クラウドコンピューティング環境における HDD 故障率モデルの提案

樋渡 仁　　岩村 相哲　　新井 克也（日本電信電話株式会社）

概要　本論文では，我々のクラウドコンピューティング環境において，多数の HDD が故障した事例を取り上げ，故障実績を調査した．この故障実績をもとに，HDD の寿命分布を求め，代表的な寿命分布をあてはめた．赤池の情報量基準（AIC）を用い，適合度を評価すると，ワイブル分布（形状母数 α=1.508）が最も良くあてはまることが分かった．この結果から，交換用 HDD の必要数を推定する手法を提案した．さらに，RAID コントローラが保持する HDD の状態と HDD の平均寿命について調査した．コマンド中断数，メディア故障数，接続失敗数と HDD の平均寿命に有意な相関関係があり，幾つかの仮定の下で我々の HDD の故障物理が速度過程モデルに従うことが分かった．この結果から，メンテナンス時に HDD の余命を推定し，予防的に交換する手法を提案した．

1. はじめに

近年，大量の計算資源をネットワーク経由で利用するクラウドコンピューティング（Cloud Computing）が注目を集めている．NTT においても，クラウドコンピューティング技術 CBoC（Common IT Base over Cloud Computing）の研究開発を推進している[1,2]. 我々は, 2008 年より，CBoC で利用するクラウドコンピューティング環境の構築・運用に従事してきた. CBoC の研究開発は，仮想化運用管理技術[3]と大規模分散処理システム[4]が並行して進められており，クラウドコンピューティング環境も仮想化運用管理技術向け環境と大規模分散処理システム向け環境の 2 系が併存し，運用されている．クラウドコンピューティング環境は，多数のサーバから構成されるため，サーバ故障が頻発する．このような特徴を持つ環境を効率的に運用するには，故障率等の基礎情報を把握することが重要である．

本論文では，サーバ故障の主な原因であるハード・ディスク・ドライブ（HDD: Hard Disk Drive）の故障に着目し，HDD 故障率をモデル化することを試みる．特に，大規模分散処理システム向け環境のうち, HDD の故障が多く発生した 2 クラスタ（以下，クラスタ#0,#1）の故障実績を解析し, HDD 故障率を予測する 2 手法を提案する（2 章，3 章）．2 章では，故障が発生する機器が HDD である事実は用いずに，全ての機器に適用できる統計的な寿命分布を用いた手法について述べ，交換用の HDD を調達する時に，ある時期までに必要になる数量を算出できることを示す. 3 章では，故障が発生した機器が HDD である事実を用い, HDD が持つ状態からより高い精度で故障率を予測する手法について述べ，サーバを保守する時に，故障率が高い HDD を発見し，予防的に交換できることを示す．4 章では，2 章，3 章の議論を受け，提案手法と既存手法を比較し，提案手法を評価する．最後に，5 章では，本論文をまとめる．

2. 寿命分布による故障率モデル

HDD に限らず，故障率を求めるためには，故障寿命がどのような分布形に従っているかを調べる必要がある. 2 章では，HDD 故障実績から，寿命分布を求め，代表的な寿命分布[5]である指数分布（Exponential distribution），ワイブル分布（Weibull distribution），対数正規分布（Log-normal distribution）へあてはめる．さらに，どの寿命分布への適合度が良いかを，赤池の情報量基準（AIC）により評価する．

2.1 HDD 故障実績の測定

クラスタ#0,#1 は, 同じサーバから構成されるクラスタで，同じ用途で一体として運用されている．クラスタ #0,#1 を運用した結果，HDD 故障が多く発生した．故障実績は, HDD を交換した事実と日時を記録したものである．表 1 に故障実績の概要を示す．ここで注意すべきことは，故障実績が，HDD を交換した事実のみを記録している点である．すなわち，交換した HDD が，本当に故障しているかは未確認となる．このため，本論文において，故障実績と述べているものは，厳密には交換実績と述べるべきであり，故障率（Failure Rate）と述べているものは，厳密には交換率（Replacement Rate）と述べるべきである．

次に，故障寿命を求める．故障寿命は，2010 年 1 月 1 日（基準日）時点で使用中の HDD（1,140 台）を対象にして求めるため，表 1 の故障実績には，この期間中に故障が発生して交換した HDD が，再度故障した HDD（4 台）は含んでいない．このため，現実のクラスタ#0,#1 は修理系だが，以下の議論では非修理系として扱うことができる．

表 1. 故障実績の概要

対象	クラスタ#0,#1
期間	2010 年 1 月 1 日～2010 年 10 月 4 日
サーバ数	285 台
HDD 総数	1,140 台（HDD 故障数: 112 台）
HDD 仕様	SATA 1TB（コンシューマ用途）

表 1 の故障実績から図示した HDD 寿命分布を，図 1 に示す．

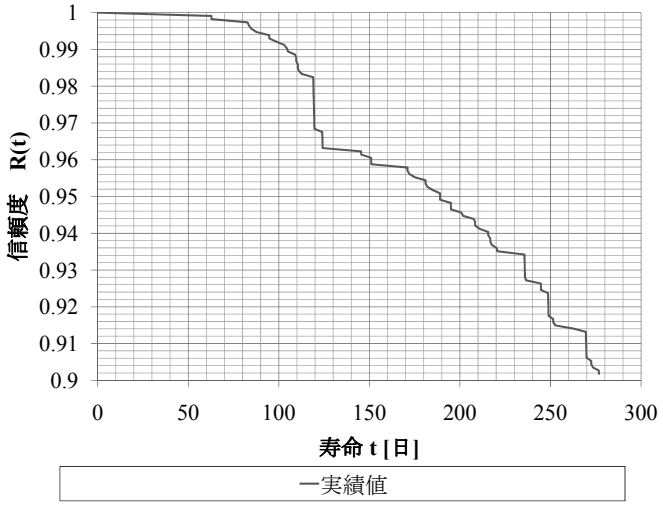


図 1. 故障実績による HDD 信頼度

図 1 において, 横軸は基準日からの経過日数, 縦軸は, 基準日に使用中の HDD に対する信頼度である．また，経過日数が 120 日周辺で，急激に故障割合が増加しているのは，2010 年 4 月 30 日より 5 月 6 日に実施したメンテナンス中に故障・縮退運転している HDD を発見し，交換したためである．

2.2 寿命分布へのあてはめ

図 1 の HDD 寿命分布にあてはめる予測モデルを選定する．代表的な故障寿命分布[5]として，指数分布，ワイブル分布，対数正規分布を選定した．寿命分布のパラメータ推定には，メンテナンス以後の故障実績を用いた．これは，メンテナンス以前の故障実績は，故障発生時期が不正確であることが想定されるためである．

2.2.1 指数分布へのあてはめ

指数分布は, 故障率 $\lambda(t)$が一定となり, 信頼度 R(t)が, 式(1)で表される分布である．

$$R(t)= \exp(-\lambda_0 t) \quad (1)$$

次に，故障率 λ_0 を推定する．指数分布関数は，横軸に t をとり，縦軸に ln R(t)をとったグラフ上で，傾きが故障率 λ_0 の値の直線となる．この性質を利用して，メンテナンス以後の故障実績をグラフ化し, 線形回帰分析により, 故障率 λ_0 を推定した．結果は，λ_0=0.0003 となった．

2.2.2 ワイブル分布へのあてはめ

ワイブル分布は，故障率が形状母数（Shape parameter）α，尺度母数（Scale parameter）β により変化し，信頼度 R(t)が，式(2)で表される分布である．

$$R(t)= \exp\left[-\left(\frac{t}{\beta}\right)^{\alpha}\right] \quad (\alpha>0, \beta>0) \quad (2)$$

故障率は，α>1 で増加，α＜1 で減少，α=1 の場合，一定となる．すなわち，α=1 のワイブル分布は指数分布と等価になる．

次に，α，β の値を推定する．ワイブル分布関数は，横軸に ln t をとり，縦軸に ln ln (1/R(t))をとったグラフ上で，傾きが α の直線となる．この時，F(t) =1-R(t) =1-(1/e) = 0.632 となる時の t が β となる．この性質を利用して，メンテナンス以後の故障実績をグラフ化し，線形回帰分析により，α，β を推定した．結果は，α=1.508，β=1315 となった．

2.2.3 対数正規分布へのあてはめ

対数正規分布は，寿命 t の対数が正規分布に従う時に現れ，信頼度 R(t)，故障率 $\lambda(t)$が，式(3)で表される分布である．

$$R(t)= \Phi\left(\frac{\mu_{Le}-\ln t}{\sigma_{Le}}\right) \quad (3)$$

ここで，$\Phi(u)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u}\exp\left(-\frac{u^2}{2}\right)du$

次に，μ_{Le}，σ_{Le} の値を推定する. 対数正規分布関数は，横軸に ln t をとり，縦軸に $\Phi^{-1}(1-R(t))$ をとったグラフ上で, 直線となる. この時, F(t)=0.5 に対する ln t が μ_{Le} に, F(t)= Φ(1)= 0.832，もしくは，F(t)=Φ(-1) =0.168 に対する ln t と μ_{Le} の差が σ_{Le} となる．この性質を利用して，メンテナンス以後の故障実績をグラフ化し，線形回帰分析により，μ_{Le}，σ_{Le} を推定した．結果は，μ_{Le}=7.433，σ_{Le}=1.301 となった．

2.3 適合度の評価

図 1 に対し，推定したパラメータを用い，指数分布，ワイブル分布, 対数正規分布の予測値を図示したものを，図 2 に示す.

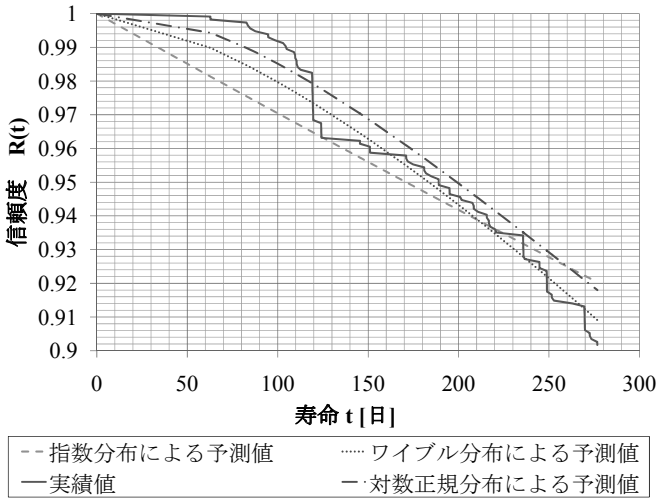


図 2. 各寿命分布による HDD 信頼度

次に，各寿命分布のいずれが最も適合度が高いかを，赤池の情報量基準（AIC）の最小値に基づき決定する[6]. 各寿命分布に対して，AIC は，式(4)で求められる.

$$AIC = n \times \left\{ \ln\left(2\pi \frac{S_e}{n} \right) + 1 \right\} + 2(p+2) \qquad (4)$$

ここで，n はサンプル数，S_e は残差平方和，p は説明変数の数であり，指数分布の場合には，p=1 とし，ワイブル分布，対数正規分布の場合には，p=2 を用いる．各寿命分布に対する AIC の結果を表 2 に示す.

表 2. 各寿命分布に対する AIC

寿命分布	AIC	推定された母数（再掲）
指数分布	-475.9	λ_0=0.0003
ワイブル分布	-591.7	α=1.508，β=1315
対数正規分布	-567.0	μ_{Le}=7.433，σ_{Le}=1.301

表 2 から, ワイブル分布が最もよく故障実績に適合し，対数正規分布がその次によく適合することが分かる.

この結果から，本論文では，HDD 寿命分布には，ワイブル分布へのあてはめを提案する．なお，ワイブル分布の定義より，故障率 $\lambda(t)$ は式(5)で表される.

$$\lambda(t) = \frac{\alpha t^{\alpha-1}}{\beta^{\alpha}} \qquad (5)$$

2.4 HDD 故障数の予測

クラウドコンピューティング環境を運用する上で，交換用の部品を調達する必要が頻繁に生じる. その際には，HDD 故障数を予測し，調達数を算出する必要がある．上記の結果を用いると，HDD 故障数を予測できる．式(6)に，予測式を示す.

$$HDD\text{故障数} = n \times \{ R(t_2) - R(t_1) \} \qquad (6)$$

ここで，n は HDD 総数，t_1，t_2 は，基準日から予測する期間の開始日，終了日までの日数であり，R(t)は式(2)を用いる．クラスタ#0,#1 であれば，α=1.508，β=1315 を代入することで，HDD 故障数が予測できる.

ただし，HDD 寿命分布がワイブル分布（α=0.7～0.8）に従うという報告[8]もあり，ワイブル係数については，HDD 製品で異なることが予想される．実務的には，直近 6 ヵ月～1 年間の故障実績にワイブル分布にあてはめて，母数を推定し, HDD の故障数を予測していくものと考えている.

3. 故障物理による故障率モデル

2 章では，寿命分布による故障率モデルを提案し，故障数の予測ができることを示した．寿命分布による故障率モデルは，統計的な手法を用いるため，個々の HDD に対して予測される故障率が低く, HDD を予防的に交換する運用には適さない. 一方, 実環境では, 温度・湿度, 振動・衝撃の条件で, HDD 毎に故障率が異なることが想定される．3 章では，このような使用条件も含む故障物理の観点から，故障率モデルの精度を向上させる.

3.1 速度過程モデル

速度過程モデル[5]は，寿命 t が式(7)で表され，電子部品の加速試験に広く用いられるモデルである.

$$t = BS^{-n} \cdot \exp\left(\frac{U}{kT} \right) \qquad (7)$$

ここで，S は負荷（応力や電圧），U は活性化エネルギー，k はボルツマン定数，T は絶対温度，B，n は定数である.

3.2 速度過程モデルへのあてはめ

クラスタ#0, #1 において，HDD の故障物理が速度過程モデルに従うかを検証する．そのためには，絶対温度 T 及び，負荷 S，寿命 t について仮定を置き，式(7)の関係が成り立つか確かめればよい.

3.2.1 絶対温度 T に関する仮定

クラスタ#0,#1 では, 各サーバの吸入温度が一定となるように空調制御しているため，温度一定と仮定する.

なお，式(7)において，温度一定と置くと，U，k，B は定数であるため，$B \cdot \exp(U/\ kT)$も定数となる．この定数を C と置くと，式(7)は式(8)で表される．

$$t = CS^{-n} \quad (\delta > 0) \tag{8}$$

さらに, 式(8)の両辺で対数をとると, 式(9)が得られる．

$$\ln t \propto -n \cdot \ln S \tag{9}$$

すなわち，温度一定の条件下では，式(9)が成り立つかを確かめればよい．

3.2.2 負荷 S に関する仮定

HDD に加えられる負荷 S（応力・電圧）には，応力としては振動・衝撃が，電圧としては過電圧がある．例えば，HDD に振動・衝撃が加えられると，メディアに傷が付き，メディア故障が発生する．代替セクタが割当てられれば，RAID カードにメディア故障数が記録され，代替セクタが割当てられなければ，HDD 故障となる．このように，負荷 S が加えられると，障害が発生し，障害から回復できれば，その結果が HDD の状態として記録され，障害から回復できなければ，故障となると考えられる．

このことから，負荷 S に関しては，RAID コントローラ経由で計測できる HDD の状態により近似でき, 式(10)に示すように, HDD 状態の間に線形関係が成り立つことを仮定する．

$$\text{負荷} S \propto HDD\text{状態} \tag{10}$$

3.2.3 寿命 t に関する仮定

また, 任意の負荷 S が加えられた HDD に対する寿命 t は，クラスタ内のすべての HDD が故障した時点では実測できるが，その時点では寿命を予測すること自体に意味がない．そこで，寿命 t に関しては，統計的な手法を用い，加えられた負荷 S がある区間に含まれる HDD を集団化し，その集団での HDD の信頼度 R(t)から求めた平均寿命 μ に等しいと仮定する．

頻度分布から密度関数を推定する場合と同様に，区間を広くすると，各 HDD の負荷 S のばらつきが大きくなり，負荷 S の値が不正確になる．また，区間を狭くすると，HDD 数が少なくなり，平均寿命 μ の値が不正確になる．すなわち，区間の幅を適切に設定すれば，この仮定は妥当性を持つと考えられる．

2 章で提案した故障率モデル（ワイブル分布）では，平均寿命 μ は，式(11)で表される[5].

$$\mu = \beta \cdot \Gamma\left(\frac{1}{\alpha} + 1\right) \tag{11}$$

ここで，Γ(x)は，ガンマ関数である．

ワイブル分布に従う電子部品の加速試験では，故障メカニズムが変化しないように α が一定になるように負荷を調整する[7]．クラスタ#0,#1 では，加速試験より弱い負荷しか加えていないため，α は一定と仮定する．

3.3 HDD 状態と平均寿命との相関関係

上記の仮定の下で，クラスタ#0,#1 の HDD の故障物理が速度過程モデルに従うかを確かめるには, HDD 状態を計測し，ある区間の代表値と，HDD 状態の値がその区間に含まれる HDD の平均寿命 μ の間に, 式(12)の関係が成り立つことを確かめればよい．

$$\ln \mu \propto -n \cdot \ln(HDD\text{状態の代表値}) \tag{12}$$

ここで，n は定数である．

3.3.1 HDD 状態の測定

まず，表 3，表 4 に示す測定条件により，HDD 状態を測定する．

表 3. HDD 状態の測定条件

対象クラスタ	クラスタ#0,#1
対象 HDD	基準日に稼働していた HDD（総数：1,140 台，故障数：83 台）
測定日	2010 年 9 月 1 日
測定値	測定項目（表 4 参照）毎に，直近に記録された値

表 4. HDD 状態の測定項目

項目名	説明
コマンド中断	RAID コントローラが発行したコマンドが中断された回数
メディア故障	HDD でメディア故障が発生し，代替セクタが割当てられた回数
接続失敗	RAID コントローラが HDD との接続に失敗した回数
パリティエラー	パリティエラーが発生した回数
ハードウェアエラー	ハードウェアエラーが発生した回数
S.M.A.R.T.警告	S.M.A.R.T.パラメータに警告が発生した回数

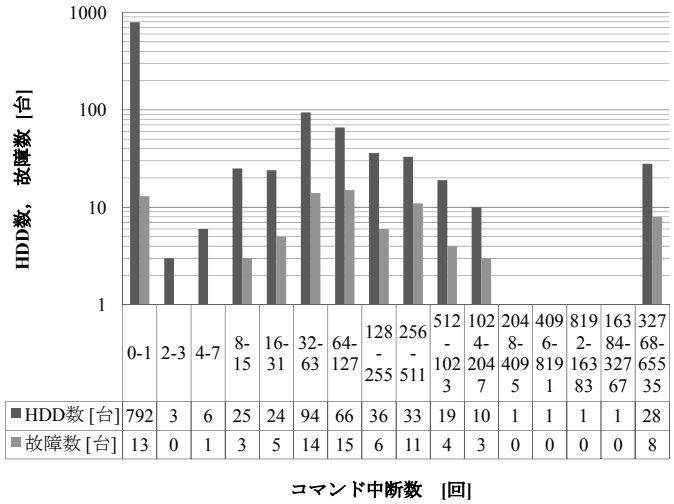
次に，測定結果より，HDD 状態と寿命の相関関係を確かめる．HDD 状態の値がある区間に含まれる HDD の頻度分布（以下，HDD 数）と，その中で故障した HDD の

頻度分布（以下，故障数）を求める．ここで，区間は，対数化を意識し，0～1，2^{n-1}～2^n-1（n は 2 以上の整数）とする．

さらに， HDD 数，故障数から，ある区間における信頼度 R(t)を求め，式(2)より β を，式(11)より平均寿命 μ を求める．ここで，α は一定との仮定のもと，2 章で求めた値（α=1.508）を用いる．

3.3.2 コマンド中断数と平均寿命との相関関係

図 3 に，コマンド中断数に対する HDD 数・故障数の頻度分布を示す．



	0-1	2-3	4-7	8-15	16-31	32-63	64-127	128-255	256-511	512-1023	1024-2047	2048-4095	4096-8191	8192-16383	16384-32767	32768-65535
HDD数 [台]	792	3	6	25	24	94	66	36	33	19	10	1	1	1	1	28
故障数 [台]	13	0	1	3	5	14	15	6	11	4	3	0	0	0	0	8

図 3. コマンド中断数と HDD 数・故障数の頻度分布

図 3 において，横軸は，コマンド中断数であり，縦軸は，コマンド中断数がその区間に含まれる HDD の頻度分布と，その中で故障した HDD の頻度分布を表している．

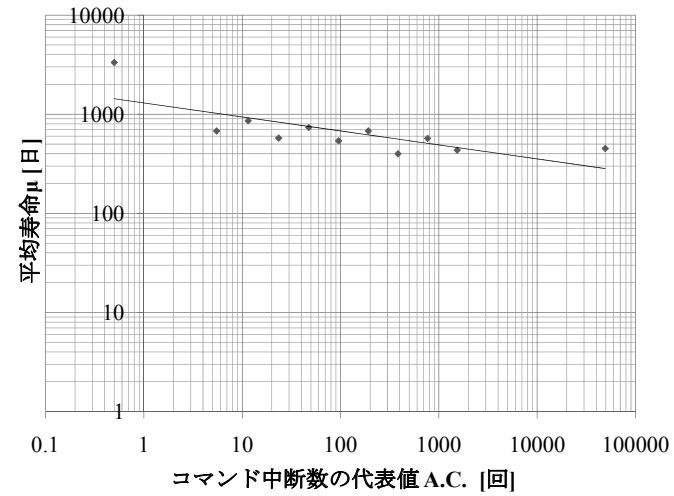


図 4. コマンド中断数と平均寿命との相関関係

次に，図 3 の結果から，平均寿命 μ を求める．コマンド中断数の区間の代表値 A.C.（Aborted Commands）と平均寿命 μ の相関関係を図 4 に示す．ここで，区間の代表値は，最小値と最大値の相加平均を利用する．例えば，最小値が 8，最大値が 15 という間隔の代表値は，11.5 である．

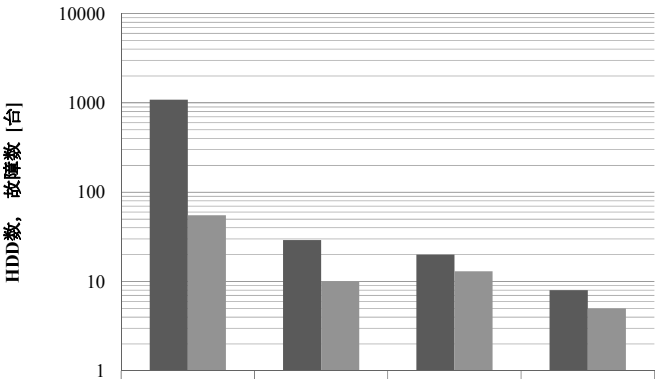
式(12)より，A.C.の対数と μ の対数には，負の線形関係が期待されるため，線形回帰分析を行なう．近似曲線として，式(13)が得られる．

$$\ln\mu = -0.1410 \times \ln A.C. + 7.169 \qquad (13)$$

なお，ピアソンの積率相関係数（以下，相関係数）は，R=-0.7555 である．

3.3.3 メディア故障数と平均寿命との相関関係

図 5 に，メディア故障数に対する HDD 数・故障数の頻度分布を示す．



	0-1	2-3	4-7	8-15
HDD数 [台]	1083	29	20	8
故障数 [台]	55	10	13	5

図 5. メディア故障数と HDD 数・故障数の頻度分布

図 5 において，横軸は，メディア故障数であり，縦軸は，メディア故障数がその区間に含まれる HDD の頻度分布と，その中で故障した HDD の頻度分布を表している．

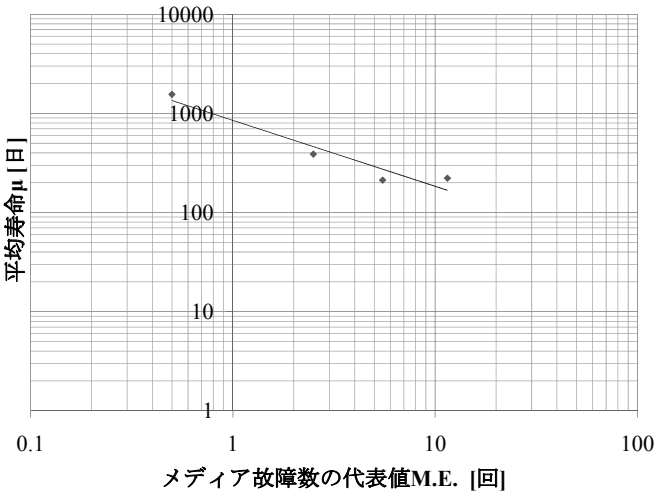


図 6. メディア故障数と平均寿命との相関関係

次に，図 5 の結果から，平均寿命 μ を求める．メディア故障数の区間の代表値 M.E.（Medium Errors）と平均寿命 μ の相関関係を図 6 に示す．

式(12)より，M.E.の対数と μ の対数には，負の線形関係が期待されるため，線形回帰分析を行なう．近似曲線として，式(14)が得られる．

$$\ln\mu = -0.6649 \times \ln M.E. + 6.744 \quad (14)$$

なお，相関係数は，R=-0.9616 である．

3.3.4 接続失敗数と平均寿命との相関関係

図 7 に，接続失敗数と HDD 数・故障数の頻度分布を示す．

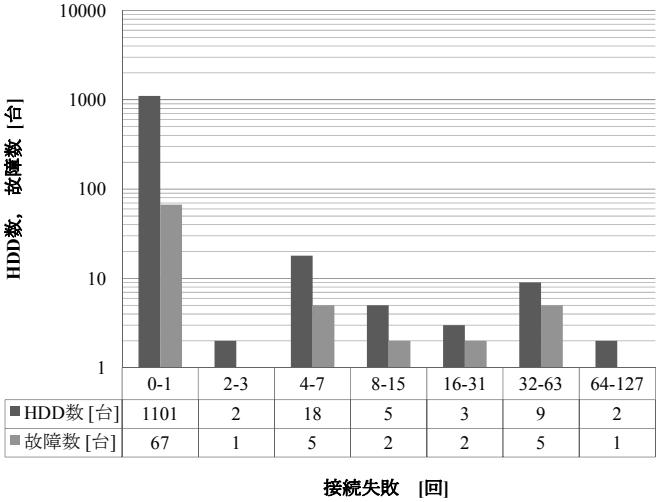


図 7. 接続失敗数に対する HDD 数・故障数の頻度分布

図 7 において，横軸は，接続失敗数であり，縦軸は，接続失敗数がその区間に含まれる HDD の頻度分布と，その中で故障した HDD の頻度分布を表している．

次に，図 7 の結果から，平均寿命 μ を求める．メディア故障数の区間の代表値 L.F.（Link Failures）と平均寿命 μ の相関関係を図 8 に示す．

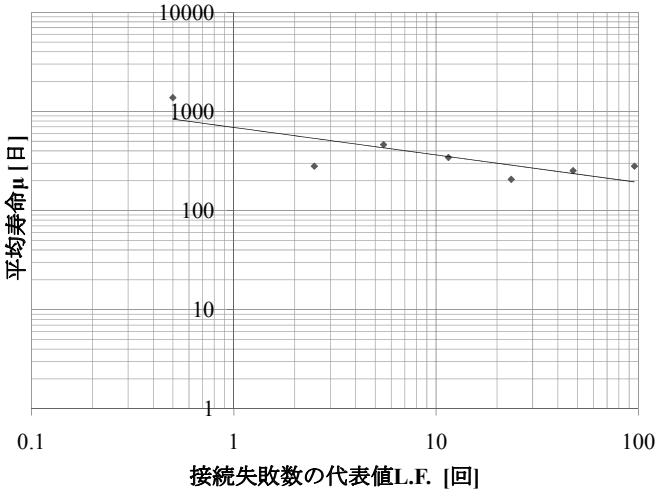


図 8. 接続失敗数と平均寿命との相関関係

式(12)より，L.F.の対数と μ の対数には，負の線形関係が期待されるため，線形回帰分析を行なう．近似曲線として，式(15)が得られる．

$$\ln\mu = -0.2776 \times \ln L.F. + 6.535 \quad (15)$$

なお，相関係数は，R=-0.7879 である．

3.3.5 パリティエラー数，ハードウェアエラー数，S.M.A.R.T.警告数と平均寿命との相関関係

パリティエラー数, ハードウェアエラー数, S.M.A.R.T. 警告数は，全ての HDD で 0 であるため，調査対象から外す．

3.4 評価

3.4.1 相関関係の検定

A.C.，M.E.，L.F.の対数と μ の対数に，有意な相関関係があるか，ピアソンの積率相関係数の有意性検定で確かめる．この有意性検定において，検定統計量 t_0 は式(16)で表される．

$$t_0 = \frac{|R|\sqrt{n-2}}{\sqrt{1-R^2}} \quad (16)$$

ここで，n は標本の大きさ（データの組数），R は標本相関係数である．t_0 が自由度 n-2 の t 分布に従うため，有意確率は $P = \Pr\{|t| \geq t_0\}$ として求められる．A.C., M.E., L.F.の対数と μ の対数の相関関係について，各変数を求めた結果を表 5 に示す．

表 5. 有意性検定の各変数値

検定対象	標本数 n	相関係数 R	検定統計量 t_0	有意確率 P
A.C.	11	-0.7555	3.460	0.03837
M.E.	4	-0.9616	4.957	0.007162
L.F.	7	-0.7879	2.861	0.03538

有意水準 5%（P≦0.05）として検定すると, A.C., M.E., L.F.の対数と μ の対数に，有意な相関関係がある．すなわち，3.2 節の仮定の下では，クラスタ#0,#1 の HDD の故障物理は，速度過程モデルに従う．

さらに, HDD の加速試験において速度過程モデルが用いられること[11]等から，HDD 一般の故障物理も速度過程モデルに従うと考えられる．この仮定の下で，A.C., M.E., L.F.を HDD 状態 H.S.（Hdd Status）と示すと，式(13)，式(14)，式(15)を一般化し，式(17)を導出できる．

$$\ln\mu = -n \times \ln H.S. + m \quad (17)$$

ここで，H.S.は，A.C.，M.E.，L.F.のいずれかの値で，

n, m は定数である. また, 式(17)の両辺で指数をとると, 式(18)が得られる.

$$\mu = \exp\{-n \times \ln H.S. + m\} \quad (18)$$

3.4.2 HDD 故障の診断

クラウドコンピューティング環境では，故障対応の稼働を極力減らしたいという要望は多い．この要望に応えるため，メンテナンス時に HDD の故障を診断し，予防的に交換する運用に対し，本論文で提案した手法が適用できる．

まず，対象とするクラスタ固有の故障実績から，3.3 節で述べた手法により, A.C., M.E., L.F.に対する定数 n, m を推定する．次に，3.4.1 項で述べた手法で，A.C.，M.E.，L.F.の対数と μ の対数に，有意な相関関係があるか推定する．A.C.，M.E.，L.F.のうち，有意性が認められるものから，1 つを選択する．ある HDD を交換すべきか判断するには，A.C.，M.E.，L.F.の値，及び，使用期間 t_3 を調査し, 式(18)から余命 $\mu - t_3$ を求める. 最終的に，この値がある基準値以内であれば，交換すべきと診断する．

例えば，クラスタ#0,#1 において，ある HDD（M.E.が 10 で，使用期間が 150 日）を交換すべきか診断することを考えてみる. 余命が 60 日以内であれば交換が必要とする. クラスタ#0,#1 の M.E.に関する n, m は, 式(14)より, n=0.6649, m=6.744 と分かる. 式(18)に, M.E.=10, n=0.6649, m=6.744 を代入すると，μ が 183.6 日と求まる．183.6 日から使用期間の 150 日を引くと，余命は 33.6 日となり，60 日以内のため，交換が必要と診断される．

4 関連研究

4.1 寿命分布による故障率モデル

実環境における HDD の故障実績を解析した研究は多くない．幾つかの研究は，HDD の年間交換率（ARR: Annual Replacement Rate）を報告している．ここで，年間交換率とは，HDD 総数に対し，1 年間で交換した HDD 数の比を百分率で示したものを意味する．Schwarz 他[9]は， HDD の年間交換率を，2%～6%と報告しており，Pinheiro 他[10]は， HDD の年間交換率を 1.7%～8.6%と報告している．

一方, HDD の寿命分布については, Schroeder 他[8]が, 70,000 台以上の HDD の故障実績を調査し，HDD 寿命分布が指数分布に従うという定説が誤りで，指数分布，ワイブル分布，ガンマ分布，対数正規分布では，ワイブル分布（形状母数 α=0.7～0.8）に良くあてはまることを報告している．

Scheroeder 他[8]の報告は, HDD 寿命分布がワイブル分布に従うという点では，本論文の結果と一致したが，適合度の評価は行なっていない．しかし，本論文では，赤池の情報量基準により，HDD 寿命分布には，ワイブル分布が最も良くあてはまることを確かめたことが，関連研究とは異なる点である．

4.2 故障物理による故障率モデル

HDD の故障物理に関する研究は, ベンダによる技術論文とユーザ経験に基づくものに大別できる[10]．前者の研究としては，SeaGate 社の Cole 他[11]が，速度過程モデルで負荷が一定と過程をおいたアレニウスモデル（Arrhenius Model）を根拠とし，HDD を運用する温度が 25℃から 42℃になると，平均故障間隔（MTBF: Mean Time Between Failures）が 50%以上短くなると報告している．なお，アレニウスモデルに負荷 S のパラメータを追加したモデルが，速度過程モデル，もしくは，アイリングモデル（Eyring Model）である. また, NetApp 社の Elerath と Shah[12]は，アレニウスモデルによる加速試験について言及し，現場での MTBF は，HDD ベンダが期待するほど，高くはないことを指摘している．

一方，後者の研究としては，S.M.A.R.T.パラメータと故障統計との関連を調べた研究が幾つかある．Hughes 他[13] は， S.M.A.R.T. （ Self-Monitoring, Analysis and Reporting Technology）アルゴリズムを改良し，3,744 台の HDD に対して適用した結果，3～4 倍の精度で故障を予測できることを報告している．また，Google 社の Pinheiro 他[10]は，HDD の S.M.A.R.T. パラメータ，温度と故障との相関を調査している．故障と高い相関を示す S.M.A.R.T.パラメータもあるが，HDD 故障の予測には有効ではないと結論付け，温度と故障の相関は，以前の報告に比べ，少ないと報告している．

Cole 他[11]と Elerath 他[12]の報告は, HDD の故障物理として，速度過程モデルを根拠にしている点が，本論文と一致している．しかし，これらの報告が，負荷 S を一定にして，絶対温度 T が寿命 t に与える影響を調査しているのに対し，本論文では，HDD の状態と寿命 t との相関を明らかにするために，絶対温度 T を一定にして，負荷 S が寿命 t に与える影響を調査している．

また，Pinheiro 他[10]の報告は，広い意味で HDD の状態を用いて故障を予測している点が，本論文と一致している．しかし，この報告が，HDD 自体に保持される S.M.A.R.T.パラメータを用いるのに対し，本論文では，S.M.A.R.T.パラメータ非対応の HDD を対象とする等, 適

用領域を広げるために，RAID コントローラが保持する HDD の状態を用いている．

5. おわりに

本論文では，我々のクラウドコンピューティング環境において，特に HDD 故障が多い 2 クラスタ（クラスタ #0,#1）の事例について，HDD 故障の実績を調査した．故障実績から HDD の寿命分布を求め，代表的な寿命分布をあてはめた．赤池の情報量基準により，各寿命分布への適合度を評価し，ワイブル分布（形状母数 α=1.508）が最も良くあてはまることを示した．本論文では，HDD 寿命分布として，ワイブル分布へのあてはめを提案し，交換用 HDD の調達数の見積りに適用できることを述べた．

次に，同じ故障実績より，RAID コントローラが出力するコマンド中断数，メディア故障数，接続失敗数と平均寿命の相関関係を調べ，ピアソンの積率相関係数の有意性検定により，有意な相関関係があることを示した．さらに，HDD の故障物理が，代表的な故障モデルである速度過程モデルに従うことを示した．本論文では，HDD の故障物理へ速度過程モデルをあてはめることを提案し，予防的に HDD を交換する運用に適用できることを述べた．

最後に，寿命分布，故障物理による HDD 故障率モデルという視点から，関連研究について述べた．さらに，本論文が提案した 2 手法が，関連研究に対し，どのような差異化点があるかを指摘した．

なお，本論文で提案した 2 手法を実環境に適用するには，事前に故障実績を測定する必要があるため，実環境を運用開始した時点では利用できない．この点は，今後の課題としたい．また，メモリ故障には，HDD 故障と同様の課題があり，今後調査し報告したい．

謝辞 本論文の作成にあたり，貴重なコメントを頂きました，NTT サービスインテグレーション基盤研究所の船越裕介様，高橋玲様に深謝します．

参考文献

1) 後藤厚宏, 西原琢夫, クラウドコンピューティング技術 CBoC の構想, NTT 技術ジャーナル, vol.21, No.9, pp.64-69 (2009).
2) 佐藤賢一，横関大子郎，新井克也: クラウドコンピューティング技術 CBoC の研究開発, NTT 技術ジャーナル, vol.21, No.9, pp.70-74 (2009).
3) 白石正裕，菅沼毅，宮田俊介: CBoC の仮想化運用管理技術による SaaS の実現, NTT 技術ジャーナル, vol.21, No.9, pp.75-79 (2009).
4) 高倉健，空一弘，天海良治，鷲坂光一，冨田清次: CBoC による大規模分散処理システムの実現, vol.21, No.9, pp.75-79 (2009).
5) 市川昌弘: 信頼性工学 (機械工学選書), 裳華房, ISBN4-7853-6506-6 (1990).
6) J.P.クライン, M.L.メシュベルガー著（打波守訳）: 生存時間解析，シュプリンガー・ジャパン, ISBN978-4-431-10071-3 (2009).
7) 鈴木秀人: 実用電子部品の疲労信頼性評価, (株)リアライズ社, ISBN4-947655-91-7 (1996).
8) Bianca Schroeder and Garth A. Gibson: Disk Failures in the Real World: What Does an MTTF of 1,000,000 Hours Mean to You?, Proceedings of the 5th USENIX Conference on File and Storage Technologies (FAST'07), pp.1-16, February 2007.
9) T. Schwarz, M. Baker, S. Bassi, B. Baumgart, W. Flagg, C. van Ingen, K. Joste, M. Manasse, and M. Shah. Disk failure investigations at the internet archive. In Work-in-Progess session, NASA/IEEE Conference on Mass Storage Systems and Technologies (MSST2006), 2006.
10) Eduardo Pinheiro, Wolf-Dietrich Weber and Luiz Andr´e Barroso: Failure Trends in a Large Disk Drive Population, Proceedings of the 5th USENIX Conference on File and Storage Technologies (FAST'07), February 2007.
11) Gerry Cole. Estimating drive reliability in desktop computers and consumer electronics systems. Seagate Technology Paper TP-338.1, November 2000.
12) Jon G. Elerath and Sandeep Shah. Server class disk drives: How reliable are they? In Proceedings of the Annual Symposium on Reliability and Maintainability, pp.151-156, January 2004.
13) Gordon F. Hughes, Joseph F. Murray, Kenneth Kreutz-Delgado, and Charles Elkan. Improved disk-drive failure warnings. IEEE Transactions on Reliability, 51(3):350-357, September 2002.

樋渡 仁（正会員）
NTT 情報流通プラットフォーム研究所 IT アーキテクチャプロジェクト 研究主任．Web 検索，クラウドコンピューティングの研究開発に従事. 1995 年 早稲田大学大学院修士課程修了.
岩村 相哲（正会員）
NTT 情報流通プラットフォーム研究所 IT アーキテクチャプロジェクト 主幹研究員．モバイル通信，クラウド技術の研究開発に従事. 1994 年 東京大学大学院博士課程修了.
新井 克也（正会員）
NTT 情報流通プラットフォーム研究所 IT アーキテクチャプロジェクト 主幹研究員．並列処理ソフトウェア，分散処理ソフトウェア，Web サービスの研究開発に従事. 1988 年 電気通信大学大学院修士課程修了．情報処理学会，ACM，IEEE 各会員.

投稿受付：2011 年 2 月 15 日
採録決定：2011 年 6 月 3 日
編集担当：東野 輝夫（大阪大学）