３．１１　サイレント障害に関する教訓（Ｔ１１）

[教訓T11]
サイレント障害を検知するには、適切なサービス監視が重要

問題

A 社の、ある Web サービスにおいて、明示的に障害が検出されていないにも拘らず性能が劣化する、いわゆる「サイレント障害」が発生した。障害が解消されるまでの間、利用者は応答速度低下の影響を受けた。

ネットワークシステムにおいて、通常の運用監視の仕組みからは検出されないまま、性能劣化等の現象が発生することを「サイレント障害」と呼ぶ。放置すれば最終的にネットワークに接続できなくなるなど、システム全体に影響が及んで大規模な障害につながることもある。

サイレント障害では、発生個所、原因の特定に多くの時間と労力を要することが多い。そのため利用者は不利益を、サービス提供者は機会損失や風評リスクを長時間にわたり蒙ることになりがちである。

システムの概要を図３．１１－１に示す。

・負荷分散のため、1 台の負荷分散装置下に 3～6 台のサーバを配したセットを、複数配置している。

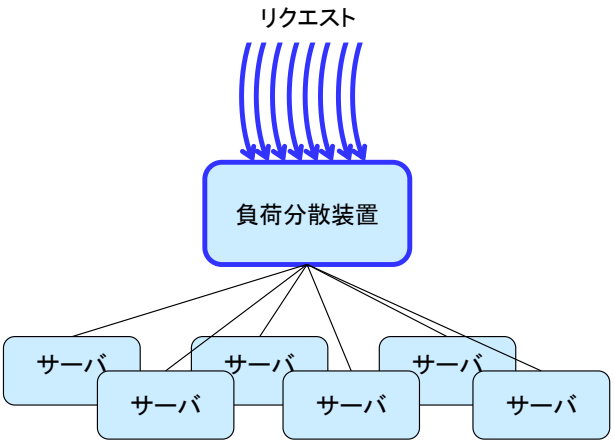
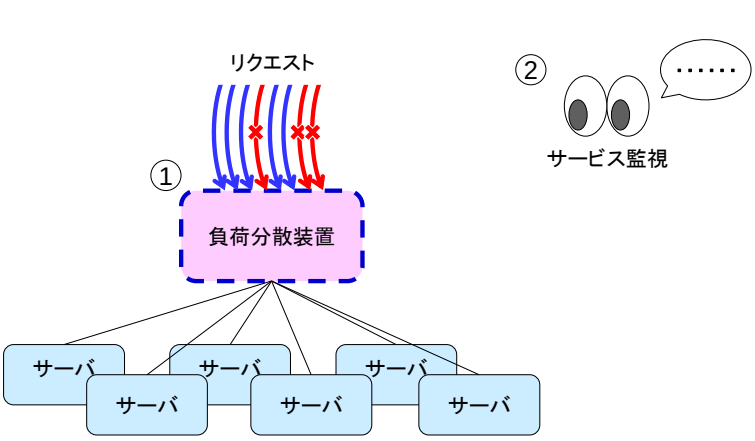


図３．１１－１　システム構成のイメージ

障害発生の経緯を図３．１１－２に示す。



①負荷分散装置において、セッション数オーバーによるリクエスト廃棄・再送が発生。応答速度が低下した。
②サービス監視からの通知はなく、性能が低下した状態が続いた。
③外部(社員)から指摘を受けるまで発見できなかった。

図３．１１－２　障害発生の経緯

原因

直接的な原因は次の通り。

・負荷分散装置のファームウェアの不具合があった。障害発生時に使用中のバージョンにおいては、受付けるセッション数の上限が、設定値の 1/4 迄しか許容されない「仕様」とされていた。これを認識しないまま上限値を設定したため、想定を遥かに下回るトラフィック数であったにも拘わらずリクエスト廃棄が発生した。
後のバージョンで改善(修正)されているので現在は問題ないとメーカは説明している。

・サービス監視はリクエストが一定回数連続して廃棄された場合にアラームを発するよう設定されていたが、閾値を超えるまでには至らなかったため、サービス監視からの通知はなかった。

根本的な原因は、サービス監視条件が必ずしも最適ではなかったことにあるともいえるが、ファームウェアの不具合がなければ、特段問題視されることはなかった可能性もある。

対策

本事例において、実際に採られた対策は次の通り。

・直接的な原因への対策
- 負荷分散装置のファームウェア更新
- サービス監視条件を変更。リクエストが一定時間内に一定回数以上廃棄されたら通知するよう改めた。

・類似障害の再発防止策は次の通り。
- 取扱ベンダーを一本化し、製品の情報を直接入手できるようにした
- 本事例の発生以前から、定期的なサービスへのアクセス、目視による SNS 上の「つぶやき」監視等を行っている。つぶやきから監視機能よりも早く障害を検知できる場合もあり、有効であると認識している。

- 更に「インバリアント分析」を用いた早期検知の試みを開始した。

本事例においては、サービス監視条件の設定が妥当であれば、より早い段階で障害の兆候を把握できた可能性はある。

運用中にサイレント障害の発生を速やかに検知し、分析・対処につなげるための基本的な取組みは、サービス監視機能を用いる、或いは実際にサービスにアクセスしてみる等の方法により、性能劣化が見られないかをチェックすることである。

基本的な取組みを実践した上で、更なる早期発見、対処を望むサービス提供者においては、監視機能に加えて障害検知、分析のための技術・製品が用いられるようになっている。

参考に、最近登場している、サイレント障害の検知や分析を行うための技術の例をいくつか紹介する。

(1)サイレント障害切分けシステム

IP ルータ網監視システムのサブシステムとして開発された。従来の障害検出技術によってサイレント障害を発見するには、高度な技能を持つ保守担当者による長時間の探索が必要であったが、当該システムが提供する「サイレント障害検出機能」と「サイレント障害発生区間特定機能」により、サイレント障害の検出及び障害個所の特定が容易になった。

「大規模な IP ネットワークにおける高精度な障害切り分けシステムの開発」(NTT DOCOMO テクニカル・ジャーナル Vol18 No.1)

(2)インバリアント分析技術の応用

性能情報間の相関関係のうち、平常時に変化しない関係（インバリアント）を自動的に学習してモデル化し、そのモデルと一致しない「いつもと違う」挙動をサイレント障害として検知する。自動的に性能情報を収集し、リアルタイムで分析を行い、障害を検知した時点で警告メッセージを通知する機能もある。

「WebSAM の分析技術と応用例～インバリアント分析の特長と適用領域～」(NEC 技報 Vol65 No.2/2012)

(3)ビッグデータ分析技術の応用

ネットワーク装置のログ(Syslog)や SNS 上の「つぶやき」情報といったビッグデータをリアルタイムに分析し、障害の早期発見、ひいては予兆検知に利用する技術が研究されている。

「Syslog と SNS 分析によるネットワーク故障検知・原因分析技術」(NTT 技術ジャーナル 2013.07)

効果

・直接的な原因への対策により、システム障害に至る前の段階で、性能劣化を検知できる確率が高まった。

・更に SNS 監視や、各種技法を用いることにより、サイレント障害をより早期に検知できるようになると期待できる。

教訓

サービス中断、サーバ停止といった「分かり易い」障害と異なり、サイレント障害はその発生を検知することが難しく、解決にも時間を要しがちである。

サイレント障害検知のためには適切なサービス監視が最も重要であり、基本的な取組みである。更なる早期検知、分析のための技術、製品等の利用も始まっている。