Vol.2015-ARC-215 No.9
Vol.2015-OS-133 No.9
2015/5/27

Markov 連鎖を用いたデータセンタ High Availability システムの信頼性評価方法

川口智大[†1]　斎藤秀雄[†2]

基幹業務では IT システム停止が莫大なビジネス損失に繋がる。災害に伴うシステム停止防止のため,2 拠点間データ二重化と,別拠点のクォーラムを利用したデータセンタ HA(High Availability)システムが利用されている。クォーラムはデータ格納やアプリケーション実行に寄与しない IT 機器であり,導入コスト削減のため安価な機器を利用したいとの要望が強い。本研究ではデータセンタ HA 環境での障害発生時の動作をモデル化し,クォーラム信頼性と業務継続期待時間の相関について Markov 連鎖を利用して導出した。震災を想定した場合,平均寿命 3 年の機器をクォーラムに利用しても,障害後 3 日以内に交換する運用を行う事で,高信頼装置利用時と比較した業務継続期待時間の低下を 8.4%程度に抑制可能であることが分かった。

1. 序論

エンタープライズシステムは，企業がビジネスを遂行するために不可欠な業務を処理するために用いられる。このようなシステムの多くでは，地震・火災・水没といったデータセンタそのものが被災しシステム停止に陥り，業務停止となることを防ぐため，データセンタ間でサーバやストレージを二重化し，障害発生時に自動的に故障した装置を切り離して正常な装置で処理を継続するデータセンタ High Availability (以下 HA と略記)システムが広く使われてきた。近年では，システム全体でのリソース有効活用や負荷分散を目的として，何れのデータセンタでも業務を可能とするために，二重化されたデータ両方に対して同時にアクセス可能な Active-Active 型 HA システムも利用される。

一般に，複数の計算機ノードを相互接続してクラスタを構成すると，相互接続ネットワーク障害時にスプリットブレイン問題が発生し，分断されたノードで同一のサービスが起動してしまい，両者間でデータ不一致が発生する。これを防止するために，何れのノードを生存させるかを調停するクォーラムが利用される。

クォーラムは業務そのものには寄与せず，限られた障害ケースにおいて不整合を防止するための機構であるため，極力安価な IT 機器を利用したいとユーザは考える。しかし，このような安価な信頼性の低い機器を利用した場合，HA システムの信頼性へどの程度影響を与えるか不明であった。

本稿では, HA システムの Markov 連鎖を用いた信頼性評価方法，及び地震を想定した場合でのクォーラムの信頼性への影響の検証結果について述べる。

2. 関連研究

本研究は，大規模システムのシステム継続性・信頼性の評価方法に関するものである。関連する研究として，システムの継続性の評価方法に関しては RAID (Redundant Array of Inexpensive Disks)のシステム継続性を評価した Greenan[1]が挙げられる。Greenan は多数のドライブを搭載したストレージシステムが RAID を構成する場合に，個々のドライブの故障率，データ読み出し不可の発生確率，データを回復するまでに要する期待時間より Markov 連鎖を利用して MTTDL (Mean Time To Data Loss) を導出する方法について述べている。また，震災などの広域災害に備えて，データを遠く離れた安全な拠点との間で二重化し，システム継続性を高める研究を Matsumoto[2]が行っている。Matsumoto は多数のデータ格納拠点がシステムに存在する場合，数理選択モデルを利用して同時に被災する可能性の低い拠点間でデータ複製することでシステム継続性を向上する方法について述べている。

3. HA システム

3.1 概要

HA システムは IT 業務を実施する(A)プライマリデータセンタ･(B)セカンダリデータセンタ, A と B を接続する(X)広域ネットワーク, X 障害時に A と B 何れかを生存させるかを判定する(Q)クォーラム，A と Q を接続する(Y)広域ネットワーク及び, B と Q を接続する(Z)広域ネットワークより構成される。

A・B それぞれはサーバ・ストレージ・データセンタ内ネットワークより構成されており，内部機器の単一障害ではシステム停止とならない程度の冗長性を持つ。

A・B に配置されたストレージは何れか一方にデータ更新が発生した場合，X を介して他方に同期更新を行う。また，A・B のストレージは定期的に Y・Z を介して各自の生存情報を Q に通知を行う。

尚，A・B・Q は同一の災害で同時に停止とならないように拠点間の距離や地形が選別される。また，A・B 間のデ

†1 (株)日立製作所
Hitachi Ltd.
†2 日立アメリカ社
Hitachi America, Ltd.

ータ同期に要する遅延時間が A・B 上で実行されるアプリケーションに影響ないように A・B 間の距離や X の通信方法が選択される。Y・Z 距離や通信方法は Q への生存情報の通知周期に影響しない程度の遅延時間内に収まるように選択される。

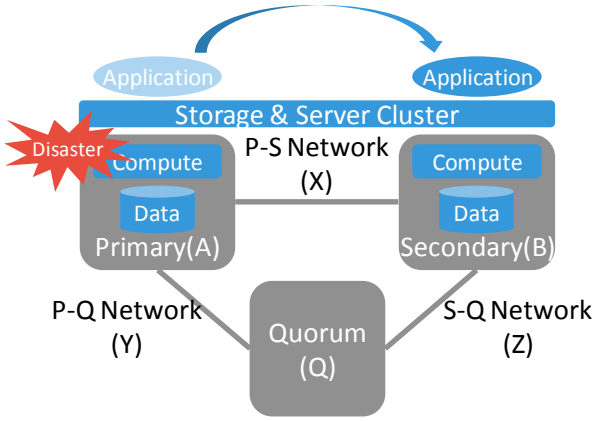


図 1 HA システム概要

A・B のデータセンタでは，A・B 間のストレージ装置間でデータを二重化するストレージクラスタ機能及び，A・B のデータセンタに障害が発生した場合に他方のデータセンタにアプリケーションを切り替えるサーバクラスタ機能が動作する。もしも A で動作するアプリケーションがあり，A が障害となった場合，B のサーバクラスタ機能により業務は B に引継がれるのが期待である。A が障害となる直前まで A・B 間でデータを同期しているため，業務無停止が発生した場合でもアプリケーションを引継ぐことが可能である。しかし，これを自動化しようとした場合，スプリットブレイン問題が発生する。B はネットワークの不通を契機に，B へアプリケーションを引継ごうとするが，ネットワーク不通の情報だけでは A 障害か X 障害か判別できない。この問題に対処するため，定期的に A より Q へ送付される生存情報に更新がないことを併せてチェックすることで，A 障害であると判断する。

3.2 他の障害

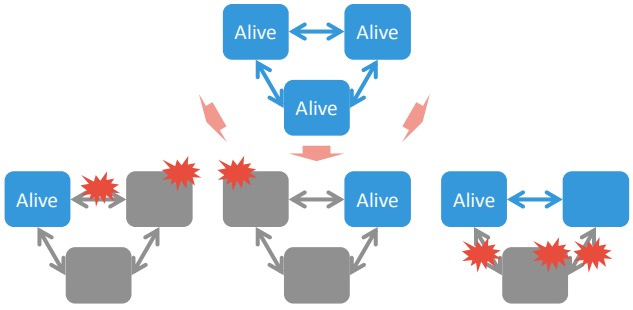


図 2 起こりうる障害及び，その対処法

3.1 では A 障害となるケースを示したが，障害は B・Q・X・Y・Z にも発生し得る。このような場合でも，少なくとも A・B の何れか一方で業務継続できるようにシステムは制御を行う。

4. HA システムのモデル化

4.1 コンポーネントの状態遷移モデル

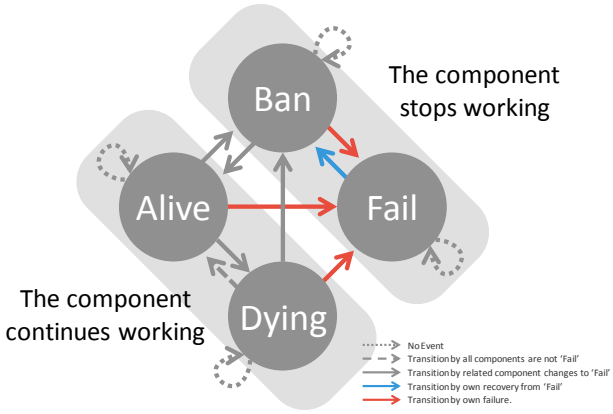


図 3 コンポーネントの状態遷移

このような HA システムの状態についてモデルを構築する。システム全体の状態遷移モデル考案に先立って，各コンポーネントの状態遷移モデルを考案する。

A・B・Q・X・Y・Z は Alive・Ban・Dying・Fail の何れかの状態を取るとする。Alive では対象コンポーネントが正常稼動状態であり HA システムに関与している，Ban では対象コンポーネントは正常稼動状態であるが HA システムに関与していない(e.g. B 障害後の Q はそれ以上調停する必要はなくなる)，Dying では HA となっているが障害によってはシステム停止となる(e.g. Q 障害後に A 障害となった場合，調停が働かないのでシステムは停止する。但し Q 障害後の B 障害は A を優先的に生存させるためシステムは継続する)，Fail では対象コンポーネントに障害が発生している，ことを意味する。

Fail 状態へは Alive・Ban・Dying 全てから状態遷移可能であり，Dying や Ban は他のコンポーネントの障害を契機に Alive から状態遷移する(表 1)。また，Alive への状態遷移は他のコンポーネント全てが Fail でなくなったことを契機に状態遷移する。

Vol.2015-ARC-215 No.9
Vol.2015-OS-133 No.9
2015/5/27

表 1 Fail 化に伴って発生する他コンポーネント状態遷移

		Fail 遷移に伴い発生する状態遷移とコンポーネント		
		Alive →Ban	Alive →Dying	Dying →Ban
Fail となるコンポーネント	A	Q, X, Y, Z	-	B, X
	B	Q, X, Y, Z	-	X
	Q	Y, Z	B, X	-
	X	B, Q, Y, Z	-	B
	Y	Q, Z	B, X	-
	Z	Q, Y	B, X	-

4.2 システムの状態遷移表現

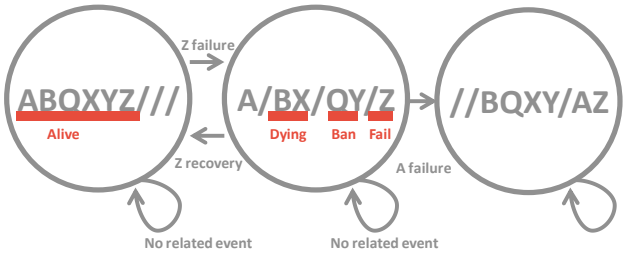


図 4 システムの状態遷移と表記方法

システム全体の状態は，4.1 で利用した各コンポーネントの状態を利用して表記する。「Alive コンポーネント / Dying コンポーネント / Ban コンポーネント / Fail コンポーネント」と各状態のコンポーネントをリストアップし，「/(スラッシュ)」を区切り文字として一列に記載する。

例えば，全コンポーネントが正常状態である環境(初期状態)は「ABQXYZ///」と記載できる。Z に障害が発生した場合，表 1 に基づき，Q・Y が Ban に，B・X は Dying に遷移するため，「A/BX/QY/Z」状態となる。更に A に障害が発生すると「//BQXY/AZ」となる。このとき，Alive 状態であるコンポーネントがなくなる状態となり，システムは停止する。また，「A/BX/QY/Z」状態から Z の障害が回復すると，Fail 状態のコンポーネントがなくなるため，システムは初期状態「ABQXYZ///」に戻る。

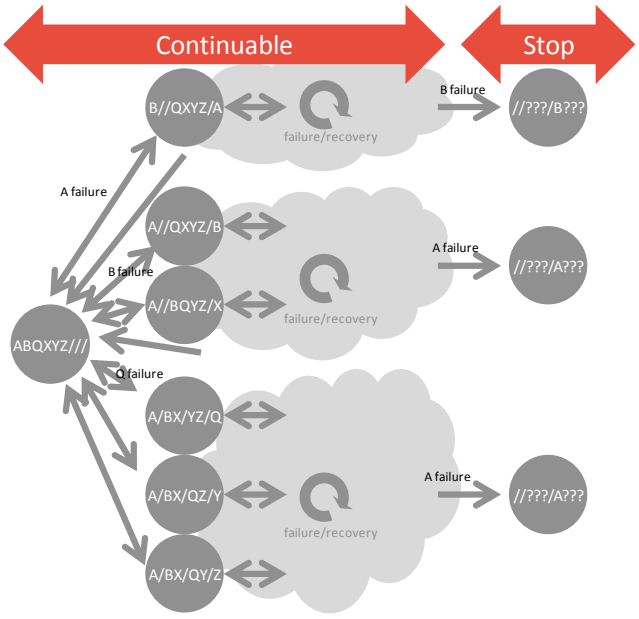


図 5 システム状態遷移

この状態遷移表現を用いて，システム停止とならず，且つ，コンポーネントの障害・回復のイベント発生によって起こりうる状態は全部で 70 通りである。

表 2 システム停止とならない状態一覧

ABQXYZ///	A///BQXYZ	A//Y/BQXZ	A//X/BQYZ	A//B/QXYZ
A//Z/BQXY	A//Q/BXYZ	A//XY/BQZ	A//BY/QXZ	A//YZ/BQX
A//QY/BXZ	A//BX/QYZ	A//XZ/BQY	A//QX/BYZ	A//BZ/QXY
A//BQ/XYZ	A//QZ/BXY	A//BXY/QZ	A//XYZ/BQ	A//QXY/BZ
A//BYZ/QX	A//BQY/XZ	A//QYZ/BX	A//BXZ/QY	A//BQX/YZ
A//QXZ/BY	A//BQZ/XY	A//BXYZ/Q	A//BQXY/Z	A//QXYZ/B
A//BQYZ/X	A//BQXZ/Y	B///AQXYZ	B//Y/AQXZ	B//X/AQYZ
B//Z/AQXY	B//A/QXYZ	B//Q/AXYZ	B//XY/AQZ	B//YZ/AQX
B//AY/QXZ	B//QY/AXZ	B//XZ/AQY	B//AX/QYZ	B//QX/AYZ
B//AZ/QXY	B//QZ/AXY	B//AQ/XYZ	B//XYZ/AQ	B//AXY/QZ
B//QXY/AZ	B//AYZ/QX	B//QYZ/AX	B//AQY/XZ	B//AXZ/QY
B//QXZ/AY	B//AQX/YZ	B//AQZ/XY	B//AXYZ/Q	B//QXYZ/A
B//AQXY/Z	B//AQYZ/X	B//AQXZ/Y	A/BX//QYZ	A/BX/Y/QZ
A/BX/Z/QY	A/BX/Q/YZ	A/BX/YZ/Q	A/BX/QY/Z	A/BX/QZ/Y

5. Markov 連鎖を用いた HA システムの信頼性評価法

5.1 吸収的 Markov 連鎖を用いた定式化

本項では一般的な吸収的 Markov 連鎖を用いた平均吸収時間の導出方法について説明する。

一般に Markov 連鎖では，単位時間当たりの状態遷移確率Pは，システム稼動継続確率(一時的状態に留まる確率)Q，システム停止確率(吸収状態に陥る確率)U，基本行列Iを用いて以下のように表現することができる。

Vol.2015-ARC-215 No.9
Vol.2015-OS-133 No.9
2015/5/27

$$P = \begin{pmatrix} Q & U \\ 0 & I \end{pmatrix}$$

時刻nにおける状態分布σ_nが与えられた場合，次ステップにおける状態分布σ_{n+1}はPを用いて以下のように表現する。

$$\sigma_{n+1} = \sigma_n P$$

P が時間に依らず一定であるならば，初期状態σ_0から時間n経過後の状態分布σ_nは下式となる。

$$\sigma_n = \sigma_0 P^n$$

当式の極限を求めることで，各状態への吸収確率MUを求めることが可能である。

$$\lim_{n\to\infty} \sigma_n = \sigma_0 \lim_{n\to\infty} P^n$$

$$\lim_{n\to\infty} P^n = \begin{pmatrix} Q & U \\ 0 & I \end{pmatrix}\begin{pmatrix} Q & U \\ 0 & I \end{pmatrix}\cdots = \begin{pmatrix} 0 & MU \\ 0 & I \end{pmatrix}$$

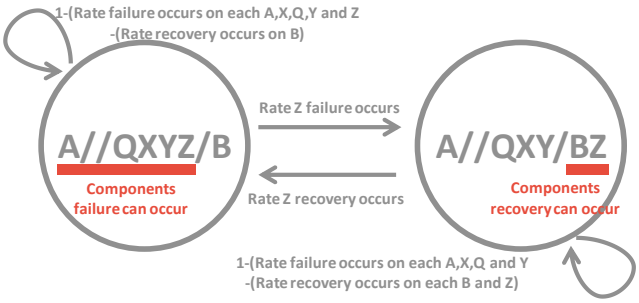
平均訪問回数を要素に持つ行列MはQを用いて以下のように計算可能である。

$$M \equiv I + Q + Q^2 + \cdots = (I - Q)^{-1}$$

任意状態iからスタートし，何れかの吸収状態に陥るまでの時間(平均システム稼働時間)は，要素が 1 の列ベクトルgを用いて，Mgを求めることで計算可能である。

行列Qはm次の正方行列であり，mはシステム継続可能である状態数に等しい。4.2 にてm = 70であり，それほど大きくないことが判明しているため，現実的な計算時間でMgを求めることが可能である。また，Mの導出には高次の逆行列計算が必要であり、比較的長い計算時間を必要とするが，M の値が不要であるならば$(I - Q)^{-1}$を求めず，$(I - Q)x = g$と連立方程式化してxを算出することで，効率的に計算することが可能である。

5.2 状態遷移確率の計算方法



システムの状態遷移は，コンポーネントの障害・回復により起きるとする。障害は Alive・Dying・Ban 何れかであるコンポーネントに発生し，回復は Fail であるコンポーネントに発生する。また，離散時間間隔は短時間であり，複数のコンポーネントに障害・回復が同時に発生する確率はきわめて低いと仮定し，状態遷移が起き得るコンポーネントは 1 個のみとする。

この時，状態xから状態yに遷移する状態遷移確率R_{xy}は以下のように計算できる。

- xとyを比較し Fail から遷移した，若しくは Fail へ遷移したコンポーネントが 2 個以上ある場合には，$R_{xy} \approx 0$
- xとyを比較し，Fail に遷移したコンポーネントが(1個のみ)ある場合には，対象コンポーネントαの障害確率f_αを用いて$R_{xy} = f_\alpha$
- xとyを比較し，Fail から遷移したコンポーネントが(1個のみ)ある場合には，対象コンポーネントαの回復確率r_αを用いて$R_{xy} = r_\alpha$
- x = yならば Active(Alive・Dying・Ban 何れか)の状態にある全コンポーネントの障害発生確率の総和$\sum_i^{Active} f_i$と，Fail 状態にある全コンポーネントの回復発生確率の総和$\sum_j^{Failed} r_j$を用いて$R_{xy} \approx 1 - \sum_i^{Active} f_i - \sum_j^{Failed} r_j$

6. 信頼性評価

6.1 震災を想定した場合の評価パラメタ

5.2 で導出した式を利用して震災を想定した場合の信頼性を評価する。一般にデータセンタや広域ネットワークは冗長構成を採るため，信頼性・可用性が高く，IT 機器の単一故障程度ではコンポーネント障害とみなされない。一方でクォーラムには一般的なサーバやストレージ装置が利用されるため，対象の機器が故障した場合でも障害発生となる。

このとき，以下に記す事実より得られた情報をパラメタとして使用する。

- 日本においては 20～200 年周期でマグニチュード 7 以上の地震が各地で発生している[3]。内閣府資料の南海トラフ巨大地震の発生確率予測によると，前回地震発生から 50 年経過時点での以後 30 年間以内の地震発生確率は 45～50%と見積もられている[4]。また，南海トラフのみならず日本全国各地で高い確率で大地震が発生する可能性が示唆されている[5]。
- 東日本大震災の際には SINET 東北大ノードは 96 時間で復旧した[6]。
- データセンタの構築において，企画からシステム稼動開始までに要する期間は 102～156 週間(平均値 129 週間)である[7]。

ある事象の 1 日当たり平均発生確率μ_{Daily}と，その事象の長期発生確率$\mu_{LongTerm}$及び期間dには下式の関係がある。

$$1 - \mu_{LongTerm} = (1 - \mu_{Daily})^d$$

これより$\mu_{Daily} = 1 - (1 - \mu_{LongTerm})^{1/d}$が導出される。当式を利用して求めた，コンポーネントαの 1 日当たり障害発生確率f_αを表 3 に，回復発生確率r_αを表 4 に記す。

表 3 想定する障害発生確率

Component	主となる障害要因	障害発生率	1 日当たり障害発生率
A	地震	50%@50 年	3.80E-05
B	地震	50%@50 年	3.80E-05
Q	サーバ障害	サーバの信頼性に依存	
X	地震	50%@50 年	3.80E-05
Y	地震	50%@50 年	3.80E-05
Z	地震	50%@50 年	3.80E-05

表 4 想定する回復発生確率

Component	回復方法	障害発生率	1 日当たり障害発生率
A	データセンタ再建	50%@129 週	7.67E-04
B	データセンタ再建	50%@129 週	7.67E-04
Q	サーバ交換	運用管理方法に依存	
X	ネットワークノード交換	50%@96h	1.59E-01
Y	ネットワークノード交換	50%@96h	1.59E-01
Z	ネットワークノード交換	50%@96h	1.59E-01

6.2 クォーラムの故障時間・回復時間とシステム信頼性の相関

クォーラムに使用する機器の平均寿命及び，機器故障から交換が完了するまでに要する平均時間を変数として，震災を想定としたケースでのシステム信頼性の計算結果を図 6 と図 7 に記す。

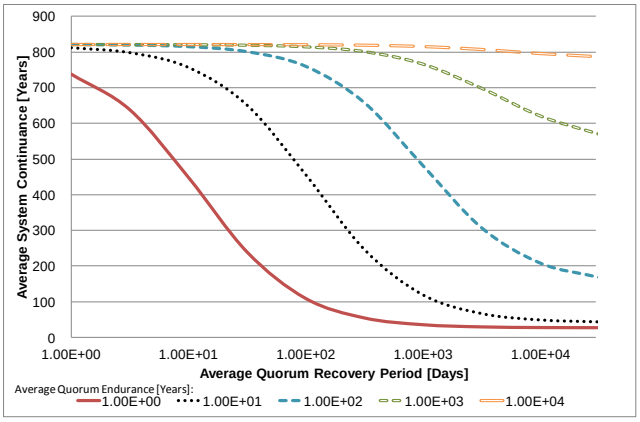


図 6 クォーラム障害確率・回復時間と HA システム信頼性の相関 1

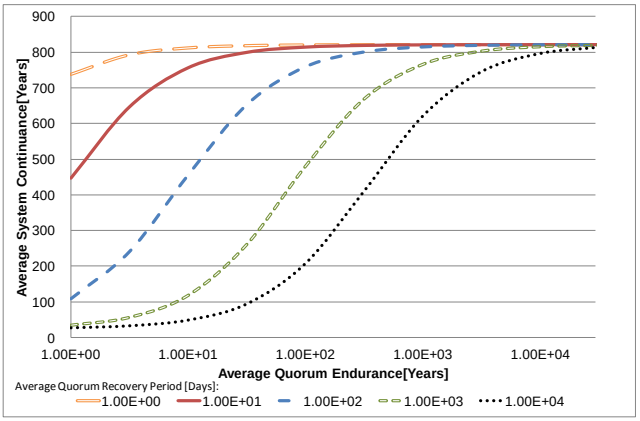


図 7 クォーラム障害確率・回復時間と HA システム信頼性の相関 2

6.3 考察

6.2 の計算結果から，以下の事実が分かった。

i. データセンタ HA システムにおいて，システム稼動平均時間の限界は約 821 年である。

ii. 高信頼(寿命の長い)装置をクォーラムとして使用することでシステム稼動平均時間を限界に近づけることができるが，クォーラム故障時の回復(交換)を早期に行わないならば，システム稼動平均時間が著しく低下する。

iii. 低信頼(寿命の短い)装置をクォーラムとして使用する場合であっても，クォーラム故障時の回復(交換)を早期に行う運用を採るならばシステム稼動平均時間の限界に近づけることができる。例えば，平均寿命 3 年の装置をクォーラムに使用し，故障時には 3 日以内に回復する運用であるならばシステム稼動平均時間を約 751 年とすることが可能である。(限界からの低下量は約 8.4%)

特に iii を考慮すると，パブリッククラウドにクォーラム機能を配置するアイデアが提案できる。一般にパブリッククラウドは CPU 不可やアクセス量といった利用量に応じて課金されるため，定期的に生存情報を受信するのみのクォーラムでの利用ならば安価に運用可能である。また，パブリッククラウド内には多数のサーバ・ストレージをシステムに持ち，サービスを提供しているため，一部の機器に故障が発生しても代替機器を即座に用意することができる。

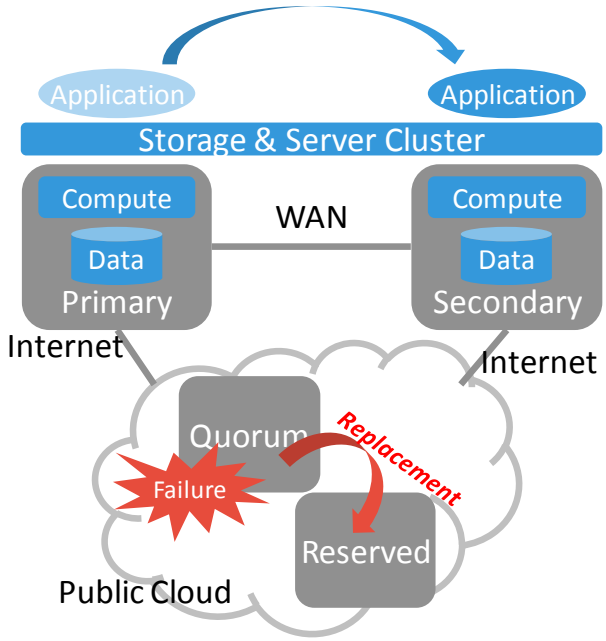


図 8 パブリッククラウドを利用したクォーラム

7. まとめと今後の課題

7.1 まとめ

本稿では，HA システムに関して，コンポーネントの障害及び回復に伴う状態遷移モデルを構築し，その状態遷移モデルに基づいて Markov 連鎖を適用することで HA システムの信頼性評価を定式化した。加えて，震災を想定した条件でクォーラムの障害発生確率及び障害からの回復期間と HA システムの信頼性の相関について調査した。その結果，低信頼の装置をクォーラムとして使用する場合であってもクォーラム故障時の回復(交換)を早期に行う運用を採るならばシステム稼動平均時間の限界に近づけることができることがわかった。

7.2 今後の課題

今回の計算の想定では，地震が発生する確率は前回の地震発生からの経過時間に依らず一定であると仮定した。実際は地震発生からの経過と共に確率が増加する。このような時間変動を加味した計算モデルの確立が重要である。

また，災害は地震だけでなく火災・水害(津波・洪水)・人災(操作ミス)も含まれる。このような各種災害の被災率も考慮したパラメタを用いた計算を行うべきである。

参考文献

1) K. M. Greenan, J. S. Plank, Jay. J. Wylie: “Mean time to meaningless: MTTDL, Markov models, and storage system reliability,” 2nd USENIX Workshop on Hot Topics in Storage and File Systems, pp. 1-5, 2010
2) T. N. H. M. Shinya Matsumoto: “Risk-aware Data Replication to Massively Multi-sites against Widespread Disasters,” ACIS, 2013
3) 地震調査研究推進本部, “活断層及び海溝型地震の長期評価結果一覧,” 2015
4) 地震調査研究推進本部地震調査委員会, “南海トラフの地震活動の長期評価（第二版）について,” 2013
5) 地震調査研究推進本部地震調査委員会, “全国地震動予測地図２０１４年版～全国の地震動ハザードを概観して～,” : http://www.jishin.go.jp/main/chousa/14_yosokuchizu/index.htm
6) 学術情報ネットワーク, “SINET4 の概要,” 2011. http://www.sinet.ad.jp/storage/cloud.pdf
7) Roel W. de Both: Master Thesis, Master of Science in Business Administration, University of Twente

正誤表

大変申し訳ございません。お詫びの上、訂正させていただきます。

1 ページ 脚注 3 行目
(誤)「日立アメリカ社」
(正)「(株)日立製作所」

1 ページ 脚注 4 行目
(誤)「Hitachi America, Ltd.」
(正)「Hitachi, Ltd.」