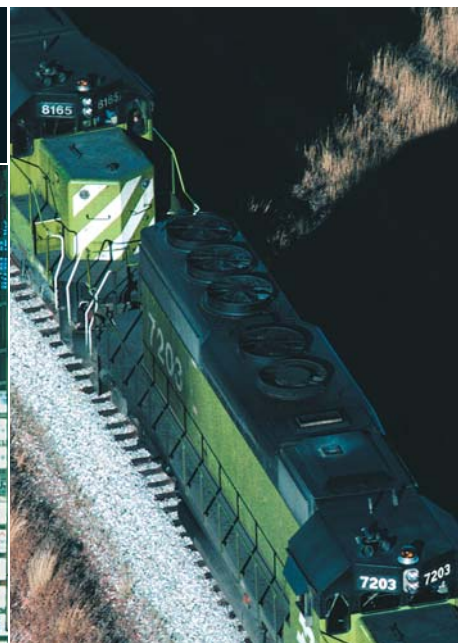


Développement Offshore

PAGE 6



Microsoft BizTalk Server 2004

PAGE 28

Stockage : le temps de la consolidation

PAGE 45



Les plateformes de travail collaboratif

PAGE 19



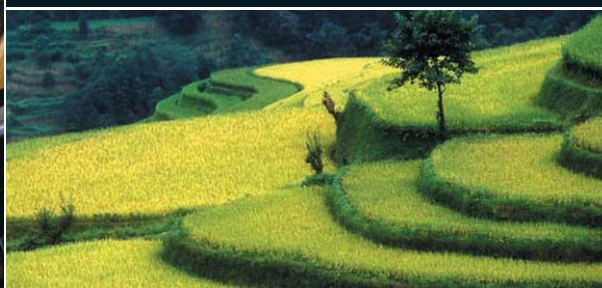
IT-expert n° 48 - Mars/Avril 2004

L 18746 - 48 - F: 9,00 € - RD



Wi-Fi

PAGE 34



Sommaire



6 Dossier

Développement Offshore

Les développements offshore connaissent un intérêt croissant de la part des entreprises qui y voient un moyen de réduire leur coût. De nombreux pays sont aujourd'hui sur ce marché avec des atouts propres à chacun. Si la comparaison des coûts salariaux est évidente, il convient de ne pas oublier les aspects méthodologiques et humains qui doivent sous-tendre le choix d'une telle démarche, faute de quoi cette dernière peu s'avérer catastrophique.



19 Technique

Les plateformes de travail collaboratif

Les plates-formes de travail collaboratif constituent des outils puissants qui contribuent à la réduction des coûts et à un meilleur partage de l'information entre les différents acteurs d'un projet. Cet article présente un panorama de ce qu'il convient d'attendre de ce type d'outils en plein essor.



34

Comment ça marche Wi-Fi ?

28 Quoi de neuf Docteur ?

Microsoft BizTalk Server 2004

La nouvelle version du logiciel d'intégration d'applications d'entreprise de Microsoft a fait l'objet d'un lancement mondial le mois dernier. Nous faisons le point dans cet article sur les nouvelles fonctionnalités du produit qui ciblent aussi bien les développeurs et administrateurs que les analystes.

34 Comment ça marche ?

Wi-Fi

Le Wi-Fi promet l'internet sans fil. Ce qui n'était au début qu'une facilité pour quelques informaticiens désireux de surfer depuis leur jardin ou leur baignoire entre de plein pied dans les entreprises et allèche les opérateurs. Petit descriptif du mode de fonctionnement de cette technologie.

41 Fenêtre sur cour

Interview UNILOG

Monsieur Xhaard, Directeur Général Adjoint d'Unilog France.

Unilog, acteur européen majeur de l'ingénierie, du conseil et de la formation voit dans l'outsourcing et l'offshoring deux axes importants de son développement comme nous l'explique son Directeur Général Adjoint.

45 Stockage : le temps de la consolidation



45 Rubrique à bras

Stockage : le temps de la consolidation

Les dépenses d'infrastructure stagnent, tandis que celles pour la gestion du stockage, du cycle de vie des données et de la virtualisation explosent. Voici donc venu le temps de la consolidation et de l'optimisation de l'existant.

50 Livres

Maximizing Performance and Scalability with IBM Websphere et **Pro Jakarta Struts** sont les deux ouvrages que nous vous recommandons ce mois-ci.



Stockage : le temps de la consolidation



Les réseaux de stockage, soumis à de fortes contraintes de haute disponibilité, de performances et de retour sur investissement sont particulièrement visés par les procédures de maîtrise des coûts de stockage, éventuellement leur réduction, et l'établissement d'une véritable qualité de service. D'où une transformation annoncée (par les prédictions de Gartner DataQuest et d'IDC) de la structure des investissements dans le stockage. Gartner Dataquest prévoit par exemple que les dépenses mondiales consacrées au SRM (Storage Ressource Management) en 2005 seront de 577M\$, alors que les dépenses de base d'infrastructures (baies, commutateurs, etc .) stagneront à 691M\$ (*). Les évolutions clés prévues par les instituts de conjecture concernent la convergence du stockage vers IP, la gestion des ressources de stockage (SRM), la virtualisation et la gestion du cycle de vie des données (ILM - Information Life Management).

* Source Gartner Dataquest - avril 2003



Convergences SAN-NAS et coexistence Fiber Channel IP

Les équipements FC (Fiber Channel) d'un SAN représentent une part considérable du coût total d'une infrastructure de stockage. Par ailleurs, toutes les études montrent que leur coût d'administration dépasse celui de l'acquisition des matériels. Dans la mesure où certaines convergences SAN-NAS, jointes à des solutions mixtes IP-FC existent sur le marché, il est donc légitime de s'interroger sur la pertinence du tout FC dans les réseaux de stockage, voire même de sa simple présence. Outre leurs coûts d'acquisition élevés et leurs nombreux problèmes de compatibilité, les équipements de la Fabric devront, en effet, être gérés par un personnel d'une qualification différente de celle nécessaire à l'administration des réseaux IP.

Il est indispensable de distinguer deux problématiques avant de faire valoir les arguments des deux parties :

- **SAN et NAS :** Un grand nombre d'entreprises font coexister des NAS (Networked Attached Storage) avec des SAN. Les NAS sont des serveurs de fichiers (d'où le nom « filer »), basés sur NFS et CIFS. Ces deux protocoles s'appuient eux-mêmes sur TCP/IP. Les NAS présentent donc l'avantage d'un routage simple, leur administration est généralement relativement basique, les possibilités de « tuning » sont assez limitées, et le partage de fichiers entre plusieurs hôtes est par nature disponible. A contrario, les SAN travaillent en mode block entre les serveurs et les baies de stockage. La représentation des blocks en fichiers est réalisée par le serveur lui-même. Bien que des éditeurs proposent des systèmes de fichiers qui permettent à plusieurs hôtes de visualiser le même fichier à travers un SAN, comme ADIC avec StorNext File System, le partage de fichier sur un SAN est encore relativement confidentiel. En d'autres termes, sur un plan conceptuel, les LUNs proposés par une Fabric à un hôte ne sont pas par nature différents de ceux vus à travers un gestionnaire de volumes sur une baie connectée en attachement direct. Enfin, les SAN sont en grande majorité au moins en partie basés sur le protocole FC.
- **SAN sur IP :** à la différence de la représentation SAN/NAS, le débat FC-IP ne repose que sur des questions protocolaires : vu d'un système d'exploitation, un SAN basé sur IP ne diffère du SAN FC que par les interfaces par lesquelles transitent les blocks de données : HBA (Host Bus Adapter) pour FC, NIC (Network Interface Card) pour IP.

La coexistence entre les deux types de réseau est avant tout liée à des raisons historiques. FC a initialement été développé afin de passer outre aux limitations d'IP et du SCSI. Avant l'avènement du Gigabit Ethernet, les débits des premiers réseaux FC (100 Mo/s) étaient en effet nettement supérieurs à ceux permis par Ethernet. Autre argument fréquemment avancé : les sauvegardes peuvent saturer le LAN de production, tandis que le réseau FC dispose d'un accès direct aux périphériques de sauvegarde. Un VLAN Gigabit Ethernet dédié aux sauvegardes permet aujourd'hui de contourner ce problème. Les principales limitations de SCSI n'ont pas disparu : la taille maximale des câbles est de 18 mètres, contre 10 km pour les câbles FC,

et le nombre d'adresses disponibles sur un contrôleur est beaucoup plus faible. Les avantages historiques du protocole FC sur IP ont donc tendance à disparaître. La mainmise d'IP sur les échanges réseau et l'avènement d'Internet donnent à penser que la R&D consacrée à IP est très largement supérieure aux efforts de développement consacrés au Fiber Channel (même si le dépassement de la limite actuelle de FC de 2Gb/s reste un sujet peu important pour les performances du stockage – les baies ont des points de contention en amont de cette limite -, le 10GbE est une réalité en laboratoire. Les fournisseurs de matériels FC n'annoncent rien de comparable).

Quelques points techniques méritent d'être relevés :

- Performances comparées des SAN et des NAS : bien que le débat FC/IP ne permette pas, à priori, de départager les protagonistes sur le plan protocolaire, le SAN (FC ou IP) conserve une longueur d'avance en termes de performances pour les applications de milieu et de haut de gamme. Il reste que les exigences techniques en termes de performances peuvent considérablement varier d'une application à l'autre (dominante lectures/écritures, débit par transactions, débit total, etc.) et que les caches des baies de stockage sont bien plus importants en termes d'écritures de logs de bases de données (les transactions des utilisateurs sont synchrones de ces écritures) que le mode d'accès (NFS ou mode block). En résumé, les offres haut de gamme NAS peuvent parfois concurrencer les performances des SANs pour certaines applications. Des tests sont présentés par tous les acteurs du stockage impliqués dans ces débats visant à démontrer la supériorité de l'un ou de l'autre. Ces tests sont en général assez éloignés des applications clientes. Il est donc préférable de cerner les besoins des applications (débits, temps de réponses unitaires...) et de tester par soi-même.
- La stack TCP/IP est gérée sur les NIC par les CPUs du serveur. Le remplacement d'un lien HBA par des NICs passe par le déport du traitement de la stack sur des composants dédiés.
- La sécurité des réseaux FC n'est pas structurellement meilleure que celle des réseaux IP, mais les premiers sont par nature isolés, non connectés sur le monde extérieur. Ceci étant, de nombreux NAS hébergeant des données sensibles sont protégés par l'arsenal des logiciels de sécurité des réseaux.
- Les technologies IP appliquées au stockage sont encore relativement jeunes. Les produits proposés par les constructeurs disposent en général d'une base installée peu importante.

Plusieurs protocoles basés sur IP sont disponibles sur le marché :

- **iSCSI** – Protocole présenté par l'IETF (Internet Engineering Task Force), auteur de nombreuses normes reconnues par les fabricants et par le consortium des constructeurs et des éditeurs du stockage SNIA (Storage Networking Industry Association), qui a fusionné avec la FCIA (Fiber Channel Industry Association). iSCSI est une alternative IP à FC. Ce protocole ne s'appuie pas sur FC : il supporte IP sur un sous-jacent SCSI.
- **FCIP** – Encapsulation de FC dans IP. Ce protocole est



aujourd'hui utilisé pour interconnecter sur de grandes distances des îlots SANs. Dans la mesure où il s'agit d'une encapsulation, (« des trames dans les trames »), ce protocole présente un intérêt limité dans les SANs localisés sur un périmètre géographique limité.

- **iFCP** – Ce protocole s'apparente dans son utilisation à celui de FCIP. Mais au lieu de créer un « tunnel », iFCP présente les matériels FC comme des nœuds IP, afin d'interconnecter des SANs distants. iFCP demande des routeurs spécifiques (gateways).
- **SOIP** (Storage over IP) propose sur le même matériel iFCP et iSCSI.

D'autres protocoles basés sur IP, souvent développés par de petits acteurs, émergent également. Enfin, Infiniband, une technologie non IP à l'origine développée pour les bus internes des ordinateurs Intel rencontre des difficultés d'acceptation par le marché.

Virtualisation

Il n'est pas simple de se retrouver dans le discours marketing des différents protagonistes sur ce marché. Prise dans son sens le plus large, la virtualisation appliquée au stockage est l'ensemble des techniques qui permettent de présenter à des hôtes des espaces de stockage, sans que ces derniers ne se préoccupent de la nature ou de l'organisation de ces espaces.

On trouve en réalité ici trois concepts :

- La virtualisation permet tout d'abord de gérer de la capacité de stockage sous la forme de disques peu performants et d'un faible coût (ATA ou Serial-ATA) ou de bandes. Il s'agit ici de placer des données sur ces périphériques et de permettre aux ordinateurs d'y accéder de façon relativement transparente (temps d'accès dégradé). Ces supports peuvent être utilisés sans nécessairement impliquer la notion de virtualisation : c'est l'agrégation de ces zones de stockage par un HSM (Hierarchical System Management) qui permet de déplacer les fichiers d'un support à l'autre en masquant à l'utilisateur le processus.
- La vision traditionnelle correspond peu ou prou à celles de la gestion des volumes, c'est-à-dire un agrégat de partitions de disques physiques, partitions présentées comme un disque au système d'exploitation. Les vendeurs de baies de stockage implémentent la virtualisation au niveau block (dans la baie elle-même). Des sociétés telles que FalconStor, DataCore ou StoraAge proposent de virtualiser les espaces de stockage sous la forme d'Appliances au niveau de la Fabric elle-même, en mettant l'accent sur l'hétérogénéité possible des réseaux de stockage sous-jacents. Enfin, plusieurs vendeurs de commutateurs et de directeurs prévoient dès 2004 la mise à disposition de services de virtualisation dans leurs équipements (cf. figure 1).

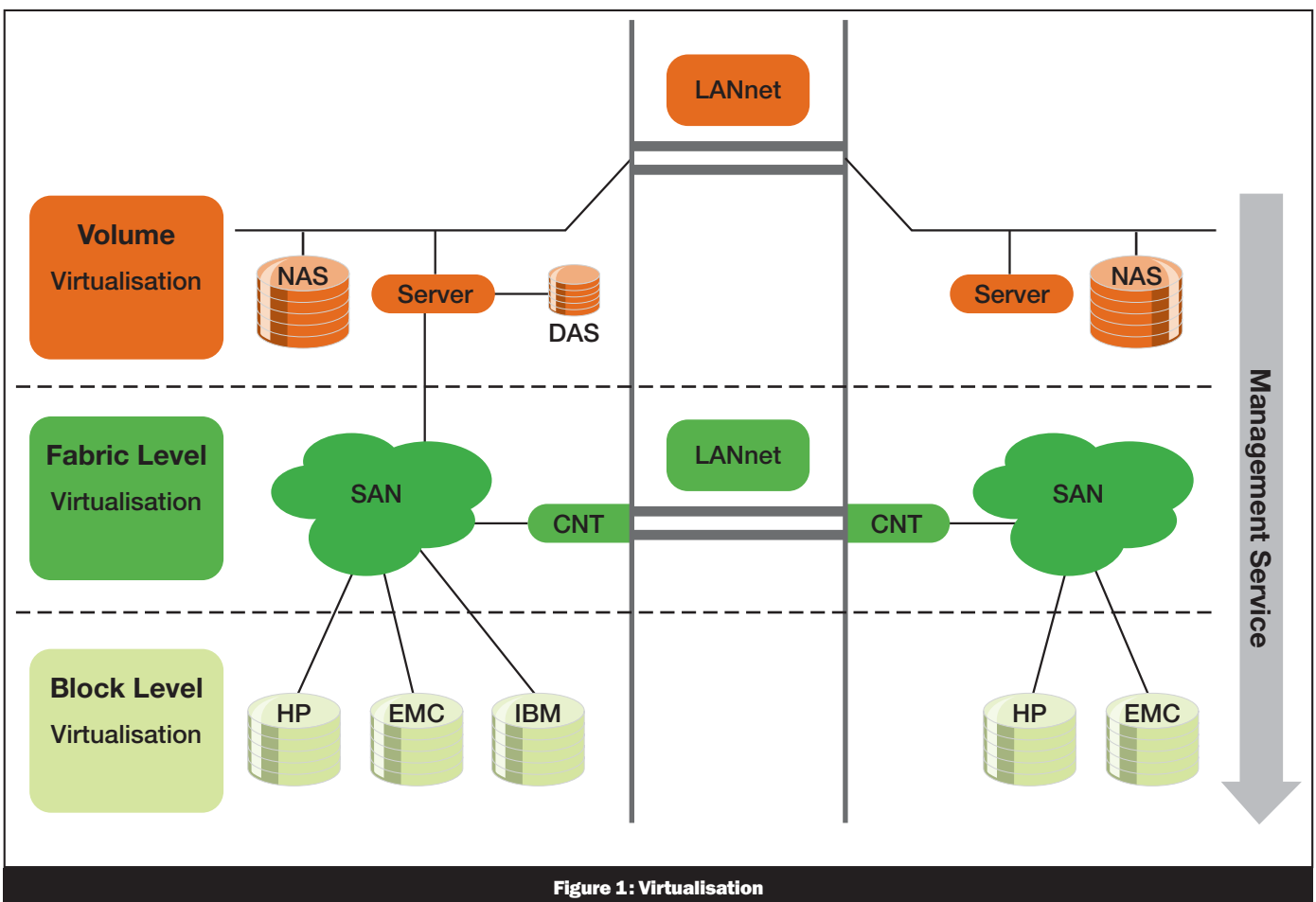


Figure 1: Virtualisation



- Certains considèrent enfin que les technologies de snapshots entrent dans le cadre de la virtualisation. Les vendeurs de baies de stockage proposent tous la possibilité de créer une copie datée d'un volume, vers lequel seraient éventuellement dirigées les E/S. Les snapshots peuvent également être disponibles au niveau des gestionnaires de volumes ou des systèmes de fichiers.



Les arguments avancés sont essentiellement tournés vers la productivité et la baisse du coût d'administration :

- Gestion centralisée en environnement hétérogène : des systèmes équipés de systèmes d'exploitation différents accèdent à un pool de stockage constitué de baies hétérogènes.
- Allocation dynamique d'espace (sans arrêt de production), mais aussi suppression et retour au pool de stockage.
- Meilleur taux d'occupation des espaces de stockage disponibles.
- Utilisation graduée de la capacité en fonction des temps d'accès requis.
- Automatisation de la restauration de fichiers archivés.
- Gestion simplifiée des opérations de copies et de migrations de données.
- Amélioration des processus de continuité de service par les systèmes de snapshots, de chemins redondants, etc.

Niveaux de virtualisation

Equipements et configurations	Mise en œuvre
Gestionnaire de volumes	Système d'exploitation (LVM - HP et IBM - VM et SDS - SUN...)
HBA (Host Bus Adapter)	Vendeurs de cartes HBA
Lun masking et zoning	Commutateurs et baies de stockage
Lun virtuels	Appliances de virtualisation et commutateurs
Stockage physique	Vendeurs de baies

Niveaux de virtualisation

SRM (Storage Resource Management)

Si le SRM a pour vocation première la gestion des ressources physiques d'un SAN, il doit également fournir une vue orientée métier de la vie des données : volumétrie, fréquence de changement, archivage, indisponibilité et in fine TCO (Total Cost of Ownership). Le SRM donne une vue physique et logique des performances, des taux de remplissage et de la configuration du réseau de stockage. L'information ainsi collectée n'a d'intérêt réel que lorsque les données produites sont agrégées et corrélées afin de pouvoir définir des pratiques et un niveau d'automatisation qui assurera au final à la fois une élévation de la qualité de service et une réduction des coûts de stockage.

Toutes les sociétés ne mettent pas en place des procédures de SRM dans les mêmes buts, mais la démarche associée puise dans les pratiques ci-dessous :

- Regroupement des ressources physiques au sein d'applications ou de fonctions métiers.
- Modèle prédictif de croissance de ressources pour ces groupes (par le suivi comportemental des utilisateurs - en particulier de leur boîte à lettres de messagerie - et des applications).
- Suivi des évolutions des ressources physiques et logiques.
- Automatisation de la gestion des ressources physiques en fonction des SLA (performance, disponibilité et fiabilité des données).
- Rapport de consommation de ressources (performances et volumétries).
- Si possible, procédures de contrôle des consommations et réduction des volumes de stockage.

La mise en lumière des comportements des utilisateurs permet très fréquemment de réduire la volumétrie de façon significative : stockages de fichiers étrangers à l'entreprise, boîtes aux lettres de messagerie énormes, etc. Il est également fréquent de pouvoir réduire les volumes des sauvegardes en identifiant les principaux contributeurs, en archivant les boîtes aux lettres, etc.

Outre la stagnation des dépenses informatiques consacrées au stockage et l'absorption des surinvestissements effectués



dans la bulle Internet, le principal obstacle du SRM était jusqu'à 2003 le manque d'hétérogénéité des outils. Des efforts notables ont été effectués par les constructeurs et par Veritas dans ce domaine dans les dernières versions de leurs logiciels de SRM. La gamme des logiciels intégrés à ECC V5 chez EMC apporte par exemple une amélioration importante dans ce domaine. Il reste tout de même que la collaboration entre constructeurs concurrents est visiblement limitée. Il est en général possible de découvrir les espaces de stockage configurés, mais il est rarement possible d'en modifier la configuration.

Cycle de vie des données (ILM)

La gestion du cycle de vie des données consiste à formaliser la vie d'une information numérique, de sa création à son archivage ou à sa destruction. La réduction du TCO apparaît de nouveau en filigrane : la mise en adéquation de la valeur intrinsèque de l'information avec le support qui la contient permet de réduire les coûts de stockage. Le HSM (Hierarchical System Management), bien connu dans le monde du Mainframe, permet essentiellement de choisir un support de stockage en se basant sur une vision liée au nom du fichier, à sa date de création, à son répertoire, etc. L'ambition des logiciels de gestion de contenu est autre : attribuer un cycle de vie basé sur des politiques parfois complexes à des informations indépendamment de leur contenant. Le terme « gestion de contenu » couvre un large spectre fonctionnel : gestion de l'archivage, gestion documentaire, gestion de contenu web. L'année 2003 a vu plusieurs rapprochements s'opérer entre les éditeurs de gestion de contenu et les constructeurs : EMC a acheté Documentum, tandis qu'Ixos, acquis par Opentext, est commercialisé par HDS et StorageTek. Ces deux fronts s'opposeront de plus à IBM, qui dispose de l'intégralité des briques. Décliner les règles métiers au niveau du stockage passe par un HSM. Les logiciels de ce type ne sont pas nombreux sur le marché, et le leader Disc Xtender est passé à travers le rachat de Legato dans l'escarcelle d'EMC. Le mariage entre les éditeurs de gestion de contenu et les constructeurs de baies de stockage risque de surcroît de ressembler à celui de la carpe et du lapin : les interlocuteurs ne sont jamais les mêmes chez les clients. Outre Documentum, EMC devra de plus digérer les rachats de Legato et de VMWare.

Conclusion

Les grandes manœuvres dans le domaine des logiciels périphériques au stockage permettent aux constructeurs d'espérer investir de nouveaux marchés après des années de

marasme. L'argumentaire de vente est toujours orienté vers la baisse des coûts d'administration et la rationalisation des coûts de stockage. Outre le coût d'achat généralement élevé des licences, les technologies mises sur le marché se signalent également par des problèmes de compatibilités en environnement hétérogène : rien de bien surprenant à ce que chacun cherche à protéger son pré carré. ■

Quelques acteurs du marché :

Généralités

Snia (Storage Networking Industry Association) : www.snia.org

SRM

Veritas : www.veritas.com

StorageTek : www.storageitek.com

EMC : www.emc.com

BMC : www.bmc.com

Virtualisation

Datacore : www.datacore.com

Hewlett Packard : www.hp.com

EMC : www.emc.com

FalconStor : www.falconstor.com

IBM : www.ibm.com

StoreAge : www.storeage.com

StorageTek : www.storageitek.com

Tricord : www.tricord.com

Veritas : www.veritas.com

XIOtech : www.xiootech.com

ILM et HSM

E-Foundation/Storage Foundation :
www.e-foundation.com

ADIC : www.adic.com

EMC : www.emc.com

StorageTek : www.storageitek.com

Hewlett Packard : www.hp.com

Baies de Stockage

Hitachi Data Systems : www.hds.com

IBM : www.ibm.com

Hewlett Packard : www.hp.com

Sun Micro Systems : www.sun.com

NetApp : www.netapp.com

Fabrics

McData : www.mcdata.com

Brocade : www.brocade.com

IBM : www.ibm.com

Cisco : www.cisco.com



Chistian Bilien

Ingénieur de l'Ecole Supérieure d'Electronique de l'Ouest et diplômé de l'IEFSI-EDHEC, Chistian Bilien a 13 ans d'expérience en ingénierie des systèmes et des bases de données, acquise chez les constructeurs et dans des grands comptes. Il s'est en particulier spécialisé dans les performances des systèmes et des bases de données. Il est actuellement directeur des opérations de Storage Foundation, une société spécialisée dans le conseil en infrastructures de stockage et de bases de données.

cbilien@storage-foundation.com