

## ファイル回復機構に対する古文書コピーの最適世代数と最適生成間隔

6 Z-6

中川 さより† 福本 聡‡ 石井 直宏†

†名古屋工業大学知能情報システム学科

‡愛知工業大学情報通信工学科

## 1 はじめに

データベース・システムでは、障害に備えて、二次記憶装置内に古文書コピー (Archive Copy) を定期的に生成し [1], データが損傷したら、これらの古文書コピーを用いてファイルを回復する。その際、エラーが発生してから発見するまでに長い時間がかかる場合もあるため、古文書コピーはいくつかの世代を保持していなければならない。

ここでは、回復処理オーバーヘッドを最小にする古文書コピーの最適間隔  $T^*$  と、回復までの全費用を最小にする最適世代数  $n^*$  について、数値例を示し議論する。

## 2 回復処理オーバーヘッド

$n$  個の古文書コピーは一定間隔  $T$  で生成され、フォールト発生から障害までのフォールト潜在時間  $X$  は、平均  $1/\mu$  をもつ指数分布  $(1 - e^{-\mu x})$  に従う。文献 [1] では、障害は古文書コピー生成間隔の中点、 $T/2$  で起こると仮定されていたが、ここではランダムに発生するとする。

このモデルでは、障害から次の仮の古文書コピーまでの時間  $Y$  は確率分布  $G(y)$  を持ち、 $Y$  は  $[0, T]$  における一様分布であると仮定する。いわば、

$$G(y) = \begin{cases} y/T & 0 \leq y \leq T, \\ 1 & y > T. \end{cases} \quad (1)$$

この仮定のもとで、最も近い古文書コピーからロールフォワード、あるいは、ロールバックすることにより、フォールト発生まで回復する。その際、回復処理オーバーヘッド  $r(X, Y)$  は、 $n \geq 2$  ( $1 \leq$

$k \leq n-1$ ) に対して、

$$r(X, Y) = \begin{cases} X & 0 \leq X < \frac{T-Y}{2}, \\ T - (X+Y) & \frac{T-Y}{2} \leq X < T-Y, \\ X+Y - kT & kT \leq X+Y < kT + \frac{T}{2}, \\ (k+1)T - (X+Y) & kT + \frac{T}{2} \leq X+Y < (k+1)T, \\ X+Y - nT & nT \leq X+Y < \infty. \end{cases} \quad (2)$$

モデルの簡単化のため、古文書コピーを作るためのオーバーヘッドは無視し、ロールフォワードのオーバーヘッドもロールバックのオーバーヘッドも処理時間と等しいとする。

回復処理オーバーヘッドの期待値  $R_n(T)$  は、 $X$  が指数分布  $(1 - e^{-\mu x})$  に従い、 $Y$  が式 (1) で与えられる一様分布に従うとき、

$$\begin{aligned} R_n(T) &\equiv E[r(X, Y)] \\ &= \frac{1}{\mu} - \frac{2}{\mu^2 T} \left(1 - e^{-\frac{\mu T}{2}}\right) \left(1 - e^{-\mu(n-\frac{1}{2})T}\right) \\ &\quad (n = 1, 2, \dots). \end{aligned} \quad (3)$$

3 最適生成間隔  $T^*$ 

回復処理オーバーヘッドの期待値  $R_n(T)$  を最小にする古文書コピー生成間隔  $T^*$  を求める。式 (3) で与えられる  $R_n(T)$  を  $T$  で微分して、0 とおくと、

$$\begin{aligned} 1 &- \left(1 + \frac{\mu T}{2}\right) e^{-\frac{\mu T}{2}} \\ &- \left[1 + \mu \left(n - \frac{1}{2}\right) T\right] e^{-\mu(n-\frac{1}{2})T} \\ &+ (1 + \mu n T) e^{-\mu n T} = 0. \end{aligned} \quad (4)$$

式 (4) の左辺を  $Q(T)$  で表すと、 $Q(T)$  は  $T = 0$

Optimal Number and Interval of Archive Copies for a File Recovery

Sayori Nakagawa†, Satoshi Fukumoto‡, Naohiro Ishii†  
†Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466, Japan

‡Aichi Institute of Technology, 1247 Yachigusa, Yakusa-cho, Toyota 470-03, Japan

のとき0で、しばらく減少して、その後1まで増加する。よって、式(4)を満たす唯一で有限の $T^*$  ( $0 < T^* < \infty$ )が存在する。

4 最適世代数  $n^*$

古文書コピーのそれぞれの世代の生成に要する費用を  $c_1$ , 単位時間当たりの回復処理オーバーヘッドに必要な費用を  $c_2$  とする。

回復に必要な全費用は,

$$C_n(T) = nc_1 + c_2 R_n(T) = nc_1 + \frac{c_2}{\mu} \left[ 1 - \frac{1}{\mu T} \left( 1 - e^{-\frac{\mu T}{2}} \right) \left( 1 - e^{-(n-\frac{1}{2})\mu T} \right) \right] \quad (n = 1, 2, \dots). \quad (5)$$

ある与えられた生成間隔  $T$  ( $> 0$ ) に対して,  $C_n(T)$  を最小にする最適世代数  $n^*$  を求める。明らかに,

$$C_1(T) = c_1 + \frac{c_2}{\mu} \left[ 1 - \frac{2}{\mu T} \left( 1 - e^{-\frac{\mu T}{2}} \right)^2 \right], \quad C_\infty(T) \equiv \lim_{n \rightarrow \infty} C_n(T) = \infty. \quad (6)$$

よって,  $C_n(T)$  を最小にする少なくとも有限の  $n^*$  ( $1 \leq n^* < \infty$ ) が存在する。  $C_{n+1}(T) - C_n(T) \geq 0$  より,

$$e^{n\mu T} \geq \frac{c_2}{c_1 \mu} \frac{2}{\mu T} \left( 1 - e^{-\frac{\mu T}{2}} \right) \left( e^{\frac{\mu T}{2}} - e^{-\frac{\mu T}{2}} \right). \quad (7)$$

いわば,

$$n \geq \frac{1}{\mu T} \ln \left( \frac{Ac_2}{c_1 \mu} \right). \quad (8)$$

ここで,

$$A \equiv \frac{2}{\mu T} \left( 1 - e^{-\frac{\mu T}{2}} \right) \left( e^{\frac{\mu T}{2}} - e^{-\frac{\mu T}{2}} \right).$$

よって, 最適世代数  $n^*$  は, 式(7)か式(8)を満たす最小の正の整数として求められる。

5 数値例

表1は  $\mu T$  と  $c_2/(c_1 \mu)$  に対する最適世代数  $n^*$  を示す。ここで,  $c_2/\mu$  は平均フォールト潜在時間の回復処理オーバーヘッドに対する費用を表し,  $c_1$  は古文書コピーの1世代に対する費用を表す。これらの  $n^*$  は  $\mu T \rightarrow \infty$  と  $\mu T \rightarrow 0$  の両方の場合,

1に近付くので,  $c_2/(c_1 \mu)$  が増加すると増加するが,  $\mu T$  が増加しても必ずしも減少しない。

表2は  $n$  と  $c_2/(c_1 \mu)$  に対する最適間隔  $\mu T^*$  と全費用  $C_n(T^*)/c_1$  を与える。これら  $\mu T^*$  の値は  $n$  とともに減少し,  $C_n(T^*)$  は凹型の関数であることが分かる。したがって,  $c_2/(c_1 \mu)$  に対して最小の  $n^*$  が存在する。アンダーラインは  $c_2/(c_1 \mu)$  に対する最小費用を表す。

この表から, 多くの古文書コピーを生成するのは不要であるということが分かる。

表1.  $\mu T$  と  $c_2/(c_1 \mu)$  に対する古文書コピーの最適世代数  $n^*$

$\mu T$	$c_2/(c_1 \mu)$				
	5	10	20	50	100
0.1	1	1	7	16	23
0.5	2	3	5	7	8
1.0	2	3	3	4	5
1.5	2	2	3	3	4
2.0	2	2	2	3	3
2.5	1	2	2	2	3
3.0	1	2	2	2	2

表2.  $n$  と  $c_2/(c_1 \mu)$  に対する最適間隔  $\mu T^*$  と全費用  $C_n(T^*)/c_1$

$n$	$\mu T^*$	$c_2/(c_1 \mu)$				
		5	10	20	50	100
1	2.51	<u>4.98</u>	8.96	16.93	40.82	80.64
2	1.37	5.42	<u>8.84</u>	15.69	36.22	70.44
4	0.79	7.06	10.13	16.26	<u>34.65</u>	65.29
6	0.58	8.92	11.84	17.68	35.19	<u>64.38</u>

6 まとめ

障害が  $(0, T]$  間の一様分布で発生するときの回復処理オーバーヘッドの期待値と費用を求め, 古文書コピーの最適生成間隔  $T^*$  と最適世代数  $n^*$  を解析した。最適値は簡単な式の唯一の解で与えられるということを示した。また, 数値例で, 全費用を最小にする最適な組  $(n^*, T^*)$  を計算した。

参考文献

[1] 福本聡, 中川 暉夫, “古文書コピーによるファイル回復機構の解析,” 信学論, J79-D-1, no. 4, pp. 206-211, 1996.