３．９　不測事態発生への備えに関する教訓（Ｔ９）

[教訓T9]
検証は万全？それでもシステム障害は起こる。回避策を準備しておくこと

問題

A 社の、あるアプリーションサービスを提供する 2 つの設備間において、通信障害時にタイムアウト検出機能の潜在不良が顕在化し、装置のバッファオーバーフロー、再起動につながった。サービス再開までの間、利用者からのリクエスト処理が停滞した。

システムの概要は次の通り(図３．９－１)。

・アプリケーションサービス用のサーバを冗長化。設備 1 でリクエストを受付け、設備 2 で処理を行う。
・設備 1 には外部からのリクエストの振分制御、処理待ちリクエストのバッファ、サーバ#1、#2 それぞれのクライアントプロセス#1、#2 が実装される。
・設備 2 には 2 つのリクエスト処理用サーバ(サーバ#1、#2)が実装される。
・通常、設備 1 で受付けたリクエストは振分制御を介してプロセス#1 へ受渡され、設備 2 のサーバ#1 により処理され、終了するとプロセス#1 へ通知される。

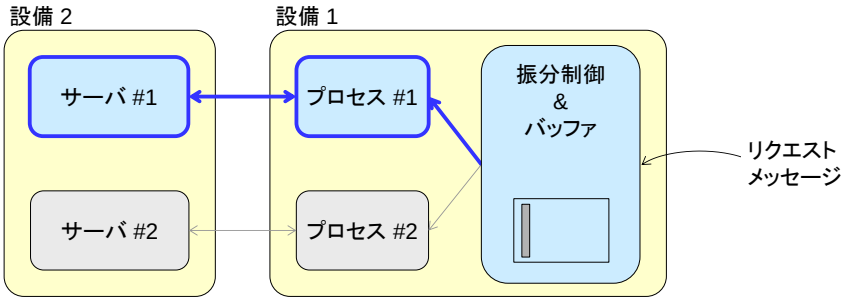


図３．９－１　システムの構成イメージ。通常はサーバ#1 の系統が稼働

・サーバ#1 の障害をプロセス#1 が検知すると、リクエストは振分制御によりプロセス#2 へ受渡され、設備 2 のサーバ#2 により処理され、終了するとプロセス#2 へ通知される(図３．９－２)。
・サーバの処理待ち、障害時切替え等の間に受付けたメッセージは、設備 1 の処理待ちバッファに蓄積される。
・処理待ちバッファがオーバーフローした場合、設備 1 を再起動、バッファはクリアされる。

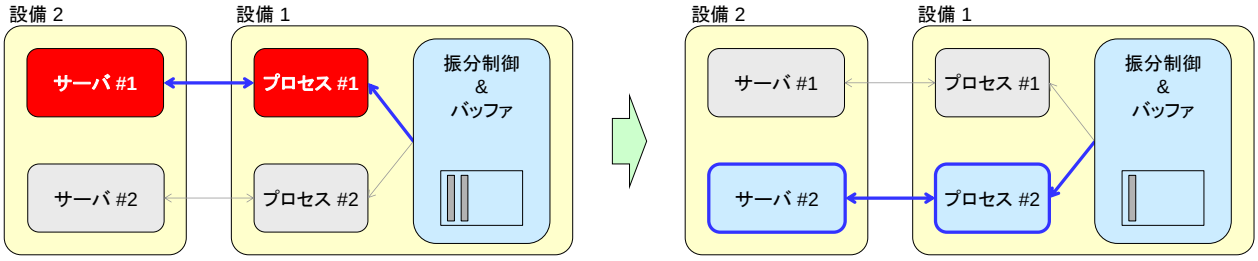
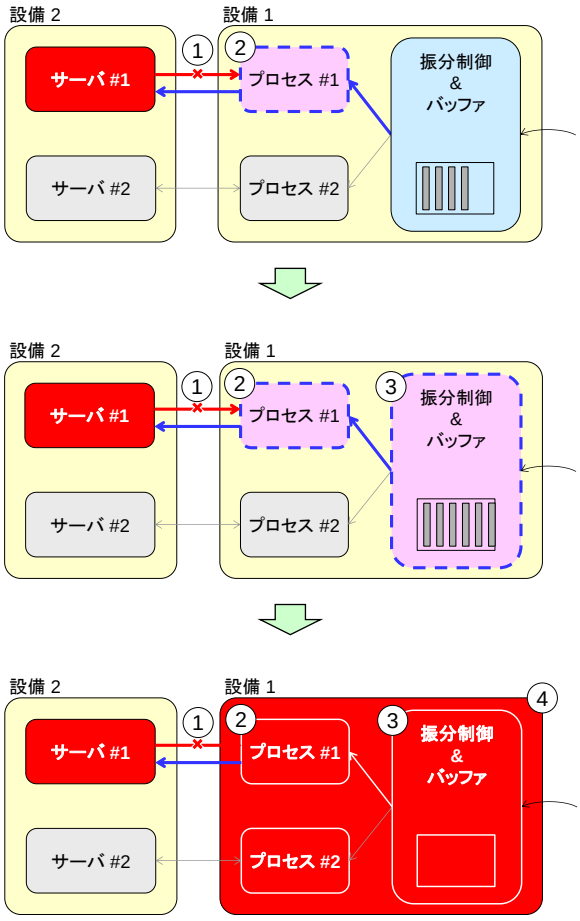


図３．９－２　障害時の自動振分

障害発生の経緯を図３．９－３に示す。



①設備 2 のサーバ#1 に障害が発生し、TCP レイヤの一方が途絶した。
②設備 1 のプロセス#1 はセッション断を検知できず、サーバ#1 からの応答待ち状態のままとなった。

③このため振分制御側で異常を検知してプロセス#2 へ切替えることができず、リクエストはバッファに蓄積され続けた。

④バッファが満杯となり、設備 1 がリブートされた。

図３．９－３　障害発生の経緯

原因

直接的な原因は次の通り。

・設備 1 のプロセス#1 における、アプリケーションレイヤのタイマ制御処理ソフトウェアのバグにより、TCP レイヤでの切断時にタイムアウトを検知できなかった。

根本的な原因は、本事例の契機となった通信障害のケースについて事前に検証できていなかったことにあるといえるが、その責を当該システムの調達側である A 社の不備に帰することは現実的ではないと考える。

・ネットワーク機器の不具合に起因する障害は、検証のため再現することさえ難しい場合も多く、取りうるすべての状況を事前に確認することは事実上不可能である。
・当該システムは全体を外部から調達したもので、タイムアウト検出機能の動作についても、本来製造元で担保すべきものである。更に本事例の原因となった通信途絶は希少なケースであり、検収時点においては調達側、供給側の双方とも当該ケースの存在を認識していなかった。

対策

本事例において実際に行われた対策は次の通り。

・暫定処置
  - 手動によりプロセス#2 へ強制的に切替え、業務を継続した。
・直接的な原因への対策
  - 設備 1 クライアントプロセスのタイマ制御不具合を修正した。
・類似障害の再発防止策
  - タイムアウト前にオーバーフローによるリブートが発生しないよう、バッファサイズを拡大した。
  - サービス復旧のための作業優先順位を定め、障害時の強制切替え手順を明記した。

当該システムがとり得る全ての状態の組合せを検証できていなかったことについては、対策として、設計・検証を徹底的に厳密化すべしと謳うことは、些か実現性に乏しいと考える。

現場においては、対象システムの重要性に応じて、検証に費やすことのできる時間と労力は制約される。発生確率の低い、希少なケースについて、完全に洗出し検証することは難しい。

当該システムの供給者の立場においては、システムの価値に応じて要求される品質水準を満たすよう、必要十分な検証を行ったことが確認できれば、責任は果たしていると考えるのが一般的であろう。

本事例で見ると、次のようにいえる。

・プロセス#1 とサーバ#1 の間にクローズドループ(指示に対して必ずフィードバックが得られる関係)が成立していることを確認(保証)するために、各レイヤ毎の応答有無の全ての組合せについて動作を想定し、検証しておけば、本事例の不具合が事前に検出された可能性はある。
・テストスタブを用いて、各レイヤの応答が得られなかった場合の動作を検証することは理論的には可能だが、本事例のような不具合を検出するためには、少なくとも各レイヤの応答有無の組合せを網羅する必要がある。
・上記検証方法が、当該システムの品質要求に照らして必要、または所要コストからみて許容できるものであった場合には、実施されるべきであった。

従って、当該システムを用いて業務機能を提供するサービス提供者の立場においては、常に不具合が潜在しているとの前提に立ち、検知と障害個所の回避により業務の継続性を確保することが、障害発生の予防策に劣らず重要である。

効果

・不具合修正により、同様の通信途絶時にタイムアウト検知が可能となった。
・バッファサイズ拡大により、設備 1 がリブートされる前に人手による強制切替えを実行できる時間的余裕が確保された。
・今後製造側では設計、検証段階において、調達側では検収、統合テスト段階において、より注意深いチェックが行われるようになると期待される。

教訓

要件定義、設計、構築等の各段階(工程)において、十分な検証が行われたとことが関係者間で確認・合意されていても、運用開始後の障害につながる不測の事態(要因)が完全に排除されたと保証できる訳ではない。

サービス提供者は、システムには常に不具合が潜在するとの前提に立ち、業務継続性を担保できるよう、障害回避策等を準備しておく必要がある。

・調達(購買、内製の双方を含む)局面において、当該システムが要求する品質水準を確保できるよう、必要十分な範囲で障害発生条件を網羅検証し、想定し得る障害の発生を予防する。
・運用局面において、システム障害に備えて、設備を冗長化し、障害を検知したら発生個所を切離して運用を継続する等、当該システムが要求するサービス継続性を確保できる体制、手続きを整える。