

選択的多重ピギーバックを用いたジョブ配送法について

7U-1

染葉佳代子 渡辺尚 太田剛 水野忠則

静岡大学 工学部

1 はじめに

個々に資源を有する計算機をネットワークで接続した分散計算機システムについての主要な論点の1つは、負荷分散とシステム性能との関係である。これまでに、負荷分散方式について多くの研究が行われている [1]。ジョブの配送先を決定するディスパッチャの形態は、1つのディスパッチャですべてのジョブを扱う集中型と、各計算機ごとにディスパッチャを持つ分散型に分けられる。本研究では、ユーザの要求への柔軟な対処、大きなネットワークへの対応といった点から分散型ディスパッチャについて考える [2]。

我々はこれまでに、双方向ピギーバックを用いて環境観測を行い予想応答時間が最小であるサーバをジョブの配送先に決定する LR-PB 方式を提案し、不均質なシステムでは閾値制御を行うランダム方式やサイクリック方式より有効であることを示している [3][4]。本研究では、LR-PB 方式を拡張し、情報を選択してピギーバックする LR-SMP 方式を提案する。そして、その有効性を検討する。

2 モデル

対象とするシステムを次のようにモデル化する。

- システムは n 個のディスパッチャとサーバの組から構成される。サーバ S_i をディスパッチャ D_i のローカルサーバ、 S_j を $D_i (i \neq j)$ のリモートサーバと呼ぶ。
- S_i の処理速度を c_i 命令/秒、 $\vec{C} = (c_1, c_2, \dots, c_n)$ とする。
- ディスパッチャに要する時間は 0 秒とする。
- D_i はすべての S_j と全二重回線で接続されている。
- 通信遅延はローカル間を 0 秒、リモート間を一定値 d 秒とする。
- ディスパッチャはローカルサーバの負荷状況が瞬時にわかる。
- D_i にはジョブが平均 λ_i 個/秒、 $\vec{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_n)$ でポアソン到着する。
- ジョブの処理要求量は平均 $\frac{1}{\mu}$ 命令の指数分布に従う。
- サーバに割り当てられたジョブはそのサーバごとに待ち行列を形成し、先着順に処理される。

3 ジョブ配送アルゴリズム

分散型ディスパッチャを用いたソースイニシアチブなジョブ配送では、次の 2 点について検討する必要がある。

On a load balancing method with selective multi piggybacking
 Kayoko Someha, Takashi Watanabe,
 Tsuyoshi Ohta, Tadanori Mizuno,
 Faculty of engineering, Shizuoka University
 3-5-1, Johoku, Hamamatsu, 432 Japan

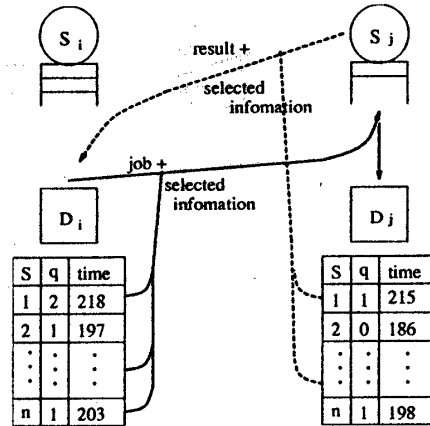


図 1: 選択的多重ピギーバック

- システム内の環境 (各サーバの負荷状況) の予測方法
- ジョブの配送先の決定方法

本研究ではそれぞれの点に対して、次に説明する選択的多重ピギーバック方式と最小応答時間方式を用いる。

3.1 選択的多重ピギーバック (SMP) 方式

システム環境を予測するための情報の収集方法として、サーバが自分の情報をブロードキャストするという方法が考えられる。しかし、この方法にはネットワークトラフィックが増大するという欠点がある。また、ジョブの特性を考慮した配送へ拡張することを考え、実行して始めてわかるジョブ量などの情報が得られるようにしたい。そこで本研究では、以下に説明する選択的多重ピギーバック (図 1) を用いる。

本研究では、直接ピギーバックと間接ピギーバックにより情報伝達を行う。直接ピギーバックは、 S_j が処理し終わったジョブの結果にローカルな D_j が持っている情報を付けて D_i に返送する方法である。一方、間接ピギーバックは D_i がリモート S_j へジョブを配送する時、持っている情報をジョブに付けて配送し、 S_j に届いた情報をそのローカルな D_j が得るという方法である。

我々が以前に提案した LR-PB 方式において、ピギーバックする情報はローカルサーバの情報のみである。そのため、ジョブのやり取りのないサーバの情報得られないという欠点があった。本稿で提案する SMP (Selective Multi Piggybacking) 方式では、ピギーバックによって得たリモートサーバの負荷情報 (q) とその時刻 ($time$) を記録しておく。ピギーバックで他のディスパッチャに情報を送る場合、自分のサーバの情報だけでなく記録した情報も合わせて送る。これを受け取ったディスパッチャは、自分の保持している情報と比較して、より新しいも

	通信遅延	サーバ数	ジョブ到着率
ホップ数	小	中	小
経過時間	大	大	大

表 1: 情報の選択方式とシステム環境の依存度

のであれば情報の更新を行う。このとき各サーバが、知り得たすべての情報をピギーバックしたのでは、ネットワークトラフィックを増大させるだけである。そこで、有効と思われる情報を選択してピギーバックする必要がある。

本研究では、情報の選択方式として、サーバから情報が発信されてから何台のディスパッチャを経由して届いたのか(ホップ数: 0は自分の情報のみを選択)、サーバから情報が発信されてから何秒経過して届いたのか(経過時間)の2つを検討した。

まず、知り得たすべての情報をピギーバックし、受け取ったディスパッチャにとって有益な情報の、ホップ数と経過時間の傾向を調査した。有益な情報の95%が含まれるようなホップ数の範囲を検討した結果、表1のような依存関係が得られた。ネットの混み具合が大きく変化し通信遅延のばらつきが大きいシステムや、ジョブ到着率が大きく変化するシステムでは、情報が有効な経過時間はシステム状態によって変化する。これに対して情報が有効なホップ数は、システム内のサーバ数に依存するものの、その他の環境にはほとんど影響されない。しかも、通信遅延やジョブ到着率に比べサーバ数はほぼ固定であるため、選択する情報の条件としてホップ数を用いるのが適当であるといえる。

3.2 最小応答時間(LR)方式

ジョブの配送先は次のように決定する。 D_i にジョブが到着したら、 D_i はピギーバックで得た環境情報に基づいて、各サーバへジョブを配送した場合の予想応答時間 R_j , $j = 1, 2, \dots, n$, を求め、それが最小のサーバをジョブの配送先に決定する。予想応答時間は次式で求められる。

$$R_j = \frac{(q_{ij} + 1) * \mu}{c_j} + 2 * d$$

ここで、 q_{ij} は D_i がピギーバックによって得た情報に自分が配送しているジョブを考慮した S_j の予想内容数である。 q_{ii} は S_i がローカルサーバであるため常に正確な値である。 R が最小のサーバが複数ある場合は、その中にローカルサーバが含まれていればローカルサーバへ、含まれていなければ R となる任意のサーバへ配送する。

4 シミュレーション

シミュレーションにより平均応答時間を測定し、LR-SMP方式の性能を評価する。システムは、 $N = 10$, $c_i = 1$, $d = 0$, $\lambda_i = \lambda$, $\mu = 1$ と仮定する。

このシステムにおいて、ホップ数1以下の情報に限定したLR-SMP方式が最適であった。これは、ローカルサーバの情報と直接相手から受け取った情報のみをピギーバックするものである。

次に、このLR-SMP方式を他の方式と比較する。比較するのは以下の方式である。

- ・ 経由回数1以下としたLR-SMP方式
- ・ ローカルサーバの情報のみをピギーバックするLR-PB方式

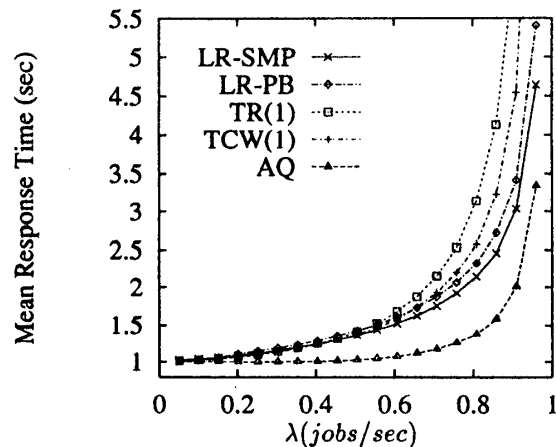


図 2: 各方式の比較

- ・ 閾値制御を行うランダム方式(TR方式)
- ・ 閾値制御を行うサイクリック方式(TCW方式)
- ・ 正確な環境情報を用いるAQ方式

シミュレーション結果を図2に示す。これから、特に中・高負荷領域ではLR-SMP方式が他の方式よりも良い性能であることがわかる。

5 結論

本研究では、分散計算機システムにおける負荷分散方式として、選択的多重ピギーバックを用いたLR-SMP方式を提案し、その性能について検討した。ピギーバックする情報の選択方式としては情報が有効なホップ数を設定する方法が適していることがわかった。また、すでに提案されている他の方式と比較した結果、SMP方式の有効性も確認できた。本研究では、サーバの負荷情報をジョブや処理結果に付加させて配送するため、情報のやり取りによるネットワークトラフィックの増加はほとんどないものと考えている。しかし、情報が有効なホップ数はシステム内のサーバ数により変化するもので、サーバ数が多い場合には、この点について検討する必要がある。

参考文献

- [1] Niranjana G. Shivaratri, Phillip Krueger and Mukesh Singhal, "Load Distributing for Locally Distributed Systems," IEEE Computer, Dec 1992.
- [2] 染葉佳代子, 渡辺尚, 太田剛, 水野忠則, "分散型ディスパッチャを用いたジョブ配送法について," 情報処理学会第48回全国大会 7D-3, 1994.
- [3] 染葉佳代子, 渡辺尚, 太田剛, 水野忠則, "双方向ピギーバックを用いたジョブ配送法について," 情報処理学会研究報告 Vol.94, No.56.
- [4] T. Ohta, T. Watanabe and T. Mizuno, "A Job Dependent Dispatching Scheme in a Heterogeneous Multiserver Network," IEICE TRANS. COMMUN., VOL.E77-B, NO.11 NOVEMBER 1994.