

データベースプロセッサ RINDA のシステム構成法

野瀬 純郎† 速水 治夫††

データベースプロセッサ RINDA に関し、システム構成と性能の関係について考察する。RINDA はリレーショナル・データベースにおいて、インデックスの利用が困難な非定型の検索を高速に処理する専用ハードウェアで、I/O インタフェースを介して汎用計算機に接続される付加型プロセッサであり、種々のワークロードに対して柔軟に適応可能とするため、ビルディングブロック構成をとっている。通常の定型検索および更新は汎用計算機上で走行するデータベース管理システム(DBMS)のみで処理を行い、非定型検索についてはDBMSの制御下で RINDA が処理を行う。本論文では、定型検索と非定型検索が混在して処理される場合に、種々のワークロード特性に対して、所要性能を達成するために必要なプロセッサ台数、接続方式、バッファ容量等のリソース条件を分析することにより、RINDA の適用領域と最適システム構成を求める手法を示すとともに、ビルディングブロック構成法の有効性を示す。

System Configuration of a Database Processor RINDA

JUNRO NOSE† and HARUO HAYAMI††

This paper investigates parameters for configuring RINDA, a database processor system. RINDA is an attached processor used as dedicated hardware for high-performance processing of non-indexed queries, for cases where indexing is not feasible in a relational database. This processor connects to a general purpose computer via an I/O interface. A building block approach is adopted for flexible adaptation to a variety of workloads. Ordinary indexed queries and updates are handled by the database management system (DBMS) alone, running on the general purpose computer, whereas RINDA processes non-indexed queries under DBMS control. This paper clarifies the areas of RINDA applicability and the optimal system configuration, by determining the required number of processors, connection method, buffer capacity and other resource conditions for realizing the required performance under various workloads, when processing includes both indexed and non-indexed queries.

1. はじめに

リレーショナル・データベース (RDB) は、データモデルの仮想化の水準が高く、柔軟性に優れている反面、アクセス経路があらかじめ決められない非定型検索では多大な処理時間を要し、汎用計算機上のソフトウェアのみでは、性能面から見て大規模データベースへの適用には限界がある。このため、大量のデータ処理を汎用計算機から切り離し、専用ハードウェアで高速に行う、データベースプロセッサの開発・導入が進められている^{1)~3)}。

データベースプロセッサは、汎用計算機との機能分担という観点から後置型と付加型の2つに分類できる。後置型はデータベース処理のすべてを汎用計算機

とは独立に、かつ自己完結的に行うもので、ホスト計算機のアーキテクチャに依存しないため、種々の汎用計算機に接続できるが、コストが大きくなるという特徴を持つ^{4), 5)}。一方、付加型はデータベース処理機能の一部を専用化し、特定の汎用計算機に付加され、ホスト計算機と協調して動作するもので、ホスト計算機のアーキテクチャに依存するため、接続関係には制限があるが、比較的小さな追加コストで性能向上が実現できるという利点がある^{6), 7)}。

著者らは、RDB において、インデックスの利用が困難な非定型の検索処理を高速化するため、汎用計算機 DIPS シリーズ^{8), 9)}用の付加型データベースプロセッサ RINDA (Relational database processor)¹⁰⁾を開発した。RINDA は適用システムのデータベース容量、検索トラフィック、検索条件適合率などの種々のワークロード特性に対して、経済的に所要性能を達成可能とするために、ビルディングブロック構成をとっている。具体的にはサーチ処理を受け持つ内容検索

† NTT ソフトウェア株式会社
NTT Software Corporation

†† NTT 情報通信網研究所
NTT Network Information Systems Laboratories

プロセッサ CSP (Content Search Processor)¹¹⁾と、ソート処理および結合演算を受け持つ関係演算プロセッサ ROP (Relational Operation Accelerating Processor)¹²⁾で構成される。それぞれ I/O インタフェースを介して DIPS に接続され、ホスト計算機上で走行するデータベース管理システム (DBMS) によって制御される。

本論文では、種々のワークロード特性に対する RINDA の適用領域と最適システム構成を求める手法を示すとともに、採用したビルディングブロック構成法の有効性を示す。2章で RINDA の概要と特徴を紹介し、3章で評価の観点と評価項目、4章で評価モデルについて述べ、5章でシミュレーションによる評価結果を示す。

2. RINDA の概要

2.1 RDB の問題点と RINDA での解決法

RDB における主要な処理は、定型の検索・更新処理と非定型の検索処理である。定型の検索処理の代表例は、表 (データ) の中からユニークなキーを用いて単一の行を選択する処理であり、インデックスを使用することにより、ソフトウェアのみでも十分に実用的な性能が実現される。

他方、非定型の検索処理には、あらかじめインデックスが作成されていない列に対して、条件を指定した検索や、複数の表を結合する処理などがあり、これらを従来のソフトウェアのみで実行すると多大な処理時間を要する。これは、インデックスが使用できないため、表中のすべての行を逐次読み出して検索条件を判定するサーチ処理や、結合演算の前処理としてのソート処理などに多大な CPU 時間と I/O 時間を必要とすることによる。

上記の問題を解決するために、サーチ処理とソート処理および結合演算処理を、それぞれ専用のハードウェアである CSP と ROP で実行するデータベースプロセッサ RINDA を開発した。サーチ処理については、ディスク (DK) からのデータ読み出しと検索条件の判定をパイプライン処理し、条件に合致した行の中の必要な列のみをホスト計算機に転送する。したがって、CSP の動作時間はほぼ DK のデータ転送時間となる。また、ソート処理については、ホスト計

算機からの表の受け取りとソート・結合演算をパイプライン処理し、その結果をホスト計算機に返送する。したがって、ROP の動作時間はほぼ I/O チャンネルのデータ転送時間となる。

2.2 ハードウェア

RINDA のシステム構成例を図 1 に示す。CSP, ROP とホスト計算機間、および CSP とディスク制御装置 (DKC) 間にはともに I/O インタフェースで接続する。

CSP の検索処理は表全体を読み出すため、表のデータ量に比例した処理時間を要する。表を複数 (n 個) の DK に分割して格納し、 n 台の CSP で並列検索すると、実行時間をほぼ $1/n$ に短縮することが可能である。RINDA はユーザプログラムに表の分割を意識させることなく、複数の CSP, DK を使用した並列検索を実行する。なお、ソートは全データの比較処理を必要とするため、分割することなく、ある 1 トランザクションの検索処理に関しては、割り当てられた 1 台の ROP で処理する。

CSP と ROP の 2 種のハードウェア構成にし、おのこの台数、接続方式を変えたことにより、システムごとに異なるワークロード特性に対して、所要性能を達成するための最適なシステム構成をとることが可能となる。

2.3 ソフトウェア

ホスト計算機上で RINDA を制御するソフトウェアの構成を図 2 に示す。制御ソフトウェアは既存の DBMS に、RINDA 対応の制御機能を追加した構成とした。これは、同一のデータベースに対して RINDA を使用したアクセス (非定型検索) と、使用しないアクセス (定型検索) を同時に行うことを可能とするためである。

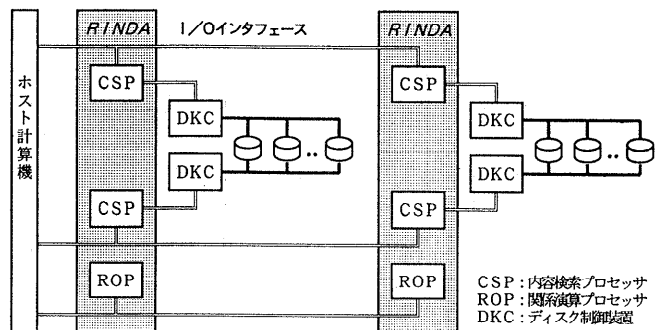


図 1 RINDA システムの構成例

Fig. 1 Typical RINDA system organization.

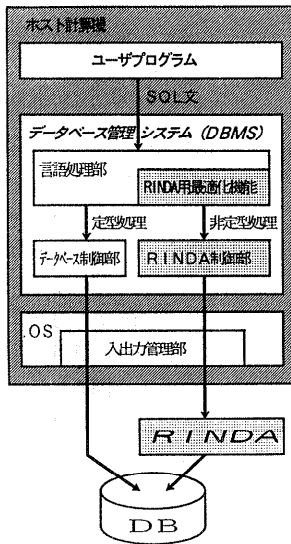


図 2 データベース管理システムの構成

Fig. 2 Database management system organization.

ユーザプログラムから発行された SQL 文は、言語処理部により構文解析、意味解析、最適化が行われた後に実行される。RINDA 制御部は RINDA の実行制御、入出力管理などを行う。データベース制御部は、RINDA を使用しないでソフトウェアのみによる検索・更新処理を実行するほか、データベースの生成、トランザクション管理、排他制御、資源管理などの、データベース処理として一元管理が必要な制御機能もつかさどる。

3. 評価項目

システムの性能は、システムを構成するリソースの条件（リソース個々の性能、台数、構成など）と、システムにかかるワークロードの特性（処理量、トラヒックなど）の関係で決まる。

本論文では、定型検索と非定型検索が混在する場合において、種々のワークロード特性に対して所要の性能を得るために必要な RINDA のリソース条件は何かを求める。これにより、RINDA の適用領域と最適システム構成を明らかにする。ここでとりあげる性能の評価尺度、リソース条件およびワークロード特性は以下のとおりとする。

(1) 性能評価尺度

①スループット

- 定型検索のスループット： P_S
- 非定型検索のスループット： P_N

②ターンアラウンドタイム (TAT)

- 定型検索の TAT： T_S
- 非定型検索の TAT： T_N

(2) リソース条件

- ①CSP の台数： N_C
- ②ROP の台数： N_R
- ③CSP, ROP の接続方式
 - 並列接続：それぞれ独立の I/O インタフェース線で接続
 - 直列接続：1本の I/O インタフェース線に複数台のプロセッサを接続

④単位検索範囲： A_U

⑤ホスト計算機のバッファの容量： B

⑥ワーク用 DK の台数： N_W

(3) ワークロード特性

①全検索範囲： A_T

②定型検索トラヒック： R_S

③非定型検索トラヒック： R_N

- 単純検索トラヒック： R_{NS}

- グループ化検索トラヒック： R_{NG}

- 結合検索トラヒック： R_{NJ}

$$(R_N = R_{NS} + R_{NG} + R_{NJ})$$

④ CSP の検索条件適合率： λ_C

⑤ ROP の検索条件適合率： λ_R

ここで、ホスト計算機のバッファとは、CSP からワーク用 DK へのデータ転送、およびワーク用 DK と ROP 間のデータ授受に使用するもので、CSP が複数台の場合、各 CSP は B/N_C の容量を使用する。

全検索範囲とは、あるトランザクションがデータベースにアクセスする際、検索する表の全容量であり、単位検索範囲とは、CSP が1回の処理当たりに検索する容量である。CSP が複数台の場合、各 CSP はそれぞれ A_U/N_C の範囲を検索する。また、検索条件適合率とはある検索条件でデータベースにアクセスした際、条件に適合するデータが抽出される割合のことである。

4. 評価モデル

定型検索と非定型検索の混在処理モデルとして、ここでは単一の表に対してアクセスが競合する場合を考える。すなわち、DK 内の単一の表に対して、定型検索および非定型検索の処理が、それぞれ所定のトラヒックで要求される場合に、両者のスループットを確保し、かつ満足すべき TAT を得るために必要十分

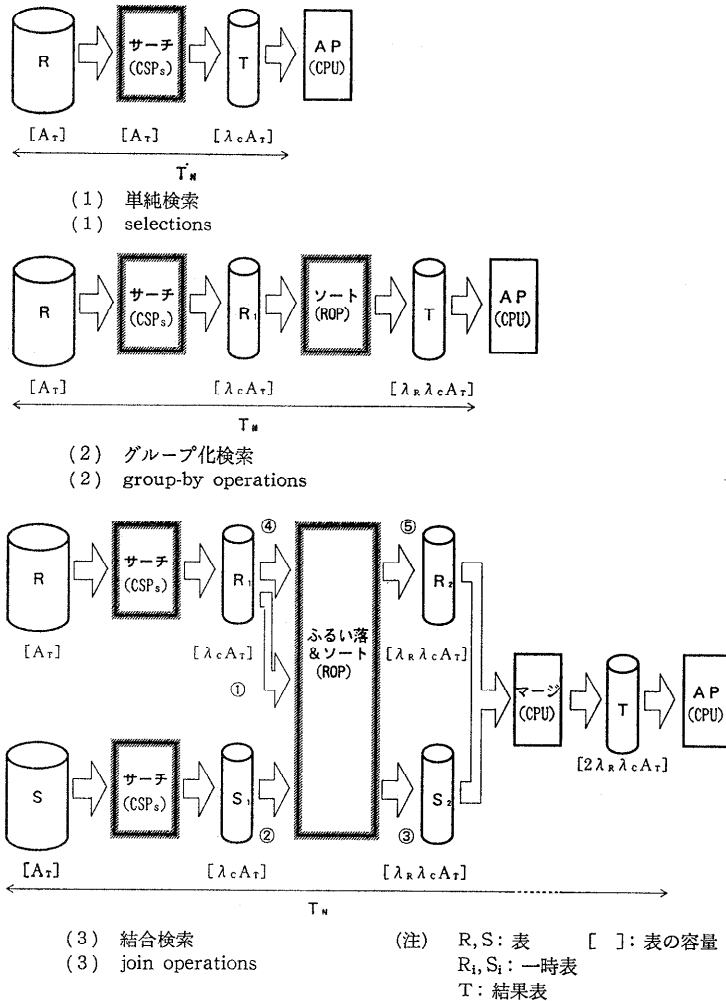


図3 非定型検索処理モデルの処理概要
Fig. 3 RINDA processing model.

なりリソースの条件をシミュレーションにより求める。この際、性能上の問題の焦点は DK および I/O インタフェースのアクセス競合にあるので、ホスト計算機での処理内容は評価モデルの対象外とし、CSP, ROPを含む I/O 系の処理に着目して評価する。

定型検索および非定型検索は、以下の処理で代表させた(図3参照)。

①定型検索

インデックスを索引し、該当ページから1行のデータを読み出す。上段のインデックスはメモリ上にあり、最下段は表と同一のDKに格納されているとし、1検索当たり同一DKを2回アクセスする。

②非定型検索

(i)単純検索: 表全体を読み出し、検索条件に適合するデータを抽出する。

(ii)グループ化検索: 単純検索の結果の一時表を、ある列の値で分類(グループ化)する。

(iii)結合検索: 2つの表を単純検索した後、両表に共通する列の値で結合する。

評価モデルを表1に、システム構成例を図4に、処理の概略フローを図5に示す。また、図4には定型検索時および非定型検索時のデータベース(インデックス、表)へのアクセスルートもあわせて示す¹³⁾。

5. 評価結果

すでに述べたように、性能はリソース条件とワークロード特性の関係で決まるが、非常に微妙なケースもある。すなわち、あるワークロード特性に対して所要の性能を達成するためのリソース条件には、複数の組み合わせが考えられる場合がある。しかし、これらのすべてについての比較検討には膨大な工数を要し、現実的でないため、評価にあたっては、まず一次評価として個々の特性のうちのひとつ

の変化が性能に及ぼす影響を調べ、定性的な傾向を明らかにし、次にそれらを総合する手順をとる。ここで、所要性能の基準は、非定型検索に関する要求条件が比較的緩い会話処理~バッチ処理を想定して、ウィスコンシン・ベンチマークの100万行相当の表に対し、発生した検索要求をシステム内に滞留させることなく処理をすべて完了させ、かつ定型検索のTATは0.2秒以下、非定型検索のそれは300秒以下とした。これらの値は以下の評価を行うために一例として設定したものである。

5.1 一次評価

(1) CSP, ROPの台数(N_C, N_R)と接続方式 N_Cの増加に伴う性能改善は、非定型検索に対して

表 1 評価モデル
Table 1 Evaluation model.

項	目	内 容
リソース条件	CSP, ROP の台数	①CSP: 1, 2, 4 台 ②ROP: 1, 2 台
	CSP, ROP の接続方式	①直列 (いもづる) 方式 ②並列 (独立) 方式
	ホスト計算機のバッファ容量	2, 4, 8, <u>32</u> ~160 ページ
	ワーク用 DK の台数	1, 2 台
	DK の諸元	①平均シーク時間: 12.0 ms ②平均回転待ち時間: 8.3 ms ③データ転送時間: 4.4 MB/s
ワークロード特性	全検索範囲	12, 800, 25, 600, <u>51, 200</u> , 76, 800 ページ
	単位検索範囲	64, 128, 192, <u>256</u> ~1,152 ページ
	検索トラヒック	①定型検索 : 1, 2, <u>3</u> , 4, 5 万件/時 ②非定型検索 ・単純検索 : 10, 15, <u>20</u> , 25, 30 件/時 ・グループ化検索: 10, 15, <u>20</u> , 25, 30 件/時 ・結合検索 : 10, 15, <u>20</u> , 25, 30 件/時
	検索条件適合率	①CSP: <u>1</u> , 5, 10% ②ROP: 25, 50, <u>75</u> , 100%

注 1) 1 ページ=4KB.

注 2) 下線付き: 本評価において標準値と想定したもの.

注 3) 全検索範囲=51,200 ページは, ウィスコンシン・ベンチマークの 100 万行のデータベースに相当.

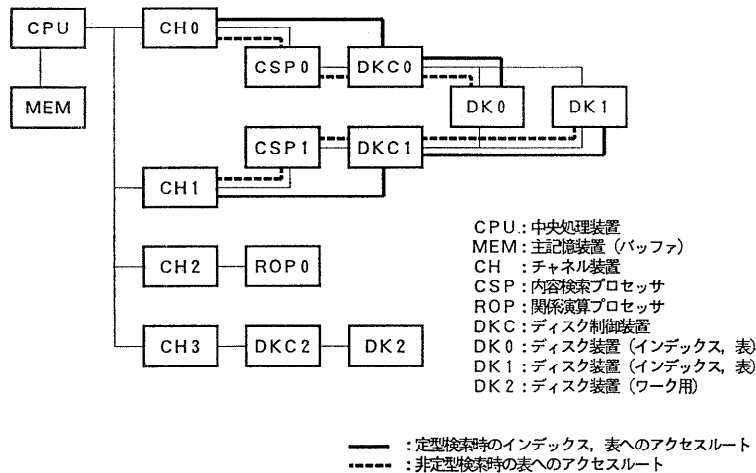


図 4 シミュレーションモデルのシステム構成例
Fig. 4 Typical system organization at simulation.

のみならず, 定型検索にも有効であり, その効果は顕著である (図 6 参照). 4 章で述べた標準のワークロード特性 (表の容量はウィスコンシン・ベンチマークの 100 万行に相当) においては, 少なくとも N_c は 2 台

必要とし, N_r は 1 台で十分である. また, それらの接続方式は性能にほとんど影響せず, すべて直列接続で十分である (図 7 参照). CSP, ROP はホスト計算機間とのデータ転送時のみチャネル装置と結合する.

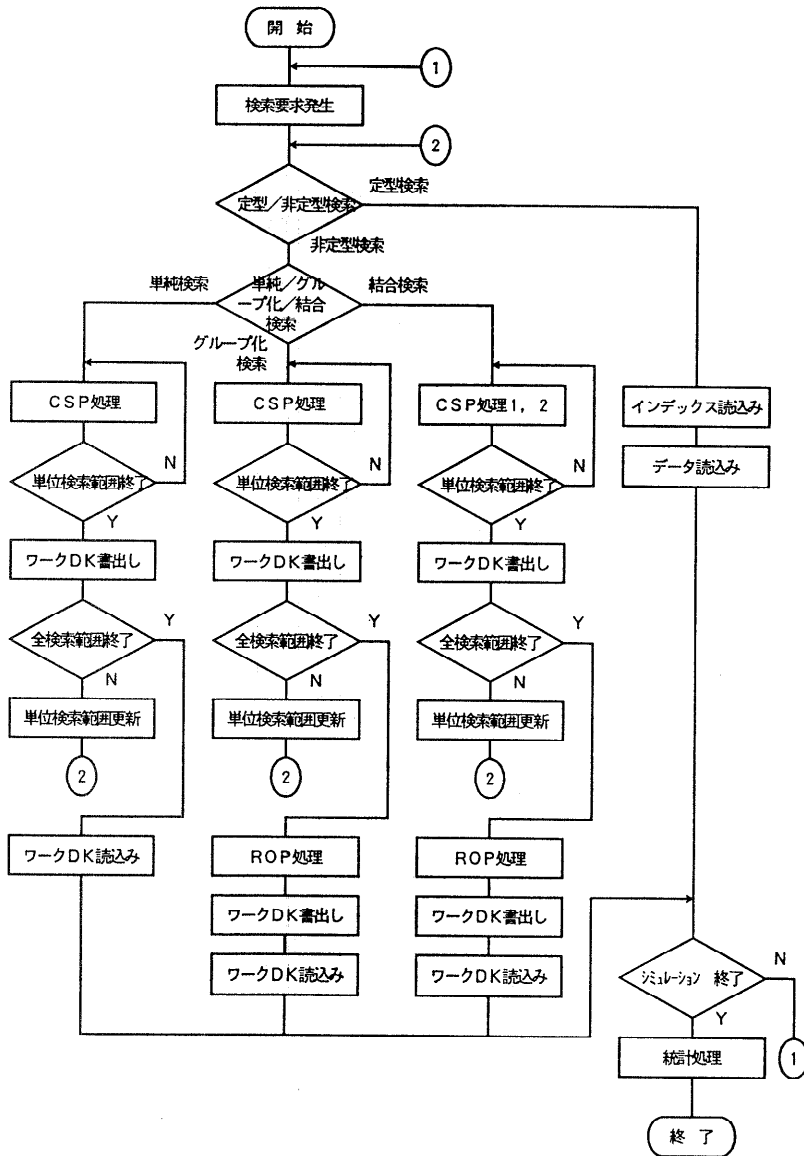


図 5 シミュレーションの概要
Fig. 5 Outline of simulation flowchart.

したがって、 λc が低い (1%) 場合は直列接続で十分であることが確認できた。これにより、チャネル装置の節減が期待できる。

(2) ホスト計算機のバッファ容量 (B)

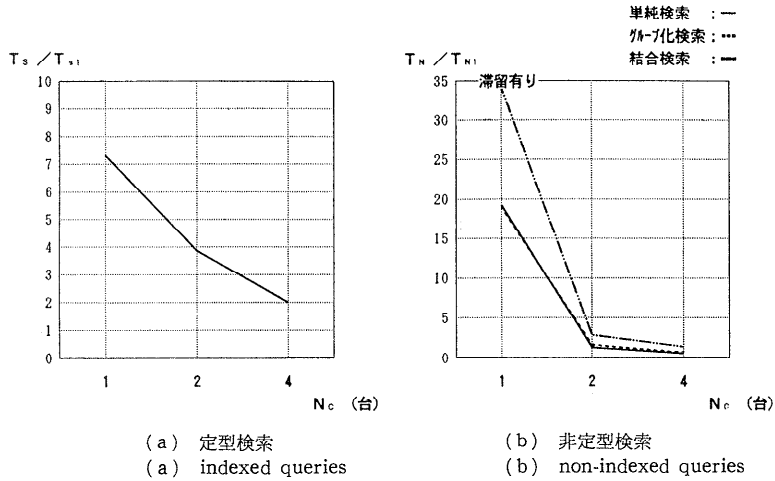
単位検索範囲の処理が終了するか、バッファが満杯になると、CSP は検索の途中で処理を中断し、その時点で定型検索の要求があれば、それらを優先処理するため¹³⁾、 B は以下が目安となる。

$$B \cong A_u \times \lambda c$$

シミュレーション結果からも B はある程度以上大きくしても効果はなく (図 8 参照)、むしろリソースの無駄となることがわかる。

(3) 全検索範囲 (A_T)

A_T はワークロード特性の最も基本となるもので、その大きさにほぼ比例して T_s , T_N は増大する。ただし、 T_N への影響の方が顕著であり、 $A_T=76,800$



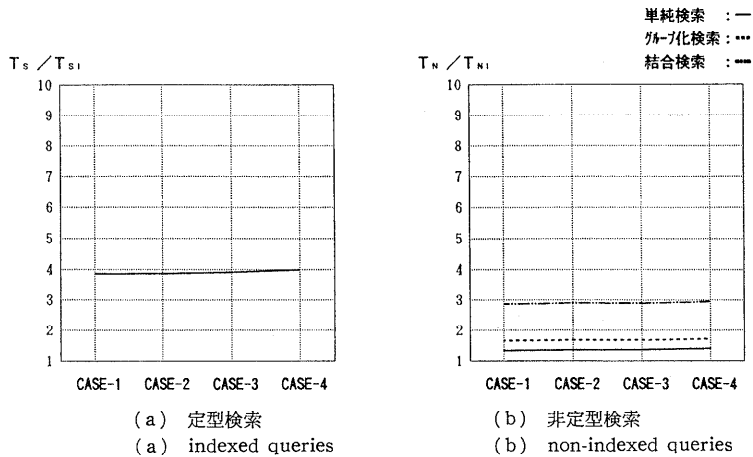
(a) 定型検索
(a) indexed queries

(b) 非定型検索
(b) non-indexed queries

《評価条件》 N_R : 1台 R_S : 3万件/時
 B : 32ページ R_{NS} : 20件/時
 A_T : 51,200ページ R_{NG} : 20件/時
 A_U : 256ページ R_{NC} : 20件/時
 λ_C : 1%
 λ_R : 75%

《注》 T_{s1} : $R_N=0$ での定型検索1件の TAT
 T_{n1} : $R_S=0$ での非定型検索 (単純) 1件の TAT
 (ただし, $N_C=N_R=N_W=1$, $B=32$ ページ, $A_T=51,200$ ページ, $A_U=256$ ページ, $\lambda_C=1\%$, CSP, ROP は直列)

図 6 CSP 台数と性能との関係
 Fig. 6 Relationship between number of CSPs and performance.



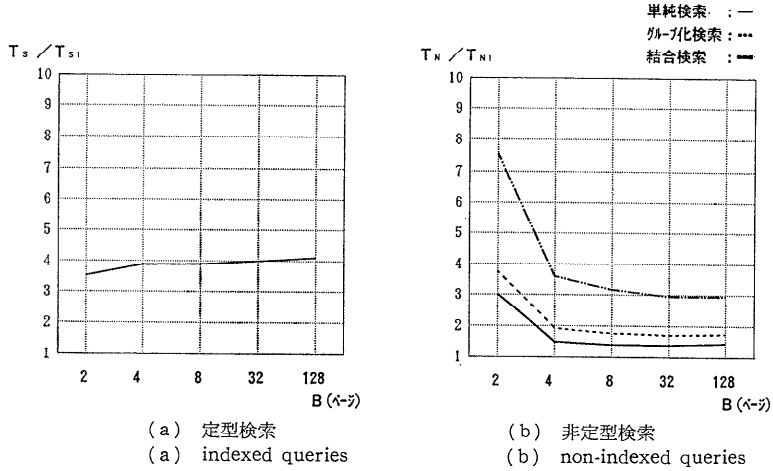
(a) 定型検索
(a) indexed queries

(b) 非定型検索
(b) non-indexed queries

《評価条件》 N_C : 2台 R_S : 3万件/時
 N_R : 1台 R_{NS} : 20件/時
 B : 32ページ R_{NG} : 20件/時
 A_T : 51,200ページ R_{NC} : 20件/時
 A_U : 256ページ λ_C : 1%
 λ_R : 75%

《注》 T_{s1} : $R_N=0$ での定型検索1件の TAT
 T_{n1} : $R_S=0$ での非定型検索 (単純) 1件の TAT
 (ただし, $N_C=N_R=N_W=1$, $B=32$ ページ, $A_T=51,200$ ページ, $A_U=256$ ページ, $\lambda_C=1\%$, CSP, ROP は直列)

図 7 CSP, ROP の接続方式と性能との関係
 Fig. 7 Relationship between system configuration and performance.

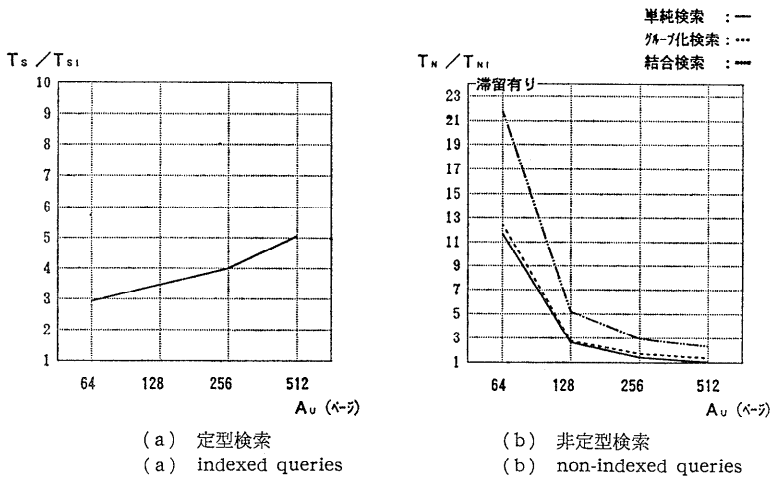


《評価条件》 N_C : 2台 R_S : 3万件/時
 N_R : 1台 R_{NS} : 20件/時
 A_T : 51,200ページ R_{NG} : 20件/時
 A_U : 256ページ R_{NC} : 20件/時
 λ_C : 1%
 λ_R : 75%

CSP, ROP の接続方式: 直列接続

《注》 T_{s1} : $R_N=0$ での定型検索 1 件の TAT
 T_{n1} : $R_S=0$ での非定型検索 (単純) 1 件の TAT
 (ただし, $N_C=N_R=N_W=1$, $B=32$ ページ, $A_T=51,200$ ページ, $A_U=256$ ページ, $\lambda_C=1\%$, CSP, ROP は直列)

図 8 ホスト計算機のバッファ容量と性能との関係
 Fig. 8 Relationship between buffer capacity and performance.

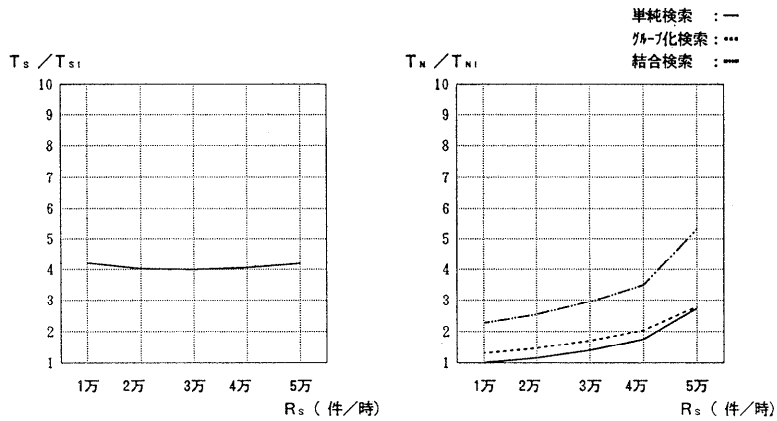


《評価条件》 N_C : 2台 R_S : 3万件/時
 N_R : 1台 R_{NS} : 20件/時
 B : 32ページ R_{NG} : 20件/時
 A_T : 51,200ページ R_{NC} : 20件/時
 λ_C : 1%
 λ_R : 75%

CSP, ROP の接続方式: 直列接続

《注》 T_{s1} : $R_N=0$ での定型検索 1 件の TAT
 T_{n1} : $R_S=0$ での非定型検索 (単純) 1 件の TAT
 (ただし, $N_C=N_R=N_W=1$, $B=32$ ページ, $A_T=51,200$ ページ, $A_U=256$ ページ, $\lambda_C=1\%$, CSP, ROP は直列)

図 9 単位検索範囲と性能との関係
 Fig. 9 Relationship between single search extent and performance.



(a) 定型検索
(a) indexed queries

(b) 非定型検索
(b) non-indexed queries

《評価条件》 N_C : 2台 R_{NS} : 20件/時
 N_R : 1台 R_{NG} : 20件/時
 B : 32ページ R_{NC} : 20件/時
 A_T : 51,200ページ λ_C : 1%
 A_U : 256ページ λ_R : 75%

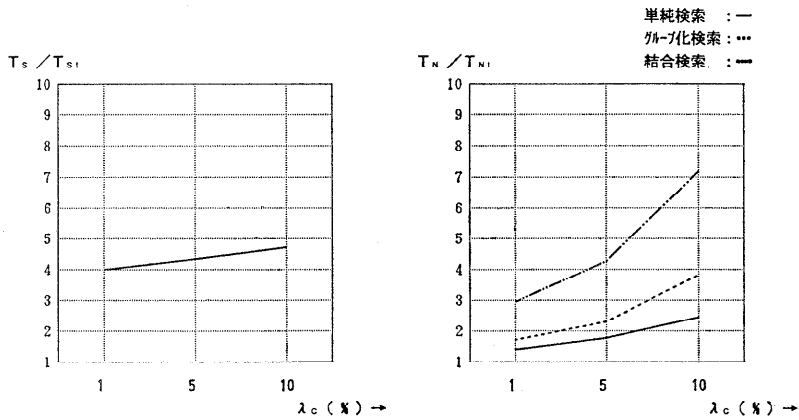
CSP, ROP の接続方式: 直列接続

《注》 T_{s1} : $R_N=0$ での定型検索 1 件の TAT

T_{N1} : $R_S=0$ での非定型検索 (単純) 1 件の TAT

(ただし, $N_C=N_R=N_W=1$, $B=32$ ページ, $A_T=51,200$ ページ, $A_U=256$ ページ, $\lambda_C=1\%$, CSP, ROP は直列)

図 10 定型検索のトラフィックと性能との関係
 Fig. 10 Relationship between traffic and performance.



(a) 定型検索
(a) indexed queries

(b) 非定型検索
(b) non-indexed queries

《評価条件》 N_C : 2台 R_S : 3万件/時
 N_R : 1台 R_{NS} : 20件/時
 B : 32ページ R_{NG} : 20件/時
 A_T : 51,200ページ R_{NC} : 20件/時
 A_U : 256ページ λ_R : 75%

CSP, ROP の接続方式: 直列接続

《注》 T_{s1} : $R_N=0$ での定型検索 1 件の TAT

T_{N1} : $R_S=0$ での非定型検索 (単純) 1 件の TAT

(ただし, $N_C=N_R=N_W=1$, $B=32$ ページ, $A_T=51,200$ ページ, $A_U=256$ ページ, $\lambda_C=1\%$, CSP, ROP は直列)

図 11 CSP 検索条件適合率と性能との関係
 Fig. 11 Relationship between CSP hit ratio and performance.

ページ (300 MB) では、 $N_c = 2$ 台の場合システム内にトランザクションの滞留が生じ (スループットが低下)、 T_N が急増する。

(4) 単位検索範囲 (A_U)

CSP が 1 回の検索処理でサーチする範囲 A_U を大きくすると、それに伴い T_N は小さくなり、逆に T_s は大きくなるという、逆の関係にある (図 9 参照)。これは、CSP が DK をマルチトラック・リードするため、1 回の検索範囲が大きいほどヘッドのシーク、サーチのオーバヘッドが小さくなることと、第 2 項で述べたように、CSP の 1 回当たりの処理が終了する度に、待ち状態にある定型検索を優先して実行するためである。 A_U はシステムごとに任意に決定することができるので、適切な選択によりバランスのとれた性能の達成が可能である。標準のワークロード特性においては、 $A_U = 256$ ページ (1 MB) 程度が適当である。

(5) 定型検索トラヒック (R_s)

R_s が増加しても、定型検索の処理が優先するため、 $R_s = 5$ 万件/時程度までは T_s への影響は全くない。その反面、 T_N は増大するので、ある限度を超えると非定型検索の所要性能の確保のために N_c の追加などの対処が必要となる (図 10 参照)。

(6) 非定型検索トラヒック (R_N)

R_N の増加に伴い、 T_s は若干増加する程度であるが、 T_N への影響は大きく、 $N_c = 2$ 台では $R_N = 90$ 件/時は処理不能である。

なお、本評価モデルでは非定型検索のうち結合検索は CSP 処理を 2 回行うため、 T_N は単純検索、グループ化検索のほぼ 2 倍になる。

(7) CSP の検索条件適合率 (λ_c)

λ_c は非定型検索の処理量の基本となる CSP の出力量を決めるため、 T_N への影響は大きい、 T_s への影響は小さい (図 11 参照)。

(8) ROP の検索条件適合率 (λ_r)

λ_r は λ_c で 1 次的に選択されたデータ量に対して効

A_T	R_N R_s	λ_c	30 件/時 (R_{Ns}, R_{Ng}, R_{Nj} は各 10 件/時)				
			1 万件/時	2 万件/時	3 万件/時	4 万件/時	5 万件/時
25,600 ページ		1%	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]
		5%	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]
		10%	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]
51,200 ページ		1%	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]
		5%	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]
		10%	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [384; 64]
76,800 ページ		1%	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [256; 32]
		5%	[1; 1] [256; 32]	[1; 1] [256; 32]	[1; 1] [384; 32]	[1; 1] [640; 64]	[1; 1] [640; 64]
		10%	[1; 1] [384; 64]	[1; 1] [512; 64]	[1; 1] [512; 64]	[1; 1] [640; 96]	[1; 1] [640; 96]

《評価条件》 $\lambda_r = 75\%$ 《凡例》 $[N_r; N_w]$, : $N_c = 1$
 : $N_c = 2$
 : $N_c = 4$
 : CSP と ROP は直列接続

図 12 ワークロード特性とリソース条件の関係 (1/3)
 Fig. 12 Relationship between workload and resources (1/3).

いてくるもので、 λ_c が小さい場合は定型検索、非定型検索ともに性能上の影響はほとんどない。

5.2 総合評価

前節の内容をまとめると、定性的には以下のようになる。

① システムの性能に最も寄与するリソース条件は CSP の台数であり、他は副次的である。

② CSP, ROP の接続方式は、CSP の処理能力に余裕がある状況では直列接続で十分であるが、高負荷時には並列接続することにより、CSP の台数を増加させなくても対応可能な領域がある。

③ ホスト計算機のバッファ容量は、単位検索範囲の大きさと CSP の検索条件適合率の積を目安としてよ

く、それ以上大きくしても効果はない。

④ワークロード特性のうち、性能に及ぼす影響という観点から基本的なものは、全検索範囲および非定型検索トラヒックであり、逆に前項①からこの組み合わせで必要な CSP の台数は決まると予想される。

⑤単位検索範囲の大きさは、定型検索の性能と非定型検索の性能に対して逆の作用をするので、両性能のバランスを考慮して決定する必要がある。

以上の RINDA の基本的な特性を基に、種々のワークロード特性に対して、所要の性能を達成するために必要なリソース条件を求めた結果を図 12～図 14 にまとめて示す。これらの情報を分析することにより以下が明らかとなった。

①CSP の必要台数を決定するにあたり、考慮すべきワークロード特性は何かを重回帰分析¹⁴⁾により求めると、上位 3 つの t 検定値は以下ようになる (決定係数=0.81, 有意水準=1%)。

- (i) 全検索範囲 : $t=18.6$
- (ii) 非定型検索トラヒック : $t=13.8$
- (iii) CSP 検索条件適合率 : $t=2.7$

すなわち、CSP の必要台数は全検索範囲と非定型検索トラヒックをもとに決定すればよく、5.2 節で述べた内容が定量的に裏付けられた。

②同様に、CSP と ROP の接続方式を決定するにあたり、考慮すべきワークロード特性は何かを判別分析¹⁴⁾により求めると、上位 3 つの F 検定値は以下ようになる (正答率=0.90, 有意水準=1%)。

- (i) CSP 検索条件適合率 : $F=31.0$
- (ii) 全検索範囲 : $F=22.6$
- (iii) 非定型検索トラヒック : $F=16.9$

すなわち、CSP と ROP の接続方式は CSP 検索条件適合率と全検索範囲をもとに決定すればほぼ十分である。

A _T	R _w R _s λ _c	60 件/時 (R _{NS} ; R _{NG} ; R _{NJ} は 各 20 件/時)				
		1万件/時	2万件/時	3万件/時	4万件/時	5万件/時
25,600 ページ	1%	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]
	5%	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]
	10%	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]
51,200 ページ	1%	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [384 ; 32]
	5%	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [384 ; 32]	[1 ; 1] [512 ; 32]
	10%	[1 ; 1] [384 ; 64]	[1 ; 1] [384 ; 64]	[1 ; 1] [384 ; 64]	[1 ; 1] [512 ; 64]	[1 ; 1] [640 ; 96]
76,800 ページ	1%	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]
	5%	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [256 ; 32]	[1 ; 1] [384 ; 32]	[1 ; 1] [384 ; 32]
	10%	[1 ; 2] [640 ; 96]	[1 ; 2] [768 ; 96]	[1 ; 2] [640 ; 96]	[1 ; 2] [640 ; 96]	[1 ; 2] [640 ; 96]

《評価条件》 λ_R = 75 % 《凡例》 [N_R ; N_w] , [A_U ; B] ,

[〇 ; 〇] : N_c = 1
 [〇 ; 〇] : N_c = 2
 [〇 ; 〇] : N_c = 4

□ : CSPとROPは直列接続
 □ : CSPとROPは並列接続

図 13 ワークロード特性とリソース条件の関係 (2/3)
 Fig. 13 Relationship between workload and resources (2/3).

③ROP とワーク用 DK の必要台数を決定するにあたり、考慮すべきワークロード特性を重回帰分析により求めると、決定係数が 0.4 程度と低く、今回分析した範囲では支配的な説明変量は不定である。したがって、図 12～図 14 を参考に、負荷が相当に大きくなる場合は両者の増設を判断することになる。

6. おわりに

本論文では、データベースプロセッサ RINDA の適用領域と最適システム構成を明らかにすることを目的として行った、シミュレーションによる性能評価手

$\begin{matrix} R_w \\ R_s \\ A_T \\ \lambda_c \end{matrix}$		90 件/時 (R_{us}, R_{uc}, R_w は各30件/時)				
		1万件/時	2万件/時	3万件/時	4万件/時	5万件/時
25,600 ページ	1%	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$
	5%	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$
	10%	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$
51,200 ページ	1%	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$
	5%	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$
	10%	$\begin{matrix} [1; 2] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 2] \\ [384, 64] \end{matrix}$	$\begin{matrix} [1; 2] \\ [384, 64] \end{matrix}$	$\begin{matrix} [1; 2] \\ [512; 64] \end{matrix}$	$\begin{matrix} [1; 2] \\ [512; 64] \end{matrix}$
76,800 ページ	1%	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [256; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [384; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [512; 32] \end{matrix}$	$\begin{matrix} [1; 1] \\ [640; 32] \end{matrix}$
	5%	$\begin{matrix} [1; 1] \\ [640; 64] \end{matrix}$	$\begin{matrix} [1; 1] \\ [640; 64] \end{matrix}$	$\begin{matrix} [1; 1] \\ [640; 64] \end{matrix}$	$\begin{matrix} [1; 1] \\ [896; 64] \end{matrix}$	$\begin{matrix} [1; 1] \\ [640; 64] \end{matrix}$
	10%	$\begin{matrix} [2; 2] \\ [896; 128] \end{matrix}$	$\begin{matrix} [2; 2] \\ [896; 128] \end{matrix}$	$\begin{matrix} [2; 2] \\ [896; 128] \end{matrix}$	$\begin{matrix} [2; 2] \\ [1152; 160] \end{matrix}$	$\begin{matrix} [2; 2] \\ [1152; 160] \end{matrix}$

《評価条件》 $\lambda_r = 75\%$ 《凡例》 $[N_R; N_W]$, $\begin{matrix} \text{斜線} \\ \text{斜線} \end{matrix} : N_c = 2$
 $[A_U; B]$, $\begin{matrix} \text{斜線} \\ \text{斜線} \end{matrix} : N_c = 4$
 \square : CSPとROPは直列接続
 \blacksquare : CSPとROPは並列接続

図 14 ワークロード特性とリソース条件の関係 (3/3)
 Fig. 14 Relationship between workload and resources (3/3).

法について述べた。ワークロード特性とリソース条件および達成性能の関係を定量的に明らかにすることができたので、具体的なシステム条件に対して、各パラメータの決定が可能である。概略は以下のように要約できる。

(1) 性能に最も寄与するリソース条件は、CSPの台数であり、ROPの台数、ホスト計算機のバッファ容量等は副次的である。

(2) ある性能を達成するために必要な CSP の台数は、全検索範囲と非定型検索トラヒックの組み合わせで求まる。

(3) 単位検索範囲の大きさは、定型検索と非定型検索の性能に逆に作用するので、どちらを優先するのか、両者のバランスをみて決定する必要がある。

(4) ホスト計算機のバッファ容量は単位検索範囲の大きさと CSP 検索条件適合率の積を目安としてよく、それ以上大きくしても効果はない。

RINDA の開発にあたり、適用システムごとに異なるデータベース容量、検索トラヒック、検索条件適合率などの種々のワークロード特性に対して、最適なシステム構成を経済的に実現可能とするため、ビルディングブロック構成をとったが、その有効性を定量的に

実証できた。

謝辞 シミュレーションの実施とデータの分析に多大な協力をいただいた、NTT 情報通信研究所の高橋良英主任研究員、日本情報通信コンサルティング株式会社の稲葉哲也氏に深く感謝いたします。

参考文献

- 1) 喜連川優, 伏見信也: データベースマシン, 情報処理, Vol. 28, No. 1, pp. 56-67 (1987).
- 2) 清水 康: データベースマシンの動向, アドバンスト・データベース・システムシンポジウム論文集, pp. 31-40 (1987).
- 3) 拜原正人, 鈴木健司: データベースプロセッサの動向, NTT R & D, Vol. 38, No. 8, pp. 843-849 (1989).
- 4) Britton Lee, Inc.: The Intelligent Database Machine—Product Description (1984).
- 5) Teradata Corp.: DBC/1012 Data Base Computer System—Introduction (1986).
- 6) Babb, E.: Implementing a Relational Database by Means of Specialized Hardware, *ACM Trans. Database Syst.*, Vol. 4, No. 1, pp. 1-29 (1979).
- 7) Kojima, K., Torii, S. and Yoshizumi, S.: IDP—A Main Storage Based Vector Database Processor, *Database Machines and Knowledge Base Machines* (Kitsuregawa, M. and Tanaka, H. eds.), pp. 47-60, Kluwer Academic (1988).
- 8) 小柳津育郎, 塩川鎮雄, 木ノ内康夫, 安保 進: DIPS-11/5E シリーズの実用化, NTT 研究実用化報告, Vol. 36, No. 1, pp. 49-56 (1986).
- 9) 矢沢良一, 平野正則, 山口利和, 岡田靖史: DIPS-V 30 E のハードウェア構成, NTT 研究実用化報告, Vol. 37, No. 9, pp. 523-532 (1988).
- 10) 井上 潮, 速水治夫, 福岡秀樹, 鈴木健司, 松永俊雄: データベースプロセッサ RINDA の設計と実現, 情報処理学会論文誌, Vol. 31, No. 3, pp. 373-380 (1990).
- 11) 速水治夫, 佐藤哲司, 中村敏夫, 黒岩淳一, 武

田英昭: リレーショナルデータベースマシン RINDA におけるサーチ処理方式, 電子情報通信学会論文誌, Vol. J75-D-I, No. 4, pp. 232-240 (1992).

- 12) 武田英昭, 佐藤哲司, 中村敏夫, 速水治夫: 関係演算高速化プロセッサ, 情報処理学会論文誌, Vol. 31, No. 8, pp. 1230-1241 (1990).
- 13) 速水治夫: データベースプロセッサ RINDA におけるディスク装置アクセス競合制御方式, 情報処理学会論文誌, Vol. 33, No. 9, pp. 1123-1132 (1992).
- 14) 石原辰雄, 長谷川勝也, 川口輝久: Lotus 1-2-3 活用多変量解析, 共立出版株式会社, 東京(1990).
(平成4年11月13日受付)
(平成5年9月8日採録)



野瀬 純郎 (正会員)

1945年生。1967年神戸大学工学部電気工学科卒業。1969年同大学院工学研究科電気工学専攻修士課程修了。同年、日本電信電話公社入社。以来、論理回路の故障診断、大型情報処理システム、通信処理システム、システム評価方式の研究実用化に従事。現在、NTT ソフトウェア株式会社開発合理化推進本部担当部長。



速水 治夫 (正会員)

昭和22年生。昭和45年名古屋大学工学部応用物理学科卒業。昭和47年同大学院工学研究科応用物理学専攻修士課程修了。同年、日本電信電話公社入社。現在、NTT 情報通信研究所基本アーキテクチャ研究部主幹研究員。工学博士。DIPS ハードウェアシステム、データベースマシン、情報検索システムの研究実用化に従事。電子情報通信学会会員。