Chapitre 12

La gestion de la disponibilité

Système d'assistance respiratoire, contrôle aérien, assistance au pilotage et électronique embarqué, distributeur automatique de billets, logiciel de gestion d'entreprise, site de commerce électronique… Ces exemples de la vie quotidienne illustrent bien le caractère parfois vital de la disponibilité des systèmes. Cette problématique se présente dans les mêmes termes dans le cas de l'entreprise. En effet, les progrès continus de la technologie et les gains de performance qui y sont associés ont considérablement modifié notre utilisation de l'informatique. Malheureusement, cela nous rend également de plus en plus dépendants de cet outil. En conséquence, la disponibilité de ces services est désormais un facteur économique différentiateur. Si une entreprise ne peut répondre aux commandes de ses clients, elle perd son marché.

En corollaire, on comprend aisément qu'une bonne gestion de la disponibilité a un impact non négligeable sur l'image du service informatique, ainsi que sur la satisfaction et la confiance des utilisateurs… Par extension, cela conditionne également l'image de l'entreprise sur le marché, ainsi que la satisfaction et la confiance de ses clients.

Vue d'ensemble

Ce processus est responsable de l'atteinte des objectifs négociés dans les contrats de service (SLA), c'est-à-dire qu'il a pour mission d'assurer un niveau de disponibilité compatible avec les impératifs métier tout en restant rentable financièrement.

Pour atteindre ce but, il est indispensable de déterminer précisément les besoins métier, ou « business », et en particulier ceux qui couvrent les fonctions vitales de l'entreprise. À ces *Vital Business Functions*(VBF), identifiées au sein du SLA, doit correspondre une réponse technique et organisationnelle.

La démarche de mise en place de ce processus passe donc par une optimisation de l'infrastructure informatique et du support du service.

La mise en place de ce processus dépend dans une large mesure de la qualité et de la complexité de cette infrastructure existante (fiabilité, redondance, maintenance, etc.), mais également, et surtout, de la qualité et de la maîtrise des processus et des procédures utilisés par les services informatiques.

Périmètre

Ce processus est responsable de la conception, de l'implémentation et du contrôle de la disponibilité des services informatiques.

En conséquence, il doit être appliqué à tous les services pour lesquels un SLA a été signé, qu'il s'agisse d'un service existant ou à venir. Dans le cas d'un nouveau service, la réflexion sur la disponibilité doit être menée dès le départ afin de minimiser l'impact de ce coût sur le service.

Par extension, le processus de gestion doit également être pris en compte pour des services dépourvus de SLA, mais dont le caractère critique au niveau informatique, peut entraîner des conséquences négatives pour le SLA d'autres services ou pour l'ensemble de l'entreprise. On peut citer par exemple le service de résolution de nom DNS qui, s'il est défaillant, peut gravement perturber l'accès à l'ensemble des ressources TIC de l'entreprise.

En revanche, ce processus n'est pas responsable de la gestion de la continuité de service du métier, ni de la résolution des incidents après un désastre majeur. Cependant, il fournit des éléments importants aux processus de gestion de la continuité de service, de gestion des incidents, ainsi qu'à la gestion des problèmes.

Une gestion de la disponibilité efficace peut faire la différence et sera reconnue comme pertinente par les directions métier si son déploiement au sein de l'organisation informatique est en phase avec les besoins du métier et des utilisateurs.

Trois grands principes doivent rester à l'esprit en permanence :

- La disponibilité est au cœur du business et de la satisfaction du client.

  La course à l'innovation dans les services ou les produits amène parfois le service informatique à se focaliser sur la technologie. Il risque alors de perdre de vue qu'il a pour but de satisfaire les besoins métier de l'entreprise.

  En conséquence, peu importe la technologie sous-jacente. Elle est inutile si les utilisateurs ne peuvent accéder à leurs données.

- Quand les choses vont mal, il reste toujours la possibilité de satisfaire le client et de remplir ses obligations.

Pour le service informatique, c'est le moment de vérité, et la manière de gérer et de résoudre la crise est l'élément principal qui sera retenu par les utilisateurs et la direction de l'entreprise pour juger de la qualité de leur organisation informatique.

Il est évident que la durée de la panne n'est qu'un élément de la réponse de l'organisation informatique. Il est vrai que cette valeur est un révélateur de la vitesse de réaction de l'organisation, mais elle doit être pondérée par son impact sur le métier.

- L'anticipation de la panne et l'utilisation de solutions de contournement identifiées et testées au préalable sont également des critères importants dans la vision que le métier retiendra de l'efficacité du service.

  Améliorer la disponibilité ne peut se faire sans comprendre comment l'informatique supporte le métier.

  La gestion de la disponibilité du système ne se limite pas à la compréhension technique de chaque élément du système, mais bien à sa relation avec le processus métier.

  L'optimisation de l'infrastructure informatique et de son support ne doit se faire que dans le but de délivrer le niveau de disponibilité requis par le métier.

Terminologie

La terminologie utilisée dans le processus de gestion de la disponibilité est présentée dans le tableau 13-1, où un système peut représenter le matériel, le logiciel, le service ou un composé des trois.

Disponibilité (Availability)

Indique la capacité d'un système à remplir son rôle pendant un temps donné. Cette valeur s'exprime le plus souvent sous la forme d'un pourcentage de temps pendant lequel le système est disponible par rapport à l'horaire de service négocié lors du SLA.

Fiabilité (Reliability)

Correspond à l'aptitude d'un système à fonctionner durablement avec un minimum d'incidents ou d'interruptions.

Capacité de maintenance (Maintainability)

Représente le potentiel de maintenance externe et indique la capacité d'un fournisseur externe à remettre un système en état de fonctionnement normal. La maintenabilité d'un composant particulier peut être décomposée en 7 étapes distinctes :

- l'anticipation de la panne ;
- la détection de la panne ;

- le diagnostic de la panne ;
- la résolution de la panne ;
- le retour à la normale après une panne ;
- la restauration des données et du service ;
- le niveau de maintenance préventive appliqué pour prévenir cette panne.

Résilience (Resilience)

Correspond à la faculté d'un système à continuer de fonctionner même si une partie des éléments du système est défaillante.

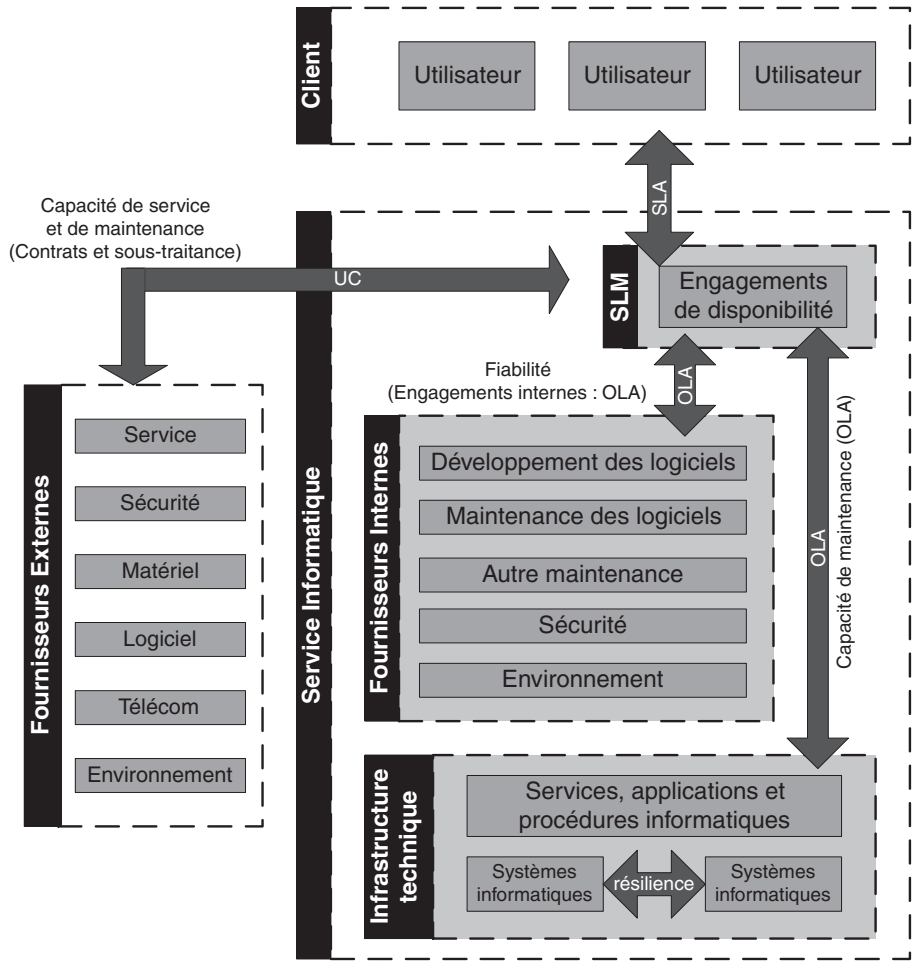

Figure 12-1 : Relations avec les organisations de support

Capacité de service (Serviceability)

Représente le potentiel de service externe et indique le niveau de service que l'on peut trouver auprès des fournisseurs tiers pour certaines parties de l'infrastructure informatique (exemple : contrat de support, contrat de maintenance).

Sécurité (Security)

Correspond à la mise en œuvre de contrôles visant à assurer un service continu en respectant des paramètres sûrs, à savoir :

- la confidentialité (interdire les accès non autorisés) ;
- l'intégrité (garantir la validité des données) ;
- la disponibilité (permettre l'accès aux données).

La disponibilité d'un élément de l'infrastructure informatique qui procure un service au métier et aux utilisateurs est influencée par :

- la complexité de l'infrastructure informatique et du service ;
- la fiabilité des composants et de l'environnement de l'infrastructure ;
- la capacité de l'organisation informatique à maintenir et à supporter cette infrastructure ;
- le niveau et la qualité de la maintenance apportée par les fournisseurs ;
- la qualité et l'étendue du déploiement des procédures.

Un moyen mnémotechnique comme un autre pour s'en souvenir : confidentialité (C), intégrité (I), disponibilité (A)... A pour Availability

Le service informatique est responsable devant le métier de la fourniture du système d'information. Cela implique que le niveau de disponibilité de chaque service doit être formalisé dans le SLA. De plus, ainsi que l'illustre la figure 12-1, les relations avec chaque fournisseur de service extérieur doivent également faire l'objet d'une documentation formelle.

Description du processus

La gestion de la disponibilité débute aussitôt que les exigences métier pour un service sont formalisées. C'est un processus permanent, se terminant uniquement lorsque le service disparaît ou n'est plus utilisé.

Flux du processus

Le flux interne du processus doit prendre en compte un nombre important d'informations en entrée, et fournir un volume également très important d'informations en sortie.

Mais globalement, la mission de la gestion de la disponibilité peut se résumer à définir les besoins exprimés par les directions métier, évaluer les risques et proposer des solutions y répondant, puis formaliser un engagement opérationnel, qui doit servir à la rédaction du SLA pour le service évalué (voir figure 12-2).

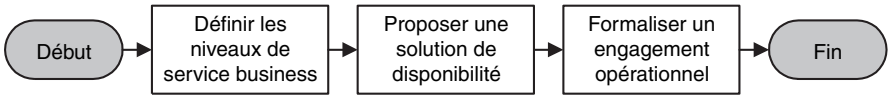

Figure 12-2 : Flux simplifié de la gestion de la disponibilité

Le périmètre de la gestion de la disponibilité couvre la conception, la mise en œuvre, la mesure et la gestion de la disponibilité de l'infrastructure. Cela implique qu'il existe un nombre important de liens entre ce processus et les autres processus ITIL. De plus, ces informations sont extrêmement diverses, ce qui complique largement la tâche du responsable de ce domaine, en lui imposant de disposer de compétences très différentes.

Parmi ces entrées, on trouve notamment des documents exprimant les besoins du métier et les risques associés à chacun de ces besoins. On trouve également un ensemble de rapports d'incident et de rapports de performance qu'il convient d'analyser avec attention. Enfin, s'ajoute à cela la connaissance fine de l'architecture, des technologies employées et des conséquences d'une défaillance technique sur un environnement métier.

La figure 12-3 reprend dans le détail les principales actions à prendre en compte dans le processus de gestion de la disponibilité.

La complexité relative de ce processus provient pour l'essentiel du nombre d'informations en entrée et en sortie interagissant les unes avec les autres.

De plus, il ne faut pas oublier qu'un lien fondamental a été retiré du diagramme afin d'en faciliter la lecture. Il s'agit de la boucle indiquant que ce processus est en perpétuelle réévaluation.

En résumé, les tâches que l'on peut identifier pour ce processus sont l'étude et la conception, l'implémentation et l'exploitation, et enfin le contrôle et la gestion.

La tâche d'étude et de conception regroupe les actions suivantes :

- analyser le besoin en terme de disponibilité, c'est-à-dire :
  - définir les besoins métier ;
  - définir les conditions optimales de disponibilité.
- analyser les tendances et les solutions proposées ;
- analyser les pannes fréquentes (gestion des problèmes) ;
- élaborer le plan de disponibilité (*Availability Plan*) ;
  - établir les niveaux de disponibilité convenus ;
  - établir les objectifs de chaque niveau ;
  - établir les plannings des interruptions (*DownTime*).
- rédiger les procédures et les documentations.

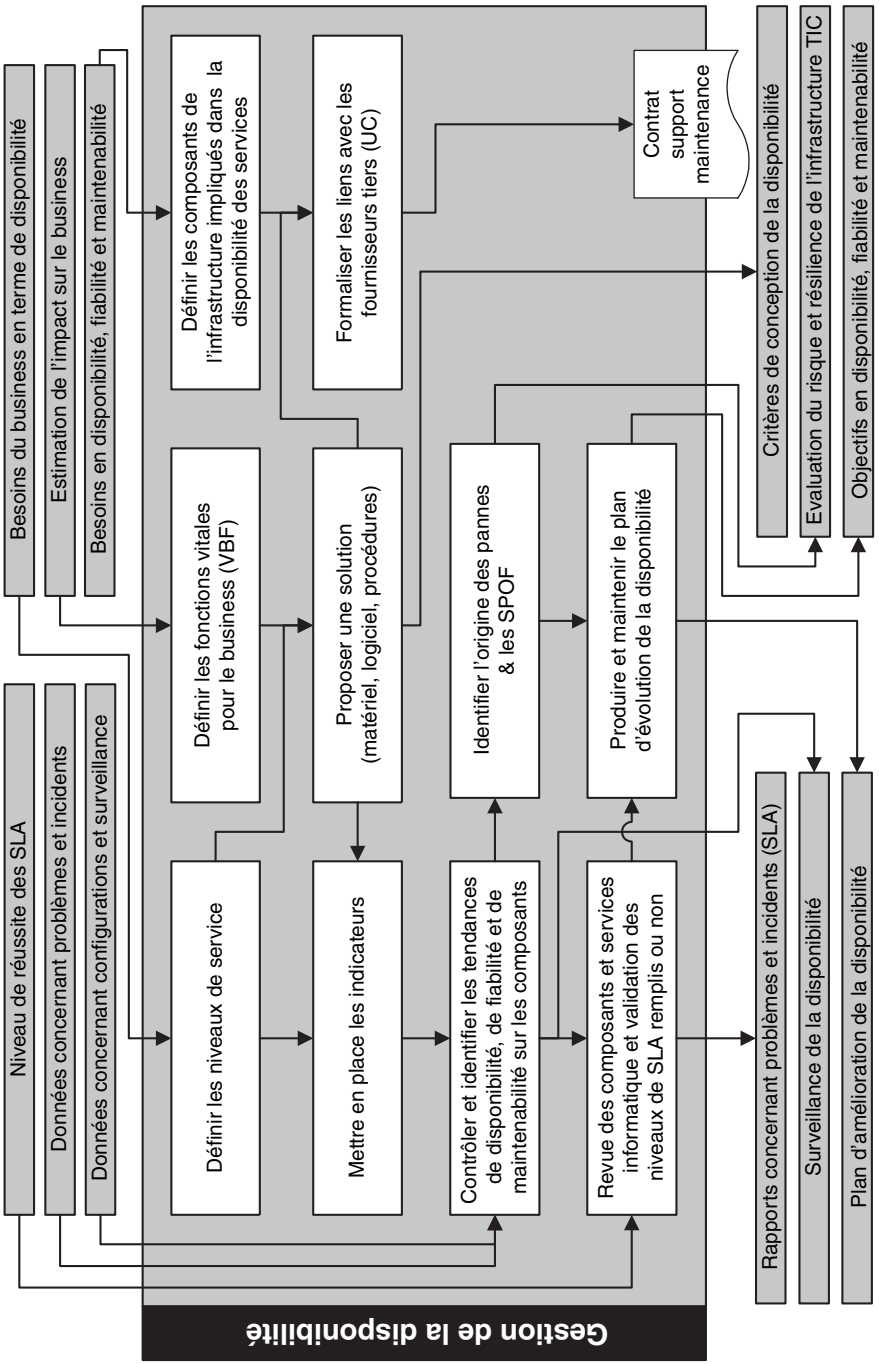

Figure 12-3 : Relations internes des activités du processus

La tâche d'implémentation et d'exploitation comprend les actions suivantes :

- assurer et optimiser les niveaux attendus de disponibilité en surveillant les éléments essentiels ;
- évaluer les options techniques et les procédures de retour au niveau précédant l'arrêt ;
- rechercher et éliminer les points uniques de panne (*Single Point Of Failure* ou SPOF) ;
- gérer la capacité de maintenance et de service.

Enfin, la tâche de contrôle et de gestion comprend les actions suivantes :

- s'assurer que les niveaux de service sont atteints par comparaison du réalisé avec les valeurs des SLA, et surveiller le respect des objectifs des OLA et les performances des fournisseurs tiers en matière de service (utilisabilité) ;
- collecter, analyser et maintenir à jour les données de disponibilité et produire des rapports sur ces données ;
- réexaminer et améliorer continuellement la disponibilité.

Documents

Le plan de disponibilité

Le plan de disponibilité (*Availability Plan*) correspond à une planification sur le long terme des buts, des objectifs et des documents de ce processus.

Ce planning définit les étapes et les moyens envisagés, tels que les ressources humaines, les processus, les outils et les techniques, dans le but d'améliorer de façon proactive la disponibilité des services informatiques tout en respectant les contraintes de coût imposées.

Comptes-rendus sur la partie disponibilité du SLA

Il s'agit de produire les tableaux, les mesures, les rapports, les commentaires et les explications sur les niveaux de SLA en exploitation.

Critères de disponibilité

Ces critères représentent des documents et des « normes » de conception à prendre en compte dans le domaine de la disponibilité et du « retour à la normale » lors de la mise en place de nouveaux services ou la modification de services existants.

Améliorer la disponibilité

La capacité du processus de gestion de la disponibilité est fortement influencée par la gamme et la qualité des méthodes et des techniques disponibles pour son déploiement et son exploitation.

Dans l'absolu, il convient d'appliquer des méthodes simples pour y parvenir.

On peut tout d'abord réduire le temps d'arrêt (*downtime*) en procédant aux actions suivantes :

- réduire le temps de détection par la supervision du système et la mise en place d'alarmes ;
- réduire le temps de réaction par la gestion des incidents et la mise en place de procédures ;
- réduire le temps de réparation en disposant de configurations standards, de dossiers système à jour et de matériels de rechange disponibles (*spare*) ;
- réduire les temps de restauration (procédures, sauvegardes, données d'origine, etc.).

On peut ensuite réduire la fréquence ou l'impact des incidents par les actions suivantes :

- investir dans des systèmes à tolérance de panne (améliorer la résilience) ;
- réaliser des maintenances préventives permettant d'éviter un grand nombre d'incidents (gestion de la capacité, gestion des configuration, etc.) ;
- mettre en place la duplication des services ou des environnements, de façon synchrone ou asynchrone en fonction de la criticité du service.

Enfin, pour réduire les arrêts de service on peut : planifier les interventions préventives en dehors des horaires de service (gestion du changement, gestion des mises en production).

Il existe plusieurs techniques et méthodes associées à la gestion de la disponibilité. Parmi celles que préconise ITIL, on trouve notamment les méthodes suivantes :

- CFIA : évaluation de l'impact de la défaillance d'un composant.

  La CFIA peut être utilisée pour prévoir et évaluer l'impact sur le service résultant de la défaillance d'un ou plusieurs composants de l'infrastructure, et en particulier pour identifier les endroits où une infrastructure redondante permettrait de réduire au minimum cet impact.

- FTA : analyse d'arbre des fautes.

  La FTA est une technique qui peut être utilisée pour déterminer la chaîne des événements qui entraîne une rupture des services informatiques. En conjonction avec des méthodes de calcul, elle peut offrir les modèles détaillés de disponibilité.

- CRAMM : méthode de gestion et d'analyse du risque.

  CRAMM est utilisé pour identifier de nouveaux risques et fournir les contre-mesures appropriées lors de l'évolution des besoins de disponibilité exprimés par le métier.

- SOA : analyse des défaillances système.

  SOA est une technique conçue pour fournir une approche structurée à l'identification des causes sous-jacentes d'interruption de service.

Mesures de la disponibilité

Disponibilité des applications *vs* disponibilité des données.
Il est intéressant de noter que selon le métier ou la fonction que l'on considère, l'importance que l'on accorde à la disponibilité des applications/services est plus ou moins forte. Il en est de même pour la disponibilité des données.
Exemple : la disponibilité des données transportées est moins importante que la disponibilité du service dans le cas de la voix sur le protocole Internet, où on accepte de perdre quelques paquets. En revanche, l'indisponibilité d'une application comptable est moins critique que l'impossibilité d'accéder aux données, ou le fait que celles-ci soient altérées...
Tout cela est à prendre en compte lors de l'établissement du SLA.

Les comptes-rendus destinés au service commercial ne peuvent pas être les mêmes que ceux qu'utilise en interne le support informatique. Afin de satisfaire ces différents points de vue, il est nécessaire de proposer un ensemble de métriques qui reflètent ces perspectives concernant la disponibilité des composants et des services informatiques. Tant sur le fond que sur la forme, les informations présentées sur le rapport doivent être adaptées au public visé.

Le modèle de mesures de la disponibilité informatique (ITAMM, IT Availability Metrics Model) regroupe un ensemble de métriques et de perspectives, qui peuvent être utilisées dans la réalisation des mesures de disponibilité, puis dans la publication des rapports associés.

Il est nécessaire de déterminer ce que l'on désire inclure dans ce type de rapport en fonction du destinataire de ce document. La liste suivante présente les paramètres importants à prendre en compte dans la mise en place de ce modèle.

- Fonctions métier critiques (VBF)

  Mesures et rapports démontrant les conséquences (contribution et impact) de la disponibilité du système d'information sur les fonctions métier importantes d'un point de vue opérationnel.

- Applications

  Mesures et rapports concernant la disponibilité des applications ou des services nécessaires au fonctionnement opérationnel du métier.

- Données

  Mesures et rapports concernant la disponibilité des données nécessaires au fonctionnement opérationnel du métier.

- Réseau

  Mesures et rapports concernant la disponibilité du réseau et sa capacité de transport effective.

- Composants clés

  Mesures et rapports concernant la disponibilité, la fiabilité et la capacité de maintenance des composants de l'infrastructure TIC fournis et maintenus par le support informatique.

- Plate-forme

  Mesures et rapports concernant la plate-forme qui supporte le fonctionnement des applications métier.

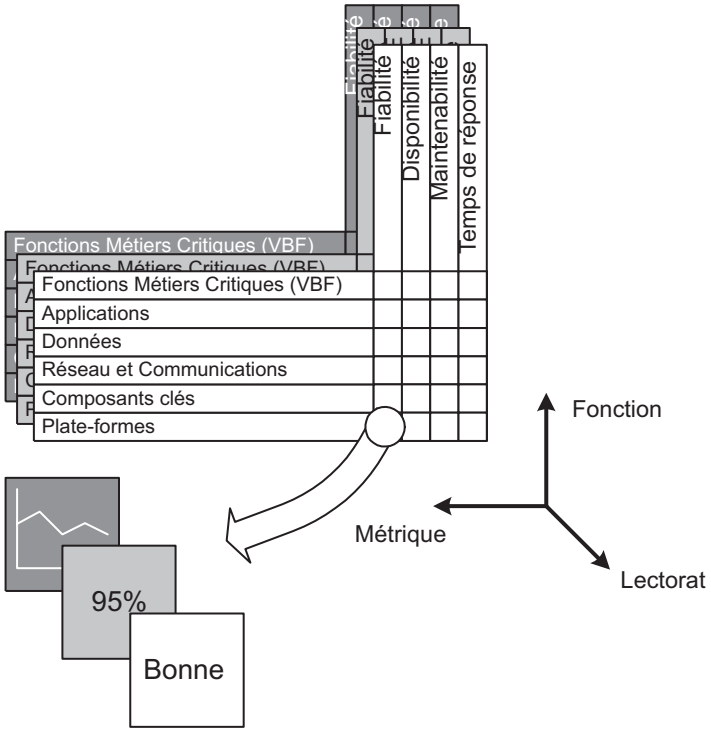

Figure 12-4 : ITAMM, modèle de mesures de la disponibilité des technologies de l'information

Comme l'indique la figure 12-4, chaque case à l'intersection d'une ligne et d'une colonne représente une métrique du modèle ITAMM. En fonction du public visé, chaque métrique peut être représentée comme une valeur discrète communiquée sous la forme d'une moyenne, d'un pourcentage (80 %), ou d'une appréciation globale (très bonne/bonne/moyenne...). Cependant, elle peut également être présentée sous la forme d'un sous-tableau ou d'un graphique plus complet mais moins rapide à lire et à interpréter. Le modèle ITAMM représente donc une mesure de disponibilité à trois dimensions qui sont la fonction, la métrique et le lectorat visé par la mesure.

Pour fournir une vue équilibrée et significative de la disponibilité d'un service ou d'un composant, le rapport doit, pour chaque élément évoqué ci-dessus, considérer les dimensions suivantes :

- Disponibilité.

  Mesures et rapports exprimant le taux de disponibilité en fonctions des objectifs du SLA.

- Fiabilité.

  Mesures et rapports exprimant la fréquence des pannes.

- Maintenabilité.

  Mesures et rapports exprimant la durée des pannes (c'est-à-dire le temps de reprise total depuis la détection jusqu'au retour du service pour l'utilisateur.

- Temps de réponse.

  Mesures et rapports exprimant le niveau de performance du système d'un point de vue utilisateur.

Calcul de la disponibilité

Du point de vue technique, la mesure de la disponibilité s'établit sur la base du ratio entre les temps des disponibilités réelles et convenues sur une période donnée (annuelle, mensuelle, hebdomadaire, etc.) :

$$\text{Disponibilité (en pourcentage)} = \frac{(\text{AST} - \text{DT})}{\text{AST}} \times 100$$

AST = A*greed Service Time* = temps de service convenu.

DT = D*own Time* = durée d'interruption du service.

Exemple : soit un SLA indiquant une période de disponibilité de 8 heures sur 5 jours sur le service A.

Si A est indisponible pendant 4 heures sur la semaine, la disponibilité est d(A) = (40 - 4) / 40 _ 100 = 90 %.

Cela semble simple mais en réalité tout dépend de ce que l'on mesure et comment on le mesure. En effet, si on envisage que l'indisponibilité du service ci-dessus ne s'est manifestée que sur une dizaine d'utilisateurs parmi 1 000 personnes, doit-on toujours considérer une disponibilité de 90 % pour l'ensemble de l'entreprise ?

De plus, et en dehors de toute recherche de responsabilité, il convient de valider que la valeur publiée correspond bien au « ressenti utilisateur » et pas uniquement à une mesure issue d'un calcul.

C'est le cas, par exemple, quand une application distante n'est plus accessible du fait d'une panne de télécommunication dont la responsabilité ne se trouve pas au sein de l'équipe informatique interne. Pour l'équipe interne, tout se déroule bien, alors que l'utilisateur ne peut plus travailler.

Un service informatique est indisponible pour un client lorsque la fonction requise à un horaire et en un lieu ne peut être utilisée dans les termes du SLA.

Dans la situation où plusieurs composants liés sont impliqués dans une indisponibilité, le calcul doit prendre en compte l'aspect séquentiel ou parallèle des ressources sur l'ensemble de l'infrastructure concernée.

Prenons le cas d'un service commercial réalisant des commandes par téléphone. Une base de données est hébergée sur un serveur distant, et une application de saisie est présente sur les postes de travail. Pour atteindre

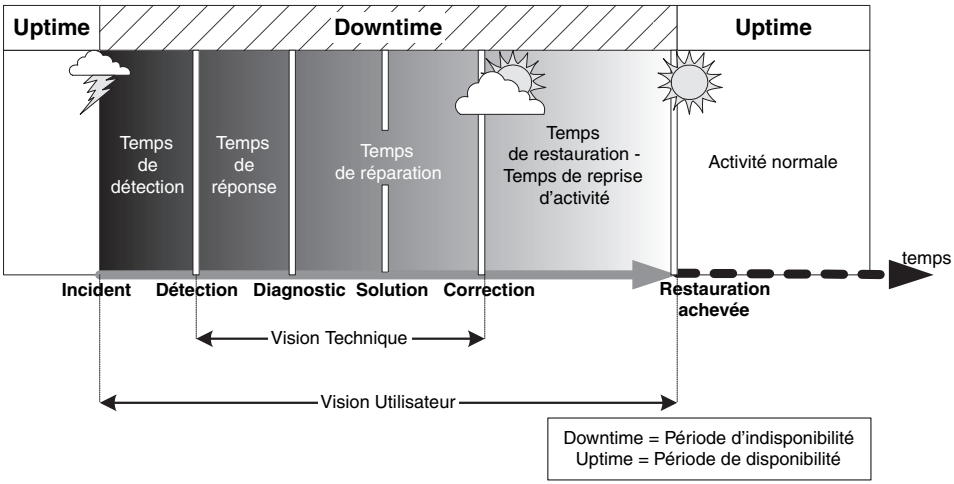

Figure 12-5 : Perception des différentes étapes de la panne

le serveur depuis le poste de travail, l'application transite via le réseau. C'est le cas typique d'une chaîne de traitement en série.

Si le serveur affiche un taux de disponibilité de 95 % mais que le réseau n'atteint en revanche que 90 % alors la disponibilité de l'ensemble ne représente plus que $0.95 \times 0{,}90 = 85{,}5$ %. Sur une semaine, dans le cadre d'un SLA réclamant une disponibilité de 8 heures sur 5 jours, le serveur est indisponible moins de 2 heures, alors que l'utilisateur ne peut travailler pendant près de 6 heures… Une journée de perdue.

Encore une fois, on comprend que la qualité d'une chaîne est celle de son maillon le plus faible. Ce qui explique l'importance de ce concept dans le ressenti de l'utilisateur, et qui permet d'insister sur la prise en compte de la disponibilité de chacun des éléments de la chaîne dans la vision de la disponibilité globale du système pour l'utilisateur.

Afin d'améliorer la disponibilité du service, une des solutions consiste à dupliquer certaines ressources afin de paralléliser les accès. Imaginons que la ressource réseau soit désormais équipée d'un système de secours qui propose un second passage. La disponibilité des deux équipements n'est pas parfaite, mais nous obtenons à présent 95 % pour le premier et 90 % pour le second. Dans ce cas la règle de calcul est obtenue par la durée de l'indisponibilité globale, ce qui donne :

Disponibilité du système complet = 1 – indisponibilité du système A × indisponibilité du système B.

On obtient alors un taux de disponibilité de $1 - (1 - 0{,}95) \times (1 - 0{,}90) = 99{,}5$ %.

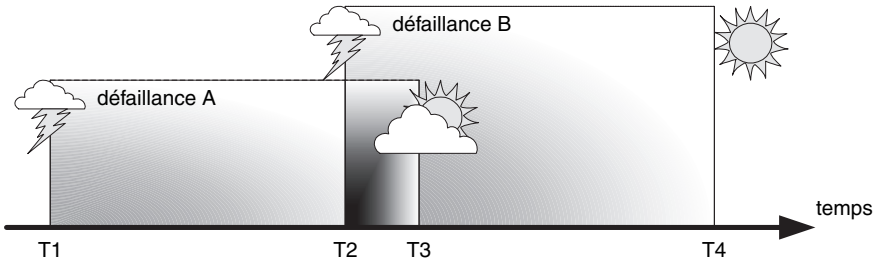

Figure 12-6 : Défaillance conjointe de deux équipements redondants

Attention, nous ne prenons pas en compte le fait que les deux équipements peuvent subir une défaillance simultanément ou en léger décalage. Dans ce cas, ce calcul est incorrect.

Dans l'exemple de la figure 12-6, on constate que l'équipement A subit une défaillance à T1, mais que le service est toujours disponible puisque l'équipement B prend le relais. Malheureusement, à T2, l'équipement B tombe également en panne alors que A n'est toujours pas disponible puisque la restauration des données se termine à T3…

Dans ce cas, le calcul de disponibilité du service d'un point de vue technique est sensiblement différent des deux calculs précédents. En effet, on peut considérer que pendant les phases T1 à T2 et T3 à T4 les équipements se portent secours mutuellement. Pendant ces périodes, le calcul des systèmes en parallèle s'applique. Mais durant l'intervalle T2 à T3, où plus aucun service n'est disponible, c'est bien le calcul des systèmes en série qui s'applique.

Il est aisé de comprendre qu'un rapport qui présente des taux de disponibilité sans autres explications et dans lequel ne sont pas précisées les règles de calcul, ne présente pas un grand intérêt. De ce point de vue, l'idéal est de fixer ces règles de calcul et d'interprétation dans le SLA.

D'un point de vue présentation, la méthode utilisée et les chiffres retenus ont également leur importance.

Si nous considérons un service stratégique pour l'entreprise et sur lequel un SLA de disponibilité à 99 % est signé, nous avons trois possibilités de présentation de la valeur mesurée :

- Pourcentage de disponibilité : 98,8 %.

  Ce chiffre, qui doit être le plus grand possible, donne le pourcentage de temps où le service est disponible durant la période considérée. Souvent assénée comme une vérité absolue, cette valeur peut induire en erreur et donner un sentiment de confiance excessive chez le personnel informatique. La différence entre 98,8 et 99 % ne semble guère importante.

C'est oublier bien vite que pour certaines entreprises, un dixième équivaut à des millions d'euros de pertes.

- Pourcentage d'indisponibilité : 1,2 %.

  Cette valeur qui met l'accent sur l'indisponibilité semble plus efficace dans la prise en compte par l'équipe informatique. Paradoxalement, le besoin d'une baisse de la valeur de quelques dixièmes de cet indice est mieux vécu par les personnes chargées de sa mise en œuvre que le besoin d'une hausse dans le cas précédent.

- Durée de l'indisponibilité : 105 heures par an en moyenne ; 9 heures par mois en moyenne ; 2 heures par semaine en moyenne.

Cette méthode de publication est de loin la plus intéressante pour les utilisateurs à qui elle « parle » vraiment. En effet, il est facile de mettre un coût en face de cette durée d'indisponibilité.

Cependant, pour les utilisateurs, l'indice de disponibilité le plus pertinent est une valeur composée de trois facteurs qui sont la fréquence, la durée et l'ampleur de la panne ou l'impact sur le métier.

C'est d'ailleurs l'impact qui doit être utilisé dans une communication de la disponibilité vers la direction générale ou les directions métier. Cet impact est plus complexe à calculer et est en partie fonction de l'importance que l'on donne à la disponibilité d'un service dans un métier (exemple : disponibilité de l'outil CRM de prise de commande dans un service commercial).

La valeur de l'impact de la disponibilité doit être communiquée en positif, et on parle alors de valeur ajoutée ou de création de valeur, ou en négatif, dans le cas où le niveau de SLA n'est pas atteint.

Enfin, il est indispensable de déterminer correctement les périodes à prendre en compte dans l'évaluation du taux de disponibilité. En effet, les événements tels que l'indisponibilité planifiée pour la réalisation d'une maintenance, ou le délai de débordement post-maintenance utilisé pour terminer les actions planifiées précédemment, sont des périodes que l'on doit distinguer de la période classique d'indisponibilité, c'est-à-dire la durée pendant laquelle un service défaillant n'est plus disponible.

Coûts

La gestion de la disponibilité est en permanence confrontée au dilemme entre le coût de l'investissement et le coût associé au risque en cas de défaillance du système d'information.

À l'image d'une police d'assurance que l'on peut souscrire ou non, il est du ressort de la direction de l'entreprise de décider de la somme qu'elle est prête à consacrer à la sauvegarde de son activité. Après avoir évalué les pertes potentielles, il est théoriquement assez simple d'estimer de quel budget peut disposer le processus de gestion de la disponibilité dans l'entreprise.

Figure 12-7 : Le fragile équilibre entre le coût des investissements et le coût du risque

Le coût de la disponibilité

Puisque le taux de disponibilité d'un système équivaut à la disponibilité de son maillon le plus faible, la méthode la plus classique pour augmenter ce taux est de fiabiliser chaque maillon en le dupliquant ou en appliquant une procédure de fiabilisation adaptée. On réduit ainsi le nombre de SPOF (points sensibles) du système. Le revers de la médaille est que l'on risque fort de doubler le coût de l'infrastructure. On peut également appliquer des procédures spécifiques permettant de s'affranchir de la défaillance d'un composant quelconque du système d'information. C'est le cas des procédures palliatives dans lesquelles sont expliquées les méthodes de contournement.

Mais, en observant la figure 12-8, on constate qu'au-delà d'un certain seuil, l'investissement supplémentaire n'est plus autant valorisé que dans les débuts de la courbe. Celle-ci, de forme logarithmique, nous indique que passé

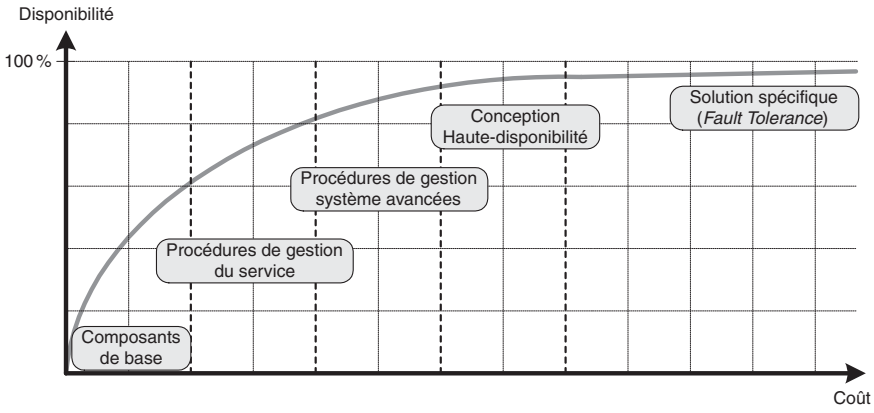

Figure 12-8 : Relation entre coût global et niveau de disponibilité

un certain niveau (le palier) chaque dixième de point de disponibilité gagné le sera au prix d'un investissement matériel ou humain de plus en plus important.

Cependant, cette courbe a aussi un grand intérêt. Elle confirme que l'investissement de base permettant d'atteindre une bonne disponibilité est loin d'être inaccessible. Cela implique qu'avec un peu de bon sens dans la conception de l'infrastructure et en mettant en place des procédures adaptées le plus tôt possible, tout système d'information peut atteindre des taux de disponibilité très respectables, sans investissement excessif par ailleurs. Rappelons également que le coût de la disponibilité est moindre lorsque l'étude et l'implantation se font en début de projet.

Le coût de l'indisponibilité

Les données du coût de l'indisponibilité sont calculées à partir des informations provenant de la gestion de la continuité de service qui détermine les fonctions vitales pour le métier, de la gestion des niveaux de service qui exprime les besoins métier de disponibilité sur ces fonctions vitales, et la gestion financière qui valorise ces demandes en termes monétaires.

L'indisponibilité du système engendre des coûts divers qui peuvent être des coûts immédiats et tangibles, mais également diffus ou différés, et par nature intangibles.

Coûts tangibles

Les coûts tangibles sont ceux que l'on peut calculer aisément. Ils représentent une part importante des pertes provoquées par une indisponibilité du système :

- perte sur les salaires des employés qu'il faut verser alors que ceux-ci ne produisent plus ;
- perte de productivité des utilisateurs qu'il faut bien rattraper par la suite ;
- perte de revenus directe si l'activité touchée génère un revenu immédiat ;
- perte de matière première ;
- pénalités de retard ;
- pertes financières sur les fluctuations du cours de l'action.

Coûts intangibles

Les coûts intangibles sont bien plus complexes à évaluer. Mais il n'en reste pas moins qu'il s'agit de sommes importantes en termes de manque à gagner. Il est fondamental de ne pas faire l'impasse sur ces coûts « cachés » :

- perte de confiance des clients envers une société à la fiabilité douteuse ;
- perte de confiance des analystes et des marchés financiers ;
- perte directe des clients qui se dirigent alors vers la concurrence ;

Les pertes de revenus directes se produisent par exemple dans les caisses des magasins de détail, ou dans celles de la restauration. Les pertes de matière première concernent les matières périssables, l'industrie plastique et chimique, l'imprimerie, etc.

- perte d'opportunités, de marchés, de contrat, etc. ;
- problème de réputation à long terme sur le marché ;
- perte de confiance des utilisateurs envers le service informatique.

Scénario

L'ERP d'une entreprise industrielle de 1 500 salariés réalisant un CA annuel de 200 millions d'euros est indisponible pendant 5 heures. La production est stoppée, les commandes ne passent plus, les livraisons ne peuvent pas partir à l'heure et sont retardées de plusieurs jours... Le salaire moyen dans l'entreprise est de 20 €/heure. Le CA horaire moyen par employé est de 83 €.

Les pertes directes en salaire sont de 150 k€, alors que les pertes en CA représentent plus de 620 k€.

Ces données issues d'un cas réel sont édifiantes. N'oublions pas qu'elles ne représentent qu'une partie des pertes potentielles, puisque les pertes en matières premières ne sont pas indiquées et que les pertes intangibles sont également à prendre en compte, dont notamment l'impact sur la confiance des clients.

Ces 770 k€ (ou au moins une partie) ne seraient-ils pas mieux employés dans le budget de la gestion de la disponibilité ?

Conséquences du processus de gestion de la disponibilité

Les conséquences du processus sont multiples. Mais ce que l'on attend en premier lieu de ce processus est bien une amélioration en terme de satisfaction client ou tout du moins, que la disponibilité des services ne soit pas un facteur d'insatisfaction.

Avantages

Parmi les conséquences de la mise en place de la gestion de la disponibilité, les avantages sont nombreux et parfois gratifiants.

En premier lieu, cela officialise un point unique de responsabilité pour cette lourde tâche. Jusqu'alors, chaque responsable d'application, d'infrastructure ou de service devait assumer la responsabilité de la mise en place de la disponibilité. Ce qui est loin d'être simple et n'est pas nécessairement dans les compétences de la personne. Mais le pire est qu'ensuite, le suivi et la responsabilité de cette disponibilité incombait au responsable de l'exploitation qui n'avait souvent pas participé à l'élaboration du service…

En second lieu, ce processus permet de répondre aux besoins élémentaires exprimés par le métier en termes de disponibilité. Ce qui implique que la vision de son implantation évolue d'un point de vue technique vers une vision plus centrée sur le métier. La conséquence est que les services sont conçus et gérés pour répondre de manière optimale sur le plan financier à des besoins spécifiques de l'entreprise.

De plus, on découvre rapidement une évolution de l'esprit du personnel informatique, d'une attitude réactive à une attitude proactive avec en corollaire la réduction de la fréquence et la durée des « pannes » informatiques.

L'organisation informatique est alors perçue comme « apportant de la valeur » au métier.

Problèmes possibles

Les principales difficultés de ce processus proviennent essentiellement du volume et de la complexité des informations en entrée.

Le problème le plus fréquent est la difficulté à déterminer les besoins en disponibilité du métier et du client, c'est-à-dire traduire sur le terrain l'expression d'une demande en architecture, en procédures et en métriques. Malheureusement, cela s'accompagne inévitablement par la conception d'une infrastructure décalée en regard de l'attente du client, entraînant parfois des surcoûts inutiles. Mais cela peut également produire des mesures de disponibilité qui ne présentent pas d'intérêt pour le métier ou pire, une absence totale de mesure concernant la fourniture du service en question.

Le second problème réside dans la difficulté à trouver les moyens nécessaires à la mise en place du processus, tel, que les personnes qualifiées et expérimentées, ou les logiciels appropriés. Cela peut d'ailleurs entraîner une certaine dépendance vis-à-vis des fournisseurs avec d'éventuelles conséquences sur la qualité de leur service, comme la fiabilité et la maintenabilité.

Enfin, lorsque le processus n'est pas correctement expliqué ou interprété, la justification des dépenses auprès de la direction semble être le « souci » le plus fréquent et s'accompagne inévitablement d'un manque d'engagement de celle-ci.

Rôles et compétences du responsable du processus

Le rôle principal de ce responsable est de définir et de déployer la gestion de la disponibilité dans l'entreprise. Pour cela, il doit s'assurer que les services informatiques sont conçus pour livrer les niveaux requis de disponibilité demandés par le métier.

Il doit identifier les procédures ou les éléments de l'infrastructure susceptibles de fragiliser les services informatiques et doit alors apporter les actions préventives adaptées.

Le responsable doit également définir les métriques indispensables au contrôle du processus.

Il doit fournir une gamme de rapports permettant en premier lieu de valider que les niveaux de disponibilité, de fiabilité et d'aptitude à la réparation sont bien remplis (SLA), mais permettant également de valider l'optimisation de l'infrastructure afin d'apporter des améliorations financièrement avantageuses pour l'entreprise.

Après analyse des informations, il doit également identifier les pannes et apporter l'action corrective visant à réduire la durée et la fréquence de ces pannes.

Enfin, il doit créer et maintenir le plan de disponibilité destiné à mettre en phase les services informatiques et les composants de l'infrastructure avec les besoins actuels et futurs de l'entreprise.

Compétences clés

Le responsable doit avoir une bonne expérience de la gestion des processus ainsi qu'une bonne compréhension des « disciplines » ITIL.

Il doit avoir une bonne compréhension de principes statistiques et analytiques.

Il doit posséder de bonnes compétences pour la communication écrite et orale, ainsi que des compétences dans le domaine de la négociation.

Il doit bien sûr avoir une bonne compréhension des concepts de la disponibilité et de ces techniques, et comprendre comment les technologies « informatiques » supportent le métier.

Enfin, il doit avoir une compréhension raisonnable de la gestion des coûts.