2.17 重要サービスの運用に関する教訓（G17）

教訓 G17 サービスの重要度を識別し、それに応じた連絡体制や障害検知のしくみを作れ

問題

（システムの概要）

A 社の B サービスは、本社で作成・編集された情報を、通信回線を通じて全国のグループ企業の関連部署に配信するグループ企業内のシステムサービスである。配信される情報は音声や動画コンテンツなども含め多岐にわたっており、また重要性・緊急性が異なる情報が混在している。そのため、通信回線が混雑している場合でも優先的に重要情報を配信できるように帯域制御装置を使用している。配信される情報は重要度に応じて多段階に分けられ、最重要レベルの情報は、リアルタイムに伝達・処理されるべき情報として、関連部署に送られる。

また、システムは二重化されており、片方の帯域制御装置が故障した場合には、もう一方の側に寄せてサービスが継続できる構成になっている。ただし、最重要レベルの情報は図 2.17－1 のシステム構成概要図に示す帯域制御装置 #1 にだけ流れる運用になっていた。

（障害内容）

この帯域制御装置を駆動するソフトウェアは定期的なライセンス更新作業が必要であり、過去数度にわたり更新を実施しているが、特段の問題は発生していなかった。今回の故障が発生した際も、これまでと同じ手順で作業を実施し、更新そのものは正常に終了した。ところが、作業実施後すぐに B サービスの利用者から通信ができないという連絡があった。確認したところ、一部のグループ企業の拠点に対し、最重要レベルの情報が不通になっていることが判明した。帯域制御装置からはアラートは発出されておらず、また最重要レベル以外の情報配信は正常に行われていた。このため運用管理担当者は故障に気づくのが遅れ、約 30 分にわたり該当の拠点への最重要レベル情報の通信断絶が続いた。

ライセンス更新作業は、最重要レベルの情報配信に影響を与えないように、先に帯域制御装置 #2 において実施し、問題が発生しなかったことを確認したのちに、帯域制御装置 #1 において実施したが、その際に #1 で上記の故障が発生した。既に両装置とも同じ更新を実施済であったことから、帯域制御装置 #2 側に切り替えても同様の故障が再発する可能性があると考え、通信を帯域制御装置 #2 に片寄せすることはせず、他の策による回復を選択することにした。最終的に帯域制御機能を停止することで最重要レベルの情報配信を回復した。

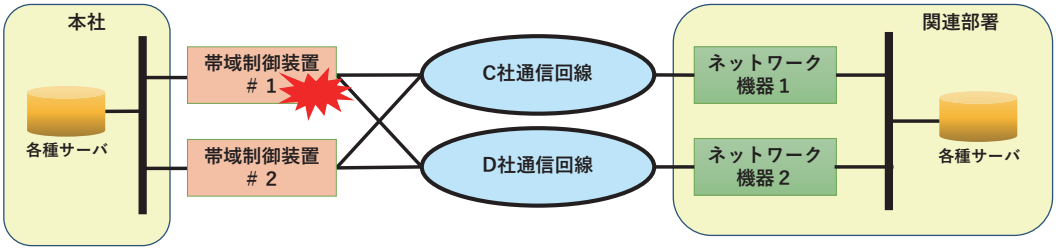

図 2.17－1　システム構成概要

原因

（故障の原因）

障害の状況から帯域制御装置の故障が疑われたが、原因の解明に結びつく特段のログが出力されていなかったことから、発生原因は特定できなかった。当該装置は海外製品であり、原因分析の対応に時間がかかるなど、故障発生時に十分なサポートが得られなかった。また、後日再現テストを実施したが現象は再現せず、原因は不明なままとなっている。

（対応が遅れた原因）

過去の同様の保守では問題が起きておらず、今回の作業においてもアラートは発生していなかったため、運用管理担当者は故障発生に気づかなかった（原因①）。

また、最初に障害に気づいたのはシステムを利用する業務担当者であったが、この時点では業務担当者から運用管理担当者への連絡ルートは確立しておらず、初動の遅れの一因となった（原因②）。これは、少しの遅れも許されない重要サービスの運用管理体制としては、不十分な状態であった（原因③）。

さらに、過去に実績のある作業であったため、関連部署への事前の作業実施連絡は不要として実施された（原因④）。事前に作業連絡があれば、もっと早く故障を発見する、あるいは現場での運用対処による影響回避策がとれた可能性がある。

対策

重要サービスに対する保守作業においては、特段の考慮をした作業環境やルールを用意する必要があると考え、以下の追加対策を実施した。

対策①　代替アラートを発出させる仕組みの追加

装置本体が持つアラート機能だけでは検知できない場合があるため、データ流量の下限しきい値監視など、他の機能や周辺機材を活用して代替アラートを発出させる仕組みを構築し、俯瞰的に故障発生を検知できるようにした。

対策② 連絡系統の見直し

重要サービスに関しては、迅速に運用管理担当者に連絡が入るよう、通常とは別の連絡系統を追加した。

対策③ 重要作業の分類

すべての保守作業項目（約 100 件）を抽出し、その作業が影響を与える各サービスの重要度に応じて、連絡要・不要などの分類を行い、関連部署と共有した。

対策④ 事前の作業連絡

たとえ過去に実績のある作業であっても、重要サービスに対する作業を実施する場合には、関連部署への作業連絡を必ず実施するようルール化した。

効果

重要サービスに障害が発生した場合、障害発生を早期に検知し、迅速に復旧対応が開始できるようになった。仮に今回の事例と同様の障害が発生した場合には、10 分以内に回復が可能となった。

教訓

本件の原因や対策を総括すると、「重要サービスに対しては特別な環境やルールを設定し、遵守を徹底する」という教訓事例になる。この事例の重要サービスは、重要と位置づけられるものの中でも、止まると社会的にも影響を与えかねないレベルのサービスであった。そのようなシステム / サービスに対しては、極めて厚い障害対応への備えが求められる。

より詳細にこの事例を見ていくと、以下のような点が見受けられた。

- 何度も実績がある作業だから今回も大丈夫だろう、という判断があった。また製品の輸入代理店から「ライセンス更新作業は動作に影響を与えない」という情報も得ていた。しかし影響を与える結果となった。重要サービスに関係する作業に対しては、「実績のある作業であっても、想定外の事象は起こり得る」と考え方を変えて、二重三重の備えをしておくべきであった。
- システムのアラートが発生した場合は、システム設置拠点から運用管理担当者にメールで通知が届くフローになっていた。運用管理担当者への通知が一刻を争う場合、より迅速に伝えるため、また確実に伝わったことを確認するという意味でも、電話を使った緊急連絡という手段・体制を加えておくべきであった。
- この業種・業務に特化して実績のある製品は、海外製品しか選択肢がなかった。しかしそうであればサポート面の不足は避けられないため、国内製品の場合とは異なる特別な運用ルールや体制の整備を行うことにより、サポートが不足する部分を補っておくべきであった。「場合によっては自組織だけではなく、海外製品製造元と国内代理店ベンダとの間の連携体制にまで踏み込み、改善に取り組む必要がある」という意識にまでは至っていなかった。

これらは、従来から当たり前に実施していたやり方や考え方には不備があり、障害からの回復の長時間化のリスクが内在していた、ということを表している。重要サービスの提供にあたっては、通常サービスでは必要十分とみなされ、見落としがちな仔細な部分まで深く考慮し、徹底的にリスクを排除した運用環境を整備していく必要がある。

その際には、極力思い込みを排除し、「大丈夫」と判断する根拠に曖昧な点がないか、をチェックすることが重要となる。例えば上記に「輸入代理店から影響を与えないと聞いた」とあるが、なぜ影響を与えないと言えるのか、根拠が曖昧な点が残る。こうした曖昧な点がある場合、「そうではないかもしれない」「仮に影響を与えたらどういうことが起こるのか」と想定を変えて現状のルールの検証及び事前対策を実施することにより、影響を軽減できる可能性がある。重要サービスに対しては、そのような特段深い考慮に基づいたルールを整備することが求められる。

本事例から得られたことを中心に、教訓は「サービスの重要度を識別し、それに応じた連絡体制や障害検知の仕組みを作れ」としたが、リスクについてさらに深く考えることで、この 2 点以外の課題や改善策にも気づくことができるであろう。