T1		
0003	BOTR	T2		
0004	WR-BEFR	T1	B10	(¥60,000)
0005	WR-AFTR	T1	B10	(¥80,000)
0006	COMMIT	T1		
0007	WR-BEFR	T2	B10	(¥80,000)
0008	WR-AFTR	T2	B10	(¥60,000)
0008	WR-BEFR	T3	B10	(¥60,000)
0008	WR-AFTR	T3	B10	(¥40,000)

ロールバック方式

- ・直前のCOMMITまで、
- ・更新前レコード使い戻す作業



3. 4 使用法2 (ロールフォワード)

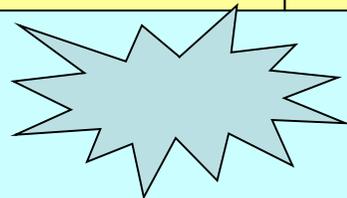
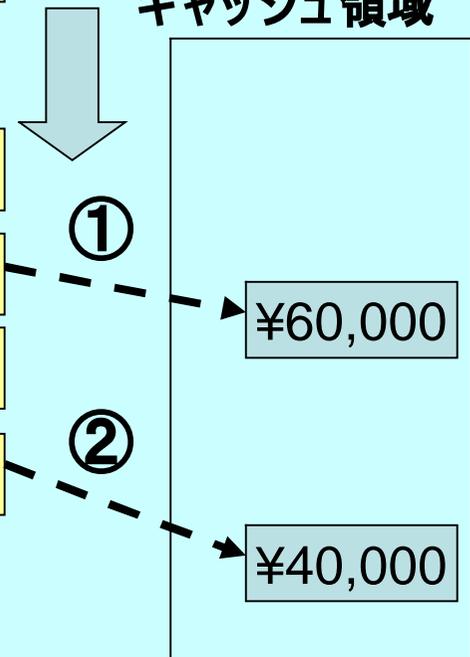
事例:

0001	CHK-PONT	001		
0002	BOTR	T1		
0003	BOTR	T2		
0004	WR-BEFR	T1	B10	(¥60,000)
0005	WR-AFTR	T1	B10	(¥80,000)
0006	COMMIT	T1		
0007	WR-BEFR	T2	B10	(¥80,000)
0008	WR-AFTR	T2	B10	(¥60,000)
0008	WR-BEFR	T3	B10	(¥60,000)
0008	WR-AFTR	T3	B10	(¥40,000)

ロールフォワード方式

- ・直前のCOMMIT以降の
- ・更新後レコードを使い戻す作業

キャッシュ領域



4. 障害回復方法

4. 1 トランザクション障害時の回復

- ①通常の**ロールバック命令**により、開始前の状態に戻す
 - ・SQLのROLLBACK命令が実行される。
 - ・キャッシュ領域の更新結果は破棄される。
 - ・コミット前であり、データベースは、開始前の状態のまま
 - ・(ミス訂正後に、該当トランザクションのリラン)
- ②コミット後でも、トランザクション開始前に戻したい場合は、
 - ・ログファイルを使用して
 - ・更新前レコードを使い、
 - ・トランザクション開始まで戻す(**ロールバック方式**)

4. 2 システム障害時の回復

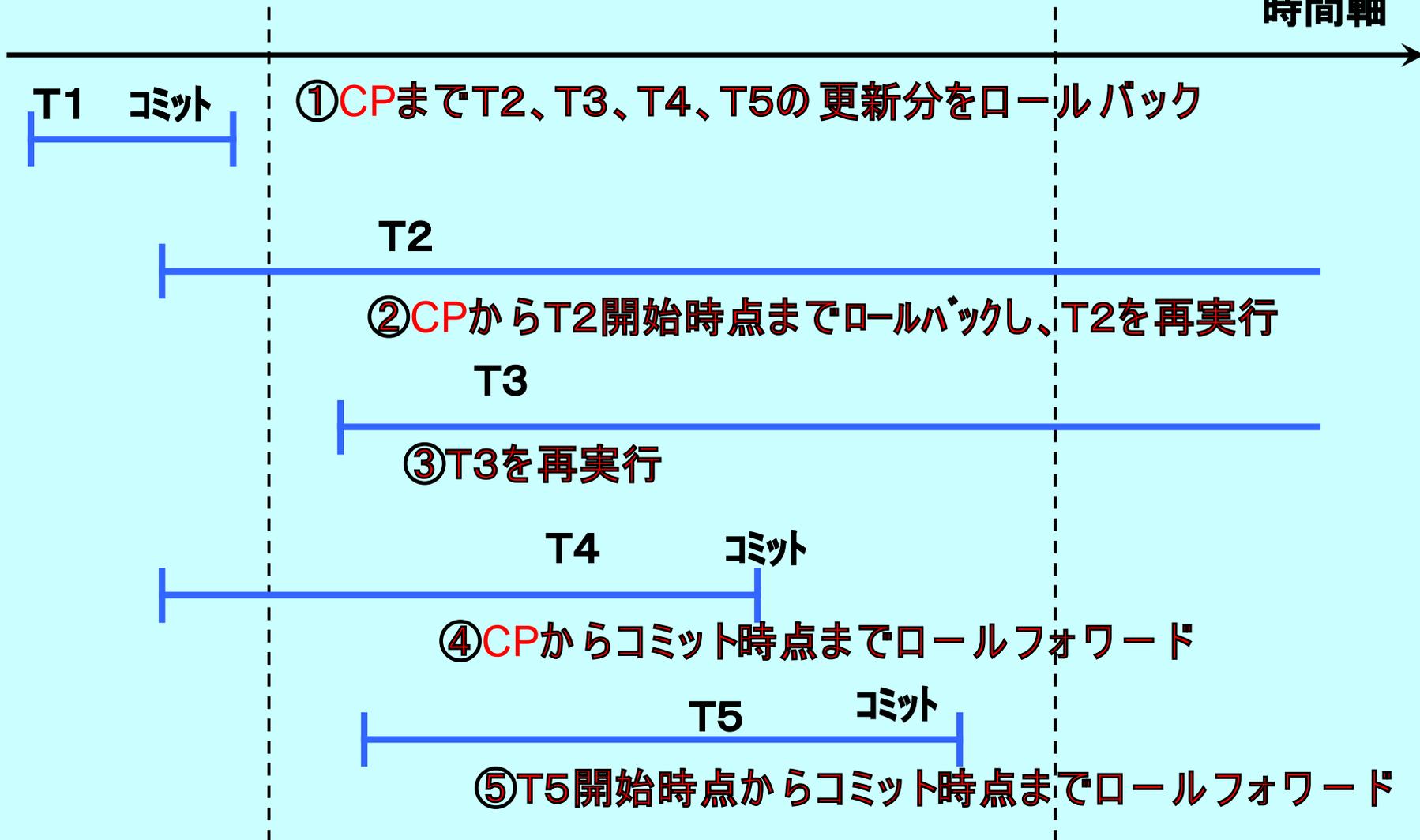
- ① 障害直前のチェックポイントまで、一旦戻す
- ② 同時並行処理されているトランザクションの状況により回復方法が異なる。
 - チェックポイントの使用
 - ロールバック方式の使用
 - ロールフォワード方式の使用
 - 再実行

4.2 (詳細)システム障害時の回復

事例: チェックポイント(CP)

障害発生

時間軸



4.3 媒体障害時の回復

- ① 媒体の修理・交換
- ② 障害DBに関わるトランザクションをすべて停止する。(新たなものも)
- ③ 最新のバックアップ・ファイルからデータベースにリストア(復元)
- ④ ログファイル元に、障害発生時点までDB反映済についてロールフォワード(T1、T4、T5)
- ⑤ 障害発生時点で実行中のトランザクションの処置(T2、T3)

5. DB保守管理

①障害の予防保守

- ・DASDスペースの使用状況を調査
(プライマリー領域、あふれ領域のパンク予防)
- ・残スペースが少なくなれば、拡張する

②DBの再編成

- ・データ更新で発生した「あふれ領域」や再利用できない未使用領域を解消し、物理ファイルの再編成
- ・ディスクスペースの効率利用、アクセス効率向上
- ・データレコードの初期設定、INDEX生成の作業

③DBの再構成

- ・データ項目の変更・追加
- ・テーブルの追加、INDEXの変更・追加

5. (続き)DB保守管理

④ 災害対策(Disaster Recovery Plan)

- ・天災、長期停電、破壊行為などに備える
- ・主要な情報資産のバックアップ、疎開

対象: 情報機器

: データベース、プログラム、マニュアル

方法: バックアップ・ファイル

時期: 定期的

保管: 遠隔地

機器: 遠隔地の機器を借用

: 同一仕様の機器

管理: 手順化、定期的な監査

6. レポート課題

- ①データベース障害の分類と原因をあげてください。
- ②障害回復のために必要な準備項目をあげてください。
- ③媒体障害時のデータベースの障害回復手順の概要を述べよ。

- ①レポートの内容レベルは、A4x1枚程度。
- ②次回の授業開始時に、提出して下さい。
(ただし、それ以前に提出する場合は、
メールで願います。
アドレス: fwhy6454@mb.infoweb.ne.jp)

7. 参考書ほか

- **大木幹雄「データベース設計の基礎」**(日本理工出版会)
- **北川博之「データベースシステム」**(昭晃堂)
- **山田精一「Oracleのデータベース」**(翔泳社)
- **松崎為豁「データベースの基礎の基礎
SQL Server2000」**(ソフトバンク)
- **織田敬三「ビジネス・パソコンユーザーのための
ネットワーク対応、データベース構築ガイド」**
(電波新聞社)
- **山田照吉ほか「テクニカルエンジニアデータベース 2003年版」**
(日経BP社)