

分散処理システムによるデータバックアップ モデルのペトリネット表現と性能評価

5R-1

亀井 光雄, 菅澤 喜男
(日本大学)

1. まえがき

近年、高度情報化社会の進展に伴い、情報伝達を担う情報通信ネットワークの役割はますます重要になっている。通信の中断あるいは停止は大きな社会的混乱を引き起こす要因ともなる。また、企業内あるいは企業間で連結した情報通信ネットワークについても、企業の利益あるいは情報管理への影響は多大であり無視出来ない。

情報通信ネットワークの障害あるいは停止による損害を最小限に食い止める為に、分散処理形態を有するシステムの在り方とその構築方法も重要な課題となっている。

更に、分散処理形態を有するシステムのデータをいかにバックアップしデータを安全かつ正確に保存するかはシステムの信頼性・安全性に多大な影響を及ぼす。

ここでは、基本的な分散処理形態を有するシステムのデータをバックアップするモデルを取り上げ、システムの性能評価をするものである。システムのモデル化はペトリネット (Petri Net: 以後PNと記す) を用いて行なう。PNでモデル化されたシステムは、マルコフ再生過程 (Markov renewal process: 以後MRPと記す) によりシステムの性能評価を行なう。

2. 分散システムのモデル化

分散システムのモデル化をPNを用いて表現する。それは○で示される場所 (Place) とIで示される遷移 (Transition) そして有向枝→から成る図で分散システムのモデル化が行なわれる。

2つのプロセッサ、主 (メイン) プロセッサと従 (サブ) プロセッサからなる分散システムの主プロセッサは、一定時間になるとデータのバックアップを受ける。主プロセッサのデータをバックアップするために、ここで取り上げる分散システムでは、2つのバックアップが用意されているものとする。つまり、主プロセッサのデータはバックアップ1でセーブされ、バックアップ1が一杯になったらバックアップ2がデータを確実にセーブする。また、バックアップ2が一杯になればバックアップ1が主プロセッサのデータをバックアップする。ただし、バックアップは主プロセッサに対してだけ行なわれ、かつ主プロセッサがバックアップしている間は分散システムは故障あるいは他の通信障害などは起きないものとする。

2.1 バックアップ方針

分散システムがバックアップを行なう条件としては、主プロセッサはオンライン処理を行ない、従プロセッサはローカルを処理を行ない、主従プロセッサが交替を行なって

いる間に交替処理システムが通信障害等でシステム故障する場合にも行なわれる。主プロセッサに一定時間 (定時) でデータバックアップを施す。

2.2 ペトリネットモデル

このような挙動をする分散システムと2つのバックアップ機能を有するシステムをPNによりモデル化すると図-1となり、各場所 P_i ($i=1, 2, \dots, 8$) は次の通りである。

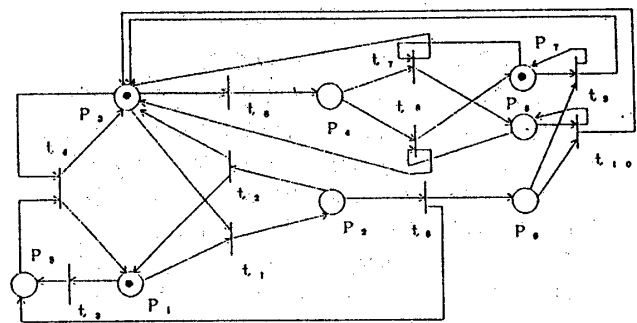


図1. 分散システムのバックアップを考慮したPNモデル

- P_1 : 従プロセッサがローカル処理中。
- P_2 : 従プロセッサが主プロセッサと交替中。
- P_3 : 主プロセッサがオンライン処理中。
- P_4 : 主プロセッサがバックアップ中。
- P_5 : 従プロセッサの処理機能が停止。
- P_6 : 交替処理システムがシステム故障中。
- P_7 : 主プロセッサのデータがバックアップ1ある。
- P_8 : 主プロセッサのデータがバックアップ2ある。

各遷移 t_j ($j=1, 2, \dots, 10$) は次の通りである。

- t_1 : 従プロセッサがローカルからオンライン処理に切り換えられ、主従プロセッサの交替が始まる。
- t_2 : 主従プロセッサの交替が終了し、従プロセッサがローカル処理に戻る。
- t_3 : 通信障害等により従プロセッサが処理不能なる。
- t_4 : 主プロセッサの機能により従プロセッサの機能が回復する。
- t_5 : 主従プロセッサの交替中に通信障害が発生、交替処理システムがシステム故障する。
- t_6 : 主プロセッサがバックアップを開始する
- t_7 : 主プロセッサがバックアップ1を終了し、バックアップ2を更新する。
- t_8 : 主プロセッサがバックアップ2を終了し、バックアップ1を更新する。
- t_9 : バックアップ1により主プロセッサの機能を回復し、主プロセッサがオンライン処理に戻る。
- t_{10} : バックアップ2により主プロセッサの機能を回復し、主プロセッサがオンライン処理に戻る。

図1において、場所 P_1, P_3, P_7 の中に●が示してあるのは、初期刻印であり、システムの初期状態を示す。つまり、入力場所にある●印で示される標号 (Token) を遷移の発火 (Fire) により出力場所に移動することで刻印 (Marking) がなされ、システムの状態を捕えることができる。図1はシステムの静的な関係だけを示したものである。

図2は図1を到達可能木で表したものであり、木に開始節点の初期刻印を置き、木の節点は刻印、木の枝は刻印間の直接到達可能であることを示している。つまり、システムの挙動を動的変化として把握できる。

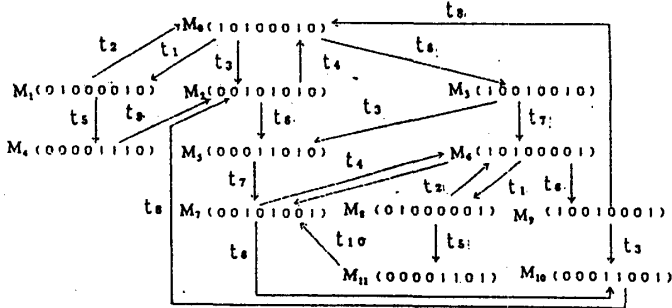


図2. 図1の到達可能木

3. MRPによるシステムの挙動

ここではMRPをモデルに適用して分散システムの挙動を確率量を求めて、性能評価を行なうために必要な尺度を明かにする。図2で示したシステムの動的な性質を捕えてMRPを適用するには、各遷移が発火する時点に着目し刻印によりシステムの状態を把握する。従って、図2で示した到達可能木における $M_i (i=0,1,2,\dots,11)$ をシステムの状態として定義すれば、同じ状態が同時に生じることはない。よって、図2における節点がMRPにおける状態を表し、木の枝が状態間の推移を表すので、各遷移が発火する時点に着目し、刻印によってシステムの状態を捕えることでMRPを適用することができる。

ここで、時刻 $t=0$ で状態 $M_i (i=0,1,2,\dots,11)$ に推移した後に、時刻 t までに $t_j (j=1,2,\dots,10)$ が発火するまでの時間を発火時間分布 (Fire Time Distribution) $F_j(t) (j=1,2,\dots,10)$ とする。例えば、図2において時刻 $t=0$ で状態 M_0 にシステムがある時に、時刻 t までに遷移 t_1, t_3, t_6 が発火する時間は発火時間分布 $F_1(t), F_3(t), F_6(t)$ のいずれかに従うことを意味している。よって、PN表現されたシステムが各状態 $M_i (i=0,1,2,\dots,11)$ に推移する時点に着目し、それにより各時点に対応する状態を定義することになる。図2で示した各状態は次の通りである。

- M_0 : 従プロセッサはローカル処理をし、主プロセッサがオンライン処理をする。主プロセッサのデータはバックアップ1にある。
- M_1 : 従プロセッサは主プロセッサと交替し、主プロセッサのデータはバックアップ1にある。
- M_2 : 主プロセッサはオンライン処理をし、従プロセッサの機能が停止する。主プロセッサのデータはバックアップ1にある。
- M_3 : 従プロセッサはローカル処理をし、主プロセッサはバックアップする。主プロセッサのデータはバックアップ1にある。
- M_4 : 従プロセッサの処理機能が停止し、交替処理システムがシステム故障する。主プロセッサのデータはバックアップ1にある。
- M_5 : 主プロセッサはバックアップをし、従プロセッサの処理機能が停止する。主プロセッサのデータはバックアップ1にある。
- M_6 : 従プロセッサはローカル処理をし、主プロセッサはオンライン処理をする。主プロセッサのデータはバックアップ2にある。
- M_7 : 主プロセッサはオンライン処理をし、従プロセッサの処理機能が停止する。主プロセッサのデータはバックアップ2にある。
- M_8 : 従プロセッサが主プロセッサと交替する。主プロセッサのデータはバックアップ2にある。

- M_9 : 従プロセッサはローカル処理をし、主プロセッサはバックアップをする。主プロセッサのデータはバックアップ2にある。
- M_{10} : 主プロセッサはバックアップをし、従プロセッサの処理機能が停止する。主プロセッサのデータはバックアップ2にある。
- M_{11} : 従プロセッサの処理機能が停止し、交替処理システムがシステム故障する。主プロセッサのデータはバックアップ2にある。

3.1 MRPによる挙動解析

図2で示した、 M_i の各状態から状態へと推移する発火は瞬間的かつ完全である。二つ以上の状態が同時に生じする確率はない。従って、 $M_i (i=1,2,3,5,6,7,8,9,10)$ 状態に推移をした後に、時間 t 以内に状態 $M_j (j=1,2,3,5,6,7,8,9,10)$ の状態に推移する状態間の1ステップ推移時間分布を $Q_{ij}(t)$ としたときのラプラス・ステュルチェス変換を $q_{ij}(s)$ とする。

例えば、 M_0 から M_1 への推移時間分布 $Q_{01}(t)$ は式(1)で求められ、 M_1 から M_2 の非再生点 M_1 を含む推移時間分布 $Q_{12}^{(1)}(t)$ は式(2)から求められる。

$$Q_{01}(t) = \int_0^t F_3(v) F_6(t-v) dF_1(v) \quad (1)$$

$$Q_{12}^{(1)}(t) = \int_0^t \int_0^{t-u} dF_1(v) dF_2(w) dF_1(u) \quad (2)$$

他の推移時間分布 $Q_{ij}(t)$ も同様に求めることができる。

3.2 極限確率

分散システムの確率的挙動をシステムが各状態に留まる極限確率を求めて明かにする。そこで、本論文では、埋め込まれたマルコフ連鎖 (Embedder Markov Chain)を導入して求める。従って、MRPにおけるマルコフ連鎖は、単位時間当たりにある状態から他のある状態へ推移する推移確率に着目する方法である。

4. おわりに

PN表現を利用して分散処理形態を有するシステムをモデル化し、マルコフ再生過程と到達可能木と対応づけ分散システムの性能評価の方法を示した。ここでは、求め得る極限確率に関して、具体的な数値例は示されていないが、各発火時間分布に具体的な分布を導入することで数値例を示すことが可能である。特に、遷移 t_i は定時に発火すべく発火時間分布を対応させ、また、遷移 t_j は指数分布以外の分布としてガンマ分布あるいはワイブル分布等を適用する事を試みる。

参考文献

- (1) Peterson J.L.: "Petri Net Theory and the Modeling of systems", Prentice-Hall, Englewood Cliffs, NJ (1981)
- (2) Dugan J.B. and Triveri K.S.: "Coverage Modeling for Dependability Analysis of Fault-Tolerant Systems", IEEE Trans. of Computers, 38, 6, pp.775-787 (1989).
- (3) Qun JIN, Mituo KAMEI and Yoshio SUGASAWA.: "Behavioral Analysis and Performance Evaluation of a Shift Processing System by an Extended Stochastic Petri Net", The Institute of Electronics, Information and Communication Engineers Vol.E75A, pp.1378-1384 (1992)
- (4) 金, 菅澤, 瀬谷: "分散処理におけるシンク発生の確率モデルと挙動解析", 信学論, J75A, 3, pp.658-660 (1992).
- (5) 菅澤: "ロボットとコンベアによる単独・強調動作の確率的挙動", 日本経営工学会誌, 39, 6 pp.378-383 (1989).