

OLTP のための並列コンピュータシステムにおける 機器並列度の応答性能に及ぼす効果

河野 健一郎[†] 亀田 壽夫^{††}

本論文はオンライントランザクション処理 (OLTP) のための並列コンピュータシステムの機器構成がシステムの性能に及ぼす効果を検討したものである。並列コンピュータによりオンライントランザクション処理システムを構成する場合、並列に動作する機器をどれだけ組み込むかがシステムの性能を左右する要因のひとつとなる。本論文ではこれによるシステムの平均応答時間をシミュレーションにより評価した。その結果、並列度を高めるように機器を増設しコストをかけたにもかかわらず、かえって性能が低下する場合さえありうることを観測した。これは機器を増設したことにより、機器を接続している相互結合網への転送要求量が増加し、その結果、相互結合網の転送遅延を増大させてしまったことがその原因と分かった。いうまでもなく並列コンピュータは多種多数の機器により構成される複雑なシステムであり、機器構成にも非常に多くの可能性がある。良い性能を無駄なコストをかけずに達成するために適切な機器構成を選択することが重要であることを確認した。

Effects on Responsiveness of the Degree of Parallelism in Parallel Computer Systems for On-Line Transaction Processing

KEN'ICHIRO KOHNO[†] and HISAO KAMEDA^{††}

This paper presents the effects of system configurations on performance of parallel computers for On-Line Transaction Processing (OLTP). In OLTP systems that consist of parallel computers, system performance will depend on the number of parallel computers, that is, the degree of machine parallelism installed in the system. We examined the effects of the degree of parallelism on the response time by simulation. We observed that higher degree of parallelism obtained by attaching additional computers does not usually lead to shorter responses. Furthermore, we also observed that higher degree of parallelism in database processors may sometimes give longer response time. We found that this is mainly because the transmission delay in the interconnection network increases due to the increase in the number of the requests issued by the database processors. Those results seem to emphasize the importance of choosing appropriate configurations to achieve good performance with small cost.

1. はじめに

オンライントランザクション処理 (On-Line Transaction Processing: 以下 OLTP と略する) は銀行の勘定システム、航空/鉄道の座席予約システムや企業における販売/在庫管理システムなどに代表されるように、社会の基幹システムであり日常生活を営むうえで必要不可欠なものである。OLTP は現在のコンピュータ応用全体の中の半分以上を占めるともいわれている。

OLTP システムには高い信頼性が要求されるので、システムの構成機器 (CPU やディスク等) は冗長化されてきたが、最近では冗長化を含めた並列化システムとして構成されるものがでてきた¹⁾。代表的なものに各メインフレームが相次いで発表した並列汎用機がある。これら並列コンピュータによる OLTP システム (以下、並列 OLTP システムと略する) では、トランザクション処理におけるある特定の処理を並列に実行するために、並列に動作する複数の機器を備えている。そしてこれらの機器は相互結合網で結合され、処理の実行要求や結果の返送などを転送し合う形態をとっている。

このような並列 OLTP システムにおいては、並列に動作する機器数とそれらを結合する相互結合網の種類、すなわちシステムの機器構成がその性能に影響を

[†] 筑波大学大学院工学研究科

Doctrial Program in Engineering, Graduate School of University of Tsukuba

^{††} 筑波大学電子・情報工学系

Institute of Information Sciences and Electronics, University of Tsukuba

与えると考えられる²⁾。本論文では、並列 OLTP システムのある機能を、並列に実行する機器の数とそれによる性能の差異という点に着目し、機器の並列度によりトランザクションの応答時間がどのように異なるかをシミュレーションにより評価した。その結果、機器を増加させて並列度を上げたにもかかわらず、性能が向上せず、かえって悪化させてしまう場合さえあることが分かった。

現在までに発表されている OLTP システムの性能問題に関する研究は、いくつかのアーキテクチャの比較⁸⁾、システムの入出力部のメモリ階層化による高速化と全体の性能向上の度合⁷⁾、データベースのバッファ管理の諸方式の性能比較⁹⁾、データの共有方式による比較¹⁰⁾、同時実行制御の改善による性能向上¹¹⁾など多岐にわたっている。本研究は上記の中でのアーキテクチャの比較に近いが、並列 OLTP システムの性能を機器の並列度という観点から評価した例は今のところ見当たらないのではないと思われる。

本論文の構成は次のとおりである。2章で並列 OLTP システムの機器の種類と本論文で対象とする機器構成モデル、およびワークロードなどのシミュレーションモデルについて述べる。3章でシミュレーション結果を示し、考察を行う。4章でまとめと今後の課題を述べる。

2. シミュレーションモデル

本章では 2.1 節で並列 OLTP システムを構成する機器の種類をその機能により分類する。2.2 節ではそれらの機器による機器構成モデルを、2.3 節では性能評価に用いるワークロードモデルを示す。2.4 節ではシミュレーション環境と実験に用いたパラメータの値を示す。

2.1 並列 OLTP システムの構成機器

並列 OLTP システムを構成する機器の種類は多種にわたると考えられる。我々は最新の並列汎用機のシステム構成などから、各システムに共通の機能と考えられる以下 4 種類の機器に分類し、それらの機器により並列 OLTP システムを構成するものとした。

- (1) 通信制御プロセッサ … 伝送制御手順による通信制御やデータの蓄積/振り分け処理を行うプロセッサ。本論文では CP (Communication Processors) と表記する。
- (2) アプリケーションプロセッサ … トランザクション処理のユーザアプリケーションを実行するプロセッサ。AP (Application Processors) と表記する。

- (3) データベースプロセッサ … データベースアクセスを実行するプロセッサ。データベースのレコードを保持するバッファやキャッシュのページを持っていることを仮定する。DP (Database Processors) と表記する。
- (4) データベース格納機器 … データベースを格納しているディスク群とそのコントローラ。DK (Disks) と表記する。

これら 4 つの機器によるトランザクション処理は次のように設定している。

- (1) WAN (Wide Area Network) などを通してトランザクションが CP に到着する。
- (2) CP で必要な処理を行った後、トランザクションは AP へ移される。AP でトランザクション処理のアプリケーションが実行される。
- (3) AP での実行中データベースの処理が必要になった場合、処理要求が DP に出される。DP は内蔵のキャッシュや DK にアクセスしてトランザクションが要求するレコードの読み書きを行う。
- (4) 処理がすべて終了したトランザクションは AP から CP へ移される。CP は必要な処理を行った後トランザクションの発生元へ応答を返す。

上に示した 4 種類の各機器において、並列に動作する 1 つ 1 つの機器をモジュールと呼ぶことにする。CP, AP, DP などのプロセッサ機器では 1 つのモジュールに 1 台以上の CPU が組み込まれる。DK では 1 つのコントローラで制御されるディスク群をモジュールとする。

2.2 機器構成モデル

前節で述べた各種類の機器を用いて構成した並列 OLTP システムを図 1 に示す。図中“Center”と記されている部分がシミュレーションの対象である。本モデルは AT&T システム 3600³⁾、および富士通の MISSION D/C¹⁾を参考にした。図 1 で CP-M は CP のモジュールを表す。同様に AP-M, DP-M, DK-M はそれぞれ AP, DP, DK のモジュールを表す。ICN は CP, AP, DP の各モジュールを結合する相互結合網である。

2.3 ワークロードモデル

ワークロードとしてのトランザクション処理は TPC-A ベンチマークを用いた。TPC-A の詳細については文献 4), 5) を参照されたい。このベンチマークは銀行の口座取引を単純化したもので、アクセス対象ファイル (データベース) は ACCOUNT (預金者ファイル), HISTORY (履歴ファイル), TELLER (窓口担当者ファイル), BRANCH (支店ファイル) の 4 つ

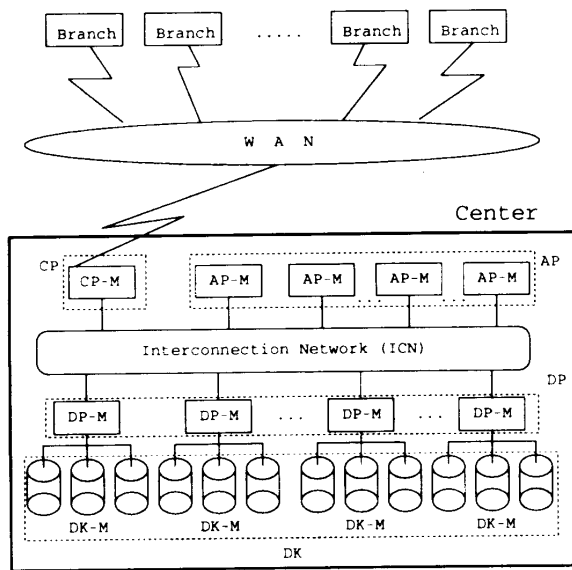


図1 機器構成モデル

Fig. 1 System configuration model.

である。ACCOUNT, BRANCH, TELLER ファイルに対しては更新 (UPDATE) を行い, HISTORY ファイルには追加 (INSERT) を行う。我々はさらにコミット時の LOG ファイルを追加している。

トランザクション処理のシステム内での流れは TMF における方式⁴⁾をもとに, 以下のように設定している。

- (1) トランザクションは図 1 における各支店 (Branch) で発生する。TPC-A の各ファイルは全 DK-M にわたって水平分割される。各 DP-M はこれら DK-M のいくつかを重複することなく制御, 管理する。発生したトランザクションは CP-M で所定の処理を受ける。なお本シミュレーションでは Center と各支店間の通信時間は考慮していない。
- (2) CP-M で処理を受けたトランザクションはランダムに AP-M が割り当てられ, ICN を経由してその AP-M へ転送される。
- (3) AP-M で TPC-A プログラムが実行される。我々はどの DP-M が各ファイルのどのレコードを保持しているかという情報 (索引) を AP-M は持っていないと仮定している^{*}。このため AP-M は ICN を経由した一斉通報により全 DP-M に処理要求を出す。まず新規のトランザクシ

^{*} AP-M が索引を持っていないとしたのは次の理由による。索引にはいくつかの方式¹³⁾があり, どの方式をとるか, およびそれに対するパラメータの設定でシステムの性能に影響を及ぼすことが考えられる。本論文では索引の方式によらない OLTP 用の並列アーキテクチャの性能評価として, AP-M に索引を持たせず AP-M の要求に対して全 DP-M が並列に動作することとした。

ンの到着に対して AP-M から全 DP-M にトランザクションの処理開始が ICN を経由して一斉通報される。AP-M は全 DP-M からトランザクション開始承諾の応答を待つ。

- (4) TPC-A プログラム実行中のファイル処理要求 (UPDATE や INSERT) は ICN を経由して全 DP-M に渡され, 各 DP-M は当該 AP-M の要求レコードが自分の制御下にあるかどうかを調べる。要求レコードの存在する DP-M はファイルアクセスと必要な処理を行う。アクセス時は 2 フェーズロックを行う。処理結果は ICN を経由して該当する AP-M へ返される。
- (5) コミット時には AP-M が全 DP-M にトランザクションがコミット処理に入ることを ICN を経由して一斉通報する。このとき 2 フェーズコミットが行われる。ファイル処理を行った DP-M はデータを LOG ファイルに書き込みロックを解除する。
- (6) コミット終了後, 処理結果が ICN を経由して CP-M へ転送され, 所定の処理を受けた後に各支店へ出力される。

2.4 シミュレータと実験環境

シミュレータは C 言語とシミュレーションライブラリ *smpl*⁶⁾ で開発した。開発規模は 13 モジュールで約 6300 ステップである。シミュレーションの実行は SUN4 SPARC Station 上で行った。シミュレーションに用いた各種のパラメータは次のとおりである。

- (1) モジュール…各機器のモジュールには以下のような性能値を設定している。

- CP-M…50 MIPS のプロセッサを 4 本搭載
- AP-M…20 MIPS のプロセッサを 4 本搭載
- DP-M…40 MIPS のプロセッサを 4 本搭載
- DK-M…平均アクセスタイム 10.5 ms の IBM 社 0061 3.5 インチ SCSI ディスクを仮定

上記のプロセッサの MIPS 値に関しては参考文献 1) で述べられている並列トランザクションサーバや並列データベースサーバの値を参考にした。

- (2) 相互結合網…CP, AP, DP の各モジュールを結合する相互結合網としてトークンリング型の結合網を用いた。これは富士通 MISSION D/C¹⁾, Gamma データベース (バージョン 1)¹²⁾ など, OLTP システムや並列データベースシステムに使用されており一般的な結合網のひとつと考えられる。プロトコルは Timed Token Protocol

で TTRT (Target Token Rotation Time) は 8 ms に設定した. チャネル数は 1 でバンド幅は 10 Mbps と 100 Mbps の双方で実験した.

- (3) ダイナミックステップ…Rahm⁷⁾の Debit-Credit のシミュレーションパラメータをもとに以下のように設定した (単位: キロステップ). いずれも幾何分布に従うとする.

- CP-M…プロトコル変換等通信処理 20
- AP-M…トランザクション開始処理 70, 通信処理 35, ファイルアクセスの間隔 5, コミット処理 14
- DP-M…トランザクション開始処理 20, 通信処理 35, 1 回あたりのファイル処理 25, コミット処理 60

- (4) ファイル

- レコード数…ACCOUNT 50,000,000 件, TELLER 5,000 件, BRANCH 500 件
- DK への配置…全 256 台のディスクドライブに分散配置
- ページ数…2000 ページ/DP-M
- ページサイズ…4 KB
- ページ内レコード数 … ACCOUNT, TELLER, BRANCH ファイル 10, HISTORY ファイル 20
- 同時実行制御…レコード単位
- アクセスパターン…85-15 ルール⁵⁾に従う.

3. 実験結果

実験結果を以下に示す. 結果の図においては, シミュレーションにより計測されたトランザクションの各到着率に対する平均応答時間は記号 (□ や ○, -) で, 平均応答時間の 95% の信頼区間の幅は平均応答時間の記号を中心とした上下の記号 (T と ⊥) の幅で表示してある.

なお, ここで示す結果は 2.4 節で述べたように相互結合網としてトークンリングを採用した場合のものであるが, 類似の現象が他の相互結合網 (1-Persistent CSMA/CD プロトコルを用いたバス型の結合網) を採用した場合でも観測されている.

3.1 AP のモジュール数と平均応答時間

図 2 と図 3 に AP のモジュール数を 1, 2, 4, 8, 16 と増加させていった場合のトランザクション平均応答時間をトランザクションの到着率 (以下, 記号 λ で表す. 単位はトランザクション/秒) ごとに示す. 図 2 は相互結合網のバンド幅が 10 Mbps, 図 3 では 100 Mbps の場合である. DP のモジュール数はいずれ

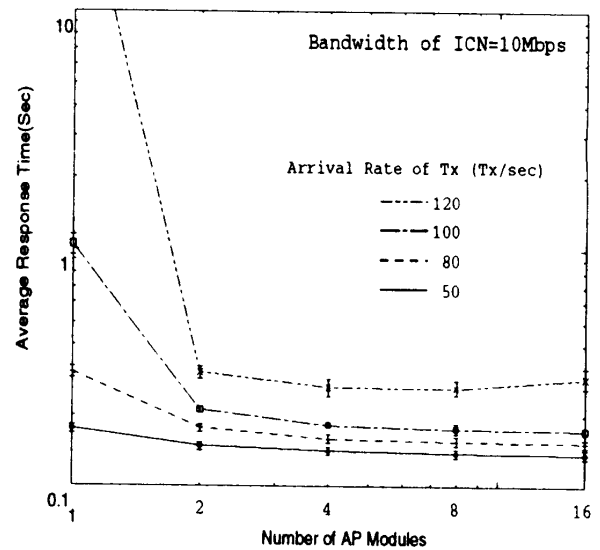


図 2 AP のモジュール数とトランザクションの平均応答時間 (1)
Fig. 2 Number of AP modules vs. average response time of transactions (1).

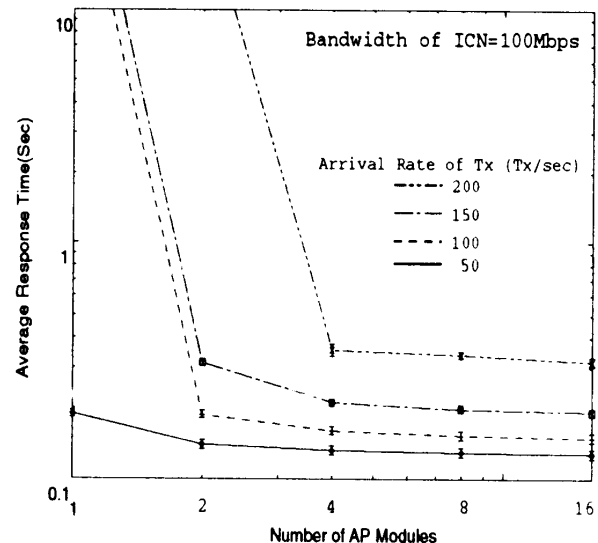


図 3 AP のモジュール数とトランザクションの平均応答時間 (2)
Fig. 3 Number of AP modules vs. average response time of transactions (2).

の実験も 8 に固定した. なお, 図 2 において λ が 120 を超えると相互結合網での転送待ちが非常に大きくなり, 実行環境のメモリ容量の関係で当該部分のシミュレーションが実行できなくなるので計測を打ち切っている.

図 2 よりトランザクションの各到着率に対して AP のモジュール数をある数以上に増やしても, すなわち AP の機器並列度を上げて平均応答時間を向上させることにはなっていないことが分かる. $\lambda = 50$ の場合, どのモジュール数でもほとんど差はなく, それ以上の到着率ではモジュール数を 1 から 2 に増やすことで応答性は改善されるが, 2 以上ではほとんど差がな

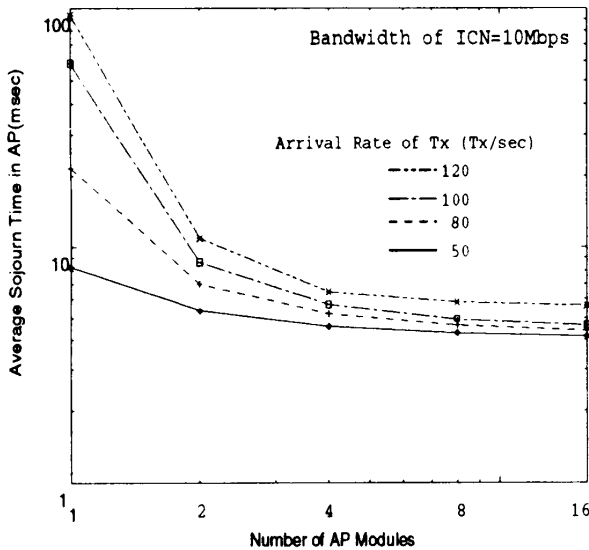


図4 APでの平均滞在時間(1)

Fig. 4 Average sojourn time in AP (1).

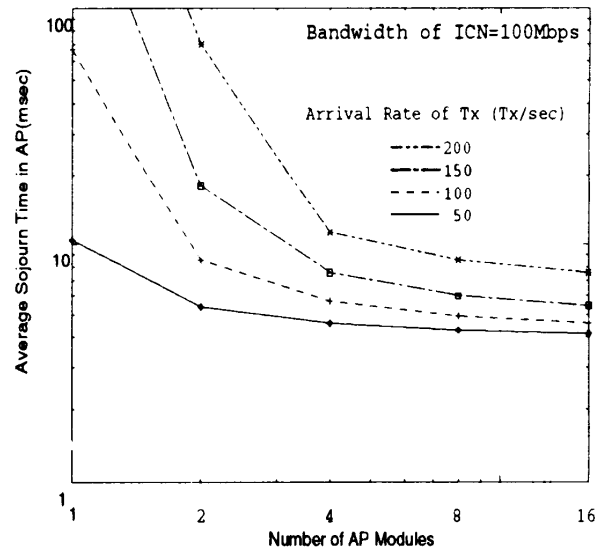


図5 APでの平均滞在時間(2)

Fig. 5 Average sojourn time in AP (2).

い。同様に図3では $\lambda = 50$ だとどのモジュール数でも応答性に差はなく、 $\lambda = 100$ の場合はモジュール数を1から2に増やすことで応答性は大きく改善されるが、2以上では差がなくなる。 λ が100を超えるとモジュール数が4までは応答性の改善が見られるが、それ以上のモジュール数では応答性に差がでない。

この原因を考えるためトランザクションのAPのモジュールでの平均滞在時間を調べてみる。図4(相互結合網のバンド幅10Mbps)、図5(同100Mbps)にこれを示す。図4、図5から分かるように相互結合網のバンド幅が10Mbpsの場合はAPモジュール数2、相互結合網のバンド幅が100Mbpsの場合は λ が100以下ではモジュール数2から、それ以上の到着率ではモジュール数4からAPでの平均滞在時間に大きな差がでなくなっている。これらAPのモジュール数のときのトランザクションに平均応答時間に占めるAPの滞在時間の割合をみると、バンド幅が10Mbpsの場合APのモジュール数が2以上では30%前後である。バンド幅が100Mbpsの場合、 λ が100以下では、APのモジュール数が1の場合には45%以上と高い割合を占めるが、2以上では25%前後となる。それ以上の到着率ではモジュール数が2までだと45%以上であるが、4以上では25%以下となる。このようにAPのモジュール数を増やしていくと、ある数以上ではトランザクション処理に対するAPの能力が十分になり、APのモジュール数の増加が応答性の向上に結び付かなくなると考えられる。

APのモジュールの能力を下げてどうなるかを図6に示す。相互結合網のバンド幅は100Mbpsで、AP

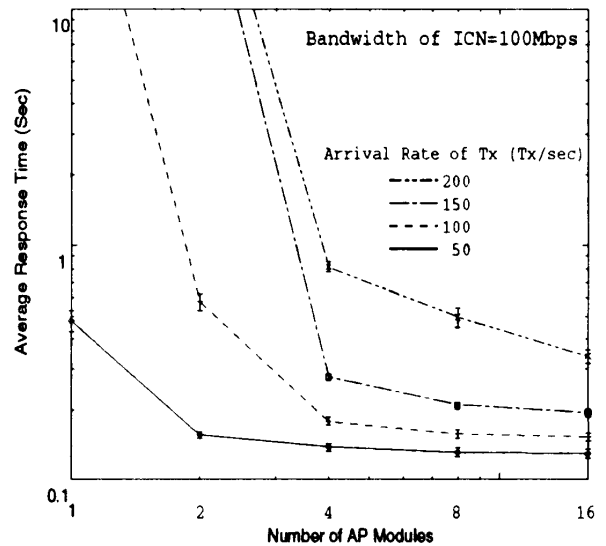


図6 APのモジュール数とトランザクションの平均応答時間 (APのプロセッサ数=2)

Fig. 6 Number of AP modules vs. average response time of transactions (number of processors in each AP module=2).

の各モジュールに組み込むプロセッサ数を4から2に減らしている。この場合、 λ が50ではモジュール数2、 $\lambda = 100, 150$ ではモジュール数4、 $\lambda = 200$ ではモジュール数を8まで増加させると応答性を向上させるが、それ以上では改善がみられなくなる。

これも各到着率に対して前述したモジュール数以下の場合、トランザクションの平均応答時間に占めるAPでの平均滞在時間の割合が40%以上となり、モジュール数を増加することによりこの割合を下げるができる。しかし前述したモジュール数以上では30%以

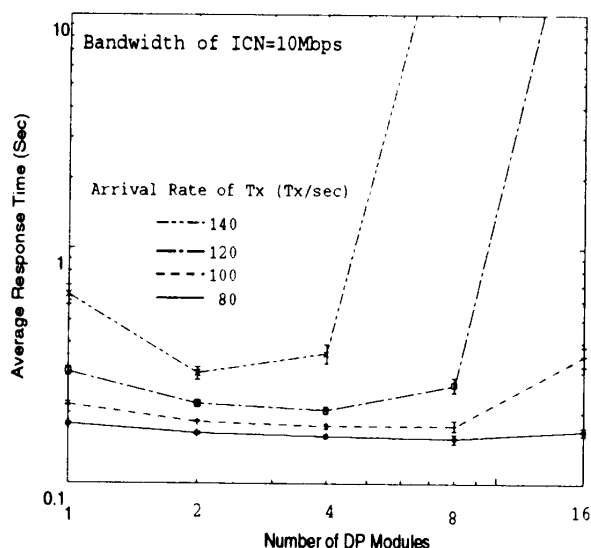


図7 DPのモジュール数とトランザクションの平均応答時間(1)
Fig. 7 Number of DP modules vs. average response time (1).

下となり、応答性に影響を与えなくなる。このようにAPの各モジュールに組み込むプロセッサ数を減らす(=能力を下げる)ことで応答性を確保するのに必要なモジュール数は増加するが、それ以上のモジュール数、すなわちより高い機器の並列度が応答性の改善に寄与しないという点は同じである。

3.2 DPのモジュール数とトランザクションの平均応答時間

図7と図8にDPのモジュール数を1, 2, 4, 8, 16と増加させていった場合のトランザクション平均応答時間を示す。図7は相互結合網のバンド幅が10 Mbps, 図8では100 Mbpsの場合である。APのモジュール数はいずれの実験も8に固定した。なお、図7において λ が140を超えるものについては前節と同様の理由で計測を打ち切っている。

この結果において注目すべきことは、図7の相互結合網のバンド幅が10 Mbpsの場合、DPのモジュール数に適正值(たとえば $\lambda = 120$ の場合には4, $\lambda = 140$ の場合には2)がありそれよりもモジュール数を増やすと逆に応答性を悪化させてしまっている点である。すなわちデータベース処理の並列度がある数以上に上げることがかえって性能の劣化を招くというものである。一方、図8の相互結合網のバンド幅が100 Mbpsの場合においては前節でみたのと同様の現象があり、DPのモジュール数が各到着率に対してある数以上になると応答性能に差がでなくなっている。

図7におけるDPのモジュール数の増加による性能劣化の原因を調べるため、各DPモジュール数における各トランザクションの到着率に対する相互結合網

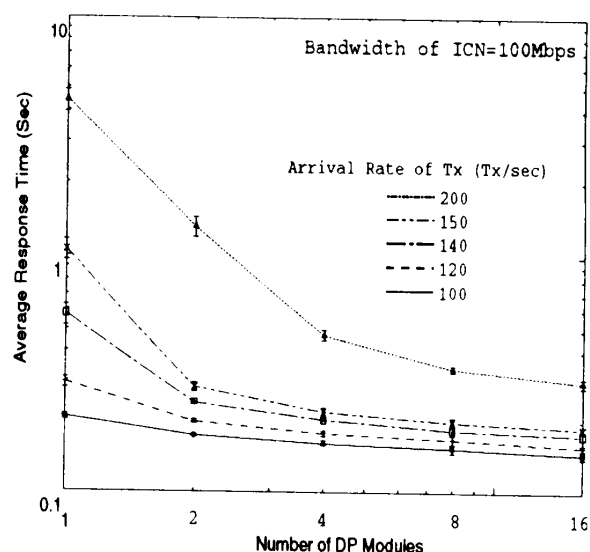


図8 DPのモジュール数とトランザクションの平均応答時間(2)
Fig. 8 Number of DP modules vs. average response time (2).

への転送要求率と転送遅延を調べてみる。これは仮定しているトランザクション処理においてAPの各モジュールは相互結合網に対して一斉通報の packets を送信して全DPモジュールを駆動するが、一部の処理は全DPのモジュールが当該APに対し応答を返すようになっているため、トランザクションの到着率は同じでもDPのモジュール数が増加することで相互結合網への packets 転送要求が増加して転送遅延を増大させてしまうことが考えられるからである。図9にDPのモジュール数と packets の転送要求率、図10にDPのモジュール数と packets の転送遅延時間の関係を示す。両図よりトランザクションの到着率があってもDPのモジュール数を増やすことで packets の転送要求率が増加し、また転送遅延が増大していくことが分かる。

一方、相互結合網のバンド幅が100 Mbpsの場合のDPのモジュール数と packets の転送要求率、および packets の転送遅延時間の関係を示したのが図11と図12である。両図より相互結合網のバンド幅が100 Mbpsの場合、DPのモジュール数を増やすことによる packets の転送要求率の増加に対して転送遅延の大幅な増大はみられないことが分かる。

図13には相互結合網のバンド幅が10 Mbpsで $\lambda = 120$ のとき、DPのモジュール数を増加していった場合の、各トランザクションあたりのAP, CP, DKの各機器における平均滞在時間、DPにおけるCPUにかかわる処理時間(待ち時間を含む)の平均(図中ではDP(1))、DPにおけるディスク処理にかかわる時間(待ち時間を含む)の平均(図中ではDP(2))、および

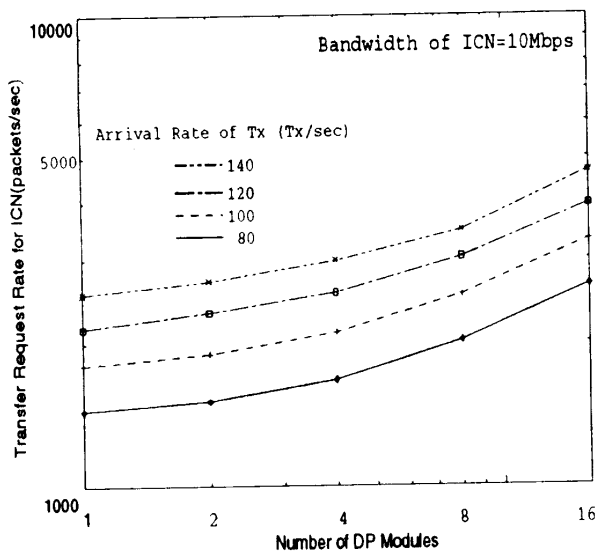


図9 DP のモジュール数とパケット転送要求率 (1)
Fig. 9 Number of DP modules vs. transfer request rate of packets (1).

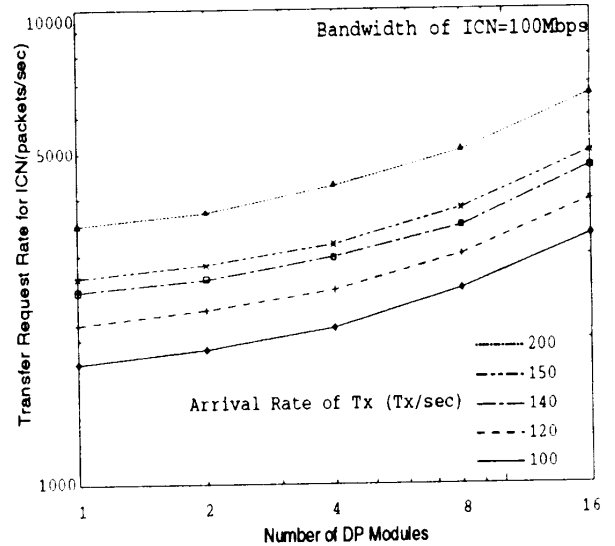


図11 DP のモジュール数とパケット転送要求率 (2)
Fig. 11 Number of DP modules vs. transfer request rate of packets (2).

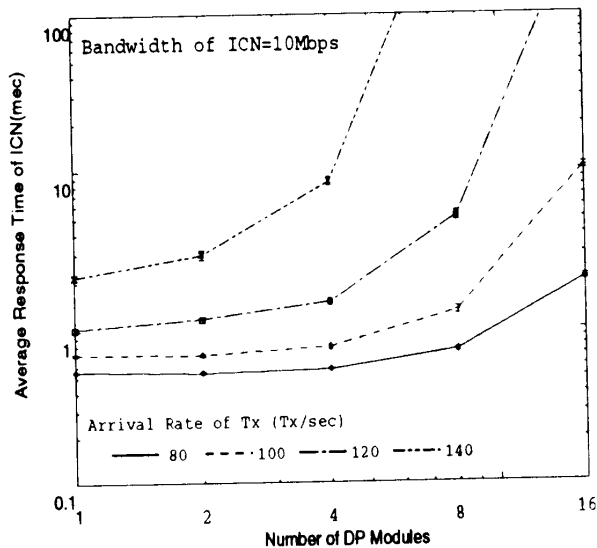


図10 DP のモジュール数とパケット転送遅延時間 (1)
Fig. 10 Number of DP modules vs. communication delay time of packets (1).

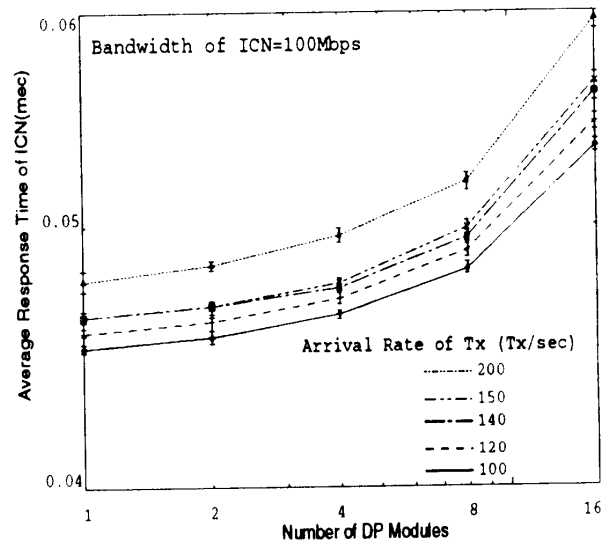


図12 DP のモジュール数とパケット転送遅延時間 (2)
Fig. 12 Number of DP modules vs. communication delay time of packets (2).

パケットの平均転送遅延時間を示す。これらの総和がトランザクションの平均応答時間になり、同図にはこれもあわせて示してある。図13よりDPのモジュール数の増加にともないパケットの転送遅延とDPにおけるディスク処理にかかわる時間が増大している。これはパケットの転送遅延に急激な増大があると、特に同時実行制御によるオーバーヘッド(レコードにロックをかけている時間)も大きくなり、データベースに関する処理能力が性能上の大きな隘路になるためと考えられる。パケットの転送遅延と同時実行制御によるオーバーヘッドの関係の詳細については別途検討したい。

3.3 考 察

前2節より得られた結果をまとめると、

- 相互結合網のバンド幅が10Mbps/100Mbpsにかかわらず、APのモジュール数を増やしていくとある数以上では応答性が改善されなくなる、
- 相互結合網のバンド幅が10Mbpsの場合はDPのモジュール数に適正值があり、それ以上増やすと逆に応答性を悪化させてしまう。100Mbpsではある数以上にDPのモジュール数を増やしても応答性が改善されなくなる、

となる。ここでAPの並列化とDPの並列化の違いを検討してみる。APにおいては1つ1つの処理要求

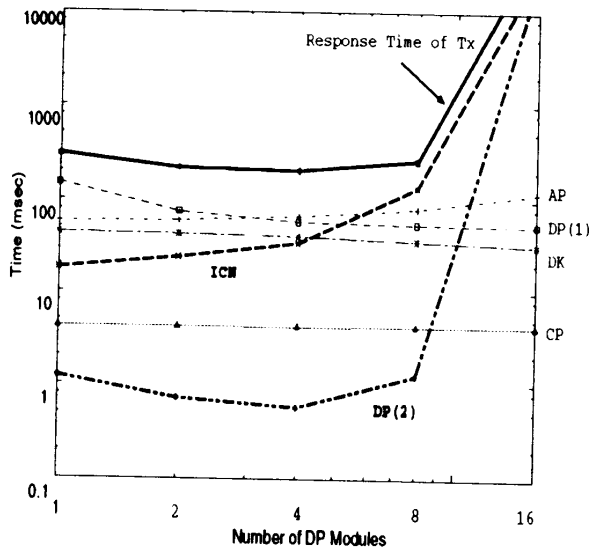


図 13 DP のモジュール数と各機器における処理時間の平均、および ICM の平均転送遅延時間

Fig. 13 Number of DP modules vs. average processing time in each machine and delay of ICM.

(ユーザアプリケーションの実行要求)は各モジュールで独立に処理できる。1つの処理自体は比較的負荷が軽く、並列性を持つわけではないので1つの要求の処理に対応するモジュール数はただか1である。同じトランザクションの到着率に対して AP のモジュール数を増加させることは各モジュールの負荷を減少させることになるが、モジュール数がある数以上になると十分な処理能力となるので応答時間に影響を与えなくなる。また AP の各モジュールでの処理の性格上、モジュール数を増やしても相互結合網への転送要求量は変化しない。

一方 DP は 1つの処理 (データベース処理) が大きな負荷を持つため各モジュールに処理を分割して並列に実行させる。1つの処理要求が DP の複数のモジュールに分割されているため処理結果や後続の処理に必要な応答を複数のモジュールが返すことになる。この場合、DP のモジュール数が増すと各モジュールの負荷は減少するが、処理結果や応答を返すための転送要求量が増加することになり相互結合網への負荷を増大させることになる。その結果、相互結合網が性能上の隘路になり、またその結果同時実行制御によるオーバーヘッドがさらに増大し、応答性を悪化させてしまうことがありうる。また相互結合網が性能上の隘路にならない場合には、AP を増やす場合と同様に、ある数以上の DP のモジュールを設置すると処理能力が十分になり、応答性の改善に寄与しなくなると考えられる。

モジュール数を増やす、すなわち機器の並列度を上げることはコストをかけることになるが、以上みたよ

うにそれに見合うだけの性能向上が得られるとは限らない。かえって悪化させてしまう場合もありうる。機器の並列度の設定は、各機器の持つ並列性の違いや各機器間の通信方式、相互結合網も含めた性能のバランスを十分に考慮することが重要と考えられる。

4. おわりに

本論文では、並列コンピュータにより OLTP システムを構成する場合においてある機能を並列に実行する機器の数、すなわち機器並列度を変えることによりその応答性能がどのように変化するかを、実システムを参考にしたモデルでシミュレーションによって評価した。その結果、機器並列度を上げることが必ずしも応答性の向上に寄与しないことや逆に応答性を悪化させてしまうことさえあることを示した。

並列コンピュータは多種多数の機器から構成される複雑なシステムであるので、その機器構成の選択肢は数え切れないほど多い。その中から適切な機器構成を選び出すことが良い性能を無駄なく実現させるために重要であることがあらためて確認されたと考えられる。

以上、本論文ではシステムの性能指標としてトランザクションの平均応答時間を用いた。しかしスループットを単位時間あたり処理できる最大のトランザクション数と定義すると、スループットも重要な性能指標と考えられる。スループットはトランザクションの到着率を少しずつ変えて実験して求めることになり、膨大なシミュレーション時間を要するので本論文では割愛した。別途検討したい。

今後の課題として、

- 並列度増加に対して性能劣化が生じにくいデータベース処理方式の検討、
 - 相互結合網の性能と同時実行制御によるオーバーヘッドの関係の評価、
 - 他の機器構成モデルや TPC-C などのワークロードでの評価、
 - スループットなど、他の性能指標による評価、
- などを行う予定である。

謝辞 本研究に関してデータベースの技術面でご指導いただいた(株)システムクリエートセンターの関係各位に深く感謝いたします。また有益な情報、助言をいただいた本学電子情報工学系の李頡講師、および亀田研究室の各位にも深く感謝いたします。

参考文献

- 1) 日経ウォッチャー：並列汎用機の全貌，日経 BP 社 (1994)。

- 2) 河野健一郎, 亀田壽夫: 並列コンピュータによるオンライントランザクション処理システムの機器構成の性能評価, 並列処理シンポジウム *JSP'95*, pp.145-152 (1995).
- 3) 日本 NCR: NCR3600 概説書 (1992).
- 4) Gray, J.: *On-Line Transaction Processing Systems*, McGrawHill (1994). 渡辺栄一 (訳): OLTP システム (増補版), マグロウヒル (1994).
- 5) Gray, J.: *The Benchmark Handbook for Database and Transaction Processing Systems*, Morgan Kaufmann (1991). 喜連川優, 渡辺栄一 (訳): データベースベンチマーキング, 日経 BP 社 (1992).
- 6) MacDougall, M.H.: *Simulating Computer Systems: Techniques and Tools*, MIT Press (1987). 小林誠 (訳): シミュレーションによるコンピュータシステムの性能評価: テクニックとツール, 工学社 (1990).
- 7) Rahm, E.: Performance Evaluation of Extended Storage Architecture for Transaction Processing, *ACM SIGMOD*, pp.308-317 (1992).
- 8) Yu, P.S. and Dan, A.: Performance Evaluation of Transaction Processing Coupling Architectures for Handling System Dynamics, *IEEE Trans. Parallel and Distributed Systems*, Vol.5, No.2, pp.139-153 (1994).
- 9) Dan, A. and Yu, P.S.: Performance Analysis of Buffer Coherency Policies in a Multi-system Data Sharing Environment, *IEEE Trans. Parallel and Distributed Systems*, Vol.4, No.3, pp.289-305 (1993).
- 10) Wolf, J.L., Dias, D.M., Iyer, B.R. and Yu, P.S.: Multi-system Coupling by a Combination of Data Sharing and Data Partitioning, *IEEE Trans. Softw. Eng.*, Vol.15, No.7, pp.854-860 (1989).
- 11) Thomasian, A. and Rahm, E.: A New Distributed Optimistic Concurrency Control Method and A Comparison of Its Performance with Two-Phase Locking, *Proc. 10th Distributed Computing Conference*, pp.294-301 (1990).
- 12) DeWitt, D.J., Ghandeharizadeh, S., Schneider, D.A., Bricker, A., Hsiao, H. and Rasmussen, R.: The Gamma Database Machine Project, *IEEE Trans. Knowledge and Data Engineering*, Vol.2, No.1, pp.44-61 (1990).
- 13) Gray, J. and Reuter, A.: *Transaction Processing: Concepts and Techniques*, Morgan Kaufmann Publisher (1994).

(平成 7 年 9 月 4 日受付)

(平成 8 年 3 月 12 日採録)

河野健一郎

1961 年生。1984 同志社大学工学部電気科卒業。1984~89 年 (株) 日立ソフトウェアエンジニアリングに勤務。オペレーティングシステムの設計・開発に従事。1992 年電気通信大学博士前期課程修了 (情報工学)。1993 年より筑波大学工学研究科博士後期課程に在学。主な研究テーマは並列/分散システムの性能評価, オペレーティングシステムやデータベースシステムの設計方式, 相互結合網や通信プロトコルの評価など。

亀田 壽夫 (正会員)

1942 年生。1965 年東京大学理学部物理学科卒業。1970 年同大学院理学系研究科博士課程修了。理学博士。同年東京大学理学部助手。1971 年電気通信大学講師。同大学助教授・教授を経て、1992 年より筑波大学電子・情報工学系教授。この間、1973~74 年 IBM ワトソン研究所、1974~75 年トロント大学において研究に従事。オペレーティングシステム, 分散/並列処理, システム性能評価, 等の研究教育を行ってきた。本会編集委員, OS 研究会主査等を歴任。現在, 本会システム評価研究グループ主査, ACM, 電子情報通信学会, ソフトウェア科学会, 日本 OR 学会, 日本応用数理学会各会員。