

マルチ・スレッド・アンローダー

Oracle用高速CSV出カツール

製品紹介サイト <http://www.mtu-accelerates.jp/>

開発／販売：株式会社プラムシックス

アンローダーの用途

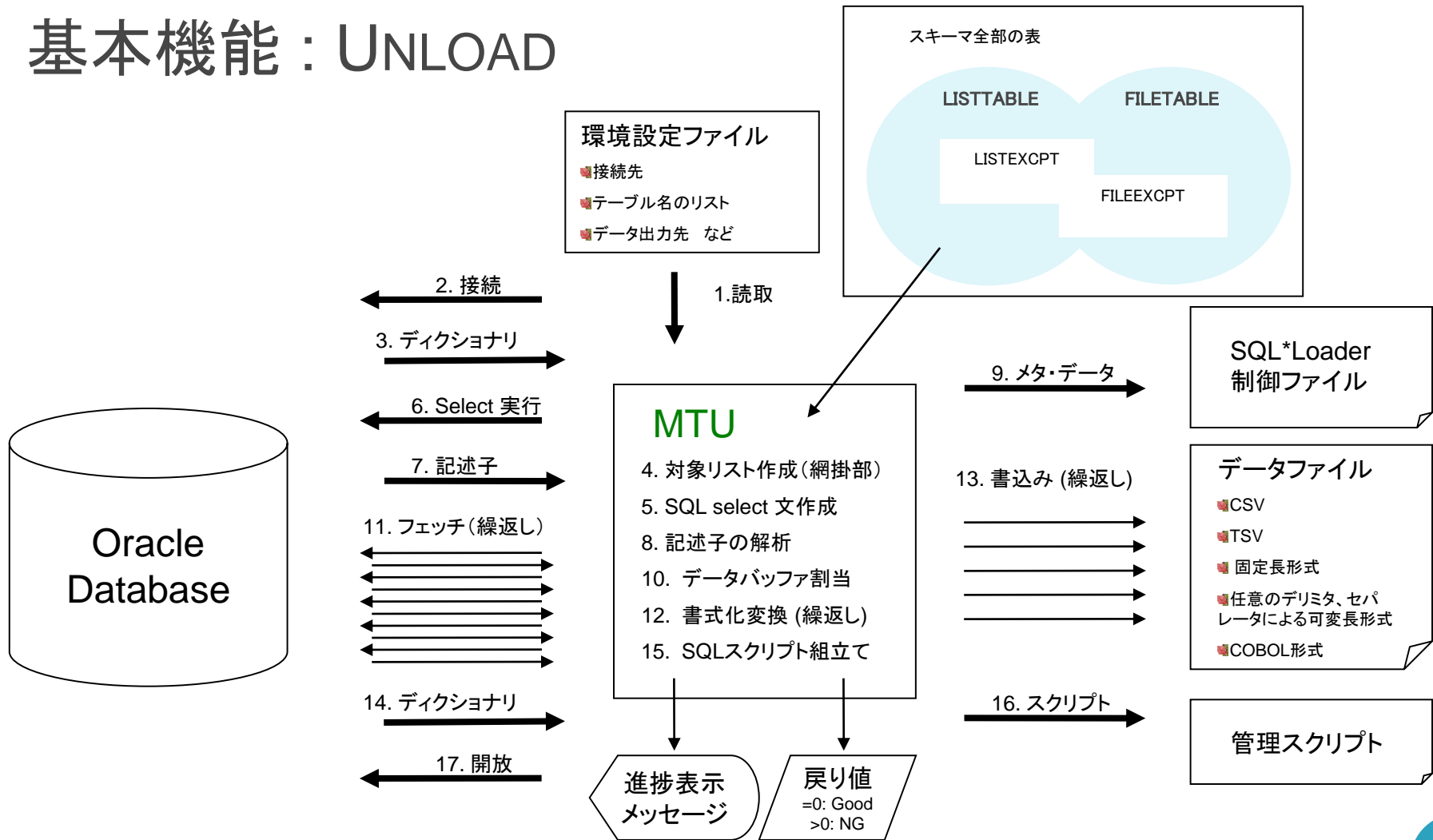
- 論理バックアップ
 - 物理(RMAN)バックアップのみでは回復できない(*1)プログラムのバグやヒューマンエラーによるデータ消失に対する備え
- システム間インターフェース
 - ホストコンピュータによるバッチ処理、帳票出力へのデータ供給
- データ移行
 - 限られた時間内での速やかなデータ移動(精度の高い所要時間見積)
- システムテストデータの準備
 - 条件を変えて繰り返しテストする為データ・セットの容易で迅速な巻き戻し
- データ比較／分析
 - NoSQLを使ってリソース集約的な計算を行う為のデータ・ステージング
- 表や索引のチューニング
 - 逆順キーの追加や行移行により生じた非効率な領域使用の解消

*1: 但しフラッシュバック・データベースやフラッシュバック・トランザクション等Enterprise Editionで利用できる有償オプション、あるいはPoint in-time Recoveryを使うと回復できる場合があります。

汎用化と大量データ対応への課題

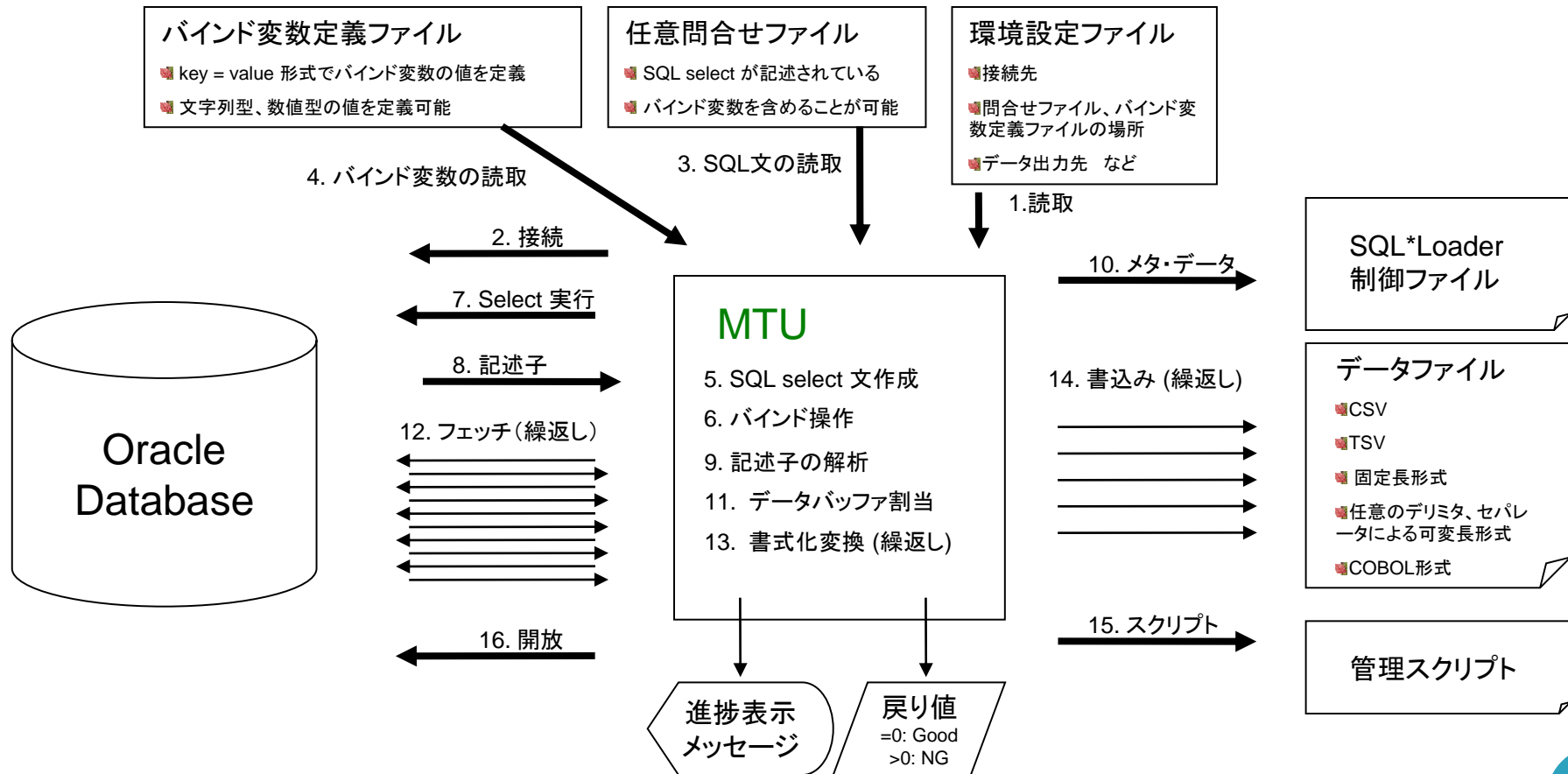
- 読取り一貫性(分離レベル)
 - 大量データのアンロードは時間が掛かります。開始と終了の間にトランザクションがあった場合、デフォルトの設定では問合せ相互の読取り一貫性が保たれません。
- 整合性制約、索引、データベース・トリガー、REDOログの発生
 - 上記はアンロードしたデータを再ロードする際の所要時間長期化や空き領域を逼迫させる原因です。これらを回避するにはOracle Databaseの専門的な知識が要求されます。
- データ用途の多様性
 - さまざまなアプリケーションへの入力にアンロードデータを使いたい場合には書式や入出力方法を自由に選べると便利です。
- データ型の多様性
 - LOBやタイムスタンプ型、間隔型といった比較的新しくサポートされたOracleデータ型にも対応できると更に用途が広がります。
- 言語や地域毎のデータ表現の多様性
 - Unicodeや日本語以外の文字セットの出力、数値や通貨、日付時刻の文字列表現をパラメータで自由に選択できると便利です。

基本機能：UNLOAD



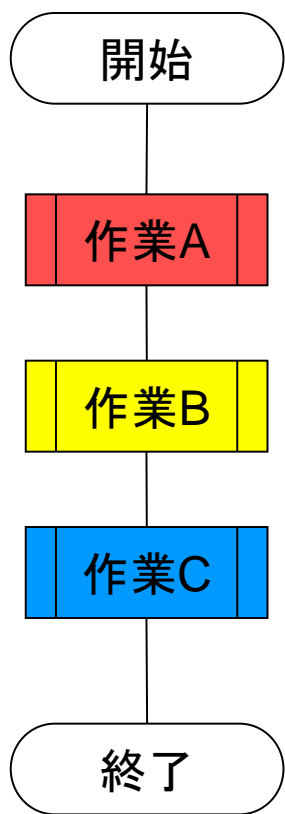
Unload は表データのフルセットを戻す機能です。

基本機能：QUERY

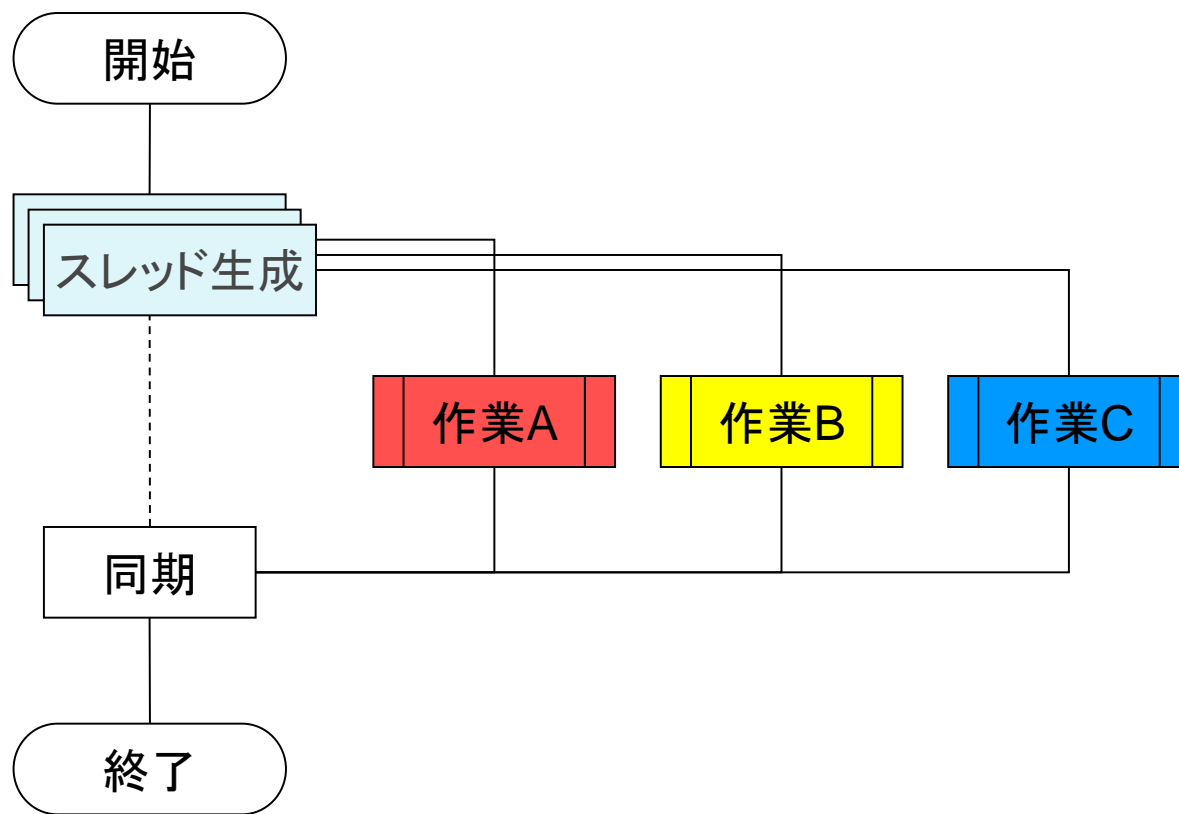


Query は SQL Select 文で表現可能なあらゆる問い合わせの結果を戻す機能です。

複数作業を同時にこなす

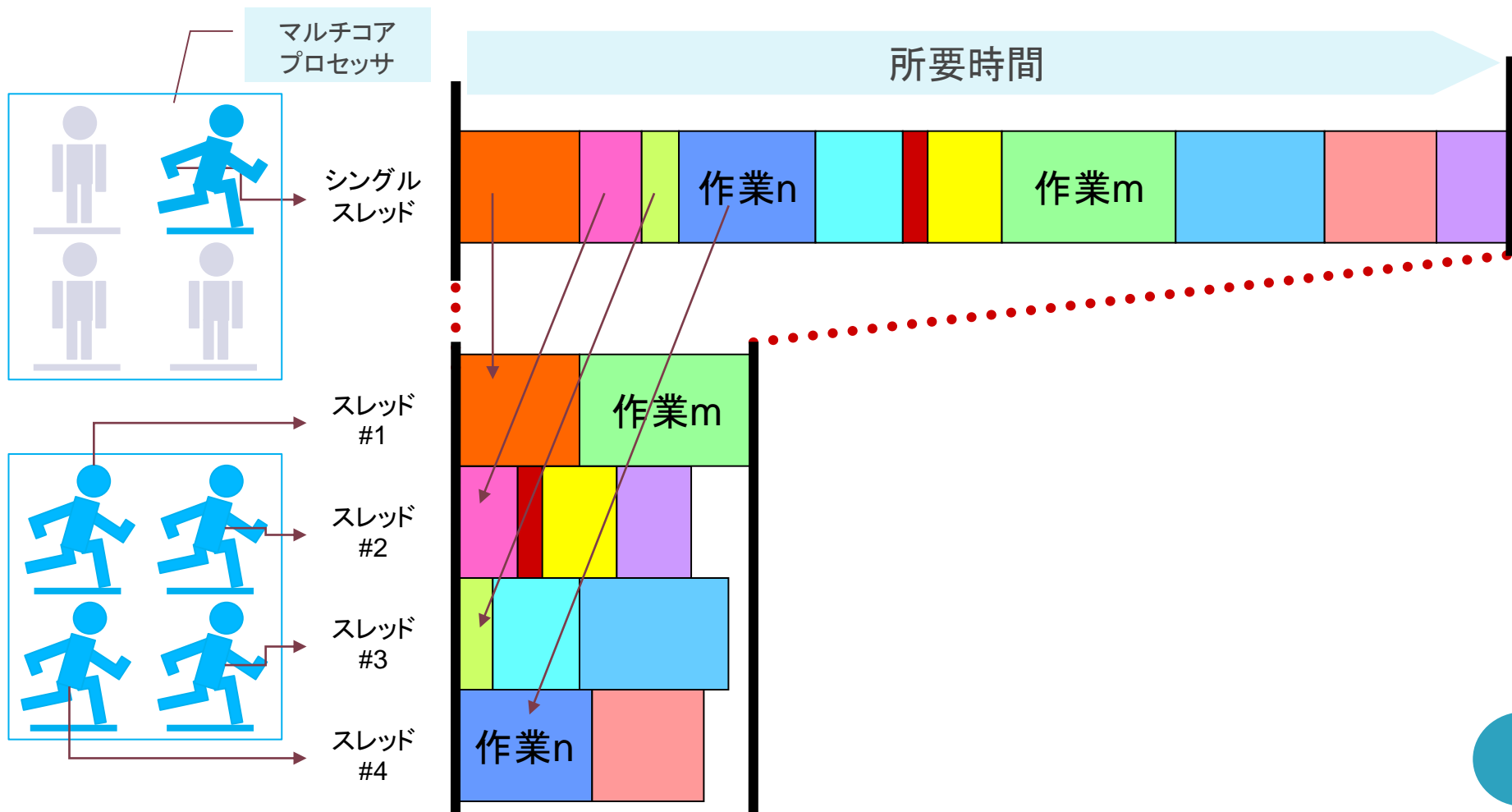


シングル・スレッド・プロセス

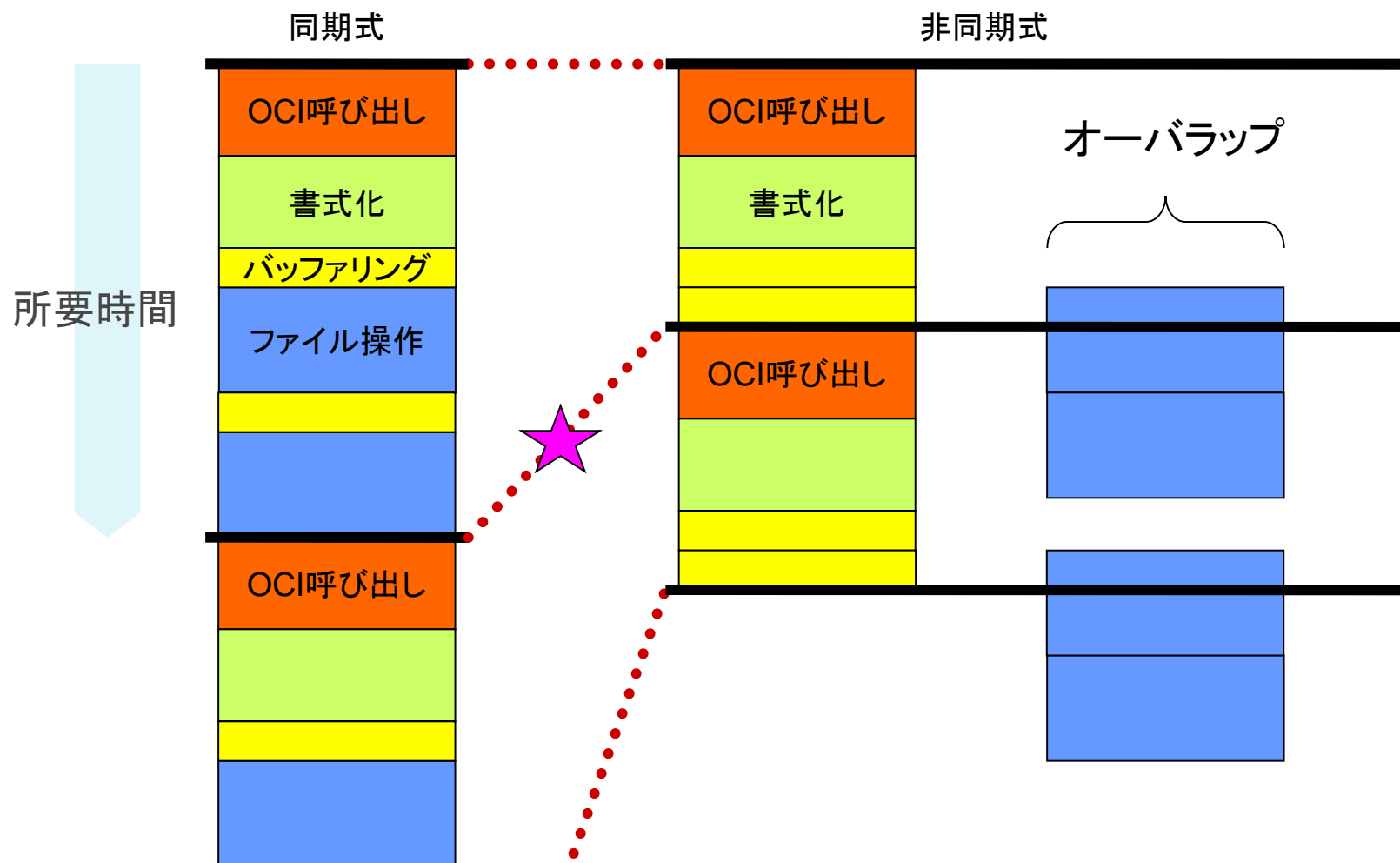


マルチ・スレッド・プロセス

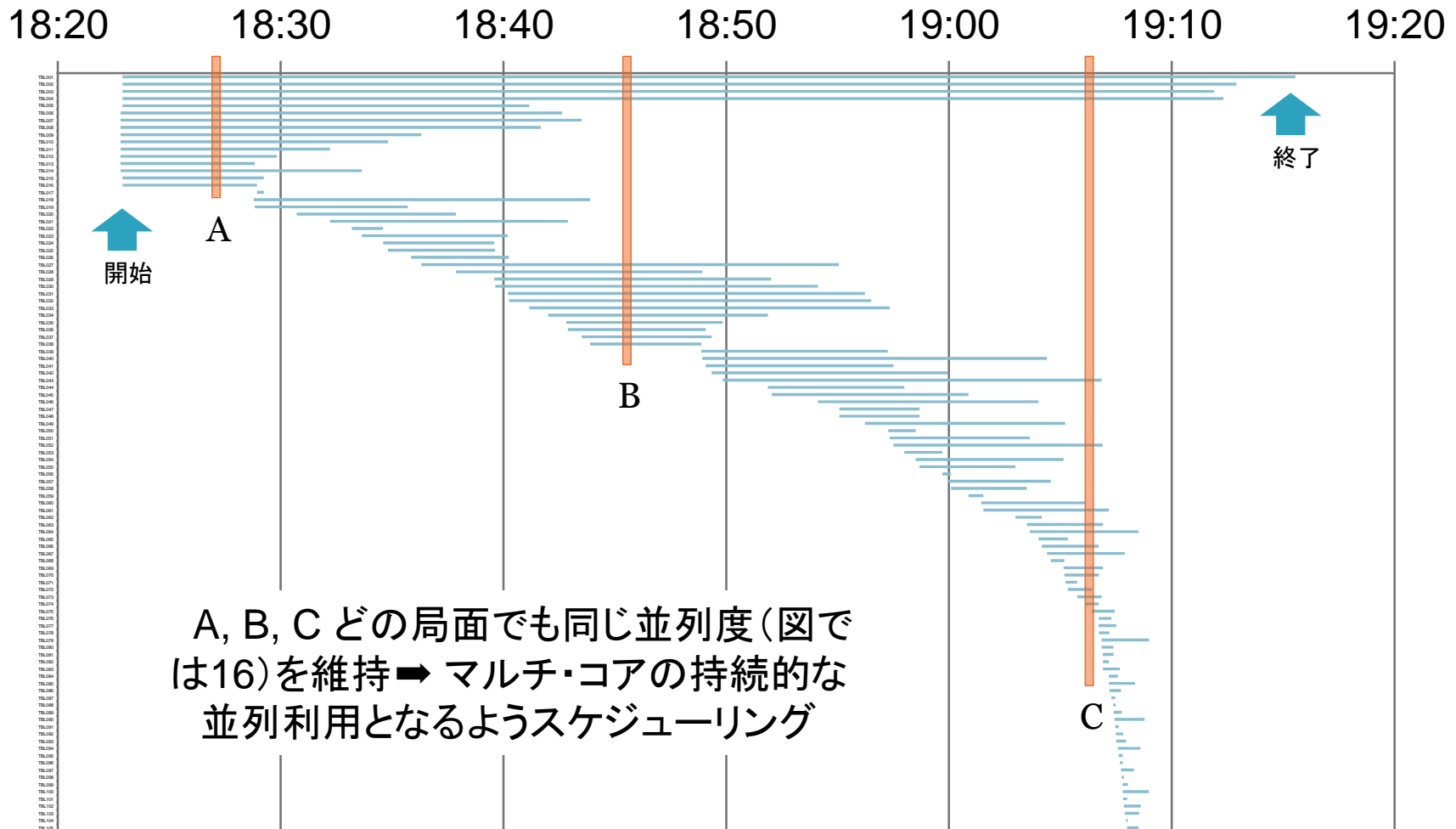
マルチコアを有効に使う



非同期ファイル操作

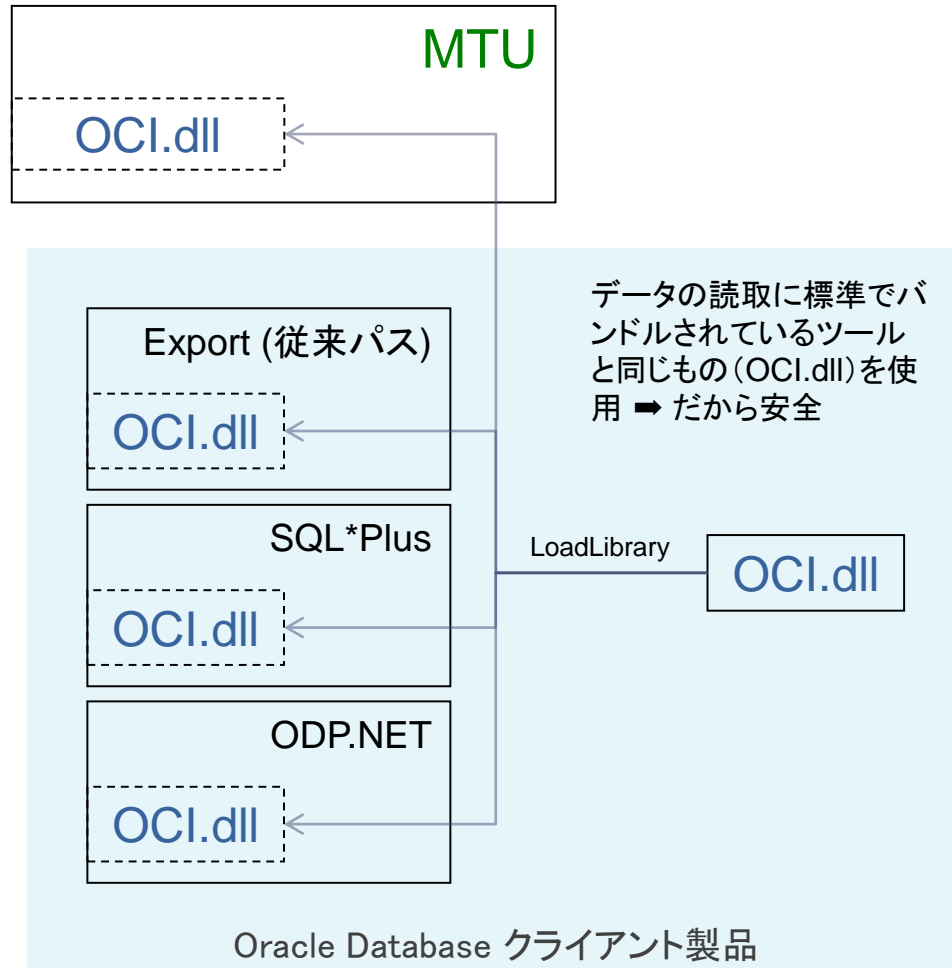
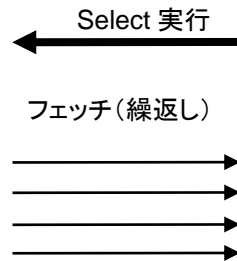
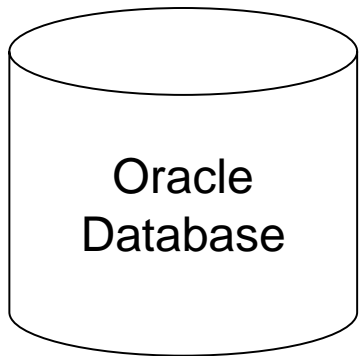


タスク制御

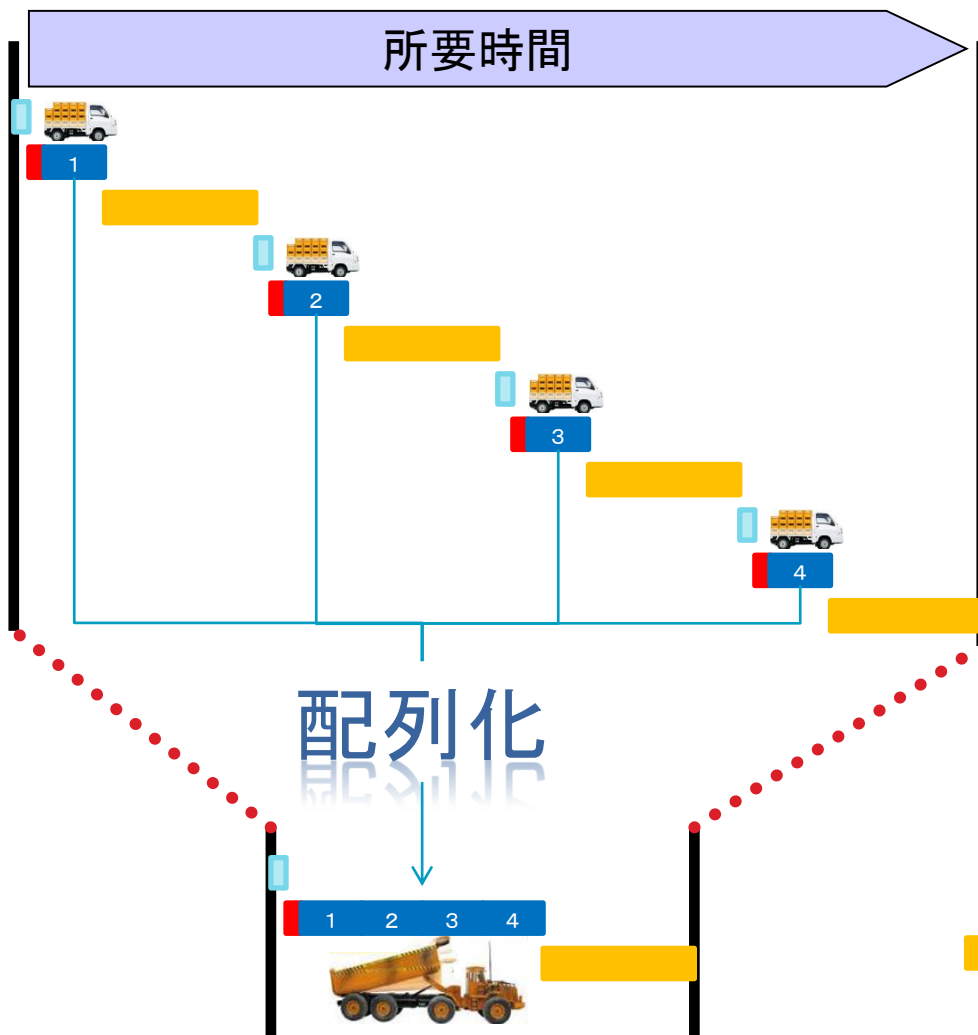


データの安全性

OCI = Oracle Call Interface



何故速いのか？



1行ずつ受け取って処理を行うプログラムは簡単に作れる半面、クライアント側のオーバーヘッドが大きい

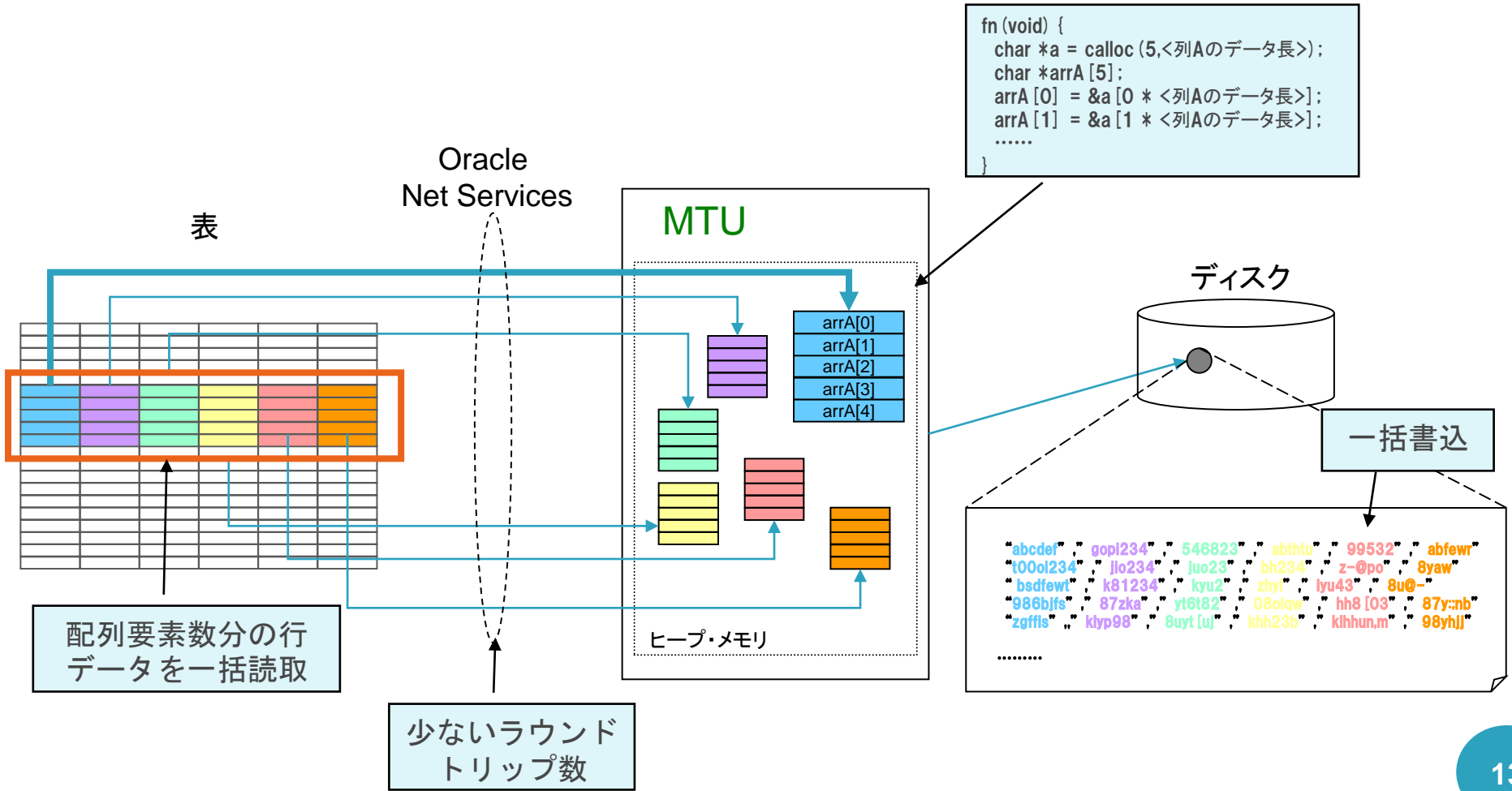
行をまとめて一括処理

- 要求パケットの到達時間
- 応答パケットの到達時間
- 3 データの転送時間
- 1往復毎のクライアント処理時間

その他の特長

- OCI配列インタフェース処理
- 製品固有コードの最適化
- パーティショニング対応
- Unicode出力対応
- 名前付きパイプへの出力
- SAM (Sequential Access Method) 出力
- SQL*Loader用制御ファイルの自動生成
- 索引、整合性制約、データベーストリガの無効化有効化スクリプト作成
- 再ロード時の並列処理

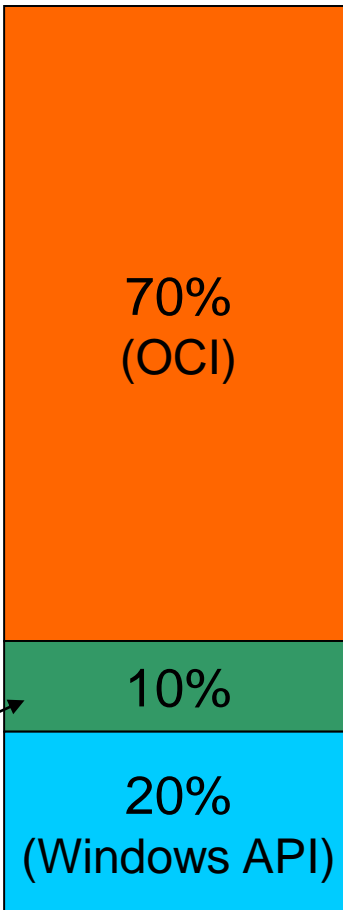
OCI配列インタフェース処理



製品固有コードの最適化

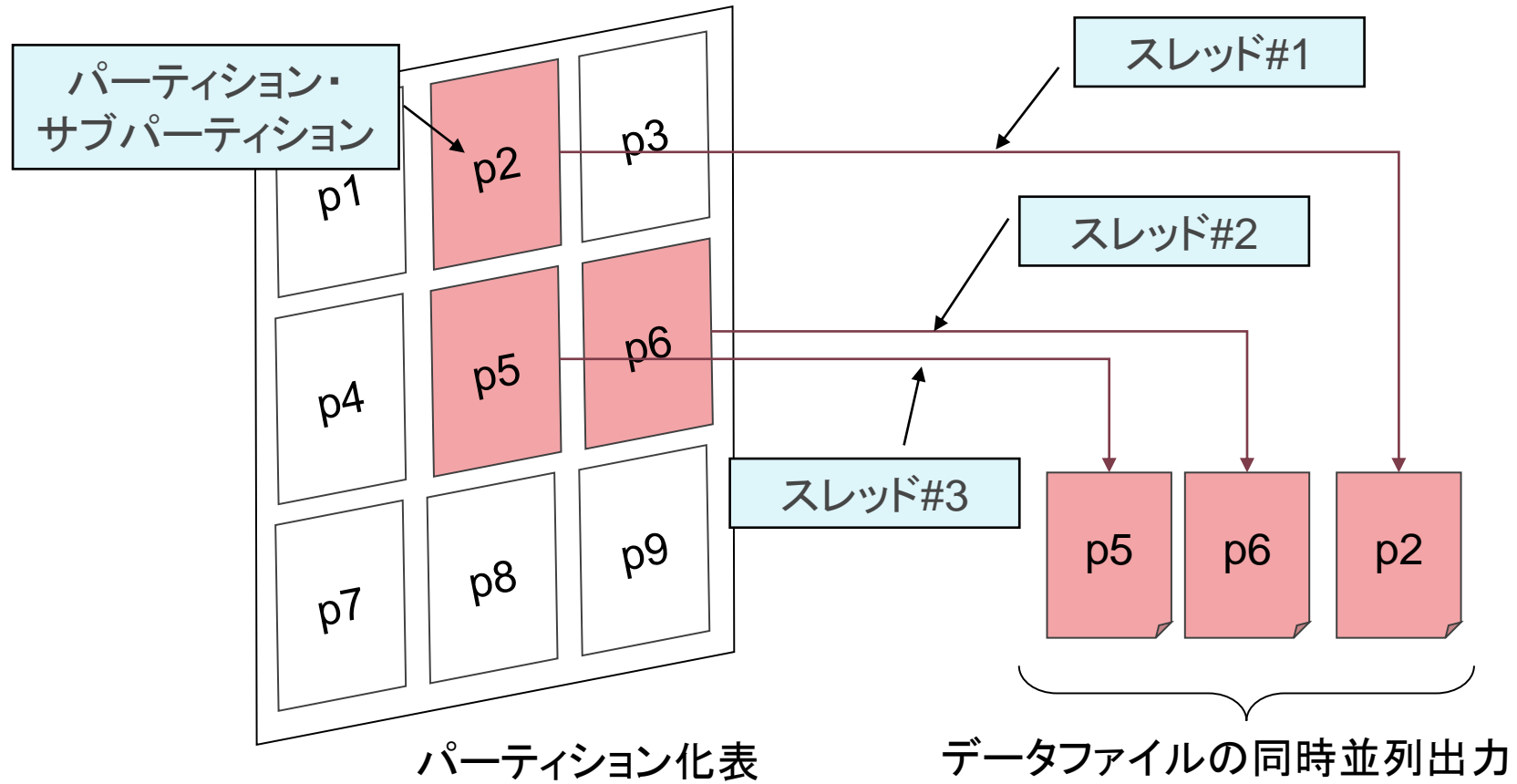
```
void func(void)
{
  sword i0ciRtn;
  i0ciRtn = OCISstmtPrepare (stmtp, ...);
  /*(ホスト入力変数のバインド)*/
  i0ciRtn = OCISstmtExecute (svchp, stmtp, ...);
  /*(暗黙的パラメータ記述の解析)*/
  /*(ホスト出力変数の定義)*/
  /*(SQL*Loser制御ファイル出力)*/
  while (i0ciRtn != OCI_NO_DATA)
  {
    i0ciRtn = OCISstmtFetch2 (svchp, stmtp, ...);
    /*ホスト出力変数をCSV or 固定長データ形式へ変換*/
    WriteFileEx (hFile, lpBuffer, ...);
  }
}
```

MTU待機要因

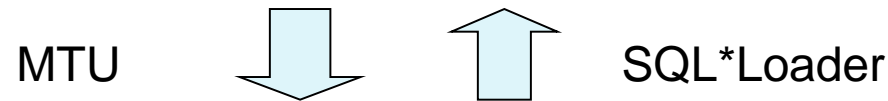
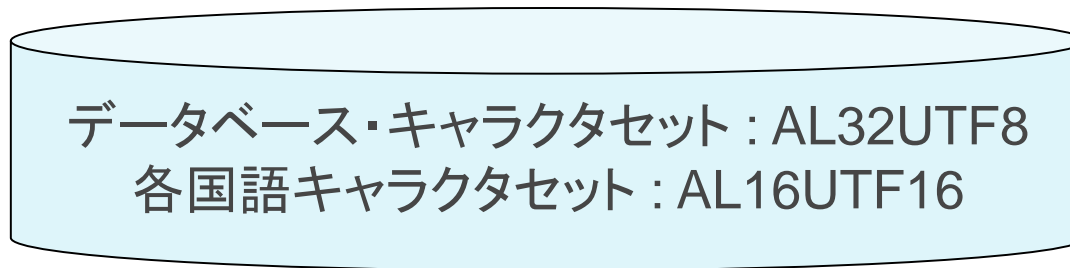


製品固有コード

パーティショニング対応

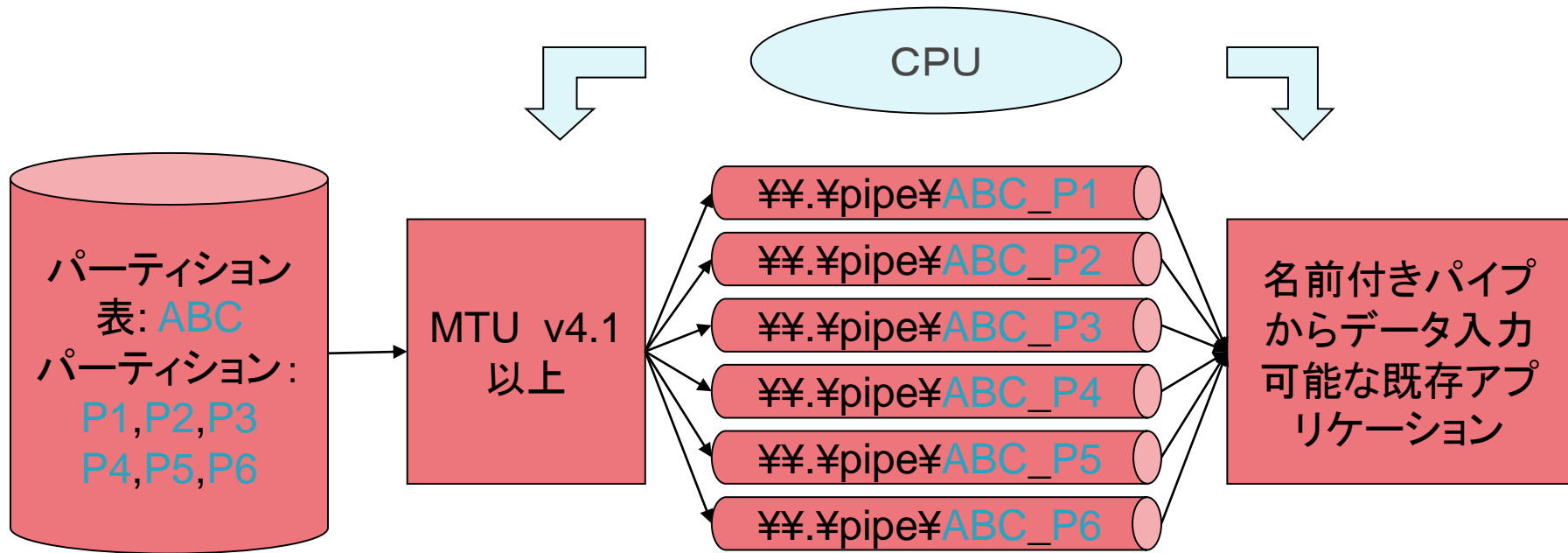


UNICODE出力対応



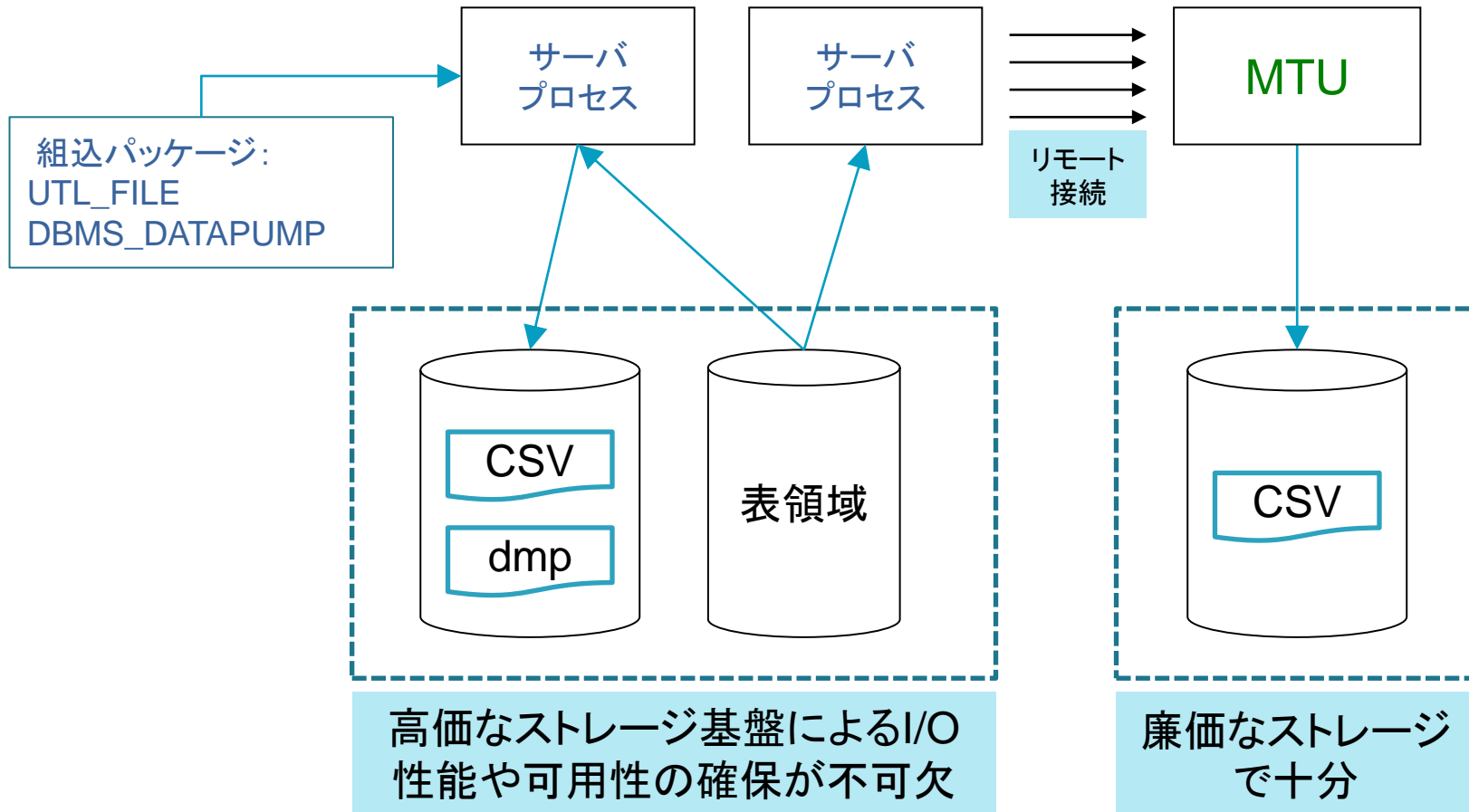
Oracleデータ型	MTU出力文字エンコーディング
CHAR、VARCHAR2、NCHAR、NVARCHAR2、LONG、CLOB	UTF-8 (BOM付／BOM無選択可能)
NCLOB	UTF-16 BE

名前付きパイプへの出力



H/Wプラットフォームのスケラビリティに適した数の名前付きパイプを配置してパイプライン化

PL/SQL フリー



SAM出力

- IBM互換メインフレーム用のCOBOL言語で扱うSAMと互換性のある形式でファイル出力可能です。
- NUMBER型 : COMP-3 (パック10進数)で出力できます。
 - DISPLAY SIGN LEADING SEPARATE 形式のBCD (Binary Coded Decimal)を選ぶ事も出来ます。
- NCHAR / NVARCHAR2 型 : マルチバイト文字の出力に UTF-16LE、UTF-16BEいずれかを選択できます。

競合製品との機能比較

	MTU v4.2	FUO r11.2
Oracleへのリモート接続	○	×
64bit 版Windows	○	×
再ロード用スクリプト作成	○	×
再ロード時のパラレル処理	○	×
表ロックを掛けない一貫性読取り	○	×
実行時の強力なOSのシステム権限	不要	必要
接続時フェイルオーバ	○	×
透過的アプリケーションフェイルオーバ	○	×
ASM上のデータ抽出	○	×
並列度制限	無	有
データの出力先	ローカル、共有フォルダ	ローカルのみ
領域あふれ時のフェイルオーバ	○	×
並列化可能なパイプ処理	○	×
ヒープ構成表の並列アンロード	○	○

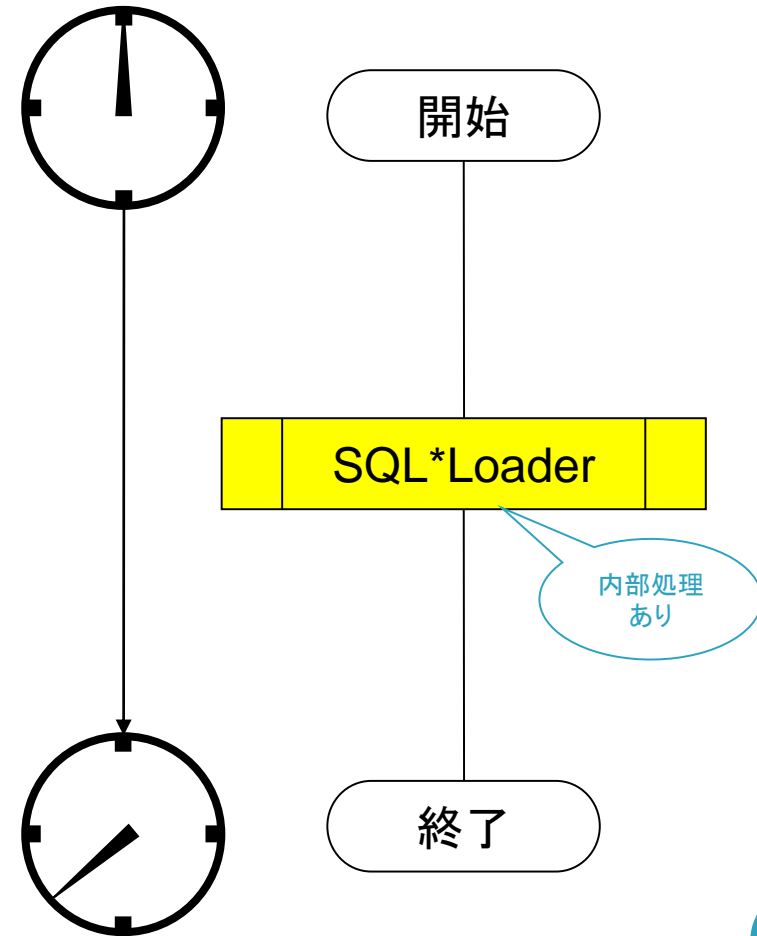
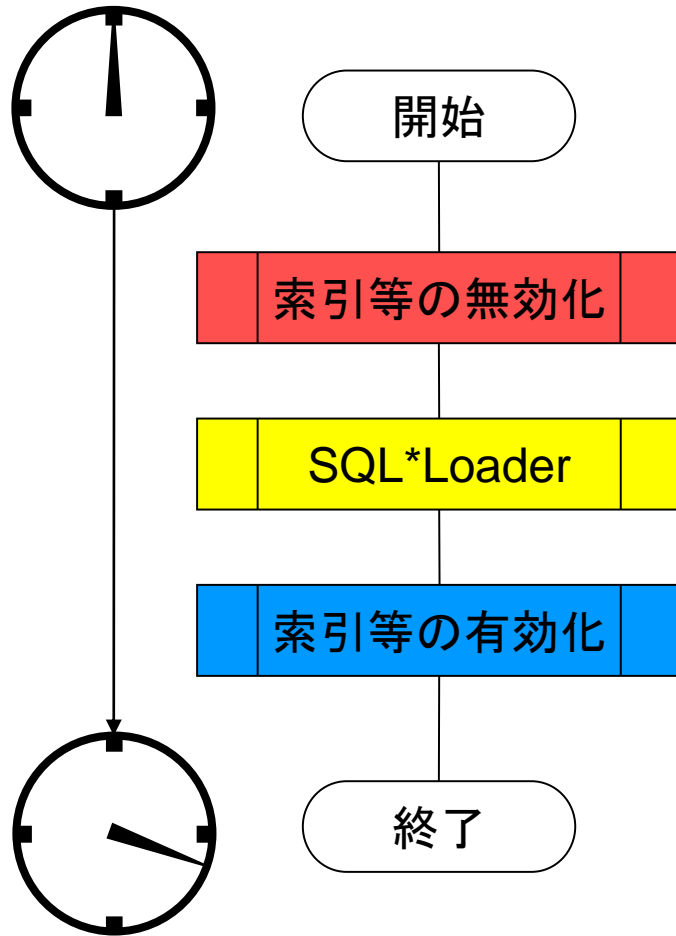
SQL*LOADER用制御ファイルの自動生成

```

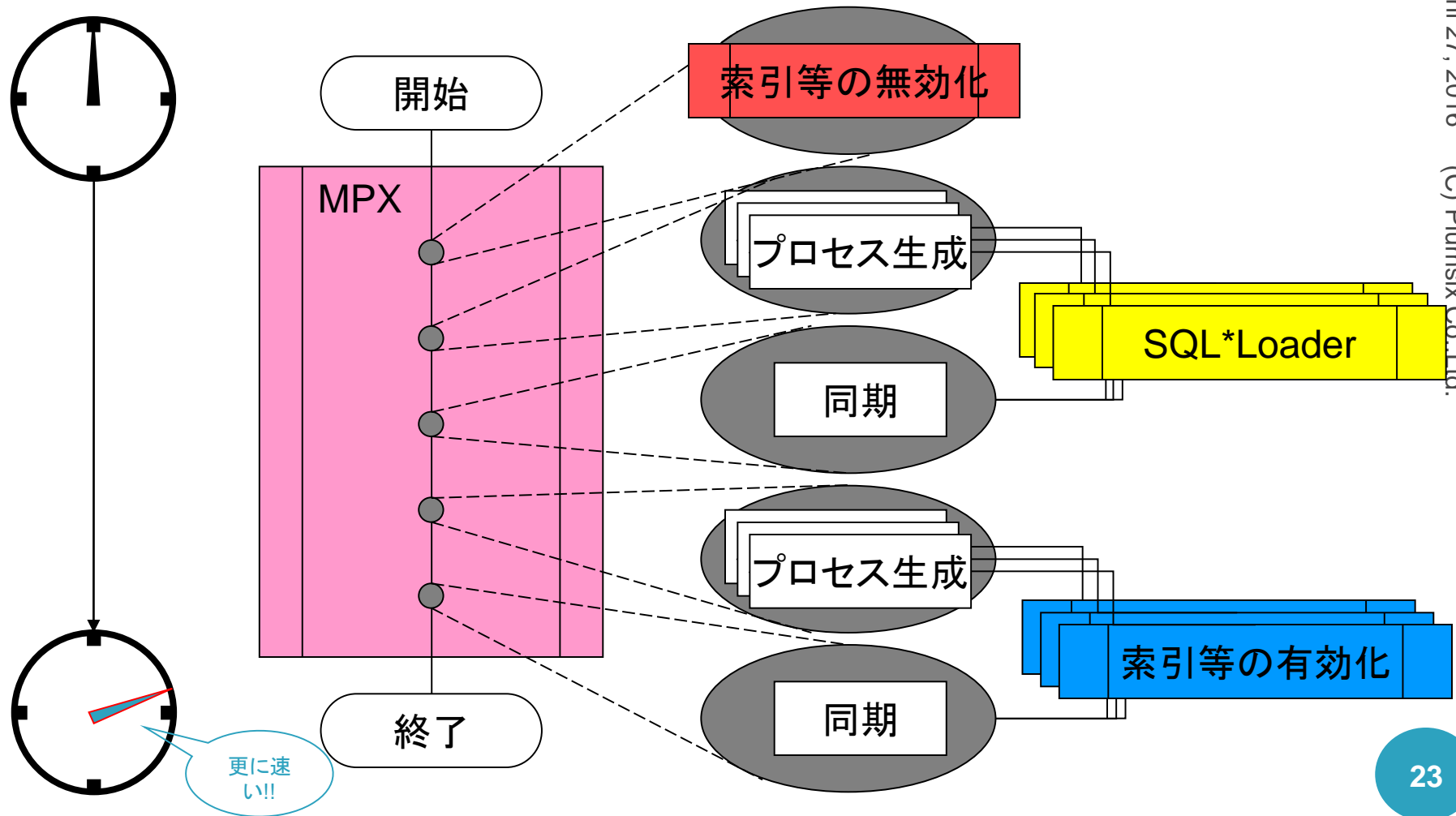
UNRECOVERABLE LOAD INFILE "EMP" "VAR 10"
INTO TABLE "EMP"
TRUNCATE REENABLE FIELDS TERMINATED BY ','
( "EMPNO"      DECIMAL EXTERNAL (5)
, "ENAME"     CHAR (10) ENCLOSED BY '"'
, "JOB"       CHAR (9) ENCLOSED BY '"'
, "MGR"       DECIMAL EXTERNAL (5)
, "HIREDATE"  DATE (14) "YYYYMMDDHH24MISS" ENCLOSED BY '"'
, "SAL"       DECIMAL EXTERNAL (9)
, "COMM"      DECIMAL EXTERNAL (9)
, "DEPTNO"    DECIMAL EXTERNAL (3)
)

```

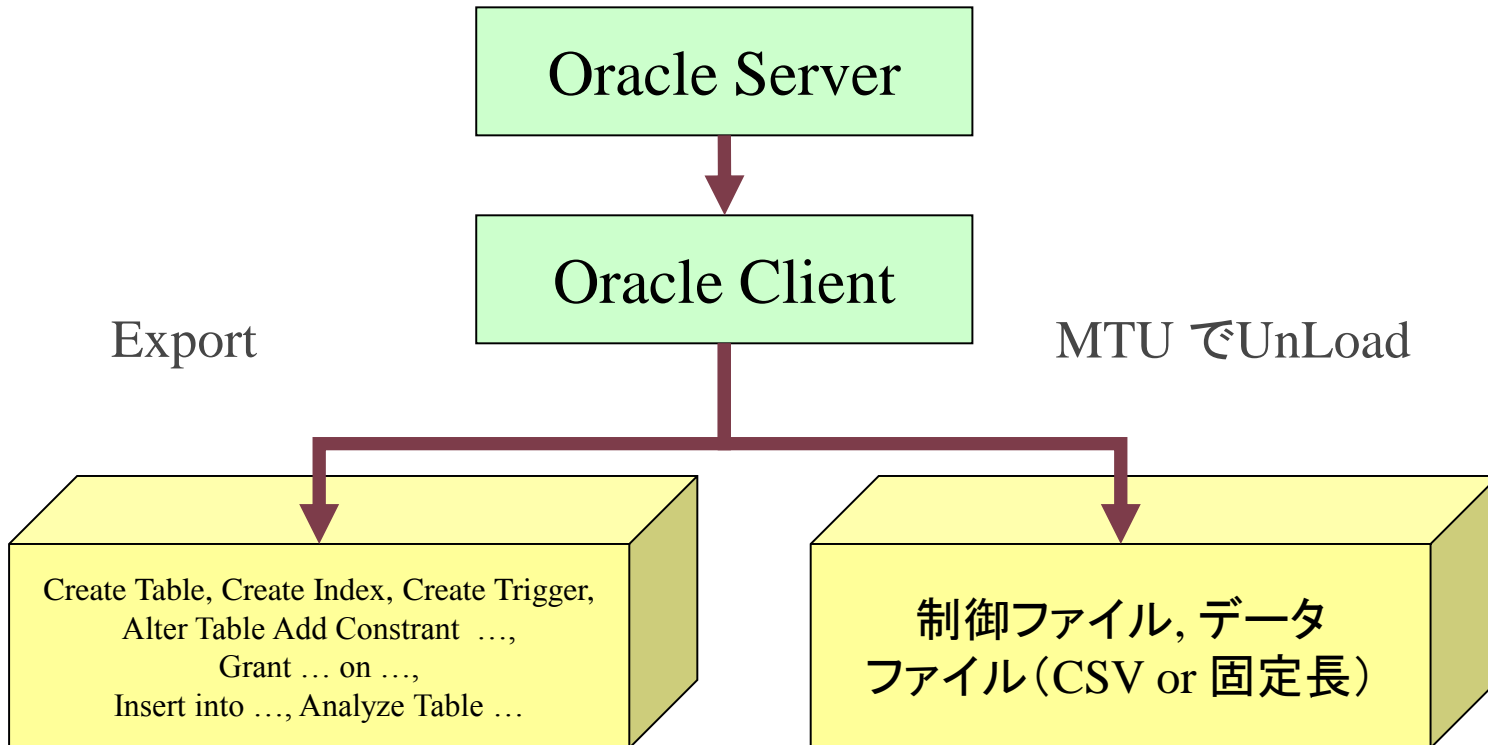
索引、整合性制約、データベーストリガの無効化有効化スクリプト作成



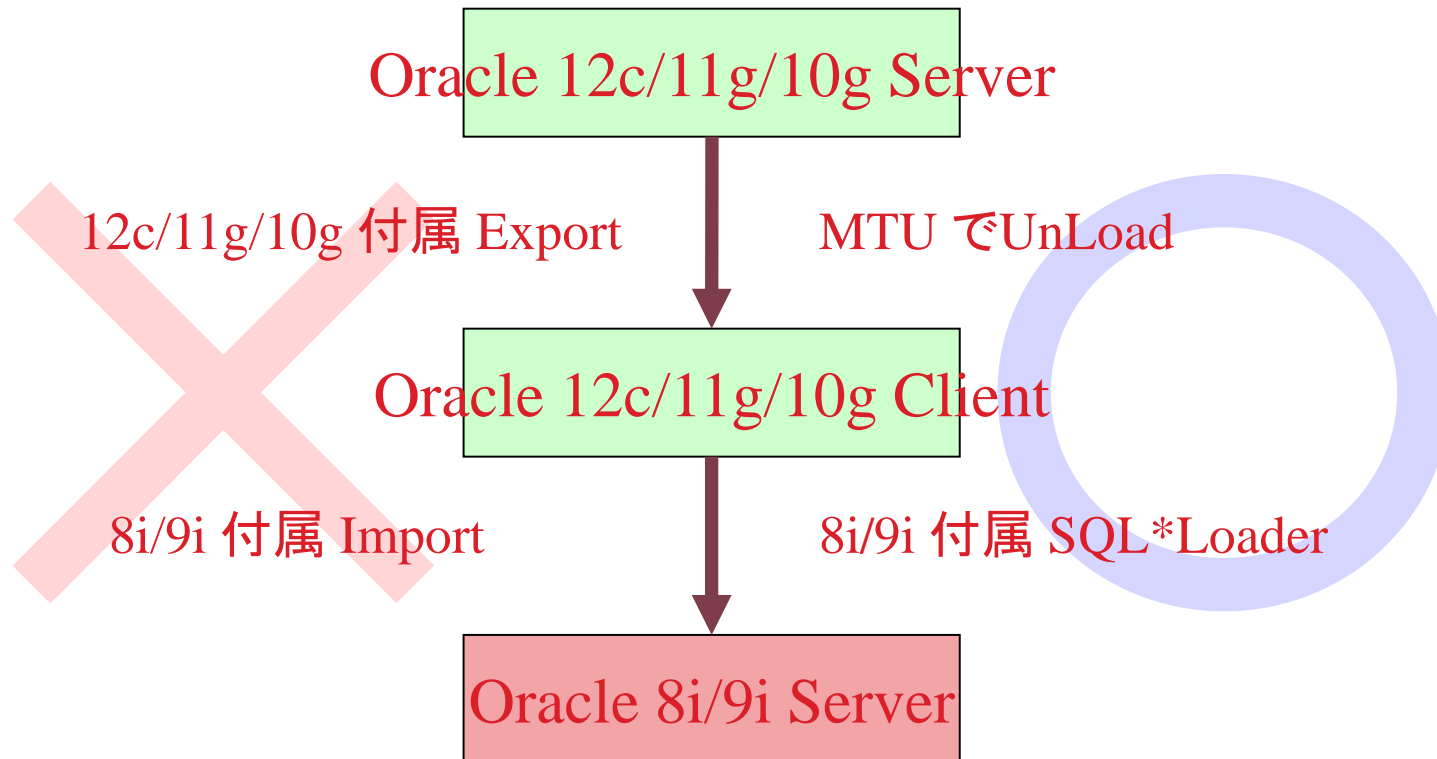
再ロード時の並列処理



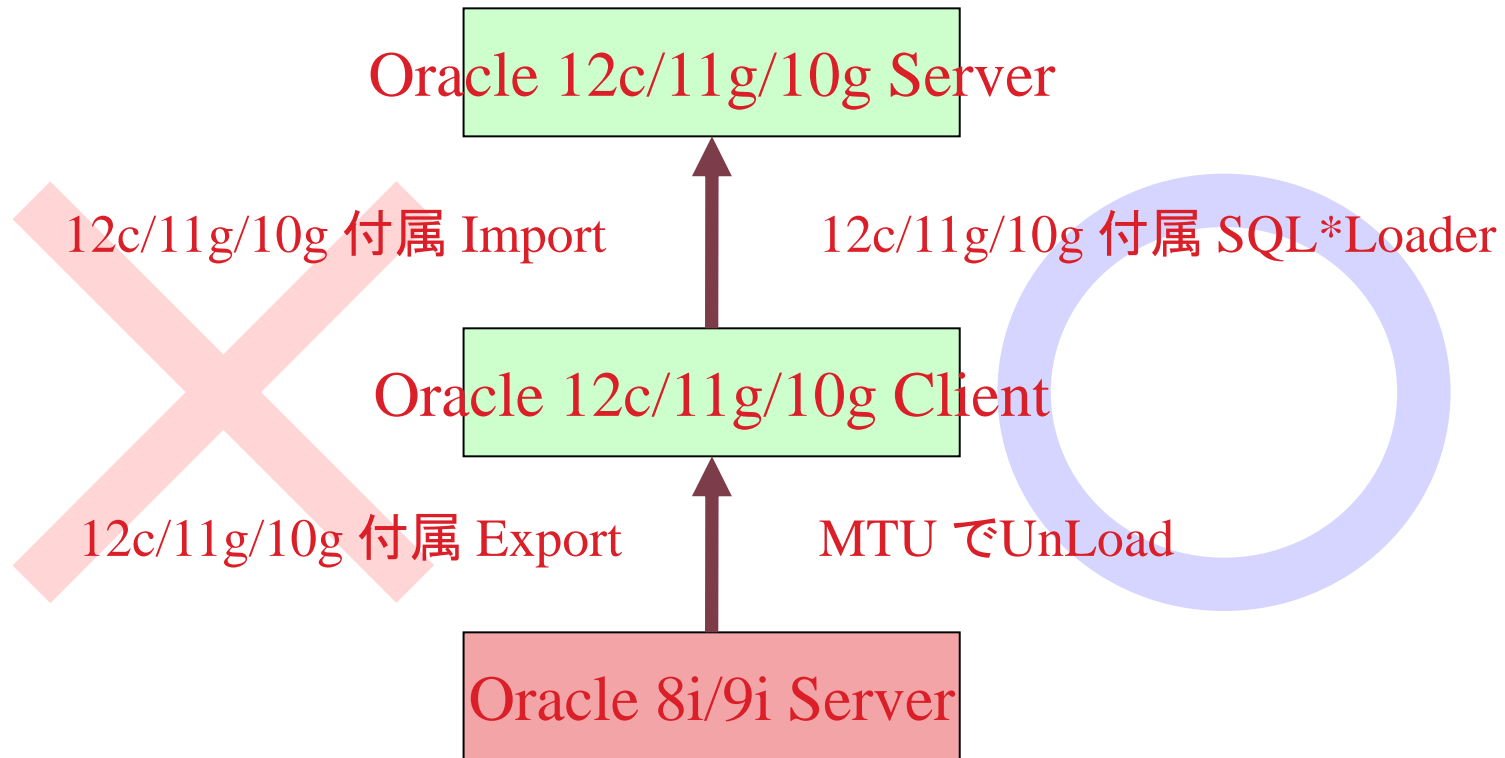
用途例：単純なデータだけの移行



用途例：下方データ移行



用途例：上方データ移行



ありがとうございました

提供：株式会社プラムシックス
<http://www.mtu-accelerates.jp/>

マルチ・スレッド・アンローダー、Multi-threaded Unloaderは株式会社プラムシックスの登録商標です。

オラクル、および ORACLE は米国Oracle Corporationの登録商標です。

Microsoft Windows Windows Serverは、米国Microsoft Corporationの米国およびその他の国における登録商標または商標です。

他のブランドおよび製品名は、各社の商標および登録商標です。