

分散システムでの負荷分散に向けたシステム状態の把握法

40-8

青木 義則

谷口 秀夫

牛島 和夫

九州大学大学院システム情報科学研究科

1 はじめに

負荷分散を行なうことで、分散システムでの効率的なサービス処理を実現できる。そのため、様々な負荷の分散法が提案されている^[1, 2]。負荷分散を行なうには、メッセージ通信により集めた各計算機の負荷情報に基づき、負荷の分散形態を決定する。しかし、時間の経過と共に、集めた情報と実際の計算機の状態が異なってくことがある。以降、これを情報の劣化と呼ぶ。劣化した情報を用いた負荷分散により、逆に特定計算機に負荷が集中してしまうこともある。

我々は、サービス処理の特徴と計算機や通信路の特徴を考慮した、分散システムでの負荷分散法を提案している^[2]。我々の提案方式では、プロセス分散後の各計算機の予測負荷を基に分散形態を決定する。以降、各計算機の負荷をシステム状態と呼ぶ。

本稿では、負荷の予測値を用いることで、情報の劣化を防ぐシステム状態の把握法を示す。

2 システム状態の把握法

2.1 把握の形態

システム状態を把握する方法は、把握の形態により問合せ方式と通知方式に分類することができる。

2.1.1 問合せ方式

問合せ方式とは、負荷分散を行なうときに、他の計算機にシステム状態を問い合わせる方法である。この方式では、システムを構成する計算機数（以降、システム規模と呼ぶ）が大きいと、システム状態の収集時間も大きくなり、起動までのオーバーヘッドが大きくなる。

2.1.2 通知方式

通知方式とは、何らかの契機により特定の計算機（以降、収集計算機と呼ぶ）にシステム状態に関する情報を通知し、負荷分散を行なう際に収集した情報を用いる方法である。既に収集されている情報を用いて負荷分散を行なうので、システム状態の収集にかかる時間は小さい。

通知方式は、収集計算機の集中度と通知契機により分類できる。

収集計算機の集中度 システム状態を収集する計算機の形態により、集中型と分散型に分類できる。集中型は、一部の計算機にシステム状態を通知するので通信量が少ない。しかし、耐故障性に問題がある。この問題は、収集計算機に冗長度を持たせることにより回避できる^[1]。一方、全ての計算機にシステム状態を通知する分散型では、システム規模が大きくなると、通信量の増加が問題となる。

通知契機 システム状態を収集計算機に通知する契機により、定期方式と非定期方式に分類できる。定期方式では定期的に収集計算機にシステム状態を通知し、非定期方式ではシステム状態が大きく変化したときに収集計算機に通知する。非定期方式では、プロセスの起動や終了が頻発すると、通信量が増加する。一方、定期方式では通信量は安定している。しかし、周期が長いと情報の劣化が生じ、特定計算機に負荷が集中する恐れがある。これを避けるには短い周期で情報を更新すればよいが、それでは通信量が増加してしまう。

2.2 把握方式の提案

システム状態把握のための通信量の増加を防ぐには、集中型の通知方式の使用が効果的であることは既に述べた。また、通知契機は、定期方式の方が通信量は安定している。しかし、情報の劣化が問題となる。そこで、予測値を用いた集中型定期通知方式を提案する。

提案方式の手順を以下に示す。

- (1) サービス到着後、収集計算機が持つ各計算機のシステム状態を基に分散形態を決定する。
- (2) 提案した分散法^[2]で得た起動後の計算機の予測負荷とプロセスの起動メッセージを、プロセスを起動する計算機に送る。
- (3) プロセスを起動する計算機は、プロセスを起動し、予測負荷に基づき予測値を更新する。
- (4) 通知の周期毎に、各計算機はシステム状態を調べ、予測値と実測値の差が大きい場合のみ、実測値を収集計算機に通知する。

本方式は、手順(2)において、プロセスを分散するときに収集計算機に予測負荷を通知する。これより、情報の劣化を防いでいる。

2.3 評価

各方式の通信量を比較する。計算機 i の単位時間のプロセスの到着数と終了数をそれぞれ λ_i 、 μ_i とし、計算機 i のシステム状態把握のための通信量を ξ_i バイト、システム規模を n とすると、システム全体の単位時間の通信量 msg は、次のようになる。

Collection of System Load Information for Load Balancing in Distributed System.

Yoshinori AOKI, Hideo TANIGUCHI, and

Kazuo USHIJIMA

Kyushu University

E-mail: aoki, tani, ushijima@csce.kyushu-u.ac.jp

問合せ方式 各計算機は、サービスが到着する度に全ての計算機に問い合わせる。ここで、計算機*i*の単位時間のサービスの到着数を λ'_i 、問い合わせ通信量を ξ'_i とすると以下ようになる。

$$msg = \sum_i \lambda'_i \sum_i (\xi'_i + \xi_i) \quad (1)$$

集中型定期通知方式 各計算機は、周期 τ 毎に収集計算機にシステム状態を通知する。そのため以下のようになる。

$$msg = \frac{\sum_i \xi_i}{\tau} \quad (2)$$

集中型非定期通知方式 各計算機は、プロセスの到着または終了時に収集計算機にシステム状態を通知する。そのため以下のようになる。

$$msg = \sum_i \xi_i (\lambda_i + \mu_i) \quad (3)$$

提案方式 1 プロセスの起動後に予測値が正解する確率を $\alpha\%$ とする。提案した分散法で求める予測値は、プロセスの終了を考慮していないので、プロセスが1つでも終了したら、予測値が正解する確率は0になる。単位時間に1つ以上のプロセスが終了する確率を φ とすると、

$$\begin{cases} \varphi_i = 1 & \mu_i \geq 1 \\ \varphi_i = \mu_i & \mu_i < 1 \end{cases}$$

となり、以下のようになる。

$$msg = \frac{\sum_i \xi_i \left\{ 1 - \left(\frac{\alpha}{100} \right)^{\tau \lambda_i} (1 - \varphi_i)^{\tau} \right\}}{\tau} \quad (4)$$

式(1)～(4)より、以下のことがわかる。

- (A) 式(1)より、問合せ方式の通信量はシステム規模の2乗に比例して増加する。これに対し、式(2)(3)(4)より、通知方式の通信量はシステム規模に比例して増加する。したがって、通信量へのシステム規模の影響は、通知方式の方が小さい。

次に、通知方式において、プロセス到着数と通信量、および周期と通信量の関係を考察する。式(2)～(4)において、全ての*i*で $\lambda_i = \lambda$ 、 $\mu_i = \mu$ とすると、 $\varphi_i = \varphi$ となる。また、全ての*i*で $\xi_i = \xi$ 、 $\alpha = 80\%$ とする。 $\lambda = \mu$ 、 $\tau = 4$ のときのプロセス到着数と通信量の関係を図1に示し、 $\lambda = \mu = 0.2$ のときの周期と通信量の関係を図2に示す。図1と図2より、以下のことがわかる。

- (B) 提案方式は常に通信量が最小である。
(C) プロセス到着数が大きくなっても、提案方式の通信量は大きくならない。
(D) 周期を小さくしても、提案方式の通信量は大きくならない。

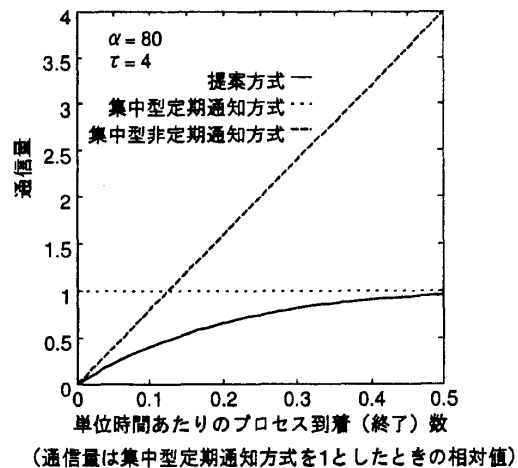


図1 プロセス到着数と通信量の関係

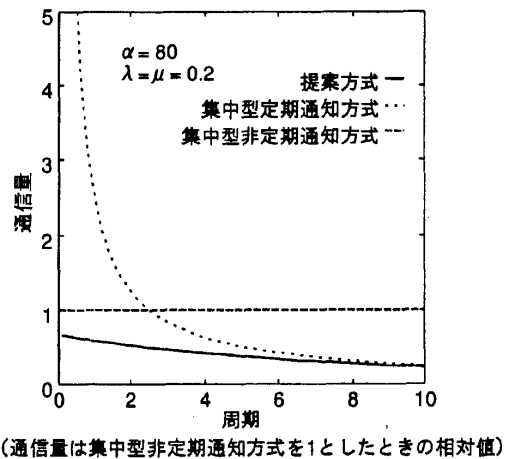


図2 周期と通信量の関係

- (E) プロセス到着数や周期が大きくなると、提案方式の通信量は集中型定期通知方式に近づく。上記(A)～(D)より、提案方式は通信量が少なく、しかも安定していることがわかる。また、(E)より、プロセス到着数や周期が大きいと予測値の効果が小さくなることがわかる。

3 おわりに

本稿では、システム状態を把握する方式として、予測値を用いる方式を提案した。提案方式は、非常に少ない通信量でシステム状態を把握することができ、更に従来の方法では困難だった情報の劣化を防ぐことができることを示した。今後は、有効な周期の決定法について研究する予定である。

参考文献

- [1] Niranjan G. Shivaratri, P. K. and Singhal, M.: Load Distributing for Locally Distributed Systems, *IEEE Computer*, pp. 33-44 (1992).
- [2] 青木義則, 谷口秀夫, 牛島和夫: 応答時間に着目した静的な処理分散法の実現と評価, 情報処理学会シンポジウム論文集, Vol. 95, No. 7, pp. 117-124 (1995).