Storage Networking Industry Association

SNIA EUROPE

SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION

Technologies de Protection de Données de Nouvelles Générations (NGDP)

SNIA Europe
Comité Régional SNIA France
24 Avril 2007
IBM Forum

Agenda

- 9h30 A propos de la SNIA
  « Next Generation Data Protection »
- 9h45 Protection Continue de Données (CDP)
- 10h30 Virtualisation de Librairies de Bandes (VTL)
- *11h* *Pause*
- 11h15 Le Guide Stockage Tarsus
- 11h25 Réduction de Données
- 12h10 Protection des Postes et Bureaux Distants
- 12h40 Conclusion
- *13h* *Fin*

SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION

A propos de la SNIA

Philippe Nicolas
Chairman & Founder
francechair@snia.org

Historique, Mission et Activités SNIA

- **SNIA fondée en 1997, SNIA Europe en 2001**
  - « Association de l'Industrie du Stockage Réseau »
  - ~460 membres WW, ~100 Europe, ~7000 individus

- **Mission**
  - **Contribuer au développement et à l'adoption de solutions de stockage en réseau <u>efficaces, complètes et fiables</u>**
- **Activités**
  - Standardisation/Normalisation (SMI-S, XAM)
  - Initiatives Technologiques (SMF, DMF, IPS, Sécurité)
  - Education/Formation/Certification
  - Conférence SNW, Website & Newsletter
  - Technology Center & Workshops/benchmarks
  - Comités Régionaux

SNIA Data Management Forum

STORAGE & SAN CERTIFICATION — SNIA CERTIFIED

COMPUTERWORLD STORAGE NETWORKING WORLD — SNIA

Quelques annonces et actions récentes

- SMI-S nouveau standard international ISO et IEC, suite ANSI
- Alliance SNIA et Eclipse Foundation autour de SMI-S et du projet Aperi
- Développement de XAM (eXtensible Access Method) avec XAM SDK, Démo réelle pour Automne 2007 (SNW) et soumission ANSI et ISO
- Création de 2 Task-Forces: Grid et FAN
- Enquête Utilisateur « Meilleur Interopérabilité »
  - http://advisorygroups.snia.org/home/SNIA_End_User_Council_2006_Survey_Results_Report.pdf

Conférence SNW @San Diego 16-19 Avril 2007

- >2000 visiteurs
- 153 intervenants et 143 sessions éducatives (~60 Tutoriaux SNIA)
- 120 exposants dont 83 sponsors

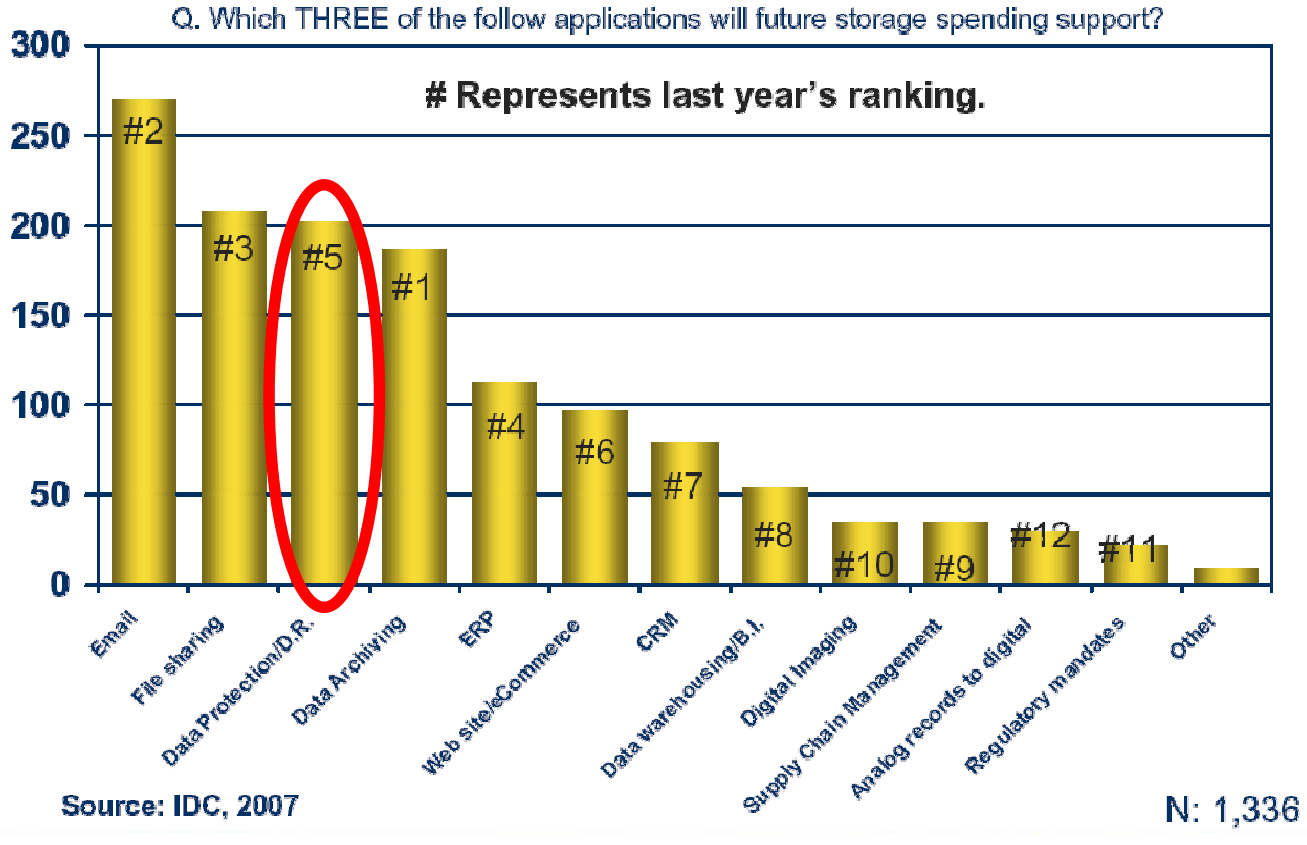
- Sujets: Virtualisation, DP (Rec. Mgmt., DéDup.), Grid/Cluster Storage, FC/IP/Eth/IB…, Green IT

- Prochaines dates
  - USA: 15-18 Oct. 2007 à Dallas (www.snwusa.com)
  - Europe: 29-31 Oct. 2007 à Francfort (www.snweurope.com)

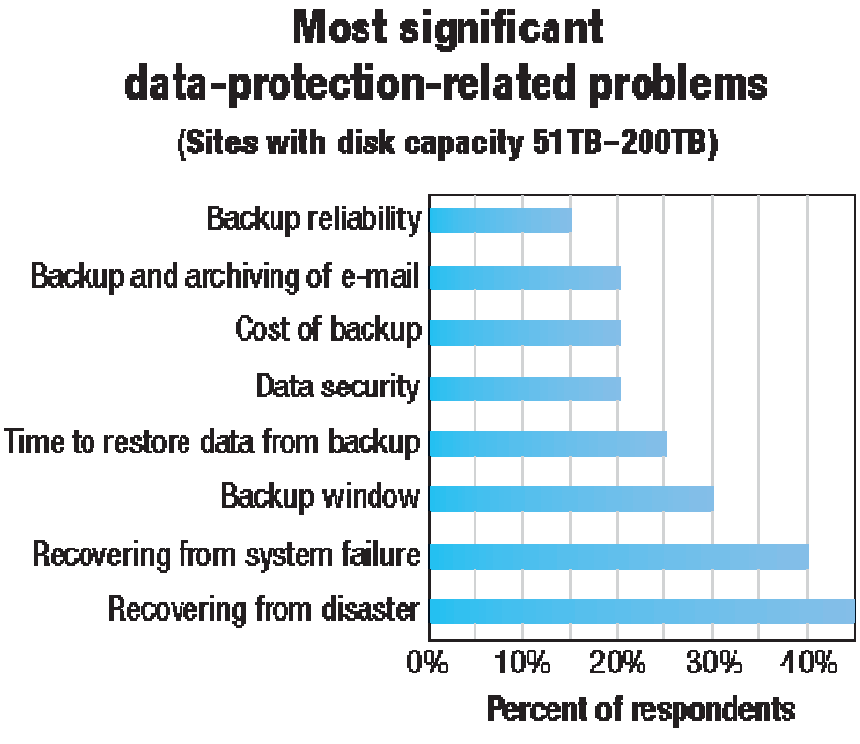
SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION

Technologies de Protection de Données de Nouvelles Générations (NGDP)

Dépense Européenne Future pour le Stockage



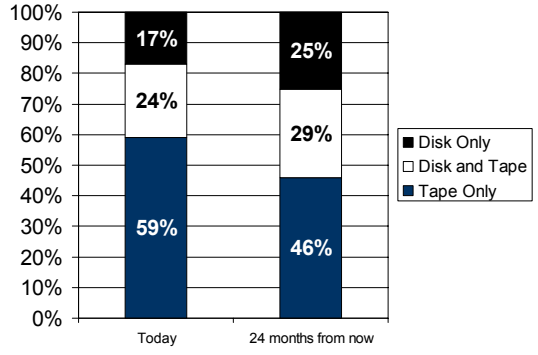
Besoins Utilisateurs



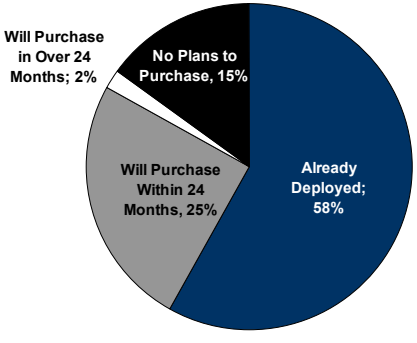
Source: Peripheral Concepts

- Protéger un volume toujours croissant (GB, TB) dans le Data Center et au-delà (bureaux et postes distants ou nomades)
- Augmenter la fraicheur des données (RPO)
- Réduire le temps de reprise (RTO)
- Réduire la fenêtre de sauvegarde (DPW)
- Chiffrer/Crypter les données
- « Externaliser » les données pour éviter l'impact d'un désastre
- Et bien sûr réduire le coût de la protection ($£€)
- Dans un contexte de contraintes de Conformité Réglementaire (SOX, SEC 17a-4, HIPPA, FDA, EU Data Retention Directive…)…

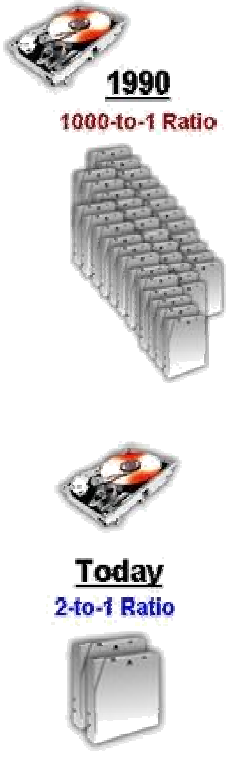
Le Disque comme Média de Protection

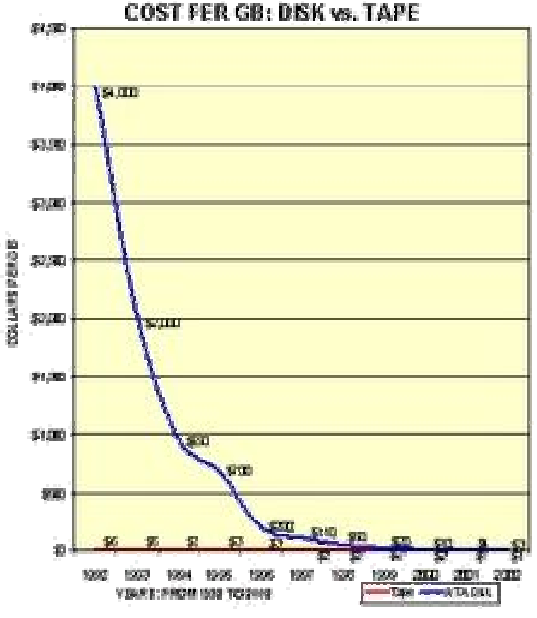


Data backed up to disk only, disk + tape, tape only



DBDP Adoption Rates



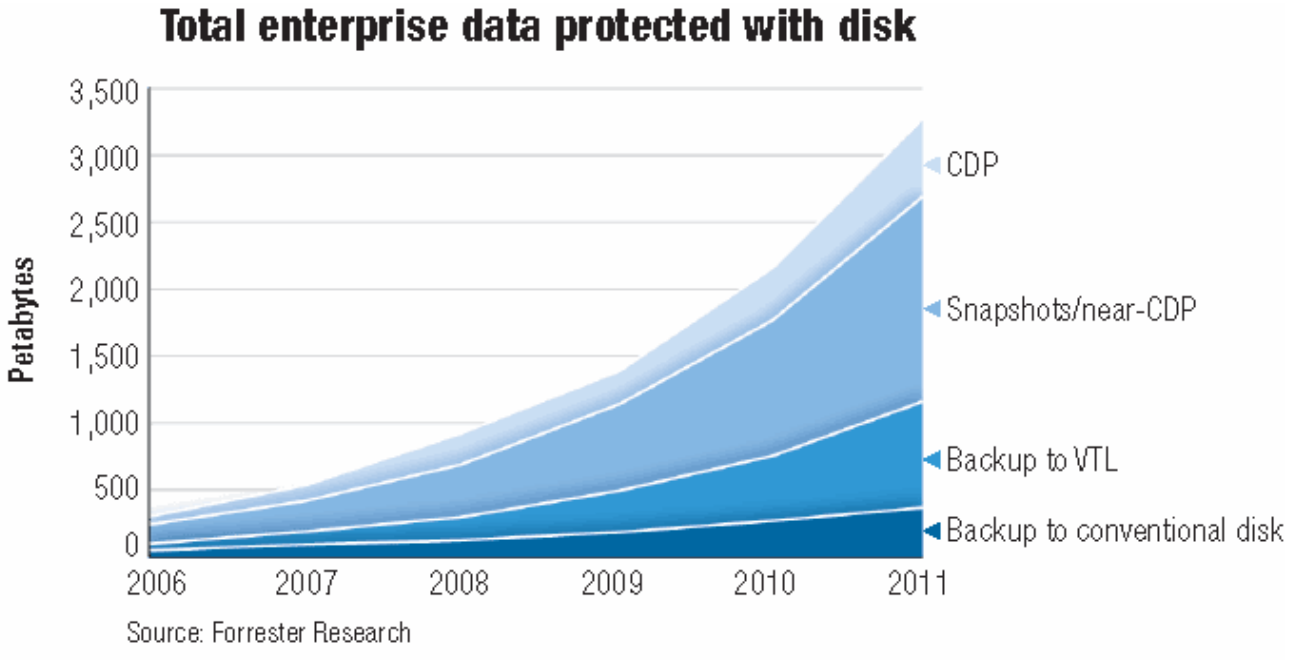


Source: Enterprise Strategy Group, The Evolution Of Enterprise Data Protection: User Adoption Of Next-Generation Backup & Recovery Technologies

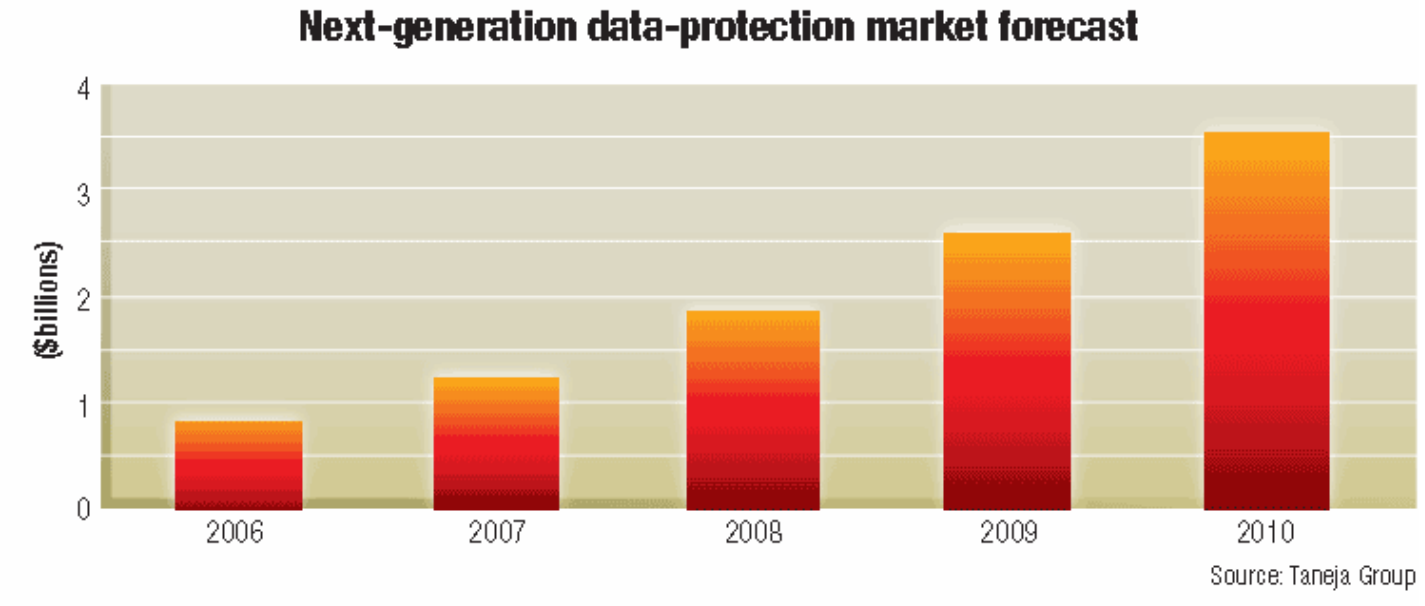
Tendances et Réponses du Marché

- Apparition du disque dans le processus de backup (staging, **VTL**…) pour réduire le RTO
- Couplage de technologies de **snapshot**
  données pour réduire la
- **Protection**
  Nouvelle et Prochaine Génération de Solutions de Protection de Données (NGDP)
  de Données
  graphie)
  pour les environnements distants et
  nodes
- **Chiffrement** des données externalisées
- Techniques de classification de données, ILM
- Nouveaux supports pour l'archivage (approche contenu)

L’évolution de l’adoption par le marché



Evolution du Marché des NGDPs



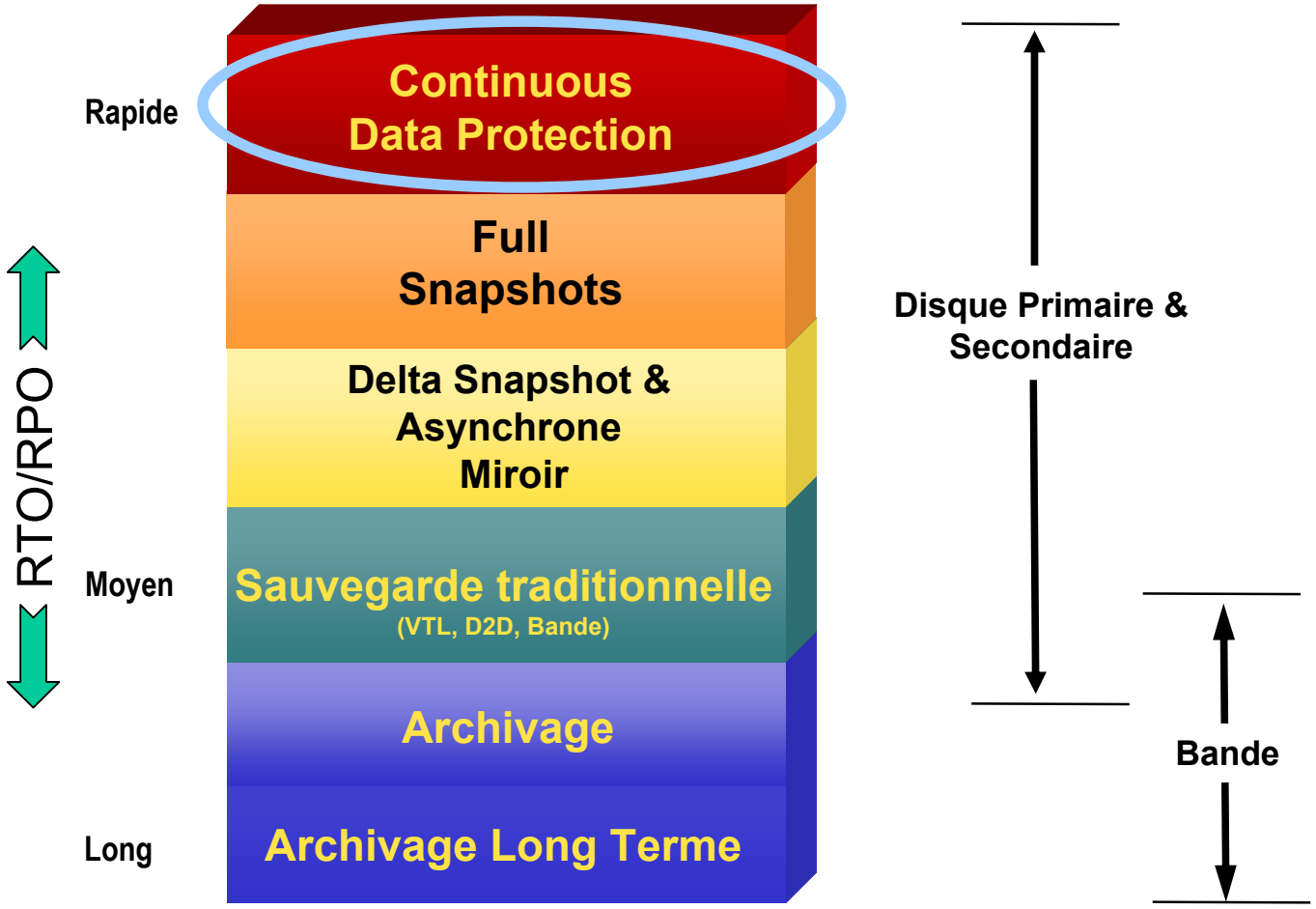
x4 en 3 ans (2006/2009 – 2007/2010)

SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION

Protection Continue de Données (CDP)

Philippe Nicolas
Chairman & Founder
francechair@snia.org

Les différentes solutions de protection

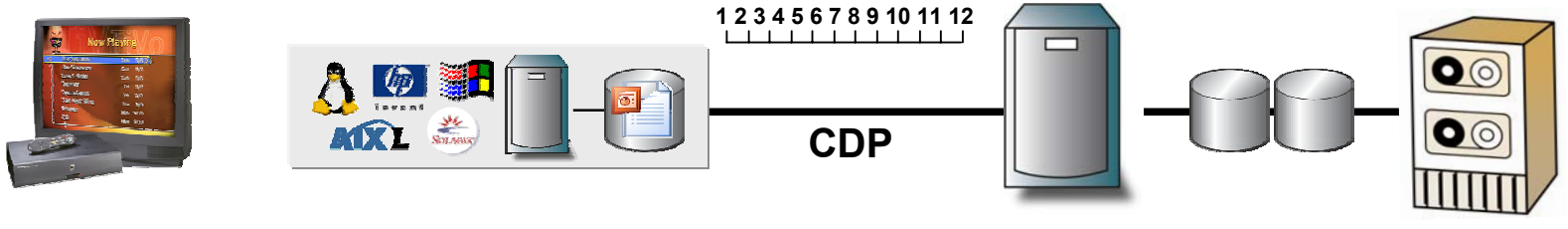


RTO Objectif de Temps de Reprise – **RPO** Objectif de Point de Reprise

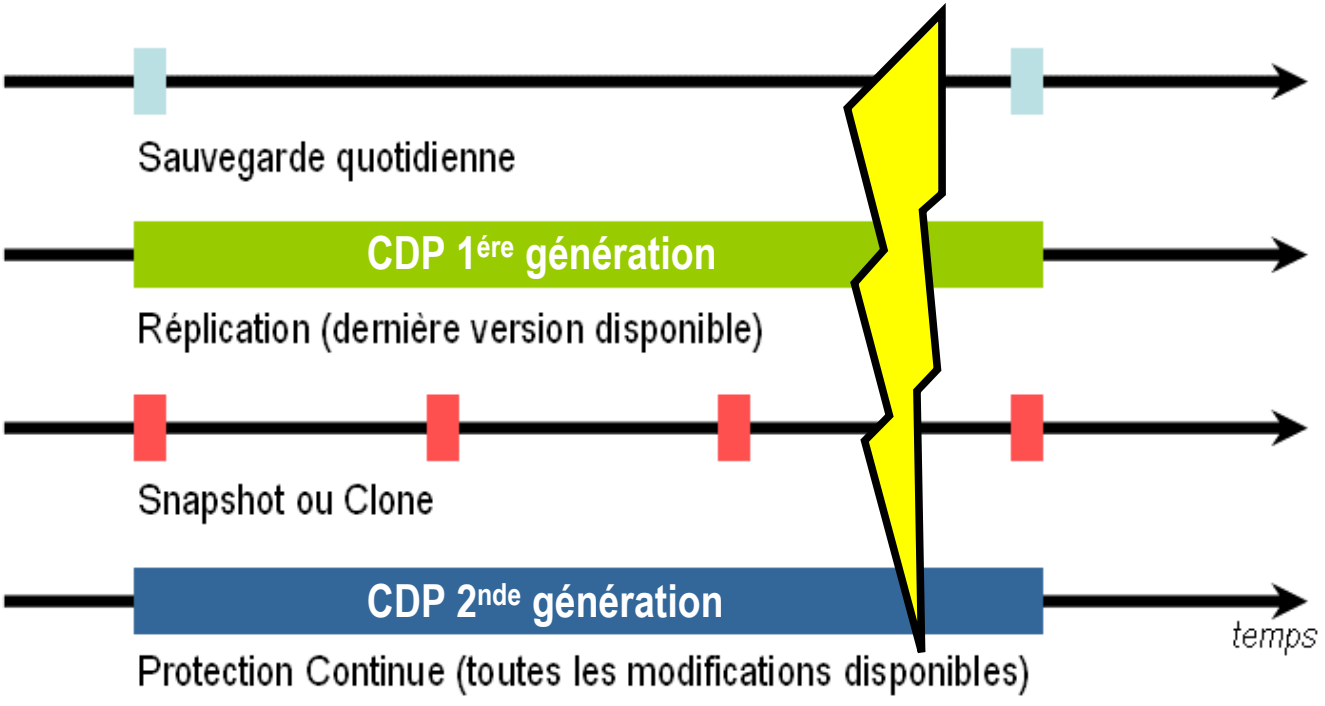
Définition du CDP par le SNIA DMF

"Continuous data protection (CDP) is a methodology that continuously captures or tracks data modifications and stores changes independent of the primary data, enabling recovery points from any point in the past.
CDP systems may be block-, file- or application-based and can provide fine granularities of restorable objects to infinitely variable recovery points."

http://www.snia-dmf.org/cdp/index.shtml



Capture de données, RPO et RTO



Le CDP en détail

- CDP impose que l'on parle de Protection de Données et non de Sauvegarde de Données
  - Les outils de sauvegarde peuvent offrir la fonction CDP

- Notion de True/Near/Dual CDP avec RPO/RTO très faible ou nul
  - Granularité importante
  - RPO – APIT* et non SPIT**
    - Dégradation possible en SPIT
    - Notion d'Event Addressable Storage (EAS)
  - RTO – Concept de TAS (Time-Addressable Storage)
    - « Accès aux données – stockage – sur critère de temps »

- Elimination de la fenêtre de sauvegarde

* APIT Any Point In Time image
** SPIT Single Point In Time image

Le CDP en détail (2)

- Cas d'utilisations
  - Serveur (fichiers, applications, SGBD…)
  - Poste de travail
  - Local ou distante
- Trés varié en stockage, connectivité, protocoles…
  - primaire, secondaire…, FC/xATA…
- Niveau bloc, fichier ou application

Le CDP en détail (3)

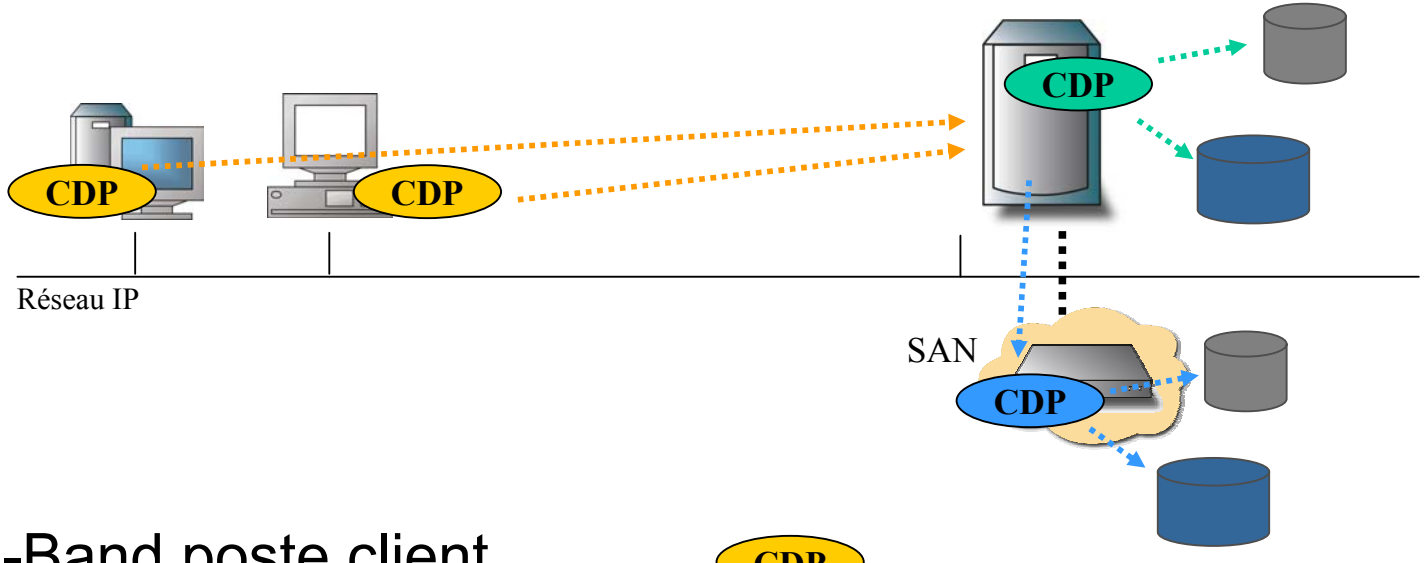
- Approches
  - Logicielle (application ou système)
  - Matérielle (switch, appliance, array/contrôleur)

- Méthodes
  - « Image avant » (CoW, RoW, Write-anywhere)
  - « Image après » (journalisation)
  - Méthodes mixtes

Le CDP en détail (4)

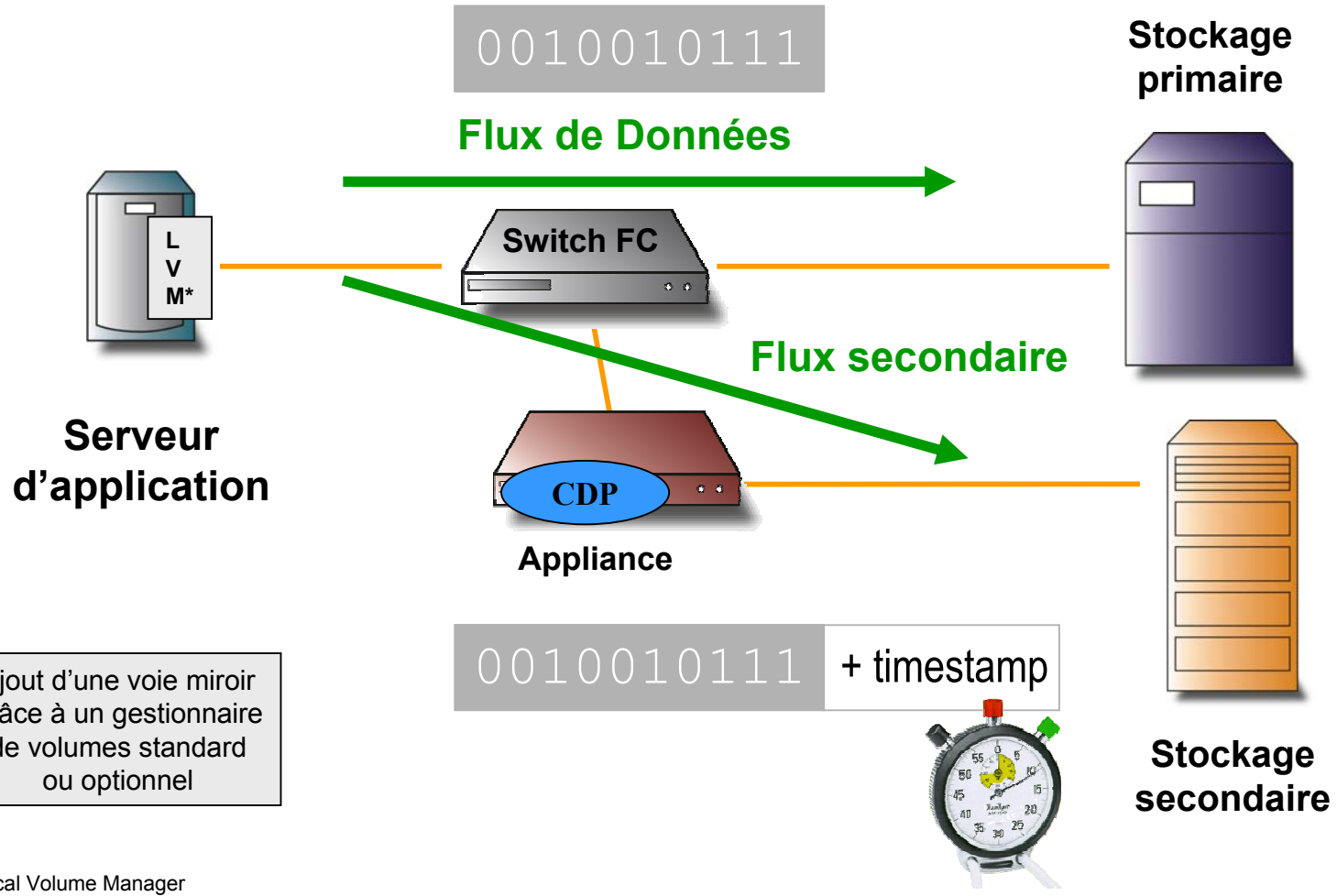
- Architectures
  - In-Band (dans le chemin de données – 1 seul chemin)
    - Intelligence CDP essentiellement embarquée dans un serveur ou poste client, éventuellement dans un switch
  - Side-Band (« forme de In-Band »)
    - Copie des données au travers de la couche CDP en *parallèle* du stockage primaire (miroir RW ou WO) – au moins 2 chemins
    - Plutôt une offre périphérique (switch, appliance)
  - Out-of-Band (2 alternatives)
    - Agent de duplication d'opération sur host avec envoi des données vers le volume primaire et le moteur CDP ou
    - Pas d'agent sur host mais intégration au plateforme switch évoluée qui joue le rôle d'agent de duplication des données/opérations
    - Écriture synchrone sur tous les miroirs (primaire + log des data sur secondaire et meta-data gérée par le moteur CDP)

Architecture CDP In-Band



- In-Band poste client CDP
- In-Band serveur CDP
- In-Band appliance/switch… CDP

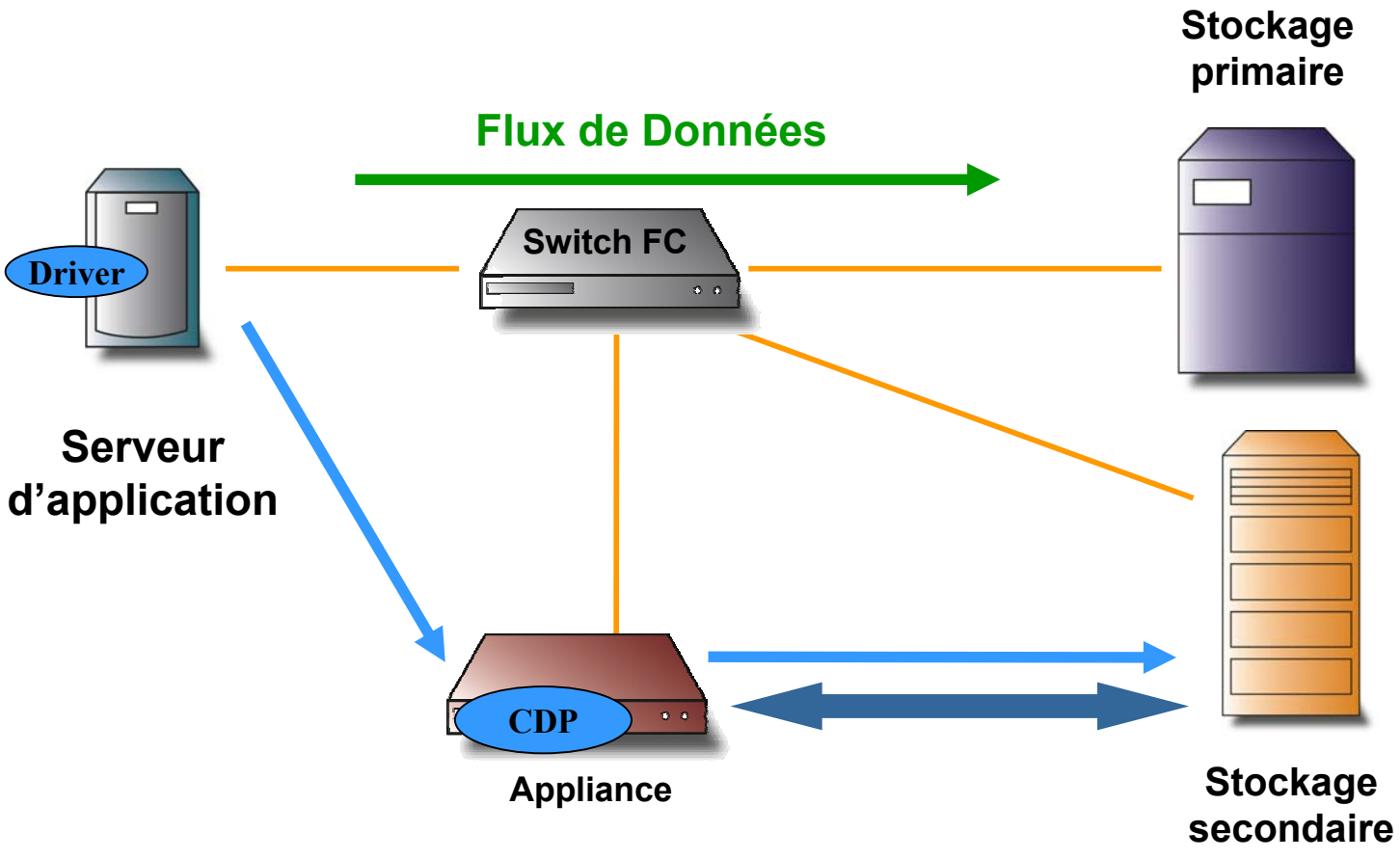
Architecture CDP Side-Band



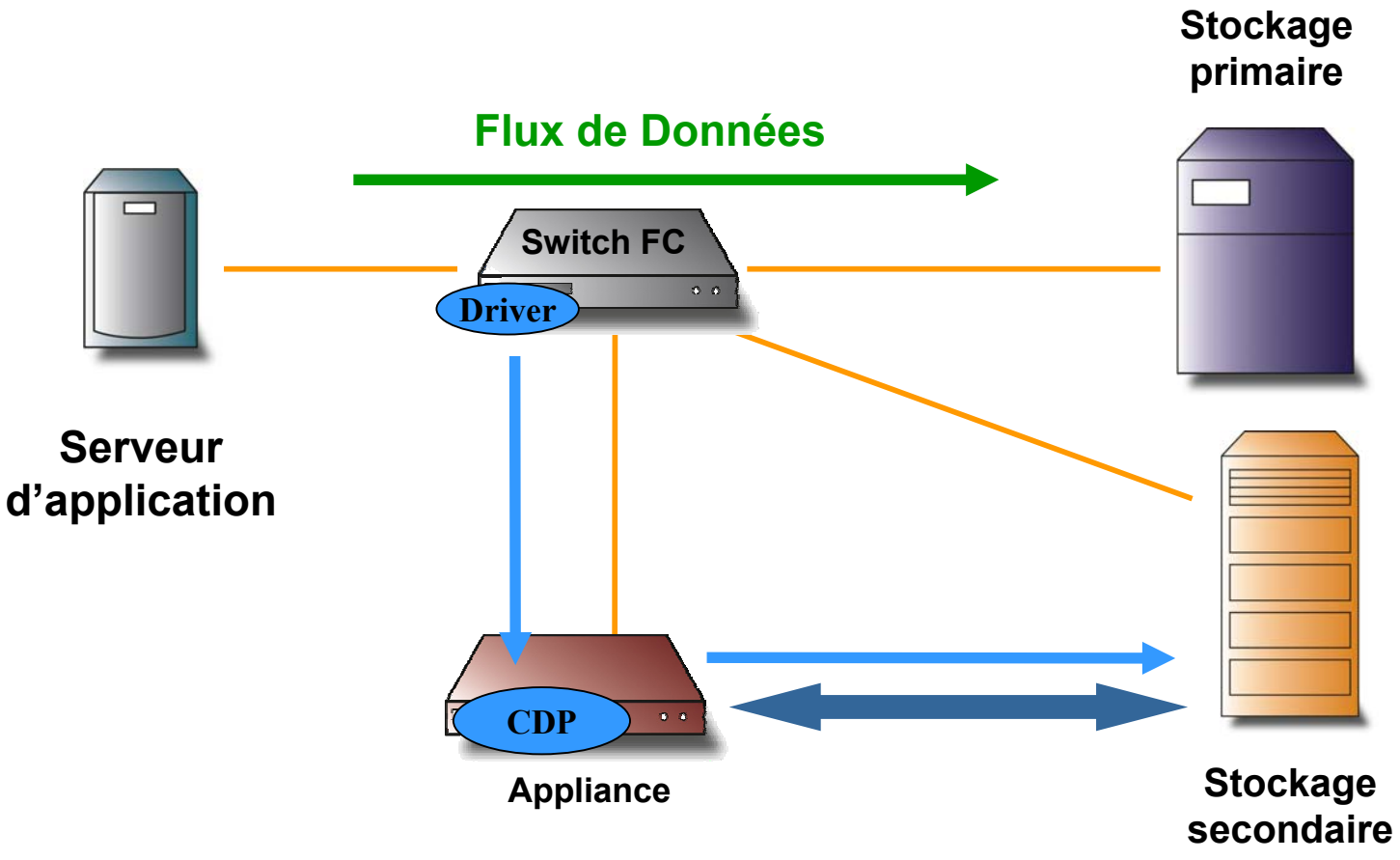
Ajout d'une voie miroir grâce à un gestionnaire de volumes standard ou optionnel

* Logical Volume Manager

Architecture CDP Out-of-Band



Architecture CDP Out-of-Band (2)



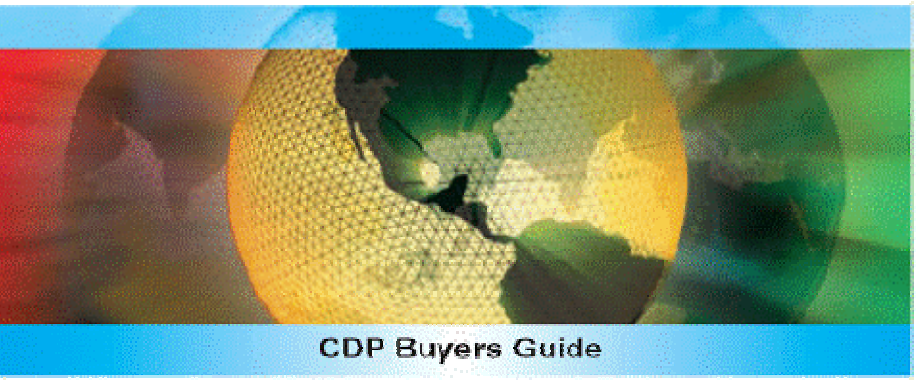
Un groupe spécialisé et dédié au CDP

The SNIA Data Management Forum Establishes
Continuous Data Protection Special Interest Group

14 sociétés fondatrices:
Alacritus (NetApp), EMC, Hitachi Data Systems, InMage Systems, Mendocino Software, Mimosa Systems, NetApp, TimeSpring, Revivio (Symantec), Scentric, Storactive (Atempo), Sun Microsystems, VERITAS Software (Symantec) et XOsoft (CA)

"Over the years, data protection technologies have evolved from simple backups to advanced techniques like mirroring, snap-shots, remote replication, disaster recovery and archiving," said DMF Chair Sheila Childs. "As the newest weapon in the data protection arsenal, Continuous Data Protection can offer additional protection capabilities that eliminate backup windows and provide exceptional recovery performance and granularity."

www.snia-dmf.org/cdp/index.shtml



An overview of today's Continuous Data Protection (CDP) solutions

September 2006 – Second Edition

CDP Buyer's Guide
Second edition – September 2006

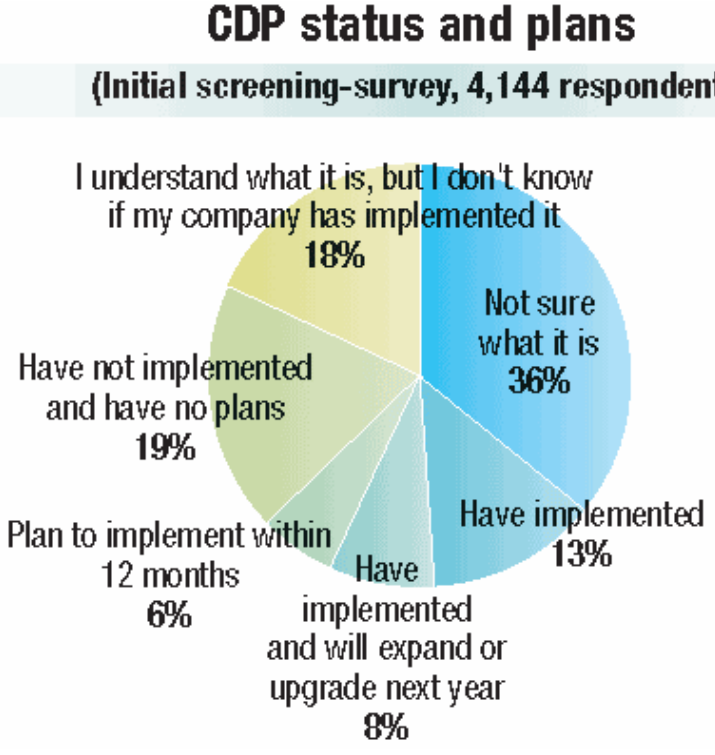
PUBLISHED BY
The Storage Networking Industry Association
SNIA DMF CDP Special Interest Group
500 Sansome Street, Suite #504
San Francisco, CA 94111

Project Editor: Brian J. Olson (brian.olson@atempo.com)

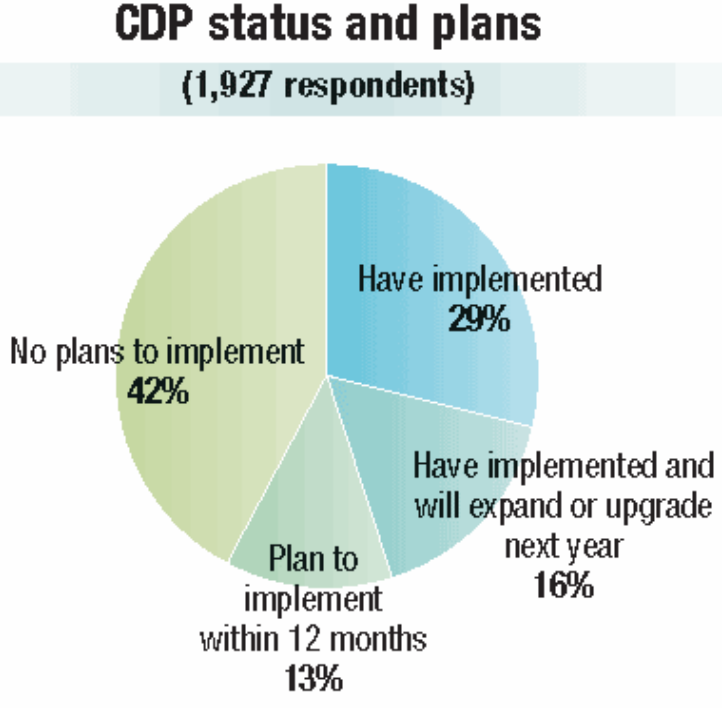
Copyright © 2006 by the Storage Networking Industry Association (SNIA), USA. All rights reserved. No part of the contents of this guide may be reproduced or transmitted in any form or by any means without the written permission of the publisher. Created in USA. The SNIA logo is a trademark of SNIA. Other features or product names mentioned are trademarks of SNIA or other vendors or manufacturers.

Disponible sur www.snia-dmf.org/cdp/index.shtml

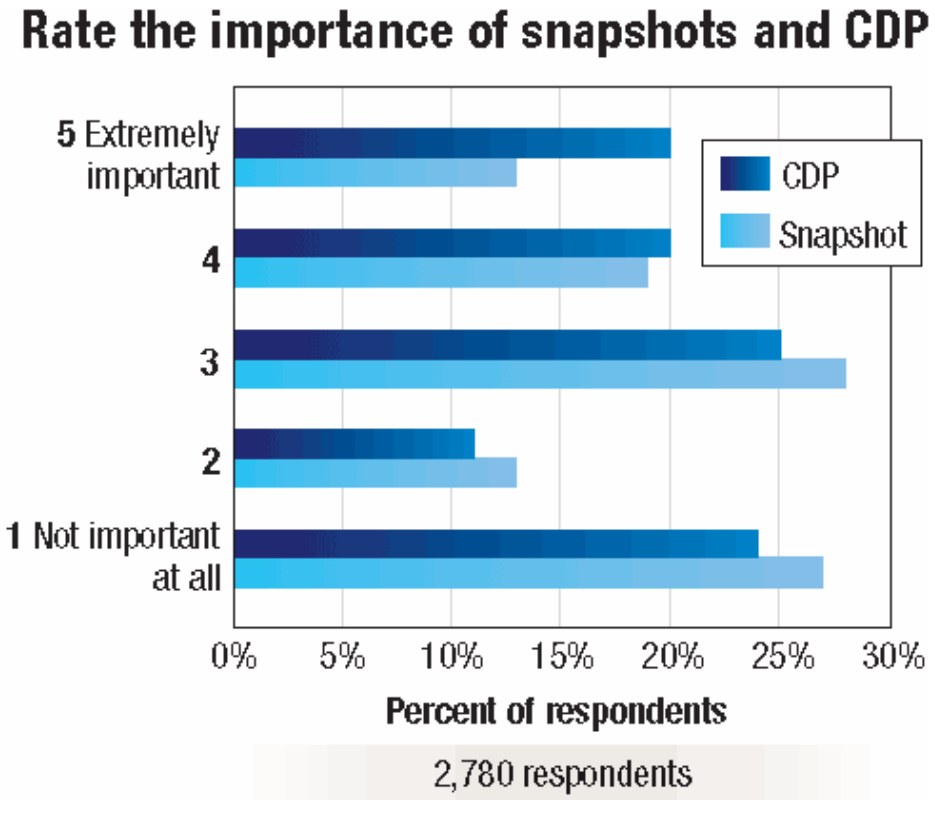
Etat de l'adoption



Source Peripheral Concepts

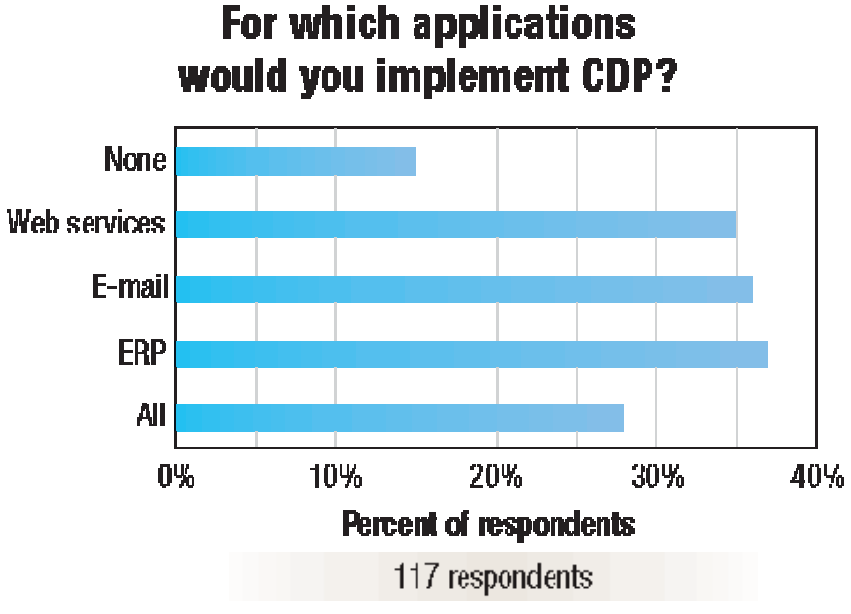


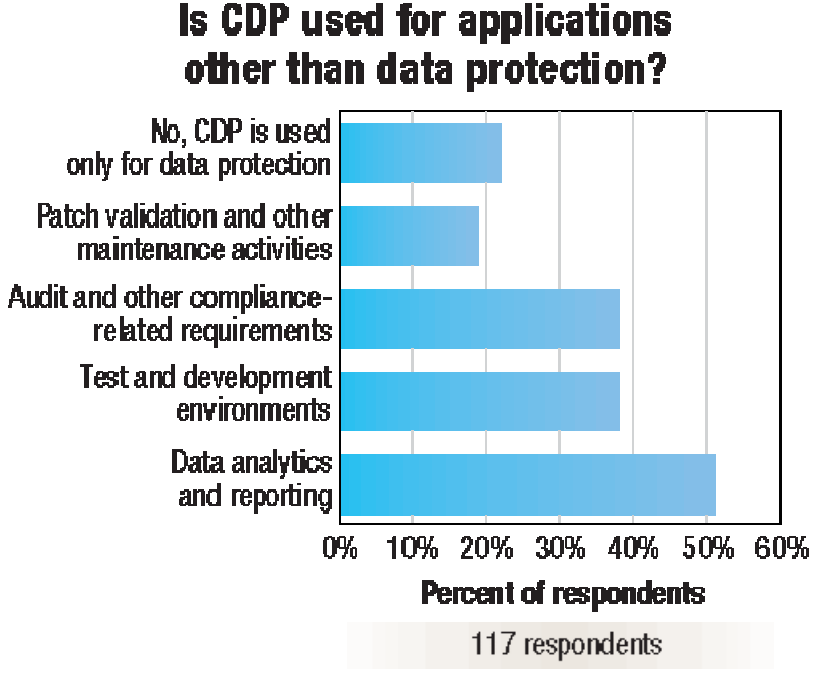
Importance du CDP et du Snapshot



Source Peripheral Concepts

Le CDP, pourquoi faire ?





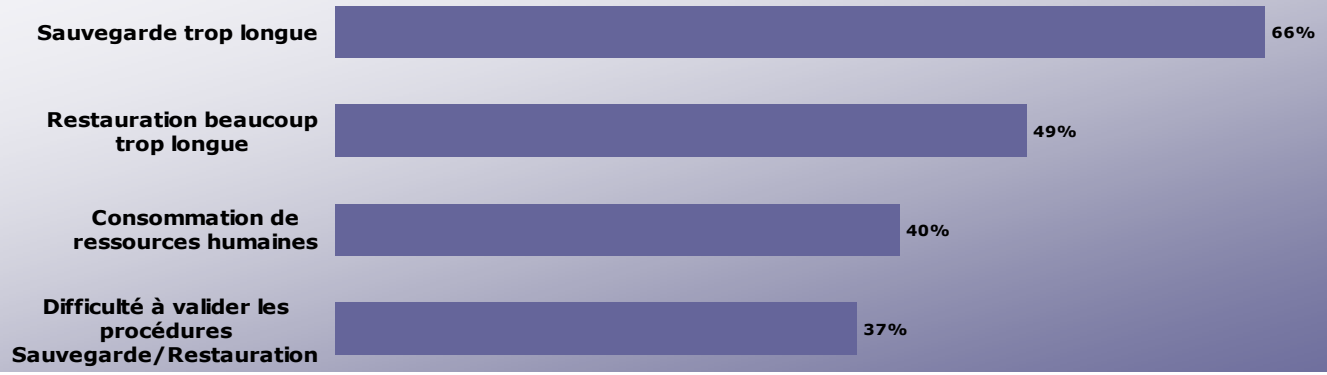
Source Peripheral Concepts

SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION

Virtualisation de Librairies et Lecteurs de bandes (VTL)

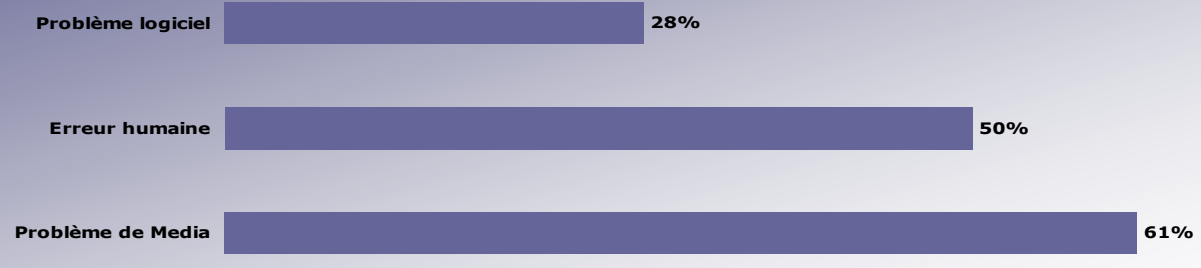
Jean-Baptiste Fuster
Technical Chair – SNIA Europe Comité France

Principaux problèmes des solutions de Sauvegarde/Restauration



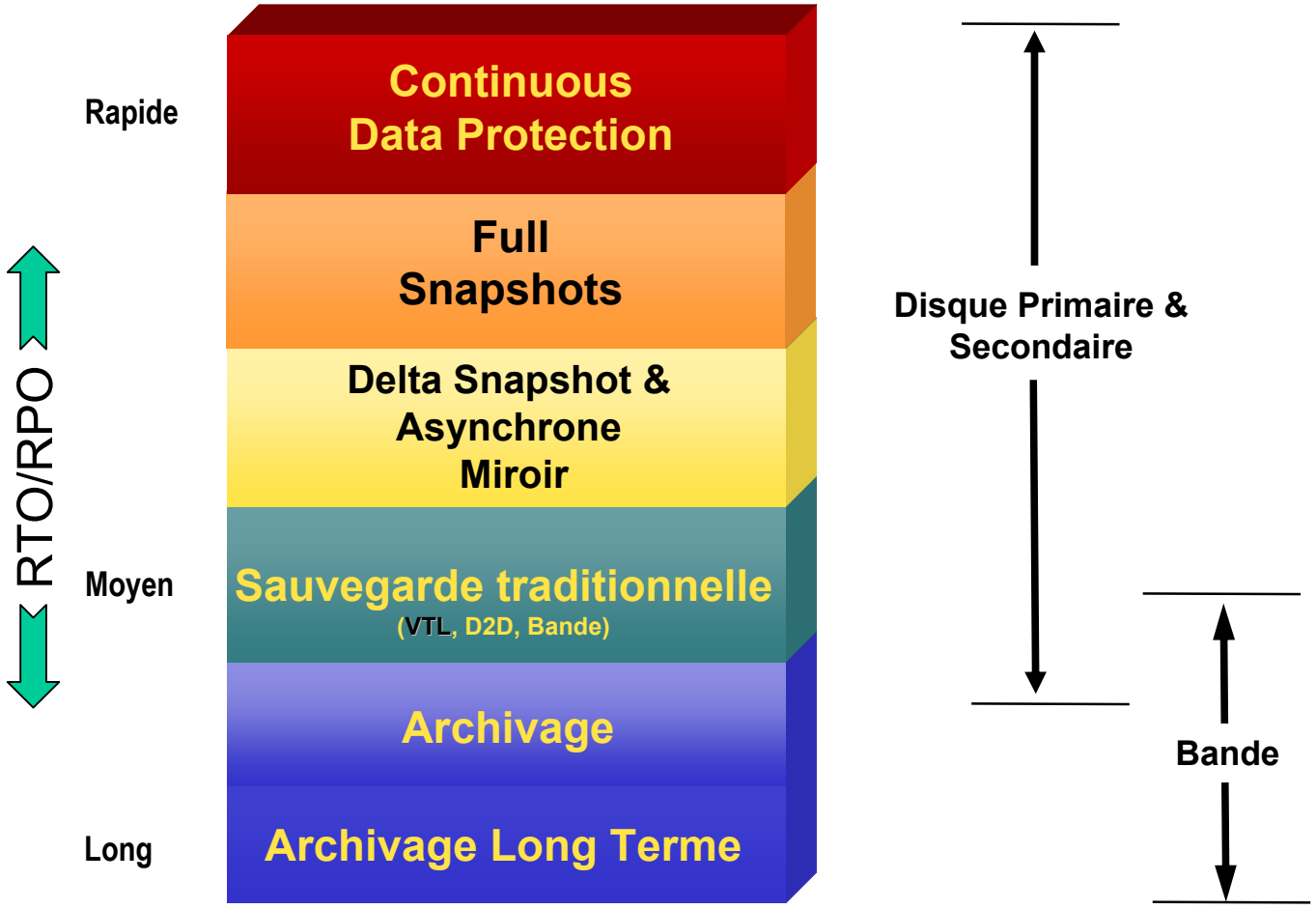
Echantillon de 222 utilisateurs
Plusieurs réponses acceptées
Source:

Même si vous avez défini un PSI, pouvez-vous vraiment restaurer vos données?
Ci-dessous les causes principales d'echec de restauration



Echantillon de 135 utilisateurs
Plusieurs réponses acceptées
Source:

Les différentes solutions de protection



RTO Objectif de Temps de Reprise – **RPO** Objectif de Point de Reprise

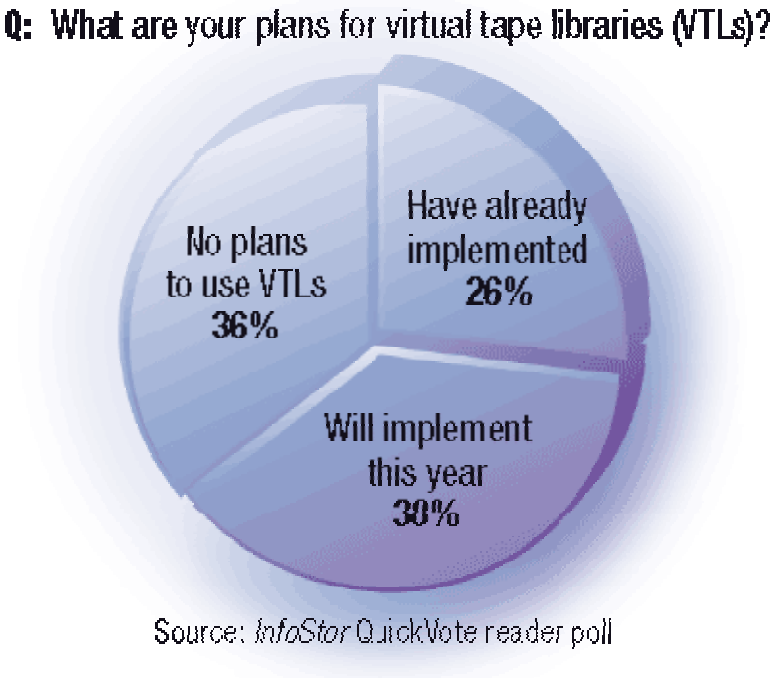
Variété des supports de sauvegardes …

- **Traditionnelles sur dispositifs de bande**
  Les logiciels de sauvegarde sont installés sur les serveurs (dits serveurs de sauvegarde) et utilisent des dispositifs de bande physique

- **Disques cibles (disk targets ou espace de staging)**
  La majeure partie des éditeurs de sauvegarde fournissent des APIs de sauvegarde sur disque et certains fabricants fournissent même des technologies de réduction de données basées sur l'analyse de contenu

- **VTLs (librairies de bandes virtuelles)**
  Certains fabricants fournissent des émulations de librairies de bandes au travers de SCSI, SAN-FC et SAN-IP. Ces VTLs sont indépendantes des serveurs de sauvegarde

Réalité chez les utilisateurs



64% d'utilisateurs d'ici 12 mois

Estimation du marché

- Marché VTL = 1,4 Mds $ en 2011
  - Doublement en 5 ans
- Capacité totale des unités VTL vendues
  - 90325 To en 2006, ~700000 To en 2011
- Décollage du monde Open en premier puis VTL grand système

Source IDC

La Virtualisation de bandes…

The act of creating abstracted tape devices by applying **virtualization** to tape drives or tape libraries

Source: SNIA Dictionary

"The act of integrating one or more (back end) services or functions with additional (front end) functionality for the purpose of providing useful abstractions

Typically virtualization hides some of the back end complexity, or adds or integrates new functionality with existing back end services. Examples of virtualization are the aggregation of multiple instances of a service into one virtualized service, or to add security to an otherwise insecure service. Virtualization can be nested or applied to multiple layers of a system"

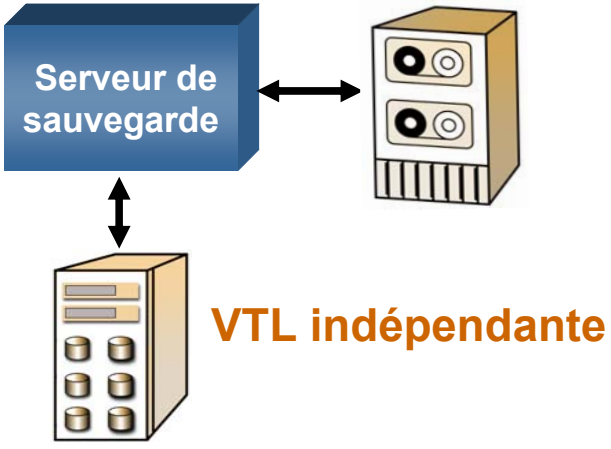
La Virtualisation de bandes…

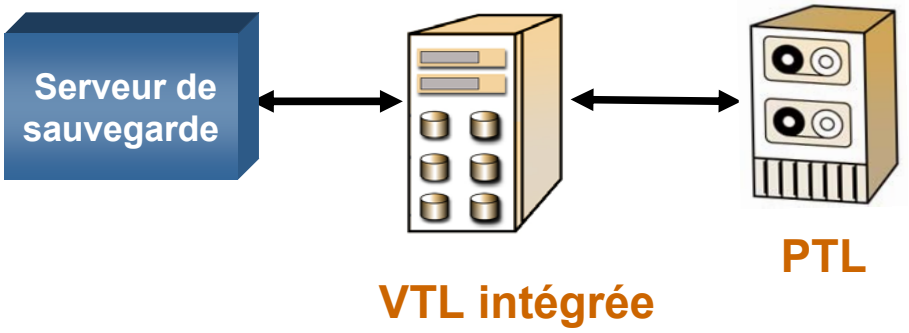
- Courante dans les environnements mainframe
  - Utilisations multiples
  - Plusieurs ensembles de données ayant un cycle de vie de longue durée

- Aujourd'hui disponible pour les environnements ouverts
  - Evite les changements de logiciels de sauvegarde (mais peut nécessiter des licences logicielles supplémentaires)
  - Intégrée, configurée et optimisée

- Problèmes des bandes à résoudre dans les environnements ouverts :
  - **Performances** : Réduire les fenêtres de sauvegarde et restauration
  - **Gestion** des logiciels de sauvegarde et restauration (temps et coûts)
  - **Coût** avantageux des bandes par rapport aux disques se réduit

VTL indépendante ou intégrée





Indépendante

- “se fait passer” pour une autre librairie de bandes
- Les logiciels de sauvegarde utilisent des bandes virtuelles comme des bandes physiques
- Les logiciels de sauvegarde continuent de cataloguer les données et les bandes. Ils sont seuls à gérer les « Meta Data »

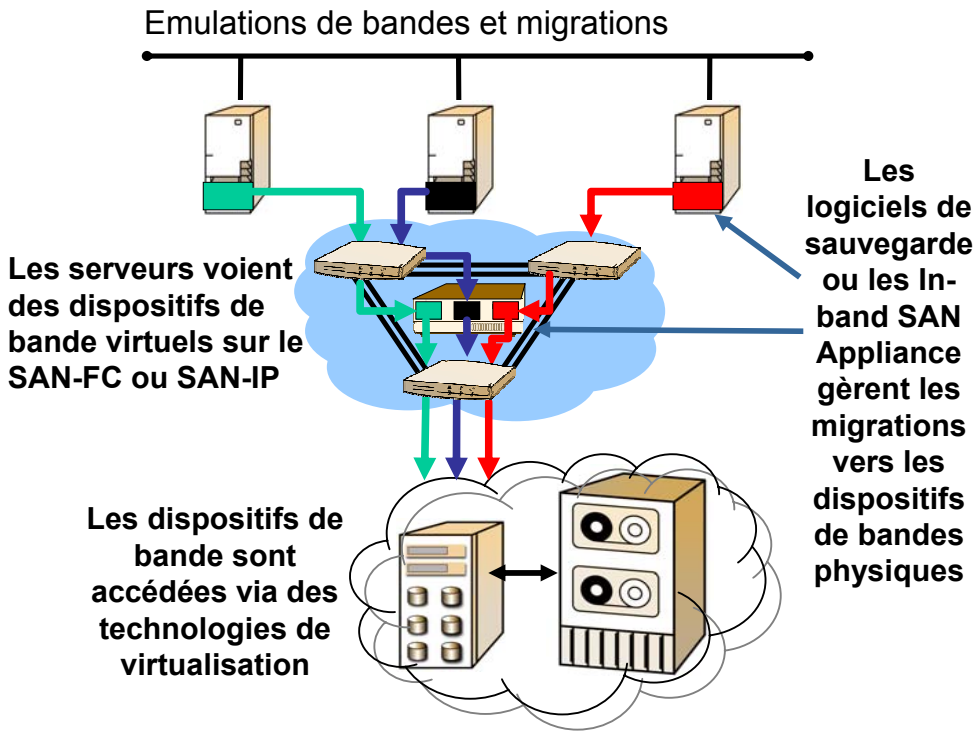
Intégrée

- “se fait passer” pour votre librairie de bande physique (PTL)
- Sauvegarde sur les bandes virtuelles
- Migration et copie automatique des media virtuels vers les media physiques avec cohérence des code barres ou copie physique

Comparaison d'une bande physique et virtuelle

Avantages d'une bande virtuelle	Avantages d'une bande physique
• Nombre illimité de… – Librairies Virtuelles – Slots Virtuels – Media Virtuels – Lecteurs Virtuels • Technologies de formats d'écritures transparents • Parallélisation des tâches de sauvegarde et restauration • Performances accrues de sauvegarde/restauration	• Amovible et transportable pour… – Stockage externalisé – Distribution du contenu – Archivage long terme • Compression hardware intégrée • Contraintes de normalisation – Media non réinscriptible WORM – Non modifiable • Supposé robuste

Fonctionnalités des VTLs (1)

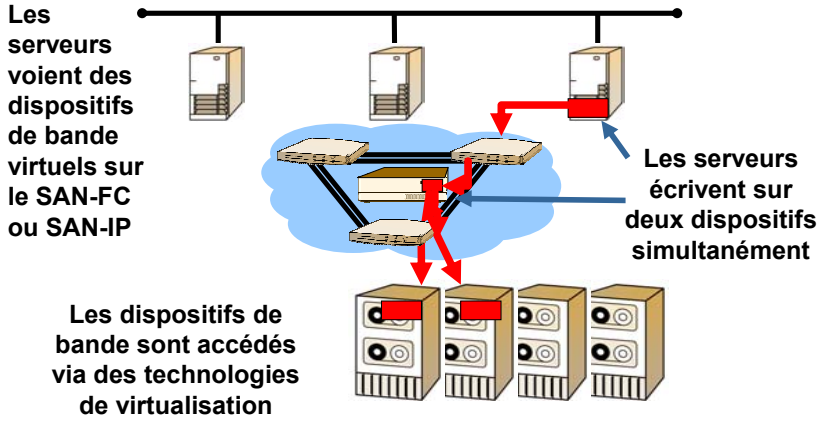


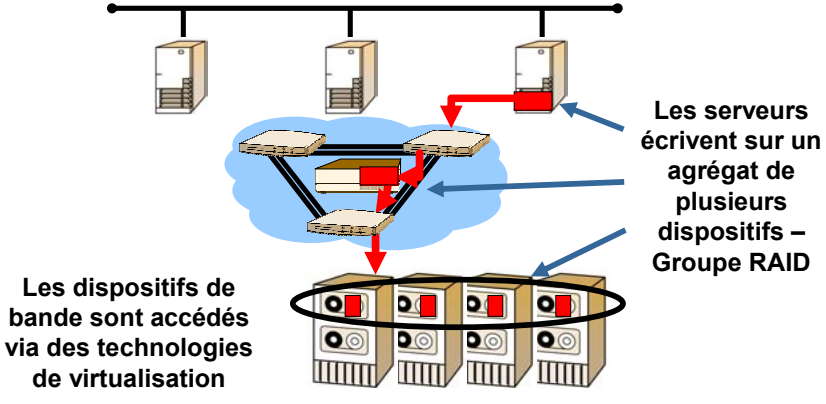
- **Emulation de dispositifs bande sur disque**
  - Le stockage disque héberge les volumes de bandes émulées
  - **Elimine la surcharge** des opérations mécaniques
- **Migration automatique de volumes bandes virtuelles vers des bandes physiques**
  - Dés que la politique de rétention mise en place le requiert
  - Si certains seuils de stockage sont atteints
  - Pour répondre aux politiques d'externalisation ou aux procédures de Disaster Recovery

Fonctionnalités des VTLs (2)

Duplication de bandes/Externalisation

- Améliore la disponibilité des données
- Ecriture synchrone ou asynchrone sur plusieurs dispositifs
  - Assure la disponibilité d'au moins une bonne copie
  - Permet de répondre aux procédures de Disaster Recovery
  - Crée automatiquement une copie pour externalisation





Ecriture sur un agrégat de bandes

- Améliore les performances
- Ecriture simultanée sur plusieurs dispositifs
- Création de groupes RAID de bandes pour améliorer l'intégrité et la disponibilité des données

Bénéfices des VTLs …

- Améliore le taux de réussite des opérations de sauvegarde et restauration en éliminant les erreurs de media et les pannes de robotique
- Augmente la vitesse de sauvegarde/restauration
- Réduit les coûts matériels et de Supervision en centralisant la gestion et la fourniture de ressources de sauvegarde
- Simplifie le déploiement de Services de sauvegarde
- Facilite l'externalisation des sauvegardes (via réplication IP ou FC)
- Préserve les investissements existants en matière de politique de sauvegarde et de connaissances des équipes techniques opérationnelles

Tendances

- Couplage de la fonction VTL à
  - La réplication
  - La déduplication
  - Au chiffrement…

- Volonté d'agréger les fonctions au sein d'une couche intelligente de traitement de la donnée

SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION

Pause 15’

Agenda

- 9h30 SNIA - Rappel Mission et Actualités « Next Generation Data Protection »
- 9h45 Protection Continue de Données (CDP)
- 10h30 Virtualisation de Librairies de Bandes (VTL)
- *11h* *Pause*
- 11h15 Le Guide Stockage Tarsus
- 11h25 Réduction de Données
- 12h10 Protection des Postes et Bureaux Distants
- 12h40 Conclusion
- *13h* *Fin*

SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION

« Le Guide Storage Tarsus »

Yannick Villain
Tarsus

Le guide Storage 2008

Le guide des p~~rofessionnels du stockage~~ des données

TARSUS FRANCE

Un rédactionnel sur les tendances et perspectives du marché réalisé par le comité régional SNIA France

La présentation [illegible] secteur

Les acteurs du marché classés par rubrique

La présentation des produits et des services des prestataires référencés

Un panorama général des prestataires

Comment est-il structuré ?

19 Salons professionnels

14 Guides professionnels

Le guide Storage 2008

Le guide des professionnels du stockage des données

- Éditeur
- Administration / Supervision
- Archivage
- Constructeur
- Gestion des données
- SSII / Intégrateur
- Connectivité
- Audit / Conseil / Expertise
- Périphériques de stockage
- Prestataires de services de stockage
- Continuité d'activité
- Formation
- Service associé au stockage
- Sécurité et Chiffrement

La présentation des développeurs et des produits

Les services des prestataires

Optimisez votre parc

Publiez vos références

Le guide Storage 2008

Le guide des professionnels du stockage des données

Le guide Storage est remis en mains propres à 6000 exemplaires aux visiteurs du salon Storage Expo, les 21 et 22 novembre 2007 et sur les conférences du salon Solutions LINUX, les 29, 30 et 31 janvier 2008 (également diffusé sur le salon Infosecurity)

Comment est-il diffusé ?

SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION

La DéDuplication de Données

Alexandre Delcayre

SNIA
Storage Networking Industry Association

Agenda

- Les besoins de « réduction »
- Définition et concept général
- Les différentes approches et méthodes
- Les configurations possibles
- Quelques avantages de cette technologie

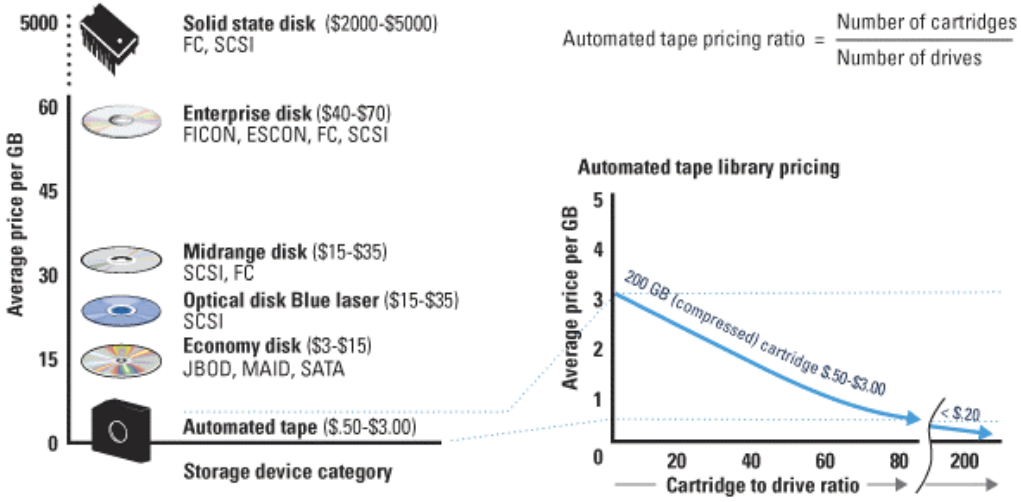
Les besoins de « réduction »

- Besoin #1 – Efficacité Technique
  - **Arrêter de stocker toujours les mêmes données** sur les unités de stockage générées par les procédures de sauvegarde ou le service fichiers
    - Occupation disque trop importante en volume et durée engendrant une explosion de l'espace de stockage secondaire: complexité de l'environnement de protection, fenêtre de sauvegarde trop longue (DPW)…
    - Croissance forte des volumes de données pour les bureaux distants

« Le parc installé de librairie de bandes stockent 4 à 15 fois le volume des baies de disques type Entreprise. »
Univ. Berkeley/Horison Information Strategies

Les besoins de « réduction »

- Besoin #2 – Efficacité Economique
  - **Obtenir un coût disque au TB inférieur au coût bande** pour dépasser l'intérêt technique de la solution (RPO, RTO, DPW petits) grâce aux différentes technologies disque de type ATA



Source: Horison Information Strategies (Juin 2005)

Les besoins de « réduction »

- Besoin #3 – Disaster Recovery
  - **Permettre un Recouvrement de Désastre** complet et rapide à coût raisonnable par élimination de l'occupation et l'utilisation réseau trop importante
    - Réplication possible sur réseaux à débits faibles et/ou longue distance

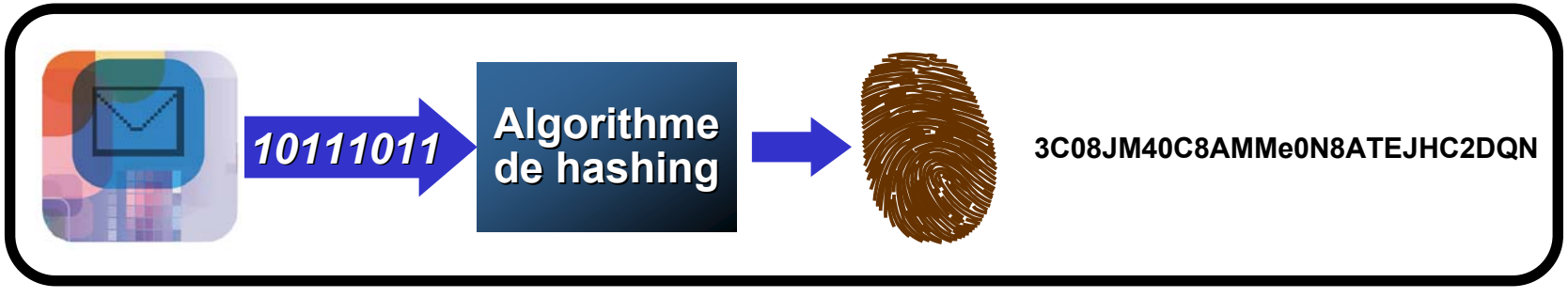
Le défi est donc de trouver une solution permettant de stocker un SEUL exemplaire d'une donnée pour permettre la MAITRISE du volume de données

Définition

- La DéDuplication de Données est une technique permettant de ne stocker qu'un **SEUL** exemplaire d'un jeu de données sur l'unité de stockage
  - Cette technique n'est pas une technique de compression de données, la compression pouvant être couplée à la DéDuplication
  - Les termes suivants sont synonymes dans la littérature de l'Industrie du Stockage:
    - Réduction, Factorisation (Commonality Factoring), Dé-Duplication, Optimisation de la Capacité (COS pour Capacity Optimized Storage), Elimination de la redondance, Coalescence de Données ou Instance de Données Unique (Single Instance Storage/Single Instance Repository)

Le Concept général

- Dé-doublonner les données entre elles grâce au calcul de leur signature établie par un algorithme de *hashing* et par la comparaison de ces signatures pour garantir l'unicité du jeu de données à stocker



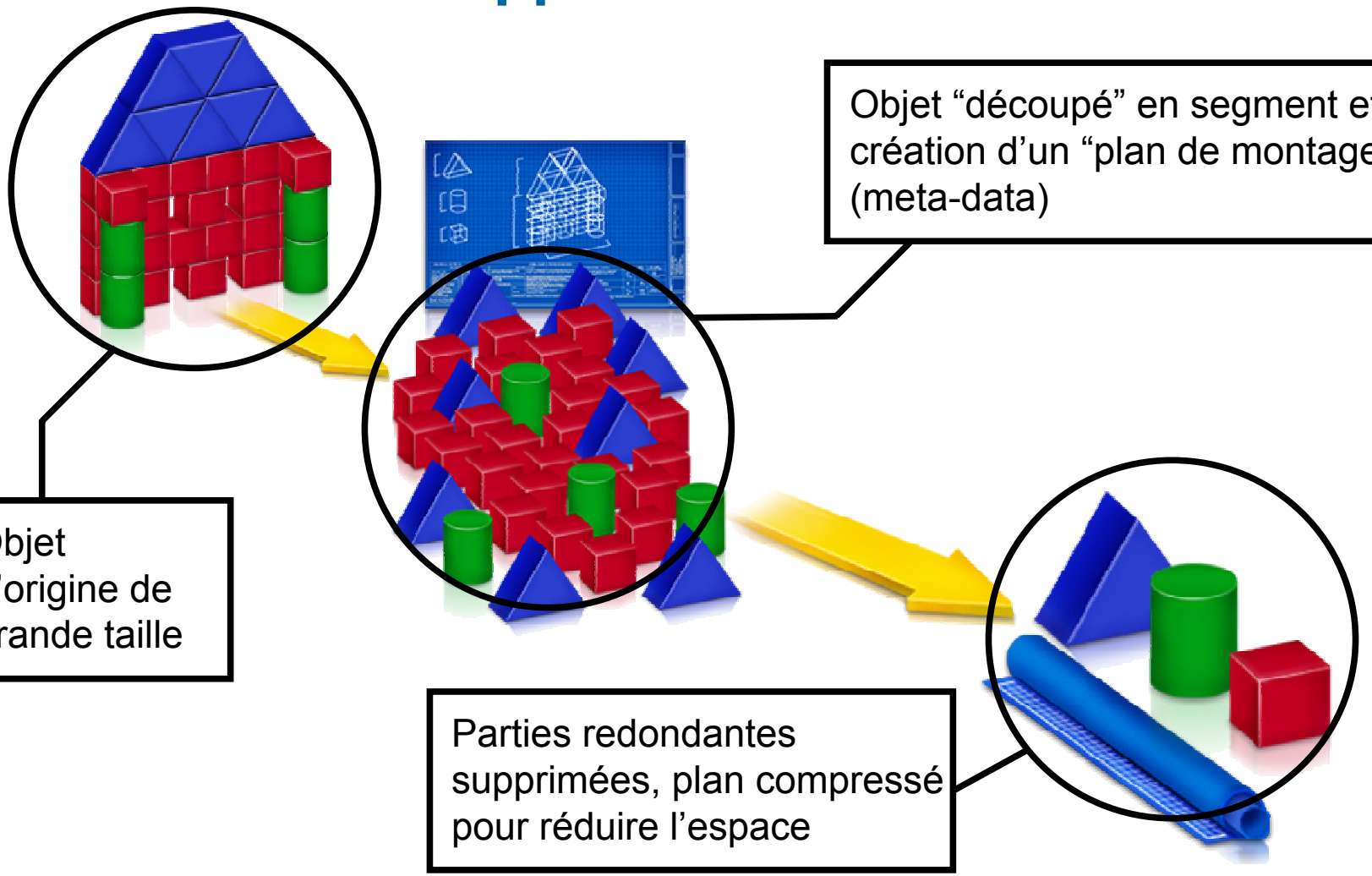
Les différentes approches et méthodes

- Calcul de l'unicité
  - Par plusieurs algorithmes de *hashing* (cryptographie)
    - MD5 (126 bits), SHA1 (160 bits)…SHA-512
    - Autres méthodes propres
  - Appliqués au fichier complet ou par découpage au segment de données (octet, bloc)
    - Longueur fixe ou variable

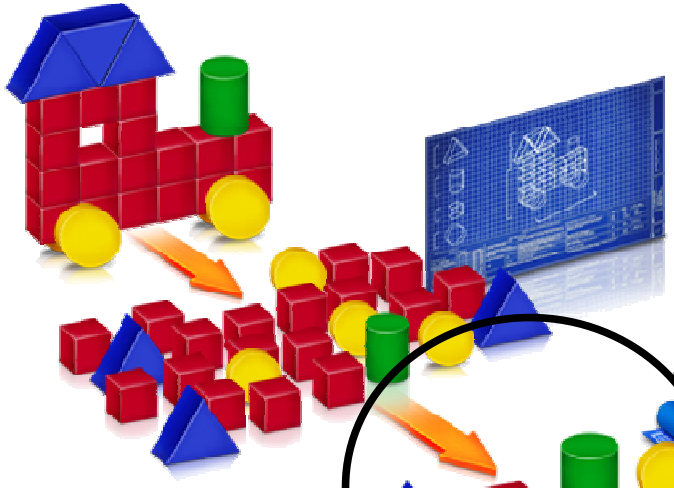
Les différentes approches et méthodes

- *hashing* puis comparaison des signatures et gestion de la « collision » statistique mais réelle
  - Ecriture du segment de données si sa signature est unique
  - Ajout d'un pointeur sur le segment s'il est déjà présent (signature identique) avec les alternatives suivantes
    - Alternative 1: Comparaison des jeux de données octet par octet
    - Alternative 2: Utilisation d'un second algorithme pour garantir l'unicité du segment (SHA-1 suivi de MD5…)
    - Alternative 3: Méthode propre

Les différentes approches et méthodes



Les différentes approches et méthodes



Objet “découpé” en 4 segments, comme en première phase.

Plusieurs objets peuvent être ainsi assemblés avec les mêmes 4 pièces de base tant que le plan d’assemblage est conservé.

3 des 4 pièces sont déja présentes (stockées) sur le support, seul l’ajout des nouvelles est nécessaire.

Les différentes approches et méthodes

- Notion de pointeurs (*reverse/forward*) et de référencement (direct/indirect)
- Traitement Synchrone ou Asynchrone
  - Synchrone ou In-line ou dans le flux de données
  - Asynchrone ou post-process ou à posteriori
- « A la source » ou « à la cible »
  - Source: Essentiellement présent sous la forme d'une solution logicielle qui réduit les données avant le transfert sur le réseau
  - Cible: Envoi complet du jeu de données vers une unité de DéDuplication
    - Unité disque cible intelligente (IDT – Intelligent Disk Target)
    - VTL (cas particulier d'une IDT)

Les différentes approches et méthodes

Synchrone

- DéDuplication à l'arrivée ou à l'envoi des données
- Peut utiliser 1 CPU par processus
- Peu de charge I/O (écriture seule)
- Doit être rapide pour ne pas impacter le processus de backup

Asynchrone

- Ecriture complète du backup sur disque, ensuite lecture et DéDuplication des données
- Peut utiliser plusieurs CPU par processus
- Charge I/O plus importante que mode Synchrone (écriture/lecture/écriture)
- Peut consommer du temps

Les différentes approches et méthodes

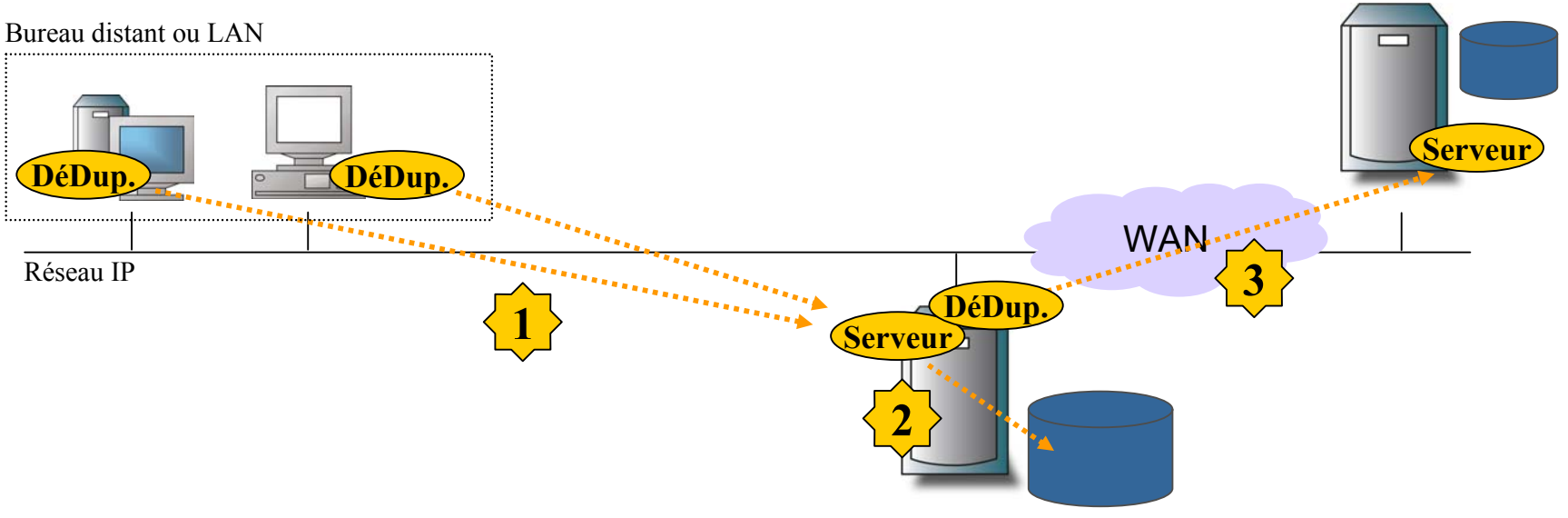
« A la source »

- DéDuplication du volume de données protégé
- DéDuplication par logiciel client
- Impact CPU sur client
- Nouvelle solution de backup
- Site distant sans nouveau matériel
- Protection de volume important avec impact sur le débit
- Très adapté au data center « moyen » et tout type de site distant

« A la cible »

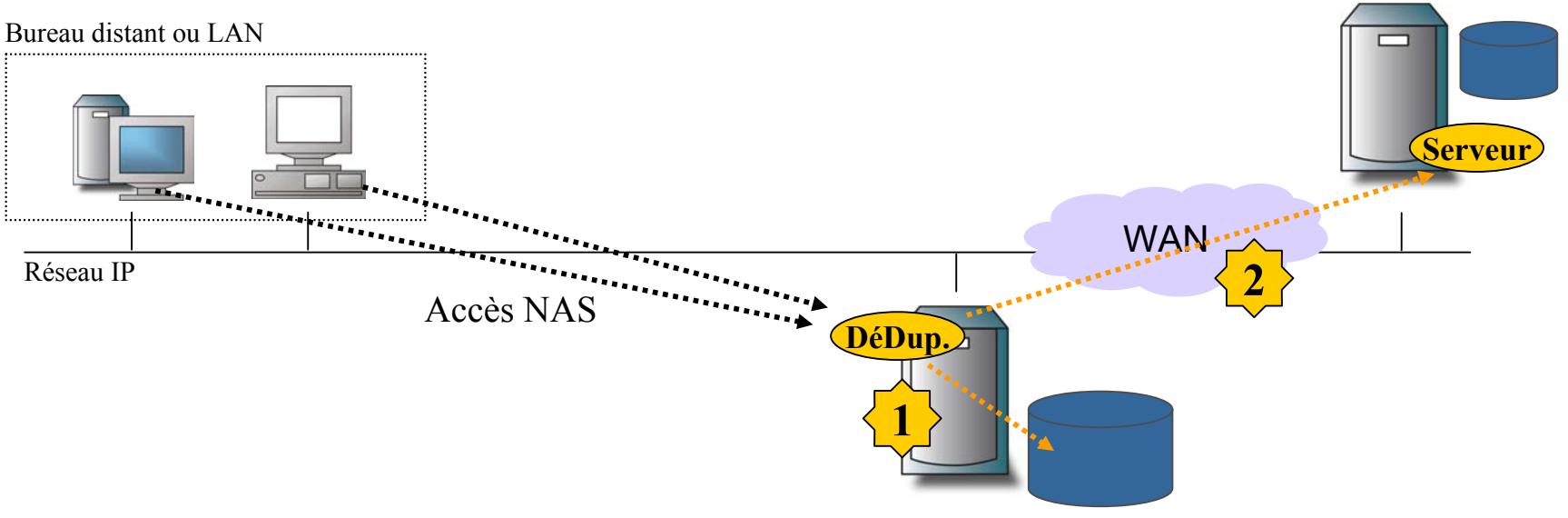
- DéDuplication du volume de données protégé
- Pas de DéDuplication avec logiciel client
- Pas d'impact CPU sur le client
- Solution de backup inchangé
- Site distant équipé
- Protection de volume important sans impact sur le débit
- Très adapté au data center de grand taille et site distant important

Exemples de configuration



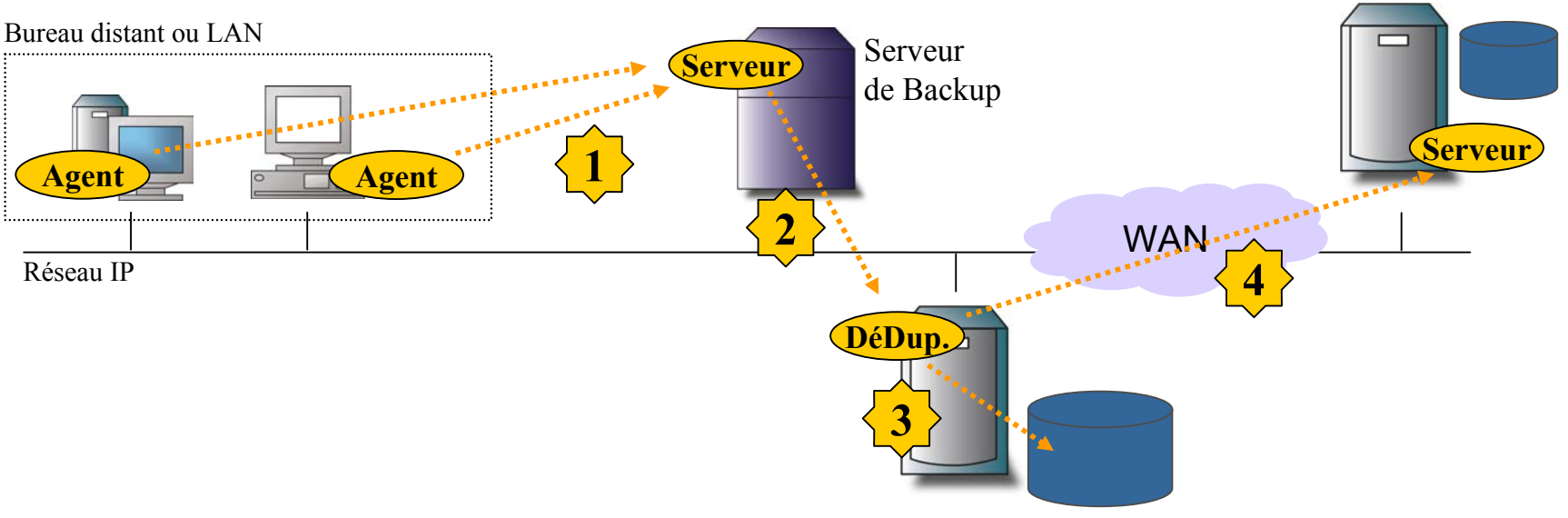
- Agent de DéDuplication installé sur les postes, transit des données uniques sur le réseau local ou distant
- Serveur stocke les données sur le stockage local, SAN…
- Réplication possible avec action de réduction entre les serveurs

Exemples de configuration (2)



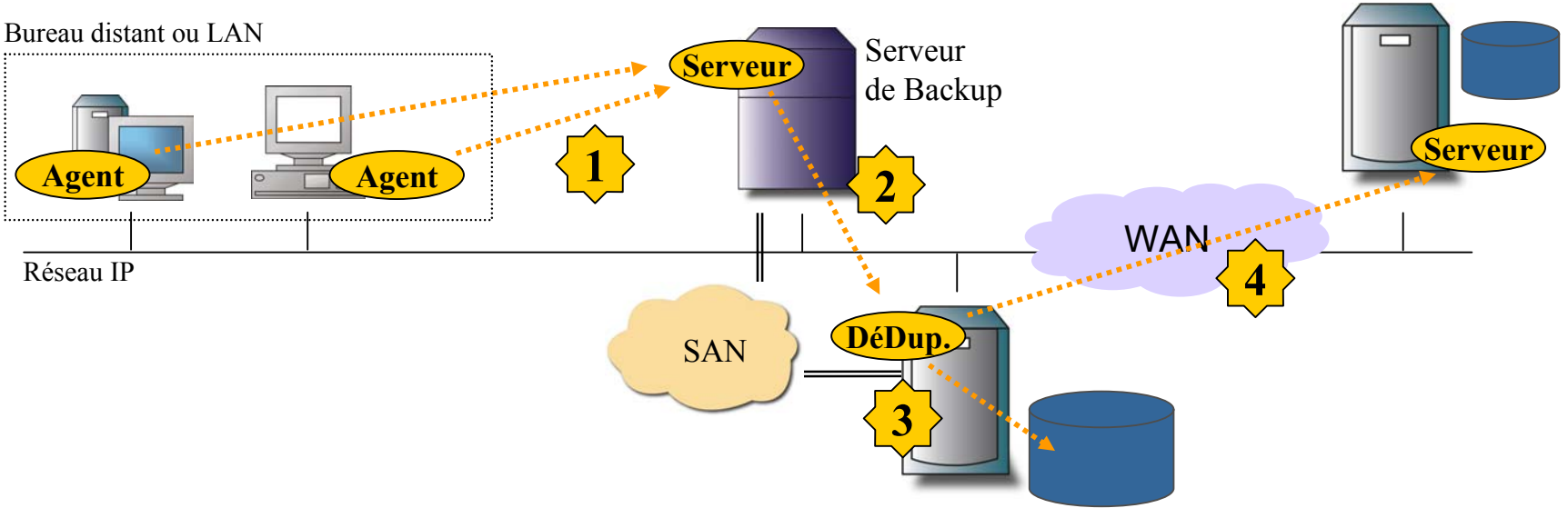
- Serveur dédié (serveur + logiciel « DéDup. » ou appliance) accédé par protocole NAS comme serveur de fichiers
- DéDuplication interne au serveur dédié
- Réplication possible avec action de réduction entre les serveurs

Exemples de configuration (3)



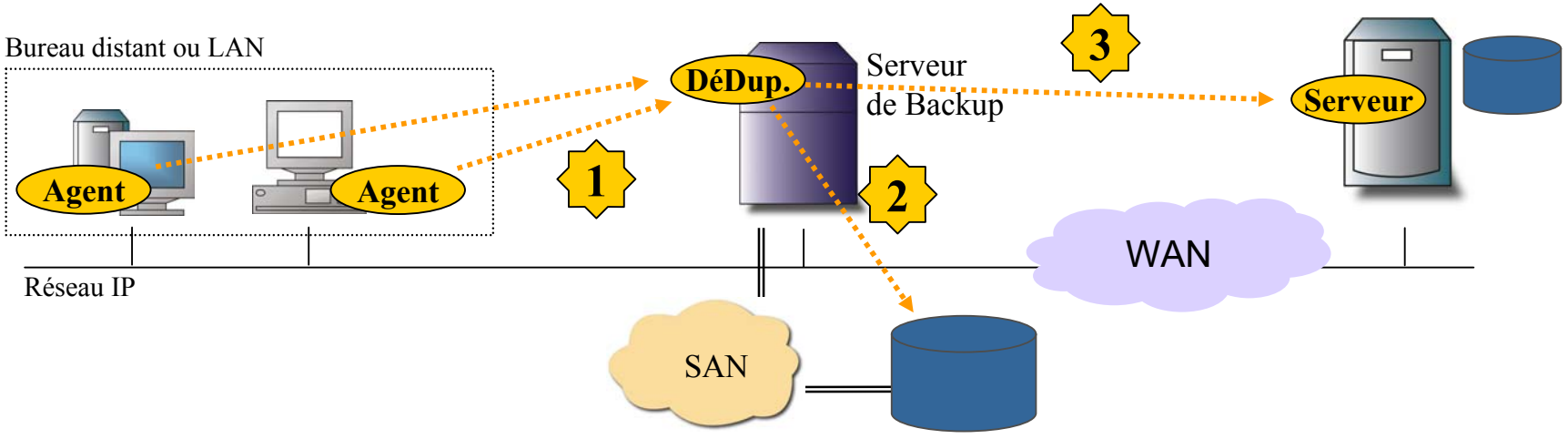
- Serveur dédié (serveur + logiciel « DéDup. » ou appliance) accédé par protocole NAS comme unité de stockage du serveur de backup
- DéDuplication interne au serveur dédié
- Réplication possible avec action de réduction entre les serveurs

Exemples de configuration (4)



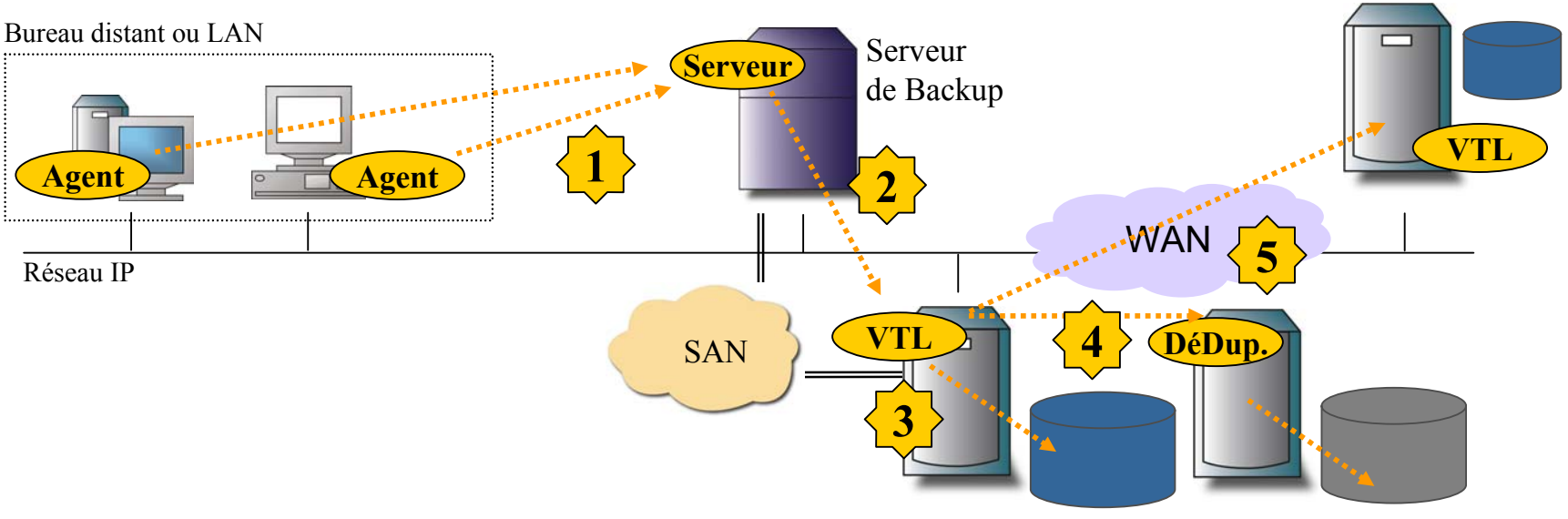
- Serveur dédié (serveur + logiciel « DéDup. » ou appliance) accédé en mode bloc (FC, iSCSI) comme unité de stockage du serveur de backup
- DéDuplication interne au serveur dédié
- Réplication possible avec action de réduction entre les serveurs

Exemples de configuration (5)



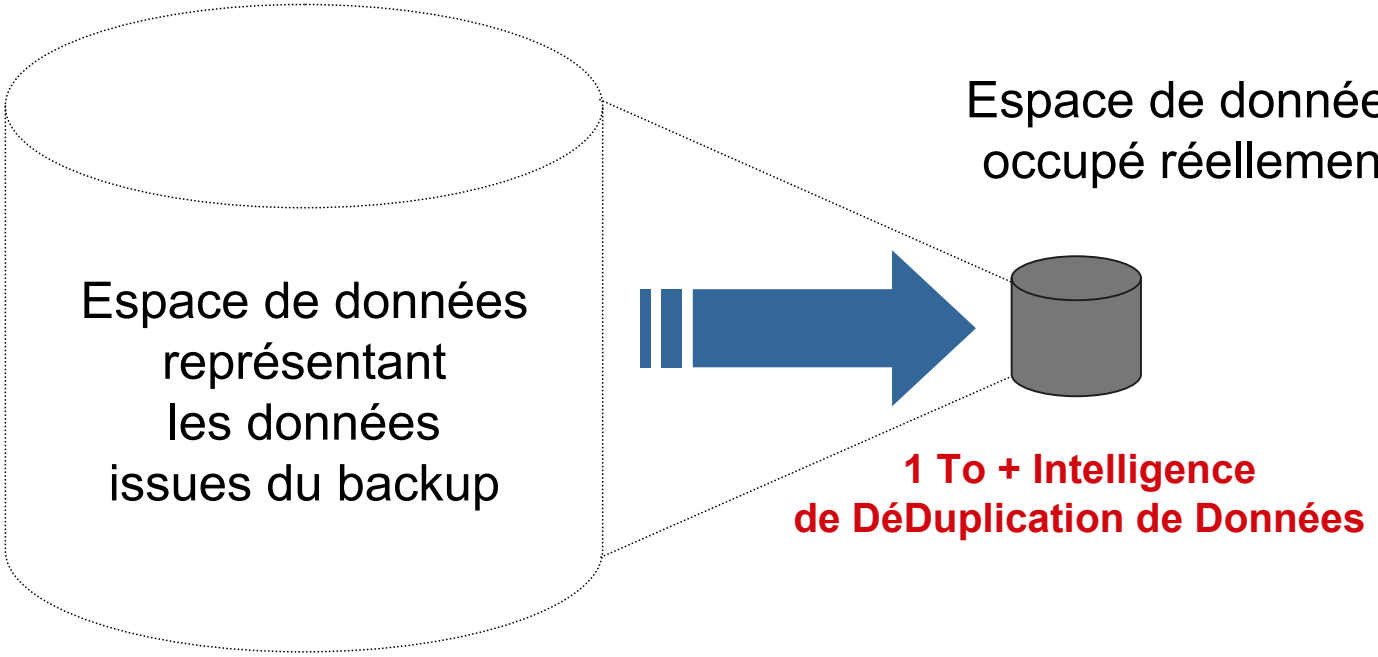
- Serveur de backup intégrant la fonction de DéDuplication
- Réplication possible avec action de réduction entre les serveurs

Exemples de configuration (6)



- Serveur dédié (appliance) accédé en mode bloc (FC, iSCSI) comme unité de stockage dédiée pour la DéDuplication du VTL
- DéDuplication interne au serveur dédié de concert avec le VTL
- DéDuplication aussi possible sur le VTL lui-même (même appliance), à la volée (efficacité moindre)
- Réplication possible avec action de réduction entre les VTL

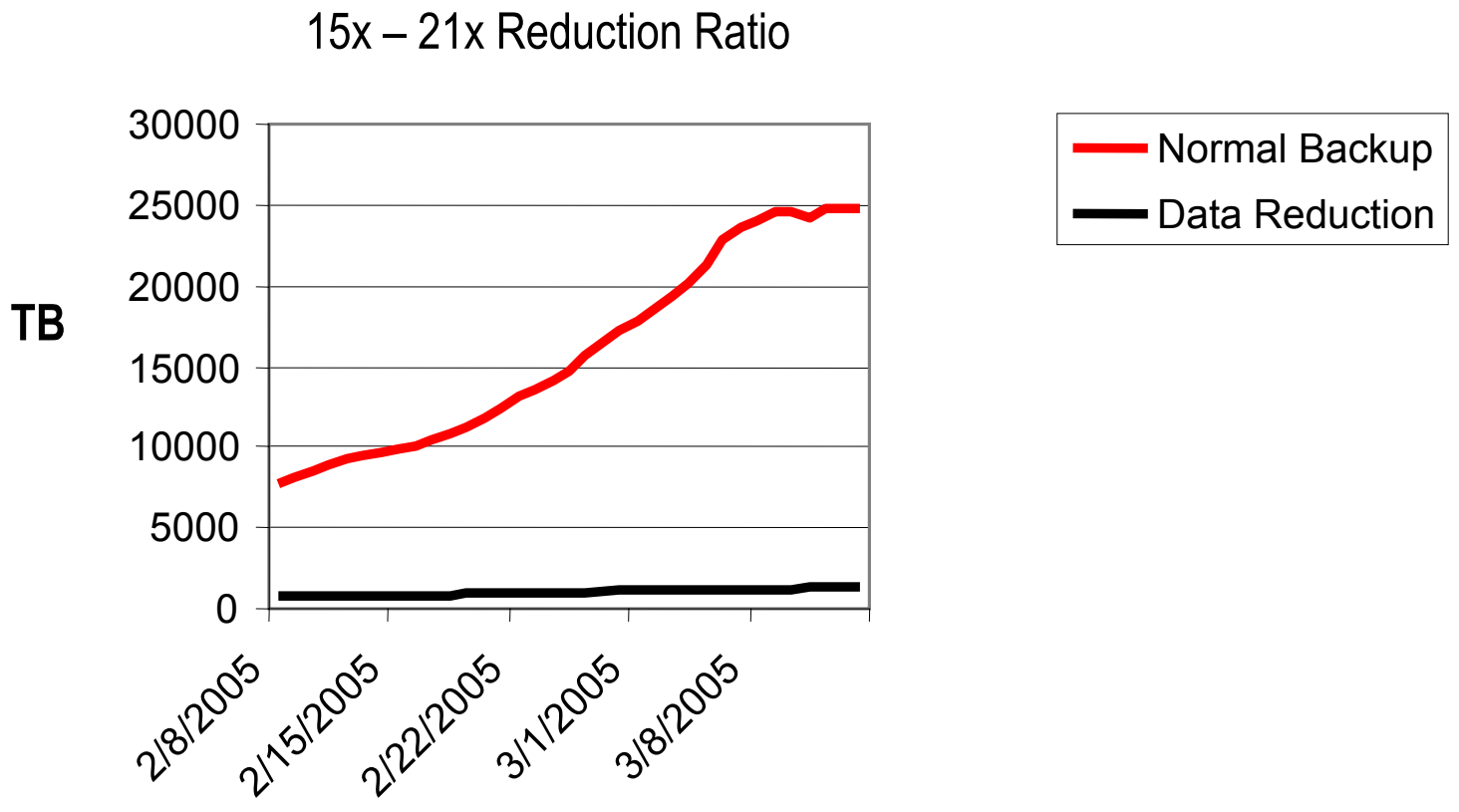
Mesure de l'économie



Baie de disque RAID SATA…
20 To utiles RAID 5/6/DP

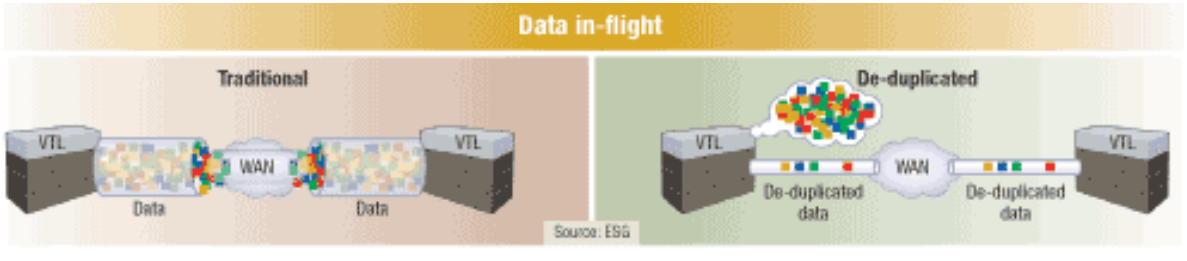
Rapport de 20 à 25 sur la capacité de stockage pour une utilisation en backup

Retour d'expérience – Scénario de Backup



Quelques bénéfices au-delà du TCO Stockage

- Rétention plus longue
  - Réduction de la « perte » d'informations, disponibilité plus longue
- Sauvegarde quotidienne
  - Amélioration du RTO
- Disaster Recovery par réplication « fine »
  - Meilleure qualité du PRA



SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION

Protection des Postes et Bureaux Distants

Miguel Dos Santos Lopes

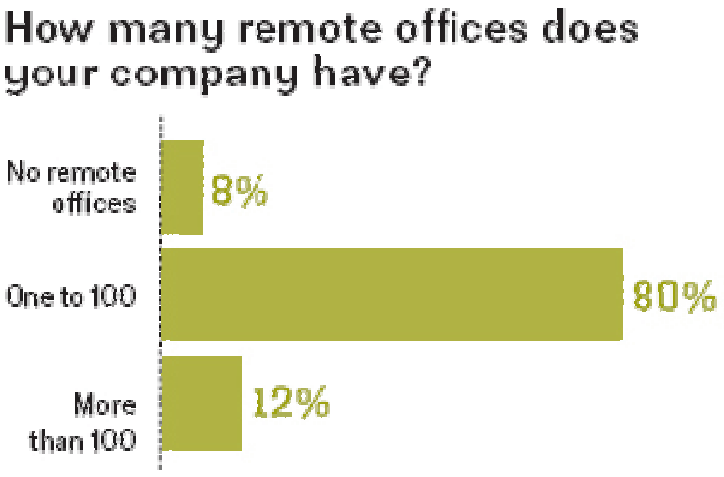
Constat et Chiffres….

- 87% des employés travaillent « détachés » de la maison mère (Nemertes Research)
- Plus de 75% des données résident en dehors du Datacenter (ESG)
- 70% du capital intellectuel des entreprises est stocké sur des PC et des portables

- Une quantité importante de données distribuées dans les bureaux distants ne sont jamais protégées
  - 90% des données stockées sur des PCs et ordinateurs portables ne sont jamais sauvegardées
  - 52% des managers les estiment « très vulnérables »
- Seulement 30 à 40% des backups distants sont efficaces et se terminent bien (ESG)

- 18 milliards de $ perdus par an

Vers une démarche nouvelle…



Top 5 problems users have with remote-office backups

1. Remote site backups to tape are costly and/or unreliable
2. Remote site backups are handled by inexperienced staff and/or done irregularly
3. Sending data across the WAN creates intolerable packet loss/throughput levels
4. The added backup data from remote sites makes central backups more difficult
5. Remote site backup tapes aren't sent offsite

Source: Enquête Storage Magazine Avril 2007

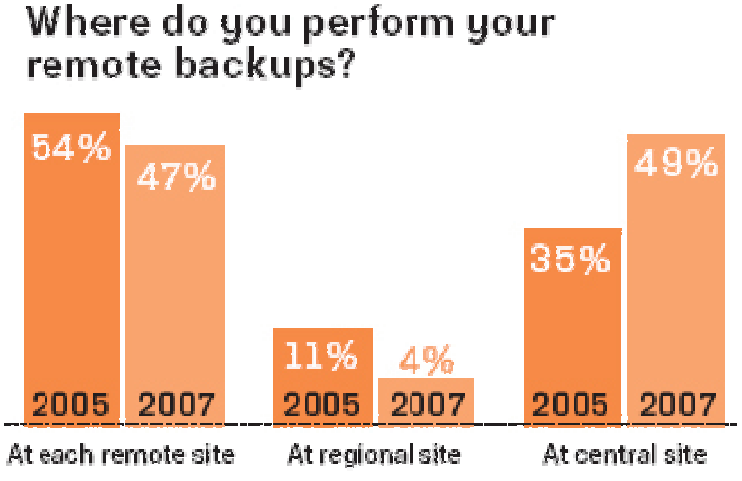
Vers une démarche nouvelle…

62%

Users who plan to centralize backups

38%

Users who don't plan to centralize backups



When will you complete your centralization project?

43%
End of 2007

14%
End of 2008

43%
Unknown

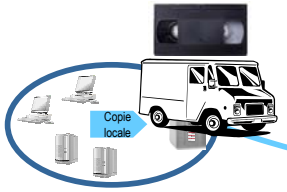
Source: Enquête Storage Magazine Avril 2007

Approche traditionnelle… aujourd'hui limitée

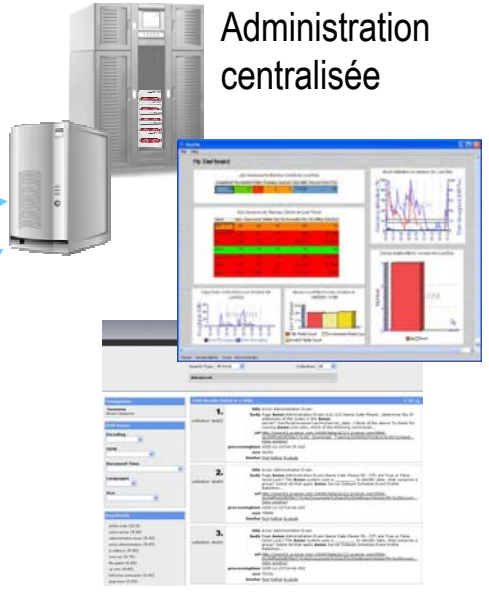
- Sauvegarde locale au site avec administration centralisée
  - Stockage local type librairie ou même disque
  - Externalisation des médias vers un site tierce ou central
  - Coût important
  - DR limité
  - RTO et RPO non optimal
- 69% des sociétés équipées d'une solution de backup distant ne sont pas satisfaites

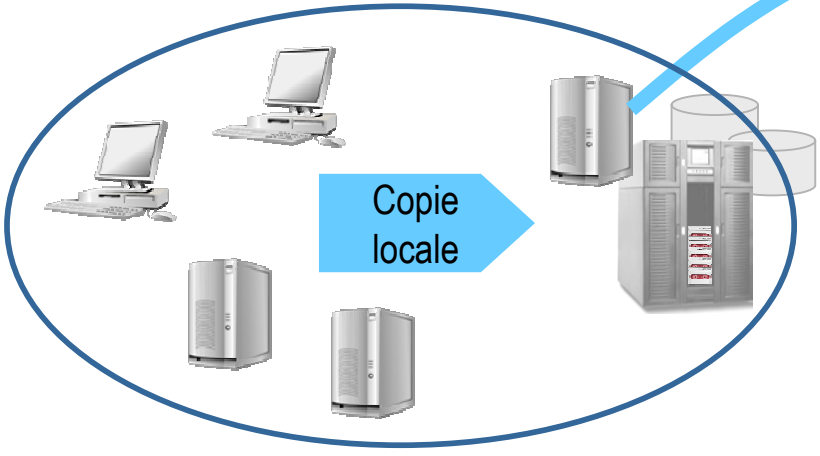
Approche traditionnelle… aujourd'hui limitée

Sauvegarde locale sur bandes
Externalisation des médias



Administration
centralisée





Sauvegarde locale sur disques ou bandes
Réplication « standard » éventuelle

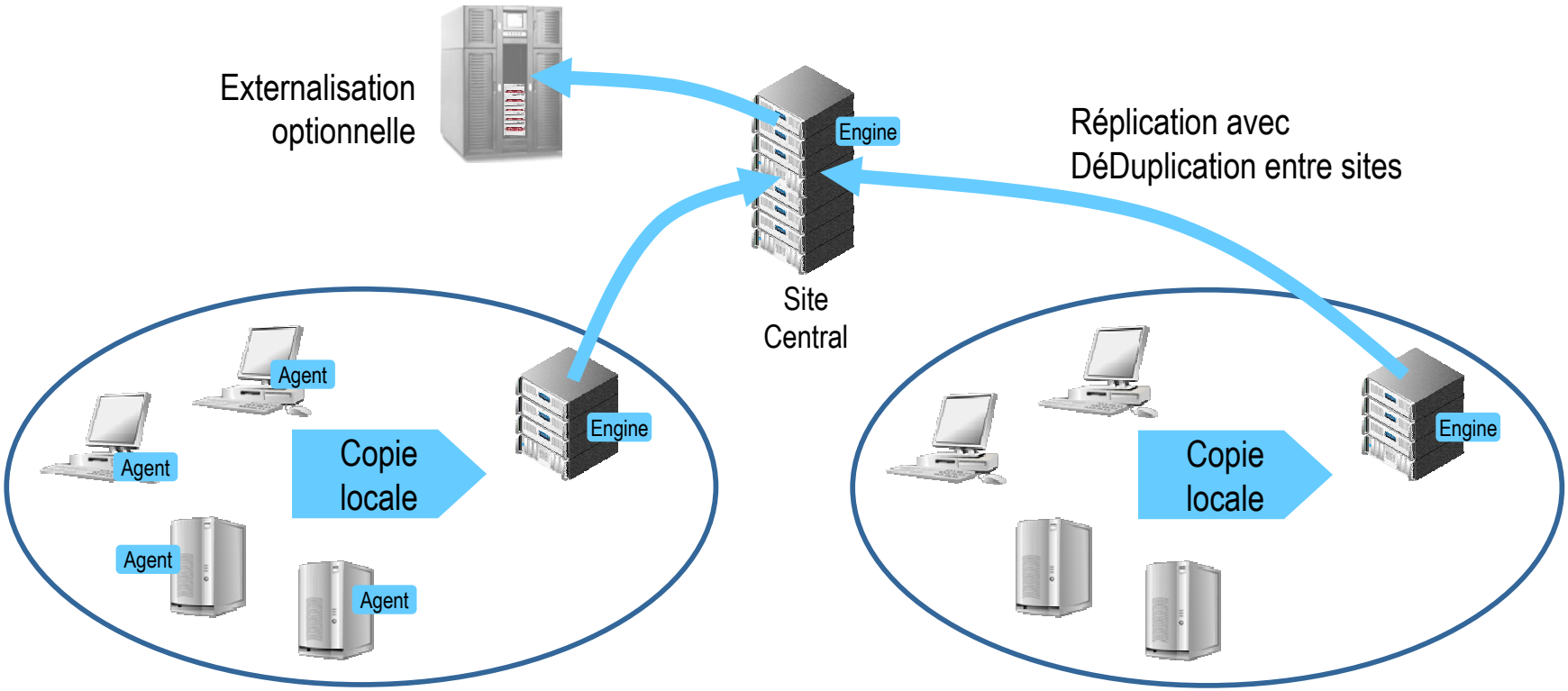
Sauvegarde distante
directe sur site central

Intégration des technologies d'aujourd'hui

- Disque SATA et dérivés
  - RAID 6/DP….
- CDP
- DéDuplication « globale » (entre les sites et sessions…)
  - À la source ou à la cible sur site et vers le site DR
  - Compression éventuelle suite à la DéDuplication
- Chiffrement
- Réplication entre les sites
  - Cascade et mode DR
- Intégration potentielle de VTL sur le site central/distant
- Indexation des Meta-Data et potentiellement du Contenu
- Restauration directe par les utilisateurs grâce à l'accès direct aux données sauvegardées (Virtual File System basé sur CIFS)

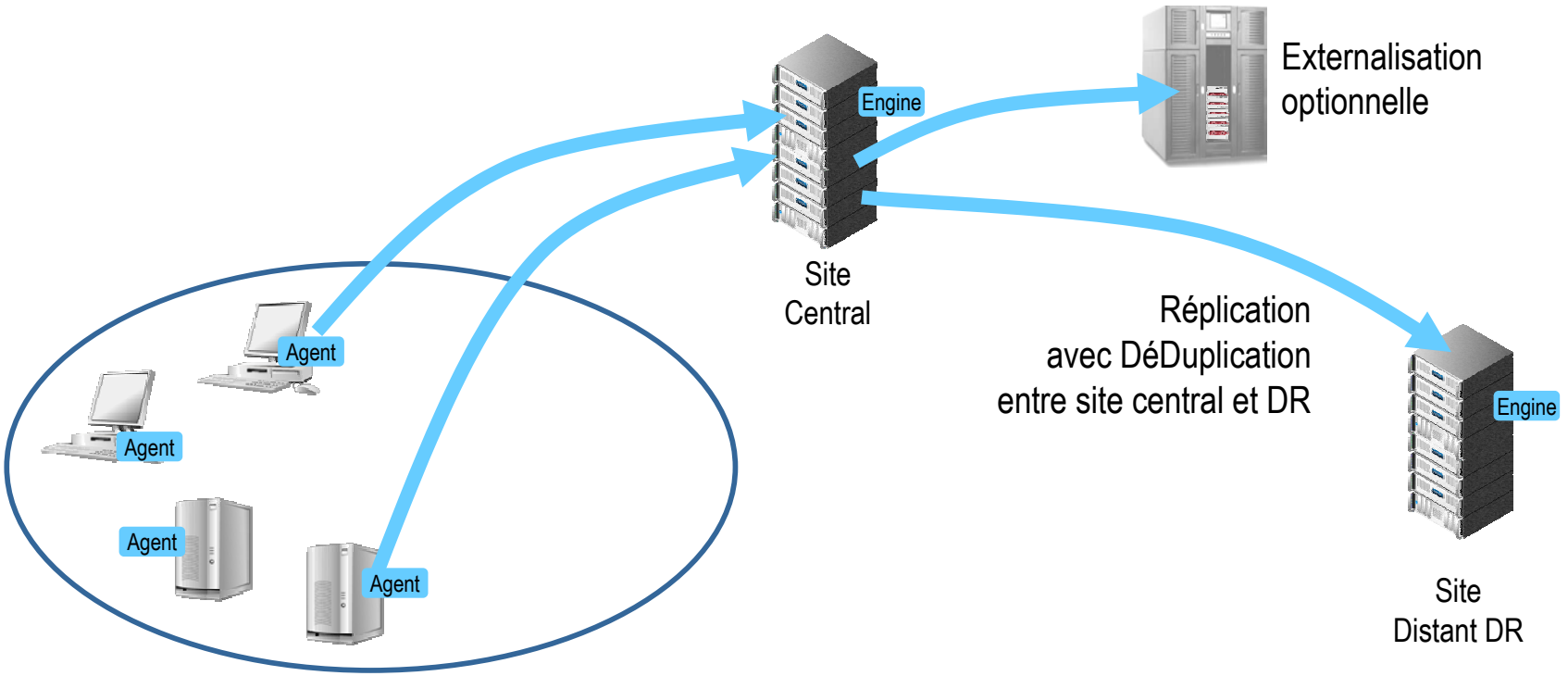
Sauvegarde locale évoluée

- DéDuplication à la source
- DéDuplication à la cible



Sauvegarde distante évoluée

- DéDuplication à la source
- Réplication sur site DR



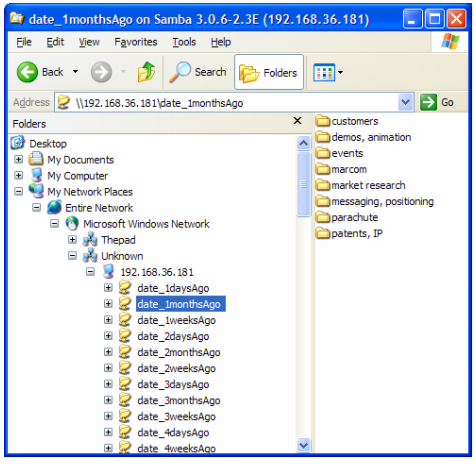
SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION

Sécurité des données

- Chiffrement des données à la source
  - Si agent présent sinon sur serveur
- Transport des données chiffrées
- Stockage des données chiffrées

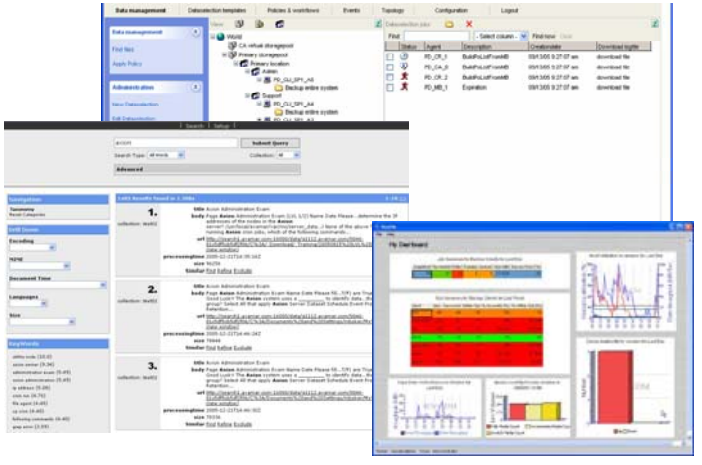
Accès utilisateur / Restauration

- Accès en read-only au share des images – notion de disque virtuel par interface CIFS
  - « Montage » ou accès UNC (« \\.... »)
- Accès en read-only via WebDAV (http) et lecteur local virtuel
- Accès par Browser
  - Pas d'agent nécessaire pour la restauration
  - Téléchargement http
- Navigation dans l'arborescence
  - « Drag and Restore »



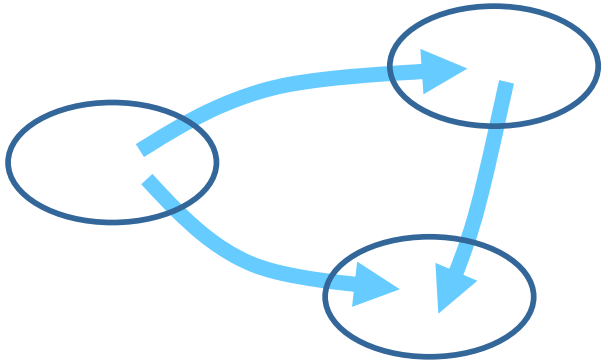
Administration

- Par Gui spécialisé (Java/Windows…)
- Par interface Web sécurisé

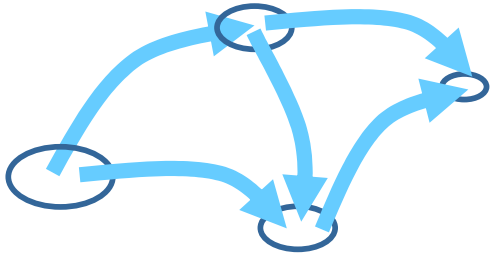


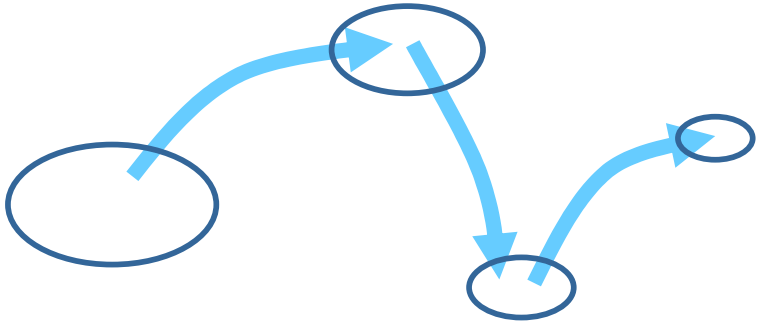
- Sécurisé par https, SSL…
- Notion de rôle, groupe…

Autres Alternatives



- Grid Storage répartie
- Peer-to-Peer
- Cellulaire
- RAIN* distribué
- ASP et récemment SaaS (Software as a Service) dédié au Stockage/Backup…
- Et Wafs (Wide Area File Services)

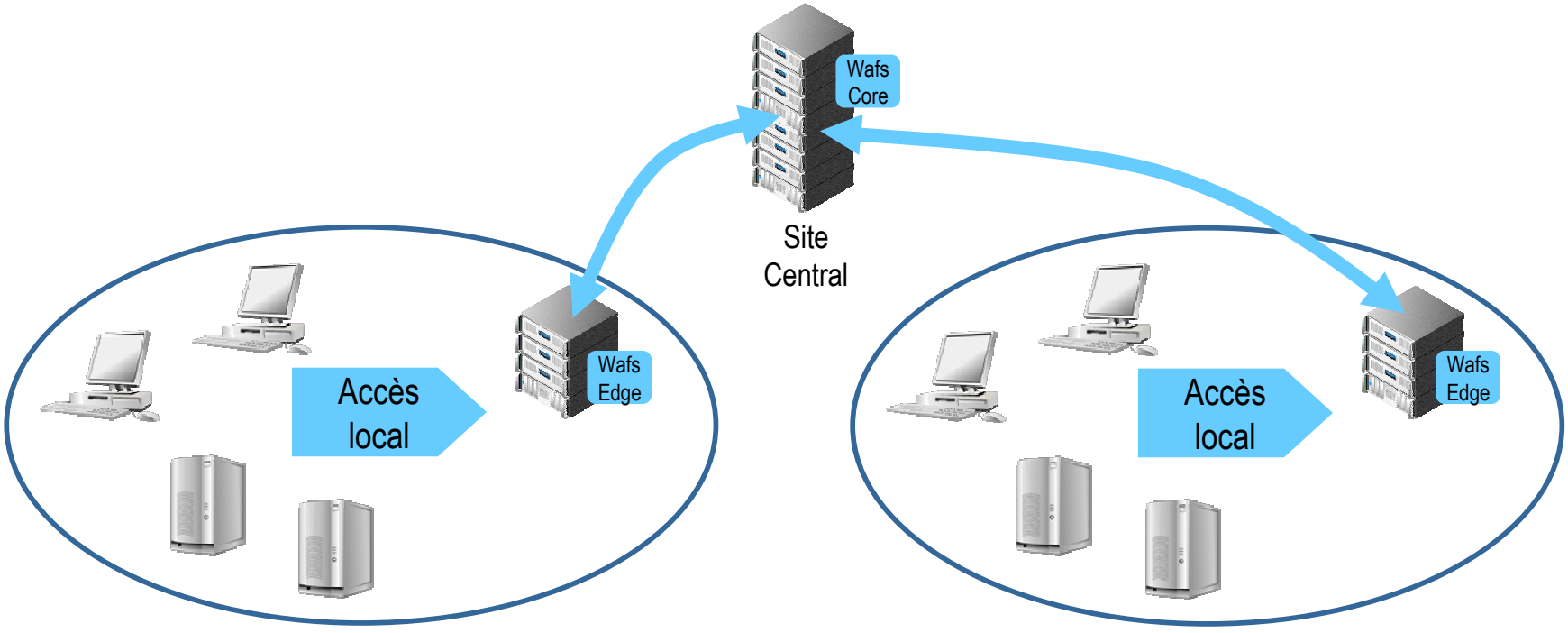




RAIN: Reliable|Redundant|Random Array of Inexpensive|Independant Nodes

Pas de sauvegarde mais accès distant

- Serveur de fichiers centralisé à accès local en Wafs
  - Sauvegarde centrale du serveur



Conclusion

- Quelques éléments clés à considérer
  - Protection orientée disque
  - Réduction de Données (DéDuplication) semble indispensable
  - Facilité de Sauvegarde/Restauration
  - Chiffrement
  - Réplication
  - Administration centralisée
  - Evolutivité importante
  - Intégration éventuelle à une solution Data Center

SNIA
STORAGE NETWORKING INDUSTRY ASSOCIATION

Conclusion

Conclusion

- Réalité du Backup sur Disque
- **Next Generation Data Protection (NGDP)**
  Révolution de la Protection de Données avec plus de 80 produits ou sociétés actives en VTL, CDP, Réduction de Données, Chiffrement, Backup distant…
  - Importance de l'agrégation fonctionnelle
- La SNIA (DMF-DPI) au cœur de la standardisation (www.snia.org/dmf)

Remerciements

TARSUS FRANCE

REFLEXION Publique
COMVITAMINE
Relations Presse & Événements

et tous les membres du comité régional SNIA France

Storage Networking Industry Association

SNIA EUROPE