

分散システムにおける独立性の観点からの資源分散方法の提案

7D-1

宮西 洋太郎 高品 智一 渡辺 尚 水野 忠則
三菱電機 静岡大学工学部 静岡大学工学部 静岡大学工学部

1. はじめに

ある分散システムにおいて、正常時には個々のノードは自己以外のノードと協調しながら正常な動作を行っている。自己と他とを接続する回線や他のノード(特に中央ホスト)等が故障した場合に、ある程度サービス品質が低下しても、自己ノード単独で運転を継続することが望ましい応用分野もある。従来は、切り離され孤立したノードの代りに別のノードで処理を継続するのが一般的な方策であった。ここでは、孤立したノードで運転を継続するという観点、即ち独立性の観点からのシステム構成方式を確立することを目指した提案を行う。

2. 実現方法と検討課題

あるノードが孤立した場合にも、その分担業務を継続するために、自己ノード内に「全体」を代行する資源即ち「分散代行システム」を配置する。資源を、ここではデータに限定し、「分散代行データ」と称する。分散代行データは全体データの複製である。このようなシステム構築方法において検討すべき課題は、

- ・複製データの分散配置方式
- ・孤立運転時のサービス品質
- ・サービス品質と複製データ維持負担とのトレードオフ

等についての定量的判断基準を確立していくことが検討課題である。

3. 複製データの分散配置方式

(1) 対象データのモデル

対象となるデータは、データアイテム単位にアクセスされ、また1回のアクセスは地域的に分散した複数ノードの内の一つを用いて行われるものとする。

また、データベース(以下DBと略)はセンターに全体のデータを持ち、分散したノードには、複製データを持つものとする。

(2) 対象データの特性

データの複製を行う際、対象データの特性を利用する。データ特性は、処理対象によって様々な特性を持つであろうが、一般的には地域的特性及び時間的特性を有すると考えられ、その統計的特性を利用することを提案する。

・地域的特性(局所性)

地域的特性の表現方法として、データアイテム毎の、データアクセスを行ったノードとの関連の強さ、即ち条件付確率が考えられる。

・時間的特性

時間的特性の表現方法として、データアイテム毎の到着率、更新周期毎(時、日、月等)のデータ、時系列履歴データ、バージョン管理等が考えられる。ここでは到着率を考察の対象とする。

次に、上記のデータ特性を数式的に表現する。

単位時間内[Sec]に発生するデータアクセスにおいて、それがデータアイテム i であり、ノード j に到着する到着率 $\lambda(i, j)$ [1/Sec]をここでは一致到着率と称することとする。これは、

$$\lambda(i, j) = P(j|i) \cdot \lambda(i) \quad (1)$$

であり、一致到着率は、データアイテム i の到着率 $\lambda(i)$ [1/Sec]とそのアクセスが j で行われるという条件付き確率 $P(j|i)$ [1]との積で表さ

A Proposal for a Resource Distribution Method from a Viewpoint of Isolation
Yohtaro Miyanishi*, Tomokazu Takashina#, Takashi Watanabe#, Tadanori Mizuno#
*Mitsubishi Electric Corp.
7-10-4 Nishigotanda Shinagawa-ku, Tokyo, 141
#Shizuoka University, Faculty of Engineering
3-5-1 Johoku, Hamamatsu, 432, Japan

れる。ただし $[\]$ はディメンションを表す。

地域的特性を $P(j|i)$ で、アイテム自体の発生度合を $\lambda(i)$ で把握することとする。

(3) データの分割、配置方式

対象データの地域的特性に応じて、DBを分割し、ノードに配置するものとする。分割方式、配置方式に、いくつかの方式が考えられる。

・DB分割方式

分割無し、重複を許容する分割、しない分割

・ノードへの配置方式

全てのノードに配置、ランダムなノードに配置

特定のノードに配置、判断によるノード配置

これらの方式の組み合わせが考えられる。その選択は、データ特性、要求サービス品質、複製維持負担の総合的評価で分割配置方式を選択することが望ましいが、簡略のためデータ特性のみによる方式選択、特にここでは「重複許容の判断ノード配置方式」について検討し、判断方法を提案する。

(4) 重複許容の判断ノード配置方式選択方法

当該方式を選択する条件を次の条件とする。

・対象データに地域的特性が大きい、

$$\sigma_i = \sqrt{(\sum_{j=1}^N P(j|i) - (1/N))^2 / N} \quad (2)$$

$$\sigma = \sum_{i=1}^I \sigma_i \cdot \lambda(i) / \sum_{i=1}^I \lambda(i) \quad (3)$$

このようにして求めた σ が N 、 I に応じた閾値を越える場合には地域的特性ありと判断する。

・アクセスが複数ノードに分散している

全ての i について、最大の $P(j|i)$ が一定値を越えない場合には、分散していると判断する。

(5) 複製データの配置ノードの決定

(a) 複製データを持たせるか否かの判断。

アクセス到着率 $\lambda(i)$ が閾値を越えるデータアイテム i については複製をもつこととする。

(b) 複製を配置するノードの決定。

各 i について、 $P(j|i)$ の大きい順に j を並べ直す。並べ直した j を j_{11}, j_{21}, \dots とする。従って大きい順に並べ直した $P(j|i)$ は、 $P(j_{11})$, $P(j_{21}), \dots$ となる。上位のものから、一定の閾値 P_0 を越えるまで r を増加して加算する。

$$\lambda(i) \cdot \sum_{r=1}^R P(j_{r1}) \geq \lambda_m \cdot P_0 \quad (4)$$

$$\lambda_m = (1/I) \cdot \sum_{i=1}^I \lambda(i) \quad (5)$$

上記の式が成立する最初の r を R_i とし、このときのノード $j_{11}, j_{21}, \dots, j_{R_i1}$ に複製データを配置する。

(6) 複製維持のための負担

複製を維持するための記憶容量、処理資源を評価する。

4. 孤立運転時のサービス品質の低下

ここでは一つの評価方法として、特定ノードが孤立運転時に、発生したアクセスが当該ノードに複製データを持つアイテムならばサービスを受けることができ、持たないならばサービスを受けることができない、という状況を評価する。

ノード j にアイテム i の複製を持っている確率 q_{ij} は、

$$q_{ij} = \sum_{r=1}^R \delta_{jj_{r1}} \quad (6)$$

であり、孤立運転時に特定ノード j においてデータアイテム i がサービスを受けられる率 $\lambda'(i, j)$ は、複製が存在する場合には λ がそのまま維持され、複製が存在しない場合には λ は 0 となりアクセスが発生してもサービスは行われぬ。従って、

$$\lambda'(i, j) = \sum_{r=1}^R \delta_{jj_{r1}} \cdot \lambda(i, j) \quad (7)$$

ここで、 $\delta_{jj_{r1}} = 1$ ($j = j_{r1}$ の場合のみ) である。これを全ての i について評価するため、サービス継続比率 SL として次ぎの評価指数を提案する。

$$SL = \sum_{i=1}^I \lambda'(i, j) / \sum_{i=1}^I \lambda(i, j) \quad (8)$$

前述の一定閾値 P_0 を大きくすれば R_i が小さくなり、それだけ複製を持たせるノードの数が少なくなり、その結果 $\lambda' = 0$ の i が多くなり、 SL は小さくなる。

5. 今後の課題

今後は上記の方法の詳細検討と妥当性の評価、複製維持負担の評価、一貫性の検討、さらには他の分割配置方式の検討等を行っていきたい。