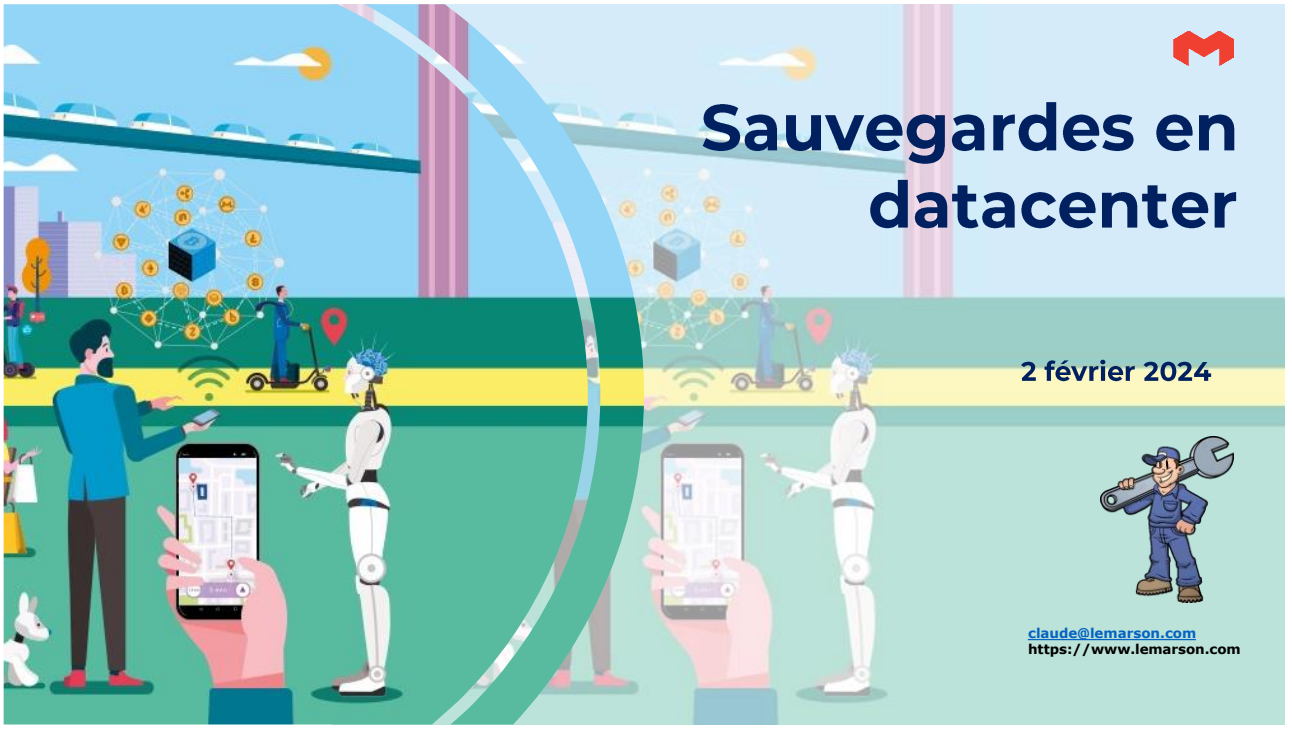


Sauvegardes en datacenter

2 février 2024

claude@lemarson.com
https://www.lemarson.com

SOMMAIRE

Sauvegardes du datacenter : un problème jamais résolu

- Coût d'une panne : 6 500 $/mn (Gartner).
- Taux de survie : seules 6% des entreprises qui ont subi un désastre sans stratégie de sauvegarde, ont survécu au-delà de 2 ans.

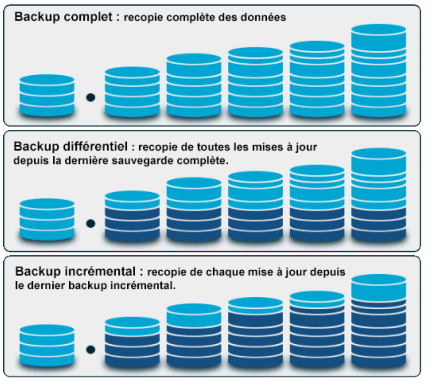
Les sauvegardes, un vieux problème

- Depuis les débuts des mainframes (1964), les sauvegardes ont été un problème majeur
- Bandes magnétiques et disques amovibles
- Le temps d'exécution était important, incompatible avec une production simultanée
- Dans le même temps où les supports et technologies de sauvegarde ont fortement progressé, les architectures de SI se sont complexifiées : ce que l'on a gagné avec les supports, on l'a perdu en complexité

- Contrairement à ce que l'on croit :
- Il y a toujours une production en salle machine
- Des opérateurs qui lancent des commandes et surveillent…les robots
- Des personnels chargés de la maintenance
- Des spécialistes de l'exploitation : sauvegardes, optimisations, abreuvés de "dashboards"
- On n'a pas encore atteint la sérénité (d'où l'envie d'aller dans le Cloud)

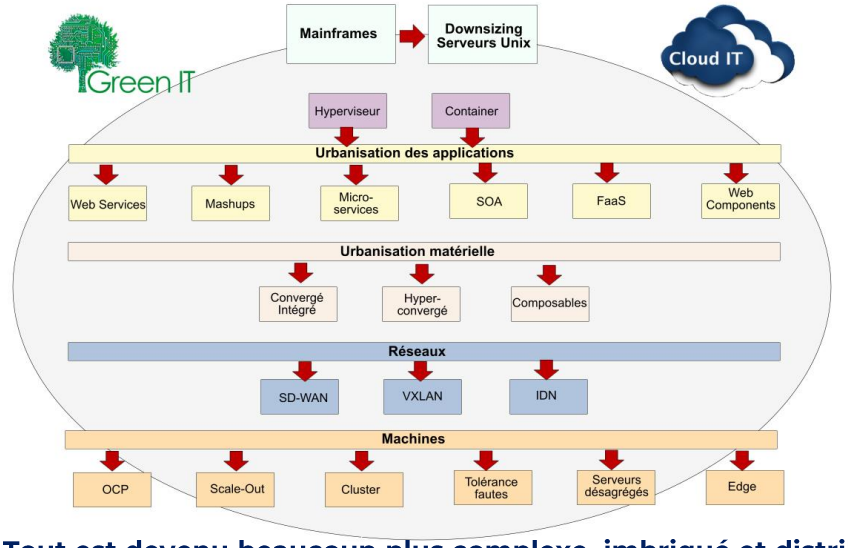
Sauvegardes : de quoi parlons-nous ?

- La sauvegarde ("backup") est une activité de production qui consiste à combiner différentes technologies serveurs, réseaux, logicielles, etc, pour copier certaines ressources de l'entreprise, pour pouvoir les récupérer en cas de sinistre et redémarrer ses métiers selon un périmètre donné.
- Les ressources ne se limitent pas aux données sensibles :
  - Applications
  - Bases de données : sensibles, moyennement sensibles, sans intérêt stratégique.
  - Images systèmes : machines virtuelles, conteneurs.
  - Systèmes d'exploitation.
  - API et frameworks utilisés dans les développements.
  - Registres.
  - Fichiers, documents.
- La stratégie consiste à savoir ce que l'on va sauvegarder, par quels moyens et à quelles fréquences.
- On ne confondra pas sauvegardes, migration de données et réplications.
  - La migration de données revient à effectuer une copie complète d'un fichier, bases de données, sur un autre support. Elle est généralement planifiée à une fréquence longue : 24 h ou plus.
  - La réplication est une procédure logicielle, accessible par un module système ou les applications (triggers, CDP : Continuous Data Protection), qui recopie une ou plusieurs données en temps réel sur un autre support, quand elles sont mises à jour. La réplication se fait en temps réel, alors que la migration est décalée.
- Une sauvegarde peut être complète, incrémentale ou différentielle.
  - Mode incrémental : on ne recopie que les données modifiées depuis la dernière sauvegarde.
  - Mode différentiel : on recopie toutes les données nouvelles depuis la dernière sauvegarde complète.
  - Mode "reverse incremental" : on part d'une copie que l'on synchronise avec la version courante. On sauvegarde tous les incréments, pour faire revenir la copie aux versions précédentes.



Nouvelles architectures, nouvelles contraintes

Les architectures modernes sont un défi au maintien de l'activité



Tout est devenu beaucoup plus complexe, imbriqué et distribué

Les niveaux de sévérité

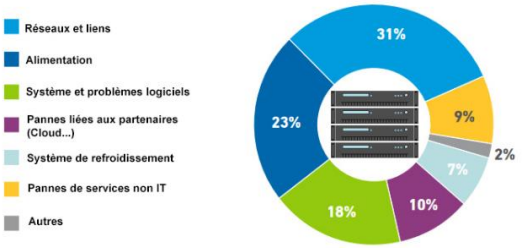
1	Négligeable	Peu ou pas d'impacts sur les services fournis. Les inconvénients sont faciles à réparer. Il suffit de repartir d'une situation saine.
2	Minimum	Interruptions faibles de services. Peu de conséquences chez les usagers, ni pour l'image de marque.
3	Significative	Interruptions significatives, mais limitées en conséquences. Faible impact financier et perte ponctuelle de clients. Mauvais pour l'image.
4	Sérieuse	Interruptions de services, pertes financières, pertes de clients. Image écornée. Situation rattrapée au bout d'une durée limitée.
5	Grave	Destruction d'installations et interruptions longues de services. Perte de nombreux clients et graves conséquences sur l'image.

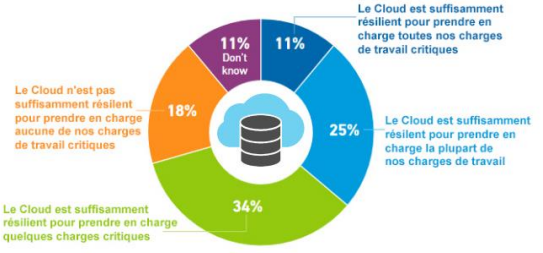
Même les plus grosses structures...

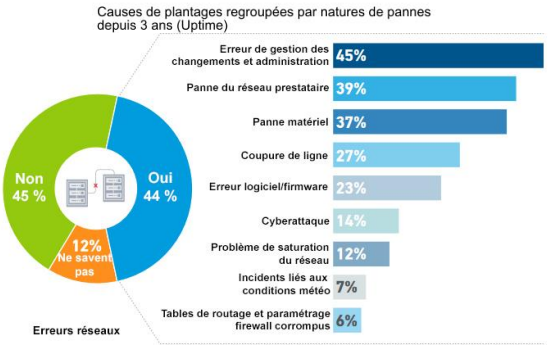
		Sévérité	
	**FAA (Federal Aviation Admnistration)**	**5**	Erreurs de configurations llogicielles, erreurs dans les bases de données. Tous les vols américains ont été interrompus, graves désordres dans les autres vols.
	**Kakao Entertainment** (Corée du Sud)	**5**	Feux de batteries. Les usages et services ont été interrompus pendant 8 heurs. Nombreux procès.
	**KDDI** (Japon)	**5**	Panique dans la configuration des réseaux. 39 mlillions d'usagers sont perturbés pendant 86 h. Des services critiques d'entreprises ne fonctionnent plus.
	**Google**	**4**	Erreurs de paramétrages logiciel. Interruption ou dégradation de Google Search et des applications connectées dessus, pendant 40 mn.
	**Common Spirit Health**	**4**	Cyberattaques, ransomwares. Les services du 2ème plus grand hôpital américain sont interrompus 1 semaine et de grosses pertes de données. 150 M$ de pertes.
	**AWS**	**4**	Une zone AWS est mise off line, ce qui affecte des milliers d'usagers.

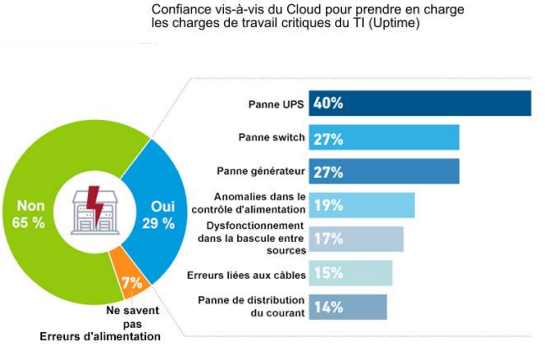
Les raisons des plantages

Des statistiques évocatrices



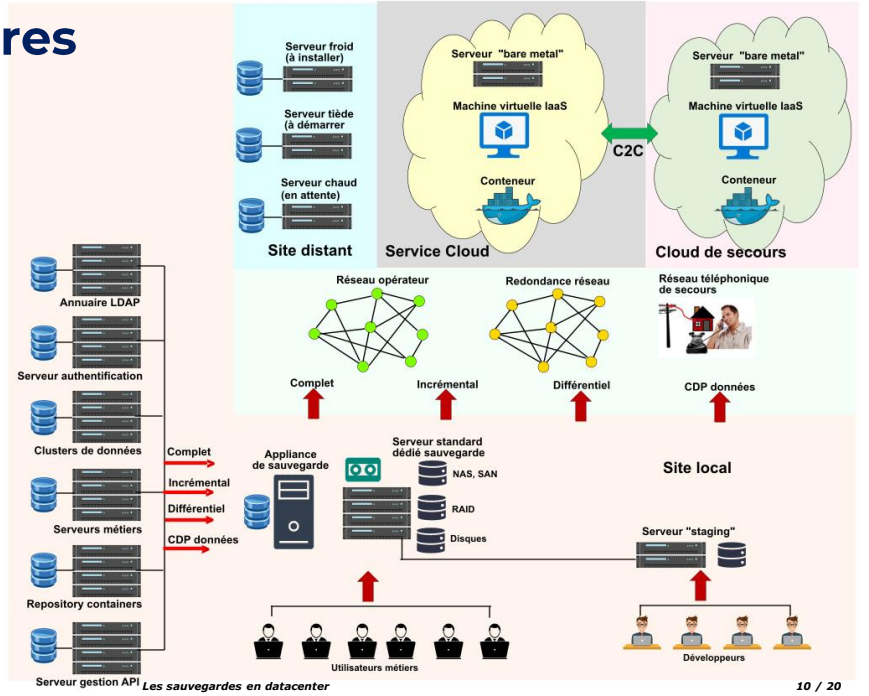






Les architectures possibles

- Il existe un grand nombre de solutions qui peuvent être exploitées dans un contexte de sauvegarde
- Serveurs locaux avec logiciels de backup
- Appliances locales : serveurs dédié aux sauvegardes
- Supports amovibles : DVD, disques
- Disques redondants : jusque RAID-5
- Disques externes
- NAS, SAN, serveurs de données objets
- Cartouches LTO
- Services de Cloud



Les services attendus

- La plupart des prestataires ont des solutions mixtes : logiciel spécifique, sur serveur dédié et Cloud.
- Certains sont ouverts à tout ou partie des Clouds, d'autres sont plus limités.
- Les solutions doivent être adaptables aux besoins exprimés en RTO/RPO.
- Elles doivent être simples à mettre en œuvre.
- Tous les modes de sauvegardes doivent être assurés : complets, incrémentaux, différentiels et s'appliquer à tous les types de fichiers et répertoires, système, données.
- Respect de la règle 3-2-1.
- Le logiciel de sauvegarde doit s'appliquer à tous les serveurs, locaux, Cloud et à tous les dispositifs : NAS, disques partagés…
- La reprise en cas de dysfonctionnement doit être très rapide pour les application et fichiers sensibles : 5 secondes sont une éternité…
- La restauration doit être granulaire et s'appliquer au niveau fichier, dossier, volumes disques, messagerie.
- Automaticité paramétrable, en cas de panne système, bascule sur une ressource en attente, sans intervention.
- Capacité de chiffrement, de garantie d'intégrité (immutabilité des données), protections contre les accès non autorisés.
- Respect des contraintes réglementaires sur la protection des données personnelles (RGPD…).
- Evolutivité ("scaling") : les solutions doivent être compatibles avec les différents types d'extensibilité : horizontale et verticale.
- Rapidité des sauvegardes. Elles ne doivent pas bloquer la production au-delà d'un certain temps.
- Possibilités d'archivage et de stockage à long terme.
- Documentation à jour, complète, compréhensible.

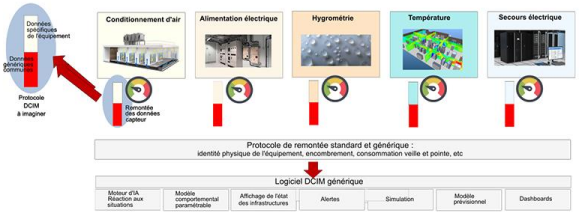
Les sauvegardes dans le Cloud
(avantages et inconvénients)

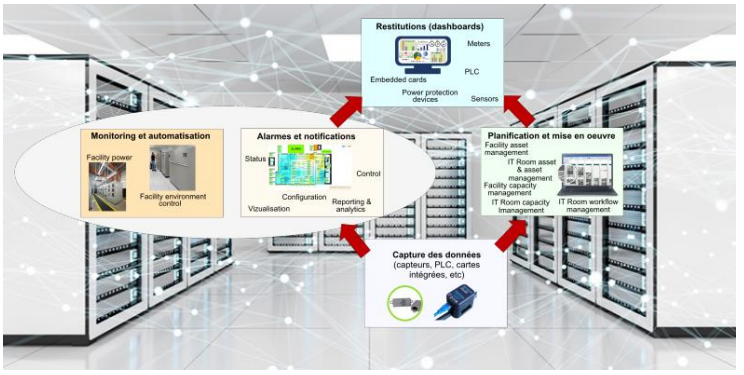
	Sauvegarde dans le Cloud	Sauvegarde locale
Nature du service	Les ressources sont copiées sur un serveur externe, porté par le Cloud (IaaS) ou "bare metal".	Les ressources sont copiées sur un serveur local, appliance ou serveur standard doté d'un logiciel dédié.
Coût	Les coûts sont abordables pour de petits volumes. Il y a risque d'inflation si on ne maîtrise pas ces volumes.	Les solutions "appliance" sont intéressantes techniquement, mais sont coûteuses. Les solutions standard sont abordables, mais les licences logicielles peuvent être élevées.
Evolutivité	L'évolutivité est sous la responsabilité des prestataires de Cloud. Elle est garantie, mais il faut être vigilant sur les ressources.	Plus difficile à garantir. On ne dispose pas nécessairement en local des moyens réseaux et serveurs.
Accessibilité	L'accès est assuré à partir d'une liaison Internet. Mais il convient de calibrer la bande passante de manière "pessimiste" (règle de 3 pour 1)	L'accès aux serveurs et "appliances" est assuré en interne et dépend de la bande passante du réseau d'entreprise : Ethernet sur fibre optique.
Sécurité	La sécurité est garantie de bout en bout, avec des procédures inter-clouds C2C.	Les ressources locales, quelles qu'elles soient font l'objet d'attaques permanentes. Les clients ont moins de moyens internes pour s'y opposer.
Management	La gestion des services est assurée par le prestataire, le client n'ayant à mobiliser moins de moyens humains pour cela.	L'équipe à mettre en place est importante, qui doit disposer de capacités de décision élevées. Une sous-traitance contrôlée peut être envisagée.

Les solutions du marché

- Prestataires de Cloud (IaaS)
- BaaS : Backup as a Service,
- Logiciels de backups
- Appliances (tout compris)

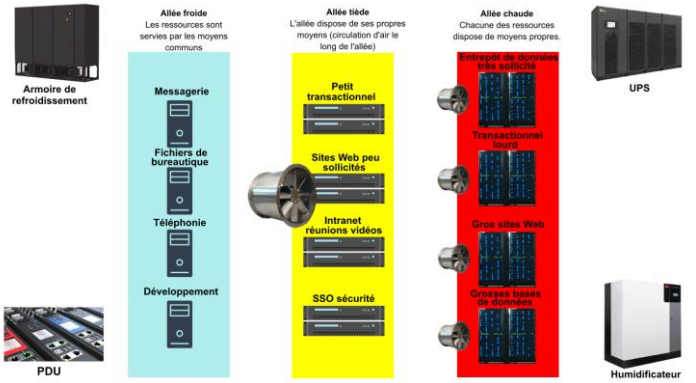
L'ingénierie DCIM





L'organisation intelligente des allées

- Les salles machines doivent être organisées en fonction de la consommation et des besoins énergétiques des charges de travail.
- C'est parfois le motif pour reporter ces charges partiellement ou totalement dans le Cloud.
- Globalement le PUE : Power Usage Effectiveness, mesure l'efficacité du datacenter en termes énergétiques.
- Le PUE permet de séparer la consommation électrique des serveurs et des baies de stockage, des écrans, postes de travail, climatisation, ventilation, alimentation des connexions réseaux, éclairage…
- Se situe généralement entre 1,5 et 2,5 : un tel centre consomme plus en dehors des équipements informatiques que pour les équipements eux-mêmes.



PCA/PRA : Répartition entre sites

- Pour élaborer un PCA, les entreprises doivent situer leurs applications en termes de « criticité »
- Même si le réflexe veut que toutes estiment devoir bénéficier d'un PRA temps réel
- Il n'y en a que très peu qui en ont réellement besoin…mais c'est très compliqué de le leur faire admettre
- Il existe des palliatifs (solutions dégradées) qui suffisent amplement

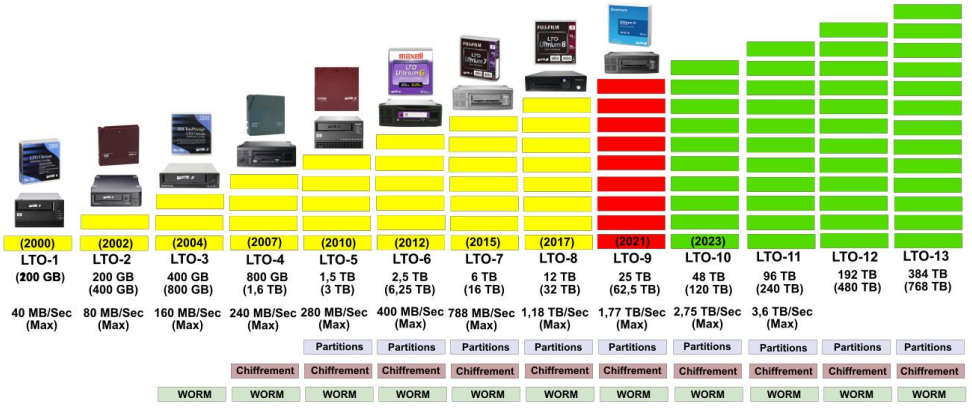
	Niveau 1	Niveau 2	Niveau 3	Niveau 4
Criticité	Négligeable	Non vital	Vital	Critique
Impact sur l'entreprise	Pas d'effet	Effet compensable non immédiat	Effet immédiat compensable	Arrêt de l'entreprise
Site principal	Configuration unique	Configuration de secours en plus	Clustering	Tolérance aux fautes, avec disponibilités temps réel (Cloud)…
RTO	> 8 h	30 mn - 1 h	5 - 30 mn	0
RPO	24 h	1 mn - 24 h	< 1 mn	0
Technologies utilisées	Pas de redondance machine. Sauvegardes classique, locales, Cloud, serveurs dédiés. Procédures de reprise simples.	Disques miroirs et RAID 5. Sauvegardes Cloud, serveurs dédiés.Procédures de reprise rigoureuse, PCA minimum	Clusters avec copies complètes planifiées. Fréquence élevéd de copies. Disques RAID 5, redondance réseau, Cloud, PCA/PRA	Ressources de secours prêtes à l'emploi, redondance totale, réseau, téléphone, alimentations. Tolérance aux fautes.
Services associés	Maintenance matérielle	Maintenance avec engagement de reprise	Audits de disponibilité. Surveillance proactive. Contrat de maintenance ("Business Critical")	Contrat de reprise PCA/PRA. SLA : serveurs, temps de réponse clients ("elapse").
Site de secours	Aucun	Ressources disponibles. MV sur Cloud, ressources locales.	Configurations dupliquées localement, 3-2-1 et Cloud.	Ressources prêtes à l'emploi, aucune latence. Tolérance de faute, tolérance aux fautes.
RTO	72 h	8 - 72 h	1 h	0
RPO	Plusieurs jours	24 h	< 5 mn	0
Technologies utilisées	Sauvegardes classiques, disques, bandes, Cloud	Données répliquées en mode asynchrone. Sauvegardes classiques disques, bandes, Cloud	Redondances de ressources, failover de bases de données, réplication synchrone de certaines données	Redondance de ressources. Tolérance aux fautes. Réplication synchrone des données sensibles, réplication

RTO : "Recovery Time Objective", durée acceptable pour relancer le service
RPO : Recovery Point Objective, délai de pertes de données que l'on ne doit pas dépasser

PCA/PRA : 1 sur 2 à la poubelle
(ce qu'il faudrait faire)

- Disposer de l'historique des problèmes rencontrés.
- Qui est en charge de quoi ?
- Identifier les applications et données prioritaires.
- Délimiter le périmètre PCA/PRA.
- Préciser les conditions de reprise, éventuellement dégradées.
- Définir une architecture globale de sauvegarde, avec des fréquences adaptées.
- Appliquer la règle 3-2-1, quand c'est possible.
- Tester le PCA, dans des conditions proches du réel.
- Construire un site de référence.
- Formaliser le PCA/PRA, dont le plan de sauvegarde et faire en sorte que le document soit à jour et accessible.
    - Les objectifs attendus.
    - Le périmètre et les personnes impliquées.
    - Procédures de récupération.
    - Liste exhaustive des tests à effectuer pour valider la reprise.
- Réévaluation régulière du PCA/PRA et des conditions de sauvegardes.

Et les bandes magnétiques...
Toujours présentes



Les évolutions prévisibles

Sauvegardes hybrides
Les sauvegardes locales, dupliquées sur un Cloud public, deviennent la règle. La capacité à sauvegarder de manière tranparente les données, quelle que soit leur localisation, est une nécessité.

Mise en adéquation des réseaux
L'une des améliorations va porter sur l'adéquation de la bande passante réseaux avec les liens entre le local et le Cloud. Règle de 3.

Pannes des grands acteurs
La confiance envers les grands acteurs : Microsoft, AWS, Google, etc, n'est plus totale. Même avec des niveaux de disponibilité très élevés, les acteurs du Cloud ne sont plus considérés comme une protection à 100 % (OVH...)

Apport de l'Intelligence Artificielle
Anticipation sur les besoins des usagers. Amélioration du RTO par analyse des conditions de sauvegardes. Meilleure compréhension des incidents. Allocation intelligente des ressources.

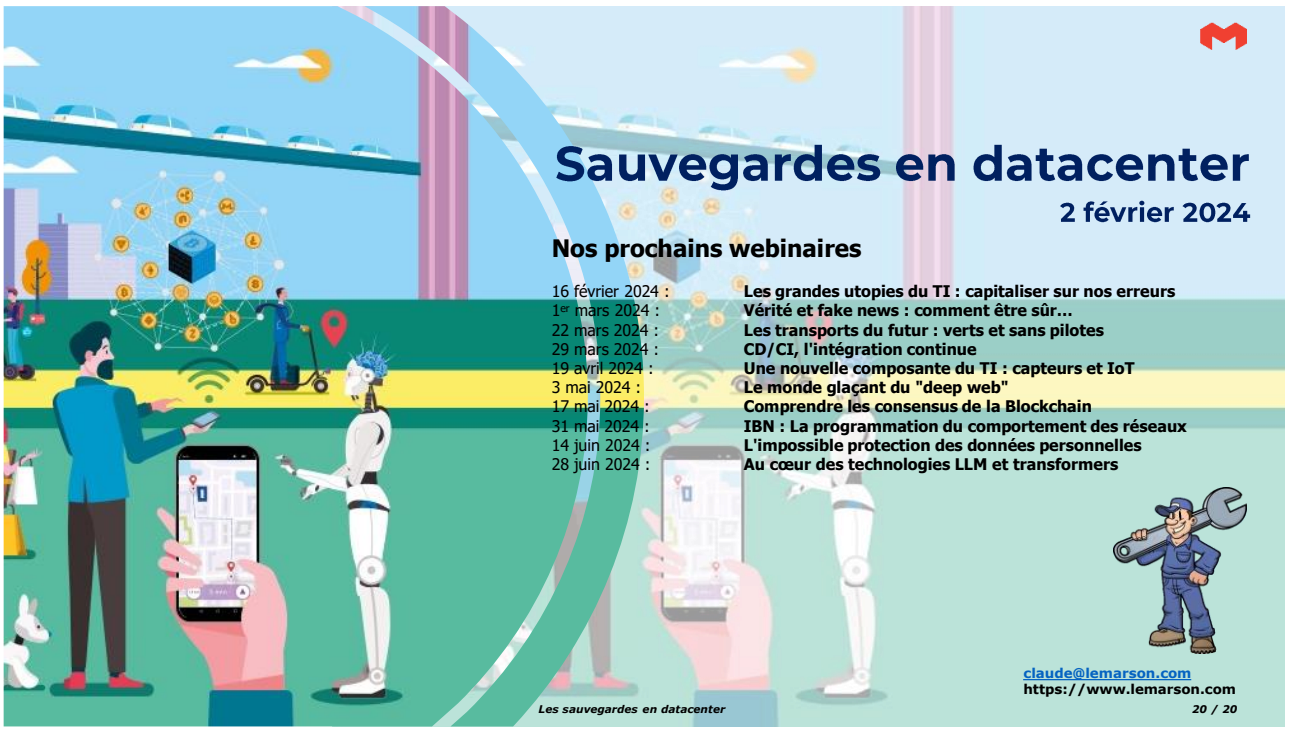
Les menaces de cybersécurité
Les attaques vont cibler de plus en plus les sauvegardes elles-mêmes. Le processus est largerment enclenché. Il va se poursuivre et s'amplifier, d'où la nécessité d'appliquer une stratégie de type zero-trust.

Retour en force des bandes
Rapport performance/prix imbattable. Très grandes capacités et progrès sur la conservation à long terme. Le mode "off line" déconnecte les données des cybercriminels. Efficacité énergétique.

Résilence (réactivité) maximum
L'objectif sera de se rapprocher de la tolérance zéro pour les applications et données critiques et donc de la résilience maximum. Des techniques telles que le CDP, vont y contribuer.

Conformité des sauvegardes
De gros progrès seront accomplis en matière de conformité aux règlementations de type RGPD.

Forte extension du mode BaaS.
La quasi-totalité des datacenters vont exploiter des services BaaS, le plus souvent pour améliorer les performances des plans de PCA/PRA.



Sauvegardes en datacenter

2 février 2024

Nos prochains webinaires

16 février 2024 :	**Les grandes utopies du TI : capitaliser sur nos erreurs**
1er mars 2024 :	**Vérité et fake news : comment être sûr...**
22 mars 2024 :	**Les transports du futur : verts et sans pilotes**
29 mars 2024 :	**CD/CI, l'intégration continue**
19 avril 2024 :	**Une nouvelle composante du TI : capteurs et IoT**
3 mai 2024 :	**Le monde glaçant du "deep web"**
17 mai 2024 :	**Comprendre les consensus de la Blockchain**
31 mai 2024 :	**IBN : La programmation du comportement des réseaux**
14 juin 2024 :	**L'impossible protection des données personnelles**
28 juin 2024 :	**Au cœur des technologies LLM et transformers**

claude@lemarson.com
https://www.lemarson.com