Vol.2010-SE-167 No.31
2010/3/19

サーバ仮想化技術を利用した障害試験自動化手法の提案

花崎 芳彦[†]　伊藤 孝之[†]　國分 俊介[†]
片山 吉章[†]　岡部 亮[†]　飯塚 剛[†]

高信頼なシステムを実現するためには，十分な障害試験を行う必要がある．しかし障害試験は，試験作業に人手と時間を要するため，開発コスト増加の一因となっている．本稿では，サーバ仮想化技術を利用し，障害試験を自動化するとともに試験時間を短縮する手法を提案する．

A Proposal of the Automated Testing Method Applying Server Virtualization Technology

Yoshihiko Hanazaki[†], Takayuki Ito[†], Shunsuke Kokubu[†], Toshiaki Katayama[†], Ryo Okabe[†] and Tsuyoshi Iizuka[†]

To achive high-availability system, it is necessary to test failure detection and recovery mechanism of the system sufficiently. Such tests is a cause of increase of system cost, because it requires manual operations for a long term. This paper describes a proposal of the automated testing method applying server virtualization technology, which automates fault test and reduce the time required for testing.

1. はじめに

近年，社会インフラ向けシステムなどの高信頼を要求される大規模システムにおいても，汎用サーバや仮想化技術を用いて低コストでシステムを開発/構築することが求められている．汎用サーバを用いて高信頼システムを構築する場合，ソフトウェアによる冗長化を行う方法が主流であり，故障発生時のシステム動作を検証する障害試験を十分に行うことが必要である．しかしながら障害試験には人手と時間を要し，開発/構築コスト増加の一因となる．このため，障害試験における試験作業の効率化・省力化が求められている．

このような背景のもと，本稿ではサーバ仮想化技術を利用し，障害試験における試験作業を自動化するとともに，試験実施に要する時間を短縮する障害試験自動化手法を提案する．

2. 背景と課題

高信頼システムを実現するためには，特殊な故障防止機構を備えた専用機を用いる方法と（メインフレーム・フォールトトレラントサーバなど），汎用サーバを用いる方法がある．汎用サーバを用いる場合，クラスタミドルウェアなどのソフトウェアによる冗長化を行い，故障検知/系切り替えなどの故障対策処理を実現する．このようなシステムでは，故障の発生部位や種別によって故障検知や切り替え方法が異なる．またハードウェア構成等に応じて，システム独自の故障対策処理を追加する場合もある．したがって十分な障害試験を行い，システムで発生し得る故障に対して，故障対策処理が期待通りの動作を行うことを確認することが重要である．

一般に障害試験では，試験項目ごとに以下の作業を行う．システム復旧は，試験のために発生させた故障によって異常状態となったシステムを，次の試験項目を実施するために正常状態に戻す作業である．

- 試験開始状態への移行
- 故障発生と動作確認後の故障解除
- 故障対策処理の動作確認（故障検知/系切り替えなど）
- システム復旧（サーバ再起動・データ復旧など）

従来はこれらの作業を人手で行っており，試験実施にあたって以下のような課題があった．

(1) 試験範囲に限界がある

試験のつどハードウェアを破壊することは現実的ではなく，何らかの方法で模擬的に故障を発生させる必要がある．しかし人手で模擬可能なハードウェア故障はケーブ

[†] 三菱電機株式会社 情報技術総合研究所
Information Technology R&D Center, Mitsubishi Electric Corporation

ルの抜き差しや電源の OFF/ON 程度にとどまり，実施可能な試験範囲が限られている．

(2) 試験要員を長時間拘束する

故障の発生/解除及びシステム復旧に人手を要するため，実施に時間を要する障害試験のほぼ全期間にわたって試験要員を拘束している．

(3) 試験実施に時間を要する

上述したように，試験を行うために発生させた故障により異常状態となったシステムを，次の試験項目を行うために正常状態に復旧しなければならない．このためにサーバ再起動やデータ復旧などの作業を人手で行っており，故障発生と動作確認という試験本来の作業に比べて，復旧作業に大きな時間をかけている．

3. 障害試験自動化手法

本章では，本稿で提案する障害試験自動化手法の概要と，自動化の実現にあたってポイントとなる機能の実現方式について述べる．

3.1 概要

3.1.1 全体構成

図 1 に本手法の全体構成を示す．本手法では，試験対象のサーバを，仮想化ソフトウェアであるXen[1]を用いて仮想化されたサーバのゲストドメインとして配置する．検証端末上で実行する検証スクリプトから，以下の機能を制御することによって障害試験の自動化を実現する．検証スクリプトは試験手続きを記述したスクリプトであり，各機能が検証端末上に提供する制御コマンドを呼び出して試験を自動実行する．

(1) 前処理実行機能

試験項目を実行する前に行う必要がある，システム特有の処理を行う機能である．二重系システムにおける主系/待機系の設定など，下記(4)の状態保存/復元機能を用いて正常状態を復元した後，試験開始状態への移行操作が必要な場合に使用する．

(2) 故障発生/解除機能

ゲストドメインに対して，模擬的なハードウェア故障の発生/解除を行う機能である．ゲストドメインとハードウェアの間に位置する仮想化レイヤで故障を注入することにより，ゲストドメインに対するディスク/LAN/CPU/メモリ等の故障模擬を実現する．なお本機能は，これまで提案/開発を行ってきた故障模擬手法[2][3]を用いて実現する．

(3) 状態情報収集機能

模擬故障を発生させた後に，試験結果の確認に必要な情報（状態情報）を，自動的に検証端末上に収集･蓄積する機能である．状態情報としてはゲストドメイン上の OS，ミドルウェア，アプリケーションが出力するログファイルを想定する．仮想化環境に特有の管理ドメインから情報収集，仮想化ソフトウェアのスナップショット機能を用いた仮想 CPU/メモリ/ディスクの状態を保存により，障害試験において要求される異

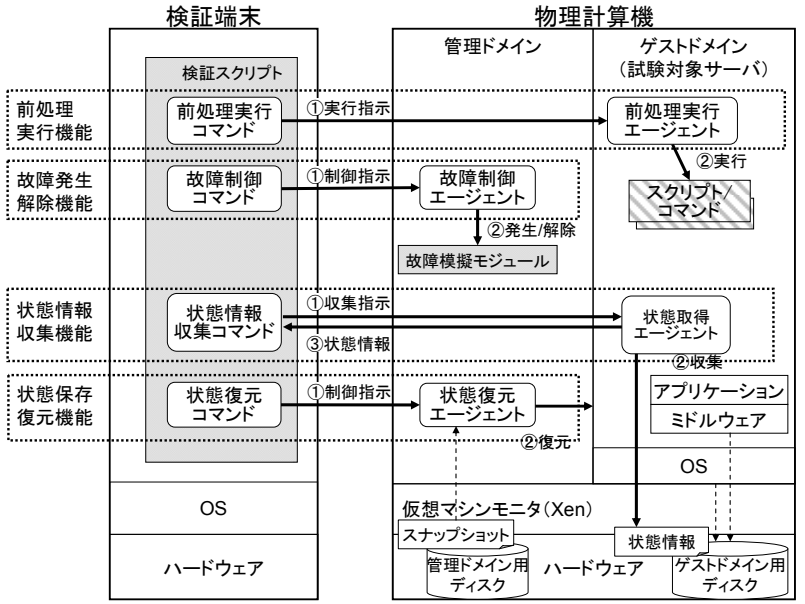

図 1 全体構成

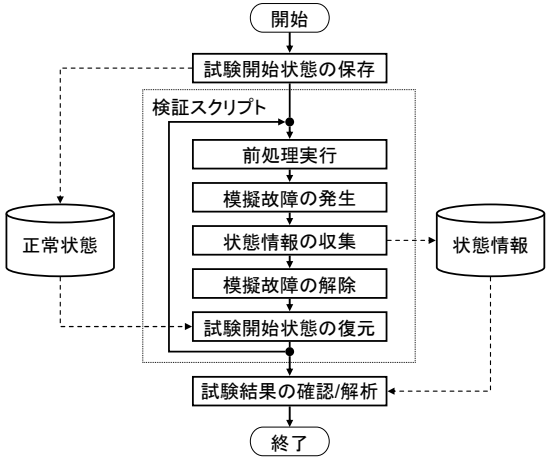

図 2 試験フロー

常状態にあるシステムからの状態情報収集を実現する．

(4) 状態保存/復元機能

1 つの試験項目の実行後，模擬故障により異常状態となったゲストドメインを，次の試験項目を実施可能な正常状態に復元する機能である．全試験項目の実施に先立って正常状態を保存しておき，試験項目実行のつど，保存した正常状態を復元する．仮想化ソフトウェアのスナップショット機能を利用し，ゲストドメインの仮想 CPU/メモリ/ディスクの状態を保存/復元することにより，正常状態の自動的かつ高速な復元を実現する．

3.1.2 試験フロー

図 2 に，本手法における障害試験のフローを示す．あらかじめ正常状態を保存しておき，検証スクリプトを実行することによって複数の試験項目を自動実行する．全試験項目の完了後，試験実行中に蓄積される状態情報を用いて試験結果の確認・解析を行う．検証スクリプトは試験項目ごとに，前処理，模擬故障の発生，状態情報の収集，模擬故障の解除，正常状態の復元の各処理を実行する．

3.2 故障発生/解除機能

Xen に故障模擬モジュールを組み込むことにより，ゲストドメインに対する模擬的なハードウェア故障の発生/解除を実現する．図 3 に故障模擬方式の概要を示す．

Xenでは基本的に，ハードウェア入出力は管理ドメインが一括して行う構成である．ゲストドメインからのハードウェア入出力は，ゲストドメイン上のフロントエンドドライバを通して，管理ドメイン上のバックエンドドライバへ送られる．故障模擬モジュールは管理ドメイン上のドライバとして組み込まれ，ゲストドメインからの入出力要求をフックし，エラー応答や遅延を発生させることによって故障模擬を実現している[2][3].

表 1 ハードウェア模擬故障一覧

対象ハードウェア	模擬内容
LAN	物理リンク断
	VM 間リンク断
	無応答
	応答遅延
ディスク	バスエラー
	HDD メディアエラー
	無応答
	応答遅延
CPU	実行停止
メモリ	ECC/Parity エラー

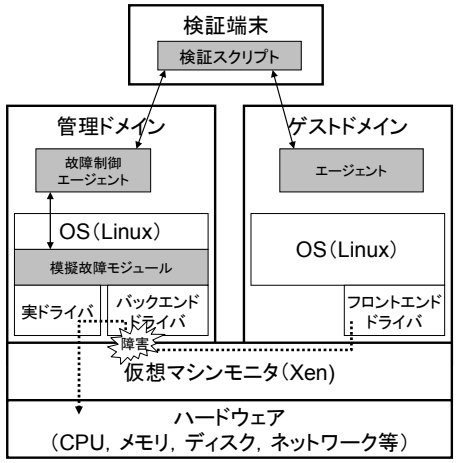

図 3 故障模擬方式

3.3 状態情報収集機能

障害試験においては，試験のために発生させるによって情報収集が行えない場合がある（LAN 故障など）．また試験時の動作が，試験内容から期待されるものと異なる場合，収集対象として指定した情報のみでは原因解析を行えないことも考えられる．

これに対して本機能では，試験実行時に発生させる模擬故障の内容と，試験項目実行後のゲストドメインの動作状態に応じて収集手段を切り替えることにより，試験後の動作状態によらない状態情報の収集を実現する[4].

3.3.1 収集方式

本システムでは，以下に示す 3 つの収集方式を用いる（収集方式の選択方法は後述する）．図 4 に，状態情報の収集方式の概要を示す．

(1) 直接アクセスによる収集

ゲストドメイン上に配置するエージェントが状態情報を収集し，ネットワークを経由して検証端末へ送信・蓄積する．ゲストドメインが動作しており，ネットワーク通信が可能な場合にこの方法を用いる．

(2) 間接アクセスによる収集

管理ドメイン上に配置するエージェントが状態情報を収集し，ネットワークを経由して検証端末へ送信・蓄積する．ゲストドメインがネットワーク通信を行えない場合にこの方法を用いる．この方式では，ゲストドメインの仮想ディスクの内容を，以下の手順で管理ドメインから読み出す．

- 試験実行に先立って，正常に動作しているゲストドメインから，仮想ディスクの

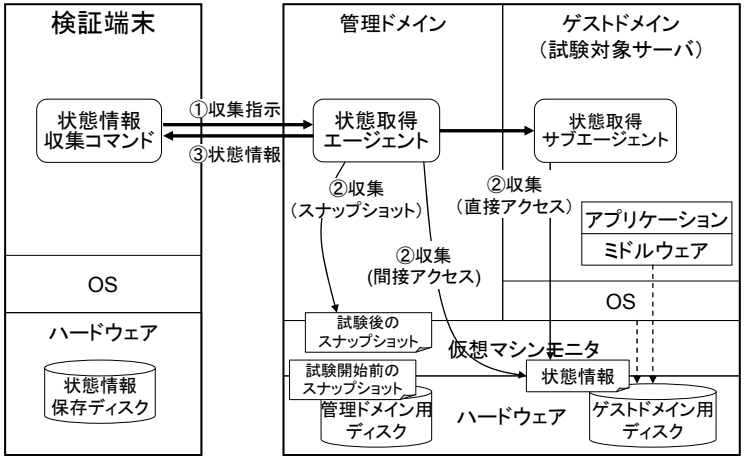

図 4 状態情報収集方式

マウント情報を取得しておく．

- 状態情報の収集時には，ゲストドメインが動作している場合は強制停止し，管理ドメインからゲストドメインの仮想ディスクをマウントする．次に，上記で取得したマウント情報を用いて，収集対象のファイルパス名を管理ドメイン上のパス名に変換し，管理ドメイン上で検索・収集を行う．

(3) スナップショットの収集

後述する状態保存・復元機能を用いて，試験項目の完了時点のゲストドメインの動作状態を保存する（3.4.2節を参照）．この方法では，収集対象として指定された状態情報に加えて，ゲストドメインのメモリ及び仮想ディスクの内容が全て保存される．試験実行後のゲストドメインの動作状態が，試験内容（発生させる模擬故障の内容）から期待されるものと異なる場合にこの方法を用いる．

3.3.2 収集方式の制御

状態情報の収集時は，試験実行後のゲストドメインの動作状態に応じて，管理ドメイン上のエージェントが3.3.1節に示した 3 つの収集方式を切り替える．エージェントは試験時に発生させる模擬故障内容と，模擬故障発生後のゲストドメインの稼動状況を取得し，模擬故障内容から想定される稼動状況と比較することで収集方式を切り替える．表 2に，模擬故障内容と想定動作を示す．

(1) 稼動状況が想定と一致する場合

ログ書き込みが可能な場合は，ネットワーク通信が可能な場合は直接アクセスによる収集を，通信不可能な場合は間接アクセスによる収集を行う．ログ書き込みが不可能な場合は，このゲストドメインからの収集は行わない．

(2) 稼動状況が想定と一致しない場合

この場合は試験結果が FAIL であることは当然であるが，試験後の原因解析のための情報を収集しておくことが望ましい．ただし想定とは異なる動作をしているため，収集対象として指定した情報のみでは情報が不足する可能性がある．したがって，このケースではスナップショットの収集を行う．

表 2 模擬故障と想定動作

模擬故障	ログ書込可否	通信可否	稼動状況
LAN 無応答	可	不可	稼動
LAN 応答遅延	可	可	稼動
ディスク無応答	不可	可	稼動
ディスク応答遅延	可	可	稼動
CPU 停止	不可	不可	停止
・・・	・・・	・・・	・・・

3.4 状態保存/復元機能

ゲストドメインの仮想ディスクを管理ドメイン上の論理ボリュームとして確保しておき，仮想メモリ，仮想CPU及び仮想ディスクの状態をディスク上に保存・復元することでゲストドメインの動作状態の保存･復元を実現する. 仮想メモリ及び仮想CPUの状態保存・復元にはXenのスナップショット機能を，仮想ディスクに関しては論理ボリュームマネージャのスナップショット機能を用いる[5].

本システムでは，試験実行時の正常状態の保存・復元と，試験結果が期待したものと異なる場合の結果状態の保存・復元を行う．以下，それぞれについて実現方式を示す.

3.4.1 正常状態の保存・復元

全試験項目の実行に先立って正常状態の保存を行い，試験項目ごとに状態復元と試験実行を繰り返す. 全試験項目の実行完了後, 保存した正常状態を削除する. 図 5 に，正常状態の保存・復元の概要を示す.

(1) 保存

以下に示すように, 仮想メモリ･仮想 CPU → 仮想ディスクの順に状態保存を行う. 状態保存が完了すると，保存時の状態を復元可能となる.

- 仮想メモリ・仮想 CPU

ゲストドメインの仮想メモリと仮想 CPU のレジスタ内容を, Xen のスナップショッ

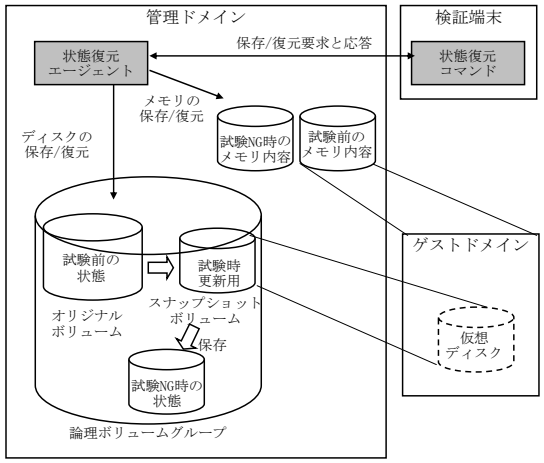

図 5 状態保存・復元方式

ト機能を用いて，管理ドメイン上のファイルとしてに保存する．

- 仮想ディスク

以下では，ゲストドメインの仮想ディスクが定義された論理ボリュームをオリジナルボリュームと呼ぶ．仮想ディスクの状態保存時には，オリジナルボリュームのボリューム名を別名に変更する．さらに名前を変更したオリジナルボリュームに対して，元のオリジナルボリューム名を持つスナップショットボリュームを作成する．

(2) **復元**

状態保存後に試験を実行すると，仮想メモリと仮想 CPU のレジスタ内容が変更される．また試験実行中に行われた仮想ディスクの更新内容は，上記で作成したスナップショットボリュームに更新差分として保持されており，オリジナルボリュームの内容は変更されない．

試験実施後に，仮想ディスク → 仮想メモリ・仮想 CPU の順に状態復元を行う．状態復元を行うことにより，ゲストドメインは状態保存時の動作状態に復元される．

- 仮想ディスク

ゲストドメインを強制停止し，スナップショットボリュームの削除と再作成を行う．削除・作成により，スナップショットボリューム中に保持されている試験実行中の更新内容がキャンセルされ，状態保存時の状態が復元される．

- 仮想メモリ・仮想 CPU

Xen のスナップショット機能を用いて，管理ドメインに保存した仮想メモリと仮想 CPU のレジスタ内容を復元し，ゲストドメインの動作を再開する．

(3) **削除**

スナップショットボリュームを削除し，オリジナルボリュームを元のボリューム名に変更する．さらに，管理ドメインのファイルシステム上に保存した，メモリ・CPU 状態ファイルを削除する．

3.4.2 試験結果状態の保存・復元

前述したように，試験項目完了時点の動作状態が期待したものと異なる場合，ログファイルだけではその原因を解析できない場合がある．このような場合は試験項目完了時点の動作状態を保存しておき，全試験項目の完了後に状態を復元して試験結果の確認・解析を行う．

(1) **保存**

正常状態と同様に仮想メモリ・仮想 CPU の状態を保存し，スナップショットボリュームを別名に変更する．ここで退避したスナップショットボリュームには，試験実行中の更新内容が書き込まれている．

(2) **復元**

正常状態と同様にスナップショットボリュームの再作成を行い，作成したボリュームに上記で退避したスナップショットボリュームの内容をコピーする．次に仮想メモリ・仮想 CPU の状態を復元し，ゲストドメインの動作を再開する．

4. 二重系システムへの適用と効果

4.1 二重系システムへの適用

本手法を二重系システムに適用する場合，主系/待機系となる各ゲストドメインに対して，それぞれ正常状態の保存・復元を行う．基本的には，各ゲストドメインが主系/待機系モードで動作している状態で状態保存を行い，主系 → 待機系の順に復元を行えばよい．

しかしシステムによっては，待機系が稼動していない状態を異常状態とみなす場合もあり得る．このようなシステムに適用する場合は，正常状態として各ゲストドメインが準備モードにある状態を保存・復元する．1 つの試験項目の完了後に次の試験項目を実行する際は，準備モードを復元した後に，前処理として準備モードから主系/待機系モードへの遷移処理を実行する．図 6 に処理フローを示す．

4.2 適用の効果

本手法を用いることにより，障害試験の試験範囲の拡大，試験作業の省力化，試験実施時間の短縮という効果が得られる．

(1) **試験範囲の拡大**

故障発生/解除機能により，人手による従来の試験では発生させることができなかっ

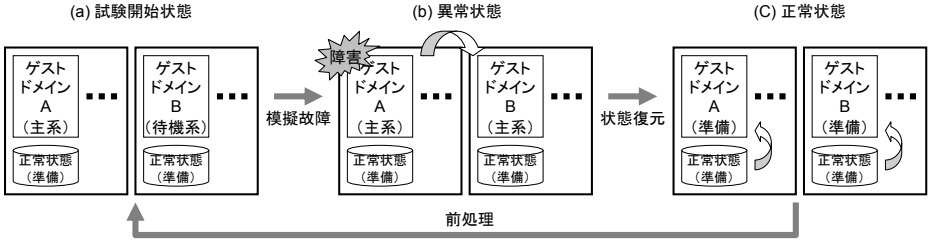

図 6 試験開始状態の復元

たハードウェア故障に対しても，障害試験を行うことが可能になる．

(2) 試験作業の省力化

故障発生/解除機能による模擬的なハードウェア故障の発生/解除，状態情報収集機能によるサーバ動作状態によらない状態情報の収集，状態保存/復元機能による正常状態の自動復元，前処理実行機能による試験開始状態への自動移行により，障害試験の自動化が実現される．自動化することにより，試験作業が省力化されるとともに，夜間実行等によって試験期間を短縮することが可能になる．

(3) 試験実施時間の短縮

状態保存/復元機能による正常状態の復元時間の短縮により，障害試験の実施に要する時間が短縮される．

人手による従来の試験方法では，システム復旧のためのサーバ再起動に概ね表 3 に示す時間が必要となる．試験開始状態の準備には，再起動に加えてデータ復旧も行う必要があり，システム構成/規模やミドルウェア/アプリケーションによっては，試験項目ごとの準備処理に 20 分以上を要する場合もあり得る．これに対して，本手法を用いた場合，状態復元に要する時間は表 4のとおりである（Xeon5160 クラスのCPU搭載サーバの場合）．復元時間はメモリサイズのみに依存し，ディスク容量にはよらない．メモリサイズが 4GBの場合では 1 分以内に状態復元が完了する．

表 3 サーバ再起動に要する時間

復元項目	所要時間
OS シャットダウン	1～2 分
BIOS 処理	1～2 分
OS 起動	2～3 分
ミドルウェア/アプリケーション起動	数分以上

表 4 状態復元に要する時間

復元項目	所要時間
CPU/メモリ	メモリ 1GB あたり 9 秒
仮想ディスク	サイズに関わらず 1 秒

本手法が対象とする大規模な高信頼システムの開発では，障害試験の項目数は数 100～1000 項目に及ぶ．障害試験の実施に大きなコストをかけており，障害試験の実施時間の短縮が，試験コストの低減に大きな効果を持つことを期待できる．例えば障害試験が 1000 項目あり，各項目の実施に 25 分を要していると仮定すると，全項目の実施には約 417 時間（52 日）が必要になる．これに対して，本手法の適用により各項目の実施時間が 5 分に短縮された場合，全項目の実施に要する時間は約 83 時間（10 日）に短縮されることになる．

5. おわりに

高信頼性を要求されるシステムでは，故障対策処理の動作を検証するために，十分なの障害試験を行うことが要求される．しかし人手による従来の障害試験には，試験範囲に限界がある，試験要員を長時間拘束する，試験実施に時間を要するという課題があった．

この課題に対して本稿では，故障発生/解除・状態情報収集・状態保存/復元および前処理実行の 4 機能からなる障害試験自動化手法を提案した．故障発生/解除機能は，仮想化レイヤで故障注入を行うことにより，人手による発生が困難なハードウェア故障を模擬可能とする．また状態情報収集機能は，仮想化されたサーバが持つ管理ドメインとスナップショット機能を利用し，障害試験の実施によって異常状態となったサーバからの試験結果情報の自動収集を実現する．状態保存/復元機能は，仮想化ソフトウェアのスナップショット機能を利用し，サーバの正常動作状態の自動的かつ高速な復元を実現する．状態保存/復元機能と前処理実行機能により，障害試験項目の実施により異常状態となったサーバを，次の試験項目の開始状態に自動的に移行することが可能になる．

本手法は上記 4 機能により，障害試験における試験範囲の拡大と試験作業の省力化，試験実施時間の短縮を実現し，低コストでの高信頼システム開発/構築に寄与する．今後は，実際のシステム開発への適用に向け，本手法の評価を行っていく予定である．

Vol.2010-SE-167 No.31
2010/3/19

参考文献

1) Xen ：http://www.xen.org/
2) 片山 他，“仮想マシン環境における障害模擬手法～仮想障害発生機構”，電子情報通信学会 2008 年総合大会，D-10-6，March.2008
3) 國分, 他 : 仮想化環境におけるハードウェア障害模擬と HA クラスタシステム試験への適用，信学技報 DC2008-19，pp.1～7(2008)
4) 國分，他：仮想化技術を利用した異常処理試験自動化手法 - 状態情報収集方式 - ，情報処理学会第 72 回全国大会(2010)
5) 伊藤，他：仮想化技術を利用した異常処理試験自動化手法 - 試験開始状態の復元 - ，情報処理学会第 72 回全国大会(2010)