

汎用DBMSの性能評価システムについて

米田 茂、山口康隆、吉田郁三 (日立 シ研)
仁平博三 (日立ソフトウェア工場)

1. 緒言

データベースを中心としたアプリケーション・システムの普及に伴ない、データベース・マネジメント・システム (DBMS) に対する評価において、機能的な要求ばかりでなく、適応性の技術、特に性能に関する評価が重要視されている。特に、計算機システム建設段階で、多様なデータベース利用への対応を実現するため、DBMS性能の評価、検討を実施することが、データベース・システム構築上の必須手続きであると思われられて来ている。DBMSの性能には、システムを構成する多様な要素が密接に影響し合っているが、それらを整理すると、図1のようになる。

日立製作所においては、Adaptable Data Manager (ADM)、Practical Data Manager (PDM) の2種のDBMSを開発している。ここでは、PDMを対象として、上記データベース・システム建設に際しての基本的技術課題を解決するため開発した、性能評価システム PDM Performance Estimation and Administration Library (PEAL) について、解析モデルを中心に報告するものである。

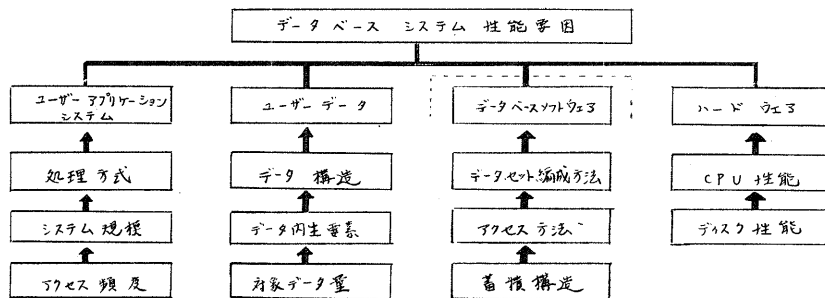


図1 データベース・システムの性能要因

2. 解析モデルの建設

2.1 DBMS

幾々の対象としたDBMSは、マスタ・データ・セット (MDS)、とバリエーブル・データ・セット (VDS) と云う2つのデータ・セット編成方法を持っている。MDSは、アプリケーション・プログラム・キーを用いダイレクトにアクセスできるデータ・セットで、複数のVDSと接続可能である。一方、VDSは、MDSを経由してアクセスできるデータ・セットで、複数のMDSと接続可能である。

この2つの型のデータ・セットMDSとVDSとの接続を図ることによってPDMは、データ・セット間ネットワーク構造のデータベースを形成する。

以下では、今回の解析モデル構成の主対象である、MDSについて、解析モデル構成の立場から説明を行なう。

MDSは、ランダムアクセス法を用いレコードをアクセスするデータ・セットである。PDMのランダムアクセスアルゴリズムは、指定されたキーの値に変換を施し、データ・セット中の個々のレコードの格納エリアを表わす相対レコード番号(RR

N) を出力する。ランダムイズの結果、異なるキーの値が等しい R R N を持つことが発生する。これをシノニムが発生したと云い、ランダムイズ手法のデータ、セット編成においては、シノニム、レコードの管理と、シノニム、レコード格納エリアを決定するスペース管理の方法が性能上重要な意味を持っている。

(1) シノニム、レコードの管理

シノニム、レコードの管理方法には、大別して、Open 方法と Chaining 方法の 2 種がある。M D S では、Chaining 方法を採用しているが、通常の Chaining 方法のシノニム、レコードの管理で設定されるシノニム専用のオーバーロ、エリアを持たない、と云う Chaining 方法の中に Open 方法の持長を併用した管理を行なっている。また、シノニム、レコードを結び付けるアドレス、チェーンとしては、2 方向チェーンを有している。

(2) スペース管理と登録処理

M D S のスペース管理は、データ、セットの各シリンドラの先頭にある C C R (Cylinder Control Record) によって行われる。C C R には、どこから空エリアを探索すべきかという情報が入っており、空エリアの探索に際しては、C C R の示すアドレスより上昇順にシリアルに行なり。ここで C C R のスペース、ポインタが意味を持つのは、そのシリンドラにシノニム、レコードが 1 つ以上存在するときである。また、その時の C C R の内容は、最も新しく登録されたシノニム、レコードが格納されている R R N である。

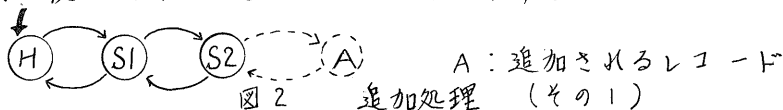
以下、M D S への登録手順を、新レコードが入るべきエリアの状態に場合分けし、それぞれについて説明する。

a) ホーム、アドレス (ランダムイズの結果出力される R R N のエリア) がスペース、エリアの場合。

追加すべきレコードをそのままホーム、レコードとして格納する。

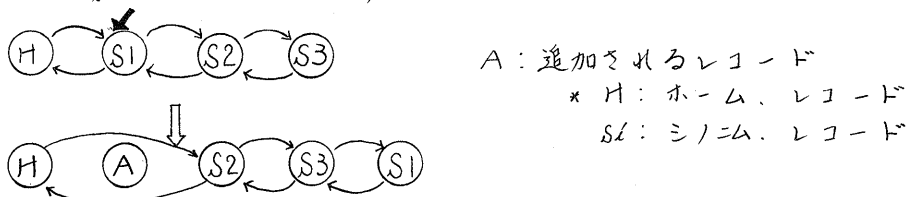
b) ホーム、アドレスにホーム、レコードが格納されていた場合。

そのホーム、レコードを出発点とするチェーンの最後のレコード迄たどり、その後追加されるレコードを結びつける。



c) ホーム、アドレスにシノニム、レコードが格納されていた場合。

これは、追加しようとするレコード A に対して、ホーム、アドレスを求め、そこを見た時、そこをホーム、アドレスとしはレコードがすでに格納されている場合である。(図3参照)



ホーム、アドレスにあるレコードを一時メモリ上にセーブして、追加されるレコードをそのエリアにホーム、レコードとして登録する。その後、セーブしたレコードを入れるべき空エリアを探索し、その空エリアにセー

アした、レコードを再登録し、シノニム、チェーンを更新する。

2. 2 解析モデル

チェーンング方法を用いたランダムイズ法法の解析に関し、L.R. Johnson, Van der Pool 等種々の論文が出されている。ここでは、既に出されているランダムイズ法法の解析に対し、我々の解析モデルの特徴と考えている、レコードの登録処理に関する解析モデルを、まず、ブロッキング、ファクタ (BF) が 1 の場合について説明し、任意の BF に拡張するに際しての要領を説明する。

2. 2. 1 記号の定義

N : データ、セットに収容可能な最大レコード数

n : データ、セットに現在収容されているレコード数

S : ブロッキング、ファクタ ($S = \text{レコード数} / 1 \text{ ブロック}$)

f : ローディング、ファクタ ($f = n / N$)

m : $m = f \times S$

2. 2. 2 BF = 1 の解析モデル

解析モデルは、アプリケーション、プログラムよりレコードの追加要求が発行された時、DBMS によって行なわれる平均の入出力回数を求めることを目的とする。

(1) 確率空間の設定

(A) P 空間

N 個のアドレス、スペースを持つデータ、セットに n 個のレコードをランダムイズして格納するとする。このとき、1 個のアドレスに着目すれば、そこに k 個のレコードがランダムイズされる確率 P_k は、各レコードが確率的に互いに独立に、ランダムイズされると仮定すると、

$$P_k = n C_k \left(\frac{1}{N}\right)^k \left(1 - \frac{1}{N}\right)^{n-k} \quad (1)$$

と云う 2 項分布で与えられる。

よく知られているように $n/N = f$ を一定として、 $N \rightarrow \infty$ 、 $n \rightarrow \infty$ とすると、2 項分布 (1) は、 f もポアッソン、パラメータとするポアッソン分布 (2) に漸近的に収束する。

$$P_k(f) = \frac{f^k \times e^{-f}}{k!} \quad (2)$$

式 (2) より明らかなことは、ランダムイズ法によれば、シノニムの発生個数は、 N と云うようなデータ、セットの大きさを表わすようなものには独立で、 f によってのみ決定されると云うことである。

式 (2) は、 X を任意のアドレス、スペースにランダムイズされているレコードの数を値としてとる確率変数とし、ポアッソン、パラメータを陽に書くと、式 (3) と表現できる。

$$P_f(f, X=k) = \frac{f^k \times e^{-f}}{k!} \quad (3)$$

確率分布 $P_f(f, X=k)$ の期待値 $E\{X\}$ は f であり、これは、1 つのアドレス、スペース当たり平均 f 個のデータ、レコードがランダムイズされていることを表わす。

次に、新には確率分布 $P'_f(f, X=k)$ を次のように定める。

$$P'_f(f, X=k) = P_f(X=k | X \neq 0) \quad (4)$$

式(4)は、 $X \neq 0$ と云う条件のもとで、 $X = k$ となる条件付確率を表す。

$$P_r(\beta, X=k) = \frac{P_r(\beta, X=k)}{\sum_{i=1}^{\infty} P_r(\beta, X=i)} \quad (5)$$

確率分布 $P_r(\beta, X=k)$ の期待値 $E[X]$ が、1個のホーレ・レコード当たりの平均のミニム・レコード・リスト長となる。

$$E[X] = \sum_{k=1}^{\infty} k P_r(\beta, X=k) = \frac{\beta}{1 - P_r(\beta, X=0)} = \frac{\beta}{1 - e^{-\beta}} \quad (6)$$

(b) Q空間

Q空間は、P空間より直接導かれる確率空間である。

今、ランダムサイズされたレコードに対し、次の操作を行なうものとする。

- i) ホーレ・レコードには1、スペース・エリアには0のラベルをつける。
- ii) ミニム・レコードには、ミニム・チェーンの逆方向から1、2、3……とラベルをつける。
- iii) ミニム・レコードをラベル0のついているエリアに移す。(図4参照)

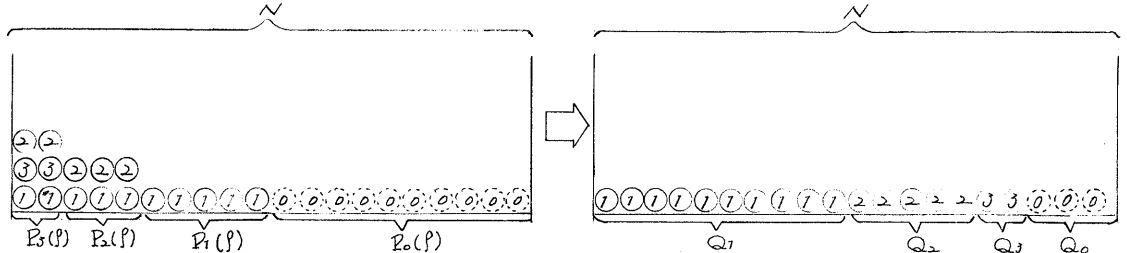


図4. Q空間の設定

前述の操作を行なった後で、 Y を、任意のアドレス・スペースに割り当てられたラベルの値をとる確率変数とし、 Y が値 k をとる確率を、 Q_k とすると、

$$Q_k = P_r(Y=k) \quad (7)$$

となる。

この Q_k は、以下の確率を表現する。

$$\begin{cases} Q_0 = P_r(Y=0): \text{スペースの存在する確率} \\ Q_1 = P_r(Y=1): \text{ホーレ・レコードの存在する確率} \\ \sum_{i=2}^{\infty} Q_i = P_r(Y>1): \text{ミニム・レコードの存在する確率} \\ Q_{i+1} (i>1): \text{ミニム・チェーンの後ろから数えて} i \text{ 番目に登録されたミニム・レコードの存在する確率} \end{cases}$$

図4からも推察されるように、 Q_i と P_i は次式で結ばれている。

$$\left. \begin{aligned} Q_0 &= 1 - \beta \\ Q_1 &= \sum_{i=1}^{\infty} P_i(\beta) \\ &\vdots \\ Q_k &= \sum_{i=k}^{\infty} P_i(\beta) \end{aligned} \right\} \quad (8)$$

(c) 探索空間

データ・セットの先頭から、アドレスの上昇順にシリアルに空エリアを探索する時、平均何回探索を行なえば、空エリアを見つけることが出来るかを考える。

Σをスペースが見つかる迄に必要な探索回数と表す確率変数とすると、 n 回の探索でスペースを見つかる確率 G_n は、次の幾何分布で表現される。

$$G_n = P(Z=n) = \sigma^{n-1}(1-\sigma) \quad (9)$$

但し $\sigma = Q_1$

式(9)の期待値が、今考えている探索回数を与えており、この期待値 $E[Z]$ は、次式で表現される。

$$E[Z] = \sum_{k=1}^{\infty} k G_k = \frac{1}{1-\sigma} = \frac{1}{1-Q_1} = \frac{1}{P_0(\beta)} \quad (10)$$

我々のDBMSのスペース探索は、CCRのアドレスに基づいて行われるため、探索するエリアには、シリアル・レコード一般的には存在せず、上記 G 空間が、スペース探索回数設定の確率空間として使用できることが分かる。また、前述の説明によって、DBMSのスペース探索の平均回数は、式(10)で表現できることが分かる。

(2) 解析モデルの設定

P , Q , G 各空間設定に際し行なった考察を適用し、データ追加処理手順に沿って、I/O回数の推定経過を図5に示す。

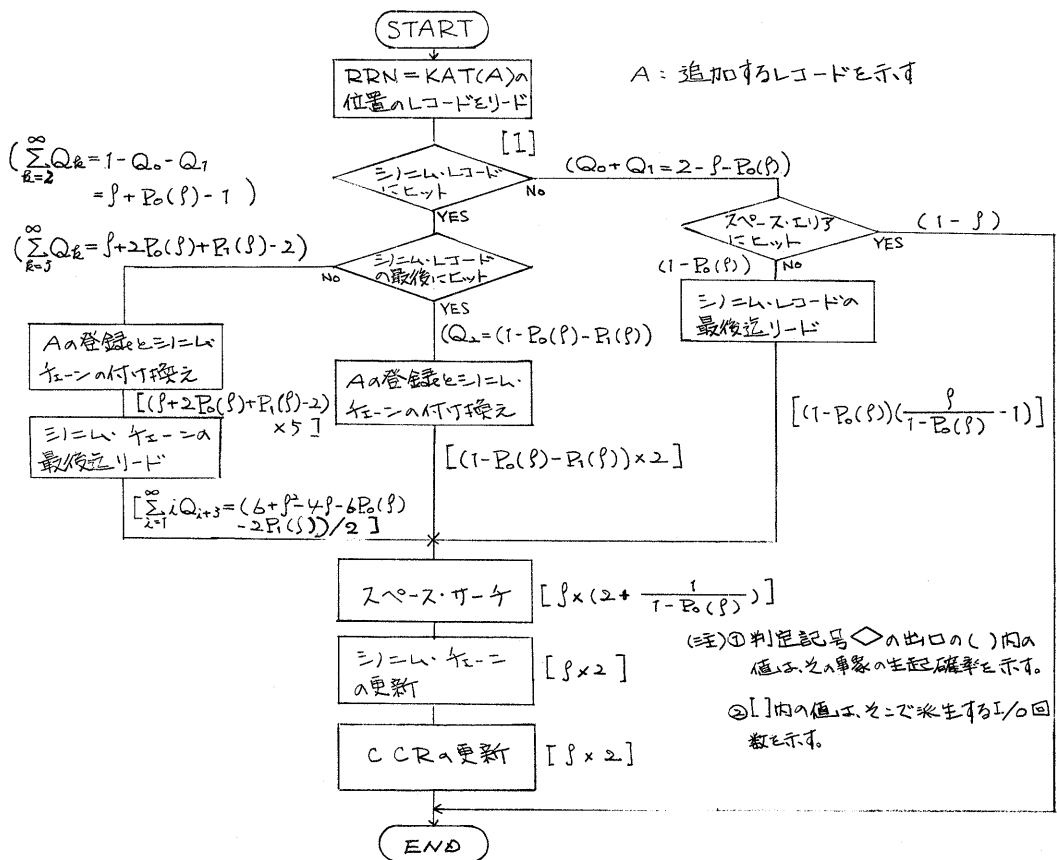


図5 MDSへのデータ追加手順とI/O回数

データ追加処理に必要なI/O回数は、図5で示した各処理区分ごとの項を足し合わせたものになるが、これは β のみの関数 $f(\beta)$ として、式(11)で表現できる。

$$f(\rho) = \frac{1}{2} \rho^2 + 10\rho - 6 P_0(\rho) + 2 P_1(\rho) - 4 + \frac{\rho}{1 - P_0(\rho)} \quad (11)$$

式(11)を用いて、多量の追加処理を行なう場合の I/O 回数の合計 TA は、次式で求められる。

$$TA = \sum_{i=1}^{an} f\left(\frac{i}{N}\right) \quad (12)$$

但し、a は追加するレコード数

2. 2...3 任意の BF の解析モデル

(1) フィジカル I/O 回数の設定

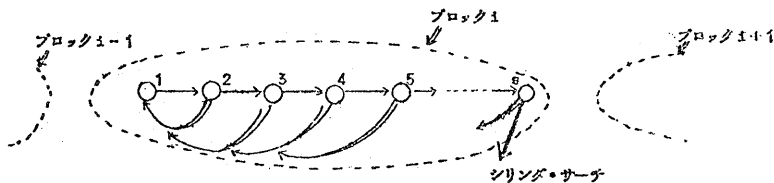
$m = \rho \times \delta$ をポアソン・パラメータとするポアソン分布を基に、I/O 回数を算定する。BF が δ の場合、任意のブロックに少なくとも 1 レコード分 m の空エリアが存在する確率 V は、次式で表現される。

$$V = \sum_{i=0}^{\delta-1} P_i(m) = \sum_{i=0}^{\delta-1} \frac{m^i \times e^{-m}}{i!} \quad (13)$$

フィジカル I/O 回数算定のロジックは、 $\delta = 1$ の場合と殆んど同様であり、本報告では省略する。

(2) ロジカル I/O 回数の設定

図 6 に、ブロック内スペース・サーチの手続きを示す。図 6 から推測できるように、1 ブロック中のデータ・エリアの位置する順序によって、空エリアの存在する確率に偏りがある。この偏りを反映し、ロジカル I/O 回数を見積るため、 ρ をポアソン・パラメータとしたポアソン分布を用いる。



ノシノニ最終レコードのエリアに、空きが無かった場合、当該ブロックの先頭に戻ることを示す。

→ シノニ最終レコードのエリアに、空きが無かった場合のスペース・サーチの順序を示す。

※ 当該ブロックに空きが無かった場合、CCR の示す、エリアからスペース・サーチに入ることを示す。

図 6. ブロック内スペース・サーチの手続き図

$\delta = 1$ の場合には、ロジカル I/O 回数 = フィジカル I/O 回数となるが、 $\delta \neq 1$ の場合には、この値が大きく異なっており、CPI 時間算出のため、ロジカル I/O 回数を算定することは、解析モデルの重要な要素となる。

ここでは、ロジカル I/O 回数推定用解析モデルの特徴を最もよく表現している、スペース・サーチに際し発生する、ロジカル I/O 回数推定方法の要点について記す。

ブロックの先頭から数えまわ ($j = 1, 2, \dots, \delta$) 番目のデータ・エリアが空である確率 V_j を求める。

i) $j = 1$ が定である確率 (V_1)

V_1 は、次の条件が満足される確率である。当該エリアに、レコードバラングマイズされている数が 0 で、かつ、本ブロック内の他のエリアには、高々 1 個しかレコードバラングマイズされていない。

$$\begin{aligned} \therefore V_1 = P_0(p) \{ & P_0(p)^{S-1} + {}_{S-1}C_1 P_0(p)^{S-2} P_1(p) + \dots \\ & + {}_{S-1}C_{S-2} P_0(p) P_1(p)^{S-2} + P_1(p)^{S-1} \} \end{aligned} \quad (14)$$

ii) $j = 2$ が定である確率 (V_2)

V_2 は、次の条件が満足される確率である。当該エリアに、レコードバラングマイズされている数が 0 で、かつ、本ブロック内の他のエリアに高々 1 個しかレコードバラングマイズされていない場合、および、 $j = 1$ のエリアの $P_0(p)$ の状態、かつ、 $j \geq 3$ のエリアのどれか一つに 2 個のレコードバラングマイズされたり他のエリアには高々 1 個しかレコードバラングマイズされていない。

$$\begin{aligned} \therefore V_2 = V_1 + P_0(p) \times P_2(p) \times (S-2) \{ & P_0(p)^{S-3} + {}_{S-3}C_1 P_0(p)^{S-4} P_1(p) \\ & + {}_{S-3}C_{S-4} P_0(p) P_1(p)^{S-4} + P_1(p)^{S-3} \} \end{aligned} \quad (15)$$

iii) $j = 3, 4, \dots, S$ も同じ同様の考察によって算出可能であるが、本稿では省略する。

スペース・サーチによって、空エリアを探し出すために必要がロジカル I/O 回数 L を示す。 C, R のアドレス・ポインタ、ブロック中の i 番目のデータ・エリアを増し示す確率を T_i とした時、 L は式 (16) で表わされる。

$$\begin{aligned} L = 1 + \sum_{i=1}^S T_i \left[\left\{ \sum_{k=1}^i \frac{V_k}{\pi (1-V_k)} \left(\frac{S\alpha}{(1-\alpha)} + \frac{1-i+k}{(1-\alpha)} \alpha \right) \right\} \right. \\ \left. + V_{i+1} \left(\frac{S\alpha}{(1-\alpha)^2} + \frac{2}{(1-\alpha)} \right) \right. \\ \left. + \left\{ \sum_{j=i+2}^S V_j \frac{j-1}{\pi (1-V_j)} \left(\frac{S\alpha}{(1-\alpha)^2} + \frac{1-i+j}{(1-\alpha)} \right) \right\} \right] \end{aligned} \quad (16)$$

$$\text{但し、} \alpha = (1-V_1)(1-V_2)\dots(1-V_S) = \prod_{i=1}^S (1-V_i), \sum_{i=1}^S T_i = 1$$

3. 性能評価システム PEAL

3.1 開発の目標

2 章で述べた解析モデルを利用し、処理能力を推定するシステム PEAL を開発している。図 7 に、PEAL 開発に際して設定した基本目標について示す。

3.2 概略仕様

図 8 に PEAL の全体概念図、図 9 に制御の流れの例を示す。

以下 PEAL 開発の内容について、入出力情報を中心に説明する。

3.2.1 PEAL のインプット

図 8 に示したように、インプットとしては、データベース定義情報、業務プログラム (UAP) モデルの 2 種があり、図 10、図 11 に、定義例、UAP モデル記述例を示す。二の中で、特に重要と考えられる UAP モデルに関し、それを記述可

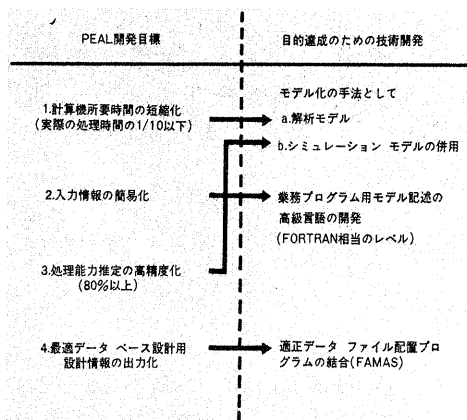


図7 PEAL開発の目標 PEAL開発に際しての開発目標と、目標を実現するために使用した技術を表わす。

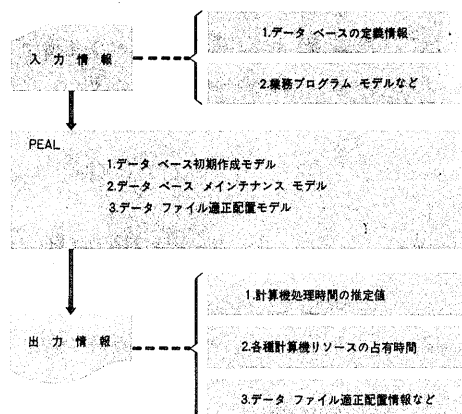


図8 PEALの概要 データベースの定義情報、業務プログラム モデルなどの記述を入力すると、PEAL処理の結果として処理能力に関するデータが出力される。

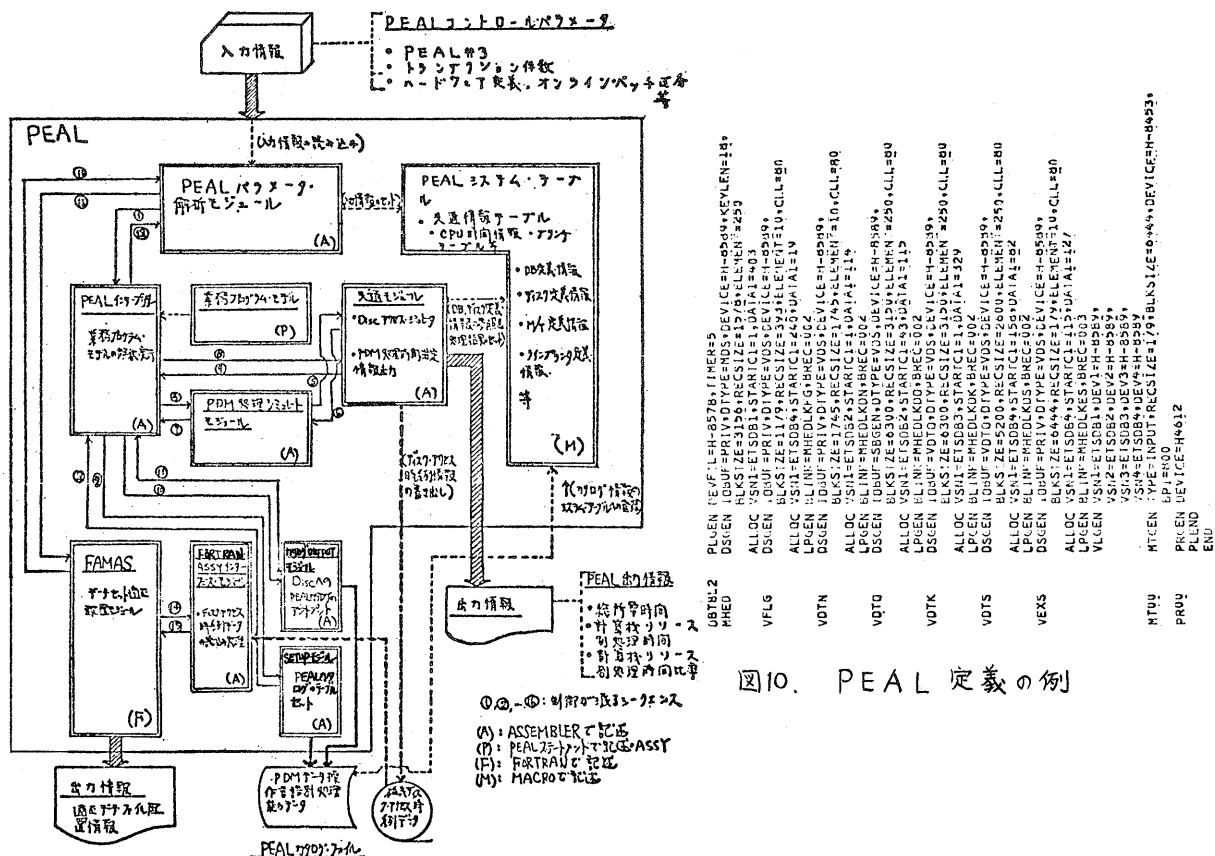


図9. PEALの制御の流れ
(DB検索・更新の場合)

[illegible]

図10. P E A L 定義の例

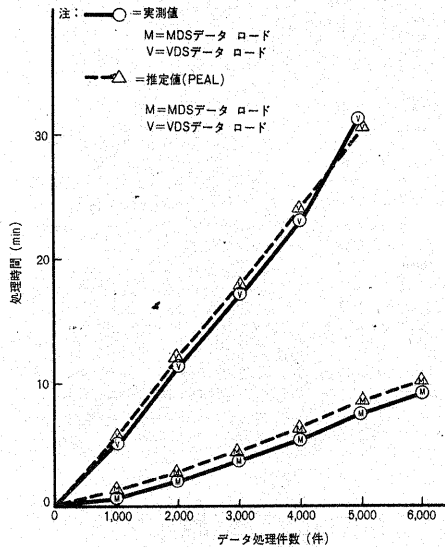


図13 PDM処理時間の推定精度 マスタ及びバリエーション各データ セットの初期作成(データ ロード)を行なうに際しての実測値データとPEALの出力結果の推定時間の比較である。

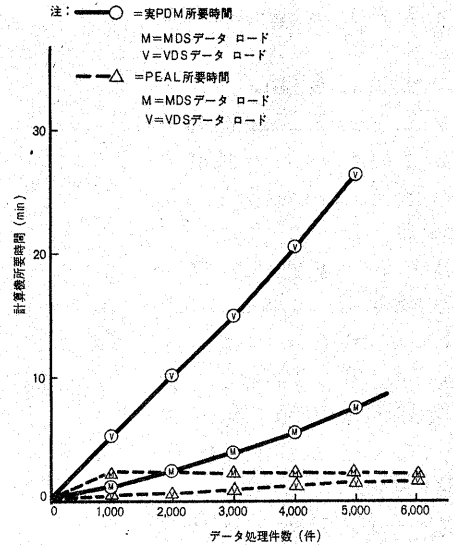


図14 PEALの計算機所要時間 PDM実データ処理の計算機所要時間と、PEALでそのモデルにおける計算機所要時間を推定する際に必要なPEAL計算機所要時間との比較を表す。

力することによって

(d) データ・セットの適正配置に関する情報が得られる。

(e) PEAL出力結果の推定精度、および、計算機所要時間。

図13、図14にPEALと実測データとの比較結果を示す。図13では、PEALの推定精度は約85%に近づいていることが分かる。また、他の種々の利用状況でも、本図と同等の結果を得ており、精度として85%～95%である、と考えられる。

図14によって、PEAL計算機所要時間として、以下の点が把握できる。

(a) MDSの初期作成では、データ処理件数が多い程PEALの効果は大きい。(高々1000点の代表点によって推定処理を行なっている理由による。)

(b) VDSの初期作成では、PEALは実PDMのデータ処理時間の $\frac{1}{10}$ 程度の計算機所要時間である。

また、データベースの検索・更新処理では、PEALの計算機所要時間の、実PDMの $\frac{1}{20} \sim \frac{1}{30}$ であることを確認している。

4. 結言

本PEALの開発によってHITAC-8000シリーズ用PDMの処理能力評価に関し、次の結果を得た。

(a) オンライン、バッチ両環境における、処理能力の推定が行なえる。

(b) 適正データ・セット配置プログラムFAMASとの結合によってデータ・セットの適正配置情報が得られる。

(c) 以上の結果、データベース設計の妥当性の検証が可能となる。

今回説明した解析モデル、性能評価ツールPEALの開発がみられるように、データベース・システムの建設、運用段階でのシステム性能、および、操作性の向上に重点を置いて、その改善を続けている。その他の性能評価の面では、シミュレート、ソフトウェアモック、解析ツールなどの評価手法が蓄積されているので、これから適じ汎用的なシステム設計手法の確立を着実に推進していきたい。