Continuité d'activité : le choix des armes

Beaucoup de Plans de Recouvrement d'Activité (PRA) furent conçus dans le but de parer à des désastres tels que les incendies, les inondations ou encore des pannes électriques prolongées. Il s'agissait d'une assurance contre des évènements aussi graves qu'improbables pour les entreprises souscripteurs. Les plans de Continuité d'Activité (Plan de Continuité de Service), initiés par les banques et les acteurs des marchés financiers ont pris le relais des PRA de base : il ne s'agit plus seulement de concevoir des infrastructures redondantes, mais également de concevoir des procédures qui réduisent dans le pire des cas les interruptions de service à quelques minutes. La concentration des infrastructures dans les grandes entreprises sur des serveurs toujours plus puissants, la puissance « à la demande » et les réseaux de stockage, au détriment du modèle client serveur et des serveurs distribués facilite d'ailleurs grandement la mise en place de PCS.

Les procédures de continuité d'activité appartiennent rarement exclusivement aux informaticiens. Constater la dépendance à l'informatique des entreprises relève du lieu commun. On trouvera naturellement un volet technique très important dans la mise en place de plans de continuité de services car les technologies fleurissent en ce domaine : réplications de bases de données et des matériels de stockage, clusters, miroirs de volumes, appliances, snapshots de baies, etc. La mise en correspondance des risques et des parades requiert une bonne connaissance du panel des offres du marché. Une fois exploré, ce maquis de solutions techniques permet des couvertures de risques hétérogènes, adaptées à la majorité des situations.

Du Plan de Recouvrement d'Activités au Plan de Continuité de Service

Le PRA traditionnel s'adresse principalement à la conservation et à l'intégrité des données. Le développement des réseaux de stockage et les possibilités de miroirs offerts par les baies de stockage et par les appliances spécialisées dans la réplication sur IP ont popularisés dans les grandes entreprises la réplication synchrone sans perte de données. Le PCS englobe certes le PRA, mais il impose par ailleurs une forme d'automatisme dans la mise à disposition des bases de données ainsi que des procédures adaptées à une remise en service dans des délais ne dépassant pas quelques minutes, voire quelques secondes.

Un PCS classique (figure 1) est composé de quatre volets qui traitent chacun une catégorie d'évènements : désastres physiques, incidents matériels graves mais localisés, incidents logiciels entraînant une indisponibilité des applications, incidents externes (chez un fournisseur par exemple).

matières de dysfonctionnement afin de prioriser les investissements. Cette classification a le mérite d'identifier d'une part les applications indispensables à l'activité de l'entreprise (centre d'appel des régies publicitaires, connexion à la bourse pour des salles de marchés, etc.), mais également celles dont l'indisponibilité entraînerait une perte d'image ou de clientèle (achats en ligne). Chaque groupe d'applications se voit autoriser un temps maximal d'indisponibilité, une tolérance à la perte de données, éventuellement une classe de risque à couvrir. Cette phase est souvent délicate à mettre en œuvre dans les grandes structures car les visions des acteurs tendent à diverger. La phase d'architecture permet ensuite de choisir les technologies qui seront mises en œuvre à l'issue de la première phase. Pour certaines applications, les exigences initiales sont d'ailleurs fréquemment révisées à la baisse lorsque les montants des investissements paraissent trop élevés.

Après réalisation, un PCS ne restera viable que s'il est régulièrement maintenu. Les configurations de clusters doivent par exemple être maintenues à jour lors de l'ajout de volumes.

	Désastres	Incidents matériels localisés	Indisponibilité applicative	Autres incidents
Périmètre	Interruption illimitée de l'activité de l'entreprise	Interruption temporaire de l'activité de l'entreprise	Système, bases de données, bogues, erreurs humaines	Evènement externe imprévisible
Cible à traiter	Indisponibilité illimité d'un composant critique	Indisponibilité temporaire d'un composant critique	Dysfonctionnement d'une application	Dépendances externes de l'activité de l'entreprise
Exemples d'évènements	Feu, inondation, panne d'un serveur consolidé	Panne de Backbone réseau	Corruptions de bases de données, dégradations importantes des performances	Interruption des flux d'informations financières dans une salle de marchés
Exemples de parades	Transfert de l'activité sur un autre site (PRA)	Basculement sur un réseau de secours	Procédure manuelle	Disponiblité d'une seconde source d'information

Figure 1 : Les quatre composants d'un plan de continuité de service

La mise en oeuvre d'un PCS passe d'abord par une phase d'analyse de risque (figure 2) ou par la production d'un cahier des charges de continuité de service qui permet surtout de grouper et de hiérarchiser les besoins des applications en

Enfin, le leitmotiv sera « test, test, test ». Quelque soit la qualité de la mise en place initiale et celle de la maintenance, les tests font régulièrement apparaître des oublis qui compromettent parfois gravement les objectifs affichés.

Catégories	Fonctions	Niveau de service requis
Criticité 1	• Interface clientèle • Applications indispensables à l'activité (réservations aériennes, achats en ligne, etc.)	• 24x5 ou 24x7 • 5'/mois d'arrêt • Aucune perte de données
Criticité 2	• Approvisionnements • Distribution	• 24x5 • 1h à 4h d'arrêt • Pertes de données < 1h
Criticité 3	• Middle Office • Back office	• 18x5 • 4h à 10h/mois d'arrêt • Pertes de données < 4h
Criticité 4	• Ressources humaines • Gestion de projets	• 12x5 • 1 à 3 j/mois d'arrêt • Pertes de données < 12h

Figure 2 : Contrats de services HD par types d'applications

Les niveaux de services requis par classes d'applications fluctuent. Une comptabilité est beaucoup plus sensible aux interruptions de services lors des arrêtés comptables qu'en début de mois. Une application peut devenir critique à la suite d'un changement de périmètre. En bref, un PCS est voué dès sa conception à être fréquemment remanié.

De quoi veut-on se protéger ?

La réponse à cette question conditionne l'organisation du système de haute disponibilité. Un système d'envoi d'ordre électronique dans une salle de marché ne souffre pas d'arrêt, et la sauvegarde/restauration des bases de données telles qu'elles se présentaient 10' avant le désastre n'a que peu d'intérêt. En revanche, certaines applications sont plus exposées aux erreurs des utilisateurs (suppressions ou altérations de données après une clôture comptable par exemple), voire des administrateurs (suppression d'un fichier de bases de données). La première catégorie sera protégée par un cluster, la seconde par un mécanisme de réplication de bases de données différé dans le temps (réplication H-1 par exemple). Dans la pratique, plusieurs mécanismes sont souvent utilisés en parallèle. Une réplication « Data Guard » (pour Oracle) ou « log shipping » (pour MS Sql Server) alimente une base de données maintenue avec un décalage dans le temps, tandis qu'un cluster reposant sur une réplication synchrone entre baies distantes permet un basculement sans perte de données en cas de sinistre sur le site primaire. Les serveurs applicatifs (2nd tiers), ne stockant pas de données, disposent d'un système d'équilibrage de charge qui assure de facto la redondance des accès.

Des outils adaptés

Les exigences des responsables de périmètres fonctionnels et les outils disponibles peuvent être positionnés sur deux axes (figure 3). Le premier caractérise le temps d'indisponibilité, le second la perte de données – sur le risque couvert - exprimée sur une échelle de temps.

Clusters et réplications hardware (stockages et appliances)

Les clusters locaux ou géographiques permettent de minimiser le temps de basculement des applications à quelques secondes (1 à 2 minutes au maximum) sans perte de données, mais tous les clusters n'ont pas le même niveau de cou-

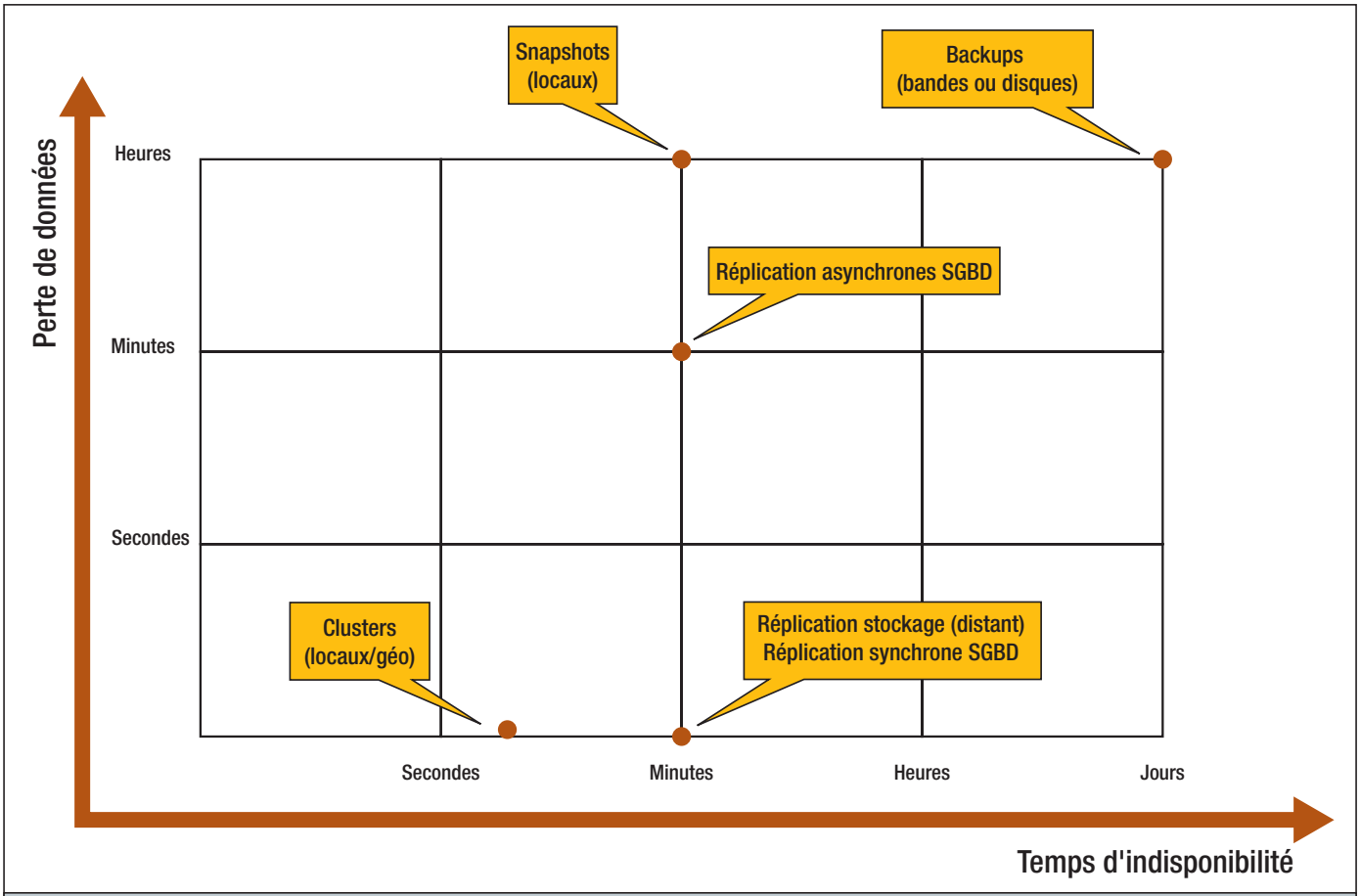

Figure 3 : Des technologies adaptées au couple temps d'indisponibilité/temps de recouvrement

verture du risque. En schématisant, on trouve d'une part les clusters locaux (N machines accèdent aux même disques) et d'autre part les clusters distants ou géographiques (figure 4) : machines et ressources de stockage sont dupliquées sur des sites différents, les disques étant mis en miroir d'un site vers l'autre.

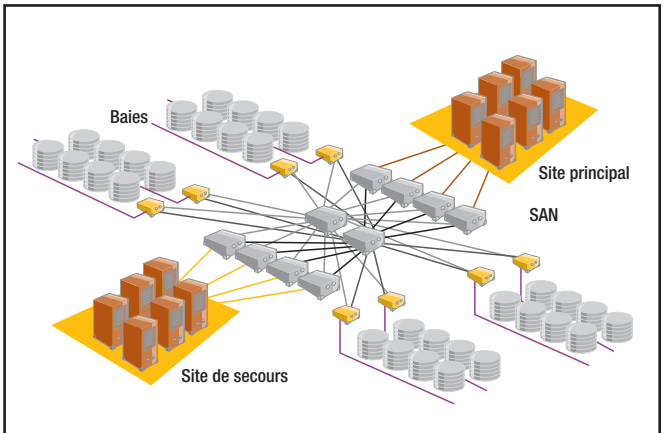

Figure 4 : SAN (Fabric) réparti

S'ils peuvent s'appuyer sur des systèmes de réplication basés sur le stockage, la gestion de volumes ou encore la réplication IP (basée sur des appliances), la mise en œuvre de clusters n'est pas toujours indispensable. Il est possible de basculer par script les volumes d'un site sur l'autre, puis de remonter les systèmes de fichiers et les bases de données sur le site de secours. Aucune perte de données n'est à craindre, mais la détection et le déclenchement des opérations de basculements demanderont l'intervention d'opérateurs, ce qui augmentera le temps d'indisponibilité des applications.

Oracle est le principal éditeur de SGBD qui investisse dans les technologies de clusters actifs/actifs. Le RAC (Real Application Cluster) permet de monter plusieurs instances sur une même base de données physiques : l'espace de stockage est partagé entre plusieurs nœuds. Cette technique offre plusieurs avantages : ajout de puissance par ajout de machines, limitation des impacts des applications entre elles, etc. En ce qui concerne la haute disponibilité, l'intérêt du cluster RAC réside principalement dans la disponibilité des instances. Les clients ou les serveurs d'applications passeront d'un nœud à l'autre par reconnexion, sans démarrage d'instance. A l'opposé, un cluster « failover » nécessite en effet une action d'activation des disques du côté du nœud de secours. Cette activation est pilotée par un cluster qui devra exporter les groupes de volumes, activer en R/W les volumes des baies du coté secondaire, importer les groupes de volumes, monter les systèmes de fichiers pour ensuite démarrer l'instance. L'ensemble de ces manipulations est soumis à des aléas de mauvaise configura-

tion (en cas d'ajout de volumes par exemple) et à des problèmes de basculement des volumes.

Mécanismes de réplication des bases de données

D'autres systèmes de réplications, plus ou moins encapsulés (comme Data Guard d'Oracle), peuvent être mis en œuvre dans les bases de données. Ces fonctionnalités peuvent être classées en deux catégories : synchrones ou asynchrones (cf. encadré). La première est adaptée à des volumes de mises à jour peu importants, mais présente l'avantage de n'occasionner aucune perte de données. Les réplications asynchrones basées sur des copies d'archive log (Oracle – Bases standby) ou des troncatures de fichiers logs de transaction (MS Sql Server - log shipping) peuvent quant à elle fonctionner sans pénalisation de performances, mais la perte de données ira de plusieurs minutes à plusieurs heures selon la fréquence des permutations de logs (Oracle) ou des backups de transaction logs (MS Sql Server). Les bases de données offrent également des mécanismes synchrones ou asynchrones de réplications basés sur le transfert des ordres SQL sur la base de secours (transactions différées). Enfin, Oracle dispose d'un système de « vues matérialisées ». A la différence des vues classiques qui doivent être exécutées à chaque invocation, les vues matérialisées stockent des données rafraîchies périodiquement.

Snapshots

Le principe du snapshot est de créer une copie datée de volumes (par exemple de bases de données) sur disques. L'intérêt de cette technique par rapport à une sauvegarde classique réside principalement dans la vitesse de mise à disposition des volumes sauvegardés. L'inconvénient majeur des snapshots est la multiplicité des espaces disques qui seront générés. Il est difficile en effet de s'offrir la possibilité de faire plusieurs copies physiques d'une base de 1To !

Le terme « snapshot » recouvre deux mécanismes : une copie datée d'un volume (par exemple sur une baie), ou une duplication à la volée des blocks du volume source au fur et à mesure des modifications vers le volume de copie. La copie à la volée a pour principal avantage de n'occasionner qu'une courte indisponibilité du volume, mais les blocks du volume d'origine devront être dupliqués sur la copie lors de leur première modification. Cette solution est donc adaptée à des volumes peu mis à jour. Un autre avantage réside également dans le gain d'espace disque dans la mesure où l'espace à réserver sera rarement aussi important que celui du volume source. La duplication intégrale d'un volume ne permet bien entendu ni de limiter l'espace consacré au volume de secours ni d'en réduire le temps d'indisponibilité. Le volume doit être désactivé, copié, puis réactivé. Cette technique a pour principal avantage de ne pas causer de dégradations sur les écritures qui seront ensuite effectuées sur le volume source.

Les snapshots ont surtout été popularisés par les vendeurs des baies de stockage qui combinent ces deux techniques avec des logiciels de gestion des volumes créés. Ces techniques existent également dans les gestionnaires de volumes traditionnels ou dans certains systèmes de fichiers.

Et les sauvegardes ?

Le gigantisme des bases de données, qui commence à prévaloir dans de nombreuses entreprises, conduit à la remise en question du principe même des sauvegardes. La majorité des environnements de production refuserait un arrêt supérieur à la journée. Le rôle des sauvegardes en tant que garant de l'intégrité des données et de la continuité de service s'est amenuisé. La réplication des bases de données en mode différé est par exemple un substitut à la restauration. Dans certains cas il suffit d'appliquer les logs de recouvrement avec un intervalle pour disposer d'une fonctionnalité de basculement sur une image de l'environnement de production à un point dans le temps. Dans la mesure où les copies des fichiers logs sont dis-

Réplications des bases de données : synchrones ou asynchrones

	asynchrones	**synchrones**
Principe	• Copies de fichiers logs • Réplication des ordres de modification SQL asynchrones • Vues matérialisées (Oracle)	Réplication des ordres de modification SQL synchrone (two phase commit)
Avantages	Impact de performances limité sur la base source	• Bases utilisables simultanément : répartition de charge • Pas de pertes de données • Basculement automatisable
Inconvénients	• Base non accessible (en recouvrement dans le cas des copies de logs), pas à jour dans le cas des réplications d'ordres SQL • Réplication SQL : probabilité de conflit forte en cas de mises à jour sur les 2 sites des mêmes bases	Inadapté aux mises à jour transactionnelles importantes

ponibles sur le site de secours, on peut amener la base de données à un point dans le temps situé entre la dernière mise à jour de la base de secours et le moment du crash (ou celui de la corruption, ou du « drop table » intempestif). Les sauvegardes restent cependant indispensables, et cela pour au moins deux raisons : elles contiennent la « mémoire » des informations stockées sur plusieurs années, et elles sauvegardent la totalité d'un environnement à un point dans le temps (la réplication différée associée au « roll forward » ne concerne que les bases de données). Mais le couple temps de recouvrement/perte de données est fréquemment situé en dehors des limites acceptables.

Des investissements à la hauteur du risque

Sans entrer dans des calculs de probabilité d'occurrence de tel ou tel risque, au moins deux catégories d'incidents doivent être envisagées : les incidents physiques et ceux liés aux logiciels. La première catégorie comprend les incidents qui auront un impact sur des équipements : machines, baies de stockage, commutateurs, salles informatiques, climatisation, etc. L'impact n'est d'ailleurs pas obligatoirement physique : une zone déconfigurée dans la fabric, un volume de baie attribué par erreur à deux serveurs, voire un « system panic » rendront le composant aussi inaccessible qu'une panne électrique. La seconde catégorie concerne les applications, en particulier les structures de stockage comme les bases ou les fichiers de données. Les modifications intempestives de données et surtout les erreurs humaines telles que les suppressions de tables ou de fichiers de données en constituent la trame.

On se protégera des incidents physiques en s'équipant de matériels redondants et de salles de secours. Les incidents liés aux applications doivent être corrigés par un retour des données dans le temps : snapshots, réplications de bases et sauvegardes.

In fine, les vieux principes restent valables : un PCS n'est que l'image d'un triptyque risque/investissement/impact... comme l'assurance d'une voiture. ■

Chistian Bilien

Ingénieur de l'Ecole Supérieure d'Electronique de l'Ouest et diplomé de l'IEFSI-EDHEC, Christian Bilien est spécialisé dans les performances et la haute disponibilité.

cbilien@storage-foundation.com

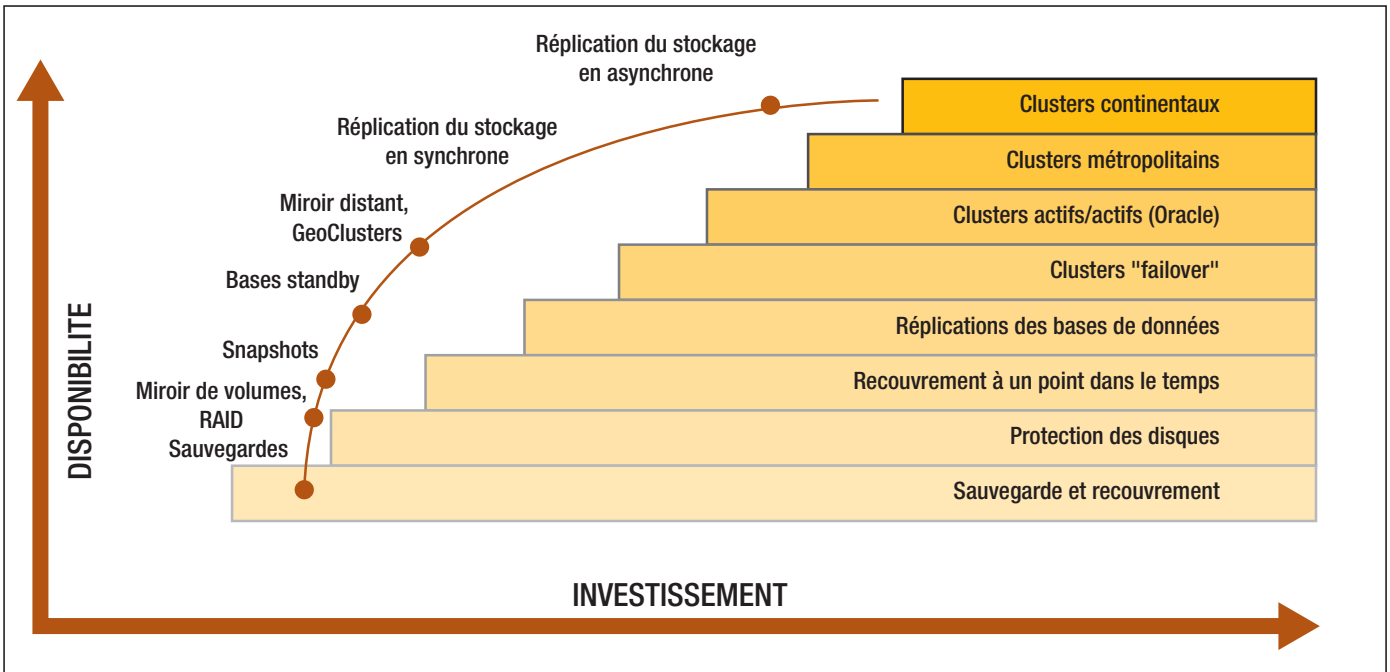

Figure 5 : Investissements en fonction des risques encourus