1995年度日本オペレーションズ・リサーチ学会
春季研究発表会

最適チェックポインティング方策におけるシステム障害の影響について

小田切政徳，土肥正，海生直人[†]，尾崎俊治
広島大学工学部，[†]広島修道大学商学部

1 はじめに

チェックポイント技法とは，予め定められた時点ごとに主記憶上の情報を安定な二次記憶媒体に保存する作業のことである．チェックポイントの生成をあまり行わない場合，チェックポインティング自体にかかる費用は少なくてすむが，システム障害が発生したときに失われるデータの量が非常に多くなる．これに対して，チェックポイントの生成を頻繁に行なう場合，システム障害が発生したときに失われるデータの量は，最新のチェックポイントからシステム障害発生までの量で済み，チェックポイントの生成をあまり行わない場合と比較して失われるデータの量は少なくなる．しかしながら，チェックポインティング自体にかかる費用が多くなることは明かである．そこで，チェックポイントの生成頻度・時機を如何に決定するかが重要な問題となる．

一方，障害発生によって失われたデータを回復する技法としてジャーナル技法[1,2]がある．これは，システムに影響を与える全ての処理をジャーナルとして記録しておき，システム障害が発生した際には記録されたジャーナルに従って障害が発生した直前までシステムを回復させるものである．

本報告では，二次記憶媒体上のチェックポインティングに起因するファイル破壊が発生する(すなわち以前のチェックポイント生成時のデータが全て失われてしまう)場合について，ジャーナルを使う場合と使わない場合の二つの場合について最適チェックポインティング方策を導出し，最適チェックポインティング方策におけるシステム障害の影響について言及する．

2 モデルと仮定

文献[3, 4]に従って，本報告で扱うモデルに対して以下の仮定をおく．

1) 単一コンピュータシステムを扱う．システムは時刻0で動作を開始し，計画期間は無限大とする．
2) システムはチェックポイント t_i $(i=1,2,3,\cdots)$ で瞬時にチェックポインティングされる $(t_0=0)$．
3) システム障害は直ちに発見され，同時にシステムのリカバリが瞬時に施される．システムの動作開始からシステムの障害発生・リカバリまでの期間(ジャーナルを使う場合はリカバリ完了後ジャーナル回復が終了するまで)を1サイクルとし，同様なサイクルを繰り返す．但し，ジャーナルを使う場合にはジャーナル回復後すぐに障害発生を伴わないチェックポインティングがなされる．
4) システム固有の障害発生時間は有限な平均 $1/\lambda$ を持つ任意の確率分布 $F(t)$ に従う．
5) 時刻 t_{i-1} $(i=1,2,3,\cdots)$ においてシステムが正常に動作しているとき，時間間隔 $(t_{i-1},t_i]$ の間にシステム障害が発生する確率を p $(0<p<1)$，即ち，
$$\frac{F(t_i)-F(t_{i-1})}{\overline{F}(t_{i-1})}=p \tag{1}$$
とする．
6) システム障害は処理期間中だけでなく，チェックポインティングの際にもチェックポインティングに起因して確率 a $(0\le a\le 1)$ で発生する．
7) チェックポインティングによるシステム障害の際に確率 b $(0\le b\le 1)$ でファイルが破壊される．ファイルの破壊は処理期間中には発生しないものとする．

本稿では，確率 a と b を共にチェックポインティングに起因する障害発生確率と呼ぶことにする．

さらに以下の諸量を定義する．

c_c: チェックポインティング1回当たりの費用．これはチェックポイントでシステム障害が発生した場合に対しても同様にかかるものとする．

k_l: システム障害発生直前のチェックポイントからシステム障害発生時点まで，あるいは動作開始からシステム障害発生までの損失期間に対する単位時間当たりの損失費用．

k_r: ジャーナル回復にかかる単位時間当たりの処理費用．

c_r: システム障害に対するリカバリ1回当たりの費用．

$E_C(p)$: 1サイクル当たりの期待費用．

$E_T(p)$: 1サイクルの平均時間．

$C(p)$: 単位時間当たりの期待費用．

3 期待費用関数の導出

3.1 ジャーナルを使わない場合

1サイクル当たりの期待費用 $E_C(p)$ はチェックポイント生成にかかる費用，システム障害発生による時間損失にかか

る費用，システム障害に対するリカバリにかかる費用の和として次式のように求められる．

$$E_C(p) = c_c\left[\sum_{i=1}^{\infty}(i-1)(\bar{p}\bar{a})^{i-1}p + \sum_{i=1}^{\infty}i(\bar{p}\bar{a})^{i-1}\bar{p}a\right] + k_l\left[\sum_{i=1}^{\infty}\bar{a}\int_{t_{i-1}}^{t_i}(t-t_{i-1})dF(t) + \sum_{i=1}^{\infty}(t_i-t_{i-1})(\bar{p}\bar{a})^{i-1}\bar{p}a\bar{b} + \sum_{i=1}^{\infty}t_i(\bar{p}\bar{a})^{i-1}\bar{p}ab\right] + c_r. \quad (2)$$

ここで，一般に$\bar{\psi}=1-\psi$．式(2)は次のように展開することができる．

$$E_C(p) = \frac{c_c\bar{p}}{1-\bar{p}\bar{a}} + k_l\left[\bar{a}\left(\frac{1}{\lambda} - p\sum_{i=1}^{\infty}t_{i-1}\bar{p}^{i-1}\right) + \bar{p}a\bar{b}\sum_{i=1}^{\infty}(t_i-t_{i-1})(\bar{p}\bar{a})^{i-1} + \bar{p}ab\sum_{i=1}^{\infty}t_i(\bar{p}\bar{a})^{i-1}\right] + c_r. \quad (3)$$

式(3)においてpを変数とし，$E_C(p)$を最小化することによって最適チェックポイント時刻列を求めることができる．また，1サイクルの平均時間$E_T(p)$は

$$E_T(p) = \frac{\bar{a}}{\lambda} + \bar{p}a\sum_{i=1}^{\infty}t_i(\bar{p}\bar{a})^{i-1} \quad (4)$$

となる．これより，単位時間当たりの期待損失費用$C(p) = E_C(p)/E_T(p)$を解析的に求めることができる．

3.2 ジャーナルを使う場合

この場合，1サイクル当たりの期待費用$E_C(p)$はチェックポイント生成にかかる費用，ジャーナル回復にかかる費用，システム障害に対するリカバリにかかる費用の和となるので

$$E_C(p) = c_c\left[\sum_{i=1}^{\infty}i(\bar{p}\bar{a})^{i-1}p + \sum_{i=1}^{\infty}(i+1)(\bar{p}\bar{a})^{i-1}\bar{p}a\right] + k_r\frac{\kappa}{\mu}\left[\sum_{i=1}^{\infty}\bar{a}\int_{t_{i-1}}^{t_i}(t-t_{i-1})dF(t) + \sum_{i=1}^{\infty}(t_i-t_{i-1})(\bar{p}\bar{a})^{i-1}\bar{p}a\bar{b} + \sum_{i=1}^{\infty}t_i(\bar{p}\bar{a})^{i-1}\bar{p}ab\right] + c_r \quad (5)$$

となり，整理すると，

$$E_C(p) = \frac{c_c\bar{p}}{1-\bar{p}\bar{a}} + c_c + k_r\frac{\kappa}{\mu}\left[\bar{a}\left(\frac{1}{\lambda} - p\sum_{i=1}^{\infty}t_{i-1}\bar{p}^{i-1}\right) + \bar{p}a\bar{b}\sum_{i=1}^{\infty}(t_i-t_{i-1})(\bar{p}\bar{a})^{i-1} + \bar{p}ab\sum_{i=1}^{\infty}t_i(\bar{p}\bar{a})^{i-1}\right] + c_r \quad (6)$$

となる．ここでκ, μはそれぞれ処理するデータの到着率およびサービス率である．また，1サイクルの平均時間は

$$E_T(p) = \bar{a}\left[\frac{1}{\lambda} + \frac{\kappa}{\mu}\left\{\frac{1}{\lambda} - p\sum_{i=1}^{\infty}t_{i-1}\bar{p}^{i-1}\right\}\right] + \bar{p}a\bar{b}\sum_{i=1}^{\infty}\{t_i + \frac{\kappa}{\mu}(t_i-t_{i-1})\}(\bar{p}\bar{a})^{i-1} + \bar{p}ab(1+\frac{\kappa}{\mu})\sum_{i=1}^{\infty}t_i(\bar{p}\bar{a})^{i-1} \quad (7)$$

となり，ジャーナルを使わない場合と同様にして単位時間当たりの期待損失費用$C(p)$を求めることができる．

4 数値例と考察

システム固有の障害発生時間が指数分布に従う場合において，チェックポインティングに起因する障害発生確率に対する最適チェックポイント時刻列の従属性について調べた．障害発生時間分布の無記憶性から，最適チェックポイント時刻列は一定間隔になるが，確率aの増大に伴いチェックポイント間隔は減少する一方で，確率bが大きくなればその間隔は増加傾向を示すことが確認された．これは，処理期間中以外の障害発生確率が小さくなるとチェックポインティングを頻繁に行わなくて済む反面，ファイルが破壊される可能性が大きくなれば，チェックポインティングを頻繁に行うことが無意味になることを示唆している．

参考文献

[1] 当麻喜弘，"コンピュータシステムの高信頼化技術入門"，日本規格協会，(1988).

[2] 福本聡，海生直人，尾崎俊治，"チェックポイント濃度を使用した最適チェックポインティング方策"，情報処理学会論文誌，第31巻，第6号，(1990).

[3] M. Odagiri, N. Kaio and S. Osaki, "A Note on Optimal Checkpoint Sequence Taking Account of Preventive Maintenance", *IEICE Trans. Fundamentals*, Vol. E77-A, No. 1, (1994).

[4] 柳原正義，小田切政徳，海生直人，尾崎俊治，"チェックポインティングに起因するシステム障害を考慮した最適チェックポインティング方策"，電子情報通信学会論文誌(採録予定).