

ハードウェアマージソータの
駆動系設計

1D-9

鈴木 孝[†] 伏見 信也[‡] 喜連川 優[‡] 田中 英彦[‡] 元岡 達[‡]

([†]三菱電機 [‡]東京大学工学部 [‡]東京大学生産技術研究所)

1. はじめに

ハードウェアソータは、データベースマシンGRACEにおけるプロセッシングモジュールの基本要素となるものである。本ソータは多種の拡張機能を有しており、その制御は個々のレコードに付加された各種のフラグによって実行される。本ソータには、これらフラグの生成等に加えて、ソートユニットのデータ入出力制御や、同期制御を行うハードウェア環境が外部に必要であり、それらの要求を満たすソータ駆動系としてのソートドライバを設計した。ソートドライバによって外部インターフェイスは簡略化され、汎用性の高いソートシステムが実現できる。

2. プロセッシングモジュールの構成

Fig.1に、GRACEにおけるプロセッシングモジュール(PM)の構成図を示す。

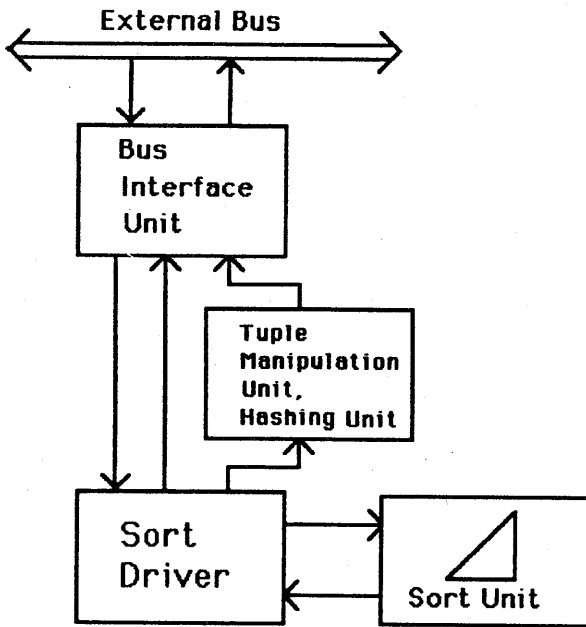


Fig.1 Processing Module

PMは、関係データベース処理におけるSort、Join、Aggregation等の演算を行うモジュールである。PMは大きく次のユニットから構成される。

(1) Bus Interface Unit

外部のバスに接続され、PMの制御情報の受信、リレクションのタプルデータ入力、ソート結果の出力、関係代数演算後のタプルデータの出力を行う。

(2) Sort Driver

ソートユニットを制御するハードウェアドライバで、入力レコードに付加されたフラグの生成、データフォーマット変換、ソートユニットに対するデータ入出力の同期制御等を行う。

(3) Sort Unit

パイプラインマージソートを行うハードウェアソータである。本ソータはレコード長、レコード数等の変動に対し、極めて柔軟に対処することができる。[1]

(4) Tuple Manipulation Unit、Hashing Unit

ソートされたタプルに対し、JoinやAggregation演算を行い、次JoinのためのHashingを行うユニットであり、現在68000システムを用いた設計を進めている。

3. ハードウェアソータの制御方式

Fig.2にソートユニットの入出力信号線を示す。

ソートユニットは、クロック信号に同期して動作するが、その外部環境においては、データ転送は非同期に行われることが一般的である。そのため外部からのデータは、DATA VALIDによって同期化され入力される。ソートユニットは4MB/S程度の処理速度を持っており、各ソートプロセッサはデータ入力に待たされることなく処理を実行できる。それに対しデータ出力に関しては、送出先がBus Busy等でデータを送れない場合があるため、HALT信号でソートユニットの機能を一時的に停止することで同期を取る方式とした。

またソート処理における昇順、降順を決めるA/D信号線によって、1レコード内の複数のKeyごとにバイト単位で昇順、降順ソートが可能となっている。

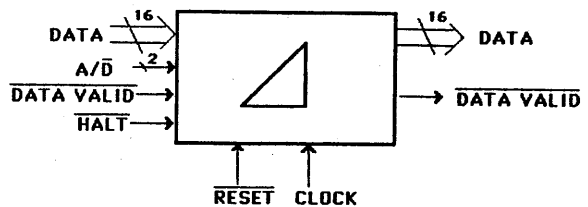


Fig.2 I/O Signals of Sort Unit

Fig.3に入力データストリームの一般的フォーマットを示す。

Fig.3で示されるように、入力データストリームには1レコードごとの各種フラグが付加され、それによって拡張機能が制御される。

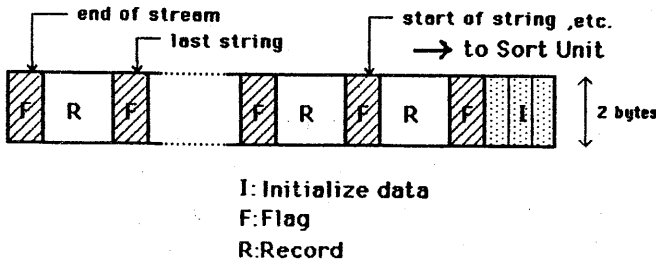
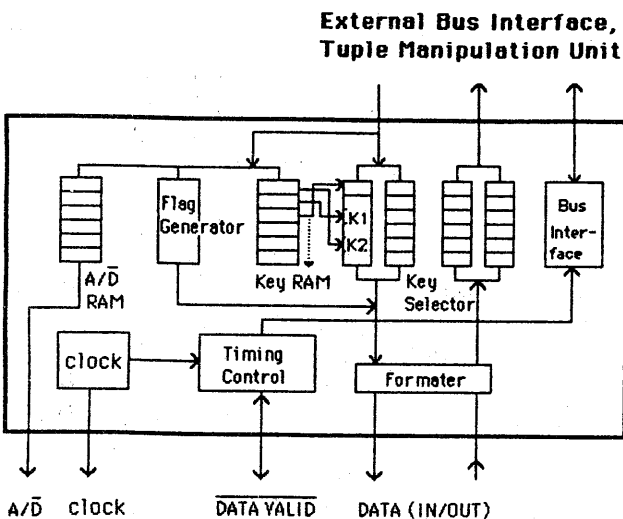


Fig.3 Input Data Stream Format

4. ソートドライバ

3. で述べたソートユニットの制御とデータ転送を実行するユニットがソートドライバである。Fig. 4 にその内部構成を示し、以下にその概要について述べる。



(Sort Unit Interface)

Fig. 4 Sort Driver

4-1 初期化

外部バスインターフェイスよりデータ転送に先立って送出されてくる初期化命令によって、初期化データを生成し、ソートユニットに送出する機能である。

4-2 入力データ処理部

以下の機能があり、ハードウェアで制御される。

Key フィールド抽出機能は、1レコード内の任意個のKey フィールドを抽出し、レコードの先頭に置き換えて出力する機能である。これをデータストリームを乱すことなく実時間で実行する回路が、Fig. 4 のKey パターンRAM と1レコード分のダブルバッファ部である。初期化命令によって1レコード内のKey の位置（バッファRAM のアドレス）が書き込まれたRAM テーブルによって参照される順にソートユニットへとレコードデータストリームが出力されていく。

Formaterは、Binary、Character、10進演算におけるZone.Pack 表現の各データタイプをソート可能なフォーマットに変換する機能である。

Flag付加機能は、初期化命令によって、入力データストリームのフラグパターンを計算し、その結果によりデータストリームの1レコードごとにフラグデータを挿入していく機能である。

A/D 信号付加機能は、初期化命令によって前もってA/D RAM テーブルに書き込まれた昇順、降順パターンに従い、レコードデータ転送に同期して送出する機能である。

これらの機能は、ソフトウェアで実現する方法も考えられるが、ソフトウェアの負荷をできるだけ軽くすること、できる限り高いレベルの外部インターフェイスを与えるため、今回の設計ではハードウェア制御方式とした。そのため余分な制御オーバーヘッドが抑えられ、高スループットが期待できる。

4-3 出力データ処理部

ソートユニットから出力されてくるソートされたレコードに対し、以下の処理を行う機能である。

Formaterは、入力処理によって変換されたデータフォーマットを元の形式に再変換する機能である。

Key フィールド変換機能としては、入力処理によってレコード先頭に配置されたKey フィールドを、そのままの形で送出するか、あるいは元の形式に再変換した後に送出するか選択できる設計となっている。

また、データ入力時に各レコードに付加されたフラグは、フラグ除去部で除かれる。

4-4 外部インターフェイス部

ソートドライバは、ソートユニットのポートの他に、External Bus Interface、Tuple Manipulation Unit へのデータ転送ポートを持つ。

ソートされたデータストリームは、一般にTuple Manipulation Unit に送られ、Join等の処理が行われる。一方データのソートのみを実行する場合は、ソートデータはTuple Manipulation Unit を介さず、直接ソートドライバからExternal Bus Interfaceへ返すことができる。

5. おわりに

以上のようなソートドライバの設計により、ソータを駆動、制御するためのソフトウェアの負荷が軽減され、高レベルのインターフェイスを持ち、かつ高スループットのソート処理能力を持つソートシステムが実現可能となる。

現在ソートドライバの論理設計、プロセッシングモジュールのアーキテクチャの詳細化を進めている。

<参考文献>

[1] 楊、鈴木、林、喜連川、田中、元岡：
「Length Tuning機構を有するハードウェアマージソータの設計」
情報処理学会第30回全国大会 1D-8, '85