2.18 障害対策を立案する際に利用部門と取り決めるべき事項に関する教訓（G18）

教訓 G18　障害対策とは許容時間内の回復や停止中の業務継続まで具体化すること

問題

A 社で保有する社内特定業務向け基幹業務システムは、現場部門に対して 24 時間のオンライン業務サポートを実施している。業務は日中だけでなく夜間も継続しており、かつシステム停止による現場業務への影響が大きいことから、運用サポート部門も 24 時間体制でシステムの運用をサポートしている。業務システムは当該用途向けのパッケージを購入して構築されており、アプリケーションサーバ、DB サーバ、他システムとの連携サーバなど各層のサーバはそれぞれ冗長構成になっている。システムは本稼働後 3 年の間にサービスの一時停止に至る障害を 3 回経験していた。

A 社では、おのおのの障害の発生時点で原因分析と再発防止策を適切に実施し、それぞれ効果を上げてきた。しかしながら、このシステムが A 社業務の中核を担うサービスを提供するものであり、サービスが一時的にでも停止すると A 社の業務に甚大な影響をもたらすことから、A 社は障害発生時に早期にシステムを回復させる施策と、システムが停止中でも最低限の業務を継続できる施策の検討を開始した。

原因

A社が直面したサーバダウンと、その原因、主な再発防止策の概略は、以下のとおりである。

表 2.18－1　A 社システムにおけるサーバ停止の概要

No.	内容	停止時間	発生時期	直接の原因	影響	個別の再発防止策
①	アプリケーションサーバダウン	8 時間	本稼働 2 カ月後	パッケージの潜在的な不具合	大	開発者による 24 時間直接サポート
②	アプリケーションサーバ応答なし	3 時間	メジャーバージョンアップ 2 週間後	バージョンアップ時に実施されたリリースノート未記載のアプリケーション修正の不具合	中	すべての修正に対するレビュー会の実施 回帰テストシナリオの作成と自動実行環境の構築
③	他システムとのデータ連携サーバの二重起動（運用系と待機系）	1 時間	②の 1 カ月後	通信制御ミドルウェアの不具合による待機系切替えの誤起動	小	異常検知時のメッセージの整理と、障害発生時の運用要員の対応範囲変更

障害を経験し、対策を早期に実施することにより、発生から回復までの時間は確実に短縮できていることが、この表から読み取れる。

なお、上記の障害のうち①と②については、それぞれ教訓「T27 パッケージはサポートを買え」、「T28 パッケージを更新するときは、変更内容の詳細確認と回帰テストで二重に安全を確保せよ」としてまとめたので、詳細はそちらを参照していただきたい。

対策

上記の各障害に対する個別の再発防止策は確実に効果を上げているが、A 社では、このシステムが一時的にせよ停止した場合の自社業務への影響が看過できないことから、何らかの障害で一時的に停止したときに早期に回復する手段の構築と、システム停止期間中でも業務の停止を最低限に止める方法を検討した。

【システムの早期回復策】業務システムの二重化と常時並行稼働によるシステム切替えの迅速化

A 社の本番系システムはもともと各層のサーバが冗長化されており、運用系に何らかの障害があった際には待機系に切り替わるようになっていた。ところが、③の障害のように、運用系と待機系を巻き込んだ障害が発生した際には、通常の運用系と待機系の組み合わせだけでは、短時間での回復には不十分である事態が想定される。そのため、これを重視した A 社では、テスト用に保有していた別環境を増強して第二本番系として整備し、常時、本番系と同じ処理を並列で実行させておく運用を開始した。

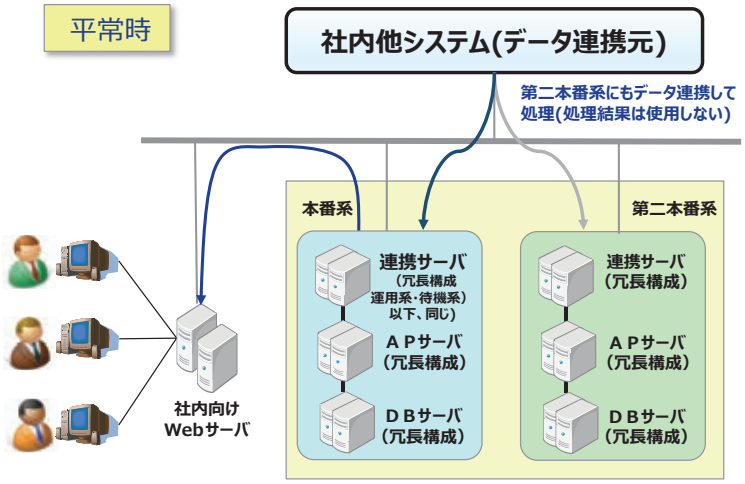

図 2.18－1　本番系と第二本番系での運用イメージ

図 2.18－1 の構成で第二本番系を構築し、他システムから連携されてくるデータはそれぞれで受信して、第二本番系でも同一の処理をさせ続ける運用をとることにより、本番系でトラブルが発生した際に、並行して稼働している第二本番系に短時間で切り替えられるようにした。これによりシステム運用者は、障害発生時の対応が長期化しそうな場合には、調査を打ち切って直ちに第二本番系に切り替えることによりサービスを再開し、その間に本番系での調査を継続できるようになった。

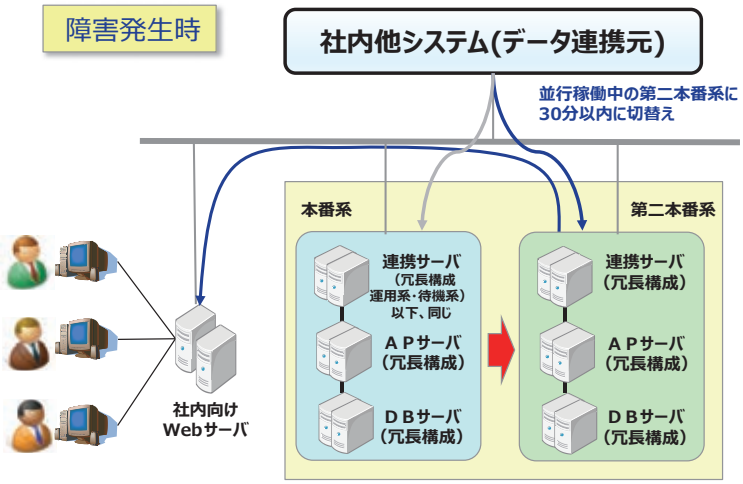

図 2.18－2　障害発生時の切替えイメージ

【業務を停止させないための対策】障害時対応をサービス利用者と明確化

また A 社では、このシステムで提供されるサービスの利用部門に対してこのサービスが使用できない場合の影響について以下の点を確認した。

① サービス停止から業務に影響を与えない（またはリカバリ可能な）復旧までの許容時間

② サービスが上記時間内に回復しない場合の業務側での代替策の有無と実現性

その結果、A 社ではこのサービスの停止後からサービス回復までの許容時間（業務の現場に対する SLA）を 30 分とした。また、サービスが回復しない場合の現場での最低限の業務継続施策を決定してもらい、それらを手順化した。作成した手順は、定期的に訓練を行い、継続的に現場への浸透を図ることとした。

効果

上記の施策により、この事例では以下の効果を得た。

(1) トラブル発生時の回復までの時間短縮

第二本番系システムでの並行稼働の開始と、サービスの利用者と合意したサービス再開までの許容時間にあわせた回復手順の構築により、これまでに発生したことがない障害に遭遇した場合も含めて、トラブルが発生しても利用者（さらにこのサービスにより提供されるエンドユーザへのサービス）への影響を極小化することができた。今後、障害が発生した際に、本番系での障害調査・対応に手間取り、本番系ではサービス利用部門と決定した 30 分以内にサービスを再開できない見通しになった場合には、そこで調査を中断して第二本番系に切り替えてサービスを再開することにより、現場への影響を最小限に止めることができるようになった。

(2) サービスがどうしても再開できない場合の代替手段の構築

上記 (1) のような施策を講じていてもサービスが再開できない万が一の事態が生じた際、その時点でこのサービスを使用しないでできる現場の業務を検討するのではなく、事前に現場で実施できることを想定して準備することにより、現場において、サービスを使用しないで維持する業務の優先度、実施範囲、実施手順をあらかじめ用意する契機になった。

教訓

基幹業務の一時停止は、多かれ少なかれ、自社の業務に何らかの影響を与える。そのため、IT システム／サービスの運用者は、システムが停止しないようにサービスの品質を安定化させることが最優先であることは言うまでもない。しかし、どのように対策してもサービスの停止は発生し得る。

そのため、サービス運用においては、その際の備えとして、以下の事項について準備をしておくことが、障害の影響を縮小させる施策として有効に機能すると思われる。

(1) サービス復旧までの許容時間は影響度に応じて異なる。サービスごとに停止不可、1 時間以内など許容時間を明確化し、それらを目標としてシステム／サービス復旧策を具体化する。

(2) サービスの復旧までの間、現場で実施できる施策を整理しておく。その上で、用意した施策を定期的に現場で訓練し、周知徹底させる。

昨今、IT システムおよびそれにより提供されるサービスは、業務の支援レベルにとどまらず、それ自身が業務そのものになりつつある。システム／サービスが停止した際には、それによって実施するビジネスそのものが停止することがほとんどである。ところが、IT システムは加速度的に複雑性を増しており、システム停止のリスクは増加傾向にある。そのため、IT システム／サービス運用においては、停止させないことの追求だけでなく、もし停止した場合に、早期にサービスを復旧させることがこれまでにも増して重要視されてきている。

この事例では、パッケージを導入して構築したシステムに対して、障害発生を機に運用品質を向上させると同時に、新たなパターンの障害に直面したときの対策として、回復目標時間をサービス利用部門と合意し、それを達成できる早期復旧策と、施策を実施しても万一復旧しなかったときにサービス利用の現場で採り得る対策の両方を構築して、万全の備えとしている。高度の運用品質を確保するために、運用系と待機系の組み合わせをさらに二重化するという、一般にはそう多くは実施されない方法を採った事例であるが、そのサービスレベルの維持に対する姿勢は、すべてのシステム運用に共通するものであると考える。障害発生を糧として、システム運用の品質をあらゆる方面から向上させる、それを継続する姿勢を教訓として発信したい。

教訓タイトルは、「障害対策とは許容時間内の回復や停止中の業務継続まで具体化すること」とした。