

Université IBM i 2018

16 et 17 mai

IBM Client Center Paris



Session S46 – GDPR et IBM i Et si l'article 32 repositionnait la réplication logicielle ?

Didier Feraud

Syncsort

didier.feraud@syncsort.com



Au sommaire ...

RGPD et Résiliences

- Résilience des applications
- Résilience des environnements
- Résilience des données

RGPD et Répliquions logicielles

- Syncsort pour la GDPR
- Mimix 9 et Quick-EDD/HA
- La continuité pour chaque situation

En synthèse

GDPR: 173 Considérants et 99 Articles

Section dans le RGPD	Catégorie
Considérant (6)	Protection
Considérant (7)	Protection
Considérant (15)	Protection
Considérant (28)	Chiffrement
Considérant (39)	Vie privée, Contrôle d'accès
Considérant (49)	Dommages, Intégrité, Contrôle d'accès
Considérant (71)	Vie privée, Risques
Considérant (78)	Conformité, Chiffrement
Considérant (83)	Risques, Chiffrement, Dommages, Vie privée, Intégrité, Contrôle d'accès
Considérant (84)	Risques
Considérant (87)	Paramètres, Journal/Audit
Considérant (90)	Risques, Paramètres, Protection, Conformité
Considérant (108)	Protection
Considérant (162)	Contrôle d'accès
Article 5 1(f)	Intégrité, Vie privée, Dommages
Article 6 4(e)	Chiffrement
Article 24	Toutes les catégories
Article 25	Protection, Chiffrement, Contrôle d'accès
Article 28	Toutes les catégories
Article 30	Journal/Audit, Paramètres de sécurité
Article 32	Risques, Chiffrement, Intégrité, Paramètres de sécurité, Dommages, Vie privée, Contrôle d'accès
Article 34	Chiffrement
Article 35	Risques, Paramètres
Article 47	Intégrité, Protection
Article 71	Paramètres de sécurité

- ✓ 14 Considérants (motivation) et 11 articles mentionnent ou impliquent le besoin en **technologies de protection des données**.
- ✓ Technologies de protection des données regroupées dans les catégories ci-dessous:
 - **Protection des données**
 - Vie privée et confidentialité des données
 - Intégrité des données
 - Chiffrement et pseudonymisation
 - Contrôle d'accès
 - Dommages accidentels et malveillants
 - Conformités aux règlements
 - Evaluation des risques
 - **Journalisation et vérification**
 - Paramètre et politique de sécurité

GDPR= General Data Protection Regulation



Article 32 Sécurité du traitement



2.1 Obligations de disponibilité (art. 32)

- **Disponibilité & résilience constantes** des systèmes et des services de traitement
- Moyens permettant de **rétablir la disponibilité** des données et leur accès dans des délais appropriés en cas d'incident physique ou technique
- Procédure de **tests réguliers de l'efficacité** des mesures techniques et organisationnelles



Où il est question pour le responsable du traitement et son sous-traitant de mettre en œuvre les mesures techniques et organisationnelles afin de garantir:

- la confidentialité,
- l'intégrité,
- la disponibilité
- la résilience constantes des systèmes et des services de traitement



Article 32 ... Sécurité du traitement



2.1 Obligations de disponibilité (art. 32)

- **Disponibilité & résilience constantes** des systèmes et des services de traitement
- Moyens permettant de **rétablir la disponibilité** des données et leur accès dans des délais appropriés en cas d'incident physique ou technique
- Procédure de **tests réguliers de l'efficacité** des mesures techniques et organisationnelles



Où il est question pour le responsable du traitement et son sous-traitant de mettre en œuvre les moyens techniques et organisationnelles afin de garantir:

- la disponibilité des données à caractère personnel
- l'accès à celles-ci dans **les délais appropriés** en cas d'incident physique ou technique
- une **procédure fiable de test et d'évaluation**

Résilience des données



Définition de la résilience

La résilience c'est l'aptitude ou la capacité pour un **système en général** à **répondre, absorber, s'adapter** à tout **changement dynamique** interne ou externe et à revenir à un **état stable**.



Résilience des applications

- ✓ Doivent être pensées et développées pour redémarrer automatiquement en ayant mémorisé la dernière frontière de transaction (*Commitment Control*) pour se repositionner d'elles-mêmes au bon endroit
- ✓ Si transparence totale au niveau utilisateur au temps de réponse près, elles doivent être de type client/serveur ou trois tiers afin que la disponibilité des applications soient décorrélée de la disponibilité des données

Résilience de l'environnement

- ✓ Définir une **Infrastructure Essentielle** afin de maintenir ou rétablir un fonctionnement acceptable, idéalement identique, voire plus ou moins dégradé.
- ✓ Protéger cette **Infrastructure Essentielle** en élaborant une stratégie de protection adaptée.



Résilience des données et métadonnées

- ✓ *Paradoxe*: la donnée n'a pas ou peu de valeur mais les données en ont. La force des données et de ce qu'elles peuvent produire c'est la **quantité**. C'est une source de valeur considérable !
- ✓ De plus en plus de domaines d'activités imposent des **règles de gestion** bien précises incluses dans des normes: PCI-DSS, Bâle II et III, Solvency II, HIPAA, loi SOX et norme COBIT, RGPD....
- ✓ La **sécurité**, l'**intégrité**, la **qualité** et la **conformité**, la **disponibilité** ainsi que la **traçabilité** des données et des métadonnées (*données complémentaires contextuelles qui permettent de comprendre les propriétés des données brutes*) occupent une place prépondérante dans ces textes qui préconisent une **protection continue** de ces dernières (*Continuous Data Protection*)
 - Politique de sauvegarde régulière plus ou moins élaborée et automatisée
 - Plusieurs technologies de haute disponibilité (HA) multi-systèmes permettent de satisfaire aux besoins de résilience des données et métadonnées:
 - la réplication logicielle des données indépendante du type de stockage
 - les répliques matérielles des entités de stockage des données

Il est possible de combiner plusieurs technologies afin de renforcer leur résilience

Quelles que soient les solutions technologiques et organisationnelles retenues pour assurer ces différentes résiliences, il y a obligation de procéder à des tests réguliers pour en mesurer l'efficacité et de s'assurer qu'on est en conformité avec la réglementation.





RTO, RPO et NRO

- **RTO** ou **DMIA** ... *Durée Maximum d'Indisponibilité Admissible*
- **RPO** ou **PDMA** ... *Perte de Données Maximum Admissible*
- **NRO** ou **IRMA** ... *Indisponibilité Réseau Maximum Admissible*
 - ✓ **Indicateurs d'objectif** et métriques de continuité d'activité
 - ✓ Utilisés pour **définir les attentes** de remise à disposition des applications et donc des données

SLA

- **SLA** ou **Contrat/Entente de Niveau de Service**
 - ✓ **Indicateurs d'engagement** de continuité opérationnelle
 - ✓ **Contrat** passé entre un fournisseur de service et ses clients internes ou externes
(responsable du SI et organisations fonctionnelles par exemple)
 - ✓ Les métriques définies telles que RPO et RTO sont considérées comme des **composants** du SLA et non comme le SLA lui-même



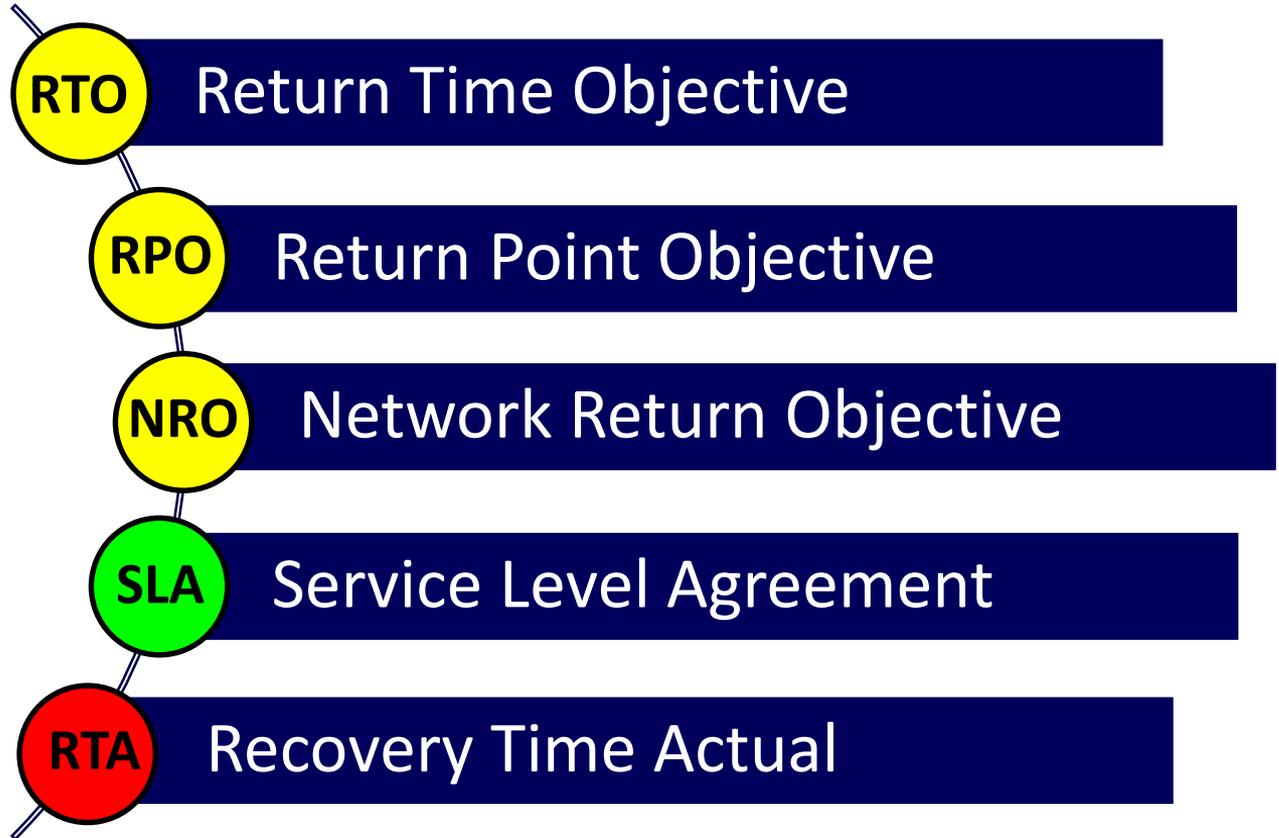
RTA

- **RTA (Recovery Time Actual) ou Temps de Reprise Effectif**
 - ✓ **Indicateur factuel**, en général validé dans un contexte de reprise planifiée
 - ✓ *Ne tient pas compte de la reprise applicative complète et de l'identification du point de reprise des données*
 - ✓ Permet de mesurer l'efficacité de la stratégie de secours mise en place.
 - ✓ *On peut le comparer au RTO pour l'ajuster si besoin ou revoir la stratégie de secours pour raccourcir le RTA*

L'objectif final sera de mettre en perspective RTO, RPO, NRO et RTA afin de s'assurer que le SLA publié est cohérent et qu'on est en capacité de l'atteindre



Résilience... Les principaux indicateurs



Et si on parlait de gouvernance de la donnée ?

La donnée est au centre de toutes les attentions dans l'entreprise mais aussi de la réglementation.

En conséquence, pour beaucoup d'organisations, la mise en place d'un système de gouvernance de la donnée s'impose afin d'assurer un traitement qualitatif des informations et d'en assurer leur sécurité.

Définition de la gouvernance des données

La gouvernance des données est une collection de bonnes pratiques qui considère l'information comme une ressource à part entière de l'entreprise: l'information doit être gérée avec des règles précises, des processus "responsabilités" clairement établis et véhiculés au sein de l'organisation dans le respect de leur standards technologiques.

Elle regroupe tous les moyens permettant la gestion optimale des dimensions qualité, disponibilité, sécurité et conformité réglementaire, et a pour objectif la valorisation de la donnée dans le temps.

Source <http://fracademic.com/dic.nsf/frwiki/718438>

Définition du Chief Data Officer

La responsabilité du **Chief Data Officer** (Directeur des Données) concerne les données et leur état (riches, fraîches, fiables et cohérentes) ainsi que leur gouvernance. Les données peuvent être d'origine interne ou externe, privée ou publique (*Open Data...*).

Le Chief Data Officer peut être chargé de l'acquisition de données:

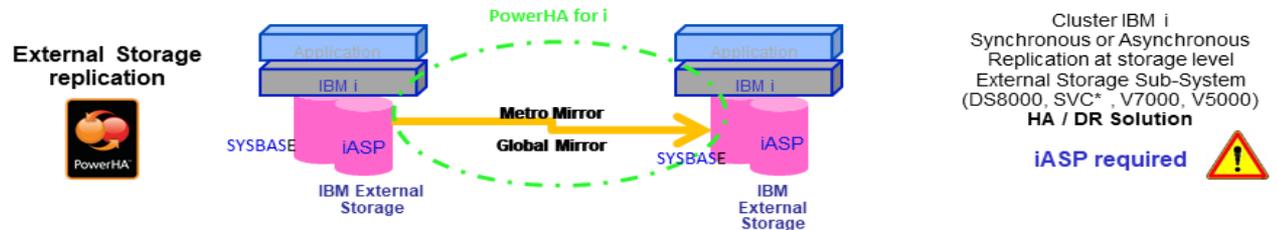
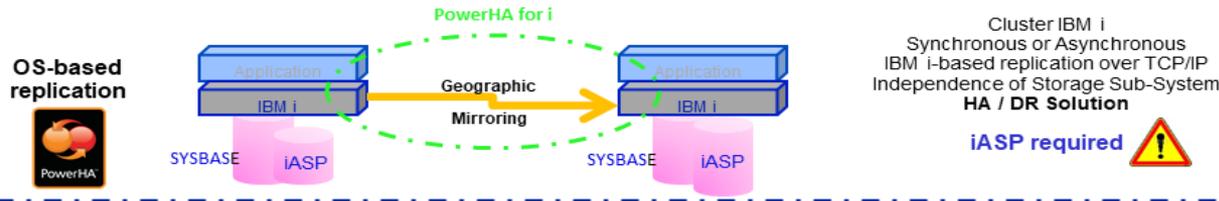
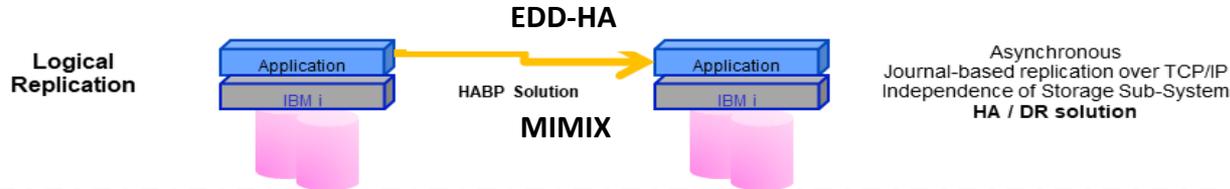
- échanges, négociation pour fourniture par des fournisseurs, partenaires ou clients de l'entreprise
- identification de fournisseurs possibles de données, sensibilisation, négociation et acquisition
- action auprès d'un régulateur si les données s'y prêtent et si elles ne sont pas disponibles à l'achat

La fonction pourra éventuellement inclure le rôle de **Data Protection Officer** (Directeur de la protection des données personnelles) inscrit dans le RGPD. Il est le garant de la protection des données personnelles devant le régulateur.

Source https://fr.wikipedia.org/wiki/Chief_Data_Officer



Résilience des Données ... Quelles Solutions ?



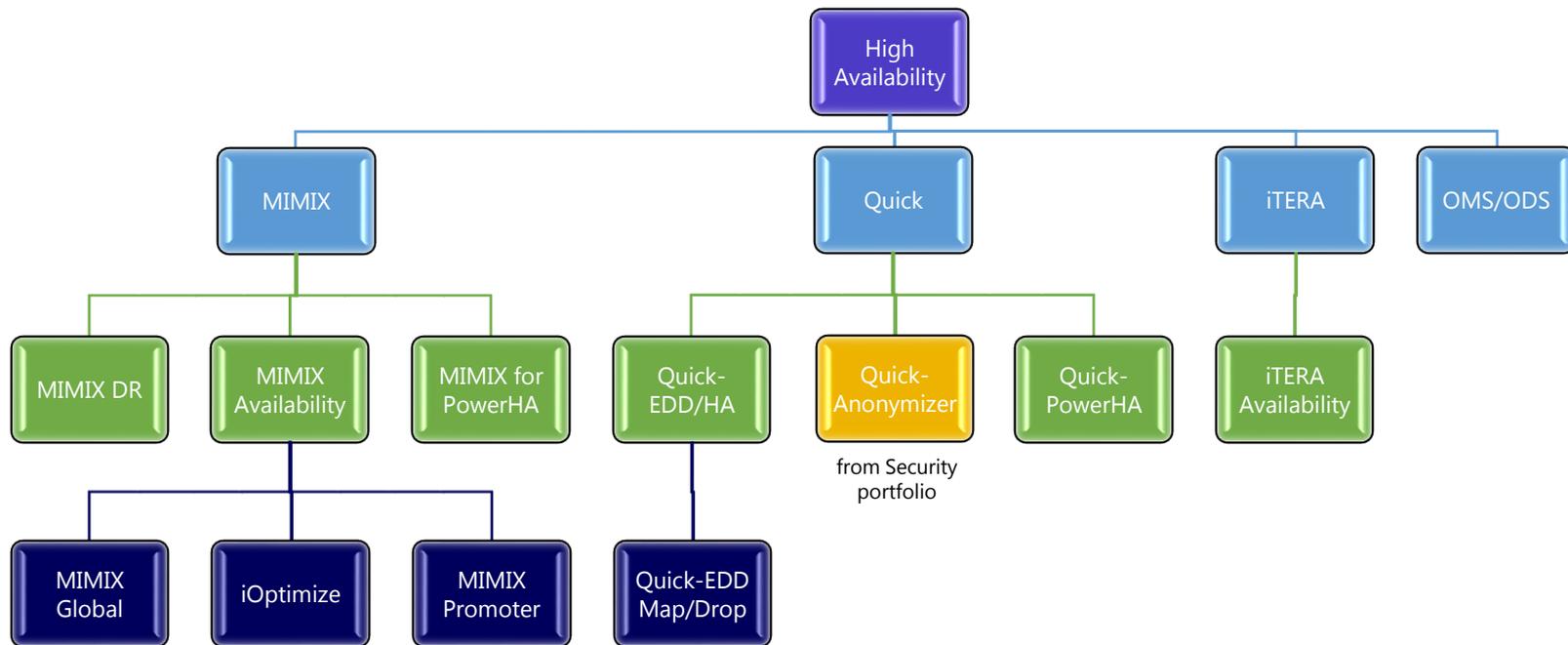
Syncsort's High Availability Portfolio

Portfolio

Brands

Products

Branded Features



Innovation et Intégration

Le support et le développement sont assurés pour les produits actuels

- Quick-EDD/HA
- MIMIX Availability
- iTERA Availability
- OMS/ODS

Innovation et intégration pour fournir de nouvelles solutions

- Etendre dans la mesure du possible les solutions Syncsort non-IBM i au marché de l'IBM i
- Fournir de nouvelles solutions pour répondre aux besoins des clients actuels et futurs
- Intégrer le meilleur de chaque produit HA pour tous les clients Syncsort



Réplication évolutive en temps réel MIMIX

Basée sur la journalisation distante d'IBM

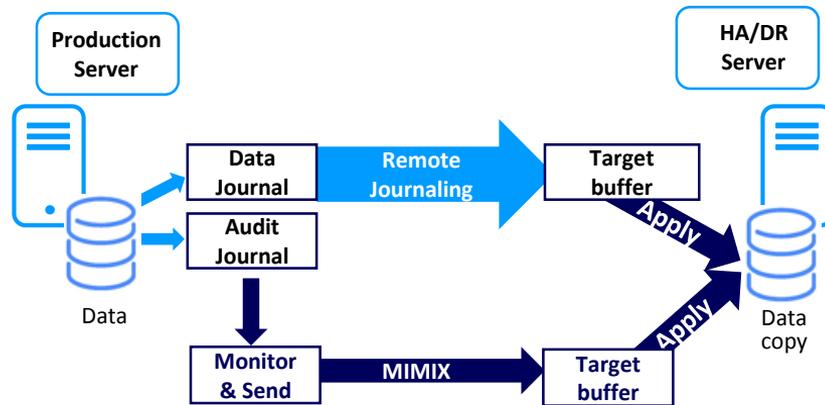
- Réplique les données modifiées en temps réel
- Assure l'intégrité des données et élimine leur perte
- Minimise l'utilisation de la bande passante
- La fonction de capture des données modifiées utilisée par MIMIX assure la réplication des données non journalisées

Flexible et indépendant de l'architecture de la plate-forme

- Réplique entre serveurs locaux, distants ou dans le Cloud
- Prend en charge une grande variété de configurations de réplication, y compris bidirectionnelles
- Supporte des configurations de serveur, de stockage et versions de système d'exploitation différentes
- Permet un accès en lecture aux données répliquées par Query, générer des rapports, voire réaliser des sauvegardes, etc...

Optimisé pour la performance

- Assure un RPO proche de zéro (zéro avec le Remote Journaling synchrone)
- Assure un RTO proche de zéro en maintenant un serveur de secours prêt pour un changement de rôle
- S'adapte aussi bien aux petites entreprises qu'aux grandes sociétés les plus exigeantes en utilisant un véritable multithreading
- Les fonctions d'optimisation de performance adaptent l'application à votre environnement (par exemple, applications parallèles sur la cible et mise en cache des accès base de données)
- La réplication active-active avec résolution de conflit intégrée délivre un RTO de quelques minutes (aussi rapidement que le temps nécessaire au déplacement des utilisateurs vers un autre serveur actif)



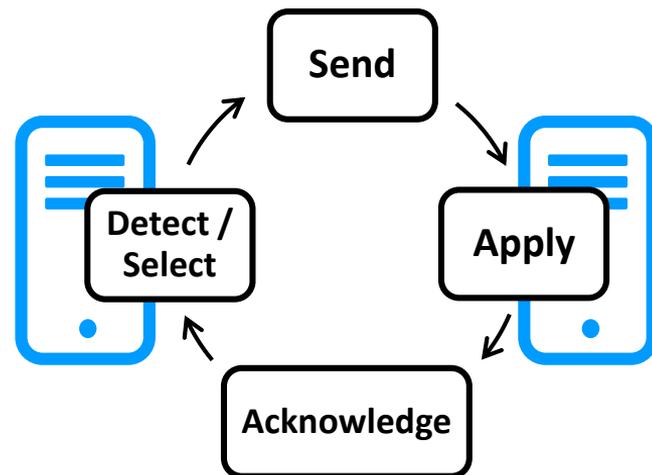
Réplication en temps réel efficace et rapide EDD-HA

Rapide et Efficace

- Assure l'intégrité des transactions grâce à l'utilisation du journal local de l'IBM i
- Permet un RTO au niveau minutes et un RPO sub-seconde
- S'adapte aussi bien aux petites entreprises qu'aux grands groupes
- Minimise l'utilisation de la bande passante
- Permet la compression et l'encryption au niveau des communications inter-serveurs
- Permet l'anonymisation des données personnelles lors de la réplication avec le module optionnel Quick-Anonymizer

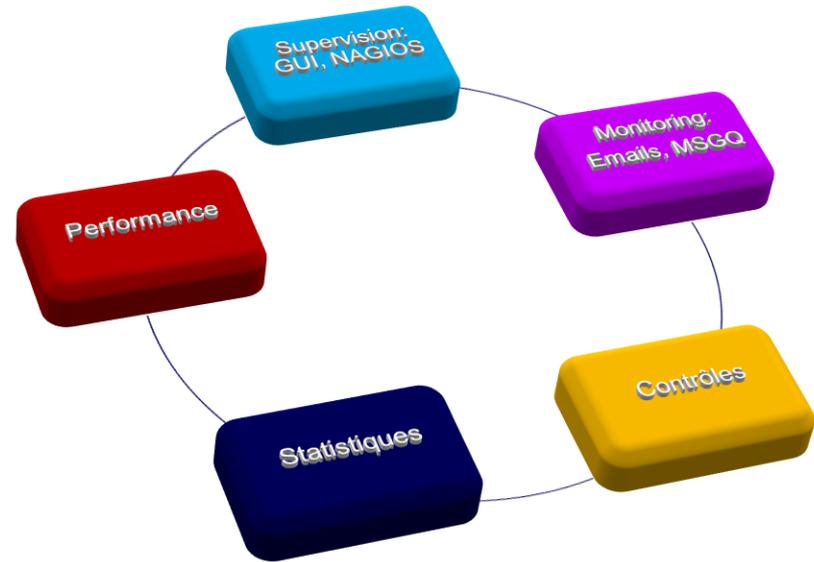
Flexible

- Replicates between mixed hardware, storage and IBM i OS versions
- Supports a variety of configurations, including bi-directional and broadcast
- Permet une utilisation en mode multi-nœuds
- Fournit les fonctions nécessaires pour la protection continue des données (Continuous Data Protection)
- Permet l'accès aux données répliquées pour les opérations du quotidien ou la maintenance de l'infrastructure



Outillage livré avec la réplication logicielle

- ✓ Les outils constituent les briques complémentaires permettant de s'intégrer à toutes les organisations
- ✓ Des fonctions issues de l'expérience des clients et partenaires



La continuité pour chaque situation

- ✓ **Sinistres (RTO & RPO) ou Problèmes Hardware ...** *réplication active/active et intégrité de la cible*
- ✓ **Tout type de stockage et toute distance inter-sites ...** *réplication indépendante du matériel*
- ✓ **Maintenance des systèmes et OS ...** *bascule scriptée*
- ✓ **Sauvegardes déportées ...** *sauvegardes des données sensibles sur la cible sans arrêt de réplication*
- ✓ **Continuous Data Protection ...** *protection des données sensibles sur la cible*
- ✓ **Alimentation d'InfoCentres, Entrepôts de Données, Environnements de Développement** *réplication/synchronisation en temps réel ou par vacations*
- ✓ **Erreurs Humaines ...** *réplication en Y dont une en décalé*
- ✓ **Sécurité à deux niveaux ...** *réplication en rebond ou en Y (HA + DR)*
- ✓ **Contraintes d'Exploitation pour changement de serveur ..** *Migration Kit*
- ✓ **Upgrades applicatifs ...** *alimentation des fichiers en cas de modifications de structure liées à un nouveau design applicatif pour tests, validation et migration applicative*





- (1) Possible mais avec configuration spécifique
- (2) Pas d'accès possible sur la baie répliquée, même si IASP
- (3) Réplication synchrone limitée à 40/50Kms maxi, au-delà réplication asynchrone avec groupes de consistance et augmentation du risque de perte de données
- (4) Fonction dépendante du produit.
- (5) Possible sans retard de la réplication car stockage des Mises à Jour sur la cible et temps de retard de l'application différée paramétrable dans la configuration du lien
- (6) Fonction optimisée pour traiter un très grand nombre d'objets et/ou de données dans un minimum de temps en utilisant jusqu'à 24 services multithreads en parallèle

Fonction	Réplication HWR	Réplication SFW
Journal <i>RPO=journal</i>	✗	✗
Audit Journal <i>QSYS/QAUDJRN</i>	✗	✗
Remote Journal <i>RMT JRN</i>	Inutile	✗
Suivi des travaux sur la cible	✗	✗
Copie Receveurs sur cible <i>JRN / RMTJRN</i>	✗	✗
Copie Receveurs sur cible <i>Audit Journal</i>	✗	✗ (1)
<i>SYSDASE (IASP)</i> Synchro des objets	Partiel <i>Domain Administrative Cluster</i>	✗
Sauvegarde et Query sur la cible <i>Pas d'arrêt de la Prod</i>	✗ (2)	✗
Configuration HA/DR	Dépendant du type de baie (<i>DS8K</i>) Architecture MM/GM	✗
Dispersion géographique	✗ (3)	✗
Migration technologique & Map/Drop (Data Migration)	✗	✗ (1) (4)
Possibilité de retarder la réplication sur la cible	✗	✗ (4) (5)
Verify While Active <i>Comparaison des objets cible/source pour en vérifier l'intégrité</i>	✗	✗ (6)



Differentiated Product Portfolio & Technical Expertise

Data Infrastructure Optimization

Best-in-class resource utilization and performance, on premise or in the cloud

- MFX® for z/OS
- ZPSaver Suite
- EZ-DB2
- EZ-IDMS
- DMX & DMX-h
- DMX AppMod
- DL/2
- Zen Suite
- **athene®**
- athene SaaS®

Data Availability

#1 in high availability for IBM i and AIX Power Systems

- MIMIX Availability & DR
- MIMIX Move
- MIMIX Share
- Quick-EDD – Quick CSI
- iTera Availability
- Enforcive IBM i Security
- Cilasoft

Data Integration

Industry-leading mainframe data access and highest performing ETL

- **Ironstream®**
- Ironstream® Transaction Tracing
- DMX & DMX-h
- DMX Change Data Capture

Data Quality

Market-leading data quality capability

- Trillium Software System
- Trillium Quality for Big Data
- Trillium Precise
- Trillium Cloud
- Trillium Global Locator

Big Iron to Big Data

A fast-growing market segment composed of solutions that **optimize** traditional data systems and **deliver** mission-critical data from these systems to **next-generation analytic environments**.



En conclusion ...



Pour la majorité des entreprises, un long chemin reste à parcourir avant que les organisations ne deviennent, à l'image des GAFAs, des entreprises "Data Centric" où la donnée est un élément structurant de la culture de l'entreprise et où elle est gérée comme un **actif clé**.

<http://www.cigref.fr/wp/wp-content/uploads/2016/11/CIGREF-Valorisation-des-donnees-Pratiques-Modele-2016.pdf>



Questions ?



